Analysis of MRI-derived spleen iron in the UK Biobank identifies genetic variation linked to iron homeostasis and hemolysis

Authors

Elena P. Sorokin, Nicolas Basty, Brandon Whitcher, ..., Robert L. Cohen, Madeleine Cule, E. Louise Thomas

Correspondence

sorokin@calicolabs.com **(E.P.S.)**,
l.thomas3@westminster.ac.uk **(E.L.T.)**

Sorokin et al., 2022, The American Journal of Human Genetics *109*, 1092–1104
June 2, 2022 © 2022 The Author(s).
https://doi.org/10.1016/j.ajhg.2022.04.013

CellPress

ARTICLE

Analysis of MRI-derived spleen iron in the UK Biobank identifies genetic variation linked to iron homeostasis and hemolysis

Elena P. Sorokin,[1,3,*] Nicolas Basty,[2,3] Brandon Whitcher,[2] Yi Liu,[1] Jimmy D. Bell,[2] Robert L. Cohen,[1] Madeleine Cule,[1,3] and E. Louise Thomas[2,3,*]

Summary

The spleen plays a key role in iron homeostasis. It is the largest filter of the blood and performs iron reuptake from old or damaged erythrocytes. Despite this role, spleen iron concentration has not been measured in a large, population-based cohort. In this study, we quantify spleen iron in 41,764 participants of the UK Biobank by using magnetic resonance imaging and provide a reference range for spleen iron in an unselected population. Through genome-wide association study, we identify associations between spleen iron and regulatory variation at two hereditary spherocytosis genes, *ANK1* and *SPTA1*. Spherocytosis-causing coding mutations in these genes are associated with lower reticulocyte volume and increased reticulocyte percentage, while these common alleles are associated with increased expression of *ANK1* and *SPTA1* in blood and with larger reticulocyte volume and reduced reticulocyte percentage. As genetic modifiers, these common alleles may explain mild spherocytosis phenotypes that have been observed clinically. Our genetic study also identifies a signal that co-localizes with a splicing quantitative trait locus for *MS4A7*, and we show this gene is abundantly expressed in the spleen and in macrophages. The combination of deep learning and efficient image processing enables non-invasive measurement of spleen iron and, in turn, characterization of genetic factors related to the lytic phase of the erythrocyte life cycle and iron reuptake in the spleen.

Introduction

In normal human physiology, iron is recycled much faster than new dietary iron is absorbed, and this iron economy is regulated by the spleen.[1–3] The spleen plays a critical role in removing senescent erythrocytes from the blood and does so via a population of splenic macrophages as well as by the action of the protein ferroportin, which transports iron back to the plasma.[4] Given the large iron flux through the normal spleen due to erythrocyte recycling, measurement of spleen iron has the potential to reflect activity of iron salvage pathways.

Spleen iron also may reflect erythrocyte biology and dysfunction. Hereditary spherocytosis (HS [MIM: 270970 and 182900]) is a relatively common hemolytic anemia, occurring at a prevalence of 1:1,000–2,500 in European populations.[5] Affected individuals have defects in membrane and cytoskeletal genes that contribute to erythrocyte membrane integrity and deformability, including *SPTA1* (encoding the filamentous protein alpha-spectrin [MIM: 182860]) and *ANK1* (encoding a protein that tethers spectrin filaments to erythrocyte membranes [MIM: 612641]).[6,7] The protein products of these genes interact in the formation of the mature erythrocyte cytoskeleton, a process accompanying cellular remodeling during reticulocytosis.[8] In HS, the resulting erythrocytes are spherically shaped and lose deformability and are ultimately trapped and ingested by macrophages within the red pulp of the spleen, resulting in an enlarged spleen and anemia despite brisk reticulocytosis. When measured directly with radiolabeled cells, erythrocyte turnover is dramatically accelerated in individuals with severe HS.[9] While laboratory measures such as sphered cell and reticulocyte volume can diagnose HS,[10] less is understood about the phenotypic and genetic heterogeneity of the disease.[11–13]

Although spleen iron has been investigated in specific disease groups,[14–20] few studies were performed in unascertained cohorts, partly because of limitations in detecting and quantifying low levels of tissue iron in non-overloaded populations.[21–23] The UK Biobank (UKBB) is a prospective study of half a million adults in the UK[24] that has genetic and phenotypic data including magnetic resonance imaging (MRI).[25] In this study, we applied computer vision techniques to quantify spleen iron non-invasively and at scale by repurposing the dedicated liver MRI acquisition. Spleen iron is only moderately correlated with other measures of iron stores in the body. Through genome-wide association study (GWAS), we characterized associations between spleen iron and genes involved in the human iron economy including common regulatory variation in the ferroportin gene, *SLC40A1*. Our GWAS identified a signal that co-localizes with a splicing quantitative trait locus for *MS4A7*, which we found to be abundantly expressed in the spleen and in macrophages. Regulatory variation in the genes encoding alpha-spectrin (*SPTA1*) and ankyrin (*ANK1*) were also linked to spleen iron and

[1]Calico Life Sciences LLC, South San Francisco, CA, USA; [2]Research Centre for Optimal Health, School of Life Sciences, University of Westminster, London, UK
[3]These authors contributed equally
*Correspondence: sorokin@calicolabs.com (E.P.S.), l.thomas3@westminster.ac.uk (E.L.T.)
https://doi.org/10.1016/j.ajhg.2022.04.013.
© 2022 The Author(s). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

increased mRNA expression of these genes and furthermore exhibited effects on reticulocyte and erythrocyte parameters opposite to the effects observed in HS-affected individuals. Non-invasive imaging coupled with genetic analysis enabled us to develop a quantitative trait that illuminated aspects of iron homeostasis relevant to hemolysis and iron recycling in the spleen.

Subjects and methods

Image analysis

We trained a convolutional neural network based on the U-net architecture[26] for 3D organ segmentation from the neck-to-knee Dixon MRI data by using 119 manual spleen annotations. The dice score on an out-of-sample held-out test set was 0.922. For the quantitative MRI data, we estimated the proton density fat fraction (PDFF) and R2* from the single-slice liver multi-echo data by using the PRESCO (phase regularized estimation using smoothing and constrained optimization) algorithm.[27] We converted R2* into iron concentration (mg/g) by using the widely used formula proposed by Wood et al.: iron concentration = $0.202 + 0.0254 \times R2^*$.[28,29] We extracted 2D masks from the 3D spleen segmentations at their intersection with the liver acquisition (Figure 1A).[30] We applied one-pixel erosion before computing the median value within that mask. We excluded 2D masks that had <1% of 3D volume or <20 voxels. Additional methods, including image acquisition details, are described in the supplemental information.

Epidemiological modeling of spleen iron risk factors

Associations of spleen iron with age, genetic sex, and self-reported ethnicity were performed in R v3.6.3. with the linear and logistic models for spleen iron as a quantitative or binary trait, respectively, after adjusting for covariates including imaging center, date, and time.

Phenome-wide association study

We generated a list of variables derived from raw data by using PHESANT[31] and removed procedural metrics (e.g., measurement date), duplicates, and raw measures, resulting in 1,824 traits (Table S2). We used PheWAS[32] to combine ICD10 codes (Field 41270) into distinct phenotype codes or phecodes (Table S3). In addition, we included 11 quantitative traits defined in our previous study.[33] We performed linear (quantitative traits) or logistic regression (binary traits) on spleen iron, adjusting for imaging center, date, scan time, age, sex, BMI, height, and ethnicity.

GWAS

We performed a GWAS in n = 35,324 participants as described[34] by using UKBB-imputed genotypes[24] version 3 for our GWAS, excluding single-nucleotide polymorphisms (SNPs) with minor allele frequency < 1% or imputation quality (INFO score) < 0.9. We included participants identified as White British by using UK Biobank Field 22006, which is based on self-reported ancestry and genetic ancestry based on principal-component analysis.[24] We excluded participants exhibiting sex chromosome aneuploidy, participants with a discrepancy between genetic and self-reported sex, heterozygosity and missingness outliers, and genotype call rate outliers.[24] 9,911,384 SNPs passed quality control (QC). We used BOLT-LMM[35] v2.3.2 to conduct the genetic association study. We included age at imaging, age^2, sex, imaging center, scan date and time, and genotyping batch as fixed-effect covariates and genetic relatedness derived from genotyped SNPs as a random effect to control for population structure and relatedness. We normalized the outcome variable by using inverse-rank normalization. In the genetic association study, we found no evidence for global inflation of test statistics ($\lambda_{gc} = 1.035$; linkage disequilibrium [LD] score regression intercept 1.027 [SE 0.0072]).

Replication analysis

We replicated blood trait findings by using the non-European meta-analyses from the Blood Cell Consortium Phase 2, which investigated the genetic basis of 15 blood cell traits in cohorts of diverse ancestries, including African ancestry (n = 15,171), East Asian (n = 151,807), Hispanic/Latino (n = 9,368), and South Asian participants (n = 8,180) in addition to European ancestry (n = 563,946).[36,37] To avoid overlap with UK Biobank Europeans, we focused on summary statistics from the non-European cohorts. We downloaded summary statistics for mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), and monocyte count from http://www.mhi-humangenetics.org/en/resources/ on November 14, 2021, and harmonized them by using dbSNP build 151 (GRCh37). We tested for replication of MCHC and MCV at rs4737010[A] (*ANK1*) and rs2479868[T] (*SPTA1*) and replication of monocyte count at rs950802[A] (*MS4A7*).

Conditional analysis and fine-mapping

We performed conditional analysis by using GCTA,[38] considering variants within 500 kb of an index variant. We constructed a reference LD panel of 5,000 randomly selected, unrelated European UKBB participants.[24] We excluded the major histocompatibility complex region because of the complexity of LD structure at this locus (GRCh37:6: 28,477,797–33,448,354; see https://www.ncbi.nlm.nih.gov/grc/human/regions/MHC). For each locus, we considered variants with locus-wide evidence of association ($p_{joint} < 10^{-6}$) to be conditionally independent. We followed an iterative procedure to determine credible sets of causal variants with 95% coverage.[34]

Colocalization studies

For gene expression studies, we used summary statistics from GTEx v8.[39] For disease and quantitative trait studies, we used UKBB summary statistics of phecodes,[40] normalized quantitative traits (http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank). We selected phenotypes with $p < 5 \times 10^{-8}$ within 500 kb of the index variant. We performed colocalization analysis by using coloc[41] with default priors and considered variants within 500 kb of the index variant. We considered two genetic signals to have strong evidence of colocalization if PP3 + PP4 ≥ 0.99 and PP4/PP3 ≥ 5.[42]

Heritability estimates

We estimated the heritability of each trait by using the restricted maximum likelihood method,[43] as implemented in BOLT-LMM.

Genetic correlation

We computed genetic correlation by using bivariate LD score regression (LDSC).[44] We computed the genetic correlation between spleen iron and 288 complex traits with a heritability of

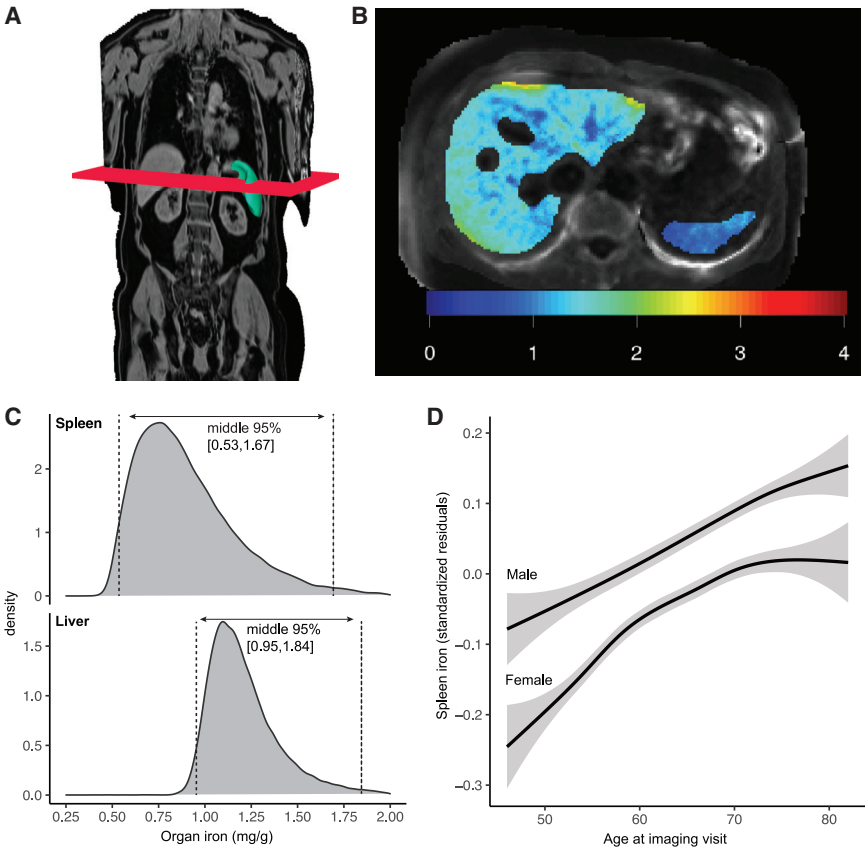

Figure 1. Image-based spleen iron quantification in the UK Biobank

(A) Example neck-to-knee abdominal MRI acquisition; the liver acquisition is shown as a red plane and deep-learning-derived spleen segmentation is shown in green.

(B) Axial view of iron concentration (mg/g) obtained through deep learning and image processing; the liver is shown at left and the spleen is shown at right.

(C) Distribution of spleen and liver iron. The middle 95% of the distribution is labeled.

(D) Spleen iron increases with age for both males and females. Standardized residuals are shown, adjusted for study center, scan date, and scan time.

at least 5% from the Neale Lab http://www.nealelab.is/uk-biobank/, plus organ iron measurements and volumes,[33] and blood iron biomarkers.[45] Following the recommendations of the developers, we (1) removed variants with imputation quality (INFO) < 0.9 because the INFO value is correlated with the LD score and could introduce bias, (2) excluded the major histocompatibility complex (MHC) region because of the complexity of LD structure at this locus (GRCh37:6: 28,477,797–33,448,354; see https://www.ncbi.nlm.nih.gov/grc/human/regions/MHC), and (3) restricted to HapMap3 SNPs.[46] We used a Bonferroni-corrected p value of 1.7e−4 as the significance threshold to identify traits with a significant genetic correlation.

Exome sequence quality control

We performed quality control of n = 200,643 whole exomes from the UKBB. Raw genotype calls were filtered genotype-level quality metrics to identify quality outliers for a given site and remove poor-quality individual-level genotypes. Similarly to the functionally equivalent (FE) pipeline,[47] we removed genotypes below a minimum read depth (for SNPs: 7 and for indels: 10) and genotypes below a minimum Phred-scaled genotype quality of 20. We removed genotypes where minor allele allelic balance < 0.15 for SNPs and 0.2 for indels. Supplementing FE filters with additional filters, we performed per-SNP QC, requiring the average genotype quality to be at least 30 and per-SNP depth of coverage to be at least 15, to filter out badly captured sites. Additionally, we removed variants with genotype missingness > 10% or that deviated meaningfully from Hardy-Weinberg equilibrium in a European ancestry cohort (HWE p < 1e−10). Of 17,981,897 total variants, 13,907,865 variants passed QC in the European exome cohort with MRI data (n = 18,240).

Exome variant annotation

We performed annotation by using VEPv100, LOFTEE,[48] CADD,[49] and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/, downloaded on September 27, 2020) with a custom pipeline to select variants meeting high-confidence loss-of-function criteria, filtered for rare variants (defined as cohort-specific minor allele frequency < 0.001). Of variants passing quality control, we subset 286,456 high-confidence, rare loss-of-function variants, 2,919,962 rare missense variants (CADD score ≥ 20), and 13,705 rare clinical pathogenic variants in 19,992 European ancestry samples for further analysis.

Rare variant association study

We performed rare variant burden and SKAT testing in SAIGE-GENE by using a mixed-effects model.[50] A kinship matrix was built in SAIGE from a filtered set of 354,878 genotyped variants (r2 < 0.2, minor allele frequency > 0.05, Hardy-Weinberg p value > 1e−10, excluding known regions of long-range LD). The linear mixed model regression equation was as follows:

$$y_i = \alpha + X\beta + G_i\beta + b_i + \varepsilon_i.$$

In the model, y_i is inverse-rank normalized spleen iron, and X represents fixed-effect covariates: age at imaging visit, age^2, chromosomally determined sex expressed as a binary indicator variable, study center, standardized scan date, standardized scan time, and the first five principal components of European genetic ancestry. G_i represents allele counts (0, 1, 2) for q variants in each gene to test. We then performed SKAT and burden tests in SAIGE-GENE and reported the p value from SKAT-O.[50] To inform directionality of effect, we reported the betas from the burden test. To avoid unstable results at low sample size, we calculated cumulative minor allele count and thresholded at ≥5 minor alleles per gene, including singletons and doubletons.

Genetic identification of hereditary spherocytosis alleles

From the exome cohort filtered for European ancestry (n = 167,246), we annotated variants by clinical assertion as pathogenic according to ClinVar, downloaded from https://ftp.ncbi.nlm.nih.gov/pub/clinvar/ on September 27, 2020, and called

Table 1. Summary of the UKBB cohort at recruitment, image acquisition, and quantification of liver iron and spleen iron

	UK Biobank cohort (at time of baseline visit)	Imaging cohort (at time of imaging visit)	Identified in combined liver slices	
			Liver	Spleen
N	502,520	44,265	44,265	41,764
% female	54.4 [54.2–54.5]	51.8 [51.2–52.1]	51.8 [51.3–52.3]	52.1 [51.6–52.6]
Age	56.5 (8.1)	64.2 (7.73)	63.7 (7.56)	64.1 (7.71)
BMI (kg/m^2)	27.4 (4.8)	26.5 (4.37)	26.5 (4.39)	26.4 (4.32)
Height (cm)	168 (9.28)	169 (9.27)	169 (9.29)	169 (9.26)
% Caucasian	81.4 [81.3–81.5]	84.8 [84.4–85.1]	85.2 [84.4–85.5]	84.7 [84.4–85.0]
Iron concentration (mg/g)			1.24 (0.29) [0.203–6.96]	0.915 (0.318) [0.172–6.68]

Spleen iron measurements (mg/g) are provided as mean (standard deviation) and range. For other quantitative values, the mean (standard deviation) is given. For binary values, mean [95% confidence interval] is given.

predicted loss-of-function alleles by using LOFTEE.[48] HS alleles were defined as clinical pathogenic variants or high-confidence putative loss-of-function alleles in one of six genes (*SPTA1*, *SPTB*, *SLC4A1*, *ANK1*, *EPB41*, or *EPB42*).

Ethics statement

The UK Biobank has approval from the North West Multi-centre Research Ethics Committee (MREC) to obtain and disseminate data and samples from the participants (http://www.ukbiobank.ac.uk/ethics/), and these ethical regulations cover the work in this study. Written informed consent was obtained from all participants.

Results

Characterization of spleen iron in a large, population-based cohort

We quantified spleen iron concentration (spleen iron hereafter) in 41,764 UKBB participants with both the 3D neck-to-knee and the quantitative liver single-slice MRI sequences available. We opportunistically measured spleen iron concentration by segmenting the spleen from the neck-to-knee image[33] (Figure 1A) and subsequently extracting a 2D mask where the spleen volume intersects with the quantitative 2D liver slice (Figure 1B).[30] The average spleen iron was 0.92 ± 0.32 mg/g, significantly lower than liver iron of 1.24 ± 0.29 mg/g (Table 1, Figure 1C; paired t test $p < 2.2e{-}16$). While there is no accepted normal range of spleen iron, 1.98 mg/g has been suggested as an upper cut-off, and 2.74 mg/g is reported to be pathological.[51] Using 1.98 mg/g as the threshold, 1.04% (n = 435) of this cohort had elevated spleen iron, while 0.32% (n = 137) had spleen iron above the 2.74 mg/g threshold. 95% of the population fell into the range of 0.54 to 1.69 mg/g, and we propose this as a possible reference range in an unselected population (Table S2).

Spleen iron differed by age and sex. Men had higher spleen iron than women (men: 0.96 ± 0.34 mg/g, women: 0.87 ± 0.29 mg/g) ($p = 6.3e{-}219$; Tables S2 and S3). Increasing spleen iron was associated with age (0.0044 mg/g/year or 0.012 SD/year) (Figure 1D). In women, menopause was associated with 0.12 mg/g higher spleen iron [95% CI 0.08–0.16].

In a phenome-wide association study with over 3,200 quantitative traits and disease outcomes, spleen iron was correlated with erythrocyte parameters: reticulocyte percentage (beta = 0.091; $p = 3.7e{-}66$), reticulocyte count (beta = 0.087; $p = 6.6e{-}64$), and high light scatter reticulocyte count (beta = 0.089; $p = 5.8e{-}64$). Spleen iron was also associated with lifestyle factors, including consumption of lamb (beta = 0.150; $p = 2.1e{-}44$) and beef (beta = 0.143; $p = 1.3e{-}43$), and negatively associated with alcohol consumption (beta = −0.087; $p = 1.7e{-}18$). Spleen iron correlated with liver iron (beta = 0.409, $p < 1e{-}300$), brain iron content, specifically with T2* (inversely proportional to iron) in the caudate (beta = −0.063; $p = 4.3e{-}16$) and putamen (beta = −0.061, $p = 4.1e{-}15$).[52] Spleen iron was associated with myeloid leukemia (beta = 0.386; $p = 9.6e{-}10$), chronic dermatitis (beta = 0.328; $p = 4.1e{-}07$), hypokalemia (beta = 0.286, $p = 5.8e{-}6$), and glaucoma (beta = 0.186, $p = 1.0e{-}6$) (Figure 2, Tables S4 and S5). Spleen iron was negatively correlated with iron-deficiency anemia, but this did not achieve Bonferroni significance (beta = −0.134, p = 0.002), although this diagnosis may not have been fully captured by medical billing codes (n = 580 affected individuals, n = 35,316 control individuals).

GWAS of spleen iron identifies DNA polymorphisms linked to global iron homeostasis

In a common variant GWAS of spleen iron, seven loci reached genome-wide significance ($p < 5e{-}8$; Figure 3, Table 2). Conditional analysis yielded no secondary signals. We estimated the narrow-sense heritability of spleen iron to be 16.7% (SE 1.64%). Spleen iron was moderately genetically correlated with ferritin ($r_g = 0.56$) and MCHC ($r_g = 0.42$), but the genetic correlations with other iron measures, including liver iron, were not significant (Tables S6 and S7).

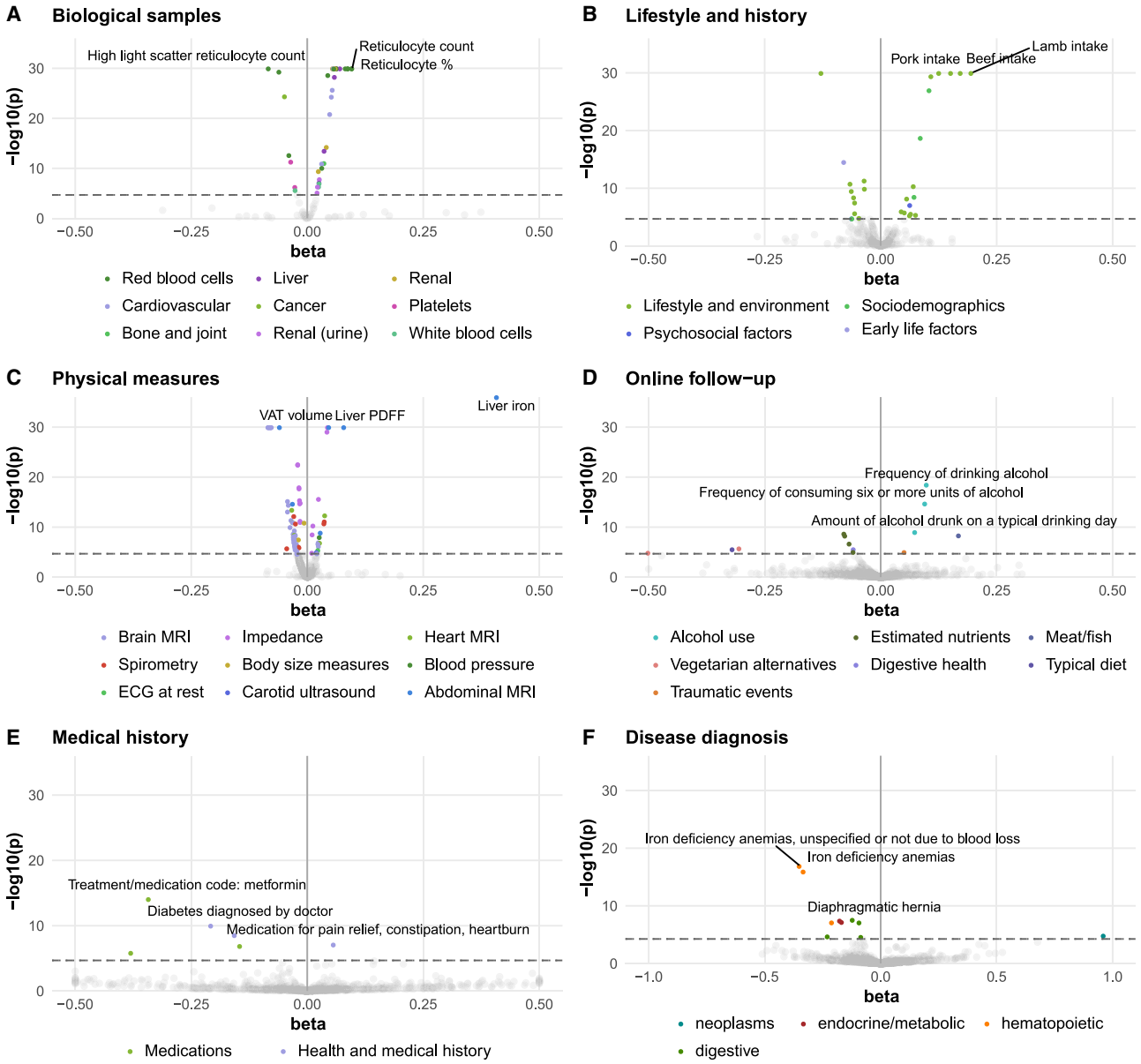

Figure 2. Phenome-wide association between spleen iron and complex traits in the UK Biobank
(A–F) Traits are organized by category: (A) biological samples, (B) lifestyle and history, (C) physical measures, (D) online follow up, (E) medical history, and (F) disease diagnosis. Bonferroni significance threshold is shown by horizontal dashed line. The top three associations in each category are annotated. ML, myeloid leukemia; HLS reticulocyte, high light scatter reticulocyte.

We observed a signal on chromosome 2 at *SLC40A1*, which encodes the iron transporter ferroportin (lead variant: rs13008848[G], beta = −0.057, p = 2.9e−10, Figure 4A). To test whether the spleen signal was shared across other body iron traits, we re-analyzed previous genetic studies of serum ferritin, serum iron, and liver iron[33,45] and observed evidence of co-localization at *SLC40A1* (posterior probability ≥ 0.99) (Table S8). Since the lead variant lies upstream of the *SLC40A1* open reading frame, we tested for a shared effect between spleen iron and *SLC40A1* expression. Across 54 human tissues of the GTEx Consortium,[53] we observed evidence for regional co-localization with a quantitative trait locus for *SLC40A1* mRNA in many tissues, including whole blood (posterior probability ≥ 0.99) (Figure 4D; Table S9). As expected, spleen iron was associated with *SLC40A1*, and furthermore, we showed that this locus most likely influences tissue iron levels through *SLC40A1* mRNA abundance.

The GWAS of spleen iron identified other loci relevant to iron homeostasis, including a locus on chromosome 9 (Figure S5B). The lead SNP rs41276777[A] (beta = 0.17, p = 2.1e−9) occurred in the 5′ untranslated region of *PRPF4* and *CDC26* (Figure S5A) and was associated with a regulatory locus affecting expression of both *PRPF4* and *CDC26* mRNA in whole blood and the spleen (Table S9), suggesting bi-directional regulation of gene expression. This region colocalized with signals for serum ferritin, serum iron, and other blood iron traits (Table S8),

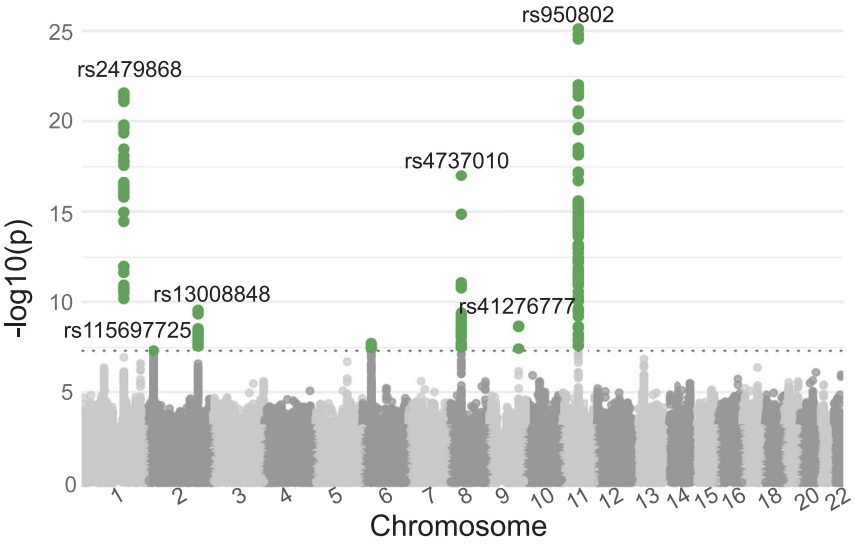

Figure 3. GWAS of spleen iron
Genome-wide significant signals are annotated with the variant at each locus. Dotted horizontal line marks genome-wide significance (p = 5e−8).

suggesting roles not only in spleen iron but body iron more broadly.

Elevated spleen iron colocalizes with a splicing quantitative trait locus for *MS4A7* and monocyte traits

Our GWAS with spleen iron found a significant association at the *MS4A7* locus (lead variant: rs950802[A], beta = 0.086, p = 7.7e−26; Table 2, Figure 5A). rs950802[A] causes a synonymous mutation at Leu57 in the third exon of *MS4A7* and is also a variant in the first intron of *MS4A14*. This signal colocalized with a signal of serum ferritin[45] (posterior probability ≥ 0.99). Using splicing quantitative trait data,[53] we identified an alternative splicing event in *MS4A7*, which colocalized with spleen iron (Figure 5B). Based on an analysis of the open reading frame of *MS4A7,* this alternative splicing event was predicted to increase skipping of the second exon and therefore interrupt a conserved CD20-like transmembrane domain in *MS4A7* (Figure 5C). We thus identified an association between *MS4A7* and spleen iron and found a plausible molecular mechanism by which this variant disrupts *MS4A7* function.

To explore the functional consequences of regulatory variation in *MS4A7*, we examined its association with hematological parameters in the UKBB and observed a significant association with monocyte count and percentage as well as platelet count and crit (Figure 5D). Using meta-analysis in non-European populations from the Blood Cell Consortium, a large study of blood cell traits in diverse cohorts,[36,37] we replicated associations and directions of effect between this locus and monocyte count in an East Asian cohort (n = 151,807) at Bonferroni significance (p < 0.017) and observed the same direction of effect in African ancestry (n = 15,171) and Hispanic (n = 9,367) cohorts (Table S11). Further, we found enrichment of *MS4A7* mRNA in monocytes compared to sixteen other hematopoietic cell types[55] (Figure 5E) and also higher expression of *MS4A7* mRNA in the spleen relative to other tissues[53] (Figure 5F). A common allele at *MS4A7* was thus associated with elevated spleen iron and monocyte traits. *MS4A7* mRNA was enriched in the spleen and in monocytes, suggesting a role for this gene in iron recycling in the spleen.

GWAS of spleen iron identifies common alleles in *SPTA1* and *ANK1* linked to increased gene expression and erythrocyte function

In addition to signals linked to iron homeostasis, the GWAS of spleen iron revealed associations in *SPTA1* and *ANK1*, encoding structural components of erythrocytes (Table 2, Figure S6). The lead variant rs2479868[T] (beta = −0.083, p = 2.5e−22) on chromosome 1 was located in the 3′ untranslated region of *SPTA1*, and the lead SNP rs4737010[A] (beta = −0.077, p = 1.0e−17) on chromosome 8 was the first intron of *ANK1* (Figures 4B and 4C). We tested for a shared signal with mRNA expression levels of *SPTA1* and *ANK1* and observed co-localization with *cis*-regulatory variation for *ANK1* in multiple tissues (Figures 4E and 4F).

Variation at each locus was associated with increased mRNA expression and mean reticulocyte volume and decreased MCHC, reticulocyte percentage, and spleen volume (Figure 6). The magnitude and directionality for the changes in red cell parameters associated with the lead SNPs at *SPTA1* and *ANK1* were similar, but a regression model testing for interaction between these two loci found that the effects are independent (beta = 0.016, p = 0.28). These signals did not co-localize with iron traits in other tissues such as serum or liver, suggesting specific effects to the spleen (Table S7).

Using non-European meta-analyses from the Blood Cell Consortium as independent cohorts,[36,37] we replicated the associations and directions of effect between *SPTA1* and *ANK1* lead variants and both MCHC and mean corpuscular volume in an East Asian cohort at Bonferroni significance (p < 0.0042). In a cohort of African ancestry and in a cohort of Hispanic ancestry, we replicated directions of effect (Table S11).

Common alleles in *SPTA1* and *ANK1* show effects on erythrocytes opposite to effects observed with rare deleterious alleles and in hereditary spherocytosis

Given the known associations of both *SPTA1* and *ANK1* gene defects with hereditary spherocytosis (HS), we analyzed rare variation (minor allele frequency < 0.001) predicted to cause loss of function. Starting with 167,246 exomes of European ancestry, we conducted rare variant association studies for

Table 2. Fine-mapped lead SNPs from GWAS of spleen iron

Lead SNP	Locus	Lead SNP consequence	Effect allele	Other allele	Beta	p	Minor allele frequency
rs950802	*MS4A7/MS4A14*	synonymous variant	A	G	0.086	7.7e−26	0.307
rs2479868	*SPTA1*	intron variant	T	C	−0.083	2.5e−22	0.266
rs4737010	*ANK1*	intron variant	A	G	−0.077	1.0e−17	0.228
rs13008848	*SLC40A1*	5' UTR variant	G	C	−0.057	2.9e−10	0.226
rs41276777	*PRPF4*	5' UTR variant	A	G	0.175	2.1e−09	0.019
rs115697725	*KLHL29*	intron variant	G	C	0.059	4.9e−08	0.146

Genome-wide significant associations ($p < 5e{-}8$) are shown by locus and lead SNP after fine-mapping. A seventh association at the major histocompatibility locus (MHC) could not be fine-mapped.

reticulocyte percentage and volume and identified one significant gene, *SPTA1*, in both studies ($p_{SKAT\text{-}O} < 1e{-}24$; Figure S7). Performing a scan of the same loss-of-function rare variation in *SPTA1* across a hematology panel, we recapitulated clinical hallmarks of spherocytosis: *SPTA1* loss of function was significantly associated with increased reticulocyte percentage and increased MCHC, increased bilirubin, decreased mean spherical cell volume, and decreased mean reticulocyte volume (Figure S7). We did not observe a significant association with spleen iron in the imaging subcohort for either putative loss of function or deleterious missense variation, perhaps as a result of reduced statistical power of the imaging subcohort (n = 18,420) compared to the exome cohort (n = 167,246) (Figure S8).

Finally, we genetically identified HS in the UKBB exome cohort by using clinical assertions of pathogenicity and predicted high confidence loss of function in one of six HS genes and estimated prevalence to be 1:389 [95% CI 1:427–1:354]. We asked whether the common alleles found via GWAS could modify the effects of rare deleterious alleles for hematology parameters relevant for HS, including mean reticulocyte volume. We estimated that carrying either of the *SPTA1* or *ANK1* lead GWAS SNPs skewed erythrocyte parameters toward beneficial effects in deleterious allele carriers and noncarriers, suggesting that the common GWAS alleles could modify the effects of the rare deleterious alleles (Figure S10).

Discussion

Most studies of spleen iron have not included healthy volunteers or methods that reliably assess organ iron[15,56–58] or only provided qualitative histological grading[58,54] (Table S11).[14,17–19] In this study, we combined deep-learning algorithms and efficient image processing to quantify spleen iron in 41,764 participants of the UKBB. This enabled us to estimate a reference range for spleen iron in a large, unselected population.

Spleen iron was higher in men, and our analysis of pre- and post-menopausal women is consistent with prior reports of iron stores in other tissues.[59] Spleen iron increased with age, is associated with red meat intake, and is inversely associated with alcohol consumption, extending prior observations in the liver.[28,60,61] Further, spleen iron was only modestly associated with measures of iron in other tissues. Un-like other iron-rich organs such as the liver, we discovered that spleen iron was associated with indicators of reticulocytosis and reticulocyte turnover.

Our GWAS of spleen iron identifies regulatory loci in *SPTA1* and *ANK1*, which when combined with other evidence, suggests a model of low splenic turnover due to relatively large, long-lived erythrocytes. These alleles are associated with decreased spleen iron and increased mRNA expression of their respective *cis* gene, as well as larger reticulocyte volume and reduced measures of reticulocytosis (Figure 6). Since the spleen is the major route of erythrocyte clearance and iron salvage, lower levels of reticulocytosis would be expected to result in lower spleen iron at steady state. Neither variant is associated with anemia, suggesting that they are not pathogenic. This reduced turnover model also predicts that iron levels in other organs not involved in erythrocyte clearance (such as the pancreas) would be unaffected by these two loci, consistent with the data.

As genetic modifiers, these alleles in *SPTA1* and *ANK1*, which segregate frequently across global populations, may explain the variable penetrance and expressivity observed in HS. In a recent study of affected individuals with identified HS mutations, 64% involved *SPTA1* or *ANK1*, and the investigators observed multiple HS families with broad phenotypic variability, including a compelling example of dizygotic twins sharing the same pathogenic *ANK1* mutation presenting with mild disease in one affected individual and severe disease requiring splenectomy in the other.[11] The variation could not be explained, leading to speculation that yet-unknown genetic factors may be contributing. We were able to identify the hallmarks of HS through rare loss-of-function variation in *SPTA1*, even within a cohort not ascertained for hemolytic anemias. Our analysis suggested that common genetic variants may modify the effects of rare deleterious alleles. The expression-increasing variants identified here may help to explain the heterogeneity that has been observed in HS patients.

In addition to genes specific to red cell turnover in the spleen, our genetic study of spleen iron also pointed to regulators of the human body's iron economy. We characterized a spleen iron signal at *MS4A7*, which belongs to the CD20

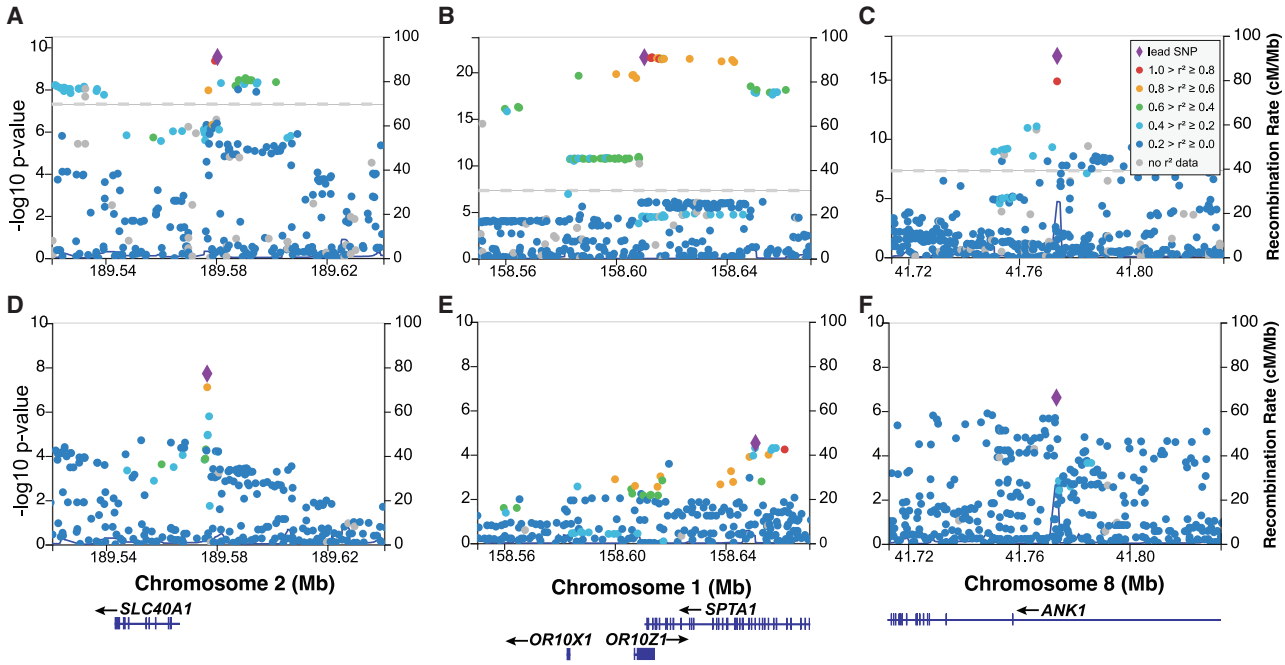

Figure 4. Common variants in *SLC40A1*, *SPTA1*, and *ANK1* loci are associated with spleen iron and colocalize with *cis*-regulatory variation

(A) Fine-mapped locus near *SLC40A1* is associated with spleen iron (lead SNP: rs13008848).
(B) Fine-mapped locus near *SPTA1* is associated with spleen iron (lead SNP: rs2479868).
(C) Fine-mapped locus near *ANK1* is associated with spleen iron (lead SNP: 4373010).
(D) Co-localization of *cis*-regulatory variation in the *SLC40A1* locus and spleen iron. Co-localization at this locus was observed in multiple tissues (posterior probability $\geq$ 0.99 for each). Blood *cis*-eQTL data are shown.
(E) *Cis*-regulatory variation in the *SPTA1* locus co-localizes with spleen iron. A signal was observed in multiple tissues but did not meet significance threshold in whole blood (posterior probability = 0.72).
(F) *Cis*-regulatory variation at the *ANK1* locus co-localizes with spleen iron. Co-localization was observed in multiple tissues (posterior probability $\geq$ 0.99). Gray dashed line indicates genome-wide significance ($p = 5e{-}8$). For all eQTL signals, a threshold of FDR < 5% was used in GTEx.[54] Linkage disequilibrium was calculated with 1000 Genomes Phase 3. Gene models are shown at bottom in GRCh38 coordinates.

family of membrane proteins, which are expressed within the hematopoietic lineage and largely uncharacterized.[62,63] Here, we linked a splicing quantitative trait locus in *MS4A7* to spleen iron and macrophage abundance and even found that this gene was enriched in the spleen and in macrophages. It is possible that excess circulating monocytes can be recruited to provide an expanded reservoir in the spleen, contributing to the body's iron economy in addition to splenic macrophages.[64] Our findings regarding *MS4A7* potentially illuminate additional details of splenic erythrocyte clearance mechanisms. For example, the specific functional roles played by macrophages and monocytes in the splenic red pulp remain incompletely understood.[9,64]

This study has limitations. First, as the only study to quantify spleen iron in a large cohort, no replication cohort is available, though we are able to replicate all our findings on relevant blood cell traits in independent cohorts.[36,37] Second, while we found limited evidence in the UK Biobank that spleen iron varies by ethnicity (Figures S3 and S4), additional imaging studies are warranted to quantify spleen iron across populations. Third, experimental studies in the hematopoietic lineage of a model system will be needed to test the functional consequences of the regulatory variation observed here. Fourth, while we genetically identified HS in the UKBB and showed that the common alleles in *SPTA1* and *ANK1* act as genetic modifiers of erythrocyte parameters, additional clinical validation is needed to substantiate these findings in an HS cohort.

In summary, we have quantified spleen iron by repurposing liver MRI acquisitions, maximizing use of the data at no extra cost to scanning or participant time. Our findings suggest that steady-state levels of spleen iron are sensitive to alterations in erythrocyte structure affecting cell turnover as well as alterations in iron transport by macrophages. We identified common regulatory variation in HS genes showing effects on erythrocytes that are opposite of the effects observed in HS-affected individuals. Imaging-derived spleen iron is thus a quantitative biomarker reflecting hemolysis and iron reuptake by the spleen, is tractable for genetic analysis, and has potential to contribute to future clinical characterization of hemolytic anemias.

Data and code availability

The derived datasets generated during this study are available from the UK Biobank (https://www.ukbiobank.ac.uk). Summary

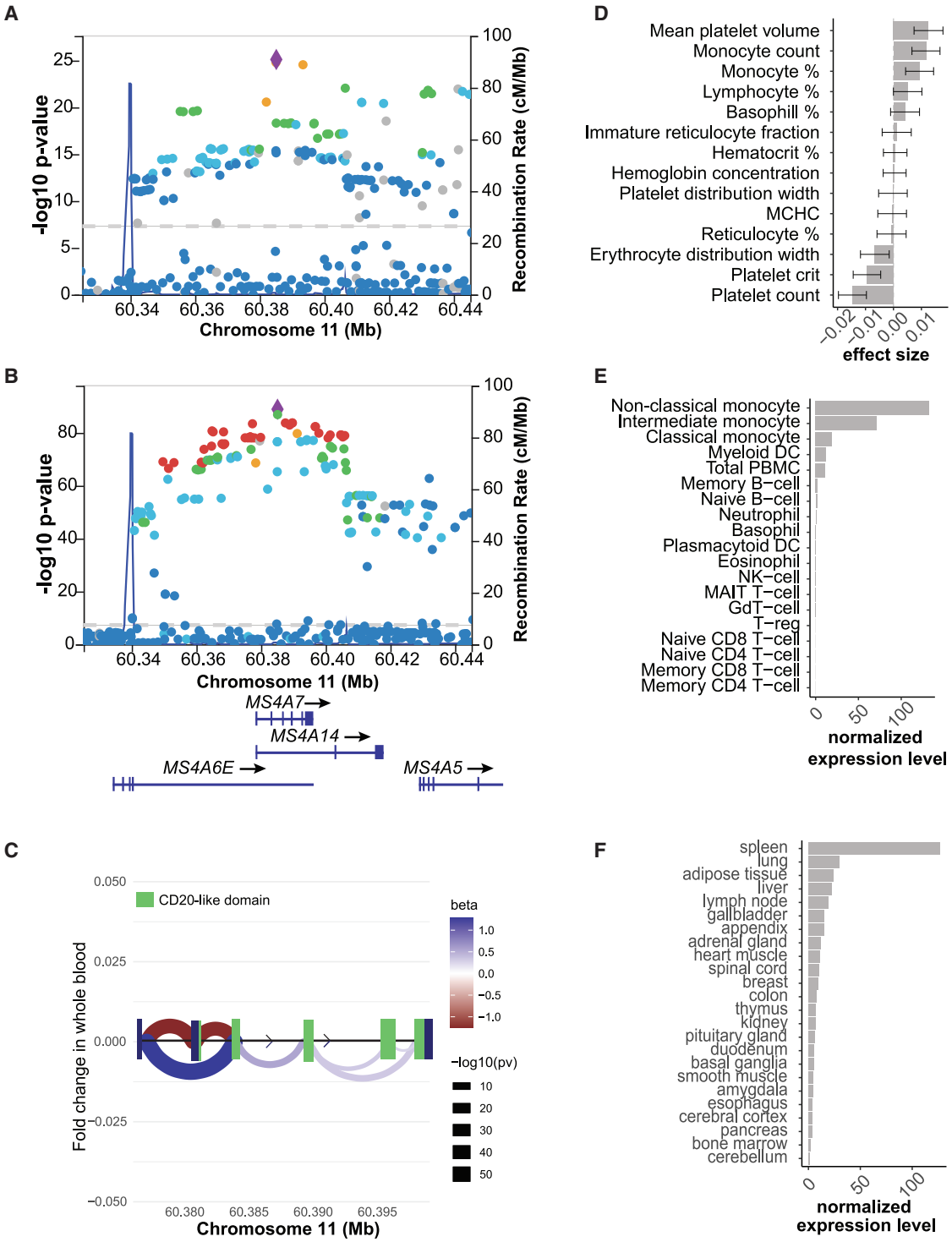

Figure 5. Characterization of the MS4A7 locus

(A) The MS4A7 locus on chromosome 11 is associated with an increase in spleen iron. Purple triangle displays the lead SNP rs950802[G].

(B) This locus co-localizes with splicing quantitative trait locus of *MS4A7* in whole blood (posterior probability ≥ 0.99).

(C) This splicing quantitative trait locus is associated with exon skipping in the *MS4A7* locus. Corresponding betas and p value are shown. The *MS4A7* gene model is displayed with the conserved CD20-like domain shown in green.

(D) Genetic associations with a panel of blood cell traits show associations with monocyte count, monocyte percentage, and mean platelet volume and negative associations with platelet count and platelet crit.

(E) Expression of *MS4A7* was enriched in monocytes, including non-classical, intermediate, and classical forms.

(F) Expression of *MS4A7* in human tissues displayed enrichment for lymphoid tissue, notably spleen and lymph nodes. DC, dendritic cell; MCHC, mean corpuscular hemoglobin concentration.

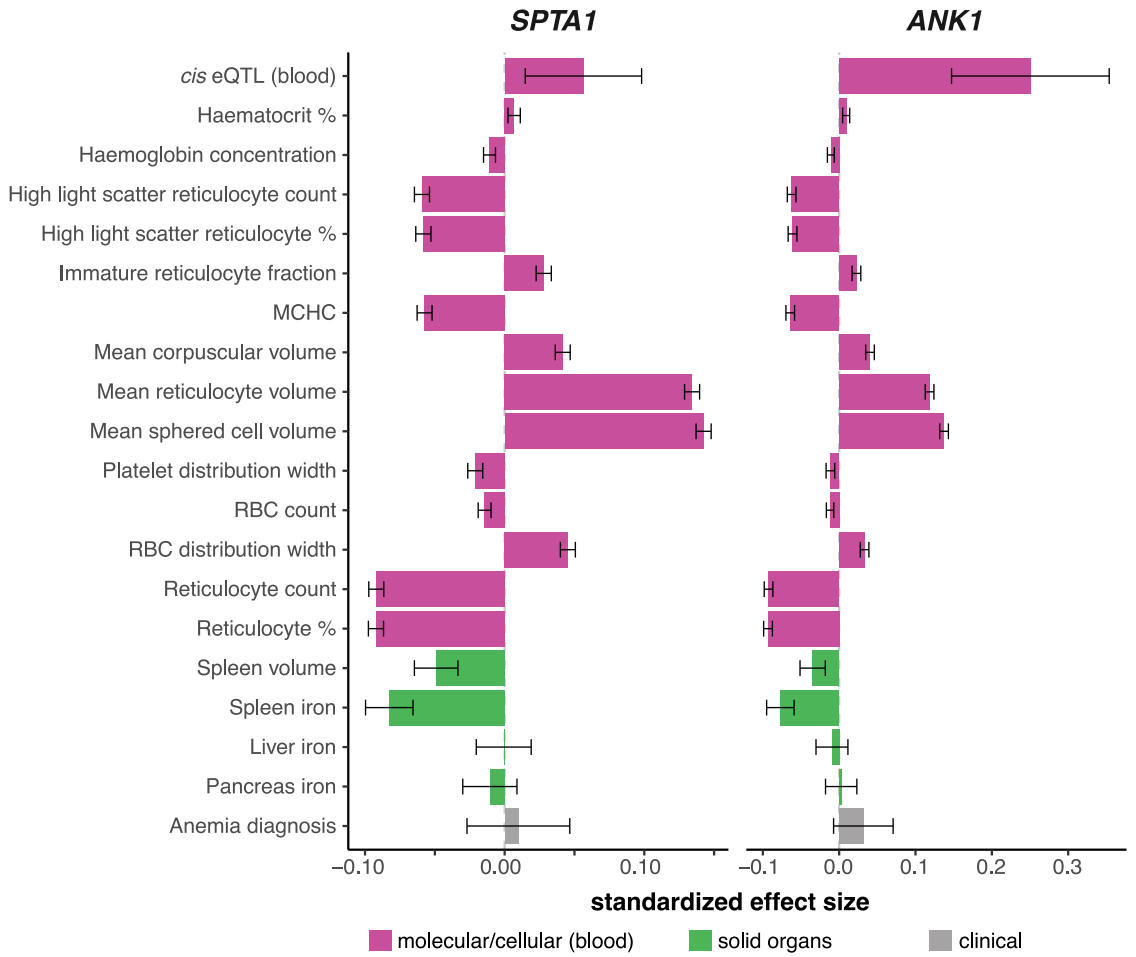

Figure 6. Associations between molecular, cellular, organ, and tissue traits and lead SNPs
Associations of molecular, cellular, organ, and tissue traits with lead SNPs at *SPTA1* and *ANK1* loci indicate protective effects on red blood cell function. Both *SPTA1* and *ANK1* loci were associated with increased mRNA expression, beneficial effects on red blood cell parameters, decreases in spleen iron and spleen volume, and no effects on other measures of body iron or anemia diagnosis. Standardized effect sizes are shown. Error bars represent 95% confidence intervals. Lead SNP at *SPTA1* locus: rs2479868; lead SNP at *ANK1* locus: rs4737010.

statistics are available from the GWAS catalog (https://www.ebi.ac.uk/gwas) at accession number GCST90101831. The code generated during this study is publicly available at https://github.com/calico/ukbb-mri-sseg and www.github.com/recoh/pipeline.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.04.013.

Acknowledgments

The authors would like to thank David Botstein, Marcia Paddock, Amoolya Singh, Kevin Wright, and Bert van der Zwaag for helpful feedback and discussion. This work was made possible by the UK Biobank, including staff, funders, and study volunteers. This research has been conducted with the UKBB Resource under application number 44584 and was funded by Calico Life Sciences LLC.

Author contributions

N.B., M.C., E.P.S., J.D.B., and E.L.T. designed the study. N.B., B.W., and Y.L. implemented the image processing methods. M.C. performed the data processing. E.P.S. and M.C. performed genetic analysis. E.L.T., E.P.S., R.L.C., M.C., and N.B. drafted the manuscript. All authors edited, read, and approved the manuscript.

Declaration of interests

E.P.S., R.L.C., and M.C. are employees of Calico Life Sciences LLC. Y.L. is a former employee of Calico Life Sciences LLC.

Received: January 19, 2022
Accepted: April 19, 2022
Published: May 13, 2022

Web resources

OMIM, https://www.omim.org

References

1. Cesta, M.F. (2006). Normal structure, function, and histology of the spleen. Toxicol. Pathol. *34*, 455–465. https://doi.org/10.1080/01926230600867743.

2. Dzierzak, E., and Philipsen, S. (2013). Erythropoiesis: development and differentiation. Cold Spring Harb. Perspect. Med. *3*, a011601. https://doi.org/10.1101/cshperspect.a011601.
3. Mebius, R.E., and Kraal, G. (2005). Structure and function of the spleen. Nat. Rev. Immunol. *5*, 606–616. https://doi.org/10.1038/nri1669.
4. Ganz, T. (2013). Systemic iron homeostasis. Physiol. Rev. *93*, 1721–1741. https://doi.org/10.1152/physrev.00008.2013.
5. Gallagher, P.G. (2005). Red cell membrane disorders. Hematol. Am. Soc. Hematol. Educ. Program, 13–18. https://doi.org/10.1182/asheducation-2005.1.13.
6. Delaunay, J. (2002). Molecular basis of red cell membrane disorders. Acta Haematol. *108*, 210–218. https://doi.org/10.1159/000065657.
7. Bennett, V., and Baines, A.J. (2001). Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. Physiol. Rev. *81*, 1353–1392. https://doi.org/10.1152/physrev.2001.81.3.1353.
8. Ovchynnikova, E., Aglialoro, F., von Lindern, M., and van den Akker, E. (2018). The shape shifting story of reticulocyte maturation. Front. Physiol. *9*, 829. https://doi.org/10.3389/fphys.2018.00829.
9. Klei, T., Van Bruggen, R., Dalimot, J., Veldthuis, M., Mul, E., Rademakers, T., Hoogenboezem, M., Nagelkerke, S.Q., Van Ijcken, W., Moestrup, S.K., et al. (2019). Hemolysis in the spleen drives erythrocyte turnover. Blood *134*, 946. https://doi.org/10.1182/blood-2019-124342.
10. Arora, R.D., Dass, J., Maydeo, S., Arya, V., Kotwal, J., and Bhargava, M. (2018). Utility of mean sphered cell volume and mean reticulocyte volume for the diagnosis of hereditary spherocytosis. Hematology *23*, 413–416. https://doi.org/10.1080/10245332.2018.1423879.
11. van Vuren, A., van der Zwaag, B., Huisjes, R., Lak, N., Bierings, M., Gerritsen, E., van Beers, E., Bartels, M., and van Wijk, R. (2019). The complexity of genotype-phenotype correlations in hereditary spherocytosis: a cohort of 95 patients. HemaSphere *3*, e276. https://doi.org/10.1097/hs9.0000000000000276.
12. Chonat, S., Risinger, M., Sakthivel, H., Niss, O., Rothman, J.A., Hsieh, L., Chou, S.T., Kwiatkowski, J.L., Khandros, E., Gorman, M.F., et al. (2019). The spectrum of SPTA1-associated hereditary spherocytosis. Front. Physiol. *10*, 815. https://doi.org/10.3389/fphys.2019.00815.
13. Tole, S., Dhir, P., Pugi, J., Drury, L.J., Butchart, S., Fantauzzi, M., Langer, J.C., Baker, J.M., Blanchette, V.S., Kirby-Allen, M., and Carcao, M.D. (2020). Genotype–phenotype correlation in children with hereditary spherocytosis. Br. J. Haematol. *191*, 486–496. https://doi.org/10.1111/bjh.16750.
14. Kolnagou, A., Michaelides, Y., Kontoghiorghe, C.N., and Kontoghiorghes, G.J. (2013). The importance of spleen, spleen iron, and splenectomy for determining total body iron load, ferrikinetics, and iron toxicity in thalassemia major patients. Toxicol. Mech. Methods *23*, 34–41. https://doi.org/10.3109/15376516.2012.735278.
15. Brewer, C.J., Coates, T.D., and Wood, J.C. (2009). Spleen R2 and R2* in iron-overloaded patients with sickle cell disease and thalassemia major. J. Magn. Reson. Imaging *29*, 357–364. https://doi.org/10.1002/jmri.21666.
16. Papakonstantinou, O., Alexopoulou, E., Economopoulos, N., Benekos, O., Kattamis, A., Kostaridou, S., Ladis, V., Efstathopoulos, E., Gouliamos, A., and Kelekis, N.L. (2009). Assessment of iron distribution between liver, spleen, pancreas, bone marrow, and myocardium by means of R2 relaxometry with MRI in patients with beta-thalassemia major. J. Magn. Reson. Imaging *29*, 853–859. https://doi.org/10.1002/jmri.21707.
17. Aslan, E., Luo, J.W., Lesage, A., Paquin, P., Cerny, M., Chin, A.S.-L., Olivié, D., Gilbert, G., Soulières, D., and Tang, A. (2021). MRI-based R2* mapping in patients with suspected or known iron overload. Abdom. Radiol. *46*, 2505–2515. https://doi.org/10.1007/s00261-020-02912-w.
18. Çetinçakmak, M.G., Hattapoğlu, S., Söker, M., Ekici, F., Yilmaz, K., Göya, C., and Hamidi, C. (2020). Evaluation of the relationship between splenic iron overload and liver, heart and muscle features evident on T2*-weighted magnetic resonance imaging. Adv. Clin. Exp. Med. *29*, 475–480. https://doi.org/10.17219/acem/116758.
19. França, M., Martí-Bonmatí, L., Porto, G., Silva, S., Guimarães, S., Alberich-Bayarri, Á., Vizcaíno, J.R., and Pessegueiro Miranda, H. (2018). Tissue iron quantification in chronic liver diseases using MRI shows a relationship between iron accumulation in liver, spleen, and bone marrow. Clin. Radiol. *73*, 215.e1–215.e9. https://doi.org/10.1016/j.crad.2017.07.022.
20. Adler, D.D., Glazer, G.M., and Aisen, A.M. (1986). MRI of the spleen: normal appearance and findings in sickle-cell anemia. Am. J. Roentgenol. *147*, 843–845. https://doi.org/10.2214/ajr.147.4.843.
21. Angelucci, E., Giovagnoni, A., Valeri, G., Paci, E., Ripalti, M., Muretto, P., McLaren, C., Brittenham, G.M., and Lucarelli, G. (1997). Limitations of magnetic resonance imaging in measurement of hepatic iron. Blood *90*, 4736–4742. https://doi.org/10.1182/blood.v90.12.4736.4736_4736_4742.
22. Saito, H. (2014). Metabolism OF iron stores. Nagoya J. Med. Sci. *76*, 235–254.
23. Schwenzer, N.F., Machann, J., Haap, M.M., Martirosian, P., Schraml, C., Liebig, G., Stefan, N., Häring, H.-U., Claussen, C.D., Fritsche, A., and Schick, F. (2008). T2* relaxometry in liver, pancreas, and spleen in a healthy cohort of one hundred twenty-nine subjects-correlation with age, gender, and serum ferritin. Invest. Radiol. *43*, 854–860. https://doi.org/10.1097/rli.0b013e3181862413.
24. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209. https://doi.org/10.1038/s41586-018-0579-z.
25. Littlejohns, T.J., Holliday, J., Gibson, L.M., Garratt, S., Oesingmann, N., Alfaro-Almagro, F., Bell, J.D., Boultwood, C., Collins, R., Conroy, M.C., et al. (2020). The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. Nat. Commun. *11*, 2624. https://doi.org/10.1038/s41467-020-15948-9.
26. Ronneberger, Olaf, Fischer, Phillip, and Brox Thomas. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science *9351*. https://doi.org/10.1007/978-3-319-24574-4_28.
27. Bydder, M., Ghodrati, V., Gao, Y., Robson, M.D., Yang, Y., and Hu, P. (2020). Constraints in estimating the proton density fat fraction. Magn. Reson. Imaging *66*, 1–8. https://doi.org/10.1016/j.mri.2019.11.009.
28. McKay, A., Wilman, H.R., Dennis, A., Kelly, M., Gyngell, M.L., Neubauer, S., Bell, J.D., Banerjee, R., and Thomas, E.L. (2018). Measurement of liver iron by magnetic resonance imaging in

the UK Biobank population. PLoS One *13*, e0209340. https://doi.org/10.1371/journal.pone.0209340.
29. Wood, J.C., Enriquez, C., Ghugre, N., Tyzka, J.M., Carson, S., Nelson, M.D., and Coates, T.D. (2005). MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. Blood *106*, 1460–1465. https://doi.org/10.1182/blood-2004-10-3982.
30. Basty, N., Liu, Y., Cule, M., Louise Thomas, E., Bell, J.D., and Whitcher, B. (2020). Automated measurement of pancreatic fat and iron concentration using multi-echo and T1-weighted MRI data. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (IEEE), pp. 345–348.
31. Millard, L.A.C., Davies, N.M., Gaunt, T.R., Davey Smith, G., and Tilling, K. (2018). Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. Int. J. Epidemiol. *47*, 29–35. https://doi.org/10.1093/ije/dyx204.
32. Carroll, R.J., Bastarache, L., and Denny, J.C. (2014). R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. Bioinformatics *30*, 2375–2376. https://doi.org/10.1093/bioinformatics/btu197.
33. Liu, Y., Basty, N., Whitcher, B., Bell, J.D., Sorokin, E.P., van Bruggen, N., Thomas, E.L., and Cule, M. (2021). Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. eLife *10*, e65554. https://doi.org/10.7554/elife.65554.
34. Sethi, A., Taylor, L., Graham Ruby, J., Venkataraman, J., Cule, M., and Melamud, E. Calcification of abdominal aorta is an underappreciated cardiovascular disease risk factor.Preprint at medRxiv https://doi.org/10.1101/2020.05.07.20094706
35. Loh, P.-R., Tucker, G., Bulik-Sullivan, B.K., Vilhjálmsson, B.J., Finucane, H.K., Salem, R.M., Chasman, D.I., Ridker, P.M., Neale, B.M., Berger, B., et al. (2015). Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat. Genet. *47*, 284–290. https://doi.org/10.1038/ng.3190.
36. Vuckovic, D., Bao, E.L., Akbari, P., Lareau, C.A., Mousas, A., Jiang, T., Chen, M.-H., Raffield, L.M., Tardaguila, M., Huffman, J.E., et al. (2020). The polygenic and monogenic basis of blood traits and diseases. Cell *182*, 1214–1231.e11. https://doi.org/10.1016/j.cell.2020.08.008.e11.
37. Chen, M.-H., Raffield, L.M., Mousas, A., Sakaue, S., Huffman, J.E., Moscati, A., Trivedi, B., Jiang, T., Akbari, P., Vuckovic, D., et al. (2020). Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. Cell *182*, 1198–1213.e14. https://doi.org/10.1016/j.cell.2020.06.045.e14.
38. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of ANthropometric Traits (GIANT) Consortium; and DIAbetes Genetics Replication and Meta-Analysis (DIAGRAM) Consortium, Madden, P.A.F., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. *44*, 369–375. S1–S3.
39. GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, et al. (2017). Genetic effects on gene expression across human tissues. Nature 550, 204–213.
40. Zhou, W., Nielsen, J.B., Fritsche, L.G., Dey, R., Gabrielsen, M.E., Wolford, B.N., LeFaive, J., VandeHaar, P., Gagliano, S.A., Gifford, A., et al. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat. Genet. *50*, 1335–1341. https://doi.org/10.1038/s41588-018-0184-y.
41. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. *10*, e1004383. https://doi.org/10.1371/journal.pgen.1004383.
42. Guo, H., Fortune, M.D., Burren, O.S., Schofield, E., Todd, J.A., and Wallace, C. (2015). Integration of disease association and eQTL data using a Bayesian colocalisation approach highlights six candidate causal genes in immune-mediated diseases. Hum. Mol. Genet. *24*, 3305–3313. https://doi.org/10.1093/hmg/ddv077.
43. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569. https://doi.org/10.1038/ng.608.
44. Bulik-Sullivan, B.K., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., ReproGen Consortium; Psychiatric Genomics Consortium; and Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, and Duncan, L., et al. (2015). An atlas of genetic correlations across human diseases and traits. Nat. Genet. *47*, 1236–1241.
45. Bell, S., Rigas, A.S., Magnusson, M.K., Ferkingstad, E., Allara, E., Bjornsdottir, G., Ramond, A., Sørensen, E., Halldorsson, G.H., Paul, D.S., et al. (2021). A genome-wide meta-analysis yields 46 new loci associating with biomarkers of iron homeostasis. Commun. Biol. *4*, 156. https://doi.org/10.1038/s42003-020-01575-z.
46. International HapMap 3 Consortium, Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., et al. (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58.
47. Regier, A.A., Farjoun, Y., Larson, D.E., Krasheninina, O., Kang, H.M., Howrigan, D.P., Chen, B.-J., Kher, M., Banks, E., Ames, D.C., et al. (2018). Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. Nat. Commun. *9*, 4038. https://doi.org/10.1038/s41467-018-06159-4.
48. Karczewski, K.J., Francioli, L.C., MacArthur, D.G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443. https://doi.org/10.1530/ey.17.14.3.
49. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315. https://doi.org/10.1038/ng.2892.
50. Zhou, W., Zhao, Z., Nielsen, J.B., Fritsche, L.G., LeFaive, J., Gagliano Taliun, S.A., Bi, W., Gabrielsen, M.E., Daly, M.J., Neale, B.M., et al. (2020). Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. Nat. Genet. *52*, 634–639. https://doi.org/10.1038/s41588-020-0621-6.
51. Henninger, B., Alustiza, J., Garbowski, M., and Gandon, Y. (2020). Practical guide to quantification of hepatic iron with

MRI. Eur. Radiol. *30*, 383–393. https://doi.org/10.1007/s00330-019-06380-9.

52. Elliott, L.T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K.L., Douaud, G., Marchini, J., and Smith, S.M. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature *562*, 210–216. https://doi.org/10.1038/s41586-018-0571-7.
53. The GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.
54. Ooi, G.C., Chen, F.E., Chan, K.N., Tsang, K.W., Wong, Y.H., Liang, R., Chan, V., and Ngan, H. (1999). Qualitative and quantitative magnetic resonance imaging in haemoglobin H disease: screening for iron overload. Clin. Radiol. *54*, 98–102. https://doi.org/10.1016/s0009-9260(99)91068-1.
55. Uhlen, M., Karlsson, M.J., Zhong, W., Tebani, A., Pou, C., Mikes, J., Lakshmikanth, T., Forsström, B., Edfors, F., Odeberg, J., et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. Science *366*, eaax9198. https://doi.org/10.1126/science.aax9198.
56. Sirlin, C.B., and Reeder, S.B. (2010). Magnetic resonance imaging quantification of liver iron. Magn. Reson. Imaging Clin. N. Am. *18*, 359–381. https://doi.org/10.1016/j.mric.2010.08.014.
57. Fernandes, J.L. (2018). MRI for iron overload in thalassemia. Hematol. Oncol. Clin. North. Am. *32*, 277–295. https://doi.org/10.1016/j.hoc.2017.11.012.
58. Arrivé, L., Thurnher, S., Hricak, H., and Price, D.C. (1990). Magnetic resonance imaging of splenic iron overload. Eur. J. Radiol. *10*, 98–104. https://doi.org/10.1016/0720-048x(90)90115-r.
59. Rushton, D.H., and Barth, J.H. (2010). What is the evidence for gender differences in ferritin and haemoglobin? Crit. Rev. Oncol. Hematol. *73*, 1–9. https://doi.org/10.1016/j.critrevonc.2009.03.010.
60. Timmers, P.R.H.J., Wilson, J.F., Joshi, P.K., and Deelen, J. (2020). Multivariate genomic scan implicates novel loci and haem metabolism in human ageing. Nat. Commun. *11*, 3570. https://doi.org/10.1038/s41467-020-17312-3.
61. Harrison-Findik, D.D. (2007). Role of alcohol in the regulation of iron metabolism. World J. Gastroenterol. *13*, 4925. https://doi.org/10.3748/wjg.v13.i37.4925.
62. Silva-Gomes, R., Mapelli, S.N., Boutet, M.-A., Mattiola, I., Sironi, M., Grizzi, F., Colombo, F., Supino, D., Carnevale, S., Pasqualini, F., et al. (2021). Differential expression and regulation of MS4A family members in myeloid cells in physiological and pathological conditions. J. Leukoc. Biol. *111*, 817–836. https://doi.org/10.1002/JLB.2A0421-200R.
63. Eon Kuek, L., Leffler, M., Mackay, G.A., and Hulett, M.D. (2016). The MS4A family: counting past 1, 2 and 3. Immunol. Cel Biol. *94*, 11–23. https://doi.org/10.1038/icb.2015.48.
64. Swirski, F.K., Nahrendorf, M., Etzrodt, M., Wildgruber, M., Cortez-Retamozo, V., Panizzi, P., Figueiredo, J.-L., Kohler, R.H., Chudnovskiy, A., Waterman, P., et al. (2009). Identification of splenic reservoir monocytes and their deployment to inflammatory sites. Science *325*, 612–616. https://doi.org/10.1126/science.1175202.