LETTER

doi:10.1038/nature09672

Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability

Daniel Cortázar[1]*, Christophe Kunz[1]*, Jim Selfridge[2], Teresa Lettieri[3]†, Yusuke Saito[1], Eilidh MacDougall[2], Annika Wirz[1], David Schuermann[1], Angelika L. Jacobs[1], Fredy Siegrist[4], Roland Steinacher[1]†, Josef Jiricny[3], Adrian Bird[2] & Primo Schär[1]

Thymine DNA glycosylase (TDG) is a member of the uracil DNA glycosylase (UDG) superfamily of DNA repair enzymes. Owing to its ability to excise thymine when mispaired with guanine, it was proposed to act against the mutability of 5-methylcytosine (5-mC) deamination in mammalian DNA[1]. However, TDG was also found to interact with transcription factors[2,3], histone acetyltransferases[4] and *de novo* DNA methyltransferases[5,6], and it has been associated with DNA demethylation in gene promoters following activation of transcription[7–9], altogether implicating an engagement in gene regulation rather than DNA repair. Here we use a mouse genetic approach to determine the biological function of this multifaceted DNA repair enzyme. We find that, unlike other DNA glycosylases, TDG is essential for embryonic development, and that this phenotype is associated with epigenetic aberrations affecting the expression of developmental genes. Fibroblasts derived from *Tdg* null embryos (mouse embryonic fibroblasts, MEFs) show impaired gene regulation, coincident with imbalanced histone modification and CpG methylation at promoters of affected genes. TDG associates with the promoters of such genes both in fibroblasts and in embryonic stem cells (ESCs), but epigenetic aberrations only appear upon cell lineage commitment. We show that TDG contributes to the maintenance of active and bivalent chromatin throughout cell differentiation, facilitating a proper assembly of chromatin-modifying complexes and initiating base excision repair to counter aberrant *de novo* methylation. We thus conclude that TDG-dependent DNA repair has evolved to provide epigenetic stability in lineage committed cells.

TDG is one of four enzymes with UDG activity in mammalian cells, but its biological function has remained enigmatic[10]. We thus set out to generate and phenotypically investigate a *Tdg* knockout mouse (Supplementary Fig. 1a–c). ESC clones carrying the targeted allele gave rise to healthy heterozygous *Tdg* knockout mice but attempts to generate homozygous null mutants failed, indicating that TDG-deficiency may cause embryonic lethality. This was unexpected, given the generally mild phenotype of other DNA glycosylase knockouts[11]. In timed matings, *Tdg* null embryos isolated up to embryonic day (E) 10.5 appeared alive and normal, whereas those isolated at E12.5 were dead, and none were detectable at E16.5 (Fig. 1a and Supplementary Fig. 1d). *Tdg* null embryos at E10.5 produced viable fibroblasts (MEFs) but only a third of E11.5 embryos did so, suggesting that by this stage most of them were dead. We thus concluded that lethality in *Tdg* null embryos occurs around E11.5. For the actual cause of lethality, closer examination of the *Tdg* null embryos at E10.5 indicated internal haemorrhage, and evidence for haemorrhagic necrosis (data not shown), but otherwise did not reveal an informative pathology.

We then explored the essential function of TDG in MEFs and ESCs, first addressing a potential DNA repair defect by classical genotoxicity and mutator analyses. The TDG status did not affect cell survival following ionizing radiation or H_2O_2 exposure, both of which induce DNA base lesions processed by TDG *in vitro*[10], nor did it affect mutation frequencies in a Big Blue transgenic mutation assay (Supplementary Fig. 2). We therefore concluded that the role of TDG in the repair of canonical base damage is minor and therefore unlikely to account for its essential function in mouse embryogenesis.

We next investigated a possible involvement of TDG in gene regulation by expression profiling of TDG-proficient and -deficient MEFs. To limit potential clonal biases, we compared the transcriptomes of early passages of litter-matched populations of SV40 immortalized MEFs. This identified 461 differentially transcribed genes ($P \leq 0.05$, fold change (FC) ≥ 1.5, Fig. 1b), comprising many transcription factors and, thus, likely reflecting both direct and indirect consequences of TDG loss. Global pathway analyses revealed gene networks associated with embryogenesis and development as being most significantly misregulated in the absence of TDG (Supplementary Fig. 3a). Four out of six target genes analysed showed TDG-dependent differential expression also in independently isolated primary MEFs (Supplementary Fig. 3b).

Considering its putative involvement in DNA demethylation[7–9], we next investigated a possible occurrence of aberrant promoter methylation in TDG-deficient cells. We examined the CpG islands in the promoters of *Hoxa10*, *Hoxd13*, *Sfrp2*, *Twist2* and *Rarb*, all of which were downregulated in TDG-deficient MEFs (Fig. 1b and Supplementary Fig. 3a). These genes are developmentally regulated by the polycomb repressive system[12] and their promoter CpG islands are unmethylated in most normal tissues but subject to aberrant *de novo* methylation in human cancers[13,14]. Na-bisulphite sequencing of the respective CpG islands revealed an increased occurrence of *de novo* methylation in the TDG-deficient MEFs (Fig. 1c and Supplementary Figs 4 and 5a). The patterns and frequency of these methylation events indicated that the loss of TDG generates hotspots of *de novo* methylation in certain gene promoters. We then used chromatin immunoprecipitation (ChIP) to examine a possible association of TDG with the promoters of these and additional differentially expressed genes. Compared with a randomly chosen intergenic sequence or the silent promoters of *Oct4* and *Tuba3*, DNA fragments surrounding the promoters of all genes examined were significantly enriched in the TDG precipitates (Fig. 1d). This indicated that TDG is targeted to specific gene promoters, possibly to protect them from acquiring aberrant CpG methylation and eventual epigenetic silencing. Consistently, further examination of the chromatin status revealed a general loss of activating (H3K4me2) and a concomitant increase of repressive histone marks (H3K27me3, H3K9me3) in TDG-deficient cells with promoter-specific patterns (Fig. 1e): a complete loss of H3K4 dimethylation was accompanied by a strong increase of H3K27 and/or H3K9 trimethylation at the *Hoxd13* and *Hoxa10* promoters; a partial reduction of H3K4me2 coincided with an enrichment of H3K27me3 but not H3K9me3 at

[1]Institute of Biochemistry and Genetics, Department of Biomedicine, University of Basel, 4048 Basel, Switzerland. [2]The Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK. [3]Institute of Molecular Cancer Research, University of Zürich, 8057 Zürich, Switzerland. [4]Pharmaceutical Research, Global Preclinical Safety, F. Hoffmann-La Roche Ltd., 4058 Basel, Switzerland.
†Present addresses: European Commission, Joint Research Centre, Institute for Environment and Sustainability, 21027 Ispra, Italy (T.L.); Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK (R.S.).
*These authors contributed equally to this work.

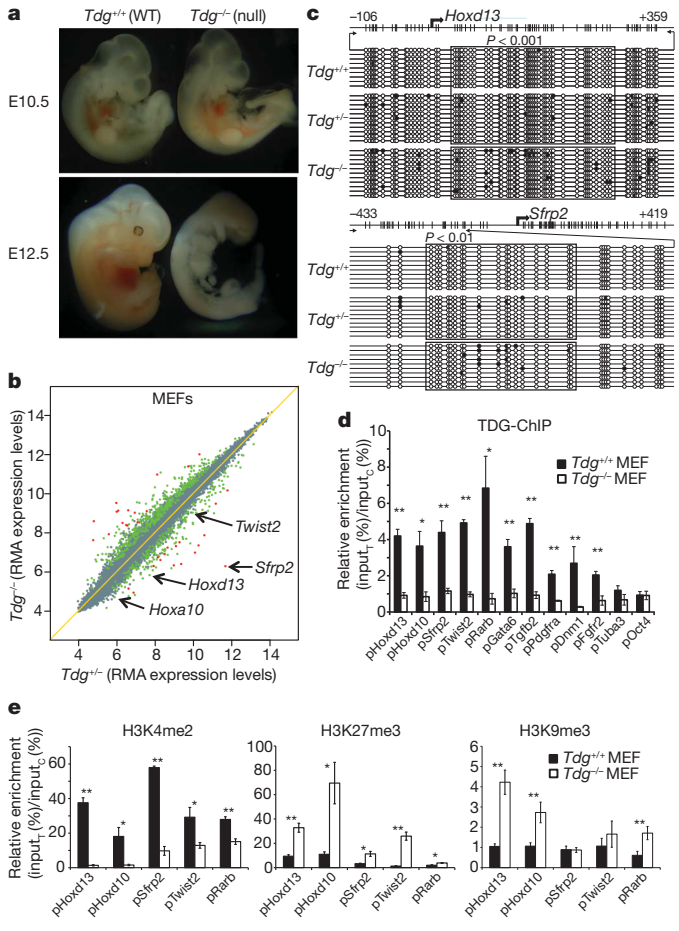

Figure 1 | Embryonic essential function of *Tdg* in epigenetic gene regulation. **a**, Whole mount images of typical examples of $Tdg^{+/+}$ and $Tdg^{-/-}$ littermate embryos taken at E10.5 and E12.5. **b**, Scatter plot comparing gene expression levels of matched $Tdg^{+/-}$ and $Tdg^{-/-}$ MEFs. Differentially expressed genes at $P < 0.05$ and $P < 0.01$ are indicated by green and red dots, respectively, and examples of developmental genes affected are denoted. **c**, Na-bisulphite sequencing of the *Hoxd13* and *Sfrp2* promoters in $Tdg^{+/+}$, $Tdg^{+/-}$ and $Tdg^{-/-}$ MEFs. White and black circles indicate unmethylated and methylated CpGs, respectively. *P* values indicate statistical difference of methylation frequencies as determined by contingency tables and χ^2 test. **d**, ChIP–quantitative PCR (qPCR) analysis of TDG association with the promoters of the genes indicated in chromatin from $Tdg^{+/+}$ and $Tdg^{-/-}$ MEFs. Shown are relative enrichments of TDG at these promoters normalized to a randomly chosen intergenic control region (means ± s.e.m.; $n \geq 3$; $^*P < 0.05$; $^{**}P < 0.01$, unpaired Student's *t*-test). **e**, ChIP–qPCR analyses in $Tdg^{+/+}$ and $Tdg^{-/-}$ MEFs to assess the presence of activating (H3K4me2) and repressive (H3K9me3, H3K27me3) histone modifications at the promoter regions indicated. Shown are enrichments relative to appropriate negative controls: intracisternal A-particle (*Iap*) transposon for active chromatin marks and the *Hprt* promoter for silencing marks (means ± s.e.m.; $n \geq 3$; $^*P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test). Subscript $_T$, target region; subscript $_C$, control region.

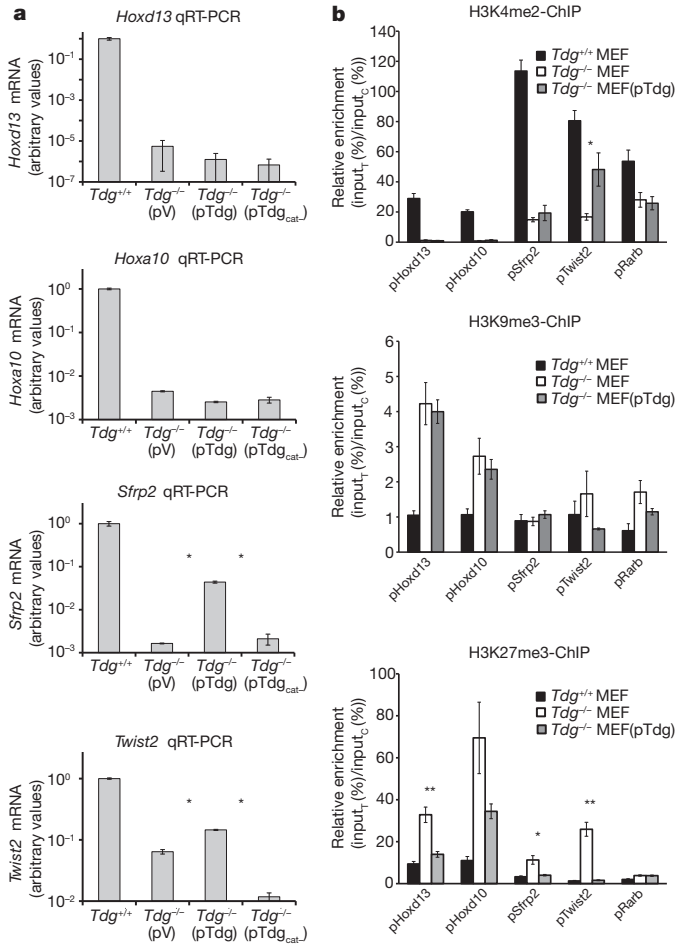

Figure 2 | Complementation of the loss of gene expression depends on the chromatin state of the promoter. **a**, *Hoxd13*, *Hoxa10*, *Sfrp2* and *Twist2* expression in $Tdg^{+/+}$ and $Tdg^{-/-}$ MEFs complemented with vectors expressing either a wild-type (pTdg) or a catalytically deficient *Tdg* ($pTdg_{cat-}$, N151A), or a vector control (pV). Target-specific messenger RNA (mRNA) levels were assessed by qRT–PCR and normalized to *Gapdh* mRNA; values represent arbitrary units (means ± s.d.; $n \geq 3$; $^*P < 0.05$; unpaired Student's *t*-test). **b**, ChIP–qPCR analyses to detect H3K27me3 and H3K4me2 marks at the gene promoters indicated in chromatin of $Tdg^{+/+}$, $Tdg^{-/-}$ and $Tdg^{-/-}$ MEFs complemented with a wild-type *Tdg* cDNA. *IAP* and the *Hprt* promoter were used as normalizers for active and repressive chromatin marks, respectively (means ± s.e.m.; $n = 3$; $^*P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test).

the *Sfrp2* and *Twist2* promoters; and reduction of H3K4me2 was coupled with an increase in H3K9me3 but not H3K27me3 at the *Rarb* promoter. Thus, promoter *de novo* methylation in TDG-deficient cells is associated with a loss of H3K4 dimethylation and a concomitant increase in trimethylation of H3K27 more than H3K9.

Stable expression of a TDG encoding complementary DNA (cDNA) in $Tdg^{-/-}$ MEFs (Supplementary Fig. 1f) restored activity to the *Sfrp2* and *Twist2* genes (Fig. 2a). This correlated with a loss of H3K27 trimethylation in their promoters and an increase in H3K4 dimethylation in the case of *Twist2* (Fig. 2b). Expression of *Hoxd13* and *Hoxa10*, however, was not restored although a partial reduction of H3K27 trimethylation also occurred. This indicated that, once H3K4 methylation is lost (*Hoxd13*, *Hoxa10*), the repressive chromatin maintained by H3K9 and H3K27 methylation and aberrant CpG methylation cannot be reversed to an active state by re-expression of *Tdg*. If residual H3K4 methylation is present, however, promoter reactivation is possible, and this requires the catalytic function of TDG[15] as shown for *Sfrp2* and *Twist2* (Fig. 2a).

To address the origin of the epigenetic aberrations in *Tdg* null MEFs, we investigated gene expression and chromatin states in TDG-proficient and -deficient ESCs before and after retinoic-acid-induced *in vitro* differentiation to neuronal progenitor cells[16] (Supplementary Fig. 6a). Strikingly, gene expression differences were minor in ESCs (16 genes, $P \leq 0.05$, FC ≥ 1.5) but increased significantly upon differentiation to neuronal progenitor cells (297 genes, $P \leq 0.05$, FC ≥ 1.5) (Fig. 3a). This was not due to an inability of TDG-deficient ESCs to respond transcriptionally to retinoic acid (Supplementary Fig. 6b), although they showed somewhat faster kinetics of silencing pluripotency genes (*Oct4*, *Nanog*) and activating developmental genes (for example, *Gata6*, *Pax6*) (Supplementary Fig. 6c). Similar to the situation in MEFs, the genes most significantly misregulated in TDG-deficient neuronal progenitor cells control developmental functions, most of

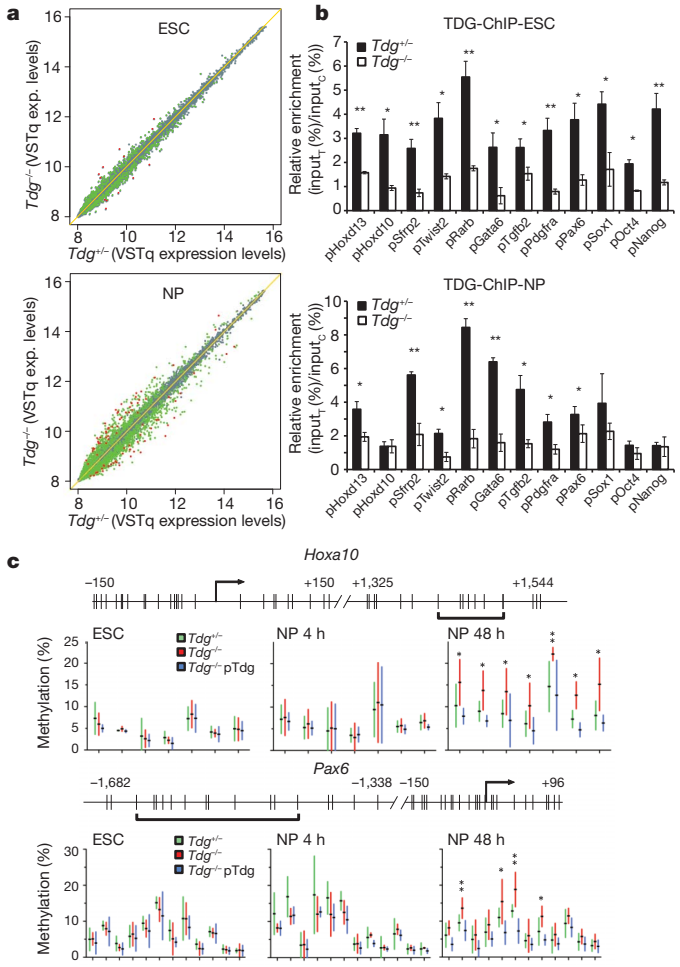

Figure 3 | TDG-dependent differences in gene expression and chromatin status arise during cell differentiation. **a**, Scatter plots comparing gene expression profiles of $Tdg^{+/-}$ and $Tdg^{-/-}$ ESCs or *in vitro* differentiated neuronal progenitors. Green and red dots indicate differentially expressed genes at $P < 0.05$ and $P < 0.01$, respectively. **b**, ChIP–qPCR analysis of TDG association with the gene promoters indicated in chromatin from $Tdg^{+/-}$ and $Tdg^{-/-}$ ESCs and neuronal progenitor (NP) cells. Shown is the relative enrichment of TDG at these promoters normalized to a randomly chosen intergenic control region (means ± s.e.m.; ESCs, $n = 3$; neuronal progenitor cells, $n = 3$; $^{*}P < 0.05$, $^{**}P < 0.01$; unpaired Student's *t*-test). **c**, DNA methylation states at the *Hoxa10* and *Pax6* promoters in TDG-deficient ESCs and neuronal progenitor cells analysed by bisulphite pyrosequencing. Promoter regions are depicted schematically at the top. Vertical tick marks indicate CpG sites, bent arrows transcription start sites, and horizontal brackets the CpGs for which methylation data are presented in the graphs. Methylation levels are given as the percentage of methylated cytosines at each CpG analysed. Shown are means with the 95% confidence intervals (bars) obtained from three differentiation experiments. $^{*}P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test.

them having CpG islands in their promoters and being targets of the polycomb repressive system (Supplementary Fig. 7a). Using ChIP, we confirmed an enrichment of TDG at the promoters of differentially expressed genes both in ESCs and in neuronal progenitor cells (Fig. 3b). This also revealed that TDG associates with the promoters of *Oct4* and *Nanog* in ESCs but not in neuronal progenitor cells and MEFs (Fig. 3b and Supplementary Fig. 6d), suggesting that its interaction is lost upon heterochromatinization of these promoters. Notably, the inability to associate with heterochromatized promoters may explain why re-expression of TDG in *Tdg* null MEFs failed to restore *Hoxd13* and *Hoxa10* transcription (Fig. 2).

Next, we examined the status of CpG methylation in gene promoters downregulated in TDG-deficient neuronal progenitor cells, making use of Na-bisulphite (pyro)sequencing and methylated DNA-immunoprecipitation (MeDIP). Although MeDIP only detected trends for methylation differences at specific promoters (Supplementary Fig. 7b and unpublished observations), pyrosequencing revealed significantly increased DNA methylation in *Tdg* null neuronal progenitor cells at three out of five gene promoters tested (*Hoxa10*, *Pax6*, *Tgfb2*). Notably, these methylation differences were not present in ESCs nor in freshly dissociated embryonic bodies, they arose within 48 h of cultivation of the neuronal progenitor cells in progenitor medium (Fig. 3c and Supplementary Fig. 7c), and the phenotype was complemented by ectopic expression of *Tdg* during cell differentiation. Similarly, histone methylation marks were not different between TDG-proficient and -deficient ESCs but arose in neuronal progenitor cells with an enrichment of H3K27me3 at the promoters of *Hoxd13*, *Hoxa10* (Supplementary Fig. 8) and *Pdgfra* (unpublished observations). Thus, differences in DNA methylation and histone modifications became apparent at the neuronal progenitor cell stage but were not as pronounced as in MEFs, indicating an epigenetic phenotype that may progress upon further differentiation and/or cultivation. Consistently, attempts to differentiate TDG-deficient neuronal progenitor cells to terminal neurons failed because of a rapid loss of cell viability in neuronal-rich medium.

We then wondered whether this epigenetic function of TDG involves active DNA repair, as implicated by the inability of a catalytic-dead TDG (N151A) to complement the loss of *Sfrp2* and *Twist2* expression in *Tdg* null MEFs (Fig. 2). To monitor a possible engagement of downstream base excision repair, we first performed ChIP for XRCC1[17]. This revealed a specific, TDG-dependent enrichment of this critical base excision repair protein at the *Hoxd13*, *Hoxa10*, *Sfrp2* and *Twist2* promoters in MEFs but not in ESCs (Fig. 4a and Supplementary Fig. 5b). Hence, in MEFs, where TDG helps maintain these promoters in an active state, its presence correlates with an enrichment of XRCC1. In ESCs, however, where TDG also associates with these promoters but does not affect their chromatin status, XRCC1 enrichment is not observed. Besides XRCC1, we also found APE1, another component of base excision repair, to associate with these promoters in a TDG dependent manner in MEFs (Fig. 4a). Moreover, retinoic acid treatment of ESCs for 8 h increased the number of chromatin-associated XRCC1 foci in the presence but not in the absence of TDG (Supplementary Fig. 9), and TDG-proficient cells were significantly more sensitive to poly(ADP-ribose) polymerase (PARP) inhibition than TDG-deficient cells upon retinoic-acid-induced differentiation (Supplementary Fig. 10). These observations strongly suggest that cell differentiation-induced TDG activity feeds into PARP and XRCC1-dependent DNA single-strand break repair[18].

Addressing the phenotype on histone modifications, we then found by ChIP that the absence of TDG also compromises the association of the H3K4-specific methyltransferase MLL1 with the promoters of *Hoxd13*, *Hoxa10*, *Sfrp2* and *Twist2* (Fig. 4b). This was apparent in TDG-deficient MEFs but not in ESCs, with the former indeed showing a loss of H3K4 methylation and an occurrence of aberrant CpG methylation at gene promoters reminiscent of the phenotype of MLL defects[19–21]. Similar to MLL, the binding of CBP/p300 to these promoters was significantly reduced in the *Tdg* null MEFs (Fig. 4b). CBP/p300 is a transcription-activating histone acetyltransferase known to interact with TDG[4] and, notably, its association with gene promoters has been reported to protect from polycomb-mediated H3K27 trimethylation[22].

Taken together, our data suggest structural and catalytic functions of TDG in epigenetic maintenance (Fig. 4c). As a structural component, TDG complexes with activating histone modifiers (for example, MLL, CBP/p300) to maintain states of active (H3K4me2) and bivalent (H3K4me2, H3K27me3) chromatin during cell differentiation. In the absence of TDG, the assembly and function of such complexes is distorted and, consequently, chromatin modifications imbalanced towards repressive states. TDG also provides DNA repair capacity to erase CpG methylation locally. Aberrant methylation arises at GC-rich promoters in TDG-deficient cells following lineage commitment, and the frequencies and patterns of these events indicate an underlying

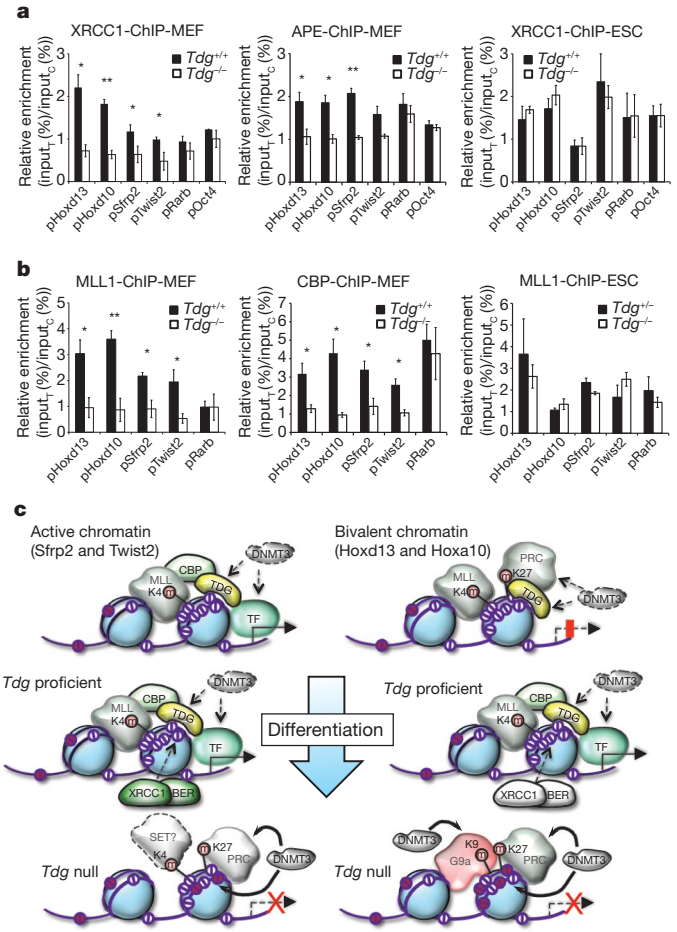

Figure 4 | Structural and catalytic functions of TDG in epigenetic maintenance. a, ChIP–qPCR analysis of XRCC1 and APE1 association with the gene promoters indicated in chromatin of TDG-proficient and -deficient MEFs and ESCs. Shown are relative enrichments of XRCC1 and APE1 at these promoters normalized to a randomly chosen intergenic control region (means ± s.e.m.; $n \geq 3$; $^*P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test). **b**, ChIP–qPCR analysis of MLL1 and CBP/p300 association with the gene promoters indicated in chromatin of TDG-proficient and -deficient MEFs and ESCs. Shown are relative enrichments of MLL1 and CBP/p300 at these promoters normalized to a randomly chosen intergenic control region (means ± s.e.m.; $n \geq 3$; $^*P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test). **c**, Model summarizing epigenetic aberrations and implicated functions observed in the absence of TDG. In ESCs TDG associates with gene promoters in an active 'open' (H3K4me2, for example; *Sfrp2* and *Twist2*, left side) or transiently silent 'bivalent' chromatin conformation (H3K4me2 and H3K27me3, for example; *Hoxd13* and *Hoxa10*, right side). In active chromatin, the lack of TDG results in a partial loss of H3K4 dimethylation and a gain of H3K27 trimethylation as well as in sporadic DNA hypermethylation (red balls) upon cell differentiation. Differentiation-associated activation of promoters in 'bivalent' chromatin involves the demethylation of H3K27me3 and transcription factor binding. The absence of TDG results in an aberrant loss of H3K4 dimethylation accompanied by a gain in repressive H3K9 and H3K27 trimethylation and in DNA methylation, eventually directing irreversible transcriptional silencing. In both cases, the loss of active and the gain in repressive histone marks can be accounted for by a failure of TDG-deficient cells to target MLL and CBP to these promoters. We propose that TDG, as part of transcription regulatory complexes, assures the establishment and the maintenance of proper epigenetic states at developmentally regulated gene promoters. As a DNA glycosylase, it protects these regions from aberrant CpG methylation in a process that engages XRCC1 and APE1, factors essential for downstream base excision repair.

stochastic process of *de novo* methylation. Hence, TDG keeps *de novo* DNMT activities in check to avoid erroneous methylation, and the engagement of XRCC1 and APE1 suggests that it operates through base excision repair. Several previous studies have implicated TDG in active DNA demethylation[8,9,23]. Mechanistically, it may do so on its own, acting as a 5-mC DNA glycosylase[23], or it may cooperate with a 5-mC deaminase (for example, AID/Apobec[24,25] or DNMTs[8]), or a 5-mC hydroxylase (for example, TET1[26,27]) that would convert 5-mC into a favourable substrate for TDG. Numerous efforts to reproduce 5-mC glycosylase activity for mouse and human TDG have failed (Supplementary Fig. 11 and unpublished observations). We therefore consider a deamination or hydroxylation-mediated, TDG-dependent repair process a preferable model for active cytosine demethylation. The mouse *Tdg* knockout phenotype shows that such an epigenetic control system has evolved to protect critical DNA sequences from *de novo* methylation and heterochromatinization during development.

METHODS SUMMARY

***Tdg* knockout mouse and cell lines.** The *Tdg*-targeting construct (Supplementary Fig. 1) was generated by replacement of a NarI–PacI fragment enclosing exons 6 and 7 by a neomycin resistance cassette in a cloned fragment spanning exons 5–10 of the *Tdg* locus. This construct was used to target the *Tdg* allele in 129 mouse ESCs, which were then used to generate chimaeras and, ultimately, $Tdg^{+/-}$ heterozygotes by backcrossing to C57BL/6. The generation and establishment of MEFs and $Tdg^{-/-}$ ESCs was previously described[28].

***In vitro* differentiation.** *In vitro* differentiation of ESCs was performed essentially according to the protocol published in ref. 16. RNA isolation for transcriptome analysis of MEFs or ESCs and neuronal progenitor cells was performed using the RNeasy Mini Kit (Qiagen) or the Trizol reagent (Invitrogen), respectively. Antibodies and sequences of oligonucleotides used for PCR with reverse transcription (RT–PCR), bisulphite sequencing and ChIP are listed in Supplementary Tables 1–4.

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

Received 22 February; accepted 17 November 2010.
Published online 30 January 2011.

1. Gallinari, P. & Jiricny, J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* **383,** 735–738 (1996).
2. Um, S. *et al.* Retinoic acid receptors interact physically and functionally with the T:G mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* **273,** 20728–20736 (1998).
3. Chen, D. *et al.* T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. *J. Biol. Chem.* **278,** 38586–38592 (2003).
4. Tini, M. *et al.* Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Mol. Cell* **9,** 265–277 (2002).
5. Li, Y. Q., Zhou, P. Z., Zheng, X. D., Walsh, C. P. & Xu, G. L. Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. *Nucleic Acids Res.* **35,** 390–400 (2007).
6. Gallais, R. *et al.* Deoxyribonucleic acid methyl transferases 3a and 3b associate with the nuclear orphan receptor COUP-TFI during gene activation. *Mol. Endocrinol.* **21,** 2085–2098 (2007).
7. Zhu, B. *et al.* Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. *Proc. Natl Acad. Sci. USA* **98,** 5031–5036 (2001).
8. Metivier, R. *et al.* Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452,** 45–50 (2008).
9. Kangaspeska, S. *et al.* Transient cyclical methylation of promoter DNA. *Nature* **452,** 112–115 (2008).
10. Cortazar, D., Kunz, C., Saito, Y., Steinacher, R. & Schär, P. The enigmatic thymine DNA glycosylase. *DNA Repair (Amst.)* **6,** 489–504 (2007).
11. Robertson, A. B., Klungland, A., Rognes, T. & Leiros, I. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell. Mol. Life Sci.* **66,** 981–993 (2009).
12. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441,** 349–353 (2006).
13. Furuta, J. *et al.* Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. *Cancer Res.* **66,** 6080–6086 (2006).
14. Cheng, Y. Y. *et al.* Frequent epigenetic inactivation of secreted frizzled-related protein 2 (SFRP2) by promoter methylation in human gastric cancer. *Br. J. Cancer* **97,** 895–901 (2007).
15. Hardeland, U., Bentele, M., Jiricny, J. & Schär, P. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J. Biol. Chem.* **275,** 33449–33456 (2000).
16. Bibel, M., Richter, J., Lacroix, E. & Barde, Y. A. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nature Protocols* **2,** 1034–1043 (2007).

17. Caldecott, K. W. XRCC1 and DNA strand break repair. *DNA Repair (Amst.)* **2,** 955–969 (2003).
18. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434,** 913–917 (2005).
19. Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E. & Korsmeyer, S. J. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc. Natl Acad. Sci. USA* **95,** 10632–10636 (1998).
20. Erfurth, F. E. *et al.* MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. *Proc. Natl Acad. Sci. USA* **105,** 7517–7522 (2008).
21. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* **29,** 6074–6085 (2009).
22. Pasini, D. *et al.* Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* **38,** 4958–4969 (2010).
23. Zhu, B. *et al.* 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *Proc. Natl Acad. Sci. USA* **97,** 5135–5139 (2000).
24. Morgan, H. D., Dean, W., Coker, H. A., Reik, W. & Petersen-Mahrt, S. K. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* **279,** 52353–52360 (2004).
25. Rai, K. *et al.* DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135,** 1201–1212 (2008).
26. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324,** 930–935 (2009).
27. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466,** 1129–1133 (2010).
28. Kunz, C. *et al.* Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. *PLoS Biol.* **7,** e91 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Klewe-Nebenius for preparations of mouse primary fibroblast, and F. Mohn and D. Schübeler for discussions and assistance in the setup and evaluation of the ChIP experiments. The work was supported by project grants from the Swiss National Science Foundation (3100A0-108436; 3100A0-122574/) and the Association For International Cancer Research (01-330).

Author Contributions D.C. established and performed the ChIP and MeDIP analyses and the *in vitro* differentiation experiments, and contributed to writing the paper; C.K. established and characterized MEF lines, designed and performed gene expression and DNA methylation analyses, and contributed to writing the paper; J.S. did blastocyst injections, established the first heterozygous *Tdg* knockout mice, characterized the lethal phenotype of the *Tdg* null embryos and provided SV40-immortalized MEFs; T.L. and Y.S. generated *Tdg*-targeting constructs and established heterozygous and homozygous *Tdg* knockout ES cell lines; Y.S. established *in vitro* differentiation protocols; E.MacD. performed the Big Blue mutation assays; A.W. performed ChIP experiments; D.S. isolated primary MEFs and performed RT–PCR validations of gene expression differences and the PARP inhibitor experiments; A.L.J. established and performed immunofluorescence experiments including XRCC1 foci analyses; F.S. performed bioinformatic analyses of gene expression array data; R.S. affinity-purified anti-TDG antibodies for ChIP; J.J. contributed genomic *Tdg* clones and supervised initial work of T.L.; A.B. was involved in study design (mutation analyses) and supervised the work of J.S. and E.MacD.; P.S. designed, coordinated and supervised the study, analysed the data and wrote the paper. All authors discussed the results and commented on the paper.

Author Information Gene expression array data have been deposited in the NCBI Gene Expression Omnibus under accession number GSE20693. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.S. (primo.schaer@unibas.ch).

METHODS

***Tdg* knockout strategy.** The *Tdg*-targeting construct (Supplementary Fig. 1) was generated by replacement of a NarI–PacI fragment enclosing exons 6 and 7 by a neomycin resistance cassette in a cloned fragment spanning exons 5–10 of the *Tdg* locus. This construct was used to target the *Tdg* allele in 129 mouse ESCs, which were then used to generate chimaeras and, ultimately, $Tdg^{+/-}$ heterozygotes by backcrossing to C57BL/6. The generation and establishment of MEFs and $Tdg^{-/-}$ ESCs was previously described[28].

Cell culture and ESC differentiation. SV40-immortalized MEF cell lines were previously described[29] and cultivated in growth medium (DMEM, 10% FCS, 2 mM L-glutamine) at 37 °C in a humidified atmosphere containing 5% CO_2. For growth of cell lines complemented with Tdg-expressing vectors, the growth medium was additionally supplemented with 1 μg ml^{-1} puromycin.

For isolation of primary MEFs, 10.5 days post-coitum embryos were dissected, homogenized and cells dissociated in 0.05% trypsin-EDTA for 5 min before plating in modified ES cell medium without LIF (DMEM, 10% FCS seraplus, 1× non-essential amino acids, 1 mM sodium pyruvate, 2 mM L-glutamine and 50 μM β-mercaptoethanol, 1× penicillin/streptomycin) and cultivation for 10 days.

ESCs were grown in the presence of feeder cells at 37 °C in ES cell medium (ECM: DMEM, 15% heat-inactivated FCS, LIF (1,000 U ml^{-1}), 1× non-essential amino acids, 1 mM Na-pyruvate, 2 mM L-glutamine and 90 μM β-mercaptoethanol) in a humidified atmosphere containing 5% CO_2.

Before starting retinoic-acid-induced differentiation , ESCs were grown in the absence of feeder cells for two passages. For embryoid body formation during neuronal differentiation, 4×10^6 $Tdg^{+/-}$ or $Tdg^{-/-}$ ESCs were plated into non-adherent bacterial dishes (Greiner Bio-one) in differentiation medium (ECM without LIF and with 10% FCS) and grown at 37 °C with a medium exchange after 2 days. After 4 days, 5 μM all-*trans* retinoic acid was added and cells were further incubated for 4 days with a medium exchange after 2 days. Embryoid bodies were washed twice with 1× PBS and dissociated with freshly prepared trypsin solution (0.05% TPCK-treated trypsin in 0.05% EDTA/1× PBS) at 37 °C for 3 min. Dissociated embryoid bodies were re-suspended in 10 ml differentiation medium and collected by centrifugation at 700*g* for 5 min at room temperature. The pellet was re-suspended in N2 medium (DMEM-F12 nutrient mixture 1:1, 1× N2 supplement) and the cell suspension filtered through a 40-μm nylon cell strainer (BD). Filtered cells were immediately plated onto poly-L-lysine and laminin-coated dishes at a density of 5×10^6 cells per 60-mm dish or 1.5×10^7 cells per 100-mm dish. The N2 medium was exchanged 2 and 24 h after plating, and cells were collected after 4 and 48 h for further analysis.

Retinoic-acid-induced differentiation of ESCs for time course, PARP inhibitor and immunofluorescence experiments was induced in ECM without LIF in the presence of 1 or 5 μM retinoic acid. The retinoic-acid-containing medium was exchanged every 24 h, and cells were collected at the indicated time points. For immunofluorescence experiments, 10^5 ESCs were seeded onto gelatin-coated cover slips 1 day before differentiation. For the analysis of PARP inhibition on cell survival during differentiation, 10^5 ESCs were seeded into gelatin-coated 12-well dishes, 1 day before the addition of 5 μM retinoic acid or further cultivation in ECM. After 24 h, increasing concentrations of the PARP inhibitor AG-014699 (a gift of SelleckChem) were added and cell numbers determined 24 h later with the CASYcell counter. The 50% lethal dose of the inhibitor and statistical differences between Tdg-proficient and -deficient cells were calculated on triplicate experiments by linear regression with 95% confidence intervals using GraphPad Prism software.

Microarray gene expression analysis. For the analysis of differential gene expression between $Tdg^{+/-}$ and $Tdg^{-/-}$ MEFs, total RNA was isolated from three independent cultures of each cell line using the RNeasy Mini Kit (Qiagen), cDNA synthesized from 13 μg RNA with the SuperScript double-Stranded cDNA Synthesis Kit (Invitrogen) followed by *in vitro* transcription reactions using the MEGA Script T7 Kit (Ambion) supplemented with 1.5 mM Bio-11-CTP and Bio-16-UTP (Enzo Life Sciences). cDNAs and cRNAs were purified using the GeneChip Sample Cleanup Module (Qiagen). cRNA (15 μg) was fragmented and hybridized to GeneChip Mouse Expression Arrays 430A (Affymetrix). Hybridized arrays were stained and washed according to the manufacturer's protocol and scanned with an Affymetrix Scanner 3000 7G. Scanned images were processed with Microarray Suite software and obtained 'cel'-files used for further data analysis.

For gene expression analysis of ESCs and *in vitro* differentiated neuronal progenitor cells, total RNA was extracted from independent triplicates using the Trizol reagent (Invitrogen). RNA was quantified using the Quant-iT RiboGreen RNA Assay (Invitrogen) and 500 ng of total RNA subjected to cDNA synthesis and subsequent *in vitro* transcription to biotiylated cRNA using the Illumina TotalPrep RNA Amplification Kit (Ambion, USA). cRNA (1.5 μg) was hybridized to MouseWG-6v2 slides (Illumina) according to the manufacturer's protocol. Bead arrays were washed and stained using FluoroLink Cy3 Streptavidin (GE Healthcare). Fluorescent signals were imaged using the iScan system (Illumina). Scanner images files were processed to probe intensity files by the manufacturer's software and further processed with the genome studio software (Illumina) without normalization and background correction.

Affymetrix data and Illumina probe intensity data were either processed by robust multi-array average or variance stabilization transformation, respectively, using R/Bioconductor software and 'affy' or 'lumi' libraries, followed by quantile normalization. Significance of effects for probes (Illumina) or probe-sets (Affymetrix) was tested in R/Bioconductor ('limma' library) using a moderated *t*-test and the false discovery rate (=5%) method of Benjamini and Hochberg for multiple testing correction. No unspecific filter was applied and multiple probe-sets per gene or probe-sets with ambiguous genomic targets were retained.

Methylation analyses. Genomic DNA from MEFs, ESCs and neuronal progenitor cells was isolated with the QIAamp DNA Mini Kit (Qiagen). DNA (2 μg) was subjected to bisulphite conversion using the EZ DNA Methylation Kit (Zymo Research). Respective target regions were amplified from bisulphite-treated DNA with TrueStart Taq polymerase (New England Biolabs). For conventional bisulphite sequencing, *Hoxd13* or *Sfrp2* promoter regions were amplified from converted DNA and cloned into the XhoI and BamHI restriction sites of pBluescript SK- before sequencing of individual clones. For pyrosequencing, potential regions of hypermethylation were first identified by COBRA. Pyrosequencing primers (Supplementary Table 1) were designed using the PyroMark Assays Design software (version 2.0.1.15, Qiagen). Primer pairs included either one biotinylated primer or one primer containing a universal region. In the latter case, products were subjected to a second amplification using a biotinylated universal primer and Phusion Hot Start High-Fidelity DNA Polymerase (Finnzymes). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen), quantified and 300–500 ng were used for pyrosequencing in a PyroMark Q24 (Qiagen). Reactions were analysed using PyroMark Q24 software (version 2.0.6, Qiagen). Significance of methylation differences between different Tdg-proficient and -deficient cell lines at individual CpG sites was evaluated by unpaired, two-tailed *t*-tests.

ChIP. To crosslink protein-bound DNA, MEFs, ESCs and neuronal progenitor cells were incubated in freshly prepared crosslinking solution (PBS pH 7.4, 1% formaldehyde) at room temperature. The reaction was quenched after 10 min by addition of glycine to a final concentration of 125 mM. After washing twice with cold PBS, cells were collected using a cell scraper and subsequent centrifugation at 600*g* and 4 °C. Nuclei were isolated by incubation in 200 μl of cold ChIP Buffer I (10 mM HEPES pH 6.5, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100) for 5 min on ice followed by two incubations of 5 min on ice in 200 μl cold ChIP buffer II (10 mM HEPES pH 6.5, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl). Pelleted nuclei were lysed in 400 μl ChIP buffer III (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5% Triton X-100, 1% SDS, 1 mM PMSF) for 10 min on ice followed by sonication for 15 min (15 s on, 30 s off, power high) using a Bioruptor sonicator (Diagenode) to produce random chromatin fragments ranging from 300 to 1,000 base pairs. The solution was cleared by centrifugation at 14,000*g* and 4 °C for 10 min and the concentration of chromatin was estimated by absorbance at 260 nm. For ChIP of TDG, MLL and APE1 100–150 μg of chromatin were diluted ten times in ChIP dilution buffer I (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 0.1% Triton X-100, 1 mM PMSF). For histone ChIPs, 25–75 μg of chromatin were diluted in ChIP dilution Buffer II (16.7 mM Tris-HCl pH 8.0, 1.2 mM EDTA, 167 mM NaCl, 1.1% Triton X-100, 0.01% SDS, 1 mM PMSF). Diluted chromatin was pre-cleared at 4 °C for 1 h with 40 μl of a 50% slurry of magnetic Protein G beads (Invitrogen) preblocked with 1 mg ml^{-1} BSA and 1 mg ml^{-1} tRNA (TDG, XRCC1, APE1 and MLL-ChIPs) or salmon sperm single-stranded DNA (histone ChIPs). Precleared chromatin was incubated with 2–5 μg of the respective antibody (Supplementary Table 2) overnight at 4 °C under slow rotation. Immunocomplexes were precipitated with 40 μl of a 50% slurry of blocked Protein G beads and further incubated at 4 °C for 2 h. Beads were then serially washed with 500 μl ChIP wash buffer I (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.1% SDS, 1% Triton X-100), 500 μl ChIP wash buffer II (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 0.1% SDS, 1% Triton X-100) and 500 μl ChIP wash buffer III (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 250 mM LiCl, 1% sodium deoxycholate, 1% NP-40). For TDG, APE1 and MLL ChIPs, beads were washed once with 500 μl ChIP wash buffer I and twice with 500 μl ChIP wash buffer II. After two additional washes with 500 μl TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), bound complexes were eluted by two sequential incubations with 150 μl elution buffer (1% SDS, 0.1 M $NaHCO_3$) at 65 °C for 10 min. Crosslink reversal of eluates and respective input samples (1% of chromatin used for ChIP) was done in the presence of 200 mM NaCl at 65 °C for 4 h followed by proteinase K digestion (50 μg ml^{-1}) in the presence of 10 mM EDTA at 45 °C for 1 h. DNA was purified by phenol/chloroform extraction and Na-acetate/ethanol precipitation, and re-suspended in 10 mM Tris-HCl pH 8.0. qPCR analysis with target specific primers (Supplementary Table 3) was performed using Quantitect SYBR Green (Qiagen) with

a Rotor-Gene 3000 thermocycler (Qiagen). The significance of different ChIP efficiencies among Tdg-proficient and -deficient cell lines was evaluated from triplicate experiments by non-paired, two-tailed *t*-tests.

MeDIP. MeDIP assays were performed as described in ref. 30. Briefly, genomic DNA was prepared from 5×10^6 cells by incubation in lysis buffer (20 mM Tris-HCl pH 8.0, 4 mM EDTA, 20 mM NaCl, 1% SDS and 1 mg ml^{-1} proteinase K) at 55 °C for 5 h and subsequent phenol/chloroform extraction and Na-acetate/ethanol precipitation. DNA pellets were re-suspended in TE containing 20 μg ml^{-1} RNase. DNA was sonicated as described for ChIP followed by NaCl (400 mM)/EtOH precipitation in the presence of glycogen-carrier. Fragmented DNA (4 μg) in 450 μl TE was denatured at 95 °C for 10 min and immediately chilled on ice. After addition of 10× immunoprecipitation buffer (100 mM sodium phosphate pH 7.0, 1.4 M NaCl, 0.5% Triton X-100), the DNA was incubated with 10 μg of a monoclonal anti 5-methylcytidine antibody (clone 33D2, Eurogentec) at 4 °C for 2 h. Immuno-complexes were precipitated by the addition of 40 μl M-280 sheep anti mouse IgG antibody coupled Dynabeads (Invitrogen) and incubation at 4 °C for 2 h followed by three washes in 700 μl IP buffer. Bound material was treated with 250 μl proteinase K digestion buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS and 0.25 mg ml^{-1} proteinase K) at 50 °C for 3 h. Immunoprecipitated methylated DNA was purified by phenol/chloroform extraction followed by Na-acetate/ethanol precipitation and re-suspended in TE. qPCR analysis of sonicated genomic input DNA and MeDIP DNA with target specific primers (Supplementary Table 3) was performed as described for ChIP, and significance of MeDIP efficiencies tested by non-paired, two-tailed *t*-tests.

Quantitative RT–PCR analyses. Total RNA (2–4 μg) extracted by RNeasy Mini Kit or by Trizol methods was reverse transcribed with the RevertAid H Minus M-MuLV Kit (Fermetas) according to the manufacturer's protocol. qPCR with target specific primers (Supplementary Table 4) was performed using Power SYBR Green Master Mix (Applied Biosystems) with a Rotor-Gene 3000 thermocycler. Conditions for each target were validated by standard and melting curve analyses. Target-specific amplifications were normalized to a GAPDH control and data of at least three independent experiments were analysed by unpaired, two-tailed *t*-tests. *Tdg* genotype-specific target gene expression in primary MEFs was analysed by the non-parametric Kruskal–Wallis test and *post hoc* Dunn's multiple comparison.

Western blot analyses. Whole-cell extracts were prepared by cell lysis in lysis buffer (50 mM Na-phosphate pH 8.0, 125 mM NaCl, 1% NP-40, 0.5 mM EDTA, 1 mM PMSF, 1 mM DTT, 1× complete protease inhibitor, 2× phosphatase inhibitor cocktail 1 and 2) on ice for 30 min and clarification by centrifugation (15 min, 20,000*g*, 4 °C). Chromatin extracts were isolated as described for ChIP assays. Soluble proteins (50 μg) were separated on 7% or 10% SDS–polyacrylamide gels and transferred to a nitrocellulose membrane (Millipore). Membranes were washed once with TBS-T (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% Tween-20), blocked with blocking buffer (TBS-T, 5% dry milk) at room temperature for 1 h and incubated with the primary antibody at 33 °C (anti-mTDG) or room temperature (anti-DNMT1, anti-DNMT3a, anti-XRCC1, anti-APE1, anti-MLL, anti-β-actin) for 1 h in blocking buffer. Dilutions were 1:10,000 for the rabbit anti-mTDG, the mouse anti-β-actin and the anti-DNMT1 antibodies; 1:1,000 for the anti-DNMT3a and anti-XRCC1 antibodies; 1:500 for the anti-APE1 and anti-MLL antibodies. Washing steps after hybridization were once at 33 °C and twice at room temperature for 15 min for anti-mTDG, or three times at room temperature for 10 min for all other antibodies. Membranes were incubated with secondary HRP-conjugated antibodies diluted 1:5,000 in blocking buffer and at room temperature for 1 h. After three washing steps of 10 min at room temperature, detection of the signals was performed using the Immobilon Western Chemiluminescent HRP Substrate (Millipore).

Cytotoxicity assays. For measurement of γ-ray sensitivity, MEF single-cell suspensions at a cell density of 2×10^5 cells ml^{-1} in PBS were irradiated with the indicated doses in a Gammacell 40 irradiator using ^{137}Cs as a radioactive source. Irradiated cells were plated in 96-well microtitre plates at a density of 1000 cells per well in growth medium, and survival was measured after 3 days using the Cell Counting Kit-8 (Dojindo). Alternatively, survival was determined by clonogenic growth by plating 500–2000 cells in triplicate in 10-cm dishes containing growth medium and counting of Giemsa-stained colonies after 10 days. To measure sensitivity to H_2O_2, cells were plated at 2,500 cells per well in 96-well plates. After 24 h cells were treated for 15 min with the indicated concentrations of H_2O_2, washed with PBS and incubated in fresh growth medium for a further 24 h before measurement of survival with the Cell Counting Kit-8. Survival was determined as the percentage of mock-treated cells.

Base release assay. For base release assays, 25–50 μg of ESC whole-cell extracts were incubated with 0.5 pmol of a fluorescein-labelled GC/TG, GCm/CG or GCm/mCG DNA substrate in reaction buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM DTT, 1 mg ml^{-1} BSA) at 37 °C for 1 h (GC/TG) or overnight (methylated substrates). Resulting AP-sites were cleaved by the addition of NaOH to a final concentration of 100 mM and heating to 95 °C for 10 min. Subsequently, DNA was ethanol-precipitated overnight at −20 °C in the presence of 0.3 M Na-acetate pH 5.2 and 0.4 mg ml^{-1} carrier t-RNA. DNA was collected by centrifugation (20 min, 20,000*g*, 4 °C) and washed with 80% ethanol. Air-dried pellets were re-suspended in loading buffer (1× TBE, 90% formamide), heated at 95 °C for 5 min and immediately chilled on ice. Reaction products were separated on denaturing 8 M urea/15% polyacrylamide gels in 1× TBE. The fluorescein-labelled DNA was visualized with a Typhoon 9400 and quantified using ImageQuant TL software (GE Healthcare).

Immunofluorescence. For detection of XRCC1 foci during retinoic acid stimulation, cells were fixed in ice-cold methanol for 5 min, then permeated in 0.2% Triton X-100/PBS pH 7.4 and 0.2% Triton X-100/0.2% $NaBH_4$/PBS pH 7.4 on ice for 5 min each. The induceability of XRCC1 foci formation in ESCs was tested by incubation with H_2O_2 (50 μM in PBS) or PBS for 15 min at 37 °C and an additional 5 min in ECM with LIF before further processing. Coverslips were blocked in blocking buffer (1% BSA/0.05% Tween20/PBS pH 7.4), stained with rabbit anti-XRCC1 antibody (1:100 in blocking buffer) at room temperature for 1 h and washed three times for 10 min with blocking buffer before labelling with goat anti-rabbit Alexa Fluor 594 (1:200 in blocking buffer) for 30–60 min. After two washes of 10 min with blocking buffer, cells were again fixed in −20 °C cold methanol, incubated in blocking buffer for 1 h and stained with a mouse monoclonal anti-PCNA antibody (1:100 dilution) in blocking buffer overnight at 4 °C. Slides were counterstained for DNA with 50 ng ml^{-1} DAPI and mounted in VectaShield mounting medium (Vector Lab). Slides were randomized and blinded before z-stacks were acquired on a Leica SP5 with the 405-nm diode, argon 488 nm and He–Ne 594-nm laser lines. XRCC1 foci numbers for individual cells were determined by visual inspection of the three-dimensional stacks. One hundred and fifty (retinoic acid stimulation) or 50 (H_2O_2) cells per sample were analysed. For co-staining of PAR and XRCC1 during retinoic acid differentiation, cells were fixed with 2% formaldehyde/PBS at room temperature for 30 min and permeabilized with PBS/0.2% Triton-X100 for 30 min. Antigene detection was done with a 1:250 diluted monoclonal α-PAR antibody 10H (Enzo Life Sciences) and a polyclonal α-XRCC1 as described above, but using 1:250 diluted anti-rabbit Alexa Fluor 594 and anti-mouse Alexa Fluor 488 secondary antibodies (Invitrogen). Pictures were acquired with a Nikon Diaphot 300 epifluorescence microscope using identical settings for all slides.

29. Kunz, C. *et al.* Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. *PLoS Biol.* **7,** e91 (2009).
30. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genet.* **37,** 853–862 (2005).