Appendix e-1: Supplemental information on methods

Replication AD case-control sample

For the AD case-control data derived from the Swedish Twin Register, blood samples were collected as part of the studies and genome-wide data were produced from biobanked DNA using two microarray platforms from Illumina (San Diego, USA), the Infinium HumanOmniExpress and CardioMetabochip. AD ascertainment was available for over 99.9% of the sample via linkage to total population register data. This was determined either by a registered diagnosis of AD or mixed AD/vascular dementia in the National Patient Register (covering in- and out-patient settings within the public healthcare system), in the Cause of Death Register for deceased participants (ICD-7 codes 304-305, ICD-8 290, ICD-9 290A-B and 331A, and ICD-10 F00 and G30), and/or AD-specific medication prescriptions recorded in the Prescribed Drug Register (ATC code group N06D). At the time of analyses, AD ascertainment from prescriptions data was covered until the end of 2015; ascertainment coverage from other registers ceased at the end of 2014. Additionally, participants in three of the sub-studies were screened, and clinically evaluated, for signs of dementia/AD using a similar protocol [1]. These were mostly in the sample genotyped with the CardioMetabochip array. We excluded cases that had a first record of dementia (either from register data or clinical assessment) under the age of 65 years to remove those that may have had an early-onset forms of AD from the sample. We limited the entire case-control sample to those aged at least 65 years by censoring of AD ascertainment in 2015, or at time of death. Those with diagnoses of other forms of dementia but not AD were also excluded, to avoid misclassified AD cases being present among the control samples. The derivation of the samples is depicted in figure e-1. Sample characteristics are shown in table e-1 below.

Statistical analysis

For the replication analysis, mixed-effects logistic regression models were conducted for replication analysis of AD risk according to allele dosages of *FOXO3* variant rs2153960 (using A as the coded allele) in cases and controls derived from the STR. A SNP in high linkage disequilibrium with rs2153960 ($r^2 = 1$) was directly genotyped on the CardioMetabochip, so we used this variant (rs3800229) as the independent variable in the analysis of the sample with genotyping from that platform. This was coded on its G allele, which corresponds to allele A of rs2153960. Analyses were stratified by genotyping platform used to produce SNP data on the samples. The two separate estimates of SNP-AD associations and the standard errors were then combined alongside the estimate from IGAP using a fixed-effects meta-analysis with inverse-variance weighting. Models for both samples included covariates for age, sex and five principal components derived from each genotyping platform. For the sample typed on the CardioMetaboChip, an indicator variable for the cohort study that individuals were enrolled in was also included, since the assaying was performed on samples stored from three different studies or follow-ups (this was not necessary for the individuals typed with the HumanOmniExpress, whom were all enrolled in the same cohort study). A random-effects term was also included in models for twin pair status to control for clustering where both co-twins within a pair were in the sample.

References

1. Gatz M, Fratiglioni L, Johansson B, et al. Complete ascertainment of dementia in the Swedish Twin Registry: the HARMONY study. Neurobiology of Aging 2005;26:439-447.
2. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45:1452-1458.
3. Teumer A, Qi Q, Nethander M, et al. Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. Aging Cell 2016;15:811-824.
4. Kaplan RC, Petersen A-K, Chen M-H, et al. A genome-wide association study identifies novel loci associated with circulating IGF-I and IGFBP-3. Human molecular genetics 2011;20:1241-1251.
5. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res 2015.