

Aberrant splicing contributes to severe α -spectrin–linked congenital hemolytic anemia

Patrick G. Gallagher, ... , Susan J. Baserga, Vincent P. Schulz

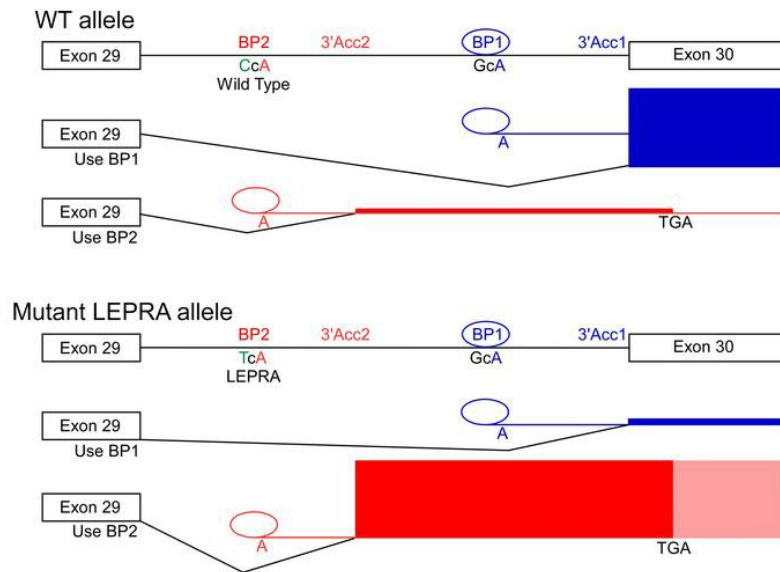
J Clin Invest. 2019;129(7):2878-2887. <https://doi.org/10.1172/JCI127195>.

Research Article

Genetics

Hematology

Graphical abstract



Find the latest version:

<https://jci.me/127195/pdf>



Aberrant splicing contributes to severe α -spectrin-linked congenital hemolytic anemia

Patrick G. Gallagher,^{1,2,3} Yelena Maksimova,¹ Kimberly Lezon-Geyda,¹ Peter E. Newburger,⁴ Desiree Medeiros,⁵ Robin D. Hanson,⁶ Jennifer Rothman,⁷ Sara Israels,⁸ Donna A. Wall,⁹ Robert F. Sidonio Jr.,¹⁰ Colin Sieff,¹¹ L. Kate Gowans,¹² Nupur Mittal,¹³ Roland Rivera-Santiago,¹⁴ David W. Speicher,¹⁴ Susan J. Baserga,^{2,15} and Vincent P. Schulz¹

¹Department of Pediatrics, ²Department of Genetics, and ³Department of Pathology, Yale University School of Medicine, New Haven, Connecticut, USA. ⁴Department of Pediatrics, University of Massachusetts Medical School, Worcester, Massachusetts, USA. ⁵Kapiolani Medical Center for Women & Children, Honolulu, Hawaii, USA. ⁶Cardinal's Kids Cancer Center, St. Louis, Missouri, USA. ⁷Department of Pediatrics, Duke University Medical Center, Durham, North Carolina, USA. ⁸Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Manitoba, Canada. ⁹Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada. ¹⁰Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia, USA. ¹¹Harvard Medical School, Dana-Farber and Boston Children's, Cancer and Blood Disorders Center, Boston, Massachusetts, USA. ¹²Beaumont Children's Hospital, Royal Oak, Michigan, USA. ¹³Department of Pediatrics, Rush University Medical Center, Chicago, Illinois, USA. ¹⁴The Center for Systems and Computational Biology and Molecular and Cellular Oncogenesis Program, The Wistar Institute, Philadelphia, Pennsylvania, USA. ¹⁵Departments of Molecular Biophysics and Biochemistry and Therapeutic Radiology, Yale University School of Medicine, New Haven, Connecticut, USA.

The etiology of severe hemolytic anemia in most patients with recessive hereditary spherocytosis (rHS) and the related disorder hereditary pyropoikilocytosis (HPP) is unknown. Whole-exome sequencing of DNA from probands of 24 rHS or HPP kindreds identified numerous mutations in erythrocyte membrane α -spectrin (*SPTA1*). Twenty-eight mutations were novel, with null alleles frequently found in *trans* to missense mutations. No mutations were identified in a third of *SPTA1* alleles (17/48). WGS revealed linkage disequilibrium between the common rHS-linked α^{BH} polymorphism and a rare intron 30 variant in all 17 mutation-negative alleles. In vitro minigene studies and in vivo splicing analyses revealed the intron 30 variant changes a weak alternate branch point (BP) to a strong BP. This change leads to increased utilization of an alternate 3' splice acceptor site, perturbing normal α -spectrin mRNA splicing and creating an elongated mRNA transcript. In vivo mRNA stability studies revealed the newly created termination codon in the elongated transcript activates nonsense-mediated decay leading to spectrin deficiency. These results demonstrate that a unique mechanism of human genetic disease contributes to the etiology of a third of rHS cases, facilitating diagnosis and treatment of severe anemia and identifying a new target for therapeutic manipulation.

Introduction

Hereditary spherocytosis (HS), the most common inherited hemolytic anemia in Northern Europeans, affects people of every ethnic background worldwide. A quarter of HS cases exhibit recessive inheritance (rHS) (1–5). Clinically, rHS patients are more severely affected than patients with typical, dominant HS, often presenting in infancy or early childhood with life-threatening, hemolytic anemia (1, 2, 5). Many of these patients are transfusion-dependent. Erythrocyte membranes from patients with rHS and the related disorder hereditary pyropoikilocytosis (HPP) are spectrin-deficient, which leads to destabilization of the lipid bilayer (1, 2, 5–7). The degree of spectrin deficiency correlates with the degree of hemolysis and the clinical response to splenectomy (6). Biochemical and genetic studies have implicated defects of α -spectrin in most rHS and HPP patients, but in most, the molecular basis of disease is unknown (1, 5, 7–9).

In normal erythroid cells, α -spectrin is synthesized in large excess, with α -spectrin synthesized in excess of β -spectrin by a 3:1 to 4:1 ratio (10–12). Thus patients with 1 normal and 1 production-defective α -spectrin allele are biochemically and clinically normal, because normal α -spectrin production remains in excess and normal amounts of $\alpha\beta$ -spectrin heterodimers are assembled on the membrane. Patients who are homozygotes or compound heterozygotes for defects of α -spectrin production will suffer from rHS (13). Similarly, patients who are compound heterozygotes for a structurally defective α -spectrin allele and a defective α -spectrin production allele will suffer from HPP (8, 9, 14).

We have studied the genetic basis of α -spectrin-linked anemia in 24 kindreds with rHS or HPP, including the first α -spectrin-deficient rHS kindreds described by Agre et al., as well as a group of transfusion-dependent patients, using whole-exome sequencing (WES) (1, 2, 15). We identified numerous novel mutant α -spectrin (*SPTA1*) alleles, including null alleles and missense mutations, primarily in the $\alpha\beta$ -spectrin self-association site. In 17 of 48 alleles, only 1 or no deleterious *SPTA1* mutations were identified. Whole-genome sequencing (WGS) of selected individuals identified the α^{LEPRA} allele, a rare, poorly characterized variant identified in a patient with rHS associated with an elongated α -spectrin mRNA transcript, in *trans* to other mutations (13). Because WT

► Related Commentary: p. 2655

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2019, American Society for Clinical Investigation.

Submitted: January 3, 2019; **Accepted:** April 23, 2019; **Published:** June 4, 2019.

Reference information: *J Clin Invest.* 2019;129(7):2878–2887.

<https://doi.org/10.1172/JCI127195>.

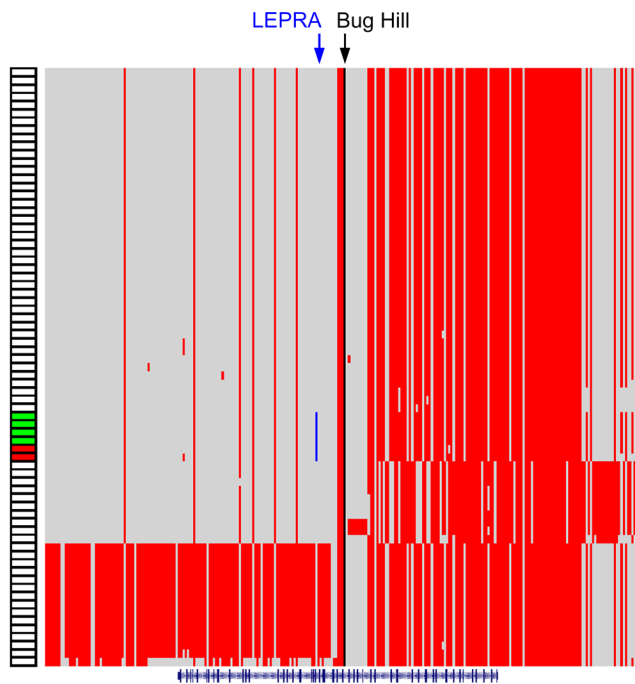


Figure 1. Haplotyping at the SPTA1 locus. Haplotype analysis around the *SPTA1* locus of all 71 alleles with the α^{BH} variant, Ala970Asp, in the 1000 Genomes 2015 database and 2 rHS patients heterozygous for the α^{BH} variant. The α^{BH} allele is denoted in black and the rare intron 30 variant, the α^{LEPRA} allele, is denoted in blue. The HG19 reference sequence is denoted in gray and the nonreference sequence is shown in red. The *SPTA1* gene is shown at the bottom. The 4 individuals in the 2015 1000 Genomes database and both rHS patients who carry the α^{LEPRA} allele are shown in green and red, respectively.

erythroid cells also carry the elongated transcript described with the α^{LEPRA} allele, its role in the pathogenesis of α -spectrin deficiency and inherited anemia has been unclear.

In a series of in vitro minigene splicing studies and in vivo splicing assays, we demonstrate that the α^{LEPRA} variant changes a weak alternate branch point (BP) to a strong BP. This change leads to increased utilization of an alternate 3' splice acceptor site, perturbing normal α -spectrin mRNA splicing and creating the elongated mRNA transcript. In vivo mRNA stability studies in erythroid cells rendered homozygous for the α^{LEPRA} variant via CRISPR/Cas9-based gene editing revealed that a newly created termination codon in the elongated transcript activates nonsense-mediated decay leading to spectrin deficiency. These studies indicate that a novel mechanism of human genetic disease contributes to the etiology of a third of rHS cases, facilitating disease diagnosis and treatment and identifying a new target for therapeutic manipulation.

Results

Individuals from 24 kindreds from 3 groups with inherited hemolytic anemia were studied using WES. One group was individuals from 5 spectrin-deficient rHS kindreds from the original reports of HS by Agre et al. and 2 spectrin-deficient rHS kindreds described by Tse et al. (1–3, 15). Clinical and laboratory details of these patients have been previously described. A second group was from 7 kindreds with the clinical diagnosis of rHS or HPP referred to the Yale Center for Blood Disorders (YCBD). Clinical and laboratory details of these patients are provided in Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/JCI127195DS1>. The unifying features of this group are that the patients had severe hemolytic anemia and their erythrocyte membranes were spectrin-deficient. The third group was composed of 10 patients from 9 kindreds with transfusion-dependent (TD) anemia referred to the YCBD. Clinical details of these patients are provided in Supplemental Table 2. Prenatal onset

of severe hemolytic anemia or onset in the immediate neonatal period, requiring in utero or early postnatal transfusion therapy, respectively, were common. All TD patients required chelation for early onset iron overload. Two patients in this group died. One died from complications of anemia in the neonatal period, while a similarly affected sibling remains transfusion dependent. The other patient died from hepatic complications of transfusion-related iron overload at 17 months of age.

WES identified numerous mutations in α -spectrin (*SPTA1*), with 28 of 48 alleles carrying novel variants (Table 1). No deleterious mutations were identified in other erythrocyte membrane protein genes. Null *SPTA1* alleles were frequently found in *trans* to missense mutations. Two TD patients had deleterious mutations in both *SPTA1* alleles; one with nonsense mutations in *trans* died of liver failure associated with iron overload (patient 16), the other with nonsense and splicing mutations in *trans* remains transfusion dependent. Six missense mutations were identified, with 4 in the α -spectrin self-association site, critical for membrane stability. Modeling studies predicted 3 mutations (R45G, L49R, R28H) disrupt tetramer formation while a fourth, M70R, although not directly located in the tetramer interface, creates steric hindrance with the laterally associated spectrin chain in bivalent tetramers (Supplemental Figure 1 and Supplemental Figure 2). One TD kindred patient homozygous for the R45G mutation had a sibling homozygous for the same mutation die in the immediate perinatal period due to complications of anemia (patient 21).

In 17 of 48 alleles, no deleterious mutations were identified. All 17 alleles carried the common α -spectrin^{Bug Hill} polymorphism, (rs35948326, NM_003126.2:c.2909C>A, NP_003117.2:p. Ala970Asp, α^{BH}), previously associated with α -spectrin-linked rHS (Supplemental Figure 3) (3). WGS of 2 rHS patients heterozygous for the α^{BH} variant was performed. Comparison to samples in the 1000 Genomes database revealed a single, common *SPTA1* variant, α^{LEPRA} , chr1:158613314C>T, present on only 4 of 4610 alleles in 1000 Genomes. In all 4, it was heterozygous and in *cis* to the α^{BH} allele. This rare intron 30 variant was previously described in a rHS patient in *trans* to a *SPTA1* nonsense mutation (13). Analysis of α^{BH} haplotypes revealed 3 predominant patterns (Figure 1) with the haplotype of the rHS patients identical to the 4 heterozygous α^{LEPRA} individuals. All 17 mutation-negative alleles carried the α^{LEPRA} mutation, in *trans* to other mutations in 16 patients (Table 1). The proband lacking any deleterious *SPTA1* alleles was homozygous for the α^{LEPRA} allele. Thus, a third of all mutant *SPTA1* alleles associated with severe hemolytic anemia carried the α^{LEPRA} variant.

In the original description of the α^{LEPRA} variant, RT-PCR of reticulocyte RNA demonstrated the α^{LEPRA} allele was associated with an elongated α -spectrin mRNA transcript (13). This transcript

Table 1. Patient diagnosis and genetic characteristics

| Patient number/ group | Diagnosis | | Mutation type | Mutation location | Mutation | 1000G ^A | ExAc ^B | CADD phred ^C | Mutation taster ^D |
|--------------------------|-----------|----------|------------------|----------------------|----------------------------|--------------------|-------------------------|-------------------------|---------------------------------|
| 1/1 | rHS | Allele 1 | Splicing | Intron 44 | c.6418-1G>A | Not found | Not found | 27.20 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 2/1 | rHS | Allele 1 | Missense | Exon 2 | c.209T>G;p.M70R | Not found | Not found | 9.85 | 0.97 |
| | | Allele 2 | Missense | Exon 2 | c.209T>G;p.M70R | Not found | Not found | 9.85 | 0.97 |
| 3/1 | rHS | Allele 1 | Splicing | Intron 30 | c.4339-2A>C | Not found | 8.40 × 10 ⁻⁶ | 25.40 | 1.00 |
| | | Allele 2 | Missense | Exon 2 | c.209T>G;p.M70R | Not found | Not found | 9.85 | 0.97 |
| 4/1 | rHS | Allele 1 | Nonsense | Exon 37 | c.5203C>T;p.R1735* | Not found | Not found | 37.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 5/1 | rHS | Allele 1 | Nonsense | Exon 22 | c.3139C>T;p.R1044* | Not found | Not found | 35.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 6/1 | rHS | Allele 1 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 7/1 | rHS | Allele 1 | Splicing | Exon 4 | c.531+1G>A | Not found | Not found | 25.90 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 8/2 | rHS | Allele 1 | Deletion | Exon 48 | c.6756delG;p.M2252fs | Not found | Not found | NA | NA |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 9/2 | rHS | Allele 1 | Nonsense | Exon 44 | c.6335T>G;p.L2112* | Not found | Not found | 46.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 10/2 | rHS | Allele 1 | Nonsense | Exon 39 | c.5530C>T;p.R1844* | Not found | Not found | 39.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 11/2 | rHS | Allele 1 | Splicing | Intron 37 | c.5311-2A>G | Not found | Not found | 24.80 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 12/2 | HPP | Allele 1 | Insertion | Exon 4 | c.460_461insTGC;p.L154insL | Not found | 8.28 × 10 ⁻⁶ | NA | NA |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 13/2 | HPP | Allele 1 | Missense | Exon 2 ^E | c.83G>A;p.R28H | 0.000399 | 2.49 × 10 ⁻⁵ | 34.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 14/2 | HPP | Allele 1 | Nonsense | Exon 41 | c.5804G>A;p.W1935* | Not found | Not found | 50.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 15/3 | TD | Allele 1 | deletion | Exon 52 | c.7199delG;p.G2400fs | Not found | Not found | NA | NA |
| | | Allele 2 | Nonsense | Exon 37 | c.5215C>T;p.Q1739* | Not found | Not found | 38.00 | 1.00 |
| 16/3 | TD | Allele 1 | Nonsense | Exon 2 | c.178C>T;p.R60* | Not found | 8.31 × 10 ⁻⁶ | 34.00 | 1.00 |
| | | Allele 2 | Nonsense | Exon 41 | c.5804G>A;p.W1935* | Not found | Not found | 50.00 | 1.00 |
| 17/3 | TD | Allele 1 | Nonsense | Exon 31 | c.4398T>G;p.Y1466* | Not found | Not found | 42.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 18/3 | TD | Allele 1 | Deletion | Exon 41 | c.5814delT;p.D1938fs | Not found | Not found | NA | NA |
| | | Allele 2 | Missense | Exon 8 | c.1037A>G;p.E346G | Not found | Not found | 22.70 | 0.84 |
| 19/3 | TD | Allele 1 | Nonsense | Exon 36 | c.5102T>A;p.L1701* | Not found | Not found | 43.00 | 1.00 |
| | | Allele 2 | Missense | Exon 6 | c.803G>A;p.R268Q | Not found | Not found | 28.00 | 1.00 |
| 20/3 | TD | Allele 1 | Missense | Exon 2 | c.146T>G;p.L49R | Not found | Not found | 28.50 | 1.00 |
| | | Allele 2 | Splicing | Intron 45 | c.6531-12C>T | 0.227835 | 0.2564 | NA | NA |
| 21/3 | TD | Allele 1 | Missense | Exon 2 | c.133A>G;p.R45G | Not found | Not found | 25.90 | 1.00 |
| | | Allele 2 | Missense | Exon 2 | c.133A>G;p.R45G | Not found | Not found | 25.90 | 1.00 |
| 22/3 | TD | Allele 1 | Nonsense | Exon 27 | c.3823C>T;p.Q1275* | Not found | Not found | 35.00 | 1.00 |
| | | Allele 2 | Missense | Exon 2 ^E | c.83G>A;p.R28H | 0.000399 | 2.49 × 10 ⁻⁵ | 34.00 | 1.00 |
| 23/3 | TD | Allele 1 | Nonsense | Exon 23 | c.3267T>A;p.Y1089* | Not found | Not found | 40.00 | 1.00 |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |
| 24/3 | TD | Allele 1 | Deletion | | c.340delA;p.Thr114Glnfs | Not found | Not found | NA | NA |
| | | Allele 2 | Splicing | Intron 30 | c.4339-99C>T | 0.000799 | Not found | NA | NA |

^AFrequency in 1000 Genomes database. ^BFrequency in Exome Aggregation Consortium (ExAc) database. ^CPredicted score of the deleteriousness of mutations by the CADD algorithm. ^DScore of the disease-causing potential of DNA variants. ^EPreviously reported (1–3).

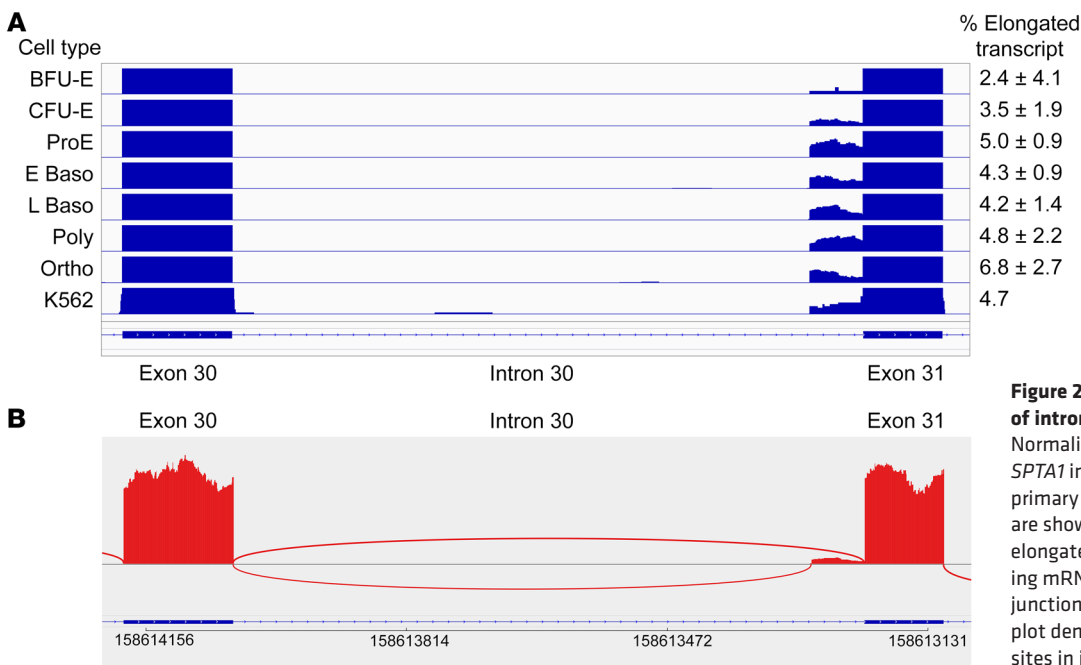


Figure 2. An elongated transcript of intron 30 of the SPTA1 gene. (A) Normalized RNA-seq profiles of the SPTA1 intron 30 region in human WT, primary erythroid cells, and K562 cells are shown. Estimated percentages of elongated, partial intron 30-containing mRNA transcripts based on splice junction reads are shown. (B) A Sashimi plot demonstrates utilization of splice sites in intron-spanning reads.

contained 70 nt from the 3' end of intron 30 and was predicted to lead to frameshift and premature chain termination (13). Located 29 nt upstream of the alternate acceptor splice, it was unclear if or how the α^{LEPRA} variant influenced α -spectrin pre-mRNA splicing. Examination of this region in transcriptomes derived from RNA-seq of primary human WT erythroid cells at varying stages of development and differentiation reveals that an alternately spliced, elongated transcript identical to that associated with α^{LEPRA} is produced on the background of WT SPTA1 alleles, ranging from 2.4%–6.8% of total α -spectrin transcripts (Figure 2, A and B). K562 cells, which do not carry the α^{LEPRA} variant, also produce this transcript. Because WT erythroid cells carry the same elongated transcript described with the α^{LEPRA} allele, its role in the pathogenesis of α -spectrin deficiency and inherited anemia has been unclear, with some speculating it is a single nucleotide polymorphism of no functional importance.

Attempts to quantitate the elongated α -spectrin transcript in reticulocyte RNA from affected patients led to inconsistent results among patients and controls, and among individual samples from the same patient, with elongated/WT α -spectrin mRNA ratios ranging from 3.8%–5.4% for controls and 3.5%–9.7% for heterozygous α^{LEPRA} patients (not shown). This variability is likely multifactorial, reflecting variability of the contribution of the mutant SPTA1 allele in *trans* (nonsense, splicing, other mutations) due to mRNA stability, instability of the elongated mRNA transcripts, variability in amounts and differentiation stages of circulating reticulocyte mRNA (early vs. late) in individual patients, shipping-induced changes in mRNA, and limitations of PCR-based assays.

To gain insight into potential effects of the α^{LEPRA} allele on mRNA processing, a series of bioinformatic studies was performed. Analysis of SPTA1 intron 30 using Sroogle, which analyzes splicing signals and the influence of mutations on splicing (16), predicted: (a) a branch point (BP1) at the expected location 31 bp 5' of the 3' acceptor site (3'Acc1) of exon 31; (b) an alternate

upstream BP centered on an A 2 bp 3' of the α^{LEPRA} mutation (BP2); (c) a novel alternate 3' acceptor site downstream of BP2 (3'Acc2); and (d) the α^{LEPRA} variant substantially improves the probability of using BP2 (schematic in Figure 3, A and B). These findings predict that alternate BP usage between WT and α^{LEPRA} alleles influences transcript composition and abundance. Branchpointer, an algorithm that predicts and analyzes BP (17), indicated that WT BP1 and BP2 have low BP probability scores (0.51 and 0.67, respectively), whereas the α^{LEPRA} variant markedly improves the probability score of BP2 (0.94, Figure 3C), supporting the Sroogle predictions. These observations are true when comparing intron 30 BP probability scores to predicted BPs of 8180 introns of 1000 highly expressed erythroid genes (Figure 3D).

To experimentally validate the bioinformatic predictions, functional analyses were performed in minigene assays, a valuable technique for the study of splicing (18–20), and in cells manipulated via CRISPR/Cas9-based gene editing. A set of minigenes with mutation of predicted splicing elements on the background of WT and α^{LEPRA} alleles was analyzed in K562 cells (Figure 4, A and B). A WT minigene yielded only 0.78% of the elongated α -spectrin transcript (from use of BP2). In contrast, a minigene containing the α^{LEPRA} mutation yielded 42% of the elongated transcript (Figure 4C). To address the critical question of whether BP2 is a functional BP, we mutated the invariant A of BP2 to G (Figure 4C). This minigene yielded 100% WT and no elongated α -spectrin mRNA transcript. To further validate BP2 as a functional BP, we created K562 cells homozygous for G in BP2 (from central A) using gene editing. Mutation of BP2 A also eliminated the elongated α -spectrin mRNA transcript. These data support the prediction that BP2 is a functional BP.

Additional aspects of α -spectrin mRNA processing were also assessed (Figure 4). Mutation of the alternate 3' splice acceptor site led to abolition of the elongated transcript. Changing the poor BP consensus of BP1 to a U2 binding consensus sequence (21) yielded 100% WT and no elongated α -spectrin transcript. These

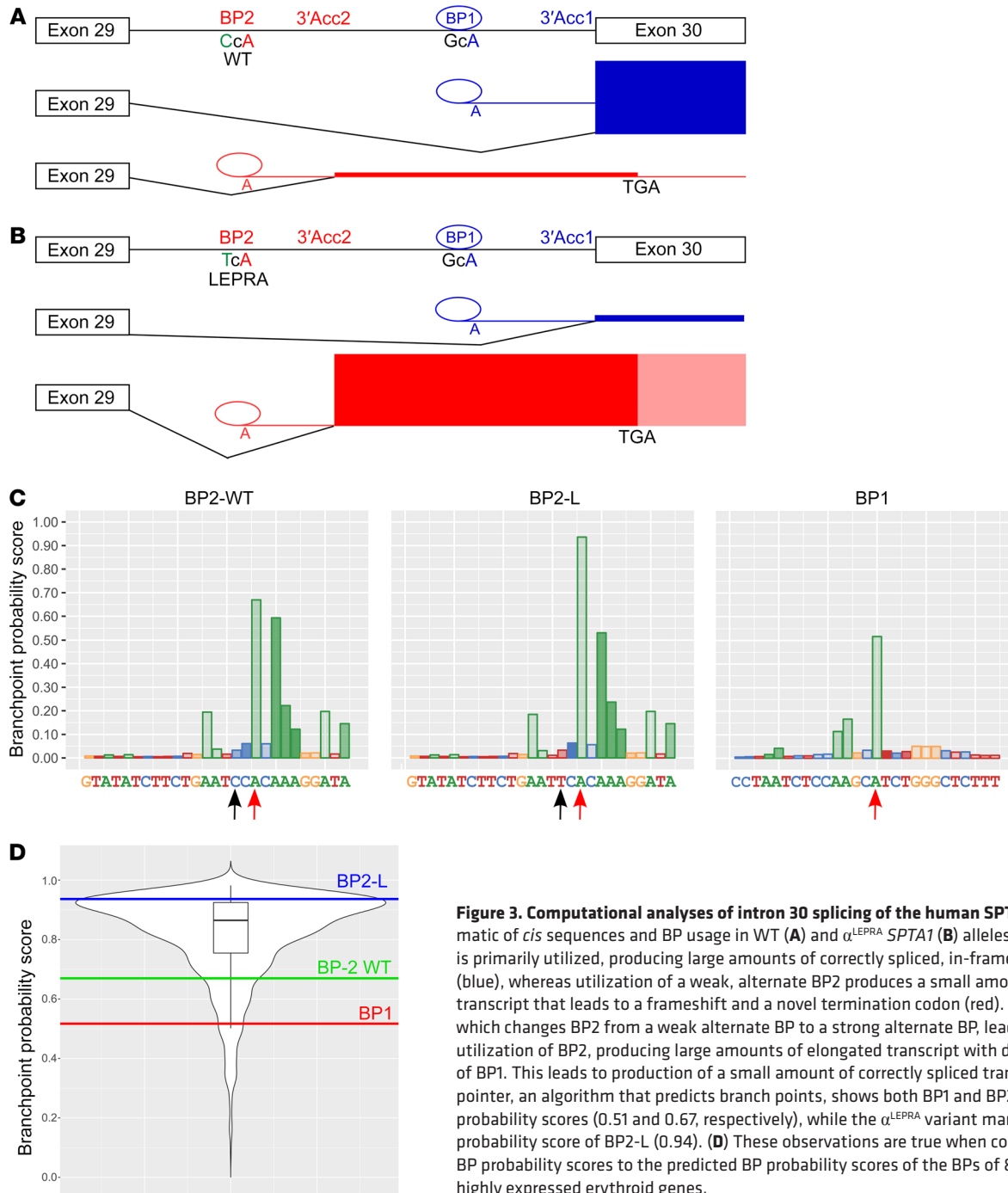


Figure 3. Computational analyses of intron 30 splicing of the human SPTA1 gene. A schematic of *cis* sequences and BP usage in WT (A) and α^{LEPRA} SPTA1 (B) alleles. In WT cells, BP1 is primarily utilized, producing large amounts of correctly spliced, in-frame mRNA transcripts (blue), whereas utilization of a weak, alternate BP2 produces a small amount of elongated transcript that leads to a frameshift and a novel termination codon (red). The α^{LEPRA} mutation, which changes BP2 from a weak alternate BP to a strong alternate BP, leads to increased utilization of BP2, producing large amounts of elongated transcript with decreased utilization of BP1. This leads to production of a small amount of correctly spliced transcript. (C) Branchpointer, an algorithm that predicts branch points, shows both BP1 and BP2-WT have low BP probability scores (0.51 and 0.67, respectively), while the α^{LEPRA} variant markedly improves the probability score of BP2-L (0.94). (D) These observations are true when comparing intron 30 BP probability scores to the predicted BP probability scores of the BPs of 8180 introns of 1000 highly expressed erythroid genes.

studies indicate the α^{LEPRA} mutation, located upstream of an alternate 3' splice acceptor site, changes a weak alternate BP to a strong BP (BP2) in the context of a poor primary BP (BP1). These changes lead to increased usage of the alternate 3' splice acceptor site, which is utilized at a low level in WT cells, creating an alternate elongated α -spectrin mRNA transcript that leads to frameshift and a novel termination codon.

The novel termination codon, 87 nt upstream of an exon-exon junction, is in a position predicted to activate nonsense-mediated decay (NMD) (22, 23). To address whether NMD of the elongated transcript is the mechanism of α -spectrin deficiency,

we created K562 cells homozygous for the α^{LEPRA} allele using gene editing (Supplemental Data) and treated them with the NMD inhibitors emetine or cycloheximide (24, 25). In homozygous α^{LEPRA} and to a lesser extent in WT cells, the total amount of elongated transcript increased relative to total α -spectrin transcript after NMD inhibition (Figure 5A).

To further validate these observations, we inserted 2 nt into the polypyrimidine tract of the 3' splice acceptor site of exon 30 (3'Acc1), putting the elongated transcript in-frame in our minigene model. On both WT and α^{LEPRA} backgrounds, the amount of the elongated α -spectrin transcript increased (Figure 5B), sup-

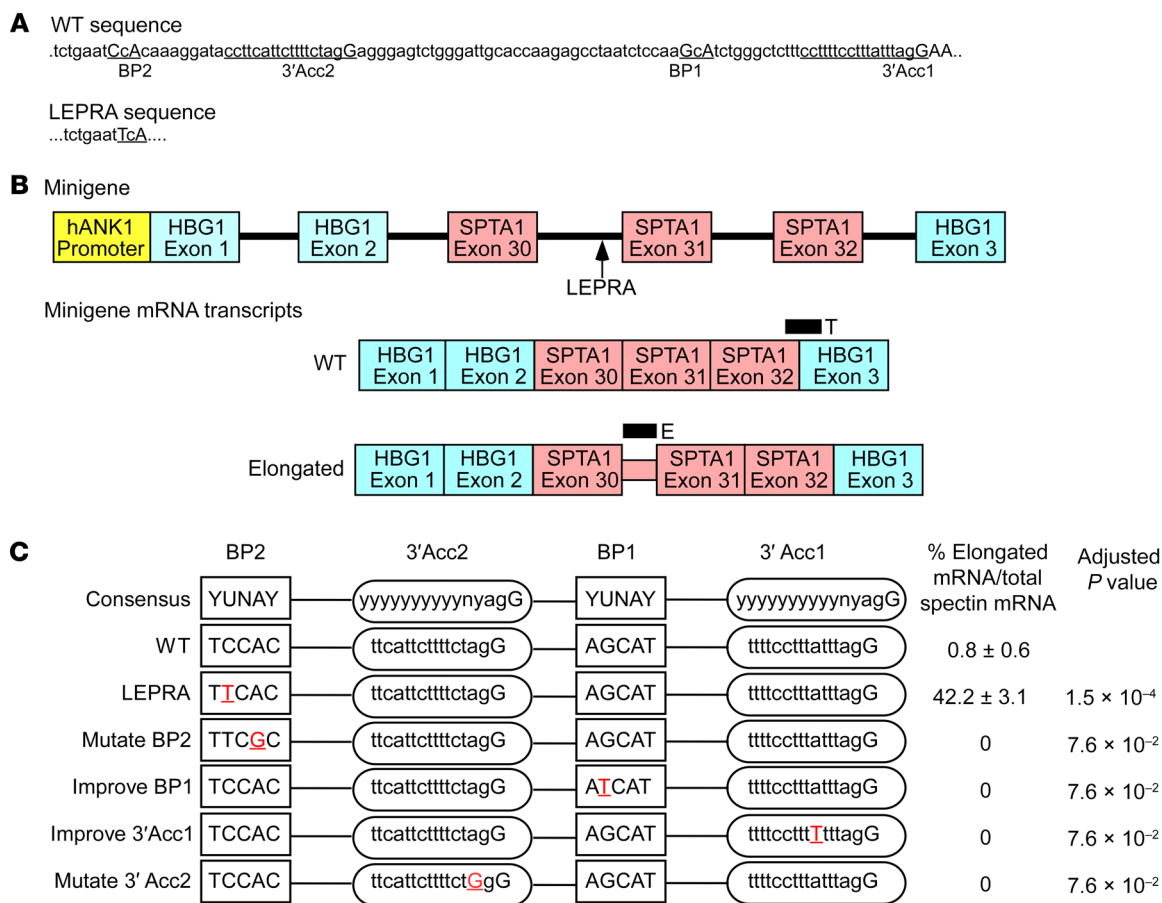


Figure 4. Minigene studies of α -spectrin intron 30 and the α^{LEPRA} variant. (A) Partial sequence of intron 30 of the *SPTA1* gene, showing the location of the α^{LEPRA} variant. (B) Each minigene construct used in minigene assays includes the *ANK1* erythroid promoter, a fragment of *SPTA1* genomic DNA inserted into intron 2 of the *HBG1* gene, and the *HBG1* 3' untranslated region and polyA signal. The hybrid *HBG1-SPTA1* transcripts derived from minigenes are shown, either WT or elongated, with the locations of Taqman probes utilized to detect total spectrin (T bar) or the unique insert of the elongated transcript (E bar). (C) Minigene results. The specific sequences utilized in minigene constructs are shown. Percentages of elongated α -spectrin transcript over total α -spectrin transcript are shown in the second column from right. The adjusted *P* value of the difference from WT is shown on the right.

porting the conclusion that NMD is responsible for low levels of α -spectrin mRNA associated with utilization of BP2.

Discussion

Mutations leading to alterations in normal mRNA splicing are a major cause of genetic disease (26–28). Most are *cis*-acting mutations located in canonical donor or acceptor splice site sequences, with others in proteins involved in the splicing process (26, 28). Only a small number of *cis* mutations in BP consensus sequences have been reported, associated with mutation or deletion of the canonical A or mutation of the -2U (29, 30). With α^{LEPRA} , strengthening of an alternate BP in the appropriate *cis* context creates a series of aberrant splicing events that lead to a decreased amount of normal mRNA transcripts. As shown here, application of WGS to the study of inherited disorders is expected to increase the number of disease-associated intronic mutations influencing splicing, including abnormalities of intronic BPs (31).

These studies identify a novel mechanism of disease and reveal the etiology underlying many cases of rHS, HPP, and TD anemia. They also resolve an unanswered question of the role of the α^{LEPRA} allele in cases of severe anemia where its contribution to

α -spectrin deficiency and disease pathogenesis had been unknown (13, 32–35). Correlating genotype with erythrocyte α -spectrin protein amounts determined by radioimmunoassay, a patient homozygous for the α^{LEPRA} allele exhibited 59% α -spectrin protein on the membrane and a compound heterozygote with α^{LEPRA} in *trans* to a null allele exhibited 34% to α -spectrin protein on the membrane (1, 2). This leads to speculation that α^{LEPRA} allele spectrin deficiency is associated with approximately 30% of normal α -spectrin protein assembled on the mature erythrocyte membrane.

These data demonstrate marked genetic heterogeneity in α -spectrin-linked hemolytic anemia, supporting the use of genomic strategies for disease diagnosis, particularly in severe and TD cases. These observations will facilitate disease diagnosis, as most diagnostic gene panels for hereditary hemolytic anemias do not include the intronic region containing the α^{LEPRA} variant nor do most whole-exome-sequencing capture panels.

When identified in affected patients, the findings are actionable, allowing the design of appropriate therapy. In our cohort, 3 patients became transfusion independent after splenectomy and 2 became transfusion independent after hematopoietic stem cell transplantation.

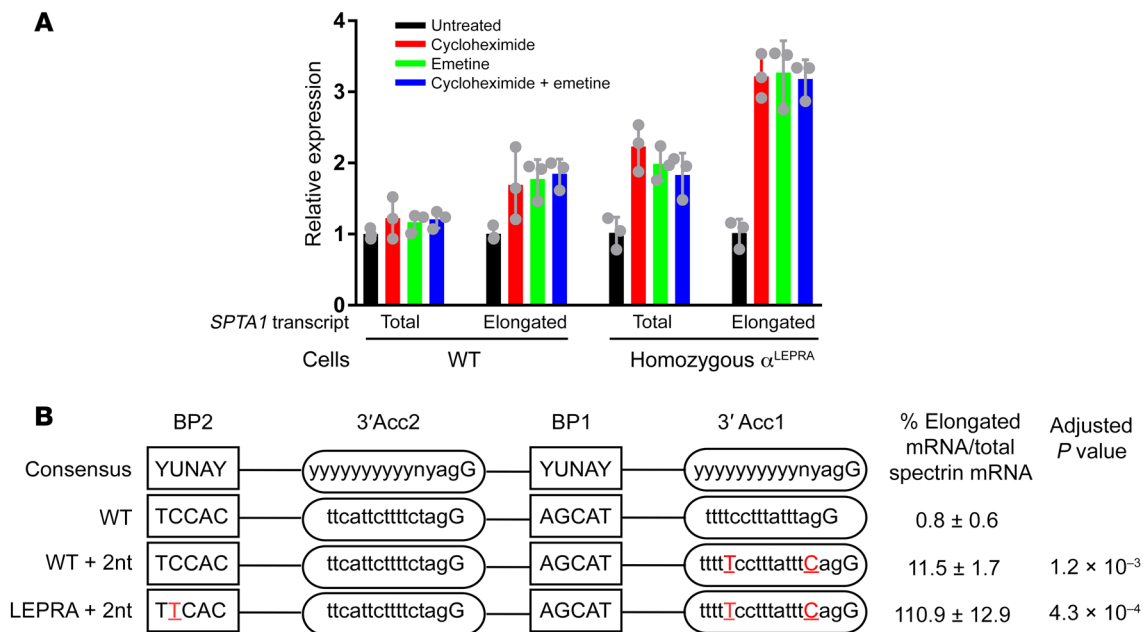


Figure 5. Analyses of nonsense-mediated decay. The influence of NMD on the stability of the elongated α -spectrin transcript in WT K562 cells and in K562 cells rendered homozygous for the α^{LEPRA} allele. **(A)** After treatment with the NMD inhibitor emetine (E) or cycloheximide (C), or both, the amounts of total α -spectrin and the elongated α -spectrin transcript were determined by real-time RT-PCR with fluorescent Taqman probes. After NMD inhibition, amounts of elongated α -spectrin transcript were increased in α^{LEPRA} cells (adjusted P values 0.0014, 0.0027, and 0.0014 for WT vs. cycloheximide, emetine, and cycloheximide + emetine, respectively) and to a lesser extent in WT cells (adjusted P values 0.084, 0.012, and 0.0047, respectively). **(B)** Minigene assay of NMD. Two nucleotides were inserted into the polypyrimidine tract of the primary 3' acceptor site of exon 30 in the minigene model to place the elongated transcript in-frame. As predicted, NMD was abrogated and the amounts of elongated α -spectrin transcript from both the WT and the α^{LEPRA} minigene constructs significantly increased. Three biologic samples each were tested in 3 independent experiments with the mean calculated from 9 values. Statistical significance was determined by 2-tailed Student's t test.

The α^{LEPRA} allele is a new target for therapeutic gene manipulation (26, 36–38). Therapies targeting splicing, particularly antisense oligonucleotides which recognize *cis* RNA regulatory elements and alter splicing or target transcripts for degradation and small molecule-based strategies that target splicing factors or RNA sequences and/or RNA structure, are showing significant promise (26, 36). The α^{LEPRA} allele is also an attractive candidate for therapeutic gene editing, as our data indicate alteration of the α^{LEPRA} allele variant itself or the alternate 3' acceptor site will abolish production of the mutant, elongated α -spectrin mRNA transcript.

The data demonstrate how the synthesis of complementary lines of investigation, clinical, laboratory, biochemical, and genetic data can be leveraged to define and advance our understanding, diagnosis, and treatment of inherited disease.

Methods

Erythrocyte membrane preparation, quantitation of spectrin content, and limited tryptic digestion of spectrin. Erythrocyte membranes were prepared from peripheral blood as previously described (39). Membrane proteins were analyzed and spectrin content determined as described (40). Membrane proteins were separated by SDS-PAGE on 3.5%–17% gradient polyacrylamide gels and stained with Coomassie blue. The spectrin/band 3 ratio was quantified by densitometric scanning of the stained gels at 540 nm and integration of the surface area under the spectrin and band 3 peaks (41). Spectrin was extracted by incubating ghosts overnight at 4°C in low ionic strength buffer. Limited tryptic digests of spectrin extracts were prepared as described (39) and sepa-

rated by 2-dimensional gel electrophoresis with isoelectric focusing (IEF) as modified by Speicher et al. (42).

Exon capture and WES. Targeted regions of genomic DNA were captured using a SeqCap EZExome V2.0 (Roche) solution-based capture system according to the manufacturer's protocol. The captured, purified, and amplified libraries targeting exomes from patients were sequenced on a HiSeq Analyzer with paired-end sequencing at 75-bp read length. Sequencing reads were processed and analyzed as described (8).

Fastq sequence reads were aligned to the human genome (hg19/build37/GRCh37) using BWA mem 0.7.9a software (9). Variant analysis was performed using Genome Analysis Toolkit (GATK) analysis software v.3.1-1 (10, 11). Sequencing reads in the region of known insertion/deletions in the 1000 Genomes database were realigned to reduce false-positive variants. Individual base quality scores were empirically recalibrated using covariate information in combination with the variation in dbSNP build 137 and from 629 human genome sequences from the 1000 Genomes project (12, 13). Base alignment quality scores were determined then recalibrated with GATK to further reduce false-positive calls (14). SNPs and insertions/deletions (indels) were called using the GATK HaplotypeCaller Bayesian algorithm for variant discovery and genotyping, which uses base qualities and allele counts to determine probabilities for called variants. Indels and novel SNPs were annotated using Annovar (43), which determines if a SNP or indel changes a protein sequence or splice site, and provides a variety of predictions, including the combined annotation dependent depletion (CADD) algorithm (44) and Mutation Taster (45), to assess the likely effect of a mutation on protein function.

Whole-genome sequencing. Genomic DNA was prepared using the TruSeq PCR-free DNA HT sample preparation kit (Illumina) with 450 bp insert size. Intact genomic DNA was sheared, followed by end-repair and bead-based size selection of fragmented molecules. Adenines were added to the 3' ends of the DNA size-selected fragments followed by ligation of Illumina sequence adaptors ligated onto the fragments and PCR. Library quality control included a measurement of the average size of library fragments using a BioAnalyzer (Agilent), estimation of the total concentration of DNA by PicoGreen (Thermo Fisher Scientific), and a measurement of the yield and efficiency of the adaptor ligation process via a quantitative PCR assay using primers specific to the adaptor sequence.

Sequencing was performed on a HiSeq X instrument (Illumina) generating 2×150 -bp read lengths. After alignment and duplicate removal, this equated to $\times 30$ mean genome coverage (for the gender-specific ~ 2.85 gigabase mappable human genome). WGS data were processed through an automated pipeline at New York Genome Center's high-performance computational facility. Paired-end 150-bp reads from the WGS were aligned to the GRCh37 human reference using the Burrows-Wheeler Aligner (BWA-MEM v0.78) (46) and processed using the best practices pipeline that includes marking of duplicate reads by the use of Picard tools (v1.83, <http://picard.sourceforge.net>), realignment around indels, and base recalibration via GATK v3.2.2 (47).

Two patients (nos. 1 and 23) were selected for WGS, both heterozygous for the α^{BH} variant. Analyzing WGS data, 122 variants in the *SPTA1* region (chr1:158545000-158690000) present in both samples were selected for further analysis using Annovar software to annotate allele frequencies and predict functional consequences. In addition, we identified samples in the 1000 Genomes database of healthy individuals who were homozygous for α^{BH} and downloaded genotype data for these individuals. Variants present in the 2 patient samples that were homozygous in the 1000 Genomes samples were excluded from further analysis, leaving 16 candidate variants, of which 4 were excluded because they were present at high frequency ($>1\%$) in the 1000 Genomes database. Eleven remaining variants were excluded because they were homopolymeric or short tandem repeat size variations that are common and often not accurately genotyped. The sole remaining variant, rs200830867, chr1:158613314 G to A, is the α^{LEPRA} allele (13). This variant was initially described in an rHS patient in *trans* to an *SPTA1* nonsense mutation associated with an elongated α -spectrin mRNA transcript containing 70 nt from the 3' end of intron 30 hypothesized to lead to frameshift and premature chain termination (13).

Mutation validation. Variants identified by WES or WGS were validated by Sanger sequencing. Sanger sequencing was performed on an 3130XL capillary sequencer (Applied Biosystems). *SPTA1* variants were classified as *cis* or *trans* by study of the proband's parents. Variants were classified as novel if not found in the 1000 Genomes or Exome Aggregation Consortium (ExAc) databases.

Molecular modeling. The effects of α -spectrin missense mutations on tetramer structure and their likely functional consequences were evaluated using the structures of the univalent α - β tetramer complex (48), closed spectrin dimers (49, 50), and the divalent tetramer complex as templates (50).

Minigene analyses. Minigene splicing assays were designed as described (18). Plasmids were constructed containing the human erythroid *ANK1* promoter, the human *HBG1* gene including the *HBG1* ATG/Kozak consensus, and the region of the *SPTA1* gene containing either WT

or α^{LEPRA} sequence for use in minigene splicing assays. Briefly, a SmaI-BglII fragment of the *ANK1* gene promoter was linked to a BglII-HindIII fragment containing exons 1–3 of the *HBG1* gene and flanking 3' region. A 1684-bp XbaI fragment of the *SPTA1* gene from intron 29 to intron 32, containing either the WT or α^{LEPRA} sequence, was cloned into intron 2 of the *HBG1* gene. These plasmids contained a termination codon and the polyadenylation signal from the *HBG1* gene.

Wild-type and α^{LEPRA} spectrin minigene plasmids containing fragments of the *SPTA1* gene from intron 29 to intron 32 were prepared and transfected into 10^7 K562 cells (ATCC, CCL-243) using nucleofection program T-016, buffer V (Amaxa). Cells were harvested after 48 hours, RNA was prepared then reverse transcribed using oligo d(T) and random hexamers, and cDNA was amplified with primers that included sequences from flanking *HBG1* and internal *SPTA1* exons. These primers allow differentiation of minigene amplification products from endogenous *SPTA1* transcripts.

Fluorescent Taqman probes (Applied Biosystems) corresponding to total *SPTA1* and the elongated α -spectrin transcript were added to amplification reactions, allowing determination of the contribution of elongated α -spectrin transcripts after normalization. To assay the total minigene *SPTA1* transcript, PCR was performed using primers 5'-CTTGGAGACTATGCCAACCTAAA-3' (sense) and 5'-CACATTTCCCAGGAGCTGAA-3' (antisense) with transcript quantitation by Taqman probe 5'-AATGGATCAGTGAGATGCTGCCCA-3' (sense). To assay elongated α -spectrin transcripts, PCR was performed using primers G4833, 5'-GAGAATTCCTGAGGT-CAGATG-3' (sense) and G44820, 5'-ATCCCAGACTCCCCTCTG-3' (antisense) with transcript quantitation by Taqman probe 5'-AGGCTTTGTATGAAGAAACGGGACGA-3' (sense). All analyses included 3 biologic and 2 technical replicates.

CRISPR/Cas9 gene editing analyses. Four guide RNAs (gRNAs) in the regions flanking the predicted BP2/ α^{LEPRA} region were designed, synthesized, and tested for use in gene editing experiments. gRNAs were synthesized using HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs). Testing was done using the Guide-it sgRNA Screening Kit (Takara). One gRNA was selected and used for all experiments (Target sequence: GGATTCAGAAGATATACTCA). Cas9 protein (15 μg , PNAbio) was complexed with 200 pmol gRNA. The gRNA-Cas9-RNP complex and 200 pmol donor ssDNA (BP2 A to G donor ssDNA: g(s)g(s)a(s)actctgtaccacacaagtagccattattagatgtttctctcctattaagtgtgaaaccacctcctgtgaagcatatattttgacctgagatatactctctgaatccGcaaaggatcctt(s)c(s)a(s) or α^{LEPRA} donor ssDNA: g(s)g(s)a(s)actctgtaccacacaagtagccattattagatgtttctcctcctattaagtgtgaaaccacctcctgtgaagcatatattttgacctgagatatactctctgaatTcacaaggatcctt(s)c(s)a(s)) were nucleofected (Amaxa) into K562 cells using the T-016 nucleofector program. Cells were placed into 96-well plates 48 hours after nucleofection. Clones were expanded for 10–14 days and then screened for either homozygous substitution of the BP2 A to G or homozygous alteration of the WT nucleotide α^{LEPRA} substitution C to T, by PCR amplification followed by Sanger nucleotide sequencing. Homozygous mutant BP2 and α^{LEPRA} clones were identified and used in additional studies (Supplemental Figure 4).

mRNA stability assays. The influence of NMD on the stability of WT α -spectrin and α^{LEPRA} mRNA was examined in WT K562 cells and in K562 cells rendered homozygous for the α^{LEPRA} allele. Cells were incubated at 37°C with the NMD inhibitor emetine (100 $\mu\text{g}/\text{ml}$, emetine dihydrochloride hydrate, MilliporeSigma, E2375) for 8 hours or

with the NMD inhibitor cycloheximide (100 µg/ml, MilliporeSigma, C1988) for 4 hours, or with a combination of both. The amount of total α -spectrin mRNA transcripts and α^{LEPRA} mRNA transcripts were determined by real-time RT-PCR with fluorescent probes as described above. Analyses included 3 biologic and 3 technical replicates.

RNA-seq data. RNA-seq data sets accessed from the Gene Expression Omnibus database were GSE61566 and GSE53983 for human erythroid cells and GSM958729 for K562 cells.

Statistics. GraphPad Prism 8 software (Graph Pad Software) was used for statistical analyses. All data were reported as mean \pm SD in tables or as error bars. Comparisons between 2 groups were performed using the 2-tailed Student's *t* test. Multiple test correction was performed using the Benjamini and Hochberg method (51). Significance was set at *P* less than 0.05.

Human subjects. All human studies were approved by the Yale University Human Investigation Committee review board. Written informed consent was received from participants or their parents, as appropriate, prior to inclusion in the study.

Author contributions

PGG designed experiments, analyzed data, and wrote the manuscript. KLG and YM designed and performed experiments. PEN, DM, RDH, JR, SI, DAW, RFS, CS, LKG, and NM phenotyped, diagnosed, and collected additional information on patients. RRS, DWS, SJB, and VPS designed experiments and analyzed data.

Acknowledgments

We thank all patients and their family members for participation. This work was supported in part by grants from the National Institutes of Health (HL65448, DK106857, DK032094, CA009171, GM115710, and GM122926) and the Arnold J. Alderman family. We dedicate this report to the late Bernard G. Forget.

Address correspondence to: Patrick G. Gallagher, Department of Pediatrics, Yale University School of Medicine, 333 Cedar Street, PO Box 208064, New Haven, Connecticut 06520-8064, USA. Phone: 203.688.2896; Email: patrick.gallagher@yale.edu.

- Agre P, Orringer EP, Bennett V. Deficient red-cell spectrin in severe, recessively inherited spherocytosis. *N Engl J Med.* 1982;306(19):1155–1161.
- Agre P, Casella JF, Zinkham WH, McMillan C, Bennett V. Partial deficiency of erythrocyte spectrin in hereditary spherocytosis. *Nature.* 1985;314(6009):380–383.
- Tse WT, et al. Amino-acid substitution in alpha-spectrin commonly coinherited with non-dominant hereditary spherocytosis. *Am J Hematol.* 1997;54(3):233–241.
- Mohandas N. Inherited hemolytic anemia: a possessive beginner's guide. *Hematology Am Soc Hematol Educ Program.* 2018;2018(1):377–381.
- Gallagher PG. Abnormalities of the erythrocyte membrane. *Pediatr Clin North Am.* 2013;60(6):1349–1362.
- Chasis JA, Agre P, Mohandas N. Decreased membrane mechanical stability and in vivo loss of surface area reflect spectrin deficiencies in hereditary spherocytosis. *J Clin Invest.* 1988;82(2):617–623.
- Zarkowsky HS, Mohandas N, Speaker CB, Shohet SB. A congenital haemolytic anaemia with thermal sensitivity of the erythrocyte membrane. *Br J Haematol.* 1975;29(4):537–543.
- Gallagher PG, et al. A common type of the spectrin alpha I 46-50a-kD peptide abnormality in hereditary elliptocytosis and pyropoikilocytosis is associated with a mutation distant from the proteolytic cleavage site. Evidence for the functional importance of the triple helical model of spectrin. *J Clin Invest.* 1992;89(3):892–898.
- Costa DB, Lozovatsky L, Gallagher PG, Forget BG. A novel splicing mutation of the alpha-spectrin gene in the original hereditary pyropoikilocytosis kindred. *Blood.* 2005;106(13):4367–4369.
- Hanspal M, et al. Asynchronous synthesis of membrane skeletal proteins during terminal maturation of murine erythroblasts. *Blood.* 1992;80(2):530–539.
- Hanspal M, Palek J. Biogenesis of normal and abnormal red blood cell membrane skeleton. *Semin Hematol.* 1992;29(4):305–319.
- Hanspal M, Palek J. Synthesis and assembly of membrane skeletal proteins in mammalian red cell precursors. *J Cell Biol.* 1987;105(3):1417–1424.
- Wichterle H, Hanspal M, Palek J, Jarolim P. Combination of two mutant alpha spectrin alleles underlies a severe spherocytic hemolytic anemia. *J Clin Invest.* 1996;98(10):2300–2307.
- Hanspal M, Hanspal JS, Sahr KE, Fibach E, Nachman J, Palek J. Molecular basis of spectrin deficiency in hereditary pyropoikilocytosis. *Blood.* 1993;82(5):1652–1660.
- Agre P, Asimos A, Casella JF, McMillan C. Inheritance pattern and clinical response to splenectomy as a reflection of erythrocyte spectrin deficiency in hereditary spherocytosis. *N Engl J Med.* 1986;315(25):1579–1583.
- Schwartz S, Hall E, Ast G. SROOGLE: webserver for integrative, user-friendly visualization of splicing signals. *Nucleic Acids Res.* 2009;37(Web Server issue):W189–W192.
- Signal B, Gloss BS, Dinger ME, Mercer TR. Machine learning annotation of human branchpoints. *Bioinformatics.* 2018;34(6):920–927.
- Cooper TA. Use of minigene systems to dissect alternative splicing elements. *Methods.* 2005;37(4):331–340.
- Edelman EJ, Maksimova Y, Duru F, Altay C, Gallagher PG. A complex splicing defect associated with homozygous ankyrin-deficient hereditary spherocytosis. *Blood.* 2007;109(12):5491–5493.
- Wong MS, Kinney JB, Krainer AR. Quantitative activity profile and context dependence of all human 5' splice sites. *Mol Cell.* 2018;71(6):1012–1026.e3.
- Pascolo E, Séraphin B. The branchpoint residue is recognized during commitment complex formation before being bulged out of the U2 snRNA-pre-mRNA duplex. *Mol Cell Biol.* 1997;17(7):3469–3476.
- Kurosaki T, Maquat LE. Nonsense-mediated mRNA decay in humans at a glance. *J Cell Sci.* 2016;129(3):461–467.
- Popp MW, Maquat LE. Leveraging rules of nonsense-mediated mRNA decay for genome engineering and personalized medicine. *Cell.* 2016;165(6):1319–1322.
- Conboy JG. RNA splicing during terminal erythropoiesis. *Curr Opin Hematol.* 2017;24(3):215–221.
- Pimentel H, et al. A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res.* 2014;42(6):4031–4042.
- Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet.* 2016;17(1):19–32.
- Sterne-Weiler T, Sanford JR. Exon identity crisis: disease-causing mutations that disrupt the splicing code. *Genome Biol.* 2014;15(1):201.
- Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease. *J Cell Biol.* 2016;212(1):13–27.
- Bishop DF, Schneider-Yin X, Clavero S, Yoo HW, Minder EI, Desnick RJ. Congenital erythropoietic porphyria: a novel uroporphyrinogen III synthase branchpoint mutation reveals underlying wild-type alternatively spliced transcripts. *Blood.* 2010;115(5):1062–1069.
- Lewandowska MA. The missing puzzle piece: splicing mutations. *Int J Clin Exp Pathol.* 2013;6(12):2675–2682.
- Vaz-Drago R, Custódio N, Carmo-Fonseca M. Deep intronic mutations and human disease. *Hum Genet.* 2017;136(9):1093–1111.
- Delaunay J, et al. Different impacts of alleles alphaLEPRA and alphaLELY as assessed versus a novel, virtually null allele of the SPTA1 gene in trans. *Br J Haematol.* 2004;127(1):118–122.
- Miraglia del Giudice E, et al. Clinical and molecular evaluation of non-dominant hereditary spherocytosis. *Br J Haematol.* 2001;112(1):42–47.
- Dhermy D, et al. Coinheritance of two alpha-spectrin gene defects in a recessive spherocytosis family. *Clin Lab Haematol.* 2000;22(6):329–336.
- Nussenzveig RH, Christensen RD, Prchal JT, Yaish HM, Agarwal AM. Novel α -spectrin mutation in trans with α -spectrin causing severe neonatal jaundice from hereditary spherocytosis. *Neonatology.* 2014;106(4):355–357.

36. Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med*. 2012;18(8):472–482.
37. Baralle D, Buratti E. RNA splicing in human disease and in the clinic. *Clin Sci*. 2017;131(5):355–368.
38. Havens MA, Duelli DM, Hastings ML. Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev RNA*. 2013;4(3):247–266.
39. Marchesi SL, Knowles WJ, Morrow JS, Bologna M, Marchesi VT. Abnormal spectrin in hereditary elliptocytosis. *Blood*. 1986;67(1):141–151.
40. Gallagher PG, Zhang Z, Morrow JS, Forget BG. Mutation of a highly conserved isoleucine disrupts hydrophobic interactions in the alpha beta spectrin self-association binding site. *Lab Invest*. 2004;84(2):229–234.
41. Fairbanks G, Steck TL, Wallach DF. Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry*. 1971;10(13):2606–2617.
42. Speicher DW, Morrow JS, Knowles WJ, Marchesi VT. Identification of proteolytically resistant domains of human erythrocyte spectrin. *Proc Natl Acad Sci U S A*. 1980;77(10):5673–5677.
43. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
44. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310–315.
45. Schwarz JM, Rödelberger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7(8):575–576.
46. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–1760.
47. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491–498.
48. Ipsaro JJ, Harper SL, Messick TE, Marmorstein R, Mondragón A, Speicher DW. Crystal structure and functional interpretation of the erythrocyte spectrin tetramerization domain complex. *Blood*. 2010;115(23):4843–4852.
49. Harper SL, Sriswasdi S, Tang HY, Gaetani M, Gallagher PG, Speicher DW. The common hereditary elliptocytosis-associated α -spectrin L260P mutation perturbs erythrocyte membranes by stabilizing spectrin in the closed dimer conformation. *Blood*. 2013;122(17):3045–3053.
50. Sriswasdi S, Harper SL, Tang HY, Gallagher PG, Speicher DW. Probing large conformational rearrangements in wild-type and mutant spectrin using structural mass spectrometry. *Proc Natl Acad Sci U S A*. 2014;111(5):1801–1806.
51. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stats Soc Series B Stat Methodol*. 1995;57(1):289–300.