*Genetic data*

*Quality control*

The genome-wide genotyping was performed at University College London Genomics in 2013-2014 with the funding from the Economic and Social Research Council. This involved genotyping ELSA participants of European ancestry using the Illumina HumanOmni2.5 BeadChips (HumanOmni2.5-4v1, HumanOmni2.5-8v1.3). A detailed description of the quality control employed of the genotyped data is provided in Supplementary Information. Quality control was performed using PLINK[1], R and VCFtools[2]. Samples were removed based on call rate (<0.99), suspected non-European ancestry, autosomal heterozygosity deviation ($|F_{het}|<0.2$), and relatedness. SNPs were excluded if the minor allele frequency (MAF) was <0.01%, if more than 2% of genotype data were missing and if the Hardy-Weinberg Equilibrium (HWE) $P$-value<$10^{-4}$. Non-autosomal markers were also removed. We excluded regions that are known to contain clusters of highly correlated SNPs. These were the Lactase Gene (chromosome 6: 12578740-135837195bp), human leukocyte antigen (chromosome 2: 2550000-3350000bp), two inversion regions located on 8p23.1 (chromosome 8: 81305000-1200000bp) and 17q21.31 (chromosome 17, 40900000-45000000bp), Major histocompatibility complex (chromosome 6: 26,000,000-34,000,000bp) region, as outliers can overly influence the analyses[3, 4]. The indels and chromosome X were also excluded. .In total, 7,183 samples (96.9% of 7,412 original cohort) and 1,372,240 (61.5% of 2,230,767) variants remained after quality control.

*Genetic imputation*

To estimate genotypes that were not assayed, imputation was performed on the Michigan Imputation Server[5] running SHAPEIT for pre-phasing[6], and Minimac3 for imputation[7, 8] using the Haplotype Reference Consortium (HRC.r1-1.GRCh37)[5, 9] as the reference panel. All variants align to human genome build 19 (hg19). After imputation, we required very high

imputation quality (INFO>0.95), low missingness (<1%) for further quality control. We limited our analyses to variants genotyped or imputed with HWE *P*-value>$10^{-5}$. We further applied stringent pruning to remove markers in high linkage disequilibrium ($r^2$>0.1) and excluding high linkage disequilibrium genomic regions. In order to investigate population structure, we chose less correlated SNPs for principal components analysis. The SNP pruning was performed following the procedure: i) consider a window of 50 SNPs, ii) calculate LD between each pair of SNPs in the window, iii) remove one of a pair of SNPs if the LD is greater than 0.5, iv) shift the window 5 SNPs forward and v) repeat the procedure. Altogether, 1,083,252 autosomal SNPs remained after the SNP pruning and were used to run principal components analysis. As a result, the top 10 principal components retained to account for any ancestry differences in genetic structures that could potentially bias the results[10]. After the sample quality control, 7,179,780 variants and 7183 samples were kept.

*References*


1. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.
2. Danecek, P., et al., *The variant call format and VCFtools.* Bioinformatics, 2011. **27**(15): p. 2156-8.
3. Kunkle, B.W., et al., *Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing.* Nat Genet, 2019. **51**(3): p. 414-430.
4. Novembre, J., et al., *Genes mirror geography within Europe.* Nature, 2008. **456**(7218): p. 98-101.
5. Das, S., et al., *Next-generation genotype imputation service and methods.* Nat Genet, 2016. **48**(10): p. 1284-1287.
6. Delaneau, O., J.F. Zagury, and J. Marchini, *Improved whole-chromosome phasing for disease and population genetic studies.* Nat Methods, 2013. **10**(1): p. 5-6.
7. Fuchsberger, C., G.R. Abecasis, and D.A. Hinds, *minimac2: faster genotype imputation.* Bioinformatics, 2015. **31**(5): p. 782-4.
8. Howie, B., et al., *Fast and accurate genotype imputation in genome-wide association studies through pre-phasing.* Nat Genet, 2012. **44**(8): p. 955-9.
9. McCarthy, S., et al., *A reference panel of 64,976 haplotypes for genotype imputation.* Nat Genet, 2016. **48**(10): p. 1279-83.
10. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies.* Nat Genet, 2006. **38**(8): p. 904-9.