CYBRD1 as a modifier gene that modulates iron phenotype in *HFE* p.C282Y homozygous patients

Sara Pelucchi,[1] Raffaella Mariani,[2] Stefano Calza,[3,4] Anna Ludovica Fracanzani,[5] Giulia Litta Modignani,[1] Francesca Bertola,[6] Fabiana Busti,[7] Paola Trombini,[2] Mirella Fraquelli,[8] Gian Luca Forni,[9] Domenico Girelli,[7] Silvia Fargion,[5] Claudia Specchia,[3] and Alberto Piperno[1,2,6]

[1]Department of Clinical Medicine and Prevention, University of Milano-Bicocca, Monza, Italy; [2]Center for Diagnosis and Treatment of Hemochromatosis, S. Gerardo Hospital, Monza, Italy; [3]Department of Biomedical Sciences and Biotechnologies, University of Brescia, Brescia, Italy; [4]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [5]Department of Medicine, IRCCS, Ospedale Maggiore Policlinico, University of Milano, Milano, Italy; [6]Consortium for Human Molecular Genetics, Monza, Italy; [7]Department of Medicine Policlinico GB Rossi, University of Verona, Verona, Italy; [8]Second Division of Gastroenterology, IRCCS, Ospedale Maggiore Policlinico, University of Milano, Milano, Italy; and [9]Centre of Microcitemia and Congenital Anemias, Galliera Hospital, Genova, Italy

ABSTRACT

Background
Most patients with hereditary hemochromatosis in the Caucasian population are homozygous for the p.C282Y mutation in the *HFE* gene. The penetrance and expression of hereditary hemochromatosis differ largely among cases of homozygous p.C282Y. Genetic factors might be involved in addition to environmental factors.

Design and Methods
In the present study, we analyzed 50 candidate genes involved in iron metabolism and evaluated the association between 214 single nucleotide polymorphisms in these genes and three phenotypic outcomes of iron overload (serum ferritin, iron removed and transferrin saturation) in a large group of 296 p.C282Y homozygous Italians. Polymorphisms were tested for genetic association with each single outcome using linear regression models adjusted for age, sex and alcohol consumption.

Results
We found a series of 17 genetic variants located in different genes with possible additive effects on the studied outcomes. In order to evaluate whether the selected polymorphisms could provide a predictive signature for adverse phenotype, we re-evaluated data by dividing patients in two extreme phenotype classes based on the three phenotypic outcomes. We found that only a small improvement in prediction could be achieved by adding genetic information to clinical data. Among the selected polymorphisms, a significant association was observed between rs3806562, located in the 5'UTR of *CYBRD1*, and transferrin saturation. This variant belongs to the same haplotype block that contains the *CYBRD1* polymorphism rs884409, found to be associated with serum ferritin in another population of p.C282Y homozygotes, and able to modulate promoter activity. A luciferase assay indicated that rs3806562 does not have a significant functional role, suggesting that it is a genetic marker linked to the putative genetic modifier rs884409.

Conclusions
While our results support the hypothesis that polymorphisms in genes regulating iron metabolism may modulate penetrance of *HFE*-hereditary hemochromatosis, with emphasis on *CYBRD1*, they strengthen the notion that none of these polymorphisms alone is a major modifier of the phenotype of hereditary hemochromatosis.

Key words: hemochromatosis, SNP, iron, gene, ferritin, transferrin saturation.

Citation: Pelucchi S, Mariani R, Calza S, Fracanzani AL, Modignani GL, Bertola F, Busti F, Trombini P, Fraquelli M, Forni GL, Girelli D, Fargion S, Specchia C, and Piperno A. CYBRD1 *as a modifier gene that modulates iron phenotype in* HFE *p.C282Y homozygous patients. Haematologica 2012;97(12):1818-1825. doi:10.3324/haematol.2012.062661*

Funding: this work was supported by Progetto Quadro Regione Lombardia funding (CUP D41J10000390002) to AP. SP is supported by a research fellowship from the University of Milano-Bicocca and by the Associazione per lo Studio dell'Emocromatosi +Fe – ONLUS, Monza, Italy.

Manuscript received on January 23, 2012. Revised version arrived on July 3, 2012. Manuscript accepted on July 3, 2012.

Correspondence:
Alberto Piperno, MD, Department of Clinical Medicine and Prevention, Centre for Diagnosis and Treatment of Hemochromatosis S.Gerardo Hospital, via Pergolesi 33, 20900, Monza, Italy.
Phone: international +39.039. 2339710.
E-mail: alberto.piperno@unimib.it

The online version of this article has a Supplementary Appendix.

Introduction

Hereditary hemochromatosis is a heterogeneous disorder at both genetic and phenotypic levels. *HFE*-hereditary hemochromatosis is the most common form in Caucasian populations and homozygosity for p.C282Y mutation is the genotype most frequently associated with iron overload and related complications.[1] *HFE*-hereditary hemochromatosis may lead to accumulation of iron in tissues and iron-related complications, but penetrance and expression vary greatly among p.C282Y homozygotes. Environmental and genetic factors have been implicated: blood loss, alcohol intake, coexistence of chronic hepatitis B and C, and non-alcoholic fatty liver disease can influence clinical manifestations in humans.[1,2] Animal studies showed that the genetic background modulates the expression of the disease[3] and identified several candidate modifier regions in *HFE*-knock-out mice.[4] In humans, a high concordance of iron indices and/or iron-related disease has been found among related patients with *HFE*-hereditary hemochromatosis, supporting the existence of genetic modifiers influencing phenotype expression.[5,6] Association studies between genetic markers and disease phenotype have given conflicting results.[7-14] More recently, candidate gene studies revealed significant associations between single nucleotide polymorphisms (SNPs) in genes involved in iron metabolism and indices of iron overload in p.C282Y homozygotes. Milet *et al.* focused on two biologically relevant gene categories: genes involved in non-*HFE*-hereditary hemochromatosis (*TFR2*, *HAMP*, and *SLC40A1*) and genes involved in the regulation of hepcidin expression (*BMP2*, *BMP4*, *HJV*, *SMAD1*, *4* and *5*, and *IL6*). They found an association between the SNP rs235756 of the bone morphogenetic protein-2 (*BMP2*) gene and serum ferritin in a large series of French patients.[15] Constantine *et al.* found that the SNP rs884409 in *CYBRD1* was a possible novel modifier specific to *HFE*-hereditary hemochromatosis, but were unable to confirm the association with the *BMP2* rs235756[16]. Moreover, genome-wide association studies performed in the general population in recent years have shown associations between SNPs of transmembrane protease, serine 6 (*TMPRSS6*) and serum iron[17,18] and transferrin saturation,[19] suggesting a relevant involvement of TMPRSS6 in the control of iron homeostasis. In the present study, we evaluated the association between several SNPs in genes involved in iron metabolism and three phenotypic outcomes of iron overload (serum ferritin, iron removed and transferrin saturation) in a group of p.C282Y homozygous Italians under the assumption that these SNPs may act as modifiers of their iron phenotype.

Design and Methods

Subjects

We enrolled a group of 306 unrelated patients with *HFE* in the study. All these patients were p.C282Y homozygotes attending four centers in northern Italy (Milan, Verona, Genoa and Monza). There was no selection based on disease severity. From the whole database of each center, patients were selected based on the following inclusion and exclusion criteria. Inclusion criteria were: p.C282Y homozygosity, availability of data on serum ferritin and transferrin saturation before iron depletion, good quality DNA, and information about age, sex and alcohol intake. Patients with a previous history of regular blood donations were excluded from the study. Data on total iron removed were available for 211 patients. A group of 114 healthy controls were recruited among blood donors from the same geographic area of the patients only to further validate the quality of the genotype data (calculation of the Hardy-Weinberg equilibrium to check for bias and mistakes in genotyping). Patients and controls gave their informed consent to participation in this study. The Lombardy Region and University of Milano-Bicocca research fellowship committees approved the study.

Iron indices

Serum ferritin levels, total iron removed and percentage of transferrin saturation were used as markers of the expression of hereditary hemochromatosis among patients. Serum ferritin and transferrin saturation were measured at the time of diagnosis before phlebotomy therapy; total iron removed was calculated based on the number of phlebotomies performed to achieve iron depletion, as previously reported.[20]

Extraction of DNA from blood

Blood samples for DNA extraction were collected from all subjects into tubes containing EDTA. Genomic DNA was extracted from the whole blood of each subject using the Wizard® Genomic DNA Purification kit (Promega, Madison, WI, USA), and stored at -20°C before use. DNA samples were adjusted to a concentration of 50 ng/μL.

Selection of single nucleotide polymorphisms

Two hundred and fourteen TagSNPs within 50 candidate genes were analyzed. Two hundred and eleven were selected by Haploview Tagger[21] for the CEU population to be screened with a custom-designed 384-plex VeraCode GoldenGate genotyping assay on an Illumina BeadXpress Reader platform (Illumina Inc. San Diego, CA, USA). Three (rs855791 in *TMPRSS6*, rs884409 and rs3806566 in *CYBRD1*) were selected *a posteriori* based on previous results showing significant correlations with iron status in the general population[18] and in patients with *HFE*-hereditary hemochromatosis,[16] respectively, and analyzed by direct sequencing. More information on the SNP selection criteria is reported in the *Online Supplementary Design and Methods* section, and the whole list of SNPs analyzed is given in the *Online Supplementary Table S1*.

Genotyping of single nucleotide polymorphisms

SNP genotyping was performed using the GoldenGate Genotyping assay on an Illumina BeadXpress Reader platform according to the manufacturer's protocol. GenomeStudio software was used to call genotypes (see *Online Supplementary Design and Methods* section). Illumina results were further validated by sequencing 100 samples for 12 SNPs on an ABI Prism 3130 Avant Automatic Sequencer (Applied Biosystems, Foster City, CA, USA).

Dual-luciferase reporter assay

DNA fragments (415 bp) of the *CYBRD1* gene were obtained from genomic DNA of two individuals, each one carrying the TT and CC genotype of the SNP rs3806562, by polymerase chain reaction (forward primer: 5'-ggCTCgAgggCTggACCAgATCAAAgAA-3'; reverse primer: 5'-gggATATCgCCTgCCCTCTTTCCAACATT-3'). The polymerase chain reaction products, which did not include either rs3806566 or rs884409, were cloned into the pGL4.13 plasmid vector (Promega Corp., Madison, WI, USA) upstream of the firefly luciferase gene, by digestion with XhoI and EcoRV. Plasmid constructs were verified by direct sequencing. Plasmid DNA was isolated by Pure Yield™ Plasmid Miniprep System kit (Promega Corp., Madison,

WI, USA) for transfection. The recombinant plasmids were co-transfected with pGL4.74 plasmid (carrying the hRluc/Tk reporter gene as an expression control) into hepatoma carcinoma cells (Huh-7) by Lipofectamine 2000 (Life-Technologies Corp., Carlsbad, CA, USA). The pGL4.13 basic plasmid was also co-transfected with pGL4.74 as a negative control. After incubation for 48 h, cells were lysed and firefly and renilla luciferase activities were measured by a Glomax Multi JR luminometer according to manufacturer's protocols (Promega Corp., Madison, WI, USA). Each construct was tested in triplicate, and the transfection experiments were performed three times independently. Data are expressed as mean ± standard deviation. Luciferase activities were compared by the Mann-Whitney test, using Prism 3.2 software (GraphPad Software, San Diego, CA, USA).

Statistical methods

Stringent quality control criteria were applied to all samples and genotype data. A per-SNP genotype rate threshold of 95% was used. Identity-by-state values were calculated for pairs of subjects to identify duplicates or possibly related subjects. For any pair with more than 95% identical SNP genotypes, the sample with the lower call rate was removed from the analysis. SNPs with a minor allele frequency less than 0.6% were also removed. To check for genotyping errors, the Hardy-Weinberg equilibrium for each SNP was tested among the controls with Pearson's χ^2 test statistic. SNPs that deviated from the Hardy-Weinberg equilibrium were excluded from the analysis. Iron indices were analyzed as continuously distributed outcomes and normalized using a log transformation. A linear regression model was fitted to evaluate the effect on the three outcomes of age, sex and alcohol consumption (g/day). Each SNP was tested for association with each single outcome by using separate linear regression models adjusted for age, sex and alcohol consumption. Genotypic association was considered and an additive genetic model was assumed. SNPs were ranked according to their uncorrected *P*-value and top ranked SNPs for each outcome were defined as those with $P<0.05$. The false discovery rate, as computed by the q value, was applied to adjust for multiple comparisons, using a threshold rate of <0.2. Interactions between pairs of top ranked SNPs with each outcome were evaluated by adding product terms to a multiple regression linear model, adjusted for age, sex and alcohol consumption. Multiple comparison adjustment was ignored while assessing the significance of the interaction term. In order to evaluate whether SNPs could provide a predictive signature for adverse phenotype, we re-evaluated data by dividing patients with *HFE*-hereditary hemochromatosis into two extreme iron-related phenotype classes. Identification of an extreme adverse phenotype, based on ferritin, iron removed and transferrin saturation values, was performed using principal component analysis, based on a correlation matrix, fitted on 209 patients with non-missing values for all three hematologic parameters. The component retaining the highest proportion of variance (principal component analysis first component) was used as a pseudo-marker and cut into tertiles, using the first and third ones to define an extreme binary phenotype. Top ranked SNPs were included in classification procedures using four different algorithms; support vector machine (SVM[22]), random forest (RF[22]), ridge penalized logistic regression (PEN[22]) and K-nearest neighbors (KNN[22]). The classification performances of all algorithms were evaluated using the area under the receiver operating characteristic curve (AUC). The AUC for each procedure was computed using class probability i.e. the estimated probability of being a "case" (i.e. of being in the highest tertile of the combined hematologic parameters estimated via leave-one-out cross validation (LOOCV[22]). Class probability was estimated on the out of the bag sample to reduce the bias of evaluating a classification model on the same data used to build it. In order to estimate the best number of SNPs to use, the cross-validation procedure was repeated for varying number of SNPs. For every run of the LOOCV algorithm, all SNPs are ranked according to the size of the odds ratio estimated by a logistic model accounting for the SNP itself, age, sex and alcohol consumption. A classification model was then built for increasing number of SNPs, from a model with only one SNP (the most associated) to a model with all SNPs. The optimal number of SNPs was then selected as the one with the highest AUC. All methods require some fine tuning of parameters (e.g. the k value in KNN). This was performed with cross-validation using the whole set of SNPs. Tests for the equality of the AUC were performed based on the method described by Delong *et al.*[23] while model prediction improvement was evaluated as suggested by Pencina *et al.*[24] More details are given in the *Online Supplementary Design and Methods*. The analyses were performed using R and the library GenABEL (genome-wide SNP association analysis R package version 1.7)[25].

Results

After imposing the quality control measures, 22 subjects (10 *HFE* patients and 12 healthy controls) and 10 SNPs were excluded from the analysis because of low call rate and high frequency of SNPs that were identical-by-state. In addition, sequence analysis of the SNPs rs3806566 and rs884409 in the 5'UTR of *CYBRD1*, which were recently found to be significantly associated with serum ferritin in patients with *HFE*-hereditary hemochromatosis,[16] confirmed that they are in complete linkage with rs3806562, the SNPs we previously selected and validated. For this reason, only the latter SNP was retained for statistical analysis. A total of 296 p.C282Y homozygous patients (220 men, 76 women) and 102 healthy controls (86 men, 16 women) for 202 SNPs were finally considered. None of these SNPs deviated from the Hardy-Weinberg equilibrium. The patients' demographic data, alcohol intake, hemoglobin and iron indices are reported in Table 1. Age and sex were significantly associated with serum ferritin, total iron removed and transferrin saturation ($P<0.001$), while alcohol daily intake was significantly associated with serum ferritin and total iron removed ($P<0.001$).

Seventeen SNPs were associated with the iron indices among patients (top ranked SNPs, uncorrected $P<0.05$). Their allele frequencies were compared to those in the CEU population (Table 2) and resulted similar with the exception of the *HFE* SNP.

For each single outcome, the top ranked associated SNPs with their location within the gene and the worst allele in terms of more severe phenotype are reported in Table 3.

Table 1. Median and (range) of age, alcohol intake, hemoglobin and serum iron indices in p.C282Y homozygotes divided by sex.

	Missing (%)	All patients (N=296)	Males (N=220)	Females (N=76)
Age (years)	0	45.5 (11-77)	43 (11-76)	56 (21-77)
Alcohol intake (g/day)	0	10 (0-250)	10 (0-250)	5 (0-60)
Hemoglobin (g/dL)	22.6	14.8 (9.0-18.7)	15.0 (9.0-18.7)	13 (9.4-16.9)
Transferrin saturation (%)	2.4	85 (30-100)	86 (41-100)	80 (30-100)
Serum ferritin (μg/L)	0	1060 (31-13136)	1209 (32-13136)	552.5 (31-5089)
Iron removed (g)	28.7	7 (0.5-41)	8 (1.6-41)	4.4 (0.5-25)

Serum ferritin was associated with rs12467409, rs9325886 and rs17804636, belonging to candidate genes *BMPR2*, *BMP9* and *SMAD8*, respectively. Total iron removed was associated with rs3178250, rs762642, rs12467409, rs11204215, rs1800702, rs149411, rs11684885 and rs3780474, belonging to candidate genes *BMP2*, *BMP4*, *BMPR2*, *BMP9*, *HFE*, *DMT1*, *HIF2A* and *IRP1*, respectively. Transferrin saturation was associated with rs4401458, rs2292915, rs701753, rs773050, rs701754, rs17554 and rs3806562, belonging to candidate genes *BMPR1B*, *NEO1*, *CP* and *CYBRD1*, respectively. The effect of the top ranked SNP is expressed in terms of fold-change induced on the estimated value of the outcome by adding a single worst allele and after adjusting for age, sex and alcohol consumption. After correction for multiple comparisons, a significant association, adjusted for age, sex and alcohol consumption, was detected between a SNP located in *CYBRD1*, rs3806562, and transferrin saturation (uncorrected $P<0.001$, false discovery rate=0.07). Observed mean transferrin saturation levels were 0.93 among TT genotypes, 0.90 among TC genotype and 0.85 among CC genotypes.

In vitro functional assays of rs3806562 did not show significant differences between constructs carrying the C or the T allele (CC: 1.26±0.28, TT: 2.11±1.48 Relative Luciferase Unit, P=NS).

We found two suggestive but not significant interactions between pairs of the 17 top ranked SNPs, one associated with total iron removed (rs149411 and rs762642, uncorrected P=0.027) and one associated with transferrin saturation (rs773050 and rs4401458, uncorrected P=0.02). Figure 1 shows interaction plots of the estimated levels of these indices as a function of SNP genotypes.

In order to evaluate whether SNPs could provide a predictive signature for adverse phenotype, we re-evaluated data using principal component analysis as described in the *Design and Methods* section. The principal component analysis first component (hereafter called 'pseudo-marker') accounted for 62.4% of the total variance of iron indices and reflected the overall iron load (*loadings*: serum ferritin=0.64, transferrin saturation=0.43, total iron removed=0.64). High levels of the three indices, therefore, resulted in high values of the pseudo-marker and *vice versa*. Based on the first and third tertiles of the pseudo-marker we categorized patients as controls (mild phenotype) or cases (adverse phenotype). In this way, the sample was

Table 2. Allelic frequencies in p.C282Y homozygotes and the CEU population of the 17 SNPs associated with the outcomes (top ranked SNPs, uncorrected $P<0.05$).

SNP	CEU allelic frequencies	Allelic frequencies in p.C282Y homozygotes
rs17554 (G/A)	G: 0.55; A: 0.45	G: 0.67; A: 0.33
rs149411 (C/T)	C: 0.58; T: 0.42	C: 0.63; T: 0.37
rs701753 (A/T)	A: 0.92 ; T: 0.08	A: 0.94 ; T: 0.06
rs701754 (A/T)	A: 0.85 ; T: 0.15	A: 0.84 ; T: 0.16
rs762642 (T/G)	T: 0.61; G: 0.39	T: 0.56; G: 0.44
rs1800702 (C/G)*	C: 0.61; G:0.39	C: 0.01; G:0.99
rs2292915 (C/T)	C: 0.66; T: 0.34	C: 0.65; T: 0.35
rs3178250 (T/C)	T: 0.81; C: 0.19	T: 0.72; C: 0.28
rs3780474 (A/C)	A: 0.64; C:0.36	A: 0.64; C:0.36
rs3806562 (T/C)	T: 0.85; C: 0.15	T: 0.84; C: 0.16
rs4401458 (T/C)	T: 0.52; C: 0.48	T: 0.46; C: 0.54
rs9325886 (C/A)	C: 0.93; A: 0.07	C: 0.94; A: 0.06
rs11204215 (A/C)	A: 0.87; C: 0.13	A: 0.83; C: 0.17
rs11684885 (G/T)	G: 0.66; T: 0.34	G: 0.62; T: 0.38
rs12467409 (G/T)	G: 0.86; T: 0.14	G: 0.87; T: 0.13
rs17804636 (A/G)	A: 0.92; G: 0.08	A: 0.93; G: 0.07
rs773050 (A/G)	A: 0.94; G: 0.06	A: 0.94; G: 0.06

**HFE SNP*

Table 3. SNPs associated with variability of serum ferritin, iron removed and transferrin saturation.

Outcome	SNP	Gene symbol	Location	Worst allele	N.	Fold change* (95% CI)	*P* value°
Serum ferritin	rs12467409 (G/T)	*BMPR2*	Intron	T	293	1.26 (1.02-1.58)	0.036
	rs9325886 (C/A)	*BMP9*	3'UTR	C	296	1.39 (1.04-1.86)	0.026
	rs17804636 (A/G)	*SMAD8*	3'UTR	G	296	1.43 (1.06-1.92)	0.019
Iron removed	rs3178250 (T/C)	*BMP2*	3'UTR	T	211	1.19 (1.02 -1.37)	0.023
	rs762642 (T/G)	*BMP4*	Intron	G	210	1.16 (1.02 -1.33)	0.030
	rs12467409 (G/T)	*BMPR2*	Intron	T	208	1.30 (1.03 -1.64)	0.026
	rs11204215 (A/C)	*BMP9*	5'UTR	A	211	1.22 (1.02 -1.46)	0.028
	rs1800702 (C/G)	*HFE*	5'UTR	C	211	1.56 (1.01-2.38)	0.043
	rs149411 (C/T)	*DMT1*	3'UTR	T	211	1.18 (1.02 -1.35)	0.025
	rs11684885 (G/T)	*HIF2A*	Intron	A	211	1.16 (1.01 - 1.33)	0.041
	rs3780474 (A/C)	*IRP1*	Intron	C	211	1.21 (1.06-1.38)	0.005
Transferrin saturation	rs4401458 (T/C)	*BMPR1B*	Intron	T	289	1.23 (1.02 - 1.49)	0.031
	rs2292915 (C/T)	*NEO1*	Intron	C	287	1.21 (1.00-1.47)	0.046
	rs701753 (A/T)	*CP*	Coding	A	288	1.56 (1.06-2.30)	0.023
	rs773050 (A/G)	*CP*	Intron	G	289	1.45 (1.002-2.09)	0.048
	rs701754 (A/T)	*CP*	Intron	A	288	1.32 (1.04-1.67)	0.022
	rs17554 (G/A)	*CYBRD1*	Intron	G	289	1.23 (1.02-1.48)	0.034
	rs3806562 (T/C)	*CYBRD1*	5'UTR	T	288	1.54 (1.21-1.96)	<0.001

**Adjusted for age, gender and alcohol intake in a multiple linear regression model; ° P value uncorrected for multiple testing.*

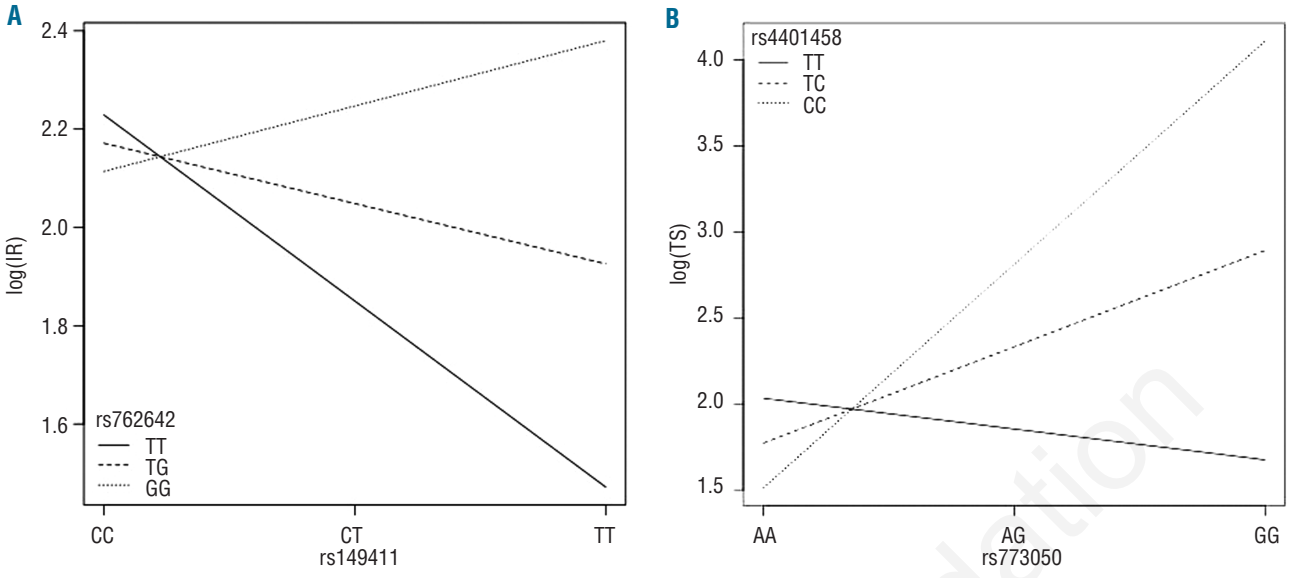

Figure 1. Boxplots of the estimated total iron removed (IR) levels (log scale) according to rs149411 (*DMT1*) and rs762642 (*BMP4*) genotypes (A) and of the estimated transferrin saturation (TS) levels (logit scale) according to rs773050 (*CP*) and rs4401458 (*BMPR1B*) genotypes (B).

reduced to 140 patients (70 controls and 70 cases). The average values of serum ferritin in the two groups were 619 μg/L in controls and 3718 μg/L in cases, with 85.5% of cases and no controls having values 2000 μg/L and 91.3% of controls and only 1.5% of cases having values below 1000 μg/L. The average levels for total iron removed were 3.73 g and 16.88 g in controls and cases, respectively, while the average percentages for transferrin saturation were 67.2% in controls and 89.8% in cases. In order to evaluate whether the top ranked SNPs could provide a predictive signature for adverse phenotype, we performed a classification procedure based on four different algorithms. For the classification algorithms we had to remove subjects with at least one missing value in any of the 17 selected SNPs. Overall we excluded two patients and, therefore, performed the classification procedure on 69 controls and 69 cases. All models performed reasonably well with an AUC greater than 65%. The best model was the random forest one with an AUC of 76.8% using 14 SNPs (KNN 67.6% with 6 SNPs; PEN 67.0% with 16 SNPs; SVM 71.9% with 15 SNP). A predictive model based on logistic regression including only age, sex and alcohol consumption was also fitted. The resulting AUC (70.0%) was lower than the best model with genetic effect, but the difference was not statistically significant (*P* value of the test for equality of AUC = 0.12). We then evaluated the change in predicted probability using the additional information from SNPs and we found that it increased the estimated probability of being an event among cases (sensitivity) to 2.97%, while reducing it among controls (specificity) to 3%. To combine the latter two quantities we computed the integrated discrimination improvement which was close to, but still did not reach, statistical significance (integrated discrimination improvement, *P*-value = 0.06). These results suggest that the 17 selected SNPs provide little additional predictive power for phenotype classification to the known clinical features.

Discussion

Homozygosity for the *HFE* p.C282Y mutation is necessary but not sufficient to manifest a disease phenotype in *HFE*-related hereditary hemochromatosis. The present study shows that SNPs in several genes involved in iron metabolism may modulate the expression of the disease in p.C282Y homozygous patients. We analyzed three different outcomes: transferrin saturation, serum ferritin and total iron removed. Transferrin saturation is not a quantitative index of iron overload, but might represent a qualitative index of alteration of iron homeostasis characterized by increased intestinal iron absorption and iron release from macrophages and storage cells.[26] Accordingly, very high transferrin saturation is usually found in the most severe forms of hereditary hemochromatosis and in patients with ineffective erythropoiesis, both characterized by absent or very low level of hepcidin production and high iron absorption.[27] Serum ferritin is generally considered a good index of iron stores in hereditary hemochromatosis[28] and in our series serum ferritin correlated significantly with the amount of iron removed (r=0.613, *P*<0.0001). However, serum ferritin can be influenced by hepatocellular necrosis, inflammation and alcohol intake, which may increase its concentration disproportionately to the amount of iron overload. Phlebotomy, with careful measurement of the amount of iron in the blood removed, is the most accurate means of measuring total body iron stores.[29,30] However, it was not available for all patients studied, this being a limitation when considering a large cohort of patients.

Our sample is largely representative of the local population of p.C282Y homozygotes since it covered a diverse range of phenotypes, spanning from mild through moderate to severe. It differed from the sample recruited in the study by Milet *et al.* which, although from a region in north-western Europe, was not fully representative of the

population of p.C282Y homozygotes from which it was drawn since it was, by the authors' admission, rich in individuals with severe symptoms.[31] In the present study we found a significant association between a variant in *CYBRD1*, rs3806562, and transferrin saturation. Moreover we suggested a series of 17 SNPs which could have a possible additive effect on the studied outcomes. The SNP rs3806562 of *CBRYD1* is located in the 5'UTR of the gene and is, therefore, likely to be functional. HapMap shows that this SNP is in linkage disequilibrium with rs3806566 and rs884409 previously found to be associated with serum ferritin in Australian p.C282Y homozygotes.[16] We, therefore, analyzed SNPs rs3806566 and rs884409 in our cohort and confirmed that the three SNPs are in complete linkage. However, luciferase assays did not show significant differences in promoter activity between different alleles of rs3806562. This suggests that SNP rs3806562 is a genetic marker, located in the 5'UTR of *CYBRD1*, linked to rs884409, a polymorphism able to modulate *CYBRD1* promoter activity.[16] High *CYBRD1* activity might lead to increased amounts of iron available for divalent metal transporter 1 (DMT1) at the epithelial intestinal mucosa and, in turn, to increased iron absorption and transferrin saturation. In contrast to the Australian study, we were not able to find correlations with quantitative indices of iron overload (serum ferritin and total iron removed) and we have no clear explanation for this discrepancy. Serum ferritin and iron removed have their own intrinsic limits (see above) and it is possible that acquired factors such as dietary habits might contribute to modify the whole amount of body iron. However, our results support the hypothesis that *CYBRD1* could be a modifier gene of iron phenotype in patients with *HFE*-hereditary hemochromatosis.

Our results also suggest that genes coding for the bone morphogenetic proteins may be involved in the modulation of iron overload in p.C282Y homozygotes. In particular, we found an association between three SNPs in *BMP9*, *SMAD8* and *BMPR2* and serum ferritin. In addition, iron removed was associated with other polymorphisms present in *BMP2*, *BMP4*, *BMP9*, and *BMPR2* (the same SNP as associated with serum ferritin), and transferrin saturation with another component of this pathway: *BMPR1B*. Although we included rs235756 in *BMP2* (previously reported as a genetic modifier in French patients with *HFE*-hereditary hemochromatosis) in the analysis, we could not confirm this result in our series, and the rs3178250 in *BMP2* that we found associated with iron removed was not in linkage with rs235756. This could be due to a difference in sample sizes between studies, but also to the inherent heterogeneity related to disease, as reported in other genetic studies of complex traits.[32] All these findings suggest that the expression of disease is not only related to the impairment of HFE function but also depends on the modulation exerted by the functional bone morphogenetic protein on the expression of hepatic hepcidin. BMP9 has been shown to be the most potent inducer of hepcidin *in vitro* and also in mice.[33] *In vitro* studies showed that SMAD8 and receptors type I and type II (BMPR1A, BMPR1B, and BMPR2) are also hepcidin modulators,[34] and that *Bmpr1a* is critically responsible for basal hepcidin expression and is required (together with *Acvr1*) for regulation of hepcidin in response to iron and bone morphogenetic protein signaling in mice.[35]

Besides genes of the BMP/SMAD pathways and *CYBRD1*, other genes involved in iron homeostasis emerged from our study. Transferrin saturation was associated with SNPs in genes involved in iron release from storage cells (*Cp*). Previous studies showed the existence of a complex interaction between Cp and Hfe in transgenic mice, suggesting that *Cp* is a modifier gene with a protective effect on *HFE*-hereditary hemochromatosis. Iron removed correlated with SNPs in *DMT1*, *HIF2A*, *IRP1*, and *HFE* itself. Previous studies showed that *DMT1* is over-expressed in *HFE*-hereditary hemochromatosis[36] and that genetic loss of DMT1 modulates iron overload in *Hfe* knockout mice.[37] Although the great majority of our patients carry identical *HFE* SNPs according to the observation that c.845G>A (p.C282Y) mutation is in complete linkage disequilibrium with a unique haplotype,[38] we found that one *HFE* SNP (rs1800702) was associated with variable iron removed. This finding is quite unexpected because it indirectly suggests that the p.C282Y mutation does not completely abolish HFE function. Although this hypothesis needs to be validated, it is to be noted that patients carrying null *HFE* mutations and *Hfe* knockout mice have a more severe iron phenotype than their counterparts carrying p.C282Y or a p.C282Y ortholog mutation.[3] Among the SNPs emerging from the analysis, some (rs9325886, rs17804636, rs3178250, rs11204215, rs1800702, rs149411, rs3806562) were in regulatory 5' and 3' UTR and one in the coding region (rs701753), suggesting that all currently described genes in these pathways might be candidates as modifier genes in p.C282Y homozygotes and opening the way to functional studies to confirm the effects of these variants on gene expression. We were also able to suggest some interactive effects of a couple of different SNPs on transferrin saturation and iron removed. This result should be considered with caution because the interaction analysis was not corrected for multiple comparisons. However, the interactive effect on iron removed of SNPs in *BMP4* and *DMT1*, which are involved in hepcidin regulation and intestinal iron absorption, respectively, is intriguing because it suggests that two different pathways regulating iron status might cooperate in modulating iron overload in p.C282Y homozygotes.

In order to evaluate whether the selected SNPs could provide a predictive signature for adverse phenotype, we re-evaluated our data by dividing our patients with *HFE*-hereditary hemochromatosis into two extreme iron-related phenotype classes. As extensively described in the *Design and Methods* section here and in the *Online Supplement*, we used principal component analysis to extract relevant information from the set of data including all three iron indices: serum ferritin, iron removed and transferrin saturation. We compared the predictive performance, as measured by a cross-validated AUC, between various algorithms considering variable number of SNPs with a model with only clinical characteristics and found that only a small, but still not significant improvement in prediction was achieved by adding genetic information. This could be due either to an inappropriate definition of the phenotype or, more probably, to a strong association between the binary phenotype and clinical variables, namely alcohol consumption.

We did not observe associations with SNPs of *TF* and *TMPRSS6*, which emerged as modulators of some indices of iron status in the general population in a recent genome-wide association study.[17] Transferrin heritability ranged from 0.2 to 0.5 in different isolated populations in

Italy[39] and it seems reasonable to hypothesize that genetic factors might influence transferrin saturation in hereditary hemochromatosis by modulating transferrin levels, thus increasing iron deposition and storage.[19] A decreased serum transferrin level is a common observation in hereditary hemochromatosis and it is generally considered secondary to hepatocellular iron overload,[40] which our results seem to confirm. We tested nine SNPs of *TMPRSS6* including the common SNPs (rs4820268 and rs855791) associated with serum iron and transferrin saturation in the general population.[17,18] Recent studies suggested that rs855791 is a *TMPRSS6* functional variant able to modulate hepcidin production,[41] and that genetic loss of *Tmprss6* in *Hfe* knock-out mice reduces systemic iron overload by increasing Bmp/Smad signaling in an *Hfe*-independent manner.[42] Although these findings suggest that natural genetic variation in the human ortholog *TMPRSS6* might modify the clinical penetrance of *HFE*-associated hereditary hemochromatosis, our results indicate that the SNPs studied did not have enough power to modify the iron phenotype in our series.

In conclusion, the present study suggests that SNPs in genes regulating iron metabolism may modulate the penetrance of *HFE*-hereditary hemochromatosis. These results support the role of bone morphogenetic proteins as possible modifiers of *HFE*-hereditary hemochromatosis phenotype and further expand the observations on a larger number of genes involved in iron absorption and release, with emphasis on the 5' UTR region in *CYBRD1*. Our results also strengthen the notion that none of these polymorphisms alone is a major modifier of the hereditary hemochromatosis phenotype, suggesting that the iron phenotype in this disease is the result of a complex interaction between a major gene defect, genetic background and environmental factors (alcohol intake in particular), thus supporting the idea that *HFE*-hereditary hemochromatosis is a multifactorial disease. From a practical point of view the identification of all these factors, including one or more risk SNPs might add information on patients' susceptibility to fully penetrant *HFE*-hereditary hemochromatosis and would help in adapting the clinical approach better, defining follow-up and therapeutic strategies.

Authorship and Disclosures

The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.

Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.

References

1. Pietrangelo A. Hereditary hemochromatosis: pathogenesis, diagnosis, and treatment. Gastroenterology. 2010;139(2):393-408, e1-2.
2. Deugnier Y, Brissot P, Loreal O. Iron and the liver: update 2008. J Hepatol. 2008;48 (Suppl 1):S113-23.
3. Fleming RE, Holden CC, Tomatsu S, Waheed A, Brunt EM, Britton RS, et al. Mouse strain differences determine severity of iron accumulation in Hfe knockout model of hereditary hemochromatosis. Proc Natl Acad Sci USA. 2001;98(5):2707-11.
4. Bensaid M, Fruchon S, Mazeres C, Bahram S, Roth MP, Coppin H. Multigenic control of hepatic iron loading in a murine model of hemochromatosis. Gastroenterology. 2004; 126(5):1400-8.
5. 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(5):833-7.
6. Whiting PW, Fletcher LM, Dixon JK, Gochee P, Powell LW, Crawford DH. Concordance of iron indices in homozygote and heterozygote sibling pairs in hemochromatosis families: implications for family screening. J Hepatol. 2002;37(3):309-14.
7. Distante S, Elmberg M, Foss Haug KB, Ovstebo R, Berg JP, Kierulf P, et al. Tumour necrosis factor alpha and its promoter polymorphisms' role in the phenotypic expression of hemochromatosis. Scand J Gastroenterol. 2003;38(8):871-7.
8. Mura C, Raguenes O, Ferec C. HFE mutations analysis in 711 hemochromatosis probands: evidence for S65C implication in mild form of hemochromatosis. Blood. 1999;93(8):2502-5.
9. van Dijk BA, Kemna EH, Tjalsma H, Klaver SM, Wiegerinck ET, Goossens JP, et al. Effect of the new HJV-L165X mutation on penetrance of HFE. Blood. 2007;109(12):5525-6.
10. Lee PL, Gelbart T, West C, Halloran C, Felitti V, Beutler E. A study of genes that may modulate the expression of hereditary hemochromatosis: transferrin receptor-1, ferroportin, ceruloplasmin, ferritin light and heavy chains, iron regulatory proteins (IRP)-1 and -2, and hepcidin. Blood Cells Mol Dis. 2001;27(5):783-802.
11. Jacolot S, Le Gac G, Scotet V, Quere I, Mura C, Ferec C. HAMP as a modifier gene that increases the phenotypic expression of the HFE pC282Y homozygous genotype. Blood. 2004;103(7):2835-40.
12. Krayenbuehl PA, Maly FE, Hersberger M, Wiesli P, Himmelmann A, Eid K, et al. Tumor necrosis factor-alpha -308G>A allelic variant modulates iron accumulation in patients with hereditary hemochromatosis. Clin Chem. 2006;52(8):1552-8.
13. Tolosano E, Fagoonee S, Garuti C, Valli L, Andrews NC, Altruda F, et al. Haptoglobin modifies the hemochromatosis phenotype in mice. Blood. 2005;105(8):3353-5.
14. Gouya L, Muzeau F, Robreau AM, Letteron P, Couchi E, Lyoumi S, et al. Genetic study of variation in normal mouse iron homeostasis reveals ceruloplasmin as an HFE-hemochromatosis modifier gene. Gastroenterology. 2007;132(2):679-86.
15. Milet J, Le Gac G, Scotet V, Gourlaouen I, Theze C, Mosser J, et al. A common SNP near BMP2 is associated with severity of the iron burden in HFE p.C282Y homozygous patients: a follow-up study. Blood Cells Mol Dis. 2010;44(1):34-7.
16. Constantine CC, Anderson GJ, Vulpe CD, McLaren CE, Bahlo M, Yeap HL, et al. A novel association between a SNP in CYBRD1 and serum ferritin levels in a cohort study of HFE hereditary haemochromatosis. Br J Haematol. 2009;147(1):140-9.
17. Tanaka T, Roy CN, Yao W, Matteini A, Semba RD, Arking D, et al. A genome-wide association analysis of serum iron concentrations. Blood. 2010;115(1):94-6.
18. Benyamin B, Ferreira MA, Willemsen G, Gordon S, Middelberg RP, McEvoy BP, et al. Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. Nat Genet. 2009;41(11):1173-5.
19. Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A, et al. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. Am J Hum Genet. 2009;84(1):60-5.
20. Piperno A. Classification and diagnosis of iron overload. Haematologica. 1998;83(5): 447-55.
21. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet. 2005;37(11):1217-23.
22. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning, 2nd ed. In: Springer, ed, 2009.
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.
24. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-72; discussion 207-12.
25. Team RDC. R: A language and environment for statistical computing. 2012.
26. Mariani R, Pelucchi S, Arosio C, Coletti S, Pozzi M, Paolini V, et al. Genetic and metabolic factors are associated with increased

hepatic iron stores in a selected population of p.Cys282Tyr heterozygotes. Blood Cells Mol Dis. 2010;44(3):159-63.
27. Pietrangelo A, Caleffi A, Corradini E. Non-HFE hepatic iron overload. Semin Liver Dis. 2011;31(3):302-18.
28. Camaschella C, Poggiali E. Towards explaining "unexplained hyperferritinemia". Haematologica. 2009;94(3):307-9.
29. Worwood M. Laboratory determination of iron status. In: Brock JH HJ, Pippard MJ, Powell LW, ed. Iron Metabolism in Health and Disease. W.B. Saunders, 1994:449-76.
30. Piperno A, Arosio C, Fargion S, Roetto A, Nicoli C, Girelli D, et al. The ancestral hemochromatosis haplotype is associated with a severe phenotype expression in Italian patients. Hepatology. 1996;24(1):43-6.
31. Milet J, Dehais V, Bourgain C, Jouanolle AM, Mosser A, Perrin M, et al. Common variants in the BMP2, BMP4, and HJV genes of the hepcidin regulation pathway modulate HFE hemochromatosis penetrance. Am J Hum Genet. 2007;81(4):799-807.
32. Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype-phenotype associations. Nature. 2007;447(7145):655-60.
33. Truksa J, Peng H, Lee P, Beutler E. Bone morphogenetic proteins 2, 4, and 9 stimulate murine hepcidin 1 expression independently of Hfe, transferrin receptor 2 (Tfr2), and IL-6. Proc Natl Acad Sci USA. 2006;103(27):10289-93.
34. Babitt JL, Huang FW, Xia Y, Sidis Y, Andrews NC, Lin HY. Modulation of bone morphogenetic protein signaling in vivo regulates systemic iron balance. J Clin Invest. 2007;117(7):1933-9.
35. Steinbicker AU, Bartnikas TB, Lohmeyer LK, Leyton P, Mayeur C, Kao SM, et al. Perturbation of hepcidin expression by BMP type I receptor deletion induces iron overload in mice. Blood. 2011;118(15):4224-30.
36. Griffiths WJ, Sly WS, Cox TM. Intestinal iron uptake determined by divalent metal transporter is enhanced in HFE-deficient mice with hemochromatosis. Gastroenterology. 2001;120(6):1420-9.
37. Levy JE, Montross LK, Andrews NC. Genes that modify the hemochromatosis phenotype in mice. J Clin Invest. 2000;105(9):1209-16.
38. Yang Y, Ferec C, Mura C. SNP and haplotype analysis reveals new HFE variants associated with iron overload trait. Hum Mutat. 2011;32(4):E2104-17.
39. Traglia M, Sala C, Masciullo C, Cverhova V, Lori F, Pistis G, 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(10):e7554.
40. Morton AG, Tavill AS. The role of iron in the regulation of hepatic transferrin synthesis. Br J Haematol. 1977;36(3):383-94.
41. Nai A, Pagani A, Silvestri L, Campostrini N, Corbella M, Girelli D, et al. TMPRSS6 rs855791 modulates hepcidin transcription in vitro and serum hepcidin levels in normal individuals. Blood. 2011;118(16):4459-62.
42. Finberg KE, Whittlesey RL, Andrews NC. Tmprss6 is a genetic modifier of the Hfe-hemochromatosis phenotype in mice. Blood. 2011;117(17):4590-9.