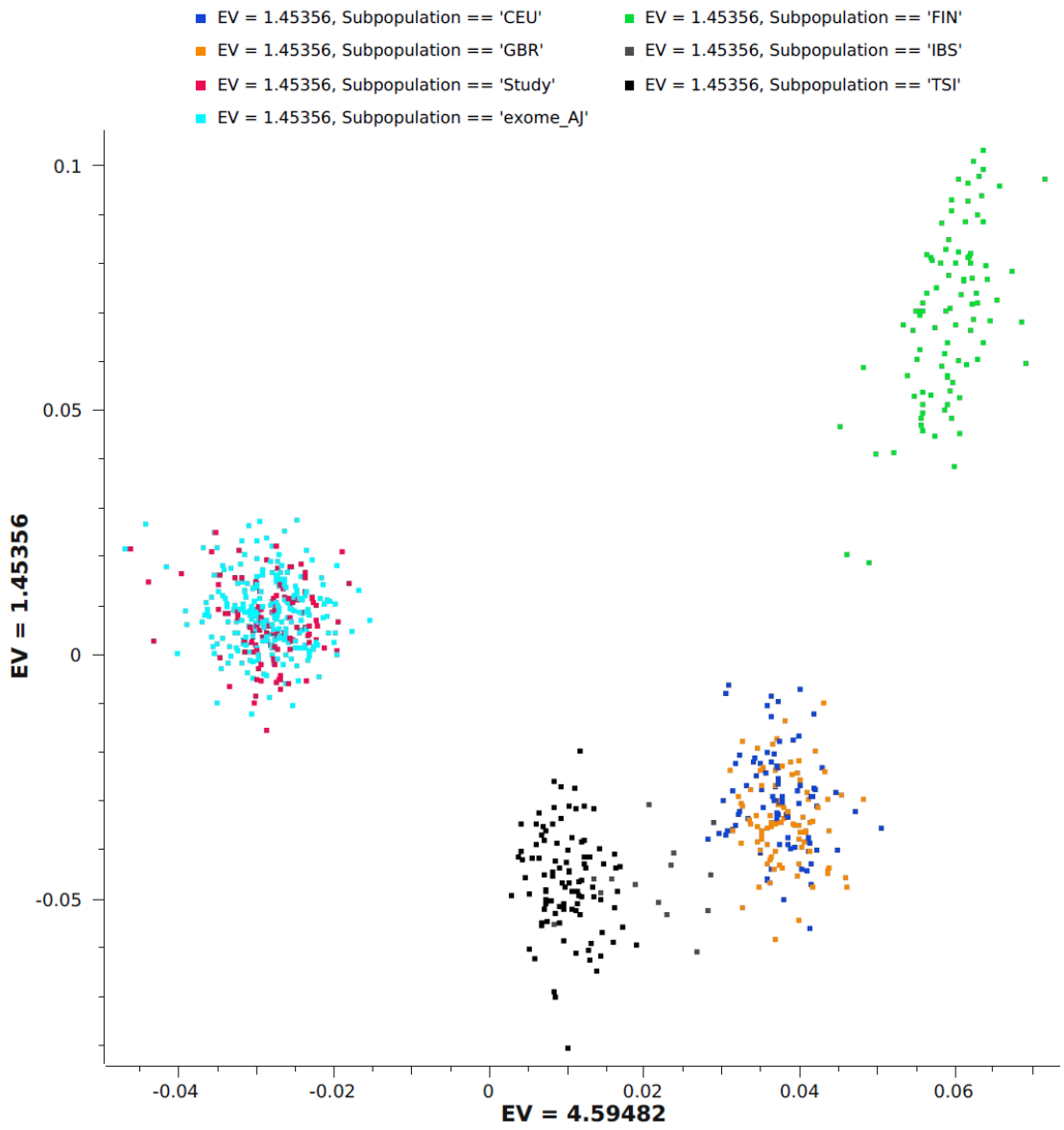
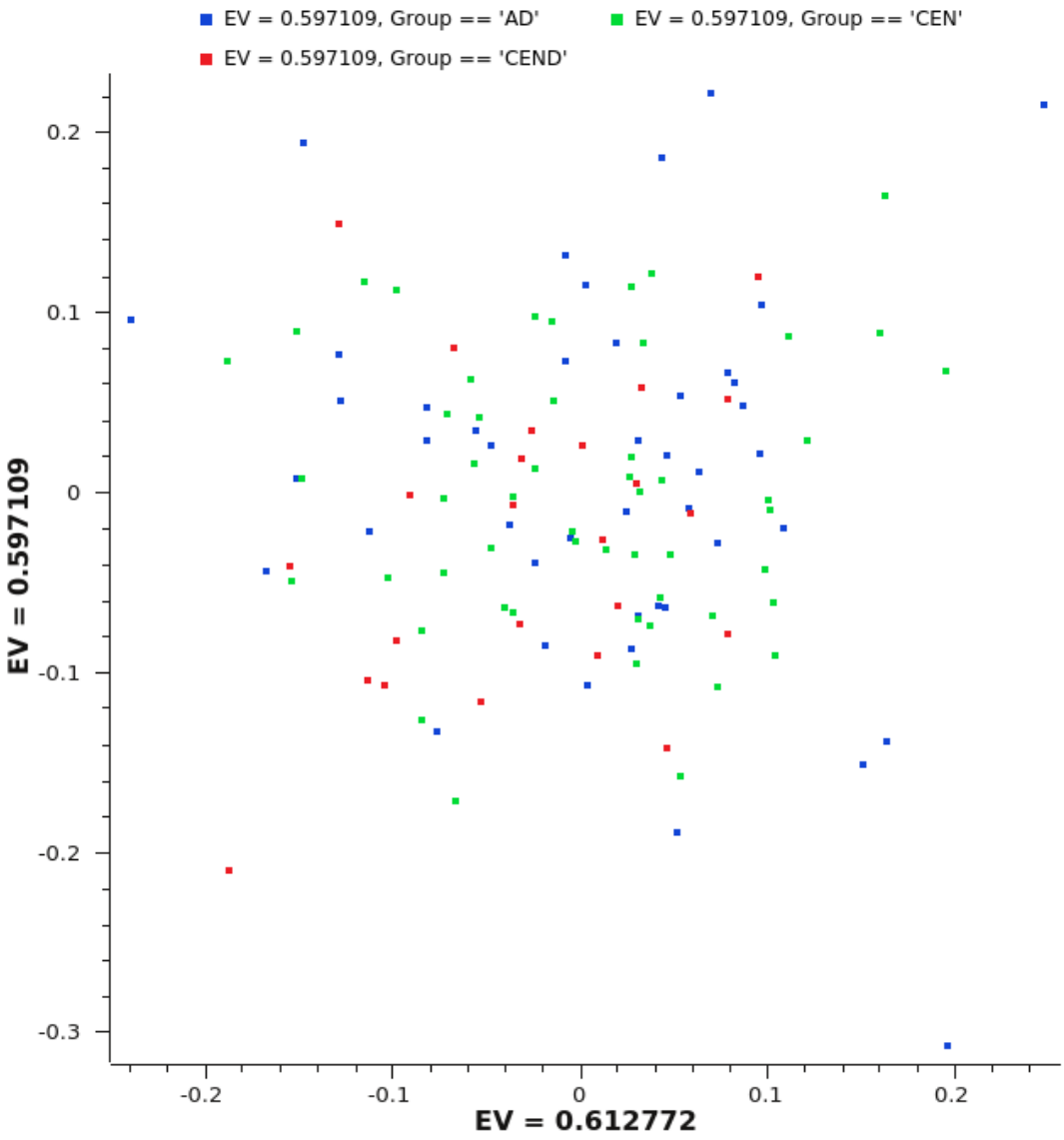
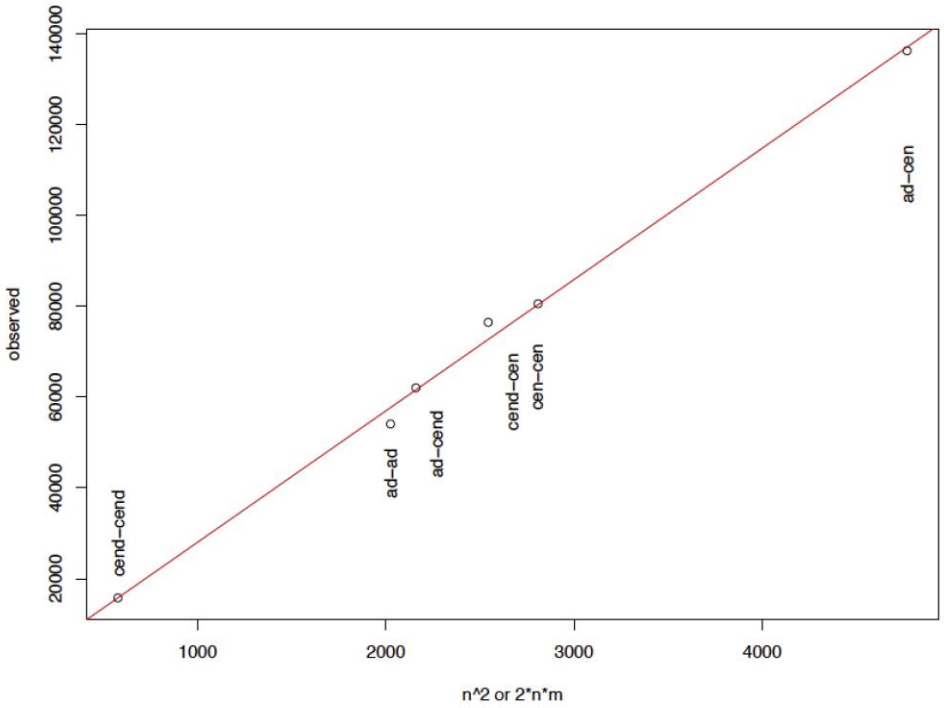


Supplementary Figure S1: Principle Component Analysis (PCA) of AJ Whole-Genome-Sequencing sample of AD and centenarian subjects (red) and younger AJ control subjects from Whole-Exome-Sequencing dataset from the T2D GENES Consortium (exome_AJ; blue). PCA was performed using LD-pruned 1000 Genomes Project Phase 1 v3 variants with MAF>=5%. CEU: Utah Residents (CEPH) with Northern and Western Ancestry; FIN: Finnish in Finland; GBR: British in England and Scotland; IBS: Iberian Population in Spain; TSI: Toscani in Italia.



Supplementary Figure S2: Principle Component Analysis (PCA) of AD (blue), centenarian (green) and demented centenarians (red) individuals. PCA was performed using 13,501 ancestry informative markers from the Illumina Immunochip.

Comparison of Doubleton and Singleton Counts in Different Groups



Supplementary Figure S3: Number of different doubleton configurations (AD-AD, CEN-CEN, CEND-CEND, AD-CEN, AD-CEND, CEN-CEND) observed versus expected. The close fit between observed and expected counts supports the homogeneity of the sample. Plot of the actual number of doubleton pairs vs. "expectation", which is the square of the group size (n^2) for pairs within the same group and 2*n*m for pairs between two groups where n and m are the two group sizes. “ad”: Subjects with Alzheiemer’s disease; “cen”: centenarians; “cend”: centenarians with nonspecific dementia.

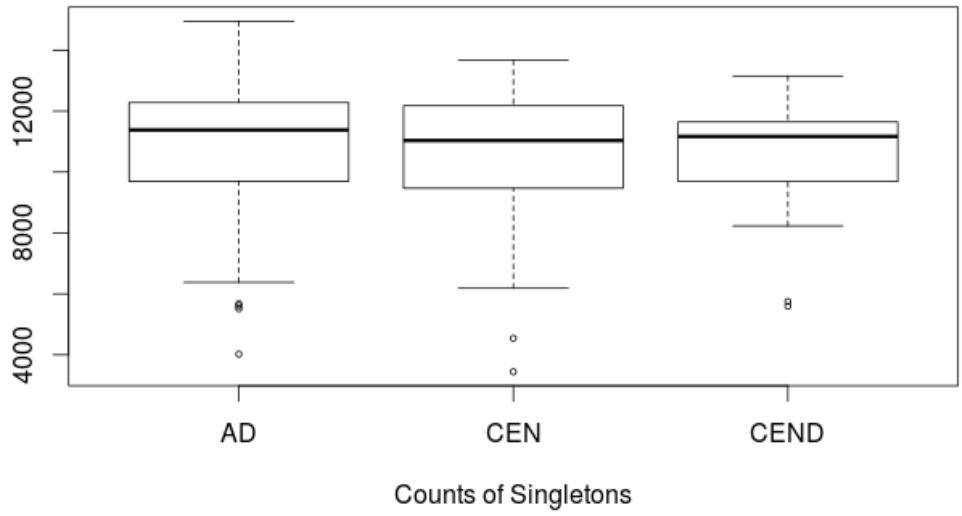
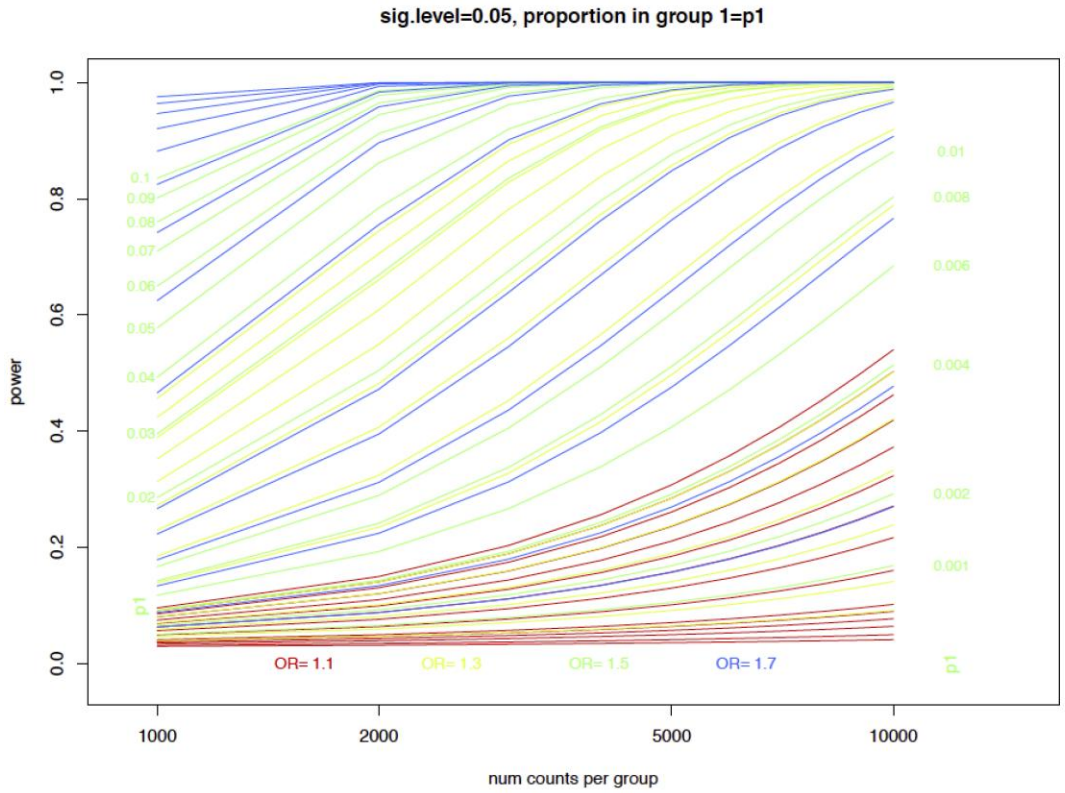


Figure S4: Boxplot of number of singletons per individual as stratified by AD, CEN, CEND. Supplementary Figure : Distributions of singletons between the three AJ sample groups. Median singleton counts for CEN, AD, and CEND are 11,032, 11,379 and 11,169 and means are 10,527, 10,717 and 10,568 respectively.



Supplementary Figure S5: Power analysis based on proportion test for variant counts For aggregation tests, the sample size may be viewed as the number of (rare, protein altering) variants instead of the number of individuals genotyped. LOF variants are a proportion of all PAVs. With LOFs being enriched in AD versus centenarians, the proportion of LOFs among the PAVs is different between the groups. We have used a proportion test power analysis as implemented in R (power.prop.test) to evaluate how power changes with sample size as defined by the number of PAVs. Based on the observed the proportions of LOFs among PAVs in centenarians versus AD, we assume the true LOF/PAV frequency being 0.06 vs 0.077, and fix power=0.8. The sample size (number of PAVs) per group required as a function of significance level is plotted for different significance levels. For instance, at sig.level=0.05, 3465 PAVs per group are required. The actual sample size observed in AD and centenarians in the sequencing data set is much larger than needed.