

# Genetic determinants of haemolysis in sickle cell anaemia

Jacqueline N. Milton,<sup>1</sup> Helen Rooks,<sup>2</sup> Emma Drasar,<sup>2,3</sup> Elizabeth L. McCabe,<sup>1</sup> Clinton T. Baldwin,<sup>4</sup> Efi Melista,<sup>5</sup> Victor R. Gordeuk,<sup>6</sup> Mehdi Nouraie,<sup>6</sup> Gregory R. Kato,<sup>7</sup> Caterina Minniti,<sup>7</sup> James Taylor,<sup>7</sup> Andrew Campbell,<sup>8</sup> Lori Luchtman-Jones,<sup>9</sup> Sohail Rana,<sup>6</sup> Oswaldo Castro,<sup>7,10</sup> Yingze Zhang,<sup>11</sup> Swee Lay Thein,<sup>2,3</sup> Paola Sebastiani,<sup>1</sup> Mark T. Gladwin,<sup>11</sup> for the Walk-PHAASST Investigators<sup>11</sup> and Martin H. Steinberg<sup>4</sup>

<sup>1</sup>Biostatistics, Boston University, Boston, MA, USA, <sup>2</sup>King's College School of Medicine, London, UK, <sup>3</sup>Molecular Haematology, King's College London, London, UK, <sup>4</sup>Department of Medicine, Boston University School of Medicine, <sup>5</sup>Center for Human Genetics, Boston University, Boston, MA, <sup>6</sup>Center for Sickle Cell Disease, Howard University, Washington, DC, <sup>7</sup>Sickle Cell Vascular Disease Section, Hematology Branch, National Heart, Lung, and Blood Institute, Bethesda, MD, <sup>8</sup>Pediatric Hematology Oncology, University of Michigan, Ann Arbor, MI, <sup>9</sup>Hematology, Children's National Medical Center, <sup>10</sup>Center For Sickle Cell Disease, Howard University, Washington, DC, and <sup>11</sup>Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, USA

Received 23 October 2012; accepted for publication 29 December 2012

Correspondence: Martin H. Steinberg, Department of Medicine, Boston University School of Medicine, 72 E. Concord St., Boston, MA 02118, USA.

E-mail: mhsteinb@bu.edu

Drs. Gladwin and Steinberg contributed equally to this manuscript.

For the Walk-PHAASST Investigators detail see Appendix I.

The phenotype of sickle cell anaemia is caused by sickle vaso-occlusion and haemolytic anaemia (Kato *et al*, 2007). Haemolysis in this disease has been associated with complications that could result in part from vascular nitric oxide (NO) depletion due to scavenging by free plasma haemoglobin

## Summary

Haemolytic anaemia is variable among patients with sickle cell anaemia and can be estimated by reticulocyte count, lactate dehydrogenase, aspartate aminotransferase and bilirubin levels. Using principal component analysis of these measurements we computed a haemolytic score that we used as a subphenotype in a genome-wide association study. We identified in one cohort and replicated in two additional cohorts the association of a single nucleotide polymorphism in *NPRL3* (rs7203560; chr16p13.3) ( $P = 6.04 \times 10^{-07}$ ). This association was validated by targeted genotyping in a fourth independent cohort. The *HBA1/HBA2* regulatory elements, hypersensitive sites (HS)-33, HS-40 and HS-48 are located in introns of *NPRL3*. Rs7203560 was in perfect linkage disequilibrium (LD) with rs9926112 ( $r^2 = 1$ ) and in strong LD with rs7197554 ( $r^2 = 0.75$ ) and rs13336641 ( $r^2 = 0.77$ ); the latter is located between HS-33 and HS-40 sites and next to a CTCF binding site. The minor allele for rs7203560 was associated with the  $-\alpha^{3.7}$  thalassaemia gene deletion. When adjusting for HbF and  $\alpha$  thalassaemia, the association of *NPRL3* with the haemolytic score was significant ( $P = 0.00375$ ) and remained significant when examining only cases without gene deletion  $\alpha$  thalassaemia ( $P = 0.02463$ ). Perhaps by independently down-regulating expression of the *HBA1/HBA2* genes, variants of the *HBA1/HBA2* gene regulatory loci, tagged by rs7203560, reduce haemolysis in sickle cell anaemia.

**Keywords:** haemolysis, sickle cell anaemia, haemolytic anaemia, genetic analysis, thalassaemia.

(Gladwin *et al*, 2004; Morris *et al*, 2005; Taylor *et al*, 2008). Plasma haemoglobin is a specific marker of intravascular haemolysis and red cell survival studies are the definitive measurement of haemolysis. Nevertheless, these tests are rarely done and are not available in large cohorts. However,

haemolysis can be estimated by the reticulocyte count, lactate dehydrogenase (LDH), aspartate aminotransferase (AST) and bilirubin levels, all of which are commonly measured in cohort studies, although none is specific for haemolysis (Kato *et al*, 2007; Hebbel, 2011).

High concentrations of foetal haemoglobin (HbF) decrease the polymerization tendency of sickle haemoglobin (HbS) (reviewed in Akinsheye *et al*, 2011). In sickle cell anaemia with concurrent  $\alpha$  thalassaemia, the concentration of HbS in sickle erythrocytes is reduced, decreasing its polymerization tendency (reviewed in Steinberg & Embury, 1986). Both high HbF and  $\alpha$  thalassaemia increase the lifespan of the sickle erythrocyte (De Ceulaer *et al*, 1983; Steinberg & Sebastiani, 2012). However, other genes might also affect cell life span.

We estimated the severity of haemolysis using a principal component analysis of the commonly measured markers of haemolysis (Gordeuk *et al*, 2009; Minniti *et al*, 2009). The development of such a component resolves the problem of dealing with correlated predictors in multivariate analyses, allows for adjustment of important confounders, such as gender, hydroxycarbamide use, age, haemoglobin levels, HbF,  $\alpha$  thalassaemia and site variability in laboratory assay protocols and standards. The haemolytic score was associated with intravascular haemolysis as measured by plasma haemoglobin and red cell micro particles and had a significant inverse relationship to total haemoglobin levels, HbF and  $\alpha$  thalassaemia (Nouraei *et al*, 2012). Therefore, the score can be used as a robust population-based continuous measure of haemolytic rate and is suited for studies of genetic factors that regulate the intensity of haemolytic anaemia in sickle cell anaemia. We used this score as a phenotype in a genome-wide association study (GWAS).

## Methods

### Study participants

The patient cohort used for discovery consisted of 1 117 patients from the Cooperative Study of Sickle Cell Disease (CSSCD; NCT00005277) homozygous for the HbS gene or with HbS- $\beta^0$  thalassaemia. Four-hundred and forty-nine patients from the Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy (Walk-PHaSST) study (NCT00492531) and 296 patients from the Pulmonary Hypertension and the Hypoxic Response in Sickle Cell Disease (PUSH) study (NCT 00495638) with similar haemoglobin phenotypes as the discovery set were used for replication of the discovery findings. The demographics of these studies have been described (Gaston & Rosse, 1982; Dham *et al*, 2009; Machado *et al*, 2011). For further validation, targeted genotyping was performed in a third cohort of 213 additional patients with sickle cell disease from London, UK. These studies were approved by the Institutional Review Boards of each participating institution.

### Phenotype

We used principal component analysis to derive a haemolytic score from reticulocyte count, LDH, AST and bilirubin using the most informative component as previously suggested (Gordeuk *et al*, 2009; Minniti *et al*, 2009). This score is a linear combination of the 4 haemolytic variables with mean of 0.

Serum total bilirubin, LDH, AST, and reticulocyte counts were measured using automated chemical and haematologic analysers. For the CSSCD patients, longitudinal bilirubin measurements were collected from phases 1, 2, and 3 of the study. Only steady state measurements 4 months removed from blood transfusion were used. The longitudinal measurements of these study patients were analysed using a Bayesian hierarchical mixed model that included a random effect per patient to account for the repeated measurements, as well as random intercept and age effects that were allowed to vary with the clinics. The random intercept and age effects were used to remove the between-site systematic differences. Markov Chain Monte Carlo method in Open bugs was used to estimate the predicted values, and log-transformed median predicted values for each of the 4 haemolytic markers were used in the principal component analysis (Milton *et al*, 2012). For the PUSH and Walk-PHaSST patients, log transformed baseline values were used in the principal component analysis. Patient characteristics for the discovery, replication and validation cohorts are shown in Table I.

### Heritability of haemolysis

To examine heritability of the 4 haemolytic markers and the haemolytic score in the CSSCD population, we examined their correlation in 90 sibling pairs that could be identified by identity by descent (IBD) analysis in PLINK (Purcell *et al*, 2007) using the genome-wide single nucleotide polymorphism (SNP) data. As a comparison, we randomly selected 200 unrelated pairs 1 000 times and computed average correlation the 4 haemolytic markers and the haemolytic score among unrelated individuals.

### Genotyping

DNA from the CSSCD, PUSH and Walk-PHaSST samples that formed the discovery and replication cohorts were genotyped using Illumina Human610-Quad SNP arrays (Illumina, San Diego, CA, USA) with approximately 600 000 SNPs. All samples were processed according the manufacturer's protocol and Bead Studio Software was used to make genotype calls utilizing the Illumina pre-defined clusters. Samples with less than a 95% call rate were removed and SNPs with a call rate <97.5% were re-clustered. After re-clustering, SNPs with call rates >97.5%, cluster separation score >0.25, excess heterozygosity between -0.10 and 0.10, and minor allele frequency >5% were retained in the analysis. We used the

**Table I.** Patient characteristics.

Variable	CSSCD ( <i>n</i> = 1,117)			Walk-PHaSST* ( <i>n</i> = 449)		
	Overall	M ( <i>n</i> = 585)	F ( <i>n</i> = 532)	Overall	M ( <i>n</i> = 207)	F ( <i>n</i> = 242)
Age	17.86 (11.71)	18.85 (11.98)	16.79 (11.31)	37.20 (13.20)	35.81 (12.94)	38.35 (13.34)
Haemoglobin	2.11 (0.14)	2.10 (0.13)	2.12 (0.15)	2.24 (0.69)	2.26 (0.80)	2.23 (0.55)
Haematocrit	3.18 (0.15)	3.17 (0.14)	3.18 (0.17)	3.08 (2.51)	3.12 (2.54)	3.01 (2.51)
HbF	1.80 (0.52)	1.86 (0.51)	1.73 (0.52)	1.64 (1.48)	1.57 (1.49)	1.69 (1.46)
α thalassaemia	370 (33.3%)	189 (32.3%)	181 (34%)	129 (30.1%)	56 (28.4%)	73 (31.6%)
AST	3.74 (0.29)	3.81 (0.27)	3.67 (0.20)	3.80 (1.35)	3.91 (0.53)	3.70 (0.50)
LDH	6.02 (0.22)	5.98 (0.21)	6.06 (0.23)	6.08 (0.59)	6.00 (0.59)	6.16 (0.96)
Reticulocyte	2.36 (0.40)	2.37 (0.39)	2.35 (0.40)	2.11 (0.73)	2.15 (0.73)	2.08 (0.73)
Bilirubin	1.13 (0.51)	1.19 (0.51)	1.08 (0.51)	2.48 (1.35)	2.67 (1.38)	2.35 (1.32)
Haemolytic Score†	$-2.97 \times 10^{-18}$ (1.08)	0.20 (1.53)	-0.22 (1.49)	$-2.18 \times 10^{-16}$ (1.00)	0.29 (1.59)	-0.23 (1.51)

Variable	PUSH ( <i>n</i> = 296)			London ( <i>n</i> = 213)		
	Overall	M ( <i>n</i> = 140)	F ( <i>n</i> = 156)	Overall	M ( <i>n</i> = 94)	F ( <i>n</i> = 119)
Age	11.85 (5.42)	12.04 (5.25)	11.65 (5.61)	33.02 (10.89)	32.30 (10.55)	33.59 (11.17)
Haemoglobin	2.22 (0.57)	2.23 (0.69)	2.22 (0.39)	2.11 (0.18)	2.17 (0.18)	2.08 (0.16)
Haematocrit	3.28 (1.69)	3.28 (1.81)	3.28 (1.52)			
HbF	2.20 (1.98)	2.09 (1.96)	2.32 (1.97)	1.79 (0.48)	1.67 (0.47)	1.89 (0.46)
α thalassaemia	38 (32.2%)	51 (43.22%)	29 (24.58%)	75 (35%)	28 (30%)	47 (39%)
AST	3.85 (0.40)	3.89 (0.40)	3.81 (0.39)	3.71 (0.32)	3.78 (0.33)	3.66 (0.31)
LDH	6.17 (0.48)	6.17 (0.43)	6.17 (0.53)	6.04 (0.32)	6.04 (0.35)	6.04 (0.31)
Reticulocyte	2.17 (0.73)	2.17 (0.80)	2.17 (0.67)	5.81 (0.43)	5.89 (0.32)	5.74 (0.48)
Bilirubin	0.99 (0.65)	1.08 (0.70)	0.91 (0.58)	3.86 (0.56)	4.00 (0.58)	3.75 (0.52)
Haemolytic Score	$-7.17 \times 10^{-16}$ (1.08)	0.21 (1.62)	-0.22 (1.50)	$1.82 \times 10^{-16}$ (1.06)	0.29 (1.37)	-0.23 (1.32)

Summary statistics of patient characteristics in the CSSCD discovery cohort, the PUSH and WALK-PHaSST replication cohorts and in the London validation cohort. For each cohort, the first column reports statistics (mean and standard deviate or frequencies) for all patients included in the analysis and the second and third columns report statistics stratified by gender. Aspartate transaminase (AST), lactate dehydrogenase (LDH), reticulocyte count, bilirubin, haemoglobin and haematocrit are reported as log transformed values in the Walk-PHaSST and PUSH cohorts and sex, age and clinic adjusted values for the CSSCD cohort. M-male, F-female. HbF values are reported as cubic root transformed values in the Walk-PHaSST and PUSH cohorts and sex, age and clinic adjusted values for the CSSCD cohort.

\*Walk-PHaSST variables measured in SI units.

†The haemolytic score is a linear combination of the 4 haemolytic variables with mean of 0. Males have higher values than females in all 4 cohorts studied.

genome-wide identity by descent analysis in PLINK to discover unknown relatedness. Pairs with IBD measurements greater than 0.2 were deemed to be related and related subjects within individual or different studies were removed. We also removed samples with inconsistent gender findings defined by heterozygosity of the X chromosome and gender recorded in the database.

Samples from the validation cohort were genotyped using predesigned TaqMan SNP genotyping assays according to standard Applied Biosystems protocol (Applied Biosystems, Carlsbad, CA, USA). For each sample, 10 ng DNA was dried in a 384-well plate, upon which was added a 5 µl reaction volume (2.5 µl TaqMan universal polymerase chain reaction-master mix, 2.375 µl Sigma water, 0.125 µl Assay). The samples were run on the ABI7900 under the following thermal cycling conditions: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 1 min. For each assay run, 12 control samples and 3 non template controls were

included. Genotype data was exported from Sequence Detection System (SDS) 2.2 software (Applied Biosystems).

In all studies, the presence of gene deletion α thalassaemia was directly ascertained by restriction endonuclease analysis or by multiplex polymerase chain reactions (Dozy *et al*, 1979; Liu *et al*, 2000).

### Analysis

The association between haemolytic score and each SNP was tested with multiple linear regression adjusting for age and sex using an additive model in PLINK (Purcell *et al*, 2007). Age and gender were both included as covariates as both were significantly associated with the haemolytic score (age *P*-value =  $3.40 \times 10^{-11}$ , sex *P*-value =  $1.74 \times 10^{-11}$ ). The minor allele was used as the coded allele in the additive model. To assess population stratification in the CSSCD cohort the genomic control lambda factor was calculated

using PLINK. A meta-analysis was also performed with all 4 cohorts with Meta Analysis Helper (METAL) using the inverse variance weighting method (<http://www.sph.umich.edu/csg/abecasis/metal>). To determine if there was an association between our top genetic variant and  $\alpha$  thalassaemia, a  $\chi^2$  test was implemented in the CSSCD cohort.

To determine the exact coordinates of the hypersensitive site region in chromosome 16, we used the Basic Local Alignment Search Tool (BLAST; Altschul *et al*, 1990) and VISTA (Mayor *et al*, 2000) to compare and align sequences in regions containing the regulatory regions of the  $\alpha$ -globin cluster using methods described by Hughes *et al* (2005). Phylogenetic conservation at, or near, the hypersensitive sites was also determined (Hughes *et al*, 2005).

## Results

### Patient characteristics

Table I shows demographics, haematological data, laboratory measures of haemolysis and the derived haemolytic score for all 4 cohorts. Males had a higher haemolytic score compared with females. Participants in the Walk-PHaSST study were adults and had the highest mean age, patients in the PUSH study were children with the lowest mean age while CSSCD cases included both adults and children and had intermediate ages. There was a significant association between age and the haemolytic score ( $r = 0.195$ ,  $P$ -value =  $3.40 \times 10^{-11}$ ) and this was adjusted for in the analysis. Cooperative study of sickle cell disease cases did not take hydroxycarbamide, 58% of Walk-PHaSST and 44% PUSH patients were treated with hydroxycarbamide.

The first principal component that was calculated explained 67.4% of the total variance and this measurement was used as the haemolytic score in the CSSCD cohort. When comparing the association between the haemolytic score and each of the 4 markers of haemolysis, as each of the 4 haemolytic markers increased the haemolytic score increased (Fig S1). This consistency of effect across the 4 haemolytic markers indicates that the first principal component is a good marker of haemolysis.

### Heritability analysis

In the 90 sibling pairs identified in the CSSCD there was a significant positive correlation ( $r = 0.24$ ,  $P$ -value = 0.02) of haemolytic score compared with 200 unrelated subjects, where there was no correlation, suggesting that the haemolytic score is heritable. Table II shows similar heritability for LDH, AST, bilirubin and reticulocyte count.

### Genome-wide association study and targeted genotyping

After quality control there were 569, 554 SNPs left for the GWAS. There was no significant correlation between the top

Table II. Heritability analysis for haemolytic markers and haemolytic score.

Clinical Variable	Sibling Pairs		Unrelated Pairs	
	Correlation	$P$ -value	Correlation	$P$ -value
AST	0.559	$3.94 \times 10^{-10}$	0.002	0.4982
LDH	0.307	0.0048	-0.0013	0.4978
Serum bilirubin	0.304	0.0036	-0.0054	0.4866
Reticulocyte count	0.495	$6.01 \times 10^{-08}$	0.0013	0.512
Haemolytic score	0.24	0.02271	0.0002	0.5051

AST, aspartate transaminase; LDH, lactate dehydrogenase.

10 principal components and the haemolytic score, suggesting no evidence for confounding by population stratification.

The Manhattan plot summarizing the results of the GWAS is shown in Fig 1 and the results of the most significantly associated SNPs summarized in Table III. The QQ plot in Fig 2 shows no inflation and a genomic lambda factor of 1.01 was calculated, indicating that there is no confounding due to population stratification. Although no SNP met the  $10^{-8}$  level of genome-wide significance, rs7203560 in *NPRL3* was associated with haemolytic score with a  $P$  value of  $6.04 \times 10^{-07}$ . Rs7203560 lies ~30 kb upstream from HS-33. The minor allele of this *NPRL3* SNP shows a protective effect; as the number of minor alleles increases the haemolytic score decreases, reflecting decreased haemolysis. These results were replicated in the Walk-PHaSST cohort ( $\beta = -0.52$ ,  $P$ -value = 0.0143).

Genome-wide data was not available in the London cohort who were largely of African/West African/Afrocaribbean origin. In this validation cohort, rs7203560 was associated with

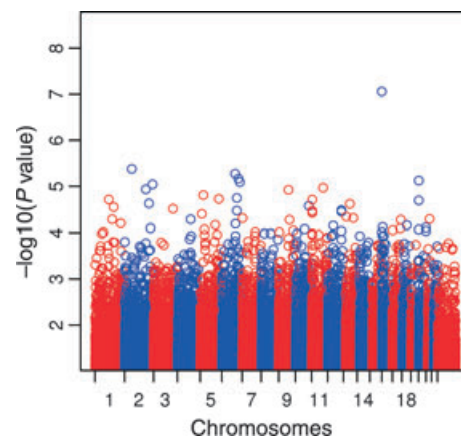


Fig 1. Manhattan Plot of Haemolytic Score. Manhattan plot summarizing the results of the genome-wide association study of haemolytic score with minor allele frequency >0.05.

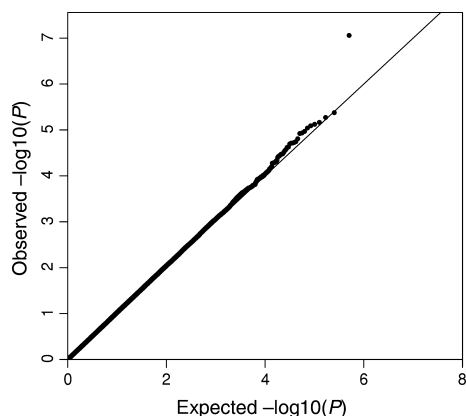
**Table III.** SNPs associated with haemolytic score.

Variant information							Meta- analysis				CSSCD		
SNP	Chr	Gene	bp	Coded allele	Non-coded allele	MAF	$\beta$	SE	<i>P</i> -value	$\beta$	SE	<i>P</i> -value	
rs7203560	16	<i>NPRL3</i>	184390	C	A	0.07	-0.44	0.07	$2.06 \times 10^{-09}$	-0.44	0.09	$6.04 \times 10^{-07}$	
rs7948471	11	<i>OR51I2,OR51I1</i>	5471746	A	G	0.21	-0.26	0.04	$3.03 \times 10^{-10}$	-0.21	0.05	$5.87 \times 10^{-05}$	
rs7938426	11	<i>OR51I2,OR51I1</i>	5471832	G	A	0.21	-0.25	0.04	$1.09 \times 10^{-08}$	-0.21	0.05	$6.08 \times 10^{-05}$	
rs2445284	11	<i>OR51L1</i>	5029703	G	A	0.05	-0.82	0.07	$1.34 \times 10^{-29}$	-0.42	0.10	$7.39 \times 10^{-05}$	

Variant information							Walk-PHaSST			PUSH			London		
SNP	Chr	Gene	bp	Coded allele	Non-coded allele	MAF	$\beta$	SE	<i>P</i> -value	$\beta$	SE	<i>P</i> -value	$\beta$	SE	<i>P</i> -value
rs7203560	16	<i>NPRL3</i>	184390	C	A	0.07	-0.53	0.22	0.014	-0.13	0.31	0.672	-0.42	0.20	0.0367
rs7948471	11	<i>OR51I2, OR51I1</i>	5471746	A	G	0.21	-0.35	0.11	0.002	-0.39	0.14	0.004	-0.46	0.15	0.017
rs7938426	11	<i>OR51I2, OR51I1</i>	5471832	G	A	0.21	-0.37	0.11	0.0012	-0.39	0.14	0.004	NA	NA	NA
rs2445284	11	<i>OR51L1</i>	5029703	G	A	0.05	-1.32	0.15	$1.60 \times 10^{-17}$	-1.75	0.17	$3.01 \times 10^{-21}$	0.41	0.28	0.146

Single nucleotide polymorphisms (SNPs) meeting significance threshold ( $5 \times 10^{-4}$ ) in the CSSCD study that replicate in the three independent cohorts. The table reports the SNP identifier from dbSNP, chromosome (Chr), physical coordinates [human genome (hg)19], the coded allele in PLINK (also minor allele) and the non-coded allele, the minor allele frequency (MAF), the gene clusters where the SNP is located, and regression coefficient ( $\beta$ ), standard error (SE) and *P*-value in each study. Additive models of association were used in all studies adjusting for age and gender. bp denotes position of the SNP according to hg19 coordinates.



**Fig 2.** QQ plot from CSSCD for single nucleotide polymorphisms with minor allele frequency >0.05.

lower haemolytic score (*P*-value = 0.03674). Rs7203560 was not as highly associated with the 4 haemolytic markers individually after adjusting for age and gender in the CSSCD cohort (Table IV). These results suggest that using a summary trait (haemolytic score) allows us to have more power to detect traits that are associated with all 4 of the haemolytic variables with consistent effects in comparison to performing 4 individual GWAS.

Four SNPs in olfactory receptor (OR) genes on chromosome chr11p; and *OR51L2* (rs1391617:  $\beta = -0.17$ ,

**Table IV.** Association of rs7203560 with four haemolytic markers.

Haemolytic Marker	$\beta$	Standard Error	<i>P</i> -value
Reticulocyte count	-0.15	0.03	$5.12 \times 10^{-06}$
Serum bilirubin	-0.09	0.04	0.0099
LDH	-0.07	0.02	0.0003
AST	-0.07	0.02	0.0026

The association of rs7203560 with each of the 4 haemolytic markers individually using an additive model and adjusting for age and gender using data from the CSSCD.

AST, aspartate transaminase; LDH, lactate dehydrogenase.

*P*-value = 0.0003), *OR51L1* (rs2445284:  $\beta = -0.42$ , *P*-value =  $7.39 \times 10^{-05}$ ), *OR51L1/OR51L2* rs7938426:  $\beta = -0.21$ , *P*-value =  $6.08 \times 10^{-05}$ , rs7948471:  $\beta = -0.21$ , *P*-value =  $5.87 \times 10^{-05}$ ) were associated with the haemolytic score although below the generally accepted level for genome-wide significance. These results were replicated in the Walk-PHaSST and PUSH cohorts and rs794847 in *OR51L1/OR51L2* was validated in the London cohort (Table III).

When a meta-analysis of genome-wide data was done, rs7203560 and OR receptor polymorphisms all met genome-wide significance levels (Table III).

Two SNPs located in *BCL11A* (rs766432:  $\beta = 0.22$ , *P*-value =  $9.27 \times 10^{-07}$ ; rs10195871:  $\beta = 0.18$ , *P*-value =

$9.87 \times 10^{-05}$ ) that were associated with HbF in sickle cell anaemia in a previous GWAS were also associated with haemolytic score (Solovieff *et al*, 2010).

Eight SNPs had regression coefficients in the same direction in the CSSCD cohort and the Walk-PHaSST and PUSH cohorts. The results for the annotated SNPs are shown in Table III.

*Linkage disequilibrium patterns*

We hypothesized that rs7203560 is a marker for genetic variants in or near the hypersensitive site regions that are known to regulate expression of the *HBA1/HBA2* genes (Fig 3).

Using Haploview, Barrett *et al* (2005), rs7203560 was in perfect linkage disequilibrium (LD) with rs9926112 ( $r^2 = 1$ ) and in strong LD with rs7197554 ( $r^2 = 0.75$ ) and rs13336641 ( $r^2 = 0.77$ ). SNPs rs7197554 and rs9926112 are located approximately 10 kb upstream from the HS-33 site and rs13336641 is located approximately 3–4 kb upstream from HS-40 and right at the edge of CTCF and other transcription factor binding sites (ENCODE data). While we were able to determine that rs7203560 was in high LD with SNPs near the HS regions, we were unable to determine if our top SNP was in LD with SNPs in the HS regions due to the low minor allele frequencies of these SNPs. Resequencing is required to determine if the haemolytic score is associated with novel variants in the HS region.

*Effects of  $\alpha$  thalassaemia and HbF*

As the number of minor alleles of rs7203560 increases from 0 to 1 to 2, the odds of at least one *HBA1/HBA2* deletion ( $-\alpha^{3.7}$ ) increases (0.33 to 2.18 to 3, respectively;  $\chi^2_2 = 97.32$ ,  $P$ -value  $<0.0001$ ). The odds of having at least one *HBA1/HBA2* deletion is 0.39 for individuals without a copy of the rs7203560 protective allele and is 2.23 for individuals with a copy of the protective allele ( $\chi^2_1 = 106.15$ ,  $P$ -value  $<0.0001$ ). To determine if the association seen between rs7203560 and the haemolytic score was a result of gene deletion  $\alpha$  thalassaemia, a 2-way analysis of variance was performed including the interaction between rs7203560 and  $\alpha$  thalassaemia. The interaction was not significant ( $\chi^2_1 = 0.23$ ,  $P$ -value = 0.4623) and the lack of interaction shows that  $\alpha$  thalassaemia does not modify the association between the haemolytic score and rs7203560.

There was a significant association between HbF and the haemolytic score after adjusting for  $\alpha$  thalassaemia status ( $\beta = -0.04$ ,  $P$ -value =  $4.00 \times 10^{-13}$ ). After adjusting for HbF and  $\alpha$  thalassaemia in a GWAS, rs7203560 remained significantly associated with the haemolytic score ( $\beta = -0.28$ ,  $P$ -value = 0.00375) indicating that the association between this SNP and the haemolytic score is not solely due to gene deletion  $\alpha$  thalassaemia and HbF. When we removed patients who had at least one *HBA1/HBA2* deletion the association between rs7203560 and the haemolytic score remained significant ( $P$ -value = 0.02463).

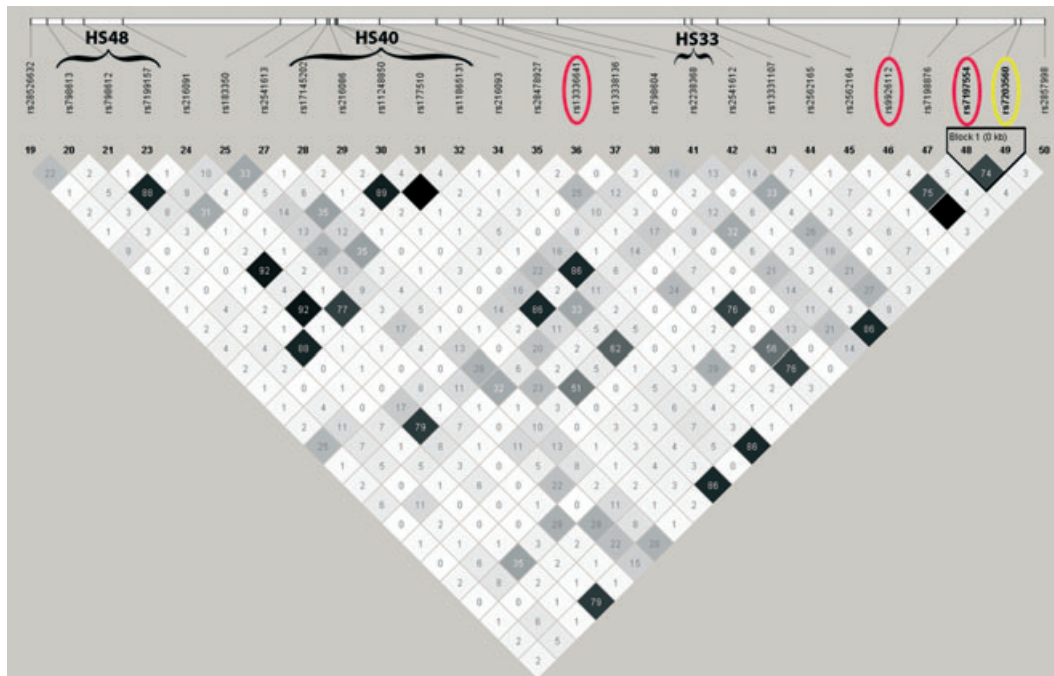


Fig 3. Location of the  $\alpha$ -globin regulatory elements (blue brackets) within introns of *NPRL3* and the linkage disequilibrium patterns of rs7203560 according to HapMap. Rs7203560 is circled in yellow. Circled in red are single nucleotide polymorphisms (SNPs) for which it was possible to calculate linkage disequilibrium. The allele frequencies for the SNPs are as follows: 0.139 for rs7203560, 0.175 for rs13336641, 0.175 for rs9926112 and 0.219 for rs7197554.

## Discussion

A composite variable derived from readily available markers of haemolysis was used as a subphenotype of sickle cell anaemia in a GWAS. This approach to estimating haemolysis was recently validated by showing that the haemolytic score was significantly associated with plasma red blood cell micro particles and cell-free haemoglobin concentration, both of which are more direct markers of intravascular haemolysis in sickle cell disease (Cannan, 1958; Reiter *et al*, 2002; Nouraei *et al*, 2012). As the baseline rate of haemolysis in an individual with sickle cell anaemia is stable over time, it is a suitable subphenotype in a genetic association study (Taylor *et al*, 2008).

Some complications of sickle cell anaemia more often appear in patients with high rates of haemolysis (Kato *et al*, 2007; Taylor *et al*, 2008). Among these are stroke, leg ulcers, sickle nephropathy, priapism and sickle vasculopathy, estimated by the tricuspid regurgitant jet velocity (TRV). All have been used as subphenotypes in genetic association studies (reviewed in Steinberg & Sebastiani, 2012).  $\alpha$  Thalassaemia reduces the risk of all these complications, while the degree of protection afforded by a high HbF level is less clear (Steinberg & Sebastiani, 2012). Insufficient power was present to detect an association of rs7203560 with individual disease subphenotypes associated with haemolysis because of the very small numbers of patients having these complications who were homozygous for the minor allele (data not shown).

Focused genotyping of candidate genetic modifiers of selected subphenotypes have shown that polymorphisms in *TGFBR3* and a few other genes have been associated with the haemolysis-associated subphenotypes. These results have not been replicated by GWAS. This is likely to be a result of the small effects of these polymorphisms on a subphenotype that then requires a very large number of cases to meet the stringent significance levels needed when hundreds of thousands of comparisons are being made. It has not yet been possible to assemble cohorts of the requisite size in a rare disorder like sickle cell anaemia.

Rs7203560, a SNP in the first intron of *NPRL3* (Nitrogen Permease Regulator-Like 3; C16orf35; chr 16p13.3) was most significantly associated with haemolytic score. *NPRL3* is highly conserved and upstream of the human *HBA1/HBA2* gene cluster in all vertebrates examined. (Hughes *et al*, 2005) Its functions are unknown, although, its deletion causes embryonic lethality from multiple cardiovascular defects (Kowalczyk *et al*, 2012). Within introns of *NPRL3* lie the *HBA1/HBA2* gene regulatory elements, HS-48, HS-40 and HS-33 that are required for *HBA1/HBA2* gene expression (Hughes *et al*, 2005). SNPs very close to the *HBA1/HBA2* regulatory elements are in strong LD with rs7203560. Also located between HS-33 and HS-40 adjacent to rs1333664 is a CTCF binding site. CTCF is a conserved zinc finger protein with many regulatory functions (Phillips & Corces, 2009).

We therefore hypothesize that rs7203560 is a marker for one or more variants in or near the *HBA1/HBA2* gene regulatory elements that down-regulate *HBA1/HBA2* gene expression causing a mild  $\alpha$  thalassaemia-like effect. The functional basis for this effect is unknown although it has recently become apparent that polymorphisms in DNase I hypersensitive sites and other regulatory regions play a critical role in defining gene expression (Maurano *et al*, 2012). The minor allele frequency (MAF) of rs7203560 is 0.139 in the HapMap Yoruban population; however, it has a MAF of 0 in the HapMap Central European population. This suggests that, like gene deletion  $\alpha$  thalassaemia, genetic elements marked by rs7203560 SNP could confer a selective advantage in populations with a high prevalence of malaria.

A meta-analysis of 24 167 individuals of European ancestry and 14 700 Japanese, found SNPs in *ITFG3* that were associated with mean corpuscular volume (MCV) and mean corpuscular haemoglobin (MCH) (Ganesh *et al*, 2009; Kamatani *et al*, 2010). *ITFG3* (integrin alpha FG-GAP repeat containing 3; C16orf9; 16p13.3) is downstream of *HBA1* and 41.2 kb from rs7203560. Down-regulating *HBA1/HBA2* gene expression should reduce MCV; however, because of the relatively small number of patients we studied we were unable to show a relationship between rs7203560 and MCV.

Six haplotypes of the HS-40 region have been described and haplotypes A and D are most prevalent in individuals of African origin (Harteveld *et al*, 2002). The common  $-\alpha^{3-7}$  deletion is found on several of these haplotypes. Whether or not HS-40 haplotypes affect *HBA1/HBA2* gene expression, as reflected by the  $\alpha/\beta$  globin biosynthesis ratios, or alter the properties of erythrocytes, estimated by MCV and MCH, have not been definitively determined. Expression of a luciferase reporter in K562 cells was greatest when the enhancer was the 350 bp core element of the A haplotype of HS-40 (Ribeiro *et al*, 2009). The relationship of haplotypes of HS-40—haplotypes of other *HBA1/HBA2* gene regulatory elements have yet to be described—to rs7203560 is unknown.

In conclusion, a SNP in *NPRL3* was independently associated with haemolysis in sickle cell anaemia and was in LD with regions of the genome that regulate *HBA1/HBA2* gene expression. Perhaps by independently down-regulating expression of the *HBA1/HBA2* genes, variants of the *HBA1/HBA2* gene regulatory loci, tagged by the *NPRL3* SNP, rs7203560, reduce haemolysis in sickle cell anaemia. As intravascular haemolysis plays an important role in the pathophysiology of sickle cell disease and other diseases where chronic haemolysis is present, like malaria, and in the lesion associated with blood storage, understanding further the genetic basis of red cell destruction has potential therapeutic implications (Rother *et al*, 2005; Gladwin *et al*, 2012).

## Acknowledgements

This work was supported by National Institutes of Health Grants R01 HL87681 (MHS), RC2 L101212 (MHS), 5T32

HL007501 (JNM), 2R25 HL003679-8 (VRG), R01 HL079912 (VRG), 2M01 RR10284-10 (VRG), R01HL098032 (MTG), R01HL096973 (MTG), P01HL103455 (MTG), the Institute for Transfusion Medicine and the Hemophilia Center of Western Pennsylvania (MTG), Medical Research Grant (UK) G0001249 ID 56477 (SLT). Monika Kowalczyk and Douglas Higgs kindly provided their unpublished results and very important input.

### Conflict-of interest disclosure

The authors declare no competing financial interests.

### Author contributions

JNM and PS analysed the data, CTB, EM, ED, HR performed laboratory analyses, VRG, GRK, CM, JT, AC, LL-J, SR, OC contributed patients to the studies, JNM, EM, YZ, MN, PS, SLT, MTG and MHS analysed and interpreted data and wrote and edited the manuscript.

### Appendix I

Walk-PHAASST Investigators: DB. Badesch<sup>1</sup>, RJ Barst<sup>2</sup>, OL Castro<sup>3</sup>, JSR Gibbs<sup>4</sup>, RE Girgis<sup>5</sup>, MT Gladwin<sup>6, 7</sup>, JC Goldsmith<sup>8</sup>, VR Gordeuk<sup>3</sup>, KL Hassell<sup>1</sup>, GJ Kato<sup>9</sup>, L Krishnamurti<sup>10</sup>, S Lanzkron<sup>5</sup>, JA. Little<sup>11</sup>, R F Machado<sup>12</sup>, CR. Morris<sup>13</sup>, M Nouraie<sup>3</sup>, O Onyekwere<sup>3</sup>, EB Rosenzweig<sup>2</sup>, V Sachdev<sup>14</sup>, DE Schraufnagel<sup>12</sup>, MA Waclawiw<sup>15</sup>, R Woolson<sup>16</sup>, NA Yovetich<sup>16</sup>.

<sup>1</sup>University of Colorado HSC, Denver, CO; <sup>2</sup>Columbia University, New York, NY; <sup>3</sup>Howard University, Washington, DC; <sup>4</sup>National Heart & Lung Institute, Imperial College London, and Hammersmith Hospital, London; <sup>5</sup>Johns Hopkins University, Baltimore, MD; <sup>6</sup>Division of Pulmonary, Allergy

and Critical Care Medicine and the <sup>7</sup>Vascular Medicine Institute, at the University of Pittsburgh, Pittsburgh, PA; <sup>8</sup>National Heart Lung and Blood Institute/NIH, Bethesda, MD (personal views that do not represent the Government); <sup>9</sup>Cardiovascular and Pulmonary Branch, NHLBI, Bethesda, MD; <sup>10</sup>Children's Hospital of Pittsburgh, Pittsburgh, PA; <sup>11</sup>Albert Einstein College of Medicine, Bronx, NY; <sup>12</sup>University of Illinois, Chicago, IL <sup>13</sup>Children's Hospital & Research Center Oakland, Oakland, CA; <sup>14</sup>Translational Medicine Branch, NHLBI, Bethesda, MD; <sup>15</sup>Office of Biostatistics Research, NHLBI, Bethesda, MD (personal views that do not represent the Government) <sup>16</sup>Rho, Inc., Chapel Hill, NC.

Walk-PHAASST Intramural NHLBI staff: Research nurses: C Seamon; A Chi; W Coles; Pulmonologists: S Alam; Haematologists: J Taylor; C Minniti; Protocol Management: MK. Hall;

Walk-PHAASST Children's Hospital & Research Center Oakland, Oakland, CA staff: L Lavrishia; W Hagar; H Rosenfeld; Echocardiography lab: C Brenneman; S Sidenko; C Birdsall; W Li, M St. Peter; C Brenneman.

National Heart & Lung Institute, Imperial College London and Hammersmith Hospital, London: Central Middlesex Hospital staff: Ki Anie, G Cho; S Davies; A Gilmore; Hammersmith Hospital staff: M Layton; I Cabrita; G Mahalingam; S Meehan; G Addis; St. Thomas Hospital staff: J Howard; C Woodley.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Fig S1.** Plot of the haemolysis score versus each of the four haemolytic variable.

### References

- Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H. & Steinberg, M.H. (2011) Fetal hemoglobin in sickle cell anemia. *Blood*, **118**, 19–27.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of molecular biology*, **215**, 403–410.
- Barrett, J.C., Fry, B., Maller, J.D.M.J. & Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
- Cannan, R.K. (1958) Proposal for a certified standard for use in hemoglobinometry; second and final report. *The Journal of laboratory and clinical medicine*, **52**, 471–476.
- De Ceulaer, K., Higgs, D.R., Weatherall, D.J., Hayes, R.J., Serjeant, B.E. & Serjeant, G.R. (1983) Alpha-Thalassemia reduces the hemolytic rate in homozygous sickle-cell disease [letter]. *New England Journal of Medicine*, **309**, 189–190.
- Dham, N., Ensing, G., Minniti, C., Campbell, A., Arteta, M., Rana, S., Darbari, D., Nouraie, M., Onyekwere, O., Lasota, M., Kato, G.J., Gladwin, M.T., Castro, O., Gordeuk, V. & Sable, C. (2009) Prospective echocardiography assessment of pulmonary hypertension and its potential etiologies in children with sickle cell disease. *American Journal of Cardiology*, **104**, 713–720.
- Dozy, A.M., Kan, Y.W., Embury, S.H., Mentzer, W.C., Wang, W.C., Lubin, B., Davis, J.R. Jr & Koenig, H.M. (1979)  $\alpha$ -Globin gene organisation in blacks precludes the severe form of  $\alpha$ -thalassaemia. *Nature*, **280**, 605–607.
- Ganesh, S.K., Zakai, N.A., van Rooij, F.J., Soranzo, N., Smith, A.V., Nalls, M.A., Chen, M.H., Kottgen, A., Glazer, N.L., DeGhghan, A., Kuhnel, B., Aspelund, T., Yang, Q., Tanaka, T., Jaffe, A., Bis, J.C., Verwoert, G.C., Teumer, A., Fox, C.S., Guralnik, J.M., Ehret, G.B., Rice, K., Felix, J.F., Rendon, A., Eiriksdottir, G., Levy, D., Patel, K.V., Boerwinkle, E., Rotter, J.I., Hofman, A., Sambrook, J.G., Hernandez, D.G., Zheng, G., Bandinelli, S., Singleton, A.B., Coresh, J., Lumley, T., Uitterlinden, A.G., Vangils, J.M., Launer, L.J., Cupples, L.A., Oostra, B.A., Zwavinga, J.J., Ouwendijk, W.H., Thein, S.L., Meisinger, C., Deloukas, P., Nauck, M., Spector, T.D., Gieger, C., Gudnason, V., van Duijn, C.M., Psaty, B.M., Ferrucci, L., Chakravarti, A., Greinacher, A., O'Donnell, C.J., Witteman, J.C., Furth, S., Cushman, M., Harris, T.B. & Lin, J.P. (2009) Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nature Genetics*, **41**, 1191–1198.
- Gaston, M. & Rosse, W.F. (1982) The cooperative study of sickle cell disease: review of study design and objectives. *The American Journal of Pediatric Hematology-Oncology*, **4**, 197–201.
- Gladwin, M.T., Crawford, J.H. & Patel, R.P. (2004) The biochemistry of nitric oxide, nitrite, and hemoglobin: role in blood flow regulation. *Free Radical Biology and Medicine*, **36**, 707–717.
- Gladwin, M.T., Kanias, T. & Kim-Shapiro, D.B. (2012) Hemolysis and cell-free hemoglobin drive an intrinsic mechanism for human disease. *The Journal of Clinical Investigation*, **122**, 1205–1208.
- Gordeuk, V.R., Campbell, A., Rana, S., Nouraie, M., Niu, X., Minniti, C.P., Sable, C., Darbari, D., Dham, N., Onyekwere, O., Ammosova, T.,



- Nekhai, S., Kato, G.J., Gladwin, M.T. & Castro, O.L. (2009) Relationship of erythropoietin, fetal hemoglobin, and hydroxyurea treatment to tricuspid regurgitation velocity in children with sickle cell disease. *Blood*, **114**, 4639–4644.
- Hartevelde, C.L., Muglia, M., Passarino, G., Kielman, M.F. & Bernini, L.F. (2002) Genetic polymorphism of the major regulatory element HS-40 upstream of the human alpha-globin gene cluster. *British Journal of Haematology*, **119**, 848–854.
- Hebbel, R.P. (2011) Reconstructing sickle cell disease: a data-based analysis of the “hyperhemolysis paradigm” for pulmonary hypertension from the perspective of evidence-based medicine. *American Journal of Hematology*, **86**, 123–154.
- Hughes, J.R., Cheng, J.F., Ventress, N., Prabhakar, S., Clark, K., Anguita, E., De Gobbi, M., de Jong, P., Rubin, E. & Higgs, D.R. (2005) Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 9830–9835.
- Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Nakamura, Y. & Kamatani, N. (2010) Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nature Genetics*, **42**, 210–215.
- Kato, G.J., Gladwin, M.T. & Steinberg, M.H. (2007) Deconstructing sickle cell disease: reappraisal of the role of hemolysis in the development of clinical subphenotypes. *Blood Reviews*, **21**, 37–47.
- Kowalczyk, M.S., Hughes, J.R., Babbs, C., Sanchez-Pulido, L., Szumska, D., Sharpe, J.A., Sloane-Stanley, J.A., Morriss-Kay, G.M., Smoot, L.B., Roberts, A.E., Watkins, H., Bhattacharya, S., Gibbons, R.J., Ponting, C.P., Wood, W.G. & Higgs, D.R. (2012) Nprl3 is required for normal development of the cardiovascular system. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, **7–8**, 404–415.
- Liu, Y.T., Old, J.M., Miles, K., Fisher, C.A., Weatherall, D.J. & Clegg, J.B. (2000) Rapid detection of  $\alpha$ -thalassaemia deletions and  $\alpha$ -globin gene triplication by multiplex polymerase chain reactions. *British Journal of Haematology*, **108**, 295–299.
- Machado, R.F., Barst, R.J., Yovetich, N.A., Hassell, K.L., Kato, G.J., Gordeuk, V.R., Gibbs, J.S., Little, J.A., Schraufnagel, D.E., Krishnamurti, L., Girgis, R.E., Morris, C.R., Rosenzweig, E.B., Badesch, D.B., Lanzkron, S., Onyekwere, O., Castro, O.L., Sachdev, V., Waclawiw, M.A., Woolson, R., Goldsmith, J.C. & Gladwin, M.T. (2011) Hospitalization for pain in patients with sickle cell disease treated with sildenafil for elevated TRV and low exercise capacity. *Blood*, **118**, 855–864.
- Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, F., Kutayvin, T., Stehling-Sun, S., Johnson, A.K., Canfield, T.K., Giste, E., Diegel, M., Bates, D., Hansen, R.S., Neph, S., Sabo, P.J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S.R., Kaul, R. & Stamatoyannopoulos, J.A. (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science*, **337**, 1190–1195.
- Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S. & Dubchak, I. (2000) VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*, **16**, 1046–1047.
- Milton, J.N., Sebastiani, P., Solovieff, N., Hartley, S.W., Bhatnagar, P., Arking, D.E., Dworkis, D.A., Casella, J.F., Barron-Casella, E., Bean, C.J., Hooper, W.C., DeBaun, M.R., Garrett, M.E., Soldano, K., Telen, M.J., Ashley-Koch, A., Gladwin, M.T., Baldwin, C.T., Steinberg, M.H. & Klings, E.S. (2012) A genome-wide association study of total bilirubin and cholelithiasis risk in sickle cell anemia. *PLoS ONE*, **7**, e34741.
- Minniti, C.P., Sable, C., Campbell, A., Rana, S., Ensing, G., Dham, N., Onyekwere, O., Nouraie, M., Kato, G.J., Gladwin, M.T., Castro, O.L. & Gordeuk, V.R. (2009) Elevated tricuspid regurgitant jet velocity in children and adolescents with sickle cell disease: association with hemolysis and hemoglobin oxygen desaturation. *Haematologica*, **94**, 340–347.
- Morris, C.R., Kato, G.J., Poljakovic, M., Wang, X., Blackwelder, W.C., Sachdev, V., Hazen, S.L., Vichinsky, E.P., Morris, S.M. & Gladwin, M.T. (2005) Dysregulated arginine metabolism, hemolysis-associated pulmonary hypertension and mortality in sickle cell disease. *Journal of the American Medical Association*, **294**, 81–90.
- Nouraie, M., Lee, J.S., Zhang, Y., Kanias, T., Zhao, X., Xiong, Z., Oriss, T.B., Zeng, Q., Kato, G.J., Gibbs, J.S., Hildesheim, M.E., Sachdev, V., Barst, R., Machado, R., Hassell, K.L., Little, J.A., Schraufnagel, D.E., Krishnamurti, L., Novelli, E., Girgis, R.E., Morris, C., Rosenzweig, E.B., Badesch, D.B., Lanzkron, S., Castro, O., Goldsmith, J.C., Gordeuk, V. & Gladwin, M.T. (2012) The relationship between the severity of hemolysis, clinical manifestations and risk of death in 415 patients with sickle cell anemia in the US and Europe. *Haematologica*, (Epub ahead of print).
- Phillips, J.E. & Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.
- Reiter, C.D., Wang, X., Tanus-Santos, J.E., Hogg, N., Cannon, R.O. III, Schechter, A.N. & Gladwin, M.T. (2002) Cell-free hemoglobin limits nitric oxide bioavailability in sickle-cell disease. *Nature Medicine*, **8**, 1383–1389.
- Ribeiro, D.M., Zaccariotto, T.R., Santos, M.N., Costa, F.F. & Sonati, M.F. (2009) Influence of the polymorphisms of the alpha-major regulatory element HS-40 on *in vitro* gene expression. *Brazilian Journal of Medical and Biological Research*, **42**, 783–786.
- Rother, R.P., Bell, L., Hillmen, P. & Gladwin, M.T. (2005) The clinical sequelae of intravascular hemolysis and extracellular plasma hemoglobin: a novel mechanism of human disease. *Journal of the American Medical Association*, **293**, 1653–1662.
- Solovieff, N., Milton, J.N., Hartley, S.W., Sherva, R., Sebastiani, P., Dworkis, D.A., Klings, E.S., Farrer, L.A., Garrett, M.E., Shley-Koch, A., Telen, M.J., Fucharoen, S., Ha, S.Y., Li, C.K., Chui, D.H., Baldwin, C.T. & Steinberg, M.H. (2010) Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. *Blood*, **115**, 1815–1822.
- Steinberg, M.H. & Embury, S.H. (1986) Alpha-thalassemia in blacks: genetic and clinical aspects and interactions with the sickle hemoglobin gene. [Review]. *Blood*, **68**, 985–990.
- Steinberg, M.H. & Sebastiani, P. (2012) Genetic modifiers of sickle cell disease. *American Journal of Hematology*, **87**, 824–826.
- Taylor, J.G., Nolan, V.G., Mendelsohn, L., Kato, G.J., Gladwin, M.T. & Steinberg, M.H. (2008) Chronic hyper-hemolysis in sickle cell anemia: association of vascular complications and mortality with infrequent vasoocclusive pain. *PLoS ONE*, **3**, e2095.