



Exome sequencing identifies *GATA1* mutations resulting in Diamond-Blackfan anemia

Vijay G. Sankaran,^{1,2,3,4} Roxanne Ghazvinian,⁵ Ron Do,¹ Prathapan Thiru,² Jo-Anne Vergilio,⁶ Alan H. Beggs,^{4,5} Colin A. Sieff,^{4,7,8} Stuart H. Orkin,^{4,7,8} David G. Nathan,^{4,7,8} Eric S. Lander,^{1,9,10} and Hanna T. Gazda^{4,5}

¹Broad Institute, Cambridge, Massachusetts, USA. ²Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, USA.

³Department of Medicine, Children's Hospital Boston, Boston, Massachusetts, USA. ⁴Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA. ⁵Division of Genetics and Program in Genomics, The Manton Center for Orphan Disease Research, ⁶Department of Pathology, and ⁷Division of Hematology/Oncology, Children's Hospital Boston, Boston, Massachusetts, USA. ⁸Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. ⁹Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, USA.

¹⁰Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Diamond-Blackfan anemia (DBA) is a hypoplastic anemia characterized by impaired production of red blood cells, with approximately half of all cases attributed to ribosomal protein gene mutations. We performed exome sequencing on two siblings who had no known pathogenic mutations for DBA and identified a mutation in the gene encoding the hematopoietic transcription factor *GATA1*. This mutation, which occurred at a splice site of the *GATA1* gene, impaired production of the full-length form of the protein. We further identified an additional patient carrying a distinct mutation at the same splice site of the *GATA1* gene. These findings provide insight into the pathogenesis of DBA, showing that the reduction in erythropoiesis associated with the disease can arise from causes other than defects in ribosomal protein genes. These results also illustrate the multifactorial role of *GATA1* in human hematopoiesis.

Introduction

Advances in genomic sequencing technology promise to provide a more complete picture of the genetic basis of human disease and to make new connections between molecular pathways involved in disease pathophysiology. Diamond-Blackfan anemia (DBA; OMIM 105650) is attributable to reduced proliferation and survival of erythroid progenitors leading to hypoproliferative anemia (1, 2). The anemia in DBA is characterized by the production of enlarged (macrocytic) erythrocytes and, unlike other anemias, by therapeutic response to corticosteroids. Over the past decade, the elucidation of mutations in the ribosomal protein gene *RPS19* (3), followed by the discovery of mutations in 9 other ribosomal protein genes, has led to the hypothesis that DBA is a disorder of ribosomal biogenesis (1, 2). However, approximately 50% of DBA cases have as-yet-unidentified molecular mutations, despite systematic sequencing of all ribosomal protein and other candidate genes in these cases (1, 2, 4, 5). Recent work has highlighted the role of deletions of ribosomal protein genes that may be responsible for a subset of cases without previously identified mutations (6, 7). However, it is clear that the molecular etiology of many DBA cases remains to be uncovered. In light of this observation, we examined cases without previously identified pathogenic mutations in an attempt to identify new mutations that can cause DBA.

Results and Discussion

We focused on a family in which two male siblings were diagnosed with DBA (Figure 1A). While DBA generally demonstrates auto-

somal dominant inheritance (1), both parents in the family had entirely normal hematological analyses; assuming that the disease has full penetrance, this suggests X-linked or autosomal recessive inheritance. The affected siblings (II-1 and II-3) remain alive and most recently required chronic red blood cell transfusions for treatment (Table 1). Both siblings showed robust clinical responses to corticosteroid therapy for a period of 4 (II-1) and 6 (II-3) years, as is typical for DBA (Table 1). The siblings have macrocytic anemia, consistently low reticulocyte counts, and a modest elevation of fetal hemoglobin levels — features characteristic of DBA (Table 1 and Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI63597DS1). Sibling II-1 did not have an elevation of erythrocyte adenosine deaminase levels (Supplemental Table 1), which is seen in some DBA patients (1). Interestingly, sibling II-1 has also had a mildly low platelet count beginning at the age of 17 years, and both siblings were noted to have had occasional mild reductions in the neutrophil count, with those in sibling II-1 being consistently lower (Table 1). However, these siblings showed neither clinical signs of abnormal bleeding nor an increased propensity for infections. To confirm the clinical diagnosis of DBA, we reevaluated the original bone marrow aspirates and biopsies from these patients. Independent evaluation by hematopathologists agreed with the diagnosis of DBA. The bone marrow was noted to have erythroid hypoplasia without abnormalities of the other hematopoietic lineages (Figure 1, D–F).

We performed whole-exome sequencing on the 2 siblings with DBA, obtaining at least 10-fold coverage for more than 93% of the target bases (Supplemental Table 2 and refs. 8, 9). We reasoned that pathogenic mutations leading to DBA were likely to be rare in unaffected populations and therefore filtered out all variants identified from the latest draft of the 1000 Genomes Project, dbSNP

Authorship note: Eric S. Lander and Hanna T. Gazda contributed equally to this work.

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

Citation for this article: *J Clin Invest.* 2012;122(7):2439–2443. doi:10.1172/JCI63597.



Table 1
Hematologic parameters for index DBA cases

Age	wbc ($\times 10^9/l$)	Hgb (g/dl)	HCT (%)	Platelets ($\times 10^9/l$)	Abs. retic. ($\times 10^9/l$)	MCV (fl)	MCH (pg)	MCHC (g/dl)	RDW (%)	MPV (fl)	ANC ($\times 10^9/l$)
Individual II-1											
Birth ^A	27.4	9	27	572	5	106.2	35.4	33.4	23		3.01
2 yr, 11.8 mo	4.6	8	23.9	751	33	94.9	31.7	33.5			2.3
3 yr, 4.7 mo ^B	4.3	6.7	19.3	987	29	92.4	32.1	34.7			
3 yr, 5.6 mo	8.6	7.7	23.6	1,122	127	98.9	32.2	32.6			
3 yr, 6.3 mo	5.3	9.9	30.8	916	110	103.6	33.3	32.2			4.08
6 yr, 3.6 mo	7.5	7.1	22	545	30	94.8	30.6	32.2			
7 yr, 3.3 mo ^C	2.1	5.4	15.8	233	68						
14 yr, 0.8 mo ^D	2.5	6.9	19.9	116	5						0.46
16 yr, 11.3 mo ^D	2.5	7.9	22.2	79							0.51
19 yr, 3.5 mo ^D	2.8	7.5	20.9	102	24	79.1	28.4	35.8	16.3	10	0.67
21 yr, 4.2 mo ^D	1.7	8.6	24.3	67	24						0.45
22 yr, 11.9 mo ^D	3.1	8.4	23.7	77	27	79.5	28	35.3	14.3	8.8	0.56
23 yr, 11.9 mo ^D	3.1	8.4	23.3	71	21	79	28.5	36.1	15.2	8.1	0.83
Individual II-3											
1.6 mo	7.7	10.5	30.3	803	54						
3 yr, 10.8 mo	4.4	11.2	33.1	403	73	90.3	30.5	33.7	16.9	7	1.8
10 yr, 4 mo	3.6	8.7	25.9	291	37						1.6
11 yr, 5.6 mo ^B	2.6	5.8	17.4	300	15						0.67
11 yr, 5.8 mo	4	6.8	19.8	622	103	94.4					1.8
11 yr, 7.3 mo	2.6	9	26.3	409	147	100.6	34.5	34.3	21	6.5	1.8
15 yr, 2.2 mo	2.4	8.8	26.3	218	62	93.3	31	33.2	18.8	7.7	0.72
15 yr, 10.2 mo	2.3	9.9	29	220	78						1.2
17 yr, 5.2 mo	2.1	7.1	22	177	44						1.4
17 yr, 5.7 mo ^C	3.7	7.2	22.2	209	43						0.79
19 yr, 0.6 mo ^D	4.5	8.3	24.4	201	68	100	34.1	34.1	20.2	7.7	2.1

^AComplete blood count from birth. Cord blood hemoglobin (Hgb) was 7.5 g/dl. ^BStarted on prednisone. ^CFirst transfusion. ^DBlood count was determined before transfusion. Abs. retic., absolute reticulocyte count; ANC, absolute neutrophil count; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular Hgb; MCHC, mean corpuscular Hgb concentration; MPV, mean platelet volume; RDW, red cell distribution width.

build 132, and 95 exomes sequenced for the National Institute of Environmental Health Sciences Environmental Genome Project (Supplemental Table 2, Methods, and ref. 9). After filtering, a total of 74 variants were identified as being shared by the 2 affected siblings. We genotyped these mutations in the other family members (Figure 1A). Of the 74 mutations, 31 (42%) were found in the 2 affected siblings but not in an unaffected sibling (Supplemental Table 3). No variants were identified that would fit an autosomal recessive model of inheritance. Only a single variant on the X chromosome, within the *GATA1* gene, showed appropriate segregation for an X-linked disorder with full penetrance (Figure 1, A and B, and Supplemental Table 3).

GATA1 encodes a transcription factor necessary for erythroid differentiation (10), and therefore it is biologically plausible that this gene is involved in DBA. The mutation in the *GATA1* gene is a G→C transversion at position 48,649,736 on the X chromosome (hg19 coordinates) (Figure 1B) and results in the substitution of leucine for valine at amino acid 74 of the *GATA1* protein. The mutation occurs at the last nucleotide of the exon 2 donor splice site and therefore would also be predicted to affect splicing of *GATA1*.

RT-PCR on peripheral blood-derived RNA samples from control individuals confirmed prior findings (11) that there are normally 2 splice variants of *GATA1* produced: a full-length form involving splicing of exons 1, 2, and 3 with subsequent exons and a shorter *GATA1s* form involving splicing of exons 1 to 3 directly, with skipping of exon 2 (Figure 1C). By contrast, RT-PCR analysis of samples

from the patients showed that the *GATA1* mutation greatly favors the production of *GATA1s* mRNA, which lacks exon 2 (Figure 1C). Quantitative RT-PCR analysis of *GATA1* exon 2 demonstrated that individuals II-1 and II-3 had only trace amounts of mRNA containing this exon (3%–5% of control levels), while their mother (I-2), who carries the *GATA1* mutation, had a level at 53% of controls (Figure 2). This suggests that trace amounts of properly spliced full-length *GATA1* mRNA may possibly be produced with this mutation. A lack of exon 2 would only allow translation to initiate at codon 84, resulting in the formation of the *GATA1s* protein that lacks the first 83 amino acids, which contain the transactivation domain of this transcription factor (Figure 3 and refs. 12, 13).

We then sought to identify additional DBA patients carrying mutations in *GATA1* by screening 62 additional male DBA patients without known pathogenic mutations. We identified one patient with a deletion of one of 2 adjacent G nucleotides (X chromosome positions 48,649,736–48,649,737) at the same genomic position as the *GATA1* mutation found in the 2 brothers above (Figure 1G). This mutation would also be predicted to favor production of *GATA1s*, as a result of impaired splicing and frameshift of the full-length *GATA1* open reading frame (Figure 1G). This patient has anemia that has responded to treatment with corticosteroids and has not had other hematologic abnormalities (Supplemental Table 4).

Interestingly, a mutation identical to the G→C transversion in exon 2 has been reported to result in dyserythropoietic anemia in humans, and other *GATA1* germline mutations are associated

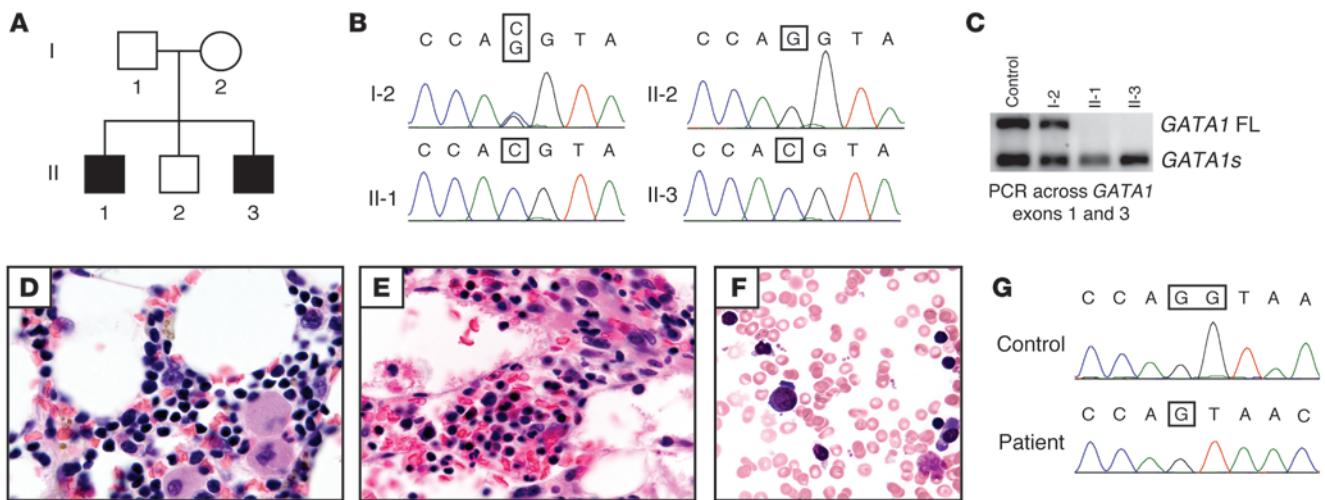


Figure 1

Identification of *GATA1* mutations in DBA. (A) Pedigree of a family with DBA in 2 male children and normal hematological findings in the parents. (B) Sanger sequencing chromatograms from the 3' end of exon 2 and 5' end of intron 2 on *GATA1* in the mother (I-2) and 3 children from the family shown in A. (C) RT-PCR analysis of *GATA1* exons 1 and 3 from peripheral blood RNA samples from the patients (II-1 and II-3); their mother (I-2), who carries the *GATA1* mutation; and a normal control. FL, full-length. (D and E) Bone marrow core sections from II-1 and II-3 shown at $\times 100$ magnification demonstrate erythroid hypoplasia. (F) Rare erythroblasts with normal morphology were noted on bone marrow aspirates (from II-1; shown at $\times 100$ magnification). (G) Sanger sequencing chromatograms show a *GATA1* G nucleotide deletion (at the exon 2–intron 2 junction) in an independent male patient with DBA.

with variable types of anemias and thrombocytopenias (11, 14, 15). These latter cases are due to missense mutations in the zinc fingers of *GATA1* and are distinct from the mutations affecting the production of different isoforms (Figure 3). The variability among phenotypes seen in the different mutations favoring *GATA1*s production may be attributable to differences in the levels of *GATA1* expressed. This phenomenon has been seen with mouse hypomorphic mutations, where even slight differences in *Gata1* levels can lead to variable phenotypes involving survival defects, unrestrained proliferation, or impaired differentiation of erythroid progenitors (16, 17). We speculate that alterations in *GATA1* expression may also underlie the phenotypic variability seen over time in the DBA patients. In addition, similar mutations that lead to the production of *GATA1*s alone are acquired somatically in all cases of Down syndrome–associated acute megakaryoblastic leukemia and transient myeloproliferative disease (12, 13). Mice with a mutation resulting in expression of only *GATA1*s have apparently normal erythropoiesis (18), which emphasizes the species-divergent functions of *GATA1*. Specifically, these findings demonstrate that the full-length form of *GATA1* is required for normal erythropoiesis in humans, but not in mice.

Systematic sequencing of *GATA1* mutations in other cases of DBA will likely unveil similar mutations and reveal the extent to which such mutations contribute to this disease. All 3 of the patients with *GATA1* mutations meet all of the current clinical diagnostic criteria for DBA and many of the supportive criteria, with the exception of elevations in erythrocyte adenosine deaminase levels (19). However, discovery and further phenotypic analysis of DBA patients with *GATA1* mutations may uncover unique differences between this set of patients and cases due to ribosome protein gene mutations, which may lead to revision of the current diagnostic criteria for DBA (19). While the majority of studies on DBA pathogenesis have been focused on the role of ribosomal biogenesis, the find-

ing of *GATA1* mutations in DBA opens new avenues for studying the underlying basis of this disorder. These findings may provide insight into the erythroid specificity of this disease, which remains an enigma. Additionally, these DBA cases, coupled with the phenotypes described for other human *GATA1* mutations (11, 14, 15), increase our understanding of how this transcription factor plays a role in specifying human erythropoiesis.

Methods

Exome sequencing and analysis. The DBA patient cohort has been described previously (4, 5). All patients were selected based on the diagnosis of DBA,

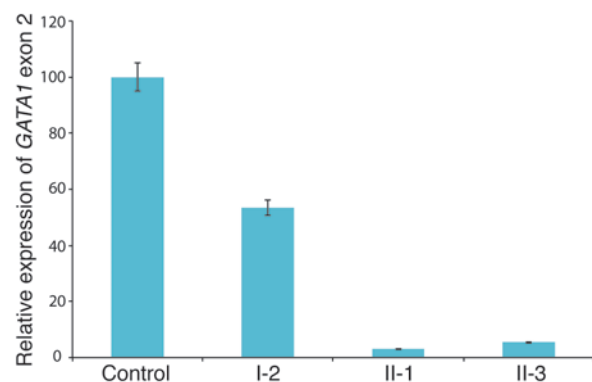


Figure 2

GATA1 exon 2 mutation results in trace amounts of *GATA1* mRNA containing exon 2. Quantitative RT-PCR was performed on peripheral blood RNA samples from individuals I-2, II-1, and II-3, as well as a normal control individual. Data are shown as the mean \pm SD. Trace amounts of exon 2 are present in the male carriers of the *GATA1* exon 2 mutation (3% and 5%, respectively), and a female carrier (I-2) with this mutation has half the level of mRNA compared with controls (53%).

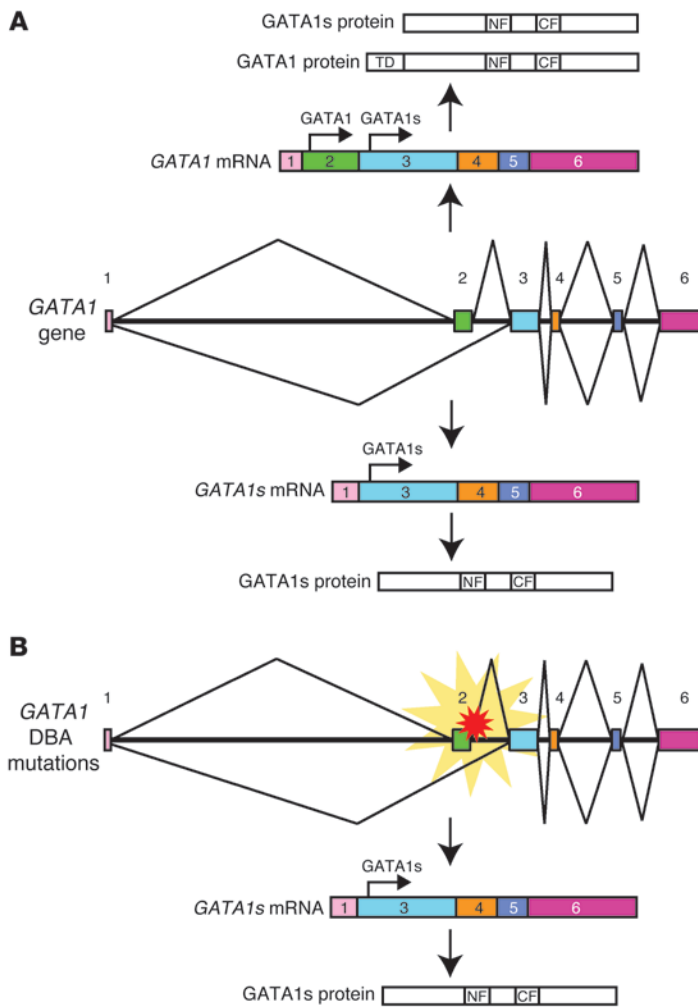


Figure 3

A model of how *GATA1* mutations in DBA favor production of *GATA1s* alone. **(A)** Normal splicing pattern of *GATA1* in humans and the resulting mRNA and protein products. **(B)** Aberrant splicing of exons 2 and 3 (highlighted by red star at mutation sites) leads to production of *GATA1s* alone. For the protein products, the transactivation domain (TD), N-terminal zinc finger (NF), and C-terminal zinc finger (CF) are shown.

(21) (<http://www.1000genomes.org/>), dbSNP build 132 (<http://www.ncbi.nlm.nih.gov/projects/SNP/>), and 95 exomes available from the National Institute of Environmental Health Sciences Environmental Genome Project (<http://evs.gs.washington.edu/niehsExome/>). We filtered variants in the patient samples using all of these datasets with the GATK Variant Filtration and Select Variants packages (20). Metrics from this process are shown in Supplemental Table 2. We developed custom Perl scripts to assess the number of variants in various functional classes following variant filtration. We also developed a custom Perl script to assess common variants that were identified in both siblings affected with DBA.

Sanger sequencing-based genotyping of variants. A custom Perl script was written to design primer pairs for all variants identified from the analysis described above to genotype all affected and unaffected family members (5 total individuals). We used the script to get the genomic sequence surrounding the variant and then used the Primer3 program (<http://frodo.wi.mit.edu/primer3/>) to design primers to amplify a fragment of 200–300 bp centered around each variant. The same set of primers was used for both PCR amplification from genomic DNA samples and for Sanger sequencing. Results from the sequencing-based genotyping were curated using automated and manual approaches. The results of the genotyping for all 5 family members are listed in Supplemental Table 3. Sixty-two additional male patients with DBA were screened for *GATA1* mutations. All 6 exons of *GATA1* were screened for mutations using standard PCR-based Sanger sequencing.

RT-PCR analysis. Total RNA was isolated from peripheral blood using a Paxgene Blood RNA kit (QIAGEN/BD) and treated with DNase (QIAGEN). RT-PCR was performed using a One-Step RT-PCR kit (QIAGEN). RT-PCR primers were designed in exon 1 (forward, 5'-ACACTGAGCTT-GCCACATC-3') and in exon 3 (reverse, 5'-CACAGTTGAGGCAGGG-TAGAG-3') of *GATA1*. RT-PCR was performed for 30 cycles with an annealing temperature of 57°C, and the products were analyzed by agarose gel electrophoresis.

Quantitative RT-PCR. To detect expression of *GATA1* exon 2, RNA was isolated from peripheral blood as above. cDNA was prepared using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). Real-time PCR for *GATA1* exon 2 was performed using the SYBR Green PCR Master Mix (Applied Biosystems) and Applied Biosystems Real-Time PCR 7300 System. *GAPDH* was used as the control to normalize *GATA1* exon 2 expression levels. *GATA1* expression quantification was performed using the $\Delta\Delta C_t$ method (22). The following primers were used for quantitative RT-PCR: *GATA1* exon 2, forward 5'-CCCCAGTTTGTGGATCCTG-3', reverse 5'-ACCCCTGATTCTGGTGTGG-3'; *GAPDH*, 5'-TGCACCAC-CACTGCTTAGC-3', reverse 5'-GGCATGGACTGTGGTCATGAG-3'.

Statistics. All pairwise comparisons were assessed using unpaired 2-tailed Student *t* test. Results were considered significant if the *P* value was less than 0.05.

Study approval. All patients or their families had provided written informed consent to participate in this study. The institutional review

using criteria outlined previously (1). Exome sequencing was performed at the Broad Institute (8, 9). Briefly, oligonucleotides with 170 bp of target sequence flanked by 15 bp of universal primer sequence were synthesized in parallel on an Agilent microarray and subsequently cleaved. The oligonucleotides were PCR amplified and transcribed in vitro with biotinylated UTP to generate single-stranded RNA “bait.” The genomic DNA was sheared, ligated to sequencing adapters, and selected for lengths between 200 and 350 bp. This “pond” of DNA was hybridized with an excess of bait in solution. The captured material was pulled down by magnetic beads coated with streptavidin and subsequently eluted. Afterward, each sample was pooled, and sufficient sequencing was performed so that each sample had an average coverage equivalent to one lane of sequencing with 76-bp paired-end reads using an Illumina HiSeq 2000 sequencer.

All sequencing data were processed through an automated analysis pipeline built around the Picard suite (<http://picard.sourceforge.net/>), and BAM files were exported from this process (using the hg19 human genome draft). These BAM files were then run through a variant caller pipeline utilizing the Genome Analysis Toolkit (GATK) (20). This pipeline implements best-practice approaches for SNP and indel calling, and variants for each individual sample were compiled into a single variant call format (VCF) file.

Once all variants were obtained, annotated, and assessed from the exome sequencing process, variant filtration was performed. We gathered data for all variants from the June 2011 data release of the 1000 Genomes Project



boards at the Massachusetts Institute of Technology/Broad Institute and the Children's Hospital Boston approved the study protocols.

Acknowledgments

We are grateful to the families involved in this work for their courage and commitment to research and to the clinicians caring for these patients. We thank M. Ilzarbe, N. Gupta, J. Murphy, and S. Flynn for their assistance in study organization; M. DePristo, C. Hartl, and K. Shakir for assistance with genetic analysis; H. Lodish, L. Ludwig, O. Zuk, M. Garber, and M. Schnall-Levin for valuable comments and assistance; and T. DiCesare for assistance with illustrations. This work was supported by NIH grant T32 HL007574-30 (to V.G. Sankaran); NIH grant R01 HL107558 and

grants from the Manton Center for Orphan Disease Research and the Diamond-Blackfan Anemia Foundation (to H.T. Gazda); and NIH grant U54 HG003067-09 (to E.S. Lander).

Received for publication February 28, 2012, and accepted in revised form May 2, 2012.

Address correspondence to: Hanna T. Gazda, Children's Hospital Boston, 300 Longwood Ave., Boston, Massachusetts 02115, USA. Phone: 617.919.4587; Fax: 617.730.0253; E-mail: hanna.gazda@childrens.harvard.edu. Or to: Eric S. Lander, Broad Institute, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. Phone: 617.714.7011; Fax: 617.714.8992; E-mail: lander@broadinstitute.org.

1. Boria I, et al. The ribosomal basis of Diamond-Blackfan Anemia: mutation and database update. *Hum Mutat.* 2010;31(12):1269-1279.
2. Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood.* 2010; 115(16):3196-3205.
3. Draptchinskaia N, et al. The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat Genet.* 1999;21(2):169-175.
4. Doherty L, et al. Ribosomal protein genes RPS10 and RPS26 are commonly mutated in Diamond-Blackfan anemia. *Am J Hum Genet.* 2010;86(2):222-228.
5. Gazda HT, et al. Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am J Hum Genet.* 2008;83(6):769-780.
6. Farrar JE, et al. Ribosomal protein gene deletions in Diamond-Blackfan anemia. *Blood.* 2011; 118(26):6943-6951.
7. Kuramitsu M, et al. Extensive gene deletions in Japanese patients with Diamond-Blackfan anemia. *Blood.* 2012;119(10):2376-2384.
8. Gnirke A, et al. Solution hybrid selection with ultralong oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol.* 2009;27(2):182-189.
9. Musunuru K, et al. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N Engl J Med.* 2010;363(23):2220-2227.
10. Tsai SF, Martin DI, Zon LI, D'Andrea AD, Wong GG, Orkin SH. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature.* 1989; 339(6224):446-451.
11. Hollanda LM, et al. An inherited mutation leading to production of only the short isoform of GATA-1 is associated with impaired erythropoiesis. *Nat Genet.* 2006;38(7):807-812.
12. Alford KA, et al. Analysis of GATA1 mutations in Down syndrome transient myeloproliferative disorder and myeloid leukemia. *Blood.* 2011; 118(8):2222-2238.
13. Wechsler J, et al. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 2002;32(1):148-152.
14. Nichols KE, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet.* 2000;24(3):266-270.
15. Kacena MA, Chou ST, Weiss MJ, Raskind WH. GATA1-related X-linked cytopenia. In: Pagon RA, Bird TD, Dolan CR, Stephens K, Adam MP, eds. *GeneReviews* [Internet]. Seattle, Washington, USA; 1993.
16. McDevitt MA, Shivdasani RA, Fujiwara Y, Yang H, Orkin SH. A "knockdown" mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proc Natl Acad Sci U S A.* 1997;94(13):6781-6785.
17. Pan X, et al. Graded levels of GATA-1 expression modulate survival, proliferation, and differentiation of erythroid progenitors. *J Biol Chem.* 2005;280(23):22385-22394.
18. Li Z, Godinho FJ, Klusmann JH, Garriga-Canut M, Yu C, Orkin SH. Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1. *Nat Genet.* 2005; 37(6):613-619.
19. Vlachos A, et al. Diagnosing and treating Diamond Blackfan anaemia: results of an international clinical consensus conference. *Br J Haematol.* 2008; 142(6):859-876.
20. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498.
21. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467(7319):1061-1073.
22. Sankaran VG, Orkin SH, Walkley CR. Rb intrinsically promotes erythropoiesis by coupling cell cycle exit with mitochondrial biogenesis. *Genes Dev.* 2008; 22(4):463-475.