Blood samples were collected from _N_ children and _N_ mothers. Samples were genotyped in two batches. The first batch genotyped _N_ children using the Illumina Infinium Omni2.5-8 v1.1 BeadChip. The second batch genotyped _N_ children using Illumina Multi-Ethnic Global Array. In preparation for imputation by the Michigan Imputation Server, genotypes measured by each array underwent quality control to retain only high-quality samples and variants.

Prior to any other quality control procedures, variants genotyped on the Omni2.5-8 v1.1 BeadChip were filtered to retain variants that are represented at the same chromosome and position on the most up-to-date version of the Omni 2.5 array (Omni2.5-8 v1.4). 97.8% of Omni 2.5-8 v1.1 variants were represented on the Omni 2.5-8 v1.4 BeadChip at the same genomic location. Beyond this initial filtering step, the same quality control procedures were applied to both the Omni 2.5 and MEGA data sets.

Strict variant- and sample-level call rate thresholds were applied to the Omni 2.5 and MEGA genotype data sets, respectively. Individuals with less than 5% missing genotypes were retained in the sample to be imputed. Variants with less than 2% missing genotypes across individuals were retained to inform the imputation.

For each array, samples with acceptable call rates were examined for unexpected levels of heterozygosity, as well as discrepancies between genetically predicted and phenotypic sex. Identity-by-descent statistics were calculated to check for the presence of undocumented relatedness within the EPOCH cohort. Principal components for ancestry were calculated and plotted to identify outliers relative to reported race/ethnicity.

Principal components for ancestry were calculated using the overlap of quality-controlled variants with greater than 5% minor allele frequency across all samples. Variants that met these criteria underwent linkage disequilibrium pruning to retain a subset independent variants with a maximum pairwise correlation of 0.2. The resulting principal components represent major axes of genetic variation captured by both arrays, permitting downstream adjustment for both population stratification and batch effects resulting from the use of two genotyping arrays.

From the Omni 2.5 genotype data, _N_ samples and _N_ variants passed pre-imputation quality control. _N_ samples and _N_ variants passed quality control for the MEGA.

Quality-controlled, unphased genotypes from each array were uploaded to the Michigan Imputation Server for separate imputation. Prior to imputation, each array was subjected to the Imputation Server's internal quality control pipeline. Each array was then imputed separately using the 1000 Genomes reference panel.

Principal component analysis of genetic ancestry was conducted to identify ancestry outliers and to produce covariates for use downstream genome-wide association analyses. In preparation for PCA, variants that passed the quality control standards outlined above underwent linkage disequilibrium (LD) pruning to identify a subset of pairwise-independent variants with a maximum pairwise correlation of 0.2.