A nonsense mutation in the DNA repair factor Hebo causes mild bone marrow failure and microcephaly

Shu Zhang,1 Corinne Pondarre,2 Gaelle Pennarun,3,4 Helene Labussiere-Wallet,6 Gabriella Vera,1 Benoît France,1 Marie Chansel,1 Isabelle Rouvet,7 Patrick Revy,1 Bernard Lopez,5 Jean Soulier,8 Pascale Bertrand,3 Isabelle Callebaut,9 and Jean-Pierre de Villartay1

Inherited bone marrow failure syndromes are human conditions in which one or several cell lineages of the hematopoietic system are affected. They are present at birth or may develop progressively. They are sometimes accompanied by other developmental anomalies. Three main molecular causes have been recognized to result in bone marrow failure syndromes: (1) defects in the Fanconi anemia (FA)/BRCA DNA repair pathway, (2) defects in telomere maintenance, and (3) abnormal ribosome biogenesis.

We analyzed a patient with mild bone marrow failure and microcephaly who did not present with the typical FA phenotype. Cells from this patient showed increased sensitivity to ionizing radiation and phleomycin, attesting to a probable DNA double strand break (dsb) repair defect. Linkage analysis and whole exome sequencing revealed a homozygous nonsense mutation in the ERCC6L2 gene. We identified a new ERCC6L2 alternative transcript encoding the DNA repair factor Hebo, which is critical for complementation of the patient's DNAdsb repair defect. Sequence analysis revealed three structured regions within Hebo:

Cells from living organisms are facing DNA damages all through their lives, originating from either internal sources or inflicted by external genotoxics (Sancar et al., 2004). DNA double strand breaks (dsb's) are considered to be the most toxic DNA lesion, which results in cell death if not properly repaired. Two main DNA repair mechanisms are used to cope with DNAdsb's. Homologous recombination (HR), considered to be the most faithful process because it uses the homologous sister chromatid as a template, is restricted to the S/G2 phases of the cell cycle. The nonhomologous end-joining (NHEJ) pathway, which does not rely on a DNA template, is considered more error prone. NHEJ can cope with DNAdsb's happening in all phases of the cell cycle. One prototypical example of a programmed DNA-damaging process in the hematopoietic system is the somatic rearrangement of T cell receptors and immunoglobulin genes during the maturation of immature lymphocytes, V(D)J recombination. The efficient repair of DNAdsb's introduced during V(D)J recombination is required to maintain genome integrity, thus preventing the development of cancer and other DNA instability disorders (Alt et al., 2013). Defects in critical NHEJ factors are associated with the stalling of V(D)J recombination and the resulting arrest in lymphoid cell development leading to SCID both in human conditions.

Correspondence to Jean-Pierre de Villartay: devillartay@gmail.com

C. Pondarre’s present address is Centre Hospitalier Intercommunal de Créteil, Paris XII University, 34000 Créteil, France.

Abbreviations used: 3D, three-dimensional; AT, ataxia telangiectasia; ATM, AT mutated; CBC, cell blood count; DDR, DNA damage response; dsb, double strand break; FA, Fanconi anemia; gRNA, guide RNA; HR, homologous recombination; HSC, hematopoietic stem cell; IBMPs, inherited bone marrow failure syndrome; IR, ionizing radiation; IRIF, irradiation-induced foci; MMC, mitomycin C; NHEJ, nonhomologous end-joining; nls, nuclear localization signal; NSV, nonsynonymous variant; Orf, open reading frame; PGM, personal genome machine; SA, senescence-associated; SF, short isoform; SJ, signal joint; WES, whole exome sequencing; WGHM, whole genome homozygosity mapping; Xif, XRCC4-like factor.
and animal models (de Villartay et al., 2003). In some instances, such as Cernunnos/XRCC4-like factor (Xlf) deficiency, SCID is also associated with extraintestinal manifestations, such as microcephaly (Buck et al., 2006a), owing to the critical role of NHEJ in the central nervous system. DNA damage and repair is also an important aspect of the homeostasis of hematopoietic stem cells (HSCs) in the bone marrow (Walter et al., 2015). Indeed, defects in the NHEJ factors DNA-dependent protein kinase catalytic subunit and Cernunnos/Xlf are accompanied by HSC dysfunction, resulting in progressive bone marrow failure (Zhang et al., 2011; Avagyan et al., 2014).

Inherited bone marrow failure syndrome (IBMFS) is a generic term for clinically heterogeneous syndromes that have in common defects in the development and/or survival of various hematopoietic cell lineages (Dokal and Vulliamy, 2008). IBMFS can affect all types of blood cells, thus resembling aplastic anemia, or be restricted to only a few cell subsets like the specific red cell aplasia characteristic of Diamond–Blackfan anemia. IBMFS can appear very early in childhood or develop more progressively and is often accompanied by a series of other developmental manifestations such as facial features, skeletal anomalies, skin pigmentation, pulmonary diseases, leukoplakia, or nail dystrophy. Fanconi anemia (FA) is the most frequent cause of IBMFS and is characterized by progressive bone marrow failure often accompanied by onset of hematological malignancies (Moldovan and D’Andrea, 2009). Mutations identified in >18 different genes are responsible for FA (FANC-A to FANC-T; Hira et al., 2015; Rickman et al., 2015; Virts et al., 2015; Wang and Smogorzewska, 2015). The various FANC proteins, which are at the crossroad of several DNA repair pathways including HR, nucleotide excision repair, and trans-lesion synthesis, are of critical importance to removing interstrand DNA cross-links during replication. Indeed, cells from FA patients present with increased sensitivity to DNA cross-linking agents such as mitomycin C (MMC) and elevated spontaneous or MMC-induced chromosomal aberrations. Besides impairment of the FA DNA repair pathway, other molecular anomalies can lead to IBM FS. Telomere dysfunction, which often but not always result in telomere shortening, as observed in dyskeratosis congenita or its more severe form Hoyeraal-Hreidarsson syndrome, are responsible for early and severe IBMFS observed in pediatric patients (Glousker et al., 2015). Lastly, abnormal ribosome biogenesis is a recognized cause of a certain type of IBMFS such as observed in Diamond–Blackfan anemia, Shwachman-Diamond Syndrome, and dyskeratosis congenita (Ruggero and Shimamura, 2014). For most of these human conditions, the extent to which the bone marrow is affected often requires the transplantation of HSCs as the sole curative treatment.

Here, we analyzed a patient, born from consanguineous parents, who presented with a mild IBMFS, not necessitating HSC transplantation, which mostly affects the generation of platelets, and a slight microcephaly. Clinical and biological findings excluded known causes of IBMFS. Instead, we found that the patient’s cells were sensitive to the DNAdsb–damaging agents ionizing radiations (IRs) and phleomycin (phleo) and only weakly sensitive to MMC. Whole exome sequencing (WES) revealed a homozygous nonsense mutation in Hebo, a new DNA repair factor encoded by the ERCC6L2 gene.

RESULTS

A patient with mild IBMFS and microcephaly

A 13-yr-old boy was admitted to the pediatric emergency department for diffuse petechiae and bruises developing a few days after a febrile episode of unknown etiology. The cell blood count (CBC) revealed profound thrombocytopenia (platelet count of 4 G/liter) associated with mild anemia (hemoglobin level of 90 g/liter), macrocytosis (mean red cell volume of 104 fl), and moderate neutropenia (neutrophil count of 0.7 × 109/liter). Bone marrow smears and biopsies showed reduced cellularity (<15%) without myelodysplastic features, consistent with aplastic anemia. No cytogenetic abnormalities were found. The patient was the fourth of five children of a healthy consanguineous (first cousins) couple. One of his sisters died at 7 mo of unexplained sudden infant death, and his youngest brother was affected by developmental delay consecutive to perinatal hypoxia. All other children and both parents had a normal CBC. The patient’s past medical history was unremarkable except for learning difficulties and attention deficit disorder. Birth weight (3,070 g) and height (49 cm) were normal and appropriate for gestational age (39 wk), but a mild microcephaly was noted at birth (head circumference of 34 cm). Additional findings included increased hemoglobin F levels (7%) suggesting mild dyserythropoiesis, vascular abnormalities in the right frontal lobe, and bilateral pyleurotetal junction abnormalities. There were no cardiopulmonary defects or mucocutaneous changes. A few weeks later, CBC parameters improved spontaneously and stabilized (Table S1). The patient is now 20 yr old, and CBCs show a hemoglobin level of 120 g/liter, stable macrocytosis (mean red cell volume of 104 fl), reticulocyte count of 75 × 109/liter, moderate neutropenia (neutrophil count of 1.04 × 109/liter), thrombocytopenia (platelet count of 59 × 109/liter), and a subnormal lymphocyte count (1.47 × 109/liter). Previous immunophenotyping had revealed a slight diminution in T lymphocytes (871/µl), with a significant decrease in naïve CD4 T cells (49% CD45RA+CD4+) as well as a B lymphocyte deficiency (92/µl). Immunoglobulin levels were not affected (14 g/liter IgG, 1.06 g/liter IgA, and 1.17 g/liter IgM). There has been no clonal evolution or progression to myelodysplasia. He achieved normal puberty development but remained <1 SD for weight and head circumference and <0.5 SD for height. We found neither gastrointestinal nor skeletal abnormalities, nor fatty infiltration of the pancreas suggestive of Shwachman-Diamond syndrome. However, the association of the bone marrow failure and the microcephaly/dysmorphosis suggested a possible FA in this patient.

Exclusion of classical FA or telomeropathy syndromes

One hallmark of FA-deficient cells is a pronounced sensitivity to the DNA cross-linking agent MMC (Pinto et al., 2009).
The sensitivity of the patient's SV40-transformed fibroblasts to increasing doses of MMC, although significantly higher than that of control cells, was by no means comparable to that of cells from an FA patient (Fig. 1A). Moreover, an EBV lymphoblastoid cell line from the patient did not demonstrate the typical G2/M accumulation after MMC treatment observed in FA cells and only arrested at the highest dose (50 ng/ml), similarly to lymphoblastoid cells from his father and an unrelated control (Fig. 1B). Lastly, although a large subset of FA patients with mutations in the FA core proteins unveiled an impaired ubiquitination of FANCD2 after MMC treatment (García-Higuerá et al., 2001), this was again not the case with our patient's cells (Fig. 1C). Altogether, these analyses ruled out a major defect in the Fanconi pathway as the primary cause of the bone marrow failure identified in this patient. Nevertheless, a significant increase in chromosomal aberrations (i.e., chromosomal breakages and radial chromosomes) was found in the patient's cells upon exposure with MMC (Fig. 1D), aberrations that are reminiscent of those associated with FA (Auerbach, 2009), which together with the slight increase in MMC sensitivity (Fig. 1A) raised the possibility of either a hypomorphic FA condition or the existence of a different type of DNA repair defect responsible for the patient's clinical condition. Telomere dysfunction constitutes another likely cause of early bone marrow failure (Touzot et al., 2012). However, analysis of telomere length in the peripheral blood lymphocytes of the patient, compared with that of three siblings and his parents, did not demonstrate any abnormal shortening of telomeres in these cells with a mean telomere length of ~11–14 kb in all individuals (Fig. 1E). In accord with the apparent normal telomere function, primary fibroblasts from the patient did not experience premature replicative senescence as judged by senescence-associated (SA) β-galactosidase (β-gal) staining (Fig. 1F).

**Patient cells have defects in DNAdsb repair**

To get further insight into the DNA repair pathway that might be affected in this patient, we analyzed his SV40-transformed fibroblasts and EBV B cells for sensitivity against a panel of genotoxic drugs causing various types of DNA lesions (Fig. 2A). Although the patient's fibroblasts were not overtly sensitive to methyl methanesulfonate, UV light, or poly–ADP-ribose polymerase inhibitor, a sharp sensitivity to both phleomycin and IR was noticed, almost comparable to that observed with NHEJ-deficient Cernunnos/Xlf mutant cells. The sensitivity to phleomycin was not restricted to fibroblasts, as the EBV B cells demonstrated a sensitivity similar to that of cells from a DNA ligase IV patient, in contrast to the normal sensitivity observed in EBV B cells from the two parents and an unrelated control (Fig. 2A). Increased sensitivity to both phleomycin and IR suggested a possible defect in DNAdsb repair, as observed in various NHEJ-deficient conditions. DNAdsb repair capacity of the patient's primary fibroblasts was analyzed through the kinetics of 53BP1 irradiation-induced foci (IRIF) formation 1 h and 24 h after IR (Fig. 2B). γ-H2AX or 53BP1 foci were used as surrogate markers of DNAdsb (Rogakou et al., 1998). Although 53BP1 IRIF, which form upon DNA damages, were no longer present after 24 h in control cells, attesting for the efficient DNA repair process, they persisted in Cernunnos/Xlf mutant cells, attesting for the DNAdsb repair deficiency in these cells. Likewise, a statistically significant retention of 53BP1 IRIF was noted in the patient's cells, although to a lesser extent as compared with the Cernunnos-deficient situation.

From this series of analyses in the patient's cells, we conclude that a general DNA repair defect affecting resolution of DNAdsb's is the probable cause of the bone marrow failure and other developmental defects presented by this child.

**A nonsense mutation in the ERCC6L2 gene**

To firmly exclude an FA molecular defect, we performed WES and initially filtered the results on the 18 known FA-causing genes (Wang and Smogorzewska, 2015) and did not identify variants with frequencies <1% in the public and in-house databases for these genes. The patient being born to consanguineous parents, we also undertook whole genome homozygosity mapping (WGHM) from DNA in the family, which highlighted 14 compatible chromosomal regions (Fig. 3A). These regions were subsequently used to filter nonsynonymous variants (NSVs) identified through WES performed on the patient's DNA after an initial filtering on homozygous variants with frequencies <1% in single nucleotide polymorphism and in-house databases (Fig. 3B). A nonsense NSV in the ERCC6L2 gene (Chr9 [GRCh37]:g.98718278C > T, c.1963C > T, and p.Arg655*) was selected as the best candidate given the plausible implication of this factor in the DNA damage response (DDR; Fig. 3B). Although this variant was already described in a database (Single Nucleotide Polymorphism database accession no. rs147948835), its overall frequency of 1.6 × 10^{-4} and of 2.6 × 10^{-4} in the European population (The Exome Aggregation Consortium) was low enough to consider it as a possible pathological NSV. Moreover, this NSV had a very high PHRED score of 49 using combined annotation–dependent depletion (Kircher et al., 2014). Sanger resequencing confirmed the appropriate inheritance of the R655X mutation in the patient's family with respect to WGHM results (Fig. 3C). The same mutation was initially described by Tummal et al. (2014).**

**Identification of Hebo**

Analyzing the assembly of the ERCC6L2 gene in various species through the Ensembl database (Fig. 4A), we were puzzled by the difference between the longer transcript annotated in the rat (ENSRNOT00000025965) coding for a 1,533-aa protein and the 712-aa human annotated protein, raising the possibility of an error in the automatic annotation of the human counterpart. Blast analysis of the rat ERCC6L2 open reading frame (orf) over the human genomic region spanning the ERCC6L2 gene established that the first 14 exons correspond to the human ERCC6L2-001 annotated...
transcript, whereas the additional five exons present in the rat matched short alternative human annotated transcripts (ERCC6L2-006 and ERCC6L2-007) but with a gap of 347 bp within the putative exon 16. Altogether, this suggested the possible expression of a longer ERCC6L2 transcript in humans, comparable with the rat situation but not annotated as such in the Ensembl database. We designed PCR primers in the first exon of transcript ERCC6L2-001 and in the last exon of transcript ERCC6L2-006 to amplify cDNA from human activated peripheral blood leukocytes. This resulted in a 4.7-kb specific band attesting to the existence of an expressed ERCC6L2 long form transcript in humans (unpublished data). Sequencing the PCR product established the 4,686-bp orf encoding a 1,561-aa protein showing 73% identity with the rat ERCC6L2 counterpart. The alternative splicing event between the short and long isoforms intervenes through the usage of a splice donor site within exon 14 after amino acid E711. This ERCC6L2 long isoform is now quoted in Gen-
Bank (accession no. NM_020207) after a recent update of human genome sequence annotation but without experimental evidence of its expression. For simplicity, we propose to rename this new isoform of the ERCC6L2 gene Hebo (helicase mutated in bone marrow failure). Hebo differs from the ERCC6L2 short isoform (SF) by the replacement of the last V712 residue with a new 850-aa peptide. Both isoforms include the R655X mutation identified in the patient.

We next evaluated the pattern of expression of the short and long ERCC6L2 transcripts in various tissues. cDNA from a panel of tissues was analyzed by real-time quantitative PCR using pairs of oligonucleotide primers and probes specific for the ERCC6L2-SF and Hebo isoform. The expression of ERCC6L2 is ubiquitous without evidence for a tissue-specific regulation of the splicing between the two isoforms (Fig. 4 B). Finally, the subcellular localization of ERCC6L2-SF and Hebo was determined after transfection of 293 T cells with ERCC6L2-SF and Hebo fused to GFP. Although ERCC6L2-SF is mostly localized in the cytoplasm, Hebo localization is restricted to the nucleus as expected for a DNA repair factor (Fig. 4 C).

**Functional validation of Hebo mutation**

As a first mean to corroborate the possible implication of ERCC6L2 in the DDR after DNA dsb’s, we disrupted the ERCC6L2 gene in U2OS cells through CRISPR/Cas9 mutagenesis. Single molecule sequencing of PCR products spanning the guide RNA (gRNA) site from U2OS transfected with the dual spCas9 and gRNA pX330 expression plasmid revealed a level of 82% indels of various sizes (unpublished data). The bulk mutagenized U2OS cells presented increased phleo sensitivity comparable with that of U2OS cells in which DNA ligase IV had been disrupted by CRISPR/Cas9, well above the sensitivity observed in isogenic WT U2OS cells, attesting for the critical role of ERCC6L2 in DNA repair after phleo-induced DNA damage (Fig. 5 A).

---

**Figure 2. DNA repair defect in patient’s cells.** (A) Survival of SV40-transformed and EBV B cell lines from the patient and appropriate negative controls upon treatment with increasing doses of various genotoxic agents. These experiments were performed between one and four times depending on the drug. PARPi, poly–ADP-ribose polymerase inhibitor. MMS, methyl methanesulfonate. ATR, AT and rad3-related protein. (B) Kinetics of irradiation-induced 53BP1 foci formation in primary fibroblasts from the patient, a healthy control, and a Cernunnos-mutated patient. A representative image of cells treated with 2 Gy IR and quantification of IRIF is shown. Data are mean ± SD. ****, P < 0.0001 for a Mann-Whitney test. This experiment was performed two times. Bar, 10 µm. Ctrl, control. Cernu, Cernunnos.
To further validate the R655X mutation in the ERCC6L2 gene as being responsible for the DNA repair deficiency observed in the patient’s cells, we PCR amplified both the ERCC6L2-SF (Ensembl transcript ENST00000288985) and Hebo orf, cloned them into an Ires-GFP lentiviral vector, and transduced the patient’s fibroblasts. The resulting mix populations of transduced (GFP+) and untransduced (GFP−) cells were analyzed through multicolor competition assay after phleo treatment (Smogorzewska et al., 2007). Neither transduction with the empty vector nor with the ERCRC6L2-SF–expressing vector conferred any selective advantage in culture, as evaluated by the GFP index of transduced over nontransduced cells upon treatment with 100 ng/ml phleo (Fig. 5 B). In contrast, introduction of WT Hebo into the patient’s fibroblasts conferred a strong selective advantage over cells transduced with the empty vector as judged by the 2.5-fold increase in GFP-expressing cells 3 wk after phleo treatment. In addition, GFP-sorted Hebo–transduced cells recovered an overall phleo sensitivity comparable with WT control cells, whereas the patient’s fibroblasts transduced with ERCRC6L2-SF remained as sensitive as their untransduced counterparts (Fig. 5 C). The absence of complementation by ERCC6L2-SF was not the consequence of a reduced expression of the protein compared with Hebo (Fig. 5 D).

Altogether, we conclude that the DNA repair deficiency identified in the patient’s cells is caused by the R655X homozygous nonsense mutation in the ERCC6L2 gene that results in the truncation of about half of the new DNA repair factor Hebo.

**Structure of Hebo**

Analysis of the Hebo protein sequence revealed three structured domains, the first two of which are common with ERCRC6L2-SF (Ensembl transcript ENST00000288985) and Hebo orf, cloned them into an Ires-GFP lentiviral vector, and transduced the patient’s fibroblasts. The resulting mix populations of transduced (GFP+) and untransduced (GFP−) cells were analyzed through multicolor competition assay after phleo treatment (Smogorzewska et al., 2007).
Moreover, highly conserved amino acids are observed in a few positions (Y, H, N, V, G, E, S, P, Q, M, A, L, and Y), suggesting that they play key roles in the function and/or structure of this domain. The HEBO domain matches the shorter profile described in the Pfam database (accession no. PF14773). Apart from the new HEBO domain, the N-terminal half of Hebo contains a Tudor domain (K26–C89) that shares the most striking sequence similarities with PHF20, TDRD3, SMN, and SPF30 (Fig. 7), which contain similar methyl-binding cages (Tripsianes et al., 2011; Cui et al., 2012; Liu et al., 2012). Three aromatic residues (Fig. 7, labeled 1–3) constitute Tudor domains, with a fourth additional one (Fig. 7, labeled 4’) existing in TDR3, SMN, and SPF30, whereas this last position is occupied by an aspartic acid in PHF20 (Fig. 7, labeled 4). The first two aromatic residues in positions 1 and 2, however, are not present in the Hebo Tudor domain, these being replaced by a serine (strand $\beta_1$) and a cysteine (strand $\beta_2$) residue, whereas an aromatic side chain is still present in positions 3 and 4’. This suggests that the Hebo Tudor domain, which is not conserved in yeast species, has evolved toward a different function than methyl binding. This contrasts with the Tudor domain present in other DNA repair factors such as 53BP1, which acts as a histone methylation reader at sites of DNAds’.
The Tudor domain is followed by a helicase/ATPase domain (T129-H693), the 3D putative structure of which is represented in Fig. 8. The core of the ATPase domain contains, like Rad54, two RecA-like core domains, with an interface between them forming a deep groove accommodating the nucleotide-binding pocket and part of a DNA-binding site. Amino acids typical of the SF2 family, as well as of the SWI2/SNF2 family, are well conserved in Hebo and ERCC6L2-SF. Three large insertions (Fig. 8, A and B, labeled 1–3), however, are found in Hebo/ERCC6L2-SF relative to the Rad54 structures with which they were aligned, one of which (labeled 3) is oriented toward the groove. The R655X mutation shortens the ATPase domain of its last 38 residues. This fragment encompasses almost all of the second part of the discontinuous HD2 domain (Fig. 8 C), which participates in the structural integrity of the whole HD2 domain.

Hebo is recruited to DNA damages

As a first step in the description of the precise function of Hebo during DDR, we analyzed its recruitment at sites of DNA damage after laser microirradiation. GFP-Hebo and GFP-NBS1 fusion proteins were transfected into HeLa cells. 48 h after transfection, DNA damage was inflicted through laser irradiation, and the recruitment of GFP-Hebo and GFP-NBS1 was followed by live imaging. As shown in Fig. 9 A and quantified in Fig. 9 B, Hebo is recruited at the site of DNA damage with the same kinetics observed with NBS1. Interestingly, the forced expression of GFP–ERCC6L2-SF into the nucleus through the addition of the SV40–nuclear localization signal (nls) did not lead to the recruitment of this isoform to the DNA lesion. Altogether, these results further strengthen the notion that Hebo acts as a bona fide DNA repair factor and strongly suggests that the C-terminal–specific region of Hebo is critical for its localization in the nucleus and its recruitment to DNA lesions. The recruitment of NBS1 to DNAdsb's represents one of the earliest events in the process of DDR (Lukas et al., 2003). The comparable early kinetics of Hebo and NBS1 recruitment after laser microirradiation prompted us to search for a possible functional relationship between these two events. For this, we transfected NBS1-deficient fibroblasts obtained from a Nijmegen breakage syndrome patient with GFP-Hebo and found that the absence of NBS1 precluded the recruitment of Hebo at the site of DNA lesions (Fig. 9 C). Conversely, the cotransfection of DsRed-NBS1 and GFP-Hebo in the NBS1-defective cells restored the Hebo recruitment to DNA damage. In contrast, the localization of Hebo to DNA damage does not depend on ataxia telangiectasia (AT) mutated (ATM) as shown by its detectable recruitment in fibroblasts from an AT patient. We conclude that the recruitment of Hebo at DNA breaks is a very initial event during DDR and is dependent on the presence of NBS1.
Figure 6. The HEBO domain. (A) Schematic representation of the Hebo and ERCC6L2-SF domain architecture. The asterisk indicates the stop codon at R655 in the patient. (B) Multiple alignment of the HEBO domain sequence of Hebo proteins from representative species. Identical amino acids are represented as white on a black background, and amino acid similarities are boxed. Lines indicate the position of hydrophobic clusters, which are conserved within the whole family (Table S1). GenBank/NCBI RefSeq accession nos. are NP_064592.2 (Homo sapiens), AAH91795.1 (Danio rerio), XP_011663528 (Strongylocentrotus purpuratus), XP_002151909.1 (Talaromyces marneffei), CDP26308.1 (Podospora anserina), XP_001391908.1 (Aspergillus niger), and CAD37001.1 (Neurospora crassa). (C) Hydrophobic cluster analysis of the C-terminal region of ERCC6L2 reveals a globular domain (boxed), which is conserved in the ERCC6L2 orthologues. The sequences are shown on duplicated α-helical nets, on which hydrophobic amino acids are contoured and form clusters, which mainly correspond to the regular secondary structures forming globular domains. Conservation of hydrophobic clusters is indicated in green, whereas sequence identities are reported in orange. A more variable region (arrows) is highlighted in the middle of the domain, likely corresponding either to a large loop or to a linker separating two distinct subdomains. S, T, P, and G stand for serine, threonine, proline, and glycine, respectively. Orange and yellow colorings are used to designate identical and similar amino acids, respectively.
Hebo is not essential for DNAdsb repair by NHEJ or HR

We analyzed the implication of Hebo during NHEJ using a V(D)J recombination assay, as this mechanism relies exclusively on NHEJ. The formation of signal joints (SJs) during V(D)J recombination consists of the error-free fusion of recombination sequences in WT cells but is highly imprecise in cells harboring NHEJ defects such as Cernunnos or DNA ligase IV mutants (Buck et al., 2006a,b). Fibroblasts from the Hebo patient were cotransfected with the RAG1/2 expression constructs together with the V(D)J reporter substrate pRec-SJ. SJs were PCR amplified and analyzed by single molecule sequencing on a personal genome machine (PGM; IonTorrent). As shown in Fig. 10 A, whereas 40% of SJs are imprecise in the case of Cernunnos cells with various lengths of deletions as previously described (Buck et al., 2006a), 92% of SJs recovered from Hebo cells were precise compared with the 95% in control cells, arguing against a major role of Hebo during NHEJ. The function of Hebo during HR was analyzed using the well established U20S-DRGFP experimental system (Pierce and Jasin, 2005). Expression of Hebo and RAD51 as controls was down-regulated by means of siRNA, and levels of HR were determined after introduction of DNAdsb’s through I-SceI. As shown in Fig. 10 B, whereas extinction of RAD51 severely impaired HR (1.97% GFP+ vs. 4.29% GFP+ for control; P < 0.0001), the efficacy of HR was only slightly (not statistically significant) diminished (3.60% GFP+; P = 0.054) in the condition with siRNA against Hebo. This suggests that Hebo is not critically required for HR, either, although it may participate to some extent. These two results have to be considered in light of the mild clinical presentation of the patient, which does not support, despite the existence of a homozygous nonsense mutation, a major role of Hebo during NHEJ or HR. Further studies are now required to better understand the implication of Hebo during DNA repair in general.

DISCUSSION

We analyzed the cells of a young patient presenting with mild bone marrow failure and microcephaly to demonstrate the existence of a general DNA repair defect in these cells that affects resolution of DNAdsb’s as judged by specific phleomycin and IR sensitivity. Genetic analyses uncovered a nonsense mutation (R655X) in the ERCC6L2 gene as the molecular cause of this deficiency. Tummala et al. (2014) previously reported the same mutation and concluded that the 712-aa protein encoded by ERCC6L2 participates in DDR. Our results only partly support this conclusion, primarily based on the fact that the 712-aa ERCC6L2-encoded protein does not complement the DNA repair defect in our patient’s cells, a point that was not addressed in Tummala’s study. Instead, we identified a new protein, Hebo, encoded by an alternative ERCC6L2 transcript, in which an additional 850-aa peptide is appended to the previously annotated ERCC6L2-SF. Hebo presents features of a bona fide DNA repair factor. It is ubiquitously expressed, localized primarily in the nucleus, and rapidly recruited to the site of DNA lesions in an NBS1-dependent manner. Most importantly, the introduction of WT Hebo in the patient’s cells rescues their DNA repair defi-
Figure 8. **ERCC6L2/Hebo ATPase domain.** (A) Sequence alignment of the human ERCC6L2 SWI/SNF2 ATPase domain with those of the Rad54 proteins of known 3D structures (D. rerio, Protein Data Bank accession no. 1z3i [Thomä et al., 2005]; and Sulfolobus solfataricus, Protein Data Bank accession no. 1z63 [Dürr et al., 2005]). SWI2/SNF2-specific elements are depicted in magenta (HD1) and green (HD2), whereas the RecA-like helicase domains are blue (lobe1) and red (lobe2). Helicase-specific domains are shown with numbers, and SWI-SNF2–specific elements are depicted with letters (Thomä et al., 2005). Yellow boxes (alignment) indicate the positions of insertions in the ERCC6L2 structure relative to the Rad54 templates. (B) Global view of the 3D structure model (ribbon representation) of the human ERCC6L2 SWI/SNF2 ATPase domain, constructed on the basis of its alignment with sequences of known 3D structures (see A). The same colors are used as in A. Circles (3D structure) indicate the positions of insertions in the ERCC6L2 structure relative to the Rad54 templates (yellow boxes in A). (C) Focus on the HD2 domain showing the location of arginine 655. N-ter, N-terminal. C-ter, C-terminal.
ciency as judged by the recovery of a normal sensitivity to the DNAdsb inducer phleo. Whether the ERCC6L2-SF exerts any biological function or represents a nonfunctional alternative transcript remains to be investigated. Tummala et al. (2014) suggested a possible link to mitochondrial function, a facet not addressed in our study. Alternative splicing events have been previously described in the case of the RecQ5 helicase, for which the long form is localized in the nucleus, whereas the SF remains in the cytoplasm (Shimamoto et al., 2000). Both isoforms include the ATPase domain like in ERCC6L2-SF/Hebo. Although in the case of RecQ5 the different splicing forms appear to be tissue specific, we did not find strong evidence for such a regulation in the case of ERCC6L2/Hebo.

Hebo carries three noticeable structured regions, a Tudor domain, an ATPase domain, and HEBO, a new globular domain conserved through evolution but only present in direct homologues of Hebo. The presence of a Tudor domain adjacent to the ATPase fold suggested that this region might be responsible for the recruitment of Hebo to chromatin after DNA damage. One of the recognized functions of the Tudor domain is its interaction with methylated proteins, which makes it a histone mark reader, particularly within the 53BP1 DNA repair factor (Charier et al., 2004; Botuyan et al., 2006; Faure and Callebaut, 2013; Fradet-Turcotte et al., 2013). Indeed, the binding of the 53BP1 Tudor domain to histone H4 dimethylated on lysine 20 (H4K20me2) is essential during DDR (Mallette et al., 2012; Tang et al., 2013). Unexpectedly, the SF of ERCC6L2 containing only the Tudor domain and the ATPase domain is not recruited to DNA damage, even when its nuclear localization is enforced by the addition of the SV40-nls, which argues against a critical role of the Hebo Tudor domain in direct chromatin binding during DNA repair, as opposed to 53BP1. Structural differences in the Hebo Tudor domain, in particular the replacement of two and four aromatic residues critical for methyl binding by serine and cysteine, suggest that it has probably evolved to serve a different function. Although this would be largely unprecedented, a few examples exist, such as the Tudor domain in the Tejas (Tej) protein in Drosophila melanogaster, which drives interaction with its partner Aubergine (Aub) independently of arginine methylation (Patil and Kai, 2010). The search for specific protein partners of the Hebo Tudor domain should help clarify this interesting issue. If the Hebo Tudor domain is not directly involved in chromatin binding, what other regions of the protein are critical for Hebo’s recruitment to DNA lesions? Because only the full-length Hebo, as opposed to the ERCC6L2-SF, is recruited to DNA damage, this suggests that the C-terminal half of the protein, perhaps the newly described HEBO domain, is critical for this function.
this implies a direct binding of Hebo to chromatin or its conveyance through interaction with a cargo protein remains to be determined. The second noticeable region of Hebo is the ATPase domain adjacent to the Tudor domain, which suggests a possible helicase function for Hebo. DNA/RNA helicases are classified into six superfamilies that perform various functions (Singleton et al., 2007). The SF2 family is the largest and includes the DEAD-box helicases, the RecQ family, and the Snf2-related enzymes. Defects in the members of the RecQ family Bloom syndrome helicase, Werner syndrome helicase, and RECQL4, known to be critical for the repair of DNAdsb’s through HR, result in Bloom, Werner, and Rothmund-Thomson syndromes, respectively, which associate developmental anomalies and a clear cancer predisposition (Hickson, 2003; Suhasini and Brosh, 2013). Another family of DNA helicase, the iron-sulfur (Fe-S) cluster helicases, includes the XPD and XPB factors mutated in xeroderma pigmentosum, FANCJ, the mutations of which are responsible for one group of FA patients, the DDX11 factor involved in Warsaw breakage syndrome, and the more recently described RTEL1 helicase with a more dedicated function to telomere maintenance (Le Guen et al., 2013; Suhasini and Brosh, 2013). In contrast to RecQ and Fe-S helicases that exert their function directly on DNA through their ATP- and Mg2+-dependent unwinding activity, the helicases of the Snf2/Swi2 family lack classical helicase activity in the sense of DNA strand displacement but are motor proteins that act as chromatin remodelers (Ceballos and Heyer, 2011). One prototypical member of this family is the RAD54 protein, which together with RAD51 plays a major role during various phases of HR. Hebo most likely belongs to the Snf2/Swi2 family of chromatin remodelers given its sequence/structure similarities with RAD54.

The third domain identified in Hebo is new. The 200-aa HEBO domain is well conserved during evolution but is only present in direct homologues of Hebo. To date, no clear function can be ascribed to this domain outside the fact that the 850-aa peptide to which it belongs is required for both the nuclear localization and the recruitment of Hebo to DNA damage. Identification of putative interacting partners of the HEBO domain should help clarify its function in the future.

One important question to understand is what the specific function of Hebo is among the panel of other chromatin remodelers in general and during the repair of DNAdsb’s in particular, and how and why its defect translates...
into clinical manifestations observed in the described patient here. The patient suffers from mild bone marrow failure that did not necessitate HSC transplantation up to now, in contrast to the much more severe aplasia conditions observed in FA or Hoyeraal-Hreidarsson syndrome. The premium physiological causes underlying IBMFS in FA and other related disorders have remained poorly understood, in particular given the fact that several KO mouse models do not recapitulate this phenotypic trait spontaneously (Parmar et al., 2009). Nevertheless, it was proposed that a chronic activation of the TP53 DDR was involved in the decline of HSCs in FA patients (Ceccaldi et al., 2012), although the source of DNA damages in HSCs remained unclear. Walter et al. (2015) recently demonstrated that DNA damages in HSCs is the direct consequence of their exit from the quiescent state as experimentally provoked by injection of polynucleotides: polycytidyl acid to mimic a viral infection and induce a type I IFN response. The same observations were made using more physiological stimuli such as IFN-α, G-CSF or serial bleeding. Perhaps most importantly, Fanca<sup>−/−</sup> mice, but not WT or untreated Fanca<sup>+/−</sup> mice, subjected to several rounds of polynucleotides: polycytidyl acid–induced HSC activation succumbed to severe aplastic anemia, thus mimicking the bone marrow failure observed in human FA patients. In this context, it is of interest to note that the clinical symptoms of the patient we studied, in particular the thrombocytopenia, arose subsequent to a febrile episode of unknown etiology. It is therefore tempting to speculate that an acute triggering of HSCs during that episode resulted in DNA damages that require Hebo for their proper repair. IBMFS is not restricted to patients (or mouse models) harboring gene defects in the Fanconi pathway. Indeed, a recent study noticed a cytopenia requiring platelet transfusion in 80% of a dozen of DNA ligase IV–mutated patients (Murray et al., 2014). Likewise, one of the patients we initially described with a Cernunnos/Xlf deficiency (Buck et al., 2006a) subsequently developed bone marrow failure (unpublished data), a characteristic that can be explained in light of a study demonstrating a premature aging of HSCs in a Cernunnos/Xlf KO mouse model (Avagyan et al., 2014). Lastly, congenital bone marrow failure caused the premature death of mice harboring three alanine substitutions in the DNA-dependent protein kinase catalytic subunit in a cluster of threonine residues known to be phosphorylated as part of DDR (Zhang et al., 2011). One could argue that Hebo deficiency may be even more connected to these three examples of NHEJ defect–related bone marrow failure, as these conditions are associated with a DNA DSB repair defect (phleo) like in Hebo-deficient cells rather than the inability to cope with interstrand cross-links (MMC) as observed in FA cells.

FA patients, as other conditions of DNA repair deficiency, present with a high risk of developing hematopoietic malignancies and other types of solid cancers. It is therefore legitimate to consider that Hebo may represent yet another critical genome caretaker, although this question cannot obviously be addressed at the present time with so few affected patients. The development of a Hebo–deficient mouse model should be very valuable in the future to address these important questions by providing the possibility to analyze a large series of affected individuals, whereas our case is unique.

**MATERIALS AND METHODS**

**Cells.** Informed consent for our study was obtained from the family in accordance with the Helsinki Declaration. The Institut National de la Santé and de la Recherche Médicale Institutional Review Board also approved this study. Primary fibroblasts were obtained through culture of the patient’s skin biopsy and further transfected by transfection with SV40 large T expression plasmids. Control primary fibroblasts (GM05757 and GM08399) were from apparently healthy donors of 7 and 19 yr, respectively (Coriell Cell Repositories). A-T cells (GM05823) were from a 18-yr-old patient suffering from AT (Coriell Cell Repositories). Cernunnos-deficient cells were described previously (Buck et al., 2006a). DNA ligase IV–deficient EBV B cells were from our undescribed patient. FA fibroblasts were from our undescribed FANC-G-mutated patient. AT and rad3–related protein–deficient fibroblasts were from patient F02-98 described previously (O’Driscoll et al., 2003).

**Sensitivity to genotoxics and complementation survival assay.** Cellular sensitivity was performed as previously described (Buck et al., 2006a). In brief, cells were seeded at low density in complete culture medium and incubated with increasing doses of genotoxic drugs or subjected to ionizing or UV radiations. Surviving cells were counted using flow cytometry with a constant amount of fluorescent beads. After 7 d, survival fraction was calculated relative to untreated cells. IRIF were analyzed using anti-53BP1 staining as described previously (Buck et al., 2006a). In brief, early passaged primary fibroblasts from the healthy control, patient, and Cernunnos/Xlf patient were X irradiated (2 Gy) on coverslips. Cells were harvested 1 and 24 h after irradiation, fixed in 4% paraformaldehyde, and permeabilized in 1% PBS, 1% BSA, 0.1% Triton X-100, and 1 mM EDTA before staining for 1 h with anti-53BP1 (Novus Biologicals). Cells were further incubated with Alexa Fluor 488 secondary antibodies (Molecular Probes and Invitrogen), stained with DAPI (Invitrogen), and mounted in FluorSave (EMD Millipore). Images were taken under the same quantity of laser power and exposure time with epifluorescence microscopy (Axioplan; ZEISS). Foci numbers were quantified for each cell with ImageJ software (National Institutes of Health). For complementation of phleo sensitivity, patients’ SV 40–transformed fibroblasts were transduced with a bicistronic ER.CC6fl2-Ires-GFP–expressing lentivirus or empty vector and treated or not with 100 ng/ml phleo for up to 3 wk. The number of GFP-positive cells was scored at various time points to assess for relative selective advantage of transduced (GFP<sup>+</sup>) over untransduced (GFP<sup>−</sup>) cells. The absolute sensitivity of transduced cells was
also determined by incubating cell-sorted GFP-positive transduced cells with increasing doses of phleo and determining the surviving fraction after 7 d. For G2M block analysis, EBV-transformed B cells were treated with increasing doses of MMC and collected 48 h later. The cells were washed in PBS, resuspended in 100 µl PBS, fixed in 1 ml of 70% EtOH, and kept overnight at ~20°C. Fixed cells were washed in PBS, resuspended in 300 µl PBS with 5% Triton X-100 and 25 µg/ml propidium iodide, and analyzed by FACS.

Cytenetic analysis. Primary fibroblasts of similar cell culture passages were seeded at a density of 4 × 10^5 cells per 75-cm² flask. After adhesion of the cells, 50 ng/ml MMC was added for 24 h to the cell medium. Then, the medium was changed, and the cells were grown for another 28 h. During the last 8 h, 0.1 µg/ml colchicine (Sigma-Aldrich) was added to the cells. Trypsinized cell pellet was swollen in hypotonic KCl solution and incubated at 37°C for 20 min. Cells were fixed in ethanol/acetic acid (3:1) and spread on glass slides. After air drying, the slides were stained with 4% Giemsa for 10 min. After mounting, metaphase spreads were captured in a blind fashion using bright field microscopy (AX70; Olympus) with a 63× objective lens. Statistical analyses were performed with Prism software (GraphPad Software) using an unpaired Student’s t test.

SAβ-gal staining. SAβ-gal-positive cells were detected as previously described (Dimri et al., 1995). In brief, monolayers of cells were fixed with 4% paraformaldehyde, incubated at 37°C, and stained overnight at 37°C with a staining solution mix at pH 6.0 using a β-gal detection kit (Invitrogen). After overnight incubation, cells were counterstained with propidium iodide before observation under a fluorescence microscope (B51; Olympus) with a 20× objective lens. The number of positive cells with blue color was counted and normalized to the number of total cells in the same field. For each condition, cells from more than five randomly chosen fields were counted (to total at least 100 cells counted per condition). Similar results were obtained in two independent experiments.

Telomere length determination. Telomere length was determined using terminal restriction fragment analysis by Southern blotting of DNA extracted from whole blood from the patient, his siblings, and parents as described previously (Touzot et al., 2010).

WGHM genotyping and WES. In accordance with the Helsinki Declaration, informed consent for the genetic and functional investigations was obtained from the parents and investigated family members. WGHM was determined using a BeadChip (Human Linkage-24; Illumina) according to standard methodology as previously described (Le Guen et al., 2013). WES was performed through 100-bp paired-end reads on Illumina HiSeq after capture using an exome enrichment kit (TruSeq; Illumina) as previously described (Le Guen et al., 2013).

Real-time quantitative RT-PCR. 2 µg RNA from a panel of tissues (Human Total RNA Panel II; Takara Bio Inc.) was treated with DNasel (amplification grade; Invitrogen) before reverse transcription. 80 ng cDNA was used for each Taqman reaction (ViiA7) using sets of primers and probes specifically designed to distinguish ERC6L2-SF and Hebo alternative transcripts. The expression level was normalized with the housekeeping gene PSMB2 (Hs01002946; Thermo Fisher Scientific) sequence of the following primers: ERC6L2-SF forward, 5′-AGGCAGTTCAAGGACATTG-3′; ERC6L2-SF reverse, 5′-TTCACCTCGGTCAAGATG-3′; ERC6L2-SF Taqman MGB Probe, 5′-TGTCCTTTGA GGAACCTTA-3′; Hebo forward, 5′-TGGATCAGGCAA AGCTGAAA-3′; Hebo reverse, 5′-CTGCTTTCAACCT AAATAGCTCATGT-3′; and Hebo Taqman MGB Probe, 5′-CACATGACCGCATGGG-3′.

ERC6L2 long form (Hebo) identification and cloning. Based on human (accession no. ENSG00000182150) and rat (accession no. ENSRNOG00000019175) genome annotation in the Ensembl database for the ERC6L2 gene, primers were designed to amplify a putative transcript linking human transcripts 001, 006, and 007 (Fig. 4) as follows: forward exon 12, 5′-TGGACTAGGCTCAATTTTGTCGTGTC-3′; reverse exon 15, 5′-GGCTCTTCTCATCAGTAACTCA CTGCCAGAG-3′; forward exon 16, 5′-CAACAGAGA CCAAAGATGCTGGTTGGAG-3′; and reverse exon 16, 5′-ACTAGTTGATCTTGCTGCTGCTGTG-3′. The expression of a full-length ERC6L2 long form was further validated by cDNA PCR amplification using the following primers: forward, 5′-TTTGTGGGTCATCCCCCTC TTCATCATCT-3′; and reverse, 5′-GGCAGACAGGGTC CACACGATCATTATGTCG-3′.

ERC6L2-SF and ERC6L2 long form (or HEBO) orf’s were PCR amplified and cloned into a bicistronic (IresGFP) lentiviral vector for complementation experiments and in fusion with GFP into EGF-P-N2 vector for cellular localization and microirradiation experiments. To drive the ERC6L2-SF in the nucleus, the SV40 large T antigen nls PPKKKRKSV sequence was added at the N terminus of ERC6L2-SF.

Laser microirradiation. NBS1-EGFP and DsRed plasmids were provided by C. Lukas (University of Copenhagen, Copenhagen, Denmark). HeLa cells were transfected with ERC6L2-SF-GFP and NBS1-GFP constructs using Jet Prime reagent (Polyplus transfection). 24 h after transfection, cells were presensitized with BrdU (BD) for another 28 h. Microirradiation was performed using a 405-nm UV laser focused through a 63× objective lens in a confocal microscope system (TCS sp8; Leica Biosystems). Laser settings (100% laser power for 10% of laser duration) were used to generate DNA damage specifically along the laser path (defined by a series of dots) in a BrdU-dependent manner. The cells were analyzed in time course mode with a 0.329-s interval between pictures.
for up to 215 pictures. The interval between pictures was then shifted to 30 s for another 15 pictures. Images were exported as LIF files and analyzed using ImageJ.

CRISPR/Cas9 gene inactivation. The 5’-TTGAGAGACTAC CAAAGAGA-3’ located in ERCC6L2 exon 2 was used as gRNA and cloned into the pX330–U6-Chimeric_BB-CBh-hSpCas9 plasmid (a gift from F. Zhang, the Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA; plasmid 42230; Addgene; Ran et al., 2013) for transfection into U2OS cells. The extent of mutagenesis was evaluated by single molecule sequencing using PGM (IonTorrent) after amplification of genomic DNA spanning the CRISPR target site.

Sequence analysis and molecular modeling. Phyre2 was used to select the best templates for the modeling of the 3D structure of the ATPase domain, which was then performed using Modeller (v9.10; Martí-Renom et al., 2000). The analysis of the C-terminal domain was made using hydrophobic cluster analysis (Callebaut et al., 1997; Faure and Callebaut, 2013). Profile–profile search methods, such as HH-PRED (Söding et al., 2005), were also used together with standard sequence similarity searches (Altschul et al., 1997) to highlight putative relationships of the HEBO domain with other protein sequences.

V(DJ) and HR assays. V(DJ) recombination assays were performed as previously described (Buck et al., 2006a). Szs were sequenced using single molecule sequencing on a PGM (IonTorrent). For HR assay, 5 µg I-SceI–expressing plasmid was electroporated in 10^6 U2OS-DRGFP cells (a gift from M. Jasin, Memorial Sloan Kettering Cancer Center, New York, NY). Cells were then retransfected with scramble, Hebo, and Rad51 siRNAs (SMARTpool ON-TARGET plus siRNA; GE Healthcare) according to manufacturer recommendations. The percentage of GFP-positive cells was determined by FACS analysis 72 h after transfection. The efficacy of siRNA was verified by Taqman real-time PCR.

Online supplemental material. Table S1 shows biological characteristics of the patient. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20151183/DC1.

ACKNOWLEDGMENTS

We thank the patient and his parents for participating in this study, the imaging facility of Institut Imagine for help with microlaser irradiation, the Bioinformatic Department of Institut Imagine for help with WES analysis, and the genomic facility of Institut Imagine for single molecule sequencing on the PGM. We thank Dr. Claudia Lukas for the pEGFP-NBS1 plasmid and Dr. Feng Zhang for the pX330 plasmid.

This work was supported by institutional grants from the Institut National de la Santé et de la Recherche Médicale, the Institut National du Cancer (PIBIO11-151), Imagine Institut (ANR-10-1AHU-01), the European Research Council (ERC-2009-AAdG_2000506 n°FP7-249816), and Ligue Contre le Cancer (Equipe Labelisée 1A Legue).

The authors declare no competing financial interests.

REFERENCES


Submitted: 20 July 2015

Accepted: 12 April 2016


