Abstract

IMPORTANCE Sickle cell disease (SCD) is a monogenic disorder, yet clinical outcomes are influenced by additional genetic factors. Despite decades of research, the genetics of SCD remain poorly understood.

OBJECTIVE To assess all reported genetic modifiers of SCD, evaluate the design of associated studies, and provide guidelines for future analyses according to modern genetic study recommendations.

DATA SOURCES PubMed, Web of Science, and Scopus were searched through May 16, 2023, identifying 5290 publications.

STUDY SELECTION At least 2 reviewers identified 571 original, peer-reviewed English-language publications reporting genetic modifiers of human SCD phenotypes, wherein the outcome was not treatment response, and the comparison was not between SCD subtypes or including healthy controls.

DATA EXTRACTION AND SYNTHESIS Data relevant to all genetic modifiers of SCD were extracted, evaluated, and presented following STREGA and PRISMA guidelines. Weighted z score meta-analyses and pathway analyses were conducted.

MAIN OUTCOMES AND MEASURES Outcomes were aggregated into 25 categories, grouped as acute complications, chronic conditions, hematologic parameters or biomarkers, and general or mixed measures of SCD severity.

RESULTS The 571 included studies reported on 29 670 unique individuals (50% ≤ 18 years of age) from 43 countries. Of the 17 757 extracted results (4890 significant) in 1552 genes, 3675 results met the study criteria for meta-analysis: reported phenotype and genotype, association size and direction, variability measure, sample size, and statistical test. Only 173 results for 62 associations could be cross-study combined. The remaining associations could not be aggregated because they were only reported once or methods (eg, study design, reporting practice) and genotype or phenotype definitions were insufficiently harmonized. Gene variants regulating fetal hemoglobin and α-thalassemia (important markers for SCD severity) were frequently identified: 19 single-nucleotide variants in BCL11A, HBS1L-MYB, and HBG2 were significantly associated with fetal hemoglobin (absolute value of Z = 4.00 to 20.66; P = 8.63 × 10⁻⁶⁰⁰⁰ to 6.19 × 10⁻⁵), and α-thalassemia deletions were significantly associated with increased hemoglobin level and reduced risk of albuminuria, abnormal transcranial Doppler velocity, and stroke (absolute value of Z = 3.43 to 5.16; P = 2.42 × 10⁻⁷⁰⁰⁰ to 6.00 × 10⁻⁷⁰⁰⁰). However, other associations remain unconfirmed. Pathway analyses of significant genes highlighted the importance of cellular adhesion, inflammation, oxidative and (continued)
Abstract (continued)
toxic stress, and blood vessel regulation in SCD (23 of the top 25 Gene Ontology pathways involve these processes) and suggested future research areas.

CONCLUSIONS AND RELEVANCE The findings of this comprehensive systematic review and meta-analysis of all published genetic modifiers of SCD indicated that implementation of standardized phenotypes, statistical methods, and reporting practices should accelerate discovery and validation of genetic modifiers and development of clinically actionable genetic profiles.


Introduction

Sickle cell disease (SCD) is the most common monogenic disorder in the world due to the protection that heterozygosity affords against malaria. Although SCD most heavily impacts sub-Saharan Africa, population migration and relocation have resulted in 1 in 2000 infants born in the United States with SCD, and 1 in 67 infants will be heterozygous carriers. Demographic trends and widespread improvements in clinical care will result in an increase in the proportion of the world's population affected by SCD. An improved understanding of the pathophysiology of SCD and the environmental and genetic drivers of disease severity is essential to improve the lives of individuals with this disease.

Most cases of SCD are caused by a homozygous variation in the HBB gene (p.Glu6Val) encoding the β-globin subunit of adult hemoglobin tetramer (α2β2). At low oxygen concentrations in venous capillaries, sickle hemoglobin (α2βS2) forms rigid polymers, causing circulating red blood cells to become stiff, sticky, and brittle, triggering a complex pathophysiology including hemolysis, vascular occlusion, and inflammation. Clinical manifestations include severe acute and chronic pain, immunodeficiency, multiorgan damage, and early mortality. Hemolysis-related cellular injury, partly mediated by circulating free heme released from red blood cells, is thought to drive progression of cerebrovascular disease, kidney disease, pulmonary hypertension, priapism, and leg ulcers, whereas vaso-occlusion is thought to precipitate acute pain episodes, acute chest syndrome, and avascular necrosis.

Despite being a monogenic disorder, the symptoms of SCD vary between affected individuals. The influence of environment on SCD is illustrated by markedly different outcomes between sub-Saharan Africa, where approximately half of affected children die before 5 years of age, and high-income countries, where enhanced medical support extends patient lifespan, although most patients still suffer considerably and die prematurely.

Manifestations of SCD are also influenced by genetic factors. For example, residual expression of fetal hemoglobin (Hbf, α2γ2) in postnatal red blood cells, which reduces SCD severity by interfering with polymerization of sickle hemoglobin, is largely determined genetically. Coinherited hereditary persistence of fetal hemoglobin, caused by variants in the extended β-like globin locus, results in extremely high levels of Hbf, eliminating many symptoms of SCD. Genome-wide association studies have shown that 20%-50% of the variation in Hbf can be explained by single-nucleotide variants (SNVs) in 3 loci: BCL11A, HBS1L-MYB, and the extended β-like globin locus. This discovery led to gene therapy strategies aimed at reducing erythroid BCL11A expression, some of which are showing early signs of efficacy in clinical trials, illustrating how understanding the genetic modifiers of SCD can have profound therapeutic implications.

The genetic contributions to SCD-related complications are poorly defined, despite a preponderance of publications on this topic. As a motivating example, a recent polygenic score incorporating 21 SNVs in 9 genetic loci, including HbF modifiers, explained only 3.5% of the variation in acute pain episodes. A more complete understanding of how genetics influences

pathophysiology could improve therapy by providing tools to predict outcomes and identifying new modes for therapeutic intervention. To assess current knowledge, we performed a systematic review of, to our knowledge, all publications reporting genetic modifiers of SCD, cataloged the findings by subdividing genotype-phenotype associations by quality of data analysis and reporting, and performed meta-analyses and pathway analyses. Based on our findings and current guidelines in the field of human genetics, we provide recommendations for analytical approaches and reporting to enhance scientific rigor, reduce spurious results, and facilitate cross-study data synthesis.

Methods

Article Search and Abstract Screening
This systematic review was prospectively registered with PROSPERO (No. CRD42021274466) and was reported following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline and the Strengthening the Reporting of Genetic Association Studies (STREGA) reporting guideline. We searched PubMed, Web of Science, and Scopus for all studies reporting genetic modifiers of SCD, irrespective of SCD subtype, published before May 16, 2023 (search terms in eMethods in Supplement 1). A total of 8892 studies were identified (eFigure 1 in Supplement 1): 3132 from PubMed, 2443 from Web of Science, and 3317 from Scopus. After deduplication, 5290 unique manuscripts remained.

Abstracts were screened by 2 independent reviewers (J.K.K., S.R.R.), with a third reviewer (J.H.E.) blinded to other screening opinions resolving disagreements. To comprehensively aggregate all published mutations associated with SCD-related outcomes and to avoid excluding important genetic modifiers due to incorrect phenotypic or genotypic attribution, we included all reported phenotypes and genetic polymorphisms. Studies were excluded if the manuscript was unavailable in English, the research was not conducted in humans, individuals without SCD were included, the only analysis was nongenetic, the only comparison was between SCD subtypes, the outcome was treatment response only, the manuscript was not peer reviewed, or no original research was included. A total of 571 publications passed this screening (eFigure 1 in Supplement 1, eTable 1 in Supplement 2). Data from these studies were extracted (eMethods, eTable 2, and eTable 3 in Supplement 2). Following extraction, we standardized gene annotation and phenotype categories to facilitate results tabulation (eMethods, eTable 4 in Supplement 1). No individual-level participant data were used.

Risk of Bias
Evolving approaches to genetic studies and clinical care over time, combined with variability in study design, phenotype definitions, and reporting practices, resulted in highly heterogeneous data, even within a single publication. Rather than determining the risk of bias for each publication, we created 3 categories into which all results were assigned using the STREGA guidelines584 (eMethods in Supplement 1). Briefly, exploratory results were evaluated statistically but lacked information required for cross-study meta-analysis. Meta-suitable results contained the minimum information to allow for meta-analysis: clearly defined outcome and genetic variants, sample size, statistical test, association size, direction, and measure of variability. Contemporary results contained all requirements for the meta-suitable category plus further elements crucial for genetic association studies (ie, quality control checks, accounting for population stratification and relatedness, covariate adjustment, and external validation). We did not exclude results based on these categories; however, some sections only used meta-suitable and contemporary results (Figure 1; eMethods in Supplement 1).

Meta-Analysis
We conducted meta-analyses using a weighted z score-based approach (eMethods in Supplement 1) on all SNV-phenotype pairs with meta-suitable and contemporary results from at least 2 cohorts.
reported in at least 2 manuscripts in which the phenotype was the same, the same genotype comparison was done, and the same statistical test was performed. When a manuscript reported multiple results for the same cohort, we selected the one most similar to the other results being used in terms of adjustment for other covariates. If multiple studies reported results for the same cohort, we selected the result using the largest sample size.

**Beyond Meta-Analysis**

Our meta-analyses included only associations in which phenotypes and genotypes were defined consistently across studies. While statistically rigorous, this approach omitted biologically relevant associations established through repeated linking of loci to related phenotypes. As variability in study design and reporting prevented meta-analysis of most results, including well-established modifiers, we performed further data interrogations, as described here and in the eMethods in Supplement 1.

We examined genes with variants significantly associated with HbF in at least 3 manuscripts because HbF is a well-known disease modifier. Similarly, co-inherited deletional α-thalassemia is common in SCD populations and modifies SCD pathophysiology, although the association with SCD varies across phenotypes and studies. To illustrate this comprehensively, we compared all meta-suitable and contemporary associations of SCD phenotypes with α-thalassemia deletions (eMethods in Supplement 1). Finally, we conducted pathway analyses to align significant findings with biological functions from the curated Gene Ontology (GO) and Reactome databases (eMethods).

---

**GWSS indicates genome-wide summary statistics; HbF, fetal hemoglobin.**
in Supplement 1). This approach is used to analyze lists of important genes to facilitate interpretation and hypothesis generation.

**Statistical Analysis**

Unless stated otherwise, analyses were conducted with R, version 4.2.1 (CRAN). For the meta-analyses, a Bonferroni-corrected 2-sided $P < 8.1 \times 10^{-4}$ ($0.05/62$) was considered statistically significant. For pathway analyses, an adjusted $P < .05$ was considered significant.

**Results**

We identified 571 manuscripts published before May 16, 2023, reporting genotype-phenotype associations across 29,670 unique individuals (50% ≤ 18 years of age) from 43 countries (eMethods, eFigures 1 and 2, and eTable 5 in Supplement 1; eTable 1 in Supplement 2). Approximately 52% of individuals resided in USA, Canada, France, or Brazil, while just 6628 individuals studied (22%) were from African cohorts (eFigure 2 and eTable 5 in Supplement 1). Fifty-five manuscripts (10%) assessed individuals from more than 1 country. At least 14,970 individuals were included in studies as children (eTable 1 in Supplement 2 and eTable 5 in Supplement 1), but some children were also included in studies as adults in longitudinal cohorts.

**Risk of Bias**

Of 17,757 reported associations, 3631 (20%) were meta-suitable, containing all elements required to perform a meta-analysis, and only 44 results from 2 studies met all contemporary criteria (eTable 2 in Supplement 2). The proportion of studies using a more rudimentary exploratory approach to a more rigorous meta-suitable approach did not appear to change over time (eFigure 3 in Supplement 1). For the rest of our analysis, we grouped the contemporary results with the meta-suitable results (3675 total results).

**Genotype-Phenotype Associations**

Across the 571 publications analyzed, 17,757 association results in 1552 unique genes were reported, along with 249,078 genome-wide association summary statistics (Figure 1, Table 1, and eTables 2 and 3 in Supplement 2). The number of genes interrogated varied by phenotype category (Table 1), ranging from 9 (retinopathy) to 452 (HbF). Of the 2,399 unique gene- or polygene-phenotype category pairs (eTable 6 in Supplement 2), there was a median (IQR) of 2 (1-4) results, with 1976 (82%) limited to a single study. Overall, 4,890 (28%) extracted results were statistically significant (eMethods in Supplement 1), but this does not account for varying significance thresholds (ie, 0.05, $5 \times 10^{-8}$, etc).

**Meta-Analysis**

While 3675 of the 17,757 total results (21%) were categorized as meta-suitable, due to differences in specific phenotypes, genotypes, or statistical methods, studies analyzing similar outcomes were often insufficiently harmonized for cross-comparison or were not replicated (Figure 1; eMethods in Supplement 1). Only 173 of 17,757 results (1%) plus 2 replication results and 18 genome-wide association summary statistics (Figure 1, Table 2). Of these 193 results, 111 (58%) matched direction and significance with the meta-analysis results; of the remaining 82, 54 (66%) were in directional agreement but differed in statistical significance.

Meta-analysis results indicated that α-thalassemia deletions were significantly associated with increased hemoglobin level ($Z = 5.12; P = 3.11 \times 10^{-7}$) and reduced risk of albuminuria ($Z = -3.54; P = 4.10 \times 10^{-4}$), abnormal transcranial Doppler velocity ($Z = -5.16; P = 2.42 \times 10^{-7}$), and stroke ($Z = -5.12; P = 2.97 \times 10^{-7}$ for occurrence; $Z = -3.43; P = 6.00 \times 10^{-4}$ for time to event) (Table 2). Ten SNVs in **BCL11A**, 8 in **HBS1L-MYB**, and 1 in the γ-globin gene (rs7482144, the XmnI site of **HBG2** were...
significantly associated with HbF \((\text{absolute value of } Z = 4.00 \text{ to } 20.66; P = 8.63 \times 10^{-95} \text{ to } 6.19 \times 10^{-5})\). An increased number of \(UGT1A1\) promoter repeats was associated with increased risk for cholelithiasis \((Z = 5.09, P = 3.57 \times 10^{-7} \text{ for } (TA)7/(TA)7; Z = 4.27, P = 1.93 \times 10^{-5} \text{ for } (TA)7/(TA)8\); compared with \((TA)6/(TA)6\). Single-nucleotide variants in \(RPS24\) \((\text{rs7899453-A, } Z = 4.70; P = 2.61 \times 10^{-6})\) and \(TBC1D1\) \((\text{rs6858735-T, } Z = 4.57; P = 4.81 \times 10^{-6})\) were significantly associated with increased rate of vaso-occlusive crisis.

While high-risk \(G1/G2\) \(APOL1\) variants were frequently associated with numerous markers for kidney dysfunction (eTables 2 and 6 in Supplement 2), only 2 results could be meta-analyzed, showing nominal association with increased risk for albuminuria \((Z = 2.83, P = .005)\) (Table 2). Similarly, 1SNV in \(COMT\) \((\text{rs4680})\) was nominally associated with vaso-occlusive crisis in our meta-analyses \((Z = 3.30, P = 9.60 \times 10^{-4})\) (Table 2), but numerous SNVs and haplotypes within this gene have been associated with acute pain outcomes (eTables 2 and 6 in Supplement 2).

**HbF**

The genetics of HbF expression was more widely studied than any other SCD phenotype, with 240 of 571 studies (42%) reporting a total of 3952 associations involving 452 genes (Table 1; eTables 2 and 6 in Supplement 2). Significant associations with HbF were reported for 140 genes in 144 studies, yet 1220 of 1347 (91%) of these associations were exploratory (eFigure 4 in Supplement 1 and eTables 2 and 6 in Supplement 2). Most significant results identified the extended \(\beta\)-like globin locus.

<table>
<thead>
<tr>
<th>Table 1. Number Of Studies, Total Results, and Unique Genes Reported for Each Phenotype Category</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phenotype category</strong></td>
</tr>
<tr>
<td>Acute SCD-related complications</td>
</tr>
<tr>
<td>Acute pain episode</td>
</tr>
<tr>
<td>ACS, pneumonia, or respiratory infection</td>
</tr>
<tr>
<td>Infection (excludes respiratory infection)</td>
</tr>
<tr>
<td>Priapism</td>
</tr>
<tr>
<td>Acute splenic sequestration</td>
</tr>
<tr>
<td>Other acute phenotype</td>
</tr>
<tr>
<td>Chronic SCD-related complications</td>
</tr>
<tr>
<td>Allo- or autoantibody or transfusion reaction</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
</tr>
<tr>
<td>Kidney dysfunction</td>
</tr>
<tr>
<td>Cardiopulmonary dysfunction</td>
</tr>
<tr>
<td>Hyperbilirubinemia, cholelithiasis, cholecystitis, or cholecystectomy</td>
</tr>
<tr>
<td>Osteonecrosis</td>
</tr>
<tr>
<td>Leg ulcers</td>
</tr>
<tr>
<td>Iron overload</td>
</tr>
<tr>
<td>Chronic pain</td>
</tr>
<tr>
<td>Splenic dysfunction</td>
</tr>
<tr>
<td>Retinopathy</td>
</tr>
<tr>
<td>Other chronic phenotype</td>
</tr>
<tr>
<td>Hematologic parameters and biomarkers of disease severity</td>
</tr>
<tr>
<td>HbF</td>
</tr>
<tr>
<td>Hemolysis</td>
</tr>
<tr>
<td>Anemia</td>
</tr>
<tr>
<td>Oxidative stress</td>
</tr>
<tr>
<td>Other hematologic parameter</td>
</tr>
<tr>
<td>Other parameter or biomarker</td>
</tr>
<tr>
<td>General or mixed measurement of SCD severity</td>
</tr>
</tbody>
</table>

Abbreviations: ACS, acute chest syndrome; HbF, fetal hemoglobin; NA, not applicable; SCD, sickle cell disease.

* Within each subset, phenotype categories are ordered by decreasing number of total unique genes, excepting “other” categories, which are listed last.

* Complication prevalence rates were obtained from published estimates among adults of all SCD subtypes within the United States, when available (eMethods in Supplement 1).

* Excludes 249 078 genome-wide summary statistics from 2 publications to avoid count distortion.

* Among male participants.
<table>
<thead>
<tr>
<th>Phenotype category</th>
<th>Specific outcome</th>
<th>Variant</th>
<th>Gene</th>
<th>EA</th>
<th>OA</th>
<th>No. of studies</th>
<th>Total sample size, No.</th>
<th>z Score</th>
<th>P value</th>
<th>Direction</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS, pneumonia, or respiratory infections</td>
<td>ACS (event occurrence)</td>
<td>rs2070744</td>
<td>NOS3</td>
<td>CC</td>
<td>TT or CT</td>
<td>2</td>
<td>273</td>
<td>1.4</td>
<td>.16</td>
<td>+</td>
<td>SN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>3</td>
<td>435</td>
<td>2.18</td>
<td>.03</td>
<td>+++</td>
<td>NNN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event occurrence)</td>
<td>rs5030737, rs1800450, rs1800451</td>
<td>MBL2</td>
<td>AA</td>
<td>AO or OD</td>
<td>2</td>
<td>155</td>
<td>−0.62</td>
<td>.54</td>
<td>+</td>
<td>NS</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1042713</td>
<td>ADRB2</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>463</td>
<td>−1.16</td>
<td>.25</td>
<td>--</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1042713</td>
<td>ADRB2</td>
<td>AA or AG</td>
<td>GG</td>
<td>2</td>
<td>463</td>
<td>−2.55</td>
<td>.01</td>
<td>--</td>
<td>SN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs10483639</td>
<td>MBL2</td>
<td>AA AO or OO</td>
<td>GG</td>
<td>2</td>
<td>155</td>
<td>−0.62</td>
<td>.54</td>
<td>+</td>
<td>NS</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs10483639</td>
<td>MBL2</td>
<td>AA or AG</td>
<td>GG</td>
<td>2</td>
<td>442</td>
<td>2.92</td>
<td>.01</td>
<td>+</td>
<td>SN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs10483639</td>
<td>MBL2</td>
<td>CC or CG</td>
<td>GG</td>
<td>2</td>
<td>458</td>
<td>1.61</td>
<td>.11</td>
<td>+</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1800587</td>
<td>IL1A</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>442</td>
<td>2.19</td>
<td>.03</td>
<td>+</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1800587</td>
<td>IL1A</td>
<td>AA or AG</td>
<td>GG</td>
<td>2</td>
<td>442</td>
<td>2.19</td>
<td>.03</td>
<td>+</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1947913</td>
<td>TRPA1</td>
<td>A</td>
<td>T</td>
<td>2</td>
<td>459</td>
<td>−1.67</td>
<td>.10</td>
<td>--</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs1947913</td>
<td>TRPA1</td>
<td>AA or AT</td>
<td>TT</td>
<td>2</td>
<td>459</td>
<td>0.017</td>
<td>.99</td>
<td>--</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs2963155</td>
<td>NOS3</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>463</td>
<td>0.87</td>
<td>.38</td>
<td>++</td>
<td>SS</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs2963155</td>
<td>NOS3</td>
<td>AA or AG</td>
<td>GG or GA</td>
<td>2</td>
<td>463</td>
<td>0.71</td>
<td>.48</td>
<td>+</td>
<td>SS</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs4680</td>
<td>COMT</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>457</td>
<td>3.3</td>
<td>9.60 × 10^-4</td>
<td>+</td>
<td>NN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs6858735</td>
<td>TBC1D1</td>
<td>T</td>
<td>C</td>
<td>3</td>
<td>2228</td>
<td>4.57</td>
<td>4.81 × 10^-6</td>
<td>++</td>
<td>SNN</td>
</tr>
<tr>
<td>Acute pain episode</td>
<td>VOC (event rate)</td>
<td>rs7899453</td>
<td>TBC1D1</td>
<td>A</td>
<td>C</td>
<td>3</td>
<td>2228</td>
<td>4.7</td>
<td>2.61 × 10^-6</td>
<td>+++</td>
<td>NNN</td>
</tr>
<tr>
<td>Acute splenic sequestration</td>
<td>Acute splenic sequestration (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>239</td>
<td>1.59</td>
<td>.11</td>
<td>--</td>
<td>NS</td>
</tr>
<tr>
<td>Anemia</td>
<td>Blood transfusion (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>225</td>
<td>−1.81</td>
<td>.07</td>
<td>--</td>
<td>NN</td>
</tr>
<tr>
<td>Anemia</td>
<td>Hemoglobin (continuous)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>1193</td>
<td>5.12</td>
<td>3.11 × 10^-7</td>
<td>++</td>
<td>SS</td>
</tr>
<tr>
<td>Anemia</td>
<td>Hemoglobin (continuous)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>453</td>
<td>1.3</td>
<td>2.69 × 10^-4</td>
<td>++</td>
<td>SS</td>
</tr>
<tr>
<td>Anemia</td>
<td>Hemoglobin (continuous)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>1193</td>
<td>5.12</td>
<td>3.11 × 10^-7</td>
<td>++</td>
<td>SS</td>
</tr>
<tr>
<td>Anemia</td>
<td>Hemoglobin (continuous)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>453</td>
<td>1.3</td>
<td>2.69 × 10^-4</td>
<td>++</td>
<td>SS</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>Abnormal TCD result (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>366</td>
<td>−5.16</td>
<td>2.42 × 10^-7</td>
<td>--</td>
<td>SS</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>Stroke (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>6</td>
<td>3655</td>
<td>−5.12</td>
<td>2.97 × 10^-7</td>
<td>--</td>
<td>SS</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>Stroke (event occurrence)</td>
<td>GT repeats</td>
<td>AGT</td>
<td>A3 and/or A4</td>
<td>Other</td>
<td>2</td>
<td>219</td>
<td>−0.35</td>
<td>0.73</td>
<td>--</td>
<td>NS</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>Stroke (event occurrence)</td>
<td>rs1800629</td>
<td>TNF</td>
<td>AA or AG</td>
<td>GG</td>
<td>3</td>
<td>599</td>
<td>−1.83</td>
<td>.07</td>
<td>--</td>
<td>NS</td>
</tr>
<tr>
<td>Cholelithiasis, cholecystitis, or cholecystectomy</td>
<td>Hyperbilirubinemia, cholelithiasis, cholecystitis, or cholecystectomy</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>242</td>
<td>−2.48</td>
<td>.01</td>
<td>--</td>
<td>NS</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Phenotype category, Specific outcome</th>
<th>Variant</th>
<th>Gene</th>
<th>EA</th>
<th>OA</th>
<th>No. of studies*</th>
<th>Total sample size, No.</th>
<th>z Score</th>
<th>P value</th>
<th>Direction*</th>
<th>Significance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperbilirubinemia, cholelithiasis, cholecystitis, or cholecystectomy</td>
<td>TA repeats</td>
<td>UGT1A locus</td>
<td>(6/6)</td>
<td>(7/7)</td>
<td>4</td>
<td>821</td>
<td>−5.09</td>
<td>3.57 × 10^{-7}</td>
<td>−−−−</td>
<td>NNSS</td>
</tr>
<tr>
<td>Hyperbilirubinemia, cholelithiasis, cholecystitis, or cholecystectomy</td>
<td>TA repeats</td>
<td>UGT1A locus</td>
<td>(6/6)</td>
<td>(6/7)</td>
<td>3</td>
<td>719</td>
<td>−1.92</td>
<td>.06</td>
<td>−−</td>
<td>NNN</td>
</tr>
<tr>
<td>No. of studiesb</td>
<td>Total samplesize, No.</td>
<td>z Score</td>
<td>P value</td>
<td>Direction*</td>
<td>Significance*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperbilirubinemia, cholelithiasis, cholecystitis, or cholecystectomy</td>
<td>TA repeats</td>
<td>UGT1A locus</td>
<td>(6/6)</td>
<td>(7/8)</td>
<td>3</td>
<td>668</td>
<td>−4.27</td>
<td>1.93 × 10^{-5}</td>
<td>−−−−</td>
<td>NNS</td>
</tr>
<tr>
<td>General or mixed measurement of SCD severity</td>
<td>α-Thalassemia α-Globin 1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>225</td>
<td>−0.9</td>
<td>.37</td>
<td>−−</td>
<td>NN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbF17,115,122,163,207,269,270,312</td>
<td>HbF (continuous)</td>
<td>rs11886888</td>
<td>BCL11A</td>
<td>T</td>
<td>C</td>
<td>7</td>
<td>2339</td>
<td>−15.3</td>
<td>7.08 × 10^{-53}</td>
<td>−−−−−−−</td>
</tr>
<tr>
<td>HbF17,115,132,163,202,234,269,270</td>
<td>HbF (continuous)</td>
<td>rs1427407</td>
<td>BCL11A</td>
<td>T</td>
<td>G</td>
<td>10</td>
<td>3394</td>
<td>20.66</td>
<td>8.63 × 10^{-95}</td>
<td>++++++++</td>
</tr>
<tr>
<td>HbF17,228,312</td>
<td>HbF (continuous)</td>
<td>rs28384513</td>
<td>HBS1L-MYB</td>
<td>A</td>
<td>C</td>
<td>4</td>
<td>1947</td>
<td>5.26</td>
<td>1.43 × 10^{-7}</td>
<td>+−</td>
</tr>
<tr>
<td>HbF17,214</td>
<td>HbF (continuous)</td>
<td>rs35786788</td>
<td>HBS1L-MYB</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>1606</td>
<td>7.11</td>
<td>1.20 × 10^{-12}</td>
<td>+</td>
</tr>
<tr>
<td>HbF17,37,115,163,269,270,312,320</td>
<td>HbF (continuous)</td>
<td>rs4671393</td>
<td>BCL11A</td>
<td>A</td>
<td>G</td>
<td>7</td>
<td>2256</td>
<td>16.25</td>
<td>2.19 × 10^{-59}</td>
<td>++++++++</td>
</tr>
<tr>
<td>HbF17,57,210</td>
<td>HbF (continuous)</td>
<td>rs4895441</td>
<td>BCL11A</td>
<td>A</td>
<td>G</td>
<td>6</td>
<td>2221</td>
<td>−7.55</td>
<td>4.24 × 10^{-14}</td>
<td>−−−−−−−−−</td>
</tr>
<tr>
<td>HbF17,115,122,210</td>
<td>HbF (continuous)</td>
<td>rs6545816</td>
<td>BCL11A</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>841</td>
<td>−4</td>
<td>6.19 × 10^{-5}</td>
<td>+−</td>
</tr>
<tr>
<td>HbF17,115,132,210</td>
<td>HbF (continuous)</td>
<td>rs66650371</td>
<td>HBS1L-MYB</td>
<td>D</td>
<td>I</td>
<td>4</td>
<td>2447</td>
<td>9.52</td>
<td>1.82 × 10^{-21}</td>
<td>++++</td>
</tr>
<tr>
<td>HbF17,214</td>
<td>HbF (continuous)</td>
<td>rs6706648</td>
<td>BCL11A</td>
<td>T</td>
<td>C</td>
<td>5</td>
<td>1728</td>
<td>−12.57</td>
<td>3.01 × 10^{-36}</td>
<td>−−−−−−−−</td>
</tr>
<tr>
<td>HbF17,269,270</td>
<td>HbF (continuous)</td>
<td>rs6729815</td>
<td>BCL11A</td>
<td>T</td>
<td>C</td>
<td>2</td>
<td>198</td>
<td>−0.12</td>
<td>.91</td>
<td>−</td>
</tr>
<tr>
<td>HbF17,269,270</td>
<td>HbF (continuous)</td>
<td>rs6732518</td>
<td>BCL11A</td>
<td>T</td>
<td>C</td>
<td>3</td>
<td>782</td>
<td>−4.55</td>
<td>5.27 × 10^{-6}</td>
<td>+−</td>
</tr>
<tr>
<td>HbF17,115,122,210</td>
<td>HbF (continuous)</td>
<td>rs6738440</td>
<td>BCL11A</td>
<td>A</td>
<td>G</td>
<td>5</td>
<td>1728</td>
<td>9.7</td>
<td>3.10 × 10^{-22}</td>
<td>++++</td>
</tr>
<tr>
<td>HbF17,214</td>
<td>HbF (continuous)</td>
<td>rs73555746</td>
<td>HBS1L-MYB</td>
<td>A</td>
<td>C</td>
<td>2</td>
<td>198</td>
<td>0.39</td>
<td>.69</td>
<td>+−</td>
</tr>
<tr>
<td>HbF17,37,115,163,269,270,312,320</td>
<td>HbF (continuous)</td>
<td>rs7482144</td>
<td>Extended β-globin locus</td>
<td>A</td>
<td>G</td>
<td>5</td>
<td>2637</td>
<td>7.14</td>
<td>9.27 × 10^{-13}</td>
<td>++++</td>
</tr>
<tr>
<td>HbF17,57,320</td>
<td>HbF (continuous)</td>
<td>rs7557939</td>
<td>BCL11A</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>834</td>
<td>−9.19</td>
<td>3.95 × 10^{-20}</td>
<td>−</td>
</tr>
<tr>
<td>HbF17,202,234</td>
<td>HbF (continuous)</td>
<td>rs7599488</td>
<td>BCL11A</td>
<td>T</td>
<td>C</td>
<td>4</td>
<td>1298</td>
<td>0.47</td>
<td>.63</td>
<td>−−−−−</td>
</tr>
<tr>
<td>HbF17,269,270</td>
<td>HbF (continuous)</td>
<td>rs7606173</td>
<td>BCL11A</td>
<td>C</td>
<td>G</td>
<td>6</td>
<td>2354</td>
<td>−13.68</td>
<td>1.41 × 10^{-42}</td>
<td>−−−−−−−</td>
</tr>
<tr>
<td>HbF17,27,115,163,269,270,312</td>
<td>HbF (continuous)</td>
<td>rs766432</td>
<td>BCL11A</td>
<td>A</td>
<td>C</td>
<td>3</td>
<td>782</td>
<td>−10.13</td>
<td>4.04 × 10^{-24}</td>
<td>−−</td>
</tr>
<tr>
<td>HbF17,269,270</td>
<td>HbF (continuous)</td>
<td>rs7775698</td>
<td>HBS1L-MYB</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>198</td>
<td>0.53</td>
<td>.60</td>
<td>+−</td>
</tr>
<tr>
<td>HbF17,115,132,163,269,270,312</td>
<td>HbF (continuous)</td>
<td>rs9399137</td>
<td>HBS1L-MYB</td>
<td>T</td>
<td>C</td>
<td>7</td>
<td>2256</td>
<td>−7.24</td>
<td>3.33 × 10^{-15}</td>
<td>−−−−−−−−</td>
</tr>
<tr>
<td>HbF17,37,115,163,269,270,312</td>
<td>HbF (continuous)</td>
<td>rs9402686</td>
<td>HBS1L-MYB</td>
<td>A</td>
<td>G</td>
<td>7</td>
<td>2349</td>
<td>8.44</td>
<td>3.11 × 10^{-17}</td>
<td>++++++</td>
</tr>
<tr>
<td>HbF17,37,115,163,269,270,312</td>
<td>HbF (continuous)</td>
<td>rs9404142</td>
<td>HBS1L-MYB</td>
<td>T</td>
<td>C</td>
<td>3</td>
<td>1780</td>
<td>−3.51</td>
<td>1.07 × 10^{-7}</td>
<td>−−−−−</td>
</tr>
<tr>
<td>HbF17,57,210</td>
<td>HbF (continuous)</td>
<td>rs9494145</td>
<td>HBS1L-MYB</td>
<td>T</td>
<td>C</td>
<td>3</td>
<td>1856</td>
<td>−7.88</td>
<td>3.33 × 10^{-15}</td>
<td>−−−−−−−</td>
</tr>
<tr>
<td>Leg ulcers</td>
<td>Leg ulcers (event occurrence)</td>
<td>α-Thalassemia α-Globin 1 or 2 Deletions</td>
<td>No deletions</td>
<td>2</td>
<td>2336</td>
<td>−2.05</td>
<td>.04</td>
<td>−−−−−</td>
<td>NN</td>
<td></td>
</tr>
<tr>
<td>Other hematologic parameter</td>
<td>PLT (continuous)</td>
<td>rs66650371</td>
<td>HBS1L-MYB</td>
<td>D</td>
<td>I</td>
<td>2</td>
<td>986</td>
<td>−0.96</td>
<td>.34</td>
<td>−+</td>
</tr>
<tr>
<td>Other hematologic parameter</td>
<td>PLT (continuous)</td>
<td>rs7482144</td>
<td>Extended β-globin locus</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>986</td>
<td>−1.97</td>
<td>.05</td>
<td>−−−−−</td>
</tr>
</tbody>
</table>

*Table 2. Meta-Analysis Results*
### Table 2. Meta-Analysis Results (continued)

<table>
<thead>
<tr>
<th>Phenotype category</th>
<th>Specific outcome</th>
<th>Variant</th>
<th>Gene</th>
<th>EA</th>
<th>OA</th>
<th>No. of studies</th>
<th>Total sample size, No.</th>
<th>z Score</th>
<th>P value</th>
<th>Direction</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other hematologic parameter115,223</td>
<td>RBC count (continuous)</td>
<td>rs66650371</td>
<td>HBS1L-MYB</td>
<td>D</td>
<td>I</td>
<td>2</td>
<td>986</td>
<td>1.89</td>
<td>.06</td>
<td>++</td>
<td>SN</td>
</tr>
<tr>
<td>Other hematologic parameter115,223</td>
<td>RBC count (continuous)</td>
<td>rs7482144</td>
<td>Extended β-globin locus</td>
<td>A</td>
<td>G</td>
<td>2</td>
<td>986</td>
<td>1.66</td>
<td>.10</td>
<td>++</td>
<td>NN</td>
</tr>
<tr>
<td>Priapism11,118,148</td>
<td>Priapism (event occurrence)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>3</td>
<td>1745</td>
<td>−2.64</td>
<td>.008</td>
<td>−−</td>
<td>SNN</td>
</tr>
<tr>
<td>Kidney dysfunction70,320,534</td>
<td>Albuminuria (time to event)</td>
<td>α-Thalassemia</td>
<td>α-Globin</td>
<td>1 or 2 Deletions</td>
<td>No deletions</td>
<td>3</td>
<td>978</td>
<td>−3.54</td>
<td>4.10 × 10&lt;sup&gt;−4&lt;/sup&gt;</td>
<td>−−−</td>
<td>NSS</td>
</tr>
<tr>
<td>Kidney dysfunction770,541</td>
<td>Albuminuria (event occurrence)</td>
<td>G1/G2</td>
<td>APOL1</td>
<td>Homozygous G1 or G2 or compound heterozygous</td>
<td>Other</td>
<td>2</td>
<td>433</td>
<td>2.83</td>
<td>.005</td>
<td>++</td>
<td>NS</td>
</tr>
</tbody>
</table>

Abbreviations. ACS, acute chest syndrome; EA, effect allele or genotype; HbF, fetal hemoglobin; OA, other allele or genotype; PLT, platelets; RBC, red blood cell; TCD, transcranial doppler; VOC, vaso-occlusive crisis.

*Weighted z score meta-analyses were conducted for all single-nucleotide variant-phenotype pairs with meta-appropriate or contemporary results from at least 2 cohorts that were reported in at least 2 manuscripts, where the phenotype was the same, the same genotype comparison was made, and the same statistical test was performed.

*Significance for each respective association is labeled S when it was reported as significant and N otherwise.

*Other results are considered NS.
**BCL11A**, or **HBS1L-MYB**, with few repeated or meta-suitable results outside of these regions (eFigure 4 in Supplement 1). Numerous studies linked these same HbF modifier genes to specific complications of SCD, including acute pain, anemia, cerebrovascular disease, and hemolysis (eTables 2 and 6 in Supplement 2).

**α-Thalassemia**

Concurrent α-thalassemia was consistently associated with increased hemoglobin levels (eg, continuous \( \beta \), 0.39; 95% CI, 0.24-0.53) and reduced risk for elevated markers of hemolysis (eg, hemolytic component \( \beta \), −0.70; 95% CI, −1.26 to −0.14), hepatomegaly (≤4 cm vs >4 cm \( \beta \), 1.82; 95% CI, 0.51-3.31), biliary dysfunction (eg, bilirubin levels as high vs low \( \beta \), −1.32; 95% CI, −2.51 to −0.13), stroke (eg, occurrence \( \beta \), −0.85; 95% CI, −1.18 to −0.52), and kidney dysfunction (eg, albuminuria occurrence \( \beta \), −1.10; 95% CI, −1.74 to −0.47) (Figure 2). While less clear, there may be increased risk of acute pain crisis (eg, vaso-occlusive crisis events per year \( \beta \), 0.29; 95% CI, −0.16 to 0.74), acute splenic sequestration (\( \beta \), 0.72; 95% CI, −0.19 to 1.62), and osteonecrosis (\( \beta \), 1.06; 95% CI, −0.22 to 2.36) and reduced risk of leg ulcers (\( \beta \), −0.30; 95% CI, −0.59 to −0.01) and priapism (\( \beta \), −0.42; 95% CI, −0.75 to −0.08).

**Pathway Analysis**

Consistent with the known pathophysiology of SCD, genes with at least 1 significant result for any outcome were enriched for cellular adhesion, oxidative and toxic stress, inflammation, and blood vessel regulation, with 23 of the 25 most enriched GO pathways representing these processes (Figure 3; eTables 7 and 8 in Supplement 2). Phenotype category-specific pathway analyses (eTables 7 and 8 in Supplement 2) identified numerous other pathway enrichments, including the flavonoid metabolic GO pathway, for which genes associated with hyperbilirubinemia or biliary dysfunction, hemolysis, and anemia were enriched.

**Discussion**

This comprehensive systematic review and meta-analysis consolidates current knowledge of SCD genetic modifiers, incorporating data from 571 publications from 1981 to 2023 and describing at least 29 670 unique individuals residing in 43 countries. These 571 studies were assessed for quality of study design and reporting, according to STREGA guidelines. Remarkably, only 1% of results reported across the last 43 years of work met minimum standards for cross-study meta-analysis due to variability in study designs, reporting practices, and phenotype or genotype definitions.

The differing methodologies and reporting practices between screened studies limited our ability to include, aggregate, or analyze published data. Many manuscripts used individuals without SCD as controls and were excluded, as this does not assess the genetic modifiers of SCD severity. Similarly, studies using biochemical measures as surrogates for genotypes (ie, glucose-6-phosphate dehydrogenase levels) were excluded because biochemical measurements do not always align with genotype. Other limitations included use of rudimentary statistical tests, rather than regression-based techniques adjusting for confounders, and inconsistent phenotype or genotype definitions. Despite these methodological challenges, we resolved some previous discrepant results and provide guidelines for future studies.

Across 571 manuscripts, we, as in a recent review, identified several genes associated with SCD complications in at least 2 studies. However, among those validated in meta-analysis, most were related to HbF levels. Specifically, meta-analyses confirmed the polygenic regulation of HbF: 10 SNVs in **BCL11A**, 8 in **HBS1L-MYB**, and 1 **HGB2** were significantly associated with HbF. Remarkably, 137 additional genes were reportedly associated with HbF but have not been confirmed due to lack of validation in a separate cohort or insufficient harmonization of phenotype (ie, dichotomous vs continuous, F cell percentage vs HbF percentage, unclear units) or genotype. Only 20%-50% of the genetic variability in HbF can be explained by currently validated variants (extended β-like globin locus, **BCL11A**, and **HBS1L-MYB**), which have relatively large association sizes. Most likely, many

---

**JAMA Network Open | Genetics and Genomics**

Genetic Variation and Sickle Cell Disease Severity

---

Downloaded From: https://jamanetwork.com/ on 10/23/2023
### Figure 2. Published α-Thalassemia Associations Across All Meta-Suitable Results

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Total sample size, No.</th>
<th>Negative association</th>
<th>Positive association</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS, occurrence</td>
<td>83</td>
<td>0.18 (-1.02 to 1.04)</td>
<td></td>
</tr>
<tr>
<td>Pneumonia, occurrence</td>
<td>210</td>
<td>0.44 (-0.30 to 1.17)</td>
<td></td>
</tr>
<tr>
<td>Respiratory infections, occurrence</td>
<td>210</td>
<td>-0.20 (-1.11 to 0.72)</td>
<td></td>
</tr>
<tr>
<td>VOC, No. events≥3</td>
<td>327</td>
<td>0.29 (-0.16 to 0.74)</td>
<td></td>
</tr>
<tr>
<td>VOC, ≥3 episodes vs none</td>
<td>234</td>
<td>-0.60 (-1.17 to -0.04)</td>
<td></td>
</tr>
<tr>
<td>VOC, occurrence</td>
<td>435</td>
<td>0.65 (-0.03 to 1.33)</td>
<td></td>
</tr>
<tr>
<td>Acute splenic sequestration, occurrence</td>
<td>239</td>
<td>0.72 (-0.19 to 1.62)</td>
<td></td>
</tr>
<tr>
<td>Blood transfusion, occurrence</td>
<td>225</td>
<td>-0.60 (-1.28 to 0.07)</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin, low vs high</td>
<td>580</td>
<td>-1.36 (-1.97 to -0.74)</td>
<td></td>
</tr>
<tr>
<td>Mean overnight SpO2</td>
<td>30</td>
<td>-2.63 (-4.65 to -0.60)</td>
<td></td>
</tr>
<tr>
<td>Pulmonary hypertension, occurrence</td>
<td>188</td>
<td>-0.06 (-0.78 to 0.65)</td>
<td></td>
</tr>
<tr>
<td>Abnormal TCD, occurrence</td>
<td>366</td>
<td>-1.20 (-1.66 to -0.74)</td>
<td></td>
</tr>
<tr>
<td>Intracranial stenosis, time to event</td>
<td>375</td>
<td>-0.83 (-1.66 to 0.002)</td>
<td></td>
</tr>
<tr>
<td>Stroke, time to event</td>
<td>3655</td>
<td>-0.85 (-1.18 to -0.52)</td>
<td></td>
</tr>
<tr>
<td>Vasculopathy, occurrence</td>
<td>191</td>
<td>-0.51 (-1.56 to 0.57)</td>
<td></td>
</tr>
<tr>
<td>Hospitalization, occurrence</td>
<td>225</td>
<td>-0.34 (-1.12 to 0.43)</td>
<td></td>
</tr>
<tr>
<td>HbF</td>
<td>105</td>
<td>-0.21 (-0.47 to 0.05)</td>
<td></td>
</tr>
<tr>
<td>Hemolytic component, continuous score</td>
<td>150</td>
<td>-0.70 (-1.26 to -0.14)</td>
<td></td>
</tr>
<tr>
<td>Red cell survival</td>
<td>62</td>
<td>0.31 (0.12 to 0.49)</td>
<td></td>
</tr>
<tr>
<td>Reticulocyte count, high vs low</td>
<td>95</td>
<td>-3.54 (-9.21 to 1.59)</td>
<td></td>
</tr>
<tr>
<td>Bilirubin, high vs low</td>
<td>93</td>
<td>-1.32 (-2.51 to -0.13)</td>
<td></td>
</tr>
<tr>
<td>Cholecystectomy, occurrence</td>
<td>83</td>
<td>-1.97 (-3.51 to -0.51)</td>
<td></td>
</tr>
<tr>
<td>Cholecystitis, occurrence</td>
<td>83</td>
<td>-2.66 (-4.63 to -0.62)</td>
<td></td>
</tr>
<tr>
<td>Cholelithiasis, occurrence</td>
<td>242</td>
<td>-0.87 (-1.54 to -0.19)</td>
<td></td>
</tr>
<tr>
<td>Cholelithiasis, time to event</td>
<td>158</td>
<td>-0.90 (-1.61 to -0.20)</td>
<td></td>
</tr>
<tr>
<td>Bacteremia, occurrence</td>
<td>1473</td>
<td>-0.03 (-0.14 to 0.10)</td>
<td></td>
</tr>
<tr>
<td>Fetal infection, occurrence</td>
<td>83</td>
<td>0.54 (0.66 to 1.75)</td>
<td></td>
</tr>
<tr>
<td>Urinary infection, occurrence</td>
<td>210</td>
<td>-0.36 (-2.53 to 1.82)</td>
<td></td>
</tr>
<tr>
<td>Viral infection, occurrence</td>
<td>83</td>
<td>-0.02 (-1.47 to 1.39)</td>
<td></td>
</tr>
<tr>
<td>Leg ulcers, occurrence</td>
<td>2336</td>
<td>-0.30 (-0.59 to -0.01)</td>
<td></td>
</tr>
<tr>
<td>Osteonecrosis, occurrence</td>
<td>83</td>
<td>1.06 (-0.22 to 2.36)</td>
<td></td>
</tr>
<tr>
<td>Hepatomegaly, ≥4 cm vs &gt;4 cm</td>
<td>98</td>
<td>1.82 (0.51 to 3.11)</td>
<td></td>
</tr>
<tr>
<td>Hematocrit</td>
<td>1131</td>
<td>1.57 (1.09 to 2.04)</td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td>918</td>
<td>-2.69 (-3.04 to -2.35)</td>
<td></td>
</tr>
<tr>
<td>MCV</td>
<td>1034</td>
<td>-7.54 (-8.51 to -6.58)</td>
<td></td>
</tr>
<tr>
<td>RBC count</td>
<td>1034</td>
<td>0.42 (0.35 to 0.49)</td>
<td></td>
</tr>
<tr>
<td>Priapism, occurrence</td>
<td>1745</td>
<td>-0.42 (-0.75 to -0.08)</td>
<td></td>
</tr>
<tr>
<td>UACR&lt;70</td>
<td>205</td>
<td>-1.00 (-1.50 to -0.60)</td>
<td></td>
</tr>
<tr>
<td>Albuminuria, occurrence</td>
<td>205</td>
<td>-1.10 (-1.74 to -0.47)</td>
<td></td>
</tr>
<tr>
<td>Albuminuria, time to event</td>
<td>978</td>
<td>-0.51 (-0.79 to -0.24)</td>
<td></td>
</tr>
<tr>
<td>eGFR, adult cohort</td>
<td>262</td>
<td>8.80 (1.70 to 15.90)</td>
<td></td>
</tr>
<tr>
<td>eGFR, pediatric cohort, Schwartz equation</td>
<td>521</td>
<td>-11.81 (-20.24 to -3.38)</td>
<td></td>
</tr>
<tr>
<td>eGFR, &lt;60 vs ≥60</td>
<td>262</td>
<td>-1.57 (-3.02 to -0.10)</td>
<td></td>
</tr>
<tr>
<td>Glomerular hyperfiltration, time to event</td>
<td>521</td>
<td>-0.80 (-1.35 to -0.24)</td>
<td></td>
</tr>
<tr>
<td>Glomerular hyperfiltration, mGFR&gt;110 vs ≤110</td>
<td>119</td>
<td>-1.65 (-2.74 to -0.55)</td>
<td></td>
</tr>
<tr>
<td>CKD progression, time to event</td>
<td>241</td>
<td>-0.18 (-1.16 to 0.82)</td>
<td></td>
</tr>
<tr>
<td>Renal length</td>
<td>372</td>
<td>-0.20 (-1.60 to 1.20)</td>
<td></td>
</tr>
<tr>
<td>Border classification, occurrence</td>
<td>52</td>
<td>0.18 (-1.20 to 1.63)</td>
<td></td>
</tr>
<tr>
<td>Patches, occurrence</td>
<td>52</td>
<td>0.96 (-0.36 to 2.21)</td>
<td></td>
</tr>
<tr>
<td>Posterior border, occurrence</td>
<td>52</td>
<td>0.00 (-1.20 to 1.34)</td>
<td></td>
</tr>
<tr>
<td>Splenectomy, occurrence</td>
<td>83</td>
<td>-1.35 (-3.51 to 0.80)</td>
<td></td>
</tr>
</tbody>
</table>

All associations are comparing 1 or 2 deletions vs no deletions, except fetal hemoglobin (HbF) and cholelithiasis (time to event), where only additive genotype coding was reported. Where appropriate, odds ratios and hazard ratios were transformed to β-scale (log odds ratio or log hazard ratio) for purposes of plotting. Results studied in more than 1 study were first combined via fixed-effects meta-analysis. Significant results as originally reported are shown in blue and gold for positive and negative association directions, respectively. For those meta-analyzed, significance was determined as P < 8.1 × 10^-4 (.05/62). ACS indicates acute chest syndrome; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; HbF, fetal hemoglobin; MCH, mean corpuscular hemoglobin; MCV, mean corpuscular volume; mGFR, measured glomerular filtration rate; RBC, red blood cell; SCD, sickle cell disease; SpO2, hemoglobin oxygen saturation; TCD, transcranial Doppler; UACR, urine albumin-to-creatinine ratio; and VOC, vaso-occlusive crisis.
Gene Ontology- and Reactome-curated pathways were examined for any pathways with significant enrichment among all genes with at least 1 reported significant result. The top 25 and all 22 with adjusted P < .05 are reported for Gene Ontology and Reactome, respectively. Circle size represents the number of genes in that pathway out of the total submitted genes (gene ratio); circle color, degree of significance after adjusting for multiple testing (adjusted P value, using the Benjamini-Hochberg method). ADME indicates absorption, distribution, metabolism, and excretion; BMP, bone morphogenetic proteins; CO2, carbon dioxide; O2, oxygen; PpS, Peters-plus syndrome; and TSR, thrombospondin type 1 repeat.
other loci with small association sizes or low frequency account for the remaining heritability. These additional modifier variants may be represented among the genes that lack validation.

This review also clarifies the role of α-thalassemia as a modifier of SCD severity by demonstrating that α-thalassemia was associated with reduced risk of clinically relevant SCD symptoms thought to be driven by hemolysis, including severe anemia, hyperbilirubinemia or gallstones, kidney dysfunction, and stroke. While this was a commonly accepted belief, our analysis consolidated the data to identify consistent trends among nonharmonized phenotypes with conflicting results.

To date, most validated studies of SCD modifiers have identified common variants with large association sizes (ie, “low-hanging fruit”) in relatively small cohorts. However, most genetic variation in health-related traits is driven by the interplay of many variants with small association sizes or low allele frequencies. Discovering such modifiers of SCD will require well-designed studies in larger cohorts using modern approaches to genetic analyses, including genome-wide association studies, adjustment for covariates, and, as discussed by Pincez, et al., multiomics. Future studies following best practices may also confirm candidate associations that have been reported in only 1 study. Moreover, studies in African cohorts could identify heretofore undiscovered variants with different frequencies in European and admixed cohorts. Most high-income countries, such as the US, have relatively few patients with SCD available for genetic studies. By contrast, while only 22% of individuals studied to date were from African cohorts, millions of individuals with SCD reside in sub-Saharan Africa, reflecting a fertile region for future research.

Advanced statistical and machine learning approaches, will also prove beneficial. For example, polygenic scores combining variants of small association sizes have been generated, including for HbF, pain, kidney, and cerebrovascular outcomes in SCD. However, those scores generally account for a small fraction of heritability and thus lack clinical utility. Improving polygenic scores for SCD phenotypes will require identifying more variants, validation, and rigorous testing, all of which would benefit from larger, more diverse cohorts and could be informed by analogous studies in non-SCD cohorts. Similarly, mendelian randomization, a method to explore causal relationships, has been used infrequently in SCD cohorts. In addition to increasing studies in Africa of both individuals with or without SCD, local ancestry inference in admixed individuals may help deconstruct associations driven by African ancestry.

Pathway analysis is another avenue for identifying potential candidate genes and generating clinically relevant hypotheses, even when traditional meta-analysis is not possible, as it can identify biologically meaningful pathways based on genes with significant associations and indicate other genes in these pathways that may contribute to disease risk. Among genes significantly associated with any SCD outcome in at least 1 study, we found enrichment in pathways controlling cellular adhesion, inflammation, response to toxic and oxidative stress, and blood vessel regulation, aligning with known disease pathophysiology. Other genes in those pathways are potential candidates for future investigation. We also identified potential therapeutic targets, such as flavonoid metabolic processes (of interest generally and in SCD), which were enriched for genes associated with hemolysis, anemia, and hyperbilirubinemia or biliary complications.

Limitations
This study has limitations that may confound or reduce the generalizability of our results. Subtype of SCD, ancestry, and hydroxyurea treatment status were not often reported in detail or, in the case of SCD subtype and ancestry, determined genetically; thus, we made no attempt to assess the difference between SCD subtypes or ancestries or to examine treatment response. Because most studies used a candidate gene approach, our results may be biased toward genes or pathways that were historically of high interest. Similarly, our analysis could be affected by unreported negative or contradictory results arising from positive publication bias. There are some methods, such as bayesian approaches, that cannot be integrated into a meta-analysis, resulting in some high-quality results being classified as exploratory. Finally, while our analysis categories allowed for a measure of study design and reporting rigor, they did not constitute a formal risk of bias assessment.
Conclusions

Although this systematic review and meta-analysis assessed 571 manuscripts that collectively reported 17,757 genetic associations with outcomes related to SCD severity, those associations validated in cross-study meta-analysis were largely related to HbF. To accelerate the understanding of the genetic etiology of SCD, future genetic association studies should report sufficient information for results to be included in meta-analyses. Elements of contemporary study design and international collaborations will improve scientific rigor, reduce the risk of false positives, and expand generalizability of study results. To facilitate cross-study analysis, the use of consensus measures is recommended for phenotypes and exposures. Combined, these steps will generate the high quality results necessary to develop clinically actionable genetic tools.
REFERENCES


43. Brewin JN, Smith AE, Cook R, et al. Genetic analysis of patients with sickle cell anemia and stroke before 4 years of age suggest an important role for apolipoprotein E. Circ Genom Precis Med. 2020;13(5):531-540. doi:10.1161/CIRCGEN.120.003025


123. Bhanushali AA, Himani K, Patra PK, Das BR. Hb F levels in Indian sickle cell patients and association with the HBB locus variant rs1012856 (C>T), and the HBG Xmnl (Arab-Indian) variant. *Hum Genet*. 2017;41(4-6):317-320. doi:10.1007/s00439-017-14059


Abu-Duhier F, Mir R. GSTT1(rs4025935)nullgenotype is associated with increased risk of sickle cell disease in the populations of Tabuk-Northwestern region of Saudi Arabia. *Hematology*. 2017;22(3):172-177. doi:10.1080/10245332.2016.1253253


211. Ellithy HN, Youssi S, Shahin GH. Relation between glutathione S-transferase genes (GSTM1, GSTT1, and GSTP1) polymorphisms and clinical manifestation of sickle cell disease in Egyptian patients. Hematology. 2015;20(10):598-606. doi:10.1179/1607845415Y.0000000013


SUPPLEMENT 1.
eMethods.
eFigure 1. PRISMA Flow Diagram For Manuscript Identification, Screening, and Inclusion
eFigure 2. Map of Included Patient Cohort Locations by Number of Individuals
eFigure 3. Data Strength Categories by Time
eFigure 4. Number of Papers Reporting Significant Associations With Fetal Hemoglobin (HbF) and Number of Unique Variants for Genes With Significant Results Reported in at Least Three Manuscripts
eTable 4. Phenotype Categories With Major Constituent Phenotypes
eTable 5. Number of Unique Patients and Studies by Country
eReferences.

SUPPLEMENT 2.
eTable 1. Manuscripts Included in This Systematic Review and Meta-Analysis
eTable 2. Data Extracted and Utilized in This Systematic Review and Meta-Analysis, Excluding Genome-Wide Summary Statistics
eTable 3. Genome-Wide Summary Statistics Extracted and Utilized in This Systematic Review and Meta-Analysis
eTable 6. Full List of Gene-Phenotype Category Associations
eTable 7. Gene Ontology Pathway Analysis Results
eTable 8. Reactome Pathway Analysis Results

SUPPLEMENT 3.
Data Sharing Statement