A novel association between a SNP in *CYBRD1* and serum ferritin levels in a cohort study of *HFE* Hereditary Haemochromatosis

C. C. Constantine[1], **G. J. Anderson**[2], **C. D. Vulpe**[3], **C. E. McLaren**[4], **M. Bahlo**[5], **H. L. Yeap**[1], **D. M. Gertig**[6], **N. J. Osborne**[1,7,8], **N. A. Bertalli**[1], **K. B. Beckman**[9], **V. Chen**[9], **P. Matak**[10], **A. T. McKie**[10], **M. B. Delatycki**[7], **J. K. Olynyk**[11], **D. R. English**[1], **M. C. Southey**[12], **G. G. Giles**[13], **J. L. Hopper**[1], **K. J. Allen**[7,8], and **L. C. Gurrin**[1]

[1]Centre for MEGA Epidemiology, School of Population Health, University of Melbourne, Carlton, Victoria 3053, Australia

[2]The Queensland Institute of Medical Research, Herston, Queensland 4029, Australia

[3]Department of Nutritional Sciences and Toxicology, University of California, Berkeley, California 94720, USA

[4]Department of Epidemiology, College of Health Sciences, University of California, Irvine, California 92697, USA

[5]The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3050, Australia

[6]Victorian Cervical Cytology Registry, East Melbourne, Victoria 3002, Australia

[7]Murdoch Childrens Research Institute, Parkville, Victoria 3052, Australia

[8]Department of Paediatrics, University of Melbourne, Parkville, Victoria 3052, Australia

[9]Children's Hospital Oakland Research Institute, Oakland, California 94609, USA

[10]Nutritional Sciences Division, King's College London, London SE1 9NH, UK

[11]School of Medicine & Pharmacology, University of Western Australia and Western, Australian Institute of Medical Research, Nedlands, Australia

[12]Department of Pathology, University of Melbourne, Carlton, Victoria 3053, Australia

[13]The Cancer Council Victoria, Carlton, Victoria 3053, Australia

Abstract

There is emerging evidence that there are genetic modifiers of iron indices for *HFE* gene mutation carriers at risk of hereditary hemochromatosis. A random sample stratified by *HFE* genotype of 863 from a cohort of 31,192 people of northern European descent provided blood samples for genotyping of 476 SNPs in 44 genes involved in iron metabolism. Single SNP association testing using linear regression models adjusted for sex, menopause and *HFE* genotype was conducted for four continuously distributed outcomes: serum ferritin (log transformed), transferrin saturation, serum transferrin, and serum iron. The SNP rs884409 in *CYBRD1* is a novel modifier specific to *HFE* C282Y homozygotes. Median unadjusted serum ferritin concentration decreased from 1194 µg/L (N=27) to 387 µg/L (N=16) for male C282Y homozygotes and from 357 µg/L (N=42) to 69 µg/L (N=12) for females, comparing those with no copies to those with one copy of rs884409. Functional testing of this *CYBRD1* promoter polymorphism using a heterologous expression assay

Correspondence: Lyle Gurrin lgurrin@unimelb.edu.au Phone: +61 (0)3 8344 0743 Fax: +61 (0)3 9349 5815.

resulted in a 30% decrease in basal promoter activity relative to the common genotype (p=0.004). This putative genetic modifier of iron overload expression accounts for 11% (95% CI 0.4%, 22.6%) of the variance in serum ferritin levels of C282Y homozygotes.

Keywords

Introduction

More than 80% of Caucasians with hereditary haemochromatosis (HH) are homozygous for the C282Y mutation in the *HFE* gene, but only 28% of men and 1% of women with that genotype develop iron overload-related disease (Allen, *et al* 2008). There is strong evidence that the moderate penetrance of C282Y homozygosity for iron-overload disease is due in part to modifying genetic factors. A high concordance of iron indices and/or iron-related disease has been found between related individuals (Crawford, *et al* 1993, McLaren, *et al* 2008, Njajou, *et al* 2006, Whiting, *et al* 2002), although this is also consistent with the presence of shared environmental factors that modify the iron phenotype. Bulaj *et al.* (Bulaj, *et al* 2000) reported that C282Y homozygotes who were relatives of symptomatic probands were more likely to have symptoms of HH disease (defined as cirrhosis, fibrosis grade 2 or above, aminotransferase above 1.2 of upper reference range in the absence of other cause, and radiographically confirmed arthropathy of the metacarpophalangeal joints). Two other studies of relatives of C282Y homozygotes showed at least the same prevalence of elevated iron indices and disease symptoms for those C282Y homozygote relatives identified through family screening as for those C282Y homozygotes ascertained via a proband presenting with symptoms (McCune, *et al* 2006, Powell, *et al* 2006). Stronger evidence supporting the presence of genetic modifiers was reported by Whitfield *et al.* (Whitfield, *et al* 2000) who studied a sample of both monozygotic and dizygotic twin pairs and showed that the pattern of residual variation in serum iron indices, after adjusting for an effect of the C282Y mutation, was consistent with the additive effects of multiple genes. After correcting for age and body-mass index they estimated the proportion of variance explained by additive genetic factors, for men and women respectively, was 23% and 31% for iron, 66% and 49% for transferrin, 33% and 47% for transferrin saturation, and 47% and 47% for ferritin.

Studies of inbred mice strains provide direct evidence of genetic contributions to iron homeostasis; more than two fold variation in biochemical measurements of iron status, including serum iron, transferrin saturation, and hepatic iron, was found when mice were fed a basal iron diet (Leboeuf, *et al* 1995). Data from studies of *Hfe* knockout mice have shown that the degree of hepatic iron loading varies considerably depending on the genetic background on which the knockout is placed (Leboeuf, *et al* 1995, Sproule, *et al* 2001).

Since the *HFE* gene was identified in 1996 there have been substantial advances in the understanding of the molecular processes that control body iron loading. The liver-derived peptide hepcidin plays a central role in this process and has been shown to act on the small intestine to limit iron intake, and most iron loading disorders share the common feature of inappropriately low hepcidin levels (Ganz 2008). HFE has been shown to be an upstream regulator of hepcidin, as have transferrin receptor 2 (TfR2), haemojuvelin (HJV), members of the BMP-SMAD signalling pathway and proinflammatory cytokines (Darshan and Anderson 2009). Mutations in any of the genes encoding components of this regulatory network, or those facilitating iron transport across the intestinal epithelium, have the potential to alter iron loading.

A recent study by Milet *et al.* (Milet, *et al* 2007) of a cohort of 592 unrelated C282Y homozygous probands who attended the Liver Unit in Rennes, France since 1990 was the first to show strong evidence for an association between a measured genetic variant (a SNP in the *BMP2* gene; BMP2 is a stimulator of hepcidin expression) and the serum ferritin levels of C282Y homozygotes. BMP6 has recently been shown to be the key endogenous regulator of hepcidin (Andriopoulos, *et al* 2009, Camaschella 2009, Meynard, *et al* 2009). Mutations in the TMPRSS6 gene (another upstream regulator of hepcidin) have been implicated in iron refractory iron deficient anaemia (Finberg, *et al* 2008, Guillem, *et al* 2008, Melis, *et al* 2008) through linkage studies although these results are based on a few extended pedigrees and may have limited relevance at the population level. Mutations in many other genes are known to cause serious disruption of normal iron metabolism (e.g. *HFE2*, *HAMP*, *SLC40A1, TFR2*) but the causal mutations are very rare (Wallace and Subramaniam 2007).

A recent study (Benyamin, *et al* 2009) investigated genetic influences on markers of iron status using a cohort of twins and their siblings and conducting a genome-wide association study (GWAS) on four serum markers of iron status (iron, transferrin, transferrin saturation and ferritin). As well as confirming previously reported associations of *HFE* C282Y on all four markers, they found strong associations between a *TMPRSS6* SNP (rs4820268) and serum iron, and several *TF* SNPs (rs3811647, rs1358024, rs4525863 and rs6794945 in adjacent *SRPRB* gene) and serum transferrin. Although this study identified SNP associations worthy of further investigation, the analysis was unable to adjust for menopause status in women or to examine associations within all *HFE* genotype groups due to small sample sizes.

A major limitation to the understanding of genetic modifiers of iron indices has been a paucity of data derived from large, prospective population-based studies. Here we present results from a candidate gene study using a prospective cohort design where selection was stratified on the two most important *HFE* mutations, C282Y and H63D. A series of genes involved in iron metabolism were selected assuming that SNPs in these genes are likely to act as modifiers of iron phenotypes. SNPs from these genes were genotyped, and biochemical and clinical phenotypes were recorded prospectively before many participants were aware of their *HFE* status, in an *HFE*-stratified random sample of participants of northern European descent (1).

Material and Methods

Study subjects

Between 1990 and 1994, 41,528 people (24,479 women) aged between 27 and 75 years were recruited for the Melbourne Collaborative Cohort Study (MCCS)(Giles and English 2002); 99.3% were 40 to 69 years old. Recruitment was via the Australian Electoral Roll (both registration and voting are compulsory in Australia), advertisements and community announcements in local media. Participants attended a study centre where they were interviewed about a range of lifestyle and dietary factors, had physical measurements taken and provided a blood sample. For the present study, known as "HealthIron", participants born in southern Europe were excluded, leaving 31,192 participants born in Australia, the United Kingdom, Ireland or New Zealand (i.e. almost exclusively of northern European ancestry).

Preliminary HFE genotyping of baseline samples

DNA from stored baseline samples was extracted from Guthrie cards (n=23,484) using a Chelex method or from buffy coats (Corbett Buffy Coat CorProtocol™ 14102) which had

been stored in liquid nitrogen (n=7,708) and genotyped for the C282Y and H63D alleles of the *HFE* gene using Taqman ® real-time PCR probes. Genotyping for the H63D allele was performed only on samples from people heterozygous for the C282Y allele. The overall successful genotyping percentage (including missing samples) was 95% (29,676/31,192).

HealthIron Clinics

Between 2004 and 2006, letters of invitation to participate in a HealthIron clinical assessment were sent to a sample of 1438 participants which included all C282Y homozygotes and a stratified random sample of participants from the remaining *HFE* genotype groups. Participants completed a computer-assisted personal interview and provided a cheekbrush sample for confirmatory *HFE* genotyping (Applied Biosystem 7000 real-time PCR with Taqman ® probes). Stored frozen serum collected in the morning at baseline MCCS attendance was used for measurement of serum iron indices (serum ferritin, serum transferrin, transferrin saturation and serum iron).

All participants gave written, informed consent to participate in the MCCS and the HealthIron study. Both study protocols were approved by The Cancer Council Victoria's Human Research Ethics Committee. Participants who provided a blood sample at the HealthIron clinic were genotyped for selected SNPs using DNA extracted from buffy coats.

Genotyping of SNPs from candidate genes

Tag SNPs were selected based on a candidate gene approach using resequencing data from a variety of sources and HapMap (Constantine, *et al* 2008). Resequencing of exonic regions in 11 genes was carried out on samples from 94 randomly selected (54 wild type, 22 H63D simple heterozygotes, 18 C282Y simple heterozygotes, two compound heterozygotes, two H63D homozygotes, and one C282Y homozygote) and 94 C282Y homozygous participants in HealthIron. Two other resequencing data sets were included: (A) the HEIRS ancillary study resequencing, completed by the Resequencing and Genotyping Service of the National Heart Lung and Blood Institute of samples from five populations for 14 genes (EMBL accession numbers are *TFR2*:DQ496110, *TF*: DQ525716, *HFE2*: DQ309445, *HCP1*: DQ496103, *TFRC*: DQ496099, *PGRMC2*: DQ496105, *PGRMC1*: DQ496104, *IREB2*: DQ496102, *HEPH*: DQ496100, *HAMP*: DQ496109, *FTH1*: DQ496108, *FLVCR*: DQ496107, *CYBRD1*: DQ496101, and *ACO1*: DQ496106), and (B) the Seattle SNPs database which has data for three genes of interest (*TF*, *HMOX1*, *TNFα*). HapMap has SNP information across the entire human genome, although only data from phase 1 of the HapMap was available at the time of SNP selection. For genes that appeared in more than one of these data resources, sets of tag SNPs were identified separately from each resource and the union of these tag SNPs was selected for genotyping, providing a high level of redundancy to overcome the possibility of assay failure.

We combined these data from Caucasians to select 384 SNPs for further genotyping. The first round of genotyping was conducted using the Illumina Golden Gate platform for 384 SNPs. The second round SNPs were selected for new candidate genes along with some SNPs in high linkage disequilibrium (LD) with significant associations from the first round, as well as repeating five SNPs from the first round. These second round SNP assays were conducted using the Sequenom iPlex platform with 4 multiplexes of 35, 34, 33 and 23 SNPs. SNPs which failed to form good clusters, deviated from Hardy-Weinberg equilibrium (HWE) at p<0.01, were monomorphic, or had less than five heterozygotes were excluded from further analysis.

Statistical methods

Serum ferritin (SF), transferrin saturation (TS), serum transferrin, and serum iron concentrations were analysed as continuously distributed outcomes, with serum ferritin transformed to the natural logarithm scale. Associations between serum iron indices and measured SNP genotypes were determined using separate linear regression models for each SNP, which was included as a three category exposure variable (zero, one, or two copies of the minor allele), along with three other covariates: *HFE* genotype (6 categories: wildtype, C282Y simple heterozygotes, H63D simple heterozygotes, C282Y/H63D compound heterozygotes, H63D homozygotes, C282Y homozygotes), sex and menopause status for women (pre- or post-menopausal at baseline, post-menopausal included those who had a hysterectomy). Age was not included in these analyses since we found, as have others, that after adjusting for women's menopausal status it had little association with iron indices (Adams 2008, Gurrin, *et al* 2008, Koziol, *et al* 2001).

For each candidate SNP, we compared each of the regression models including that candidate SNP with the corresponding model that included only the three other covariates using the likelihood ratio test to determine if that SNP was associated with the outcome (Lettre, *et al* 2007). SNP genotype groups with five or fewer participants were excluded, so many of these tests were effectively comparisons between two predominant genotype groups with exclusion of the minor allele homozygous group. SNPs for which this comparison generated a p-value of less than 0.01, chosen to keep the expected number of false positives associations to less than five for each iron index, were retained for further investigation. Percent of variation in an iron index explained by a given SNP was calculated as the difference in r^2 values from linear regression models with and without that SNP as a covariate; 95% confidence intervals for these quantities were calculated using 1000 bootstrap samples.

We used the Beagle (Browning and Browning 2007) program to test for associations between two binary outcomes (classifying serum ferritin or transferrin saturation as above or below median values specific for groups defined by *HFE* genotype, sex and menopause) and SNP haplotypes (Appendix 2).

Functional CYBRD1 promoter testing

A 486 base pairs fragment of the proximal human *CYBRD1* promoter was amplified from genomic DNA derived from intestinal Caco-2 cells and cloned into the pGL3-basic luciferase reporter. The SNP rs884409 was generated by site-directed mutagenesis as per manufacturer's protocol (QuickChange®, Stratagene, La Jolla, USA), whereby the nucleotide guanine was substituted for thymine at position −326 relative to the transcriptional start site (+1). All constructs were sequenced. Wild-type and mutant reporter constructs were transiently transfected (Fugene 6, Roche) into Caco-2 cells and luciferase activity was measured 48 hours later using the Dual-Luciferase Kit (Promega, Wisconsin, USA) and a luminometer.

Results

Genotyping

We genotyped SNPs in 44 candidate genes (see Table II) previously reported to have a known or hypothesized role in iron homeostasis. DNAs from 863 participants were genotyped: 121 C282Y homozygotes (56 male, 65 female), 148 C282Y/H63D compound heterozygotes (68 male, 80 female), 204 C282Y heterozygotes (96 male, 108 female), 104 H63D heterozygotes (41 male, 63 female), 12 H63D homozygotes (4 male, 8 female), and 274 with neither C282Y nor H63D (122 male, 152 female). These participants were

genotyped for 384 SNPs using GoldenGate (Constantine, *et al* 2008) and 125 SNPs in 4 multiplexes using iPlex. Twenty nine SNPs were excluded due to: ten SNPs in the Golden Gate assay and 10 from iPlex did not form genotype clusters and thus were not able to be scored; two Golden Gate and one iPlex SNP were near monomorphic (less than 5 heterozygous participants) in our sample; three GoldenGate were monomorphic, and three iPlex SNPs deviated from HWE with p<0.01. Four SNPs from round one were successfully repeated in the second round, so a total of 476 unique SNPs were scored. Table II shows the number of first and second round SNPs per gene, and the full SNP list is in Supplemental data 1. Appendix 1 describes some minor genotyping errors discovered during the study affecting a small proportion (3–11%) of individuals for five SNPs (*TFRC* rs9846149; *TNFα* rs1800630; *HMOX1* rs9607267; *TF* rs4241357; *CUBN* rs7094474).

Single SNP association with four outcomes

At the significance level of p<0.01, there were four SNPs associated with serum ferritin, four with transferrin saturation, 24 with serum transferrin, and four with serum iron (Table III, Table IV, Table V, Table VI). Since some of the genotyped SNPs are in strong linkage disequilibrium with other nearby SNPs, for each iron index we looked at SNPs in the same gene with p-values for association with that index of less than 0.01, and grouped those SNPs that were highly correlated with each other (r^2>0.8). When each of these groups was considered to represent a single association, the number of unique associations with TS dropped from four to three and for serum transferrin dropped from 24 to 13. There were no changes in the number of associations for the other two iron indices.

The results from the Beagle analysis (Appendix 2) found 11 associations with permutation p-value <0.2. Four of the seven unique associations found in our regression analysis for TS and SF were also detected using Beagle.

C282Y homozygote-specific modifier

The SNP rs884409 is in the *CYBRD1* promoter, 326 base pairs upstream from the transcription start site. A first round SNP rs3806566 was found to be associated with serum ferritin, examination of data from NHLBI and HapMap show that it is in strong linkage disequilibrium (r^2>0.8) with 17 other variants across the gene. Within this group rs884409 is located in the promoter and therefore likely to be functional and thus was selected for second round typing. Although the summary statistic in Table I suggests an association of rs884409 with SF that is consistent in direction across *HFE* genotype groups, there is evidence that the magnitude of association varies across *HFE*-genotype and sex groups (p=0.0006) with the strongest association for C282Y homozygotes. For C282Y homozygotes the model r^2's were 41.2% (95% CI (25.4%, 57.0%)) and 30.6% (95% CI (13.9%, 47.3%)) with and without the SNP respectively, so this SNP explains 10.6% (95% CI (0.4%, 22.6%)) of the variation in SF.

Functional testing of this promoter polymorphism using a heterologous expression assay found significantly (p = 0.004) decreased promoter activity compared with the more common genotype (Fig 1). A further SNP combination (rs2356782 and rs3731976 which occur together in exon 1 of *CYBRD1* (Lee, *et al* 2002)) was also tested and showed an even greater decrease in promoter activity, but our data revealed no phenotypic association with these two SNPs.

Replication of previous studies

Our results confirm and extend the findings of Benyamin *et al* (Benyamin, *et al* 2009) which showed that *TMPRSS6* rs4820268 was associated with lower transferrin saturation and lower serum iron, and rs3811647 and rs1358024, both in *TF* with r^2=0.6, were associated with

greater serum transferrin (Table II–Table V, Supplemental data 2). By contrast, we did not replicate the previous finding of Milet *et al.* (Milet, *et al* 2007) which reported an association with *BMP2* rs235756 and SF for C282Y homozygotes, (Table I, Supplemental data 2).

Discussion

In this paper we have provided evidence of a novel association between serum ferritin and *CYBRD1* rs884409 using a prospective cohort study in which participants were unselected for either iron levels or disease status. CYBRD1 is a ferric reductase known to contribute to iron uptake (Latunde-Dada, *et al* 2008). Participants with one or two copies of the minor allele for this SNP had lower mean levels of serum ferritin than those with the more common genotype, after adjusting for *HFE*-genotype, sex and menopausal status for women. This association was determined using a regression modelling approach that effectively averaged the magnitude of the associations of this SNP on serum ferritin levels across *HFE*-genotype groups. The association was, however, stronger for C282Y homozygotes than for the other *HFE*-genotype groups. Results from a functional luciferase reporter assay showed that this SNP, which is localized to the proximal region of the *CYBRD1* promoter, resulted in a 30% decrease in its basal activity. This SNP might thus decrease the basal expression of CYBRD1, and lower iron absorption by the gut, therefore protecting against the iron loading that would otherwise occur for C282Y homozygotes as they absorb excess iron.

The discovery of this modifier of the haemochromatosis-related phenotype has potential clinical relevance since serum ferritin is strongly correlated with body iron stores and is used routinely in clinical practice to indicate the extent of iron loading and to determine whether patients should be referred for liver biopsy (typically if SF > 1000 µg/L). Of male C282Y homozygotes with no copies of rs884409, 67% recorded a serum ferritin above 1000 µg/L, while for those with one copy of rs884409 only 21% had serum ferritin levels above 1000 µg/L. While a SNP in high linkage disequilibrium with rs884409 (rs13009270, in our population $r^2=0.93$) is on the Illumina 300K chip used by Benyamin *et al.* (Benyamin, *et al* 2009), the paucity of *HFE* C282Y homozygotes in their sample might explain why this SNP did not appear in their top ten SNPs associated with serum ferritin. The Benyamin *et al.* (Benyamin, *et al* 2009) GWAS study found the strongest associations were between SNPs in the transferrin gene and serum transferrin. Serum transferrin is not usually considered a sensitive marker of iron status in itself, but is combined with measured serum iron to calculate the transferrin saturation. With the exception of variants in the *HFE* gene, the associations reported by Benyamin *et al.* (2009) between genetic variants and transferrin saturation were much weaker than those revealed for serum transferrin.

Our results confirm those of Benyamin *et al.* (Benyamin, *et al* 2009) in two respects. First, we confirmed two of the main novel findings of Benyamin *et al.* (Benyamin, *et al* 2009), namely that rs4820268 in *TMPRSS6* is associated with lower serum iron and transferrin saturation and that two SNPs in *TF* (rs3811647 and rs1358024 with $r^2=0.6$) are associated with increased mean levels of serum transferrin. We have extended their conclusions by showing that these associations occur within *HFE* genotype groups, in particular C282Y homozygotes. The previously reported association of rs1800562 (C282Y) with these four iron indices was also demonstrated by our study. None of the other top ten SNPs of Benyamin *et al.* (Benyamin, *et al* 2009) for their four outcomes (Supplemental Table IV) were tested in this study. Second, we found a number of associations between variants in the transferrin gene and serum transferrin. We took a different approach to determining how many of these were effectively repeated occurrences of a single association due to strong linkage disequilibrium between typed SNPs, by grouping SNPs with pairwise r^2 values

greater than 0.80 rather than invoking a stepwise regression with multiple SNPs. We found nine separate associations between SNPs in *TF* and serum transferrin from 20 associations in total, whereas Benyamin *et al.* (Benyamin, *et al* 2009) found just three from 16 initial associations.

Adjustment for multiple testing is complicated by the relatively high level of linkage disequilibrium that remains for the SNPs we typed. One approach would be to calculate the effective number of independent or unique SNPs using methods based on principal components analysis (PCA) (Gauderman, *et al* 2007, Nyholt 2004), which in our case reduces the number of SNPs from 476 to an effective 229. Bonferroni correction would result in no significant associations (with overall p value of 0.05) for SF, one for TS, three for serum transferrin, and one for serum iron. Screening at p=0.01 for associations between each outcome and the almost 500 SNPs we genotyped, would generate on average $500 \times 0.01 = 5$ associations assuming independent tests even if the global null hypothesis of no association with any SNP was true. The number of putative associations exceeded this benchmark for only one of the iron indices, serum transferrin.

Results from our study did not support the association reported by Milet (Milet, *et al* 2007) between *BMP2* rs235756 and SF. However, the two sample ascertainments and study designs are very different; Milet focussed on unrelated probands, generating a sample that they recognised "is not strictly representative of the population of C282Y homozygotes" and "is rich in individuals with serious symptoms". This suggests that the SNP identified in BMP2 may modify the serum ferritin levels of those C282Y homozygotes who have already progressed to clinical symptoms of disease.

In conclusion, evidence continues to accumulate indicating that SNPs in genes involved in the regulation of iron metabolism potentially modify the biochemical penetrance of the C282Y homozygous phenotype and explaining, at least in part, the large diversity of expression of iron indices for C282Y homozygotes. As further information on candidate genes is generated, these genetic modifiers may result in stratification of people homozygous for C282Y into those who are unlikely to ever progress to serum ferritin levels sufficiently high to cause iron-overload related disease through to those for whom continued monitoring is of benefit until the age of 55–60 years.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Adams PC. Natural history of hemochromatosis: heading down the up escalator? Gastroenterology. 2008; 135:1855–1857. [PubMed: 19000683]

Allen KJ, Gurrin LC, Constantine CC, Osborne NJ, Delatycki MB, Nicoll AJ, McLaren CE, Bahlo M, Nisselle AE, Vulpe CD, Anderson GJ, Southey MC, Giles GG, English DR, Hopper JL, Olynyk JK, Powell LW, Gertig DM. Iron-overload-related disease in HFE hereditary hemochromatosis. New England Journal of Medicine. 2008; 358:221–230. [PubMed: 18199861]

Andriopoulos B Jr, Corradini E, Xia Y, Faasse SA, Chen S, Grgurevic L, Knutson MD, Pietrangelo A, Vukicevic S, Lin HY, Babitt JL. BMP6 is a key endogenous regulator of hepcidin expression and iron metabolism. Nature Genetics. 2009; 41:482–487. [PubMed: 19252486]

Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A, Peltonen L, Martin NG, Montgomery GW, Whitfield JB, Visscher PM. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. American Journal of Human Genetics. 2009; 84:60–65. [PubMed: 19084217]

Browning BL, Browning SR. Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. Genetic Epidemiology. 2007; 31:365–375. [PubMed: 17326099]

Bulaj ZJ, Ajioka RS, Phillips JD, LaSalle BA, Jorde LB, Griffen LM, Edwards CQ, Kushner JP. Disease-Related Conditions in Relatives of Patients with Hemochromatosis. New England Journal of Medicine. 2000; 343:1529–1535. [PubMed: 11087882]

Camaschella C. BMP6 orchestrates iron metabolism. Nature Genetics. 2009; 41:386–388. [PubMed: 19338078]

Constantine CC, Gurrin LC, McLaren CE, Bahlo M, Anderson GJ, Vulpe CD, Forrest SM, Allen KJ, Gertig DM. SNP selection for genes of iron metabolism in a study of genetic modifiers of hemochromatosis. BMC Medical Genetics. 2008; 9:18. [PubMed: 18366708]

Crawford DH, Halliday JW, Summers KM, Bourke MJ, Powell LW. Concordance of iron storage in siblings with genetic hemochromatosis: evidence for a predominantly genetic effect on iron storage. Hepatology. 1993; 17:833–837. [PubMed: 8491452]

Darshan D, Anderson GJ. Interacting signals in the control of hepcidin expression. Biometals. 2009; 22:77–87. [PubMed: 19130266]

Finberg KE, Heeney MM, Campagna DR, Aydinok Y, Pearson HA, Hartman KR, Mayo MM, Samuel SM, Strouse JJ, Markianos K, Andrews NC, Fleming MD. Mutations in TMPRSS6 cause iron-refractory iron deficiency anemia (IRIDA). Nature Genetics. 2008; 40:569–571. [PubMed: 18408718]

Ganz T. Iron homeostasis: fitting the puzzle pieces together. Cell Metabolism. 2008; 7:288–290. [PubMed: 18396134]

Gauderman WJ, Murcray C, Gilliland F, Conti DV. Testing association between disease and multiple SNPs in a candidate gene. Genetic Epidemiology. 2007; 31:383–395. [PubMed: 17410554]

Giles GG, English DR. The Melbourne Collaborative Cohort Study. IARC Scientific Publications. 2002; 156:69–70. [PubMed: 12484128]

Guillem F, Lawson S, Kannengiesser C, Westerman M, Beaumont C, Grandchamp B. Two nonsense mutations in the TMPRSS6 gene in a patient with microcytic anemia and iron deficiency. Blood. 2008; 112:2089–2091. [PubMed: 18596229]

Gurrin LC, Osborne NJ, Constantine CC, McLaren CE, English DR, Gertig DM, Delatycki MB, Southey MC, Hopper JL, Giles GG, Anderson GJ, Olynyk JK, Powell LW, Allen KJ. The natural history of serum iron indices for HFE C282Y homozygosity associated with hereditary hemochromatosis. Gastroenterology. 2008; 135:1945–1952. [PubMed: 18848943]

Koziol JA, Ho NJ, Felitti VJ, Beutler E. Reference centiles for serum ferritin and percentage of transferrin saturation, with application to mutations of the hfe geneq. Clinical Chemistry. 2001; 47:1804–1810. [PubMed: 11568090]

Latunde-Dada GO, Simpson RJ, McKie AT. Duodenal cytochrome B expression stimulates iron uptake by human intestinal epithelial cells. Journal of Nutrition. 2008; 138:991–995. [PubMed: 18492824]

Leboeuf RC, Tolson D, Heinecke JW. Dissociation between tissue iron concentrations and transferrin saturation among inbred mouse strains. Journal of Laboratory and Clinical Medicine. 1995; 126:128–136. [PubMed: 7636385]

Lee P, Gelbart T, West C, Halloran C, Beutler E. Seeking candidate mutations that affect iron homeostasis. Blood Cells Molecules and Diseases. 2002; 29:471–487.

Lettre G, Lange C, Hirschhorn JN. Genetic model testing and statistical power in population-based association studies of quantitative traits. Genetic Epidemiology. 2007; 31:358–362. [PubMed: 17352422]

McCune CA, Ravine D, Carter K, Jackson HA, Hutton D, Hedderich J, Krawczak M, Worwood M. Iron loading and morbidity among relatives of HFE C282Y homozygotes identified either by population genetic testing or presenting as patients. Gut. 2006; 55:554–562. [PubMed: 16174659]

McLaren CE, Barton JC, McLaren GD, Acton RT, Adams PC, Henkin LF, Gordeuk VR, Eckfeldt JH, Vulpe CD, Harrison BW, Reiss JA, Snively BM. Heritability of Iron Status Phenotype in the Hemochromatosis and Iron Overload Screening (HEIRS) Family Study. Blood (ASH Annual Meeting Abstracts). 2008; 112:115.

Melis MA, Cau M, Congiu R, Sole G, Barella S, Cao A, Westerman M, Cazzola M, Galanello R. A mutation in the TMPRSS6 gene, encoding a transmembrane serine protease that suppresses hepcidin production, in familial iron deficiency anemia refractory to oral iron. Haematologica. 2008; 93:1473–1479. [PubMed: 18603562]

Meynard D, Kautz L, Darnaud V, Canonne-Hergaux F, Coppin H, Roth MP. Lack of the bone morphogenetic protein BMP6 induces massive iron overload. Nature Genetics. 2009; 41:478–481. [PubMed: 19252488]

Milet J, Dehais V, Bourgain C, Jouanolle AM, Mosser A, Perrin M, Morcet J, Brissot P, David V, Deugnier Y, Mosser J. Common variants in the BMP2, BMP4, and HJV genes of the hepcidin regulation pathway modulate HFE hemochromatosis penetrance. American Journal of Human Genetics. 2007; 81:799–807. [PubMed: 17847004]

Njajou OT, Alizadeh BZ, Aulchenko Y, Zillikens MC, Pols HAP, Oostra BA, Swinkels DW, van Duijn CM. Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) study. Human Heredity. 2006; 61:222–228. [PubMed: 16877869]

Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. American Journal of Human Genetics. 2004; 74:765–769. [PubMed: 14997420]

Powell LW, Dixon JL, Ramm GA, Purdie DM, Lincoln DJ, Anderson GJ, Subramaniam VN, Hewett DG, Searle JW, Fletcher LM, Crawford DH, Rodgers H, Allen KJ, Cavanaugh JA, Bassett ML. Screening for hemochromatosis in asymptomatic subjects with or without a family history. Archives of Internal Medicine. 2006; 166:294–301. [PubMed: 16476869]

Sproule TJ, Jazwinska EC, Britton RS, Bacon BR, Fleming RE, Sly WS, Roopenian DC. Naturally variant autosomal and sex-linked loci determine the severity of iron overload in beta 2-microglobulin-deficient mice. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:5170–5174. [PubMed: 11309500]

Wallace DF, Subramaniam VN. Non-HFE haemochromatosis. World Journal of Gastroenterology. 2007; 13:4690–4698. [PubMed: 17729390]

Whitfield JB, Cullen LM, Jazwinska EC, Powell LW, Heath AC, Zhu G, Duffy DL, Martin NG. Effects of HFE C282Y and H63D polymorphisms and polygenic background on iron stores in a large community sample of twins. American Journal of Human Genetics. 2000; 66:1246–1258. [PubMed: 10739755]

Whiting PW, Fletcher LM, Dixon JK, Gochee P, Powell LW, Crawford DHG. Concordance of iron indices in homozygote and heterozygote sibling pairs in hemochromatosis families: implications for family screening. Journal of Hepatology. 2002; 37:309–314. [PubMed: 12175625]

Appendix 1 - Genotyping discrepancies

Examination of all pair wise combinations of SNPs within each gene (total 4516 pairs) found three pairs of SNPs with no compound heterozygotes, despite minor allele frequencies >0.05: *TFRC* rs9846149 and TFRC019335; *TNFα* rs1800630 and rs1800610 (genotype counts shown in Table I); *HMOX1* rs9607267 and rs11912889. Further evaluation of the genotype counts suggests the genotyping error (for compound heterozygotes only) occurred in the first SNP listed.

We hypothesise that this is due to an adjacent untyped SNP that is in high linkage disequilibrium with the second of the pair of SNPs, which interferes with the assay to act as a null allele when it occurred on the same chromosome as the tested SNP. This was

supported in two of the three pairs where the minor allele homozygotes of the second SNP failed to be typed for the first SNP (N=3 in *TFRC*; N=5 in *TNFα*). In the second round for *TNFα* we typed a SNP (rs1799724) that is 6 base pairs from rs1800630 and as expected by the interference hypothesis it was in high LD with rs1800610 and showed no compound heterozygotes with rs1800630. For the two pairs with data available, HapMap data obtained from the Genome Variation Server also showed a deficit of compound heterozygotes (rs1800630 and rs1799724 observed 0 expected 6.2, and 7 rs1799724 rare homozygous failed to type for rs1800630; rs9607267 and rs11912889 observed 1, expected 12.5).

Examination of the distribution of the p-values for all SNP pairs reflecting the difference between the observed and expected number of compound heterozygotes showed excesses at each end of this distribution. For p-values above 0.95 the excess is expected due to linkage disequilibrium, however for p-values below 0.05 there was an excess of ~ 400 SNP pairs (9%) which show significant deficiencies of compound heterozygotes.

There were 46 SNPs (43 Golden Gate, 4 iPlex, 1 SNP both) which had some resequencing data (on a maximum of 188 participants) as well as custom genotyping. The highest number of discrepancies was 15/176 comparisons for rs4241357 in *TF*. Inspection of the Golden Gate scatter plot showed these 15 individuals were all within a second heterozygous cluster, however from resequencing they are minor allele homozygotes. The rs7094474 SNP was assayed twice (once by Golden Gate and once by iPlex) and showed a group of 96 discordant genotypes (heterozygotes by Golden Gate and common homozygotes by iPlex). This was sufficient to cause deviation from Hardy-Weinberg equilibrium (HWE) of the iPlex genotypes to fall into our exclusion criteria of p<0.01. The other four described genotyping errors in subgroups did not cause HWE deviation to reach our exclusion criteria.

Overall five SNPs (4/269 Golden Gate and 1/107 iPlex) showed mild discrepancies (involving 24 to 96 samples out of 863, i.e. 3 to 11%).

Appendix 1 Table I

Observed number of genotypes at a pair of TNFα SNPs showing missing compound heterozygotes and failure of the rs1800630 assay for minor allele homozygotes at the rs1800630 locus.

TNFα	rs1800610				Total
rs1800630	Wild type	Heterozygotes	Minor allele homozygotes	missing	
Wild type	536	96	0	0	632
Heterozygotes	193	0	0	3	196
Minor allele homozygotes	19	11	0	0	30
missing	0	0	5	0	5
Total	748	107	5	3	863

Appendix 2 – Beagle analysis

Haplotype analysis

We used the Beagle software (Browning and Browning 2007), which implements a graphical model based on variable length Markov chains to represent linkage disequilibrium between SNPs, to test for association between a binary outcome and SNP haplotypes.

Currently Beagle accepts only binary outcome measures, so we defined a binary outcome for SF and TS by comparing the observed iron index to a subgroup specific median iron index, with subgroups defined by the same three covariates used in the regression modelling above, namely *HFE*-genotype, sex and menopausal status. The settings used were scale 3 and shift 0.2.

The analysis using Beagle supported three of the seven of single SNP associations from regression models described in the text (one for serum ferritin, three for transferrin saturation) by generating p-values of less than 0.20 based on permutation tests (Table I).

Appendix 2 Table I

Association of SNP haplotypes with a binary outcome classifying serum ferritin or transferrin saturation as above or below median values specific for groups defined by *HFE* genotype, sex and menopause (in women).

Outcome	Gene	SNP	p-valu	Model	permutation p-value	Other SNPs in high LD $r^2>0.8$	Single SNP associations from regression models
TS	HCP1	rs2239908	0.00073	Recessive	0.020	rs2239907	
TS	HCP1	hcp1000942	0.00940	Allelic	0.139	rs9905973	
TS	CUBN	rs11254389	0.00172	Recessive	0.199		yes
TS[*]	EXOC6	rs787625	0.00453	Overdominant	0.069		
TS[*]	SLC11A2	rs427020	0.00395	Recessive	0.078		
TS	TMPRSS6	rs48020268	0.00338	Allelic	0.126		yes
TS	TMPRSS6	rs5756506	0.00448	Recessive	0.161		yes
SF	CYBRD1	rs884409	0.00030	Allelic	0.027	rs3806566 rs12621138 rs13009270 rs17288294	yes
SF[*]	HEPHL1	rs1518563	0.00763	Dominant	0.180		
SF	HFE2	rs16827043	0.02639	Dominant	0.140		
SF	HFE2	hfe2006712	0.03137	Allelic	0.192		

[*] These represent haplotype associations rather than single SNP associations.
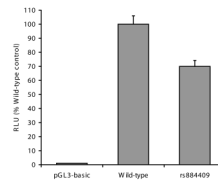
Figure 1.
Promoter activity without (wild-type) and with rs884409 SNP. Wild-type or mutant rs884409 SNP luciferase reporter constructs were transiently transfected into Caco-2 cells and relative luciferase activity (RLU) was measured 48 hours later. rs884409 SNP is localized to nucleotide −326 relative to the TSS-1 and had a significant negative association with basal promoter activity (p=0.0043). Data are representative of two independent experiments measured in triplicate, presented as means ±SD and analyzed using 1-way ANOVA and Tukey's post hoc test. Data is presented as % of Wild-type promoter activity set as 100.

Table I

Mean serum ferritin by *CYBRD1* rs884409 genotype for six subgroups based on sex, menopause status and *HFE* C282Y homozygous or not.

Copies of rs884409	Serum ferritin (µg/L) geometric mean and 95% CI		
	No copies	One copy	Two copies
C282Y homozygous men	1047 (775,1415) n=27	315 (134, 739) n=16	
C282Y homozygous women post-menopause	324 (223, 469) n=28	150 (30, 753) n=6	
C282Y homozygous women pre-menopause	100 (45, 221) n=14	24 (5, 118) n=6	8 n=1
non-C282Y homozygous men	177 (156, 202) n=201	183 (146, 230) n=72	152 (68, 343) n=7
non-C282Y homozygous women post-menopause	106 (89, 125) n=114	89 (68, 117) n=48	83 (34, 201) n=4
non-C282Y homozygous women pre-menopause	43 (35, 52) n=111	34 (25, 46) n=39	27 (17, 44) n=10

Table II

Number of SNPs successfully genotyped per gene, the complete list of SNPs in Supplemental data 1.

	First round	Second round		First round	Second round		First round	Second round
ACO1	15	-	GAST	2	-	PLEKHB2	-	2
BMP2	-	26	GDF15	-	4	SLC11A2	13	-
BMP4	-	17	GSTP1	1	-	SLC40A1	12	-
BTBD9	-	1	HAMP	5	-	SMAD1	-	6
CALR	1	-	HCP1	5	-	SMAD4	-	10
CD163	14	-	HEPH	17	-	SMAD5	-	4
CP	25	-	HEPHL1	9	-	STEAP3[a]	10	-
CUBN[a]	29	-	HFE	6	1	TF	43	2
CYBRD1	26	8	HFE2	4	1	TFR2	4	1
DHCR7	5	-	HMOX1	10	1	TFRC	29	1
EXOC6[a]	3	-	HMOX2	1	1	TMPRSS6	-	8
FLVCR	16	2	HP	8	-	TNF	2	8
Frascati	19	-	IREB2	17	-	Total	369	107
FTH1	7	-	MON1A	-	2			
FTL	1	-	PGRMC1	1	1			
FXN	6	-	PGRMC2	3	-			

[a]There are known HapMap phase 1 SNPs in these genes which were not genotyped, nor were they in strong linkage disequilibrium ($r^2 > 0.8$) with any of the typed SNPs, i.e. tagSNP coverage was incomplete for these genes.

Table III

Single SNPs associated with serum ferritin in a regression model adjusting for sex, menopause and *HFE* genotype. Difference from mean (estimated regression coefficient) for common homozygotes is shown for heterozygotes and minor allele homozygotes (where N>5).

SNP	Gene	Minor allele frequency	Difference heterozygote	Difference rare homozygote	Likelihood ratio test p-value
rs7294582	*CD163*	0.101	0.26 ±0.10	−0.78 ±0.37	0.0021
rs1284859[a]	*FLVCR*	0.057	−0.34 ±0.13	-	0.0075
rs884409	*CYBRD1*	0.165	−0.24 ±0.09	−0.40 ±0.23	0.0080

[a] r^2 between rs1284859 and rs931591 is 0.99

Table IV

Single SNPs associated with transferrin saturation in a regression model adjusting for sex, menopause and *HFE* genotype. Difference from mean (estimated regression coefficient) for common homozygotes is shown for heterozygotes and minor allele homozygotes (where N>5).

SNP	Gene	Minor allele frequency	Difference heterozygote	Difference rare homozygote	Likelihood ratio test p-value
rs4820268[a] D521D	*TMPRSS6*	0.462	−3.98 ±1.12	−5.71 ±1.39	0.0001
rs11254389	*CUBN*	0.167	−2.39 ±1.07	7.18 ±2.77	0.0014
rs5756506[a]	*TMPRSS6*	0.372	0.52 ±1.04	4.71 ±1.49	0.0032
rs1049296 S589P	*TF*	0.163	−0.21 ±1.08	8.65 ±2.84	0.0083

[a] r^2 between rs5756506 and rs4820268 is −0.70

Table V

Single SNPs associated with serum transferrin in a regression model adjusting for sex, menopause and *HFE* genotype. Difference from mean (estimated regression coefficient) for common homozygotes is shown for heterozygotes and minor allele homozygotes (if N>5).

SNP (Group with r^2>0.8) [a]	Gene	Minor allele frequency	Difference heterozygote	Difference rare homozygote	Likelihood ratio test p-value
rs1880669 [a]	TF	0.401	−0.10 ± 0.03	−0.21 ± 0.04	<0.0001
rs3811647 [b]	TF	0.321	0.12 ± 0.03	0.33 ± 0.05	<0.0001
rs1358024 [c]	TF	0.154	0.09 ± 0.03	0.36 ± 0.08	<0.0001
rs1405023 [d]	TF	0.477	−0.08 ± 0.03	−0.15 ± 0.04	0.0004
rs6488340	CD163	0.084	−0.00 ± 0.04	0.54 ± 0.15	0.0014
rs12493168	TF	0.142	0.08 ± 0.03	0.22 ± 0.08	0.0021
rs1799852 L247L	TF	0.099	−0.11 ± 0.04	-	0.0023
rs8177326	TF	0.022	0.21 ± 0.07	-	0.0024
rs8177215 [e]	TF	0.052	−0.13 ± 0.05	-	0.0039
rs11254389	CUBN	0.167	0.10 ± 0.03	−0.03 ± 0.08	0.0040
rs838102	STEAP3	0.441	−0.10 ±0.03	−0.04 ± 0.04	0.0056
rs1045537	HFE	0.066	0.12 ± 0.04	−0.17 ± 0.14	0.0083
rs1049296 S589P	TF	0.163	−0.08 ± 0.03	−0.17 ± 0.08	0.0093

[a]For groups of SNPs with r^2>0.8, only one SNP per group is shown in the table:

[a]rs1880669, rs2692695.

[b]rs3811647, rs3811658, rs8177240, rs8177260, rs8177178, rs4459901, rs3811656.

[c]rs1358024, rs8177297.

[d]rs8177237, rs1405023.

[e]rs8177185, rs8177215, rs8177235.

Table VI

Single SNPs associated with serum iron in a regression model adjusting for sex, menopause and *HFE* genotype. Difference from mean (estimated regression coefficient) for common homozygotes is shown for heterozygotes and minor allele homozygotes (where N>5).

SNP	Gene	Minor allele frequency	Difference heterozygote	Difference rare homozygote	Likelihood ratio test p-value
rs4820268 [a] D521D	*TMPRSS6*	0.462	−1.89 ±0.60	−3.10 ±0.75	0.0001
rs5756506 [a]	*TMPRSS6*	0.372	0.02 ±0.56	2.53 ±0.80	0.0028
rs7094474	*CUBN*	0.426	0.24 ±0.57	−1.92 ±0.75	0.0071
rs7219746	*GAST*	0.422	−0.14 ±0.57	−2.11 ±0.74	0.0086

[a] r^2 between rs5756506 and rs4820268 was −0.70