A Comprehensive, Ethnically Diverse Library of Sickle Cell Disease-Specific Induced Pluripotent Stem Cells

Seonmi Park,1,2 Andrea Gianotti-Sommer,1,2 Francisco Javier Molina-Estevez,1,2 Kim Vanuytsel,1,3 Nick Skvir,1,3 Amy Leung,1,3 Sarah S. Rozelle,1,3 Elmutaz Mohammed Shaikho,4 Isabelle Weir,5 Zhihua Jiang,6 Hong-Yuan Luo,6 David H.K. Chui,5,6 Maria Stella Figueiredo,7 Abdulraham Alsultan,8 Amein Al-Al,9 Paola Sebastiani,5 Martin H. Steinberg,3,10 Gustavo Mostoslavsky,1,2,11,* and George J. Murphy 1,3,11,*

1Department of Medicine, Center for Regenerative Medicine (CReM), Boston University School of Medicine, 670 Albany Street, 2nd Floor, Boston, MA 02118, USA
2Section of Gastroenterology
3Section of Hematology-Oncology
Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA
4Bioinformatics Program, Boston University, Boston, MA 02215, USA
5Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA
6Hemoglobin Diagnostic Reference Laboratory, Boston Medical Center, Boston, MA 02118, USA
7Hematology and Blood Transfusion Division, Escola Paulista de Medicina, São Paulo 04023-062, Brazil
8Department of Pediatrics, Sickle Cell Disease Research Center, College of Medicine, King Saud University, Riyadh 12372, Saudi Arabia
9Center for Research & Medical Consultation, University of Dammam, Dammam 34212, Saudi Arabia
10Center of Excellence in Sickle Cell Disease, Boston Medical Center, Boston, MA 02118, USA
11Co-senior author
*Correspondence: gmostosl@bu.edu (G.M.), gjmurphy@bu.edu (G.J.M.)

http://dx.doi.org/10.1016/j.stemcr.2016.12.017

SUMMARY

Sickle cell anemia affects millions of people worldwide and is an emerging global health burden. As part of a large NIH-funded NextGen Consortium, we generated a diverse, comprehensive, and fully characterized library of sickle-cell-disease-specific induced pluripotent stem cells (iPSCs) from patients of different ethnicities, β-globin gene (HBB) haplotypes, and fetal hemoglobin (HbF) levels. iPSCs stand to revolutionize the way we study human development, model disease, and perhaps eventually, treat patients. Here, we describe this unique resource for the study of sickle cell disease, including novel haplotype-specific polymorphisms that affect disease severity, as well as for the development of patient-specific therapeutics for this phenotypically diverse disorder. As a complement to this library, and as proof of principle for future cell- and gene-based therapies, we also designed and employed CRISPR/Cas gene editing tools to correct the sickle hemoglobin (HbS) mutation.

INTRODUCTION

Sickle cell anemia, one of humankind’s most common hereditary monogenic diseases, is an emerging global health burden. In the United States, approximately 100,000 people are affected, annual mortality approaches 4%, and the costs of medical care exceed $1.1 billion (Kauf et al., 2009). Moreover, sickle cell disease is designated by the World Health Organization as a public health priority, with 300,000 births yearly, and it is estimated that 10 million African, Arab, and Indian individuals will be living with this disease in the future (Piel et al., 2013a, 2013b). In underdeveloped countries, this is a disease of childhood where most of the affected die young. With access to high-quality medical care, survival into the seventh and eighth decades is possible. Hydroxyurea is the sole approved drug treatment that alters disease pathophysiology by increasing the level of fetal hemoglobin (HbF). HbF has the property of inhibiting the polymerization of deoxy sickle hemoglobin (HbS), which is the proximal driver of disease pathophysiology (Steinberg et al., 2009).

As part of a large NIH-funded NextGen Consortium, we generated a comprehensive library of sickle-cell-disease-specific induced pluripotent stem cells (iPSCs) from patients of different ethnicities, β-globin gene (HBB) haplotypes, and HbF levels. iPSCs stand to revolutionize the way we study human development, model disease, and perhaps eventually, treat patients. Access to a genetically diverse cohort of sickle-cell-disease-specific iPSCs provides a unique resource for the study of novel haplotype-specific polymorphisms that affect disease severity as well as the development of novel patient-specific therapeutics for this phenotypically diverse disorder. As a complement to this library, and as proof of principle for future cell and gene-based therapies, we also designed and employed CRISPR/Cas gene editing tools to correct the sickle hemoglobin (HbS) mutation.
| Name of Line | Gender | Nationality of Origin | Age | Haplotype               |
|--------------|--------|-----------------------|-----|-------------------------|
| SS2-1        | female | US                    | 32  | atypical/indeterminate  |
| SS2-1GAG (CRISPR corrected) | female | US                    | 32  | atypical/indeterminate  |
| SS4-1        | male   | US                    | 30  | Benin/Senegal           |
| SS5-1        | male   | US                    | 32  | Benin/Benin             |
| SS8-2        | female | US                    | 31  | Senegal/Senegal         |
| SS9-1        | female | US                    | 29  | Benin/Bantu             |
| SS12-1       | female | US                    | 27  | atypical/indeterminate  |
| SS13-1       | female | US                    | 25  | Benin/Bantu             |
| SS14-1       | female | US                    | 39  | Benin/Benin             |
| SS15-1       | female | US                    | 28  | Benin/Bantu             |
| SS16-1       | female | US                    | 36  | Benin/Benin             |
| SS18-1       | female | US                    | 23  | atypical/indeterminate  |
| SS19-1       | male   | US                    | 30  | Benin/Benin             |
| SS24-1       | male   | US                    | 24  | Bantu/Bantu             |
| SS25-1       | female | US                    | 22  | Bantu/Bantu             |
| SS28-1       | female | US                    | 25  | atypical/indeterminate  |
| SS29-1       | female | US                    | 32  | Benin/Benin             |
| SS32         | female | US                    | 33  | Benin/Benin             |
| SS35         | male   | US                    | 50  | Benin/Bantu             |
| SS36         | male   | US                    | 38  | atypical/indeterminate  |
| SS37         | female | US                    | 37  | Benin/Benin             |
| SS38         | male   | US                    | 26  | Benin/Benin             |
| SS41-1       | male   | US                    | 21  | atypical/indeterminate  |
| SS43-1       | female | US                    | 32  | Senegal/Senegal         |
| SS44-1       | female | US                    | 23  | Benin/Benin             |
| SS45-1       | female | US                    | 37  | atypical/indeterminate  |
| SS47-1       | female | US                    | 42  | atypical/indeterminate  |
| SS48-1       | male   | US                    | 30  | atypical/indeterminate  |
| SS49-1       | male   | US                    | 31  | Benin/Benin             |
| SA5-1        | female | Saudi Arabia          | 9   | atypical/indeterminate  |
| SA36         | female | Saudi Arabia          | 26  | Benin/Benin             |
| SA40-1       | male   | Saudi Arabia          | 20  | Benin/Benin             |
| SA53-1       | male   | Saudi Arabia          | 14  | atypical/indeterminate  |
| SA64         | male   | Saudi Arabia          | 14  | Benin/Benin             |

(Continued on next page)
RESULTS

Establishment of an Ethnically Diverse, Sickle-Cell-Disease-Specific iPSC Library Representing Multiple HbS Gene-Associated Haplotypes

Peripheral blood samples were procured from three geographical locations to obtain a wide representation of four common HbS gene haplotypes. This starting material was used to generate 54 independent iPSC lines from individuals of African American, Brazilian, and Saudi Arabian descent (Table 1). Three independent clones were generated from each patient, with starting samples representative of both genders and a wide range of ages (3–53 years old). All lines were created using previously described methodologies and met stringent quality control parameters for pluripotency and functionality (Sommer et al., 2009, 2012; Somers et al., 2010) (Figure 1). Most lines in the library have been adapted to grow under feeder-free conditions with the entire library available through WiCell (https://www.wicell.org/) using simplified material transfer agreements (MTAs).

Genetic Characterization of Banked iPSC Lines

Detailed genotyping of all samples included confirmation of homozygosity for the HbS gene mutation, focused SNP analyses of sentinel SNPs of BCL11A and MYB and five SNPs in the HBB gene cluster used to determine classical haplotypes of this gene cluster (Tables 1 and 2). These analyses confirmed the diverse representation of HBB haplotypes within our cohort, including African American homozygotes and compound heterozygotes for the Senegal, Benin, and Bantu haplotypes; Saudi patients with the Arab-Indian haplotype and Saudi patients with the Benin haplotype whose HbF levels are about twice as high as in the African Benin haplotype; and Brazilian patients homozygous and heterozygous for the Bantu haplotype that is typically associated with the lowest HbF of all HBB haplotypes. In addition, and as part of the NIH NextGen

Table 1. Continued

| Name of Line | Gender | Nationality of Origin | Age | Haplotype         |
|--------------|--------|-----------------------|-----|-------------------|
| SA82-2       | male   | Saudi Arabia          | 24  | Benin/Benin       |
| SA108        | male   | Saudi Arabia          | 9   | Arab-Indian/Arab-Indian |
| SA208        | male   | Saudi Arabia          | 7   | atypical/indeterminate |
| SA209-1      | male   | Saudi Arabia          | 12  | Benin/Benin       |
| SA210-1      | male   | Saudi Arabia          | 9   | Benin/Benin       |
| SA50-1       | female | Saudi Arabia          | NA  | Arab-Indian/Arab-Indian |
| SA106-1      | female | Saudi Arabia          | NA  | Arab-Indian/Arab-Indian |
| SA138-1      | male   | Saudi Arabia          | 16  | Atypical/Indeterminate |
| SA170-1      | male   | Saudi Arabia          | 3   | Arab-Indian/Arab-Indian |
| BR-SP-3-1    | female | Brazil                | 34  | Bantu/Bantu       |
| BR-SP-21-1   | female | Brazil                | 20  | atypical/indeterminate |
| BR-SP-23-1   | female | Brazil                | 23  | Bantu/Bantu       |
| BR-SP-25-1   | male   | Brazil                | 34  | Bantu/Bantu       |
| BR-SP-29-1   | male   | Brazil                | 20  | Benin/Bantu       |
| BR-SP-31-1   | male   | Brazil                | 35  | Benin/Benin       |
| BR-SP-33-1   | female | Brazil                | 53  | Benin/Bantu       |
| BR-SP-37-1   | female | Brazil                | 20  | atypical/indeterminate |
| BR-SP-39-1   | male   | Brazil                | 22  | Benin/Bantu       |
| BR-SP-41-1   | male   | Brazil                | 22  | Bantu/Bantu       |
| BR-SP-43-1   | male   | Brazil                | 21  | Bantu/Bantu       |
| BR-SP-45-1   | female | Brazil                | 20  | Atypical/Indeterminate |
Consortium effort, all of our lines have undergone whole-genome array (Illumina Infinium Expanded Multi-Ethnic Genotyping Array “MEGAEX”) to establish fingerprinting-based identity (by comparing with the donor sample material) and genomic integrity. These data are also deposited in dbGaP.

Characterization of Capacity for Directed Differentiation into the Erythroid Lineage

Differentiation capacity characterization of banked lines was performed as we previously described using a 2D, feeder-free, and chemically defined erythroid specification protocol (Smith et al., 2013). On erythroid specification, iPSC-derived erythroblasts were uniform in morphology and robustly expressed hemoglobin as assessed by Wright-Giemsa and benzidine staining, respectively (Figure 2 A). Cell pellets from iPSC-derived erythroblasts demonstrated increased accumulation of hemoglobin as differentiation proceeded (Figure 2 B). Fluorescence-activated cell sorting (FACS) analysis of erythroid specification using representative iPSC lines from the four major haplotypes of sickle cell anemia demonstrated efficient specification as judged by robust coexpression of CD71 (transferrin receptor) and CD235 (glycoporphin A), two markers of the erythroid lineage (Figure 2 C).

Transcriptional Profiling of Normal and Sickle Cell iPSC upon Differentiation into Erythroid Progeny

To assess the gene expression signature of our cells in the context of directed differentiation, we employed digital gene expression (DGE) (Cacchiarelli et al., 2015), which provides a relatively inexpensive approach to perform high-fidelity differential RNA sequencing. For these DGE-based studies, we analyzed 12 samples representative of all the geographic locations, including three control lines and nine sickle cell lines. We collected RNA at day 0, day 20, and day 25 of differentiation and submitted samples in duplicate for sequencing. We analyzed the data using the multitest package in R after rescaling to identify genes with differential expression and to remove samples of poor quality. Differentially expressed genes were selected if the Bonferroni corrected p value from the t test was <0.05, and the fold change was greater than 2. The heat-map in Figure 3 A shows a clear expression signature emerging over time during differentiation. At days 15 and 20, we found 867 genes differentially expressed compared with day 0, whereas only 7 genes were differentially expressed comparing day 20 with day 25. By focusing on a specific set of genes, we noted that all differentiated lines were capable of upregulating erythroid-relevant genes, including KLF1, GATA1 and GATA2, FTH1, TAL1, and several globin genes (Figure 3 B).
Sickle Glu6Val Reversion Using CRISPR/CAS9

As a proof of principle for future gene/cell therapy for sickle cell disease and taking advantage of the fact that all patients with sickle cell disease share the same exact single point mutation, we designed and constructed universal CRISPRs to target and correct the A > T mutation present in the sixth codon of the beta globin coding sequence. Previous studies have shown that the HBB locus is susceptible to being modified by nucleases helped by inclusion of drug resistance during selection (Huang et al., 2015). We decided to establish CRISPR/CAS9 nucleases targeting the sequences in the closest vicinity to the Glu6Val mutation, as sequences downstream within the first HBB exon have high homology with HBD (Craddock et al., 2013). We constructed two guide RNAs targeting positions –13 and +2nt from the mutation (Figure 4A). To obtain correction of the sickle mutation in the absence of selection, we designed single-strand donor oligonucleotides (ssODN) that included the normal HBB sequence and discrete silent mutations facilitating quick screening and preventing re-binding of the guide RNA to the corrected DNA strand, as shown in Figure 4B. Using this approach, we found an overall 40% efficiency of CRISPR-mediated indels, with the vast majority being deletions as reported by others (Figures 4C and 4D). We were able to create a corrected clone by homology-directed repair, and biallelic sequencing confirmed correction of one of the sickle mutant alleles, with an out-of-frame deletion in the other allele, creating a corrected sickle cell clone, named SCD iPSC SS.2-1-GAG (Figure 4E). Characterization of the corrected SCD patient-derived iPSC line completely mirrored parental features in terms of morphology and growth, with pluripotency markers and a normal karyotype maintained (Figures 4F and 4G). Furthermore, the corrected clone showed the same erythroid specification efficiency with a similar pattern of erythroid-specific marker expression as compared with the original sickle cell parental iPSC line (Figure 4H).

Patient Consent and Global Distribution of Created Lines

All the iPSC lines in this bank were created from patients using a progressive state-of-the-art consent form under the Boston University Institutional Review Board (H32506). This consent form includes a comprehensive template that allows for the unrestricted sharing of created lines, including potential commercialization and sharing of lines with commercial entities. As a resource to investigators, this consent form has been included as a Supplemental Information. In addition, all cell lines have been deposited with the WiCell Stem Cell Bank for distribution to the scientific community. Investigators may request specific lines directly from the WiCell website (www.wicell.org). Once the requesting investigator has executed the appropriate MTA and provided the designated transfer fee, the lines are shipped directly to the requesting investigator from WiCell’s facilities. Fees collected from investigators fund the re-banking and characterization of in-demand lines and support the continued preservation and availability of all materials deposited with WiCell. WiCell’s established record in contract management and domestic and international shipping ensure that the lines will be distributed widely and without impediment. WiCell also offers ongoing customer and technical service to support investigators. This guarantees that questions are answered in a timely manner and helps to ensure the success of investigators using lines provided through WiCell.

DISCUSSION

We generated a library of iPSCs from patients with sickle cell anemia of diverse ethnicities and HBB haplotypes to study the biology of these cells and the feasibility of their generation from blood samples collected from patients in distant locations and shipped frozen to our laboratory. These fully characterized lines, along with accompanying genetic and hematologic data, are now freely available.

Drug development is an expensive and time-consuming process that requires stringent specificity, potency, and toxicity validations of potential novel therapeutics. Traditionally, drug discovery proceeds from testing in in vitro cell-based assays in the laboratory to in vivo animal models, followed by three phases of clinical testing. Unfortunately, potential therapeutics usually are not extensively tested in humans until phase II clinical trials, which can occur many years after initial drug discovery. If in vitro testing is performed on human cells before clinical trials, these cells are typically immortalized cell lines, which have undergone genetic alterations to ensure their immortalization, possibly altering the fidelity of the drug screens. Use of immortalized cell lines is a common cause of high
attrition rates for drug development, as what works in vitro and subsequently in animal models may not always translate to the clinic (Kola and Landis, 2004). Pluripotent stem cells, and in particular iPSCs, have the opportunity to revolutionize preclinical drug screening. iPSC technology offers the prospect of an unlimited supply of material and is ideal for screening drugs against the genetic variations found in a patient population, such as those suffering from sickle cell disease for which there is currently only a single FDA-approved drug. Sickle cell disease is phenotypically diverse, a quality that arises primarily from the known and unknown quantitative trait loci that regulate HbF expression and are polymorphic in diverse patient populations. This variance has led to many discoveries regarding transcriptional regulation of HbF and further elucidated the complexities of hemoglobin switching. Since there are still many unknown regulators of HbF expression, finding drugs that will be efficacious in patients with a variety of genetic backgrounds would be ideal, and the creation of the described iPSC bank may contribute to this effort.

Cell-based treatments for sickle cell disease include blood transfusion, hematopoietic stem cell transplantation, and nascent trials of gene therapy. It is hoped that the gene editing tools described in this work, coupled with corrected sickle-cell-disease-specific iPSCs could one day provide a functional cure for the disorder. Erythroid-progenitor-derived iPSCs also hold promise for development as a potential, autologous, cellular therapeutic due to their constitutive HbF expression without progression to an adult globin phenotype (Smith et al., 2013). An autologously derived erythroid progenitor that makes high concentrations of HbF should render any remaining HbS incapable of damaging the sickle erythrocyte (Ngo et al., 2012).

**EXPERIMENTAL PROCEDURES**

**Patient Samples**
To capture the phenotypic diversity of this complex disease, samples were procured from three geographical locations in an attempt to obtain a wide representation of HbF-related haplotypes. We
collected peripheral blood from individuals at the Center of Excellence for Sickle Cell Disease at Boston Medical Center representative of Africa Americans with sickle cell anemia, the Sickle Cell Disease Research Center in Riyadh and Center for Research and Medical Consultation in Dammam, both in Saudi Arabia, and from the Escola Paulista de Medicina in São Paulo, Brazil. Reprogramming of material from the Boston location was performed on fresh samples immediately following collection, while samples sourced from Saudi Arabia and Brazil required the shipment of frozen mononuclear cells to Boston for reprogramming.

**iPSC Generation**

Derivation of our entire iPSC library was performed as described (Sommer et al., 2009, 2012). Briefly, 4 mL of peripheral blood was collected from all participating individuals, and the mononuclear cells (either fresh or frozen) were expanded in vitro and reprogrammed using the STEMCCA vector. Although all lines in the bank are currently unexcised, the STEMCCA vector used in these studies is equipped with a reprogramming cassette flanked by *LoxP* sites that allows for the excision of reprogramming genes. At least three independent clones were established, expanded, and banked from each individual. For all studies described here, cells were maintained either on inactivated murine embryonic fibroblast feeders with knockout serum replacement supplemented media, or under feeder-free conditions using mTeSR1 media. All iPSC lines are available for distribution through WiCell (http://www.wicell.org/). These studies were approved by the institutional review boards of the participating institutions.

**Immunofluorescence Staining**

Cells were fixed in 4% paraformaldehyde/PBS and stained with mouse anti-human TRA-1-81 (EMD Millipore, MAB4381) followed by secondary antibody, Alexa Fluor 488 conjugated goat anti-mouse immunoglobulin M (Thermo Fisher, A21042).

**FACS Analysis**

Cells were stained with phycoerythrin-conjugated mouse anti-Human CD235a (BD Pharmingen, 555570) and allophycocyanin-conjugated mouse anti-Human CD71 (BD Pharmingen, 555570) and analyzed on a Stratedigm S1000EXi flow cytometer. FlowJo software was used for data analysis.

**Erythroid-Lineage Specification**

In vitro differentiation of iPSC into erythroid-lineage cells was done as previously described (Smith et al., 2013).

**Genotyping and Haplotype Analysis**

Homozygosity for the HbS gene was confirmed using amplification refractory mutation system analysis (Little, 2001). HBB gene cluster haplotypes were ascertained by analysis of five SNPs: rs7482144 (Xmn1 C-T restriction site 158 bp 5' to HBG2), rs10128556 (HincII restriction site in HBBP1), rs28440105 (HindIII restriction site in HBG1), rs3834466 (Hinc2 restriction site 5' to HBE1), and rs968857 (HincII restriction site in HBD) as described in Alsultan et al., 2012). Targeted genotyping of SNPs was done with tetra-primer ARMS-PCR, TaqMan assays. Haplotypes were ascertained using combinations of the five SNPs (rs28440105, rs10128556, rs7482144, rs3834466, rs968857) as shown in http://www.ncbi.nlm.nih.gov/pubmed/21440534. Included in the genetic studies...
Figure 4. CRISPR/Cas9 Correction of the SCD Mutation in iPSCs
(A) Genomic Glu6Val mutated HBB gene and position of the engineered guide RNAs HBBg_1 and HBBg_2 targeting the vicinity of the mutation.
(B) Alignment of donor ssODNs: 140 bp ssODN were used as donor templates to restore normal HBB sequence. Normal A at position +20 is highlighted in green, sickle mutation (T) in red, and additional mutations within the PAM sequence to prevent recutting after editing are indicated in yellow.
(C) Distribution of CRISPR/CAS9 genetic modifications in clones electroporated with HBBg_2/ssODN_2. This chart represents clones from two independent SCD iPSC lines.
(legend continued on next page)
completed on these lines was genome-wide SNP analysis using whole-genome arrays (Illumina Infinium Expanded Multi-Ethnic Genotyping Array “MEGAEx”). These arrays contained ~1 million SNPs.

**DGE Analysis**

For DGE, RNA was collected from all samples using the RNeasy Mini Kit (QIAGEN) and sent out for library construction and sequencing at the Broad Institute (Cacchiarelli et al., 2015). After removal of poor-quality samples, samples were rescaled to a common target across plates, duplicate samples were averaged, and cubic root transformed to follow approximately normal distribution. Differential expression was tested using t tests and levels of significance were corrected using Bonferroni correction. Data are displayed using heatmaps generated with the Heatplus package in R v3. Enrichment analysis was conducted using the David portal (https://david.ncifcrf.gov/home.jsp).

**CRISPR/CAS9 Gene Editing**

CRISPR/CAS9 was done as described (Ding et al., 2013). To minimize the chances of off-target cleavage, we cross-matched two independent online resources for in silico CRISPR evaluation (Heigwer et al., 2014; Montague et al., 2014) to identify several candidate guide RNAs in close proximity to the mutation, +2 and −13 nt from the mutation for HBBg_1 and HBBg_2, respectively. The best two candidates were cotransfected in sickle cell iPSCs together with Cas9 and the ssODN donor using electroporation. Efficiency of cutting by different guide RNA/CAS9 combinations was ranked by Cel-I surveyor assay using genomic DNA isolated from SCD iPSC (see Figure S1). The two best candidates were cotransfected in iPSiCC S2.2-1 together with Cas9 and the ssODN donor. Positive Cas9-GFP iPSCs were sort purified (Figure S2) and plated into 100 mm plates at a density of 500–2000 cells/cm² to facilitate the selection of clonal iPSC lines. After 16–20 days, when the average colony size was consistent, individual clones were manually picked and seeded into 96-well plates for screening and expansion.

**ACCESSION NUMBERS**

SNP analyses, transcriptomic data, along with associated clinical and phenotypic data of all iPSC clones are available through dbGaP (Study NHLBI Globin Sickle Cell iPSC; accession phs001212.v1.p1).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes two figures and the research consent form and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.12.017.

**AUTHOR CONTRIBUTIONS**

G.J.M., G.M., D.H.K.C., and M.H.S conceived and designed the study; S.P., A.G.S., F.J.M., K.V., A.L., S.S.R., Z.J., H.Y.L., M.S.F., A.A., and A.A.A. collected data; N.S., E.M.S., and I.W. performed statistical analyses under the supervision of P.S.; and G.J.M., G.M., and M.H.S wrote the manuscript with input from all authors.

**ACKNOWLEDGMENTS**

This work was funded by the NextGen Consortium U01HL107443 from the NIH/NHLBI, the University of Dammam, SP 11/2011, Office of Collaboration and Knowledge Exchange, University of Dammam, Training grant for Biostatistics ST32GM074905, and Training Grant for Hematology ST32 HL007501. Anthony Akim-bami, MD, MPH, assisted with sample collection.

Received: July 19, 2016
Revised: December 13, 2016
Accepted: December 16, 2016
Published: January 19, 2017

**REFERENCES**

Alsultan, A., Ngo, D.A., Farrell, J.J., Akinsheyne, L., Solovieff, N., Ghabbour, H.A., Al-Alli, A., Alsuliman, A., Al-Baghshi, M., Albu-Ali, W., et al. (2012). A functional promoter polymorphism of the delta-globin gene is a specific marker of the Arab-Indian haplotype. Am. J. Hematol. 87, 824–826.

Cacchiarelli, D., Trapnell, C., Ziller, M.J., Kere, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. Cell 162, 412–424.

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. (2013). CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 41, 9584–9592.

Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. (2013). Enhanced efficiency of human pluripotent stem cell...
cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell 12, 393–394.

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. Nat. Methods 11, 122–123.

Huang, X., Wang, Y., Yan, W., Smith, C., Ye, Z., Wang, J., Gao, Y., Mendelsohn, L., and Cheng, L. (2015). Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. Stem Cells 33, 1470–1479.

Kauf, T.L., Coates, T.D., Huazhi, L., Mody-Patel, N., and Hartzema, A.G. (2009). The cost of health care for children and adults with sickle cell disease. Am. J. Hematol. 84, 323–327.

Kola, I., and Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? Nat. Rev. Drug Discov. 3, 711–715.

Little, S. (2001). Amplification-refractory mutation system (ARMS) analysis of point mutations. Curr. Protoc. Hum. Genet Chapter 9, Unit 9.8.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 42 (Web Server issue), W401–W407.

Ngo, D.A., Aygun, B., Akinsheye, I., Hankins, J.S., Bhan, I., Luo, H.Y., Steinberg, M.H., and Chui, D.H. (2012). Fetal haemoglobin levels and haematological characteristics of compound heterozygotes for haemoglobin S and deleterious hereditary persistence of fetal haemoglobin. Br. J. Haematol. 156, 259–264.

Piel, F.B., Hay, S.L., Gupta, S., Weatherall, D.J., and Williams, T.N. (2013a). Global burden of sickle cell anaemia in children under five, 2010-2050: modelling based on demographics, excess mortality, and interventions. PLoS Med. 10, e1001484.

Piel, F.B., Patil, A.P., Howes, R.E., Nyangiri, O.A., Gething, P.W., Dewi, M., Temperley, W.H., Williams, T.N., Weatherall, D.J., and Hay, S.I. (2013b). Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet 381, 142–151.

Smith, B.W., Rozelle, S.S., Leung, A., Ubellacker, J., Parks, A., Nah, S.K., French, D., Gadue, P., Monti, S., Chui, D.H., et al. (2013). The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. Blood 122, 376–385.

Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean, J.M., Smith, B.W., et al. (2010). Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells 28, 1728–1740.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells 27, 543–549.

Sommer, A.G., Rozelle, S.S., Sullivan, S., Mills, J.A., Park, S.M., Smith, B.W., Iyer, A.M., French, D.L., Kotton, D.N., Gadue, P., et al. (2012). Generation of human induced pluripotent stem cells from peripheral blood using the STEMCCA lentiviral vector. J. Vis. Exp. http://dx.doi.org/10.3791/4327.

Steinberg, M.H., Forget, B.G., Higgs, D.R., and Weatherall, D.J. (2009). Disorders of Hemoglobin: Genetics, Pathophysiology, Clinical Management (Cambridge University Press).