This Month in *The Journal*

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

Genetic Substructure of the Han Chinese Population

Xu et al., page 762 Chen et al., page 775

The People's Republic of China constitutes an enormous country rich in history and diverse in culture. The Han Chinese represent the largest ethnic group of people within China. Because of the tremendous population of China, Han people comprise ~20% of the total human population. However, the Han Chinese are not a uniform people. Rather, there are regional, cultural, and linguistic differences within the Han. Thus, there are probably genetic differences to consider when performing GWAS of Chinese people. Some of these differences have been noted in the different HapMap populations, which include Chinese people from Beijing (CHB) and from Denver (CHD). Here, two different groups present their comprehensive analyses of the genetic structure of Han Chinese from different Chinese provinces. Through genotyping and analyzing thousands of Han Chinese from many different regions of China, both Chen and colleagues and Xu and colleagues note a north-south division in Chinese population structure. These data corroborate previous studies noting genetic differences between northern and southern Chinese. Both groups advise caution when designing a GWAS of Chinese people because findings can be confounded by the differences in genetic makeup between these regions. In addition, the population substructure within individuals from large metropolises can be somewhat complex as a result of migration events. However, as a whole, the HapMap CHB population corresponds more to a northern Chinese population, whereas the CHD people correspond more closely to a southern Chinese population. Together, these analyses will contribute to future GWAS for complex diseases; an endeavor in which China has pledged to invest.

Short Telomeres and Age-Related Disease

Armanios et al., page 823

The telomeres at the end of chromosomes are known to shorten as DNA replication and cell division proceed. When the ends pass a certain point in length, the cell cycle stops and cells die. To counteract this shortening process, the enzyme, telomerase, acts to replenish the telomeres, and when mutations affect this function, diseases associated with premature aging and early mortality are the result. Despite these correlations between telomere shortening and aging, it has not been possible to establish whether telomere shortening is sufficient to cause aging in cells in which telomerase is functioning normally. Here, Armanios et al. study a unique strain of mice in an effort to observe a direct relationship between telomere shortening and phenotypes associated with aging. The mice are the wild-type late-generation offspring of mice who are heterozygous for a deficiency in telomerase. Because telomere length is heritable, these mice have the abnormally short telomeres of their parents, but do not have anything wrong with their own telomerase enzyme. The authors examine hematopoietic and immune function in these mice and determine that short telomeres alone are enough to cause aging phenotypes.

GWAS with Overlapping Subjects

Lin and Sullivan, page 862

Given the large number of samples that are often needed to generate enough power to observe significant associations of small effect sizes in genome-wide association studies, combining data in meta-analyses has emerged as a critical method for association analysis. One issue to keep in mind, though, when performing meta-analyses is that sometimes independent studies use overlapping samples. This can often be the case when studies have utilized publicly available control data. For minimizing the effect of these shared samples on the meta-analysis, it is possible to simply designate that some controls be considered to be part of one study and that others be relegated to another study. Unfortunately, such splitting isn't always possible if only summary data are available, and even when using individual data, power loss and bias can result. Ideally, it would be best to be able to use all data and to mathematically account for the lack of independence between the observations based on the shared samples. Here, Lin and Sullivan present their method for handling these correlated observations in meta-analysis. They demonstrate that they are able to preserve the false-positive rate while benefiting from maximum power and lack of bias and show the advantages of their method over the splitting techniques on WTCCC rheumatoid arthritis and type-1 diabetes data as well as on a meta-analysis of schizophrenia data.

[1]Science Editor, *AJHG*; [2]Deputy Editor, *AJHG*
DOI 10.1016/j.ajhg.2009.11.013.

TRAPPC9 Mutations in Autosomal-Recessive Mental Retardation

Mochida et al., page 897 Philippe et al., page 903 Mir et al., page 909

Mental retardation (MR) is a common type of developmental disability affecting 2%–3% of the world population. Although both pre- and postnatal environmental factors can contribute to the development of intellectual disabilities, many cases of MR are thought to have a genetic foundation. Ranging from mild to severe, MR can present alone or as part of a syndrome and can be inherited in a recessive or dominant fashion. Autosomal recessive (AR) MR is estimated to account for ~25% of all cases; however, only a handful of genetic mutations have thus far been associated with this type of MR. Here, three groups independently identify mutations in *TRAPPC9* as causative for ARMR associated with microcephaly. Mochida and colleagues and Mir and colleagues use homozygosity mapping and gene sequencing to identify *TRAPPC9* mutations in their respective families, whereas Philippe and colleagues combine autozygosity mapping and RNA expression profiling to identify *TRAPPC9* as a candidate, followed by sequencing to identify mutations. *TRAPPC9* encodes a protein that goes by many names, including IKK-β binding protein, NIK binding protein, and trafficking protein particle complex 9 (TRAPCC9). As the names indicate, TRAPPC9 binds IKK-β and NIK, both involved in the NF-κB signaling pathway, and is thought to be involved in protein trafficking. NF-κB signaling regulates transcription of genes involved in cell proliferation and cell survival. Through binding NF-κB-activating proteins, TRAPPC9 enhances NF-κB activation. Both the Mochida and Philippe groups show impaired NF-κB signaling in their MR patients harboring *TRAPPC9* mutations. Together, these three groups provide evidence of NF-κB activity in the brain and a possible mechanism for ARMR.

Purifying Selection in the Y Chromosome

Rozen et al., page 923

The male-specific region of the Y chromosome (MSY) contains 16 genes that are called X degenerate genes because they no longer have a homolog on the X chromosome. On the basis of the facts that many genes have been lost during Y chromosome evolution and that, since the chimpanzee-human divergence, the chimpanzee genome has lost functional versions of four of these X degenerate genes, there has been speculation that the human Y chromosome genes are on there way out and that the human Y chromosome will eventually decay away. In this issue, Rozen and colleagues gather additional information about the human X degenerate genes in an effort to assess this prediction. The authors sequence the 16 X degenerate genes from 105 Y chromosomes that were selected to represent maximum population diversity. When they compare the sequences at a protein level, they find very little variability. In fact, the authors report that the sequences of the MSY degenerate genes of two randomly selected Y chromosomes differ, on average, by a single amino acid. This lack of diversity at the amino acid level, along with the finding that synonymous changes and intronic changes are more frequent, leads the authors to predict that selection has acted to maintain the protein sequence of the X degenerate genes and that the genes have not changed much in the last 100,000 years. Such strict maintenance of gene function is an indication of the importance of these genes, and Rozen and colleagues suggest that this serves as evidence against the hypothesis that the Y chromosome will become obsolete in future human generations.