Original Research *Highlight article*

Whole-exome sequencing indicates *FLG2* variant associated with leg ulcers in Brazilian sickle cell anemia patients

Gabriela Queila de Carvalho-Siqueira[1], Galina Ananina[1], Bruno Batista de Souza[1], Murilo Guimarães Borges[2], Mirta Tomie Ito[1], Sueli Matilde da Silva-Costa[1], Igor de Farias Domingos[3], Diego Arruda Falcão[3], Iscia Lopes-Cendes[2], Marcos André Cavalcanti Bezerra[3], Aderson da Silva Araújo[4], Antônio Roberto Lucena-Araújo[3], Marilda de Souza Gonçalves[5], Sara Teresinha Olalla Saad[6], Fernando Ferreira Costa[6] and Mônica Barbosa de Melo[1]

[1]Center for Molecular Biology and Genetic Engineering (CBMEG), University of Campinas, Campinas, SP 13083-875, Brazil; [2]Department of Medical Genetics and Genome Medicine, Faculty of Medical Sciences, University of Campinas, Campinas, SP 13083-887, Brazil; [3]Genetics Postgraduate Program, Federal University of Pernambuco, Recife, PE 50670-901, Brazil; [4]Hematology and Hemotherapy Foundation of Pernambuco, Recife, PE 52011-000, Brazil; [5]Gonçalo Moniz Institute (Fiocruz), Salvador, BA 40296-710, Brazil; [6]Hematology and Hemotherapy Center, University of Campinas, Campinas, SP 13083-878, Brazil
Corresponding author: Mônica Barbosa de Melo. Email: melomb@uol.com.br

Impact statement

To our knowledge, the present study is the first to use whole-exome sequencing based on extreme phenotypes to identify new candidate genes associated with leg ulcers in sickle cell anemia patients. There are few studies about this complication; the pathogenesis remains complex and has yet to be fully elucidated. We identified interesting associations in genes never related with this complication to our knowledge, especially the variant in the *FLG2* gene. The knowledge of variants related with leg ulcer in sickle cell anemia may lead to a better comprehension of the disease's etiology, allowing prevention and early treatment options in risk genotypes while improving quality of life for these patients.

Abstract

Although sickle cell anemia results from homozygosity for a single mutation at position 7 of the β-globin chain, the clinical aspects of this condition are very heterogeneous. Complications include leg ulcers, which have a negative impact on patients' quality of life and are related to the severity of the disease. Nevertheless, the complex pathogenesis of this complication has yet to be elucidated. To identify novel genes associated with leg ulcers in sickle cell anemia, we performed whole-exome sequencing of extreme phenotypes in a sample of Brazilian sickle cell anemia patients and validated our findings in another sample. Our discovery cohort consisted of 40 unrelated sickle cell anemia patients selected based on extreme phenotypes: 20 patients without leg ulcers, aged from 40 to 61 years, and 20 with chronic leg ulcers. DNA was extracted from peripheral blood leukocytes and used for whole-exome sequencing. After the bioinformatics analysis, eight variants were selected for validation by Sanger sequencing and TaqMan® genotyping in 293 sickle cell anemia patients (153 without leg ulcers) from two different locations in Brazil. After the validation, Fisher's exact test revealed a statistically significant difference in a stop codon variant (rs12568784 G/T) in the *FLG2* gene between the GT and GG genotypes ($P = 0.035$). We highlight the importance of rs12568784 in leg ulcer development as this variant of the *FLG2* gene results in impairment of the skin barrier, predisposing the individual to inflammation and infection. Additionally, we suggest that the remaining seven variants and the genes in which they occur could be strong candidates for leg ulcers in sickle cell anemia.

Keywords: Sickle cell anemia, whole-exome sequencing, leg ulcer, association study, complex disease, variants

Experimental Biology and Medicine **2019; 244: 932–939. DOI: 10.1177/1535370219849592**

Introduction

Sickle cell anemia (SCA) is an inherited blood disease which results from a homozygous point mutation of the β-globin chain at position 7, leading to substitution of glutamic acid for valine (Human Genome Variation Society, HGVS, <https://www.ncbi.nlm.nih.gov/projects/SNP/

snp_ref.cgi?rs=334>).[1] Vaso-occlusion events and hemolytic anemia occur as a result of this amino acid substitution[2] and can affect different organs. Although SCA is a monogenic disorder, individuals with the condition (homozygous for hemoglobin $\beta^S\beta^S$) show a wide phenotypic variability.

The causal relationship between sickle cell disease (SCD, which includes SCA) and leg ulcers was not established until the publication of a paper by Cummer and LaRocco in 1940.[3] Ulcerations of the skin and underlying tissues may occur in sickle cell patients as well as in those with other hematological disorders and affect the medial and lateral regions of the ankle. The pathogenesis of chronic leg ulcers in sickle cell patients is complex and has yet to be fully elucidated. Factors predisposing to chronic ulceration in these patients include poor skin perfusion, increased local edema due to venous incompetence or abnormal autonomic vascular control, microvascular thrombosis, reduced oxygenation, reduced bioavailability of nitric oxide, and mild cutaneous trauma.[4]

Ulceration is more common in patients with the $\beta^S\beta^S$ genotype than in those who are heterozygous for Hb S.[5,6] Its incidence is lower in patients with low lactate dehydrogenase levels,[7] higher hemoglobin concentrations or high levels of fetal hemoglobin.[8] Age-adjusted comparisons showed that the frequency of leg ulcers in SCA patients with α-thalassemia trait and HbSC patients was lower than in homozygotes for Hb S, suggesting that the probability of having leg ulcers could be associated with the intensity of hemolysis.[9] Similarly, other studies have observed an association between hemolytic markers and leg ulcers.[5,9–11] Cutaneous leg ulcers are also seen in other forms of hemolytic anemia without sickling, including hereditary spherocytosis, pyruvate kinase deficiency, and thalassemia.[12–22]

Hydroxyurea is one of the few FDA-approved therapies to treat SCD. This therapy is used to induce the maximal fetal hemoglobin response through different molecular pathways, reducing the severity of the disease in many patients.[23] Limited evidence suggest that hydroxyurea is not associated with leg ulcers in SCD.[24,25] Interestingly, Minniti *et al.*,[26] in a study comprising 505 patients at the National Institutes of Health, did not observe differences in ulcer prevalence between SCD patients that were taking or not taking hydroxyurea. Besides hydroxyurea therapy, frequent transfusions of red blood cells may greatly decrease disease's severity. This therapy reduces the Hb S concentration and increases hemoglobin levels improving tissue perfusion.

The prevalence of leg ulcers depends on geographic location and environmental factors. In the United States, for example, about 2.5% of people with SCD have leg ulcers, while in Jamaica, the corresponding figure is over 40%.[4] Traumas have been reported to play an important role in the development of these lesions, which are predominantly a result of mosquito bites, a common occurrence in areas with a tropical climate.[7] In Brazil, the prevalence of leg ulcers among adult patients with SCD varies from 23.4% to 50.0%,[27] making this complication an important public health problem.

Studies with twins have shown strong heritability of venous functions in humans,[28] suggesting that genetic modifiers that modulate the risk of vaso-occlusion play a part in leg ulceration in SCA patients. However, few studies have examined the association between genetic factors and leg ulcers.[9,29,30] In one such study, a strong relationship between this complication and single nucleotide variants (SNVs) in candidate genes that could affect sickle vaso-occlusion was reported. Associations were observed between leg ulcers in SCA and SNVs in Klotho (*KL*), TEK receptor tyrosine kinase (*TEK*), and genes from the *TGF-beta/BMP* pathway.[9] Additionally, Ofosu *et al*.[29] showed that HLA-35 and Cw4 genotypes were more present in patients with leg ulcers history and Vicari *et al*.[30] associated some variants in interleukin-1β and interleukin-6 genes with complications in SCA, including leg ulcers.

The exome represents about 1% of the genome, with approximately 30 million base pairs, but accounts for about 85% of the mutations identified in Mendelian diseases.[31] Family-based phenotypes and extreme phenotypes are often used as part of sampling strategies in whole-exome sequencing (WES). In an extreme phenotype study design, the frequency of the alleles that contribute to the characteristic is enriched at the end of the distribution. It should therefore be possible to identify new candidate alleles using a relatively small sample.[32] In an attempt to identify novel candidate genes associated with leg ulcers in SCA, we performed WES of extreme phenotypes in a sample of Brazilian SCA patients and validated our findings in another cohort.

Materials and methods

Patients' recruitment

Patients have been accompanied in routine appointments and interviewed regarding their leg ulcers history, data confirmed by their medical records. As a result, cases (with active or already healed leg ulcers) and controls (without a history of leg ulcers until blood collection) have been distinguished.

Selection of patients with extreme phenotypes in the discovery cohort

The discovery cohort consisted of 40 unrelated SCA patients from the Hematology and Hemotherapy Center, University of Campinas, Campinas, SP, Brazil. The enrollment period was between January 2015 and August 2016. We compared the most severe cases in the cohort (with active or already healed chronic leg ulcers) with the mild cases (without a history of leg ulcers and over 40 years of age as the appearance of this complication after the age of 30 is unusual[33]—we tried to be as stringent as possible in this case). None of them presented α-thalassemia trait. Twenty cases with leg ulcers (average age 43.55 ± 8.94 years) and 20 controls without this complication (average age 45.85 ± 6.67 years) were included at this stage.

Selection of the validation cohort

The validation cohort comprised 293 unrelated patients with SCA: 153 aged over 21 years without leg ulcers (average age 35.8 ± 8.58 years, since the development of the first-time leg ulcers in Brazilian SCA patients is usual during adolescence, before 20 years of age) and 140 with active or already healed leg ulcers (average age 37.44 ± 10.8 years). Of the 293 patients, 116 were from the Bahia Hematology and Hemotherapy Foundation (HEMOBA), Salvador, BA, Brazil, while 177 were from the Hematology and Hemotherapy Foundation of Pernambuco (HEMOPE), Recife, PE, Brazil. We did not exclude the patients with α-thalassemia trait, once the number of patients with and without this trait was not significantly different between cases and controls ($P > 0.05$).

Laboratorial parameters

Medical records of the entire discovery cohort ($n = 38$) and of the most representative part of the validation cohort (from the Hematology and Hemotherapy Foundation of Pernambuco–HEMOPE, Recife, PE, Brazil; $n = 177$) have been consulted. Lactate dehydrogenase levels, hemoglobin concentrations, and fetal hemoglobin levels have been compared between cases and controls through Mann–Whitney-Wilcoxon Test. The values were collected before the use of hydroxyurea or red blood cell transfusion therapy. We emphasize that the number of patients who took hydroxyurea in case and control groups was similar in both cohorts, and no statistically significant difference was observed ($P > 0.05$).

DNA extraction

DNA was extracted from peripheral blood. This was done with the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) for the discovery cohort and with in-house reagents by the phenol–chloroform method for the validation cohort.

Ethics statement

All patients signed a voluntary informed consent form for collection, storage, and sequencing of DNA samples. The study protocol complied with the principles of the Declaration of Helsinki and was approved by the Research Ethics Committees of the Faculty of Medical Sciences, Unicamp, HEMOBA/Fiocruz and HEMOPE (ref. no. 53001315.6.1001.5404).

Whole-exome sequencing

DNA concentrations were quantified by the fluorescence-detection method (Qubit, Life Technologies Corp., Carlsbad, CA, USA). Exome libraries were prepared for each DNA sample from the discovery cohort with Nextera Rapid Capture Exome (Illumina Inc., San Diego, CA, USA), and then sequenced (2×100 bp per read) on seven flowcell lanes on the Illumina HiSeq 2500 platform at the central high-performance technologies laboratory (LaCTAD, Campinas, SP, Brazil).

Bioinformatics analysis

Sequencing reads were mapped to the GRCh38 reference using BWA (<http://bio-bwa.sourceforge.net>)[34] followed by GATK Pipeline (<https://software.broadinstitute.org/gatk>)[34,35] which generated a gVCF file for each sample. Joint genotyping was then performed for all the samples, and a single VCF file was produced as output. VCFtools 0.1.15(<http://vcftools.sourceforge.net>)[36] was used to exclude variants with score quality <30[37–39] or read depth <10; multiallelic variants; variants present in sex chromosomes or in mitochondrial DNA; sites with >5 % missing genotypes;[40] variants with MAF <0.1; and variants not in Hardy–Weinberg equilibrium (HWE).[39] To complete the analysis, we used PLINK 1.9 (<http://pngu.mgh.harvard.edu/purcell/plink>)[41] to perform identity-by-descent analysis and Fisher's exact test. The remaining variants were extracted from the original VCF file and annotated with wANNOVAR (<http://wannovar.wglab.org>).[42] The synonymous variants were excluded, deleterious variants and variants that had not been classified by SIFT (<http://sift.jcvi.org>), PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2>)[43] or FATHMM (<http://fathmm.biocompute.org.uk/>)[44] were selected from the CSV file.

Validation

All variants selected were confirmed by Sanger sequencing in the discovery cohort. The validation cohort was genotyped for these variants using Sanger sequencing or the TaqMan® genotyping assay (Applied Biosystems, Foster City, CA, USA). R software (<https://www.R-project.org/>)[45] was used to calculate HWE and to determine statistical significances in genotype and allele frequencies between cases and controls using Fisher's exact test with Bonferroni corrections to *P*-values.

Results

Laboratorial parameters

We compared the lactate dehydrogenase levels, hemoglobin concentrations, and fetal hemoglobin levels between 38 patients from the discovery cohort, and 177 patients from the validation cohort, separated in cases and controls. The fetal hemoglobin levels were different between cases and controls from both cohorts (for the discovery cohort: mean $= 10.31\%$ in cases and 14.46% in controls, P-value $= 0.046$; for the validation cohort: mean $= 5.21\%$ in cases and 8.71% in controls, P-value < 0.001). The lactate dehydrogenase levels were similar between cases and controls for both cohorts (means varying from 906.33 to 1069.77 U/L), and *P*-values greater than 0.05. The hemoglobin concentrations were not different between cases and controls in the discovery cohort (mean = 8.49 g/dL for cases and 9.06 g/dL for controls), with P-value $= 0.11$, whereas for the validation cohort, the difference observed was statistically significant, with P-value $= 0.02$ (mean = 7.62 g/dL for cases and 8.04 g/dL for controls).

WES quality control

The samples from the discovery cohort (40 SCA patients) had an average depth of coverage of $45.21 \pm 12.98\times$. The average number of reads for the samples was 365,898,080 $\pm$ 54,334,241. Identity-by-descent analysis revealed two possibly blood-related individuals in each group, and one of each pair was randomly excluded from the subsequent analysis.

Selection of variants for validation

After the bioinformatics analysis ($n = 38$) eight variants with *P*-values less than or near 0.05 were chosen (Table 1).

All the selected variants are the result of missense mutations (Table 1) apart from rs11454536 (insertion) and rs12568784 (stop codon). All the variants were confirmed by Sanger sequencing in the discovery cohort (87.5% consistency) except for rs13428956 (in the *FOXD4L1* gene), which was localized to a homologous region. As the exome reads are shorter (100 bp), they were probably aligned to another region of the genome. This variant was therefore excluded from the validation analysis.

The seven remaining variants were genotyped by Sanger sequencing (rs1975937 and rs12568784) or with the TaqMan® assay (rs4857302, rs3782489, rs11800462, rs201853154, and rs11454536) in the independent validation cohort of 293 SCA patients (Table 2).

No homozygotes for the variant allele were found for rs201853154. We believe this can be attributed to a genotyping error although the results of the TaqMan® genotyping assay were confirmed by Sanger sequencing in more than 10% of the samples. This could explain the *P*-value for HWE being less than 0.05 (Table 2).

Only variant rs12568784 ($P = 0.03$) showed a *P*-value of less than 0.05 for the GT and GG genotypes. The difference between cases and controls for allele distribution (G and T) bordered significance ($P = 0.07$) (Table 2).

For rs12568784, a similar distribution for alleles and genotypes was observed when the discovery and validation cohorts were compared except for the GG and GT genotypes in the cases (Figure 1).

Table 1. Variants selected by whole-exome sequencing in the discovery cohort.

			Minor allele frequency-MAF			Variant allele		
Gene	**Function**	**SNV ID**	**This study**	**ExAC**	**BIPMed**	**Cases**	**Controls**	***P*-value**
CRYBG3	A-kinase anchoring protein	rs4857302	0.486	0.508	0.382	23 (67.65%)	12 (31.58%)	0.004**
KRT77	Cornification and keratinization	rs3782489	0.486	0.369	0.361	11 (30.56%)	24 (66.67%)	0.004**
TNFRSF25	Apoptosis and lymphocyte homeostasis	rs11800462	0.176	0.206	0.080	11 (30.56%)	2 (5.26%)	0.005**
FOXD4L1	Morphogenesis, cell differentiation and transcription	rs13428956	0.145	0.017	0.049	10 (26.32%)	1 (2.63%)	0.007**
ZNF540	Transcriptional repressor	rs1975937	0.250	0.305	0.269	4 (10.53%)	15 (39.47%)	0.007**
UBTFL1	Proliferation of inner cell mass and trophectoderm cells	rs201853154	0.276	0.188	0.194	5 (13.16%)	16 (42.11%)	0.009**
VWDE	Coagulation	rs11454536	0.171	0.124	0.137	2 (5.26%)	11 (28.95%)	0.013*
FLG2	Establishment of skin barrier	rs12568784	0.237	0.232	0.175	14 (36.8%)	5 (13.16%)	0.057

$n = 38$ patients; Function data retrieved from Uniprot (<https://www.uniprot.org>).
ExAC: exome aggregation consortium; BIPMed: the Brazilian initiative on precision medicine; SNV: single nucleotide variant.
P-value: obtained by Fisher's exact test; **P*-value <0.05; ***P*-value <0.01.

Table 2. Variants genotyped in the validation cohort.

SNV ID	**Gene**	***P*-value for HWE**	***P*-value for genotypes**		***P*-value for alleles**
rs4857302 (A/C)	*CRYBG3*	0.40	AC vs. AA	0.10	0.78
			CC vs. AA	0.89	
rs3782489 (G/A)	*KRT77*	0.05*	GA vs. AA	0.71	0.27
			GG vs. AA	0.45	
rs11800462 (T/C)	*TNFRSF25*	0.15	TC vs. CC	0.96	0.38
			TT vs. CC	0.95	
rs1975937 (A/T)	*ZNF540*	0.11	AT vs. AA	0.47	0.48
			TT vs. AA	0.61	
rs201853154 (A/T)	*UBTFL1*	0.0005**	AT vs. AA	0.51	0.73
			–	–	–
rs11454536 (−/A)	*VWDE*	0.77	−/A vs. −/−	0.37	0.26
			A/A vs. −/−	0.86	
rs12568784 (G/T)	*FLG2*	0.08	GT vs. GG	0.03*	0.07
			TT vs. GG	0.84	

$n = 293$ patients; HWE: Hardy–Weinberg equilibrium; *P*-value: obtained with Fisher's exact test; **P*-value <0.05; ***P*-value <0.01. *P*-values for genotypes are two-sided. *P*-values for alleles are one-sided.

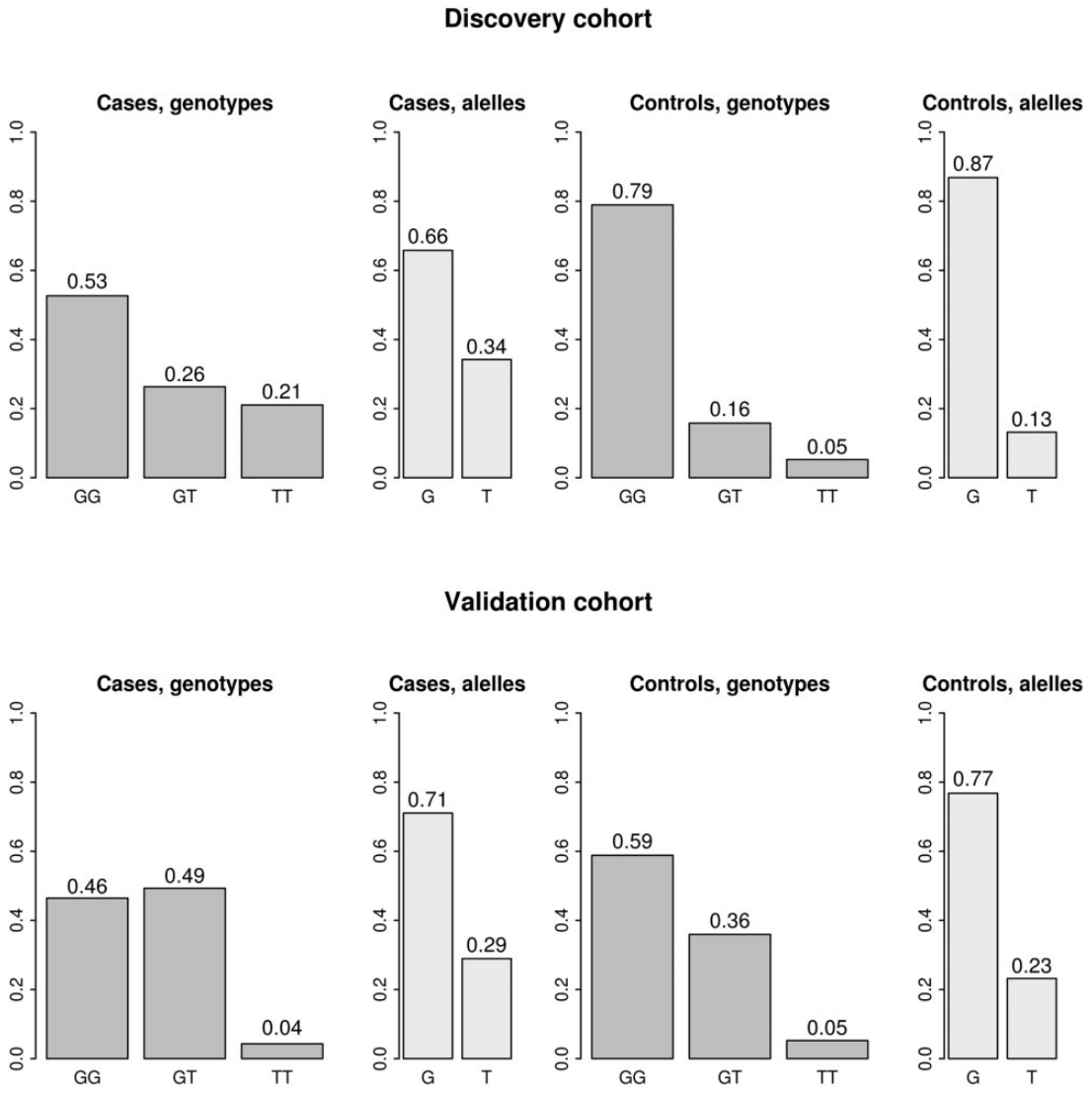

Figure 1. Genotypic and allelic distribution of rs12568784 (*FLG2* gene) in the discovery and validation cohorts. Graphs produced with R (<https://www.R-project.org/>).[45]

Discussion

The etiology of leg ulcers in SCA is complex and multifactorial. Mechanical obstruction by dense sickled red cells, infections, abnormal autonomic control with excessive vasoconstriction, venous disease, thrombosis, reduced oxygen-carrying capacity due to anemia and impaired endothelial function have all been suggested in the literature as contributing factors.[46,47] Leg ulcers can be disabling and become a chronic complication, usually affecting younger patients.[33] They can reduce patient quality of life as they cause pain, poor physical function and social isolation and hinder the development of social relationships.[48]

The literature data showed differences between lactate dehydrogenase levels,[7] hemoglobin concentrations, and fetal hemoglobin levels[8] in SCA patients with and without leg ulcers. Our results showed statistically significant differences in the percentage of fetal hemoglobin in both cohorts, statistically significant difference among total hemoglobin levels only in the validation cohort, and no difference in lactate dehydrogenase levels between case and control groups in both cohorts. Despite the differences observed, the values between cases and controls were similar, and probably it would be very difficult to explain the higher frequency of leg ulcers in the case group based on these values.

Studies on the association between genetic variants and susceptibility to leg ulcers in SCA are scarce. In one of the few such studies, Nolan *et al.*[9] genotyped candidate genes previously implicated in sickle vaso-occlusion to examine the relationship between leg ulcers and single nucleotide polymorphisms in these genes.[9] To our knowledge, the present study is the first to have used WES based on extreme phenotypes to identify new candidate genes associated with leg ulcers in SCA. By allowing other genes in addition to the usual set of candidate genes for this phenotype to be analyzed, this approach can reveal new associations.

In the discovery cohort, we identified seven variants in different genes that may be associated with leg ulcers in SCA as the *P*-values in the discovery cohort showed a potentially strong association (Table 1). Curiously, just

one variant (rs12568784 in the *FLG*2 gene) has been associated with an epidermal disease.[49] This variant showed a significant *P*-value when we compared the genotypes (GT and GG) between cases and controls in our validation cohort and a *P*-value bordering on significance when the alleles (G and T) were compared in both cohorts (Tables 1 and 2, Figure 1). Although associations with the other six variants were not confirmed in our validation sample, the genes where these variants were found are related to proliferation [*CRYBG3*[50,51] and *UBTFL1*[52]], epidermal morphogenesis (*KRT77*),[53] inflammation (*TNFRSF*25),[54] transcription (*ZNF*540),[55] and coagulation (*VWDE)*, processes that can be involved in the development of leg ulcers in SCA.

Although not previously implicated in the development of leg ulcers in SCA patients, rs12568784 (in the *FLG*2 gene) has been associated with persistent atopic dermatitis in African Americans.[49] This variant results in a premature stop codon in exon 3, which according to some studies could cause a loss of function in the protein filaggrin-2[56–59] and consequent impairment in skin barrier function in the stratum corneum, predisposing the individual to inflammation and infection and potentially explaining the persistence of atopic dermatitis in African-American patients.[49] In addition, it has been shown that filaggrin-2 expression in the skin is reduced in atopic dermatitis.[60]

These findings may help to explain the development of leg ulcers in SCA since sickle-cell individuals are of African descent.[61–63] Loss-of-function mutations in the *FLG*2 gene are not commonly found in people of African ancestry.[57,64,65] However, our results show that rs12568784 was present in more than 23% of our discovery cohort (Table 1), a finding consistent with a study of persistent atopic dermatitis in African Americans by Margolis *et al*.[49] who found this variant in 22% of their cases.[49]

Leg ulcers in SCA patients are typically chronic, lasting at least six months.[4] As rs12568784 has been associated with persistent atopic dermatitis,[49] it is reasonable to suggest that this variant can affect wound healing in SCA patients. Traumas have been reported to play an important role in the development of leg ulcers, which are predominantly a result of mosquito bites, a common occurrence in areas with a tropical climate.[7] Hence, we suggest that the rs12568784 variant is a probable contributor to a genetic predisposition to leg ulcers in SCA because it weakens the protective barrier of the skin, compromising healing. Traumas to the malleolar region may therefore result in chronic ulceration in patients with this variant.

The discovery cohort of extreme phenotypes comprised all the patients available at the Hematology and Hemotherapy Center, University of Campinas, Campinas, SP, Brazil during the enrollment period who fulfilled our inclusion/exclusion criteria of extreme phenotypes. According to the literature, it is possible to identify new candidate genes in a small sample using extreme phenotypes when they are well characterized,[32] as supported by the study conducted by Emond *et al*.[39] with a complex trait in a Mendelian disease and Shtir *et al.*[66] in a complication of a complex disease.

Some associations identified in the discovery cohort may have been overlooked in this study because of the small size of the validation cohort. A nominal association was found; however, this association was not statistically significant after Bonferroni correction (after the correction only *P*-values lower than 0.00625 were considered statistically significant).

Nevertheless, all the seven variants identified in the discovery cohort and their respective genes warrant further investigation as they may play a role in the etiology of leg ulcers in SCA as genetic modifiers. In particular, the rs12568784 variant deserves attention, since the literature showed evidences to support this association. An association between these seven variants and leg ulcers in SCA may be confirmed in studies with larger cohorts.

Our findings may provide new insights into the pathophysiology of chronic leg ulcers in SCA, potentially helping in the development of measures to prevent this complication of SCA in patients with risk genotypes and allowing early treatment options, more individualized treatment and improved quality of life for these patients.

Authors' contributions: All authors participated in the design, interpretation of the studies and analysis of the data and review of the manuscript; GQCS, MTI, IFD, DAF, MACB, ASA, ARLA, MSG and STOS helped in sample collection; GQCS conducted the experiments; GA, BBS, MGB and ILC conducted the bioinformatic and statistical analysis; SMSC, FFC and MBM supervised the whole work; GQCS, SMSC and MBM wrote the manuscript.

ACKNOWLEDGEMENTS

We thank all the patients and the research staff at the Hematology and Hemotherapy Center, University of Campinas, Campinas, SP, Brazil; the Bahia Hematology and Hemotherapy Foundation (HEMOBA), Salvador, BA, Brazil; and the Hematology and Hemotherapy Foundation of Pernambuco (HEMOPE), Recife, PE, Brazil.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by The São Paulo Research Foundation [grants 2014/00984–3 and 2015/24029–3].

ORCID iD

Murilo Guimarães Borges https://orcid.org/0000-0002-9854-5783

Mônica Barbosa de Melo https://orcid.org/0000-0002-1801-5441

REFERENCES

1. Frenette PS, Atweh GF. Sickle cell disease: old discoveries, new concepts, and future promise. *J Clin Invest* 2007;**117**:850–8
2. Odièvre MH, Verger E, Silva-Pinto AC, Elion J. Pathophysiological insights in sickle cell disease. *Ind J Med Res* 2011;**134**:532–7
3. Cummer CL, LaRocco CG. Ulcers of the legs in sickle cell anemia. *Arch Dermatol Syphilol* 1940;**42**:1015–39
4. Ballas SK, Kesen MR, Goldberg MF, Lutty GA, Dampier C, Osunkwo I, Wang WC, Hoppe C, Hagar W, Darbari DS, Malik P. Beyond the definitions of the phenotypic complications of sickle cell disease: an update on management. *Sci World J* [Internet] 2012;**2012**:949535
5. Minniti CP, Eckman J, Sebastiani P, Steinberg MH, Ballas SK. Leg ulcers in sickle cell disease. *Am J Hematol* 2010;**85**:831–3
6. Delaney KMH, Axelrod KC, Buscetta A, Hassell KL, Adams-Graves PE, Seamon C, Kato GJ, Minniti CP. Leg ulcers in sickle cell disease: current patterns and practices. *Hemoglobin* 2013;**37**:325–32
7. Cumming V, King L, Fraser R, Serjeant G, Reid M. Venous incompetence, poverty and lactate dehydrogenase in Jamaica are important predictors of leg ulceration in sickle cell anaemia. *Br J Haematol* 2008;**142**:119–25
8. Dubert M, Elion J, Tolo A, Diallo DA, Diop S, Diagne I, Sanogo I, Belinga S, Guifo O, Wamba G, Ngo Sack F, Boidy K, Kamara I, Traore Y, Diakite CO, Gbonon V, Faye BF, Seck M, Deme Ly I, Chelo D, N'Guetta R, Diop IB, Gaye B, Jouven X, Ranque B. Degree of anemia, indirect markers of hemolysis, and vascular complications of sickle cell disease in Africa. *Blood* 2017;**130**:2215–23
9. Nolan VG, Adewoye A, Baldwin C, Wang L, Ma Q, Wyszynski DF, Farrell JJ, Sebastiani P, Farrer LA, Steinberg MH. Sickle cell leg ulcers: associations with haemolysis and SNPs in Klotho, TEK and genes of the TGF-beta/BMP pathway. *Br J Haematol* 2006;**133**:570–8
10. Connes P, Lamarre Y, Hardy-Dessources MD, Lemonne N, Waltz X, Mougenel D, Mukisi-Mukaza M, Lalanne-Mistrih ML, Tarer V, Tressières B, Etienne-Julan M, Romana M. Decreased hematocrit-to-viscosity ratio and increased lactate dehydrogenase level in patients with sickle cell anemia and recurrent leg ulcers. *PLoS One* 2013;**8**:e79680
11. Connes P, Lamarre Y, Waltz X, Ballas SK, Lemonne N, Etienne-Julan M, Hue O, Hardy-Dessources MD, Romana M. Haemolysis and abnormal haemorheology in sickle cell anaemia. *Br J Haematol* 2014;**165**:564–72
12. Pope FM, Hodgson GA. Le ulceration and thalassaemia. *Br J Dermatol* 1968;**80**:840
13. Tanaka KR, Paglia DE. Pyruvate kinase deficiency. *Semin Hematol* 1971;**8**:367–96
14. Müller-Soyano A, Tovar de Roura E, Duke PR, De Acquatella GC, Arends T, Guinto E. Beutler Pyruvate kinase deficiency and leg ulcers. *Blood* 1976;**47**:807–13
15. Stevens DM, Shupack JL, Javid J, Silber R. Ulcers of the leg in thalassemia. *Arch Dermatol* 1977;**113**:1558–60
16. Peachey RD. Leg ulceration and haemolytic anaemia: an hypothesis. *Br J Dermatol* 1978;**98**:245–9
17. Gimmon Z, Wexler MR, Rachmilewitz EA. Juvenile leg ulceration in beta-thalassemia major and intermedia. *Plast Reconstr Surg* 1982;**69**:320–5 Feb
18. Vanscheidt W, Leder O, Vanscheidt E, Wokalek H, Wilmer J, Hofmann E, Schöpf E. Leg ulcers in a patient with spherocytosis: a clinicopathological report. *Dermatologica* 1990;**181**:56–9
19. Lawrence P, Aronson I, Saxe N, Jacobs P. Leg ulcers in hereditary spherocytosis. *Clin Exp Dermatol* 1991;**16**:28–30
20. Bordi B, Rosaria D'Amico M, Guariglia R, Capobianco G, Bordi E, Tirelli A. A case of congenital dyserythropoietic anemia type II, Gilbert's syndrome and malleolar trophic ulcers. *Hematology* 2002;**7**:197–9
21. Giraldi S, Abbage KT, Marinoni LP, Oliveira V, Pianowski M/A, Lehmkuhl AE, Fillus Neto J. Leg ulcer in hereditary spherocytosis. *Pediatr Dermatol* 2003;**20**:427–8
22. Borgna-Pignatti C, Marsella M, Zanforlin N. The natural history of thalassemia intermedia. *Ann N Y Acad Sci* 2010;**1202**:214–20
23. Pule GD, Mowla S, Novitzky N, Wiysonge CS, Wonkam A. A systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of sickle cell disease. *Expert Rev Hematol* 2015;**8**:669–79
24. Lanzkron S, Haywood C, Hassell KL, Rand C. Provider barriers to hydroxyurea use in adults with sickle cell disease: a survey of the Sickle Cell Disease Adult Provider Network. *J Natl Med Assoc* 2008;**100**:968–73
25. Minniti CP, Kato GJ. Critical reviews: how we treat sickle cell patients with leg ulcers. *Am J Hematol* 2016;**91**:22–30
26. Minniti CP, Taylor JG, Hildesheim M, O'Neal P, Wilson J, Castro O, Gordeuk VR, Kato GJ. Laboratory and echocardiography markers in sickle cell patients with leg ulcers. *Am J Hematol* 2011;**86**:705–8
27. Belini Junior E, Silva DGH, Torres L, de S, Okumura JV, Lobo CL, de C, Bonini-Domingos CR. Severity of Brazilian sickle cell disease patients: severity scores and feasibility of the Bayesian network model use. *Blood Cells Mol Dis* 2015;**54**:321–7
28. Brinsuk M, Tank J, Luft FC, Busjahn A, Jordan J. Heritability of venous function in humans. *Arterioscler Thromb Vasc Biol* 2004;**24**:207–11
29. Ofosu MD, Castro O, Alarif L. Sickle cell leg ulcers are associated with HLA-B35 and Cw4. *Arch Dermatol* 1987;**123**:482–4
30. Vicari P, Adegoke SA, Mazzotti DR, Cançado RD, Nogutti MAE, Figueiredo MS. Interleukin-1β and interleukin-6 gene polymorphisms are associated with manifestations of sickle cell anemia. *Blood Cells Mol Dis* 2015;**54**:244–9
31. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 2009;**461**:272–6
32. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;**12**:745–55
33. Koshy M, Entsuah R, Koranda A, Kraus AP, Johnson R, Bellvue R, Flournoy-Gill Z, Levy P. Leg ulcers in patients with sickle cell disease. *Blood* 1989;**74**:1403–8
34. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;**25**:1754–60
35. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;**20**:1297–303
36. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics* 2011;**27**:2156–81
37. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Zhang J, Guo Y, Feng B, Li H, Lu Y, Fang X, Liang H, Du Z, Li D, Zhao Y, Hu Y, Yang Z, Zheng H, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan J, Zhou Y, Qin J, Ma L, Li G, Yang Z, Zhang G, Yang B, Yu C, Liang F, Li W, Li S, Li D, Ni P, Ruan J, Li Q, Zhu H, Liu D, Lu Z, Li N, Guo G, Zhang J, Ye J, Fang L, Hao Q, Chen Q, Liang Y, Su Y, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Zheng H, Ren Y, Yang L, Gao Y, Yang G, Li Z, Feng X, Kristiansen K, Wong GK, Nielsen R, Durbin R, Bolund L, Zhang X, Li S, Yang H, Wang J. The diploid genome sequence of an Asian individual. *Nature* 2008;**456**:60–5
38. Ahn SM, Kim TH, Lee S, Kim D, Ghang H, Kim DS, Kim BC, Kim SY, Kim WY, Kim C, Park D, Lee YS, Kim S, Reja R, Jho S, Kim CG, Cha JY, Kim KH, Lee B, Bhak J, Kim SJ. The first Korean genome sequence and analysis: full genome sequencing for a socio-ethnic group. *Genome Res* 2009;**19**:1622–9
39. Emond MJ, Louie T, Emerson J, Zhao W, Mathias RA, Knowles MR, Wright FA, Rieder MJ, Tabor HK, Nickerson DA, Barnes KC, National Heart, Lung, and Blood(NHLBI) GO, Exome Sequencing Project Lung; GO, Gibson RL. Bamshad MJ. Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nat Genet* 2012;**44**:886–9
40. Sheehan VA, Crosby JR, Sabo A, Mortier NA, Howard TA, Muzny DM, Dugan-Perez S, Aygun B, Nottage KA, Boerwinkle E, Gibbs RA, Ware RE, Flanagan JM. Whole exome sequencing identifies novel genes for fetal hemoglobin response to hydroxyurea in children with sickle cell anemia. *PLoS One* 2014;**9**:e110740

41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75
42. Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet* 2012;**49**:433–6
43. Flanagan SE, Patch A-M, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomark* 2010;**14**:533–7
44. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, Day IN, Gaunt TR. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat* 2013;**34**:57–65
45. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Core Team, 2018, www.R-project.org/ (accessed 29 April 2019)
46. Aguilar Martinez P, Angastiniotis M, Eleftheriou A, Gulbis B, Mañú Pereira MDM, Petrova-Benedict R, Corrons JL. Haemoglobinopathies in Europe: health & migration policy perspectives. *Orphanet J Rare Dis* 2014;**9**:97
47. Steinberg MH, Brugnara C. Pathophysiological-based approaches to treatment of sickle cell disease. *Annu Rev Med* 2003;**54**:89–112
48. Umeh NI, Ajegba B, Buscetta AJ, Abdallah KE, Minniti CP, Bonham VL. The psychosocial impact of leg ulcers in patients with sickle cell disease: i don't want them to know my little secret. *PLoS One* 2017;**12**:e0186270
49. Margolis DJ, Gupta J, Apter AJ, Ganguly T, Hoffstad O, Papadopoulos M, Rebbeck TR, Mitra N. Filaggrin-2 variation is associated with more persistent atopic dermatitis in African American subjects. *J Allergy Clin Immunol* 2014;**133**:784–9
50. Chen H, Pei H, Hu W, Ma J, Zhang J, Mao W, Nie J, Xu C, Li B, Hei TK, Wang C, Zhou G. Long non-coding RNA CRYBG3 regulates glycolysis of lung cancer cells by interacting with lactate dehydrogenase A. *J Cancer* 2018;**9**:2580–8
51. Pei H, Hu W, Guo Z, Chen H, Ma J, Mao W, Li B, Wang A, Wan J, Zhang J, Nie J, Zhou G, Hei TK. Long noncoding RNA CRYBG3 blocks cytokinesis by directly binding G-actin. *Cancer Res* 2018;**78**:4563–72
52. Yamada M, Hamatani T, Akutsu H, Chikazawa N, Kuji N, Yoshimura Y, Umezawa A. Involvement of a novel preimplantation-specific gene encoding the high mobility group box protein Hmgpi in early embryonic development. *Hum Mol Genet* 2010;**19**:480–93
53. Bazzi H, Fantauzzo KA, Richardson GD, Jahoda CAB, Christiano AM. Transcriptional profiling of developing mouse epidermis reveals novel patterns of coordinated gene expression. *Dev Dyn off Publ Am Assoc Anat* 2007;**236**:961–70
54. Perks WV, Singh RK, Jones GW, Twohig JP, Williams AS, Humphreys IR, Taylor PR, Jones SA, Wang E. Death receptor 3 promotes chemokine-directed leukocyte recruitment in acute resolving inflammation and is essential for pathological development of mesothelial fibrosis in chronic disease. *Am J Pathol* 2016;**186**:2813–23
55. Xiang Z, Yuan W, Luo N, Wang Y, Tan K, Deng Y, Zhou X, Zhu C, Li Y, Liu M, Wu X, Li Y. A novel human zinc finger protein ZNF540 interacts with MVP and inhibits transcriptional activities of the ERK signal pathway. *Biochem Biophys Res Commun* 2006;**347**:288–96
56. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, Carrick T, Evans AT, Liao H, Zhao Y, Campbell LE, Schmuth M, Gruber R, Janecke AR, Elias PM, van Steensel MA, Nagtzaam I, van Geel M, Steijlen PM, Munro CS, Bradley DG, Palmer CN, Smith FJ, McLean WH, Irvine AD. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 2007;**39**:650–4
57. Winge MCG, Bilcha KD, Liedén A, Shibeshi D, Sandilands A, Wahlgren CF, McLean WH, Nordenskjöld M, Bradley M. Novel filaggrin mutation but no other loss-of-function variants found in Ethiopian patients with atopic dermatitis. *Br J Dermatol* 2011;**165**:1074–80
58. Marenholz I, Rivera VAG, Esparza-Gordillo J, Bauerfeind A, Lee-Kirsch M-A, Ciechanowicz A, Kurek M, Piskackova T, Macek M, Lee YA. Association screening in the Epidermal Differentiation Complex (EDC) identifies an SPRR3 repeat number variant as a risk factor for eczema. *J Invest Dermatol* 2011;**131**:1644–9
59. Henry J, Toulza E, Hsu C-Y, Pellerin L, Balica S, Mazereeuw-Hautier J, Paul C, Serre G, Jonca N, Simon M. Update on the epidermal differentiation complex. *Front Biosci* 2012;**17**:1517–32
60. Pellerin L, Henry J, Hsu C-Y, Balica S, Jean-Decoster C, Méchin M-C, Hansmann B, Rodriguez E, Weindinger S, Schmitt AM, Serre G, Paul C, Simon M. Defects of filaggrin-like proteins in both lesional and nonlesional atopic skin. *J Allergy Clin Immunol* 2013;**131**:1094–102
61. Gelpi AP. Migrant populations and the diffusion of the sickle-cell gene. *Ann Intern Med* 1973;**79**:258–64
62. Lehmann H. Distribution of the sickle cell gene: a new light on the origin of the East Africans. *Eugen Rev* 1954;**46**:101–21
63. Livingstone FB. Who gave whom hemoglobin S: the use of restriction site haplotype variation for the interpretation of the evolution of the βS -globin gene. *Am J Hum Biol off Biol* 1989;**1**:289–302
64. Margolis DJ, Apter AJ, Gupta J, Hoffstad O, Papadopoulos M, Campbell LE, Sandilands A, McLean WH, Rebbeck TR, Mitra N. The persistence of atopic dermatitis and filaggrin (FLG) mutations in a US longitudinal cohort. *J Allergy Clin Immunol* 2012;**130**:912–7
65. Gao P-S, Rafaels NM, Hand T, Murray T, Boguniewicz M, Hata T, Schneider L, Hanifin JM, Gallo RL, Gao L, Beaty TH, Beck LA, Barnes KC, Leung DY. Filaggrin mutations that confer risk of atopic dermatitis confer greater risk for eczema herpeticum. *J Allergy Clin Immunol* 2009;**124**:507–13
66. Shtir C, Aldahmesh MA, Al-Dahmash S, Abboud E, Alkuraya H, Abouammoh MA, Nowailaty SR, Al-Thubaiti G, Naim EA, ALYounes B, Binhumaid FS, ALOtaibi AB, Altamimi AS, Alamer FH, Hashem M, Abouelhoda M, Monies D, Alkuraya FS. Exome-based case-control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum Genet* 2016;**135**:193–200

(Received February 7, 2019, Accepted April 17, 2019)