Gene 497 (2012) 330–335

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

# Deep sequencing analysis of small non-coding RNAs reveals the diversity of microRNAs and piRNAs in the human epididymis

Yan Li [a,1], Hai-Yan Wang [a,1], Feng-Chun Wan [a], Fu-Jun Liu [a], Juan Liu [a], Na Zhang [c], Shao-Hua Jin [a], Jian-Yuan Li [a,b,c,*]

[a] Yuhuangding Hospital, Yantai, Shandong Province, 264000, China
[b] National Research Institute for Family Planning, Beijing, 100081, China
[c] Yantai University, Yantai, Shandong Province, 264000, China

ARTICLE INFO

*Article history:*
Accepted 18 January 2012
Available online 31 January 2012

*Keywords:*
Epididymis
Small non-coding RNAs
miRNA
piRNA
Deep sequencing

ABSTRACT

The epididymis plays a crucial role in regulating the development of sperm motility and fertilizing capacity. Small non-coding RNAs (sncRNAs), especially microRNAs (miRNAs), can participate in the regulation of various physiological pathways. However, their abundance and whether they are involved in the regulation of gene expression in the human epididymis are unknown. By adopting the Solexa deep sequencing approach, we systematically investigated the sncRNAs in the adult human epididymis. A total of 4903 unique sequences representing 527 known miRNA were discovered. Eighteen novel miRNA genes encoding 23 mature miRNAs were also identified and the expression of some of them was confirmed by qRT-PCR. The presence of Piwi-interacting RNAs (piRNAs) in the library also adds to the diversity of the sncRNA population in the human epididymis. This research will contribute to a preliminary database for their functional study in male reproductive system.



## 1. Introduction

The epididymis performs an important role in sperm maturation, transport, concentration and storage (Jones, 1999; Turner, 1995). Immature spermatozoa leaving the seminiferous tubule of the testis acquire their progressive motility and fertilizing ability partly by interaction with proteins synthesized and secreted by the epididymal epithelium. The human epididymis, as that in other mammals, comprises the caput, corpus and cauda regions on the basis of histological and ultrastructural differences, with the caput and corpus carrying out early and late stages of sperm maturation, respectively, while the cauda region serves as the storage area for mature spermatozoa before ejaculation (Cornwall, 2009). All these functions are carried out in the highly diverse luminal environment along the length of the epididymal duct; thus region- and cell-specific gene expression, as well as their temporal regulation is considered crucial to the maintenance of a functional epididymis (Rodriguez et al., 2001).

The discovery of various kinds of small non-coding RNAs (sncRNAs) in recent years has revealed the complexity of the regulation of gene expression at both transcriptional and post-transcriptional levels (Huttenhofer et al., 2005; Szymanski and Barciszewski, 2002). Of the numerous sncRNAs, microRNAs (miRNAs) constitute a large family which are 19–23 nucleotides in length and widely distributed in plants, animals, microorganisms and even in unicellular organisms and viruses (Alvarez-Garcia and Miska, 2005; Bartel, 2004, 2009; Cullen, 2006; Molnar et al., 2007; Spizzo et al., 2009; Zhao et al., 2007). After transcription by Pol II or Pol III, the primary transcripts (pri-miRNAs) are subsequently processed by Drosha and Dicer endonucleases to produce miRNA/miRNA* duplexes (Borchert et al., 2006; Lee et al., 2004). On the basis of the thermodynamic stability of the duplexes, the miRNA* strand, which is the reverse sequence complementary to the miRNA strand, is typically degraded, although under certain circumstances it accumulates at a lower level, whereas the miRNA strand is selectively taken up into the micro-ribonucleoprotein complex to form an RNA-induced silencing complex (RISC). In metazoans, miRNAs tend to inhibit gene expression after transcription by binding to the partial complementary site of the 3′UTR in the target mRNA, resulting in translational repression or to a lesser extent, mRNA degradation (Bartel, 2009). Increasing evidence suggests that miRNAs are associated with diverse biological pathways, including development, differentiation, apoptosis, metabolism, and stress response, and are involved in the

*Abbreviations:* AMV, avian myeloblastosis virus; cDNA, DNA complementary to RNA; miRNA, microRNA; nt, nucleotide; PAGE, PA-gel electrophoresis; PCR, polymerase chain reaction; piRNA, Piwi-interacting RNA; Pol II, RNA polymerase II enzyme; Pol III, RNA polymerase III enzyme; pri-miRNA, primary miRNA; qRT-PCR, quantitative real time-PCR; RISC, RNA-induced silencing complex; RNA, ribonucleic acid; RNase, ribonuclease; rRNA, ribosomal RNA; snRNA, small nuclear RNA; sncRNA, small non-coding RNA; snoRNA, small nucleolar RNA; srpRNA, signal recognition particle RNA; tRNA, transfer RNA; UTR, untranslated region.

* Corresponding author at: Yuhuangding Hospital, Yantai, Shandong Province, 264000, China. Tel./fax: +86 535 6696695.
*E-mail address:* sdscli@126.com (J.-Y. Li).

[1] Yan Li and Hai-Yan Wang contributed equally to this work.

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.01.038

occurrence and development of various diseases (Plasterk, 2006; Schickel et al., 2008).

The second generation of deep sequencing technology provides a powerful strategy to explore various sncRNA populations with high sensitivity and specificity (Mardis, 2008; Morozova and Marra, 2008; Ruan et al., 2009; Zhang et al., 2011). Identification of the comprehensive sets of sncRNA species and their abundance in the human epididymis is the first and critical step towards understanding the regulatory networks involved in sperm maturation and storage. Previous work has systematically presented the profile of known miRNAs expressed in the newborn, young adult and aged human epididymides (Zhang et al., 2010). However, by employing microarray assays in that report, only known miRNAs and their abundance were studied and so the presence of other kinds of sncRNAs remains unclarified. In the present study, by employing Illumina's high-throughput sequencing technology, we identified 527 known miRNAs and 18 novel miRNA candidates in the human epididymis. The expression of nine miRNA candidates was also confirmed by stem-loop qRT-PCR. In contrast to studies in the mouse which failed to show piRNAs in the epididymis (Girard et al., 2006; Grivna et al., 2006), our results revealed that the human epididymis expresses large numbers of piRNAs. This work provides basic insights into sncRNA populations in the human epididymis and provides a preliminary database for further research clarifying their identities and regulatory roles in epididymal function.

## 2. Results and discussion

### 2.1. Overall analysis of sncRNAs in the human epididymal small RNA library

To identify sncRNA populations in the human epididymis, a library of small RNAs was constructed. By using Solexa high-throughput sequencing technology, which produces highly accurate and quantitative assessment of small RNAs, a total of 11,801,000 sequences were obtained (Supplementary data 1). After filtering out low quality data and sequences shorter than 18 nt, trimming the adaptor/acceptor sequences and cleaning up contaminants formed by adaptor–adaptor ligation, a total of 11,362,936 clean reads were obtained, representing 519,361 unique sequences. Among them, 9,122,364 (80.28%) total small RNAs, representing 303,285 (58.40%) unique sequences, were successfully mapped to the genome by using the short oligonucleotide alignment program (SOAP) (Li et al., 2008). Their expression and distribution on the genome are shown in Supplementary data 2.

The length distribution of sRNAs between 18 and 32 nt is summarized in Fig. 1. The majority of the sRNAs from the epididymal library were 22 nt in size, followed by 20 nt, 23 nt and 21 nt, corresponding well with the length of vertebrate miRNAs (Plasterk, 2006). By annotating the clean data with Genbank and Rfam databases, 6,633,813 (58.38% of the total clean data) sequences, of which 4903 (1.94%) unique were known miRNAs present in the miRBase 17.0. The rest of them were other types of RNAs, such as tRNA, rRNA, snRNA (small nuclear RNA), snoRNA (small nucleolar RNA), piRNA and srpRNA (signal recognition particle RNA), as demonstrated in Table 1. This result suggests that like other tissues, the human epididymis contains large numbers of various kinds of sncRNAs, which may contribute to the regulation of gene expression.

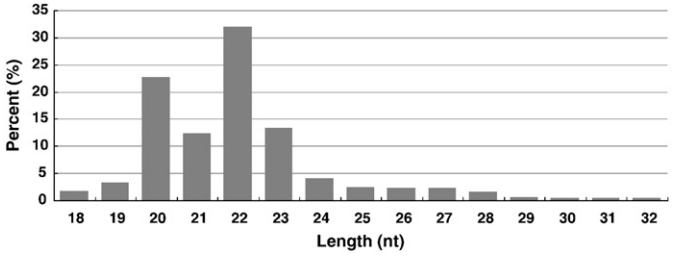


**Fig. 1.** Length distribution of sRNA tags between 18 and 32 nt.

**Table 1**
The categories of small RNA (sRNA) tags in the human epididymal library.

| Category | Unique sRNAs | Percentage of total | Total sRNAs | Percentage of total |
|---|---|---|---|---|
| Total | 519,361 | 100.00 | 11,362,936 | 100.00 |
| exon_antisense | 2125 | 0.41 | 2882 | 0.03 |
| exon_sense | 99,229 | 19.11 | 168,412 | 1.48 |
| intron_antisense | 6368 | 1.23 | 6701 | 0.06 |
| intron_sense | 42,333 | 8.15 | 44,312 | 0.39 |
| miRNA | 4903 | 0.94 | 6,633,813 | 58.38 |
| piRNA | 991 | 0.19 | 6032 | 0.05 |
| rRNA | 130,407 | 25.11 | 2,061,625 | 18.14 |
| repeat | 39,931 | 7.69 | 52,710 | 0.46 |
| scRNA | 2096 | 0.40 | 8934 | 0.08 |
| snRNA | 5574 | 1.07 | 52,826 | 0.46 |
| snoRNA | 3815 | 0.73 | 26,607 | 0.23 |
| srpRNA | 497 | 0.10 | 5611 | 0.05 |
| tRNA | 20,145 | 3.88 | 246,172 | 2.17 |
| Unannotated | 160,947 | 30.99 | 2,046,299 | 18.01 |

### 2.2. Identification of conserved miRNAs

To identify conserved miRNAs, all small RNA sequences were Blastn-searched against miRBase 17.0. Up to April 27, 2011, a total of 1424 miRNA and 194 miRNA* sequences from *Homo sapiens* have been included in miRBase17.0. In a comprehensive survey of human brain miRNAs, 602 out of 721 registered miRNA in miRBase 14.0 have been discovered (Shao et al., 2010). In the deep sequencing study of the human stomach, 404 known mature miRNAs (970 in total recruited by miRBase 15.0) were identified (Ribeiro-dos-Santos et al., 2010). Compared with those studies, there were 527 known mature miRNAs in the epididymal small RNA library. The details of all the sequenced miRNA tags are listed in Supplementary data 3. Besides these, there were abundant sequences representing 147 unique miRNA*s in the library (Supplementary data 4). In most cases, miRNA*s are undetectable by conventional methods owing to their rapid turnover rate. The presence of such large number of miRNA* indicates the high efficiency of the deep-sequencing procedure. One tag matched two miRNAs (hsa-let-7a-1 and let-7a-3), one matched three (hsa-miR-9-1, miR-9-2 and miR-9-3), while one tag matched four miRNAs (hsa-miR-518e, miR-519a-1, miR-522 and miR-523). On the whole, the abundance of miRNA*s was much lower than that of their corresponding miRNAs, consistent with the instability of miRNA* during biogenesis.

Since the sequencing frequency of the miRNAs generally reflects their relative expression abundance in the sample (Tian et al., 2010), we compared this for individual miRNAs. Of the 527 miRNAs in the library, hsa-miR-143 and hsa-let-7 family members represented the most frequent. As reported by Zhang et al. (2010), these two miRNAs are widely distributed in newborn, young adult and aged human epididymides, with hsa-miR-143 the adult-abundant form (Zhang et al., 2010). One of the target genes for miR-143 is ERK5 in human colon cancer DLD-1 cells and it may also be involved in adipocyte differentiation through targeting ERK5 (Akao et al., 2007; Esau et al., 2004). It is known that ERK5 is involved in the signaling processes downstream of various receptors including receptor tyrosine kinases and G protein-coupled receptors and is known to promote cell growth and proliferation in response to tyrosine kinase signaling (Kato et al., 1998). The large proportion (31.8% of the total miRNA sequences) of miR-143 suggests ERK5 is finely regulated and the relevant signaling pathways are quite active in the human epididymis. The top 10 abundant miRNA families and their experimentally