**Methods**

Genotyping, quality control and imputation.

In total, DNA from 71,860 HUNT samples was genotyped using one of three different Illumina HumanCoreExome arrays (HumanCoreExome12 v1.0, HumanCoreExome12 v1.1 and UM HUNT Biobank v1.0). Samples that failed to reach a 99% call rate, had contamination > 2.5% as estimated with BAF Regress (Jun et al., 2012), large chromosomal copy number variants, lower call rate of a technical duplicate pair and twins, gonosomal constellations other than XX and XY, or whose inferred sex contradicted the reported gender, were excluded. Samples that passed quality control were analysed in a second round of genotype calling following the Genome Studio quality control protocol described elsewhere (Guo et al., 2014). Genomic position, strand orientation and the reference allele of genotyped variants were determined by aligning their probe sequences against the human genome (Genome Reference Consortium Human genome build 37 and revised Cambridge Reference Sequence of the human mitochondrial DNA; <http://genome.ucsc.edu>) using BLAT (Dunham et al., 2012). Variants were excluded if (1) their probe sequences could not be perfectly mapped to the reference genome, cluster separation was < 0.3, Gentrain score was < 0.15, showed deviations from Hardy Weinberg equilibrium in unrelated samples of European ancestry with p-value < 0.0001), their call rate was < 99%, or another assay with higher call rate genotyped the same variant. Ancestry of all samples was inferred by projecting all genotyped samples into the space of the principal components of the Human Genome Diversity Project (HGDP) reference panel (938 unrelated individuals; downloaded from <http://csg.sph.umich.edu/chaolong/LASER/>) (Li et al., 2008; Wang et al., 2014), using PLINK v1.90 (Chang et al., 2015). Recent European ancestry was defined as samples that fell into an ellipsoid spanning exclusively European populations of the HGDP panel. The different arrays were harmonized by reducing to a set of overlapping variants and excluding variants that showed frequency differences > 15% between data sets, or that were monomorphic in one and had MAF > 1% in another data set. The resulting genotype data were phased using Eagle2 v2.3 (Loh et al., 2016).

Imputation was performed on the 69,716 samples of recent European ancestry using Minimac3 (v2.0.1, <http://genome.sph.umich.edu/wiki/Minimac3>) (Das et al., 2016) with default settings (2.5 Mb reference based chunking with 500kb windows) and a customized Haplotype Reference consortium release 1.1 (HRC v1.1) for autosomal variants and HRC v1.1 for chromosome X variants (McCarthy et al., 2016). The customized reference panel represented the merged panel of two reciprocally imputed reference panels: (1) 2,201 low-coverage whole-genome sequences samples from the HUNT study and (2) HRC v1.1 with 1,023 HUNT WGS samples removed before merging. We excluded imputed variants with Rsq < 0.3 resulting in over 24.9 million well-imputed variants.

Single marker association testing

We used a linear mixed effects model approach to account for cryptic population structure and relatedness, to model the association between genotyped variants (when available for a particular position) or imputed variants (dosages) and the trait of interest, as implemented in BOLT-LMM ([https://data.broadinstitute.org/alkesgroup/BOLT-LMM/](https://data.broadinstitute.org/alkesgroup/BOLT-LMM%20/) )(Loh et al., 2016). We assumed an additive allelic effect and we only looked at variants with minor allele frequencies >0.5%. All models were adjusted for birth year, sex, genotyping batch, and PCs 1-4.

References

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8

Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A. E., Kwong, A., . . . Fuchsberger, C. (2016). Next-generation genotype imputation service and methods. *Nat Genet*. doi:10.1038/ng.3656

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., . . . Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*(7414), 57-74. doi:nature11247 [pii]

10.1038/nature11247

Guo, Y., He, J., Zhao, S., Wu, H., Zhong, X., Sheng, Q., . . . Long, J. (2014). Illumina human exome genotyping array clustering and quality control. *Nat Protoc, 9*(11), 2643-2662.

Jun, G., Flickinger, M., Hetrick, K. N., Romm, J. M., Doheny, K. F., Abecasis, G. R., . . . Kang, H. M. (2012). Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet, 91*(5), 839-848. doi:10.1016/j.ajhg.2012.09.004

Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., . . . Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science, 319*(5866), 1100-1104. doi:10.1126/science.1153717

Loh, P.-R., Danecek, P., Palamara, P. F., Fuchsberger, C., Reshef, Y. A., Finucane, H. K., . . . Price, A. L. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *bioRxiv*. doi:<http://dx.doi.org/10.1101/052308>

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., . . . Haplotype Reference, C. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet, 48*(10), 1279-1283. doi:10.1038/ng.3643

Wang, C., Zhan, X., Bragg-Gresham, J., Kang, H. M., Stambolian, D., Chew, E. Y., . . . Abecasis, G. R. (2014). Ancestry estimation and control of population stratification for sequence-based association studies. *Nat Genet, 46*(4), 409-415. doi:10.1038/ng.2924