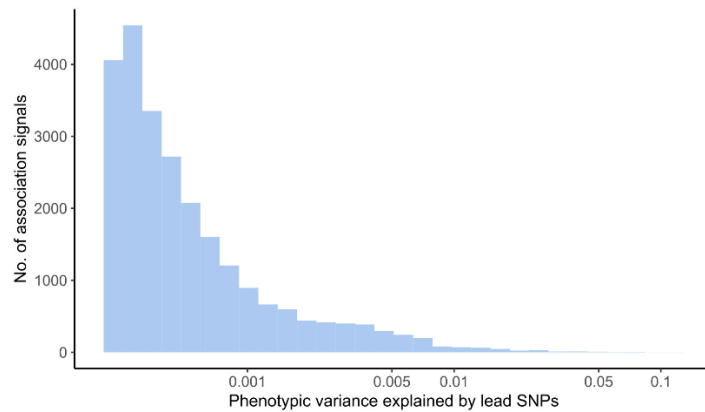
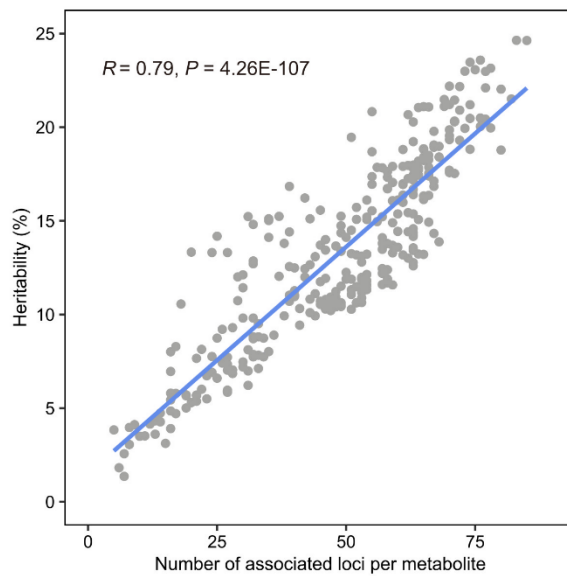


Supplementary Figures



Supplementary Figure 1. Distribution of metabolite phenotypic variance explained by lead variants.

The distribution is plotted on the log scale.



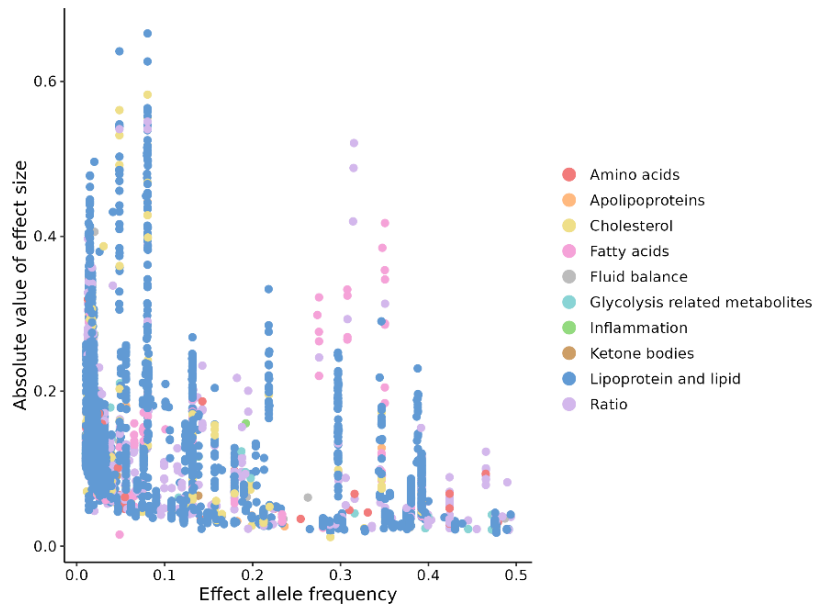
Supplementary Figure 2. Scatterplot illustrating the relationship between the heritability of metabolites and the number of associated genetic loci.

The fitted line was derived from a linear regression, with corresponding Spearman's correlation coefficient and two-sided P -value provided.

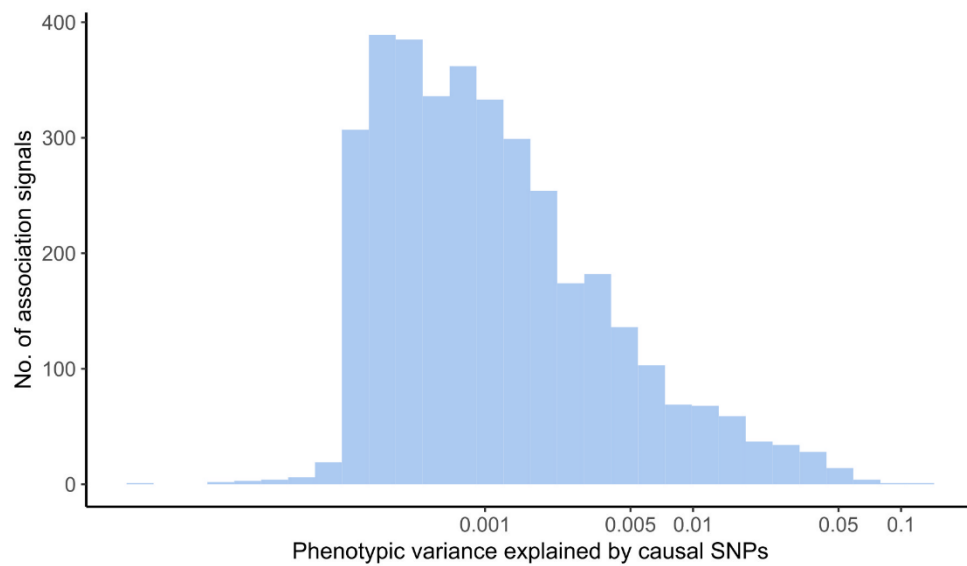


Supplementary Figure 3. Heat map illustrating the phenotypic (top left inset) and genetic (bottom right inset) correlation structure of the metabolic phenotypes.

The hierarchical clustering tree displays the results of hierarchical clustering based on the genetic correlation of the metabolic traits. The colors on the horizontal and vertical axes indicate the metabolite groups.

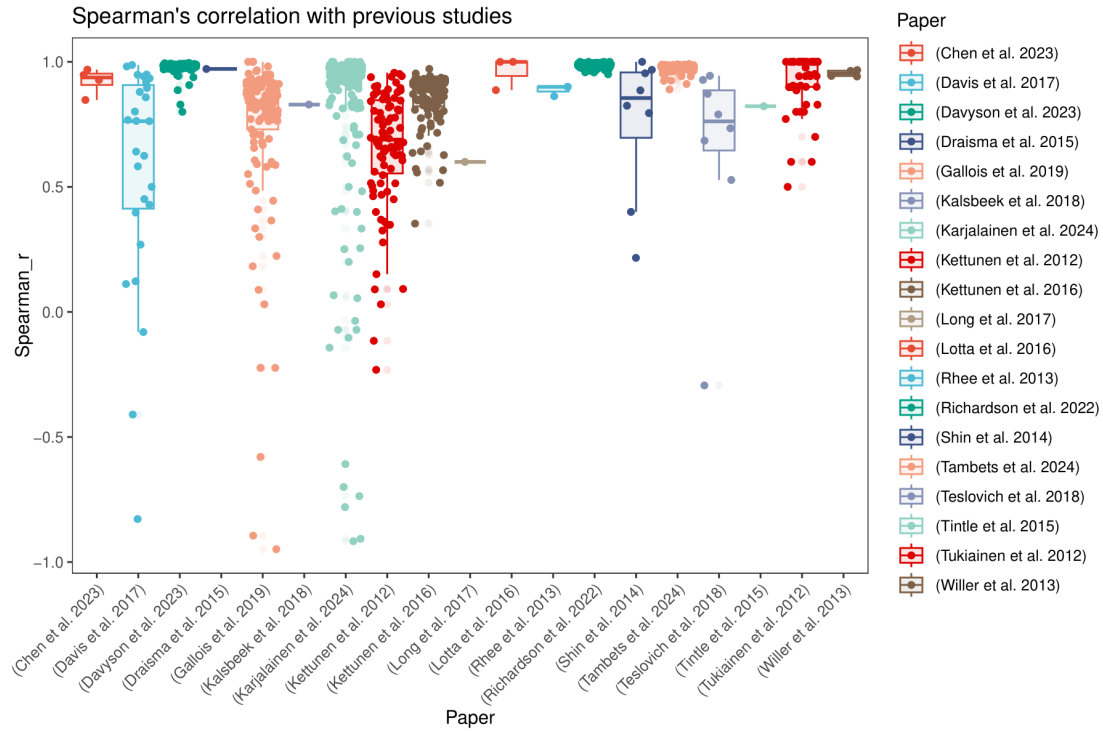


Supplementary Figure 4. Absolute effect size plotted against effect allele frequency for variants in 3,610 putative causal associations.



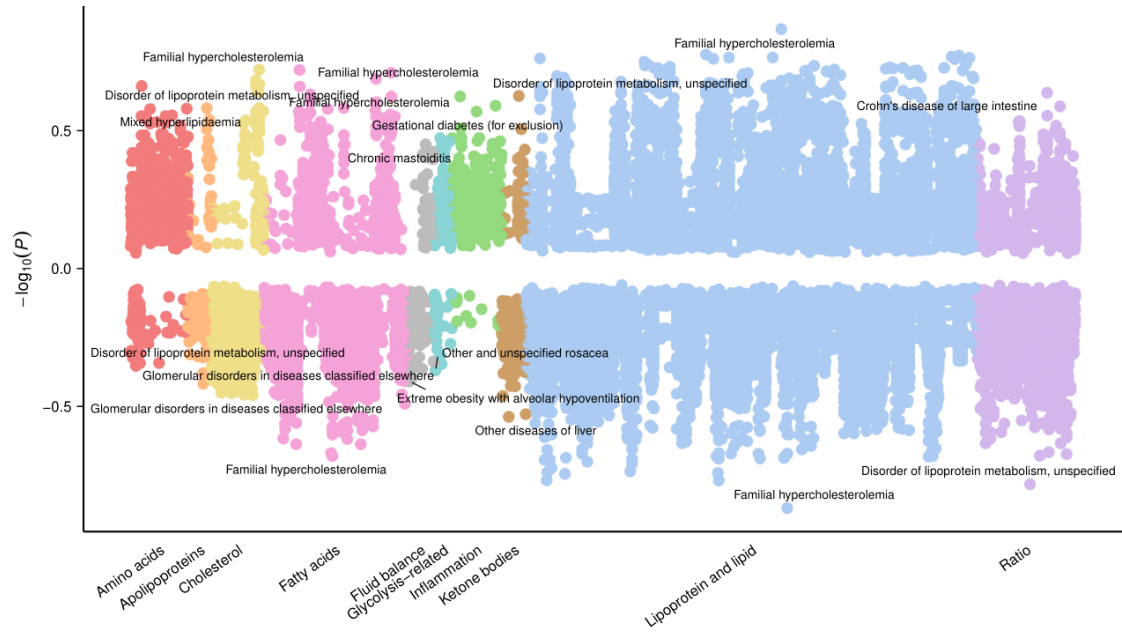
Supplementary Figure 5. Distribution of metabolite phenotypic variance explained by fine-mapped causal variants.

The distribution is plotted on the log scale.



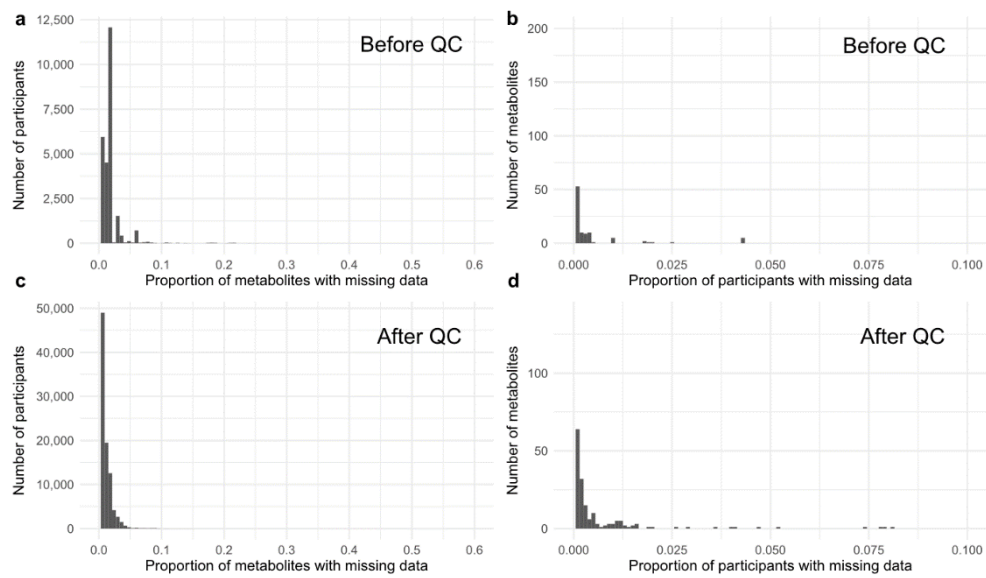
Supplementary Figure 6. Spearman's correlation of GWAS effect estimates across 19 studies

Spearman's correlation coefficients between the effect estimates in the current study and the effect estimates from 19 previous studies. Each point represents a correlation, with different colors indicating different papers. Data are presented as box plots, with the horizontal line indicating the median, boxes spanning the IQR, whiskers extending to the most extreme data point that is no more than $1.5 \times \text{IQR}$ from the box edges, and points beyond representing outliers.

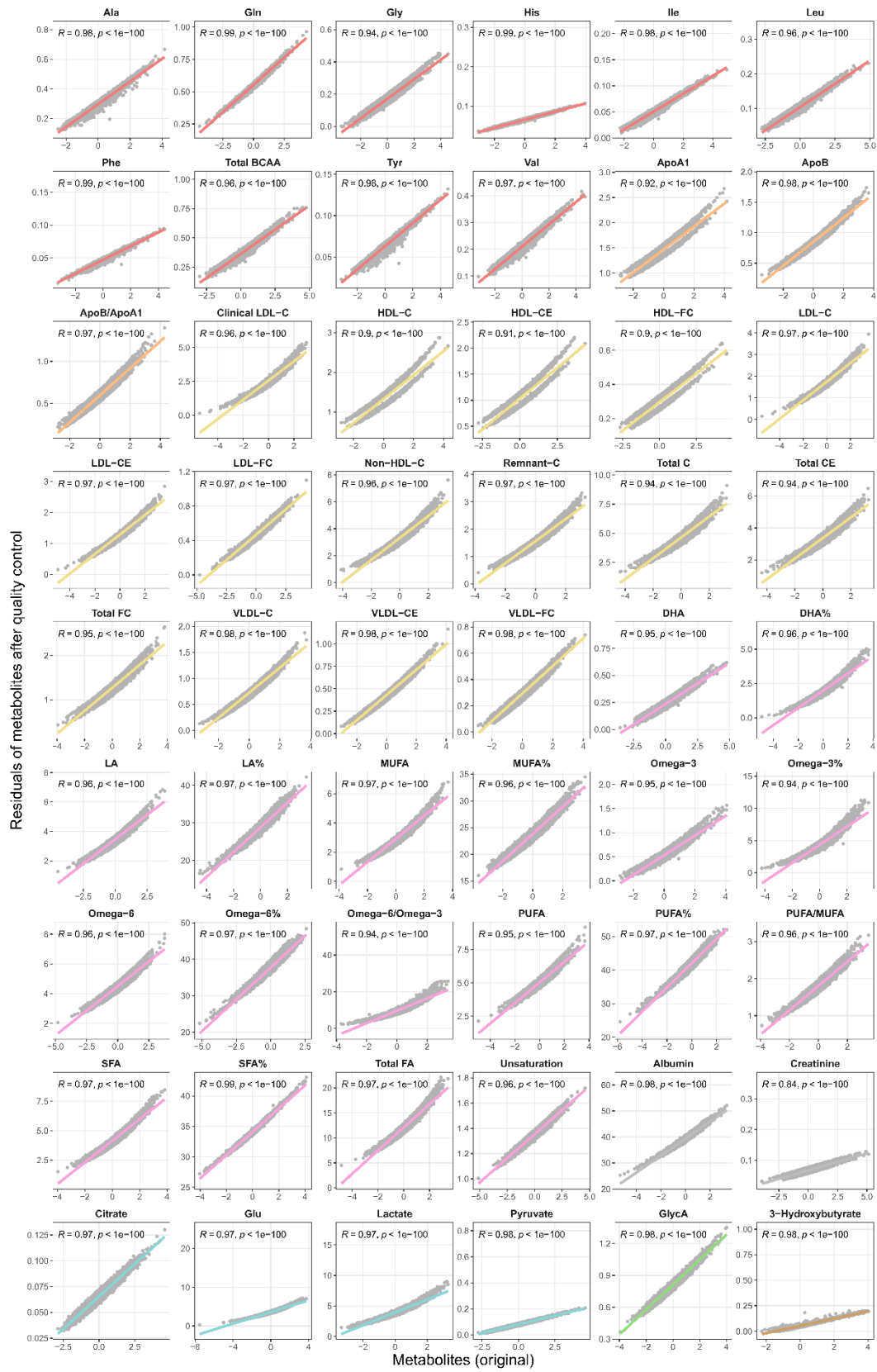


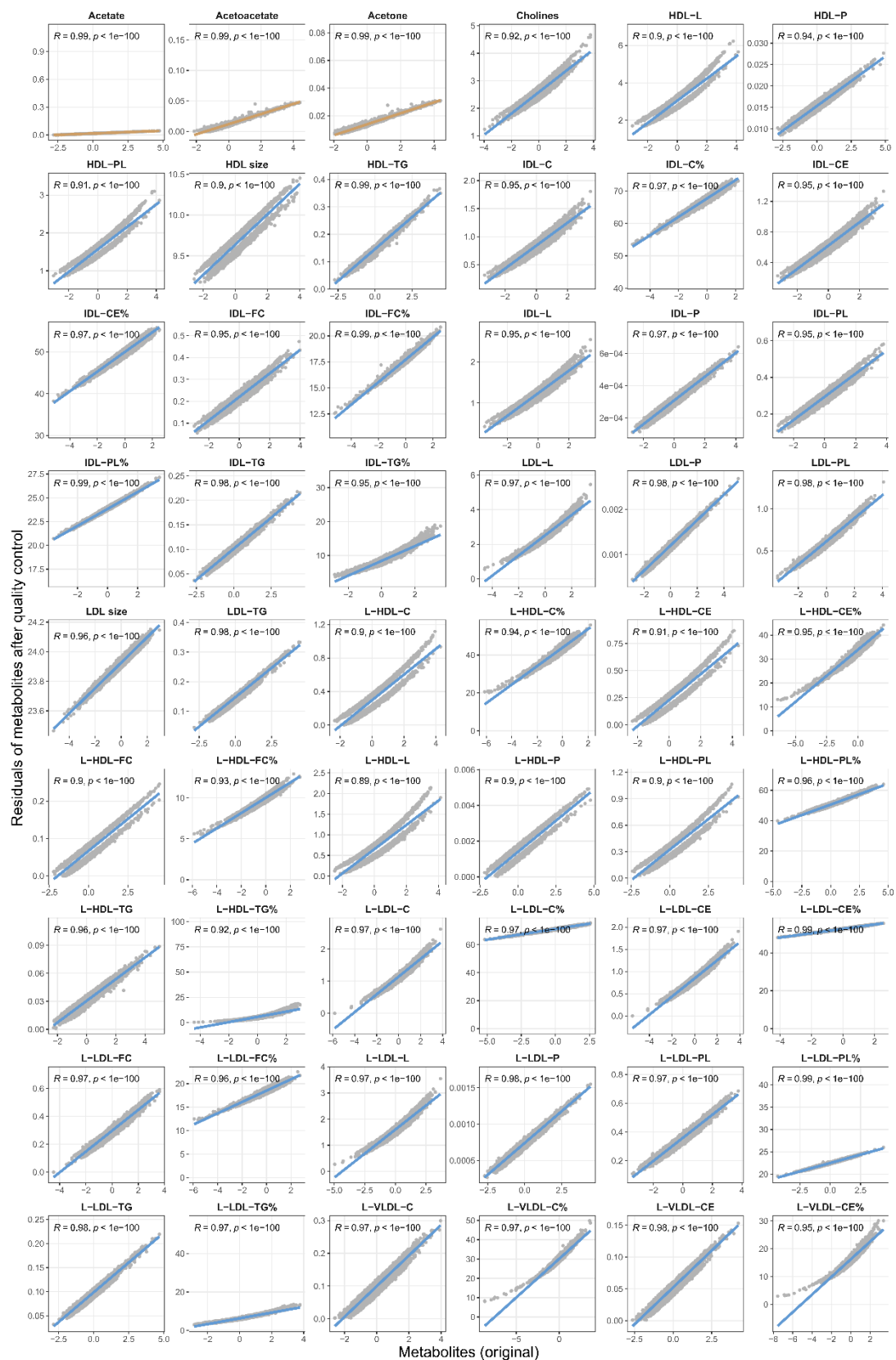
Supplementary Figure 8. Scatter plot displaying genetic correlations estimated for metabolic trait-disease pairs.

For each metabolite category, the diseases with the strongest positive or negative genetic correlations (above 0.3) are labeled. P -values are derived from genetic correlation analyses using LDSC. Only pairs with a $P < 0.01$ are included.



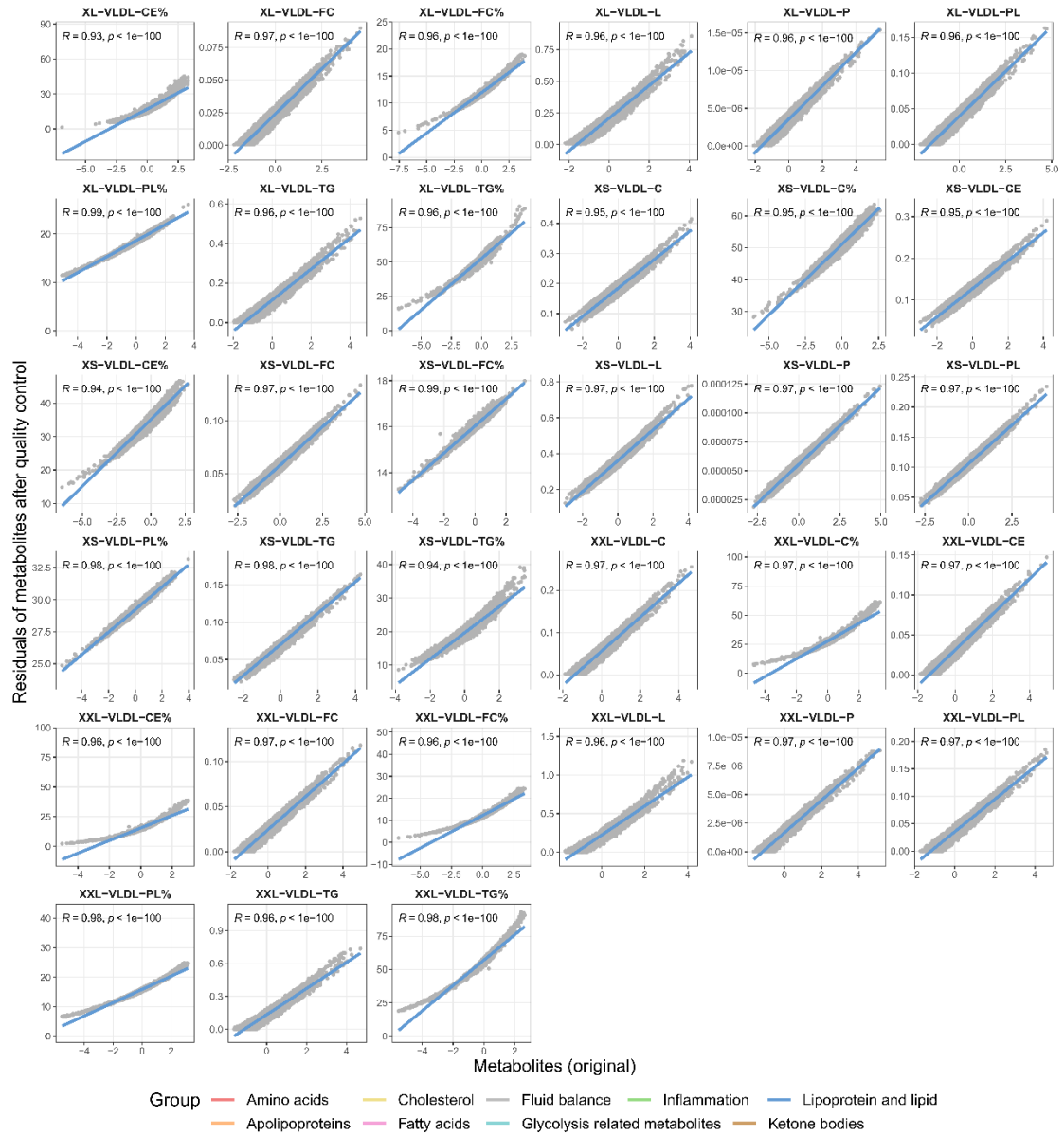
Supplementary Figure 9. Distribution of missing data across participants and metabolites before (a and b) and after (c and d) the quality control process.











Supplementary Figure 10. Correlations between each of the raw and quality-controlled metabolites.

Pearson correlation coefficients and corresponding *P*-values are labeled.