

Letter to the Editor

RNA sequencing reveals the consequences of a novel insertion in dedicator of cytokinesis-8

To the Editor:

Primary immunodeficiency diseases (PIDs) are a diverse group of genetic disorders of immune function in humans with varied clinical heterogeneity. Recent clinical approaches to identify novel mutations in genes causing PIDs have included both whole-genome sequencing and whole-exome sequencing.¹ Yet, these technologies cannot uncover all potential causes of PIDs, especially given that 70% of the human genome is transcribed into noncoding RNAs. The most recent advance in the characterization of the transcriptome is RNA sequencing. This technique has the advantage of providing quantitative measures of both coding and noncoding RNA species. We assessed the value of this approach in identifying and characterizing the genetic causes leading to PIDs.

Two siblings, born to consanguineous parents (Fig 1, A), were diagnosed with a PID on the basis of clinical and laboratory findings (Fig 1, A; see Tables E1 and E2 in this article's Online Repository at www.jacionline.org). The parents and another sibling were healthy (Fig 1, A). Both affected children had a T-cell lymphopenia, with reduced numbers of CD4⁺ and CD8⁺ T cells. The proliferative response of T cells to mitogens and antigens was significantly reduced compared with controls, and there was evidence of an increased susceptibility to activation-induced cell death in response to T-cell receptor stimulation. This clinical presentation prompted us to determine whether the cells from these children had an increased susceptibility to DNA damage. A fibroblast cell line established from patient 1 and a B-cell lymphoblastoid cell line from patient 2 exhibited intermediate cell survival in cell-colony survival assays after radiation² (see Fig E1 in this article's Online Repository at www.jacionline.org). However, known genes defective in classic radiosensitive PID were normal. To identify candidate genes whose mutations could lead to immunodeficiency, exome sequencing and RNA sequencing were performed on samples from the affected siblings, 1 unaffected sister, and their father (Fig 1, B). A series of filtering strategies were introduced with the exome-sequencing data (see this article's Methods section in the Online Repository at www.jacionline.org). A candidate list of 24 variants was obtained. Most notable was the presence of a previously unreported, homozygous 2 base-pair cytosine-adenosine insertion, in the 3' splice acceptor site region of the dedicator of cytokinesis-8 (*DOCK8*) gene, just upstream of exon 27 (Fig 1, C and D). The father and the unaffected sibling were heterozygous for this mutation. *DOCK8* is a guanine nucleotide exchange factor belonging to the DOCK180 family of proteins. This family has a phosphoinositide-binding domain, termed the dock homology region 1 and a GTPase interaction site (dock homology region 2). These domains enable *DOCK8* to regulate actin cytoskeleton rearrangements by functioning as a guanine nucleotide exchange factor for the Rho-GTPases Rac and CDC42. *DOCK8* also couples Toll-like receptor 9 to MyD88.³ Loss-of-function mutations in *DOCK8* causes a combined immunodeficiency, with patients having varied clinical presentations including recurrent viral and bacterial infections, elevated IgE

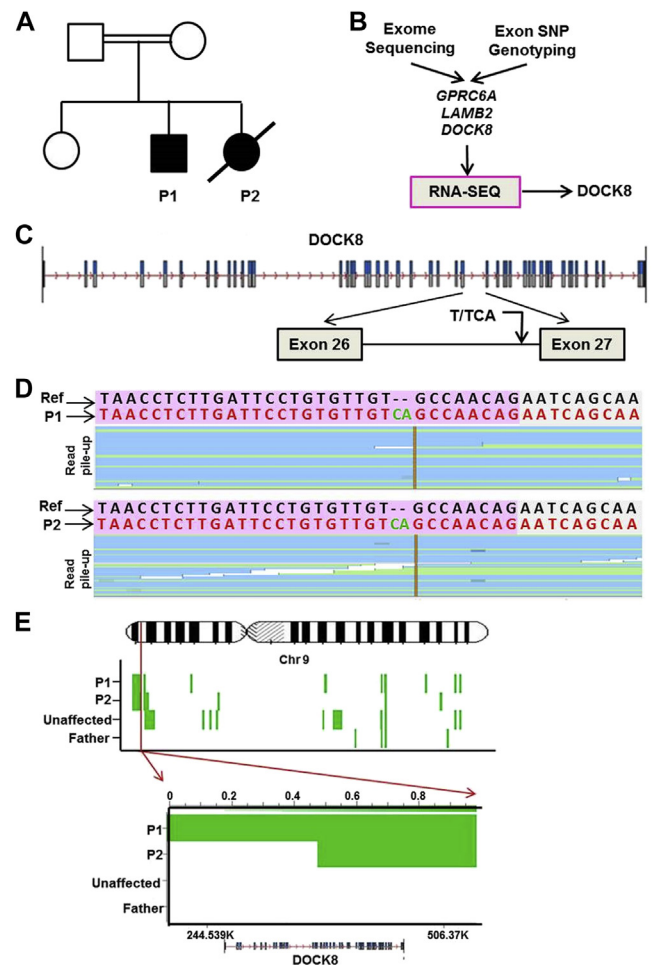


FIG 1. Two siblings with PIDs have mutations in *DOCK8*. **A**, Family pedigree with 2 affected siblings (black filled) and an unaffected sister. **B**, Sequencing and filtering strategy used to identify candidate genes. **C**, Schematic representation of intronic-exonic organization of *DOCK8* with a 2 nucleotide (CA) insertion detected at the splice junction preceding exon 27. **D**, The sequence read number is displayed as a read pile-up for P1 and P2. CA insertion noted in green. **E**, The regions of homozygosity on chromosome 9 are shown as green segments. The red line expands the area encompassing an extensive region of homozygosity unique to the affected siblings. CA, Cytosine-adenosine; P1, patient 1; P2, patient 2; SEQ, sequencing; SNP, single nucleotide polymorphism.

levels, and eczema.^{4,5} Although *DOCK8* was the most likely candidate in our screen, there were 23 other candidate variants. To determine whether these were coupled to the clinical phenotypes, homozygosity mapping was performed. Fifty-two runs of homozygosity were identified. Among all novel variants, *DOCK8*, *GPRC6A*, and *LAMB2* were selectively homozygous in the 2 affected children but not in the unaffected sibling or their father (Fig 1, E). To determine whether these genes were expressed in the relevant cell types, we performed RNA sequencing on whole blood of the 2 affected siblings, the unaffected sister, and their father. Reads per kilobase per million values from the RNA-sequencing data were used to define the levels of *DOCK8*, *LAMB2*, and *GPRC6A*. Only transcripts for *DOCK8*

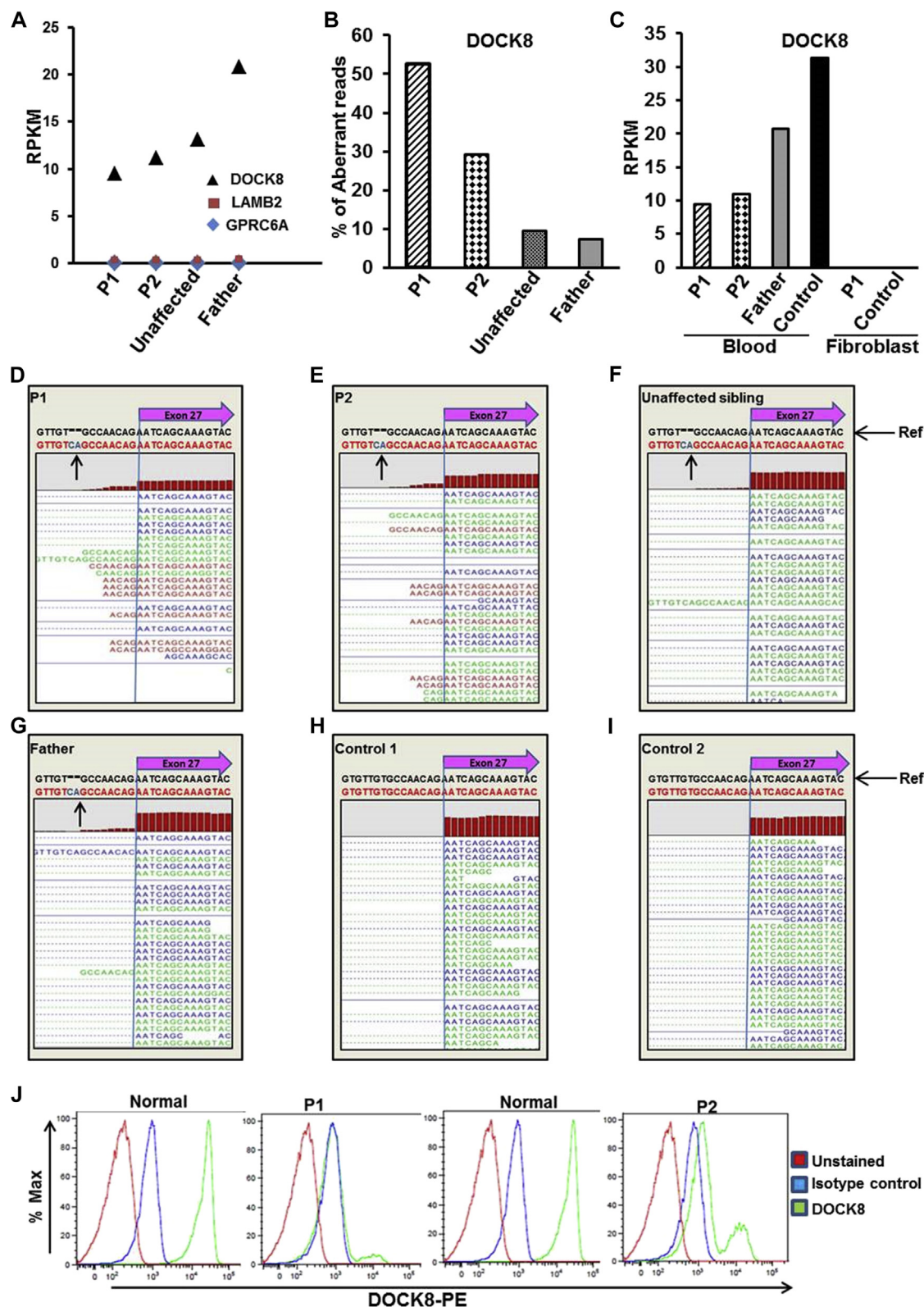


FIG 2. RNA sequencing uncovers abnormal *DOCK8* transcripts. **A**, RNA sequencing was used to assess the expression levels of *DOCK8*, *LAMB2*, and *GPRC38* transcripts in the peripheral blood of the patients and family members. **B**, The percentage of aberrant reads in the region spanning the exon-intron boundaries was determined for the 2 patients, the unaffected sibling, and their father from the RNA-sequencing data. **C**, The RPKM values for *DOCK8* transcripts were compared in peripheral blood and skin fibroblasts from the indicated patients and control. The presence and type of aberrant sequence read preceding exon 27 is shown for P1 (**D**), P2 (**E**), the unaffected sibling (**F**), and their father (**G**) and compared with that for 2 independently isolated normal controls (**H** and **I**). **J**, Flow cytometry analysis of lymphocytes isolated from the affected siblings (P1 and P2) is compared with that for normal controls. Cells were analyzed for forward and side scatter properties, and for the appearance of DOCK8 protein, as detected by intracellular staining. P1, Patient 1; P2, patient 2; RPKM, reads per kilobase per million.

were identified, eliminating *LAMB2* and *GPRC6A* as candidates (Fig 2, A). Half of the mapped sequence reads spanned the intron-exon junction in the affected patients, whereas the unaffected sibling and the father contained 10% of aberrant reads (Fig 2, B). Both affected siblings had a 2-fold reduction in total *DOCK8* mRNA levels compared with the heterozygous father and 3-fold less than control (Fig 2, C). The RNA-sequencing data further revealed the consequence of these 2 base-pair cytosine-adenosine insertion. The aberrant sequence reads generated nonfunctional *DOCK8* transcripts of varying lengths, with disrupted codons and premature stop sites (Fig 2, D-G). Two unrelated normal controls showed no aberrant transcripts (Fig 2, H and I). The severely disrupted *DOCK8* transcripts suggested that the full-length protein would be missing. To examine this possibility, peripheral blood cells were subjected to intracellular staining with antibodies specific to *DOCK8*. Most of the T- and B-cell populations lacked *DOCK8* expression (Fig 2, J). Interestingly, a small percentage of CD4, CD8, and CD56 natural killer cell subsets retained some *DOCK8* expression in patient 2 (see Fig E2 in this article's Online Repository at www.jacionline.org). These could have resulted from either alternative splicing or somatic reversions, which has been reported in certain lymphocyte populations, as shown previously.⁶

There is immense clinical and immunologic heterogeneity in patients with *DOCK8* deficiency.⁷⁻¹⁰ In fact, cells from the 2 siblings described in the current report had intermediate survival after radiation, a phenotype not previously reported. Neither sibling had a separate mutation in known genes involved in double-stranded DNA repair. Of note, no *DOCK8* transcripts were present in skin fibroblasts from either the patient or a control (Fig 2, C). This suggests that the intermediate cell survival exhibited by the patient fibroblasts after radiation is not a consequence of mutations in *DOCK8*. Instead, the intermediate colony survival results are likely independent of DNA repair.

In conclusion, this report reveals the clinical applicability of RNA sequencing as a critical tool in the diagnosis of PID presentations, complementing the more widely used exome sequencing. Within the last several years, RNA-sequencing costs have come down 5- to 10-fold. Laboratories holding Clinical Laboratory Improvement Amendments of 1988 (CLIA)-certification are currently offering this service at a cost of \$500 to \$600 per sample. Multiplexing capabilities and technology advances will reduce these costs even further in the near future. This will enable clinicians to assess the impact of incorrectly spliced, aberrant RNA species that are generated from any gene mutation in the clinical phenotypes of patients.

We thank the patients and their family members for their participation in the study.

Shaheen Khan, PhD^a
Merin Kuruvilla, MD^{b,c}
David Hagin, MD, PhD^f

Benjamin Wakeland, BSc^a
Chaoying Liang, BSc^a
Kasthuribhai Vishwanathan, MS^a
Richard A. Gatti, MD^g
Troy R. Torgersen, MD, PhD^f
Roshini S. Abraham, PhD^h
Edward K. Wakeland, PhD^a
Nicolai S. C. van Oers, PhD^{a,c,d}
M. Teresa de la Morena, MD^{b,c,e,i}

From ^athe Departments of Immunology, ^bInternal Medicine, ^cPediatrics, and ^dMicrobiology, ^eDivision of Allergy and Immunology, University of Texas Southwestern Medical Center, Dallas, Tex; ^fthe Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, Wash; ^gthe Departments of Pathology & Laboratory Medicine and Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Calif; ^hthe Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minn; and ⁱChildren's Medical Center, Children's Heath, Dallas, Tex. E-mail: maite.delamorena@utsouthwestern.edu.

This study was supported by funds from the Jeffrey Modell Foundation (M.T. de la Morena) and Children's Medical Center Research Foundation in Dallas (N.S.C. van Oers). Disclosure of potential conflict of interest: T. Torgersen receives consultant fees from Baxalta Biosciences, CLS Behring, and ADMA Biosciences and grant funding from Baxalta Biosciences, CSL Behring, and the National Institutes of Health; has also received payment for the development of educational presentations from Baxalta Biosciences, CSL Behring, Questcor Pharmaceuticals, and RWJF. M. T. de la Morena receives grant support from the Jeffrey Modell Foundation. The rest of the authors declare that they have no relevant conflicts of interest.

REFERENCES

1. Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD. Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. *J Exp Med* 2014;211:2137-49.
2. Sun X, Becker-Catania SG, Chun HH, Hwang MJ, Huo Y, Wang Z, et al. Early diagnosis of ataxia-telangiectasia using radiosensitivity testing. *J Pediatr* 2002; 140:724-31.
3. Jabara HH, McDonald DR, Janssen E, Massaad MJ, Ramesh N, Borzutzky A, et al. *DOCK8* functions as an adaptor that links TLR-MyD88 signaling to B cell activation. *Nat Immunol* 2012;13:612-20.
4. Dinwiddie DL, Kingsmore SF, Caracciolo S, Rossi G, Moratto D, Mazza C, et al. Combined *DOCK8* and *CLEC7A* mutations causing immunodeficiency in 3 brothers with diarrhea, eczema, and infections. *J Allergy Clin Immunol* 2013; 131:594-7, e591-3.
5. Engelhardt KR, McGhee S, Winkler S, Sassi A, Woellner C, Lopez-Herrera G, et al. Large deletions and point mutations involving the dedicator of cytokinesis 8 (*DOCK8*) in the autosomal-recessive form of hyper-IgE syndrome. *J Allergy Clin Immunol* 2009;124:1289-302, e1284.
6. Jing H, Zhang Q, Zhang Y, Hill BJ, Dove CG, Gelfand EW, et al. Somatic reversion in dedicator of cytokinesis 8 immunodeficiency modulates disease phenotype. *J Allergy Clin Immunol* 2014;133:1667-75.
7. Al-Herz W, Ragupathy R, Massaad MJ, Al-Attayah R, Nanda A, Engelhardt KR, et al. Clinical, immunologic and genetic profiles of *DOCK8*-deficient patients in Kuwait. *Clin Immunol* 2012;143:266-72.
8. Sanal O, Jing H, Ozgur T, Ayvaz D, Strauss-Albee DM, Ersoy-Evans S, et al. Additional diverse findings expand the clinical presentation of *DOCK8* deficiency. *J Clin Immunol* 2012;32:698-708.
9. Alsum Z, Hawwari A, Alsmadi O, Al-Hissi S, Borrero E, Abu-Staiteh A, et al. Clinical, immunological and molecular characterization of *DOCK8* and *DOCK8*-like deficient patients: single center experience of twenty-five patients. *J Clin Immunol* 2013;33:55-67.
10. Engelhardt KR, Gertz ME, Keles S, Schaffer AA, Sigmund EC, Glocker C, et al. The extended clinical phenotype of 64 patients with dedicator of cytokinesis 8 deficiency. *J Allergy Clin Immunol* 2015;136:402-12.

<http://dx.doi.org/10.1016/j.jaci.2015.11.033>

METHODS

Patients

The patients were referred to the Immunology Clinics at Children's Medical Center, Dallas, Texas (Children's Health). Informed consent was obtained from the family for this study in accordance with institutional review board (IRB) approval at The University of Texas Southwestern Medical Center, Dallas (IRB no. STU112010-013). The sampling of RNA, DNA, and peripheral blood from the patients and family members was done under IRB-approved protocols. Buccal swabs were used for DNA preparations, whereas peripheral blood samples were used for flow cytometry and RNA isolation. RNA was isolated following blood sampling with PAXgen isolation tubes and RNA isolation kits (Qiagen Corp, Valencia, Calif).

Detailed clinical presentation

Patient 1 was a female who moved to the United States at age 7 years. Reported clinical history included frequent upper respiratory tract infections. At age 5 years, this child was diagnosed with failure to thrive and severe malnutrition. At age 6 years, documented pneumonia with associated pleural effusions resulted in a hospitalization in El Salvador. This child presented clinically in the United States at age 8 years with persistent cough and weight loss of 5 to 6 lb (below the 5th percentile for height and weight), scleral icterus, and hepatosplenomegaly. In addition, she had planar and classic verruca over trunk and extremities. She was subsequently diagnosed with *Candida* esophagitis, portal hypertension, and sclerosing cholangitis with portal fibrosis. The clinical course progressed to liver failure requiring liver transplantation. The liver explant revealed *Cryptosporidia*, despite repeated stool tests negative for cryptosporidiosis and *Giardia* antigen for diarrhea. She received a matched unrelated bone marrow transplant 1 year after liver transplant and subsequently died from disseminated histoplasmosis complicating graft versus host disease. The older male sibling (patient 2) also presented at age 8 years with a history of multiple respiratory tract infections (otitis media and pneumonias), thrush, chronic diarrhea, failure to thrive, atopic dermatitis, and disseminated cutaneous viral infections including planar and classic verruca along with molluscum contagiosum (Table E1). Chronic cryptosporidiosis and *Giardia* infection were identified in stool samples. He also had chronic transaminitis and sclerosing cholangitis with mild portal fibrosis. He is currently being treated with antibiotic prophylaxis and gammaglobulin therapy. Both siblings received routine vaccinations including measles, mumps, and rubella and varicella during immigration without complications. Neither patient displayed significant rhinitis, asthma, or food allergy. There was no history of skin abscesses or neurological or autoimmune disease. There was no evidence of developmental delay in either patient (Table I). CD8⁺ T-cell and natural killer cell cytotoxicity appeared intact, by *in vitro* assays, indicating normal cytotoxic granule exocytosis and perforin/granzyme B content (Table E2). Both patients had elevated levels of IgG and IgA (Table E2). Despite the increased levels of IgG, specific antibody responses were defective to pneumococcal polysaccharide immunizations. Circulating B cells were present in normal numbers, though most were of the naive phenotype with significantly decreased switched memory B cells (CD27⁺ B cells < 2%). Both children had IgE levels below 100 IU/mL (Table E2).

Colony survival assay and immunoblotting

Colony survival assays and western blotting were used to evaluate the repair of double-strand DNA breaks as a workup for the 2 siblings manifesting lymphopenia and immune deficiency. A fibroblast cell line was established from patient 1 because of failure of expansion of a lymphoblastoid line. A lymphoblastoid cell line was established from patient 2. The cells were plated at various concentrations and subject to 1, 2, and 5 Gy of radiation; control plates were not irradiated. After 2 weeks of culture, the number of surviving colonies in the plates was scored. Positive and negative controls were included in each batch, including individuals with known DNA repair disorders. To detect specific proteins associated with similar syndromes, immunoblots were loaded with 50 µg of radioimmunoprecipitation assay buffer whole-cell lysates from lymphoblastoid cell lines or fibroblasts. The electrophoresed proteins were transferred from gels to nylon membranes so

that these could be visualized with commercial antibodies as indicated (Abcam, Cambridge, Mass). Specific protein levels were semi-quantitated by densitometry using a GelDoc (BioRad, Hercules, Calif). Antibodies to aprataxin, nibrin, or Mre11 were used as loading controls.

Exome sequencing

Genomic DNA was isolated from buccal swabs with a QIAamp DNA kit (Qiagen). The DNA (3 µg) was sheared with a S220 Ultrasonicator (Covaris, Woburn, Mass) and subsequently ligated to the Illumina sequencing paired-end adaptors. These were amplified by PCR and purified. The exon-containing fragments from the PCR libraries were hybridized to the SureSelect Human All Exon Kit from Agilent Technologies (Santa Clara, Calif) and/or the SeqCap EZ Human Exome Library v3.0 from NimbleGen (Madison, Wis). After hybrid capture, amplification, and quality assessment, the samples were sequenced on Illumina GAIIx sequencer using a standard 75-base-pair (bp) paired-end protocol or Hiseq 2000 using a 100-bp paired-end protocol. The image analysis and base calling were performed with an Illumina pipeline, using default settings.

Sequence alignment, variant calling, and annotation

Sequence reads were aligned to human reference genome (release hg19, February 2009) from University of California – Santa Cruz with the Burrows-Wheeler Aligner version 0.6.1-r104. A Genome Analysis Toolkit (version v1.0.577) was used for base quality score recalibration, INDEL realignment, and duplicate removal. Single nucleotide polymorphism (SNP) and INDEL discovery and genotyping were performed simultaneously using variant quality recalibration according to Genome Analysis Toolkit Best Practices recommendations. All the variations were annotated, and their effects on genes were predicted using variant annotation and effect prediction tool SnpEff 3.0a. Candidate single nucleotide variants were also analyzed using Sorting Intolerant from Tolerant and polyphen-2 to assess their impact on protein structure and function.

RNA sequencing

RNA was isolated from the whole blood of all family members as well as normal controls using PAXgene blood RNA isolation and purification kits as per the manufacturer's protocol. RNA was treated with DNase (Qiagen) to remove genomic DNA, and RNA quality was verified with a bioanalyzer. One microgram of total RNA was used to perform RNA sequencing using TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, Calif) as per the manufacturer's instructions. The library from each sample was sequenced on Illumina Hiseq 2000 using a standard 100-bp paired-end protocol. Reads were mapped to the University of California – Santa Cruz Homo sapiens reference genome hg19, and their relative expression values were calculated in reads per kilobase per million using CLC Biosystems Genomic Workbench software (CLC Bio, a Qiagen company, Waltham, Mass).

Exon chip genotyping

Genomic DNA (500 ng) from each sample was used for exon SNP genotyping analysis. Whole-exome SNP genotyping using Infinium HumanCoreExome BeadChips from Illumina was performed at the University of Texas Southwestern Microarray and Genomics Core facility following manufacturer's protocols. The HumanCoreExome BeadChip consists of all the tag SNPs found on the HumanCore BeadChip, plus over 240,000 markers from Illumina's Human Exome BeadChip with minimum allele frequency of more than 0.05 (mostly cSNPs or few intronic variants known to have been previously associated with clinical phenotypes), representing diverse populations including European, African, Chinese, and Hispanic individuals. The whole-exome SNP data were first analyzed using Genomestudio v3 software (Illumina) to generate SNP calls. The quality control for the recalling process was performed with a call rate cutoff set at 99% per SNP call rate per

individual, along with a 95% threshold for identity by state and a cutoff of 0.2 for the false-discovery rate applied to Hardy-Weinberg equilibrium tests. The SNP data were further analyzed using Golden Helix software (Bozeman, Mont) for statistical analysis.

Runs of homozygosity analysis

Runs of homozygosity (ROH) analysis was performed by combining all the SNPs from the exome-sequencing data and the exon SNP data with Golden Helix software. ROH analysis identifies patterned clusters of SNPs demonstrating extended homozygosity. The parameters were set for ROHs to be greater or equal to 500 kb in length with minimum of 25 SNPs. ROHs across all chromosomes present in the 2 affected siblings but not present in their father or the unaffected child were identified.

Flow cytometry

Flow cytometry was carried on whole blood and fractionated lymphocyte subpopulations. The intracellular staining of DOCK8 and cell surface staining of the indicated cell surface markers was listed as previously described.^{E1}

REFERENCES

- E1. Jing H, Zhang Q, Zhang Y, Hill BJ, Dove CG, Gelfand EW, et al. Somatic reversion in dedicator of cytokinesis 8 immunodeficiency modulates disease phenotype. *J Allergy Clin Immunol* 2014;133:1667-75.
- E2. Sun X, Becker-Catania SG, Chun HH, Hwang MJ, Huo Y, Wang Z, et al. Early diagnosis of ataxia-telangiectasia using radiosensitivity testing. *J Pediatr* 2002;140:724-31.

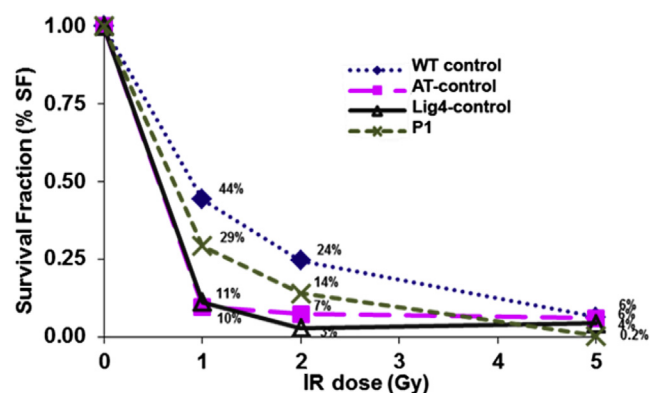


FIG E1. Two siblings have similar clinical presentations including radiosensitivity. Dose-response curve (%SF) for fibroblasts of patient 1 (green X) scored in an intermediate range of radiosensitivity, when compared with fibroblasts from a wild-type control (blue diamonds), AT control (pink squares), and a Ligase-4 deficiency control (black triangle), studied in the same experiment. *AT*, Ataxia telangiectasia; *IR*, ionizing radiation; *P1*, patient 1; *SF*, survival fraction; *WT*, wild-type.

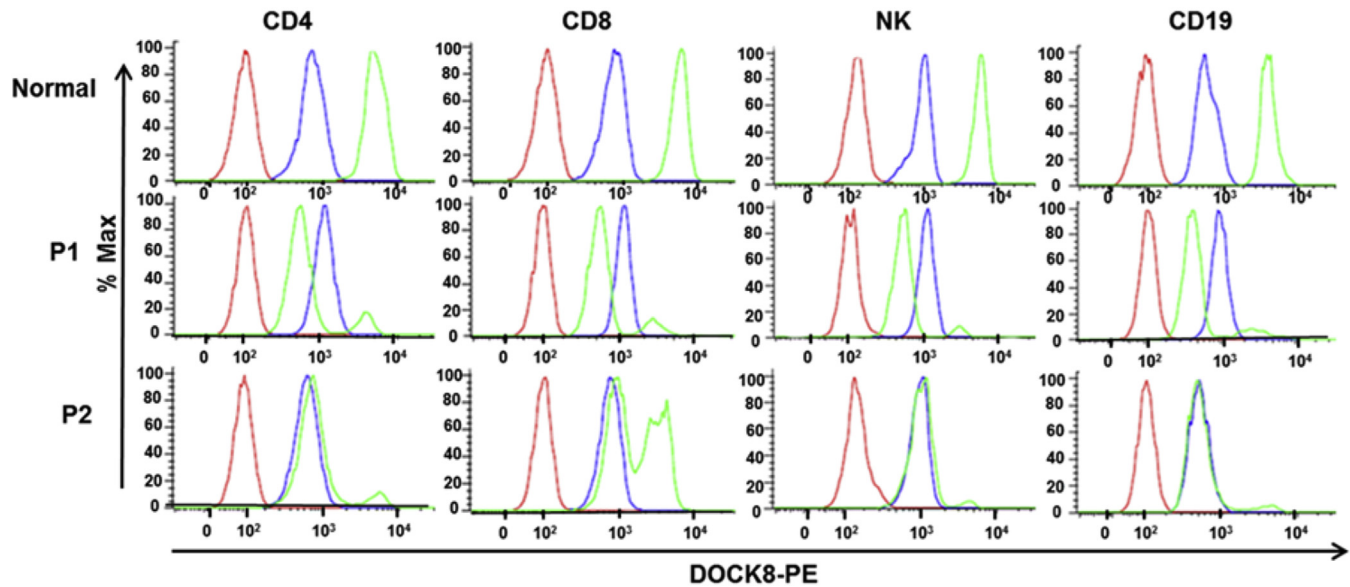


FIG E2. DOCK8 is significantly reduced in CD4, CD8, NK, and B cells from patients with a 2-bp insertion in the intronic region of DOCK8. Histogram analysis of intracellular DOCK8 staining (*green line*) compared with unstained samples (*red*) and isotype controls (*blue*). Electronic gating was used to select CD4, CD8, and natural killer (CD56) cells and B cells (CD19), which were then analyzed for the intracellular levels of DOCK8. *NK*, Natural killer.

TABLE E1. Clinical features of patients at diagnosis

Characteristic	Patient 1 (P1)	Patient 2 (P2)
Sex	Female	Male
Eczema	+	+
Cutaneous viral infections (planar warts, classic verruca, molluscum)	+++	+++
Allergic rhinitis	—	—
Asthma	—	—
Food allergies	—	—
Respiratory infections	+++	+++
Enteritis	++	++
Sclerosing cholangitis	++	+++
Fungal infections	++	++
NIH HIES score	23	22

HIES, Hyper IgE syndrome; *NIH*, National Institutes of Health.

TABLE E2. Immunologic features of patients at diagnosis

Characteristic	Patient 1 (P1)	Patient 2 (P2)
Age at evaluation (y)	8	8
White cell count (1000/mm ³)	11.1	6.0
Eosinophil	0.80	0.18
Absolute neutrophil count	7.6	4.0
Immunoglobulins (mg/dL)		
IgA	504	516
IgG	2350	1400
IgM	83	32
IgE (IU/mL)	67	97
Specific antibody responses		
Antitetanus (kU/L)	Normal	Normal
Antidiphtheria (kU/L)	Normal	Normal
Anti- <i>Streptococcus pneumoniae</i> (μg of antibody/mL)	Near absent	Near absent
Lymphocyte subsets (%/absolute counts [cells/mm ³])		
Absolute lymphocyte count (cells/μL)	976	1344
CD3 ⁺	32.9/321	55.9/751
CD4 ⁺	19.7/192	27/363
CD8 ⁺	12.6/123	24.2/325
CD19 ⁺	57/556	31.2/419
CD16 ⁺ /56 ⁺	5.7/56	10.4/140
T-cell proliferation		
PHA	Diminished	Diminished
Tetanus	Diminished	Diminished
Candida	Diminished	Diminished
Radiosensitivity assay measured by colony survival	Intermediate	22% SF (intermediate)*

SF, Survival fraction.

*Ranges given in Sun et al.^{E2}