

## This Month in *The Journal*

Kathryn D. Bungartz<sup>1</sup> and Robin E. Williamson<sup>2</sup>

### He Ain't Heavy, He's My Brother

**Pemberton et al., page 457**

HapMap has grown. With the third release, The International HapMap Consortium has analyzed more samples from the original four populations, and they've also studied samples from seven additional populations. The study design involved collecting samples from known related individuals for some of the population groups, but with such a large number of participants, there are also likely to be unknown relative pairs included as well. Because the performance of certain analyses on the HapMap 3 samples might be confounded by unidentified relationships, recognition of all closely related individuals is important. In this issue, Pemberton and colleagues use the genome-wide SNP data of the population samples to look for unknown pairs of close relatives. The authors stick to predicting relationships closer than first cousins, because assignment of more distant relationships can be less reliable. Following their analysis, they report that the sex of some of the samples is likely to have been mislabeled, a few of the relationships reported by HapMap are incorrect, and two of the samples are either monozygotic twins or duplicates. They then delineate all the unknown relationships that they infer from the data. In the end, Pemberton and colleagues present two lists that will aid researchers using these HapMap 3 samples: one list is of the samples that should be removed if a data set of individuals without any parent-offspring or full-sibling pairs is needed, and the second list is of the samples to exclude if half sibling, grandparent-grandchild, and avuncular pairs are to be removed as well.

### Move Over, GT, There's a New Splice Site in Town

**Hartmann et al., page 480**

Fanconi anemia (FA) is a genetic disorder of blood cells that leads to bone marrow failure. Rather than producing enough new blood cells to maintain the body, the bone marrow of FA patients often produces abnormal blood cells, which results in the development of leukemia. This serious condition is caused by mutations in at least 13 genes involved in DNA repair. FA can be classified into three distinct categories, which differ in severity and genetic cause. FA type C (FANCC) is caused by mutations in *FANCC* and belongs to the category with the least severe

phenotype. In this issue, Harmann and colleagues investigate the disease mechanism of patients with FANCC. They identify the same c.456+4A>T mutation in nine FANCC patients from three different pedigrees. The mutation changes a 5' (donor) splice site from "GT" to "TT" at the DNA level. Although around 99% of all 5' splice sites use the "GT" consensus sequence, "AT" can also serve as a splice donor site. "TT" has not been considered a donor splice site sequence at all. Thus, the c.456+4A>T mutation has been predicted to prevent correct splicing in FA patients. However, Hartmann and colleagues find the "TT" sequence to act as a 5' splice donor site some of the time, resulting in a small amount of correctly spliced *FANCC*. This correctly spliced product results in low levels of normal posttranscriptionally processed protein, accounting for the milder phenotype of these patients. In addition to providing a mechanism for the phenotypic difference in FANCC patients, this work suggests that gene therapy for splice-site mutations may be able to correct mRNA processing in FA and other diseases.

### Searching for CNV Hotspots

**Fu et al., page 494**

In the last few years, there has been a growing recognition of the contribution of copy-number variants (CNVs) to the risk of developing both Mendelian and complex diseases. Previous work has examined the mutation rate of some CNV loci, and results suggest that these sites mutate more frequently than do single-nucleotide sites and that there is a great deal of variation in the mutation rates of CNV loci. Therefore, investigating CNVs involves not only looking for an association with disease but also determining why CNVs form in the genome where they do and why some unstable regions are affected more frequently than others. In this issue, Fu and colleagues endeavor to estimate the mutation rate of CNV loci throughout the genome and establish whether CNV hotspots exist. The first step in their study is to develop methodology that they call CNVMut, which allows them to predict CNV mutation rates by using SNP data. The authors then apply CNVMut to data from CNV loci in the YRI, CEU, and JPT+CHB HapMap populations. Along with observing diversity in the values of mutation rates at these loci, they also identify sites that are predicted to be potential CNV mutational hotspots. The genomic architecture of

<sup>1</sup>Science Editor, *AJHG*; <sup>2</sup>Deputy Editor, *AJHG*

DOI 10.1016/j.ajhg.2010.09.012. ©2010 by The American Society of Human Genetics. All rights reserved.

these CNV hotspots is compared to that of regions with lower mutation rates in an effort to figure out which features might contribute to their increased instability. As the importance of CNVs in disease increases, so too does the need to understand the mechanisms behind the creation of CNVs.

### A Genetic Contribution to Male Infertility

#### Bashamboo et al., page 505

Infertility can be a huge obstacle for couples trying to have children, and even with improvements in assisted reproductive techniques, many couples struggle to conceive. Previous work supports genetic factors as a contribution to a percentage of cases of male infertility: chromosomal abnormalities are responsible for some cases, and mutations in single genes have been found in others. One gene in particular, *NR5A1*, encodes a protein known to be involved in gonadal development, and mutations can cause complex phenotypes that include infertility. Here, Bashamboo report their findings that *NR5A1* mutations can also cause infertility in men who are otherwise healthy. The men carrying these mutations are all affected with at least moderate oligozoospermia, and all six mutations fall within the same region of the protein. The authors perform functional studies on mutations that were heterozygous in the seven infertile men. They find that, although the mutant proteins share a localization pattern with that of the wild-type protein, their ability to activate transcription of target genes is significantly disrupted. On the basis of these data, *NR5A1* could

be added to the list of genes that, when mutated, cause male infertility.

### Sphingolipids and Neurodegeneration

#### Rotthier et al., page 513

Sphingolipids are proteins found within the plasma membrane and are thought to play a protective role against harmful environmental factors. They have also been shown to be involved in cell signaling pathways important for apoptosis, proliferation, and stress responses. Sphingolipids are particularly important in nervous tissue, because mutations in many of the genes involved in sphingolipid biosynthesis result in neurodegenerative diseases. Hereditary sensory and autonomic neuropathy type I (HSAN-I) is one such disorder. HSAN-I is the most common type of HSAN and is transmitted in an autosomal-dominant manner. HSAN-I is characterized by a lack of feeling in the lower limbs, chronic ulcerations of the feet, and progressive destruction of affected bones. About 19% of HSAN-I patients are known to have mutations in *SPTLC1*, encoding the first subunit of serine palmitoyltransferase (SPT). *SPTLC1* dimerizes with either *SPTLC2* or *SPTLC3* to form an active enzyme that catalyzes the first (and rate-limiting) step of sphingolipid biosynthesis. Here, Rotthier and colleagues sequence *SPTLC2* and *SPTLC3* in a cohort of HSAN patients who do not harbor *SPTLC1* mutations. They identify three different missense mutations in *SPTLC2* in four index patients. Their functional work supports the pathogenicity of these mutations and highlights *SPTLC2* mutations as causative for HSAN-I.