

MEETING REPORT

Meeting on big mutations addresses big questions in human genetics

Donald F Conrad*

Abstract

A report on the Keystone Symposium 'Functional Consequences of Structural Variation in the Genome', Steamboat Springs, Colorado, USA, 8-13 January 2011.

Introduction

Low-cost technologies for variation discovery that can screen entire genomes (such as oligonucleotide arrays) have been available to researchers studying copy number variations (CNVs) for several years. These technologies allow them to zero in on extremely rare, large-effect mutations that could not be found by endeavors based on single-nucleotide polymorphisms (SNPs), such as the International HapMap Project (<http://hapmap.ncbi.nlm.nih.gov/>) and the hundreds of genome-wide association studies (GWASs) that it enabled. As the high-profile questions in human genetics have turned from the biology of common variation to the properties of rare mutations and the process of mutation itself, those scientists who initially set out to study CNV itself now find themselves in a position to address some of the biggest current questions. This was manifest in January's Keystone Symposium on structural variation, which featured talks dissecting the extent of rare variation and somatic mosaicism and their impact on health, presentations on measurement and analysis of mutation and mutation rates, and talks about phenotypes of *in vivo* cellular models derived using induced pluripotent stem (iPS) cell technology. Here, I cover some of the highlights, ranging from well established areas (see section 'Common CNVs'), present-day darlings ('Rare CNVs'), and the impressive foundations that are being put in place to translate genetic research into clinical tools ('Towards genomic medicine').

Common CNVs

The mapping and characterization of common germline structural variation is quickly becoming a mature field. Ryan Mills (Harvard Medical School and Brigham and Women's Hospital, Boston, USA) discussed the progress that the 1000 Genomes Project (<http://www.1000genomes.org>) has made towards mapping all structural variations (SVs) with over 1% frequency to base-pair level resolution in populations from three continents. In the pilot phase of the project, over 28,000 CNVs have been discovered, ranging in size from 50 bp to over 500 kb. These numbers are consistent with previous analyses of the common CNV universe, which estimate that there are 4,000 to 5,000 common deletions over 450 bp in size in European populations. Importantly, the 1000 Genomes Project dataset contributes several forms of SV mutation that had been poorly documented until now - mobile elements, such as long interspersed nuclear elements (LINEs) and Alus, and insertions of sequence not seen in the human reference genome.

Although many people expected that, compared with SNPs, common CNVs would contribute disproportionately to the variation in risk for common diseases, the past 3 years of CNV GWASs has provided no evidence to support this hypothesis, as summarized by Matthew Hurles (Wellcome Trust Sanger Institute, Hinxton, UK). One potential explanation for this counterintuitive finding is that the traits studied so far have been prioritized for their medical interest, and common CNVs, whose locations are biased toward specific genes and genomic regions, are likely to have an impact on more superficial traits, such as olfaction, digestion, taste, and pigmentation.

Rare CNVs

The same array technologies that have enabled whole-genome mapping of common variants have also been successful in helping researchers identify many previously unknown genomic disorders and to characterize CNV mutation processes. In the keynote address, James Lupski (Baylor College of Medicine, Houston, USA) reviewed his contributions to the field, from the description of non-allelic homologous recombination as the mutation process forming the recurrent aneuploidies

*Correspondence: dconrad@genetics.wustl.edu
Department of Genetics, Washington University School of Medicine, St Louis, MO, USA

that are the basis of genomic disorders to the recent introduction of a mutation model called microhomology-mediated break-induced replication that can form CNVs of arbitrary complexity. He then announced the exciting discovery of a new, as yet unnamed CNV mutation process, which produces a nested duplication-triplication-duplication structure.

Evan Eichler (University of Washington, Seattle, USA) presented a summary of the 17 rare, recurrent microduplications and microdeletions that his laboratory has discovered to be associated with intellectual disability, autism, and epilepsy; all this from an explosion of work in just the past 5 years. Rare variant analyses also featured prominently in a talk on autism by Dalila Pinto (Hospital for Sick Children, Toronto, Canada), who reported an increased rate of *de novo* CNVs in children with autism compared with controls, and much larger *de novo* CNVs in children from simplex families (with only one affected member) than from multiplex families (with more than one).

The final frontier in rare variant analysis, the rarest of the rare, is the somatic mutations that segregate among different cells of the same individual. SNP arrays are an important new tool for the analysis of mosaicism: because of their sensitive quantification of allele-specific copy number, it is now possible to reliably detect mosaic aneuploidies and uniparental disomies with frequencies down to 5% or less in a population of cells, allowing researchers to describe both the normal distribution of mosaicism in healthy individuals and previously cryptic pathologies. Nancy Spinner (Children's Hospital of Philadelphia, USA) reviewed results from 459 patients with cytogenomic abnormalities visible on Illumina Quad2 arrays; 60 (13%) of these cases were mosaic. In his talk on somatic mutation and aging, Jan Dumanski (Uppsala University, Uppsala, Sweden) described an increase in the number somatic mutations in apparently healthy monozygotic twins over 60 years old compared with younger twin pairs, and estimated that 3 to 4% of the population over 60 years old are mosaic for a large CNV that is detectable with current technology.

Towards genomic medicine

Two challenges for genomic medicine are *in vivo* functional characterization of mutations in human cells and quantification of the probability that a known or novel variant is causally related to disease. Several talks at the meeting addressed these challenges.

Perhaps the most exciting new development at this meeting were the first reports from groups that are starting to investigate the functional impact of structural variation in human iPS cells, *in vivo* systems that can be used to directly model disease states. Kristen Brennand (The Salk Institute, La Jolla, USA) presented impressive results from her studies of neurons derived from iPS cells of individuals

with apparently familial cases of schizophrenia and controls. Compared with controls, cells from people with schizophrenia demonstrated lower neural connectivity, reduced neurite outgrowth, gene expression changes, and impaired synaptic transport, the latter of which improved after treatment with the antipsychotic drug loxapine. Jonathan Sebat (University of California, San Diego, USA) reported an exciting new association between duplication of the vasoactive intestinal peptide receptor 2 (*VIPR2*) gene and schizophrenia; experiments are under way to test the impact of *VIPR2* antagonists in schizophrenia-derived iPS cell systems. Ira Hall (University of Virginia, Charlottesville, USA) presented for the first time experiments investigating the impact of inducible reprogramming on genome stability. There is a concern that the reprogramming factors used by iPS cell technology may accelerate the rate of somatic mutation, adding noise to assays of single mutations; fortunately, Ira's preliminary results suggest that this effect is likely to be modest.

The identification of causal mutations in severe cases of pediatric disease is a rapidly changing area of medical inference. As the size spectrum of unresolved disease-causing mutations gets smaller, the overlap with benign variation increases, and the calls for 'evidence-based' assessment of causality are mounting. As a result, the leaders in medical genetics across the world have embraced statistical thinking, as reflected in massive array datasets described by Lisa Schaffer (Signature Genomic Laboratories, Spokane, USA), presenting the Signature Genomics database, and David Ledbetter (Geisinger Health System, Danville, USA), presenting analysis of CNV data from the International Standards for Cytogenomic Arrays (ISCA) Consortium. With sample sizes over 45,000 and over 15,000, respectively, these medical genetics projects have moved into a scale previously only obtained by the largest GWAS meta-analyses. Both speakers estimate that causal mutations have been identified for 15 to 25% of the cases in each database and highlighted cases in which previously unknown disease mutations were identified by rigorous statistical analysis. Although medical genetics will no doubt continue to identify new causal variants with the same paradigm of case-control association used by GWASs, it is important to set our sights higher and develop new methods for assessing causality for mutations that have never been seen before. Conceptually this might be done by finding features generally enriched in known disease-causing mutations and then assessing the evidence, by analogy with confirmed disease variants, that an unclassified variant is causal.

The vanguard of human genetics

It is clear that leaders in the CNV field are now addressing questions of broad interest across human genetics,

having mastered new technologies for variation discovery and embraced statistical thinking. Investigators of CNV have broadened their experimental toolkit and talks routinely navigated data from fluorescence *in situ* hybridization experiments, SNP arrays, array comparative genomic hybridization, and next generation sequencing. As next-generation sequencing matures, and the integrated analysis of all human variation becomes routine, I predict that the scientific stereotype currently recognized as 'the CNV guy' will start to be referred to as 'the human geneticist'.

Abbreviations

CNV, copy number variation; GWAS, genome-wide association study; iPS cell, induced pluripotent stem cell; SNP, single nucleotide polymorphism; SV, structural variation.

Published: 22 February 2011

doi:10.1186/gm226

Cite this article as: Conrad DF: Meeting on big mutations addresses big questions in human genetics. *Genome Medicine* 2011, **3**:12.