Blood Cells, Molecules and Diseases 65 (2017) 60–65

Contents lists available at ScienceDirect

Blood Cells, Molecules and Diseases

journal homepage: www.elsevier.com/locate/bcmd

Genome-wide association study of erythrocyte density in sickle cell disease patients

Yann Ilboudo[a,b], Pablo Bartolucci[c], Alicia Rivera[d,h], Josepha-Clara Sedzro[f], Mélissa Beaudoin[b], Marie Trudel[e,f], Seth L. Alper[d,h], Carlo Brugnara[g], Frédéric Galactéros[c], Guillaume Lettre[a,b,*]

[a] Faculty of Medicine, Program in Bioinformatics, Université de Montréal, Montreal, Quebec, Canada
[b] Montreal Heart Institute, Montreal, Quebec, Canada
[c] Red Cell Genetic Disease Unit, Hôpital Henri-Mondor, Assistance Publique–Hôpitaux de Paris (AP-HP), Université Paris Est, IMRB - U955 - Equipe no 2, Créteil, France
[d] Division of Nephrology and Vascular Biology Research Center, Beth Israel Deaconess Medical Center, Boston, USA
[e] Faculty of Medicine, Department of Medicine and Department of Biochemistry, Université de Montréal, Montreal, Quebec, Canada
[f] Institut de recherches cliniques de Montréal, Montréal, Québec, Canada
[g] Department of Laboratory Medicine, Boston Children's Hospital, Boston, MA, USA
[h] Department of Medicine, Harvard Medical School, Boston, USA

ARTICLE INFO

Editor: Mohandas Narla

Keywords:
Dense red blood cells
Sickle cell disease
Genetic association study
GWAS
ATP2B4

ABSTRACT

Deoxy-hemoglobin S polymerization into rigid fibers is the direct cause of the clinical sequelae observed in sickle cell disease (SCD). The rate of polymerization of sickle hemoglobin is determined primarily by intracellular hemoglobin concentration, itself dependent on the amount of sickle hemoglobin and on red blood cell (RBC) volume. Dense, dehydrated RBC (DRBC) are observed in SCD patients, and their number correlates with hemolytic parameters and complications such as renal dysfunction, leg ulcers and priapism. To identify new genes involved in RBC hydration in SCD, we performed the first genome-wide association study for DRBC in 374 sickle cell anemia (HbSS) patients. We did not find genome-wide significant results, indicating that variants that modulate DRBC have modest-to-weak effects. A secondary analysis demonstrated a nominal association ($P = 0.003$) between DRBC in SCD patients and a variant associated with mean corpuscular hemoglobin concentration (MCHC) in non-anemic individuals. This intronic variant controls the expression of *ATP2B4*, the main calcium pump in erythrocytes. Our study highlights ATP2B4 as a promising target for modulation of RBC hydration in SCD patients.

1. Introduction

Sickle cell disease (SCD) is one of the most common monogenic diseases in the world. It is caused by a single mutation in the gene that encodes the β-chain of hemoglobin. Despite this genetic homogeneity, SCD patients are characterized by extreme clinical heterogeneity, ranging in presentation from benign mild anemia to devastating cerebrovascular events. Studies of the natural history of this blood disorder have improved clinical care such that most SCD patients in North America and Europe can now expect to reach middle age. Despite this progress, the life expectancy and quality-of-life of SCD patients is reduced, as treatment options remain limited, and no widely accessible curative therapy is available. Moreover, universal genetic screening and improved care for SCD have been slow to reach the sub-Saharan region, where the vast majority of SCD patients reside.

Results of seminal observational, epidemiological, biochemical, and genetic experiments have led to the emergence of fetal hemoglobin (HbF) as a key genetic modifier of severity in SCD [1]. The beneficial effects of hydroxyurea (HU), the only drug currently approved to treat SCD, are mediated in part by increasing HbF production. Dense, dehydrated erythrocytes are a hallmark of SCD patients, and red blood cell density (DRBC) has been investigated as a potential modifier of patient-to-patient clinical variability in SCD. Patients with elevated numbers of dense erythrocytes are expected to have clinical courses of greater severity, because the intracellular concentration of sickle hemoglobin (HbS) influences its rate of polymerization after deoxygenation [2]. Indeed, a study carried out in ~500 SCD patients showed that higher DRBC was associated with increased risk of leg ulcer, priapism, heart remodeling and renal dysfunction [3,4]. Interestingly, DRBC is not correlated with HbF, suggesting that therapeutic modulation of DRBC could further reduce complications when combined with HbF-stimulating agents such as HU.

* Corresponding author at: Montreal Heart Institute, 5000 Bélanger St, Montreal, Quebec H1T 1C8, Canada.
E-mail address: guillaume.lettre@umontreal.ca (G. Lettre).

http://dx.doi.org/10.1016/j.bcmd.2017.05.005
Received 25 March 2017; Received in revised form 10 May 2017; Accepted 10 May 2017
Available online 13 May 2017
1079-9796/

Several ion transporters and channels can control directly or indirectly RBC hydration (and thus density) [5]. Senicapoc, a selective inhibitor of the calcium-activated potassium Gardos channel, was shown in a mouse model of SCD to reduce the number of DRBC [6]. A phase III clinical trial of senicapoc in SCD patients similarly decreased the number of dense red blood cells, but failed to reduce the number of painful vaso-occlusive crises [7]. Strong interest nonetheless persists in the identification of novel drug targets, inhibition of which would selectively re-hydrate erythrocytes in SCD patients. Evidence of the pathologic importance of dehydration in SCD erythrocytes continues to accumulate [8]. Human genetics can provide an unbiased approach to discover the role of proteins and biological pathways in RBC hydration. In this article, we describe results from the first genome-wide association study (GWAS) to identify DNA sequence variants associated with DRBC in SCD patients.

2. Methods

2.1. Ethics statement

Informed consent was obtained for all participants in accordance with the Declaration of Helsinki. This project was also reviewed and approved by the Montreal Heart Institute Ethics Committee and the different recruiting centers.

2.2. Samples and DNA genotyping

The GEN-MOD study, a cohort of sickle cell anemia (HbSS) patients recruited in Paris, France, has been described elsewhere [3]. Hemoglobin profile, complete blood count, and red blood cell density (DRBC) determination using the phthalate density-distribution technique were all obtained at steady-state on the same day. 408 GEN-MOD participants were available for our genetic investigation. The DNA of the GEN-MOD participants was genotyped on the Illumina Infinium HumanOmni2.5Exome-8v1.1 array at the Montreal Heart Institute Pharmacogenomics Center. We used PLINK [9] and other custom scripts to control the quality of the genotyping dataset: we excluded samples and markers with genotyping success rate $< 95\%$, markers out of Hardy-Weinberg Equilibrium ($P < 1 \times 10^{-7}$) and markers with extreme (high or low) heterozygosity. We performed multidimensional scaling (MDS) in PLINK, anchoring these results on projections obtained using reference populations from the 1000 Genomes Project, to detect and remove (after visual inspection) population outliers. The Cooperative Study of Sickle Cell Disease (CSSCD) has been described extensively elsewhere [10–12]. Genome-wide genotype data generated with the Illumina Human610-Quad array was available for 1279 CSSCD participants. We conducted genotype imputation using Minimac3 (v1.0.11) [13] and reference haplotypes from phase 3 of the 1000 Genomes Project. We restricted association testing to markers with an imputation quality $r^2 > 0.3$.

2.3. Statistical analyses

The descriptive statistics of the participants analyzed in this study are presented in Table 1. Continuous phenotypes (DRBC and mean corpuscular hemoglobin concentration (MCHC)) were adjusted for sex and age, and the residuals were normalized using inverse normal transformation. Because low MCHC can be confounded by the thalassemia trait, we excluded from the analyses participants with α-thalassemia or a mean corpuscular hemoglobin (MCH) < 26 pg. We used linear regression for association testing between single variants and continuous traits, as implemented in RVtests (v.20140416) [14]. We used Sequence Kernel Association Test (SKAT) [15] and Variable Threshold (VT) [16] for our gene-based testing using rareMETALS (v.6.3) [17]. For gene-based testing, we focused our analysis on genotyped variants with minor allele frequency (MAF) $< 5\%$. We ran two sets of gene-based analyses: broad set (missense, nonsense, splice-site, frameshift and stop codon) and strict set (all of the above except missense variants). All genetic association analyses presented in this study were adjusted for the ten first principal components. Furthermore, we applied a genomic control correction to the DRBC GWAS results.

We defined genome-wide significance as $\alpha = 5 \times 10^{-8}$ and $\alpha = 2.5 \times 10^{-6}$ for single-variant and gene-based tests, respectively. In the post-hoc prioritization analyses (see below), we considered 12,360 erythroid enhancers ($\alpha = 4 \times 10^{-6}$ after Bonferroni correction) or expression quantitative trait loci (eQTL) for 66 candidate genes ($\alpha = 8 \times 10^{-4}$ after Bonferroni correction). For the 84 variants previously associated with MCHC by GWAS, and their linkage disequilibrium (LD) proxies, we highlighted variants with nominal significance ($\alpha = 0.05$) given the strong prior probability of these loci contributing to RBC hydration.

Table 1
Descriptive statistics of the GEN-MOD and CSSCD sickle cell disease participants analyzed in this study. For continuous variables, we provide the mean ± standard deviation and the number of participants with available data. NA, not available.

Phenotype	GEN-MOD (N = 408)	CSSCD (N = 1279)
Males/females	185/223	616/663
Age, years	30 ± 9	13 ± 12
DRBC, %	13.1 ± 8.6	NA
MCHC, g/dL	34.5 ± 1.8	34.6 ± 1.16

2.4. Genetic and functional prioritization of genetic variants

Given the limited statistical power offered by our sample size, we sought to prioritize variants using independent genetic and functional genomic information. In GEN-MOD, DRBC is strongly correlated with MCHC (Pearson's $r = 0.63$, $P = 7 \times 10^{-41}$, Supplementary Figs. 1–2). Although MCHC is not a perfect proxy for DRBC, variants associated with RBC dehydration are expected to result in increased MCHC. Since hemoglobin concentration is one of the major factors influencing sickle hemoglobin (HbS) polymerization [2], we tested the association of the top-scoring DRBC variants ($P_{DRBC} < 1 \times 10^{-6}$) for association with MCHC in GEN-MOD and the CSSCD. We also tested whether the variants associated with MCHC in a large genome-wide association study (GWAS) of European-ancestry non-anemic individuals [18] are associated with DRBC in SCD participants from GEN-MOD. For this lookup, we considered not only the sentinel MCHC GWAS variants, but also all variants in strong LD ($r^2 > 0.8$) in European populations from the 1000 Genomes Project.

We also prioritized variants that map to erythroid enhancers defined using DNase I hypersensitive sites and histone tail modifications [19]. Finally, we queried the GTEx database [20] to retrieve eQTL for 66 candidate genes. These genes were pre-selected based on their known and suspected roles in erythrocyte hydration. Supplementary Table 1 lists these candidate genes and rationales for their inclusion in the study.

3. Results

3.1. Genome-wide association study of red blood cell density

After quality-control and genotype imputation, we performed a genome-wide association study (GWAS) between ~31 million DNA sequence variants and red blood cell density (DRBC) in 374 sickle cell disease (SCD) patients from the GEN-MOD cohort (Table 1). DRBC was measured at steady-state during recruitment. For 26 GEN-MOD participants, we also repeated the measure on average 3.2 years later (range: 0.5–9 years) and found no significant difference (paired *t*-test $P = 0.79$, $r^2 = 0.65$), suggesting that DRBC is a stable phenotype. Although our

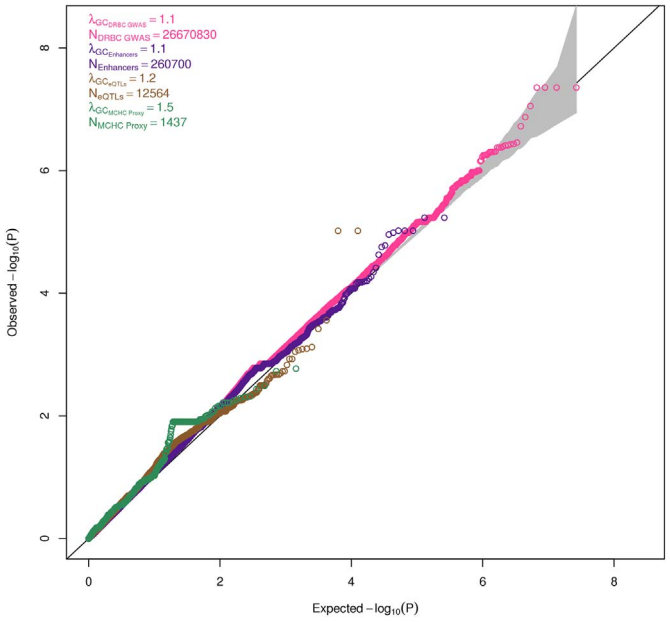

Fig. 1. Distribution of genome-wide association results with red blood cell density (DRBC) in 374 sickle cell disease patients. We present results for all imputed markers (pink), markers that map to erythroid enhancers (purple), markers that are expression quantitative trait loci (eQTL) for 66 candidate genes implicated in red blood cell hydration (brown), and markers associated with mean corpuscular hemoglobin concentration (MCHC) from previous genome-wide association studies (green). The grey area corresponds to the 95% confidence interval. λ_{GC}, genomic inflation factor.

single variant analysis was adjusted using principal components, we noted a modest inflation of the test statistics ($\lambda_{GC} = 1.1$, Fig. 1). For this reason, we corrected the test statistics using genomic control. Table 2 presents results for loci and associated variants with $P_{DRBC} < 5 \times 10^{-6}$. Gene-based testing focused on directly genotyped coding variants with minor allele frequencies (MAF) < 5% identified no significant association with DRBC.

The gold standard validation of genetic association studies requires replication of the initially observed associations for the same phenotype and variant in an independent cohort. Unfortunately, we are unaware of any SCD cohorts of sufficient size to replicate our DRBC genetic results. For this reason, we explored the use of mean corpuscular hemoglobin concentration (MCHC) as a surrogate phenotype. DRBC and MCHC are highly correlated in SCD patients (Supplementary Fig. 1), and high DRBC and MCHC each can reflect erythrocyte dehydration. Thus, a variant associated with DRBC might be predicted also to associate with MCHC.

First, we tested within GEN-MOD itself the association between the top-scoring variants associated with DRBC and MCHC. As expected for two correlated traits tested in the same individuals, several variants are associated with both DRBC and MCHC in GEN-MOD (Table 2). As an independent validation step, we performed the MCHC genetic association analysis in the Cooperative Study of Sickle Cell Disease (CSSCD). After excluding participants with α-thalassemia, which may independently affect MCHC, we identified 584 CSSCD participants with baseline MCHC and genotype data available. Only one of the 15 tested variants with $P_{DRBC} < 5 \times 10^{-6}$ in GEN-MOD also had a $P_{MCHC} < 0.05$ and consistent direction of effect in the CSSCD: this variant, rs59264502, is common (MAF = 46%) and intergenic (Table 2).

3.2. Variant prioritization

We implemented three strategies to increase the probability of finding robust genetic associations with DRBC. First, we considered potentially regulatory variants that map to erythroid enhancers, as defined by DNase I hypersensitive sites and histone modifications [19]. Among the 12,360 regulatory elements tested, we found no variants more strongly associated with DRBC than would be expected by chance (Fig. 1). Second, we retrieved from the GTEx resource [20] expression quantitative trait loci (eQTL) for 66 candidate genes, selected because they encode proteins with direct or indirect effects on red blood cell hydration (Supplementary Table 1). Three of these genes had eQTLs that were also associated with DRBC in SCD patients from GEN-MOD (at $P_{DRBC} < 8 \times 10^{-4}$, Bonferroni correction for 66 genes), although none were significantly associated with MCHC in the CSSCD (Table 2). These three promising variants control the expression of the Mg^{2+} transporter *SLC41A3*, the cytoskeletal protein *SPTB* (β-spectrin), and the mechanosensitive cation channel *PIEZO1*.

Our final strategy to prioritize variants was to exploit the physiological link between DRBC and MCHC. We reasoned that some variants previously associated with MCHC by GWAS could also influence DRBC. A recent meta-analysis carried out in 173,480 participants of European ancestry identified 84 DNA sequence variants robustly associated with MCHC [18]. To accommodate ethnicity difference, we retrieved DRBC results for these 84 variants as well as for all variants in strong linkage disequilibrium (LD, $r^2 > 0.8$ in European-ancestry individuals from the 1000 Genomes Project). This query highlighted eight variants with $P_{DRBC} < 0.05$ (Table 3).

One of these eight variants, rs1203972, is located near the α-globin locus on chromosome 16. This is promising since the presence of α-thalassemia is associated with fewer DRBC [3], although it is unknown whether this specific SNP is in LD with an α-thalassemia mutation. The most common cause of α-thalassemia in individuals of African ancestry is a 3.7-kb deletion that encompasses one of the genes (*HBA2*) encoding the α-chain of hemoglobin. Analyses of whole-genome sequence data from African populations in the 1000 Genomes Project showed this deletion is in LD with rs13335629 [21]. However, rs1203972 and rs13335629 are not in LD in GEN-MOD ($r^2 = 0.02$), nor is rs13335629 associated with DRBC ($P_{DRBC} = 0.24$).

3.3. ATP2B4 and DRBC in SCD patients

The second interesting result arising from this analysis of MCHC-associated SNPs in the DRBC GWAS data is an intronic SNP at the *ATP2B4* locus. *ATP2B4*, also known as PMCA4, encodes the main calcium pump of erythrocytes. We recently showed that this SNP, rs10751450, strongly associated with MCHC in European populations [18] and with malaria susceptibility in African populations [22], is an erythroid-specific eQTL for *ATP2B4* [23].

4. Discussion

DRBC has emerged as a promising biomarker of SCD clinical severity, but also an interesting target for therapeutic interventions [3,7,8]. Motivated by the recent genetic success that led to the identification of BCL11A and its role in controlling HbF production [24], we performed the first GWAS for DRBC in SCD patients. Given its limited sample size, our study design would have allowed us to identify only genetic variants of strong phenotypic effect. In the absence of such results, we can conclude that inter-individual variation in DRBC levels is controlled by common variants of weak effects and/or rare variants of modest-to-strong effects, a prediction that is consistent with the genetic architecture of most complex human traits. Performing genetic investigations with an increased number of SCD patients with DRBC measures available represents the most straightforward strategy to identify genetic variants that modulate erythrocyte density.

Despite the absence of genome-wide significant associations, biologically-informed *post hoc* analyses allowed us to prioritize candidate variants for future investigations. Using the GTEx resource, we found two common variants that are associated with DRBC in SCD patients and the expression of two strong candidate genes (Table 2). Indeed, rare

Table 2
Top single variant association results with red blood cell density (DRBC) in 374 participants from GEN-MOD. We included in this table variants with $P_{DRBC} < 5 \times 10^{-6}$ or variants that are expression quantitative trait loci (eQTL) for candidate genes in GTEx and have a $P_{DRBC} < 8 \times 10^{-4}$ (Methods). Chr:Pos, genomic coordinates on build hg19; REF/ALT, reference and alternate alleles; AF, frequency of the alternate allele; BETA(SE), effect size (for the alternate allele) and standard error in standard deviation units.

rsID	Chr:Pos	REF/ALT	GEN-MOD, DRBC (N = 374)			GEN-MOD, MCHC (N = 317)		CSSCD, MCHC (N = 584)			Gene	Annotation
			AF	BETA (SE)	P-value	BETA (SE)	P-value	AF	BETA (SE)	P-value		
Top association results												
rs4234795	4:7210802	A/G	0.94	− 0.84 (0.15)	1.99×10^{-7}	− 0.46 (0.17)	0.0062	0.94	0.03 (0.13)	0.80	*SORCS2*	Intron
rs9714060	3:195487476	A/G	0.43	− 0.39 (0.08)	7.43×10^{-7}	− 0.11 (0.08)	0.19	0.44	− 0.01 (0.07)	0.93	*MUC4*	Intron
rs146893001	9:112181617	T/C	0.01	− 2.04 (0.4)	1.29×10^{-6}	− 1.33 (0.6)	0.028	0.004	− 0.06 (0.48)	0.90	*PTPN3*	Intron
rs7216169	17:5219511	C/T	0.22	0.45 (0.09)	1.36×10^{-7}	0.25 (0.1)	0.0087	0.22	0.05 (0.08)	0.54	*RABEP1*	Intron
rs543023132	6:155973785	GTTTT/G	0.02	− 1.54 (0.3)	1.37×10^{-6}	− 0.68 (0.36)	0.061	0.022	− 0.15 (0.21)	0.47	–	Intergenic
rs144995469	14:57199082	C/T	0.03	− 1.15 (0.23)	1.48×10^{-6}	− 0.32 (0.26)	0.22	0.033	0.36 (0.18)	0.039	–	Intergenic
rs74989317	21:35296139	T/A	0.04	− 0.99 (0.2)	1.53×10^{-6}	− 0.42 (0.2)	0.041	0.045	− 0.21 (0.15)	0.17	*LINC00649*	Intron
rs73108077	20:30006859	T/C	0.06	− 0.83 (0.17)	1.75×10^{-6}	− 0.22 (0.2)	0.27	0.063	− 0.07 (0.13)	0.58	*DEFB122*	Downstream
rs114402357	13:22493635	C/T	0.01	2.03 (0.4)	1.78×10^{-6}	1.1 (0.45)	0.015	0.016	0.28 (0.25)	0.26	–	Intergenic
rs77141833	1:159825190	T/C	0.03	− 1.12 (0.22)	$1.80 \times 10–^{6}$	− 0.44 (0.28)	0.12	0.032	0.08 (0.18)	0.66	*VSIG8*	Intron
rs62015549	15:71671418	C/T	0.01	− 2.44 (0.49)	1.89×10^{-6}	− 1.07 (0.73)	0.15	0.015	− 0.25 (0.26)	0.33	*THSD4*	Intron
rs76513454	1:218861569	G/C	0.01	− 2.17 (0.43)	1.97×10^{-6}	− 0.93 (0.73)	0.20	NA	NA	NA	–	Intergenic
rs139628543	2:239053045	A/C	0.06	0.75 (0.15)	1.99×10^{-6}	0.22 (0.17)	0.19	0.05	0.08 (0.14)	0.59	*KLHL30*	Intron
rs59264502	13:106846272	AT/A	0.46	0.37 (0.08)	2.39×10^{-6}	0.21 (0.08)	0.011	0.47	0.14 (0.06)	0.030	–	Intergenic
rs147900370	1:115552925	A/C	0.04	− 0.92 (0.18)	2.44×10^{-6}	− 0.35 (0.21)	0.0905849	0.038	− 0.09 (0.17)	0.59	–	Intergenic
eQTL for candidate genes												
rs62270871	3:125672365	G/A	0.51	0.33 (0.07)	2.60×10^{-5}	0.14 (0.08)	0.11	0.47	0.02 (0.07)	0.71	*ALG1L*	Intron; eQTL for *SLC41A3*
rs146977005	14:65305030	G/GA	0.28	− 0.33 (0.09)	5.5×10^{-4}	− 0.10 (0.09)	0.26	0.75	0.11 (0.07)	0.15	*SPTB*	Intron; eQTL for *SPTB*
rs8048714	16:88809773	G/C	0.72	− 0.3 (0.08)	7.3×10^{-4}	0.08 (0.1)	0.45	0.25	0.14 (0.07)	0.056	*PIEZO1*	Intron; eQTL for *PIEZO1*

Table 3
Top association results between variants previously associated with mean corpuscular hemoglobin concentration (MCHC) in non-anemic European-ancestry individuals and red blood cell density in 374 sickle cell disease patients. We included in this table variants with nominal $P_{DRBC} < 0.05$. Chr:Pos, genomic coordinates on build hg19; REF/ALT, reference and alternate alleles; AF, frequency of the alternate allele; BETA (SE), effect size (for the alternate allele) and standard error in standard deviation units.

rsID	Chr:Pos	REF/ALT	GEN-MOD, DRBC (N = 374)			Gene	Annotation
			AF	BETA (SE)	P-value		
rs144514173	1:205246482	TTTTG/T	0.11	0.37 (0.12)	0.0029	*TMCC2*	Downstream
rs10751450	1:203650945	C/T	0.64	− 0.25 (0.08)	0.0031	*ATP2B4*	Intron
rs148303943	6:16263455	T/C	0.85	− 0.32 (0.11)	0.0057	*GMPR*	Intron
rs11421513	6:13901073	G/GT	0.69	− 0.23 (0.08)	0.0074	–	Intergenic
rs1203972	16:283232	T/C	0.66	− 0.22 (0.08)	0.0082	*LUC7L*	Upstream
rs201794926	8:145710909	G/GA	0.49	0.18 (0.07)	0.021	*PPP1R16A*	Intron
rs34514965	19:13071559	T/TG	0.83	0.21 (0.1)	0.043	*GADD45GIP1*	Upstream
rs5875087	6:26118437	CA/C	0.91	− 0.27 (0.13)	0.045	*HIST1H2BC*	Intron

Mendelian mutations in *SPTB* and PIEZO1 cause hereditary spherocytosis and xerocytosis, respectively, two disorders of RBC volume homeostasis. Whether these two promising associations represent true signals will require replication in independent samples.

Additionally, our study showed that DRBC and HbF are independent traits. Indeed, they were not correlated (Pearson's $r = -0.09$, $P = 0.088$). Therefore, when we tested common DNA sequence variants at the *BCL11A*, *HBS1L-MYB* and the *β-globin* loci known to strongly influence HbF levels [1], we found no significant associations with DRBC ($P > 0.05$). However, we noticed a nominal association between a low-frequency intronic variant in *HBS1L* (rs116460276, MAF = 4%) and DRBC ($P = 0.009$). This association might simply reflect the pleiotropy of this locus, which is associated with HbF but also with several other RBC parameters [18], likely through an effect on *MYB*, a key transcription factor during erythropoiesis [25]. Thus, it remains entirely possible that a dual therapeutic intervention that would aim to increase HbF as well as reduce DRBC would show additive benefits in improving SCD clinical outcomes.

Interestingly, we found an association between an intronic variant in *ATP2B4* and DRBC in SCD patients. The same association signal is strongly associated with MCHC in non-anemic Europeans and malaria susceptibility in Africans [18,22]. Furthermore, we have functionally validated that this variant controls *ATP2B4* expression in human erythroblasts [23]. *ATP2B4* encodes for PMCA4, the main RBC plasma membrane ATPase pump that transports Ca^{2+} ions outside the cells. Intracellular calcium concentration is important for RBC volume control since the Gardos channel, which is the principal channel that regulates potassium and water loss, is calcium-dependent. A clinical trial of the Gardos channel inhibitor, senicapoc, in SCD patients improved RBC hydration but failed to meet the mandated clinical endpoint of decreased pain crisis [7]. Our results highlight ATP2B4/PMCA4 as an alternate target for strategies to minimize RBC dehydration by regulating Ca^{2+} concentration without directly interfering with the Gardos channel itself. This would require the development of an ATP2B4/PMCA4-specific activator, or a strategy to augment *ATP2B4* expression in RBC. It might also be worth exploring how genetic variation at the *ATP2B4* locus, which is associated with DRBC in SCD patients, could be used as a companion pharmacogenetic test to identify patients more likely to benefit from new drugs that would increase RBC hydration. This could easily be implemented within future clinical trials.

Funding sources

G.L. is funded by the Biogen Sickle Cell Disease Consortium, the Canadian Institutes of Health Research (CIHR, MOP #123382), the Doris Duke Charitable Foundation, and the Canada Research Chair program. S.L.A. is funded by the Doris Duke Charitable Foundation. M.T. is funded by the CIHR/Canadian Blood Services (MOP #3251163).

Author contributions

Y.I. and G.L. conceived and designed the experiments; Y.I., A.R., J.-C.S. and M.B. performed the experiments; P.B., M.T., S.L.A., C.B., and F.G. contributed DNA samples, clinical information, and expert knowledge; Y.I. and G.L. analyzed the results; Y.I. and G.L. wrote the manuscript with contributions from all authors.

Conflict of interest

The authors declare no competing financial interests.

Acknowledgments

We thank all participants for their contribution to this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bcmd.2017.05.005.

References

[1] G. Lettre, D.E. Bauer, Fetal haemoglobin in sickle-cell disease: from genetic epidemiology to new therapeutic strategies, Lancet 387 (2016) 2554–2564.
[2] W.A. Eaton, J. Hofrichter, Sickle cell hemoglobin polymerization, Adv. Protein Chem. 40 (1990) 63–279.
[3] P. Bartolucci, C. Brugnara, A. Teixeira-Pinto, S. Pissard, K. Moradkhani, H. Jouault, F. Galacteros, Erythrocyte density in sickle cell syndromes is associated with specific clinical manifestations and hemolysis, Blood 120 (2012) 3136–3141.
[4] T. Damy, D. Bodez, A. Habibi, A. Guellich, S. Rappeneau, J. Inamo, S. Guendouz, J. Gellen-Dautremer, S. Pissard, S. Loric, O. Wagner-Ballon, B. Godeau, S. Adnot, J.L. Dubois-Rande, L. Hittinger, F. Galacteros, P. Bartolucci, Haematological determinants of cardiac involvement in adults with sickle cell disease, Eur. Heart J. 37 (2016) 1158–1167.
[5] P.G. Gallagher, Disorders of red cell volume regulation, Curr. Opin. Hematol. 20 (2013) 201–207.
[6] J.W. Stocker, L. De Franceschi, G.A. McNaughton-Smith, R. Corrocher, Y. Beuzard, C. Brugnara, ICA-17043, a novel Gardos channel blocker, prevents sickled red blood cell dehydration in vitro and in vivo in SAD mice, Blood 101 (2003) 2412–2418.
[7] K.I. Ataga, M. Reid, S.K. Ballas, Z. Yasin, C. Bigelow, L.S. James, W.R. Smith, F. Galacteros, A. Kutlar, J.H. Hull, J.W. Stocker, I.C.A.S. Investigators, Improvements in haemolysis and indicators of erythrocyte survival do not correlate with acute vaso-occlusive crises in patients with sickle cell disease: a phase III randomized, placebo-controlled, double-blind study of the Gardos channel blocker senicapoc (ICA-17043), Br. J. Haematol. 153 (2011) 92–104.
[8] Q. Li, E.R. Henry, J. Hofrichter, J.F. Smith, T. Cellmer, E.B. Dunkelberger, B.B. Metaferia, S. Jones-Straehle, S. Boutom, G.W. Christoph, T.H. Wakefield, M.E. Link, D. Staton, E.R. Vass, J.L. Miller, M.M. Hsieh, J.F. Tisdale, W.A. Eaton, Kinetic assay shows that increasing red cell volume could be a treatment for sickle cell disease, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E689–E696.
[9] S. Purcell, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. de Bakker, M.J. Daly, P.C. Sham, PLINK: a toolset for whole-genome association and population-based linkage analysis, Am. J. Hum. Genet. 81 (2007).
[10] M. Gaston, J. Smith, D. Gallagher, Z. Flournoy-Gill, S. West, R. Bellevue, M. Farber, R. Grover, M. Koshy, A.K. Ritchey, et al., Recruitment in the cooperative study of sickle cell disease (CSSCD), Control. Clin. Trials 8 (1987) 131S–140S.
[11] M. Gaston, W.F. Rosse, The cooperative study of sickle cell disease: review of study

design and objectives, Am. J. Pediatr. Hematol. Oncol. 4 (1982) 197–201.
[12] N. Solovieff, J.N. Milton, S.W. Hartley, R. Sherva, P. Sebastiani, D.A. Dworkis, E.S. Klings, L.A. Farrer, M.E. Garrett, A. Ashley-Koch, M.J. Telen, S. Fucharoen, S.Y. Ha, C.K. Li, D.H.K. Chui, C.T. Baldwin, M.H. Steinberg, Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster, Blood 115 (2010) 1815–1822.
[13] S. Das, L. Forer, S. Schonherr, C. Sidore, A.E. Locke, A. Kwong, S.I. Vrieze, E.Y. Chew, S. Levy, M. McGue, D. Schlessinger, D. Stambolian, P.R. Loh, W.G. Iacono, A. Swaroop, L.J. Scott, F. Cucca, F. Kronenberg, M. Boehnke, G.R. Abecasis, C. Fuchsberger, Next-generation genotype imputation service and methods, Nat. Genet. 48 (2016) 1284–1287.
[14] X. Zhan, Y. Hu, B. Li, G.R. Abecasis, D.J. Liu, RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data, Bioinformatics 32 (2016) 1423–1426.
[15] M.C. Wu, S. Lee, T. Cai, Y. Li, M. Boehnke, X. Lin, Rare-variant association testing for sequencing data with the sequence kernel association test, Am. J. Hum. Genet. 89 (2011) 82–93.
[16] A.L. Price, G.V. Kryukov, P.I. de Bakker, S.M. Purcell, J. Staples, L.J. Wei, S.R. Sunyaev, Pooled association tests for rare variants in exon-resequencing studies, Am. J. Hum. Genet. 86 (2010) 832–838.
[17] D.J. Liu, G.M. Peloso, X. Zhan, O.L. Holmen, M. Zawistowski, S. Feng, M. Nikpay, P.L. Auer, A. Goel, H. Zhang, U. Peters, M. Farrall, M. Orho-Melander, C. Kooperberg, R. McPherson, H. Watkins, C.J. Willer, K. Hveem, O. Melander, S. Kathiresan, G.R. Abecasis, Meta-analysis of gene-level tests for rare variant association, Nat. Genet. 46 (2014) 200–204.
[18] W.J. Astle, H. Elding, T. Jiang, D. Allen, D. Ruklisa, A.L. Mann, D. Mead, H. Bouman, F. Riveros-Mckay, M.A. Kostadima, J.J. Lambourne, S. Sivapalaratnam, K. Downes, K. Kundu, L. Bomba, K. Berentsen, J.R. Bradley, L.C. Daugherty, O. Delaneau, K. Freson, S.F. Garner, L. Grassi, J. Guerrero, M. Haimel, E.M. Janssen-Megens, A. Kaan, M. Kamat, B. Kim, A. Mandoli, J. Marchini, J.H. Martens, S. Meacham, K. Megy, J. O'Connell, R. Petersen, N. Sharifi, S.M. Sheard, J.R. Staley, S. Tuna, M. van der Ent, K. Walter, S.Y. Wang, E. Wheeler, S.P. Wilder, V. Iotchkova, C. Moore, J. Sambrook, H.G. Stunnenberg, E. Di Angelantonio, S. Kaptoge, T.W. Kuijpers, E. Carrillo-de-Santa-Pau, D. Juan, D. Rico, A. Valencia, L. Chen, B. Ge, L. Vasquez, T. Kwan, D. Garrido-Martin, S. Watt, Y. Yang, R. Guigo, S. Beck, D.S. Paul, T. Pastinen, D. Bujold, G. Bourque, M. Frontini, J. Danesh, D.J. Roberts, W.H. Ouwehand, A.S. Butterworth, N. Soranzo, The allelic landscape of human blood cell trait variation and links to common complex disease, Cell 167 (2016) 1415–1429e1419.
[19] J. Xu, Z. Shao, K. Glass, D.E. Bauer, L. Pinello, B. Van Handel, S. Hou, J.A. Stamatoyannopoulos, H.K. Mikkola, G.C. Yuan, S.H. Orkin, Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis, Dev. Cell 23 (2012) 796–811.
[20] G.T. Consortium, The genotype-tissue expression (GTEx) project, Nat. Genet. 45 (2013) 580–585.
[21] Z. Chen, H. Tang, R. Qayyum, U.M. Schick, M.A. Nalls, R. Handsaker, J. Li, Y. Lu, L.R. Yanek, B. Keating, Y. Meng, F.J. van Rooij, Y. Okada, M. Kubo, L. Rasmussen-Torvik, M.F. Keller, L. Lange, M. Evans, E.P. Bottinger, M.D. Linderman, D.M. Ruderfer, H. Hakonarson, G. Papanicolaou, A.B. Zonderman, O. Gottesman, C. Thomson, E. Ziv, A.B. Singleton, R.J. Loos, P.M. Sleiman, S. Ganesh, S. McCarroll, D.M. Becker, J.G. Wilson, G. Lettre, A.P. Reiner, Genome-wide association analysis of red blood cell traits in African Americans: the COGENT Network, Hum. Mol. Genet. 22 (2013) 2529–2538.
[22] C. Timmann, T. Thye, M. Vens, J. Evans, J. May, C. Ehmen, J. Sievertsen, B. Muntau, G. Ruge, W. Loag, D. Ansong, S. Antwi, E. Asafo-Adjei, S.B. Nguah, K.O. Kwakye, A.O. Akoto, J. Sylverken, M. Brendel, K. Schuldt, C. Loley, A. Franke, C.G. Meyer, T. Agbenyega, A. Ziegler, R.D. Horstmann, Genome-wide association study indicates two novel resistance loci for severe malaria, Nature 489 (2012) 443–446.
[23] E.N. Stern, S. Lessard, P.G. Schupp, F. Sher, G. Lettre, D.E. Bauer, An essential erythroid-specific enhancer of ATP2B4 associated with red blood cell traits and malaria susceptibility, Blood 128 (2016) 1250.
[24] G. Lettre, V.G. Sankaran, M.A. Bezerra, A.S. Araujo, M. Uda, S. Sanna, A. Cao, D. Schlessinger, F.F. Costa, J.N. Hirschhorn, S.H. Orkin, DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 11869–11874.
[25] J. Palis, Primitive and definitive erythropoiesis in mammals, Front. Physiol. 5 (2014) 3.