**Title**: Price et al. (2006)

**Background**: Ancestry differences can create the appearance of false associations in genome-wide association studies (GWAS). Previously, researchers used genomic control and structured association to manage stratification by ancestry – the former by uniformly inflating association statistics, and the latter by clustering subpopulations into discrete groups. These face multiple problems, including inability to handle SNP-specific adjustment, intensive computational costs, and sensitivity to model parameters. Price et al. propose a novel method (EIGENSTRAT) to find and correct for population stratification that avoids these issues.

**Findings**: EIGENSTRAT uses principle components analysis (PCA) to extract the principle components (PCs) of greatest variance. It is presumed that these PCs, or axes, capture ancestry information that should be excised from the data. EIGENSTRAT uses the first few PCs to continuously adjust genotypes and phenotypes, and final independence is confirmed using association statistics. EIGENSTRAT is tested on simulated population stratified genotype data and a real genotype data set of 116,204 SNPs across 488 European Americans. The simulated data set had only 1 statistically significant axis (capturing 99.9% of ancestry information), while the real data set had 10. Visually, the 2D plot of PC1 and PC2 revealed discrete and continuous genetic variation across geographic subpopulations in Europe. Chi-sq statistics were collected for genomic control, EIGENSTRAT, and an uncorrected control under diverse conditions: differentiated and causal disease SNPs, in discrete and admixed populations, with varying degree of ancestry differences. EIGENSTRAT cleanly outperforms genomic control: producing fewer false positives at differentiated SNPs and providing higher power for causal disease SNPs under all conditions, in less than 0.5% of the computational time. EIGENSTRAT is largely insensitive to the number of samples, PCs applied, and inclusion of candidate SNPs.

**Conclusions**: EIGENSTRAT is a specific, robust, efficient, and effective method to explicitly model and control for ancestry differences along continuous axes of variation, and has been demonstrated to outperform previous methods on both real and simulated data sets.