I have my notochord cells, but where did I leave my tail: A notochord gene expression analysis of tailed, tailless and hybrid Molgulids.

Elijah Lowe[1, 2], Billie J. Swalla[1, 3]

Evolution and Development of Metazons 2011
Summer 2011

[1]Friday Harbor Laboratories, University of Washington, Friday Harbor, WA 98250
[2]Department of Computer Science and Engineering & Quantitative Biology, Michigan State University, East Lansing, MI 48823
[3]Department of Biology, University of Washington, Seattle, WA 98195

Contact Information:
Elijah Lowe
Computer Science & Quantitative Biology
Michigan State University
GED Lab
5156 Biomedical and Physical Science Building, East Lansing, MI 48823
elijahlowe@gmail.com

Abstract

In the molgulid clade of ascidians several species have individually evolved tail loss during their larval stage of development. The notochord, one of the key features of chordates, is found in the center of the tail in most ascidian larvae. In almost all solitary ascidians, including *M. oculata*, 40 notochord cells have been found extended in the tail. However, in a closely related species, *M. occulta* (tailless), only 20 notochord cells were found, and these express *brachyury*. These two species are only 10% divergent at a transcriptomic level and are able to cross hybridize. Some of the hybrid also have 20 notochord cells, however the notochord cells extend and converge in a shortened version of *M. oculata* tail. Through the use of high throughput sequencing technologies and experimental techniques, we are able to investigate the express of genes associated with notochord development in the parent species and in the hybrid. Several genes—*FGF9/16/20, prickle (pk), noto6, leprecan*, *merlin,* and *noto17*—were analyzed for presence, temporal and spatial expression. *FGF9/16/20, pk, noto6, leprecan*, *merlin,* and *noto1* were found to be present in both species. *Pk* expression pattern was found to be similar to *Ci-pk2*.

Introduction

Although vertebrates and tunicates differ with their adult body plan, they share a similar developmental stage where their embryos display tadpole-like organization (Lemaire *et. al*, 2008). During the tadpole stage, ascidians have both a hollow dorsal nerve cord and a notochord (Swalla *et. al*, 1999). In ascidians, the larvae stage is the motile phase of the life cycle. During this time, the notochord helps structure the larval tail (Stemple, 2005), which allows for locomotion (Satoh, 2003). Interestingly, several ascidian species have lost their otolith and tail during development, which presumably limits their ability to disperse (Berrill, 1931). One such species is *Molgula occulta*, which can be found off the shores of Roscoff, France *M. occulta* has lost several chordate features including the notochord and tail muscle cells (Swalla, 1993). *M. oculata* is a closely related species in the same clade, also found in Roscoff. *M. oculata* has retained all of its chordate features in the tadpole larvae.In the tailed species—*M. oculata*—there are 40 notochord genes that are intercalated and extended throughout the tail . *M. occulta*, the tailless species only has 20 notochord cells; however, they do not intercalate or extend (Swalla and Jeffery, 1990).

Interspecies hybrids (*M. occulta* eggs and *M. oculata* sperm) are able to regain a shortened notochord—containing 20 notochord cells—along with other chordate features (Swalla and Jeffery, 1996). Several genes have been found to affect the development of tail loss and the notochord. *Brachyury,* a T-box transcription factor exclusive to the notochord in ascidians, is known to be necessary for notochord development (Swalla, 2006). *FGF9/16/20* has been found to induce *brachyury* expression, and a morpholino knockdown of *FGF9/16/20* led to almost a complete loss of *brachyury* expression at 110-cell stage in *Halocynthia.* Complete loss of *brachyury* expression is in contrast with that of *Ciona intestinalis*, which regain *brachyury* expression by tailbud. In Halocynthia, FGF9/16/20 is fully responsible for notochord induction (Kumano, 2006). *Prickle*, a downstream gene to *brachyury,* has been found as an initiating factor for ascidian tail extension and notochord polarity (Lemaire, 2008).

With advances in high throughput sequencing technologies, gene expression of *M. occulta*, *M. oculata*, and hybrid species can be analyzed (Gyoja et al., 2007; Pickrell et al., 2010). The transcriptome of the *M. occulta*, *M. oculata*, and hybrid species have been sequence by Titus Brown at Michigan State University in collaboration with the Swalla lab at the University of Washington. The transcriptomes were assembled *de novo* with no reference genome. There is not currently a genome sequenced for *M. oculata* or *M. occulta*. The three transcriptomes have been used to identify the presence or absence of known notochord genes using *Ciona* and *Halocynthia* data in the NCBI database. A BLAST search was conducted on ~40 known notochord genes, and several of them were selected for further analysis. *FGF9/16/20, prickle (pk),* and several other downstream *brachyury* factors—*noto6, leprecan*, *merlin,* and *noto17—*were analyzed for presence, temporal and spatial expression using *in situ* hybridizations. This investigation gives us a better understanding of the gene network for notochord specification in chordate development.

Methods and Materials

Animals and embryos

Adult *M. oculata* and *M. occulta* were collected in Station Biologique de Roscoff by Billie Swalla, and kept alive in tanks during the spawning season. Sperm and eggs were collected from both species and fertilized and developed in seawater. *M. oculata* eggs were fertilized using *M. oculata* sperm. *M. occulta* eggs were fertilized using sperm *M. occulta* sperm and also *M. oculata* sperm for hybrid embryos. Embryos were fixed and stored in ethanol.

Sequencing, alignments and gene identification

mRNA was collected at gastrula, neurula, and tailbud stages for *M. oculata*, *M. occulta*, and hybrid embryos. Sequencing was done using Illumina sequencing at Michigan State University and assembled my C. Titus Brown.

Known notochord genes from *C. intestinalis* and *Molgula tectiformis* were used to BLAST against the assembled *M. occulta*, *M. oculata*, and hybrid sequences. BLAST version 2.2.25 used for this analysis. The standalone BLAST program tblastn was used with assembled *M. occulta*, *M. oculata*, and hybrid transcriptome as a database. *C. intestinalis* and *M. tetiformis* gene sequences were retrieved from the NCBI protein database (http://www.ncbi.nlm.nih.gov/). A reciprocal BLAST was preformed using transfer cDNA Alignments of known homologues were performed using MAFFT v5 (Katoh, 2004).

Isolation of cDNA

FGF9/16/20 primers (forward: 5'-caattgtactgtcggacgggg-3' reverse 5'-cagaaattcacacatttttta-3') were designed using multiple sequence alignment output from MAFFT v5 software. cDNA libraries from *M. occulta* and *M. oculata* gastrula were used to conduct PCR. cDNA was be isolated using the methods in Swalla *et al,* (1993). cDNA for *ci-prickle, ci-noto6, ci-leprecan*, *mt-merlin, and mt-noto17* have previously been isolated by Andrykovich, and Swalla (2011).

In situ hybridization

The methods of Swalla *et al* (1993) were used to perform *in situ* hybridization. *In situ* hybridization were preformed on *M. occulta* and *hybrid* embryos at the gastrula, mid-gastrula and tailbud stages.

Tree generation

A gene trees was created for *prickle* using data from the sequenced M. oculata transcriptome. First, *M. oculata* cDNA sequences were translated using ExPASy (Boeckmann*, et. al*, 2003), and the translation with the least amount of stop codons was selected. Next a BLAST search for each gene sequence was performed using the NCBI database. All tunicate hits were included in addition to several vertebrates, and an organism from each of the deuterostome phylums—cephalochordate, hemichordate, and echinoderm. The homologous gene set was then realigned using MAFFT v5, followed by Gblocks (Castresana, *et al*, 2000; Talavera, *et al*, 2007) to remove highly unconserved regions. Maximum likely-hood trees, with 100 bootstrap were created using MEGA5 (Tamura, *et al*, 2011).

Results

Noto6, leprecan, *merlin,* and *noto17*

Andrykovich and Swalla (2011) amplified *Noto6, leprecan*, *merlin,* and *noto17* through PCR. All genes were found to be present in both *M. oculata* and *M. occulta* (data not shown). This shows the presence of said genes in both species. PCR bands were cut and purified and minipreped. Many of the vectors contained no inserts, because Xgal was not used for blue/white colony staining, and colonies had to be chosen at random.

FGF9/16/20

The effect of *FGF9/16/20* has been studied in *Halocynthia roretzi* (Kumano, *et. al*, 2006), and two species of *Ciona* (Imai, et. al, 2002; Imai, et. al, 2004; Yasuo and Hudson, 2007), all showed effects of notochord development. In *Ciona*, *FGF9/16/20* was shown to only be an initiation factor and other mechanisms were involved in later notochord development. However, in *H. roretzi*, *FGF9/16/20* was found to be fully responsible for notochord induction. *M. oculata* and *M. occulta* more closely resembles *H. roretzi* than *Ciona (*figure 1).

FGF9/16/20 was only found in *M. oculata* during the transcriptome analysis. However, using primers design from the *M. oculata* transcriptome sequence, *FGF9/16/20* was amplified in both *M. oculata* and *M. occulta* with bands of size 373bp.

Prickle

Di Jiang *et al,* (2008) analyzed the *aimless* mutant in *Ciona savignyi*, which contained a deletion in the *pk* gene. *C. savignyi* with *aimless* mutations contained 40 notochord cells with dramatically shortened tails. No other defects were seen. The lack of *prickle* expression created two columns of notochord cells instead of the normal single column (figure 2). This was a morphogenetic defect, rather than a tissue differentiation defect (Jiang, *et al*, 2008).

Previously, *prickle 1* (*pk1*) and *prickle 2* (*pk2*) homologues were found in *C. intestinalis*. A PET domain and three LIM domains—which are protein-protein interaction domains—are conserved in both *ci-pk1* and *ci-pk2*. *Ci-pk1* and *ci-pk2* are identical in the 5' region and differ in the 3' prime region (Hotta, K., *et. al*, 2000). The transcriptome data for *M. oculata* and *M. occulta* shows preservation of PET and the three LIM conserved domains (Andrykovich and Swalla, 2011). Unfortunately, the transcriptome data does not contain the entire 3' region of the *M. occulta* or *M. oculata prickle* protein sequence as a signal contig and render it difficult to identify whether the discovered *pk* is more similar to *Ci-pk1* or *Ci-pk2*. The gene tree for *M. oculata pk* (*Mocu-pk*) shows *Mocu-pk* to be more closely related to *Ci-pk2*. *Mocu-pk* is closer to *Ci-pk2* in the gene tree (figure 3) because *Ci-pk2* is a shorter protein sequence than *Ci-pk1*; this artifact also occurs in BLAST searches because of sequence length (figure 4).

The wholemount *in situs* for both *M. occulta* and the *M. oculata* x *M. occulta* hybrid showed expression in the notochord cell line (figure 5). The expression patterns resembled that of *Ci-pk2* done by Hotta, *et. al* (2000). Strong staining could not be obtained using the *Prickle* probes so new probes are needed. Probes with a higher concentration of RNA would be ideal.

Discussion

Starting with the transcriptome analysis, the expression of known notochord genes were analyzed in *M. oculata, M. occulta*, and the *M. oculata x M. occulta* hybrid. The transcriptome was assembled without a reference genome. To test the validity of the transcriptome assembly, *M. oculata and M. occulta* found on then NCBI database were used. Ninety-six out of ninety-eight Sanger sequences were found within the transcriptome (data not shown).

Initially ~40 notochord genes were searched using BLAST. Several genes were selected because of their interaction with *brachyury,* a T-box transcription factor that has been found to induce the notochord. *FGF9/16/20* signaling gene was chosen because of its activation of *brachyury* and was identified as inducer of the notochord (Kumano, 2006; Hotta, 2008). In both *Halocynthia* and *Ciona*, the MO knockdown of *FGF9/16/20* completely eliminated and temporarily caused *brachyury* not to express, respectively.

During BLAST screening of the *Molgula* transcripts *ci-fgf9/16/20* was only found in *M. oculata*. The presence of *FGF9/16/20* in only *M. oculata* was proven to be incorrect through the PCR amplification of *FGF9/16/20* in *M. occulta* gastrula cDNA library. *Mt-noto17* was also only found in *M. oculata* during the BLAST screening and was able to be amplified in *M. occulta*.

Several downstream genes—*ci-prickle, ci-noto6, ci-leprecan*, and *mt-merlin*—were all found to have sequence matching for both *M. oculata* and *M. occluta* using BLAST search. These genes could also be amplified using PCR in both *M. oculata* and *M. occulta*. *Prickle* was expressed in the primary notochord lineage (A7.3 & A7.7 lines). It was not computationally possible to identify *prickle* as either *prickle 1* or *prickle 2* because of the contig size; this can be further investigated with a complete genome. RNA probes for the remaining genes—*noto6, leprecan, merlin, noto17* and *FGF9/16/20*—could not be made because of time constraints. *Leprecan* is predicted to have ubiquitous expression throughout the embryo. As for *FGF9/16/20*, expression in the a8.7, a8.8, a8.15, a8.16, b8.7, b8.8 line cells in *M. oculata* and in *M. occulta* is expected. The predictions for *leprecan* and *FGF9/16/20* are based on staining from the ANISEED database (http://www.aniseed.cnrs.fr/). The remaining three genes—*noto6, noto17 and merlin* did not have staining in the ANISEED database. I predict that the remaining genes—*noto6, noto17 and merlin*—will be expressed in the notochord lineage.

The absence of FGF9/16/20 in the computational screening of *M. occulta* negatively correlates with the PCR amplification of FGF9/16/20 in *M. occulta*. The negative correlation with the BLAST search support two points—computational analysis is a tool for hypothesis generation, not conclusive evidence, and a complete genome is needed to better analyze computational data.

Future work

In addition to the work already completed, more work needs to be done both computationally and experimentally. On the computational side, bio-replicates need to be sequence to test for the validity of the transcriptome. There is a need for a complete genome, and transcriptomes at various stages—16 cell stage to late tailbud—to model the gene network. The next step for the experimental analyzsis is to create probes for the remaining genes—*FGF9/16/20*, *noto6*, *leprecan*, *merlin,* and *noto17*—and analyze their expression patterns. In addition to *in situ* hybridizations, quantitative PCR is needed to examine the expression levels of the genes. Also, western blots are needed to determine which *prickle* protein was found in the *Molgula* species. The gene set will be increased to include several more interacting genes such as *dsh* and *stbm*—which interacts with *prickle*.

Acknowledgments

I would like to thank Billie Swalla and C. Titus Brown for giving me the opportunity to work on such an exciting project and for their expertise. I would also like to thank Ken Halanych for his expertise. Thanks Kevin Kocot for all of his help and support. Special thanks to Joie Cannon for her assistance in lab and helping in coming up with various protocol dances. I would also like to thank my classmates from the Evolution and Development of Metazoans class, especially Nathan Farrar, and Tasneem Pierce. I would like to acknowledge Ryan Gray for dinning hall humor. The Fernald Fellowship and the BEACON center for the study of evolution in action provided funding.

Reference:

Andrykovich, K., and Swalla, B.J. (2011). Investigating the molecular basis of notochord loss in Molgula occulta via transcriptome sequencing. FHL, 1-36

Berrill, N.J. (1931). Studies in Tunicate Development. Part II. Abbreviation of Development in the Molgulidae. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character. 219, 281-346.

Boeckmann B., Bairoch A., Apweiler R., Blatter M.-C., Estreicher A., Gasteiger E., Martin M.J., Michoud K., O'Donovan C., Phan I., Pilbout S. and Schneider M., The Swiss-Prot Protein Knowledgebase and its supplement TrEMBL (2003). Nucleic Acids Res. 31, 365-370.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17, 540-552.

Gyoja, F., Satou, Y., Shin-i, T., Kohara, Y., Swalla, B.J., and Satoh, N. (2007). Analysis of large scale expression sequenced tags (ESTs) from the anural ascidian, Molgula tectiformis. Developmental Biology. 307, 460-482.

Hotta, K., et al. 2000. Characterization of Brachyury-downstream notochord genes in the Ciona intestinalis embryo. Dev Biol 224, 69–80.

Hotta, K., Takahashi, H., Satoh, N., and Gojobori, T. (2008). Brachyury-downstream gene sets in a chordate, Ciona intestinalis: integrating notochord specification, morphogenesis and chordate evolution. Evo. & Dev. 10, 37-51.

Imai, K.S., Satoh, N., Satou, Y., (2002). Early embryonic expression of FGF4/6/9 gene and its role in the induction of mesenchyme and notchord in Ciona savignyi embryos. Development. 129, 1729-1738.

Imai, S., Kino, K., Yagi, K., Satoh, N., Satou, Y., (2004). Gene expression profiles of transcrition factrors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. Development. 131, 4047-4058.

Jeffery, W.R. (2002). Ascidian gene-expression profiles. Genome Biology. 3, 1030.1-1030.4

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2004). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucl. Acids Res. 33 (2), 511-518.

Kugler, J.E., Kerner, P., Bouquet, J., Jiang, D., and Gregorio, A. (2011). Evolutionary changes in the notochord genetic toolkit: a comparative analysis of notochord genes in the ascidian Ciona and the larvacean Oikopleura. BMC Evo. Bio. 11:20

Kumano, G. et al. (2006). Overlapping expression of FoxA and Zic confers responsiveness to FGF signaling to specify notochord in ascidian embryos. Developmental biology. 30, 770-784.

Lemaire, P., Smith, S.C., and Nishida, H. (2008). Ascidians and the Plasticity of the Chordate Developmental Program. Current Biology. 18, R620-R631.

Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J., Stephens, M., Gilad, Y., and Pritchard, J.K. (2010). Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature. 464, 768-772.

Sotah, N. (2003). The ascidian tadpole larva: comparative molecular development and genomics. Nat. Rev. Genet. 4, 285-295.

Stemple, D.L. (2005). Structure and function of the notochord: an essential organ for chordate development. Development. 132, 2503-2515

Swalla, B.J., Just, M.A., Pederson, E.L., and Jeffery, W.R. (1999). A multigene locus containing the Manx and bobcat genes is required for development of chordate features in the ascidian tadpole larva. Development. 126, 1643-1653.

Swalla, BJ. (1993). Novel genes expressed differentially in ascidians with alternate modes of development. Development. 19, 301-318

Swalla, BJ. (2006). Building divergent body plans with similar genetic pathways. Heredity. 97, 235-243

Swalla, B. J., and Jeffery, W. R. (1996). Interspecific hybridization between an anural and urodele ascidian: differential expression of urodele features suggests multiple mechanisms control anural development. Dev. Biol. 142, 319-334.

Swalla, B. J., and Jeffery, W. R. (1996). Requirement of the Manx gene for the expression of chordate features in a tailless ascidian larva. Science 274, 1205–1208.

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology 56, 564-577.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution doi: 10.1093/molbev/msr121.

Veeman, M., Nakatani, Y., Hendrickson, C., Ericson, V., Lin, C., and Smith., W. (2007). *chongmague* reveals an essential role for laminin-mediated boundary formation in chordate convergence and extension movements. Development. 135, 33-41.

Yasuo, H., and Hudson, C. (2007). FGF8/17/18 functions together with *FGF9/16/20* during formation of the notochord in Ciona embryos. Dev. Biol. 302, 92-103.

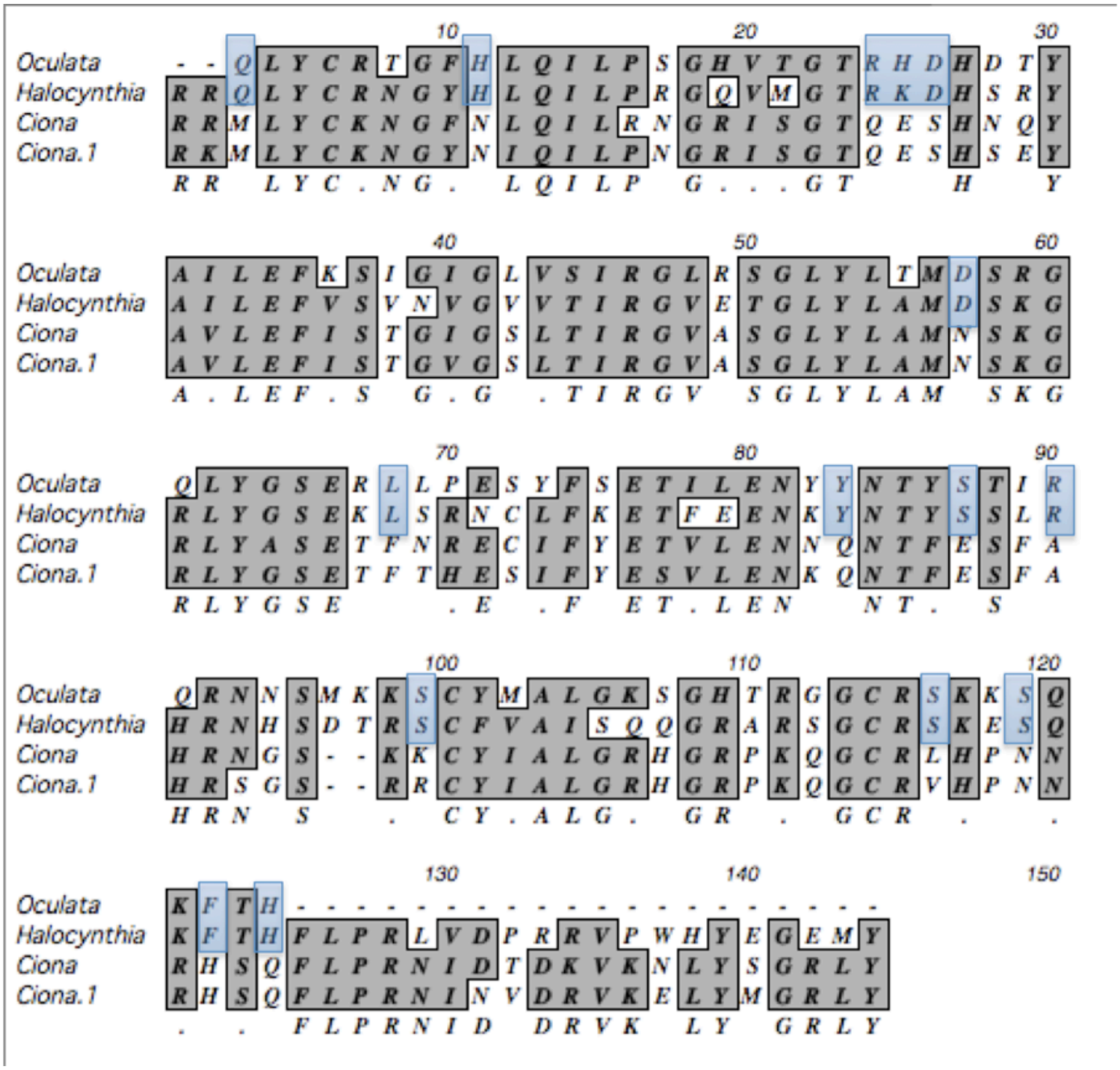


Figure 1 Sequence comparison of *M. oculata, Halocynthia roretzi*, *Ciona intestinalis* (Ciona), and *Ciona savignyi* (Ciona 1). The blue regions show areas where *M. oculata* and *H. roretzi* are more conserved.

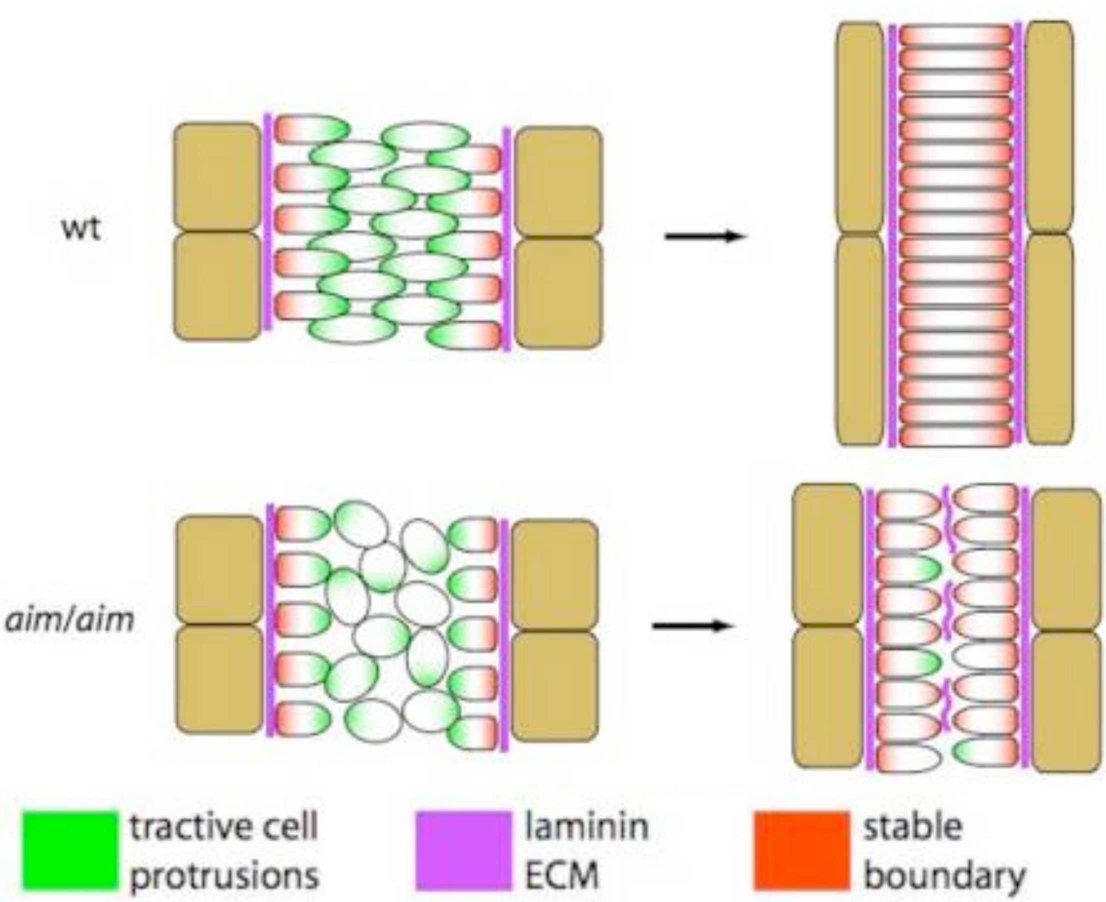
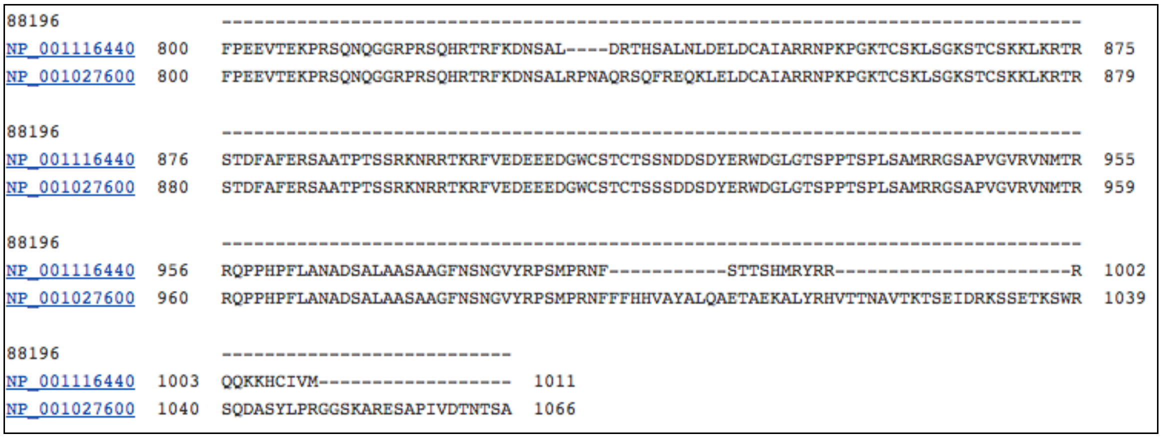


Figure 2 figure that from (Veeman et al, 2007). Aim/aim shows the effect of mutant prickle; polarity is loss and two columns form.



```
88196                    --------------------------------------------------------------------------------
NP_001116440   800    FPEEVTEKPRSQNQGGRPRSQHRTRFKDNSAL----DRTHSALNLDELDCAIARRNPKPGKTCSKLSGKSTCSKKLKRTR  875
NP_001027600   800    FPEEVTEKPRSQNQGGRPRSQHRTRFKDNSALRPNAQRSQFREQKLELDCAIARRNPKPGKTCSKLSGKSTCSKKLKRTR  879

88196                    --------------------------------------------------------------------------------
NP_001116440   876    STDFAFERSAATPTSSRKNRRTKRFVEDEEEDGWCSTCTSSNDDSDYERWDGLGTSPPTSPLSAMRRGSAPVGVRVNMTR  955
NP_001027600   880    STDFAFERSAATPTSSRKNRRTKRFVEDEEEDGWCSTCTSSSDDSDYERWDGLGTSPPTSPLSAMRRGSAPVGVRVNMTR  959

88196                    --------------------------------------------------------------------------------
NP_001116440   956    RQPPHPFLANADSALAASAAGFNSNGVYRPSMPRNF-----------STTSHMRYRR----------------------R  1002
NP_001027600   960    RQPPHPFLANADSALAASAAGFNSNGVYRPSMPRNFFFHHVAYALQAETAEKALYRHVTTNAVTKTSEIDRKSSETKSWR  1039

88196                    ---------------------------
NP_001116440   1003   QQKKHCIVM------------------  1011
NP_001027600   1040   SQDASYLPRGGSKARESAPIVDTNTSA  1066
```

Figure 3 *Ci-pk2* (NP_001116440) 3' end is shorter than *Ci-pk1* (NP_001027600), which causes the shorter *M. oculata* contig (88196) to have a high BLAST score when align with Ci-pk2.

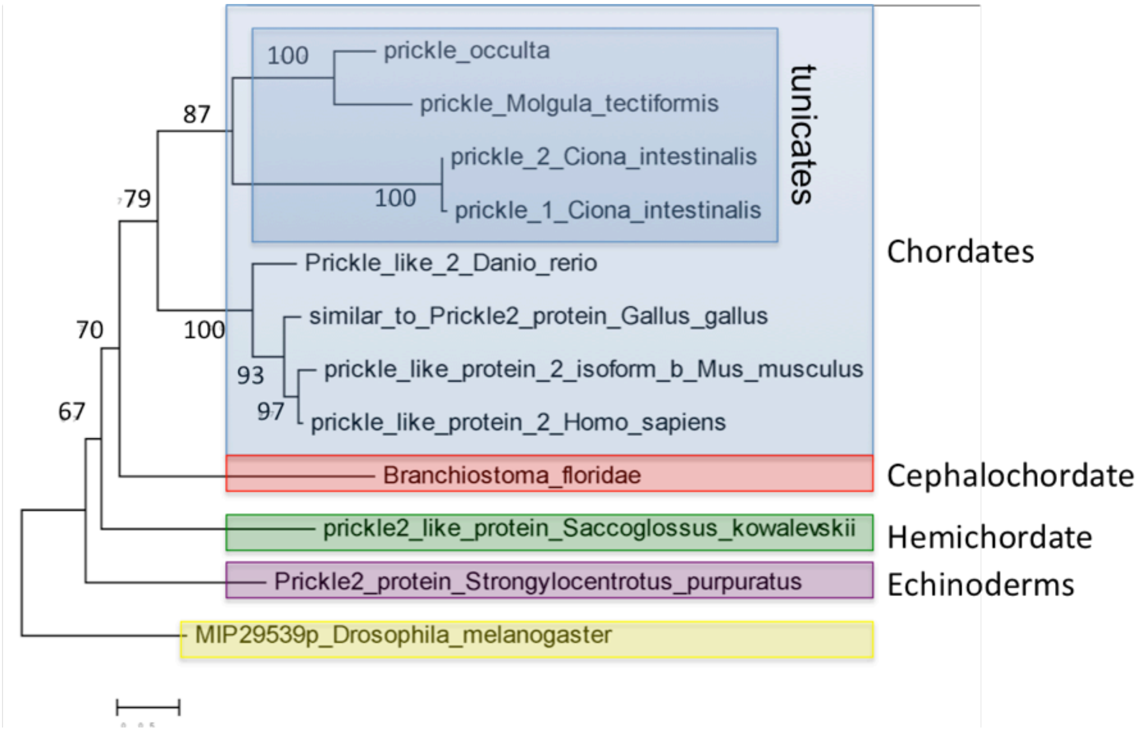


Figure 4 Prickle gene tree. The Chordate clade is conserved, and the tree closely resembles current phylogenetic trees.

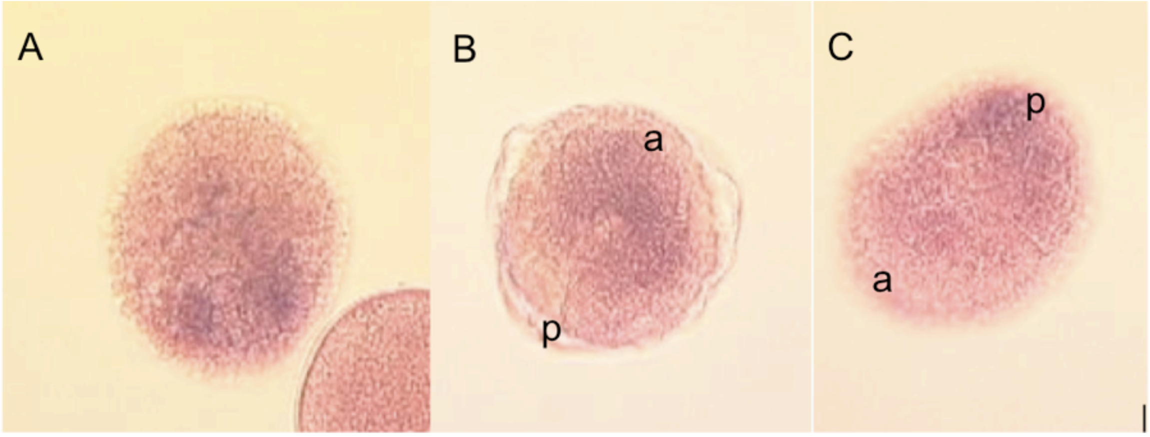


Figure 5 Prickle WMIH expression patterns. A) Hybrid embryo at tailbud. B) *M. occulta* embryo at gastrula stage. Expression is show in the notochord cell line. C) *M. occulta* embryo at tailbud stage. Expression is shown in the tailed region.

```
lcl|7259         -PGKRINIEEVESLPRKARLRSH--TRSRQMHKKLPT---------FDDVESPKILDENT
gi|110592138     WSGSRIPVNELDKLPISAEATQHQNTRSRILSIRSKTRRPTDVRSLSDDNSIPQRLTDST
gi|80751119      --------HGLESF--AH--GSH-------------------------------------
gi|118090462     --------GGLDGL--GQPVGAH-------------------------------------
gi|300793744     --------GGLEGL--GQQVGSH-------------------------------------
gi|74096169      SSMSRTEAAKVQSV-LSAKLGRA--------------RR--K-NSLSVDSSTPQKLTEET
gi|19912822      NSMPRIEAAKVRSL-FPAKLGRR--------------RR--QPSSLESNSPVPQKLTEDT
                           :  .

lcl|7259         MPLVHDRLPDPELWKKLDFDGEESDED----------SNDKELNL-EVKKTHR-VARS--
gi|110592138     VPIVHDRLPDPKLWEKLNFGTEVEVDDDTLTDLSPQTEEDSSYSLPESSSTHKITKRS--
gi|80751119      -------------------------------FVLTPPGDSGALMNEHLVSAER-LSRS--
gi|118090462     -------------------------------FLLPPAGERPALLGEQRGPAER-AA----
gi|300793744     -------------------------------FLLPPAGERPPLLGERRSAAER-GARG--
gi|74096169      MPVVHDRLPDLSLWSKIDEELKEQ-------------EKSSANGLLS-STSAR-SKRNT-
gi|19912822      MPVVHDRFPDLTLWKKIEDDMKVQ-------------EKKLSLGLPSHSQNTR-SKRNIY
                                                       .             :

lcl|7259         ---VS---PVTRGVVRYKQLYCRTGFHLQILPSGHVTGTRHDHDSYAILEFKSISIGVVS
gi|110592138     ---IS---AGLEGIMRRRQLYCRNGYHLQILPRGQVMGTRKDHSRYAILEFVSVNVGVVT
gi|80751119      ---TSPDLTHLKGILRRRQLYCRTGYHLEILPDGSVQGTRKDHSRFGILEFISLAVGMVS
gi|118090462     ---AAAELAHLQGILRRRQLYCRTGFHLQILPDGSVRGTRQDHSLFGILEFISVAVGLVS
gi|300793744     -GPGAAELAHLHGFLRRRQLYCRTGFHLQILPDGSVQGTRQDHSLFGILEFISVAVGLVS
gi|74096169      -GMSAYDYDTQETTMRRRMLYCKNGFNLQILRNGRISGTQESHNQYAVLEFISTGIGSLT
gi|19912822      PFMQKYDKDTPQVTMRRKMLYCKNGYNIQILPNGRISGTQESHSEYAVLEFISTGVGSLT
                            .  :* : ***:.*::::**  * : **:..*. :.:*** *  :* ::

lcl|7259         IRGLRSGLYLTMDSRGQLYGSERLLPESYFSETILENYYNTYSTIRQRNNSMKKSCYLAL
gi|110592138     IRGVETGLYLAMDSKGRLYGSEKLSRNCLFKETFEENKYNTYSSLRHRNHSDTRSCFVAI
gi|80751119      IRGVDSGLYLGMNSKGELYGSEKLSAECVFREQFEENWYNTYSSNLYRHGERGSHYYVAL
gi|118090462     IRGVDSGLYLGMNEKGELYGSEKLTSECIFREQFEENWYNTYSSNVYKHGDSGRRYFVAL
gi|300793744     IRGVDSGLYLGMNDKGELYGSEKLTSECIFREQFEENWYNTYSSNIYKHGDTGRRYFVAL
gi|74096169      IRGVASGLYLAMNSKGRLYASETFNRECIFYETVLENNQNTFESFAHRNGS--KKCYIAL
gi|19912822      IRGVASGLYLAMNSKGRLYGSETFTHESIFYESVLENKQNTFESFAHRSGS--RRCYIAL
                 ***: :**** *:.:*.**.** :  :. * * . **  **:.:   :  .     ::*:

lcl|7259         GKSGHTRGGCRSKKSQKFTHFLPRDVYRHKVPHLYTGVD--Y
gi|110592138     SQQGRARSGCRSKESQKFTHFLPRLVDPRRVPWHYEGEM--Y
gi|80751119      NKDGTSRDGARSRRHQRFTHFLPRPVDPERVPELYKDVLGHS
gi|118090462     NKDGTPRDGARSKRHQKFTHFLPRPVDPERVPELYKDVLGYS
gi|300793744     NKDGTPRDGARSKRHQKFTHFLPRPVDPERVPELYKDLLMYS
gi|74096169      GRHGRPKQGCRLHPNNRHSQFLPRNIDTDKVKNLYSGRL--Y
gi|19912822      GRHGRPKQGCRVHPNNRHSQFLPRNINVDRVKELYMGRL--Y
                 .: * .: *.* :  ::.::**** :   :*   * .
```

Figure 6 Multiple sequence alignment of—*M. oculata* (lcl|7259), *Danio rerio* (gi|110592138), *Gallus gallus* (gi|118090464), *Bos taurus* (gi|300793744), *C. intestinalis* (gi|74096169) and *C. savignyi*—protein sequences for FGF9/16/20 used to identify conserved regions. Complete conservation is marked by a star (*), high conservation is marked by a colon (:), and low conservation is marked by a period (.). The box shows the area primers where designed in *M. oculata* for PCR amplification.

Figure 7 A) Forward and reverse primers for FGF9/16/20. B) Translated cDNA sequences with primers highlighted in red.

```
Mofgf9/16/20 F1
      caattgtactgtcggacgggg
      Q  L  Y  C  R  T  G

Mofgf9/16/20 R1
      cagaaattcacacatttttta
      Q  K  F  T  H  F  L

RC:taaaaatgtgtgaatttctg

    6 atgtcgcttagtcacgtcattccgacagctagcgaaagcaatata
      M  S  L  S  H  V  I  P  T  A  S  E  S  N  I
   51 gcatcagtactaatatcaatgctattgaaagcaaagaaatattca
      A  S  V  L  I  S  M  L  L  K  A  K  K  Y  S
   96 ccattcagtaatacggataatacatcatcacaacttccagggaaa
      P  F  S  N  T  D  N  T  S  S  Q  L  P  G  K
  141 aggataaatattgaagaagttgaatctttaccaagaaaagcaaga
      R  I  N  I  E  E  V  E  S  L  P  R  K  A  R
  186 ctgagaagccatacaagaagcagacaaatgcacaagaaattgcca
      L  R  S  H  T  R  S  R  Q  M  H  K  K  L  P
  231 acatttgacgacgtcgaaagtccaaaaatactcgacgaaaacaca
      T  F  D  D  V  E  S  P  K  I  L  D  E  N  T
  276 atgccactcgtacatgatcgtcttcctgatcctgagttgtggaag
      M  P  L  V  H  D  R  L  P  D  P  E  L  W  K
  321 aaattagactttgacggagaagaatctgatgaagatagcaatgat
      K  L  D  F  D  G  E  E  S  D  E  D  S  N  D
  366 aaagaattaaatttggaagtaaagaaaacgcatcgcgttgcacga
      K  E  L  N  L  E  V  K  K  T  H  R  V  A  R
  411 agtgtatcgccagtaacaagaggagtggttcgctacaaacaattg
      S  V  S  P  V  T  R  G  V  V  R  Y  K  Q  L
  456 tactgtcggacggggtttcatcttcaaattttaccgagcggacac
      Y  C  R  T  G  F  H  L  Q  I  L  P  S  G  H
  501 gtaacaggcacccgccacgatcatgattcttatgccattcttgaa
      V  T  G  T  R  H  D  H  D  S  Y  A  I  L  E
  546 tttaaatccatcagtatcggggtggtcagcatacgtggtcttcga
      F  K  S  I  S  I  G  V  V  S  I  R  G  L  R
  591 agtggactttatttaacaatggattccagaggtcaactttacggc
      S  G  L  Y  L  T  M  D  S  R  G  Q  L  Y  G
  636 tcggaacggctgttacctgaaagttacttcagtgaaactatattg
      S  E  R  L  L  P  E  S  Y  F  S  E  T  I  L
  681 gaaaactattacaacacatattcaacaattcgacaaaggaataat
      E  N  Y  Y  N  T  Y  S  T  I  R  Q  R  N  N
  726 tctatgaagaaatcctgttatttggctcttggaaaatctggccac
      S  M  K  K  S  C  Y  L  A  L  G  K  S  G  H
  771 acacgaggaggatgcagatcgaaaaaatcacagaaattcacacat
      T  R  G  G  C  R  S  K  K  S  Q  K  F  T  H
  816 tttttaccgcgggacgtatatcgacataaagtgcctcacttgtac
      F  L  P  R  D  V  Y  R  H  K  V  P  H  L  Y
  861 acaggggttgattattga 878
      T  G  V  D  Y  *
```