

# Identification of a novel quantitative trait nucleotide related to iron status in a calcium channel gene

Carlos Baeza-Richer<sup>a,\*</sup>, Ruth Blanco-Rojo<sup>b</sup>, Ana M. López-Parra<sup>a</sup>, Anna Brichs<sup>c</sup>, Stefania Bertoncini<sup>d</sup>, Ana M. Pérez-Granados<sup>b</sup>, Alfonso Buil<sup>c</sup>, José M. Soria<sup>c</sup>, Eduardo Arroyo-Pardo<sup>a</sup> and M. Pilar Vaquero<sup>b</sup>

<sup>a</sup>*Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain*

<sup>b</sup>*Department of Metabolism and Nutrition, Institute of Food Science, Food Technology and Nutrition, Spanish National Research Council, Madrid, Spain*

<sup>c</sup>*Unit of Genomic of Complex Diseases, Institute of Biomedical Research, Barcelona, Spain*

<sup>d</sup>*Department of Biology, University of Pisa, Pisa, Italy*

**Abstract.** Several iron-related parameters have been reported to show significant heritability, and thus, seemed to be genetically regulated. A genome wide family-based study revealed two regions that showed a linkage signal with transferrin receptor levels. The aim of the study was to identify genetic markers associated with iron status biomarkers. Ten SNPs selected from the literature were tested, and parameters related to iron metabolism were analysed, in a group ( $n = 284$ ) of Spanish women. Data were analyzed using Bayesian Model Averaging (BMA) test and decision trees. The rs1375515, located in an intronic region of the calcium channel gene *CACNA2D3*, showed strong associations with levels of mean corpuscular volume according to BMA test, and with levels of haemoglobin and ferritin according to decision trees. The allele G was associated to low levels of these parameters which suggests higher iron deficiency anaemia risk. This SNP along with the C282Y mutation explained significant differences in the distribution of individuals in three iron-related clinical phenotypes (normal, iron deficient and iron deficiency anaemic). In conclusion, the rs1375515, or other genetic polymorphisms in linkage, may play important roles in iron status, probably by affecting the function of a calcium channel. These findings may be useful for further investigation in the etiology of iron diseases.

Keywords: Quantitative trait nucleotide, iron deficiency anaemia, calcium channel gene, SNP, association study

## 1. Introduction

Specific levels of iron in biological fluids are essential for normal body function, in oxygen transport and for other important metabolic reactions. Thus, fine regulation of this element is required since departures from its optimal levels produce severe alterations. For example, iron overload triggers the Fenton reaction in

which the generation of hydroxyl radicals, causes may- or tissue damages [1]. By contrast, iron deficiency may cause anaemia. As in mammals there are no active mechanisms to eliminate iron, its absorption and storage regulation are crucial for maintaining appropriate levels of this mineral.

Among iron related diseases, anaemia is the most wide world spread, being considered as a pandemic. According to World Health Organization (WHO), a quarter of the world population suffers from anaemia, being women at fertile age the most affected group [2]. The prevalence of iron deficiency has been estimated to be between 8% and 33% of young women in Europe,

---

\*Corresponding author: Carlos Baeza-Richer, Department of Toxicology and Health Legislation, Faculty of Medicine, Complutense University of Madrid, Complutense Avenue, 28040 Madrid, Spain. Tel./Fax: +34 913 941 576; E-mail: cbaezaricher@med.ucm.es.

10%–16% in the United States, and 42% in developing countries [2].

Not only traditional dietary and host-related factors are determinants of iron levels, but also genetic factors play an important role. Mutations in key proteins involved in iron metabolism, red cell stability, and in iron absorption, have proved to generate severe anaemia and haemochromatosis [3–8]. Haemoglobin, serum ferritin, transferrin saturation with iron, which are parameters employed to assess iron status, show significant levels of heritability, and thus, seem to be genetically regulated [9,10]. In this context, recent studies have demonstrated that common allelic variants of certain genes are associated with iron related phenotypes [11,12], or account for a remarkable percentage of the genetic variation in the levels of iron-related parameters [13,14]. A genome wide linkage study revealed two regions that were related to the levels of transferrin receptor [15]: the higher was located on Chromosome 2 (lod score: 2.64) and the lower on Chromosome 3 (lod score: 1.94).

Therefore, by genotyping genetic determinants in a risk population for anaemia this study was aimed at two objectives: First, confirm in an independent sample, the importance of a specific group of genetic markers that has been already reported to be associated with several parameters related to iron status. Second, to follow up our previously observed linkage signals from a genome-wide linkage study [15] by testing three SNPs selected from the two regions (Chromosome 2 and 3) that showed the before mentioned signals. In summary, the main goal was to identify common genetic variants associated with iron status especially with anaemia.

## 2. Material and methods

### 2.1. Subjects

A total of 284 subjects were recruited as part of a wide project aimed at understanding the interactions among iron, genes, nutrition and disease in menstruating women. As described in [16], the subjects were Caucasian women born in Europe, between 18 and 45 years of age, menstruating, non-pregnant, non-lactating and non-smoking. In this sample, 43.7% of the total individuals had normal haemoglobin (Hb) and ferritin (Ft) levels ( $Hb \geq 12$ ,  $Ft \geq 20$ ), 45.8% were iron deficient ( $Hb \geq 12$  and  $Ft < 20$  or  $Hb \leq 12$  and  $Ft > 20$ ), and 10.6% were iron deficient anaemic ( $Hb < 12$  and  $Ft < 20$ ). All women underwent a health questionnaire and blood test. The study was approved by the ethics committees of the *Clínica Puerta de Hierro* Hospital and the Spanish National Research Council, Madrid, Spain.

### 2.2. Hematological and biochemical determinations

Venous fasting blood samples (5 mL) were collected in EDTA tubes to analyse haemoglobin(g/dl), haematocrit (%), and mean corpuscular volume (MCV)(fl.) and in tubes with Gel + Clot activator to determine serum iron ( $\mu\text{g/dl}$ ) and serum ferritin (ng/ml) by automatic standardized methods.

### 2.3. Genotyping

DNA was extracted from whole blood using standard phenol-chloroform methodology with proteinase K. Genotyping of 10 SNPs was carried out by a minisequencing method (Table 1), described in [16]. Criteria for selecting the 10 SNPs for minisequencing were reported in the same study. Briefly, these criteria were to choose those SNPs which were reported to be associated with biochemical parameters related to iron metabolism. A group of SNPs (rs3811647, rs1799852, rs2280673 rs1800562, rs855791 and rs1799945) were selected due to their association with haemoglobin levels and other iron-related parameters [17–20]. Another SNP (rs4820268) was selected because of its association with serum iron [13]. Finally three more SNPs were chosen from two linkage signals found in the Remacha et al. 2006 study [15]. These three tagSNPs were selected, by means of the The HapMap Project (<http://www.hapmap.org/>) [21]: two from the linkage signal located on Chromosome 2 (rs16826756 and rs2673289), and another one from Chromosome 3 (rs1375515).

### 2.4. Statistical analysis

SNP allele frequencies assessment, genotype frequencies and Hardy-Weinberg equilibrium (HWE) were carried out. Normal distribution of hematological and biochemical parameters was assessed. All analysis was carried out using the R package (<http://www.Rproject.org/>) [22]. Serum ferritin was log transformed before it was analyzed, because it showed a skewed distribution.

The association between genotype and phenotype variables was studied by decision trees using SPSS vs. 17 statistic package, and Bayesian Model Averaging (BMA) and linear regression models using R statistic package.

A decision tree was tested for each haematological and biochemical parameter, as well as for the distribution of the three clinical phenotypic groups (normal,

Table 1  
Descriptive analysis of the 10 studied SNPs

| SNP        | Allele frequencies |         | Genotype frequencies |          |          | p-value, $X^2$ (HWE) | Location     |
|------------|--------------------|---------|----------------------|----------|----------|----------------------|--------------|
| rs4820268  | 0.54(A)            | 0.46(G) | 0.28(AA)             | 0.51(AG) | 0.21(GG) | 0.871493             | TMPRSS6      |
| rs855791   | 0.56(C)            | 0.44(T) | 0.32(CC)             | 0.49(CT) | 0.19(TT) | 0.515545             | TMPRSS6      |
| rs1799852  | 0.88(C)            | 0.12(T) | 0.76(CC)             | 0.24(CT) | 0(TT)    | 0.125421             | TRANSFERRIN  |
| rs2280673  | 0.62(C)            | 0.38(A) | 0.38(CC)             | 0.49(CA) | 0.13(AA) | 0.607449             | TRANSFERRIN  |
| rs1800562  | 0.97(G)            | 0.03(A) | 0.93(GG)             | 0.07(GA) | 0(AA)    | 0.505265             | HFE          |
| rs3811647  | 0.67(G)            | 0.33(A) | 0.44(GG)             | 0.47(GA) | 0.1(AA)  | 0.059705             | TRANSFERRIN  |
| rs2673289  | 0.61(C)            | 0.39(T) | 0.38(CC)             | 0.47(CT) | 0.15(TT) | 0.906312             | CHROMOSOME 2 |
| rs1375515  | 0.67(A)            | 0.33(G) | 0.44(AA)             | 0.45(AG) | 0.1(GG)  | 0.655416             | CACNA2D3     |
| rs1799945  | 0.82(C)            | 0.18(G) | 0.67(CC)             | 0.29(CG) | 0.04(GG) | 0.777062             | HFE          |
| rs16826756 | 0.81(A)            | 0.19(G) | 0.65(AA)             | 0.31(AG) | 0.04(GG) | 0.960139             | CHROMOSOME 2 |

iron deficient and anaemic) considering the latter as dependent variables and the 10 SNPs as independent variables. The decision trees are built by means of a non-parametric method that clusters the observations according to a factor or predictor that better explains the differences in the studied variable. Each generated subdivision is again partitioned according to the existence of new predictors with a significant effect, thus, the data are presented in a hierarchical manner [23]. This method extracts information by finding effects of the factors which are not homogeneous at all the levels of the dependent variable, and by discovering specific interactions among variables or data mining [24, 25]. From all the different analyses available for this partitioning approach, those algorithms that apply non-binary divisions are more convenient. In our study, the CHAID (Chi-square Automatic Interaction Detection) exhaustive method was selected [26]. The significant threshold was  $p < 0.05$ , the minimum size of a node that could be divided consisted of 20 individuals and the minimum size of a child node consisted of 5 individuals. Finally this method applied corrections due to multiple testing through the Bonferroni method.

An analysis of the association between each of the biochemical parameters and the 10 SNPs through a BMA was carried out. BMA test accounts for the model uncertainty in the variable selection by averaging over the best models according to a posterior probability model. BMA test estimates all of the possible models and calculates a probabilistic average of the effect of each SNP.

In our study, using 10 SNPs, we considered  $2^{10} = 1024$  possible models. The BMA posterior distribution of the effect of each SNP is a weighted average of its posterior distribution under all of the considered models [27].

Because results obtained with BMA and decision trees varied, linear regression models were calculated for the rs1375515 with all the biochemical parameters studied, testing for recessive and dominance models.

HapMap project (release #27) [21] was employed to describe the genomic location of rs1375515. Haploview 4.1 used the information of a 100kb region, extracted from the HapMap, to describe in detail the linkage disequilibrium patterns of the surroundings of rs1375515. Haplotype blocks were defined using a confidence interval method describe elsewhere [28].

### 3. Results

Table 1 lists the allele and genotype frequencies, as well as the p-values for HWE test for the 10 studied SNPs. The 10 SNPs presented a Minimum Allele Frequency (MAF) greater than 0.01, and were in HWE.

#### 3.1. BMA test

Table 2 shows the results of the BMA test for MCV. The rs1375515 showed a significant association with MCV yielding a 55.4% probability value of being associated with this parameter, and appeared in the model which presented the highest posterior probability (0.225). The rest of the BMA tests did not include any of the SNPs in the best model according to its posterior probability and thus they are not presented.

#### 3.2. Decision trees

Only those decision trees that showed at least one significant subdivision ( $p$ -value  $< 0.05$ ) are described.

Figure 1 depicts the decision tree for haemoglobin. It is rs1375515 which better explains this variable, partitioning the sample ( $P = 0.019$ ) into two subgroups. Those individuals genotyped as A/A and those as A/G were clustered together with a mean haemoglobin level of 13.189 g/dl. The rest of the sample, those typed as G/G homozygous were included in a different node with a mean value of 12.641 g/dl. This suggests that

Table 2  
Association analysis between the 10 SNPs and Mean Corpuscular Volume (MCV) using BMA test

|            | $p! = 0$ | EV         | SD      | Model1  | Model2 | Model3  | Model4  | Model5  |
|------------|----------|------------|---------|---------|--------|---------|---------|---------|
| Intercept  | 100      | 87.7448111 | 0.60796 | 87.9715 | 87.242 | 88.3481 | 87.6379 | 88.3896 |
| rs4820268  | 1.3      | 0.0002043  | 0.04914 |         |        |         |         |         |
| rs855791   | 3.4      | 0.0116184  | 0.09927 |         |        |         |         |         |
| rs1799852  | 7        | 0.0629203  | 0.29384 |         |        |         |         |         |
| rs2280673  | 3.2      | 0.0099255  | 0.09574 |         |        |         |         |         |
| rs1800562  | 7.1      | 0.1100278  | 0.5081  |         |        |         |         |         |
| rs3811647  | 11.1     | -0.0787761 | 0.2702  |         |        |         |         | -0.6728 |
| rs2673289  | 5        | -0.0251253 | 0.14551 |         |        |         |         |         |
| rs1375515  | 55.4     | -0.6066949 | 0.63944 | -1.1079 |        | -1.0922 |         | -1.07   |
| rs1799945  | 5.1      | 0.0328943  | 0.18588 |         |        |         |         |         |
| rs16826756 | 21.5     | -0.2214793 | 0.49003 |         |        | -1.0174 | -1.0412 |         |
| nVar       |          |            |         | 1       | 0      | 2       | 1       | 2       |
| r2         |          |            |         | 0.021   | 0      | 0.034   | 0.013   | 0.029   |
| BIC        |          |            |         | -0.3901 | 0      | 1.5651  | 1.8608  | 3.0805  |
| postprob   |          |            |         | 0.225   | 0.186  | 0.085   | 0.073   | 0.04    |

$p! = 0$ : Probability of the SNP to be associated with the variable. EV: Posterior mean of the beta parameter. SD: Standard deviation of each beta. Model1...5: Most probable multiple-SNP models. Post prob: Posterior probability of the model. nVar: Number of SNPs included in the model. BIC: Bayesian Information Criterion. r2: Coefficient of determination.

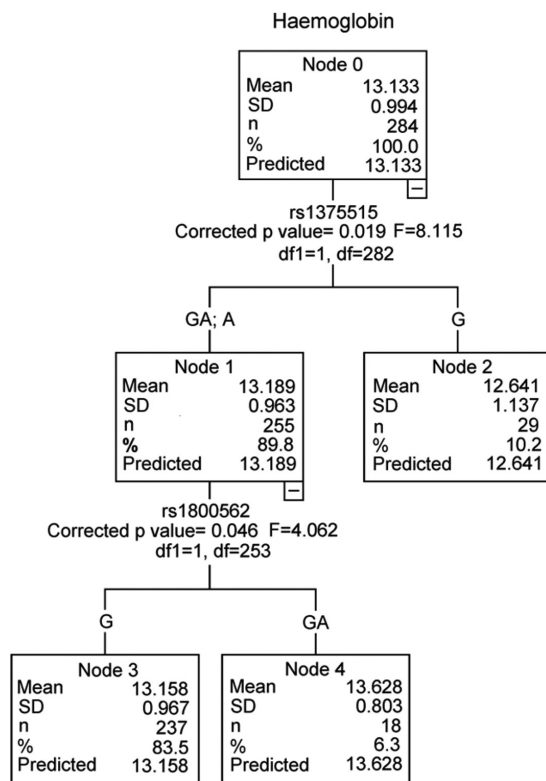


Fig. 1. Decision tree built with Haemoglobin (g/dl) as dependent variable and the 10 SNPs as independent variables or factors. %: Percentage of the total sample included in each node. Predicted: Predicted mean value assuming no effect for the factors.

the absence of the A variant reduces the levels of haemoglobin by approximately 0.54 g/dl. As for the first node, it was split into two other subgroups ( $P = 0.045$ ), the rs1800562 (corresponding to the C282Y

non-synonymous amino acid change) being the one producing this division. These two subgroups correspond to those individuals homozygous G/G and those heterozygous G/A. Thus, individuals carrying the allele A tend to express higher levels of haemoglobin (13.628 g/dl vs. 13.156 g/dl).

The tree built for  $\log_{10}$ ferritin is shown in Fig. 2. It seems very similar to that of the haemoglobin, thus, the subdivisions have occurred in the same manner being the rs1375515 the one which produces the first split ( $P = 0.048$ ), and rs1800562 the one that produces the second ( $P = 0.038$ ). In the first subdivision, individuals A/A and A/G were clustered together yielding a mean of 1.284 (Ft values: 25.894 ng/ml) whereas individuals G/G presented a mean of 1.112 (Ft values: 18.033 ng/ml). The second division occurs in the group with higher mean, as in the previous tree. Considering this is a hierarchical method, the rs1375515, accounts for the variability of this parameter better than the rest of the SNPs, but with a lower significance level ( $P = 0.048$ ) in comparison with the haemoglobin tree.

Figure 3 shows the decision tree that assumes the distribution of the individuals in three phenotypic groups (anaemic, deficient and normal) as the dependent variable. As in the previous trees the first subdivision is caused by the SNP rs1375515 generating two nodes ( $P = 0.02$ ). In node 2, which includes all the individuals homozygous for G/G, the distribution of phenotypes is significantly different from that of the whole sample (node 0), where the anaemic group having a sharp increase (27.6% vs. 10.6%) and both the normal and the deficient groups a considerable downscale. In node 2, only one out of 29 individuals presents the

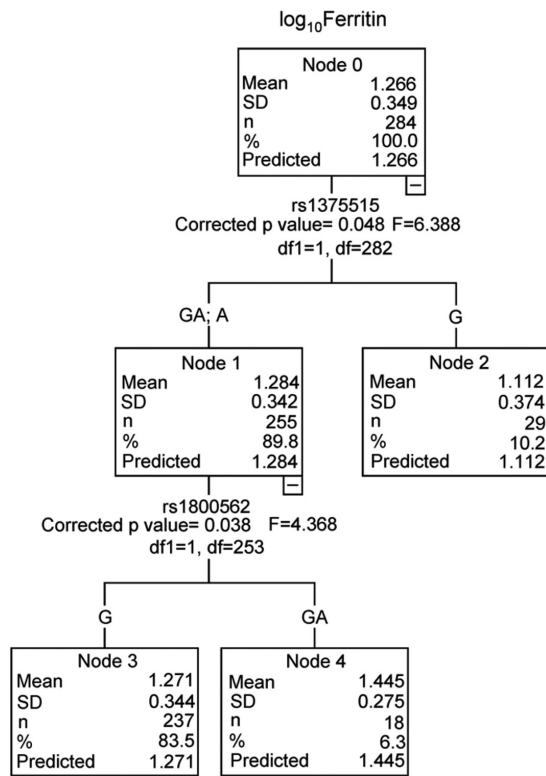


Fig. 2. Decision tree built with  $\log_{10}$ Ferritin (ng/ml) as dependent variable and the 10 SNPs as independent variables or factors. %: Percentage of the total sample included in each node. Predicted: Predicted mean value assuming no effect for the factors.

A mutation for rs1800562. Node 1 is subdivided into two groups by the SNP rs1800562. The resulting node 4, which contains those individuals heterozygous G/A, has a significantly different distribution of phenotypes, where there were no anaemic individuals (0%) and the proportion of normal individuals was highly increased compared to that of the total sample (77.8% in node 4 vs. 43.7% in the complete sample). This tree shows the significant combined effect of rs1375515 and rs1800562 over iron status, thus, in node 2, the effect of being homozygous G/G (rs1375515) and not carrying the A mutation of rs1800562 raises significantly the risk of anaemia, while in node 4 the combination of not being homozygous for the G mutation of rs1375515 and presenting the A mutation in rs1800562 produces a very low probability of anaemia and therefore increases the probability of being included in the normal group.

### 3.3. Dominance and recessive models

According to the decision trees, the association of the rs1375515 with haemoglobin levels and to a lesser

### Iron-clinical phenotypes

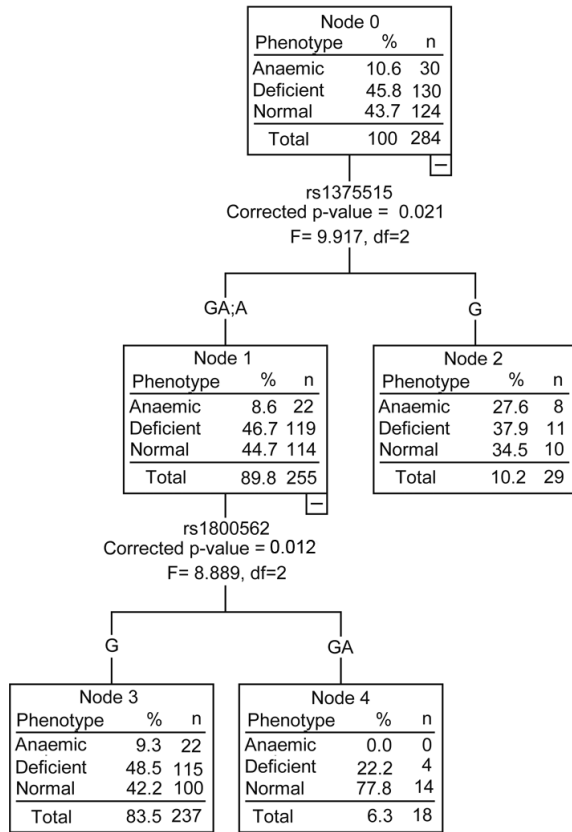


Fig. 3. Decision tree built with the distribution of iron clinical phenotypes (Iron deficient anaemic, iron deficient, and normal) as dependent categorical variable and the 10 SNPs as independent variables or factors. %: Percentage of each phenotype included in each node.

extent with  $\log_{10}$ ferritin levels, raised questions about why this SNP did not appear significantly associated in the BMA test, for the same parameters.

The pattern the rs1375515 shows in the trees when producing the subdivisions, suggests that this SNP does not follow a co-dominance effect model. Thus, homozygous A/A and heterozygous A/G cluster in the same node, whereas G/G homozygous express a significant lower mean. This suggests that a recessive model may apply to these data. As BMA test assumes a co-dominant model, this could be the cause of the discrepancy between the two statistics.

Therefore, dominant and recessive models were tested for the rs1375515 with all the variables. The results presented in Table 3 show that this SNP generates significant p-values for regression models considering recessive models, for the variables haemoglobin,  $\log_{10}$ ferritin, haematocrit and MCV.

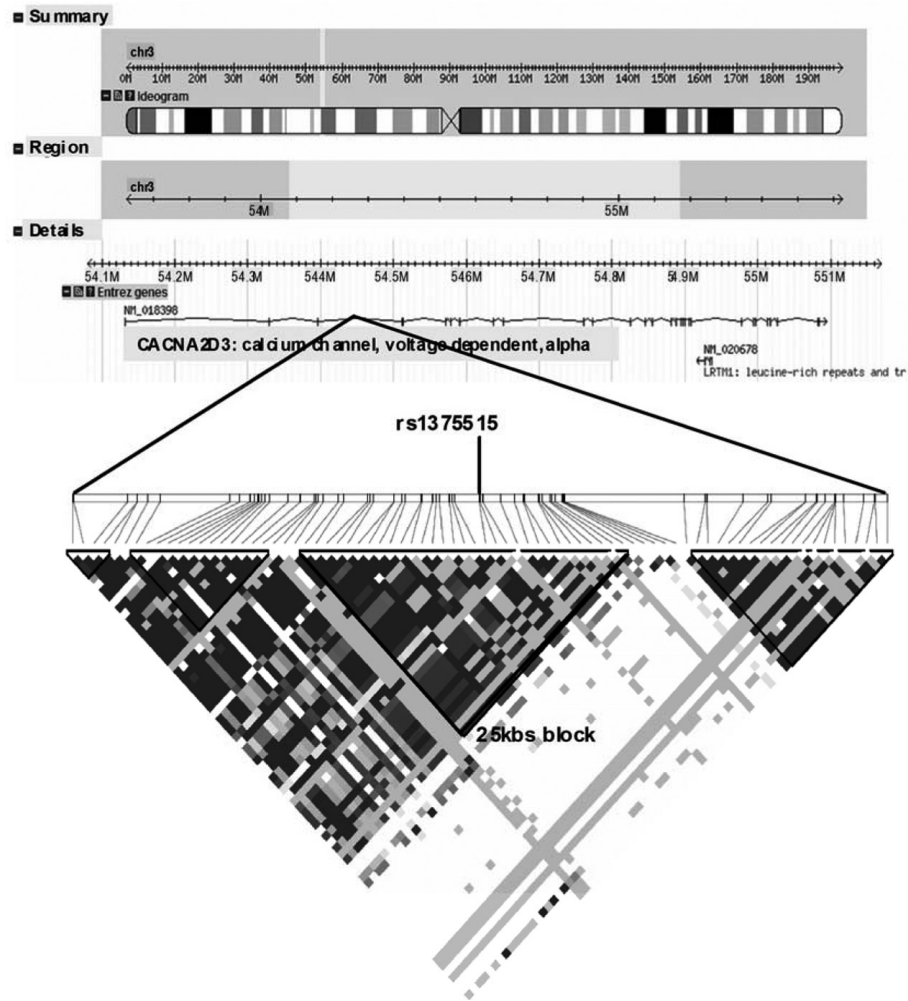


Fig. 4. Plots modified from HapMap (release # 27) and Haploview vs. 4.1. HapMap plot shows a region of 1 Mb in which the *CACNA2D3* is included. The linkage disequilibrium plot (values of  $r^2$ ) was constructed from a 100 kbs region in which rs1375515 is included. Darker diamonds indicate higher  $r^2$  values.

Table 3

Simple linear regression analysis of variables versus rs1375515 assuming a recessive model

|             | Effect  | Std. Error | t value | P-value | Variable             |
|-------------|---------|------------|---------|---------|----------------------|
| (Intercept) | 13.1894 | 0.0615     | 214.55  | 0       | Haemoglobin          |
| rs1375515   | -0.548  | 0.1924     | -2.85   | 0.0047  |                      |
| (Intercept) | 2.9555  | 0.0499     | 59.28   | 0       | $\log_{10}$ Ferritin |
| rs1375515   | -0.3943 | 0.156      | -2.53   | 0.012   |                      |
| (Intercept) | 39.3071 | 0.169      | 232.56  | 0       | Haematocrit          |
| rs1375515   | -1.2588 | 0.5289     | -2.38   | 0.018   |                      |
| (Intercept) | 87.448  | 0.3121     | 280.19  | 0       | MCV                  |
| rs1375515   | -2.0169 | 0.9767     | -2.07   | 0.0398  |                      |

### 3.4. HapMap and Haploview analysis

Figure 4 describes the genomic location of rs1375515, on Chromosome 3 at position 54451680. This position maps in an intronic region of a gene that

expresses an alpha voltage dependant calcium channel (*CACNA2D3*). The LD pattern of a surrounding region of 100 kb is showed in Fig. 4. According to Haploview rs1375515 is located in a 25 kb haplotype block.

## 4. Discussion

The main goal of our study was to identify common genetic variants associated with general iron status. We show for the first time that rs1375515 is associated with the levels of iron-related biomarkers, as well as with iron clinical phenotypes (normal, iron deficient and anaemic). According to BMA test, this SNP showed significant association with the levels of MCV. It also

showed significant association with the levels of ferritin and haemoglobin according to decision trees. This is significant, since these two variables (haemoglobin and ferritin) are employed by the WHO to define healthy and pathological groups regarding iron status. The variant G in homozygosis was found in the current study to be significantly associated with higher risk of belonging to the anaemic group.

Although the two main statistics employed (BMA test and decision trees) did not present the same results, we should highlight that rs1375515 behaves in the same manner if we focus on its observed effect on iron status. Even though the results were different between the two statistical methods used (BMA and decision trees), the observed iron status was similar as regards of rs1375515. Regarding the decision trees, the allele G, or minor allele, is associated with low levels of both haemoglobin and ferritin which points to a lower iron status. Thus, for the BMA test for MCV, the coefficient of association in model 1 shows that G allele is related to low levels of this parameter and it is known that low MCV together with low ferritin and haemoglobin values indicate iron deficiency anaemia risk. Therefore, from a physiological point of view, the presence of the G allele seems to be always associated with lower iron status.

Moreover, decision trees results suggest that rs1375515 had a recessive model effect on haemoglobin, ferritin, and on the clinical phenotypes. The results obtained using linear regression models were also consistent with the recessive effect hypothesis. Therefore, this could be the cause of not having found those associations in the BMA test given that this analysis only assumes codominant model effects.

Since rs1375515 is a tagSNP (according to HapMap criteria), it captures a great amount of the genetic variation in the region where it is placed. This SNP is located in the intronic region of a gene that expresses an alpha voltage dependant calcium channel (*CACNA2D3*). The significant associations that were found could mean either that this SNP itself affects the parameters levels, or that there is another SNP in linkage disequilibrium (LD) with this one, which is responsible for the statistical association. The calcium channel gene, in which this SNP is located, spans approximately 1Mbs on chromosome 3, with rs1375515 separated approximately 750 kbs from the 5' end and 300 kbs from the 3' end of this gene. Linkage disequilibrium values in the genome can show different patterns among populations, and decrease with genetic distance, thus LD blocks greater than 100 kbs are rare [29]. In this regard, rs1375515

is located, according to Haploview analysis, in a 25 kb haplotype block which suggests that if there is another SNP, in LD with rs1375515 responsible for the statistical associations, it would be located most probably in the same block or in adjacent blocks within the calcium channel gene. Therefore, the data are pointing to a possible relation between variants of *CACNA2D3* gene and several iron status biomarkers (MCV, haemoglobin, ferritin) as well as with iron clinical phenotypes.

Regarding the association with MCV, given the position of rs1375515 in a calcium gene, there are several evidences in the bibliography that relate cell volume and ionic channels with iron metabolism pathologies that should be discussed as follows. Fine regulation of the cell volume, associated with ionic exchange [30, 31], is essential for normal cellular function, and consequently, alterations in cell volume participate in the physiopathology of disorders such as liver insufficiency, fibrosing disease, and sickle anaemia [33]. Moreover, red cells are extremely sensitive to volume changes that may easily produce haemolytic related anaemias [33]. Recently, Sebastiani et al. [34] have discovered some genetic modifiers of the severity of sickle cell anaemia, some of them within a  $K^+$  channel *KCNK6*. Although this association is not fully understood, this work may support the fact that cationic exchange is somehow associated with erythrocytes volume and therefore with deficiency anaemia or other anaemias.

According to the bibliography, calcium channels are reported, to be associated with cell volume and iron status as well. In Lew et al. [33] it was shown how blocking these channels affected red cell volume by dehydration. This resulted, as well, in changes in cell volume and anaemia. In the same line, it has been demonstrated [35] that blocking calcium channels had beneficial effects on an iron overload cardiomyopathy. Other works show that these channels could be directly responsible for iron uptake, namely the L-type voltage dependent calcium channel could be an alternate route for iron to enter the cardiomyocytes [36].

All these evidences may suggest that the G variant of the rs1375515, or other variants in LD, can be involved in hampering the calcium channel function which could increase the risk of anaemia. The intronic location of this SNP should not be an argument to discard its possible effect, as other SNPs located in intronic regions have been proved to influence the levels of the protein expression [13,34].

We found that the C282Y substitution (rs1800562) is related to high levels of haemoglobin and ferritin, and

to low probabilities of being anaemic, counterbalancing to some extent the effect of the G allele of rs1375515. In fact, the results have shown the combined effect of these two SNPs can significantly vary the distribution of the iron clinical phenotypes. This SNP (rs1800562) located at the *HFE* gene, is associated with haemochromatosis [19,37,38] and has been reported to show significant effect on haemoglobin levels as well as other parameters related to iron status such as serum ferritin, serum iron, serum transferrin and transferrin saturation [13,17]. This study supports other recent findings in which common variants the *HFE* gene [18], may be used as predictor factors of iron status.

The rs855791 and the rs4820268, both located in *TMPRSS6* gene, were analyzed because they have recently been associated with haemoglobin and iron status, MCV [13,18,39], serum iron [17] and to some iron-related diseases such as iron refractory anaemia [5]. In our study, these SNPs had no significant effect on the levels of any of the parameters regardless of the statistic employed (data not shown). A possible explanation could be that the relatively small size of our sample ( $n = 284$ ) hampers to find associations for SNPs that do not show very strong effect. In this respect the effect of the recessive homozygous of rs855791 on haemoglobin levels in Chambers et al. [18] was a 0.21 g/dl reduction (in a larger sample), whereas in the same case for rs1375515 (present study) the reduction was of 0.54 g/dl, which is notably higher.

The region in which rs2673289 and rs16826756 are located (Chromosome 2), showed a linkage signal with transferrin receptor in a previous family-based study [15], however, in the present work these SNPs have not showed any association with the studied parameters. An explanation for this could be that the markers that produced the linkage signal, may be rare or specific of some families, and hence, are not present in our sample.

It is important to note that our sample was not chosen to be representative of the general population. The individuals were chosen to be women at fertile age since this is the population group most at risk of suffering from anaemia. Following the WHO criteria for haemoglobin and ferritin levels, in our sample the proportion of those defined as iron deficient and anaemic (45.8% and 10.6%) compared with those defined as healthy (43.7%) could be overrepresented compared with the general population.

In summary, our results show that the rs1375515 is associated with haematological and biochemical parameters used to assess iron status. The location of this

SNP in a calcium channel gene (*CACNA2D3*), suggests that the functionality of this channel regarding its relation to iron related parameters and anaemia should be further investigated. The combined effect of this SNP together with the C282Y substitution (rs1800562) explains significant differences in the risk of developing anaemia. Thus, our study has broad implications for future studies that focus on the basis of genetic variation within iron-related traits.

### Acknowledgments

This study was supported partially by projects AGL2009-11437, N8/2006-4130063, PI-08/0420 and PI-08/0756. R Blanco-Rojo was supported by a JAE-predoc grant from CSIC and Social European Found, S Bertoncini was supported by Grupo Santander 2009 (Estancia doctores y tecnólogos UCM, and J.M. Soria by "Programa d'Estabilització d'Investigadors de la Direcció d'Estratègia i Coordinació del Departament de Salut" (Generalitat de Catalunya).

### References

- [1] M.W. Hentze, M.U. Muckenthaler and N.C. Andrews, Balancing acts: molecular control of mammalian iron metabolism, *Cell* **117** (2004), 285-297.
- [2] B. de Benoist, E. McLean, I. Egli and M. Cogswell, Worldwide prevalence of anaemia 1993-2005: WHO Global Database on Anaemia. Geneva, World Health Organization, 2008.
- [3] K.E. Finberg, Iron-refractory iron deficiency anemia, *Semin Hematol* **46** (2009), 378-386.
- [4] S.G. Gehrke, A. Pietrangelo, M. Kascák, A. Braner, M. Eisold, H. Kulaksiz, T. Herrmann, U. Hebling, K. Bents, R. Gugler and W. Stremmel, HJV gene mutations in European patients with juvenile hemochromatosis, *67 Clin Genet* (2005), 425-428.
- [5] A.J. Ramsay, V. Quesada, M. Sanchez, C. Garabaya, M.P. Sardà, M. Baiget, A. Remacha, G. Velasco and C. López-Otín, Matriptase-2 mutations in iron-refractory iron deficiency anemia patients provide new insights into protease activation mechanisms, *Hum Mol Genet* **18** (2009), 3673-3683.
- [6] A. Roetto, F. Daraio, F. Alberti, P. Porporato, A. Calì, M. De Gobbi and C. Camaschella, Hemochromatosis due to mutations in transferrin receptor 2, *Blood Cells Mol Dis* **29** (2002), 465-470.
- [7] B. Sarria, S. Navas-Carretero, A.M. Lopez-Parra, A.M. Pérez-Granados, E. Arroyo-Pardo, M.A. Roe, B. Teucher, M.P. Vaquero and S.J. Fairweather-Tait, The G277S transferrin mutation does not affect iron absorption in iron deficient women, *Eur J Nutr* **46** (2007), 57-60.
- [8] D.J. Weatherall, Pathophysiology of thalassaemia, *Baillieres Clin Haematol* **11** (1998), 127-146.
- [9] O.T. Njajou, B.Z. Alizadeh, Y. Aulchenko, M.C. Zillicens, H.A. Pols, B.A. Oostra, D.W. Swinkels and C.M. van



- Duijn, Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study, *Hum Hered* **61** (2006), 222-228.
- [10] J.B. Whitfield, L.M. Cullen, E.C. Jazwinska, L.W. Powell, A.C. Heath, G. Zhu, D.L. Duffy and N.G. Martin, Effects of HFE C282Y and H63D polymorphisms and polygenic background on iron stores in a large community sample of twins, *Am J Hum Genet* **66** (2000), 1246-1258.
- [11] H.K. Bayele and S.K. Sräi, Genetic variation in hepcidin expression and its implications for phenotypic differences in iron metabolism, *Haematologica* **94** (2009), 1185-1188.
- [12] C.E. McLaren, C.P. Garner, C.C. Constantine, S. McLachlan, C.D. Vulpe, B.M. Snively, V.R. Gordeuk, D.A. Nickerson, J.D. Cook, C. Leiendecker-Foster, K.B. Beckman, J.H. Eckfeldt, L.F. Barcellos, J.A. Murray, P.C. Adams, R.T. Acton, A.A. Killeen and G.D. McLaren, Genome-Wide Association Study Identifies Genetic Loci Associated with Iron Deficiency, *Plos One* (2011), 6: e17390. doi:10.1371/journal.pone.0017390.
- [13] B. Benyamin, A.F. McRae, G. Zhu, S. Gordon, A.K. Henders, A. Palotie, L. Peltonen, N.G. Martin, G.W. Montgomery, J.B. Whitfield and P.M. Visscher, Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels, *Am J Hum Genet* **84** (2009), 60-65.
- [14] N. Soranzo, T.D. Spector, M. Mangino et al., A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium, *Nat Genet* **41** (2009), 1182-1190.
- [15] A.F. Remacha, J.C. Souto, J.M. Soria, A. Buil, M.P. Sardà, M. Lathrop, J. Blangero, L. Almasy and J. Fontcuberta, Genomewide linkage analysis of soluble transferrin receptor plasma levels, *Ann Hematol* **85** (2006), 25-28.
- [16] S. Bertoni, R. Blanco-Rojo, C. Baeza, E. Arroyo-Pardo, M.P. Vaquero and A.M. López-Parra, A novel SNaPshot assay to detect genetic mutations related to iron metabolism, *Genet Test Mol Biomark* **15** (2011), 173-179.
- [17] B. Benyamin, M.A. Ferreira, G. Willemsen, S. Gordon, R.P. Middelberg, B.P. McEvoy, J.J. Hottenga, A.K. Henders, M.J. Campbell, L. Wallace, I.H. Frazer, A.C. Heath, E.J. de Geus, D.R. Nyholt, P.M. Visscher, B.W. Penninx, D.I. Boomsma, N.G. Martin, G.W. Montgomery and J.B. Whitfield, Common variants in TMPRSS6 are associated with iron status and erythrocyte volume, *Nat Genet* **41** (2009), 1173-1175.
- [18] J.C. Chambers, W. Zhang, Y. Li, J. Sehmi, M.N. Wass, D. Zabaneh, C. Hoggart, H. Bayele, M.I. McCarthy, L. Peltonen, N.B. Freimer, S.K. Sräi, P.H. Maxwell, M.J. Sternberg, A. Ruokonen, G. Abecasis, M.R. Jarvelin, J. Scott, P. Elliott and J.S. Kooner, Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels, *Nat Genet* **41** (2009), 1170-1172.
- [19] J.N. Feder, A. Gnirke, W. Thomas et al., A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis, *Nat Genet* **13** (1996), 399-408.
- [20] S.K. Ganesh, N.A. Zakai, F.J. van Rooij et al., Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium, *Nat Genet* **41** (2009), 1191-1198.
- [21] The International HapMap Project, *Nature* **426** (2003), 789-796.
- [22] R.D.C. Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.Rproject.org/>.
- [23] E. Marvez, S.J. Weiss, D.E. Houry and A.A. Ernst, Predicting adverse outcomes in a diagnosis-based protocol system for rapid sequence intubation, *Am J Emerg Med* **21** (2003), 23-29.
- [24] D. Delen, G. Walker and A. Kadam, Predicting breast cancer survivability: a comparison of three data mining methods, *Artif Intell Med* **34** (2005), 113-127.
- [25] I.S. Ríos, C. Bielza and A. Mateos, Fundamentos de los Sistemas de Ayuda a la Decisión, ed., Ra- Ma, Madrid, España, 2001.
- [26] M.T. Martín, M.V. Román and J.P. Lévy, Criterios de diferenciación de segmentos de predisposición tecnológica en el ámbito hospitalario público, *Ciencia Ergo Sum* **12** (2005), 125-132.
- [27] A.E. Raftery, Bayesian model selection in social research, *Sociological Methodology* **25** (1995), 111-163.
- [28] S.B. Gabriel, S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Fagart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S. Lander, M.J. Daly and D. Altshuler, The structure of haplotype blocks in the human genome, *Science* **296** (2002), 2225-2229.
- [29] M.J. Daly, J.D. Rioux, S.F. Schaffner, T.J. Hudson and E.S. Lander, High-resolution haplotype structure in the human genome, *Nat Genet* **29** (2001), 229-232.
- [30] L. McManus, K.B. Churchwell and K. Strange, Regulation of cell volume in health and disease, *N Engl J Med* **333** (1995), 1260-1266.
- [31] T. Tiffert, N. Daw, Z. Etzion, R.M. Bookchin and V.L. Lew, Age decline in the activity of the Ca<sup>2+</sup>-sensitive K<sup>+</sup> channel of human red blood cells, *J Gen Physiol* **129** (2007), 429-436.
- [32] F. Lang, Mechanisms and significance of cell volume regulation, *J Am Coll Nutr* **26** (2007), 613S-623S.
- [33] V.L. Lew, Z. Etzion and R.M. Bookchin, Dehydration response of sickle cells to sickling-induced Ca(++) permeabilization, *Blood* **99** (2002), 2578-2585.
- [34] P. Sebastiani, N. Solovieff, S.W. Hartley, J.N. Milton, A. Riva, D.A. Dworkis, E. Melista, E.S. Klings, M.E. Garrett, M.J. Telen, A. Ashley-Koch, C.T. Baldwin and M.H. Steinberg, Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study, *Am J Hematol* **85** (2010), 29-35.
- [35] K. Sugishita, M. Asakawa, S. Usui and T. Takahashi, A case of iron overload cardiomyopathy: beneficial effects of iron chelating agent and calcium channel blocker on left ventricular dysfunction, *Int Heart J* **50** (2009), 829-838.
- [36] J.A. Gaasch, W.J. Geldenhuys, P.R. Lockman, D.D. Allen and C.J. Van der Schyf, Voltage-gated Calcium Channels Provide an Alternate Route for Iron Uptake in Neuronal Cell Cultures, *Neurochem Res* **32** (2007), 1686-1693.
- [37] A.F. Remacha, M.P. Sarda, M.J. Barcelo et al., Genotype-phenotype correlation in a Spanish population homozygous for the H63D mutation of the HFE gene, *Ann Hematol* **85** (2006), 340-342.
- [38] M. Sánchez, M. Bruguera, J. Bosch, J. Rodes, F. Ballesta and R. Oliva, Prevalence of the Cys282Tyr and His63Asp HFE gene mutations in Spanish patients with hereditary hemochromatosis and in controls, *J Hepatol* **29** (1998), 725-728.
- [39] X. Du, E. She, T. Gelbart et al., The serine protease TMPRSS6 is required to sense iron deficiency, *Science* **320** (2008), 1088-1092.