Comparative Transcriptome Analysis of the Accessory Sex Gland and Testis from the Chinese Mitten Crab (Eriocheir sinensis)

Lin He1, Hui Jiang1, Dandan Cao2, Lihua Liu1, Songnian Hu2, Qun Wang1*

1 School of Life Science, East China Normal University, Shanghai, China, 2 Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China

Abstract

The accessory sex gland (ASG) is an important component of the male reproductive system, which functions to enhance the fertility of spermatozoa during male reproduction. Certain proteins secreted by the ASG are known to bind to the spermatozoa membrane and affect its function. The ASG gene expression profile in Chinese mitten crab (Eriocheir sinensis) has not been extensively studied, and limited genetic research has been conducted on this species. The advent of high-throughput sequencing technologies enables the generation of genomic resources within a short period of time and at minimal cost. In the present study, we performed de novo transcriptome sequencing to produce a comprehensive transcript dataset for the ASG of E. sinensis using Illumina sequencing technology. This analysis yielded a total of 33,221,284 sequencing reads, including 2.6 Gb of total nucleotides. Reads were assembled into 85,913 contigs (average 218 bp), or 58,567 scaffold sequences (average 292 bp), that identified 37,955 unigenes (average 385 bp). We assembled all unigenes and compared them with the published testis transcriptome from E. sinensis. In order to identify which genes may be involved in ASG function, as it pertains to modification of spermatozoa, we compared the ASG and testis transcriptome of E. sinensis. Our analysis identified specific genes with both higher and lower tissue expression levels in the two tissues, and the functions of these genes were analyzed to elucidate their potential roles during maturation of spermatozoa. Availability of detailed transcriptome data from ASG and testis in E. sinensis can assist our understanding of the molecular mechanisms involved with spermatozoa conservation, transport, maturation and capacitation and potentially acrossome activation.

Citation: He L, Jiang H, Cao D, Liu L, Hu S, et al. (2013) Comparative Transcriptome Analysis of the Accessory Sex Gland and Testis from the Chinese Mitten Crab (Eriocheir sinensis). PLOS ONE 8(1): e53915. doi:10.1371/journal.pone.0053915

Editor: Zhanjiang Liu, Auburn University, United States of America

Received August 24, 2012; Accepted December 4, 2012; Published January 16, 2013

Copyright: © 2013 He et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by grants from National Natural Science Foundation of China (No. 31172393 and No. 31201974) and the Science and Technology Commission of Shanghai Municipality (No. 12ZR1408900). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: qun_300@hotmail.com

Introduction

The product of spermatogenesis is a genetically unique male gamete that can fertilize an ovum and produce offspring. Spermatogenesis and the accumulation of spermatozoa occur in the unique tissues of the testis, in a process that involves a series of intricate, cellular, proliferative and developmental phases. Spermatozoa are not capable of fertilizing an oocyte immediately after completing spermatogenesis and spermiation in the testis, though transport through the accessory sex glands (ASG) changes the activity of spermatozoa [1]. The testis and epididymis are the two male reproductive glands that produce spermatozoa and secrete androgens with the testis being responsible for continuous production of spermatozoa, and the epididymis ensuring production of a heterogeneous sperm population capable of fertilizing an oocyte and also acting as a reservoir for male gametes [2]. In mammals, it is well established that some important sperm attributes are acquired during epididymal transit, including motility, oocyte binding, and penetrating capacity, but there is also evidence that secretions from the ASG influence other aspects of sperm physiology and fertilization [3]. Insects and crustaceans have no additional accessoryial glands, and the function of the ASG corresponds with the function of the epididymis in mammals. In most species, sperm maturation studies have focused on secretions from the ASG, and have reported that these secretions are able to enhance fertilizing capacity of sperm collected from the cauda epididymis [4].

As stated above, sperm maturation and fertilizing capacity are not intrinsic to sperm themselves but are acquired during their transit through the epididymis [5]. Post-meiotic haploid spermatids differentiate into mature spermatozoa via highly specialized processes, this modification of spermatozoa can occur in the epididymis or ASG [6]. The ASG is known to have a significant function in mammals, and its secretions contain a variety of bioactive molecules that exert wide-ranging effects on female reproductive activity, they also improve the male’s chances of successful reproduction [7]. In addition, some ASG proteins provide nutritional factors to newly developed spermatozoa, and other yet unidentified factors are capable of inducing a cascade of spermatozoa membrane alterations that exert an influence on spermatozoa vitality [8], physiological state, motility and capacitation [9], as well as fertilization capacity [10]. A delicate reorientation and modification of sperm surface molecules takes place when sperm are activated by capacitation factors. These
surface changes are probably required to enable the sperm to bind to the extracellular matrix of the oocyte (the zona pellucida, ZP) [11]. For example, sperm surface coating protein that normally prevent adhesion are lost during transit of sperm in the uterus and are recoated in the oviduct. The surface of the sperm cell may also be modified by the oviduct epithelium that adsorbs proteins from the sperm surface and also secretes glycoproteins with an unknown function in sperm-ZP binding [12].

The Chinese mitten crab (Eriocheir sinensis) (Henri Milne Edwards 1854) is one of the most important aquaculture species in China and has high commercial value as a food source [13]. E. sinensis is a catadromous crustacean with a life-span of about two years. During its complex life cycle, the crab spends most of its life in rivers and lakes [14]. Adults migrate downstream towards estuarine waters, where they reach maturity and mate from November to March before moving into high salinity regions in estuaries where they release the larvae during early spring [15]. This species reproduces only once and dies shortly afterwards. Relative to mammals, E. sinensis require more complex environments to induce mating and spawning, and unique regulatory mechanisms are involved in crustacean reproduction. Sexual precocity has been reported in cultured Chinese E. sinensis populations since development of their intensive aquaculture in the early 1980s [16]. Precocious crabs mature and die prematurely at a small size, where this occurs it can lead to catastrophic losses for farmers and this problem seriously impacts development of crab aquaculture. The molecular mechanisms underlying E. sinensis sexual precocity remain unclear. As a consequence, genetic mechanisms involved in growth, reproduction and immune response of E. sinensis are currently an active research area for this economically important species.

Recently, the focus of E. sinensis research in reproductive and developmental biology has shifted from histological and biochemical analyses to genetic and molecular studies [17]. In this regard, genes crucial for reproduction and development need to be identified and their regulatory mechanisms elucidated. Transcriptome sequencing yields a subset of genes from the genome that are functionally active in selected tissues and species of interest. In nonstandard model organisms where genomic resources are lacking, such as a fully sequenced genome, obtaining a transcriptome is an effective way to evaluate gene expression and to perform comparative studies at the whole genome level [18]. In order to study gene expression profiles during spermatogenesis, we previously performed de novo transcriptome sequencing to produce a comprehensive transcript dataset for E. sinensis testis, that produced 25,698,778 sequencing reads corresponding with 2.31 Gb of total nucleotides. Reads were assembled into 342,753 contigs or 141,861 scaffold sequences, that identified 96,311 unigenes [19]. In the above mentioned study, we identified several sperm membrane proteins, that may be modified by ASG proteins during maturation, which we later identified as ASG proteins involved in spermatophore rupture [20]. In a continuation of our previous studies, we have performed a de novo transcriptome analysis for the E. sinensis ASG, and present a comparative analysis of the transcriptome for both the ASG and testis in E. sinensis in order to elucidate ASG function in sperm maturation. The analysis was based on construction of annotated ASG and testis transcriptome libraries by de novo assembly of short raw reads generated by high-throughput technology (Illumina Solexa sequencing) without genomic sequence information. We believe global approaches of this type will pave the way to allow development of a more complete understanding of the complex gene and protein networks that drive the biological and reproductive processes of spermatogenesis. The goal of this research is to provide a general overview of the potential molecular mechanisms that are involved in E. sinensis reproduction and to find key genes or pathways that function in the process of fertilization and spermatogenesis. Furthermore, we hope to provide fundamental and significant information about the sperm maturation process during transport through the ASG in E. sinensis, and elucidate sperm modification mechanisms during the acrosome reaction and sperm-oocyte interactions.

Materials and Methods

Tissue Sampling, cDNA Library Creation, and Sequencing

All animal investigations were carried out according to Animal Care and Use of Science and Technology guidelines. Healthy, sexually mature, male mitten crabs (E. sinensis, weighing 150 to 200 g) that had reached the stage of rapid ASG development were obtained from a commercial crab farm (Caojing Town special aquaculture farm in Jinshan District) near Shanghai, China between October and December in 2010. Male crabs were dissected on ice, the ASGs were removed immediately and tissues were flash frozen in liquid nitrogen. ASG tissues from three different individuals were taken on three occasions, and the nine pairs of ASG tissue were pooled as a single sample for RNA extraction. Total RNA was isolated using TRizol reagent (Invitrogen, Shanghai, China). The RNA integrity score and quantity were determined using an Agilent 2100 Bioanalyzer (Agilent, Shanghai, China) before cDNA synthesis. RNA extraction, cDNA synthesis, cDNA library normalization, and Illumina sequencing were performed according to published methods [19].

Transcriptome Assembly

Transcriptome de novo assembly was carried out with the short read assembling program SOAPdenovo-v1.03 [21]. All subsequent analyses were based on clean reads. Reads with certain lengths of overlap and no uncalled bases (N) were combined as contigs to form longer fragments. Contigs were then connected using N to represent the unknown sequence between each pair of contigs to form scaffolds. Paired-end reads were used for gap filling of scaffolds to obtain sequences with the smallest number of N’s. These sequences were defined as unigenes. In the final step, Blastx alignments (E-value $\leq 10^{-5}$) between unigenes and sequences in protein databases, including the National Center for Biotechnology Information (NCBI) non-redundant (nr) database, Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/) and Clusters of Orthologous Groups (COG) were performed to identify the sequence direction of unigenes. If results of different databases were conflicting, a priority order of alignments from the nr, Swiss-Prot, KEGG and COG databases was followed to decide the sequence direction. When a unigene happened to be unaligned to any sequence in the above databases, the software program ESTScan [22] was used to define the sequence direction. For unigenes with determined sequence directions, we identified their sequences from the 5’ to 3’ end and for those with undetermined directions, we provided their sequence based on the assembly software. When multiple samples from the same species are sequenced, unigenes from each sample’s assembly can be further processed for sequence splicing and removal of redundancy with sequence clustering software to acquire the longest reads of nr unigenes (Fig. 1).

Homology Searches and Functional Unigene Annotation

Annotation provides information on expression and function of a unigene. In our functional annotation, unigene sequences were first aligned using Blastx to the nr, Swiss-Prot, KEGG and COG
protein databases (E-value $<10^{-5}$), to retrieve proteins with the highest sequence similarity to *E. sinensis* unigenes along with their protein functional annotations. Homology searches were carried out by query of the NCBI nr protein database using the Blastx algorithm (E-value $<10^{-5}$) [23]. After nr annotation, we used the Blast2GO program [24] to obtain Gene Ontology (GO) annotations, and WEGO software [25] was used to perform GO functional classification of all unigenes in order to understand the distribution of gene functions at the macro level.

Using EC (Enzyme Commission number) terms, biochemical pathway information was generated by downloading relevant maps from the KEGG database [26]. This database contains systematic analysis of inner-cell metabolic pathways and functions of individual gene products. Here we identified the biological pathways that were active in *E. sinensis* ASG and assessed up or down regulation of key genes involved in the relevant pathways. After obtaining the KEGG pathway annotations, unigenes were aligned to the COG database to predict and classify potential functions based on known orthologous gene products. Every protein in COG is assumed to evolve from an ancestor protein, and the whole database is built on coding proteins with complete genomes as well as systematic evolutionary relationships among bacteria, algae and eukaryotic organisms [27].

Unigene Expression Difference Analysis

Unigene expression was calculated using the reads per kb per million reads method (RPKM), for which the formula is shown below:

$$\text{RPKM} = \frac{10^6 C}{N L / 10^3}$$

Where RPKM is the expression of unigene A, and C is the number of reads that uniquely aligned to unigene A. N is the total number of reads that uniquely aligned to all unigenes, and L is the base number in the CDS of unigene A. The RPKM method is able to eliminate the influence of different gene length and sequencing level on the calculation of gene expression. Therefore the calculated gene expression level can be used directly for comparing difference in gene expression between samples [28].

Data Deposition

*De novo* assembly sequence data from *E. sinensis* were deposited in the National Center for Biotechnology Information (NCBI, USA, http://www.ncbi.nlm.nih.gov/), while *de novo* assembly of sequence data from the ASG and testis in *E. sinensis* were deposited in the Transcriptome Shotgun Assembly (TSA) database with accession numbers KA660105–KA728674.

Results

General Features of the ASG Transcriptome in *E. sinensis*

Illumina high-throughput second generation sequencing produced 33,221,284 clean reads representing a total of 2,657,702,720 (2.66 Gb) nucleotides. Average read size, Q20 percentage and GC content were 90 bp, 91.06%, and 55.19%, respectively. From these short reads, 85,913 contigs were assembled, with a median length of 218 bp. From the contigs, 58,567 scaffolds were constructed using SOAPdenovo, with a median length of 292 bp, and 37,955 unigenes were obtained with a median length of 385 bp (Table 1). The quality of Illumina short read sequence assemblies results are shown in Figure 2.

Unigene Annotation and GO Assignment

Functional annotation consisted of protein functional annotation, pathway annotation, GO assignments and COG functional annotation. Distinct gene sequence analysis identified 27,541 unigene annotations (37.2% of all unigenes) above the preset cutoff value; similarly, 6,350 (8.6%) unigenes were annotated via ESTscan analysis. Based on similarity searches with known proteins, 33,891 unigenes were annotated based on having a Blast hit in the nr database or ESTscan results (Table S1). Since no genome or EST information existed previously for *Eriocheir* species, 54.2% of the unigenes could not be matched to known genes,
though it is likely that many of the genes of unknown function and/or unknown protein product would share common functions with known genes within the same cluster in the GO clustering analysis. Annotation analysis was used to provide information on gene expression and functional annotation of all unigenes from ASG and testis from *E. sinensis* resulted in 74,049 distinct events (Table 1). This number does not necessarily reflect the real transcriptome complexity, as many of the assembled sequences may represent distinct non-overlapping regions of the same transcripts. Thus, the final number of unique transcripts covered by our data would probably be lower.

**Table 1. Summary of transcriptomes from the accessory sex gland (ASG) and testis (T) in *E. sinensis***

|                  | ASG    | T      | ASG & T |
|------------------|--------|--------|---------|
| Total Reads      | 33,221,284 | 25,698,778 | —       |
| Total base pair(bp) | 2,657,702,720 | 2,312,890,020 | —       |
| Q20 percentage   | 91.06% | 91.3%  | —       |
| N percentage     | 0.15%  | 0.01%  | —       |
| GC percentage    | 55.19% | 49.17% | —       |
| Total number of contigs | 85,913 | 342,753 | —       |
| Mean length of contigs (bp) | 218 | 191 | —       |
| Total number of scaffolds | 58,567 | 141,861 | —       |
| Mean length of scaffolds (bp) | 292 | 300 | —       |
| The number of unigenes | 37,955 | 96311 | —       |
| Mean length of unigenes | 385 | 382 | —       |
| The number of all-unigenes | — | — | 74049 |
| Mean length of all-unigenes | — | — | 512 |

doi:10.1371/journal.pone.0053915.t001

GO assignments were used to classify the functions of the predicted genes. Based on sequence homology, sequences can be categorized into 43 functional groups; the best hits from this query were extracted for GO classification using Uniprot2GO; each sequence was assigned at least one GO term. Second-level GO terms were used to classify the sequences in terms of their involvement as cellular components, in molecular functions, and in biological processes (Fig. 3). In total, 44,144 unigenes were clustered in three assignments; 15,261 were categorized as “Cellular Component” (34.6%), 21,745 as “Biological Process” (49.3%) and 7,138 as “Molecular Function” (16.2%).

Using COG functional annotation, 27,657 unigenes were assigned into 25 function classes. 4,062 unigenes (14.7%) were assigned into “General function prediction only”, 3,030 unigenes were annotated into “Translation, ribosomal structure and biogenesis” and 2,524 unigenes were related to “Transcription”. The most abundantly represented biological processes were “Cell wall/membrane/envelope biogenesis”, “Cell cycle control, cell division, chromosome partitioning” and “Replication, recombination and repair” which comprised 2,171, 1,790 and 1,788 unigenes respectively, of the biological process sequences (Fig. 4).

**KEGG Pathway Assignment**

We mapped the 17,645 annotated sequences to the reference canonical pathways in the KEGG database to identify the biological pathways involved. A total of 17,645 unigenes were associated with 223 predicted KEGG metabolic pathways, and the number of different expressed genes (DEG) with pathway annotation was 11,962 (Table S3). The top two most prominent pathways (metabolic pathways and regulation of actin cytoskeleton) included over 1,510 unigenes. The most important pathways that may be relevant to spermatogenesis or reproduction included regulation of actin cytoskeleton (1,146 unigenes), DNA replication (90 unigenes), spliceosome (1,007 unigenes), RNA polymerase (234 unigenes),...
unigenes), Mismatch repair (60 unigenes), purine metabolism (468 unigenes), adherens junction (769 unigenes), cell cycle (272 unigenes), Fc gamma R-mediated phagocytosis (708 unigenes), pyrimidine metabolism (357 unigenes) and other anti-hyperthermia stress and anti-oxidative stress pathways or gene families such as the proteasome (53 unigenes). The top 30 pathways with highest DEGs genes are shown in Table 2. These predicted pathways are likely to be useful in future investigations that focus on their functions in *E. sinensis*. Using KEGG, 1,704 unigenes (14.99%) were included in basic metabolism process specific pathways; most of these were involved in carbohydrate, energy, and amino acid metabolism.

**Tissue-specific Analysis for Differentially Expressed Genes**

With regard to tissue specific analysis of differentially regulated genes, numerous genes crucial for reproduction and development were identified, including fertilin, serine proteinase inhibitor, Sperm antigen P26h, Sperm protamine and bovine seminal plasma protein BSP (Table 3). Identification of these essential genes and their regulatory mechanisms provided new understanding about the complex processes of reproduction and development. We believe information gained about these genes in *E. Sinensis* can be applied to this species to improve industrial aquaculture.

Here, we investigated differentially expressed genes identified in our transcriptome analysis of ASG and testis tissues in *E. sinensis*. Comparison of gene expression using DEGseq produced a total of 68,412 unigenes expressed in the testis at a significantly higher level than that in the ASG, and 5,174 unigenes were down regulated in testis compared with the ASG. On the other hand, of all the unigenes identified, 26,653 unigenes were expressed in the testis, but not expressed in the ASG, and 631 unigenes were expressed in the ASG, but not in the testis (Fig. 5 and Table S2).

**Discussion**

Descriptive and quantitative transcriptome analyses are important for interpreting the functional elements of the genome and revealing the molecular constituents of cells and tissues. It is known that sperm function can be affected by ASG proteins, including the processes of capacitation and the acrosome reaction, as well as sperm motility, DNA integrity and interaction with the oocyte. Here we identified many ASG secreted proteins that function in the modification of sperm and in sperm maturation (Table 3), including P26h (L-xylulose reductase, unigene 4,288), BSPs (bovine seminal plasma protein, unigene 64,588 and unigene 69,768), fertilin (unigene 17,270 and unigene 27,136), ACE (Angiotensin converting enzyme, unigene 7,164 and unigene 21,069), GPX5 (glutathione peroxidase, epididymal secretory glutathione peroxidase, unigene15860), Spermadhesin-1 (Acidic seminal fluid protein, aSFP). The reproduction-related transcripts identified in the ASG and testis transcriptomes in *E. sinensis*, with a special focus on the process of sperm transit through the ASG...
and the proteins involved in sperm membrane modification will be discussed in detail in the following section.

Proteins Involved in the Acrosome Reaction and Sperm-oocyte Interaction

P26h (L-xylulose reductase) catalyzes the NADPH-dependent reduction of several pentoses, tetroses, trioses, alpha-dicarbonyl compounds and L-xylulose. Functionally, P26h is involved in sperm–oocyte binding and its presence on sperm is an absolute prerequisite for fertilization [29]. Here we identified that Unigene4288 annotated as L-xylulose reductase (gi|229365856|gb|ACQ57908.1| Anoplopoma fimbria), was expressed equally in the ASG (RPKM 27.5693) and testis (RPKM 19.8661). During epididymal transit, P26h accumulates on the acrosomal cap of spermatozoa. Moreover, P26h is found in epididymosomes and becomes GPI-anchored to the sperm surface of the acrosomal region during epididymal transit, via an as yet unknown mechanism. Similarly, PH-20 (Sperm adhesion molecule 1, SPAM1) is a glycoprotein synthesized by the principal cells that associates with epididymosomes [27]. PH-20 is located on the sperm surface and in the acrosome, where it is bound to the inner acrosomal membrane. PH-20 is a multifunctional protein which can serve as a hyaluronidase, a receptor for HA-induced cell signaling, and a receptor for ZP binding [30].

In the bull (Bos taurus), the seminal plasma contains a group of four closely related acidic proteins called Bovine seminal plasma protein (BSP) BSP-A1, BSP-A2, BSP-A5, and BSP-30-kDa that bind to sperm plasma membranes after ejaculation by specific interaction with phospholipids [31]. Here we identified two BSP unigenes (Unigene64588, Unigene69768) that were only expressed in testis (RPKM 6.192 and 4.7583 respectively). The BSP-A1 and BSP-A2 mixture referred to as PDC-109, constitutes the major protein fraction in bovine seminal plasma and contains two tandem repeat fibronectin type-II (Fn II) domains, each of which can bind to a choline phospholipid on the sperm plasma membrane by its specific interaction with the phosphocholine headgroup [32]. This interaction of PDC-109 with the sperm cell membrane results in an efflux of cholesterol and choline phospholipids, that appears to be important for capacitation.

The main changes in spermatozoa that occur during epididymal maturation are the ability to move, recognize and bind to the ZP, and to fuse with the plasma membrane of the oocyte. The cellular processes responsible for these new properties of the sperm are probably related to changes in the surface of the plasma membrane itself [33]. In all species studied to date, it appears that specific testicular sperm surface proteins are removed or processed further as gametes pass through the epididymis [34]. Disappearance of some of these proteins is clearly related to a specific proteolytic mechanism during epididymal transit. For most proteins, proteolysis induces either a change in their membrane domain distribution, as has been shown for fertilin/PH30, or a release of a cleaved protein in the epididymal medium, as is the case for ACE.

Figure 4. COG classification of all unigenes from ASG and testis in E. sinensis.
doi:10.1371/journal.pone.0053915.g004
Among spermatozoa surface proteins, fertilin, a heterodimeric complex composed of two integral membrane glycoproteins named α-fertilin (ADAM-1) and β-fertilin (ADAM-2), as well as several other ADAMs have been reported to be involved in sperm-oocyte recognition and in membrane fusion [35]. Here we identified five unigenes (Unigene17270, 18613, 62731, 71804, and 27136) annotated as fertilin α subunits but we did not identify β subunit in our annotation results. These unigenes all showed significantly higher expression in testis (shown in table 3). The fertilin αβ complex shares traits with certain viral adhesion/fusion proteins, notably the presence of a candidate fusion peptide [36]. Both proteins are members of the ADAM (a disintegrin and metalloprotease) domain protein family with sequences containing a pro-domain, a metalloprotease, a disintegrin and a cysteine-rich domain, EGF-like repeats, a transmembrane domain and a carboxy-terminal cytosolic tail [37]. The β subunit is present as a full length protein on the testicular sperm surface and is proteolytically transformed during the passage of spermatozoa through the caput [38], and cleaved into a 35 kDa form in spermatozoa [39]. This proteolytic processing results in the removal of the pro- and metalloprotease-like domains, with only the full or part of the disintegrin domain, the cysteine-rich domain, the EGF repeat, the transmembrane and the cytoplasmic domains remaining on the sperm cell. This processing also induces a relocation of the fertilin complex to a different plasma membrane domain on the mature spermatozoa [40].

Proteins Associated with Sperm Motility

Little is known about the impact of ASG secretions on sperm motility. Semenogelins proteins are mainly synthesized in the seminal vesicles and are believed to have an inhibitory effect on the ability of sperm to move [41]. In contrast, another vesicle product, fructose, has been reported to be the main source of energy for spermatozoa [42]. Enzymes in the polyol pathway, including aldose reductase and sorbitol dehydrogenase, have been identified in epididymosomes [43] and appear to be involved in a mechanism for modulating sperm motility during epididymal transit. Patel et al. demonstrated a positive correlation between

Table 2. All unigenes KEGG metabolic pathway analysis in E. sinensis.

| No. | Pathway                                      | DEGs genes with pathway annotation (11,962) | All genes with pathway annotation (17,645) | P value | Q value   | Pathway ID |
|-----|----------------------------------------------|---------------------------------------------|---------------------------------------------|---------|-----------|------------|
| 1   | Vibrio cholerae infection                    | 657 (5.49%)                                 | 851 (4.82%)                                 | 3.70E-10| 8.33E-08  | ko05110    |
| 2   | Phototransduction                            | 505 (4.22%)                                 | 659 (3.73%)                                 | 2.36E-07| 2.66E-05  | ko04744    |
| 3   | Olfactory transduction                       | 506 (4.23%)                                 | 666 (3.77%)                                 | 1.47E-06| 1.10E-04  | ko04740    |
| 4   | DNA replication                              | 80 (0.67%)                                  | 90 (0.51%)                                  | 2.73E-06| 1.54E-04  | ko03030    |
| 5   | Pyrimidine metabolism                        | 279 (2.33%)                                 | 357 (2.02%)                                 | 8.20E-06| 3.69E-04  | ko00240    |
| 6   | Amoebiasis                                   | 710 (5.94%)                                 | 962 (5.45%)                                 | 1.77E-05| 6.66E-04  | ko05146    |
| 7   | Spliceosome                                  | 739 (6.18%)                                 | 1007 (5.71%)                                | 4.16E-05| 1.22E-03  | ko03040    |
| 8   | RNA polymerase                               | 186 (1.55%)                                 | 234 (1.33%)                                 | 4.35E-05| 1.22E-03  | ko03020    |
| 9   | Amyotrophic lateral sclerosis (ALS)          | 217 (1.81%)                                 | 277 (1.57%)                                 | 6.09E-05| 1.52E-03  | ko05014    |
| 10  | Homologous recombination                     | 46 (0.38%)                                  | 51 (0.29%)                                  | 0.000174216| 3.92E-03  | ko03440    |
| 11  | Mismatch repair                              | 53 (0.44%)                                  | 60 (0.34%)                                  | 0.000207332| 4.24E-03  | ko03430    |
| 12  | Purine metabolism                            | 349 (2.92%)                                 | 468 (2.65%)                                 | 0.000714386| 1.34E-02  | ko00230    |
| 13  | Base excision repair                         | 72 (0.6%)                                   | 87 (0.49%)                                  | 0.0001272401| 2.20E-02  | ko03410    |
| 14  | Nucleotide excision repair                   | 89 (0.74%)                                  | 110 (0.62%)                                 | 0.000151106| 2.43E-02  | ko03420    |
| 15  | Regulation of actin cytoskeleton             | 821 (6.86%)                                 | 1146 (6.49%)                                | 0.000201154| 3.02E-02  | ko04810    |
| 16  | Pathogenic Escherichia coli infection        | 455 (3.83%)                                 | 624 (3.54%)                                 | 0.000271319| 3.82E-02  | ko05130    |
| 17  | Neuroactive ligand-receptor interaction      | 183 (1.53%)                                 | 241 (1.37%)                                 | 0.003346124| 4.43E-02  | ko04080    |
| 18  | Proteasome                                   | 45 (0.38%)                                  | 53 (0.3%)                                   | 0.000390126| 4.88E-02  | ko03050    |
| 19  | Cardiac muscle contraction                   | 167 (1.4%)                                  | 220 (1.25%)                                 | 0.000501368| 5.96E-02  | ko04260    |
| 20  | Adherens junction                            | 552 (4.61%)                                 | 769 (4.36%)                                 | 0.000813791| 9.15E-02  | ko04520    |
| 21  | Bacterial invasion of epithelial cells       | 477 (3.99%)                                 | 662 (3.75%)                                 | 0.008830244| 9.46E-02  | ko05100    |
| 22  | Fc gamma R-mediated phagocytosis             | 508 (4.25%)                                 | 708 (4.01%)                                 | 0.01130496| 1.16E-01  | ko04666    |
| 23  | Shigellosis                                  | 477 (3.99%)                                 | 666 (3.77%)                                 | 0.01654301| 1.62E-01  | ko05131    |
| 24  | Cell cycle                                   | 200 (1.67%)                                 | 272 (1.54%)                                 | 0.02275106| 2.13E-01  | ko04110    |
| 25  | Viral myocarditis                            | 150 (1.25%)                                 | 202 (1.14%)                                 | 0.02689615| 2.42E-01  | ko05416    |
| 26  | Huntington’s disease                         | 404 (3.38%)                                 | 566 (3.21%)                                 | 0.03420199| 2.96E-01  | ko05016    |
| 27  | SNARE interactions in vesicular transport    | 36 (0.3%)                                   | 45 (0.26%)                                  | 0.0512175| 4.12E-01  | ko04130    |
| 28  | Staphylococcus aureus infection              | 36 (0.3%)                                   | 45 (0.26%)                                  | 0.0512175| 4.12E-01  | ko05150    |
| 29  | Dorso-ventral axis formation                 | 225 (1.88%)                                 | 312 (1.77%)                                 | 0.05467221| 4.26E-01  | ko04320    |
| 30  | Riboflavin metabolism                       | 25 (0.21%)                                  | 31 (0.18%)                                  | 0.08619657| 6.46E-01  | ko00740    |

doi:10.1371/journal.pone.0053915.t002
Table 3. The reproduction-related unigenes identified in the accessory sex gland (ASG) and testis (T) transcriptomes during the sexual maturation stage in *E. sinensis*.

| Unigene No. | Unigene name of top BLASTX hit (Accession no; species) | Length (bp) | E-value | ASG RPKM | Testis RPKM | Log2(Testis RPKM/ASG RPKM) |
|-------------|--------------------------------------------------------|-------------|---------|----------|-------------|-----------------------------|
| **Sperm antigen P26h (L-xylulose reductase)** | | | | | | |
| Unigene4288_All | L-xylulose reductase (gi|229365856|gb|ACQ57908.1;|Anoplopoma fimбриa) | 989 | 5.00E-82 | 27.5693 | 19.8661 | −0.4728 |
| **BSP (bovine seminal plasma protein)** | | | | | | |
| Unigene64588_All | surface antigen BspA-like (gi|123302396|ref|XP_001291104.1;|Trichomonas vaginalis G3) | 384 | 8.00E-06 | 0 | 6.192 | 12.5962 |
| Unigene69768_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 625 | 5.00E-06 | 0 | 16.619 | 14.0206 |
| **Fertilin** | | | | | | |
| Unigene17270_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 625 | 5.00E-06 | 0 | 16.619 | 14.0206 |
| Unigene18613_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 580 | 9.00E-09 | 0.8666 | 8.1991 | 3.242 |
| Unigene62731_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 356 | 2.00E-09 | 0 | 9.4912 | 13.2124 |
| Unigene71804_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 356 | 2.00E-09 | 0 | 9.4912 | 13.2124 |
| Unigene27136_All | similar to fertilin alpha-I (gi|126324185|ref|XP_001371111.1;|Monodelphis domestica) | 356 | 2.00E-09 | 0 | 9.4912 | 13.2124 |
| **Immunoglobulin** | | | | | | |
| Unigene2775_All | leucine-rich repeats and immunoglobulin-like domains 2-like (gi|291228204|ref|XP_002734069.1;|Saccoglossus kowalevskii) | 1394 | 6.00E-06 | 0 | 4.0398 | 22.0699 |
| Unigene21708_All | leucine-rich repeats and immunoglobulin-like domains protein 3 (gi|296471357|gb|EFN85034.1;|Harpegnathos saltator) | 1627 | 6.00E-09 | 0 | 25.2287 | 14.6228 |
| Unigene24011_All | immunoglobulin superfamily DCC subclass member 4 (gi|292616070|ref|XP_002662886.1;|Danio rerio) | 512 | 4.00E-09 | 3.9266 | 10.510 | 1.4204 |
| Unigene28499_All | immunoglobulin mu binding protein 2 (gi|296471357|gb|EFN85034.1;|Harpegnathos saltator) | 442 | 2.00E-03 | 0.5886 | 6.2289 | 3.4535 |
| Unigene67254_All | leucine-rich repeats and immunoglobulin-like domains protein 3 (gi|296471357|gb|EFN85034.1;|Harpegnathos saltator) | 440 | 9.00E-21 | 0.2856 | 13.0832 | 5.5176 |
| Unigene67826_All | leucine-rich repeats and immunoglobulin-like domains protein 3 (gi|296471357|gb|EFN85034.1;|Harpegnathos saltator) | 455 | 1.00E-11 | 0 | 39.6059 | 15.2734 |
| **Serine proteinase inhibitor** | | | | | | |
| Unigene12520_All | serine proteinase inhibitor 6 (gi|288188852|gb|ADC42876.1;|Penaeus monodon) | 796 | 4.00E-36 | 0.3157 | 0.4716 | 0.579 |
| Unigene13284_All | serine proteinase inhibitor 6 (gi|288188852|gb|ADC42876.1;|Penaeus monodon) | 441 | 1.00E-16 | 57.8388 | 33.769 | −0.7763 |
| Unigene13350_All | serine proteinase inhibitor (gi|33590491|gb|AAQ22771.1;|Procambarus clarkii) | 653 | 3.00E-38 | 15.2011 | 0.9582 | −3.9877 |
| Unigene14659_All | serine proteinase inhibitor (gi|33590491|gb|AAQ22771.1;|Procambarus clarkii) | 327 | 4.00E−18 | 3.8425 | 11.6838 | 1.6264 |
| Unigene1959_All | serine proteinase inhibitor 6 (gi|288188852|gb|ADC42876.1;|Penaeus monodon) | 1229 | 3.00E-79 | 5.6231 | 10.3862 | 0.8852 |
| Unigene24440_All | serine proteinase inhibitor (gi|33590491|gb|AAQ22771.1;|Procambarus clarkii) | 654 | 2.00E-21 | 0.9606 | 23.5362 | 4.6148 |
| Unigene31267_All | serine proteinase inhibitor (gi|33590491|gb|AAQ22771.1;|Procambarus clarkii) | 659 | 4.00E−26 | 0.3813 | 12.3435 | 5.0167 |
| Unigene32663_All | Kazal-type serine proteinase inhibitor 1 (gi|219809644|gb|ACL36280.1;|Fenneropenaeus chinensis) | 1828 | 2.00E-04 | 0.6186 | 28.616 | 5.5317 |
| Unigene3605_All | serine proteinase inhibitor (gi|33590491|gb|AAQ22771.1;|Procambarus clarkii) | 531 | 2.00E-16 | 23.3501 | 10.8411 | −0.9085 |
| Unigene5046_All | serine proteinase inhibitor 6 (gi|288188852|gb|ADC42876.1;|Penaeus monodon) | 316 | 7.00E−11 | 338.7773 | 116.0352 | −1.5458 |
| Unigene67995_All | serine proteinase inhibitor 6 (gi|288188852|gb|ADC42876.1;|Penaeus monodon) | 459 | 8.00E-17 | 19.436 | 12.5416 | −0.632 |
| Unigene No. | Unigene name of top BLASTX hit (Accession no; species) | Length (bp) | E-value | ASG RPKM | Testis RPKM | Log2(Testis RPKM/ASG RPKM) |
|-------------|-----------------------------------------------------|------------|----------|-----------|-------------|-----------------------------|
| Unigene43562_All | trypsin-like serine protease (gi|26085386|gb|ACG55893.1;Eriocheir sinensis) | 233 | 3.00E-39 | 2.6963 | 2.6855 | 2.0058 |
| Unigene49929_All | trypsin-like serine protease (gi|124188546|gb|ACG55893.1;Eriocheir sinensis) | 261 | 3.00E-07 | 0 | 7.6716 | 7.1375 |
| Unigene50348_All | trypsin-like serine protease (gi|26085386|gb|ACG55893.1;Eriocheir sinensis) | 263 | 6.00E-25 | 1.4333 | 7.1375 | 2.3161 |
| Unigene7579_All | sorbitol dehydrogenase (gi|3580322|gb|ABN13876.1;Locusta migratoria manilensis) | 950 | 1.00E-113 | 12.0359 | 8.6942 | 2.0046 |
| Unigene7617_All | Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial (gi|160276311|gb|EFN70179.1;Camponotus floridanus) | 422 | 2.00E-33 | 5.0617 | 16.0137 | 1.6616 |
| Unigene7691_All | D-beta-hydroxybutyrate dehydrogenase, mitochondrial (gi|147899736|gb|NP_001082978.1;Danio rerio) | 667 | 5.00E-50 | 21.287 | 8.8182 | 2.1714 |
| Unigene7769_All | Zinc-type alcohol dehydrogenase-like protein C1773.06c (gi|307174541|gb|EFN64991.1;Camponotus floridanus) | 4041 | 3.00E-99 | 8.0222 | 11.6751 | 0.5414 |
| Unigene8005_All | glucose-6-phosphate dehydrogenase isoform B (gi|61394184|gb|AAX45785.1;Ips typographus) | 1730 | 0 | 14.526 | 24.9564 | 0.7808 |
| Unigene8007_All | aldehyde dehydrogenase family 6, subfamily A1 (gi|224051481|gb|XP_002199925.1;Taeniopygia guttata) | 2071 | 0 | 0 | 40.5889 | 49.7312 |
| Unigene8101_All | isovaleryl coenzyme A dehydrogenase (gi|209571446|ref|NP_001129356.1;Bombyx mori) | 625 | 6.00E-95 | 6.8353 | 7.4085 | 0.1162 |
| Unigene8139_All | glyceraldehyde 3-phosphate dehydrogenase (gi|296785436|gb|ADH43624.1;Eriocheir sinensis) | 1401 | 0 | 204.8423 | 351.2243 | 0.7779 |
| Unigene8170_All | NADH dehydrogenase (ubiquinone) Fe-S protein 1 isoform 1 and 2 (gi|72133227|ref|XP_780124.1;Strongylocentrotus purpuratus) | 2322 | 0 | 15.0975 | 17.5697 | 0.2188 |
| Unigene819_All | NADH dehydrogenase flavoprotein 2, mitochondrial (gi|189237290|ref|NP_001082978.1;Camponotus floridanus) | 512 | 2.00E-71 | 11.288 | 27.8644 | 1.3035 |
| Unigene8242_All | PREDICTED: similar to isocitrate dehydrogenase (gi|189235505|ref|XP_974070.2;Tribolium castaneum) | 786 | 1.00E-101 | 5.9148 | 15.7624 | 1.4141 |
| Unigene8243_All | 15-hydroxyprostaglandin dehydrogenase [NAD[P] (gi|307184287|gb|EFN70745.1;Camponotus floridanus) | 1168 | 6.00E-52 | 5.4864 | 12.1072 | 0.2931 |
| Unigene8353_All | NADH dehydrogenase subunit 1 (gi|63025123|ref|YP_232831.1;Eriocheir sinensis) | 685 | 1.00E-96 | 6.00E-05 | 6.6343 | 7.5122 |
| Unigene8999_All | hydroxyacyl dehydrogenase (gi|157122882|ref|XP_001659938.1;Aedes aegypti) | 818 | 5.00E-64 | 5.5298 | 7.3434 | 0.4692 |
| Unigene64835_All | PREDICTED: similar to glycosyl-phosphatidyl inositol-specific phospholipidase C (gi|91088447|ref|XP_968769.1;Tribolium castaneum) | 388 | 1.00E-18 | 0 | 11.288 | 13.4626 |
| Unigene71361_All | Glycosyl-phosphatidyl inositol anchor attachment 1 protein (gi|307188892|gb|EFN73441.1;Camponotus floridanus) | 621 | 7.00E-24 | 3.8444 | 11.899 | 16.098 |
| Unigene67401_All | ligand-independent activating molecule for estrogen receptor-like (gi|293351305|ref|NP_002727750.1;Rattus norvegicus) | 444 | 5.00E-06 | 0 | 7.00E-02 | 2.3697 |
| Unigene7117_All | breast cancer anti-estrogen resistance protein (gi|107187906|gb|NP_001082978.1;Camponotus floridanus) | 3294 | 1.00E-37 | 1.1449 | 7.5178 | 1.4499 |
| Unigene73158_All | estrogen sulfotransferase-like (gi|110764250|ref|XP_394850.3;Apis mellifera) | 308 | 1.00E-18 | 3.8444 | 11.899 | 16.098 |
| Unigene73891_All | estrogen sulfotransferase-like (gi|110764250|ref|XP_394850.3;Apis mellifera) | 1682 | 3.00E-47 | 4.2726 | 11.049 |
| Unigene74177_All | estrogen sulfotransferase-like (gi|110764250|ref|XP_394850.3;Apis mellifera) | 631 | 1.00E-37 | 1.1449 | 7.5178 | 1.4499 |
| Unigene74677_All | estrogen sulfotransferase-like (gi|110764250|ref|XP_394850.3;Apis mellifera) | 439 | 1.00E-06 | 0 | 4.7915 | 12.2823 |
| Unigene75117_All | estrogen receptor binding protein (gi|351313510|ref|NP_002727750.1;Rattus norvegicus) | 444 | 5.00E-06 | 0 | 7.00E-02 | 2.3697 |
| Unigene75502_All | estrogen receptor binding protein (gi|351313510|ref|NP_002727750.1;Rattus norvegicus) | 3294 | 1.00E-37 | 1.1449 | 7.5178 | 1.4499 |
| Unigene11682_All | estrogen receptor binding protein (gi|351313510|ref|NP_002727750.1;Rattus norvegicus) | 903 | 1.00E-37 | 1.1449 | 7.5178 | 1.4499 |
Table 3. Cont.

| Unigene No.     | Unigene name of top BLASTX hit (Accession no; species)                                                                 | Length (bp) | E-value        | ASG RPKM | Testis RPKM | Log2(Testis RPKM/ASG RPKM) |
|-----------------|------------------------------------------------------------------------------------------------------------------------|-------------|----------------|----------|-------------|-----------------------------|
| Unigene47381_All| estrogen-related receptor beta like 1-like (gi|291225239|ref|XP_002732609.1|Saccoglossus kowalevskii)   | 249        | 7.00E-18      | 0        | 7.0362  | 12.7806                     |
| Unigene55150_All| ras-related and estrogen-regulated growth inhibitor-like (gi|296210885|ref|XP_002752248.1|Callithrix jacchus)          | 288        | 5.00E-14      | 1.7451   | 3.4762  | 0.9942                      |
| Unigene59552_All| ras-related and estrogen-regulated growth inhibitor (gi|223649254|gb|ACN11385.1;Salmo salar) | 322        | 9.00E-16      | 3.9022   | 4.2751  | 0.1317                      |
| Unigene61508_All| estrogen-related receptor beta like 1-like (gi|291225239|ref|XP_002732609.1|Saccoglossus kowalevskii)   | 341        | 3.00E-20      | 7.0333   | 6.6058  | 3.3921                      |
| Epididymal secretory glutathione peroxidase | epididymal secretory glutathione peroxidase precursor (gi|47523090|ref|NP_999051.1|Sus scrofa) | 381 | 4.00E-13 | 0 | 4.27 | 12.056 |
| Unigene27168_All| glutathione peroxidase (gi|17118951|gb|ACB42236.1;Metapenaeus ensis) | 413 | 2.00E-39 | 2.7381 | 7.8783 | 1.5247 |
| Unigene27684_All| glutathione peroxidase 7 (gi|148236625|ref|NP_00108904.1|Xenopus laevis) | 879 | 6.00E-40 | 3.5737 | 19.6471 | 2.4588 |
| Unigene62805_All| selenium-dependent glutathione peroxidase (gi|222873557|gb|ACM68948.1;Macrobrachium rosenbergii) | 357 | 7.00E-29 | 2.4637 | 2.1033 | 0.2282 |
| Unigene63410_All| phospholipid-hydroperoxide glutathione peroxidase (gi|164608818|gb|ABY62740.1|Artemia franciscana) | 365 | 2.00E-49 | 11.0159 | 128.2295 | 3.5411 |
| Unigene65933_All| glutathione peroxidase (gi|17118951|gb|ACB42236.1;Metapenaeus ensis) | 409 | 3.00E-12 | 0 | 7.9553 | 12.9577 |
| Unigene7003_All| phospholipid-hydroperoxide glutathione peroxidase (gi|164608818|gb|ABY62740.1|Artemia franciscana) | 459 | 9.00E-29 | 7.3912 | 94.8802 | 3.6822 |
| Unigene7010_All| dehydrogenase/reductase SDR family member 12-like (gi|292611020|ref|XP_002660947.1;Danio rerio) | 663 | 2.00E-64 | 70.3108 | 41.1483 | 2.2771 |
| Unigene70401_All| FAD-dependent oxidoreductase domain containing 1 (gi|156717942|ref|NP_001096513.1;Xenopus (Silurana) tropicalis) | 561 | 9.00E-48 | 0 | 13.6074 | 13.3721 |
| Unigene71193_All| NADH-Ubiquinone oxidoreductase AGGG subunit (gi|242017690|ref|XP_002429320.1;Pediculus humanus corporis) | 614 | 2.00E-16 | 3.0696 | 14.8787 | 2.2771 |
| Sperm protamine | sperm protamine P1(sp|P83321|HSP1_MURBR; Murex brandaris) | 239 | 7.00E-07 | 3.6080 | 14.1376 | 1.9417 |
| Unigene7164_All| angiotensin converting enzyme (ACE) (gi|224028155|emb|CAX48990.1;Pontastacus leptodactylus) | 3013 | 2.00E-16 | 6.1303 | 4.4027 | 0.2776 |
| Unigene50101_All| angiotensin converting enzyme (gi|224028155|emb|CAX48990.1;Pontastacus leptodactylus) | 329 | 2.00E-18 | 2.6734 | 5.3253 | 0.9942 |
FDR ≤ 0.001 AND |log2Ratio| ≥ 1

- up-regulated genes
- down-regulated genes
- Not DEGs

Figure 5. DEGs analysis of unigenes in the ASG and testis from E. sinensis.
doi:10.1371/journal.pone.0053915.g005

...semenal levels of fructose and the relative proportion of motile sperm [44], but other studies could not find such a correlation.
Prostate-specific antigen has been reported to be involved in degradation of semenogelins and may therefore be expected to have a positive impact on sperm motility.

The Serpin (serine proteinase inhibitor) family is exclusively expressed in the rat cauda epididymis and up-regulated by androgens, and is secreted into the lumen to cover the sperm head [45]. Zhao et al. identified a Serpin family protein (As_SRPI) that is secreted from spermatids during nematode Ascaris suum spermiogenesis (also called sperm activation) and showed that As_SRPI has two major functions. First, As_SRPI functions in cis to support major sperm protein-based cytoskeletal assembly in the spermatid that releases it, thereby facilitating sperm motility acquisition. Second, As_SRPI released from activated sperm inhibits in trans the activation of surrounding spermatids by inhibiting vas deferens-derived AsTRY-5, a trypsin-like serine protease necessary for sperm activation. Here we identified eleven unigenes, including: unigene12520, 13284, 13350, 14659, 1959, 24440, 31267, 32663, 3605, 5046 and 67995 that were annotated as serine proteinase inhibitors, which were differently expressed in ASG and testis (Table 3). On the other hand, vesicular exocytosis is necessary to create fertilization-competent sperm in many animal species, components released during this process could be more important modulators of the physiology and behavior of surrounding sperm than was previously appreciated [46].

Another factor that is implicated in the process of semen viscosity is zinc, which primarily originates from the prostate. This metal may be crucial for modulation of the three-dimensional structure of Sgl and SgII rendering them more susceptible to proteolytic breakdown by seminal proteases [47]. Additionally, immunoglobulin G, which is a luminal protein in the epididymis, was present only in the epididymal fluid. Caveolin-1, previously found in prostasomes, which are membranous vesicles similar to epididymosomes, has also been detected in epididymal vesicles. Here we identified four unigenes annotated as Sperm protamine P1 (Unigene13877, 23278, 29647 and 59678) with higher expression in testis, as sperm nuclear proteins, specifically protamine 2 which is a zinc-finger protein [48]. Interestingly, Zinc binding to the sperm nucleus varies proportionately with the zinc content of protamine 2 in sperm chromatin [49]. A previous report indicated that an abnormally high contribution of seminal vesicular fluid to sperm-rich fractions of the ejaculate creates a risk of depleting chromatin zinc and thereby impairing zinc-dependent chromatin stability [50]. Some of the enzymes important for the function of sperm are zinc metallo-enzymes and can thus become dysfunctional when zinc is deficient. One of these, sorbitol dehydrogenase (SoDH), utilizes sorbitol to provide sperm with fructose for energy, so that SoDH activity is correlated with sperm motility. Similarly, lactate dehydrogenase-X, another zinc metallo-enzyme, has also been reported to have some relationship with sperm motility [9]. To our knowledge, this is the first presentation of strong evidence for protamine gene expression in E. sinensis testis and ASG.

Proteins Involved in Protection of Sperm

We discussed the sperm protection mechanism in testis during spermatogenesis, but sperm have a long journey after leaving the testis and before it arrives at the oocyte for fertilization. Here we discuss the ASG proteins involved in protection of sperm during epididymal transition. GPX5 (Type 5 glutathione peroxidase, Epididymal secretory glutathione peroxidase) is a protein secreted by the caput epididymis in an epididymosome-associated form and is thought to be involved in protecting epididymal sperm against oxidative stress. Here we identified 10 epididymal secretory glutathione peroxidases (Table 3), most of them were higher expressed in testis and only unigene62805 and 7010 were slightly higher expressed in ASG. GPX5 protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxides, by glutathione. It may constitute a glutathione peroxidase-like protective system against peroxide damage in sperm membrane lipids [51]. MIF (Macrophage migration inhibitory factor) is a protein found in rat, human and bovine epididymis and epididymal sperm [52]. MIF has been localized within apical protrusions of epithelial cells, in epididymosomes and associated with sperm in the epididymal lumen, thereby supporting the hypothesis of apocrine secretion mediated protein transfer via epididymosomes.

Epididymosome Associated Transportation

Frenette and Sullivan proposed that the transfer of epididymal proteins to the sub-cellular compartments of the sperm is mediated by small membranous vesicles, known as epididymosomes [53]. Epididymosomes are electron dense vesicles secreted in an apocrine fashion that range between 50 and 500 nm in diameter. Proteins associated with epididymosomes are not processed through the endoplasmic reticulum and Golgi apparatus and are characterized by unusual glycosylation patterns. Epididymosomes are rich in cholesterol, with cholesterol: phospholipid ratios as high as 2:1, and have sphingomyelin as their major phospholipid. Epididymosomes contain lipid rafts, i.e. cholesterol and phospholipid-enriched microdomains [54]. These microdomains contain GPI-anchored and transmembrane proteins, as well as signaling molecules including protein tyrosine kinases, and may serve as a platform for transferring the proteins from the epididymal epithelium to a maturing sperm.

In vitro and in vivo studies have shown that these vesicles, which are present in the cauda epididymis and seminal plasma, transfer a number of proteins to sperm. Additionally, some of these proteins have been shown to be essential for sperm motility and fertility [55]. We observed two sizes of ASG vesicles referred to small and large vesicles that were thought to play a key role in E. sinensis similar to the described previously epididymosomes [20]. Furthermore, these vesicles, when observed under transmission and scanning electron microscopy, were thought to contain the enzymatic proteins or other activation factors required for
spERMatoSPhore rupture, that were released immediately during homogenate isolation and processing. We hypothesize that in a natural mating context environmental parameters, including pH or spermasthma-produced factors, may induce the slow release of the vesicle contained proteins or factors [20]. In crabs, the ASG is an important component of the male reproductive system that opens at the junction of the seminal vesicle and ejaculatory duct. Secretions from the ASG, along with spermastophores from the seminal vesicle and spermatheca, enter the female spermastheca through the ejaculatory duct during mating. In Brachyura, spermastophores are delivered into the spermastheca of the female during mating and gradually are broken down to release free sperm into the spermastheca, thus facilitating spermatozoa and egg fusion to complete fertilization [56]. Given this important process, we focused on the ASG functions of spermastrophe rupture and sperm maturation, in order to identify secreted proteins from the ASG that may be important in these processes.

Important Signaling Pathways in the Testis and the ASG

We listed the top 30 pathways in Table 2, showing the number of differently expressed genes and all genes with pathway annotations. In our analysis, classes of genes that maintain relatively steady-state levels of gene expression included those controlling tissue remodeling, immunoregulation, cell-cycle progression, apoptosis, and growth. Development of reproductive tissue is a dynamic process involving coordinated interactions between regulators that assemble or edit the cellular constituents that support developing gametes [12]. The regulation of actin cytoskeleton, proteasome, adherens junction, cell cycle and SNARE interactions in vesicular transport pathways were identified and are all thought to be involved in spermatogenesis and sperm maturation.

The central importance of cAMP and PKA in driving tyrosine phosphorylation events associated with capacitation is well established [57]. Interestingly, the key components of the MAPK signaling pathway including MAP kinas, ERK1/2, and MEK, which were identified in our dataset, are implicated in various aspects of capacitation in human spermatozoa [58]. It is thought that sperm cells may also have unique signaling pathways. For example, the small GTPases in the Rap family are important for many aspects of cytoplasmic signaling. In sperm cells, some complicated signaling cascades may be simplified. For example, mitogen-activated protein kinase (MAPK) cascades are central to many signaling pathways in animals, and there is often cross talk between different members in different signaling pathways [59].

Cell cycle transitions may be controlled by regulation of the ubiquitin carrier and cyclin ligase destruction machinery. To date, our lab has reported detailed cDNA expression of some components of the ubiquitin-proteasome involved in reproduction in E. sinensis, including Eo-UbS27, Eo-UbL40, Eo-SUMO, Eo-Aos1/Eo-Uba2 and Eo-Ubc9, that were widely observed in the testis and ovary [60,61]. We also identified the ubiquitin mediated



References

1. Qiu GF, Ramachandra RK, Rexroad Iii CE, Yao J (2008) Molecular characterization and expression profiles of cyclin B1, B2 and Cdc2 kinase during oogenesis and spermatogenesis in rainbow trout (Oncorhynchus mykiss). Animal Reproduction Science 105: 209–223.
2. Caballero J, Frenette G, Sullivan R (2011) Post Testicular Sperm Maturational Changes in the Bull: Important Role of the Epididymosomes and Prostasomes. Veterinary Medicine International 2011.
3. Moura AA, Chapman DA, Koc H, Killian GJ (2007) A comprehensive proteomic analysis of the accessory sex gland fluid from mature Holstein bulls. Animal Reproduction Science 98: 169–188.
4. Moura AA, Koc H, Chapman DA, Killian GJ (2006) Identification of Proteins in the Accessory Sex Gland Fluid Associated With Fertility Indexes of Dairy Bulls: A Proteomic Approach. J Androl 27: 201–211.
5. Frenette G, Lessard C, Sullivan R (2002) Selected Proteins of “Prostasome-Like Particles” from Epididymal Cauda Fluid Are Transferred to Epididymal Caput Spermatozoa in Bull. Biology of Reproduction 67: 308–313.
6. Moura AA, Chapman DA, Killian GJ (2007) Proteins of the accessory sex glands associated with the oocyte-penetrating capacity of cauda epididymal sperm from holstein bulls of documented fertility. Molecular Reproduction and Development 74: 214–222.
7. Gillott C (2003) Male Accessory Gland Secretions: Modulators of Female Reproductive Physiology and Behavior. Annual Review of Entomology 48: 163–184.

8. Sostaric E, Aalberts M, Gadella BM, Stout TAE (2008) The roles of the epimastigates and prosomas in the attainment of fertilizing capacity by stallion sperm. Animal Reproduction Science 107: 237–248.

9. Elzanaty S, Richthoff J, Malin, J, Giwercman A (2002) The impact of epidymal and accessory sex gland function on sperm motility. Human Reproduction 17: 143–151.

10. Heraught MA, Killian GJ, Kangawa JF, Grier LC (1995) Effect of accessory sex gland fluid from bull of differing fertilities on the ability of cauda epidymal sperm to penetrate zona-free bovine oocytes. Biology of Reproduction 52: 380–386.

11. Gadella BM, van Gestel RA (2004) Bicarbonate and its role in mammalian sperm function. Animal Reproduction Science 82: 307–319.

12. Sostaric E, Dielemans SJ, van de Les CH, Cohenbrander B, Vos PLAM, et al. (2008) Sperm binding properties and secretory activity of the bovine epididymis immediately before and after ovulation. Molecular Reproduction and Development 76: 70–74.

13. Wang HZ, Wang HJ, Liang XM, Cui YD (2006) Stocking models of Chinese mitten crab (Eriocheir japonicus sinensis) in Yangtze lakes. Aquaculture 255: 456–465.

14. Sui L, Wille M, Cheng Y, Sorensen P (2007) The effect of dietary n-3 HUFA levels and DHA/EPA ratios on growth, survival and osmotic stress tolerance of Chinese mitten crab Eriocheir japonicus larvae. Aquaculture 273: 139–150.

15. Herborg LM, Bentley MG, Clare AS, Last KS (2006) Mating behaviour and chemical communication in the invasive Chinese mitten crab Eriocheir sinensis. Journal of Experimental Marine Biology and Ecology 329: 1–10.

16. Rudnick DA, Hsib K, Grimmer FK, Resh VH (2003) Patterns and processes of biological invasion: The Chinese mitten crab in San Francisco Bay. Basic and Applied Ecology 4: 249–262.

17. Suzuki K, Sato Y, Li L, Ohta S, Utsunomiya T, et al. (2003) Developmental biology meets with reproductive engineering: interdisciplinary science area as a breakthrough. Cell Mol Biol (Noisy-le-Grand) 49: 633–660.

18. Hao DC, Ge G, Xiao P, Zhang Y, Yang L (2011) The First Insight into the Escherichia coli-specific Transposase via Illumina Second Generation Sequencing. PLoS ONE 6: e21220.

19. He L, Wang Q, Jin X, Wang Y, Chen L, et al. (2012) Transcriptional Profiling of Testis during Sexual Maturation Stages in Erischeir sinensis using Illumina Sequencing. PLoS ONE 7: e33735.

20. Hou XL, Mao Q, He L, Gong YN, Qu D, et al. (2010) Accessory Sex Gland Proteins Affect Spermatozoa Digestion Rate and Spermatozoa Acrosin Activity in Erischeir sinensis. Journal of Crustacean Biology. 455–440.

21. Li Z, Zha JQ, Yang WX (2010) Acrosome reaction in Octopus tchangkii induced by calcium ionophore A23187 and a possible role of the acrosomal screw. Journal of Crustacean Biology. 435–440.

22. Iseli C, Jongeneel CV, Bucher P (1999) ESTScan: A Program for Detecting, Evaluating, and Reconstructing Potential Coding Regions in EST Sequences. by calcium ionophore A23187 and a possible role of the acrosomal screw. Biology of Reproduction 52: 390–397.

23. Patel SM, Skarnhafn KF, Mehta YB (1988) Seminal plasma fructose and glucose levels and DHA/EPA ratios on growth, survival and osmotic stress tolerance of Chinese mitten crab Eriocheir japonicus larvae. Aquaculture 273: 139–150.

24. Zhou Y, Zhang M, Shi Q, Zhang L, Chen W, et al. (2008) An Epididymis-Specific Secretory Protein HongESI Critically Regulates Sperm Capacitation and Acrosome Reaction. Mol Fertil. PLoS 8: e11605.

25. Zhao Y, Sun W, Zhang P, Chi H, Zhang M-J, et al. (2012) Nematode sperm maturation triggered by protease involves sperm-secreted serine protease inhibitor (Serpin). Proceedings of the National Academy of Sciences 109: 10122–10127.

26. Li T, Liao H (1985) A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. J Clin Invest 76(3): 1899–1903.

27. Bal W, Dyba M, Szwarcz Z, Zejowska Bojeuk M, Lasko J, et al. (2001) Differential zinc and DNA binding by partial peptides of human prostate HP3. Molecular and Cellular Biochemistry 227: 97–106.

28. Bench G, Corzett MH, Kramer CE, Grant PG, Balhorn R (2000) Zinc is a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

29. Nixon B, Bielanowicz A, Anderson AL, Walsh A, Hall T, et al. (2010) Sperm binding properties and secretory activity of the bovine epididymis immediately before and after ovulation. Molecular Reproduction and Development 76: 70–74.

30. Oh J, Han C, Cho C (2009) ADAM7 is associated with epididymosomes and involved in sperm to penetrate zona-free bovine oocytes. Biology of Reproduction 52: 390–397.

31. Eickhoff R, Jennemann G, Schuring MP, Kaltner H, et al. (2006) Immunohistochemical Detection of Macrophage Migration Inhibitory Factor in Fetal and Adult Bovine Epididymis: Release by the Apocrine Secretion Mode. Cells Tissues Organs 182: 22–31.

32. Fabrega A, Guyonnet B, Dacheux JL, Gatti JL, Puigmule M, et al. (2011) Differential zinc and DNA binding by partial peptides of human prostate HP3. Molecular and Cellular Biochemistry 227: 97–106.

33. Bench G, Corzett MH, Kramer CE, Grant PG, Balhorn R (2000) Zinc is a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

34. Lum L, Blobel CP (1997) Evidence for Distinct Serine Protease Activities with a Rylcholine Complex. Structure 10: 505–514.

35. Nishimura H, Kim E, Nakanishi T, Baba T (2004) Possible Function of the Nematode sperm protein involved in sperm-egg fusion correlates with acquisition of fertilization competence. The Journal of Cell Biology 166: 69–78.

36. Fabrega A, Guyonnet B, Dacheux JL, Gatti JL, Puigmule M, et al. (2011) Immunohistochemical Detection of Macrophage Migration Inhibitory Factor in Fetal and Adult Bovine Epididymis: Release by the Apocrine Secretion Mode. Cells Tissues Organs 182: 22–31.

37. bench G, Corzett MH, Kramer CE, Grant PG, Balhorn R (2000) Zinc is a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

38. Blobel C (2000) Functional processing of fertilin: evidence for a critical role of proline: in sperm maturation and activation. Rev Reprod 5: 75–83.

39. Balhorn R, Schuring MP, Kaltner H, et al. (2006) Immunohistochemical Detection of Macrophage Migration Inhibitory Factor in Fetal and Adult Bovine Epididymis: Release by the Apocrine Secretion Mode. Cells Tissues Organs 182: 22–31.

40. Fabrega A, Guyonnet B, Dacheux JL, Gatti JL, Puigmule M, et al. (2011) Immunohistochemical Detection of Macrophage Migration Inhibitory Factor in Fetal and Adult Bovine Epididymis: Release by the Apocrine Secretion Mode. Cells Tissues Organs 182: 22–31.