Supplementary Note

Normalized phenotypes

To check the effect of performing transformations on the non-binary phenotypes, we repeated the analysis by using a rank-based normal transformation of the phenotypes which contain at least 500 different values or categories (Because the statistical power may differ depending on the particular phenotype, the results with transformed phenotypes are also available at the GeneATLAS web). The correlation between the original and transformed phenotypes for the estimated effects and p-values were large (Supplementary Fig. 20 and Supplementary Table 11), with the exception of one phenotype, Nucleated red blood cell percentage. Approximately 96% of the individuals have a value of zero for Nucleated red blood cell percentage, whilst ~4% have a value above zero, hence when using a ranked normal transformation, the individuals with a phenotype of zero are randomly ordered and this leads to effectively random results when performing a GWAS with the transformed phenotype.

Comparisons of effect sizes and heritability with previous studies

We compared the effects sizes and Odd Ratios[1] from our results with those published in the GWAS Catalog (URLs section)[2] in cohorts of white ancestry. The comparison included 18 binary and non-binary traits. For the comparison we kept those variants from GWAS Catalog studies including a discovery cohort with at least 500 cases, and which were significantly associated ($P<10^{-8}$) both in GWAS Catalog and GeneATLAS. The effects for the quantitative traits, were normalized using the standard deviation of the trait. Overall, the results were in good agreement with those previously published (Supplementary Figs. 10-18). However, we observed an apparent bias for the effect sizes for some non-binary traits. This may be due to the combination of different effects such as the Winner's Curse[3] (Fig. 1) or the effect of performing different data transformations of the traits[4] (e.g. BMI is one of the traits for which many results were reported in specified units in GWAS Catalog, instead of some sort of standardization, and does not show a clear bias). There are a few associated variants within the binary traits which did not show a good agreement between our results and previous studies,

which may be related to different definition of the phenotypes or to differences in sample size (e.g. number of cases).

The comparison between our heritability estimations with previously published heritabilities for ten traits showed a good agreement (Supplementary Fig. 21 and Supplementary Table 12).

Advantages of large sample sizes

We investigated the increase in number of identified variants as a function of the sample size by reanalyzing the genotyped variants in subsets of individuals of increasing size using 692 non-gender specific traits to maximize power. Our results show a large increase in the level of significance achieved by associated genetic variants (defined as reaching a p-value $<10^{-8}$ in the whole cohort) as sample size increases, and a large increase in the total number of detected variants as sample size increases (Fig. 1). The plot of the number of hits as a function of sample size does not show any sign of saturation thus suggesting that increasing the size of cohorts like UK Biobank would continue to yield new discoveries.

When comparing effect sizes estimated with the whole cohort and with subsets of it of different size, the direction of the effect was always the same when the genetic variant was significant in both datasets (that is, the genetic variant was significant in the small subset of the cohort and the whole cohort). A number of the variants, which were not significant when using smaller cohorts, but were significant in the whole cohort, changed their sign (Supplementary Table 13).

We investigated how GWAS discoveries from small datasets can lead to the inflation of the genetic effects because this has important implications for the missing heritability problem or designing replication studies. We selected the five traits with the largest number of hits in the whole cohort and rerun the GWAS of genotyped variants with decreasing number of samples. Then, we selected the GWAS hits in the smallest cohort and regressed their effects on the effects estimated in the whole cohort. Our results show a clear inflation of the effect sizes of genetic variants identified with small cohorts, which is in agreement with a Winner's Curse effect[3] (Fig. 1).

Q-Q Plots and inflation factors

We computed the Q-Q plots and the inflation factor λ_{median} (λ_{median} = 1 under no inflation) for each trait using the methods described by Yang et al.[5] Standing height and hypertension were the non-binary and binary traits with largest inflation factors: λ_{median} = 2.43 and λ_{median} = 1.49, respectively (Supplementary Fig. 22). These two traits are known to be highly polygenic (Fig. 5).

As observed previously[5], the inflation factors were highly correlated with the heritability (Supplementary Fig. 19). On average the inflation factor estimated from imputed variants was larger than when estimated with the genotyped variants.

Areas depleted of associations

We checked whether there were any areas that showed few significant associations. Specifically, we computed the distance in base pairs between adjacent loci, that are lead variants clumped across all traits. We chose to consider the distance between loci rather than say variants with a significant association with any trait, in order to implicitly adjust for differences in linkage disequilibrium structure and density of genetic variants across the genome. We found only 313 regions larger than 1Mb and only 3 larger than 10Mb without an association (Supplementary Fig. 23). The 3 largest regions all spanned the centromere of a chromosome (Supplementary Fig. 24).


References

1.	Pirinen, M., Donnelly, P. & Spencer, C. C. A. Efficient computation with a linear mixed model on large-scale data sets with applications to genetic studies. *Ann. Appl. Stat.* **7,** 369–390 (2013).

2.	MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45,** D896–D901 (2017).

3.	Palmer, C. & Pe'er, I. Statistical correction of the Winner's Curse explains replication variability in quantitative trait genome-wide association studies. *PLOS Genet.* **13,** e1006916 (2017).

4.	Rodríguez-Barranco, M., Tobías, A., Redondo, D., Molina-Portillo, E. & Sánchez, M. J. Standardizing effect size from linear regression models with log-transformed variables for meta-analysis. *BMC Med. Res. Methodol.* **17,** 44 (2017).

5.	Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19,** 807–12 (2011).

6.   Canela-Xandri, O., Rawlik, K., Woolliams, J. A. & Tenesa, A. Improved Genetic Profiling of Anthropometric Traits Using a Big Data Approach. *PLoS One* **11,** e0166755 (2016).
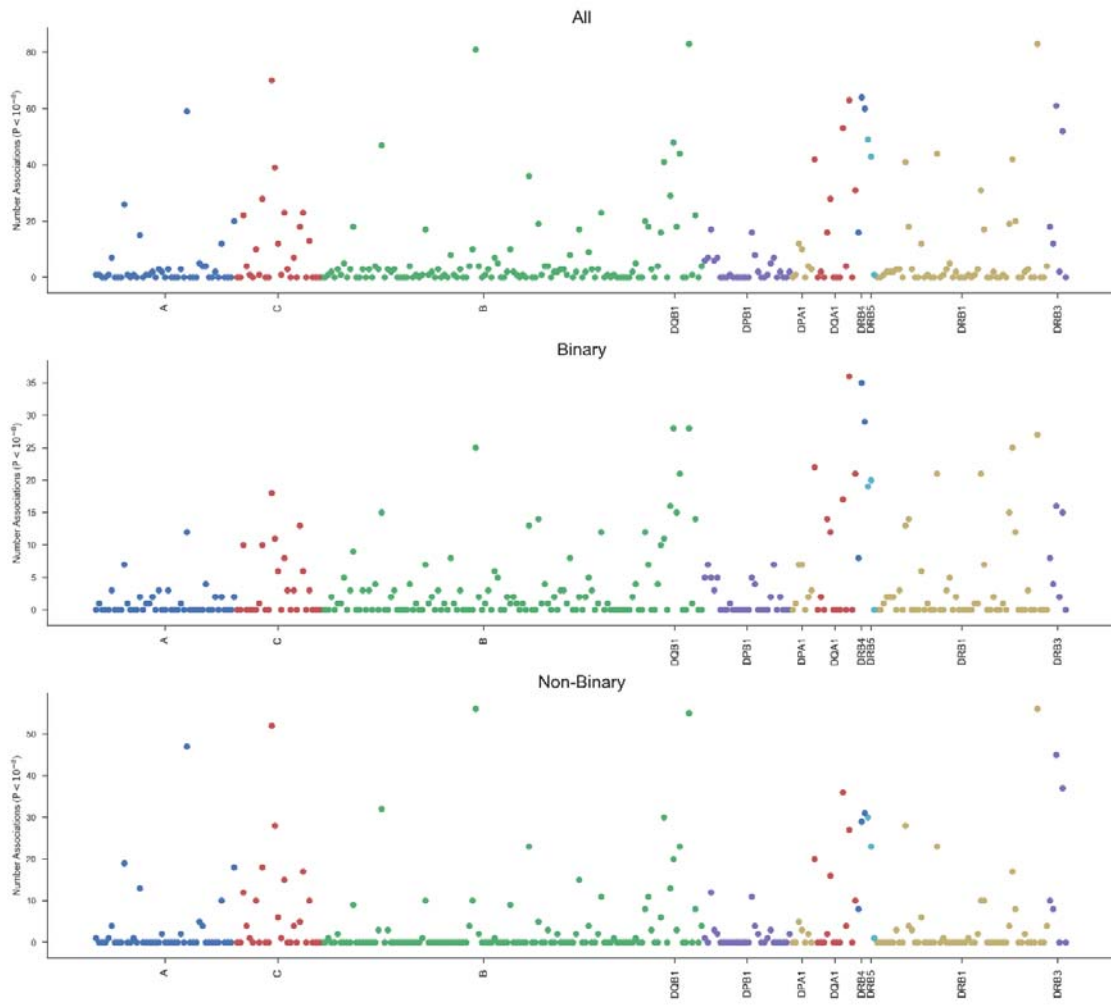
Supplementary Figures

Supplementary Fig. 1: Histograms of numbers of significant associations (two-sided t-test, P < 10⁻⁸) for each phenotype (left) and tested variant (right) for non-binary (top) and binary (bottom) phenotypes.
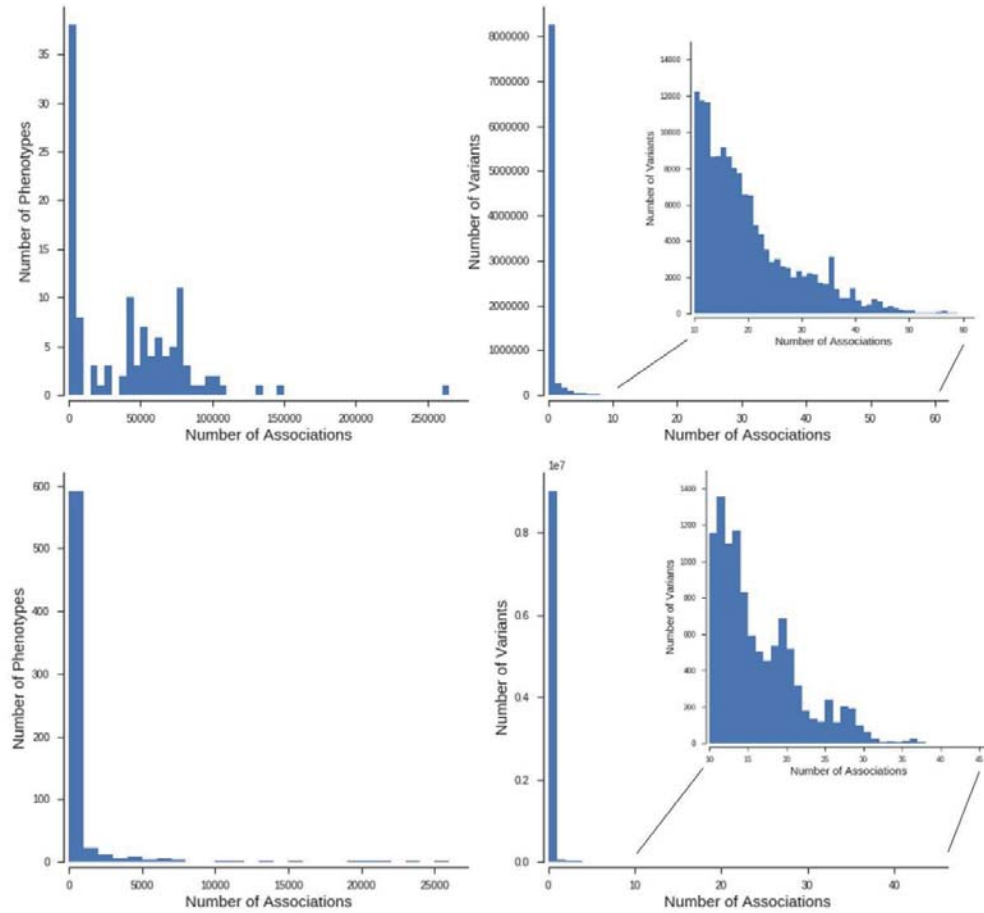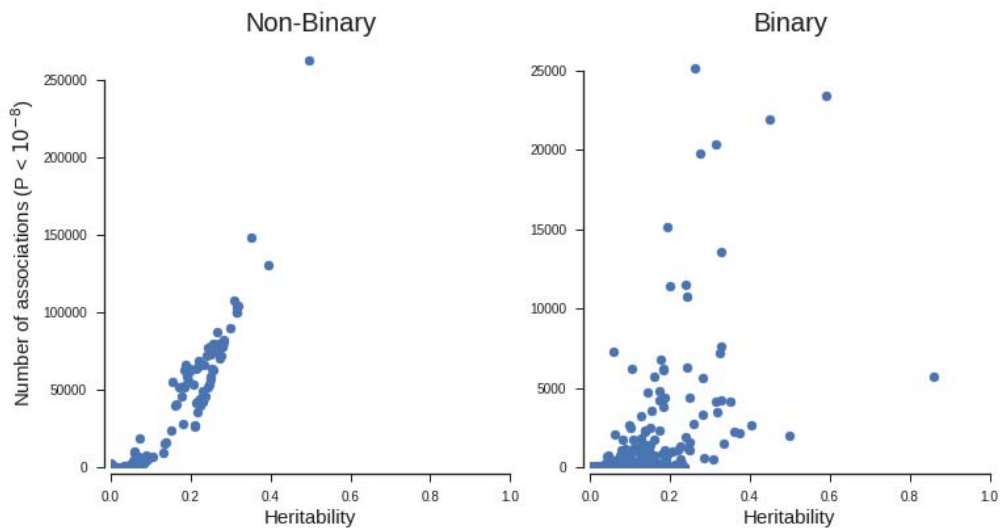
Supplementary Fig. 2: Manhattan plots of median and mean log$_{10}$ p-values (two-sided t-test) across all 118 non-binary phenotypes considered.
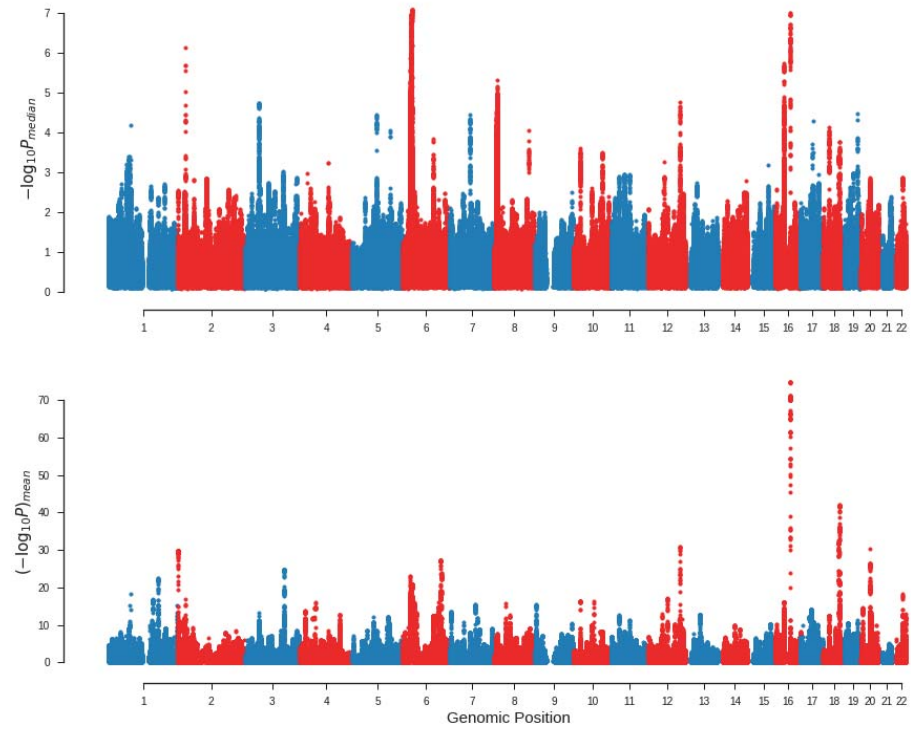
Supplementary Fig. 3: Number of significant associations (F-test, P < 10⁻⁸) at each tested imputed HLA allele for all traits, non-binary, and binary phenotypes.

Supplementary Fig. 4: Relationship between estimated SNP heritability (MAF > 5%) and total numbers of genome wide significant associations (two-sided t-test, $P < 10^{-8}$) for non-binary and binary phenotypes, respectively.

Supplementary Fig. 5: Number of significant (two-sided t-test, $P < 10^{-8}$) associations as a function of the chromosomal length covered by genotyped variants for non-binary and binary phenotypes. For chromosome 6 we excluded the HLA and surrounding 10Mb region.
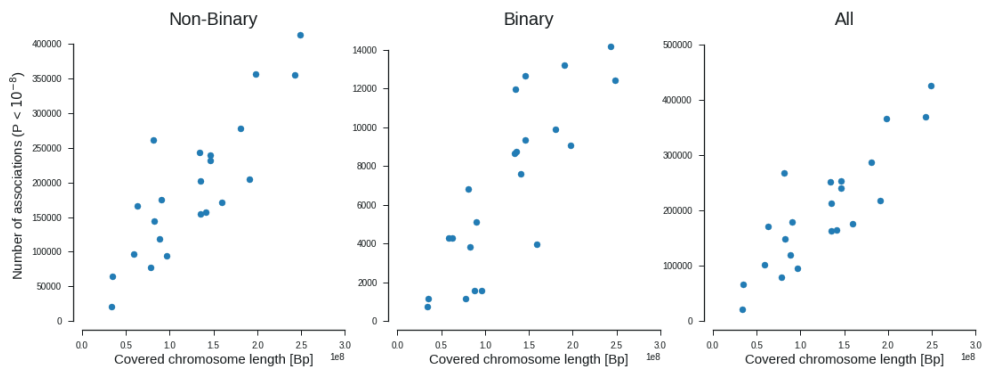
Supplementary Fig. 6: Manhattan plots of median and mean log$_{10}$ p-values (two-sided t-test) at genetic variants for the 660 binary phenotypes considered.
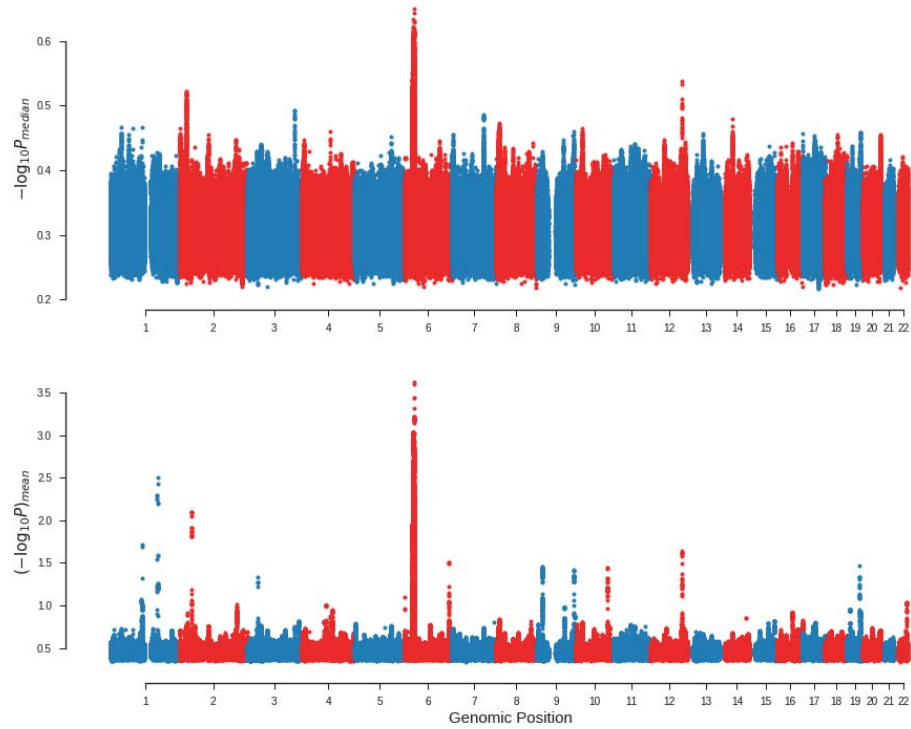
Supplementary Fig. 7: Manhattan plots for traits with an estimated heritability of zero which have at least 10 GWAS hits (two-sided t-test, P < 10⁻⁸).
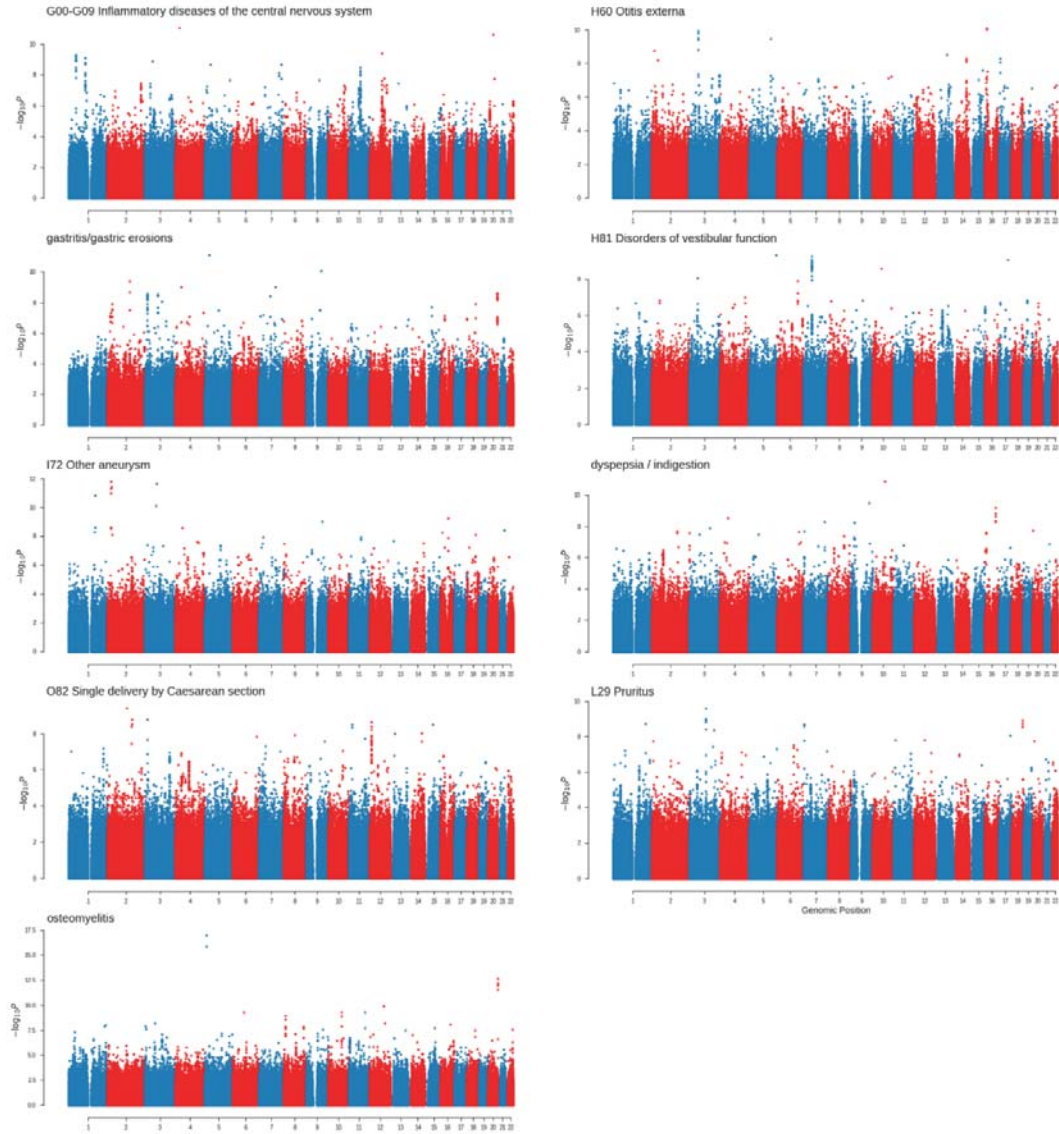
Supplementary Fig. 8: Estimated genetic (below diagonal) and environmental (above diagonal) correlations for non-binary phenotypes.
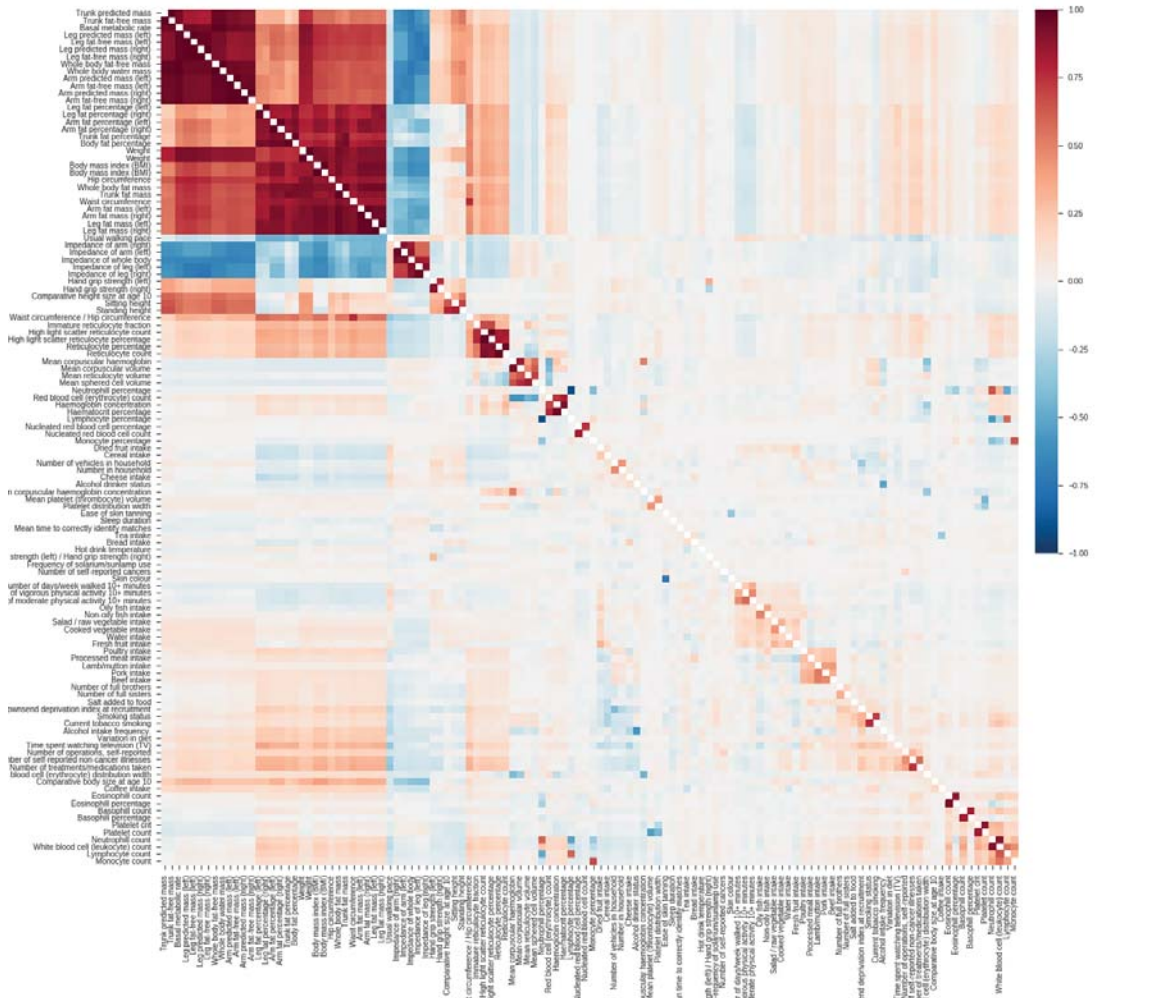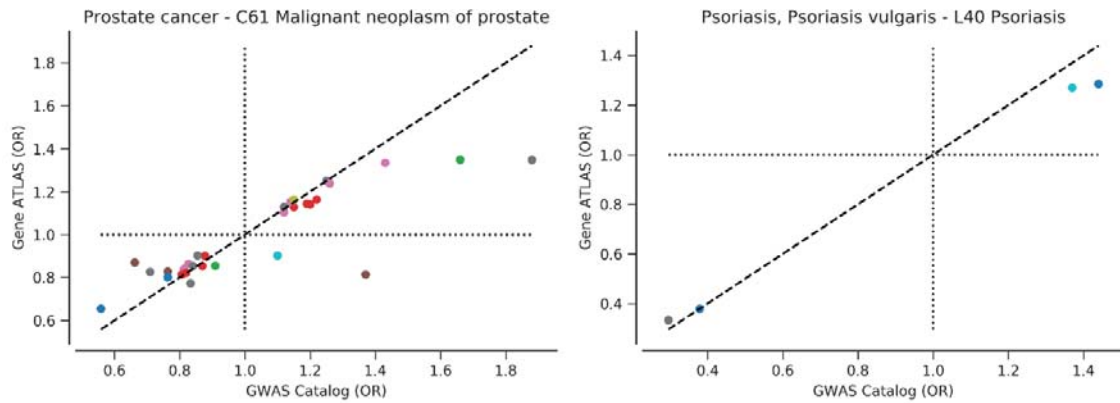
Supplementary Fig. 9: Comparison between the accuracy of phenotypic prediction obtained when training the models with 452,264 white participants of UK Biobank and expected accuracies we predicted in Canela-Xandri 2016[6]. The comparison includes Standing Height (Height), Body Mass Index (BMI), Basal Metabolic Rate (BMR), Body Fat Percentage, and Waist to Hip Ratio (WHR). Error bars indicate the 95% confidence interval for the estimate of prediction accuracy.
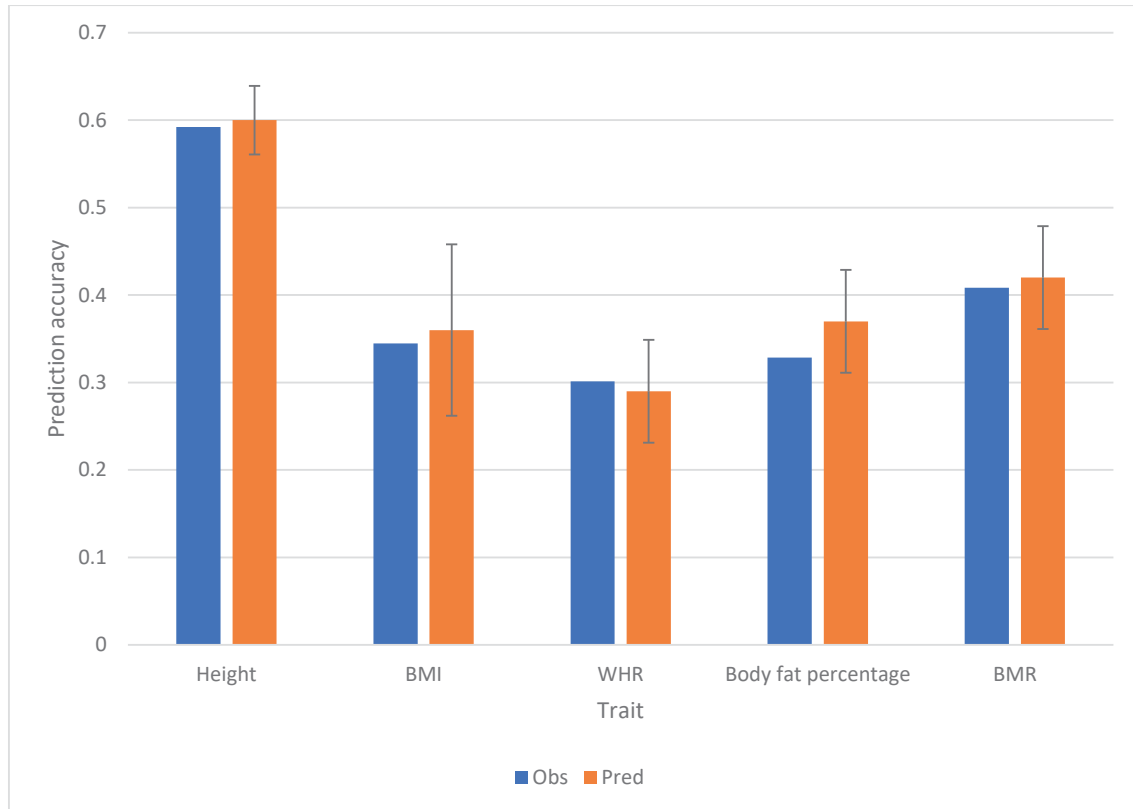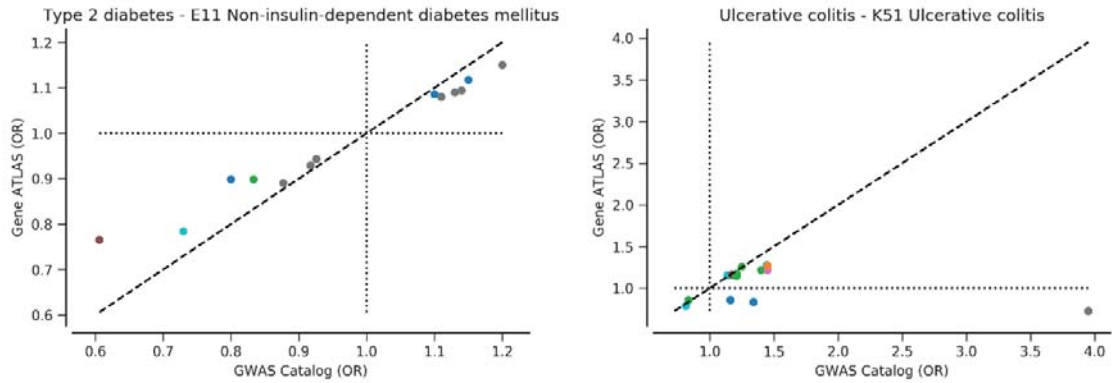
Supplementary Fig. 10: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
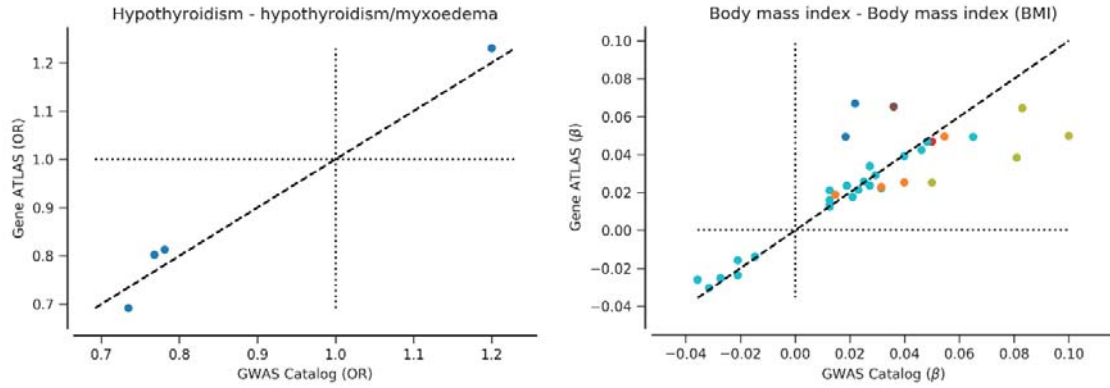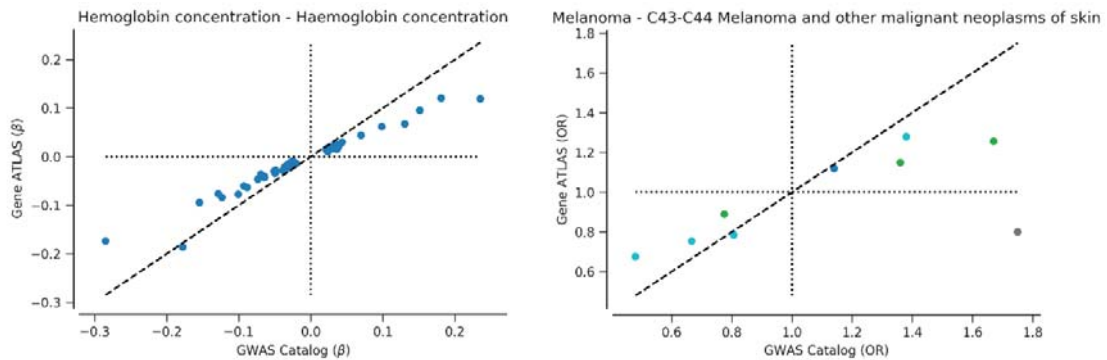

Supplementary Fig. 11: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
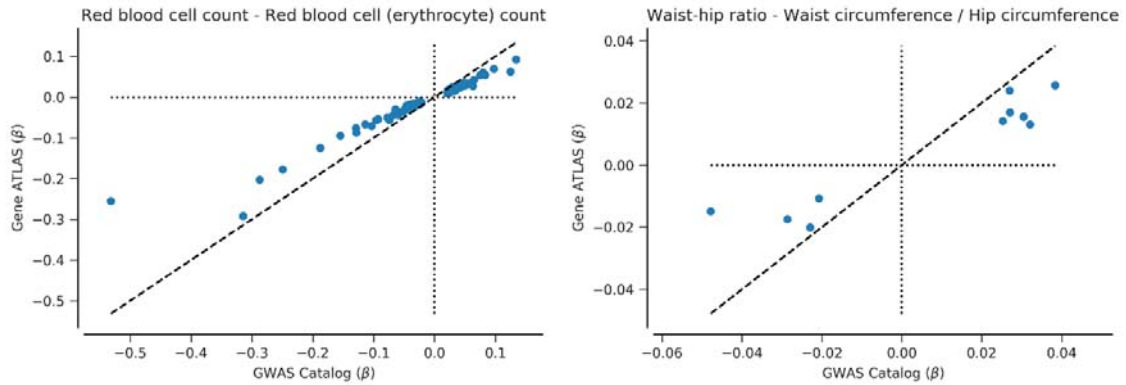
Supplementary Fig. 12: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
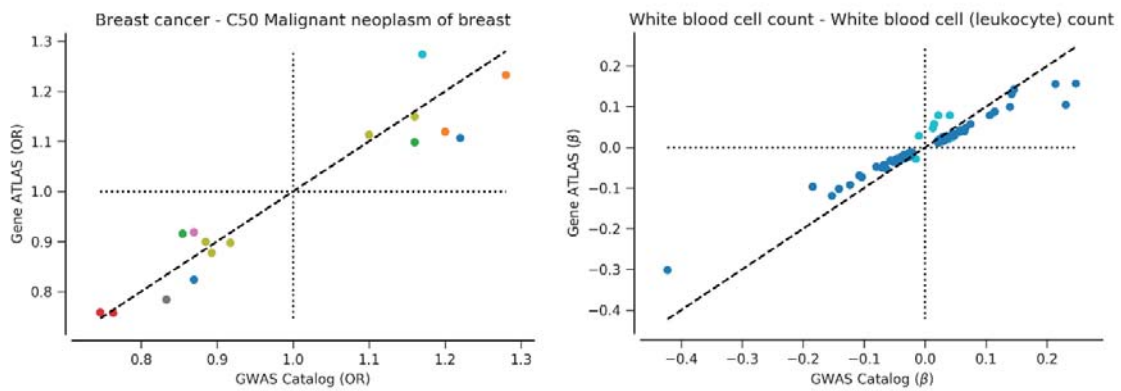

Supplementary Fig. 13: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
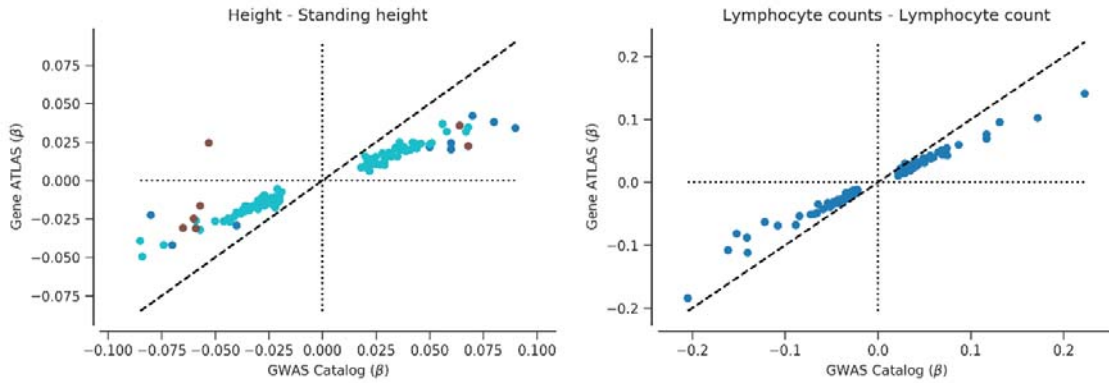
Supplementary Fig. 14: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
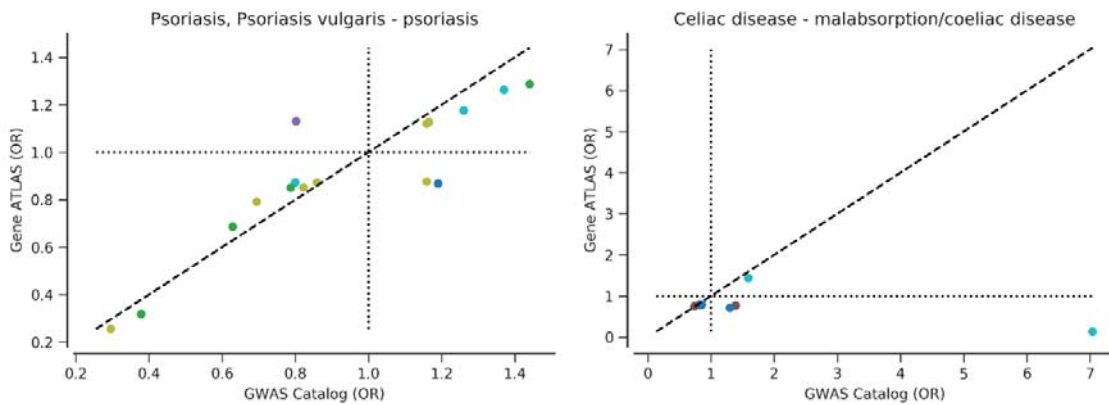

Supplementary Fig. 15: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.

Supplementary Fig. 16: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
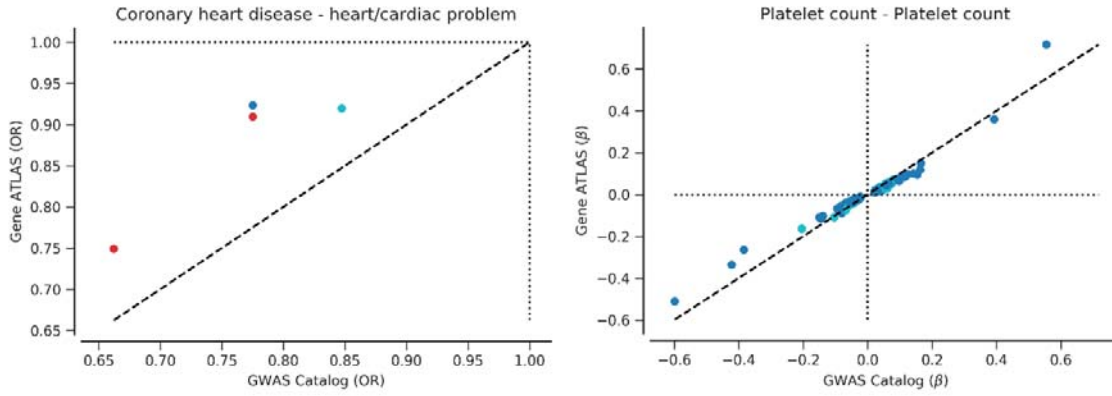

Supplementary Fig. 17: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.
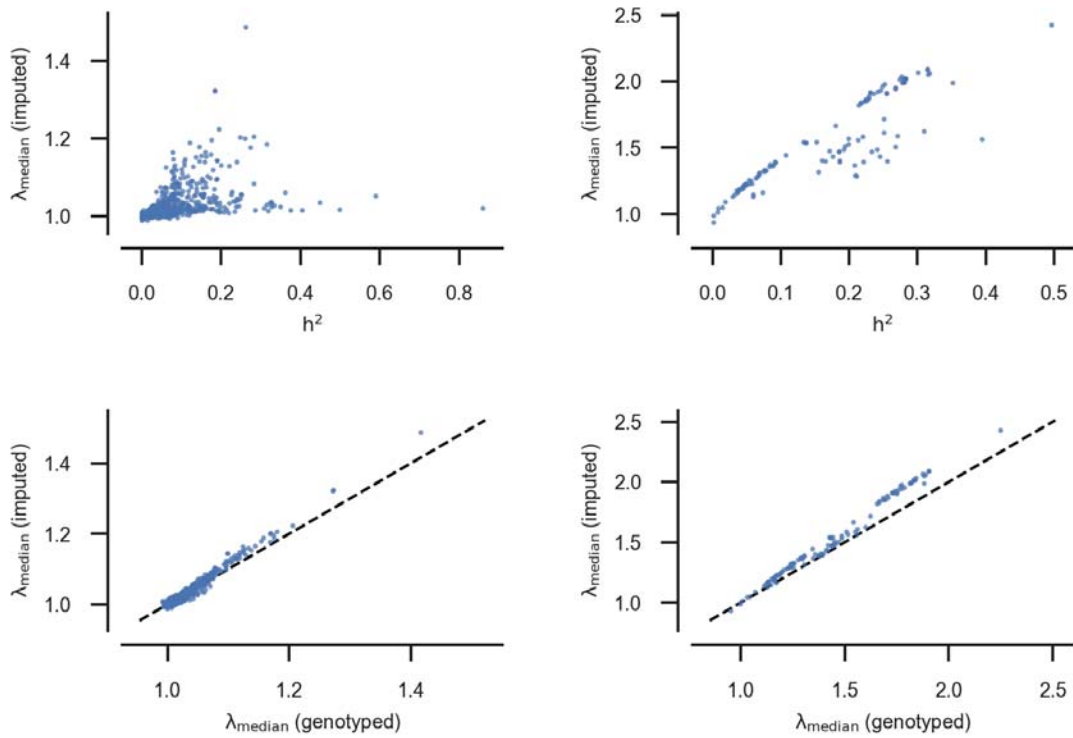
Supplementary Fig. 18: Comparison between effects sizes in GeneATLAS and previously published GWAS hits in GWAS Catalog. Each plot title indicates the name of the trait in the GWAS Catalog and the name of the trait in the GeneATLAS separated by a dash. Only genetic variants significant (two-sided t-test, $P < 10^{-8}$) in both GWAS Catalog and GeneATLAS were included. Units are in Odds Ratio (OR) or as expressed in the axes.

Supplementary Fig. 19: Relationship between inflation factor and heritability. (Top) Inflation factor λ_{median} of imputed variants as a function of the estimated heritability for binary (left) and non-binary (right) traits. (Bottom) Comparison between the inflation factors from imputed and genotyped variants for binary (left) and non-binary (right) traits.


Supplementary Fig. 20: Histogram of slopes comparing results from untransformed phenotypes and normal rank normalized phenotypes. Histogram of the Pearson correlation between (**a**) p-values (two-sided t-test) and (**b**) effect sizes obtained using normal rank phenotypes and the untransformed phenotypes for 9,113,133 genetic variants.
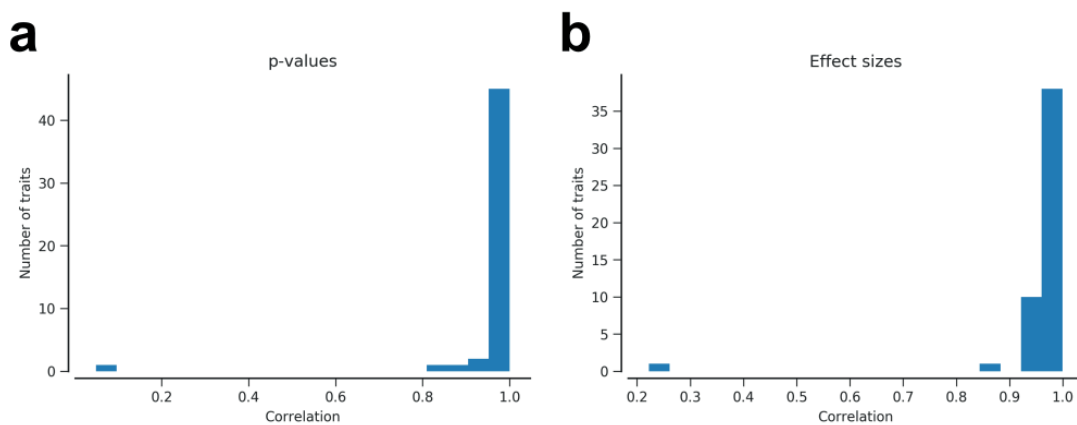
Supplementary Fig. 21: Comparison between the heritability estimates in UK Biobank and estimates in previous studies.
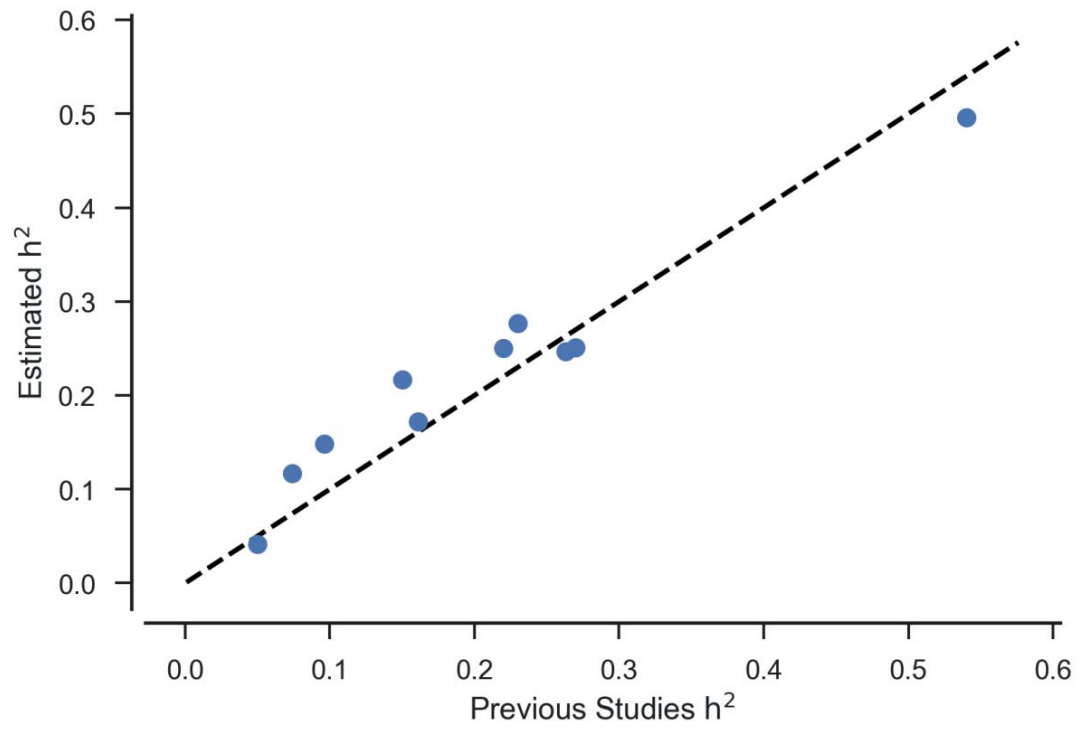
Supplementary Fig. 22: Q-Q plots (p-values obtained from Chi-squared test) for the phenotypes with largest (top) and smallest (bottom) inflation factors, λ_{median}, for non-binary traits (left) and binary traits (right). Blue dots indicate the outlier points (i.e. those larger the 5 standard deviations from the median). The dashed black line indicates the diagonal. Blue solid line indicated the regression line fitted to the non-outlier points (green points).
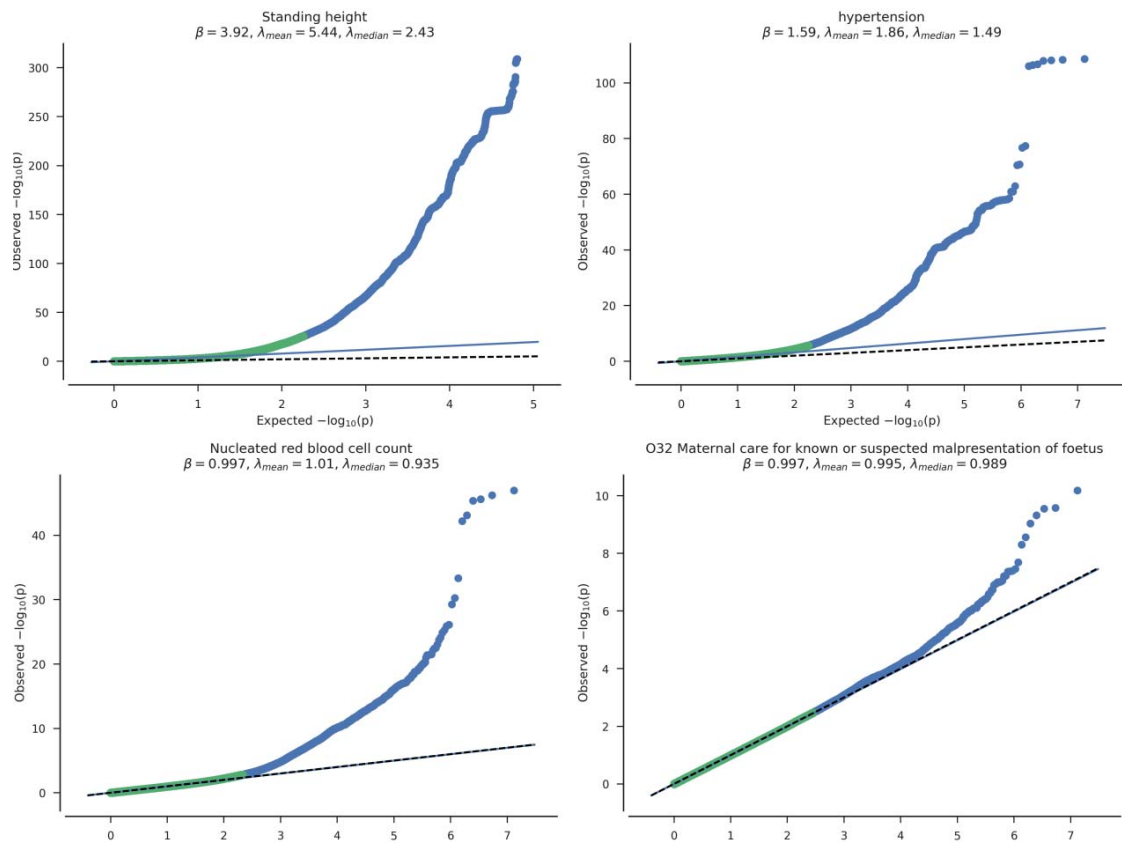
Supplementary Fig. 23: Histogram of distances between successive significant loci (two-sided t-test, P < 10⁻⁸) across all traits. On the left, the full histogram; on the right, with the y truncated to 10 to improve visibility for longer distances. There are 313 regions longer than 1Mb with no significant genetic variant and 3 larger than 10Mb.
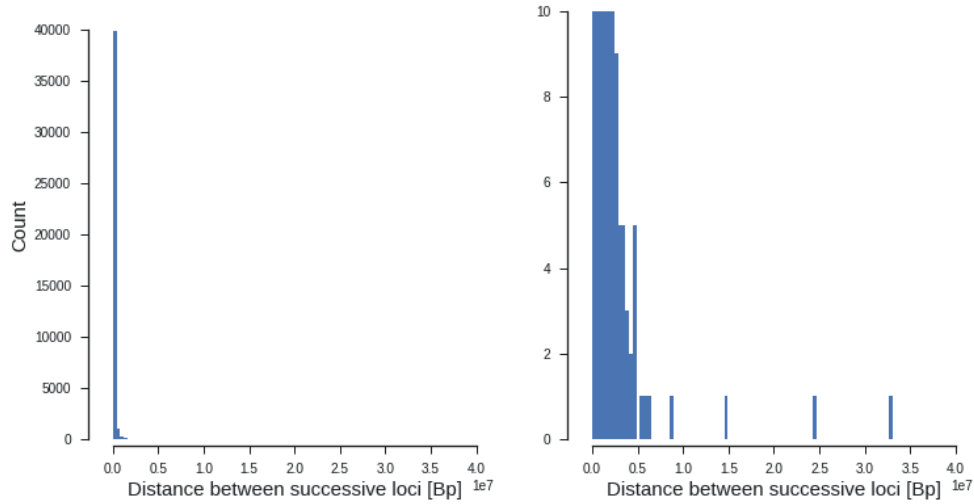
Supplementary Fig. 24: Distance between adjacent significant loci (two-sided t-test, P < 10⁻⁸) for any trait as a function of the position, indicating the size of the areas without hits.
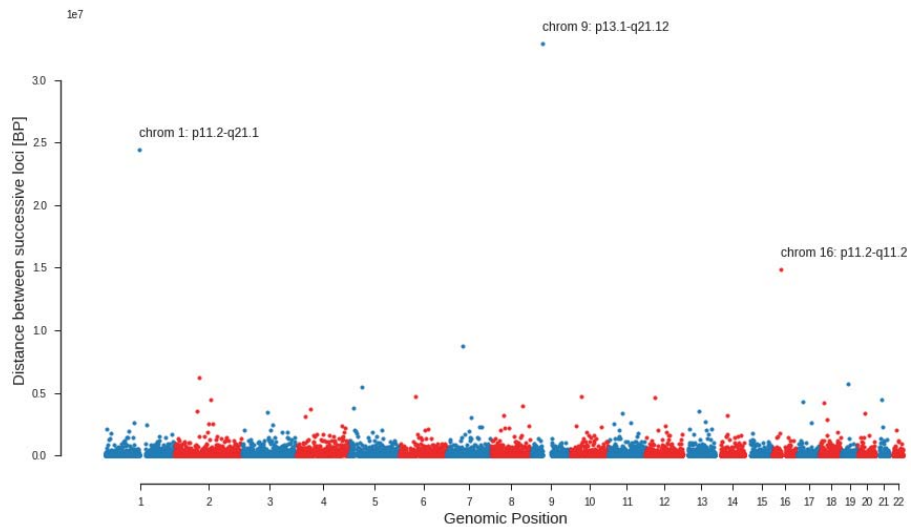
Supplementary Tables Captions

Supplementary Table 1: Summary information of considered phenotypes.

Supplementary Table 2: Total numbers and fractions of significant associations (two-sided t-test) amongst tested genotype-phenotype pairs by genomic region and category of phenotype for different significance thresholds.

Supplementary Table 3: List of lead variants for each phenotype. The p-values were calculated using a two-sided t-test.

Supplementary Table 4: Total numbers and fraction of genetic variants with at least one significant phenotype association (two-sided t-test) by genomic region and phenotype category.

Supplementary Table 5: Total numbers and fraction of genetic variants with at least one significant phenotype association (two-sided t-test, $P < 10^{-8}$) by genomic region and phenotype category for each UKB Axiom array inclusion category.

Supplementary Table 6: Number of base pairs covered by genetic variants and numbers of significant genotype-phenotype associations (two-sided t-test) for different significance thresholds in each chromosome.

Supplementary Table 7: Number and percentage of disease associated variants (two-sided t-test, $P < 10^{-8}$) with significant associations with Height and BMI at different thresholds.

Supplementary Table 8: Accuracy of prediction from genetic markers measured as AUC for binary and Pearson's Correlation for non-binary phenotypes.

Supplementary Table 9: Substitute values for special values used by UKB in the coding of non-binary phenotypes.

Supplementary Table 10: Values used to encode ordering of ordinal codings in the UKB and the phenotypes using the coding.

Supplementary Table 11: Correlation between the original and rank normalised phenotypes for the estimated effects and p-values (two-sided t-test).

Supplementary Table 12: Comparison of heritability estimates with previously published estimates.

Supplementary Table 13: Comparison of the direction of effect when estimated in the full cohort or subsets of it of varying size. Only significant genetic variants in the full cohort were considered.