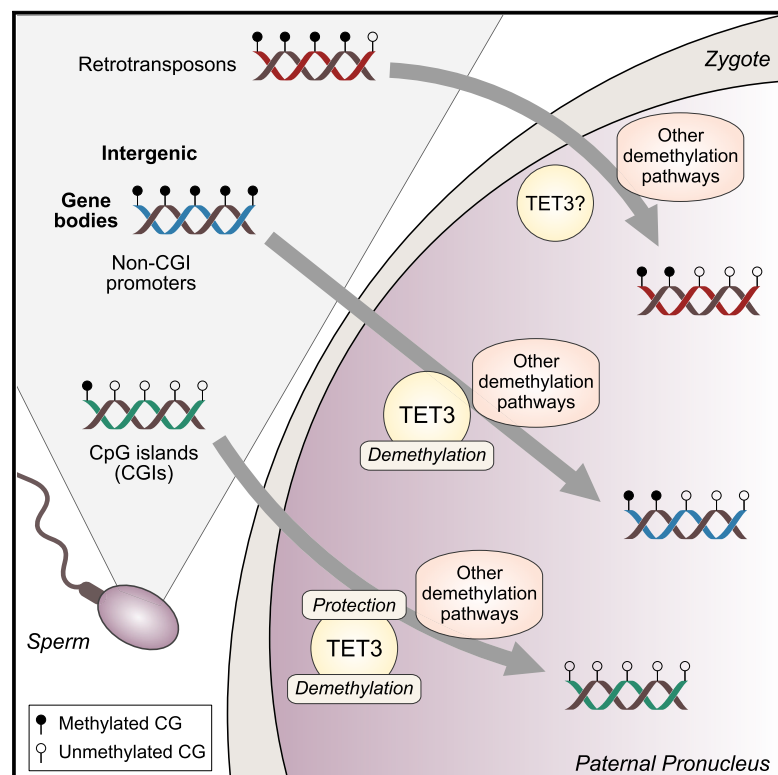# Cell Reports

Genome-wide Bisulfite Sequencing in Zygotes Identifies Demethylation Targets and Maps the Contribution of TET3 Oxidation

Graphical Abstract

Authors

Julian R. Peat, Wendy Dean, ...,
Timothy A. Hore, Wolf Reik

Correspondence

julian.peat@babraham.ac.uk (J.R.P.),
wolf.reik@babraham.ac.uk (W.R.)

In Brief

Peat et al. employ whole-genome bisulfite sequencing to generate a global picture of the profound DNA methylation reprogramming that is triggered by fertilization. This identifies gene bodies as a major demethylation target where TET3 oxidation plays an important role and uncovers a protective TET3 function as well as redundancy in the demethylation machinery.

Highlights

- An enhanced bisulfite strategy allows genome-wide methylation profiling of zygotes

- Gene bodies constitute a major target of zygotic demethylation and TET3 oxidation

- The impact of TET3 loss is moderate and implicates redundant demethylation pathways

- Protective TET3 activity shields certain CpG islands against methylation buildup

Accession Numbers

GSE63417

CrossMark

CellPress

Genome-wide Bisulfite Sequencing in Zygotes Identifies Demethylation Targets and Maps the Contribution of TET3 Oxidation

Julian R. Peat,[1,*] Wendy Dean,[1] Stephen J. Clark,[1] Felix Krueger,[2] Sébastien A. Smallwood,[1] Gabriella Ficz,[1,6] Jong Kyoung Kim,[3] John C. Marioni,[3,4] Timothy A. Hore,[1] and Wolf Reik[1,4,5,*]

[1]Epigenetics Programme, The Babraham Institute, Cambridge CB22 3AT, UK
[2]Bioinformatics Group, The Babraham Institute, Cambridge CB22 3AT, UK
[3]European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI), Hinxton CB10 1SD, UK
[4]Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK
[5]Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK
[6]Present address: Barts Cancer Institute, Queen Mary University of London, London EC1M 6BQ, UK
*Correspondence: julian.peat@babraham.ac.uk (J.R.P.), wolf.reik@babraham.ac.uk (W.R.)
http://dx.doi.org/10.1016/j.celrep.2014.11.034

SUMMARY

Fertilization triggers global erasure of paternal 5-methylcytosine as part of epigenetic reprogramming during the transition from gametic specialization to totipotency. This involves oxidation by TET3, but our understanding of its targets and the wider context of demethylation is limited to a small fraction of the genome. We employed an optimized bisulfite strategy to generate genome-wide methylation profiles of control and TET3-deficient zygotes, using SNPs to access paternal alleles. This revealed that in addition to pervasive removal from intergenic sequences and most retrotransposons, gene bodies constitute a major target of zygotic demethylation. Methylation loss is associated with zygotic genome activation and at gene bodies is also linked to increased transcriptional noise in early development. Our data map the primary contribution of oxidative demethylation to a subset of gene bodies and intergenic sequences and implicate redundant pathways at many loci. Unexpectedly, we demonstrate that TET3 activity also protects certain CpG islands against methylation buildup.

INTRODUCTION

Over a decade ago, pioneering studies demonstrated that fertilization triggers a global and active loss of DNA methylation from the paternal genome, but not its maternal counterpart (Mayer et al., 2000; Oswald et al., 2000). This methylation reprogramming occurs at a pivotal developmental time point, when a global transcriptional transition is required for the genesis of totipotency from specialized germ cell states (Hemberger et al., 2009). The role of methylation changes is still unclear, and a "resetting" of the epigenome may be necessary to generate a blank canvas on which to paint new regulatory marks for the development of the embryo.

The wave of zygotic erasure does not affect all regions of the paternal genome equally, as methylation at certain sequences, such as parental imprints and active retrotransposons (e.g., intracisternal A particles [IAPs]), must be maintained for embryonic viability (Seisenberger et al., 2012). Recent genome-scale profiling by reduced-representation bisulfite sequencing (RRBS-seq) extended observations of variation in methylation dynamics across different classes of repetitive elements and even within their component families (Smith et al., 2012). However, as the great majority of 5-methylcytosine (5mC) in sperm lies outside the CpG islands (CGIs) enriched in RRBS-seq (Kobayashi et al., 2012), which are largely hypomethylated (Smallwood et al., 2011), a comprehensive characterization of zygotic demethylation targets is still lacking. Such a profile would be particularly instructive in the context of a major transcriptional and fate transition, and would complement recent genome-wide studies in gametes and later preimplantation stages (Kobayashi et al., 2012; Wang et al., 2014).

We and others have demonstrated that oxidation of the paternal genome by TET3 constitutes an important mechanism for the active removal of its methylation in the zygote (Gu et al., 2011; Santos et al., 2013; Wossidlo et al., 2011), adding to the small repertoire of factors that are known to contribute to demethylation in the zygote, including the elongator complex (Okada et al., 2010) and the deaminase AID (Santos et al., 2013). However, the knowledge of these factors' involvement has been set against a limited understanding of normal demethylation dynamics across the genome, and an appreciation of the role of specific pathways that is restricted to a handful of loci (Gu et al., 2011; Hajkova et al., 2010; Okada et al., 2010). Therefore, the architecture of the demethylation machinery and the relationship between different mechanisms (synergy, redundancy, or specificity) remain almost entirely unknown.

Here, we set out to produce a comprehensive picture of methylation dynamics in the paternal pronucleus and thus
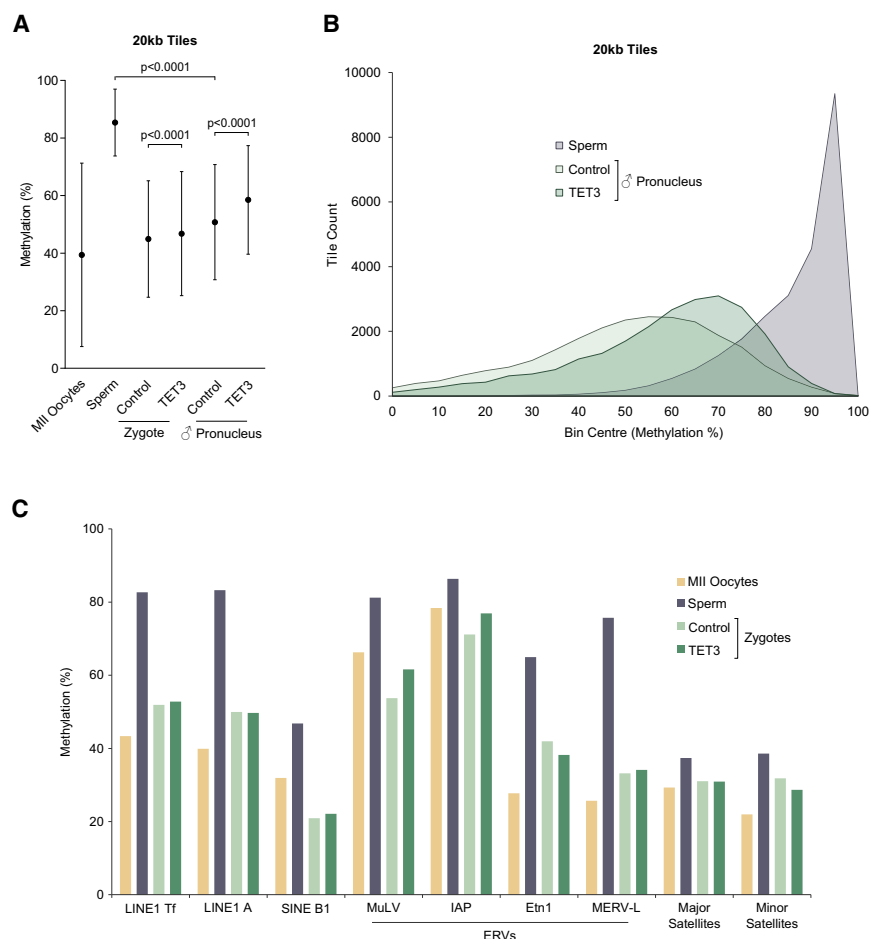
A

B

C

Figure 1. The Zygotic Methylation Landscape

(A) Global methylation levels in gametes and zygotes. The graph shows the mean and SD of the methylation values of 20 kb tiles across the genome.

(B) Frequency distribution of 20 kb tile methylation values. Tiles were allocated to bins centered every 5%.

(C) Mean methylation levels in repetitive-element classes. "Zygotes" refers to overall levels, as natural sequence variability prohibits assignment of parental alleles using SNPs. LINE, long interspersed nuclear element; SINE, short interspersed nuclear element; MuLV, murine leukemia virus; IAP, intracisternal A particle; Etn, early transposon; MERV-L, murine endogenous retrovirus with leucine tRNA primer; ERV, endogenous retrovirus. See also Figure S1.

provide the foundation on which to map the contribution of oxidation using genetic disruption of TET3 activity in the zygote.

RESULTS

In order to obtain genome-wide profiles from limited material, we optimized a strategy for generating whole-genome bisulfite sequencing (WGBS-seq) libraries wherein bisulfite treatment is used to both convert cytosines and fragment DNA at the start of the protocol (Miura et al., 2012). This enhances yield by eliminating the need for fragmentation by sonication and avoiding the degradation of adaptor-tagged sequences. We made substantial modifications (for the detailed protocol, see Supplemental Experimental Procedures) that allowed us to consistently generate high-quality libraries from fewer than 300 zygotes.

This enhanced WGBS-seq protocol was applied to zygotes collected from control females and females carrying a conditional deletion that abolishes TET3 activity in the germline (referred to as TET3 zygotes; see Supplemental Experimental Procedures and sequencing statistics in Table S1). 129S2/SvHsd studs were used in matings in order to specifically trace the fate of paternal methylation in the zygote with SNPs. To permit examination of demethylation by oxidation as well as dilution at DNA replication, we collected late-stage zygotes

that had completed S phase, after the global loss of 5mC and gain in 5-hydroxymethylcytosine (5hmC) had occurred (Santos et al., 2013; Wossidlo et al., 2011). The concordance of imprint control region methylation levels with expected values indicated that the data sets generated by this strategy were of high quality (Figure S1A).

The Zygotic Methylation Landscape

We first used 20 kb tiles to assess global 5mC values (Figure 1A). The results demonstrated a large decrease in mean methylation from sperm (85.4%) to the paternal pronucleus (50.8%), in line with genome-wide demethylation. Methylation levels were slightly but significantly increased in TET3 zygotes (44.9% versus 46.8%) and this effect was more marked in paternal DNA (50.8% versus 58.5%), consistent with the role of TET3 in promoting paternal demethylation as shown by immunofluorescence and locus-specific bisulfite analysis (Gu et al., 2011; Santos et al., 2013; Wossidlo et al., 2011).

The distribution of methylation values in sperm and paternal pronuclear DNA further established a dramatic reprogramming of methylation patterns at fertilization: whereas the majority of sequences were highly methylated in sperm, few tiles (7.5%) remained over 80% methylated in the zygote (Figure 1B), demonstrating that most of the paternal genome was affected by some methylation loss. The bulk of sequences in the paternal pronucleus possessed intermediate methylation in a broad distribution that shifted upward significantly in the absence of TET3 oxidation.

Examination of major repetitive-element classes confirmed previous observations of a transition to a more hypomethylated state in zygotes relative to sperm, with the established exception of IAPs (Figure 1C; Lane et al., 2003; Smith et al., 2012; Wossidlo et al., 2010). The use of WGBS-seq allowed this analysis to be extended to largely uncharacterized elements that are poorly
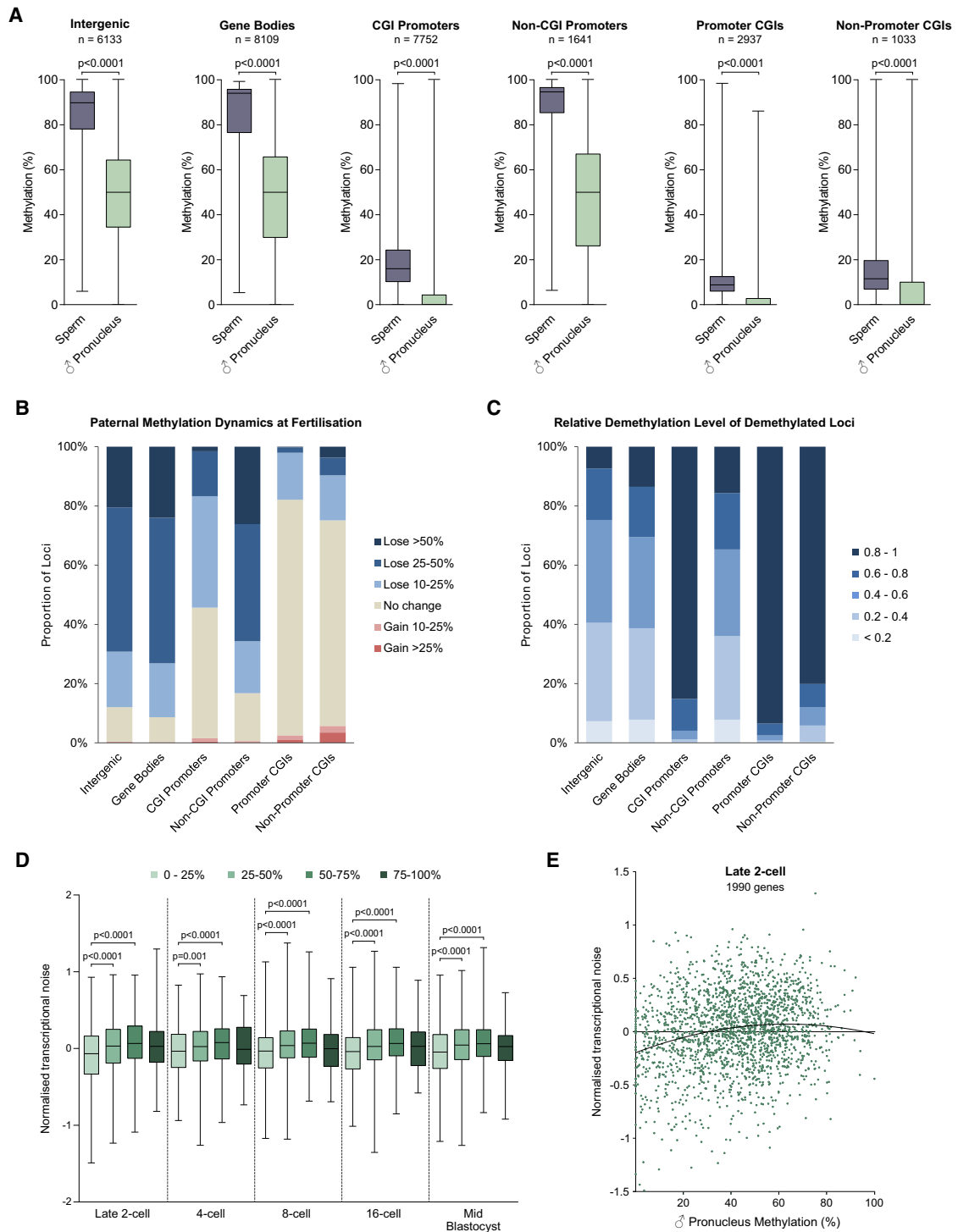
Figure 2. Methylation Trajectories in Different Genomic Contexts

(A) Distribution of methylation values at sequence features in sperm and the paternal component of control zygotes. The plot displays the median (bar), inter-quartile range (box), and maximum and minimum (whiskers). The p values shown are the result of a paired ANOVA multiple-comparison test with Sidak correction; "n" denotes the number of sequences that met quantification criteria.

(B) Absolute methylation change between sperm and the paternal component of control zygotes. Only changes ≥ 10% and significant according to a chi-square test (corrected p value < 0.05) are shown; all others are recorded as "no change."

(C) RDL at identified demethylated loci, calculated by dividing the absolute paternal methylation change (as in B) by the sperm methylation level. Demethylated loci are defined as those that had an absolute paternal methylation loss of ≥10% and were significant according to a chi-square test (corrected

(legend continued on next page)

covered by RRBS-seq due to the paucity of CCGG sites used for enrichment. This demonstrated that MERV-L elements (which contain only four CCGG sites in 6.5 kb) undergo substantial demethylation in the zygote. Removal of methylation may be important for the strong transcriptional activation of these repeats at fertilization (Kigami et al., 2003; Macfarlan et al., 2012). Furthermore, our data show that methylation at satellite repeats, which lack CCGG sites altogether, is relatively stable during zygotic reprogramming. This maintenance may be important for the heterochromatinization that is required for the first mitosis to proceed (Fadloun et al., 2013; Probst and Almouzni, 2011). These insights into methylation trajectories in the zygote underscore the value of the WGBS-seq approach.

The absence of oxidation by TET3 had a limited impact on the methylation levels of repetitive elements (none were altered by more than 10%). This indicates either that TET3 is not targeted to these sequences or that alternative demethylation pathways operate redundantly or are recruited to noncanonical targets to compensate for the absence of TET3 oxidation.

These analyses paint a picture of a significant but moderate impact of TET3 at a global level, consistent with immunofluorescence data (Santos et al., 2013), and suggest that the major role of oxidative demethylation is at single-copy loci.

Methylation Trajectories in Different Genomic Contexts

We next examined methylation dynamics at specific genomic features more closely. Given the particular biology of methylation at CGIs and transcription start sites (Jones, 2012), we further subcategorized the CGIs and promoters based on whether they overlapped with a promoter region or CGI, respectively.

An examination of methylation patterns first established that loss of paternal methylation at fertilization was significant at all sequence features, but occurred to varying degrees (Figure 2A). A quantitative comparison of methylation levels in sperm and paternal DNA revealed that ~90% of intergenic and gene body sequences were significantly demethylated (decrease ≥ 10%, $p < 0.05$), with ~70% losing >25% methylation (Figure 2B). As these regions contribute the great majority of paternally methylated cytosines (Figure S1B), this demonstrates the pervasive removal of methylation from single-copy sequences and further identifies gene bodies as a major target of zygotic demethylation. In contrast, little demethylation was observed at CGI-associated sequences (Figure 2B), reflecting the general hypomethylation of these regions in sperm (Figures 2A and S2A). The minority that were methylated underwent a similar degree of demethylation compared with other sequences (Figure S2B), indicating that

paternal methylation is generally targeted for removal regardless of the context. Although they were a less prevalent feature and thus not a large carrier of paternal methylation, non-CGI promoters followed a trajectory similar to that observed for intergenic and gene body sequences (Figures 2A and 2B).

In order to delineate the contribution of replication to the observed demethylation, we expressed absolute demethylation as a proportion of the initial methylation in sperm to determine the relative demethylation level (RDL) for each locus (Figure 2C). An RDL > 0.5 cannot be explained simply by a 2-fold reduction in methylation at replication, and therefore implies the action of active removal pathways. Demethylation at loci with RDL < 0.5 may also be active but cannot be discriminated from replicative dilution; therefore, this analysis measures the minimum number of actively demethylated loci. We applied a conservative threshold of RDL > 0.6 and found that the degree of methylation loss at 25% and 30% of demethylated intergenic and gene body sequences, respectively, which constitute the majority of demethylation targets, implicates an active process. This demonstrates that while lack of maintenance at replication may provide an important mechanism for loss of methylation, active demethylation of the paternal genome is common. Strikingly, although few CGI-associated sequences lost 5mC, where demethylation did occur, it was almost exclusively active.

By annotating each demethylated locus with the nearest gene, we examined the relationship between zygotic demethylation and transcription in 2-cell embryos using published mRNA-seq data (Park et al., 2013), and searched for functional enrichment in Gene Ontology (GO) databases. With the exception of nonpromoter CGIs, genes associated with demethylated loci have elevated expression at the 2-cell stage (Figure S2C) and as a group are enriched for cytoskeletal, ion transport, signaling, protein modification, and RNA processing terms (Table S2). This expression profile and the enrichment for several known ZGA transcriptional modules, including GTPase signaling and RNA processing (Xue et al., 2013), suggest a role for demethylation in preparing for transcriptional activation in the early embryo. The prevalence of cytoskeletal and cell junction terms may be linked to the function of these components during embryo compaction (Ducibella and Anderson, 1975; Fierro-González et al., 2013). Genes associated with demethylated nonpromoter CGIs are not highly expressed at ZGA and are enriched for a largely nonoverlapping set of GO terms, including many neural functions.

We next questioned whether pervasive loss of intragenic methylation at fertilization was linked to variation in transcription

---

p value < 0.05). n = 5,393 intergenic sequences, 7,405 gene body sequences, 4,213 CGI promoters, 1,366 non-CGI promoters, 526 promoter CGIs, and 257 nonpromoter CGIs.

(D) Distribution of early embryo paternal transcriptional noise levels at genes allocated into quartiles according to average intragenic methylation levels in the paternal pronucleus. Normalized transcriptional noise refers to expression-corrected variation in transcription from the paternal allele between single cells as calculated in Supplemental Experimental Procedures. The p values shown are the result of an ANOVA multiple-comparison test between indicated data sets with Sidak correction. The differences between second and third quartiles, and third and fourth quartiles are not significant at any stage according to this test. The total number of genes that fell into each quartile was as follows: 0%–25%, 1,379; 25%–50%, 2,468; 50%–75%, 2,605; and 75%–100%, 270. Genes that were not expressed at a particular stage were excluded from analysis of that stage. The number of genes analyzed at each stage is given in Supplemental Experimental Procedures.

(E) Relationship between the average paternal pronuclear gene body methylation and paternal transcriptional noise at the late 2-cell stage. A quadratic equation is a significantly better fit to the data than either a linear or no relationship ($p < 0.0001$ in both cases, extra sum-of-squares F test).
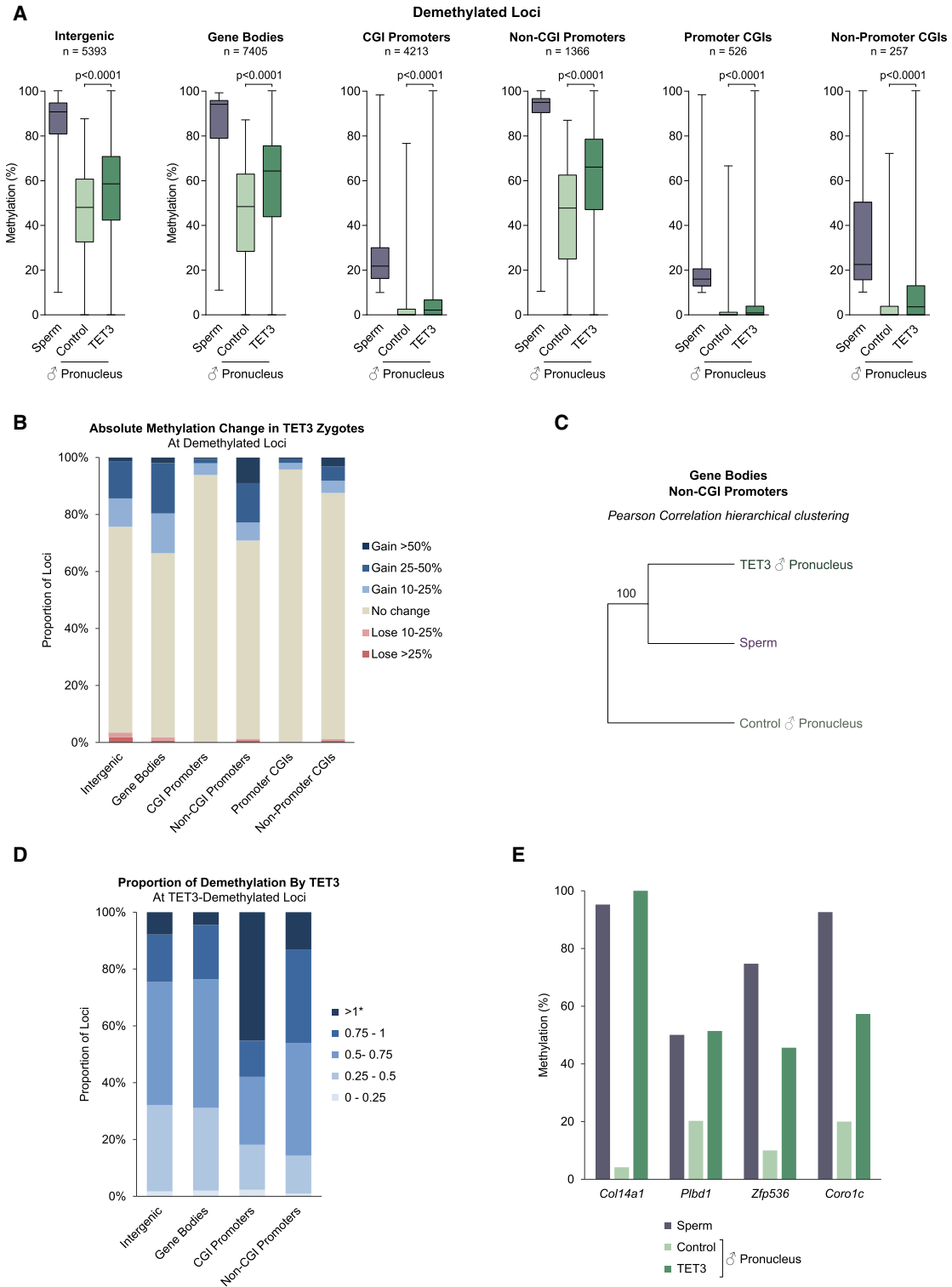
See also Figure S2 and Table S2.

Figure 3. Impact of TET3 Deletion on Methylation Dynamics

(A) Distribution of methylation values at demethylated loci, defined as in Figure 2C. The p values shown are the result of a paired ANOVA multiple-comparison test with Sidak correction.

(B) Impact of TET3 deletion on demethylated loci. The ranges indicate the absolute change in methylation level between the paternal component of control and TET3 zygotes. Only changes ≥ 10% and significant according to a chi-square test (corrected p value < 0.05) are shown; all others are recorded as "no change."

(legend continued on next page)

between cells at subsequent stages. Using published allele-specific, single-cell RNA-seq data (Deng et al., 2014), we generated a metric of cell-to-cell transcriptional noise at paternal alleles for each expressed gene during early embryonic development (see Supplemental Experimental Procedures). We then examined the distribution of noise signals at genes whose average intragenic methylation levels in the paternal pronucleus fell into each of four quartiles (Figure 2D). Strikingly, this revealed that intermediate levels of gene body methylation (25%–75%) are associated with significantly higher transcriptional noise than hypomethylated genes (<25%) throughout the preimplantation embryo. Although the low number of hypermethylated genes (>75%) in the paternal pronucleus precludes establishment of statistical significance, their transcriptional noise levels appear substantially lower than those of intermediate-methylated genes. Indeed, the relationship between intragenic methylation and transcriptional noise is significantly better modeled by a quadratic equation than a straight line at all developmental stages examined (Figures 2E and S2D). The widespread demethylation of paternally hypermethylated gene bodies to intermediate methylation levels in the zygote (Figure 2A) may therefore promote transcriptional noise in the early embryo. As our noise metric normalizes for expression levels, this association is independent of any influence of intragenic methylation on overall transcription. This analysis uncovers a connection between methylation and early embryonic transcription at a major demethylation target, further emphasizing the importance of genome-wide approaches to explore methylation reprogramming and its functional impact.

The Absence of TET3 Activity Disrupts Normal Methylation Dynamics

Having established a detailed picture of zygote methylation, we sought to understand the contribution of oxidation by TET3 to this landscape.

The absence of TET3 oxidation significantly elevated the paternal pronuclear methylation levels of all features at identified demethylation targets (Figure 3A) as well as these groups as a whole (Figure S3A). By quantifying the impact of TET3 deletion on demethylated loci, we established that loss of methylation was significantly impaired (gain $\leq$ 10%, $p < 0.05$) at 25%–30% of intergenic, gene body, and non-CGI promoter sequences, but found relatively little effect on the small number of demethylated CGI-associated sequences (Figure 3B). As intergenic and gene body sequences constitute the bulk of demethylated loci, this identifies these regions as

the major target of oxidative demethylation by TET3. Hierarchical clustering analysis of all methylation values showed that for gene body and non-CGI promoter sequences, but not other features, paternal pronuclear DNA from zygotes lacking TET3 activity formed a strongly supported clade with sperm DNA rather than paternal pronuclear DNA from control zygotes (Figures 3C and S3B). This underlines the importance of TET3 for reprogramming of sperm methylation patterns at these features.

Nevertheless, our data demonstrate that although TET3 makes a significant contribution, demethylation proceeds normally at the majority of loci in its absence. This indicates that either most demethylated loci are not targeted by TET3 or redundant or compensatory mechanisms minimize the impact of its absence. At CGI-associated sequences, where demethylation is almost exclusively active (Figure 2C), this implicates the activity of alternative enzymatic pathways such as the elongator complex or oxidation-independent BER.

In general, the targets of TET3 demethylation reflected the characteristics of demethylated loci with regard to their association with elevated expression in 2-cell embryos (Figure S3C) and, where there were a sufficient number to test, enrichment for similar GO terms (Table S2). This suggests that although the importance of oxidative 5mC removal varies with the sequence feature, demethylation by TET3 is not associated with specific biological functions.

We next focused on the loci that were demethylated by TET3 and examined the degree to which its absence affected normal demethylation in order to gauge the contribution of oxidation to their methylation trajectory (Figure 3D). At a substantial number of loci, particularly the small number of targeted CGI promoters (e.g., the *Col14a1* intragenic sequence and *Plbd1* CGI promoter), TET3 was responsible for all or most of the observed demethylation (Figure 3E). However, at most loci, TET3 deletion resulted in partial impairment of demethylation (e.g., the *Zfp356* intragenic sequence and *Coro1c* non-CGI promoter; Figure 3E), indicating that multiple pathways can act at these sequences—either in synergy with TET3 oxidation or to compensate (incompletely) for its loss. The prevalence of partial impairment demonstrates that such redundant targeting or compensatory ability is a widespread feature of the demethylation machinery.

This analysis further revealed that at some demethylated loci, paternal methylation in TET3 zygotes actually increased to higher levels than in sperm (i.e., value > 1 on Figure 3D), implying that beyond its function in demethylation, TET3 is required to

(C) Dendrogram derived from hierarchical clustering using the Pearson correlation distance for the indicated sequence features. The number above the branch denotes the approximately unbiased p values computed by multiscale bootstrap resampling (n = 10,000) using the pvclust package for R (Suzuki and Shimodaira, 2006). This was identical for both features.
(D) Proportion of normal demethylation affected by the loss of TET3 at TET3-demethylated loci, calculated by dividing the absolute paternal methylation change between control and TET3 zygotes (as in B) by the normal absolute demethylation in control zygotes (as in Figure 2B). TET3-demethylated loci are defined as demethylated loci that underwent an absolute paternal methylation gain of ≥10% in the absence of TET3 and were significant according to a chi-square test (corrected p value < 0.05). n = 1,310 intergenic sequences, 2,487 gene body sequences, 259 CGI promoters, and 398 non-CGI promoters. The number of promoter and nonpromoter CGIs that met these criteria was insufficient for meaningful analysis. *Proportional change > 1 indicates that paternal methylation in TET3 zygotes increased to higher levels than in sperm.
(E) Examples of complete (*Col14a1* intragenic sequence and *Plbd1* CGI promoter) and partial (*Zfp356* intragenic sequence and *Coro1c* non-CGI promoter) impairment of demethylation due to loss of TET3 oxidation.
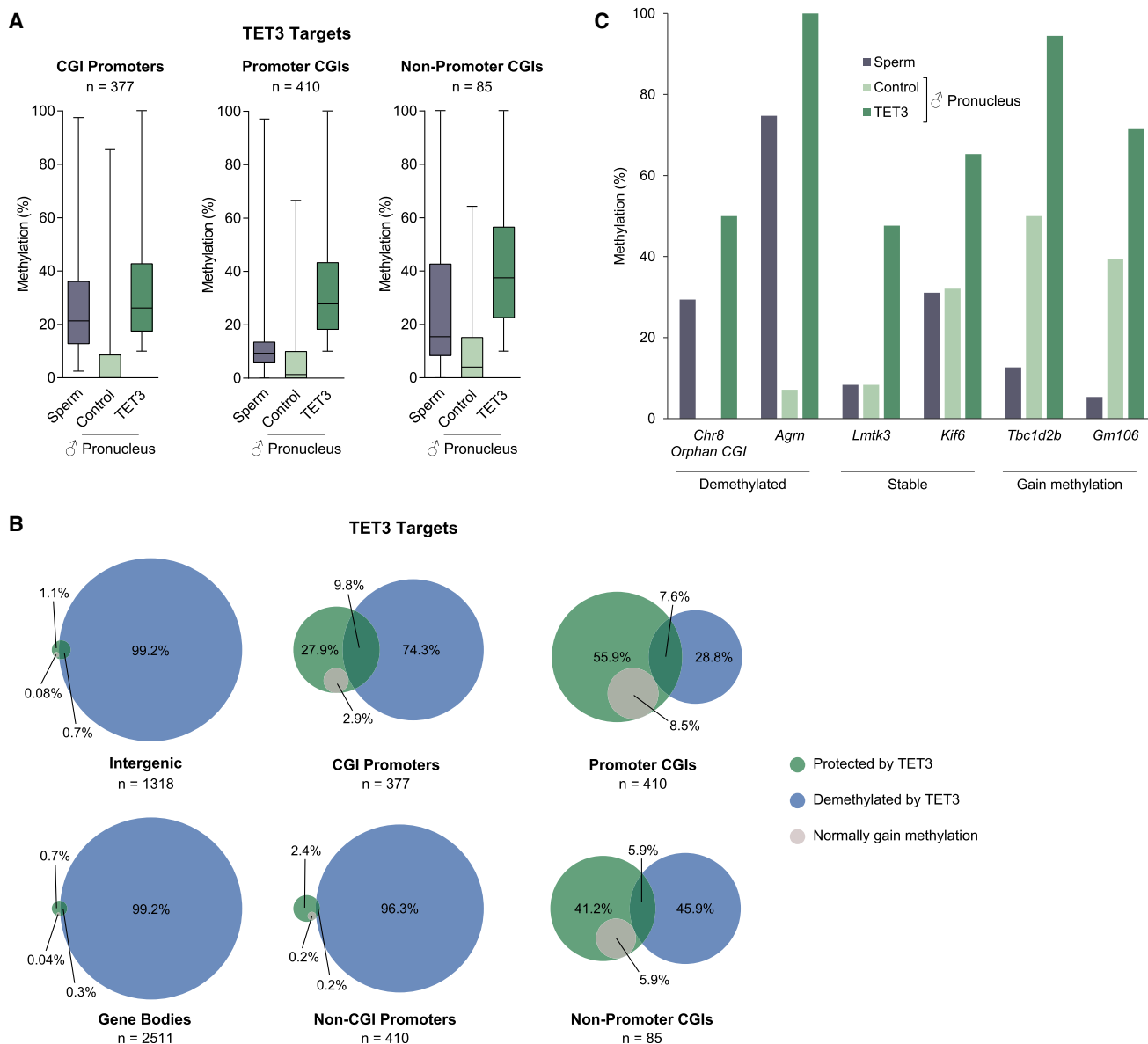See also Figure S3 and Table S2.

A

TET3 Targets

B

TET3 Targets

C

Figure 4. A Protective Role for TET3

(A) Distribution of methylation values for TET3 targets at CGI-associated sequences. TET3 targets are defined as loci where the absence of TET3 activity resulted in a paternal methylation gain of $\geq 10\%$, which is significant according to a chi-square test (corrected p value < 0.05).

(B) Function of TET3 at its targets and their methylation behavior. The area of each circle is proportional to the percentage of loci in that category. TET3 targets are defined as in (A). Protected, paternal methylation $\geq 10\%$ higher in TET3 zygotes than in sperm and significant according to a chi-square test (corrected p value < 0.05); demethylated, criteria as in Figure 2C; normal gain methylation, paternal methylation $\geq 10\%$ higher in control zygotes than in sperm and significant according to a chi-square test (corrected p value < 0.05). Percentages do not necessarily add to 100 due to the application of statistical criteria.

(C) Examples of loci protected by TET3 classified according to whether they are normally demethylated, stable, or gain methylation. *Agrn* and *Gm106* are CGI promoters, *Lmtk3* is an intragenic CGI, *Kif6* is a non-CGI promoter, and *Tbc1d2b* is a promoter CGI. Chr8 Orphan CGI is located at chromosome 8 29088955-29089878 (NCBI37 assembly).

See also Figure S4.

protect certain sequences from aberrant de novo methylation in the zygote. Indeed, an examination of all loci that gained paternal methylation in TET3 zygotes ("TET3 targets") suggests that this phenomenon occurs frequently at CGI-associated sequences (Figures 4A and S4A).

By determining whether paternal methylation in TET3 zygotes increased to significantly higher levels than in sperm ($\geq 10\%$, p < 0.05), we generated a list of loci protected by TET3 to assess their normal methylation behavior and establish the relative importance of this protective function (Figures 4B and 4C).

Although a subset of protected sequences are demethylation targets (e.g., *Agrn* CGI promoter and orphan CGI), the majority do not normally undergo methylation loss at fertilization and instead are stably maintained (e.g., *Lmtk3* intragenic CGI and *Kif6* non-CGI promoter; Figure 4C) or gain some methylation (e.g., *Tbc1d2b* promoter CGI and *Gm106* CGI promoter; Figure 4C). Thus, TET3 targets a distinct set of loci specifically to prevent accumulation of methylation. Although they constitute a small fraction (2.4%) of total TET3 targets, almost twice as many promoter CGIs are protected than demethylated. A roughly equal number of nonpromoter CGIs (which account for 1.8% of TET3 targets) fall into each category (Figure 4B). Protection is therefore an important function at these sequence features, in contrast to intergenic, gene body, and non-CGI promoter sequences, where demethylation is the sole function of TET3 at >95% of loci (Figure 4B).

Like its demethylation targets, the loci that are protected by TET3 are associated with elevated expression in the early embryo (Figure S4B), suggesting that TET3 prevents accumulation of methylation to safeguard transcriptional activation.

DISCUSSION

WGBS-Seq Characterizes Methylation Trajectories across the Genome

Here, we present genome-wide methylation profiles of zygotes, allowing a characterization of the extensive methylation reprogramming that occurs at this profound developmental transition. This fills an important gap in WGBS-seq studies tracking methylation dynamics from gametes to the blastocyst stage (Kobayashi et al., 2012; Wang et al., 2014), and expands the perspective from enrichment-based profiling of zygotes, including two reports published while this study was under review (Guo et al., 2014; Shen et al., 2014; Smith et al., 2012). By extensively optimizing a noncanonical WGBS-seq strategy, we were able to overcome the technical barrier of low cell numbers, which enabled us to access previously poorly characterized regions of the genome.

Our data paint a picture of global remodeling of sperm methylation patterns in which the major carriers of 5mC (repetitive elements, intergenic regions, and gene bodies) are extensively demethylated. The observation of pervasive removal of 5mC from gene bodies adds a new dimension to our understanding of methylation reprogramming, which up to now has centered on promoters and CGIs. Furthermore, the WGBS-seq profiles allowed us to establish the behavior of functionally important repetitive sequences lacking the CCGG sites used in RRBS-seq, including the demethylation of MERV-L elements and stable maintenance of satellite methylation.

The degree of methylation loss from many sequences necessarily implicates active removal processes and shows that such mechanisms operate at a large number of loci. The role of active and passive mechanisms at sequences undergoing lower proportional demethylation (RDL < 0.5) is unclear. Dissecting their relative contribution across the genome will necessitate inhibition of replication coupled with WGBS-seq to examine the generality of observations of a large role for passive demethylation from RRBS-seq (Guo et al., 2014; Shen et al., 2014).

We find that loss of 5mC from all functional sequence features except nonpromoter CGIs is associated with transcriptional activation in the early embryo, as assessed by expression levels and functional enrichment. Clearly, this relationship requires experimental testing. Given the existence of multiple demethylation pathways in the zygote that may act redundantly, including replicative dilution, it may be challenging to perturb demethylation sufficiently to evaluate any transcriptional effect. The potential involvement of regulatory mechanisms in addition to DNA methylation that operate to promote transcriptional activation could further complicate interpretation. Indeed, maternal deletion of TET3 had little impact on 2-cell and inner cell mass transcriptomes (Shen et al., 2014).

A connection between zygotic demethylation of gene bodies and ZGA seems at odds with the positive correlation between intragenic methylation and active transcription observed in many tissues, including gametes (Jones, 2012; Kobayashi et al., 2012). However, recent studies suggest that this relationship is more nuanced. A meta-analysis in human cell lines found that intragenic methylation was highest at moderately expressed genes and low at both weakly and strongly expressed genes (Jjingo et al., 2012). Functionally, methylation can modulate the use of alternative promoters that initiate within the gene body (Maunakea et al., 2010), and the methylation of exons regulates their incorporation during splicing in a mechanism that involves the opposing effects of CTCF and MeCP2 on RNA polymerase II (Maunakea et al., 2013; Shukla et al., 2011). This raises the possibility that demethylation of gene bodies following fertilization is involved in the transition from gamete-specific promoters and exons to those employed in the early embryo.

Beyond its impact on overall expression levels, we uncovered a connection between intragenic methylation in the paternal pronucleus and transcriptional variability that endures throughout the preimplantation embryo. In contrast to the negative correlation reported in some somatic cells (Huh et al., 2013), intermediate methylation in the paternal pronucleus was consistently associated with higher noise levels relative to hypomethylated and hypermethylated gene bodies. Paternal demethylation at fertilization results in a massive conversion of hypermethylation to intermediate methylation levels at gene bodies, and is therefore linked to elevated transcriptional noise. Whether this relationship is causal will need to be examined experimentally, which will involve challenges similar to those faced in attempting to understand the role of demethylation in ZGA. Intermediate methylation levels in postreplicative zygotes could be the result of averaging differentially methylated strands or chromatids generated by processes such as replicative dilution and uneven active demethylation. Such differences would create a variation in methylation between cells upon subsequent divisions that could drive transcriptional noise. The use of strand-specific and single-cell methylation profiling technologies will provide insight into this possibility.

It has been posited that methylation heterogeneity provides a means of breaking symmetry during fate allocation through its influence on transcriptional activity (Lee et al., 2014), and indeed the loss of methylation greatly increases the symmetry of embryonic stem cell divisions (Jasnos et al., 2013). The association of zygotic demethylation with transcriptional noise invites

speculation that this serves to generate heterogeneity in the early embryo that contributes to the first lineage decisions. We look forward to further investigations into the relationship among epigenetic reprogramming, heterogeneity, and lineage specification.

Dissecting the Contribution of TET3 to the Zygotic Methylation Landscape

Oxidation by TET3 has been shown to provide an important pathway for removal of 5mC from the paternal genome at fertilization (Gu et al., 2011; Santos et al., 2013; Wossidlo et al., 2011), adding to other identified active pathways such as BER (Hajkova et al., 2010; Santos et al., 2013) and the elongator complex (Okada et al., 2010), and the potential for passive demethylation at the first DNA replication. However, the targets and interplay of these pathways have remained almost entirely uncharacterized. By combining our whole-genome methylation profiling with maternal deletion of zygotic TET3 activity, we were able to examine the contribution of TET3 to paternal methylation trajectories across the genome. This broadens findings from recently published RRBS-seq studies using similar genetic deletion approaches (Guo et al., 2014; Shen et al., 2014) and represents an important step toward mapping the division of labor among components of the demethylation machinery.

Although loss of TET3 activity disrupts methylation patterns at all single-copy features, the major impact is seen at intergenic, gene body, and non-CGI promoter sequences, where demethylation is impaired at more than a quarter of loci. As these regions constitute the bulk of demethylated loci, this establishes oxidation as a prevalent pathway for removal of 5mC across the genome. However, these data imply that TET3 is not required for demethylation of most loci, particularly among CGI-associated sequences, and this is also true of repetitive elements, which are largely unaffected in deletion zygotes. TET3 activity may be restricted to this limited number of loci; alternatively, it is possible that TET3 in fact targets additional loci, but that other demethylation pathways operate redundantly or are recruited in compensation and thereby neutralize the effect of its absence. Discriminating among these possibilities will require the simultaneous disruption of multiple demethylation pathways or the ability to generate detailed maps of TET3 binding to paternal pronuclear DNA. However, the observation that loss of TET3 frequently results in partial impairment of demethylation already demonstrates redundancy or compensation at a large number of loci. This is an important insight into the architecture of the demethylation machinery. Such redundancy could serve to limit the impact of the loss of a single pathway on methylation reprogramming and may be especially prevalent at sequences where demethylation must be ensured for developmental integrity. This could include LINE1 and MERV-L retrotransposons, whose activation is critical for early cleavage divisions (Beraldi et al., 2006; Kigami et al., 2003).

In addition to its contribution to zygotic demethylation, we uncovered a role for TET3 in protecting against the accumulation of aberrant de novo methylation at CGI-associated sequences. Mechanistically, it seems plausible that the same catalytic activity achieves both functions, i.e., TET3 provides protection through the oxidation of any deposited methyl group—a mecha-

nism for active maintenance of a hypomethylated state in the zygote. Alternatively, TET3 could prevent accumulation of methylation independently of its oxidase function, perhaps by inhibiting the de novo methylation machinery or recruiting other chromatin modifiers. The elevated expression in 2-cell embryos of both demethylation and protection targets of TET3 suggests that these roles serve a common function in facilitating transcriptional activation.

Because bisulfite treatment cannot discriminate between 5mC and 5hmC (Huang et al., 2010), it is important to consider that the extent of demethylation may be underestimated when assessed with this technique. Furthermore, conventional bisulfite analysis is not informative regarding events downstream of the initial 5mC oxidation, as 5fC, 5caC, and unmodified cytosine behave alike (He et al., 2011). Previous work has shown that oxidized bases may be diluted passively due to a lack of maintenance at DNA synthesis (Hashimoto et al., 2012), and 5fC and 5caC can also be actively processed to unmodified cytosine by excision and BER, or decarboxylation (He et al., 2011; Ito et al., 2011; Schiesser et al., 2012, 2013). Although some studies have used immunofluorescence or locus-specific modified bisulfite techniques to examine the use of these demethylation pathways (Guo et al., 2014; Inoue and Zhang, 2011; Inoue et al., 2011), it will be important for future studies to employ quantitative and genome-wide techniques that are capable of discriminating each cytosine modification in order to assess the full extent of 5mC removal and track the fate of oxidation products across the genomes of zygotes and early embryos.

EXPERIMENTAL PROCEDURES

Mouse Lines and Zygote Collection

All experimental procedures were approved by the Animal Welfare, Experimentation and Ethics Committee at the Babraham Institute and were performed under licenses by the Home Office (UK) in accordance with the Animals (Scientific Procedures) Act 1986. Mice with a conditional deletion in the *Tet3* gene were generated for a previous study (Santos et al., 2013; details in Supplemental Experimental Procedures).

Female mice were superovulated by intraperitoneal injection of pregnant mare's serum followed by intraperitoneal injection of human chorionic gonadotrophin (hCG) 48 hr later. Superovulated females were naturally mated with a stud from a 129S2/SvHsd or C57BL/6J background and examined for a vaginal plug the following day. Zygotes were collected 12–13 hr after the presumed time of insemination (i.e., 24–25 hr after hCG injection), at which stage zygotes have completed S phase (Santos et al., 2005). Verification of the pronuclear stage in batches of zygotes was performed manually.

Preparation of WGBS-Seq Libraries

Three independent collections of zygotes were obtained for each genotype (control and TET3; see Supplemental Experimental Procedures for details), giving a total of 225 control and 237 TET3 zygotes, of which 120 and 129 were derived from 129S2/SvHsd studs for control and TET3 samples, respectively. Zygotes were pooled and whole-genome bisulfite libraries were prepared using a post-bisulfite adaptor tagging (PBAT) strategy optimized from Miura et al. (2012). An outline of this approach and a detailed protocol are provided in Supplemental Experimental Procedures.

Analysis of WGBS-Seq Data Sets

Details regarding analysis of the WGBS-seq data sets are provided in Supplemental Experimental Procedures.

Bisulfite data for C57BL/6J MII oocytes and sperm were drawn from Smallwood et al. (2014) (bulk and all single cells merged; Gene Expression Omnibus accession number GSE56879) and Kobayashi et al. (2012) (DNA Data Bank of Japan accession number DRA000484), respectively.

For data sets generated in this study, raw reads were trimmed using Trim Galore, aligned, and deduplicated, and methylation calls were extracted with Bismark (Krueger and Andrews, 2011) using custom pipelines for non-allele-specific analysis (alignment to NCBIM37) and paternal-specific analysis (alignment to both NCBIM37 and a 129S1/SvImJ genome).

For defined genomic features, probes that did not meet the threshold coverage for reliability were excluded. The methylation level was expressed as the mean of all sufficiently covered individual CG sites within the defined region. Methylation dynamics were calculated by comparing data sets using the SeqMonk chi-square filter with a p value requirement of <0.05 after multiple-test correction, and a minimum of ten observations and 10% difference.

For repetitive elements, Bismark was used to map all reads from each data set against consensus sequences constructed from Repbase (Jurka et al., 2005). The methylation level was expressed as the mean of individual CG sites.

Functional enrichment was assessed by analyzing gene lists with the DAVID web tool (Huang et al., 2009) using a background of all mouse genes. This was performed with the Functional Annotation Table feature for level 5 terms in the GOTERM databases. Terms with p < 0.05 after Benjamini correction were considered to be significant.

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the WGBS-seq data reported here is GSE63417.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.11.034.

AUTHOR CONTRIBUTIONS

J.R.P., W.D. and W.R. conceived and designed the study. J.R.P. managed mouse colonies, undertook technical development, prepared libraries, analyzed data, and wrote the paper. W.D. collected zygote samples. S.J.C. contributed to technical development and prepared libraries. F.K. performed bioinformatic processing. S.A.S. contributed to technical development. G.F. was involved in designing and establishing the TET3 mouse line. J.K.K. and J.C.M. contributed to transcriptional noise analysis. T.H. contributed to the study strategy and provided data analysis guidance. W.D. and W.R. supervised the study.

ACKNOWLEDGMENTS

REFERENCES

Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E., and Spadafora, C. (2006). Expression of LINE-1 retroposons is essential for murine preimplantation development. Mol. Reprod. Dev. 73, 279–287.

Deng, Q., Ramsköld, D., Reinius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 343, 193–196.

Ducibella, T., and Anderson, E. (1975). Cell shape and membrane changes in the eight-cell mouse embryo: prerequisites for morphogenesis of the blastocyst. Dev. Biol. 47, 45–58.

Fadloun, A., Eid, A., and Torres-Padilla, M.-E. (2013). Mechanisms and dynamics of heterochromatin formation during mammalian development: closed paths and open questions. Curr. Top. Dev. Biol. 104, 1–45.

Fierro-González, J.C., White, M.D., Silva, J.C., and Plachta, N. (2013). Cadherin-dependent filopodia control preimplantation embryo compaction. Nat. Cell Biol. 15, 1424–1433.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S.G., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477, 606–610.

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.-P., Hu, B., et al. (2014). Active and passive demethylation of male and female pronuclear DNA in the Mammalian zygote. Cell Stem Cell 15, 447–458.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science 329, 78–82.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. 40, 4841–4849.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303–1307.

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat. Rev. Mol. Cell Biol. 10, 526–537.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R., and Rao, A. (2010). The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS ONE 5, e8888.

Huh, I., Zeng, J., Park, T., and Yi, S.V. (2013). DNA methylation and transcriptional noise. Epigenetics Chromatin 6, 9.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science 334, 194.

Inoue, A., Shen, L., Dai, Q., He, C., and Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res. 21, 1670–1676.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 333, 1300–1303.

Jasnos, L., Aksoy, F.B., Hersi, H.M., Wantuch, S., and Sawado, T. (2013). Identifying division symmetry of mouse embryonic stem cells: negative impact of DNA methyltransferases on symmetric self-renewal. Stem Cell Rep. 1, 360–369.

Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V., and Jordan, I.K. (2012). On the presence and role of human gene-body DNA methylation. Oncotarget 3, 462–474.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Wali-chiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. *110*, 462–467.

Kigami, D., Minami, N., Takayama, H., and Imai, H. (2003). MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. Biol. Reprod. *68*, 651–654.

Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O., Sato, S., Nakabayashi, K., Hata, K., Sotomaru, Y., et al. (2012). Contribution of intra-genic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet. *8*, e1002440.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methyl-ation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., and Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis *35*, 88–93.

Lee, H.J., Hore, T.A., and Reik, W. (2014). Reprogramming the methylome: erasing memory and creating diversity. Cell Stem Cell *14*, 710–719.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative pro-moters. Nature *466*, 253–257.

Maunakea, A.K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. *23*, 1256–1269.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethy-lation of the zygotic paternal genome. Nature *403*, 501–502.

Miura, F., Enomoto, Y., Dairiki, R., and Ito, T. (2012). Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res. *40*, e136.

Okada, Y., Yamagata, K., Hong, K., Wakayama, T., and Zhang, Y. (2010). A role for the elongator complex in zygotic paternal genome demethylation. Nature *463*, 554–558.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Park, S.-J., Komata, M., Inoue, F., Yamada, K., Nakai, K., Ohsugi, M., and Shirahige, K. (2013). Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. Genes Dev. *27*, 2736–2748.

Probst, A.V., and Almouzni, G. (2011). Heterochromatin establishment in the context of genome-wide epigenetic reprogramming. Trends Genet. *27*, 177–185.

Santos, F., Peters, A.H., Otte, A.P., Reik, W., and Dean, W. (2005). Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. Dev. Biol. *280*, 225–236.

Santos, F., Peat, J., Burgess, H., Rada, C., Reik, W., and Dean, W. (2013). Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. Epigenetics Chromatin *6*, 39.

Schiesser, S., Hackner, B., Pfaffeneder, T., Müller, M., Hagemeier, C., Truss, M., and Carell, T. (2012). Mechanism and stem-cell activity of 5-carboxycyto-sine decarboxylation determined by isotope tracing. Angew. Chem. Int. Ed. Engl. *51*, 6516–6520.

Schiesser, S., Pfaffeneder, T., Sadeghian, K., Hackner, B., Steigenberger, B., Schröder, A.S., Steinbacher, J., Kashiwazaki, G., Höfner, G., Wanner, K.T., et al. (2013). Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. J. Am. Chem. Soc. *135*, 14593–14599.

Seisenberger, S., Peat, J.R., Hore, T.A., Santos, F., Dean, W., and Reik, W. (2012). Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos. Trans. R. Soc. B Biol. Sci. *368*, 20110330.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014). Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell *15*, 459–470.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature *479*, 74–79.

Smallwood, S.A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat. Genet. *43*, 811–814.

Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat. Methods *11*, 817–820.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339–344.

Suzuki, R., and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics *22*, 1540–1542.

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Zhang, J., Li, G., Ci, W., et al. (2014). Programming and inheritance of parental DNA methylomes in mammals. Cell *157*, 979–991.

Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Schöler, H., and Walter, J. (2010). Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. EMBO J. *29*, 1877–1888.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxy-methylcytosine in the mammalian zygote is linked with epigenetic reprogram-ming. Nat. Commun. *2*, 241.

Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C.Y., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature *500*, 593–597.