Transposons and the PIWI pathway: genome defense in gametes and embryos

Stewart J Russell¹ and Jonathan LaMarre²

¹CReATe Fertility Centre, Toronto, Ontario, Canada and ²Department of Biomedical Sciences, University of Guelph, Guelph, Ontario, Canada

Correspondence should be addressed to J LaMarre; Email: jlamarre@uoguelph.ca

Abstract

Hiding in plain sight within the genome of virtually every eukaryotic organism are large numbers of sequences known as transposable elements (TEs). These sequences often comprise 50% or more of the DNA in many mammals and are transcriptionally constrained by DNA methylation and repressive chromatin marks. Individual TEs, when relieved of these epigenetic constraints, can readily move from one genomic location to another, either directly or through RNA intermediates. Demethylation and removal of repressive histone marks during epigenetic reprogramming stages of gametogenesis and embryogenesis render the genome particularly susceptible to increased TE mobilization, which has significant implications for the fidelity of genome replication and subsequent viability of the progeny. Importantly, however, TEs have functionally integrated themselves into developmental events to the extent that complete suppression precludes normal gamete and embryo development. Consequently, multiple mechanisms have evolved to limit the extent of TE expression and mobilization during reprogramming without completely suppressing it. One of the most important TE repression mechanisms is the PIWI/piRNA pathway, in which 25–32 nucleotide RNA molecules known as piRNAs associate with Argonaute proteins from the PIWI clade to form piRISC complexes. These complexes target and silence TEs post-transcriptionally and through the induction of epigenetic changes at the loci from which they are expressed. This review will briefly discuss the intricate molecular détente between TE expression and its suppression by the PIWI pathway, with particular emphasis on mammalian species including human, bovine and murine.

Reproduction (2018) **156** R111–R124

Genomic assailants from within: transposable elements

The trans-generational propagation of genes from parent to offspring is an indispensable component of the process of reproduction. Faithful genome replication is threatened in eukaryotic organisms by viral elements that have the potential to cause mutations and compromise the fitness of subsequent generations or damage the genome such that reproduction is no longer possible. Surprisingly perhaps, the richest source of these viral elements lies within an organism's own DNA in the form of integrated sequences known as transposable elements (TEs) that are often virus derived.

TEs are mobile DNA sequences present within the eukaryotic genome (Fedoroff 2012). They were discovered in plants by Barbara McClintock where they were observed to move within and between chromosomes and subsequently named 'jumping genes' (McClintock 1950). In higher eukaryotes, these sequences constitute a significant fraction of the genome; approximately 50% of human and bovine genomes are composed of TEs and their derivatives (Lander *et al.* 2001). While a large number of these elements are evolutionarily conserved,

more than one quarter appear to be species specific (Adelson *et al.* 2009). Most TEs are fragmented and cannot undergo independent transposition, but some remain intact and competent to mobilize within the genome. The frequency of activation is reflected in the observation that approximately 10% of all spontaneous genome mutations in rodents appear to result from TE mobilization (Kazazian 1998). Surprisingly perhaps, mechanisms to completely eliminate or permanently suppress transposons within the genome have not evolved, raising important questions with respect to the high potential fitness costs of continually replicating large amounts of DNA with no obvious coding or regulatory functions. In particular, the possibility must be considered that TEs provide unrecognized advantages to host genomes.

Transposable elements are categorized in two classes and a number of subclasses based on the mechanism of transposition. Class I elements, also known as retrotransposons, (Wicker *et al.* 2007, 2008) move through RNA intermediates that are transcribed in a manner similar to coding genes. Intact retrotransposons can mobilize through transcription then re-integrate into the genome through the activities of reverse transcriptase

(RT) and endonuclease enzymes that are encoded by the TE (Wicker *et al.* 2007). Common examples of this class are long terminal repeat (LTR)/endogenous retroviruses (ERV) and non-LTR retrotransposons, including the long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (Fig. 1A).

Within the LINE retrotransposons, LINE1 (L1) is an autonomous element (encoding the proteins required for retrotransposition) normally consisting of a few thousand base pairs (bp) (Swergold 1990). It is transcribed similarly to genic mRNA, utilizing RNA Polymerase II (Pol II) followed by 5' capping and polyadenylation. Two open reading frames (ORFs) are present in a complete L1 transcript; ORF1 encodes an RNA-binding protein and ORF2 encodes a protein with RT and endonuclease activities (Ostertag & Kazazian 2001, Martin 2006). Together, ORF1 and ORF2 facilitate L1 RNA translocation to the nucleus, reverse transcription and reintegration into DNA (Ostertag & Kazazian 2001) (Fig. 1B). The L1 element has been described as the only autonomous non-LTR retrotransposon in mammals, with the exception of marsupials and ruminants which have LINE RTE repeats in addition to LINE1 (Deininger *et al.* 2003). The bovine genome encodes nine complete copies of the LINE RTE BovB, which includes a reverse transcriptase similar to L1 (Adelson *et al.* 2009) suggesting that the L1 class of retrotransposons may be active in this species.

In contrast to LINEs, SINEs are non-autonomous and transcribed by Pol III (Kroutter *et al.* 2009). SINEs (and some fragmented LINEs) capitalize on the L1 retrotransposition machinery, specifically requiring ORF2 from L1 for successful retrotransposition (Kroutter *et al.* 2009). There is greater variation in SINE classes between species, with bovine SINEs dominated by BOV-A2, Bov-tA and tRNA insertions (Adelson *et al.* 2009) and human by Alu and SVA type SINEs (Ostertag *et al.* 2003, An *et al.* 2004). SINE elements range from 85 to 500 base pairs and contain an internal Pol III promoter (Deininger 2002). However, the majority of both LINEs and SINEs in mammalian genomes have lost their functional promoters (Deininger 2002).

Endogenous retroviruses (ERVs) are less abundant in the bovine genome, at between 0.4 and 4.3% of total DNA and approximately 13,000 copies, depending on the detection algorithm applied (Garcia-Etxebarria & Jugo 2010). Structurally, ERVs encode the canonical retroviral GAG, POL and ENV proteins over 5–10 kilobases (kb), flanked by LTRs of between 300 and 1200bps. The *gag* gene encodes the core viral proteins that bind the retroviral RNA genome, the *pol* region encodes the RT and integrase responsible for genomic reintegration and the *env* region encodes surface and transmembrane proteins that facilitate receptor binding and membrane fusion (Ryan 2004). Importantly, ERVs show considerable variability in sequence due to the accumulation of mutations, insertions and deletions

Figure 1 Structure and activity of common mammalian transposable elements (TEs). Note: Due to the generally dispensable nature of most TEs, there are many permutations of the structures described here. (A) Long and short interspersed nuclear elements (LINEs and SINEs) are flanked by target site duplications due to the mechanism of LINE1 retrotransposition. Transcription of LINE1 is initiated after association of RNA polymerase (RNA pol) II with the TE promoter sequence (black bent arrow) in the 5' UTR. Complete copies encode two open reading frames (ORF1 and ORF2), followed by the 3' UTR and poly(A) tail. Both ORF1 and ORF2 are required for LINE1 retrotransposition; the former encoding a nucleic acid chaperone and the latter a protein with endonuclease (EN) and reverse transcriptase (RT) domains. SINEs are dependent on LINE1 ORF2 for retrotransposition, and do not encode proteins. Their structures are highly variable, but two conserved primate elements are presented here: Alu and SVA. Alu is transcribed by RNA pol III from a promoter (gray bent arrow) in the left monomer, which terminates downstream of the element at the closest genomic terminator sequence (poly(T)). Alu contains two monomers that form independent secondary RNA structures, an A-rich spacer (A_n), and a poly(A) tail. A full length SVA is a 'repeat of repeats', composed of a CCCTCT repeat, an Alu-like domain, a GC-rich variable number of tandem repeats, and SINE-R derived from the HERV-K LTR element, followed by a poly(A) tail. It is transcribed by both RNA pol II and pol III at either internal or external promoters, and terminates similar to Alu. Finally, ERVs are transcribed by RNA pol II and encode 2–3 proteins: gag (viral core protein), pol (RT, integrase, protease) and sometimes env (viral envelope protein). ERVs are flanked by long terminal repeats (LTRs), a hallmark of their mechanism of retrotransposition. (B) Prior to reprogramming or specific cell stress, TEs are packaged into heterochromatin and their activity is suppressed, primarily through repressive DNA and histone methylation. During reprogramming, repressive marks are removed and TEs may be transcribed. A complete LINE1 element contains two coding domains, which produce ORF1 and ORF2, forming a ribonucleoprotein complex with LINE1 transcripts. Upon translocation to the nucleus, ORF2 creates a single-stranded nick in the DNA, which allows target primed reverse transcription of a new complete or partial (depicted here), LINE1 element in the genome. In contrast, SINEs only require ORF2 to replicate. Adapted from (Ostertag *et al.* 2003, Beck *et al.* 2011, Deininger 2011).

(Stoye 2012). Moreover, recombination events involving LTRs of ERVs can lead to excision of the coding regions, leaving behind 'Solo LTRs' that are much more abundant than full or partial ERVs (Benachenhou *et al.* 2009). Recent studies by Svoboda and colleagues (Franke *et al.* 2017) have demonstrated that LTRs have strong potential to play important roles in gene evolution and expression control in oocytes and early embryos of many different species including bovine.

The second major class of TEs consists of Class II elements, also called DNA-transposons, which replicate through a DNA intermediate, by rolling circle replication or through other unknown mechanisms (Wicker *et al.* 2007). Class II TEs comprise only about 2% of the bovine genome and will not be discussed further here. For a recent discussion of DNA TEs in mammals (see Hickman & Dyda (2015)).

Transposable element expression

The expression of TEs is normally suppressed epigenetically by methylation of cytosine and guanine residues in DNA and repressive marks on histones (Huda *et al.* 2010, Zamudio *et al.* 2015). However, as described in more detail below, TEs become transcribed at higher frequency during periods of genome demethylation, increasing the probability of transposition events. In principle, transposition leads to one of three outcomes: positive, negative or neutral effect on host viability. The net effect depends on the integration site and downstream impacts of the new TE on the expression or function of nearby genes. TEs have contributed substantially to the evolution of the bovine genome in which an average of 18 TE insertions per gene (in regulatory and coding regions) have been identified (Almeida *et al.* 2007). While no formal associations between TEs and heritable diseases in cattle are apparent in the literature, many examples of diseases resulting from retrotransposition are present in humans (Kazazian *et al.* 1988). TE mobilization is also observed in different human cancers, but whether TE expression is the cause, or a consequence, of genomic instability is not clear (reviewed in Belancio *et al.* 2010). Mechanistically, disease-causing TE insertions typically disrupt the coding sequence or processing (i.e. splicing) of mRNAs (Maksakova *et al.* 2006). As with any mutation, TE insertion would be predicted to cause more negative than positive effects on the host cell, and therefore, most cells or progeny in which retrotransposition occur are likely subject to negative selection.

TE expression in embryos

Despite the deleterious potential of TE expression and reintegration, specific TEs and their truncated transcripts appears to be beneficial, or at least required, during embryogenesis where they are frequently expressed

(Beraldi *et al.* 2006). In support of a functional requirement for specific TE expression, suppression of the L1 ORF1 in murine zygotes increases the rate of embryo arrest (Beraldi *et al.* 2006). A recent investigation by Jachowicz *et al.* demonstrated that not only is LINE1 expression required for early mouse embryogenesis, but that it must be subsequently silenced for successful progression to the blastocyst stage (Jachowicz *et al.* 2017). Furthermore knockdown of MuERV-L, one of the first transcripts generated from the quiescent mouse genome after fertilization, leads to a developmental block at the four-cell stage (Kigami *et al.* 2003). Moreover, ample evidence suggests that specific TE transcripts are actively regulated throughout development in many species (Ge 2017).

Evolutionary selection for TE expression may have resulted from the tendency of these elements to contain regulatory domains that are frequently co-opted by genes near the insertion sites (Töhönen *et al.* 2015, Franke *et al.* 2017). The regulatory mechanisms in which they participate can be particularly complex, making them difficult to fully elucidate (Sokol *et al.* 2015). The contribution of TEs to gene regulation in mouse and human genomes has been investigated at the genome level (Faulkner *et al.* 2009). TEs situated upstream of 5' untranslated regions (UTRs) tend to function as enhancers or alternate promoters (Bejerano *et al.* 2006, Franke *et al.* 2017). The promoters for L1 and LINE-like elements are internal to the transcription start site, which prevents promoter loss and provides potential alternative start sites for genes into which they transpose (Swergold 1990). In addition, a large fraction of Refseq genes (~25%) contain one or more TE insertions in their 3' UTRs (Faulkner *et al.* 2009). Globally, these transcripts show decreased expression as TE coverage in their 3' UTRs increases, suggesting that negative regulatory functions can often be ascribed to repeat sequences in this region. In support of this, SINE *A/u* elements contain a poly(T) sequence that contributes approximately 40% of the sequence present in known 3'UTR AU-rich elements (ARE) (An *et al.* 2004). AREs are critical for mRNA stability and regulation through binding of proteins that facilitate or impair mRNA degradation (reviewed in Barreau *et al.* 2005). In addition to altering the regulatory regions of existing genes, TEs participate in the generation of retrogenes – host mRNAs that have undergone reverse transcription, with the resultant cDNAs incorporated into the genome as intron-less genes (Buzdin *et al.* 2003). Recent work by Hendrickson *et al.* (2017) demonstrates that the retrogene paralogs DUX/DUX4 are indispensable for embryo cleavage in mice and humans, highlighting a particularly relevant example of this process in embryos. In a fascinating twist on this observation, transcripts induced by the DUX retrogene product (a transcription factor) include the endogenous retroviruses M/HERVL, directly suggesting

an explanation for the persistence of this TE in the genome. Finally, expression of the maternally inherited factor Stella in oocytes is required for embryo progression past the 4-cell stage (Payer *et al.* 2003). Recently, Stella has been shown to induce the expression of ERVs during the maternal-to-zygotic transition (MZT), which appears to be critical for mouse embryo development (Huang *et al.* 2017), reinforcing the importance of regulated TE expression during development.

The advantages of TE expression

Increased awareness that TEs may confer fitness advantages to their hosts has contributed to a significant shift in the perception of their roles and importance. While previously considered parasitic 'passengers' in the genome that were occasionally responsible for disease, TEs are now more commonly considered 'evolutionary drivers' and 'exaptation mediators' that exist symbiotically in the genome (Elbarbary *et al.* 2016).

In addition to their roles in altering the landscape of regulatory regions in coding genes, a number of studies have demonstrated that TE expression is increased in response to stress. This association is most firmly established in plants, where tissue damage, pathogen exposure or heat stress increases TE expression (reviewed in Wessler 1996, Capy *et al.* 2000). In yeast, Ty retrotransposon expression is increased after ethanol exposure (Stanley *et al.* 2010). TE expression changes in response to stress are not as widely documented in mammals. Interestingly, higher numbers of Alu elements are present in the regulatory elements of genes involved in stress and immune responses (van de Lagemaat *et al.* 2003), suggesting that common regulatory elements may control their expression. Furthermore, SINE-mediated mRNA regulation has been observed in mammals under stressful conditions (reviewed in Elbarbary *et al.* 2016). In addition to stress, hormonal stimuli can drive TE activation, HERV expression in human breast cancer cells and LTR VL30 expression in mouse reproductive organs are increased in response to steroid hormones (Ono *et al.* 1987, Schiff *et al.* 1991). Based on these observations, it is reasonable to speculate that TE expression also depends on environmental stimuli in the context of early mammalian development. Regardless of where TEs transpose within the genome or the stimuli that cause them to do so, selection for the accumulated beneficial changes they induce has clearly contributed to their persistence.

TE mobilization during reprogramming

While emerging evidence suggests that regulated TE expression is associated with normal embryogenesis, unregulated, widespread expression and reintegration of mobile elements during reprogramming poses

significant risks to the genome. Reprogramming refers to the erasure and re-writing of specific epigenetic marks, namely DNA methylation and histone modifications. In the context of reproduction, reprogramming occurs twice: once during gametogenesis and again during early embryogenesis (Morgan *et al.* 2005). The resultant chromatin state leads to TE de-repression (reviewed in Leung & Lorincz 2012). Human embryos express TEs from various classes including L1, Alu, SVA and HERV-K (Adjaye *et al.* 1997, Guo *et al.* 2014, Grow *et al.* 2015). HERV-K transcription is driven by OCT4 and produces functional proteins that are present in the blastocyst and capable of RNA binding, although their developmental roles are unknown (Grow *et al.* 2015).

Two published studies have investigated the expression of TEs during bovine reprogramming. In both *in vitro*-produced (IVP) and -cloned (SCNT) embryos, marked induction of three ERVs was observed in one study (Bui *et al.* 2009). A 1000-fold increase in ERV1_1_BT was seen between the four-cell and morula stages for both the IVP and SCNT groups. Independently, seven TE families were examined between the four-cell and blastocyst stages (Li *et al.* 2016). Low or highly variable expression was observed in four of the seven families, and L1, BovB and ERV1_1_BT showed increased expression from the four-cell to blastocyst stages, with ERV1_1_BT showing a similar increase to that seen in the Bui study. These findings highlight marked changes in TE expression in the bovine embryo during the MZT.

Some inference regarding TE activity can be obtained from evaluating the overall abundance of specific TEs in the genome; L1 BT and BovB comprise 11 and 10% of the bovine genome respectively (Adelson *et al.* 2009). Seventy-three potentially active copies of L1 BT and nine of BovB are present. The SINE element Bov-tA is substantially more abundant, with approximately 1.5 million copies. Based on these data, the L1 BT family of TEs is likely the most active, along with SINE elements that exploit the L1 ORF genes in their replication/reintegration cycles. In the endogenous retrovirus (ERV) category, 24 families and ~13,622 total ERV elements are present (Garcia-Etxebarria & Jugo 2010) suggesting substantial recent activity in evolutionary terms.

Studies on TEs during reprogramming in gametes have focused primarily on the mechanisms responsible for TE suppression, particularly in the mouse, where knockout of several different repressive pathways arrests meiosis through genomic instability in both sexes (Aravin *et al.* 2007a, Carmell *et al.* 2007, Kuramochi-Miyagawa *et al.* 2008, Flemr *et al.* 2013). In particular, widespread activation of the L1 and ERV families during spermatogenesis is observed when repressive mechanisms are inhibited (Aravin *et al.* 2007a, Flemr *et al.* 2013). Activation of TE transcription and its consequences during oogenesis is less extensively studied, although recent studies employing oocyte-

specific deletion of the histone methyltransferase Setdb1 highlight the need for some level of epigenetic suppression of TE expression (Eymery *et al.* 2016). In one example of the requirement for TE activation, LINE1 ORF1 protein is required for exit from meiotic arrest in mouse oocytes, while overexpression inhibits progression to metaphase II (Luo *et al.* 2016). ORF1 is also expressed at variable levels in fetal oocytes, with overexpression leading to meiotic prophase I defects (Malki *et al.* 2014). Studies on the murine LTR TE mouse transcript (MT) expressed in neonatal to mature oocytes (Park *et al.* 2004, Holt *et al.* 2006) suggest that it accounts for 10% of total oocyte transcripts (Peaston *et al.* 2004) although no function is recognized. In support of its importance, RNAi-mediated inhibition of MT expression in GV oocytes inhibited germinal vesicle breakdown (GVBD), implicating MT in oocyte maturation (Peaston *et al.* 2004). Overall, TE regulation appears important to ensure that some level of TE expression occurs during the reprogramming stage of gametogenesis without excessive activation. The functions of TE expression in this context and the evolutionary forces that drive their selection are not known.

Control of TE expression during reprogramming

Given the delicate balance between the positive effects of limited TE expression during reprogramming, and the severe genomic damage that results from TE overexpression, it is perhaps not surprising that a number of distinct pathways have evolved to control their expression. Endogenous siRNAs appear to be particularly important in mouse oocytes, where suppression of Dicer or Ago2 increases the levels of MT transcripts and abnormalities in spindle formation during meiosis I (Watanabe *et al.* 2008, Stein *et al.* 2015). This pathway may be most important in the mouse and rat, which have acquired a unique oocyte-specific variant of Dicer (Dicer_o) (Flemr *et al.* 2013). However, recent studies on L1 targeting endo-siRNAs in the pig suggest that Dicer and Ago2 are also important in other species (Zhang *et al.* 2017). Small RNA pathways in which targeting RNAs derived from tRNAs (known as tRFs) have recently been shown to suppress LTR TEs in mice (Schorn *et al.* 2017). In this pathway, 18 nt tRFs selectively antagonize RT action (an obligate step in replication for this TE class) while 22nt tRFs act as endo-siRNAs, participating in RISC-dependent destruction of TE transcripts. Additional pathways involving the proteins MARF1 (Su *et al.* 2012), Stella (Huang *et al.* 2017), RIF-1 (Li *et al.* 2017) and TRIM28 (Tao *et al.* 2018) are also important in TE repression and act through epigenetic and transcriptional modulation, but lie beyond the scope of this review. The most widely studied pathway involved in the control of TE expression is the PIWI/piRNA pathway.

PIWI proteins and piRNAs – permissive, targeted TE defense

A clade of Argonaute family members known as PIWI proteins and their associated small (25–32 nt) RNAs form the central components of a genome protection mechanism known as the PIWI/piRNA pathway (Aravin *et al.* 2006, Saito *et al.* 2006, Vagin *et al.* 2006, Brennecke *et al.* 2007). The pathway was originally described in *Drosophila* gametes where its role in limiting TE expression was first recognized (Lin & Spradling 1997, Aravin *et al.* 2001, Deng & Lin 2002, Vagin *et al.* 2006). The name is derived from the phenotype ‘*P*-element Induced WImpy testis’ (PIWI), where a mutation in the PIWI gene of *Drosophila* was induced by a Class II transposon (P-element) to which some strains are particularly susceptible (Lin & Spradling 1997). The importance in reproduction was immediately evident as both male and female progeny are sterile (Lin & Spradling 1997). The first identified mechanism of action involved the interaction of PIWI proteins with piRNAs of specific sequences complementary to RNA targets (i.e. expressed TEs) to form a piRNA-induced silencing complex (piRISC) that specifically cleaves the RNA target through RNase activity of the complex (Aravin *et al.* 2006, Girard *et al.* 2006, Gunawardane *et al.* 2007). Additional potential roles in targeting mRNA transcripts have recently been reported in germline and somatic cells (Rajaseethupathy *et al.* 2012, Chen *et al.* 2013, Kwon *et al.* 2014, Nandi *et al.* 2016, Russell *et al.* 2017), although the extent and importance of this targeting is not yet clear. This review will focus primarily on the established roles of PIWI proteins and their associated piRNAs in the control of TE expression.

A systematic examination of mutations that impacted germline stem cell viability led to the identification of the first PIWI gene in *Drosophila* (Lin & Spradling 1997) where PIWI gene mutations result in cellular loss (Cox *et al.* 1998). The *Drosophila* PIWI proteins, Argonaute3 (Ago3) and Aubergine (Aub), were subsequently characterized based on their structural similarity and co-localization with PIWI in the pole plasm (Wilson *et al.* 1996, Harris & Macdonald 2001, Williams & Rubin 2002). Structurally all proteins of the PIWI family contain a PIWI domain, which has RNase activity, in addition to N-terminal, PIWI/Argonaute/Zwille and MID domains (Cora *et al.* 2014). The importance of the PIWI pathway in mammalian reproduction was first demonstrated in the mouse, where the PIWI homolog MIWI (PIWI-like 1 or PIWIL1) was shown to be essential for spermatogenesis (Deng & Lin 2002). Mouse PIWIL1 is expressed during spermatogenesis between the pachytene stage and the round spermatid stage where it silences TEs and specific mRNAs (Reuter *et al.* 2011, Gou *et al.* 2014). A similar expression pattern is observed in the bull (Russell *et al.* 2016) and dog (Stalker *et al.* 2016). The key mechanism of PIWI action was identified when it was shown to

bind a class of small RNAs that guide complementary interactions with target RNA transcripts including TEs (Aravin *et al.* 2006, Girard *et al.* 2006, Lau *et al.* 2006, Brennecke *et al.* 2007, Carmell *et al.* 2007).

Two additional murine PIWI homologs were identified in subsequent studies: MILI (PIWIL2) and MIWI2 (PIWIL4) (Aravin *et al.* 2006, Carmell *et al.* 2007). PIWIL2 is thought to be the oldest PIWI protein in evolutionary terms (Kuramochi-Miyagawa *et al.* 2008) and has the most extensive period of expression throughout spermatogenesis, starting at the primordial germ cell stage *in utero* and persisting to the pachytene stage (Kuramochi-Miyagawa *et al.* 2004, Aravin *et al.* 2006). Mice lacking PIWIL2 develop genomic instability and elevated TE expression after which spermatogenesis becomes blocked in prophase of meiosis I (Kuramochi-Miyagawa *et al.* 2004, Reuter *et al.* 2009). PIWIL2 is localized to the cytoplasm, primarily in perinuclear granules, a known site for RNA processing (Aravin *et al.* 2008). PIWIL4 is found in the pre-natal testis and is expressed from 15.5 days post-coitum until 3 days after birth in prospermatogonia (Aravin *et al.* 2008) coinciding with an important period of *de novo* DNA methylation (Lees-Murdock *et al.* 2003, Maatouk *et al.* 2006). PIWIL4 distribution is primarily nuclear, with limited expression in cytoplasmic granules (Aravin *et al.* 2008). One potentially important but less thoroughly studied member of the PIWI clade is PIWIL3, which is encoded in human, bovine (Roovers *et al.* 2015) and other genomes, but absent in the mouse and rat where most functional studies have been performed. Recent evidence suggests that PIWIL3 expression is normally limited to oocytes and embryos (Roovers *et al.* 2015, Virant-Klun *et al.* 2016, S J Russell and J LaMarre, unpublished observations). The roles and importance of this protein are largely unknown.

Regulation of piRNA precursor expression from genomic 'clusters'

Most mature piRNAs are generated from longer RNA transcripts originating in specific piRNA-rich genomic regions called piRNA clusters (Brennecke *et al.* 2007). piRNA clusters consist mostly of repetitive element and transposon fragments that have integrated into the genome throughout the course of evolution and which reflect the history of transposition and virus integration into the genome of a species and its evolutionary ancestors (Aravin *et al.* 2007a). Several hundred to a few thousand clusters are present in the mammalian genomes examined (Lau *et al.* 2006, Aravin *et al.* 2007a, Brennecke *et al.* 2007, Gebert *et al.* 2015, Russell *et al.* 2017). The factors that regulate transcription from these clusters (enhancers, epigenetic status) remain poorly characterized. Transcription of piRNA precursor RNAs

can occur from a single, or from both strands of DNA, and are respectively classified as uni- or dual-strand clusters (Aravin *et al.* 2006, Lau *et al.* 2006, Brennecke *et al.* 2007, Carmell *et al.* 2007). Uni- or bi-directional transcription from uni-strand clusters is mediated by RNA polymerase II (Pol II) from one transcription start site, precursors are then polyadenylated, 5' capped and sometimes subjected to alternative splicing (Goriaux *et al.* 2014). Expression from dual-strand clusters may be due in part to overlapping, run-through transcription from nearby genes (Mohn *et al.* 2014). However, recent studies in flies have identified a novel heterochromatin-dependent transcriptional process through which the basal transcription machinery containing a variant of transcription factor II A named 'Moonshiner' associates with another protein known as Rhino to license transcription from heterochromatin at piRNA clusters (Andersen *et al.* 2017).

Two distinct 'waves' of piRNA expression occur in murine testes (Vourekas *et al.* 2012) and the piRNAs from these waves are named based on the timing of their expression. Pre-pachytene piRNAs are present during early spermatogenesis and are primarily targeted at TEs and associate with PIWIL2 and PIWIL4 (Aravin *et al.* 2008, Li *et al.* 2013). Pachytene piRNAs are abundantly expressed in pachytene spermatocytes and round spermatids in adult mouse testes and physically associate with PIWIL1 and PIWIL2 (Aravin *et al.* 2006, Girard *et al.* 2006). Transcription of pachytene piRNA precursors is driven by the transcription factor A-MYB, which also facilitates PIWIL1 transcription, increasing processing 'capacity' for the cluster-derived precursor piRNAs induced at that time (Li *et al.* 2013). Interestingly, in addition to TEs, pachytene piRNAs may also target mRNAs and have been strongly implicated in the widespread active decay of mRNAs that occurs late in spermatogenesis (Gou *et al.* 2014). PIWI pathway targeting of mRNAs may not be limited to testes, as piRNAs that appear to target mRNAs have also been recognized in bovine oocytes (Russell *et al.* 2017).

Export and processing – primary piRNA biogenesis

The generation of piRNAs from primary cluster transcripts is termed *primary* piRNA biogenesis (Aravin *et al.* 2007b). After transcription, piRNA transcripts move from the nucleus to the nuage, a peri-nuclear 'cloud' containing structures also known as germ granules, Yb bodies, chromatoid bodies, pi-bodies and piP-bodies depending on the cell type and content (Buchan & Parker 2009, Ishizu *et al.* 2012, Pillai & Chuma 2012, Meikar *et al.* 2014). The nuage contains mRNA processing machinery and is a known site of decay, storage and quality assessment for RNA (Buchan & Parker 2009, Kulkarni *et al.* 2010, Meikar *et al.* 2014). Transcription of piRNA precursors is

coupled to the processing machinery in the nuage by the DEAD box helicase UAP56 (Shen 2009, Zhang *et al.* 2012). Precursors to piRNAs remain single stranded (in contrast to siRNAs and miRNAs) or assume unrecognized structures prior to processing (Houwing *et al.* 2007). Nucleo-cytoplasmic shuttling is facilitated by the protein Maelstrom (Mael) in *Drosophila* and mice, which interestingly plays different roles in each species (Aravin *et al.* 2009, Sienski *et al.* 2012). In *Drosophila*, Mael participates in nuclear silencing of transposons in conjunction with PIWI but is not required for piRNA biogenesis (Sienski *et al.* 2012). In contrast, Mael is required for piRNA biogenesis during pachytene spermatogenesis in mice and binds piRNA precursors that associate with the nuage. Inactivation of Mael results in meiotic arrest and the dissociation of perinuclear granules containing PIWIL4, leading to failure of LINE1 repression (Aravin *et al.* 2009). These findings implicate Maelstrom in both pre-pachytene PIWIL4 functions and pachytene piRNA biogenesis.

Exported piRNA precursors are processed in several subsequent steps in the nuage, all of which appear closely associated with the mitochondria. An endonuclease known as Zucchini (Zuc; mZuc or mitoPLD in the mouse) processes precursors into intermediate fragments (Pane *et al.* 2007, Watanabe *et al.* 2011a, Ipsaro *et al.* 2012, Nishimasu *et al.* 2012) after the RNA helicase Armitage (Armi; MOV10L1 in the mouse), unwinds RNA secondary structure (Frost *et al.* 2010, Vourekas *et al.* 2015). Physical loading of piRNA intermediates onto PIWI proteins in the nuage requires the co-chaperones shutdown (Shu) and heat shock protein 90 (HSP90) (Olivieri *et al.* 2012, Izumi *et al.* 2013).

The next step in piRNA biogenesis is 3' trimming, which occurs after intermediates are loaded onto PIWI proteins. Mature piRNA length (extent of trimming) appears to depend on the specific PIWI protein to which it is bound – the modal size of mature piRNAs bound to PIWI1, PIWI2 and PIWIL4 are 29–30, 26–27, 27–28 nt respectively (Kuramochi-Miyagawa *et al.* 2008, Vourekas *et al.* 2012). An enzyme known as Nibbler is a mut-7 homolog in *Drosophila* which likely mediates the 3'–5' exonuclease activity (Han *et al.* 2011, Kawaoka *et al.* 2011, Wang *et al.* 2016) for piRNAs bound to PIWI proteins, however the mammalian homolog is unidentified. Nibbler activity depends on TDRKH (also called Papi), which is present at the mitochondrial membrane (Chen *et al.* 2009, Honda *et al.* 2013, Saxe *et al.* 2013). In the final processing step, fully cleaved piRNAs, when bound to PIWI proteins, become 3'-methylated through the activity of Hen1/Pimet, which produces a 2'-O-methyl modification (Horwich *et al.* 2007, Saito *et al.* 2007). At the completion of biogenesis, the mature piRNA–PIWI protein complex is considered a 'piRNA-induced silencing complex' (piRISC).

The ping-pong amplification loop – generation of secondary piRNAs

One of the most unique elements of the piRNA pathway is a secondary 'feed-forward' biogenesis process known as the 'Ping-Pong' amplification loop. This loop is considered the secondary piRNA biogenesis pathway (Aravin *et al.* 2008). In *Drosophila*, piRNAs generated through primary biogenesis and bound to Aubergine target TEs in the cytoplasm leading to cleavage of the target RNA through RNase slicer activity (Brennecke *et al.* 2007, Gunawardane *et al.* 2007). Rather than becoming degraded with the target, the resulting cleaved RNA products load directly onto another PIWI protein, Ago3, as intermediate piRNAs that then become trimmed and 2'-O-Me modified as described above (Han *et al.* 2015, Wang *et al.* 2016). These secondary Ago3 piRISC complexes resulting from this 'ping' step may then reciprocally target primary piRNA precursors ('pong' step) and amplify the number of targeting piRNAs available for silencing through the PIWI pathway (Brennecke *et al.* 2007). The reciprocal cleavage events, combined with the positional nature of cleavage through this pathway result in a ten nucleotide overlap in piRNA populations that have undergone ping-pong amplification (Aravin *et al.* 2008). This ping-pong 'signature' of overlapping complementary nucleotides is visible as a defined peak when plotted (for examples see Russell *et al.* 2017).

One important consequence of secondary piRNA generation through the ping-pong cycle is higher sequence diversity in piRISC complexes. Cleavage products from Ago3 in *Drosophila* when loaded onto PIWI, can direct transcriptional gene silencing (TGS) (Fig. 2; Senti *et al.* 2015). Similarly, mouse PIWIL2-cleaved targets can associate with PIWIL4, which also direct TGS in the nucleus after translocation (De Fazio *et al.* 2011). Primary and secondary piRNA biogenesis require a number of cofactors. Anchoring PIWI proteins to the scaffold in the nuage are a family of Tudor domain-containing proteins (TDRDs), TDRD9, TDRD1 and TDRDKH, that bind to symmetric dimethylarginines (sDMAs) on the PIWI proteins (Reuter *et al.* 2009, Shoji *et al.* 2009, Saxe *et al.* 2013). In the *Drosophila* nuage, the proteins Spindle-E, Krimper, Tejas, Tapas and Qin (Lim & Kai 2007, Malone *et al.* 2009, Patil & Kai 2010, Zhang *et al.* 2011) are involved as reviewed extensively elsewhere (Czech & Hannon 2016). In the mouse, a gamete-specific protein known as GTSF1 (Yoshimura *et al.* 2018) that was previously implicated in other aspects of PIWI/piRNA function (Dönertas *et al.* 2013, Ohtani *et al.* 2013) has very recently been shown to be required for secondary piRNA generation. The protein Vasa appears to play an important part in piRNA biogenesis in the silkworm model of the PIWI pathway. Here, piRNA intermediates generated by PIWIL1 bind to Vasa and are protected from degradation, facilitating

Figure 2 Mechanisms of regulation by the PIWI pathway. (A) Post-transcriptional gene regulation (PTGS) of transposable elements (TE) or mRNA by piRNA-induced silencing complexes (piRISC). piRISCs can regulate target RNA by recruiting the CAF1 deadenylation complex or through endonuclease activity and target degradation. (B) Transcriptional gene silencing is mediated by recruitment of DNA methyltransferases or chromatin components to maintain repression of TE loci.

their transfer to Ago3 (Xiol *et al.* 2014). Similarly, in mice, the Mouse Vasa Homolog (MVH) participates in retrotransposon suppression during fetal male and female gametogenesis; targeted MVH knockout generates a phenotype remarkably similar to PIWIL2 or PIWIL4 (Kuramochi-Miyagawa *et al.* 2010, Lim *et al.* 2013) supporting a conserved role across species.

An elegant series of recent studies (Han *et al.* 2015, Mohn *et al.* 2015) have identified mechanisms by which a subset of piRNAs are 'phased' during the biogenesis process, further increasing the variability of targeting piRNAs generated by the pathway. Phased piRNA generation proceeds when PIWI proteins (PIWI in the fly and PIWIL2 in the mouse) are recruited to the 5' end of precursor transcripts and define a cleavage position 3' to the complex. The endonuclease Zucchini (MitoPLD in the mouse) is then recruited and cleaves the precursor transcript at this 3' position, defining the 3' end of the piRNA and leaving a new 5' terminus on the precursor, upon which the process repeats. In the mouse the piRNAs generated are further processed by Nibbler in association with TDRKH (Hayashi *et al.* 2016). The multiple pathways described highlight the complexity of biogenesis and suggest that additional factors, which participate in piRNA biogenesis and TE silencing are likely to be identified as the pathway continues to be characterized.

Mechanisms of piRNA suppression of TEs

Through the generation of piRISC complexes, the PIWI pathway suppresses TE and gene expression through

several distinct mechanisms. Complementary RNA targets such as retrotransposon-encoded RNAs are cleaved by RNase slicer activity (Fig. 1A) (reviewed in Hiraoka & Siomi 2016). Although less extensively studied, piRISC complexes also help induce epigenetic changes in chromatin conformation and DNA methylation patterns at specific loci in flies (Fig. 1B; Sienski *et al.* 2012) and mice (Kuramochi-Miyagawa *et al.* 2008). Cleavage of TE RNA transcripts by piRNA-guided slicer activity of piRISC complexes is the best-studied and first-recognized function of the PIWI pathway (Reuter *et al.* 2011, Zhang *et al.* 2015). This post-transcriptional silencing pathway mediated by PIWI proteins participates in both TEs and mRNA cleavage. Targeting and cleavage by piRISC complexes appears more stringent than for some other RISC complexes, requiring a minimum of 16–22 continuous bases for efficient target decay (Zhang *et al.* 2015, Yuan *et al.* 2016). In spite of this, the canonical post-transcriptional silencing by PIWI proteins has many similarities with other Argonaute-dependent RISC pathways that have been reviewed extensively elsewhere (Wilczynska & Bushell 2015). In addition to mediating classical RNA cleavage events, some studies have suggested that murine PIWIL1 can also direct target decay through recruitment of the deadenylase CAF1 during spermiogenesis (Fig. 2; Gou *et al.* 2014). For targeting through this mechanism, the requirement for piRISC-target complementarity is not strict and mismatches similar to those observed in miRNA-target interactions are possible (Gou *et al.* 2014). Follow-up studies will be required to determine the importance and scope of this pathway in cells where PIWIL1 is expressed.

Although it has been most widely studied in the context of post-transcriptional destruction of RNAs originating from genomic TE sequences, the PIWI pathway has also been implicated in TGS through directed DNA methylation and histone modifications (Fig. 1B; Carmell *et al.* 2007, Klenov *et al.* 2014). In studies with *Drosophila*, Piwi localization in both the nucleus and cytoplasm was demonstrated and Piwi depletion changed the pattern of chromatin silencing of specific TEs (Klenov *et al.* 2011). Subsequently, Piwi binding to nascent RNAs transcribed from target loci was shown to result in the deposition of repressive H3K9me3 marks on nearby histones (Le Thomas *et al.* 2013, Post *et al.* 2014). Further epigenetic influence was observed with a mutated, slicer-incompetent Piwi (incapable of endonucleolytic cleavage) which retained the ability to transcriptionally silence TE expression through the prevention of RNA Pol II occupancy at specific loci in conjunction with the cofactor Mael (Klenov *et al.* 2011, Sienski *et al.* 2012). Mammalian PIWIL4 has similar roles in transcriptional repression, but mediates changes in DNA methylation rather than altering histones (Aravin *et al.* 2008). One early hypothesis in which PIWIL2 was thought to generate piRNAs that subsequently

bind to PIWI4 and direct targeted methylation has recently been challenged by reports showing substantial differences in TE repression in mouse testes by PIWI2 and PIWI4 (Manakov *et al.* 2015). While abundant evidence suggests that the PIWI pathway participates in both post-transcriptional and transcriptional TE silencing, it is important to note that the relative importance of the two processes is likely to vary with the stage of gamete differentiation or embryo development during which it is expressed, as demonstrated very recently during spermatogenesis (Inoue *et al.* 2017). The mechanisms and consequences of epigenetic changes directed by the PIWI pathway remain very active areas of investigation and are likely to yield fascinating insight into reprogramming and development.

Our understanding of the PIWI pathway has been most strongly and consistently informed by the many studies of its different roles in TE suppression. However, any review of this pathway would be incomplete without briefly mentioning several emerging roles for the pathway in the control of gene expression during gamete and embryo development. One of the first described examples of this phenomenon is presented in studies demonstrating PIWI-dependent, piRNA-directed DNA methylation of the paternally imprinted *Rasgrf1* locus in mice (Watanabe *et al.* 2011b). Through targeting TE sequences within a differentially methylated region adjacent to this locus, piRNAs direct methylation and silencing. Importantly however, piRNAs do not direct all imprinting events, many of which are unaffected after PIWI pathway suppression (Watanabe *et al.* 2011b). While piRNA-directed imprinting processes appear to require TE sequences, several recent studies in gametes suggest that mRNAs can also act as direct PIWI pathway targets, particularly during spermatogenesis (Gou *et al.* 2014) and, more recently, in oocytes (Russell *et al.* 2017). The mechanisms through which such targeting has evolved and its overall importance in fertility and gamete development remain active fields of study.

Conclusions

The structure and function of non-coding regions of the genome, much of which consists of TE sequences, has only recently become the focus of intensive study. Normally, TEs are epigenetically silenced, but it has become apparent that their limited activation during reprogramming is an essential feature of gametogenesis and development, contributing directly to these processes through mechanisms that remain largely uncharacterized. The evolutionary advantages selecting for some level of TE expression may lie in the expanded regulatory potential they confer after reintegration within regulatory domains, a process that has contributed sequences to more than one quarter of human coding genes (Faulkner *et al.* 2009) and driven adaptive change.

However, excessive TE activation is clearly detrimental and leads to infertility and developmental problems. These combined observations suggest that, rather than completely eliminating TE expression, control systems have evolved to promote a delicate balance between permission and suppression of TE activation during reprogramming. The characteristics of the PIWI pathway are consistent with this – TEs are suppressed only after the piRNAs that target them are transcribed and processed. A secondary amplification loop that depends on the presence of TE transcripts enhances suppression once initiated. Phasing and trimming of piRNAs during the biogenesis process expands the repertoire of targets and helps specify the nature of target suppression (post-transcriptional vs transcriptional). Evolutionary expansion of the PIWI pathway to target coding sequences, possibly as a consequence of TE integration into coding and transcriptional regulatory regions of the genome, have expanded functionality in ways we are only just beginning to understand. Future studies on the PIWI pathway will need to resolve a number of key issues; the identity of many cofactors in mammalian species, the factors that initiate piRNA biogenesis, particularly in somatic cells and the mechanisms by which they do so are particularly important questions. With respect to reproduction in particular, the overall importance of the PIWI pathway in mammalian oogenesis and embryogenesis remain unknown. However, regardless of the specific importance of the PIWI pathway itself in any one species, the ongoing tension between ancient virus sequences that hide within host DNA, and the defense pathways that keep them in check, continues to shape the genomic landscape and profoundly influence the control of gene expression that guides gametogenesis and the earliest phases of embryonic development.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this review.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Acknowledgements

The authors wish to thank Clifford Librach, Leanne Stalker, Allison Tscherner, Graham Gilchrist, Pavneesh Madan and Allan King for helpful discussions. Studies on the PIWI pathway in gametes and embryos in the LaMarre laboratory have been supported by NSERC (Canada), the Ontario Ministry of Agriculture Food and Rural Affairs and by a Michelson Grant in Reproductive Biology from the Michelson Foundation.

References

- Adelson DL, Raison JM & Edgar RC 2009 Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *PNAS* **106** 12855–12860. (<https://doi.org/10.1073/pnas.0901282106>)
- Adjaye J, Daniels R, Bolton V & Monk M 1997 cDNA libraries from single human preimplantation embryos. *Genomics* **46** 337–344. (<https://doi.org/10.1006/geno.1997.5117>)
- Almeida LM, Silva IT, Silva WA Jr, Castro JP, Riggs PK, Carareto CM & Amaral MEJ 2007 The contribution of transposable elements to Bos taurus gene structure. *Gene* **390** 180–189. (<https://doi.org/10.1016/j.gene.2006.10.012>)
- An HJ, Lee D, Lee KH & Bhak J 2004 The association of Alu repeats with the generation of potential AU-rich elements (ARE) at 3' untranslated regions. *BMC Genomics* **5** 97. (<https://doi.org/10.1186/1471-2164-5-97>)
- Andersen PR, Tirian L, Vunjak M & Brennecke J 2017 A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549** 54–59. (<https://doi.org/10.1038/nature23482>)
- Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM & Gvozdev VA 2001 Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current Biology* **11** 1017–1027. ([https://doi.org/10.1016/S0960-9822\(01\)00299-8](https://doi.org/10.1016/S0960-9822(01)00299-8))
- Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T *et al.* 2006 A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442** 203–207. (<https://doi.org/10.1038/nature04916>)
- Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K & Hannon GJ 2007a Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316** 744–747. (<https://doi.org/10.1126/science.1142612>)
- Aravin AA, Hannon GJ & Brennecke J 2007b The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318** 761–764. (<https://doi.org/10.1126/science.1146484>)
- Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T & Hannon GJ 2008 A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Molecular Cell* **31** 785–799. (<https://doi.org/10.1016/j.molcel.2008.09.003>)
- Aravin AA, van der Heijden GW, Castañeda J, Vagin VV, Hannon GJ & Bortvin A 2009 Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genetics* **5** e1000764. (<https://doi.org/10.1371/journal.pgen.1000764>)
- Barreau C, Paillard L & Osborne HB 2005 AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Research* **33** 7138–7150. (<https://doi.org/10.1093/nar/gki1012>)
- Beck CR, Garcia-Perez JL, Badge RM & Moran JV 2011 LINE-1 elements in structural variation and disease. *Annual Review of Genomics and Human Genetics* **12** 187–215. (<https://doi.org/10.1146/annurev-genom-082509-141802>)
- Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, Rubin EM, Kent WJ & Haussler D 2006 A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **441** 87–90. (<https://doi.org/10.1038/nature04696>)
- Belancio VP, Roy-Engel AM & Deininger PL 2010 All y'all need to know 'bout retroelements in cancer. *Seminars in Cancer Biology* **20** 200–210. (<https://doi.org/10.1016/j.semcancer.2010.06.001>)
- Benachou F, Jern P, Oja M, Sperber G, Blikstad V, Somervuo P, Kaski S & Blomberg J 2009 Evolutionary conservation of orthoretroviral long terminal repeats (LTRs) and ab initio detection of single LTRs in genomic data. *PLoS ONE* **4** e5179. (<https://doi.org/10.1371/journal.pone.0005179>)
- Beraldi R, Pittoggi C, Sciamanna I, Mattei E & Spadafora C 2006 Expression of LINE-1 retrotransposons is essential for murine preimplantation development. *Molecular Reproduction and Development* **73** 279–287. (<https://doi.org/10.1002/mrd.20423>)
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R & Hannon GJ 2007 Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128** 1089–1103. (<https://doi.org/10.1016/j.cell.2007.01.043>)
- Buchan JR & Parker R 2009 Eukaryotic stress granules: the ins and outs of translation. *Molecular Cell* **36** 932–941. (<https://doi.org/10.1016/j.molcel.2009.11.020>)
- Bui LC, Evsikov AV, Khan DR, Archilla C, Peynot N, Hénaut A, Le Bourhis D, Vignon X, Renard JP & Duranthon V 2009 Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. *Reproduction* **138** 289–299. (<https://doi.org/10.1530/REP-09-0042>)
- Buzdin A, Gogvadze E, Kovalskaya E, Volchikov P, Ustyugova S, Illarionova A, Fushan A, Vinogradova T & Sverdlov E 2003 The human genome contains many types of chimeric retrogenes generated through in vivo RNA recombination. *Nucleic Acids Research* **31** 4385–4390. (<https://doi.org/10.1093/nar/gkg496>)
- Capy P, Gasperi G, Biémont C & Bazin C 2000 Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85** 101–106. (<https://doi.org/10.1046/j.1365-2540.2000.00751.x>)
- Carmell MA, Girard A, van de Kant HJG, Bourc'his D, Bestor TH, de Rooij DG & Hannon GJ 2007 MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Developmental Cell* **12** 503–514. (<https://doi.org/10.1016/j.devcel.2007.03.001>)
- Chen C, Jin J, James DA, Adams-Cioaba MA, Park JG, Guo Y, Tenaglia E, Xu C, Gish G, Min J *et al.* 2009 Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *PNAS* **106** 20336–20341. (<https://doi.org/10.1073/pnas.0911640106>)
- Chen C, Liu J & Xu G 2013 Overexpression of PIWI proteins in human stage III epithelial ovarian cancer with lymph node metastasis. *Cancer Biomarkers: Section A of Disease Markers* **13** 315–321. (<https://doi.org/10.3233/CBM-130360>)
- Cora E, Pandey RR, Xiol J, Taylor J, Sachidanandam R, McCarthy AA & Pillai RS 2014 The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. *RNA* **20** 773–781. (<https://doi.org/10.1261/rna.044701.114>)
- Cox DN, Chao A, Baker J, Chang L, Qiao D & Lin H 1998 A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes and Development* **12** 3715–3727. (<https://doi.org/10.1101/gad.12.23.3715>)
- Czech B & Hannon GJ 2016 One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends in Biochemical Sciences* **41** 324–337. (<https://doi.org/10.1016/j.tibs.2015.12.008>)
- De Fazio S, Bartonicek N, Di Giacomo M, Abreu-Goodger C, Sankar A, Funaya C, Antony C, Moreira PN, Enright AJ & O'Carroll D 2011 The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **480** 259–263. (<https://doi.org/10.1038/nature10547>)
- Deininger PL 2002 Mammalian retroelements. *Genome Research* **12** 1455–1465. (<https://doi.org/10.1101/gr.282402>)
- Deininger PL 2011 Alu elements: know the SINEs. *Genome Biology* **12** 236. (<https://doi.org/10.1186/gb-2011-12-12-236>)
- Deininger PL, Moran JV, Batzer MA & Kazazian HH Jr 2003 Mobile elements and mammalian genome evolution. *Current Opinion in Genetics and Development* **13** 651–658. (<https://doi.org/10.1016/j.gde.2003.10.013>)
- Deng W & Lin H 2002 Miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Developmental Cell* **2** 819–830. ([https://doi.org/10.1016/S1534-5807\(02\)00165-X](https://doi.org/10.1016/S1534-5807(02)00165-X))
- Dönertas D, Sienski G & Brennecke J 2013 Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes and Development* **27** 1693–1705. (<https://doi.org/10.1101/gad.221150.113>)
- Elbarbary RA, Lucas BA & Maquat LE 2016 Retrotransposons as regulators of gene expression. *Science* **351** aac7247. (<https://doi.org/10.1126/science.aac7247>)
- Eymery A, Liu Z, Ozonov EA, Stadler MB & Peters AHFM 2016 The methyltransferase Setdb1 is essential for meiosis and mitosis in mouse oocytes and early embryos. *Development* **143** 2767–2779. (<https://doi.org/10.1242/dev.132746>)
- Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T *et al.* 2009 The regulated retrotransposon transcriptome of mammalian cells. *Nature Genetics* **41** 563–571. (<https://doi.org/10.1038/ng.368>)
- Fedoroff NV 2012 Presidential address. Transposable elements, epigenetics, and genome evolution. *Science* **338** 758–767. (<https://doi.org/10.1126/science.1238610>)
- Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovick K & Svoboda P 2013 A retrotransposon-driven dicer isoform directs

- endogenous small interfering RNA production in mouse oocytes. *Cell* **155** 807–816. (<https://doi.org/10.1016/j.cell.2013.10.001>)
- Franke V, Ganesh S, Karlic R, Malik R, Pasulka J, Horvat F, Kuzman M, Fulka H, Cernohorska M, Urbanova J *et al.* 2017 Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Research* **27** 1384–1394. (<https://doi.org/10.1101/gr.216150.116>)
- Frost RJA, Hamra FK, Richardson JA, Qi X, Bassel-Duby R & Olson EN 2010 MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *PNAS* **107** 11847–11852. (<https://doi.org/10.1073/pnas.1007158107>)
- García-Etxebarria K & Jugo BM 2010 Genome-wide detection and characterization of endogenous retroviruses in *Bos taurus*. *Journal of Virology* **84** 10852–10862. (<https://doi.org/10.1128/JVI.00106-10>)
- Ge SX 2017 Exploratory bioinformatics investigation reveals importance of 'junk' DNA in early embryo development. *BMC Genomics* **18** 200. (<https://doi.org/10.1186/s12864-017-3566-0>)
- Gebert D, Ketting RF, Zischler H & Rosenkranz D 2015 piRNAs from pig testis provide evidence for a conserved role of the piwi pathway in post-transcriptional gene regulation in mammals. *PLoS ONE* **10** e0124860. (<https://doi.org/10.1371/journal.pone.0124860>)
- Girard A, Sachidanandam R, Hannon GJ & Carmell MA 2006 A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442** 199–202.
- Goriaux C, Desset S, Renaud Y, Vaury C & Brasset E 2014 Transcriptional properties and splicing of the flamenco piRNA cluster. *EMBO Reports* **15** 411–418. (<https://doi.org/10.1002/embr.201337898>)
- Gou L-T, Dai P, Yang J-H, Xue Y, Hu Y-P, Zhou Y, Kang J-Y, Wang X, Li H, Hua M-M *et al.* 2014 Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Research* **24** 680–700. (<https://doi.org/10.1038/cr.2014.41>)
- Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, Martin L, Ware CB, Blish CA, Chang HY *et al.* 2015 Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **522** 221–225. (<https://doi.org/10.1038/nature14308>)
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H & Siomi MC 2007 A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* **315** 1587–1590. (<https://doi.org/10.1126/science.1140494>)
- Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J *et al.* 2014 The DNA methylation landscape of human early embryos. *Nature* **511** 606–610. (<https://doi.org/10.1038/nature13544>)
- Han BW, Hung J-H, Weng Z, Zamore PD & Ameres SL 2011 The 3'-to-5' exonuclease Nibbler shapes the 3' ends of microRNAs bound to *Drosophila* Argonaute. *Current Biology* **21** 1878–1887. (<https://doi.org/10.1016/j.cub.2011.09.034>)
- Han BW, Wang W, Li C, Weng Z & Zamore PD 2015 Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348** 817–821. (<https://doi.org/10.1126/science.aaa1264>)
- Harris AN & Macdonald PM 2001 Aubergine encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* **128** 2823–2832.
- Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL & Brennecke J 2016 Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature* **539** 588–592. (<https://doi.org/10.1038/nature20162>)
- Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL *et al.* 2017 Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nature Genetics* **49** 925–934. (<https://doi.org/10.1038/ng.3844>)
- Hickman AB & Dyda F 2015 Mechanisms of DNA transposition. In *Mobile DNA III*. Eds N Craig, M Chandler, A Lambowitz, P Rice and S Sandmeyer; ASM Press: Washington DC, USA. (<https://doi.org/10.1128/microbiolspec.MDNA3-0034-2014>)
- Hirakata S & Siomi MC 2016 piRNA biogenesis in the germline: from transcription of piRNA genomic sources to piRNA maturation. *Biochimica et Biophysica Acta* **1859** 82–92. (<https://doi.org/10.1016/j.bbagr.2015.09.002>)
- Holt JE, Roman SD, Aitken RJ & McLaughlin EA 2006 Identification and characterization of a novel Mt-retrotransposon highly represented in the female mouse germline. *Genomics* **87** 490–499. (<https://doi.org/10.1016/j.ygeno.2005.08.015>)
- Honda S, Kirino Y, Maragkakis M, Alexiou P, Ohtaki A, Murali R, Mourelatos Z & Kirino Y 2013 Mitochondrial protein BmPAPI modulates the length of mature piRNAs. *RNA* **19** 1405–1418. (<https://doi.org/10.1261/rna.040428.113>)
- Horwich MD, Li C, Matraga C, Vagin V, Farley G, Wang P & Zamore PD 2007 The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Current Biology* **17** 1265–1272. (<https://doi.org/10.1016/j.cub.2007.06.030>)
- Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB *et al.* 2007 A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129** 69–82. (<https://doi.org/10.1016/j.cell.2007.03.026>)
- Huang Y, Kim JK, Do DV, Lee C, Penfold CA, Zyllics JJ, Marioni JC, Hackett JA & Surani MA 2017 Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. *eLife* **6** e22345. (<https://doi.org/10.7554/eLife.22345>)
- Huda A, Mariño-Ramírez L & Jordan IK 2010 Epigenetic histone modifications of human transposable elements: genome defense versus exaptation. *Mobile DNA* **1** 2. (<https://doi.org/10.1186/1759-8753-1-2>)
- Inoue K, Ichihayashi K, Fukuda K, Glinka M & Sasaki H 2017 Switching of dominant retrotransposon silencing strategies from posttranscriptional to transcriptional mechanisms during male germ-cell development in mice. *PLoS Genetics* **13** e1006926. (<https://doi.org/10.1371/journal.pgen.1006926>)
- Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L & Hannon GJ 2012 The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **491** 279–283. (<https://doi.org/10.1038/nature11502>)
- Ishizu H, Siomi H & Siomi MC 2012 Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes and Development* **26** 2361–2373. (<https://doi.org/10.1101/gad.203786.112>)
- Izumi N, Kawaoka S, Yasuhara S, Suzuki Y, Sugano S, Katsuma S & Tomari Y 2013 Hsp90 facilitates accurate loading of precursor piRNAs into PIWI proteins. *RNA* **19** 896–901. (<https://doi.org/10.1261/rna.037200.112>)
- Jachowicz JW, Bing X, Pontabry J, Bošković A, Rando OJ & Torres-Padilla M-E 2017 LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nature Genetics* **49** 1502–1510. (<https://doi.org/10.1038/ng.3945>)
- Kawaoka S, Izumi N, Katsuma S & Tomari Y 2011 3' end formation of PIWI-interacting RNAs in vitro. *Molecular Cell* **43** 1015–1022. (<https://doi.org/10.1016/j.molcel.2011.07.029>)
- Kazazian HH 1998 Mobile elements and disease. *Current Opinion in Genetics and Development* **8** 343–350. ([https://doi.org/10.1016/S0959-437X\(98\)80092-0](https://doi.org/10.1016/S0959-437X(98)80092-0))
- Kazazian HH Jr, Wong C, Yousoufian H, Scott AF, Phillips DG & Antonarakis SE 1988 Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* **332** 164–166. (<https://doi.org/10.1038/332164a0>)
- Kigami D, Minami N, Takayama H & Imai H 2003 MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biology of Reproduction* **68** 651–654. (<https://doi.org/10.1095/biolreprod.102.007906>)
- Klenov MS, Sokolova OA, Yakushev EY, Stolyarenko AD, Mikhaleva EA, Lavrov SA & Gvozdev VA 2011 Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *PNAS* **108** 18760–18765. (<https://doi.org/10.1073/pnas.1106676108>)
- Klenov MS, Lavrov SA, Korbut AP, Stolyarenko AD, Yakushev EY, Reuter M, Pillai RS & Gvozdev VA 2014 Impact of nuclear Piwi elimination on chromatin state in *Drosophila melanogaster* ovaries. *Nucleic Acids Research* **42** 6208–6218. (<https://doi.org/10.1093/nar/gku268>)
- Kroutter EN, Belancio VP, Wagstaff BJ & Roy-Engel AM 2009 The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. *PLoS Genetics* **5** e1000458. (<https://doi.org/10.1371/journal.pgen.1000458>)
- Kulkarni M, Ozgur S & Stoecklin G 2010 On track with P-bodies. *Biochemical Society Transactions* **38** 242–251. (<https://doi.org/10.1042/BST0380242>)
- Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W *et al.* 2004 Mili, a mammalian

- member of piwi family gene, is essential for spermatogenesis. *Development* **131** 839–849. (<https://doi.org/10.1242/dev.00973>)
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW *et al.* 2008 DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes and Development* **22** 908–917. (<https://doi.org/10.1101/gad.1640708>)
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Takamatsu K, Chuma S, Kojima-Kita K, Shiromoto Y, Asada N, Toyoda A, Fujiyama A *et al.* 2010 MVH in piRNA processing and gene silencing of retrotransposons. *Genes and Development* **24** 887–892. (<https://doi.org/10.1101/gad.1902110>)
- Kwon C, Tak H, Rho M, Chang HR, Kim YH, Kim KT, Balch C, Lee EK & Nam S 2014 Detection of PIWI and piRNAs in the mitochondria of mammalian cancer cells. *Biochemical and Biophysical Research Communications* **446** 218–223. (<https://doi.org/10.1016/j.bbrc.2014.02.112>)
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al.* 2001 Initial sequencing and analysis of the human genome. *Nature* **409** 860–921. (<https://doi.org/10.1038/35057062>)
- Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP & Kingston RE 2006 Characterization of the piRNA complex from rat testes. *Science* **313** 363–367. (<https://doi.org/10.1126/science.1130164>)
- Lees-Murdock DJ, De Felici M & Walsh CP 2003 Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics* **82** 230–237. ([https://doi.org/10.1016/S0888-7543\(03\)00105-8](https://doi.org/10.1016/S0888-7543(03)00105-8))
- Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA & Tóth KF 2013 Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes and Development* **27** 390–399. (<https://doi.org/10.1101/gad.209841.112>)
- Leung DC & Lorincz MC 2012 Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends in Biochemical Sciences* **37** 127–133. (<https://doi.org/10.1016/j.tibs.2011.11.006>)
- Li XZ, Roy CK, Dong X, Bolcun-Filas E, Wang J, Han BW, Xu J, Moore MJ, Schimenti JC, Weng Z *et al.* 2013 An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Molecular Cell* **50** 67–81. (<https://doi.org/10.1016/j.molcel.2013.02.016>)
- Li W, Goossens K, Van Poucke M, Forier K, Braeckmans K, Van Soom A & Peelam LJ 2016 High oxygen tension increases global methylation in bovine 4-cell embryos and blastocysts but does not affect general retrotransposon expression. *Reproduction, Fertility, and Development* **28** 948–959. (<https://doi.org/10.1071/RD14133>)
- Li P, Wang L, Bennett BD, Wang J, Li J, Qin Y, Takaku M, Wade PA, Wong J & Hu G 2017 Rif1 promotes a repressive chromatin state to safeguard against endogenous retrovirus activation. *Nucleic Acids Research* **45** 12723–12738. (<https://doi.org/10.1093/nar/gkx884>)
- Lim AK & Kai T 2007 Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *PNAS* **104** 6714–6719. (<https://doi.org/10.1073/pnas.0701920104>)
- Lim AK, Lorthongpanich C, Chew TG, Tan CWC, Shue YT, Balu S, Gounko N, Kuramochi-Miyagawa S, Matzuk MM, Chuma S *et al.* 2013 The nuage mediates retrotransposon silencing in mouse primordial ovarian follicles. *Development* **140** 3819–3825. (<https://doi.org/10.1242/dev.099184>)
- Lin H & Spradling AC 1997 A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development* **124** 2463–2476.
- Luo Y-B, Zhang L, Lin Z-L, Ma J-Y, Jia J, Namgoong S & Sun Q-Y 2016 Distinct subcellular localization and potential role of LINE1-ORF1P in meiotic oocytes. *Histochemistry and Cell Biology* **145** 93–104. (<https://doi.org/10.1007/s00418-015-1369-4>)
- Maatouk DM, Kellam LD, Mann MRW, Lei H, Li E, Bartolomei MS & Resnick JL 2006 DNA methylation is a primary mechanism for silencing postmitrotic primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133** 3411–3418. (<https://doi.org/10.1242/dev.02500>)
- Maksakova IA, Romanish MT, Gagnier L, Dunn CA, Van De Lagemaat LN & Mager DL 2006 Retroviral elements and their hosts: Insertional mutagenesis in the mouse germ line. *PLoS Genetics* **2** 1–10. (<https://doi.org/10.1371/journal.pgen.0020001>)
- Malki S, van der Heijden GW, O'Donnell KA, Martin SL & Bortvin A 2014 A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Developmental Cell* **29** 521–533. (<https://doi.org/10.1016/j.devcel.2014.04.027>)
- Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R & Hannon GJ 2009 Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* **137** 522–535. (<https://doi.org/10.1016/j.cell.2009.03.040>)
- Manakov SA, Pezic D, Marinov GK, Pastor WA, Sachidanandam R & Aravin AA 2015 MIWI2 and MILI have differential effects on piRNA biogenesis and DNA methylation. *Cell Reports* **12** 1234–1243. (<https://doi.org/10.1016/j.celrep.2015.07.036>)
- Martin SL 2006 The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. *Journal of Biomedicine and Biotechnology* **2006** 45621. (<https://doi.org/10.1155/JBB/2006/45621>)
- McClintock B 1950 The origin and behavior of mutable loci in maize. *PNAS* **36** 344–355. (<https://doi.org/10.1073/pnas.36.6.344>)
- Meikar O, Vagin VV, Chalmel F, Söstar K, Lardenois A, Hammell M, Jin Y, Da Ros M, Wasik KA, Toppari J *et al.* 2014 An atlas of chromatoid body components. *RNA* **20** 483–495. (<https://doi.org/10.1261/rna.043729.113>)
- Mohn F, Sienski G, Handler D & Brennecke J 2014 The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **157** 1364–1379. (<https://doi.org/10.1016/j.cell.2014.04.031>)
- Mohn F, Handler D & Brennecke J 2015 Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* **348** 812–817. (<https://doi.org/10.1126/science.aaa1039>)
- Morgan HD, Santos F, Green K, Dean W & Reik W 2005 Epigenetic reprogramming in mammals. *Human Molecular Genetics* **14** R47–R58. (<https://doi.org/10.1093/hmg/ddi114>)
- Nandi S, Chandramohan D, Fioriti L, Melnick AM, Hébert JM, Mason CE, Rajasethupathy P & Kandel ER 2016 Roles for small noncoding RNAs in silencing of retrotransposons in the mammalian brain. *PNAS* **113** 12697–12702. (<https://doi.org/10.1073/pnas.1609287113>)
- Nishimura H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, Bonnefond L, Matsumoto N, Nishizawa T, Nakanaga K, Aoki J *et al.* 2012 Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* **491** 284–287. (<https://doi.org/10.1038/nature11509>)
- Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC & Saito K 2013 DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the *Drosophila* ovary. *Genes and Development* **27** 1656–1661. (<https://doi.org/10.1101/gad.221515.113>)
- Olivieri D, Senti K-A, Subramanian S, Sachidanandam R & Brennecke J 2012 The co-chaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in *Drosophila*. *Molecular Cell* **47** 954–969. (<https://doi.org/10.1016/j.molcel.2012.07.021>)
- Ono M, Kawakami M & Ushikubo H 1987 Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *Journal of Virology* **61** 2059–2062.
- Ostertag EM & Kazazian HH Jr 2001 Biology of mammalian L1 retrotransposons. *Annual Review of Genetics* **35** 501–538. (<https://doi.org/10.1146/annurev.genet.35.102401.091032>)
- Ostertag EM, Goodier JL, Zhang Y & Kazazian HH Jr 2003 SVA elements are nonautonomous retrotransposons that cause disease in humans. *American Journal of Human Genetics* **73** 1444–1451. (<https://doi.org/10.1086/380207>)
- Pane A, Wehr K & Schüpbach T 2007 zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. *Developmental Cell* **12** 851–862. (<https://doi.org/10.1016/j.devcel.2007.03.022>)
- Park C-E, Shin M-R, Jeon E-H, Lee S-H, Cha K-Y, Kim K, Kim N-H & Lee K-A 2004 Oocyte-selective expression of MT transposon-like element, clone MT17 and its role in oocyte maturation and embryo development. *Molecular Reproduction and Development* **69** 365–374. (<https://doi.org/10.1002/mrd.20179>)
- Patil VS & Kai T 2010 Repression of retroelements in *Drosophila* germline via piRNA pathway by the Tudor domain protein Tejas. *Current Biology* **20** 724–730. (<https://doi.org/10.1016/j.cub.2010.02.046>)
- Payer B, Saitou M, Barton SC, Thresher R, Dixon JPC, Zahn D, Colledge WH, Carlton MBL, Nakano T & Surani MA 2003 Stella is a maternal effect gene required for normal early development in mice. *Current Biology* **13** 2110–2117. (<https://doi.org/10.1016/j.cub.2003.11.026>)

- Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D & Knowles BB 2004 Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental Cell* **7** 597–606. (<https://doi.org/10.1016/j.devcel.2004.09.004>)
- Pillai RS & Chuma S 2012 piRNAs and their involvement in male germline development in mice. *Development, Growth and Differentiation* **54** 78–92. (<https://doi.org/10.1111/j.1440-169X.2011.01320.x>)
- Post C, Clark JP, Sytnikova YA, Chirn G-W & Lau NC 2014 The capacity of target silencing by Drosophila PIWI and piRNAs. *RNA* **20** 1977–1986. (<https://doi.org/10.1261/rna.046300.114>)
- Rajasethupathy P, Antonov I, Sheridan R, Frey S, Sander C, Tuschl T & Kandel ER 2012 A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* **149** 693–707. (<https://doi.org/10.1016/j.cell.2012.02.057>)
- Reuter M, Chuma S, Tanaka T, Franz T, Stark A & Pillai RS 2009 Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nature Structural and Molecular Biology* **16** 639–646. (<https://doi.org/10.1038/nsmb.1615>)
- Reuter M, Berninger P, Chuma S, Shah H, Hosokawa M, Funaya C, Antony C, Sachidanandam R & Pillai RS 2011 Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* **480** 264–267. (<https://doi.org/10.1038/nature10672>)
- Roovers EF, Rosenkranz D, Mahdipour M, Han C-T, He N, Chuva de Sousa Lopes SM, van der Westerlaken LAJ, Zischler H, Butter F, Roelen BAJ et al. 2015 Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Reports* **10** 2069–2082. (<https://doi.org/10.1016/j.celrep.2015.02.062>)
- Russell SJ, Stalker L, Gilchrist G, Backx A, Molledo G, Foster RA & LaMarre J 2016 Identification of PIWIL1 isoforms and their expression in bovine testes, oocytes, and early embryos. *Biology of Reproduction* **94** 75. (<https://doi.org/10.1095/biolreprod.115.136721>)
- Russell S, Patel M, Gilchrist G, Stalker L, Gillis D, Rosenkranz D & LaMarre J 2017 Bovine piRNA-like RNAs are associated with both transposable elements and mRNAs. *Reproduction* **153** 305–318. (<https://doi.org/10.1530/REP-16-0620>)
- Ryan FP 2004 Human endogenous retroviruses in health and disease: a symbiotic perspective. *Journal of the Royal Society of Medicine* **97** 560–565. (<https://doi.org/10.1177/014107680409701202>)
- Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H & Siomi MC 2006 Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes and Development* **20** 2214–2222. (<https://doi.org/10.1101/gad.1454806>)
- Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H & Siomi MC 2007 Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes and Development* **21** 1603–1608. (<https://doi.org/10.1101/gad.1563607>)
- Saxe JP, Chen M, Zhao H & Lin H 2013 Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO Journal* **32** 1869–1885. (<https://doi.org/10.1038/emboj.2013.121>)
- Schiff R, Itin A & Keshet E 1991 Transcriptional activation of mouse retrotransposons in vivo: specific expression in steroidogenic cells in response to trophic hormones. *Genes and Development* **5** 521–532. (<https://doi.org/10.1101/gad.5.4.521>)
- Schorn AJ, Guthrod MJ, LeBlanc C & Martienssen R 2017 LTR-retrotransposon control by tRNA-derived small RNAs. *Cell* **170** 61.e11–71.e11. (<https://doi.org/10.1016/j.cell.2017.06.013>)
- Senti K-A, Jurczak D, Sachidanandam R & Brennecke J 2015 piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes and Development* **29** 1747–1762. (<https://doi.org/10.1101/gad.267252.115>)
- Shen H 2009 UAP56-a key player with surprisingly diverse roles in pre-mRNA splicing and nuclear export. *BMB Reports* **42** 185–188. (<https://doi.org/10.5483/BMBRep.2009.42.4.185>)
- Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, Kondoh G, Okawa K, Chujo T, Suzuki T et al. 2009 The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Developmental Cell* **17** 775–787. (<https://doi.org/10.1016/j.devcel.2009.10.012>)
- Sienski G, Dönertas D & Brennecke J 2012 Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151** 964–980. (<https://doi.org/10.1016/j.cell.2012.10.040>)
- Sokol M, Jessen KM & Pedersen FS 2015 Human endogenous retroviruses sustain complex and cooperative regulation of gene-containing loci and unannotated megabase-sized regions. *Retrovirology* **12** 32. (<https://doi.org/10.1186/s12977-015-0161-9>)
- Stalker L, Russell SJ, Co C, Foster RA & LaMarre J 2016 PIWIL1 is expressed in the canine testis, increases with sexual maturity, and binds small RNAs. *Biology of Reproduction* **94** 17. (<https://doi.org/10.1095/biolreprod.115.131854>)
- Stanley D, Fraser S, Stanley GA & Chambers PJ 2010 Retrotransposon expression in ethanol-stressed *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology* **87** 1447–1454. (<https://doi.org/10.1007/s00253-010-2562-y>)
- Stein P, Rozhkov NV, Li F, Cárdenas FL, Davydenko O, Davydenko O, Vandivier LE, Gregory BD, Hannon GJ & Schultz RM 2015 Essential role for endogenous siRNAs during meiosis in mouse oocytes. *PLoS Genetics* **11** e1005013. (<https://doi.org/10.1371/journal.pgen.1005013>)
- Stoye JP 2012 Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nature Reviews: Microbiology* **10** 395–406.
- Su Y-Q, Sugiura K, Sun F, Pendola JK, Cox GA, Handel MA, Schimenti JC & Eppig JJ 2012 MARF1 regulates essential oogenic processes in mice. *Science* **335** 1496–1499. (<https://doi.org/10.1126/science.1214680>)
- Swergold GD 1990 Identification, characterization, and cell specificity of a human LINE-1 promoter. *Molecular and Cellular Biology* **10** 6718–6729. (<https://doi.org/10.1128/MCB.10.12.6718>)
- Tao Y, Yen M-R, Chitashvili T, Nakano H, Kim R, Hosohama L, Tan YC, Nakano A, Chen P-Y & Clark AT 2018 TRIM28-regulated transposon repression is required for human germline competency and not primed or naive human pluripotency. *Stem Cell Reports* **10** 243–256. (<https://doi.org/10.1016/j.stemcr.2017.11.020>)
- Töhönen V, Katayama S, Vesterlund L, Jouhilahti E-M, Sheikhi M, Madissoon E, Filippini-Cattaneo G, Jaconi M, Johnsson A, Bürglin TR et al. 2015 Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nature Communications* **6** 8207. (<https://doi.org/10.1038/ncomms9207>)
- Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V & Zamore PD 2006 A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313** 320–324. (<https://doi.org/10.1126/science.1129333>)
- van de Lagemat LN, Landry J-R, Mager DL & Medstrand P 2003 Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends in Genetics* **19** 530–536. (<https://doi.org/10.1016/j.tig.2003.08.004>)
- Virant-Klun I, Leicht S, Hughes C & Krigsveld J 2016 Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Molecular and Cellular Proteomics* **15** 2616–2627. (<https://doi.org/10.1074/mcp.M115.056887>)
- Vourekas A, Zheng Q, Alexiou P, Maragkakis M, Kirino Y, Gregory BD & Mourelatos Z 2012 Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nature Structural and Molecular Biology* **19** 773–781. (<https://doi.org/10.1038/nsmb.2347>)
- Vourekas A, Zheng K, Fu Q, Maragkakis M, Alexiou P, Ma J, Pillai RS, Mourelatos Z & Wang PJ 2015 The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. *Genes and Development* **29** 617–629. (<https://doi.org/10.1101/gad.254631.114>)
- Wang H, Ma Z, Niu K, Xiao Y, Wu X, Pan C, Zhao Y, Wang K, Zhang Y & Liu N 2016 Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in Drosophila. *Development* **143** 530–539. (<https://doi.org/10.1242/dev.128116>)
- Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T et al. 2008 Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453** 539–543. (<https://doi.org/10.1038/nature06908>)
- Watanabe T, Chuma S, Yamamoto Y, Kuramochi-Miyagawa S, Totoki Y, Toyoda A, Hoki Y, Fujiyama A, Shibata T, Sado T et al. 2011a MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. *Developmental Cell* **20** 364–375. (<https://doi.org/10.1016/j.devcel.2011.01.005>)
- Watanabe T, Tomizawa S-I, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S, Iida N, Hoki Y, Murphy PJ, Toyoda A et al. 2011b Role for piRNAs and noncoding RNA in de novo DNA methylation of the

- imprinted mouse Rasgrf1 locus. *Science* **332** 848–852. (<https://doi.org/10.1126/science.1203919>)
- Wessler SR 1996 Turned on by stress. Plant retrotransposons. *Current Biology* **6** 959–961. ([https://doi.org/10.1016/S0960-9822\(02\)00638-3](https://doi.org/10.1016/S0960-9822(02)00638-3))
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O *et al.* 2007 A unified classification system for eukaryotic transposable elements. *Nature Reviews: Genetics* **8** 973–982. (<https://doi.org/10.1038/nrg2165>)
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O *et al.* 2008 A universal classification of eukaryotic transposable elements implemented in Repbase. *Nature Reviews: Genetics* **9** 414–414. (<https://doi.org/10.1038/nrg2165-c2>)
- Wilczynska A & Bushell M 2015 The complexity of miRNA-mediated repression. *Cell Death and Differentiation* **22** 22–33. (<https://doi.org/10.1038/cdd.2014.112>)
- Williams RW & Rubin GM 2002 ARGONAUTE1 is required for efficient RNA interference in *Drosophila* embryos. *PNAS* **99** 6889–6894. (<https://doi.org/10.1073/pnas.072190799>)
- Wilson JE, Connell JE & Macdonald PM 1996 Aubergine enhances oskar translation in the *Drosophila* ovary. *Development* **122** 1631–1639.
- Xiol J, Spinelli P, Laussmann MA, Homolka D, Yang Z, Cora E, Couté Y, Conn S, Kadlec J, Sachidanandam R *et al.* 2014 RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. *Cell* **157** 1698–1711. (<https://doi.org/10.1016/j.cell.2014.05.018>)
- Yoshimura T, Watanabe T, Kuramochi-Miyagawa S, Takemoto N, Shiromoto Y, Kudo A, Kanai-Azuma M, Tashiro F, Miyazaki S, Katanaya A *et al.* 2018 Mouse GTSF1 is an essential factor for secondary piRNA biogenesis. *EMBO Reports* **19** e42054. (<https://doi.org/10.15252/embr.201642054>)
- Yuan J, Zhang P, Cui Y, Wang J, Skogerboe G, Huang D-W, Chen R & He S 2016 Computational identification of piRNA targets on mouse mRNAs. *Bioinformatics* **32** 1170–1177. (<https://doi.org/10.1093/bioinformatics/btv729>)
- Zamudio N, Barau J, Teissandier A, Walter M, Borsos M, Servant N & Bourc'his D 2015 DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes and Development* **29** 1256–1270. (<https://doi.org/10.1101/gad.257840.114>)
- Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma S, Weng Z, Theurkauf WE & Zamore PD 2011 Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. *Molecular Cell* **44** 572–584. (<https://doi.org/10.1016/j.molcel.2011.10.011>)
- Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD *et al.* 2012 UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151** 871–884. (<https://doi.org/10.1016/j.cell.2012.09.040>)
- Zhang P, Kang J-Y, Gou L-T, Wang J, Xue Y, Skogerboe G, Dai P, Huang D-W, Chen R, Fu X-D *et al.* 2015 MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Research* **25** 193–207. (<https://doi.org/10.1038/cr.2015.4>)
- Zhang H, Liu J, Tai Y, Zhang X, Zhang J, Liu S, Lv J, Liu Z & Kong Q 2017 Identification and characterization of L1-specific endo-siRNAs essential for early embryonic development in pig. *Oncotarget* **8** 23167–23176. (<https://doi.org/10.18632/oncotarget.15517>)

Received 27 April 2018

First decision 18 June 2018

Revised manuscript received 16 July 2018

Accepted 19 July 2018