

Metamorphosis genes are expressed early during *Molgula occulta* and *Molgula oculata* development

Paul Hausch^{a,c} and Billie J. Swalla^{a,b}

BEACON Research Fellowship 2011
Summer 2011

^aFriday Harbor Laboratories, University of Washington, Friday Harbor, WA 98250

^bDepartment of Biology, University of Washington, Seattle, WA 98195

^cDepartment of Biology, Ripon College, Ripon, WI 54971

Key Words: *Molgula*, metamorphosis, genes, *Cornichon*, *Mannose-binding lectin*, *Programmed cell death 2 protein*, *Programmed cell death 6 interacting protein*, *Selectin-like protein 2*, *Tolloid*

Contact Information:

Paul Hausch
Ripon College
Ripon, WI 54971
Phone: (920) 748-8115
hauschp@ripon.edu

Introduction

Ascidians and vertebrates share very similar chordate characteristics during their developmental stages. Since both lineages have highly conserved genomes we are able to make evolutionary comparisons between the two subphyla at the genome level (Dehal et al, 2002). By further exploring metamorphosis in the ascidians, we may be able to develop insight into the evolution of their unusual adult body plan.

Molgula oculata and *Molgula occulta* belong to the class Ascidiacea, or ascidians. Ascidians (sea squirts) belong to the subphylum Tunicata (Urochordata) and are known for their hard cellulose tunic that covers them. Ascidians are part of the clade Chordata, however, they possess most of the chordate characteristics in the larval stage of development. Ascidian larvae swim in the larval stage until they find a suitable substrate to attach themselves. Then, metamorphosis transforms a larval tadpole into a juvenile. Finally, a stationary juvenile ascidian has become a filter feeder invertebrate with two siphons, an incurrent and excurrent one (Davidson and Swalla, 2002).

There are a number of genes associated with the process of metamorphosis in other ascidians, specifically *Boltenia villosa* (Davidson and Swalla, 2002; Davidson et al, 2003). We found that *Molgula oculata* and *Molgula occulta* have very similar genes that in *Boltenia villosa* are expressed during metamorphosis. By selecting various metamorphosis genes in *Boltenia villosa*, we can observe how and where they are expressed in the similarly sequenced *Molgula oculata* and *Molgula occulta* by using whole-mount *in situ* hybridizations for larvae and juveniles. This will reveal how similar or different the process of metamorphosis is between the two species.

The genes I chose are all known to be involved in various events in the metamorphosis of *Boltenia villosa*. The first three are *Cornichon (Bv-Cni)*, *Programmed cell death protein 2 (Bv-PCD2)*, and *Programmed cell death 6 interacting protein (Bv-Aip)*. The last three immune-related genes are *Selectin-like protein 2 (Bv-Sccp2)*, *Mannose-binding lectin (MBL)*, and *Tolloid protein (Ci-BMP1)*, most similar to the *Ciona intestinalis* gene.

In the fly *Drosophila melanogaster*, the *Cni* gene along with *gurken*, determines the dorsal/ventral and anterior/posterior position of the oocyte in development and may also be involved in epidermal growth factor (EGF) signaling (Roth et al, 1995). In *Boltenia villosa*, *Cni* is most active in the anterior region during larval metamorphosis where the expression can be seen one hour into metamorphosis. The expression stayed in the papillary (anterior) region about three days, then the *Bv-Cni* expression ceased (Davidson and Swalla, 2001). Also *Bv-Cni* may help to promote Humps, which induces settlement of larvae (Eri et al. 1999; Davidson and Swalla, 2001). The gene sequence is also very similar to a number of organisms including *Homo sapiens*.

Programmed cell death 2 or the similar *Rp8* in rats has been identified to be involved in programmed cell death (Owens and Cohen, 1992).

Bv-Aip is known to partially regulate apoptosis, or programmed cell death (Vito et al. 1999). *AIP1* working with *ALG-2* are found in the cytosol of cells and depend on calcium to become active in apoptosis. *AIP1* may control the expression of *ALG-2* in metamorphic apoptosis.

Sccp2 is responsible for signaling lectin proteins to adhere and destroy unidentified pathogens in inflamed tissue (Ley, 2001).

MBL aids in the protection from various microbial invaders and is responsible for activating *MBL-activated serine protease (MASP)* by binding to pathogens. *MASP* then sends signals to encode for immune-related proteins active in the lectin pathway (Takahashi et al, 2008).

The *Tolloid* protein (*Ci-BMP1*) function is not well known in *Ciona intestinalis*. In *Drosophila melanogaster* however, *tolloid* activates the Decapentaplegic (DPP) complex. The DPP complex in *Drosophila melanogaster* helps regulate the growth of the imaginal discs that form the limbs and organs (Finelli et al, 1994). Finelli and colleagues also explain that the *Drosophila toloid* gene is similar to the human gene *bone morphogenetic protein 1 (BMP-1)*, and may interact with DPP (*BMP-2* in humans). This similarity of genes and uses of them shows a high rate of homology and conservation between organisms.

This study takes metamorphic genes from *Boltenia villosa*, and observes expression in *Molgula oculata* and *Molgula occulta*, very close relatives with a similar genome. By looking where and when these genes appear during development, we can determine what role they play in this developmental stage for tailless ascidians. We will discuss why these processes occur in ascidians and how these two species have such different larval morphologies. We also will briefly take a look at what these genes do in other organisms, specifically deuterostomes, and discuss why they have been conserved in more specialized organisms over millions of years. Studying the genes of *Molgula oculata* and *Molgula occulta* is important because we can potentially uncover genomic homologies that will further our understanding about the origins of vertebrates.

Methods

Molgula oculata and *Molgula occulta* samples from the Station Biologique in Roscoff, France were used for sequencing. Genes known to be involved in *Boltenia villosa* metamorphosis were selected for *Molgula oculata* and *Molgula occulta* testing (Davidson and Swalla, 2001). Samples were collected in the Swalla lab for the transcriptomes of the metamorphic genes in *Molgula oculata*, *Molgula occulta*, and hybrids. Sequencing was done at Michigan State University in collaboration with C. Titus Brown. We could find areas that were highly conserved and sequences for different alleles by alignment of the gene sequences in Clustal W2 at www.ebi.ac.uk.

The most complete nucleotide sequences of genes were put through BLASTp, to find the protein sequences. By searching 'Multiple alignment' with the protein sequence, organisms with high conservation of the gene could be identified. These organisms could be used to compare function of certain genes.

We made half of the primers using the program primer3, and manually making the other half by choosing from highly conserved sections of the transcriptome sequences. These primers will be subcloned next and used to make probes for in situ.

Results

Cornichon (Cni)

Hybrid.38411

In Figure 1, we show the protein alignment between the *Molgula cornichon* sequence and other deuterostomes. The alignment shows that this gene is highly conserved between hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis* and *Boltenia villosa*), and mammals (*Mus musculus* and *Homo sapiens*). The time of *cornichon* expression shows that there is early expression in *Molgula oculata*, *Mogula occulta*, their hybrid, and *Boltenia villosa*. In *B. villosa*, *cornichon* expression is turned on at the beginning of the embryonic stage, turns off at the end, and then turns back on in the larval stage for metamorphosis (Table 1). Studies by Castro and colleagues show that *cornichon* is involved in TGF-alpha trafficking within the cell and that it is not involved in transmembrane operations (Figure 2, Castro et al, 2007).

Mannose-binding lectin-like protein (MBL)

M.oculata.3951 M-SL-LP.SOC

The *Molgula* gene transcriptome of *Mannose-binding lectin-like protein*, protein alignment compares ascidians (*Polyandrocarpa misakiensis*), mammals (*Monodelphis domestica*, *Mus musculus*, and *Homo sapiens*), and Aves (*Meleagris gallopavo*) to the *molgula* gene sequence and shows that there is relatively high conservation among the listed deuterostomes (Figure 5). The time of *Mannose-binding lectin-like protein* expression shows that there is some late expression in the embryonic stage for only in the hybrid of *Molgula oculata*, *Mogula occulta*. In *Molgula oculata* and

Molgula occulta there is no expression seen in the embryonic stage from data we have at this time (Table 4). The *Endoplasmic reticulum Golgi intermediate compartment 53-kD protein (ERGIC-53)* is the homologous gene found in humans, to the molgulid *MBL* gene. *ERGIC-53* is a L-type lectin because it shares homology to leguminous lectins. It is a mammalian type 1 membrane protein involved in the secretory pathway by exporting of a subset of glycoproteins (Kamiya et al, 2008; Nyfeler et al, 2008).

Programmed cell death 2 protein (PDCD2)

oculata.11977 PCD2.SOC

Molgula gene transcriptome of *Programmed cell death 2 protein*, protein alignment compares hemichordates (*Saccoglossus kowalevskii*), actinopterygii (*Danio rerio*), ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*), to the *molgula* gene sequence. Excluding areas of insertions or deletions, it shows a relatively high conservation of the transcriptome among the listed deuterostomes (Figure 7). The time of *Programmed cell death 2 protein* expression shows that there is some early expression in *Molgula oculata*, and the hybrid in the embryonic stage. However, in *Molgula occulta* there is no seen expression in the data we have at this time (Table 4). In humans and other mammals, this gene is expressed in many tissues. It induces apoptosis in human cells and may protect against lymphoma development (Baron et al, 2010).

Programmed cell death 6 interacting protein (PDCD6IP)

hybrid.63042 PCD6IP.S

Molgula gene transcriptome of *Programmed cell death 6 interacting protein*, protein alignment shows a relatively high conservation of the transcriptome among hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis*), reptilia (*Anolis carolinensis*), and mammals (*Mus musculus* and *Homo sapiens*) (Figure 9). The time of *Programmed cell death 6 interacting protein* expression shows that there is some early expression in *Molgula oculata*, and the hybrid in the embryonic stage. However, in *Molgula occulta* there is no seen expression in the data we have at this time (Table 4).

Selectin-like protein 2 (Sccp2)

hybrid.28816 SLP2.SOC

Molgula gene transcriptome of *Selectin-like protein 2*, protein alignment compares ascidians (*Ciona intestinalis*), amphibians (*Xenopus (Silurana) tropicalis*), and actinopterygii (*Salmo salar*), to the *molgula* gene sequence. This alignment is less conserved among the listed deuterostomes (Figure 11). The time of *Selectin-like protein 2* expression shows that there is early expression in *Molgula oculata*, *Mogula occulta*, and their hybrid. However, in *B. villosa*, *Selectin-like protein 2* expression is off in the embryonic stage and only turns on in the larval stage for metamorphosis (Table 2). In *Ciona intestinalis*, the Carbohydrate-binding domain in this gene has various functions such as enzymatic activity, combating toxicity and signaling transduction (GenPept).

oculata.29132 SLP2.SOC

Tolloid

M.ocu.TolloidSOC

Molgula gene transcriptome of *Tolloid* (*TLD*), protein alignment compares ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*), to the *molgula* gene sequence. This alignment showed the highest level of homology compared to the other alignments of the *molgula* metamorphosis genes (Figure 13). The time of *Tolloid* expression shows that there is early expression in *Molgula oculata*, and the hybrid. In *Molgula occulta* there is no early expression seen in the data we have so far, and in *B. villosa*, *Tolloid* expression is off in the embryonic stage and only turns on in the larval stage for metamorphosis (Table 3). In vertebrates, the *TLD*-like metalloproteinases are involved in making the extracellular matrix (ECM). They are important in the forming the ECM and signaling by TGF-beta-like proteins in morphogenesis. Precursor proteins are made into functional enzymes, structural proteins, and ECM mineralization proteins in hard tissues. The vertebrate transforming growth factor-beta (TGF-beta) and GDF11 (BMP11), can be activated by *TLD*-like metalloproteinases as well (Ge and Greenspan, 2006).

Discussion

The early expression of metamorphosis genes in the ascidians *Molgula occulta* and *Molgula oculata* is an unusual result compared to the ascidians *Boltenia villosa* and *Ciona intestinalis*.

We see early expression of the metamorphosis gene *cornichon* in the embryonic stage for the *molgulids* as well as *Boltenia villosa* (Table 1). In *Boltenia villosa*, *cornichon* expression turns on at the beginning of the embryonic stage, turns off at the end of the embryonic stage, and then turns back on in the larval stage when metamorphosis begins (Table 1, Davidson et al, 2003). *Cornichon* is involved in Epidermal growth factor (EGF) signaling, which lets larvae react to cues for metamorphosis in *Boltenia villosa* (Davidson and Swalla, 2001). We may predict that it has a similar function in other ascidians. *Cornichon* is involved in TGF- α trafficking within the cell in vertebrates, but not in transmembrane operations (Figure 2, Castro et al, 2007). For this gene we also see high homology among deuterostomes, which may point toward more similar function among them to be investigated in the future (Figure 2).

We see early expression of the immune-related gene *Mannose-binding lectin-like protein* in the embryonic stage only for the hybrid of the two *molgulids*. In *M. oculata* there is no expression seen, and for *M. occulta* we also do not see expression for the data we have at this time (Table 4). *Mannose-binding lectin-like protein* binds to pathogens, which activates MBL-activated serine protease (MASP) to send signals to encode for immune-related proteins active in the lectin pathway (Takahashi et al, 2008). We may predict that it has a similar function in other ascidians. For this gene we also see

somewhat high homology among deuterostomes, which may point toward some similar function among them to still be investigated (Figure 5).

We see early expression of the cell-death-related gene *Programmed cell death 2 protein* in the embryonic stage for *M. oculata* and the hybrid of the two *molgulids*. In *M. oculata* there is expression seen near the end of the embryonic stage while expression for the hybrid begins earlier near the middle of the embryonic stage. For *M. occulta* we do not see expression in the data we have at this time (Table 4). *Programmed cell death 2 protein* is involved in human apoptosis and may protect against lymphoma development (Baron et al, 2010). Since we see relatively high homology in this gene among deuterostomes in Figure 7, this may point toward some similar function within deuterostomes.

We see early expression of the cell-death-related gene *Programmed cell death 6 interacting protein* in the embryonic stage for *M. oculata* and the hybrid of the two *molgulids*. In *M. oculata* there is expression seen near the end of the embryonic stage while expression for the hybrid begins earlier near the middle of the embryonic stage. For *M. occulta* we do not see expression for the data we have at this time (Table 4). *Programmed cell death 6 interacting protein* partially regulates apoptosis in *Boltenia villosa* (Davidson and Swalla, 2002). Again we see relatively high homology in this gene among deuterostomes in Figure 9, which may point toward some similar function within deuterostomes.

We see early expression of the immune-related gene *Selectin-like protein 2*, in the embryonic stage for all the *molgulids* but not in *Boltenia villosa*, which is unusual compared to other ascidians (Table 2). In *Boltenia villosa*, *Selectin-like protein 2*

expression is off in the embryonic stage, but turns on in the larval stage for metamorphosis (Table 2, Davidson and Swalla, 2002). *Selectin-like protein 2* is involved in fighting toxins, and is involved in transduction signaling in *Ciona intestinalis* (GenPept). This function in *C. intestinalis* may be similar to the function in the *molgulids* as well and other ascidians. For this gene we see a lower homology, compared to the other metamorphosis genes in this study, among deuterostomes (Figure 11).

We see early expression of the immune-related gene *Tolloid*, in the embryonic stage for *M. oculata* and the hybrid of the two *molgulids*. In *M. oculata* there is expression seen near the beginning of development (3 hours after fertilization) and at six hours after fertilization, while expression for the hybrid is seen throughout the embryonic stage. For *M. occulta* we do not see expression in the data we have at this time (Table 3). In *Boltenia villosa*, *Tolloid* expression is off in the embryonic stage, but turns on in the larval stage for metamorphosis (Table 3, Davidson and Swalla, 2002). *Tolloid* activates the Decapentaplegic (DPP) complex in *Drosophila melanogaster*, which is partially responsible for forming limbs and wings. Also this gene is similar to the human bone morphogenetic protein (BMP). (Finelli et al, 1994). This gene has the highest homology, compared to the other metamorphosis genes in this study, among deuterostomes leading us to think that the function may be very well conserved among many deuterostomes (Figure 13).

Early expression of metamorphosis of this gene could imply that the *molgulid* ascidians are entering metamorphosis early. An early metamorphosis means the larval stage would be cut short which could lead to many events including the loss of a tail. With a shortened larval stage, there would be less time for a tail to develop or for gene

expression of tail/notochord genes. Another idea on tail loss is that cell-death-related genes that are being expressed early which means they are responsible for the death of cells that normally form the tail.

Future studies

There are a number of studies that should be done in the future to uncover more evidence surrounding the question of tail loss in the *molgulid* ascidians. The first step in future studies should be to get the correct size of PCR products. The sizes of the fragments need to be much closer to the sequence sizes we are expecting to see. Next, PCR fragments need to be subcloned and linearized to make probes that will be used for *in situ*. Then finally get to *in situ* hybridizations, which will tell us the location of the genes in the body at a specific time.

Another test to be done would be to inhibit cell-death-related genes in the *molgulids*. If we inhibited the cell-death-related genes in the tailless *molgulids*, then we could observe if the tail would develop. If the tail did not form then we could say that these certain cell-death-related genes are not involved in tail loss of the *molgulids*. However, if a tail formed then we could say that they are responsible because they are not destroying the cells that normally form the tail.

We also need to collect the data for the *molgulid* metamorphosis gene expressions for after the embryonic stage so we can compare it to gene expression in other ascidians.

Finally we need to collect gene expression data for the genes *Programmed cell death protein 2*, *Programmed cell death 6 interacting protein*, and *Mannose-binding lectin* in other ascidians so we can compare this data to the *molgulid* gene expressions.

Acknowledgements

First I'd like to give thanks to Dr. Billie Swalla for inviting me to participate in this REU project in her lab, Sophie George and Adam Summers for orchestrating the REU program this summer, Kristin Andrykovich and Peter X. Wu for teaching me techniques in the lab, Jaquan Horton for helping me prepare my presentation, Brad Shuster for teaching me techniques for observing larva during development, Dr. Titus Brown and Max Maliska for collaboration on this project and providing critical results on gene expression. I also would like to give thanks to ASCB (American Society for Cell Biology) and the BEACON foundation for funding this opportunity. Lastly I would like to thank everyone from Friday Harbor Laboratories for making this such a great learning experience.

References

- Aderem A., R. J. Ulevitch. 2000. Toll-like receptors in the induction of the innate immune response. *Nature* 406:782-787
- Baron B. W., E. Hyjek, B. Gladstone, M. J. Thirman, J. M. Baron. 2010. PDCD2, a protein whose expression is repressed by BCL6, induces apoptosis in human cells by activation of the caspase cascade. *Blood Cells, Molecules, and Diseases* 45:169–175
- Castro C. P., D. Piscopo, T. Nakagawa, and R. Derynck. 2007. Cornichon regulates transport and secretion of TGF- α -related proteins in metazoan cells. *Journal of Cell Science* 120:2454-2466
- Cloney, R. A. 1982. Ascidian larvae and the events of metamorphosis. *American Zoologist* 22:817-826
- Corbo J. C., A. Di Gregorio, M. Levine. 2001. The ascidian as a model organism in developmental and evolutionary biology. *Cell* 106:535-538
- Davidson B., and B. J. Swalla. 2001. Isolation of genes involved in ascidian metamorphosis: epidermal growth factor signaling and metamorphic competence. *Development Genes and Evolution* 211:190-194

Davidson B, and B. J. Swalla. 2002. A molecular analysis of ascidian metamorphosis reveals activation of an innate immune response. *Development* 129:4739-4751

Davidson B., SE. S. Wallace, R. A. Howsmon, B. J. Swalla. 2003. A morphological and genetic characterization of metamorphosis in the ascidian *Boltenia villosa*. *Development Genes and Evolution* 213:601-611

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM et al. 2002. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298:2157-2167

Eri R., J. M. Arnold, V. F. Hinman, K. M. Green, M. K. Jones, B. M. Degnan, and M. F. Lavin. 1999. Hemps, a novel EGF-like protein, plays a central role in ascidian metamorphosis. *Development* 126:5809-5818

Finelli A. L., C. A. Bossie, T. Xie, and R. W. Padgett. 1994. Mutational analysis of the *Drosophila* tolloid gene, a human BMP-1 homolog. *Development* 120:861-870

Ge G. and D. S. Greenspan. 2006. Developmental Roles of the BMP1/TLD Metalloproteinases. *Birth Defects Research (Part C)* 78:47-68

- Kamiya Y., D. Kamiya, K. Yamamoto, B. Nyfeler, H.-P. Hauri, and K. Kato. 2008. Molecular Basis of Sugar Recognition by the Human L-type Lectins ERGIC-53, VIPL, and VIP36. *The Journal of Biological Chemistry* 283:1857–1861
- Ley K. 2001. Functions of Selectins. *Cell Differentiation* 33:177-200
- Nyfeler B., V. Reiterer, M. W. Wendeler, E. Stefan, B. Zhang, S. W. Michnick, and H.-P. Hauri. 2008. Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin. *The Journal of Cell Biology* 180:705-712
- Owens G. P., J. J. Cohen. 1992. Identification of genes involved in programmed cell death. *Cancer Metastasis Reviews* 11:149-156
- Roth S., F. S. Neuman-Silberberg, G. Barcelo, and T. Schüpbach. 1995. *Cornichon* and the EGF Receptor Signaling Process Are Necessary for Both Anterior-Posterior and Dorsal-Ventral Pattern Formation in *Drosophila*. *Cell* 81:967-978
- Takahashi M., D. Iwaki, K. Kanno, Y. Ishida, J. Xiong, M. Matsushita, Y. Endo, S. Miura, N. Ishii, K. Sugamura, and T. Fujita. 2008. *Mannose-Binding Lectin (MBL)-Associated Serine Protease (MASP)-1* Contributes to Activation of the Lectin Complement Pathway. *Journal of Immunology* 180:6132-6138

Vito P., L. Pellegrini, C. Guiet, and L. D'Adamio. 1999. Cloning of *AIP1*, a Novel Protein That Associates with the Apoptosis-linked Gene *ALG-2* in a Ca^{2+} -dependent Reaction. *Journal of Biological Chemistry* 274:1533-1540

Figure Legends

Figure 1. Molgula gene transcriptome of *Cornichon*, protein alignment compares hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis* and *Boltenia villosa*), and mammals (*Mus musculus* and *Homo sapiens*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of deuterostomes, are listed on the left hand side while their protein sequence is to the right of them. Predicted transmembrane domains are underlined. The conserved *Cornichon* domain is boxed in red.

Figure 2. Molgula gene transcriptome of *Cornichon*, protein alignment comparing hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis* and *Boltenia villosa*), and mammals (*Mus musculus* and *Homo sapiens*). Highlighted sections in green show conserved motifs in the transcriptome among all the listed deuterostomes.

Figure 3. Drawing of *Cornichon* location in vertebrates. *Cornichon* is involved in moving TGF- α proteins from the endoplasmic reticulum (ER) to the Golgi apparatus. The cornichon product stays within the cell and is not involved in transmembrane operations.

Figure 4. Molgula gene transcriptome of *Mannose-binding lectin-like protein*, protein alignment compares ascidians (*Polyandrocarpa misakiensis*), mammals (*Monodelphis domestica*, *Mus musculus*, and *Homo sapiens*), and Aves (*Meleagris gallopavo*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of deuterostomes, are listed on the left hand side while their protein sequence is to the right of them.

Figure 5. Molgula gene transcriptome of *Mannose-binding lectin-like protein*, protein alignment comparing ascidians (*Polyandrocarpa misakiensis*), mammals (*Monodelphis domestica*, *Mus musculus*, and *Homo sapiens*), and Aves (*Meleagris gallopavo*).

Highlighted sections in green show conserved motifs in the transcriptome among all the listed dueterostomes.

Figure 6. Molgula gene transcriptome of *Programmed cell death 2 protein*, protein alignment compares hemichordates (*Saccoglossus kowalevskii*), actinopterygii (*Danio rerio*), ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of dueterostomes, are listed on the left hand side while their protein sequence is to the right of them.

Figure 7. Figure 2. Molgula gene transcriptome of *Programmed cell death 2 protein*, protein alignment comparing hemichordates (*Saccoglossus kowalevskii*), actinopterygii (*Danio rerio*), ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*). Highlighted sections in green show conserved motifs in the transcriptome among all the listed

Figure 8. Molgula gene transcriptome of *Programmed cell death 6 interacting protein*, protein alignment compares hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis*), reptilia (*Anolis carolinensis*), and mammals (*Mus musculus* and *Homo sapiens*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of dueterostomes, are listed on the left hand side while their protein sequence is to the right of them.

Figure 9. Molgula gene transcriptome of *Programmed cell death 6 interacting protein*, protein alignment comparing hemichordates (*Saccoglossus kowalevskii*), ascidians (*Ciona intestinalis*), reptilia (*Anolis carolinensis*), and mammals (*Mus musculus* and *Homo sapiens*). Highlighted sections in green show conserved motifs in the transcriptome among all the listed dueterostomes.

Figure 10. Molgula gene transcriptome of *Selectin-like protein 2*, protein alignment compares ascidians (*Ciona intestinalis*), amphibians (*Xenopus (Silurana) tropicalis*), and actinopterygii (*Salmo salar*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of dueterostomes, are listed on the left hand side while their protein sequence is to the right of them.

Figure 11. Molgula gene transcriptome of *Selectin-like protein 2*, protein alignment comparing ascidians (*Ciona intestinalis*), amphibians (*Xenopus (Silurana) tropicalis*), and actinopterygii (*Salmo salar*). Highlighted sections in green show conserved motifs in the transcriptome among all the listed dueterostomes.

Figure 12. Molgula gene transcriptome of *Tolloid*, protein alignment compares ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*), to the “unnamed protein” which is the *molgula* gene sequence, selected as the subject of the alignment. The list of dueterostomes, are listed on the left hand side while their protein sequence is to the right of them.

Figure 13. Molgula gene transcriptome of *Tolloid*, protein alignment comparing ascidians (*Ciona intestinalis*), and mammals (*Mus musculus* and *Homo sapiens*). Highlighted sections in green show conserved motifs in the transcriptome among all the listed dueterostomes.

Table 1. Chart comparing *cornichon* expression in the *molgulids* to *Boltenia villosa*. The Y-axis lists the names of ascidians being compared. The X-axis shows the hours during development, and is also split into three stages, embryonic, larval, and juvenile. “F+#” stands for the number of hours past fertilization. “L+#” stands for the number of hours into the larval stage. “J+#” stands for the number of hours into the juvenile stage. There are also two events that occur “hatching”, between the embryonic and larval stages, and “metamorphosis” between the larval and juvenile stages. The green checks mean expression is found at a certain time. The red dash mean no hits were found when the gene was being transcribed, more tests need to be completed to be sure that there is actually no expression at a certain time. That red X means that there is no expression found at a certain time. The darker circled question marks represent areas where data is expected to come soon to be added to the chart, and the lighter un-circled question marks represent a future area of study.

Table 2. Chart comparing *Selectin-like protein 2* expression in the *molgulids* to *Boltenia villosa*. The Y-axis lists the names of ascidians being compared. The X-axis shows the hours during development, and is also split into three stages, embryonic, larval, and juvenile. “F+#” stands for the number of hours past fertilization. “L+#” stands for the number of hours into the larval stage. “J+#” stands for the number of hours into the juvenile stage. There are also two events that occur “hatching”, between the embryonic and larval stages, and “metamorphosis” between the larval and juvenile stages. The green checks mean expression is found at a certain time. The red dash mean no hits were found when the gene was being transcribed, more tests need to be completed to be sure that there is actually no expression at a certain time. That red X means that there is no

expression found at a certain time. The darker circled question marks represent areas where data is expected to come soon to be added to the chart, and the lighter un-circled question marks represent a future area of study.

Table 3. Chart comparing *Tolloid* expression in the *molgulids* to *Boltenia villosa*. The Y-axis lists the names of ascidians being compared. The X-axis shows the hours during development, and is also split into three stages, embryonic, larval, and juvenile. “F+#” stands for the number of hours past fertilization. “L+#” stands for the number of hours into the larval stage. “J+#” stands for the number of hours into the juvenile stage. There are also two events that occur “hatching”, between the embryonic and larval stages, and “metamorphosis” between the larval and juvenile stages. The green checks mean expression is found at a certain time. The red dash mean no hits were found when the gene was being transcribed, more tests need to be completed to be sure that there is actually no expression at a certain time. That red X means that there is no expression found at a certain time. The darker circled question marks represent areas where data is expected to come soon to be added to the chart, and the lighter un-circled question marks represent a future area of study.

Table 4. Chart comparing *Programmed cell death 2 protein*, *Programmed cell death 6 interacting protein*, and *Mannose-binding lectin-like protein* expression to each other in the *molgulids*. The Y-axis lists the names of ascidians. The X-axis shows the hours during development in the embryonic stage. “F+#” stands for the number of hours past fertilization, and hatching occurs at the end of the embryonic stage. The green checks mean expression is found at a certain time. The red dash mean no hits were found when the gene was being transcribed, more tests need to be completed to be sure that there is

actually no expression at a certain time. That red X means that there is no expression found at a certain time. The circled question marks represent areas where data is expected to come soon to be added to the chart.

Figures

unnamed protein	1	-----SIIILNLP	7
Ciona intestinalis	1	MAFTFAAFCYLALVLTAFILFFAIWQIIAFDELKTDYKNPIDQCNSLNP	80
Boltenia villosa	1	MAFTFAAFCYIAAIIILTAFLIFFAIWQIIAFDELKTDYKNPIDQCNSLNP	80
Saccoglossus kowalevskii	1	MAFTFVAFCYLLAMILSAVLIFFAIYHIIAFDELKTDYKNPIDQCNSLNP	80
Mus musculus	1	MAFTFAAFCYMLALLITAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNP	80
Homo sapiens	1	MAFTFAAFCYMLALLITAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNP	80

unnamed protein	8	LIYHIYRYANRPVMSVTPGLYDPTTVLNADILSYCMKEGWGKLAIFYILSFFYYLYRMIYVLT---	70
Ciona intestinalis	81	LIAYNVYRYSKRPVMSGPLYDPTTIMNADILSYCMKEGWGKLAIFYLLSFFYYLYRMIYVLTVY--	144
Boltenia villosa	81	LIGYNVWRYLHRPVMSAPGLYDPTVMNADVLTTCMREGWCKLAFYLISFFYYLYRMIYVLTVY--	144
Saccoglossus kowalevskii	81	VIGYNIYRYANRPVMSQPLYDPTTIMNADVLSRCMKEGWMKLGFFLLSFFYYLYSMIYVLVSSA	146
Mus musculus	81	LLAYHIWRYMSRPVMSGPLYDPTTIMNADILAYCQKEGWCKLAFYLLAFFYYLYGMIYVLVSS--	144
Homo sapiens	81	LLAYHIWRYMSRPVMSGPLYDPTTIMNADILAYCQKEGWCKLAFYLLAFFYYLYGMIYVLVSS--	144

Figure 1.

Molgula_protein_product	-----	-----SIIILNLP	27
Ciona_intestinalis	1	MAFTFAAFCYLALVLTAFILFFAIWQIIAFDELKTDYKNPIDQCNSLNP	50
Boltenia_villosa	1	MAFTFAAFCYIAAIIILTAFLIFFAIWQIIAFDELKTDYKNPIDQCNSLNP	50
Mus_musculus	1	MAFTFAAFCYMLALLITAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNP	50
Homo_sapiens	1	MAFTFAAFCYMLALLITAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNP	50
Saccoglossus_kowalevskii	1	MAFTFVAFCYLLAMILSAVLIFFAIYHIIAFDELKTDYKNPIDQCNSLNP	50

Molgula_protein_product	-----SIIILNLP	27
Ciona_intestinalis	LVLPEYAIHVFNLLFLWAFEWLTVL	100
Boltenia_villosa	LVLPEYAIHIFFTLLFLWGGWVTVL	100
Mus_musculus	LVLPEYLIHAFFFCVMFLCAAEWLTLG	100
Homo_sapiens	LVLPEYLIHAFFFCVMFLCAAEWLTLG	100
Saccoglossus_kowalevskii	LVLPEYGLHMFFTILFLLAGQFVTL	100

Molgula_protein_product	YDPTTVLNADILSYCMKEGWGKLAIFYILSFFYYLYRMIYVLT---	70
Ciona_intestinalis	YDPTTIMNADILSYCMKEGWGKLAIFYLLSFFYYLYRMIYVLTVY--	144
Boltenia_villosa	YDPTVMNADVLTTCMREGWCKLAFYLISFFYYLYRMIYVLTVY--	144
Mus_musculus	YDPTTIMNADILAYCQKEGWCKLAFYLLAFFYYLYGMIYVLVSS--	144
Homo_sapiens	YDPTTIMNADILAYCQKEGWCKLAFYLLAFFYYLYGMIYVLVSS--	144
Saccoglossus_kowalevskii	YDPTTIMNADVLSRCMKEGWMKLGFFLLSFFYYLYSMIYVLVSSA	146

Figure 2.

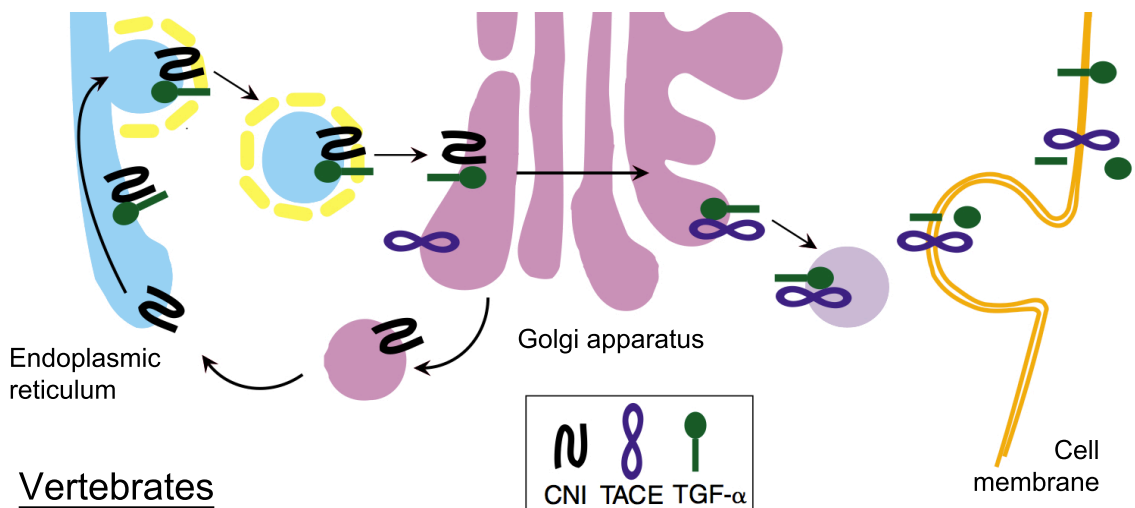


Figure 3.

unnamed protein		-----	
Polyandrocarya misakiensis	1	M-----VRDLV--LKCLLLFLAAVL-----PANENAI-----RRRFEYKHSFKGPHLIQEDLTVPF	50
Monodelphis domestica	1	MAGSRRGGPRAGTRAGPGLLCVLLLSLRSFSSAQGAGDPAAGGGAPAAATVPVQPHRRFEYKYSFKGPHLVQSDGTVPF	80
Meleagris gallopavo		-----	
Mus musculus	1	MAVSRRRVPQAGAR---SFFCALLLSFSQFTGSDGTGGDAAAPGAA--GTQAEHPHRRFEYKYSFKGPHLVQSDGTVPF	74
Homo sapiens	1	MAGSRQRGLRARVR----PLFCALLLSLGRFVRGDGVGGDPAVA-----LPHRRFEYKYSFKGPHLVQSDGTVPF	66
unnamed protein		-----	
Polyandrocarya misakiensis	51	WTL LGDTVAGDDQIRLVPSIRDKGLAQSHYPPHQNWEVEVHIRVEGRGKVGADGLAIWYTSAPPELGSVFGSNNHWTG	130
Monodelphis domestica	81	WAHAGNAIPSSDQIRIAPSLKSKGKSVWTKTKAAFENWEVEVTRVTRGRGRIGADGLAIWYTEGQGLEGPVFGSADNWNW	160
Meleagris gallopavo	1	-----AIP SADQIRITTSLSKSKGKSVWTKNKSIF EYWEVEVTRVTRGRGRIGADGLAIWYTEGQMEGPVFGAADKWNW	74
Mus musculus	75	WAHAGNAIP SADQIRIAPSLKSKGKSVWTKAKA AFENWEVEVTRVTRGRGRIGADGLAIWYTENQGLDGPVFGSADTWNW	154
Homo sapiens	67	WAHAGNAIPSSDQIRVAPSLKSKGKSVWTKTKA AFENWEVEVTRVTRGRGRIGADGLAIWYAENQGLEGPVFGSADLWNW	146
unnamed protein		-----	
Polyandrocarya misakiensis	131	LGVL LDFDND SKLDNPLISVVYNDGTMEFQHGNDGINQYSGSCRFDFRNKPYPVRIKLTYYQEVLT VSYTSGNLPG--DD	209
Monodelphis domestica	161	VGIF FDSFDNDGKKNPAVVIIGNNGQIQYDHQNDGANQALSSCQRDFRNKPYPVRAKITYYQKTLTVMINNGFTPDKED	240
Meleagris gallopavo	75	VGIF FDSFDNDGKXXXXXXXXX-----KITYYQKTLTVMINNGFTPDKED	119
Mus musculus	155	VGIF FDSFDNDGKKNPAIVVIGNNGQIN YDHQNDGATQALASCQRDFRNKPYPVRAKITYYQKTLTVMINNGFTPDKND	234
Homo sapiens	147	VGIF FDSFDNDGKKNPAIVIIIGNNGQIHYDHQNDGASQALASCHRDFRNKPYPVRAKITYYQNTL TVMINNGFTPDKND	226
unnamed protein	1	-----AATGAVADDHVDLSFVTWSILD---VQPEDAGLSPEENQQ--QEEYKNVVEEFNKMHED	54
Polyandrocarya misakiensis	210	FVTC TTVENI VIPPNYFVGSAAATGGLADDDHDLKFLTWSISD--KNMTEESEGLQEDQRQELMDDY EKNL EYDNRMQDE	287
Monodelphis domestica	241	YELCAKVENMVI PAQGHFGISAAATGGLADDDHDLVSLTFQLIEPGKETPAPDKEIPQKDKKEYQEEFEHFQQLDKKKEA	320
Meleagris gallopavo	120	YEFCAKVEDMVLPSQGYFGISAATGGLADDDHDLVSLTFQLTEPGKEVPTSDAEIPEKDKKEYQEEFEHFQQLDKKKEE	199
Mus musculus	235	YEFCAKVENMVIPTQGHFGISAAATGGLADDDHDLVSLTFQLTEPGKEPPTAEKDI SEKEKEYQEEFEHFQQLDKKKEE	314
Homo sapiens	227	YEFCAKVENMIIP AQGHFGISAAATGGLADDDHDLVSLTFQLTEPGKEPPTDKEI SEKEKEYQEEFEHFQQLDKKKEE	306
unnamed protein	55	YKEQHPDQG---IRQPLLEEISGPELQAI FDIQSNVHTDVRSIH--SLM LSLQQQQQT VTNLLQQMDARSSTTTQQQQVP	129
Polyandrocarya misakiensis	288	YKSENFDSYQEQRSPEDLYGGAATQELRSVFEIQSAIKKEIRNLAEVVGHMLNQNNLYEKLDN--ANIGATADGQ---P	362
Monodelphis domestica	321	FQKDHPI--QGQPAEEIFESVSDRELRFIFEGQNRHLEIKQLNRQLDMILDEQRRYVSALTEELSKRGPSPAQ---L	395
Meleagris gallopavo	200	FQKEHPDV--QGQPADDPFETVSDRELRFIFEGQNRHLEIKQLNRQLDMILDEQRRYVS AVTEEIAKRGAAVSGQ---Q	274
Mus musculus	315	FQKGHPDL--QGQPADDIFESIGDRELRFVFEQNRHLEIKQLNRQLDMILDEQRRYVSSLTEEISRRGAGTPGQ---P	389
Homo sapiens	307	FQKGHPDL--QGQPAEEIFESVSDRELRFVFEQNRHLEIKQLNRQLDMILDEQRRYVSSLTEEISKRGA MPGQ---H	381
unnamed protein	130	PSGGDDNAKNYHIENVMLQREVIDSLNVRSTMDTVQSSIIAVSQVSSSKGTGQQQATGGAGDQYAMQQ---VQRQL	205
Polyandrocarya misakiensis	363	PAGADDNAKKYHIDNVLNMQRETYDQVKGMRTELGRFEQIMNMAVQLQRRPKIEGGSNTGTQNS--YEVVQQQAVNEKL	441
Monodelphis domestica	396	-----GQTPPQELDTVVKTQEQVELRQVNMKNMSETLRL-----VSGSQHTGTAGSVYETTQHFNDIKEHL	457
Meleagris gallopavo	275	-----GQVSQREIDTVVKTQEEVIKQVNMKNMADTLRL-----ISGAQHPGSAGVGVYETTQHFNDIKEHL	336
Mus musculus	390	-----GQVSQELDTVVKSQEILRQVNEVKNMSETVRL-----VSGIQHPGSAG--VYETTQHFNDIKEHL	450
Homo sapiens	382	-----GQITQEQELDTVVKTQHEILRQVNMKNMSETVRL-----VSGMQHPGSAGGVYETTQHFNDIKEHL	443
unnamed protein	206	TAVQGD LQALANKKQPT--PQLRCPE--PEMNCSSIPAVVLIIVQIIILVIYHQIKSSRENAAKKFF	270
Polyandrocarya misakiensis	442	HRIQNDLTVMMQRQPPT--PKLVCPE--PEMNCISIGVFVA VAVVQVILIIYHSTRTRSREDAAKKFF	506
Monodelphis domestica	458	HVVKRDIEHLIQRNVPSNERPKPCPELPPFPSCSLSTHFLIFVAVQTMFLIGYIMYRTQQEAAAKKFF	524
Meleagris gallopavo	337	HVVKRDIEHLVQRNMP SNEKPKCPDLPFPSCSLSTMHFFIFVAVQTVL FVAYIMYRSQQEAAAKKFF	403
Mus musculus	451	HVVKRDIDSLAQRSMPSNEKPKCPDLPFPSCSLSTIHVFI FVVVQTVL FVGYIMYRTQQEAAAKKFF	517
Homo sapiens	444	HIVKRDIDNLVQRNMP SNEKPKCPDLPFPSCSLSTVHFI FVVVQTVL FVGYIMYRSQQEAAAKKFF	510

Figure 4.

unnamed protein product	1	-----TDVDLGF AEPVKS PWQLHRCFFPSKIGKPSWLNPDNLPTVSELQCEICSNPLVFLQVYAPDENSEN	68
Danio rerio	1	MAANTLKESNTHSDVVLGFLEEAES-WQLLSDQFPFSKVGGRPAWLSQSDLPVAVSELQCEECKLPAVFLQVYAPVTEYDR	79
Ciona intestinalis	1	-----MELGFVEEPENTWRLLRCYFPSKVGGKPAWLDPEHLPSVDDMECGCCHKPLVFLQVYSPNDDKLD	66
Saccoglossus kowalevskii	1	MAS----PI----DTELGFVEKVDC-WRLNSKFFPSKVGGKPSWLRRLKNLPAENLLQCHNCKKQCVFLQVYAPVTEKES	71
Homo sapiens	1	MAAAGARP-----VELGFAESAPA-WRLRSEQFPSKVGGRPAWLGAAGLPGPQALACELCGRPLSFLQVYAPLPGRPD	73
Mus musculus	1	MAAAPGP-----VELGFAEEAPA-WRLRSEQFPSKVGGRPAWLGAAELPGPGALACARCGRPLAFLQVYAPLPGRDD	73
unnamed protein product	69	AFHRTIFIFCCKTDCYKNSAGVVPVKVFRQCLSKNNFYPEDAPKYDDKNYISTIKTT----SLKLCDCVCGCRGNHCS	143
Danio rerio	80	CFHRTL FVFCCKTPACYTR-NDSKCFKVFRLPRKNEFYFNFPPDEKPEQPVHDAQV-LGSGLKLCLRCGLGQKACS	157
Ciona intestinalis	67	SFHRTIFVFCCTNGTCYSSENDAPFKVFRSMLPKENDYYPADPPDYDEPNAIRESQS----TNLCGLCGNHGVKCS	141
Saccoglossus kowalevskii	72	CFHRTVIFCCRNPQCYSS-SNSPPIVVFRCQLPRINEFYDFEPPEEKENGGLSPDDF----GQKLCVVCAGPKQCG	146
Homo sapiens	74	AFHRCIFLFCCEQPCCAG-----LRVFRNQLPRKNDFYSYEPPESENPPPETGESVCLQLKSGAHLRCVCGCLGPKTCS	147
Mus musculus	74	AFHRSFLFCCEQPCCAG-----LRVFRNQLPRNNAFYSYEPPESEALGT-ECVCLQLKSGAHLRCVCGCLAPMTCS	146
unnamed protein product	144	RCKKVNYCSKHLHQLHWK-QHKKVCSEISSSEAE-NDLNGLNFATFEIVMEPEIKINEAESKR-----Q	205
Danio rerio	158	RCHSVTYCCKEHQTTDQKQRHKKELAEASQVSG--ELNSFLFPEWELVTEP--EVI PA-----KDELQESPSLDQ	224
Ciona intestinalis	142	RCREVQYCSKDHQVFWHKLHXYQCSKPKDENNEVEMEKSLFSEKELVIEPEPTAPKKENTE-----SEFEN	209
Saccoglossus kowalevskii	147	KCHKVTYCSRHDHTVDRWGHKINCGNQCEMNV---SKLLFPEYEVTEAEEYEEPEENEKCEADKMKYAEFMESEE	222
Homo sapiens	148	RCHKAYYCSKEHQTLDWRLGHKQAC-AQPDHLDHIPDNHFLFPEFEIVETETEIMPE-----VVEKE----DY	212
Mus musculus	147	RCKQAHYCSKEHQTLDWRLGHKQAC-TQSDKIDHMPDHNFLFPEFEIVTETEDELPE-----VVEME----DY	211
unnamed protein product	206	DELPNNSASLSAADLEQFANAGNKADKYFIEFKEKLNNGHEQVLRVYHKSQEPLWCACDGKLTEDVPLCGSGGSRIFEF	285
Danio rerio	225	ENIASLNSGLEDESELSMALHETLDSKVFQFKQRIANEPQVLRVYCRGGSPLWVTAEHVPREEEVPEC-TCGAKRLPEF	303
Ciona intestinalis	210	LKNHIQSGDYTEELND-AVANVKSDFKQKQIAVEPEQVIRYQREGSVLWCSSENIPSK--IPSC-ECGAQRQFEF	285
Saccoglossus kowalevskii	223	AKSVMKDRGLTDHDLERMAHE--NDEQFTVFKRINHEPEQVLRVYCRGGSPLWVSSSENIPKDEEIPAC-QCGAKRKFPEF	299
Homo sapiens	213	SEIIGSMGEALEEELDSMAKHESREDKIQFKFTQIALEPEQILRYGRGIAPWIWISGENIPQEKDIPDC-PCGAKRILEF	291
Mus musculus	212	SEVTGSMGGIPEEELDSMAKHESKEDHIFQFKSKIALEPEQILRYGRGIKPIWISGENIPQEKDIPDC-PCGAKRIFEF	290
unnamed protein product	286	EIMPQLLNLYLNIDQGF-TDGDWGVITITCKNSCLIKSDGYTSEYAHVQNFS----- 337	
Danio rerio	304	QIMPQLLNHLKVDST--DASIDWGTVAIYTCAESCDQ-GNKYSPEFIWKQDFSGDQAE 358	
Ciona intestinalis	286	QIMPQLLNHLKVDHSIEGPTIDWGTVAIYTHDDCNI--SPYIKYCWLQHFSEKAI- 340	
Saccoglossus kowalevskii	300	QIMPQLLIHLNVDSI--GESLDWGTGIIYTCSDSCD--GISYHQEYVWKQDYSGNK- 352	
Homo sapiens	292	QVMPQLLNLYLKADRL--GKSIDWGI LAVFTCAESCSL-GTGYTEEFVWKQDVDTTP-- 344	
Mus musculus	291	QVMPQLLNHLKADRL--GRSIDWGLAVFTCAESCSL-GSGYTEEFVWKQDVDTTP-- 343	

Figure 6.

```

Homo_sapiens      MAAAGARP-----VELGFAESAPA-WRLRSEQFPSKVGGRFAWLGAAGL 43
Mus_musculus     MAAAAPGP-----VELGFAEEAPA-WRLRSEQFPSKVGGRFAWLGLAEL 43
Danio_rerio      MAANTLKESNTHSDVVLGFLEEAES-WQLLSDQFPSKVGGRFAWLSQSDL 49
Saccoglossus_kowalevskii MAS-----PI-----DTELGFVEKVDC-WRLNSKFFPSKVGGRFSWLRLKNL 41
Ciona_intestinalis -----MELGFVEEPENTWRLLRCYFPSKVGGRFAWLDPEHL 36
Molgula_protein_product -----TDVDLGFAEPVKSPWQLLRCFFPSKIGGRFSWLNPDNL 38
                *** *      *:*      *****:***:*** *

Homo_sapiens      PGPQALACELGRPLSFLLQVYAFLPGRPDAFHRCIFLFCCREQPCCAG- 92
Mus_musculus     PGPALACARGRPLAFLLQVYAFLPGRDDAFHRSLFLFCCREPPCCAG- 92
Danio_rerio      PAVSELQCEEKLPAVFLLQVYAFLVTEYDRCFHRTLFLFCCKTPACYTR- 98
Saccoglossus_kowalevskii PAENLLQCHNCKQCVFLLQVYAFLVTEKEESCFHRTVFLFCCRNPQCYSS- 90
Ciona_intestinalis PSVDDMECGCHKPLVFLLQLYSENDDKLDSFHRTIFLFCCTNGTCYSSE 86
Molgula_protein_product PTVSELQCEISNPLVFLLQVYAFLEDENSENAFHRTIFLFCCKTDSCYKN- 87
                * : * *      *****:***      .*** :*:*** *

Homo_sapiens      -----LRVFRNLPRKNDFYSYEPPSENPPPETGESVCLQLKSGAHLCRV 137
Mus_musculus     -----LRVFRNLPRNAFYSYEPPSETEALGT-ECVCLQLKSGAHLCRV 136
Danio_rerio      NDSKCFKVFRSQLPRKNEFYPFNPPPDEKPEQPVHDAQV-LGSGLKLCRL 147
Saccoglossus_kowalevskii SNSPPIVFRCQLPRINEFYDFEPPEEKENGLLSPDDF----GQKLCVV 136
Ciona_intestinalis NDAFPFKVFRSMLPKENDYYPADPPDYDEPNAIRECSQF----TNLCGL 131
Molgula_protein_product SAGVPVKVFRCQLSKKNFYPEDAPKYDDKNYISTIKTT----SLKLCDV 133
                . *** *.: * : * :.*      :** :

Homo_sapiens      CGCLGPKTSRCHKAYYCSKEHQTLDWRLGHKQAC-AQPDHLDHIPDHN 186
Mus_musculus     CGCLAPMTSRCQAHYYCSKEHQTLDWRLGHKQAC-TQSDKIDHMVPDHN 185
Danio_rerio      CGCLGQKACSRCHSVTYCKEHQTDWKQRHKKECLAEASQSG--ELNS 195
Saccoglossus_kowalevskii CGCAGPQKCGKCHKVTYCSRDHQTVDWRDGHKINCGNQGCEMNV----SK 182
Ciona_intestinalis CGNHGVKKSRCREVQYCSKDHQVFHWKLLHKYQCSKPKDENNEVEMEKS 181
Molgula_protein_product CGCRGPNHCSRCKKVNYCSKLHQTLHWK-QHKKVCSEISSEAE-NDLNG 181
                ** . *.:**.. **.:**.*.:** *

Homo_sapiens      FLFPEFEIVIITEDEIMPE-----VVEKE-----DY 212
Mus_musculus     FLFPEFEIVITEDEILPE-----VVEME-----DY 211
Danio_rerio      FLFPEWELVTP--EVIPA-----KDELQES-----PSLDQ 224
Saccoglossus_kowalevskii LLFPEYEIVTEAEYEEPEENEKCEADKMKEYAEF-----MESE 222
Ciona_intestinalis ILFSEKELVIEP-PTAPK-----KENTE-----SEFEN 209
Molgula_protein_product LNFATFEIVMEPEIKINEA-----ESKR-----Q 205
                : * . ** * * .

Homo_sapiens      SEIIGSMGEALEEELDSMAKHESREDKIQKFKTQIALEPEILRYGRGI 262
Mus_musculus     SEVTGSMGGIPEEELDSMAKHESKEDHIQKFKSKIALEPEILRYGRGI 261
Danio_rerio      ENIASLNSGLEDSELSMALHETLDSKVQKFKQRIANEPQVLRYCRGG 274
Saccoglossus_kowalevskii AKSVMKDRGLTDHDLERMALHE--NDEQFTVFQKRINHEPEVLRYNRGG 270
Ciona_intestinalis LKNHIQSGDYTEELND-AVANVKSDKAFEKFQKIAVEPEVIRYQREG 258
Molgula_protein_product DELPNNSASLSAADLEQFANAGNKADKYEIEKEKLNGNHEVLRYHKSG 255
                :      :*: *      .. * *:: :*:*** :

Homo_sapiens      APIISGENIPQEKDIFDC-PCGAKRILEFQVMPQLLNYLKADRL--GS 309
Mus_musculus     KPIISGENIPQEKDIFDC-PCGAKRIFEFQVMPQLLNHLKADRL--GR 308
Danio_rerio      SPLWVTAEHVPREEVEFC-TCGAKRLFEFQIMPQLLNHLKVDST--DA 321
Saccoglossus_kowalevskii KPLWVSSENIPKDEEIFAC-QCGAKRKFEFQIMPQLLIHLNVDSI--GE 317
Ciona_intestinalis SVLWCSSENIPSK--IPSC-ECGAQRQFEFQIMPQLLNHLKVDHSIEGPT 305
Molgula_protein_product EPLWCADGLKTDVLCGSCGSRIFEFEIMPQLLNYLNIDQFTDG 304
                :* : :      * *      **.* :*:*** :*: *

Homo_sapiens      IDWGILAVFTCAESCSL-GTGYTEFVWKDVTDTP-- 344
Mus_musculus     IDWGVLAVFTCAESCSL-GSGYTEFVWKDVTDTP-- 343
Danio_rerio      IDWGTVAIYTCAESCDQ-GNKYSPEFIWKDFSGDQAE 358
Saccoglossus_kowalevskii LDWGTLGIYTCSDSCD--GISYHQEYVWKDYSNGNK- 352
Ciona_intestinalis IDWGTVAIYTCHDDCNI--SPYIKEYCWLHFSKEAI- 340
Molgula_protein_product IDWGVITIYTCKNSCLIKSDGYTSEYAHVDNFS---- 337
                :*** : :*: * .      * * : * . :

```

Figure 7.

unnamed protein product	1	MAYTVDVGRFIVVPLKKTDDVDLVKPLQTFIKNTY----DKVDDDLKXHKDFNSLRKTAIVYKAQDKSAAGLNVLQYY	75
Ciona intestinalis	1	MSYEVELSRFFNIPIKRSHDVFVKPLETFIKVTF----DKTDDDLKXHKDFNSLRKTAIVYKAQDKSAAGLNVLQYY	75
Anolis carolinensis	1	MASP---SSFITVQLKASEVDLAKPLSKFIQQTYY--P-GGESQAHECRAAEELNKLKRSALGRSLDKHESSELTLLRY	74
Homo sapiens	1	M-----ATFISVQLKKTSEVDLAKPLVKFIQQTY--PSGEEQAQYCRAAEELSKLRRAAVGRPLDKHGALETLLRY	72
Mus musculus	1	M-----ASFIVVQLKKTSEVDLAKPLVKFIQQTY--PSGEEQAQYCRAAEELSKLRRALGRPLDKHGALETLLRY	72
Saccoglossus kowalevskii	1	MA----TSFVSVPLKVSSEVDLNRPFKTFIENTYGVVEGGED---YGQAIKEFTKRLTRNVCKALDKHESLDSLHRY	72
unnamed protein product	76	DQFSSLNRKVPFCSENGVHVSTFKWDSLEKGLSFGSGKITSASGEFERLMTIYINIAALKSQIASEANLSNDELKSAK	155
Ciona intestinalis	76	DQLVAASRKFPPREANGVIVNFTWKDSLQKGLSSTPKLSIADGDFERLCILYINIAALMSQVGEANLQTDGKLSAAK	155
Anolis carolinensis	75	DQLCSIEPKFPFSE-NQICVTFWTKDAFDKGLSFGSVKLALASLYEKTCLVFNCAALASQIASEQNLDSEGLKAAK	153
Homo sapiens	73	DQICISIEPKFPFSE-NQICLFTWTKDAFDKGLSFGSVKLALASLYEKSCLVFNCAALASQIASEQNLDNDEGLKIAAK	151
Mus musculus	73	DQICISIEPKFPFSE-NQICLFTWTKDAFDKGLSFGSVKLALASLYEKSCLVFNCAALASQIASEQNLDNDEGLKTAAK	151
Saccoglossus kowalevskii	73	DQLTAIDAKLPITE-NQVXKTFWTKDAFDKGSFFGAKKQSGSTAGFERKGVLFNIGAMHSQIASVQSELEDEGLKKAAR	151
unnamed protein product	156	HFEEAAGMLDFIKDNLIASI-NQNPTSDLSFNVLAFRCILIAQSQECFYKASSDK---MKPDILAKVANQAASLYADA	231
Ciona intestinalis	156	YFQEAAGILTFIKERVMSVVGNSNITSDLAPEVLNLYRCVLGQAQECFYDKACTDASLNRKPEVLAKVAMQASLYSEA	235
Anolis carolinensis	154	HYQFASGAFQHIKDTVLSL-NREPTVDIAPDTVGTLSLIMLAQAQEVFFLKATRDK---MKDAI IAKLGNQAADYFGDA	229
Homo sapiens	152	HYQFASGAFQHIKDTVLSL-SREPTVDIAPDTVGTLSLIMLAQAQEVFFLKATRDK---MKDAI IAKLANQAADYFGDA	227
Mus musculus	152	YQFASGAFQHIKDTVLSL-SREPTVDIAPDTVGTLSLIMLAQAQEVFFLKATRDK---MKDAI IAKLANQAADYFGDA	227
Saccoglossus kowalevskii	152	HFQLSAGIYSYLRDHYVTLT-PTIRTPDIEPDLVLSLAIMLAQAQEVFFKAVMDK---MKSVVIAKIANQAADLYADA	227
unnamed protein product	232	VSALSEAGQ-MLPSNVSTMCNTKKSLEYVLHFMASVAHNARKFGEERGLQKIEPSLKLPGKHPQLIQGFNM---KDLA	308
Ciona intestinalis	236	SKAFTEAESYSLAGVQAMCGAKKELFECYQYHQATAAQEGKRFGEAIARYQIVPEPMARNLHNKNAEYL-VNFKGKQFI	314
Anolis carolinensis	230	FKQCQYKDT-LPK-EVFPFLAAKHCIQMANAEYHQSLAKQKQKFGEEIARLQHAADLVKTVASRYDEY---INI--KDMA	303
Homo sapiens	228	FKQCQYKDT-LPK-EVFPVLAAKHCIMQANAEYHQSLAKQKQKFGEEIARLQHAALIKTVASRYDEY---VNV--KDFS	301
Mus musculus	228	FKQCQYKDT-LPK-EVFPFLAAKHCIMQANAEYHQSLAKQKQKFGEEIARLQHAALIKTVASRYDEY---VNV--KDFS	301
Saccoglossus kowalevskii	228	LKSC--NET-IPK-DWYSILAGKTPYFQGEAEYHQSLVCKAANKFGEESRLRHAELIEAHRNGGLS---VSF--KDTA	299
unnamed protein product	309	LKVSQVLEKAEQENDFIYHDPPI--KELSAIGKAAIKAMPFNATEALGGKYVDVFTSLVPLAVHTAIESYENRRKQIIE	386
Ciona intestinalis	315	EMVSTIRGTTEKDNFIYNDVVPKTSLSLPIGKAPIAKAKAFNEKENLGAFTDMFSSLVPLDVHNALEAYEGRRKELVE	394
Anolis carolinensis	304	DKTRNLAATAAKDNDFIYHDIRPDLKLEPIGKASLVKPTQAVP--LSQKFTDLFEKMPVLAQVALQTLTYNQRKADLVN	381
Homo sapiens	302	DKINRDLAAAKDNDFIYHDIRPDLKLDPIGKATLVKSTPVNVP--ISQKFTDLFEKMPVSVQQLAAVYQNRKADLVN	379
Mus musculus	302	DKINRDLAAAKDNDFIYHDIRPDLKLDPIGKATLVKSTPVNVP--VSQKFTDLFEKMPVSVQQLAVFSQRKADLVN	379
Saccoglossus kowalevskii	300	AKIRKRYDESKDNDFIYHERVDPVGLAAIGKAPIAKAPISPTP--MGAQFTDLFEKLVPLSVHQSFTSFENRKAIEVN	377
unnamed protein product	387	REVARIRELNTLINGMSSLNLPAALEDISGDSVPQSIIEKQSAIVISKGLQSIENLLQELPELSTRNQQLLENCMRMMN	466
Ciona intestinalis	395	KEVERIRELNNLVNGVMSSFNLPAALEDMSGDLPSPSILEKSRVTIQKGSAPIDQLDLPESLNRNQQLLETCVRILN	474
Anolis carolinensis	382	RSIAQMREATNLAVGLASLNLPAAIEDVSGDSVPQSIILQKSKSVIEQGGIQTIDQLMKDLPPELLQRNREILDESRLID	461
Homo sapiens	380	RSIAQMREATNLAVGLASLNLPAAIEDVSGDTPVQSIILTKSRSVIEQGGIQTVDQLIKELPELLQRNREILDESRLID	459
Mus musculus	380	RSIAQMREATNLAVGLASLNLPAAIEDVSGDTPVQSIILTKSTSVEQGGIQTVDQLIKELPELLQRNREILDESRLID	459
Saccoglossus kowalevskii	378	REIKGLRSKTEFLNGVLSFNLPAAIEDISGNVPPSLEKSRNMQQGGLHALDQMISELPPELLQRNREILDESIRLLD	457
unnamed protein product	467	EEQSTDELRKFKGATWNRTPS-----	488
Ciona intestinalis	475	DEQSSDAELREKFGDKWRKPSAELTVPLRQEMSKYQTIENAQKADAIIRGKYEDNKEGISLCLPPDQLKNAIPSSSP	554
Anolis carolinensis	462	EEEATDNLKKNFKERQRTPSNDLYKPLRTEGNSYRNI LDKAVQADKQKERYQSHRDTIISLCKPEALNAAIPSNAP	541
Homo sapiens	460	EEEATDNLRAKFKERQRTPSNELYKPLRAEGTNFRVLDKAVQADGQVKECYQSHRDTIIVLCKPEPELNAIPSNAP	539
Mus musculus	460	EEEATDNLRAKFKERQRTPSNDLYKPLRAEGAKFRAVLDKAVQADGQVKECYQSHRDTIIVLCKPEPELNAIPSNAP	539
Saccoglossus kowalevskii	458	KEQSSDDMRGQFKERWTRVPSQQLTEPLRAEAAKYGILDNAIQADHIVKEKYNAHKDGFILMNKSESDLAGIPTAGP	537
unnamed protein product	555	VAAKLGHPVQELKQCEQVNTLKAERVMETELRSASVDMTDKFMRLVAEEGCINEESISPGLDRIYQPLIEQVNDNS	634
Ciona intestinalis	542	AKTMQGSSEVNVLSKLLNNDVKKQREQLENDLKSVMFDMTKFLTAQAQDGALENEAISVTELDRIYGLLTHKVOEQL	621
Anolis carolinensis	540	AKTMQGSSEVNVLSKLLSNLDEVKERELENDLKSVMFDMTKFLTAQAQDGVINEEALSVELDRYVGLLTHKVOEQL	619
Homo sapiens	540	AKTMQGSSEVNVLSKLLSNLDEIKKERESLENDLKSVMFDMTKFLTAQAQDGVINEEALSVELDRYVGLLTHKVOEQL	619
Mus musculus	540	AKTMQGSSEVNVLSKLLSNLDEIKKERESLENDLKSVMFDMTKFLTAQAQDGVINEEALSVELDRYVGLLTHKVOEQL	619
Saccoglossus kowalevskii	538	TATPQGSSEVNVLSKLLSNLDEIKKERESLENDLKSVMFDMTKFLTAQAQDGVINEEALSVELDRYVGLLTHKVOEQL	617
unnamed protein product	635	GRQESVLANLQRANEQPAVLK-SGGGQNKREDVLKKSIAAFDRYEELTAHLREGNKFYNDLTELKLPKCSDFVFARK	713
Ciona intestinalis	622	KKQEELNLIQNSHQEFSKMKQSNNEANLREEVLNKLAVANDNFVELVANLKEGTFKYNELTEILLKFNKCSDFVFARK	701
Anolis carolinensis	620	KKQEGLLKNIQNSHQEFSKMKQSNNEANLREEVLNKLATAYDNFVELVANLKEGTFKYNELTEILLVRFNKCSDIVFARK	699
Homo sapiens	620	KKQEGLLKNIQNSHQEFSKMKQSNNEANLREEVLNKLATAYDNFVELVANLKEGTFKYNELTEILLVRFNKCSDIVFARK	699
Mus musculus	620	KKQEGLLKNIQNSHQEFSKMKQSNNEANLREEVLNKLATAYDNFVELVANLKEGTFKYNELTEILLVRFNKCSDIVFARK	699
Saccoglossus kowalevskii	618	ERQDALLANIERTNQPCKEKQSNAGAQRREMLRKLAAAYDAFMELLANLKEGTFKYNELTEILLVRFNKCSDIVFARK	697
unnamed protein product	714	TEKEELLKDVQQTIVGTPASTAPNVPAYHQQQDVYAEVKKPPARPPPTAPATQAPPTAQQTQVQPKPAPRKSAPAPPT	793
Ciona intestinalis	702	TERDELLKDLQOSIAREP-----SAPSFPSVPAY--QTGPAGNKPPASS	744
Anolis carolinensis	700	TERDELLKDLQOSIAREP-----SAPSIPT-PAY--QSSPAGGH---APT	738
Homo sapiens	700	TERDELLKDLQOSIAREP-----SAPSIPT-PAY--QSSPAGGH---APT	738
Mus musculus	700	TERDELLKDLQOSIAREP-----SAPSIPT-PAY--QSSPAGGH---APT	738
Saccoglossus kowalevskii	698	TEKEELMKDLQQINQIP-----TGPTPAA-PSH--HGGQASRQ---PP	736
unnamed protein product	794	PNAPAAQGPASAPGAPGAPGAQGGPPQYAHQQPPYGYGGAPPPQGYAPFPVGYGAPAAAPYQPPYGYPPQGGYAP	873
Ciona intestinalis	745	VPTPAPRTSVANKQPAPRPPVISA--ANTASAPG--PAAAGTAPSASSAMGTA-PSQAQ-----GPPYPTYPGY	812
Anolis carolinensis	739	PPTPAPRTMPTKQPAPRPPVLPANRAPSATAPS-PVAGATAAPAPSQTGPSAPPPQAQ-----GPPYPTYPGY	809
Homo sapiens	739	PPTPAPRTMPTKQPAPRPPVLPANRAPSATAPS-PVAGATAAPAPSQTGPSAPPPQAQ-----GPPYPTYPGY	810
Mus musculus	739	PPTPAPRTMPTKQPAPRPPVLPANRAPSATAPS-PVAGATAAPAPSQTGPSAPPPQAQ-----GPPYPTYPGY	810
Saccoglossus kowalevskii	737	PRPPPSVQPPYQVQQQQPQPPYAPA---YSQYTPPMPGTGYNPAPMNYPAQPPYPPQ-----GQP-----	797
unnamed protein product	874	QGYGAYPPQPGY-----PTQPGGYPPQPPQPP-----QNYGNYGQPPY-YYPPQ	917
Ciona intestinalis	813	PGYQMPMPMGYNPYAYGQYNNPY-PPVYHQSPGQAPYFGPQQPSYFFPPQPSYFFQ	869
Anolis carolinensis	810	PGYQMPMPMGYNPYAYGQYNNPY-PPVYHQSPGQAPYFGPQQPSYFFPPQPSYFFQ	868
Homo sapiens	810	PGYQMPMPMGYNPYAYGQYNNPY-PPVYHQSPGQAPYFGPQQPSYFFPPQPSYFFQ	869
Mus musculus	811	PGYQMPMPMGYNPYAYGQYNNPY-PPVYHQSPGQAPYFGPQQPSYFFPPQPSYFFQ	869
Saccoglossus kowalevskii	798	-----PQQYQVQQQQYPP-QQPPYQYGNPP-----YPRQ	827

Figure 8.

unnamed protein product	-----	
Ciona intestinalis	1	MDKRTSHVATVAVFKLIFCLGLEISFSDDEINVAPEYHCTCSSEWMSDTPDDGSEIETPNEGRKNGYSFKSPV 80
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1	-----EYKRVLCYCPPEPIGPKFQIESQNGLSGLSAPKSPQSTTT 44
Ciona intestinalis	81	DIQCYNLALGRITIKEDTAAGCEVDFGSCNPLCDDYKIRVACCYCPFPVIGAPAFIQSQTG--GACLAVS--PFSBVL 157
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	45	VSVQCTDQD--YQNEWYSKYRINKTQSQCLSV--TADDITITENDC-----SNDNQNKYHKRSV 106
Ciona intestinalis	158	VSRNCHNDNEISQLMNMWNSNQIMQHNQICISVGTAVSSPVLTEDCD-----ASDTKQKMSYD--KRTF 223
Xenopus (Silurana) tropicalis	1	-----MI-----SLPYCLTAFVLA-----GVQGNYSYHS--DTNMTYEDA--RFFC 38
Salmo salar	1	-----MLPWSVLAQSLYKRLVILVLLTITYDGLTGGTGVQAKTYHSVQGLDWDGA--RQWC 57
unnamed protein product	107	LQYHNKSEINTRKDNQKRLRHSRGTSEWLA--LDVQTT--SLRNLTQPKCEGT--DMANPHGAIKCSL----- 174
Ciona intestinalis	222	TLSSSKYL-----DSTTATINLPSNSHWIA--KYDEGTA--DMNMQQCCSNAPD--DPR--FGSVCSF----- 283
Xenopus (Silurana) tropicalis	39	QRYTALVAIQHREEDLNSILPFPAYTWIGRLIEMONTWEGTKVLEAKHWALHEPNKKSREDCVHYVYKRE 118
Salmo salar	58	RQHTYDVAIQQEEIIVLWNLKPKMGTWIGLAKVHEWTFWGTNKLTKAEAWATGEPHNVGDDQCVETIKKQK 137
unnamed protein product	175	--GSYSSVCQ-----FSCDFGLPDAVTFKFCDD--GFVPEIIP--PTKCIIVLQPLSAPGNITCTNA--N 240
Ciona intestinalis	282	--GSYSSVCQ-----FSCDFGLPDAVTFKFCDD--GFVPEIIP--PTKCIIVLQPLSAPGNITCTNA--N 345
Xenopus (Silurana) tropicalis	119	DGKRWDEPKRKKKVALCTAACFSCSGSRECEITINNYTCSYEGFSGACHEVVDCLRVLPVGLSHECKDYKGF 198
Salmo salar	138	DEAKWDEPKRKKKVALCTAACFSCSGSRECEITINNYTCSYEGFSGACHEVVDCLRVLPVGLSHECKDYKGF 217
unnamed protein product	241	NLYSTCTDCLGKMSSTESTLSEYETKRW--HDPTECRITCQPMITSDGLTSGAICTSANSYVSKYSCDVGYEI 319
Ciona intestinalis	346	NQDSNCLYECDLGKMSSTESTLSEYETKRW--HDPTECRITCQPMITSDGLTSGAICTSANSYVSKYSCDVGYEI 425
Xenopus (Silurana) tropicalis	199	QYNSKCFSCNIGFTLGP--NSLQCISSGWS----- 229
Salmo salar	218	AFNSCCFPCACQGLLGA--SLLCQATGPD----- 248
unnamed protein product	320	IGAAHKECLGKMSSTESTLSEYETKRW--HDPTECRITCQPMITSDGLTSGAICTSANSYVSKYSCDVGYEI 395
Ciona intestinalis	426	IGLVLVLCQSSAMNATKPCIKIKCPLPLA--LQNGTVCVLG--NSLSECFSCHEGTFMVGSGSTCTCHVDTWSS 501
Xenopus (Silurana) tropicalis	230	-----SDVPACKAVCPSPV--PEHSHKEDDYEFYVSHKCFSCHEGTFMVGSGSTCTCHVDTWSS 293
Salmo salar	249	-----HFAAGCQVKKCSLDRGPHRGTNHCENFIALRSTHSCAFSCDLGFLQDFRQCHNQTGWGTC 313
unnamed protein product	396	IPVCRQECTALDIPDNGVHTSD--NPFYSSCQFTEYKLEGE--PTLCKTEKMDKGVPTCIKIVSLNLTQI-- 469
Ciona intestinalis	502	LPDQQRCTNKLTAIDNGHLCSLSD--SYSSYSSVCDFRGEIDGGQAALCTDEGNYWSPQCTRIKCPNHRKLV 578
Xenopus (Silurana) tropicalis	294	VPTCEAVVCFITLCPHSGKEDDGFQYNSKCFSCHEGTFMVGSGSTCTCHVDTWSS 348
Salmo salar	314	VYCTVYVRCGLPFPAMGNMCDVPLGASFTSSCQFCEGFLRGLD--INLTCFLS----- 368
unnamed protein product	470	--SGKVSCTDNEYSVYTYACDVGKFTFGPMELECGDEGNMVPACQKMSCPVANPANGMLNCV--AGQYNSIC 545
Ciona intestinalis	579	WGAICHNDTDFASLCTFTCTDGYEVQSL--AVICGSGNDTASTCKVKTCTPLANPKNQVWCV--GNEFNSLC 652
Xenopus (Silurana) tropicalis	349	-----SGMSSDVPACKAVCPSPV--PEHSHKEDDYEFYVSHKCFSCHEGTFMVGSGSTCTCHVDTWSS 399
Salmo salar	369	-----TGQWNTHTPACPLQCDVLTSPPHGSMNCSDLGQFHYSGQC 410
unnamed protein product	546	PLTCDYGYKTS--TRPFIQCHNEQMSQAVPCKKIQCKKLSRLRSGVTCSL--GNFNTCTNCFKRSYVNGFSEKA 623
Ciona intestinalis	653	PYSCDEGPKSHAGGPPFKLNSQAMSGPFDCKRIRKRTIKRHWKVCSD--SNNYSNCTFRCKDDPTLYGYTVA 729
Xenopus (Silurana) tropicalis	391	SFSCKEGFTLT--EENSLSQTSGMSSDVPACKAVCPSPV--PEHSHKEDDYEFYVSHKCFSCHEGTFMVGSG 468
Salmo salar	411	LFSCEGFLN--GLADTCTSLGTSRRAPLCLAQVCDVLTSPPHGSMNCSDLGQFHYSGQCFLSCEGFLNGLAD 488
unnamed protein product	622	TCR--INGWDEMPKAKRSCTPLARPYNGIINC--DTPGAPCTFKNTGYLIGRAKVCQAD--EWSNTDVRTC 694
Ciona intestinalis	730	TCG--INGWDEMPKAKRSCTPLARPYNGIINC--DTPGAPCTFKNTGYLIGRAKVCQAD--EWSNTDVRTC 803
Xenopus (Silurana) tropicalis	469	OCTSSGMSDVPACKAVCP--PSLPLPEHSGKEDDYEFYVSHKCFSCHEGTFMVGSGSTCTCHVDTWSS 545
Salmo salar	489	ECTSLGTSRRAPLCLVRC--GPLSPAMGNMCDVPLGASFTSSCQFCEGFLRGLD--INLTCFLS--GQWN--HTPAC 565
unnamed protein product	695	ARIECTSVVRSFANGVSCSD--GKHYGSKCFSCDVGVFLEGESEITCTDKRTWOGKAATCRIS----- 759
Ciona intestinalis	804	ORICLVAQKRSRSHGSKCTD--GRMFSQCNQYTCNRGSLNSHGLTCTADGTWDRKRPVPCQVRF----- 868
Xenopus (Silurana) tropicalis	546	KAVDCPHLV--VPELGSHECKDETFRQYNSKCFSCHEGTFMVGSGSTCTCHVDTWSS 624
Salmo salar	566	EPLQCDLLT--SPPHSGMNCSDLHGQFHYSGQCFLSCEGFLNGLADTCTSLGTSRRAPLCLVRC--GQWN--HTPAC 632
unnamed protein product	760	PEKIQVINGIDRDCSD--NYMSKCAVTCNGYKLVGQITSTCTQISMSAAEPVCEKICKPPNPSHGVNFCSDNN 836
Ciona intestinalis	869	PNLFCNLNGETQCTD--SYGSRCFSCQYGLIETVETNCLNRTMSKQSPYCEKICKPPNPSHGVNFCSDNN 945
Xenopus (Silurana) tropicalis	625	ETLKDPEHGVKCFER--SEYHCTCYTCARGYLVGLSEVQCLASGDMSPFPICEVVECK-----VP 686
Salmo salar	633	PLLVFWRGWNCTHPSFSPYSBCLLQCNQEGFLTGPATMHCNAGVNSQDLFSCQVQCNELHRLHPLPFPSPST 712
unnamed protein product	837	N----- 837
Ciona intestinalis	946	NFGSRNCFHCGFLNGLALHTCNADGSGVEAETPKTICTEANRLMSTSCDRSNEYSLQCNAGYLVGTQ 1025
Xenopus (Silurana) tropicalis	687	FIPDMGNMCSHFP----- 700
Salmo salar	713	PPFLGSMNCSHSL----- 726
unnamed protein product	838	N-----YGLSCEFCVCEGFTLEGSKETCG--GHWSTAKAPTKCK 875
Ciona intestinalis	1026	FORCTNDGEMSAKPIQEIKCPVLRGIPRQDVKCGSGVSECVCDPQIRKNSRVRCTDKKHPEDVPTICEL 1105
Xenopus (Silurana) tropicalis	701	-----GDFKYGVCVCKFCEGDRLLMGLTCTHCESTGTSSEV--PTCEA 742
Salmo salar	727	-----GDFSSSLCTCSKGYLMTGTELESTCEGQWSDLV--PSCOR 768
unnamed protein product	1106	S--PTIPSTTAPPITLSTTTTERTTTTTTTPPTTTTTTTP-----FSTLPTPTTIPFPVVGAGTDFPPEVAPVFAK 1180
Ciona intestinalis	1106	S--PTIPSTTAPPITLSTTTTERTTTTTTTPPTTTTTTTP-----FSTLPTPTTIPFPVVGAGTDFPPEVAPVFAK 1180
Xenopus (Silurana) tropicalis	743	--FRVDSGVVHV--AVGVAATGALLSLSLMLKVLRLKS-----AKKFTFSR----- 789
Salmo salar	769	NITSEVQSSRLVAVVGAASAGLLLLGLGHWVAKLSKGDVFNANLWEEERMPAFET----- 831
unnamed protein product	1181	QPLDGTHTPEPPVAPGAGIPIPIAVSIVVVLLVIGITLFIWVRRSRKPKFKNGLSNRIYNSDQLQKLVPHGSSN 1260
Ciona intestinalis	1181	QPLDGTHTPEPPVAPGAGIPIPIAVSIVVVLLVIGITLFIWVRRSRKPKFKNGLSNRIYNSDQLQKLVPHGSSN 1260
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1261	KRYPAIDRLSEPTFRHADDKLFREYSNIPGVSGRDGGKSHNKKRHTVHIFPDRHVRVLPKSTHTDDDYINA 1340
Ciona intestinalis	1261	KRYPAIDRLSEPTFRHADDKLFREYSNIPGVSGRDGGKSHNKKRHTVHIFPDRHVRVLPKSTHTDDDYINA 1340
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1341	SPVDGYRHKGFIAAQGPKDGTINDPWRMIWQNVTTIMVMTLKENNPKCAQVYQSGSIVGLSDVLLGENHLDVY 1420
Ciona intestinalis	1341	SPVDGYRHKGFIAAQGPKDGTINDPWRMIWQNVTTIMVMTLKENNPKCAQVYQSGSIVGLSDVLLGENHLDVY 1420
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1421	TRKFTIQCCGEATLSVRRNMIQYHPTSPDPPVPSKPSGLKFNKIKMGSPTQYGVVHSCAGVGRPTCYICIDAM 1500
Ciona intestinalis	1421	TRKFTIQCCGEATLSVRRNMIQYHPTSPDPPVPSKPSGLKFNKIKMGSPTQYGVVHSCAGVGRPTCYICIDAM 1500
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1501	VDMHREGQVGVYVTVQHRNRPVMTQYQYFIYQALLEHLVGDTEATEVNNHIDELGQHPGMVGHMERHFKK 1580
Ciona intestinalis	1501	VDMHREGQVGVYVTVQHRNRPVMTQYQYFIYQALLEHLVGDTEATEVNNHIDELGQHPGMVGHMERHFKK 1580
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1581	LTTRIQDQNRAGNRPANMKNRNVLLLPYDWNATLVPBRGLENSDYINASYIDGVYQDAYIATQGLPHTIDPWR 1660
Ciona intestinalis	1581	LTTRIQDQNRAGNRPANMKNRNVLLLPYDWNATLVPBRGLENSDYINASYIDGVYQDAYIATQGLPHTIDPWR 1660
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1661	LWETKASIVMLTELVERGQAKCEYVNRHCTVYGLQVQLKIEEELDNYTDFIITNNTDTERKQVRFYHGH 1740
Ciona intestinalis	1661	LWETKASIVMLTELVERGQAKCEYVNRHCTVYGLQVQLKIEEELDNYTDFIITNNTDTERKQVRFYHGH 1740
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1741	PTGTFPSGFMISLYEQVQKQSSGNPITITCGSAGRTGAFCALSTALEGKVAEGLVDFIVKLCNRPYHQN 1820
Ciona intestinalis	1741	PTGTFPSGFMISLYEQVQKQSSGNPITITCGSAGRTGAFCALSTALEGKVAEGLVDFIVKLCNRPYHQN 1820
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----
unnamed protein product	1821	LEQYFCTYRANQEIYDMSSEYFKK 1845 1845
Ciona intestinalis	1821	LEQYFCTYRANQEIYDMSSEYFKK 1845 1845
Xenopus (Silurana) tropicalis		-----
Salmo salar		-----

Figure 10.

```

Xenopus_silurana_tropicalis QTWGTTKVLTEAKNWLN----EPNNKKSNEDCVENVK--RKEDG 120
Salmo_salar MWTMGTKNLTKEAENWATG----EPNVGDGQDCVRIYIK--RGKDE 139
Molgula_protein_product NKKYHKSSVLYQMNKKSINYRKNQKLRHTSRGTDSEWLA--LLDVG 144
Ciona_intestinalis KSYD-KRTFTLSSSSMYL-----DSTTATITNLPNSHWIA--KVDEG 253
*.: . . . . . : :

Xenopus_silurana_tropicalis GK-----WNDEPGRKKV-----ALCYTAA 141
Salmo_salar AK-----WNDERTKKKG-----ALCYTAS 160
Molgula_protein_product TT--SLRNLQTPKPEEGT-DMANPHGAIDCSL-----GSVSSVC 182
Ciona_intestinalis TA--DMNIMQGQQSNAPD-DPN--PGSVSCSF-----GSYQGSEC 289
*.: . . . . . : :

Xenopus_silurana_tropicalis S-----PSSCSGHGETETINNYTCSYEGYVGSABHVVDCC 178
Salmo_salar S-----EHCCSTHAELENIGNYTCQCDPGFGPRQEAAC 197
Molgula_protein_product Q---FSCFEGYSLPEDAVTFKEDD-GEFVPEIIP---PTCEKIVCL 223
Ciona_intestinalis H---FSCDYGYRLPSQAIYPAICMRN-GLFNASA-----PTKEIICN 328
*.: . . . . . : :

Xenopus_silurana_tropicalis PRLVVRELSSMEKDEYKFOYNNISFSENEPFLTGP-NSLQISSGS 227
Salmo_salar DTVSDPDCQSVNEDHPHGLFAPNSQDFHAQYQLLGA-SOLLQATGF 246
Molgula_protein_product POLSAQHONITINA---NNLYTQTYDALCYKMSSTELSYETKK 270
Ciona_intestinalis PTLTSANRQKVEITNS---NNQDSNQLYEDLYKLNSSQOQLMENGH 375
*.: . . . . . : :

Xenopus_silurana_tropicalis S----- 229
Salmo_salar D----- 248
Molgula_protein_product D-HDPPTCERITCQPNITSDGLTSGAISCTSANFYSSVCKYSCDVGYEI 319
Ciona_intestinalis DDPYPPPTCQKITCLPDPITTEGINSRNVCSMKNSFASKCIYCTDGYEL 425
*.: . . . . . : :

Xenopus_silurana_tropicalis -----SDVPAKAVVPSIPV-PEHSMKEDDYEEFQ 263
Salmo_salar -----HPAAQVQKQSSLNGPHRMTNNSNPIALHSN 283
Molgula_protein_product IGAANHECQLSGKNDTKPAQKIQPPLPV-PSNKMSDDN---QLD 365
Ciona_intestinalis IGLVLRQCQSSAWNATKPIQKIMPRPLA-LQNTVTLG--NSL 471
*.: . . . . . : :

Xenopus_silurana_tropicalis NISSENEFTLASNSLQTSSSGSSSNVPTKAVVPPITCEHSGK 313
Salmo_salar TIASDLFDLIPDRIOYNTGQYTGIVSTTVVRGPLSPAMNM 333
Molgula_protein_product SKKEDNMVYMGSGSATGSDRKTSEIPVQRQETALDIDNVM 415
Ciona_intestinalis SRRSHHEIYTHVSGSTINVDRTGSLPDKRKTNKLTAIDNML 521
*.: . . . . . : :

Xenopus_silurana_tropicalis EEEIDPGEFOYNSSESEFILTSS-NSQTS----- 348
Salmo_salar SEVPLGASSPTSGSSEEEYLLRQD-INITLS----- 368
Molgula_protein_product TSS---NNFYSSTOITTEYKLEE--PTKTEKGDWGPVTCIKI 460
Ciona_intestinalis SSS---SSYYSVISELVKVEIDGGQAAITDEGWNYSPQCTRI 568
*.: . . . . . : :

Xenopus_silurana_tropicalis -----SG 350
Salmo_salar -----TG 370
Molgula_protein_product VCSRLNTIQ--SGKVSCTDNEYKSVCTYACDVGFKITPGMPNELECGDE 508
Ciona_intestinalis SCPMHNKLVPWNGAHCNDTDFASLCTPTCDTYEYQVGGSL--AVICGSS 616
*.: . . . . . : :

Xenopus_silurana_tropicalis SSSDVPKAVDPHLVVELSMEKDEHGEPKYNISFSEKFTL 400
Salmo_salar QVNTHTPAEPLQDLVLSPHSMMNSDLHGQPHYGSOLFSSEFLL 420
Molgula_protein_product GNNPVPVAKMSFVVMANMLNIV--AGQLYNIPLTLDVYKT 555
Ciona_intestinalis GNDTASTVKKRFTLAMRNVKVM---GNEFMLEPYDEPFMK 662
*.: . . . . . : :

Xenopus_silurana_tropicalis T--ESNSLQTSSSGSSSDVPAKAVVPSIPVPEHSMKEDDYEEFQY 448
Salmo_salar N--GLADTEITSLGTERRAPLQAOQDLVLSPPHSMNSLYGQFQY 468
Molgula_protein_product S-TSRPIIQMENEQSAVPTKKIQKRLSSRSVTSI---GHNF 601
Ciona_intestinalis SHAGPPITLSNGASGPPDTRRIKRTIKRMVYRSL---SNWY 709
*.: . . . . . : :

Xenopus_silurana_tropicalis NKKSSSNEGPIASNSLQTSSSGSSSDVPAKAVVPSIPLHEG 497
Salmo_salar GQQLSSEEGFLNGLADTEITSLGTERRAPLQVLRG-GPLSPAM 517
Molgula_protein_product NPTINTEKSYNFTSEKATTA-INGDEMTAKKSSPPLANVM 650
Ciona_intestinalis NNDTERKDDFTYETVATG-INGDIQITTSRAVSTTPSKING 758
*.: . . . . . : :

Xenopus_silurana_tropicalis SKEENDYGEFOYNSSESEFTITCPNSLQTSSSGSSDVPK 545
Salmo_salar NMSVDPGLGASSPTSGSSEEEYLRDITNLTLST-GQTNHTPA 565
Molgula_protein_product QIN-----DPTCAPETKKNTYTIIEAKSVQAD-ETSDNDVRC 694
Ciona_intestinalis QITL-----SGKSTVKLLSNGQFRILQNETHLSNTEITSTEAPL 803
*.: . . . . . : :

Xenopus_silurana_tropicalis KAVDPHLV-VPELSMEKDEYGRFOYNNISFSEKFTITGNSLQ 594
Salmo_salar EPLQDLLT-SPPHSMNSLHGQPHYGSOLFSSEFLLNGLADTE 614
Molgula_protein_product ARIEKTVVRSPANQVSSH---GRMGKSPSSDVPTEEEIIT 741
Ciona_intestinalis ORIQLEVAQRSSHVSKVTL---GRMPGQNYTNRVYSLSNSLTH 850
*.: . . . . . : :

Xenopus_silurana_tropicalis SSSGSSSDVPTREGIAILLIKTIFVSVQETLKDPEHVKPEK---S 641
Salmo_salar ELGTERRAPL-----LVR-----HPLVPEPVMNTHPSFP 652
Molgula_protein_product DKRTDGKAATTRIS-----PEKIQVNIISDSDG--N 776
Ciona_intestinalis ADGTDKRPVVRV-----PNLTGLNYETQIDG--S 885
*.: . . . . . : :

Xenopus_silurana_tropicalis ENNTSYTCAEYVNLSEVQLASGDMSPPLEEVVEK----- 684
Salmo_salar SSSRLLQNEFLTEPMMHALQVSOGLSQQVQCEWLLHRL 701
Molgula_protein_product YNKRKAVTNGYKVGQITSTQINTSAAEVEERIKPK----- 820
Ciona_intestinalis YNGRRRFSQYELIEVETNINLNRYSKQVYERITFP----- 929
*.: . . . . . : :

Xenopus_silurana_tropicalis -----VFPIDMKNSEPP----- 700
Salmo_salar PSLPTPSSPTPLPSMNSL----- 726
Molgula_protein_product -----VFTNSNCKVTSDNN----- 837
Ciona_intestinalis -----RFSNSNCKVSDNSN----- 946
*.: . . . . . : :

Xenopus_silurana_tropicalis -----
Salmo_salar -----
Molgula_protein_product -----
Ciona_intestinalis -----

Xenopus_silurana_tropicalis -----GDFKVG 708
Salmo_salar -----GEPFSS 734
Molgula_protein_product -----YGS 841
Ciona_intestinalis -----FGS 950
*.: . . . . . : :

Xenopus_silurana_tropicalis CPECDRLNNTLEESTGTSEVTEA--PRVPSGVVVN-A 754
Salmo_salar CLTSKGYLNTLEELTSEGOQDLV-SCQRNITSVQSSRMLVYA 783
Molgula_protein_product CESVEGPTLESKISITG-GNCKATAK----- 875
Ciona_intestinalis NHEGGFRNALHTTIN-ADGAVGEATTKITCPIANRLLHGTE 999
*.: . . . . . : :

```

Figure 11.

unnamed protein product	-----	-----
Ciona intestinalis	1	-HVVTTHHTGLVKEANNCKCKQFRKTETFYSGQNFCSAELKSKVSKSVTKTIYETVQVTSNKNRHRISKTNTRMPFS 79
Mus musculus	1	-----LGTIRMSPT-----TEPWKAR-----RARSSSC---IT---TTPA-- 30
Homo sapiens	1	ASRLPRRRQAPRSACSA-----PEPCPA--LCNHRVPAGADRPVRRLSAELSFARNSLC---PSRPRATMP-- 65
unnamed protein product	-----	-----
Ciona intestinalis	80	MRIRRRELSAAVAVILMIIPREFSYVGVAAQVSPPTQLRSSPTALFEDLPEESLANTKDPCKARLFLGDIALDEEDV 159
Mus musculus	31	---KLLSSGEILLWKMIM-----SSYFTSTRL--RTGPSHPLTN-----QDM 67
Homo sapiens	66	---RATLALGALVSLLLLPFRGAG---GLGERPDATADYSELDD--EEGTEQQLEHYHDPCKAAVFWGDIALDEDDL 135
unnamed protein product	-----	-----
Ciona intestinalis	160	ENIFGETADDFIQNVGKAPATPTRAPIGDVMSSPPSPSSNPNRRRRKGRKRRRRNRKRNKRVQRQATLDFDVASPSN 239
Mus musculus	68	TLASL-----EETSARWPNDTAS-----NASIQAPRKDKDATFLPNPGTS-----N 110
Homo sapiens	136	KLFHIDKARDWTKQTVGATGHSTGGLEEQAESSPDTTAM-----DTGTEAGKDGRENTLLHSPGTL-----H 200
unnamed protein product	-----	-----
Ciona intestinalis	240	MTEVFASYNSSQGNRTNTPVRTTGDTLRTGIPNLRLESLVLDREVSRRRSVMDIDVARAGRGRGRNRKRKRKRGR 319
Mus musculus	111	TTAKTFS----- 118
Homo sapiens	201	AAAKTFSP----- 208
unnamed protein product	-----	-----
Ciona intestinalis	320	RNRHRLRYKSRQRAEEVFKRTRRAATARQERKWPBGVPIPYTISANFTGSQRAMFKQAMRHWEQGTCLTFIERTNEDNY 399
Mus musculus	119	-----RVRRTTSRTERIWPGGVPIYVIGGNFTGQRAIFKQAMRHWEKHTCVTFVERTDEESF 177
Homo sapiens	209	-----RVRRTTSRTERIWPGGVPIYVIGGNFTGQRAIFKQAMRHWEKHTCVTFIERTDEESF 267
unnamed protein product	-----	-----
Ciona intestinalis	400	IRFTYRPGCCSVVGRKGTGPQAISIGKNCDFGIVVHELGHVIGFWEHTRPDRDDHIEIYKNIQAGQEYNFQEMDSS 479
Mus musculus	178	IVFSYRTCGCCSVVGRGGGQAISIGKNCDFGIVVHELGHVIGFWEHTRPDRDQHVITIRENIQPGQEYNFLKMEAG 257
Homo sapiens	268	IVFSYRTCGCCSVVGRGGGQAISIGKNCDFGIVVHELGHVIGFWEHTRPDRDQHVITIRENIQPGQEYNFLKMEAG 347
unnamed protein product	-----	-----
Ciona intestinalis	480	-----TGIRPSIGQRTLRSEGDIAQAKKLYQCPKCGRTLQDTSGNIT 42
Mus musculus	258	EINSLGEKYDYYSIMHYARNTFSGMFLDTRPMVDQETGMRPSIQRTQLSEGDVIQANKLYSCPTCGSTMQSTGNIS 559
Homo sapiens	348	EVSSLGETYDFDSIMHYARNTFSRGVFLDTILPRD--DNGVRPTIGQVRLSQGDIAQARKLYKCPACGETLQDITGNFS 336
unnamed protein product	43	SPDWPETYFNYAHCEWRISVTPGEKIVLDFTTMDIYLSRGCWYDFVEVRDGHWNDSPIGRYCGTTLPAQITSSDSRLWI 122
Ciona intestinalis	560	SPNWPASYPAYSCNEWRISVTPGEKIVMDFTSLNIFRSRGCWYDIEIRDGHWERSPLIGRYCGNLLPQITSTDSRILWM 639
Mus musculus	337	APGFPNGYPSYSHCVWRISVTPGEKILNFTSMDLFPKSRLCWYDVEIRDGYWRKAPLLGRFCGDKIPESLVSDSRLLWV 416
Homo sapiens	427	APGFPNGYPSYSHCVWRISVTPGEKIVLFTSMDLFPKSRLCWYDVEVRDGYWRKAPLLGRFCGDKIPESLVSDSRLLWV 506
unnamed protein product	123	KFRSTTNHEGRGFALKYEAMCGGDIVKSSGQIQSPNYDDYRPSKECVWRVVPEGYTVGLNFQVFEIERHDTCAVDYLE 202
Ciona intestinalis	640	KFRSTNFNGSGFLRLYEAVCGDIRRNSGQIQSPNYDDYRPSKECIWTVVVDGYQVGLSFQAFKVERHDTCSVDYLE 719
Mus musculus	417	EFRSSSSSLGKGFFAVYEAMCGDITKDAQIQSPNYDDYRPSKECVWRITVDPGFHVGLTFQSFIEERHDSCAVDYLE 496
Homo sapiens	507	EFRSSSNILGKGFFAAYEATCGDMNKDAQIQSPNYDDYRPSKECVWRIMVSEGFHVGLTFQAFIEERHDSCAVDYLE 586
unnamed protein product	203	IRDGDSNDKLGRCGVEPEKIQSKYNRLWIKFVSDGVSINKAGFAANFFKEIDECELEGENE--CAQVCVNTLGSYKCEC 280
Ciona intestinalis	720	VRDGNATSELIGRCGSDRPPDIKATNTLWIKFVSDASVKNAGFAANFFKERDECALESNGGCQKICINTLGSFKAC 799
Mus musculus	497	IRDGPTEDSTLGHFCGVEKPEAVKSSANRLWVKFVSDGINKAGFAANFFKEVDECSWPDHGGEQRCVNTLGSYKCEC 576
Homo sapiens	587	VRDGPTEESALIGHFCGVEKPEAVKSSNRLWVKFVSDGINKAGFAANFFKEVDECSWPDHGGEHRCVNTLGSYKCAC 666
unnamed protein product	281	NPGFELNSNGRTCDAACGGFLSNLDGYITSPNWPSEYFPNKQCTQWIVAPLQHKISIKFEEFDVEGHVCKYDVEVRS 360
Ciona intestinalis	800	YPGYELSDGMTCEAACGGVTPDGEITSPNWPREYPTNKQCIWIVAPPQHRITIEFDKFELEGNEVCKYDVEVRS 879
Mus musculus	577	DGYELAADKKEVACGGFITKLNGTITSPGWPKYPTNKCVWQVAVPQYRISLQFEAFLEGNVCKYDVEVRS 656
Homo sapiens	667	DGYELAADKKEVACGGFITKLNGTITSPGWPKYPTNKCVWQVAVPQYRISLQFEVFELEGNDVCKYDVEVRS 746
unnamed protein product	361	LEEDSPINGKYCGKTLPPITTSGNQMRITFKSDDTVAQRGFRIFYSDMDECAADNGGCQDCVNTIGSYSCSRNGFV 440
Ciona intestinalis	880	LTDESTFHGKFCGTELPPIVITSTINQMRIFKNSDDTVAQRGFRIRFSDRDECAVRNGGCMHVCVNTVGSYMSCSRNGY 959
Mus musculus	657	LSPDAKLHGKFCGSETPEVITQSNNMRVEFKSDNTVSKRGFRAHFFSKDECAKNDGGCQCEVNTFGSYLRCRNGY 736
Homo sapiens	747	LSPDAKLHGRFCGSETPEVITQSNNMRVEFKSDNTVSKRGFRAHFFSKDECAKNDGGCQHECVNTFGSYLRCRNGY 826
unnamed protein product	441	IHDGHGCKEAGCENEIVAAVEITSPNWPNKYPSRRECTWHLSTPGHRVKVVFNEFDVEFHDECTYDHLLEYDGNAD 520
Ciona intestinalis	960	LHNSGHGCKEAGCEHDVTSYVGEITSPNWPNKYPSRRECTWHISTIAGHRVKLVFNEFLESHPDCAYDHLLEYDGNST 1039
Mus musculus	737	LHENGHDCKEAGCAYKISSAEGTLMSPNWPDKYPSRRECTWNISSTAGHRVKITFEFEIEQHCEYDHLLEYDGTDSL 816
Homo sapiens	827	LHENGHDCKEAGCAHKISSVEGTLASPNWPKYPSRRECTWNISSTAGHRVKLTFNEFEIEQHCEYDHLLEYDGTDSL 906
unnamed protein product	521	NESLGKFCGKTKPPTVVATGNEMFMKFFSDASVTKRGSATHATECGGELTAGVRMOKIYSHFPQYGDNNMYSIGQDCDWI 600
Ciona intestinalis	1040	APILGRYCGSKKPSVVASGIHMFVKFFSDASVQKRGFSASHTVTCGGSLNASARTKDLFSHPQYGDITVYVNGRECDW 1119
Mus musculus	817	APILGRFCGSKKPPDVVATGSSLFRLFYSDASVQRKGFQAVHSTECGGRLKAEVQTKELYSHAQFGDNNYPSQARCDW 896
Homo sapiens	907	APILGRFCGSKKPPDTVASGSMFLRFYSDASVQRKGFQAVHSTECGGRLKAEVQTKELYSHAQFGDNNYPSQARCDW 986
unnamed protein product	601	SSDEAYGVGIELDFALFIEGEDDCSYDFVEYDGDSDSASRMGRYCGSDSPRLLSSGRSIRIFHSDDTINKKGF 680
Ciona intestinalis	1120	TASNGY--GVELMFRAFEVEDETDSCYDFVEVDPDTSAVRFRGRCGYSPDITFSSGDSLLVRFHSDDTVNNKGFQAS 1197
Mus musculus	897	VAEDGY--GVELIFRTFVEEEDACGYDFMEAYDGYDSSAPRLGRFCGSGPLEEIIYSAGDSLMIRFHTDDTINKKGF 974
Homo sapiens	987	VAEDGY--GVELTRTFVEEEDACGYDFMEAYDGYDSSAPRLGRFCGSGPLEEIIYSAGDSLMIRFHTDDTINKKGF 1064
unnamed protein product	681	YTTTQIEDVNSI----- 692
Ciona intestinalis	1198	YSSTVLQDTSVSLVQVNA 1215
Mus musculus	975	YTSTKFDALHMKR---- 988
Homo sapiens	1065	YTSTKFDALHMKK---- 1078

Figure 12.

Cornichon Expression

<i>M. oculata</i>	✓	✓	—	?	?	?	?	?	?
hybrid	✓	—	—	?	?	?	?	?	?
<i>M. occulta</i>	✓	?	?	?	?	?	?	?	?
<i>B. villosa</i>	✓	✓	✓	✗	✓	✓	✓	✓	✓
	F+3	F+4	F+6	L+2	L+11	J+1	J+96	J+192	J+480
	Embryonic			Larval		Juvenile			
	↑			↑					
	Hatching			Metamorphosis					
	Hours								

✓ = Expression
 — = No hits
 ✗ = No Expression

Table 1.

Selectin-like protein 2 Expression

<i>M. oculata</i>	✓	✓	✓	?	?	?	?	?	?
hybrid	✓	✓	✓	?	?	?	?	?	?
<i>M. occulta</i>	✓	?	?	?	?	?	?	?	?
<i>B. villosa</i>	✗	✗	✗	✓	✓	✓	✓	✓	✓
	F+3	F+4	F+6	L+2	L+11	J+1	J+96	J+192	J+480
	Embryonic			Larval		Juvenile			
	↑			↑					
	Hatching			Metamorphosis					
	Hours								

✓ = Expression
 — = No hits
 ✗ = No Expression

Table 2.

Tolloid Expression

<i>M. oculata</i>	✓	—	✓	?	?	?	?	?	?
hybrid	✓	✓	✓	?	?	?	?	?	?
<i>M. occulta</i>	—	?	?	?	?	?	?	?	?
<i>B. villosa</i>	✗	✗	✗	✓	✓	✓	✓	✓	✓
	F+3	F+4	F+6	L+2	L+11	J+1	J+96	J+192	J+480
	Embryonic			Larval		Juvenile			
				↑	↑				
				Hatching		Metamorphosis			
	Hours								

✓ = Expression
 — = No hits
 ✗ = No Expression
 ? = Unknown

Table 3.

	<i>Programmed cell death 2</i> Expression			<i>Programmed cell death 6</i> Expression			<i>Mannose-binding lectin</i> Expression		
<i>M. oculata</i>	—	—	✓	—	—	✓	—	—	—
hybrid	—	✓	✓	—	✓	✓	—	—	✓
<i>M. occulta</i>	—	?	?	—	?	?	—	?	?
	F+3	F+4	F+6	F+3	F+4	F+6	F+3	F+4	F+6
	Embryonic			Embryonic			Embryonic		
	Hours								

✓ = Expression
 — = No hits

Table 4.