

## SCIENCE – IN THE NEWS

# Mining the X-Chromosome for Disease Genes by Deep Resequencing

A team of 79 authors representing 18 research groups from Africa, Europe, and North America investigated 208 families with X-linked mental retardation (XLMR) in what the investigators describe as “the largest direct screen for constitutional disease-causing mutations thus far reported” (1). In this extensive resequencing study, the group sequenced 718 of the somewhat more than 800 genes on the X chromosome, and found 1,858 different variants in the coding regions.

This screen identified nine additional genes in which mutations appear to be associated with XLMR, adding approximately 10% to the roughly 80 genes previously implicated in XLMR. Among these new XLMR-related genes are *SYP*, a protein in the membranes of small synaptic vesicles; *ZNF711*, a protein with zinc-finger characteristics that presumably interacts like two interlocking combs with another zinc-finger protein; and *CASK*, a protein kinase in the postsynaptic membrane of CNS synapses.

Perhaps more important than the specific genes identified in this massive study of XLMR are the lessons learned that will have relevance to future whole-genome sequencing investigations (1). For example, the majority of the families in this study had previously been classified phenotypically as having nonsyndromic XLMR, but when the families were subdivided genotypically, seven of nine genes were associated with mental retardation (MR) and additional phenotypic features, such as seizures with *SYP* mutations and nystagmus with *CASK* mutations. Therefore, this sequencing strategy resolved what appeared to be a large group of phenotypically homogeneous families into a group of heterogeneous subsets that were genotypically and phenotypically consistent.

The investigators identified “a likely genetic basis” for XLMR in 25% (53 of 208 families) of the subjects screened based on mutations identifiable by their strategy (1). They attribute the lack of an identified genetic basis for the remaining 75% to a variety of possibilities, including inability to amplify some of the X-chromosomal genes, mutations in noncoding regions that were not sequenced, interstitial rearrangements such as copy number variations and inversions that would not be detected by their resequencing strategy, and contamination of the sample with individuals evidencing mental retardation that did not result from X-chromosomal mutations.

A striking observation was that loss-of-function mutations, including gene truncations, were observed among controls and/or did not

segregate in a familial manner with the XLMR in 1% or more of sequences encoding proteins on the X chromosome (1). The investigators speculated that this might be due to genomic duplication followed by truncation to degrade and eventually remove unnecessary, redundant copies. This is a well-recognized evolutionary mechanism in other eukaryotic organisms in which duplicated genes may take on novel expression patterns by variations in promoter sequences, but, if not beneficial to the organism, then they will be lost (2).

Tarpey *et al.* (1) have provided us a glimpse into the power of large-scale resequencing investigations that will soon include the full human genome, and they have also introduced us to the challenges involved in this strategy. Some disorders will be the consequence of mutations in a small set of genes, whereas a significant number will be, like XLMR, due to mutations in a large number of genes, with each contributing to only a small proportion of those with the familial risk of disease. The observation that  $\geq 1\%$  of apparent protein-encoding sequences on the X-chromosome tolerate loss-of-function mutations, suggesting they have no obvious biologic function influencing normal well-being or cognition, shows us the uncertainty that will be inherent in sequence-based medicine, in which a sequence may not be as it appears. Although extensive resequencing appears to be an extremely powerful technology, we must enter this exciting era with an open mind, recognizing we will most definitely experience consequences we have not anticipated. – Edward R. B. McCabe

## REFERENCES

1. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'Meara S, Latimer C, Dicks E, Menzies A, Stephens P, Blow M, Greenman C, Xue Y, Tyler-Smith C, Thompson D, Gray K, Andrews J, Barthorpe S, Buck G, Cole J, Dunmore R, Jones D, Maddison M, Mironenko T, Turner R, Turrell K, Varian J, West S, Widaa S, Wray P, Teague J, Butler A, Jenkinson A, Jia M, Richardson D, Shepherd R, Wooster R, Tejada MI, Martinez F, Carvill G, Goliath R, de Brouwer AP, van Bokhoven H, Van Esch H, Chelly J, Raynaud M, Ropers HH, Abidi FE, Srivastava AK, Cox J, Luo Y, Mallya U, Moon J, Parnau J, Mohammed S, Tolmie JL, Shoubbridge C, Corbett M, Gardner A, Haan E, Rujirabanjerd S, Shaw M, Vandeleur L, Fullston T, Easton DF, Boyle J, Partington M, Hackett A, Field M, Skinner C, Stevenson RE, Bobrow M, Turner G, Schwartz CE, Gecz J, Raymond FL, Futreal PA, Stratton MR 2009 A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 41:535–543
2. McCabe L, McCabe ERB 2008 DNA: Promise and Peril. University of California Press, Berkeley, pp 57–65