Expression of Pax-2/5/8 in *Helobdella* sp. (Austin)

Presented by Brennan Lang

In partial fulfillment of the requirements for graduation with the
Dean's Scholars Honor's Degree in the School of Biological Sciences

_____ _____

Marty Shankland Date
Supervising Professor

_____ _____

Ruth Buskirk Date
Honor's Adviser in the School of Biological Sciences

I grant the Dean's Scholars Program permission to (check all that apply):

_____Post a copy of my thesis on the Dean's Scholars website

_____Keep a copy of my thesis in the Dean's Scholars office for Dean's Scholars to read

Expression of Pax-2/5/8 in *Helobdella* sp. (Austin)

Department: School of Biological Sciences

_____
Author's Printed Name

_____ _____
Signature Date

_____
Supervising Professor's Printed Name

_____ _____
Signature Date

Expression of Pax-2/5/8 in *Helobdella* sp. (Austin)

Brennan Lang* and Marty Shankland

Section of Molecular Cell and Developmental Biology, Institute of Cellular and Molecular Biology, University of Texas, Austin, TX 78712, USA

*Author of correspondence (email: brennan.lang@yahoo.com)

ABSTRACT

Pax-2/5/8 is a member of the Group II Pax genes and, like its relatives, is an important transcription factor most likely involved in cellular differentiation during embryonic development. This highly conserved gene can be found throughout the three superphyla of all bilaterally symmetric animals: Deuterostomes, Ecdysozoans, and Lophotrochozoans. Although Group II Pax genes are involved in the organogenesis and neural patterning of both Deuterostomes and Ecdysozoans, it has yet been unclear what particular role Pax-2/5/8 plays in Lophotrochozoan development. This study, of annelid leech Pax-2/5/8 expression, should help to further elucidate the role of Pax-2/5/8 expression in Lophotrochozoans. More specifically, this project's aim was to fully sequence the Pax-2/5/8 gene from *Helobdella* sp. (Austin), and to then begin the characterization of its expression pattern in the developing leech embryo. This preliminary study into annelid Pax-2/5/8 expression has verified the role of Pax-2/5/8 in organogenesis and more specifically in the formation of the leech excretory system. However, further functional studies are necessary to determine if Pax-2/5/8 plays a further role in leech nervous system patterning as well as the specific phenotypic effects that result from disrupted Pax-2/5/8 expression.

BACKGROUND

Pax Genes

Embryonic development is a highly complex process. In the leech and other segmented organisms, it involves the coordination of many different genes that regulate the segmentation and subsequent cellular differentiation of those segments. The Pax family of transcription factors and their role in early development has been largely conserved over millions of years of evolution, and as a result, the Pax family plays a major role in a wide variety of multicellular organisms. As a highly conserved family of genes, the Pax family has been studied throughout the three superphyla of bilaterally symmetrical animals: Deuterostomes, which include humans and other mammals, Ecdysozoans, which include insects, and Lophotrochozoans, which include mollusks and annelids like the leech.

The Pax family of transcription factors act as tissue-specific transcriptional regulators and play very important roles in development and cell-type differentiation (Chi and Epstein 2002; Lang et al. 2007). In mammals, there are a total of nine Pax genes that are divided into four subgroups based on sequence similarity and gene structure (Walther et al. 1991). According to molecular phylogenetic studies the last common ancestor of bilaterian animals had five Pax genes, roughly corresponding to the four mammalian subgroups: Pax-1/9, Pax-2/5/8, Pax-3/7, Pax-4/6, and Pox neuro, which is missing in chordates (Balczarek et al. 1997; Sun et al. 1997). Only recently has a novel Pax gene, Pax-β, been discovered that appears to be restricted to the bilaterian superphylum Lophotrochozoa (Schmerer et al. 2009). The role of Pax genes in bilaterian development ranges from eye formation via Pax-6, to the role of Pax-3/7 in segmentation, to the roles of Pox neuro and Pax-2/5/8 in neural patterning (Matus et al. 2007).

All Pax genes are characterized by a highly conserved 128 amino acid DNA-binding motif known as a Paired domain. Many Pax genes also contain a complete or truncated homeodomain and/or an eight amino acid octapeptide region (Balczarek et al. 1997). The octapeptide region, in particular, has also been highly conserved throughout the Pax gene family as well. It has also been examined as a potential protein-binding site of directional modification by methylation or demethylation (Ziman and Kay 1998). Ongoing studies suggest that these Pax transcription factors function primarily by binding enhancer regions of genes and modifying their transcriptional activity at multiple stages of development (Chi and Epstein 2002).

Embryonic Development of *Helobdella*

Helobdella sp. (Austin), like *Helobdella robusta*, is a small (2-4 cm) glossiphoniid leech and a member of Lophotrochozoa. Over the past few decades *Helobdella* has become an increasingly useful model for both annelid and lophotrochozoan development. Its usefulness stems from its comparatively small genome size, with nearly 95% of the approximately 350 Mbp of genomic DNA sequence made available in 2007, as well as the relative ease with which it can be maintained in the laboratory environment. The hermaphroditic leech is capable of both self- and cross-fertilization year round. It produces several external membranous sacs totaling approximately 20-100 embryos, each reproductive cycle. These embryos can be easily removed from the underside of brooding adults and cultured in simple saline solutions. Observation and experimentation on all stages of leech development can be achieved by culturing embryos from single cell zygotes through their juvenile stages. *Helobdella*, like other annelids, also follows an invariable sequence of embryonic cleavages. Unlike insects and many

other invertebrates, the highly stereotyped cell-lineage development program of the leech results in its segmental body plan (Shankland 1991).

After a mother has deposited her eggs within membranous sacs attached to her ventral side, the first six stages of embryonic cleavage will occur within 24 hours. After five more stages (7-11) of gastrulation, segmentation, and organogenesis the resulting juvenile will have been formed, taking in total approximately 10 days.

Shortly before the first cleavage event of the leech zygote, two pools of teloplasm, transparent yolk-free cytoplasm rich in maternal RNAs, begin to form at the animal and vegetal poles. The first asymmetric cleavage event begins with the formation of a cleavage furrow along the animal-vegetal axis. As a result of this first cleavage, the zygote is divided into two unequal daughter cells, with the larger of the two cells containing the two pools of teloplasm. After the second round of cleavage, four macromeres have been formed (A, B, C, D), the largest of which (D) has received the vast majority of the teloplasm. After this four-cell stage the embryo enters into a series of spiral cleavage events, which results in the formation of two distinct populations of cells. The micromeres are a group of cells descendant from the A, B, and C macromeres, with a single micromere coming from the D macromere during the formation of the first micromere quartet. As the embryo develops the micromeres eventually develop into the unsegmented head region of the leech as well as the provisional integument, a thin temporary layer of epidermal cells that expands during stage 8 and is eventually discarded during the epiboly and germinal band migration events of stage 10. The second group of cells is called teloblasts and is composed of large embryonic stem cells that are primarily descendant

from the D macromere. These teloblasts will eventually become the mesoderm and ectoderm of the mature segmented adult leech.

During embryonic development the leech will develop exactly 32 body segments. These segments are generated in an anterior to posterior progression from a posterior growth zone, composed of five bilateral pairs of teloblasts (M, N, O, P, Q) formed from the D teloblast (Fig. 1). As a result of the fourth round of cell division the D macromere is divided into cells DNOPQ and DM. The DNOPQ cell will give rise to the segmented trunk ectoderm via teloblasts N, O, P, and Q, while the vegetally located DM cell will give rise to the segmented trunk mesoderm via the M teloblast.

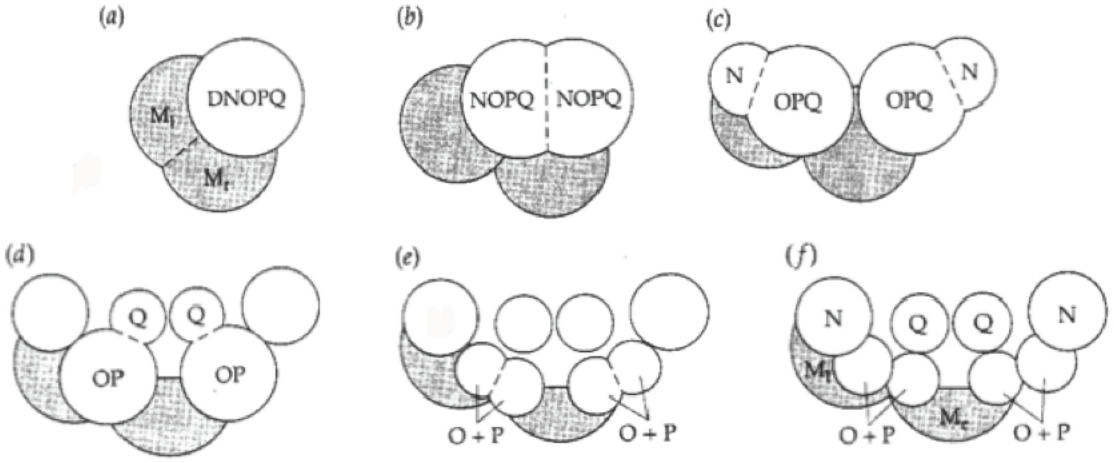

Fig. 1. Formation of D-derived teloblasts (Adapted from Shankland and Savage, 1997)

After all five pairs of teloblast have been formed, each teloblast undergoes a series of highly asymmetrical cell divisions. The result of this repetitive cellular division is the production of ten chains of blast cells that coalesce into separate right and left germinal bands (Fig. 2).

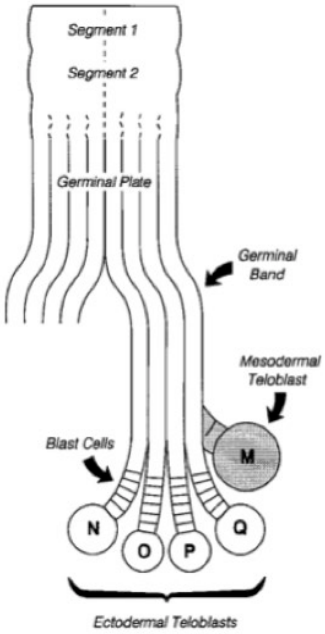

Fig. 2. Stylized rendition of leech development (adapted from Nardelli-Haeiger et al. 1994)

With their leading ends anchored at the micromere region of the animal pole, the elongating germinal bands spread over the surface of the underlying macromeres; 'older' blast cells are pushed towards the anterior end of the embryo while new cells are continually generated at the posterior end. As the germinal bands continue to loop around the embryo, in a process called epiboly, they begin to meet and subsequently fuse together along the ventral midline, forming the germinal plate. At this time segmental differentiation begins. During stage 9, the embryo hatches from a vitelline membrane, which had previously protected it as it developed, then begins to elongate while the germinal plate begins to expand dorsally from the ventral midline. During stage 10, fusion of the germinal plate along the dorsal midline is completed and organogenesis continues. After stage 11, embryonic organ systems have completely differentiated and the embryo is able to grasp substrate with its two suckers and eat its first meal.

Pax genes are expressed during the morphogenesis of a variety of organs, including those of the vertebrate excretory and immune systems (Dahl et al. 1997). In addition to their

diverse roles in developmental patterning and cell-type specification, they are also involved in the development of the nervous system (Stoykova and Gruss 1994). To date, Group II Pax genes have been examined at length in both Deuterostomes and Ecdysozoans, including Xenopus, Drospohila, and Mouse (Heller and Brandli 1999; Mansouri et al. 1999; Dressler and Woolfe 1999). However, significantly less expression analysis of Pax-2/5/8 has been done in Lophotrochozoa (O'Brien and Degnan 2000; O'Brien and Degnan 2003).

Here we provide a preliminary molecular and expression analysis of Pax-2/5/8 in *Helobdella* sp. (Austin). We place specific emphasis on elucidating the genetic structure of *Hau-Pax2/5/8* and to what extent it is expressed during the developmental program of the leech embryo.

MATERIALS AND METHODS

Leech Embryos

Embryos were obtained from a laboratory based breeding colony of *Helobdella* sp. (Austin), established with individuals collected from Shoal Creek, Austin, TX, USA. This colony was originally thought to be *Helobdella robusta*, but molecular phylogenetic analysis of mitochondrial DNA (Bely and Weisblat 2006) identified them as a distinct and yet unnamed species, *Helobdella* sp. (Austin). Adults were maintained at room temperature in 1% artificial seawater and fed three times a week on pond snails collected at the University of Texas Brackenridge Field Laboratory. Embryos were raised at 24°C in a defined saline solution and staged according to Stent et al. (1992).

Molecular Characterization of *Hau-Pax2/5/8*

Ian Quigley initially obtained a fragment of the Hau-Pax2/5/8 transcript, from leech embryonic cDNA, by designing primers for the predicted Pax-2/5/8 Paired domain within the Helobdella robusta genome sequence, provided by the Joint Genome Institute (JGI). The remainder of the transcript was obtained by Rapid Amplification of cDNA Ends (RACE) using both the Ambion FirstChoice RLM-RACE kit and Invitrogen GeneRacer kit (List of primers used is included in Appendix 1). Gene-specific PCR was used to clone and verify the complete cDNA transcript. PCR products were cloned using Invitrogen's TOPO TA Cloning Kit (pCR II-TOPO Vector). Sequencing was performed at the UT Core DNA Sequencing Facility. Multiple overlapping 3' and 5' RACE products were aligned and used to predict the full-length sequence of *Hau-Pax2/5/8*. The inferred *Hro-Pax2/5/8* sequence was constructed by BLAST searching for the *Hau-Pax-2/5/8* sequence within the *Helobdella robusta* genome (JGI) and aligning the exon fragments that corresponded with the *Hau-Pax2/5/8* open reading frame. Phylogenetic analysis was done using PUAP software Version 4.0b10. The constructed Phylogram is the consensus of 1000 replicates, with bootstrap values given as percentages, and collapse of branches with < 50% support.

Whole-mount *In Situ* Hybridization

Whole-mount in situ hybridization was performed on leech embryos in a similar manner as described by Nardelli-Haefliger and Shankland (1992): After being harvested, leech embryos younger than stage 9 were fixed for 1 hour in a heptane-enriched formaldehyde solution and embryos stage 9 and older for 30 minutes, in 4% formaldehyde solution. They were then washed and stored in methanol for over 24hrs at -20°C. After being rehydrated in PBS [130 mM NaCl + 10 mM phosphate buffer, pH 7.4], embryos stage 9 and younger, were devitellinized and treated for 2-

2.5 minutes in 0.5 mg/ml Pronase E (Sigma) in 50 mM Tris and 5 mM EDTA (pH 8.0). Older embryos had already hatched from the vitelline membrane and were treated for 18-20 minutes with Pronase. Treatment was stopped with 2 mg/ml glycine in PBS and the embryos were transferred to 2 ml of 0.1 M tri-ethanolamine buffer (pH 8) and acetylated by two sequential 5 minutes treatments of 4.5 µl acetic anhydride. Embryos were then returned to PBS, and post-fixed for 30 minutes with 4% formaldehyde in PBT [PBS + 0.1% Tween-20], followed by washing in PBT and storage in 2X SSC + 0.3% CHAPS (Sigma).

Embryos were then washed in fresh hybridization solution [50% de-ionized formamide, 5× SSC, 0.2 µg/ml tRNA, 1× Denhart's solution, 0.1 µg/ml heparin, 0.1% Tween-20, and 0.1% CHAPS (Sigma)] for six hours at 60°C. Following this blocking step, embryos were transferred to the hybridization solution containing a full-length digoxigenin-labeled (Roche Applied Science) antisense riboprobe, generated from a 1475 bp cDNA fragment of *Hau-Pax2/5/8*, synthesized using T7 RNA polymerase. Optimal staining was obtained after a trial series of probe dilutions. Labeled embryos were then washed at 60°C for five hours with fresh hybridization solution followed by 2× SSC with 0.3% CHAPS.

Washed embryos were transferred to room temperature PBT, and bound probe was visualized using immunochemistry, as follows. Nonspecific binding was reduced using a 1-hour incubation with 10% normal sheep serum. Embryos were then washed and incubated overnight at 4°C in a 1:5000 dilution of alkaline phosphatase (AP)-conjugated anti-digoxigenin Fab fragments (Roche Applied Science). Unbound antibody was removed with several washes, totaling 5 hours, in room temperature PBT. Embryos were then transferred to AP buffer [100 mM NaCl + 100 mM Tris + 50 mM MgCl2, pH 9.5] with 0.1% Tween-20 for five minutes which was then replaced with the developing solution [3 ml filtered AP buffer + 15 µl of 100 mg/ml NBT (4-nitro blue tetrazolium

chloride) and 10 μl of 50 mg/ml BCIP (5-bromo-4-chloro-3-indolyl-phosphate)]. Reactions were incubated in the dark at either room temperature or 4°C and checked routinely as color developed. Once staining was sufficiently dark, the reaction was terminated with PBT. Stained embryos were viewed on a Nikon E800 microscope and imaged with a Diagnostic Instruments Spot Flex 1520 CCD camera. Images were processed w/ Adobe Photoshop (CS3 Version 10.0).

Reverse Transcriptase PCR (RT-PCR)

RNA was extracted from several stages of leech embryos using the RNAqueous kit (Ambion), and reverse- transcribed into cDNA using the Invitrogen GeneRacer kit. RT-PCR was then carried out on the staged embryonic leech cDNA. Each reaction contained gene-specific primers and control rRNA primers (QuantumRNA 18S Internal Standards; Ambion), and was optimized according to kit protocols. The primers used (Pax 258-24 and Pax 258-26 at 58°C for 35 cycles) were designed so that the amplicon spanned a single intron.

RESULTS

Molecular characterization of *Hau-Pax2/5/8*

The *Hau-Pax2/5/8* gene product isolated from 5' and 3' RACE of late stage *Helobdella* sp. (Austin) cDNA is a 2370 bp sequence containing a 5' UTR, a 3' UTR, and a 1647 bp open reading frame (ORF) (Fig. 3). Confirmation of true 3' and 5' ends was established after consistent replicate cloning and subsequent sequencing. There are 4 possible in-frame ATG start codons between the Paired domain and the nearest stop codon upstream of the Paired

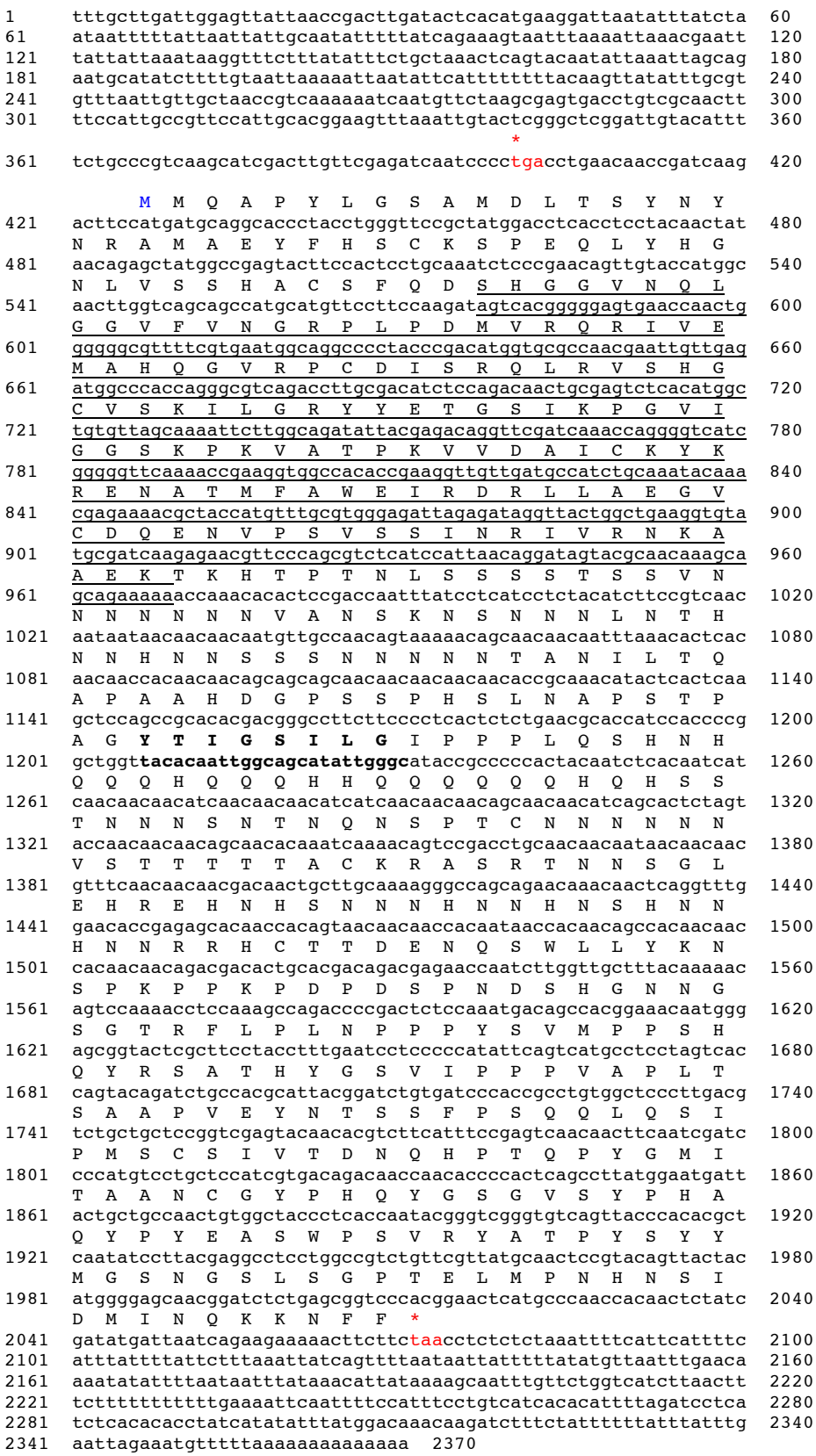

Fig. 3. Nucleotide and deduced amino acid sequence of *Hau-Pax2/5/8*. The Paired domain region is denoted with a single underline. The Octapeptide region is denoted by bolded text. The start methionine is indicated by a blue M. Stop codons of the termination site are indicated by a red asterisk..

domain. The most 5' of these start sites predicts the ORF of 1647 bp. A BL2SEQ (Altschul et al. 1997) revealed a 99% similarity between the sequenced *Helobdella* sp. (Austin) Pax-2/5/8 cDNA sequence and the *Helobdella robusta* genome sequence (JGI).

The *Hau-Pax2/5/8* Paired domain region is 444 bp long and located 150 bp from the proposed start codon. BL2SEQ comparison (Altschul et al. 1997) of the *Hau-Pax2/5/8* Paired domain against the Paired domains of other Group II Pax genes also revealed significant sequence similarity (Fig. 4): 100% similarity with *Hro-Pax2/5/8*, 94% similarity with *Pdu-Pax2/5/8*, 89% similarity with *Dme-shaven*, and 87% similarity with both *Hsa-Pax2* and *Has-Pax2/5/8*. *Hro-Pax2/5/8*, *Pdu-Pax2/5/8*, *Has-Pax2/5/8* are all Pax 2/5/8 genes from leech and other Lophotrochozoans. *Dme-shaven* is the Drosophila, an Ecdysozoan, equivalent of Pax2 and *Hsa-Pax2* is the Human, a Deuterostome, Pax2 gene.

```
Hau-Pax2/5/8        SHGGVNQLGGVFVNGRPLPDMVRQRIVEMAHQGVRPCDISRQLRVSHGCV
Hro-Pax2/5/8        SHGGVNQLGGVFVNGRPLPDMVRQRIVEMAHQGVRPCDISRQLRVSHGCV
Pdu-Pax2/5/8        -HGGVNQLGGVFVNGRPLPDVVRTRIVELAHQGVRPCDISRQLRVSHGCV
Dme-Shaven          CHGGVNQLGGVFVNGRPLPDVVRQRIVELAHNGVRPCDISRQLRVSHGCV
Hsa-Pax2            GHGGVNQLGGVFVNGRPLPDVVRQRIVELAHQGVRPCDISRQLRVSHGCV
Has-Pax2/5/8        ---------------RPLPDVVRTRIVDLAHQGVRPCDISRQLRVSHGCV

Hau-Pax2/5/8        SKILGRYYETGSIKPGVIGGSKPKVATPKVVDAICKYKRENATMFAWEIR
Hro-Pax2/5/8        SKILGRYYETGSIKPGVIGGSKPKVATPKVVDAICKYKRENATMFAWEIR
Pdu-Pax2/5/8        SKILGRYYETGSVRPGVIGGSKPKVATPKVVGAICKYKRENPTMFAWEIR
Dme-Shaven          SKILSRYYETGSFKAGVIGGSKPKVATPPVVDAIANYKRENPTMFAWEIR
Hsa-Pax2            SKILGRYYETGSIKPGVIGGSKPKVATPKVVDKIAEYKRQNPTMFAWEIR
Has-Pax2/5/8        SKILGRYYETGSIKPGVIGGSKPKVATPKVVDAITRYKQDNPTI------


Hau-Pax2/5/8        DRLLAEGVCDQENVPSVSSINRIVRNKAAEK
Hro-Pax2/5/8        DRLLAEGVCDQENVPSVSSINRIVRNKAAEK
Pdu-Pax2/5/8        DRLLAEGVCDQENVPSVSSINRIVRNKAAEK
Dme-Shaven          DRLLAEAICSQDNVPSVSSINRIVRNKAAEK
Hsa-Pax2            DRLLAEGICDNDTVPSVSSINRIIRTKVQQP
Has-Pax2/5/8        --------------------------------
```

Fig. 4. Conserved Paired domain amino acid sequence encoded by various Group 2 Pax genes. Highlighted amino acids are not conserved across all species. [Hau-Pax2/5/8: *Helobdella* sp. (Austin), sequence determined experimentally; Hro-Pax2/5/8: *Helobdella robusta*, sequence from JGI BLAST Alignment Search; Pdu-Pax2/5/8: *Platynereis dumerilii*, AJ505023; Dme-Shaven: *Drosophila Melanogaster*, AF016888; Has-Pax2: *Homo sapiens*, NM000278; Has-Pax2/5/8: *Haliotis asinina*, AAK26112]

The *Hau-Pax2/5/8 gene,* like many Pax genes, also contains an octapeptide region, downstream of the Paired domain. Further BL2SEQ comparison (Altschul et al. 1997) of the *Hau-Pax2/5/8* octapeptide region against the octapeptide region of other Group II Pax genes also revealed significant sequence similarity (Fig. 5): 100% similarity with *Hro-Pax2/5/8*, 75% similarity with *Pdu-Pax2/5/8*, and 62% with both *Dme-shaven*, and *Hsa-Pax2*.

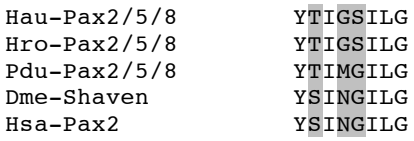

Fig. 5. Conserved octapeptide amino acid sequence encoded by various Group 2 Pax genes. Highlighted amino acids are not conserved across all species. [Hau-Pax2/5/8: *Helobdella* sp. (Austin), sequence determined experimentally; Hro-Pax2/5/8: *Helobdella robusta*, sequence from JGI BLAST Alignment Search; Pdu-Pax2/5/8: *Platynereis dumerilii*, AJ505023; Dme-Shaven: *Drosophila Melanogaster*, AF016888; Has-Pax2: *Homo sapiens*, NM000278]

Many Pax genes also contain a partial, if not complete, N-terminal homeodomain sequence downstream of both the Paired domain and the octapeptide region. BL2SEQ comparison (Altschul et al. 1997) of *Hau-Pax2/5/8*, against the homeodomains of known Group II Pax genes (Fu and Noll 1997), failed to identify any remnant of an N-terminal homeodomain, downstream of the Paired domain and octapeptide regions.

A phylogenetic analysis of the Pax 2/5/8 Paired domain, based on parsimony, was constructed (Fig. 6) to determine if *Hau-Pax2/5/8* is indeed a Pax2/5/8 gene. Upon comparison with members of the five Pax gene groups (Pax1/9, -2/5/8, -3/7, -4/6, and Pox neuro), as well as Paxβ it can be seen that, with a bootstrapping value of 72, *Hau-Pax2/5/8* in fact does cluster with other known group II Pax genes.

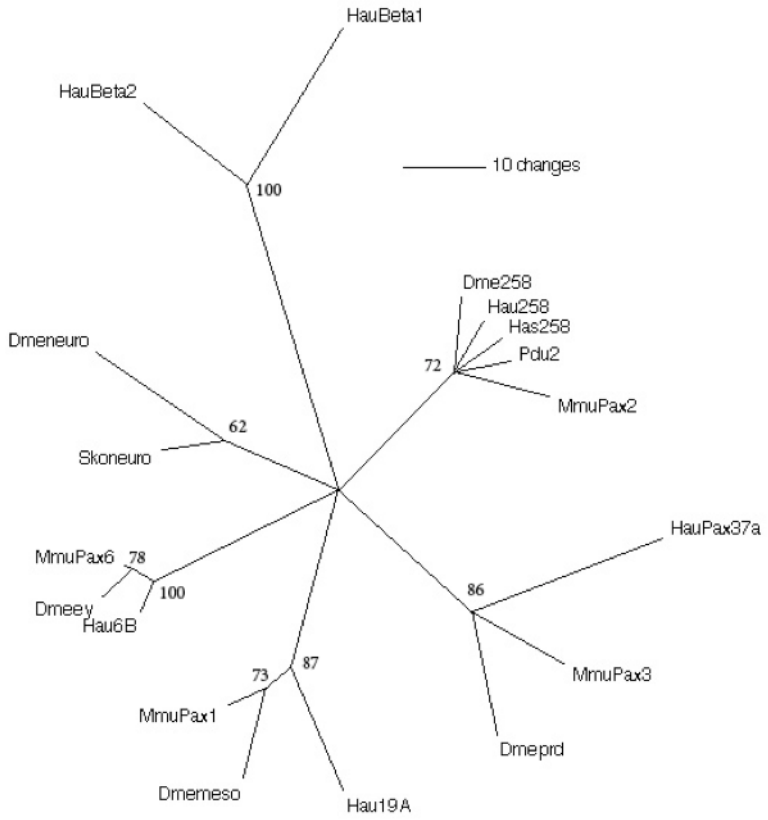

Fig. 6. Phylogenetic analysis of Pax 2/5/8 Paired domain based on parsimony, compared with members of the five Pax gene groups (Pax1/9, -2/5/8, -3/7, -4/6, and Pox neuro), as well as Paxβ. The phylogram is the consensus of 1000 replicates, with bootstrap values given as percentages. Branches with < 50% support were collapsed. *Hau, Helobdella* sp. (Austin); *Dme, Drosophila melanogaster; Sko, Saccoglossus koalevskii; Mmu, Mus musculus; Pdu, Playnereis dumerilii*.

Expression of *Hau-Pax2/5/8*

The expression of *Hau-Pax2/5/8* during *Helobdella* sp. Austin embryonic development was characterized using both in-situ hybridization and RT-PCR at multiple stages of development. From preliminary RT-PCR experiments it appears that *Hau-Pax2/5/8* transcripts are present in the *Helobdella* zygote (presumably maternally deposited), and expression is maintained throughout embryonic development (Fig. 7).

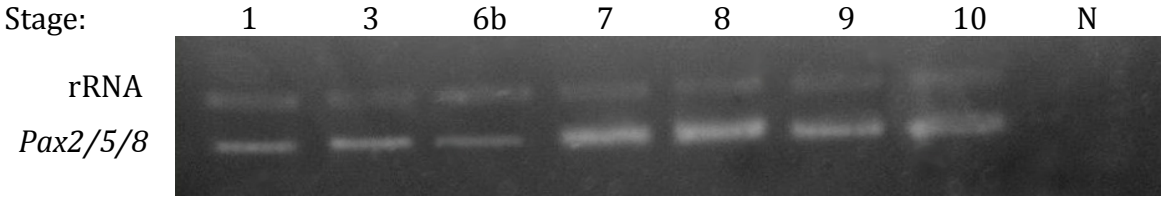

Fig. 7. Developmental reverse transcriptase – polymerase chain reaction of *Hau-Pax2/5/8*. *Pax2/5/8* is expressed throughout embryogenesis. In earlier stages expression is most likely a result of the presence of maternal transcripts while later stages reflect zygotic expression. 18S rRNA was co-amplified as a loading control. Lane N is a no template negative control.

It can be speculated that *Hau-Pax2/5/8* transcript present throughout stages 1-6, however, beginning in stage 7 the relative amount of RNA present in the developing leech embryo increases quite substantially. Expression then appears to taper off only slightly, with RNA levels staying relatively high up through stage 10, the last stage of development examined. The expression patterns seen here appear to be relatively consistent with those visualized using whole-mount in situ hybridization.

After conducting whole-mount in situ hybridization it can be seen (Fig. 8) that the *Hau-Pax2/5/8* transcript is present even during the first cleavage of the leech zygote. As a result of the first asymmetric cleavage event along the animal-vegetal axis, the larger of the two cells, CD, contains the two pools of teloplasm. *Hau-Pax2/5/8* transcript can be seen in both the animal and vegetal pools of teloplasm, of the stage 2 embryo, as well as in the perinuclear cytoplasm of both the AB and CD cells.

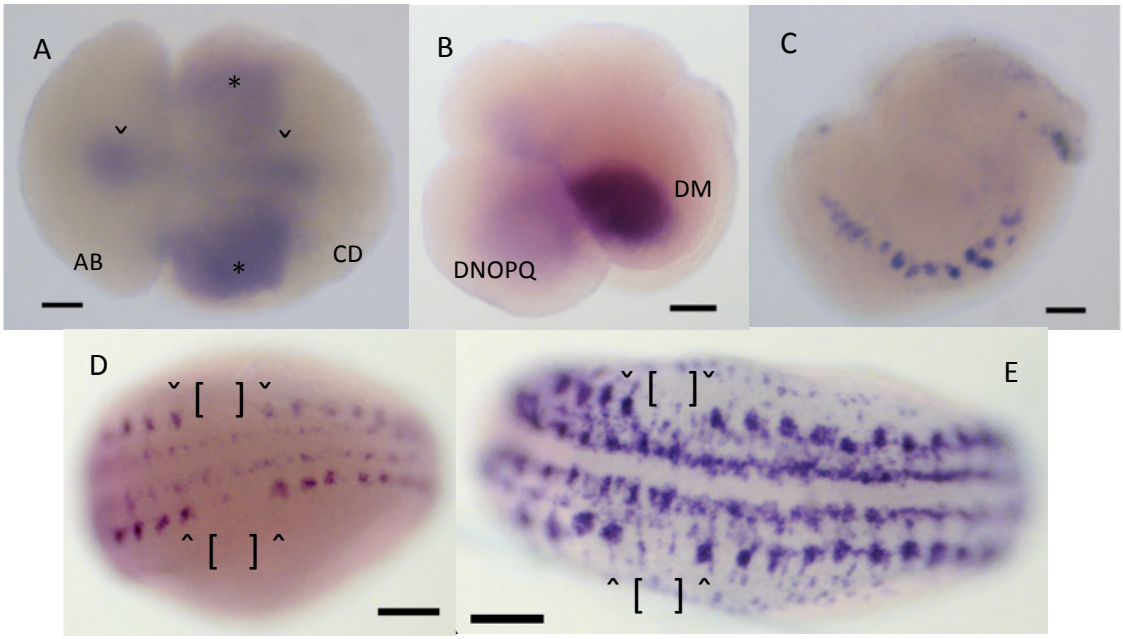

Fig. 8. Whole-mount in-situ hybridization of *Hau-Pax2/5/8* expression. (A) Side view of Stage 2 *Hau-Pax2/5/8* RNA in two pools of teloplasm (asterisks) and perinuclear cytoplasm (arrowheads). (B) Ventral view of Stage 4, "5-cell" *Hau-Pax2/5/8* RNA in both DM and DNOPQ cells, strongest in DM. (C) Side view of Stage 8 *Hau-Pax2/5/8* expression in the germinal bands. (D and E) Ventral view of developmental progression of *Hau-Pax2/5/8* expression in the germinal plate of Stage 9 embryos. Segments with under developed nephridia are brackets, with neighboring nephridia designated by arrowheads. Scale bars: 100μm.

As the embryo continues to divide the majority of the *Hau-Pax2/5/8* transcript continues to segregate with the teloplasm. During the fourth round of cell divisions as the D macromere divides into cells DNOPQ and DM, *Hau-Pax2/5/8* accumulation is greatest in DM and to a lesser extent in DNOPQ. Then, during stages 7 and 8, as the germinal bands are forming and migrating around the embryo, *Hau-Pax2/5/8* expression persists in both left and right germinal bands. During stage 9, when the embryo hatches from the vitelline membrane and begins to elongate, the germinal plate begins to expand dorsally and *Hau-Pax2/5/8* expression is seen along the ventral midline in what will become the nephridia, mesodermally derived excretory organs. As fusion of the germinal plate along the dorsal midline is completed and organogenesis continues, the nephridia continue to develop during stage 10. The nephridia develop in 15 of the 21 mid-body segments (M2-5, M8-18). The gap in nephridia-developing segments, corresponding with the two mid-body segments that contain the leech's reproductive organs, is located one-third of the way from the anterior end of the embryo and exhibits noticeably less *Hau-Pax2/5/8* expression. This most likely points to *Hau-Pax2/5/8* as playing a role in the development of the leech excretory system. Cross-sections of Stage 10 in situ hybridized embryos (Fig. 9) confirm *Hau-Pax2/5/8* expression in the nephridia and also indicate expression in the lateral regions of the developing central nervous system (CNS), as well as at the lateral-most portions of the germinal plate.

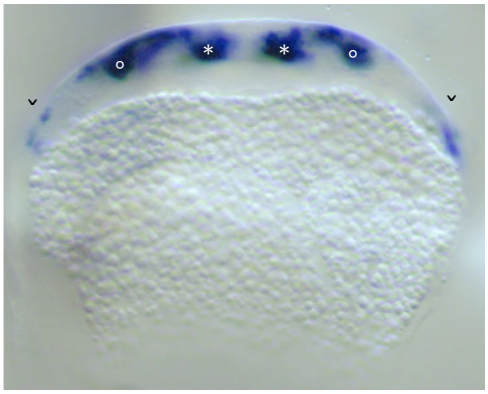

Fig. 9. Whole-mount in-situ hybridization of *Hau-Pax2/5/8* expression. Cross section of Stage 10 *H.* sp. (Austin). Developing nephridia denoted by circles. Central nervous system expression denoted by asterisks. Lateral germinal plate expression

DISCUSSION

Pax genes have been coordinating the embryonic developmental patterns of multicellular organisms over hundreds of millions of years. Ever since they diverged from their fellow Pax genes, Group II Pax genes have been working specifically to regulate tissue-specific transcription in the nervous, excretory, and immune systems (Chi and Epstein 2007). *Hau-Pax2/5/8* is the first Group II Pax gene to have been described in the leech *Helobdella* sp. (Austin), and the third in the superphylum Lophotrochozoa (O'Brien and Degnan 2000; O'Brien and Degnan 2003; Denes et al. 2007).

We believe that the sequence presented in Figure 3 represents the true Pax-2/5/8 sequence of the annelid leech. As expected, Pax-2/5/8 transcript from *Helobdella* sp. (Austin) contains both the Paired domain characteristic of all Pax genes as well as the octapeptide region chracteristic of previously identified Group II Pax genes (Chi and Epstein 2002). The Paired domain identified shares 94% similarity in the amino acid sequence as *Platynereis dumerilii*, a polychaetous annelid, and 87% with the Pax2 gene of *Homo sapiens*. Most importantly, as seen in Figure 6, a parsimonious phylogenetic comparison of known Pax genes groups *Hau-Pax2/5/8* with other known group II Pax genes and yields a bootstrapping value of 72. Thus, making a strong case for its identity as the annelid leech equivalent of a Pax-2/5/8 class gene.

The *Hau-Pax2/5/8* octapeptide sequence, however, is noticeably deviant from the conserved octapeptide sequence deduced by Ziman and Kay (1998), which contains a TN_8TCCT motif in the nucleotide sequence. The corresponding octapeptide-encoding sequence of *Hau-Pax2/5/8* is TN_8TATT. While there is a distinct deviation from Ziman's and Kay's consensus

sequence it does maintain its directional nature (the single T nucleotide located upstream of the TCCT sequence), thought to direct methylation to upstream regions of the gene. The *Hau-Pax2/5/8* octapeptide amino acid sequence, YTIGSILG, is also notably absent of at least one of the functionally significant amino acids, Serine at position 2 (Ziman and Kay 1998). However, the *Hau-Pax2/5/8* is noticeably absent the partial homeodomain motif, present in other species (Chi and Epstein 2002). Because of this, the function of the octapeptide region, in the proposed process of directed hypomethylation of the Paired domain and hypermethylation of the homeodomain region (Ziman and Kay 1998), may not have been as strictly maintained in the leech over evolutionary time, allowing the sequence to diverge from that of other species.

In previous studies Group II Pax genes have been shown to be predominantly involved in the development of the nervous and sensory systems of both Deuterostomes and Ecdysozoans (Stoykova and Gruss 1994; Mansouri et al. 1999) as well as an playing an important role in the development in the excretory system of vertebrates (Dahl et al. 1997). The preliminary expression analysis from this study suggests that *Hau-Pax2/5/8* does play a role in the development of the excretory and nervous systems in the developing leech embryo, though its involvement in segmentation of the leech body plan has not been conclusively determined. This can be seen in the largely mesodermal expression of *Hau-Pax2/5/8* throughout embryogenesis and especially in the development of the nephridia and portions of the CNS in stages 9 and 10. Further analysis is required to determine to what extent Pax-2/5/8 plays in the development of the leech nervous and excretory systems, including the specific phenotypic effects that result from disrupted Pax-2/5/8 expression. However, this study provides an excellent foundation for

future analysis of the functional role of the Pax-2/5/8 gene in the developmental patterning and organogenesis of the annelid leech.

ACKNOWLEDGMENTS

The authors wish to thank the members of the Shankland and Wallingford labs for their advice and assistance on this project, as well as Ian Quigley for cloning the first *Hau-Pax2/5/8* fragment. Work was funded by a grant from the National Science Foundation awarded to Marty Shankland.

REFERENCES

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res*. 25: 3389-3402.

Balczarek, K. A., Lai, Z., and Kumar, S. 1997. Evolution and functional diversification of the paired box (Pax) DNA-binding domains. *Mol. Biol. Evol*. 14: 829-849.

Bely, A. E., and Weisblat, D. A. 2006. Lessons from leeches: a call for DNA barcoding in the lab. *Evol. Dev.* 8: 491-501.

Chi, N., Epstein, J. A., 2002. Getting your Pax straight: Pax proteins in development and disease. *Trends Genet.* 18: 41-47.

Dahl, E., Koseki, H., Balling, R. 1997. Pax genes and organogenesis. *Bioessays* 19: 755-765.

Denes, A. S., Jekely G., Steinmetz, P. R. H., Raible, F., Snyman, H., Prud'homme B., Ferrier, D. E. K., Balavoine, G, and Arendt, D. 2007. Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilaterian. *Cell*. 129: 277-288.

Dressler, G. R. and Woolf, A. S. 1999. Pax2 in development and renal disease. *Int. J. Dev. Bio.* 43: 463-468.

Fu, W. and Noll, M. 1997. The Pax2 homolog sparkling is required for development of cone and pigment cells in the Drosophila eye. *Genes and Dev.* 11: 2066-2078.

Heller, N. and Brandli, A. W. 1999. Xenopus Pax-2/5/8 Orthologues: Novel Insights Into Pax Gene Evolution and Identification of Pax-8 as the Earliest Marker for Otic and Pronephric Cell Lineages. *Dev. Gen.* 24: 208–219.

Lang, D., Powell, S. K., Plummer, R. S., Young, K. p., and Ruggeri, B. A. 2007 PAX genes: roles in development, pathophysiology, and cancer. *Biochem. Pharmacol.* 73: 1-14.

Mansouri, A., Goudreau, G., and Gruss, P. 1999. Pax genes and their role in organogenesis. *Cancer Research* 59: 1707-1710

Matus, D. Q., Pang, K., Daly, M., and Martindale, M. Q. 2007 Expression of Pax gene family members in the anthozoan cnidarian, *Nematostella vectensis. Evol. Dev. 9: 25-38.*

Nardelli-Haefliger, D., and Shankland, M. 1992. *Lox2*, a putative leech segment identity gene, is expressed in the same segemental domain in different stem cell lineages. *Development* 116: 697-710.

O'Brien, E. K., and Degnan, B. M., 2000. Expression of POU, Sox, and Pax genes in the brain ganglia of the tropical abalone *Haliotis asinine*. *Mar. Biotechnol.* 2: 545-557.

O'Brien, E. K., and Degnan, B. M., 2003. Expression of Pax2/5/8 in the gastropod statocyst: insights into the antiquity of metazoan geosensory organs. *Evol. Dev.* 5: 572-578.

Shankland, M. 1991. Leech segmentation: cell lineage and the formation of complex body patterns. *Dev. Biol.* 144: 221-331.

Shankland, M. and Savage, R. M. 1997. "Annelids, the Segmented Worms." *Embryology, Constructing the Organism*, Sinauer Associates, Sunderland, MA, p. 219-235.

Stent, G. S., Kristan, W. B. Jr., Torrence, S. A., French, K. A., and Weisblat, D. A. 1992. Development of the leech nervous system. *Int. Rev. Neurobiol.* 33: 109-193.

Stoykova, A., Gruss, P. 1994. Roles of Pax genes in developing and adult brain as suggested by expression patterns. J. Neurosci. 14: 1395-1412.

Sun, H., Rodin, A., Zhou, Y., Dickinson, D. P., Harper, D. E., Hewett-Emmett, D., and Li, W. 1997. Evolution of paired domains: Isolation and sequencing of jellyfish and hydra Pax genes related to Pax-5 and Pax-6. Proc. Natl. Acad. Sci. USA 94: 5156-5161.

Wada, H., Saiga, H., Satoh, N., Holland, P. W. H. 1998. Tripartite organization of the ancestral chordate brain and the antiquity of placodes: insights from ascidian Pax-2/5/8, Hox and Otx genes. Development 125: 1113-1122.

Walther, C., Guenet, J. L., Simon, D., Deutsch, U., Jostes, B., Goulding, M. D., Plachov, D., Balling, R., and Gruss, P. 1991. Pax: a murine multigene family of paired box-containing genes. Genomics 11: 424-434.

Weisblat, D.A., Shankland, M., 1985. Cell lineage and segmentation in the leech. Philos. Trans. R. Soc. Lond., B 312, 39–56.

Ziman, M. R., Kay, P. H. 1998. A conserved TN_8TCCT motif in the octapeptide region of *Pax* genes which has the potential to direct cytosine methylation. Gene 223: 303-308.

APPENDIX I: PRIMER SEQUENCES (Listed in the form: 5' to 3')

Name	*Sequence*	*Tm*
Pax 258-1	GGGCGTTTTCGTGAATGGCAG	59.6°C
Pax 258-2	GGCCCCTACCCGACATGGT	62.2°C
Pax 258-3	CGCAGTTGTCTGGAGATGTCGCA	61.5°C
Pax 258-4	CCAAGAATTTTGCTAACACAGCCATG	58.3°C
Pax 258-5	TTGGCGCACCATGTCGGGTAGG	64.2°C
Pax 258-6	GACATCTCCAGACAACTGC	53.1°C
Pax 258-7	TTCTCGTTTGTATTTGCAGATGGC	56.1°C
Pax 258-8	CCACGCAAACATGGTAGC	54.4°C
Pax 258-11	GCAAATACAAACGAGAAAACGCTACC	57.0°C
Pax 258-12	GCGTGGGAGATTAGAGATAGG	54.2°C
Pax 258-13	CGCAACAAAGCAGCAGAAAAAACC	58.3°C
Pax 258-14	CCGCAAACATACTCACTCAAGC	56.4°C
Pax 258-15	GCCTTCTTCCCCTCACTC	55.0°C
Pax 258-16	GCCCCCACTACAATCTCAC	55.4°C
Pax 258-17	GCATAACGAACAGACGGC	54.1°C
Pax 258-18	GCGTGTGGGTAACTGAC	53.3°C
Pax 258-19	CCGTATTGGGTGAGGGTAGC	55.1°C
Pax 258-20	CCCAACCACAACTCTATCG	52.8°C
Pax 258-21	GCCGTCTGTTCGTTATGC	54.1°C
Pax 258-22	CTACCCTCACCAATACGG	51.8°C
Pax 258-23	GGGGTCATCGGGGGTTC	58.3°C
Pax 258-24	TTGCGTGGGAGATTAGAGATAGG	55.9°C
Pax 258-25	TGCTGCTGCTGTTGTTGTGG	59.0°C
Pax 258-26	GAGAGTGAGGGGAAGAAGGC	57.2°C
Pax 258-27	TGGCATCAACAACCTTCGG	55.8°C
Pax 258-28	TTTTGAACCCCCGATGACC	55.3°C
Pax 258-29	CCCGATGACCCCTGG	55.2°C
Pax 258-30	GCCGTCTGTTCGTTATGC	54.1°C
Pax 258-31	GTCAGTTACCCACACGCTC	55.7°C
Pax 258-32	ATAGGTGCGTGAGATGAGG	53.9°C