Table 2. Summary of Phenotypes of CS Cells

	Individual		
	CS20LO	**CS1USAU**	**XPCS1CD**
RRS	–	–	–
UDS	–	–	–
Virus complementation with *ERCC1* wild-type cDNA (RRS)[a]	+	–	–
Virus complementation with *ERCC4* wild-type cDNA (RRS)[a]	–	+	+
XPF expression	–	+/–	+/–
ERCC1 expression	–	+/–	+/–
Protein alteration	ERCC1 p.Phe231Leu	XPF p.Cys236Arg and p.Tyr577*	XPF p.Cys236Arg and p.Arg589Trp
ERCC1-XPF IP	+	+	+ (p.Cys236Arg)
p89 IP	+	+/–	+/– (p.Cys236Arg)
Virus complementation with *ERCC1* c.693C>G (p.Phe231Leu) cDNA (RRS)[a]	+	ND	ND
Virus complementation with *ERCC4* c.706T>C (p.Cys236Arg) cDNA (RRS)[a]	ND	–	ND
ERCC1-XPF nuclease activity	ND	– (p.Cys236Arg)	– (p.Cys236Arg)
Mutant mRNA level	V low	+/–	ND

Abbreviations are as follows: +, normal; –, defective or absent; +/–, intermediate; IP, immunoprecipitation; ND, not done.
[a]Effect of the indicated cDNA on the RRS response.

the deletion of exon 3) had severe skin symptoms and neurodegeneration.[19] The reason why a partially functional allele (c.1765C>T [p.Arg589Trp]) in XPCS1CD results in a more severe phenotype than the null allele in CS1USAU is currently unclear.

Mice homozygous for a frameshift mutation in *Ercc4* (*Xpf*) die by 3 weeks of age.[31] Likewise, mice in which *Ercc1* has been deleted die soon after birth.[32,33] They develop progressive neurodegeneration, leucopenia and thrombocytopenia,[22] and impaired liver function.[32] Many of these characteristics resemble those of XPCS1CD. Pancytopenia is a hallmark of FA, a disorder associated with cellular hypersensitivity to DNA-ICL-inducing agents.[34] Recent data have indicated that endogenous aldehydes generate DNA damage that requires the FA pathway for resolution.[35] We therefore favor the explanation that pancytopenia arises from a defect in the response to such crosslinkers; this defect is common to the FA individuals and XPCS1CD (Figure 1G). Support for this idea comes from Bogliolo et al.,[36] who report on two FA individuals with mutations in *ERCC4* but without features of CS or XP.

Finally, our results demonstrate that *ERCC4* and *ERCC1* must be added to the list of genes in which defects can result in either CS or the combined XP-CS-FA phenotype.

Acknowledgments

This work was supported by KAKENHI Grants-in-Aid for Young Scientists A (24681008), Exploratory Research (24659533) from the Japan Society for the Promotion of Science, a science research grant from Inamori Foundation, a medical research grant from Mochida Memorial Founds for Medical and Pharmaceutical Research, a medical research grant from the Daiichi-Sankyo Foundation of Life Science, a grant for basic science research from The Sumitomo Foundation, and a medical research grant from Takeda Science Foundation to T.O.; Special Coordination Funds for Promoting Science and Technology from the Japan Science and Technology Agency to Y.N.; and a grant from the Associazione Italiana per la Ricerca sul Cancro (AIRC) to M.S. We are grateful to the National Commissioning Group of the UK National Health Service (NHS) for funding the xeroderma pigmentosum service. In addition, the authors acknowledge financial support from the UK Department of Health via the National Institute for Health Research comprehensive Biomedical Research Centre award to Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust. Some early experiments were carried out by Di J. Sun.

pathogenic mutant (c.706T>C) allele (XPF-p.Cys236Arg and XPF-Rv1, middle panel), and both alleles at once (XPF-com1 and XPF-Rv1, right panel).
(C) Selective quantitative amplification of the wild-type and the c.1730_1731insA (p.Tyr577*) *ERCC4* alleles. Allele-specific primers selectively amplified the wild-type (c.1731C) allele (XPF-WT2 and XPF-Rv2, left panel), the CS pathogenic mutant (c.1730_1731insA) allele (XPF-p.Tyr577* and XPF-Rv2, middle panel), and both alleles at once (XPF-com2 and XPF-Rv2). Locations of the primer sets used for the qRT-PCR experiments are depicted on the right-hand side. Transcripts from *HPRT1* were used as a quantification control. Primers used for the qPCR are listed in Table 1.

We are grateful to the families of the affected individuals for their helpful cooperation with these studies.

Received: November 20, 2012
Revised: March 14, 2013
Accepted: April 9, 2013
Published: April 25, 2013

Web Resources

The URLs for data presented herein are as follows:

GeneReviews, Laugel, V. (1993). Cockayne Syndrome, http://www.ncbi.nlm.nih.gov/books/NBK1342/
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

References

1. Kleijer, W.J., Laugel, V., Berneburg, M., Nardo, T., Fawcett, H., Gratchev, A., Jaspers, N.G., Sarasin, A., Stefanini, M., and Lehmann, A.R. (2008). Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. DNA Repair (Amst.) *7*, 744–750.
2. Laugel, V., Dalloz, C., Tobias, E.S., Tolmie, J.L., Martin-Coignard, D., Drouin-Garraud, V., Valayannopoulos, V., Sarasin, A., and Dollfus, H. (2008). Cerebro-oculo-facio-skeletal syndrome: three additional cases with CSB mutations, new diagnostic criteria and an approach to investigation. J. Med. Genet. *45*, 564–571.
3. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., and Ellenberger, T. (2006). DNA Repair and Mutagenesis, Second Edition (Washington, DC: ASM Press).
4. Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. Nat. Rev. Mol. Cell Biol. *9*, 958–970.
5. Laugel, V., Dalloz, C., Durand, M., Sauvanaud, F., Kristensen, U., Vincent, M.C., Pasquier, L., Odent, S., Cormier-Daire, V., Gener, B., et al. (2010). Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. Hum. Mutat. *31*, 113–126.
6. Nakazawa, Y., Sasaki, K., Mitsutake, N., Matsuse, M., Shimada, M., Nardo, T., Takahashi, Y., Ohyama, K., Ito, K., Mishima, H., et al. (2012). Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair. Nat. Genet. *44*, 586–592.
7. Zhang, X., Horibata, K., Saijo, M., Ishigami, C., Ukai, A., Kanno, S., Tahara, H., Neilan, E.G., Honma, M., Nohmi, T., et al. (2012). Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. Nat. Genet. *44*, 593–597.
8. Schwertman, P., Lagarou, A., Dekkers, D.H., Raams, A., van der Hoek, A.C., Laffeber, C., Hoeijmakers, J.H., Demmers, J.A., Fousteri, M., Vermeulen, W., and Marteijn, J.A. (2012). UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. Nat. Genet. *44*, 598–602.
9. Limsirichaikul, S., Niimi, A., Fawcett, H., Lehmann, A., Yamashita, S., and Ogi, T. (2009). A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). Nucleic Acids Res. *37*, e31.
10. Nakazawa, Y., Yamashita, S., Lehmann, A.R., and Ogi, T. (2010). A semi-automated non-radioactive system for measuring recovery of RNA synthesis and unscheduled DNA synthesis using ethynyluracil derivatives. DNA Repair (Amst.) *9*, 506–516.
11. Nance, M.A., and Berry, S.A. (1992). Cockayne syndrome: review of 140 cases. Am. J. Med. Genet. *42*, 68–84.
12. Lehmann, A.R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie *85*, 1101–1111.
13. Hoy, C.A., Thompson, L.H., Mooney, C.L., and Salazar, E.P. (1985). Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. Cancer Res. *45*, 1737–1743.
14. Wood, R.D. (2010). Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ. Mol. Mutagen. *51*, 520–526.
15. Fujiwara, Y. (1982). Defective repair of mitomycin C crosslinks in Fanconi's anemia and loss in confluent normal human and xeroderma pigmentosum cells. Biochim. Biophys. Acta *699*, 217–225.
16. van Vuuren, A.J., Appeldoorn, E., Odijk, H., Yasui, A., Jaspers, N.G., Bootsma, D., and Hoeijmakers, J.H. (1993). Evidence for a repair enzyme complex involving ERCC1 and complementing activities of ERCC4, ERCC11 and xeroderma pigmentosum group F. EMBO J. *12*, 3693–3701.
17. Biggerstaff, M., Szymkowski, D.E., and Wood, R.D. (1993). Co-correction of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro. EMBO J. *12*, 3685–3692.
18. Gregg, S.Q., Robinson, A.R., and Niedernhofer, L.J. (2011). Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. DNA Repair (Amst.) *10*, 781–791.
19. Ahmad, A., Enzlin, J.H., Bhagwat, N.R., Wijgers, N., Raams, A., Appledoorn, E., Theil, A.F., J Hoeijmakers, J.H., Vermeulen, W., J Jaspers, N.G., et al. (2010). Mislocalization of XPF-ERCC1 nuclease contributes to reduced DNA repair in XP-F patients. PLoS Genet. *6*, e1000871.
20. Matsumura, Y., Nishigori, C., Yagi, T., Imamura, S., and Takebe, H. (1998). Characterization of molecular defects in xeroderma pigmentosum group F in relation to its clinically mild symptoms. Hum. Mol. Genet. *7*, 969–974.
21. Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., and Vermeulen, W. (2007). First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. Am. J. Hum. Genet. *80*, 457–466.
22. Niedernhofer, L.J., Garinis, G.A., Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., et al. (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature *444*, 1038–1043.
23. Bardwell, A.J., Bardwell, L., Tomkinson, A.E., and Friedberg, E.C. (1994). Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. Science *265*, 2082–2085.
24. O'Donovan, A., Davies, A.A., Moggs, J.G., West, S.C., and Wood, R.D. (1994). XPG endonuclease makes the 3′ incision in human DNA nucleotide excision repair. Nature *371*, 432–435.

25. Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protić, M., Hübscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859–868.
26. Araújo, S.J., Nigg, E.A., and Wood, R.D. (2001). Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. Mol. Cell. Biol. *21*, 2281–2291.
27. Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nat. Genet. *43*, 142–146.
28. Crossan, G.P., van der Weyden, L., Rosado, I.V., Langevin, F., Gaillard, P.H., McIntyre, R.E., Gallagher, F., Kettunen, M.I., Lewis, D.Y., Brindle, K., et al.; Sanger Mouse Genetics Project. (2011). Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. Nat. Genet. *43*, 147–152.
29. Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W., et al. (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nat. Genet. *43*, 138–141.
30. de Laat, W.L., Appeldoorn, E., Jaspers, N.G., and Hoeijmakers, J.H. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. J. Biol. Chem. *273*, 7835–7842.
31. Tian, M., Shinkura, R., Shinkura, N., and Alt, F.W. (2004). Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. Mol. Cell. Biol. *24*, 1200–1205.
32. McWhir, J., Selfridge, J., Harrison, D.J., Squires, S., and Melton, D.W. (1993). Mice with DNA repair gene *(ERCC-1)* deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. Nat. Genet. *5*, 217–224.
33. Weeda, G., Donker, I., de Wit, J., Morreau, H., Janssens, R., Vissers, C.J., Nigg, A., van Steeg, H., Bootsma, D., and Hoeijmakers, J.H.J. (1997). Disruption of mouse *ERCC1* results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. Curr. Biol. *7*, 427–439.
34. Kim, H., and D'Andrea, A.D. (2012). Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev. *26*, 1393–1408.
35. Garaycoechea, J.I., Crossan, G.P., Langevin, F., Daly, M., Arends, M.J., and Patel, K.J. (2012). Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. Nature *489*, 571–575.
36. Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillón, J., Ramírez, M.J., Pujol, R., et al. (2013). Mutations in *ERCC4*, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia. Am. J. Hum. Genet. *92*. Published online April 25, 2013. http://dx.doi.org/10.1016/j.ajhg.2013.04.002.

PRKDC mutations in a SCID patient with profound neurological abnormalities

Lisa Woodbine,[1] Jessica A. Neal,[2] Nanda-Kumar Sasi,[2] Mayuko Shimada,[3] Karen Deem,[4] Helen Coleman,[5] William B. Dobyns,[6] Tomoo Ogi,[3,7] Katheryn Meek,[2] E. Graham Davies,[8] and Penny A. Jeggo[1]

[1]Genome Damage and Stability Centre, University of Sussex, Brighton, United Kingdom. [2]College of Veterinary Medicine, Department of Microbiology and Molecular Genetics, and Department of Pathobiology and Diagnostic Investigation, Michigan State University, East Lansing, Michigan, USA. [3]Nagasaki University Research Centre for Genomic Instability and Carcinogenesis (NRGIC), Nagasaki, Japan. [4]Department of Paediatrics, Queen Alexandra Hospital, Portsmouth, United Kingdom. [5]Solent NHS Trust, Portsmouth, United Kingdom. [6]Center for Integrative Brain Research, Seattle Children's Hospital, Seattle, Washington, USA. [7]Department of Molecular Medicine, Atomic Bomb Disease Institute, Graduate School of Biomedical Sciences, Nagasaki, Japan. [8]Centre for Immunodeficiency, Great Ormond Street Hospital and Institute of Child Health, London, United Kingdom.

The DNA-dependent protein kinase catalytic subunit (DNA-PKcs; encoded by *PRKDC*) functions in DNA non-homologous end-joining (NHEJ), the major DNA double strand break (DSB) rejoining pathway. NHEJ also functions during lymphocyte development, joining V(D)J recombination intermediates during antigen receptor gene assembly. Here, we describe a patient with compound heterozygous mutations in *PRKDC*, low DNA-PKcs expression, barely detectable DNA-PK kinase activity, and impaired DSB repair. In a heterologous expression system, we found that one of the *PRKDC* mutations inactivated DNA-PKcs, while the other resulted in dramatically diminished but detectable residual function. The patient suffered SCID with reduced or absent T and B cells, as predicted from *PRKDC*-deficient animal models. Unexpectedly, the patient was also dysmorphic; showed severe growth failure, microcephaly, and seizures; and had profound, globally impaired neurological function. MRI scans revealed microcephaly-associated cortical and hippocampal dysplasia and progressive atrophy over 2 years of life. These neurological features were markedly more severe than those observed in patients with deficiencies in other NHEJ proteins. Although loss of DNA-PKcs in mice, dogs, and horses was previously shown not to impair neuronal development, our findings demonstrate a stringent requirement for DNA-PKcs during human neuronal development and suggest that high DNA-PK protein expression is required to sustain efficient pre- and postnatal neurogenesis.

Introduction

DNA nonhomologous end-joining (NHEJ) represents the major DNA double strand break (DSB) repair process in mammalian cells (1, 2). Thus, cells, mice, and patients defective in NHEJ proteins show markedly reduced DSB repair and radiosensitivity. NHEJ also functions during immune development due to its requisite role in V(D)J recombination (3). V(D)J recombination mediates immunoglobulin and T cell receptor gene assembly from variable (V), diversity (D), and joining (J) gene segments. Recombination is initiated by the RAG1/2 endonuclease, which introduces DSBs at recombination signal sequences (RSSs). The DNA ends, including the sequences encoding the antigen receptors (termed *coding ends*), have hairpinned termini; the RSSs are blunt ended. Rejoining of these recombination intermediates requires NHEJ proteins. Consequently, NHEJ defects in animals and humans causes SCID with reduced or absent T and B cells.

6 core NHEJ components assemble as 2 complexes (1). The Ku70/80 heterodimer is the DNA double-stranded (ds) end recognition protein, binding DNA ds ends with high affinity. Ku recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs; encoded by *PRKDC*), generating the DNA-PK holoenzyme. The DNA-PK complex prevents unbridled exonuclease digestion of DNA ends, enhances appropriate end-processing, and recruits the ligation complex, which encompasses XRCC4, DNA ligase IV, and XLF (also known as cernunnos). Ionizing radiation–induced (IR-induced) DSBs are rejoined with 2 component kinetics: a fast process rejoins approximately 85% of DSBs, whereas the remaining 15% are rejoined with slower kinetics (4). Core NHEJ proteins (Ku, DNA-PKcs, DNA ligase IV, XRCC4, and XLF) are required for both processes in G0/G1 phase. However, the nuclease, Artemis, functions uniquely in the slow process. Artemis is also required to cleave hairpin coding ends during V(D)J recombination (5). DNA ligase IV and XRCC4 are essential for NHEJ, whereas XLF enhances the process, but is not essential. Similarly, DNA-PKcs is facilitating, but not essential, for the fast process (6, 7). However, there is a more stringent requirement for DNA-PKcs for the slow DSB repair process, similar to the need for Artemis (8). Although DNA-PKcs has Artemis-independent functions, an important role of DNA-PKcs lies in activating Artemis (5, 9).

The NHEJ genes have been inactivated in mice with differing consequences. Ku-defective mice are viable but have SCID, severe growth retardation, and premature aging (10). Mice lacking XRCC4 or DNA ligase IV are embryonic lethal due to increased neuronal apoptosis (11, 12). Mice lacking XLF are, in contrast, viable with mild immunodeficiency and have no marked neuronal phenotype (13). Finally, mice lacking DNA-PKcs are viable, with no overt phenotype other than SCID (14). Spontaneous mutations in *PRKDC* have also been described in inbred horses and dogs (15–17). Such animals have SCID, but otherwise develop normally, although DNA-PKcs deficiency in dogs causes a mild

Conflict of interest: The authors have declared that no conflict of interest exists.
Citation for this article: *J Clin Invest.* 2013;123(7):2969–2980. doi:10.1172/JCI67349.

intrauterine growth delay and cellular proliferation deficits (17). Importantly, neurological development is normal. The mutational changes in these animal models confer null phenotypes in all assays used, which suggests either that DNA-PKcs is nonessential or that less than 1% activity suffices for viability and development in these animal models.

Approximately 30% of SCID patients have V(D)J recombination defects, with approximately 70% having mutations in *RAG1* or *RAG2* (18, 19). Less commonly, SCID can be caused by mutations in NHEJ genes. Importantly, LIGIV syndrome and XLF patients harbor hypomorphic mutations in the genes encoding DNA ligase IV or XLF, respectively, and display T and B cell lymphocytopenia, growth delay, and microcephaly (20, 21). Generally, these patients do not have profound neurodevelopmental problems, and the neuronal deficit is not progressive postnatally. A mouse model for LIGIV syndrome shows small growth and elevated apoptosis in the embryonic neuronal stem and progenitor cell compartment (22). One LIGIV patient who received radiotherapy died of radiation morbidity (23). Collectively, these findings suggest that reduced DNA ligase IV activity confers clinical radiosensitivity and additionally that NHEJ functions during embryonic neuronal development. In contrast, mutations in the gene encoding Artemis are often null and confer radiosensitive SCID (RS-SCID), but no overt reduced growth or microcephaly (24). A single patient with a homozygous missense mutation in *PRKDC* has been previously reported (25). The mutational change did not affect DNA-PKcs expression nor enzymatic activity, but impaired Artemis activation. Like Artemis-defective patients, this DNA-PKcs–deficient patient had SCID but was otherwise developmentally normal.

Given the occurrence of spontaneous mutations in *PRKDC* in dogs and horses, it is surprising that patients with mutations in *PRKDC* have not been more frequently reported. However, recent work has strongly suggested that the DNA-PK complex is essential in humans. Elegant work from Hendrickson and colleagues has demonstrated that homozygous loss of Ku in human cells causes loss of viability due to rapid telomere shortening (26–28). In contrast, disruption of *PRKDC* in the same genetic background had a more modest impact on telomere maintenance and did not preclude viability, although it did confer profound proliferation and genomic stability defects (29). It has therefore been suggested that a *PRKDC* null mutation would be incompatible with life in humans (16, 25, 30).

Here, we identified a SCID patient with substantially reduced DNA-PKcs protein levels, undetectable DNA-PK activity, and impaired DSB rejoining attributable to loss of DNA-PK activity. The patient had a point mutational change in 1 allele that severely impaired but did not abrogate DNA-PK activity, strongly suggesting that it represents a hymophric mutation. A cDNA product resulting from aberrant splicing was also detected. The patient's cells had substantially reduced DNA-PKcs protein levels and impaired DSB rejoining that was attributable to loss of DNA-PK activity. The mutational change in 1 allele resulted in the loss of an exon and appeared to be inactivating; the other mutational change, however, was hypomorphic. The patient had marked neurological abnormalities, remarkably more severe than observed in hitherto described NHEJ-defective SCID patients. Strikingly, neuronal atrophy was observed postnatally. Our findings highlight an important role for DNA-PKcs during neuronal development and maintenance in humans.

Results

Clinical features. Male patient NM720, the first child of nonconsanguineous parents, exhibited poor intrauterine growth on antenatal ultrasounds. There was prolonged rupture of the membranes and an abnormal cardiotocograph, leading to delivery by caesarean section at 37 weeks. Birth weight was 2.1 kg (0.4th–2nd percentile), and occipitofrontal circumference (OFC) was 32.2 cm (9th–25th percentile). He was admitted to the hospital at 3 weeks of age with suspected sepsis and persistent oral and perineal candidiasis. A diagnosis of SCID was made at 5 weeks with complete absence of T and B cells, normal NK cell numbers, and undetectable IgA and IgM levels ($T^-B^-NK^+$). At 2 years of age, there were still no detectable circulating T cells (0.02×10^9 cells/l) and no B cells. The patient was noted to be microcephalic (OFC, 31.5 cm; <0.4th percentile) and dysmorphic with prominent forehead, wide nasal bridge, deep-set eyes, long philtrum with thin upper lip, small chin, low-set ears with overfolded helices, overlapping fingers, and postaxial polysyndactyly of the right foot. He had micropenis, but with normal descended testes and normal pituitary function. At gastrostomy insertion, the stomach was noted to be very small and nondistensible. A spleen was detectable by ultrasound. Neurologically, the patient showed little developmental progress with very poor head growth (OFC, 35.2, 36.8, 38.6, and 39.2 cm at 4, 6.5, 11, and 30 months, respectively). He developed seizures. Investigations including cerebrospinal fluid examination failed to reveal any microbial pathogens. He had bilateral profound sensorineural hearing loss on brain stem–evoked responses and severe visual impairment, with electroretinograms showing markedly degraded responses and occipital flash and binocular occipital pattern visual evoked potentials showing degraded, delayed, broadened, and inconsistent responses. At 29 months of age, his neurodevelopmental level was that of a 3- to 4-month-old. He died at 31 months of age with intractable seizures.

Neurological imaging of the patient at 3 months and 2 years of age. An MRI scan undertaken at 3 months of age revealed microcephaly and an atypical cortical malformation consisting of a diffusely reduced number and depth of sulci (i.e., simplified gyral pattern), mild frontal pachygyria with the frontal cortex relatively thicker than the posterior cortex (4 mm versus 2–3 mm), and a blurred cortical–white matter border (Figure 1). The hippocampi were small and dysplastic, and the posterior pituitary bright spot was absent, suggestive of disruption of the posterior pituitary-hypothalamic axis. Additional changes included hypomyelinated white matter, mildly enlarged lateral ventricles, thin and mildly short corpus callosum, cavum septi pellucidi and cavum vergae, and moderate cerebellar vermis hypoplasia. In summary, the brain images demonstrated microcephaly-associated cortical and hippocampal dysplasia that differed from lissencephaly. The irregular or pebbled surface and microgyri that characterize typical or "coarse" polymicrogyria were not seen, although the overall immature appearance might have obscured the more subtle or "delicate" form of polymicrogyria (31). A second MRI scan at 2.4 years showed striking progressive atrophy involving all brain regions compared with the first scan (Figure 1), corresponding to one of the major microcephaly groups (group 3; microcephaly with simplified gyri and enlarged extra-axial space) defined in a recent paper (32).

Defective DSB repair in patient fibroblasts. Patient fibroblasts were established to assess the possibility of a RS-SCID disorder. Poor cloning efficiency of the primary fibroblasts precluded clonogenic radiosensitivity analysis. An hTERT immortalized line, NM720

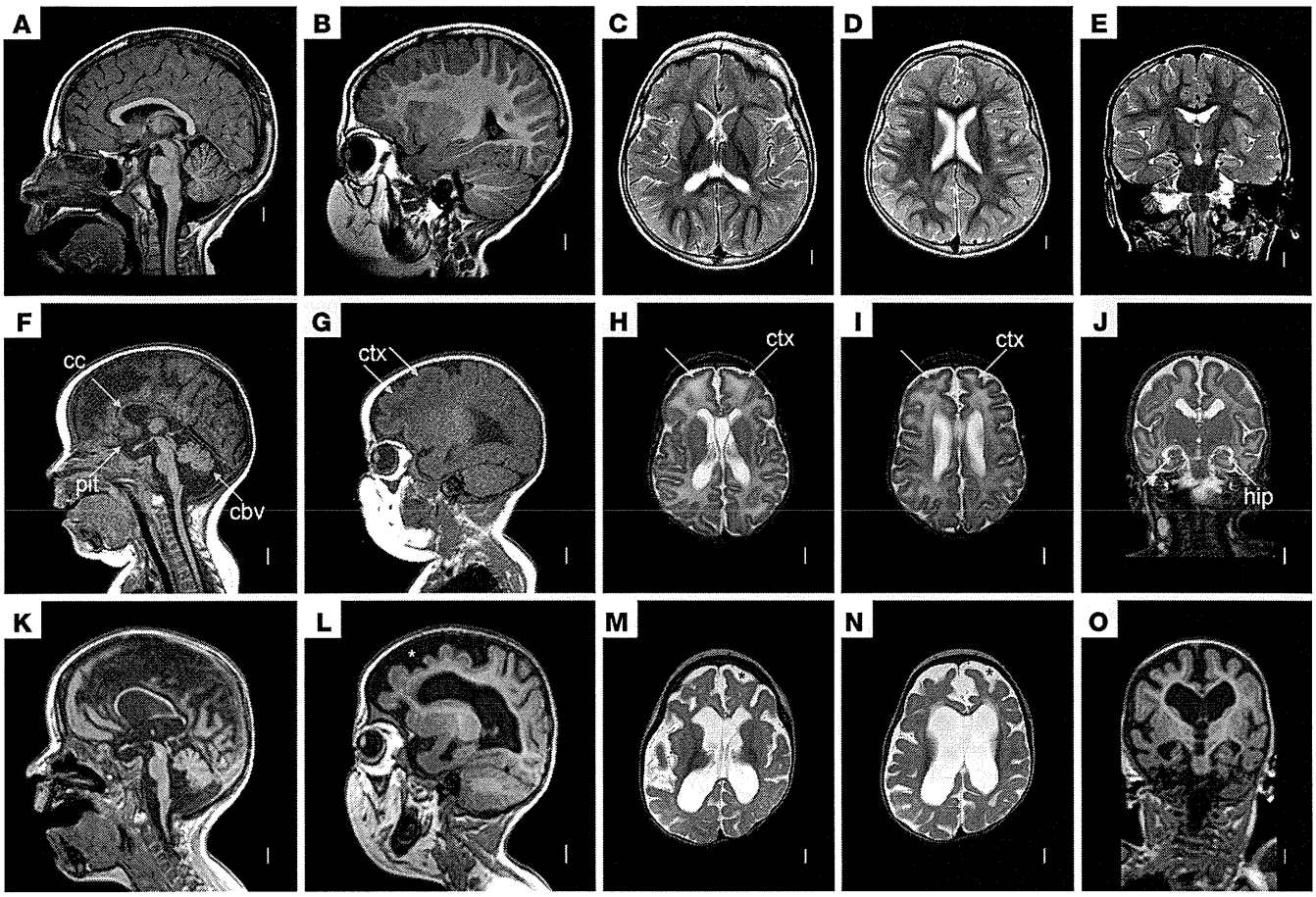

Figure 1
MRI scan images of patient NM720. Brain MRI in a normal 3-year-old child (**A**–**E**) and in patient NM720 at 3 months (**F**–**J**) and 2 years (**K**–**O**). In the patient, midline sagittal T1-weighted images (**F** and **K**) showed an absent pituitary bright spot (pit), thin and short corpus callosum (cc), and small uprotated cerebellar vermis (cbv). Parasagittal (**G** and **L**) and axial T2-weighted (**H**, **I**, **M**, and **N**) images showed reduced number of gyri, and mildly thick cortex with blurred gray-white border (ctx). Coronal images (**J** and **O**) showed the same changes in the cortex, as well as mildly small hippocampi. Striking progressive atrophy was seen involving all brain regions on the second scan (asterisks in **L**–**N**).

hTERT, was established, but these cells also plated poorly, precluding analysis. We have previously exploited γH2AX foci analysis to monitor DSB repair (8). This assay uses nondividing cells, avoiding complications from γH2AX formation during replication, and can be used with poorly growing cells. Enumeration of γH2AX formation and loss after 3 Gy IR revealed a pronounced DSB repair defect, which was distinct from that observed in either LIGIV syndrome or Artemis-defective patient cells (Figure 2A). LIGIV syndrome cells (lines 180BR and 495BR) displayed a pronounced DSB repair defect at 2 and 6 hours after IR. However, the mutational changes identified in patients are hypomorphic, and cells retain residual ligase activity, consistent with the findings in mice that DNA ligase IV is essential. Thus, DSB repair ensues slowly, as evidenced by the diminished levels of residual DSBs from 24 to 48 hours (see also ref. 33). Interestingly, the mutational changes in 495BR cells were more severe than those in 180BR cells, resulting in a greater DSB repair defect, which also correlates with more severe clinical features. A similar profile is observed for XLF-deficient cells (6). In contrast, the repair defect in an Artemis-null cell line, F02/385, was persistent but only evident at later times (beginning at 24 hours), since Artemis is dispensable for the fast DSB repair process but essential for the slow DSB repair component (8). In contrast to these phenotypes, patient cells displayed a small defect at 2 hours, but the defect was retained at 8 days after IR (Figure 2, A and B). 709BR cells, derived from the patient's mother, showed normal repair; cells from the father were not available. This profile closely corresponded to, but was milder than, that observed in *Prkdc*$^{-/-}$ mouse embryonic fibroblasts (data not shown) and in control human cells treated with the DNA-PK–inhibiting drug (DNA-PKi) NU7441 (Figure 2B). Furthermore, addition of DNA-PKi to patient cells was not additive to this repair defect and was indistinguishable from DNA-PKi–treated control cells. These findings are consistent with an epistatic relationship and raise the possibility that there could be a hypomorphic defect in DNA-PKcs in patient cells. Collectively, these findings suggest that the patient is defective in an NHEJ protein and implicate DNA-PKcs as a candidate defect.

Severely reduced DNA-PKcs expression and activity in patient cells. Next, we carried out immunoblotting of the 7 core NHEJ proteins as well as LIS1, a protein frequently deficient in patients displaying lissencephaly, which was also evaluated as a causal defect (Figure 3A). Strikingly, although most NHEJ proteins were expressed at

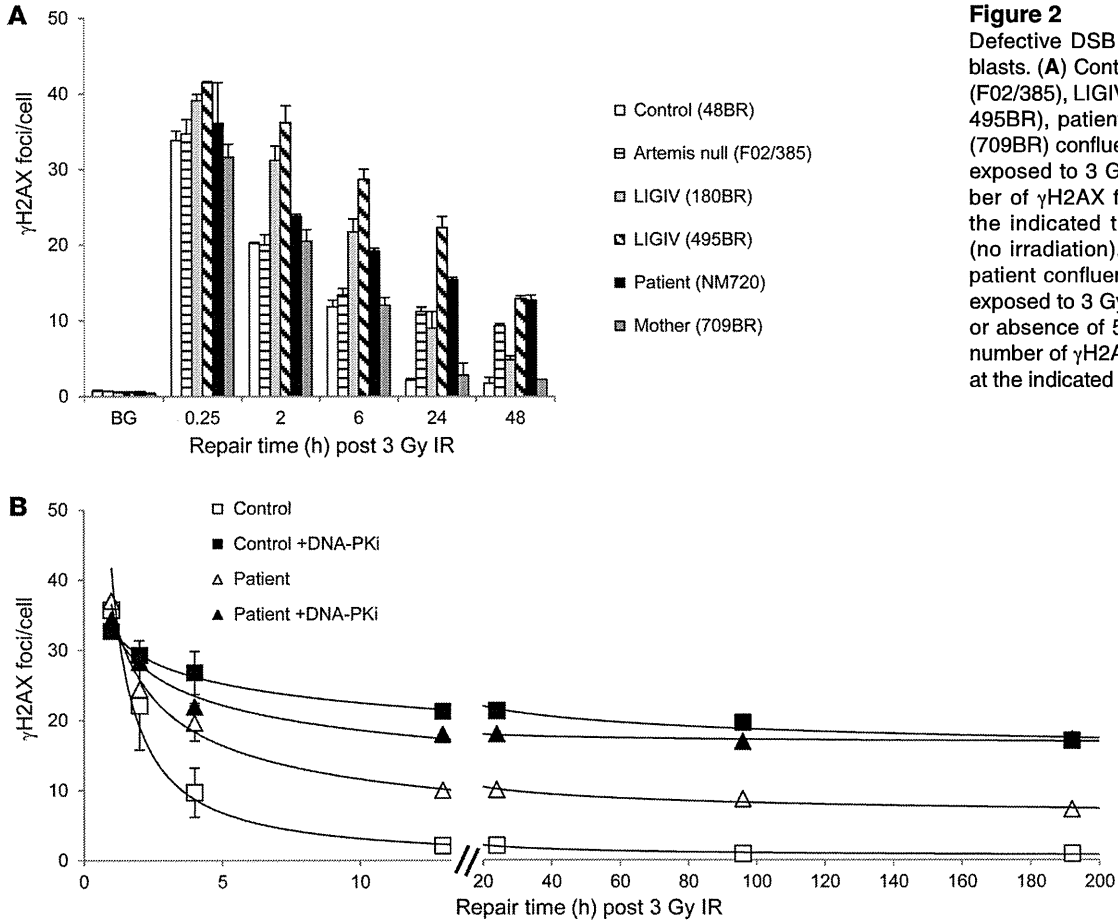

Figure 2
Defective DSB repair in patient fibroblasts. (**A**) Control (48BR), Artemis null (F02/385), LIGIV syndrome (180BR and 495BR), patient (NM720), and mother (709BR) confluent G0 phase cells were exposed to 3 Gy γ-rays, and the number of γH2AX foci was determined at the indicated times. BG, background (no irradiation). (**B**) 1BR3 control and patient confluent G0 phase cells were exposed to 3 Gy γ-rays in the presence or absence of 5 μM DNA-PKi, and the number of γH2AX foci was enumerated at the indicated times.

normal levels, DNA-PKcs expression was dramatically reduced in patient cells. However, residual DNA-PKcs protein was detectable and verified by comparison with M059J, a human tumor cell line that lacks DNA-PKcs expression (34). Furthermore, the patient's mother's cells showed approximately 50% of the level of DNA-PKcs observed in control cells (Figure 3B).

Additionally, quantitative analysis of DNA-PKcs activity present in hTERT fibroblasts derived from control (1BR3), patient, and mother cells revealed no detectable DNA-PK activity in patient hTERT cells and approximately 50% of the level of WT activity in the mother's hTERT cells (Figure 3C). This assay is extremely sensitive, and we estimate that 1% of the control level would be detectable. These findings confirmed a profound defect in DNA-PK activity in patient cells.

Mutational changes in PRKDC. DNA-PKcs is one of the largest known cDNAs (35). To search for mutational changes, 9 overlapping PCR fragments from DNA-PKcs cDNA were amplified and sequenced from control and patient, revealing a c.10721C>T mutational change in 1 allele of the patient creating an alanine to valine amino acid substitution at the highly conserved residue 3574 (p.A3574V) (Figure 4, A and E). The mother also carried this mutational change.

Additionally, we observed a double sequence commencing at c.1624, which represents the boundary between exons 15 and 16, suggesting that exon 16 might be deleted in the patient (Figure 4B). To substantiate this, primers were designed in exons 15 and 17. RT-PCR using cDNA derived from control cells revealed a single product, whereas 2 discrete bands were observed using patient cells (Figure 4C). The RT-PCR amplification products were cloned, and individual clones were sequenced. Whereas all clones from the control cell line showed the anticipated sequence encompassing exon 16, approximately 50% (10 of 20) of the clones derived from the patient lacked exon 16 (the other 50% included exon 16). Importantly, this deletion did not affect the reading frame. This is consistent with the finding that 50% of the RT-PCR amplification products were derived from the exon 16–deleted cDNA, demonstrating that there was no substantial nonsense-mediated decay (NMD). To confirm normal expression of the exon 16–deleted allele, we performed quantitative RT-PCR analyses to determine the transcript levels of the full-length WT, exon 16–deleted, and mutant *PRKDC* alleles in the patient, the mother, and a normal control (Figure 4D). The mutant c.10721C>T allele was expressed to similar extents in the patient and mother, while no mutant cDNA was detected in the normal control. Importantly, the WT c.10721C allele was expressed nearly 2-fold higher in the normal control relative to the patient or the mother. This strongly indicates that the exon 16–deleted, c.10721C, and c.10721C>T alleles are expressed at similar levels and thus are equally stable.

Finally, we examined DNA-PKcs cDNA from 15 normal individuals and did not observe any abnormal splicing involving exon 16, which suggests that this is not a common splice variant. Sequencing of exon 16 from the patient's genomic

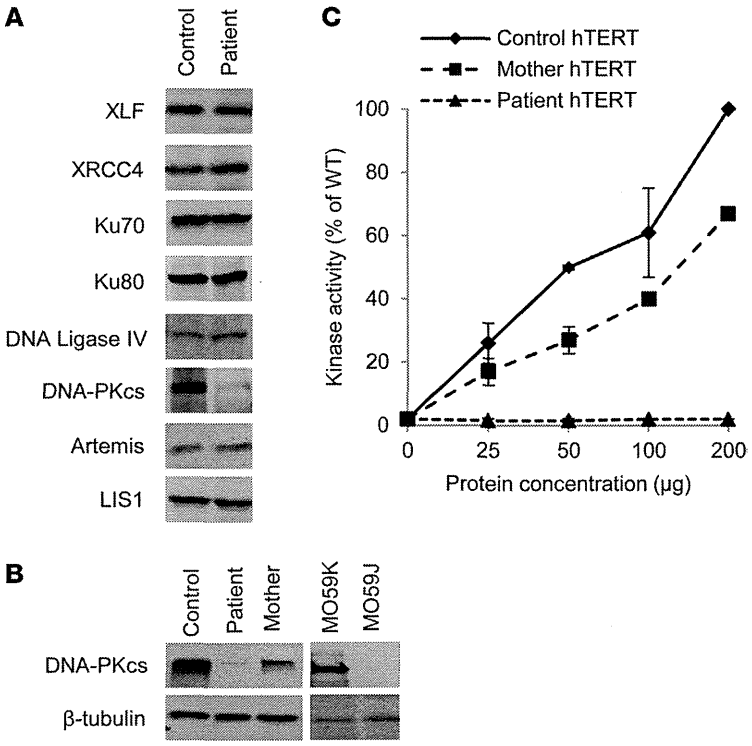

Figure 3

Reduced DNA-PKcs expression and activity in patient cells. (**A**) Western blotting analysis shows reduced DNA-PKcs protein in patient cells. Whole cell extracts (50 μg) from 1BR3 control or patient cells were processed for Western blotting using the indicated antibodies. (**B**) Western blotting analysis confirmed that residual DNA-PKcs protein was validated against human tumor cells, M059K and M059J. M059J cells are null for DNA-PKcs (53). (**C**) DNA-PK–dependent kinase activity was examined using whole cell lysates from 1BR3 control, mother, and patient immortalized primary hTERT fibroblasts.

DNA revealed a normal sequence. Additionally, intron 16 was sequenced from the patient, and an IVS16+1510insA mutation 700 bp upstream of the splice site was identified. This change was not present in the mother's genomic DNA (Supplemental Figure 1; available online with this article; doi:10.1172/JCI67349DS1). However, considering the substantial distance of this change from the exon 15/16 splice site, it is unclear whether this represents the causal mutational change causing exon skipping. This change, however, was not a reported polymorphism or SNP (Supplemental Figure 1).

In summary, patient cells harbored mutational changes in *PRKDC*. The maternally derived allele generated a p.A3574A mutant protein. Additionally, we identified a splicing abnormality generating a cDNA lacking exon 16. The location of these mutational changes and critical domains in *PRKDC* are shown in Figure 4E.

PRKDC complements the DSB repair defect observed in patient cells. To test complementation, a previously constructed GFP-tagged full-length DNA-PKcs cDNA was used (36). 1BR3 control and patient hTERT fibroblasts were transfected with DNA-PKcs cDNA and exposed to 3 Gy IR 48 hours later, and γH2AX foci were enumerated 0.25, 8, and 16 hours after IR in cells staining positively for GFP. GFP staining was used to identify and analyze only transfected cells, which represented less than 10% of the cell population (expected given the size of the DNA-PKcs cDNA and use of patient fibroblasts). Strikingly, we observed almost full correction of the DSB repair defect in patient hTERT cells expressing DNA-PKcs (Figure 5). As a control, the DNA-PKcs cDNA was also transfected into 180BR LIGIV syndrome TERT cells; as expected, no complementation of the DNA repair defect was observed. Similarly, transfection of 1BR3 control hTERT cells or the patient's mother's hTERT cells with DNA-PKcs cDNA did not affect the cells' rate of DSB repair. Collectively, these data provide compelling evidence that the DSB repair defect observed in patient cells is a consequence of impaired DNA-PKcs function.

p.A3574V or loss of exon 16 dramatically impairs DNA-PK function and activity. A3574, which lies outside the kinase domain within the large FAT domain, is a highly conserved residue in DNA-PKcs (Figure 4E). Nonetheless, an alanine to valine substitution does not represent a dramatic change. To gain further insight into whether p.A3574V affects DNA-PKcs function, this substitution was introduced into DNA-PKcs cDNA by site-directed mutagenesis and cloning. WT and mutant *PRKDC* constructs were introduced into V3, a CHO cell line that lacks functional DNA-PKcs (37). Stable clones were isolated, and those that expressed equivalent levels of DNA-PKcs (as determined by immunoblotting) were selected for further study. To examine whether the p.A3574V substitution affects DNA-PKcs function in response to damage-induced DSBs, clonogenic survival assays were performed after exposure to the DSB-inducing agent zeocin (Figure 6A). Whereas WT human DNA-PKcs substantially reversed the zeocin sensitivity of V3 cells, cells expressing the p.A3574V mutant were as sensitive to zeocin as those transfected with empty vector. We conclude that A3574V DNA-PKcs substantially disrupts DNA-PKcs function in response to damage-induced DSBs.

We next tested whether A3574V DNA-PKcs could function to repair RAG-induced DSBs during V(D)J recombination. V(D)J recombination can be assessed in somatic cells using plasmid substrates introduced into cultured cells together with the RAG recombination genes (38). V3 cells were transfected with RAG expression vectors, a recombination substrate that assesses coding end-joining, as well as WT, mutant, or no DNA-PKcs cDNA. As expected, WT DNA-PKcs supported substantial levels of coding end-joining (Figure 6B). Although A3574V DNA-PKcs cDNA supported considerably less coding end-joining than WT DNA-PKcs, a low but measurable level of coding joints were recovered, which suggests that A3574V DNA-PKcs retains some functional capacity. We considered that this low level of coding end-joining might be analogous to the "leaky" joining observed in mouse or hamster cells that lack DNA-PKcs, where the leaky joints are thought to be mediated by an alternative NHEJ pathway. Characteristics of leaky joining include excessive nt loss and long P elements. To examine the junctions formed from the rejoined joints, DNA was prepared from the rejoined plasmid substrates and sequenced. Surprisingly, coding joints mediated by A3574V DNA-PKcs were indistinguishable from those formed by WT DNA-PKcs (Table 1). We conclude that although A3574V DNA-PKcs cannot support normal levels of end-joining, end process-

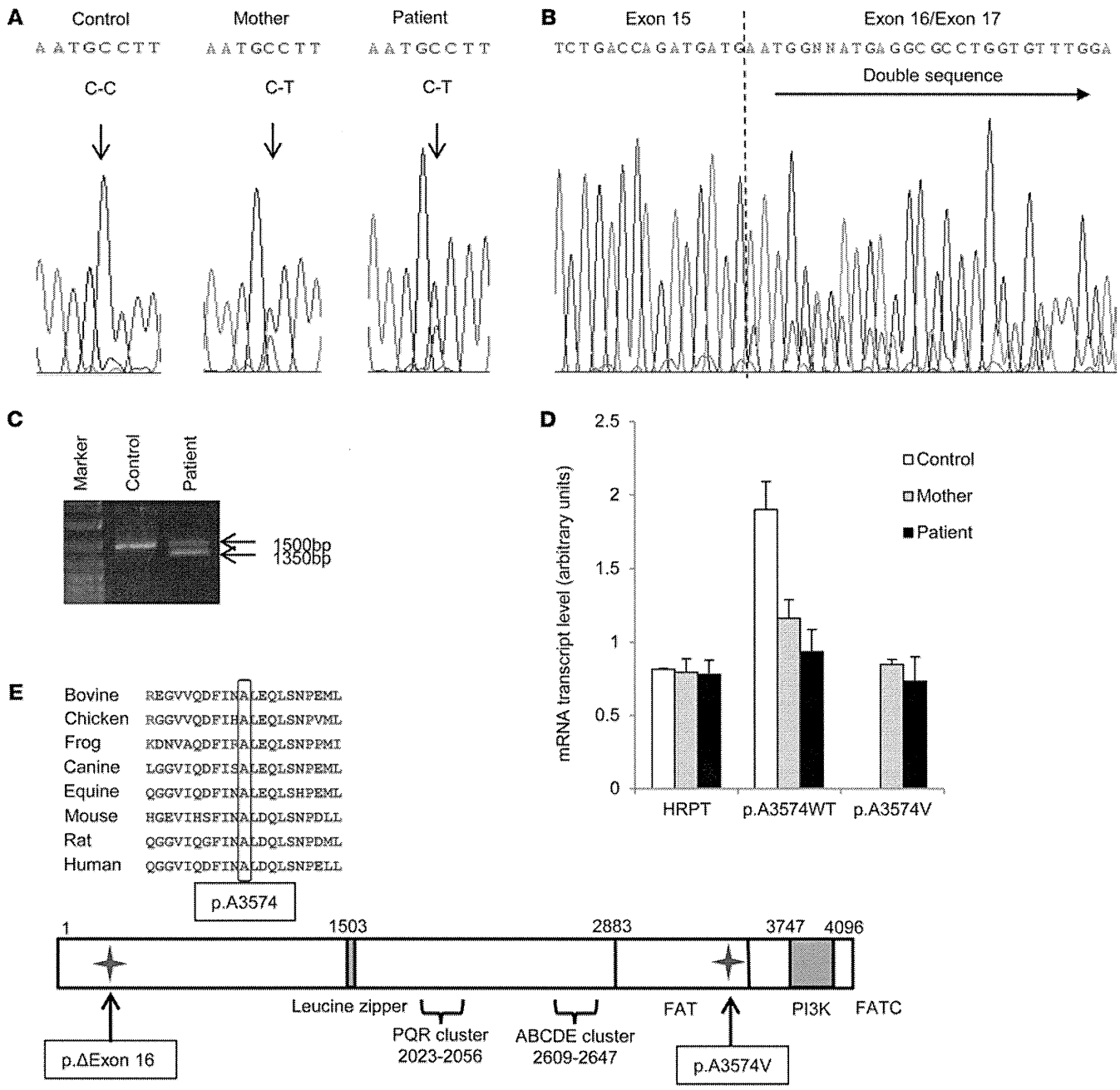

Figure 4

Identification of mutational changes in *PRKDC* cDNA. (**A**) Dye-terminator sequence figures illustrating the c.10721C>T mutational change in 1 allele of the patient and mother. (**B**) Dye-terminator sequence figures illustrating a double sequence commencing at c.1624, the bp at the boundary between exons 15 and 16. (**C**) RT-PCR amplification products using primers located in exons 15 and 17. Control cells yielded a single product of the size anticipated for the presence of exon 16. Patient cells yielded 2 bands, the smaller of which was the size expected for a product lacking exon 16. The greater signal could be the result of enhanced amplification of a smaller fragment. (**D**) mRNA transcript levels, determined using primers specific for HPRT, the WT p.A3574 allele (*PRKDC* c.10712C), or the mutant p.A3574V allele (*PRKDC* c.10712C>T). All cell lines showed equal expression of HPRT. cDNA of 48BR control cells had 2-fold greater levels of the WT allele compared with cDNA from mother or patient cells, which expressed the same level of the p.A3574 and p.A3574V mutant alleles. (**E**) Conservation of the A3574 residue between species and location of the identified mutational changes in *PRKDC* in relation to important domains.

ing of DNA ends is normal, consistent with the notion that the A3574V mutant supports reduced levels of classical NHEJ.

Finally, we examined whether A3574V DNA-PKcs retains enzymatic activity. DNA-PK from V3 extracts was pulled down onto DNA cellulose beads, and enzymatic activity was measured using a biotinylated p53 peptide as a substrate. Virtually no activity was detectable from cells expressing A3574V DNA-PKcs (Figure 6C). This was somewhat surprising, since kinase activity has previously

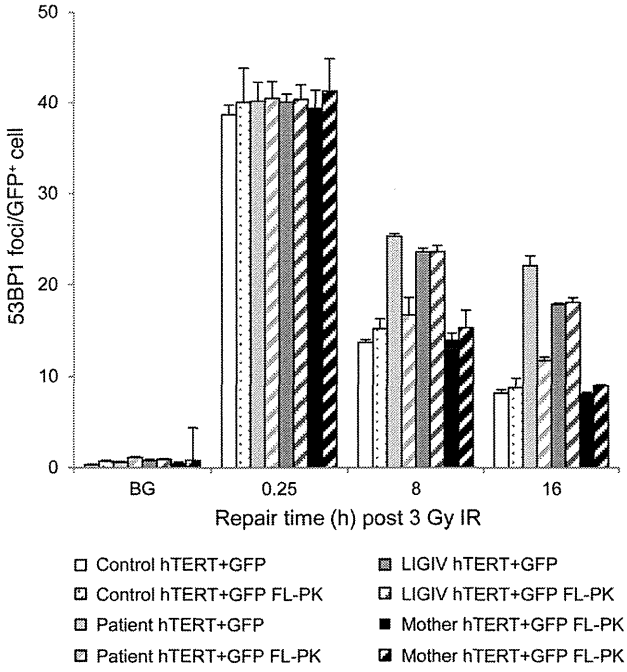

Figure 5
DNA-PKcs cDNA complements the DSB repair defect observed in patient cells. GFP-tagged full-length DNA-PKcs cDNA (GFP FL-PK) or GFP empty vector (GFP) was transfected into 1BR3 control, patient, 180BR LIGIV syndrome, and mother hTERT cells. 24 hours after transfection, cells were irradiated with 3 Gy γ-rays, and 53BP1 foci in GFP+ cells were enumerated at 8 and 16 hours. The identification of GFP+ cells was critical, since the transfection frequency using the large DNA-PKcs cDNA and human fibroblasts was low (<1%). Substantial correction of the DSB repair defect of patient cells was observed upon expression of full-length DNA-PKcs cDNA. No correction was observed after transfection of full-length DNA-PKcs cDNA into LIGIV syndrome cells.

been shown to be essential for the ability of DNA-PKcs to support coding end-joining. Thus, DNA-PKcs autophosphorylation was also assessed as an alternative measure of enzymatic activity, using a phosphospecific antibody to pS2056 DNA-PKcs, a well-studied DNA-PKcs autophosphorylation site (39). Immunoblotting of in vitro kinase reaction proteins (after 5 or 30 minutes) revealed that A3574V DNA-PKcs underwent autophosphorylation, albeit with slower kinetics and with less efficiency than the WT protein (Figure 6D). We conclude that A3574V DNA-PKcs has residual but substantially impaired catalytic activity.

Loss of exon 16 did not cause a frameshift, but the encoded protein lacked amino acids 541–592. Our previous studies have shown that *PRKDC* does not tolerate deletions (40, 41). Furthermore, the deleted region is highly conserved in vertebrate *PRKDC* and, from the available structural studies, is predicted to reside within the N-terminal "pincer" domain. DNA-PKcs cDNA lacking exon 16 was generated, and its function was assessed in V3 cells. DNA-PKcs expression in V3 clones expressing the exon 16 deletion was considerably lower than in cells expressing WT protein (Figure 6). Moreover, the clones did not reverse the zeocin-sensitive phenotype of V3 cells, failed to support VDJ coding end resolution, and yielded no detectable DNA-PKcs activity using the assays described above. Collectively, these findings provide strong evidence that loss of exon 16 dramatically affects function and likely represents a null mutational change. This finding is consistent with our previous DNA-PKcs mutational studies showing that small N-terminal deletions completely ablate function and that other exon-skipped DNA-PKcs variants are similarly nonfunctional (40, 41).

In summary, although cells expressing A3574V DNA-PKcs were similarly sensitive to zeocin-induced DSBs as cells expressing no DNA-PKcs, the mutant protein did retain residual catalytic activity, supporting a low level of coding end-joining. This is consistent with previous suggestions that very low levels of DNA-PKcs activity are sufficient to support V(D)J recombination but insufficient to reverse the radiosensitive phenotype caused by DNA-PKcs deficiency (30, 42). Finally, since neither the A3574V DNA-PKcs nor the exon 16 deletion increased zeocin sensitivity in cells completely lacking DNA-PKcs, we conclude that the phenotype cannot be attributed to a dominant negative effect on NHEJ but results from loss of DNA-PK activity. The lack of a dominant negative impact is further substantiated by the complementation of DSB repair in patient cells by GFP-tagged DNA-PKcs.

Discussion

Here, we identified an immunodeficient patient with impaired DSB repair, affecting mutations in *PRKDC*, and dramatically reduced DNA-PK activity. Although a previous patient with mutations in *PRKDC* has been reported, the homozygous mutational change exerted a subtle impact that specifically affected Artemis activation and conferred no overt clinical features other than combined immunodeficiency (25). Thus, our findings represent the first description of a patient with substantially reduced DNA-PK activity that impairs core NHEJ. The patient displayed a marked neurological phenotype substantially more severe than that of previously described LIGIV syndrome or XLF-deficient patients, who normally show mild microcephaly (20, 21). The cortical disorganization bears some resemblance to the malformation described in a fetus carrying a balanced translocation that disrupted the gene encoding XLF (43). The fetus had posterior predominant polymicrogyria with extension to the frontal lobe, and small groups of heterotopic neurons in the white matter. While the typical changes of polymicrogyria were not seen in the patient reported here, the frontal cortex was clearly abnormal, potentially consistent with the "delicate" form of polymicrogyria. Further, the blurred cortical–white matter border was also abnormal and would be consistent with heterotopic neurons in the white matter. No mutations were identified in XLF, LIGIV, or XRCC4, however. Additionally, the MRI images and the neurological features in this patient bore similarities to those previously seen in children with PNKP mutations (44). DNA-PK, as well as ataxia-telangiectasia mutated (ATM), phosphorylate PNKP (45), and PNKP functions during NHEJ, although it also functions in single-strand break repair. Further work is required to determine whether the overlapping clinical manifestations could have a common basis.

The DSB repair phenotypes are distinct from those of previously described RS-SCID patients. Cells from DNA ligase IV– and/or XLF-deficient patients have a hypomorphic phenotype that allows slow rejoining of all DSBs (8). Thus, they manifest a marked defect at early times after irradiation (e.g., 2–6 hours), but the defect becomes milder at later times, and finally all DSBs are rejoined. Artemis-defective cells show no defect at early times, since Artemis is dispensable for the fast DSB rejoining process, but a 10%–15%

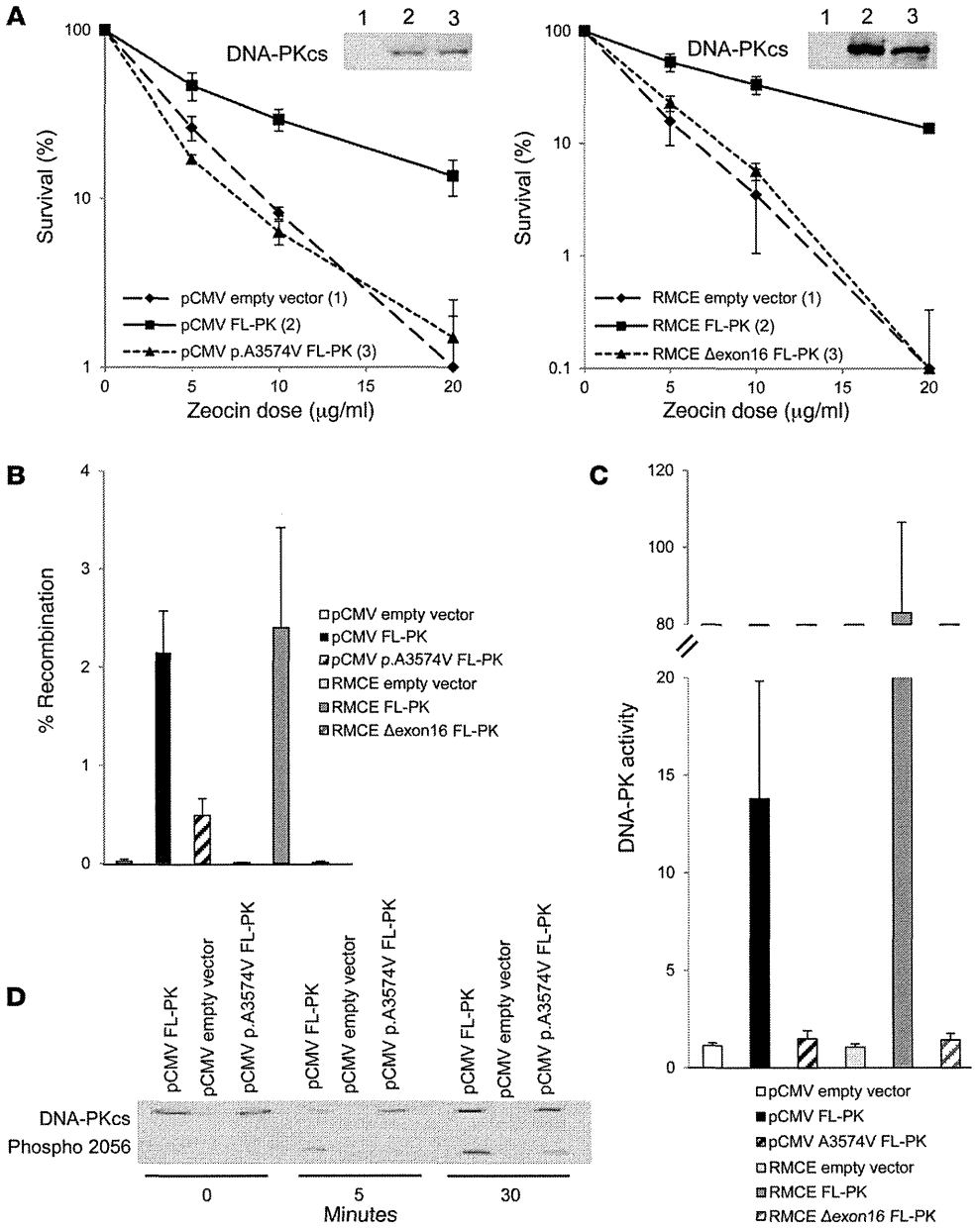

Figure 6
A3574V and exon 16–deleted DNA-PKcs impair DNA-PKcs kinase activity and DNA-PKcs function in the response to DNA damage and during V(D)J recombination. (**A**) Full-length, A3574V, or exon 16–deleted DNA-PKcs or empty vector was transiently transfected into DNA-PKcs–defective hamster V3 cells. 48 hours after transfection, cells were exposed to zeocin as indicated, and survival was estimated 7 days later. Expression of DNA-PKcs following Western blotting is also shown. Note that the exon 16–deleted DNA-PKcs cDNA was cloned into a distinct vector (RMCE) compared with A3574V DNA-PKcs cDNA (pCMV); the former is an improved vector for large inserts and was used in later studies. Survival analyses using these 2 vectors are presented separately. (**B**) Estimated percent recombination of coding junctions in V3 cells after expression of full-length, A3574V, or exon 16–deleted DNA-PKcs or empty vector together with plasmids encoding the RAG recombinases and a recombination substrate that assesses coding end-joining. (**C** and **D**) Estimation of DNA-PK activity in whole cell extracts from V3 cells expressing full-length, A3574V, or exon 16–deleted DNA-PKcs or empty vector. DNA-PK was extracted using DNA–cellulose beads, and activity was assessed using a biotinylated p53 peptide as a substrate (**C**) or by examining autophosphorylation after Western blotting and analysis using DNA-PKcs or phospho-2056–DNA-PKcs antibodies (**D**).

subfraction of DSBs remain unrejoined for prolonged times (8). The patient cells, in contrast, had a defect at early times after IR, although this was milder compared with DNA ligase IV–deficient cells; however, at later times, the defect remained marked and similar to that observed in Artemis-defective cells. Thus, patient cells cannot be described as being more or less defective than those of LIGIV syndrome patients — rather, the phenotype is distinct. The repair deficiency in patient cells appears less severe than that observed after treatment with DNA-PKi, potentially attributable to residual DNA-PK activity. Indeed, our functional analysis provides evidence that the change in the maternal allele (p.A3574V) confers low but detectable residual activity, predicting a hypomorphic phenotype. Nonetheless, despite residual activity, unrepaired DSBs were detectable at 8 days after 3 Gy in patient cells, which suggests that the slow component of DSB repair is almost fully inhibited. These findings are consistent with the notion that the different processes in which DNA-PKcs functions — e.g., fast versus slow DSB repair and V(D)J recombination — have different requirements for DNA-PK activity, and that the slow DSB repair process is dramatically affected by low DNA-PKcs activity.

In mice, loss of DNA ligase IV or XRCC4, both essential NHEJ proteins, is embryonic lethal, with marked apoptosis in the embryonic neuronal cells (11, 12). In humans, LIGIV syndrome and XLF-deficient patients, who harbor hypomorphic mutations in DNA ligase IV and XLF, respectively, display growth delay and microcephaly at birth, but the effect is not progressive postnatally (20, 21). A mouse model for LIGIV syndrome, *LigIV*^{Y288C}, which, like LIGIV syndrome patients, has a hypomorphic mutation in DNA ligase IV, also shows delayed growth, a small head, and elevated apoptosis in the embryonic neuronal stem cells (22, 46). As in

Table 1
Coding joints mediated by WT and A3574V DNA-PKcs

DNA-PKcs	Sequences	bp deletion		Junction change frequency		
	(*n*)	Mean	Median	Complete ends	P nt insertions	Short sequence homologies
WT	40	5.8	5.5	31% (25 of 80)	12% (3 of 25)	63% (25 of 40)
A3574V	55	4.3	4	34% (37 of 110)	35% (13 of 37)	36% (20 of 55)

Full-length WT and mutant A3574V DNA-PKcs cDNA was transfected into V3 cells, followed by transfection of plasmids encompassing coding junctions and the RAG1/2 nucleases. 48 hours after transfection, PCR amplification was carried out using primers that allow amplification of the rejoined coding junctions. Individual clones were sequenced. The changes observed at the junctions were similar following expression of WT and A3574V DNA-PKcs.

humans, the effect is not progressive postnatally. These findings suggest that there is a stringent requirement for the NHEJ ligation complex during embryonic neuronal development in mice and humans. Additionally, Ku-defective mice are viable, but display impaired embryonic neurogenesis (10, 47). In contrast, the animal models for DNA-PKcs loss (mice, dogs, and horses) have no evident neurological phenotype, despite having no detectable residual DNA-PKcs function (14–16, 47). Thus, the marked neuronal abnormalities in patient NM720 were completely unexpected based on findings from these animal models. Indeed, they showed not only that there is a stringent requirement for DNA-PKcs during embryonic neuronal development in humans, but that DNA-PKcs is also required postnatally.

At least 3 possible models underlying this marked clinical impact can be considered: (a) there is a more stringent requirement for DNA-PKcs for NHEJ in humans; (b) the clinical features arise from a non-NHEJ function of DNA-PKcs in humans, such as a role in telomere maintenance; or (c) there is a greater requirement for DSB repair during neuronal embryogenesis in humans. We and others have observed that mice and humans have a similar requirement for each NHEJ protein for DSB repair (8, 22). In both mouse and human cells, loss or inhibition of DNA-PKcs has a modest effect on the fast DSB repair component, but dramatically impairs the slow process (Figure 3). One explanation for residual DSB rejoining in the absence of DNA-PKcs is that an alternative form of end-joining can occur, whereas loss of DNA ligase IV results in almost complete loss of NHEJ in G0 phase mouse and human cells (48). Importantly, however, these requirements and alternative processes appear to function at similar levels in mice and in human cells (48); thus, the first possibility appears unlikely. The DNA-PK complex has also been shown to function in telomere maintenance, with this function being more important in humans than in mice (26, 28, 29, 49). However, if the clinical effect was predominantly due to telomere shortening, one might expect developmental defects in many tissues and cell types. Furthermore, since neurons are predominantly nonreplicating, the postnatal impact cannot readily be attributed to telomere shortening. An alternative possibility is that DNA-PK activity has a distinct role in the brain, potentially activating a response distinct to its role in DSB repair or activating a response that enhances survival to endogenously arising DSBs. However, although these possibilities cannot be eliminated, we favor the explanation that the clinical features are caused predominantly by the DSB repair defect and that neuronal development and maintenance in humans places a high demand on DSB repair capacity. The lack of a postnatal effect in LIGIV syndrome or XLF-deficient patients could be a consequence of their hypomorphic mutational changes.

Importantly, a distinguishing characteristic of NHEJ in primates versus all other vertebrates is remarkably high expression levels of all 3 DNA-PK component polypeptides. Indeed, DNA-PK protein and enzymatic activity is approximately 50-fold higher in human versus mouse cells (16, 50). This, however, does not affect DSB repair rates or radioresistance, which, as discussed above, are similar in rodents and humans. DNA-PK avidly binds to DSB ends, preventing their resection by exonucleases. Examination of DNA-PK dosage in cell line models has suggested that there is an inverse correlation between DNA-PKcs expression and the use of homologous recombination (HR), an alternative process of DSB repair (30). Consistent with this model, Hendrickson and colleagues have observed that loss of telomeres in Ku- or DNA-PKcs–deficient cells is mediated by HR and is dosage dependent (e.g., haplodeficient cells have a modest telomere loss phenotype; refs. 28, 29). Similarly, targeting efficiency, which is mediated by HR, is increased in a dose-dependent manner (51). We have suggested that differing requirements for DNA-PK levels in humans may reflect a more stringent requirement to regulate HR at hyper-recombinogenic genomic regions (30, 52). Thus, the high expression of DNA-PK in primates may suggest a unique necessity to restrain HR or other DSB repair pathways, which may be particularly important during neuronal development. In this context, it is notable that our analysis of *LigIV*^{Y288C} embryos revealed that significant levels of DSBs arise in the embryonic neuronal stem and early progenitor cells in a manner temporally related to the stage of rapid replication (22). Although previous studies have suggested that HR repairs replication-associated DSBs, DNA ligase IV is clearly required to repair the endogenous DSBs that arise in a manner associated with rapid replication. Importantly, the demand for efficient repair in this tissue is due not only to high DSB formation, but also to a low threshold for activating apoptosis from unrepaired DBSs (22). Thus, we propose that the high expression of DNA-PK proteins in primates is due, at least in part, to ensure efficient repair by NHEJ rather than HR during neuronal development. The embryonic neuronal stem and early progenitor cells may be particularly vulnerable to DNA-PKcs loss due to the high frequency of S/G2 phase cells, the cell cycle phase where HR functions, and where resection, the initiating step of HR, occurs avidly.

In summary, we provide the first identification of a patient with a profound defect in DNA-PKcs levels and activity. In addition to the expected SCID phenotype, the patient showed dramatically impaired neurological development and progressive brain atrophy, which was not anticipated based on animal models. Our findings suggest that high expression of DNA-PK component proteins in primates is required, at least in part, to ensure efficient embryonic neuronal development, which we postulate might be necessary to avoid DSB repair by HR.

Methods

Cell culture. Primary fibroblast lines used were 48BR and 1BR3 (control), F02/385 (Artemis null), 180BR and 495BR (LIGIV syndrome), NM720 (patient), and 709BR (mother) (8). hTERT represents immortalized derivatives. MO59K and MO59J are DNA-PK–expressing or –nonexpressing, respectively, glioblastoma cell lines (53). Cells were grown in MEM with 15% FCS, penicillin/streptomycin, and L-glutamine at 37°C and 5% CO_2. V3 (DNA-PKcs–deficient CHO cells) were cultured in α-MEM supplemented with 10% FCS, penicillin/streptomycin, and ciprofloxacin. DSB repair analysis was as described previously (8).

Immunofluorescence and immunoblotting. For immunofluorescence, anti-53BP1 or anti-γH2AX antibodies were from Millipore and Bethyl, respectively. Secondary antibodies were from Dako. Cells were fixed in 3% paraformaldehyde, 2% sucrose PBS, for 10 minutes at room temperature (RT) and permeabilized in 20 mM HEPES (pH 7.4), 50 mM NaCl, 3 mM $MgCl_2$, 300 mM sucrose, and 0.5% Triton X-100 (Sigma-Aldrich) for 2 minutes at 4°C. Primary antibody incubations were for 40 minutes at 37°C in PBS supplemented with 2% BSA (Sigma-Aldrich). Secondary incubations with anti-mouse FITC or anti-rabbit Cy3 secondary antibodies (Sigma-Aldrich) were performed at 37°C in 2% BSA for 20 minutes. Nuclei were counterstained with DAPI (Sigma-Aldrich), and coverslips were mounted in Vectashield mounting medium (Vector Laboratories). For immunoblotting, 25 or 50 μg of whole cell extracts were resolved on 7.5% PAGE gels. After transfer to PVDF membranes and blocking with 25% skim milk powder in TTBS buffer (20 mM Tris Base–HCl, pH 7.5; 150 mM NaCl; and 0.1% Tween-20), blots were incubated overnight at 4°C with anti-XLF (Eurogentec), anti-XRCC4 (Serotec), anti-KU70 (Santa Cruz), anti-KU80 (Cell Signaling), anti–DNA-LIGIV (Serotec), anti-Lis1 (Santa Cruz), or anti–DNA-PK polyclonal antibody (DPK-1; provided by S. Lees-Miller, University of Calgary, Calgary, Alberta, Canada). Blots were incubated with either rabbit, goat, or mouse secondary HRP antibodies (Dako).

DNA-PKcs kinase assay. DNA-PK kinase activity was measured using a pull-down assay (54). Briefly, whole-cell extracts from hTERT fibroblasts were mixed with 5 mg dsDNA cellulose (Sigma-Aldrich) and incubated on ice for 30 minutes at 4°C, followed by centrifugation. The pellets containing DNA cellulose–bound DNA-PKcs were washed, resuspended, and incubated with biotinylated peptide substrate supplied in the SigmaTECT DNA-PK kinase assay kit (Promega) in the presence of [γ-^{32}P] ATP for 10 minutes at 30°C. Reactions were stopped by addition of CH_3COOH, spotted onto 1.5 cm^2 phosphocellulose paper (cationic exchange paper grade P91; Whatman) and washed 3–4 times in 200 ml 15% CH_3COOH. After air drying, the signal was quantified using a STORM PhosphoImager (Molecular Dynamics/GE Healthcare).

DNA-PK sequencing. cDNA was synthesized from total RNA using random hexamer primers and SUPERSCRIPT II reverse transcriptase (Invitrogen). 9 overlapping fragments were PCR amplified using KOD polymerase (Merck Biosciences) as follows: 94°C for 2 minutes; 35 cycles of 94°C for 15 seconds, 54°C for 30 seconds, and 68°C for 90 seconds; and 68°C for 10 minutes. Each fragment was amplified 3 times in separate PCR reactions, gel purified, and sequenced in both directions using ABI BigDye Terminator Cycle Sequencing Kit V1.1 (Applied Biosystems). To confirm the exon 16 deletion, the PCR fragment covering exons 15–17 was cloned into a pGEM-5 vector, and individual clones were sequenced. Genomic DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen) and amplified in 3 overlapping fragments.

qPCR analysis. mRNA transcript levels of *PRKDC* c.10721C and c.10721T alleles in patient cells, mother cells, and 48BR control cells were quantified by the cycleave real-time PCR assay (Cycleave-qPCR; TaKaRa Co. Ltd.). *HPRT1* served as a quantification control. RNaseH-sensitive 2-color-labeled fluorescent probes that selectively recognize the *PRKDC* c.10721C and c.10721T alleles were used. qPCR results were analyzed by the ΔΔCt method. qPCR primers and probes were as follows: *PRKDC* c.10721 forward, 5′-ATAAGGAGTTTGTGGCAAGG-3′; *PRKDC* c.10721 reverse, 5′-CCAAGGCTGCATACATTC-3′; *PRKDC* c.10721C WT, (Eclipse) 5′-dAdAdG(rG)dCdAdTdTdAdAdTdA-3′ (FAM); *PRKDC* c.10721T mutant, (Eclipse) 5′-dAdAdG(rA)dCdAdTdTdAdAdTdAdAdA-3′ (ROX); *HPRT1* forward, 5′-CAGGCAGTATAATCCAAAGATG-3′; *HPRT1* reverse, 5′-ACTGGCGATGTCAATAGGA-3′; *HPRT1* probe, (Eclipse) 5′-dCdAdGdCdA(A)dGdCdT-3′ (FAM).

Plasmids and transfectants. WT DNA-PKcs expression plasmids were as described previously (42). A3574V was generated by synthesizing a portion of DNA-PKcs cDNA by multiplex PCR between the BglII (nt 10,455) and PmlI (nt 11,146) sites, including 3 nt changes within codons 3,574–3,475 (AATGCC>AACGTT) that introduce the A3574V substitution as well as noncoding substitutions that generate a novel AclI restriction site. This BglII-PmlI fragment was subcloned into a fragment of DNA-PKcs spanning the SalI site (nt 8,000) to the termination codon at nt 12,384. An Fse1 (nt 8,160) to Pml1 (nt 11,146) restriction fragment was subcloned into the complete DNA-PKcs.

Manipulation of DNA-PKcs cDNA is hampered by its size and instability. Recently, we developed an alternative expression strategy using a smaller, very low–copy number expression plasmid and the CAGG promoter. Briefly, GFP and a cloning cassette were assembled into the low–copy number plasmid pSM7 (gift of K. Yu, Michigan State University, East Lansing, Michigan, USA) between the CAGG promoter (isolated from pTriEX-2 neo; Millipore) and the bovine growth hormone poly adenylation site (isolated from PefV5His; Invitrogen). Human DNA-PKcs was cloned between unique Nhe1 and Not1 sites. cDNA was prepared from immortalized patient fibroblasts. RT-PCR was performed using primers spanning nt 1,000–2,360; 2 fragments were amplified (with and without exon 16, similar to Figure 3C). The smaller fragment was subcloned into pcr2.1 and sequenced, confirming deletion of exon 16. Plasmid DNA from this subclone as well as the CAGG-WT DNA-PKcs expression construct were prepared from bacteria lacking dcm methylase. A restriction fragment (BstZ171, nt 1,293; to SexAI, nt 1,875) spanning exon 16 was cloned from the pcr2.1 RT-PCR subclone into the CAGG-WT DNA-PKcs expression plasmid.

Complementation of primary human fibroblasts. 1BR3 control, patient, 180BR LIGIV syndrome, and mother hTERT cells were transfected with GFP-tagged full-length WT DNA-PK using Metafectine Pro Transfection Reagent (Biotex) following the manufacturer's protocol. 24 hours after transfection, cells were untreated (background) or exposed to 3 Gy IR, harvested 8 and 16 hours later, and stained for 53BP1. Only GFP^+ cells were analyzed.

Clonogenic survival assays. V3 transfectants were plated at cloning densities estimated to give scoreable colonies in the indicated dosage of zeocin. After 7 days, cell colonies were stained with 1% (w/v) crystal violet in ethanol and colony numbers were assessed. Survival was plotted as percent relative to untreated cells.

VDJ recombination assays. Extrachromosomal VDJ recombination assays were performed using the coding joint substrate pJH290 (38). V3 cells were transiently transfected with 1 μg substrate, 6 μg WT or mutant DNA-PKcs or pCMV6 vector, and 3 μg each RAG1 and RAG2 using Fugene6 transfection reagent. 48 hours later, substrate plasmids were isolated by alkaline lysis and digested with DpnI for 1 hour. Digested DNA was transformed into competent DH5α cells (Invitrogen) according to the manufacturer's instructions. Transformed cells were spread onto LB Agar plates containing 100 μg/ml ampicillin, alone or combined with 22 μg/ml chloramphenicol. Plasmid DNA from recombined pJH290 (coding joints) was isolated and sequenced.

Immunoblotting and measurement of protein kinase activity in V3 transfectants. Cell extracts were separated using 4.5% SDS polyacrylamide gels and transferred to PVDF membranes. DNA-PKcs was detected using monoclonal antibody r42-27 (gift from T. Carter, St. Johns University, New York, New York, USA). DNA-PK activity was assessed as described previously (42). Briefly, DNA-PK was pulled down onto DNA cellulose. Phosphorylation of a biotinylated peptide was assessed by capture onto SAM2 membranes (Promega Corp). Enzymatic activity was expressed as fold P-32 incorporation (bound peptide) relative to no peptide control. To assess autophosphorylation, 100 μg whole cell extracts were incubated in kinase buffer at room temperature for 5 or 30 minutes and analyzed by immunoblotting. A phosphospecific antibody was used to detect autophosphorylation at S2056 (Abcam).

Statistics. In figures where error bars are shown, results represent mean ± SD of 3 independent experiments.

Study approval. Human studies were reviewed and approved by the University of Sussex Research Governance Committee. Written informed consent was obtained from participants or their guardians.

Acknowledgments

The P.A. Jeggo laboratory is supported by the Medical Research Council, Association for International Cancer Research, and Wellcome Research Trust. The K. Meek laboratory is supported by Public Health Service grant AI048758 (to K. Meek).

Received for publication October 15, 2012, and accepted in revised form March 28, 2013.

Address correspondence to: Penny A. Jeggo, Genome Damage and Stability Centre, University of Sussex, Science Park Road, Brighton BN1 9RQ, United Kingdom. Phone: 44.273.678482; Fax: 44.273.678121; E-mail: p.a.jeggo@sussex.ac.uk. Or to: E. Graham Davies, Centre for Immunodeficiency, Great Ormond Street Hospital and Institute of Child Health, Great Ormond Street, London WC1N 3JH, United Kingdom. Phone: 44.20.7829.8834; Fax: 44.20.7813.8552; E-mail: Graham.Davies@gosh.nhs.uk. Or to: Katheryn Meek, College of Veterinary Medicine, Department of Microbiology and Molecular Genetics, and Department of Pathobiology and Diagnostic Investigation, Michigan State University, 5109 Biomedical and Physical Sciences Building, East Lansing, Michigan 48823, USA. Phone: 517.884.5361; Fax: 517.353.8957; E-mail: kmeek@msu.edu.

1. Jeggo P, Lavin MF. Cellular radiosensitivity: how much better do we understand it? *Int J Radiat Biol.* 2009;85(12):1061–1081.
2. Mahaney BL, Meek K, Lees-Miller SP. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem J.* 2009;417(3):639–650.
3. Soulas-Sprauel P, et al. V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining. *Oncogene.* 2007;26(56):7780–7791.
4. Goodarzi AA, et al. ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. *Mol Cell.* 2008;31(2):167–177.
5. Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell.* 2002;108(6):781–794.
6. Riballo E, Woodbine L, Stiff T, Walker SA, Goodarzi AA, Jeggo PA. XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. *Nucleic Acids Res.* 2009;37(2):482–492.
7. Singh SK, Wu W, Wu W, Wang M, Iliakis G. Extensive repair of DNA double-strand breaks in cells deficient in the DNA-PK-dependent pathway of NHEJ after exclusion of heat-labile sites. *Radiat Res.* 2009;172(2):152–164.
8. Riballo E, et al. A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Mol Cell.* 2004;16(5):715–724.
9. Goodarzi AA, et al. Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase 2A. *EMBO J.* 2004;23(22):4451–4461.
10. Nussenzweig A, et al. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature.* 1996;382(6591):551–555.
11. Frank KM, et al. DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Mol Cell.* 2000;5(6):993–1002.
12. Barnes DE, 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.* 1998;8(25):1395–1398.
13. Li G, et al. Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. *Mol Cell.* 2008 Sep 5;31(5):631–640.
14. Taccioli GE, et al. Targeted disruption of the catalytic subunit of the DNA-PK gene in mice confers severe combined immunodeficiency and radiosensitivity. *Immunity.* 1998;9(3):355–366.
15. Wiler R, Leber R, Moore BB, VanDyk LF, Perryman LE, Meek K. Equine severe combined immunodeficiency — a defect in V(D)J recombination and DNA-dependent protein-kinase activity. *Proc Natl Acad Sci U S A.* 1995;92(25):11485–11489.
16. Meek K, et al. SCID in Jack Russell terriers: a new animal model of DNA-PKcs deficiency. *J Immunol.* 2001;167(4):2142–2150.
17. Meek K, et al. SCID dogs: similar transplant potential but distinct intra-uterine growth defects and premature replicative senescence compared with SCID mice. *J Immunol.* 2009;183(4):2529–2536.
18. Schwarz K, et al. RAG mutations in human B cell-negative SCID. *Science.* 1996;274(5284):97–99.
19. Schatz DG. V(D)J recombination. *Immunol Rev.* 2004;200:5–11.
20. O'Driscoll M, et al. DNA Ligase IV mutations identified in patients exhibiting development delay and immunodeficiency. *Mol Cell.* 2001;8(6):1175–1185.
21. Buck D, et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell.* 2006;124(2):287–299.
22. Gatz SA, et al. Requirement for DNA ligase IV during embryonic neuronal development. *J Neurosci.* 2011;31(27):10088–10100.
23. Riballo E, et al. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Curr Biol.* 1999;9(13):699–702.
24. Moshous D, et al. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell.* 2001;105(2):177–186.
25. van der Burg M, et al. A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining. *J Clin Invest.* 2009;119(1):91–98.
26. Li G, Nelsen C, Hendrickson EA. Ku86 is essential in human somatic cells. *Proc Natl Acad Sci U S A.* 2002;99(2):832–837.
27. Wang Y, Ghosh G, Hendrickson EA. Ku86 represses lethal telomere deletion events in human somatic cells. *Proc Natl Acad Sci U S A.* 2009;106(30):12430–12435.
28. Myung K, et al. Regulation of telomere length and suppression of genomic instability in human somatic cells by Ku86. *Mol Cell Biol.* 2004;24(11):5050–5059.
29. Ruis BL, Fattah KR, Hendrickson EA. The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells. *Mol Cell Biol.* 2008;28(20):6182–6195.
30. Neal JA, Dang V, Douglas P, Wold MS, Lees-Miller SP, Meek K. Inhibition of homologous recombination by DNA-dependent protein kinase requires kinase activity, is titratable, and is modulated by autophosphorylation. *Mol Cell Biol.* 2011;31(8):1719–1733.
31. Barkovich AJ. MRI analysis of sulcation morphology in polymicrogyria. *Epilepsia.* 2010;51(suppl 1):17–22.
32. Basel-Vanagaite L, Dobyns WB. Clinical and brain imaging heterogeneity of severe microcephaly. *Pediatr Neurol.* 2010;43(1):7–16.
33. Girard PM, Kysela B, Harer CJ, Doherty AJ, Jeggo PA. Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms. *Hum Mol Genet.* 2004;13(20):2369–2376.
34. Lees-Miller SP, et al. Absence of p350 subunit of DNA-activated protein kinase from a radiosensitive human cell line. *Science.* 1995;267(5201):1183–1185.
35. Hartley KO, et al. DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell.* 1995;82(5):849–856.
36. Shin EK, Rijkers T, Pastink A, Meek K. Analyses of TCRB rearrangements substantiate a profound deficit in recombination signal sequence joining in SCID foals: implications for the role of DNA-dependent protein kinase in V(D)J recombination. *J Immunol.* 2000;164(3):1416–1424.
37. Ding Q, et al. Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Mol Cell Biol.* 2003;23(16):5836–5848.
38. Hesse JE, Lieber MR, Gellert M, Mizuuchi K. Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell.* 1987;49(6):775–783.
39. Cui X, Yu Y, Gupta S, Cho YM, Lees-Miller SP, Meek K. Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice. *Mol Cell Biol.* 2005;25(24):10842–10852.
40. Douglas P, et al. The DNA-dependent protein

kinase catalytic subunit is phosphorylated in vivo on threonine 3950, a highly conserved amino acid in the protein kinase domain. *Mol Cell Biol.* 2007;27(5):1581–1591.
41. Convery E, et al. Inhibition of homologous recombination by variants of the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs). *Proc Natl Acad Sci U S A.* 2005;102(5):1345–1350.
42. Kienker LJ, Shin EK, Meek K. Both V(D)J recombination and radioresistance require DNA-PK kinase activity, though minimal levels suffice for V(D)J recombination. *Nucleic Acids Res.* 2000;28(14):2752–2761.
43. Cantagrel V, et al. Truncation of NHEJ1 in a patient with polymicrogyria. *Hum Mutat.* 2007; 28(4):356–364.
44. Shen J, et al. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nat Genet.* 2010;42(3):245–249.
45. Zolner AE, et al. Phosphorylation of polynucleotide kinase/ phosphatase by DNA-dependent protein kinase and ataxia-telangiectasia mutated regulates its association with sites of DNA damage. *Nucleic Acids Res.* 2011;39(21):9224–9237.
46. Nijnik A, et al. DNA repair is limiting for haematopoietic stem cells during ageing. *Nature.* 2007;447(7145):686–690.
47. Gu Y, et al. Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc Natl Acad Sci U S A.* 2000;97(6):2668–2673.
48. Iliakis G. Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence. *Radiother Oncol.* 2009;92(3):310–315.
49. Ghosh G, Li G, Myung K, Hendrickson EA. The lethality of Ku86 (XRCC5) loss-of-function mutations in human cells is independent of p53 (TP53). *Radiat Res.* 2007;167(1):66–79.
50. Finnie NJ, Gottlieb TM, Blunt T, Jeggo PA, Jackson SP. DNA-dependent protein-kinase activity is absent in xrs-6 cells – implications for site-specific recombination and dna double-strand break repair. *Proc Natl Acad Sci U S A.* 1995;92(1):320–324.
51. Fattah KR, Ruis BL, Hendrickson EA. Mutations to Ku reveal differences in human somatic cell lines. *DNA Repair (Amst).* 2008;7(5):762–774.
52. Neal JA, Meek K. Choosing the right path: does DNA-PK help make the decision? *Mutat Res.* 2011;711(1–2):73–86.
53. Allalunis-Turner MJ, Lintott LG, Barron GM, Day RS, Lees-Miller SP. Lack of correlation between DNA-dependent protein kinase activity and tumour-cell radiosensitivity. *Cancer Res.* 1995;55(22):5200–5202.
54. Dai Y, et al. Non-homologous end joining and V(D) J recombination require an additional factor. *Proc Natl Acad Sci U S A.* 2003;100(5):2462–2467.

Gene 512 (2013) 267–274

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Uniparental disomy analysis in trios using genome-wide SNP array and whole-genome sequencing data imply segmental uniparental isodisomy in general populations

Kensaku Sasaki [a], Hiroyuki Mishima [a], Kiyonori Miura [b], Koh-ichiro Yoshiura [a,*]

[a] Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[b] Department of Obstetrics and Gynecology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

ARTICLE INFO

Article history:
Accepted 19 October 2012
Available online 27 October 2012

Keywords:
Human genome
Genomic integrity
DNA repair
Gene conversion
International HapMap Project
1000 Genomes Project

ABSTRACT

Whole chromosomal and segmental uniparental disomy (UPD) is one of the causes of imprinting disorder and other recessive disorders. Most investigations of UPD were performed only using cases with relevant phenotypic features and included few markers. However, the diagnosis of cases with segmental UPD requires a large number of molecular investigations. Currently, the accurate frequency of whole chromosomal and segmental UPD in a normal developing embryo is not well understood. Here, we present whole chromosome and segmental UPD analysis using single nucleotide polymorphism (SNP) microarray data of 173 mother-father-child trios (519 individuals) from six populations (including 170 HapMap trios). For two of these trios, we also investigated the possibility of shorter segmental UPD as a consequence of homologous recombination repair (HR) for DNA double strand breaks (DSBs) during the early developing stage using high-coverage whole-genome sequencing (WGS) data from 1000 Genomes Project. This could be overlooked by SNP microarray. We identified one obvious segmental paternal uniparental isodisomy (iUPD) (8.2 mega bases) in one HapMap sample from 173 trios using Genome-Wide Human SNP Array 6.0 (SNP6.0 array) data. However, we could not identify shorter segmental iUPD in two trios using WGS data. Finally, we estimated the rate of segmental UPD to be one per 173 births (0.578%) based on the UPD screening for 173 trios in general populations. Based on the autosomal chromosome pairs investigated, we estimate the rate of segmental UPD to be one per 3806 chromosome pairs (0.026%). These data imply the possibility of hidden segmental UPD in normal individuals.

1. Introduction

Uniparental disomy (UPD) is defined as the inheritance of a chromosome pair derived only from one parent (Engel, 1980). Chromosomal UPD can occur because of gamete complementation, trisomic rescue, monosomic rescue and postfertilization error (Robinson, 2000). Uniparental heterodisomy (hUPD) is defined as the inheritance of both homologous chromosomes from one parent and occurs when bivalent chromatids fail to separate during meiosis I. Uniparental isodisomy (iUPD) is defined as the inheritance of two copies of one chromosome from one parent and may occur when sister chromatids fail to separate during meiosis II. The region of UPD may extend over an entire or segmental (interstitial or telomeric) chromosome. Segmental UPD is defined as UPD of one part of a chromosome (Kotzot, 2008), and occurs due to postzygotic somatic recombination between maternal and paternal homologues (Kotzot, 2008). Problems associated with UPD include aberrant genomic imprinting and homozygosity of autosomal recessively inherited mutations.

To maintain genome integrity, cells repair DNA damage including DNA double strand breaks (DSBs), by one of two major pathways, non-homologous end-joining (NHEJ) and homologous recombination (HR) (Wyman and Kanaar, 2006). NHEJ repair performs error-prone repair by joining DNA ends directly, independent of extensive DNA sequence homology, while HR repair performs error-free repair by utilizing the undamaged homologous sequence as the template for repair (Hartlerode and Scully, 2009). DNA damage during DNA replication can be repaired by HR using the intact sister chromatid (Sonoda et al., 2006) and inter-sister chromatid HR during S phase

Abbreviations: DSBs, double strand breaks; HR, homologous recombination; NHEJ, non-homologous end joining; UPD, uniparental disomy; hUPD, uniparental heterodisomy; iUPD, uniparental isodisomy; NGS, next-generation sequencing; WGS, whole-genome sequencing; LCLs, lymphoblastoid cell lines; SNPs, single-nucleotide polymorphisms; SNP6.0 array, Genome-Wide Human SNP Array 6.0; PartekGS, Partek Genomics Suite; INDELs, short insertions and deletions; SVs, structural variants; GATK, Genome Analysis Toolkit; CNVs, copy number variants; LTA, loss of transmitted allele; LOH, loss of heterozygosity; ROH, runs of homozygosity; QPCR, quantitative polymerase chain reaction; ESCs, embryonic stem cells.

* Corresponding author at: Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan. Tel.: +81 95 819 7118; fax: +81 95 819 7121.
E-mail addresses: nell2gene@yahoo.co.jp (K. Sasaki), hmishima@nagasaki-u.ac.jp (H. Mishima), kiyonori@nagasaki-u.ac.jp (K. Miura), kyoshi@nagasaki-u.ac.jp (K. Yoshiura).

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2012.10.035

will not result in segmental iUPD. However, several imprinting disorders such as Beckwith–Wiedemann syndrome (BWS; OMIM #130650), Prader Willi syndrome (PWS; OMIM #176270), and Angelman syndrome (AS; OMIM #105830) can be caused by UPD. In BWS almost all patients with UPD have segmental UPD; in contrast, in PWS/AS patients mostly have UPD of the whole chromosome. In addition to those imprinting disorders, recessive hereditary disease can be caused by segmental iUPD (Kotzot, 2001; Pérez et al., 2011). Because segmental iUPD can be found in some disorders, it is possible that segmental UPD can occur in normal development without any disease phenotype. Segmental iUPD could be considered the signature of HR between maternal and paternal homologues during the early stages of embryogenesis.

UPD can be detected using microsatellite analysis (Hannula et al., 2000) and methylation testing (Baumer et al., 2001), based on a limited number of markers in the chromosomal region of interest. The advent of high throughput single nucleotide polymorphism (SNP) microarray technology has recently permitted the identification of UPD in DNA samples from clinically affected individuals (Altug-Teber et al., 2005; Pérez et al., 2011), and the number of UPD case reports is increasing (Pérez et al., 2011). To assess the clinical significance of UPD, it is necessary to document the frequency and nature of UPD in the general population. Recently, several studies reported mosaic genomic variations (copy-neutral loss of heterozygosity (LOH) or acquired UPDs, trisomies and CNVs) in blood and buccal genomic DNA samples from cancer cases and controls (Jacobs et al., 2012; Laurie et al., 2012; Rodríguez-Santiago et al., 2010). However, assessing the segmental UPD in general populations using trios and genome-wide SNP array has not been performed to date.

Two thousand cases of UPD have been reported thus far (http://www.fish.uniklinikum-jena.de/UPD.html). UPD is one of the causes of "imprinting disorders" and is found at a high rate (7% for AS and 25% for PWS: Amor and Halliday, 2008). BWS has segmental UPD11p in 20% of cases (Amor and Halliday, 2008). Until 2010, 122 cases were reported as segmental UPD, and ~65% of those cases were due to BWS and segmental paternal UPD 11p (Liehr, 2010). However, segmental UPD of other chromosomes not associated with a cytogenetically abnormal karyotype is extremely rare (Kotzot, 2001), and UPD has no effect on phenotype at many chromosomal region. Although UPD cases without clinical abnormalities have been reported in the literature, they were found by chance or were due to repeated abortions in a family with chromosomal rearrangement (Liehr, 2010). Thus, despite the increasing importance of UPD as a disease causing mechanism, the precise UPD rate, including segmental UPD, in the general population is unknown.

Little information is available regarding DNA repair in the early development of zygotes. But it is clear that segmental iUPD detected systemically in adult can be the result from inter-allelic HR during the postzygotic period to the early embryonic stage. Therefore, we attempted to identify segmental iUPD in individuals without an abnormal phenotype. To this aim, we analyzed parent-offspring trios from SNP microarray data and also whole-genome sequencing (WGS) data of genomic DNA from two trios derived from lymphoblastoid cell lines (LCLs) during the pilot 2 data of the 1000 Genomes Project (http://www.1000genomes.org/) (Altshuler et al., 2010). WGS data was used to identify shorter iUPD, because it is difficult to identify shorter segmental iUPD by SNP microarray due to limited SNP information of the whole genome.

In this paper, we evaluated the frequency of UPD in healthy normal development.

2. Materials and methods

2.1. HapMap 3 samples

We downloaded and studied a set of 170 trios (510 samples) data from SNP6.0 arrays from 5 populations in HapMap 3 (ftp://ftp.ncbi.nlm.nih.gov/hapmap/raw_data/hapmap3_affy6.0/); 159 individuals from the Centre d'Etude du Polymorphisme Humain collected in Utah, USA, with ancestry from northern and western Europe (CEU); 33 Africans with ancestry in the southwestern USA (ASW); 81 Maasai in Kinyawa, Kenya (MKK); 174 Yoruba in Ibadan, Nigeria (YRI); and 63 Mexicans with ancestry in Los Angeles, California (MXL) (Supplementary Table 1).

2.2. Genomic DNA

We attempted to identify UPD in 173 trios. Three trios (original trio 1, trio 2 and trio 3) in this study were Japanese (JPT) and were healthy volunteers (not included in HapMap samples). The three trio's genomic DNA was extracted from peripheral blood following standard protocols. Genomic DNA for two HapMap trios (CEU family ID (FID) 1463 and YRI Y117 trio) was obtained from the Coriell Institute (http://ccr.coriell.org/sections/collections/NHGRI/?SsId=11).

2.3. Microarray analysis

We performed high-resolution genome-wide SNP genotyping and DNA copy number detection using Genome-Wide Human SNP Array 6.0 (SNP6.0 array) following the manufacturer's instructions (Affymetrix, Inc., Santa Clara, California, USA). Genotyping were performed using the default parameters in the Birdseed v2 algorithm of Genotyping Console (GTC) 4.1 software (Affymetrix). As a quality control for the genotyping, Contrast QC values were calculated as implemented in the GTC 4.1, and samples used passed the recommended values for contrast QC > 0.4. Genomic positions of the SNPs corresponded to the March 2006 human genome (hg18). Copy number and allele ratio analysis was performed by Partek Genomics Suite (PartekGS) version 6.5 (Partek Inc., St. Louis, Missouri, USA). For 3 trios of healthy volunteers and 170 trios from HapMap, the copy number reference generated from the intensities of 20 normal sample profiles in our laboratory and 100 HapMap sample profiles (no overlapping 170 trios) were used, respectively. The Hidden Markov Model (HMM) method was used to detect amplified or deleted regions using PartekGS with default parameters, and required at least 5 genomic markers to obtain CNVs call. We considered the 27 possible combinations of genotypes when each of the mother/father/child in a trio had a biallelic genotype (Supplementary Table 2). UPD genotypes were identified using in-house Ruby script from trio genotyping information exported from GTC. A UPD region was defined as a set of consecutive SNPs, where all plots had the same type (paternal and maternal UPD segment) and occurred along a chromosome. We used the criteria of a minimum of 6 consecutive UPD SNPs, with segments extending over 200 kilo bases (kb). In this study, we focused on the autosome, and chromosome X only when the offspring in the trio was a daughter. We visualized tracts of paternal uniparental inheritance (UPI-P), maternal uniparental inheritance (UPI-M), biparental inheritance (BPI), MI-S, single Mendelian inconsistencies (MI-S), double Mendelian inconsistencies (MI-D) and not informative (NI) in biallelic SNP data from trios using PartekGS'SNPtrio. Current software does not distinguish between homozygous and hemizygous states. In addition, it is known that UPD type genotypes can result from the loss of transmitted allele (LTA) (Ting et al., 2007). LTA was defined as a phenomenon in which the transmitted allele is lost (due to deletion or UPD) in the parent after the transmission to a normal child (Redon et al., 2006). Therefore, the putative UPD genotype overlapping with CNVs in trios confirmed using BEDtools (version 2.12.0) (Quinlan and Hall, 2010). The distinction between the segmental UPD as opposed to homozygosity due to small deletion is difficult to determine just by inspection of the SNP array data alone. To exclude false segmental UPD due to undetectable small CNVs, we adopted a cutoff value of a length of 200 kb or smaller. Finally, we confirmed whether known imprinting genes were present in the identified segmental UPD region. Imprinting

genes are based on Geneimprint (http://www.geneimprint.com/site/genes-by-species.Homo+sapiens.imprinted-All).

2.4. Next-generation sequencing (NGS) data

HapMap CEU 1463 and YRI Y117 trios were sequenced using multiple platforms, as described elsewhere (Altshuler et al., 2010). We downloaded BAM files (aligned to the NCBI36 reference genome using Maq v0.7) of two trios (CEU and YRI) sequenced using Illumina Genome Analyzer I, II and IIx in the 1000 Genomes Project pilot 2 (ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/pilot_data/data/) with high coverage. We focused on the autosomal and X chromosomes. Each included one offspring (daughter), father and mother: CEU daughter NA12878, father NA12891 and mother NA12892; and YRI daughter NA19240, father NA19239 and mother NA19238.

2.5. NGS bioinformatics

After downloading the BAM files, duplicate reads from samples were identified and removed using Picard (version 1.38) (http://picard.sourceforge.net/). Base quality scores were recalibrated and reads were locally realigned with the Genome Analysis Toolkit (GATK) (version 1.0.5974) (DePristo et al., 2011; McKenna et al., 2010). Coverage statistics were calculated as default using GATK'DepthOfCoverageWalker. The diploid consensus sequences and variants for autosomal and X chromosomes were obtained by the 'EMIT_ALL_CONFIDENT_SITES (using -stand_call_conf 50.0 and -stand_emit_conf 10.0)' command of the GATK'UnifiedGenotyper. SNPs and short insertions and deletions (INDELs) were detected with the GATK's UnifiedGenotyper according to the Best Practice Variant Detection with the GATK v2 (http://www.broadinstitute.org/gsa/wiki/index.php/Best_Practice_Variant_Detection_with_the_GATK_v2). SNPs and INDELs were then filtered for the removal of low quality variants with GATK's VariantFiltrationWalker tools. We filtered out any SNPs matching the following criteria: greater than 10% of aligned reads included at the site have a mapping quality of 0 (MAPQ0), or overlaps INDELs, or DP>100 || MQ0>40 || SB>−0.10. We filtered out any INDELs matching the following criteria: greater than 10% of aligned reads included at the site have a mapping quality of 0 (MAPQ0), or SB>= −1.0, QUAL<10. Identified SNPs were annotated based on the dbSNP132 with ANNOVAR (Wang et al., 2010). Once the trio genotypes were determined, we extracted any iUPD genotypes that did not comply with the rules of Mendelian inheritance.

Filters were applied to exclude genomic regions in which false positive iUPD calls might be picked up. Since some genome regions are problematic for mapping and assembly, including regions of CNV in the each daughter, a putative iUPD call was not attempted in these regions (Altshuler et al., 2010; Conrad et al., 2011). We used the following filters: Simple Repeats, Segmental Duplications, CNV regions (Conrad et al., 2010; Kidd et al., 2008; McCarroll et al., 2008; Mills et al., 2011), and read depth (sites where at least one trio member has no mapped Illumina reads). BEDTools was used to confirm the intersections between putative iUPD genotypes and above-mentioned regions (Quinlan and Hall, 2010). Other annotations are based on The National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) and The University of California Santa Cruz (UCSC; http://genome.ucsc.edu/) databases. Finally, we required each genotype in a trio to have qualities GQ40 or greater for more efficient identification of the true iUPD genotypes.

2.6. Capillary sequencing

Validation experiments were performed on the DNA extracted from LCLs in each trio by a standard capillary sequencing approach. For CEU 1463 and YRI Y117 trios, primers were designed for 140 and 178 sites, respectively. We designed PCR primers using PrimerZ (http://genepipe.ngc.sinica.edu.tw/primerz/beginDesign.do) (Tsai et al., 2007) or Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) (Untergasser et al., 2007). Primers for each data set are provided in Supplementary Table 3.

2.7. Quantitative polymerase chain reaction (qPCR)

qPCR analysis was performed to measure the genomic copy number using a LightCycler 480 (Roche Diagnostics, Basel, Switzerland) and the Thunderbird SYBR qPCR Mix (Toyobo Co., Ltd.) according to the manufacturer's experimental protocol. Two sets of primers, zinc finger protein 80 (*ZNF80*) and G protein-coupled receptor 15 (*GPR15*) (D'haene et al., 2010), were used as references for quantification. Data analysis was performed with the second derivative maximum method of LightCycler 480 software (version 1.5.0.39) (Roche Diagnostics). qPCR amplification was carried out in triplicate. Primers for target regions were designed to surround the putative iUPD genotype by PrimerZ. Primers for each data set are provided in Supplementary Table 3.

3. Results

3.1. A whole chromosomal and segmental UPD analysis in 173 trios using SNP6.0 array

To investigate whole chromosome and segmental UPD in general populations using SNP6.0 array data, we examined the genotypes of the 173 trios that included 3 JPT trios in Nagasaki and 170 HapMap trios. Screening of UPD segments identified 46 putative segments (Table 1). A whole chromosomal UPD was not found in any chromosome except the Y chromosome in all samples tested. To rule out false segmental UPD due to CNVs and LTAs, we performed CNV analysis (Supplementary Table 4) and then cross-referenced with regions of putative segmental UPD in each trio (Table 1). As a result, we identified 24 CNVs, 21 LTAs (18 results from CNV and 3 possible copy number neutral LOH in the investigated parent's genome) (Supplementary Figs. 1–3) and 1 obvious segmental iUPD (Table 1). This one segmental iUPD indicated a paternal iUPD range from p-terminal to physical position 8,202,065 on chromosome 17p13.3-13.1 (about 8.2 mega bases (Mb)) in NA19918 (HapMap ASW FID 2431) (Fig. 1).

3.2. Base calling and detection of iUPD genotype

We investigated the possibility of the shorter segmental iUPD being undetectable by SNP6.0 array in the human genome using sequence data with a high coverage by Illumina platform during the pilot phase 2 of the 1000 Genomes Project. For each of the trios, we called the genotype of the three genomes independently using the GATK framework. In the CEU trio (NA12878, NA12891 and NA12892), mapped sequence coverage of 31.9×, 30.3× and 25.6×, respectively, and 2.32, 2.33 and 2.31 gigabases (Gb) of accessible genome included 2.85, 2.85, 2.79 million SNPs. In the YRI trio (NA19240, NA19239 and NA19238), mapped sequence coverage of 33.4×, 24.5× and 20.6×, respectively, and 2.36, 2.30 and 2.23 Gb of accessible genome included 3.60, 3.40 and 3.10 million SNPs. The accessible genome per CEU and YRI trio set were 2.24 Gb and 2.14 Gb, respectively. Statistics for each data set are provided in Table 2. Of these accessible genomes in each trio set, in the CEU 1463 and YRI Y117 trios, 1,094 and 1,474 putative iUPD genotypes were selected, respectively (Fig. 2). To exclude false iUPD genotypes, we filtered out the putative iUPD genotypes overlapping with regions of the simple repeats and segmental duplications and previously reported CNVs in the trio's daughter (Supplementary Tables 5 and 6). As a result, we identified 502 and 965 putative iUPD genotypes in the CEU 1463 and YRI Y117 trio, respectively (Fig. 2).

Table 1
Summary of putative segmental UPD segments in 173 trios detected by SNP6.0 array data analysis. Chr, chromosome; CNV, copy number variant; ND, not detectable; LTA, loss of transmitted allele; iUPD, uniparental isodisomy; LOH, loss of herozygosity.

Chr	Start position	End position	Population	HapMap FID	Sex	UPD type	Mother CNV	Father CNV	Child CNV	Result	UPD probe number	Length (bp)
11	51,078,178	51,359,581	ASW	2368	XY	Paternal	CNV	ND	ND	LTA	7	281,404
17	6,689	8,202,065	ASW	2431	XY	Paternal	ND	ND	ND	Paternal iUPD	301	8,195,377
7	119,133,278	119,393,868	ASW	2427	XY	Maternal	ND	CNV	CNV	CNV	8	260,591
6	137,300,451	143,369,018	CEU	1423	XX	Paternal	CNV	ND	ND	LTA	122	6,068,568
5	107,513,060	107,716,753	CEU	1350	XY	Paternal	ND	ND	CNV	CNV	15	203,694
8	14,677,944	15,701,490	CEU	1375	XX	Paternal	CNV	ND	CNV	CNV	64	1,023,547
7	88,496,196	88,887,792	CEU	1330	XY	Paternal	CNV	ND	CNV	CNV	15	391,597
2	85,751,279	88,861,509	CEU	1330	XX	Paternal	CNV	ND	ND	LTA	34	3,110,231
12	129,502	131,942,726	CEU	1444	XY	Paternal	ND	ND	CNV	CNV	825	131,813,225
11	81,131,219	81,387,538	CEU	1447	XX	Paternal	CNV	ND	CNV	CNV	11	256,320
22	20,825,481	21,201,922	CEU	1459	XY	Paternal	CNV	ND	ND	LTA	7	376,442
X	28,498,460	31,437,190	CEU	1463	XX	Paternal	CNV	CNV	ND	LTA	34	2,938,731
4	118,785,685	119,509,766	CEU	1340	XX	Maternal	ND	CNV	ND	LTA	16	724,082
12	33,468,716	34,188,071	CEU	1345	XX	Maternal	ND	CNV	ND	LTA	9	719,356
22	20,784,680	21,191,527	CEU	1420	XX	Maternal	ND	CNV	ND	CNV	11	406,848
X	276,282	154,127,693	CEU	1349	XX	Maternal	ND	ND	CNV	CNV	2,170	153,851,412
15	21,205,648	45,731,444	CEU	1377	XY	Maternal	ND	CNV	ND	LTA	85	24,525,797
18	65,224,346	76,085,336	CEU	1328	XX	Maternal	ND	CNV	ND	LTA	369	10,860,991
22	20,927,130	21,243,931	CEU	1330	XY	Maternal	ND	CNV	ND	LTA	8	316,802
1	206,304,300	246,785,226	CEU	1330	XX	Maternal	ND	CNV	ND	LTA	144	40,480,927
17	69,586,313	70,111,013	CEU	13281	XY	Maternal	CNV	ND	CNV	CNV	12	524,701
X	140,182,100	140,575,068	CEU	1354	XX	Maternal	ND	CNV	CNV	CNV	18	392,969
22	20,718,086	21,107,920	CEU	1358	XY	Maternal	ND	ND	CNV	CNV	15	389,835
1	144,979,429	145,700,719	CEU	1459	XX	Maternal	ND	ND	ND	LTA (putative LOH in father)	51	721,291
1	236,789,304	246,590,204	CEU	1459	XX	Maternal	ND	ND	ND	LTA (putative LOH in father)	396	9,800,901
11	114,231,222	134,235,117	CEU	1463	XY	Maternal	ND	ND	ND	LTA (putative LOH in father)	543	20,003,896
13	82,217,462	83,042,185	MXL	M019	XX	Maternal	ND	ND	CNV	CNV	9	824,724
6	140,718,454	141,182,824	MXL	M027	XX	Maternal	ND	CNV	CNV	CNV	9	464,371
X	4,726,561	145,198,977	MKK	2596	XX	Paternal	CNV	CNV	ND	LTA	611	140,472,417
22	20,909,341	21,181,447	MKK	2699	XY	Paternal	ND	ND	CNV	CNV	7	272,107
11	55,060,441	55,440,561	MKK	2588	XX	Maternal	ND	CNV	CNV	CNV	8	380,121
5	110,517,276	110,787,436	MKK	2634	XY	Maternal	ND	CNV	CNV	CNV	17	270,161
9	11,947,750	12,155,758	MKK	2634	XY	Maternal	ND	CNV	CNV	CNV	11	208,009
22	24,042,173	24,292,988	MKK	2634	XY	Maternal	ND	CNV	ND	LTA	9	250,816
1	245,373,155	247,137,334	YRI	Y014	XX	Paternal	ND	ND	CNV	CNV	66	1,764,180
3	136,397,878	137,151,871	YRI	Y014	XX	Paternal	CNV	ND	ND	LTA	10	753,994
11	329,969	26,983,000	YRI	Y014	XX	Paternal	ND	ND	CNV	CNV	263	26,653,032
X	2,386,344	25,622,488	YRI	Y014	XX	Paternal	CNV	ND	ND	LTA	81	23,236,145
19	22,821,274	23,413,380	YRI	Y074	XY	Paternal	ND	ND	CNV	CNV	7	592,107
7	119,175,698	119,393,868	YRI	Y038	XX	Paternal	CNV	ND	CNV	CNV	8	218,171
12	73,201,200	91,388,277	YRI	Y112	XY	Paternal	CNV	ND	ND	LTA	70	18,187,078
1	22,392,010	28,325,476	YRI	Y003	XY	Maternal	ND	ND	CNV	CNV	97	5,933,467
15	20,318,185	20,773,725	YRI	Y009	XY	Maternal	ND	CNV	CNV	CNV	11	455,541
22	24,012,780	24,238,616	YRI	Y071	XY	Maternal	ND	CNV	CNV	CNV	10	225,837
13	18,759,817	19,002,511	YRI	Y039	XY	Maternal	ND	CNV	ND	LTA	9	242,695
2	151,462,078	153,347,758	YRI	Y048	XY	Maternal	ND	CNV	ND	LTA	36	1,885,681

3.3. GQ threshold and filtering for iUPD genotypes

Our approach was simple and would allow false iUPD candidates in the initial screening. Therefore, of 502 putative iUPD genotypes in the CEU 1463 trio, 100 candidate sites (300 genotypes in the trio) were selected at random, and validated by capillary sequencing on the LCLs DNA. We used this data to estimate the accuracy of the genotype and to determine the threshold quality more efficiently for identification of the true iUPD genotype. Of the 300 validated genotypes, the correct and incorrect genotypes were 189 (63%) and 111 (37%), respectively, and true iUPD genotype was not confirmed (Supplementary Table 7). For more efficient screening, we focused on genotype quality (GQ), encoded as a Phred quality and read depth (DP) at genotype position. The 300 genotypes validated had a mean GQ of 71.13 (from a minimum of 1.61 to a maximum of 99.00) and a mean DP of 31.79 (from a minimum of 8.00 to a maximum of 75.00), respectively. Studying the relationship between GQ and accuracy of the genotypes with GQ10 or more, the correct genotype rate was 64.9% (189/291), 72.5% (182/251) with GQ40 or more, 91.0% (162/178) with GQ60 or more and 99.3% (150/151) with GQ80 or more. Thus, a higher GQ showed a higher reliability (Supplementary Fig. 4A). In contrast, increasing DP simply did not have much power to remove incorrect genotypes (Supplementary Fig. 4B). Furthermore, the majority of false positives for putative iUPD genotypes arose from an inaccuracy of genotyping in any one of the trio (81.1%, 90/111). Therefore, we required all genotypes in the trio with GQ40 or greater for identification of the true iUPD genotype. After filtering with a threshold GQ40, we identified 100 and 178 putative iUPD genotypes in the CEU 1463 and YRI Y117 trio, respectively (Fig. 2, Supplementary Tables 8 and 9).

3.4. Validation of the putative iUPD genotypes by capillary sequencing and qPCR

We attempted to validate these candidates by capillary sequencing. Of these, only 1 putative iUPD genotype (Validation ID C1383 and Y3887, respectively), in the CEU 1463 and YRI Y117 trio was confirmed as a true iUPD genotype (Fig. 2, Supplementary Tables 8 and 9, Supplementary Fig. 5A and B). Although iUPD candidates were not present in the known CNVs regions in the daughter, qPCR analysis with DNA from each trio was performed with primers C1383 and Y3887 to confirm the copy number on the putative iUPD loci. The results revealed a deletion

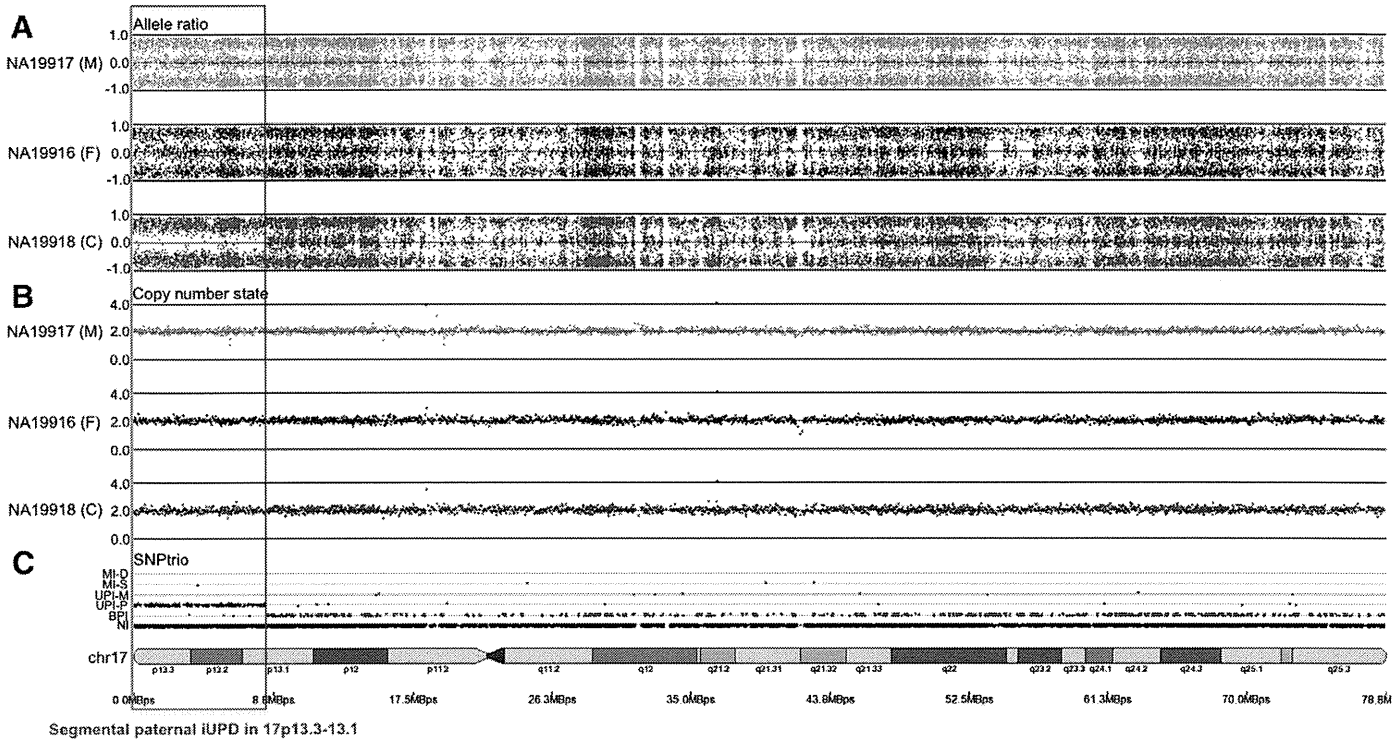

Fig. 1. Segmental paternal iUPD in HapMap ASW sample (NA19918). SNP6.0 data analyzed with PartekGS software shows the plots for the allele ratio, copy number state, and inheritance pattern by SNPtrio on chromosome 17 in HapMap ASW trio (FID 2431) (M, mother; F, father; C, child). (A) The allele ratio graph represents the genotypes for each individual single nucleotide polymorphism (SNP). Dots with a value of 1, −1, and 0 represent SNPs with AA, BB, and AB genotypes, respectively. (B) Plots represent chromosome copy number state (0.0–4.0). (C) SNPtrio displayed five classes of inheritance pattern. The five classes are 1) double Mendelian inconsistency (MI-D); 2) single Mendelian inconsistency (MI-S); 3) maternal uniparental inheritance (UPI-M); 4) paternal uniparental inheritance (UPI-P); 5) biparental inheritance (BPI). NI indicates not informative. The BPI plots represent the biparental inheritance SNPs, in which the parents have AA and BB calls and the child has an AB call. A red box indicates the segmental paternal iUPD locus.

on the C1383 locus in the daughter (NA12878) and mother (NA12892). Similarly, the results revealed a deletion on the Y3887 locus in the daughter (NA19240) and father (NA19239) (Supplementary Fig. 5C). In our investigation, we could not identify shorter segmental iUPD in the daughters from the two trios (Fig. 2).

3.5. Genes in identified segmental UPD regions in normal individuals

Finally, we identified one segmental paternal iUPD on 17p13.3-13.1 from 173 individuals. This segmental UPD region was included in the 233 RefSeq genes (Supplementary Table 10), but which are not "imprinted genes." According to the conventional concept, UPD has no practical impact on phenotypes with the exception of the disruption of imprinting and homozygosity for recessive mutations.

4. Discussion

At any stage of the life cycle, from gamete formation to fetal postnatal life, exposure to genotoxic stress may affect the genomic integrity and fate of the organism (Jaroudi and SenGupta, 2006; Vinson and Hales, 2002). In undifferentiated cells, such as the embryo and progenitor cells, mutations are propagated to multiple differentiated cell types

Table 2
Summary of alignment and base calling in two trios. AC+X: autosomal chromosome and X chromosome (exclude gap) = total length 2,706,959,439 bases (about 2.71 Gb).

Family	CEU 1463			YRI Y117		
Sample	NA12878	NA12891	NA12892	NA19240	NA19239	NA19238
Relation	Daughter	Father	Mother	Daughter	Father	Mother
Total bases (Gb)	102.24 Gb	100.89 Gb	85.68 Gb	108.25 Gb	84.03 Gb	71.56 Gb
Mapped bases (Gb)	99.63 Gb	97.25 Gb	80.2 Gb	104.25 Gb	79.94 Gb	65.24 Gb
Total reads	2,507,012,490	2,264,396,064	2,051,935,811	2,738,304,812	2,296,647,842	1,971,737,379
Mapped reads	2,443,207,477	2,189,660,230	1,952,966,402	2,632,175,898	2,184,252,515	1,796,606,841
Mean mapped depth	31.9	30.3	25.6	33.4	24.5	20.6
Accessible genome (Gb)	2.32 Gb	2.33 Gb	2.31 Gb	2.36 Gb	2.30 Gb	2.23 Gb
Accessible genome (% of AC+X)	85.61 (%)	85.98 (%)	85.24 (%)	87.08 (%)	84.87 (%)	82.29 (%)
Accessible genome with trio (Gb)	2.24 Gb			2.14 Gb		
Accessible genome with trio (Gb) (% of AC+X)	82.66 (%)			78.97 (%)		
SNPs (N)	2,854,439	2,846,437	2,785,908	3,602,569	3,395,713	3,090,355
SNPs in dbSNP132 (N)	2,838,282	2,831,464	2,773,304	3,576,164	3,371,626	3,070,416
SNPs in dbSNP132 (%)	99.43 (%)	99.47 (%)	99.55 (%)	99.27 (%)	99.29 (%)	99.35 (%)