nature genetics

Supplementary information

https://doi.org/10.1038/s41588-024-01913-5

Population-specific putative causal variants shape quantitative traits

In the format provided by the authors and unedited

Contents

Supplementary Notes 3

1 Study cohorts 3

1.1 Study cohorts (BBJ and NCGG) ····· 3

1.2 Study cohort (ToMMo) ····· 3

1.3 Genotyping and imputation in ToMMo ····· 3

2 Association studies 4

2.1 Association study methods ····· 4

2.2 Replication rates in ToMMo ····· 4

2.3 Significant associations ····· 5

3 Statistical finemapping 5

3.1 Rationale of statistical finemapping ····· 5

3.2 Selection of statistical model ····· 6

3.3 Sample size and finemapping resolution ····· 7

3.4 Harmonized approach for statistical finemapping ····· 8

3.5 Quality control of finemapping results ····· 10

3.6 Threshold of PPI to regard variants as causal ····· 10

3.7 The inference of putative causal variants ····· 11

3.8 Statistical finemapping in the rare variants ····· 12

3.9 Prioritize associated variants by posterior probability ····· 14

3.10 PPI of lead variants and multiple independent signals in associations ····· 15

3.11 Associations where lead variants do not show PPI>0.1 ····· 16

3.12 Non-lead variants with PPI >0.9 ····· 17

3.13 Applying other finemapping methods ····· 17

3.14 Finemapping results in the UK Biobank data ····· 18

3.15 Potential extension of the analyses to binary traits ····· 18

3.16 Improved transferability of polygenic risk score by statistical finemapping ···· 19

4 Novel associations 20

4.1 Novel associations between phenotypes and genes ····· 21

4.2 Novel variants in known associations between phenotypes and genes ····· 28

5 Analyses of High PPI variants 29

5.1 High PPI variants and rare variants with strong effect sizes ····· 29

5.2 Non-coding variants with high PPI ····· 30

5.3 Enrichment for pathogenic variants in ClinVar ····· 30

6 Functional enrichment of non-coding variants 31

6.1 Enrichment for causal eQTL variants in the GTEx ····· 31

6.2 Functional enrichment using the UKB data ····· 32

6.3 Functional consequences of rs13306436 at 3'UTR in *IL6* ····· 32

6.4 Enrichment for target genes of Regnase-1 ····· 33

References 34

Supplementary Tables 39

Supplementary Notes

1 Study cohorts

1.1 Study cohorts (BBJ and NCGG)

We included three different datasets constructed from the contemporary Japanese population [Biobank Japan (BBJ) 1st cohort, BBJ 2nd cohort, and National Center for Geriatrics and Gerontology (NCGG) cohort] and meta-analyzed the results. BBJ is a nationwide hospital-based biobank with 12 collaborating medical institutions. The first cohort targeted 47 diseases and recruited 200,000 people between 2003 and 2013, and the second cohort targeted 38 diseases and recruited 67,000 people between 2013 and 2018 (https://biobankjp.org/en/index.html). This study included 12,098 people with available genotypes from BBJ 2nd cohort. The NCGG Biobank is a hospital-based biobank maintained by NCGG since 2012. The participants were recruited from NCGG hospital, Obu City, Aichi prefecture, and nearby medical institutes (https://www.ncgg.go.jp/english). Written informed consent was obtained from all the participants. Trait information, the number of participants, and its distributions are summarized in Supplementary Tables 1, 2 and 16.

1.2 Study cohort (ToMMo)

We additionally analyzed quantitative trait data of 53,365 subjects from the Tohoku Medical Megabank Organization (ToMMo)'s Community-based Cohort study (TMM-CommCohort 67K)[1]. These subjects were recruited from the health checkups conducted in two prefectures of Northeastern Japan: Miyagi (N = 32,459) and Iwate (N = 20,906). The data of a total of 26 phenotypes was obtained which includes alanine transaminase (ALT), aspartate transaminase (AST), basophil count (BASO), body mass index (BMI), blood sugar (BS), blood urea nitrogen (BUN), diastolic blood pressure (DBP), estimated glomerular filtration rate (EGFR), eosinophil count (EOSINO), gamma glutamyl transpeptidase (GTP), hemoglobin (HB), hemoglobin A1c (HBA1C), high-density lipoprotein cholesterol (HDLC), hematocrit (HT), low-density lipoprotein cholesterol (LDL), lymphocyte count (LYMPH), monocyte count (MONO), neutrophil (NEUTRO), platelet count (PLT), red blood cell count (RBC), systolic blood pressure (SBP), serum creatinine (SCR), total cholesterol (TC), triglyceride (TG), uric acid (UA), and white blood cell count (WBC). The number of participants available for each phenotype is shown in Supplementary Table 1. The study was approved by the Ethics Committee of RIKEN (17-17-16(16)) and the ToMMo (2019-0075-1).

1.3 Genotyping and imputation in ToMMo

The array dataset in PLINK binary format (659,326 SNPs) and the imputed genotype dataset in the Oxford BGEN format (54,041,917 variants) for 53,365 subjects was obtained from ToMMo. The procedures of

genotyping and imputation have been described elsewhere[2]. All samples were genotyped by the Affymetrix Axiom Japonica array. After quality control, autosomal variants were phased using SHAPEIT2 (v2. R837) and subsequently imputed using the ToMMo 3.5KJPNv2 reference panel[3] by IMPUTE2 (ver. 2.3.2). We conducted further quality control and excluded samples with (i) array call rate <97%; (ii) non-Japanese identified by principal component analysis (PCA) with all samples from the 1000 Genomes Phase III dataset. For variants, we excluded variants with an imputation INFO score < 0.3 from the downstream analysis. The final dataset consisted of 37,167,587 variants for 53,083 subjects.

2 Association studies

2.1 Association study methods

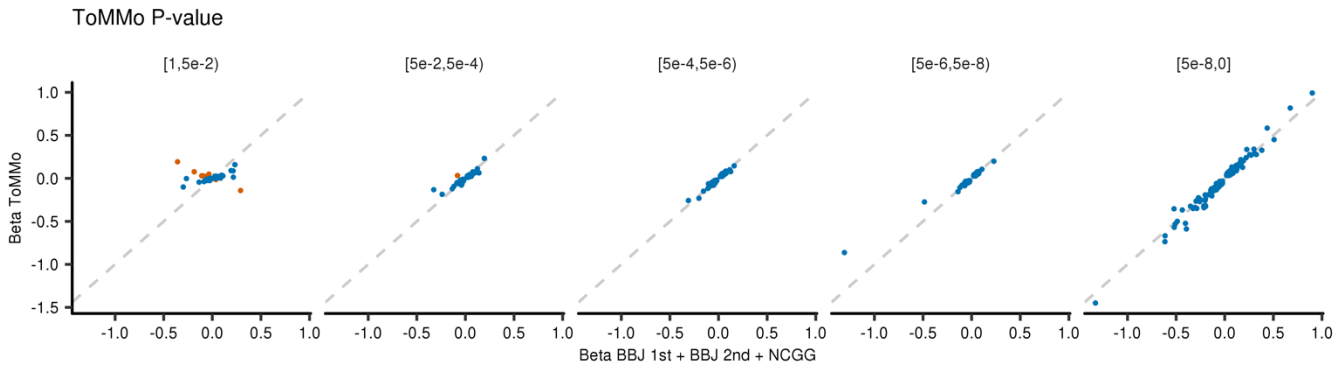
The linear mixed model analysis was performed using BOLT-LMM v2.3.956. The raw phenotypes were regressed and residualized by age, sex, and PC1-10 used as covariates. Additionally, we introduced 47 target disease statuses for BBJ 1st cohort, 38 target disease statuses for BBJ 2nd cohort, and prefecture of enrollment for the ToMMo cohort into the model. Then, the residuals were rank-based inverse normalized and used as quantitative phenotypes. After normalization, we conducted association analysis by BOLT-LMM without covariates.

2.2 Replication rates in ToMMo

We assessed the reproducibility of our primary meta-analysis (BBJ 1st, BBJ 2nd, and NCGG) using the ToMMo study as a replication cohort for 26 traits available in both studies. We observed highly concordant allelic effect sizes between primary meta-analysis and the ToMMo study. 1,505 (98.5%) of 1,528 lead variants ($P < 5 \times 10^{-8}$ in the primary meta-analysis) and 1304 (99.9%) of 1305 variants associated with nominal significance in ToMMo study ($P < 5 \times 10^{-8}$ in the primary meta-analysis and $P < 0.05$ in the ToMMo study) showed concordant allelic effect direction in these studies (SI 2.2 and Supplementary Table 17).

To assess the consistency of other variants not restricted to those with strong associations, we evaluated genetic correlations between the additional data set and meta-analysis of the three data sets by linkage disequilibrium (LD) score regression (LDSC). As a result, we found strong genetic correlations between the two results (mean r_g:0.957, SD 0.046). To note, these are consistent results of variants showing strong associations and entire common variants represented by variants in Hapmap3 [since LDSC regression only uses variants in Hapmap3 or 1000 genomes project (1KG)]. However, these results did not ensure the same genetic architecture between the two data sets, especially for rare variants. Thus, we did not include the additional data set (the ToMMo data) for statistical finemapping (see below) since the major aim of statistical

finemapping is to find convincing causal variants, including rare variants in the association results generated by the harmonized data processing.



SI 2.2 | Replication rate in ToMMo dataset

The points represent the lead variant in the primary meta-analysis (BBJ 1st, BBJ 2nd, and NCGG). The horizontal axes show beta coefficient in primary meta-analysis and the vertical axes show beta coefficient in ToMMo study. The blue and red points show the variants with concordant and discordant directions of effects in primary meta-analysis and ToMMo study.

2.3 Significant associations

We found a total of 4,423 significant associations, among which 601 were novel with *P*-values less than 5 x 10^{-8}. When we applied more strict thresholds to the results, 3,349 associations and 283 novels with $P < 3.1 \times 10^{-9} = 0.05/15{,}907{,}072$; 2,462 associations and 110 novels with $P < 5.0 \times 10^{-11} = 0.05/(15{,}907{,}072 \times 63)$. We compared the minor allele frequencies (MAF) of the lead variants driving the novel associations between the Japanese population and the European population, using the frequencies in the East Asian (EAS) and Non-Finnish European (NFE) of gnomAD. We found that 12.0% of the lead variants in the novel loci were not found in NFE and were either EAS- or Japanese-specific (not found in gnomAD EAS data either). Additionally, 18.1% were very rare (MAF<1%), and 10.0% were rare [1%≤ MAF <5%] and more frequent in EAS (median 47.5 times and 5.7 times, respectively). The remaining 60.0% of the novel lead variants with MAF ≥ 5% in NFE tended to be more frequent in EAS (median 1.46 times and median allele frequency difference 9.8%). These results support that the current study taking advantage of large-scale Japanese data could identify EAS-specific or EAS-frequent associations.

3 Statistical finemapping

3.1 Rationale of statistical finemapping

Genome-wide association studies (GWAS) are increasing in their study size and resolution. High-dense genotyping based on the combination of high-performance microarray genotyping and genotype imputation with a large-scale, high-depth imputation reference panel enabled us to directly observe significant causal associations in a genome-wide fashion. However, pinpointing such putative causal variants within the extensive list of variants with statistical significance is challenging due to the nature of the linkage between germline variants and multiple signals in the loci. To tackle these challenges, LD-informed Bayesian statistical finemapping (after this statistical finemapping) has actively been developed[4]. Taking advantage of the densely imputed genotypes and their LD information, we applied statistical finemapping in this study. We summarized the advantages of statistical finemapping and our approaches below.

Historically, statistically significant signals detected by GWAS were annotated by the variants with the highest statistical significance (lead variants). However, such variants are not always functional but often are tagging another/other causal variants; therefore, we cannot fully interpret the GWAS associations only by functional annotations of lead variants. To interpret the association signals in GWAS more flexibly, the previous studies have applied either of two approaches: the credible set estimation assuming a single causal variant[5], and the conditional analysis assuming multiple causal variants[6,7].

The credible set approach provides a small set of variants that could be causal in the locus assuming a single association. The credible set approach increases the sensitivity of detection of candidates of causal variant allowing fluctuations in test statistics. However, the traditional credible set approach could not handle multiple, independent causal variants in a single locus.

On the other hand, conditional analysis provides a set of variants with independent associations in the locus. While traditional conditional analysis is still a popular approach, the variant prioritization focusing on the strongest conditional *P*-values in each iteration would have a high chance of missing true causal variants due to the fluctuation of the statistics arising from various reasons including polygenic architecture (described below). Here, statistical finemapping integrates the strengths of these two powerful approaches.

Considering the existence of multiple signals, statistical finemapping estimates a tractable measure of causality – the posterior probability of inclusion (PPI) for each variant in the locus. PPI is particularly important when we want to detect more interpretable, functional associations such as coding variants with low frequency under the selection in the population. Lower frequencies are detrimental in achieving strong nominal/conditional *P*-values than common variants. Moreover, due to the polygenic nature of complex traits, common variants have the advantage over rare variants in obtaining strong *P*-values. Even under such condition, statistical finemapping capture the independent, causal associations driven by rare variants.

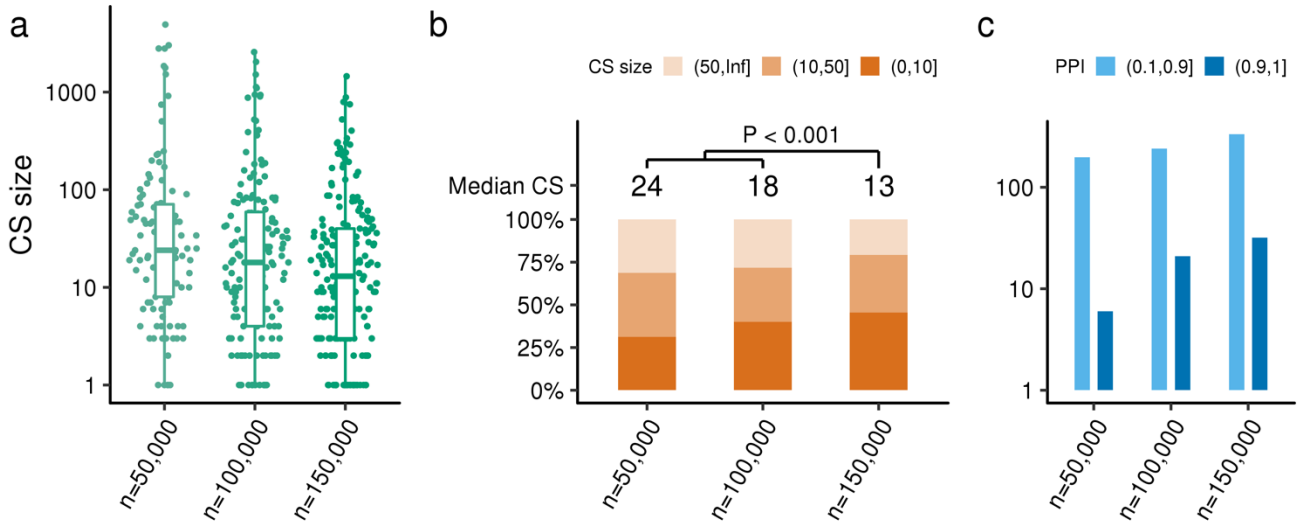
3.2 Selection of statistical model

To conduct statistical finemapping, we applied FINEMAP software which takes the Bayesian approach to estimate variants' posterior probability of being causal. FINEMAP used a shotgun stochastic search algorithm to test various configurations (combinations of causal variants). This approach can reduce the

computational burden of testing possible scenarios with an expanded number of causal variants in a single locus. Since this approach prioritizes configurations maximizing variance explained by variants in the configurations and is conservative for rare variants, which often explain small variance despite large effect sizes, most of the configurations are composed of common variants. Thus, we used this procedure not to overestimate possible rare causal associations. We also compared results between FINEMAP and other finemapping methods (see 3.12 below) and obtained comparable results of enrichment.

3.3 Sample size and finemapping resolution

To empirically estimate the impact of sample size in the finemapping resolution, we performed a down-sampling analysis for height as a representative trait. We randomly took 50k, 100k, and 150k subjects in the BBJ 1st cohort whose height data was available and conducted association analysis followed by statistical finemapping. We restricted the analysis to the significant regions found in the 50k analysis. We compared credible set sizes and the number of variants in the bins of PPI. As a result, we observed that the more samples, the smaller credible set sizes, and the more variants with high PPI.



SI 3.3 | Increased sample size resulting in small credible sets and more variants with high PPI in statistical finemapping

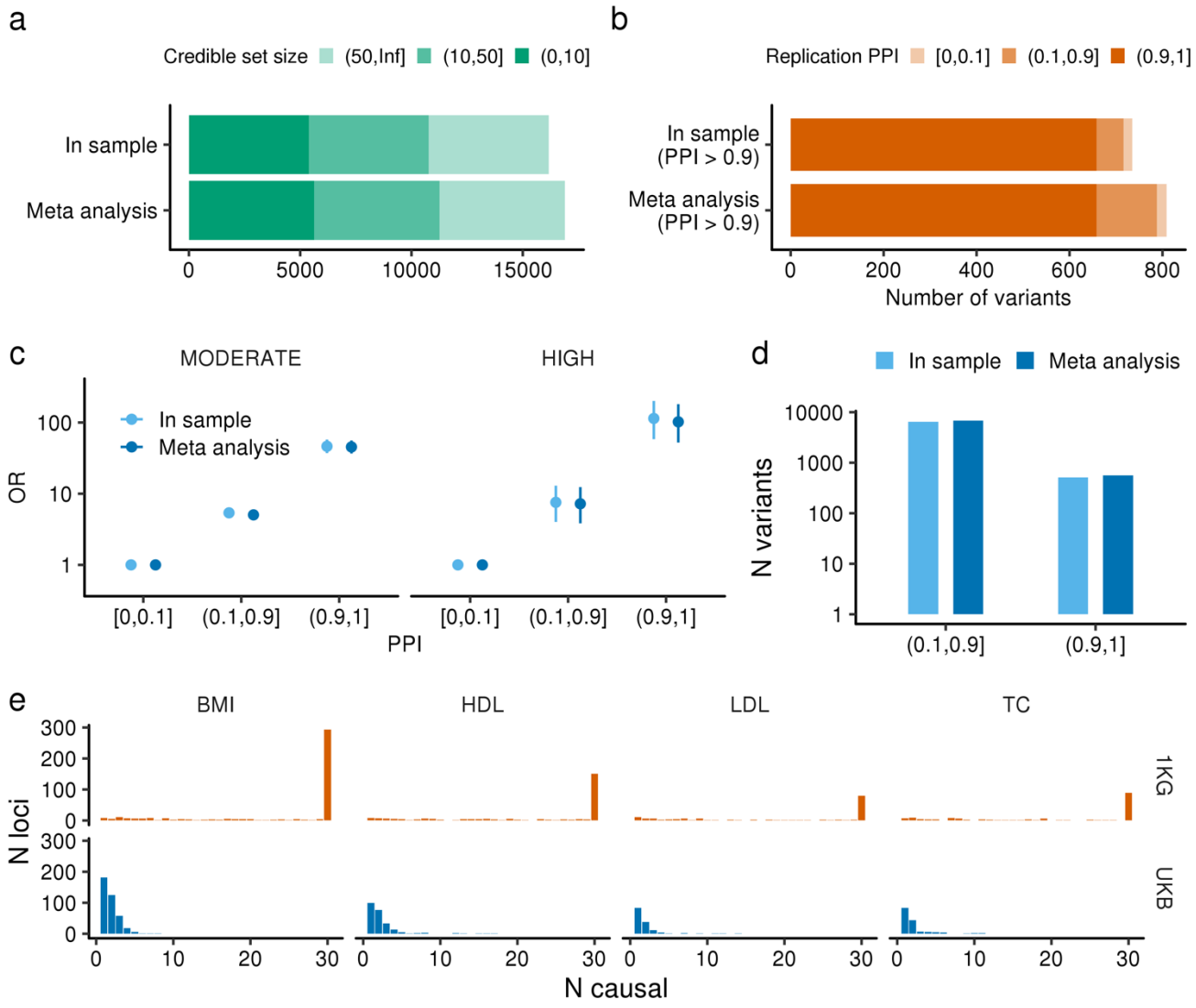
a. Credible set (CS) size and sample size in the down sample analysis (height). Each dot shows CS and the Y axis shows number of variants in each CS. Box plots show median and interquartile ranges of CS size. **b.** CS size categories and sample size. *P*-value was calculated by Wilcoxon rank-sum test. **c.** Number of variants with high PPI. CS, credible set; PPI, posterior probability of inclusion.

3.4 Harmonized approach for statistical finemapping

To maximize the power of statistical finemapping, we compiled large Japanese dataset under the harmonized data processing (quality control, imputation, and association analysis) and meta-analyzed. However, finemapping in the meta-analyzed data is recently suggested to result in the bias[8]. To confirm the feasibility of our strategy, we conducted sensitivity analysis to confirm concordance and performance of the finemapping in in-sample approach and meta-analysis approach in our data. As in-sample approach, we explicitly used BBJ 1st cohort for association analysis and generation of LD matrices. As meta-analysis approach, we used meta-analyzed summary statistics (BBJ 1st, BBJ 2nd and NCGG), and generated LD matrices from the BBJ 1st cohort. We observed better resolution in meta-analysis approach, high replication rate of high PPI variants and equivalent functional enrichment. From these results, we decided to use meta-analysis approach in this study (SI 3.4a-d).

In addition, we confirmed that the use of external LD structure in the statistical finemapping resulted in unstable and noisy results. This is often the case in the efforts of meta-analysis where summary statistics are available, and researchers take LD structure of external data such as 1KG which is not used for imputation. When we applied FINEMAP to the summary statistics of UK Biobank (UKB, imputed by Haplotype Reference Consortium panel) with the use of 1KG European (EUR) LD structure, we failed to resolve associations in most of the regions. On the contrary, when we applied FINEMAP to the summary statistics with the use of LD structure of UKB (inferred by imputation data), we could resolve most of the regions.

These results highlighted importance of the careful harmonization of the input data in the statistical finemapping. Thus, we processed the three data sets (BBJ 1st, BBJ 2nd and NCGG) using harmonized data process, namely, the same quality control, the same imputation reference panel with the same imputation protocol. These efforts resulted in stable and clean results of finemapping (see SI 3.4e) despite meta-analysis. The additional data set (ToMMo cohort) was based on different imputation panel and we did not include the additional data set for statistical finemapping (we applied FINEMAP to the summary statistics of meta-analysis of the three data sets).



SI 3.4 | Sensitivity analysis for finemapping strategy

a. The number of credible sets identified by in-sample approach and meta-analysis approach. The colors show the credible set size (the number of variants included in each credible set). **b.** The number of variants with high PPI (> 0.9). The colors show the PPI in the replication dataset. **c.** The functional enrichment by PPI. The points show odds ratio and error bars show 95% credible sets estimated by Fisher's exact test. **d.** The number of variants with high PPI. PPI, posterior probability of inclusion; OR, odds ratio. **e.** The distributions of inferred number of causal variants in UKB dataset using different LD data sources. The X-axes show the number of inferred causal variants and Y-axes show the number of loci.

3.5 Quality control of finemapping results

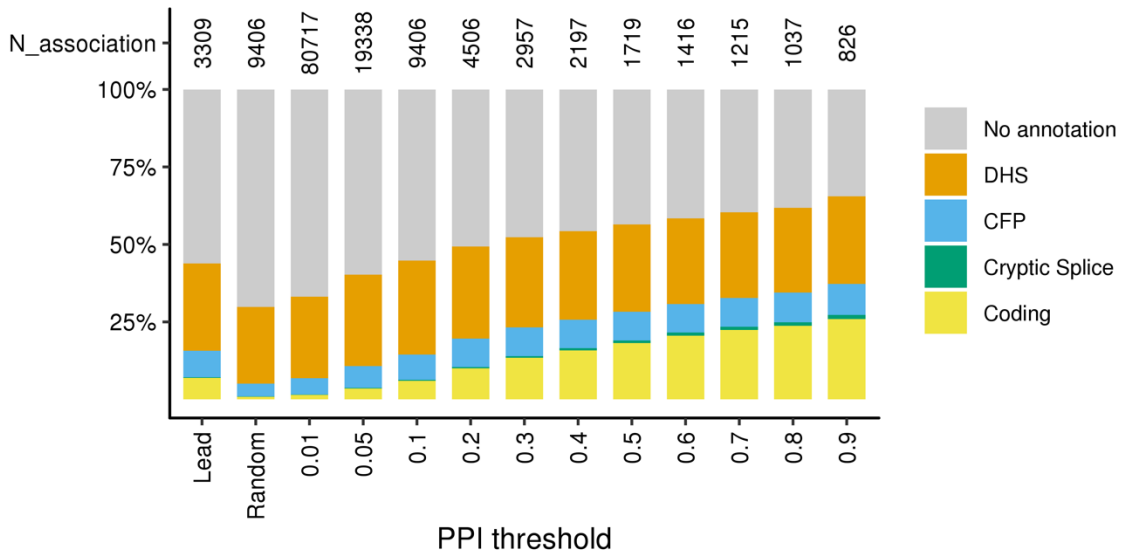
We applied the following approach to confirm the causality inferred by statistical finemapping. First, we extract the genotype dosage of the *k* variants, which were estimated to be causal by FINEMAP. Using the dosages extracted, we conducted multivariate linear regression analysis against the normalized phenotype, excluding related individuals in the BBJ 1st cohort.

$$Phenotype \sim \beta_1 dosage_1 + \beta_2 dosage_2 + \beta_3 dosage_3 + \ldots + \beta_k dosage_k$$

Then, we evaluated the independence of associations of these putative causal variants by the significance of beta coefficients of these dosage terms. For the loci in the X chromosome, we conducted this regression in males and females separately, then meta-analyzed using a fixed-effect model to obtain test statistics. We noted that a small part of putative causal variants did not attain conditionally significant associations. To control the quality of the result of finemapping, we excluded such unsolved loci (16 loci, 0.48%) harboring even one putative causal variant with conditioned *P*-value < 1×10^{-4} from downstream analysis.

3.6 Threshold of PPI to regard variants as causal

In this study, we used marginal PPI as a uniform index of causality of variants, which is a continuous value from 0 to 1 assigned to each phenotype-variant association. PPI represents the overall probability at which the variant of interest is causal in the locus. To determine a threshold of PPI with which we regarded variants tested as causal, we assessed distributions of functional annotations in variants with different thresholds of PPI. Not surprisingly, the distribution of functional annotations in variants randomly selected regardless of PPI was quite different from that of lead variants (as shown below). Lead variants contained more coding variants (~ 10 times), cryptic splice variants[9] (~ 3 times) and variants in consensus footprints (CFP, ~ 2 times)[7] and DNase I hypersensitivity sites (DHS, ~ 1.25 times)[8]. Thus, we put a reference distribution of functional annotations in lead variants. We found that variants with PPI > 0.1 showed almost the same distribution of functional annotations as lead variants (as shown below). We confirmed that variants with PPI > 0.9 showed enhanced enrichment in the functional annotations of coding variants, cryptic splice variants, and variants in CFP and DHS, in line with the theory of Bayesian inference that the higher the PPI, the more convincingly causal the variants are. Based on these results and findings, we adopted PPI > 0.9 and PPI > 0.1 as the main thresholds to include variants in the current study.



SI 3.6 | Distributions of variants in functional annotations based on different thresholds of PPI

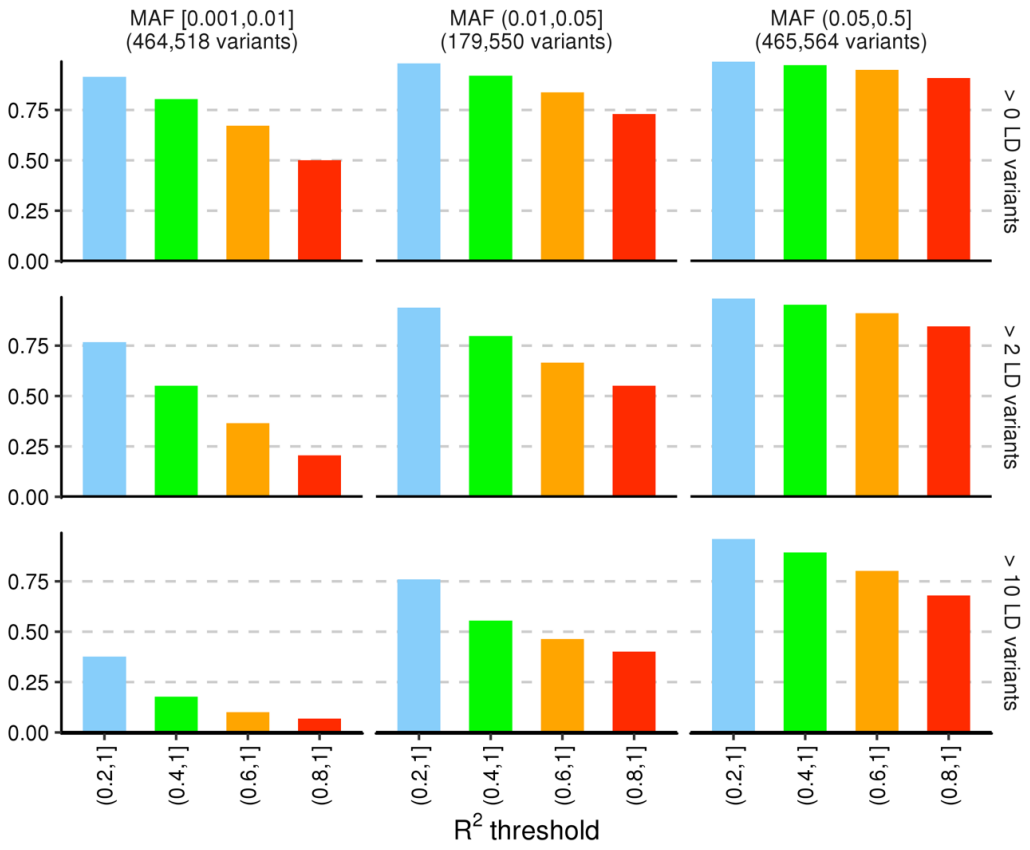
The bar charts show the composition of the variant annotation of lead association that attained the strongest association in the locus (n = 3,309), randomly selected associations (n = 9,406, same as associations with PPI > 0.1), associations with designated PPI bins. DHS, DNase I hypersensitivity site; CFP, consensus footprint. PPI, posterior probability of inclusion.

3.7 The inference of putative causal variants

To determine the putative causal variants in these associated loci in an unbiased manner, we applied an LD-informed Bayesian statistical finemapping using FINEMAP[4], taking advantage of the fine LD structure derived from the dense reference panel. We computed a marginal PPI, the probability that the variant of interest is causal, for each variant as a scalar value ranging from 0 to 1. Applying the algorithm to the 3,309 significant loci excluding major histocompatibility complex region (MHC), We identified 5,673 credible sets and 72.4% of them were solved with ≤ 50 variants in their 95% credible sets. Among them, we identified 826 phenotype-variant pairs (associations) with PPI > 0.9 and 9,406 with PPI > 0.1, a threshold we found comparable to lead associations that attained the smallest *P*-value in the loci. 25.1% (2,358/9,406) of associations with PPI > 0.1, and 52.8% (436/826) of associations with PPI > 0.9 were lead associations. Conversely, 71.3% (2,358/3,309) of lead associations attained PPI > 0.1, and 13.2% of lead associations (436/3,309) attained PPI > 0.9. We observed multiple independent associations at hundreds of loci. 40.1% (1,352/3,309) of loci contain ≥ 2 credible sets, and 3.5% of loci (116/3,309) contain ≥ 2 associations with PPI > 0.9. These results indicate that a single genome-wide significant locus often contains independent-, multiple causal signals and that densely imputed and well-powered GWAS can efficiently detect these multiple causal variants in each locus.

3.8 Statistical finemapping in the rare variants

First, we assessed the number of variants in LD by MAF bins (SI 3.8.1). We calculated R^2 value for imputed genotypes in study samples for all variants on the chromosome 1 with minor allele frequency ≥ 0.001. We observed around half of rare variants (0.001 ≤ MAF < 0.01) have at least one LD variants ($0.8 < R^2 \leq 1.0$) which suggest the utility of statistical finemapping approach for the variants in this range of MAF. However, the number of variants in LD ($0.2 < R^2 \leq 1.0$) for rare variants (median 7 variants) are limited compared to low frequency (34 for 0.01 ≤ MAF < 0.05) or common variants (134 for 0.05 ≤ MAF < 0.5).



SI 3.8.1 | LD patterns for 1,109,632 variants on Chromosome 1 in 169,020 unrelated Japanese
Each bars indicates the fraction of the variants (Left Y-axes) with at least one, three, and 10 LD variants (Right Y-axes) with designated R^2 values (X-axes). R^2 was calculated for all the variants in 1MB window (the mean number of variants in each window was 10,206). We stratified the results by MAF of the variant of interest (Top row).

From the observation above, we recognized significantly different LD patterns in rare, low-frequency, and common variants. To confirm the validity of the statistical finemapping for the rare variants, we performed

finemapping study in the simulated causal variants and phenotypes using study genotypes. We randomly picked the *I* causal variants (*I* = 1, 3, 5) in the randomly chosen 1MB loci in the chromosome 1. For each causal variant *i*, we set effect size β_i to attain 80% power at its minor allele frequency and sample size (n = 169,020). We generated quantitative phenotypes Y_j by multiplying genotype dosage (G_{ij}) by corresponding effect size β_i for all the causal variants. We added error term e_j that follows normal distribution with mean 0 and standard deviation 1 - h^2 (h^2 as trait heritability).

$$Y_j = \sum_i^I \beta_i G_{ij} + e_j$$

$$e = norm\,(0, 1 - h^2)$$

We evaluated the effectiveness of statistical finemapping for identifying causal variants across different MAF bins by using simulated phenotype data. To accomplish this, we conducted an association analysis in the locus using a generalized linear model by PLINK2, excluding related samples, and then performed FINEMAP, assuming ten causal variants. We used imputed genotypes from 169,020 unrelated individuals and ran 1,000 simulations for *I* = 1, 3, 5, and h^2 = 0.1. To measure performance, for each PPI bin, we calculated Positive Predictive Value ($PPV_{Finemapping}$) and False Negative Rate ($FNR_{Finemapping}$). We selected these metrics due to the overwhelming number of variants in the true negative. For comparison, we also calculated lead variants' PPV ($PPV_{Lead\text{-}variant)}$ and FNR ($FNR_{Lead\text{-}variant}$).

$$PPV_{Finemapping} = \frac{N_{High-PPI\ and\ True-causal}}{N_{High-PPI\ and\ True-causal} + N_{High-PPI\ and\ Non-causal}}$$

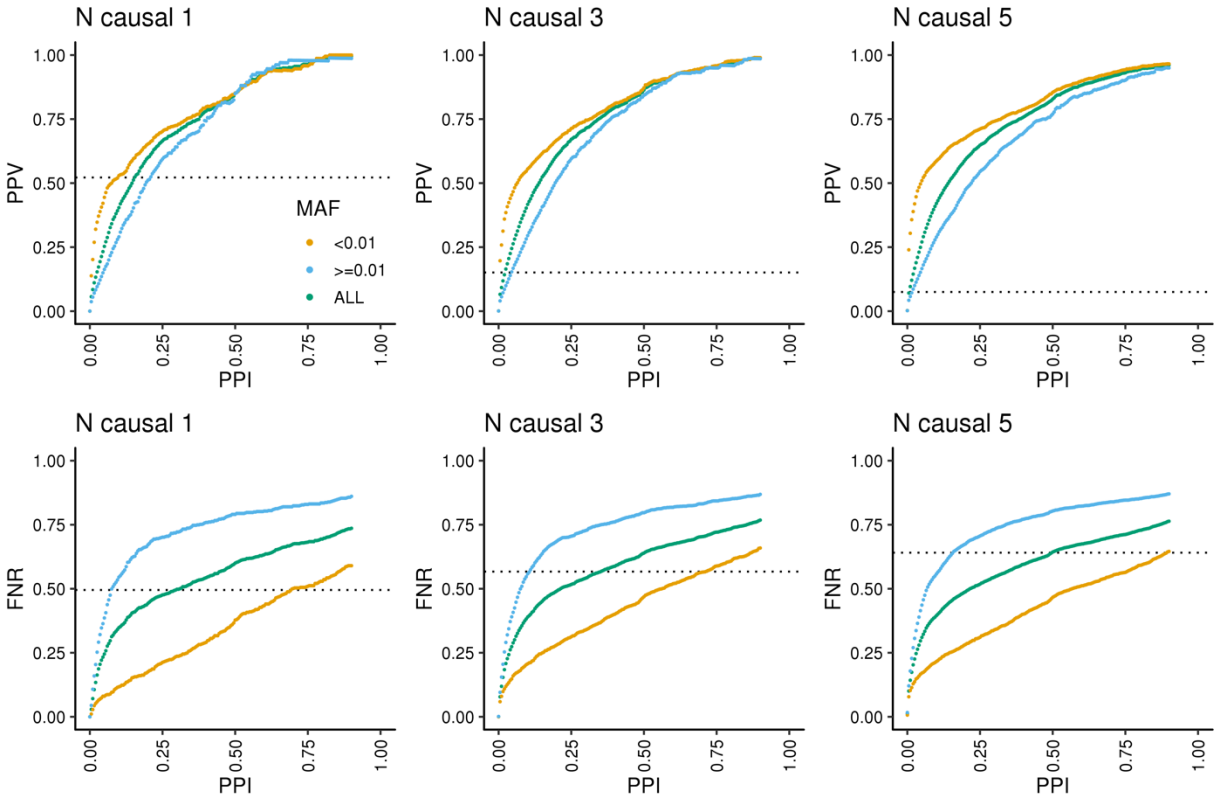
$$FNR_{Finemapping} = \frac{N_{Low-PPI\ and\ True-causal}}{N_{Low-PPI\ and\ True-causal} + N_{High-PPI\ and\ True-causal}}$$

$$PPV_{Leadvariant} = \frac{N_{Lead\ and\ True-causal}}{N_{Lead\ and\ True-causal} + N_{Lead\ and\ Non-causal}}$$

$$FNR_{Leadvariant} = \frac{N_{Non-lead\ and\ True-causal}}{N_{Non-lead\ and\ True-causal} + N_{Lead\ and\ True-causal}}$$

We observed comparable and slightly superior finemapping performance in rare variants (MAF [0.1%, 1%]), indicating that the current statistical finemapping approach is reliable for variants within this allele frequency range (SI 3.8.2). Our simulation showed that lead variants exhibited similar PPV to variants with PPI = 0.155 in the case of N_{Causal} = 1, PPI = 0.025 in the case of N_{Causal} = 3, and PPI =0.010 in the case of N_{Causal} = 5. Considering that we detected multiple credible sets (N_{Causal} > 1) in 28.8% of the loci analyzed in this study,

these results align with our finding that PPI = 0.1 serves as a threshold equivalent to lead variants in terms of causal variant detection.



SI 3.8.2 | Performance of statistical finemapping in rare and common variants

We display the performance of statistical finemapping in rare and common variants. The dots on the graph represent the outcomes of statistical finemapping using simulated data. We randomly selected causal variants and generated simulated phenotypes based on these variants. PPV and FNR were calculated based on the true causal variant assignment and the corresponding PPI estimated by statistical finemapping. The dotted lines indicate the PPV/FNR by lead variants in the loci. We conducted 1,000 simulations on 1MB loci in chromosome 1 for each situation (N causal 1, 3, and 5). PPV, positive predictive value; FNR, false negative rate; PPI, posterior probability of inclusion.

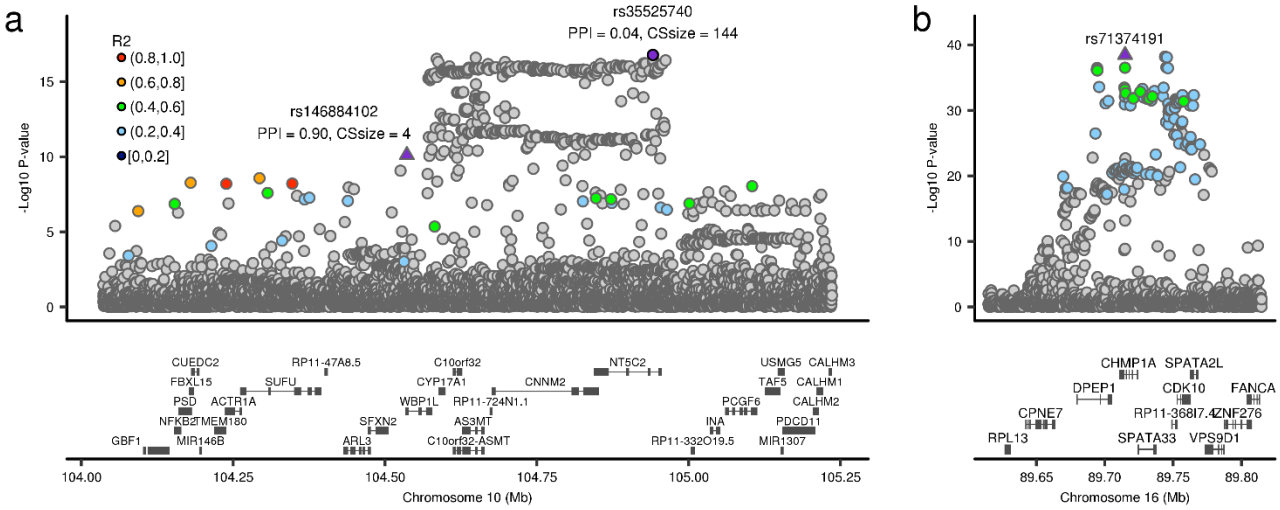
3.9 Prioritize associated variants by posterior probability

Statistical finemapping can solve the complex LD structure and identify putatively causal variants or narrow down candidates of causal variants in the loci.

One example was found at the *NT5C2* locus on chromosome 10 associated with sodium levels. rs35525740 was the lead variant in this locus and tagged by numerous variants. Accordingly, the PPI of rs35525740 was 0.04 and the size of the credible set containing rs35525740 was 144. Statistical finemapping suggested an

additional signal in this locus. rs146884102, not in LD with rs35525740, showed high PPI of 0.90 and its credible set only contained 4 variants (SI 3.9a). As shown by this example, even in loci where the lead variants are in LD with a large number of variants, LD-informed statistical finemapping identify independent putative causal variants.

We also show the region where many variants are in strong LD with a lead variant and statistical finemapping successfully resolved the association and narrowed down candidates of a causal variant depending on sample sizes. Finemapping showed one causal variant in the *CHMP1A* region on chromosome 16 associated with height. While the lead variant tags many variants in strong LD and the down-sampling analysis of 50k samples resulted in the credible set of the 50 variants, the current study could narrow down the credible set of the 6 variants (SI 3.9b).



SI 3.9 | Usefulness of statistical finemapping to identify an additional causal variant or to narrow down candidates of causal variants in regions with different LD structure

X-axes show the genomic coordinate. Y-axis in the top panel shows -$\log_{10}$ *P*-value. The bottom panel shows genes located in the loci. **a.**rs35525740 is the lead variant in this locus. rs146884102 is the variant with high PPI. The color of points indicates linkage disequilibrium R^2 to **a.**rs146884102 and **b.**rs71374191.
PPI, posterior probability of inclusion; CS, credible set.

3.10 PPI of lead variants and multiple independent signals in associations

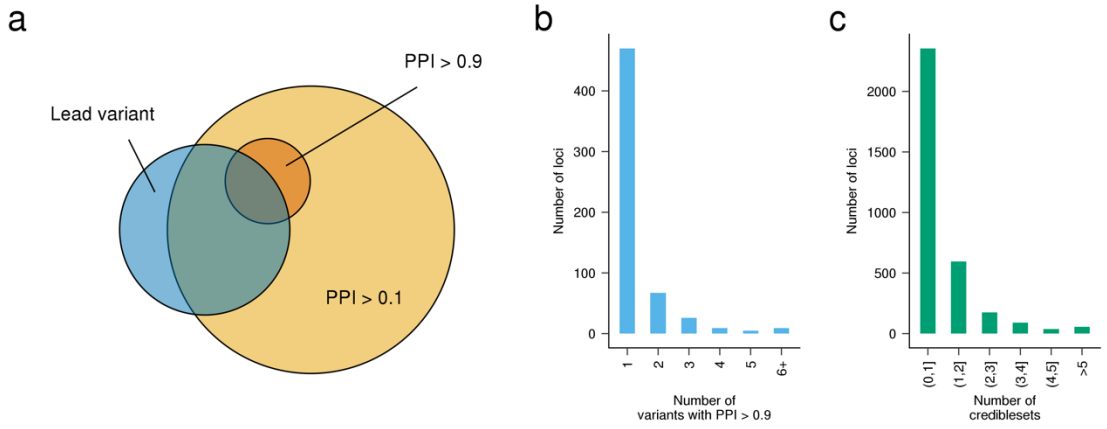
We found 71.3% (2,358/3,309) of lead associations (used for the finemapping) attained PPI > 0.1, and 13.2% of lead associations (436/3,309) attained PPI > 0.9. We observed multiple independent associations at hundreds of loci. 28.8% (954/3,309) of loci contain ≥ 2 credible sets, and 3.5% of loci (116/3,309) contain

≥ 2 associations with PPI > 0.9. These results indicate that a single genome-wide significant locus often contains independent, multiple causal signals and that densely imputed and well-powered GWAS can efficiently detect these multiple causal variants in each locus.

3.11 Associations where lead variants do not show PPI>0.1

In the current study, we found that about 951 (28.7%) of 3,309 lead variants did not show PPI more than 0.1, which is the threshold we found to obtain comparable functional enrichment to lead variants (discussed in **3.6 Threshold of PPI to regard variants as causal**).

Among the 951 loci, we found that we did not find any variants with PPI more than 0.1 in the 653 loci (68.7%). This indicates that lead variants in the 653 loci are in strong LD with many other variants, suggesting that still 200k subjects are underpowered to resolve the associations (if these variants are not in complete LD with each other - when many variants are in complete LD, we cannot resolve these associations regardless of sample sizes). For the remaining 298 loci (31.3% of the 951 loci) where lead variants do not have PPI more than 0.1, the other variants have PPI more than 0.1. Among the 298 loci, we noticed that majority of regions, 273 loci (91.6%), have multiple causal variants (expected number of causal variants more than 1). There might be several possibilities for these loci, including the presence of strong LD and independent signals with lead variants at the same locus (as exemplified in SI 3.9a) and lead variants tagging the multiple causal variants (and tagging multiple causal variants would result in the strongest associations), but these possibilities remain to be further explored in a more powered study.



SI 3.11 | Distribution of variants with high PPI

a. The overlap of lead variants, variants with PPI > 0.1, and PPI > 0.9. **b.** The bar charts show the number of loci including putative causal variants shown in the X-axis. **c.** The bar charts show the number of loci including credible sets shown in the X-axis.

3.12 Non-lead variants with PPI >0.9

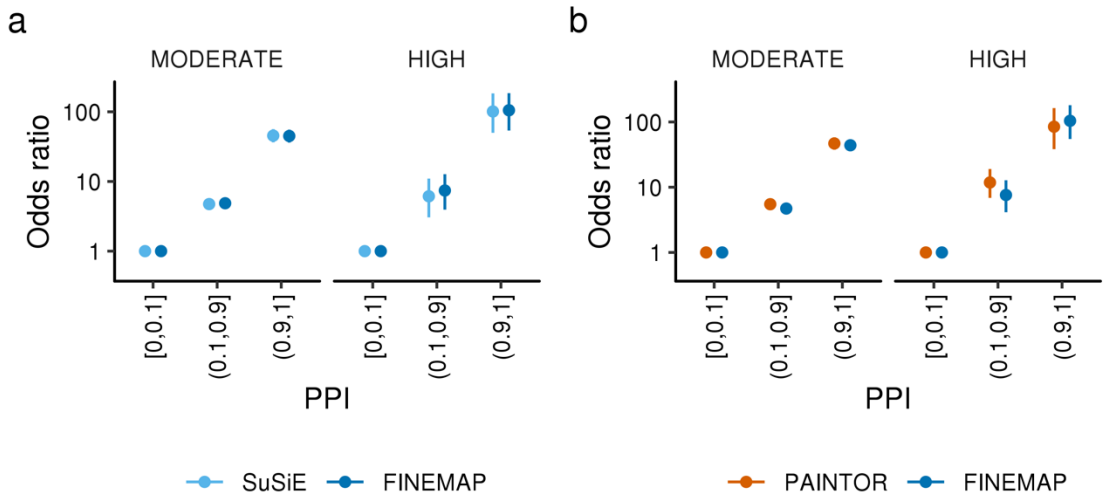
We found a total of 390 non-lead variants with PPI more than 0.9 in the 221 associations (phenotype-region pairs). All of the 221 associations have multiple causal variants. Regarding MAF, many of these variants are rare. Among them, 53.6 and 34.9% have MAF less than 0.05 and 0.01, respectively, in contrast to 9.1 and 3.4% in the lead variants. Supplementary Table 12 summarizes non-lead variants with PPI > 0.9 while lead variants in these regions show PPI < 0.1.

3.13 Applying other finemapping methods

We confirmed that our finemapping strategy was not dependent on a program we used. We applied SuSiE, another finemapping method, to our summary statistics and compared results between the two methods. We used R package “susieR” (version 0.12.19) and conducted finemapping using the same summary statistics and LD matrix as FINEMAP. As a result, we found that SuSiE also showed comparable results with those in the FINEMAP under the viewpoint of PPI and enrichment for functional annotations (Supplementary Table 19 and SI 3.13a).

Since we observed tissue-specific enrichment of putative causal variants, we ran another finemapping method, PAINTOR (v3.0), leveraging functional annotations. Again, we used the same summary statistics and LD matrix as FINEMAP. We put functional annotations defined by VEP (Supplementary Table 15) and Disease impact score into the model to assess enrichment for functional annotations. As a result, we found that PAINTOR showed reasonable slight increase in variants with high PPI (0.1 < PPI, due to taking advantage of functional annotations) and almost comparable (slightly increased) enrichment of annotations analyzed. This slight increased enrichment in PAINTOR is reasonable since PAINTOR prioritizes variants in functional annotations (unlike FINEMAP) for which we found enrichment of causal variants (SI 3.13b). These results support enrichment of functional annotations in our results and indicate that FINEMAP could successfully prioritize causal variants.

The annotation agnostic finemapping methods, including FINEMAP, have several advantages. One is that these approaches only require summary statistics and LD information, and the quality of functional annotation does not affect the results. The second is that the list of finemapped variants or its PPI could be used as unbiased barometer of pathogenicity of variants. For example, we found strong enrichment of high PPI in cryptic splice variants or variants with high disease impact scores in this study. This observation could provide the evidence of the functionality of these variants.



SI 3.13 | Comparable enrichment of causal variants for functional annotations by SuSiE, PAINTOR, and FINEMAP

The X axes show PPI bins in BBJ GWAS, and the Y axis shows odds ratio in reference to the lowest PPI bins. Error bars show 95% confidence intervals estimated by the Fisher's exact test.

3.14 Finemapping results in the UK Biobank data

We compared finemapping results from BBJ with the results from UKB. We found 49% of variants with high PPI (> 0.9) in BBJ showed PPI more than 0.1 in UKB. We also found 27% of high PPI (> 0.9) variants in UKB showed PPI more than 0.1 in BBJ. As recently suggested[10], we found limited replication rate of high PPI variant – indicating that even using biobank-scale, we failed to identify concordance of causal genetic variants (variant-level concordance). However, we found similar patterns of enrichment of causal variants in functional annotations between Japanese and Europeans, including frameshift, 3'UTR, 5'UTR, and other non-coding regions (this is also discussed in SI 6.2).

These indicate that even using biobank-scale GWAS which could maximize significant hits, GWAS significant regions are not frequently driven by causal variants common across populations. Since we found genetic correlations of associations with the same phenotypes between populations, similar regions (including polygenetic architectures not reaching GWAS significant levels), the same genes can globally explain the associations with the phenotypes, but locally causal variants are not the same in many cases. Especially, high PPI variants tend to be population specific and less frequent in the other population. These results suggest the importance of further causal variant inference in the various populations.

3.15 Potential extension of the analyses to binary traits

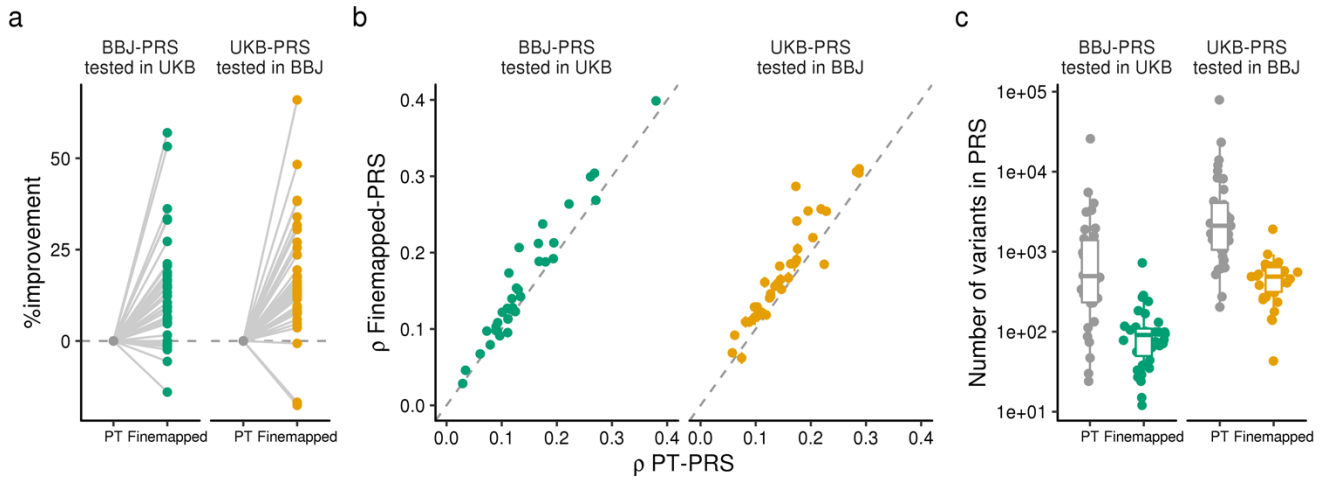
While it would be an option to extend the current analyses to binary traits, in the power analyses described in Supplementary Note 3.3, we noticed that statistical finemapping requires enough statistical power supported by large sample sizes. The additional data of 53k subjects are from a population cohort in which sample sizes of binary traits are limited due to the nature of the population cohort and the data was imputed by different imputation panel. Since the remaining 200K samples are not a very large number of subjects, we decided to focus on quantitative traits for which we could maximize statistical power.

3.16 Improved transferability of polygenic risk score by statistical finemapping

Previous studies have suggested that statistical finemapping can improve the performance of polygenic risk prediction. To evaluate the effectiveness of statistical finemapping in our framework, we generated PRS models based on putatively causal variants (highest PPI in each credible set) and LD-independent lead variants selected using the conventional pruning and thresholding (PT) method. We then tested these models in the white-British population in UKBB and in the BBJ dataset. The PT was performed on the 1KG EAS (n=504), EUR (n=503) samples. We scored individuals in the testing dataset using the selected variants. We optimized the parameters for PT method using 25% of randomly selected individuals and subsequently scored the remaining 75% of individuals using the best performing PT-PRS and finemapped PRS. For this testing population, we calculated the rank-based Spearman's correlation between PRS and normalized phenotypes (linearly regressed on age, sex, and genetic principal components. In BBJ disease status at enrollment was also included) in unrelated samples. We tested the significance of the relative improvement in correlation coefficients using the paired Wilcoxon rank-sum test.

As demonstrated in SI 3.16, both experiments (PRS models derived in BBJ and tested in UKB, or derived in UKB and tested in BBJ) resulted in a significant improvement in prediction accuracy. The finemapping-based PRS achieved better performance with significantly fewer variants than the conventional PT method (Supplementary Table 19, 20). We also observed better directional concordance of variant associations between BBJ and UKB GWAS (Supplementary Table 21) in the finemapping-based PRS. Based on these observations, we hypothesize that statistical finemapping can capture causal variants that share effect direction across different ancestries, resulting in improved prediction performance of genetic risk.

We additionally compared the performance of PRS between the conventional (1KG) and new (1KG + 3,256 Japanese) reference panels and observed modest improvement in the novel reference panel (Supplementary Table 22).



SI 3.16 | Improved transferability of PRS by statistical finemapping

a, The ranked correlation coefficients (ρ values) between PRS and phenotypes were compared in PT-PRS and Finemapped PRS. The performances were normalized to PT-PRS, and the relative improvements by Finemapped-PRS are shown as color dots. The performance of PRS was evaluated by Spearman's rank correlation between PRS and quantitative phenotypes after adjusting for covariates (age, sex and first ten genetic principal components). **b,** The dots and error bars represent the ranked correlation coefficients (ρ) and their 95% CI, respectively. The X-axes display the performance of PT-PRS and the Y-axes display Finemapped-PRS. **c,** The dots indicate the number of variants in the best performing PT-PRS and Finemapped-PRS for each phenotype. PRS, polygenic risk score; BBJ, BioBank Japan; UKB, UK-Biobank; PT, pruning and thresholding.

4 Novel associations

We listed examples of novel associations between phenotypes and genes in the current study which draw our attention. We distinguish novel associations between novel gene-phenotype pairs and variant-phenotype pairs (unreported variants in known gene-phenotype pairs). Since the full table and summary statistics will be released, these results offer examples and further opportunities to expand analyses based on the current findings (focusing on the phenotype-genes, phenotype-variants, and gene-variants not restricted to the pairs below). We noticed that novel associations include not only variants that are rare and specific to Japanese or EAS, but variants that are globally common (such as those with MAF of more than 5% in the other populations) and more prevalent in EAS or Japanese (also discussed in 2.3 Significant associations).

We also found that novel associations include variants that are rare, specifically in Europeans. This is consistent with the current situation of Europeans as major sources of genetic association studies and

underscores the importance of extending genetic association studies to non-European populations to understand the genetics-supported biology that underlies complex traits.

4.1 Novel associations between phenotypes and genes

Associations described below are shown in Supplementary Tables 4-11.

BMA-BAFF interaction

We found novel associations of a rare missense variant (rs150352299) in *TNFRSF17* which encodes BMA. This rare variant is highly Japanese-specific (AAF_{BBJ} = 0.38%, $AAF_{gnomAD\ EAS}$ = 0.0642%, $AAF_{gnomAD\ FIN}$ = 0.0288% not observed in other populations). This Japanese-specific missense variant was significantly associated with a higher albumin-goblin (AG) ratio (β_{AG} = 0.306, $P_{AG} = 3.9\times10^{-22}$), lower non-albumin protein (NAP, β_{NAP} = -0.327, $P_{NAP} = 3.3\times10^{-25}$), and lower total protein (TP, β_{TP} = -0.183, $P_{NAP} = 6.5\times10^{-10}$). *TNFRSF17* encodes B-cell maturation antigen (BMA) which is specifically expressed in mature B-cell and responsible for antibody production (Extended Fig. 2a). BMA is known to interact with B-cell activating factor (BAFF) encoded by *TNFRSF13B* in which we also identified a rare loss of function variant (rs769165409) associated with AG ratio with high PPI (AAF_{BBJ} = 0.1%, β_{AG} = 0.353, $P_{AG} = 3.9\times10^{-7}$, Supplementary Table 7) in addition to the known common missense variant (rs34562254, AAF_{BBJ} = 36%, Supplementary Table 6). This rare missense was also Japanese-specific and not observed in gnomAD. These findings genetically support BMA-BAFF interaction in B-cells since rare missense variants in both genes converge into the same phenotype (AG) with high PPI and both higher AG and lower NAP suggest decreased levels of immunoglobulin.

USP47

We observed a significant association between a missense variant in *USP47* and glucose levels (rs138329346). This variant is a Japanese-specific (3% in Japanese, and 0.8% in Korean) rare deleterious missense variant (SIFT deleterious; PolyPhen probably damaging), and the variant allele associated with increased glucose levels. We found two supporting evidence of our finding. First, rare genetic aggregating testing in the European population (accessed via https://app.genebass.org) showed a strong association between increased glucose levels and the burden of protein-truncating variants in the *USP47* gene ($P = 7.3 \times 10^{-7}$). Second, the murine knockout model of Usp47 showed increased glucose levels [data was obtained from International Mouse Phenotyping Consortium database].

MLLT10

We found an EAS-specific variant in the *MLLT10* (Mixed lineage leukemia T10) gene associated with serum creatinine level and eGFR. This region was identified for these phenotypes for the first time. While *MLLT10* is known for its function in fusion genes and hematopoietic malignancy[11], this association suggests an

unknown mechanism between AP10 and kidney function. In line with this association, *MLLT10* is expressed in the kidney based on the GTEx data.

ETV6

We found a novel association between an intronic variant in *ETV6* and eosinophil count. *ETV6* is known to be associated with hematopoietic malignancy, and the fusion gene of *ETV6-RUNX1* is frequently found in childhood B- cell acute lymphocytic leukemia[12]. Fusion gene *ETV6-PDGFRB* is also known to be associated with chronic myeloid malignancies with eosinophilia and this fusion gene stimulates hematopoietic progenitors to differentiate into eosinophils[13]. This association may provide a direct link between eosinophils and *ETV6* function (without fusion with other genes).

DMXL1

We found a novel association between BUN and *DMXL1* region (an upstream variant). Interestingly, while *DMXL1* expresses in various tissues and has many regulatory functions, this gene encodes a protein one of which functions is a regulator of the V-ATPase proton pump[14]. Thus, this association may provide the first evidence of a connection between renal function and a genetic variant in the *DMXL1* region, possibly through affecting the regulation of the V-ATPase proton pump.

EVC

EVC encodes a protein whose dysfunction leads to Ellis-van Creveld syndrome[15]. We found a missense variant in *EVC* associated with the eGFR level. Since patients with Ellis-van Creveld syndrome were not reported to have renal dysfunction[15], this association may provide novel evidence of the broader function of *EVC*.

STAB2

We found a novel association between *STAB2* missense variant rs185968359 *and APTT*. By in-silico functional prediction, rs185968359 is predicted to be deleterious (SIFT deleterious; Polyphen probably damaging). This likely causal variant (PPI = 0.93) was an East Asian specific rare variant (MAF = 0.48%), and this variant was associated with shortened APTT (Beta -0.252) – indicating enhanced coagulation. Consistent with our observation, other deleterious missense variants in *STAB2* were associated with an increased risk of venous thromboembolism in the European population[16]. *STAB2* is reported to be a clearance receptor for von Willebrand factor (vWF), and functional loss of *STAB2* was associated with increased vWF levels and enhanced coagulation. Our study linked a blood biomarker and the pathophysiology of VTE through genetic variants.

EGLN1

We found novel associations between EAS-specific missense variant rs186996510 in *EGLN1* and EGFR. *EGLN1* encodes PHD2 and regulates hypoxia inducing factor (HIF) pathways. rs186996510 is known to be under strong selection in the Tibetan population[17] ($AAF_{Tibetan}$ = 70.9%, AAF_{BBJ} = 5.2%, $AAF_{gnomAD\ NFE}$ =

0.025%) suggesting high altitude adaptation. *EGLN1* locus is an established hemoglobin/hematocrit-associated locus and *EGLN1* has been attracting attention in erythropoiesis. Recently, however, the effects of the HIF pathway in the progression of chronic kidney disease have received increased attention, and this genetic association may have implications for the impact of *EGLN1* on renal function.

PAX4

We found a novel association between glucose level and rs3757787 an upstream variant of *PAX4*. This variant is highly EAS-specific (AAF_{EAS} = 11%, AAF < 0.1% in other populations in gnomAD). Since *PAX4* is a master regulator of beta cell development in the pancreas, the direction of the effect suggests this variant decreases the development of beta cells via dysfunction of *PAX4*, resulting in increased glucose levels even in non-diabetic subjects.

LINC01094

We found an association of the LINC01094 region with TC and HDLC (Supplementary Table 4). This variant is highly specific to EAS. LINC01094 is reported to be related to various cancers[18,19] including pancreatic cancer and breast cancer, in which cholesterol levels play important roles in cancer development. While a direct link between LINC01094 and cholesterol was not previously reported, the current findings might provide human genetic evidence between them for the first time.

IL20RA

We found an association between an intronic variant in *IL20RA* and BUN. *IL20RA* encodes IL20 receptor alpha subunit. IL20 is a proinflammatory cytokine associated with renal fibrosis, and an increased level of IL20 is observed in acute and chronic kidney diseases[20]. The previous study reported that IL20 treatment induced apoptosis of tubular epithelial cells via caspase-3 activation in rats[21], suggesting a mechanism by which IL20 is associated with renal diseases. Thus, our finding genetically connects this gene with kidney function in humans for the first time.

GOLM1

We found novel associations between intronic variants in the *GOLM1* and neutrophil counts and basophil counts. *GOLM1* encodes Golgi membrane protein GP73. GP73 is known to be associated with cell growth and cancers especially hepatocellular carcinoma, but never known for QTL associations with blood cells. A previous report showed that this locus was associated with QTL of IL6 production and was modulated in response to the stimulus of multiple pathogens[22] which might suggest a context-dependent effect of this region on blood cell counts.

MYCT1

A deleterious missense variant highly specific to EAS in *MYCT1* showed an association with RBC. *MYCT1* is a myc-related gene and is known as a possible tumor suppressor gene[23]. Since MYC is deeply involved

with cell expansion, this association seems reasonable and provides a possible novel biological mechanism of erythropoiesis supported by genetic findings.

JAZF1

We found an intronic variant in *JAZF1* was associated with RBC count. *JAZF1* is known to be associated with T2D, but its functional role in erythrocytes had not been reported at all. A recent paper suggested *JAZF1* as a potential transcriptional repressor of γ-globin in adult erythroid cells[24]. The current study is the first study to provide human genetic evidence for the link between them.

ALDOB

We found a novel association between the missense deleterious variant in *ALDOB* and platelet count. The variant is highly specific to EAS. Since *ALDOB* encodes *Fructose*-bisphosphate aldolase B; this association might suggest an unknown link between platelet and fructose.

CD163

Macrophage marker *CD163* scavenger receptor for the hemoglobin-haptoglobin complex. An intronic variant, much more prevalent in EAS than in EUR, showed an association with AST for the first time. This genetic association would directly connect macrophage and liver function.

PGM2L1

Intronic variants in the *PGM2L1* gene, showing large differences in allele frequency between EAS and EUR, showed associations with eGFR, sCr and BUN. *PGM2L1* is associated with glucose metabolism, but detailed function is not known.

CYP19A1

An intronic variant in the *CYP19A1* is associated with HDLC. *CYP19A1* is a member of the cytochrome P450 superfamily. P450 is known to be associated with cholesterol synthesis[25], but this is the first time to identify the association between *CYP19A1* and HDLC. This novel association may provide genetic evidence of the function of *CYP19A1* on cholesterol production, while how a causal variant shows its effect on *CYP19A1* should be further addressed.

SIRT1

An intronic variant in *SIRT1* was shown to be associated with hemoglobin levels for the first time. This variant showed quite a different allele frequency between EAS and EUR. Previous studies have shown an association between *SIRT1* and hematopoiesis. Activation of *SIRT1* is shown to result in increased expression of fetal hemoglobin level[26].

ATM

An intronic variant in *ATM* showed an association with hematocrit and hemoglobin. While *ATM* is well known

for its function on hematopoiesis via stabilizing p53 and contributing to the survival of hematopoietic stem cell[27], this is the first time to show genetic evidence between *ATM* and hematocrit or hemoglobin.

DOCK5

We found a novel association between an intron variant in *DOCK5* and hemoglobin. *DOCK5*, known as the dedicator of cytokine 5, is an atypical guanine nucleotide exchange factor (GEF) known for its function on B cells via regulating BCR signaling and acting reorganization[28]. Since the link between *DOCK5* and erythrocytes is not known, the current finding may suggest DOCK5's role in the hematopoietic lineage.

DOCK3

We found associations of the *DOCK3* region with hemoglobin and HT (hematocrit). *DOCK3* is well known for its role in immune surveillance. Since DOCK3 is also a member of the dedicator of cytokines (DOCK) like *DOCK5*, the association between *DOCK3* and hemoglobin and hematocrit might be interpreted in the same context as the association between *DOCK5* and hemoglobin. Interestingly, this gene is found under selection pressure in Japanese (Liu et al. manuscript submitted).

ARHGAP36

A very rare missense variant in *ARHGAP36* showed an association with sodium level. *ARHGAP36* encodes Rho GTPase-activating protein 36. The functions of this protein in the electrolyte are almost not known. Since *ARHGAP36* expresses mainly in the adrenal gland, the association might suggest its function on sodium level via humoral regulation.

RRAS2

A rare intronic variant in the *RRAS2* gene was associated with platelet count. This gene encodes Ras-related protein. While the involvement of the RAS pathway in platelet is reported[29], detailed analyses were not conducted so far. The current finding for the first time provides human genetic evidence to link *RRAS2* and platelet count.

S1PR4

A rare missense variant, rs3746072, showed an association with the segmented neutrophil count. *S1PR4* encodes sphingosine-1-phosphate receptor 4 and sphingosine-1-phosphate signaling is involved with leukocyte trafficking. A previous study (not containing the Japanese population) showed an association of this rare missense variant in *S1PR4* with white blood cells or neutrophil count (but not segmented neutrophil count). Furthermore, finemapping results supported this variant as causal (PPI_{SEG} = 0.733). Our finding for the first time provides the association of *S1PR4* with segmented neutrophil count (rather than band neutrophil count) and further reinforces the link between this variant in *S1PR4* and white blood cells, especially segmented neutrophils[30].

PDE10A

PDE10A encodes an enzyme that hydrolyzes cAMP/cGMP, the second messenger in cells. We found an association between chloride levels and a variant in *PDE10A*. Since chloride and second messengers are functionally tightly linked, this association seems reasonable. Since *PDE10A* is a potential drug target such as schizophrenia[31], this association would also raise the possibility of potential side effects of a drug to be monitored.

SLC12A3

We found multiple rare (0.27% - 1.35%) missense variants in *SLC12A3* are associated with chloride levels (rs146158333, rs369344478, rs185927948). Among them, rs146158333 was previously reported as a causal variant for Familial hypokalemia hypomagnesemia. All these variants are not found or are very rare in other populations (< 0.1% in the gnomAD database). *SLC12A3* encodes renal thiazide-sensitive sodium-chloride cotransporter. These variants are putatively causal (PPI 0.38 - 1.00) and functional follow-up especially conformational analyses warrant a further functional understanding of this transporter.

STAT3

STAT3 region is associated with CK. *STAT3* is a critical transcription factor with various functions on a wide variety of cells. *STAT3* is also known to play a role on muscles such as skeletal muscle in the context of response to mechanical overload[32], hypertrophy in cardiac muscle[33] and vascular smooth muscle[34]. Interestingly, recent studies suggest that induction of STAT3 phosphorylation (activation of STAT3) is associated with muscle wasting and that suppression of STAT3 phosphorylation via treatment or resistant training might be an option to prevent sarcopenia[35,36].

SMAD3

SMAD3 region is associated with CK. SMAD3 transduces TGF-beta signaling in various cells. TGF-beta/SMAD3 pathway has been shown to play roles in the proliferation of smooth muscle, cardiac muscle and skeletal muscle[37,38]. A recent study shows that a specific inhibitor of p-Smad3 canceled out induced suppression of myogenesis in vitro, suggesting that phosphorylation of SMAD3 is associated with myogenesis[39]. GDF11 is shown to induce muscle atrophy in human iPSC via activation of SMAD3[40]. These support the association found in the current study.

ZNF365

An intronic variant in *ZNF365* showed an association with HDLC. This variant was much more frequent in EAS than in EUR. Another intronic variant in *ZNF365* showed an association with basophil count. This variant also showed a large difference in allele frequencies. The *ZNF365* region is associated with metabolic and immune-related diseases including type 2 diabetes[41], inflammatory bowel diseases[42] and breast cancer[43], the current findings would provide clues of biological mechanisms underlying the associations between this region and the diseases.

ZNF468

An intronic variant in *ZNF468* and highly prevalent in EAS is associated with eGFR and BUN. ZNF468 is a zinc finger protein expressed in a wide variety of cells except for the brain and heart. This gene's function is not explored at all.

PTGER1

We found an association between eGFR or UA and an intronic variant of *PTGER1* which encodes the Prostaglandin E receptor (EP1). EP1 is known for its important role in kidney function in which antagonist of EP1 resulted in anti-renal fibrosis via decreased fibronectin alpha-smooth muscle actin[44] and deletion of this gene led to attenuation of renal injury in diabetic model[45]. Thus, our finding provides human genetic evidence between kidney function and EP1 for the first time.

RFWD2

We found an association between basophil counts and a rare missense variant of *RFWD2*, also known as *COP1*, which is specific to EAS. *RFWD2* encodes E3 ubiquitin-protein ligase which interacts with c-Jun and has many immunological functions. While previous studies reported a biological link between *RFWD2* and monocytes, microglia or B-cells[46], this is the first genetic evidence to show *RFWD2* and basophil.

SPRED2

We found an association between BUN and a rare intronic variant in *SPRED2* highly specific to EAS. We also found an association between eGFR and an upstream variant. Since *SPRED2* is known as an immune-mediator and a link between *SPRED2* and kidney function is not reported, this association suggests immunological involvement of SPRED2 in kidney function.

STK3

We found a novel association between the intron variant in *STK3* and total bilirubin level. This variant is highly specific to EAS. While direct involvement of STK3 with lipid metabolism is not shown, STK3 (also known as MST2) is a protein kinase in the Hippo pathway, which cooperates with p53 to fine-tune the sterol regulatory element-binding protein activity and regulate lipid levels[47]. These suggest a link between Serine-threonine kinase and total bilirubin to be further elucidated.

AQP1

We found a novel association between the intronic variant in *AQP1* and eGFR and serum creatinine level. AQP1 is a widely expressed water channel in the kidney. AQP1 facilitates fast water exchange across cell membranes. The carriers of the *AQP1* gene in humans and the Aqp1 knockout murine model consistently showed the impaired formation of concentrated urine. Recently, another common variant in the AQP1 promoter region was reported to be associated with decreased ultrafiltration and with a risk of death or transfer to hemodialysis due to technical failure in patients with peritoneal dialysis[48]. Although its

involvement in renal function has also been suggested, this study is the first study to find these genetic associations.

4.2 Novel variants in known associations between phenotypes and genes

PCSK9

The *PCSK9* locus, a very well-known locus for low-density lipoprotein cholesterol (LDLC) levels, showed multiple associations with high PPI including novel associations (novel variant-level associations, Extended Data Fig 7). We identified seven variants with PPI > 0.1 in this locus and as expected these variants are mutually LD-independent. Among them, 3 are missense variants and 4 are non-coding variants. All the missense variants are rare and exhibited large effect sizes (rs564427867, PCSK9 p.E32K, MAF 1.13%, β_{LDLC} = 0.51 ± 0.022, P_{LDLC} = 4.2 × 10^{-117}, PPI_{LDLC} = 1.00; rs151193009, PCSK9 p.R93C, MAF 0.93%, β_{LDLC} = -0.48 ± 0.022, P_{LDLC} = 1.3 × 10^{-105}, PPI_{LDLC} = 1.00; rs759250273, PCSK9 p.I424V, MAF 0.8%, β_{LDLC} = -0.23 ± 0.029, P_{LDLC} = 4.6 × 10^{-15}, PPI_{LDLC} = 0.79); three out of the four non-coding variants are common [MAF > 0.1]; three non-coding variants are in candidate cis-regulatory region determined by active chromatin mark[7] supporting the functionality of these variants. Among them, rs188211891 is very rare and associated with serum LDLC levels with a larger effect size than missense variants in *PCSK9* (MAF 0.001, β_{LDLC} = -0.70 ± 0.082, P_{LDLC} = 1.5 × 10^{-17}, PPI_{LDLC} = 1.00). Also, we noted these four rare coding- or non-coding putative causal variants are either not found or are, at most, rare in other populations[49], indicating that these variants are specific to the Japanese population. A total of these four rare and Japanese or EAS-specific variants (three missense and one non-coding) are novel associations.

FLT3

A rare intronic variant rs76080105 (ENST00000241453.12:c.2208-14A>G, MAF = 0.76%), specific to individuals of Asian ancestry and predicted to cause an acceptor loss of *FLT3*, was significantly and highly confidently associated with various immunological traits including leukocyte counts and clinical indicators of immunoglobulin levels ($P_{\text{White blood cell count (WBC)}}$ = 1.9 × 10^{-14}, PPI_{WBC} = 1.00, $P_{\text{Monocyte percentage (MONO)}}$ = 2.0 × 10^{-50}, PPI_{MONO} = 1.00, P_{NAP} = 7.4 × 10^{-11}, PPI_{NAP} = 1.00). *FLT3* encodes a tyrosine kinase critical for the expansion and proliferation of hematopoietic stem cells and was recently reported as a causal gene for autoimmune thyroid disease in Europeans through a distinct cryptic splicing variant. We also identified that rs76080105 was also associated with rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) in the Japanese population[43] [OR_{RA} = 1.63 (1.26 - 2.08), P_{RA} = 1.5 × 10^{-4}, OR_{SLE} = 2.16 (1.58 - 2.96), P_{SLE} = 1.6 × 10^{-6}]. We conducted a minigene assay to validate this predicted cryptic splice alteration (Fig. 3b). We confirmed that the alternate allele at rs76080105 significantly decreased the proportion of normal splicing in this junction (Fig. 3c), compared to the reference allele (P = 1.26 × 10^{-6}, Fisher's Exact test). Because of the Asian specificity of this allele, we were not able to find rs76080105 in the GTEx database.

MMP2

A rare missense variant in *MMP2* [ENST00000219070:c.1453A>T(I485F), MAF = 0.63%], specific to Japanese and associated with height ($P = 6.9 \times 10^{-9}$, PPI= 0.15). This missense variant is predicted to introduce donor gain, which resulted in 25 base-pair frameshift deletion. This predicted abnormal splicing was also experimentally validated (Fig. 3e). The alternate allele at rs141440582 significantly induced this aberrant splicing by 66-fold ($P = 2.9 \times 10^{-11}$, Fisher's Exact test). As a result, more than half of spliced transcripts from the alternate construct were affected by abnormal splicing.

G6PC2

We found novel (variant level) associations between blood glucose levels and three deleterious variants with high PPI in the *G6PC2* gene which is a very known gene and catalyzes Glucose-6-Phosphatase into glucose. We found that all the variants are rare (2.2% or lower in the current data set) and show high specificity in EAS. Since all the missense variants showed decreased levels in glucose levels, the results are reasonable (altered amino acids resulted in decreased enzymatic activity) and help us to understand G6PC2 topological functions (amino acid positions at which variants alter amino acid sequences are strongly suggested to have fundamental effects on enzymatic activity).

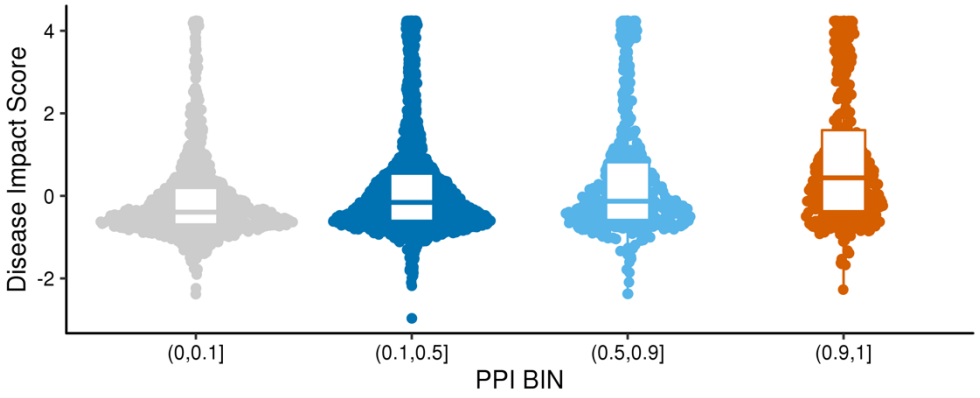
LDHB

We found novel associations between LDH and variants absent in EUR in *LDHB* which encodes Lactate Dehydrogenase B. rs542962114, a rare variant (MAF 1.13%) upstream of *LDHB*, showed a strong decreasing association with LDH level (β_{LDH} = 32% of SD, PPI_{LDH} = 1.0). Previously, the pathogenicity of this mutation has been discussed but inconclusive in ClinVar. Additionally, we found another rare variant (rs185436061, MAF 0.11%) in the CTCF region in this locus showed a strong decreasing association with LDH (β_{LDH} = 36% of SD, PPI_{LDH} = 0.998).

5 Analyses of High PPI variants

5.1 High PPI and strong effect sizes

We derived machine learning-based pathogenicity estimation for non-coding variants (Disease Impact Score). We found a significant positive association between PPI and Disease Impact Score (SI 5.1) after controlling local LD structure and minor allele frequency. This suggests that variants with high PPI have a high possibility of having biological impacts (supported by Disease Impact Score). This is also supported by the enrichment of ClinVar coding variants for high PPI (as discussed in section 5.3).



SI 5.1 | Non-coding variants with high PPI showing high Disease Impact Score

The association between PPI bins and Disease Impact Score. The disease impact score was estimated by machine learning model[50]. To avoid overestimation by linkage, we restricted the analysis to the variants with the highest PPI in each locus. The boxplots show median and interquartile range.

5.2 Non-coding variants with high PPI

If we consider the absolute effect size 0.261 (median absolute effect size of 15 protein-truncating associations with PPI > 0.9, 0.2% of all protein-truncating associations) as a threshold for the high-impact association, we identified 45 protein-altering high-impact associations (22% of protein-altering associations with PPI > 0.9, 0.018% of all protein-altering associations) and more than twice as many non-coding high-impact associations (99 associations, 16% of non-coding associations with PPI > 0.9, 0.00043% of all non-coding associations). These indicate that reflecting the majority of associations are in non-coding regions, more than two third of variants with high PPI are non-coding.

5.3 Enrichment for pathogenic variants in ClinVar

F7

rs387906507, the missense variant on *F7* (factor seven), which we found associated with prothrombin time (PT, MAF 0.12%, β_{PT} = 0.569 ± 0.089, P_{PT} = 1.3 × 10^{-10}, PPI_{PT} = 1.0), a common clinical indicator of blood coagulation. This variant is known as a causal variant for Factor VII deficiency.

ALPL

Another sample is rs387906525, a frameshift deletion of the *ALPL* (alkaline phosphatase) gene. rs387906525 is very rare with the second largest effect size in this study on the serum alkaline phosphatase (ALP, MAF 0.26%, β_{ALP} = -1.65 ± 0.050, P_{ALP} = 8.7 × 10^{-239}, PPI_{ALP} = 1.00). This variant is a causal variant of infantile hypophosphatasia, which is manifested by rickets due to hypomineralization[51,52].

CD36

We identified rs75326924, a missense variant in the *CD36* gene, as a putative causal variant for multiple

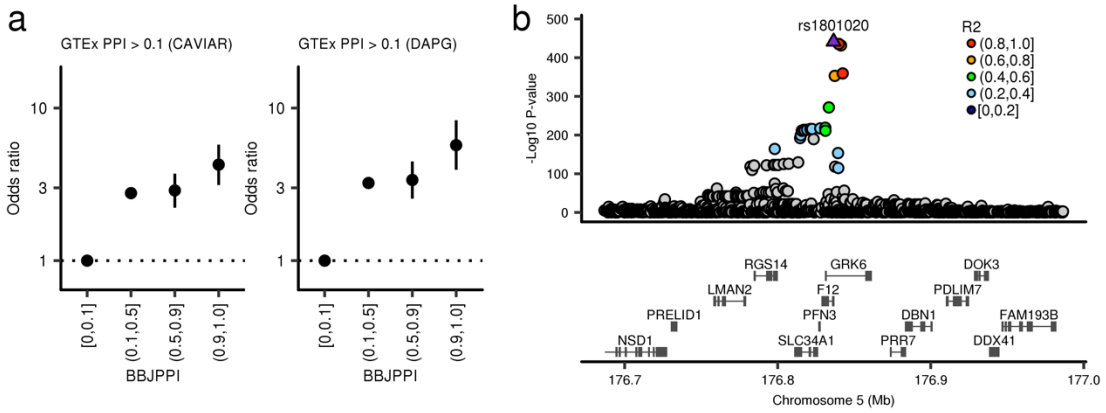
quantitative traits including platelet counts (PLT, MAF = 4.6% β_{PLT} = -0.102 ± 0.0081, P_{PLT} = 4.8 × 10^{-36}, PPI_{PLT} = 1.0, Fig. 3d). This variant is known to be causal for *CD36* deficiency[53], a disease of genetic platelet glycoprotein IV deficiency and relatively common in Asians. *CD36* deficiency is also known as a cause of dilated cardiomyopathy through dysregulated fatty acid metabolism[54]. In concordance with these clinical findings, we also found that the carriers with this variant showed lower cardiac contractile function [ejection fraction (EF)] and an increased risk of heart failure [β_{EF} = -0.059 ± 0.021, P_{EF} = 4.9 × 10^{-3}, $OR_{\text{Heart Failure}}$ = 1.14 (1.05 - 1.23), $P_{\text{Heart Failure}}$ = 1.0 × 10^{-3}].

6 Functional enrichment of non-coding variants

6.1 Enrichment for causal eQTL variants in the GTEx

We addressed the overlap of variants with high PPI between our results and GTEx to show putative causal variants in our study may be partly explained by altering gene expression.

We observed the significant overlap of finemapped variants by our study and GTEx (PPI > 0.1, by CAVIAR or DAPG, available in the GTEx portal), indicating the contribution of eQTL variants in the causal variants associated with quantitative traits. This enrichment was consistent between CAVIAR and DAPG (SI 6.1a). As an example, we showed that the causal variant in the *F12* gene at 5'UTR associated with APTT (PPI: 0.9999 and effect size of -0.26 SD, SI 6.1b) is also a causal eQTL variant of this gene in the liver (SI 6.1b, Supplementary Table 23). Since this variant (G allele) is associated with decreased APTT (indicating over coagulation) and increases expression of *F12*, one of the coagulation factors which facilitate blood coagulation, this overlap can explain the variant-phenotype association via altered expression of *F12*.



SI 6.1 | Enrichment of causal variants in eQTL for finemapped variants

a. Enrichment of causal variants in eQTL for finemapped variants The X axes show PPI bins in BBJ GWAS.

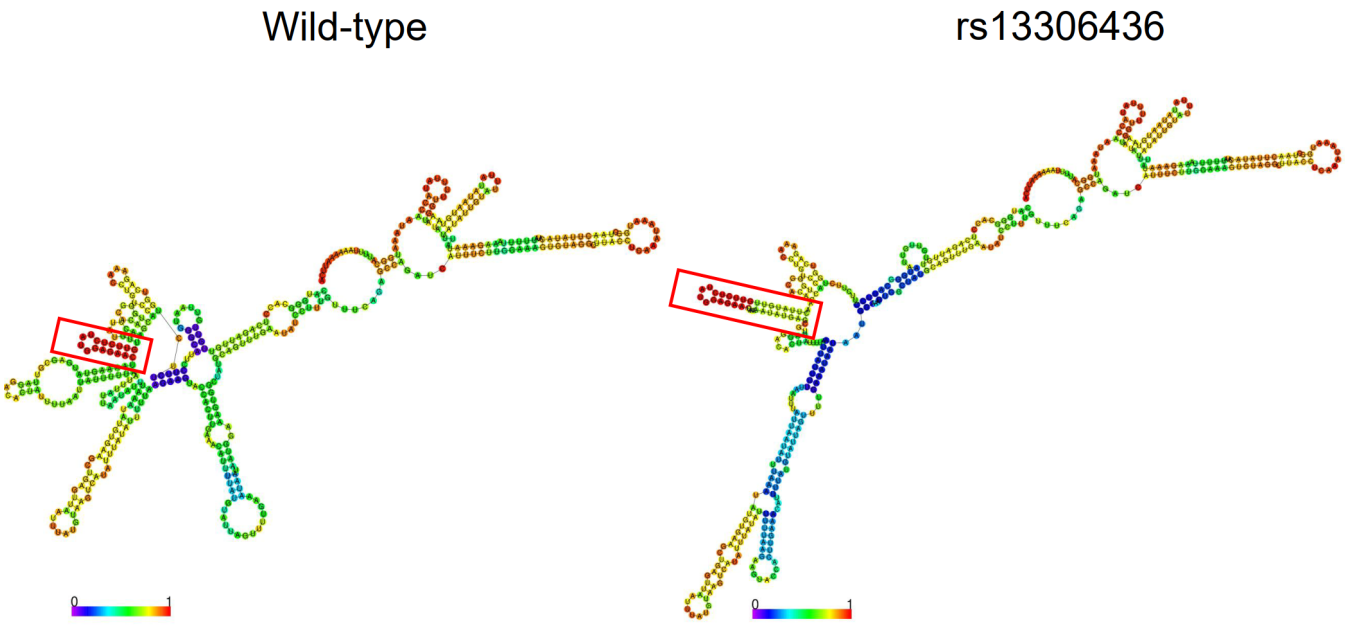
The Y axes show odds ratio and error bars show 95% confidence interval estimated by Fisher's exact test. **b.** X-axes show the genomic coordinate. Y-axis in the top panel shows -$\log_{10}$ *P*-value. The bottom panel shows genes located in the loci.

6.2 Functional enrichment using the UKB data

Since we obtained enrichment of putative causal variants for functional annotations in the Japanese data, we performed finemapping in the UKB dataset to test generalizability of the enrichments. We took summary statistics of UKB for a total of 30 traits available in both populations (http://www.nealelab.is/uk-biobank). We took advantages of LD structure of imputed UKB genotypes. We applied the same methods to UKB finemapped variants as the Japanese results. As a result, we found comparable enrichment of causal variants for the functional annotations (Extended Data Fig. 9c).

6.3 Functional consequences of rs13306436 at 3'UTR in *IL6*

We observed a significant increase in luciferase activity in rs13306436 alternative allele in comparison with the reference allele ($P < 0.05$), suggesting more stable mRNA in rs13306436 (since only the difference in mRNA between the two conditions is the presence of this variant in 3'UTR sequences). Since *IL6* is a known target of Regnase-1 which interacts with 3'UTR of mRNA and degrades mRNA, we hypothesized that rs13306436 would alter the stability of mRNA and lead to resistance to Regnase-1-mediated degradation. We experimentally observed increased resistance to mRNA degradation measured by luciferase activity in rs13306436 alternative allele in the condition of overexpression of Regnase-1 (Fig 4c). In line with these findings, the RNAfold (version 2.4.18)[55,56] predicted the altered structure of stem-loop in mRNA of *IL6* in rs13306436 (SI 6.3).



SI 6.3 | Altered stem-loop structure of *IL6* mRNA in rs13306436

We computed RNA secondary structure of 3'UTR of *IL6* by RNAfold software[55,56].

6.4 Enrichment for target genes of Regnase-1

Since we observed the causal non-coding variant at 3'UTR in IL6 showed the decreased effect of Regnase-1, we analyzed whether other genes targeted by Regnase-1 are present in our finemapping results. We found a total of 136 genes where causal variants are present in 3'UTR. Experimental results validated a total of 56 target genes of Regnase-1 among 18.8k genes in multiple cell lines[57]. We found a significant enrichment of the 136 genes for the Regnase-1 target genes (OR = 10.6, $P = 5.2 \times 10^{-5}$, hypergeometric test).

References

1 Hozawa, A. *et al.* Study Profile of the Tohoku Medical Megabank Community-Based Cohort Study. *J Epidemiol* **31**, 65-76 (2021). https://doi.org:10.2188/jea.JE20190271

2 Fuse, N. *et al.* Genome-wide Association Study of Axial Length in Population-based Cohorts in Japan: The Tohoku Medical Megabank Organization Eye Study. *Ophthalmology Science* **2**, 100113 (2022). https://doi.org:https://doi.org/10.1016/j.xops.2022.100113

3 Tadaka, S. *et al.* 3.5KJPNv2: an allele frequency panel of 3552 Japanese individuals including the X chromosome. *Hum Genome Var* **6**, 28 (2019). https://doi.org:10.1038/s41439-019-0059-5

4 Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493-1501 (2016). https://doi.org:10.1093/bioinformatics/btw018

5 Wellcome Trust Case Control, C. *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet* **44**, 1294-1301 (2012). https://doi.org:10.1038/ng.2435

6 Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-838 (2010). https://doi.org:10.1038/nature09410

7 Schizophrenia Psychiatric Genome-Wide Association Study, C. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* **43**, 969-976 (2011). https://doi.org:10.1038/ng.940

8 Kanai, M. *et al.* Meta-analysis fine-mapping is often miscalibrated at single-variant resolution. *medRxiv* (2022).

9 Jaganathan, K. *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **176**, 535-548 e524 (2019). https://doi.org:10.1016/j.cell.2018.12.015

10 Kanai, M. *et al.* Insights from complex trait fine-mapping across diverse populations. *medRxiv* (2021).

11 Van Limbergen, H. *et al.* Molecular cytogenetic analysis of 10;11 rearrangements in acute myeloid leukemia. *Leukemia* **16**, 344-351 (2002). https://doi.org:10.1038/sj.leu.2402397

12 Jakobczyk, H. *et al.* ETV6-RUNX1 and RUNX1 directly regulate RAG1 expression: one more step in the understanding of childhood B-cell acute lymphoblastic leukemia leukemogenesis. *Leukemia* **36**, 549-554 (2022). https://doi.org:10.1038/s41375-021-01409-9

13 Montano-Almendras, C. P. *et al.* ETV6-PDGFRB and FIP1L1-PDGFRA stimulate human hematopoietic progenitor cell proliferation and differentiation into eosinophils: the role of nuclear factor-kappaB. *Haematologica* **97**, 1064-1072 (2012). https://doi.org:10.3324/haematol.2011.047530

14 Breyer, F. *et al.* TPL-2 kinase induces phagosome acidification to promote macrophage killing of bacteria. *EMBO J* **40**, e106188 (2021). https://doi.org:10.15252/embj.2020106188

15 Louie, K. W., Mishina, Y. & Zhang, H. Molecular and Cellular Pathogenesis of Ellis-van Creveld Syndrome: Lessons from Targeted and Natural Mutations in Animal Models. *J Dev Biol* **8** (2020). https://doi.org:10.3390/jdb8040025

16 Desch, K. C. *et al.* Whole-exome sequencing identifies rare variants in STAB2 associated with venous thromboembolic disease. *Blood* **136**, 533-541 (2020). https://doi.org:10.1182/blood.2019004161

17 Xiang, K. *et al.* Identification of a Tibetan-specific mutation in the hypoxic gene EGLN1 and its contribution to high-altitude adaptation. *Mol Biol Evol* **30**, 1889-1898 (2013). https://doi.org:10.1093/molbev/mst090

18 Luo, C. *et al.* LINC01094 promotes pancreatic cancer progression by sponging miR-577 to regulate LIN28B expression and the PI3K/AKT pathway. *Mol Ther Nucleic Acids* **26**, 523-535 (2021). https://doi.org:10.1016/j.omtn.2021.08.024

19 Wu, X., Kong, C. & Wu, Y. Long intergenic non-protein coding RNA 1094 (LINC01094) promotes the progression of breast cancer (BC) by regulating the microRNA-340-5p (miR-340-5p)/E2F transcription factor 3 (E2F3) axis. *Bioengineered* **12**, 9046-9057 (2021). https://doi.org:10.1080/21655979.2021.1993715

20 Chang, M. S. & Hsu, Y. H. The role of IL-20 in chronic kidney disease and diabetic nephropathy: Pathogenic and therapeutic implications. *J Leukoc Biol* **104**, 919-923 (2018). https://doi.org:10.1002/JLB.MR1217-489R

21 Wei, C. C. *et al.* Interleukin-20 targets renal cells and is associated with chronic kidney disease. *Biochem Biophys Res Commun* **374**, 448-453 (2008). https://doi.org:10.1016/j.bbrc.2008.07.039

22 Li, Y. *et al.* Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi. *Nat Med* **22**, 952-960 (2016). https://doi.org:10.1038/nm.4139

23 Fu, S. *et al.* Overexpression of MYCT1 Inhibits Proliferation and Induces Apoptosis in Human Acute Myeloid Leukemia HL-60 and KG-1a Cells in vitro and in vivo. *Front Pharmacol* **9**, 1045 (2018). https://doi.org:10.3389/fphar.2018.01045

24 Wongborisuth, C. *et al.* Down-regulation of the transcriptional repressor ZNF802 (JAZF1) reactivates fetal hemoglobin in beta(0)-thalassemia/HbE. *Sci Rep* **12**, 4952 (2022). https://doi.org:10.1038/s41598-022-08920-8

25 Rowland, A. & Mangoni, A. A. Cytochrome P450 and ischemic heart disease: current concepts and future directions. *Expert Opin Drug Metab Toxicol* **10**, 191-213 (2014). https://doi.org:10.1517/17425255.2014.859675

26 Dai, Y., Chen, T., Ijaz, H., Cho, E. H. & Steinberg, M. H. SIRT1 activates the expression of fetal hemoglobin genes. *Am J Hematol* **92**, 1177-1186 (2017). https://doi.org:10.1002/ajh.24879

27 Tu, Z. *et al.* The chromatin remodeler CHD8 governs hematopoietic stem/progenitor survival by regulating ATM-mediated P53 protein stability. *Blood* **138**, 221-233 (2021). https://doi.org:10.1182/blood.2020009997

28 Chen, A. *et al.* Dock5 controls the peripheral B cell differentiation via regulating BCR signaling and actin reorganization. *Cell Immunol* **337**, 15-21 (2019). https://doi.org:10.1016/j.cellimm.2019.01.001

29 Tulasne, D., Bori, T. & Watson, S. P. Regulation of RAS in human platelets. Evidence that activation of RAS is not sufficient to lead to ERK1-2 phosphorylation. *Eur J Biochem* **269**, 1511-1517 (2002). https://doi.org:10.1046/j.1432-1033.2002.02798.x

30 Group, C. C. H. W. Meta-analysis of rare and common exome chip variants identifies S1PR4 and other loci influencing blood cell traits. *Nat Genet* **48**, 867-876 (2016). https://doi.org:10.1038/ng.3607

31 Menniti, F. S., Chappie, T. A., Humphrey, J. M. & Schmidt, C. J. Phosphodiesterase 10A inhibitors: a novel approach to the treatment of the symptoms of schizophrenia. *Curr Opin Investig Drugs* **8**, 54-59 (2007).

32 Laskin, G. R. & Gordon, B. S. The influence of nutrients on mechanical overload-induced changes to skeletal muscle mRNA content. *Physiol Genomics* **54**, 360-369 (2022). https://doi.org:10.1152/physiolgenomics.00075.2022

33 Mao, L. *et al.* Heart-targeting exosomes from human cardiosphere-derived cells improve the therapeutic effect on cardiac hypertrophy. *J Nanobiotechnology* **20**, 435 (2022). https://doi.org:10.1186/s12951-022-01630-3

34 Wang, L. *et al.* Adropin Inhibits Vascular Smooth Muscle Cell Osteogenic Differentiation to Alleviate Vascular Calcification via the JAK2/STAT3 Signaling Pathway. *Biomed Res Int* **2022**, 9122264 (2022). https://doi.org:10.1155/2022/9122264

35 Wijaya, Y. T. *et al.* Ginsenoside Rd ameliorates muscle wasting by suppressing the signal transducer and activator of transcription 3 pathway. *J Cachexia Sarcopenia Muscle* (2022). https://doi.org:10.1002/jcsm.13084

36 Testa, M. T. J. *et al.* Resistance Training Attenuates Activation of STAT3 and Muscle Atrophy in Tumor-Bearing Mice. *Front Oncol* **12**, 880787 (2022). https://doi.org:10.3389/fonc.2022.880787

37 Hitachi, K. & Tsuchida, K. Role of microRNAs in skeletal muscle hypertrophy. *Front Physiol* **4**, 408 (2013). https://doi.org:10.3389/fphys.2013.00408

38 Guo, Y. *et al.* Entanglement of GSK-3beta, beta-catenin and TGF-beta1 signaling network to regulate myocardial fibrosis. *J Mol Cell Cardiol* **110**, 109-120 (2017). https://doi.org:10.1016/j.yjmcc.2017.07.011

39 Zhang, P. *et al.* Salidroside Inhibits Myogenesis by Modulating p-Smad3-Induced Myf5 Transcription. *Front Pharmacol* **9**, 209 (2018). https://doi.org:10.3389/fphar.2018.00209

40 Honda, M. *et al.* Pathophysiological levels of GDF11 activate Smad2/Smad3 signaling and induce muscle atrophy in human iPSC-derived myocytes. *Am J Physiol Cell Physiol* (2022). https://doi.org:10.1152/ajpcell.00341.2022

41 Kim, H., Bae, J. H., Park, K. S., Sung, J. & Kwak, S. H. DNA Methylation Changes Associated With Type 2 Diabetes and Diabetic Kidney Disease in an East Asian Population. *J Clin Endocrinol Metab* **106**, e3837-e3851 (2021). https://doi.org:10.1210/clinem/dgab488

42 Umeno, J. *et al.* Meta-analysis of published studies identified eight additional common susceptibility loci for Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* **17**, 2407-2415 (2011). https://doi.org:10.1002/ibd.21651

43 Lindstrom, S. *et al.* Common variants in ZNF365 are associated with both mammographic density and breast cancer risk. *Nat Genet* **43**, 185-187 (2011). https://doi.org:10.1038/ng.760

44 Kresse, J. C. *et al.* EP1 receptor antagonism mitigates early and late stage renal fibrosis. *Acta Physiol (Oxf)* **234**, e13780 (2022). https://doi.org:10.1111/apha.13780

45 Thibodeau, J. F. *et al.* PTGER1 deletion attenuates renal injury in diabetic mouse models. *Am J Pathol* **183**, 1789-1802 (2013). https://doi.org:10.1016/j.ajpath.2013.08.022

46 Ndoja, A. *et al.* Ubiquitin Ligase COP1 Suppresses Neuroinflammation by Degrading c/EBPbeta in Microglia. *Cell* **182**, 1156-1169 e1112 (2020). https://doi.org:10.1016/j.cell.2020.07.011

47 Aylon, Y. & Oren, M. The Hippo pathway, p53 and cholesterol. *Cell Cycle* **15**, 2248-2255 (2016). https://doi.org:10.1080/15384101.2016.1207840

48 Morelle, J. *et al.* AQP1 Promoter Variant, Water Transport, and Outcomes in Peritoneal Dialysis. *N Engl J Med* **385**, 1570-1580 (2021). https://doi.org:10.1056/NEJMoa2034279

49 Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434-443 (2020). https://doi.org:10.1038/s41586-020-2308-7

50 Zhou, J. *et al.* Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. *Nat Genet* **51**, 973-980 (2019). https://doi.org:10.1038/s41588-019-0420-0

51 Komaru, K. *et al.* Novel aggregate formation of a frame-shift mutant protein of tissue-nonspecific alkaline phosphatase is ascribed to three cysteine residues in the C-terminal extension. Retarded secretion and proteasomal degradation. *FEBS J* **272**, 1704-1717 (2005). https://doi.org:10.1111/j.1742-4658.2005.04597.x

52 Michigami, T. *et al.* Common mutations F310L and T1559del in the tissue-nonspecific alkaline phosphatase gene are related to distinct phenotypes in Japanese patients with hypophosphatasia. *Eur J Pediatr* **164**, 277-282 (2005). https://doi.org:10.1007/s00431-004-1612-9

53 Hanawa, H. *et al.* Identification of cryptic splice site, exon skipping, and novel point mutations in type I CD36 deficiency. *J Med Genet* **39**, 286-291 (2002). https://doi.org:10.1136/jmg.39.4.286

54 Shu, H. *et al.* The role of CD36 in cardiovascular disease. *Cardiovasc Res* **118**, 115-129 (2022). https://doi.org:10.1093/cvr/cvaa319

55 Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26 (2011). https://doi.org:10.1186/1748-7188-6-26

56 Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res* **36**, W70-74 (2008). https://doi.org:10.1093/nar/gkn188

57 Mino, T. *et al.* Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms. *Cell* **161**, 1058-1073 (2015). https://doi.org:10.1016/j.cell.2015.04.029

Supplementary Tables

(Supplementary Tables 1-15 are provided in a excel file)

Supplementary Table 16 | Summary of the three data sets using the same imputation reference panel in the current study.

Cohort	Gender	N	Mean age	SD age
BBJ 1st	Male	94,304	63.58	13.37
	Female	82,590	62.60	14.95
BBJ 2nd	Male	7,647	70.45	9.18
	Female	4,451	65.22	12.36
NCGG	Male	6,296	69.17	14.15
	Female	7,928	71.96	12.86

BBJ, Biobank Japan; NCGG, National Center for Geriatrics and Gerontology

Supplementary Table 17 | Consistent associations between the three data sets and the ToMMo.

Lead variants in the BBJ 1st + BBJ 2nd + NCGG	Count	Beta consistent
ToMMo $P \leq 1$	1,528 (100%)	1,505 (98.5%)
ToMMo $P < 5 \times 10^{-2}$	1,305 (85.4%)	1,304 (99.9%)
ToMMo $P < 5 \times 10^{-4}$	941 (64.7%)	941 (100%)
ToMMo $P < 5 \times 10^{-6}$	647 (42.3%)	674 (100%)
ToMMo $P < 5 \times 10^{-8}$	464 (30.4%)	464 (100%)

Supplementary Table 18 | The comparison of the PPI inferred by FINEMAP and SuSiE.

		PPI by FINEMAP		
		[0, 0.1]	(0.1,0.9]	(0.9,1.0]
PPI by SuSiE	[0, 0.1]	22,255,416	1,049	56
	(0.1,0.9]	1,224	7,238	66

	(0.9,1.0]	51	57	649

Supplementary Table 19 | Improved performance of transancestral PRS by statistical finemapping

	Median Rho [1st - 3rd quantile]		*P*
	PT PRS	Finemapped PRS	
Derived from BBJ_{1st} tested in BBJ_{2nd}	0.191[0.142 - 0.278]	0.209[0.151 - 0.278]	3.7×10^{-7}
Derived from BBJ_{meta}, tested in UKB	0.120[0.095 - 0.171]	0.128[0.100 - 0.199]	1.4×10^{-5}
Derived from UKB, tested in BBJ_{1st}	0.144[0.108 - 0.175]	0.161[0.125 - 0.212]	5.9×10^{-6}

PRS performances as median Rho and its interquartile ranges in 34 (intra Japanese analysis) and 35 (trans ancestry analysis) quantitative phenotypes are shown. P values were calculated by Paired Wilcoxon rank sum test. PRS, polygenic risk score; PT, pruning and thresholding; BBJ, Biobank Japan; UKB, UK-Biobank.

Supplementary Table 20 | The trans ancestral performance of PT-PRS and finemapped PRS

Phenotype	Derived from	Tested in	Pruning and thresholding PRS			Finemapped PRS	
			N variants	P;R^2	Rho[95%CI]	N variants	Rho[95%CI]
ALB	BBJ	UKB	478	5e-4;0.7	0.079[0.073-0.086]	33	0.079[0.073-0.086]
	UKB	BBJ	631	5e-8;0.2	0.088[0.082-0.095]	356	0.109[0.103-0.116]
ALP	BBJ	UKB	1325	5e-5;0.8	0.127[0.121-0.133]	75	0.154[0.148-0.160]
	UKB	BBJ	2608	5e-7;0.3	0.173[0.166-0.179]	547	0.287[0.281-0.293]
ALT	BBJ	UKB	897	5e-4;0.5	0.093[0.087-0.099]	67	0.108[0.102-0.114]
	UKB	BBJ	782	5e-6;0.1	0.103[0.097-0.108]	352	0.129[0.123-0.135]
AST	BBJ	UKB	691	5e-5;0.5	0.073[0.067-0.079]	71	0.097[0.091-0.103]
	UKB	BBJ	8365	5e-3;0.3	0.098[0.092-0.104]	408	0.129[0.123-0.135]
BASO	BBJ	UKB	87	5e-8;0.1	0.029[0.023-0.035]	79	0.029[0.022-0.035]
	UKB	BBJ	203	5e-8;0.7	0.074[0.067-0.082]	43	0.062[0.054-0.070]
BMI	BBJ	UKB	1197	5e-5;0.7	0.097[0.091-0.102]	117	0.091[0.085-0.097]
	UKB	BBJ	2263	5e-5;0.1	0.119[0.114-0.125]	697	0.118[0.113-0.124]
BS	BBJ	UKB	47	5e-8;0.1	0.111[0.104-0.117]	35	0.095[0.089-0.102]
	UKB	BBJ	1063	5e-6;0.6	0.112[0.105-0.120]	143	0.121[0.114-0.128]
BUN	BBJ	UKB	338	5e-7;0.4	0.090[0.084-0.096]	99	0.098[0.092-0.104]
	UKB	BBJ	518	5e-8;0.2	0.127[0.121-0.133]	271	0.143[0.137-0.149]
CA	BBJ	UKB	30	5e-8;0.1	0.111[0.105-0.118]	27	0.113[0.107-0.119]
	UKB	BBJ	1689	5e-6;0.4	0.103[0.096-0.111]	322	0.117[0.109-0.124]
CRP	BBJ	UKB	227	5e-7;0.9	0.179[0.173-0.185]	15	0.188[0.182-0.194]

	UKB	BBJ	1376	5e-6;0.2	0.082[0.074-0.089]	327	0.109[0.102-0.117]
DBP	BBJ	UKB	262	5e-5;0.7	0.034[0.028-0.041]	24	0.046[0.040-0.052]
	UKB	BBJ	2258	5e-4;0.2	0.058[0.052-0.064]	307	0.069[0.063-0.075]
EOSINO	BBJ	UKB	4029	5e-3;0.9	0.130[0.124-0.136]	97	0.151[0.145-0.157]
	UKB	BBJ	2904	5e-5;0.2	0.174[0.167-0.181]	575	0.191[0.183-0.198]
GTP	BBJ	UKB	5513	5e-3;0.9	0.113[0.107-0.119]	93	0.173[0.167-0.179]
	UKB	BBJ	2110	5e-6;0.2	0.175[0.169-0.181]	523	0.241[0.235-0.247]
HBA1C	BBJ	UKB	74	5e-8;0.2	0.134[0.128-0.140]	56	0.142[0.136-0.148]
	UKB	BBJ	1047	5e-8;0.1	0.170[0.160-0.181]	616	0.186[0.175-0.196]
HB	BBJ	UKB	324	5e-6;0.4	0.125[0.119-0.131]	91	0.123[0.117-0.129]
	UKB	BBJ	1317	5e-7;0.2	0.125[0.119-0.131]	554	0.141[0.135-0.146]
HDLC	BBJ	UKB	1962	5e-5;0.7	0.222[0.216-0.228]	105	0.264[0.258-0.270]
	UKB	BBJ	4269	5e-5;0.4	0.218[0.212-0.225]	488	0.257[0.251-0.264]
HEIGHT	BBJ	UKB	25800	5e-3;0.8	0.271[0.265-0.276]	724	0.269[0.263-0.274]
	UKB	BBJ	13979	5e-5;0.3	0.288[0.283-0.293]	1909	0.310[0.305-0.315]
HT	BBJ	UKB	505	5e-5;0.4	0.125[0.119-0.131]	101	0.123[0.117-0.129]
	UKB	BBJ	23290	5e-3;0.8	0.127[0.121-0.133]	497	0.146[0.140-0.152]
LDL	BBJ	UKB	976	5e-7;0.9	0.132[0.126-0.138]	63	0.207[0.201-0.212]
	UKB	BBJ	2343	5e-5;0.4	0.116[0.109-0.124]	234	0.161[0.154-0.168]
LYMPH	BBJ	UKB	1570	5e-4;0.9	0.116[0.110-0.122]	63	0.126[0.119-0.132]
	UKB	BBJ	8168	5e-4;0.4	0.145[0.138-0.152]	651	0.165[0.157-0.172]
MCHC	BBJ	UKB	450	5e-7;0.3	0.090[0.084-0.096]	114	0.103[0.097-0.109]
	UKB	BBJ	274	5e-7;0.1	0.147[0.141-0.152]	139	0.152[0.146-0.158]
MCH	BBJ	UKB	1899	5e-5;0.4	0.268[0.263-0.274]	239	0.304[0.298-0.309]
	UKB	BBJ	2844	5e-7;0.3	0.283[0.277-0.288]	631	0.306[0.300-0.311]
MCV	BBJ	UKB	3134	5e-5;0.6	0.261[0.256-0.267]	284	0.299[0.294-0.305]
	UKB	BBJ	3874	5e-6;0.3	0.287[0.282-0.293]	696	0.304[0.299-0.310]
MONO	BBJ	UKB	133	5e-8;0.1	0.194[0.188-0.199]	99	0.192[0.186-0.198]
	UKB	BBJ	2420	5e-7;0.3	0.176[0.169-0.183]	660	0.205[0.198-0.212]
NEUTRO	BBJ	UKB	112	5e-7;0.2	0.101[0.095-0.107]	44	0.122[0.116-0.128]
	UKB	BBJ	1686	5e-6;0.2	0.159[0.151-0.167]	533	0.167[0.159-0.175]
P	BBJ	UKB	24	5e-8;0.5	0.092[0.086-0.098]	12	0.097[0.090-0.103]
	UKB	BBJ	598	5e-6;0.1	0.098[0.088-0.109]	252	0.115[0.105-0.126]
PLT	BBJ	UKB	1456	5e-5;0.3	0.174[0.169-0.180]	270	0.237[0.232-0.243]
	UKB	BBJ	10217	5e-4;0.3	0.228[0.222-0.234]	928	0.254[0.249-0.260]
RBC	BBJ	UKB	832	5e-7;0.5	0.195[0.189-0.200]	183	0.213[0.207-0.219]
	UKB	BBJ	2001	5e-7;0.2	0.204[0.198-0.209]	738	0.220[0.214-0.225]
SBP	BBJ	UKB	577	5e-6;0.9	0.061[0.055-0.067]	38	0.067[0.061-0.074]
	UKB	BBJ	78953	5e-1;0.4	0.062[0.056-0.068]	381	0.092[0.086-0.098]

SCR	BBJ	UKB	292	5e-7;0.3	0.118[0.113-0.124]	131	0.140[0.134-0.146]
	UKB	BBJ	1646	5e-6;0.1	0.140[0.135-0.146]	648	0.160[0.154-0.166]
TBIL	BBJ	UKB	338	5e-8;0.6	0.380[0.375-0.385]	29	0.399[0.394-0.404]
	UKB	BBJ	12054	5e-3;0.9	0.224[0.218-0.230]	178	0.185[0.178-0.191]
TC	BBJ	UKB	230	5e-7;0.1	0.167[0.161-0.173]	99	0.212[0.206-0.218]
	UKB	BBJ	5997	5e-5;0.8	0.143[0.138-0.149]	267	0.182[0.176-0.188]
TG	BBJ	UKB	495	5e-5;0.2	0.168[0.162-0.174]	78	0.189[0.182-0.194]
	UKB	BBJ	1419	5e-8;0.3	0.195[0.189-0.201]	414	0.255[0.248-0.261]
TP	BBJ	UKB	3299	5e-3;0.8	0.110[0.104-0.117]	94	0.127[0.121-0.133]
	UKB	BBJ	613	5e-8;0.1	0.135[0.129-0.141]	457	0.156[0.150-0.162]
WBC	BBJ	UKB	234	5e-7;0.1	0.120[0.114-0.126]	168	0.128[0.122-0.134]
	UKB	BBJ	876	5e-8;0.1	0.164[0.158-0.169]	648	0.185[0.180-0.191]

N variants show the number of variants in PRS. P/R^2 shows the optimized threshold for PT-PRS in validation dataset (25% of dataset). We tested a range of P-value and R^2 threshold (*P* = 0.5, 0.05, 5e-3, 5e-4, 5e-5, 5e-6, 5e-7, 5e-8, and R^2 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9). Rho[95%CI] show Spearman's rank correlation coefficients and 95% confidence intervals in the testing dataset (75% withheld dataset). PRS, polygenic risk score; PT, pruning and thresholding; BBJ, Biobank Japan; UKB, UK-Biobank.

Supplementary Table 21 | Concordant allelic effect of putatively causal variants between BBJ and UKB

Derived	Tested	Median sign concordance [1st - 3rd quantile]		*P*
		Pruning and thresholding PRS	Finemapped PRS	
BBJ	UKB	0.858[0.805-0.884]	0.899[0.884-0.924]	8.7×10^{-5}
UKB	BBJ	0.784[0.756-0.797]	0.821[0.790-0.848]	5.0×10^{-6}

For 35 studied traits, the median [1st-3rd quantile] sign concordances were shown. The sign concordances of effects of variants included in each PRS were calculated by the comparison of effect direction between BBJ and UKB summary statistics. For PT-PRS, we used the best concordant PRS irrespective to prediction performance. The differences were tested by the paired Wilcoxon rank sum test.

Supplementary Table 22 | Improved PRS transferability in the new reference panel

Derived	Tested	Median Rho [1st - 3rd quantile]		*P*
		1KG reference panel	BBJ reference panel	
BBJ	UKB	0.123[0.093-0.196]	0.132[0.102-0.204]	0.0014

We conducted GWAS, Finemapping, and PRS construction from the different imputation reference panel. 1KG reference panel indicates panel generated from 1KG entire population (n = 2,504). BBJ reference panel indicates panel generated in this study (n = 3,256 Japanese + 2,504 1KG participants). For 35 traits validated in UKB, the median [1st-3rd quantile] correlation coefficients are shown. The difference was tested

by the paired Wilcoxon rank sum test. GWAS, genome wide association study; 1KG, 1000 genomes project; BBJ, BioBank Japan; UKB, UK-Biobank.

Supplementary Table 23 | The causal variant of *F12* gene expression in the liver is causal for the APTT association

	PPI	Effect size	P
APTT GWAS	99.99%	-0.26	3.4×10^{-442}
F12 Liver eQTL	98.40%	0.74	8.60×10^{-37}

Supplementary Table 24 | Oligonucleotide sequence used for splicing assay

FLT3	Ref	TTACGCCAAGTTATTTAGGTGACACCAATTTACTTGATTTTTGAATACTGTTGCTATGGTGATC TTCTCAACTATCTAAGAAGTAAAAGAGAAAAATTTCACAGGACTTGGACAGAGATTTTCAAGGA ACACAATTTCAGTTTTTACCCCACTTTCCAATCACATCCAAATTCCAGGTAAGAGGCTGGGTCA GGGTTTCGTAATTACACATCATAGAACGTAGGCATACACAAAAAATGGACAAATGGATGGTGTG CTTTTAATTTAAAAGTCAATTTATAACTGGAATCCATTTATTTGATGAATTACATTTTTAAAAT AAAATGTCTGCTGAGAAGAATAATGTAAATAATGAAAAAACAAAATTTTTTAATAGCATGCCTG GTTCAAGAGAAGTTCAGATACACCCGGACTCGGATCAAATCTCAGGGCTTCATGGGAATTCATT TCACTCTGAAGAAATCTAGATAACTGATCATAATCAGCCATACCACATTTGT
	ALT	TTACGCCAAGTTATTTAGGTGACACCAATTTACTTGATTTTTGAATACTGTTGCTATGGTGATC TTCTCAACTATCTAAGAAGTAAAAGAGAAAAATTTCACAGGACTTGGACAGAGATTTTCAAGGA ACACAATTTCAGTTTTTACCCCACTTTCCAATCACATCCAAATTCCAGGTAAGAGGCTGGGTCA GGGTTTCGTAATTACACATCATAGAACGTAGGCATACACAAAAAATGGACAAATGGATGGTGTG CTTTTAATTTAAAAGTCAATTTATAACTGGAATCCATTTATTTGATGAATTACATTTTTAAAAT AAAATGTCTGCTGAGAAGAATAATGTAAATAATGAAAAAACAGAATTTTTTAATAGCATGCCTG GTTCAAGAGAAGTTCAGATACACCCGGACTCGGATCAAATCTCAGGGCTTCATGGGAATTCATT TCACTCTGAAGGCATCTAGATAACTGATCATAATCAGCCATACCACATTTGT
MMP2	REF	TTACGCCAAGTTATTTAGGTGACAGGGCCTCTCCTGACATTGACCTTGGCACCGGCCCCACCCC CACGCTGGGCCCTGTCACTCCTGAGATCTGCAAACAGGACATTGTATTTGATGGCATCGCTCAG ATCCGTGGTGAGATCTTCTTCTTCAAGGACCGGTGAGTGCAGGAGCTTGCTTCTTGTCCTCCTT GTCTCCTGTCCTCTGCTCTTATACCATTATTCTTTTCCCTCACTCTTCGCTGAAGACTCCGCCA AATGCTTCCCAGAGGTGGGTTTGGGGGTGTGTGTGGTTCGAGCTGCAGGGTGACTGAAGATGTG GTTTCCTGTGCCCCCTTGCCTCCTGCCAGGTTCATTTGGCGGACTGTGACGCCACGTGACAAGC CCATGGGGCCCCTGCTGGTGGCCACATTCTGTATCTAGATAACTGATCATAATCAGCCATACCA CATTTGT
	ALT	TTACGCCAAGTTATTTAGGTGACAGGGCCTCTCCTGACATTGACCTTGGCACCGGCCCCACCCC CACGCTGGGCCCTGTCACTCCTGAGATCTGCAAACAGGACATTGTATTTGATGGCATCGCTCAG ATCCGTGGTGAGTTCTTCTTCTTCAAGGACCGGTGAGTGCAGGAGCTTGCTTCTTGTCCTCCTT GTCTCCTGTCCTCTGCTCTTATACCATTATTCTTTTCCCTCACTCTTCGCTGAAGACTCCGCCA AATGCTTCCCAGAGGTGGGTTTGGGGGTGTGTGTGGTTCGAGCTGCAGGGTGACTGAAGATGTG GTTTCCTGTGCCCCCTTGCCTCCTGCCAGGTTCATTTGGCGGACTGTGACGCCACGTGACAAGC CCATGGGGCCCCTGCTGGTGGCCACATTCTCGATCTAGATAACTGATCATAATCAGCCATACCA CATTTGT

BioBank Japan Project Consortium

Koichi Matsuda. Department of Computational Biology and Medical Sciences, Graduate school of Frontier Sciences, The University of Tokyo, Tokyo, Japan.

Yuji Yamanashi. Division of Genetics, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Yoichi Furukawa. Division of Clinical Genome Research, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Takayuki Morisaki. Division of Molecular Pathology IMSUT Hospital, Department of Internal Medicine Project Division of Genomic Medicine and Disease Prevention The Institute of Medical Science The University of Tokyo, Tokyo, Japan.

Yoshinori Murakami. Department of Cancer Biology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Kaori Muto. Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Akiko Nagai. Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Wataru Obara. Department of Urology, Iwate Medical University, Iwate, Japan.

Ken Yamaji. Department of Internal Medicine and Rheumatology, Juntendo University Graduate School of Medicine, Tokyo, Japan.

Kazuhisa Takahashi. Department of Respiratory Medicine, Juntendo University Graduate School of Medicine, Tokyo, Japan.

Satoshi Asai. Division of Pharmacology, Department of Biomedical Science, Nihon University School of Medicine, Tokyo, Japan.
Division of Genomic Epidemiology and Clinical Trials, Clinical Trials Research Center, Nihon University. School of Medicine, Tokyo, Japan.

Yasuo Takahashi. Division of Genomic Epidemiology and Clinical Trials, Clinical Trials Research Center, Nihon University School of Medicine, Tokyo, Japan.

Takao Suzuki. Tokushukai Group, Tokyo, Japan.

Nobuaki Sinozaki. Tokushukai Group, Tokyo, Japan.

Hiroki Yamaguchi. Departmentof Hematology, Nippon Medical School, Tokyo, Japan.

Shiro Minami. Department of Bioregulation, Nippon Medical School, Kawasaki, Japan.

Shigeo Murayama. Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology, Tokyo, Japan.

Kozo Yoshimori. Fukujuji Hospital, Japan Anti-Tuberculosis Association, Tokyo, Japan.

Satoshi Nagayama. The Cancer Institute Hospital of the Japanese Foundation for Cancer Research, Tokyo, Japan.

Daisuke Obata. Center for Clinical Research and Advanced Medicine, Shiga University of Medical Science, Shiga, Japan.

Masahiko Higashiyama. Department of General Thoracic Surgery, Osaka International Cancer Institute, Osaka, Japan.

Akihide Masumoto. IIZUKA HOSPITAL, Fukuoka,Japan.

Yukihiro Koretsune. National Hospital Organization Osaka National Hospital, Osaka, Japan.