Two novel germline *DDX41* mutations in a family with inherited myelodysplasia/acute myeloid leukemia

Ruijuan Li,[1,2] Nara Sobreira,[3,4] P. Dane Witmer,[5] Keith W. Pratz,[6] and Evan M. Braunstein[1]

[1]Division of Hematology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [2]Department of Hematology/Institute of Molecular Hematology, The Second Xiang-Ya Hospital, Central South University, Changsha, Hunan, China; [3]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [4]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [5]Center for Inherited Disease Research, Johns Hopkins University School of Medicine, Baltimore, MD, USA; and [6]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Correspondence: ebrauns3@jhmi.edu
doi:10.3324/haematol.2015.139790

Supplemental Methods:

Patient tissue acquisition

All research participants were counseled regarding the possible outcomes of WES and signed a consent form approved by the Johns Hopkins University School of Medicine Institutional Review Board. DNA samples were obtained via a skin biopsy from the proband, buccal swabs for individuals post-bone marrow transplant, and saliva was used for unaffected individuals using the Gentra Puregene kit (Qiagen). Samples were de-identified and entered into the PhenoDB database (www.mendeliangenomics.org) for the Baylor-Hopkins Center for Mendelian Genomics [1, 2].

Whole exome sequencing

The Agilent SureSelect HumanAllExonV4_51MbKit_S03723314 was used for exome capture. For library preparation and enrichment: 1ug of genomic DNA was sheared using the Covaris E210 instrument (Covaris, Woburn, MA). A hybrid protocol for library preparation and whole exome enrichment was developed at the Center for Inherited Disease Research (CIDR, unpublished) based on methods and parameters from Fisher et al. [3], applied to the reagents, volumes and parameters from the Agilent SureSelect XT kit (Santa Clara, CA) and automated protocol (p/n G7550-90000 revision B). All processing was done in 96 well plate formats using robotics (Beckman F/X, Perkin-Elmer Multiprobe II, Agilent Bravo, Beckman Biomek 2000). These reactions were carried out using the XT reagent kits, volumes and conditions described in the Agilent protocol. At pre-capture PCR the entire product was amplified, and 750ng of amplified library was used in an enrichment reaction following Agilent protocols (24 hour hybridization). Modifications include the use of KAPA HiFi HotStart ReadyMix for amplification (Kapa Biosystems, Wilmington, MA), increasing the number of pre-capture PCR cycles from 6-8 cycles and the 'off-bead' catch process from Fisher et al., was incorporated following hybridization washes. Libraries were sequenced on the HiSeq2500 platform using 100 bp paired end runs and sequencing chemistry kits TruSeq Rapid PE Cluster Kit-HS and TruSeq Rapid SBS-HS.

Primary and Secondary Analysis

Intensity analysis and base calling were performed through the Illumina Real Time Analysis (RTA) software (version 1.17.20).). Basecall files were demultiplexed from a binary format (BCL) to single sample fastq files using a demultiplexer written at CIDR as part of CIDRSeqSuite version 6.1 (unpublished). Fastq files were aligned with BWA [4] version 0.5.10-tpx to the 1000 genomes phase 2 (GRCh37/hg19) human genome reference. Duplicate molecules were flagged with Picard version 1.74. Local realignment around indels and base call quality score recalibration were performed using the Genome Analysis Toolkit (GATK) [5] version v3.1-1-g07a4bf8 or v3.3-0-g37228af. Unified Genotyper (GATK) was used for multi-sample calling (all BHCMG study samples run at CIDR to date plus HapMap and 1000 Genome control samples) of SNVs and indels using Reduced BAMs (GATK). Variant filtering was done using the Variant Quality Score Recalibration (VQSR) method (4). For SNVs the annotations of MQRankSum, HaplotypeScore, QD, FS, MQ, ReadPosRankSum were used in the adaptive error model (6 max Gaussians allowed, worst 3% used for training the negative model). HapMap3.3 and Omni2.5 were used as training sites with HapMap3.3 used as the truth set. SNVs were filtered to obtain all variants up to the 99th percentile of truth sites (1% false negative rate). For indels the annotations of QD, FS, ReadPosRankSum were used in the adaptive error model (4 max Gaussians allowed, worst 12% used for training the negative model, indels that had annotations more than 10 standard deviations from the mean were excluded from the Gaussian mixture model). A set of curated indels obtained from the GATK resource bundle (Mills_and_1000G_gold_standard.indels.b37.vcf) were used as training and truth sites. Indels were filtered to obtain all variants up to the 99th percentile of truth sites.

Identification of variants and Segregation analysis

We used the PhenoDB Variant Analysis tool to design the prioritization strategy [1]. We applied a filtering designed to prioritize heterozygous rare functional variants (missense, nonsense, splice site variants, and indels) that fit the autosomal dominant mode of inheritance and were present in the proband and 2 affected siblings. We excluded variants with a MAF > 0.01 in dbSNP 126, 129, and 131 or in the Exome Variant Server (release ESP6500SI-V2) or 1000 Genomes Proje**c**t [6].

We also excluded all variants found in our in-house controls (CIDRVar 51Mb). Reference sequences used throughout for DDX41 are NM_016222.2 and NP_057306.2, and genomic coordinates refer to hg19.

Variant confirmation

PCR for *DDX41* was performed using the following primers: 5' CAATGGCTAGGTGTCCTTCA 3' and 5' GGGCAGATGATGAGTCCATA 3'. PCR products were purified prior to Sanger sequencing using ExoSAP-IT (Affymetrix). Pyrosequencing was performed using the Qiagen advanced chemistry protocol. Custom primers were ordered from Qiagen: PCR primers 5' TCTGCAGTACACCGGATTGG 3' and 5' GATGAAGGTAGTGGCGATGC 3' and sequencing primer 5' CCTGTGTTTCCCGAG 3'. PCR products were cloned using Zero Blunt TOPO Kit (Invitrogen).

Cloning and Transfection of DDX41

Total RNA was isolated from TF-1 cells and converted to cDNA using SuperScript III (Invitrogen). The full DDX41 coding sequence was cloned into pcDNA3.1 (Invitrogen) with an EGFP c-terminal insert. Site directed mutagenesis was performed using the Q5 kit (NEB) to generate the mutant DDX41 construct. 293T cells were transfected with either WT or mutant DDX41 using Lipofectamine 2000 (Invitrogen) per manufacturer's protocol.

Western Blot

293T cells were harvested 24 hours post transfection using 0.05% Trypsin, washed twice in PBS and protein lysates were isolate using RIPA buffer (Sigma) containing protein and phosphatase inhibitors. Lysates were quantitated using BCA Protein Assay (Pierce) and 4ug total lysate was run on NuPage TrisAcetate gels (Invitrogen). Transfers were performed using the iBlot System (Invitrogen), and membranes were probed overnight at 4°C with DDX41 primary antibody (sc-98410, Santa Cruz Biotechnology) at the recommended concentration. Visualization was done using chemiluminescent substrate.

DDX41 sequencing primers

exon1-3 forward TCTCCGGAACAAATTTGTCC
exon1-3 reverse AAAACGGTGTACCAGGCTCA
exon4-6 forward CACCGTTTTGGTATCTGTCTG
exon4-6 reverse CATAACCTCACAGGCATTTGA
exon7-9 forward TTGGGGCTCTGAGAAAGTACC
exon7-9 reverse TGGAGTGGCTAAGGTAAAGGG
exon10-11 forward TGATGAGAAGGGGACAGACAA
exon10-11 reverse CAATCAGCTTCAGGGAGACTT
exon12-14 forward TGTGGCGGGTAAGCCTAAT
exon12-14 reverse TTTTCCGGCCTAACCCAT
exon15-16 forward ATGGGTTAGGCCGGAAAA
exon15-16 reverse TTGGTCTGCATAGCCTCGAGT
exon17 forward TGCTGGAAGTCAGAAACGTG
exon17 reverse TGACGTCTGTCCAGGAACCA

References

1. Sobreira N, Schiettecatte F, Boehm C, Valle D, Hamosh A. New tools for mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. Hum Mutat. 2015; 36:425-431.

2. Hamosh A, Sobreira N, Hoover-Fong J, et al. PhenoDB: A new web-based tool for the collection, storage, and analysis of phenotypic features. Hum Mutat. 2013; 34:566-571.

3. Fisher S, Barry A, Abreu J, et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol. 2011; 12:R1-2011-12-1-r1. Epub 2011 Jan 4.

4. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM arXiv:1303.3997v1 [q-bio.GN]., 2013.

5. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297-1303.

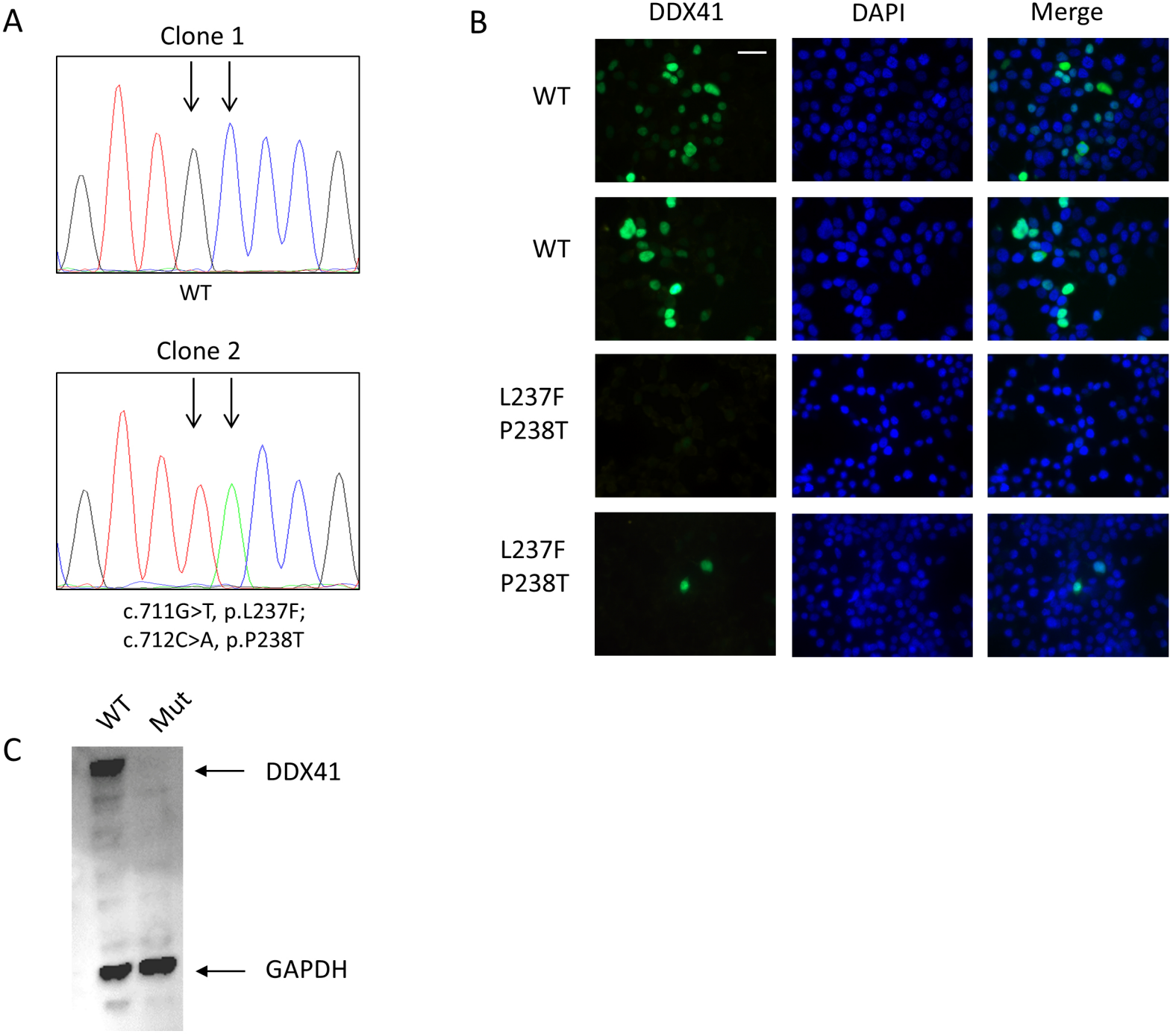
6. 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56-65.

Table S1: Candidate variants from whole exome sequencing of affected family members

Chromosome	band	RefgeneGeneName	Genotype	TotalDepth	VariantType	RefgeneGeneLocation	RefgeneExonFunction	% above 20x	StartPosition
1	1p36.13	PADI3	het	17	Snv	exonic	nonsynonymous SNV	93.1	
1	1q24.1	ALDH9A1	het	33	Snv	exonic	nonsynonymous SNV	100	
2	2p22.1	THUMPD2	het	92	Snv	exonic	nonsynonymous SNV	84.1	
4	4q12	SPATA18	het	22	Snv	exonic	nonsynonymous SNV	88.2	
5	5p13.2	C5orf42	het	78	Snv	exonic	nonsynonymous SNV	99.7	
5	5q31.2	KIF20A	het	32	Snv	exonic	nonsynonymous SNV	97.7	
5	5q34	ODZ2	het	21	Snv	exonic	nonsynonymous SNV	98.2	
5	5q35.1	LCP2;LCP2	het	59	Snv	exonic;splicing	nonsynonymous SNV	94.6	
5	5q35.2	UNC5A	het	88	Snv	exonic	nonsynonymous SNV	97.3	
5	5q35.3	DDX41	het	130	Snv	exonic	nonsynonymous SNV	99.3	176942003
5	5q35.3	DDX41	het	132	Snv	exonic	nonsynonymous SNV	99.3	176942004
5	5q35.3	AGXT2L2	het	148	Snv	exonic	nonsynonymous SNV	93.3	
5	5q35.3	SQSTM1	het	63	Snv	exonic	nonsynonymous SNV	97.4	
5	5q35.3	TRIM7	het	100	Snv	exonic	nonsynonymous SNV	89.8	
7	7p22.3	SUN1	het	89	Snv	exonic	nonsynonymous SNV	100	
7	7p14.3	PRR15	het	77	Indel	exonic	nonframeshift deletion	99.3	
7	7p14.3	WIPF3	het	65	Snv	exonic	nonsynonymous SNV	100	
7	7q31.33	ZNF800	het	124	Snv	exonic	nonsynonymous SNV	100	
9	9q21.2	VPS13A	het	31	Snv	exonic	nonsynonymous SNV	91.9	
9	9q21.2	GNA14	het	71	Snv	exonic	nonsynonymous SNV	100	
9	9q31.3	SVEP1	het	60	Snv	exonic	nonsynonymous SNV	99.2	
9	9q32	ZFP37	het	107	Snv	exonic	nonsynonymous SNV	100	
11	11p11.2	MYBPC3	het	134	Snv	exonic	nonsynonymous SNV	96.8	
11	11q12.3	AHNAK	het	122	Snv	exonic	nonsynonymous SNV	98.4	
11	11q12.3	WDR74	het	46	Snv	exonic	nonsynonymous SNV	100	
11	11q13.1	CAPN1	het	66	Snv	exonic	nonsynonymous SNV	99.9	
11	11q13.4	CLPB	het	99	Snv	exonic	nonsynonymous SNV	98.9	
12	12p13.31	SLC2A3	het	111	Snv	exonic	nonsynonymous SNV	69.8	
12	12q13.2	OR6C68	het	73	Snv	exonic	nonsynonymous SNV	99.5	
13	13q34	MCF2L	het	54	Snv	exonic	nonsynonymous SNV	94.5	
14	14q23.2	SYNE2	het	71	Snv	exonic	nonsynonymous SNV	97.3	
14	14q32.33	KIF26A	het	56	Snv	exonic	nonsynonymous SNV	94	
14	14q32.33	NUDT14	het	83	Snv	exonic	nonsynonymous SNV	87.9	
15	15q15.3	ZSCAN29	het	62	Snv	exonic	stopgain SNV	99.8	
15	15q22.31	FBXL22	het	38	Snv	exonic	nonsynonymous SNV	100	
15	15q24.3	TSPAN3	het	10	Snv	exonic	nonsynonymous SNV	81	
16	16p13.12	SNX29	het	98	Snv	exonic	nonsynonymous SNV	64	
17	17p13.3	WDR81	het	18	Snv	exonic	nonsynonymous SNV	98.7	
17	17p13.1	DNAH2	het	49	Snv	exonic	nonsynonymous SNV	99.2	
17	17p13.1	DNAH2	het	64	Snv	exonic	nonsynonymous SNV	99.2	
17	17q11.2	DHRS13	het	70	Snv	exonic	nonsynonymous SNV	100	
17	17q12	UNC45B	het	30	Snv	exonic	nonsynonymous SNV	97.5	
17	17q21.2	CNP	het	87	Snv	exonic	nonsynonymous SNV	99.8	
17	17q21.2	KCNH4	het	58	Snv	exonic	nonsynonymous SNV	95	
17	17q21.31	SPPL2C	het	109	Snv	exonic	nonsynonymous SNV	100	
17	17q23.2	MED13	het	115	Snv	exonic	nonsynonymous SNV	97.2	
17	17q25.3	RBFOX3	het	48	Snv	exonic	nonsynonymous SNV	80	
19	19p13.3	ADAMTSL5	het	44	Snv	exonic	nonsynonymous SNV	84.3	
19	19p13.3	MEX3D	het	42	Snv	exonic	nonsynonymous SNV	64.5	
19	19p13.11	SUGP2	het	66	Snv	exonic	nonsynonymous SNV	98.2	
19	19p13.11	TM6SF2	het	79	Snv	exonic	nonsynonymous SNV	93	
19	19q13.12	CAPNS1	het	15	Snv	exonic	nonsynonymous SNV	73.9	
19	19q13.2	GRIK5	het	58	Snv	exonic	nonsynonymous SNV	85.3	
19	19q13.32	BCAM	het	123	Snv	exonic	nonsynonymous SNV	95.6	
19	19q13.32	VASP	het	19	Snv	exonic	nonsynonymous SNV	92.5	
20	20q11.21	DEFB121	het	44	Snv	exonic	nonsynonymous SNV	86.6	
20	20q11.21	TM9SF4	het	39	Snv	exonic	nonsynonymous SNV	100	
20	20q13.13	MOCS3	het	126	Snv	exonic	nonsynonymous SNV	100	
20	20q13.33	ARFGAP1	het	74	Snv	exonic	nonsynonymous SNV	100	

Table S2: Exome sequencing coverage

Sample	UPIN5_1	UPIN4_1	UPIN1_1
MEAN_TARGET_COVERAGE	77.002247	88.297451	74.261767
ZERO_CVG_TARGETS_PCT	0.001137	0.001321	0.000911
PCT_TARGET_BASES_8X	0.988604	0.987208	0.990994
PCT_TARGET_BASES_10X	0.983558	0.98256	0.986979
PCT_TARGET_BASES_20X	0.941497	0.946867	0.950982
PCT_TARGET_BASES_30X	0.872543	0.891347	0.885769


A
Clone 1
WT
Clone 2
c.711G>T, p.L237F;
c.712C>A, p.P238T
B
DDX41
DAPI
Merge
WT
WT
L237F
P238T
L237F
P238T
C
WT
Mut
DDX41
GAPDH

Supplemental Figure 1. (A) The *DDX41* variants are present in the cis configuration. Representative chromatograms from Sanger sequencing of PCR clones generated from amplification of genomic DNA of the proband. In the top panel, both reference alleles are present, while in the bottom panel, both variant alleles are present. (B) Immnunofluorescence of DDX41-EGFP WT or mutant expression (green) vectors transfected into 293T cells. Nuclei are detected by DAPI staining (blue). Scale bar is 60μM. (C) Western blot using whole cell lysates of the cells transfected in (B). DDX41 protein expression is absent in cells transfected with a DDX41 mutant expression vector.