A Genome-Wide Survey of MicroRNA Expression Associated with Sexual Maturation in Macaque Testis

Guodong Zhao\textsuperscript{a}, Zheng Yan\textsuperscript{a}, Ning Fu\textsuperscript{b}, Haiyang Hu\textsuperscript{a} and Philip Khaytovich\textsuperscript{a,b,c}

\textsuperscript{a}Shanghai Institute for Biological Sciences, CAS, Shanghai, CHINA; \textsuperscript{b}Skolkovo Institute for Science and Technology, Skolkovo, RUSSIA; \textsuperscript{c}TheImmanuel Kant Baltic Federal University, Kaliningrad, RUSSIA.

ABSTRACT
Small RNA including microRNA (miRNA) and Piwi-interacting RNA (piRNA) play important roles in germline maintenance and maturation in a wide variety of species. However, relatively little is known about the role of miRNA in male germline maturation in humans and other primates. Here, we focused on rhesus macaques, as a model primate species closely related to humans, to study small RNA expression in testis samples throughout postnatal development and maturation. During testis development, we observe a sharp transition in overall miRNA and piRNA expression levels resulting from a rise in piRNA and a corresponding decrease in miRNA abundance. Unexpectedly, this transition takes place at approximately one year of age – far earlier than rhesus macaque sexual maturation at 4-5 years of age. Although the overall trend is for declining expression, we identified a group of 29 miRNA that show a marked increase in expression in macaque testis at approximately five years of age - the time interval associated with macaque sexual maturation. These miRNA are preferentially expressed in testis and tend to target genes involved in transcriptional regulation, DNA binding, cell migration, embryonic morphogenesis and variety of other functions associated with growth and development. In addition, 14 of these 29 miRNA originate from one primate-specific miRNA cluster.

Introduction
Small RNA, a heterogeneous RNA population with a length distribution of between 18 and 40 nucleotides (nt) play important roles in epigenetic, transcriptional and posttranscriptional regulation, and the maintenance of genomic stability. Among small RNA, two classes of single-stranded RNA molecules are particularly abundant: microRNA (miRNA) and Piwi-interacting RNA (piRNA). MiRNA are typically 21-23 nt long and are generated from non-perfectly base-paired hairpin precursors. MiRNA function as negative post-transcriptional regulators of gene expression in a broad range of species. In animals, miRNA recognize their target genes through complementary binding of the miRNA seed region to a target sequence commonly located within the 3’ untranslated region (3’UTR) of long protein-coding transcripts, resulting in target transcript decay or translation repression (Bartel, 2009; Carthew

KEYWORDS
Testis; miRNA; rhesus macaque; sexual maturity; male reproduction.

ARTICLE HISTORY
Received 19 May 2016
Revised 25 October 2016
Accepted 16 December 2016

© 2016 Guodong Zhao et al. Open Access terms of the Creative Commons Attribution 4.0 International License apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes. (http://creativecommons.org/licenses/by/4.0/)
Numerous miRNA have been demonstrated to play roles in the regulation of various biological processes, including tissue and organ development, metabolism, cell apoptosis and cancer (Ambros, 2004; Miska, 2005).

Another recently discovered class of small non-coding RNA is piRNA. piRNA are 26-30 nt in length and are predominantly expressed in the animal germline. While the exact piRNA biogenesis pathway is currently unknown, piRNA have been proposed to amplify themselves by a self-amplification mechanism called the “ping-pong model”, which is capable of generating millions of sequences that are complementary to actively transcribed transposable elements (Klattenhoff & Theurkauf, 2008). Functionally, piRNA have been implicated in transposable element silencing in germ cells, which is important in maintaining genomic stability (Senti & Brennecke, 2010; Patil & Kai, 2010).

Multiple studies have carried out extensive characterization of the anatomical, histological and physiological aspects of animal testes development and maturation, revealing discrete stages of spermatogenesis: from spermatogonial stem cells to the mature spermatids (Harrison & Weiner, 1949; Heller & Clermont, 1963; Lockwood, 1888; Franca et al., 1998). At the molecular level, expression of multiple protein-coding genes has been associated with specific stages of spermatogenesis (Sluka et al., 2002; Yomogida et al., 1994; Yu et al., 2003). Over the last decade, a number of studies have demonstrated important roles for miRNA and piRNA in male germline maintenance and sperm maturation in model organisms such as fruit flies and mice (Klattenhoff & Theurkauf, 2008; Houwing et al., 2007; Bjork et al., 2010; Boerke et al., 2007). Yet in primates, expression profiles of small RNA, such as miRNA and piRNA, within the male gonad, as well as their contribution to testis development and maturation remain largely unexplored. To fill in this blank, we investigated small genome-wide RNA expression in testis samples from 12 rhesus macaque individuals with ages ranging from 16 days to 26 years using high-throughput RNA sequencing (RNA-seq).

**Materials and Methods**

**Ethics Statement**

All macaques used in this study suffered sudden deaths for reasons other than their participation in this study and without any relation to the tissue used. The Biomedical Research Ethics Committee of the Shanghai Institutes for Biological Sciences completed a review of the use and care of the animals in this research project (approval ID: ER-SIBS-260802P).

**Sample collection and sequencing**

We collected 14 postmortem samples of testicular tissues from healthy rhesus macaque individuals with postnatal ages ranging from 16 days to 26 years. Two individuals were sampled twice to generate technical replicates (For details, see Table S1). All samples were obtained from the Suzhou Experimental Animal Center (China). All subjects were healthy and died from a sudden death unrelated to the tissue used. No subjects suffered a prolonged agonal state. The small RNA sequencing procedure was conducted following (Hu et al., 2011). In brief, we followed Illumina’s Small RNA preparation protocol to isolate low molecular weight RNA from total RNA. After adapter ligation and RNA amplification, deep sequencing was conducted using the Illumina Genome Analyzer I system. All sequencing data is deposited in NCBI GEO database (GSEXXXX)

**Mapping sequence reads to the rhesus macaque genome**
To remove the adapter sequence from the sequence reads, we first used the DUST algorithm (Mount, 2007) to remove reads with low complexity. We then trimmed reads by matching the adapter sequence to 3'-ends allowing up to three mismatches. We aligned the trimmed reads to the rhesus macaque genome (rheMac2) using the Bowtie algorithm (Langmead et al., 2009). Approximately 43% of reads mapped perfectly to the genome. Up to 50% of reads could be mapped to the genome when we allowed for one mismatch (See Table S2). Perfectly aligned reads were used for miRNA analyses. We allowed one mismatch in the aligned reads for piRNA analyses, as piRNA have a greater length. All procedures were applied to each sequencing library separately.

**MiRNA expression quantification**

Rhesus macaque miRNA annotation was defined using miRNA orthologs based on human miRNA annotation. Specifically, we downloaded human miRNA annotations from miRBase (Release 12) (Griffiths-Jones et al., 2006) and used the reciprocal BLAST method to obtain the genomic location of human miRNA orthologs in the rhesus macaque genome. MiRNA quantification was performed following (Hu et al., 2011). We normalized miRNA expression, using the total number of mapped reads for each sample, as TPM (Transcript per Million).

**Identification and clustering of age-related miRNA**

Polynomial regression models were used to identify age-related miRNA. Specifically, for each miRNA, we used families of polynomial regression models (linear, quadratic and cubic terms with individual age) to test the significance of expression level changes with age with the F-test with R “anova” function. MiRNA with p<0.01 were kept as age-related miRNA for downstream analysis. To estimate the false discovery rate of age-related miRNA, we shuffled the age labels and repeated the age test described above 1000 times to calculate the FDR under the cut off of p<0.01. To cluster age-related miRNA based on their expression change pattern with age, we used the k-means method with k=2. Use of greater cluster numbers (k) did not reveal any additional expression patterns.

**MiRNA tissue-specificity analysis**

miRNA expression values for 40 human tissues were downloaded from (Liang et al., 2007). To calculate miRNA tissue specificity, Cohen’s d (mean difference between groups normalized to the combined standard deviation) was used as a measure of effect size. Effect size was used to quantify tissue-specificity for each miRNA. The Wilcoxon signed rank test was further used to compare testis tissue-specificity between Cluster 1 and Cluster 2 miRNA.

**MiRNA target gene analysis**

We used TargetScan (Lewis et al., 2005) to predict target genes for 29 miRNAs from Cluster 1. GO (Gene Ontology) and KEGG pathway enrichment analysis were conducted using the DAVID Gene-annotation Enrichment Analysis tool (Huang da et al., 2009).

**PiRNA cluster identification**

The piRNA cluster identification method was adapted from (Yan et al., 2010). piRNA clusters were required to contain at least 30 uniquely mapped reads and show at least 60% 5’ uridine bias, piRNA clusters were merged if they were located within 1kb region.
Results and Discussion

The majority of small RNA sequences present in macaque testis correspond to miRNA and piRNA

To characterize the small RNA transcriptome of rhesus macaque testis and to determine the changes associated with development and maturation, we sequenced RNA of 18 to 32 nt in length—a size range that includes both miRNA and piRNA—in 12 rhesus macaque testis samples using the Illumina high-throughput sequencing platform. The ages of sequenced individuals ranged from 16 days postnatal until 26 years of age, and included 12 samples spanning the rhesus macaque sexual maturation period that occurs between 3 and 6 years of age (Table 1; Figure 1A). To estimate sequencing quality, we performed two technical replicates for the 16 days-old individual and a 4 year-old individual, adding up to a total of 14 separate sequencing libraries. Notably, for all individuals except one (a 3 year-old individual), testes weight was well correlated with the samples’ age (Figure 1A). This irregular sample was excluded from further analyses. In total, we obtained approximately 137 million unique sequence reads, with each sample covered by approximately 9.8 million reads on average. Approximately 50% of these reads mapped to the rhesus macaque genome when we allowed up to one mismatch (Table 2).

A length distribution analysis of the 69 million reads that mapped to the rhesus macaque genome showed two distinct peaks: one at 22 nt and another at 26-28 nt (fig. 1B). Based on previous studies, these peaks respectively represent miRNA and piRNA populations (Bartel, 2009; Carthew & Sontheimer, 2009; Klattenhoff & Theurkauf, 2008). Further, in agreement with published studies, both miRNA and piRNA populations showed a strong bias for uridine (U) at the 5’ most position (Seitz et al., 2011).

Based on these data, 538 annotated miRNA were expressed in the rhesus macaque testis. Using methodology described elsewhere (Yan et al., 2010), we also identified 5,484 piRNA clusters. Together, miRNA and piRNA constituted 78% of the total reads mapped to the rhesus macaque genome, indicating that these two types of small RNA molecules play fundamental roles in rhesus macaque testis development and maturation (Figure 2).
Figure 1. Rhesus macaque testes sample information and small RNA reads distribution. (A) The correlation of testicular weight (gram) and the respective rhesus macaque individual age (log transformed). Each circle represents one individual. The empty circle represents an individual with unusually large testes for its age, possibly developed pre-maturely. (B) Sequence reads length distribution. The x-axis shows the sequence read length in nucleotides; the y-axis shows the proportion of reads with particular length among all the mapped reads.

Table 2. Alignment statistics of the 14 sequencing libraries

<table>
<thead>
<tr>
<th>Library ID</th>
<th>Sample ID</th>
<th>Total reads</th>
<th>Perfectly mapped reads</th>
<th>Reads mapped allowing one mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>lib_01</td>
<td>mml_01</td>
<td>10436026</td>
<td>4582217</td>
<td>4924887</td>
</tr>
<tr>
<td>lib_02</td>
<td>mml_02</td>
<td>10068712</td>
<td>5094842</td>
<td>5253505</td>
</tr>
<tr>
<td>lib_03</td>
<td>mml_04</td>
<td>9568772</td>
<td>4343638</td>
<td>5215977</td>
</tr>
<tr>
<td>lib_04</td>
<td>mml_06</td>
<td>9835093</td>
<td>4497649</td>
<td>5450371</td>
</tr>
<tr>
<td>lib_05</td>
<td>mml_08</td>
<td>8930711</td>
<td>4064417</td>
<td>4799574</td>
</tr>
<tr>
<td>lib_06</td>
<td>mml_10</td>
<td>10555222</td>
<td>4193558</td>
<td>5154335</td>
</tr>
<tr>
<td>lib_07</td>
<td>mml_12</td>
<td>9545355</td>
<td>3760389</td>
<td>4599057</td>
</tr>
<tr>
<td>lib_08</td>
<td>mml_01</td>
<td>10436026</td>
<td>4006015</td>
<td>4381953</td>
</tr>
<tr>
<td>lib_09</td>
<td>mml_03</td>
<td>10068712</td>
<td>4463384</td>
<td>5048774</td>
</tr>
<tr>
<td>lib_10</td>
<td>mml_05</td>
<td>9568772</td>
<td>4423968</td>
<td>5249675</td>
</tr>
<tr>
<td>lib_11</td>
<td>mml_07</td>
<td>9835093</td>
<td>3958729</td>
<td>4671416</td>
</tr>
<tr>
<td>lib_12</td>
<td>mml_08</td>
<td>8930711</td>
<td>4709507</td>
<td>5607623</td>
</tr>
<tr>
<td>lib_13</td>
<td>mml_09</td>
<td>10555222</td>
<td>3072888</td>
<td>339245</td>
</tr>
<tr>
<td>lib_14</td>
<td>mml_11</td>
<td>9545355</td>
<td>4540392</td>
<td>5620503</td>
</tr>
</tbody>
</table>
Proportion of sequence reads that mapped to annotated miRNA or identified piRNA clusters among all mapped reads.

**Macaque testes development is characterized by sharp transition in relative abundance of miRNA and piRNA**

When we looked further into the effect of age on the small RNA transcriptome, we observed that the miRNA and piRNA levels showed quite different patterns of change across macaque lifespan. MiRNA tend to be highly expressed during the adolescent period and drop sharply at approximately one year of age and then again – although less drastically – at approximately 6.5 years of age (Figure 3A). By contrast, piRNA show the opposite pattern of expression changes with a sharp increase at approximately one year of age, followed by an additional increase at approximately 6.5 years (Figure 3A). The largest change in miRNA and piRNA expression unexpectedly takes place as early as one year of age – far earlier than rhesus macaque sexual maturation, which takes place at approximately 4 years of age (Kornack & Rakic, 1999). The premature state of testes development at one year of age can be clearly seen using the testes size distributions shown in Figure 1.

![Figure 3A, B. MiRNA expression changes with age. (A) The overall pattern of age-related abundance changes for miRNA and piRNA clusters during macaque testis development. Each time point represents an individual. One individual with suspected pre-mature testes development (see Figure 1) was excluded from further analysis. For the two individuals with technical duplicates, the average of the duplicates is shown. (B) miRNA expression patterns during macaque testis development. Each triangle represents the mean standard-normalized expression value of all miRNA in the cluster. The lines represent spline curves fitted to the data allowing three degrees of freedom. The numbers on top of the cluster panels show numbers of miRNA within the cluster.](image-url)
To identify miRNA showing significant age-related expression level changes, we used a statistical test based on a series of polynomial regressions, as described in (Somel et al., 2010). By permuting age information among samples, we further estimated the false discovery rate of the test. In total, 221 out of 538 expressed miRNA showed significant age-related expression profiles changes (p<0.01, FDR=7%) (Table 3). Of these miRNA, 192 (87%) showed a sharp expression decline with age consistent with the overall miRNA expression trajectory. Notably, 29 miRNA (13%) showed an opposite pattern: an expression increase with age especially pronounced between 4 and 6 years – the time period corresponding to rhesus macaque sexual maturation (Figures 3, B and C).

Table 3. Age test result for miRNAs

<table>
<thead>
<tr>
<th>Total quantified miRNA</th>
<th>Age-related miRNA</th>
<th>p value cutoff</th>
<th>False discovery rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>538</td>
<td>221</td>
<td>0.01</td>
<td>7.30%</td>
</tr>
</tbody>
</table>

**MiRNA highly expressed in adult testes trend to be tissue-specific and regulate genes in specific biological processes**

To further characterize the miRNA expressed in macaque testes, we investigated the tissue-specificity of the miRNA that increased (Cluster 1) and declined (Cluster 2) during testicular maturation. We used miRNA expression data from 40 healthy human tissues to calculate a tissue-specificity index for each of the tissues (Liang, et al., 2007). Using these data, we found that expression of 29 of the miRNA showing expression increases during rhesus macaque sexual maturation are significantly more specific to testis than the 192 miRNA that showed expression decline (Wilcoxon test, p=0.0015) (fig. 3D).

We next predicted the putative targets of the 29 miRNA that showed expression increases during rhesus macaque sexual maturation using the TargetScan algorithm (Lewis et al., 2005). A total of 3,940 genes were predicted to be targeted by these miRNA. Functional analysis of these genes based on human functional annotation revealed significant enrichment in many biological processes, including those associated with transcriptional regulation, DNA binding, cell migration, and embryonic morphogenesis.
Conclusion

Using deep sequencing technologies, we conducted a comprehensive analysis of small RNA changes during rhesus macaque testicular development. We assayed the expression levels of 538 annotated miRNA and identified and quantified the expression of 5,484 piRNA clusters at different stages of rhesus macaque testis developmental and maturation. We discovered a distinctive pattern of expression change for these two major types of small RNA: expression levels of miRNA tended to drop with age with the largest drop taking place at approximately one year of age and a smaller drop occurring at 5-7 years of age – the age that rhesus macaques reach sexual maturity. On the other hand, the expression of piRNA showed a reciprocal pattern with concentrations rising at one and at 5-7 years of age. In this high-throughput sequencing experiment, increases in piRNA abundance will cause apparent reciprocal decreases in miRNA levels (or vice versa) simply due to the limited number of molecules sequenced in each sequencing reaction. While increases in piRNA expression related to sexual maturation are documented (Grivna et al., 2006), our study does not provide information for whether the observed decrease in miRNA concentration is absolute or simply relative to piRNA expression levels. Still, our study clearly demonstrates that the main transition in relative quantities of miRNA and piRNA present in the macaque testis takes place at approximately one year of age. This observation is highly unexpected, as it precedes the age of sexual maturation by several years. Additionally, during this period macaque testes retain their immature size (figure 1A), indicating that the relative change in miRNA and piRNA expression levels cannot be explained by tissue growth or male germ line proliferation.

Even though the overall trend of miRNA expression decline with advanced age could be caused by increase in piRNA expression, our analysis demonstrates that a specific group of miRNA consisting of 29 members shows a reverse ontogenetic trajectory. This indicates that expression levels of these miRNA increase dramatically in sexually mature rhesus macaque individuals. This observation suggests that these miRNAs play important roles in testicular functions. Notably, some of these 29 miRNA have been shown to be associated with testis development. For instance, miR-372 and miR-373 act as oncogenes in the development of human testicular germ cell tumors by direct inhibition of the tumor-suppressor gene LATS2 (Voorhoeve et al., 2006). Further, miR-34c has been shown to be specifically expressed in male germ cells and to play pivotal regulatory roles in the late stages of spermatogenesis through its interaction with TGIF2 and NOTCH2 genes (Bouhallier et al., 2010). What’s more, 14 of the 29 miRNA up-regulated in macaque testis maturation are members of a primate-specific miRNA cluster C19MC (Bentwich et al., 2005). These miRNAs were previously shown to originate from an Alu transposable element expansion on the primate lineage and are highly expressed in human placenta and fetal brain tissues (Berezikov et al., 2006; Zhang et al., 2008). Furthermore, expression of these miRNA has been shown to be imprinted, with only the paternal allele expressed in the placenta (Noguer-Dance et al., 2010). Increased expression of these miRNA indicates that primate-specific miRNA are involved in male germline maturation. Therefore, taken together, our results strongly suggest an important role for miRNA in the testicular maturation of primates.

Acknowledgement
We would like to thank Jerome Boyd-Kirkup for help with manuscript preparation; the Suzhou Experimental Animal Center (China) for providing samples; and all members of the Comparative Biology Group in Shanghai for helpful discussions and suggestions. We also thank the Max Planck-Society, and the Bundesministerium fuer Bildung und Forschung for financial support.

Thanks. This study was supported by The Federal Targeted Programme for Research and Development in Priority Areas of Advancement of the Russian Scientific and Technological Complex for 2014-2020 (the Ministry of Education and Science of the Russian Federation), grant No 14.615.21.0002, the Unique identifier of the agreement: RFMEFI61515X0002

Disclosure statement
No potential conflict of interest was reported by the authors.

Notes on contributors
Guodong Zhao, PhD, specialist at the Shanghai Institute for Biological Sciences, CAS, Shanghai, China.
Zheng Yan, specialist at the Shanghai Institute for Biological Sciences, CAS, Shanghai, China.
Ning Fu, PhD, specialist at the Skolkovo Institute for Science and Technology, Skolkovo, Russia.
Haiyang Hu, PhD, specialist at the Shanghai Institute for Biological Sciences, CAS, Shanghai, China.
Philip Khaytovich, PhD, specialist at the Shanghai Institute for Biological Sciences, CAS, Shanghai, China; Skolkovo Institute for Science and Technology, Skolkovo, Russia; The Immanuel Kant Baltic Federal University, Kaliningrad, Russia.

References


