Article

Genome-Wide Analysis of Exertional Rhabdomyolysis in Sickle Cell Trait Positive African Americans

Mingqiang Ren [1,2,*], Nyamkhishig Sambuughin [1,2], Ognoon Mungunshukh [1,2,3], Daniel Baxter Edgeworth [1,2], Daniel Hupalo [3], Xijun Zhang [3], Matthew D. Wilkerson [3], Clifton L. Dalgard [3], Francis G. O'Connor [1] and Patricia A. Deuster [1]

1. Consortium for Health and Military Performance, Department of Military and Emergency Medicine, F. Edward Hébert School of Medicine, Uniformed Services University, Bethesda, MD 20814, USA; daniel.edgeworth.ctr@usuhs.edu (D.B.E.); francis.oconnor@usuhs.edu (F.G.O.)
2. Henry M Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD 20817, USA
3. Department of Anatomy, Physiology, and Genetics, Center for Military Precision Health, Uniformed Services University, Bethesda, MD 20814, USA
* Correspondence: mingqiang.ren.ctr@usuhs.edu

Abstract: Sickle cell trait (SCT), although generally a benign carrier state of hemoglobin S (HbAS), is a risk factor for exertional rhabdomyolysis (ERM), a rare but potentially fatal consequence of highly intense physical exercise, particularly among active-duty military personnel and high-performance athletes. The association between SCT and ERM is poorly understood. The objective of this study was to elucidate the genetic basis of ERM in an SCT-positive African American cohort. SCT-positive African Americans with a personal history of ERM (cases, n = 30) and without history of ERM (controls, n = 53) were enrolled in this study. Whole-genome sequencing was performed on DNA samples isolated from peripheral white blood cells. Participants' demographic, behavioral, and medical history information was obtained. An additional 131 controls were extracted from SCT-positive subjects of African descent from the 1000 Genomes Project. SCT carriers with ERM were characterized by myotoxicity features, significant muscle involvement dominated by muscle weakness, and severe pain and substantial increase in serum creatine kinase, with a mean value of 50,480 U/L. A distinctive feature of the SCT individuals with ERM was exertional collapse, which was reported in 53.3% of the cases in the study cohort. An important factor for the development of ERM was the duration and frequency of strenuous physical activity in the cases compared to the controls. Whole-genome sequencing identified 79,696 protein-coding variants. Genome-wide association analysis revealed that the p.C477R, rs115958260 variant in the SLC44A3 gene was significantly associated with ERM event in SCT-positive African Americans. The study results suggest that a combination of vigorous exercise and a genetic predisposing factor is involved in ERM.

Keywords: sickle cell trait; exertional rhabdomyolysis; genome-wide association study; SCL44A3 gene; whole-genome sequencing

1. Introduction

Approximately 300 million people worldwide [1] and nearly 7.8% of African Americans in the US [2] have sickle cell trait (SCT), a carrier state of sickle hemoglobin S (HbS) that, in HbSS form, causes Sickle Cell Disease (SCD) [3]. Although SCT carrier status is largely considered benign, individuals with SCT are at increased risk for negative health outcomes during intense physical exertion, especially in hot climates and high-elevation locations, or when dehydrated [4]. Our previous epidemiological study with a large cohort (n = 47,944) of active-duty African Americans showed that SCT is associated with a significantly higher risk of exertional rhabdomyolysis (ERM) [2], and this conclusion is also supported by a recent study from the UK Biobank [5]. In addition, cohort studies have demonstrated that military recruits and college athletes with SCT have increased

risk of exercise-related death compared to peers without SCT [6,7]. In 2019, a consensus conference we had the opportunity to lead, explored this increased risk of death in those with SCT, outlined critical research gaps [8].

Rhabdomyolysis is an acute and potentially fatal clinical syndrome, which can lead to a variety of complications, including acute kidney injury [9,10]. There are a number of potential etiologies, including exertion, extreme temperature changes, ischemia, infections, immobility, drugs, toxins, endocrine causes, autoimmune reactions, trauma, or genetic conditions [11–13]. Inherited muscle disorders associated with rhabdomyolysis are heterogeneous and rare. Genetic variants associated with rhabdomyolysis have been reviewed by Scalco et al. [14]. However, these or other variants are still insufficient to predict an individual's clinical susceptibility to rhabdomyolysis. Currently, there are limited genomic studies on rhabdomyolysis [12,15,16]. The underlying causes of rhabdomyolysis remain unknown in many cases [12,16,17]. This suggests that additional studies are needed to identify causative disease genes and variants associated with rhabdomyolysis.

ERM is a serious consequence of intense, repetitive, and prolonged physical exercise, particularly among active-duty military personnel and high-performance athletes [18,19]. From 2019 to 2020, the national estimate of hospital visits in the US for ERM was 40,654, which translates to 0.66 visits per 100,000 population. The highest risk groups for ERM are young males and African Americans [20]. Clinically, ERM is most commonly defined as severe myalgia, muscle weakness, and swelling, with sudden elevation of serum creatine kinase (CK) levels with or without the presence of myoglobinuria upon significant exertion [12]. Unaccustomed exercise or exercise in combination with other individual and environmental factors such as dehydration, drugs and toxins, infections, and hyperthermia is a frequent cause of ERM. On the other hand, ERM, especially recurrent ERM, can also be a presentation of inherited metabolic myopathies, intramuscular calcium release, mitochondrial cytopathy, and muscular dystrophies [12,21,22]. However, these rare inherited muscle disorders are not likely a shared cause of ERM associated with SCT. The underlying cause of the variable degree of susceptibility to ERM in SCT individuals remains unknown.

In this study, we aimed to characterize shared clinical features of a series of SCT/ERM cases and to determine genetic markers contributing to ERM in SCT individuals by using a case-control study with whole-genome sequencing (WGS) and exploring genome-wide association approach. To increase the statistical power of the genetic association, we took advantage of SCT-positive genomes from the 1000 Genomes Project data that are widely used in genome-wide association studies (GWAS) [23,24]. We reasoned that the exertion level of our enrolled controls would be comparable to that of SCT individuals from the 1000 Genomes Project, which were added to the control group in our GWAS analysis. We found that a non-synonymous variant leading to c.T1429C:p.C477R (rs115958260) in the SLC44A3 gene was significantly associated with ERM occurrence in SCT-positive subjects. Our study suggests that abnormal function of choline-like transporter 3 (solute carrier transporters) could potentially be associated with ERM development.

2. Materials and Methods

2.1. Study Participants and Clinical Information

The study protocol was approved by the Uniformed Services University (USU) Institutional Review Board. An informed consent document was obtained from each participant prior to enrollment. We enrolled African American carriers of SCT with a history of ERM (case) and with no history of ERM (control). The study participants were recruited nationwide via referral from military and civilian health care providers. The participants were also invited through advertisements at various treatment facilities, social media, and verbal communications. Following enrollment, participant demographics (age and gender), physical characteristics (height and weight), and behavioral patterns (physical activity, drug and supplement usage), as well as medical and family histories, were obtained through self-reported questionnaires and medical records when available. All cases were evaluated by study physicians in person or via phone interview to obtain clinical history and test

results during or after rhabdomyolysis episodes. The enrollment period for this study is from 14 December 2016 to 21 September 2021.

2.2. DNA Sample Handling and Whole-Genome Sequencing

Genomic DNA was isolated from peripheral white blood cells by using the QIAamp DNA blood mini kit (Qiagen) following the manufacturer's instructions. Genomic DNA quantitation was performed using a Qubit® dsDNA HS Assay Kit (Molecular Probes) and measured by an Invitrogen™ Qubit™ 4 Fluorometer. Carrier status of SCT was first validated by genotyping the E7V mutation in the HBB gene by restriction enzyme DdeI. Genotyping of HBB was finally confirmed by Sanger DNA sequencing. DNA library preparation and whole-genome sequencing (WGS) were described per Soltis et al. [25]. WGS was performed on an Illumina HiSeq X System with 151 + 7 + 151 cycle parameters by using HiSeq X HD SBS Kit reagents (Illumina, San Diego, CA, USA).

2.3. Germline Variant Calling

Raw sequencing reads were first converted to standard FASTQ format by using bcl2fastq2 conversion software v2.20 (Illumina). Quality control of the raw reads was performed with FastQC. The GATK package was used for variant calling. The GATK pipeline was based on the best practices workflow for germlines established by the Broad Institute [26]. Briefly, reads were aligned to the reference genome (NCBI GRCh38) using BWA-MEM. Picard tools were then used to sort and mark PCR duplicate reads and generate BAM files. GATK version 4.2.0 was used to recalibrate BAM files with BaseRecalibrator and ApplyBQSR, and to generate VCF and GVCF files with HaplotypeCaller. The VCF files were then used for annotation and analysis.

2.4. Variant Annotation and Filtering

Variant annotation was performed by the Ensembl VEP [27]. Variants were then filtered with various criteria: (1) single nucleotide polymorphism (SNP) call rate $\geq$ 95%, (2) a minor allele frequency (MAF) > 1% (estimated using combined datasets from the 1000 Genomes Project and dbSNP154), (3) Hardy-Weinberg equilibrium p-value $\geq 1 \times 10^{-4}$, (4) non-synonymous and splice variants, and (5) protein-coding genes (excluded noncoding genes). Variants located in introns, 3' or 5' untranslated regions (UTRs), upstream, downstream, or noncoding regions, or those that would cause synonymous mutations and in-frame indels, were mostly not included as the majority of variations occurring in these regions have insufficient evidence. In addition, samples were checked for relatedness and sex mismatch. The relationship was estimated through identity-by-descent (IBD) segments detection with Plink (v1.9) software [28] and further confirmed by using KING (v2.3.0) and PRIMUS (v1.9.0) software [29,30]. Three related subjects with pi-hat threshold > 0.2 were removed from the control group [31].

2.5. Case-Control Genome-Wide Association Analysis

PLINK (v1.9) software [28] was used to perform the basic genome-wide association study with sex as a covariate. Additional controls (n = 131) of African descent were extracted from the 1000 Genomes Project, and their genomic variant data were merged with our cohort data for a final GWAS. In total, 106 females and 101 males, in the age range 18–75 years, were included in this study. Considering our small sample size, we intentionally set the significant threshold p-values at 5×10^{-8}, which is more strenuous than Bonferroni correction p-value (0.05/79,696 = 6.27×10^{-7}). The qqman R package was used for graphical representation of the GWAS single-locus analysis results (Manhattan plot). Genetic variants were annotated using the Variant Effect Predictor tool (version 104) [27].

2.6. Gene-Based Association Study and Enrichment Analyses

MAGMA software (version 1.10) [32] was used to calculate gene-wise statistics. This software detects multi-marker effects, taking into account the physical distance and linkage

disequilibrium between SNPs [32]. These analyses used a 50 kb upstream and downstream window around each gene to capture potential regulatory variants of these genes. Bonferroni correction of *p*-values was performed on the resultant genes (15,694 genes). Finally, 3.2×10^{-6} (0.05/15,694) was the *p*-value threshold established for this study for statistical significance.

3. Results

3.1. Study Participant Characteristics

The demographic and behavioral characteristics of the SCT individuals who participated in this study are presented in Table 1. Male subjects comprised 86.7% of the cases, whereas 30.4% of the controls were male. The median age was also significantly different between cases and controls. Cases also showed significantly more strenuous physical activity compared to controls, with cases tending to perform activities frequently and for longer periods of time. These differences between the cases and controls were likely due to the fact that most of the cases (70%) were active-duty Service Members. Cases had lower BMI compared to controls, although the differences were not significant. There were no differences in medical and family histories between cases and controls. All the participants denied frequent use of drugs associated with rhabdomyolysis, such as statin, opioids, and antipsychotics.

Table 1. General characteristics of SCT-positive subjects enrolled in the study.

Characteristics	Cases (n = 30)	Controls (n = 56)	*p*-Value
Gender (n, %)			<0.001
Male	26 (86.7)	17 (30.4)	
Female	4 (13.3)	39 (69.6)	
Age at enrollment (years)			
Median (range)	29.5 (18–69)	45 (18–75)	
Mean (SD)	33.8 (14.2)	43.6 (15.7)	0.003
BMI (n, %)			
Mean (SD)	27.2 (6.5)	29.5 (6.6)	0.104
Physical activity %			
Strenuous activity low (≤75%)	45	79	<0.001
Strenuous activity high (≥75%)	55	21	<0.001
Smoking (n, %)	3 (10%)	1 (1.8%)	0.14

Clinical characteristics of the ERM cases are shown in Table 2. Clinical data included peak creatine kinase (CK) levels of greater than 5 times the normal upper limit (200 U/L), presenting symptoms during or after rhabdomyolysis episodes, and other co-morbidities. The mean peak CK was 50,480 U/L, with large variability in range (989–400,000 U/L) among cases. Severe muscle pain was the most common complaint during an episode, followed by weakness/fatigue. Almost half (53.3%) of the cases also reported exertional collapse. Other complications were compartment syndrome, acute kidney injury, and exertional heat illness. Recurrent episodes were observed in 39% of the cases, and the remaining cases were subjects with single episodes.

Table 2. Clinical characteristics of SCT subjects with ERM.

Characteristics	Reported (%)
Creatine kinase (U/L)	
Median (range)	11,000 (989–400,000)
"Coca-Cola"-colored urine	26.7%
Number of episodes	
Single	61%
Recurrent (two or more)	39%

Table 2. *Cont.*

Characteristics	Reported (%)
Muscle symptoms during event (n)	
Weakness/fatigue	60%
Pain	88.9%
Pain (mean on a scale of 1–10)	7.4 ± 2.6
Swelling	23.3%
Cramps	30%
Exertional collapse	53.3%
Exertional heat illness	10%
Acute kidney injury	13.3%
Compartment syndrome	16.7%

3.2. Single Nucleotide Polymorphisms (SNP) Based on GWAS of ERM in Subjects with SCT

We performed WGS in 23 of 30 cases and 53 controls (3 of 56 family-related subjects were removed in the final analysis). Variant calling was performed using the GATK best-practices workflow (GATK V4.2.0). We first combined our genomic variant data with subset data of SCT carriers (n = 131) extracted from the 1000 Genomes Project and then performed a genome-wide association study (Figure 1). Because the downstream effect of coding region variants is largely considered more deleterious, we focused only on protein-coding variants in the analysis. A total of 79,696 protein-coding SNPs, which are shared among all samples, were tested for association with ERM, and the results are summarized in Figure 1. The non-synonymous variant of rs115958260 (located in the SLC44A3 gene, exon13:c.T1429C:p.C477R) and both synonymous variants of rs11570544 (located in the CDC27 gene) and rs1340610489 (located in AHNAK2) were associated with ERM at the genome-wide significance threshold ($p < 5 \times 10^{-8}$) (Figure 1C). These two synonymous variants were removed from further analysis based on our variant filter criteria (Section 2) and low functional scores of CADD (Combined Annotation Dependent Depletion) [33] and Fathmm-XF (Functional Analysis Through Hidden Markov Models with an eXtended Feature set) [34] (Figure 1D).

Detailed analyses of minor allele frequencies (MAF) of rs115958260 demonstrated at least tenfold difference between the cases and controls, both from the enrolled subjects and from presumably healthy SCT individuals from the 1000 Genomes Project (Figure 1D). The overall MAF of this variant in the African-ancestry population is also low (1–2%) in both the 1000 Genomes and gnomAD databases (Supplementary Figure). Functional annotation with in silico scores showed both high CADD and Fathmm-XF integrative scores for rs115958260: 26.4 and 21.53, respectively (Figure 1D and Supplementary Figure), which suggests potentially deleterious effects of the variant. We further screened the remaining 7 SCT/ERM cases, for which whole-genome sequencing was not performed, and found that one case also carried rs115958260. Next, we searched for tissue-specific expression of SLC44A3 protein by exploring data from the Human Protein Atlas project [35]. As illustrated in Figure 1E, high protein expression scores in both skeletal muscle and kidney tissues, which are frequently damaged in ERM, suggests that the SLC44A3 may play an important role in transport of inorganic cations/anions (choline-like transporter) during exercise and muscle contraction.
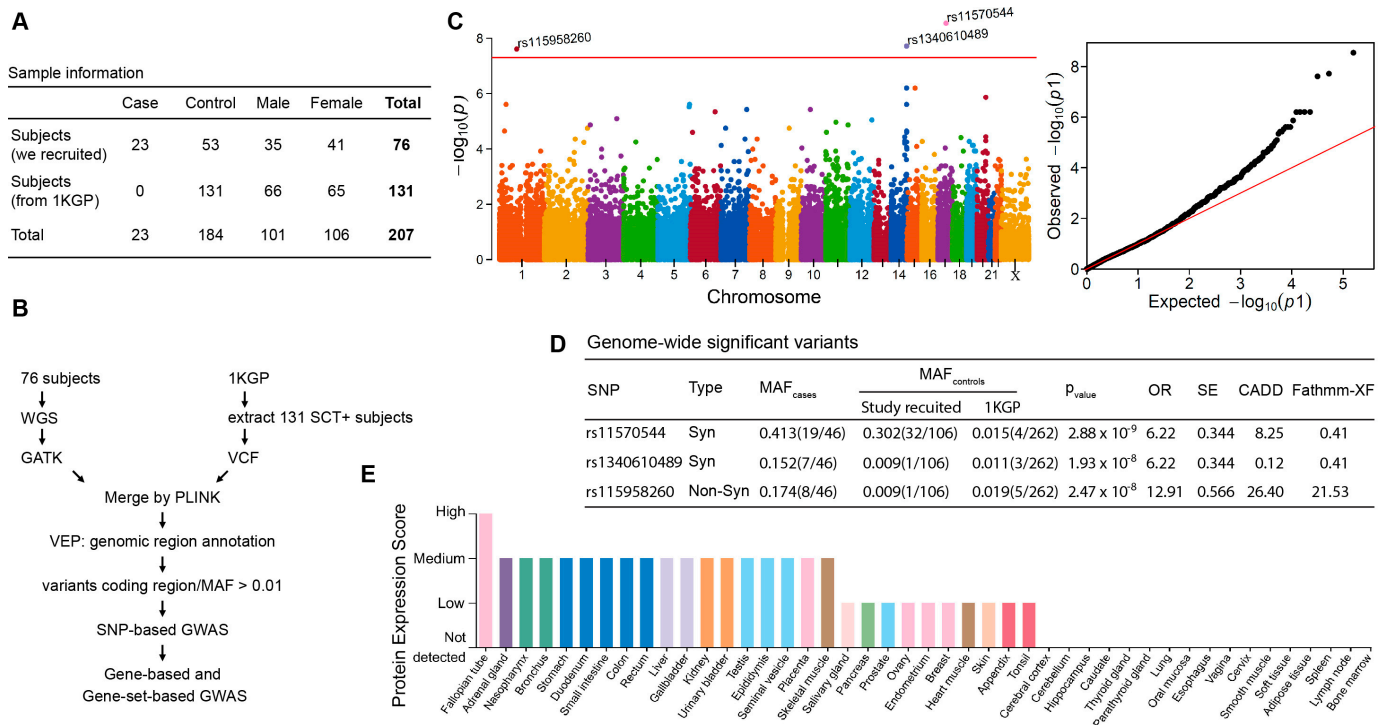
Figure 1. Summary of genome-wide association study of exertion-related rhabdomyolysis in subjects with sickle cell trait (SCT). (**A**) Sample information from enrolled subjects and African-ancestry SCT-positive subjects extracted from the 1000 Genomes Project (1KGP). (**B**) Workflow of data analysis. (**C**) SNP-based Manhattan plot (left, $p < 5 \times 10^{-8}$) and Q–Q plot (right). The *y*-axis of the former shows the $-\log10$ *p*-values for SNP in the GWAS. The horizontal red line represents the threshold of genome-wide significance ($p = 5 \times 10^{-8}$). (**D**) The minor allele frequency (MAF) of variants in the case and control groups (number of effect alleles vs. reference alleles in parentheses). Syn:synonymous; Non-Syn: nonsynonymous; CADD: Combined Annotation Dependent Depletion Score (phred); Fathmm-XF: Functional Analysis Through Hidden Markov Models with an eXtended Feature set. (**E**) Tissue-specific patterns of protein expression of SLC44A3 from the Human Protein Atlas project [35].

3.3. Gene-Based and Gene-Set-Based GWAS Analyses Show SLC44A3 Significantly Associated with ERM in SCT Carriers

Next, we performed gene-based genome-wide association analysis by using a multiple linear principal components regression model (MAGMA) [32]. MAGMA analysis demonstrated again that SLC44A3 is significantly ($p = 1.84 \times 10^{-6}$) associated with ERM, but not for CDC27 (the Bonferroni correction *p*-value = 3.089×10^{-6}; Figure 2). We further performed gene-set-based genome-wide analysis by using the Reactome pathway database [36]. Solute carrier (SLC)-mediated membrane transport was among the top 5 of 15 pathways (Table 3), which suggests that abnormal function of small molecular transport may serve a role in the development of ERM.

Table 3. Top 15 pathways from gene-set-based genome-wide association analysis using MAGMA with the Reactome database.

	Pathways	Genes	β	SE	*p*-Value
1	SYNTHESIS_OF_PIPS_AT_THE_EARLY_ENDOSOME_MEMBRANE	12	0.923	0.268	0.0003
2	REGULATION_OF_IFNG_SIGNALING	9	0.978	0.308	0.0007
3	**SLC-MEDIATED_TRANSMEMBRANE_TRANSPORT**	**200**	**0.216**	**0.069**	**0.0009**
4	SYNTHESIS_OF_PIPS_AT_THE_GOLGI_MEMBRANE	14	0.704	0.228	0.0010
5	INTERFERON_γ_SIGNALING	43	0.438	0.156	0.0025
6	**TRANSPORT_OF_VITAMINS,_NUCLEOSIDES,_AND_RELATED_MOLECULES**	**22**	**0.592**	**0.213**	**0.0028**
7	PI_METABOLISM	42	0.381	0.142	0.0036

Table 3. *Cont.*

|	Pathways	Genes	β	SE	*p*-Value
8	**TRANSPORT_OF_NUCLEOTIDE_SUGARS**	**5**	**1.136**	**0.426**	**0.0039**
9	LECTIN_PATHWAY_OF_COMPLEMENT_ACTIVATION	5	1.156	0.444	0.0047
10	GRB2_EVENTS_IN_ERBB2_SIGNALING	16	0.611	0.248	0.0069
11	SYNTHESIS_OF_PIPS_AT_THE_LATE_ENDOSOME_MEMBRANE	9	0.750	0.311	0.0080
12	**TRANSPORT_OF_INORGANIC_CATIONS/ANIONS_AND_AMINO_ACIDS/OLIGOPEPTIDES**	**78**	**0.264**	**0.110**	**0.0083**
13	POST_NMDA_RECEPTOR_ACTIVATION_EVENTS	21	0.495	0.207	0.0085
14	KSRP_DESTABILIZES_MRNA	9	0.792	0.339	0.0097
15	**TRANSMEMBRANE_TRANSPORT_OF_SMALL_MOLECULES**	**406**	**0.115**	**0.049**	**0.0098**
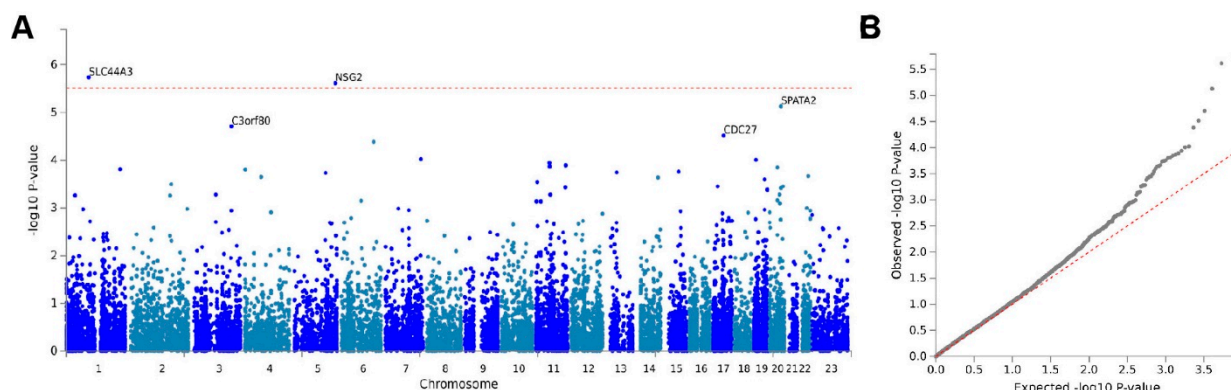
Figure 2. Gene-based genome-wide association analysis. (**A**) Manhattan plot of the gene-based genome-wide association analysis by using a multiple linear principal components regression model (MAGMA). Input SNPs were mapped to 16,187 protein-coding genes. Genome-wide significance (red dashed line) was defined at $p = 0.05/16187 = 3.089 \times 10^{-6}$. (**B**) Gene-based Q–Q plot.

4. Discussion

Although largely considered a benign carrier state, SCT is of particular interest to all military and civilian providers due to its association with exertion-related injuries such as ERM, exertional collapse associated with SCT, and sudden death [37–40]. Fundamental questions addressing etiology, as well as the potential variable risk to various health co-morbidities, such as exertional injuries among those with SCT, remain unanswered. An increasing number of studies, including this one, have shown that certain genetic risk factors contribute to an individual's variability in the development of ERM [21,41]. However, no studies have been conducted to determine genetic risk factors other than SCT that might predict which SCT carriers are at greatest risk for developing ERM. In the current study, we used whole-genome sequencing to perform GWAS to uncover any genetic factors linked to this association. Both SNP- and gene-based GWAS analyses revealed significant association of the SLC44A3 gene variant rs115958260 with ERM in the SCT-positive African American population. We also characterized shared clinical features of ERM in the series of SCT individuals. SCT carriers with ERM presented with features of myotoxicity symptoms of significant muscle involvement dominated by muscle weakness and severe pain and substantial increase in serum CK, with a mean value of 50,480 U/L. A distinctive feature of SCT individuals with ERM was exertional collapse, which was reported in 53.3% of the cases presented in this study. An important factor for the development of ERM was duration and frequency of strenuous physical activity in cases compared to controls, which suggests that a combination of vigorous exercise and a genetic predisposing factor is involved in ERM.

The SLC44A3 gene is one of the SLC44 family members (SLC44A1-5), also called the choline-like transporter family [42]. Choline is important for the formation of membrane phospholipids and the neurotransmitter acetylcholine [43]. A choline deficiency caused by mutations in the SCL44A1 gene, a close homolog of SLC44A3, is known to cause neurodegenerative disease, with abnormal regulation of the genes involved in mitochondrial fatty

acid transport and β-oxidation, and evidence of disturbances in fatty acid homeostasis [43]. Likewise, a dietary choline deficiency has been associated with muscle damage [44,45]. Further larger-scale genetic association between ERM and SCT, as well as the functional role of the SLC44A3 gene in the development of ERM, will be critically important.

ERM is a complex syndrome with multiple environmental (unaccustomed exercise, fasting, and hypokalemia) and genetic contributing factors. Abnormalities in the genes encoding proteins in the glycolytic and fatty acid oxidation pathways, calcium dysregulation, and mitochondrial respiratory chains are thought to be involved in the pathogenesis of ERM [12,14,19]. No mitochondrial DNA analysis was performed in this study. However, these inherited disorders are rare and likely not responsible for most cases of ERM. Also, our previous study showed the same hazard ratio as being a SCT carrier. However, the percentage of participants in the present study with smoking history (as shown in Table 1) is small, with no difference between the case and control groups. It is unlikely that smoking status impacted ERM development in this study. Common genetic variants in angiotensin 1 converting enzyme (ACE1), muscle isoform of creatine kinase (CKM), and fast-fiber specific α-actinin 3 (ACTN3) genes have been associated with ERM or muscle symptoms related to ERM. However, these proposed variants of ERM-related genes were not confirmed by this study, similar to the results of other association studies [46,47]. Moreover, it is worth noting that glucose-6-phosphate dehydrogenase (G6PD) deficiency is prevalent among male African Americans, ranging from 10% to 14% [48]. G6PD deficiency in concert with SCT likely precipitated rhabdomyolysis, albeit very rarely [49]. Our current study did not reveal any pathogenic mutation associated with G6PD deficiency or accumulation of deleterious G6PD variants. As a result, our gene-level association analysis found no significant ($p = 0.68$) association between G6PD and ERM.

The percentage of HbS, which ranges from 25% to 45% [50], may affect the hematological and clinical characteristics of SCT, including ERM [51]. However, the Kennedy et al. study suggested that HbS levels in SCT were very unlikely to have any clinical significance [52]. Upon reviewing the literature, we noticed that the majority of studies compared SCT carriers with healthy controls (SCT-negative) without addressing the variability in HbS percentage among SCT individuals [51]. Nevertheless, we cannot exclude the effect of HbS percentage on ERM development without examining it in our study. Additionally, the influence of α-thalassemia or α-globin gene triplication on ERM cannot be excluded since these mutations have been shown to affect HbS concentration in the African American population [53].

The association between ERM and SCT has mostly been described in clinical case studies and or epidemiological studies of SCT-associated co-morbidities. To date, there are no GWAS studies on ERM. Clinically significant ERM is rare, and collection of large study cases remains challenging. Another challenge is the phenotypic ascertainment of individuals. The clinical presentation in most cases is nonspecific, but usually includes elevated CK, dark urine secondary to myoglobinuria, and various other health complications. To date, most of the available data regarding the genetic basis of ERM are based on individual gene-specific case reports, which makes it difficult to conclude whether the variants identified are pathogenic.

Finally, there were no GWAS summary statistics results of ERM/RM in the public databases, such as the NIH GWAS catalog and the UK BioBank. These challenges are also relevant to the limitations of the current study, where sample size was not sufficiently powered for genome-wide variants-based GWAS. This led us to focus only on protein-coding region variants, although filtration of other types of variants certainly impacted the gene-based GWAS. This association study does not confirm the genomic variants previously identified as risk factors for susceptibility to ERM. However, the study design could not provide complete confidence in the control groups not including some individuals with a propensity for ERM or simply having never been challenged with extreme exertion. As a result, there is a possible bias toward "false negatives" in the control groups, and our study could not completely refute those other genomic variants. In addition, our study

cannot exclude that rare variants (MAF < 1%) might contribute to ERM development due to the limited sample size. Nevertheless, we clinically characterized African American SCT individuals with ERM and used an unbiased WGS approach to gain a deeper understanding of the association between SCT and exertion-related injuries.

5. Conclusions

The study identified a rare, non-synonymous rs115958260 variant in the SLC44A3 gene, which is significantly associated with ERM in SCT-positive subjects of African American descent. Our study suggests that a combination of vigorous exercise and a genetic predisposing factor is involved in ERM.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes15040408/s1, Supplementary Figure Functional annotation results of rs115958260 variant.

Author Contributions: Conceptualization, M.R., N.S., F.G.O. and P.A.D.; Methodology, D.H.; Formal analysis, M.R. and N.S.; Investigation, M.R., N.S., O.M., D.B.E. and F.G.O.; Data curation, M.R., N.S., D.H., X.Z., M.D.W. and C.L.D.; Writing—original draft, M.R. and N.S.; Writing—review & editing, M.D.W., C.L.D., F.G.O. and P.A.D.; Supervision, F.G.O. and P.A.D.; Funding acquisition, P.A.D. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Uniformed Services University (protocol: MEM-91-3139-S1, date of approval: 12 December 2016).

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

Data Availability Statement: All relevant data are within the manuscript and its Supplementary Materials. The data isolated from the 1000 Genomes Project are publicly available from https://www.internationalgenome.org/data, and other genotype data will be available from https://www.ebi.ac.uk/eva (after 30 September 2025) (submission #701876).

Conflicts of Interest: The authors declare no conflicts of interest. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Disclaimers and Disclosures: The opinions and assertions expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University or the Department of Defense. The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views, opinions, or policies of The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. Mention of trade names, commercial products, or organizations does not imply endorsement by the US Government.

References

1. Xu, J.Z.; Thein, S.L. The carrier state for sickle cell disease is not completely harmless. *Haematologica* **2019**, *104*, 1106–1111. [CrossRef] [PubMed]
2. Nelson, D.A.; Deuster, P.A.; Carter, R., 3rd; Hill, O.T.; Wolcott, V.L.; Kurina, L.M. Sickle Cell Trait, Rhabdomyolysis, and Mortality among U.S. Army Soldiers. *N. Engl. J. Med.* **2016**, *375*, 435–442. [CrossRef]
3. Ojodu, J.; Hulihan, M.M.; Pope, S.N.; Grant, A.M.; Centers for Disease, C.; Centers for Disease Control and Prevention. Incidence of sickle cell trait—United States, 2010. *MMWR Morb. Mortal. Wkly. Rep.* **2014**, *63*, 1155–1158.
4. Naik, R.P.; Smith-Whitley, K.; Hassell, K.L.; Umeh, N.I.; de Montalembert, M.; Sahota, P.; Haywood, C., Jr.; Jenkins, J.; Lloyd-Puryear, M.A.; Joiner, C.H.; et al. Clinical Outcomes Associated with Sickle Cell Trait: A Systematic Review. *Ann. Intern. Med.* **2018**, *169*, 619–627. [CrossRef] [PubMed]
5. Hulsizer, J.; Resurreccion, W.K.; Shi, Z.; Wei, J.; Ladson-Gary, S.; Zheng, S.L.; Helfand, B.T.; Billings, L.; Caplan, M.S.; Xu, J. Sickle Cell Trait and Risk for Common Diseases: Evidence from the UK Biobank. *Am. J. Med.* **2022**, *135*, e279–e287. [CrossRef]
6. Kerle, K.K.; Nishimura, K.D. Exertional collapse and sudden death associated with sickle cell trait. *Mil. Med.* **1996**, *161*, 766–767. [CrossRef] [PubMed]

7. Harmon, K.G.; Drezner, J.A.; Klossner, D.; Asif, I.M. Sickle cell trait associated with a RR of death of 37 times in national collegiate athletic association football athletes: A database with 2 million athlete-years as the denominator. *Br. J. Sports Med.* **2012**, *46*, 325–330. [CrossRef] [PubMed]

8. O'Connor, F.G.; Franzos, M.A.; Nye, N.S.; Nelson, D.A.; Shell, D.; Voss, J.D.; Anderson, S.A.; Coleman, N.J.; Thompson, A.A.; Harmon, K.G.; et al. Summit on Exercise Collapse Associated with Sickle Cell Trait: Finding the "Way Ahead". *Optom. Vis. Sci.* **2021**, *20*, 47–56. [CrossRef] [PubMed]

9. O'Connnor, F.G.; Deuster, P.A. Rhabdomyolysis. In *Goldman-Cecil Medicine*, 27th ed.; Goldman, L., Cooney, K.A., Eds.; Elsevier Inc: New York, NY, USA, 2023; pp. 714–717.

10. Chavez, L.O.; Leon, M.; Einav, S.; Varon, J. Beyond muscle destruction: A systematic review of rhabdomyolysis for clinical practice. *Crit. Care* **2016**, *20*, 135. [CrossRef]

11. Nance, J.R.; Mammen, A.L. Diagnostic evaluation of rhabdomyolysis. *Muscle Nerve* **2015**, *51*, 793–810. [CrossRef]

12. Kruijt, N.; van den Bersselaar, L.R.; Kamsteeg, E.J.; Verbeeck, W.; Snoeck, M.M.J.; Everaerd, D.S.; Abdo, W.F.; Jansen, D.R.M.; Erasmus, C.E.; Jungbluth, H.; et al. The etiology of rhabdomyolysis: An interaction between genetic susceptibility and external triggers. *Eur. J. Neurol.* **2021**, *28*, 647–659. [CrossRef]

13. Long, B.; Koyfman, A.; Gottlieb, M. An evidence-based narrative review of the emergency department evaluation and management of rhabdomyolysis. *Am. J. Emerg. Med.* **2019**, *37*, 518–523. [CrossRef] [PubMed]

14. Scalco, R.S.; Gardiner, A.R.; Pitceathly, R.D.; Zanoteli, E.; Becker, J.; Holton, J.L.; Houlden, H.; Jungbluth, H.; Quinlivan, R. Rhabdomyolysis: A genetic perspective. *Orphanet J. Rare Dis.* **2015**, *10*, 51. [CrossRef]

15. Marciante, K.D.; Durda, J.P.; Heckbert, S.R.; Lumley, T.; Rice, K.; McKnight, B.; Totah, R.A.; Tamraz, B.; Kroetz, D.L.; Fukushima, H.; et al. Cerivastatin, genetic variants, and the risk of rhabdomyolysis. *Pharm. Genom.* **2011**, *21*, 280–288. [CrossRef] [PubMed]

16. Vivante, A.; Ityel, H.; Pode-Shakked, B.; Chen, J.; Shril, S.; van der Ven, A.T.; Mann, N.; Schmidt, J.M.; Segel, R.; Aran, A.; et al. Exome sequencing in Jewish and Arab patients with rhabdomyolysis reveals single-gene etiology in 43% of cases. *Pediatr. Nephrol.* **2017**, *32*, 2273–2282. [CrossRef]

17. Wu, L.; Brady, L.; Shoffner, J.; Tarnopolsky, M.A. Next-Generation Sequencing to Diagnose Muscular Dystrophy, Rhabdomyolysis, and HyperCKemia. *Can. J. Neurol. Sci.* **2018**, *45*, 262–268. [CrossRef] [PubMed]

18. Daniele, D.O.; Murray, J. Update: Exertional rhabdomyolysis, active component, U.S. Armed Forces, 2017–2021. *MSMR* **2022**, *29*, 15–20. [PubMed]

19. Carneiro, A.; Viana-Gomes, D.; Macedo-Da-Silva, J.; Lima, G.H.O.; Mitri, S.; Alves, S.R.; Kolliari-Turner, A.; Zanoteli, E.; Neto, F.R.A.; Palmisano, G.; et al. Risk factors and future directions for preventing and diagnosing exertional rhabdomyolysis. *Neuromuscul. Disord.* **2021**, *31*, 583–595. [CrossRef] [PubMed]

20. Boden, B.P.; Isaacs, D.J.; Ahmed, A.E.; Anderson, S.A. Epidemiology of Exertional Rhabdomyolysis in the United States: Analysis of NEISS Database 2000 to 2019. *Physician Sportsmed.* **2021**, *50*, 486–493. [CrossRef]

21. Deuster, P.A.; Contreras-Sesvold, C.L.; O'Connor, F.G.; Campbell, W.W.; Kenney, K.; Capacchione, J.F.; Landau, M.E.; Muldoon, S.M.; Rushing, E.J.; Heled, Y. Genetic polymorphisms associated with exertional rhabdomyolysis. *Eur. J. Appl. Physiol.* **2013**, *113*, 1997–2004. [CrossRef]

22. Norton, E.M.; Mickelson, J.R.; Binns, M.M.; Blott, S.C.; Caputo, P.; Isgren, C.M.; McCoy, A.M.; Moore, A.; Piercy, R.J.; Swinburne, J.E.; et al. Heritability of Recurrent Exertional Rhabdomyolysis in Standardbred and Thoroughbred Racehorses Derived from SNP Genotyping Data. *J. Hered.* **2016**, *107*, 537–543. [CrossRef] [PubMed]

23. Ware, S.M.; Bhatnagar, S.; Dexheimer, P.J.; Wilkinson, J.D.; Sridhar, A.; Fan, X.; Shen, Y.; Tariq, M.; Schubert, J.A.; Colan, S.D.; et al. The genetic architecture of pediatric cardiomyopathy. *Am. J. Hum. Genet.* **2022**, *109*, 282–298. [CrossRef] [PubMed]

24. Gorsi, B.; Hernandez, E.; Moore, M.B.; Moriwaki, M.; Chow, C.Y.; Coelho, E.; Taylor, E.; Lu, C.; Walker, A.; Touraine, P.; et al. Causal and Candidate Gene Variants in a Large Cohort of Women with Primary Ovarian Insufficiency. *J. Clin. Endocrinol. Metab.* **2022**, *107*, 685–714. [CrossRef] [PubMed]

25. Soltis, A.R.; Bateman, N.W.; Liu, J.; Nguyen, T.; Franks, T.J.; Zhang, X.; Dalgard, C.L.; Viollet, C.; Somiari, S.; Yan, C.; et al. Proteogenomic analysis of lung adenocarcinoma reveals tumor heterogeneity, survival determinants, and therapeutically relevant pathways. *Cell Rep. Med.* **2022**, *3*, 100819. [CrossRef] [PubMed]

26. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From fastq data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 11.10.1–11.10.33. [CrossRef] [PubMed]

27. Hunt, S.E.; Moore, B.; Amode, R.M.; Armean, I.M.; Lemos, D.; Mushtaq, A.; Parton, A.; Schuilenburg, H.; Szpak, M.; Thormann, A.; et al. Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor—A tutorial. *Hum. Mutat.* **2022**, *43*, 986–997. [CrossRef] [PubMed]

28. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]

29. Manichaikul, A.; Mychaleckyj, J.C.; Rich, S.S.; Daly, K.; Sale, M.; Chen, W.-M. Robust relationship inference in genome-wide association studies. *Bioinformatics* **2010**, *26*, 2867–2873. [CrossRef] [PubMed]

30. Staples, J.; Qiao, D.; Cho, M.H.; Silverman, E.K.; University of Washington Center for Mendelian; Nickerson, D.A.; Below, J.E. PRIMUS: Rapid Reconstruction of Pedigrees from Genome-wide Estimates of Identity by Descent. *Am. J. Hum. Genet.* **2014**, *95*, 553–564. [CrossRef]

31. Marees, A.T.; de Kluiver, H.; Stringer, S.; Vorspan, F.; Curis, E.; Marie-Claire, C.; Derks, E.M. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int. J. Methods Psychiatr. Res.* **2018**, *27*, e1608. [CrossRef]

32. de Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **2015**, *11*, e1004219. [CrossRef] [PubMed]

33. Kircher, M.; Witten, D.M.; Jain, P.; O'Roak, B.J.; Cooper, G.M.; Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **2014**, *46*, 310–315. [CrossRef] [PubMed]

34. Rogers, M.F.; Shihab, H.A.; Mort, M.; Cooper, D.N.; Gaunt, T.R.; Campbell, C. FATHMM-XF: Accurate prediction of pathogenic point mutations via extended features. *Bioinformatics* **2017**, *34*, 511–513. [CrossRef] [PubMed]

35. Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Proteomics. Tissue-Based Map of the Human Proteome. *Science* **2015**, *347*, 1260419. [CrossRef] [PubMed]

36. Gillespie, M.; Jassal, B.; Stephan, R.; Milacic, M.; Rothfels, K.; Senff-Ribeiro, A.; Griss, J.; Sevilla, C.; Matthews, L.; Gong, C.; et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **2022**, *50*, D687–D692. [CrossRef] [PubMed]

37. O'Connor, F.G.; Bergeron, M.F.; Cantrell, J.; Connes, P.; Harmon, K.G.; Ivy, E.; Kark, J.; Klossner, D.; Lisman, P.; Meyers, B.K.; et al. ACSM and CHAMP summit on sickle cell trait: Mitigating risks for warfighters and athletes. *Med. Sci. Sports Exerc.* **2012**, *44*, 2045–2056. [CrossRef] [PubMed]

38. Anzalone, M.L.; Green, V.S.; Buja, M.; Sanchez, L.A.; Harrykissoon, R.I.; Eichner, E.R. Sickle cell trait and fatal rhabdomyolysis in football training: A case study. *Med. Sci. Sports Exerc.* **2010**, *42*, 3–7. [CrossRef] [PubMed]

39. Makaryus, J.N.; Catanzaro, J.N.; Katona, K.C. Exertional rhabdomyolysis and renal failure in patients with sickle cell trait: Is it time to change our approach? *Hematology* **2007**, *12*, 349–352. [CrossRef] [PubMed]

40. Tsaras, G.; Owusu-Ansah, A.; Boateng, F.O.; Amoateng-Adjepong, Y. Complications associated with sickle cell trait: A brief narrative review. *Am. J. Med.* **2009**, *122*, 507–512. [CrossRef]

41. Sambuughin, N.; Capacchione, J.; Blokhin, A.; Bayarsaikhan, M.; Bina, S.; Muldoon, S. The ryanodine receptor type 1 gene variants in African American men with exertional rhabdomyolysis and malignant hyperthermia susceptibility. *Clin. Genet.* **2009**, *76*, 564–568. [CrossRef]

42. Traiffort, E.; O'regan, S.; Ruat, M. The choline transporter-like family SLC44: Properties and roles in human diseases. *Mol. Asp. Med.* **2013**, *34*, 646–654. [CrossRef] [PubMed]

43. Fagerberg, C.R.; Taylor, A.; Distelmaier, F.; Schröder, H.D.; Kibæk, M.; Wieczorek, D.; Tarnopolsky, M.; Brady, L.; Larsen, M.J.; A Jamra, R.; et al. Choline transporter-like 1 deficiency causes a new type of childhood-onset neurodegeneration. *Brain* **2020**, *143*, 94–111. [CrossRef] [PubMed]

44. da Costa, K.-A.; Badea, M.; Fischer, L.M.; Zeisel, S.H. Elevated serum creatine phosphokinase in choline-deficient humans: Mechanistic studies in C2C12 mouse myoblasts. *Am. J. Clin. Nutr.* **2004**, *80*, 163–170. [CrossRef] [PubMed]

45. Wallace, T.C.; Blusztajn, J.K.; Caudill, M.A.; Klatt, K.C.; Natker, E.; Zeisel, S.H.; Zelman, K.M. Choline: The Underconsumed and Underappreciated Essential Nutrient. *Nutr. Today* **2018**, *53*, 240–253. [CrossRef] [PubMed]

46. Ma, F.; Yang, Y.; Li, X.; Zhou, F.; Gao, C.; Li, M.; Gao, L. The association of sport performance with ACE and ACTN3 genetic polymorphisms: A systematic review and meta-analysis. *PLoS ONE* **2013**, *8*, e54685. [CrossRef] [PubMed]

47. Sonna, L.A.; Sharp, M.A.; Knapik, J.J.; Cullivan, M.; Angel, K.C.; Patton, J.F.; Lilly, C.M. Angiotensin-converting enzyme genotype and physical performance during US Army basic training. *J. Appl. Physiol.* **2001**, *91*, 1355–1363. [CrossRef] [PubMed]

48. Gómez-Manzo, S.; Marcial-Quino, J.; Vanoye-Carlo, A.; Serrano-Posada, H.; Ortega-Cuellar, D.; González-Valdez, A.; Castillo-Rodríguez, R.A.; Hernández-Ochoa, B.; Sierra-Palacios, E.; Rodríguez-Bustamante, E.; et al. Glucose-6-Phosphate Dehydrogenase: Update and Analysis of New Mutations around the World. *Int. J. Mol. Sci.* **2016**, *17*, 2069. [CrossRef] [PubMed]

49. Eziokwu, A.S.; Angelini, D. New Diagnosis of G6PD Deficiency Presenting as Severe Rhabdomyolysis. *Cureus* **2018**, *10*, e2387. [CrossRef] [PubMed]

50. Khaled, S.A.; Ahmed, H.A.; Elbadry, M.I.; NasrEldin, E.; Hassany, S.M.; Ahmed, S.A. Hematological, Biochemical Properties, and Clinical Correlates of Hemoglobin S Variant Disorder: A New Insight into Sickle Cell Trait. *J. Hematol.* **2022**, *11*, 92–108. [CrossRef]

51. Tripette, J.; Hardy-Dessources, M.-D.; Romana, M.; Hue, O.; Diaw, M.; Samb, A.; Diop, S.; Connes, P. Exercise-related complications in sickle cell trait. *Clin. Hemorheol. Microcirc.* **2013**, *55*, 29–37. [CrossRef]

52. Kennedy, A.P.; Walsh, D.A.; Nicholson, R.; Adams, J.G., 3rd; Steinberg, M.H. Influence of HbS levels upon the hematological and clinical characteristics of sickle cell trait. *Am. J. Hematol.* **1986**, *22*, 51–54. [CrossRef] [PubMed]

53. Pinto, V.M.; De Franceschi, L.; Gianesin, B.; Gigante, A.; Graziadei, G.; Lombardini, L.; Palazzi, G.; Quota, A.; Russo, R.; Sainati, L.; et al. Management of the Sickle Cell Trait: An Opinion by Expert Panel Members. *J. Clin. Med.* **2023**, *12*, 3441. [CrossRef] [PubMed]