As siRNAs derive from such long double-stranded RNAs, one could expect siRNA precursors to constitutively activate the interferon response… except that the whole pathway seems to be inactive in mouse oocytes [15]. Identifying the mouse tissues devoid of an interferon response could prove helpful in the search for endogenous siRNA function.

It can be expected, at least, that one particular feature of siRNAs will not make that task easy: endogenous siRNAs are expressed at very low levels — much lower than typical miRNAs. This may be linked to the activity of siRNAs: as guides for the endonucleolytic cleavage of their targets, they act catalytically, whereas most animal miRNAs act stoichiometrically (the target needs to be bound by the miRNA to be repressed, until it is degraded).

These observations suggest an impressive, almost scary, possibility. After the miRNA era (where so many functions have been ascribed to miRNAs, in so many physiological processes), we may be entering the 'siRNA era'. How many biological pathways will involve siRNAs? As siRNAs can act catalytically, minute amounts of these novel regulators could have tremendous effects; undetected small RNAs may lie behind unexplained phenomena. Clearly, the exploration of small regulatory RNAs is not over.

References

1. 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.
2. Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., *et al*. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534–538.
3. Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev. *22*, 2773–2785.
4. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853–858.
5. Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. Nat. Rev. Mol. Cell. Biol. *9*, 219–230.
6. Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nat. Genet. *35*, 215–217.
7. Suh, N., Baehner, L., Melton, C., Shenoy, A., Chen, J., and Blelloch, R. (2010). MicroRNA function is globally suppressed during mouse oocyte maturation and preimplantation development. Curr. Biol. *20*, 271–277.
8. Ma, J., Flemr, M., Stein, P., Berninger, P., Malik, R., Zavolan, M., Svoboda, P., and Schultz, R.M. (2010). Loss of microRNA function during oocyte growth in mouse. Curr. Biol. *20*, 265–270.
9. Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. Genes Dev. *21*, 682–693.
10. Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S.C., Sun, Y.A., Lee, C., Tarakhovsky, A., Lao, K., and Surani, M.A. (2007). Maternal microRNAs are essential for mouse zygotic development. Genes Dev. *21*, 644–648.
11. Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. PLoS Biol. *4*, e210.
12. Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol. Cell. Biol. *27*, 3970–3981.
13. Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van, Wolfswinkel, J.C., Chaves, D.A., Shirayama, M., Mitani, S., Ketting, R.F., *et al*. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell *139*, 123–134.
14. Bushati, N., and Cohen, S.M. (2007). microRNA functions. Annu. Rev. Cell. Dev. Biol. *23*, 175–205.
15. Stein, P., Zeng, F., Pan, H., and Schultz, R.M. (2005). Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. Dev. Biol. *286*, 464–471.

Université de Toulouse, UPS, Laboratoire de Biologie Moléculaire Eucaryote, F-31000, France, and CNRS, Laboratoire de Biologie Moléculaire Eucaryote, F-31000 Toulouse, France.
E-mail: seitz@ibcg.biotoul.fr

DOI: 10.1016/j.cub.2009.12.027

Short RNAs: How Big Is This Iceberg?

Recent studies have reported the identification of piwi-associated RNAs (piRNAs) in *Drosophila* somatic cells. Interestingly, these piRNAs derive from the 3′ untranslated regions of a subset of transcribed protein-coding genes and, experimentation suggests, might control the expression of other protein-coding transcripts. Studies of additional organisms support the new pathway's presence across animals.

Isidore Rigoutsos

The RNA revolution is upon us. The last decade and a half has witnessed a flurry of research activity revolving around RNA, a molecule that for many years was thought to play only ancillary roles in the workings of a cell. Aided by technological and scientific advances, the repertoire of non-coding RNAs, i.e. RNAs that do not code for proteins, has been expanding steadily [1].

A particular class of short RNAs, microRNAs, has so far commanded the lion's share of researchers' attention. Approximately 22 nucleotides long, microRNAs have been shown to post-transcriptionally regulate their targets in a sequence-dependent manner [2]. Originally believed to act through the 3′ untranslated region (3′UTR) of the targeted mRNAs, microRNAs were recently shown to target the mRNA's amino-acid coding sequence (CDS) as well [3,4]. Recently, piwi-associated RNAs (piRNAs), another intriguing class of short RNAs with lengths between 23 and 30 nucleotides, burst onto the stage [5–8]. To date, piRNAs have been reported in *Drosophila*, human, mouse, rat, *Xenopus*, zebrafish and the worm, and their known numbers are in the hundreds of thousands.

Drosophila piRNAs deriving from heterochromatic loci that are rich in nested, truncated, or damaged repeat elements, termed piRNA 'clusters', have been among the better studied. Such piRNAs participate in the 'ping-pong' cycle, an auto-amplification loop that is conserved in many metazoans and constitutes an adaptive immune response that maintains genomic integrity by suppressing the expression of transposons and other repetitive sequences [9,10].

In addition to piRNAs generated via the ping-pong cycle, other piRNAs that do not depend on auto-amplification have also been reported. The specifics of their generation had remained elusive for a while, but recent studies by the laboratories of Mikiko and

Haruhiko Siomi, and Eric Lai, have broken new ground, revealing surprising results with potentially very important ramifications [11–13].

A common element in the work by both teams is their use of a *Drosophila* ovarian somatic cell line (OSS), a homogeneous and stable cell line derived from a parental cell line (fGS/OSS) comprising a mix of female germline stem cells and sheets of somatic cells. This was a very important design choice in that it removed one confounding factor of previous piRNA studies, namely the use of systems comprising heterogeneous cell types.

A combination of qRT-PCR and Western blot analysis by the two teams established the presence of Piwi in OSS, and the absence of both Aubergine (Aub) and Argonaute 3 (AGO3), the latter two being required for the production of piRNAs through the ping-pong cycle. Immunostaining of OSS revealed Piwi's nuclear localization but left open the question as to whether such localization was required for the production and loading of piRNAs. The Siomi team [13] designed and studied three mutants: a Piwi mutant lacking its nuclear localization signal expectedly localized to the cytoplasm but could still load piRNAs, just like the wild type; a second Piwi mutant that was deprived of its Slicing faculty through mutations in its PIWI domain also maintained the ability to load piRNAs; and finally, a double mutant that was Slicer-deficient and also lacked its nuclear localization signal continued to load piRNAs even when endogenous Piwi was depleted. These findings suggest that the process at hand, unlike the ping-pong cycle, does not depend on Piwi's Slicing ability and that piRNA production and loading on Piwi occur in the cytoplasm. In parallel work, the Lai team [12] analyzed newly generated and previously published deep sequencing data and found no evidence of the sequence biases that are the hallmarks of the ping-pong cycle, namely a preponderance of uridine in the first piRNA position and adenine in the tenth, and the existence of piRNA pairs containing complementary sequences in the first ten positions from their 5′ end. Combined, these results indicate that the piRNAs present in the OSS cells are generated independently of Aub/AGO3 and the auto-amplification loop. But where do these piRNAs come from?

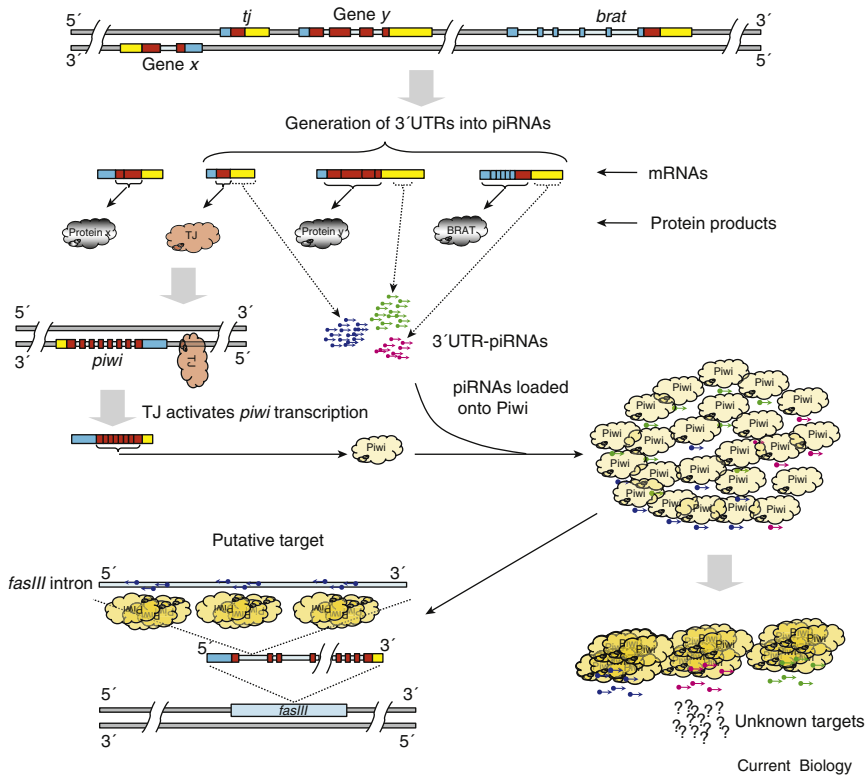

Figure 1. Summary depiction of the findings described in [11–13].

Among the transcribed genes, a subset (e.g. *tj*, *brat*, hypothetical gene *y* and others) gives rise to piRNAs through the 3′UTRs of their mRNAs. Abundant transcription does not appear to be a prerequisite, and not every transcribed gene's 3′UTR will generate piRNAs (e.g., hypothetical gene *x*). The generation of these piRNAs that are sense to the source 3′UTR occurs in parallel with the translation of the mRNA. Among the piRNA-producing genes in the OSS cell line, *tj* and *brat* are the two most prolific ones. Once made, the TJ protein appears to facilitate/initiate transcription of *piwi* and the production of Piwi. The generated piRNAs are subsequently loaded on Piwi in the cytoplasm. Preliminary *indirect* evidence suggests that the complex carrying some of the *tj*-piRNAs may target the primary *fasIII* transcript within its first intron. Given the high number of the generated distinct piRNAs, it is reasonable to assume that there is an abundance of targets that remain to be elucidated. Color convention: blue, 5′UTR; red, CDS; yellow, 3′UTR; cyan, intronic sequence; gray, intergenic sequence.

Mapping of the piRNAs on the *Drosophila* genome by both teams revealed the *flamenco* locus as one of the sources [12,13]. In addition to *flamenco*, and rather unexpectedly, the two teams discovered that many of the sequenced piRNAs were *sense* to, and derived from, the 3′UTRs of *Drosophila* mRNAs (Figure 1). The deeper sequencing carried out by Lai's team [11], as reported in a recent issue of *Current Biology*, allowed them to identify many tens of genes generating greater than 1,000 piRNAs from their 3′UTRs. In fact, these mRNA-derived piRNAs were comparable in abundance to microRNAs in OSS. It is worth emphasizing here that despite using different platforms, sequencing at different depths, and employing different thresholds, the two teams' lists of top-ranking mRNAs, vis-à-vis the piRNAs that the mRNAs generated, were in agreement. The highest-ranking mRNA in both lists corresponded to *traffic jam* (*tj*), a gene encoding a Maf-bZIP transcription factor that is necessary for controlling gonad morphogenesis in the fruit fly. Another high-ranking piRNA-generating mRNA was *brat*.

Given that *tj*'s mRNA apparently serves two purposes, namely protein synthesis and piRNA production, is the mRNA pre-processed in any way prior to entering the primary piRNA pathway? The answer appears to be negative. Indeed, Northern blot analysis by the Siomi team [13] using probes targeting, respectively, *tj*'s CDS and 3′UTR revealed a single band of the same length, suggesting no

separation of the primary mRNA into segments.

Further analysis of data from recent studies of *Drosophila* mutants [14,15], and of newly generated data [11] in conjunction with RNAi studies [13], allowed the two teams to establish that *zucchini* is involved in the 3′UTR-piRNA pathway. Also, study of larval ovaries of *tj* mutants by the Siomi team [13] showed TJ to be required for *piwi* activation in *Drosophila* gonadal somatic cells. Notably, there are several potential Maf binding sites near *piwi*'s putative transcription start site: for some of them, chromatin immunoprecipitation experiments indicated a direct association with TJ (Figure 1).

Once produced and loaded on Piwi, the *tj*-derived piRNAs have the ability to suppress targets. By searching for targets with (near) complementarity to *tj*'s 3′UTR piRNAs, the Siomi team [13] was able to identify *fasIII* as one such candidate target gene. Interestingly, all three putative sites of interaction with the sequenced piRNAs were in *fasIII* introns (Figure 1). Quantitative RT-PCR analysis showed that *fasIII* is upregulated in *piwi*- and *tj*-mutants, supporting the conjecture that *fasIII* is a piRNA target.

However, are the above results a *tj*-specific or, perhaps, a *Drosophila*-specific peculiarity? A wide-ranging investigation by Lai's team [11] provides the answer. Analysis of public data and additional primary data that they generated revealed strong 3′UTR piRNA production in mouse testis as well. These mouse piRNAs were proportionally highest in pre-pachytene testis (10 days postpartum) wherein Mili was found to be the predominant carrier of 3′UTR-piRNAs, in agreement with earlier observations [16]. Notably, despite the difference in the carrier molecule, the identities of the gene sources of the Mili- and Miwi-loaded 3′UTR-piRNAs were relatively similar. Further analysis also allowed Lai's team to reclassify previously published piRNA clusters as 3′UTR-derived ones. The team also analyzed piRNAs from *X. tropicalis* eggs that associated with the Y12 monoclonal antibody and identified numerous 3′UTR-piRNAs there as well, albeit they were not as abundant as in the mouse and in *Drosophila*. The findings of the Lai group [11] together with corroborating results by the Siomi team [13] (obtained by studying *piwi* expression in mouse *tj*-mutant testis) suggest conservation of the identified pathway across animals.

Do all 3′UTRs give rise to piRNAs? Interestingly, by studying sequenced reads in conjunction with gene expression data in OSS, Lai's team showed that high-expression is not a sufficient condition for the generation of piRNAs. Instead, they found preference for mRNA substrates belonging to select gene ontology (GO) categories such as post-transcriptional gene regulation, regulation of structure and organ development, morphogenesis, cell differentiation, and others. A similar analysis of mouse testis data revealed analogous GO biases.

These above results provide evidence in support of a previously reported, computationally derived framework that linked intergenic/intronic regions with exons through short motifs termed pyknons [17]. Pyknons are organism-specific, variable-length, statistically significant nucleotide motifs that were arrived at using purely computational means. By definition, the pyknons of a given genome occur multiple times in intergenic and/or intronic regions, as well as at least once in an exon. Despite a pyknon's multiple occurrences in a genome, not all instances of the motif necessarily coincide with retroelements [17]. They were originally a genomic curiosity [18] but after piRNAs identified by deep sequencing became available [5–8], it was apparent that pyknons had computationally presaged the existence of piRNAs, both in terms of length and sequence. At the same time pyknons predicted putative regulatory connections between exonic and non-exonic sequences, similar to that reported by the Siomi group [13]. Notably, several of the GO categories that were found enriched in the original pyknon analysis [17] agree with those identified by Lai's team [11] as associated with 3′UTR-piRNAs. Subsequent analysis showed that human and mouse pyknons are over-represented in the introns of genes belonging to specific GO categories, and that the categories are the same in the two genomes despite the absence of conservation of the underlying intronic sequences [19]. Additionally, several examples of pyknons and published piRNAs that are *antisense* to introns have been reported [19], underscoring the connection between pyknons and piRNAs and in accordance with the putative *fasIII* findings by Siomi's team [13].

The pyknon framework anticipates that each mRNA region (5′UTR, CDS, 3′UTR) is associated (as source, target, or perhaps both) to distinguishable-by-sequence collections of short RNAs with each region being linked to distinct GO categories: glimpses of this are evident in the results generated by Lai's team [11]. If we now combine this multitude of connections with a previous report on the differential expression of a human pyknon containing genomic region in a disease context [20], it becomes increasingly likely that these findings by the Siomi and Lai teams [11–13] represent further evidence that the community may have stumbled upon the tip of a regulatory iceberg of potentially staggering proportions.

References

1. Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. Hum Mol. Genet. *15* Spec No 1, R17–R29.
2. Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215–233.
3. Rigoutsos, I. (2009). New tricks for animal microRNAS: targeting of amino acid coding regions at conserved and nonconserved sites. Cancer Res. *69*, 3245–3248.
4. Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature *460*, 479–486.
5. Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363–367.
6. Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.
7. Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., 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.
8. Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709–1714.
9. Tushir, J.S., Zamore, P.D., and Zhang, Z. (2009). SnapShot: Fly piRNAs, PIWI proteins, and the ping-pong cycle. Cell *139*, 634–634. e1.
10. Tushir, J.S., Zamore, P.D., and Zhang, Z. (2009). SnapShot: mouse piRNAs, PIWI proteins, and the ping-pong cycle. Cell *139*, 830–830. e1.
11. Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3′UTR-directed primary piRNAs. Curr. Biol. *19*, 2066–2076.
12. Lau, N.C., Robine, N., Martin, R., Chung, W.J., Niki, Y., Berezikov, E., and Lai, E.C. (2009). Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. Genome Res. *19*, 1776–1785.

13. Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature *461*, 1296–1299.
14. Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., *et al.* (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.
15. Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522–535.
16. Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell *31*, 785–799.
17. Rigoutsos, I., Huynh, T., Miranda, K., Tsirigos, A., McHardy, A., and Platt, D. (2006). Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. Proc. Natl. Acad. Sci. USA *103*, 6605–6610.
18. Meynert, A., and Birney, E. (2006). Picking pyknons out of the human genome. Cell *125*, 836–838.
19. Tsirigos, A., and Rigoutsos, I. (2008). Human and mouse introns are linked to the same processes and functions through each genome's most frequent non-conserved motifs. Nucleic Acids Res. *36*, 3484–3493.
20. Calin, G.A., Liu, C.G., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E.J., Wojcik, S.E., *et al.* (2007). Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell *12*, 215–229.

Bioinformatics and pattern Discovery Group, Computation Biology Center, IBM Thomas J. Watson Research Center, PO BOX 218, Yonktown Heights, NY 10598, USA.
E-mail: rigoutso@us.ibm.com

DOI: 10.1016/j.cub.2009.12.036