



OPEN

## Gene expression changes in sickle cell reticulocytes and their clinical associations

Xu Zhang<sup>1</sup>, Jihyun Song<sup>2</sup>, Binal N. Shah<sup>1</sup>, Jin Han<sup>3</sup>, Taif Hassan<sup>1</sup>, Galina Miasniakova<sup>4</sup>, Adelina Sergueeva<sup>5</sup>, Sergei Nekhai<sup>6</sup>, Roberto F. Machado<sup>7</sup>, Mark T. Gladwin<sup>8</sup>, Santosh L. Saraf<sup>1</sup>, Josef T. Prchal<sup>2,9</sup> & Victor R. Gordeuk<sup>1,9</sup>

Transcriptional changes in compensatory erythropoiesis in sickle cell anemia (SCA) and their disease modulation are unclear. We detected 1226 differentially expressed genes in hemoglobin SS reticulocytes compared to non-anemic hemoglobin AA controls. Assessing developmental expression changes in hemoglobin AA erythroblasts for these genes suggests heightened terminal differentiation in early erythroblasts in SCA that diminishes toward the polychromatic to orthochromatic stage transition. Comparison of reticulocyte gene expression changes in SCA with that in Chuvash erythrocytosis, a non-anemic disorder of increased erythropoiesis due to constitutive activation of hypoxia inducible factors, identified 453 SCA-specific changes attributable to compensatory erythropoiesis. Peripheral blood mononuclear cells (PBMCs) in SCA contain elevated proportions of erythroid progenitors due to heightened erythropoiesis. Deconvolution analysis in PBMCs from 131 SCA patients detected 54 genes whose erythroid expression correlated with erythropoiesis efficiency, which were enriched with SCA-specific changes (OR = 2.9, P = 0.00063) and annotation keyword “ubiquitin-dependent protein catabolic process”, “protein ubiquitination”, and “protein polyubiquitination” (OR = 4.2, P = 7.5 × 10<sup>-5</sup>). An erythroid expression quantitative trait locus of one of these genes, *LNX2* encoding an E3 ubiquitin ligase, associated with severe pain episodes in 774 SCA patients (OR = 1.7, P = 3.9 × 10<sup>-5</sup>). Thus, erythroid gene transcription responds to unique conditions within SCA erythroblasts and these changes potentially correspond to vaso-occlusive manifestations.

In sickle cell anemia (SCA), polymerization of deoxygenated hemoglobin S results in erythrocytes with reduced deformability<sup>1</sup>, altered properties of flow and aggregation<sup>2</sup>, and high membrane conductance of cations<sup>3</sup>. Abnormal display of adhesive surface molecules induces erythrocyte interaction with endothelial cells and other blood cells to occlude microvasculature<sup>4</sup>. Damaged erythrocytes rupture or are removed from circulation by phagocytes at a much younger age of 17–35 days than 100–120 days in normal individuals<sup>5</sup>. Compensatory erythropoiesis occurs in response to hemolytic anemia and tissue hypoxia in SCA<sup>6</sup>. Reticulocytes in SCA constitute 4–15% or more of total erythrocytes compared to 0.5–1.5% in normal individuals.

Polymerization of hemoglobin S also occurs in bone marrow erythroblasts that synthesize and accumulate hemoglobin<sup>7</sup>, thereby impacting erythropoiesis<sup>8</sup>. High cell death rates occur between the polychromatic and orthochromatic maturation stages of in vitro and in vivo derived SCA erythroblasts<sup>9</sup>. The phenomenon is reminiscent of ineffective erythropoiesis in β-thalassemia major and intermedia, characterized by accelerated erythroid differentiation, maturation arrest, and apoptosis at the polychromatic stage<sup>10</sup>, although ineffective erythropoiesis is less marked in SCA versus β-thalassemia.

The pattern and regulation of transcription changes characterizing compensatory erythropoiesis in SCA are largely unknown. In derived human erythroblasts, global transcriptional activation peaks in proerythroblasts and reduces to 40% of the peak in polychromatic and orthochromatic stages<sup>11</sup>. Distinct sets of genes are activated at

<sup>1</sup>Department of Medicine, University of Illinois at Chicago, Chicago, IL, USA. <sup>2</sup>Department of Medicine, University of Utah, Salt Lake City, UT, USA. <sup>3</sup>College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA. <sup>4</sup>Chuvash Republic Clinical Hospital 2, Cheboksary, Russia. <sup>5</sup>Cheboksary Children’s Hospital, Cheboksary, Russia. <sup>6</sup>Center for Sickle Cell Disease, Howard University, Washington, DC, USA. <sup>7</sup>Division of Pulmonary, Critical Care, Sleep, and Occupational Medicine, Department of Medicine, Indiana University, Indianapolis, IN, USA. <sup>8</sup>Division of Pulmonary, Allergy, and Critical Care Medicine, Vascular Medicine Institute, University of Pittsburgh, Pittsburgh, PA, USA. <sup>9</sup>These authors contributed equally: Josef T. Prchal and Victor R. Gordeuk. ✉email: zhangxu@uic.edu; josef.prchal@hsc.utah.edu; vgordeuk@uic.edu

stage transitions throughout erythroid differentiation<sup>12,13</sup>. Reticulocytes retain transcripts from erythroblasts<sup>14</sup>; therefore differential gene expression observed in reticulocytes between SCA patients and healthy individuals can reflect cumulative variation in the preceding stages. Importantly, some observations suggest contribution of reticulocytes to vaso-occlusion manifested as acute painful episodes<sup>4,15,16</sup>. We postulated that reticulocyte transcriptional variation may give rise to certain disease manifestations.

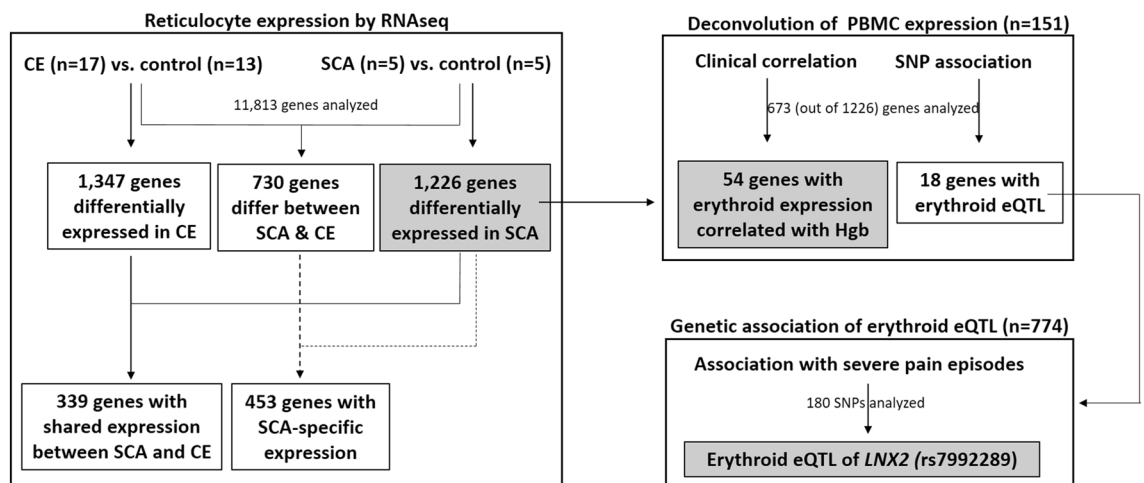
We applied RNA-Seq to reveal gene expression changes in reticulocytes from SCA patients compared to healthy African Americans. We delineated SCA-specific regulation of reticulocyte transcripts using gene expression changes in Chuvash erythrocytosis (CE) as a comparator<sup>17</sup>. CE is a monogenic disorder characterized by homozygosity for hypomorphic *VHL*<sup>R200W</sup> leading to constitutive hypoxia inducible factor (HIF) activation that results in increased erythropoietin, erythrocytosis and a propensity to thrombosis<sup>18</sup>. We assessed clinical correlations with erythroid expression variation employing deconvolution analysis<sup>19</sup> of archived gene expression data from SCA peripheral blood mononuclear cells (PBMCs), which contain elevated erythroid progenitors due to heightened erythropoiesis<sup>20</sup>. We further assessed genetic association between erythroid expression quantitative trait loci (eQTL) and clinical manifestations.

## Results

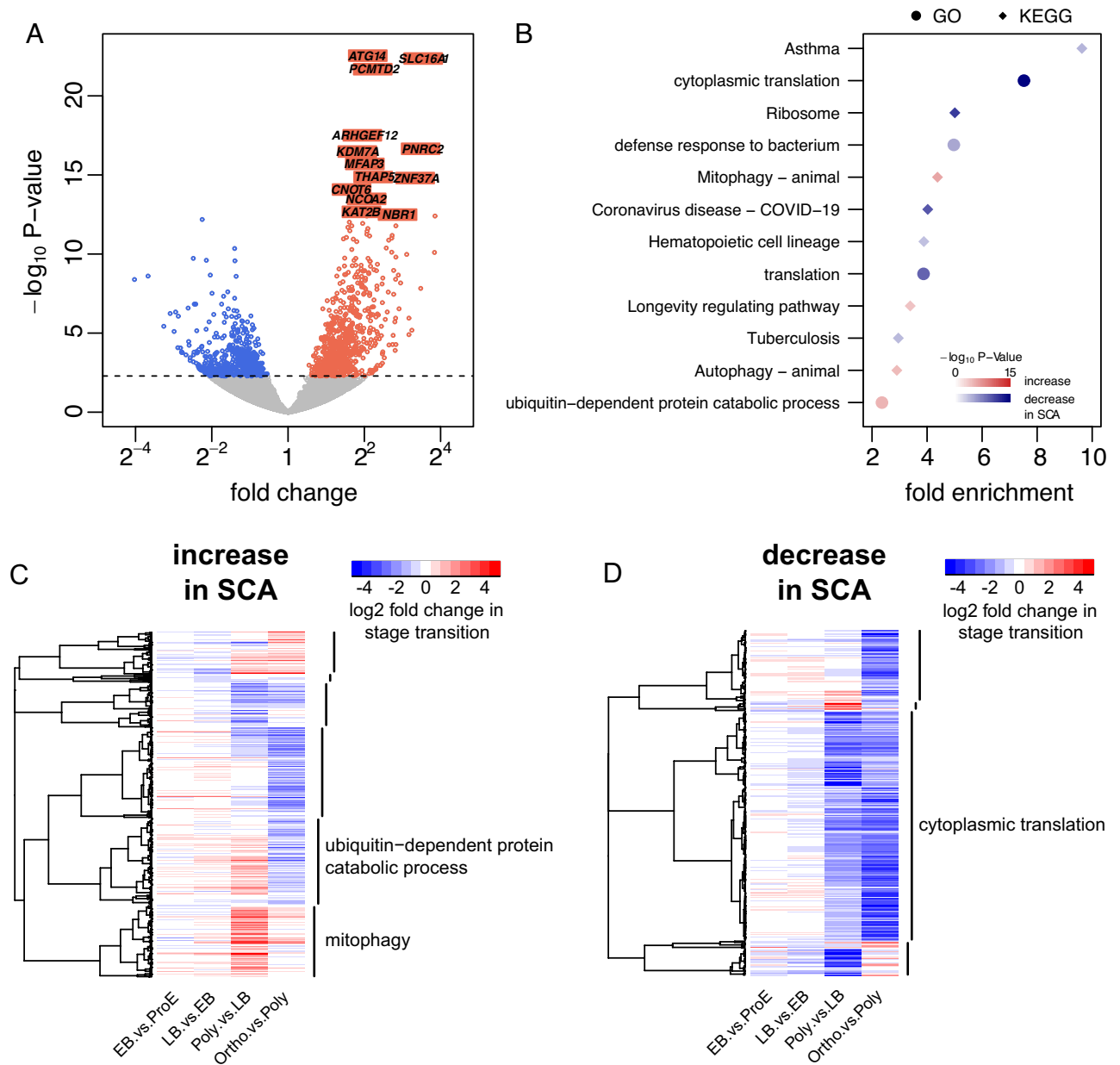
A flow chart of data analyses is presented in Fig. 1.

**Gene expression changes in SCA reticulocytes.** Five SCA patients (mean  $\pm$  standard deviation [SD] age  $43 \pm 8$  years; 80% female; mean  $\pm$  SD hemoglobin concentration  $73 \pm 13$  g/L) and 5 African-American control individuals without *HBB* mutation (mean  $\pm$  SD age  $44 \pm 8$  years; 100% female; mean  $\pm$  SD hemoglobin concentration  $130 \pm 7.8$  g/L) from the University of Illinois at Chicago (UIC) were studied. All were under steady state. SCA patients received neither hydroxyurea treatment nor blood transfusion in the three months prior to participating in the study. Principal components analysis revealed *HBB* genotype as the primary segregating variable of the studied transcriptomes (SFig. 1). At 5% false discovery rate (FDR), 1226 of 11,813 analyzed genes (10.4%) were differentially expressed in SCA compared to controls, 741 with increased expression (STable 1) and 485 with decreased expression (STable 2) in SCA. The top one percent most significant differential expressions all showed increase in SCA (Fig. 2A). Genes with increased expression in SCA are enriched in Gene Ontology (GO) biological process “ubiquitin-dependent protein catabolic process” and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway “mitophagy—animal”, “autophagy—animal”, and “longevity regulating pathway”. Genes with decreased expression in SCA are enriched in GO “cytoplasmic translation”, “translation”, and “defense response to bacterium”, as well as KEGG “ribosome”, “coronavirus disease—COVID-19”, “asthma”, “tuberculosis”, and “hematopoietic cell lineage” (STable 3, Fig. 2B). Upregulation of “mitophagy” and downregulation of “cytoplasmic translation” are consistent with clearance of mitochondria and ribosomes in terminal differentiation<sup>21</sup>.

Transcripts in reticulocytes are remnants of preceding erythroblasts. To assess the developmental context of the genes differentially expressed in SCA (hemoglobin SS) reticulocytes, we examined how their expressions changed in stage transitions in human hemoglobin AA erythroblasts as reported by An and colleagues<sup>13</sup>. An’s study applied fluorescence-activated sorting-based methods to purify erythroblasts and examined transitions spanning the proerythroblastic (ProE), early basophilic (EB), late basophilic (LB), polychromatic (Poly), and orthochromatic (Ortho) stages. Across the genes analyzed in both studies, the median reads per kilobase million in An et al. were correlated with the median fragments per kilobase million in this study with Spearman’s rho 0.40, 0.44, 0.50, 0.65, 0.68 for ProE, EB, LB, Poly, and Ortho, respectively. The increasing correlation of gene expression profile between reticulocytes and differentiating erythroblasts is consistent with the position of reticulocytes in erythroblast lineage. It also indicates that our fractioning method of reticulocyte transcripts (“Methods”) captured relevant variation in erythroblasts.



**Figure 1.** Flow chart of data analyses.



**Figure 2.** Differential genes detected at 5% FDR in hemoglobin SS reticulocytes from 5 SCA patients compared to hemoglobin AA reticulocytes from 5 African American control individuals. **(A)** Volcano plot of differential genes. The top one percent of the most significant genes were labeled. *ATG14* autophagy related 14, *SLC16A1* solute carrier family 16 member 1, *PCMTD2* protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 2, *ARHGEF12* Rho guanine nucleotide exchange factor 12, *PNRC2* proline rich nuclear receptor coactivator 2, *KDM7A* lysine demethylase 7A, *MFAP3* microfibrillar associated protein 3, *THAP5* THAP domain containing 5, *ZNF37A* zinc finger protein 37A, *CNOT6* CCR4-NOT transcription complex subunit 6, *NCOA2* nuclear receptor coactivator 2, *KAT2B* lysine acetyltransferase 2B, *NBR1* NBR1 autophagy cargo receptor. **(B)** GO biological processes and KEGG pathways enriched with differential genes. **(C,D)** Hierarchical clustering of genes whose expressions were **(C)** increased or **(D)** decreased in hemoglobin SS reticulocytes, using correlation distance of expression changes across stage transitions in hemoglobin AA erythroblasts by An et al. Enriched GO biological processes with adjusted P-value < 0.05 are labelled per gene cluster.

Of the 1226 differentially expressed genes in hemoglobin SS reticulocytes, 910 exhibited expression change at 5% FDR from an earlier stage to a consecutive later stage in hemoglobin AA erythroblasts in An et al.<sup>13</sup> (STable 4). The direction of gene expression change in hemoglobin SS reticulocytes vs hemoglobin AA controls was consistent with the direction of expression change in the maturation of hemoglobin AA erythroblasts for decreasing proportions of genes: 82%, 83%, 70%, and 52% for stage transitions from ProE through Ortho (STable 4, test for trend in proportions P-value =  $2.3 \times 10^{-16}$ ). We assumed here that upregulation in SCA relative to controls of genes normally upregulated in an erythroid stage transition, or vice versa, indicates accentuated erythroid

differentiation in SCA. The observed trend therefore suggests heightened erythroid differentiation in SCA relative to controls in ProE, which gradually diminishes toward Ortho. This diminishing erythroid differentiation appeared to be driven by genes upregulated in SCA. Figure 2C shows clustering according to stage transitions of genes whose expressions were increased in SCA. Of these genes, 338 (56%) also showed increased expression from LB to Poly while only 167 (28%) had increased expression from Poly to Ortho. As shown in Fig. 2C, genes whose expressions were increased consistently from LB through Ortho are enriched in GO “mitophagy”, whereas genes whose expressions were increased from LB to Poly but then decreased from Poly to Ortho are enriched in GO “ubiquitin-dependent protein catabolic process”. Figure 2D shows clustering according to stage transitions of genes whose expressions were decreased in SCA. Of these genes, 161 (52%) also showed decreased expression from LB to Poly and 213 (69%) from Poly to Ortho. Genes whose expressions were decreased consistently from LB through Ortho are enriched in GO “cytoplasmic translation” (Fig. 2D).

**SCA-specific gene expression changes versus expression changes shared with CE.** Heightened erythropoiesis occurs in both SCA and CE, the latter a condition of constitutive HIF activation in the absence of anemia. A comparison of gene expression changes in SCA (versus healthy African-American individuals) to gene expression changes in CE (versus healthy Chuvash individuals) allowed us to classify SCA-specific changes attributable to compensatory erythropoiesis versus changes not necessarily induced by anemia. Seventeen *VHL*<sup>R200W</sup> homozygotes (mean  $\pm$  SD age  $48 \pm 14$  years; 71% female; mean  $\pm$  SD hemoglobin concentration  $190 \pm 33$  g/L) and 13 *VHL* wild-type individuals (mean  $\pm$  SD age  $47 \pm 15$  years; 77% female; mean  $\pm$  SD hemoglobin concentration  $129 \pm 9$  g/L) from Chuvashia, Russia were studied. We analyzed in CE reticulocytes the expression of the same 11,813 genes as in SCA reticulocytes. Of the 1347 differentially expressed genes detected in CE reticulocytes at 5% FDR, 339 shared the same direction of expression change with SCA reticulocytes, 220 having increased and 119 decreased expression (STables 1 and 2). Figure 3A shows the top 20 most significant differential genes in CE whose expression changes were shared in SCA.

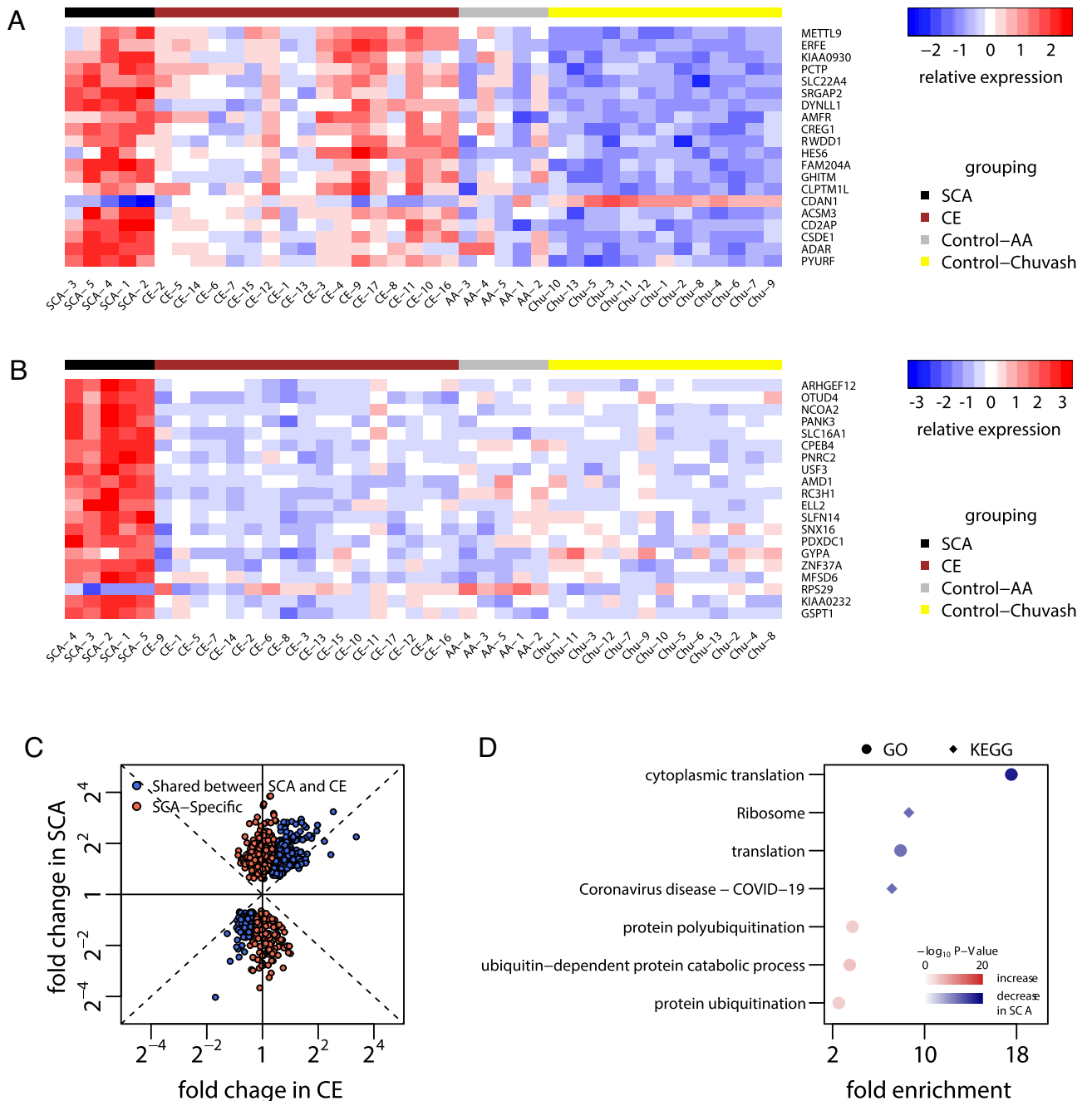
We performed a formal genome-wide comparison between SCA and CE of reticulocyte gene expression changes relative to their respective control individuals, using the likelihood ratio test of the interaction effect between disease status (patients versus healthy individuals) and population (African Americans versus Chuvash people). We detected 730 genes exhibiting different expression changes between SCA and CE at 5% FDR, among which we further identified 453 genes that were among the 1226 differentially expressed genes in SCA reticulocytes and that showed no evidence of similarly altered expression in CE as defined by nominal P-value  $< 0.05$ . Of these 453 genes with “SCA-specific” expression changes, 300 had increased expression and 153 had decreased expression in SCA reticulocytes (STables 1 and 2). Figure 3B shows the top 20 most significant genes whose expression changes differed between SCA and CE.

The magnitude of expression changes that were shared between SCA and CE reticulocytes was in general greater in SCA (Fig. 3C), likely reflecting a robust hypoxic response to severe anemia in SCA. Genes with shared expression changes between SCA and CE had no enriched GO or KEGG at 5% FDR. Genes with SCA-specific increased expression are enriched in GO “ubiquitin-dependent protein catabolic process”, “protein polyubiquitination”, and “protein ubiquitination”. Genes with SCA-specific decreased expression are enriched in GO “cytoplasmic translation” and “translation” and KEGG “ribosome” and “coronavirus disease—COVID-19” (STable 5, Fig. 3D).

**Erythroid expression variation that correlates with hemoglobin concentration and hematocrit in SCA.** PBMCs are easily obtainable peripheral blood cells normally comprised of lymphocytes, monocytes, dendritic cells<sup>22</sup> and a small amount of erythroid progenitors<sup>23,24</sup>. PBMCs from SCA contain elevated amounts of erythroid progenitors due to accentuated erythropoiesis<sup>9,17</sup>. We correlated erythroid expression variation with clinical variation employing deconvolution analysis on PBMC expression data. Deconvolution analysis may use cell/tissue-specific marker genes to dissect gene expression variation in samples with cell/tissue heterogeneity<sup>19</sup>. The study included 151 SCA patients 19–72 years old under steady state (mean  $\pm$  SD age  $37 \pm 12$  years; 53% female; mean  $\pm$  SD hemoglobin concentration  $86 \pm 14$  g/L). For the 1226 genes differentially expressed in SCA reticulocytes, we first assessed their PBMC expression correlation with a 16-gene signature of erythroid progenitors<sup>20</sup> in the 151 SCA patients. We identified 673 genes whose PBMC expression correlated with the erythroid gene signature, i.e., genes suitable for deconvolution analysis.

We correlated erythroid expression of the 673 genes in PBMCs with quartiles of hemoglobin concentration or hematocrit, adjusting for hemolysis defined as percent reticulocytes (“Methods”). When hemolysis is accounted for, more abundant hemoglobin or higher hematocrit signify more successful terminal differentiation and thus serve as proxies of more efficient erythropoiesis. The analysis was stratified by hydroxyurea treatment and restricted to 131 patients without blood transfusion. We detected 54 genes whose increased erythroid expression consistently correlated with higher hemoglobin concentration and higher hematocrit at 5% FDR (STable 6, Fig. 4). The expression of these genes was all increased in SCA reticulocytes relative to healthy individuals. Of the 673 genes subjected to deconvolution analysis, the 54 genes correlated with hemoglobin concentration and hematocrit are enriched with SCA-specific, increased expression (OR = 2.9, Fisher’s exact test P-value = 0.00063). The 54 genes are also enriched with GO annotations with key word “protein ubiquitination” (*DCAF10*, *FBXO30*, *FBXO9*, *GCLC*, *LNX2*, *NEDD4L*, *PINK1*, *UBR2*), “protein polyubiquitination” (*MARCHF8*, *NEDD4L*, *TRIP12*, *UBE2H*), and “ubiquitin-dependent protein catabolic process” (*FBXL4*, *FBXO9*, *GCLC*, *NEDD4L*, *PINK1*, *TBL1XR1*, *TRIP12*, *UBAP1*, *UBE2H*, *UBR2*, *ZRANB1*) (OR = 4.2, P-value =  $7.5 \times 10^{-5}$ ).

**An erythroid eQTL associating with severe pain episodes in SCA.** eQTL are genetic polymorphisms associated with gene expression variation. We mapped erythroid eQTL on the same PBMC expression

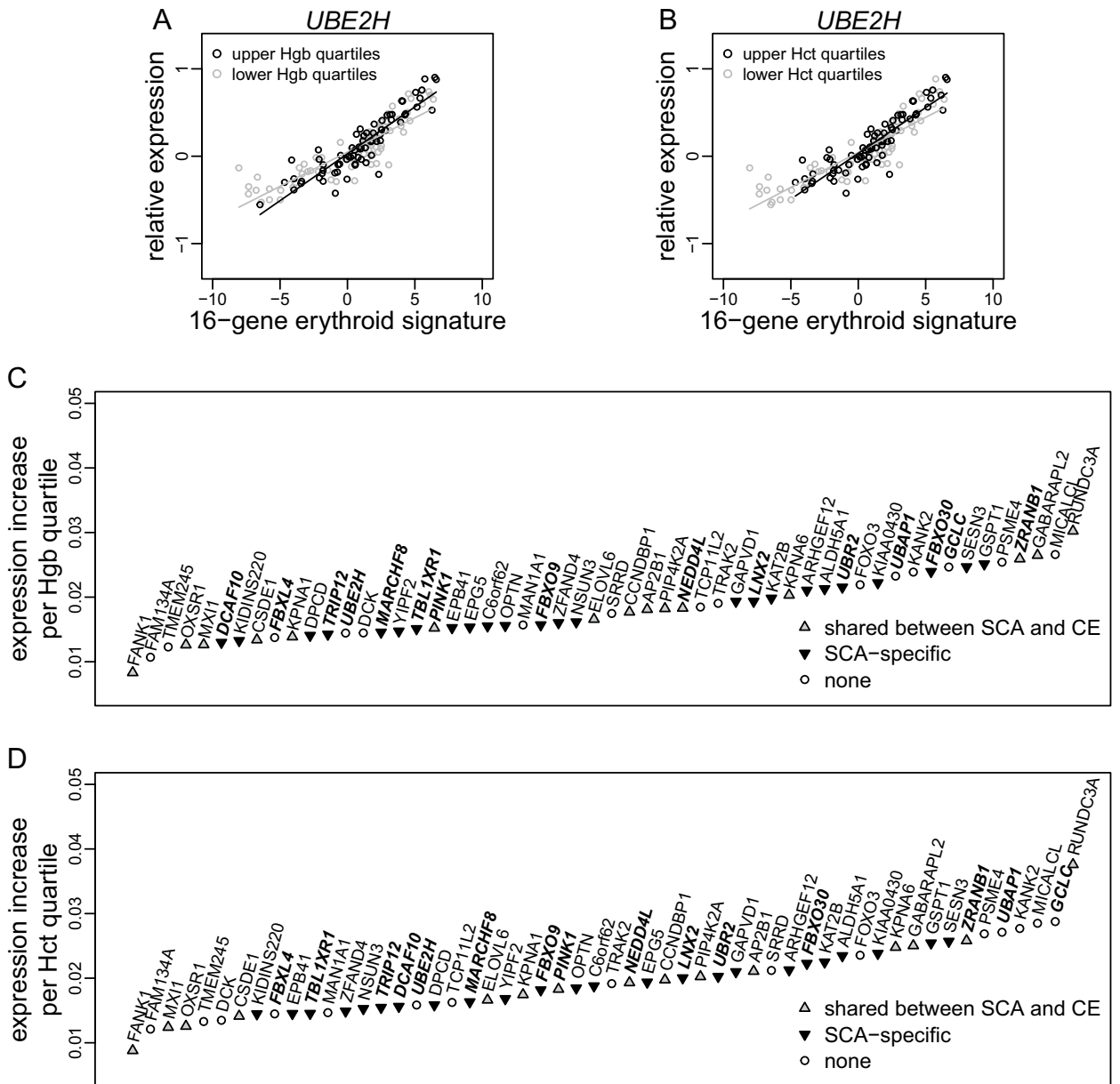


**Figure 3.** SCA-specific expression changes. (A,B) Hierarchical clustering of individuals within groups, using Euclidean distance of gene expression levels of (A) the top 20 most significant differential genes in CE whose expression changes were shared in SCA and (B) the top 20 most significant genes whose expression changes differed between SCA and CE. Individuals are grouped as SCA, CE, African American controls (Control-AA), and Chuvash controls (Control-Chuvash). (C) Scatter plot of fold change in CE (x-axis) and SCA (y-axis). (D) GO biological processes and KEGG pathways enriched with SCA-specific differential genes.

data from the 151 SCA patients for the 673 genes suitable for deconvolution analysis as in the previous section (“Methods”). At 5% FDR, we detected 180 associations between 180 local single nucleotide polymorphisms (SNPs), i.e., SNPs located less than 100 Kb away from gene ends, and erythroid expression of 18 genes. The Genotype-Tissue Expression (GTEx) project generates categories of eQTL across diverse tissues<sup>25</sup> which currently do not cover erythroblasts. Because of substantial sharing of eQTL across tissues<sup>26</sup>, a comparison of erythroid eQTL with GTEx eQTL from diverse tissues may assure the validity of the deconvolution analysis. Of the 180 gene-SNP associations, 117 (65%) overlapped with eQTL associations from GTEx analysis version 8 for at least one tissue, representing a tenfold enrichment and supporting our erythroid eQTL results.

We tested genetic association of the 180 SNPs with severe pain episodes, defined as having three or more pain episodes that required Emergency Room visit or hospitalization in the past 12 months<sup>27</sup>, in 774 SCA patients





**Figure 4.** Correlation of erythroid expression of 54 genes with hemoglobin concentration (Hgb) and hematocrit (Hct). (A,B) Illustration for the *UBE2H* gene as an example, where the slope of PBMC expression of *UBE2H* versus the 16-gene erythroid signature was greater in patients with higher levels of (A) Hgb and (B) Hct. The significant erythroid signature by Hgb or Hct interaction effect (see “Methods”) indicates elevated erythroid expression of *UBE2H* with higher Hgb or Hct. (C,D) Increase in erythroid expression per quartile increase in (C) Hgb and (D) Hct. Gene symbols in bold text refer to genes annotated by GO keyword “protein ubiquitination”, “protein polyubiquitination”, and “ubiquitin-dependent protein catabolic process”.

from the Walk-PHaST<sup>28</sup>, PUSH<sup>29</sup>, and UIC cohorts (STable 7). The genetic association results are present in STable 8. The most significant association, between an eQTL of *LNX2* and severe pain episodes, reached Bonferroni-corrected  $P < 0.05$  (STable 8). *LNX2* encodes ligand of Numb-protein X 2 and is annotated by GO “protein ubiquitination”. Expression of *LNX2* was specifically increased in SCA reticulocytes (STable 1). Erythroid expression of *LNX2* correlated with higher hemoglobin concentration and hematocrit (STable 6). The eQTL, rs7992289, is 83,620 bp down stream of *LNX2*. The G allele of the eQTL (frequency = 0.29; imputation quality  $r^2 = 0.97$ ) associated with greater expression level of *LNX2* ( $\beta = 0.043$ , 95% CI 0.024–0.062,  $P = 2.4 \times 10^{-5}$ ) and with severe pain episodes (OR = 1.7, 95 CI 1.3–2.2,  $P = 3.9 \times 10^{-5}$ ) (Table 1).

## Discussion

Analysis of the reticulocyte transcriptome in SCA suggests heightened differentiation in early erythroblasts that diminishes toward transition to orthochromatic stage. Comparison with reticulocyte transcriptome in CE reveals SCA-specific upregulation of genes involved in ubiquitin-dependent protein degradation and protein ubiquitination. Deconvolution analysis in SCA PBMCs identified 54 genes whose erythroid expression correlated with erythropoiesis efficiency, which are enriched with SCA-specific upregulation and with functional annotation of ubiquitin-dependent protein degradation and protein ubiquitination. An erythroid eQTL of one such gene, *LNX2*, was found to associate with severe pain episodes in 774 SCA patients. Our findings raise a hypothesis that increased ubiquitin-dependent protein degradation and protein ubiquitination in late erythroid differentiation may link improved erythroblast survival with altered presentation of surface proteins in SCA erythroid cells that favors adherence to the microvasculature, a feature of SCA pain crisis.

A marked number of genes whose expressions are normally decreased in transition from polychromatic to orthochromatic stage had increased expression in SCA reticulocytes (Fig. 2C), implying distinct cellular conditions in the latest stages of SCA erythroid differentiation. These include genes that normally showed transiently increased expression in the polychromatic stage and that are enriched in GO “ubiquitin-dependent protein catabolic process” (Fig. 2C). Transient transcriptional activation of genes encoding ubiquitin conjugation enzymes has been observed in normal erythropoiesis<sup>30</sup>, potentially to eliminate misfolded proteins at the peak of hemoglobin synthesis<sup>31</sup>. Polymerization of HbS can occur in the hemoglobin synthesis phase and induce apoptosis in erythroblasts derived from SCA patients<sup>9,32</sup>. Apoptosis of erythroblasts in SCA was associated with cytoplasmic sequestration of HSP70 similar to  $\beta$ -thalassemia<sup>33</sup>, which can be rescued by induction of HbF synthesis to interfere with HbS polymerization<sup>9</sup>. In  $\beta$ -thalassemia, the ubiquitin–proteasome system plays a critical role in proteolytic control of globin chain balance<sup>34</sup>. Our study identified 54 genes whose erythroid expression correlated with more effective erythropoiesis as reflected by higher hemoglobin concentration and hematocrit. The 54 genes are enriched with annotation keyword “ubiquitin-dependent protein catabolic process”, “protein polyubiquitination”, and “protein ubiquitination”, suggesting that upregulation of these processes may alleviate oxidative damages<sup>35</sup> resulted from HbS polymerization in SCA erythroblasts.

The ubiquitin system contributes to membrane remodeling, in addition to protein quality control and autophagy which eliminates mitochondria and ribosomes<sup>36</sup>. Ubiquitination of plasma membrane aquaporin-1 is required for its endocytosis, sorting, and secretion<sup>37</sup>. In the presence of bulk protein damage, the ubiquitin–proteasome system can be overloaded and impact the turnover of plasma membrane proteins<sup>38</sup>. In  $\beta$ -thalassemia, membrane assembly of spectrin, band 3, and band 4.1 in erythroblasts is disrupted<sup>39</sup>. In SCA, the adhesive markers glycoprotein IV<sup>40</sup> and  $\alpha$ 4 $\beta$ 1 integrin receptor<sup>15</sup> are uniquely displayed on the reticulocyte surface<sup>4</sup>. Hydroxyurea treatment, which reduces HbS polymerization by elevating HbF level, decreases these markers on the surface of reticulocytes and erythroblasts in children with SCA<sup>41</sup> and decreases the adhesion of sickle erythrocytes to thrombospondin and laminin in vitro<sup>42</sup>. We found that the G allele of SNP rs7992289 associates with increased erythroid expression of *LNX2* and increased risk for severe pain episodes in SCA patients. The *LNX2* gene encodes an E3 ubiquitin ligase that was shown to promote CD8 $\alpha$  ubiquitination, endocytosis, and degradation in lysosomes<sup>43</sup>. The results imply that the *LNX2* gene may have contrary impacts in SCA erythroblast survival and reticulocyte maturation. *LNX2* was also shown to modulate the ubiquitination status of the neuronal glycine transport GlyT2<sup>44</sup>, the gene expression level of which in PBMCs was found to positively correlate with pain frequency in our patients<sup>45</sup>. Future studies are needed to confirm these associations and elucidate *LNX2* substrates in SCA erythroblasts.

There are several limitations of this study. First, it compared relatively small number of hemoglobin SS patients and hemoglobin AA controls using reticulocyte RNA-Seq. The detection of 10.4% of the analyzed genes as differentially expressed genes with this small sample size, however, suggests that large transcriptional differences accumulate during SCA erythropoiesis. The transcriptional differences can be attributed to distinct selection pressure in SCA bone marrow<sup>9</sup>, disturbed progression of erythroblast maturation, and varied rates of RNA degradation, in addition to transcriptional changes at aligned differentiating stages. As such, the second limitation of this study is the lack of biological resolution with the current experimental assays. Our reticulocyte purification protocol inclusively collects reticulocytes in peripheral blood (“Methods”). A finer delineation of the heterogeneous reticulocytes may be accomplished by single cell RNA-Seq. A more resource-demanding approach using in vitro derived erythroblasts may further supplement the study. Third, the comparison between SCA and CE was confounded with both ethnicity and experimental batch effects. Our study design ameliorates such confounding in that patients were paired with control individuals matched for ethnicity, age, gender, and experimental batch. Despite of these limitations, our results suggest that a deeper understanding of transcriptome dysregulation of erythroid cells in SCA has potential to shed novel insights on SCA pathogenesis.

Trait	Walk-PHaSST			Push			UIC			Combined		
	OR (95% CI)	P-value	N	OR (95% CI)	P-value	N	$\beta$ or OR (95% CI)	P-value	N	OR (95% CI)	P-value	N
Erythroid Expression of <i>LNX2</i>							0.043 (0.024–0.062)	$2.4 \times 10^{-5}$	151			
Severe pain episodes $\geq 3$	1.4 (0.98–1.9)	0.062	399	2.5 (1.5–4.2)	0.00048	207	1.9 (1.1–3.4)	0.022	168	1.7 (1.3–2.2)	$3.9 \times 10^{-5}$	774

**Table 1.** Association of the G allele of rs7992289 with erythroid gene expression of *LNX2* using linear model and with severe pain episodes  $\geq 3$  using binary logistic model. OR odds ratio.

## Methods

**Ethics approval and consent to participate.** Institutional Review Board (IRB) of UIC approved the study “Genomic and Biomarker Studies in Sickle Cell Disease” for the UIC cohort (protocol # 2010-1125). IRBs of Albert Einstein College of Medicine, Columbia University, Children’s Hospital Oakland, Children’s Hospital Pittsburgh, Hammersmith Hospital London, Howard University, Johns Hopkins University, NHLBI, University of Colorado, and UIC approved the study “Treatment of Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy Walk-PHaSST” for Walk-PHaSST (protocol # 2007-0936). IRBs of Children’s National Medical Center, Howard University Hospital, University of Michigan Ann Arbor, and Pulmonary and Vascular Medicine Branch of the NHLBI approved the study “Pulmonary Hypertension and the Hypoxic Response in SCD” for PUSH (protocol # 2011-1071). All subjects provided written informed consent to participate. This study was designed and carried out in accordance with the Declaration of Helsinki.

**Patients and patient cohorts.** Five SCA patients and 5 African-American control individuals from UIC were studied for differential gene expression in SCA reticulocytes. Seventeen CE patients and 13 control individuals from Chuvashia, Russia were studied for differential gene expression in CE reticulocytes. To assess clinical correlation with erythroid gene expression we studied 151 SCA patients from UIC whose PBMC expressions were profiled by arrays previously<sup>46</sup>. To assess genetic association with disease manifestations, we studied 774 SCA patients: 399 patients from Walk-PHaSST, a screening study that enrolled patients  $\geq 12$  years old at 10 centers from 2008–2009<sup>28</sup>, 207 patients from PUSH, which enrolled patients 3–20 years old at four centers from 2006 to 2010<sup>29,47</sup>, and 168 UIC patients including the 151 selected for PBMC expression analysis. All patients were in their usual state of health.

**Reticulocyte RNA-Seq data and analysis.** Whole blood (8–16 mL) was collected using ACD as an anticoagulant. Plasma and platelets were isolated by gentle centrifugation at 400 g. Blood was further layered on ficoll-histopaque (density 1.077 g/mL) for separation of mononuclear cells and lymphocytes. The lower layer of packed erythrocytes were transferred to a clean tube and preferentially lysed with 4 volumes of 0.144 M ammonium chloride, 1 mM ammonium bicarbonate solution. Granulocytes were separated by centrifugation and the supernatant containing reticulocyte RNA was precipitated by acid titration to pH 5.1. Precipitated RNA was preserved in Trizol and stored at  $-80\text{ }^{\circ}\text{C}$ <sup>48</sup>. Library construction was performed using the Illumina TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero Human/Mouse/Rat. Strand-specific libraries were constructed for 100 or 125 bp paired-end sequencing to 30–45 million read pairs per sample using Illumina HiSeq 2500 or 4000 platform. RNA extraction, library construction, and sequencing were performed at the University of Chicago for SCA and control samples in one batch and at the University of Utah for CE and control samples in three batches. Sequencing data were mapped to human reference genome version GRCh37 annotated by Gencode version 24, using STAR<sup>49</sup>. The effect of SCA or CE on autosomal gene expression was assessed by the likelihood ratio test based on the negative binomial distribution as implemented in DESeq2<sup>50</sup>. Batch effects were adjusted in the analysis of CE and controls. Genome-wide comparison of SCA with CE was assessed by the likelihood ratio test of the interaction effect between disease status (patients vs. healthy individuals) and population (African Americans versus Chuvash people), controlling for batch effects. NIH DAVID<sup>51</sup> was used for gene enrichment in GO<sup>52</sup> and KEGG<sup>53</sup>.

**Deconvolution of erythroid gene expression using PBMC data.** Messenger RNA was purified from PBMCs from 151 UIC SCA patients and profiled on Affymetrix Gene 2.0 ST Array. Probes with unique perfect alignment to human genome assembly GRCh37 were selected. Probes that interrogate transcripts from multiple genes were removed, as well those that contain SNPs with  $\geq 1\%$  minor allele frequency (MAF) in dbSNP release 149 to avoid signal noise. Probe intensities were log<sub>2</sub> transformed, background corrected, and quantile normalized. The corresponding probe mean across samples was then subtracted from the probe intensity. Gene expression level was summarized as average intensity across probes according to Gencode version 24. Experimental batch effects were adjusted using an empirical Bayes method<sup>54</sup>.

We assessed correlation of erythroid gene expression in PBMCs with hemoglobin concentration using the model:

$$y_{ij} = u_i + x_j^{\text{ery}} \cdot \beta_i^{\text{ery}} + x_j^{\text{hgb}} \cdot \beta_i^{\text{hgb}} + x_j^{\text{hemo}} \cdot \beta_i^{\text{hemo}} + x_j^{\text{int\_hgb}} \cdot \beta_i^{\text{int\_hgb}} + x_j^{\text{int\_hemo}} \cdot \beta_i^{\text{int\_hemo}} + \varepsilon_{ij}.$$

Here  $y_{ij}$  denotes expression of gene  $i$  for patient  $j$ .  $u_i$  denotes mean expression of gene  $i$ .  $x_j^{\text{ery}}$ ,  $x_j^{\text{hgb}}$ ,  $x_j^{\text{hemo}}$ ,  $x_j^{\text{int\_hgb}}$ ,  $x_j^{\text{int\_hemo}}$  denote the number of erythroid progenitors, hemoglobin concentration in quartiles, level of hemolysis reflected by square root transformed percent reticulocytes, erythroid progenitor by hemoglobin concentration interaction, and erythroid progenitor by hemolysis interaction, respectively for patient  $j$ .  $\beta$  s denotes the corresponding coefficients for gene  $i$ .  $\varepsilon_{ij}$  denotes residual error. The P-value of the erythroid progenitor by hemoglobin concentration interaction term, representing erythroid gene expression correlation with hemoglobin concentration, was estimated by the F test. Correlation of erythroid gene expression with hematocrit was similarly assessed. The number of erythroid progenitors was approximated as the first principal component of PBMC gene expression of 16 marker genes<sup>20</sup>.

For erythroid eQTL mapping, PBMC gene expression levels were analyzed using the model:

$$y_{ij} = u_i + x_j^{\text{ery}} \cdot \beta_i^{\text{ery}} + x_j^{\text{dose}} \cdot \beta_i^{\text{dose}} + x_j^{\text{int}} \cdot \beta_i^{\text{int}} + x_j^{\text{sv}} \cdot \beta_i^{\text{sv}} + \varepsilon_{ij}.$$

Here  $x_j^{\text{dose}}$  denotes SNP allele dosage,  $x_j^{\text{int}}$  denotes erythroid progenitor by allele dosage interaction, and  $x_j^{\text{sv}}$  denotes surrogate variables<sup>55</sup> for patient  $j$ . The P-value of the erythroid progenitor by allele dosage interaction term, representing erythroid gene expression correlation with allele dosage, was estimated by the F test.



**Genetic association between erythroid eQTL and clinical manifestations.** Genomic DNA isolated from PBMCs was genotyped using Illumina Human 610-Quad SNP array for Walk-PHaST and PUSH patients and Affymetrix Axiom Pan African array for UIC patients. Samples having a genotype rate < 95% were removed. SNPs deviating from Hardy Weinberg equilibrium ( $P$ -value < 0.0001) or with MAF < 0.01 were removed. Population outliers were identified by principal components analysis using PLINK 2.0<sup>56</sup> and removed from analysis. Kinship coefficients were estimated using KING-robust<sup>57</sup>. First to third-degree relatives as well as duplicates were randomly thinned to one individual per family. Genotypes were phased using SHAPIT2<sup>58</sup> and imputed to 1000 genomes project phase 3 data with African reference panels using IMPUTE2<sup>57</sup>. To provide an empirical estimation of the imputation quality, chromosome 22 from the UIC patients was re-imputed with 5% of array-genotyped SNPs randomly masked, using the same parameters for whole genome imputation. The median (interquartile range) of concordance rate between imputation and array-genotyping is 0.97 (0.95–0.99).

SNPs with MAF > 0.1 and imputation  $r^2$  > 0.9 were further tested for erythroid expression association in UIC patients. Erythroid eQTL were assessed for association with severe pain episodes. We examined binary outcome of having 3 or more pain episodes requiring emergency room visits or hospital admission over 12 months prior to a baseline visit<sup>27</sup>, using logistic regression with covariates of age, gender,  $\alpha$ -thalassemia as copies of deletion, and population stratification.  $P$ -value was estimated by Wald test. Meta-analysis used inverse variance based method<sup>58</sup>.

R version 4.2.1<sup>59</sup> was used in regression analyses. In all of the analyses, multiple comparison was controlled by Benjamini–Hochberg procedure<sup>60</sup> and significance was defined as adjusted  $P$ -value < 0.05, unless specified otherwise.

## Data availability

The datasets generated during the current study are available in the GEO repository with accession number GSE232221: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232221>.

Received: 5 May 2023; Accepted: 3 August 2023

Published online: 08 August 2023

## References

- Pauling, L. *et al.* Sickle cell anemia a molecular disease. *Science (New York)* **110**, 543–8 (1949).
- Connes, P. *et al.* The role of blood rheology in sickle cell disease. *Blood Rev.* **30**, 111–118 (2016).
- Browning, J. A. *et al.* The effect of deoxygenation on whole-cell conductance of red blood cells from healthy individuals and patients with sickle cell disease. *Blood* **109**, 2622–2629 (2007).
- Carden, M. A., Fasano, R. M. & Meier, E. R. Not all red cells sickle the same: Contributions of the reticulocyte to disease pathology in sickle cell anemia. *Blood Rev.* **40**, 100637 (2020).
- Ballas, S. K. & Marcolina, M. J. Hyperhemolysis during the evolution of uncomplicated acute painful episodes in patients with sickle cell anemia. *Transfusion* **46**, 105–110 (2006).
- Kato, G. J., Steinberg, M. H. & Gladwin, M. T. Intravascular hemolysis and the pathophysiology of sickle cell disease. *J. Clin. Investig.* **127**, 750–760 (2017).
- Hasegawa, S. *et al.* Sickling of nucleated erythroid precursors from patients with sickle cell anemia. *Exp. Hematol.* **26**, 314–319 (1998).
- Wu, C. J. *et al.* Evidence for ineffective erythropoiesis in severe sickle cell disease. *Blood* **106**, 3639–3645 (2005).
- El Hoss, S. *et al.* Fetal hemoglobin rescues ineffective erythropoiesis in sickle cell disease. *Haematologica* **106**, 2707–2719 (2021).
- Mathias, L. A. *et al.* Ineffective erythropoiesis in beta-thalassemia major is due to apoptosis at the polychromatophilic normoblast stage. *Exp. Hematol.* **28**, 1343–1353 (2000).
- Ludwig, L. S. *et al.* Transcriptional states and chromatin accessibility underlying human erythropoiesis. *Cell. Rep.* **27**, 3228–40 e7 (2019).
- Yan, H. *et al.* Developmental differences between neonatal and adult human erythropoiesis. *Am. J. Hematol.* **93**, 494–503 (2018).
- An, X. *et al.* Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood* **123**, 3466–3477 (2014).
- Goh, S. H. *et al.* The human reticulocyte transcriptome. *Physiol. Genom.* **30**, 172–178 (2007).
- Swerlick, R. A., Eckman, J. R., Kumar, A., Jeitler, M. & Wick, T. M. Alpha 4 beta 1-integrin expression on sickle reticulocytes: Vascular cell adhesion molecule-1-dependent binding to endothelium. *Blood* **82**, 1891–1899 (1993).
- Papageorgiou, D. P. *et al.* Simultaneous polymerization and adhesion under hypoxia in sickle cell disease. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9473–9478 (2018).
- Zhang, X. *et al.* Hypoxic response contributes to altered gene expression and precapillary pulmonary hypertension in patients with sickle cell disease. *Circulation* **129**, 1650–1658 (2014).
- Ang, S. O. *et al.* Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. *Nat. Genet.* **32**, 614–621 (2002).
- Zhao, Y. & Simon, R. Gene expression deconvolution in clinical samples. *Genome Med.* **2**, 93 (2010).
- Zhang, X. *et al.* Peripheral blood mononuclear cells show prominent gene expression by erythroid progenitors in diseases characterized by heightened erythropoiesis. *Br. J. Haematol.* **190**, e42–e45 (2020).
- Barminko, J., Reinholt, B. & Baron, M. H. Development and differentiation of the erythroid lineage in mammals. *Dev. Comp. Immunol.* **58**, 18–29 (2016).
- Kleiveland, C. R. Peripheral blood mononuclear cells. In *The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models* (eds Verhoeckx, K., Cotter, P., Lopez-Exposito, I. *et al.*) 161–7 (Springer International Publishing, 2015).
- Cheadle, C. *et al.* Erythroid-specific transcriptional changes in PBMCs from pulmonary hypertension patients. *PLoS ONE* **7**, e34951 (2012).
- Prchal, J. T. (ed.) *Chapter 58: Primary and Secondary Erythrocytoses/Polycythemia* 10th edn. (McGraw Hill, 2021).
- GTE Consortium. The Genotype-Tissue Expression (GTEx) project. *Nature genetics* **2013**;45:580–5.
- GTE Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (New York, NY)* **2015**;348:648–60.
- Darbari, D. S. *et al.* Markers of severe vaso-occlusive painful episode frequency in children and adolescents with sickle cell anemia. *J. Pediatr.* **160**, 286–290 (2012).
- Machado, R. F. *et al.* Hospitalization for pain in patients with sickle cell disease treated with sildenafil for elevated TRV and low exercise capacity. *Blood* **118**, 855–864 (2011).

29. Minniti, C. P. *et al.* Elevated tricuspid regurgitant jet velocity in children and adolescents with sickle cell disease: association with hemolysis and hemoglobin oxygen desaturation. *Haematologica* **94**, 340–347 (2009).
30. Wefes, I. *et al.* Induction of ubiquitin-conjugating enzymes during terminal erythroid differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 4982–4986 (1995).
31. Casale, G. P., Khairallah, E. A. & Grasso, J. A. An analysis of hemoglobin synthesis in erythropoietic cells. *Dev. Biol.* **80**, 107–119 (1980).
32. Park, S. Y. *et al.* Pathologic angiogenesis in the bone marrow of humanized sickle cell mice is reversed by blood transfusion. *Blood* **135**, 2071–2084 (2020).
33. Arlet, J. B. *et al.* HSP70 sequestration by free alpha-globin promotes ineffective erythropoiesis in beta-thalassaemia. *Nature* **514**, 242–246 (2014).
34. Shaeffer, J. R. Turnover of excess hemoglobin alpha chains in beta-thalassaemic cells is ATP-dependent. *J. Biol. Chem.* **258**, 13172–13177 (1983).
35. Amer, J. *et al.* Red blood cells, platelets and polymorphonuclear neutrophils of patients with sickle cell disease exhibit oxidative stress that can be ameliorated by antioxidants. *Br. J. Haematol.* **132**, 108–113 (2006).
36. Narendra, D. P. *et al.* PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS Biol.* **8**, e1000298 (2010).
37. Blanc, L. *et al.* The water channel aquaporin-1 partitions into exosomes during reticulocyte maturation: Implication for the regulation of cell volume. *Blood* **114**, 3928–3934 (2009).
38. Khandros, E. & Weiss, M. J. Protein quality control during erythropoiesis and hemoglobin synthesis. *Hematol. Oncol. Clin. North Am.* **24**, 1071–1088 (2010).
39. Aljurf, M. *et al.* Abnormal assembly of membrane proteins in erythroid progenitors of patients with beta-thalassemia major. *Blood* **87**, 2049–2056 (1996).
40. Sugihara, K., Sugihara, T., Mohandas, N. & Hebbel, R. P. Thrombospondin mediates adherence of CD36+ sickle reticulocytes to endothelial cells. *Blood* **80**, 2634–2642 (1992).
41. Odievre, M. H. *et al.* Modulation of erythroid adhesion receptor expression by hydroxyurea in children with sickle cell disease. *Haematologica* **93**, 502–510 (2008).
42. Hillery, C. A., Du, M. C., Wang, W. C. & Scott, J. P. Hydroxyurea therapy decreases the in vitro adhesion of sickle erythrocytes to thrombospondin and laminin. *Br. J. Haematol.* **109**, 322–327 (2000).
43. D'Agostino, M. *et al.* Ligand of Numb proteins LNX1p80 and LNX2 interact with the human glycoprotein CD8alpha and promote its ubiquitylation and endocytosis. *J. Cell Sci.* **124**, 3545–3556 (2011).
44. de la Rocha-Munoz, A. *et al.* E3 ubiquitin ligases LNX1 and LNX2 are major regulators of the presynaptic glycine transporter GlyT2. *Sci. Rep.* **9**, 14944 (2019).
45. Han, J. *et al.* Risk factors for vitamin D deficiency in sickle cell disease. *Br. J. Haematol.* **181**, 828–835 (2018).
46. Desai, A. A. *et al.* Association of circulating transcriptomic profiles with mortality in sickle cell disease. *Blood* **129**, 3009–3016 (2017).
47. Nouraei, M. *et al.* Tricuspid regurgitation velocity and other biomarkers of mortality in children, adolescents and young adults with sickle cell disease in the United States: The PUSH study. *Am. J. Hematol.* **95**, 766–774 (2020).
48. Prchal, J. T., Cashman, D. P. & Kan, Y. W. Hemoglobin Long Island is caused by a single mutation (adenine to cytosine) resulting in a failure to cleave amino-terminal methionine. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 24–27 (1986).
49. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**, 15–21 (2013).
50. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
51. Dennis, G. Jr. *et al.* DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3 (2003).
52. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
53. Ogata, H., Goto, S., Fujibuchi, W. & Kanehisa, M. Computation with the KEGG pathway database. *Biosystems* **47**, 119–128 (1998).
54. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics (Oxford, England)* **8**, 118–127 (2007).
55. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England)* **28**, 882–883 (2012).
56. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
57. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics (Oxford, England)* **26**, 2867–2873 (2010).
58. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics (Oxford, England)* **26**, 2190–2191 (2010).
59. R CORE Team. R: A language and environment for statistical computing. 2021.
60. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B (Methodol.)* **57**, 289–300 (1995).

## Acknowledgements

This work was supported in part by Amgen Corporation, Incyte Corporation and by NIH Research Grants P01CA108671, 1P50HL118006, 1R01HL125005 and R01HL111656. The content is solely the responsibility of the authors and does not necessarily represent the official view of NHLBI or NIH.

## Author contributions

X.Z., J.T.P., and V.R.G. conceived and designed study; X.Z. designed and carried out data analysis; T.H., G.M., A.S. and V.R.G. collected patient samples; J.S. generated CE transcriptome library and with B.N.S. purified reticulocytes and coordinated RNA-seq experiments; S.N., R.F.M., M.T.G., and V.R.G. contributed cohort data from UIC, PUSH and Walk-PHaSST; X.Z., J.T.P. and V.R.G. interpreted data; X.Z., J.T.P. and V.R.G. wrote paper with all authors providing critical comments.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-40039-2>.

**Correspondence** and requests for materials should be addressed to X.Z., J.T.P. or V.R.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023, corrected publication 2023