OPEN ACCESS Freely available online

PLOS | GENETICS

RNAi-Dependent and Independent Control of LINE1 Accumulation and Mobility in Mouse Embryonic Stem Cells

Constance Ciaudo[1,2¤a], Florence Jay[1,3], Ikuhiro Okamoto[2¤b], Chong-Jian Chen[2,4], Alexis Sarazin[1], Nicolas Servant[4,5,6], Emmanuel Barillot[4,5,6], Edith Heard[2], Olivier Voinnet[1,3]*

1 Swiss Federal Institute of Technology Zurich, Department of Biology, Chair of RNA biology, Zurich, Switzerland, **2** Institut Curie, CNRS UMR3215, Paris, France, **3** Life Science Zurich Graduate School, Plant Sciences program, University of Zurich, Zurich, Switzerland, **4** Institut Curie, Paris, France, **5** INSERM U900, Paris, France, **6** Mines ParisTech, Fontainebleau, France

Abstract

In most mouse tissues, long-interspersed elements-1 (L1s) are silenced *via* methylation of their 5′-untranslated regions (5′-UTR). A gradual loss-of-methylation in pre-implantation embryos coincides with L1 retrotransposition in blastocysts, generating potentially harmful mutations. Here, we show that Dicer- and Ago2-dependent RNAi restricts L1 accumulation and retrotransposition in undifferentiated mouse embryonic stem cells (mESCs), derived from blastocysts. RNAi correlates with production of Dicer-dependent 22-nt small RNAs mapping to overlapping sense/antisense transcripts produced from the L1 5′-UTR. However, RNA-surveillance pathways simultaneously degrade these transcripts and, consequently, confound the anti-L1 RNAi response. In *Dicer*$^{-/-}$ mESC complementation experiments involving ectopic Dicer expression, L1 silencing was rescued in cells in which microRNAs remained strongly depleted. Furthermore, these cells proliferated and differentiated normally, unlike their non-complemented counterparts. These results shed new light on L1 biology, uncover defensive, in addition to regulatory roles for RNAi, and raise questions on the differentiation defects of *Dicer*$^{-/-}$ mESCs.

Citation: Ciaudo C, Jay F, Okamoto I, Chen C-J, Sarazin A, et al. (2013) RNAi-Dependent and Independent Control of LINE1 Accumulation and Mobility in Mouse Embryonic Stem Cells. PLoS Genet 9(11): e1003791. doi:10.1371/journal.pgen.1003791

Editor: Wolf Reik, The Babraham Institute, United Kingdom

Received June 3, 2013; **Accepted** July 29, 2013; **Published** November 7, 2013

Funding: This work was supported by the ANR grant "RNA ES"; to both EH and OV. OV's laboratory was further supported by an ERC Young Investigator Award (Grant ERC-210890 "Frontiers of RNAi") and Prix Liliane Bettencourt for Life Sciences. EH is a member of the European Network of excellence "Epigenesis". CC was supported by a post-doctoral fellowship from the Federation of European Biochemical Societies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: voinneto@ethz.ch

¤a Current address: Swiss Federal Institute of Technology Zurich, Department of Biology, Chair of RNAi and Genome Integrity, Zurich, Switzerland.
¤b Current address: Kyoto University, Kyoto, Japan.

Introduction

Long-interspersed elements-1 (LINE-1 or L1) belong to the most abundant class of autonomous transposable elements (TEs) in mammalian genomes. While most L1s are truncated and unable to transcribe or retrotranspose, a fraction of young, full-length L1s are capable of mobilization [1]. Active and inactive L1s influence the evolution of mammalian genomes, yet L1 insertions are also linked to disease [1], raising the issue of how L1 expression and retrotransposition are controlled. In plants, fungi and metazoans, silencing small (s)RNAs suppress TEs at both transcriptional and post-transcriptional levels [2]. In mice, germline-specific, 26–31-nt PIWI-associated RNAs (piRNAs) derived from TE-enriched clusters are loaded into ARGONAUTE-like PIWI proteins directing *de-novo* cytosine methylation and RNA degradation of active TEs, including L1 [3]. In most healthy somatic tissues, L1s are silenced via 5′-UTR promoter methylation, established from 7.5 days of embryogenesis [4]. In pre-implantation embryos, by contrast, L1 methylation progressively decreases, to reach 13–23% in blastocysts [5], which accumulate full-length L1 transcripts and undergo mosaic retrotransposition, a potential source of heritable and non-heritable mutations [6,7]. Pre-implantation embryogenesis thus defines a critical window during which L1s should be tightly controlled despite their hypo-methylated status and the lack of piRNAs.

In plants, RNA interference (RNAi) at the post-transcriptional level can operate as a surrogate to cytosine methylation and heterochromatinization in TE-silencing [8,9]. RNAi relies on populations of small interfering (si)RNAs, processed sequentially by the RNase-III Dicer (DCR) from long, perfectly double-stranded (ds)RNA precursors [10]; these are commonly produced by TEs due to their complex insertion patterns or intrinsic bi-directional transcription. Processed siRNAs load into ARGONAUTE (Ago)-family effector proteins and guide sequence-specific degradation of complementary target transcripts. The existence of an endogenous (endo)-siRNA pathway in mammals has been debated, notably because long dsRNA triggers the non-specific interferon (INF) response in most cells [11]. In mouse oocytes, which lack an INF response, heterogeneous sRNA populations map to L1 and LTR elements, among other loci, but their DCR-dependency is unknown; additionally, L1 accumulation is unchanged in oocytes of conditional *Dcr*$^{-/-}$ animals [12]. Mouse

Author Summary

A basal network of gene regulation orchestrates the processes ensuring maintenance of genome integrity. Eukaryotic small RNAs generated by the RNAse-III Dicer have emerged as central players in this network, by mediating gene silencing at the transcriptional or post-transcriptional level via RNA interference (RNAi). To gain insight into their potential developmental functions in mammals, we have characterized small RNA expression profiles during mouse Embryonic Stem Cell (mESCs) differentiation, a model for early mammalian development. Long interspersed elements 1 (L1) are non-long-terminal-repeat retrotransposons that dominate the mouse genomic landscape, and are expressed in germ cells or during early development and mESCs. Based on clear precedents in plants and fission yeast, we investigated a role for RNAi and other RNA-based pathways in the regulation of L1 transcription and mobilization. Our work uncovered the existence of small (s)RNAs that map to active L1 elements. Some have characteristics of cognate siRNA produced by Dicer, while others display strand biases and length heterogeneity that evoke their biogenesis through RNA surveillance pathways, in a Dicer-independent manner. Furthermore, genetic ablation of DICER or of ARGONAUTE proteins has complex and profound consequences on L1 transcription and mobilization, indicating that endogenous RNAi do indeed maintain genomic integrity against L1 proliferation.

Embryonic Stem Cells (mESCs) also lack an INF response and their ability to produce DCR-dependent endo-siRNAs was clearly established genetically [13]. Being isolated from the blastocyst's inner mass, cultured mESCs are thus potentially suited to study the mechanism(s) that might restrict L1 retrotransposition during pre-implantation, including, possibly, RNAi. Supporting this view, several classes of young, full-length endogenous L1 are hypomethylated and transcriptionally active in undifferentiated mESCs [14,15] but become re-methylated and silenced upon differentiation [15,16]. Moreover, substantially increased L1 transcript levels were reported in undifferentiated $Dcr^{-/-}$ mESCs [17], although this was not confirmed in separate analyses of a distinct KO cell line [18]. Shallow RNA sequencing (15–50-nt size-range) in undifferentiated mESCs revealed that L1 transcription correlates with accumulation of sense and antisense sRNAs of undetermined nature/function, mapping mostly to the L1_5′-UTR [15]. In Human L1s, this region displays overlapping sense-antisense transcription with the potential to form dsRNA and, as such, was proposed to generate anti-L1 endo-siRNAs [19,20]. In a pioneering study, attempts to substantiate this idea in somatic human cells yielded, however, indecisive conclusions: discrete 21–23-nt L1-derived sRNAs could indeed be detected in some cell lines but not others, and their DCR-dependency was not established; moreover, knocking-down human *Dcr-1* caused only marginal increases in endogenous L1 transcription and retrotransposition [20,21].

Here, we have investigated the possible link between RNAi and endogenous L1 regulation in undifferentiated mESCs. Uniquely, these cells can withstand full genetic ablation of DCR or the AGO proteins, albeit at the cost of proliferation and differentiation defects tentatively ascribed, at least partly, to an inability of $Dcr^{-/-}$ mESCs to produce micro- (mi)RNAs [17,22]. Unlike siRNA populations, DCR-dependent miRNAs accumulate as discrete, imperfect duplexes excised from stem-loop-containing precursor transcripts produced from numerous independent transcription units. Mature miRNAs are thought to regulate hundreds of cellular transcripts displaying partial miRNA-complementarity, which include mRNAs important for cell fate specification but also pluripotency [23]. Undifferentiated mESCs contain relatively few, albeit highly abundant miRNAs, that can be genetically discriminated from endo-siRNAs and other rare DCR-dependent sRNAs using mutations in the generic miRNA biogenesis factor DGCR8; *Dgcr8*_KO mESCs contain, nonetheless, few non-canonical miRNAs produced by diverse means [13]. Combining the use of deep-sequencing and cell lines carrying null mutations in *Dcr*, *Ago*s and *Dgcr8*, we have investigated the distribution, biochemical origin(s) and ability of L1-derived sRNAs to silence L1 transcript accumulation and retro-transposition in undifferentiated mESCs. Our study reveals an unexpected level of complexity in L1 silencing in these cells, where siRNA-directed RNAi processes are confounded by the overlapping effects of general RNA-surveillance pathways. These findings reveal a novel level of mammalian L1 regulation and shed new light on the proliferation defects and inability of $Dcr^{-/-}$ mESCs to differentiate.

Results

LINE-1 mRNA and proteins overaccumulate in Dicer knockout mESCs

To further explore the L1-derived sRNAs in undifferentiated mESCs, we combined ILLUMINA deep-sequencing and the use of the ncPRO pipeline [24] enabling genomic mapping of repeat-derived sRNAs. A population of abundant, sense and antisense sRNAs was detected, mapping as a majority to the L1-Tf_5′-UTR, consistent with our previous observations (Figure 1A and Figure S1A) [15]. As seen previously with human L1 [19], strand-specific RT-PCR revealed that the 5′-UTR of the L1-Tf subfamily [25] displays overlapping sense-antisense transcription (Figure S1B) with the potential, therefore, to generate dsRNA as a possible source of DCR-dependent siRNAs. Because constitutive DCR depletion is detrimental to cultured mESCs [22], we pursued the above idea by generating inducible Cre-ERT2 *Dcr* knockouts. Although *Dcr* deletion was already achieved 24 h post-tamoxifen treatment (Figure 1B), reduced accumulation of miR-295, one of the most abundant mESC miRNAs, was only visible 6 days post-tamoxifen treatment, presumably reflecting the high DCR protein stability [22]. By 12 d post-tamoxifen treatment, miR-295 was below detection levels of quantitative qRT-PCR, indicating full depletion of DCR activity, also confirmed by quantitation of previously validated mESC miRNA target transcripts (Figure 1C and Figure S1C). Strikingly, decreased DCR levels were inversely correlated with accumulation of mRNA and ORF1 protein derived from all L1 classes (Figure 1D and 1E) or from distinct L1-subtypes displaying 5′_UTR polymorphisms [25] (Figure S1D). Analysis of a specific, polymorphic L1 element on chromosome 17 [15] yielded similar results (Figure 1F). $Dicer^{-/-}$ mESCs have been reported to display hypomethylation due to decrease levels in DNA methy-transferases (DNMTs) [26]. However, DNMT1 and DNMT3b proteins were expressed to the same levels in wild type and $Dicer^{-/-}$ cells (Figure S2A). In addition the L1 mRNA was not up-regulated in a cell line carrying a triple-KO for DNMT1, 3a and 3b (Figure S2B) [27]. Investigating the methylation status of the L1_5′-UTR through bisulfite sequencing revealed nonetheless that $Dicer^{-/-}$ mESCs are hypomethylated, which could contribute to the observed up-regulation of the L1 mRNA (Figure S2C).

Retrotransposition of LINE-1 in Dicer knockout mESCs

Moreover, this increase in L1 transcript/protein in $Dcr^{-/-}$, but not $Dcr^{Flx/Flx}$ mESCs, was paralleled by a marked gain in

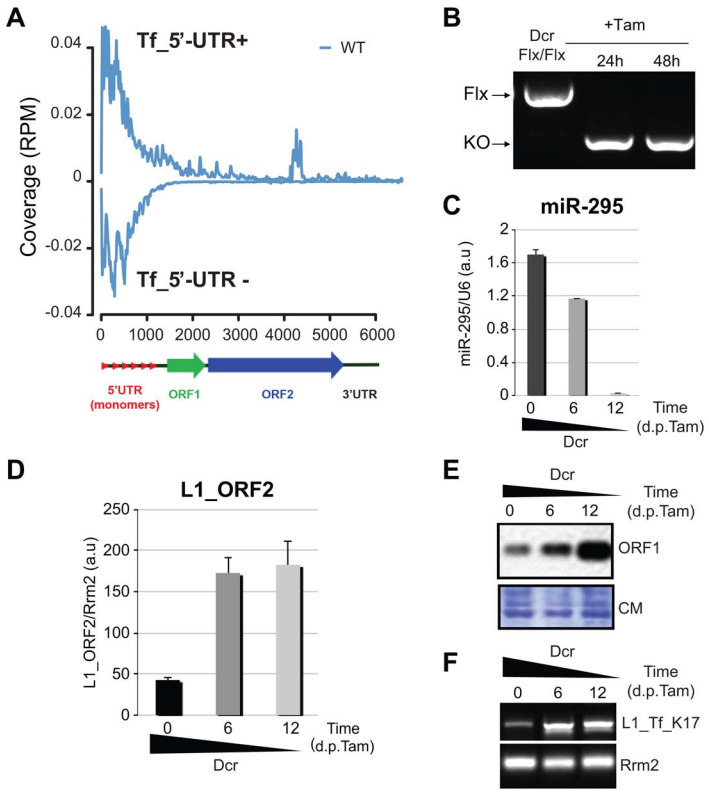

Figure 1. L1 elements are up-regulated in *Dcr*$^{-/-}$ mESCs. A. Sequencing reads, from WT mESCs, within the 19–32-nt range were aligned against the mouse genome (version mm9). The distinct sequences coverage (Reads per Million (RPM) normalized) is depicted for the full length L1Md_Tf L1 [24]. B. PCR-based genotyping of the *Dcr* deletion 24 h and 48 h post-tamoxifen (Tam) treatment. C. qRT-PCR analysis of miR-295 levels in the tamoxifen treated mESCs, as depicted in (B). D. L1_ORF2 mRNA accumulation detected by qRT-PCR before and after *Dcr* deletion. E. Western analysis of L1_ORF1 protein levels before and after *Dcr* deletion; CM: Coomassie staining of total protein. F. Semi-quantitative RT-PCR analysis of RNA levels from a single L1-Tf copy on chromosome 17 before and after *Dcr* deletion.
doi:10.1371/journal.pgen.1003791.g001

endogenous L1 copy-number, estimated by Q-PCR using PCR primers specific for the L1_Tf subfamily. The promoter activity of mouse L1 elements lies in tandemly repeated, 200 bp monomeric units within the 5′-UTR. These monomers are distinct between different LINE-1 families [28]. We used promoter-specific primers to discriminate, by Q-PCR, the three active families of murine L1 elements designated Tf-, Gf-, and A-type (Table S1). For copy number analysis, we focused exclusively on L1_Tf, which was the subfamily we found mostly associated with small RNAs accumulation (Figure S1A) [15]. Using RepeatMasker (AFA. Smit and R. Hubley. RepeatModeler Open-1.0. http://www.repeatmasker.org, 2008–2010), we identified 22,506 sequences annotated L1Md_T (L1_Tf), 15,286 annotated L1Md_A and 819 annotated L1Md_Gf. Among these "fragment" population of L1 elements, we identified 2,291 L1Md_T, 1,338 L1Md_A and 35 L1Md_Gf, which have a length matching at least 95% of their corresponding L1 reference sequence. For 1 512 L1Md_T, we were able to identify at least one amplicon using the L1_Tf specific PCR primers (See Materials & Methods and Table S1), with an average of 3.9 amplicons per elements. We thus calculated 2,770 full length L1_Tf elements in the mm9 genome, which is in line with the 2000–3000 L1_Tf elements previously estimated by Naas *et al.* [29], of which 60% are putatively active. To evaluate the gain in copy-number, *Dcr*$^{Flx/Flx}$ mESCs were analysed at passage 10, upon which the *Dcr* deletion was induced; the L1_Tf copy number was then re-assessed after 20 additional passages in *Dcr*$^{Flx/Flx}$ and *Dcr*$^{-/-}$ background (Figure 2A). About 2,452 active L1_Tf copies were found in *Dcr*$^{Flx/Flx}$ mESCs at P10 and 2,707 copies at P30. A gain of approximately 860 new copies was detected in the *Dcr*$^{-/-}$ cell line after 20 passages. Therefore, we estimate that between 1 and 20 new active copies of L1_Tf were generated per day (i.e. 2 cell divisions) in the *Dcr*$^{-/-}$ background.

To ascertain the above PCR-based results, we adapted the gain-of-GFP retrotransposition assay previously developed by Prak and colleagues [30] and validated in WT mESCs [31]. To score *de novo* L1 retrotransposition events in the *Dcr*$^{Flx/Flx}$ and *Dcr*$^{-/-}$ mESC lines, a human L1 modified to contain an intronic, split eGFP reporter was stably integrated into the genome of *Dcr*$^{Flx/Flx}$ and *Dcr*$^{-/-}$ mESCs. During early propagation (passages 4 to 6 after selection of puromycin-resistant cells), only in *Dcr*$^{-/-}$ cells was the eGFP mRNA significantly increased, although the same quantity of plasmid was transfected in each cell lines and was similarly expressed 24 h post-transfection (Figure 2B and S2D); moreover, this increase was not visible in *Dcr*$^{-/-}$ cells transformed with RT-deficient point-mutation alleles of the eGFP-tagged L1 (Figure S2D) or in WT cells (data not shown). Amplification of *eGfp* DNA was also only observed in *Dcr*$^{-/-}$ cells transformed with WT eGFP-tagged L1 (passage 6; Figure 2C), a method also previously employed to validate active retrotransposition [32]. Finally, similar results were also obtained using a gain-of-Luciferase retrotransposition assay [33] (data not shown). We conclude that DCR negatively controls L1 transcript accumulation and retrotransposition in undifferentiated mESCs.

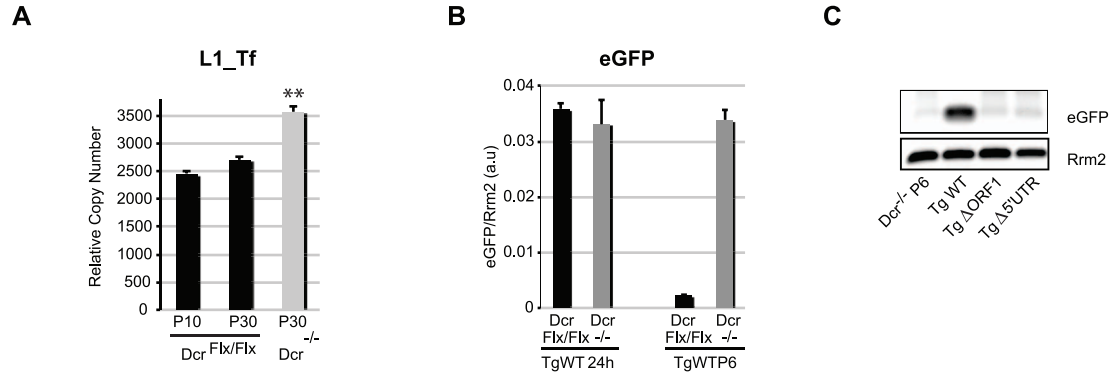

Figure 2. L1 elements retrotranspose in $Dcr^{-/-}$ mESCs. A. qPCR-based copy-number analysis of L1_Tf elements in $Dcr^{Flx/Flx}$ P10 (5 replicates), P30 (2 replicates) and $Dcr^{-/-}$ P30 (5 replicates) mESCs. **: p-value<0.01. B. qRT-PCR-based analysis of eGFP mRNA levels in WT and $Dcr^{-/-}$ mESCs stably transformed with a WT human eGFP-tagged L1 transgene, 24 h post-transfection and 6 passages (P6) after puromycin selection. C. Integrated eGFP as a diagnostic of retrotransposition detected by PCR in the genomic DNA of $Dcr^{-/-}$ mESCs carrying the human eGFP-tagged L1 transgene after 6 passages post-puromycin treatment.
doi:10.1371/journal.pgen.1003791.g002

L1 sRNAs form overlapping populations of DCR-dependent and -independent species

Given the DCR-dependent control of L1, we next investigated whether L1_5′-UTR sRNAs are DCR products. Total RNA from $Dcr^{-/-}$ mESCs was subjected to ILLUMINA sequencing. As expected, loss-of-DCR activity caused a dramatic decline in cellular 21–23-nt RNAs including, chiefly, miRNAs, representing most sRNAs in WT mESCs. Consequently, the relative proportion of repeat-derived sRNAs was seemingly increased in $Dcr^{-/-}$ compared to WT cells (Figure S3A–C). However, read-size analysis and genomic mapping revealed a specific depletion in sense and antisense L1-derived 22-nt sRNAs in $Dcr^{-/-}$ mESCs (Figure 3A); the remaining, abundant DCR-independent L1-derived sRNAs were heterogeneous in size, ranging from 19- to 32-nt, had both sense and antisense orientations and mapped mostly to the L1_5′-UTR, as in WT cells (Figure 3B and S3D). To test if some L1-derived sRNAs were effectively loaded into cognate RNA silencing effectors, we analysed the sRNA content of immunoprecipitates from endogenous AGO1 and AGO2, the only Agos we found significantly expressed at the protein level in undifferentiated mESCs, in agreement with available mESC RNA-seq data [34]. Using qRT-PCR, we found that the most abundant sense and antisense L1_5′-UTR sRNAs were specifically loaded into AGO2, as were several abundant miRNAs tested (Figure 3C). To obtain a comprehensive and unbiased view of AGO2-loaded L1-derived sRNAs, we used a mESC line overexpressing Flag-HA-tagged human Ago2 (FHA-hAgo2), and subjected RNA isolated from anti-Flag immunoprecipitates to ILLUMINA sequencing (Figure 3E–G). The fraction of sense and antisense L1-derived sRNAs loaded into FHA-hAgo2 had a genomic distribution resembling that of total RNA from WT or $Dcr^{-/-}$ cells (Figure 3B); it was, however, clearly enriched in 22-nt sRNAs, the size range of cognate DCR products (Figure 3A). Sequential processing, by DCR, of long dsRNA substrates produces perfect siRNA duplexes with diagnostic 2-nt 3′ overhangs. These features were indeed displayed by approx. 10% of all 22-nt sRNA sequences derived from L1-Tf in WT mESCs, as illustrated in Figure 3D with the 5′_UTR region of a single, near-consensus L1-Tf locus on the X chromosome. The abundance of these species was reduced by approx. 20% in $Dcr^{-/-}$ compared to WT mESCs, consistent with the read-size analysis in Figure 3A. Furthermore, they showed a 7-fold enrichment in IPs of AGO2, the cognate effector of siRNAs (Figure 3D). The most straightforward interpretation of these results is that L1-derived sRNAs form overlapping and complex populations of DCR-dependent and -independent molecules. These likely derive from longer sense- and antisense-RNA, or long dsRNA hybrids thereof, which are produced mostly from the L1_5′-UTR. While DCR-independent sRNAs dominate the overall population and display broader and heterogeneous size ranges, DCR-dependent sRNAs with cognate features of siRNAs load preferentially into AGO2, owing to the known 22-nt size preference of this silencing effector protein [35].

Neither the RNA surveillance nor the microRNA pathway regulate L1 mobilization in mESCs

To gain further insights into the nature and origin of DCR-independent L1 sRNAs and their possible involvement in L1 control, we tested the impact of 5′-3′ and 3′-5′ RNA-surveillance pathways operated by the Xrn class of exoribonucleases and the exosome, respectively. In diverse organisms, these pathways degrade aberrant pre-mRNA, transposon-derived and non-coding RNA, including antisense RNA [33,34]. We thus generated stable mESC lines displaying knockdown (KD) of nuclear Xrn2, cytoplasmic Xrn1 or exosome co-factor Rrp6, using new and previously established shRNA constructs [36] (Figure 4A and S4A). The results presented are for Xrn2; similar effects were observed in *Rrp6*_KD cells (Figure S4A and S4C), while they were much less pronounced in *Xrn1*_KD cells (data not shown). L1 transcripts and ORF1 protein were found significantly up-regulated in *Xrn2*_KD mESCs, (Figure 4A–B and S4D), correlating with reduced accumulation of the most abundant sense and antisense L1-derived sRNAs (Figure S4B). Copy-number analysis (as in Figure 2A) showed, however, that neither the *Xrn2*_KD nor the *Rrp6*_KD cell lines showed enhanced L1 mobilization (Figure 4C and S4C). Thus, RNA-surveillance pathways likely contributed to the heterogeneous mESC L1-derived sRNAs, which are presumably degradation intermediates of exonucleolysis of the longer sense and antisense transcripts derived from the L1_5′-UTR. Loss of these RNA-surveillance pathways did not, however, impact on L1 retrotransposition. These results thus support a role for siRNA-mediated RNAi in the control of L1 mobility, although we could not rule out an indirect contribution of DCR-dependent miRNAs, which regulate hundreds of cellular transcripts. To address this, we examined *Dgcr8*_KO mESCs, in which production of even the most abundant mESC miRNAs,

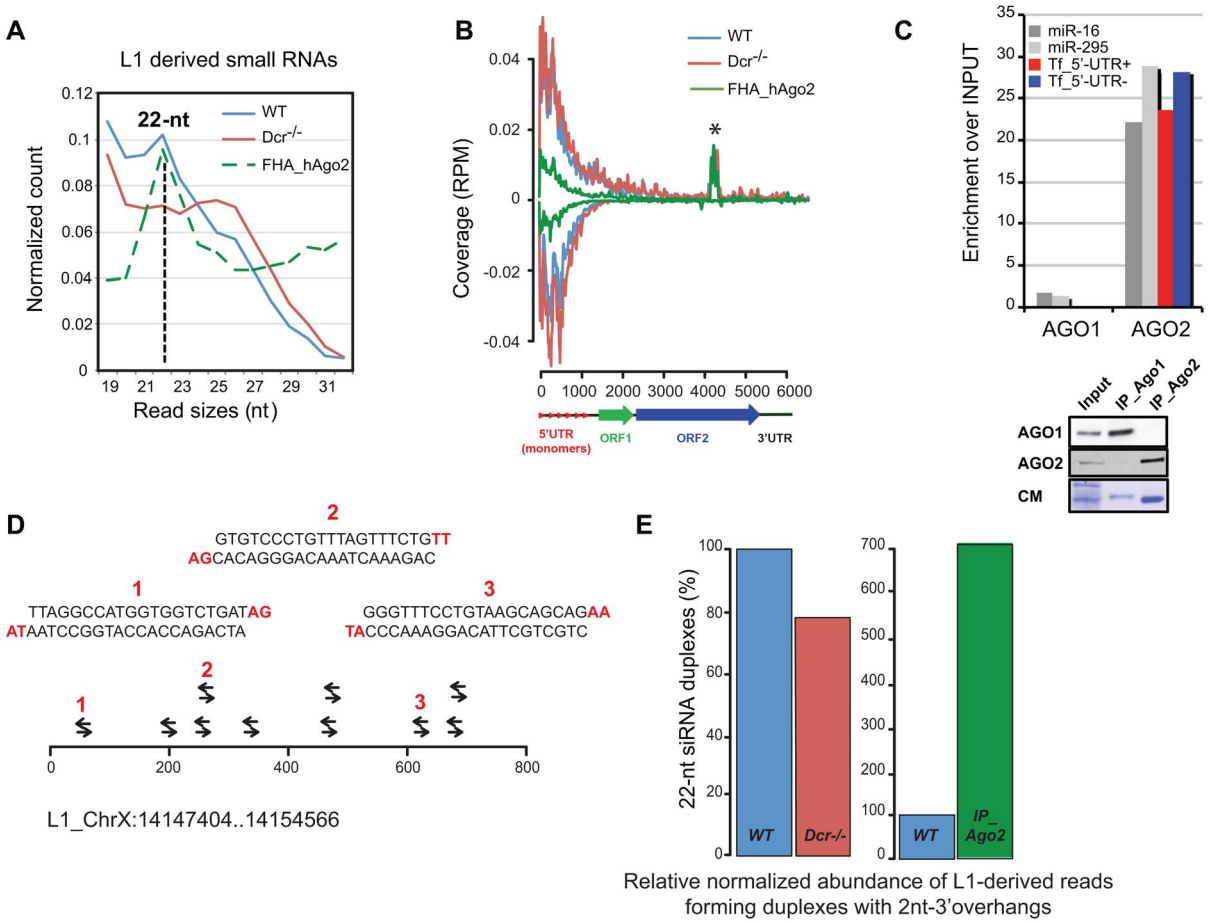

Figure 3. L1 sRNAs are partially produced by DCR and 22-nt L1-sRNAs loaded into AGO2. A. Size distribution of sequencing reads mapping to L1_Tf elements in $Dcr^{-/-}$ compared to WT mESCs. Note the deficit in 21–23-nt sRNAs in the former, and their enrichment in immunoprecipitated FHA-hAgo2, used in (B). B. Sequence coverages of L1_Tf elements from WT, $Dcr^{-/-}$ and immunoprecipitated E14-FHA-hAgo2 mESCs, as in Figure 1A. * this peak is an artefact found in all libraries sequenced. C. qRT-PCR analysis of the most abundant sense and antisense L1_5′-UTR-derived sRNAs and of two mESC miRNAs in immunoprecipitates of endogenous AGO1 and AGO2. Lower panel: control Western analysis of endogenous AGO1 and AGO2 levels after immunoprecipitation; CM: Coomassie staining of total protein. D. Snapshot representation of perfect, 22-nt-long sRNA duplexes with 2-nt 3′ overhangs (in red) mapped on the 5′-UTR of a single L1-Tf element located on chromosome X. E. Relative proportions of all L1-Tf-derived sequences forming perfect duplexes with 2-nt overhangs in WT versus $Dcr^{-/-}$ mESCs (left panel) and their enrichment in AGO2 immunoprecipitates (right panel). The numbers of 22-nt distinct sequences were normalized by the L1-Tf family coverage of RPM and expressed as proportion compare to WT.
doi:10.1371/journal.pgen.1003791.g003

including miR-295, is abrogated [37] (Figure 4D). Accumulation of endo-siRNAs and potential DGCR8-independent (i.e. non-canonical) miRNAs should remain in these cells [13,35] (Figure S4E). As shown in Figure 4E and 4F, *Dgcr8*_KO cells unexpectedly displayed enhanced L1 mRNA accumulation, possibly explained by the hypomethylation status of L1_5′-UTR in these cells (Figure S2A–C). However, there was no increase in L1 copy number (Figure 4F), ruling out the contribution of canonical miRNAs to the observed DCR-dependent control of L1 retrotransposition.

AGO2 is crucial for L1 silencing and strongly destabilized in $Dcr^{-/-}$ mESCs

Although we could not formally exclude a role for some unknown DGCR8-independent miRNAs in L1 retrotransposition control, the above results pointed to the likely contribution of DCR-dependent, 22-nt siRNAs derived from the L1_5′-UTR region of overlapping sense-antisense transcription (Figure 3A and S3B); their strong loading-bias in AGO2 thus predicted a crucial role for this silencing effector in L1 regulation (Figure 3C and 3E). To test this idea and avoid functional redundancy with AGO1 as previously observed with miRNAs [38,39], we used an established quadruple *Ago1,2,3,4*_KO mESC line, in which a stably expressed h*Ago2* transgene can be deleted upon tamoxifen treatment [38] (Figure S4F). MiR-295 levels were strongly reduced at 2 d, and at 5 d post-tamoxifen treatment; a corresponding increase in microRNA target levels confirmed cellular depletion of hAgo2, as reported [38] (Figure 4G and S4G). As in $Dcr^{-/-}$ cells, tamoxifen-induced *Ago1,2,3,4*_KO mESCs displayed strong up-regulation of L1 transcripts and increased L1 copy-number, unlike their untreated counterparts (Figure 4H–I and S4H–I), supporting a key contribution of AGO2 in L1 silencing in undifferentiated mESCs.

The L1 copy-number increase in *Ago1,2,3,4*_KO cells was noticeably less pronounced than in $Dcr^{-/-}$ cells (Figure 4I and 2A) however, possibly reflecting intrinsic differences in the relative initial L1 copy-number of non-treated *Ago1,2,3,4_KO* and $Dcr^{Flx/Flx}$ cell lines. Alternatively, cellular depletion of *Dcr* may have had additional, unanticipated effects that would lead to more potent

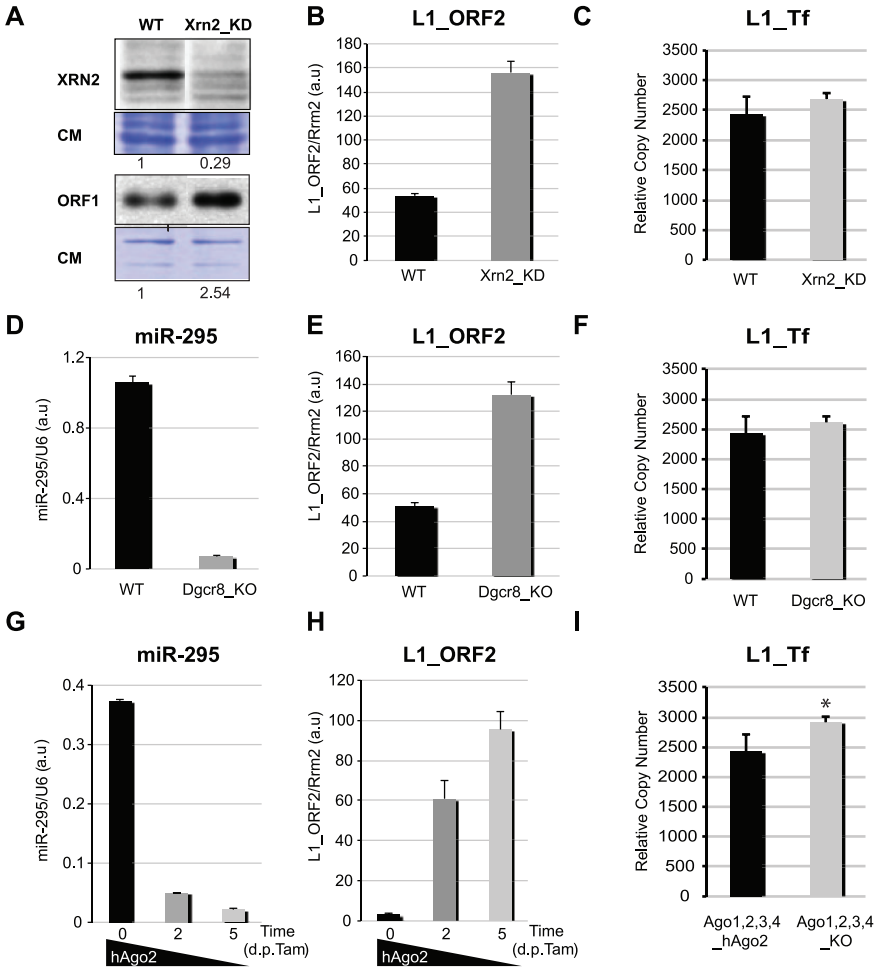

Figure 4. L1 mRNA levels and genomic copy-number in various knock-out and knock-down mESC lines. A. Western analysis of XRN2 and L1_ORF1 accumulation in WT and *Xrn2*_KD mESCs; CM: Coomassie staining of total protein. B. qRT-PCR analysis of L1_ORF2 mRNA levels in WT and *Xrn2*_KD mESCs. C. qPCR analysis of L1_Tf copy-number in WT and *Xrn2*_KD mESCs. D–E. qRT-PCR analysis of miR-295 (D) and L1_ORF2 mRNA (E) levels in WT and *Dgcr8*_KO mESCs. F. qPCR analysis of L1_Tf copy-number in WT and *Dgcr8*_KO mESCs. G–H. qRT-PCR analysis of miR-295 (G) and L1_ORF2 mRNA (H) levels upon hAgo2 deletion in Tamoxifen-treated *Ago1,2,3,4_KO* mESCs. I. qPCR analysis of L1_Tf copy-number in *Ago1,2,3,4_KO_hAgo2* mESCs before and after hAgo2 deletion. *: p-value$<$0.1.
doi:10.1371/journal.pgen.1003791.g004

L1 retrotransposition (Figure 2A). A survey of several key RNAi components indeed revealed a specific and dramatic reduction of AGO2 in $Dcr^{-/-}$ mESCs in multiple experiments, also reported recently in *Dgcr8*_KO mESCs [40] (Figure 5A). This effect was observed only at the protein level, as Ago2 mRNA remained expressed (Figure S5A); it was also specific, since AGO1 protein levels remained unchanged (Figure 5A). As the prevalent effector of DCR-dependent miRNAs (Figure 3C), representing alone up to 70% of all mESC miRNAs (Figure S3G), we reasoned that AGO2 might have been destabilized and degraded in $Dcr^{-/-}$ mESCs due to the loss of its main sRNA cargoes; such an effect was documented for the Arabidopsis miRNA-effector AGO1 [41]. Indeed, we found AGO2 levels to be significantly up-regulated upon treatment of $Dcr^{-/-}$ mESCs with the 26S-proteasome inhibitor MG132, an effect previously reported for miRNA-depleted hAgo2 [40,42] (Figure S5B). Also consistent with a loss-of-miRNA-dependent effect, AGO2 levels, unlike those of AGO1, were also reduced in *Dgcr8*_KO mESCs, albeit consistently less than in $Dcr^{-/-}$ cells (Figure 5A). This, incidentally, possibly explained the results on L1 silencing obtained in *Dgcr8*_KO mESCs (Figure 4E–F): while a deficit in the main effector of 22-nt L1-derived siRNAs likely increased L1 mRNA accumulation, the remaining AGO2 levels were presumably still sufficient to restrict L1 retrotransposition. We further infer that, in the absence of canonical miRNAs, loading of DCR-dependent endo-siRNAs (including L1-derived siRNAs) and possibly non-canonical miRNAs, explains the residual levels of AGO2 in *Dgcr8*_KO mESCs (Figure 5A). These levels are considerably lower in $Dcr^{-/-}$ mESCs, because neither sRNA class is produced in this background.

Rescue of AGO2 levels and L1 silencing in miRNA-depleted $Dcr^{-/-}$ mESCs

The above results prompted us to assess further the extent to which L1 silencing could be rescued in $Dcr^{-/-}$ mESCs in a miRNA-independent manner. We repeatedly failed to reintroduce hAgo2 transgenically into $Dcr^{-/-}$ mESCs, including a catalytic null (slicer deficient) allele of hAgo2, in our initial attempts to differentiate potential siRNA-mediated effects (slicer-dependent) and miRNA-mediated (slicer-independent) on L1 silencing. We ascribe this failure to an inability to sufficiently re-stabilize Ago2 in

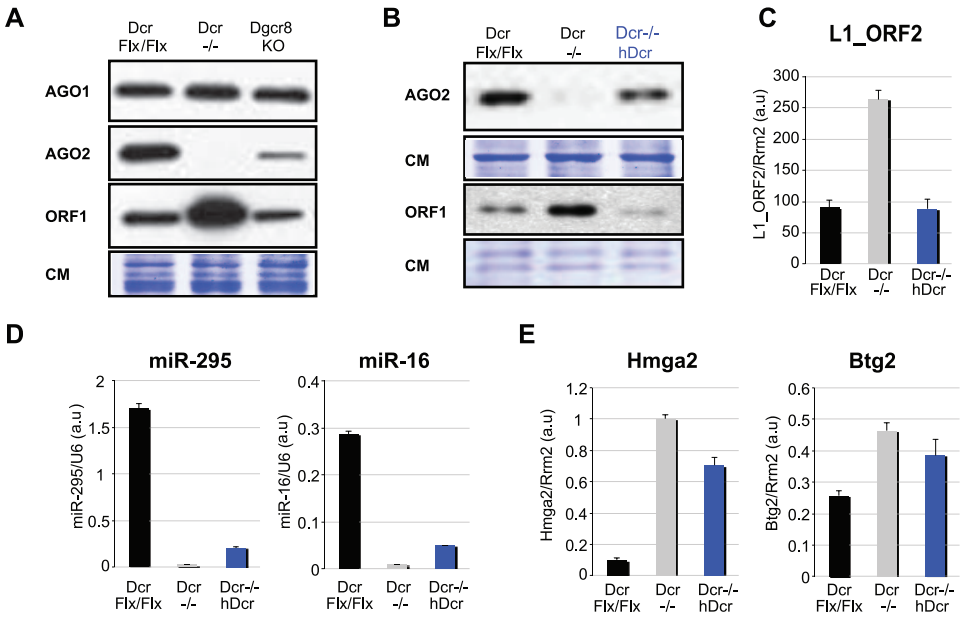

Figure 5. Rescue of L1 silencing in *hDcr*-complemented $Dcr^{-/-}$ mESCs. A. Western analysis of endogenous AGO1, AGO2 and L1_ORF1 protein levels in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and *Dgcr8*_KO ESCs; CM: Coomassie staining of total protein. B. Western analysis of endogenous AGO2 and L1_ORF1 protein levels in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ mESCs and one representative stable line of h*Dcr*-complemented $Dcr^{-/-}$ mESC; CM: Coomassie staining of total protein. C–E. qRT-PCR analysis of L1_ORF2 mRNA levels (C), miR-295 and miR-16 levels (D), and Hmga2 and Btg2 mRNA levels (established targets of mmu-miR-196a and mmu-let-7a and mmu-miR-132, respectively) in the various cell lines depicted in (E).
doi:10.1371/journal.pgen.1003791.g005

the absence of its main cargoes, the miRNAs [40]. We thus resorted to stably complement $Dcr^{-/-}$ mESCs with a human *Dcr* (h*Dcr*) transgene. During their early propagation, several independent puromycin-selected clones displayed endogenous AGO2 levels consistently higher than in non-complemented $Dcr^{-/-}$ mESCs; moreover, L1 silencing, measured by ORF1 and mRNA accumulation, was restored in these cells to almost the levels seen in $Dcr^{Flx/Flx}$ cells (Figure 5B and 5C). Strikingly, however, in nearly all these clones, mature miRNA levels were only rescued to approximately 10% of WT levels, which is likely below physiological significance, because all validated miRNA targets tested accumulated ectopically in these cells (Figure 5D, 5E and Figure S5C). The levels of miRNAs and their targets were eventually restored to WT levels, as previously described [17], but only after extended periods of culture involving more than ten cell passages, during which the poor fitness of cultured $Dcr^{-/-}$ mESCs [22] presumably resulted in selection for cell variants with higher DCR levels. Nonetheless, the early passage data demonstrate that L1 silencing could be achieved in undifferentiated mESCs displaying as little as 10% total miRNAs, suggesting that RNAi via endo-siRNAs, including L1-derived siRNAs, is sufficient to silence L1s.

Complemented, miRNA-defective $Dcr^{-/-}$ mESCs differentiate normally

These data prompted us to re-evaluate the proliferation and differentiation defects of $Dcr^{-/-}$ cells [17,22,37]. Having now uncovered a potential role for DCR and AGO2 in L1 silencing in addition to their known regulatory functions via miRNAs, we explored to what extent the $Dcr^{-/-}$ cell defects were attributable to defective miRNA, as opposed to siRNA, biogenesis or action. $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and h*Dcr*-complemented $Dcr^{-/-}$ cells (early passages) could all undergo the formation of Embryoid Bodies (EBs), although this was achieved with a 1–2 d delay in $Dcr^{-/-}$ cells. 6 d after the onset of differentiation, EBs were plated onto adherent flasks and monitored until d10 of differentiation. Microscopy and quantitation of key pluripotency markers confirmed that $Dcr^{-/-}$ cell-derived EBs were unable to differentiate, even if they attached to the flasks' surface [17] (Figure 6A, 6B and Figure S6A). In contrast, the growth rate and morphology of cells differentiated around attached EBs were similar in $Dcr^{Flx/Flx}$ and h*Dcr*-complemented $Dcr^{-/-}$ cells (Figure 6A). Furthermore, 10 d post-differentiation of h*Dcr*-complemented $Dcr^{-/-}$ cells, AGO2 accumulation was partially rescued, and the levels of pluripotency and differentiation markers were similar to those of $Dcr^{Flx/Flx}$ cells (Figure 6B and Figure S6A). MiRNA accumulated to only 5–7% of WT levels in h*Dcr*-complemented $Dcr^{-/-}$ cells, which displayed, accordingly, ectopic miRNA target accumulation (Figure S6A and S6B). However, L1 transcript accumulation in these cells was as low as in differentiated $Dcr^{Flx/Flx}$ cells (Figure S6A), demonstrating rescue of L1 silencing. These results strongly suggest that miRNAs alone are unlikely to underpin mESC differentiation in the above experimental setting. Additional processes likely entail the production of DCR-dependent endo-siRNAs, of which some might contribute to protecting the mESC genome integrity by silencing active retrotransposons, including L1. Consistent with this idea, *Dgcr8*_KO cells that lack all canonical miRNAs but, unlike $Dcr^{-/-}$ cells, suppress L1 mobilization, could partially differentiate in the same experimental setting, agreeing with previous findings [37]; *Xrn2*_KD cells, which also are L1-silencing proficient, differentiated similarly to $Dcr^{Flx/Flx}$ cells (Figure 6C).

Discussion

Our results support a role for DCR-dependent L1-derived siRNAs in taming endogenous L1 retrotransposition in undifferentiated mESCs. They are thus consistent with RNAi safeguarding genome integrity during a time window of mouse development

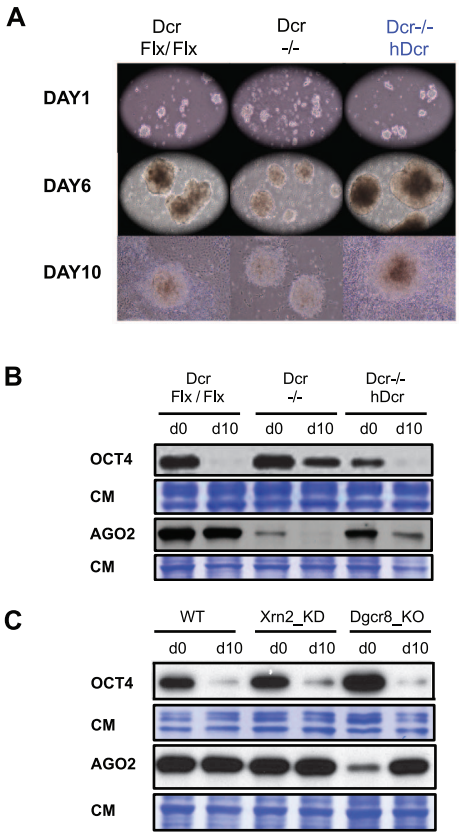

Figure 6. h*Dcr*-complemented $Dcr^{-/-}$ ESCs differentiate normally despite accumulating 5–7% total miRNAs compared to WT. A. Visualization of Embryoid bodies from $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and h*Dcr*-complemented $Dcr^{-/-}$ mESCs after 1, 6 and 10 days of differentiation. B. Western analysis of OCT4 and endogenous AGO2 protein levels in the cells depicted in (A) before (d0) and after 10 days of differentiation (d10); CM: Coomassie staining of total protein. C. Same as (B) but in WT, *Xrn2*_KD and *Dgcr8*_KO mESCs.
doi:10.1371/journal.pgen.1003791.g006

when DNA hypo-methylation coincides with L1 mobilization [7]. A role for RNAi in correcting DNA methylation defects of TEs is fully supported by previous work in Arabidopsis [8,9], in which intricate and opposing interactions between the RNAi and RNA-surveillance pathways have also been documented [43]. A parallel can be further established between our results and those of recent work in *S. pombe*, showing that RNAi at several loci, including Tf2 retroelements, is confounded by the 3′-5′ exonuclease activity of the exosome; genetic ablation of Rrp6 was, accordingly, sufficient to uncover siRNAs accumulating at these loci, showing, in that case, selective competition between the two pathways [44]. Human L1 transcription has been associated with the production of abundant non-polyadenylated and possibly uncapped RNAs of both strands accumulating in the nucleus [45]. Similar RNAs produced from the mouse L1_5′-UTR region likely provide the bulk of templates for Xrn2 and the Rrp6-associated exosome, known to degrade aberrant transcripts in the nucleus [46]; the resulting degradation intermediates form an important source of sRNA, heterogeneous in size, that confound detection of *bona fide* siRNAs. Full-length L1 transcripts may also undergo degradation by both exonucleases since a gain in ORF1 protein accumulation was observed in *Xrn2_KD* and *Rrp6_KD* mESCs. Alternatively, it is possible that different subsets of LINE1 elements are affected differently by the various genetic backgrounds tested here Figure S4D). For instance, highly transcribed but transposition-deficient L1 variants may accumulate elevated levels of specific transcripts that are more sensitive to the action of Xrn2 and/or Rrp6. This would rationalize why elevated L1 transcription, as seen in the *Xrn2_KD* and *Rrp6_KD* backgrounds, does not necessarily correlate with retrotransposition. The overall complexity uncovered here with L1 probably explains, more generally, why the existence and distribution endo-siRNAs at the whole-genome scale has been difficult to establish in mammalian cells so far, despite the pervasive nature of aberrant and antisense transcription in mammalian genomes [47].

A role for DCR in silencing L1 and perhaps other TEs could also shed light on the previously reported failure of $Dcr^{-/-}$ mESCs to differentiate. The h*Dcr* rescue experiment (Figure 6A) shows, notably, that miRNA biogenesis/activity is unlikely to contribute alone to mESC differentiation in this experimental setting and that endo-siRNAs are likely also important. Production of TE-derived siRNAs, in particular, might be crucial in preventing widespread double-strand break lesions and insertional mutagenesis, foreseeably detrimental to mESC proliferation and differentiation. This might also explain why *Dgcr8*_KO cells, unlike $Dcr^{-/-}$ cells, retain an ability to differentiate partially [17,37] (Figure 6C). Indeed, *Dgcr8*_KO cells, unlike their $Dcr^{-/-}$ counterparts, exhibit detectable levels of AGO2 (Figure 5A), which must be stabilized by the loading of endo-siRNA including L1-derived siRNAs and/or non-canonical miRNAs. Transposon taming and/or endogenous regulation by these molecules might be sufficient to rescue, at least partly, the differentiation defects of *Dgcr8*_KO cells. Assessing the extent to which RNAi-dependent control of active TEs, including L1s, contributes to the integrity of mESCs proliferation and differentiation, and thus to early mammalian embryogenesis, is an attractive prospect for future investigation.

Conclusions

This study unravels how multiple RNA silencing pathways might cooperate to dampen the expression and mobilization of an active family of transposable element family in mammalian cells. It not only echoes previously findings made in plants, fungi and invertebrates [44,48,49], but also rationalizes the complex patterns of small RNAs uncovered in studies originally conducted in mouse oocytes [12,50] and, more recently, in Human stem cells [51]. On a final note, although the mechanisms uncovered here in mESCs were tied in within the frame of early embryonic development, they may well also apply to other stem cell niches formed post-embryonically and present in many tissues of adult mammals. The states of pluri- and multi-potency seem generally associated with a deficit or absence of protein-based immunity, which includes the IFN response to exogenous and endogenous dsRNA. This might explain, at least partly, why these cells, unlike many other somatic cells, appear to accommodate RNAi triggered by long dsRNA [52,53]. In this context, we contend that siRNA-based RNAi has persisted in vertebrates as a primordial mechanism that protects progenitor cells of developing and adult organisms against the harmful effects of transposons and exogenous viruses [54]. This proposed RNAi-based defence is anticipated to be important, because genomic instability or viral infections in progenitor cells would have long-lasting detrimental consequences throughout the entire lineages derived from them. Defensive, as opposed to regulatory roles for mammalian RNAi have been somewhat overlooked thus far, but we are optimistic that the work reported here and elsewhere [54] will shed new light on this specific and fascinating aspect of RNA silencing.

Materials and Methods

Culture and in vitro differentiation of ESC

mESCs were cultured in Dulbecco's Modified Eagle Media (DMEM) (Invitrogen), containing 15% of a special selected batch of fetal bovine serum (FBS; Life Technologies) tested for optimal growth of mESCs, 1000 U/ml LIF (Millipore), 0.1 mM 2-β mercaptoethanol (Life Technologies), 0.05 mg/mL of streptomycin, and 50 U/mL of penicillin (Sigma) on a gelatin coated support in the absence of feeder cells. Embryoid body cultures were established by aggregation of mESCs in a low-adherent tissue culture dish into LIF-free DMEM, 10% FBS medium, from day 1 until day 6 and reattached on adherent flasks at day 10 of differentiation. The culture medium was changed daily. All cells were grown at 37°C in 8% CO_2.

The male E14 mESC line (129/Ola background) [55] was used for Illumina deep sequencing of WT mESC. E14_FHA-hAgo2 cells was created by stable transfection of plasmid pIRESneo-FLAG/HA Ago2 corrected (Addgene plasmid 10822) [56] and selection on G418-containing medium. TKO mESCs were described in ref. [27]. *Dgcr8*_KO mESCs were purchased from Novus Biologicals (NBA1-19349). New CreERT2-*Dcr*$^{Flx/Flx}$ mESCs were isolated from the cross of floxed *Dcr*$^{Flx/Flx}$ mice [22,57] and ROSA-CreERT2 mice [58]. The *Dcr*$^{-/-}$ mESC line was established from *Dcr*$^{Flx/Flx}$ mESC after induction with Tamoxifen (more than 15 days) and routinely control for their loss of microRNA accumulation. Genotyping primers used for the characterization of these cell lines are presented in Table S1. The inducible mESC line deficient for the four mouse Argonautes and carrying a floxed human Ago2 transgene (*Ago1,2,3,4*_KO) was described previously [38] and validated in the laboratory according to the author instructions. *Dcr* and *hAgo2* deletions were induced with 4-OHT (Tam) stock solution (1 mM, dissolved in 100% ethanol) diluted 1:1000 in cell culture medium to a final concentration of 1 µM for 6 and 12 days (*Dcr*) and 2 and 5 days (*hAgo2*). Transgenesis of WT and mutated version of the human L1-eGFP (RP = TgWT, JM111 = Tg**Δ**ORF1 and 2980 = Tg**Δ**5′UTR) [30] was carried out using 5 µg of each plasmids and Lipofectamine 2000 (Life Technologies) in E14 and *Dcr*$^{-/-}$ mESCs. Stable clones were selected on puromycin-containing medium (1 µg/mL; Sigma) 24 h post-transfection and analysed after 4 and 6 passages of the cells. *Xrn2*_KD mESC lines were generated using the pSUPER-puro vector (OligoEngine, http://www.oligoengine.com) engineered to produce the active shRNA 5′- CTCCAGAAGAGAA-CAGGAGAAAT-3′. Upon transfection of PGK mESCs [59], cells were selected on puromycin-containing medium. Control cell lines were generated by integration of an empty pSUPER-puro vector into PGK mESCs. *Dgcr8*_KO/*Xrn2*_KD mESC lines were generated using the pSUPER-puro vector engineered to produce the active shRNA 5′-CTCCAGAAGAGAACAGGAGAAAT-3′ and 5′- CTCGGGAAGATACAGTTGGAATT -3′. Upon transfection of the *Dgcr8*_KO lines, cells were selected on puromycin-containing medium. The h*Dcr*-complemented *Dcr*$^{-/-}$ cell line was created by the transfection of plasmid pDESTmycDicer (Addgene plasmid 19873) [60] into *Dcr*$^{-/-}$ mESCs. Stable *Dcr*$^{-/-}$h*Dcr* mESCs were selected on G418-containing medium and analysed at early passages (<P5) and late passages (>P10). MG132 (Z-Leu-Leu-Leu-al; Sigma, C2211) was dissolved DMSO and added to the cells for 7 h to a final concentration of 0.5 µM.

Deep-sequencing

Total cellular RNA (5 µg), extracted using Isol-RNA Lysis Reagent (5PRIME) was processed into sequencing libraries using adapted Illumina protocols and sequenced at Fasteris (http://www.fasteris.com, Switzerland) using the HiSeQ 2000 sequencer. All next-generation sequencing data have been deposited to the NCBI Gene Expression Omnibus (GEO) and are accessible with the accession n° GSE43110 (WT and *Dcr*$^{-/-}$) and GSE43153 (IP_FHA-hAgo2).

sRNA analysis

The sRNA-seq analyses were performed using the ncPRO pipeline [24]. Briefly, the reads were aligned on the mm9 genome using the Bowtie software and allowing multiple matches. Profiling of repeats was estimated from the intersection of the mapped reads with the RepeatMasker annotation. As annotated L1Md_Tf L1 repeats are often truncated or have different full length, the median size of full length was considered, and all LdMd_Tf L1 repeats were scaled to this median size when computing positional read coverage. The positional read coverage was computed by summing up the normalized counts (RPM) of reads covering each position, which was further normalized to the number of L1Md_Tf L1 repeats in the genome containing the position.

PCR

Strand-specific RT-PCR was performed using the Transcriptor Reverse Transcriptase kit (Roche) using 1 µg total RNA and following the manufacturer's instructions. PCR using primer for the specific Tf LINE from chromosome 17 were conducted at 95°C for 10 min, followed by 35 cycles at 95°C for 15 s, 60°C for 30 s, 72°C for 30 s and 10 min at 72°C and revealed on 1% agarose gel. Real-time PCR reagents for miRNAs, 5′_UTR sRNAs and control U6 snRNA were from Qiagen. 5′_UTR sRNAs sense and antisense discrete sequences have been extracted from deep-sequencing data and chose because their higher level of expression. For RT reactions, 1 µg total RNA was reverse-transcribed using the miScript Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions. Following the RT reactions, cDNA products were diluted five times in distilled water, and 2 µL of the diluted cDNAs was used for PCR using QuantiTect SYBR Green PCR Master Mix and miScript Universal Primer (Qiagen). PCR reactions were conducted at 95°C for 10 min, followed by 40 cycles at 95°C for 15 s and 60°C for 30 s on a LightCycler 480 real-time PCR machine (Roche). Real-time PCR for mRNAs was performed as described in [36] using the *Rrm2* as a reference gene. L1 copy-number analysis was conducted on 50 ng of DNA and normalized with *Rrm2* gene (a single-copy gene). Differences between samples and controls were calculated based on the $2^{-\Delta CT}$ method. Each Real-time PCR reaction was carried out in triplicates using samples from three or five independent differentiation events or cell lines for all mESC experiments, unless otherwise stated. Student's T-Test was used to evaluate the statistical significance of Q-PCR analysis of L1 copy Number. Primers used in this study are all listed in Table S1.

L1 copy number analysis

Using the ePCR package from NCBI (http://www.ncbi.nlm.nih.gov/sutils/e-pcr/) we identified 10 806 hits on the mouse mm9 reference genome using the L1_Tf specific primers as designed on the L1spa N°. AF016099 [29]. These primers generate 67 bp amplicons, present in the L1_Tf 5′_UTR repeated regions. Real-time quantitative PCR analysis of the L1_Tf copy number in *Dicer*$^{Flx/Flx}$ mESCs provided a figure of 9,478 at P10 and 10,465 at P30 PCR hits, remarkably close to the ePCR estimation (10,806). The small difference could be explained by the hybrid background of the Dicer mESC line used compared to the genome reference in mm9. The copy number assay shown in Figure 2A, involved a comparison of *Dcr*$^{Flx/Flx}$ mESCs sampled at

passage 10 (9,478 amplicons detected) and at passage 30 (10,465 amplicons detected) with $Dcr^{-/-}$ mESCs sampled at passage 30 (13,821 amplicons detected). Therefore, we estimate that (13,821-10,465) = 3,356 new PCR amplicons were generated, corresponding to 3,356/3.9 = 860 new full length insertion after 20 passages in the $Dcr^{-/-}$ background. Since each passage represents 2 days of culture (40 days in total), 860/40 = 21.51 full length L1_Tf insertions were generated on average every day in $Dcr^{-/-}$ mESCs, although the fraction of active copies among these insertions is unknown. We conclude, therefore, that $Dcr^{-/-}$ mESCs undergo between 1 and 20 L1_Tf retrotransposition events per day.

Cell lysates and immunoprecipitations

E14_FHA-hAgo2 mESCs were scraped in cell lysis buffer (25 mM Tris, pH 7.9, 250 mM KCl, 0.2 mM EDTA, 20% glycerol and Roche Complete Protease Inhibitor without EDTA). Cells were lysed 10 min on ice, sonicated and centrifuged (10 000 rpm, 10 min at 4°C) before Western analysis or immunoprecipitation. Lysates were incubated at 4°C with 20 µL of FLAG-beads (Invitrogen) for 12 h. Beads were collected by centrifugation (2,000 rpm, 1 min). After at least three washes in 1 mL lysis buffer, beads were incubated with 100 µL 0.1 M glycine pH 2.5 for 10 min RT on a shaker. Ten µL 1 M Tris–HCl pH 8 was added to neutralize the elution buffer. Immunoprecipitated RNAs was then extracted from eluted proteins with Isol-RNA Lysis Reagent (5PRIME).

Bisulfite sequencing-based DNA methylation analysis

Genomic DNA was extracted using Isol-RNA Lysis Reagent (5PRIME). Bisulfite treatment was performed using the EpiTect Bisulfite Kit (Qiagen). Bisulfite-treated DNA was then amplified using the DreamTaq DNA Polymerase and primers listed in see accompanying primer list. PCR cycling conditions and primers design were made following the recommendations in [61]. PCR fragments were purified and cloned into pGEM-T Easy (Promega) and individual colonies were sequenced using M13 primers. Sequences were then analysed using Kismeth and BISMA softwares [62,63] to obtain the percentage of methylated sites for each sequence context. Results shown were obtained in two independent experiments.

Antibodies

The following antibodies were used: anti-L1_ORF1 (gift of Dr Alex Bortvin, Carnegie Institution for Science, USA), anti-AGO1 (D84G10, Cell Signaling Technology, Beverly, MA, USA), anti-mouse AGO2 (clone 6F4, gift of Dr Gunter Meister, University of Regensburg, Germany), anti-XRN2 (A301-101A, Lubio Science, Switzerland), anti-EXOSC10 (Rrp6) (ab50558, Abcam, Cambridge, UK) and anti-OCT4 (ab19857, Abcam, Cambridge, UK).

Supporting Information

Figure S1 L1 elements are up-regulated in $Dcr^{-/-}$ mESCs. A. Number of sRNA reads and distinct sequences matching full length retrotransposon of LINE1 from L1Md_T, L1Md_A and L1Md_Gf families. B. Detection of overlapping sense and antisense L1 transcription at the L1_5′-UTR region using strand-specific RT-PCR in WT mESCs. The primer sets used are depicted. C. Accumulation of the Hmga2 and Btg2 mRNAs, respectively known targets for mmu-miR-196a and mmu-let-7a/mmu-miR-132, analyzed by qRT-PCR before and after *Dcr* deletion. D. L1_Tf, Gf and A sub-type mRNAs accumulation detected by qRT-PCR before and after *Dcr* deletion. Polymorphism in the repeated region indicated in the scheme was used to distinguish subtypes.
(EPS)

Figure S2 Methylation and retrotransposition in $Dcr^{-/-}$ ESCs. A. Western analysis of DNMT1 & 3b proteins levels in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and *Dgcr8*_KO mESCs; CM: Coomassie staining of total protein. B. L1_ORF2 mRNA accumulation detected by qRT-PCR in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$, *Dgcr8*_KO and TKO mESCs. C. Bisulfite sequencing-based methylation analysis at the L1 5′_UTR in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$, *Dgcr8*_KO and TKO mESCs. Data were analysed with the Kismeth and BISMA online softwares [62,63]. D. Expression of eGFP detected by qRT-PCR in WT and $Dcr^{-/-}$ mESCs carrying the human eGFP-tagged L1 transgene after 4 (P4) and 6 (P6) passages post-puromycin treatment for selection of stable transformants. L1 constructs lacking 5′UTR (TgΔ5′UTR) or ORF1 (TgΔORF1) were used as negative controls for retrotransposition. Note that TgWT = RP, TgΔORF1 = JM111 and TgΔ5′UTR = 2980 according to the previous nomenclature established in [30].
(EPS)

Figure S3 Deep-sequencing analysis of small RNA libraries. A. Compared size distribution of all deep sequencing reads mapping to the mm9 genome in WT and $Dcr^{-/-}$ sRNAs libraries. B. Pie chart distributions of non-coding RNAs, as annotated by the ncPRO pipeline, in WT and $Dcr^{-/-}$ sRNAs libraries. C. Relative proportions of reads mapping to pre-miRNAs in WT and $Dcr^{-/-}$ sRNAs libraries, as annotated by the ncPRO pipeline. D. 22-nt sequence coverages of L1_Tf elements from WT, $Dcr^{-/-}$ and immunoprecipitated E14-FHA-hAgo2 mESCs, normalized to the total amount of 22-nt reads from corresponding library. E. Size distribution of all reads of RNA isolated from *hAgo2* immunoprecipitates mapping to the mm9 genome. F. Same as in (B) for *hAgo2*-bound sRNAs. G. Same as in (C) for *hAgo2*-bound sRNAs.
(EPS)

Figure S4 L1 expression and genomic copy-number in various knock-out and knock-down mESC lines. A. Western analysis of RRP6 and L1_ORF1 accumulation in WT and *Rrp6*_KD mESCs; CM: Coomassie staining of total protein. B. Accumulation of Tf_5′-UTR (+) and (−) sRNAs detected by qRT-PCR in WT and *Xrn2*_KD mESCs. C. qPCR analysis of L1_Tf copy-number in WT and *Rrp6*_KD mESCs. D. L1_ORF2, Tf, Gf and A sub-type mRNAs accumulation detected by qRT-PCR in *Xrn2*_KD and *Rrp6*_KD mESCs. E. Accumulation of miR-320 detected by qRT-PCR in WT and *Dgcr8*_KO mESCs. F. Western analysis of AGO2 accumulation in WT and *Ago1,2,3,4_KO_hAgo2* mESCs before and after hAgo2 deletion induced by tamoxifen; CM: Coomassie staining of total protein. G. Accumulation of the Hmga2 and Btg2 mRNAs, respectively targeted by mmu-miR-196a and mmu-let-7a/mmu-miR-132, analyzed by qRT-PCR before and after deletion of h*Ago2*. H. mRNA accumulation of L1_Tf, _Gf and _A sub-types detected by qRT-PCR before and after h*Ago2* deletion. I. mRNA accumulation of a single Tf_L1 subtype located on chromosome 17, analyzed by semi-quantitative RT-PCR before and after h*Ago2* deletion.
(EPS)

Figure S5 Expression of AGO2 in $Dcr^{-/-}$ ESCs and microRNA expression in h*Dcr-c*omplemented $Dcr^{-/-}$ ESCs. A. Accumulation of the Ago2 mRNA analyzed by qRT-PCR in WT, $Dcr^{-/-}$ and *Dgcr8*_KO mESCs. B. Endogenous AGO2 protein accumulation in DMSO- and MG132-treated in $Dcr^{-/-}$ mESCs. The data depicted are from two independent treatments. C. MiR-302d and

miR-21 accumulation detected by qRT-PCR in $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and h*Dcr*-complemented $Dcr^{-/-}$ mESCs.
(EPS)

Figure S6 mRNA and microRNA expression in h*Dcr*-complemented $Dcr^{-/-}$ ESCs before and after differentiation. A. Accumulation of Fgf5 (ectoderm marker), Hmga2, Sox2 and L1_ORF2 mRNAs detected by qRT-PCR before (d0) and 10 days after differentiation (d10) of $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and h*Dcr*-complemented $Dcr^{-/-}$ mESCs. B. Accumulation of miR-295, miR-302d, miR-21 and miR-16 analyzed by qRT-PCR before and 10 days after differentiation of $Dcr^{Flx/Flx}$, $Dcr^{-/-}$ and h*Dcr*-complemented $Dcr^{-/-}$ mESCs.
(EPS)

Table S1 Primers table.
(DOCX)

Acknowledgments

We thank members of the Voinnet laboratory for critical reading of the manuscript and for stimulating discussions, Dr. G. Meister and Dr. A. Bortvin for the gift of the mouse AGO2 (6F4) antibody, and L1_ORF1 antibodies, respectively, Dr. X. Wang for the gift of the inducible *Ago1,2,3,4*_KO mESC line and Dr. G. Hannon for the gift of $Dcr^{Flx/Flx}$ mice. Isabelle Grandjean and Patricia Diabangouaya are deeply acknowledged for their key contribution to mice crosses and genotyping at the Curie Institute.

Author Contributions

Conceived and designed the experiments: CC OV. Performed the experiments: CC FJ IO. Analyzed the data: CC FJ CJC AS NS OV. Contributed reagents/materials/analysis tools: EH EB. Wrote the paper: CC OV.

References

1. Beck CR, Garcia-Perez JL, Badge RM, Moran J V (2011) LINE-1 elements in structural variation and disease. Annu Rev Genomics Hum Genet 12: 187–215.
2. Zaratiegui M, Irvine D V, Martienssen R (2007) Noncoding RNAs and gene silencing. Cell 128: 763–776.
3. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, et al. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31: 785–799.
4. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, et al. (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484: 339–344.
5. Howlett SK, Reik W (1991) Methylation levels of maternal and paternal genomes during preimplantation development. Development 113: 119–127.
6. Packer AI, Manova K, Bachvarova RF (1993) A Discrete LINE-1 Transcript in Mouse Blastocysts. Dev Biol 157: 281–283.
7. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, et al. (2009) L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev 23: 1303–1312.
8. Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, et al. (2009) A role for RNAi in the selective correction of DNA methylation defects. Science 323: 1600–1604.
9. Slotkin RK, Vaughn M, Borges F, Tanurdzić M, Becker JD, et al. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell 136: 461–472.
10. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409: 363–366.
11. Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD (1998) How cells respond to interferons. Annu Rev Biochem 67: 227–264.
12. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, et al. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453: 539–543.
13. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev 22: 2773–2785.
14. Arand J, Spieler D, Karius T, Branco MR, Meilinger D, et al. (2012) In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases. PLoS Genet 8: e1002750.
15. Chow JC, Ciaudo C, Fazzari MJ, Mise N, Servant N, et al. (2010) LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell 141: 956–969.
16. Weng A, Magnuson T, Storb U (1995) Strain-specific transgene methylation occurs early in mouse development and can be recapitulated in embryonic stem cells. Development 121: 2853–2859.
17. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, et al. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19: 489–501.
18. Calabrese JM, Seila AC, Yeo GW, Sharp P a (2007) RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. PNAS 104: 18097–18102.
19. Mätlik K, Redik K, Speek M (2006) L1 antisense promoter drives tissue-specific transcription of human genes. J Biomed Biotech 2006: 1–16.
20. Yang N, Kazazian HH (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 13: 763–771.
21. Kazazian H (2011) Mobile DNA: Finding Treasure in Junk. 1st ed. FT Press Science.
22. Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ (2005) Characterization of Dicer-deficient murine embryonic stem cells. PNAS 102: 12135–12140.
23. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 455: 1124–1128.
24. Chen C, Servant N, Toedling J, Sarazin A, Marchais A, et al. (2012) ncPRO-seq: a tool for annotation and profiling analysis of ncRNAs from small RNA-seq. Bioinformatics 28:3–5.
25. Goodier JL, Ostertag EM, Du K, Kazazian HH (2001) A novel active L1 retrotransposon subfamily in the mouse. Genome Res 11: 1677–1685.
26. Ip J, Canham P, Choo KHA, Inaba Y, Jacobs S a, et al. (2012) Normal DNA Methylation Dynamics in DICER1-Deficient Mouse Embryonic Stem Cells. PLoS Genet 8: e1002919.
27. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, et al. (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells 11: 805–814.
28. Ostertag EM, Kazazian HHJ (2001) BIOLOGY OF MAMMALIAN L1 RETROTRANSPOSONS. Annual review of genetics 35: 501–538.
29. Naas TP, DeBerardinis RJ, Moran J V, Ostertag EM, Kingsmore SF, et al. (1998) An actively retrotransposing, novel subfamily of mouse L1 elements. The EMBO journal 17: 590–597.
30. Prak ET, Dodson AW, Farkash EA, Kazazian Jr HH (2003) Tracking an embryonic L1 retrotransposition event. PNAS 100: 1832–1837.
31. Ostertag EM, Prak ET, DeBerardinis RJ, Moran J V, Kazazian HH (2000) Determination of L1 retrotransposition kinetics in cultured cells. NAR 28: 1418–1423.
32. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, et al. (2002) A mouse model of human L1 retrotransposition. Nat Genet 32: 655–660.
33. Xie Y, Rosser JM, Thompson TL, Boeke JD, An W (2011) Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. NAR 39: e16.
34. Xiao S, Xie D, Cao X, Yu P, Xing X, et al. (2012) Comparative epigenomic annotation of regulatory DNA. Cell 149: 1381–1392.
35. Dueck A, Ziegler C, Eichner A, Berezikov E, Meister G (2012) microRNAs associated with the different human Argonaute proteins. NAR 40: 9850–9862.
36. Ciaudo C, Bourdet A, Cohen-Tannoudji M, Dietz HC, Rougeulle C, et al. (2006) Nuclear mRNA degradation pathway(s) are implicated in Xist regulation and X chromosome inactivation. PLoS Genet 2: e94.
37. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet 39: 380–385.
38. Su H, Trombly MI, Chen J, Wang X (2009) Essential and overlapping functions for mammalian Argonautes in microRNA silencing. Genes Dev 23: 304–317.
39. Wang D, Zhang Z, O'Loughlin E, Lee T, Houel S, et al. (2012) Quantitative functions of Argonaute proteins in mammalian development. Genes Dev 26: 693–704.
40. Smibert P, Yang J-S, Azzam G, Liu J-L, Lai EC (2013) Homeostatic control of Argonaute stability by microRNA availability. NSMB: 1–9.
41. Derrien B, Baumberger N, Schepetilnikov M, Viotti C, Cillia J De (2012) Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. PNAS 109: 15942–15946.
42. Johnston M, Geoffroy M, Sobala A, Hay R, Hutvagner G (2010) HSP90 Protein Stabilizes Unloaded Argonaute Complexes and Microscopic P-bodies in Human Cells. Mol Biol Cell 21: 1462–1469.
43. Gy I, Gasciolli V, Lauressergues D, Morel J-B, Gombert J, et al. (2007) Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. The Plant cell 19: 3451–3461.
44. Yamanaka S, Mehta S, Reyes-Turcu FE, Zhuang F, Fuchs RT, et al. (2013) RNAi triggered by specialized machinery silences developmental genes and retrotransposons. Nature 493: 557–560.
45. Swergold GD (1990) Identification, Characterization, and Cell Specificity of a Human LINE-1 Promoter. Mol Cell Biol 10: 6718–6729.
46. Houseley J, Tollervey D (2009) The many pathways of RNA degradation. Cell 136: 763–776.
47. Rosenbloom KR, Sloan C a, Malladi VS, Dreszer TR, Learned K, et al. (2012) ENCODE Data in the UCSC Genome Browser: year 5 update. NAR 41: 56–63.

48. Fagegaltier D, Bouge AL, Berry B, Poisot E, Sismeiro O, et al. (2009) The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. PNAS 106: 21258–21263.
49. Moreno AB, Martínez de Alba AE, Bardou F, Crespi MD, Vaucheret H, et al. (2013) Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. NAR 41: 4699–4708.
50. Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, et al. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453: 534–538.
51. Hu Q, Tanasa B, Trabucchi M, Li W, Zhang J, et al. (2012) DICER- and AGO3-dependent generation of retinoic acid-induced DR2 Alu RNAs regulates human stem cell proliferation. NSMB 19: 1168–1175.
52. Paddison PJ, Caudy A a, Hannon GJ (2002) Stable suppression of gene expression by RNAi in mammalian cells. PNAS 99: 1443–1448.
53. Billy E, Brondani V, Zhang H, Müller U, Filipowicz W (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. PNAS 98: 11443–14428.
54. Maillart P, Ciaudo C, Marchais A, Li Y, Jay F, et al. (2013) Antiviral RNA interference in mammalian cells. In press.
55. Hooper M, Hardy K, Handyside A, Hunter S, Monk M (1987) HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature 326: 292–295.
56. Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, et al. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell 15: 185–197.
57. Bernstein E, Kim SY, Carmell M a, Murchison EP, Alcorn H, et al. (2003) Dicer is essential for mouse development. Nat Genet 35: 215–217.
58. Vooijs M, Jonkers J, Berns a (2001) A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. EMBO reports 2: 292–297.
59. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N, et al. (1996) Requirement for Xist in X chromosome inactivation. Nature 379: 131–137.
60. Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, et al. (2008) Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA 14: 2580–2596.
61. Henderson IR, Chan SR, Cao X, Johnson L, Jacobsen SE (2010) Accurate sodium bisulfite sequencing in plants. Epigenetics 5: 47–49.
62. Gruntman E, Qi Y, Slotkin RK, Roeder T, Martienssen R a, et al. (2008) Kismeth: analyzer of plant methylation states through bisulfite sequencing. BMC bioinformatics 9: 371.
63. Rohde C, Zhang Y, Reinhardt R, Jeltsch A (2010) BISMA–fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC bioinformatics 11: 230.