Title: Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis.

Authors: Benyamin B, Esko T, Ried JS, Radhakrishnan A, Vermeulen SH, Traglia M, Gögele M, Anderson D, Broer L, Podmore C, Luan J, Kutalik Z, Sanna S, van der Meer P, Tanaka T, Wang F, Westra HJ, Franke L, Mihailov E, Milani L, Hälldin J, Winkelmann J, Meitinger T, Thiery J, Peters A, Waldenberger M, Rendon A, Jolley J, Sambrook J, Kiemeney LA, Sweep FC, Sala CF, Schwienbacher C, Pichler I, Hui J, Demirkan A, Isaacs A, Amin N, Steri M, Waeber G, Verweij N, Powell JE, Nyholt DR, Heath AC, Madden PA, Visscher PM, Wright MJ, Montgomery GW, Martin NG, Hernandez D, Bandinelli S, van der Harst P, Uda M, Vollenweider P, Scott RA, Langenberg C, Wareham NJ, InterAct Consortium., van Duijn C, Beilby J, Pramstaller PP, Hicks AA, Ouwehand WH, Oexle K, Gieger C, Metspalu A, Camaschella C, Toniolo D, Swinkels DW, Whitfield JB

HHS Public Access

Author manuscript

Nat Commun. Author manuscript; available in PMC 2015 April 29.

Published in final edited form as:

Nat Commun. ; 5: 4926. doi:10.1038/ncomms5926.

Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis

A full list of authors and affiliations appears at the end of the article.

Abstract

Variation in body iron is associated with or causes diseases, including anaemia and iron overload. Here we analyse genetic association data on biochemical markers of iron status from eleven European-population studies, with replication in eight additional cohorts (total up to 48,972 subjects). We find eleven genome-wide-significant ($p < 5 \times 10^{-8}$) loci, some including known iron-related genes (*HFE*, *SLC40A1*, *TF*, *TFR2*, *TFRC*, *TMPRSS6*) and others novel (*ABO*, *ARNTL*, *FADS2*, *NAT2*, *TEX14*). SNPs at *ARNTL*, *TF*, and *TFR2* affect iron markers in *HFE* C282Y homozygotes at risk for hemochromatosis. There is substantial overlap between our iron loci and loci affecting erythrocyte and lipid phenotypes. These results will facilitate investigation of the roles of iron in disease.

INTRODUCTION

Absorption, transport and storage of iron are tightly regulated, as expected for an element which is both essential and potentially toxic. Iron deficiency is the leading cause of anaemia[1], and it also compromises immune function[2] and cognitive development 3. Iron overload damages the liver and other organs in hereditary hemochromatosis 4, and in thalassemia patients with both transfusion and non-transfusion-related iron accumulation[5]. Excess iron has harmful effects in chronic liver diseases caused by excessive alcohol, obesity or viruses[6]. There is evidence for involvement of iron in neurodegenerative diseases[7, 8, 9], and in Type 2 diabetes[10, 11]. Variation in transferrin saturation, a biomarker of iron status, has been associated with mortality in patients with diabetes[12] and in the general population[13]. All these associations between iron and either clinical disease or pathological processes make it important to understand the causes of variation in iron status. Importantly,

*Corresponding Authors: Beben Benyamin (b.benyamin@uq.edu.au), John B Whitfield (John.Whitfield@qimrberghofer.edu.au).

Study Organization: Beben Benyamin & John B Whitfield

Manuscript Preparation: Beben Benyamin, Clara Camaschella, Tonu Esko, Dorine W Swinkels, Daniela Toniolo, John B Whitfield. All authors reviewed and had the opportunity to comment on the manuscript.

Meta-analysis and Quality Controls: Beben Benyamin

Study Analysts (Discovery Cohorts): Denise Anderson, Beben Benyamin, Linda Broer, Tonu Esko, Martin Gogele, Aparna Radhakrishnan, Janina S Ried, Michela Traglia, Sita H Vermeulen

Study Analysts (Replication Cohorts): Tonu Esko, Clara Podmore, Jian'an Luan, Serenna Sanna, Fudi Wang, Pim van der Harst, Toshiko Tanaka, Zoltan Kutalik

Contributions to Organization and Conduct of Original Studies, Discovery and Replication cohorts: All authors.

Additional Analyses: Analysis of HFE C282Y homozygotes: Beben Benyamin; Pathway analysis (IPA): John B Whitfield.

Conflict Interests: None declared.

information on genetic causes of variation can be used in Mendelian randomisation studies to test whether variation in iron status is a cause or consequence of disease[14, 15].

We have used biomarkers of iron status (serum iron, transferrin, transferrin saturation and ferritin), which are commonly used clinically and readily measurable in thousands of individuals, and carried out a meta-analysis of human genome-wide association study (GWAS) data from eleven discovery and eight replication cohorts. These phenotypes show significant heritability in normal adults[16, 17], and previous population-based studies have identified relevant SNPs and gene loci (*HFE*, *TF*, *TFR2* and *TMPRSS6*[18, 19]) for iron status biomarkers. *HFE* and *TMPRSS6* have also been shown to affect red cell count, hemoglobin and erythrocyte indices[20], most likely by affecting iron availability[20, 21, 22].

Our aims were to identify additional loci affecting markers of iron status in the general population and to relate the significant loci to information on gene expression in order to identify relevant genes. We also made an initial assessment of whether any such loci affect iron status in *HFE* C282Y homozygotes, who are at genetic risk of *HFE*-related iron overload (hereditary hemochromatosis type 1, OMIM #235200).

Combination of results from discovery and replication stages of our analysis shows significant effects on one or more of the iron biomarkers at 11 loci. Those primarily affecting serum iron and transferrin saturation include, or are close to, genes whose products have recognised roles in iron homeostasis; *HFE* (the haemochromatosis gene), *TMPRSS6* (transmembrane protease, serine 6) and *TFR2* (transferrin receptor 2). Those mainly affecting serum transferrin, apart from the *TF* (transferrin) gene itself and *TFRC* (transferrin receptor), and those mainly affecting ferritin (apart from *SLC40A1*, solute carrier family 40 (iron-regulated transporter), member 1) are unexpected. There is significant overlap between the genes or loci affecting iron biomarkers and those known to affect erythrocyte numbers or size, which is reasonable given the importance of iron for erythropoesis. We also find significant overlap between genes or loci affecting iron biomarkers and known loci affecting plasma lipids or lipoproteins, showing an unexplained link between these areas of metabolism.

RESULTS

SNP and gene associations

The combination of allelic association data from eleven discovery and eight replication cohorts (Supplementary Tables 1 – 3) showed eleven loci with significant effects on one or more of the iron-related phenotypes (Table 1, Fig. 1, Supplementary Fig. 1 – 2, Supplementary Table 4). Four of these (*HFE*, *TF*, *TFR2*, *TMPRSS6*) were previously known to affect iron biomarker variation in the general population[18, 19]. Genes at two newly significant loci, *SLC40A1* which codes for the cellular iron exporter ferroportin and *TFRC* which codes for the iron importer transferrin receptor 1, are known to be important for cellular iron homeostasis 23. The other five loci (chromosome 8 at 18.3 Mbp, nearest gene *NAT2*; chromosome 9 at 136.2 Mbp, nearest gene *ABO*; chromosome 11 at 13.4 Mbp, nearest gene *ARNTL*; chromosome 11 at 61.6 Mbp, nearest gene *FADS2*; chromosome 17 at 54.1 Mbp, nearest gene *TEX14*) were not previously known to affect any of these

phenotypes. These affect either transferrin (*NAT2*, *ARNTL*, *FADS2*) or ferritin (*ABO*, *TEX14*).

Conditional analysis on the discovery cohorts (Table 1, Supplementary Fig. 4) showed additional independent signals at the *TF* locus for transferrin and transferrin saturation and at *TMPRSS6* for iron. Gene-based analysis in the discovery cohort (Supplementary Table 5) gave significant results (critical p-value for testing of 17,000 genes < 3×10^{-6}) for ferritin in a region covering two genes (*C15orf43* and *SORD*) on chromosome 15, where individual SNPs gave only suggestive associations. Allelic associations across this region are also shown in Supplementary Fig. 2. This locus did not show any SNPs with genome-wide significance in the combined discovery + replication data.

In the replication cohorts, the lead SNPs at the eleven significant loci explained 3.4%, 7.2%, 6.7%, and 0.9% of the phenotypic variance for iron, transferrin, saturation and (log-transformed) ferritin, respectively.

Secondary analyses

In view of the known association between ferritin concentration and inflammatory conditions, we repeated the discovery meta-analysis of ferritin including C-reactive protein (CRP, a marker of inflammation) as a covariate. This resulted in a decrease in effect sizes (expressed as standardized regression slopes or betas in an additive-allelic-effect model) for the lead SNPs at significant and suggestive loci, to an average of 73% (SD 15%) of the previous betas (Supplementary Fig. 5). The p-values became less significant, partly because of the decrease in effect size and partly because the number of subjects with CRP data was less than the number available for the initial analysis.

To check whether results were similar after excluding people with iron deficiency, we removed subjects with serum ferritin concentration below 30 μg/l and repeated the meta-analyses for all four phenotypes. This decreased effect sizes for transferrin and transferrin saturation, but had negligible effects for SNPs which were significant or suggestive for ferritin or iron compared to those from the all-subjects analysis (Supplementary Table 6).

We also examined the association between serum transferrin concentration and *FADS2* variation. Because this gene is known to be associated with other phenotypes related to lipids and components of the metabolic syndrome, we included high-density lipoprotein cholesterol (HDL-C) as a covariate and repeated the association meta-analysis for transferrin and the most significant SNP at the *FADS2* locus, rs174577. (HDL-C was chosen because it was available for a greater proportion of subjects than either triglycerides or glucose, which are also associated with *FADS* polymorphisms.) This conditional analysis resulted in a 35% reduction in the effect size for this SNP, from $\beta = 0.068 \pm 0.011$ to 0.044 ± 0.009.

Effects at Significant Loci on Gene Expression and Regulation

We next checked for data which may help explain the biological role of the significant SNPs or identify the causal variants which they tag, using sources listed in the Methods. The synthesis of information from our results and external sources is exemplified in Fig. 2, which shows the alignment of data at the *TFR2* locus. The region which includes genome-

wide-significant SNPs (after replication) for serum iron contains documented eQTLs for *TFR2*, and H3K27Ac histone modification sites (documented in data from ENCODE). In this case, there is striking alignment at the region around 100.2 Mbp at one end of the *TFR2* gene, which includes the most significant SNPs at this locus, documented eQTLs for this gene, and the histone modification in K562 (erythroleukaemia) cells.

A similar approach was taken for the other significant loci, as summarized in Supplementary Table 7. SNPs identified through the GWAS had significant cis-effects on expression of *SLC40A1*, *TFRC*, *ARNTL*, and *FADS1/FADS2*. At the *C15orf43-SORD* locus on chromosome 15, rs16976620 (allelic association with ferritin $p = 4.52 \times 10^{-7}$) affected expression of *SORD* at $p = 4.02 \times 10^{-4}$. The chromosome 22 region near *TMPRSS6* contains eQTLs for hepatic expression of *TMPRSS6*[24]. However the chromosome 3 locus near *TF* contains eQTLs for *SRPRB* but not for *TF*; and SNPs at the loci identified as *TFR2*, *ABO* and *TEX14* are eQTLs for multiple other genes. The ENCODE regulatory data show potential regulatory sequences or histone marks in the regions where we found SNP associations on chromosome 2 near *SLC40A1*, chromosome 11 near the *FADS* genes, and at the chromosome 17 locus near *TEX14*.

Some lead SNPs from our significant loci also showed *trans*-effects on more distant genes (Supplementary Table 7). Most notably, the three non-synonymous coding SNPs in *HFE* and *TMPRSS6* (rs1800562, rs1799945 and rs855791) had strong effects on expression of *ALAS2* (aminolevulinate, delta-, synthase 2), which catalyses the initial step in heme synthesis in erythroid tissues.

Overlap with other phenotypes and disease associations

Because of previous data showing that iron-related loci overlap with loci affecting erythrocyte phenotypes, and because several of our significant loci have been reported to affect lipid phenotypes, we compared our results against published meta-analysis data on erythrocytes and lipids. Results are summarized in Supplementary Table 8. Among the 75 significant loci for erythrocyte phenotypes[25], we found associations with one or more of our iron phenotypes after Bonferroni correction for multiple testing at $p < 6.7 \times 10^{-4}$ ($p < 0.05$ adjusted for testing of 75 SNPs) for *ABO*, *HFE*, *TFR2*, *TFRC* and *TMPRSS6*, and additionally for *HBS1L* ($p = 9.78 \times 10^{-7}$ for transferrin saturation) and *PGS1* ($p = 1.84 \times 10^{-4}$ for ferritin). For the 157 lipid loci reported by the Global Lipids Genetics Consortium[26], two loci (*HFE* and *HBS1L*) gave $p < 3.18 \times 10^{-4}$ ($p < 0.05$ adjusted for testing of 157 loci) for iron and saturation, six (*FADS1/2/3*, *GCKR*, *HFE*, *NAT2*, *SNX5* and *TRIB1*) for transferrin, and six (*ABO*, *HFE*, *LOC84931*, *LRP1*, *PGS1* and *TRIB1*) for ferritin. Moreover, plots of observed versus expected p-value distributions for the iron phenotypes (Fig. 3) showed that even the erythrocyte and lipid loci not reaching statistical significance do affect iron biomarkers to a greater degree than can be explained by chance.

The SNP association results were also analysed using Ingenuity Pathway Analysis, selecting SNPs which showed associations at $p < 0.01$, < 0.001 and < 0.0001 for transferrin saturation, and similarly for ferritin. Results for these two phenotypes, chosen as markers of iron availability and iron stores, showed substantial overlap. The $p < 0.01$ threshold identified an excess of genes which have been reported to affect or be associated with lung cancer,

cardiovascular disease and diabetes; and also with a range of developmental and nerve cell functions (Supplementary Fig. 6). Results for the $p < 0.001$ threshold were similar but showed lesser statistical significance, as expected because of the smaller number of genes included.

Effects on iron status in HFE C282Y homozygotes

We tested whether the lead SNPs at loci which affect iron-related biomarkers in the general population also explain variation in iron status in C282Y homozygotes who are at genetic risk of *HFE*-related iron overload. These comprised 76 homozygotes from the QIMR Adult cohort (one of the discovery cohorts), plus 277 homozygotes from the HEIRS study[27]. Results are shown in Table 2 for significant associations, and more fully in Supplementary Table 10.

The strong association between rs8177240 in the *TF* gene and serum transferrin was clearly present in *HFE* YY homozygotes ($p = 1.93 \times 10^{-9}$). The YY group showed association between serum iron and rs7385804 at *TFR2* ($\beta = 0.178 \pm 0.053$, $p = 0.00076$, critical p-value = 0.005 after adjusting for testing of ten loci). The standardized beta for this SNP was approximately three times as great in the YY sample as in the overall meta-analysis (0.178 ± 0.053 against 0.055 ± 0.010). There was also a significant association ($p = 0.0022$) between rs6486121 in *ARNTL* and ferritin. When we checked for associations between a genetic risk score calculated from the significant and suggestive SNPs in the population-based meta-analysis results, and the biomarker phenotypes in the HEIRS sample, only transferrin showed a significant association and this was stronger among the men than the women (Supplementary Table 11).

DISCUSSION

Our meta-analysis of GWAS on iron-related phenotypes from up to 48,000 people of European descent showed multiple significant associations. Some increased the significance of loci known from previous studies or showed significant associations with additional phenotypes (*TF*, *TFR2*, *HFE*, *TMPRSS6*); some were at loci containing genes whose products have known roles in iron homeostasis, including the transferrin receptor *TFRC* and the iron transporter ferroportin (*SLC40A1*); and others were novel (near to *ARNTL*, *FADS2* and *NAT2* for transferrin, *ABO* and *TEX14* for ferritin). Significant associations were found for biomarkers of iron status that reflect both cellular iron metabolism and systemic regulation of iron[23].

There was variation in the phenotypes affected by the significant loci, as summarized In Supplementary Fig. 3. Three of the loci mainly affected serum ferritin (*ABO*, *SLC40A1*, *TEX14*); three others mainly affected serum iron and transferrin saturation (*HFE*, *TFR2*, *TMPRSS6*); and five mainly affected serum transferrin (*ARNTL*, *FADS2*, *NAT2*, *TF* and *TFRC*). The loci with the strongest effects on serum iron (*HFE*, *TMPRSS6*) had significant, but smaller, effects on serum ferritin and it is likely that this is due to higher circulating concentrations of iron leading over time to higher iron stores and hence higher serum ferritin.

We note that there are factors which can modify the relationships between these biomarker phenotypes and whole-body or tissue-specific iron status. Ferritin has been criticised as a marker of iron stores because it is an acute-phase protein increased by inflammation, but comparisons with independent methods[28, 29] have validated it sufficiently for use in epidemiological studies. Moreover the loci which affected ferritin in this study have not been reported in GWAS for inflammatory biomarkers or CRP[30], and *SLC40A1*, which showed a highly significant association with ferritin, has strong biological plausibility because it codes for ferroportin. Including CRP as a covariate in the ferritin association analysis changed the effect size similarly for all the significant or suggestive SNPs (Supplementary Figure 5), whereas effects related to both inflammation and iron status would be expected to alter betas for some SNPs and not others.

We also note that matching of significant loci to genes is subject to uncertainty. For some, the location of the peak association close to a gene with a known and relevant physiological function gives confidence in the gene assignment. For others, data from previous reports or databases on association between SNPs and gene expression will identify a probable gene, but in other cases expression data are consistent with any of several genes or else no relevant data are available. If so, the name of the nearest gene may be provided for identification of the locus but this may require revision as more information becomes available.

Five confirmed loci contain genes (*TF*, *TFR2*, *HFE*, *TMPRSS6*, *SLC40A1*) which were already known to affect iron homeostasis. These genes have been previously identified via monogenic diseases or from functional studies. Interestingly, no association has been identified with genes for several other important players in iron homeostasis such as ferritin, the protein that safely stores excess iron, or hepcidin and hemojuvelin, which are essential in the hepcidin signalling pathway and when mutated cause severe juvenile-onset hemochromatosis (type 2A, 2B). Mutations at the loci identified cause late-onset (*HFE*, type 1) or less severe (*TFR2*, type 3 and *SLC40A1*, type 4A) hemochromatosis.

SNPs at *HFE* and *TMPRSS6* which mainly affect iron and transferrin saturation showed interesting *trans*-effects on gene expression for *ALAS2*. As this gene is on the X chromosome and we only analysed GWAS data for autosomes, we do not know whether *ALAS2* variation affects our phenotypes. However, ALAS2 activity controls the initial and rate-limiting step in porphyrin synthesis so a co-ordinated effect on both iron and protoporphyrin availability for formation of heme is an interesting possibility.

SLC40A1 is a prime candidate for affecting iron stores, as it codes for ferroportin and mutations in this gene are associated with the autosomal dominant type 4 hemochromatosis, characterized by high ferritin levels. The most significant SNPs near *SLC40A1* in our study are about 45 and 60 kbp from the gene, but are known to affect *SLC40A1* expression. Variation near *SLC40A1* also affects transferrin, probably through an effect on cellular iron availability.

Genome-wide studies of erythrocyte traits known to vary with iron status[20, 21, 22, 25] have previously found associations with many of these loci: erythrocyte volume (MCV) and hemoglobin content (MCH) with *HFE*, *TFR2*, *TFRC* and *TMPRSS6*; hematocrit with *HFE*,

TFR2, and *TMPRSS6*; and erythrocyte count with *TFR2*[25]. The results for our iron data at loci known to affect erythrocyte phenotypes are illustrated in Fig. 3a; an unexpectedly high proportion of them affect iron, transferrin and ferritin.

New associations were found for ferritin near *ABO* and *TEX14*. The *ABO* blood group locus has shown significant associations for several phenotypes; rs651007 has a particularly strong effect on E-selectin[31] and has also been found in GWAS on low-density lipoprotein cholesterol[32], coronary artery disease[33] and red blood cell count[25]. The latter is relevant to our ferritin finding, but whether *ABO* variation primarily affects iron stores and therefore erythrocyte count, or *vice versa*, is unclear.

TEX14 codes for a testis-expressed protein, but there was no evidence for male-female heterogeneity in the effect on ferritin (p_{Het} for the lead SNP, rs368243, was 0.45). The most significant SNPs are within the *TEX14* gene but the suggestive-significance region extends across other genes. Expression data suggest that variation affecting *RAD51C* may be important, but the function of this gene (in DNA repair and meiosis) also has no obvious connection with iron status. The same holds for *SEPT4*, for which rs411988 is an expression QTL. Another gene within the LD block, *MTMR4*, deserves consideration because it changes SMAD phosphorylation, with possible effects on the BMP-SMAD pathway affecting control of hepcidin[34]. The region on chromosome 15 identified in the gene-based analysis is centered on *C15orf43* but also overlaps with *SORD* (sorbitol dehydrogenase). *SORD* has no obvious connection with iron status and the function of the protein coded by *C15orf43* is unknown, although there is some evidence that it is present in human plasma (http://pax-db.org/#!protein/986968, accessed 2014-03-27). These two loci illustrate the difficulty which may be encountered in interpreting allelic associations; in some cases the region containing the most significant results overlaps with several genes, there may be unrecognised regulatory regions with effects on more distant genes, and data on gene expression may not reflect expression in the relevant tissue. For all these reasons, assignment of significant effects to specific genes must often be provisional.

Effects on transferrin were seen for most of the loci which affect serum iron, including *HFE*, *TF*, *TFRC*, and *TMPRSS6*. Contrary to the result for *TFRC*, variation at the other transferrin receptor gene *TFR2* did not affect transferrin concentration; this may reflect the different functions of the two receptors. TfRC is involved in cellular iron uptake, which may directly affect regulation of transferrin expression. TfR2 on the other hand has been reported to be involved in hepatocyte sensing of circulating iron and signalling to hepcidin production, which may subsequently affect circulating levels of iron and the transferrin saturation. TfR2 variation could also affect these iron parameters through its effect on erythropoiesis[35].

Transferrin was also affected by SNPs near *ARNTL*, *NAT2* and *FADS2*. The role of these in iron homeostasis is uncertain; transferrin is central to iron transport and receptor-mediated uptake by cells but these loci did not affect serum iron or ferritin. *ARNTL*, and its product BMAL1, is mainly known for interactions with CLOCK genes and generation of circadian rhythm. Notably, serum iron[16, 36], liver iron[37], hepcidin[38], and *TfR1* gene expression[39] all show circadian variation. The region affecting transferrin on chromosome 8 contains the *NAT2* gene, which again has no obvious relevance for iron. It has been shown to affect

lipids[32] and cardiovascular risk (see Supplementary Table 8 of [40]). The gene product is important for xenobiotic metabolism; *NAT2* codes for an N-acetyl transferase which determines fast- or slow-acetylator status. At *FADS2*, the significant SNPs for transferrin are intronic but they affect expression of *FADS* genes. *FADS1/2/3* variation affects a wide range of phenotypes including serum lipids[32, 41], polyunsaturated fatty acid content of serum phospholipids[42]; fatty acid composition of membranes and phospholipids[43]; fasting glucose and insulin response[44, 45], and liver enzymes[46]. The most significant *FADS* SNPs for lipids are rs174546, rs174547 and rs174548[32, 41, 47] and each gave significant or near-significant p-values for transferrin in our data ($p = 7.43 \times 10^{-10}$, 8.47×10^{-10} and 7.29×10^{-8}, respectively). This, together with the decrease in the locus effect on transferrin after inclusion of HDL-C as a covariate, suggests a common basis for effects on lipids and transferrin. The pathways involved are unknown, but iron homeostasis and lipid metabolism show overlap in the literature[32, 48, 49, 50, 51] as well as in our data. It has recently been shown, for example, that signalling pathways for the protein kinase mTOR, which regulates energy metabolism and lipid synthesis among other functions[52], affect transcriptional control of hepcidin and therefore potentially affect iron uptake and distribution[53].

Despite the varied functions of these three genes (*ARNTL*, *FADS2*, *NAT2*) which unexpectedly affect transferrin, they have the common feature of significant effects on plasma triglycerides[26]. Detailed comparison of our results against published lipid loci showed that a high proportion of lipid loci (not only for triglycerides) have detectable effects on our iron phenotypes, especially on transferrin (Fig. 2b, Supplementary Table 8). The pleiotropic effects at such loci, connecting iron homeostasis not only with erythropoiesis but also with lipids and possibly with cardiovascular risk, deserve further investigation.

One important clinical question about iron overload is why some *HFE* C282Y homozygotes develop biochemical evidence of iron overload and clinical symptoms of hemochromatosis, while most do not[54]. A systematic review of longitudinal studies found that 38–76% of homozygous people have increased ferritin and transferrin saturation (biochemical penetrance)[55]. However, clinical symptoms are less common at 2–38% in men and 1–10% in women[56, 57]. We therefore evaluated the effects of our lead SNPs in *HFE* C282Y homozygotes, combining data from the largest of our Discovery cohorts with available phenotypic information and DNA from participants in the HEIRS study.

Because of limited numbers of C282Y homozygotes (total N available for data analysis was 353), we had limited power to detect relevant effects. Among our results, the association between two SNPs in *TFR2* and serum iron seems the most relevant. There is both clinical and experimental evidence for interaction between the gene products of *HFE* and *TFR2*. Severe juvenile hemochromatosis occurred in a family carrying mutations in both *HFE* and *TFR2*[58]. In mice, homozygosity for deletion of both *Hfe* and *TfR2* greatly decreases hepcidin levels[59] and causes massive iron overload[60]. These reports are consistent with evidence that TFR2 and HFE proteins interact in control of hepcidin signaling; they may form an iron-sensing complex that modulates hepcidin expression in response to blood levels of diferric transferrin[61, 62].

Overall, there was a lack of correlation between effect sizes for lead SNPs at the significant loci identified in the general population, and in the YY homozygotes. Similarly, a predictor based on allele count and effect size for SNPs taken forward for replication and genotyped in the HEIRS subjects did not significantly predict iron, saturation or ferritin in the HEIRS C282Y homozygotes (Supplementary Table 11 and Supplementary Figure 7). The exception, transferrin, was due to the strong effects at the *TF* locus.

Previous studies have proposed determinants of *HFE* clinical or biochemical penetrance. Apart from age, sex and probably alcohol intake[63], the focus has been on genetic modifiers but no candidate has been convincingly identified[64]. Since iron homeostasis involves a complex regulating network[23, 65], it seems probable that any genetic effects on penetrance are either highly polygenic (in which case large genome-wide studies on *HFE* C282Y homozygotes will be needed) or result from rare variants which have not yet been examined in sufficient detail. *TFR2* variation as a modifier of *HFE* C282Y risk has statistical support and biological plausibility but confirmation is needed.

Our results have revealed genes or loci whose effects on iron status were previously unsuspected and which need to be integrated into our understanding of iron homeostasis. Discovery of SNPs which significantly affect iron status, and compilation of genomic scores, will allow Mendelian randomisation studies on the multiple conditions associated with variation in iron load and help to clarify a potential causal role of iron in such conditions (e.g. Parkinson's[14] or Alzheimer's 66 diseases). However the existence of pleiotropic effects, with many loci affecting both iron and lipid phenotypes, shows the need for caution in selecting SNPs or scores for such applications.

METHODS

Subjects

We established the Genetics of Iron Status Consortium (GISC) to coordinate our efforts in understanding the causes and consequences of genetic variation in biochemical markers for iron status, i.e. serum iron, transferrin, transferrin saturation and ferritin. Discovery samples consisted of summary data on genome-wide allelic associations between SNP genotypes and iron markers from 23,986 subjects of European ancestry gathered from 11 cohorts in 9 participating centres (Supplementary Table 1). Replication samples to confirm suggestive and significant associations were obtained from up to 24,986 subjects of European ancestry in 8 additional cohorts (also in Supplementary Table 1). There was no systematic selection whether a cohort was allocated into the discovery or replication samples. This allocation was based on the availability of data when the analyses were conducted. Information on phenotypic means, methods for phenotype measurement, and genotyping methods for each contributing cohort are shown in Supplementary Tables 2 and 3. Each participating study was approved by the appropriate human research ethics committee, as listed for each study in Supplementary Table 1, and all subjects gave informed consent.

Genome-wide association studies

Genome-wide association tests, genotype imputation and associated quality control procedures (QCs) were performed in each cohort separately. Within each cohort, QCs were applied to individual samples and SNPs prior to imputation into HAPMAP II (Release 22, NCBI Build36, dbSNP b126) or, for InterAct, 1000 Genomes. These include removing individuals based on missingness, relatedness, population and ethnic outliers. Poor-quality SNPs were also removed based on missingness, minor allele frequency, Hardy-Weinberg equilibrium test and Mendelian errors for family data. These QCs for each cohort are summarized in Supplementary Table 3.

The association between genotyped and imputed SNPs and each iron phenotype was performed using an additive model for allelic effects, on the standardized residuals of the phenotype after adjusting for age, principal component scores and other study specific covariates, for each sex separately. The details of the association analysis and imputation method for each cohort are presented in Supplementary Table 3.

Meta-analysis

We conducted meta-analysis of GWAS results from the discovery cohorts in the Metal package[67] using a standard error based approach, which weights the SNP effect size (standardized regression slope, beta) using the inverse of the corresponding squared standard errors. SNPs were included in the meta-analysis if they met the following conditions: imputation quality score either Rsq (which estimates the squared correlation between imputed and true genotypes) for MACH software >=0.3, or the 'info' measure for IMPUTE software > 0.5; Hardy-Weinberg Equilibrium Test p-value (pHWE) $>=10^{-6}$; minor allele frequency (MAF) >=0.01; genotyping Call Rate >=0.95 and if they survived QCs in all cohorts to avoid disproportionate contribution of a single cohort to the meta-analysis. In total ~2.1 million SNPs met these conditions. A genomic control correction was applied to all cohorts. Heterogeneity of effect sizes between cohorts or between sexes was also assessed using Cochran's Q statistic within Metal. Loci containing SNPs with $p < 5 \times 10^{-6}$ were carried forward for *in silico* replication in independent samples, again using Metal for the meta-analysis. The threshold p-value for choice of SNPs for replication is conventional and based in part on data for European populations in Duggal et al[68].

Power to detect allelic effects was estimated using the Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/). Under reasonable assumptions about allele frequencies for causative and marker polymorphisms (QTL increaser allele frequency = 0.2, marker allele frequency = 0.2, linkage disequilibrium between them $d' = 0.8$, $\alpha = 5 \times 10^{-8}$), the Discovery dataset with N = 24,000 gives 77% power to detect allelic effects which each account for 0.25% of the phenotypic variance.

Gene-based analysis

Gene-based analysis considers all SNPs within a gene as a unit for the association analysis. We performed gene-based analysis on SNP association p-values from the meta-analysis of discovery samples using VEGAS (http://gump.qimr.edu.au/VEGAS/, accessed

2014-03-27)[69]. The significance of gene-based analysis was based on Bonferroni correction of testing ~17,000 genes (i.e. $p < 3 \times 10^{-6}$).

Conditional analysis

To find independent signals within each significant locus, we performed conditional analysis in each cohort by repeating the association analysis but including the most significant SNP at each significant locus (in the initial meta-analysis) as covariates. We performed meta-analysis of the conditional association results using the same approach as in the main meta-analysis.

Gene expression

The eQTL look-up was based on a meta-analysis of expression data for known disease-associated loci in non-transformed peripheral blood cells, from 5,300 samples from seven cohorts. The original analysis used HapMap2 imputed SNPs and a cis-window of +/− 250kb from the transcription start-site. More details can be found in the paper by Westra et al.[70].

Information on gene expression in macrophages and monocytes was based on results obtained by the Cardiogenics consortium, on 758 samples, as described in the Supplementary Note to [33]. Online resources for gene expression and regulation included http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/, http://genenetwork.nl/bloodeqtlbrowser/ and Schadt et al. (2008) 24 for eQTL data, http://genome.ucsc.edu/ENCODE/ for information on histone modification, and http://ecrbrowser.dcode.org/ for comparison of DNA sequences across species.

Bioinformatic analyses

Pathway analysis and assessment of known disease associations or biological functions was performed using Ingenuity Pathway Analysis (IPA; Ingenuity Inc, Redwood City, CA, 94063), selecting SNPs which showed associations at $p < 0.01$, < 0.001 and < 0.0001 for transferrin saturation, and similarly for ferritin. IPA compares the list of genes associated with the selected SNPs against a proprietary library of gene-disease and gene-function associations and tests frequencies of observed and expected occurrences.

Analysis in HFE C282Y homozygotes

Data and DNA samples from *HFE* C282Y homozygotes in the HEIRS study[27] were obtained from the NIH Biologic Specimen and Data Repository Information Coordinating Center (*BioLINCC*) (https://biolincc.nhlbi.nih.gov/home/)]. HEIRS was a population-based survey of the prevalence and effects of *HFE* polymorphisms, and subjects were not selected for having a diagnosis or positive family history of hemochromatosis. Selected SNPs (those showing significant or suggestive results in our primary meta-analysis) were genotyped by primer-extension mass spectrometry (MassArray, Sequenom Inc, San Diego CA); all samples were confirmed as being homozygous for the minor allele of rs1800562 by this method. Allelic association results for the QIMR adults and HEIRS C282Y homozygotes were combined by meta-analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Beben Benyamin[1,2,*], Tonu Esko[3,4], Janina S Ried[5], Aparna Radhakrishnan[6], Sita H Vermeulen[7,8], Michela Traglia[9,10], Martin Gögele[11], Denise Anderson[12], Linda Broer[13,14], Clara Podmore[15], Jian'an Luan[15], Zoltan Kutalik[16,17], Serena Sanna[18], Peter van der Meer[19], Toshiko Tanaka[20], Fudi Wang[21], Harm-Jan Westra[22], Lude Franke[22], Evelin Mihailov[3,23], Lili Milani[3], Jonas Häldin[3], Juliane Winkelmann[24,25,26,27], Thomas Meitinger[26,28], Joachim Thiery[29,30], Annette Peters[31,32], Melanie Waldenberger[31,32], Augusto Rendon[6,33,34], Jennifer Jolley[6,33], Jennifer Sambrook[6,33], Lambertus A Kiemeney[7,35], Fred C Sweep[36], Cinzia F Sala[37], Christine Schwienbacher[11], Irene Pichler[11], Jennie Hui[38,39], Ayse Demirkan[13,40], Aaron Isaacs[13,41], Najaf Amin[13], Maristella Steri[18], Gérard Waeber[42], Niek Verweij[19], Joseph E Powell[1,43], Dale R Nyholt[2], Andrew C Heath[44], Pamela AF Madden[44], Peter M Visscher[1,43], Margaret J Wright[2], Grant W Montgomery[2], Nicholas G Martin[2], Dena Hernandez[45], Stefania Bandinelli[46], Pim van der Harst[19,47,48], Manuela Uda[18], Peter Vollenweider[42], Robert A Scott[15], Claudia Langenberg[15], Nicholas J Wareham[15], InterAct Consortium, Cornelia van Duijn[13,41,49], John Beilby[38,39], Peter P Pramstaller[11,47], Andrew A Hicks[11], Willem H Ouwehand[6,33], Konrad Oexle[28], Christian Gieger[5,31,32], Andres Metspalu[3], Clara Camaschella[9,53], Daniela Toniolo[9,54], Dorine W Swinkels[36], and John B Whitfield[2,*]

Affiliations

[1]The University of Queensland, Queensland Brain Institute, Queensland 4072, Australia [2]QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4029, Australia [3]Estonian Genome Center, University of Tartu, Riia 23b, 51010, Tartu, Estonia [4]The Broad Institute of Massachusetts Institute of Technology and Harvard University, 320 Charles St, Cambridge, MA 02141, USA [5]Institute of Genetic Epidemiology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany [6]Department of Haematology, University of Cambridge, Cambridge Biomedical Campus Box 139, Hills Road, Cambridge CB2 0XY, United Kingdom [7]Department for Health Evidence, Radboud University Medical Centre, Geert Grooteplein-Zuid 10, 6525 GA Nijmegen, the Netherlands [8]Department of Genetics, Radboud University Medical Centre, Geert Grooteplein-Zuid 10, 6525 GA Nijmegen, the Netherlands [9]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, 20132 Milano, Italy [10]Institute for Maternal and Child Health - IRCCS "Burlo Garofolo"- Trieste, 34137 Trieste, Italy [11]Center for Biomedicine, European Academy of Bolzano/Bozen (EURAC), Via Galvani 31, Bolzano, Italy [12]Telethon Institute for Child Health Research, Centre for Child Health Research, The University of Western Australia. PO Box 855 West Perth WA 6872 Australia [13]Department of Epidemiology, Subdivision Genetic Epidemiology, Erasmus Medical Center, Postbus 2040, 3000 CA Rotterdam, The Netherlands [14]Department of Internal Medicine, Erasmus Medical Center, Postbus 2040, 3000 CA Rotterdam,

The Netherlands [15]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge CB2 0QQ UK [16]Institute of Social and Preventive Medicine (IUMSP), Centre Hospitalier Universitaire Vaudois (CHUV), Rue du Bugnon 46, 1011 Lausanne, Switzerland [17]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland [18]Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche (CNR), Monserrato, 09042, Cagliari, Italy [19]Department of Cardiology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, the Netherlands [20]Translational Gerontology Branch, National Institute on Aging, 251 Bayview Boulevard, Suite 100, Baltimore, MD 21224-6825, USA [21]School of Public Health, Zhejiang University, Hangzhou 310058, P. R. China [22]Department of Genetics, University Medical Center Groningen, P.O. Box 30001 9700 RB Groningen, The Netherlands [23]Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia [24]Department of Neurology and Neurosciences and Center for Sleep Sciences and Medicine, Stanford University, Palo Alto, CA, 94304, USA [25]Neurologische Klinik und Poliklinik, Klinikum rechts der Isar, Technische Universität München, 81675 Munich, Germany [26]Institut für Humangenetik, Helmholtz Zentrum München, 85764 Neuherberg, Germany [27]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany [28]Institute of Human Genetics, Klinikum Rechts der Isar, Technische Universität München, Ismaninger Straße 22, 81675 München, Germany [29]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, Universitätsklinikum Leipzig, Liebigstraße 27, 04103 Leipzig Germany [30]LIFE Research Center for Civilization Diseases, Medical Faculty, University Leipzig, Philipp-Rosenthal-Straße 27, 04103 Leipzig Germany [31]Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany [32]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany [33]NHS Blood and Transplant, Cambridge Biomedical Campus, Long Road, Cambridge CB2 0PT, United Kingdom [34]Medical Research Council Biostatistics Unit, Cambridge Biomedical Campus, Cambridge CB2 0SR, United Kingdom [35]Department of Urology, Radboud University Medical Centre, Geert Grooteplein-Zuid 10, 6525 GA Nijmegen, The Netherlands [36]Laboratory of Genetic, Endocrine and Metabolic Diseases (LGEM), Department of Laboratory Medicine, Radboud University Medical Centre, Geert Grooteplein-Zuid 10, 6525 GA Nijmegen, The Netherlands [37]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, 20132 Milano, Italy [38]PathWest Laboratory Medicine of WA, Nedlands WA 6009 [39]Schools of Pathology and Laboratory Medicine, and Population Health, The University of Western Australia, Nedlands WA 6009 [40]Department of Human Genetics, Leiden University Medical Center, P.O. Box 9600 2300 RC Leiden, The Netherlands [41]Center of Medical Systems Biology, Einthovenweg 20, 2333 ZC Leiden, The Netherlands [42]Department of Medicine, Internal Medicine, CHUV, University of Lausanne, Rue du Bugnon 46, 1011 Lausanne, Switzerland

[43]University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4072, Australia [44]Department of Psychiatry, Washington University, 660 S. Euclid St. Louis, MO 63110, USA [45]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA [46]Geriatric Unit, Azienda Sanitaria Firenze (ASF), Via delle Casine, 16, 50122 Firenze, Italy [47]Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, 3511 GC Utrecht, The Netherlands [48]Department of Genetics, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands [49]Member of Netherlands Consortium for Healthy Aging sponsored by Netherlands Genomics Initiative, Leiden, The Netherlands [50]Department of Neurology, General Central Hospital, Lorenz-Böhler-Straße, 5, Bolzano, Italy [51]Research Unit of Molecular Epidemiology, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany [52]Institute of Epidemiology II, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany [53]Vita-Salute University, Milan, Italy [54]Institute of Molecular Genetics, CNR, 27100 Pavia, Italy

Acknowledgments

BB was funded by an Australian National Health and Medical Research Council (NHMRC) training fellowship (552498). There was no specific funding for the meta-analysis, and sources of funding for the contributing studies are listed in Supplementary Table 1. We acknowledge the important work of the investigators of the Hemochromatosis and Iron Overload Screening Study (HEIRS), and the data and DNA from that study made available to us by the US National Institutes of Health through the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC).

References

1. Miller JL. Iron deficiency anemia: a common and curable disease. Cold Spring Harb Perspect Med. 2013; 3
2. Cherayil BJ. Iron and immunity: immunological consequences of iron deficiency and overload. Arch Immunol Ther Exp (Warsz). 2010; 58:407–415. [PubMed: 20878249]
3. McCann JC, Ames BN. An overview of evidence for a causal relation between iron deficiency during development and deficits in cognitive or behavioral function. Am J Clin Nutr. 2007; 85:931–945. [PubMed: 17413089]
4. Swinkels DW, Janssen MC, Bergmans J, Marx JJ. Hereditary hemochromatosis: genetic complexity and new diagnostic approaches. Clin Chem. 2006; 52:950–968. [PubMed: 16627556]
5. Porter JB, Shah FT. Iron overload in thalassemia and related conditions: therapeutic goals and assessment of response to chelation therapies. Hematol Oncol Clin North Am. 2010; 24:1109–1130. [PubMed: 21075283]
6. Philippe MA, Ruddell RG, Ramm GA. Role of iron in hepatic fibrosis: one piece in the puzzle. World J Gastroenterol. 2007; 13:4746–4754. [PubMed: 17729396]
7. Liu B, et al. Iron promotes the toxicity of amyloid beta peptide by impeding its ordered aggregation. J Biol Chem. 2011; 286:4248–4256. [PubMed: 21147772]
8. Liu Y, et al. Mutant HFE H63D protein is associated with prolonged endoplasmic reticulum stress and increased neuronal vulnerability. J Biol Chem. 2011; 286:13161–13170. [PubMed: 21349849]
9. Nandar W, Connor JR. HFE gene variants affect iron in the brain. J Nutr. 2011; 141:729S–739S. [PubMed: 21346098]

10. Kunutsor SK, Apekey TA, Walley J, Kain K. Ferritin levels and risk of type 2 diabetes mellitus: an updated systematic review and meta-analysis of prospective evidence. Diabetes Metab Res Rev. 2013; 29:308–318. [PubMed: 23381919]
11. Simcox JA, McClain DA. Iron and diabetes risk. Cell Metab. 2013; 17:329–341. [PubMed: 23473030]
12. Ellervik C, et al. Total Mortality by Elevated Transferrin Saturation in Patients With Diabetes. Diabetes Care. 2013
13. Ellervik C, Tybjaerg-Hansen A, Nordestgaard BG. Total Mortality by Transferrin Saturation Levels: Two General Population Studies and a Meta-Analysis. Clin Chem. 2011
14. Pichler I, et al. Serum iron levels and the risk of Parkinson disease: a mendelian randomization study. PLoS Med. 2013; 10:e1001462. [PubMed: 23750121]
15. Rhodes SL, et al. Pooled analysis of iron-related genes in Parkinson's disease: Association with transferrin. Neurobiol Dis. 2013; 62C:172–178. [PubMed: 24121126]
16. Whitfield JB, et al. Effects of HFE C282Y and H63D polymorphisms and polygenic background on iron stores in a large community sample of twins. AmJHumGenet. 2000; 66:1246–1258.
17. Traglia M, et al. Heritability and demographic analyses in the large isolated population of Val Borbera suggest advantages in mapping complex traits genes. PLoS One. 2009; 4:e7554. [PubMed: 19847309]
18. Benyamin B, et al. Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. Nat Genet. 2009; 41:1173–1175. [PubMed: 19820699]
19. Pichler I, et al. Identification of a common variant in the TFR2 gene implicated in the physiological regulation of serum iron levels. Hum Mol Genet. 2011; 20:1232–1240. [PubMed: 21208937]
20. Ganesh SK, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet. 2009; 41:1191–1198. [PubMed: 19862010]
21. Chambers JC, et al. Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat Genet. 2009; 41:1170–1172. [PubMed: 19820698]
22. Soranzo N, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet. 2009; 41:1182–1190. [PubMed: 19820697]
23. Hentze MW, Muckenthaler MU, Galy B, Camaschella C. Two to tango: regulation of Mammalian iron metabolism. Cell. 2010; 142:24–38. [PubMed: 20603012]
24. Schadt EE, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol. 2008; 6:e107. [PubMed: 18462017]
25. van der Harst P, et al. Seventy-five genetic loci influencing the human red blood cell. Nature. 2012
26. Global Lipids Genetics C, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013; 45:1274–1283. [PubMed: 24097068]
27. Adams PC, et al. Hemochromatosis and iron-overload screening in a racially diverse population. N Engl J Med. 2005; 352:1769–1778. [PubMed: 15858186]
28. Skikne BS, Flowers CH, Cook JD. Serum transferrin receptor: a quantitative measure of tissue iron deficiency. Blood. 1990; 75:1870–1876. [PubMed: 2331526]
29. Olthof AW, et al. Correlation between serum ferritin levels and liver iron concentration determined by MR imaging: impact of hematologic disease and inflammation. Magn Reson Imaging. 2007; 25:228–231. [PubMed: 17275618]
30. Dehghan A, et al. Meta-Analysis of Genome-Wide Association Studies in >80 000 Subjects Identifies Multiple Loci for C-Reactive Protein Levels. Circulation. 2011
31. Qi L, et al. Genetic variants in ABO blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes. Hum Mol Genet. 2010; 19:1856–1862. [PubMed: 20147318]
32. Teslovich TM, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010; 466:707–713. [PubMed: 20686565]
33. Schunkert H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011; 43:333–338. [PubMed: 21378990]

34. Yu J, et al. Myotubularin-related protein 4 (MTMR4) attenuates BMP/Dpp signaling by dephosphorylation of Smad proteins. J Biol Chem. 2013; 288:79–88. [PubMed: 23150675]
35. Forejtnikova H, et al. Transferrin receptor 2 is a component of the erythropoietin receptor complex and is required for efficient erythropoiesis. Blood. 2010; 116:5357–5367. [PubMed: 20826723]
36. Ridefelt P, Larsson A, Rehman JU, Axelsson J. Influences of sleep and the circadian rhythm on iron-status indices. Clin Biochem. 2010; 43:1323–1328. [PubMed: 20800590]
37. Unger EL, Earley CJ, Beard JL. Diurnal cycle influences peripheral and brain iron levels in mice. J Appl Physiol (1985). 2009; 106:187–193. [PubMed: 18988764]
38. Schaap CC, et al. Diurnal rhythm rather than dietary iron mediates daily hepcidin variations. Clin Chem. 2013; 59:527–535. [PubMed: 23232066]
39. Okazaki F, et al. Circadian rhythm of transferrin receptor 1 gene expression controlled by c-Myc in colon cancer-bearing mice. Cancer Res. 2010; 70:6238–6246. [PubMed: 20631077]
40. Suhre K, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature. 2011; 477:54–60. [PubMed: 21886157]
41. Kathiresan S, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet. 2009; 41:56–65. [PubMed: 19060906]
42. Lemaitre RN, et al. Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. PLoS Genet. 2011; 7:e1002193. [PubMed: 21829377]
43. Merino DM, Ma DW, Mutch DM. Genetic variation in lipid desaturases and its impact on the development of human disease. Lipids Health Dis. 2010; 9:63. [PubMed: 20565855]
44. Dupuis J, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010; 42:105–116. [PubMed: 20081858]
45. Ingelsson E, et al. Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. Diabetes. 2010; 59:126–275.
46. Chambers JC, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet. 2011; 43:1131–1138. [PubMed: 22001757]
47. Waterworth DM, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. Arterioscler Thromb Vasc Biol. 2010; 30:2264–2276. [PubMed: 20864672]
48. Shalev H, Kapelushnik J, Moser A, Knobler H, Tamary H. Hypocholesterolemia in chronic anemias with increased erythropoietic activity. Am J Hematol. 2007; 82:199–202. [PubMed: 17039515]
49. Saeed O, et al. Pharmacological suppression of hepcidin increases macrophage cholesterol efflux and reduces foam cell formation and atherosclerosis. Arterioscler Thromb Vasc Biol. 2012; 32:299–307. [PubMed: 22095982]
50. Finn AV, et al. Hemoglobin directs macrophage differentiation and prevents foam cell formation in human atherosclerotic plaques. J Am Coll Cardiol. 2012; 59:166–177. [PubMed: 22154776]
51. Ahmed U, Latham PS, Oates PS. Interactions between hepatic iron and lipid metabolism with possible relevance to steatohepatitis. World J Gastroenterol. 2012; 18:4651–4658. [PubMed: 23002334]
52. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell. 2012; 149:274–293. [PubMed: 22500797]
53. Mleczko-Sanecka K, et al. Unbiased RNAi screen for hepcidin regulators links hepcidin suppression to the proliferative Ras/RAF and the nutrient-dependent mTOR signaling pathways. Blood. 2014; 123:1574–1585. [PubMed: 24385536]
54. van Bokhoven MA, van Deursen CT, Swinkels DW. Diagnosis and management of hereditary haemochromatosis. BMJ. 2011; 342:c7251. [PubMed: 21248018]
55. Whitlock EP, Garlitz BA, Harris EL, Beil TL, Smith PR. Screening for hereditary hemochromatosis: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2006; 145:209–223. [PubMed: 16880463]
56. Beutler E, Felitti VJ, Koziol JA, Ho NJ, Gelbart T. Penetrance of 845G–> A (C282Y) HFE hereditary haemochromatosis mutation in the USA. Lancet. 2002; 359:211–218. [PubMed: 11812557]

57. Allen KJ, et al. Iron-overload-related disease in HFE hereditary hemochromatosis. N Engl J Med. 2008; 358:221–230. [PubMed: 18199861]
58. Pietrangelo A, et al. Juvenile hemochromatosis associated with pathogenic mutations of adult hemochromatosis genes. Gastroenterology. 2005; 128:470–479. [PubMed: 15685557]
59. Tjalsma H, et al. Mass spectrometry analysis of hepcidin peptides in experimental mouse models. PLoS One. 2011; 6:e16762. [PubMed: 21408141]
60. Wallace DF, Summerville L, Crampton EM, Frazer DM, Anderson GJ, Subramaniam VN. Combined deletion of Hfe and transferrin receptor 2 in mice leads to marked dysregulation of hepcidin and iron overload. Hepatology. 2009; 50:1992–2000. [PubMed: 19824072]
61. Goswami T, Andrews NC. Hereditary hemochromatosis protein, HFE, interaction with transferrin receptor 2 suggests a molecular mechanism for mammalian iron sensing. J Biol Chem. 2006; 281:28494–28498. [PubMed: 16893896]
62. Gao J, Chen J, Kramer M, Tsukamoto H, Zhang AS, Enns CA. Interaction of the hereditary hemochromatosis protein HFE with transferrin receptor 2 is required for transferrin-induced hepcidin expression. Cell Metab. 2009; 9:217–227. [PubMed: 19254567]
63. Scotet V, et al. Hereditary hemochromatosis: effect of excessive alcohol consumption on disease expression in patients homozygous for the C282Y mutation. Am J Epidemiol. 2003; 158:129–134. [PubMed: 12851225]
64. Weiss G. Genetic mechanisms and modifying factors in hereditary hemochromatosis. Nat Rev Gastroenterol Hepatol. 2010; 7:50–58. [PubMed: 19918260]
65. Casanovas G, Banerji A, d'Alessio F, Muckenthaler MU, Legewie S. A multi-scale model of hepcidin promoter regulation reveals factors controlling systemic iron homeostasis. PLoS computational biology. 2014; 10:e1003421. [PubMed: 24391488]
66. Crespo AC, et al. Genetic and biochemical markers in patients with Alzheimer's disease support a concerted systemic iron homeostasis dysregulation. Neurobiol Aging. 2014; 35:777–785. [PubMed: 24199959]
67. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–2191. [PubMed: 20616382]
68. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. BMC genomics. 2008; 9:516. [PubMed: 18976480]
69. Liu JZ, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010; 87:139–145. [PubMed: 20598278]
70. Westra HJ, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet. 2013; 45:1238–1243. [PubMed: 24013639]

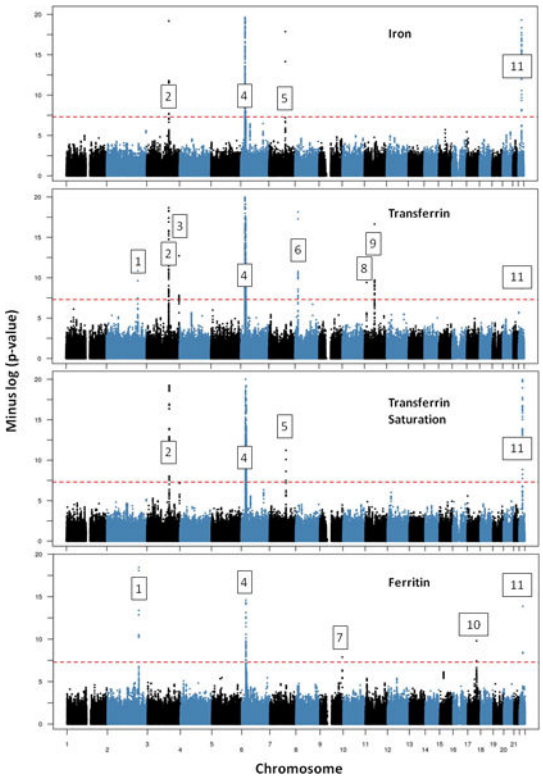

Figure 1. Manhattan plots for the associations between SNPs and iron status (serum iron, transferrin, transferrin saturation and ferritin)

The −log(p) values are from inverse-variance-weighted meta-analysis of the combined Discovery + Replication datasets for SNPs taken forward for replication, otherwise from the Discovery datasets only. Genes are assigned to the loci as follows: 1 *SLC40A1*; 2 *TF*; 3 *TFRC*; 4 *HFE*; 5 *TFR2*; 6 *NAT2*; 7 *ABO*; 8 *ARNTL*; 9 *FADS2*; 10 *TEX14*; 11 *TMPRSS6*. Note that the y-axis for −log(p) values is truncated at 20.

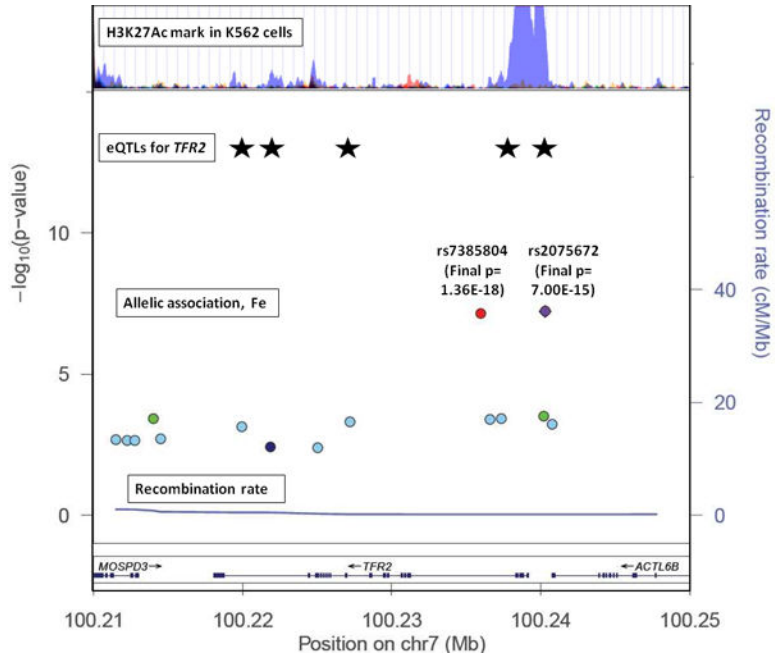

Figure 2. Comparison of results for serum iron with regulatory features at the chromosome 7 (*TFR2*) locus

From bottom: regional association plot with recombination rate and −log(p) values for serum iron; documented eQTL locations for *TFR2* expression (from left to right: rs10247962, rs4729598, rs7457868, rs4729600, rs1052897); ENCODE data on histone modification. P-values for serum iron at rs7385804 and rs2075672 are shown as text (Final p) for the Discovery + Replication dataset, but positions for all SNPs on the y-axis are determined by the Discovery dataset only.

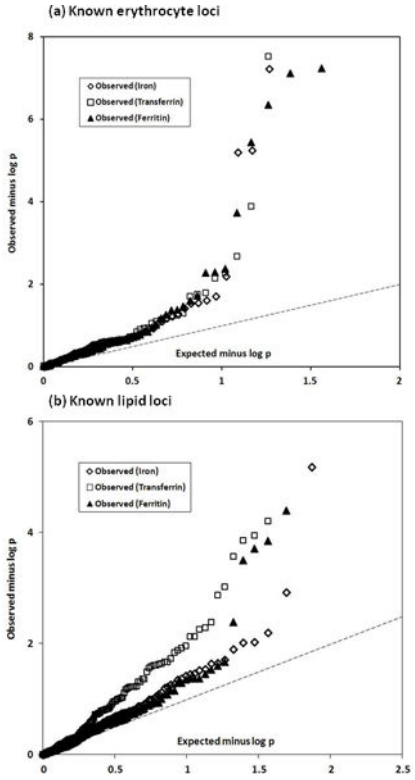

Figure 3. Q-Q plots of the association p-values for iron, transferrin and ferritin at loci previously reported to be significant for (a) erythrocyte phenotypes[25] and (b) plasma lipid phenotypes[26]

For clarity, the y-axes only extend to $p < 10^{-8}$ or $p < 10^{-6}$ so that two associations with observed $p < 10^{-8}$ for erythrocyte loci (at *HFE* and *TMPRSS6*) and four associations with $p < 10^{-6}$ for lipid loci (at *ABO*, *FADS2*, *HFE* and *NAT2*) are not plotted. The interrupted line in each plot is the line of equivalence, observed = expected.

Table 1

Results from discovery, discovery + replication and conditional analyses. For each SNP and phenotype, results are given for the Discovery cohorts only (D, N = 23,986) and for the combined Discovery and Replication cohorts (D+R, N = 48,972). p-values are from meta-analysis of covariate-adjusted standardized regression coefficients of phenotypic values on the allele count for A1.

CHR	SNP	BP (Build 37)	Nearest Gene(s)†	A1 ††	A2	Freq A1	Phenotype		Beta	SE	p
2	rs744653	190,378,750	*WDR75 – SLC40A1*	T	C	0.854	Iron	D	−0.002	0.014	0.902
								D+R	0.004	0.010	0.702
							Transferrin	D	0.092	0.014	2.00×10^{-10}
								D+R	0.068	0.010	1.35×10^{-11}
							Saturation	D	−0.037	0.014	0.0087
								D+R	−0.028	0.011	0.0084
							Ferritin (log)	D	−0.098	0.013	1.20×10^{-13}
								D+R	−0.089	0.010	8.37×10^{-19}
3	rs8177240	133,477,701	*TF*	T	G	0.669	Iron	D	−0.073	0.011	2.37×10^{-12}
								D+R	−0.066	0.007	6.65×10^{-20}
							Transferrin	D	−0.423	0.011	3.82×10^{-340}
								D+R	−0.380	0.007	3.29×10^{-615}
							Saturation	D	0.097	0.011	$5.85 \times 10-^{20}$
								D+R	0.100	0.008	7.24×10^{-38}
							Ferritin (log)	D	0.028	0.010	0.0050
								D+R	0.021	0.007	0.0039
3	rs9990333	195,827,205	*TFRC*	T	C	0.460	Iron	D	0.021	0.010	0.030
								D+R	0.017	0.007	0.014
							Transferrin	D	−0.067	0.010	3.01×10^{-11}
								D+R	−0.051	0.007	1.95×10^{-13}
							Saturation	D	0.049	0.010	7.37×10^{-7}
								D+R	0.039	0.007	7.28×10^{-8}
							Ferritin (log)	D	0.002	0.009	0.829
								D+R	0.001	0.007	0.878
6	rs1800562	26,093,141	*HFE*	A	G	0.067	Iron	D	0.372	0.020	3.96×10^{-77}

CHR	SNP	BP (Build 37)	Nearest Gene(s)†	A1 ††	A2	Freq A1	Phenotype		Beta	SE	p
			(C282Y)					D+R	0.328	0.016	2.72×10^{-97}
							Transferrin	D	−0.550	0.021	1.26×10^{-153}
								D+R	−0.479	0.016	8.90×10^{-196}
							Saturation	D	0.577	0.020	1.52×10^{-178}
								D+R	0.577	0.016	2.19×10^{-270}
							Ferritin (log)	D	0.211	0.019	1.43×10^{-29}
								D+R	0.204	0.016	1.54×10^{-38}
6	rs1799945	26,091,179	*HFE* (H63D)	C	G	0.850	Iron	D	−0.190	0.014	1.65×10^{-42}
								D+R	−0.189	0.010	1.10×10^{-81}
							Transferrin	D	0.119	0.014	5.59×10^{-17}
								D+R	0.114	0.010	9.36×10^{-30}
							Saturation	D	−0.228	0.014	2.98×10^{-60}
								D+R	−0.231	0.010	5.13×10^{-109}
							Ferritin (log)	D	−0.059	0.013	7.38×10^{-6}
								D+R	−0.065	0.010	1.71×10^{-10}
7	rs7385804	100,235,970	*TFR2*	A	C	0.621	Iron	D	0.055	0.010	7.19×10^{-8}
								D+R	0.064	0.007	1.36×10^{-18}
							Transferrin	D	−0.009	0.011	0.396
								D+R	−0.003	0.007	0.728
							Saturation	D	0.054	0.010	1.79×10^{-7}
								D+R	0.054	0.008	6.07×10^{-12}
							Ferritin (log)	D	0.022	0.010	0.0255
								D+R	0.015	0.007	0.039
8	rs4921915	18,272,466	*NAT2*	A	G	0.782	Iron	D	−0.009	0.012	0.477
								D+R	0.004	0.009	0.633
							Transferrin	D	0.082	0.012	1.74×10^{-11}
								D+R	0.079	0.009	7.05×10^{-19}
							Saturation	D	−0.034	0.012	0.0041
								D+R	−0.026	0.009	0.0036
							Ferritin (log)	D	−0.006	0.011	0.603

CHR	SNP	BP (Build 37)	Nearest Gene(s)†	A1 ††	A2	Freq A1	Phenotype		Beta	SE	p
								D+R	0.001	0.009	0.886
9	rs651007	136,153,875	*ABO*	T	C	0.202	Iron	D	−0.012	0.013	0.358
								D+R	−0.004	0.009	0.611
							Transferrin	D	0.017	0.013	0.188
								D+R	−0.001	0.009	0.916
							Saturation	D	−0.020	0.013	0.110
								D+R	−0.006	0.009	0.498
							Ferritin (log)	D	−0.060	0.012	2.54×10^{-7}
								D+R	−0.050	0.009	1.31×10^{-8}
11	rs6486121	13,355,770	*ARNTL*	T	C	0.631	Iron	D	0.001	0.010	0.898
								D+R	−0.009	0.007	0.202
							Transferrin	D	−0.056	0.011	1.04×10^{-7}
								D+R	−0.046	0.007	3.89×10^{-10}
							Saturation	D	0.026	0.010	0.0132
								D+R	0.015	0.008	0.048
							Ferritin (log)	D	0.012	0.010	0.2298
								D+R	0.006	0.007	0.424
11	rs174577	61,604,814	*FADS2*	A	C	0.330	Iron	D	0.003	0.011	0.785
								D+R	0.001	0.007	0.878
							Transferrin	D	0.068	0.011	1.90×10^{-10}
								D+R	0.062	0.007	2.28×10^{-17}
							Saturation	D	−0.023	0.011	0.029
								D+R	−0.025	0.008	0.0016
							Ferritin (log)	D	−0.020	0.010	0.040
								D+R	−0.012	0.007	0.098
17	rs411988	56,709,034	*TEX14*	A	G	0.564	Iron	D	−0.007	0.010	0.4673
								D+R	−0.002	0.007	0.770
							Transferrin	D	0.033	0.010	0.0012
								D+R	0.014	0.007	0.052
							Saturation	D	−0.021	0.010	0.036
								D+R	−0.012	0.007	0.115

CHR	SNP	BP (Build 37)	Nearest Gene(s)†	A1 ††	A2	Freq A1	Phenotype		Beta	SE	p
							Ferritin (log)	D	−0.049	0.009	1.28×10^{-7}
								D+R	−0.044	0.007	1.59×10^{-10}
22	rs855791	37,462,936	*TMPRSS6*	A	G	0.446	Iron	D	−0.187	0.010	4.31×10^{-77}
			(V736A)					D+R	−0.181	0.007	1.32×10^{-139}
							Transferrin	D	0.040	0.010	0.00013
								D+R	0.044	0.007	1.98×10^{-9}
							Saturation	D	−0.192	0.010	3.50×10^{-80}
								D+R	−0.190	0.008	6.41×10^{-137}
							Ferritin (log)	D	−0.051	0.010	5.81×10^{-8}
								D+R	−0.055	0.007	1.38×10^{-14}
Conditional analysis											
3	rs8177179	133,463,457	*TF*	A	G	0.521	Transferrin	D	−0.154	0.010	2.74×10^{-49}
3	rs1799852	133,475,722	*TF* ((L247L)	T	C	0.098	Saturation	D	0.110	0.019	7.13×10^{-9}
22	rs228916	37,505,552	*TMPRSS6*	T	C	0.875	Iron	D	−0.086	0.016	2.94×10^{-8}

†Where the SNP is a coding variant, the amino acid change is also shown.

††A1 is the effect allele for each SNP in the association analysis.

Table 2

Results for *HFE* YY subjects. For each SNP and phenotype, results are given for C282Y homozygotes (N = 353) from the QIMR Adult cohort and the HEIRS samples. p-values are from meta-analysis of covariate-adjusted standardized regression coefficients of phenotypic values on the allele count for A1. Results are shown for the most significant SNP at each locus where any SNP shows $p < 0.005$ for any of the four phenotypes.

SNP	CHR	BP (Build 37)	Nearest gene	A1	A2	Freq A1	Phenotype	Effect	SE	p
rs8177240	3	134,962,864	*TF*	T	G	0.669	Iron	−0.094	0.053	0.077
							Transferrin	−0.306	0.051	1.93×10^{-9}
							Saturation	0.017	0.054	0.752
							Ferritin	0.022	0.051	0.670
rs7385804	7	100,078,232	*TFR2*	A	C	0.621	Iron	0.178	0.053	0.00076
							Transferrin	0.038	0.054	0.485
							Saturation	0.119	0.054	0.026
							Ferritin	−0.037	0.052	0.471
rs6486121	11	13,355,770	*ARNTL*	T	C	0.631	Iron	−0.057	0.053	0.288
							Transferrin	0.029	0.053	0.588
							Saturation	−0.058	0.054	0.280
							Ferritin	−0.153	0.05	0.0022