# Identification and differential expression of piRNAs in the gonads of Amur sturgeon (*Acipenser schrenckii*)

**Lihong Yuan** Corresp., [1, 2] , **Linmiao Li** [2] , **Xiujuan Zhang** [2] , **Haiying Jiang** [2] , **Jinping Chen** Corresp. [2]

[1] School of Life sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou, China

[2] Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Institute of Applied Biological Resources, Guangzhou, China

Corresponding Authors: Lihong Yuan, Jinping Chen
Email address: ylh@gdpu.edu.cn, chenjp@giabr.gd.cn

**Objective:** Sturgeons are considered living fossils, and have a very high conservation and economic value. Studies on the molecular mechanism of sturgeon gonadal development and sex differentiation would not only aid in understanding vertebrate sex determination but also benefit sturgeon aquaculture. Piwi-interacting RNAs (piRNAs) have been shown to function in germline or gonadal development. In this study, we performed small RNA deep sequencing and microarray hybridization to identify potential sturgeon piRNAs. **Methods:** Male and female sturgeon gonads were collected and used for small RNA sequencing on an Illumina HiSeq platform with the validation of piRNA expression by microarray chip. The program Bowtie and *k*-mer scheme were performed to filter small RNA reads and discover potential sturgeon piRNAs. A known piRNA database, the coding sequence (CDS), 5’ and 3’ untranslated region (UTR) database of the *A. Schrenckii* transcriptome, Gene Ontology (GO) database and KEGG pathway database were searched subsequently to analyze the potential bio-function of sturgeon piRNAs.**Results:** A total of 875,679 putative sturgeon piRNAs were obtained, including 93 homologous to known piRNAs and hundreds showing sex-specific and sex-biased expression. Further analysis showed that they are predominant in both the ovaries and testes and those with a sex-specific expression pattern are nearly equally distribution between sexes. This may imply a relevant role in sturgeon gonadal development. KEGG pathway and GO annotation analyses indicated that they may be related to sturgeon reproductive processes. **Conclusion:** Our study provides the first insights into the gonadal piRNAs in a sturgeon species and should serve as a useful resource for further elucidation of the gene regulation involved in the sex differentiation of vertebrates. These results should also facilitate the technological development of early sex identification in sturgeon aquaculture.

# Identification and differential expression of piRNAs in the gonads of Amur sturgeon (*Acipenser schrenckii*)

**Lihong Yuan[1,2], Linmiao Li[2], Xiujuan Zhang[2], Haiying Jiang[2], Jinping Chen[2]**

1. School of Life science and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou, China
2. Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Institute of Applied Biological Resources, Guangzhou, China

*** Corresponding author:**

**Lihong Yuan**

Email: ylh@gdpu.edu.cn

**Jinping Chen**

Email: chenjp@giabr.gd.cn

**Abstract**

**Objective:** Sturgeons are considered living fossils, and have a very high conservation and economic value. Studies on the molecular mechanism of sturgeon gonadal development and sex differentiation would not only aid in understanding vertebrate sex determination but also benefit sturgeon aquaculture. Piwi-interacting RNAs (piRNAs) have been shown to function in germline or gonadal development. In this study, we performed small RNA deep sequencing and microarray hybridization to identify potential sturgeon piRNAs.

**Methods:** Male and female sturgeon gonads were collected and used for small RNA sequencing on an Illumina HiSeq platform with the validation of piRNA expression by microarray chip. The program Bowtie and *k*-mer scheme were performed to filter small RNA reads and discover potential sturgeon piRNAs. A known piRNA database, the coding sequence (CDS), 5' and 3' untranslated region (UTR) database of the *A. Schrenckii* transcriptome, Gene Ontology (GO)

database and KEGG pathway database were searched subsequently to analyze the potential bio-function of sturgeon piRNAs.

**Results:** A total of 875,679 putative sturgeon piRNAs were obtained, including 93 homologous to known piRNAs and hundreds showing sex-specific and sex-biased expression. Further analysis showed that they are predominant in both the ovaries and testes and those with a sex-specific expression pattern are nearly equally distribution between sexes. This may imply a relevant role in sturgeon gonadal development. KEGG pathway and GO annotation analyses indicated that they may be related to sturgeon reproductive processes.

**Conclusion:** Our study provides the first insights into the gonadal piRNAs in a sturgeon species and should serve as a useful resource for further elucidation of the gene regulation involved in the sex differentiation of vertebrates. These results should also facilitate the technological development of early sex identification in sturgeon aquaculture.

## INTRODUCTION

Sturgeons (order: Acipenseriformes, infraclass: Chondrostei) are referred to as living fossils and have considerable value in aquaculture as sturgeon eggs (caviar) (Bemis et al., 1997; Honkala and Lobstein, 2010). Sturgeons contain 25 caviar-producing species, 17 of which are members of the *Acipenser* genus (Bemis et al., 1997). Due to the huge profits associated with the sale of sturgeon caviar, over-exploitation of wild stocks occurred worldwide throughout the 20$^{th}$ century,

and major sturgeon fisheries are in decline (Pikitch et al., 2005). Currently, all *Acipenser* species appear on the IUCN Red List of Critically Endangered Species (Ruban, 2013) and under Appendices I or II of CITES (Convention on International Trade in Endangered Species)(Dongol, 2011). Given the decline of wild stocks and conservation, great efforts have been made to develop commercial sturgeon aquaculture to meet the demand for caviar, originally in Europe and North America but more recently in Russia, Iran and China (Raymakers and Hoover, 2010; Wei et al., 2011). Currently, it is estimated that approximately 50% of the caviar in trade is from farmed stocks (Bronzi et al., 2011).

The rapid development of sturgeon aquaculture has greatly decreased the pressure on wild stocks; however, aquaculture brings new challenges. Two of the primary challenges are performing early sex identification and culling individuals with gonad dysplasia. Due to the long culture period in sturgeon (with an estimated 15 years/generation), approximately 50% of the offspring are profitless males, and up to 30% of offspring have gonad dysplasia (Krykhtin and Svirskii, 1997). Performing early sex identification, increasing the proportion of female sturgeon and decreasing gonad dysplasia may significantly improve sturgeon farming profits. Currently, the methods used in sturgeon sex and reproductive stage determination (i.e., laparoscopy, ultrasonography, histology and sex steroid analyses (Falahatkar et al., 2011; Petochi et al., 2011)) are highly dependent on technician experience and are restricted by age and gonad maturity (Devlin and

Nagahama, 2002; Masoudifard et al., 2011). Therefore, a better understanding of the processes that regulate sexual development, especially gonadogenesis and gametogenesis, may provide novel targets in sturgeon aquaculture. Furthermore, in contrast to mammals, the sex of lower bony fish is unstable and may be affected by many factors such as the environment and hormones (Devlin and Nagahama, 2002). However, due to the complexity of the number and ploidy of their chromosomes and lack of knowledge regarding sex chromosome differentiation, the study of sex determination and differentiation in sturgeons at the molecular level is still difficult.

Piwi-interacting RNAs (piRNAs), a distinct class of 26-32 nt non-coding RNAs, have been shown to function in germline development, transposon silencing and epigenetic regulation mediated by Piwi proteins (Ashe et al., 2012; Houwing et al., 2008; Juliano et al., 2011; Malone and Hannon, 2009; Ross et al., 2014; Vagin VV, 2006). Recently the piRNA biology has been expanded rapidly in not only embryonic patterning, germ cell specification, but also in stem cell biology, neuronal activity and metabolism (Rojas-Ríos et al., 2018). Studies have shown that piRNAs are expressed in both the male and female germlines of *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio* and *Xenopus laevis* and in the male germlines of mammals and birds but have limited expression in the early female germlines of mammals (Ashe et al., 2012; Ha et al., 2014; Houwing et al., 2008; Juliano et al., 2011; Lau et al., 2009; Li et al.,

2013; Qingling Yang, 2013; Roovers et al., 2015; Wilczynska et al., 2009; Williams et al., 2015). A deficiency in the genes required for piRNA biogenesis affects the regulation of gene silencing, cell differentiation and gonadal development in animals. However, the distribution of piRNAs in sturgeon gonads and the function in sex differentiation remains unclear.

*Acipenser schrenckii* (Amur sturgeon) is an economically important sturgeon species in China, and the wild stocks are mainly distributed in the Amur River, Songhua River and Heilong River (Li et al., 2012). Sturgeon small RNA transcriptome and gene expression patterns in many tissues (including the gonads) have been assayed using RNA-Seq technology, and a batch of sex-biased RNAs have been identified (Jin et al., 2015; Yuan et al., 2014; Zhang et al., 2016a; Zhang et al., 2016b). These studies support sturgeon gonad development research. In the present study, we first analyzed the putative piRNAs of Amur sturgeon gonads using the Illumina sequencing platform. We combined these data with the results of microarray expression validation to identify the putative sturgeon-specific and/or sex-specific piRNAs and illustrate the potential role of putative piRNAs on sturgeon gonadal development and sex differentiation. This study should provide information regarding piRNAs in gonads, help to reveal the mechanism of the sex duality of sturgeon, and contribute to identifying gender-related bio-markers for use in sturgeon aquaculture.

## MATERIALS AND METHODS

### Ethics statement

The protocol was approved by the Committee on the Ethics of Animal Experiments of Guangdong Institute of Applied Biological Resources (GIABR2014008). Individual sturgeon were immersed in water with 10-4 (v/v) eugenol for approximately 1-3 minutes for euthanasia, according to the AVMA guidelines (Leary et al., 2013). All efforts were made to minimize suffering.

### Sample and RNA preparation

In this study, we used 3-year-old Amur sturgeons (*Acipenser schrenckii*), whose sex could be identified accurately by laparoscopy and histology. The animals were obtained from the Engineering and Technology Center of Sturgeon Breeding and Cultivation of the Chinese Academy of Fishery Science (Beijing, China). The testes and ovaries of six 3-year-old Amur sturgeons (three males and three females) were collected. Total RNA was extracted from tissue samples separately with RNAiso reagent (TaKaRa, Japan) according to the manufacturer's instructions. The RNA concentrations were measured using a Qubit RNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies), and RNA purity was assessed using a Nano Photometer spectrophotometer (IMPLEN). RNA integrity was inspected using an RNA Nano 6000 Assay Kit and Bioanalyzer 2100 system (Agilent Technologies).

133

134 **Small RNA library preparation and sequencing**

135 Four RNA samples (2 testes and 2 ovaries, 3 μg RNA of each) were used for the construction
136 and sequencing of small RNA libraries. In brief, NEB 3' SR Adaptor ligation, SR RT Primer
137 hybridization and NEB 5' SR Adaptor ligation were performed according to the NEBNext
138 Multiplex Small RNA Sample Preparation Set (Illumina) protocol. After first strand cDNA
139 synthesis using M-MuLVA Reverse Transcriptase (RNase $H^-$) by PCR, the 140 bp to 160 bp
140 products (with adaptors on both sides) were separated on an 8% PAGE gel and quantified using
141 an Agilent Bioanalyzer 2100 system. Then, a cluster of index-coded samples was generated
142 using a TruSeq SR Cluster Kit v3-cBot-HS (Illumina) and sequenced on an Illumina HiSeq 2000
143 platform. Finally, 50 bp single-end reads were generated.

144

145 **Small RNA annotation and piRNA identification**

146 In this study, the testis and ovary transcriptomes of *A. Schrenckii* (Jin et al., 2015) were used as
147 reference sequences for small RNA annotation. After removal of unclean reads (adapters, low
148 quality reads, reads containing 'n', and redundant reads), clean unique reads were mapped onto
149 the *A. schrenckii* transcriptome reference sequences using the program Bowtie (Langmead and
150 Pop, 2009) with no mismatches allowed. Perfectly mapped reads were filtered by 3 successive
151 steps to remove small RNA elements: 1) searched against Sanger miRBase (Release 19) to

exclude conserved miRNAs; 2) screened against Rfam (http://rfam.sanger.ac.uk/) and RepeatMasker (http://www.repeatmasker.org/) with Bowtie to filter the sequences originating from rRNA, tRNA, snRNA, snoRNA and repeats; 3) analyzed by miREvo, mirdeep2 and MirCheck to remove the potential novel miRNA reads. The detailed process of small RNA filter was described in our previous study (Yuan et al., 2014).

The remaining reads with lengths of 26-32 nt were used for piRNA discovery. A *k*-mer scheme relied on the training sets of non-piRNA and the piRNA sequences of five model species (rat, mouse, human, fruit fly and nematode), was applied as previously described (Zhang et al., 2011). Then, it was compared with the existing 'static' scheme on the basis of the position-specific base usage. Putative novel piRNAs were scanned against piRNABank, (http://pirnabank.ibab.ac.in) using Bowtie with no mismatches allowed to identify the orthologs of known piRNAs. The relative frequencies of piRNA nucleotide utility reads were analyzed. Subsequently, the putative piRNAs were functionally annotated using the coding sequence (CDS), 5' and 3' untranslated region (UTR) database of the *A. Schrenckii* transcriptome with TransDecoder software (https://github.com/trinityrnaseq/trinityrnaseq/wiki/Coding-Region-Identification-in-Trinity-Assemblies).

**piRNA microarray and data analysis**

171 To validate the expression of putative piRNAs identified by Illumina sequencing, we selected: 1)
172 1,092 that exhibited significantly different expression in ovaries vs. testes with $|\log FC| \geq 1.5$, P
173 $\leq 0.01$, Padj $\leq 0.05$, and read counts > 10; and 2) 779 highly expressed putative piRNAs with
174 read counts 50 ~ 4,646 in ovaries or 50 ~ 8177 in testes. For each of the six RNA samples (three
175 ovaries and three testes, extracted above) assessed, 4 μg of total RNA were used to hybridize
176 with the microarray chip.

177

178 A piRNA microarray was manufactured by RIBOBIO (China), and each piRNA probe had three
179 replicates. The chip was hybridized with single-color labeling (Cy5) RNAs according to the
180 manufacturer's protocol, with no modifications. Microarray results were extracted using a laser
181 scanner (GenePix 4000B, Molecular Device) and digitized using Array-Pro image analysis
182 software (Media Cybernetics). Raw data were subtracted using the background matrix, and spots
183 with CV [(standard deviation)/(signal intensity)] < 0.3 were normalized using a quantile
184 normalization method to remove system related variations, including sample amount variations
185 and signal gain differences of the scanners and to faithfully reveal the biological variations
186 (Bullard et al., 2010). The medians of repeated data (normalized intensity) were used for
187 statistical analyses with One-Way ANOVA method, and the expression of piRNAs was deemed
188 significant with the criterion $|\log FC| \geq 1$ and $P \leq 0.05$.

189

190 **piRNA-generating gene prediction and annotation**

191 Putative piRNAs were mapped onto the *A. schrenckii* transcriptome using the program Bowtie
192 with no mismatches allowed to identify piRNA-generating genes. The piRNA-generating genes
193 were mapped to Gene Ontology (GO) database (http://www.geneontology.org/) by Interproscan
194 (Zdobnov et al., 2001) and KEGG pathway database (http://www.genome.jp/kegg/) by BLASTX
195 at E values 1e-10 (Kanehisa et al., 2016), respectively by two levels, 1) total of piRNA-
196 generating genes and 2) sex-specific expression (only expressed in testes or ovaries and with a
197 read count ≥ 10) or sex-biased expression of piRNA-generating genes in ovaries and testes
198 ($|\log FC| \geq 1$ and $P \leq 0.05$ obtained by microarray). Finally, the enriched functional groups or
199 pathways were obtained with corrected $P < 0.05$.

200

201 **RESULTS**

202 **Sequencing and statistics of small RNA reads**

203 A total of $7.7 \times 10^6$ - $8.6 \times 10^6$ reads were sequenced from four small RNA libraries, with an
204 error rate of 0.01%, Q30 > 97.3% and GC content of ~48% (Table S1). Then, approximately 7.3
205 $\times 10^6$ - $8.1 \times 10^6$ high-quality small RNA reads (> 94% in each library) were obtained after
206 removal of ambiguous reads (Table S1). The size distribution and frequency percentage of small
207 RNA reads are shown in Figure S1 and, of these, the potential piRNA reads (26 - 30 nt) were the
208 major component (approximately 58%). About 3.1 - $3.4 \times 10^6$ small RNA reads mapped to the *A.*

209 *schrenckii* transcriptome reference sequences (Jin et al., 2015) were with a perfect match (Table
210 S1). A total of 3.5 - 6.7 × $10^4$ reads were mapped to at least one putative Amur sturgeon miRNA
211 precursor by searching against miRBase, corresponding to 1.0 - 2.2% perfectly matched small
212 RNA reads (Table 1). Subsequent small RNA filtering indicated that other non-coding RNAs
213 (rRNA, tRNA, snRNA and snoRNA) and repeat sequences were approximately 1.0 - 1.5% and
214 0.3 - 0.4%, respectively. Then, 5.3 - 6.6 × $10^3$ (0.2%) potential novel miRNA reads specific to *A.*
215 *schrenckii* were detected with miREvo (Table 1).

216

217 **piRNA discovery**

218 According to a *k*-mer scheme analysis which relied on the training sets of non-piRNA and the
219 piRNA sequences of five model species (Zhang et al., 2011), we identified total of 5.7 - 6.2 × $10^5$
220 piRNA reads in our sequencing results, that account for 16.6 - 19.8% of perfectly matched small
221 RNA reads (Table 1). By following steps described in the methods section, 875,679 putative
222 piRNAs from the testis and ovary libraries were predicted, including 93 piRNAs that were
223 homologous to known piRNAs (4 in fruit fly, 45 in zebra fish, 34 in human, 4 in mouse and 6 in
224 rat, see Table S2). Sturgeon putative piRNAs have a strong U bias at the extreme 5' position and
225 are enriched for adenine (A) at position 10 (Figure 1A). Further analysis indicated that their rate
226 of sex-specific expression (only expressed in ovary or testis) was up to 87% (767,805/875,679),
227 and the ratio of ovary- vs. testis-specific putative piRNAs was nearly 1:1

(43.99%,385,222/875,679 vs. 43.69%, 382,583/875,679, see Figure 1B). Moreover, we found a similar percentage of sense and antisense strand putative piRNAs (42.99%, 376,487/875,679 vs. 57.01%, 499,192/875,679, respectively). In total, approximately 42% (175,648 in ovaries and 195,552 in testes) of all putative piRNAs were derived from well-annotated gene regions, 58% were from unannotated genomic regions (Figure 1C). Of these, only 31,949 (3.65%) and 73,802 (8.43%) matched the 5' UTR and 3' UTR, respectively. Abundant piRNAs in gene regions were found in CDS, which represented 265,499 (30.32%) putative piRNAs.

**Sturgeon putative piRNAs are present in both male and female gonads**

Small RNA reads of both male and female sturgeon gonads displayed a peak in reads of 26-30 nt in length (55.8% in testes and 56.5% in ovaries, shown in Figure 2A). This is an expected length distribution for potential novel piRNA reads. Whereas Small RNA reads in 5 other somatic/non-germline tissues (pool of brain, heart, muscle, liver and spleen tissues of *A. schrenckii* (Yuan et al., 2014) showed a peak in reads of 20-24 nt in length which is an expected length distribution for miRNA reads (~52.9.4%). Moreover, both ovary and testis samples had a strong 5' U bias in reads of 20-24 nt (miRNA reads) and 26-30 nt (piRNAs reads) in length, whereas only reads of 20-24 nt among somatic samples had a 5' U bias (Figure 2B-2D). The library size of testis and ovary is nearly equal (the ratio of total read counts $\approx$1:1, Table S1), thus we further compared piRNAs vs miRNAs in sturgeon testis and ovary samples, and obtained a similar ratio of

247 4.7:1~5.5:1 (Table 1). This clearly indicates putative piRNAs are far more abundant class of
248 small RNAs in both testis and ovaries compared to miRNAs. Moreover, both piRNA (1:1) and
249 miRNA (1.2:1) showed an equal ratio in testes vs. ovaries (Table 1).

250

251 **piRNA-generating genes**

252 A total of 875,679 putative piRNAs were mapped onto the *A. schrenckii* transcriptome
253 (including 122,381 unigenes in testis and 114,527 in ovaries, Jin et al., 2015). Total 49,390
254 piRNA-generating genes were obtained (23,079 in ovaries, 26,311 in testes, and 9634 in both).
255 The mean length of piRNA-generated genes was 967 nt (median 506), with a range of 150-
256 16,256 nt. There was a mean of 18 unique piRNAs reads per gene (scale of 1 up to 4,181
257 piRNAs per gene).

258

259 To illustrate the functions of the sturgeon putative piRNAs, gene function of the piRNA-
260 generated genes was annotated by searching against GO database with Interproscan and KEGG
261 pathway database by BLASTX at E values 1e-10. The GO analysis showed 264 enriched GO
262 terms for biological process, molecular function and cellular component, in them, 13 GO terms
263 related to reproduction (such as sexual reproduction, mating/reproductive behavior, post-mating
264 regulation of female receptivity) were obtained (Table S3). Moreover, 29 enriched KEGG
265 pathways, including 6 involved in aquaculture, such as pathogen infection, multiple diseases, and

antigen processing, were identified (Table S3). We further investigated the function of piRNA-generated genes which with sex-specific piRNAs located (1,078 only detected in testis and 1,237 only detected in ovary, with read counts ≥ 10), and more GO terms related to reproductive processes were identified (Figure 3 and Table S4).

**Validation of putative piRNA expression by microarray**

We used an independent microarray platform to validate the expression of 1,871 putative piRNAs (1,092 differently expressed and 779 highly expressed piRNAs, denoted ASY-piRNA-X, where Xs are numerals) identified by Illumina sequencing. A total of 311 showed significantly biased expression between sexes (deemed sex-biased expression) and were clustered into ten clades including five clades up-regulated in ovaries and five clades up-regulated in testes (Figure S2 and Table S5). Of these, 124 were specifically up-regulated in ovaries with $1<|\log FC|<8.5$ and $P \leq 0.05$, and 187 were specifically up-regulated in testes with $1<|\log FC|<4.1$ and $P \leq 0.05$ (Table S5). In sturgeon testes, the putative piRNA with the highest expression level was ASY-piR-1255 with median of normalized intensity 4.04 in testes vs. 0.72 in ovaries (Table S5). In contrast, the most expressed in the ovaries was ASY-piR-122 with median of normalized intensity 13.33 in ovaries vs. 4.91 in testes (Table S5). Moreover, 1,335 were co-expressed in both the testes and ovaries (median > 1 in both and $P > 0.05$). Further analysis indicated no significant difference in size distribution among sex-biased piRNAs (average 27.6 nt in testes

285 and 27.8 nt in ovaries) and co-expressed piRNAs in both gonads (average 27.5 nt). The Illumina
286 sequencing and microarray expression of putative piRNAs was especially correlated for 311 sex-
287 biased piRNAs ($r$ = 0.489, $P$ = 0.000). GO and KEGG annotation indicated that the 311 sex-
288 biased expressed piRNAs mainly participate in the metabolic processes of sturgeon gonads.

289

290 **DISCUSSION**

291 Here, we report the first identification of putative piRNAs and their expression in male and
292 female sturgeon (*Acipenser schrenckii*) gonads, thus greatly increasing the knowledge of
293 piRNAs in vertebrate gonadal development and/or sex differentiation. We obtained a total of
294 875,679 putative piRNAs from sturgeon gonads by Illumina sequencing, and hundreds of them
295 showed sex-specific (only expressed in sturgeon testis or ovary) and sex-biased expressed
296 validated by microarray. In addition to basic physiological processes, many reproductive
297 processes (such as sexual reproduction, mating/reproductive behavior, post-mating regulation of
298 female receptivity) as well as pathogen infection and antigen processing are shown associated
299 with the expression of sturgeon putative piRNAs.

300

301 The development of an ultrahigh-throughput sequencing technique (RNA-seq) has allowed
302 researchers to discover and analyze the piRNAs of many organisms, and tens of thousands of
303 unique sequences have been identified (Brennecke et al., 2007; Castellano et al., 2015; Houwing

304 et al., 2007; Williams et al., 2015). In this study, we obtained a total of $7.3 \times 10^6$ - $8.1 \times 10^6$
305 high-quality small RNA reads, 58% of which were potential piRNA reads of 26-30 nt in length
306 (Table S1 and Figure S1). After mapping the reads to the *A. schrenckii* transcriptome reference
307 sequences, approximately 1.2 - 2.3% reads and 16.6 - 19.8% were identified as potential
308 miRNAs and piRNAs, respectively, whereas >75% of the samples were unannotated genomic
309 regions (Table 1). By gathering additional sturgeon genomic and transcriptomic data and
310 comparing against the known sturgeon datasets, the accuracy of sturgeon gene annotation will
311 increase, and much more detailed information on *A. schrenckii* small RNAs can be uncovered
312 from the transcriptome datasets obtained in the present study.

313

314 Our study also revealed a large number of sturgeon-specific piRNAs and provided candidates for
315 further study of sturgeon gonadal development. A total of 875,679 sturgeon putative piRNAs
316 were identified, including 93 homologous to known piRNAs. The number of sturgeon putative
317 piRNAs were 10-fold and 40-fold higher than that in zebrafish ovaries and testes (Houwing et al.,
318 2007) and normal human testes (Yang Q, 2013), respectively. The large number of putative
319 piRNAs obtained in this study was probably due to chromosome ploidy resulting from multiple
320 and independent duplication events (Fontana et al., 2008; Ludwig et al., 2001). Approximately
321 86% of reads were sequenced only once, which may have been a result of the absence of
322 genomic information and the incomplete transcriptome annotation for sturgeons.

323

324 Further analyses showed that the preference of 5'U and an adenine at position 10 (10A bias) in
325 sturgeon gonads is consistent with a known piRNA biogenesis by Piwi-mediated cleavage
326 (Brennecke et al., 2007) (Figure 1A). Moreover, Small RNA reads with characteristics typical to
327 piRNAs were found in both male and female sturgeon gonads, but not in somatic tissues (Figure
328 2A). We show that both expression of sex-specific and distribution on gene regions of putative
329 piRNAs are equally between testis and ovary (Figure 1B and 1C). The prevailing expression of
330 putative piRNAs in gonads has also been found in many other organisms (Ashe et al., 2012;
331 Houwing et al., 2007; Juliano et al., 2011; Lau et al., 2009; Wilczynska et al., 2009; Williams et
332 al., 2015) and provides insight into the possible roles of sturgeon piRNAs in gametogenesis
333 between sexes. Similar to the zebrafish, sturgeon putative piRNAs and miRNAs have equivalent
334 expression in the testes and ovaries (Table 1), indicating the important role of piRNAs and
335 miRNAs in the development of sturgeon gonads (Houwing et al., 2007). Putative piRNAs exhibit
336 significantly higher expression than putative miRNAs in both male and female sturgeon gonads
337 (Table 1), as has also been observed in both male and female zebrafish gonads, in normal adult
338 human testes and mouse testes (Aravin et al., 2006; Beyret et al., 2012; Girard et al., 2006;
339 Houwing et al., 2007; Yang Q, 2013). In addition, over 87% of piRNAs showed sex-specific
340 expression; in contrast, only approximately 30% of putative miRNAs (223/730) showed sex-
341 specific expression (Zhang et al, 2018). The predominant and equivalent male and female
342 sturgeon gonads expression of putative piRNAs suggest the key role of piRNAs (rather than

343 miRNAs) in sturgeon gonad development and sex differentiation by acting through piRNA-Piwi
344 protein compound pathways.

345 To overcome the drawbacks of RNA-seq (Illumina sequencin), such as the potentiality of
346 sequencing errors and be influenced by raw data processing before small RNA identification (Git
347 et al., 2010), the expression of 1,871 putative piRNAs was validated by an independent
348 microarray platform. Data showed significantly sex-biased expression, including 124 that were
349 specifically highly expressed in ovaries and 187 in testes (Figure S2 and Table S5). There were
350 no significant differences in the size of sex-biased sturgeon putative piRNAs. A previous study
351 has shown that hsa-piR-020485 and hsa-piR-019825 (orthologs of ASY-piR-342 and ASY-piR-
352 1706) are significantly associated with sex (Yuan et al., 2016), a result consistent with our
353 observation of ASY-piR-342 and ASY-piR-1706 up-regulation in sturgeon ovaries (Table S5).
354 Moreover, we found that 1,335 putative piRNAs were co-expressed in male and female sturgeon
355 gonads (including 4 orthologs of known piRNAs), thus suggesting the crucial role of piRNAs in
356 sturgeon physiological processes. The identification of sex-biased piRNAs combined with the
357 co-expressed piRNAs in male and female gonads provides further insight into the molecular
358 mechanisms of sturgeon gonadal development and sex differentiation.

359

360 Annotation of piRNA-generating genes suggests that putative piRNAs are involved in multiple
361 reproductive processes (i.e., sexual reproduction, mating/reproductive behavior, and post-mating

362 regulation of female receptivity). Moreover, 29 enriched KEGG pathways, including those
363 relates to common challenges in aquaculture, such as pathogen infection, multiple diseases, and
364 antigen processing, were identified (Table S3). Further enrichment analysis indicated that they
365 probably participate in reproduction, mainly via sex-specific expression (Figure 3 and Table S4).

366

367 **CONCLUSION**

368 This study greatly increase the knowledge of small regulatory RNAs in the sturgeon *A.*
369 *Schrenckii*. We have identified a large number of potential novel piRNAs and provided the first
370 description of the presence of piRNAs with likely roles in sturgeon gonadal development and sex
371 differentiation. Our data demonstrate that sturgeon putative piRNAs, similar to those of zebrafish,
372 are predominantly expressed in both male and female sturgeon gonads, thus supporting a
373 potentially conserved molecular function for piRNAs in sturgeon gametogenesis between sexes.
374 Moreover, the sex-specific expression of putative piRNAs suggests that they function in sturgeon
375 gonad development and sex differentiation. Furthermore, these gender-related piRNAs are the
376 candidates to be developed as DNA bio-marker for early sex-determination with less damaging
377 to sturgeon. Finally, information uncovered by this study aids in understanding sex determination
378 in vertebrates.

379

380

381 **REFERENCES:**

382 Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein,
383 M.J., Kuramochi-Miyagawa, S., Nakano, T., 2006. A novel class of small RNAs bind to MILI protein in
384 mouse testes. Nature 442, 203.

385 Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.L., Goldstein,
386 L.D., Lehrbach, N.J., Pen, J.L., 2012. piRNAs Can Trigger a Multigenerational Epigenetic Memory in the
387 Germline of C. elegans. Cell 150, 88-99.

388 Bemis, W.E., Findeis, E.K., Grande, L., 1997. An overview of Acipenseriformes. Environmental Biology of
389 Fishes 48, 25-71.

390 Beyret, E., Liu, N., Lin, H., 2012. piRNA biogenesis during adult spermatogenesis in mice is independent of
391 the ping-pong mechanism. Cell Research 22, 1429-1439.

392 Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., Hannon, G.J., 2007. Discrete
393 Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. Cell 128, 1089–1103.

394 Bronzi, P., Rosenthal, H., Gessner, J., 2011. Global sturgeon aquaculture production: An overview. Journal of
395 Applied Ichthyology 27, 169-175.

396 Bullard, J.H., Purdom, E., Hansen, K.D., Dudoit, S., 2010. Evaluation of statistical methods for normalization
397 and differential expression in mRNA-Seq experiments. Bmc Bioinformatics 11, 1-13.

398 Castellano, L., Rizzi, E., Krell, J., Cristina, M.D., Galizi, R., Mori, A., Tam, J., Bellis, G.D., Stebbing, J.,
399 Crisanti, A., 2015. The germline of the malaria mosquito produces abundant miRNAs, endo-siRNAs, piRNAs
400 and 29-nt small RNAs. Bmc Genomics 16, 100.

401 Devlin, R.H., Nagahama, Y., 2002. Sex determination and sex differentiation in fish: an overview of genetic,
402 physiological, and environmental influences. Aquaculture 208, 191-364.

403 Dongol, Y., 2011. A Policy Gap Analysis of Convention on International Trade in Endangered Species of Wild
404 Fauna and Flora (CITES) Implementation in Nepal.

405 Falahatkar, B., Falahatkar, S., Gilani, M.H.T., Abbasalizadeh, A., 2011. Laparoscopy, a minimally-invasive
406 technique for sex identification in cultured great sturgeon Huso huso. Aquaculture 321, 273-279.

407 Fontana, F., Congiu, L., Mudrak, V.A., Quattro, J.M., Smith, T.I., Ware, K., Doroshov, S.I., 2008. Evidence of
408 hexaploid karyotype in shortnose sturgeon. 51, 113-119.

409 Girard, A., Sachidanandam, R., Hannon, G.J., Carmell, M.A., 2006. A germline-specific class of small RNAs
410 binds mammalian Piwi proteins. Nature 442, 199-202.

411 Git A., Dvinge H., Salmon-Divon M., Osborne M., Kutter C., Hadfield J., Bertone P., Caldas C., 2010.
412 Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies
413 for measuring differential microRNA expression. Rna 16, 991–1006.

414 Ha, H., Song, J., Wang, S., Kapusta, A., Feschotte, C., Chen, K.C., Xing, J., 2014. A comprehensive analysis
415 of piRNAs from adult human testis and their relationship with genes and mobile elements. Bmc Genomics 15,
416 545.

417 Honkala, I., Lobstein, A., 2010. Identification of Acipenseriformes species in trade. Journal of Applied
418 Ichthyology 24, 2-19.

419 Houwing, S., Berezikov, E., Ketting, R.F., 2008. Zili is required for germ cell differentiation and meiosis in
420 zebrafish. Embo Journal 27, 2702-2711.

421 Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., Van, d.E.H., Filippov, D.V.,
422 Blaser, H., Raz, E., Moens, C.B., 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon
423 silencing in Zebrafish. Cell 129, 69-82.

424 Jin, S.B., Zhang, Y., Dong, X.L., Xi, Q.K., Song, D., Fu, H.T., Sun, D.J., 2015. Comparative transcriptome
425 analysis of testes and ovaries for the discovery of novel genes from Amur sturgeon (Acipenser schrenckii).
426 Genetics & Molecular Research Gmr 14, 18913.

427 Juliano, C., Wang, J., Lin, H., 2011. Uniting germline and stem cells: the function of Piwi proteins and the
428 piRNA pathway in diverse organisms. Annual Review of Genetics 45, 447.

429 Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, and
430 Hirakawa M, 2006. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res 34,
431 D354-357.

432 Krykhtin, M.L., Svirskii, V.G., 1997. Endemic sturgeons of the Amur River: kaluga, Huso dauricus, and Amur
433 sturgeon, Acipenser schrenckii. Environmental Biology of Fishes 48, 231-239.

434 Langmead, B., Pop, M., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human
435 genome. Genome Biology 10, R25.

436 Lau, N.C., Ohsumi, T., Borowsky, M., Kingston, R.E., Blower, M.D., 2009. Systematic and single cell
437 analysis of Xenopus Piwi-interacting RNAs and Xiwi. Embo Journal 28, 2945-2958.

438 Leary, S., Underwood, W., Anthony, R., Cartner, S., Corey, D., Grandin, T., Greenacre, C.B., Gwaltney-Bran,
439 S., Mccrackin, M.A., Meyer, R., 2013. AVMA Guidelines for the Euthanasia of Animals: 2013 Edition.
440 University of Alaska Anchorage.

441 Li, D., Liu, Z., Xie, C., 2012. Effect of stocking density on growth and serum concentrations of thyroid
442 hormones and cortisol in Amur sturgeon, Acipenser schrenckii. Fish Physiology & Biochemistry 38, 511-520.

443 Li, X.Z., Roy, C., Dong, X., Bolcun-Filas, E., Wang, J., Han, B., Xu, J., Moore, M., Schimenti, J., Weng, Z.,

2013. An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. Molecular Cell 50, 67-81.

Ludwig, A., Belfiore, N.M., Pitra, C., Svirsky, V., Jenneckens, I., 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (Acipenser, Huso and Scaphirhynchus). Genetics 158, 1203.

Malone, C.D., Hannon, G.J., 2009. Small RNAs as Guardians of the Genome. Cell 136, 656-668.

Masoudifard, M., Vajhi, A.R., Moghim, M., Nazari, R.M., Naghavi, A.R., Sohrabnejad, M., 2011. High validity sex determination of three years old cultured beluga sturgeon (Huso huso) using ultrasonography. Journal of Applied Ichthyology 27, 643-647.

Petochi, B.T., Di, M.P., Donadelli, V., Longobardi, A., Corsalini, I., Bertotto, D., Finoia, M.G., Marino, G., 2011. Sex and reproductive stage identification of sturgeon hybrids (Acipenser naccarii × Acipenser baerii) using different tools: ultrasounds, histology and sex steroids. Journal of Applied Ichthyology 27, 637-642.

Pikitch, E.K., Doukakis, P., Lauck, L., Chakrabarty, P., Erickson, D.L., 2005. Status, trends and management of sturgeon and paddlefish fisheries. Fish and Fisheries 6, 233-265.

Qingling Yang, J.H., Liu Wang, Bo Xu, Huan Zhang, Nan Ye, Zhiqiang Zhang, Dexin Yu, Howard J. Cooke, Yuanwei Zhang, Qinghua Shi, 2013. MicroRNA and piRNA profiles in normal human testis detected by next generation sequencing. PLoS One 8, e66809.

Raymakers, C., Hoover, C., 2010. Acipenseriformes: CITES implementation from Range States to consumer countries. Journal of Applied Ichthyology 18, 629-638.

Roovers, E., Rosenkranz, D., Mahdipour, M., Han, C.T., He, N., Chuvadesousalopes, S., Vanderwesterlaken, L.J., Zischler, H., Butter, F., Roelen, B.J., 2015. Piwi Proteins and piRNAs in Mammalian Oocytes and Early Embryos. Cell Reports 10, 2069-2082.

Ross, R.J., Weiner, M.M., Lin, H., 2014. PIWI proteins and PIWI–interacting RNAs in the soma. Nature 505, 353-359.

Ruban, G.Q., W., 2013. Acipenser schrenckii. In: IUCN Red List of Threatened Species. Version 2013.2. www.iucnredlist.org.

Rojas-Ríos, P., Simonelig, M, 2018. piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. Development 145: dev161786. doi: 10.1242/dev.161786.

Vagin VV, S.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.

Wei, Q.W., Zou, Y., Li, P., Li, L., 2011. Sturgeon aquaculture in China: progress, strategies and prospects assessed on the basis of nation-wide surveys (2007–2009). Journal of Applied Ichthyology 27, 162-168.

Wilczynska, A., Minshall, N., Armisen, J., Miska, E.A., Standart, N., 2009. Two Piwi proteins, Xiwi and Xili, are expressed in the Xenopus female germline. Rna-a Publication of the Rna Society 15, 337-345.

Williams, Z., Morozov, P., Mihailovic, A., Lin, C., Puvvula, P.K., Juranek, S., Rosenwaks, Z., Tuschl, T., 2015. Discovery and Characterization of piRNAs in the Human Fetal Ovary. Cell Reports 13, 854-863.

Yang Q, H.J., Wang L, Xu B, Zhang H, Ye N, Zhang Z, Yu D, Cooke HJ, Zhang Y, Shi Q, 2013. MicroRNA and piRNA Profiles in Normal Human Testis Detected by Next Generation Sequencing. PLoS One 8, e66809.

Yuan, L., Zhang, X., Li, L., Jiang, H., Chen, J., 2014. High-Throughput Sequencing of MicroRNA Transcriptome and Expression Assay in the Sturgeon, Acipenser schrenckii. Plos One 9, e115251.

Yuan, T., Huang, X., Woodcock, M., Du, M., Dittmar, R., Wang, Y., Tsai, S., Kohli, M., Boardman, L., Patel, T., 2016. Plasma extracellular RNA profiles in healthy and cancer patients. Scientific Reports 6, 19413.

Zdobnov EM, Apweiler R, 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17, 847–848.

Zhang, X., Yuan, L., Li, L., Jiang, H., Chen, J., 2016a. Conservation, sex-biased expression and functional annotation of microRNAs in the gonad of Amur sturgeon (Acipenser schrenckii). Comparative Biochemistry & Physiology Part D Genomics & Proteomics 18, 54-61.

Zhang, X.J., Jiang, H.Y., Li, L.M., Yuan, L.H., Chen, J.P., 2016b. Transcriptome analysis and de novo annotation of the critically endangered Amur sturgeon (Acipenser schrenckii). Genetics & Molecular Research 15.

Zhang, Y., Wang, X., Kang, L., 2011. A k-mer scheme to predict piRNAs and characterize locust piRNAs. Bioinformatics 27, 771.

Zhang, X., Li, L., Jiang, H., Ma, J., Li, J., Chen, J., 2018. Identification and differential expression of microRNAs in testis and ovary of Amur sturgeon (*Acipenser schrenckii*). Gene 658, 36-46.

## ACCESSION CODES

All small RNA data series have been submitted to the SRA database under accession numbers SRR3180645, SRR3180649, SRR3180651, and SRR3180713. The microarray data have been deposited in the ArrayExpress, GEO database under accession number GSE83840.

## Tables

**Table 1. Small RNA annotation of *Acipenser schrenckii*.**

## Figures

**Figure 1. Summary of 875,679 sturgeon putative piRNAs.** (A) Base utility analysis of putative piRNAs. Black arrow indicates the preference of 5'U and an adenine at position 10$^{th}$. (B) Graphical representation of putative piRNAs between ovary and testis libraries. (C) Pie chart summary of putative piRNA distribution.

**Figure 2. Base bias analysis of *A. schrenckii* small RNAs.** (A) Length distribution of small RNAs reads from somatic tissues, testes and ovaries of *Acipenser schrenckii*. (B) - (D) Analysis

517 of the 5' position nucleotide utility of putative miRNA reads and piRNA reads in somatic tissues,
518 testes and ovaries of *A. schrenckii*. Somatic tissues: data from our previous small RNA
519 transcriptome from a pool of five tissues (brain, heart, muscle, liver and spleen) of *A. schrenckii*
520 (Yuan et al., 2014) was used as a control.

521

522 **Figure 3. Gene ontology (GO) classification annotated for piRNA-generating genes related**
523 **to reproduction**. **: Enriched GO terms related to reproduction in both levels of sex-specific
524 expressed putative piRNAs and total putative piRNAs. Details about enriched GO terms were
525 listed in Table S3 (at total putative piRNAs level) and Table S4 (at sex-specific putative piRNAs
526 level).

527

528 **Supplementary Information**

529 **Figure S1 - The sequence length distribution and frequency percentages of small RNA**
530 **reads of *Acipenser schrenckii*.** The x-axis indicates the lengths of small RNA reads. The y-axis
531 indicates the percentages of small RNA reads with a specific length. Different colors suggest
532 different types of small RNAs. A) and B), Ovary 1 and 2; C) and D), Testis 1 and 2.

533 **Figure S2. Expression pattern of sex-biased putative piRNAs validated by microarray.** The
534 expression of 311 sex-biased putative piRNAs is reflected as Log2FC. The heat map showed the

putative piRNAs clustered into ten clades based on sexes, including five clades up-regulated in ovaries and five up-regulated in testes. For detailed information see Table S5.

**Table S1- Summary of library quality control and the ambiguous reads filter.**

**Table S2 - List of orthologs to known piRNAs.**

**Table S3 - Summary of 264 enriched GO annotations and 29 enriched KEGG pathways that may associate with sturgeon putative piRNA.** *: Enriched GO/KEGG terms may involve in reproduction and aquaculture.

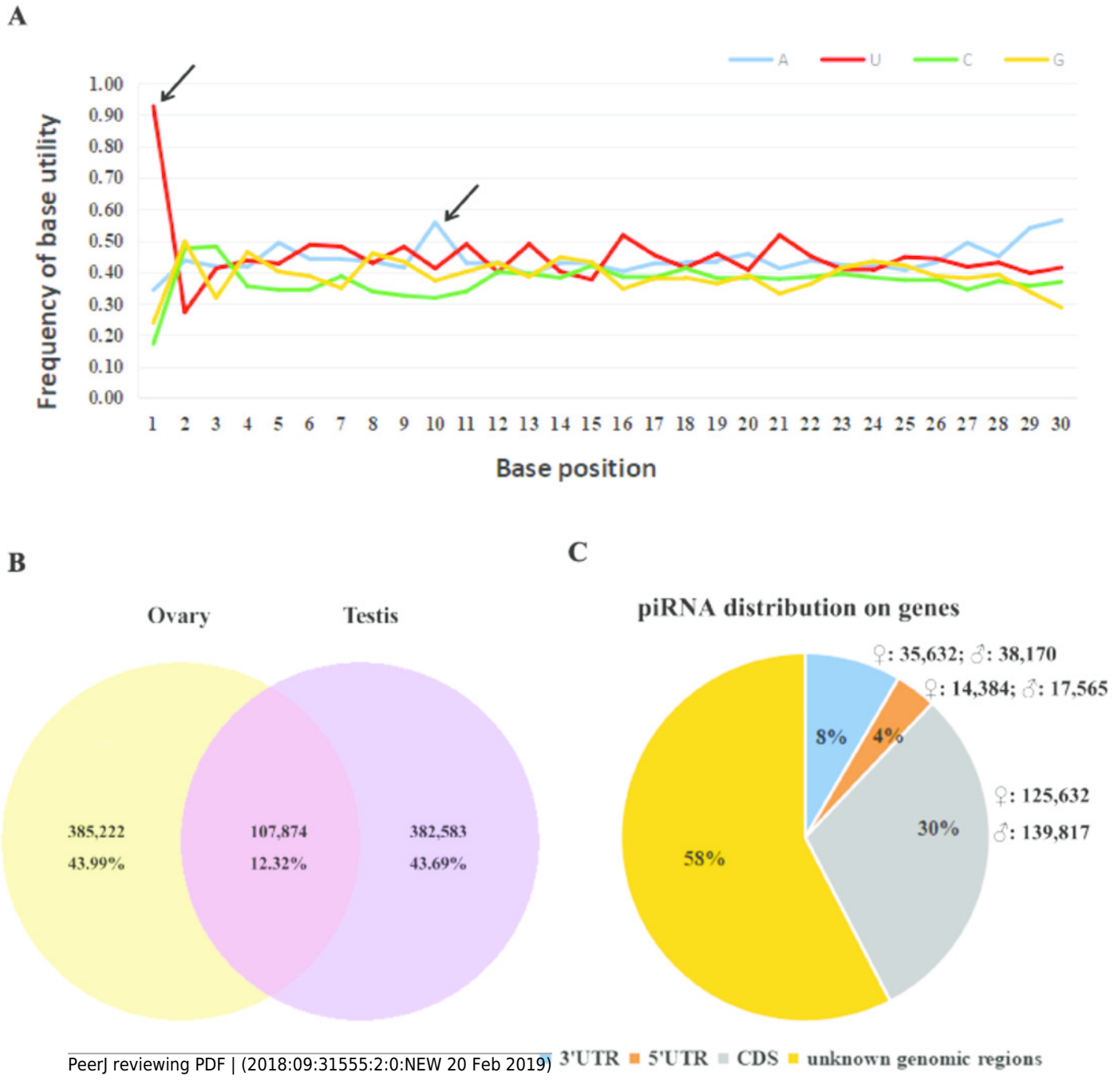
**Table S4 - Summary of 130 enriched GO annotations and 2 enriched KEGG pathways that may associate with sex-specific putative piRNA.** *: Enriched GO terms may involve in reproduction and aquaculture.

**Table S5 - Validation of the expression of 311 sex-biased putative piRNAs by miccroarray.**

# Figure 1

Summary of 875,679 sturgeon putative piRNAs.

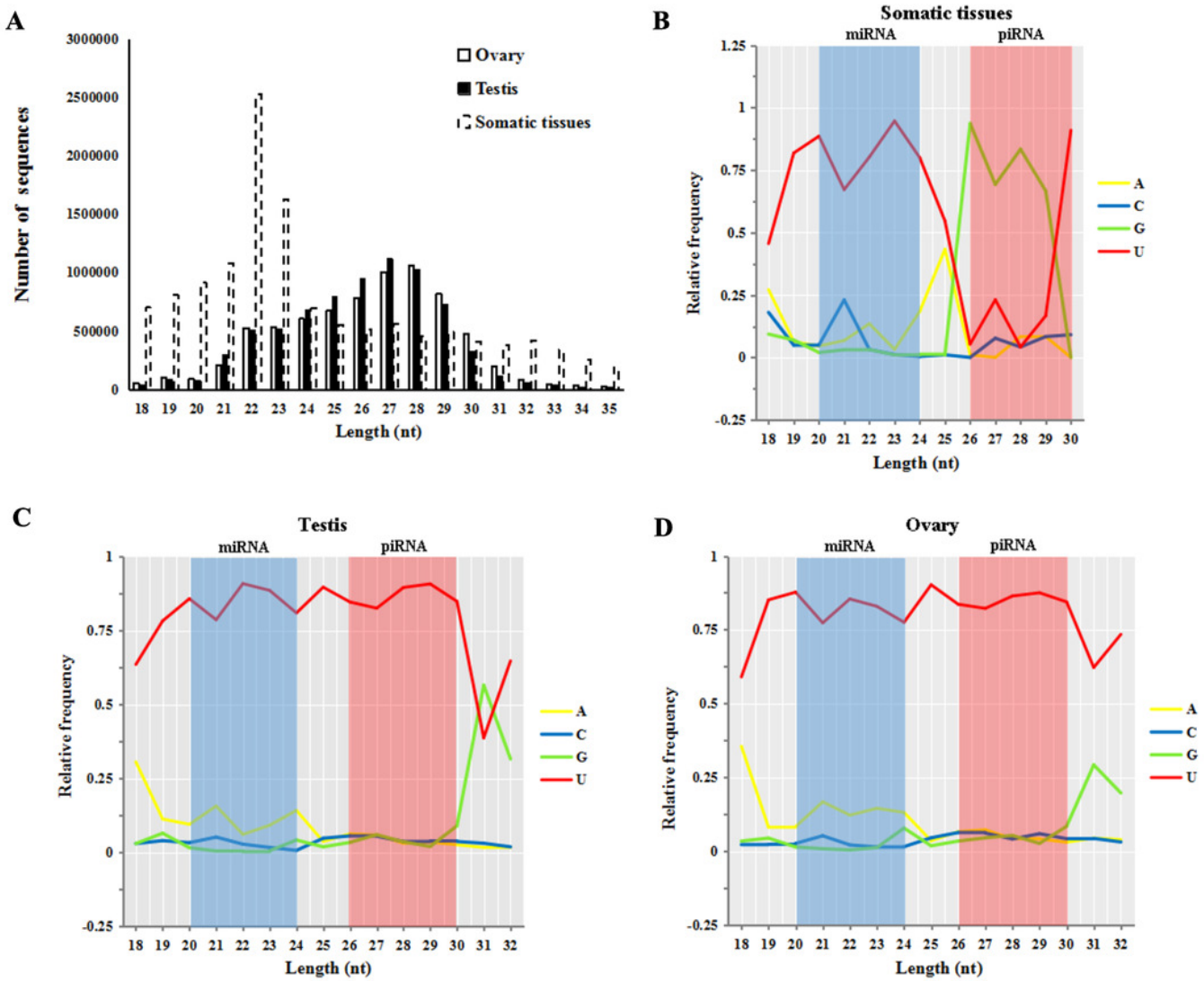
(A) Base utility analysis of putative piRNAs. Black arrow indicates the preference of 5'U and an adenine at position 10th. (B) Graphical representation of putative piRNAs between ovary and testis libraries. (C) Pie chart summary of putative piRNA distribution.



# Figure 2

Base bias analysis of A. schrenckii small RNAs.

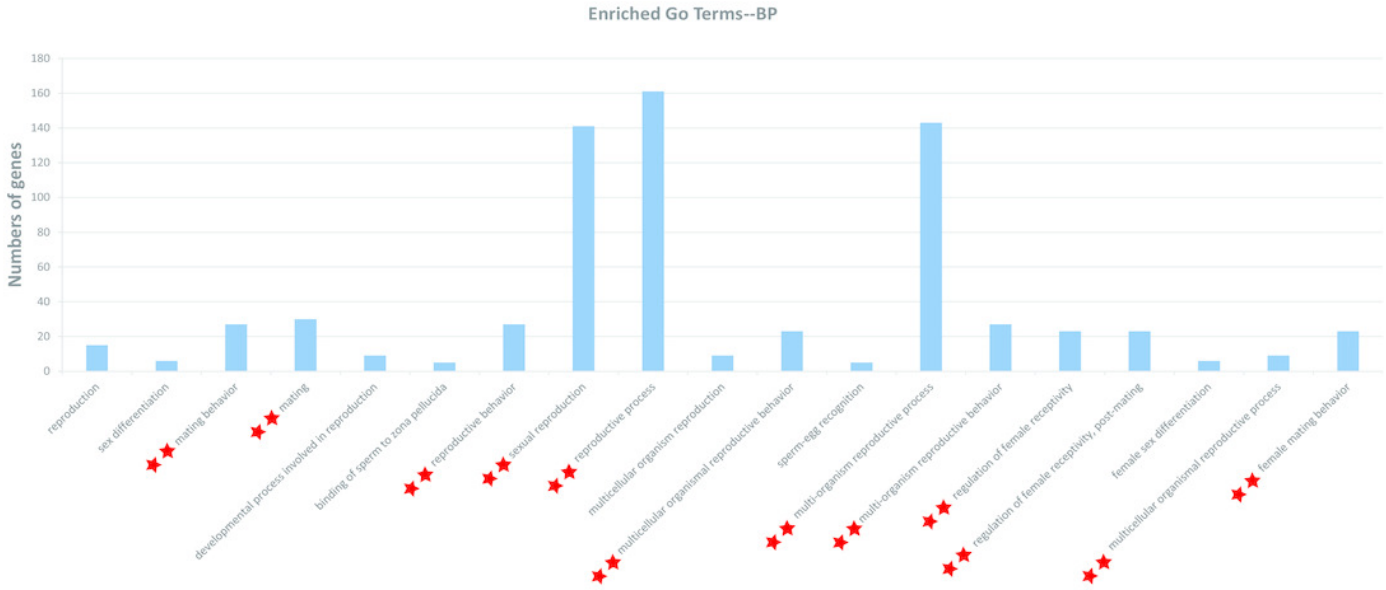
(A) Length distribution of small RNAs reads from somatic tissues, testes and ovaries of Acipenser schrenckii. (B) - (D) Analysis of the 5’ position nucleotide utility of putative miRNA reads and piRNA reads in somatic tissues, testes and ovaries of A. schrenckii. Somatic tissues: data from our previous small RNA transcriptome from a pool of five tissues (brain, heart, muscle, liver and spleen) of A. schrenckii (Yuan et al., 2014) was used as a control.



# Figure 3

Gene ontology (GO) classification annotated for piRNA-generating genes related to reproduction.

**: Enriched GO terms related to reproduction in both levels of sex-specific expressed putative piRNAs and total putative piRNAs. Details about enriched GO terms were listed in Table S3 (at total putative piRNAs level) and Table S4 (at sex-specific putative piRNAs level).



## Table 1(on next page)

Small RNA annotation of *Acipenser schrenckii*.

1 **Table 1. Small RNA annotation of *Acipenser schrenckii*.**

| Sample description | Mapped sRNA | conserved_ miRNA | rRNA | tRNA | snRNA | snoRNA | repeats | novel_m iRNA | piRNA | uniq piRNA | other |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ovary 1 | 3367185 (100.00%) | 35186 (1.04%) | 25313 (0.75%) | 1 (0.00%) | 21004 (0.62%) | 862 (0.03%) | 10175 (0.30%) | 6055 (0.18%) | 566402 (16.75%) | 272963 (8.1%) | 2704564 (80.32%) |
| Ovary 2 | 3303306 (100.00%) | 52948 (1.60%) | 28912 (0.88%) | 0 (0.00%) | 3595 (0.11%) | 826 (0.03%) | 10615 (0.32%) | 6401 (0.19%) | 595130 (18.02%) | 279287 (8.45%) | 2604879 (78.86%) |
| Testis 1 | 3164300 (100.00%) | 39703 (1.25%) | 44586 (1.41%) | 1 (0.00%) | 4478 (0.14%) | 641 (0.02%) | 10560 (0.33%) | 6552 (0.21%) | 613914 (19.40%) | 271914 (8.59%) | 2443865 (77.23%) |
| Testis 2 | 3126400 | 67386 | 39749 | 0 | 6539 | 578 | 11264 | 5340 | 620324 | 287659 | 2375220 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (100.00%) | (2.16%) | (1.27%) | (0.00%) | (0.21%) | (0.02%) | (0.36%) | (0.17%) | (19.84%) | (9.2%) | (75.97%) |

2