scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Small RNA data sets of mouse testes and ovaries before and after sexual maturity

Mailin Gan[1,2,3,4], Xingyu Wang[1,2,3,4], Jianfeng Ma[1,2,3,4], Lei Chen[1,2,3], Yan Wang[1,2,3], Linyuan Shen[1,2,3] ✉ & Li Zhu[1,2,3] ✉

For a considerable period, reproductive health, fertility, and reproductive-related diseases have posed challenges to human well-being, as well as to the conservation of endangered species and the advancement of animal husbandry. PANDORA-seq, a recently introduced sequencing technique, demonstrates heightened sensitivity towards highly modified small RNAs like tsRNA and rsRNA. In this research endeavor, we leveraged PANDORA-seq to capture the small RNA expression profiles of mouse testes and ovaries pre- and post-sexual maturation. Our investigation successfully pinpointed an array of abundantly expressed small RNAs across various tissues, encompassing tsRNA, rsRNA, piRNA, miRNA, snoRNA, and ysRNA. Next, we conducted an expression profile analysis of these small RNAs to assist researchers in screening and validating them for various areas of interest. This dataset is poised to become an invaluable resource for exploring the postnatal development of testes and ovaries, offering new insights into the epigenetic mechanisms underlying germ cell production and differentiation.

Background & Summary

Sexual maturation is a critical process in the life cycle of organisms, involving the rapid development of sexual organs and glands until reaching maturity, enabling acquisition of reproductive capabilities[1]. The testes and ovaries, pivotal reproductive organs in mammals, play key roles in sperm production and ovarian follicle development. Postnatal development of these organs depends on intricate molecular signaling cascades. Disruption of these signals can lead to adverse reproductive outcomes, including reduced primordial follicle pool, underdeveloped testes, incomplete follicle development, and impaired gender differentiation[2,3]. Biological functions and responses to external stimuli are governed by gene expression. Thus, comprehending genetic regulation at all levels during animal sexual maturation is crucial for understanding species' physiological changes and addressing diseases. Several essential functional genes — such as SRY (Sex-determining Region Y), FOXL2 (Forkhead Box L2), and DAZ (Deleted in Azoospermia) — are involved in reproductive organ development and germ cell formation and maturation[4,5]. However, numerous epigenetic regulatory mechanisms operating pre- and post-transcriptionally in this process remain unexplored.

Postnatal development of testes and ovaries, as well as germ cell formation and differentiation, is intricately regulated by the hypothalamic-pituitary-gonadal axis and testosterone secretion by testicular interstitial cells[6,7]. Numerous genes exhibit precise spatiotemporal expression during this process to ensure normal reproductive function in animals[5]. Additionally, small non-coding RNAs (sncRNAs) like PIWI-interacting RNAs (piRNAs) and microRNAs (miRNAs) play vital roles in testicular and ovarian development and germ cell formation[8–11]. SncRNAs present in sperm and follicles—such as piRNAs, tRNA-derived small RNAs (tsRNAs), and rRNA-derived small RNAs (rsRNAs)—compose a "small RNA code" carrying crucial epigenetic information[12,13]. This emerging field of transgenerational inheritance partially transfers acquired epigenetic information due to environmental stimuli or organismal changes from parents to offspring[10,14], impacting various conditions like metabolic disorders[15], inflammation[16,17], and cardiovascular diseases[18]. Hence, exploring the dynamic

[1]State Key Laboratory of Swine and Poultry Breeding Industry, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu, P. R. China. [2]Key Laboratory of Livestock and Poultry Multi-omics, Ministry of Agriculture and Rural Affairs, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu, P. R. China. [3]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu, P. R. China. [4]These authors contributed equally: Mailin Gan, Xingyu Wang, Jianfeng Ma. ✉e-mail: shenlinyuan@sicau.edu.cn; zhuli@sicau.edu.cn

Sample_ID	Raw reads	Coverage rate	Mapping rate	Clean-Reads	Match-Genome-Reads	miRNA-Match-Reads	rsRNA-Match-Reads	tsRNA-Match-Reads	mt-tsRNA-Match-Reads	piRNA-Match-Reads	ysRNA-Match-Reads
male-3w-T-1	23078490	0.639807069	0.674805103	12801025	8638197	34756	5445988	459934	80190	191929	483
male-3w-T-2	26171493	0.725554671	0.683561962	14852971	10152926	61398	5982399	865822	123330	233330	973
male-3w-T-3	25966357	0.719867667	0.685079742	15580649	10673987	76548	6036434	932774	145143	292161	1116
male-10w-T-1	21319881	0.591052992	0.651558812	12401143	8080074	55637	4703789	879761	74362	216906	723
male-10w-T-2	23370251	0.647895586	0.657622654	13606239	8947771	57031	5307689	925257	87416	226705	749
male-10w-T-3	23209913	0.643450521	0.660784939	13152939	8691264	56235	5036033	931031	85146	243235	692
female-4w-O-1	20513466	0.568696676	0.747645154	12403570	9273469	65644	5109360	704616	59371	30401	3496
female-4w-O-2	26994884	0.748381615	0.689666751	14628352	10088688	142642	6614633	692971	75676	14636	3049
female-4w-O-3	21627297	0.599575514	0.682631712	11594810	7914985	104785	4824463	657494	64921	13597	3231
female-10w-O-1	22156834	0.614255916	0.661959271	12497426	8272787	122999	5135136	844943	111813	14436	3942
female-10w-O-2	28990807	0.803714769	0.675015207	16197961	10933870	145728	6684794	1049543	156255	24345	4815
female-10w-O-3	21397043	0.593192162	0.646107381	11072223	7153845	67979	4917289	518912	88577	7728	1961

Table 1. Summary of PANDORA-Seq Reads. Raw Reads: The original sequencing reads before any quality filtering or cleaning. Coverage Rate refers to the proportion of the reference genome that is covered by the sequenced reads. Mapping Rate refers to the percentage of sequencing reads that successfully align (map) to the reference genome. Clean Reads: SPORTS 1.1 generates clean reads by removing sequence adapters, discarding reads that exceed the defined length range, and eliminating non-ATUCG sequences. Genome-Aligned Reads: The number of raw reads that align with the reference genome. miRNA-Aligned Reads: The number of reads that align with miRNA sequences. rRNA-Aligned Reads: The number of reads that align with rRNA sequences. tRNA-Aligned Reads: The number of reads that align with both mature and precursor tRNA sequences. mt-tRNA-Aligned Reads: The number of reads that align with mature and precursor tRNA sequences in mitochondria. piRNA-Aligned Reads: The number of reads that align with PIWI-interacting RNA sequences.

changes in the small RNA repertoire during sexual maturation is pivotal for understanding the genetic regulations underlying postnatal reproductive organ development and germ cell formation and maturation.

Advancements in high-throughput sequencing technologies have unveiled more sncRNAs. Traditional methods for miRNA detection, lacking RNA modifications, involve 3' and 5' adapter ligation followed by reverse transcription to construct deep sequencing cDNA libraries, offering higher accuracy. However, for sncRNAs with modifications like tsRNAs and rsRNAs, cDNA conversion may be inefficient and incomplete, challenging their detection and quantification through deep sequencing. Recently, a method named PANDORA-seq (Panoramic RNA Display by Overcoming RNA modification Aborted sequencing), proposed by Shi *et al.*, employs a blend of enzymes to eliminate RNA modifications hindering adapter ligation and reverse transcription, exhibiting enhanced detection sensitivity for highly modified small RNAs compared to conventional sequencing methods, yielding more reliable outcomes[15,19,20].

To investigate the distribution patterns and dynamic changes of sncRNAs in the testes and ovaries of mice before and after sexual maturation, we collected testes and ovaries from mice at both pre- and post-maturation stages and obtained their PANDORA-seq data. A total of 12 samples were analyzed, which were divided into four experimental groups: pre-maturation testes, pre-maturation ovaries, post-maturation testes, and post-maturation ovaries, with three biological replicates per group. Next, we generated a total of 284796716 raw reads using the Illumina NextSeq sequencing platform. The sequencing results were annotated using the small RNA annotation software SPORTS1.1, which produced 160,789,308 mapped genomic reads (Table 1). In summary, our data provide a comprehensive study that successfully outlines the expression profiles of small RNAs in the testes and ovaries of mice before and after sexual maturation, offering valuable insights and a foundation for understanding the development of reproductive organs and the differentiation of germ cells post-birth. The workflow of our study is shown in Fig. 1.

Methods

Animals and sample collection. This study utilized 12 mice (C57BL/6J strain, purchased from CHENGDUDASHUO). The mice were evenly divided into four groups based on birth time and gender: 3-week-old male mice, 4-week-old female mice, 10-week-old male mice, and female mice. To ensure animal welfare, humane euthanasia was performed on the mice. Testicular tissues from male mice and ovarian tissues from female mice were collected into cryotubes and rapidly frozen in liquid nitrogen, followed by storage at −80 °C until further use. In this study, the testes of 3-week-old male mice were used as before sexual maturity testicular samples, the ovaries of 4-week-old female mice as before sexual maturity ovarian samples, and the testes and ovaries of 10-week-old male and female mice as after sexual maturity samples, which were then used for subsequent sequencing. The choice of different ages for male and female mice was based on our experimental results showing that the testes of 3-week-old male mice have not yet formed sperm, which aligns with the definition of before sexual maturity testicular changes in our study. In contrast, the ovaries of age-matched female mice were too small, making sampling difficult. By 4 weeks of age, the ovaries of female mice reached the required weight for sampling, while still not reaching sexual maturity. Since our comparison focuses on the changes in testicular or

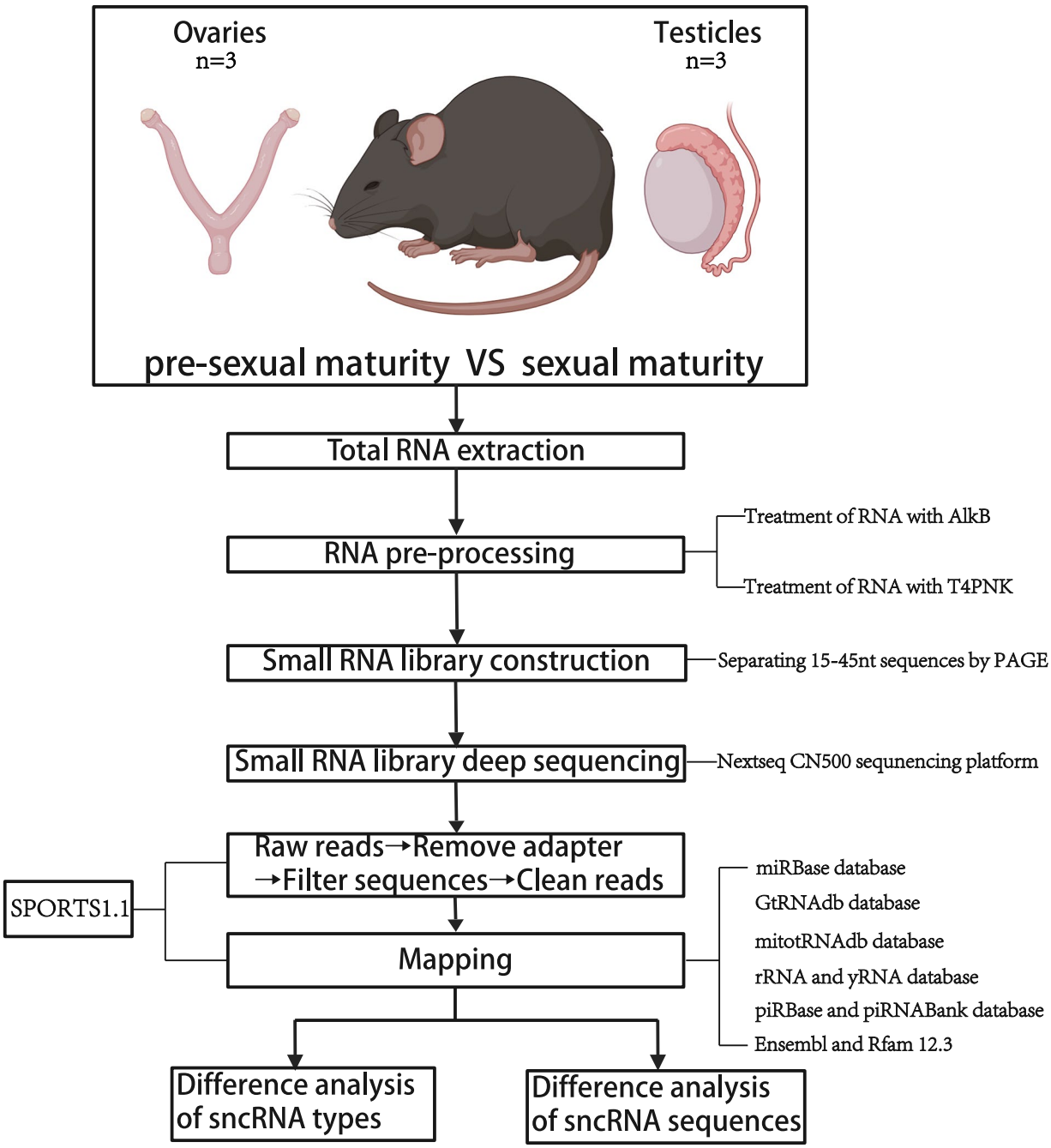

Fig. 1 Illustration of sampled animals, sample collection, sample sectioning and RNA extraction process, small RNA sequencing and data analysis.

ovarian tissues before and after sexual maturity, rather than a direct comparison between testes and ovaries at the same time point, we selected 3-week-old male mice and 4-week-old female mice for sampling.

RNA isolation. The RNA sample was mixed with an equal volume of 2 × RNA loading dye (New England Biolabs; B0363S) and incubated at 75 °C for 5 minutes. The mixture was then loaded onto a 15% (wt/vol) urea polyacrylamide gel, prepared as follows: 10 mL of the gel mixture containing 7 M urea (Invitrogen; AM9902), 3.75 mL of Acrylamide/Bis 19:1, 40% (Ambion; AM9022), 1 mL of 10 × TBE buffer (Invitrogen; AM9863), 1 g/L ammonium persulfate (Sigma-Aldrich; A3678-25G), and 1 mL/L TEMED (Thermo Fisher Scientific; BP150-100). The gel was run in 1 × TBE running buffer at 200 V until the bromophenol blue dye reached the bottom. After electrophoresis, the gel was stained with SYBR Gold solution (Invitrogen; S11494). The region containing small RNAs (15–50 nucleotides) was excised using small RNA ladders (New England Biolabs; N0364S and Takara; 3416) as markers. The excised gel slice was then eluted overnight at 4 °C in a solution containing 0.3 M sodium acetate (Invitrogen; AM9740) and 100 U/mL RNase inhibitor (New England Biolabs; M0314L). Following the

overnight incubation, the sample was centrifuged at 12,000 g for 10 minutes at 4 °C. The aqueous phase was carefully transferred to a new tube and mixed with pure ethanol, 3 M sodium acetate, and linear acrylamide (Invitrogen; AM9520) in a ratio of 3:9:0.3:0.01. The mixture was incubated at −20 °C for 2 hours to precipitate the RNA. After incubation, the sample was centrifuged at 12,000 g for 25 minutes at 4 °C. The supernatant was discarded, and the RNA pellet was resuspended in nuclease-free water. The RNA concentration was quantified, and the sample was either stored at −80 °C or used for further processing.

Treatment of RNA with AlkB. The *Escherichia coli* AlkB gene was cloned into the NdeI/BamHI restriction sites of the pET28a(+) plasmid to enable expression of the AlkB protein with a hexahistidine tag at the N-terminus. The resulting plasmid was transformed into the *E. coli* BL21(DE3) strain for protein expression. The transformed *E. coli* cells were cultured in lysogeny broth (LB) medium supplemented with 50 μg/mL kanamycin. Protein expression was induced by adding 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG), and the culture was incubated at 37 °C for 3 hours. The AlkB protein was purified using an Ni-NTA Superflow column. After purification, the protein was stored at −80 °C in a buffer containing 20 mM Tris-HCl (pH 8.0), 50% glycerol, 0.2 M NaCl, and 2 mM dithiothreitol (DTT). The purity of the AlkB protein was assessed by 12% SDS-PAGE. The RNA was incubated in 50 μl reaction mixture containing 50 mM HEPES (pH 8.0) (Gibco (15630080) and Alfa Aesar (J63578)), 75 μM ferrous ammonium sulfate (pH 5.0), 1 mM α-ketoglutaric acid (Sigma-Aldrich; K1128-25G), 2 mM sodium ascorbate, 50 mg l^{-1} bovine serum albumin (Sigma-Aldrich; A7906-500G), 4 μg ml^{-1} AlkB, 2,000 U RNase inhibitor (New England Biolabs; M0314L) and 200 ng RNA at 37 °C for 30 min. Then, the mixture was added into 500 μl TRIzol reagent to perform the RNA isolation procedure.

Treatment of RNA with T4PNK. The RNA was incubated in 50 μl reaction mixture containing 5 μl 10 × PNK buffer (New England Biolabs; B0201S), 10 mM ATP (New England Biolabs; P0756S), 10 U T4PNK (New England Biolabs;M0201L) and 200 ng RNA at 37 °C for 20 min. Then, the mixture was added into 500 μl TRIzol reagent to perform the RNA isolation procedure.

Small RNA library construction and deep sequencing. The RNA segment was separated by PAGE, then a 15- to 45-nucleotide stripe was selected and recycled. The adapters were obtained from the **QIAseq® miRNA Library Kit (QIAGEN: 331505)** and ligated sequentially. The amplified flow cell was sequenced on the Illumina system by epibiotek (Guangzhou Epibiotek Co., Ltd., Guangzhou, China).

Small RNA annotation and analyses for PANDORA-seq data. Small RNA sequences were annotated using the software SPORTS1.1 (updated from SPORTS1.0) with one mismatch tolerance (SPORTS1.1 parameter setting: -M 1)[21]. Reads were mapped to the following individual non-coding RNA databases sequentially: (1) the miRNA database miRBase[22]: In this study, we utilized all the mature.fa files available in miRbase; (2) the genomic tRNA database GtRNAdb[23]: The database version used was mm10-tRNAs.tar.gz; (3) the mitochondrial tRNA database mitotRNAdb[24]: We used the default mitotRNAdb provided in SPORTS1.1 software (https://github.com/junchaoshi/sports1.1) (https://integbio.jp/dbcatalog/en/record/nbdc01589); (4) the rRNA and YRNA databases assembled from the National Center for Biotechnology Information (NCBI) nucleotide and gene database: Detailed information about these databases can be found on the SPORTS1.1 software website (https://github.com/junchaoshi/sports1.1); (5) the piRNA databases, including piRBase and piRNABank[25]; and (6) the non-coding RNAs defined by Ensembl and Rfam 12.3[26]. The tsRNAs were annotated based on both pre-tRNA and mature tRNA sequences. Mature tRNA sequences were derived from the GtRNAdb and mitotRNAdb sequences using the following procedures: (1) predicted introns were removed; (2) a CCA sequence was added to the 3′ ends of all tRNAs; and (3) a G nucleotide was added to the 5′ end of histidine tRNAs. The tsRNAs were categorized into four types based on the origin of the tRNA loci: 5′ tsRNA (derived from the 5′ end of pre-/mature tRNA); 3′ tsRNA (derived from the 3′ end of pre-tRNA); 3′ tsRNA-CCA end (derived from the 3′ end of mature tRNA); and internal tsRNAs (not derived from 3′ or 5′ loci of tRNA). For the rsRNA annotation, we mapped the small RNAs to the parent rRNAs in an ascending order of rRNA sequence length to ensure a unique annotation of each rsRNA (for example, the rsRNAs mapped to 5.8S rRNA would not be further mapped to the genomic region overlapped by 5.8S and 45S rRNAs). Differentially expressed sncRNA analysis was performed using the R package DESeq2.

Characterisation of the expression profile of small non-coding RNAs. Based on the expression levels of various types of small RNAs, we evaluated the Spearman correlation coefficients and principal component distribution among 12 samples (Fig. 2A,B). Figure 2C presents a pie chart showing the relative expression levels of various small RNAs identified across the four sample groups. rRNA constitutes the majority (over 70%), followed by tsRNA, piRNA, and miRNA, while the expression of other sncRNAs is very low (<1%). The data reveal that, in the testes, the proportion of rRNA expression increased from 71.1% to 80.4% in post-maturation compared to pre-maturation samples, while the proportion of tsRNA decreased from 21.0% to 15.1%, piRNA from 5.8% to 3.3%, miRNA from 1.4% to 0.8%, and ysRNA from 0.68% to 0.35%. In the ovaries, the proportion of rRNA decreased from 86.0% to 83.9% in post-maturation compared to pre-maturation samples, while the proportion of tsRNA increased from 11.7% to 13.9%, and ysRNA decreased from 0.30% to 0.21% (Fig. 2C).

In Fig. 2D, we further compared the length distribution of small RNAs across the different groups using a linear graph to examine the changes in the length ratios of small RNAs before and after sexual maturation in both the testes and ovaries.

It is noteworthy that the length distribution patterns of tsRNAs in the testes and ovaries are distinct. Most of the testes tsRNAs are concentrated around 15–20 nt, while ovarian tsRNAs show two enrichment peaks, one at 15–20 nt and another at 32–37 nt. piRNAs are known for their high expression specificity in male germ cells[27]. As expected, piRNAs in the testes and ovaries exhibit significantly different length distributions. In the

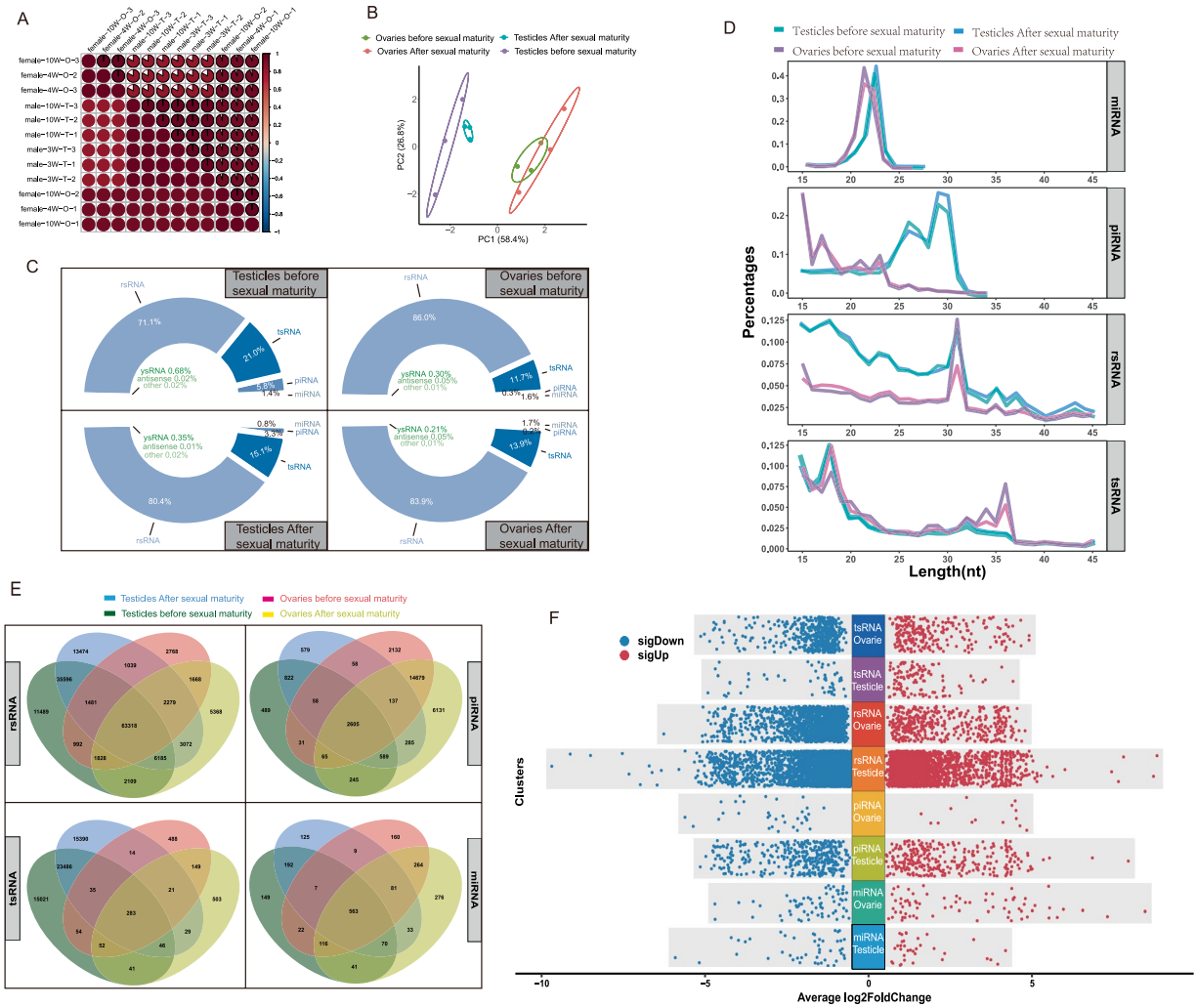

Fig. 2 Characterisation of the expression profile of small non-coding RNAs. (**A**) Spearman's correlation coefficient analysis based on the read counts of various sncRNAs. (**B**) Principal Component Analysis (PCA) based on read counts. (**C**) The proportion of reads from different types of small RNAs among the total identified small RNA reads. (**D**) The read distribution of different types of sncRNAs between 15–45 nt, with the x-axis representing the length of small RNA sequences (nt) and the y-axis indicating the proportion. (**E**) Venn diagram depicting the expression of rsRNAs, tsRNAs, piRNAs, and miRNAs across the four sample groups. (**F**) Differentially expressed small RNAs in each comparison group (after sexual maturation compared to before sexual maturation). "sigUp" and "sigDown" represent differentially expressed sncRNAs with a p-value less than 0.05, where sigUp has a Log2FC > 1 and sigDown has a Log2FC < -1.

testes, a marked change in piRNA abundance occurs between 25 and 30 nt in both pre- and post-maturation samples. Additionally, in the ovaries, rsRNA shows a clear abundance change at 31 nt between the pre- and post-maturation stages (Fig. 2D).

Next, we used Venn diagrams to display the number of rsRNAs, tsRNAs, piRNAs, and miRNAs identified in the four sample groups. The data revealed that a total of 155,666 rsRNAs, 28,905 tsRNAs, 55,612 piRNAs, and 2,108 miRNAs (RPM > 3) were identified across the four sample groups. Specifically, in the ovaries before and after sexual maturation, 16,418 and 19,864 rsRNAs, 2,279 and 7,250 tsRNAs, 591 and 619 piRNAs, and 198 and 420 miRNAs were uniquely expressed, respectively. In the testes before and after sexual maturation, 6,280 and 16,734 rsRNAs, 1,069 and 831 tsRNAs, 15,168 and 15,454 piRNAs, and 328 and 248 miRNAs were uniquely expressed, respectively (Fig. 2E).

We then compared the expression of tsRNAs, rsRNAs, piRNAs, and miRNAs in the testes before and after sexual maturation, using multi-group differential volcano plots to illustrate the distribution of differentially expressed sncRNAs during sexual maturation. The same analysis was performed for the ovaries. (Fig. 2F).

The distribution and sequence characteristics of rsRNAs, tsRNAs, piRNAs, and miRNAs. Based on the cleavage site locations, tsRNAs can be classified into tiRNA, tRF-1, tRF-2, tRF-3, tRF-5, and i-tRF[28]. SPOETS1.1 annotates these classifications using the GtRNAdb database, obtaining the categories for pre-tRNA_3' end, mature-tRNA_5' end, and mature-tRNA_3' end, which correspond to the tsRNA classifications tRF-1, tRF-5,

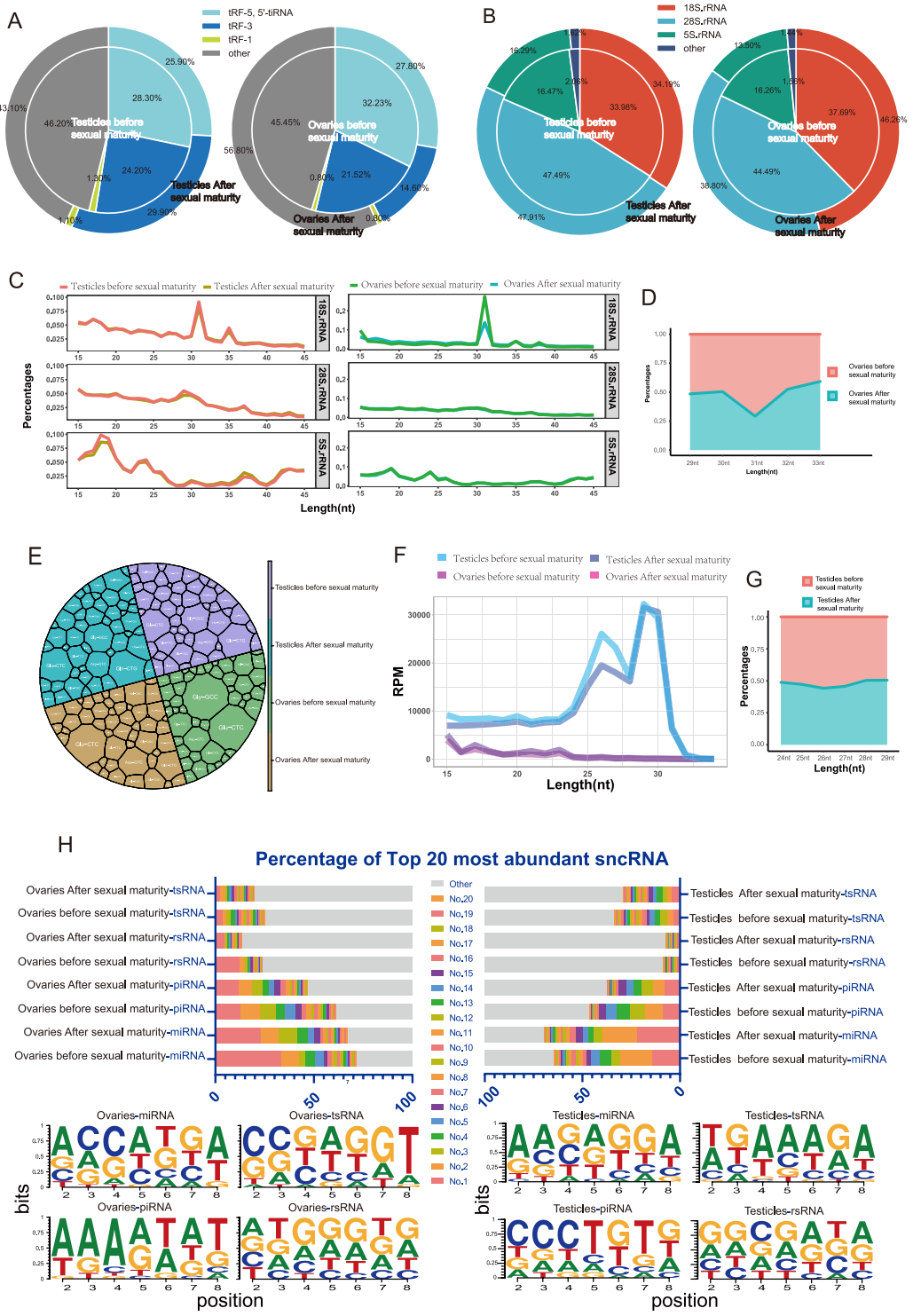

Fig. 3 The distribution and sequence characteristics of rsRNAs, tsRNAs, piRNAs, and miRNAs. (**A**) The expression proportions of tsRNAs in different categories. (**B**) The expression proportions of rsRNAs derived from different rRNA types. (**C**) The reads distribution of rsRNAs derived from 18S rRNA, 28S rRNA, and 5S rRNA in the 15–45 nt range, with the x-axis representing the length of small RNA sequences (nt) and the y-axis indicating the proportion. (**D**) Percentage of expression of 29–33 nt length rsRNA derived from 18S.rRNA in ovaries before and after sexual maturation. (**E**) Proportion of parental tRNAs of identified tsRNAs. (**H**) The abundance of the top 20 ranked small RNA reads as a percentage of the total, and the nucleotide distribution of the seed sequences of these small RNAs.

and 5' tiRNA, as well as tRF-3. Using these data, we present the changes in different tsRNA classifications in the testes after sexual maturation and in the ovaries before and after sexual maturation in the Double-layer pie chart

Fig. 4 PANDORA -seq quality score plot. The X-axis displays the position in the read, while the Y-axis represents the is quality, where = −10*log10(p), with p being the probability of error. The mean quality score is depicted as a blue line. The quality results are: Normal range (28–40),Warning range (20–28), Error range (0–20). The quality score of all samples from 1 to 75 bp falls within the Normal range.

(Fig. 3A). Additionally, we employed Voronoi Diagram to display the proportion of tsRNAs originating from tRNAs in the four identified sample groups (Fig. 3E).

Similarly, SPOETS1.1 annotates rsRNAs derived from 12S rRNA, 16S rRNA, 18S rRNA, 28S rRNA, 4.5S rRNA, 45S rRNA, 5.8S rRNA, and 5S rRNA using the rRNA database. Since the majority of rsRNAs originate from 18S rRNA, 28S rRNA, and 5S rRNA, we classified the rsRNAs from other sources as "Other." Based on this, we generated Double-layer pie chart to show the changes in the distribution of rsRNAs from different sources in the testes after sexual maturation and in the ovaries before and after sexual maturation (Fig. 3B). We further used linear plots to compare the sequence length distribution of rsRNAs derived from 18S rRNA, 28S rRNA, and 5S rRNA in the testes and ovaries before and after sexual maturation (Fig. 3C). The data show that in the ovaries before sexual maturation, the expression level of rsRNAs derived from 18S rRNA at 31 nt is significantly higher than in the ovaries after sexual maturation (Fig. 3C, D). Similarly, the expression of piRNAs in the testes before and after sexual maturation exhibits a sharp fluctuation within the 24–29 nt length range (Fig. 3F, G).

The seed sequence region is defined as the 7 nucleotides from positions 2 to 8 at the 5' end of miRNA. This region is involved in the interaction between miRNA and the 3' end of mRNA, regulating mRNA stability. Since tsRNAs, piRNAs, and rsRNAs have been reported to function similarly to miRNAs, I included these small non-coding RNAs (sncRNAs) in the analysis of their seed regions[10,16,29–31]. After organizing the sequences of different small RNAs from each group, I used WPS spreadsheet software to retrieve the seed sequences of the top 20 most expressed tsRNAs, rsRNAs, piRNAs, and miRNAs, and examined the nucleotide distribution of the seed sequences of the top 20 sncRNAs. The results show significant differences in the nucleotide distribution of seed sequences among the sncRNAs in each group (Fig. 2H).

Data Records

The RNA-Seq raw data and Processed data were deposited in the NCBI Gene Expression Omnibus (GEO) database under the accession number GSE282177[32]. That are publicly accessible at https://www.ncbi.nlm.nih.gov/geo/.

Technical Validation

Sequencing quality control. The raw data was generated using the Illumina platform. FastQC software (v0.11.7, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality scores of the raw data for each sample. The quality score graphs for each sample generated are shown in Fig. 4. The quality score Q was employed to predict the probability of base-calling error (P): $Q = -10\log10(P)$. Q30 score indicates a base calling accuracy of 99.9%, with a base calling error probability of 0.001.

Validation of experimental sample. We conducted Spearman correlation coefficient analysis on all 12 samples to assess the reproducibility of the experiment. The heatmap illustrates high correlation among the majority of biological replicates (Fig. 2A). The PCA results demonstrate that most biological samples cluster together (Fig. 2B). This undoubtedly confirms the reliability and reproducibility of our research data.

Code availability

All software utilized in this study is freely available, except for those explicitly specified in the text and methods. No custom scripts or code were employed during dataset curation and validation. The following software tools were applied:

Statistical analysis and bar chart generation were conducted using Graphpad Prism 8 (GraphPad Software Inc., USA).

Principal Component Analysis (PCA), Correlation Analysis, Pie plots, Venn plots, The seed sequence motif, and KEGG enrichment analysis were performed in the online analysis tool Bioinformatics (www.bioinformatics.com).

Linear plots, Volcano plots, and Cluster analysis heatmaps were performed in the online analysis tool omicstudio (https://www.omicstudio.cn/tool).

All analyses adhered to the methodologies outlined in the text.

Received: 10 May 2024; Accepted: 29 January 2025;
Published online: 27 February 2025

References

1. Liu, X. *et al.* Multi-omics analysis reveals changes in tryptophan and cholesterol metabolism before and after sexual maturation in captive macaques. *BMC Genomics* **24**, 308, https://doi.org/10.1186/s12864-023-09404-3 (2023).
2. Sarraj, M. A. & Drummond, A. E. Mammalian foetal ovarian development: consequences for health and disease. *Reproduction* **143**, 151–163, https://doi.org/10.1530/rep-11-0247 (2012).
3. Skakkebæk, N. E., Rajpert-De Meyts, E. & Main, K. M. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects: Opinion. *Human Reproduction* **16**, 972–978, https://doi.org/10.1093/humrep/16.5.972 (2001).
4. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G. & Clore, G. M. Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. *J Mol Biol* **312**, 481–499, https://doi.org/10.1006/jmbi.2001.4977 (2001).
5. Georges, A. *et al.* FOXL2: a central transcription factor of the ovary. *J Mol Endocrinol* **52**, R17–33, https://doi.org/10.1530/jme-13-0159 (2014).
6. Vergouwen, R. P. *et al.* Postnatal development of testicular cell populations in mice. *J Reprod Fertil* **99**, 479–485, https://doi.org/10.1530/jrf.0.0990479 (1993).
7. Wear, H. M., McPike, M. J. & Watanabe, K. H. From primordial germ cells to primordial follicles: a review and visual representation of early ovarian development in mice. *J Ovarian Res* **9**, 36, https://doi.org/10.1186/s13048-016-0246-7 (2016).
8. García-López, J. *et al.* Diversity and functional convergence of small noncoding RNAs in male germ cell differentiation and fertilization. *Rna* **21**, 946–962, https://doi.org/10.1261/rna.048215.114 (2015).
9. Yuan, S. *et al.* Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. *Development* **143**, 635–647, https://doi.org/10.1242/dev.131755 (2016).
10. Cheuquemán, C. & Maldonado, R. Non-coding RNAs and chromatin: key epigenetic factors from spermatogenesis to transgenerational inheritance. *Biological Research* **54**, 41, https://doi.org/10.1186/s40659-021-00364-0 (2021).
11. Maalouf, S. W., Liu, W. S. & Pate, J. L. MicroRNA in ovarian function. *Cell Tissue Res* **363**, 7–18, https://doi.org/10.1007/s00441-015-2307-4 (2016).
12. Zhang, Y., Shi, J., Rassoulzadegan, M., Tuorto, F. & Chen, Q. Sperm RNA code programmes the metabolic health of offspring. *Nature Reviews Endocrinology* **15**, 489–498, https://doi.org/10.1038/s41574-019-0226-2 (2019).
13. Zhang, Y. *et al.* Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs. *Nat Cell Biol* **20**, 535–540, https://doi.org/10.1038/s41556-018-0087-2 (2018).
14. Schuster, A., Skinner, M. K. & Yan, W. Ancestral vinclozolin exposure alters the epigenetic transgenerational inheritance of sperm small noncoding RNAs. *Environ Epigenet* **2**, https://doi.org/10.1093/eep/dvw001 (2016).
15. Liu, J. *et al.* Paternal phthalate exposure-elicited offspring metabolic disorders are associated with altered sperm small RNAs in mice. *Environment International* **172**, 107769, https://doi.org/10.1016/j.envint.2023.107769 (2023).
16. Chu, C. *et al.* A sequence of 28S rRNA-derived small RNAs is enriched in mature sperm and various somatic tissues and possibly associates with inflammation. *J Mol Cell Biol* **9**, 256–259, https://doi.org/10.1093/jmcb/mjx016 (2017).
17. Zhang, Y. *et al.* Angiogenin mediates paternal inflammation-induced metabolic disorders in offspring through sperm tsRNAs. *Nat Commun* **12**, 6673, https://doi.org/10.1038/s41467-021-26909-1 (2021).
18. Sui, Y., Park, S.-H., Wang, F. & Zhou, C. Perinatal Bisphenol A Exposure Increases Atherosclerosis in Adult Male PXR-Humanized Mice. *Endocrinology* **159**, 1595–1608, https://doi.org/10.1210/en.2017-03250 (2018).
19. Shi, J. *et al.* PANDORA-seq expands the repertoire of regulatory small RNAs by overcoming RNA modifications. *Nat Cell Biol* **23**, 424–436, https://doi.org/10.1038/s41556-021-00652-7 (2021).
20. Hernandez, R. *et al.* PANDORA-Seq unveils the hidden small noncoding RNA landscape in atherosclerosis of LDL receptor-deficient mice. *J Lipid Res* **64**, 100352, https://doi.org/10.1016/j.jlr.2023.100352 (2023).
21. Shi, J., Ko, E. A., Sanders, K. M., Chen, Q. & Zhou, T. SPORTS1.0: A Tool for Annotating and Profiling Non-coding RNAs Optimized for rRNA- and tRNA-derived Small RNAs. *Genomics, proteomics & bioinformatics* **16**, 144–151, https://doi.org/10.1016/j.gpb.2018.04.004 (2018).
22. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic acids research* **39**, D152–157, https://doi.org/10.1093/nar/gkq1027 (2011).
23. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic acids research* **44**, D184–189, https://doi.org/10.1093/nar/gkv1309 (2016).

24. Jühling, F. *et al*. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic acids research* **37**, D159–162, https://doi.org/10.1093/nar/gkn772 (2009).
25. Wang, J. *et al*. piRBase: integrating piRNA annotation in all aspects. *Nucleic acids research* **50**, D265–d272, https://doi.org/10.1093/nar/gkab1012 (2022).
26. Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. & Eddy, S. R. Rfam: an RNA family database. *Nucleic acids research* **31**, 439–441, https://doi.org/10.1093/nar/gkg006 (2003).
27. Ernst, C., Odom, D. T. & Kutter, C. The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. *Nature Communications* **8**, 1411, https://doi.org/10.1038/s41467-017-01049-7 (2017).
28. Liu, B. *et al*. Deciphering the tRNA-derived small RNAs: origin, development, and future. *Cell Death & Disease* **13**, 24, https://doi.org/10.1038/s41419-021-04472-3 (2021).
29. Wang, X., Ramat, A., Simonelig, M. & Liu, M.-F. Emerging roles and functional mechanisms of PIWI-interacting RNAs. *Nature Reviews Molecular Cell Biology* **24**, 123–141, https://doi.org/10.1038/s41580-022-00528-0 (2023).
30. Xiao, L., Wang, J., Ju, S., Cui, M. & Jing, R. Disorders and roles of tsRNA, snoRNA, snRNA and piRNA in cancer. *J Med Genet* **59**, 623–631, https://doi.org/10.1136/jmedgenet-2021-108327 (2022).
31. Qiao, D. *et al*. rRNA-Derived Small RNA rsRNA-28S Regulates the Chemoresistance of Prostate Cancer Cells by Targeting PTGIS. *Front Biosci (Landmark Ed)* **28**, 102, https://doi.org/10.31083/j.fbl2805102 (2023).
32. *NCBI GEO* https://identifiers.org/geo/GSE282177 (2024).

Acknowledgements

This work was supported by National Natural Science Foundation of China (32302689); China National Postdoctoral Program For Innovative Talents (BX20230250); China Postdoctoral Science Foundation (2023M732509); Sichuan Science and Technology Program (2021YFYZ0030, 2021YFYZ0007, 2024NSFSC1176); China Agriculture Research System (CARS-35).

Author contributions

L.S. and W.Y. designed the experiment. J.M and L.C. collected samples and constructed a sequencing library. M.G. and W.X. analyzed the data and wrote the paper. L.S. and L.Z. revised the paper. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.S. or L.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.