Context:
The heritability of complex traits is a critical parameter that determines power of genome-wide association studies (GWAS) and the upper bound of polygenic scores predictions. Since ~2008 (Manolio et al., Nature), heritability estimations from genetic studies (based on common variants, largely from array-based imputed genotypes) did not match the historical values obtained from pedigree-based studies. Moreover, the contribution of non-coding rare variants to human complex traits remained poorly characterized.

Overview:
We analysed whole-genome sequence (WGS) data from 347,630 European ancestry individuals in the UK Biobank to quantify the relative contribution of 40 million single nucleotide and short indel variants (with a minor allele frequency (MAF) larger than 0.01%) to the heritability of 34 complex traits and diseases. We further mapped the rare-variant heritability to single-loci associations through GWAS-by-WGS.

Methods:
We selected 34 complex phenotypes spanning a wide range of human traits and common diseases and showing a marginally significant estimate of h^2_{PED} from 171,446 pairs of relatives in the UKB, a significant $\hat{h}^2_{\mathrm{WGS}}$) and a marginally significant rare-variant heritability estimate (P<0.05). We then estimated $\hat{h}^2_{\mathrm{WGS}}$ for these 34 traits using the GREML-LDMS method implemented in MPH v0.53.217. We performed additional sensitivity analyses of the effect on $\hat{h}^2_{\mathrm{WGS}}$ of varying the number of PCs and birthplace clusters fitted as fixed effects. Finally, we performed GWAS on ~450,000 samples in the UKB and highlighted the main characteristics specific to rare-variants associations.

Results:
On average across phenotypes, we find that WGS captures approximately 88% of the pedigree-based narrow sense heritability, that is 20% from rare variants (MAF<1%) and 68% from common variants (MAF≥1%). We identified 15 traits with no significant difference between WGS-based and pedigree-based heritability estimates, suggesting their heritability is fully accounted for by WGS data, with an important contribution from non-coding rare variants. The GWAS analyses of all 34 phenotypes identified 11,243 common-variant associations and 886 rare-variant associations and evidenced a strong colocalization between common and rare variants (independent) signals, located near genic regions. For lipid traits, >25% of rare-variant heritability can be mapped to specific loci from GWAS in the UKBiobank.