

# blood

2013 122: 434-442  
Prepublished online May 29, 2013;  
doi:10.1182/blood-2013-01-478776

## Gene-centric association study of acute chest syndrome and painful crisis in sickle cell disease patients

Geneviève Galarnau, Sean Coady, Melanie E. Garrett, Neal Jeffries, Mona Puggal, Dina Paltoo, Karen Soldano, Antonio Guasch, Allison E. Ashley-Koch, Marilyn J. Telen, Abdullah Kutlar, Guillaume Lettre and George J. Papanicolaou

---

Updated information and services can be found at:

<http://bloodjournal.hematologylibrary.org/content/122/3/434.full.html>

Articles on similar topics can be found in the following Blood collections

[Free Research Articles](#) (1801 articles)  
[Red Cells, Iron, and Erythropoiesis](#) (452 articles)  
[Sickle Cell Disease](#) (61 articles)

---

Information about reproducing this article in parts or in its entirety may be found online at:

[http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub\\_requests](http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests)

Information about ordering reprints may be found online at:

<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:

<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.  
[Copyright 2011 by The American Society of Hematology; all rights reserved.](#)



## RED CELLS, IRON, AND ERYTHROPOIESIS

### Gene-centric association study of acute chest syndrome and painful crisis in sickle cell disease patients

Geneviève Galarneau,<sup>1,2</sup> Sean Coady,<sup>3</sup> Melanie E. Garrett,<sup>4</sup> Neal Jeffries,<sup>3</sup> Mona Puggal,<sup>3</sup> Dina Paltoo,<sup>3</sup> Karen Soldano,<sup>4</sup> Antonio Guasch,<sup>5</sup> Allison E. Ashley-Koch,<sup>4</sup> Marilyn J. Telen,<sup>6</sup> Abdullah Kutlar,<sup>7</sup> Guillaume Lettre,<sup>1,2</sup> and George J. Papanicolaou<sup>3</sup>

<sup>1</sup>Montreal Heart Institute, Montreal, Quebec, Canada; <sup>2</sup>Université de Montréal, Montreal, Quebec, Canada; <sup>3</sup>National Heart, Lung, and Blood Institute, Division of Cardiovascular Sciences, Bethesda, MD; <sup>4</sup>Center for Human Genetics, Department of Medicine, Duke University Medical Center, Durham, NC; <sup>5</sup>Emory University, Atlanta, GA; <sup>6</sup>Division of Hematology, Department of Medicine, Duke University Medical Center, Durham, NC; and <sup>7</sup>Department of Medicine, Sickle Cell Center, Medical College of Georgia, Augusta, GA

#### Key Points

- Using genome-wide association study, we found the first replicated genetic association with acute chest syndrome in sickle cell disease patients.
- The locus identified includes *COMMD7*, a gene highly expressed in the lung that interacts with NFκB to control inflammatory responses.

Patients with sickle cell disease (SCD) present with a wide range of clinical complications. Understanding this clinical heterogeneity offers the prospects to tailor the right treatments to the right patients and also guide the development of novel therapies. Several environmental (eg, nutrition) and nonenvironmental (eg, fetal hemoglobin levels, α-thalassemia status) factors are known to modify SCD severity. To find new genetic modifiers of SCD severity, we performed a gene-centric association study in 1514 African American participants from the Cooperative Study of Sickle Cell Disease (CSSCD) for acute chest syndrome (ACS) and painful crisis. From the initial results, we selected 36 single nucleotide polymorphism (SNPs) and genotyped them for replication in 387 independent patients from the CSSCD, 318 SCD patients recruited at Georgia Health Sciences University, and 449 patients from the Duke SCD cohort. In the combined analysis, an association between ACS and rs6141803 reached array-wide significance ( $P = 4.1 \times 10^{-7}$ ). This SNP is located 8.2 kilobases upstream of *COMMD7*, a gene highly expressed in the lung that interacts with nuclear factor-κB signaling. Our results provide new leads to gaining a better understanding of clinical variability in SCD, a “simple” monogenic disease. (*Blood*. 2013;122(3):434-442)

#### Introduction

Sickle cell disease (SCD) is among the most common Mendelian diseases worldwide and is particularly prevalent in regions where malaria is endemic.<sup>1,2</sup> SCD is caused by mutations in the *β-globin* gene that encodes 1 of the subunits of the oxygen carrier hemoglobin and is characterized by a wide spectrum of disease-specific complications. In the deoxygenated state, sickle hemoglobin forms long polymers that alter erythrocyte shape and flexibility, thus increasing hemolysis and the adherence between sickled red blood cells and the endothelium.<sup>3,4</sup> Although hemolysis and cell adherence are the main causes of complications in SCD, little is known about the additional environmental and nonenvironmental factors that may modify disease severity and therefore explain the remarkable clinical heterogeneity observed in this otherwise simple monogenic disease.

Environmental variables such as nutrition and sufficient hydration are linked to clinical heterogeneity in SCD.<sup>4,5</sup> Two nonenvironmental factors, high fetal hemoglobin (HbF) levels and concomitant α-thalassemia, also correlate with reduced morbidity and mortality in SCD.<sup>6</sup> The identification of additional disease-severity modifiers may yield novel insights into SCD pathophysiology. A genetic association study may improve understanding of SCD clinical heterogeneity

because it attempts to correlate DNA sequence variants with SCD-specific complications or relevant clinical variables (eg, HbF). Over the last 2 decades, several genetic associations in SCD have been published, but the results are questionable because of small sample size and lack of replication (reviewed in ref. 7). There are, however, 2 notable exceptions: robust associations between (1) 3 loci (*BCL11A*, *HBSIL-MYB*, and *β-globin*) and HbF levels<sup>8,9</sup> and (2) the bilirubin levels-associated *UGT1A1* locus and gallstones.<sup>10,11</sup>

To find novel genetic modifiers of SCD, we performed a gene-centric association study in 1514 participants from the Cooperative Study of Sickle Cell Disease (CSSCD) for acute chest syndrome (ACS) and painful crisis. For genotyping, we used the ITMAT-Broad-CARE (IBC) array, which covers genetic variation at ~2100 genes important for heart, lung, and blood diseases.<sup>12</sup> Although the CSSCD is one of the largest existing SCD cohorts, our discovery power is modest to detect variants of small effect on phenotypic variation. For this reason, we genotyped for replication 36 variants that reached  $P < 1.0 \times 10^{-4}$  in the CSSCD discovery sample in the DNA of 387 independent SCD patients from the CSSCD. We also genotyped markers in 318 SCD patients recruited at Georgia Health Sciences

Submitted January 28, 2013; accepted May 18, 2013. Prepublished online as *Blood* First Edition paper, May 29, 2013; DOI 10.1182/blood-2013-01-478776.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

G.L. and G.J.P. codirected the study.

© 2013 by The American Society of Hematology

University and 449 patients from Duke University. Our analysis identified 1 single nucleotide polymorphism (SNP; rs6141803) near *COMMD7* that reached array-wide significance, defined as  $P < 2.0 \times 10^{-6}$  after accounting for the number of independent SNPs present on the IBC array.<sup>13</sup> Overall, our findings prioritize DNA sequence variants and genes for future genetic and functional follow-up experiments in order to better grasp patient-to-patient clinical variability in SCD.

## Methods

### Ethics statement

Informed consent was obtained for all participants in accordance with the Declaration of Helsinki. The Candidate-gene Association Resource (CARE) Study is approved by the ethics committees of the participating studies and of the Massachusetts Institute of Technology. This project was also reviewed and approved by the Montreal Heart Institute Ethics Committee, the Duke Institutional Review Board, and the different recruiting centers.

### Samples and genotyping

The CSSCD is described in detail elsewhere.<sup>14</sup> Briefly, the CSSCD was a multicenter prospective study of the natural history of SCD; participant enrollment into phase 1 of the CSSCD began in 1978. Participant entry ended in 1981 for all patients older than age 6 months; however, infants continued to be enrolled until 1988. Both mild and hospital-based SCD patients were recruited. A total of 4085 participants, mostly African Americans and ranging in age from newborns to adults, were enrolled in phase 1 from 23 centers across the United States. Data collection for phase 1 of the CSSCD ended in 1988 (see <https://biolinc.nih.gov/studies/csscd/> for more information on the study design).

In the CSSCD, painful crisis and ACS events were defined as previously described by the CSSCD investigators and as reported in the CSSCD phase 1 clinical database.<sup>15-17</sup> Briefly, a painful crisis episode was defined as an occurrence of pain lasting  $\geq 2$  hours in the extremities, back, abdomen, chest, or head that could not be explained by a mechanism other than SCD. Pain episodes within 14 days were treated as a single episode. An episode of ACS occurred when a participant developed a new infiltrate on chest radiograph and/or had a perfusion defect detected on a lung radioisotope scan. Painful crisis and ACS were analyzed as rates by dividing the number of events by the number of patient-years.

A total of 318 patients from the Adult Sickle Cell Clinic of Georgia Health Sciences University (GHSU) Sickle Cell Center were included in this study as a validation cohort; patient age ranged from 20 to 74 years and 169 women and 149 men were included in the cohort. In the GHSU cohort, ACS was defined as a new pulmonary infiltrate involving more than 1 lung segment with fever, chest pain, and hypoxia. This excludes patients who present with classic lobar pneumonia, although the distinction may not be 100%. The Duke SCD cohort included 449 adult patients (199 men and 250 women) and used the following questions to define ACS: Have you ever experienced acute chest syndrome or pneumonia requiring hospitalization? and to define painful crisis: In the past 12 months, have you had painful episodes requiring hospitalization?. Demographics for the 3 SCD cohorts used in this study are summarized in Table 1.

DNA genotyping on the Illumina IBC array was carried out at the Broad Institute as part of the National Heart, Lung and Blood Institute (NHLBI) CARE Project. The IBC array interrogates genotypes at  $\sim 50\,000$  SNPs and captures genetic variation at  $\sim 2100$  genes relevant for heart, lung, and blood diseases.<sup>12</sup> Data quality control and genotype imputation were performed as previously described.<sup>13</sup> Imputation was performed using MACH 1.0.16.<sup>18</sup> MACH requires phased reference haplotypes to perform imputation. For the African American CSSCD participants, a combined Northern European (CEU) + Western Africans (YRI) reference panel was created using HapMap phase 2 data.<sup>19</sup> This panel includes SNPs segregating in both CEU and YRI, as well as SNPs segregating in 1 panel and monomorphic and

nonmissing in the other. Imputation was performed in 2 steps: For the first step, 300 individuals were randomly extracted to generate recombination and error rate estimates. In the second step, these rates were used to impute all individuals across the entire reference panel. Imputation results were filtered at an  $r_{sq\_hat}$  threshold  $\geq 0.6$  and a minor allele frequency (MAF) threshold  $\geq 1\%$ . For imputed markers with strong statistical association, we directly genotyped an overlapping set of CSSCD DNA samples ( $N = 777$ ) and found high concordance with the imputation results (mean Pearson correlation coefficient = 0.87; range, 0.65-1.0). The final CSSCD discovery dataset included 1514 DNA samples with a genotyping success rate  $>99.8\%$  (47 092 genotyped SNPs and 190 551 imputed SNPs). Genotyping in the CSSCD and GHSU replication cohorts was performed using the mass spectrometry-based MassArray iPLEX platform from Sequenom, removing SNPs and DNA samples with a genotyping success rate  $<95\%$  and  $<90\%$ , respectively. The concordance rate, which was estimated from replicates, was  $>99.7\%$ . Genotyping and quality-control filters for the Duke SCD cohort were described elsewhere.<sup>20</sup> We used identity-by-descent methods to identify patients who overlap in the CSSCD and Duke cohorts; these patients ( $N = 23$ ) were excluded from the analysis of the Duke cohort.

### Statistical analysis

We performed our discovery experiment in the CSSCD,<sup>14,21</sup> a large longitudinal study with hundreds of clinical variables available. Despite this wealth of phenotypic data, we focus initially on only 2 quantitative measures, ACS and painful crisis rates, for two reasons. First, statistical power to find genetic associations with quantitative phenotypes was higher than for dichotomous traits (despite remaining relatively modest), even in the large CSSCD. For instance, we estimated that our study design had 50% power to find an association between a quantitative trait and a SNP under the following assumptions: minor allele frequency = 25%, variance explained = 1%, and  $\alpha = 1 \times 10^{-4}$  (Table 1). In comparison, we only had 8% power to find an association with stroke (prevalence = 7%) under the same assumptions for a variant with an odds ratio (OR) = 1.5. Second, because the size of our replication cohorts is small, for low prevalence complications, it is likely that there would be too few affected SCD patients to robustly validate genetic associations observed in the CSSCD discovery cohort.

In the CSSCD discovery cohort, we tested associations between 237 643 genotyped (0, 1, 2) or imputed (0.0-2.0) common SNPs (MAF  $\geq 1\%$ ) and phenotypes using Poisson regression (correction for overdispersion) for painful crisis and ACS rates.<sup>17</sup> We implemented the analysis using custom scripts in the R 2.10.0 statistical package ([www.r-project.org/](http://www.r-project.org/)). We used sex, age at baseline, and the first 10 principal components as covariates. Analyses were stratified based on  $\alpha$ -thalassemia status, and association results were combined by inverse variance meta-analyses.<sup>22</sup> Analyses of global and local ancestry were performed using, respectively, the EIGENSOFT and HAPMIX software with their default parameters.<sup>23-25</sup>

Analyses in the CSSCD replication cohort were performed as for the CSSCD discovery cohort, except that we used  $\alpha$ -thalassemia status as a covariate because of the small sample size of the cohort and we did not have access to principal components. In the GHSU SCD cohort, age at baseline and painful crisis information were not available. We analyzed the association between genotypes and ACS (dichotomous) using logistic regression in PLINK<sup>26</sup> and sex and year of birth as covariates. For the Duke SCD cohort, logistic regression was used to determine the effect of genotype on a binary definition of ACS using PLINK.<sup>26</sup> To examine the impact of number of hospitalizations for painful crisis episodes, logistic regression was employed using SAS version 9.2 (SAS Systems, Cary, NC). In an attempt to reduce any population substructure that may exist, principal component analysis was performed using EIGENSOFT.<sup>23</sup> All models were adjusted for sex, age, and the first 2 principal components.

To analyze the effect of heme on the expression of *COMMD7* in pulmonary endothelial cells, we accessed the relevant gene expression dataset on the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) website (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi> accession number = GSE25014).<sup>27</sup> We performed the analyses separately for the pulmonary microvascular endothelial cells and the pulmonary artery endothelial cells using the GEO2R analytical module and default parameters

**Table 1. Description of the SCD cohorts**

Phenotype	CSSCD discovery (N = 1514)	CSSCD replication (N = 387)	Georgia Health Sciences University (N = 318)	Duke (N = 449)	Statistical power (0.5% and 1% variance explained)
Males/females	740/764	177/210	149/169	199/250	—
Age, y	14.2 ± 11.9	11.2 ± 12.5	NA	33.7 ± 12.1	—
Follow-up, y	6.6 ± 1.6	6.1 ± 2.2	NA	NA	—
α-thalassemia (N)	27.1% (410)	25.3% (98)	NA	NA	—
Acute chest syndrome (events/patient-year or affected/nonaffected)	0.12 ± 0.24	0.16 ± 0.48	52/262 (4 missing)	337/112	13%, 50%
Painful crisis (events/patient-year or any/none)	0.80 ± 1.45	0.65 ± 1.17	NA	259/171	13%, 50%

Analyses in this study were restricted to sickle cell anemia (HbSS) or HbSβ0 patients, with or without α-thalassemia. Means ± standard deviations are provided. Statistical power for the CSSCD discovery cohort was calculated using the following assumptions: minor allele frequency = 25%, effect size = 0.5% or 1% of the phenotypic variance explained, and  $\alpha = 1 \times 10^{-4}$ .

NA; not available.

(<http://www.ncbi.nlm.nih.gov/geo/info/geo2r.html>). We corrected *P* values using the Benjamini–Hochberg false-discovery rate method.

## Results

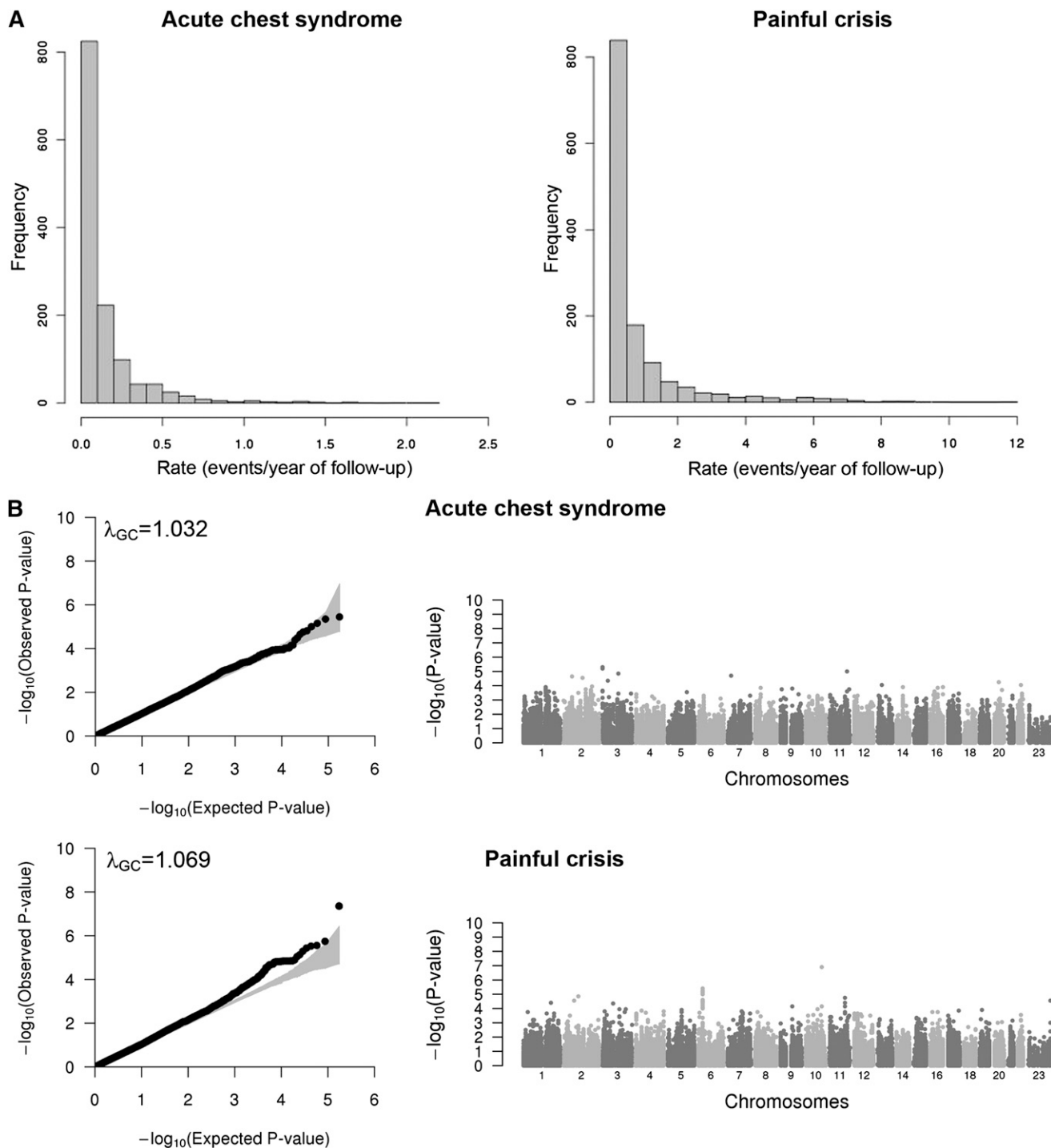
Several complications are observed in SCD and are linked to the quality of life and life expectancy of patients with this hemoglobinopathy.<sup>15,16,28–31</sup> Our goal in this study was to identify genetic variants associated with 2 of these complications, ACS and painful crisis, in order to better understand clinical heterogeneity in SCD. We modeled and analyzed ACS and painful crisis rates as previously described using Poisson regression.<sup>15–17</sup> Although for ACS the distribution of the observed test statistics does not show major departure from the expected null distribution ( $\lambda_{GC} = 1.032$ ), there was a slight inflation for the painful crisis association results ( $\lambda_{GC} = 1.069$ ; Figure 1).<sup>32</sup> For this reason, we corrected both ACS and painful crisis association results using the genomic control (GC) approach. However, overall the limited inflation that was observed indicates that our analysis was appropriate and accounted for the main possible confounders. To declare statistical significance on the IBC array, we selected a threshold of  $\alpha = 2 \times 10^{-6}$  that is sufficient to account for the number of independent tests performed.<sup>13</sup> Using this criterion, we identified a single locus that met array-wide significance: an association between rs11817401 in the SORCS1 gene and painful crisis rate ( $P = 1.2 \times 10^{-7}$ ; Figure 1 and Table 2).

To confirm this association and also identify additional loci that appeared promising but did not reach statistical significance in the CSSCD discovery cohort, we selected all SNPs with discovery  $P < 1 \times 10^{-4}$  (before GC correction; 19 SNPs for painful crisis and 17 SNPs for ACS) and genotyped them in 387 independent CSSCD participants. Replication and combined association results are presented in Table 2 and Table 3 for painful crisis and ACS, respectively. The association between SORCS1-rs11817401 and painful crisis did not replicate (replication  $P = .52$ , opposite direction of effect). Overall, we replicated in this small CSSCD cohort 2 associations at nominal level ( $P < .05$ ): an association between FAM193A-rs11732673 and painful crisis (replication  $P = .02$ , combined  $P = 9.9 \times 10^{-6}$ ) and an association between rs6141803 and ACS (replication  $P = .003$ , combined  $P = 5.2 \times 10^{-7}$ ), the latter reaching array-wide significance when combining the CSSCD discovery and replication results (Table 3). FAM193A encodes a protein with no clear biological functions. The rs6141803 SNP is located between the COMMD7 and DNMT3B genes on

chromosome 20. DNMT3B encodes a DNA methyltransferase, which is important for development, and COMMD7, a gene highly expressed in the lung, codes for an adaptor protein that interacts with subunits of the nuclear factor (NF)-κB complex.<sup>33</sup> Importantly, treating human pulmonary endothelial cells with free heme, a model that recapitulates some of the cellular responses observed when ACS is induced in a SCD mouse model,<sup>34</sup> significantly modulates the expression of COMMD7 (differential expression in pulmonary microvascular endothelial cells [ $P = 5 \times 10^{-4}$ ] and in pulmonary artery endothelial cells [ $P = 3 \times 10^{-5}$ ]).<sup>27</sup> This result, together with a role in NF-κB signaling and inflammation, adds additional evidence supporting a role for COMMD7 in ACS.

Our group had access to 2 additional SCD replication cohorts: 318 patients recruited at GHSU and 449 SCD patients from Duke University (Table 1). Painful crisis information was not available for the GHSU cohort and only available as categories for the Duke cohort. ACS information was available for both cohorts but only in the form of a binary presence/absence phenotype. In many situations, dichotomizing a quantitative trait can lead to substantial loss in statistical power as individuals with 1 or several ACS events are all labeled as affected.<sup>35</sup> For replication in the GSHU SCD cohort, we genotyped the top 17 SNPs associated with ACS (Table 3). A single variant, rs17728960 in the NFATC2 gene, was nominally significant ( $P = .05$ ), but the combined association result was not significant. The association between ACS and COMMD7-rs6141803 was not significant ( $P = .32$ ) but trended in the right direction (OR = 0.41; Table 3). Genome-wide genotype data were available for the Duke SCD cohort. After quality-control steps and genotype imputation (see Materials and Methods section), 6 painful crisis and 14 ACS SNPs were available for association testing. The association between ACS and rs6141803 near COMMD7 in the Duke cohort showed a consistent direction of effect (OR = 0.16,  $P = .08$ ; Table 3). When we combine at the *P* value level results from the CSSCD discovery, CSSCD replication, and GHSU and Duke cohorts using a Z-score method weighted based on sample size, the association between ACS and rs6141803 is array-wide significant (weighted  $P = 4.1 \times 10^{-7}$ ).

We noted that COMMD7-rs6141803 is an ancestry informative marker: the C allele has a frequency of 17% and 0% in the HapMap individuals of Northern European (CEU) and African (YRI) ancestry, respectively. This observation raises the possibility that the association between ACS and COMMD7-rs6141803 is a false-positive result owing to admixture. However, this is unlikely because the ACS rate is not correlated with the first principal component, which captures European vs African admixture (Spearman  $\rho = -0.0188$ ,  $P = .47$ ), and we used the first 10 principal components in our analysis to



**Figure 1. Association results for acute chest syndrome and painful crisis.** (A) Histograms showing the distribution of the acute chest syndrome (left) and painful crisis (right) rates in the CSSCD participants analyzed in this study (N = 1514). (B) Quantile–quantile (QQ; left column) and Manhattan (right column) plots of association results for acute chest syndrome (top) and painful crisis (bottom) analyzed in participants from the CSSCD. In the QQ plots, the gray area corresponds to the 90% CI. Results are corrected using the genomic control method.

account for global admixture. Although the association between ACS rate and genotypes at rs6141803 is not spurious because of admixture, we tried to use local ancestry to fine-map the causal variant. We inferred local European vs African ancestry at the locus and used this estimate as a covariate in our regression model.<sup>24</sup> The strength of the genetic association between ACS and rs6141803 was reduced when controlling for local ancestry but remained significant ( $P = 5.4 \times 10^{-5}$

and  $P = .001$  without and with local ancestry as covariate), suggesting that rs6141803 is unlikely to be the causal variant at the locus. rs6141803 is intergenic and not in linkage disequilibrium (LD) with any nonsynonymous DNA sequence variants identified by the 1000 Genomes Project.<sup>36</sup> Interestingly, however, it is in weak LD ( $r^2 \sim 0.2$ – $0.3$ ) with *cis*-eQTL SNPs associated with *COMMD7* transcript levels in human liver<sup>37</sup> and monocytes.<sup>38</sup>

**Table 2. Painful crisis association results**

SNP	Chromosome (position)	Effect allele/other allele	Effect allele frequency	CSSCD discovery		CSSCD replication		Combined CSSCD		Duke		Gene (annotation)
				$\beta$ (SE)	P value	$\beta$ (SE)	P value	$\beta$ (SE)	P value	OR (95% CI)	P value	
rs12720497	1 (185097043)	T/C	0.10	0.529 (0.125)	$3.9 \times 10^{-5}$	0.336 (0.194)	.08	0.470 (0.107)	$1.2 \times 10^{-5}$	NA	NA	PLA2G4A (intron)
rs440006	2 (70636930)	T/C	0.98	-0.723 (0.167)	$2.9 \times 10^{-5}$	-0.310 (0.338)	.36	-0.637 (0.154)	$3.5 \times 10^{-5}$	0.94 [0.08, 10.65]	.96	Intergenic
rs3917296	2 (102151265)	A/G	0.98	-0.773 (0.172)	$1.5 \times 10^{-5}$	1.393 (1.013)	.17	-0.708 (0.176)	$5.5 \times 10^{-5}$	$8.5 \times 10^{-8}$ [-inf, 0]	1.00	IL1R1 (intron)
rs324035	3 (115351544)	A/C	0.74	-0.303 (0.076)	$1.2 \times 10^{-4}$	0.061 (0.170)	.72	-0.239 (0.071)	$8.3 \times 10^{-4}$	NA	NA	DRD3 (intron)
rs10513478	3 (156723359)	T/C	0.02	0.836 (0.209)	$1.1 \times 10^{-4}$	-1.779 (2.450)	.47	0.816 (0.216)	$1.5 \times 10^{-4}$	0.59 [0.13, 2.72]	.50	PLCH1 (intron)
<b>rs11732673</b>	<b>4 (2652853)</b>	<b>A/G</b>	<b>0.77</b>	<b>-0.338 (0.089)</b>	<b><math>2.2 \times 10^{-4}</math></b>	<b>-0.374 (0.154)</b>	<b>.02</b>	<b>-0.347 (0.079)</b>	<b><math>9.9 \times 10^{-6}</math></b>	NA	NA	<b>FAM193A (intron)</b>
rs6858735	4 (37592622)	T/C	0.12	0.375 (0.095)	$1.4 \times 10^{-4}$	-0.165 (0.218)	.45	0.283 (0.090)	$1.6 \times 10^{-3}$	NA	NA	TBC1D1 (intron)
rs13113915	4 (187702958)	A/G	0.40	0.261 (0.067)	$1.6 \times 10^{-4}$	0.245 (0.131)	.06	0.258 (0.061)	$2.5 \times 10^{-5}$	NA	NA	MTNFR1A (intron)
rs10942625	5 (90446448)	A/G	0.25	-0.362 (0.091)	$1.2 \times 10^{-4}$	-0.329 (0.178)	.07	-0.355 (0.083)	$2.0 \times 10^{-5}$	0.93 [0.68, 1.28]	.66	GPR98 (intron)
rs1851426	7 (93220872)	A/G	0.68	-0.290 (0.074)	$1.6 \times 10^{-4}$	0.269 (0.145)	.07	-0.168 (0.068)	0.01	NA	NA	CYP3A4 (intergenic)
rs10107231	8 (18098691)	T/C	0.83	-0.375 (0.096)	$1.7 \times 10^{-4}$	-0.105 (0.203)	.61	-0.322 (0.089)	$3.1 \times 10^{-4}$	NA	NA	NAAT1 (intron)
rs7034457	9 (77807918)	A/G	0.90	-0.418 (0.102)	$7.1 \times 10^{-5}$	0.045 (0.249)	.86	-0.348 (0.097)	$3.3 \times 10^{-4}$	NA	NA	POSK5 (intron)
rs7899453	10 (79469965)	A/C	0.22	0.317 (0.079)	$1.1 \times 10^{-4}$	0.179 (0.149)	.23	0.285 (0.072)	$7.3 \times 10^{-5}$	NA	NA	RFS24 (missense)
rs11817401	10 (108823893)	T/C	0.08	0.602 (0.110)	$1.2 \times 10^{-7}$	-0.176 (0.273)	.52	0.487 (0.105)	$3.4 \times 10^{-6}$	NA	NA	SORCS1 (intron)
rs17101814	11 (103867886)	T/C	0.09	0.444 (0.100)	$1.8 \times 10^{-5}$	0.128 (0.196)	.51	0.375 (0.091)	$4.1 \times 10^{-5}$	0.82 [0.50, 1.35]	.44	PDGFD (intron)
rs9933611	16 (52331386)	A/G	0.93	-0.449 (0.116)	$1.8 \times 10^{-4}$	-0.245 (0.272)	.37	-0.415 (0.110)	$1.5 \times 10^{-4}$	NA	NA	FTO (intron)
rs445683	17 (71913117)	T/C	0.67	0.311 (0.079)	$1.5 \times 10^{-4}$	0.089 (0.142)	.53	0.256 (0.071)	$3.2 \times 10^{-4}$	0.91 [0.68, 1.23]	.56	UBE2O (intron)
rs7507634	19 (10297562)	T/C	0.02	0.899 (0.226)	$1.2 \times 10^{-4}$	-0.821 (0.674)	.22	0.714 (0.221)	$1.2 \times 10^{-3}$	NA	NA	RAVER1 (intron)
rs2872817	23 (153211612)	A/G	0.70	-0.249 (0.058)	$3.0 \times 10^{-5}$	-0.158 (0.121)	.19	-0.231 (0.053)	$1.5 \times 10^{-5}$	NA	NA	TKTL1 (3'UTR)

We genotyped 19 SNPs with  $P < 1 \times 10^{-4}$  in the CSSCD Discovery dataset (before GC correction) in an independent subset of the CSSCD. We combined results from the discovery and replication cohorts using an inverse variance meta-analysis approach. Only 1 SNP (rs11732673, in bold) replicated at nominal significance, but no SNPs reached array-wide significance ( $P < 2 \times 10^{-6}$ ) in the combined analysis. We also report results for the analysis of painful crisis as a dichotomous trait in the Duke cohort. P values for the CSSCD Discovery cohort are G-corrected. Genomic positions are on NCBI build 37.1.

NA, not available; SE, standard error.

**Table 3. Acute chest syndrome association results**

SNP	Chromosome (position)	Effect allele/other allele	Effect allele frequency	CSSCD discovery		CSSCD replication		Combined CSSCD		Georgia Health Sciences University		Duke		Gene (annotation)
				$\beta$ (SE)	P value	$\beta$ (SE)	P value	$\beta$ (SE)	P value	OR (95% CI)	P value	OR [95% CI]	P value	
rs10399947	1 (149128584)	A/G	0.57	-0.319 (0.082)	$1.2 \times 10^{-4}$	0.339 (0.156)	.03	-0.174 (0.073)	.02	0.98 [0.64, 1.50]	.93	0.89 [0.66, 1.21]	.47	Intergenic
rs34661029	2 (61002492)	C/G	0.98	-0.822 (0.191)	$2.3 \times 10^{-5}$	-0.103 (0.450)	.82	-0.709 (0.178)	$7.1 \times 10^{-5}$	0.87 [0.24, 3.19]	.83	0.75 [0.16, 3.61]	.72	REL (missense)
rs17749316	2 (191548954)	C/G	0.99	-0.891 (0.228)	$1.2 \times 10^{-4}$	0.518 (1.013)	.61	-0.821 (0.226)	$2.7 \times 10^{-4}$	$3.2 \times 10^6$ [0, inf]	1.00	$5.5 \times 10^6$ [0, inf]	1.00	STAT1 (intron)
rs13021001	2 (128125889)	A/G	0.96	-0.631 (0.149)	$3.0 \times 10^{-5}$	0.450 (0.482)	.35	-0.534 (0.144)	$2.1 \times 10^{-4}$	0.55 [0.14, 2.19]	.39	0.84 [0.3, 2.34]	.73	GPR17 (3'UTR)
rs6778854	3 (563448)	A/G	0.61	0.374 (0.081)	$5.0 \times 10^{-6}$	0.057 (0.161)	.72	0.309 (0.073)	$2.4 \times 10^{-5}$	0.98 [0.64, 1.50]	.92	NA	NA	Intergenic
rs13315133	3 (38624908)	A/C	0.79	-0.339 (0.082)	$4.3 \times 10^{-5}$	0.031 (0.197)	.88	-0.284 (0.076)	$2.0 \times 10^{-4}$	0.90 [0.53, 1.54]	.71	NA	NA	SCN5A (intron)
rs3910551	3 (106659325)	C/G	0.02	0.793 (0.180)	$1.4 \times 10^{-5}$	-0.333 (0.532)	.53	0.675 (0.172)	$9.1 \times 10^{-5}$	1.13 [0.38, 3.32]	.83	0.25 [0.08, 0.81]	.02	ALCAM (intron)
rs5030094	3 (187943852)	T/C	0.99	-0.801 (0.221)	$3.5 \times 10^{-4}$	-0.727 (0.424)	.09	-0.785 (0.198)	$7.5 \times 10^{-5}$	0.54 [0.14, 2.15]	.38	0.73 [0.11, 4.71]	.74	KNG1 (intron)
rs10478813	5 (127859164)	C/G	0.16	0.400 (0.109)	$2.8 \times 10^{-4}$	-0.002 (0.240)	.99	0.330 (0.100)	$9.9 \times 10^{-4}$	0.81 [0.38, 1.75]	.60	NA	NA	FBN2 (intron)
rs5576	7 (24297876)	A/G	0.97	-0.655 (0.152)	$2.1 \times 10^{-5}$	0.078 (0.477)	.87	-0.586 (0.146)	$6.3 \times 10^{-5}$	1.66 [0.37, 7.55]	.51	1.05 [0.37, 3.03]	.92	NPY (3'UTR)
rs1176758	11 (113275565)	A/G	0.92	-0.828 (0.117)	$9.7 \times 10^{-6}$	0.110 (0.275)	.69	-0.427 (0.109)	$9.7 \times 10^{-5}$	1.61 [0.49, 5.33]	.43	0.68 [0.33, 1.38]	.29	Intergenic
rs6309	13 (46368601)	A/G	0.04	0.631 (0.159)	$9.0 \times 10^{-5}$	-0.391 (0.477)	.41	0.526 (0.153)	$5.6 \times 10^{-4}$	1.43 [0.45, 4.56]	.55	0.68 [0.34, 1.34]	.26	HTR2A (intron)
rs9927848	16 (23740572)	A/C	0.44	-0.354 (0.099)	$4.3 \times 10^{-4}$	-0.830 (0.450)	.07	-0.141 (0.083)	.09	1.14 [0.74, 1.75]	.55	1.03 [0.75, 1.40]	.87	Intergenic
rs12447481	16 (52306470)	T/G	0.98	-0.794 (0.206)	$1.5 \times 10^{-4}$	-0.830 (0.450)	.07	-0.800 (0.190)	$2.5 \times 10^{-5}$	$3.3 \times 10^6$ [0, inf]	1	1.70 [0.31, 9.43]	.58	FTO (intron)
<b>rs6141803</b>	<b>20 (30804017)</b>	<b>T/C</b>	<b>0.98</b>	<b>-0.701 (0.171)</b>	<b><math>5.4 \times 10^{-5}</math></b>	<b>-0.797 (0.266)</b>	<b>.003</b>	<b>-0.730 (0.145)</b>	<b><math>5.2 \times 10^{-7}</math></b>	<b>0.41 [0.07, 2.34]</b>	<b>.32</b>	<b>0.16 [0.02, 1.23]</b>	<b>.08</b>	<b>Intergenic</b>
rs17728960	20 (49563123)	T/C	0.98	-0.789 (0.208)	$1.9 \times 10^{-4}$	-0.053 (0.354)	.88	-0.595 (0.182)	$1.1 \times 10^{-3}$	0.35 [0.12, 0.99]	.05	0.28 [0.04, 2.19]	.22	NFATC2 (intron)
rs16998437	22 (36418341)	T/G	0.92	-0.483 (0.1216)	$9.3 \times 10^{-5}$	-0.061 (0.233)	.79	0.390 (0.109)	$3.5 \times 10^{-4}$	NA	NA	1.07 [0.52, 2.18]	.86	NCL12 (3'UTR)

We genotyped 17 SNPs with  $P < 1 \times 10^{-4}$  in the CSSCD Discovery dataset (before GC correction) in an independent subset of the CSSCD. We combined results from the discovery and replication CSSCD cohorts using an inverse variance meta-analysis approach. Only 1 SNP (rs6141803, in bold) replicated at nominal significance and reached array-wide significance ( $P < 2 \times 10^{-6}$ ) in the combined analysis. We also report results for the analysis of acute chest syndrome as a dichotomous trait in the GHSU and Duke cohorts. For rs6141803, the association was not significant but trended in the right direction. P values for the CSSCD Discovery cohort are GC corrected. Genomic positions are on NCBI build 37.1.

NA, not available; SE, standard error.

## Discussion

Painful crisis and ACS are, respectively, the first and second most frequent causes of hospital admissions in patients with SCD.<sup>39</sup> Although they share common causes (eg, vaso-occlusion), it is also clear that some of the triggering factors are different (eg, the role of infections in ACS). We performed one of the largest genetic association experiments to date in order to identify DNA sequence variants that modify SCD clinical severity through these 2 measures of morbidity.

Our experimental design identified a single SNP, rs6141803, that reached array-wide significance ( $P = 5.2 \times 10^{-7}$  in the CSSCD discovery + replication,  $P = 4.1 \times 10^{-7}$  if we add the GHSU and Duke SCD samples). Analyzing ACS as a dichotomous phenotype in the GHSU and Duke cohorts, as we did in this study, could account for a loss of statistical power. One additional difference between the CSSCD and the other 2 cohorts is the age of the enrolled patients. The GHSU and Duke cohorts are essentially adult cohorts, whereas the CSSCD includes a large number of children (Table 1). Age is a known predictor of ACS events<sup>16</sup> and, for this reason, was included as a covariate in our statistical model. We performed a sensitivity analysis in the CSSCD discovery cohort that clearly shows that the association signal between ACS and rs6141803 is observed most often in children: the association was strong in patients recruited before they were aged 5 years ( $N = 335$ ,  $P = 4 \times 10^{-6}$ ) but not significant in older patients ( $N = 978$ ,  $P = .80$ ). Thus, this age effect might explain why we could replicate this association in the CSSCD replication cohort (which also includes children) but not in the adult GHSU and Duke cohorts. It is also possible that this association is a false-positive report. Despite allele frequency differences between the ancestral populations at rs6141803, we performed analyses that suggest that admixture is unlikely to confound this result. Additional replication attempts in large SCD cohorts with quantitative measures of ACS are needed before drawing a final conclusion on the ACS-rs6141803 genetic association.

rs6141803 is an intergenic SNP located between DNMT3B and COMMD7. DNMT3B encodes for a DNA methyltransferase that is involved in maintenance DNA methylation. Mutations in DNMT3B cause immunodeficiency-centromeric instability-facial anomalies syndrome-1 (ICF-1; MIM #242860), a very rare syndrome that has not been linked to lung-related complications. COMMD7 encodes a poorly known protein that contains a copper metabolism gene MURR1 (COMM) domain. The gene is abundantly expressed in the lung<sup>33</sup> and is overexpressed in hepatocellular carcinoma.<sup>40</sup> The knockdown of COMMD7 using short-hairpin RNA increases apoptosis and cell cycle arrest, in part, by interfering with NF- $\kappa$ B signaling.<sup>33,40</sup> NF- $\kappa$ B is a master regulator of acute inflammation; upon stimulation, it transcriptionally activates interleukins, interferon, tumor necrosis factor- $\alpha$ , and adhesion molecules. The expression of COMMD7 in pulmonary endothelial cells is also affected upon heme treatment<sup>27</sup>; this is promising as free heme can induce ACS in a SCD mouse model.<sup>34</sup> Although more work is needed to clarify the possible role of COMMD7 in ACS, these are potentially interesting observations given the importance of inflammation and free radical production in aggravating ACS episodes.<sup>41</sup>

A recent large candidate gene study in 942 SCD children identified an association between a microsatellite in the heme oxygenase-1 (HMOX1) promoter and ACS: longer alleles were associated with increased rate of hospitalization for ACS.<sup>42</sup> Results from this initial study were not replicated in an independent cohort. Therefore, we queried our own results to test if *HMOX1* SNPs were associated with

ACS rate in the CSSCD. Although the HMOX1 gene was targeted for genotyping on the IBC array (28 genotyped or imputed nearby SNPs), the microsatellite was not directly tested. Of the HMOX1 SNPs that are accessible on the IBC array, 1 SNP in the 3' untranslated region of HMOX1 is associated with ACS at nominal significance (rs12160039;  $P = .02$ ). We would need to directly genotype the promoter microsatellite to determine if this SNP captures the association signal with ACS through LD. Similarly, although an intronic sequence repeat polymorphism in the NOS1 gene has been proposed to influence ACS risk,<sup>43,44</sup> none of the NOS1 SNPs tested in our study showed significant association results with ACS.

The second most interesting association identified in our experiment is between painful crisis rate and rs12720497, an intronic SNP in the PLA2G4A gene. The association is not array-wide significant, but the directions of effect are consistent between the discovery and replication CSSCD panels (replication  $P = .08$ , combined  $P = 1.2 \times 10^{-5}$ ; Table 2). PLA2G4A encodes a cytosolic phospholipase A2, an enzyme implicated in the production of proinflammatory molecules (prostaglandins and leukotrienes) that has previously been implicated in increased sensitivity to pain (hyperalgesia) in humans.<sup>45,46</sup> Enzymes in the phospholipase A2 family can be divided into 4 groups (cytosolic, secreted, calcium-independent, and lipoprotein-associated), and high levels of secreted phospholipase A2 have been suggested to be predictive of future ACS events.<sup>47-49</sup> The link between cytosolic and secreted phospholipase A2 and their role in SCD complications is intriguing, especially because severe ACS often occurs in the course of vaso-occlusive painful crisis. In our data, however, rs12720497 in PLA2G4A (coding for cytosolic phospholipase A2) is not associated with ACS rate ( $P = .67$ ).

We performed our discovery search in 1514 participants from the CSSCD whose DNA was genotyped on gene-centric genotyping arrays.<sup>12</sup> With the caveat that this genotyping platform only captures genetic variation at a subset (~10%) of the predicted human genes, we did not identify loci with moderate to strong effect on phenotype, which is consistent with most reported genome-wide association study results.<sup>50</sup> Our results highlight promising variants for further replication in independent SCD cohorts and biologically plausible candidate genes (eg, COMMD7, PLA2G4A) to test functionally, for instance, in SCD mouse models. They are also indicative of the importance of combining genome-wide association study results through meta-analyses between SCD cohorts to gain sufficient statistical power to identify genetic associations of weak phenotypic effect.

## Acknowledgments

The authors acknowledge the contribution of MéliSSa Beaudoin for DNA genotyping and Cameron D. Palmer for genotype imputation and thank all the patients who contributed to this study.

CSSCD is supported in part by the National Institutes of Health, National Heart, Lung, and Blood Institute (N01-HB-47110). CARE is supported by the National Heart, Lung, and Blood Institute (HHSN268200625226C). A full listing of the grants and contracts that have supported CARE is provided at <http://www.nhlbi.nih.gov/resources/geneticsgenomics/programs/care.htm>. The work in the Lettre Laboratory is supported by a Innovation in Clinical Research Award grant from the Doris Duke Charitable Foundation (2009089), the Canada Research Chair Program, the Canadian Institute of Health Research (123382), and the Fonds de Recherche Santé Québec. The



work at Duke University was funded in part by the National Heart, Lung, and Blood Institute (ROI HL079915 and RC2-HL101212).

A.E.A.-K., M.J.T., A.K., G.L., and G.J.P. analyzed the results; and G.G. and G.L. wrote the manuscript with contributions from all authors.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Guillaume Lettre, Montreal Heart Institute, 5000 Belanger St, Montreal, Quebec, Canada, H1T 1C8; e-mail: guillaume.lettre@umontreal.ca; and George J. Papanicolaou, National Heart, Lung, and Blood Institute, 6701 Rockledge Dr, Bethesda, MD 20892; e-mail: gjp@mail.nih.gov.

## Authorship

Contribution: G.G., S.C., G.L., and G.J.P. conceived and designed the experiment; G.G., S.C., M.E.G., N.J., K.S., and G.L. performed experiments; G.G., S.C., M.E.G., N.J., M.P., D.P., K.S., A.G.,

## References

- Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. *Blood*. 2010;115(22):4331-4336.
- Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: an increasing global health problem. *Bull World Health Organ*. 2001;79(8):704-712.
- Bunn HF. Pathogenesis and treatment of sickle cell disease. *N Engl J Med*. 1997;337(11):762-769.
- Steinberg MH, Forget BG, Higgs DR, Weatherall D. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. 2nd ed. Cambridge, United Kingdom: Cambridge University Press; 2009.
- Serjeant GR. Natural history and determinants of clinical severity of sickle cell disease. *Curr Opin Hematol*. 1995;2(2):103-108.
- Sankaran VG, Lettre G, Orkin SH, Hirschhorn JN. Modifier genes in Mendelian disorders: the example of hemoglobin disorders. *Ann N Y Acad Sci*. 2010;1214:47-56.
- Lette G. The search for genetic modifiers of disease severity in the beta-hemoglobinopathies. *Cold Spring Harb Perspect Med*. 2012;2(10).
- Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet*. 2007;39(10):1197-1199.
- Uda M, Galanello R, Sanna S, et al. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proc Natl Acad Sci USA*. 2008;105(5):1620-1625.
- Passon RG, Howard TA, Zimmerman SA, Schultz WH, Ware RE. Influence of bilirubin uridine diphosphate-glucuronosyltransferase 1A promoter polymorphisms on serum bilirubin levels and cholelithiasis in children with sickle cell anemia. *J Pediatr Hematol Oncol*. 2001;23(7):448-451.
- Milton JN, Sebastiani P, Solovieff N, et al. A genome-wide association study of total bilirubin and cholelithiasis risk in sickle cell anemia. *PLoS ONE*. 2012;7(4):e34741.
- Keating BJ, Tischfield S, Murray SS, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS ONE*. 2008;3(10):e3583.
- Lo KS, Wilson JG, Lange LA, et al. Genetic association analysis highlights new loci that modulate hematological trait variation in Caucasians and African Americans. *Hum Genet*. 2011;129(3):307-317.
- Farber MD, Koshy M, Kinney TR. Cooperative Study of Sickle Cell Disease: Demographic and socioeconomic characteristics of patients and families with sickle cell disease. *J Chronic Dis*. 1985;38(6):495-505.
- Platt OS, Thorington BD, Brambilla DJ, Milner PF, Rosse WF, Vichinsky E, Kinney TR. Pain in sickle cell disease. Rates and risk factors. *N Engl J Med*. 1991;325(1):11-16.
- Castro O, Brambilla DJ, Thorington B, et al; The Cooperative Study of Sickle Cell Disease. The acute chest syndrome in sickle cell disease: incidence and risk factors. *Blood*. 1994;84(2):643-649.
- Lette G, Sankaran VG, Bezerra MA, et al. 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 USA*. 2008;105(33):11869-11874.
- Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet*. 2009;10:387-406.
- Kang HM, Sul JH, Service SK, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet*. 2010;42(4):348-354.
- Sebastiani P, Solovieff N, Hartley SW, et al. Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study. *Am J Hematol*. 2010;85(1):29-35.
- Gaston M, Smith J, Gallagher D, et al. Recruitment in the Cooperative Study of Sickle Cell Disease (CSSCD). *Control Clin Trials*. 1987;8(4 Suppl):131S-140S.
- Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909.
- Price AL, Tandon A, Patterson N, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet*. 2009;5(6):e1000519.
- Lette G, Palmer CD, Young T, et al. Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARE Project. *PLoS Genet*. 2011;7(2):e1001300.
- Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
- Ghosh S, Tan F, Yu T, Li Y, Adisa O, Mosunjac M, Ofori-Acquah SF. Global gene expression profiling of endothelium exposed to heme reveals an organ-specific induction of cytoprotective enzymes in sickle cell disease. *PLoS ONE*. 2011;6(3):e18399.
- Nolan VG, Wyszynski DF, Farrer LA, Steinberg MH. Hemolysis-associated priapism in sickle cell disease. *Blood*. 2005;106(9):3264-3267.
- Milner PF, Kraus AP, Sebes JI, et al. Sickle cell disease as a cause of osteonecrosis of the femoral head. *N Engl J Med*. 1991;325(21):1476-1481.
- Koshy M, Entsuaeh R, Koranda A, et al. Leg ulcers in patients with sickle cell disease. *Blood*. 1989;74(4):1403-1408.
- Ohene-Frempong K, Weiner SJ, Sleeper LA, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood*. 1998;91(1):288-294.
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004.
- Burstein E, Hoberg JE, Wilkinson AS, et al. COMMD proteins, a novel family of structural and functional homologs of MURR1. *J Biol Chem*. 2005;280(23):22222-22232.
- Ofori-Acquah S, Ghosh S, Adisa O. Acute Elevation Of Protein-Free Plasma Heme Triggers Acute Chest Syndrome In Mouse Models Of Sickle Cell Anemia [abstract]. *American Journal of Respiratory and Critical Care Medicine*. 2011;183(1 MeetingAbstracts):A3757.
- Fardo D, Celedón JC, Raby BA, Weiss ST, Lange C. On dichotomizing phenotypes in family-based association tests: quantitative phenotypes are not always the optimal choice. *Genet Epidemiol*. 2007;31(5):376-382.
- Tennessen JA, Bigham AW, O'Connor TD, et al; Broad GO; Seattle GO; NHLBI Exome Sequencing Project. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*. 2012;337(6090):64-69.
- Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008;6(5):e107.
- Zeller T, Wild P, Szymczak S, et al. Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS ONE*. 2010;5(5):e10693.
- Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. *Lancet*. 2010;376(9757):2018-2031.
- Zheng L, Liang P, Li J, et al. ShRNA-targeted COMMD7 suppresses hepatocellular carcinoma growth. *PLoS ONE*. 2012;7(9):e45412.
- Miller AC, Gladwin MT. Pulmonary complications of sickle cell disease. *Am J Respir Crit Care Med*. 2012;185(11):1154-1165.
- Bean CJ, Boulet SL, Ellingsen D, et al. Heme oxygenase-1 gene promoter polymorphism is associated with reduced incidence of acute chest syndrome among children with sickle cell disease. *Blood*. 2012;120(18):3822-3828.
- Duckworth L, Hsu L, Feng H, et al. Physician-diagnosed asthma and acute chest syndrome: associations with NOS polymorphisms. *Pediatr Pulmonol*. 2007;42(4):332-338.
- Sullivan KJ, Kissoon N, Duckworth LJ, et al. Low exhaled nitric oxide and a polymorphism in the NOS 1 gene is associated with acute chest syndrome. *Am J Respir Crit Care Med*. 2001;164(12):2186-2190.
- Svensson CI, Lucas KK, Hua XY, Powell HC, Dennis EA, Yaksh TL. Spinal phospholipase A2 in inflammatory hyperalgesia: role of the small,

- secretory phospholipase A2. *Neuroscience*. 2005; 133(2):543-553.
46. Lucas KK, Svensson CI, Hua XY, Yaksh TL, Dennis EA. Spinal phospholipase A2 in inflammatory hyperalgesia: role of group IVA cPLA2. *Br J Pharmacol*. 2005;144(7):940-952.
47. Styles LA, Schalkwijk CG, Aarsman AJ, Vichinsky EP, Lubin BH, Kuypers FA. Phospholipase A2 levels in acute chest syndrome of sickle cell disease. *Blood*. 1996;87(6):2573-2578.
48. Styles LA, Aarsman AJ, Vichinsky EP, Kuypers FA. Secretory phospholipase A(2) predicts impending acute chest syndrome in sickle cell disease. *Blood*. 2000;96(9):3276-3278.
49. Styles L, Wager CG, Labotka RJ, et al; Sickle Cell Disease Clinical Research Network (SCDCRN). Refining the value of secretory phospholipase A2 as a predictor of acute chest syndrome in sickle cell disease: results of a feasibility study (PROACTIVE). *Br J Haematol*. 2012;157(5):627-636.
50. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*. 2009;106(23):9362-9367.