

- **Data preparation:**

Study samples - 210 independent samples from the International HapMap II Project. The samples encompassed 30 African marriage pairs from Yoruba in Ibadan (YRI), 30 Caucasian marriage pairs European descent resided in Utah (CEU), 45 Han Chinese persons in Beijing (CHB), and 45 Japanese persons in Tokyo (JPT).

Genetic markers - All 210 samples were genotyped with both the Affymetrix Human Mapping 500K (500,568 SNPs) and Array 6.0 (906,600 SNPs), and gene expression (GE) were measured using Illumina's Sentrix Human-6 Expression BeadChip (47,289 transcript probes).

- **Eight combinations from four ethnic populations:** (1) four populations – CHB, JPT, CEU, and YRI, (2) three populations – Asian (CHB+JPT), Caucasian (CEU), and African (YRI), (3) CHB and JPT, (4) CHB and YRI, (5) CHB and CEU, (6) JPT and YRI, (7) JPT and CEU, and (8) YRI and CEU.

- **Quality Control (performed within each of the eight combinations):**

SNP quality control - SNPs were removed if: (a) genotyping call rate < 0.90, (b) minor allele frequency = 0, (c) Hardy-Weinberg disequilibrium ($pFDR < 0.05$), or (d) non-autosomal SNPs.

GE quality control - GE transcript probes were removed for: (a) non-RefSeq probes, (b) no-autosomal probes, or (c) probes without gene information.

- **Ten-fold cross-validation classification analysis (performed within each of the eight combinations)**

Step 1- Randomly partition samples in each population into ten subsets.

Step 2- Apply a forward selection procedure into an FDA to identify key SNP and/or GE markers with the highest training accuracy based on training samples (in 9 of ten subsets). The model is regarded as a candidate model.

Step 3- Calculate testing accuracy for the candidate model selected in Step 2 based on the testing samples (in the remaining subset).

Step 5- Pick up the model with highest testing accuracy or highest occurrence frequency among 10 candidate models.

Step 4- The first three steps were repeated until each of 10 subsets has been analyzed as a testing dataset.