Multiple Testing in the Context of Gene Discovery in Sickle Cell Disease Using Genome-Wide Association Studies

Kevin H.M. Kuo[1,2]

[1]Departments of Medical Oncology and Hematology and Medicine, University Health Network, Toronto, ON, Canada. [2]Division of Hematology, Department of Medicine, University of Toronto, Toronto, ON, Canada.

ABSTRACT: The issue of multiple testing, also termed multiplicity, is ubiquitous in studies where multiple hypotheses are tested simultaneously. Genome-wide association study (GWAS), a type of genetic association study that has gained popularity in the past decade, is most susceptible to the issue of multiple testing. Different methodologies have been employed to address the issue of multiple testing in GWAS. The purpose of the review is to examine the methodologies employed in dealing with multiple testing in the context of gene discovery using GWAS in sickle cell disease complications.

KEYWORDS: Sickle cell disease, genome-wide association study, multiple testing

Introduction

Sickle cell disease (SCD) is one of the most common monogenic diseases in the world.[1] The missense mutation from glutamic acid to valine at the sixth codon (E6V) of the β-globin gene (*HBB*) leads to the production of the sickle β-globin (βS), which heterodimerizes with α-globin (α) to form the sickle hemoglobin ($\alpha_2\beta^S_2$). The sickle hemoglobin polymerizes under lower oxygen tension and forms sickle-shaped red blood cells (RBC). Sickle cell disease possesses considerable clinical heterogeneity even though all patients with SCD have the same mutation and produce one biochemical phenotype, ie, the sickle hemoglobin. At least 23 complications have been described, including stroke, retinopathy, acute chest syndrome (ACS), pulmonary hypertension, avascular necrosis, painful vaso-occlusive episodes, nephropathy, skin ulcers, and priapism, to name a few of the more prevalent ones. Patients can have different combinations of these complications: a patient may have stroke, retinopathy, and leg ulcers, whereas another may have ACS, frequent painful vaso-occlusive episodes, nephropathy, and avascular necrosis. The severity, age of onset, and rate of progression for each clinical complication also differ from patient to patient. Some patients will suffer the most severe forms of these complications and die at a relatively young age, but a few may go through a large part of their life without knowing that they have SCD even though they are homozygous for the *HBB*E6V mutation.

Several mechanisms with their causal biochemical pathways have been proposed for these complications, including increased intravascular hemolysis of the sickle RBC, vaso-occlusion of the sickle RBC in small caliber vessels, modification of nitric oxide metabolism, and endothelial dysfunction, but the actual genetic modifiers governing these complications remained virtually unknown. Up until the late 1990s, 2 biochemical and genetic modifiers were known to affect the severity (in a very broad and general sense) of SCD: α-thalassemia mutation and fetal hemoglobin fraction. Coinheritance of SCD with α-thalassemia is associated with a reduction in the severity of some clinical presentation (stroke, proliferative retinopathy, splenic function, priapism, albuminuria, cholelithiasis, and leg ulcer)[2–4] because of its effect on sickle hemoglobin concentration[5] but has no influence on other complications (eg, painful vaso-occlusive episodes).[6] Higher fetal hemoglobin (HbF) fraction, as a result of hereditary persistence of fetal hemoglobin (HPFH) or augmented by fetal hemoglobin induction agents such as hydroxyurea,[7] is associated with a reduction in the number of painful vaso-occlusion episodes and ACS,[7] proliferative retinopathy, and improved overall survival.[8] Initial haplotype mapping of the β-globin locus in linkage studies of patients with SCD revealed 4 distinct haplotypes: Senegal, Benin, Bantu, and Arab/Indian.[9] Patients carrying the Senegal or Arab/Indian haplotype have the highest HbF levels and a milder clinical course, whereas patients with the Bantu haplotype have the lowest HbF levels and consequently the most severe clinical manifestation of SCD.[10] However, these genetic modifiers did not explain most of the clinical heterogeneity seen in SCD, and in particular, the sickle cell haplotypes alone were not sufficient in explaining the variations in HbF expression.

Genome-Wide Association Study

The completion of the human genome project and the HapMap project, a catalog of common genetic variants in the human genome known as single-nucleotide polymorphisms (SNPs),[11] provided the necessary foundation to conduct genome-wide association studies (GWAS) in search of genetic modifiers in human diseases. Prior to this, searches for disease-causing genetic variants were limited to familial linkage analysis and candidate gene studies. There are several disadvantages to these methods. Linkage analysis is designed to detect single genes with major effect, requiring large families or sibling pairs,[12] and candidate gene studies require an a priori biological hypothesis which directs the search within a small segment of the genome containing genes that may play a plausible role in the phenotype.[13] Genome-wide association study (GWAS), however, is a search within the entire genome to identify genetic variations associated with observable traits.[14] In GWAS, each SNP is examined for its association with a dichotomous phenotype (as in a case-control study) or correlation with a quantitative trait (eg, blood pressure or HbF level), and upward of several millions of SNPs are examined in each study. The number and type of SNPs tested depend on the microarray platforms employed, but they are all designed with the goal of providing maximal coverage of variations across the entire genome. Genome-wide association study rests on a hypothesis-free approach that does not presuppose that any of the positive SNPs are causal SNPs but rather they are near the causal genetic element (an SNP, copy number variation, insertion, or deletion) via linkage disequilibrium (LD). Linkage disequilibrium refers to 2 loci having a high probability of transmitting together from one generation to the next due to their close proximity to each other on the chromosome or their high likelihood of co-inheritance driven by selection or population stratification.[15] These features imply that GWAS is able to detect multiple loci-trait associations of modest effect size, ie, strength of the contribution to the trait. There are a number of association tests available in calculating the association between an SNP and the observable trait, depending on whether an a priori genetic model (dominant, recessive, co-dominant) is available or not,[14] the discussion of which is beyond the scope of this article.

Multiple Testing in GWAS

Each association test between an SNP and a trait is essentially a χ^2 test if it is a categorical trait or a linear regression test if the trait is continuous and follows a normal distribution. Thus, the testing of a million SNPs is a million χ^2 tests (or linear regressions), each with its own null hypothesis. Each test would yield its own P value. The consideration of whether a particular P value is significant is based on the significance threshold assigned a priori during study design. The significance threshold, or α level, is the probability of rejecting the null hypothesis when the null hypothesis is true. This is also termed type I error. Consider an α level of 0.05: assuming the null hypothesis is true, we can expect to find one SNP to be "significantly" associated with the trait in question simply by chance alone while examining 20 SNPs. Examining 1 million SNPs would theoretically produce 50 000 "significant" SNPs by chance alone if an α level of 0.05 is chosen. To express the same concept in terms of probability, where α = α level and k = the number of tests, the probability of finding one positive SNP = $1 - (1 - \alpha)^k$.[16] It is virtually guaranteed (the probability approaches 1) that one will find at least one false-positive SNP when examining a million SNPs, assuming that the null hypothesis is true, ie, no association between any of the SNPs examined and the trait.[17] Therefore, the significance threshold for an association test with each individual SNP (local significance level) must be much lower to take into account the possibility of false discovery. Furthermore, imputation, or an estimation of the allelic dosages of untested SNPs using genotyped SNPs, is often necessary because some SNPs may fail quality control. Meta-analyses of GWAS also use imputation to combine studies conducted on different microarray platforms. Each imputed SNP is associated with a degree of uncertainty that needs to be accounted for as well. Family-wise error fraction (also termed family-wise error rate [FWER]) is the probability of making at least one false discovery when performing multiple testing.[14] Bonferroni correction is the simplest approach where the local significance level (α_l) is the global error fraction (α_g) one aims to control divided by the number of tests performed (M), yielding $\alpha_l = \alpha_g / M$. Consider an example where the α_g is set at 0.05, and a microarray of 1 million SNPs was employed in the GWAS. This would yield an α_l of 5.0×10^{-8}. A similar method called the Sidak correction, where $\alpha_l = 1 - (1 - \alpha_g)^{1/M}$, yields similar result to the Bonferroni correction when M is large.[18] However, most methodologists consider the Bonferroni correction to be too stringent and conservative, thus inappropriate for many of the GWA studies.[14,17–23] Reasons for not using the Bonferroni correction are 2-fold. Although the Bonferroni and Sidak corrections control for false-positive discoveries very well, their inability to tolerate even one false discovery at a defined probability leads to an increase in false-negative discoveries, ie, not rejecting the null hypothesis of no association between an SNP and a trait while the null hypothesis is false.[19] This will potentially lead to missing out SNPs that may be truly associated. Also, many SNPs are in close LD to one another, but Bonferroni and Sidak corrections assume that each SNP (and thus each association test) to be independent of each other and fails to take LD into account. Methods such as the Sidak correction also assume uniformly distributed P values under the null hypothesis.[24] Population stratification and admixture can confound the genotype-phenotype association, thus creating departures from the uniform P value distribution expected by the Sidak correction,[24] making it an inappropriate choice for estimating significance threshold in GWAS.

Table 1. Using FDR to estimate significant threshold in multiple testing.

P VALUE	I	M	FDR THRESHOLD	ACCEPT/REJECT NULL
.000001	1	17	0.002941176	Reject
.000013	2	17	0.005882353	Reject
.000065	3	17	0.008823529	Reject
.00063	4	17	0.011764706	Reject
.0008	5	17	0.014705882	Reject
.0017	6	17	0.017647059	Reject
.0032	7	17	0.020588235	Reject
.0065	8	17	0.023529412	Reject
.0148	9	17	0.026470588	Reject
.049	10	17	0.029411765	Accept
.094	11	17	0.032352941	Accept
.11	12	17	0.035294118	Accept
.15	13	17	0.038235294	Accept
.24	14	17	0.041176471	Accept
.45	15	17	0.044117647	Accept
.56	16	17	0.047058824	Accept
.87	17	17	0.05	Accept

Abbreviation: FDR, false discovery rate.
In this example adapted from Benjamini et al, 17 tests were performed, and the *P* values were rank ordered from the smallest to the largest. The FDR threshold was then calculated and any *P* value that is smaller than the FDR threshold is where the null hypothesis is rejected. Adapted with permission from Benjamini et al (2001).

Other methods to estimate significance threshold in GWAS

Methodologists have proposed other methods to estimate significant threshold in GWAS. These methods center around 4 approaches: controlling the false discovery rate (FDR), estimation of the effective number of independent SNPs by accounting for LD, permutation testing, and Bayesian approach.

False discovery rate was devised by Benjamini and Hochberg in 1995 as a solution to the drawbacks faced by FWER. False discovery rate is defined as the expected proportion of false-positive associations among all associations that were declared significant[25] and can be expressed mathematically as $f_p/(f_p + t_p)$ where f_p is the number of false-positive associations and t_p is the number of true-positive associations. This is different from false-positive rate because the FDR is the proportion of false-positive associations among all associations, whereas false-positive rate is the average proportion of associations that will be declared significant assuming that the null hypothesis is true.[19] In FDR, the local significance threshold changes depending on the rank order of each SNP's *P* value while taking into the total number of SNPs examined into account. To use FDR, one first chooses a global significance level, such as 0.05. One then rank orders the *P* value of the SNPs from the smallest to the largest (Table 1, column 1). The local FDR significance threshold (FDR) for each rank (*i*) is then calculated by multiplying the global significance level (α, eg, 0.05 as in the example here) with the rank divided by the total number of tests (*m*) ($FDR_i = \alpha(i/m)$; Table 1, column 4). Finally, one compares the *P* value of each rank with the local FDR significance threshold (FDR_i), and the null hypothesis is rejected for any *P* value that is lower than the FDR_i (Table 1, column 5).[25] A modified method termed the "local FDR" estimates "the probability of a given null hypothesis to be true according to the specific *P* value of each genetic marker tested."[19] The calculation requires an estimation of the distribution of *P* values under the null hypothesis and the alternative hypothesis.[19] The advantage of the FDR and local FDR is that they are much less likely to eliminate true associations (false negatives) at the expense of having an acceptable proportion of false-positive association. What is considered "acceptable" is defined by the investigator based on the need and stage of the study. The disadvantage of the FDR and local FDR in the setting of GWAS is that it fails to account for the LD between SNPs because it considers each SNP to be independent of each other.

To account for LD between SNPs, several authors have derived significant thresholds based on an estimation of the effective number of independent SNPs. Pe'er et al[26] derived a set of "genome-wide test burden" values based on data collected by the International Haplotype Map Consortium, which when multiplied with the nominal *P* values provide "a practical, first-cut guideline" for correcting nominal *P* values. The resultant significant threshold is approximately 1×10^{-7} for the European HapMap but higher for the Yoruba HapMap (YRI; derived from an African tribe) at 1×10^{-6} because there are more SNPs and lower LD between SNPs in the YRI HapMap.[26] Dudbridge and Gusnanto provided an estimate of the local significance level at 7.2×10^{-8} via a permutation approach based on the Wellcome Trust Case Control Consortium data.[27] Gao et al[18] proposed a method of deriving the effective number of independent tests (M_{eff}) in a study by first computing the eigenvalues from the pairwise SNP correlation matrix created with composite LD correlation and then derive the M_{eff} using principal component analysis. The α_l can then be calculated by taking α_g and dividing by M_{eff}. They have shown that the α_l derived by this method is very close to permutation-based methods. Gao also compared his method of M_{eff} with SLIDE, a method that assumes a asymptotically multivariate normal distribution of commonly used association statistics, to examine whether such estimation method can provide an approximation of α_l to the computationally demanding permutation-based method when imputed SNPs are considered.[28] The author found that the M_{eff} method provided the closest approximation to the permutation method using the least computation time.[28] In a similar approach to Gao et al, Duggal et al[21] evaluated the effective number of "independent" SNPs in the Illumina 317 K and Affymetrix 500 K marker sets and derived the α_l to be at 1.21×10^{-5} and 1.49×10^{-5}, respectively, which are significantly less stringent than the conservative estimate of Bonferroni correction. Finally, the Bayesian approach of addressing the issue of multiple testing involves accepting or rejecting the null hypothesis of association between an SNP and the trait under study based on whether the posttest probabilities is above the threshold Bayes factor (BF). The posttest probabilities are derived from assumed pretest probabilities in conjunction with the data.[17] The predetermined threshold BF is determined by conducting simulations to compute the expected number of false-positive associations for different threshold BF.[29] A threshold BF is chosen such that it will reduce the probability of obtaining a false-positive association to an acceptable level.[29]

Power and sample size

Power in a GWAS depends on a number of factors, including the sample size available, the putative effect size of the associated SNP, the strength of the LD between the associated SNP and the causal locus, the minor allele frequencies of the associated SNP and the causal locus, and the number of SNPs tested.[30] Methods in determining the significance threshold can indirectly influence power substantially because a more stringent significance threshold, such as a Bonferroni correction, will require a trait-associated SNP to have a larger effect size to be declared significant or needing a larger sample size for a given effect size.[23,30] Despite the fact that SCD is the most common monogenic disease in the world, the prevalence is far lower than common disease such as heart attack or stroke. Sickle cell disease often remained underdiagnosed, especially if the complications are insidious in nature, such as osteonecrosis, silent cerebral infarcts, pulmonary hypertension, proliferative retinopathy, and obstructive lung disease. A few epidemiologic registries are underway in SCD, including the National Haemoglobinopathy Registry in the United Kingdom. The combination of such registries with genetic sampling can become a powerful tool in the search for genetic modifiers in SCD. Until such registries and biobanks come to fruition, subject availability will continue to be a major limiting factor on power in GWAS-involving SCD traits. Methods such as estimation of effective number of SNPs, FDR, permutation, and Bayesian methods provide the means to improve power with limited sample size while striking a balance between the risk of eliminating false-negative association and having an acceptable amount of false-positive associations.[17]

Multistage Design, Replication, and Meta-Analysis of GWAS

Regardless of the method used to correct for multiple testing, one cannot be certain whether the discovered SNPs in one GWAS cohort are true associations or false positives, even with highly stringent methods such as the Bonferroni correction. Therefore, any putatively associated SNPs that were discovered on the first GWAS cohort must undergo further scrutiny. There are 3 methods in which this can be achieved: multistage design, replication, and meta-analysis.

In multistage design, candidate SNPs identified from the initial stage of GWAS dictate the search for association in the next stage. The search in the next stage is restricted to genomic regions that contain these candidate SNPs so that the search can employ higher density SNPs to hone in to the interested regions. The data from the initial and subsequent stages of the GWAS are then combined to form the result. The advantages of a multistage design is 3-fold: cost is reduced by genotyping a smaller number of subjects without having to examine the entire genome at later stages, mapping of interested regions in later stages on a finer scale, and using different genotyping platforms at different stages of the study avoid false-positive reports and any technical artifacts that are specific to one platform.[14] The use of different genotyping platforms at different stages of a multistage design and in replication studies is currently considered the gold standard.[14,30]

Replication of identified associations between SNPs and the phenotypic trait under examination is essential and is considered the gold standard in differentiating true associations

from false-positive associations in GWAS.[14,31,32] There are a number of approaches to replication. Some investigators will only carry forward a small number of candidate SNPs from the initial GWAS, selected by a stringent significant threshold, to the next replication cohort. Others will use a smaller cohort in the initial GWAS and then carry forward a larger number of candidate SNPs to the next replication study using a more lax significant threshold to minimize the risk of missing any false-negative associations.[31] Irrespective of the approach chosen, in order for a replication study to be considered valid, it should have (1) sufficient sample size to convincingly distinguish the proposed effect from no effect, (2) been conducted in independent data sets, (3) same or a very similar phenotype, (4) similar population, (5) similar magnitude of effect and significance in the same direction with the same SNP or an SNP in very high LD, (6) reporting of statistical significance should be derived from the same genetic model used in the initial study (7) a joint or combined analysis should lead to a smaller *P* value than the one observed in the initial study, (8) a strong rationale for selecting SNPs for replication, based on putative functional data or published literature, and (9) the same level of detail of study design and analysis as the initial study.[33] These criteria were established by the Working Group on Replication in Association Studies from the National Cancer Institute and the National Human Genome Research Institute as a way of standardizing the interpretation of results from replication studies.

As discussed previously, increase in stringency of the significance threshold will result in the reduction in false-positive associations but concomitantly reduce power. Power loss can be compensated by increasing the number of subjects in the study, but this is often not possible in the study of rare diseases such as SCD. Meta-analysis of GWAS has been employed as a solution to overcome the limiting factor of small sample size. The purpose of performing meta-analysis in GWAS is to increase power to achieve significance that exceeds a study-wide threshold and to prioritize SNPs for subsequent studies. There are 3 approaches to perform meta-analysis in GWAS: analysis of the aggregated data from different studies at the log odds ratio level, retrospective pooled analysis of individual data from the primary studies, and prospectively planned pooled analysis of individual data from several studies.[14] Both random and fixed effects models can be employed in the analysis. Random effects model produces a more accurate estimate and is considered to be more conservative, as it takes heterogeneity between studies into account, whereas the fixed effects model can lead to false-positive associations because of overconfidence (lower *P* value) in results when there is considerable heterogeneity between studies.[34] However, Cantor et al[34] argues that it is not critical to have an accurate estimate of the association when the goal of the meta-analysis is to prioritize candidate SNPs for future investigations and they surmise that this may be one of the reasons why fixed effects model is the more popular choice.

Recently, investigators have conducted mega-analyses with multiple GWAS to increase the sample size and discovery power. Mega-anlaysis refers to the analysis of combined raw microarray data and outcomes data from multiple GWA studies. Although this method has been attempted in cardiovascular medicine[35] and psychiatry,[36] no such attempts were made in the field of SCD.

GWAS in SCD

Genetic modifiers in association with fetal hemoglobin variations

The most successful discovery of genetic modifiers using GWAS in SCD thus far has been in the realm of HbF level. Thein et al[37] conducted a candidate gene association study in 2041 non–sickle cell monozygotic and dizygotic twin pairs and unrelated individuals, which identified several SNPs within the HBS1L-MYB region as being strongly associated with F cell (HbF) levels. The first GWAS by Menzel et al genotyped using a 308 000 SNP arrays in 179 unrelated individuals, selected for extremes of F-cell distribution, identified markers near *BCL11A* (*P* values between 4.6×10^{-8} and 2.5×10^{-20}), 3 markers within 6q23 (*P* values between 8.2×10^{-6} and 2.8×10^{-27}) later confirmed to be between the *HBS1L* and *MYB* genes, and *Xmn*1 polymorphism at 158 base pairs (bp) upstream of the $^{G}\gamma$ globin (*HBG*) gene (2.0×10^{-30}) as being significantly associated with F-cell levels.[38] The findings were replicated in 90 individuals, again with extremes in F-cell distribution, and in an unselected 720 twins cohort (ie, not selected for extremes of F-cell distribution). Most of these markers retained similar levels of significance on replication. In parallel, Uda et al[39] identified a strong association between SNP rs11886868 in BCL11A gene and HbF levels in a GWAS with 362 129 SNPs in 4305 Sardinians. The C variant of this SNP was more frequent in heterocellular HPFH and in homozygous β^{0}-thalassemia with mild phenotype compared with the severe form.[39] The same study also confirmed the SNPs between *MYB* and *HBS1L* genes which were associated with elevated HbF levels. Lettre et al[40] was the first to replicate these findings in a cohort of sickle cell patients by showing that the SNPs discovered in the previous retained their significance within a cohort of 1275 North American and 350 Brazilian patients with SCD. The *BCL11A* finding was further replicated in an independent cohort of 255 SCD individuals from the United States.[41] To search for other genetic modifiers of HbF level, Solovieff et al[42] conducted a GWA study in 1153 African Americans (848 individuals in the discovery set and 305 in the validation set). Single-nucleotide polymorphisms centered around *BCL11A* again were significantly associated with HbF level, but new SNPs surrounding olfactory receptor genes on chromosome 11 (lowest *P* = 4.7×10^{-8}) were also identified. The investigators defined the significance threshold as <1.0×10^{-6} but did not explicitly state the method of deriving the

threshold. Bhatnagar et al[43] conducted a GWAS on the F-cell level of 440 African American patients with SCD from the Silent Infarct Transfusion (SIT) trial cohort using 661 000 SNPs. The investigators determined the significance threshold by permutation method, resulting in a value of 1.27×10^{-7}. This is significantly less conservative than what would be expected if Bonferroni correction was employed (7.56×10^{-8}). Also of note is that unlike other GWAS in SCD, Bhatnagar et al explicitly defined their method of deriving the significance threshold. In addition to confirming the association of *BCL11A* with F-cell variation, sex-stratified analysis also identified an SNP in the glucagon-like peptide-2 receptor (*GLP2R*) gene reaching genome-wide significance as defined by the investigators. Bae et al[44] recently conducted a meta-analysis of 7 cohorts of African Americans with SCD totaling 2040 patients in 585 563 common SNPs to locate loci of modest effect size. The investigators chose 5.0×10^{-8} as the significance threshold. Although SNPs from *BCL11A* and *HMIP* (a gene between the *HBS1L* and *MYB* genes) were successfully replicated, SNPs from the olfactory receptor genes did reach significance. Mtatiro et al[45] conducted a GWAS in sickle cell anemia (HbSS and HbS/β^0-thalassemia) patients from Tanzania and the United Kingdom. The discovery phase assayed ~2.4 million SNPs in 1213 individuals from Tanzania. The replication cohort included 321 patients from the United Kingdom and included 16 SNPs from 10 loci. The investigators used the 1000 Genomes Phase 1 release data for imputation. Although the investigators were able to replicate BCL11A and HMIP, they were not able to replicate 8 other associations with $P < 10^{-6}$ in the United Kingdom SCD replication cohort.[45] Functional study by Xu et al demonstrated that *BCL11A* encodes a zinc-finger transcription factor and is critical in HbF switching by occupying the upstream locus control region and γ-δ intergenic regions of the β-globin locus and via interaction with corepressor complexes, Mi-2/NuRD, and LSD1/CoREST, as well as the erythroid transcription factor GATA-1 and the HMG-box protein SOX6.[46] Observations and functional studies also confirmed the biologic significance of the HBS1L-MYB region on HbF expression. Sankaran et al[47] observed microRNA miR-15a and miR-16-1 to act via MYB to elevate fetal hemoglobin expression through mapping of 57 partial trisomy 13 cases in humans. Suzuki et al[48] demonstrated in a mouse model that the disruption of HBS1L-MYB locus result in HPFH, of which the downregulation of MYB suppresses the KLF1/BCL11A pathway, resulting in activation of fetal globin gene expression. Pule et al[49] has shown that the treatment of ex vivo differentiated primary erythroid cells from 7 unrelated individuals and K562 cells (immortalized erythroleukemic cells) with hydroxyurea, a known HbF inducer, resulted in downregulation of MYB, BCL11A, and KLF-1 and upregulation of γ-globin (thus HbF) expression. The discovery of *BCL11A* and the region between HBS1L-MYB as crucial regulators of HbF variation illustrates the importance of multiple replications in validating any GWAS discovery leading to identification of meaningful regulators of phenotypic variation with therapeutic potential.

In search of genetic modifiers in association with other SCD variations

The success of discoveries in genetic modifiers governing HbF variation using GWAS prompted the search for genetic modifiers in other SCD traits using the GWAS approach. A search through Medline using the Medical Subject Heading (MeSH) terms "Genome-Wide Association Study"[Mesh] AND "Anemia, Sickle Cell"[Mesh] revealed 30 citations. One article in which the search failed to locate but the author has knowledge of was also included. After excluding review articles, 8 GWAS covering 8 traits (not including the aforementioned studies in HbF variation) were examined. This included stroke, systolic blood pressure, ACS, painful crises, hemolysis, bilirubin and cholelithiasis, hemoglobin A2 (HbA2), and SCD disease severity score (Table 2).

Approximately 11% of the patients with SCD will have an overt stroke by 20 years of age.[50] Sibling-pair study had shown that a genetic component exists in SCD strokes.[51] Flanagan et al[52] conducted a GWAS in which they genotyped 512 patients (177 with stroke and 335 having no stroke as controls) from various sources including the Hustle study, SWiTCH study, Cooperative Study of Sickle Cell Disease (CSSCD), and the Comprehensive Sickle Cell Centers Collaborative Data Project.[52] They interrogated ~770 000 SNPs but found none that reached genome-wide significance, which the investigators defined as 5×10^{-8}. The investigators did not specify the method used to derive the significance threshold. The stringency of the threshold chosen by this study was quite similar to the one obtained via the Bonferroni method (6.5×10^{-8}). For purposes of illustration, if FDR was the chosen method of establishing significance, the *P* value of the top 4 SNPs versus their respective corrected FDR threshold would be 2.71×10^{-7} vs 6.49×10^{-8}, 3.13×10^{-7} vs 1.30×10^{-7}, 5.52×10^{-7} vs 1.95×10^{-7}, and 9.77×10^{-7} vs 2.60×10^{-7}. Thus, employment of the FDR method would also have resulted in the rejection of the candidate SNPs. Instead of taking the top candidate SNPs and replicating the findings in other cohorts or to perform a meta-analysis with other cohorts, the investigators took a whole-exome sequencing approach with the same cohort of patients and found 22 candidate variants. Validation study and in combination with the GWAS data resulted in isolation of 2 associated variants, but none of these 2 appeared in the top 10 SNPs identified by the initial GWAS.

The study by Bhatnagar et al[53] examined the genetic determinants of systolic blood pressure in SCD. A previous study identified a higher systolic blood pressure (not hypertension) as a risk factor for development of silent cerebral infarction in SCD children.[53] The study used 2 unrelated admixed African American SCD cohorts from 2 different studies: the

Table 2. Summary of genome-wide association studies in sickle cell disease complications, with focus on the discovery cohort size, array size, genome-wide significance level (α level), derivation of the significance threshold and the presence or absence of a multistaged design, replication studies, or meta-analysis.

STUDY	DISCOVERY COHORT SIZE	ARRAY SIZE	α LEVEL	DERIVATION METHOD	MULTISTAGED, REPLICATION, OR META-ANALYSIS	COHORT SIZE
Trait: stroke						
Flanagan, 2013		512	770 000	5.0×10^{-8}	Not stated	No
No SNPs passed the significance threshold						
Trait: systolic blood pressure (surrogate marker for silent cerebral infarction)						
Bhatnagar, 2013	573 692	661 000 600 000	5.0×10^{-8}	Bonferroni	2-staged Meta-analysis	509 1617
No SNPs passed the significance threshold						
Trait: acute chest syndrome						
Galarneau, 2013	1514	237 643	1.0×10^{-4}	FDR	Multistaged	387 318 449
17 SNPs were declared significant in the discovery cohort, 1 SNP was replicated in the combined discovery and replication cohorts, the same SNP was also replicated and reached genome-wide significance after combining the data from all 4 cohorts						
Trait: painful crises						
Galarneau, 2013	1514	237 643	1.0×10^{-4}	FDR	Multistaged	387 318 449
19 SNPs were declared significant in the discovery cohort, none of the SNPs were replicated in a combined discovery and replication CSSCD cohorts, and none of the SNPs reached genome-wide significance after combining the results from all 4 cohorts						
Trait: hemolysis (hemolytic score)						
Milton, 2013	1117	569 554	1.0×10^{-8}	Not stated	Replication Meta-analysis	745 213 2075
Although none of the SNPs reached the significance threshold stated by the investigators, 4 SNPs were very close to the significant threshold in the discovery set (5.87×10^{-5} to 6.04×10^{-7}) and remained significant in the replication sets. All 4 SNPs were significant on meta-analysis of the 3 cohorts						
Trait: bilirubin and cholelithiasis						
Milton, 2012	1117	569 615	5.0×10^{-8}	Not stated	Replication	195 522 530 905
15 SNPs had a $P < 5.0 \times 10^{-8}$ in discovery cohort and retained their significance in the all the replication cohorts						
Griffin, 2014	618	~600 000	1.0×10^{-5}	Not stated	Replication	128 45 580
14 SNPs had a $P < 1.0 \times 10^{-5}$ in discovery cohort and 2 achieved nominal replication, 1 achieved genome-wide significance						
Trait: SCD disease severity score ("mild" vs "severe" disease)						
Sebastiani, 2010	1265	600 000	>1000 BF	Bayesian	Replication	163
The investigators discovered 40 SNPs in the discovery cohort and 5 were replicated in the replication cohort. BF refers to Bayes factor because a Bayesian approach was undertaken in association tests						

Abbreviations: CSSCD, Cooperative Study of Sickle Cell Disease; FDR, false discovery rate; SNPs, single-nucleotide polymorphisms.

SIT trial cohort and the CSSCD cohort. The SIT trial cohort was divided into 2 subsets, the first subset (N = 573) was genotyped on a ~661 000 SNPs array, and the second subset (N = 509) was genotyped on a ~1 140 000 SNPs array. The

CSSCD cohort (N = 692) was genotyped on a ~600 000 SNPs array. Meta-analysis of the results from both cohorts was performed with 1 019 297 evaluable SNPs. The genome-wide significance level of 5.0×10^{-8} was determined by the Bonferroni method, and none of the SNPs examined in the discovery sets or the meta-analysis reached genome-wide significance. The top scoring SNP had a *P* value of 8.57×10^{-7} which was close to the prespecified significance threshold, but the rest were in the range of 10^{-6} to 10^{-5}.

The study by Galarneau et al[54] involved the search for genetic associations with painful crises and ACS. The CSSCD cohort (N = 1514) was used as the discovery cohort and 237 643 SNPs were tested. False discovery rate was used to correct the local *P* values of each SNP, and the authors chose a local significance level of 1×10^{-4} as this provided a 50% power for a quantitative trait (eg, frequency of painful crisis and ACS) assuming a minor allele frequency of 25% and 1% of variance explained. A total of 36 SNPs (19 SNPs for painful crises and 17 SNPs for ACS) were found to be smaller than 1×10^{-4}. As a comparison, a local significance threshold of 2.1×10^{-7} would be required to declare an SNP significant using the Bonferroni method. The investigators then genotyped these 36 candidate SNPs in 387 patients from the CSSCD who were independent from the discovery cohort, 318 patients with SCD from Georgia Health Sciences University (GHSU), and 449 patients from the Duke SCD cohort. Combining the results of the CSSCD discovery and replication cohorts with the Duke and GHSU cohorts resulted in one SNP reaching genome-wide significance ($P = 4.1 \times 10^{-7}$), whereas other SNPs failed to replicate.

Hemolysis is one of the main pathogenic mechanisms that lead to complications in SCD. Milton et al[55] conducted a GWAS where they genotyped 569 554 SNPs in 1117 patients from CSSCD against the phenotype of hemolytic score that characterized the degree of hemolysis in the discovery cohort. The investigators considered 10^{-8} as the significance threshold but did not specify how they derived it. No SNP reached the significant threshold. The significance threshold chosen by the investigators might have been too stringent considering that it was quite close to the threshold derived by Bonferroni correction (8.3×10^{-8}). The investigators then selected top 4 candidate SNPs that had $P < 5.0 \times 10^{-4}$ and successfully replicated them in the replication sets of 745 patients from the Walk-PHaSST and Pulmonary Hypertension and the Hypoxic Response in Sickle Cell Disease (PUSH) study and 213 patients from a London UK SCD cohort. The investigators also performed a meta-analysis where all 4 SNPs met genome-wide significance.

Milton et al[56] conducted a GWAS of total bilirubin and risk of cholelithiasis analyzing 569 615 SNPs, again using the 1117 patients from the CSSCD as the discovery cohort. Fifteen SNPs reached prespecified significance threshold of 5×10^{-5}. However, the 15 SNPs were in strong LD to each other, and adjustment for the first top SNP resulted in the lack of

independent association with bilirubin, suggesting that all 15 SNPs were indeed related to one another. In total, 12 of the same 15 SNPs were also associated with the risk of cholelithiasis. The replication cohort consisted of an aggregate of 2152 patients from the Duke cohort (N = 530), the MSH (N = 195), Walk-PHaSST (N = 522), and the SIT study (N = 905). All 15 SNPs isolated from the discovery cohort were successfully replicated. The fact that all 15 SNPs belong to the UGT1A1 gene which is responsible for glucuronidation of unconjugated bilirubin lends biological support to the discovery.

Hemoglobin A2 is composed of 2 α-globins and 2 δ-globins with a physiologic function indistinguishable from adult hemoglobin. Its expression is higher in the presence of HbS compared with non-SCD adults. As HbA2 has the potential of inhibiting the polymerization of HbS, understanding the genetic variability of HbA2 expression in patients with SCD may open the door to the development of antisickling therapy. Griffin et al[57] in 2014 reported a GWAS of HbA2 variability using a discovery cohort of 618 unrelated African Americans from the CSSCD study. The replication cohort consisted of 128 African American patients from the Walk-PHaSST study, 45 African Americans from the PUSH study, and 580 Chinese from the Hong Kong β-thalassemia trait study. All were genotyped using the Illumina Human610-Quad array. Replication attempts were performed on 14 SNPs that had a $P < 10 \times 10^{-5}$. Two SNPs (rs766432 and rs10195871) achieved nominal replication with one (rs766432) achieving genome-wide significance in meta-analysis, after adjusting for age and sex, but not HbF. Both of these SNPs are within BCL11A, and mediation analysis suggested that HbA2 variations are partially mediated by HbF.[57]

To quantitatively describe the burden of complications in individual patients with SCD, Sebastiani et al[29] developed a scoring system that described the risk of death within 5 years by integrating clinical and laboratory parameters. The scoring system was then validated in a cohort of European patients with SCD.[29] The investigators then genotyped ~600 000 SNPs in 1265 patients from the CSSCD cohort at the discovery stage. The evidence of association with each genotyped SNP is based on the posterior probability using Bayesian test. The significance threshold chosen by the investigators was a BF of >1000 because this level BF was expected to produce less than 1 false-positive association in 10 000 independent tests. In total, 40 candidate SNPs were strongly associated with sickle cell severity with an odds for association of >1000 isolated in the discovery set. Only 32 of the 40 SNPs could be analyzed in the replication study of 163 patients. Only 5 of the 32 SNPs replicated and 8 showed consistent effects but failed to reach the significance threshold.

Discussion

The discovery of genetic variants associated with other SCD complications has not been as successful as that for HbF

variations. Although a number of factors can contribute to the lack of success, such as case definition, sample size limitation, and population substructure, the focus of the discussion will be limited to the issue of multiple testing as a possible reason.

Only 3 of the 6 studies explicitly stated the method of deriving the significant threshold (Table 2).[29,53,54] In studies that did not specify the method of deriving the significant threshold, the values chosen by the investigators were very similar to the one derived by the Bonferroni correction.[52,55–57] Consequently, the 3 studies that failed to identify any associated SNPs (Table 2)[52–54] used the Bonferroni correction or had values similar to the Bonferroni correction, suggesting that the chosen method was overly stringent and may have resulted in false-negative findings. Conversely, the success of the GWAS conducted by Galarneau et al may be partly attributed to the investigators choosing a less stringent threshold (1.0×10^{-4}) at the discovery stage, using a different method in adjusting for multiple testing (FDR), and employed a multistage design in their study.

Even if no association was found at the discovery stage, one does not necessary have to abandon the study, especially when the significance threshold chosen was a very stringent one. In the GWAS of hemolysis by Milton et al,[55] none of the SNPs reached the prespecified significance threshold, but the investigators chose to carry on the top 4 SNPs into the replication study and successfully replicated their findings in independent cohorts. This example nicely illustrates the following principles: using the significance threshold only as a guide and not as an absolute cut-off in the initial SNP discovery stage, the value of prioritizing candidate SNPs found in discovery cohort for further examination in replication cohorts or meta-analysis, and the value of conducting replication studies in independent populations.

Study by Sebastiani et al[29] used the Bayesian approach (Table 2), but surprisingly, none of the studies used the estimation and permutation methods. Adaptation of these less stringent methods in estimating the significance threshold may aid in identifying candidate SNPs for subsequent studies, in particular, where sample size is a limiting factor.

The challenge of GWAS is to identify and separate true associations from false-positive associations. Although many factors play a role in this process, such as sample quality control, genotyping accuracy, and population substructure, correction for multiple testing plays a central role. A number of statistical methods are available in the correction of multiple testing, including FWER (Bonferroni or Sidak), FDR, estimation of effective number of SNPs by considering the LD structure, permutation testing, and Bayesian method. There is no "one-size-fits-all" approach when choosing a particular multiple testing correction method for GWAS. However, some general rules can be gleaned from the above studies in SCD, which constitute only a fraction of all GWAS conducted to date. The choice of multiple testing correction method is partly dependent on the stage of SNP discovery. At initial stages, one can be more liberal as one does not wish to eliminate any SNPs that may be truly associated (avoid false negatives). In this case, less stringent methods such as FDR or one of the estimation methods outlined above would be suitable. Regarding this, investigators conducting GWAS in SCD complications should be encouraged to consider using less stringent methods such as the ones described in this article to derive significant thresholds. In subsequent studies where the aim is to hone in on the specific causal SNPs from a handful of candidate genes, one may consider more stringent methods such as the local FDR or the Bonferroni correction. Finally, multistage design and meta-analysis provide ways to minimize the sample size required while maximizing the power and contain costs in conducting a GWAS. Replication is essential and is the gold standard in verifying any initial SNP discovery. In the GWAS of SCD stroke by Flanagan et al,[52] although the study failed to yield any associated SNPs in the discovery cohort, the significance threshold chosen was quite conservative and was close to that of Bonferroni correction. Furthermore, the *P* values of the top 10 SNPs were in the order of 10^{-7}, which might have been declared significance if a less stringent threshold was chosen. Furthermore, the investigators did not bring forward any of these candidate SNPs for replication or meta-analysis. In this case, one can consider an attempt to replicate the study in the other existing data sets.

Regardless of the statistical significance of the associated SNP or gene, a biologically plausible connection to the phenotype must be established, via new experiments, clinical studies, or prior literature, to prove that such association is a truly causal association. This is also the means in which potential therapeutic targets and novel treatment methods can be developed and arguably is the ultimate goal of conducting GWAS in human diseases.

Author Contributions

KHMK performed the literature review, abstraction of relevant data, drafting, editing, and approval of the final manuscript.

REFERENCES

1. Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: an increasing global health problem. *Bull World Health Organ*. 2001;79:704–712.
2. Steinberg MH. Predicting clinical severity in sickle cell anaemia. *Br J Haematol*. 2005;129:465–481.
3. Weatherall DJ, Clegg JB, Blankson J, McNeil JR. A new sickling disorder resulting from interaction of the genes for haemoglobin S and alpha-thalassaemia. *Br J Haematol*. 1969;17:517–526.
4. Higgs DR, Aldridge BE, Lamb J, et al. The interaction of alpha-thalassemia and homozygous sickle-cell disease. *N Engl J Med*. 1982;306:1441–1446.
5. Thein SL. Genetic modifiers of the beta-haemoglobinopathies. *Br J Haematol*. 2008;141:357–366.
6. Platt OS, Thorington BD, Brambilla DJ, et al. Pain in sickle cell disease. Rates and risk factors. *N Engl J Med*. 1991;325:11–16.

7. Charache S, Terrin ML, Moore RD, et al. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. Investigators of the multicenter study of hydroxyurea in sickle cell anemia. *N Engl J Med*. 1995;332:1317–1322.

8. Steinberg MH, McCarthy WF, Castro O, et al. The risks and benefits of long-term use of hydroxyurea in sickle cell anemia: a 17.5 year follow-up. *Am J Hematol*. 2010;85:403–408.

9. Thein SL, Menzel S. Discovering the genetics underlying foetal haemoglobin production in adults. *Br J Haematol*. 2009;145:455–467.

10. Labie D, Pagnier J, Lapoumeroulie C, et al. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. *Proc Natl Acad Sci U S A*. 1985;82:2111–2114.

11. The International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005;437:1299–1320.

12. Teare MD, Barrett JH. Genetic epidemiology 2—genetic linkage studies. *Lancet*. 2005;366:1036–1044.

13. Bras J, Guerreiro R, Hardy J. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. *Nat Rev Neurosci*. 2012;13:453–464.

14. Ziegler A, Konig IR. *A Statistical Approach to Genetic Epidemiology: Concepts and Applications*. 2nd ed. Weinheim, Germany: Wiley-Blackwell; 2010:1–489.

15. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. *Lancet*. 2005;366:941–951.

16. Streiner DL, Norman GR. Correction for multiple testing: is there a resolution? *Chest*. 2011;140:16–18.

17. Sebastiani P, Timofeev N, Dworkis DA, et al. Genome-wide association studies and the genetic dissection of complex traits. *Am J Hematol*. 2009;84:504–515.

18. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol*. 2008;32:361–369.

19. Bouaziz M, Jeanmougin M, Guedj M. Chapter 13. Multiple testing in large-scale genetic studies. In: Pompanon F and Bonin A, eds. *Data Production and Analysis in Population Genomics: Methods and Protocols, Methods in Molecular Biology*, vol. 888. New York, NY: Springer; 2012:213–233.

20. Dudbridge F, Gusnanto A, Koeleman BP. Detecting multiple associations in genome-wide studies. *Hum Genomics*. 2006;2:310–317.

21. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics*. 2008;9:516.

22. Palmer LJ, Cardon LR. Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet*. 2005;366:1223–1234.

23. Teo YY. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Curr Opin Lipidol*. 2008;19:133–143.

24. Schrodi SJ. The use of multiplicity corrections, order statistics and generalized family-wise statistics with application to genome-wide studies. *PLoS ONE*. 2016;11:e0154472.

25. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001;125:279–284.

26. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol*. 2008;32:381–385.

27. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genome-wide association scans. *Genet Epidemiol*. 2008;32:227–234.

28. Gao X. Multiple testing corrections for imputed SNPs. *Genet Epidemiol*. 2011;35:154–158.

29. 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;85:29–35.

30. Lettre G. The search for genetic modifiers of disease severity in the β-hemoglobinopathies. *Cold Spring Harb Perspect Med*. 2012;2(10) pii:a015032

31. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008;299:1335–1344.

32. Moore JH, Asselbergs FW, Williams SM. Bioinformatics challenges for genome-wide association studies. *Bioinformatics*. 2010;26:445–455.

33. Chanock SJ, Manolio T, Boehnke M, et al; National Cancer Institute–National Human Genome Research Institute (NCI-NHGRI) Working Group on Replication in Association Studies. Replicating genotype–phenotype associations—What constitutes replication of a genotype–phenotype association, and how best can it be achieved? *Nature*. 2007;447:655–660.

34. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet*. 2010;86:6–22.

35. van't Hof FN, Ruigrok YM, Lee CH, et al. Shared genetic risk factors of intracranial, abdominal, and thoracic aneurysms. *J Am Heart Assoc*. 2016;5:e002603.

36. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 2013;18:497–511.

37. Thein SL, Menzel S, Peng X, et al. Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc Natl Acad Sci U S A*. 2007;104:11346–11351.

38. 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 Gen*. 2007;39:1197–1199.

39. 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 U S A*. 2008;105:1620–1625.

40. Lettre G, Sankaran VG, Bezerra MA. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A*. 2008;105: 11869–11874.

41. Sedgewick AE, Timofeev N, Sebastiani P, et al. BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies. *Blood Cells Mol Dis*. 2008;41:255–258.

42. Solovieff N, Milton JN, Hartley SW, et al. Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. *Blood*. 2010;115:1815–1822.

43. Bhatnagar P, Purvis S, Barron-Casella E, et al. Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. *J Hum Genet*. 2011;56:316–323.

44. Bae HT, Baldwin CT, Sebastiani P, et al. Meta-analysis of 2040 sickle cell anemia patients: bCL11A and HBS1L-MYB are the major modifiers of HbF in African Americans. *Blood*. 2012;120:1961–1962.

45. Mtatiro SN, Singh T, Rooks H, et al. Genome wide association study of fetal hemoglobin in sickle cell anemia in Tanzania. *PLoS ONE*. 2014;9:e.111464.

46. Xu J, Sankaran VG, Ni M, et al. Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev*. 2010;24:783–798.

47. Sankaran VG, Menne TF, Šćepanović D, et al. MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. *Proc Natl Acad Sci U S A*. 2011;108:1519–1524.

48. Suzuki M, Yamazaki H, Mukai HY, et al. Disruption of the Hbs1l-Myb locus causes hereditary persistence of fetal hemoglobin in a mouse model. *Mol Cell Biol*. 2013;33:1687–1695.

49. Pule GD, Mowla S, Novitzky N, Wonkam A. Hydroxyurea down-regulates BCL11A, KLF-1 and MYB through miRNA-mediated actions to induce γ-globin expression: implications for new therapeutic approaches of sickle cell disease. *Clin Transl Med*. 2016;5:15

50. Bernaudin F, Verlhac S, Arnaud C, et al. Impact of early transcranial Doppler screening and intensive therapy on cerebral vasculopathy outcome in a newborn sickle cell anemia cohort. *Blood*. 2011;117:1130–1140.

51. Driscoll MC, Hurlet A, Styles L, et al. Stroke risk in siblings with sickle cell anemia. *Blood*. 2003;101:2401–2404.

52. Flanagan JM, Sheehan V, Linder H, et al. Genetic mapping and exome sequencing identify 2 mutations associated with stroke protection in pediatric patients with sickle cell anemia. *Blood*. 2013;121:3237–3245.

53. Bhatnagar P, Barron-Casella E, Bean CJ, et al. Genome-wide meta-analysis of systolic blood pressure in children with sickle cell disease. *PLoS ONE*. 2013;8:e74193.

54. Galarneau G, Coady S, Garrett ME, et al. Gene-centric association study of acute chest syndrome and painful crises in sickle cell disease patients. *Blood*. 2013;122:434–442.

55. Milton JN, Rooks H, Drasar E, et al. Genetic determinants of haemolysis in sickle cell anaemia. *Br J Haematol*. 2013;161:270–278.

56. 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:e34741.

57. Griffin PJ, Sebastiani P, Edward H, et al. The genetics of hemoglobin A2 regulation in sickle cell anemia. *Am J Hematol*. 2014;89:1019–1023.