Published online 29 October 2018 *Nucleic Acids Research, 2019, Vol. 47, Database issue* ***D175–D180***
doi: 10.1093/nar/gky1043

piRBase: a comprehensive database of piRNA sequences

Jiajia Wang[1,†], Peng Zhang[1,†], Yiping Lu[2,†], Yanyan Li[1,3], Yu Zheng[1,3], Yunchao Kan[4], Runsheng Chen[1,*] and Shunmin He[1,*]

[1]Key Laboratory of RNA Biology, Center for Big Data Research in Health, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, [2]School of Life Sciences, Zhengzhou University, Zhengzhou 450001,China, [3]University of Chinese Academy of Science, Beijing 100049, China and [4]China-UK-NYNU-RRes Joint Laboratory of insect biology, Henan Key Laboratory of Insect Biology in Funiu Mountain, Nanyang Normal University, Nanyang, Henan 473061,China

Received August 15, 2018; Revised September 23, 2018; Editorial Decision October 12, 2018; Accepted October 22, 2018

ABSTRACT

PIWI-interacting RNAs are a class of small RNAs that is most abundantly expressed in animal germline. Substantial research is going on to reveal the functions of piRNAs in the epigenetic and post-transcriptional regulation of transposons and genes. To collect and annotate these data, we developed piRBase, a database assisting piRNA functional study. Since its launch in 2014, piRBase has integrated 264 data sets from 21 organisms, and the number of collected piRNAs has reached 173 million. The latest piRBase release (v2.0, 2018) was more focused on the comprehensive annotation of piRNA sequences, as well as the increasing number of piRNAs. In addition, piRBase release v2.0 also contained the potential information of piRNA targets and disease related piRNA. All datasets in piRBase is free to access, and available for browse, search and bulk downloads at http://www.regulatoryrna.org/database/piRNA/.

INTRODUCTION

Up to now, various small RNAs can be classified into three types: microRNA (miRNA), endogenous small interfering RNA (siRNA), and PIWI-interacting RNA (piRNA), based on their biogenesis and associated proteins (1). piRNAs are mainly expressed in mammalian germline (2–5), and are generally 24–31 nucleotides in length with 2′-*O*-methylation at their 3′ ends in most animals (6–8). Compared with miRNAs and siRNAs, piRNAs are much more diverse in sequences. PIWI proteins, which are associated with piRNAs, are also enriched in germline (9–11).

piRNAs play various functions in germline and somatic tissues. PIWI/piRNA complex can repress the activity of transposon, which has a high risk of damaging the genome, through post-transcriptional silencing or heterochromatin formation (6,12–20). However, many piRNAs do not match transposon sequences, which implies additional functions of piRNAs (21–23). Recent studies have found that piRNAs can target mRNA by base pairing in mouse (24,25), *Drosophila* (26,27) and *Caenorhabditis elegans* (28–32). In silkworm (33) and *C. elegans* (31), piRNAs are reported to be involved in sex determination through down-regulation of target genes. Besides transposon and mRNA, a large number of lncRNAs are revealed to be mediated by retrotransposon derived piRNAs in mouse late spermatocytes (34).

As piRNAs are implicated in transposon and gene regulation, there has been a budding interest in defining the role of piRNAs in human disease (35,36). piRNAs are garnering more and more attention in a variety of cancers. For example, 106 piRNAs are up-regulated and 91 are down-regulated in bladder cancer (37). The abnormal expression of piRNAs are also demonstrated in other types of cancers such as breast cancer (38,39) and gastric cancer (36,40,41).

In light of the rapidly increasing studies on piRNA, several piRNA related databases, such as piRNABank (42), piRNAQuest (43), piRNA cluster database (44) and IsopiRBank (45), are generated. However, these databases only contain limited amounts of piRNAs from limited species, and the information of piRNA functions are rarely included. piRBase is the expert database about piRNA included by RNAcentral (46) and is the first database that

*To whom correspondence should be addressed. Tel: +86 10 6488 7032; Fax: +86 10 6488 7842; Email: heshunmin@ibp.ac.cn
Correspondence may also be addressed to Runsheng Chen. Tel: +86 10 6488 8543; Fax: +86 10 6487 1293; Email: crs@ibp.ac.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

systematically integrates various piRNA associated data to support piRNA functional analysis.

In the latest piRBase release, the number of unique piRNA sequences increased to 173 million, including 21 species. In addition, the piRNA target mRNA records were expanded and potential piRNA target lncRNAs were added in piRBase release v2.0. The information about eight types of cancers (breast, bladder, pancreas, gastric, liver, kidney, myeloma and colorectal cancer) related piRNAs was also added to the new version. We also provided new web tools and improved user interface in piRBase release v2.0.

DATA COLLECTION AND PROCESSING

Based on the first release of piRBase database (47), new datasets and processed piRNA sequences from the literature, supplementary files, and NCBI GEO were collected. We only included sequences that were taken as piRNA or piRNA candidates in corresponding papers. Other piRNA related information was manually extracted from the relevant articles. The contents in piRBase release v2.0 are shown in Figure 1.

Each piRNA was shown with a piRBase name and other related information such as NCBI accession number, organism of origin, sequence, length, the number of papers reporting the piRNA and the number of different methods supporting the piRNA. In the detailed information page, we also provided additional information about the piRNA, such as aliases (NCBI and RNAdb piRNA name), related datasets with tissue and method, genome location, targets and associated cancer. In piRBase, if a piRNA sequence is a subsequence of another piRNA, both of them were considered as different sequences and were assigned distinct piRBase names.

As described in the first release of piRBase, all piRNA sequences were mapped to its latest genome using Bowtie (48) with parameter '-v 1 -a –best –strata' in order to obtain the potential origin of every piRNA. piRNAs were referred to as gene- or repeat-related according to the overlapping of piRNA genome loci with RefSeq genes (49) or repeat elements.

DATABASE CONTENT

piRBase aims to provide comprehensive piRNA sequence data, annotation and targets assisting piRNA study. The first release in 2014 contains 77 million piRNA sequences from nine organisms. piRNAs in piRBase are marked as repeat- and gene-related according to their loci in genome, which imply involvement in the regulation of the corresponding elements. In additional, reported piRNA targets and piRNA related epigenetic data were collected. The information of piRNA loci, piRNA target sites, piRNA related epigenetic data, and some basic annotations like RepeatMasker annotations (50) and RefSeq genes are visualized in the UCSC genome browser (50,51) in piRBase release v1.0. Since piRBase release in 2014, we collected more piRNA related data and released in v2.0.

New piRNA records

In present release of piRBase, the number of unique piRNA sequences reaches 173 million, including 264 datasets from 21 species. The number is greatly increased in this version compared with the first release (47) and other piRNA related databases (Table 1). In order to assess credibility of piRNAs in piRBase, we also offered two evaluation score. One score is the number of papers which detected certain piRNA, and the other score is the number of different methods used to detect the sequence. These methods including protein IP, protein CLIP, chromatography, oxidization and small RNA-Seq.

piRNA targets

Many studies have found that piRNAs can regulate the expression of mRNAs and lncRNAs. In the previous version, mRNAs regulated by piRNAs in fly and mouse were included. In this version, new mRNA targets (for silkworm and *C. elegans*) extracted from the literature were added. Moreover, we predicted piRNA target lncRNAs using the same way as predicting piRNA target mRNA cleavage (25) in mouse testis. In the web page of piRNA target table, we provided the piRBase name linking to detailed information of piRNA, target gene/transcript ID, target site if provided, and other related information.

piRNA and cancer

From piRNA and cancer related literature, we manually collected cancer related piRNAs. In the piRBase release v2.0, eight types of cancers (breast, bladder, pancreas, gastric, liver, kidney, myeloma and colorectal cancer) related piRNAs were included with the information of piRNA name, piRNA expression (up-regulated, down-regulated and fold change) and functional description in cancer tissues or cell lines.

New tools

piRBase release v2.0 added 'Tools' menu in navigation bar. We incorporated pirnaPre (52) for users to predict piRNA targets. In addition, some online tools were integrated under the 'Tools' menu. 'Name Convert' enabled users to convert a list of NCBI or RNAdb piRNA names and accession numbers to piRBase names. The users can also retrieve piRBase names by a list of piRNA sequences using 'Seq To Name'. Other existing tools such as 'Reverse Complement' and 'Bowtie Online' tool are also incorporated in the 'Tools' menu. Beyond that, we also included the links of others tools for predicting piRNA-targeting sites or identifying piRNAs, such as pirScan (53) and 2L-piRNA (54).

Visualization using UCSC genome browser

In the current piRBase, the latest genome versions such as hg38 and mm10 were added in the UCSC genome browser. Many of the information in the present piRBase release

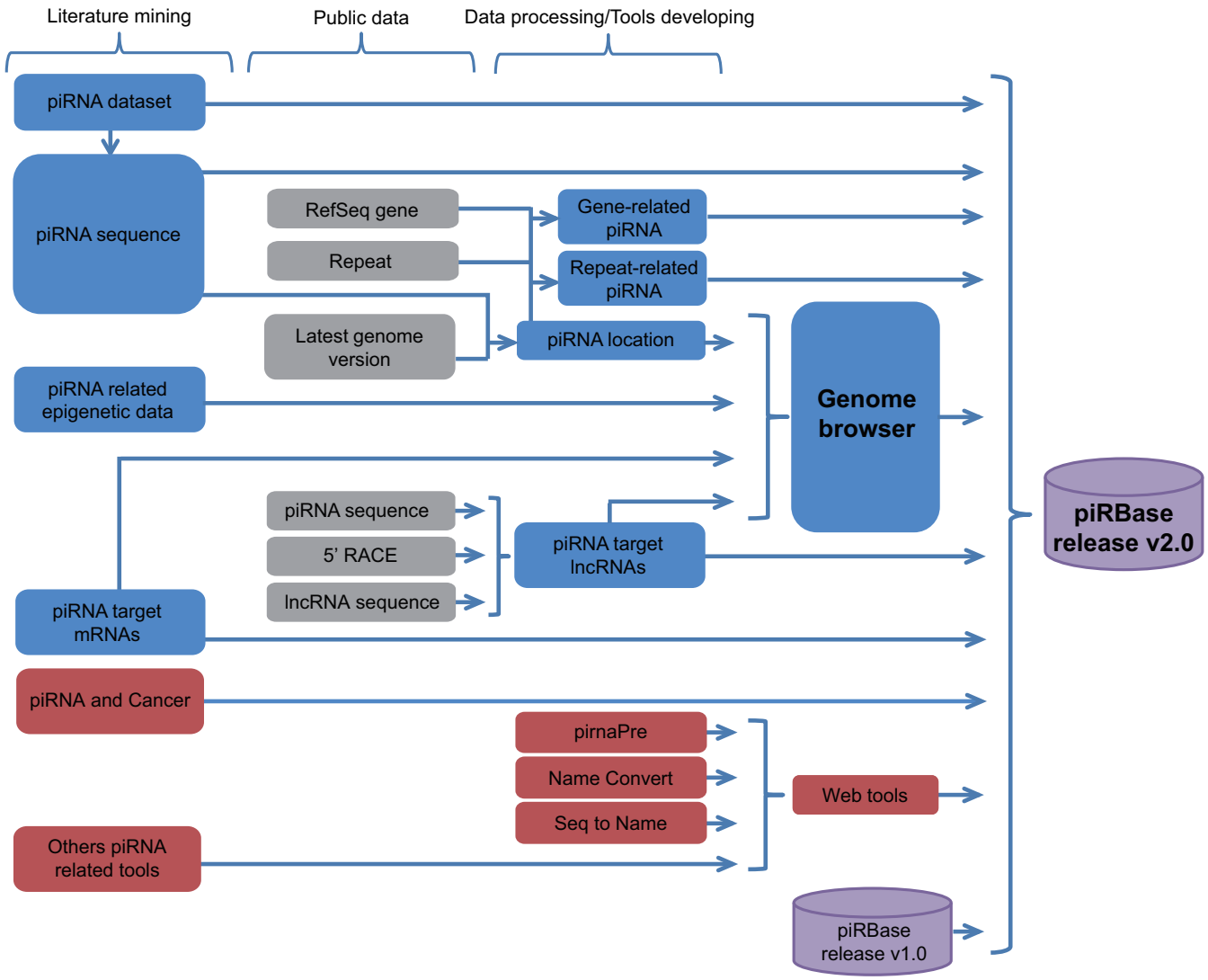

Figure 1. The contents in piRBase release v2.0. piRNA loci, piRNA target sites (mRNAs and lncRNAs) and piRNA related epigenetic data are visualized by UCSC genome browser. Blue boxes: the main updated modules; Red boxes: new modules.

v2.0, such as piRNA loci, piRNA target sites (mRNAs and lncRNAs) and piRNA related epigenetic data, have be visualized using UCSC genome browser. Moreover, we have added about 400 tracks of all related data from current version of piRBase in UCSC genome browser and users can choose to display content they are interested in.

Interface improvement

The web interface of the present piRBase database has been improved for better browse and search experience. The piRNA targets and related cancers are added in piRNA detailed information page. In the piRNA target page, users can browse and search the information by selecting organism, gene symbol and transcript ID. In piRNA search page, two more search options are added in piRNA sequence search. 'SubStr' can search all sub-sequences of input sequence. 'Extend 3′ can search 3′ extended piRNAs.

On the basis of download page in the first release, new version data were included. We provided all piRBase data in both release v1.0 and v2.0 for downloading. In addition, users can also download the related data through the 'Export' button in corresponding pages.

CONCLUSION

piRBase is the first database that systematically integrates various types of data to support piRNA functional analysis. Compared with the first release in 2014, piRBase release v2.0 is more comprehensive. The total numbers of unique piRNAs and species have been increased. Besides piRNA target mRNAs, the current piRBase also included target lncRNAs. In light of piRNAs as potential biomarkers in cancers, the information of piRNA and cancer was added to the current release. More piRNA related epigenetic data were also added in piRBase UCSC genome browser. In addition, piRBase release v2.0 added new web tools and improved the user interface. As piRNA studies expanded rapidly, we will update piRBase with more related information supporting piRNA functional analysis.

Table 1. The comparison of piRBase and other piRNA databases

				piRBase	
Species	piRNABank	piRNAQuest	IsopiRBank[c]	Release v1.0	Release v2.0
Human	35 356	41 749	4 564 080	32 826	8 438 265
Mouse	55 359	890 078	38 093 003	51 664 769	68 054 594
Rat	46 444	66 758		63 182	4 081 625
D. melanogaster	44 417[a]		15 058 093	21 027 419	41 950 613
C. elegans				28 219	28 219
Zebrafish			6 367 811	1 330 692	1 330 692
Chicken				508 437	1 716 795
X. tropicalis				6 142 904	6 142 904
Silkworm				1 174 963	1 253 987
Starlet sea anemone					11 811
Cow					19 528 252
Crab-eating macaque					5 819 946
Rhesus					6 514
Marmoset					8 311 210
Sea hare					499
Tree shrew					65 810
Pig					1 301 866
D.erecta					1 091 318
D.yakuba					1 069 918
D.virilis					907 289
Rabbit					2 816 864
Platypus	51[b]				

[a]The number of unique piRNAs may be less than shown.
[b]The number of unique piRNAs comes from statistic of download file.
[c]The number of canonical piRNAs is shown.

ACKNOWLEDGEMENTS

Data analysis and computing resource was supported by Center for Big Data Research in Health (http://bigdata.ibp.ac.cn), Institute of Biophysics, Chinese Academy of Sciences.

FUNDING

National Key R&D Program of China [2016YFC0901702]; National Natural Science Foundation of China [31871294]. Funding for open access charge: National Key R&D Program of China [2016YFC0901702].
Conflict of interest statement. None declared.

REFERENCES

1. Kim,V.N., Han,J. and Siomi,M.C. (2009) Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.*, **10**, 126–139.
2. 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.
3. 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.
4. 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.
5. 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.
6. Vagin,V.V., Sigova,A., Li,C., Seitz,H., Gvozdev,V. and Zamore,P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, **313**, 320–324.
7. Kirino,Y. and Mourelatos,Z. (2007) Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. *Nat. Struct. Mol. Biol.*, **14**, 347–348.
8. Ohara,T., Sakaguchi,Y., Suzuki,T., Ueda,H., Miyauchi,K. and Suzuki,T. (2007) The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. *Nat. Struct. Mol. Biol.*, **14**, 349–350.

9. Deng,W. and Lin,H. (2002) miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell*, **2**, 819–830.
10. Kuramochi-Miyagawa,S., Kimura,T., Ijiri,T.W., 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.
11. Carmell,M.A., Girard,A., van de Kant,H.J., Bourc'his,D., Bestor,T.H., de Rooij,D.G. and Hannon,G.J. (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell*, **12**, 503–514.
12. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089–1103.
13. Gunawardane,L.S., Saito,K., Nishida,K.M., Miyoshi,K., Kawamura,Y., Nagami,T., Siomi,H. and Siomi,M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. *Science*, **315**, 1587–1590.
14. Houwing,S., Kamminga,L.M., Berezikov,E., Cronembold,D., Girard,A., van den Elst,H., Filippov,D.V., Blaser,H., Raz,E., Moens,C.B. *et al.* (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell*, **129**, 69–82.
15. Carmell,M.A., Xuan,Z., Zhang,M.Q. and Hannon,G.J. (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.*, **16**, 2733–2742.
16. 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.
17. Saito,K., Nishida,K.M., Mori,T., Kawamura,Y., Miyoshi,K., Nagami,T., Siomi,H. and Siomi,M.C. (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev.*, **20**, 2214–2222.
18. Yin,H. and Lin,H. (2007) An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. *Nature*, **450**, 304–308.
19. 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.
20. Lim,A.K., Tao,L. and Kai,T. (2009) piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline. *J. Cell Biol.*, **186**, 333–342.
21. Batista,P.J., Ruby,J.G., Claycomb,J.M., Chiang,R., Fahlgren,N., Kasschau,K.D., Chaves,D.A., Gu,W., Vasale,J.J., Duan,S. *et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. *Mol. Cell*, **31**, 67–78.
22. Das,P.P., Bagijn,M.P., Goldstein,L.D., Woolford,J.R., Lehrbach,N.J., Sapetschnig,A., Buhecha,H.R., Gilchrist,M.J., Howe,K.L., Stark,R. *et al.* (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. *Mol. Cell*, **31**, 79–90.
23. Aravin,A.A., Hannon,G.J. and Brennecke,J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, **318**, 761–764.
24. 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 Res.*, **24**, 680–700.
25. 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 Res.*, **25**, 193–207.
26. Rouget,C., Papin,C., Boureux,A., Meunier,A.C., Franco,B., Robine,N., Lai,E.C., Pelisson,A. and Simonelig,M. (2010) Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. *Nature*, **467**, 1128–1132.
27. Vourekas,A., Alexiou,P., Vrettos,N., Maragkakis,M. and Mourelatos,Z. (2016) Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*, **531**, 390–394.
28. Lee,H.C., Gu,W., Shirayama,M., Youngman,E., Conte,D. Jr. and Mello,C.C. (2012) C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell*, **150**, 78–87.
29. Bagijn,M.P., Goldstein,L.D., Sapetschnig,A., Weick,E.M., Bouasker,S., Lehrbach,N.J., Simard,M.J. and Miska,E.A. (2012) Function, targets, and evolution of Caenorhabditis elegans piRNAs. *Science*, **337**, 574–578.
30. Shen,E.Z., Chen,H., Ozturk,A.R., Tu,S., Shirayama,M., Tang,W., Ding,Y.H., Dai,S.Y., Weng,Z. and Mello,C.C. (2018) Identification of piRNA binding sites reveals the argonaute regulatory landscape of the C. elegans Germline. *Cell*, **172**, 937–951.
31. Tang,W., Seth,M., Tu,S., Shen,E.Z., Li,Q., Shirayama,M., Weng,Z. and Mello,C.C. (2018) A sex chromosome piRNA promotes robust dosage compensation and sex determination in C. elegans. *Dev. Cell*, **44**, 762–770.
32. Zhang,D., Tu,S., Stubna,M., Wu,W.S., Huang,W.C., Weng,Z. and Lee,H.C. (2018) The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science*, **359**, 587–592.
33. Kiuchi,T., Koga,H., Kawamoto,M., Shoji,K., Sakai,H., Arai,Y., Ishihara,G., Kawaoka,S., Sugano,S., Shimada,T. *et al.* (2014) A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*, **509**, 633–636.
34. Watanabe,T., Cheng,E.C., Zhong,M. and Lin,H. (2015) Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res.*, **25**, 368–380.
35. Lu,Y., Li,C., Zhang,K., Sun,H., Tao,D., Liu,Y., Zhang,S. and Ma,Y. (2010) Identification of piRNAs in Hela cells by massive parallel sequencing. *BMB Rep.*, **43**, 635–641.
36. Cheng,J., Guo,J.M., Xiao,B.X., Miao,Y., Jiang,Z., Zhou,H. and Li,Q.N. (2011) piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clin. Chim. Acta.*, **412**, 1621–1625.
37. Chu,H., Hui,G., Yuan,L., Shi,D., Wang,Y., Du,M., Zhong,D., Ma,L., Tong,N., Qin,C. *et al.* (2015) Identification of novel piRNAs in bladder cancer. *Cancer Lett.*, **356**, 561–567.
38. Huang,G., Hu,H., Xue,X., Shen,S., Gao,E., Guo,G., Shen,X. and Zhang,X. (2013) Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. *Clin. Transl. Oncol.*, **15**, 563–568.
39. Zhang,H., Ren,Y., Xu,H., Pang,D., Duan,C. and Liu,C. (2013) The expression of stem cell protein Piwil2 and piR-932 in breast cancer. *Surg. Oncol.*, **22**, 217–223.
40. Martinez,V.D., Enfield,K.S., Rowbotham,D.A. and Lam,W.L. (2016) An atlas of gastric PIWI-interacting RNA transcriptomes and their utility for identifying signatures of gastric cancer recurrence. *Gastric Cancer*, **19**, 660–665.
41. Cheng,J., Deng,H., Xiao,B., Zhou,H., Zhou,F., Shen,Z. and Guo,J. (2012) piR-823, a novel non-coding small RNA, demonstrates in vitro and in vivo tumor suppressive activity in human gastric cancer cells. *Cancer Lett.*, **315**, 12–17.
42. Sai Lakshmi,S. and Agrawal,S. (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res.*, **36**, D173–D177.
43. Sarkar,A., Maji,R.K., Saha,S. and Ghosh,Z. (2014) piRNAQuest: searching the piRNAome for silencers. *BMC Genomics*, **15**, 555.
44. Rosenkranz,D. (2016) piRNA cluster database: a web resource for piRNA producing loci. *Nucleic Acids Res.*, **44**, D223–D230.
45. Zhang,H., Ali,A., Gao,J., Ban,R., Jiang,X., Zhang,Y. and Shi,Q. (2018) IsopiRBank: a research resource for tracking piRNA isoforms. *Database (Oxford)*, **2018**, doi:10.1093/database/bay059.
46. The, R.C. (2017) RNAcentral: a comprehensive database of non-coding RNA sequences. *Nucleic Acids Res.*, **45**, D128–D134.
47. Zhang,P., Si,X., Skogerbo,G., Wang,J., Cui,D., Li,Y., Sun,X., Liu,L., Sun,B., Chen,R. *et al.* (2014) piRBase: a web resource assisting piRNA functional study. *Database (Oxford)*, **2014**, bau110.
48. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
49. O'Leary,N.A., Wright,M.W., Brister,J.R., Ciufo,S., Haddad,D., McVeigh,R., Rajput,B., Robbertse,B., Smith-White,B., Ako-Adjei,D. *et al.* (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, **44**, D733–D745.
50. Casper,J., Zweig,A.S., Villarreal,C., Tyner,C., Speir,M.L., Rosenbloom,K.R., Raney,B.J., Lee,C.M., Lee,B.T., Karolchik,D.

et al. (2018) The UCSC Genome Browser database: 2018 update. *Nucleic Acids Res.*, **46**, D762–D769.
51. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
52. Yuan,J., Zhang,P., Cui,Y., Wang,J., Skogerbo,G., Huang,D.W., Chen,R. and He,S. (2015) Computational identification of piRNA targets on mouse mRNAs. *Bioinformatics*, **32**, 1170–1177.
53. Wu,W.S., Huang,W.C., Brown,J.S., Zhang,D., Song,X., Chen,H., Tu,S., Weng,Z. and Lee,H.C. (2018) pirScan: a webserver to predict piRNA targeting sites and to avoid transgene silencing in C. elegans. *Nucleic Acids Res.*, **46**, W43–W48.
54. Liu,B., Yang,F. and Chou,K.C. (2017) 2L-piRNA: a two-layer ensemble classifier for identifying Piwi-Interacting RNAs and their function. *Mol Ther Nucleic Acids*, **7**, 267–277.