Spring 2019 – Epigenetics and Systems Biology
Discussion Session (Epigenetics and Development)
Michael K. Skinner – Biol 476/576
Week 8 (February 26)

Epigenetics of Cell and Developmental Biology

Primary Papers

1. Hackett, et al., (2013) Science. 339:448
2. Finestra, et al., (2017) Current Opinion in Cell Biology. 46:54-61
3. Barau, et al., (2016) Science. 354:909

Discussion

Student 21 – Ref #1 above

- What stages of development and cells have reduced DNA methylation?
- What technology was used?
- What role does 5hmC have in the process?

Student 23 – Ref #2 above

- What is Xist?
- What is X inactivation and mechanism?
- What is chromosome reactivation?

Student 24 – Ref #3 above

- What is DNMT and isoform identified?
- What is a transposon and effect?
- What is the epigenetic observation on the regulation of transposon repression?

performed with *P*-element–transposon–containing Cg4-neo and DmTNP or hTh9 (Fig. 2, B and C, and fig. S7). These observations imply that the DmTNP and hTh9 proteins can nick DNA, independent of having *P*-element termini, which would lead to elevated gene transfer via DNA linearization. A similar observation was made for the SET domain and mariner transposase fusion gene–containing protein (SETMAR or Metnase protein), but this protein is inactive for transposition of HsMAR transposons (*8*, *9*). However, most important, the presence of *P*-element termini on Cg4-neo enhanced the DNA integration activity of both DmTNP and hTh9 3 to 5 times above background, which suggested transpositional DNA integration. Many G418-resistant colonies were also obtained (fig. S7) when Cg4-neo was transfected into a stable HEK293 cell line induced to express a tetracycline-inducible human THAP9 gene (Fig. 2B).

To analyze the nature of the DNA integration events in the G418-resistant colonies, genomic DNA was isolated from individual colonies obtained from DmTNP or hTh9 cotransfections with Cg4-neo, and the sites of *P*-element insertion were characterized by splinkerette polymerase chain reaction (PCR) (*17*, *18*) followed by DNA sequencing. DNA sequence analysis of PCR integration sites identified distinct integration sites with novel 8–base pair (bp) target-site duplications (TSDs) for individual integration events into the human genome that had occurred with both DmTNP and hTh9 (Table 1 and tables S1 and S2). Taken together, these data indicate that human THAP9 actively integrates genetically marked *Drosophila P*-element vectors into human cells by transposition.

The studies reported here indicate that the human THAP9 gene encodes an active DNA transposase that can mobilize *Drosophila P*-element transposons in *Drosophila* and human cells. It will be interesting to investigate the physiological relevance of THAP9's transposition function and to find out if any THAP9 recombination signal DNA elements can be found in the human genome. This is the first report, beyond the V(D)J recombination system, of an active DNA transposase in the human genome. *P* element–like transposons and THAP9-related genes are not restricted to *Drosophila* or related insect species but are widely distributed in eukaryotic genomes like *Ciona* (sea squirt), zebrafish, chicken, and *Trichomonas vaginalis* (a parasitic protozoan) (*7*, *19*). The THAP9 gene is absent and has apparently been lost from sequenced rodent genomes (*6*). Although many of the human transposase–related genes are derived from DNA transposons (43 of 47) (*2*), most have not been characterized, with the exception of the V(D)J recombinase RAG1 and RAG2 (*10*, *11*) and the SETMAR (Metnase) protein (*8*). It is possible that other human genes of this class, besides THAP9, may also encode active DNA transposases.

References and Notes

1. C. Feschotte, E. J. Pritham, *Annu. Rev. Genet.* **41**, 331 (2007).
2. E. S. Lander *et al.*, *Nature* **409**, 860 (2001).
3. S. Schaack, C. Gilbert, C. Feschotte, *Trends Ecol. Evol.* **25**, 537 (2010).
4. H. H. Kazazian Jr., *Science* **303**, 1626 (2004).
5. C. Feschotte, *Nat. Rev. Genet.* **9**, 397 (2008).
6. S. E. Hammer, S. Strehl, S. Hagemann, *Mol. Biol. Evol.* **22**, 833 (2005).
7. H. Quesneville, D. Nouaud, D. Anxolabehere, *Mol. Biol. Evol.* **22**, 741 (2005).
8. M. Shaheen, E. Williamson, J. Nickoloff, S.-H. Lee, R. Hromas, *Genetica* **138**, 559 (2010).
9. S.-H. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18075 (2005).
10. S. D. Fugmann, A. I. Lee, P. E. Shockett, I. J. Villey, D. G. Schatz, *Annu. Rev. Immunol.* **18**, 495 (2000).
11. M. Gellert, *Annu. Rev. Biochem.* **71**, 101 (2002).
12. M. Roussigne *et al.*, *Trends Biochem. Sci.* **28**, 66 (2003).
13. S. Hagemann, W. Pinsker, *Mol. Biol. Evol.* **18**, 1979 (2001).
14. A. Sabogal, A. Y. Lyubimov, J. E. Corn, J. M. Berger, D. C. Rio, *Nat. Struct. Mol. Biol.* **17**, 117 (2010).
15. Y. M. Mul, D. C. Rio, *EMBO J.* **16**, 4441 (1997).
16. P. J. Southern, P. Berg, *J. Mol. Appl. Genet.* **1**, 327 (1982).
17. C. Horn *et al.*, *Nat. Genet.* **39**, 933 (2007).
18. C. J. Potter, L. Luo, *PLoS ONE* **5**, e10168 (2010).
19. J. M. Carlton *et al.*, *Science* **315**, 207 (2007).
20. E. L. Beall, D. C. Rio, *Genes Dev.* **10**, 921 (1996).

Acknowledgments: We thank J. M. Taliaferro, J. Aspden, and M. Francis for comments and discussion; R. Rawat for technical assistance; and C. J. Potter (Johns Hopkins Medical School) for helpful suggestions about the splinkerette PCR experiments. This work was supported by NIH grants R01GM48862, R01GM61987, R01 GM097352, R01GM104385, and R01GM094890. S.M. and D.C.R. conceived the experiments. S.M., A.S., and D.C.R. performed cell culture and molecular biological experiments and analyzed data. S.M. and D.C.R. wrote the paper. The authors declare no competing financial interests. Requests for materials should be addressed to the corresponding author.

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6118/446/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References (*21*, *22*)

19 October 2012; accepted 27 November 2012
10.1126/science.1231789

Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine

Jamie A. Hackett,[1,2] Roopsha Sengupta,[1,2]* Jan J. Zylicz,[1,2,3]* Kazuhiro Murakami,[1,2]* Caroline Lee,[1,2] Thomas A. Down,[1] M. Azim Surani[1,2,3]†

Mouse primordial germ cells (PGCs) undergo sequential epigenetic changes and genome-wide DNA demethylation to reset the epigenome for totipotency. Here, we demonstrate that erasure of CpG methylation (5mC) in PGCs occurs via conversion to 5-hydroxymethylcytosine (5hmC), driven by high levels of TET1 and TET2. Global conversion to 5hmC initiates asynchronously among PGCs at embryonic day (E) 9.5 to E10.5 and accounts for the unique process of imprint erasure. Mechanistically, 5hmC enrichment is followed by its protracted decline thereafter at a rate consistent with replication-coupled dilution. The conversion to 5hmC is an important component of parallel redundant systems that drive comprehensive reprogramming in PGCs. Nonetheless, we identify rare regulatory elements that escape systematic DNA demethylation in PGCs, providing a potential mechanistic basis for transgenerational epigenetic inheritance.

Specification of primordial germ cells (PGCs) from epiblast cells at ~embryonic day (E) 6.25 is linked with extensive epigenetic reprogramming—including global DNA demethylation, chromatin reorganization, and imprint erasure—that is vital for generating totipotency (*1*, *2*). The erasure of CpG methylation (5mC) is a key component of this program, but the dynamics and underlying mechanisms of the process remain unclear (*3*). Here, we report a comprehensive analysis of PGCs by combining immunofluorescence, genome-wide 5-(hydroxy)methylcytosine DNA immunoprecipitation sequencing [(h)meDIP-seq], single-cell RNA sequencing (RNA-seq), bisulfite sequencing, and functional analyses to address the mechanistic basis of epigenetic reprogramming in PGCs.

We investigated *Tet* expression by using single-cell RNA-seq, which revealed that *Tet1* and *Tet2* are expressed in PGCs and peak between E10.5 and E11.5 but that *Tet3* is undetectable (Fig. 1A). Immunofluorescence (IF) showed that TET1 and TET2 are nuclear and expressed at significantly higher levels in PGCs than in neighboring somatic cells between E9.5 and E11.5 (Fig. 1B and figs.

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK. [2]Wellcome Trust/Medical Research Council Stem Cell Institute, University of Cambridge, Cambridge, UK. [3]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK.

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

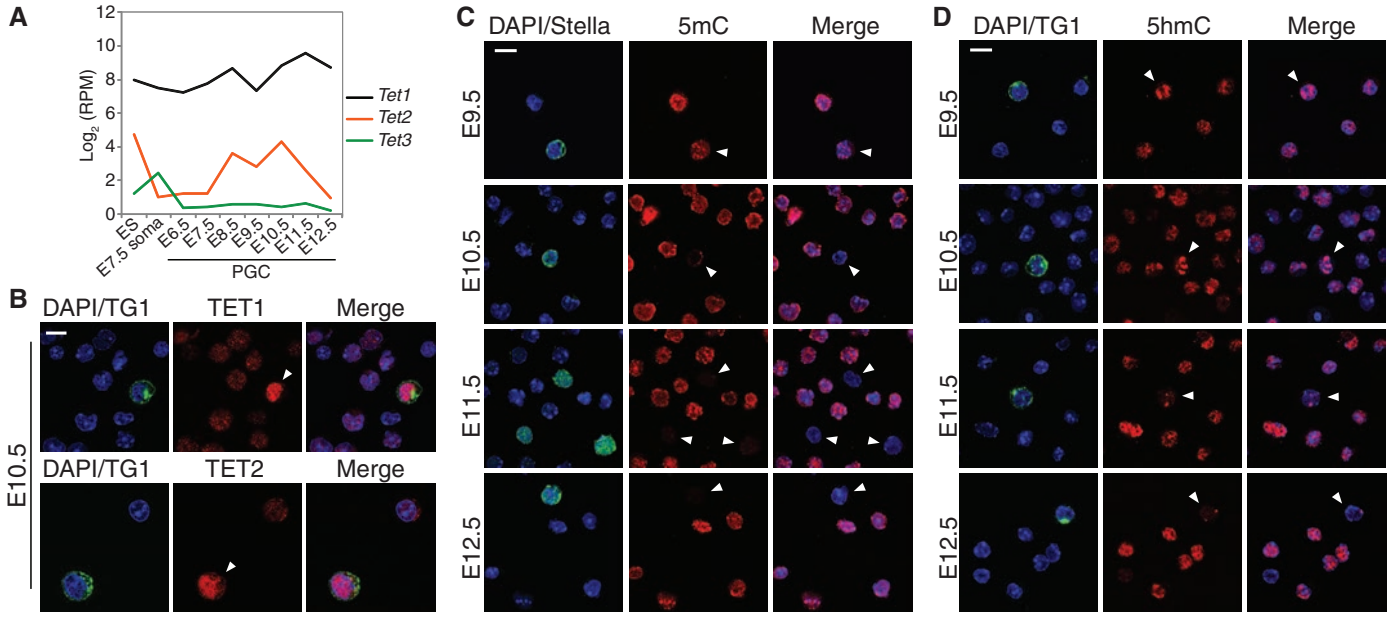


Fig. 1. Global dynamics of 5mC, 5hmC, and TETs in PGCs. (**A**) Single-cell RNA-seq analysis of *Tet1*, *Tet2*, and *Tet3* expression. Shown is $\log_2$ reads per million (RPM). (**B**) Expression of TET1 and TET2 in E10.5 PGCs (arrowheads) and soma. (**C**) Dynamics of DNA methylation (5mC) in PGCs shows 5mC erasure between E9.5 and E11.5. (**D**) Kinetics of 5hmC in PGCs. TG1/STELLA mark PGCs. Scale bars indicate 10 μm.

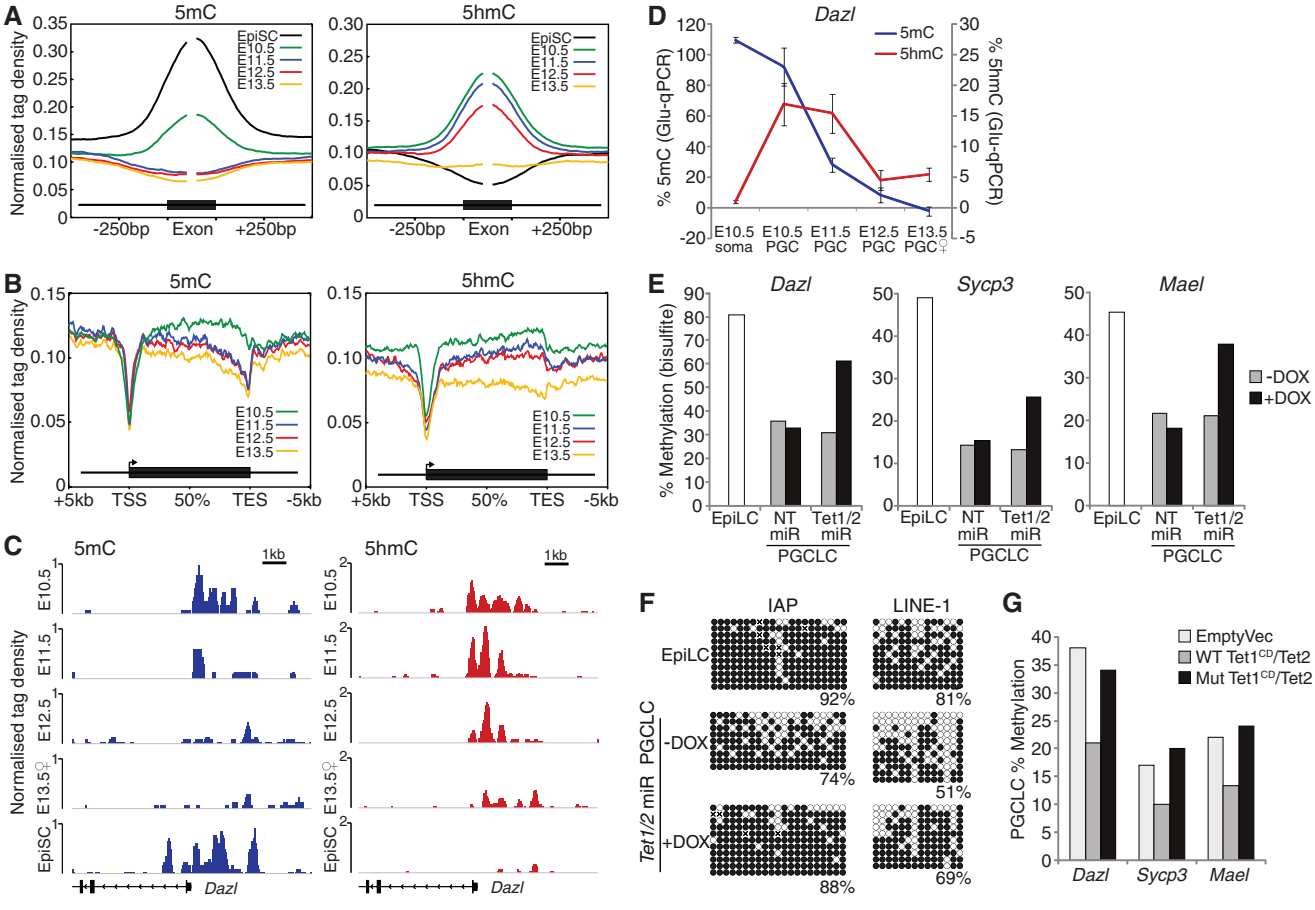


Fig. 2. Erasure of 5mC is coupled to 5hmC conversion. (**A**) Enrichment of 5mC and 5hmC in E10.5 to E13.5 PGCs and EpiSCs over internal exons. (**B**) Distributions of 5mC and 5hmC relative to a metagene. TSS, transcription start site; TES, transcription end site. (**C**) Profiles of 5mC (blue) and 5hmC (red) at the *Dazl* promoter. (**D**) Glu-qPCR showing quantitative levels of 5mC and 5hmC at a CCGG site in the *Dazl* promoter. Error bars represent SEM. (**E** and **F**) DNA methylation (%) by bisulfite sequencing of –DOX or +DOX *Tet1/Tet2* miR or NT miR PGCLCs at (E) gene promoters and (F) repeat elements. Open and solid circles represent unmethylated and methylated CpGs, respectively. (**G**) DNA methylation in PGCLCs stably expressing catalytically active (WT) or mutant (Mut) TET1 and TET2.

S1 and S2). This suggests that erasure of 5mC in PGCs could occur through conversion to 5hmC by TET1 and TET2 together (*4*, *5*).

We pursued this possibility by IF and found a progressive reduction of 5mC in PGCs between E9.5 and E10.5, until it became undetectable by E11.5 (Fig. 1C). The loss of 5mC occurs concurrently with a global enrichment of 5-hydroxymethylcytosine (5hmC) in PGCs between E9.5 and E10.5, suggesting a genome-scale conversion of 5mC to 5hmC (Fig. 1D). The global conversion to 5hmC initiates asynchronously among PGCs from E9.5, perhaps reflecting developmental heterogeneity (figs. S3 to S5). Indeed, TET1 up-regulation also initially occurs in a subset of PGCs from E9.5, which apparently also exhibit lower 5mC signal (fig. S6). In contrast to soma and embryonic stem (ES) cells (*6*), we observed that 5hmC exhibited a distinct localization in PGCs that coincided with 4′,6-diamidino-2-phenylindole (DAPI)–dense chromocenters, indicating that the conversion of 5mC to 5hmC includes heterochromatic satellite regions (fig. S7). The enrichment of 5hmC in PGCs at E10.5 is followed by its progressive reduction, suggesting that 5hmC is an intermediate toward demethylation to unmodified cytosine (C) (Fig. 1D). We checked whether 5hmC is subsequently converted to 5-formylcytosine or 5-carboxycytosine but found no detectable enrichment of these derivatives in PGCs (fig. S8) (*7*).

To gain further insight into the dynamics of 5mC to 5hmC conversion, we performed meDIP-seq and hmeDIP-seq in E10.5 to E13.5 PGCs (fig. S9). Because before E10.5 PGCs were highly limiting, we also profiled epiblast stem cells (EpiSCs), which are derived from the same post-implantation epiblast as nascent PGCs, and embryonic soma (E10.5) as references (fig. S10). Unlike bisulfite sequencing, our approach distinguishes between 5mC and 5hmC but generates a relative rather than a quantitative measure of modifications (*6*). We therefore initially examined exonic sequences, which are highly methylated and thus exhibit an informative dynamic range of relative (h)meDIP signal when they become demethylated. We found significantly reduced 5mC in E10.5 PGCs relative to EpiSC and soma and erasure by E11.5 (Fig. 2A and figs. S11 to S13). The loss of 5mC in PGCs is paralleled by a strong exonic enrichment of 5hmC, indicating 5mC-to-5hmC conversion (Fig. 2A and fig. S11). Once 5mC is converted to 5hmC, it is set on a pathway toward demethylation, because there are no 5hmC maintenance mechanisms (*6*). Consistent with this, 5hmC undergoes a progressive depletion during PGC development, which is delayed relative to loss of 5mC (Fig. 2, A and B).

Next, we examined methylation-dependent genes such as *Dazl*, which are activated by promoter demethylation in PGCs (*8*, *9*), and observed strong 5hmC enrichment coincident with loss of 5mC at their promoters (Fig. 2C and fig. S14). We confirmed that 5mC erasure is coupled to 5hmC enrichment at the *Dazl* promoter quantitatively, by using the glucosyltransferase–quantitative polymerase chain reaction (Glu-qPCR) assay (Fig. 2D). RNA-seq revealed that transcriptional activation of *Dazl* and other methylation-dependent germline genes initiates at E9.5 and increases progressively until ~E11.5 (fig. S15). This represents an important functional readout of the timing of DNA demethylation in PGCs.

To functionally link 5hmC to DNA demethylation, we used in vitro PGC-like cells (PGCLC). PGCLCs are specified from epiblast-like cells (EpiLCs) and exhibit the fundamental properties of migratory PGCs in vivo, including global DNA demethylation and chromatin reorganization (fig. S10) (*10*). TET1 and TET2 are both active in PGCs, so we generated PGCLCs carrying a doxycycline (DOX)–inducible compound microRNA (miR) knockdown of *Tet1* and *Tet2* (T-KD). We

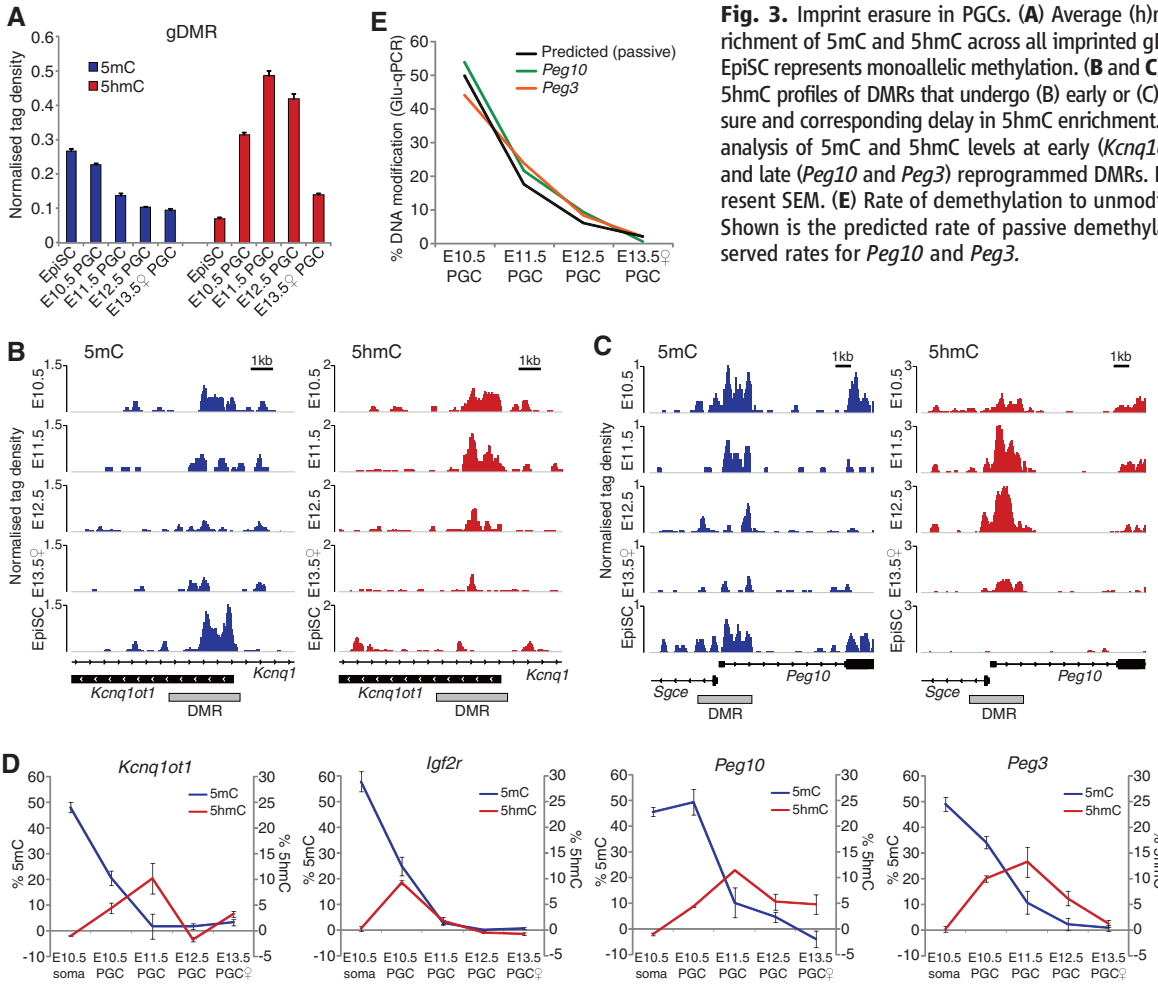


Fig. 3. Imprint erasure in PGCs. (**A**) Average (h)meDIP-seq enrichment of 5mC and 5hmC across all imprinted gDMRs in PGCs. EpiSC represents monoallelic methylation. (**B** and **C**) The 5mC and 5hmC profiles of DMRs that undergo (B) early or (C) late 5mC erasure and corresponding delay in 5hmC enrichment. (**D**) Glu-qPCR analysis of 5mC and 5hmC levels at early (*Kcnq1ot1* and *Igf2r*) and late (*Peg10* and *Peg3*) reprogrammed DMRs. Error bars represent SEM. (**E**) Rate of demethylation to unmodified cytosine. Shown is the predicted rate of passive demethylation and observed rates for *Peg10* and *Peg3*.

found that genes known to be demethylated in PGCs in vivo (*8*) also underwent DNA demethylation upon specification of control uninduced (–DOX) T-KD PGCLCs and in nontargeting (NT) miR PGCLCs (+/–DOX). In contrast, induction of *Tet1/Tet2* miR (+DOX) resulted in a substantial inhibition of DNA demethylation in PGCLCs but did not reduce the efficiency of their specification (Fig. 2E and fig. S16). Knockdown of *Tet1/Tet2* also inhibited DNA demethylation at long interspersed nuclear element 1 (LINE-1) sites and prevented the limited erasure of 5mC that occurs at intracisternal-A-particles (IAP) (Fig. 2F). These findings are important considering that both the maintenance and de novo DNA methylation systems are repressed in PGCs and PGCLCs (*10*), which likely accounts for some direct passive demethylation. Moreover, constitutive overexpression of catalytically active, but not catalytic mutant, TET1 and TET2 in PGCLCs promoted 5mC erasure to a greater extent (Fig. 2G). Thus, TET-mediated 5hmC conversion is a key event toward DNA demethylation in PGCs.

The reprogramming of gonadal PGCs in vivo uniquely entails the complete erasure of genomic imprints (*11*). Analysis of imprinted gametic differentially methylated regions (gDMRs) ($n = 21$) in PGCs revealed that erasure of 5mC is coupled to a significant increase of 5hmC enrichment (Fig. 3A). However, the precise timing of 5mC erasure is imprinted locus-specific. For example, the DMRs at *Kcnq1ot1* and *Igf2r* exhibit loss of 5mC by E10.5 relative to EpiSC (which represent ~50% allelic 5mC) and erasure by E11.5 (Fig. 3B), whereas *Peg10* and *Peg3* remain methylated until E11.5 (Fig. 3C and fig. S17). Moreover, *Kcnq1ot1* and *Igf2r* are enriched in 5hmC by E10.5, whereas 5hmC enrichment at *Peg10* and *Peg3* is delayed until E11.5, suggesting that conversion to 5hmC follows a defined temporal order at imprinted DMRs, which dictates the timing of demethylation in PGCs. Indeed, we observed that other genomic regions also exhibited differential onset of 5mC erasure (compare *Peg10* DMR versus exon, Fig. 3C). Glu-qPCR analysis confirmed that the *Peg10* and *Peg3* DMRs maintained 5mC levels of 50% and 34%, respectively, in E10.5 PGCs, whereas *Kcnq1ot1* and *Igf2r* DMRs were already reduced to 21% and 25%, respectively (Fig. 3D). Glu-qPCR also established the quantitative enrichment of 5hmC at imprinted DMRs in PGCs. The cumulative data suggest that conversion of 5mC to 5hmC by TET1 and TET2 is a general mechanism for the erasure of imprints in PGCs.

Conversion of 5mC to 5hmC at exons, promoters, and gDMRs in PGCs was followed by a protracted period of progressive 5hmC depletion between E11.5 and E13.5 (Figs. 2, A to D, and 3), suggesting a replication-coupled process (*12*). This prompted us to examine the rate of DNA demethylation between E10.5 and E13.5 quantitatively by using Glu-qPCR. Because demethylation commences asynchronously in PGCs, it is necessary to examine loci that have not initiated substantial 5mC erasure by E10.5, such as *Peg10* and *Peg3*. Because PGCs have an estimated cell cycle of ~16 hours between E10.5 and E13.5 (*13*), we would predict a reduction of DNA modification of ~threefold per 24 hours (1.5 population doublings) if the process is coupled to DNA replication. We observed that the rate of demethylation at *Peg10* ($P = 0.0022$) and *Peg3* ($P = 0.0019$) fits highly significantly with the predicted rate (Fig. 3E), suggesting that 5hmC may be removed from these loci by replication-coupled dilution. We obtained similar results for the *Dazl* promoter ($P = 0.0014$).

We next asked whether any promoters or regulatory elements can escape the comprehensive 5mC reprogramming in PGCs. We screened for CpG islands (CGI) that remain methylated in female PGCs at E13.5, because these cells represent the lowest point of global demethylation (fig. S18) (*14*). We identified 11 CGIs with significant 5mC enrichment in E13.5 PGCs (figs. S19 and S20). Validation by bisulfite sequencing showed that the promoter CGIs of *Vmn2r29* and *Sfi1* and the exonic CGI of *Srrm2* were all methylated in PGCs at E10.5 and maintained CpG methylation throughout reprogramming (Fig. 4A).

To define the extent of 5mC erasure at single-base resolution, we performed whole-genome bisulfite sequencing (WGBS), which revealed that global CpG methylation is reduced to 2.2% in female E13.5 PGCs (fig. S21). However, we identified 4730 loci that escape demethylation (>40% 5mC) in PGCs, which are predominately repeat associated (>95%). Resistant loci predominantly correspond to IAP elements, but the IAPLTR1 subclass is significantly more methylated than any other (fig. S22). IAPLTR1 is the most active and hence hazardous IAP subclass to genomic integrity, suggesting specific systems are mobilized to maintain 5mC at IAPLTR1 during reprogramming to protect genome stability (*15*). We were unable to determine any unique sequence characteristics of the 233 single-copy loci with >40% 5mC, suggesting that positional context or chromatin structure may contribute to their escape from reprogramming. Indeed, "escapees" were often adjacent to IAP elements or telomeric regions. Considered with the recent observations that many regulatory elements can evade zygotic 5mC erasure (*16*, *17*), our data suggest that rare but potentially functionally relevant 5mC epialleles could be inherited over multiple generations by evading erasure during both zygotic and PGC reprogramming.

We demonstrate here that comprehensive DNA demethylation in PGCs, including imprint erasure, entails conversion of 5mC to 5hmC, likely redundantly by TET1 and TET2. In vivo 5hmC conversion initiates asynchronously in PGCs between E9.5 and E10.5 and is largely complete by E11.5. The rate of progressive decline of 5hmC thereafter, both globally and at single-copy loci, is consistent with a replication-dependent mechanism of demethylation toward unmodified cytosines (Fig. 4B). In parallel to 5hmC conversion, repression of the de novo (*Dnmt3a/b*) and maintenance (*Uhrf1*) DNA methylation systems in PGCs prevents cyclical remethylation and simultaneously renders PGCs permissive for direct

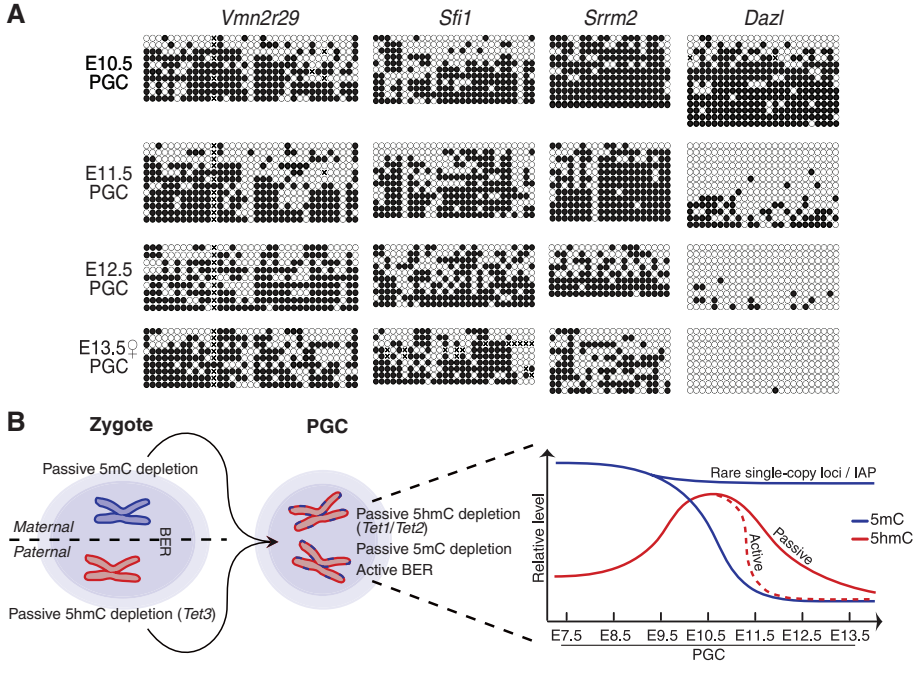


Fig. 4. Inheritance of 5mC through reprogramming. (**A**) The *Vmn2r29*, *Sfi1*, and *Srrm2* CGIs escape reprogramming in PGCs. Open and solid circles represent unmethylated and methylated CpGs, respectively. *Dazl* is representative of demethylation at most loci. (**B**) Model for the mechanisms and dynamics of DNA demethylation in PGCs.

passive 5mC depletion (fig. S23) (*18*), which may contribute to the partial demethylation observed in *Tet1* and *Tet2* knockdown PGCLCs. Thus, whereas in zygotes 5mC reprogramming is mechanistically compartmentalized into TET3-mediated 5hmC conversion of the paternal genome and direct passive 5mC depletion on the maternal genome (*12*, *19*–*21*), both of these mechanisms operate together in PGCs (Fig. 4B). In addition, up-regulation of the base excision repair (BER) pathway in PGCs may both protect against cumulative genetic damage and act as an auxiliary active demethylation mechanism, perhaps for specific loci (*22*, *23*). Reprogramming in PGCs therefore involves multiple redundant mechanisms to reset the epigenome for totipotency, which accounts for the apparent fertility (albeit subfertile) of mice lacking individual components, such as *Tet1* (*24*). The existence of multiple mechanisms may also underpin the comprehensive nature of DNA demethylation in PGCs (*3*). Nonetheless, some rare single-copy sites of CpG methylation escape from 5mC erasure (*25*), which may provide mechanistic avenues for investigations into transgenerational epigenetic inheritance.

References and Notes

1. M. A. Surani, K. Hayashi, P. Hajkova, *Cell* **128**, 747 (2007).
2. P. Hajkova *et al.*, *Nature* **452**, 877 (2008).
3. J. A. Hackett, J. J. Zylicz, M. A. Surani, *Trends Genet.* **28**, 164 (2012).
4. M. Tahiliani *et al.*, *Science* **324**, 930 (2009).
5. S. Ito *et al.*, *Nature* **466**, 1129 (2010).
6. G. Ficz *et al.*, *Nature* **473**, 398 (2011).
7. S. Ito *et al.*, *Science* **333**, 1300 (2011); 10.1126/science.1210597.
8. D. M. Maatouk *et al.*, *Development* **133**, 3411 (2006).
9. J. A. Hackett *et al.*, *Development* **139**, 3623 (2012).
10. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, *Cell* **146**, 519 (2011).
11. P. Hajkova *et al.*, *Mech. Dev.* **117**, 15 (2002).
12. A. Inoue, Y. Zhang, *Science* **334**, 194 (2011); 10.1126/science.1212483.
13. P. P. Tam, M. H. Snow, *J. Embryol. Exp. Morphol.* **64**, 133 (1981).
14. C. Popp *et al.*, *Nature* **463**, 1101 (2010).
15. C. Qin *et al.*, *Mol. Carcinog.* **49**, 54 (2010).
16. S. A. Smallwood *et al.*, *Nat. Genet.* **43**, 811 (2011).
17. J. Borgel *et al.*, *Nat. Genet.* **42**, 1093 (2010).
18. K. Kurimoto *et al.*, *Genes Dev.* **22**, 1617 (2008).
19. T. P. Gu *et al.*, *Nature* **477**, 606 (2011).
20. M. Wossidlo *et al.*, *Nat. Commun.* **2**, 241 (2011).
21. K. Iqbal, S. G. Jin, G. P. Pfeifer, P. E. Szabó, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3642 (2011).
22. P. Hajkova *et al.*, *Science* **329**, 78 (2010).
23. S. Cortellino *et al.*, *Cell* **146**, 67 (2011).
24. M. M. Dawlaty *et al.*, *Cell Stem Cell* **9**, 166 (2011).
25. S. Guibert, T. Forné, M. Weber, *Genome Res.* **22**, 633 (2012).

Acknowledgments: We thank N. Miller for fluorescence-activated cell sorting analysis, F. Tang and W. Tang for experimental support, and G-L. Xu for reagents. This work was funded by the Wellcome Trust (RG49135, RG44593, and 083563) and the Human Frontier Science Program. Sequencing data have been deposited in the Sequence Read Archive (SRA060914).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1229277/DC1
Materials and Methods
Figs. S1 to S24
References (*25*, *26*)

24 August 2012; accepted 16 November 2012
Published online 6 December 2012;
10.1126/science.1229277

Actin, Spectrin, and Associated Proteins Form a Periodic Cytoskeletal Structure in Axons

Ke Xu,[1]* Guisheng Zhong,[1]* Xiaowei Zhuang[1,2]†

Actin and spectrin play important roles in neurons, but their organization in axons and dendrites remains unclear. We used stochastic optical reconstruction microscopy to study the organization of actin, spectrin, and associated proteins in neurons. Actin formed ringlike structures that wrapped around the circumference of axons and were evenly spaced along axonal shafts with a periodicity of ~180 to 190 nanometers. This periodic structure was not observed in dendrites, which instead contained long actin filaments running along dendritic shafts. Adducin, an actin-capping protein, colocalized with the actin rings. Spectrin exhibited periodic structures alternating with those of actin and adducin, and the distance between adjacent actin-adducin rings was comparable to the length of a spectrin tetramer. Sodium channels in axons were distributed in a periodic pattern coordinated with the underlying actin-spectrin–based cytoskeleton.

Actin plays critical roles in shaping and maintaining cell morphology, as well as in supporting various cellular functions, including cell motility, cell division, and intracellular transport (*1*). In neurons, actin is essential for the establishment of neuronal polarity, cargo transport, neurite growth, and stabilization of synaptic structures (*2*–*4*). Despite its importance, our understanding of actin structures in neurons remains incomplete. Electron microscopy has shown detailed actin ultrastructure in growth cones and dendritic spines (*5*, *6*), in which actin is the dominant cytoskeletal protein, but little is known about the organization of actin in the axonal and dendritic shafts (*4*). These neurites contain a high density of different types of cytoskeletal filaments, such as microtubules and neurofilaments (*6*–*8*). Hence, the challenge of resolving the organization of actin in axons and dendrites requires imaging tools with both high spatial resolution and molecular specificity.

A prototypical actin-spectrin–based cytoskeleton structure is found in red blood cells (erythrocytes) (*9*, *10*), where actin, spectrin, and associated proteins form a two-dimensional (2D) polygonal network (mostly composed of hexagons and pentagons) underneath the erythrocyte membrane (*11*, *12*). Spectrin analogs have been found in many other animal cells (*9*, *10*), including neurons (*13*, *14*). These analogs play important roles, ranging from regulation of the heartbeat to stabilization of axons, formation of axon initial segments and nodes of Ranvier, and stabilization of synapses in neurons (*9*, *10*, *15*). An erythrocyte-like, polygonal lattice structure has been observed for spectrin in the *Drosophila* neuromuscular junction (*16*), and models similar to the erythrocyte cytoskeleton have also been proposed for other systems (*10*). However, the ultrastructural organization of spectrin in nonerythrocyte cells is largely unknown due to similar challenges in imaging.

Recent advances in superresolution fluorescence microscopy (*17*, *18*) allow resolutions down to ~10 nm to be achieved with molecular specificity, providing a promising solution to the above challenges. In particular, superresolution studies of neurons have provided valuable structural and dynamic information of actin in dendritic spines (*19*–*22*). In this work, we used a superresolution fluorescence imaging method, stochastic optical reconstruction microscopy (STORM) (*23*–*27*), to study the 3D ultrastructural organization of actin and spectrin in neurons.

To image actin in neurons, we fixed cultured rat hippocampal neurons at various days in vitro (DIV) and labeled actin filaments with phalloidin conjugated to a photoswitchable dye, Alexa Fluor 647 (Invitrogen, Carlsbad, CA) (Fig. 1) (*28*). To identify axons and dendrites, we immunolabeled MAP2, a microtubule-associated protein enriched in dendrites, or NrCAM, a cell adhesion molecule found in the initial segments of axons (*15*), using a dye of a different color (Fig. 1) (*28*). In the conventional fluorescence images (Fig. 1, A, D, and F), MAP2 specifically stained dendrites, and NrCAM specifically labeled the initial segments of axons, whereas actin was found in both dendrites and axons.

Next, we used 3D STORM (*27*) with a dual-objective astigmatism-imaging scheme (*28*, *29*)

[1]Howard Hughes Medical Institute (HHMI), Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [2]Department of Physics, Harvard University, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: zhuang@chemistry.harvard.edu

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Cell Biology

X chromosome inactivation: silencing, topology and reactivation

Teresa Robert Finestra and Joost Gribnau

To ensure X-linked gene dosage compensation between females (XX) and males (XY), one X chromosome undergoes X chromosome inactivation (XCI) in female cells. This process is tightly regulated throughout development by many different factors, with *Xist* as a key regulator, encoding a long non-coding RNA, involved in establishment of several layers of repressive epigenetic modifications. Several recent studies on XCI focusing on identification and characterization of *Xist* RNA-protein interactors, revealed new factors involved in gene silencing, genome topology and nuclear membrane attachment, amongst others. Also, new insights in higher order chromatin organization have been presented, revealing differences between the topological organization of active and inactive X chromosomes (Xa and Xi), with associated differences in gene expression. Finally, further evidence indicates that the inactive state of the Xi can be (partially) reversed, and that this X chromosome reactivation (XCR) might be associated with disease.

Address
Department of Developmental Biology, Erasmus MC, Wytemaweg 80, Rotterdam CN 3015, The Netherlands

Corresponding author: Gribnau, Joost (j.gribnau@erasmusmc.nl)

Current Opinion in Cell Biology 2017, **46**:54–61

This review comes from a themed issue on **Cell nucleus**

Edited by **Aaron F. Straight** and **Roland Foisner**

For a complete overview see the Issue and the Editorial

Available online 23rd February 2017

http://dx.doi.org/10.1016/j.ceb.2017.01.007

Introduction

In mammals, dosage compensation of X-linked gene expression between females (XX) and males (XY) is achieved through inactivation of one X chromosome in female somatic cells [1]. Two waves of X chromosome inactivation (XCI) are observed in mice, the most well-characterized XCI model system. First, the paternal X chromosome is inactivated at the 2–4 cell stage of embryonic development in female mice. This imprinted form of XCI (iXCI) is reversed in the inner cell mass, and after implantation at E5.5 random XCI (rXCI) is initiated, leading to silencing of one X chromosome with either paternal or maternal origin. After establishment of XCI, the silenced inactive X chromosome (Xi) is clonally transmitted through a near infinite number of cell divisions (reviewed in Ref. [2]).

A key regulator of XCI is the X-inactive-specific transcript (*Xist*) gene located on the X chromosome [3–5]. *Xist* encodes a long non-coding RNA (lncRNA) that, upon initiation of XCI, is monoalellically upregulated from the future Xi. *Xist* coats the X chromosome in *cis* and recruits various proteins involved in transcriptional silencing, such as chromatin remodeling factors and the DNA methylation machinery, that establish the inactive state (reviewed in Ref. [6]). *Xist* is located in the X inactivation center (Xic), which harbors *Tsix*, another lncRNA, that is transcribed antisense to *Xist*. *Tsix* in mice fully overlaps *Xist* and is involved in its repression. Despite the robustness of the process, 12–20% and 3–7% of X-linked genes escape XCI in human and mouse, respectively. Escaping genes can either escape in all cell types (constitutive escapees), or in a specific context (facultative escapees) (reviewed in Ref. [7]).

Here we review the recent advancements in understanding the newly identified *Xist* interacting proteins involved in silencing, the importance of chromosome topology in gene transcriptional regulation and X Chromosome Reactivation (XCR) after establishment of the inactive state.

Spreading of silencing along the X

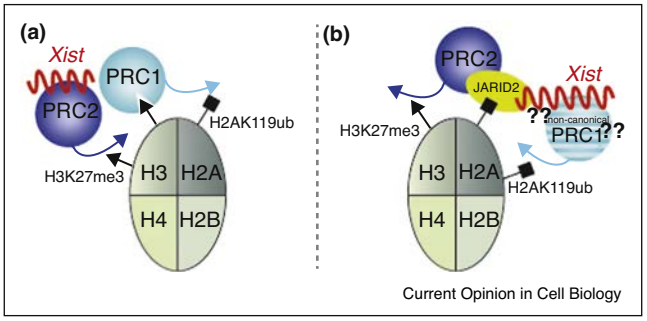
Xist has two main roles throughout establishment of the Xi, it spreads along the future Xi and recruits various proteins that via different epigenetic mechanisms render the X chromosome silent. Recently, several studies have significantly increased the number of known *Xist*-interacting proteins, that were determined using different oligonucleotide pulldown methods [8••,9••,10••]. The identified *Xist* interactors are involved in various functions, such as transcriptional silencing, chromosome architecture and nuclear membrane attachment. All studies identified SPEN (also known as SHARP) as a *Xist* interacting-protein, in accordance with two genetic screens, consisting of a forward genetic haploid screen [11] and a shRNA screen [12], aimed to identify primary X silencing factors. These studies showed that SPEN is required for transcriptional repression [8••,11,12], via interaction with the *Xist* A-repeat [9••,11], an *Xist* region that is essential for gene silencing [13–16]. SPEN might interact with Smart, an associated repressor protein, to recruit HDAC3, leading to histone deacetylation and to PolII exclusion

from the Xi territory. Also, evidence suggests that SPEN and HDAC3 are required for recruitment of Polycomb Repressive Complex 2 (PRC2) [8••] (reviewed in [17]). The exact chronology of events that lead to silencing is still controversial. It was initially proposed that PRC2 via direct interaction with the A-repeat catalyzes accumulation of tri-methylation of lysine 27 in histone H3 (H3K27me3) [15], and this modification leads to PRC1 recruitment catalyzing monoubiquitination of lysine 119 in histone H2A (H2AK119ub1) [18–20]. However, none of the oligonucleotide pulldown screens revealed an interaction of *Xist* with core-PRC2 proteins. In contrast, several canonical and non-canonical PRC1 subunits were identified, suggesting a model were first PRC1 proteins are recruited, which in turn might lead to PRC2 recruitment [21••,22,23] (Figure 1). This recruitment might involve the PRC2 cofactor JARID2, through binding H2AK119ub1 [23,24], but may also be facilitated by interaction of JARID2 with *Xist* [25]. Furthermore, the *Xist*-interactor pulldown studies also identified proteins related to late epigenetic changes in XCI, including the histone variant macroH2A [26,27] and proteins involved in CpG methylation [28,29].

Among other interesting proteins interacting with *Xist* were Lamin-B receptor (LBR), an element of the nuclear lamina, and RBM15, involved in N6-adenosine methylation (m6A). *Xist* contains many m6A RNA methylation sites, and a recent in-depth study characterizing the role of RBM15, its close homologue, RBM15B, and, the N6-adenosine-methyltransferase, METTL3, indicate that these proteins are involved in catalyzing m6A methylation of *Xist* RNA. In turn m6A methylated *Xist* is recognized by the m6A reader YTHDC1, leading to gene silencing through an unknown mechanism [30••]. The Xi preferentially locates at the nuclear periphery and the nucleolus [31–35], in agreement with the evidence that heterochromatin is found in the same locations (reviewed in Ref. [36]). The identification of LBR as an *Xist* interacting protein, together with the evidence that the Xi chromosome is associated with the nuclear envelope, suggested a mechanism where both, LBR and *Xist* are needed to position the Xi in proximity to the nuclear lamina and induce structural changes leading to proper *Xist* spreading and, in turn, gene silencing [37•]. Moreover, the Xi is often located at the periphery of the nucleolus. Recent evidence suggests that Firre, another X-linked lncRNA, has a role in anchoring the Xi to the nucleolus [38•].

Other than recruiting silencing factors, *Xist* has the ability to spread along the Xi. High-resolution *Xist* localization studies provided relevant insights on *Xist* spreading during initiation of XCI through two models of action. *Xist* RNA, transcribed from a nuclear transcriptionally active compartment, uses the three-dimensional conformation of the X chromosome to initially spread to gene-rich regions on the X chromosome, that undergo transcriptional silencing [39]. Then these genes relocate to a silent nuclear compartment, allowing the spreading of *Xist* to the rest of the chromosome via proximity transfer. In a different model *Xist* spreads via a two-step mechanism during XCI initiation, first targeting gene-rich regions followed by gene-poor regions [40]. Both studies suggest that, LINE repeats are not enriched in *Xist* binding sites, contrasting with the hypothesis that *Xist* utilizes LINE repeats to spread along the X chromosome [41–44], and suggesting that LINEs might not be the direct targets of *Xist*. The idea that *Xist* might use the spatial conformation of the X chromosome to spread and silence the future Xi [39,45], highlights the importance of the X chromosome topology in XCI.

Figure 1

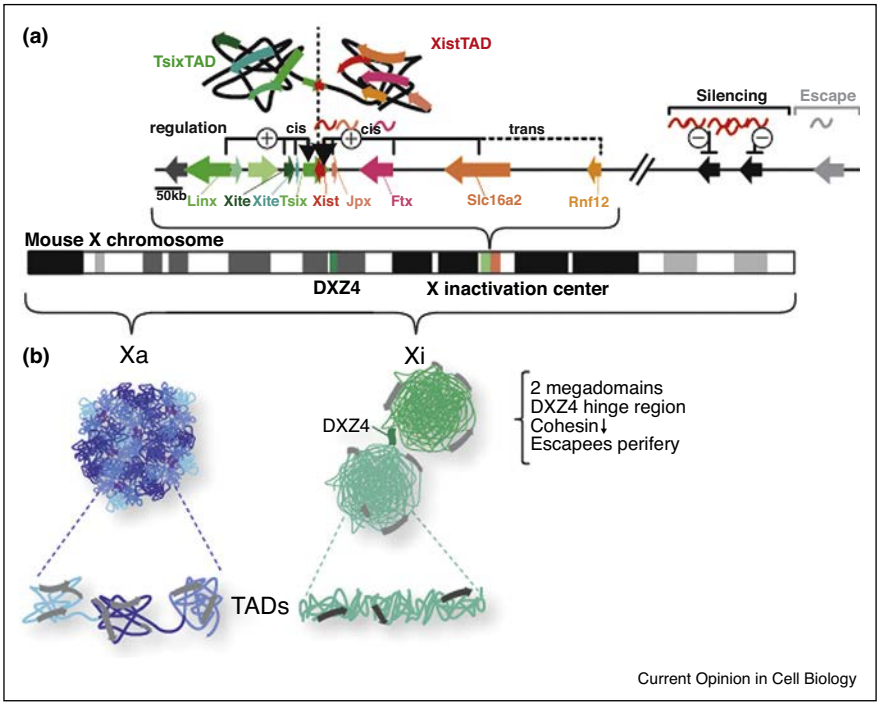


Models of *de novo* Polycomb proteins recruitment to the Xi.
(a) The classical model for *de novo* Polycomb recruitment proposes that PRC2, recruited to the Xi via *Xist*, establishes H3K27me3. This repressive mark is recognized by one of the subunits of PRC1, that in turn catalyzes H2AK119ub1.
(b) New evidence suggests that specific non-canonical PRC1 complexes interact with *Xist* [9••], establish H2AK119ub1 independently of H3K27me3 [88] and H2AK119ub1 promotes recruitment of PRC2, leading to H3K27me3 [21••,22,23]. Recruitment of PRC2 might be facilitated by JARID2 recognizing H2AK119ub1 [23,24], or may involve a mechanism independent of PRC1 through a JARID2-Xist interaction [25].

X chromosome topology, revealing the ArXISTecture

The development of genome capture technologies, from the initial Chromosome Conformation Capture (3C) technology to its posterior variants (4C, 5C, HiC, etc.), has been key to highlight the interdependence of genome topology and transcription regulation (reviewed in Ref. [46]). XCI represents a unique biological example in which the active and inactive form of the same chromosome can be compared to retrieve information on the interplay between genome topology and gene expression and identify its main players. The development of the HiC technology [47], led to the identification of Topologically Associating Domains (TADs) [48,49], as regions of the genome with more frequent physical interactions inside its boundaries, than with regions outside.

Figure 2



The transcriptional organization of the X chromosome inactivation center (Xic) and chromosomal architecture differences between the Xa and Xi. **(a)** The X chromosome inactivation center (Xic) is predominantly divided in two neighboring TADs with its boundary located between the *Xist* and *Tsix* promoter. In one TAD, elements that promote XCI via *Xist* transcription (red) are located, in the other TAD, elements that inhibit XCI via *Tsix* transcription (green) are present. Upon differentiation, the *Xist* TAD becomes actively transcribed leading to *Xist* spreading, recruitment of silencing factors and X-linked gene silencing (black), nonetheless some genes escape XCI (grey).
(b) The Xi shows an overall loss of TADs (green), in comparison to the Xa (blue). The Xi is organized in two mega-domains united via a hinge region, where the microsatellite gene DXZ4 is located. Genes that escape XCI locate in the periphery of the Xi territory. Moreover, cohesins, key players in shaping genome topology, seem to be depleted from the Xi.

Interestingly, the *Xist* and *Tsix* promoters are located in neighboring TADs together with their own *cis*-regulators, and deletion of this boundary region has two central consequences (Figure 2A), disruption of the local TAD-folding pattern and loss of long range transcriptional control [48]. This indicates a common transcriptional regulatory mechanism within these TADs and antagonistic roles between them, where elements that either activate or inhibit XCI are located. Studies addressing the topological differences between the Xi and Xa indicate that the Xi displays a more disorganized structure and lacks long-range interactions, an organization fundamentally established and maintained by *Xist* [50,51•]. Moreover, models derived from these genome topology studies show that escapees are located at the periphery of the Xi territory, in accordance to previous imaging studies [52–54]. The Xi shows an overall loss of TADs and the presence of two mega-domains that are separated by a hinge region, where the microsatellite repeat DXZ4 is located [51,55,56•,57••] (Figure 2B). DXZ4 codes for a lncRNA, that is differentially transcribed and regulated between the Xi and the Xa [58,59]. Deletion of this hinge region, both before and after initiation of XCI, leads to fusion of the two mega-domains [56•,57••]. However, only the deletion of this region before XCI initiation leads to the loss of chromatin accessibility and silencing of facultative escapees [57••], thus highlighting the relation between initiation of XCI, Xi topology and gene expression, although, the exact mode of action of DXZ4 remains unknown. Interestingly, high resolution microscopic analysis of the two X territories suggested a model where both the Xi and the Xa are organized in two segments, which in turn, are organized in smaller domains that are in closer proximity in the Xi, compared to the Xa [60]. The proximity positioning of these small domains-specific of the Xi, might translate at the HiC experiment level, into the previously explained two mega-domains conformation.

The chromatin insulator CTCF and cohesins are often located at boundaries of TADs and regulate gene expression by creating long range interactions between *cis*-acting elements [49,61–63]. CTCF and cohesins were identified as *Xist* interacting proteins in one screen, and cohesins were found to be depleted from the Xi [10••]. Interestingly, deletion of *Xist* from the Xi, resulted in restoration

of cohesin binding and TAD structure reminiscent of the Xa. These findings suggest a model were *Xist* is continuously required to repel elements that control chromosome topology, such as cohesin, to reduce physical contacts. Despite the general loss of TADs on the Xi, genes that escape XCI show local enrichment of CTCF binding sites [64], involved in preservation of the local TAD organization [57••]. Nonetheless, it is important to consider that HiC topology contacts are retrieved from cell populations. Single-cell HiC [65] and *in silico* studies using polymer modeling [66], provided new insights in the expected heterogeneity of nuclear architecture at the single cell level. All these recent findings highlight the central role for *Xist* in shaping Xi topological changes, partially associated with X-linked transcriptional silencing. Thus indicating that topological changes modulate gene expression, likely in a tight interplay where changes in gene expression lead to topological changes.

X chromosome reactivation after establishment of silencing

During female development maintaining two active X chromosomes is lethal [67,68], indicating that tight regulation of X-linked gene expression is essential. Interestingly, a recent study, involving Sox2-Cre mediated deletion of *Xist* in peri-implantation embryos challenges this view [69•]. Although *Xist* knockout mice were born at a lower than expected frequency and displayed silencing defects, dosage compensation was still observed in the absence of *Xist*. This study might implicate that loss of *Xist* mediated random XCI is less deleterious than anticipated, but could also be related to delayed onset of Cre expression around the time the Xi is established [70]. Cells with two active X chromosomes might be selected against, and the presence of cells with partial XCI or loss of XCI after *Xist* deletion might explain this unexpected phenotype.

Previous studies examining loss of *Xist* after Xi establishment indicated that reactivation is very rare in cultured fibroblasts and X-reactivation only occurs in the presence of histone deacetylases (HDAC) and DNA methylation inhibitors [71]. Although this indicates that many epigenetic layers are in place to prevent X-reactivation, loss of *Xist* in differentiating cells has a profound effect on chromosome topology of the Xi, refolding into a structure that resembles the Xa [10••,50]. In addition, an RNA interference screen identified several factors promoting X-reactivation, by partly affecting *Xist* expression. Interestingly, this study indicated that defective XCI via deletion of the XCI factor stanniocalcin 1 (Stc1), does not result in global over-expression of X-linked genes [72], in agreement with studies examining *Xist* deletion in female mouse embryos [69•]. **These findings indicate** the presence of compensatory mechanisms, as suggested in cells derived from Down syndrome patients, where only a subset of chromosome 21 genes was found to be significantly over-expressed [73].

Xi hyper mutation is frequently observed in different types of cancer [74]. In breast cancer, for example, chromosome abnormalities and the disappearance of the Barr body are frequently found [75–77], pointing toward the potential use of *Xist* and the Xi status as biomarkers for breast cancer [78,79•]. In BRCA1 deficient cancers, loss of the Xi and duplication of the Xa seems to be related to genomic instability [80–82]. However for other types of tumors, reactivation of X-linked genes has been reported to coincide with aberrant distribution of epigenetic modifications associated with the Xi [79•]. Loss of *Xist* in the murine hematopoietic lineage results in a highly penetrant hematologic phenotype [83]. Interestingly, the hematopoietic lineage might be more susceptible to X reactivation than other cell lineages, as this is the only lineage capable of initiating XCI in adult tissues [84], thus highlighting the plasticity of XCI in this linage. Additionally, mature naïve T and B cells during lymphocyte maturation display a change in *Xist* localization and loss of Xi associated chromatin marks and gene silencing, which is reversed upon lymphocyte activation. This reestablishment of the *Xist* coated Xi is mediated by the transcription factor YY1 [85], in agreement with a study that proposed a role for YY1 in tethering *Xist* to the Xi [86]. As, YY1 has also been implicated in *Xist* transcription regulation [87], follow-up experiments will be important to discriminate between effects on localization and transcription of *Xist*. Nevertheless, these studies highlight the unanticipated plasticity of XCI with potential consequences to disease.

Concluding remarks

The inactive X-chromosome undergoes chromosome-wide epigenetic and topological changes, involved in silencing of most X-linked genes. Recent publications provide new insights in the *Xist* interacting proteins, leading to the identification of new silencing factors involved in XCI. Nonetheless, additional research is needed to further characterize the exact role many of these proteins and the time window they act, in silencing and/or maintenance of silencing during XCI. Moreover, new insights in high order chromatin structure show that Xi topology is related to X-linked gene silencing and regulation of escaping genes. Future studies, examining topological changes on the Xi over time, in relation to gene expression and CTCF and cohesin binding, would provide valuable insights into the dynamics of XCI. The interplay between topology and gene expression remains unclear, representing a fundamental intriguing subject that needs further investigation. Lastly, in adult tissues, gene expression, three-dimensional organization and the over-all Xi epigenetic status are essential to maintain the Xi silent. These factors are also involved in XCR, nonetheless, it is still unclear to which extend loss of these

factors can individually or collectively lead to XCR. Further insights in this topic are specially relevant in the light of the efforts that nowadays are put into partially or totally reactivating the inactive X chromosome in order to treat X-linked diseases in women, such as Rett syndrome.

Acknowledgements

We would like to thank Sarra Merzouk and John W. M. Martens for critical reading of the manuscript and valuable feedback. We would also like to apologize to all scientists whose work has not been mentioned due to space constraints.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Lyon MF: **Gene action in the X-chromosome of the mouse (*Mus musculus* L.) [Internet]**. *Nature* 1961, **190**:372-373.

2. Schulz EG, Heard E: **Role and control of X chromosome dosage in mammalian development [Internet]**. *Curr Opin Genet Dev* 2013, **23**:109-115.

3. Borsani G, Tonlorenzi R, Simmler M, Dandolo L, Arnaud D, Capra V, Grompe M, Pizzuti A, Muzny D, Lawrence C *et al.*: **Characterization of a murine gene expressed from the inactive X chromosome**. *Nature* 1991, **354**:56-58.

4. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S: **The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus**. *Cell* 1992, **71**:515-526.

5. Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, Willard HF: **The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus**. *Cell* 1992, **71**:527-542.

6. Wutz A: **Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. [Internet]**. *Nat Rev Genet* 2011, **12**:542-553.

7. Balaton BP, Brown CJ: **Escape artists of the X chromosome [Internet]**. *Trends Genet* 2016, **32**:348-359.

8. •• McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A *et al.*: **The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3**. *Nature* 2015, **521**.

Using RAP-MS (RNA Antisense Purification followed by Mass Spectrometry) this study identified 10 proteins that specifically bind *Xist* in male cells with a doxycycline inducible system in its endogenous locus. Identifying SPEN, LBR and hnRNPK as *Xist* interactors involved in gene transcriptional silencing.

9. •• Chu C, Zhang QC, Da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY: **Systematic discovery of Xist RNA binding proteins**. *Cell* 2015, **161**:404-416.

Using ChIRP-MS (Comprehensive identification of RNA binding Proteins by Mass Spectrometry) this study identified 81*Xist* interacting proteins in male mouse ESC harboring a doxycycline inducible *Xist* transgene in chromosome 11. From all the identified *Xist* binding proteins, the authors show that PRC1 subunits (Ring1b, Rybp, Pcgf5) but not PRC2 subunits bind *Xist*. They also show that hnRNPK (also known as SAF-A) is required for establishment of histone modification marks and SPEN is involved in gene silencing via binding the A repeat.

10. •• Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Bernhard Payer, Myriam Boukhali WH, Lee JT: **A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation**. *Science (80-)* 2015:316.

This study identified 204*Xist* interacting proteins using iDRiP (identification of direct RNA interacting proteins) in female mouse fibroblasts. Identifying proteins involved in various functions, amongst them proteins involved in genome topology organization, including CTCF and members of the cohesin complex, such as SMC1a, SMC3 and RAD21. The authors show that deletion of *Xist* from the Xi restores cohesin binding and chromosome architecture, resembling the Xa.

11. Monfort A, Di Minin G, Postlmayr A, Freimann R, Arieti F, Thore S, Wutz A: **Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells**. *Cell Rep* 2015, **12**:554-561.

12. Moindrot B, Cerase A, Coker H, Masui O, Grijzenhout A, Pintacuda G, Schermelleh L, Nesterova TB, Brockdorff N: **A pooled shRNA screen identifies Rbm15 Spen, and Wtap as factors required for Xist RNA-mediated silencing [Internet]**. *Cell Rep* 2015, **12**:562-572.

13. Wutz A, Rasmussen TP, Jaenisch R: **Chromosomal silencing and localization are mediated by different domains of Xist RNA [Internet]**. *Nat Genet* 2002, **30**:167-174.

14. Kohlmaier A, Savarese F, Lachner M, Martens J, Jenuwein T, Wutz A: **A chromosomal memory triggered by Xist regulates histone methylation in X inactivation**. *PLoS Biol* 2004, **2**.

15. Zhao J, Sun BK, Erwin JA, Song J, Lee JT: **Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome**. *Science (80-)* 2008, **322**:750-756.

16. Royce-Tolland ME, Andersen AA, Koyfman HR, Talbot DJ, Wutz A, Tonks ID, Kay GF, Panning B: **The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. [Internet]**. *Nat Struct Mol Biol* 2010, **17**:948-954.

17. Mira-Bontenbal H, Gribnau J: **New Xist-interacting proteins in X-chromosome inactivation [Internet]**. *Curr Biol* 2016, **26**:R338-R342.

18. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones R, Zhiang Y: **Role of histone H3 lysine 27 methylation in Polycomb group silencing**. *Science (80-)* 2002, **298**:1039-1044.

19. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS: **Hierarchical recruitment of polycomb group silencing complexes**. *Mol Cell* 2004, **14**:637-646.

20. de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R, Nesterova TB, Silva J, Otte AP, Vidal M *et al.*: **Polycomb group proteins ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation**. *Dev Cell* 2004, **7**:663-676.

21. •• Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LLP, Ito S, Cooper S, Kondo K, Koseki Y *et al.*: **Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation [Internet]**. *Cell* 2014, **157**:1445-1459.

Using a *de novo* targeting system in mouse embryonic stem cells, the authors show that H2AK119Ub catalyzed via non-canonical PRC1 proteins (PCGF1, 2 and 3) leads to recruitment of PRC2 and deposition of H3K27me3. In contrast to the previously proposed hierarchical model, that suggests that H3K27me3 deposition is followed by H2AK119Ub

22. Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, DeMarco V, Elderkin S, Koseki H, Klose R *et al.*: **Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment [Internet]**. *Cell Rep* 2014, **7**:1456-1470.

23. Kalb R, Latwiel S, Baymaz HI, Jansen PWTC, Müller CW, Vermeulen M, Müller J: **Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression [Internet]**. *Nat Struct Mol Biol* 2014, **21**:569-571.

24. Cooper S, Grijzenhout A, Underwood E, Ancelin K, Zhang T, Nesterova TB, Anil-Kirmizitas B, Bassett A, Kooistra SM, Agger K *et al.*: **Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2 [Internet]**. *Nat Commun* 2016, **7**:13661.

25. da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, Sanulli S, Chow J, Schulz E, Picard C *et al.*: **Jarid2 is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome**. *Mol Cell* 2014, **53**: 301-316.

26. Costanzi C, Stein P, Worrad DM, Schultz RM, Pehrson JR: **Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. [Internet]**. *Development* 2000, **127**:2283-2289.

27. Mietton F, Sengupta AK, Molla A, Picchi G, Barral S, Heliot L, Grange T, Wutz A, Dimitrov S: **Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. [Internet]**. *Mol Cell Biol* 2009, **29**:150-156.

28. Kaslow DC, Migeon BR: **DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. [Internet]**. *Proc Natl Acad Sci U S A* 1987, **84**:6210-6214.

29. Lock LF, Takagi N, Martin GR: **Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation**. *Cell* 1987, **48**:39-46.

30. •• Patil DP, Chen C-K, Pickering BF, Chow A, Jackson C, Guttman M, Jaffrey SR: **m6A RNA methylation promotes XIST-mediated transcriptional repression [Internet]**. *Nature* 2016, **537**:1-25.

Xist retains at least 78 m6A residues that are involved in transcriptional X-linked gene silencing. Primarily, this study shows that knockdown of different proteins involved in m6A establishment (RBM15, RBM15B, METTL3) affects gene silencing

31. Barr M, Carr D: **Correlations between sex chromatin and sex chromosomes**. *Acta Cytol* 1962, **6**:34-45.

32. Belmont AS, Bignone F, Ts'O POP: **The relative intranuclear positions of barr bodies in XXX non-transformed human fibroblasts**. *Exp Cell Res* 1986, **165**:165-179.

33. Puck TT, Johnson R: **DNA exposure and condensation in the X and 21 chromosomes**. *Stem Cells* 1996, **14**:548-557.

34. Zhang LF, Huynh KD, Lee JT: **Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing**. *Cell* 2007, **129**:693-706.

35. Rego A, Sinclair PB, Tao W, Kireev I, Belmont AS: **The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. [Internet]**. *J Cell Sci* 2008, **121**:1119-1127.

36. Padeken J, Heun P: **Nucleolus and nuclear periphery: velcro for heterochromatin [Internet]**. *Curr Opin Cell Biol* 2014, **28**:54-60.

37. • Chen C-K, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, McDonel P, Cerase A, Guttman M: **Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. [Internet]**. *Science (80-.)* 2016, **354(6311)**:468-472 http://dx.doi.org/10.1126/science.aae0047.

This study shows that LBR directly interacts with *Xist* facilitating localization to the nuclear membrane and spreading along the Xi.

38. • Yang F, Deng X, Ma W, Berletch JB, Rabaia N, Wei G, Moore JM, Filippova GN, Xu J, Liu Y *et al.*: **The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation [Internet]**. *Genome Biol* 2015, **16**:52.

The lncRNA Firre helps position the Xi close to the nucleolus via specific CTCF binding, in contrast to the Xa. Moreover, Firre is required for H3K27me3 maintenance.

39. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES *et al.*: **The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. [Internet]**. *Science (80-)* 2013, **341**:1237973.

40. Simon MD, Pinter SF, Fang R, Sarma K, Rutenberg-Schoenberg M, Bowman SK, Kesner BA, Maier VK, Kingston RE, Lee JT: **High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation [Internet]**. *Nature* 2013, **504**:465-469.

41. Lyon MF: **X-chromosome inactivation: a repeat hypothesis**. *Cytogenet Cell Genet* 1998, **80**:133-137.

42. Bailey JA, Carrel L, Chakravarti A, Eichler EE: **Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. [Internet]**. *Proc Natl Acad Sci U S A* 2000, **97**:6634-6639.

43. Tannan NB, Brahmachary M, Garg P, Borel C, Alnefaie R, Watson CT, Simon Thomas N, Sharp AJ: **Dna methylation profiling in X;autosome translocations supports a role for L1 repeats in the spread of X chromosome inactivation**. *Hum Mol Genet* 2014, **23**:1224-1236.

44. Peeters SB, Cotton AM, Brown CJ: **Variable escape from X-chromosome inactivation: Identifying factors that tip the scales towards expression**. *BioEssays* 2014, **36**:746-756.

45. Marks H, Kerstens HHD, Barakat TS, Splinter E, Dirks RAM, van Mierlo G, Joshi O, Wang S-Y, Babak T, Albers CA *et al.*: **Dynamics of gene silencing during X inactivation using allele-specific RNA-seq [Internet]**. *Genome Biol* 2015, **16**:1-20.

46. Denker A, De Laat W: **The second decade of 3C technologies: detailed insights into nuclear organization**. *Genes Dev* 2016, **30**:1357-1382.

47. Lieberman-aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **33292**:289-294.

48. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J *et al.*: **Spatial partitioning of the regulatory landscape of the X-inactivation centre**. *Nature* 2012, **485**:381-385.

49. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions [Internet]**. *Nature* 2012, **485**:376-380.

50. Splinter E, de Wit E, Nora EP, Klous P, van de Werken HJG, Zhu Y, Kaaij LJT, van Ijcken W, Gribnau J, Heard E *et al.*: **The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA**. *Genes Dev* 2011, **25**:1371-1383.

51. • Deng X, Ma W, Ramani V, Hill A, Yang F, Ay F, Berletch JB, Blau CA, Shendure J, Duan Z *et al.*: **Bipartite structure of the inactive mouse X chromosome. [Internet]**. *Genome Biol* 2015, **16**:121-152.

Using allele-specific DNAse HiC in female mouse cells, this study proposes that the Xi is organized in two domains separated by a hinge region where the DXZ4 microsatllite is located. In human this bipartite structure is present, nonetheless the genomic content of these domains differs from mouse, although the hinge region is conserved. This hinge region binds CTCF and nucleolar proteins.

52. Dietzel S, Schiebel K, Little G, Edelmann P, Rappold GA, Eils R, Cremer C, Cremer T: **The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. [Internet]**. *Exp Cell Res* 1999, **252**:363-375.

53. Clemson CM, Hall LL, Byron M, McNeil J, Lawrence JB: **The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences**. *Proc Natl Acad Sci U S A* 2006, **103**:7688-7693.

54. Chaumeil J, Le Baccon P, Wutz A, Heard E: **A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced**. *Genes Dev* 2006, **20**:2223-2237.

55. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping [Internet]**. *Cell* 2014, **159**:1665-1680.

56. • Darrow EM, Huntley MH, Dudchenko O, Stamenova EK, Durand NC, Sun Z, Huang S-C, Sanborn AL, Machol I, Shamim M, et al: **Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture [Internet]**. *Proc Natl*

Acad Sci 2016, **113(31)**:E4504-12 http://dx.doi.org/10.1073/pnas.1609643113.
Using HiC, this work shows that the two megadomain organization of the Xi is conserved amongst eutherian mammals. Disruption of the hinge region in differentiated ESCs disrupt the two megadomain organization of the Xi, but has no effect on gene silencing or escape

57. •• Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E *et al.*: **Structural organization of the inactive X chromosome in the mouse [Internet]**. *Nature* 2016, **535**:575-579.
Using allele-specific HiC, RNA-Seq, ATAC-Seq, Chip-Seq, in combination with DNA FISH in female mouse ESCs, the authors study the Xi topological conformation and its implications in gene transcription. Showing that the Xi is organized in two megadomains separated by a hinge region, where DXZ4 is located. Deletion of this region affects the expression of constitutive escapees

58. Giacalone J, Friedes J, Francke U: **A novel GC-rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes. [Internet]**. *Nat Genet* 1992, **1**:137-143.

59. Chadwick BP: **DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts**. *Genome Res* 2008, **18**:1259-1269.

60. Teller K, Illner D, Thamm S, Casas-Delucchi CS, Versteeg R, Indemans M, Cremer T, Cremer M: **A top-down analysis of Xa- and Xi-territories reveals differences of higher order structure at ≥20 Mb genomic length scales**. *Nucleus* 2011, **2**:465-477.

61. Sofueva S, Yaffe E, Chan W-C, Georgopoulou D, Vietri Rudan M, Mira-Bontenbal H, Pollard SM, Schroth GP, Tanay A, Hadjur S: **Cohesin-mediated interactions organize chromosomal domain architecture. [Internet]**. *EMBO J* 2013, **32**:3119-3129.

62. Zuin J, Dixon JR, van der Reijden MIJA, Ye Z, Kolovos P, Brouwer RWW, van de Corput MPC, van de Werken HJG, Knoch TA, van IJcken WFJ *et al.*: **Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells [Internet]**. *Proc Natl Acad Sci U S A* 2014, **111**:996-1001.

63. Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S: **Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture [Internet]**. *Cell Rep* 2015, **10**:1297-1309.

64. Filippova GN, Cheng MK, Moore JM, Truong JP, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM: **Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development**. *Dev Cell* 2005, **8**:31-42.

65. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P: **Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. [Internet]**. *Nature* 2013, **502**:59-64.

66. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, Tiana G, Heard E: **Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription**. *Cell* 2014, **157**:950-963.

67. Takagi N, Abe K: **Detrimental effects of two active X chromosomes on early mouse development [Internet]**. *Development* 1990, **109**:189-201.

68. Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R: **Xist-deficient mice are defective in dosage compensation but not spermatogenesis**. *Genes Dev* 1997, **11**:156-166.

69. • Yang L, Kirby JE, Sunwoo H, Lee JT: **Female mice lacking Xist RNA show partial dosage compensation and survive to term**. *Genes Dev* 2016, **30**:1747-1760 http://dx.doi.org/10.1101/gad.281162.116.GENES.
This study shows that Sox2-Cre directed knock-out of *Xist* in embryonic tissues of female mice, gives rise to live born mice that die before weaning age. X-linked gene silencing is affected but present, suggesting *Xist*-independent gene dosage compensation mechanisms

70. Minocha S, Sung TL, Villeneuve D, Lammers F, Herr W: **Compensatory embryonic response to allele-specific inactivation of the murine X-linked gene Hcfc1 [Internet]**. *Dev Biol* 2016, **412**:1-17.

71. Csankovszki G, Nagy a, Jaenisch R: **Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in . . . [Internet]**. *J Cell Biol* 2001, **153**:773-783.

72. Bhatnagar S, Zhu X, Ou J, Lin L, Chamberlain L, Zhu LJ, Wajapeyee N, Green MR: **Genetic and pharmacological reactivation of the mammalian inactive X chromosome. [Internet]**. *Proc Natl Acad Sci U S A* 2014, **111**:12591-12598.

73. Aït Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G, Rossier J, Personnaz L, Creau N, Bléhaut H *et al.*: **Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. [Internet]**. *Am J Hum Genet* 2007, **81**:475-491.

74. Jäger N, Schlesner M, Jones DTW, Raffel S, Mallm JP, Junge KM, Weichenhan D, Bauer T, Ishaque N, Kool M *et al.*: **XHypermutation of the inactive X chromosome is a frequent event in cancer**. *Cell* 2013, **155**:567-581.

75. Moore KL, Barr ML: **The sex chromatin in benign tumours and related conditions in man. [Internet]**. *Br J Cancer* 1955, **9**:246-252.

76. Liao DJ, Du Q, Yu BW, Grignon D, Sarkar FH: **Novel perspective: focusing on the X chromosome in reproductive cancers**. *Cancer Invest* 2003, **21**:641-658.

77. Pageau GJ, Hall LL, Lawrence JB: **BRCA1 Does not paint the inactive X to localize XIST RNA but may contribute to broad changes in cancer that impact XIST and Xi heterochromatin**. *J Cell Biochem* 2007, **100**:835-850.

78. Rottenberg S, Vollebergh MA, De Hoon B, De Ronde J, Schouten PC, Kersbergen A, Zander SAL, Pajic M, Jaspers JE, Jonkers M *et al.*: **Impact of intertumoral heterogeneity on predicting chemotherapy response of BRCA1-deficient mammary tumors**. *Cancer Res* 2012, **72**:2350-2361.

79. • Chaligné R, Popova T, Mendoza-Parra MA, Saleem MAM, Gentien D, Ban K, Piolot T, Leroy O, Mariani O, Gronemeyer H *et al.*: **The inactive X chromosome is epigenetically unstable and transcriptionally labile in breast cancer**. *Genome Res* 2015, **25**:488-503.
This study characterizes the Xi of breast cancer cells lines and primary tumors, showing that they are epigenetically unstable, have an abnormal three-dimensional organization and display aberrant distribution of histone modification marks and DNA methylation

80. Sirchia SM, Ramoscelli L, Grati FR, Barbera F, Coradini D, Rossella F, Porta G, Lesma E, Ruggeri A, Radice P *et al.*: **Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells**. *Cancer Res* 2005.

81. Vincent-Salomon A, Ganem-Elbaz C, Manié E, Raynal V, Sastre-Garau X, Stoppa-Lyonnet D, Stern MH, Heard E: **X inactive-specific transcript RNA coating and genetic instability of the X chromosome in BRCA1 breast tumors**. *Cancer Res* 2007, **67**:5134-5140.

82. Xiao C, Sharp JA, Kawahara M, Davalos AR, Difilippantonio MJ, Hu Y, Li W, Cao L, Buetow K, Ried T *et al.*: **The XIST noncoding RNA functions independently of BRCA1 in X inactivation**. *Cell* 2007, **128**:977-989.

83. Yildirim E, Kirby JE, Brown DE, Mercier FE, Sadreyev RI, Scadden DT, Lee JT: **Xist RNA is a potent suppressor of hematologic cancer in mice [Internet]**. *Cell* 2013, **152**:727-742.

84. Savarese F, Flahndorfer K, Jaenisch R, Busslinger M, Wutz A: **Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation**. *Mol Cell Biol* 2006, **26**:7167-7177.

85. Wang J, Syrett CM, Kramer MC, Basu A, Atchison ML, Anguera MC: **Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X [Internet]**. *Proc Natl Acad Sci* 2016, **30**:1747-1760 http://dx.doi.org/10.1073/pnas.1520113113.

86. Jeon Y, Lee JT: **YY1 tethers Xist RNA to the inactive X nucleation center [Internet]**. *Cell* 2011, **146**:119-133.

87. Makhlouf M, Ouimette J-F, Oldfield A, Navarro P, Neuillet D, Rougeulle C: **A prominent and conserved role for YY1 in Xist transcriptional activation. [Internet]**. *Nat Commun* 2014, **5**:4878.

88. Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H *et al.*: **RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3**. *Cell* 2012, **148**:664-678.

DNA METHYLATION

The DNA methyltransferase DNMT3C protects male germ cells from transposon activity

Joan Barau,[1] Aurélie Teissandier,[1,2] Natasha Zamudio,[1*] Stéphanie Roy,[3] Valérie Nalesso,[4,5,6] Yann Hérault,[4,5,6] Florian Guillou,[3] Déborah Bourc'his[1†]

DNA methylation is prevalent in mammalian genomes and plays a central role in the epigenetic control of development. The mammalian DNA methylation machinery is thought to be composed of three DNA methyltransferase enzymes (DNMT1, DNMT3A, and DNMT3B) and one cofactor (DNMT3L). Here, we describe the discovery of *Dnmt3C*, a de novo DNA methyltransferase gene that evolved via a duplication of *Dnmt3B* in rodent genomes and was previously annotated as a pseudogene. We show that DNMT3C is the enzyme responsible for methylating the promoters of evolutionarily young retrotransposons in the male germ line and that this specialized activity is required for mouse fertility. DNMT3C reveals the plasticity of the mammalian DNA methylation system and expands the scope of the mechanisms involved in the epigenetic control of retrotransposons.

Genome defense via transcriptional silencing of transposable elements has been proposed to be a driving force for the evolution of DNA methylation (*1*). Retrotransposons occupy half of mammalian genomic space, and their control is of paramount importance in the germ line: Their activity can damage the hereditary material with an impact on fertility and the fitness of subsequent generations (*2*). In mammals, after germline epigenetic reprogramming, small RNA-directed DNA methylation establishes life-long epigenetic silencing of retrotransposons during the perinatal period of spermatogenesis (*3*). Piwi-interacting RNAs (piRNAs) are the cleavage products of retrotransposon transcripts and guide DNA methylation to the promoters of these elements through homology recognition (*4*, *5*). Mammals have specifically evolved a catalytically inactive DNA methyltransferase (DNMT) cofactor, DNMT3L, which acts downstream of the piRNA pathway (*6*, *7*). The inactivation of DNMT3L or PIWI-pathway proteins invariably results in hypomethylation and reactivation of retrotransposons, meiotic failure, azoospermia, and male sterility marked by small testis size (hypogonadism) (*8*).

To gain insights into the biology of retrotransposon silencing in the germ line, we screened a collection of hypogonadal male mice generated through *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis for ectopic retrotransposon activity (fig. S1, A and B). Five independent positive lines were obtained, but all showed linkage to the same genomic interval on chromosome 2 (fig. S1C), suggesting that they shared a spontaneous, ENU-independent mutation. Whole-genome sequencing revealed a de novo insertion of an IAPEz element, a subclass of inracisternal A-particle (IAP) retrotransposon, in the last intron of the *Gm14490* gene (Fig. 1A and fig. S1, D to F). *Gm14490* maps 9 kilobases (kb)

[1]Institut Curie, PSL Research University, INSERM, CNRS, 75005 Paris, France. [2]Ecole des Mines, 75005 Paris, France. [3]PRC, INRA, CNRS, IFCE, Tours University, 37380 Nouzilly, France. [4]Institut de Génétique et Biologie Moléculaire et Cellulaire, Strasbourg University, 67404 Illkirch, France. [5]Centre National de la Recherche Scientifique, UMR7104, Illkirch, France. [6]Institut National de la Santé et de la Recherche Médicale, U964, Illkirch, France.
*Present address: The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia. †Corresponding author. Email: deborah.bourchis@curie.fr

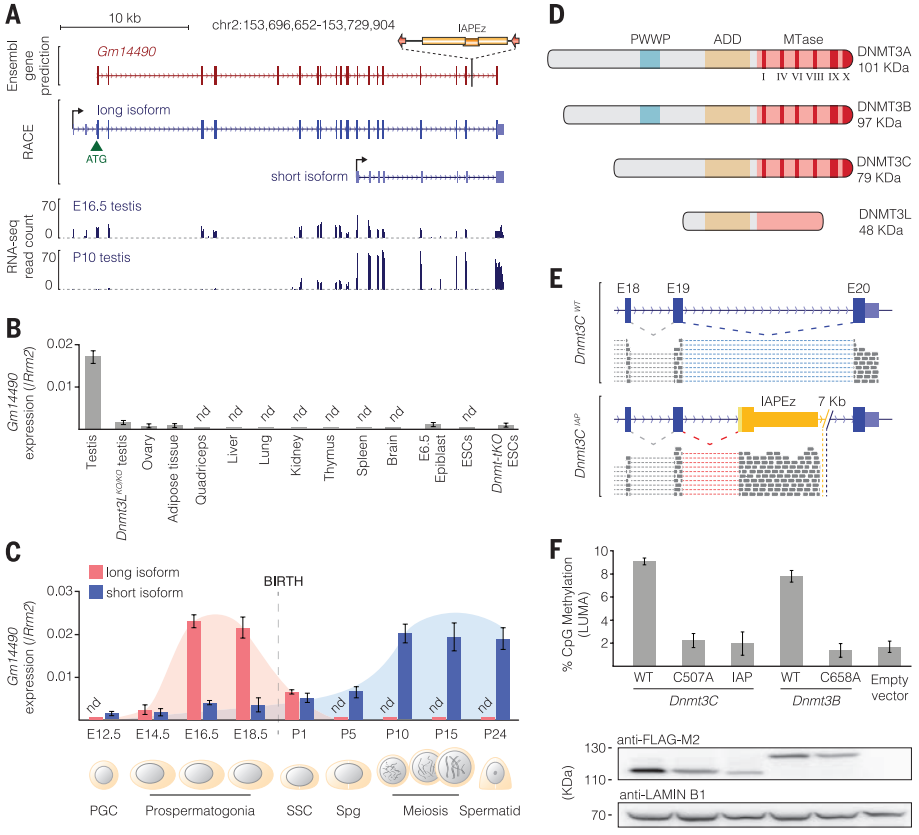


Fig. 1. *Gm14490* encodes a male germ cell–specific de novo DNA methyltransferase, DNMT3C. (**A**) Structure of *Gm14490* (*Ensembl 2011*) and position of the IAPEz insertion (antisense orientation). RACE and RNA-seq analysis of E16.5 and P10 testis identifies a long isoform with coding potential (ATG, green triangle). (**B**) *Gm14490* is detected in testis by reverse transcription–quantitative polymerase chain reaction (RT-qPCR), but not in germ cell–depleted testis from *Dnmt3L*$^{KO/KO}$ animals. Tissues from 10-week-old mice, unless otherwise specified. (**C**) RT-qPCR of the two *Gm14490* isoforms during testis development. Predominant germ cell populations are represented. Primordial germ cells (PGC), spermatogonial stem cells (SSC), and spermatogonia (Spg). (**D**) DNMT3C shows characteristics of DNMT3 proteins—conserved methyltransferase (MTase) motifs and an ADD domain, but no PWWP domain. (**E**) RNA-seq supporting wild-type and mutant *Dnmt3C* ^{IAP} splicing events in E16.5 testis. (**F**) LUminometric methylation assay (LUMA) of global CpG methylation in *Dnmt-tKO* ESCs that transiently express *Dnmt3C*- and *Dnmt3B-3XFLAG* alleles. Data are mean ± SD from three technical replicates in (B) and (C) and from three biological replicates in (F). nd, not detectable.

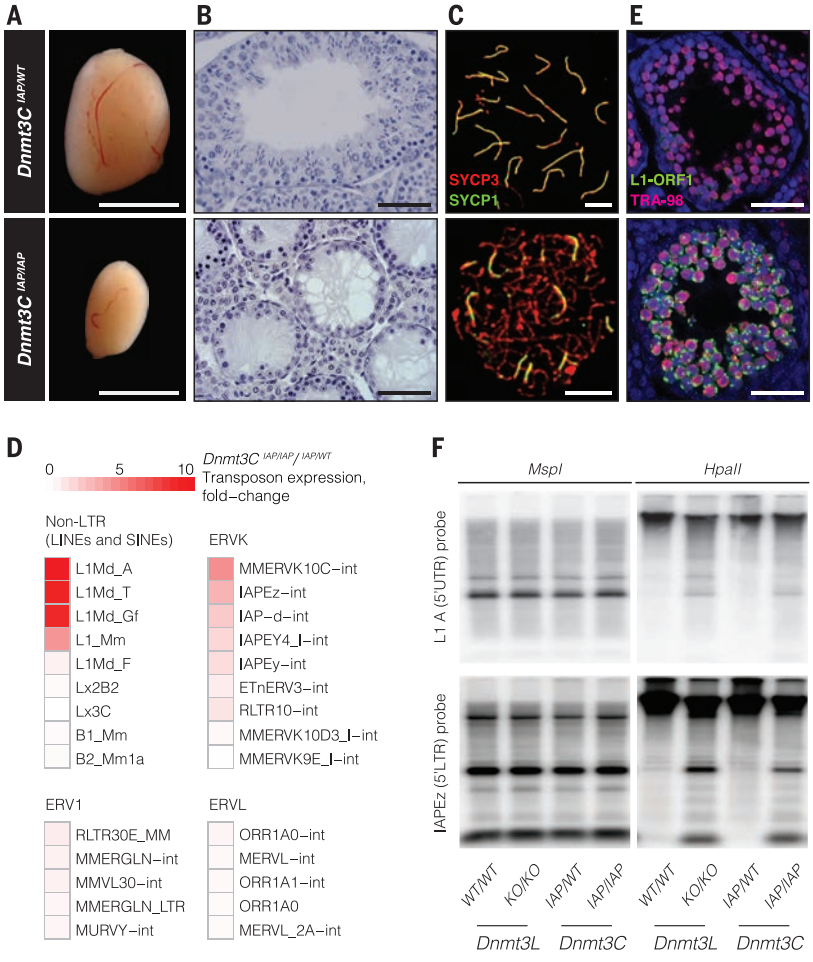


Fig. 2. Phenotype of *Dnmt3C*$^{IAP/IAP}$ males. (**A**) Hypogonadism (6-week-old mice). Scale bars, 5 mm. (**B**) Severe germ cell loss in testis sections (11-week-old mice). Scale bars, 100 μm. (**C**) Impaired chromosome synapsis at meiosis as detected by immunofluorescence against synaptonemal complex proteins (SYCP1 and SYCP3). Scale bars, 4 μm. (**D**) RNA-seq heatmap shows overexpression of young L1 and specific ERVK types in *Dnmt3C*$^{IAP/IAP}$ compared to *Dnmt3C*$^{IAP/WT}$ testis at P20. Annotations from RepeatMasker. (**E**) Aberrant expression of L1-ORF1 proteins in *Dnmt3C*$^{IAP/IAP}$ germ cells (TRA98-positive) at P20. Scale bars, 50 μm. (**F**) L1A-5′UTR and IAPEz-5′LTR are hypomethylated in *Dnmt3C*$^{IAP/IAP}$ testis DNA at P20, as in *Dnmt3L*$^{KO/KO}$ testis. Southern blot analysis after methyl-sensitive Hpa II digestion. Msp I is used as a digestion control.

downstream of the *Dnmt3B* gene and was annotated as a nonfunctional tandem duplication of *Dnmt3B*, based on lack of transcription and recognizable open reading frames (ORFs) (*9*).

We found that *Gm14490* was exclusively expressed in male germ cells (Fig. 1B and fig. S2A). Using RNA sequencing (RNA-seq) and rapid amplification of cDNA ends (RACE), we annotated two transcript isoforms, whose expression was tightly regulated during spermatogenesis (Fig. 1, A and C). The short, noncoding isoform was expressed in postnatal testes. However, the long isoform (2.8 kb) possessed a 709-codon ORF, and its expression sharply peaked around embryonic day 16.5 (E16.5), coinciding with male germline de novo DNA methylation (*5*, *10*). In comparison, *Dnmt3B* is expressed in germ cells, early embryos, and somatic tissues of both sexes (*11*, *12*). Using discriminating primers, we showed independent regulation of *Dnmt3B* and *Gm14490* during spermatogenesis (fig. S2B). The coding potential and specific developmental regulation of *Gm14490* led us to reconsider it as a functional paralog of *Dnmt3B* rather than a pseudogene, and thus we renamed it *Dnmt3C*.

The long *Dnmt3C* isoform encodes a protein with an organization characteristic of DNMT3 enzymes: six methyltransferase motifs (I, IV, VI, VIII, IX, and X) in C-terminal position and an N-terminal ATRX-DNMT3L-DNMT3A (ADD) domain, which binds unmethylated lysine 4 residues of histone H3 (H3K4) (Fig. 1D and fig. S2C) (*13*). However, DNMT3C lacks the Pro-Trp-Trp-Pro (PWWP) domain, which targets DNMT3 proteins to gene bodies through recognition of H3K36 trimethylation (H3K36me3) (*14*, *15*). Overall, DNMT3C exhibits 70% identity with DNMT3B, while DNMT3A and DNMT3B are 46% identical. In hypogonadal mutants, the IAP insertion did not affect *Dnmt3C* transcript levels but provided an alternative splice acceptor site, which led to the exclusion of *Dnmt3C* last exon in favor of the retrotransposon sequence in a chimeric *Dnmt3C-IAP* mRNA (Fig. 1E and fig. S2, D and E). Its predicted translation product lacks motifs IX and X (fig. S2C), which are essential for the AdoMet-dependent methyltransferase fold and for the binding of the methyl donor *S*-adenosyl methionine, respectively (*16*).

To demonstrate that DNMT3C is catalytically active, we performed an in vivo DNA methylation assay. *Dnmt3C* is not expressed in mouse embryonic stem cells (ESCs) (Fig. 1B). By transfecting constructs driving *Dnmt3C* expression in DNA methylation-free ESCs (*Dnmt1*, *Dnmt3A*, and *Dnmt3B* triple-knockout; *Dnmt*-tKO) (*17*), we observed a gain of CpG methylation (10 to 20%), similar to that observed upon *Dnmt3B* transfection (Fig. 1F and fig. S2, F and G). The mutant *Dnmt3C*^{IAP} allele failed to raise CpG methylation levels, as did *Dnmt3C* and *Dnmt3B* mutant alleles with a missense mutation in the catalytic site (DNMT3C C507A and DNMT3B C658A, in which cysteine at position 507 and 658 is replaced by alanine). An in vitro DNA methylation assay using the DNMT3C methyltransferase domain showed concordant results (fig. S2, H and I). These findings demonstrate that DNMT3C is an enzymatically active member of the DNMT3 family of de novo DNA methyltransferases.

Dnmt3C$^{IAP/IAP}$ animals were somatically normal, and only males were sterile (fig. S3A). Hypogonadism was linked to azoospermia with interruption of spermatogenesis at the pachytene stage of meiosis I, in the context of impaired chromosome synapsis (Fig. 2, A to C). The developmental phenotype of *Dnmt3C* mutant mice was similar to that observed in *Dnmt3L*$^{KO/KO}$ males (*7*, *18*), suggesting that DNMT3C could be involved in transposon silencing during spermatogenesis. Indeed, the same set of retrotransposons were up-regulated in *Dnmt3C*$^{IAP/IAP}$ and *Dnmt3L*$^{KO/KO}$ testes at P20 (postnatal day 20) (Fig. 2D and fig. S3B) (*18*). Long interspersed nuclear elements (LINEs or L1s) showed the strongest reactivation, and more specifically evolutionarily young subfamilies: type A, T, and Gf transcripts were increased by 10-fold in *Dnmt3C*$^{IAP/IAP}$ testes, in association with accumulation of L1-encoded ORF1 proteins (Fig. 2E). Among endogenous retroviruses (ERVs), reactivation was specific to some ERVK families (MMERVK10C, IAPEz, and IAPEy). As in the case of the *Dnmt3L* mutation, L1 and IAPEz derepression was linked to a DNA methylation defect in *Dnmt3C*$^{IAP/IAP}$ testes (Fig. 2F), despite normal expression of piRNA/DNA methylation genes and piRNA production during fetal spermatogenesis (fig. S3, C to F). Finally, we confirmed DNMT3C function by generating a *Dnmt3C* knockout mouse through CRISPR-Cas9–mediated deletion (fig. S4A). The *Dnmt3C*^{KO} allele recapitulated the *Dnmt3C*$^{IAP/IAP}$ developmental and molecular phenotypes in homozygous *Dnmt3C*$^{KO/KO}$ males and failed to complement the *Dnmt3C*^{IAP} allele in *Dnmt3C*$^{IAP/KO}$ compound heterozygous males (fig. S4, B to E).

To assess the contribution of DNMT3C to male germline methylation, we performed whole-genome bisulfite sequencing (WGBS) in sorted

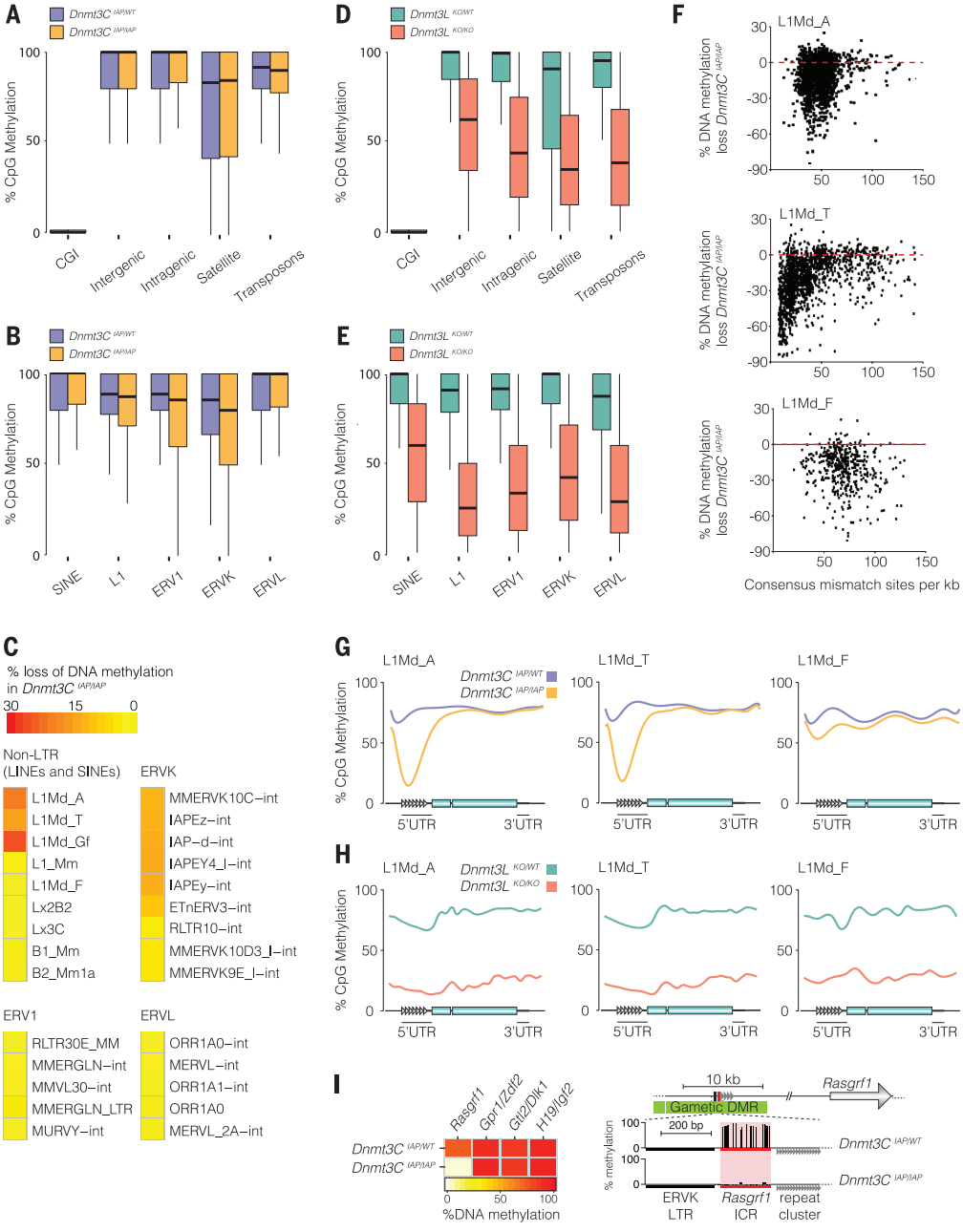


Fig. 3. DNMT3C methylates evolutionarily young retrotransposon promoters. (**A** and **B**) Tukey box-plot representation of CpG methylation content as determined by WGBS over different (A) genomic compartments and (B) retrotransposon classes in control *Dnmt3C* $^{IAP/WT}$ (purple) and mutant *Dnmt3C* $^{IAP/IAP}$ (yellow) germ cells at P10. (**C**) Percentage of DNA methylation loss within individual retrotransposon families in *Dnmt3C*$^{IAP/IAP}$ samples. (**D** and **E**) As in (A) and (B) but for *Dnmt3L*$^{KO/WT}$ (green) and *Dnmt3L*$^{KO/KO}$ (red) germ cells. (**F**) Plotting of DNA methylation loss over individual copies of L1 families according to genetic distance from consensus sequences. (**G** and **H**) Metaplots of DNA methylation over full-length L1s (>5 kb) comparing (G) *Dnmt3C*$^{IAP/IAP}$ and (*H*) *Dnmt3L*$^{KO/KO}$ versus control samples. (**I**) Left: Heatmap of *Dnmt3C*$^{IAP/IAP}$ and control DNA methylation levels across paternally imprinted loci. Right: Methylation maps at the *Rasgrf1* locus.

germ cells from testes at P10, when de novo DNA methylation is completed. Overall CpG methylation levels of *Dnmt3C*$^{IAP/IAP}$ mutant cells were not markedly different from the *Dnmt3C*$^{IAP/WT}$ control (77.7 versus 78.5%). A slight decrease was only apparent when focusing on transposons (81.5 versus 84.2%), and more specifically on LINEs, ERVK, and ERV1 (Fig. 3, A and B). Accordingly, there was only a limited number (264) of differentially methylated regions (DMRs) in *Dnmt3C*$^{IAP/IAP}$ versus *Dnmt3C*$^{IAP/WT}$ germ cells (fig. S5A); all reflected hypomethylation in the mutant, and most overlapped with LINEs (34%) and ERVs (48%). RepeatMasker annotations further highlighted that the same families that were transcriptionally derepressed were hypomethylated in *Dnmt3C*$^{IAP/IAP}$ testes; namely, young L1s and specific ERVKs (Fig. 2D and Fig. 3C). By comparison, deletion of DNMT3L had a stronger and broader impact: *Dnmt3L*$^{KO/KO}$ germ cells exhibited only 39% of global CpG methylation, and all genomic compartments and retrotransposon classes were affected (Fig. 3, D and E, and fig. S5B).

The DNA methylation defect in *Dnmt3C* mutants only reached 30% at the most for young L1s (Fig. 3C). We reasoned that DNMT3C selectively affects transcriptionally active retrotransposon copies, as these are targets of piRNA-dependent DNA methylation during fetal spermatogenesis (*4*, *5*). Indeed, individual L1-A and -T elements with a 5′ promoter (length >5 kb) and the highest similarity toward the consensus sequence showed the greatest DNA methylation loss in *Dnmt3C*$^{IAP/IAP}$ cells (Fig. 3F and fig. S4C). Additionally, DNMT3C-dependent DNA methylation was not evenly distributed across the length of these transcriptionally competent elements, but rather focalized to their promoters [5′ untranslated region (UTR)] (Fig. 3G), in a pattern previously observed in piRNA-deficient *Mili*$^{KO/KO}$ males (fig. S5D) (*4*). Older L1-F elements did not show such trends (Fig. 3F and fig. S5C). By contrast, *Dnmt3L* deficiency caused demethylation of L1s independently of their age and throughout their sequence (Fig. 3H).

ERVKs exhibited milder methylation loss (10%) than LINEs in *Dnmt3C*$^{IAP/IAP}$ cells (Fig. 3C), and this was not related to the size or the sequence conservation of individual copies (fig. S5, C and E). ERVKs partially resist the genome-wide erasure of methylation that occurs in the

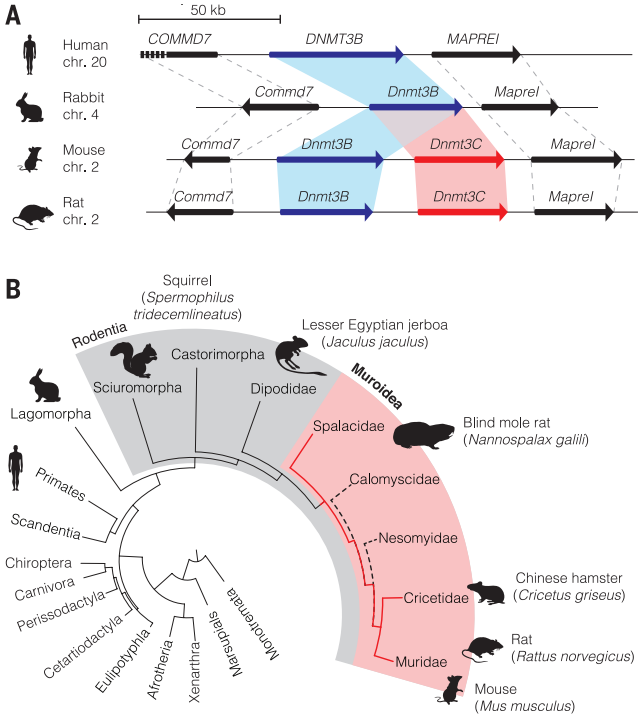


Fig. 4. Phylogenetic distribution of *Dnmt3C*. (**A**) Locus organization and synteny around the *Dnmt3B* (blue) and *Dnmt3C* (red) genes in human, rabbit, mouse, and rat genomes. (**B**) *Dnmt3C* distribution in the mammalian phylogenetic tree with a focus on the Rodentia order. Species in detail have available sequenced genomes. Red branches: Muroidea families in which *Dnmt3C* was identified; dashed branches: Muroidea families without available genome sequence.

fetal germ line (*19*, *20*); this likely explains their limited dependency toward DNMT3C remethylation activity. Nevertheless, DNMT3C dependency was still greatest over regulatory long-terminal repeat (LTR) sequences of MMERVK10C (fig. S5F) and IAPEz elements when highly conserved copies were analyzed by bisulfite pyrosequencing (fig. S4E). Finally, among paternally methylated imprinted loci, only the *Rasgfr1* imprinting control region (ICR) was hypomethylated in *Dnmt3C*$^{IAP/IAP}$ germ cells (Fig. 3I). This ICR includes an ERVK LTR fragment, which acquires DNA methylation in a piRNA-dependent manner during spermatogenesis (*21*). This highlights again the genomic convergence between piRNA and DNMT3C targeting in fetal male germ cells.

DNMT3C is highly specialized at methylating young retrotransposon promoters in the male germ line. In comparison, DNMT3B is not involved in germ cell development but establishes somatic methylation patterns genome-wide in the early embryo (*11*, *22*, *23*). DNMT3C has therefore evolved its own regulatory pattern and genomic targets, which have diverged from the ancestral DNMT3B copy. Despite the strict requirement of DNMT3C in the mouse, the *Dnmt3C* duplication is not universal among mammals but occurred ~46 million years ago in the last common ancestor of the muroid rodents (Fig. 4, A and B, and fig. S6, A and B). These represent the largest mammalian superfamily, with 1518 species accounting for 28% of all extant mammals (fig. S5C) (*24*), and include the two primary models for biomedical research, particularly in reproduction and endocrinology: the mouse and the rat. The genomes of Muroidea carry a heavy burden in young transposons: 25% have integrated in the last 25 million years with currently thousands of active copies (*25*). In comparison, most human transposons became extinct during that time (*26*).

The lack of transposon methylation defects in the germ cells of *Dnmt3A* or *Dnmt3B* mutant mice had been previously interpreted as a sign of functional redundancy between these two enzymes (*23*, *27*). We show here that this function relies instead on an additional de novo DNMT. The mammalian DNA methylation members should now be considered as a quintet: DNMT1, 3A, 3B, 3C, and 3L. The two most recent evolutionary additions are linked to reproduction: The eutherian DNMT3L cofactor stimulates germline methylation genome-wide, and the muroid DNMT3C enzyme methylates young retrotransposons during spermatogenesis. Future research should resolve the biochemical mechanism that designates specific sequences for DNMT3C-dependent DNA methylation. Its PWWP-free status may prevent it from targeting H3K36me3-enriched gene bodies and redirect it to retrotransposon promoters. In conclusion, our discovery of DNMT3C raises a new set of challenges to the current views of the remarkable evolution of DNA methyltransferases, the regulation of the de novo methylation process, and its tight links with the selective pressure to epigenetically control transposons in the germ line.

REFERENCES AND NOTES

1. J. A. Yoder, C. P. Walsh, T. H. Bestor, *Trends Genet.* **13**, 335–340 (1997).
2. J. L. Goodier, H. H. Kazazian Jr., *Cell* **135**, 23–35 (2008).
3. A. A. Aravin, D. Bourc'his, *Genes Dev.* **22**, 970–975 (2008).
4. S. A. Manakov *et al.*, *Cell Reports* **12**, 1234–1243 (2015).
5. A. Molaro *et al.*, *Genes Dev.* **28**, 1544–1549 (2014).
6. A. A. Aravin *et al.*, *Mol. Cell* **31**, 785–799 (2008).
7. D. Bourc'his, T. H. Bestor, *Nature* **431**, 96–99 (2004).
8. N. Zamudio, D. Bourc'his, *Heredity (Edinb)* **105**, 92–104 (2010).
9. D. J. Lees-Murdock *et al.*, *Genomics* **84**, 193–204 (2004).
10. H. Kobayashi *et al.*, *Genome Res.* **23**, 616–627 (2013).
11. M. Okano, D. W. Bell, D. A. Haber, E. Li, *Cell* **99**, 247–257 (1999).
12. S. La Salle *et al.*, *Dev. Biol.* **268**, 403–415 (2004).
13. S. K. T. Ooi *et al.*, *Nature* **448**, 714–717 (2007).
14. T. Baubec *et al.*, *Nature* **520**, 243–247 (2015).
15. A. Dhayalan *et al.*, *J. Biol. Chem.* **285**, 26114–26120 (2010).
16. S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng, *Cell* **76**, 357–369 (1994).
17. A. Tsumura *et al.*, *Genes Cells* **11**, 805–814 (2006).
18. N. Zamudio *et al.*, *Genes Dev.* **29**, 1256–1270 (2015).
19. P. Hajkova *et al.*, *Mech. Dev.* **117**, 15–23 (2002).
20. S. Seisenberger *et al.*, *Mol. Cell* **48**, 849–862 (2012).
21. T. Watanabe *et al.*, *Science* **332**, 848–852 (2011).
22. G. Auclair, S. Guibert, A. Bender, M. Weber, *Genome Biol.* **15**, 545 (2014).
23. M. Kaneda *et al.*, *Nature* **429**, 900–903 (2004).
24. J. Michaux, A. Reyes, F. Catzeflis, *Mol. Biol. Evol.* **18**, 2017–2031 (2001).
25. R. H. Waterston *et al.*, *Nature* **420**, 520–562 (2002).
26. R. Cordaux, M. A. Batzer, *Nat. Rev. Genet.* **10**, 691–703 (2009).
27. Y. Kato *et al.*, *Hum. Mol. Genet.* **16**, 2272–2280 (2007).

ACKNOWLEDGMENTS

We thank members of D.B.'s laboratory, especially M. Greenberg and M. Walter. We are grateful to T. H. Bestor, V. Colot, E. Heard, and E. Mulugeta for advice; S. Fouchécourt, B. Fumel, L. Mourao, and B. Piegu for technical help; A. Bortvin and S. Tajima for antibodies; C. Jouhanneau for excellent mouse husbandry; E. Desale and PHENOMIN-TAAM for the mutagenesis program; and PICT-IBiSA (France–Bioimaging, ANR-10-INSB-04), the Flow Cytometry and Animal Platforms of the Institut Curie. This research was supported by grants from European Research Council (ERC-Cog EpiRepro), Agence Nationale pour la Recherche (ANR TranspoFertil, ANR-10-LABX-0030-INRT, and ANR-10-INBS-07 PHENOMIN). D.B.'s laboratory is part of the Laboratoire d'Excellence (LABEX) DEEP (11-LBX0044). J.B. was a recipient of a postdoctoral fellowship from Agence Nationale de Sécurité Sanitaire (Anses). The Institut Curie ICGex NGS platform was supported by ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) grants and by the Canceropôle Ile-de-France. Mouse line distribution under uniform biological material transfer agreement. All sequencing data sets are available at the National Center for Biotechnology Information Gene Expression Omnibus database under GSE84141.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/354/6314/909/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S7
References (*28*–*54*)

8 July 2016; accepted 19 October 2016
10.1126/science.aah5143