nature genetics

Mutations in *PNKP* cause microcephaly, seizures and defects in DNA repair

Jun Shen[1,12], Edward C Gilmore[2,3,12], Christine A Marshall[1], Mary Haddadin[4,11], John J Reynolds[5], Wafaa Eyaid[6], Adria Bodell[1], Brenda Barry[1], Danielle Gleason[2], Kathryn Allen[1], Vijay S Ganesh[1], Bernard S Chang[1], Arthur Grix[7], R Sean Hill[2], Meral Topcu[8], Keith W Caldecott[5], A James Barkovich[9] & Christopher A Walsh[1,2,10]

Maintenance of DNA integrity is crucial for all cell types, but neurons are particularly sensitive to mutations in DNA repair genes, which lead to both abnormal development and neurodegeneration[1]. We describe a previously unknown autosomal recessive disease characterized by microcephaly, early-onset, intractable seizures and developmental delay (denoted MCSZ). Using genome-wide linkage analysis in consanguineous families, we mapped the disease locus to chromosome 19q13.33 and identified multiple mutations in *PNKP* (polynucleotide kinase 3′-phosphatase) that result in severe neurological disease; in contrast, a splicing mutation is associated with more moderate symptoms. Unexpectedly, although the cells of individuals carrying this mutation are sensitive to radiation and other DNA-damaging agents, no such individual has yet developed cancer or immunodeficiency. Unlike other DNA repair defects that affect humans, *PNKP* mutations universally cause severe seizures. The neurological abnormalities in individuals with MCSZ may reflect a role for PNKP in several DNA repair pathways.

We identified the autosomal recessive disorder MCSZ with the following features: microcephaly, infantile-onset seizures, developmental delay and variable behavioral problems, especially hyperactivity. MCSZ was observed in multiple pedigrees of Middle Eastern and European origin (**Fig. 1**). The first three pedigrees were Arabic Palestinians living in Jordan and the United States. Three other families with similar manifestations were Arabic (Kingdom of Saudi Arabia), Turkish and of mixed European ancestry (United States). We later found less severely affected individuals with MCSZ from family 7, which is also mixed European heritage and from the United States. Brain magnetic resonance imaging scans (MRIs) consistently show microcephaly with preserved brain structures, without apparent neuronal migration or other structural abnormalities, and with no evidence of degeneration (**Fig. 2**). The affected individuals did not develop ataxia or other neurological symptoms. Routine clinical genetic and metabolic screening showed no abnormalities. Despite careful inquiry, individuals with MCSZ were not found to have a higher frequency of common or uncommon infections, offering no clinical evidence of immunodeficiency. Cells from one affected individual showed sensitivity to irradiation in a standard colony-survival assay[2,3]. However, no affected person has developed cancer by age 21, and heterozygous carriers have not developed early-onset cancer or any sign of immunodeficiency (**Supplementary Note**, Clinical Information, **Supplementary Table 4** and Summary).

Genome-wide linkage screens suggested a single common homozygous region on chromosome 19q in all six individuals with MCSZ in the consanguineous families 1–3. This same locus showed homozygosity in the two affected individuals from family 4 and in the single one from family 5 whose parents were not known to be related. We observed no other regions of linkage, including at known primary microcephaly loci. Notably, all six affected individuals in the three Palestinian families (families 1–3) were homozygous for the same 3-cM (1.5-Mb) haplotype (between markers D19S879 and D19S907) at chromosome 19q13.33, suggesting a common ancestor (**Supplementary Fig. 1**). Further investigation revealed a common town of origin for these families. Linkage analysis of families 1–4 and family 6 generated a combined maximum two-point lod score of 5.60 at D19S867 with $\theta = 0$ (**Supplementary Table 1**) and a maximum multipoint lod score of 7.12 (**Supplementary Fig. 2**).

We sequenced 41 genes within the minimal region (**Supplementary Fig. 2** and **Supplementary Table 2**), and only the *PNKP* gene contained mutations. We found homozygous mutations in families 1–5

[1]Howard Hughes Medical Institute, Department of Neurology, Beth Israel Deaconess Medical Center and Program in Neuroscience, [2]Division of Genetics and The Manton Center for Orphan Disease Research, Department of Medicine, Children's Hospital Boston, and [3]Division of Child Neurology, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. [4]Department of Pathology, Cytogenetics Laboratory, Al-Bashir Hospital, Ministry of Health, Amman, Jordan. [5]Genome Damage and Stability Centre, University of Sussex, Falmer, Brighton, UK. [6]Genetics & Endocrinology, Department of Pediatrics, King Fahad National Guard Hospital, King Abdul Aziz Medical City, Saudi Arabia. [7]Department of Medical Genetics, Kaiser-Permanente Point West Medical Offices, Sacramento, California, USA. [8]Hacettepe University, Medical Faculty, Ihsan Dogramaci Children's Hospital, Department of Pediatrics, Section of Pediatric Neurology, Ankara, Turkey. [9]Department of Radiology, Department of Neurology and Department of Pediatrics, University of California at San Francisco, San Francisco, California, USA. [10]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA. [11]Present address: Quest Diagnostics, Nichols Institute, San Juan Capistrano, California, USA. [12]These authors contributed equally to the work. Correspondence should be addressed to C.A.W. (Christopher.Walsh@childrens.harvard.edu).

Received 6 August 2009; accepted 23 December 2009; published online 31 January 2010; doi:10.1038/ng.526

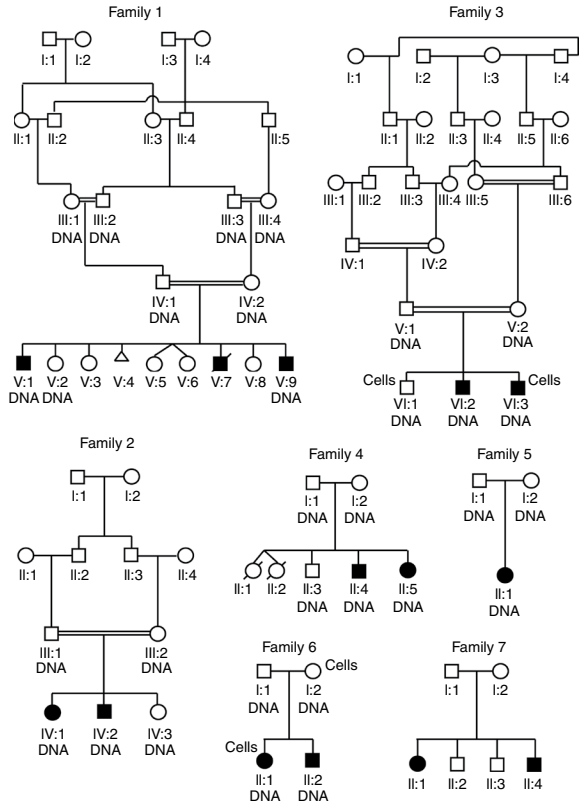

Figure 1 Pedigrees of MCSZ families. Family 1 represents a consanguineous Palestinian pedigree in Jordan. Family 2 shows another consanguineous Palestinian pedigree that is reportedly unrelated to family 1, also in Jordan. Family 3 is also consanguineous and Palestinian, but the family now resides in the United States. Family 4 is from the Kingdom of Saudi Arabia, and the parents were not known to be consanguineous. Family 5 is from Turkey, and the parents were not known to be related. Family 6 is of mixed European descent from the United States (German-Irish). Family 7 is also of mixed European (Swedish, Italian, Irish and English) heritage, from the United States. The individuals from whom samples were obtained are indicated by the label 'DNA'. The individuals from whom we established lymphoid cell lines are indicated by the label 'Cells'. Cells and DNA were available for all members of family 7.

and compound heterozygous mutations in families 6 and 7 (see **Fig. 3** and **Supplementary Fig. 3** for detailed analyses of all mutations). The three Palestinian families (families 1–3) shared a homozygous base pair substitution in exon 11 (G975A), resulting in the nonconservative amino acid change E326K. The families from the Saudi Arabia (family 4) and Turkey (family 5) had the same homozygous 17-bp duplication (17-bp dup) in exon 14 (1250_1266dup), resulting in a frameshift, T424GfsX48. Family 6 (European) had two heterozygous mutations: the same 17-bp duplication in exon 14 as the Saudi and Turkish families and a point mutation in exon 5 (C526T) that results in L176F. The moderately affected members of family 7 (European) showed compound heterozygosity, carrying the 17-bp duplication mutation in exon 14 and a 17-bp deletion in intron 15 that disrupts proper mRNA splicing.

To confirm that these mutations were pathogenic, we analyzed Epstein-Barr virus (EBV)-transformed lymphocytes derived from affected individuals from families 3 (E326K substitution) and 7 (17-bp duplication and intron 15 deletion). The samples from the severely affected individual (family 3) and mildly affected individuals (family 7) had much lower concentrations of PNKP protein than did those from unaffected or heterozygous family members (**Fig. 3b**). RT-PCR analysis showed that the intron 15 deletion in family 7 disrupts mRNA splicing and causes skipping of exon 15; a barely detectable level of properly spliced mRNA remains (**Fig. 3c**). The low concentrations of PNKP protein and the indistinguishable phenotype among severely affected individuals from families 1–6 suggest that all of these mutations impair PNKP severely or completely; in contrast, the individuals from family 7, who have a slightly milder phenotype and carry one noncoding mutation, may retain some PNKP activity.

Unexpectedly, despite their diverse ancestries (Saudi, Turkish and mixed European), all chromosomes with the 17-bp duplication mutation seem to share the same haplotype for 18 SNPs in or near the *PNKP* locus (**Supplementary Fig. 4a**). The low frequency of this mutated haplotype (**Supplementary Fig. 4b**) suggests that the 17-bp

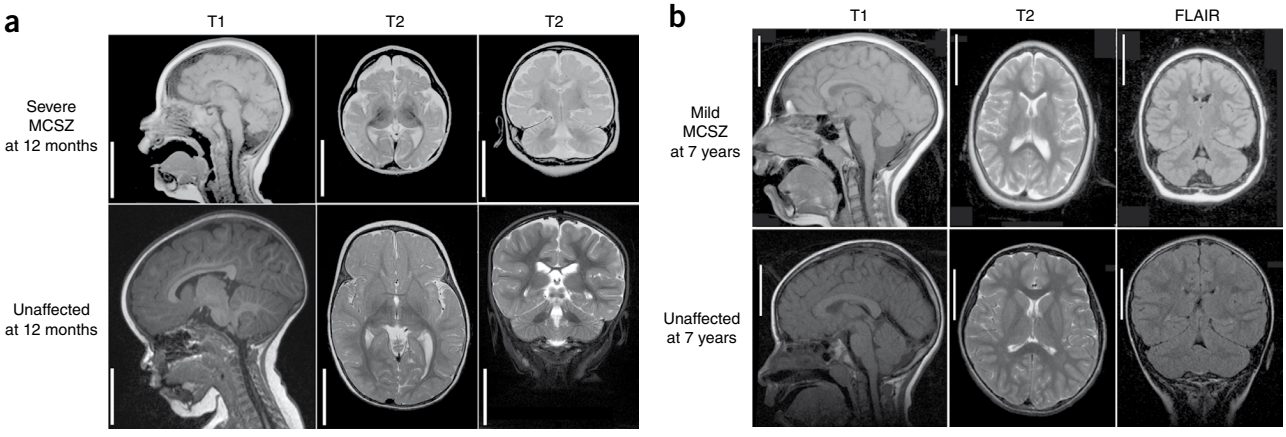

Figure 2 Brain MRIs of individuals with MCSZ. Representative MRI images are shown from families 4 (**a**; severely affected) and 7 (**b**; moderately affected) with aged-matched controls. MRIs of severely affected individuals from other families were similar to the representative images in **a**. Sagittal images are shown on the left (T1), axial images in the middle (T2) and coronal images on the right (T2 (**a**) and FLAIR (**b**)), with the MRI sequence noted above the image. The MRIs show that, despite the microencephaly (small brain), the gyral pattern is not clearly abnormal, indicating an absence of visible neuronal migration abnormality. The cerebellum is proportionately small compared to the cerebrum, and the subpallium (basal ganglia or ventral cerebrum) is proportionate to the pallium (dorsal cerebrum). There is no evidence of atrophy or glial scarring. Scale bar indicates 5 cm for both unaffected and MCSZ images.

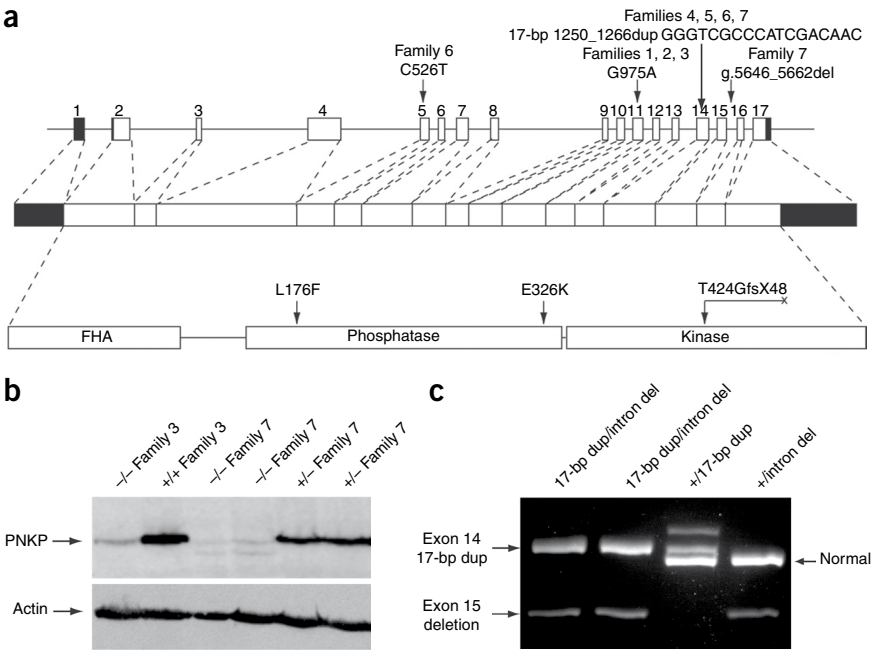

Figure 3 *PNKP* mutations in individuals with MCSZ. (**a**) The diagram shows four different mutations identified in human *PNKP* genomic DNA, mRNA and protein, including protein domains (forkhead is indicated by 'FHA'). The human *PNKP* gene consists of 17 exons (boxes) and encodes a peptide of 521 amino acids. Filled boxes represent untranslated regions and open boxes represent coding regions. Lines connecting the exons represent introns. (**b**) Protein blot for PNKP. Lanes 1 and 2 show samples from an individual with MCSZ (VI:3 with the E326K mutation) and an unaffected brother (VI:1), respectively, from family 3. Lanes 3 and 4 represent individuals with MCSZ (II:1 and II:4, with both a 17-bp duplication (dup) and a 17-bp intron 15 deletion (del)) from family 7. In contrast, lanes 5 and 6 represent the father (I:1) and brother (II:3) from this family, who are both heterozygous for the 17-bp intron 15 del. The band indicates a molecular weight of ~60 kDa (predicted size 57 kDa). Anti–β-actin is a loading control. (**c**) RT-PCR products of mRNA from members of family 7 show the expected size from the normal copy of *PNKP* cDNA (636 bp), seen in lanes 3 and 4 from unaffected carriers. The 17-bp dup results in a 653-bp fragment, seen in lanes 1, 2 and 3. The band in lanes 1, 2 and 4 corresponding to a size of 548 bp is found in samples with the intron 15 deletion lacking exon 15 (determined from sequencing; data not shown). A small amount of normal-sized transcript is seen in lanes 1 and 2 with higher exposure (data not shown), indicating that a small amount of normal PNKP mRNA can be produced.

duplication in exon 14 is probably of shared origin and very old in all or most of these families. Nonetheless, the mutation shows an extremely low carrier frequency in the general population, as the 17-bp duplication mutation was absent in 1,080 Middle Eastern or European-descended control chromosomes, and all other mutations were absent in all 280 control chromosomes screened.

The PNKP protein has been implicated in repair of both double-strand breaks (DSBs) and single-strand breaks (SSBs)[4,5], because its phosphatase domain removes 3′-phosphates and its kinase domain phosphorylates 5′-hydroxyl groups, steps that are required for DNA ligation[6]. PNKP's forkhead domain mediates interaction with the non-homologous end-joining (NHEJ) complex via XRCC4 or with the SSB and base-excision repair (BER) pathways via XRCC1 (**Supplementary Fig. 5**)[4–8]. *PNKP* has been further implicated in the repair pathway disrupted in an ataxic neurodegenerative disease, spinocerebellar ataxia with axonal neuropathy (denoted SCAN1 and involving a

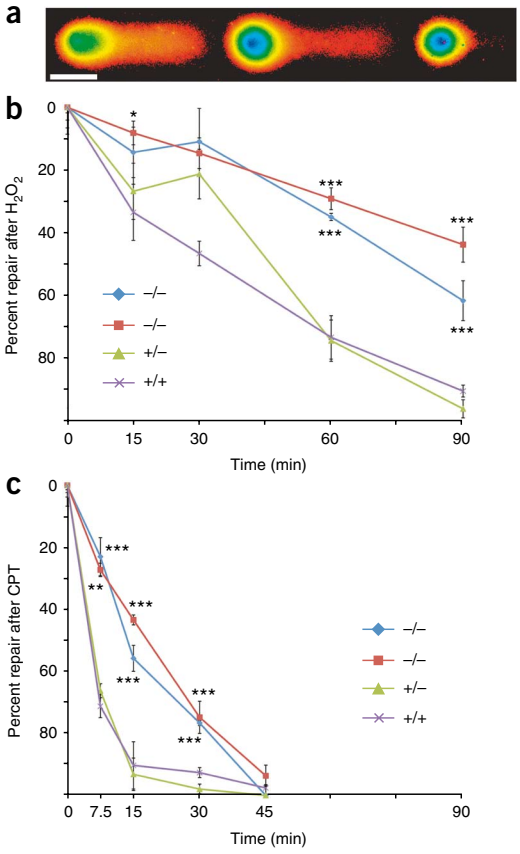

Figure 4 Lymphocytes from individuals with MCSZ show abnormal DNA repair. (**a**) Examples of comet assay results. The intensity of the fluorescence is represented in pseudocolor, as the electrical field drives damaged, loose DNA from left to right. The image at left shows a cell with 50% tail DNA with the body of the nucleus (green) on the left and the tail of the comet derived from the damaged DNA extending to the right. The images middle and right show progressively less damage, with 29% and 11% tail DNA, respectively. (**b**) After hydrogen peroxide treatment with 0 min for recovery, cells show their maximum damage. Cells derived from individuals with MCSZ (blue and red) show significant impairment in their ability to repair DNA after hydrogen peroxide was removed, whereas cells derived from unaffected family members were able to repair DNA much more efficiently. (**c**) After camptothecin (CPT) treatment, there was also statistically significantly slower repair in cells derived from individuals with MCSZ compared with those from unaffected family members (green and purple). However, after 45 min, the MCSZ-derived cells were able to repair all CPT damage, in contrast to hydrogen peroxide–treated cells. All cells were derived from family 7. Blue diamonds and red squares indicate cells from individuals with MCSZ carrying the exon 14 17-bp duplication and the intron 15 17-bp deletion (II:1 and II:4, respectively); green triangles represent an unaffected parent who is heterozygous for the intron 15 17-bp deletion (I:1); purple crosses represent an unaffected sibling with no mutation (II:2). $*P = 0.05$; $**P < 0.005$; $***P < 0.0005$. Scale bar, 50 μm.

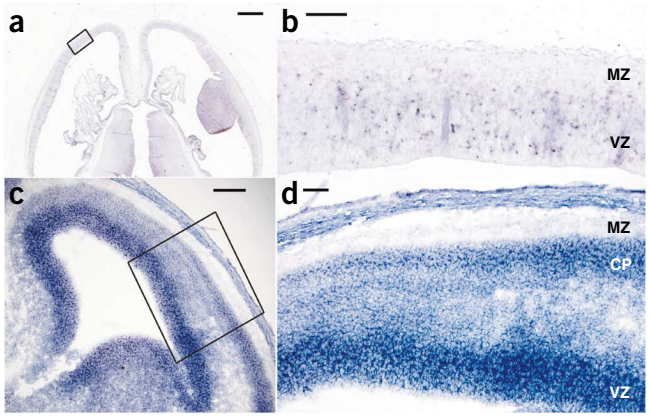

Figure 5 *PNKP in situ* hybridization. (**a**,**b**) *In situ* hybridization of Carnegie Stage 22 human embryos (~54 postovulatory days) with an antisense probe to human *PNKP* (**a**). The sense strand (not shown) showed no specific hybridization. A higher-magnification image of the boxed area in the developing cerebral cortex is shown (**b**). The ventricular zone (VZ), containing proliferating cells, shows PNKP mRNA expression, whereas the cell-sparse marginal zone (MZ) has no staining. (**c**,**d**) Mouse embryonic day 14 cerebral cortex (**c**) (with high-magnification view of the boxed region (**d**)) shows a similar staining pattern, with high expression within the proliferating VZ and lower but maintained expression within differentiated neurons of the cortical plate (CP). **a** and **b** are in the transverse plane, and **c** and **d** are coronal. Scale bars indicate 1 mm (**a**), 100 μm (**b**), 150 μm (**c**) and 75 μm (**d**).

mutation in *TDP1*) (see **Supplementary Fig. 5** for additional details). The abnormalities in a clinical irradiation–sensitivity test already reflect deficiencies in NHEJ, which is required to repair radiation-induced DSBs. Because PNKP potentially has roles in additional repair pathways, we tested cells in response to other DNA-damaging agents.

We examined the response of cells from individuals with MCSZ to free radical damage from hydrogen peroxide, which predominately requires BER, and camptothecin, which requires TDP1 via its activity on topoisomerase I, using an alkaline comet assay that detects both SSBs and DSBs by quantifying the amount of DNA that moves from a nucleus after electrophoresis. Typical nuclei with various degrees of DNA damage are shown in **Figure 4a**. We determined the relative amount of DNA damage by measuring the 'Percent tail DNA' (DNA that has moved out of the nucleus). The relative amount of DNA damage repair was quantified as the percent tail DNA after toxin exposure at various times of recovery (see **Supplementary Table 2** for all data) and normalized to the maximum and minimum damage for each cell line in each experiment. MCSZ-derived cell lines were significantly impaired in their ability to repair hydrogen peroxide–induced damage (**Fig. 4b**) and were also delayed in repairing camptothecin-induced damage (**Fig. 4c**). MCSZ-derived cells were eventually able to repair the camptothecin-induced damage after 45 min, but they could not repair all hydrogen peroxide–induced damage even after 90 min. These data suggest that camptothecin-induced damage may be easier to repair than free radical damage in the setting of *PNKP* mutations. Although individuals with MCSZ do not develop the ataxia that characteristically results from *TDP1* mutations that also cause camptothecin sensitivity[9,10], it may be that those we studied are too young (oldest 21 years) to show this characteristic, or there may be some mechanism compensating for the loss of PNKP.

Microcephaly can result from failure to produce enough neurons during development (primary microcephaly) or from degeneration after normal development. Individuals with MCSZ have no evidence of brain atrophy or clinical regression, making degeneration unlikely (see the **Supplementary Note**, Clinical Information, for further details). *In situ* hybridization indicated that human and mouse *PNKP* mRNAs are expressed in both dividing neuronal precursors in the cerebral cortical ventricular zone (VZ, **Fig. 5a–c**) and in postmitotic neurons of the cortical plate (CP, **Fig. 5d**), which is consistent with potential roles in both dividing and postmitotic neurons. In addition, we found that when we used RNAi to reduce levels of Pnkp in dissociated mouse neurons *in vitro*, there was a small but statistically significant increase in apoptosis in both neuronal precursors and differentiated neurons (**Supplementary Figs. 6** and **7** and **Supplementary Table 3**) compared to cells transfected with control plasmids. This indicates that microcephaly could result from apoptosis of precursors, differentiated neurons or both cell types.

Disruption of genes encoding NHEJ repair proteins can lead to microcephaly in humans and/or mice, as occurs in LIG4 syndrome, which involves mutations in *Lig4* (refs. 11–13), in severe combined immunodeficiency with microcephaly, or as a result of mutations in *NHEJ1* (encoding Cernunnos)[14], *Xrcc4* (ref. 15), *Xrcc6* (encoding Ku70) and *Xrcc5* (encoding Ku80)[16] and *Prkdc* (encoding DNA-PKcs)[17]. However, disruption of NHEJ function is not known to cause seizures, which are a prominent clinical feature of MCSZ. In contrast, disruption of the BER and SSB pathways by deletion of *Xrcc1* in mice produced normal brain size, but the mutant mice developed ataxia, loss of interneurons within the cerebellum and seizure-like behavior[18]. The intriguing similarity of phenotypes between humans with MCSZ who carry mutations in *PNKP* and mice with targeted deletion of *Xrcc1* does suggest a potential common mechanism, demonstrating the requirement for the BER and SSB pathways in the prevention of seizures, potentially via an interneuron-specific requirement of the BER and SSB pathway. Therefore, the unique pattern of neurological symptoms of MCSZ may reflect a requirement for PNKP activity in multiple DNA repair pathways.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

J.S. was supported by the Victoria and Stuart Quan Fellowship. Research was supported by grants from the US National Institute of Neurological Disorders and Stroke to E.C.G. (5K08NS059673-02) and to C.A.W. (NSR01-35129), the Fogarty International Center (R21 NS061772), the Manton Center for Orphan Disease Research, the Simons Foundation and by Dubai Harvard Foundation for Medical Research. E.C.G. was also supported by a K12 Child Health Research Center award to Children's Hospital Boston (1 K12 HD052896-01A1) and a Research in Training Award from the Child Neurology Foundation. J.J.R. and K.W.C. are supported by the UK Medical Research Council (G0600776 & G0400959). C.A.W. is an Investigator of the Howard Hughes Medical Institute. We thank S. Lindsay and S. Lisgo from the Medical Research Council–Wellcome Trust Human Developmental Biology Resource (HDBR), Institute of Human Genetics, International Centre for Life, Newcastle upon Tyne, UK, for performing human *in situ* hybridizations. We are grateful for genotyping services provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the US National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C and NIH N01-HG-65403. Genotyping at Children's Hospital Boston is supported by the Intellectual and Developmental Disabilities Research Centers (CHB MRDDRC, P30 HD18655).

AUTHOR CONTRIBUTIONS

J.S. helped to characterize MCSZ syndrome, identified the MCSZ locus and calculated lod scores, sequenced genes in the MCSZ locus to identify *PNKP* mutations and wrote the manuscript; E.C.G. helped to characterize MCSZ syndrome,

identified the moderately affected MCSZ family, performed RT-PCR on moderately affected family samples, performed comet assays, organized and analyzed Sequenom experiments, did analysis of PNKP mutation, performed mouse RNAi experiments, helped perform mouse *in situ* hybridizations and wrote the manuscript; C.A.M. sequenced genes in the MCSZ locus to identify *PNKP* mutations and helped perform human *in situ* hybridizations; M.H. identified affected patients and provided clinical information; J.J.R. performed PNKP protein blots and confirmatory comet assays; W.E. identified affected patients and provided clinical information; A.B. organized clinical information and patient samples; B.B. organized clinical information and patient samples; D.G. organized patient samples and helped perform Sequenom experiments; K.A. organized patient samples and helped perform sequencing experiments; V.S.G. helped analyze Sequenom experiments; B.S.C. helped organize clinical information to identify MCSZ syndrome; A.G. identified affected patients and provided clinical information; R.S.H. helped organize genetic data and calculate lod scores; M.T. identified affected patients and provided clinical information; K.W.C. advised on comet assays, supervised PNKP protein blotting and edited the manuscript; A.J.B. characterized MRIs for patient classification; C.A.W. directed the overall research and wrote the manuscript.

The genetic study was approved by Beth Israel Deaconess Medical Center and Children's Hospital Boston Institutional Review Boards. Appropriate informed consent was obtained from all involved human subjects. All animal work was approved by Harvard Medical School, Beth Israel Deaconess Medical Center and Children's Hospital Boston Institutional Animal Care and Use Committees.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. McKinnon, P.J. DNA repair deficiency and neurological disease. *Nat. Rev. Neurosci.* **10**, 100–112 (2009).
2. Huo, Y.K. *et al.* Radiosensitivity of ataxia-telangiectasia, X-linked agammaglobulinemia, and related syndromes using a modified colony survival assay. *Cancer Res.* **54**, 2544–2547 (1994).
3. Sun, X. *et al.* Early diagnosis of ataxia-telangiectasia using radiosensitivity testing. *J. Pediatr.* **140**, 724–731 (2002).
4. Jilani, A. *et al.* Molecular cloning of the human gene, *PNKP*, encoding a polynucleotide kinase 3′-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *J. Biol. Chem.* **274**, 24176–24186 (1999).
5. Karimi-Busheri, F. *et al.* Molecular characterization of a human DNA kinase. *J. Biol. Chem.* **274**, 24187–24194 (1999).
6. Chappell, C., Hanakahi, L.A., Karimi-Busheri, F., Weinfeld, M. & West, S.C. Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. *EMBO J.* **21**, 2827–2832 (2002).
7. Whitehouse, C.J. *et al.* XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. *Cell* **104**, 107–117 (2001).
8. Ali, A.A., Jukes, R.M., Pearl, L.H. & Oliver, A.W. Specific recognition of a multiply phosphorylated motif in the DNA repair scaffold XRCC1 by the FHA domain of human PNK. *Nucleic Acids Res.* **37**, 1701–1712 (2009).
9. Takashima, H. *et al.* Mutation of TDP1, encoding a topoisomerase I–dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. *Nat. Genet.* **32**, 267–272 (2002).
10. El-Khamisy, S.F. *et al.* Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature* **434**, 108–113 (2005).
11. Frank, K.M. *et al.* Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* **396**, 173–177 (1998).
12. Barnes, D.E., Stamp, G., Rosewell, I., Denzel, A. & Lindahl, T. Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr. Biol.* **8**, 1395–1398 (1998).
13. O'Driscoll, M. *et al.* DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. *Mol. Cell* **8**, 1175–1185 (2001).
14. Buck, D. *et al.* Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* **124**, 287–299 (2006).
15. Gao, Y. *et al.* A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**, 891–902 (1998).
16. Gu, Y. *et al.* Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc. Natl. Acad. Sci. USA* **97**, 2668–2673 (2000).
17. Vemuri, M.C., Schiller, E. & Naegele, J.R. Elevated DNA double strand breaks and apoptosis in the CNS of *scid* mutant mice. *Cell Death Differ.* **8**, 245–255 (2001).
18. Lee, Y. *et al.* The genesis of cerebellar interneurons and the prevention of neural DNA damage require XRCC1. *Nat. Neurosci.* **12**, 973–980 (2009).

ONLINE METHODS

Genetic screening. Family 1 underwent a genome-wide linkage screen using about 400 microsatellite markers in the ABI linkage mapping set MD v2.5 at an average density of 10 cM (Applied Biosystems) at the Children's Hospital in Boston (Genotyping Core facility). We performed genome-wide screens for families 3 and 4 at the Center for Inherited Disease Research using microsatellite markers that were also at an average spacing of 10 cM. We carried out fine mapping using polymorphic microsatellite markers from the ABI linkage mapping set HD v2.5 at an average density of 5 cM (Applied Biosystems) along with additional microsatellite markers identified using the UCSC Human Genome Browser[19] and synthesized primers (Sigma-Genosys). Two point and multipoint lod scores calculated using Allegro[20] assumed recessive inheritance, full penetrance and a disease allele frequency of 0.0001. All nucleotide numbers are in reference to cDNA, in which A (of the ATG start site) is +1, except for the intronic deletion, which is in reference to the genomic sequence, in which A of ATG of the translational start site is +1 (all from UCSC genome browser, NCBI Build 36.1).

***PNKP* protein blotting.** For *PNKP*, we grew EBV-transformed lymphocytes in DMEM with 15% (vol/vol) FCS plus normocin. We lysed 2×10^5 cells in SDS loading buffer at 90 °C and fractionated them by SDS-PAGE. We transferred the proteins to Hybond-C Extra Nitrocellulose (GE Healthcare), stained them with Ponceau S solution (Sigma) and washed them in TBST (25 mM Tris base, 150 mM NaCl, 2.7 mM KCl, 0.1% Tween-20) before carrying out immunoblotting with anti-PNK (SK3195)[21] at a dilution of 1:1,000 in 5% (wt/vol) nonfat dried milk. We then washed the blots and probed them with goat anti-rabbit IgG–horseradish peroxidase (HRP) secondary antibody (DakoCytomation), which we detected with ECL detection reagents (GE Healthcare). The blot was then washed in TBST and re-probed with monoclonal anti–β-actin (Sigma) using rabbit anti-mouse IgG HRP (DakoCytomation) as a secondary antibody.

***PNKP* reverse transcriptase polymerase chain reaction.** We grew EBV-transformed lymphocytes as described above. We isolated the RNA using the RNeasy Mini Kit (Qiagen). We used 5 μg of total RNA for first-strand synthesis with oligo(dT) primers using the SuperScript III First-Strand Synthesis SuperMix (Invitrogen). We then used 1 μl of the product of the reverse transcriptase reaction was used in PCR with primers from Exon 10 to Exon 17.

Comet assays. We carried out alkaline comet assays using the protocol from CometAssay ES unit (Trevigen). Briefly, we grew affected individual–derived EBV-transformed lymphocytes as described above. Cells were obtained before treatment or exposed to 100 μM hydrogen peroxide (Sigma) or 10 μM camptothecin (Sigma) for 30 and 60 min, respectively, at 37 °C. After exposure, we immediately collected the cells for the 0 min recovery time point or washed them once, resuspended them in growth medium and incubated at 37 °C. Cells were subsequently at the times indicated. We then embedded the cells in low-melt agarose, plated them on microscope slides, lysed them, treated them with alkaline solution and then electrophoresed the slides in alkaline solution at 1 V cm^{-1} (21 V) with ~300 mA for 30 min. We washed and dried the slides were washed and stained the DNA with SYBR green. Images of nuclei and tails were taken with a Nikon TE2000-E fluorescent microscope with CCD camera, and we determined the percent tail DNA with CometScore 1.5 software (TriTek).

***In situ* hybridization.** We performed human *in situ* hybridizations at the Human Developmental Biology Resource (Institute of Human Genetics, International Centre for Life, UK) with probes to antisense human *PNKP* cDNA (nucleotides +488–1500). We carried out mouse *in situ* hybridizations as described[22], with an antisense probe derived from mouse *Pnkp* cDNA (+386–1566).

We calculated statistics using Microsoft Excel unless otherwise described. The 'Percent DNA repair' values in **Figure 4** were determined by measuring the mean of the percent tail DNA in three separate wells (raw data are presented in **Supplementary Fig. 5**), using the combined percent tail DNA at 0 min recovery (maximum damage) and the minimum damage measured as full recovery for each individual cell line. We determined *P* values by comparing the baseline adjusted DNA repair level from three separate wells using a *t*-test that "compared" each MCSZ-deficient line to two normal control cell lines with a two-tailed distribution and homoscedastic test.

***PNKP* mutation analysis.** We performed PNKP protein alignments using MegAlign (Lasergene). We examined the three-dimensional structure of mouse Pnkp in MacPyMOL from previously published information[23]. The Branch Point analysis was done with Human Splicing Finder version 2.3 (http://www.umd.be/HSF/)

Determination of the 17-bp duplication haplotype in exon 14. We determined Sequenom SNP genotypes in samples from affected and control individuals at the Molecular Genetics Core Facility at Children's Hospital Boston. We determined haplotypes were determined with Mendelian inheritance patterns for parents and offspring and with Phase haplotype-determining software[24–26] for control samples.

PNKP RNA interference studies. pSilencer 1.0 GFP had RNAi oligos targeted to mouse *Pnkp* mRNA identified via the Broad Institute RNAi Consortium shRNA Library, ligated into EcoRI and ApaI sites (Mo-Pnkp RNAi primers, **Supplementary Table 3**; Invitrogen). We dissected the cerebral cortices from embryonic day 13.5 Swiss-Webster mice and dissociated them using the Papain Dissociation System (Worthington Biochemical). We washed the cells twice in HBSS buffer (HyClone, Thermo Scientific) and transfected them using the Amaxa Nucleofactor 96 well shuttle system (Amaxa Biosystems) with either pSilencer mouse *Pnkp*-RNAi GFP or pSilencer GFP with or without pSport-human *PNKP* (Open Biosystems, Entrez clone ID BC057659). We grew the cells on poly-L-ornithine–treated (Sigma-Aldrich) Lab-Tek chamber slides with Permox (Nalge Nunc International) in DMEM with 10% (vol/vol) FBS for the Pax6 studies. For the NeuN studies, we changed the medium of the cultures to Neurobasal complete medium after 4 h. We fixed the cultures with 4% (vol/vol) paraformaldehyde 24 h after transfection. We carried out immunostaining in PBS with 0.2% (wt/vol) Triton X-100 and 10% (vol/vol) goat serum with chick anti-GFP (Abcam), rabbit anti-cleaved caspase 3 (Asp175) (Cell Signaling Technology), mouse anti-Pax6 (The Developmental Studies Hybridoma Bank) or mouse anti-NeuN (Chemicon). We detected the primary antibodies with anti-chicken IgG Alexa 488 (Invitrogen) and donkey anti-mouse IgG Cy3 or donkey anti-rabbit IgG Cy5 (Jackson ImmunoResearch). Images were obtained by looking for GFP-positive cells with a Nikon TE2000-E fluorescent microscope and Metamorph imaging software. We counted the cells from digital images with the counter blinded as to whether the cultures were derived from *Pnkp*-RNAi or control transfections. We calculated the *P* values in **Supplementary Figure 7** using a *G*-test of goodness of fit with 5 degrees of freedom, a two-tailed *P* value and no Williams correction[27].

Mouse *Pnkp* RNA interference testing. We tested the mouse (Mo) *Pnkp* RNAi construct Mo-*Pnkp*-DsRed or the human construct Hu-*PNKP*-DsRed, both expressed by the pCAG-DsRed vector[28]. We transfected the RNAi control vector (empty vector)/Mo-*Pnkp*-DsRed fusion protein, Mo-*Pnkp*-RNAi/Mo-*Pnkp*-DsRed, RNAi-control/Hu-*PNKP*-DsRed fusion and Mo-*Pnkp*-RNAi/Hu-*PNKP*-DsRed into NIH-3T3 cells with Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. We counted the cells after 24 h in culture. Random fields were chosen by the level of GFP expression, and the number of cells with high GFP expression was counted, followed by the number of cells expressing the Ds-Red fusion protein. DsRed expression was nuclear, as expected from PNKP localization (data not shown). For protein blot analysis, we transfected NIH-3T3 cells as described above. After 24 h of culture, we lysed the cells using the protocol described above for PNKP protein blotting. After quantifying the protein using the BCA assay, we separated 25 μg of protein per lane by PAGE. We transferred the protein to a membrane, blocked it with Odyssey blocking buffer (Li-Cor Biosciences) and probed with mouse anti-actin (AC-15), rabbit anti-GFP (Molecular Probes) and rabbit anti-DsRed (Clontech). Primary antibody was detected with anti-mouse labeled with IRDye 700 or anti-rabbit IRDye 800 (Li-Cor Biosciences) and detected with Li-Cor imaging system. In the bar graph, protein amounts are normalized to RNAi control conditions.

doi:10.1038/ng.526

19. Karolchik, D. *et al*. The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.* **36**, D773–D779 (2008).
20. Gudbjartsson, D.F., Jonasson, K., Frigge, M.L. & Kong, A. Allegro, a new computer program for multipoint linkage analysis. *Nat. Genet.* **25**, 12–13 (2000).
21. Breslin, C. & Caldecott, K.W. DNA 3′-phosphatase activity is critical for rapid global rates of single-strand break repair following oxidative stress. *Mol. Cell. Biol.* **29**, 4653–4662 (2009).
22. Berger, U.V. & Hediger, M.A. Differential distribution of the glutamate transporters GLT-1 and GLAST in tanycytes of the third ventricle. *J. Comp. Neurol.* **433**, 101–114 (2001).
23. Bernstein, N.K. *et al.* The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Mol. Cell* **17**, 657–670 (2005).
24. Stephens, M., Smith, N.J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989 (2001).
25. Stephens, M. & Donnelly, P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* **73**, 1162–1169 (2003).
26. Stephens, M. & Scheet, P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* **76**, 449–462 (2005).
27. McDonald, J.H. *Handbook of Biological Statistics* (Sparky House, Baltimore, 2008).
28. Matsuda, T. & Cepko, C.L. Electroporation and RNA interference in the rodent retina *in vivo* and *in vitro*. *Proc. Natl. Acad. Sci. USA* **101**, 16–22 (2004).