S4 Text. Details of imputed genotype data.

**OE:** EGCUT genotyping and imputation using an Estonian reference panel have been described previously [1, 2]. Briefly, the OE samples analyzed in our study were genotyped using Infinium CoreExome-24 BeadChip Kit (n = 163) or Illumina HumanOmniExpress (n = 131), and their data were cleaned, phased, and imputed along with other EGCUT samples genotyped using the same platforms as described before [1]. We merged the two imputed datasets into one dataset and included genotyping platform as a covariate when performing GWAS.

**MCDS:** MCDS is part of the SIGMA Consortium, for which the participants have been genotyped using the Illumina OMNI2.5 array [3], exome sequencing [4], and/or exome array [5]. Genotype imputation of SIGMA samples have been described previously [5]. Briefly, we first divided the samples into two datasets based on genotyping platforms. Dataset 1 comprised of 4,478 samples that had been genotyped by exome chip and OMNI2.5 (MCDS: n = 66). Dataset 2 comprised of 3,732 samples genotyped by exome chip, OMNI2.5, and exome sequencing (MCDS: n = 571). For each dataset, we filtered for variants with MAF > 0.001, phased the data using SHAPEIT2 [6] (v2.5), and performed imputation using the 1000 Genomes reference panel (phase 3, release June 2014) and IMPUTE2 [7] (v2.3.2). Just as we did for OE, we merged the two imputed datasets into one dataset and included genotyping backbone (i.e. Dataset 1 or 2) as a covariate when performing GWAS.

**References**


