

## Identification of Unexpected Related Individuals in the Human Variation Panels

The NIGMS repository includes a Human Variation Collection with genotype data from 400 individuals, all of whom were reported to be unrelated by the submitters. The individuals were derived from four geographic populations: African Americans in North America (AA, n=100), Caucasians in North America (CAU, n=100), Han Chinese living in Los Angeles (CHI, n=100), and Mexicans living in Los Angeles (MEX, n=100). Jonathan Pevsner's group, at the Kennedy Krieger Institute, obtained the SNP data for these individuals from Coriell, and performed all possible pair-wise comparisons within and between the four groups using identity-by-state (IBS) (Roberson et al., manuscript in preparation). They identified 12 partial or complete nuclear families in the data set. For six of these families they reconstructed likely pedigrees based on nuclear and mitochondrial SNPs. Also, two individuals (NA17155 and NA17163) shared all alleles across the genome, consistent with the presence of an identical sample. These two samples were independently submitted to the repository six years apart, but were derived from the same individual (D. Coppock, Coriell Institute for Medical Research, personal communication).

### The Human Variation Panel (n=400 unrelated individuals each from four groups: MEX, CHI, CAU, AA) has unexpected familial relationships

