Article

Unraveling the Complex Genomic Interplay of Sickle Cell Disease Among the Saudi Population: A Case-Control GWAS Analysis

Ali Alghubayshi [1,2], Dayanjan Wijesinghe [2], Deemah Alwadaani [3,4], Farjah H. Algahtani [5], Salah Abohelaika [6,7], Mohsen Alzahrani [4,8], Hussain H. Al Saeed [9], Abdullah Al Zayed [9], Suad Alshammari [2,10], Yaseen Alhendi [4,11], Barrak Alsomaie [4,12], Abdulmonem Alsaleh [4,13] and Mohammad A. Alshabeeb [4,14,*]

[1] Department of Clinical Pharmacy, College of Pharmacy, University of Ha'il, Ha'il 55473, Saudi Arabia; a.alghubayshi@uoh.edu.sa
[2] Department of Pharmacotherapy and Outcomes Science, School of Pharmacy, Virginia Commonwealth University, Richmond, VA 23298, USA; wijesingheds@vcu.edu (D.W.); souad.hamed@nbu.edu.sa (S.A.)
[3] Medical Genomics Research Department, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia; alwadaanide@mngha.med.sa
[4] King Saud bin Abdulaziz University for Health Sciences (KSAU-HS), Ministry of National Guard Health Affairs (MNGHA), Riyadh 11426, Saudi Arabia; alzahranimo5@mngha.med.sa (M.A.); alhendiy@kaimrc.edu.sa (Y.A.); somaieb@mngha.med.sa (B.A.); alsalehabd@kaimrc.edu.sa (A.A.)
[5] Hematology/Oncology Center, King Saud University Medical City (KSUMC), Riyadh 11411, Saudi Arabia; falgahtani@ksu.edu.sa
[6] Research Department, Qatif Central Hospital (QCH), Qatif 32654, Saudi Arabia; sabohelaika@moh.gov.sa
[7] Pharmacy Department, Qatif Central Hospital (QCH), Qatif 32654, Saudi Arabia
[8] King Fahad Hospital, Ministry of National Guard Health Affairs (MNGHA), Riyadh 11426, Saudi Arabia
[9] Hematology Department, Qatif Central Hospital (QCH), Qatif 32654, Saudi Arabia; alsaeedhha@hotmail.com (H.H.A.S.); abalzayed@moh.gov.sa (A.A.Z.)
[10] Department of Clinical Pharmacy, College of Pharmacy, Northern Border University, Rafha 91911, Saudi Arabia
[11] Saudi Biobank Center, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia
[12] Operations Department, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia
[13] Blood and Cancer Research Department, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia
[14] Pharmaceutical Analysis Department, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia
* Correspondence: shabeebmo@mngha.med.sa

Abstract: Sickle cell disease (SCD) is a severe inherited blood disorder characterized by abnormal hemoglobin (HbS) that leads to varying degrees of severity, including chronic hemolysis, episodic vaso-occlusion, and damage to multiple organs, causing significant morbidity and mortality. While SCD is a monogenic disease, its complications are influenced by polygenic factors. SCD prevalence is notably high in regions including the Middle East, with Saudi Arabia reporting significant cases, particularly in the Eastern Province. Most genetic factors associated with SCD outcomes have been identified in populations predominantly from Africa or of African ancestry. This study aims to identify genetic variants that characterize Saudi SCD patients with the potential to influence disease outcomes in this population. A multicenter case-control genome-wide association study (GWAS) was conducted involving 350 adult Saudi SCD patients and 202 healthy controls. Participants were genotyped using the Affymetrix Axiom array, covering 683,030 markers. Rigorous quality control measures were applied to ensure data integrity. Fisher's exact was used to identify genetic variants with a significant difference in allele frequency ($p < 5 \times 10^{-8}$). Functional annotations and regulatory functions of variants were determined using the

Ensembl Variant Effect Predictor (VEP) and RegulomeDB databases. The GWAS identified numerous significant genetic variants characterizing SCD cases in the Saudi population. These variants, distributed across multiple chromosomes, were found in genes with known functional consequences. A substantial proportion of the markers were detected in the olfactory receptor cluster, *TRIM* family, and *HBB* locus genes. Many of the identified genes were reported in previous studies showing significant associations with various SCD outcomes, including hemoglobin regulation, inflammation, immune response, and vascular function. The findings highlight the genetic complexity underlying SCD and its clinical manifestations. The identified variants suggest potential molecular biomarkers and therapeutic targets, enhancing our understanding of the molecular basis of SCD in the Saudi population. This is the first genetic analysis characterizing SCD patients compared to healthy individuals, uncovering genetic markers that could serve as diagnostic biomarkers and therapeutic targets. Given the known molecular mechanisms of the detected genetic loci, these provide a foundation for precision medicine in SCD management, highlighting the need for further studies to validate these results and explore their clinical implications.

1. Introduction

Sickle cell disease (SCD) is an inherited red blood cell disorder in which the abnormal sickle hemoglobin leads to shortened red blood cell survival. SCD is a major global public health concern, characterized by a significant rise in prevalence worldwide. It has demonstrated a 41.4% increase from 2000 to 2021. Over this period, the number of SCD cases grew from 5.46 million to 7.74 million, causing approximately 376,000 annual deaths due to associated complications [1]. By 2050, it is predicted that the number of individuals with SCD will surge by 30% compared to current numbers [2]. SCD is prevalent among populations from Africa, India, the Caribbean, the Middle East, and the Mediterranean regions. Among the Middle Eastern Arab countries, SCD prevalence varies, with particular nations exhibiting amplified rates [3,4]. In the Gulf, specifically Saudi Arabia, SCD occurrence ranges from 2.6% to 4.2%, affecting 1 in 80 people in certain areas, with a 0.73% mortality rate owing to various complications [5–7]. However, a significant proportion of Saudis, notably those in the Eastern province (24%), either have the disease or carry one copy of the genetic mutation [8,9].

SCD arises from a non-synonymous missense single nucleotide polymorphism (SNP) [*rs334* c.20 A>T] in the *HBB* gene encoding the beta-globin protein, representing the fundamental basis underlying SCD pathology. This SNP substitution (adenine to thymine) in the sixth codon of *HBB* replaces the glutamine amino acid (encoded by GAG) with valine (encoded by GTG) in the translated protein. This missense mutation precipitates abnormal hemoglobin polymerization under low oxygen conditions, ultimately engendering the signature sickle morphology of red blood cells and manifesting the downstream clinical phenotype [10]. Sickling impedes oxygen transport and predisposes cells to hemolysis. Additionally, vascular obstructions and circulation of irregular cell adhesion molecules decrease vascular endothelial cell function, increase blood coagulation elements, and amplify inflammation biomarkers. These factors underlie classic complications such as painful ischemic crises, strokes, acute chest syndrome, and multiorgan damage, characterizing the spectrum of disease severity [11]. Despite SCD being a recognized Mendelian genetic

disorder exhibiting an autosomal recessive inheritance pattern, it demonstrates complex phenotypes shaped by multiple genes [10,12,13].

Despite the acknowledged necessity to characterize genetic variants linked to SCD outcomes across populations, the current literature remains deficient [14–21]. These variants, potentially modifying SCD severity, are integral for elucidating phenotypic diversity. As observed in hemoglobin-associated or subsidiary hemoglobin regulatory genes, they may exacerbate or mitigate clinical manifestations. Therefore, identifying mutations is imperative for detecting clinically significant markers and elucidating the biological pathways potentially linked to the varying burden and severity of SCD manifestations, stratifying patient risk and offering new therapeutic targets [5,22,23].

Earlier SCD genetics research predominantly utilized candidate gene approaches, focusing on particular populations, especially those of African ancestry [12,13,24]. This highlights a significant gap in the literature and underscores the need to identify more genetic signals potentially linked to SCD outcomes across diverse populations. Such knowledge is crucial for implementing personalized therapy for SCD patients and offers insights into the role of these variants in the disease's pathology [25]. The limited understanding of the genetic underpinnings of SCD, especially in underrepresented groups such as the Saudi population, presents a primary barrier to managing SCD complications in these individuals. This underscores the importance of identifying additional genetic markers that may influence SCD outcomes. Conducting genome-wide association studies (GWASs) is a broader genetic approach that could reveal additional insightful markers among this population compared to previously used candidate gene analyses.

Interestingly, in a previous GWAS, our research group identified several markers influencing thromboembolic risk in Saudi SCD patients, distinct from those previously identified in other populations [26]. This suggests that Saudi SCD patients may carry unique genetic markers that potentially explain the severity of complications. To our knowledge, no studies have yet compared SCD patients with healthy individuals using GWAS to identify distinctive genetic differences between cases and controls. This project aims to (i) explore the genetic makeup of the Saudi population with SCD and (ii) utilizing the literature, identify the biological pathways and molecular mechanisms associated with disease manifestations. The knowledge gained is expected to construct personalized prognostic models, identify critical targets for future drug development, and provide information on stratified therapeutic strategies tailored to Saudi patients' genetic makeup.

2. Results

2.1. Study Participant Characteristics

The demographic characteristics of the 552 enrolled subjects are summarized in Table 1. Age matching between cases and controls revealed mean values of 32.7 ± 10.2 years for cases versus 29.4 ± 8.4 years for controls. For gender, both cases and controls exhibited a male distribution of 54% and 51%, respectively. Principal component analysis revealed distinct clusters representing genetic variation between cases and controls (Figure 1). Scatter plot visualization of genotype data showed homogeneous dispersion for both groups, further validated by a permutation test (T1) yielding $p = 0.988$, which negates potential population stratification. Additionally, the genomic inflation rate was corrected using BACON, represented by a decreased lambda (λ) from 3.0 to 1.2 (Figure 2).

Table 1. Study participant characteristics.

Characteristic	Cases (n = 350)	Controls (n = 202)
Age (mean ± SD)	32.7 ± 10.2	29.4 ± 8.4
Male/female	195 (56%)/155 (44%)	103 (51%)/99 (49%)

SD = standard deviation.
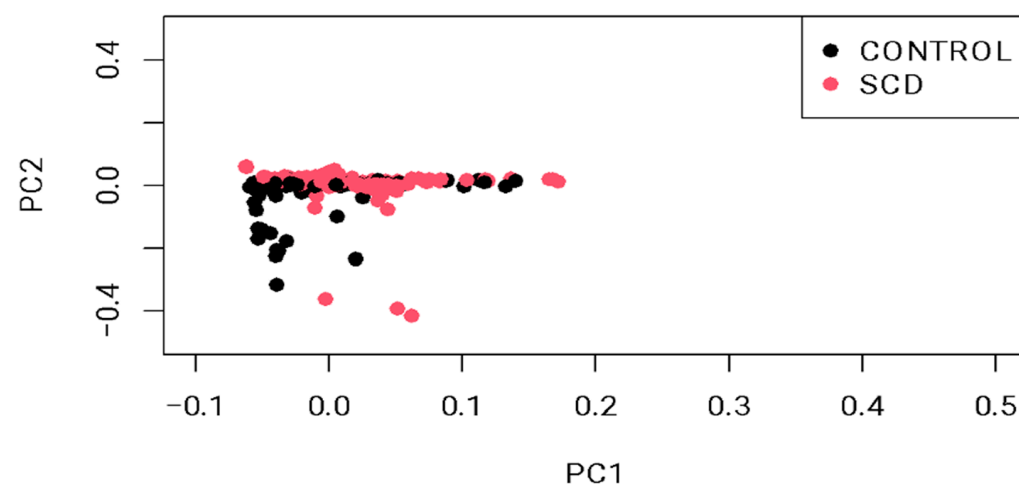
Figure 1. Principal component analysis (PCA) plot showing sample stratification. The permutation test (T1) confirmed the sample heterogeneity ($p = 0.988$).
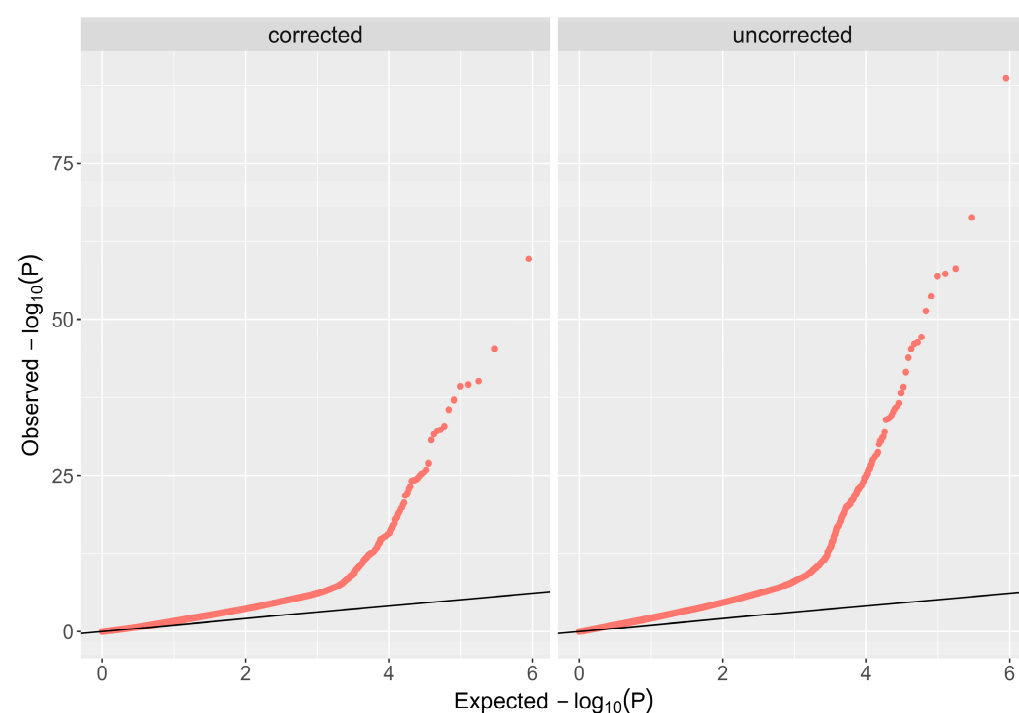
Figure 2. Corrected Q-Q plot. The Bayesian method (based on the BACON algorithm) was used to control inflation; the lambda (λ) value was reduced from 3.0 to 1.2.

2.2. Top Identified Variants with Functional Consequences

This GWAS analysis identified 48 variants with functional significance ($p < 5 \times 10^{-8}$) located within or nearby 31 known distinct genes. These variants, comprising both coding and regulatory non-coding markers, persisted after rigorous quality control measures and spanned over multiple chromosomes (Figure 3, Table 2). Notably, gene-based mapping revealed that approximately 23% of implicated variants clustered within olfactory receptor gene regions on

chromosome 11, which are known to regulate hemoglobin β expression (Supplementary Table S1). Specifically, we pinpointed high allele frequencies in coding mutations, including two missense variants: rs7933549 (MAF = 0.48) in *OR51V1*, rs2472530 (MAF = 0.45) in *OR52A5*, and a frameshift variant rs147062602 (MAF = 0.35) in *OR51B5*. In contrast, low inheritance levels were observed in a missense variant, rs12361955 (MAF = 0.077) in OR51S1, and a frameshift variant, rs112098990 (MAF = 0.029) in *OR52A1* with a statistically significant difference. Numerous other hemoglobinopathy-associated loci have emerged, such as rs2213169, rs2213170, and rs7130110 markers, which are neighbors to *HBE1* and *HBG2* genes. The frequencies of these alleles were found to be 43.7–46.6% in the SCD cohort compared to 1.5–1.8% in the healthy cohort. Additionally, a few common variants in SCD cases were mapped to tripartite motif (*TRIM*) gene family members on chromosome 11, such as *TRIM5*, with two missense mutations rs10838525 and rs11038628 (MAF = 0.27 vs. 0.05 and 0.33 vs. 0.07, respectively). Beyond these, several other markers within diverse genes across chromosome 11, including *RRM1*, *STIM1*, and *MMP26* (Table 2), are more common among SCD individuals.
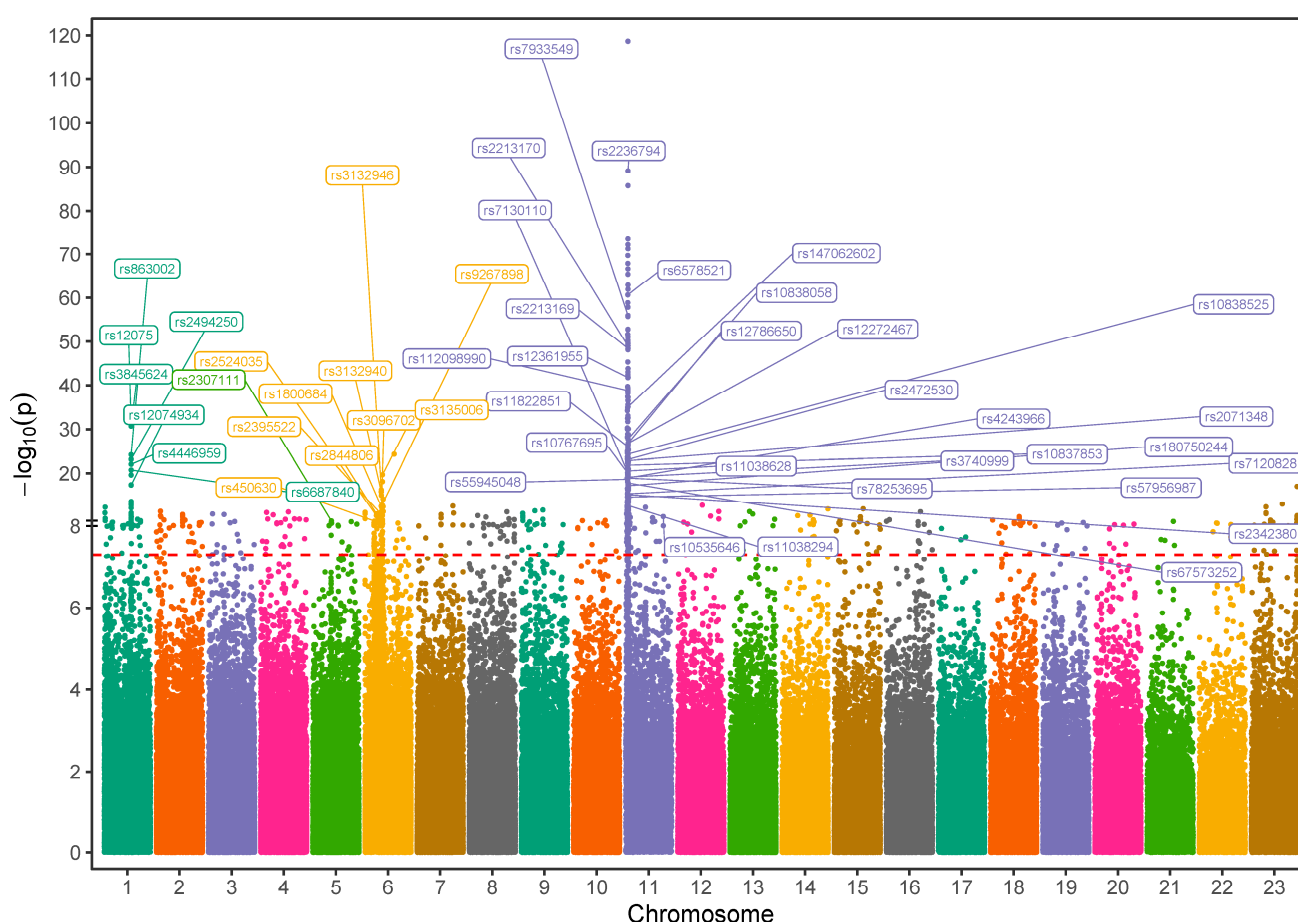

Figure 3. Manhattan plot visualization of genome-wide association study (GWAS) results in Saudi SCD cohort versus healthy Saudi individuals: The SNPs that met the genome-wide significance threshold ($p < 5 \times 10^{-8}$) are shown just above the red-dotted line. The SNPs were plotted in different colors to show a distinction between the chromosomes.

Table 2. Top variants with significant distribution difference in SCD cases than in healthy controls.

CHR	Gene	SNP	Variant Type	MAF (Cases)	MAF (Controls)	*p*-Value
1	ACKR1	rs12075 (A>G)	Missense	0.010	0.217	9.35×10^{-13}
		rs12074934 (T>G)	Intergenic	0.487	0.218	1.71×10^{-9}
		rs863002 (C>T)	Regulatory (CTCF site)	0.010	0.176	4.45×10^{-11}
6	AGER	rs1800684 (T>A)	Upstream	0.017	0.122	2.44×10^{-8}
1	FCER1A	rs2494250 (C>G)	Intron (500 B downstream)	0.011	0.174	3.31×10^{-11}
11	HBBP1, HBD	rs2071348 (T>G)	Intron	0.467	0.164	6.77×10^{-12}
11	HBE1, HBG2	rs2213170 (G>A)	Intron	0.440	0.015	1.78×10^{-11}
		rs7130110 (C>G)	Regulatory	0.466	0.179	9.21×10^{-11}
		rs2213169 (G>A)	Intron	0.437	0.015	2.17×10^{-11}
11	HBG2	rs2236794 (C>T)	Downstream	0.070	0.639	4.68×10^{-46}
6	HLA-A	rs2844806 (T>C)	Intergenic	0.330	0.536	5.92×10^{-9}
	HLA-G	rs2524035 (G>A)	Intron	0.095	0.239	2.85×10^{-8}
	HLA-DRB1/HLA-DQB1	rs3135006 (C>T)	Intergenic	0.082	0.294	8.82×10^{-14}
		rs2395522 (T>A)	Intergenic	0.409	0.617	8.00×10^{-9}
11	MMP26	rs6578521 (G>A)	Intron	0.069	0.510	1.38×10^{-33}
		rs11822851(A>G)	Intron	0.466	0.141	2.06×10^{-13}
1	MPTX1/CADM3-AS1	rs3845624 (C>A)	Intergenic	0.018	0.233	8.35×10^{-15}
6	NOTCH4	rs3132946 (G>A)	Intron	0.016	0.143	1.02×10^{-9}
		rs3132940 (G>T)	Intron	0.016	0.140	1.45×10^{-9}
		rs3096702 (G>A)	Upstream	0.077	0.222	7.92×10^{-9}
		rs9267898 (C>T)	Intergenic	0.063	0.209	1.65×10^{-9}
1	OR10J8P/OR10J9P	rs6687840 (T>C)	Intergenic	0.182	0.453	2.31×10^{-15}
		rs4446959 (C>T)	Intergenic	0.180	0.459	3.17×10^{-16}
11	OR51B5	rs147062602 (CAGCC-CCAG9GTCTGTGG>ins)	Frameshift	0.351	0.018	5.75×10^{-10}
		rs10838058 (A>G)	Intron	0.097	0.374	1.66×10^{-18}
		rs10837853 (G>A)	Intron	0.306	0.611	3.82×10^{-16}
		rs78253695 (GTC>del)	Intron	0.114	0.342	6.78×10^{-14}
		rs180750244 (G>A)	Intron	0.017	0.153	3.02×10^{-10}
11	OR51S1	rs12361955 (A>G)	Missense	0.077	0.425	2.09×10^{-25}
11	OR51V1	rs7933549 (G>A)	Missense	0.484	0.015	1.95×10^{-12}
11	OR52A1	rs112098990 (C>del)	Frameshift	0.029	0.308	9.87×10^{-20}
11	OR52A5	rs2472530 (A>G)	Missense	0.453	0.150	6.73×10^{-12}
5	POC5	rs2307111 (C>T)	Missense	0.397	0.592	4.47×10^{-8}
11	RRM1	rs55945048 (T>C)	Downstream	0.355	0.101	1.61×10^{-9}
6	SCAND3	rs450630 (A>G)	Missense	0.357	0.552	4.71×10^{-8}
11	SIDT2	rs10535646 (TGC>del)	Upstream	0.296	0.497	9.81×10^{-9}
11	STIM1	rs10767695(A>G)	Intron-NMD	0.597	0.300	4.23×10^{-11}
		rs7120828 (C>T)	Intron-NMD	0.401	0.161	2.47×10^{-8}
		rs4243966 (T>C)	Intron-NMD	0.044	0.223	1.04×10^{-12}
11	TRIM5	rs10838525 (C>T)	Missense	0.050	0.267	2.26×10^{-15}
		rs11038628 (C>T)	Missense	0.330	0.074	7.03×10^{-10}
		rs12786650 (T>C)	Intron	0.078	0.337	8.82×10^{-18}
		rs57956987 (T>C)	Intron	0.030	0.165	4.84×10^{-10}
11	TRIM6	rs3740999 (A>C)	Splice donor	0.075	0.283	1.45×10^{-13}
		rs11038294 (C>T)	Intron	0.107	0.282	4.55×10^{-10}
		rs12272467 (A>G)	Regulatory (TF site)	0.600	0.255	1.76×10^{-14}
11	TRIM22	rs67573252 (T>G)	Intron-NMD	0.050	0.228	2.54×10^{-12}
11	TRIM 34	rs2342380 (A>G)	Intron	0.136	0.345	1×10^{-11}

CHR: chromosome; SNP: single nucleotide polymorphism; MAF: minor allele frequency.

Outside chromosome 11, candidate variants showed genome-wide significance as well, including missense variants rs12075 in *ACKR1* on chromosome 1, rs2307111 in *POC5* on chromosome 5, and rs450630 in *SCAND3* on chromosome 6. Furthermore, various regulatory variants located on chromosome 6 within the *HLA* complex, *NOTCH*, and *AGER* genes were also identified. The phenotypes associated with the identified genes are described in Supplementary Table S1, as reported in the literature. This table summarizes the potential pathways linked with hemoglobinopathies for each detected gene or gene family.

3. Discussion

The application of GWAS has revolutionized our understanding of complex diseases by identifying genetic variants associated with disease risk across the genome [27]. Despite extensive efforts, the current literature on genetic factors beyond the known HBB loci that explain SCD outcomes remains insufficient [12]. Identification of these genetic variants is crucial for understanding the molecular basis of SCD and its phenotypic diversity and is essential for the implementation of personalized therapy. Previous GWASs have identified several associations between a few genetic markers and different SCD phenotypes, such as *rs3115229*, linked to vaso-occlusive pain episodes, and markers (e.g., rs766432 and rs1427407) near *BCL11A* and rs9494145 in *HBS1L-MYB*, correlated with variable fetal hemoglobin (HbF) expression [14,28]. Moreover, a GWAS analysis conducted on Saudi SCD patients identified seven markers associated with thromboembolic events, highlighting a significant risk in this population [26]. Building on these findings, the GWAS data were further analyzed to identify additional genetic variants that distinguish SCD patients.

This study pinpointed several noteworthy markers, and 62.5% of them are located on chromosome 11, with most of them (70%) mapped to the olfactory receptor (*OR*) gene cluster, *TRIM* family, and *HBB* locus genes. The *OR* genes, which constitute a significant portion of the mammalian genomes, are pseudogenes—a non-functional DNA segment [29]. Pseudogenes have no direct contribution to phenotypic traits; however, they tend to regulate the expression and function of genes. Numerous mutations in these genes may potentially affect phenotypic diversity through various mechanisms, such as impacting mRNA translation and manipulating epigenetic remodeling [26,30]. Notably, the *OR* gene family is aberrantly expressed in erythroid cells and located near the β-globin gene cluster. This proximity suggests a regulatory role in hemoglobinopathies, as these genes may modulate the chromatin structure at the CTCF binding site within the β-globin gene. This genetic linkage has been confirmed in both clinical and laboratory studies [31–34]. Also, various studies connect *OR* polymorphisms to SCD morbidity [12]. In our study, we observed a high prevalence of two missense variants: rs2472530 in *OR52A5* and rs7933549 in *OR51V1*, and a frameshift variant rs147062602 in *OR51B5* gene in SCD patients compared to the controls (MAF = 0.45, 0.48, and 0.35 among SCD patients versus 0.15, 0.015, and 0.018 in the controls, respectively). These variants have been previously reported as risk markers for thromboembolic events in Saudi SCD patients [26]. Additionally, we identified a novel frameshift variant, rs112098990, in the *OR52A1* gene, which typically introduces a stop codon, terminating the DNA sequence. Previously, an upstream variant rs4910715 in *OR52A1* was identified as part of a haplotype that characterizes classical sickle beta-globin [35]. *OR51V1* was also reported in a previous study linked to low HbA2 levels in healthy adults [36]. Another coding variant, rs12361955 in the *OR51S1* gene, was detected less frequently in our SCD cohort suggesting a protective effect. These findings are consistent with previous reports linking different mutations in OR cluster genes with multiple SCD-related phenotypes. These include hemolysis, variability in HbF levels, inflammation, hemostasis, and various hematological traits [12]. The reported data

suggests a role for OR genes in regulating β-globin gene clusters that possibly influence SCD phenotypes.

Our findings showed significant allele differences in nine variants in four *TRIM* family genes (*TRIM5*, *TRIM6*, *TRIM22*, and *TRIM34*), and two of them are missense variants: rs10838525 and rs11038628, located in *TRIM5*. The TRIM protein family is involved in immune modulation, cell-cycle progression, inflammation, oxidative stress, and hematopoietic stem-cell differentiation. All these mechanisms are critical for SCD complications. TRIM proteins play an important role in modulating TGF-β-activated kinase 1 (TAK1) to induce NF-κB and MAP kinase signaling. Kinase activation promotes NF-κB activation through IKKβ ubiquitination, regulates AP-1 signaling, and induces antiviral responses and innate immune activities [37]. *TRIM5*, for instance, which encodes immunomodulatory proteins, plays a defensive role against retroviral infections, while *TRIM6* is crucial in regulating stem-cell proliferation. Although neither *TRIM5* nor *TRIM6* are directly linked to hemoglobin expression—unlike *TRIM28*, which affects erythrocyte development as shown in animal models—their immune modulatory roles suggest they have an indirect influence on hematological outcomes. Thus, TRIM proteins were implicated in SCD pathology [12]. Furthermore, the previously reported GWAS data showed numerous associations between several loci in *TRIM* genes and a list of hematological and inflammatory traits. Notably, rs17305868 and rs11601507 in *TRIM5* indicated strong associations with increased platelet counts ($p = 9 \times 10^{-10}$) [38] and coronary artery disease ($p = 6 \times 10^{-13}$) [39], respectively. These pathophysiological phenotypes are common in SCD.

Other functional markers, with high prevalence in the SCD cohort, were detected across several genes throughout the β-globin locus on chromosome 11, such as rs2071348, rs2213169, rs2213170, and rs7130110. These variants have previously been linked to various hematological traits, such as hemostasis and thromboembolism, in SCD patients [26,40,41]. For example, rs7130110 in *HBE1* was identified as a marker impacting HbF levels and predicting the efficacy of hydroxyurea treatment in managing SCD [42,43]. Another variant (rs2213169) in *HBE1* was previously reported as a factor associated with reduced hematocrit volume [41]. These findings provide strong evidence suggesting *HBE1* as a potential pharmacogenomic marker for the treatment of SCD [43]. Moreover, two regulatory variants (rs6578521 and rs1182285) were identified in the *MMP26* gene, which encodes matrix metallopeptidase 26. This gene plays a crucial role in the degradation of extracellular matrix (ECM) components, an important process in angiogenesis, inflammation, and tissue remodeling [44]. Previous studies have shown that certain mutations in this gene are linked to hematological and neurological conditions, explained by ischemic stroke events [12]. Other related conditions, such as pneumonia risk and variability in cognitive abilities, have also been reported in the GWAS catalog database [45,46] (Supplementary Table S1). These traits are frequent too among SCD patients. Some interesting functional variants (rs55945048 and rs10535646), mapped to the *RRM1* and *SIDT2* genes, respectively, were also identified in our study. The dysregulation of these genes might have a direct or indirect impact on various SCD outcomes. For instance, RRM1 encodes ribonucleotide reductase, a rate-limiting enzyme in DNA synthesis, whereas hydroxyurea is a specific inhibitor of this enzyme [47]. Thus, significant links were found between the *RRM1* gene overexpression and hydroxyurea resistance [48,49], potentially due to increased enzyme activity overcoming hydroxyurea's inhibitory effects. A previous study reported that 18% of SCD patients discontinued taking this medication for various reasons, not explicitly indicating resistance. These reasons include the drug's immunosuppressive effects, failure to improve HbF levels, and likely also non-compliance [50,51]. Genetic factors, such as the carriage of certain *RRM1* mutations, may partially explain the variability in responses of patients to hydroxyurea. Another important marker characterizing SCD patients is

rs10767695 in the *STIM1* gene, with a double prevalence of the variant allele among the SCD group (60%) compared to the healthy cohort (30%). A previous study linked this variant to mean platelet volume [52]. The markers in *MMP26, RRM1, SIDT2,* and *STIM1* align with assorted molecular pathways linked to endothelial dysfunction, inflammation, vascular disease, and numerous SCD-associated complications [44,48,49,53,54] as described in Supplementary Table S1.

Beyond chromosome 11, three coding variants (missense) were found to be more common among the SCD cohort involving rs12075 in *ACKR1* on chromosome 1, rs2307111 in *POC5* on chromosome 5, and rs450630 in *SCAND3* on chromosome 6. The dysregulation of these genes was associated with various SCD outcomes [55–58]. The *ACKR1* gene, also known as the Duffy Antigen Receptor for Chemokines (DARC), is expressed on red blood cells and endothelial cells and plays an important role in modulating inflammation and chemokine levels. It may specifically modulate leukocyte trafficking and is predicted to be involved in various homeostasis mechanisms underlying a variety of SCD phenotypes. For example, it has been reported to potentially influence the severity of organ damage associated with SCD, including anemia, leg ulcers, priapism, and kidney dysfunction [12,55]. In addition, previous studies have demonstrated that numerous genetic variations in *ACKR1*, including the detected missense mutation rs12075 (G>A) in our study, are significantly associated with the inflammatory cascade among European populations [59]. The wild type allele (G) predisposes patients to a severe pattern of SCD. Our data revealed that this allele is more prevalent among the SCD cases. A previous study suggested a role for rs3845624 in regulating C-reactive protein levels in plasma and rs2494250 in (C-C motif ligand 2) CCL2 measurement [59,60]. CCL2 is a chemotactic factor, known as monocytic chemotactic protein 1 (*MCP-1*), which employs macrophages to induce the immune response. CCL2 and its receptors (ACKR1 and ACKR2) impact disease outcomes complementarily [61]. Additionally, the missense variant rs450630 in *SCAND3* was reported recently as a marker associated with decreased hemoglobin concentration ($p = 6 \times 10^{-11}$) [58].

Several regulatory variants that possibly contribute to variable disease phenotypes were detected in our study. These include the intergenic variant rs3845624 located between the *MPTX1* and *CADM3*-AS1 genes, linked with inflammation and adhesion regulations, and the intronic SNP rs2494250 located near *FCER1A* gene. As the *FCER1A* gene encodes the alpha chain of the high-affinity IgE receptor (FcεRI), it plays a role in allergic responses. This explains the reported association between *rs2494250* and the non-immediate urticaria/angioedema (NIUA), a type of delayed hypersensitivity reaction (DHR) that can occur hours to days after exposure to a triggering substance, particularly nonsteroidal anti-inflammatory drugs (NSAIDs) [62,63]. DHR can be life-threatening, particularly in the context of transfusion complications in SCD patients, and typically occurs 5–10 days following transfusion. These reactions often result in decreased hemoglobin levels and pain crises [64]. The incidence of DHR related to blood transfusion (alloimmunization) seems to be notably high (16.7%) in Saudi SCD patients [65]. Other identified variants with high regulatory function scores were notable across chromosome 6. Most of these markers were mapped to the *HLA* gene family, *AGER*, and *NOTCH* genes. The ontological pathways connecting these genes to SCD complications can be explained through their roles in regulating inflammation, immune responses, and endothelial function. A linkage was detected previously between several markers in these genes with various SCD phenotypes, such as vaso-occlusive crisis and organ damage [12]. Among these markers, the SNP *rs3135006* located in the *HLA-DQB1* gene was reported earlier as a factor for increased ratio levels of cell adhesion proteins [Activated Leukocyte Cell Adhesion Molecule (ALCAM)/Cadherin 5 (CDH5)] [66]. In addition, the detected variant rs2524035 in *HLA-G* was demonstrated previously as a factor associated with decreased lymphocyte levels ($p = 3 \times 10^{-23}$) [67].

Moreover, previous reports highlighted several *HLA* variants that are linked with the risk of stroke, acute chest syndrome (ACS), pain episodes, cholelithiasis, and hyperbilirubinemia in SCD patients [12,13].

Another notable marker in our study (rs1800684) in the *AGER* gene was nominated in a previous GWAS study as a factor associated with a higher expression of the Transforming Growth Factor Beta Receptor 2 (TGFBR2) ($p = 3 \times 10^{-13}$) [66]. Genetic mutations in both AGER and TGF-β receptors have been previously shown to increase the risk of various SCD complications, including stroke, priapism, infection, avascular necrosis, pulmonary hypertension, acute chest syndrome, and acute pain crises [12,13]. Additionally, four variants in *NOTCH4* on chromosome 6 were detected. This gene plays an important role in regulating endothelial function, inflammation, and hematopoiesis, protecting against endothelial dysfunction and apoptosis and balancing proinflammatory and inflammatory responses through various signaling pathways, with notable anti-inflammatory effects by reducing proinflammatory cytokine expression and co-stimulatory protein levels [68,69]. Among the identified variants, *rs3132940* was found to be associated with an increased risk of sarcoidosis disorder, specifically non-Lofgren's syndrome without extrapulmonary manifestations [70]. This serious condition was reported previously in SCD patients, noting that patients suffering from ACS and sarcoidosis had a higher mortality rate [71].

This study represents the first GWAS analysis comparing SCD patients with healthy individuals, with a particular focus on the Saudi population. It included the largest genome scan of Saudi SCD patients to date. The study identified SCD distinguishing markers, suggesting their potential as powerful diagnostic biomarkers and therapeutic targets. The identified genetic variants may influence SCD pathophysiology through potential alterations in gene regulation, including changes in transcription factor binding, gene expression levels, or mRNA splicing patterns. These molecular changes could affect vascular function, inflammatory responses, and related physiological processes relevant to SCD. The co-occurrence of potentially functional coding and non-coding variation across these priority candidate genes emphasizes their likely role in modulating SCD phenotypic expression through diverse molecular genetic mechanisms. Collectively, these data implicate a complex interplay among variants distributed across the genome and concentrated in key genes that may contribute to driving SCD pathophysiology through a multitude of effects.

4. Materials and Methods

4.1. Sample Recruitment

As summarized in Figure 4, this multicenter case-control study enrolled 350 unrelated adult SCD subjects (cases) and 202 healthy adults (controls), aged ≥18 years, encompassing both males and females. SCD diagnosis was confirmed via positive sickling assay and verification of homozygosity for the causative *rs334* A>T mutation. Participating patients routinely attended one of three hematology clinics: King Fahad Hospital (KFH) in Riyadh under the Ministry of National Guard Health Affairs (MNGHA), contributing 106 cases; King Khalid University Hospital (KKUH) in Riyadh under King Saud University Medical Center (KSUMC), contributing 82 cases; and Qatif Central Hospital (QCH) in the Eastern Province under the Ministry of Health, contributing 162 cases. The 202 control specimens were obtained from the KAIMRC biorepository. The study predominantly focused on HbSS homozygotes, excluding other SCD genotypes like HbSC, HbS-beta-thalassemia, HbSD, and HbSO, which were identified through hemoglobin electrophoresis. Notably, a majority of KKUH and KFH Riyadh patients were referrals from the Saudi Arabian southwestern or northern regions.
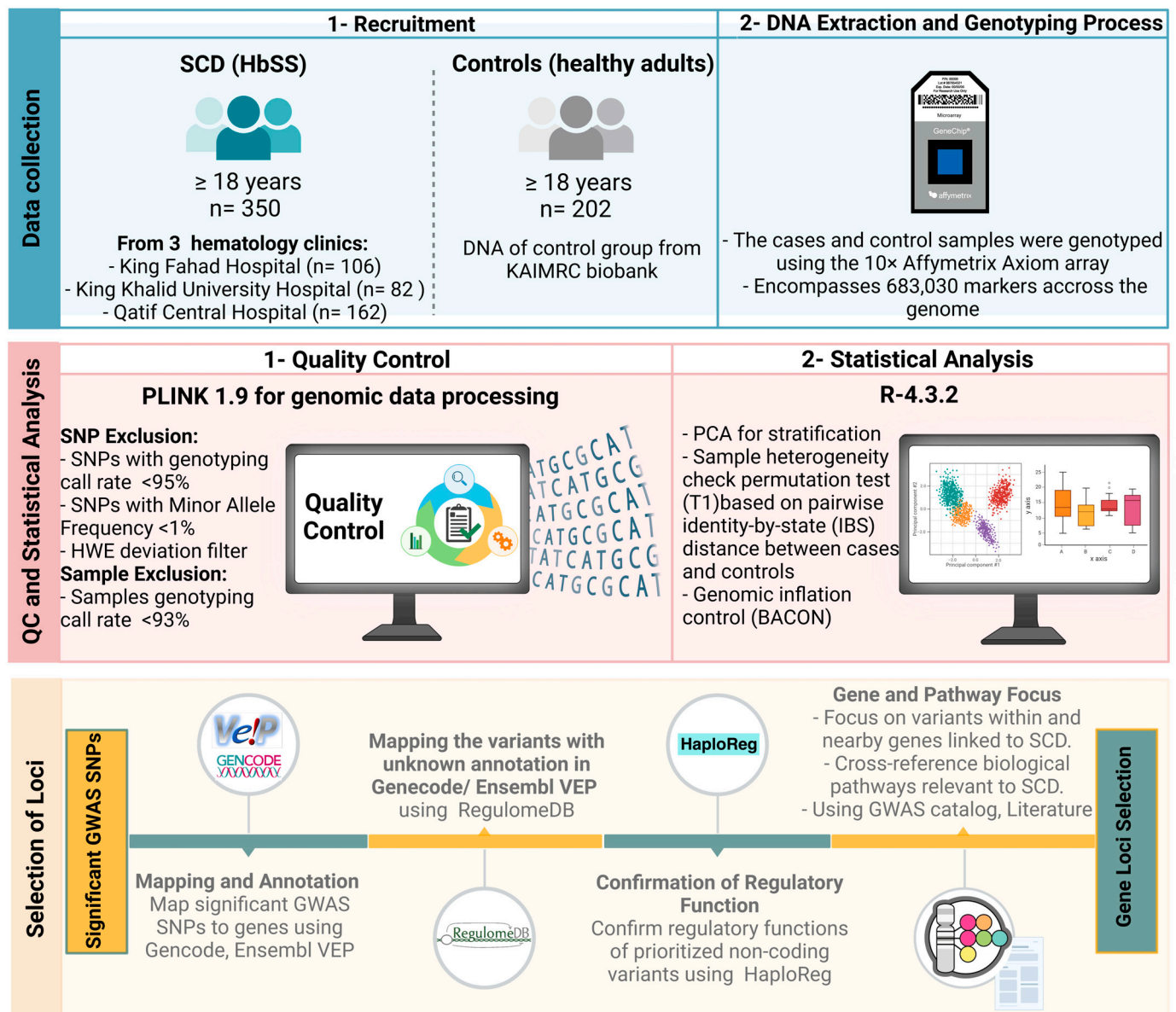
Figure 4. Integrated workflow for enrollment, DNA extraction, quality control, and statistical analysis in a GWAS of Saudi sickle cell disease.

4.2. Genomic Analysis and Quality Control (QC)

The cases and control samples were genotyped using the Affymetrix Axiom array (Axiom 2.0 reagent kit designed by Applied Biosystems[TM], Waltham, MA, USA, catalog number #901758), which encompasses 683,030 markers for the GWAS. All details regarding the genomic analysis, which includes sample processing, genotyping, quality control call rates (CRs), validity, and reliability of the platform used, have been previously published [26]. Several quality control tests were applied to filter out low-quality genotyping data. SNPs that had a genotyping call rate < 95%, a minor allele frequency (MAF) < 1%, or a Hardy–Weinberg equilibrium (HWE) p-value < 0.05 were excluded. In addition, samples with genotyping CR < 93% were omitted.

4.3. Statistical Analysis

Plink software version 1.9 was employed for genomic data processing. Genotype differences between groups were estimated using a Fisher's exact test. For GWAS, a threshold of $p < 5 \times 10^{-8}$ was established to identify loci with a statistically significant difference.

Mean and standard deviation calculations were performed to evaluate age matching. R statistical package (qqman) version R-4.3.2. was used to generate Manhattan, quantile-quantile (Q-Q), and principal component analysis (PCA) plots. Population stratification was also assessed, with sample heterogeneity evaluated through the permutation test (T1) based on pairwise identity-by-state (IBS) distance between cases and controls. Additionally, the Bayesian method, utilizing the BACON (Bayesian Analysis Computation and Optimization) standard, was used to control the genomic inflation rate.

4.4. Selection of Loci

The identified signals were mapped to genes based on their original annotations and the Gencode [72]. The Ensembl Variant Effect Predictor (VEP) tool was then used to confirm annotation and prioritize the identified markers based on their predicted effects [73]. The VEP is a robust and well-established tool that uses automated annotation pipelines based on experimental evidence highlighting consequence types and gene biotypes. The VEP tool was also used to determine the functionality of variants located in non-coding regions. We further utilized the RegulomeDB database to annotate variants that were not detected by VEP. RegulomeDB enabled us to prioritize the identified variants based on their potential regulatory function; it provides ranks for the screened variants [74]. Variants with high ranks, indicating strong evidence of functionality, were selected. Furthermore, the HaploReg database was used as a confirmatory step to identify the regulatory functions of non-coding variants [75]. The variants located within or near known genes and potentially linked to a pathological mechanism or outcomes of SCD were selected. Due to the lack of access to phenotypic data in our SCD cohort, we cross-referenced biological pathways linked to candidate genes using the literature and previously identified phenotypes indicated in the GWAS catalog website [76]. We hypothesized their relevance to SCD outcomes based on known molecular functions and ontologies curated in databases related to blood disorders and SCD phenotypes. Furthermore, genes identified with numerous associated SNPs, including both coding and non-coding variants with functional consequences, were selected as high priority markers.

5. Conclusions

This GWAS has uncovered significant genetic markers in Saudi SCD patients across multiple chromosomes. Our findings provide novel insights into SCD genetics in the Saudi population, particularly involving key gene families dominant in the disease cohort. A substantial portion of our findings aligns with genes previously implicated in various SCD complications, lending support to the validity of our approach. These results contribute to the broader understanding of SCD's molecular basis and underscore the importance of population-specific genetic factors in SCD research. The identified markers hold potential for informing future personalized medicine approaches in SCD management. Replication studies in independent cohorts are essential to validate these results and exclude potential technical or methodological biases. Furthermore, functional studies are needed to elucidate the mechanisms underlying disease severity and complications. This work represents a significant step toward comprehending the genetic landscape of SCD in the Saudi population and may guide future diagnostic and therapeutic strategies.

Limitations

While this study provides valuable insights into the genetic landscape of SCD in the Saudi population, several limitations should be noted. The sample size, although substantial, may not capture all possible genetic variations in the Saudi SCD population. The lack of detailed phenotypic data limits our ability to directly link genetic variants to specific

SCD complications or severity levels. Additionally, the focus on the Saudi population may limit the generalizability of our findings to other ethnic groups. Furthermore, while we identified potentially functional variants, experimental validation of their biological effects was beyond the scope of this study. Our GWAS approach may not capture all relevant genetic variations, particularly rare variants.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms26062817/s1. References [31–33,37,44,48,49,53–58,68,69,77–87] are cited in Supplementary Materials.

Author Contributions: Conceptualization, A.A. (Ali Alghubayshi), F.H.A., S.A. (Salah Abohelaika), M.A., H.H.A.S., A.A.Z. and M.A.A.; Methodology, A.A. (Ali Alghubayshi), D.W., D.A., Y.A., A.A. (Abdulmonem Alsaleh) and M.A.A.; Software, A.A. (Ali Alghubayshi), D.A., S.A. (Suad Alshammari) and A.A. (Abdulmonem Alsaleh); Validation, D.W.; Formal analysis, D.W. and M.A.A.; Investigation, D.A.; Resources, F.H.A., S.A. (Salah Abohelaika), M.A., H.H.A.S., A.A.Z. and B.A.; Data curation, A.A. (Ali Alghubayshi), F.H.A., S.A. (Salah Abohelaika), M.A., H.H.A.S., A.A.Z., S.A. (Suad Alshammari) and M.A.A.; Writing—original draft, A.A. (Ali Alghubayshi) and M.A.A.; Writing—review & editing, A.A. (Ali Alghubayshi), D.W. and M.A.A. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study received ethical approval from the Institutional Review Boards (IRBs) of King Abdullah International Medical Research Center (KAIMRC), Saudi Arabia (Ref: IRBC/1414/19 on 21 August 2019); Qatif Central Hospital, Saudi Arabia (QCH-SREC0216/2020 on 21 December 2020); and Virginia Commonwealth University (VCU), the United States of America (HM20030698 on 20 August 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original data of this study are included in the article. However, some detailed genotyping data are unavailable due to privacy aspects. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. GBD 2021 Sickle Cell Disease Collaborators. Global, regional, and national prevalence and mortality burden of sickle cell disease, 2000–2021: A systematic analysis from the Global Burden of Disease Study 2021. *Lancet Haematol.* **2023**, *10*, e585–e599. [CrossRef] [PubMed]
2. Piel, F.B.; Hay, S.I.; Gupta, S.; Weatherall, D.J.; Williams, T.N. Global burden of sickle cell anaemia in children under five, 2010–2050: Modelling based on demographics, excess mortality, and interventions. *PLoS Med.* **2013**, *10*, e1001484. [CrossRef] [PubMed]
3. El-Hazmi, M.A.F.; Al-Hazmi, A.M.; Warsy, A.S. Sickle cell disease in Middle East Arab countries. *Indian J. Med. Res.* **2011**, *134*, 597–610. [CrossRef]
4. Grosse, S.D.; Odame, I.; Atrash, H.K.; Amendah, D.D.; Piel, F.B.; Williams, T.N. Sickle cell disease in Africa: A neglected cause of early childhood mortality. *Am. J. Prev. Med.* **2011**, *41* (Suppl. S4), S398–S405. [CrossRef]
5. Bin Zuair, A.; Aldossari, S.; Alhumaidi, R.; Alrabiah, M.; Alshabanat, A. The Burden of Sickle Cell Disease in Saudi Arabia: A Single-Institution Large Retrospective Study. *Int. J. Gen. Med.* **2023**, *16*, 161–171. [CrossRef]
6. Alotaibi, M.M. Sickle cell disease in Saudi Arabia: A challenge or not. *J. Epidemiol. Glob. Health* **2017**, *7*, 99–101. [CrossRef]

7. Al-Ali, A.K.; Alsulaiman, A.; Alfarhan, M.; Safaya, S.; Vatte, C.B.; Albuali, W.M.; Qutub, H.O.; Alzahrani, A.J.; Milton, J.N.; Steinberg, M.H. Sickle cell disease in the Eastern Province of Saudi Arabia: Clinical and laboratory features. *Am. J. Hematol.* **2021**, *96*, E117–E121. [CrossRef]

8. Jastaniah, W. Epidemiology of sickle cell disease in Saudi Arabia. *Ann. Saudi Med.* **2011**, *31*, 289–293. [CrossRef]

9. Al-Suliman, A.; Elsarraf, N.A.; Baqishi, M.; Homrany, H.; Bousbiah, J.; Farouk, E. Patterns of mortality in adult sickle cell disease in the Al-Hasa region of Saudi Arabia. *Ann. Saudi Med.* **2006**, *26*, 487–488. [CrossRef]

10. Inusa, B.P.D.; Hsu, L.L.; Kohli, N.; Patel, A.; Ominu-Evbota, K.; Anie, K.A.; Atoyebi, W. Sickle Cell Disease-Genetics, Pathophysiology, Clinical Presentation and Treatment. *Int. J. Neonatal Screen* **2019**, *5*, 20. [CrossRef]

11. Kavanagh, P.L.; Fasipe, T.A.; Wun, T. Sickle Cell Disease: A Review. *JAMA* **2022**, *328*, 57–68. [CrossRef] [PubMed]

12. Kirkham, J.K.; Estepp, J.H.; Weiss, M.J.; Rashkin, S.R. Genetic Variation and Sickle Cell Disease Severity: A Systematic Review and Meta-Analysis. *JAMA Netw. Open* **2023**, *6*, e2337484. [CrossRef] [PubMed]

13. Driss, A.; Asare, K.O.; Hibbert, J.M.; Gee, B.E.; Adamkiewicz, T.V.; Stiles, J.K. Sickle Cell Disease in the Post Genomic Era: A Monogenic Disease with a Polygenic Phenotype. *Genom. Insights* **2009**, *2009*, 23–48.

14. Chaturvedi, S.; Bhatnagar, P.; Bean, C.J.; Steinberg, M.H.; Milton, J.N.; Casella, J.F.; Barron-Casella, E.; Arking, D.E.; DeBaun, M.R. Genome-wide association study to identify variants associated with acute severe vaso-occlusive pain in sickle cell anemia. *Blood* **2017**, *130*, 686–688. [CrossRef]

15. Griffin, P.J.; Sebastiani, P.; Edward, H.; Baldwin, C.T.; Gladwin, M.T.; Gordeuk, V.R.; Chui, D.H.; Steinberg, M.H. The genetics of hemoglobin A 2 regulation in sickle cell anemia. *Am. J. Hematol.* **2014**, *89*, 1019–1023. [CrossRef]

16. Mtatiro, S.N.; Singh, T.; Rooks, H.; Mgaya, J.; Mariki, H.; Soka, D.; Mmbando, B.; Msaki, E.; Kolder, I.; Thein, S.L.; et al. Genome Wide Association Study of Fetal Hemoglobin in Sickle Cell Anemia in Tanzania. *PLoS ONE* **2014**, *9*, e111464. [CrossRef]

17. Milton, J.N.; Sebastiani, P.; Solovieff, N.; Hartley, S.W.; Bhatnagar, P.; Arking, D.E.; Dworkis, D.A.; Casella, J.F.; Barron-Casella, E.; Bean, C.J.; et al. A Genome-Wide Association Study of Total Bilirubin and Cholelithiasis Risk in Sickle Cell Anemia. *PLoS ONE* **2012**, *7*, e34741. [CrossRef]

18. Solovieff, N.; Milton, J.N.; Hartley, S.W.; Sherva, R.; Sebastiani, P.; Dworkis, D.A.; Klings, E.S.; Farrer, L.A.; Garrett, M.E.; Ashley-Koch, A.; 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. [CrossRef]

19. Flanagan, J.M.; Frohlich, D.M.; Howard, T.A.; Schultz, W.H.; Driscoll, C.; Nagasubramanian, R.; Mortier, N.A.; Kimble, A.C.; Aygun, B.; Adams, R.J.; et al. Genetic predictors for stroke in children with sickle cell anemia. *Blood* **2011**, *117*, 6681–6684. [CrossRef]

20. Ashley-Koch, A.E.; Okocha, E.C.; Garrett, M.E.; Soldano, K.; De Castro, L.M.; Jonassaint, J.C.; Orringer, E.P.; Eckman, J.R.; Telen, M.J. MYH9 and APOL1 are both associated with sickle cell disease nephropathy. *Br. J. Haematol.* **2011**, *155*, 386–394. [CrossRef]

21. Saraf, S.L.; Zhang, X.; Shah, B.; Kanias, T.; Gudehithlu, K.P.; Kittles, R.; Machado, R.F.; Arruda, J.A.; Gladwin, M.T.; Singh, A.K.; et al. Genetic variants and cell-free hemoglobin processing in sickle cell nephropathy. *Haematologica* **2015**, *100*, 1275–1284. [CrossRef] [PubMed]

22. Sundd, P.; Gladwin, M.T.; Novelli, E.M. Pathophysiology of Sickle Cell Disease. *Annu. Rev. Pathol.* **2019**, *14*, 263–292. [CrossRef] [PubMed]

23. Alsultan, A.; Aleem, A.; Ghabbour, H.; AlGahtani, F.H.; Al-Shehri, A.; Osman, M.E.; Kurban, K.; Alsultan, M.S.; Bahakim, H.; Al-Momen, A.M. Sickle Cell Disease Subphenotypes in Patients from Southwestern Province of Saudi Arabia. *J. Pediatr. Hematol./Oncol.* **2012**, *34*, 79–84. [CrossRef] [PubMed]

24. Pincez, T.; Ashley-Koch, A.E.; Lettre, G.; Telen, M.J. Genetic Modifiers of Sickle Cell Disease. *Hematol. Oncol. Clin. N. Am.* **2022**, *36*, 1097–1124. [CrossRef]

25. Chang, A.K.; Ginter Summarell, C.C.; Birdie, P.T.; Sheehan, V.A. Genetic modifiers of severity in sickle cell disease. *Clin. Hemorheol. Microcirc.* **2018**, *68*, 147–164. [CrossRef]

26. Alshabeeb, M.A.; Alwadaani, D.; Al Qahtani, F.H.; Abohelaika, S.; Alzahrani, M.; Al Zayed, A.; Al Saeed, H.H.; Al Ajmi, H.; Alsomaie, B.; Rashid, M.; et al. Impact of Genetic Variations on Thromboembolic Risk in Saudis with Sickle Cell Disease. *Genes* **2023**, *14*, 1919. [CrossRef]

27. Witte, J.S. Genome-wide association studies and beyond. *Annu. Rev. Public Health* **2010**, *31*, 9–20. [CrossRef]

28. Bhatnagar, P.; Purvis, S.; Barron-Casella, E.; DeBaun, M.R.; Casella, J.F.; Arking, D.E.; Keefer, J.R. Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. *J. Hum. Genet.* **2011**, *56*, 316–323. [CrossRef]

29. Olender, T.; Jones, T.E.M.; Bruford, E.; Lancet, D. A unified nomenclature for vertebrate olfactory receptors. *BMC Evol. Biol.* **2020**, *20*, 42. [CrossRef]

30. Roberts, T.C.; Morris, K.V. Not so pseudo anymore: Pseudogenes as therapeutic targets. *Pharmacogenomics* **2013**, *14*, 2023–2034. [CrossRef]

31. Bulger, M.; Bender, M.A.; van Doorninck, J.H.; Wertman, B.; Farrell, C.M.; Felsenfeld, G.; Groudine, M.; Hardison, R. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14560–14565. [CrossRef] [PubMed]

32. Yang, Q.; Kathiresan, S.; Lin, J.P.; Tofler, G.H.; O'Donnell, C.J. Genome-wide association and linkage analyses of hemostatic factors and hematological phenotypes in the Framingham Heart Study. *BMC Med. Genet.* **2007**, *8* (Suppl. S1), S12. [CrossRef]

33. Feingold, E.A.; Penny, L.A.; Nienhuis, A.W.; Forget, B.G. An olfactory receptor gene is located in the extended human beta-globin gene cluster and is expressed in erythroid cells. *Genomics* **1999**, *61*, 15–23. [CrossRef]

34. Hou, C.; Dale, R.; Dean, A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 3651–3656. [CrossRef] [PubMed]

35. Hanchard, N.; Elzein, A.; Trafford, C.; Rockett, K.; Pinder, M.; Jallow, M.; Harding, R.; Kwiatkowski, D.; McKenzie, C. Classical sickle beta-globin haplotypes exhibit a high degree of long-range haplotype similarity in African and Afro-Caribbean populations. *BMC Genet.* **2007**, *8*, 52. [CrossRef]

36. Menzel, S.; Garner, C.; Rooks, H.; Spector, T.D.; Thein, S.L. HbA2 levels in normal adults are influenced by two distinct genetic mechanisms. *Br. J. Haematol.* **2013**, *160*, 101–105. [CrossRef]

37. Uchil, P.D.; Hinz, A.; Siegel, S.; Coenen-Stass, A.; Pertel, T.; Luban, J.; Mothes, W. TRIM protein-mediated regulation of inflammatory and innate immune signaling and its association with antiretroviral activity. *J. Virol.* **2013**, *87*, 257–272. [CrossRef]

38. Kanai, M.; Akiyama, M.; Takahashi, A.; Matoba, N.; Momozawa, Y.; Ikeda, M.; Iwata, N.; Ikegawa, S.; Hirata, M.; Matsuda, K.; et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat. Genet.* **2018**, *50*, 390–400. [CrossRef]

39. Van der Harst, P.; Verweij, N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ. Res.* **2018**, *122*, 433–443. [CrossRef]

40. Wojcik, G.L.; Graff, M.; Nishimura, K.K.; Tao, R.; Haessler, J.; Gignoux, C.R.; Highland, H.M.; Patel, Y.M.; Sorokin, E.P.; Avery, C.L.; et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **2019**, *570*, 514–518. [CrossRef]

41. Li, J.; Glessner, J.T.; Zhang, H.; Hou, C.; Wei, Z.; Bradfield, J.P.; Mentch, F.D.; Guo, Y.; Kim, C.; Xia, Q.; et al. GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. *Hum. Mol. Genet.* **2013**, *22*, 1457–1464. [CrossRef] [PubMed]

42. Sales, R.R.; Nogueira, B.L.; Tosatti, J.A.G.; Gomes, K.B.; Luizon, M.R. Do Genetic Polymorphisms Affect Fetal Hemoglobin (HbF) Levels in Patients with Sickle Cell Anemia Treated with Hydroxyurea? A Systematic Review and Pathway Analysis. *Front. Pharmacol.* **2021**, *12*, 779497. [CrossRef] [PubMed]

43. Ware, R.E.; Despotovic, J.M.; Mortier, N.A.; Flanagan, J.M.; He, J.; Smeltzer, M.P.; Kimble, A.C.; Aygun, B.; Wu, S.; Howard, T.; et al. Pharmacokinetics, pharmacodynamics, and pharmacogenetics of hydroxyurea treatment for children with sickle cell anemia. *Blood* **2011**, *118*, 4985–4991. [CrossRef] [PubMed]

44. Jabłońska-Trypuć, A.; Matejczyk, M.; Rosochacki, S. Matrix metalloproteinases (MMPs), the main extracellular matrix (ECM) enzymes in collagen degradation, as a target for anticancer drugs. *J. Enzyme Inhib. Med. Chem.* **2016**, *31* (Suppl. S1), 177–183. [CrossRef]

45. Ford, A.L.; Ragan, D.K.; Fellah, S.; Binkley, M.M.; Fields, M.E.; Guilliams, K.P.; An, H.; Jordan, L.C.; McKinstry, R.C.; Lee, J.M.; et al. Silent infarcts in sickle cell disease occur in the border zone region and are associated with low cerebral blood flow. *Blood* **2018**, *132*, 1714–1723. [CrossRef]

46. Young, R.C.; Castro, O.; Baxter, R.P.; Dunn, R.; Armstrong, E.M.; Cook, F.J.; Sampson, C.C. The lung in sickle cell disease: A clinical overview of common vascular, infectious, and other problems. *J. Natl. Med. Assoc.* **1981**, *73*, 19–26.

47. Machlin, E.S.; Sarnow, P.; Sagan, S.M. Masking the 5′ terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3193–3198. [CrossRef]

48. Zhan, Y.; Jiang, L.; Jin, X.; Ying, S.; Wu, Z.; Wang, L.; Yu, W.; Tong, J.; Zhang, L.; Lou, Y.; et al. Inhibiting RRM2 to enhance the anticancer activity of chemotherapy. *Biomed. Pharmacother.* **2021**, *133*, 110996. [CrossRef]

49. McClarty, G.A.; Chan, A.K.; Engstrom, Y.; Wright, J.A.; Thelander, L. Elevated expression of M1 and M2 components and drug-induced posttranscriptional modulation of ribonucleotide reductase in a hydroxyurea-resistant mouse cell line. *Biochemistry* **1987**, *26*, 8004–8011. [CrossRef]

50. Gohal, G.A.; Gosadi, I.M.; Cittana Iqbal, B.A.; Ghazwani, Y.H.; Daghriri, A.M.; Shugairi, A.A.; Daghriri, K.A.; Zurayyir, A.J.; Nemri, A.A.; Abdulhaq, M.A. Utilization of Hydroxyurea Among Patients Diagnosed with Sickle Cell Disease in Jazan, Saudi Arabia. *Patient Prefer. Adherence* **2022**, *16*, 3059–3067. [CrossRef]

51. Chand, A.R.; Xu, H.; Wells, L.G.; Clair, B.; Neunert, C.; Spellman, A.E.; Clay, L.J.; Natrajan, K.; Kutlar, A. Are There True Non-Responders to Hydroxyurea in Sickle Cell Disease? A Multiparameter Analysis. *Blood* **2014**, *124*, 4073. [CrossRef]

52. Chen, M.H.; Raffield, L.M.; Mousas, A.; Sakaue, S.; Huffman, J.E.; Moscati, A.; Trivedi, B.; Jiang, T.; Akbari, P.; Vuckovic, D.; et al. Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. *Cell* **2020**, *182*, 1198–1213.e14. [CrossRef] [PubMed]

53. Qian, D.; Cong, Y.; Wang, R.; Chen, Q.; Yan, C.; Gong, D. Structural insight into the human SID1 transmembrane family member 2 reveals its lipid hydrolytic activity. *Nat. Commun.* **2023**, *14*, 3568. [CrossRef] [PubMed]

54. Silva-Rojas, R.; Laporte, J.; Böhm, J. STIM1/ORAI1 Loss-of-Function and Gain-of-Function Mutations Inversely Impact on SOCE and Calcium Homeostasis and Cause Multi-Systemic Mirror Diseases. *Front. Physiol.* **2020**, *11*, 604941. [CrossRef]

55. Nebor, D.; Durpes, M.C.; Mougenel, D.; Mukisi-Mukaza, M.; Elion, J.; Hardy-Dessources, M.D.; Romana, M. Association between Duffy antigen receptor for chemokines expression and levels of inflammation markers in sickle cell anemia patients. *Clin. Immunol.* **2010**, *136*, 116–122. [CrossRef]

56. Drasar, E.R.; Menzel, S.; Fulford, T.; Thein, S.L. The effect of Duffy antigen receptor for chemokines on severity in sickle cell disease. *Haematologica* **2013**, *98*, e87–e89. [CrossRef]

57. Xu, F.; Zhang, L.; Xu, Y.; Song, D.; He, W.; Ji, X.; Shao, J. Hypermethylation of SCAND3 and Myo1g Gene Are Potential Diagnostic Biomarkers for Hepatocellular Carcinoma. *Cancers* **2020**, *12*, 2332. [CrossRef]

58. Timoteo, V.J.; Chiang, K.M.; Yang, H.C.; Pan, W.H. Common and ethnic-specific genetic determinants of hemoglobin concentration between Taiwanese Han Chinese and European Whites: Findings from comparative two-stage genome-wide association studies. *J. Nutr. Biochem.* **2023**, *111*, 109126. [CrossRef]

59. Naitza, S.; Porcu, E.; Steri, M.; Taub, D.D.; Mulas, A.; Xiao, X.; Strait, J.; Dei, M.; Lai, S.; Busonero, F.; et al. A genome-wide association scan on the levels of markers of inflammation in Sardinians reveals associations that underpin its complex regulation. *PLoS Genet.* **2012**, *8*, e1002480. [CrossRef]

60. Benjamin, E.J.; Dupuis, J.; Larson, M.G.; Lunetta, K.L.; Booth, S.L.; Govindaraju, D.R.; Kathiresan, S.; Keaney, J.F.; Keyes, M.J.; Lin, J.P.; et al. Genome-wide association with select biomarker traits in the Framingham Heart Study. *BMC Med. Genet.* **2007**, *8* (Suppl. S1), S11. [CrossRef]

61. Lin, Z.; Shi, J.L.; Chen, M.; Zheng, Z.M.; Li, M.Q.; Shao, J. CCL2: An important cytokine in normal and pathological pregnancies: A review. *Front. Immunol.* **2022**, *13*, 1053457. [CrossRef] [PubMed]

62. Jurado-Escobar, R.; Dona, I.; Montes, A.T.; Bartra, J.; Sanchez, N.P.; Laguna, J.; Cruz-Amaya, A.; de Santamaria, R.S.; Nuñez, R.; Salas, M.; et al. FCERIA Single Nucleotide Polymorphisms Associated with Nonsteroidal Anti-inflammatory Drug-induced Acute Urticaria/Angioedema. *J. Allergy Clin. Immunol.* **2022**, *149* (Suppl. S2), AB61. [CrossRef]

63. Doña, I.; Blanca-López, N.; Torres, M.J.; Gómez, F.; Fernández, J.; Zambonino, M.A.; Monteseirín, F.J.; Canto, G.; Blanca, M.; Cornejo-García, J.A. NSAID-induced urticaria/angioedema does not evolve into chronic urticaria: A 12-year follow-up study. *Allergy* **2014**, *69*, 438–444. [CrossRef] [PubMed]

64. Scheunemann, L.P.; Ataga, K.I. Delayed hemolytic transfusion reaction in sickle cell disease. *Am. J. Med. Sci.* **2010**, *339*, 266–269. [CrossRef]

65. Kuriri, F.A.; Ahmed, A.; Alanazi, F.; Alhumud, F.; Ageeli Hakami, M.; Atiatalla Babiker Ahmed, O. Red Blood Cell Alloimmunization and Autoimmunization in Blood Transfusion-Dependent Sickle Cell Disease and β-Thalassemia Patients in Al-Ahsa Region, Saudi Arabia. *Anemia* **2023**, *2023*, 3239960. [CrossRef]

66. Suhre, K. Genetic associations with ratios between protein levels detect new pQTLs and reveal protein-protein interactions. *Cell Genom.* **2024**, *4*, 100506. [CrossRef]

67. Astle, W.J.; Elding, H.; Jiang, T.; Allen, D.; Ruklisa, D.; Mann, A.L.; Mead, D.; Bouman, H.; Riveros-Mckay, F.; Kostadima, M.A.; et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **2016**, *167*, 1415–1429.e19. [CrossRef]

68. López-López, S.; Romero de Ávila, M.J.; Hernández de León, N.C.; Ruiz-Marcos, F.; Baladrón, V.; Nueda, M.L.; Laborda, J.; García-Ramírez, J.J.; Monsalve, E.M.; Díaz-Guerra, M.J. NOTCH4 Exhibits Anti-Inflammatory Activity in Activated Macrophages by Interfering with Interferon-γ and TLR4 Signaling. *Front. Immunol.* **2021**, *12*, 734966. [CrossRef]

69. Zhou, B.; Lin, W.; Long, Y.; Yang, Y.; Zhang, H.; Wu, K.; Chu, Q. Notch signaling pathway: Architecture, disease, and therapeutics. *Signal Transduct. Target. Ther.* **2022**, *7*, 95. [CrossRef]

70. Rivera, N.V.; Ronninger, M.; Shchetynsky, K.; Franke, A.; Nöthen, M.M.; Müller-Quernheim, J.; Schreiber, S.; Adrianto, I.; Karakaya, B.; van Moorsel, C.H.; et al. High-Density Genetic Mapping Identifies New Susceptibility Variants in Sarcoidosis Phenotypes and Shows Genomic-driven Phenotypic Differences. *Am. J. Respir. Crit. Care Med.* **2016**, *193*, 1008–1022. [CrossRef]

71. El Sharu, H.; Zweigle, J.; Jamil, M.; Singh, S.; Cowles, S.; Liles, D.K. Outcomes of Sarcoidosis on Patients with Sickle Cell Disease—A Review of the National Inpatient Database 2016–2020. *Blood* **2023**, *142* (Suppl. S1), 1135. [CrossRef]

72. Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S.; et al. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **2012**, *22*, 1760–1774. [CrossRef] [PubMed]

73. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [CrossRef]

74. RegulomeDB Search—RegulomeDB. Available online: https://regulomedb.org/regulome-search (accessed on 2 September 2024).

75. HaploReg v4.2. Available online: https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php (accessed on 2 September 2024).

76. Buniello, A.; MacArthur, J.A.L.; Cerezo, M.; Harris, L.W.; Hayhurst, J.; Malangone, C.; McMahon, A.; Morales, J.; Mountjoy, E.; Sollis, E. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **2019**, *47*, D1005–D1012.

77. Conran, N.; Belcher, J.D. Inflammation in sickle cell disease. *Clin. Hemorheol. Microcirc.* **2018**, *68*, 263–299. [CrossRef]

78. Wong, K.; Lai, W.K.; Jackson, D.E. HLA Class II regulation of immune response in sickle cell disease patients: Susceptibility to red blood cell alloimmunization (systematic review and meta-analysis). *Vox Sang.* **2022**, *117*, 1251–1261. [CrossRef]

79. Tamouza, R.; Neonato, M.G.; Busson, M.; Marzais, F.; Girot, R.; Labie, D.; Elion, J.; Charron, D. Infectious complications in sickle cell disease are influenced by HLA class II alleles. *Hum. Immunol.* **2002**, *63*, 194–199. [CrossRef]

80. Mahdi, N.; Al-Ola, K.; Al-Subaie, A.M.; Ali, M.E.; Al-Irhayim, Z.; Al-Irhayim, A.Q.; Almawi, W.Y. HLA class II haplotypes distinctly associated with vaso-occlusion in children with sickle cell disease. *Clin. Vaccine Immunol.* **2008**, *15*, 729–731. [CrossRef]

81. Martins, J.O.; Pagani, F.; Dezan, M.R.; Oliveira, V.B.; Conrado, M.; Ziza, K.C.; Gualandro, S.F.; Langui, D.M.; Bordin, J.O.; Rocha, V.; et al. Impact of HLA-G +3142C>G on the development of antibodies to blood group systems other than the Rh and Kell among sensitized patients with sickle cell disease. *Transfus. Apher. Sci.* **2022**, *61*, 103447. [CrossRef]

82. Nagata, Y.; Suzuki, R. FcεRI: A Master Regulator of Mast Cell Functions. *Cells* **2022**, *11*, 622. [CrossRef]

83. Himadewi, P.; Wang, X.Q.D.; Feng, F.; Gore, H.; Liu, Y.; Yu, L.; Kurita, R.; Nakamura, Y.; Pfeifer, G.P.; Liu, J.; et al. 3′HS1 CTCF binding site in human β-globin locus regulates fetal hemoglobin expression. *Elife* **2021**, *10*, e70557. [CrossRef] [PubMed]

84. Hassan, A.; Bagu, E.T.; Patten, S.A.; Molidperee, S.; Parent, S.; Barchi, S.; Villemure, I.; Tremblay, A.; Moldovan, F. Differential Regulation of POC5 by ERα in Human Normal and Scoliotic Cells. *Genes* **2023**, *14*, 1111. [CrossRef] [PubMed]

85. Geng, M.Y.; Wang, L.; Song, Y.Y.; Gu, J.; Hu, X.; Yuan, C.; Yang, M.; Pei, W.J.; Zhang, Y.; Gao, J.L. Sidt2 is a key protein in the autophagy-lysosomal degradation pathway and is essential for the maintenance of kidney structure and filtration function. *Cell Death Dis.* **2021**, *13*, 7. [CrossRef] [PubMed]

86. Roos, J.; DiGregorio, P.J.; Yeromin, A.V.; Ohlsen, K.; Lioudyno, M.; Zhang, S.; Safrina, O.; Kozak, J.A.; Wagner, S.L.; Cahalan, M.D.; et al. STIM1, an essential and conserved component of store-operated Ca^{2+} channel function. *J. Cell Biol.* **2005**, *169*, 435–445. [CrossRef]

87. Antwi-Boasiako, C.; Donkor, E.S.; Sey, F.; Dzudzor, B.; Dankwah, G.B.; Otu, K.H.; Doku, A.; Dale, C.A.; Ekem, I. Levels of Soluble Endothelium Adhesion Molecules and Complications among Sickle Cell Disease Patients in Ghana. *Diseases* **2018**, *6*, 29. [CrossRef]