This Month in *The Journal*

Kathryn D. Bungartz[1] and Robin E. Williamson[2]

Copy-Number Variation in X-Linked Intellectual Disability

Whibley et al., page 173

One of the methods researchers have used to narrow the search for genetic causes of intellectual disability (ID) has been to focus efforts on identifying the genes that are disrupted in cases of X-linked ID (XLID). By studying families predicted to have XLID, only the genes on the X chromosome need to be evaluated, and with technological improvements, it is now possible to screen the entire chromosome in a large number of individuals. A recent resequencing study looked for mutations in all of the X chromosome genes in XLID families, and although a number of pathogenic variants were identified, the majority of families did not carry any obvious X chromosome sequence mutations. These results led researchers to consider that there were potentially other kinds of chromosomal disruptions in these families that were not easily identified by the sequencing procedures used. In this issue, Whibley et al. describe their efforts to identify X chromosome copy-number variation (CNV) that may be responsible for causing XLID. A first step in this endeavor is the creation of an X chromosome array that has high enough resolution to detect small genomic imbalances. With this array, the authors find that some of their XLID families have disruptions in genes that have not yet been associated with ID. In other families, Whibley et al. identify structural anomalies that disrupt genes previously known to be involved in ID. The size of these CNVs varies, and in some of these cases, the disruption is large and contains a number of candidate genes, whereas in others, the CNV is predicted to affect the expression of a single gene. This comprehensive analysis of X chromosome CNVs in XLID families not only contributes to the identification of genes involved in brain development and function, but also provides additional information about the structural characteristics of the X chromosome.

Microdeletions of 3q29 and Schizophrenia

Mulle et al., page 229

Genome-wide association studies (GWAS) have been widely used to identify SNP variants associated with a broad spectrum of complex traits and disorders. This approach has been accompanied by a tremendous amount of hope in uncovering common variants conferring substantial risk. Technology has expanded as expectations have grown. Although GWAS have identified numerous variants to be associated with many different traits and disorders, the total amount of discovered heritability remains quite low. A promising addition to the genome-wide SNP association studies is to look for associated copy-number variants (CNVs). Together with the identification of rare variants, CNV associations may uncover a great deal of the missing heritability of complex diseases. Here, Mulle and colleagues identify a CNV in schizophrenia (SZ) patients conferring among the largest amount of risk of any reported variant for any complex disorder. This group analyzes SZ cases of Ashkenazi Jewish descent by using a genome-wide SNP array and identifies a significant enrichment of deletions at 3q29 compared to non-SZ controls. Adding in data from other SZ cases and controls of different ethnic backgrounds, Mulle and colleagues find a total of six 3q29 deletions in 7545 cases and only one in 39,748 controls. They calculate the odds ratio of 3q29 deletions to be 16.98. This group also confirms significant association of the well-established 22q11 deletion with SZ. Such findings provide hope that a substantial portion of missing heritability may yet be uncovered.

Detecting Heteroplasmy in Human mtDNA

Li et al., page 237

Next-generation sequencing technology has multiple applications. It can be used to sequence exomes of genomic loci associated with disease (e.g. Walsh et al., July issue) or whole exomes of patients with recessive diseases who lack abundant genetic information (e.g. Pierce et al, this issue). Here, Li and colleagues utilize next-generation sequencing to determine heteroplasmy levels of complete mtDNA genomes. mtDNA genomes are routinely used in human population- and evolutionary-genetics studies. There are also several disorders associated with mtDNA mutations. In disease, mtDNA mutations are commonly heteroplasmic, meaning that not all cells or tissues contain the mutated mtDNA. Interestingly, people harboring no known disease also harbor multiple mtDNA types. The degree of heteroplasmy, and how to calculate it, has been a subject of extensive investigation, for both disease status and forensic investigation. Because next-generation sequencing can provide quick

[1]Science Editor, *AJHG*; [2]Deputy Editor, *AJHG*
DOI 10.1016/j.ajhg.2010.07.016.

and potentially cost-effective data, Li and colleagues propose and demonstrate its usefulness in detecting mtDNA heteroplasmy levels. Using a combination of simulations and bacteriaphage genome analyses, they first establish criteria for this application. These criteria will potentially be very useful to many investigators interested in mtDNA heteroplasmy studies. Li and colleagues then apply their criteria to a study of their own. They report heteroplasmy levels, mutation rates within heteroplasmies, and evolutionary trends in five Eurasian populations.

HSD17B4 Mutations and Perrault Syndrome

Pierce et al., page 282

Exome sequencing is quickly becoming a common tool for identifying disease-causing mutations in affected individuals. In instances when linkage analysis or homozygosity mapping is used as a starting point to identify an associated locus, exome sequencing of the mapped region can be used to identify causative mutations. This technique is used in an article featured in the July issue of *AJHG* (Walsh et al.). However, it is not always feasible to limit the location of a causative mutation to a specific locus or chromosome, because disorders are often identified in single patients or in small nonconsanguineous families. In such instances, not enough genomic information is available to make linkage analysis or homozygosity mapping informative. Here, Pierce and colleagues demonstrate that whole-exome sequencing can facilitate the identification of pathogenic mutations in cases for whom little genomic information is available. By sequencing the entire exome of a patient from a small nonconsanguineous family with Perrault syndrome, this group identifies more than 200 candidate nonsense, missense, frameshift, or splice variants. Perrault syndrome is predominantly characterized by sensorineural hearing loss and ovarian dysgenesis (in females); patients may also present with neurological manifestations. Because Perrault syndrome is consistent with autosomal or X-linked inheritance in the reported family, Pierce and colleagues look for genes harboring two potentially pathogenic variants. Only one gene, *HSD17B4*, is found to have two such mutations in the proband. These two alterations segregate with disease in the family. This study identifies the first genetic mutations involved in Perrault syndrome and highlights the utility of whole-exome sequencing.

Aboriginal Australian Genetic Diversity

McEvoy et al., page 297

Following the out-of-Africa exodus, people reached Australia about 60,000 to 46,000 years ago during the Pleistocene period. At that time, Australia, Tasmania, and New Guinea were part of a single land mass called the Sahul. Information about the routes taken and the timing of the settlement of the region has been gathered from archaeological data and preliminary genetic studies, but questions remains about the circumstances of the peopling of the Sahul. One point of discussion is whether there was one major migration to the region followed by subsequent dispersion to different locations or whether there were multiple independent waves of settlers that arrived in the different Sahul locations. In this issue, McEvoy and colleagues perform genome-wide analyses of an Australian Aboriginal population and combine this with mtDNA data to learn more about the population structure of the people. They find strong evidence that a significant proportion of the male ancestry of the population is not Aboriginal Australian but is instead of European origin. Because the authors' goal of learning more about the ancestral population history would ideally be served by working with a sample with both paternal and maternal Aboriginal ancestry, McEvoy and colleagues simulate such an ideal ancestral population and, using that data, compare these ideal Aboriginal Australians with the populations from around the world. They find that the ancestral population clusters well with the other groups from the Sahul, which supports the prediction that a single migration settled the region and then dispersed throughout the Sahul. Given the high admixture in the current Aboriginal population, the authors also generate a panel of ancestry-informative markers that will be useful in future admixture studies.