

Abstract 52

A SYSTEMATIC, LARGE-SCALE RESEQUENCING AND COPY NUMBER VARIANT SCREEN OF THE X CHROMOSOME CODING EXONS IN MENTAL RETARDATION

F Lucy Raymond¹, Patrick Tarpey², Annabel Whibley¹, Mike Stratton², Charles Schwartz³, Roger Stevenson, Jozef Gecz⁴, Gillian Turner⁵ on behalf of the iGOLD study.

¹Cambridge Institute for Medical Research, Cambridge UK; ²Wellcome Trust Sanger Institute, Cambridge UK; ³JC Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, South Carolina USA; ⁴SA Pathology, Women's and Children's Hospital, North Adelaide and The University of Adelaide, Australia; ⁵GOLD Service, Hunter Genetics, Waratah, New South Wales, Australia

Large-scale systematic resequencing is being proposed as the major future strategy for the discovery of rare, disease-causing sequence variants across the spectrum of human complex disease. We have sequenced the coding exons of the X chromosome in probands from 208 families with X-linked mental retardation, the largest direct screen for constitutional disease causing mutations thus far reported. The screen has identified nine XLMR genes, including *SYP*, *ZNF711* and *CASK* confirming the power of this strategy. The study has, however, also highlighted issues confronting whole genome sequencing screens including the observation that loss of function mutations in 1% or more of X chromosome genes is compatible with apparently normal existence. To address the contribution of genomic rearrangements, a mutational class inefficiently detected by capillary re-sequencing, we also designed an X-chromosome-specific Nimblegen 385K format microarray to interrogate copy number variation by comparative genomic hybridization (CGH). Our array design targeted coding regions and evolutionarily conserved elements but also maintained high density coverage across non-coding regions. Pathogenic variants were identified in >10% of the cohort families, with aberrations ranging from 2kb to >3Mb in size. Junction fragments have been obtained where possible, providing insight into the mutational mechanism underlying these genomic imbalances. As the majority of the families in the cohort still remain unresolved, it is likely that pathogenic mutations lie in the high number of unique missense sequence variants that were identified. Distinguishing rare neutral variants from pathogenic sequence and copy number variants is now a significant challenge.