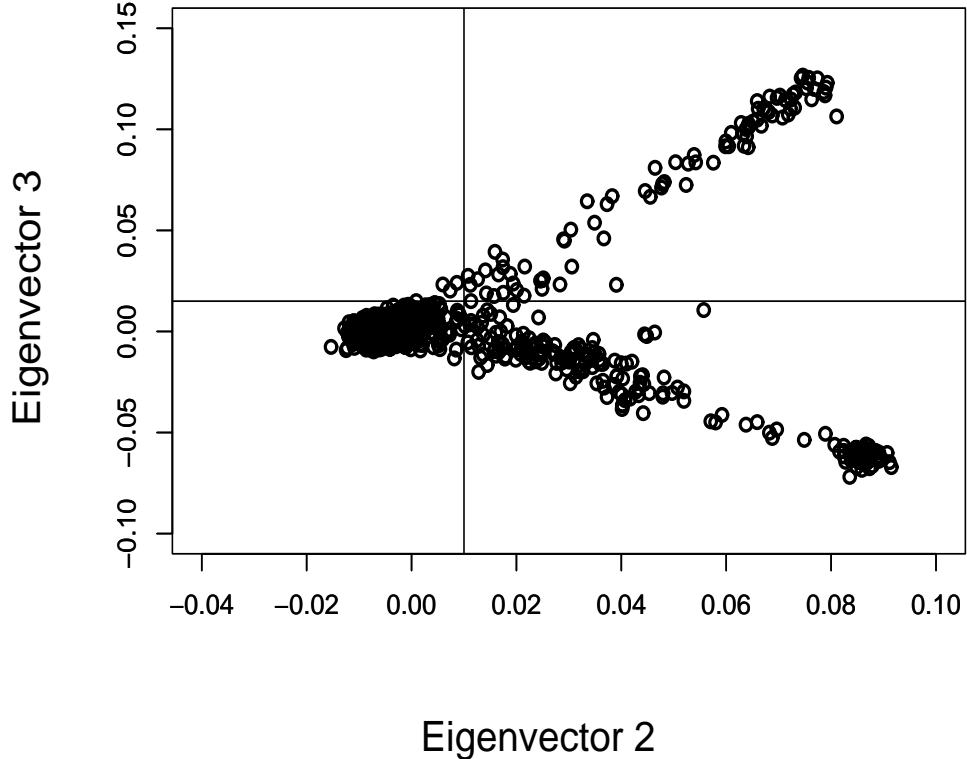


Mapping autism risk loci using genetic linkage and chromosomal rearrangements (The Autism Genome Project (AGP) Consortium)

Supplementary Figure 3. Principal component plot used to infer ancestry.



Supplementary Figure 3. Principal component plot used to infer ancestry. A simple way of inferring multidimensional ancestry for N founders of the pedigrees is to construct an N by N matrix Π of pairwise correlations of individuals i and j genotypes for all L loci. For each individual, each locus is coded as 0, 1, or 2 for the number of minor alleles carried. Then principle components can be extracted from Π that represent continuous, multidimensional estimates of ancestry for the sample (Price et al. 2006). Because correlated SNPs can affect the outcome in some settings (BD, unpublished data), we used independent or tag SNPs. Here we plot the second and third eigenvectors (E_2 , E_3) against one another. The cluster in the lower left quadrant corresponds, when we have information, to European ancestry. We apply a rule that a family is of European ancestry if all founders fall in this quadrant, 995/1168 or 85.2% of families fall in this group. Inferred ancestry was used for linkage analysis, including computing European-specific allele frequencies used for linkage analysis, for testing HW, and for relationship evaluation. Very similar clustering for European ancestry obtains from dimensions (E_1, E_2) or (E_1, E_3) (results not shown).