Determination in oocytes of the reproductive modes for the brine shrimp *Artemia parthenogenetica*

Zhong-Min DAI*, Ran LI*, Li DAI*, Jin-Shu YANG*, Su CHEN*, Qing-Guo ZENG*, Fan YANG* and Wei-Jun YANG*†

*Institute of Cell Biology and Genetics, College of Life Sciences, Zijingang Campus, Zhejiang University, Hangzhou, Zhejiang 310058, China, and †National Center for Gene Conservation of the Endangered Wild Animals and Plants, College of Life Sciences, Zijingang Campus, Zhejiang University, Hangzhou, Zhejiang 310058, China

**Synopsis**

The brine shrimp, *Artemia*, reproduces either oviparously, producing encysted embryos (diapause cysts), or ovoviviparously, producing free-swimming nauplii. Environmental factors, such as photoperiod, have been applied to control the reproduction mode of *Artemia*, but when the determination of a reproductive mode occurs remains unknown. We analysed the differential gene expression between oocytes from oviparous and ovoviviparous *Artemia* reared under different photoperiods. A total of 692 qualified cDNA clones were obtained by subtractive hybridization, 327 of which matched GenBank® Nucleotide Sequence Database entries. Gene expressions of 44 cDNAs (representing 56 clones) were analysed in oocytes using real-time PCR. Among these genes, 11 (21 clones) were significantly (*P* < 0.05) up-regulated and 7 (9 clones) down-regulated in *Artemia* oocytes that subsequently enter diapause. Remarkably, known diapause-related proteins such as ArHsp22 (*Artemia* heat-shock protein 22) and chitin-binding proteins are found to be already differentially expressed. Furthermore, RNAi (RNA interference) knockdown of a differentially expressed gene, polo-like kinase 1, in oocyte of ovoviviparous *Artemia* led to the production of white embryos rather than free-swimming nauplii. In summary, our results provide evidence at the molecular level that the reproductive mode of *Artemia* is already determined at the oocyte stage of their life cycle.

**Key words:** *Artemia*, diapause, duplex-specific nuclease, oocyte, subtractive hybridization

**INTRODUCTION**

*Artemia*, a branchiopod crustacean commonly known as the brine shrimp, can withstand the widest salinity range of any known metazoan. The genus *Artemia* contains quite a number of bisexual and asexual (parthenogenetic) species. Development of embryos of both bisexual and asexual *Artemia* follows two reproductive modes: oviparity with production of diapause cysts and ovoviviparity with release of free-swimming nauplii [1].

The cysts, which are arrested at the late gastrula stage, are capable of resisting a wide variety of extraordinary environmental stresses, including natural cycles of hydration/desiccation, extreme heat or cold treatment (as reviewed in [2]), and years of continuous anoxia [3]. Several factors or molecules are reported to be characteristic of diapause cysts and are critical for their stress tolerance, including the chitinous shell, the disaccharide trehalose [3], the sHSPs (small heat-shock proteins) p26, ArHsp21 (*Artemia* heat-shock protein 21) and ArHsp22 [4–6], and the ferritin homologue artemin [7,8]. Among them, expressions of the sHSPs and artemin increase only in diapause-destined embryos and decrease during post-diapause development [5,6,8,9]. In addition, a previous study of embryos at 2 days post-fertilization revealed a transcription cofactor, p8, which is also specifically enriched in diapause-destined embryos [4]. However, the transcriptional regulation of these genes has not been elucidated [10].

Another important question is: when does the determination of reproductive mode occur? The mode of reproduction was found to be greatly affected by photoperiod, with oviparity dominating during shorter days [11]. It has been suggested that the reproductive mode of *Artemia franciscana* may be determined early...
in oocytes before fertilization, when the shell glands are already morphologically different [1]. However, no molecular evidence has been reported for the determination of reproductive mode at such an early stage. Gene expression analysis early in this developmental process is therefore a high priority for further analysis of this issue.

To investigate whether the reproductive mode is determined at the oocyte stage and to study the molecular mechanism that determines reproductive mode, we cultured *Artemia parthenogenetica* (Gahai Lake, China) under different LD (light/dark cycles), which allowed them to reproduce differently. Oocytes from *Artemia* cultured under different photoperiods were isolated and the differentially expressed genes were cloned by a novel subtractive hybridization method called DNSH [DSN (duplex-specific nuclease)-mediated normalization and subtractive hybridization]. Many genes, including several known diapause-specific genes (e.g. ArHsp22 and chitin-binding proteins), were found to be differentially expressed at the oocyte stage. The results suggest that development modes of *Artemia* embryos may be predestined in oocytes.

**MATERIALS AND METHODS**

*Artemia* culture and manipulation

*A. parthenogenetica* from Gahai Lake, China, were a gift from Liu Fengqi (College of Life Sciences, Nankai University, Tianjin, China). Specimens were maintained at room temperature (25°C) in 50–80 g/l artificial sea water prepared by Blue Starfish (Zhejiang Blue Starfish Salt Products, Hangzhou, China) and fed once every 2 days with *Chlorella* powder. For sample collection and manipulation, *Artemia* were cultured in LD 4:20 and LD 16:8 at 25°C with an illumination intensity of approx. 2000 lx to allow most of them to reproduce oviparously and ovoviviparously (L. Fengqi, personal communication) respectively. The morphology of *Artemia* shell glands was used to differentiate oviparity and ovoviviparity as described by Liang and MacRae [1]. Oocytes tended to burst in fresh water, PBS or artificial sea water; to avoid this problem, adult females with oocyte-filled lateral pouches were fixed in 75% (v/v) ethanol in PBS for approx. 4 hours at 4°C and then stored at −20°C until dissection. The lateral pouches from *Artemia* were dissected by needles under a microscope and oocytes were collected using a 200 μl pipette. After draining the excess ethanol, oocytes were stored at −80°C until RNA extraction. Total RNA was isolated by TRIzol® (Invitrogen).

The DNSH method

The schematic procedure of the DNSH is presented in Figure 1. The first-strand cDNA was synthesized by a modified template-switching method [12]. In brief, 2 μg of total RNA was mixed with 4 μl of MMLV (Moloney-murine-leukaemia virus) reverse transcriptase 5 × reaction buffer (Promega), 2 μl of 10 mM dNTPs containing 20 mM MgCl₂ and 20 mM MnCl₂, 2 μl of TsOligo-p (Table 1), 2 μl of Tspa (Table 1), 4 μl of betaine (2.5 M) and trehalose (1 M) solution, 0.5 μl of RNase inhibitor (40 units/μl; TaKaRa) and 2 μl of MMLV reverse transcriptase (H-) point mutation (200 units/μl; Promega) in a total volume of 20 μl. The mixture was incubated at 42°C for 1 h, followed by 45°C for 0.5 h. The enzyme was inactivated by adding 1.5 μl of 0.5 M DTT. Then the mixture was divided into two reaction tubes. 3 μl of SP6 polymerase (Table 1) was added to one tube and 3 μl of T7 polymerase (Table 1) was added to the other tube. The polymerase was incubated at 42°C for 1 h, followed by 45°C for 0.5 h.
of 100 mM EDTA and incubation at 70°C for 15 min. The first-strand cDNA was then purified using the AxyPrep PCR Clean-up kit (Axygen) and eluted using 80 μl of elution buffer provided by the kit.

The double-stranded cDNA from each sample was synthesized by PCR in a total volume of 3 × 100 μl. Each 100 μl PCR mixture consisted of 10 μl of 10 × TLA buffer [500 mM Tris (pH 9.2), 200 mM (NH₄)₂SO₄ and 7.5 mM MgCl₂], 4 μl of 5 mM dNTPs (containing 20 mM MgCl₂), 2 μl of primers SP6T7 and 3 α (Table 1), 4 μl of purified first-strand cDNA and 4 units of Taq+Pfu DNA polymerase (Sangon; Taq was mixed with Pfu in a ratio of 100:3). The PCR programme was 94°C for 1 min, 24 cycles at 94°C for 30 s, 60°C for 1 min, 72°C for 1 min and finally 68°C for 10 min. The PCR products were purified using the AxyPrep PCR Clean-up kit.

Double-stranded cDNA from diapause-destined oocytes was subjected to transcription by SP6 RNA polymerase (Promega). To generate sufficient RNA driver, double-stranded cDNA from nauplii-destined oocytes was subjected to transcription by the RiboMAX™ Large Scale RNA Production System-T7 (Promega). Briefly, each 0.5 μg of double-stranded cDNA was used for in vitro transcription in a total volume of 10 μl for 4 h, followed by treatment with 0.5 unit of DNase I for 20 min. After acid phenol/chloroform extraction and ethanol precipitation, the RNA pellet was dissolved in 10 μl of DEPC (diethyl pyrocarbonate)-treated water. To generate the tester RNA–DNA complexes, 0.5 μg of SP6 transcribed RNA was mixed with 2 μl of MMLV reverse transcriptase 5 × reaction buffer (Promega), 1 μl of 10 mM dNTPs, 1 μl of T7oligodT (Table 1), 2 μl of betaine (2.5 M) and trehalose (1 M) solution, 0.25 μl of RNase inhibitor (TaKaRa), and 0.75 μl of MMLV reverse transcriptase (H-) point mutation (Promega) in a total volume of 10 μl. The reaction was conducted at 42°C for 1 h and 45°C for 0.5 h. The enzyme was heat-inactivated at 70°C for 15 min after adding 1 μl of 100 mM EDTA.

For DSN-mediated normalization [13] and subtractive hybridization (DNSH), 2 μl of tester RNA–DNA complex was mixed with 1 μl of 10 × hyb buffer (0.25 M Tris, 5 M NaCl and 0.1 mM EDTA, pH 8.0), 2 μl of DMSO, 1 μl of 2 M trehalose and 4 μl of driver RNA (approx. 150 times excess). The hybridization was conducted at 55°C overnight (>16 h) after denaturing at 88°C for 1.5 min. The hybridization mixture was then transferred to a tube containing 2 μl of 10 × DSN buffer (0.5 M Tris, 70 mM MgCl₂ and 10 mM dithiothreitol, pH 8.0), 2 μl of 2 M trehalose, and 0.2 unit of DSN (Evrogen) in a total volume of 10 μl that was preheated at 65°C. The mixture was incubated at 65°C for 30 min to allow the cDNA in the DNA–RNA hybrids to be cut, as DSN specifically digests the duplex form DNA [13,14]. Finally, 1.5 μl of 100 mM EDTA and 20 μl of water were added to the mixture to stop the reaction at 68°C for 10 min.

Three rounds of PCR were used to amplify the subtracted cDNA. The PCR-suppression effect [12] was used throughout the procedures to overcome the short-molecule biased amplification. The PCR conditions were similar to those of the double-stranded cDNA synthesis. A total volume of 100 μl was used for each PCR: (a) the primary PCR ran for 30 cycles using 2 μl of subtracted product as the template and 1 μl of SOI and 2 μl of PI (Table 1) as primers (step-out PCR [15]); (b) the secondary PCR ran for 30 cycles using 0.4 μl of the primary PCR product as the template and 4 μl of PI as the primer; and (c) the tertiary PCR ran for 25 cycles using 0.4 μl of the secondary PCR product as the template and 1.6 μl of SalT7P (Table 1) as the primer. After purification using the AxyPrep PCR Clean-up kit, 1 μg of the tertiary PCR products was digested with Sall (TaKaRa). Fragments >500 bp were purified using the AxyPrep DNA Gel Extraction kit (Axygen), ligated with Sall and CIAP (calf intestinal alkaline phosphatase; TaKaRa)-treated PUC19 plasmid and transformed into TSS-treated [16,17] TOP10 Escherichia coli (Invitrogen).

### Identification of differentially expressed genes from oocytes in oviducts of Artemia

Transformed bacteria were plated together with 20 μl of 20 mg/ml X-Gal (5-bromo-4-chloroindol-3-yl β-d-galactopyranoside) on LB (Luria–Bertani) agar containing 100 μg/ml ampicillin. In all, 768 white colonies were inoculated into 96-well PCR plates filled with 100 μl of LB+ampicillin. After being shaken overnight at moderate speed at 37°C, 50 μl of 50% (v/v) glycerol in LB was added to each well and the plate was stored at −80°C. The insert cDNAs were sequenced by vector primer M13F at ZEHENGBIO (Shanghai, China). After removal of adaptor and vector sequences, the sequence data were imported into SeqMan of Lasergene software (DNASTAR) and assembled into contiguous sequences. Blastn (megablast) and Blastx were performed against the database of the nucleotide collection (nr/nt) and non-redundant protein sequences (nr) of GenBank® respectively. Sequences matching the GenBank® database (e-value <0.0001) were categorized using COGs (clusters of orthologous groups of proteins) at http://www.ncbi.nlm.nih.gov/COG/.

Quantitative real-time RT–PCR (reverse transcription–PCR) was used to further test the expression levels of 44 selected genes. Total RNAs from oocytes of diapause- and nauplius-destined Artemia were treated with 1 unit of Dnase I (Promega) at 37°C for 15 min to remove trace amounts of genomic DNA. Followed by standard phenol/chloroform extraction and ethanol precipitation, 5 μg of total RNA was reverse transcribed in a total volume of 40 μl using the oligo-dT primer from the PrimeScript™ 1st Strand cDNA Synthesis kit (TaKaRa). The first-strand cDNA was diluted by adding 360 μl of sterile water. Real-time PCR was performed with Mini opticon PTC-0148 (Bio-Rad) using a programme of 95°C for 15 s, followed by 40 cycles of 95°C for 10 s, 56°C for 10 s, 72°C for 10 s plus plate read and 78°C for 3 s plus plate read. The 20 μl of reaction mixture contained 1 μl of the diluted cDNA or inserts containing plasmids as the template, 0.4 μM of each primer (see Supplementary Table S1 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm) and 10 μl of SYBR® Premix Ex Taq™ II (TaKaRa). After normalization by α-tubulin, the expression levels of each gene were subjected to Student’s t test analysis.
**Figure 2  Oviparous and ovoviviparous reproduction in A. parthenogenetica**

Samples were fixed in paraformaldehyde and photographed under a dissecting microscope. Oocytes released from ovaries (a and a′) were temporarily stored in oviducts for approx. 3 h (b and b′), then entered into the uterus (c and c′) and developed into either cysts (d–f) or nauplii (d′–f′). E, embryo; O, oocyte; OVI, oviduct; SH, shell gland.

**RNAi (RNA interference)**

Because cDNAs were flanked by T7 promoter sequences during the DNSH procedures, dsRNA (double-stranded RNA) could be easily produced for the RNAi assay. For this purpose, plasmids containing genes of interest were isolated and transformed into competent *E. coli* HT115 (DE3). Transformed HT115 (DE3) was inoculated into LB broth with 50 μg/ml ampicillin and 10 μg/ml tetracycline and incubated at 37°C. When the culture reached *D*₆₀₀ (attenuance at 600 nm) 0.4, IPTG was added to a final concentration of 0.4 mM to induce dsRNA production for >4 h. The cells were collected by centrifugation and total RNA was extracted by TRIzol® (Invitrogen). Total RNA was dissolved in RNase buffer (10 mM Tris/HCl, 0.5 M NaCl and 10 mM MgCl₂, pH 8.0), heated at 80°C for 5 min, cooled slowly to 37°C and treated with 0.1 mg/ml RNase A and 10 units/ml DNase I for 30 min to remove single-stranded RNA and trace amounts of DNA. The dsRNA was then purified by standard phenol/chloroform extraction and ethanol precipitation. Control dsRNA was isolated from HT115 (DE3) with the plasmid PETT7 [18] as described above. RNAi was performed by microinjection of 100 ng of dsRNA into *Artemia* at the instar XII stage using the UltraMicroPump II equipped with the Micro4™ MicroSyringe pump controller.

**RESULTS AND DISCUSSION**

*Artemia* has attracted much attention because of its ability to switch from producing free-swimming nauplii to producing encysted embryos when subjected to harsh environmental conditions. Recently, Qiu et al. [4] analysed the gene expression of *A. franciscana* embryos at 2 days post-fertilization and found that many genes were differentially expressed between diapause and nauplii-destined embryos [4]. A model involving multiple factors for initiating and maintaining diapause of *Artemia* embryos was proposed in [4]. However, when the reproductive mode of *Artemia* was determined remains unknown. Identification of differentially expressed genes at an earlier stage, such as in oocytes, would improve the understanding of the molecular changes that characterize diapause development in *Artemia*. In our present study of *A. parthenogenetica*, we found that many genes with different functions were differentially expressed between oocytes from oviparous and ovoviviparous *Artemia*. Moreover, knockdown of the expression of a PLK1 (polo-like kinase 1) led the ovoviviparous *Artemia* to produce eggs rather than nauplii.

**Oviparous and ovoviviparous development of A. parthenogenetica**

Embryos of *A. parthenogenetica* undergo either oviparous or ovoviviparous development (Figure 2) in a manner similar to that of *A. franciscana* [1]. The reproductive mode that is followed by oocytes can be identified microscopically, initially by the morphology of shell glands (Figures 2a–2c and 2a′–2c′) and later by the morphology of embryos (Figures 2d–2e and 2d′–2e′). Brown shell glands indicate oviparous development (Figures 2a–2c), whereas white or colourless shell glands indicate ovoviviparous development (Figures 2a′–2c′).

**The DNSH method**

Although the classical two-step hybridization method called suppression subtractive hybridization [19] and its derivatives such as DSN-mediated transcriptome subtraction [20], and full-length normalization subtractive hybridization [21] are efficient in enrichment of highly abundant target cDNA, they are suspected to result in low representation of rare target cDNA [22]. To investigate the differential gene expression at the oocyte stage of *Artemia* following different reproductive modes, we established a novel method called DNSH (Figure 1).
exhibits maximal activity at 65°C and specifically cleaves DNA in duplex form (either DNA–DNA or DNA–RNA), but not RNA or single-stranded DNA [13,14,23,24]. To avoid loss of transcripts due to the formation of secondary structure, we added betaine, trehalose and DMSO into the hybridization and DSN reaction mixture. These materials decreased the ‘melting’ temperature of double-stranded nucleic acids [25,26] (also see Supplementary Table S2 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm). Moreover, betaine and trehalose thermostabilize DSN, although betaine partially inhibits the activity of DSN at high concentrations (see Supplementary Figures S1 and S2 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm). Differentially expressed transcripts that comprise only approx. 0.0001% of total transcripts can be efficiently enriched using the DNSH method (see Supplementary Figure S3 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm). This indicated that cDNA of low-abundant differentially expressed genes can be efficiently obtained by DNSH.

Identification of cDNAs obtained from diapause-destined *Artemia* oocytes

Using the DNSH method, 768 colonies were randomly chosen and subjected to sequencing. Among them, 42 were of poor quality and 34 contained vector alone or small inserts (<50 bp) and thus were excluded. The cDNAs from the remaining 692 colonies (GenBank® accession numbers GH635209–GH635900) were assembled into 600 contigs, most of which were singletons (529) and duplicates (57) (see Supplementary Figure S4 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm), indicating the high efficiency of normalization and subtraction of DNSH. Due to the limitations of the knowledge of the *Artemia* and any other crustacean genome and expressed sequence tags, Blastn searches of the 692 cDNAs identified only 47 cDNAs including 1 snRNA, and Blastx searches identified 327 cDNAs. The 327 protein-encoding cDNAs were then divided into the following four functional categories: metabolism, information storage and processing, cellular processes and signalling, and poorly characterized protein (see Supplementary Table S3 at http://www.bioscirep.org/bsr/031/bsr0310017add.htm and Table 2) using COGs. More detailed information about the 327 cDNAs is listed in Supplementary Table S3.

As the diapause cysts of *Artemia* have the characteristics of stress resistance and cell cycle arrest, genes that may contribute to the characteristics of diapause cysts and several other genes were selected and subjected to quantitative real-time RT–PCR analysis. Among the 44 individual tested genes (representing 56 clones), 11 genes (21 clones) were significantly up-regulated \((P < 0.05)\) and 7 genes (9 clones) were down-regulated \((P < 0.05)\) in diapause-destined *Artemia* oocytes (Table 3 and Supplementary Table S3). The mRNA levels of these genes were quite different. The chitin-binding protein (clone 1E4) was most abundantly expressed compared with the other tested genes, and its expression level was approx. 15,000 times that of nonsense-mediated mRNA decay protein 1 (clone 4H8). Most of the 44 genes were tested to be expressed in low abundance (Table 3). The result that 18 of them were confirmed to be differentially expressed between the diapause- and nauplii-destined oocytes demonstrated that our DNSH method is highly efficient.

Reproductive mode-regulated gene expression in *Artemia* oocytes

We used DNSH to study the differential transcript profiles in *A. parthenogenetica* diapause-destined oocytes relative to nauplii-destined oocytes at the filled oviduct stage. We found that many genes whose products commonly regulate metabolism, growth, transcription and stress tolerance were already differentially expressed between diapause- and nauplii-destined oocytes. Some of these genes are discussed below; more information is listed in Table 3 and Supplementary Figure S5 (at http://www.bioscirep.org/bsr/031/bsr0310017add.htm).

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of cDNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolism</td>
<td></td>
</tr>
<tr>
<td>Energy production and conversion</td>
<td>7</td>
</tr>
<tr>
<td>Carbohydrate transport and metabolism</td>
<td>15</td>
</tr>
<tr>
<td>Amino acid transport and metabolism</td>
<td>6</td>
</tr>
<tr>
<td>Nucleotide transport and metabolism</td>
<td>6</td>
</tr>
<tr>
<td>Coenzyme transport and metabolism</td>
<td>3</td>
</tr>
<tr>
<td>Lipid transport and metabolism</td>
<td>5</td>
</tr>
<tr>
<td>Inorganic ion transport and metabolism</td>
<td>7</td>
</tr>
<tr>
<td>Secondary metabolites biosynthesis, transport and catabolism</td>
<td>8</td>
</tr>
<tr>
<td>Information storage and processing</td>
<td></td>
</tr>
<tr>
<td>Translation, ribosomal structure and biogenesis</td>
<td>18</td>
</tr>
<tr>
<td>RNA processing and modification</td>
<td>7</td>
</tr>
<tr>
<td>Transcription</td>
<td>8</td>
</tr>
<tr>
<td>Replication, recombination and repair</td>
<td>25</td>
</tr>
<tr>
<td>Cellular processes and signalling</td>
<td></td>
</tr>
<tr>
<td>Cell cycle control, cell division, chromosome partitioning</td>
<td>10</td>
</tr>
<tr>
<td>Defence mechanisms</td>
<td>4</td>
</tr>
<tr>
<td>Signal transduction mechanisms</td>
<td>36</td>
</tr>
<tr>
<td>Cell wall/membrane/envelope biogenesis</td>
<td>13</td>
</tr>
<tr>
<td>Cytoskeleton</td>
<td>12</td>
</tr>
<tr>
<td>Extracellular structures</td>
<td>18</td>
</tr>
<tr>
<td>Intracellular trafficking, secretion, and vesicular transport</td>
<td>16</td>
</tr>
<tr>
<td>Post-translational modification, protein turnover, chaperones</td>
<td>10</td>
</tr>
<tr>
<td>Poorly characterized</td>
<td></td>
</tr>
<tr>
<td>General function prediction only</td>
<td>42</td>
</tr>
<tr>
<td>Function unknown</td>
<td>51</td>
</tr>
</tbody>
</table>

Chitin-binding proteins

Chitin-binding proteins play an important role in the formation of the chitinous shell of diapause cysts. The chitinous shell,
Table 3 Expression analysis of selected protein-encoding cDNAs

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Clone number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-Value</th>
<th>Oviparous oocytes</th>
<th>Ovoviviparous oocytes</th>
<th>P-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Metabolism: Carbohydrate transport and metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A10</td>
<td>GH635218</td>
<td>1</td>
<td>Marinobacter sp. ELB17/ZP_01739187.1</td>
<td>Sugar efflux permease, MFS-like protein</td>
<td>2e-64</td>
<td>71181 (44315)</td>
<td>29818 (1133)</td>
<td>0.091</td>
</tr>
<tr>
<td>7A8</td>
<td>GH635737</td>
<td>1</td>
<td>Bombyx mori/NP_001108344</td>
<td>Trehalose transporter</td>
<td>7e-38</td>
<td>138055 (37082)</td>
<td>321412 (46276)</td>
<td><strong>0.003</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Information storage and processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4C2</td>
<td>GH635491</td>
<td>1</td>
<td>Drosophila pseudoobscura/XP_001355286.2</td>
<td>Histidine acid phosphatase domain containing protein</td>
<td>9e-14</td>
<td>1402309 (526214)</td>
<td>755688 (119044)</td>
<td>0.053</td>
</tr>
<tr>
<td>4H8</td>
<td>GH635552</td>
<td>1</td>
<td>Acyrthosiphon pisum/XP_001948916.1</td>
<td>Similar to nonsense-mediated mrna decay protein 1 (rent1)</td>
<td>1e-41</td>
<td>9164 (4831)</td>
<td>12383 (2559)</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RNA processing and modification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4D12</td>
<td>GH635512</td>
<td>1</td>
<td>Danio rerio/NP_001038397.1</td>
<td>Integrator complex subunit 3, Mediator of Small Nuclear RNA Processing</td>
<td>8e-39</td>
<td>3807917 (613021)</td>
<td>4014392 (1186986)</td>
<td>0.401</td>
</tr>
<tr>
<td>7B11</td>
<td>GH635748</td>
<td>1</td>
<td>Apis mellifera/XP_624382</td>
<td>Cleavage and polyadenylation specific factor 1</td>
<td>2e-54</td>
<td>1735556 (389168)</td>
<td>1752062 (560431)</td>
<td>0.484</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Transcription</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1F3</td>
<td>GH635264</td>
<td>1</td>
<td>Marinobacter algicola DG893/ZP_01892922.1</td>
<td>Transcriptional regulator, GntR family</td>
<td>1e-51</td>
<td>14258 (3647)</td>
<td>10337 (3185)</td>
<td>0.117</td>
</tr>
<tr>
<td>2B2</td>
<td>GH635305</td>
<td>1</td>
<td>Alteromonas macleodii ‘Deep ecotype’/YP_002126057.1</td>
<td>Transcription elongation factor NusA</td>
<td>3e-74</td>
<td>42324 (7032)</td>
<td>79551 (17515)</td>
<td><strong>0.013</strong></td>
</tr>
<tr>
<td>Clone ID</td>
<td>Accession number</td>
<td>Clone number</td>
<td>Blastx top-match in GenBank$^a$: species/accession number</td>
<td>Protein name and/or proposed function</td>
<td>e-Value</td>
<td>Oviparous oocytes$^a$</td>
<td>Ooviviparous oocytes$^a$</td>
<td>$P$-value$^b$</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------------------------------------------------</td>
<td>---------------------------------------</td>
<td>---------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>6D2</td>
<td>GH635677</td>
<td>1</td>
<td>Anopheles gambiae str. PEST/XP_315936.4</td>
<td>Mediator of RNA polymerase II transcription subunit 24</td>
<td>7e-08</td>
<td>45714263 (3304319)</td>
<td>50490809 (3252923)</td>
<td>0.074</td>
</tr>
<tr>
<td>6E6</td>
<td>GH635693</td>
<td>1</td>
<td>Danio rerio/XP_688865.3</td>
<td>Regulatory factor X domain containing 2</td>
<td>3e-32</td>
<td>8558054 (1209490)</td>
<td>8470238 (1118504)</td>
<td>0.465</td>
</tr>
<tr>
<td>7E6</td>
<td>GH635775</td>
<td>1</td>
<td>Cupiennius salei/CAD57739</td>
<td>Homothorax 1, regulate the subcellular localization of the homeotic protein cofactor Extradenticle (EXD)</td>
<td>1e-07</td>
<td>1287413 (242433)</td>
<td>1187251 (204058)</td>
<td>0.307</td>
</tr>
<tr>
<td>8F1</td>
<td>GH635867</td>
<td>1</td>
<td>Xenopus laevis/AAA49904.1</td>
<td>MYB-related protein 2, DNA binding</td>
<td>1e-46</td>
<td>11400752 (1392857)</td>
<td>8441788 (268634)</td>
<td>0.011</td>
</tr>
<tr>
<td>8H1</td>
<td>GH635890</td>
<td>1</td>
<td>Pichia stipitis CBS 6054/XP_001385384.2</td>
<td>HAP2, transcriptional activator, CCAAT-binding factor, subunit B</td>
<td>2e-17</td>
<td>48080 (14107)</td>
<td>NO$^c$</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Replication, recombination and repair

<p>| 1A2      | GH635210         | 1            | Xenopus laevis/P55861                                   | DNA replication licensing factor mcm2 (minichromosome maintenance protein 2). Acts as a factor that allows the DNA to undergo a single round of replication per cell cycle | 3e-79   | 4795326 (1035578)     | 4739264 (398060)       | 0.467       |
| 2A12     | GH635303         | 1            | Danio rerio/XP_699524.3                                 | Single-strand selective monofunctional uracil DNA glycosylase | 1e-40   | 523156 (118359)       | 739844 (45331)         | 0.021       |</p>
<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Clone number</th>
<th>Blastx top-match in GenBank®: species/ accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-Value</th>
<th>Oviparous oocytes*</th>
<th>Ooviviparous oocytes*</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E4</td>
<td>GH635339</td>
<td>2</td>
<td>Portunus pelagicus/CAP20051.1</td>
<td>Transposase</td>
<td>4e-28</td>
<td>151766 (4574)</td>
<td>189350 (26361)</td>
<td>0.036</td>
</tr>
<tr>
<td>3C6</td>
<td>GH635405</td>
<td>1</td>
<td>Bombyx mori/AAQ57129.1</td>
<td>Endonuclease and reverse transcriptase-like protein</td>
<td>8e-27</td>
<td>2471521 (394559)</td>
<td>4483323 (674743)</td>
<td>0.006</td>
</tr>
<tr>
<td>5B6</td>
<td>GH635573</td>
<td>1</td>
<td>Schistosoma mansoni/CAJ00236.1</td>
<td>Endonuclease-reverse transcriptase</td>
<td>3e-36</td>
<td>43492 (21921)</td>
<td>34889 (10208)</td>
<td>0.286</td>
</tr>
<tr>
<td>1E5</td>
<td>GH635257</td>
<td>1</td>
<td>Mus musculus/CAM13330.1</td>
<td>Dedicator of cytokinesis 7, activate some small GTPases by exchanging bound GDP for free GTP</td>
<td>9e-42</td>
<td>1429111 (132200)</td>
<td>1060023 (88038)</td>
<td>0.008</td>
</tr>
<tr>
<td>3D6</td>
<td>GH635417</td>
<td>1</td>
<td>Bombyx mori/NP_001106734.1</td>
<td>Chromodomain-helicase-DNA-binding protein, Chd1</td>
<td>4e-55</td>
<td>646212 (139984)</td>
<td>668356 (51233)</td>
<td>0.405</td>
</tr>
<tr>
<td>3E8</td>
<td>GH635430</td>
<td>2</td>
<td>Danio rerio/3D5W_A</td>
<td>Polo-like kinase 1, key regulator of mitosis</td>
<td>5e-14</td>
<td>683725 (297302)</td>
<td>1654236 (729577)</td>
<td>0.050</td>
</tr>
<tr>
<td>3H4</td>
<td>GH635459</td>
<td>1</td>
<td>Aedes aegypti/XP_001650044.1</td>
<td>Brain tumour protein, brat, plays conserved roles in regulating asymmetric cell division</td>
<td>2e-110</td>
<td>432105 (58073)</td>
<td>418845 (66371)</td>
<td>0.404</td>
</tr>
<tr>
<td>4H1</td>
<td>GH635546</td>
<td>1</td>
<td>Dictyostelium discoideum AX4/XP_636910.1</td>
<td>Cell division cycle protein 48, ATPase activity</td>
<td>2e-70</td>
<td>151011 (130802)</td>
<td>48588 (7636)</td>
<td>0.124</td>
</tr>
<tr>
<td>5H7</td>
<td>GH635638</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001851276.1</td>
<td>Maintenance of ploidy protein mob1, spindle pole body duplication and mitotic checkpoint regulation</td>
<td>7e-81</td>
<td>3326932 (355978)</td>
<td>3331466 (579226)</td>
<td>0.496</td>
</tr>
</tbody>
</table>
Table 3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Clone number</th>
<th>Blastx top-match in GenBank&lt;sup&gt;®&lt;/sup&gt;: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-Value</th>
<th>Oviparous oocytes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Ovoviviparous oocytes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P-value&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>6F11</td>
<td>GH635710</td>
<td>1</td>
<td>Marinobacter aquaeolei VT8/YP_959599.1</td>
<td>Septum site-determining protein MinD, cell division inhibitor, membrane ATPase, activates MinC</td>
<td>1e-96</td>
<td>203325 (31950)</td>
<td>167603 (39060)</td>
<td>0.144</td>
</tr>
<tr>
<td>6G4</td>
<td>GH635715</td>
<td>1</td>
<td>Tribolium castaneum/ XP_967278.1</td>
<td>Glucose inhibited division protein a</td>
<td>2e-58</td>
<td>9179540 (1800825)</td>
<td>9412445 (974140)</td>
<td>0.427</td>
</tr>
<tr>
<td>8F2</td>
<td>GH635868</td>
<td>1</td>
<td>Culex quinquefasciatus/ XP_001868862.1</td>
<td>Septin, GTP-binding proteins associated with diverse processes in dividing and non-dividing cells</td>
<td>9e-90</td>
<td>NO&lt;sup&gt;c&lt;/sup&gt;</td>
<td>NO&lt;sup&gt;c&lt;/sup&gt;</td>
<td>NO&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Signal transduction mechanisms

| 1C2      | GH635234         | 1            | Caenorhabditis elegans/NP_001122601.1                        | Multiple PDZ domain protein family member (mpz-1). PDZ domains bind C-terminal polypeptides | 1e-25   | 384951 (372239)             | 124457 (98006)              | 0.153           |
| 1F10     | GH635269         | 1            | Aedes aegypti/XP_001661429.1                                 | Serine/threonine protein phosphatase 2c | 2e-50   | 781458 (138552)            | 594286 (118186)            | 0.075           |
| 5D2      | GH635592         | 1            | Aedes aegypti/XP_001655129.1                                 | Juvenile hormone-inducible protein. Phosphotransferases of the serine or threonine-specific kinase subfamily | 1e-12   | 814885 (210772)            | 791626 (166497)            | 0.444           |
| 6F2      | GH635701         | 3            | Aedes aegypti/XP_001653183.1                                 | kek1, transmembrane protein Kekkon1, functions during oogenesis in a negative feedback loop to directly attenuate EGFR activity | 1e-65   | 4874429 (1035110)          | 2643792 (381283)           | 0.012           |

<sup>a</sup>Values estimated by qRT-PCR.

<sup>b</sup>P-values calculated by Student’s t test.

<sup>c</sup>Hypothesis rejected.
Table 3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Clone number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-Value</th>
<th>Oviparous oocytes(^a)</th>
<th>Ovoviviparous oocytes(^a)</th>
<th>P-value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8D1</td>
<td>GH635844</td>
<td>1</td>
<td><em>Dictyostelium discoideum</em> AX4/XP_637704.1</td>
<td>Extracellular-signal-regulated protein kinase</td>
<td>5e-60</td>
<td>4674108 (773778)</td>
<td>3717952 (667592)</td>
<td>0.090</td>
</tr>
<tr>
<td>Extracellular structures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1E4</td>
<td>GH635256</td>
<td>8</td>
<td>*Tribolium castaneum/NP_001073566.1</td>
<td>Chitin binding</td>
<td>3e-38</td>
<td>145004903 (17496865)</td>
<td>83435 (10291)</td>
<td>6.9e-5</td>
</tr>
<tr>
<td>5E3</td>
<td>GH635605</td>
<td>1</td>
<td>*Tribolium castaneum/NP_001073566.1</td>
<td>Chitin binding</td>
<td>2e-41</td>
<td>15137020 (2722017)</td>
<td>186539 (59612)</td>
<td>3.4e-4</td>
</tr>
<tr>
<td>Intracellular trafficking, secretion, and vesicular transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1F2</td>
<td>GH635263</td>
<td>1</td>
<td>*Rattus norvegicus/NP_114010.1</td>
<td>Nuclear receptor co-activator 2, ligand-dependent nuclear receptor transcription co-activator activity</td>
<td>6e-33</td>
<td>36555843 (2154518)</td>
<td>32220981 (2602268)</td>
<td>0.045</td>
</tr>
<tr>
<td>6D8</td>
<td>GH635683</td>
<td>1</td>
<td>*Aedes aegypti/XP_001654070.1</td>
<td>RAN binding protein</td>
<td>8e-11</td>
<td>15956226 (7616989)</td>
<td>11118037 (1423072)</td>
<td>0.170</td>
</tr>
<tr>
<td>Post-translational modification, protein turnover, chaperones</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A9</td>
<td>GH635217</td>
<td>1</td>
<td>*Artemia urmiana/ABY86224.1</td>
<td>Artemin, molecular chaperone activity</td>
<td>2e-49</td>
<td>39974 (3652)</td>
<td>68263 (13714)</td>
<td>0.013</td>
</tr>
<tr>
<td>2G6</td>
<td>GH635363</td>
<td>2</td>
<td>*Artemia franciscana/ABD19713.1</td>
<td>Small heat shock protein ArHsp22</td>
<td>2e-47</td>
<td>38109728 (4870733)</td>
<td>916717 (117775)</td>
<td>9.5e-5</td>
</tr>
<tr>
<td>5A3</td>
<td>GH635558</td>
<td>1</td>
<td>*Bombyx mori/NP_001036958.1</td>
<td>p23-like protein, p23 binds heat shock protein hsp90 and participates in the folding of a number of proteins</td>
<td>2e-15</td>
<td>13667141 (2894748)</td>
<td>22587321 (6983461)</td>
<td>0.055</td>
</tr>
<tr>
<td>5A11</td>
<td>GH635566</td>
<td>1</td>
<td><em>Dictyostelium discoideum</em> AX4/XP_635173.1</td>
<td>Chaperonin containing TCP1 eta subunit, folding of proteins, ATPase</td>
<td>5e-91</td>
<td>92113 (61186)</td>
<td>NO(^c)</td>
<td>0.030</td>
</tr>
</tbody>
</table>
### Table 3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Clone number</th>
<th>Blastx top-match in GenBank: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-Value</th>
<th>Oviparous oocytes</th>
<th>Ooviviparous oocytes</th>
<th>P-value&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>8E1</td>
<td>GH635855</td>
<td>1</td>
<td><em>Culex quinquefasciatus/XP_001845910.1</em></td>
<td>Ubiquitin-conjugating enzyme morgue, ubiquitin-mediated protein degradation pathway</td>
<td>6e-08</td>
<td>38807882 (8838408)</td>
<td>40987126 (2925164)</td>
<td>0.353</td>
</tr>
<tr>
<td>2E11</td>
<td>GH635346</td>
<td>1</td>
<td><em>Acyrthosiphon pisum/XP_001950343.1</em></td>
<td>Similar to NFX1-type zinc finger-containing protein</td>
<td>2e-7</td>
<td>1684927 (273135)</td>
<td>1661413 (471297)</td>
<td>0.472</td>
</tr>
<tr>
<td>5B2</td>
<td>GH635569</td>
<td>1</td>
<td><em>Caenorhabditis elegans/NP_001023145.1</em></td>
<td>Dauer up-regulated family member (dur-1)</td>
<td>2e-15</td>
<td>951530 (50601)</td>
<td>73024 (20448)</td>
<td><strong>4.9e-6</strong></td>
</tr>
<tr>
<td>1A5</td>
<td>GH635213</td>
<td>1</td>
<td>No</td>
<td>Unknown</td>
<td></td>
<td>No hit</td>
<td>207217 (18509)</td>
<td><strong>0.005</strong></td>
</tr>
<tr>
<td>Internal control</td>
<td></td>
<td></td>
<td>α-tubulin</td>
<td></td>
<td>21386652 (1501833)</td>
<td>21393767 (502529)</td>
<td>0.497</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Levels of mRNA expression: results are means (S.D.) (copies/μg of total RNA) of three replicates and normalized by tubulin.

<sup>b</sup>The P-values were calculated by the Student’s t test. The P-values of significantly up-regulated genes are shown in boldface and italicized, while those of significantly down-regulated genes are shown in boldface.

<sup>c</sup>Not detected by the 40 cycles of realtime PCR. Expression level was set to zero to calculate the P-value.
which is tough and impermeable to non-volatile solutes, protects the cysts from harsh conditions [27]. Five chitin-binding protein encoding cDNAs (representing 14 clones) were isolated in our subtracted library. cDNA clone 1E4 isolated in the present study is identical with a chitin-binding protein (GenBank® accession no. DW678201) isolated from *A. franciscana* embryos at 2 days post-fertilization [4], whereas clones 6G10 and 8D4 are more similar to another chitin-binding protein (GenBank® accession no. DW678180). Two tested cDNAs (clones 1E4 and 5E3) were up-regulated in diapause-destined oocytes. Our results provide evidence that those genes have already been up-regulated in diapause-destined oocytes, indicating that those proteins are critical for diapause cyst formation. However, why so many types of chitin-binding proteins are transcribed at such an early oocyte stage and how they are involved in the formation of the chitinous shell of *Artemia* cysts remain unknown. Additional work is required to elucidate how the diapause embryos are encysted.

**Major facilitator families**

Two metabolic genes, *sugar efflux permease* (clone 1A10) and *trehalose transporter* (clone 7A8), were tested. The former was not significantly up-regulated, whereas the latter was significantly down-regulated (*P* = 0.003) in diapause-destined oocytes (Table 3). Trehalose, which constitutes approx. 15% of the dry weight of cysts [3], can either alone or synergistically with p26 stabilize cells during dehydration and rehydration [28–30]. It was showed that *trehalose transporter* from the insect, *Polypedilum vanderplanki*, can transport trehalose across the cell membrane bidirectionally to balance the trehalose concentration [31]. Taking into account that trehalose plays an important role in the survival of diapause embryos of *Artemia*, the down-regulated expression of *trehalose transporter* indicates that trehalose is unlikely to be transported from females to her egg. Thus the accumulation of trehalose in cysts must be made by the egg itself. This premise is supported by the discovery that trehalose synthesis begins in diapause-destined embryos but not in nauplii-destined embryos of *A. franciscana* at approx. 2 days after fertilization (reviewed in [2]).

**Molecular chaperones**

Expressions of p26, ArHsp21, ArHsp22 and artemin were previously reported to increase at 2 days post-fertilization in *A. franciscana* [1,5,6,8]. Although we did not obtain a p26 clone from our subtracted library, we have identified the ArHsp22 and artemin encoding cDNAs therein. In contrast with the expression pattern in *A. franciscana* [6], ArHsp22 was highly expressed in diapause-destined oocytes of *A. parthenogenetica* (Table 3). This suggests that ArHsp22 may not only enhance the cysts’ stress resistance, but also function in oocyte or embryo development. The mRNA level of artemin in diapause-destined oocytes of *A. parthenogenetica* was very low, although it was up-regulated in nauplii-destined oocytes for unknown reasons.

**Cell cycle control proteins**

Proteins that play important roles in regulation of cell growth, such as dedicator of cytokinesis (clone 1E5), kekkon1 (clone 6F2) and PLK1 (clone 3E8), are significantly differentially expressed in *Artemia* (Table 3). The former two were up-regulated in diapause-destined oocytes, while the latter one was down-regulated. Dedicator of cytokinesis family proteins regulates migration, morphology, adhesion and growth of cells, whereas deficiency of dedicator of cytokinesis may lead to carcinogenesis [32]. Expression of the transmembrane protein kekkon1, which functions during oogenesis in a negative feedback loop to directly attenuate EGFR (epidermal growth factor receptor) activity, suppresses the growth of mouse mammary tumour cells derived from aberrant ErbB receptors [33]. The PLK1 is critical for regulation of cell division. It has been shown that the expression of PLK1 is inhibited by the tumour suppressor protein p53; in contrast, overexpression of PLK1 inactivated p53 (reviewed in [34]). The up-regulation of dedicator of cytokinesis and kekkon1 and the down-regulation of PLK1 in diapause-destined oocytes of *Artemia* may contribute to the cell-cycle arrest of diapause embryos.

**Transcription factor**

The transcription factor, HAP2 (haem-activated protein 2; clone 8H1), was significantly up-regulated in diapause-destined oocytes of *Artemia* (*P* = 0.002). HAP2 is a subunit of the heterotrimeric CCAAT-binding complex. The CCAAT-binding complex is conserved in eukaryotic organisms such as fungi, plants and animals [35]. It is essential to mitochondrial functions, and its activity was decreased under anoxia in yeast [36]. The CCAAT-binding transcription factor has been shown to rescue yeast and human cells from a cytochrome c oxidase defect [37]. In addition to the sHSPs and trehalose, the up-regulation of HAP2 in diapause-destined oocytes may also contribute to the survival of diapause cysts under years of anoxia.

**RNAi knockdown of PLK1**

To further understand the functions of these differentially expressed genes in *Artemia*, RNAi was used to knock down the expression of genes such as ArHsp22, dur-1 (dauer-up-regulated family member 1) and PLK1. No significant morphological changes were observed in the diapause cysts of *Artemia* when the expression of ArHsp22 and dur-1 was knocked down (results not shown). However, knockdown of the expression of PLK1 (approx. 63% knockdown) in oocytes of ovoviviparous *Artemia* led to the production of white embryos rather than free-swimming nauplii (Figure 3). This finding further supports the premise that down-regulated expression of PLK1 in diapause-destined oocytes of *Artemia* may contribute to the developmental arrest of diapause embryos.

In summary, we used a novel subtractive hybridization approach, DNSH, to study the differential transcript profiles in *A. parthenogenetica* diapause-destined oocytes relative to nauplii-destined oocytes at the filled oviduct stage. Many genes, including several known diapause-specific genes (e.g. ArHsp22 and
chitin-binding proteins), were found to be differentially expressed between diapause- and nauplii-distended oocytes. Our results suggest that the development modes of *A. parthenogenetica* embryos may be predestined in oocytes.

**ACKNOWLEDGEMENTS**

We thank Mr Chris Wood of the College of Life Sciences, Zhejiang University, for a critical reading of this paper. *E. coli* HT115 (DE3) was provided by the Caenorhabditis Genetics Center, which is funded by the NIH (National Institutes of Health) NCRR (National Center for Research Resources).

**FUNDING**

This work was supported by the National Natural Science Foundation of China [grant numbers 40730212, 40876069]; the ‘863’ Program of China [grant number 2007AA09Z2426]; the Postdoctoral Scientific Foundation of China [grant number 20090461373]; and the Scientific Research Fund of Zhejiang Provincial Education Department [grant number Y200909613].

**REFERENCES**

SUPPLEMENTARY ONLINE DATA

Determination in oocytes of the reproductive modes for the brine shrimp Artemia parthenogenetica

Zhong-Min DAI*, Ran LI*, Li DAI*, Jin-Shu YANG*, Su CHEN*, Qing-Guo ZENG*, Fan YANG* and Wei-Jun YANG*†

*Institute of Cell Biology and Genetics, College of Life Sciences, Zijingang Campus, Zhejiang University, Hangzhou, Zhejiang 310058, China, and †National Center for Gene Conservation of the Endangered Wild Animals and Plants, College of Life Sciences, Zijingang Campus, Zhejiang University, Hangzhou, Zhejiang 310058, China

Figure S1 Thermostability of DSN
DSN (0.001 unit) was incubated at 90°C for 0–120 s in 10 μl of 1 x DSN buffer with or without additives. The DSN was then mixed with 2 μg of DNA dissolved in 1 x DSN buffer and allowed to react at 65°C for 30 min. The reaction mixture was then subjected to agarose electrophoresis.

Figure S2 Activity of DSN
C0, control; C1, DSN only; B1–B4, betaine added; T1–T4, trehalose added; D1–D4, DMSO added. The same agarose gel was photographed at different times.

---

1 To whom correspondence should be addressed (email w_jyang@cls.zju.edu.cn).
Z.-M. Dai and others

Figure S3 Efficiency of DNSH
To generate ADEG (artificial differentially expressed genes), four RNA transcripts from Lambda phage DNA were transcribed in vitro by SP6 RNA polymerase, mixed together and then added to the total RNA of ovary from Portunus trituberculatus. After reverse transcription, the cDNAs from ADEG-containing RNA were mixed with 30-fold ovary RNAs. The mixture was then subjected to DSN-mediated normalization and subtraction. The final products were analysed by agarose-gel electrophoresis (A) and Southern hybridization (B). Detailed information about the procedure is available on request.

Figure S4 Identification and expression analysis of cDNAs from diapause-destined Artemia oocytes
(A) Summary of subtracted cDNA clones. (B) Functional categorization of isolated cDNA clones. (C) The 56 tested clones were distributed according to their mRNA levels.
Reproductive modes determine in *Artemia* oocytes

The GenBank® accession numbers are p26 (DQ310575), Af-Zic (AB231878) and polyubiquitin (X74405). A great diversity of transposable element-encoding cDNAs (17 cDNAs representing 19 clones) were isolated from the subtracted library (Supplementary Table S3). Expression analysis of three genes showed that two, transposase (clone 2E4) and endonuclease and reverse transcriptase-like protein (clone 3C6), were significantly down-regulated in diapause-destined oocytes (Table 3 in the main paper). Of equal interest is that 12 cDNAs (15 clones) matched at least one of the genomic DNA sequences of *A. franciscana* small heat-shock protein p26 (GenBank® accession number DQ310575), the zinc finger protein Af-Zic (GenBank® accession number AB231878) and polyubiquitin (GenBank® accession number X74405) by the Blastn search. These cDNAs are potentially from non-protein encoding RNAs, as the Blastx search found no similarity. Further analysis demonstrated that the matched regions are inverted repeats or possibly directed repeats. Some of these cDNAs are proposed to be the precursors of piRNA, a class of small RNAs that act through RNAi pathways to regulate gene expression and suppress the spread of transposable elements [1]. Moreover, the p26, Af-Zic and polyubiquitin genomic DNAs shared the inverted repeats found in the cDNAs with which clones 5A10 and 7H1 matched, and some of the repeated sequences were located in the intron regions. It is possible that the expression of p26, Af-Zic and polyubiquitin is regulated by transposable elements, as multiple repeated sequences were found in their genomic DNA.

**Figure S5** Repetitive sequences in genomic DNA of p26, zinc-finger protein and polyubiquitin

![Image of repetitive sequences](image_url)
### Table S1 Primers for real-time RT–PCR

<table>
<thead>
<tr>
<th>Clone name</th>
<th>Primer name</th>
<th>Sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>1A2F</td>
<td>TTTCCTGTATTGCCACCG</td>
</tr>
<tr>
<td>1A2R</td>
<td></td>
<td>AAGCCTATGCGCTTCTTA</td>
</tr>
<tr>
<td>1A5</td>
<td>1A5F</td>
<td>CATCTCCTTTAAGAGTTTCTC</td>
</tr>
<tr>
<td>1A5R</td>
<td></td>
<td>GAACGATCAAAACCAGT</td>
</tr>
<tr>
<td>1A9</td>
<td>1A9F</td>
<td>AGACCTCTGGAAGGACT</td>
</tr>
<tr>
<td>1A9R</td>
<td></td>
<td>CGTGGACCGCCTGAAAC</td>
</tr>
<tr>
<td>1A10</td>
<td>1A10F</td>
<td>GGTGCCTGCTTGGTGGTGC</td>
</tr>
<tr>
<td>1A10R</td>
<td></td>
<td>TAACCAGGATGCTGGTG</td>
</tr>
<tr>
<td>1C2</td>
<td>1C2F</td>
<td>TCTTCACCTCCGACATCG</td>
</tr>
<tr>
<td>1C2R</td>
<td></td>
<td>TCTTCGATATCCGGCTA</td>
</tr>
<tr>
<td>1E4</td>
<td>1E4F</td>
<td>TTTGCTGGTGGAGGTTGCG</td>
</tr>
<tr>
<td>1E4R</td>
<td></td>
<td>AAGTCAGGAGGCGTGGTATG</td>
</tr>
<tr>
<td>1E5</td>
<td>1E5F</td>
<td>GCATTGCCACAGATTCATAGTTT</td>
</tr>
<tr>
<td>1E5R</td>
<td></td>
<td>TTAGAGGGATGGGAGACG</td>
</tr>
<tr>
<td>1F2</td>
<td>1F2F</td>
<td>CAGGAGATGTTCAAGAGGGTG</td>
</tr>
<tr>
<td>1F2R</td>
<td></td>
<td>TTGCAGGAAAGAGGAAATG</td>
</tr>
<tr>
<td>1F3</td>
<td>1F3F</td>
<td>TACCCGACCCATGCTGACCTC</td>
</tr>
<tr>
<td>1F3R</td>
<td></td>
<td>GCTGCGCCTCCTGATTCCTTC</td>
</tr>
<tr>
<td>1F10</td>
<td>1F10F</td>
<td>GCACAAGATAGGTTACCACGAT</td>
</tr>
<tr>
<td>1F10R</td>
<td></td>
<td>AAATTCGCGAGATGATTCAGG</td>
</tr>
<tr>
<td>2A12</td>
<td>2A12F</td>
<td>GAACATCCTAAGCGACCAATAC</td>
</tr>
<tr>
<td>2A12R</td>
<td></td>
<td>TTCCACCTGCGAACCAATA</td>
</tr>
<tr>
<td>2B2</td>
<td>2B2F</td>
<td>TTTCCTGGTCTGCTGGT</td>
</tr>
<tr>
<td>2B2R</td>
<td></td>
<td>CATCGGACCTCGTGATG</td>
</tr>
<tr>
<td>2E4</td>
<td>2E4F</td>
<td>GCTTTCGCTTGGCTTC</td>
</tr>
<tr>
<td>2E4R</td>
<td></td>
<td>ACCCGCTACACTACCA</td>
</tr>
<tr>
<td>2E11</td>
<td>2E11F</td>
<td>CAAATTCCGTCGACACG</td>
</tr>
<tr>
<td>2E11R</td>
<td></td>
<td>GCTTCATAATCGCTTCCACT</td>
</tr>
<tr>
<td>2G6</td>
<td>2G6F</td>
<td>AGTGGACTGCTGCTGCTT</td>
</tr>
<tr>
<td>2G6R</td>
<td></td>
<td>TATCCCTGCGCTTCCG</td>
</tr>
<tr>
<td>3C6</td>
<td>3C6F</td>
<td>TGAATGCTGGTCTGGTAGTCT</td>
</tr>
<tr>
<td>3C6R</td>
<td></td>
<td>CGTTGATCTGGGACAAA</td>
</tr>
<tr>
<td>3D6</td>
<td>3D6F</td>
<td>GCCGAGGGTTGCGACTT</td>
</tr>
<tr>
<td>3D6R</td>
<td></td>
<td>CAACCGTGGCCTGATTTT</td>
</tr>
<tr>
<td>3H4</td>
<td>3H4F</td>
<td>TCACTGCGCTAAGATTCACG</td>
</tr>
<tr>
<td>3H4R</td>
<td></td>
<td>AAACAGACGCACTCATCA</td>
</tr>
<tr>
<td>4C2</td>
<td>4C2F</td>
<td>AATATCGCCCTAAGGCTACCA</td>
</tr>
<tr>
<td>4C2R</td>
<td></td>
<td>ATGCGCTTACAGGGCGAAC</td>
</tr>
<tr>
<td>4D12</td>
<td>4D12F</td>
<td>AGGCAAAGCGACTTCTT</td>
</tr>
<tr>
<td>4D12R</td>
<td></td>
<td>CCACTAAATCGCAGACAG</td>
</tr>
<tr>
<td>4H1</td>
<td>4H1F</td>
<td>AACCCGATATCTGCTGACCTG</td>
</tr>
<tr>
<td>4H1R</td>
<td></td>
<td>TTTGCCTCACTCTGCTTCC</td>
</tr>
<tr>
<td>4H8</td>
<td>4H8F</td>
<td>AGGACTGCTGTGCTAACAGAAG</td>
</tr>
<tr>
<td>4H8R</td>
<td></td>
<td>CAGCGTCTGCAACCACACT</td>
</tr>
<tr>
<td>5A3</td>
<td>5A3F</td>
<td>TCAGTGCGGAAAGCAGGAC</td>
</tr>
<tr>
<td>5A3R</td>
<td></td>
<td>TTTGTATCTAGGAGATGGTG</td>
</tr>
<tr>
<td>5A11</td>
<td>5A11F</td>
<td>TCACCGGAAAGCAAACACTCA</td>
</tr>
<tr>
<td>5A11R</td>
<td></td>
<td>CAGGCTCCCACAAACAGAAG</td>
</tr>
<tr>
<td>5B2</td>
<td>5B2F</td>
<td>GCTGAGCGGAGCTTCTTG</td>
</tr>
<tr>
<td>5B2R</td>
<td></td>
<td>TTGTTCACCTGATTCCTG</td>
</tr>
</tbody>
</table>

### Table S1 (Continued)

<table>
<thead>
<tr>
<th>Clone name</th>
<th>Primer name</th>
<th>Sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5B6</td>
<td>5B6F</td>
<td>ACTGCTTCTGCCGACCTC</td>
</tr>
<tr>
<td>5B6R</td>
<td></td>
<td>CCTGTGATCCTGCAAGTGG</td>
</tr>
<tr>
<td>5D2</td>
<td>5D2F</td>
<td>CGCTTGGCTTCAACGATT</td>
</tr>
<tr>
<td>5D2R</td>
<td></td>
<td>TTGTGCGGCCGTTGATT</td>
</tr>
<tr>
<td>5E3</td>
<td>5E3F</td>
<td>GATATGGCTAAGCGAAGT</td>
</tr>
<tr>
<td>5E3R</td>
<td></td>
<td>TCTCTGCGACCTCTTCACTT</td>
</tr>
<tr>
<td>5H7</td>
<td>5H7F</td>
<td>GAAACCGGAGGAGAACCTCA</td>
</tr>
<tr>
<td>5H7R</td>
<td></td>
<td>CAAAACACTTCGGGACAC</td>
</tr>
<tr>
<td>6D2</td>
<td>6D2F</td>
<td>CTGGCAAAGCTGGATAGG</td>
</tr>
<tr>
<td>6D2R</td>
<td></td>
<td>TGAACGCTGCAACAGGAG</td>
</tr>
<tr>
<td>6E6</td>
<td>6E6F</td>
<td>AGCAGACCTTTGGCTGAGA</td>
</tr>
<tr>
<td>6E6R</td>
<td></td>
<td>CGGATTAGGGGTCCTTGG</td>
</tr>
<tr>
<td>6F2</td>
<td>6F2F</td>
<td>GGCGCTCCCTTCTCACCG</td>
</tr>
<tr>
<td>6F2R</td>
<td></td>
<td>TCTCCGCGACATCAAC</td>
</tr>
<tr>
<td>6F11</td>
<td>6F11F</td>
<td>GCGGAAAGCCTCGGAAAC</td>
</tr>
<tr>
<td>6F11R</td>
<td></td>
<td>CCGTAGCAGCAGAAAAAG</td>
</tr>
<tr>
<td>6G4</td>
<td>6G4F</td>
<td>TACTTGGCAACAAATGG</td>
</tr>
<tr>
<td>6G4R</td>
<td></td>
<td>AAATTCGCGAGATGATTCAGG</td>
</tr>
<tr>
<td>7A8</td>
<td>7A8F</td>
<td>TTCTCCACTCGTCTCC</td>
</tr>
<tr>
<td>7A8R</td>
<td></td>
<td>GGCGATACCTCGGCGT</td>
</tr>
<tr>
<td>7B11</td>
<td>7B11F</td>
<td>CGTACGACAGCTGACATAC</td>
</tr>
<tr>
<td>7B11R</td>
<td></td>
<td>GCCGACGAGAGGAGTTT</td>
</tr>
<tr>
<td>7E6</td>
<td>7E6F</td>
<td>GGTTGACGCTTGGGAT</td>
</tr>
<tr>
<td>7E6R</td>
<td></td>
<td>CTGAAGGCGACTGAGATG</td>
</tr>
<tr>
<td>8D1</td>
<td>8D1F</td>
<td>AGTGGACGCCGACAGACG</td>
</tr>
<tr>
<td>8D1R</td>
<td></td>
<td>TGGGACAGCACTATAGGAAA</td>
</tr>
<tr>
<td>8E1</td>
<td>8E1F</td>
<td>AAGTTGGAACGCAAGAAAG</td>
</tr>
<tr>
<td>8E1R</td>
<td></td>
<td>GACCAATACTCTCATGCGCA</td>
</tr>
<tr>
<td>8F1</td>
<td>8F1F</td>
<td>TGACCACTACGACAG</td>
</tr>
<tr>
<td>8F1R</td>
<td></td>
<td>GTTATACGACGCTTCC</td>
</tr>
<tr>
<td>8F2</td>
<td>8F2F</td>
<td>AGGACAACACTCCTGTTA</td>
</tr>
<tr>
<td>8F2R</td>
<td></td>
<td>CTACGAGCAACATTGATACCC</td>
</tr>
<tr>
<td>8H1</td>
<td>8H1F</td>
<td>TTGGCAGACAGAAAGCAG</td>
</tr>
<tr>
<td>8H1R</td>
<td></td>
<td>TGAACAGCGACAGAACAAT</td>
</tr>
<tr>
<td>8H11</td>
<td>8H11F</td>
<td>CAATAGGATCAGATGGT</td>
</tr>
<tr>
<td>8H11R</td>
<td></td>
<td>AGGTATGGGACTTGAAT</td>
</tr>
</tbody>
</table>

### Table S2 Decreased ‘melting’ temperature of DNA by betaine, trehalose and DMSO

Samples are different between betaine + trehalose and DMSO experiments. Results are means (S.D.) for triplicate experiments.

<table>
<thead>
<tr>
<th>Betaine + trehalose (M)</th>
<th>Tₘ (S.D.)</th>
<th>DMSO (% v/v)</th>
<th>Tₘ (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>90 (0)</td>
<td>0</td>
<td>91 (0)</td>
</tr>
<tr>
<td>0.25/0.1</td>
<td>89 (0)</td>
<td>2</td>
<td>89.67 (0.58)</td>
</tr>
<tr>
<td>0.5/0.2</td>
<td>88 (0)</td>
<td>5</td>
<td>88 (0)</td>
</tr>
<tr>
<td>1/0.4</td>
<td>86 (0)</td>
<td>10</td>
<td>84 (1.48)</td>
</tr>
<tr>
<td>2/0.6</td>
<td>82 (0)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table S3  Clusters of orthologous groups

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Energy production and conversion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4H9</td>
<td>GH635553</td>
<td>1</td>
<td>Anopheles gambiae/XP_319323.3</td>
<td>ATP-citrate synthase</td>
<td>3e-72</td>
</tr>
<tr>
<td>8G8</td>
<td>GH635885</td>
<td>1</td>
<td>Physcomitrella patens subsp. Patens/XP_001757403.1</td>
<td>methylmalonate semialdehyde dehydrogenase, ALDH6A acting on the aldehyde or oxo group of donor with NAD⁺ or NADP⁺ as acceptor</td>
<td>8e-118</td>
</tr>
<tr>
<td>8E10</td>
<td>GH635864</td>
<td>1</td>
<td>Cryptococcus neoformans JEC21/XP_571714.1</td>
<td>electron transporter, transferring electrons within CoQH2-cytochrome c reductase complex</td>
<td>7e-53</td>
</tr>
<tr>
<td>1G6</td>
<td>GH635275</td>
<td>1</td>
<td>Smirnopsis crassicaudata/ABR32188.1</td>
<td>mitochondrial uncoupling protein 1. Mediating mitochondrial proton leak, thermogenesis</td>
<td>2e-62</td>
</tr>
<tr>
<td>7A7</td>
<td>GH635736</td>
<td>1</td>
<td>Manduca sexta/Q25531</td>
<td>vascular proton pump subunit d (V-ATPase subunit d)</td>
<td>9e-134</td>
</tr>
<tr>
<td>3A5</td>
<td>GH635383</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001867838.1</td>
<td>quinone oxidoreductase</td>
<td>1e-18</td>
</tr>
<tr>
<td>4B8</td>
<td>GH635485</td>
<td>1</td>
<td>Tetraodon nigroviridis/CAG12439.1</td>
<td>phytanoyl-CoA dioxygenase domain-containing protein 1</td>
<td>2e-61</td>
</tr>
<tr>
<td></td>
<td><strong>Carbohydrate transport and metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2G4</td>
<td>GH635361</td>
<td>1</td>
<td>Microscilla marina ATCC 23134/ZP_01690626.1</td>
<td>2,3-bisphosphoglycerate-independent phosphoglycerate mutase</td>
<td>2e-94</td>
</tr>
<tr>
<td>3E10</td>
<td>GH635432</td>
<td>1</td>
<td>Anopheles gambiae str. PEST/XP_001689393.1</td>
<td>phosphorylase kinase B, regulatory enzyme in glycogen metabolism</td>
<td>4e-15</td>
</tr>
<tr>
<td>3G5</td>
<td>GH635449</td>
<td>2</td>
<td>Aedes aegypti/XP_001662906.1</td>
<td>2-deoxyglucose-6-phosphate phosphatase</td>
<td>3e-59</td>
</tr>
<tr>
<td>8B3</td>
<td>GH635824</td>
<td>1</td>
<td>Actinobacillus succinogenes 130Z/YP_001343634.1</td>
<td>alpha amylase</td>
<td>7e-73</td>
</tr>
<tr>
<td>5G12</td>
<td>GH635633</td>
<td>1</td>
<td>Drosophila pseudoobscura/CAI63869.1</td>
<td>b3-glucuronotransferase-R galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase</td>
<td>6e-66</td>
</tr>
<tr>
<td>6B4</td>
<td>GH635658</td>
<td>1</td>
<td>Helicoverpa armigera/ACB54958.1</td>
<td>chitin deacetylase</td>
<td>8e-73</td>
</tr>
<tr>
<td>3B4</td>
<td>GH635392</td>
<td>1</td>
<td>Callinectes sapidus/ABN51214.1</td>
<td>glycosylphosphatidylinositol-linked carbonic anhydrase</td>
<td>2e-42</td>
</tr>
<tr>
<td>8D8</td>
<td>GH635850</td>
<td>1</td>
<td>Gallus gallus/NP_001073183.1</td>
<td>mannosidase, alpha, class 2C, member 1</td>
<td>4e-88</td>
</tr>
<tr>
<td>4E8</td>
<td>GH635519</td>
<td>1</td>
<td>Aedes aegypti/XP_001663852.1</td>
<td>mannosyl-oligosaccharide alpha-1,2-mannosidase</td>
<td>2e-84</td>
</tr>
<tr>
<td>1A10</td>
<td>GH635218</td>
<td>1</td>
<td>Marinobacter sp. ELB17/YP_01739187.1</td>
<td>sugar efflux permease, MFS-like protein</td>
<td>2e-64</td>
</tr>
<tr>
<td>7A8</td>
<td>GH635737</td>
<td>1</td>
<td>Bombax mori/NP_001108344</td>
<td>trehalose transporter</td>
<td>7e-38</td>
</tr>
<tr>
<td>8A8</td>
<td>GH635819</td>
<td>1</td>
<td>Drosophila erecta/XP_001975417.1</td>
<td>Major Facilitator Superfamily, MFS_1. Sugar transporter</td>
<td>6e-57</td>
</tr>
<tr>
<td>2G11</td>
<td>GH635368</td>
<td>1</td>
<td>Homo sapiens/NP_060760.2</td>
<td>PGM2 (phosphoglucomutase-2; glucose phosphomutase 2)</td>
<td>6e-37</td>
</tr>
<tr>
<td>Clone ID</td>
<td>Accession number</td>
<td>Copy number</td>
<td>Blastx top-match in GenBank: species/accession number</td>
<td>Protein name and/or proposed function</td>
<td>e-value</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>4D3</td>
<td>GH635504</td>
<td>1</td>
<td>Apis mellifera/XP_394961.3</td>
<td>glycogen debranching enzyme</td>
<td>2e-73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino acid transport and metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8E3</td>
<td>GH635857</td>
<td>1</td>
<td>Mus musculus/NP_694762</td>
<td>5-oxoprolinase (ATP-hydrolysing), ATP-hydrolysing activity</td>
<td>7e-33</td>
</tr>
<tr>
<td>1A7</td>
<td>GH635215</td>
<td>1</td>
<td>Aedes aegypti/XP_001653299.1</td>
<td>glutaminase</td>
<td>1e-155</td>
</tr>
<tr>
<td>7C3</td>
<td>GH635751</td>
<td>1</td>
<td>Apis mellifera/XP_394694</td>
<td>glutamate receptor</td>
<td>3e-81</td>
</tr>
<tr>
<td>6C8</td>
<td>GH635672</td>
<td>1</td>
<td>Tenacibaculum sp. MED152/ZP_01052844.1</td>
<td>methionine synthase</td>
<td>5e-19</td>
</tr>
<tr>
<td>5E2</td>
<td>GH635604</td>
<td>1</td>
<td>Nitrospumoccus oceani C-27/YP_342858.1</td>
<td>peptidase M50, putative membrane-associated zinc metallopeptidase</td>
<td>2e-68</td>
</tr>
<tr>
<td>8H3</td>
<td>GH635892</td>
<td>1</td>
<td>Mus musculus/NP_598422.2</td>
<td>solute carrier family 6, neurotransmitter transporter, betaine/GABA (γ-aminobutyric acid)</td>
<td>2e-25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nucleotide transport and metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2C4</td>
<td>GH635317</td>
<td>1</td>
<td>Equus caballus/XP_001490668.3</td>
<td>dithydropirimidine dehydrogenase, involved in pyrimidine degradation</td>
<td>1e-121</td>
</tr>
<tr>
<td>1C12</td>
<td>GH635243</td>
<td>1</td>
<td>Zea mays/NP_001105431.1</td>
<td>adenine nucleotide translocator, ADP ATP carrier protein 1, mitochondrial</td>
<td>2e-96</td>
</tr>
<tr>
<td>8A3</td>
<td>GH635814</td>
<td>1</td>
<td>Aedes aegypti/XP_001650904.1</td>
<td>CTP synthase</td>
<td>1e-27</td>
</tr>
<tr>
<td>2B5</td>
<td>GH635308</td>
<td>1</td>
<td>Tribolium castaneum/XP_965264.1</td>
<td>cytosolic purine 5-nucleotidase</td>
<td>2e-43</td>
</tr>
<tr>
<td>3A1</td>
<td>GH635380</td>
<td>1</td>
<td>Aedes aegypti/XP_001651486</td>
<td>uridine phosphorylase</td>
<td>2e-50</td>
</tr>
<tr>
<td>4B1</td>
<td>GH635479</td>
<td>1</td>
<td>Drosophila melanogaster/XP_723592.1</td>
<td>ribosomal pseudouridine synthase, RluA-1</td>
<td>4e-24</td>
</tr>
<tr>
<td>Coenzyme transport and metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1B11</td>
<td>GH635231</td>
<td>1</td>
<td>Xenopus laevis/AAH77971.1</td>
<td>glutathione synthetase, large subunit</td>
<td>6e-13</td>
</tr>
<tr>
<td>5C12</td>
<td>GH635591</td>
<td>1</td>
<td>Aedes aegypti/XP_001653705.1</td>
<td>glutathione synthetase</td>
<td>6e-47</td>
</tr>
<tr>
<td>5A5</td>
<td>GH635560</td>
<td>1</td>
<td>Branchiostoma belcheri/tsingtaunese/AAQ96656.1</td>
<td>S-adenosyl-L-homocysteine hydrolase, adenosylhomocysteinase</td>
<td>6e-31</td>
</tr>
<tr>
<td>Lipid transport and metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4G3</td>
<td>GH635536</td>
<td>3</td>
<td>Locusta migratoria/Q9U943.2</td>
<td>apolipopophors precursor, retinoid- and fatty acid-binding glycoprotein</td>
<td>7e-26</td>
</tr>
<tr>
<td>1B5</td>
<td>GH635225</td>
<td>1</td>
<td>Mus musculus/AAC52329.1</td>
<td>long-chain acyl-CoA dehydrogenase</td>
<td>3e-37</td>
</tr>
<tr>
<td>3H8</td>
<td>GH635463</td>
<td>1</td>
<td>Aedes aegypti/XP_001660425.1</td>
<td>myotubularin, dephosphorylates phosphatidylinositol 3-phosphate and phosphatidylinositol (3,5)-bi-phosphate</td>
<td>9e-10</td>
</tr>
</tbody>
</table>

©The Authors Journal compilation ©2011 Biochemical Society
<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®; species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A12</td>
<td>GH635220</td>
<td>2</td>
<td><em>Drosophila virilis</em> / XP_002059755.1</td>
<td>cation transport ATPase</td>
<td>1e-112</td>
</tr>
<tr>
<td>2B7</td>
<td>GH635310</td>
<td>4</td>
<td><em>Mus musculus</em> / CAB92141.1</td>
<td>chondroitin 4-sulfotransferase</td>
<td>2e-34</td>
</tr>
<tr>
<td>5F4</td>
<td>GH635615</td>
<td>1</td>
<td><em>Drosophila melanogaster</em> / NP_001097666.1</td>
<td>secretory pathway calcium ATPase CG32451-PD</td>
<td>4e-58</td>
</tr>
</tbody>
</table>

### Inorganic ion transport and metabolism

| 8F7      | GH635873         | 1           | *Bos taurus* / NP_001068926.1                         | 3-hydroxyanthranilate 3,4-dioxygenase | 5e-24   |
| 3G6      | GH635450         | 2           | *Photuris pennsylvania* / BAA05005.1                   | luciferase                           | 3e-42   |
| 2C11     | GH635324         | 2           | *Mus musculus* / Q8CIM7.1                             | cytochrome P450 2D26, involved in the oxidative degradation of various compounds | 1e-42   |
| 8C9      | GH635840         | 1           | *Streptomyces hygroscopicus* subsp. Yingchengensis / AAZ91667.2 | 2-epi-5-epi-valiolone synthase       | 2e-10   |
| 5B4      | GH635571         | 2           | *Aedes aegypti* / XP_001653883.1                       | steroid dehydrogenase               | 2e-58   |

### Secondary metabolites biosynthesis, transport and catabolism

<p>| 3E11     | GH635433         | 2           | <em>Drosophila yakuba</em> / XP_002100541.1                   | ribonuclease E, 5'-End-dependent, single-strand-specific endonuclease | 2e-06   |
| 6G12     | GH635721         | 1           | <em>Ommatostomina huwena</em> / ACH48199.1                   | 60S ribosomal protein L27            | 3e-43   |
| 3H9      | GH635464         | 1           | <em>Artemia franciscana</em> / ABK91823.1                    | 60S ribosomal protein L31            | 2e-55   |
| 6G5      | GH635716         | 1           | <em>Ommatostomina parkeri</em> / ABR23438.1                  | ribosomal protein L21                | 7e-33   |
| 3B6      | GH635394         | 2           | <em>Artemia franciscana</em> / ABY62742.1                    | ribosomal protein S23e               | 4e-78   |
| 4F2      | GH635524         | 1           | <em>Physcomitrella patens</em> subsp. Patens / XP_001753710.1 | 60S ribosomal protein L4-B (L1)      | 2e-62   |
| 5G6      | GH635627         | 1           | <em>Suberites donuncula</em> / AAX48834.1                    | ribosomal_L4                         | 4e-52   |
| 8G3      | GH635881         | 1           | <em>Arabidopsis thaliana</em> / BAB03107.1                   | pre-rRNA processing protein RRPS     | 2e-05   |
| 3G3      | GH635447         | 1           | <em>Xenopus laevis</em> / NP_001084027.1                     | ribonucleoprotein, RNA binding activity | 6e-20   |
| 1H8      | GH635288         | 1           | <em>Ostreococcus tauri</em> / CAL55825.1                     | asparaginyl-tRNA synthetase 1, asparagine-tRNA ligase activity | 2e-48   |
| 6C2      | GH635667         | 1           | <em>Spirogyra</em> sp. FWAC125 / ABQ81942.1                  | translation elongation factor 1 alpha | 4e-72   |
| 1D12     | GH635253         | 1           | <em>Brugia malayi</em> / XP_001897144.1                      | translation elongation factor aEF-2, translocation of the peptidyl-tRNA | 1e-69   |</p>
<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank(^2): species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2C1</td>
<td>GH635315</td>
<td>1</td>
<td>Bos taurus/ XP_588358.2</td>
<td>TRM1 tRNA methyltransferase 1</td>
<td>2e-27</td>
</tr>
<tr>
<td>5H4</td>
<td>GH635636</td>
<td>1</td>
<td>Algoriphagus sp. PR1/ ZP_01719006.1</td>
<td>tRNA (guanine-N(1))- methyltransferase</td>
<td>8e-89</td>
</tr>
<tr>
<td>4H8</td>
<td>GH635552</td>
<td>1</td>
<td>Acyrthosiphon pisum/ XP_001948916.1</td>
<td>similar to nonsense-mediated mRNA decay protein 1 (rent1)</td>
<td>1e-41</td>
</tr>
<tr>
<td>4C2</td>
<td>GH635491</td>
<td>1</td>
<td>Drosophila pseudoobscura/ XP_001355286.2</td>
<td>histidine acid phosphatase domain containing protein</td>
<td>9e-14</td>
</tr>
</tbody>
</table>

RNA processing and modification

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank(^2): species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G12</td>
<td>GH635281</td>
<td>1</td>
<td>Gramella forsetii KTO803/ YP_862189.1</td>
<td>ATP-dependent RNA helicase, ATP-dependent RNA unwinding</td>
<td>2e-109</td>
</tr>
<tr>
<td>7B11</td>
<td>GH635748</td>
<td>1</td>
<td>Apis mellifera/ XP_624382</td>
<td>cleavage and polyadenylation specific factor 1</td>
<td>2e-54</td>
</tr>
<tr>
<td>5C4</td>
<td>GH635583</td>
<td>3</td>
<td>Danio rerio/ NP_001017993.1</td>
<td>cleavage and polyadenylation specific factor 6</td>
<td>4e-16</td>
</tr>
<tr>
<td>4D12</td>
<td>GH635512</td>
<td>1</td>
<td>Danio rerio/ NP_001038397.1</td>
<td>integrator complex subunit 3, Mediator of Small Nuclear RNA Processing</td>
<td>8e-39</td>
</tr>
<tr>
<td>2D10</td>
<td>GH635334</td>
<td>1</td>
<td>Culex quinquefasciatus/ XP_001848576.1</td>
<td>serine/arginine-rich splicing factor</td>
<td>4e-35</td>
</tr>
</tbody>
</table>

Transcription

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank(^2): species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3A7</td>
<td>GH635385</td>
<td>1</td>
<td>Aedes aegypti/ XP_001663321.1</td>
<td>DNA-directed RNA polymerase I largest subunit</td>
<td>1e-33</td>
</tr>
<tr>
<td>2B2</td>
<td>GH635305</td>
<td>1</td>
<td>Alteromonas macleodii 'Deep ecotype'/YP_00126057.1</td>
<td>transcription elongation factor NusA</td>
<td>3e-74</td>
</tr>
<tr>
<td>8H1</td>
<td>GH635890</td>
<td>1</td>
<td>Pichia stipitis CBS 6054/XP_001385384.2</td>
<td>transcriptional activator, CCAAT-binding factor, subunit B (HAP2)</td>
<td>2e-17</td>
</tr>
<tr>
<td>1F3</td>
<td>GH635264</td>
<td>1</td>
<td>Marinobacter algicola DG893/2P_01892922.1</td>
<td>transcriptional regulator, GntR family</td>
<td>1e-51</td>
</tr>
<tr>
<td>7E6</td>
<td>GH635787</td>
<td>1</td>
<td>Cupiennius salei/ CAD57739</td>
<td>homothorax 1, regulate the subcellular localization of the homeotic protein cofactor Extradenticle (EXD)</td>
<td>1e-07</td>
</tr>
<tr>
<td>6D2</td>
<td>GH635677</td>
<td>1</td>
<td>Anopheles gambiae str. PEST/XP_315936.4</td>
<td>mediator of RNA polymerase II transcription subunit 24</td>
<td>7e-08</td>
</tr>
<tr>
<td>8F1</td>
<td>GH635867</td>
<td>1</td>
<td>Xenopus laevis/ AAA49904.1</td>
<td>myb-related protein 2, DNA binding</td>
<td>1e-46</td>
</tr>
<tr>
<td>6E6</td>
<td>GH635693</td>
<td>1</td>
<td>Danio rerio/ XP_688865.3</td>
<td>regulatory factor X domain containing 2</td>
<td>3e-32</td>
</tr>
</tbody>
</table>

Replication, recombination and repair

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank(^2): species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>7F9</td>
<td>GH635789</td>
<td>2</td>
<td>Canis lupus familiaris/XP_545850</td>
<td>DNA polymerase (DNA directed), gamma</td>
<td>3e-45</td>
</tr>
<tr>
<td>8G1</td>
<td>GH635879</td>
<td>1</td>
<td>Homo sapiens/ NP_058633.2</td>
<td>polymerase (DNA-directed), alpha</td>
<td>2e-67</td>
</tr>
<tr>
<td>1A2</td>
<td>GH635210</td>
<td>1</td>
<td>Xenopus laevis/ P55861</td>
<td>DNA replication licensing factor mcml2 (Minichromosome maintenance protein 2), Acts as a factor that allows the DNA to undergo a single round of replication per cell cycle</td>
<td>3e-79</td>
</tr>
<tr>
<td>3C6</td>
<td>GH635405</td>
<td>1</td>
<td>Bombyx mori/ AAQ57129.1</td>
<td>endonuclease and reverse transcriptase-like protein</td>
<td>8e-27</td>
</tr>
</tbody>
</table>
Table S3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>7C6</td>
<td>GH635753</td>
<td>1</td>
<td>Schistosoma mansoni/CAJ00236</td>
<td>endonuclease-reverse transcriptase</td>
<td>5e-21</td>
</tr>
<tr>
<td>6C1</td>
<td>GH635666</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/XP_001179836.1</td>
<td>endonuclease-reverse transcriptase</td>
<td>7e-06</td>
</tr>
<tr>
<td>6D5</td>
<td>GH635680</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/XP_001198883.1</td>
<td>endonuclease-reverse transcriptase</td>
<td>6e-13</td>
</tr>
<tr>
<td>6F10</td>
<td>GH635709</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/XP_001182458.1</td>
<td>endonuclease-reverse transcriptase</td>
<td>1e-12</td>
</tr>
<tr>
<td>7A2</td>
<td>GH635732</td>
<td>1</td>
<td>Zingiber officinale/ABK60177</td>
<td>reverse transcriptase</td>
<td>4e-11</td>
</tr>
<tr>
<td>3E12</td>
<td>GH635434</td>
<td>1</td>
<td>Penaeus monodon/ABB73282.1</td>
<td>reverse transcriptase</td>
<td>1e-12</td>
</tr>
<tr>
<td>4F4</td>
<td>GH635526</td>
<td>1</td>
<td>Takifugu rubripes/AAD19348.1</td>
<td>reverse transcriptase-like protein</td>
<td>2e-17</td>
</tr>
<tr>
<td>7D5</td>
<td>GH635762</td>
<td>2</td>
<td>Strongylocentrotus purpuratus/XP_001198586</td>
<td>similar to endonuclease-reverse transcriptase</td>
<td>2e-48</td>
</tr>
<tr>
<td>4E3</td>
<td>GH635515</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/XP_001198883.1</td>
<td>similar to endonuclease-reverse transcriptase</td>
<td>3e-11</td>
</tr>
<tr>
<td>7B5</td>
<td>GH635745</td>
<td>1</td>
<td>Strongylocentrotus urpuratus/XP_001182458</td>
<td>similar to endonuclease-reverse transcriptase</td>
<td>2e-17</td>
</tr>
<tr>
<td>3H5</td>
<td>GH635460</td>
<td>1</td>
<td>Strongylocentrotus/XP_001200565.1</td>
<td>similar to putative reverse transcriptase</td>
<td>6e-08</td>
</tr>
<tr>
<td>5B6</td>
<td>GH635573</td>
<td>1</td>
<td>Schistosoma mansoni/CAJ00236</td>
<td>endonuclease-reverse transcriptase</td>
<td>3e-36</td>
</tr>
<tr>
<td>5D11</td>
<td>GH635601</td>
<td>1</td>
<td>Schistosoma mansoni/CAJ00240.1</td>
<td>endonuclease-reverse transcriptase</td>
<td>2e-14</td>
</tr>
<tr>
<td>8C1</td>
<td>GH635833</td>
<td>1</td>
<td>Schistosoma mansoni/CAJ00236</td>
<td>endonuclease-reverse transcriptase</td>
<td>3e-15</td>
</tr>
<tr>
<td>2E4</td>
<td>GH635339</td>
<td>2</td>
<td>Portunus pelagicus/CAP20051.1</td>
<td>transposase</td>
<td>4e-28</td>
</tr>
<tr>
<td>8H12</td>
<td>GH635900</td>
<td>1</td>
<td>Adineta vaga/ACD54739.1</td>
<td>PIF/Harbinger-like protein, DNA transposons</td>
<td>2e-17</td>
</tr>
<tr>
<td>5F5</td>
<td>GH635616</td>
<td>1</td>
<td>Gallus gallus/NP_001026515.1</td>
<td>Twinkle protein, mitochondrial precursor (progressive external ophthalmoplegia 1 protein)</td>
<td>1e-45</td>
</tr>
<tr>
<td>2A12</td>
<td>GH635303</td>
<td>1</td>
<td>Danio rerio/XP_699524.3</td>
<td>single-strand selective monofunctional uracil DNA glycosylase</td>
<td>1e-40</td>
</tr>
</tbody>
</table>

Chromatin structure and dynamics

Cellular processes and signalling

Cell cycle control, cell division, chromosome partitioning

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3H4</td>
<td>GH635459</td>
<td>1</td>
<td>Aedes aegypti/XP_001650044.1</td>
<td>brain tumour protein, brat. play conserved roles in regulating asymmetric cell division</td>
<td>2e-110</td>
</tr>
<tr>
<td>4H1</td>
<td>GH635546</td>
<td>1</td>
<td>Dictyostelium discoideum AXA/XP_636910.1</td>
<td>cell division cycle protein 48, ATPase activity</td>
<td>2e-70</td>
</tr>
<tr>
<td>3E8</td>
<td>GH635430</td>
<td>2</td>
<td>Danio rerio/3D5W_A</td>
<td>polo-like kinase 1. Key regulator of mitosis</td>
<td>5e-14</td>
</tr>
<tr>
<td>3D6</td>
<td>GH635417</td>
<td>1</td>
<td>Bombyx mori/NP_001106734.1</td>
<td>chromodomain-helicase-DNA-binding protein, Chd1</td>
<td>4e-55</td>
</tr>
</tbody>
</table>

www.bioscirep.org / Volume 31 (1) / Pages 17–30
<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1E5</td>
<td>GH635257</td>
<td>1</td>
<td><em>Mus musculus</em>/CAM13330.1</td>
<td>dedicator of cytokinesis 7, activates some small GTPases by exchanging bound GDP for free GTP</td>
<td>9e-42</td>
</tr>
<tr>
<td>5H7</td>
<td>GH635638</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em>/XP_001851276.1</td>
<td>maintenance of ploidy protein mob1. Spindle pole body duplication and mitotic checkpoint regulation</td>
<td>7e-81</td>
</tr>
<tr>
<td>8F2</td>
<td>GH635868</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em>/XP_001868862.1</td>
<td>septin, GTP-binding proteins associated with diverse processes in dividing and non-dividing cells</td>
<td>9e-90</td>
</tr>
<tr>
<td>6F11</td>
<td>GH635710</td>
<td>1</td>
<td><em>Marinobacter aquaeolei</em>/VT8/YP_959599.1</td>
<td>septum site-determining protein MinD, cell division inhibitor, membrane ATPase, activates MinC</td>
<td>1e-96</td>
</tr>
<tr>
<td>6G4</td>
<td>GH635715</td>
<td>1</td>
<td><em>Tribolium castaneum</em>/XP_967278.1</td>
<td>glucose inhibited division protein a</td>
<td>2e-58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Nuclear structure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D7</td>
<td>GH635418</td>
<td>1</td>
<td>Trypanosoma brucei/TREU927/XP_951518.1</td>
<td>autophagocytosis associated protein, Protein transport, regulation of autophagy</td>
<td>6e-37</td>
</tr>
<tr>
<td>3F10</td>
<td>GH635442</td>
<td>1</td>
<td><em>Drosophila grimshawi</em>/XP_001985447.1</td>
<td>Autophagy-specific gene 1 CG10967-PA</td>
<td>2e-11</td>
</tr>
<tr>
<td>3F2</td>
<td>GH635436</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em>/XP_001847549.1</td>
<td>pmp22 peroxisomal membrane protein</td>
<td>4e-39</td>
</tr>
<tr>
<td>7D11</td>
<td>GH635768</td>
<td>1</td>
<td><em>Verasper variegatus</em>/BAF34210</td>
<td>serum lectin isoform 3</td>
<td>5e-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Signal transduction mechanisms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1H2</td>
<td>GH635283</td>
<td>1</td>
<td><em>Mus musculus</em>/NP_001095900.1</td>
<td>ADP-ribosylation factor guanine nucleotide-exchange factor 1 (brefeldin A-inhibited), Arfgef1. Carries the exchange factor activity</td>
<td>7e-69</td>
</tr>
<tr>
<td>8D10</td>
<td>GH635852</td>
<td>1</td>
<td><em>Aedes aegypti</em>/XP_001653141.1</td>
<td>thyroid hormone receptor-associated protein</td>
<td>1e-51</td>
</tr>
<tr>
<td>2C9</td>
<td>GH635322</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em>/XP_001866947.1</td>
<td>thyroid receptor interacting protein, formin binding protein 1-like</td>
<td>8e-39</td>
</tr>
<tr>
<td>1H12</td>
<td>GH635292</td>
<td>1</td>
<td><em>Aedes aegypti</em>/XP_001652701.1</td>
<td>attractin, involved in development</td>
<td>2e-68</td>
</tr>
<tr>
<td>1B1</td>
<td>GH635221</td>
<td>1</td>
<td><em>Gallus gallus</em>/NP_001007846.1</td>
<td>cAMP-dependent protein kinase type I-alpha regulatory subunit</td>
<td>3e-11</td>
</tr>
<tr>
<td>8F4</td>
<td>GH635870</td>
<td>1</td>
<td><em>Caenorhabditis elegans</em>/NP_508283.2</td>
<td>dopamine receptor family member (dop-4), G-protein coupled receptor protein signalling pathway</td>
<td>6e-06</td>
</tr>
<tr>
<td>8D1</td>
<td>GH635844</td>
<td>1</td>
<td><em>Dictyostelium discoideum</em>/AX4/XP_637704.1</td>
<td>extracellular signal-regulated protein kinase</td>
<td>5e-60</td>
</tr>
<tr>
<td>8H8</td>
<td>GH635897</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em>/XP_001866812.1</td>
<td>integrin-linked protein kinase, interacts with the cytoplasmic domain of beta-1 integrin</td>
<td>4e-113</td>
</tr>
<tr>
<td>5D2</td>
<td>GH635592</td>
<td>1</td>
<td><em>Aedes aegypti</em>/XP_001655129.1</td>
<td>juvenile hormone-inducible protein, Phosphotransferases of the serine- or threonine-specific kinase subfamily</td>
<td>1e-12</td>
</tr>
</tbody>
</table>
Table S3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6F2</td>
<td>GH635701</td>
<td>3</td>
<td>Aedes aegypti/XP_001653183.1</td>
<td>kek1, transmembrane protein Kekkon1; functions during oogenesis in a negative feedback loop to directly attenuate EGFR activity</td>
<td>1e-65</td>
</tr>
<tr>
<td>2F2</td>
<td>GH635349</td>
<td>1</td>
<td>Drosophila ananassae/XP_001955514.1</td>
<td>leucine-rich repeat-containing G-protein coupled receptor 4</td>
<td>1e-9</td>
</tr>
<tr>
<td>2B1</td>
<td>GH635304</td>
<td>1</td>
<td>Aedes aegypti/XP_001649206.1</td>
<td>non-receptor serine/threonine protein kinase</td>
<td>3e-21</td>
</tr>
<tr>
<td>2E5</td>
<td>GH635340</td>
<td>1</td>
<td>Mus musculus/EDK98473.1</td>
<td>lipin 1. Phosphorylated by the TOR Ser/Thr protein kinases in response to insulin stimulation</td>
<td>2e-16</td>
</tr>
<tr>
<td>4E4</td>
<td>GH635516</td>
<td>1</td>
<td>Tetrahymena thermophila SB210/XP_001020671.1</td>
<td>serine/threonine protein kinase</td>
<td>9e-59</td>
</tr>
<tr>
<td>8C2</td>
<td>GH635834</td>
<td>1</td>
<td>Drosophila ananassae/XP_001955555.1</td>
<td>serine/threonine protein kinases</td>
<td>2e-37</td>
</tr>
<tr>
<td>1F10</td>
<td>GH635269</td>
<td>1</td>
<td>Aedes aegypti/XP_001661429.1</td>
<td>serine/threonine protein phosphatase 2c</td>
<td>2e-50</td>
</tr>
<tr>
<td>4B7</td>
<td>GH635484</td>
<td>1</td>
<td>Aedes aegypti/XP_001648484.1</td>
<td>serine/threonine protein phosphatase 2c</td>
<td>2e-27</td>
</tr>
<tr>
<td>1C2</td>
<td>GH635234</td>
<td>1</td>
<td>Caenorhabditis elegans/NP_001122601.1</td>
<td>multiple PDZ domain protein family member (mpz-1); PDZ domains bind C-terminal polypeptides</td>
<td>1e-25</td>
</tr>
<tr>
<td>7C5</td>
<td>GH635752</td>
<td>1</td>
<td>Bombyx mori/NP_001127716</td>
<td>neuropeptide receptor A18</td>
<td>2e-11</td>
</tr>
<tr>
<td>7D9</td>
<td>GH635766</td>
<td>1</td>
<td>Saccoglossus kowalevskii/ABD97277</td>
<td>netrin1, involved in axon guidance</td>
<td>2e-78</td>
</tr>
<tr>
<td>2C12</td>
<td>GH635325</td>
<td>3</td>
<td>Hydractinia echinata/CAA86062.1</td>
<td>precursor protein for neuropeptide Pol-RFamide</td>
<td>2e-12</td>
</tr>
<tr>
<td>2B9</td>
<td>GH635311</td>
<td>1</td>
<td>Aedes aegypti/XP_001846225.1</td>
<td>phosphorylase b kinase beta, kpbb</td>
<td>3e-52</td>
</tr>
<tr>
<td>5F2</td>
<td>GH635613</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001850978.1</td>
<td>phosphorylase b kinase</td>
<td>1e-113</td>
</tr>
<tr>
<td>8C6</td>
<td>GH635837</td>
<td>2</td>
<td>Aedes aegypti/XP_001663674.1</td>
<td>RHO guanyl-nucleotide exchange factor</td>
<td>1e-07</td>
</tr>
<tr>
<td>6E5</td>
<td>GH635692</td>
<td>2</td>
<td>Aplysia californica/ACI23623.1</td>
<td>Src tyrosine kinase 2</td>
<td>2e-06</td>
</tr>
<tr>
<td>8A6</td>
<td>GH635817</td>
<td>1</td>
<td>Alteromonadas bacterium TW-7/ZP_01612227.1</td>
<td>sensory histidine kinase in two-component regulatory system with PhoP</td>
<td>2e-87</td>
</tr>
<tr>
<td>1H3</td>
<td>GH635284</td>
<td>2</td>
<td>Mus musculus/NP_058075.2</td>
<td>serine/arginine-rich protein specific kinase 1</td>
<td>2e-18</td>
</tr>
<tr>
<td>6F9</td>
<td>GH635708</td>
<td>1</td>
<td>Apis mellifera/XP_001121143.1</td>
<td>similar to Bicaudal C CG4824-PA</td>
<td>2e-14</td>
</tr>
<tr>
<td>7E1</td>
<td>GH635770</td>
<td>1</td>
<td>Tribolium castaneum/XP_969670</td>
<td>protein kinase C delta type homologue</td>
<td>8e-12</td>
</tr>
</tbody>
</table>

**Cell wall/membrane/envelope biogenesis**

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3B11</td>
<td>GH635398</td>
<td>1</td>
<td>Aedes aegypti/XP_001681111.1</td>
<td>c3f, membrane bound O-acyl transferase (MBOAT)</td>
<td>1e-52</td>
</tr>
<tr>
<td>2C8</td>
<td>GH635321</td>
<td>1</td>
<td>Aedes aegypti/XP_001681388.1</td>
<td>dTDP-glucose 4-6-dehydratase</td>
<td>2e-53</td>
</tr>
<tr>
<td>6E1</td>
<td>GH635688</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001863315.1</td>
<td>mannose-1-phosphate guanylytransferase</td>
<td>1e-22</td>
</tr>
</tbody>
</table>
Table S3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4G5</td>
<td>GH635538</td>
<td>1</td>
<td><em>Herpetosiphon aurantiacus</em> ATCC 23779/YP_001547947.1</td>
<td>peptidoglycan binding domain-containing protein</td>
<td>5e-08</td>
</tr>
<tr>
<td>3E9</td>
<td>GH635431</td>
<td>1</td>
<td><em>Homo sapiens</em> P32019</td>
<td>Type II inositol-1,4,5-trisphosphate 5-phosphatase precursor</td>
<td>4e-33</td>
</tr>
<tr>
<td>7G5</td>
<td>GH635795</td>
<td>2</td>
<td><em>Homo sapiens</em> ABM83104</td>
<td>transforming growth factor, beta-induced, TGFBI</td>
<td>2e-24</td>
</tr>
<tr>
<td>6B3</td>
<td>GH635657</td>
<td>1</td>
<td><em>Ovis aries</em> NP_001098400.1</td>
<td>receptor (chemosensory) transporter protein 4</td>
<td>2e-06</td>
</tr>
<tr>
<td>5C10</td>
<td>GH635589</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em> XP_001843995.1</td>
<td>tetraspanin, trans-membrane proteins with 4 trans-membrane segments</td>
<td>2e-42</td>
</tr>
<tr>
<td>3D12</td>
<td>GH635422</td>
<td>2</td>
<td><em>Drosophila ananassae</em> XP_001967549.1</td>
<td>nephrin, necessary for the renal filtration barrier</td>
<td>2e-71</td>
</tr>
<tr>
<td>3H1</td>
<td>GH635457</td>
<td>1</td>
<td><em>Caenorhabditis elegans</em> NP_495680.1</td>
<td>Nose Resistant to Fluoxetine, nrf-6, embryonic development ending in birth or egg hatching</td>
<td>3e-14</td>
</tr>
<tr>
<td>3E5</td>
<td>GH635427</td>
<td>1</td>
<td><em>Danio rerio</em> NP_001018407.1</td>
<td>cystinosis, nephropathic, membrane bound proteins</td>
<td>2e-45</td>
</tr>
</tbody>
</table>

**Cell motility**

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4H4</td>
<td>GH635549</td>
<td>1</td>
<td><em>Drosophila virilli</em> AAK25829.1</td>
<td>actin E2, structural molecule activity, protein binding</td>
<td>1e-20</td>
</tr>
<tr>
<td>2D3</td>
<td>GH635327</td>
<td>1</td>
<td><em>Nasonia vitripennis</em> XP_00160980.1</td>
<td>kelch-like 20, KLHL20. Protein binding</td>
<td>2e-40</td>
</tr>
<tr>
<td>4B10</td>
<td>GH635487</td>
<td>2</td>
<td><em>Artemia franciscana</em> CA098769.1</td>
<td>beta-actin</td>
<td>4e-68</td>
</tr>
<tr>
<td>1E2</td>
<td>GH635255</td>
<td>1</td>
<td><em>Mus musculus</em> EDL24928.1</td>
<td>dynein cytoplasmic 2 heavy chain 1, Dync2h1. ATPase activity and microtubule binding ability</td>
<td>7e-13</td>
</tr>
<tr>
<td>1C11</td>
<td>GH635242</td>
<td>1</td>
<td><em>Bombyx mori</em> ABE71569.1</td>
<td>Factin bundling protein, LIM domain containing protein</td>
<td>8e-36</td>
</tr>
<tr>
<td>4G10</td>
<td>GH635543</td>
<td>1</td>
<td><em>Aplysia californica</em> AA40189.1</td>
<td>nonmuscle myosin II</td>
<td>2e-08</td>
</tr>
<tr>
<td>1G7</td>
<td>GH635276</td>
<td>1</td>
<td><em>Bombyx mori</em> NP_001124374.1</td>
<td>Paramyosin</td>
<td>1e-74</td>
</tr>
<tr>
<td>6B9</td>
<td>GH635663</td>
<td>1</td>
<td><em>Homo sapiens</em> EAX00513.1</td>
<td>restin-like 2</td>
<td>2e-28</td>
</tr>
<tr>
<td>2H11</td>
<td>GH635378</td>
<td>1</td>
<td><em>Drosophila melanogaster</em> NP_476820.1</td>
<td>vinculin</td>
<td>4e-114</td>
</tr>
<tr>
<td>8C8</td>
<td>GH635839</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em> XP_001851358.1</td>
<td>spectrin</td>
<td>8e-80</td>
</tr>
<tr>
<td>3C4</td>
<td>GH635403</td>
<td>1</td>
<td><em>Rattus norvegicus</em> EDM04816.1</td>
<td>myosin, heavy polypeptide 10, non-muscle</td>
<td>8e-36</td>
</tr>
</tbody>
</table>

**Extracellular structures**

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1E4</td>
<td>GH635256</td>
<td>8</td>
<td><em>Trichobium castaneum</em> NP_001073566.1</td>
<td>chitin binding</td>
<td>3e-38</td>
</tr>
<tr>
<td>5E3</td>
<td>GH635605</td>
<td>1</td>
<td><em>Trichobium castaneum</em> NP_001073566.1</td>
<td>chitin binding</td>
<td>2e-41</td>
</tr>
<tr>
<td>6G10</td>
<td>GH635719</td>
<td>1</td>
<td><em>Trichobium castaneum</em> NP_001073566.1</td>
<td>chitin binding</td>
<td>5e-30</td>
</tr>
<tr>
<td>8D6</td>
<td>GH635849</td>
<td>2</td>
<td><em>Culex quinquefasciatus</em> XP_001842246.1</td>
<td>chitin binding</td>
<td>2e-37</td>
</tr>
<tr>
<td>4B3</td>
<td>GH635481</td>
<td>2</td>
<td><em>Acythosiphon pisum</em> XP_001947981.1</td>
<td>chitin binding</td>
<td>1e-29</td>
</tr>
<tr>
<td>Clone ID</td>
<td>Accession number</td>
<td>Copy number</td>
<td>Blastx top-match in GenBank®: species/accession number</td>
<td>Protein name and/or proposed function</td>
<td>e-value</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------</td>
<td>-------------</td>
<td>--------------------------------------------------------</td>
<td>---------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>1B9</td>
<td>GH635229</td>
<td>1</td>
<td>Drosophila melanogaster/ AAB58341.1</td>
<td>M-spondin, localized at the muscle attachment sites</td>
<td>1e-56</td>
</tr>
<tr>
<td>1D6</td>
<td>GH635249</td>
<td>1</td>
<td>Artemia franciscana/ ABU80629.1</td>
<td>zona pellucida-like domain-containing protein, responsible for sperm-adhesion to the zona</td>
<td>5e-11</td>
</tr>
<tr>
<td>1C10</td>
<td>GH635241</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/ XP_001190623.1</td>
<td>similar to Fat4</td>
<td>7e-16</td>
</tr>
<tr>
<td>2E8</td>
<td>GH635343</td>
<td>1</td>
<td>Strongylocentrotus purpuratus/ XP_001190623.1</td>
<td>similar to Fat4</td>
<td>2e-09</td>
</tr>
<tr>
<td>5G4</td>
<td>GH635625</td>
<td>1</td>
<td>Culex quinquefasciatus/ XP_001850042.1</td>
<td>cyclin G-associated kinase. Uncoating clathrin-coated vesicles from non-neuronal cells</td>
<td>2e-19</td>
</tr>
<tr>
<td>5D10</td>
<td>GH635600</td>
<td>2</td>
<td>Xenopus laevis/ NP_001080800.1</td>
<td>nucleoporin 155kDa</td>
<td>6e-54</td>
</tr>
<tr>
<td>1F2</td>
<td>GH635263</td>
<td>1</td>
<td>Rattus norvegicus/ NP_114010.1</td>
<td>nuclear receptor co-activator 2, ligand-dependent nuclear receptor transcription co-activator activity</td>
<td>6e-33</td>
</tr>
<tr>
<td>5F1</td>
<td>GH635612</td>
<td>1</td>
<td>Danio rerio/ NP_957229.1</td>
<td>phosphatidylinositol transfer protein, alpha</td>
<td>7e-53</td>
</tr>
<tr>
<td>6E2</td>
<td>GH635689</td>
<td>1</td>
<td>Dictyostelium discoideum AX4/XP_646687.1</td>
<td>phosphoethanolamine-cytidyltransferase</td>
<td>6e-63</td>
</tr>
<tr>
<td>8H2</td>
<td>GH635891</td>
<td>1</td>
<td>Culex quinquefasciatus/ XP_001843299.1</td>
<td>rab6 GTPase activating protein, Golgi membrane protein</td>
<td>8e-31</td>
</tr>
<tr>
<td>3H12</td>
<td>GH635467</td>
<td>1</td>
<td>Drosophila sechellia/ XP_002038498.1</td>
<td>GTPase GTP binding protein</td>
<td>1e-36</td>
</tr>
<tr>
<td>5B1</td>
<td>GH635568</td>
<td>1</td>
<td>Homo sapiens/ NP_001091967.1</td>
<td>small G protein signalling modulator 1 protein, modulate small G-protein (RAP and RAB)-mediated signalling pathway</td>
<td>8e-28</td>
</tr>
<tr>
<td>6D8</td>
<td>GH635683</td>
<td>1</td>
<td>Aedes aegypti/ XP_001654070.1</td>
<td>RAN binding protein</td>
<td>8e-11</td>
</tr>
<tr>
<td>1B10</td>
<td>GH635230</td>
<td>1</td>
<td>Aedes aegypti/ XP_0016657320.1</td>
<td>ryanodine receptor 3, brain</td>
<td>2e-56</td>
</tr>
<tr>
<td>7E7</td>
<td>GH635776</td>
<td>1</td>
<td>Xenopus laevis/ NP_001080442</td>
<td>SEC13-like 1, nuclear pore protein</td>
<td>9e-94</td>
</tr>
<tr>
<td>5D6</td>
<td>GH635596</td>
<td>2</td>
<td>Gallus gallus/ NP_001012877.1</td>
<td>vacuolar protein sorting 41</td>
<td>6e-46</td>
</tr>
<tr>
<td>7F11</td>
<td>GH635791</td>
<td>1</td>
<td>Apis mellifera/ XP_395967</td>
<td>karyopherin (importin) alpha 4, mediates the nuclear import of a large variety of substrates</td>
<td>3e-16</td>
</tr>
<tr>
<td>7G12</td>
<td>GH635802</td>
<td>1</td>
<td>Tribolium castaneum/ XP_969706.1</td>
<td>MMP37 (mitochondrial matrix protein 37)-like protein, mitochondrial</td>
<td>1e-39</td>
</tr>
</tbody>
</table>

Intracellular trafficking, secretion, and vesicular transport

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®: species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B10</td>
<td>GH635217</td>
<td>1</td>
<td>Artemia urmiana/ ABY86224.1</td>
<td>artemin, molecular chaperone activity, associates with RNA under high temperature</td>
<td>2e-49</td>
</tr>
<tr>
<td>5A11</td>
<td>GH635566</td>
<td>1</td>
<td>Dictyostelium discoideum AX4/XP_635173.1</td>
<td>chaperonin containing TCP1 eta subunit, folding of proteins, ATPase</td>
<td>5e-91</td>
</tr>
<tr>
<td>Clone ID</td>
<td>Accession number</td>
<td>Copy number</td>
<td>Blastx top-match in GenBank®: species/accession number</td>
<td>Protein name and/or proposed function</td>
<td>e-value</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------------------------------------------</td>
<td>--------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>7A4</td>
<td>GH635734</td>
<td>1</td>
<td>Tribolium castaneum/XP_966546</td>
<td>goliath E3 ubiquitin ligase</td>
<td>1e-20</td>
</tr>
<tr>
<td>8E1</td>
<td>GH635855</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001845910.1</td>
<td>ubiquitin-conjugating enzyme</td>
<td>6e-08</td>
</tr>
<tr>
<td>5A3</td>
<td>GH635558</td>
<td>1</td>
<td>Bombyx mori/XP_001036958.1</td>
<td>p23-like protein, p23 binds heat</td>
<td>2e-15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>shock protein hsp90 and participates in the folding of a number of proteins</td>
<td></td>
</tr>
<tr>
<td>2G6</td>
<td>GH635363</td>
<td>2</td>
<td>Artemia franciscana/ABD19713.1</td>
<td>small heat shock protein</td>
<td>2e-47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ahsp22, abundantly expressed in oviparous-developing embryo</td>
<td></td>
</tr>
<tr>
<td>3E3</td>
<td>GH635425</td>
<td>1</td>
<td>Laccaria bicolor S238N-HB2/XP_001876691.1</td>
<td>WASP (Wiskott–Aldrich syndrome</td>
<td>9e-10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>protein–interacting protein VP1/WIP</td>
<td></td>
</tr>
<tr>
<td>4F3</td>
<td>GH635525</td>
<td>1</td>
<td>Acrhythosiphon pisum/XP_001945864.1</td>
<td>serine protease</td>
<td>1e-72</td>
</tr>
<tr>
<td>7F10</td>
<td>GH635790</td>
<td>1</td>
<td>Tribolium castaneum/XP_969005</td>
<td>HECT (homologous with E6-associated</td>
<td>1e-81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>protein C-terminus), UBA and WWE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>domain containing protein 1, ubiquitin-protein ligases</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorly characterized</td>
<td></td>
<td></td>
<td>General function prediction only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5E11</td>
<td>GH635610</td>
<td>1</td>
<td>Anopheles gambiae str. PEST/XP_317656.3</td>
<td>AGAP007843-PA, MBOAT family, acyltransferase activity</td>
<td>1e-04</td>
</tr>
<tr>
<td>4D8</td>
<td>GH635509</td>
<td>1</td>
<td>Wolbachia endosymbiont of Drosophila ananassae/ZP_00373675.1</td>
<td>ankyrin 3, mediate protein–protein interactions</td>
<td>2e-28</td>
</tr>
<tr>
<td>6A7</td>
<td>GH635650</td>
<td>1</td>
<td>Wolbachia endosymbiont of Drosophila ananassae/ZP_00373675.1</td>
<td>ankyrin 3, mediate protein–protein interactions</td>
<td>1e-12</td>
</tr>
<tr>
<td>1A11</td>
<td>GH635219</td>
<td>2</td>
<td>Aedes aegypti/XP_001653709.1</td>
<td>breast carcinoma amplified sequence, putative spliceosome-associated protein</td>
<td>3e-26</td>
</tr>
<tr>
<td>3C9</td>
<td>GH635408</td>
<td>1</td>
<td>Thermotoga sp. RQ2/YN_001739655.1</td>
<td>chromosome segregation protein SMC</td>
<td>4e-04</td>
</tr>
<tr>
<td>5B2</td>
<td>GH635569</td>
<td>1</td>
<td>Caenorhabditis elegans/NP_001023145.1</td>
<td>Dauer up-regulated family member (dur-1). Predicted to be hydrophilic and heat-resistant, and that might participate in anhydrobiosis</td>
<td>2e-15</td>
</tr>
<tr>
<td>2G12</td>
<td>GH635369</td>
<td>3</td>
<td>Arabidopsis thaliana/NP_193834.1</td>
<td>late embryogenesis abundant (LEA) domain-containing protein</td>
<td>4e-06</td>
</tr>
<tr>
<td>4A8</td>
<td>GH635474</td>
<td>1</td>
<td>Aedes aegypti/XP_001651224.1</td>
<td>defective proboscis extension</td>
<td>2e-61</td>
</tr>
<tr>
<td>6C3</td>
<td>GH635668</td>
<td>2</td>
<td>Apis mellifera/XP_393542.2</td>
<td>falafel CG93951-PA, isofrom A, Rac protein signal transduction</td>
<td>7e-15</td>
</tr>
<tr>
<td>1B2</td>
<td>GH635222</td>
<td>1</td>
<td>Gryllus bimaculatus/BAG71783.1</td>
<td>fragile X mental retardation related protein, fmr. may play a role in mRNA trafficking and localization</td>
<td>2e-60</td>
</tr>
</tbody>
</table>

©The Authors Journal compilation ©2011 Biochemical Society
Reproductive modes determine in *Artemia* oocytes

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank®</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6A1</td>
<td>GH635644</td>
<td>1</td>
<td><em>Oryza sativa Japonica</em> Group/BAC84865.1</td>
<td>helicase-like protein</td>
<td>2e-06</td>
</tr>
<tr>
<td>3C10</td>
<td>GH635409</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em> / XP_001863023.1</td>
<td>harmonin, defect in harmonin caused human Usher syndrome</td>
<td>3e-10</td>
</tr>
<tr>
<td>8H7</td>
<td>GH635896</td>
<td>2</td>
<td><em>Rattus norvegicus</em> / NP_596912.1</td>
<td>limkain b1, peroxisomal autoantigen</td>
<td>1e-24</td>
</tr>
<tr>
<td>2H8</td>
<td>GH635375</td>
<td>3</td>
<td><em>Drosophila mojavensis</em> / XP_002010365.1</td>
<td>membrane-associating (MARVEL) domain-containing proteins</td>
<td>2e-16</td>
</tr>
<tr>
<td>1H7</td>
<td>GH635287</td>
<td>2</td>
<td><em>Mus musculus</em> / EDL00515.1</td>
<td>RNA binding motif protein 26, RBM26</td>
<td>5e-15</td>
</tr>
<tr>
<td>2G3</td>
<td>GH635360</td>
<td>1</td>
<td><em>Aedes aegypti</em> / XP_001653916.1</td>
<td>set domain protein, protein lysine methyltransferases</td>
<td>1e-31</td>
</tr>
<tr>
<td>1F11</td>
<td>GH635270</td>
<td>1</td>
<td><em>Nasonia vitripennis</em> / XP_001606723.1</td>
<td>huntingtin interacting protein</td>
<td>3e-19</td>
</tr>
<tr>
<td>1C3</td>
<td>GH635235</td>
<td>4</td>
<td><em>Cliona intestinalis</em> / NP_001071858.1</td>
<td>zinc finger protein</td>
<td>2e-17</td>
</tr>
<tr>
<td>1B6</td>
<td>GH635226</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em> / XP_001848885.1</td>
<td>zinc finger protein 383</td>
<td>4e-8</td>
</tr>
<tr>
<td>1G5</td>
<td>GH635274</td>
<td>1</td>
<td><em>Mus musculus</em> / EDL07703.1</td>
<td>zinc finger protein 39</td>
<td>6e-30</td>
</tr>
<tr>
<td>5F8</td>
<td>GH635618</td>
<td>1</td>
<td><em>Homo sapiens</em> / EAW04236.1</td>
<td>zinc finger protein 627, isoform CRA_a</td>
<td>1e-10</td>
</tr>
<tr>
<td>4E7</td>
<td>GH635518</td>
<td>1</td>
<td><em>Acyrthosiphon pisum</em> / XP_001949223.1</td>
<td>zinc finger protein</td>
<td>6e-28</td>
</tr>
<tr>
<td>4B2</td>
<td>GH635480</td>
<td>1</td>
<td><em>Drosophila pseudoobscura</em> / XP_002135598.1</td>
<td>ring finger protein</td>
<td>3e-41</td>
</tr>
<tr>
<td>2E11</td>
<td>GH635346</td>
<td>1</td>
<td><em>Acyrthosiphon pisum</em> / XP_001950343.1</td>
<td>similar to NFX1-type zinc finger-containing protein</td>
<td>2e-7</td>
</tr>
<tr>
<td>5E8</td>
<td>GH635608</td>
<td>1</td>
<td><em>Strongylocentrotus purpuratus</em> / XP_001203272.1</td>
<td>similar to zinc finger protein</td>
<td>1e-08</td>
</tr>
<tr>
<td>4F12</td>
<td>GH635533</td>
<td>1</td>
<td><em>Acyrthosiphon pisum</em> / XP_001942984.1</td>
<td>similar to PTK2 protein tyrosine kinase 2 isoform b variant</td>
<td>5e-11</td>
</tr>
<tr>
<td>4B4</td>
<td>GH635482</td>
<td>1</td>
<td><em>Tribolium castaneum</em> / XP_968391.2</td>
<td>similar to Mod(mdg4)-heS00531</td>
<td>2e-06</td>
</tr>
<tr>
<td>6B11</td>
<td>GH635664</td>
<td>1</td>
<td><em>Danio rerio</em> / XP_001341472.2</td>
<td>similar to ATPase, class I, type 8B, member 2</td>
<td>1e-16</td>
</tr>
<tr>
<td>3F9</td>
<td>GH635441</td>
<td>1</td>
<td><em>Drosophila pseudoobscura</em> pseudoobscura / XP_001353315.1</td>
<td>GA19626, transporter activity</td>
<td>1e-45</td>
</tr>
<tr>
<td>3D1</td>
<td>GH635412</td>
<td>2</td>
<td><em>Caenorhabditis elegans</em> / NP_001076772.1</td>
<td>GEX interacting protein family member (gei-3), body morphogenesis</td>
<td>4e-07</td>
</tr>
</tbody>
</table>

Function unknown

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D11</td>
<td>GH635421</td>
<td>1</td>
<td><em>Chelonus inanitus bracovirus</em> / CA098966.1</td>
<td>19.5g1 protein</td>
</tr>
<tr>
<td>8G2</td>
<td>GH635880</td>
<td>1</td>
<td><em>Chelonus inanitus bracovirus</em> / CA098966.1</td>
<td>19.5g1 protein</td>
</tr>
<tr>
<td>4E10</td>
<td>GH635521</td>
<td>2</td>
<td><em>Staphylococcus saprophyticus</em> subsp. saprophyticus ATCC 15305/YP_301831.1</td>
<td>bifunctional autolysin precursor</td>
</tr>
<tr>
<td>1H9</td>
<td>GH635289</td>
<td>1</td>
<td><em>Culex quinquefasciatus</em> / XP_001868827.1</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>5H5</td>
<td>GH635637</td>
<td>1</td>
<td><em>Tribolium castaneum</em> / XP_970889.1</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>Clone ID</td>
<td>Accession number</td>
<td>Copy number</td>
<td>Blastx top-match in GenBank\textsuperscript{®}: species/accession number</td>
<td>Protein name and/or proposed function</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>3A9</td>
<td>GH635386</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001846209.1</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>8F6</td>
<td>GH635872</td>
<td>1</td>
<td>Culex quinquefasciatus/XP_001865003.1</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>3B8</td>
<td>GH635396</td>
<td>1</td>
<td>Artemia franciscana/ABE02251.1</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>5D8</td>
<td>GH635598</td>
<td>1</td>
<td>Aedes aegypti/XP_001660983.1</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>5B11</td>
<td>GH635578</td>
<td>1</td>
<td>Dicystostelium discoideum AX4/XP_640507.1</td>
<td>hypothetical protein DDBDRAFT_0204296, Transient receptor potential (TRP) ion channel</td>
</tr>
<tr>
<td>7G8</td>
<td>GH635798</td>
<td>1</td>
<td><em>Endoriftia persephone</em> 'Hot96_1+Hot96_2'/ZP_02533583.1</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>4C11</td>
<td>GH635500</td>
<td>1</td>
<td>Paramecium tetraurelia/XP_001448738.1</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>3C1</td>
<td>GH635400</td>
<td>1</td>
<td>Xenopus tropicalis/AAl61584.1</td>
<td>LOC549629 protein</td>
</tr>
<tr>
<td>2B6</td>
<td>GH635309</td>
<td>3</td>
<td>Mus musculus/BAC97962.2</td>
<td>mKIAA0467 protein</td>
</tr>
<tr>
<td>3B12</td>
<td>GH635399</td>
<td>1</td>
<td>Danio rerio/CAK04792.2</td>
<td>novel protein</td>
</tr>
<tr>
<td>5H3</td>
<td>GH635635</td>
<td>2</td>
<td>Danio rerio/CAl20706.1</td>
<td>novel protein similar to human MDN1, midasin homologue (yeast)</td>
</tr>
<tr>
<td>6F5</td>
<td>GH635704</td>
<td>2</td>
<td>Nematostella vectensis/XP_001629094.1</td>
<td>predicted protein</td>
</tr>
<tr>
<td>4A3</td>
<td>GH635470</td>
<td>1</td>
<td>Nematostella vectensis/XP_001626044.1</td>
<td>predicted protein</td>
</tr>
<tr>
<td>8F3</td>
<td>GH635869</td>
<td>1</td>
<td>Nematostella vectensis/XP_001623510.1</td>
<td>predicted protein</td>
</tr>
<tr>
<td>7A11</td>
<td>GH635739</td>
<td>1</td>
<td>Myxococcus xanthus/AAF91388.1</td>
<td>SocE</td>
</tr>
<tr>
<td>7H4</td>
<td>GH635806</td>
<td>1</td>
<td>Myxococcus xanthus/AAF91388.1</td>
<td>SocE</td>
</tr>
<tr>
<td>7D4</td>
<td>GH635761</td>
<td>1</td>
<td><em>Vitis vinifera</em>/CA068468.1</td>
<td>unnamed protein product</td>
</tr>
<tr>
<td>2A6</td>
<td>GH635298</td>
<td>1</td>
<td>Tribolium castaneum/XP_967487.2</td>
<td>similar to WD repeat and FYVE domain containing 3</td>
</tr>
<tr>
<td>1E8</td>
<td>GH635260</td>
<td>3</td>
<td>Apis mellifera/XP_396142.3</td>
<td>PREDICTED: similar to CG11877-PA</td>
</tr>
<tr>
<td>8B4</td>
<td>GH635825</td>
<td>1</td>
<td>Tribolium castaneum/XP_968792.2</td>
<td>PREDICTED: similar to AGAP001819-PA, chitin binding domain</td>
</tr>
<tr>
<td>8C11</td>
<td>GH635842</td>
<td>1</td>
<td>Nematostella vectensis/XP_001623606.1</td>
<td>predicted protein</td>
</tr>
<tr>
<td>8E11</td>
<td>GH635865</td>
<td>1</td>
<td>Nematostella vectensis/XP_001638280.1</td>
<td>predicted protein</td>
</tr>
<tr>
<td>3F6</td>
<td>GH635438</td>
<td>2</td>
<td>Tribolium castaneum/XP_973744.2</td>
<td>similar to CG14967 CG14967-PA</td>
</tr>
<tr>
<td>4D10</td>
<td>GH635511</td>
<td>1</td>
<td>Acrythosiphon pisum/XP_001952440.1</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>2C6</td>
<td>GH635319</td>
<td>1</td>
<td>Tribolium castaneum/XP_967072.1</td>
<td>similar to CG1244 CG1244-PA</td>
</tr>
<tr>
<td>8C7</td>
<td>GH635838</td>
<td>1</td>
<td>Nasonia vitripennis/XP_001599119.1</td>
<td>hypothetical protein</td>
</tr>
</tbody>
</table>
Reproductive modes determine in *Artemia* oocytes

Table S3 Continued

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Accession number</th>
<th>Copy number</th>
<th>Blastx top-match in GenBank(^2); species/accession number</th>
<th>Protein name and/or proposed function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5E6</td>
<td>GH635607</td>
<td>1</td>
<td><em>Apis mellifera</em>/XP_394944.3</td>
<td>similar to CG3173-PA</td>
<td>2e-07</td>
</tr>
<tr>
<td>2H9</td>
<td>GH635376</td>
<td>1</td>
<td><em>Ornithorhynchus anatinus</em>/XP_001508615.1</td>
<td>similar to OTTHUMP0000017000</td>
<td>3e-06</td>
</tr>
<tr>
<td>3B1</td>
<td>GH635389</td>
<td>1</td>
<td><em>Trichomonas vaginalis</em>/G3/XP_001579764.1</td>
<td>viral A-type inclusion protein, putative</td>
<td>4e-05</td>
</tr>
<tr>
<td>3G1</td>
<td>GH635445</td>
<td>1</td>
<td><em>Nasonia vitripennis</em>/XP_001600591.1</td>
<td>similar to CG17652-PA</td>
<td>9e-06</td>
</tr>
<tr>
<td>7G10</td>
<td>GH635800</td>
<td>1</td>
<td><em>Tribolium castaneum</em>/XP_975182</td>
<td>similar to CG6312-PA</td>
<td>1e-06</td>
</tr>
<tr>
<td>8E5</td>
<td>GH635859</td>
<td>1</td>
<td><em>Tribolium castaneum</em>/XP_001810599.1</td>
<td>similar to Pol polyprotein</td>
<td>2e-22</td>
</tr>
<tr>
<td>2E9</td>
<td>GH635344</td>
<td>1</td>
<td><em>Tribolium castaneum</em>/XP_001812755.1</td>
<td>similar to rfx5</td>
<td>2e-05</td>
</tr>
<tr>
<td>7B12</td>
<td>GH635749</td>
<td>1</td>
<td><em>Culicoides sonorensis</em>/AAU06491</td>
<td>putative salivary protein</td>
<td>3e-10</td>
</tr>
<tr>
<td>5G3</td>
<td>GH635624</td>
<td>1</td>
<td><em>Drosophila pseudoobscura</em>/XP_001359275.2</td>
<td>GA15060</td>
<td>4e-07</td>
</tr>
<tr>
<td>6A3</td>
<td>GH635646</td>
<td>1</td>
<td><em>Tetrahymena thermophilia</em>/SB210/XP_001031309.1</td>
<td>GCC2 and GCC3 family protein</td>
<td>3e-21</td>
</tr>
<tr>
<td>3G4</td>
<td>GH635448</td>
<td>1</td>
<td><em>Drosophila virilis</em>/XP_002054289.1</td>
<td>Gj22884, activating signal co-integrator 1 complex subunit 3</td>
<td>6e-07</td>
</tr>
<tr>
<td>3C7</td>
<td>GH635406</td>
<td>1</td>
<td><em>Escherichia coli</em>/O157/ZP_03085932.1</td>
<td>iso-Is10R ORF</td>
<td>2e-42</td>
</tr>
</tbody>
</table>

**REFERENCE**


Received 22 October 2009/18 March 2010; accepted 30 March 2010
Published as Immediate Publication 30 March 2010, doi 10.1042/BSR20090141