

The Search for Putative Mesoderm and Muscle

Specific Genes in *Pleurobrachia bachei*

Alexander Fodor

Marine Genomics Apprenticeship
Spring 2012

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

Contact information:

Alexander Fodor
1120 Parkington ave.
Palo Alto, CA 94301
zebinini@gmail.com

Keywords: *Pleurobrachia bachei*, muscle, mesoderm, Ctenophore, *tropomyosin*, *calponin*, β -*catenin*

Abstract

There has been a long debate in the scientific community over the oldest surviving metazoan lineage. Traditionally it has been taught that Porifera, the sponges, occupy that niche possessing a diploblastic body plan without true organs; however recent phylogenetic analysis has suggests that Ctenophora may truly be the oldest lineage, free floating animals with a possible mesoderm layer and complex organs. In this study we searched for the presence of mesoderm in Ctenophora by examining the genome of *Pleurobrachia bachei* for the presence of tropomyosin, calponin, and β -catenin. Gene expression for all three were found not only in the muscular regions of *P. bachei*, but in the epidermal tissues as well, indicating there is an unknown function in the metazoan common ancestor. Homology comparisons to the rest of Metazoa reveal little about tropomyosin and calponin, however the ctenophore β -catenin protein appears to be to least derived of all metazoans and suggests Ctenophora may be the most basal extant metazoan lineage.

Introduction:

It has been a long debate in the scientific community about the body plan of the ancestor of the metazoans. Many regard Porifera, the sessile sponges, to be the oldest surviving metazoan lineage (Srivastava et al. 2010). However, phylogenetics suggest Ctenophora, not Porifera, may be the most basal of all Metazoa (Dunn et al. 2008). Ctenophores may possess a true mesoderm complete with a nervous system of sensory cells and a nerve net as well as a defined muscular system (Pang et. al. 2010). The origins of this muscular system may have come about through one of two methods: they may have been shared by a common Metazoan ancestor, or developed

in Ctenophora independently. In this study we will be examining β -Catenin, tropomyosin, and calponin genes from the ctenophore *Pleurobrachia brachei* to confirm the presence of a mesoderm and comparing their expression to the rest of the metazoa see if muscles developed before the Ctenophora/ Porifera split from their common ancestor, or if Ctenophora independently evolved a separate muscle developmental pathway.

Tropomyosin and Calponin:

Tropomyosin is a muscle specific protein comprised of two alpha helical chains in a coiled dimer present in metazoa. By linking with troponin, tropomyosin attaches to actin to control myosin heavy chain binding during muscle contraction (White, et. al. 1987). Calcium released from the sarcoplasmic reticulum reacts with troponin, releasing tropomyosin so muscle actin can bind and contract (Lehman, et. al. 1994).

Calponin is a calcium binding protein for which phosphorylation- dephosphorylation is necessary for the regulation of smooth muscle. Calponin binds to actin allowing it to be phosphorylated, but not allowing actomyosin MgATPase activity to occur. When calponin is phosphorylated by a protein phosphatase it no longer binds to actin, allowing contraction to occur (Winder and Wash, 1993). Neither tropomyosin nor Calponin is known to be used in any other function other than muscle contraction, so their presence is indicative of the presence of smooth muscle.

β -Catenin:

β -catenin is essential in the regulation of gene expression of the *WNT* signaling pathway, heavily responsible for roles in axes and cell fate determination in the cnidarians *Nematostella vectensis* and *Hydra magnipapillata*, the sponge *Amphimedon queenslandica*, and the Ctenophore

Mnemiopsis leidyi. WNT binds to frizzled, causing the phosphorylation of dishevelled, which in turn inhibits GSK-3 activity, allowing β -catenin to enter the nucleus and act as a transcription factor during development. In the absence of WNT, GSK-3 degrades β -catenin (Pang et. al. 2010; VALA Sciences, 2007). Additionally β -catenin is used in cell-cell adhesion, connecting cadherins to the *actin* cytoskeleton in muscular tissue (Sadot et. al., 2002). β -catenin is also found in large amounts in the cytoplasm of tumor cells and is largely responsible for assisting the de-differentiation of tumor cells allowing them to grow more rapidly (VALA sciences, 2007).

In the ctenophore *Mnemiopsis leidyi*, WNT's and β -catenin are largely found in the development of the mesodermal layers in embryos, the tentacle bulb, and apical organ (Pang et. al, 2010). In this study we looked for similar expression in *Pleurobrachia brachei*. We hoped to determine where β -catenin is expressed in *Pleurobrachia* adults.

Results:

β -catenin, *tropomyosin*, and *calponin* were found in the *Pleurobrachia* genome and also in Choanoflagella and a variety of other metazoa, including Porifera. The *Pleurobrachia brachei* *tropomyosin* sequence is 9389 bases long stretching from 101650 to 110803 on scaffold 24.1, containing seven exons and eight introns (Figure 1). The gene which codes for *calponin* is 4077 bases long and found on scaffold 172.1 stretching from 27355 to 23278, containing six exons and seven introns (Figure 2). The sequence coding for *β -catenin* was found to be 7710 nucleotides long, stretching from positions 26462 to 29917 and from 30232 to 34267 on scaffold 554.1, containing 10 exons and 11 introns (Figure 3). The gene trees for all three are found in Figure 7A,B,C. Best evidence found for the possible proteins of the metazoan common ancestor can be found by analyzing the protein domains.

***In Situ* Expression:**

Tropomyosin:

Expression of *tropomyosin* was found in many unique areas in *Pleurobrachia*.

Tropomyosin expression was found in the areas in which contraction is necessary: the mouth, tentacles, anal pores, and comb rows. In the mouth, *tropomyosin* RNA are found outlining the outer section, there are two strong bands, one a large circle surrounding a smaller one with not as strong expression between outlining what appears to be a lip - like structure (Figure 4E.).

Expression of *tropomyosin* also seems to form canal - like grooves leading down from the “lip” of the mouth into the digestive tract, additionally there are two “lobes” on either side of the mouth leading into it 180° apart that are labeled with *tropomyosin* expression (Figure 4E,F.).

Tropomyosin expression shows strong muscular filaments weaving through the tentacles and tentilla (Figure 4G.) possibly responsible for movement. The anal pores are perfectly outlined by a thick layer of *tropomyosin* expression with no expression in the center, but revealing small tracts leading from the stomach cavity (Figure 4D.). The comb rows have expression laterally across, a line on each side where the comb physically attaches to the *Pleurobrachia* body, with a small amount of expression in the middle of each comb, perpendicular to the directionality of all other expression in the comb. In addition the polar fields and ciliary grooves were found to have strong *tropomyosin* expression outlining their boundaries and filling in a specific circular pattern in the polar fields outlining them completely all the way back to the balancing organs where they are expressed as well (Figure 4C.). Surprisingly there is also expression scattered in individual cells throughout the ectoderm (Figure 4A.).

Calponin:

Calponin gene expression is very similar to *tropomyosin*, it is expressed in the muscular sections and in the epithelial tissue (Figure 5A,B,C,D.). In the mouth there is a very distinct band of expression outlining a lip surrounding a muscular complex leading down into the body cavity. This muscular complex looks as if it is a series of layered muscle bands which look as if they work in unison and probably assist in swallowing prey (Figure 5C,D.). There is also very strong expression in the combs where the comb plate makes physical contact with the *Pleurobrachia* body, revealing structures which appear to be cilia or possibly striated or smooth muscle (Figure 5A,B.). There is also expression in the tentacle bulbs where the muscles for contracting the tentacles are located and is expressed in the tentacles and the tenillia as well. Unexpectedly, there is also expression of calponin spread throughout the epithelial tissue of the entire animal (Figure 5A.).

***β*-Catenin:**

β-catenin is expressed in the same sections of the *Pleurobrachia* body as *calponin*. There is very strong expression in the mouth lip and the complex of muscular tissue leading down into the cavity (Figure 6.A,B,C). There is also fairly strong expression in the comb rows where the comb plates attach to the *Pleurobrachia* body (Figure 6E.). *β-catenin* expression is also found in the cytoplasm of the cells lining the tentacles (Figure 6D.), whether it assists in prey capture or is simply for cell – cell adhesion is unknown. In addition to muscular tissue, *β-catenin* is also expressed in the epithelial tissue; there is expression in seemingly every cell of the epithelial layer. In the epithelium there appears to be the highest expression of *β-catenin* in between the comb plates (Figure 6F.).

Homology Comparisons:

Tropomyosin:

Across Metazoa the tropomyosin protein consists of a PFAM tropomyosin domain, responsible for binding to actin, but the size of the domain differs across taxa. *Pleurobrachia bachei* tropomyosin has a PFAM domain of only 151 amino acids. When compared to *Mnemiopsis leidyi* with a domain of 273 amino acids, *Amphimedon queenslandica* with a domain 236 long, *Mus musculus* with a domain of 236 amino acids, and *Trichinella spiralis* with a domain of 256 amino acids. It seems that there was a major deletion after the branching of *Pleurobrachia bachei* from *Mnemiopsis leidyi* which caused a deletion in part of the domain, or there is a problem with the gene annotation in our database.

Calponin:

The calponin protein in *P. bachei* consists of a 50 amino acid non-specific region, followed by a 100 amino acid calponin homology (CH) binding domain involved in binding to filamentous actin (Castresana and Saraste, 1995). The CH domain is then followed by another 40 amino acid long non-specific region, and finally a 20 amino acid long calponin family repeat that performs an un-established role. The size of the non-specific regions differs from organism to organism, but the overall structure is conserved throughout all of metazoan with the exception of *Trichoplax adhaerens* (placozoa) where the calponin family repeat appears to have been lost. The overall conserved nature of this gene suggests that the common ancestor of all metazoan had a similar gene with a CH domain and a calponin family repeat.

β -Catenin:

The majority of metazoan β -catenin proteins contain 12 repeating armadillo (ARM) domains, 11 of them possessing 40 to 43 amino acids, and one of them (consistently the third to

last) contains 62 (Cnidaria, Vertebrata, Nematoda, Hemichordata, Arthropoda, and Porifera), all of which are positioned directly next to each other on the genome. Each one of these ARM domains is specific to an incoming WNT signal and to linking sites in cell-cell adhesion. The Porifera *Amphimedon queenslandica* possesses eleven such domains including the 62 amino acid domain positioned third to last in the classic metazoan pattern. *Amphimedon* has a 40 amino acid gap in between the first and second of its eleven domains, indicating the possibility of a loss of a domain. *Pleurobrachia* have six such repeating domains, all of which possess 40 amino acids. In *Pleurobrachia* the ARM domains are in three groups of two, each positioned 40 amino acids apart (Figure 1.).

Discussion:

Tropomyosin and Calponin:

Tropomyosin's domain organization reveals information about the common ancestor to all Metazoa. The presence of a longer domain in both *Mnemiopsis leidyi* and in *Amphimedon queenslandica* suggests that the common ancestor possessed an elongated tropomyosin domain. This longer domain seems to be partially deleted in *Pleurobrachia*, but this deletion does not appear to have changed the whole protein domain. The calponin protein appears to be highly conserved across taxa. All metazoans have the same CH domain and a PFAM calponin domain except for *Trichoplax*, as they have presumably lost their muscular tissue. This is indicative of a common ancestor with a fully formed calponin protein with both the CH and PFAM: calponin domains (Figure 2). Neither tropomyosin nor calponin have been found to function anywhere but in the muscle; therefore the presence of both proteins in all of metazoa predicts an ancestor

with a fully functional muscular system, however the existence of both proteins in the epithelial layer suggests that theory may not be entirely accurate.

The presence of the conserved domains of both tropomyosin and calponin in choanoflagellates suggest that there may be another possibility. This may indicate there is another function for both of these proteins, and doesn't prove the existence of muscular tissue in the common ancestor, but rather that the common ancestor may have had the proteins necessary for forming muscle, using them in an ancient arcane manor. However, *Trichoplax's* loss of the calponin domain may indicate it did not serve a completely vital function in the common ancestor.

The presence of tropomyosin in the polar fields and ciliary grooves has interesting implications. This suggests that tropomyosin may be used in the sensory system of *P. bachei*. This may also indicate that the sensory cells in *P. bachei* are derived from muscle cells and some expression still remains. More study should be performed on the development of *P. bachei* and more specifically of the cells of the polar fields and ciliary grooves to determine the proteins necessary for their formation and function.

β-catenin:

The β-catenin protein domain structure is very consistent across taxa. Analysis of the β-catenin domains shows the non-basal metazoans possess 12 ARM domains, 11 of which are 40 to 43 amino acids long and one of 62 (the third to last). The ctenophores *P. bachei* and *M. leidy* possess the fewest ARM domains: *P. bachei* possesses only six ARM domains and *Mnemiopsis* has eight, none of which have the characteristic metazoan 62 amino acid domain. The Porifera

A. queenslandica possesses 11 domains including the metazoan 62 amino acid long domain and a 40 amino acid gap between the first and second ARM domains.

The missing domains in Ctenophora could be the result of one of two evolutionary events: either there was a loss of three repeating ARM domains including the 62 amino acid long domain, or this is closer to the original form of the gene present in the common ancestor. *M. leidy* β -catenin contains 8 repeating ARM domains of 40-43 amino acids six of them connected to one another with a 40 amino acid gap before the next group of two. *M. leidy* only shares six of its eight repeats in common with *P. bachei* (all six *Pleurobrachia bachei* repeats) and two which look as if they were lost in *P. bachei* (there still is a great deal of homology between the ctenophores in that section). When these domains are compared to Porifera there is only one domain Ctenophora possess which *A. queenslandica* does not, and it is shared by both ctenophores. The other seven domains of *Mnemiopsis* and five of *Pleurobrachia* are found in *Amphimedon*. Interestingly neither ctenophore possesses the 62 amino acid ARM domain, indicating that the mutations causing this occurred in metazoan evolution after the branching of Ctenophora, but before the speciation of Porifera. If this is the case then the repeat of Ctenophora that Porifera is lacking were either lost in Porifera, or are domains specifically gained by the Ctenophora lineage. When Ctenophora and Porifera are compared to Hemichordata, it is found that every sequence found in either Ctenophora and Porifera is found in Hemichordata. Furthermore, the domains shared only by Porifera and Hemichordata are not very well conserved in Ctenophora and may not have been present yet in the common ancestor, however the domain area possessed by Ctenophora and not Porifera is well conserved in Porifera and may indicate the loss of that domain in Porifera (Figure 8 and supplemental 1a,b,c). All suggesting Porifera are more derived than Ctenophora.

Through the analysis of β -catenin, tropomyosin, and calponin genes, a common ancestor to metazoa starts to take form. This ancestor seems to have had the determinants for a mesoderm as well as a true muscular system. If the muscular system was present, that would be indicative of a free floating motile organism, much closer to Ctenophora than to Porifera. However, the presence of the conserved versions of the genes in Choanoflagella along with their expressing in the epithelial layers of *P. bachei* suggests that these genes may serve some other functions in single celled organisms. In order to properly assess this phylogeny, more genomes will need to be sequenced on more ctenophores, sponges, and the rest of metazoa in order to perform a proper comparison.

Analysis of the β -catenin domains across taxa draws the conclusion that the common ancestor to all Metazoa possessed at least 8 repeating ARM domains. One domain mutated to 62 amino acids in Porifera and higher Metazoa indicate Ctenophora may be the most basal surviving metazoan lineage, branching off before that event. More evidence for this is found when analyzing the domains in Porifera not present in Ctenophora: the overlapping section in both Ctenophora genomes match well to one another, but do not align well to Porifera indicating that the Ctenophore sequences are highly conserved and Porifera is more closely related to the other metazoa. Sequencing of more Porifera and Ctenophora species' genomes will have to be undertaken as well as the rest of the basal metazoans to determine the presence and conservancy of the mesodermal and muscle specific genes. In addition more data should be collected on the protein domains of β -catenin in all ctenophores to try to determine the overall domain of their common ancestor and see if it possessed the 62 base ARM domain.

Although all three proteins were found expressed in muscular tissues in the ctenophore, there is no concrete proof of a true mesoderm. *Tropomyosin*, *calponin* and β -catenin expression

were all found in the epithelial layer as well, indicating that there is an unknown ancient function to these muscle specific genes which still may be used in Ctenophora. Further studies will need to be performed to find the original function of these proteins, and a study of the expression of more mesoderm and muscle specific proteins will show definitively if there is indeed a mesodermal tissue layer.

Materials and Methods:

Gene identification:

A list of possible mesoderm and muscle specific genes was compiled using the KEGG online database of other Metazoa. The sequences of the genes were then run through a blastx search of the Moroz *Pleurobrachia bachei* sequenced database, a theoretical database of the *Pleurobrachia* genome. The significant Blast results there were then run through a tblastn in the transcriptome database for the protein sequences (supplementary information). Four genes were then selected for PCR: *Myosin heavy chain*, *tropomyosin*, *calponin*, and β -*catenin*. Forward and reverse primers were designed for the genes in the program gene runner and sent to Integrated DNA Technologies to be synthesized.

PCR was run on each gene using the forward and reverse primers, and the products separated by gel electrophoresis. Unfortunately at this stage it is believed that the *myosin heavy chain* sample was contaminated by the *tropomyosin* and, due to time constraints, was abandoned for the remainder of the experiments. The remaining three genes (*tropomyosin*, *calponin*, and β -*catenin*) were then transformed into cloning vectors, and placed into One Shot© bacteria (using the One Shot© protocol). The bacteria were then placed on antibacterial plates (to only allow the colonies which contained the vectors to survive) and allowed to grow colonies overnight. A

colony was then removed from each plate and grown in solution in an incubator at 37°C. Once in a sufficient concentration (when the solution was cloudy and viscous), the bacteria were lysed and the vectors inside were isolated for sequencing and probe making (*myosin heavy chain* was discovered to be contaminated by *tropomyosin*).

Adult *Pleurobrachia bachei* were collected off the docks of Friday Harbor Marine Laboratories, and put through a series of washes (see supplemental information). The animals were then hybridized with probes for *tropomyosin*, *calponin*, and β -*catenin* made from the bacterial cloning vectors using *P. bachei* *In situ* hybridization protocol adapted by Dr. Andrea Kohn (supplemental information). The animals were then stored in ethanol until they were ready to be mounted on slides in permount (see supplemental information).

Phylogenetic analysis:

The protein sequences found in the Moroz *Pleurobrachia bachei* database were blasted in the NCBI, JGI, and Broad Institute databases and the top hits for each phyla used. The protein sequences were put into Fasta format (with >name then a new line with the sequence) and then aligned in Mafft and trimmed using G-blocks. The trimmed protein alignments were compared in MEGA and a maximum likelihood gene tree was constructed for each gene.

Supplement

***In situ* protocol (adapted by Dr. Andrea Kohn from Derelle and Manuel 2007 and Moroz):**

Animals were fixed in 4% paraformaldehyde in Filtered Sea Water (FSW) overnight at 4°C in a 50 mL conical tube. The next day they were rinsed 3 times in PTW for 10 minutes at

room temperature. They were then rinsed in 1:1 methanol to PTW mixture for another 10 minutes at room temperature. The animals were then dehydrated in 100% methanol and stored at -20°C for at least two hours (for up to a week).

On the third day the animals were rehydrated by washing them with 3:1, 1:1, 1:3, and 1:0 solutions of methanol to PTW for ten minutes each at room temperature. They were then washed in a 1:1 ratio of pre-hybridization buffer (25 mL formamide, .5mL EDTA, 12.5mL SSC 50µL Tween 20, and fill to 50 mL with molecular water) to PTW for 15 minutes at room temperature. The animals were next incubated (pre-hybridized) in pre-hybridization buffer for an hour at 60 °C. Finally, the animals were incubated in hybridization buffer (pre-hybridization buffer with 1mL Denhart and 15mg tRNA) with DIG-RNA probes added overnight at 60 °C.

On day four the animals were washed in hybridization buffer for 30 minutes at 60 °C. They were then washed in a 1:1 ratio of hybridization buffer and PTW for 30 minutes at 60 °C, and washed for a final time in PTW for 30 minutes at room temperature. The animals were blocked in 10% goat serum to prohibit non-specific binding (at room temperature). Finally, the animals were incubated in anti-DIG 1/2000 at 4 °C over night.

On day five the animals were first washed four times in PBS for 30 minutes each, then placed in detection buffer made from the Vector Red Alkaline Phosphatase kit 1 (add two drops of reagents 1,2, and 3 to 5 mL of developing solution; it is best if the detection buffer is made directly before this step), covered in foil, and placed on ice (to eliminate background staining). Development took various amounts of time, ranging from four hours to eight and a half hours. Once the DIG staining had progressed far enough they were stopped in PBS. Then the animals were washed in 4% paraformaldehyde in methanol for 30 minutes. They were finally washed

three times in ethanol for ten minutes each at room temperature and stored in ethanol until mounting.

Pleurobrachia In situ hybridizations were “cleared” in methylsalicylate until they sunk to the bottom and mounted carefully on microscope slides using tweezers under a dissecting microscope. Permunt was then carefully added to the animals and a coverslip eased into place, slowly guided by a pair of tweezers. Pressure was then applied to the top of the coverslip to stop the formation of bubbles until the permunt was allowed enough time to dry.

Phylogenetic analysis:

Mafft instructions:

A file was created in notepad in Fasta format (fas: “>”name of sample then a new line with sequence) of the genes from the various species we wished to compare. The program was opened and file name typed in for the input. The name we wished to give the new file created with the newly aligned sequenced was typed in as the output. We choose option number 3 for a Fasta file, then option 5 for the “L” process, and the program run giving an output. We then put a .fas suffix on the file so that it could be used in other programs.

G-block instructions:

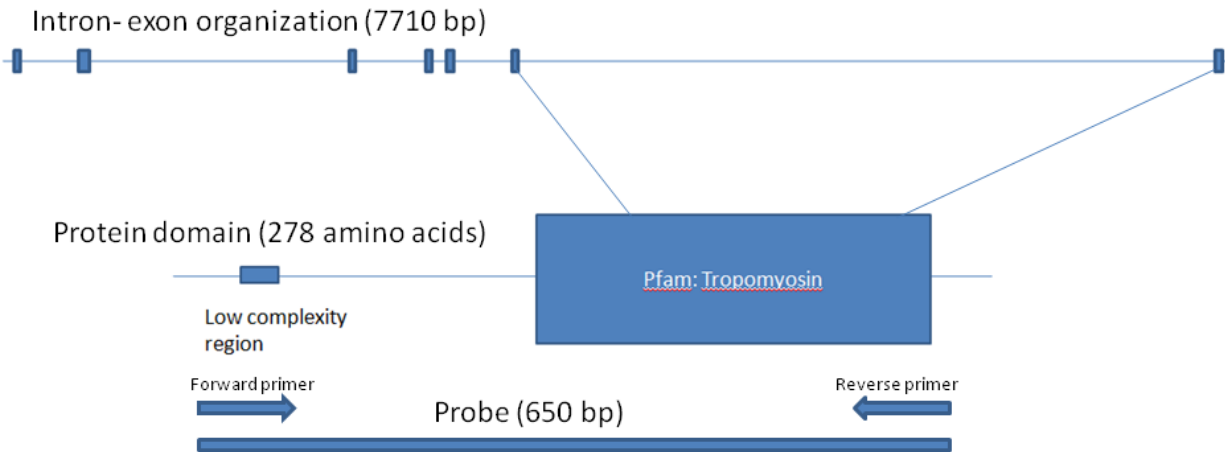
The g-blocks program was opened, selecting “o” for open, and the output file from Mafft used. Option “b” for block parameters was selected and “5” was then selected twice to change it from no gaps to all gaps. Option “g” to get the blocks was then selected and the program trimmed the sequences, and an output given. The output file suffix was then changed to .fas so that it could be opened in MEGA.

MEGA instructions:

The G-block output was selected and opened in MEGA. In the alignment file, under the tab “data” the option phylogenetic analysis was selected. We did not have nucleotide sequence so selected no when asked. In the main window we then selected the “analysis” tab, went into the phylogeny sub-section and selected construct maximum likelihood tree. The phylogeny test was changed to bootstrap analysis with 100 bootstraps, the model/ method was changed to WAG with Frequencies, and gap treatment was changed to “use all sites”, and the tree was constructed.

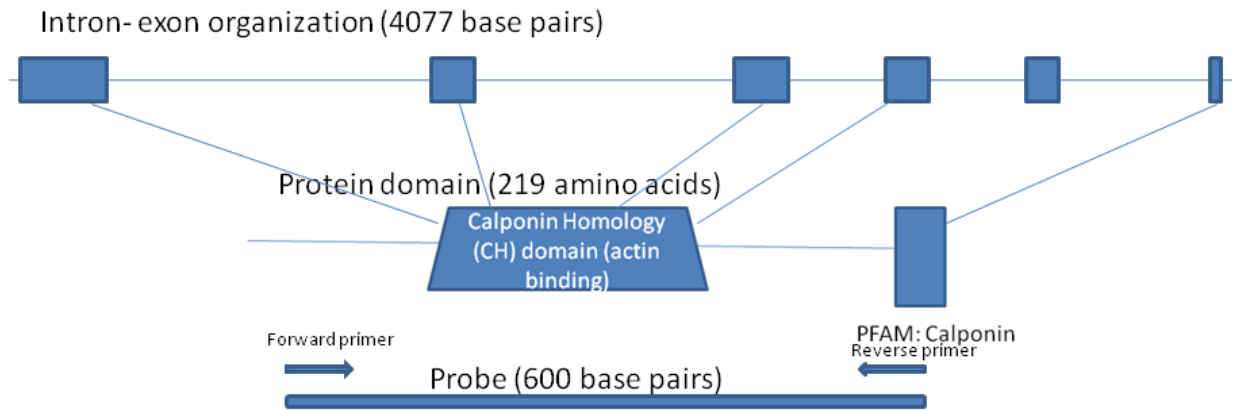
Domain alignment:

GeneDoc was opened and the G-block Fasta file imported. The entire input was selected and configure was selected under the project tab. The font settings were set to 8 points, the seq block sizing was set to fixed and changed to 80, and the consensus line was changed to no consensus. Under the edit tab, copy selected blocks to RTF file was selected so that the alignments could be viewed as a word document. The alignments were then inspected by hand and compared to the domain structure to mark where the domains of the different organisms overlap.



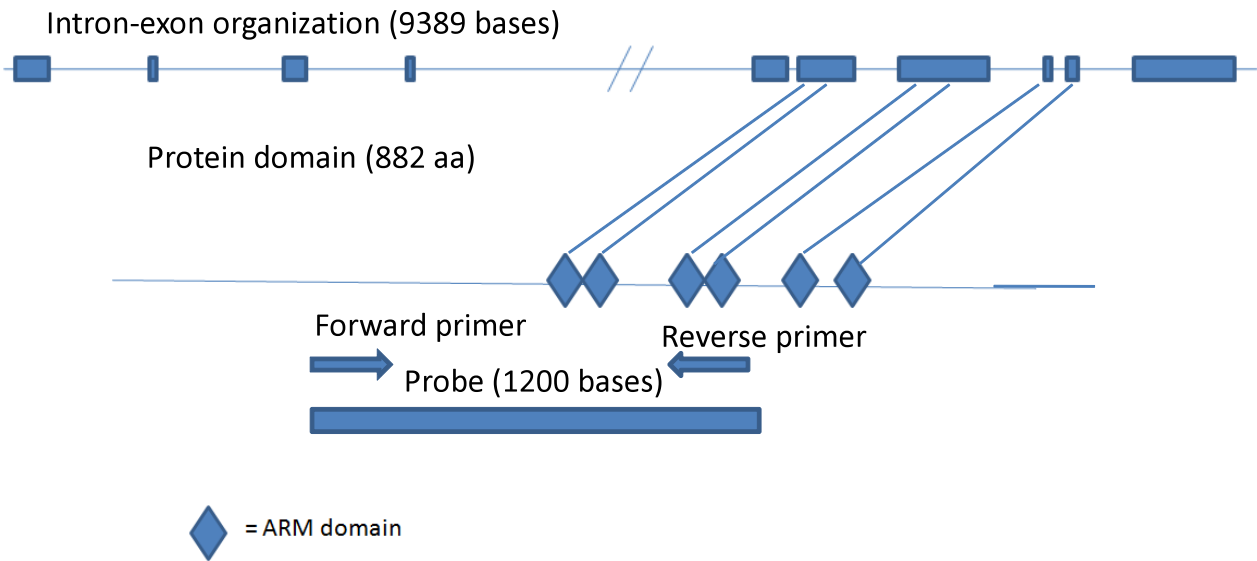
Tropomyosin

Figure 1. Intron - exon organization, protein domains, primer and probe positions of *Pleurobrachia bachei tropomyosin*. Intron - exon organization found through Neurobase BLAT, protein domains found in SMART.



Calponin

Figure 2. Intron - exon organization, protein domains, primer and probe positions of *Pleurobrachia bachei calponin*. Intron - exon organization found through Neurobase BLAT, protein domains found in SMART.



Beta-catenin

Figure 3. Intron - exon organization, protein domains, primer and probe positions of *Pleurobrachia bachei* β -catenin. Intron - exon organization found through Neurobase BLAT, protein domains found in SMART.

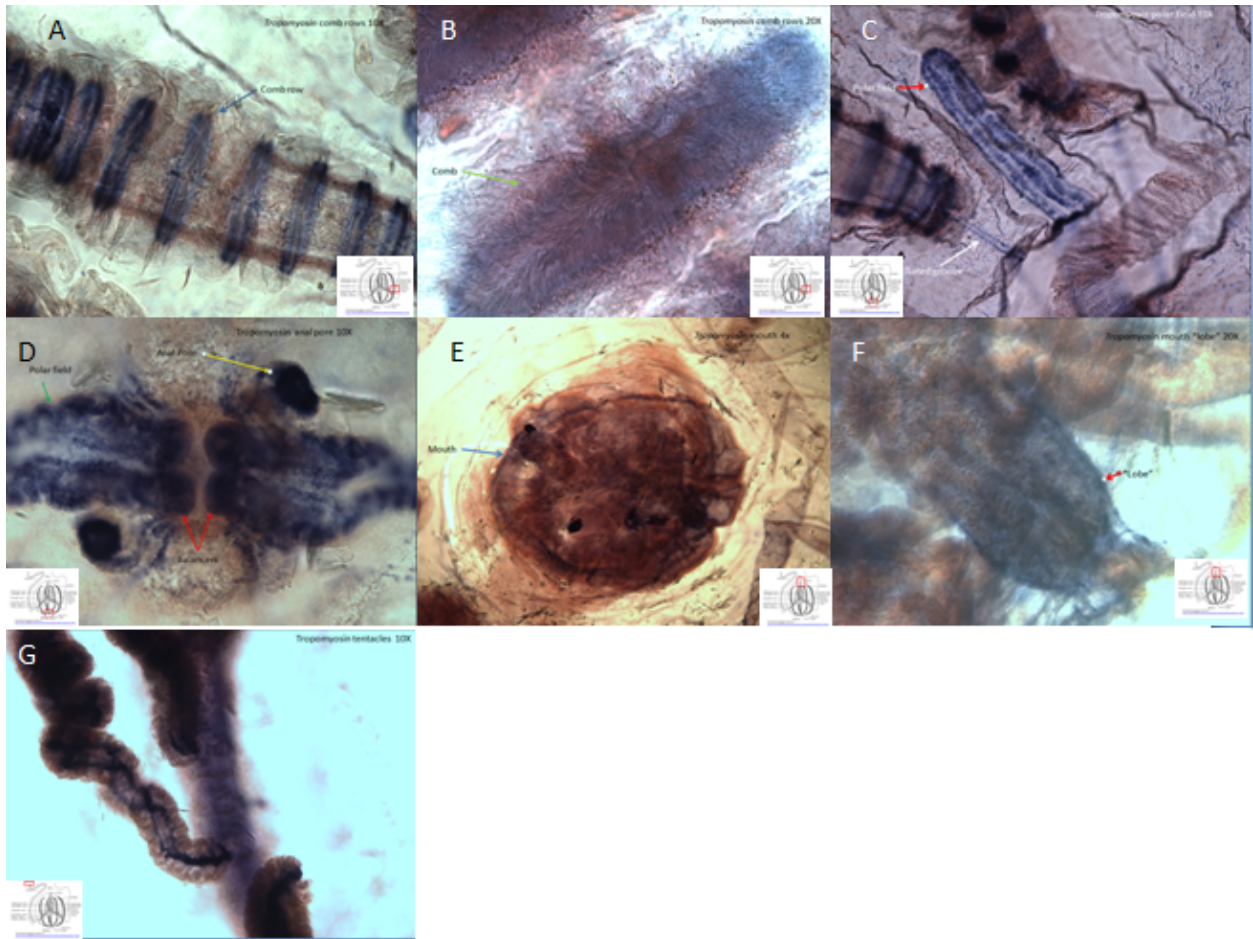


Figure 4. *Pleurobrachia bachia* *In situ* hybridization showing expression of *tropomyosin* in A, B: the comb rows and epithelial tissue; C: polar fields (red arrow) and ciliary grooves (white arrow); D: anal pores (white arrow) and balancing organs (red arrow); E: mouth; F: mouth lobes; and G: tentacles.

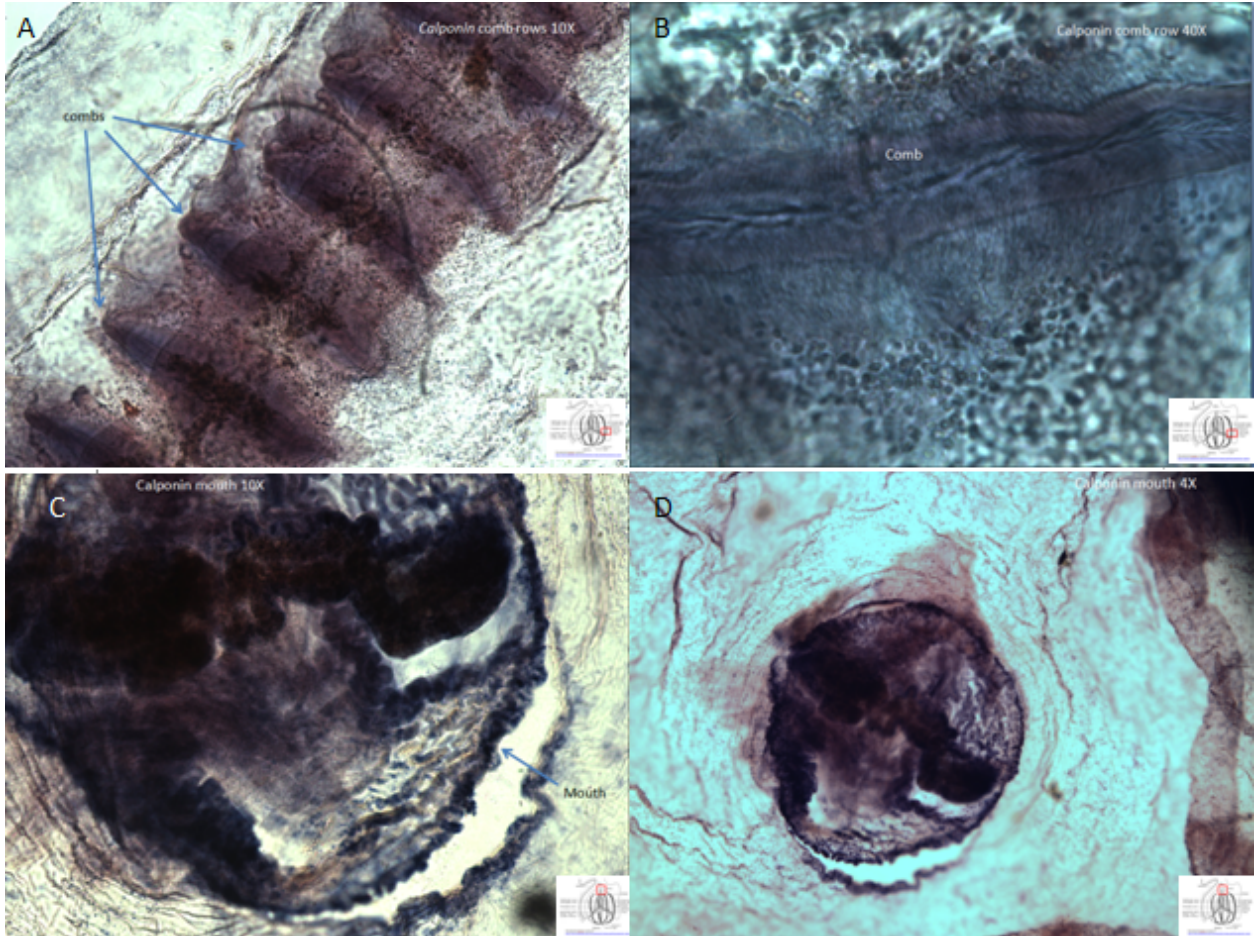


Figure 5. *Pleurobrachia bachia* *In situ* hybridization showing expression of *calponin* in A,B: the comb rows and epithelial tissue; and C,D: mouth.

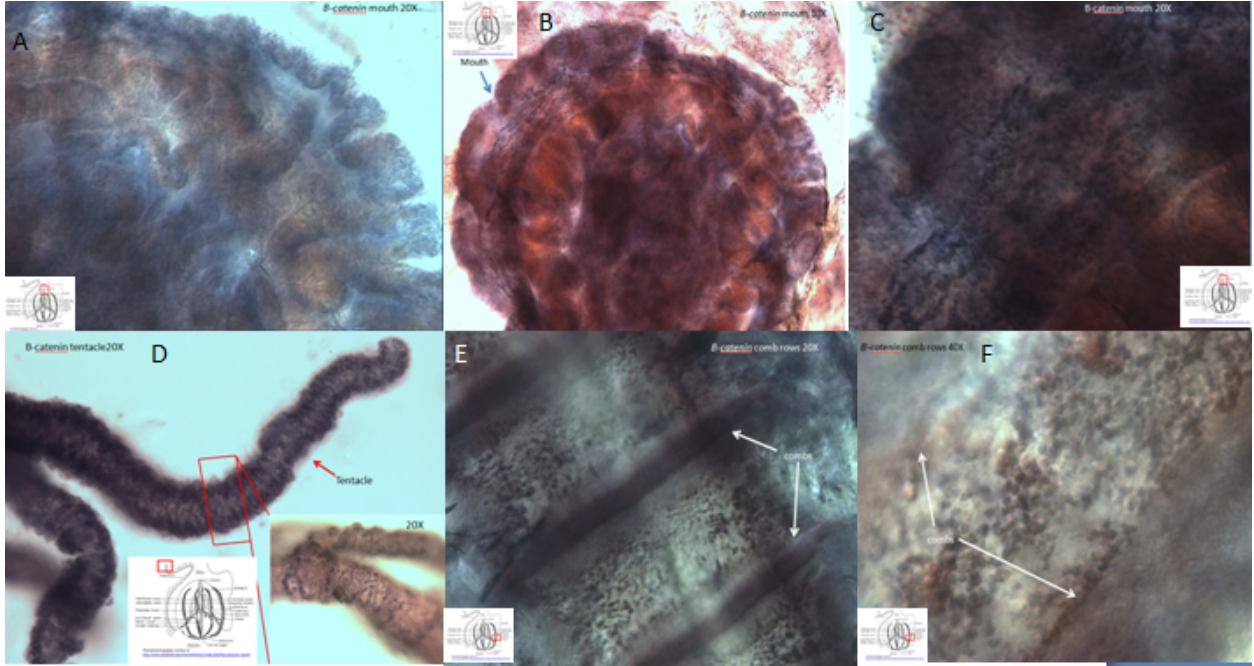


Figure 6. *Pleurobrachia bachia* *In situ* hybridization showing expression of β -catenin in the A,B,C: mouth; D: tentacles; and E,F: comb rows and epithelial tissue.

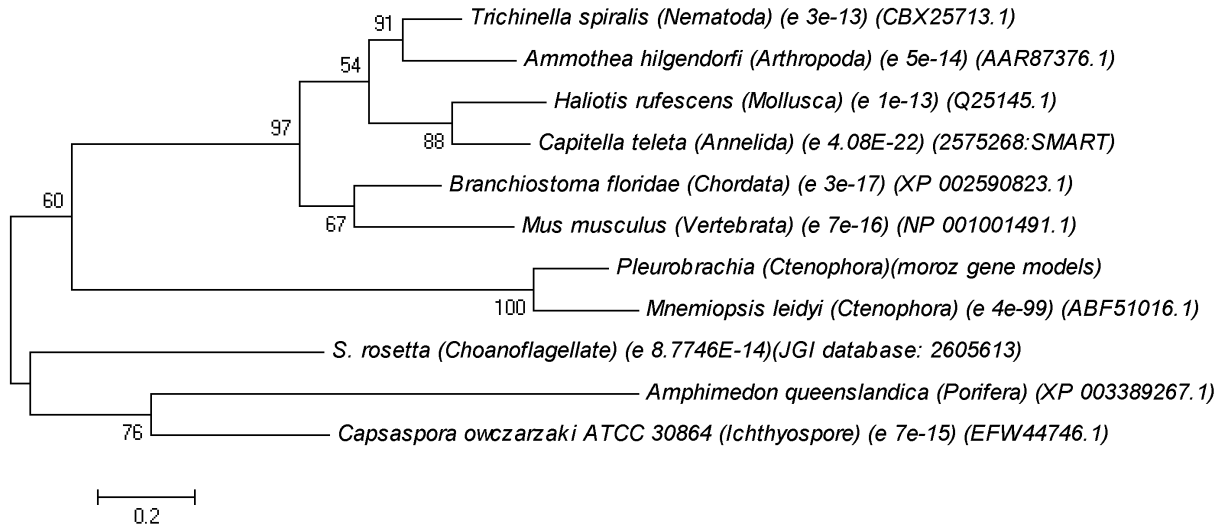


Figure 7a. *Tropomyosin* maximum likelihood gene tree, sequences found through Blast search of NCBI, alignments were made in Mafft, trimmed in G-blocks, and trees were assembled in MEGA. Bootstrap values of less than 50 were omitted. All reference numbers are NCBI unless otherwise stated. Ctenophora appears to be basal to metazoa, Porifera has grouped with the out groups, most likely a result of long branch attraction.

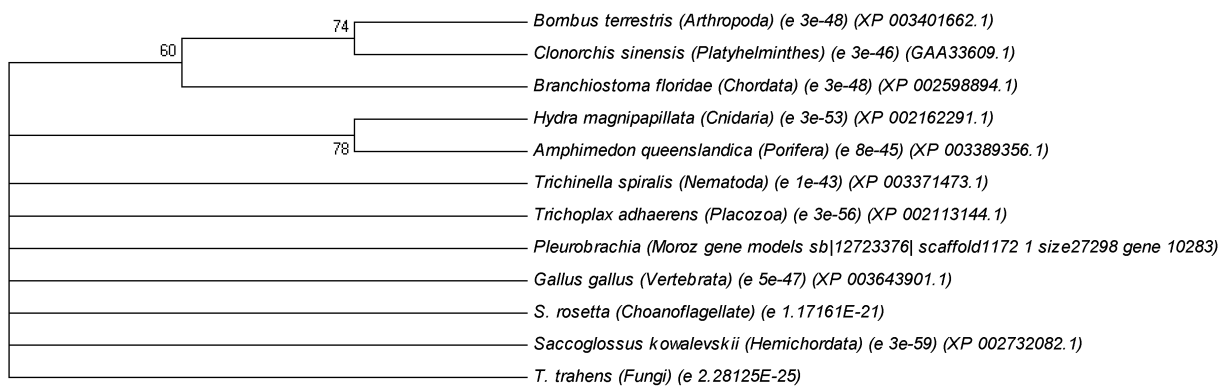


Figure 7b. *Calponin* maximum likelihood gene tree, sequences found through Blast search of NCBI, alignments were made in Mafft, trimmed in G-blocks, and trees were assembled in MEGA. Bootstrap values of less than 50 were omitted. All reference numbers are NCBI unless

otherwise stated. Lack of support is due to protein domain conservancy. The protein has been conserved across taxa and the program has trouble determining the relationship due to similarity.

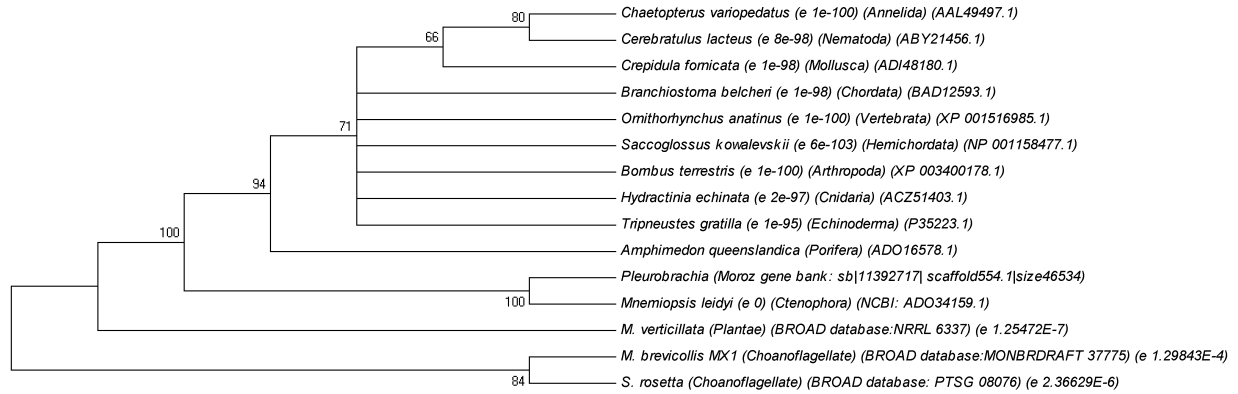
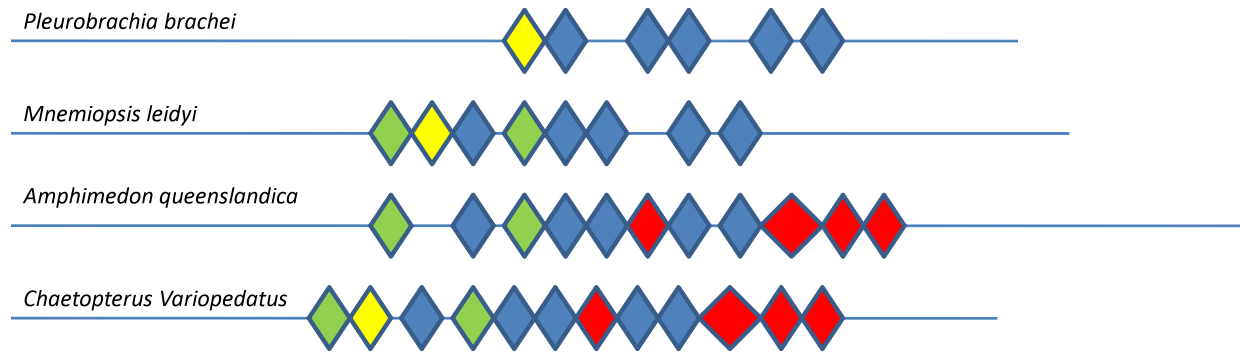


Figure 7c. β -catenin maximum likelihood gene tree, sequences found through Blast search of NCBI, alignments were made in Mafft, trimmed in G-blocks, and trees were assembled in MEGA. Bootstrap values of less than 50 were omitted. All reference numbers are NCBI unless otherwise stated. Ctenophora is more basal than Porifera which appears to be ancestral to the rest of Metazoa. Low values within the rest of Metazoa are due to protein conservancy.



B-Catenin Domain Alignment

Figure 8. β -catenin protein domains of *Pleurobrachia brachei*, *Mnemiopsis leidy*, *Amphimedon queenslandica*, and *Chaetopterus Variopedatus*. Domains found using SMART online software. Blue and yellow domains are found in both ctenophores and non – Porifera Metazoa (*P. brachei* and *M. leidy*) blue indicates it is found in Porifera as well, yellow indicates it is not; green is found in *M. leidy* *A. queenslandica* and the rest of tested metazoans except *P. brachei*; red is found in Porifera and other Metazoa but not Ctenophora.

```

MlB-Cateni : MEIPVYQELSWN-----20-----*-----40-----CKQ-KDYRLEQRRNPTIQSLQOSPNNRVLEKQLHWQ : 55
PbB-Cateni : MEIPVYQELSWNSLGGGTTLAHSIVPPSQPPNNTQTSSVNA CKRTAEYRALEQRRNVI IQRLQQT PNRVLEKQLHWQ : 80

MlB-Cateni : TEMYMYQN-----*-----100-----*-----120-----*-----140-----*-----160
PbB-Cateni : IEIYMFQSRHDSRDYGGGGDIMEKLGVPQTETGGWPKQFDSVPE SQCPNNTQTSSVNACKRTAEYRALEQRRNVI IQRLQ : 160

MlB-Cateni : -----*-----180-----*-----200-----*-----220-----*-----240
PbB-Cateni : QTPTNRVELEQELHQVQIEIYMFQSRHDSRDYGGGGDIMEKLGVPQTETGGWPKQFDSGIASGITTSPPSDDGQNDYQ : 239

MlB-Cateni : NGIDEELAMENIPSEDDVVPFNKSSFNQIQRSQRRLRAELFPDSEEQENPAPTTPSHASPTTPVITITLQTLGEPTFVSHN : 213
PbB-Cateni : AALIEEDL--ONLPEEEN---YSVTG--QSRSERLRAELFPDTSSEQ-----VNHGYSI---EPLSHL : 295

MlB-Cateni : -----*-----340-----*-----360-----*-----380-----*-----400
PbB-Cateni : TMENIAHYODDEELAMQAIIEEELKLSNDEVCVSVTCTIHELSKKEPSMHALISYPQLISITLLEHTSTETDCEVVRGC : 293
PbB-Cateni : ELQPTS--YQGETELARQAISEEKLINHSREIILLSVATCLRELSKKE--VECVNEAQLLESITLRLSTSSDREVRRCF : 372

MlB-Cateni : -----*-----420-----*-----440-----*-----460-----*-----480
PbB-Cateni : TGLIHNLSYDREGLESIQRYRGLIILCELLSNSSNVLYTITLILNMIQRHKSSQLDITCLSGGLERIVSILKYSDKKPE : 373
PbB-Cateni : SGLVLYNLKSRKGLIATQSYRGLVLCDDLLTSPSEVLTQETINILNHLIQKQCSCHYSVFLSGGLERIVGLKYYDKPOL : 452

MlB-Cateni : -----*-----500-----*-----520-----*-----540-----*-----560
PbB-Cateni : VAMVAECIHM LCHRNPDKKAVLACS EPIELVKLITQTSVEVYVAVARALKTILSVONKKECILEAGGMCLARHLDTG : 453
PbB-Cateni : AALIAECVYMICHRSPDKKSLIGONGPIILVWQENSTLEYLSAVARALKTILSVCPNKRATIEAGGMCLARHLDTG : 532

MlB-Cateni : -----*-----580-----*-----600-----*-----620-----*-----640
PbB-Cateni : SQRVTEEVLYTMRNLSNVAKGGSDVDVDPILTKIMDVLARTNINILACSVGCLSNLTCNPNKKEVFRAGGLDTLLRCL : 533
PbB-Cateni : SQRVTEEVLYTMRNLSNVAALSIETIDPILTRIMDVLARTNINILACSVGCLSNLTCNPNKRKRVFRAGGMDTLLRCL : 612

MlB-Cateni : -----*-----660-----*-----680-----*-----700-----*-----720
PbB-Cateni : LQNQAIIPVAEPLLCTLRHINYKHSEVEQTOHAFRTSGGISLIDKVISIAQHWAVIKAACGLIRNLNANMIEFKETG : 613
PbB-Cateni : VHHQSIASIAEPLLCTLRHINYKHSEVEQTOHAFRTSGGISLIVKVLMTAQHWAVIKAACGLIRNLNANMIEFKETG : 692

MlB-Cateni : -----*-----740-----*-----760-----*-----780-----*-----800
PbB-Cateni : LLLHLIQCLRITDDEYKGCINDNRDGIKLEDIMWESIKLESVSCHP--EIRAQLDNPIAHLLMKCINSNINENLWQVIALS : 692
PbB-Cateni : LMLHLIQCLRITDDEYKGCISEIDRGIKLEDIMREGINLERSHESREIR--SOLDNETIHLILRCVASEHEDLWRDALS : 772

MlB-Cateni : -----*-----820-----*-----840-----*-----860-----*-----880
PbB-Cateni : SVVQFESHENFVLIKNDPREFELICKLCHERN--PRTAQMAVLSNRINEVQVDVWPPPTAMEQPPHCTPGIFVNGGVQGE : 771
PbB-Cateni : VMILYEQRDNFANLLR--ARSELLTSLCHTHAGSOVASLATELRLQLGSKNVVPPPPPPQPPPSQ-----QQRWID : 844

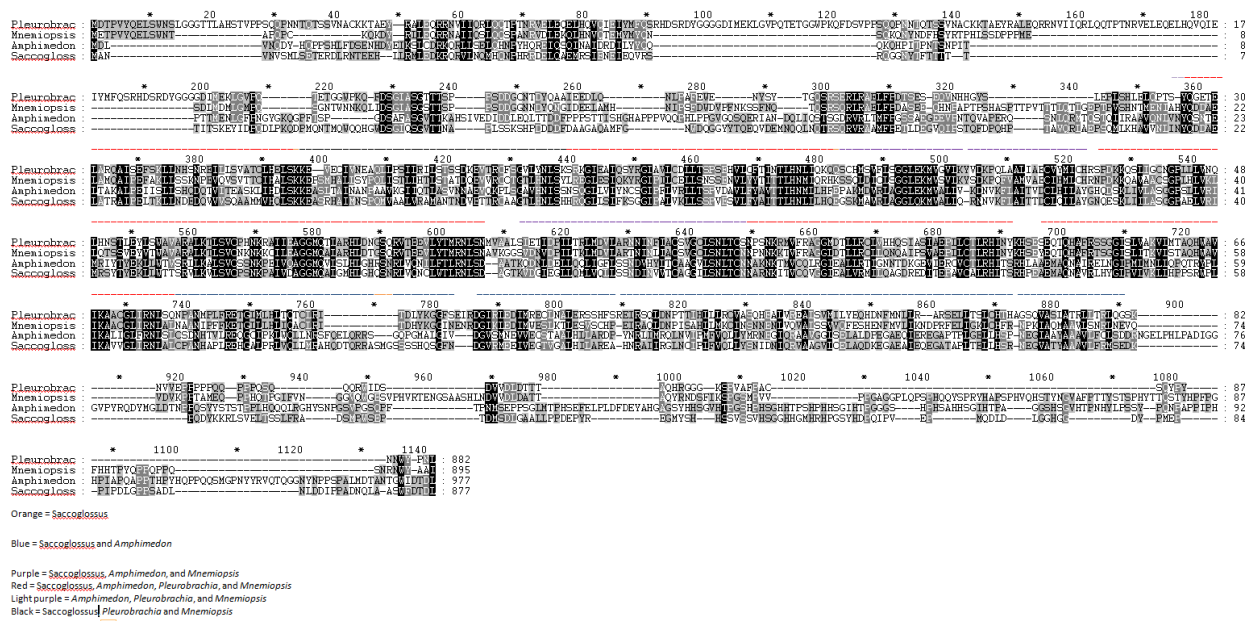
MlB-Cateni : -----*-----900-----*-----920-----*-----940-----*-----960
PbB-Cateni : SVPHVRTENGSAASHLNDVVDLDTTAQMRNDSFKSPGSMFVPPGAGGPLQPSPHQYSPRYHAPS PHVQHSIYNGVA : 851
PbB-Cateni : S-----DVVLDLTTAQHREGG--KSP-----VA : 866

MlB-Cateni : -----*-----980-----*-----1000
PbB-Cateni : FBITYSTSPHYTTQSTYHPFGFHTPYQPPQPQSNRNWYAAI : 895
PbB-Cateni : FPACS-----QYPPN-----NWYPNL : 882

```

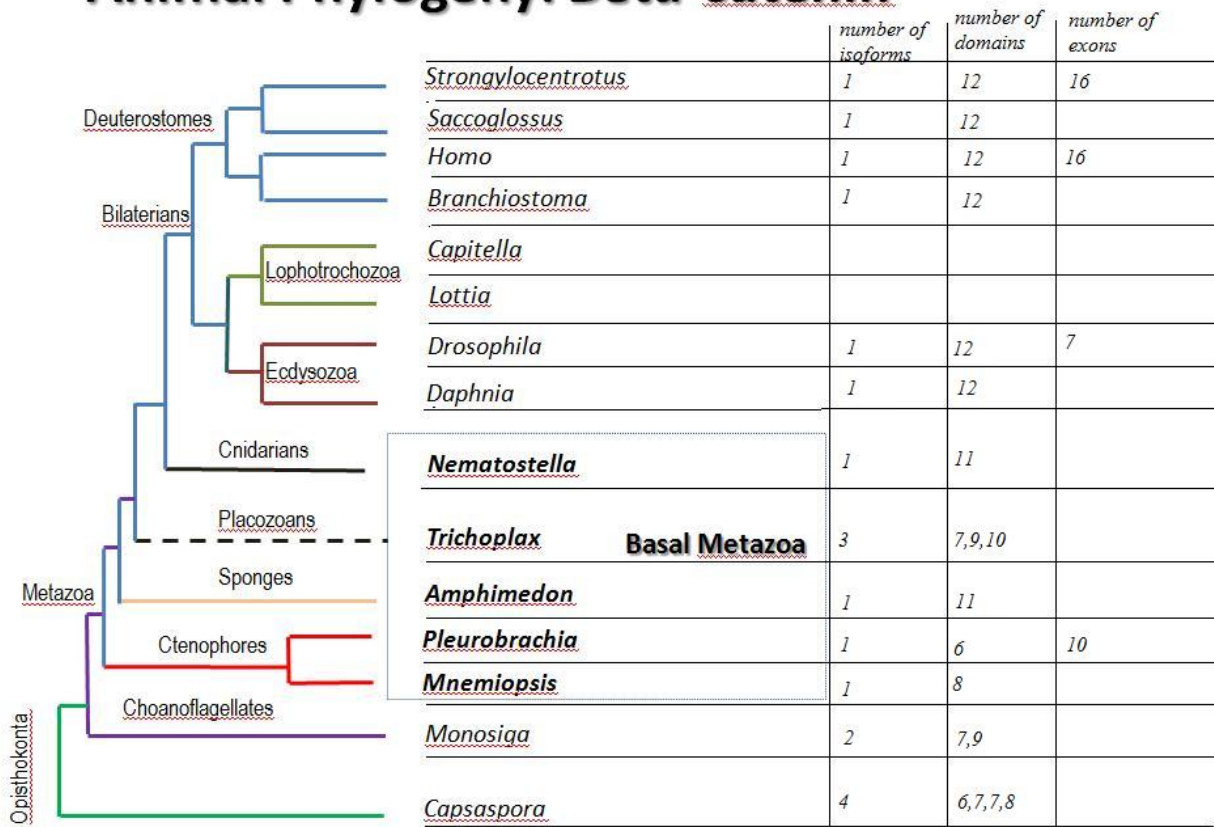
Blue is *Mnemiopsis* only
Green is *Pleurobrachia* and *Mnemiopsis*.

Supplemental Figure 1a. Alignment of β -catenin between *Mnemiopsis leidyi* and *Pleurobrachia bachei*. Alignments done in MAFFT, trimming done in G-block, and figure made in GeneDoc. Blue dashes indicate *Mnemiopsis* only, green indicates both.



Supplemental Figure 1c. Alignment of β -catenin between *Saccoglossus kowalevskii*, *Amphimedon queenslandica*, *Mnemiopsis leidyi*, and *Pleurobrachia bachei*. Alignments done in MAFFT, trimming done in G-block, and figure made in GeneDoc. The orange line is *Saccoglossus* only; blue is *Saccoglossus* and *Amphimedon*; purple is *Saccoglossus*, *Amphimedon*, and *Mnemiopsis*; red is *Saccoglossus*, *Amphimedon*, *Pleurobrachia* and *Mnemiopsis*; black is *Saccoglossus*, *Pleurobrachia* and *Mnemiopsis*.

Animal Phylogeny: *Beta-catenin*



Supplemental Figure 1. Number of isoforms, domains, and exons of β -*catenin* across phyla. Isoforms were found by NCBI search, domains were found by a SMART analysis, and number of exons was found by BLAT search. Reference numbers can be found in supplementary information.

References

- Castresana, J. (2000). "Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis." Society for Molecular Biology and Evolution: 540-552.
- Castresana, J., and Saraste, M. (1995). Does Vav bind to F-actin through a CH domain? **FEBS Letters** 374, 149-151.
- Dunn, C.W., et al. (2008) "Broad phylogenomic sampling improves resolution of the animal tree of life." *Nature* 452: 745-749.
- Einat Sadot, M. C.-S., Jacob Zhurinsky, Dalia Shnizer, Zeev Lando, Dorit Zharhary, and A. B.-Z. e. a. B. G. Zvi Kam (2002). "Regulation of S33/S37 phosphorylated β -catenin in normal and transformed cells." Journal of Cell Science **115**(13): 2771-2780.
- Information, N. C. f. B. (2012). "Blast." Retrieved May, 2012, 2012, from <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
- JORG SCHULTZ, F. M., PEER BORK, CHRIS P. PONTING (1998). "SMART, a simple modular architecture research tool: Identification of signaling domains." The National Academy of Sciences **95**: 5857-5864.
- Karl B. Nicholas, H. B. N. J., and David W. Deerfield II. (1991). GeneDoc: Analysis and Visualization of Genetic Variation Analysis and visualization of genetic variation.
- Katoh, K. (2011). MAFFT, AIST: Multiple alignment program for amino acid or nucleotide sequences.
- Letunic, I., T. Doerks, et al. (2012). "SMART 7: recent updates to the protein domain annotation resource." Nucleic Acids Res **40**(Database issue): D302-305.
- Pang, K., J. F. Ryan, et al. (2010). "Genomic insights into Wnt signaling in an early diverging metazoan, the ctenophore *Mnemiopsis leidyi*." Evodevo **1**(1): 10. Sciences, V. (2006). "Beta Catenin." Vala Sciences: 1-4.
- Simmons, D. K., K. Pang, et al. (2012). "Lim homeobox genes in the Ctenophore *Mnemiopsis leidyi*: the evolution of neural cell type specification." Evodevo **3**(1): 2.
- Srivastava, M., O. Simakov, et al. (2010). "The Amphimedon queenslandica genome and the evolution of animal complexity." Nature **466**(7307): 720-726.
- Steven P. White, C. C. (1987). "Structure of co-crystals tropomyosin and troponin." Nature **325**: 826-828.
- Talavera, G. and J. Castresana (2007). "Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments." Syst Biol **56**(4): 564-577.
- Tamura, K., D. Peterson, et al. (2011). "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Mol Biol Evol **28**(10): 2731-2739.
- THUMA, S. L. H. A. J. B. (2005). "Invertebrate Muscles: Muscle Specific Genes and Proteins." The American Physiological Society **85**: 1001-1028.
- UCSC. (2012, 18 april, 2012). "Genome Bioinformatics." Retrieved May, 2012, 2012, from <http://genome.ucsc.edu/>.
- William Lehman, R. C., Peter Vibert (1994). "Ca²⁺ induced tropomyosin movement in *Limulus* thin filaments revealed by three-dimensional reconstruction." Nature **368**: 65-67.