R-loops induce repressive chromatin marks over mammalian gene terminators

Konstantina Skourtí-Stathaki1, Kinga Kamieniarz-Gdula1 & Nicholas J. Proudfoot1

The formation of R-loops is a natural consequence of the transcription process, caused by invasion of the DNA duplex by nascent transcripts. These structures have been considered rare transcriptional by-products with potentially harmful effects on genome integrity owing to the fragility of the displaced DNA coding strand. However, R-loops may also possess beneficial effects, as their widespread formation has been detected over CpG island promoters in human genes. Furthermore, we have previously shown that R-loops are particularly enriched over G-rich terminator elements. These facilitate RNA polymerase II (Pol II) pausing before efficient termination. Here we reveal an unanticipated link between R-loops and RNA-interference (RNAi) machinery acting broad-merase II (Pol II) pausing before efficient termination. Here we reveal an unanticipated link between R-loops and RNA-interference (RNAi) machinery acting broad-

Figure 1 | The RNAi-dependent H3K9me2 repressive mark is formed over the human β-actin terminator in HeLa cells. a, RT–qPCR of β-actin antisense transcription. RT with region-specific forward primers. b, Sense and antisense transcripts levels determined by RT–qPCR from J2 immunoprecipitation (ChIP) analysis using anti-G9a antibody. 

1Sir William Dunn School of Pathology, South Parks Road, University of Oxford, Oxford OX1 3RE, UK.

DOI: 10.1038/nature13787

©2014 Macmillan Publishers Limited. All rights reserved
poly(A) signal (PAS) and G-rich pause elements. To investigate whether R-loops and the H3K9me2 mark are specific features of the pause-dependent termination mechanism, we employed cyclin B1 and akirin 1 genes that utilize alternative CoTC terminators. DNA immunoprecipitation (DIP) and H3K9me2 ChIP analyses (Extended Data Fig. 2) showed no R-loop or H3K9me2 marks over these CoTC terminators, suggesting that such features are restricted to genes possessing pause-site terminals.

We investigated whether R-loops promote the recruitment of RNAi factors and H3K9me2 formation over the termination region of the β-actin gene by testing their sensitivity to RNase H1. Overexpression of this enzyme diminished R-loop levels over both the gene body (intron 1 amplicon) and pause regions (Fig. 2a). Remarkably, antisense RNA, DICER, G9a and HP1γ occupancy were also diminished (Fig. 2b–e). To selectively remove the H3K9me2 repressive mark, we used the chemical inhibitor of G9a/GLP, BIX-01294 (BIX), which induces transient DICER, G9a and HP1γ (Extended Data Fig. 2a). Ninety per cent of HeLa cell nuclei had R-loops and dsRNA in close proximity with H3K9me2 foci. Notably, 10% of either dsRNA or R-loop foci co-localized with H3K9me2 (Fig. 2f and Extended Data Fig. 3b). These data strongly suggest that, at a cellular level, R-loops are associated with gene silencing.

We next employed cell lines derived from mouse gene knockouts for Ago2 and G9a/Glp to test their role in Pol II termination. We initially validated the data obtained with the human β-actin gene for its mouse homologue (Extended Data Figs 4–6). Notably, we observed that the repressive mark associated with the termination region of the mouse β-actin gene is specifically H3K9me2, and not H3K9me3 (Extended Data Fig. 4d). We then confirmed that Ago2, like Ago1, is specifically enriched at the termination region of the mouse β-actin gene and its recruitment is reduced to background levels in Ago2-knockout cells (Fig. 3a and Extended Data Fig. 4c). However, Ago1 recruitment is enhanced in Ago2-knockout cells, suggesting that Ago1 compensates for Ago2 depletion (Extended Data Fig. 4e). G9a (Fig. 3b) and H3K9me2 (Fig. 3c) ChIP analyses in Ago2-knockout cells showed a decrease in ChIP signals over the gene termination region, suggesting that the observed H3K9me2 mark is Ago2-dependent. However, the R-loop profile is Ago2-independent (Extended Data Fig. 7a), confirming that R-loops act upstream of the RNAi pathway. Similar results were obtained in G9a/Glp double-knockout cells (Extended Data Fig. 7c, d).

To investigate whether both R-loops and the H3K9me2 mark are needed for efficient transcriptional termination of the mouse β-actin gene, we performed Pol II ChIP in wild-type and Ago2-knockout mouse embryonic fibroblasts (MEFs) also overexpressing RNase H1. Pol II density increased, especially over termination probes C and D, indicative of a defect in transcriptional termination (Fig. 3d and Extended Data Fig. 7b). We also performed Br-UTP nuclear run-on (NRO) analysis (Fig. 3e) and detected significant enrichment of nascent read-through DNA signals over the termination region, relative to the gene body (intron 3 primer), in Ago2-knockout cells overexpressing RNase H1, as compared to wild-type cells. This suggests that R-loops and the H3K9me2 mark are both critical components of efficient pause-dependent Pol II termination. No nascent transcripts were detected over probes E and F, located 3.2 and 4 kilobases (kb) downstream of the PAS, suggesting that the effect of combined loss of the H3K9me2 mark and R-loops promotes read-through transcription up to 3 kb downstream of the PAS.

We considered the possibility that RNAi-mediated heterochromatin formation induced by R-loop formation is a general termination mechanism, at least for a subset of genes. We performed a genomic meta-analysis of high-throughput sequencing of DNA derived from ChIP (ChIP-seq) data sets to look for the co-incidence of a paused elongating form of Pol II that is phosphorylated on Ser 2 of the carboxy-terminal domain (CTD) (POLLIS2ph15), with HP1γ enrichment within termination regions (Fig. 4a). We termed such regions of overlap pause-type termination (PTT) candidate regions. HP1γ was previously implicated in transcriptional elongation12–15. Indeed, 84% of the summits of HP1γ peaks determined
Figure 3 | Ago2-dependent H3K9me2 mark and R-loop formation promote efficient termination on mouse β-actin gene. a, β-actin in wild-type and Ago2-knockout (KO) MEFs using Ago2 (a) and G9a antibodies (b). c, Ratio of H3K9me2 versus H3 ChIP in wild-type and Ago2-knockout MEFs. d, Pol II ChIP with probes downstream of the PAS with extended y axis. Experiment was performed in wild-type (grey bars), wild-type overexpressing RNase H1 (black bars), Ago2-knockout (white bars) and Ago2-knockout overexpressing RNase H1 (red bars) MEFs. Full gene profile is in Extended Data Fig. 7b. All ChIP values are ± s.d. from 3–4 biological repeats. e, Br-UTP NRO analysis in wild-type (grey bars) and Ago2-knockout MEFs overexpressing RNase H1 (red bars). Nascent Br-RNA over intron 3 probe is set as 1. Fold of enrichment of read-through transcripts for pause, pause2 and the C termination probe calculated relative to intron (in)3 signal. D, E and F indicate D, E and F termination probes, respectively. Values are ± s.d. from three biological repeats. p(A), poly(A); prom, promoter.

by ChIP-seq reside within gene bodies (Extended Data Fig. 8a)15. However, the highest-fold enrichment for HP1γ relative to genomic annotation is over termination regions and the highest density of HP1γ peak summits is detected downstream of PAS, genome wide (Extended Data Fig. 8a, b). Notably, 74% of HP1γ enriched regions in termination regions overlap with PolIIIS2ph enrichment (Fig. 4b). PTT candidate regions show a statistically significant signal enrichment of the G9a ChIP hybridized to a genomic microarray (ChIP-chip)18, both compared to randomly sampled genomic regions of the same size as well as non-PTT HP1γ peaks (Fig. 4c and Extended Data Fig. 8c), implicating H3K9 methyltransferase activity at these locations. To investigate whether PTT candidate regions are associated with R-loop formation, we compared the signal obtained by DNA-RNA-immunoprecipitation (DRIP) with that obtained by DRIP treated with RNase H1 (DRIPRH1 control) from previously published DRIP-seq data2. PTT candidate regions show a significant enrichment of DRIP signal as compared to DRIPRH1 (Extended Data Fig. 8d, e), implying R-loop formation over these regions. We conclude that PTTs associated with R-loops, G9a and HP1γ are widespread in the human genome.

Two genes, ENSA and GEMIN7, which show PolIIIS2ph pausing coincident with HP1γ and DRIP-seq signal, were used to validate our genomic analysis. R-loops, antisense transcription, DICER, H3K9me2 and HP1γ were observed over their termination regions (Extended Data Fig. 9), similar to the β-actin terminator. Finally, we performed Br-UTP NRO
analysis after BIX treatment and RNase H1 overexpression on these non-actin genes, showing that their termination requires R-loops and the H3K9me2 mark (Fig. 4d). The same effect was observed for the human β-actin gene, thus validating the data obtained in mouse β-actin (Fig. 3d, e).

Finally, to corroborate the role of R-loops and H3K9me2 on transcriptional termination genome wide, we performed ChiP-seq using an antibody against PolⅡS2ph on BIX-treated cells overexpressing RNase H1 (BIX RH1) and on untreated cells. We observe a decrease in PolⅡS2ph accumulation in the vicinity of PAS in the BIX RH1 sample versus the untreated sample (Fig. 4e, right). However, an increase in PolⅡS2ph accumulation around the transcription start site (TSS) is detected in the BIX RH1 sample versus the untreated condition (Fig. 4e, left). We then calculated the PolⅡS2ph pausing index in PTT candidate regions relative to gene bodies and observed that the BIX RH1 sample has a significantly lower value compared with the untreated sample ($P = 3.398 \times 10^{-16}$; Extended Data Fig. 8f). This implies that efficient pausing in these locations depends on the presence of R-loops and H3K9me2. By contrast, PolⅡS2ph pausing around the TSS concurrently increases in the BIX RH1 condition (Fig. 4e and Extended Data Fig. 8g), suggesting that termination and promoter pausing mechanisms are distinct. This is consistent with the specific enrichment of DICER, H3K9me2 and HP1γ over gene 3′ ends, but not promoter regions, of β-actin, ENSA and GEMIN7 (Fig. 1 and Extended Data Fig. 9). Overall, we demonstrate that a termination mechanism mediated by Pol II pausing dependent on R-loop-induced heterochromatin is shared by a subset of human genes.

We reveal a molecular link between R-loop structures and the RNAi pathway. In particular, we have uncovered an unanticipated mechanism that terminates and promotes transcriptional termination genome wide, we performed ChiP-seq using an antibody against PolⅡS2ph on BIX-treated cells overexpressing RNase H1 (BIX RH1) and on untreated cells. We observe a decrease in PolⅡS2ph accumulation in the vicinity of PAS in the BIX RH1 sample versus the untreated sample (Fig. 4e, right). However, an increase in PolⅡS2ph accumulation around the transcription start site (TSS) is detected in the BIX RH1 sample versus the untreated condition (Fig. 4e, left). We then calculated the PolⅡS2ph pausing index in PTT candidate regions relative to gene bodies and observed that the BIX RH1 sample has a significantly lower value compared with the untreated sample ($P = 3.398 \times 10^{-16}$; Extended Data Fig. 8f). This implies that efficient pausing in these locations depends on the presence of R-loops and H3K9me2. By contrast, PolⅡS2ph pausing around the TSS concurrently increases in the BIX RH1 condition (Fig. 4e and Extended Data Fig. 8g), suggesting that termination and promoter pausing mechanisms are distinct. This is consistent with the specific enrichment of DICER, H3K9me2 and HP1γ over gene 3′ ends, but not promoter regions, of β-actin, ENSA and GEMIN7 (Fig. 1 and Extended Data Fig. 9). Overall, we demonstrate that a termination mechanism mediated by Pol II pausing dependent on R-loop-induced heterochromatin is shared by a subset of human genes.

We reveal a molecular link between R-loop structures and the RNAi pathway. In particular, we have uncovered a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell 45, 814–825 (2012).

1. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. Mol. Cell 46, 115–124 (2012).

2. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chedin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell 45, 814–825 (2012).

3. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chedin, F. GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcriptional termination. Genome Res. 23, 1590–1600 (2013).

4. Skourtis-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell 42, 794–805 (2011).

5. Nakama, M., Kawakami, K., Kajitani, T., Uruno, T. & Murakami, Y. DNA–RNA hybrid formation mediates RNA-directed heterochromatin formation. Genes Cells 17, 218–233 (2012).

6. Guillerova, M. & Proudfoot, N. J. Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Mol. Cell 132, 983–995 (2008).

7. Zofall, M. et al. Histone H2AZ cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature 461, 419–422 (2009).

8. Schönborn, J. et al. Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. Nucleic Acids Res. 19, 2993–3000 (1991).

9. Jenuwein, T. The epigenetic magic of histone lysine methylation. FEBS J. 273, 3121–3135 (2006).

10. Tachibana, M. et al. Histone methyltransferases G9a and GLP form heteromorphic complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev. 19, 815–826 (2005).

11. Rice, J. C. et al. Distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell 46, 115–124 (2012).

12. Vakoc, C. R., Mandat, S. A., Olenchok, B. A. & Blobel, G. A. Histone H3 lysine 9 methylation and HP1γ are associated with transcription elongation through mammalian chromatin. Mol. Cell 19, 381–391 (2005).

13. Mateescu, B., Bourachot, B., Rachez, C., Ogryzko, V. & Muchardt, C. Regulation of an inducible promoter by an HP1γ-H1 switch. EMBO Rep. 9, 267–272 (2008).

14. Saint-André, V., Batsché, E., Rachez, C., Muchardt, C. Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. Nature Struct. Mol. Biol. 18, 337–344 (2011).

15. Smallwood, A. et al. CBX3 regulates efficient RNA processing genome-wide. Genome Res. 22, 1426–1436 (2012).

16. Nojima, T., Dienstbier, M., Murphy, S., Proudfoot, N. J. & Dye, M. J. Definition of RNA polymerase II CoTC terminator elements in the human genome. Cell Rep. 3, 1080–1092 (2013).

17. Kubicek, S. et al. Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. Mol. Cell 25, 473–481 (2007).

18. Rosenbloom, K. R. et al. ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. 41, D56–D63 (2013).

19. Frietze, S., O’Geen, H., Guck, K. R., Jin, V. X. & Farnham, P. J. JZF274 recruits the histone methyltransferase SETDB1 to the 3′ ends of ZF genes. PLoS ONE 5, e15092 (2010).

20. Alò, M. et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nature Struct. Mol. Biol. 16, 717–724 (2009).

Online Content Information is available in the online version of the paper.

Acknowledgements This work was supported by a Wellcome Trust Programme grant (091805/Z/10/2) and a European Research Council Advanced grant (339270-polyloop) to N.J.P., and a Marie Curie Actions grant from EU FP7 (REA grant agreement, 327985) to K.K.-G. We are grateful to R. Nunes Bastos for valuable help with immunofluorescence/imaging analysis and figure formatting. We also thank O. Sun for helpful discussions and advice. We thank R. J. Crouch for the gift of the GFP-RNase H1 expression plasmid, G. Hannon for the Age2-knockout mouse cell line and Y. Shinaka for the G9a/Glp double-knockout mouse embryonic stem cells.

Author Contributions K.S.-S. performed all the molecular biology and imaging experiments. K.K.-G. performed the bioinformatics analysis. K.S.-S. and N.J.P. designed the experiments and wrote the manuscript.

Author Information Sequencing data for PolⅡS2ph ChiP and inputs from BIX RH1-treated and untreated HeLa cells have been deposited in the Gene Expression Omnibus under accession number GSE39878. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.J.P. (nicolas.proudfoot@path.ox.ac.uk).
METHODS

Molecular and cell biology techniques. Transfections of GFP-RNase H1 plasmid into human HeLa and mouse MEF cells were carried out as described previously. Ago2-knockout and parental wild-type cells are MEFs. G9a/Glp double-knockout and their parental wild-type are mouse embryonic stem (mES) cells. Treatment with 10 µM of BIX-01294 inhibitor (Sigma) was performed as described. Total RNA was isolated using TRIzol reagent (Invitrogen) and reverse transcribed with SuperScript III Reverse Transcriptase (Invitrogen) using gene-specific primers. J2 dsRNA pull-down was performed as described. RT–qPCR levels are presented graphically as raw values ±1,000. ChIP and genomic DIP analyses were carried out as before. The following antibodies were used for ChIP: anti-H3K9me2 (Abcam), anti-H3K9me3 (Abcam), anti-H3 (Abcam), anti-DICER (13D6) (Abcam), anti-KMT1C/G9a (Abcam), anti-Ago2 (Millipore), anti-Ago2 (Abcam) and anti-Pol II (H-224) (Santa Cruz Biotechnology). S9.6 RNA:DNA-hybrid-specific antibody was used for DIP. DNA oligonucleotide primers employed in these studies are listed in Supplementary Information Table 1.

Genomic analysis. Genomic interval processing, overlap calculations, statistical analysis and occupancy profiles were performed using custom scripts within the R/Bioconductor environment.

The following publicly available human data sets were used: G9a ChIP-chip, Gene Expression Omnibus (GEO) accession GSE24480 (ref. 19); Pol II ChIP-seq, ENCODE-enriched 2,046 PolII S2ph peaks (enriched regions) for the untreated sample and 7,712 treated input and 113,690,799 for the BIX RH1-treated input. ELAND parameters were: unique matches, 32 base seed, 2 mismatches allowed. Tags to the hg19 human genome was performed using the CASAVA pipeline 1.8.2, ChIP and two input samples were multiplexed, using NEBNext ChIP-Seq master-mix kit to prepare the libraries. The samples were sequenced on a 50 bp single-end run on three lanes using the Illumina HiSeq 2000 platform. Alignment of the sequenced tags to the hg19 human genome was performed using the CASAVA pipeline 1.8.2. ELAND parameters were: unique matches, 32 base seed, 2 mismatches allowed. This yielded a total of 115,813,632 reads uniquely aligned to hg19 for the untreated IP sample, 126,051,048 for the BIX RH1-treated sample, 127,749,851 for the untreated input and 113,690,799 for the BIX RH1-treated input. Peak calling was performed on the IP samples versus their input controls using MACS2 with the parameters: -q 0.05—nomodel—shiftsize 100. This procedure deli- 2,046 PolII S2ph peaks (enriched regions) for the untreated sample and 7,712 peaks for the BIX RH1-treated sample. We noted that the BIX RH1 treatment resulted in a higher overall PolII S2ph enrichment, presumably reflecting a globally more open chromatin environment following the treatment. Therefore, to avoid potential bias we chose to base our analysis on pausing indices relative to gene body signal (see later).

To obtain the PolII S2ph enrichment profiles over the TSS and PAS (Fig. 4e), the distance of the PolII S2ph peaks to the nearest TSS (or PAS) was computed, retaining only distances <10 kb away from the feature of interest, which were then subject to kernel density estimation using the Gaussian smoothing kernel and plotted.

We defined the PTT PolII S2ph pausing index as a ratio of the normalized read density in PTT candidate regions to the normalized read density in its corresponding gene body. In more detail, we first re-computed the PTT candidate regions using the PolII S2ph peaks found in the untreated sample in place of the ENCODE data-set derived peaks, and their corresponding gene body coordinates were extracted. The IP and input read number overlapping with each PTT and its corresponding gene body were counted and the IP and input reads per kilobase per million mapped reads (RPKM) read density for each region were computed as follows: RPKM = (number of reads overlapping with region)/(length of region in kb)/(million mapped reads). The RPKM value for input reads was then subtracted from the RPKM value for the IP reads for each region to yield the final normalized read density (NRD) for each region. Genes with low PolII S2ph NRD over their gene body (NRD < 0.1) were considered inactive and excluded from downstream analysis. The pausing index (PI) of each PTT/gene body pair was then computed as: PI_{PTT} = PTTRPKM/NRDgene body.

The pausing index for regions surrounding the TSSs of the PTT-linked genes by ±1 kb was computed analogously: PI_{TSS} = NRD_{TSS} ± 1 kb/NRDgene body. These computations were done in parallel for the BIX RH1-treated and untreated sample, and finally the distribution of the fold change in PolII S2ph pausing index between the BIX RH1-treated and untreated samples was calculated.

For statistical tests, since the data in Extended Data Fig. 8c, e–g did not conform to a normal distribution, non-parametric tests were employed: Wilcoxon signed-rank test for the paired samples in Extended Data Fig. 8e, f and Wilcoxon Mann–Whitney for the unpaired samples in Extended Data Fig. 8c, g. In all cases two-sided tests were applied.

Immunofluorescence and imaging analysis. Fixed cell samples were prepared and imaged exactly as described. In summary, cells grown on coverslips were fixed with 2 ml of ice-cold methanol or 3% paraformaldehyde in PBS for 15 min. Cells were quenched with 2 ml of 50 mM NH4Cl in PBS for 10 min. Coverslips were washed three times in 2 ml PBS before permeabilization in 0.2% Triton X-100 for 5 min. In all cases primary and secondary antibody staining was performed in PBS for 60 min at room temperature. S9.6 antibody was used in 1:2,500 dilution, whereas commercial H3K9me2 (Cell Signaling) and J2 (Sciccon) antibodies were used as directed by the manufacturers. 4′,-Diamidino-2-phenylindole (DAPI) was added to the secondary antibody staining solution at 0.3 µg/ml. Coverslips were mounted in Mowiol 4-88 mounting medium (EMD Millipore). Fixed samples on glass slides were imaged using a ×60/NA 1.35 oil immersion objective on an upright microscope (BX61, Olympus) with filter sets for DAPI, GFP/Alexa Fluor 488, 555, 568, and 647 (Chroma Technology), a CoolSNAP HQ2 camera (Roper Scientific) and MetaMorph 7.5 imaging software (Molecular Dynamics). Co-localization foci were measured as foci <200 nm apart.

J2 dsRNA pull-down. J2 antibody (Sciccon, 10010200, diluted to 0.1 µg per 1 µg of chromatin) was incubated with total cell extracts for 1.5 h on a rotating wheel at 4 °C. Protein G-agarose beads (Millipore) were then added for an additional 1.5 h. dsRNA was then isolated from washed beads using the TRizol reagent (Invitrogen) and analysed by RT–qPCR for sense and antisense transcripts. Signals from immunoprecipitated samples were subtracted from signals arising from non-precipitated samples. V1 and S1 treatments were carried out for 2 h at 37 °C after the dsRNA isolation.

21. Gentleman, R. C. et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5, R80 (2004).
22. Karolchik, D. et al. The UCSC Table Browser data retrieval tool. Nucleic Acids Res. 32, D493–D496 (2004).
23. Bastos, R. N., Penate, X., Bates, M., Hammond, D. & Barr, F. A. CYK4 inhibits Rac1-dependent PAK1 and ARhGf7 effector pathways during cytokinesis. J. Cell Biol. 196, 865–880 (2012).
Extended Data Figure 1 | H3K9me2 and H3 levels over human β-actin gene.

a, Left, H3K9me2 ChIP on β-actin gene. Right, H3K9me2 ChIP analysis on human centromere 9 (positive control).
b, Left, H3 ChIP on β-actin gene. Right, H3 ChIP analysis on human centromere 9.
c, Left, H3K9me2 ChIP with or without BIX treatment. Right, H3 ChIP with or without BIX treatment. ChIP values are ±s.d. from three biological repeats. C and D indicate C and D termination probes, respectively. cen, centromere; in, intron; p(A), poly(A); prom, promoter.
Extended Data Figure 2 | R-loops and the H3K9me2 repressive mark are not specifically enriched over the CoTC terminators of human cyclin B1 and akirin 1 genes. a, DIP on endogenous cyclin B1 and akirin 1 genes. No detection of R-loops was observed over their CoTC terminators. Human β-actin gene was used as a positive control. For cyclin B1 and akirin 1 genes, 3′-end 1 and 3′-end 2 amplicons amplify two different regions within the CoTC terminator of each gene. 3′-end 1 and 3′-end 2 amplicons for the β-actin gene amplify the 5′ pause and pause amplicons, respectively. b, Ratio of H3K9me2 signal over the 3′ ends versus intron 3 signal in cyclin B1, akirin 1 and β-actin human genes. DIP and ChIP values are ±s.d. from three biological repeats.
Extended Data Figure 3 | Cellular localization of R-loops, dsRNA and H3K9me2. a, Immunofluorescence imaging of dsRNA (J2 antibody) and R-loops (S9.6 antibody), using paraformaldehyde (PFA) and methanol (MeOH) as fixing reagents. Fixation with methanol allowed visualization of R-loops and dsRNA in HeLa cell nuclei. Enlarged boxes (1 and 2) are shown in right panels. b, Whole-cell images showing immunofluorescence of H3K9me2 with dsRNA (J2; top) and R-loops (S9.6; bottom). Enlarged versions (1 and 2) are shown in Fig. 2h.
Extended Data Figure 4 | R-loops and RNAi promote the H3K9me2 mark over mouse β-actin terminator. a, DIP performed on mouse β-actin gene in MEFs. b, RT–qPCR of total RNA from MEF cells on β-actin gene to detect antisense transcripts with region-specific forward primers. Average RT–qPCR values are ± s.d. from four biological repeats. c, Ago1 ChIP performed on mouse β-actin gene in MEFs. ChIP signal is normalized to intron 1 signal. d, Left, ratio of H3K9me2 ChIP signal versus H3 on mouse β-actin in MEFs. Middle, normalized H3K9me3 to total H3 levels. Right, ratio of H3K9me2 and H3K9me3 signal versus H3 signal on major satellites in MEFs. e, Ago1 ChIP in wild-type (grey bars) and Ago2-knockout (KO) (white bars) cells. Ago1 recruitment over mouse β-actin is enhanced upon Ago2 depletion. f, Left, ratio of H3K9me2 ChIP signal versus total H3 on β-actin gene in wild-type and G9a/Glp double-knockout mouse embryonic stem cells. Right, H3K9me2/H3 ratio on the mouse major satellites in wild-type and G9a/Glp double-knockout cells. Average ChIP and DIP values are ± s.d. from three biological repeats.
Extended Data Figure 5 | H3K9me2, H3K9me3 and H3 levels over the endogenous mouse β-actin gene. 

a, H3K9me2 and H3K9me3 ChIP on mouse β-actin gene in MEF cells. Right, H3K9me2 and H3K9me3 ChIP on mouse major satellites (positive control). 

b, Total H3 ChIP on mouse β-actin gene. Major satellites were used as a positive control. ChIP values are ±s.d. from three biological repeats.
Extended Data Figure 6 | H3K9me2 and H3 levels over mouse β-actin gene in G9a/Glp double-knockout mouse embryonic stem cells and Ago2-knockout MEFs. a, Top and bottom, H3K9me2 and H3 ChIP performed on mouse β-actin gene in wild-type and G9a/Glp double-knockout embryonic stem cells. H3K9me2 occupancy depends on the presence of G9a/Glp HKMTs.

b, ChIP analyses using H3K9me2 (top) and H3 (bottom) antibodies performed on mouse β-actin gene in wild-type and Ago2-knockout cells. ChIP values are ± s.d. from three biological repeats.
Extended Data Figure 7 | R-loop formation and antisense transcription are Ago2- and G9a/GLP-independent. a–c, DIP performed on mouse β-actin gene in wild-type, Ago2-knockout (a) and G9a/Glp double-knockout (c) cells. b, Pol II ChIP in wild-type (grey bars), wild-type overexpressing RNase H1 (black bars), Ago2-knockout (white bars) and Ago2-knockout overexpressing RNase H1 (red bars) MEFs. Hatched box quantifies Pol II read-through transcription versus promoter signal. d, RT–qPCR analysis of total RNA from wild-type and G9a/Glp double-knockout cells for the mouse β-actin gene. RT reaction was performed with specific forward primers. Average DIP and RT–qPCR values are ± s.d. from three biological repeats.
Extended Data Figure 8 | HP1γ, G9a and R-loops are globally associated with paused Pol II over PTTs. a, Genomic annotation of HP1γ based on ChIP-seq peak summit localization (HP1γ annotation, pie chart on the left) and the fold enrichment of HP1γ over the indicated genomic regions (table on the right) as compared to their base-pair coverage in the human genome (genome annotation, pie chart in the middle). Genic regions were defined by RefSeq gene coordinates (hg19). Promoter regions were defined as regions 1 kb upstream of RefSeq gene TSS excluding intervals overlapping with any genic regions. Termination regions were defined as regions 5 kb downstream of RefSeq genes excluding intervals overlapping with any genic region or promoter. b, HP1γ ChIP-seq enrichment profile in 10-kb regions surrounding the TSS (left graph) and PAS (right graph). HP1γ peaks summit frequencies are plotted in 500-bp bins. c, Box plot showing the average log2 (G9a/input) ChIP-chip signal distribution in PTT candidate regions (right box), randomly sampled regions of the same size and number as PTT candidate regions (random regions, left box), and in HP1γ peaks outside of PTT candidate regions (non-PTT HP1γ peaks, middle box). In all box plots the horizontal line in the box shows the median, the lower and upper limits of the box show respectively the first and third quartile, and the whiskers extend to the non-outlier extreme data points. The log2 (G9a/input) signal is significantly higher in the PTT candidate regions compared to random regions (P = 0.0001067) as well as compared to non-PTT HP1γ peaks (P = 0.02213). The log2 (G9a/input) signal is also significantly higher in non-PTT HP1γ peaks compared to random regions (P = 0.0009337). The Wilcoxon Mann–Whitney test was applied in all cases. d, DRIP-seq profile over the centre of PTT candidate regions. Read frequencies of DRIP sample (black curve) and DRIP RH1 sample (red curve) are plotted in 500-bp bins, both normalized to million mapped reads. e, Box plot showing DRIP-seq read density (RPKM) of DRIP sample compared with DRIP RH1 control in PTT candidate regions. P < 2.2 × 10^{-16} determined by Wilcoxon signed-rank test. f, Box plot of PolII2ph pausing index over PTTs (relative to gene bodies) in the BIX RH1 sample (right) and the untreated sample (left). P = 3.398 × 10^{-16} using the Wilcoxon signed-rank test. g, Box plot displaying the ratio of PolII2ph pausing index in the BIX RH1-treated sample compared with the untreated sample in TSS regions (±1 kb, left) and in PTT regions (right). P = 2.468 × 10^{-18} using the Wilcoxon Mann–Whitney test.
Extended Data Figure 9 | ENSA and GEMIN7 share features of R-loop mediated PTT. a, DIP on ENSA and GEMIN7 genes. R-loops specifically enriched over 3' ends (grey bars), compared to promoter regions (white bars). Human β-actin gene is positive control. Values ± s.d. for three biological repeats. b, RT-qPCR of total RNA from HeLa cells performed on indicated gene. RT reaction was performed with promoter or 3' end-specific forward primer to detect antisense transcript. Average RT-qPCR values are ± s.d. from four biological repeats. c, DICER ChIP of ENSA and GEMIN7 genes over promoters and termination regions. d, Left, ratio of H3K9me2 ChIP signal versus H3 on GEMIN7 and β-actin genes. Right, ratio of H3K9me2 signal versus H3 on ENSA gene. e, f, H3K9me2 and H3 ChIP for ENSA and GEMIN7 genes over promoter (white bars) and pause terminators (grey bars). β-Actin gene was used as a positive control. g, HP1γ ChIP for ENSA and GEMIN7 genes over intronic and 3'-end regions. ChIP values are ± s.d. from three biological repeats.
Extended Data Figure 10 | Model for how R-loops and RNAi-dependent H3K9me2 chromatin mediate pause-type termination in mammalian genes.

Mammalian genes possessing pause elements downstream of their PAS form R-loops in termination regions. This facilitates generation of an antisense transcript that hybridizes with the sense transcript to form dsRNA. This triggers recruitment of the RNAi factors, DICER, AGO1 and AGO2. G9a/GLP HKMTs and HP1γ are then recruited, forming and maintaining H3K9me2 repressive marks. R-loops and H3K9me2 facilitate Pol II pausing before termination. DNA is shown as grey lines and RNA as a red line. Points of contact between the DNA strand and nascent RNA indicates R-loop formation, whereas points of contact between sense and antisense RNA indicate dsRNA formation. Pol II is shown as a blue icon with arrow indicating transcription direction. Nucleosomes are shown in green except over H3K9me2 region where they are coloured red.