noted, however, that the Tridimensional Personality Questionnaire has not been as robust as other measures (i.e., NEO personality inventory) in identifying genotype effects on fear and anxiety (*3*) and that the use of more robust measures in the current study may have revealed group differences in these behaviors.

34. N. Logothetis, J. Pauls, M. Augath, T. Trinath, A. Oeltermann, *Nature* **412**, 150 (2001).
35. D. G. Rainnie, *J. Neurophysiol.* **82**, 69 (1999).
36. D. Julius, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15153 (1998).
37. C. Gross *et al.*, *Nature* **416**, 396 (2002).
38. P. D. Hrdina, E. Demeter, T. B. Vu, P. Sotonyi, M. Palkovits, *Brain Res.* **614**, 37 (1993).
39. M. F. Egan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6917 (2001).
40. J. H. Callicott *et al.*, *Cereb. Cortex* **9**, 20 (1999).
41. C. R. Cloninger, *Psychiatric Dev.* **3**, 167 (1986).
42. P. Ekman, W. V. Friesen, *Pictures of Facial Affect* (Consulting Psychologists Press, Palo Alto, CA, 1976).
43. We thank S. Das, S. Lee, E. O'Hare, W. G. Smith, and R. Vakkalanka for technical assistance.

Supporting Online Material
www.sciencemag.org/cgi/content/full/297/5580/400/DC1
Materials and Methods

13 March 2002; accepted 13 June 2002

Enhanced CpG Mutability and Tumorigenesis in MBD4-Deficient Mice

Catherine B. Millar,[1]* Jacky Guy, [1]* Owen J. Sansom,[2]* Jim Selfridge,[1] Eilidh MacDougall,[1] Brian Hendrich,[1] Peter D. Keightley,[3] Stefan M. Bishop,[2] Alan R. Clarke,[2] Adrian Bird[1]†

The mammalian protein MBD4 contains a methyl-CpG binding domain and can enzymatically remove thymine (T) or uracil (U) from a mismatched CpG site in vitro. These properties suggest that MBD4 might function in vivo to minimize the mutability of 5-methylcytosine by removing its deamination product from DNA. We tested this hypothesis by analyzing *Mbd4*$^{-/-}$ mice and found that the frequency of of C → T transitions at CpG sites was increased by a factor of three. On a cancer-susceptible *Apc*$^{Min/+}$ background, *Mbd4*$^{-/-}$ mice showed accelerated tumor formation with CpG → TpG mutations in the *Apc* gene. Thus MBD4 suppresses CpG mutability and tumorigenesis in vivo.

Deamination of 5-methylcytosine (m^5C) to T at CpG sites is probably the single most important cause of point mutations in humans, accounting for more than 20% of all base substitutions that give rise to genetic disease (*1*). Estimates based on the in vitro deamination rate of m^5C (*2*) suggest that approximately four m^5C residues deaminate per diploid genome per day in the germ line. It is likely that many of the resulting mismatches are repaired, but appropriate repair must take into account that the incorrect base at the T:G mismatches is invariably T, not G. Two mismatch-specific T glycosylases that accomplish this discrimination in vitro have been discovered. Both thymine DNA glycosylase [TDG; (*3*, *4*)] and MBD4 (*5*, *6*) can specifically remove T from a T-G mismatch within a CpG context without cleaving the DNA strand. This study concerns MBD4, which contains a COOH-terminal glycosylase domain (*7*, *8*) and an NH_2-terminal methyl-CpG binding domain [MBD; (*7*)]. The MBD binds to symmetrically methylated CpG sites but has a higher affinity in vitro for m^5CpG/TpG mismatches, which arise due to deamination at methyl-CpG (*5*). Neither TDG nor MBD4 has been tested for its effect on mutations in vivo.

To investigate whether MBD4 suppresses the mutability of methylated CpG sites in vivo we generated mice bearing a mutated *Mbd4* gene by means of targeted allele replacement in embryonic stem cells (*9*). The mutated allele contained a cassette within intron 1 containing a splice acceptor and polyadenylation signal. Mice homozygous for this allele were viable and fertile. Northern blot (*10*) and reverse transcriptase–polymerase chain reaction (RT-PCR) analyses confirmed that MBD4 expression was less than 0.1% of the wild-type level in tissues from homozygous-mutant mice (fig. S1). We refer to this allele as *Mbd4*$^-$ throughout this study.

To determine whether MBD4 deficiency leads to an increase in mutations at CpG sites, we crossed mice with the "Big Blue" reporter locus, comprising 40 head-to-tail copies of a recoverable lambda transgene (*11*), onto the *Mbd4*$^{-/-}$ background and measured in vivo mutation frequencies at the λ *cII* locus (*12*). Bisulfite sequencing (*9*) of the *cII* locus from both wild-type and *Mbd4*$^{-/-}$ mice showed that, on average, 95% of the CpG sites in the *cII* gene are methylated (Fig. 1A). The frequency of *cII* mutations in liver and spleen of 105-day-old mice and in spleens of 183-day-old mice was determined by plating packaged genomic DNA (*9*, *12*). The total mutant frequency in *Mbd4*$^{-/-}$ mice (6.8×10^{-5}) was significantly higher than in wild-type animals (3.2×10^{-5}) by a maximum likelihood ratio test [$P < 0.0001$; Fig. 1B; Table 1; (*9*)]. This difference was reduced, but remained statistically significant ($P < 0.05$), when a correction for the number of independent mutations was applied [Fig. 1B; (*9*)]. The true mutation frequency is likely to lie between the uncorrected and corrected values, as correction may underestimate the real number of independent mutational events.

A striking difference between wild-type and *Mbd4*$^{-/-}$ mice emerged when the spectrum of mutations was examined. The most abundant mutational change in both wild-type and mutant mice involved CG → TA transitions, but this category was more frequent in *Mbd4*$^{-/-}$ mice than in *Mbd4*$^{+/+}$

Table 1. Analysis of mutation frequencies at the *cII* locus of a bacteriophage λ transgene in *Mbd4*$^{-/-}$ and *Mbd4*$^{+/+}$ mice. Numbers of "mutants" were derived from raw counts of mutant plaques. Numbers of "mutations" were deduced by correcting mutant numbers for likely clonal expansion of a single mutant cell (*9*). "Mutant" and "mutation" frequencies reflect the uncorrected and corrected data, respectively.

Genotype (*n*)	Age (days)/ tissue	Plaques screened (*n*)	Mutants (*n*)	Mutant frequency ($\times 10^{-5}$)	Individual mutations (*n*)	Mutation frequency ($\times 10^{-5}$)
Mbd4$^{+/+}$ (3)	105/liver	656,165	15	2.29	15	2.29
Mbd4$^{-/-}$ (3)	105/liver	991,900	39	3.93	32	3.23
Mbd4$^{+/+}$ (4)	105/spleen	1,340,250	45	3.36	38	2.84
Mbd4$^{-/-}$ (4)	105/ spleen	1,201,833	88	7.32	39	3.25
Mbd4$^{+/+}$ (1)	183/ spleen	200,500	11	5.49	11	5.49
Mbd4$^{-/-}$ (1)	183/ spleen	227,433	38	16.71	27	11.87

[1]Wellcome Trust Centre for Cell Biology, The King's Buildings, Edinburgh University, Edinburgh EH9 3JR, UK. [2]Cardiff School of Biosciences, Cardiff University, Post Office Box 911, Cardiff CF10 3US, UK. [3]Institute of Cell Animal and Population Biology, The King's Buildings, Edinburgh University, Edinburgh EH9 3JT, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: a.bird@ed.ac.uk

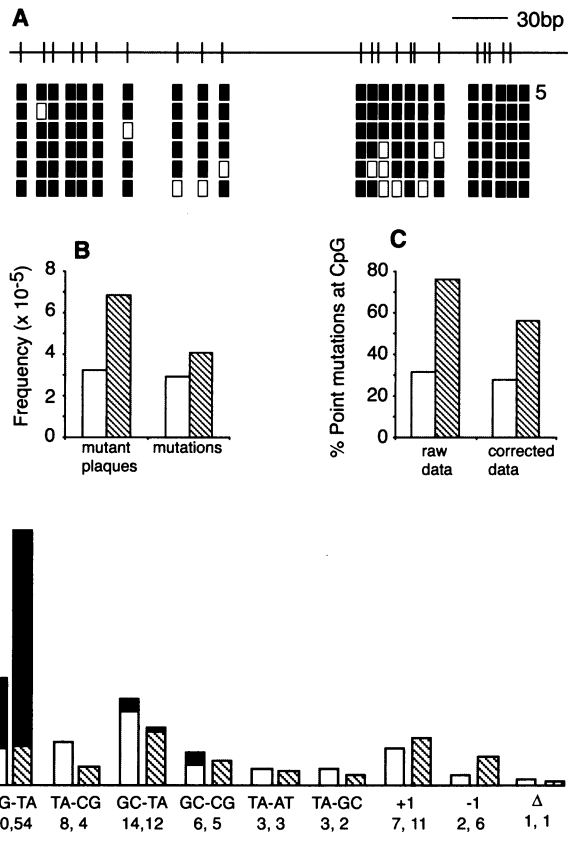

Fig. 1. MBD4-deficiency increases the frequency of CG → TA mutations at CpG sites. (**A**) The positions of the 22 CpG sites in the *cII* gene are indicated by vertical lines. The methylation status of 10 clones from $Mbd4^{-/-}$ spleen DNA is shown below, with filled boxes denoting methylated sites and unfilled boxes denoting unmethylated sites. The number of sequenced clones with identical methylation patterns is indicated on the right, where it exceeds one. A similar level of methylation was observed in $Mbd4^{+/+}$ mice, and in liver tissue from both genotypes (*21*). (**B**) $Mbd4^{-/-}$ mice (hatched bars) have a significantly higher frequency of mutant plaques ($P < 0.0001$) and of individual mutations ($P < 0.05$) than wild-type mice (white bars). These data sum all time points and tissues. (**C**) Increased percentage of all point mutations that are caused by CG → TA mutations at CpG sites in $Mbd4^{-/-}$ mice, showing both raw and corrected data. (**D**) Mutational spectra on the $Mbd4^{+/+}$ and $Mbd4^{-/-}$ backgrounds (corrected data). Mutations at CpG sites in all categories are shaded in black, and the total number of individual mutations found in each category for each genotype is indicated below the appropriate bar.

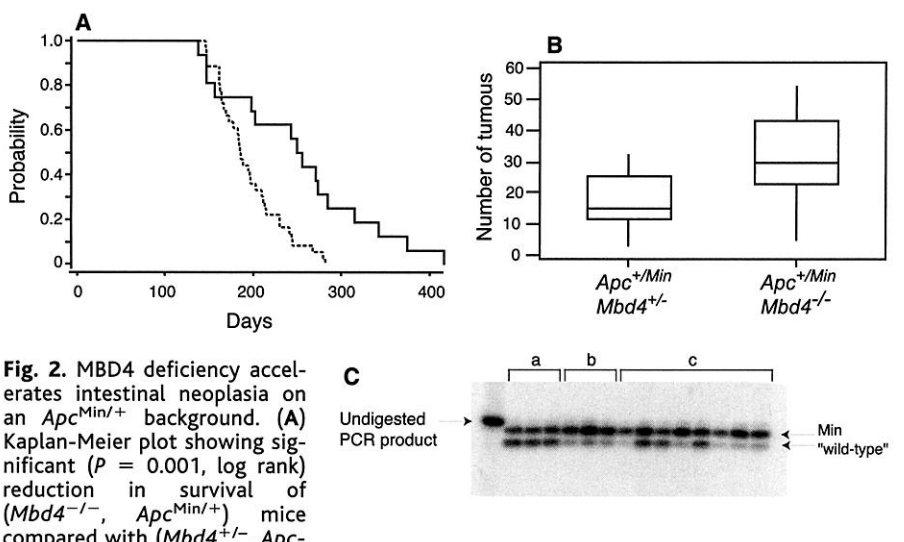

Fig. 2. MBD4 deficiency accelerates intestinal neoplasia on an $Apc^{Min/+}$ background. (**A**) Kaplan-Meier plot showing significant ($P = 0.001$, log rank) reduction in survival of ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice compared with ($Mbd4^{+/-}$, $Apc^{Min/+}$) mice. Solid black line ($Mbd4^{+/-}$, $Apc^{Min/+}$) mice ($n = 17$); broken line ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice ($n = 36$). (**B**) Tumor burden in ($Mbd4^{+/-}$, $Apc^{Min/+}$) and ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice at necroscopy. The median value for each genotype is indicated by a horizontal line in the boxplots. There was a small but statistically significant increase (Mann-Whitney U test, $P = 0.01$) in tumor number in the ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice. (**C**) *Apc* mutations in tumors of the large intestine in ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice. The upper band (Min) is the Hind III-resistant Apc^{Min} PCR product, whereas the lower band ("wild type") is the initially wild-type *Apc* allele, which can be cut by Hind III. Lanes a, nontumor control samples from intestines of $Mbd4^{+/-}$ mice; lanes b, tumors from $Mbd4^{+/-}$ mice showed 100% ($n = 20$) LOH; lanes c, tumors from $Mbd4^{-/-}$ mice showed 62% ($n = 39$) LOH, which is significantly less than $Mbd4^{+/-}$ mice ($P < 0.001$, chi-square test). The presence of a weak Apc^{+} band in tumors that have lost this allele is due to unavoidable contamination with nontumor tissue (notably, blood).

controls (Fig. 1D; corrected data were analyzed). The difference between wild-type and $Mbd4^{-/-}$ mice could be specifically attributed to a highly significant 3.3-fold increase in CG → TA transitions at CpG sites ($P < 0.0001$; Fig. 1D). Other mutation categories, including CG → TA transitions at non-CpG sites, were not significantly altered by the absence of MBD4 ($P > 0.1$). The effect on the overall mutation spectrum was apparent from the fraction of all point mutations at CpG (Fig. 1C). In wild-type mice, 28 to 31% (this study) and 26.9% (*13*) of point mutations were found to be at CpG, which is similar to the fraction deduced from human genetic diseases (23%) (*1*). In $Mbd4^{-/-}$ mice, however, the fraction rose to between 56% (corrected) and 76% (uncorrected).

We next asked whether the increased mutability of $Mbd4^{-/-}$ mice led to an increased incidence of cancer in an in vivo model of intestinal tumorigenesis. MBD4 is mutated in 26 to 43% of human colorectal tumors that show microsatellite instability (*14*, *15*), suggesting that loss of MBD4 may contribute to genome instability (*16*). To test whether loss of MBD4 increases tumorigenesis, $Mbd4^{-/-}$ mice were made heterozygous for the Min allele of the adenomatous polyposis coli gene (Apc^{Min}), which predisposes mice to the development of spontaneous intestinal neoplasia (*17*). Littermates were divided into ($Mbd4^{-/-}$, $Apc^{Min/+}$) and ($Mbd4^{+/-}$, $Apc^{Min/+}$) cohorts and killed when they became symptomatic for intestinal neoplasia (*9*). $Mbd4^{-/-}$ mice showed markedly reduced survival compared with $Mbd4^{+/-}$ controls ($P = 0.001$, Fig. 2A). Moreover, whole-mount analysis showed a small but significant ($P = 0.01$) increase in intestinal adenoma burden at death in ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice (median = 28 tumors) compared with ($Mbd4^{+/-}$, $Apc^{Min/+}$) mice (median = 14 tumors) (Fig. 2B). Histological analysis revealed no morphological differences between the $Mbd4^{+/-}$ and $Mbd4^{-/-}$ tumors.

To determine whether the loss of repair activity at CpG sites contributed to the accelerated tumor formation, we analyzed the nature of mutations causing loss of function of the wild-type *Apc* allele in intestinal tumors. On an $Apc^{Min/+}$ background in the absence of mutagenic challenge, the overwhelming cause of spontaneous tumors is complete deletion (loss of heterozygosity, or LOH) of the Apc^{+} allele (*18*). In agreement with this, we found that all 15 $Mbd4^{+/-}$ tumors in the large intestine showed loss of the Apc^{+} allele. In contrast, 15 of 39 tumors (38%) from ($Mbd4^{-/-}$, $Apc^{Min/+}$) mice retained both *Apc* alleles (Fig. 2C). Sequencing of part of the *Apc* gene (*9*) in the latter tumors revealed eight mutations. Five mutations were CpG → TpG transitions, of which four would cause

truncation of the APC protein (table S1). The complete absence of equivalent mutations in control mice indicates that the suppression of intestinal tumors by *Mbd4* is due at least in part to suppression of CpG → TpG transitions at the endogenous *Apc* gene.

Our findings clearly implicate MBD4 in the repair of m^5C deamination at methylated CpG residues, but is it solely responsible for this repair? If we assume that the in vitro deamination rate of m^5C (5.8×10^{-13}/second) (*2*) applies in vivo, we predict a C → T transition frequency of 1.25×10^{-4} at CpG sites in the *cII* locus in the absence of repair. The observed mutation frequency in wild-type mice is 4% of this calculated value, suggesting that repair of m^5C deamination is about 96% efficient. In $Mbd4^{-/-}$ mice, the CpG to TpG mutation frequency is 1.5×10^{-5}, suggesting that repair efficiency has fallen to 88% but remains moderately effective. This approximate calculation indicates that repair of deaminated m^5C is shared with factors other than MBD4, for example, TDG or equivalent activities. In theory, two independent repair activities that were each ~80% efficient could give a combined efficiency of ~96%.

Why is m^5C a mutational hotspot in spite of the presence of at least one repair mechanism? One plausible explanation is that the existence of close to 10^9 T residues in the mouse genome, each one losing its base pairing with A many times per second (*19*), renders enzymes that remove unpaired T potentially mutagenic. With respect to C deamination, this problem is thought to have been solved by replacing U with T in the ancestral RNA-derived genome, so that the deamination product of C became easily distinguishable from a normal DNA base (*20*). All known U:G mismatch glycosylases are inert against T:G mismatches, despite the structural similarity between T and U, perhaps because of the danger of T removal. By methylating C, however, the problem is recreated, as deamination once more generates a normal DNA base whose excision may be problematic. Inefficient correction of the resulting T:G mismatches may represent a compromise between the benefits of repair and the damage that could arise by removal of legitimate T residues.

In conclusion, we have shown that murine MBD4 suppresses CpG → TpG mutations in both a bacteriophage-based transgene and at the endogenous *Apc* gene locus. These findings suggest that human MBD4 plays a similarly important role in reducing inherited disease and cancer.

References and Notes

1. M. Krawczak, E. V. Ball, D. N. Cooper, *Am. J. Hum. Genet.* **63**, 474 (1998).
2. J.-C. Shen, W. M. I. Rideout, P. A. Jones, *Nucleic Acids Res.* **22**, 972 (1994).
3. K. Wiebauer, J. Jiricny, *Nature* **339**, 234 (1989).
4. P. Neddermann, J. Jiricny, *J. Biol. Chem.* **268**, 21218 (1993).
5. B. Hendrich, U. Hardeland, H.-H. Ng, J. Jiricny, A. Bird, *Nature* **401**, 301 (1999).
6. F. Petronzelli *et al.*, *J. Biol. Chem.* **275**, 32422 (2000).
7. B. Hendrich, A. Bird, *Mol. Cell. Biol.* **18**, 6538 (1998).
8. F. Petronzelli *et al.*, *J. Cell. Physiol.* **185**, 473 (2000).
9. Materials and methods are available as supporting material on *Science* Online.
10. J. Guy, unpublished observations.
11. S. W. Kohler *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7958 (1991).
12. J. L. Jakubczak *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9073 (1996).
13. P. R. Harbach, D. M. Zimmer, A. L. Filipunas, W. B. Mattes, C. S. Aaron, *Environ. Mol. Mutagen.* **33**, 132 (1999).
14. A. Riccio *et al.*, *Nature Genet.* **23**, 266 (1999).
15. S. Bader, M. Walker, D. Harrison, *Br. J. Cancer* **83**, 1646 (2000).
16. C. Lengauer, K. W. Kinzler, B. Vogelstein, *Nature* **386**, 623 (1997).
17. L. K. Su *et al.*, *Science* **256**, 668 (1992).
18. C. Luongo, A. R. Moser, S. Gledhill, W. F. Dove, *Cancer Res.* **54**, 5947 (1994).
19. J. L. Leroy, M. Kochoyan, T. Huynh-Dinh, M. Gueron, *J. Mol. Biol.* **200**, 223 (1988).
20. A. Poole, D. Penny, B. M. Sjoberg, *Nature Rev. Mol. Cell. Biol.* **2**, 147 (2001).
21. C. B. Millar, unpublished observation.

22. We thank C. Abbott for advice, D. Macleod for comments on the manuscript, H. Barr for DNA sequencing, N. Hill and the staff of the Anne Walker Building for animal husbandry. Supported by the Cancer Research Campaign and by a Programme Grant from the Wellcome Trust. C.B.M. holds a Wellcome Trust 4-year Ph.D. Studentship. A.R.C. is a Royal Society University Research Fellow.

Supporting Online Material

www.sciencemag.org/cgi/content/full/297/5580/403/DC1
Materials and Methods
Fig. S1
Table S1

29 April 2002; accepted 18 June 2002

A Role for Peroxisomes in Photomorphogenesis and Development of *Arabidopsis*

Jianping Hu,[1,2] Maria Aguirre,[1,2] Charles Peto,[3] José Alonso,[1] Joseph Ecker,[1] Joanne Chory[1,2]*

The nuclear protein DET1 is a central repressor of photomorphogenesis in plants. We have identified the molecular lesion in *ted3*, a mutation that dominantly suppresses the phenotypes of *det1-1*. *TED3* encodes a peroxisomal protein (AtPex2p) essential for *Arabidopsis* growth. Developmental defects and the abnormal expression of many genes in *det1* are rescued by *ted3*. *ted3* also partially suppresses another pleiotropic de-etiolated mutant *cop1*. Thus, peroxisomes, whose functions are still largely unexplored, play a key role in a photomorphogenetic pathway negatively regulated by the DET1 and COP proteins.

Plants use sophisticated signal transduction systems to sense and respond to environmental variation. The light signaling network, for example, consists of a complex web of interactions between multiple photoreceptors, early signaling factors, and central integrators to control the expression of hundreds of responsive genes (*1*). The *Arabidopsis* DET1, COP, and FUS proteins are proposed to be global repressors of photomorphogenesis, because their loss-of-function mutants develop as light-grown plants in the absence of light (*2*). Dark-grown *det1* mutants have short hypocotyls, opened cotyledons, and developed chloroplasts and ectopically express many light-regulated genes (*3*). They also exhibit stress symptoms, such as anthocyanin accumulation and aberrant expression of stress-related genes. Light-grown *det1* plants exhibit general growth defects, such as small, pale-green leaves and seedling lethality in strong alleles (*3*). *DET1* encodes a 62-kD nuclear protein that regulates gene expression (*4*), yet its precise mechanism of action is still unknown.

To better understand how DET1 exerts its functions on photomorphogenesis and development, we identified several extragenic suppressors of *det1-1*, an intermediate-strength allele that is impaired in the splicing of intron 1 but still produces ~2% of wild-type mRNA, and named them *ted* mutants. *ted3* is a dominant suppressor of *det1* phenotypes that does not correct the splicing defect of *det1-1* (*5*).

We fine-mapped the *TED3* gene to a ~130-kb region at the bottom of chromosome 1 (Fig. 1A). Sequencing of selected open reading frames (ORFs) in this interval in *ted3* revealed a mutation in an ORF with 333 amino acids (GenBank accession number AAG52254). Sequence comparison between the genomic fragment and the corresponding expressed sequence tag (EST) clone (N96573) suggested that this gene has eight exons (Fig. 1A) and encodes a 38-kD protein.

[1]Plant Biology Laboratory, [2]Howard Hughes Medical Institute, [3]Laboratory of Neuronal Structure and Function, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA.

*To whom correspondence should be addressed. E-mail: chory@salk.edu