



THE EXPANSION AND DIVERSIFICATION OF THE CLAUDIN GENE FAMILY: INSIGHT FROM THE LAMPREY

Christian Kabasele Mukendi

A Dissertation submitted to the Faculty of Science, University of the Witwatersrand,
Johannesburg, in fulfilment of the requirements for the degree of Master of Science

May 2015 in Johannesburg

DECLARATION

I declare that this Dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

A handwritten signature in black ink, appearing to read 'M. M. M. M.', with a long horizontal line extending to the right.

22nd day of May 2015 in Wits GH 708

ABSTRACT

Claudins are a large gene family found in all vertebrates. Claudins encode tetraspan membrane proteins, involved in the structure and function of the tight junctions. This association of cells leads to the formation of the epithelial sheet which is involved in many functions such as embryo morphogenesis. The NCBI database shows 27 claudins identified in humans; 23 in mice and 17 in *Xenopus*. This suggests that an increase in gene family size may correlate with the evolution of more complex vertebrates. In this study claudins from the most basal extant vertebrate, the sea lamprey, were investigated. RNA used to build up the lamprey genome by Jeramiah Smith (Smith et al., 2012), was used for lamprey claudin sequences. Additionally this study identified 2 more claudins (Cldn B & Cldn F). The phylogenetic tree constructed using claudins from higher vertebrate model organisms and the invertebrates *Ciona intestinalis* and *Drosophila melanogaster*; showed that lamprey claudins are evolutionarily more distantly related to their orthologs in higher vertebrates. Furthermore some claudins in lamprey did not show any homologs in higher vertebrates and vice versa, indicating the emergence of novel members in higher vertebrates. However lamprey Cldn A was found to be homologous to CLDN 3 in higher vertebrates. This is interesting since CLDN 3 is involved in the development of two vertebrate specific traits; one of which is the ear placode. Thus Cldn A (renamed Cldn 3B), was made a focus of this study. RNA *in situ* hybridization using probes designed from individual UTRs showed localised expression of Cldn 3B in the ear placode, pharyngeal pouch, pericardial cavity and the fin fold whereas Cldn B (renamed Cldn 8B) was mostly expressed in the pharyngeal pouch and ear placode much like its orthologs in higher vertebrates. Knockout experiments showed that Cldn 3B is involved in sealing and expansion of the ear placode and pharyngeal arches during development whereas Cldn 8B is involved in determining ear placode development. Thus claudins are seen to be heavily involved in the morphology of vertebrate specific traits therefore an expansion in this gene family would affect the complexity of vertebrates during evolution.

LIST OF RESEARCH OUTPUTS

Presentation		Conference	
Poster Presentation		Molecular Biosciences Research Thrust Research Day 2013	
Oral Presentation		South African Society of Biochemistry and Molecular Biology 2014	
Paper		Status	
C. K. Mukendi, N. Dean, R. Lala, S. Ntila, J. Smith, M. E. Bronner and N. V. Nikitina. Evolution of the Vertebrate Claudin Gene Family: Insights from the most basal vertebrate the sea lamprey		Manuscript in preparation	

ACKNOWLEDGEMENTS

- Dr. Natalya Nikitina for her guidance and support throughout my MSc
- The MCB staff for their help in the usage of communal equipment
- NRF for the funding, especially for the trip to Pasadena, California, USA to work with the live lamprey
- The Bronner Laboratory at the California Institute of Technology, for allowing us to use their facilities to perform the knockout experiments, In that regards special thanks to Stephen Green and Hugo Parker for their help with usage of the equipment and guidance
- The Bone Research Laboratory, a research unit of the South African Medical Research Council and Wits University. Special thanks to Ruqayya “Rooki” Parak for her assistance with usage of the equipment.
- Scott for his assistance in the bioinformatics to construct the phylogenetic tree
- My colleague in the lab, Miss Thembikile Zwane for her support
- My family for their love and support

LIST OF FIGURES

Figure 1:	Diagram Showing the Epithelial Cells Attach to Each Other Via Their Lateral Membranes, at the Tight Junction Between the Exocytoplasmic Units of Membranes from Adjacent Cells	2
Figure 2:	A Visual Representation of the Transcellular Pathway and the Paracellular Pathway	3
Figure 3:	Diagram Showing the Structure of the Claudin Protein	5
Figure 4:	Different Types of Claudin Protein Interactions	7
Figure 5:	Map of the Plasmid pCMV-Sport-6 and its Multiple Cloning Site (MCS)	23
Figure 6:	Phylogenetic tree of all Claudins in Model Organism	38
Figure 7:	Claudin 3b Expression Pattern During Lamprey Embryogenesis	41
Figure 8:	Sections Through the Lamprey Embryos Showing Claudin 3b Expression	42
Figure 9:	Claudin 8B Expression During Lamprey Embryonic Development	44
Figure 10:	Sections Through the Lamprey Embryos Showing Claudin 8b Expression	45
Figure 11:	Knockdown of Cldn 3b Results In Abnormal Development of Ear Placodes and Pharyngeal Arches	49
Figure 12:	Sections Through Stage E16 Claudin 3B Knockout Morphants Post <i>in situ</i> Hybridization	50
Figure 13:	Bar Graph Showing the Percentage of Embryos With Abnormalities in the Ear Placode at Stage E16	50
Figure 14:	Bar graph showing the percentage of embryos with abnormalities in the pharyngeal arches at stage E16	51
Figure 15:	Whole Mount <i>in situ</i> Hybridization on Stage E8 Claudin 3b Knockout Morphants	51
Figure 16:	Pie Chart Showing the Ratio of Embryos Affected By the Knockout of Claudin 3B at Stage E8	52
Figure 17:	Stage E8 Embryos Injected with 0.5mM Claudin 3B Morpholino at the Two Cell Stage	52
Figure 18:	Pie chart showing the ratio of E8 embryos affected by 50% Cldn 3B Morpholino	53

Figure 19:	Comparison Between Wild Type Embryo and Claudin 3B Morphant at Stage E6	53
Figure 20:	Claudin 3B Knockout Morphants at stage E6	54
Figure 21:	Stage E8 Claudin 8B Knockout Morphants Injected at Two Cell Stage	56
Figure 22:	Pie Chart Showing Results From the Claudin 8B Knockouts on One side of the Embryos	57
Figure 23:	Whole Mount <i>in situ</i> Hybridization Using Tbx1 on Claudin 8b Control Morpholino Injected Embryos at Stage E8	57
Figure 24:	Whole Mount <i>in situ</i> Hybridization Using Tbx1 on Claudin 8b Knockout Morphants at Stage E10	58
Figure 25:	Pie Chart Showing the Ratio of Embryos at Stage E10 Affected by Cldn 8b knockout	58
Figure 26:	Expression profile of Cldn 3B and 8B in the pharyngeal arches of the lamprey	70

LIST OF TABLES

Table 1:	Expression of <i>Claudins</i> in Different Cell Types During Lung Organogenesis	10
Table 2:	Summary of Known Expression Profile of <i>Claudin 3</i> in Vertebrate Model Organisms	15
Table 3:	Summary of Known Expression Profile of <i>Claudin 8</i> in Vertebrate Model Organisms	18
Table 4:	Morpholino and Target Sequences used in the Knockout Experiments	21
Table 5:	Internal primers used to sequence the CLDN from the T7 and SP6 direction	27
Table 6:	Ligation Mix	30
Table 7:	Linearization reaction mix	30
Table 8:	Probe Mix Composition	31
Table 9:	Summary of claudin 3b and claudin 8b expression profile during lamprey embryogenesis	46
Table 10:	The Size of the Claudin Gene Family in Different Organisms, Including <i>P. marinus</i>	60
Table 11:	Comparison of <i>P. marinus</i> claudin 3 expression to that of higher vertebrates	63
Table 12:	Comparison of <i>P. marinus</i> claudin 8 expression to that of higher vertebrates	72

Contents

DECLARATION	i
ABSTRACT.....	ii
LIST OF RESEARCH OUTPUTS	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF FIGURES	v
LIST OF TABLES.....	vii
CHAPTER ONE – INTRODUCTION	1
1.1. Tight Junctions.....	1
1.2. Claudins	5
1.3. Claudins in Embryonic Epithelia	8
1.4. Expansion and Diversification of the Claudin Gene Family in Vertebrates	10
1.5. Claudin 3.....	14
1.6. Claudin 8.....	17
1.7. Hypothesis and Aims	19
CHAPTER TWO – MATERIALS AND METHODS.....	20
2.1. Culturing Lamprey Embryos	20
2.2. Morpholino Injection	20
2.3. Embryo Fixing and Processing	21
2.4. Cloning and sequencing.....	22
2.4.1. Bacterial Cultures.....	22
2.4.2. Glycerol Stocks	23
2.4.3. Transformation of <i>E.coli</i> JM109 and Screening	24
2.4.4. DNA Plasmid Preparation and Analysis	26
2.4.5. Electrophoresis.....	26
2.4.6. Sequencing & Sequence Analysis.....	26
2.5. Phylogenetic tree construction	28
2.5.1. Sequence Assembly	28
2.5.2. Multiple Alignment and Tree Construction	28
2.6. DIG-labelled RNA Probe preparation.....	28
2.6.1. Probe Templates.....	28
2.6.2. Probe Synthesis	30
2.7. RNA <i>in situ</i> hybridization	32

2.7.1.	Embryo Preparation	32
2.7.2.	<i>In situ</i> hybridization	32
2.8.	Sectioning	34
2.8.1.	Gelatin Embedding	34
2.8.2.	Slide Mounting.....	34
CHAPTER THREE – RESULTS		35
3.1.	Phylogenetic tree indicates that claudin homologs and orthologs can be found in basal and more complex vertebrates, and paralogs can also be seen in individual lineages	35
3.2.	Lamprey Claudin 3b is expressed predominantly in the embryonic ectoderm, the neural tube, ear placodes, pharyngeal arches and the pericardial cavity.....	39
3.3.	Claudin 8b Expression is seen in the embryonic ectoderm, the ear placode and the pharyngeal arches.....	43
3.4.	Claudin 3 is involved in the neural tube, ear placode and pharyngeal arch development.....	47
3.5.	Claudin 8 is important for ear placode development	55
CHAPTER FOUR – DISCUSSION AND CONCLUSION		59
4.1.	Insight from the phylogenetic tree of evolution of claudins in vertebrates.....	59
4.1.1.	Lamprey Claudin Gene Family	59
4.1.2.	Phylogenetic Tree	60
4.2.	Claudin 3B	62
4.2.1.	Expression.....	62
4.2.2.	Function	65
4.3.	Claudin 8B	68
4.3.1.	Expression.....	68
4.3.2.	Function	73
4.4.	Conclusion	59
REFERENCES		79
APPENDIX.....		86
6.1.	APPENDIX A	86
6.2.	APPENDIX B	91
6.3.	APPENDIX C	99

CHAPTER ONE – INTRODUCTION

1.1. Tight Junctions

Over the course of evolution vertebrates have become one of the most diverse subphyla of animals in the entire kingdom. This subphylum shows a high level of adaptations which allowed for the success of different animals in various environments. The contributing factor to this phenomenon is the evolution of novel characteristics while maintaining certain traits. This led to the variations in individual lineages and ultimately gave rise to novel types of vertebrates, divided into 7 classes.

Compartmentalization of the body, genomes and other developmental structures form the basis of variation in individual properties during vertebrate evolution. Compartmentalization is an important trait because it overcomes the limitation brought about by diffusions and assists the regulation of cellular environments. This controls the division of function within the different cell lines in the compartments and ultimately maintains the function of each compartment. The supply of nutrients and information over great distances is also regulated through compartmentalization.

In multicellular organisms, tissue barriers that restrict passage of liquids, ions and larger solutes are essential for the development (Shen et al., 2011). This emphasises the importance of the epithelium, which is one of the four basic types of animal tissue, which lines the cavities and surfaces of structures all through the body. It is required to form effective diffusion barriers between the different compartments of the body thus the structure of the epithelial sheet supports the function. Epithelial cells are tightly packed to connect via cellular junctions leading to the formation of the epithelial sheet. During morphogenesis, epithelial sheets reorganize themselves to drive the morphogenetic outcomes that shape organs and the overall organism. In this way epithelial morphogenesis plays an important role in embryogenesis.

Tight junctions (TJ) are cellular connections used by the epithelia to form the continuous associations between the cells (Amasheh et al., 2009). These junctions are located at the apical regions where the membranes of two adjacent cells join together as shown in Figure 1. TJs are found only in vertebrates, however a corresponding junction, known as the septate junction, can be found in invertebrates. These junctions are dynamic with a structural component which includes branching networks of connecting strands. Under an electron microscope, the TJ appear as a set of long, parallel, linear fibrils that unite the

cells and short fibrils that connect to the main parallel array (Gumbiner, 1993). Increasing the density of fibrils produces a tighter bond, which would decrease the permeability of the junction. This means that the complexity of the tight junction components is important to the structure of the epithelia.

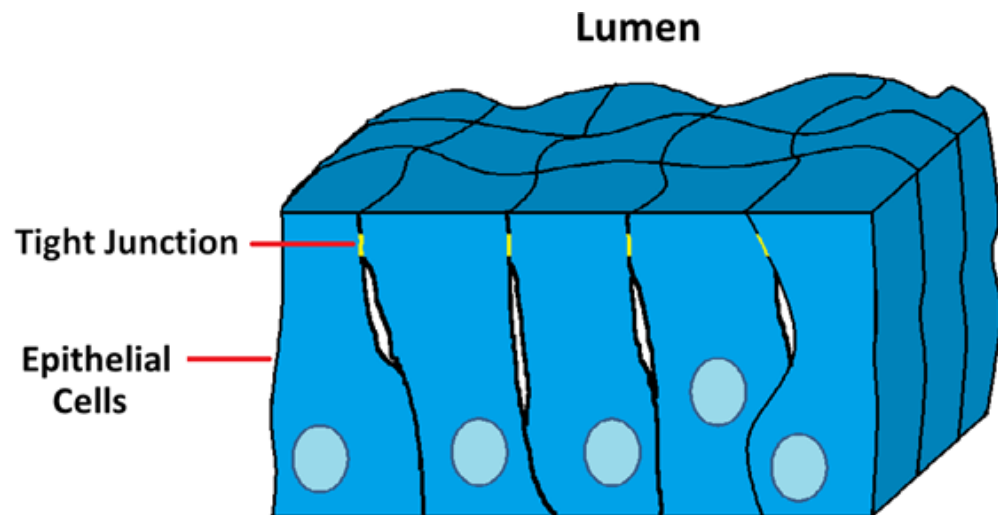


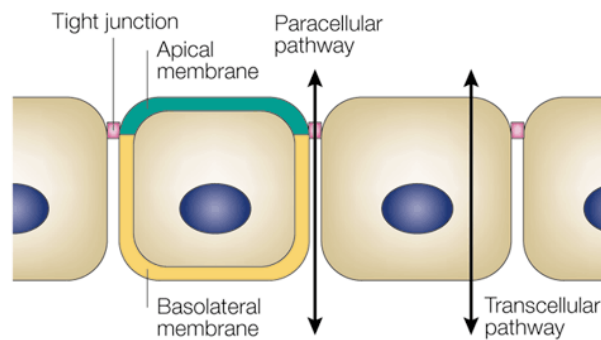
Figure 1: Diagram Showing the Epithelial Cells Attach to Each Other Via Their Lateral Membranes, at the Tight Junction Between the Exocytosmic Units of Membranes from Adjacent Cells.

TJs are multiprotein complexes, consisting of transmembrane proteins including occludins, junction adhesion proteins (JAMs) and the CLDNs (Hiroshi et al., 2012). TJ occurs between the exocytosmic units of membranes from adjacent cells (Günzel & Yu, 2013) where they are involved in many roles such as cell to cell association for sealing off the intracellular spaces and forming an intracellular signal transduction pathway (Cybulski & de Mendoza, 2011). The overall conformation of these components in the junction affects the efficiency of the TJ in sealing off the intracellular space.

TJ are crucial to the development and function of most organ systems because they enable the epithelia, as well as the endothelia, to create compositionally distinct fluid compartments (Shen et al., 2011). The barrier property of the junction in the tissue is not absolute because homeostasis in the body has to be maintained, thus selective ions must regularly travel across the tissue.

Therefore the main function of the TJ is to create the regulating pores that act as a sieve to selective molecules. The TJ allows various materials to selectively travel across the epithelial sheet by regulating the epithelial transport system which consists of two

pathways, the paracellular and the transcellular pathways (Dörfel & Huber, 2012; Nag & Morin, 2009) (Figure 2).



Nature Reviews | Molecular Cell Biology

Figure 2: A Visual Representation of the Transcellular Pathway and the Paracellular Pathway. Arrows show the movement of ions, water and solutes from the lumen and the underlying cells of the body. (This diagram was taken from Tsukita et al., 2001)

The transcellular pathway is involved in the passage of ions, solutes or water via the epithelia cells. On the other hand, the paracellular pathway is involved in the movement of the same molecules, channelled through the TJ (Cybulski & de Mendoza, 2011). The TJ also plays a role in the transcellular pathway. By sealing off the paracellular pathway the TJ forces materials to enter the cells in order to pass through the tissue via the transcellular pathway.

Another function of the tight junction is to maintain polarity of cells by creating the boundary in plasma membrane bilayer separating the cell surface into the biochemically and functionally distinct apical and basolateral membrane regions (Cybulski & de Mendoza, 2011). This is achieved by preventing lateral diffusion of integral membrane proteins between the membrane surfaces thus allowing the cells to carry out polarized transport (Shen et al., 2011; Gumbiner 1993). It also holds the cells together. TJs have intracellular membrane-associated proteins that link them to signalling pathways and the cytoskeleton (Morales-Tlalpan et al., 2012). This allows for the communication between the cell and other substratum adhesion sites (Schneeberger & Lynch, 2004). Additionally, many cytoplasmic scaffolding molecules associated with the TJ are involved in regulating diverse processes such as cell proliferation (Schneeberger & Lynch, 2004).

The TJ are not all the same in structure. Microscopic images show different TJs appearing as close associations of cells with varying pore sizes which determine the tightness of the

overall epithelial sheet. Permeability properties of the TJ are dependent on the physiological requirements of the transepithelial solute transport in different epithelia of different compartments. For this reason different classes of epithelia can be found in the body, with each class determined based on the barrier ability of the TJ. Tight epithelia have complex TJs that prevent the movement of solutes between the cells. These TJ can be found in the epithelia of distal convoluted tubule of the kidney. On the other hand leaky epithelia (as for example, TJ found in the kidney proximal tubule) have less complex TJs.

Initially it was assumed that occludins were the core units of these epithelial junctions but CLDN 1 double knockout mice maintained overall junction structure and functionally (McCarthy et al., 2000), revealing that other proteins were essential components of the TJ. In 1998 Dr Tsukita discovered CLDNs in membrane fractions from chicken liver, enriched with TJ. Over-expression of *cldns* in fibroblasts, which do not have any TJs, leads to the formation of tight-junction-like network of strands (Günzel & Yu, 2013). Henceforth these experiments disprove the primary theory of occludins being the core unit of the TJ in favour of CLDNs as the core unit of the TJ (Van Itallie & Anderson, 2006). Evidence suggests that the occludin is involved in cell signalling instead of formation of the paracellular structure (Barrios-Rodiles et al., 2005). Interestingly CLDNs are also located in the basolateral membranes, possibly as precursors to the fibrils (Peter et al., 2004).

Claudins are encoded by a superfamily of genes (*cldns*) that are found in all vertebrates (Morales-Tlalpan et al., 2012). These genes encode membrane proteins also known as the epithelial membrane proteins (EMP)/peripheral myelin protein/ epithelial membrane protein or membrane protein (MP20) (Morales-Tlalpan et al., 2012). These membrane proteins are important for the structure and function of the TJ between the epithelial cells (Krause et al., 2009). This largely contributes to the barrier property of the epithelial sheet as it blocks off the external lumen from the underlying tissues (Tsukita & Furuse, 2000). Additionally it should be noted that there is a difference between the barrier properties and the fencing characteristics of the TJ.

By means of the integral membrane domains of the claudins, TJs also act as a fence which prevents the diffusion of proteins and lipids within the plane of the lipid bilayer. Ultimately, this ensures that proteins and lipid components of the plasma membrane remain separated and integral membrane proteins do not drift between the surfaces of the cell (Cybulski & de Mendoza, 2011; Nag & Morin, 2009; Cybulski & Mendoza, 2011).

In this way the apical-basolateral gradient of epithelial cells is preserved (Günzel & Yu, 2013).

1.2. Claudins

The epithelia has numerous functions in the body, the most important function is protection (Van de Peer et al., 2009; Coulombe & Lee, 2001). The skin of vertebrates is made of the epidermis, which in itself is an external layer of epithelial tissue, and dermis, under the epidermis. This layer provides protection to the internal organs from the external environment. Embryos of vertebrates contain keratin in their epithelial cells of the skin which prevents the loss of water (Coulombe & Lee, 2001). Other functions of this tissue include secretion, absorption, excretion and sensory reception. The functional properties of the epithelia are determined by the characteristics of the TJ which in turn are characterised by the composition of the CLDNs regulating the function of the TJ.

All CLDNs have the same basic structural components as shown in Figure 3. However it should be noted that the sizes may vary.

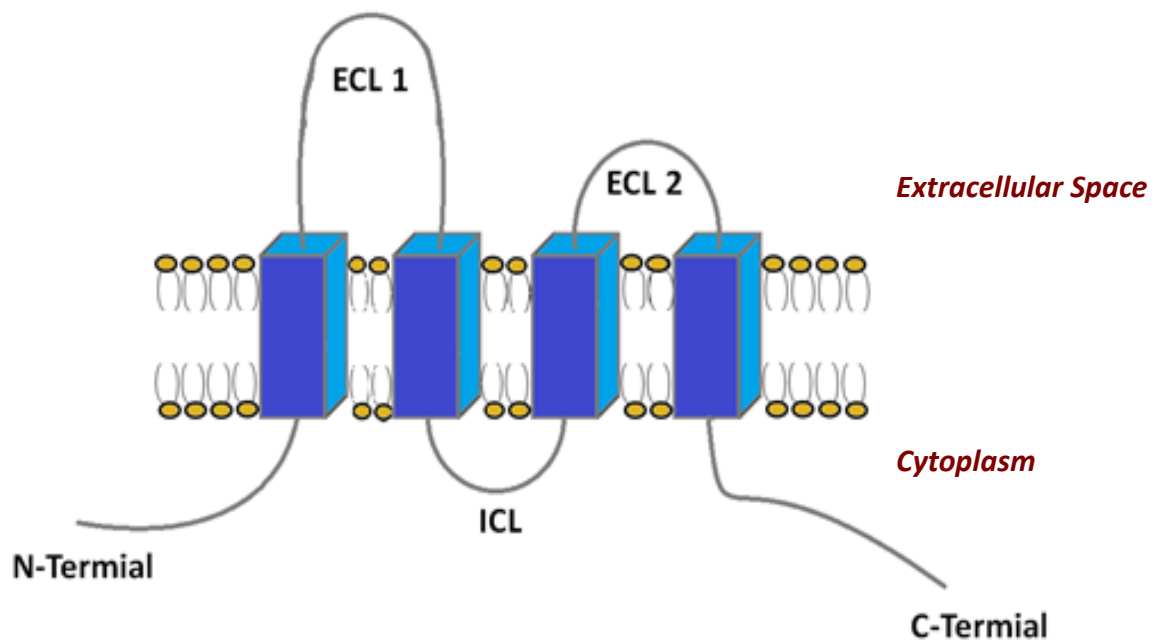


Figure 3: Diagram Showing the Structure of the Claudin Protein. The blue rectangle shapes represent the transmembrane domains (TMD); ECL 1 and ECL 2 represent the two extracellular loops; the ICL represents the intracellular loop

CLDNs have an intracellular N-terminal consisting of approximately 7 amino acids (Acloque et al., 2009). On the other end of the protein is an intracellular C terminal consisting of 25 to 55 amino acids (Acloque et al., 2009). The C-terminal has a PDZ-

domain-binding motif that allows CLDNs to interact directly with cytoplasmic scaffolding proteins, such as MUPP1; PATJ; ZO-1 and other proteins associated with the tight junction (TJ) (Morales-Tlalpan et al., 2012). By means of the PDZ domain, the C-terminal also functions in phosphorylation and signalling with molecules in the cytoplasm (Gupta & Ryan, 2010).

All CLDNs have four transmembrane domains which interact with the cytoskeleton to anchor the protein tightly in the cellular membrane. These domains also interact with two extracellular loops and one intracellular loop. The first and longer extracellular loop, the ECL1 (Shown in figure 2), contains approximately 60 amino acids with two conserved cysteines (Colegio et al., 2002); however there is great variations in the overall amino acid sequence. The ECL1 is involved in the pore forming property of the CLDNs in the TJ thus aiding the paracellular ion permeability (Krause et al., 2009). The charged amino acids found in the ECL1 affects the tightness of the pore formed at the TJ. These amino acids also generates the positive and negative electric fields which determines the types of ions that travel through the junction (Günzel & Yu, 2013). The paracellular transport system occurs as a passive process of materials crossing the electro-chemical gradient. It is produced by the transcellular pathway; formed by the Na^+/K^+ -ATPase (Bagnat et al., 2007); or other gradients created by external factors. An example of such external factors is the ingestion of solutes (Van Itallie & Anderson, 2006). Ion balance of TJs varies from tissue to tissue and the selectivity between cations and anions varies by less than 30 fold in native tissue. This suggests that the transport system is specifically designed to cater for multiple types of ions and variations in the TJ structure, ECL1 amino sequence, which increase the specificity of ions at a given site (Van Itallie & Anderson, 2006).

The second extracellular loop, the ECL 2 (also shown in Figure 2), contains approximately 24 amino acids which are important for the function of the CLDNs (Gupta & Ryan, 2010). The ECL 2, is involved in the side-to-side oligomerization of CLDNs from adjacent epithelial cells leading to the formation of the tight junction (Gupta & Ryan, 2010), which encircle the apical end of lateral surface of these cells (Dorfel & Huber, 2012; Van Itallie & Anderson, 2006). Alterations in the number of amino acids found in the ECL2 may vary slightly; for example CLDN11 has fewer amino acids than the usual 24 and is involved in tight junction formation in the basal cells of the stria vascularis in mice (Nag & Morin, 2009). The number of amino acids in this loop governs the size of the paracellular cleft.

CLDN association however is not limited to adjacent cell connection. Four possible types of interactions may occur based on the claudins involved and their location. In reference to the claudins involved, the interaction is known as homophilic (same CLDN, example CLDN 3 – CLDN 3) or the interaction is heterophilic (different CLDNs, example CLDN 1 – CLDN 3) (Krause et al., 2009). With regards to the location of the claudins involved in the oligomerization, the interactions may be either *cis*-acting (interactions between CLDNs in same plasma membrane) or *trans*-acting (CLDNs in adjacent plasma membranes). Together these lead to the four types of interactions; homophilic *cis*-interaction (same CLDN, same plasma membrane), homophilic *trans*-interaction (same CLDN, opposite plasma membranes), heterophilic *cis*-interaction (different CLDNs, same plasma membrane) and heterophilic *trans*-interaction (different CLDNs, opposite plasma membranes). These interactions are represented in Figure 4.

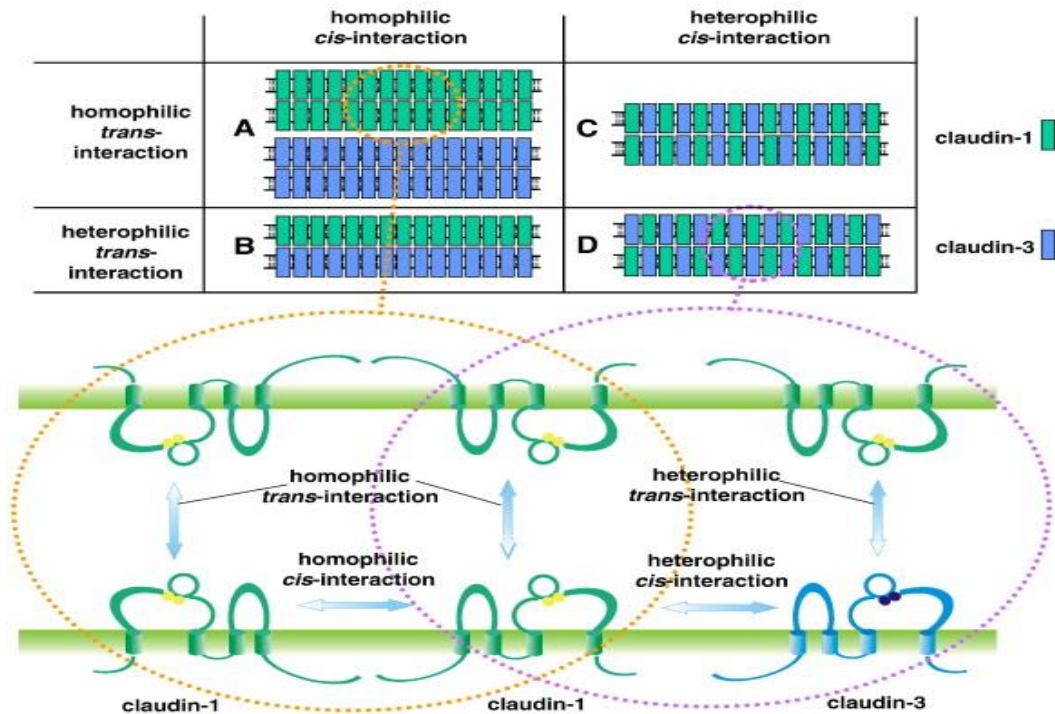


Figure 4: Different Types of Claudin Protein Interactions. A) Comparison between *cis*- & *trans*-homophilic interactions. B) *cis*-homophilic interaction and *trans*-heterophilic. C) Represents both *cis*-heterophilic interaction and *trans*-homophilic. D) Shows both *cis*- & *trans*-heterophilic interactions (Krause et al., 2008)

Additionally sequence analysis has led to the differentiation of *cldns* into two groups, classic claudins (consisting of highly conserved *cldns* [1-10; 14; 15; 17; 19]) and nonclassic claudins (consisting of *cldns* such as *cldn* 11-13; 16; 18; 20-24) (Krause et al., 2009). Classic claudins have conserved structural features in the ECL 1, however non-conserved negatively charged residues in the ECL 1 of CLDN 2 and 15 are thought to be

responsible for cation pore formation at the TJ (Krause et al., 2008). The ECL2 of classic claudins consists of approximately 25 amino acids and is of great importance in holding the cells together and narrowing the paracellular cleft. In nonclassic claudins this loop ranges from 17 to 39 amino acids, allowing a greater variation in the size of the paracellular cleft from a very narrow cleft to a more open one (Krause et al., 2009).

Both classes of *clnns* play a vital role in paracellular ion channels. CLDN composition in the TJ affects the sealing off property of the epithelial sheets. Tight epithelial sheets have a high transepithelial electrical potential and can generate ion gradients. The interactions of CLDNs at the TJ contribute to varying types of epithelial sheets and thus its effects. Leaky epithelial sheets have low transepithelial potential. The function of the two types of epithelia can be seen in the urinary bladder and the gastrointestinal tract (Van Itallie & Anderson, 2006). The former uses thick epithelia to generate intraluminal fluids which diverge from that of interstitial fluid, as the urinary bladder serves as a barrier to electrolytes in urine and water (Günzel & Yu, 2013). The latter uses leaky epithelia to function in the movement of large volumes of iso-osmotic fluids (Van Itallie & Anderson, 2006). Intestinal epithelia form the barrier to bacterial toxins in the gut. ECL 2 components of intestinal claudins may be involved in bacterial toxin recognition (Van Itallie & Anderson, 2006; Günzel & Yu, 2013). The occurrence of a specific type of epithelia is dependent on the permeability of the TJ pores within the epithelia, which in itself is dependent on the properties of the extracellular loops. These contain different electrostatic interaction sites which again contribute to specificity (Morales-Tlalpan *et al.*, 2012) and have an overall effect on epithelial morphogenesis.

1.3. Claudins in Embryonic Epithelia

Embryonic morphogenesis is the process whereby the embryo will undergo tissue differentiation and develop anatomical structures according to the genetic background of the organism. This process relies heavily on the reshaping of the epithelia covering the internal and external surfaces (Dörfel, 2012).

There are many factors which influence embryonic morphogenesis. Such factors include the arrangement of actin cytoskeleton; the orientation of cell-cell and cell-matrix adhesion or most importantly, the spatial restriction due to the extracellular materials and the hydrostatic pressures. In this manner the transcellular and paracellular pathways are required to create the hydrostatic pressures as part of the epithelial transport system accumulating fluid inside the lumen (Thompson & Barasitser et al., 1988). In this way the

expression of claudins in the TJ can ultimately have a vital role in development (Thompson & Barasitser et al., 1988).

Investigation into the gene family has identified many claudins expressed during embryo morphogenesis, such as *cldn 5a* in zebrafish. This gene has been found to be essential for the normal development of the embryonic cerebral-ventricular barrier system which forms part of the neuroepithelium lining the brain (Zhang et al., 2010; Abdelilah-Seyfreid, 2010). The gene is involved in the expansion of the ventricular lumen by blockage formed at the paracellular sites. The mutants were seen to have smaller ventricular volume without the organization of neuroepithelium being affected. Mutation in the human homolog, *cldn 5*, results in velocardiofacial syndrome. *Cldn 5* is expressed in the kidney vasculature of the human embryos. Mutations in *cldn 19* would result in eye and skeletal malfunction, an example being the Sorsby macular coloboma syndrome (Nag & Morin, 2009). The toes and the fingers do not develop normally as they may be fused or this syndrome causes formation of an underdeveloped retina (Thompson & Baraitser, 1988).

During embryonic development CLDN 1 and 2 are present at the cell boundaries in the neonatal stage of development, more specifically in the tight segments of nephron and the leaky portion (proximal).

In mice *Cldn 4* and *6* have been seen to be important in the formation of the blastocyst. These two CLDNs function in sealing off the intracellular spaces of the trophectoderm. By doing so, they contribute to the accumulation of fluid in the blastocoel cavity eventually leading to an expansion of the cavity due to hydrostatic pressure.

In zebrafish *cldn 15* is expressed in the gut where it is essential for the development of a single lumen gut during development of the larvae. Knockdown mutants of *cldn 15* results in a multi-luminate gut (Bagnat et al., 2007). In mice *cldn 15* mutants suffer from megaintestine which means that here the gene is involved in regulating the size of the intestine, a more specific role which has emerged from the rise of more claudins which all play a specific role in the development of the more complex gut. The general function of CLDN 15 is to increase the paracellular cation permeability in mammals - a feature important for the ion exchange in the gut.

Cldn 1-deficient mice were seen to suffer from dehydration due to transdermal water loss, highlighting the role of this claudin in the epithelia of adult mice (Nag & Morin, 2009). As a result these mutants die soon after birth (Furuse et al., 2002). *Cldn 14* double

knockout mice resulted in deafness possibly because of the weakened ion discrimination in the epithelial layers directly associated with the reticular lamina. Deafness was also seen in *Cldn 11* deficient mice. The cause of this was found to be due to the loss of TJ in basal cells of *stria vascularis* (Nag & Morin, 2009).

1.4. Expansion and Diversification of the Claudin Gene Family in Vertebrates

During the course of evolution of vertebrates, the chordate body plan was, in part, retained while novel structures such as the cranium, paired sensory organs and paired appendages were added (Kollmar et al., 2001).

Novel traits arise from pre-existing ones as we see in the development of the human lung and homologous swim bladder in fish which express *cldn 4, 5, 6* and *7*. The emergence of the primitive lungs developed from gas filled sacs in primitive fish. In some lineages the organ became more specialized to its function in gaseous exchange, eventually leading to the rise of new respiratory organs. Throughout evolution more complex classes of vertebrates adapted novel types of respiratory systems. The molecular bases of lung development in humans show that multiple claudins are involved in all five stages of lung organogenesis. The pseudoglandular and canalicular stages show expression of claudin 1, 3, 4, 5 and 7 in the bronchial epithelium. The full claudin protein expression profile in the different cells can be seen in Table 1.

Table 1: Expression of Claudins in Different Cell Types during Lung Organogenesis

Cells	CLDN
Bronchial Cells	1, 2, 3, 4, 5, 7
Alveolar Cells Type I	2, 3, 4, 5, 7, 18
Alveolar Cells Type II	1, 3, 4, 5, 7, 8, 18
Mesothelial Cell	1, 2, 3, 5, 7

This table represents the complexity of the human lung and how claudins are involved with the developments of the different units making up the respiratory system, from the bronchial epithelium to the alveolus. In adults the lung is lined by specialized types of epithelia organised to fit a tree-like formation, with three anatomical and functional units (Morales-Tlalpan et al., 2012). The first is the tracheobronchial structure, the second is

the bronchiolar and the third is the peripheral saccular-alveolar structure (Maeda et al., 2007).

Swim bladder development is not as complex as can be seen in zebrafish where the swim bladder first forms as a single chamber which then inflates 3 days after hatching. The anterior chamber is formed by invagination from the cranial end of the original chamber then both chambers enlarge with the ductus communicans forming a constriction between them (Robertson, 2007). From this it can be concluded that there must be expression of claudins during this point to regulate the sealing off of the epithelium and allow for the inflation and division of the chambers during swim bladder development. The zebrafish homologs *cldn* 4, 5, 6, 7 and 9 were identified in the swim bladder (Zheng 2011). This is not as many as seen in humans which have 8 claudins essential for development of the complex gas exchange system, almost double that required in the zebrafish. This further enforces the role of novel claudins in the evolution novel vertebrate characteristics.

In humans *CLDN* 9 is not expressed in lung organogenesis, instead it is highly expressed in the inner ear and in all of the major epithelial cell types that line the endolymphatic space (Zheng, 2011). *CLDN* 9 mutant mice suffer from deafness as this claudin forms the Na^+/K^+ barrier in the inner ear. This is an example of the change in claudin expression profile from the swim bladder to the ear, with the primitive *CLDN* 9 being recruited for ear development in higher vertebrates. This evolution of *CLDN* 9 from its role in development of gas exchange organ to development of the inner ear supports the theory that different epithelium permeability, by means of different *CLDN* functions, affected the vertebrate body plans leading to the rise of novel traits.

Vertebrates were affected by two types of duplications, the first was whole genome duplication and the second was gene duplications. After the divergence of the cephalochordate the vertebrate lineage experienced two whole genome duplication events (Van de Peer et al., 2009). This was further supported by the work done by Smith et al., 2012. They sequenced and partially assembled the genome of the sea lamprey, an ancient vertebrate lineage that diverged from our own approximately 500 million years ago. Analyse of the assembly indicated that two whole duplication events occurred before the divergence of lamprey and gnathostome lineages (Smith et al., 2012).

These duplication events led to more redundant genes which in specific lineages were further enriched by individual gene duplication. In mammals *CLDN* 3 and 4 are

duplicates, with both being located on the chromosome 17, in this case suggesting that duplication arose as a result of tandem repeat mutation (Takezaki et al., 2003). This can also be seen in chicken with *CLDN 3* and *4* located 5kb apart on chromosome 19. In the same species *CLDN 8* and *17* are located apart on chicken chromosome 1 (Collins et al., 2013).

Zebrafish experienced further duplication in the previously mentioned claudin 5 gene leading to the development of *cldn 5a* and *cldn 5b*. These new duplicates serve an important role in not only the blood-brain barrier but also in the blood-retinal barrier. Here both *cldns* are expressed in the hyaloid vasculature with *cldn 5b* showing expression in the vasculature system (Xie, 2010). Generally in mammals, *cldn 5* is associated with sealing off functions. The evolution of the complex system could have emerged from the evolution and specification of both *cldn 5a* and *5b*. The *Xenopus* homolog to this zebrafish claudin is important for the development of the heart. This represents more recruiting of claudins in different compartments in higher vertebrates showing new claudin expression profiles. This phenomenon is vital to the development of modern organs such as the highly efficient kidneys in mammals.

The phenomenon of having duplicated genes allowed these genes to evolve independently and acquire novel functions or performs the same function in novel fashions/combinations (Kollmar et al., 2001); this may be what has occurred with the zebrafish *cldn 9* and its homolog in humans. Many more duplicates can be seen within the claudin gene family of higher vertebrates.

The presence of claudins in all vertebrates suggests that this gene family may have emerged before the rise of vertebrates. This theory is supported by the discovery of three claudin-like gene family members found in *D. melanogaster*, the fruit fly. This contradicts the previous theory which suggested that CLDN orthologous were absent in this species of invertebrates (Yasunori et al., 2003).

The three *D. melanogaster* claudin-like genes encode proteins which are required in the paracellular transport system. The Mega (Megatrachea) is a transmembrane protein homologous to claudins and acts in septate junctions. This protein has transepithelial barrier functions like conventional claudins and is important for normal tracheal cell morphogenesis (Behr et al., 2003). Mega is not however involved in the apical-basal axis formation or the integrity of the epithelia. Sinuous (Sinu) is the second protein encoded by the claudin-like gene. Sinuous protein share several characteristics with vertebrate

claudins, sharing the amino acids and sequence similar to canonical vertebrate claudins. It also localizes to and is involved in the function of the paracellular barrier junctions (Wu *et al.*, 2004). Kune protein localizes at the septate junction where it is required to organise the junction and the paracellular barrier (Nelson *et al.*, 2010). Like Mega it is not involved in the apical-basal polarity during development of the *Drosophila* (Nelson *et al.*, 2010). Mutation in any of these two *Drosophila* claudins (*simous* and *megatrachea*) result in tracheal tube defects due to the malfunction of the septate junction (functional equivalent of the vertebrate TJ) and the barrier function in tracheal epithelia (Bagnat *et al.*, 2007).

The *C. elegans* has five claudin-related genes (Asano *et al.*, 2003; Behr *et al.*, 2003) which encode four claudin-related integral membrane proteins (CLC 1 to 5). These proteins show great similarities in their sequence to those found in vertebrates. The CLC 1 has been shown to be expressed in the pharyngeal section of the digestive tubes (Asano *et al.*, 2003). CLC 2 is expressed in hypodermis cells, where it functions in the hypodermis barrier (Asano *et al.*, 2003). Eleven true claudin members of the claudin gene family were identified in *Ciona intestinalis* genome, which belong to the invertebrate Urochordata (Tunicata) subphylum (Sasakura *et al.*, 2003). Again the sequence of the proteins encoded for by these genes showed similarities to those characterised by claudin proteins in vertebrates. Claudin genes may have been found in some lineages of invertebrates but they are found in all lineages of the vertebrata subphylum with great diversity in family size.

The NCBI database shows that mammals have over 25 claudin members with humans showing 27 members. Chick show 19 members of the claudin gene family so far but due to the position of this species the true number is expected to be closer to 20. The database also shows that 20 claudin genes have been identified in *Xenopus laevis* genome. Zebrafish have about 88 members of the claudin gene family, but it should be noted that this species is tetraploid (4n) due to an additional round of genome duplication in the teleost lineage. The puffer fish, *Takifugu rubripes*, reportedly has approximately 56 claudin genes but this, again, is due to extensive rounds of gene duplication (Loh *et al.*, 2004). This can be as a results of the different environmental conditions that fish are exposed to, i.e., seawater *versus* fresh water.

Euryhaline fish are able to regulate their internal environment to balance out the changes of intracellular ion levels, cellular volumes and/or protein functionality during salinity fluctuations. This trait is dependent on the fine tuning of the ion transport system within

the gills (Tipsmark et al., 2008), thereby allowing these fish to survive in both fresh- and seawater. This is achieved by regulating the critical proteins including specific members of the claudin gene family within vital compartments, such as the gills. In the gills of euryhaline fish, $\text{Na}^+\text{-K}^+\text{-ATPase}$ and the co-transporter $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ are up regulated in seawater and inversely affected in freshwater (Tipsmark et al., 2008). Alterations to the ion transport will depend largely on the properties of the paracellular pathway as part of the epithelial transport system in the gills. As previously mentioned claudins largely contribute to the properties of the TJ therefore in this case they would contribute to the permeability of the ion channels in the different types of water environment. A CLDN3-like protein was found localized in the gill filament of freshwater tilapia and a CLDN4-like protein was found in the filament outer epithelia (Tipsmark et al., 2008). Expression of both genes was reduced in seawater tilapia, suggesting that these genes are required for the sealing off of the ion channels thus preventing ions from entering the interior of the fish and disrupting the osmolality. This is an example of the evolution of the claudin gene family function in determining the survival capabilities of vertebrates in different environment. Differential expression of the claudins regulates the permeability of the gills to ions, in this way increasing the diversity of environments that the euryhaline fish can survive in, i.e. freshwater and salt water.

According to Crow and Wager, 2006, gene duplication would play a vital role in how an organism adjusts to its environment, this may lead to divergence of that organism possibly into a novel species as further mutations occur. The *Fugu* claudin expression profile suggests that gene duplication has led to the rise of novel functions in specific tissues (Loh et al., 2004). New traits could have been adapted after the rise of new members of the claudin gene family, leading to new genes which probably acquire new functions in the different compartments of vertebrates.

1.5. Claudin 3

One of the most widely expressed claudins in vertebrates is *CLDN 3*, which suggests a variety of novel functions that this gene plays in compartmentalization during vertebrate embryo development. The expression profile of this gene during chick development has shown an early expression of the gene during mid-gastrula stages. Table 2 contains a summary of the known expression profile of *cldn 3* during development.

Table 2: Summary of Known Expression Profile of Claudin 3 in Vertebrate Model Organisms

Organism	Region of Expression (Stage of Development)	Citation/s
<i>Xenopus laevis</i>	Brain, heart, intestine, liver, lung, and the mesonephric kidney	Yanai et al., 2011.
<i>Danio rerio</i>	No Data	---
<i>Gallus gallus</i>	Primitive Streak, Henson's node (HH4), Endoderm (HH5 – HH10), Surface Ectoderm, Head Process Head Fold; Somites, Mesoderm (HH6 – HH8), Oral Pharynx (HH10 – HH14), Ear placode (HH11 – 22), Pharyngeal arches, Stomodeum, Liver, Gut, Lung, Nasal Placode (HH13 – HH22)	Collins et al., 2013. Haddad et al., 2011. Haworth et al., 2005. Ozden et al., 2010
<i>Mus Musculus</i>	Ectoderm (E7 – E11), Submandibular gland primordium, Pancreas tip epithelium, bladder, urothelium of bladder, Nasal cavity olfactory, Thyroid gland, Lung, Ear placode, Metanephros, Trachea (E14.5), Bladder, Submandibular gland, Ear (E16 – E18)	Chihara et al., 2013. Ohazama et al., 2007. Ohta et al., 2006

Most of the expression experiments have been done in the chick, which shows a great variety of expression especially in the head. Interestingly, the expression in the ear is very persistent from stage 11 onwards. Expression also persists in other areas such as the pharyngeal arches. Expression is also consistent in the developing ear of the mouse however, at present most of the work done for this gene's expression profile is at stage 14. Still the expression is very diverse in areas such as the bladder and the submandibular gland.

No clear expression profile has been mapped out for the *Xenopus laevis*, but an experiment done by Yanai et al., 2011, where they mapped out the gene expression in the two *Xenopus* species, identified the expression of *cldn 3* from the Nieuwkoop and Faber (NF) stage 2, corresponding to an embryo at the two cell stage, right through to NF stage 33 (Yanai et al., 2011). This was done through microarray analysis of whole embryo expression thus the exact location of expression was not determined. However other research in gene expression and function has identified the expression of this gene in the brain, heart, intestine, liver, lung, and the mesonephric kidney.

Interestingly, as of yet, no *cldn 3* homologs has been identified in the zebrafish however other vertebrates, more closely related to humans, for example, the chicken, still have the homolog. However the lineage maintained the *cldn 4* which is the duplicate of *cldn 3* in vertebrate lineage. Also the zebrafish *cldn h* was seen to group with *cldn 3* and *4* on the tree constructed by Loh et al., 2004.

The expression profile of this gene is most well known in *Gallus gallus*, through use of *in situ* hybridization experiments. This is good as these vertebrates are perfectly evolutionarily placed to bridge the gap between mammals and non-amniotic vertebrates.

By stage 8 of development of the chick embryo, expression is in the developing head fold and the surface ectoderm. The gene is also expressed in the epiblast on either side of the primitive streak but not in the neural plate (Haworth et al., 2005). At stage 10, expression is still in the endoderm and the ectoderm regions encompassing a region fated to give rise to the ectoderm of the developing first pharyngeal arch. At a later stage, 11, *cldn 3* expression begins in the ear placode, which is a thickening of the ectoderm on the outer surface of a developing embryo from which the ear develops. Low level expression is also in the oral regions and later in the stomodeum and the developing pharyngeal pouches, with strong expression persisting in the otic vesicle (Haworth et al., 2005). The interesting feature in Table 2 is the persisting expression of *cldn 3* in the ear placode from stage 11 in the chick, onwards to much older embryos, stage 22. In addition both developing mice and chick embryos have shown expression of *cldn 3* in the liver during later stages of embryo development.

As previously mentioned, certain genes were further enriched in individual vertebrate lineages by individual duplication events. The phylogenetic tree constructed by Loh et al., 2004, revealed that mammalian *CLDN 3* and *CLDN 4* are indeed duplicate genes as previously mentioned. On the same tree, 17 *Fugu* CLDNs are found in a cluster and show a relationship with the mammalian CLDN 3 and 4 suggesting several additional rounds of duplication in the *Fugu* lineage (Loh et al., 2004).

The Nextprot database shows that *CLDN 3* is widely expressed in the mouth, stomach, liver and endocrine pancreas. The same data base shows expression of *CLDN 4* in the mouth, endocrine pancreas, dermis, tendon and the brain (Rossa et al., 2014) (Xu et al., 2013).

There is a difference in expression of these duplicates for different systems for example *cldn 3* is expressed in the nervous system; the brain; similar to that of *Cldn 3* in mice where it is involved in the development of the blood brain barrier (Rossa et al., 2014) (Xu et al., 2013) (Shang et al., 2013).

Human *cldn 3* is not expressed in the dermal system, whereas the *cldn 4* is expressed, thereby showing recruitment of genes for different structures. The mature protein of CLDN 3 has 220 amino acids; 11 more than that of the mature CLDN 4, suggesting a

certain domain which could have been lost during the mutation which allows for the slightly varying functions between the paralogs leading to their varying expression profile. However there is some crossing over, for example the endocrine pancreas of complex endocrine system and the bronchoalveolar system of the lung, suggesting the more complex traits would require gene expression of duplicate gene possible for proper compartmentalization of these organs.

So far, a lot of the work done for this gene has mostly been restricted to initially identifying the expression of the gene during a certain stage focusing on the embryo as a whole. However some researches have taken it a step further and focused on the expression profile with the ultimate goal being to identify the functions of this gene in their region of expression, specifically during morphogenesis.

1.6. Claudin 8

Other claudins which are paralogs are the human *CLDN 8* and *17* found on chromosome 21 (Loh et al., 2004). In humans, *CLDN 8* is expressed in the kidney where it interacts with the multi-PDZ domain protein 1 to control TJ barrier functions by regulating conductance in epithelial cells (Jeansonne et al., 2003).

Table 3: Summary of Known Expression Profile of Claudin 8 in Vertebrate Model Organisms

Organism	Region of Expression (Stage of Development)	Citation/s
Xenopus laevis	Distal Tubule, Pronephric kidney, Brain Forebrain, Midbrain, Hindbrain (NF 35 – 37)	Raciti et al., 2008. Clellan & Kelly, 2010. Thisse et al., 2004
Danio rerio	No Data	---
Gallus gallus	Epiblast, Extraembryonic ectoderm (HH4 – HH7, Non-neural ectoderm, Pharynx (HH8 – HH12), Eye epithelium, Nasal epithelium, Limb epithelium, Heart (E3 – E10)	Collins et al., 2013
Mus musculus	Endoderm (E8 – E9), Brain, Branchial arch, Ear placode, Heart, Primitive ventricle, Gut (E10 – E12), Ear, Inner ear, Middle ear, Pharynx, Submandibular gland Metanephros, Ureteric trunk, Cortical collecting duct, Distal tubule of capillary loop nephron, Kidney pelvis, Pelvic urothelial lining, Renal cortex, Renal medulla, Oral epithelium, Hindbrain, Pituitary gland, Cochlea, Developing eye, skin (E14 – E17)	Hou et al., 2010. Ohazama et al., 2007. Ohta et al., 2006. Fujita et al., 2006

Most of the expression profile for this gene has been done in mice. The developing ear seems to show most of the expression of this gene along with the submandibular gland and also the developing kidney. Experiments in *Xenopus* have shown this gene to be expressed in the brain and kidney at a later stage.

As with *cldn 3*, no *cldn 8* homologs have been unidentified in zebrafish however other vertebrates, more closely related to humans, for example chicken, still have the homolog. Yet again the duplicate of this gene can be found in this lineage, by way of *cldn 17*. Whole organism expression analysis have shown *cldn 8* to be expressed in *Xenopus* as early as NF stage 28 in areas such as the head and the intestine (Raciti *et al.*, 2008).

CLDN 8 has been identified in the *Gallus gallus* genome, the expression profile has not yet been determined especially with regards to the gene family as a whole in this species. By 2010, the NCBI database only had 6 claudins identified for this gene family. To date 19 members have been identified in the Chicken (Collins *et al.*, 2013). The *Gallus gallus* expression *in situ* hybridization analysis (Antin *et al.*, 2007) has record of the expression profiles for *CLDN 1*, *3* and *11* in embryo development.

The expression profile of *Cldn 8* in mice shows that the gene is initially expressed in the endoderm of the embryo by stage E8.25. Expression is strong in the brain by stage E10.5 at which stage expression also starts in the sensory organs; the gut and the bronchial

arches. At stage E14 expression is in the eye; the ear and the skin (Visel et al., 2004) (Reymond et al., 2002).

According to the Nextprot protein database, *CLDN 8* mRNA is highly expressed in the gastrointestinal tract structures, specifically the oesophagus; intestine and the mouth. Protein expression however is only seen in the intestine and the mouth. Its paralog, *CLDN 17*, shows mRNA expressed in the oesophagus and the mouth however. The protein's expression profile is similar to the mRNA only it also includes the intestine. This is a good example of recruitment of the CLDNs to perform novel functions in different regions by way of gain in functions made possible by the rise of novel family members.

In addition certain *cldns* have been lost in both mammalian and fish lineages throughout the course of evolution but the overall gene family size has continued to grow and this project investigated the role of the gene family in vertebrates.

Table 2 and 3 shows that although this gene family is found to be very important in many functions not just relating to the embryonic stage but also the adult stage, they are still relatively new and much work is required in order to better understand this gene family. There are still some gaps in their expression profile, particularly during embryogenesis which is important in determining their role in the vertebrate morphology as a whole.

1.7. Hypothesis and Aims

Hypothesis:

Expansion of the claudin gene superfamily correlates with the development of more complex vertebrate characteristics. We therefore hypothesize that this gene family played a vital role in the evolution of diverse vertebrates

Aims:

- Aim 1: Determine the claudin gene family size in the most basal extant vertebrate, the sea lamprey.
- Aim 2: To construct a phylogenetic tree consisting of claudin gene members from different organisms
- Aim 3: Map out the expression profile of claudins during embryonic development
- Aim 4: Determine the function of the claudins during embryonic development

CHAPTER TWO – MATERIALS AND METHODS

2.1. Culturing Lamprey Embryos

Eggs were removed from ovulating females by gently squeezing their abdomens, and deposited into a crystallization dish containing spring H₂O. Sperm from mature males was obtained in the same manner and deposited into the same dish, and the dish was swirled around several times before being allowed to set for 15 mins so that *in vitro* fertilization can occur. To prevent polyspermy, excess sperm was washed out by gently performing four distilled H₂O changes. The distilled H₂O was once again replaced with fresh spring H₂O and then the embryos were incubated at 18°C until the first cell division was underway, for approximately 5-6 hours. Within 10 hours of fertilization, the spring H₂O was replaced with sterile 0.1x MMR which contained CaCl₂. This provided the Ca²⁺ required by the cleaving embryos in order to form intercellular junctions. Embryos were checked on a daily basis. Dead/arrested embryos were removed and the solution was replaced with fresh 0.1x MMR. On the third day of incubation the embryos were spread at least 1cm apart in a glass dish to allow them to gastrulate. The embryos were left undisturbed until day 4.5 when they were brought back together and incubated in fresh 0.1x MMR at 4°C.

2.2. Morpholino Injection

Translation - blocking Fluorescein isothiocyanate (FITC) labelled morpholinos were obtained from GeneTools LLC. Custom 5 nucleotide mismatch for each morpholino were designed to serve as controls for this experiment since these oligos do not have any target sequences. The GeneTool standard control was also used in this experiment.

Both the Custom 5 nucleotide mismatch morpholino and the GeneTool standard controls are negative controls for this experiment, however the 5 nucleotide mismatch oligo is more specific for each target sequence. Since it has not been experimentally validated, it might be toxic due to off-target binding. The GeneTool standard control morpholino is a negative control oligo that targets a mutated human beta-globin intron which causes beta-thalassemia. This morpholino has been experimentally validated in lamprey embryos (Nikitina et al., 2008) (Sauka-Spengler et al., 2007) and was found to have no off-target effects.

The sequences for all morpholinos used in this experiment are shown in Table 4.

Table 4: Morpholino and Target Sequences used in the Knockout Experiments

Morpholino Name	Morpholino Sequence	Target Sequence
Cldn 3b	AAAAAACGGGAGTGTGGAAGAC GAT	ATCGTCTTCCACACTCCCGTTTT T
Cldn 3b Control	AAAAA tCcGc AGTGTG cAAc ACGA T	-
Cldn 8b	AGTTGGTTCTACGGCGTTCGCTC TC	GAGAGCGAACGCCGTAGAACCA ACT
Cldn 8b Control	AGT aGc TTCTAC cGc TT Cc CTCTC	-
Standard Control	CCTCTTACCTCAGTTACAATTTA TA	-

Wide-bore 3ml, sterile plastic pipettes were used to transfer 100 – 200 viable, one cell or two cell stage lamprey embryos into an injection dish. The excess 0.1x MMR was removed from the dish and the embryos were allowed to settle into the slots on the agarose. The dissecting dish holding the embryos was placed under a dissection microscope. Glass needles were filled with 5 – 6 µl of 0.25 – 1 mM of the morpholino. The lamprey embryos were injected on one side at the two-cell stage (for unilateral incorporation of the morpholino) and also at one-cell stage (for whole-cell embryo incorporation). The injected embryos were incubated at 18°C in 0.1x MMR until the desired developmental stage at which point the embryos were viewed under a Nikon dual excitation band DAPT-FITC filter set to select those that had integrated the morpholino. FITC has an excitation/emission spectrum peak wavelength of approximately 495nm/519nm thus these settings were used in viewing of the embryos.

2.3. Embryo Fixing and Processing

MEMFA was prepared per the protocol included in Appendix A. Freshly cultured lamprey embryos were collected in 2ml black capped clear glass bottle and fixed in MEMFA at RT for 1 hour. This was followed by three PBS/DEPC washes, for 15 min each time. The embryos were then dehydrated by washing them in 25%; 50% and 75% MeOH in 1xPBS/DEPC for 15 min each. The embryos were washed twice in 100% MeOH and then stored in fresh 100% MeOH at -20°C. The MeOH was replaced with 100% EtOH which helps better bleach the embryos during storage.

2.4. Cloning and sequencing

2.4.1. Bacterial Cultures

Luria broth (LB) (see Appendix A for the recipe) was used to grow cultures of bacterial strain JM 109 in 5ml media and similarly Luria agar (LA) were used to grow the bacteria on plates. The plasmid pCMV-SPORT 6 was used as the vector to carry the *cldn* inserted between the *Sal I* and *Not I* MCS of the plasmid. This plasmid was used because naturally the JM 109 strain does not contain an ampicillin resistance gene however the plasmid does, thus the antibiotic was used to select for the bacteria which contain the plasmid and also the insert of interest. JM 109 was used because it contains the genes *endA1* (abolishes non-specific endonuclease I activity); *recA1* (prevents recombination between introduced DNA and host DNA) and amongst others it contains *hsdR17* (rk-;mk+) (Restriction – ; modification +) transformed DNA which will not be cleaved by endogenous restriction endonuclease.

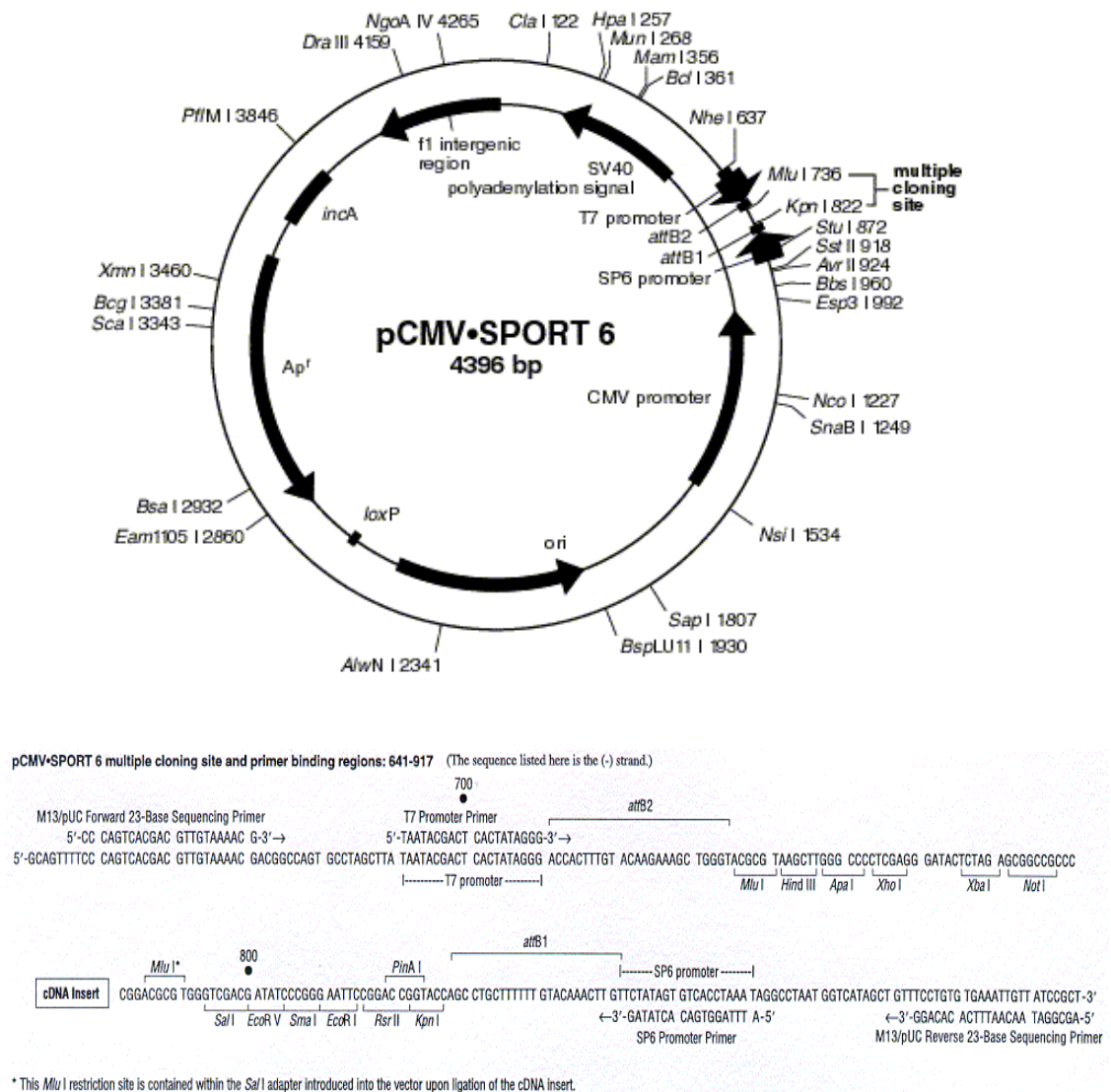


Figure 5: Map of the Plasmid pCMV-Sport-6 and its Multiple Cloning Site (MCS)

2.4.2. Glycerol Stocks

To maintain the pCMV-SPORT 6 plasmid containing the different *cldn* inserts for long-term storage, bacterial glycerol stocks were made of each clone. This was done by selecting a single colony from the plate of the desired clone that has the *cldn* inserts. The colony was grown overnight in 5ml LB ampicillin-containing media. The culture was then spun down at ~12 000rpm in 1.5 ml microcentrifuge in stages until all 5ml of culture was pelleted down in the microcentrifuge. The pellet was then resuspended in either 500 µl of fresh LB ampicillin media or 500 µl of 50% glycerol or 700 µl of fresh LB media and 300 µl of 100% glycerol. The glycerol stocks were stored at -70°C.

2.4.3. Transformation of *E.coli* JM109 and Screening

E.coli JM109 were transformed to incorporate the plasmid pCMV-SPORT 6 and thus became resistant to ampicillin and therefore LB ampicillin media and LA ampicillin plates were used to screen for and grow only colonies which contained the *clon* inserts. Different techniques were used to transform the bacteria but the screening method remained constant.

2.4.3.1. Transformation of Chemically (Calcium Chloride) Competent JM 109 Cells

LB media containing ampicillin and the LB agar plates containing ampicillin, were prepared as shown in the Appendix A.

5ml of preculture JM 109 cells were grown overnight in autoclaved LB media at 37°C. 25ml of fresh LB containing 10% glucose was then inoculated with 250 µl of the preculture. This new preculture was grown with vigorous aeration for 1.75 hours at 37°C. The culture was then incubated in ice-water slurry for 5mins before it was centrifuged at 4°C at 10 000rpm for 5min. The supernatant was discarded and the pellet of cells was resuspended in half of the original culture volume of ice-cold, sterile solution of Transformation Buffer (see Appendix A for recipe). The Calcium chloride (CaCl₂) was used as an essential component in allowing for the uptake of DNA into the bacterium. The positive calcium ions attract the negatively charged groups of lipopolysaccharide inner core and the negatively charged DNA backbone.

At this point the cells were left on ice for 5 minutes before DMSO was added to the solution to inhibit secondary structures of the DNA. DMSO is a polar aprotic solvent which dissolves both polar and nonpolar compounds thus passing through membranes and making it more permeable. The cells were incubated on ice for an additional 15 min. After which they were placed in the centrifuge and spun at 4°C at 10 000rpm for 5mins. The supernatant was discarded and the remaining pellet was resuspended gently in 1/15 transformation buffer and incubated on ice for 24 hours.

10 µl DNA sample of the plasmid, pCMV-SPORT 6 containing the respective *clon* insert were placed in 1.5ml Eppendorf tubes and left on ice for 10+ min. 47 µl of competent cells were added to each tube then gently mixed and incubated on ice for 10 min to allow diffusion. Heat shock was done at 42°C for 90 seconds in a water bath, this facilitated the intake of DNA into the cell. Following the heat shock, 0.5ml of broth was added to each

tube and then the tubes were incubated in a 37°C room with the caps left off. The new cells were spread on LA ampicillin plates and then grown at 37°C overnight.

2.4.3.2. Preparing JM 109 Competent Cells in the Lab

One colony of JM 109 was inoculated into 2ml of autoclaved LB media and then incubated overnight at 37°C. The full 2ml bacterial preculture was transferred into 50ml LB media. The bacterial cells were then grown at 37°C until the exponential phase, which was determined by a spectrometer reading of ± 0.6 . All subsequent steps were done on ice. The new LB media preculture was added to fill up 15ml pre-chilled falcon tubes and then centrifuged at 4°C for 5 min at 2 000 g. The supernatant was discarded then another 15ml of bacterial broth was centrifuged and discarded until all of the bacterial broth was spun down and the pellet was resuspended in 15ml of ice-cold 0.1M MgCl₂. The cells were incubated on ice for 30 min and then spun down at 4°C for 5 min at 2 000 g. The pellet of cells was resuspended in 2ml 0.1M CaCl₂ with 15% glycerol and then aliquoted in 50 µl samples in 2ml pre-chilled Eppendorf tubes. The competent cells were stored in -70°C freezer.

2.4.3.3. Transformation of Commercially Competent JM 109 Cells

2 µl of plasmid with the respective cldn insert was added to sterile 1.5 ml tube and then put on ice. Promega JM 109 high efficiency competent cells (or the competent JM 109 cells prepared in the lab) were placed in an ice bath until just thawed (approximately 5mins), then the cells were mixed by gently flicking the tube. 25–50 µl of the competent cells were carefully added to the tubes containing the DNA plasmids. The cells were mixed again by gentle flicking of the tube then placed on ice for 20 min. Heat-shock was done in a 42°C water bath for 45 – 50 seconds after which the tubes were immediately returned to ice for additional 2 min. SOC medium (see Appendix A for recipe) was stored at room temperature until 950 µl was poured into a culture tube and then the pre-treated bacterial cells were added to the medium. The cells were incubated for 1.5 hours at 37°C on a shaker. The LA ampicillin plates were used to plate out 100 µl of each transformation culture in duplicates. The plates were incubated at 37°C overnight. Colonies were then selected for further DNA extraction experiments.

2.4.4. DNA Plasmid Preparation and Analysis

2.4.4.1. DNA Miniprep

The plasmids were extracted from the bacterial culture by following the Fermentas GeneJET miniprep kit protocol. 50 µl was used to elute out the DNA plasmid from the column and when a higher concentration was required, 25 µl was used to elute DNA from two separate cultures of the same sample and then the eluted DNA was pooled together. Concentration was determined using the NanoDrop.

2.4.4.2. Double Digest

To confirm the presence of the inserts in the plasmids, 4-5 µl of each sample subjected to double digest using the Fermentas FastDigest enzymes, *Not I* and *Sal I*. *Xho I* could be used instead of *Not I*. The water bath temperature was increased to 65°C for 1 minute to inactivate the enzymes. The digests were then run on agarose gel at 100 volts for 40 -45 min.

2.4.5. Electrophoresis

The Thermo Scientific GeneRuler 1 kb DNA ladder was used as the molecular marker during all gel electrophoresis. GR Green was used as the fluorescent dye during gel electrophoresis because it is more sensitive than other nucleic acid dyes and non-toxic.

1% Agarose (1g/100ml) in TAE buffer was used to make up the agarose gel. Therefore 0.01% of GR green was used in the gel. The electrophoresis was run at 100 volts for 45min.

2.4.6. Sequencing & Sequence Analysis

The sequencing of the claudin clones was done at Inqaba Biotech where DMSO (Dimethyl sulfoxide) was used in order to inhibit the formation of secondary structures in the DNA template or primer, because lamprey DNA is quite GC-rich and repeat-rich, and therefore prone to secondary structure formation. Primers binding to SP6 and T7 promoter sequences, which flank the insert in pCMV-Sport6, were used in the first round of sequencing. However due to the long size of the inserts, several rounds of additional sequencing were done with multiple internal primers. The custom primers were designed by using the online program Primer 3 and the following parameters: 45-50 % GC; Size:

Min 15, Opt 18, Max 20; Melting temperature: min 53, Opt 55, Max 59. The Primers used for the sequencing of the clones can be seen in Table 5.

Table 5: Internal primers used to sequence the CLDN from the T7 and SP6 direction

Clone Name	Primer Name	Primer sequence (5' – 3')
CLDN 3B	T7	TAATACGACTCACTATAGGG
	SP6	ATTTAGGTGACACTATAG
CLDN 8B	T7	TAATACGACTCACTATAGGG
	T7 side Internal Primer 1	TTCCAGTACACGGCACA
	T7 side Internal Primer 2	AAAGAAGCTCCCACAGCA
	T7 side Internal Primer 3	TGCCAATGTTAGCTCTGC
	T7 side Internal Primer 4	GGGTGTTCTTTTGGGTTC
	T7 side Internal Primer 5	GGGTGTTCTTTGGGTTC
	SP6	ATTTAGGTGACACTATAG
	SP6 side Internal Primer 1	AGCAGCAATGCAAGGTCT
	SP6 side Internal Primer 2	TGTAGAAGTCGCGGATGA
	SP6 side Internal Primer 3	GGTGAAGATGGTTCGTGA
	SP6 side Internal Primer 4	GGGTGTTCTTTTGGGTTC
	SP6 side Internal Primer 5	CAAATCCGGAGTGTGAGA

2.4.6.1. Sequence Assembly

The sequencing results obtained from Inqaba Biotech were viewed using the Finch TV program for analysis of the quality of the sequence. The data was exported to FASTA format. DNASTAR Lasergene tool EditSeq was used to identify the *Sal I* site and delete everything before that. Similarly, the *Not I* site located just 5' to the insert was identified and everything after that was deleted. Since the sequencing was done from both ends of the gene different data was obtained for the same gene. All sequencing data sets per gene were assembled using the DNASTAR Lasergene tool SeqMan.

2.4.6.2. Blasting and Alignments

The NCBI tool Blast X was used to compare our *cldn* clones to the NCBI database of claudin sequences in higher vertebrates. The Blast X tool specifically uses a nucleotide query and searches the protein database. The clones were also compared to each other by using the NCBI multiple alignment tools.

2.5. Phylogenetic tree construction

2.5.1. Sequence Assembly

Claudin amino acid sequences from *Drosophila*, chicken, *Ciona intestinalis*, mouse, *Xenopus* and zebrafish were obtained from the NCBI and ENSEMBL database (See table in Appendix C for the full list of their accession numbers).

The ExPASy-Translate tool was used to translate claudin nucleic acid sequence into the amino acid sequence.

2.5.2. Multiple Alignment and Tree Construction

The Muscle multiple alignment tool, on the WITS wEMBOSS server, was used to align the collected sequence data. This wEMBOSS tool, performed a multiple alignment of the sequences by global optimization. As a branch of applied mathematics and numerical analysis it optimized a function or a set of functions according to certain criteria, in this case the similarities.

The complete alignment was used to construct a phylogenetic tree on the online program <http://www.phylogeny.fr/>; Phylogeny analysis – Advanced Mode. This program implemented the PhyML system to estimate the maximum likelihood phylogenies from the alignment of the amino acid sequences. This online program used the TreeDyn software to show the tree.

The wEMBOSS fproml tool; along with other online programs (https://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/; <http://www.genome.jp/tools/clustalw/>) were also used to construct more phylogenetic trees in order to get comparative trees and thus more accurate data.

2.6. DIG-labelled RNA Probe preparation

2.6.1. Probe Templates

Due to the high similarities identified in the coding regions of the *cldn* inserts (Cldn 3b and 8B) using full length probes may result in cross-hybridization of the probes to other, non-target claudin mRNA during *in situ* hybridization. Thus in order to get a specific expression profile for each gene, the RNA probe would need to be synthesised from the UTR only as the template for the DIG-labelled RNA probes.

The DNASTAR lasergene tool SeqBuilder was used to identify the coding region of *cldn A* and *cldn B* respectively. The same program was used to screen both full length sequences for restriction sites for enzymes that could be used to cut out the coding region of the genes. For the *cldn A* clone, the restriction enzymes *PvuII* and *NotI* were used to cut out the coding region leaving behind the 5'UTR. Similarly the restriction enzymes *PvuII* and *Sal I* were used to cut out the coding region from the *cldn B* clone leaving behind 3'UTR.

To confirm that the digests proceeded to completion and also to separate the pCMV-SPORT6 + UTR sequence from the coding region, the whole sample was run on a 1% agarose gel for 45mins at 100Volts. The bigger band (containing our DNA sample of interest) was extracted from the gel by using the Thermo Scientific GeneJET Gel Extraction Kit. However the enzymes resulted in liner bands with staggered ends which could not be re-ligated.

The enzyme *PvuII* cuts blunt ends, recognising the restriction site:

5'— CAG ↓ CTG — 3'

3'— GTC ↑ GAC — 5'

However *Sal I* and *Not I* cut staggered ends, recognising the following restriction sites:

5'— G ↓ TCGAC — 3' AND 5'— GC ↓ GGCCGC — 3'

3'— CAGCT ↑ G — 5' 3'— CGCCGG ↑ CG — 5' ; respectively.

Thus in order to re-circularize the plasmid the ends had to be repaired first to make the ligation possible. DNA end repair was required to fill in the staggered ends, making them compatible with the blunt ends created by the *PvuII* digest. The Thermo Scientific Fast DNA End Repair Kit was used to fill in 5' protruded DNA ends using the 5' – 3' polymerase simultaneously the 3' – 5' exonucleases degrades the 3'end overhangs making another blunt end (and so also eliminating the *Sal I* and *Not I* restriction site).

The Thermo Scientific GeneJET PCR purification kit was used to purify the DNA plasmid with the new UTR insert. The concentration of DNA was recorded and the DNA sample was stored in the -20°C freezer to be used later for re-circularization.

The following reaction mixture was used to induce self-circularization of the new plasmids containing the *cldn A* 5'UTR and *cldn B* 3'UTR respectively.

Table 6: Ligation Mix

Reagents	Quantity
Linear DNA	10-50 ng
10x T4 Ligase buffer	5 µl
T4 DNA ligase	5 µl
Nuclease free Water	To 20 µl
Total	20 µl

The mixture was incubated in a water bath at 22°C for 10mins; at that point 2.5 – 5 µl of the sample was used to transform competent *E.coli* JM 109 cells in order to clone the plasmids with the new inserts for probe synthesis.

2.6.2. Probe Synthesis

In order to synthesis RNA strands from the DNA plasmid and UTR inserts, 10ug of DNA was linearized by using *Sal I* restriction enzyme. The reaction mix composition can be seen in Table 7:

Table 7: Linearization Reaction Mix

Reagents	Quantity
DNA	10 µg
Restriction Buffer	2 – 10 µl
Restriction Enzyme	2 – 5 µl
dH ₂ O	Made up to 20 – 100 µl

The tube was incubated overnight in 37°C water bath.

The Promega Wizard SV gel and PCR clean-up kit was used to purify the linearized plasmids.

The phenol-chloroform extraction and ethanol precipitation technique were also attempted to purify the DNA.

300 µl samples were used; an equal volume of phenol: chloroform: isoamyl alcohol was mixed with the solution. Centrifugation was done at RT for 10 minutes and then the aqueous phase (containing the nucleic acid) was removed to a new tube. This process was

repeated twice before the final solution was subjected to ethanol precipitation and the new nucleic acid concentration was recorded.

The linearized purified DNA samples were used for *in vitro* transcription to synthesise the probe as shown in Table 8:

Table 8: Probe Mix Composition

Reagents	Quantity
Linearized Plasmid	1 µg
5x Transcription buffer	4 µl
10x DIG RNA labelling mix	2 µl
RNase Inhibitor	0.5 µl
Depc-Treated Water	Make up to 20 µl
Total Volume	20 µl

The probe mix was incubated for 2 hours at 37°C, after which 1 µl of the mix was run on an agarose gel.

After the 2 hours were completed, 2 µl of RNase free DNase + 2 µl DNase buffer were added to the mix. After a 1 hour incubation period, the addition of 0.5M EDTA was used to stop the reaction. Finally the Promega PCR clean up kit was used to purify the newly synthesised DIG-labelled Antisense RNA probes.

Alternatively, the DIG-Labelled RNA probe was purified by pipetting it into new 1.5 ml microcentrifuge tube, then 2.5 µl of 4M LiCl solution (DEPC-treated) and 75 µl of pre-chilled 100% EtOH, was added to the tubes. The RNA was mixed and then left for 1 hour at -20°C to precipitate. The tubes were spun at 16 000 g for 20 min at 4°C. The supernatant was discarded and the pellet was washed with 100 µl of ice-cold 70% ethanol (in DEPC H₂O) and the tubes were once again spun with the same parameters as before. The supernatant was once again discarded and then the tubes were left, lid open, to dry at RT or at 37°C incubator for 25 min. The purified RNA was dissolved by adding 50 µl of DEPC H₂O and incubating the tube in 37°C incubator for 30 mins. 1 µl of the purified RNA was run on an agarose gel to ensure the presence of the RNA post purification. 1 ml of Hybe Mix was added to the purified RNA stock and the tubes were stored at -20°C.

2.7. RNA *in situ* hybridization

2.7.1. Embryo Preparation

Embryos were staged according to the embryonic days of development at 18°C (Sauka-Spengler, 2007).

Lamprey embryos at from embryonic day 4 – 9 (E4 – E9) were re-hydrated from 100% MeOH to 0.1x PBS/DEPC in a MeOH series of 100%, 75%, 50% and 25% all in 0.1x PBS/DEPC. DEPC inactivates possible RNase in the solution. The embryos were washed in 0.1x PBS/DEPC three times for 30 min each time, this process allowed the chorion surrounding the embryo to expand. Then the embryos were placed in a Petri dish containing 1x PBS under a dissecting microscope. Fine forceps were used to remove the chorion membrane and free up the young embryo for *in situ* hybridization. Once dechoriation was complete the embryos were dehydrated back into 100% MeOH and stored at -20°C until needed.

For *in situ* hybridization, lamprey embryos between stages E4 – E31 were re-hydrated in methanol series from 100% MeOH to 1x PBS/DEPC in a MeOH series of 100%, 75%, 50% and 25% all in 1x PBS/DEPC. The embryos were bleached in freshly made bleaching solution consisting of 5% formamide, DEPC H₂O, 0.5% SSC and 10% H₂O₂.

Proteinase K in PBT Tween (DEPC) was used to partially degrade proteins to create the spaces for probes to move through in order to reach their mRNA targets. Glycine in 1x PBS/DEPC wash was done in order to stop the proteinase K activity; by binding the glycine to the active site of the proteinase and inactivating it.

The embryos were washed in 1:1 ratio of PBS-Tween (DEPC) and Hybe Mix (see Appendix A) before they were washed in 100% Hybe Mix and then stored in new Hybe Mix at -20°C.

2.7.2. *In situ* hybridization

The embryos prepared as previously described (Section 2.6.1.) were treated with fresh Hybe Mix and then incubated at 70°C for 1 – 2 hours in a hybridization oven (Mettmert). Simultaneously the RNA probe stock in Hybe Mix was diluted to 300 µl/2 ml Hybe Mix and this probe aliquot was also incubated in the hybridization oven for 1 – 2 hours. The embryos were then treated with the probe mix and then incubated again at 70°C for at least 16 hours.

At this point the probes had hybridized with the complementary mRNA in the embryos. Next, the posthybridization washes were performed to remove the unbound probe. The probe solution was removed and stored at -20°C to be used again for later *in situ* hybridization. The embryos were washed six times in new hybe mix with two washes being done for 15mins each and the remaining four being done for 30 min each. The embryos were washed in 1:1 ratio of Hybe mix: 1x MABRT, with the first wash being done for 15 min and the second wash was done for 30 min. Next, four 30 min washes were done in 1x MABT.

Bovine serum albumin (BSA) was used in the blocking solution which consisted of 3% BSA in 1x MABT. This step was necessary to block the sites for potential non-specific binding of the antibody, and thus ensure the high specificity of the next step. Anti-DIG antibody conjugated to alkaline phosphatase (anti-DIG-AP) in blocking solution was added to the embryos in order to detect the DIG-labelled probes hybridized to their complementary mRNA strand.

Post antibody washes were done in 1x MABT, twice for 5mins, twice for 30 min each then six times for one hour each. These washes were done to wash off unbound antibody. The embryos were then washed in 1x MABT at 4°C overnight as this buffer also inhibits endogenous phosphatases.

The embryos were washed with the alkaline phosphatase buffer, NTMT (Appendix A). BM Purple is a colour detection reagent, which forms purplish-blue precipitates in the presence of alkaline phosphatase, thus BM Purple was used to identify the expression profile of the claudins. The embryos were incubated in this substrate for 1 hour and then incubated at 4°C until the desired stain had been reached. The reaction was monitored regularly by visual inspection, and is stopped washing the embryos with PBST. MeOH washes were used to remove any background staining and then the embryos were fixed in 4% PFA at 4°C overnight. The embryos were washed in glycerol series and stored in 75% glycerol for photographing at a later stage.

Agar plates were used as the stage to set the embryos on for taking pictures. 25% glycerol was placed as a coating over the agar so that the 75% glycerol bathed embryos would sink to the bottom and thus no shadows would appear in the photos. The Zeiss Axio Zoom.V16 camera microscope with a Zeiss Axio Cam ICC 5 camera was used to photograph the embryos.

2.8. Sectioning

2.8.1. Gelatin Embedding

The embryos were washed three times with 1x PBS for 5 min each time. The embryos were fixed with 4% PFA for 1 hour at RT. This was followed by four more 1x PBS washes in 5 min intervals. To protect the embryos from crystal formation during the freezing process, they were washed in 5% sucrose for 5 hours at RT and then in 15% sucrose overnight at 4°C. Embryos were incubated in 37°C pre-warmed 7.5% gelatin overnight or for at least 5-6 hours. The embedding was done using 20% gelatin and TEM rubber molds. The embedded embryos were flash-frozen in liquid nitrogen and the blocks were stored at -20°C. Sectioning was done at -30°C in the Leica CM 1510 S cryostat.

2.8.2. Slide Mounting

The slides with embryo sections were incubated in pre-warmed PBS at 42°C for 10 min. This was done to remove the gelatin. The slides were washed twice in 1x PBS at RT for 5 mins each time. The slides were then removed from the slide holder one at a time, the excess PBS was shaken off and 2 drops of aqueous mounting media (Sigma CC/Mount tissue mounting medium) were applied to the slide before a 60mm glass coverslip was placed over the tissue carefully to avoid bubbles. The slides were left on the bench for a few hours to allow the mounting medium to set for viewing and then they were stored horizontally. The slides were sealed with nail polish to prevent the mounting medium from drying out. Photographs of the sections were taken using the Olympus AX70 camera microscope with an Olympus soft imaging system camera, CC-12.

CHAPTER THREE – RESULTS

3.1. Phylogenetic tree indicates that claudin homologs and orthologs can be found in basal and more complex vertebrates, and paralogs can also be seen in individual lineages

The phylogenetic tree was constructed using claudins from selected vertebrates. Ideally homologs are best to use in a phylogeny however since this tree included all the members from the claudin gene family, orthologs were also included. Nevertheless, the conservation may be more closely linked with genes in the same species due to the rate of conservation of intron and exons during speciation. The rise of novel members of the claudin gene family, as a result of duplication in certain lineages, may affect homology of the gene sequence while maintaining similar amino acid sequence thus data collection was crucial to constructing this tree.

The tree focused on elucidating the functional evolutionary relationship between the claudins, thus it was important to separate introns from the exons because of the difference in the rate of evolution between those two DNA sequences. Thus modified mRNA was used.

The tree included CLDNs from vertebrate model organisms; *Mus musculus* (26 members); *Gallus gallus* (16 members); *Xenopus laevis* (20 members) and *Danio rerio* (33 members). The sequences were collected from various online databases such as Ensemble Genome Browser and the National Center for Biotechnology Information (see Appendix A). Only the amino acid sequences were used. Where necessary the DNA sequences were translated into amino acid sequences to be used in the final tree. To identify how the claudins from the basal vertebrate fits in with that of higher vertebrates, the 20 members of the *Petromyzon marinus* claudin gene family (see Appendix A) were crucial in the construction of the phylogenetic tree.

As previously mentioned, claudins do predate vertebrates, therefore several invertebrate claudins from *Drosophila melanogaster* and *Ciona intestinalis* were included to root the tree. The *Ciona intestinalis*, belonging to the phylum Chordata, is an ascidian that has been used as a model invertebrate chordate for developmental biology and genomics. Initially it was believed that this species was a kind of mollusc. However, phylogenetic and phylogenomic studies have confirmed this sea squid to be the closest relatives of vertebrates (Putnam et al., 2008). This makes the organism a suitable outgroup for the vertebrates, especially since the CLDN in this organism have been identified. The data

was collected from Ensemble Genome Browser database, although most sequences needed to be translated to the amino acid sequences, 11 members from the *Ciona intestinalis* were used.

Drosophila melanogaster, more commonly known as the fruit fly, is a species of fly which belong to the phylum Arthropoda. This species has been used as a model organism for biological research in genetics among other fields. Although more distantly related to vertebrates than the *Ciona intestinalis*, the fruit fly has three claudin-like genes encoding proteins involved in the paracellular pathway of the septate junction. Therefore the tree was expanded with *Drosophila* was used and the amino acid sequences were added to root the tree.

The wEMBOSS tool, Muscle, was used to align the complete dataset of claudin amino acid sequences. The program is a wrapper program for the program MUSCLE of Robert Edgar (Edgar 2004). The program incorporates an input of nucleic acid or protein sequences that has a reasonable degree of similarities over their whole length and produces a multiple sequence alignment as an output. This process is carried out in 3 stages as discussed in detail below.

Stage one, the k – mer clustering, computes all pairwise similarity scores based on a fast procedure involving counting of common oligomers. The scores are transformed into distances and used to calculate a dendrogram like a phylogenetic tree using UPGMA (Unweighted Pair Group Method with Arithmetic mean), neighbour-joining or a combination called the UPGMB. Sequences are progressively aligned according to the branching order in the tree. This step is very important as it sets up the initial analysis of the tree and also the output is already based on a rooted phylogenetic-like ‘guide tree’. The input parameters in this program allows for the selection of the outgroup, in this case, the *Drosophila*.

Stage two, tree refinement, computes all pairwise distances from the sequences as taken from the multiple sequences alignment and using the method of Kimura (Edgar, 2004). This method distinguishes between transitions and transversions in DNA sequences, for correcting of multiple substitutions at one site. This step then re-computes a tree which, if different from stage one tree, re-aligns the groups of sequences at those nodes that have changed. With the large size of the data set refinement is a necessity.

Stage three, bipartite refinement, uses the final ‘guide tree’ from stage 2. It visits all the edges beginning with those most remote to the root. For each edge it makes a bipartition

of the alignment into two profiles by deleting an edge from the tree, resulting into two sets of sequences. Out of the two profiles, one then re-aligns. If the alignment has been improved it is retained, if not it is discarded and the previous version of the alignment is maintained. This continues until all edges have been visited without any further change needed to be made.

The complete alignment was used to construct a phylogenetic tree on the online program phylogeny.fr. This program implements the PhyML system for which the strength of this software is in the large number of substitution models coupled to various options to search the space of phylogenetic tree topologies.

Fig. 6 shows the final tree consisting of all CLDNs from selected model organisms. The lamprey claudins have been marked off with the blue and purple boxes. The blue boxes identify claudins with orthologs in higher vertebrates, for example lamprey *cldn 16* pairs with its orthologs in higher vertebrates thus all the CLDN 16 pair together. Furthermore the other *cldn16* in higher vertebrates were more closely related to each other compared to the lamprey *cldn 16*. The same phenomenon can be seen with *cldn 12* and *cldn 1b* in the lamprey.

The lamprey claudin initially named *cldn A*, grouped together with the CLDN 3s, suggesting that it is an ortholog. Thus *cldn A* was then renamed *cldn 3b*, shown on the tree as PmCLDN 3B_CLDN A. The Partial sequence of lamprey *cldn 3a* does not group with the other CLDN 3s. Possibly, Cldn 3a and 3b would have appeared as paralogs on the tree had the full sequence for both genes been identified.

The purple boxes on the tree indicate the paralogs which have arisen early in the vertebrate lineage, with the Sea Lamprey, in this way making it possible for the development of novel traits based on the importance of claudins during compartmentalization. The paralogs found in lamprey are *cldn 10* and *15*; *cldn 2* and *20*; and *cldn 8* and *B*. The lamprey *cldn B* was renamed *cldn 8b*, shown on the tree as PmCLDN 8B_CLDN B.

The areas marked with red brackets on the tree mark off regions with no lamprey claudins, this suggests that these genes emerged later in higher vertebrates, example is CLDN 4, in region 3, not seen in the lamprey.

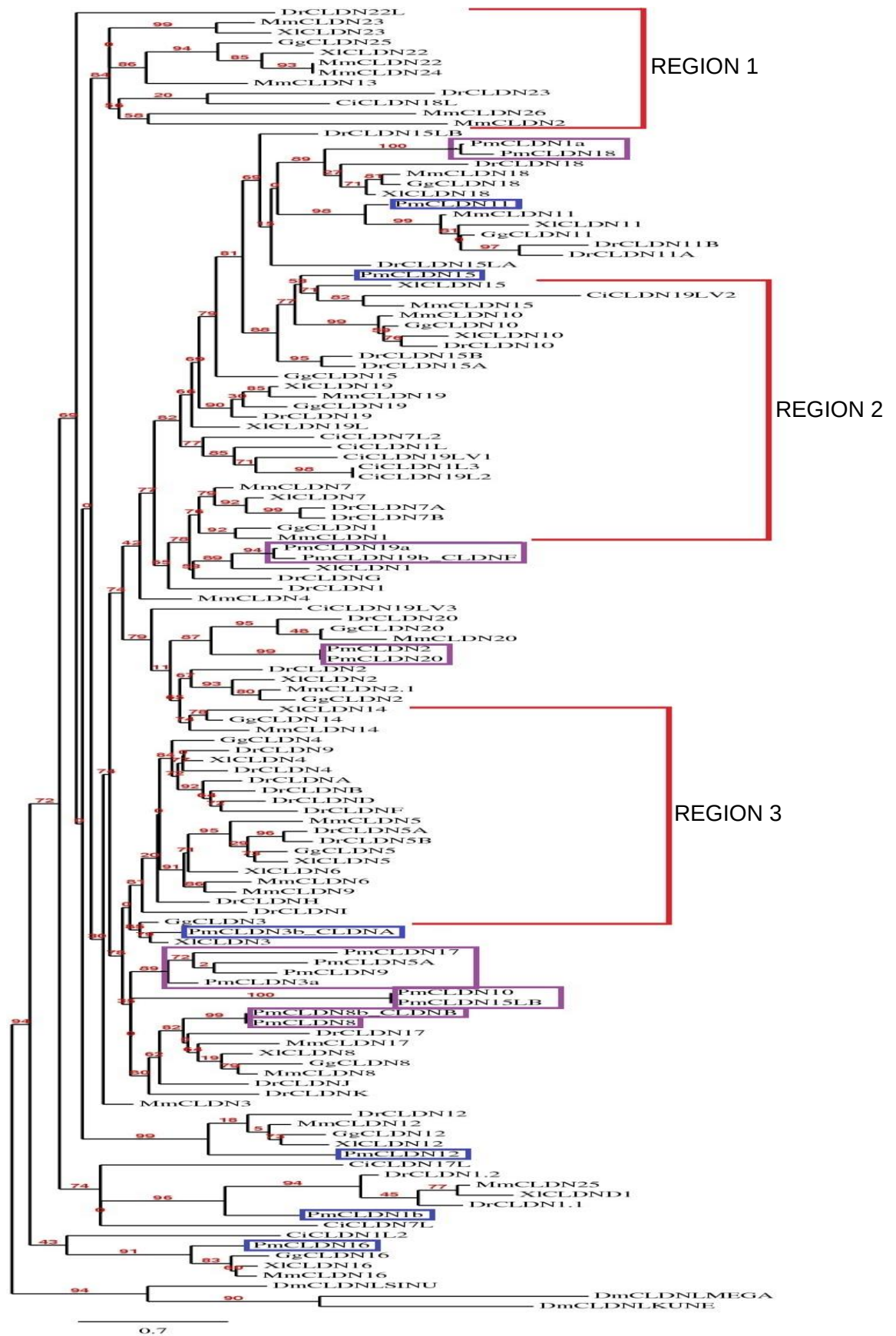


Figure 6: Phylogenetic tree of all Claudins in Model Organism. Amino acid sequences were aligned using the Wits EMBOSS serve tools Muscle and then the tree was constructed on www.phylogeny.fr. Blue and Purple boxes highlight the Sea Lamprey CLDNs with the purple boxes specifically highlighting paralogs. Red brackets highlight areas with no lamprey claudins. Abbreviations: Dm - *Drosophila melanogaster*, Ci - *Ciona intestinalis*, Pm - *Petromyzon marinus*, Dr - *Danio rerio*, Xl - *Xenopus laevis*, Gg - *Gallus gallus*, Mm - *Mus musculus*

3.2.Lamprey Claudin 3b is expressed predominantly in the embryonic ectoderm, the neural tube, ear placodes, pharyngeal arches and the pericardial cavity

RNA *in situ* hybridization using DIG-labelled antisense *cldn 3b* RNA probes was performed on lamprey embryos in order to obtain an expression profile of this gene during development. Figures 7 and 8 show the complete expression profile from stage E4 – E31, seen as blue stain.

Expression of *cldn 3b* is first seen in the ectoderm of embryos at stage E4 (Fig. 7.A, 8.A). This expression pattern continues in stage E5 embryos at which stage the gene is also expressed in the neural folds (Fig. 7.A1-B2). By stage E6 the neural folds fuse to form the neural tube, where *cldn 3b* is seen to be strongly expressed in the neural tube (Fig. 7.C.1-C.2). At this stages expression persist in the ectoderm. By stage E 7 expression begins in the frontal regions of the head as well as persisting in the neural tube. Interestingly, at this stage expression begins to localise at the ectoderm of the developing auditory placodes (Fig. 7.D.1-D.2). Expression becomes more localized in the stomodeum, auditory placodes and the neural tube (Fig. 7.E.1-F.2) Sections through the embryo shows expression in the cranial neural crest by stage E8 (Fig. 8.B’’). This is interesting since these cells will migrate dorsolateral to produce the craniofacial mesenchyme that differentiates into bone, connective tissue, cartilage and cranial neurons. Expression in the stomodeum, ear placodes and the neural tube is maintained at E10 along with expression in the pericardial cavity. Reduced expression is seen in the forebrain (Fig. 7.G.1-G.2; 8.C’’’).

In older embryos *cldn 3b* expression patterns become more consistent. By stage E11 expression is seen in the ear placodes, mouth, forebrain, pericardial cavity and the neural crest component of the developing pharyngeal arches. This pattern is seen through to stage E12 (Fig. 7.H.1-I.2). From E14–E17, expression is seen in the pharyngeal arches. Strong expression can also be seen in the fin fold. Expression also persists in the ear placode (Fig. 7.J.1-M.2). Section through a stage E17 embryo showed that *cldn 3b* expression was specific in the endoderm of the pharyngeal arches at the folds in close

contacts to the ectoderm (Fig. 8.D'). By stage E23, the *cldn 3b* expression isn't as widely spread as seen in earlier stages. Expression is restricted to the ear placode (Fig. 8.N). At E31 expression was seen the eye spot and in the gill arches (Fig. 7.O). Section through a stage E31 embryo revealed a dynamic expression in the areas connecting cartilage to connective tissue of the head and mouth and also in the neural sheath (Fig. 8.E'). Sections through the body showed expression in the dorsal fin, coelom and little in the neural tube (Fig. 8.E''-E''').

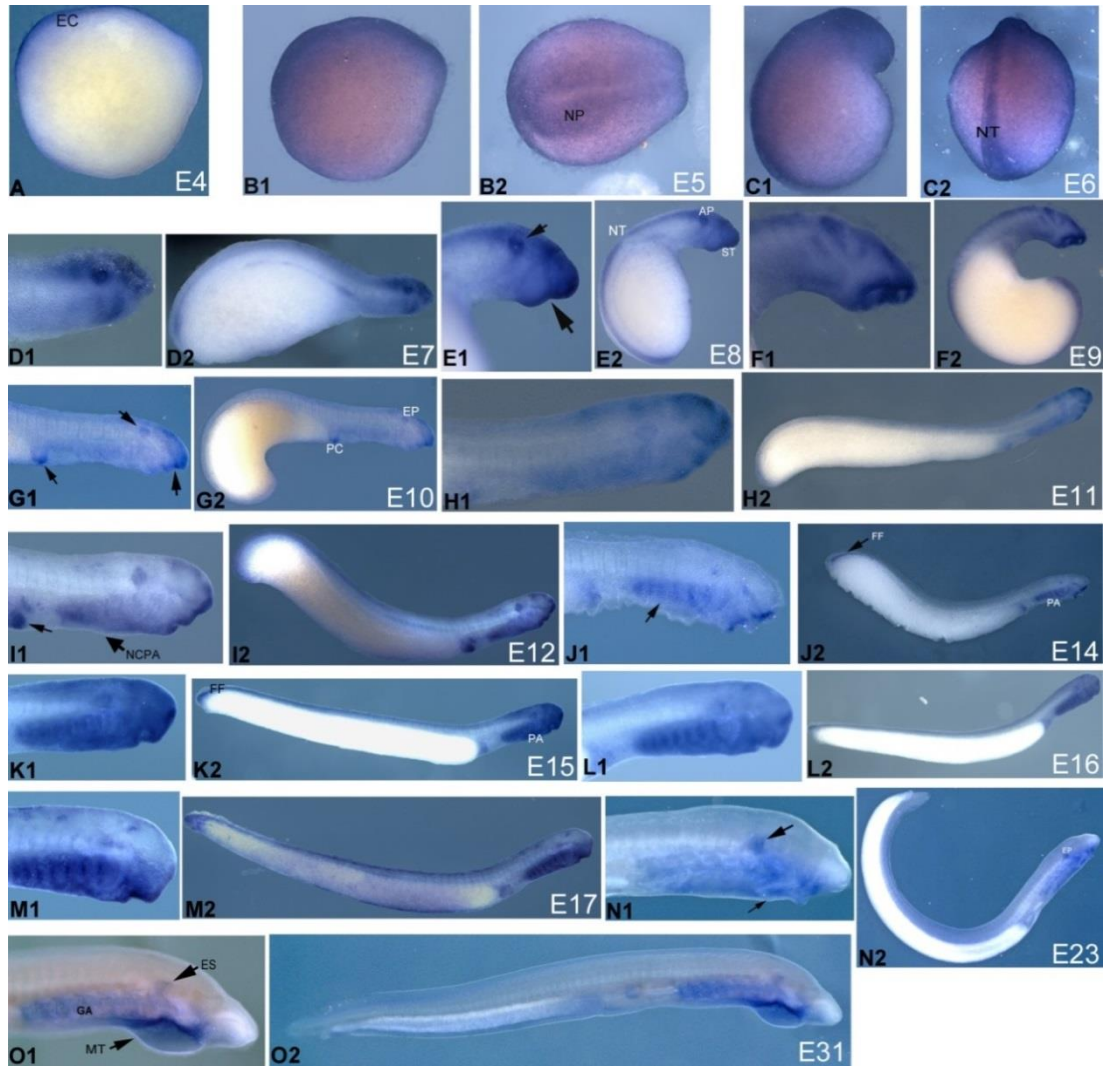


Figure 7: Claudin 3b Expression Pattern During Lamprey Embryogenesis. Expression is seen strongly throughout embryonic development. (A – C) During the early stages of development, expression is seen in the ectoderm and is strongly expressed in the neural plate/tube as development continues. The first image at each stage shows the embryo, lateral view, and the second image shows the embryo, dorsal view. (D – O) The first image shows the head at a higher magnification and the second image shows the whole body of the embryo. (D) Stage E7 strong expression starts in regions of the head including the ectoderm precursor of the auditory placodes. (E) E8 Embryo, arrows showing expression in the emerging stomodeum and the developing auditory placode. Expression is also seen in other cranial neural crest. (F, G) Stage E9 & E10 embryos showing express in the ectoderm, the ear placode, neural tube, mouth and in the pericardial cavity. (H, I) Stage E11 & E12 embryos, strong expression is seen in the forebrain and expression begins in the neural crest at the pharyngeal arches region and the pericardial cavity while continued expression is seen in the mouth and ear placode. (J – M) Stage E14 – E17, embryos indicating strong expression in the pharyngeal pouch developing into the pharyngeal arches. Ear placode, mouth and pericardial cavity also show expression along with the fin fold, FF. (N) Stage E23 shows little expression in ear placode and developing mouth. (O) Little expression in the developing eye spot and gill arches. Strong expression in the mouth. Arrows point to structures with strong expression. EC – ectoderm, NP – neural plate, NT – neural tube, AP – auditory placode, ST – stomodeum, EP – ear placode, PC – pericardial cavity, PA – pharyngeal arches, FF – fin fold, ES – eye spot, MT – mouth, GA – gill arches.

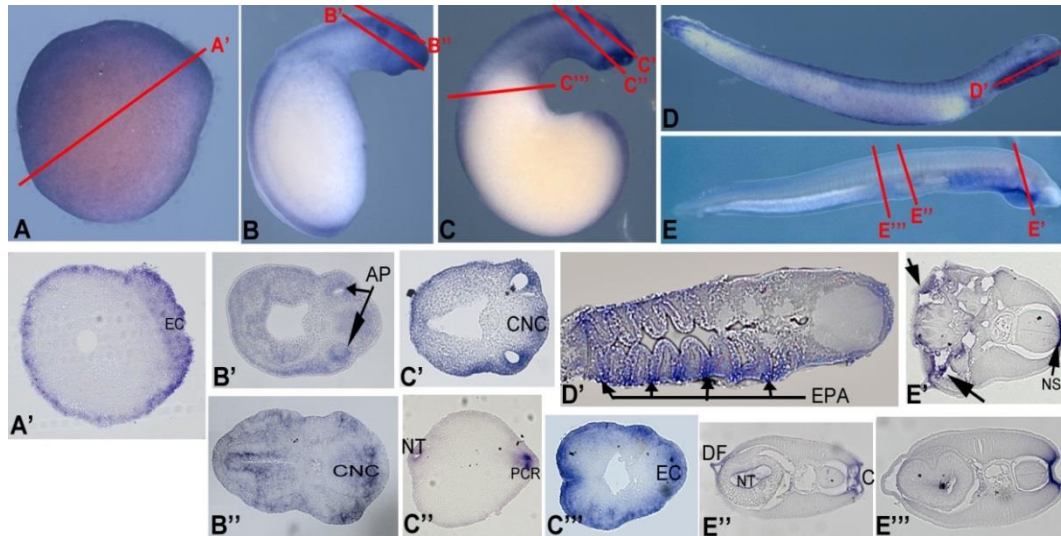


Figure 8: Sections Through the Lamprey Embryos Showing Claudin 3b Expression. Red lines indicate where the sections were taken from. (A) Stage E5 embryo cross sectioned, (A') Showing expression in the ectoderm. Original magnification, 20X (B) Stage E8 embryo sectioned twice through the head (B' & B'') Expression can be seen in the auditory placode and also the cranial neural crest. Original magnification, 20X (C) Stage E9 embryo sectioned through the head and body (C' & C'') Sections through the head showing expression of Cldn 3b in the cranial neural crest, the auditory placode and the surrounding ectoderm. Original magnification, 20X. (C''') Section through the body showing little expression in the ectoderm but strong expression in the neural tube and developing pericardial cavity region. Original magnification, 40X. (D) E17 embryo cross sectioned through the pharyngeal arches. (D') Section through the arches shows strong expression at the endoderm folds of the pharyngeal arches. (E) Stage E31 embryo sectioned through the head and the body (E') Strong expression is seen in the regions connecting cartilage to connective tissue. Expression is also seen in the dorsal neural sheath. (E'' & E''') Expression is in the dorsal fin; the coelom and little in the neural tube. Arrows point to areas with strong expression. Original magnification, 40X. EC – ectoderm, AP – auditory placode, CNC – cranial neural crest, PCR – pericardial cavity region, EPA – endoderm folds of the pharyngeal arches, NS – neural sheath, DF – dorsal fin, C – coelom, NT – neural tube

3.3.Claudin 8b Expression is seen in the embryonic ectoderm, the ear placode and the pharyngeal arches

RNA *in situ* hybridization using DIG-labelled antisense *cldn 8b* RNA probes was performed on lamprey embryos in order to obtain in expression profile of this gene during development. Figure 4 and 5 show the complete expression profile from stage E4 – E31. The preparation and treatment of these embryos was the same as performed during *cldn 3b* *in situ* hybridization.

The expression pattern of *cldn 3b* begins in the ectoderm of young embryos from stage E4 – E6. No expression is seen in the neural tube or the neural folds during these stages (Fig. 9.A.1 – C.2; 10.A'). At later stages expression becomes more restricted in the frontal region of the head and also the ectoderm of the developing auditory placodes (Fig. 9.D.1; 10.B). This expression pattern continues by stage E10 with further expression occurring in the stomodeum (Fig. 9.E.1). By stage E9 expression is limited to the ear placodes, stomodeum and cranial ectoderm. Sections through the body revealed expression in the neural tube at this stage (Fig. 9.F; 10.C' – C''). Expression continues in these structures by stage E10 with the addition of expression in the forebrain and little expression in the neural tube (Fig. 9. G.1.).

In older embryos expression of *cldn 8b* is mostly seen in the mouth and the ear placodes. From stage E11-E14 expression is localized in the ear placode; mouth, forebrain, in neural crest of the developing pharyngeal arches (Fig. 9.H.1 – J.2). The ear placodes; pharyngeal arches and neural tube continue their expression of *cldn 8b* until stage E17 (Fig. 9.K.1 – M.2). Expression in the pharyngeal arches is localised in the endoderm of these structures (Fig. 10.D'). The expression profile is different in much older embryos. Embryos at stage E23 show expression in the ear placodes and little expression in the pharyngeal pouch II (Fig. 9.N.1). By stage E31 expression is seen in the eye spot and little expression in the pharyngeal pouches (Fig. 9.O.1). Sections through the head revealed expression in the regions connecting cartilage to connective tissue of the mouth (Fig. 10.E').

The expression profile of *cldn 8b* is quite dynamic and appears to have many overlapping regions to that of *cldn 3b*, however the expression profile is quite different and in most cases *cldn 3b* seems to have a much broader range of expression during development. A comparative view of these two expressions at different stage of development can be seen in Table 9, which summarises the two profiles. Structures expressing both genes include

the vertebrate specific traits, the ear placodes and the pharyngeal arches. Both strongly expressed between stages E14 – E17.

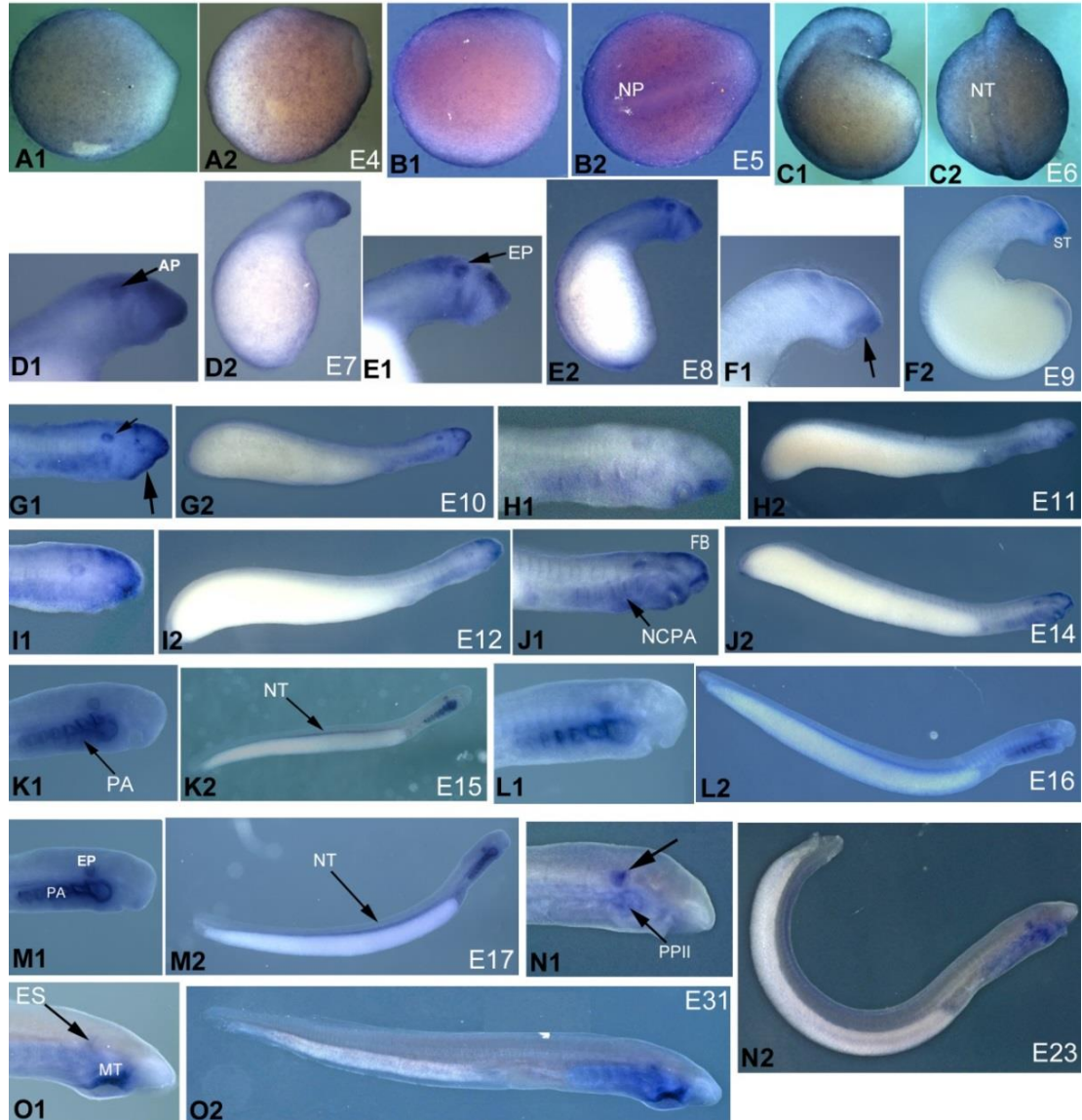


Figure 9: Claudin 8B Expression During Lamprey Embryonic Development. (A – C) Stage E4 – E6 shows expression in the ectoderm. (D, E) Expression starts very strong in the developing auditory placode and frontal regions of the head. (F) At stage E9 expression begins in the stomodeum and persists in the ear placode. (G – J) Expression in the frontal regions of the head at stage E10 – E14. Later expression begins in the neural crest of the developing pharyngeal arches. (K – L) Stage E15 – E16 show strong expression throughout the pharyngeal arches. No expression is seen in the forebrain or the mouth. (M) Stage E17 shows expression in the ear placode, pharyngeal arches and expression in the neural tube. (N) Embryo at stage E23 shows decreased expression in the arches but strong expression is still seen in the pharyngeal pouch II, Expression also persists in the ear placode. (O) At stage E31 little expression is in the eye spot; ES, but strong in the mouth, MT. Arrows point to structures with strong expression. EC – ectoderm, EP – ear placode, ST – stomodeum, EP – ear placode, NCPA – neural crest of the developing pharyngeal arches, NT – neural tube, PA – pharyngeal arches, PPII –pharyngeal pouch II.

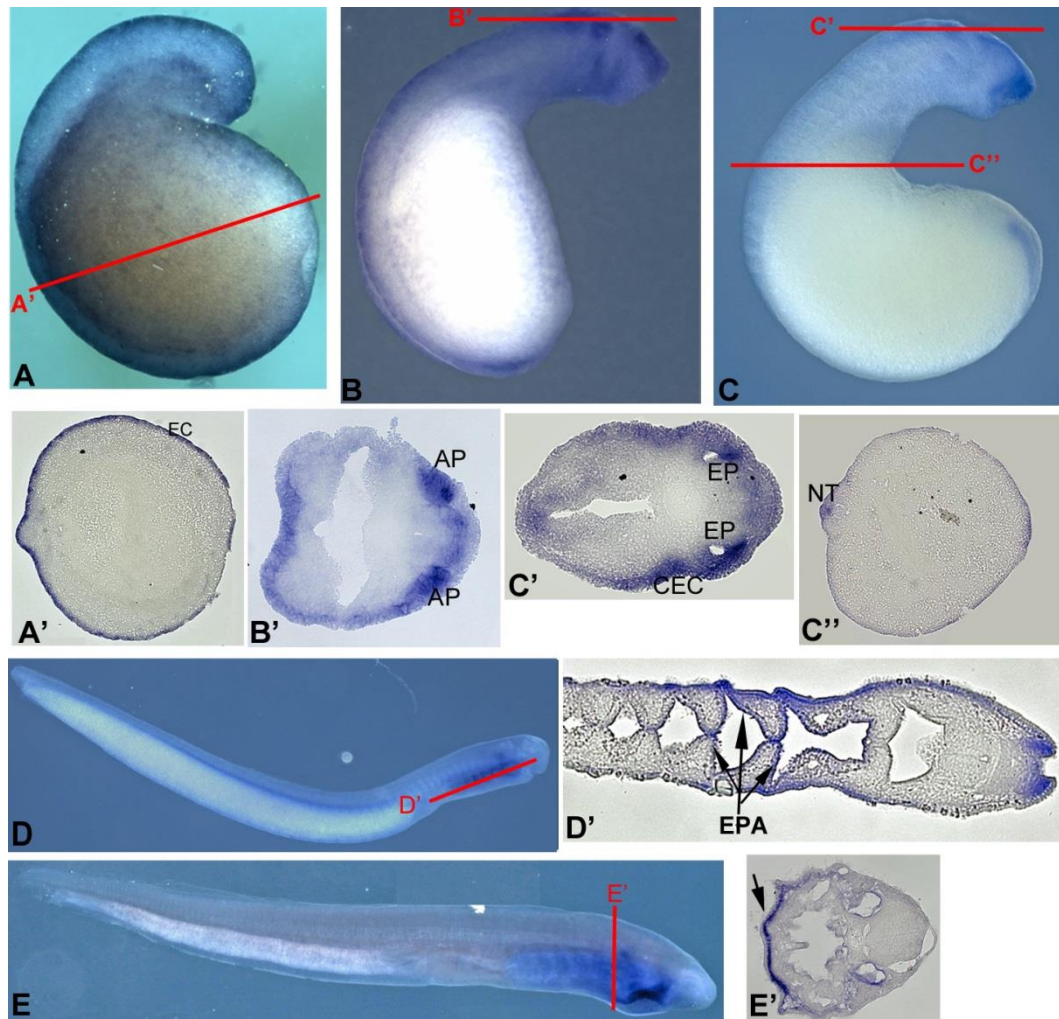


Figure 10: Sections Through the Lamprey Embryos Showing Claudin 8b Expression. Red lines indicate where the sections were taken from. (A') Section through the stage E6 lamprey embryo showing expression in the ectoderm. Original magnification, 20X. (B') Section through the head of an embryo at stage E9 shows expression in the auditory placode and the surrounding ectoderm of the head. Original magnification, 40X. (C) Stage E10 embryo sectioned through the head and the body. (C') Section through the head showing expression in the ear placode and cranial ectoderm. Original magnification, 40X. (C'') Section through the body of the embryo shows little expression in the neural tube and ectoderm. Original magnification, 20X. (D') Section through the pharyngeal arches of an embryo at stage E16. Original magnification, 20X. Expression is seen in the endoderm of the arches. (E') Section through the head of an E31 embryo. Original magnification, 40X. Expression is in areas connecting cartilage to connective tissue of the mouth. Arrows indicate areas with strong expression. AP – auditory placode, CEC – cranial ectoderm, EP – ear placode, EPA – endoderm of the arches, EC – ectoderm. NT – neural tube, PA – pharyngeal arches.

Table 9: Summary of claudin 3b and claudin 8b expression profile during lamprey embryogenesis

Embryonic Structure	<i>Claudin 3B</i>	<i>Claudin 8B</i>
Ectoderm	E4 – E9	E4 – E10
Neural Plate	E5	
Neural Tube	E6 – 9; E31	E10 – E17
Ectoderm of the developing auditory placode	E7	E7
Forebrain	E7 – E12	E7 – E14
Cranial neural crest	E8	E8 – E9
Stomodeum	E8 – E9	E8 – E9
Ear placode	E8 – E23	E8 – E23
Pericardial cavity	E10 – E17	
Mouth	E10 – E23	E10 – E14
Neural Crest (Developing Pharyngeal Arches)	E11 – E12	E11 – E12
Pharyngeal arches	E14 – E17	E14 – E17
Fin Fold	E15 – E17	
Pharyngeal Pouch II		E23
Cartilage (In Head)	E31	E31
Coelom	E31	
Connective Tissue (In Head)	E31	E31
Dorsal Fin	E31	
Eyespot	E31	E31
Gill arches	E31	E31
Neural Sheath	E31	

3.4.Claudin 3 is involved in the neural tube, ear placode and pharyngeal arch development

Knockout experiments were performed using morpholinos which are antisense fragments of RNA analogue used to bind to complementary mRNA sequences. These oligomers block access of other molecules to specific sequences. The new complex would therefore prevent the mRNA from binding to the ribosome and being translated thus down regulating the protein synthesis.

Morphological effects of Cldn3bMO on the development of neural tube, pharyngeal arches and the ear placode were examined with markers for these specific structures being used during *in situ* hybridization. Cldn 3b showed expression in the neural tube, pharyngeal arches and ear placode at different stages of development thus the morphants were collected at different stages. The markers used were specific to the structures at the explicit stage of the morphants as such, SoxE1 (marker for the ear placode and pharyngeal at stage E16 [Fig. 10.A]), Tbx1 (ear placode marker, stages E8 – E10 [Fig. 15.A]) and Sox2 (neural tube marker at stage E6 [Fig. 20.A]).

Early expression of *cldn 3b* is seen in the neural tube along with some expression in the ectoderm. Morphological effects of Cldn3bMO on embryos injected at the whole cell stage, were seen in the neural tube as early as in stage E6. Morphants showed abnormalities in the development of the neural tube, specifically the sealing off of the neural folds during neurulation (Fig. 20.B). As a result of *cldn 3b* deficiency, the morphants also showed defects in the integrity of the ectoderm as well as abnormalities in the overall embryo morphology at this stage (Fig. 19.B). Neural tube defects were not seen in Cldn19b knockout morphants, which showed wild type phenotype in the neural tube (Fig. 20.C).

Cldn3bMO was seen to greatly effect ear placode development even at a later stage. Embryos injected at the whole cell stage of development showed abnormalities in the development of the ear placode also at stage E8. At this stage 50% (5/10) of the embryos showed poor development of the ear placode. Most of the abnormal embryos (3/5) were severely affected by the lack of *cldn 3b* and completely lacked ear placode structures (Fig. 15.C, Fig. 16). The remaining 2 embryos showed defects in the expansion of the ear placode (Fig. 15.B). Half of embryos collected showed wild type phenotype in the ear placode at this stage suggesting that *cldn 3b* isn't the only gene essential for the development of ear placode (Fig. 16). The same phenotypes were seen in embryos

injected with 0.5mM Cldn3bMO at a two cell stage. The morphants exhibited defects in the ear placode formation on the side that was injected. No placodes were seen in 4 of the 8 embryos collected. Additionally, 1 showed reduced placode on the side that was injected (Fig. 17.B-C).

During the later stages of development, *cldn 3b* shows strong expression in the endoderm of the pharyngeal arches as well as the ear placode, two features which are unique to the vertebrate lineage. After the injection of Cldn3bMO in embryos at a whole cell stage, the morphants were allowed to develop to stage E16.

The expression profile suggests that *cldn 3b* may play a role in the development of these traits, thus a knockout mutant would have poorly developed ear placode or pharyngeal arches. Three distinguishing phenotypes were identified in the morphants injected with the Cldn3bMO, remarkably all phenotypes involved the formation of the ear placode and the pharyngeal arches (Fig. 11).

The first phenotype showed defects in both traits, with the embryos showing retarded growth in the development of the pharyngeal arches and the ear placode (Fig. 11.D). When compared to the wild type phenotype (Fig. 11.A) both structures appear to be small and depressed for stage E16, suggesting that *cldn 3b* may be involved in the expansion of these structures.

Fig. 11.E shows the second phenotype seen in the *cldn3b* knockout morphants. Out of the 83 morphants (including the morphants showing defects in ear placode and pharyngeal arches), 14 morphants failed to show normal development of the ear placode only. In contrast the pharyngeal arches showed normal phenotype (Fig. 11.A). In total 27 out of the 83 morphants showed defects in the ear placode, that's 32% of the morphants (Fig. 13).

Final phenotype showed defects in the development of the pharyngeal arches. Injection of Cldn3bMO resulted in arches which showed incomplete sealing of individual compartments of these structures (Fig. 11.F, Fig. 12.B). Thus suggesting that *cldn 3b* may be important in sealing off the pouches during development of the arches. From 83 morphants, 38 morphants showed defects only in the arches with normal phenotype shown in the resulting ear placode. An additional 13 embryos showed defects in the arches as well as the ear placode, thus in total 51 out of 83 embryos showed defects in the pharyngeal arches (Fig. 14). The difference in phenotype may be related to the concentration of the morpholino incorporated in the embryos. Overall 78.3 % (65/83) of

the embryos showed defects in the ear placode and/or pharyngeal arches when injected with Cldn3bMO.

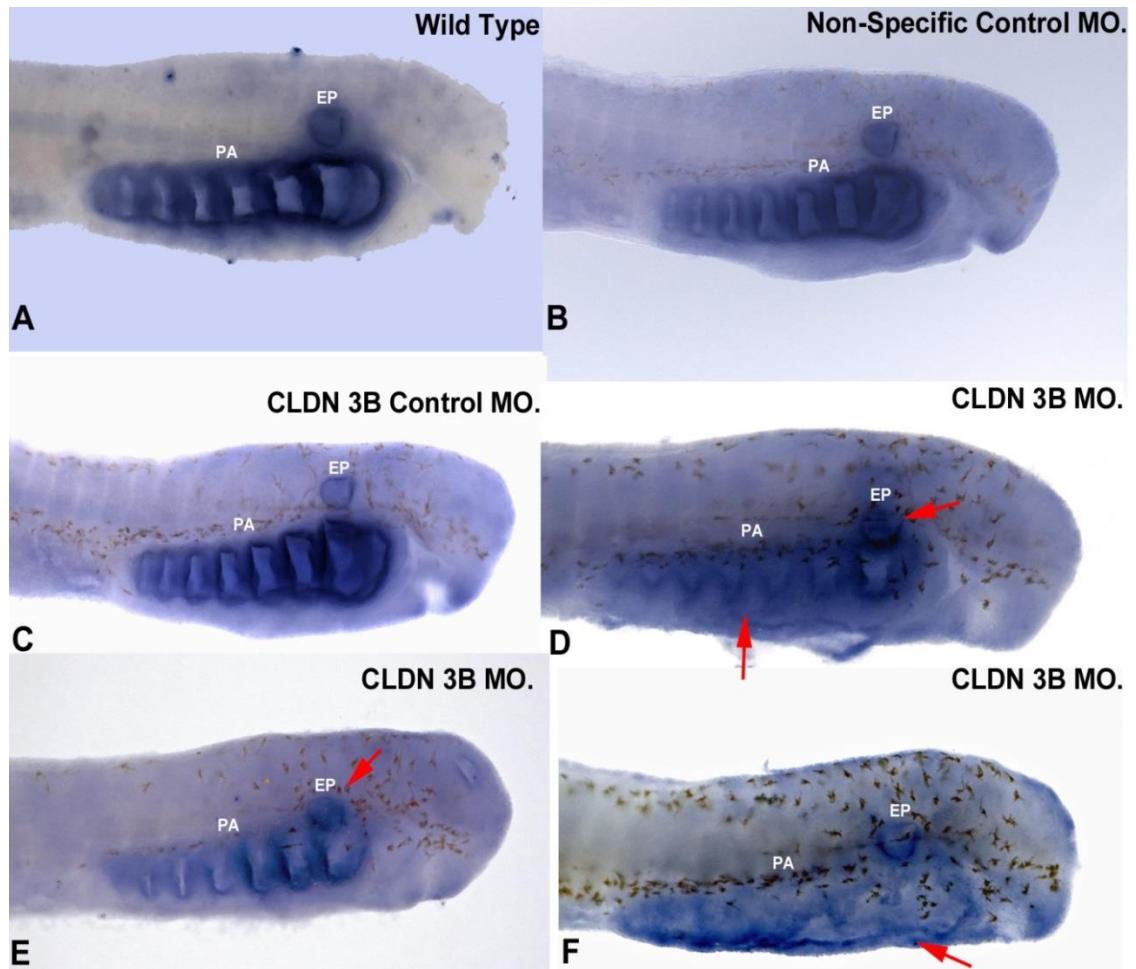


Figure 11: Knockdown of Cldn 3b Results In Abnormal Development of Ear Placodes and Pharyngeal Arches. Whole mount *in situ* hybridization performed on stage E16 Cldn 3b knockout morphants using the probe designed from the gene SoxE1. (A) Wild type embryo showing normal ear placode and pharyngeal arches development. (B) Non-specific control morpholino injected embryo showing normal EP and PA development. (C) Cldn 3b Control morpholino injected embryo showing normal growth of the placode and the arches. (D – F) Cldn 3b knockout morphants showing abnormalities in the development of the pharyngeal arches and the ear placode. D. Retardation of the expansion of the arches and the ear placode indicated by the red arrows. As a result of decreased or loss of Cldn 3b. E. Poor development of the ear placode shown by the red arrow and normal pharyngeal arches development. F. Incomplete sealing off of the first two pharyngeal arches shown by the red arrow. Normal development of the ear placode. EP – ear placode, PA – pharyngeal arches.

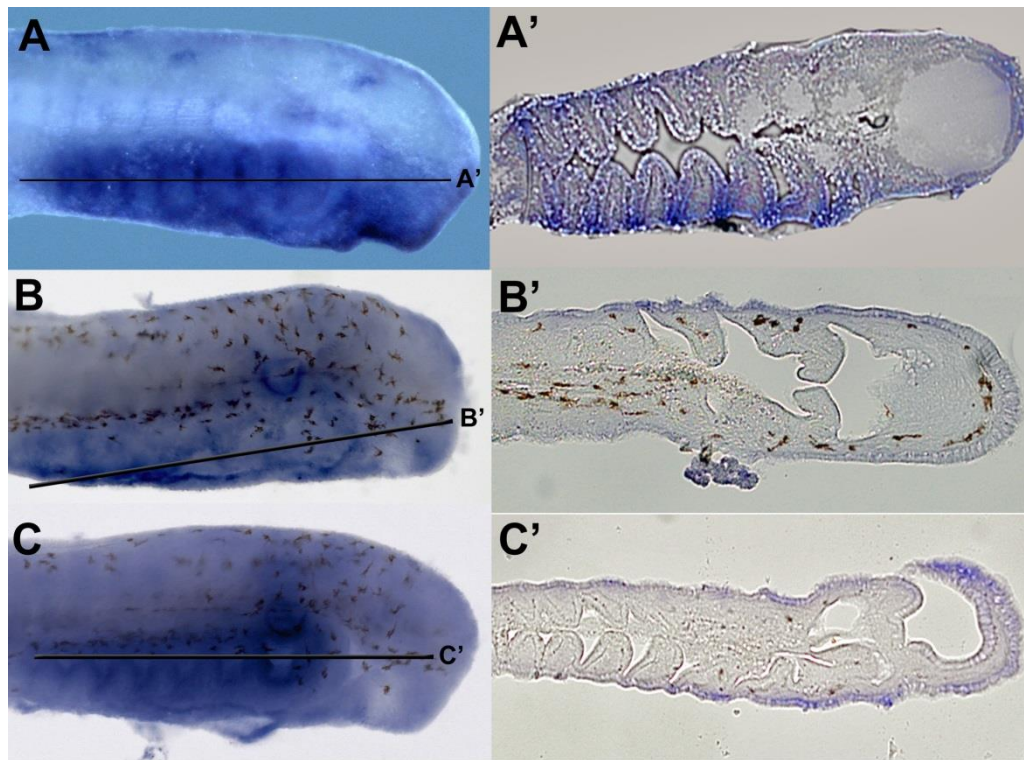


Figure 12: Sections Through Stage E16 Claudin 3B Knockout Morphants Post *in situ* Hybridization. (A) Wild type pharyngeal arches (A') Section through the pharyngeal arches. Original magnification, 20X (B) Morphant with abnormal development of the pharyngeal arches. (B') Section through the pharyngeal arches showing abnormal sealing. Original magnification, 20X (C) Morphant with abnormal development of the ear placode and pharyngeal arches. (C') Section through the pharyngeal arches showing retarded growth. Original magnification, 20X

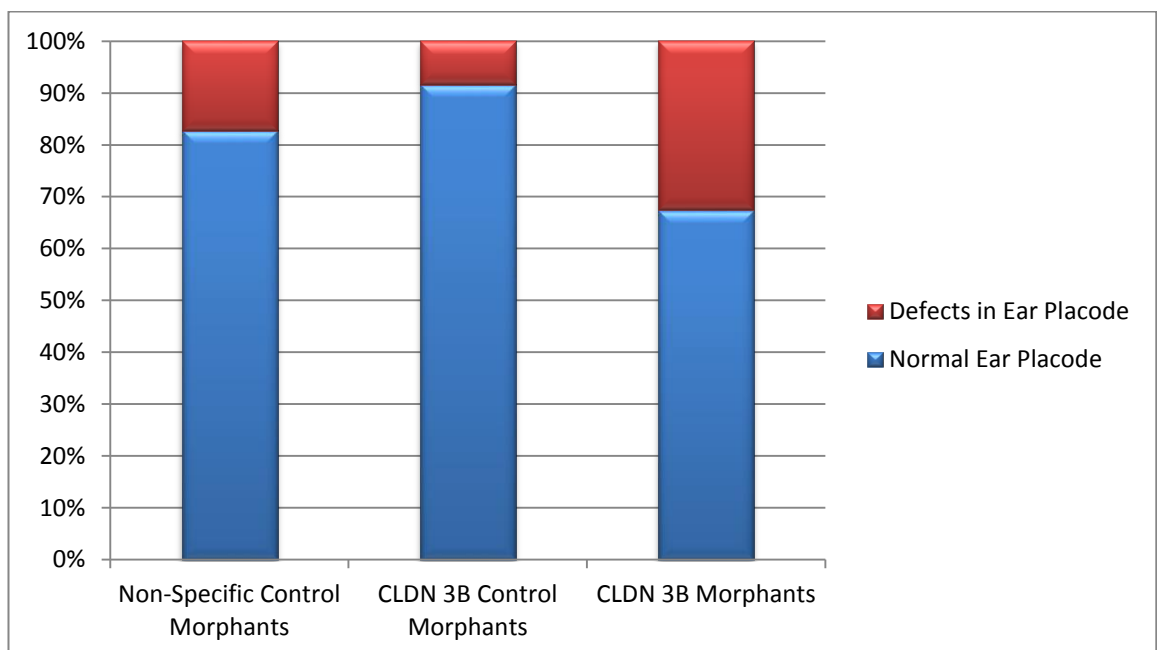


Figure 13: Bar Graph Showing the Percentage of Embryos With Abnormalities In the Ear Placode at Stage E16. Post knockout experiments using the Cldn 3b morpholino and control morpholinos. Knockdown of Cldn 3b has affected the development of ear placodes in 32.5% of the embryos injected with Cldn3bMO.

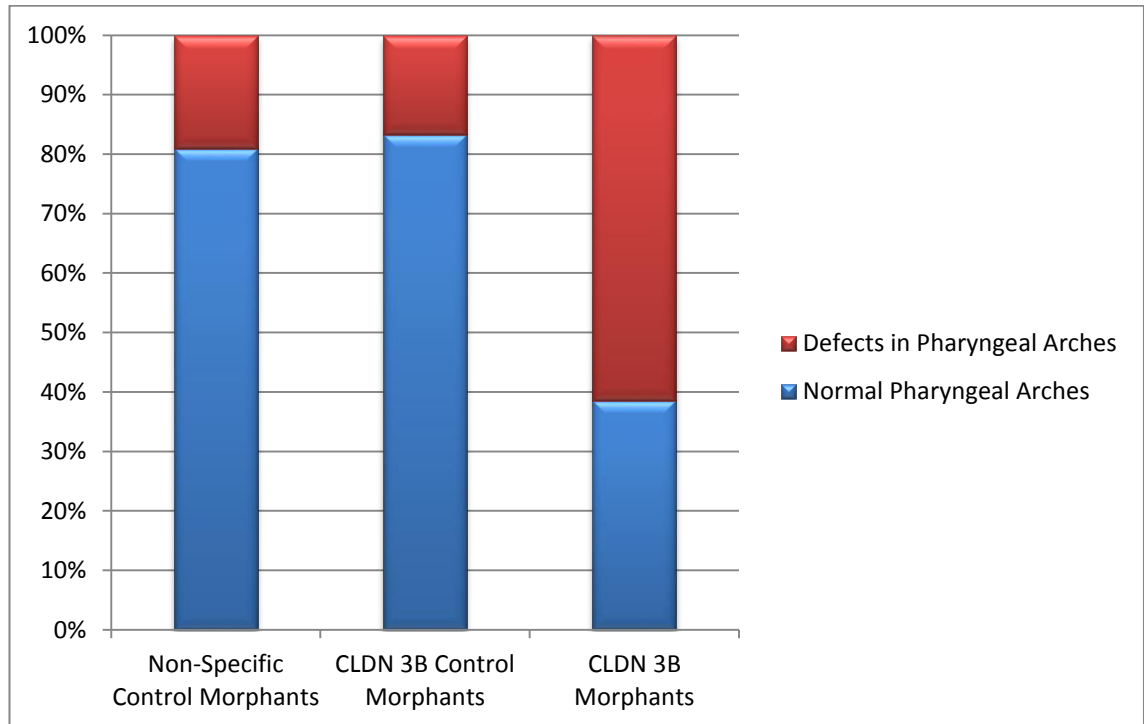


Figure 14: Bar graph showing the percentage of embryos with abnormalities in the pharyngeal arches at stage E16. Post knockout experiments using the Cldn 3B morpholino and control morpholinos. Knockdown of Cldn 3b has affected the development of pharyngeal arches in 61.4% of the embryos injected with Cldn 3B morpholino.

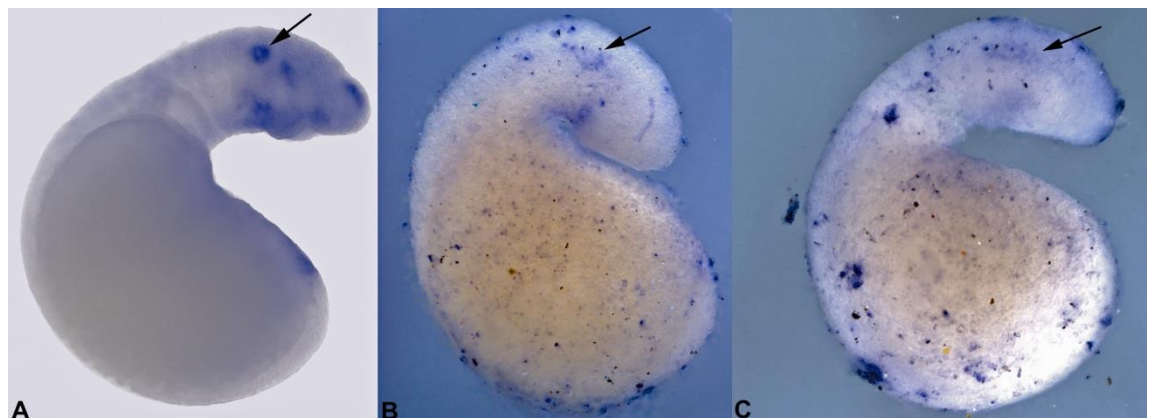


Figure 15: Whole Mount *in situ* Hybridization on Stage E8 Claudin 3b Knockout Morphants. The probes designed from Tbx1 were used as it is expressed in the ear placode . (A) Wild-type embryo showing normal development of the ear placode. (B) Cldn 3b Knockout morphant showing poor development of the ear placode. (C) Cldn 3b knockout morphant severely affected by the lack of Cldn 3b and as a result showing no placode development. Arrows indicating the ear placode.

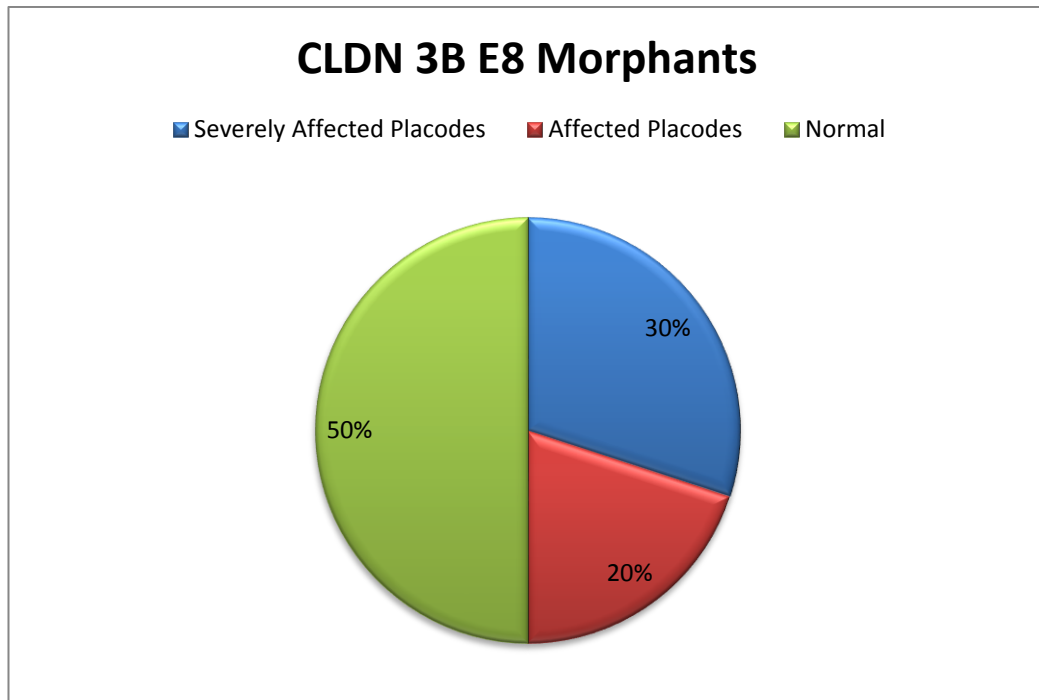


Figure 16: Pie Chart Showing the Ratio of Embryos Affected By the Knockout of Claudin 3B at Stage E8. 50% of the embryos showed the same phenotype as the wild type. 20% showed abnormalities in the ear placodes and 30% were severely affected in the ear placodes due to the lack of Cldn 3b.

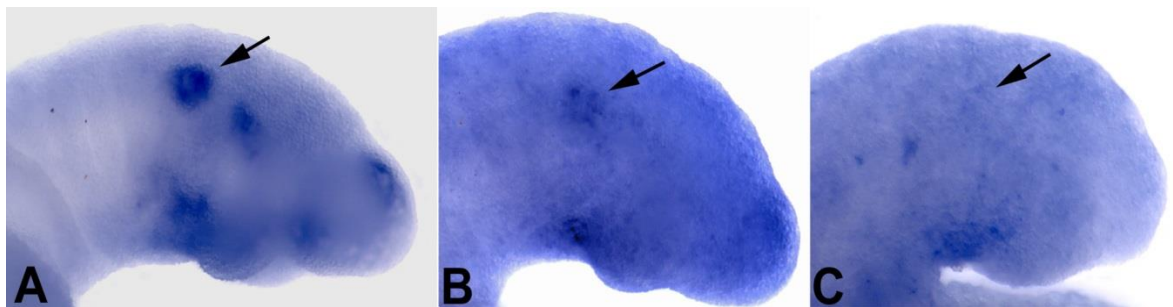


Figure 17: Stage E8 Embryos Injected with 0.5mM Claudin 3B Morpholino at the Two Cell Stage. E8 embryos post *in situ* hybridization using Tbx1. Arrows are showing the ear placode. (A) Wild type embryo. (B) Morphants showing reduced ear placode on the side that was injected with the Cldn 3b morpholino. (C) Morphants severely affected showing no placodes on the side injected with the morpholino.

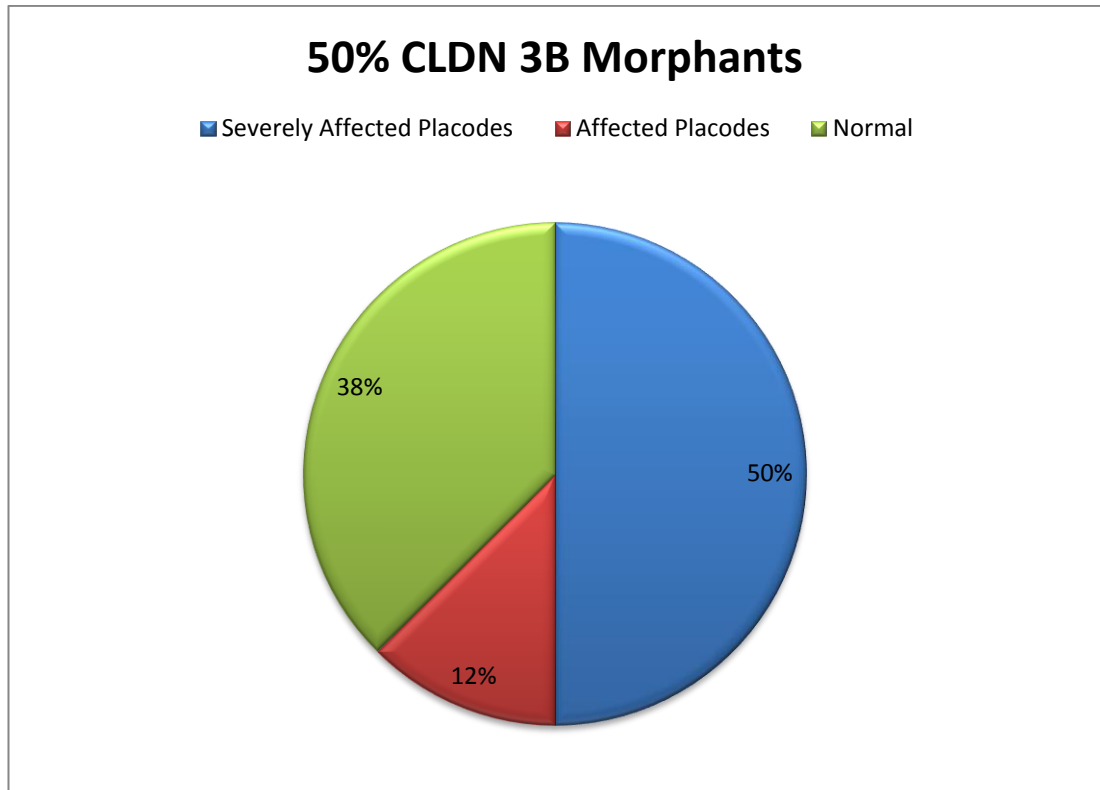


Figure 18: Pie Chart Showing the Ratio of E8 Embryos Affected by 50% Claudin 3B Morpholino. 12% of the embryos were affected in the ear placode and a further 50% of the embryos were severely affected in the ear placode due to the lack of Cldn 3b. 38% showed wild type phenotype.

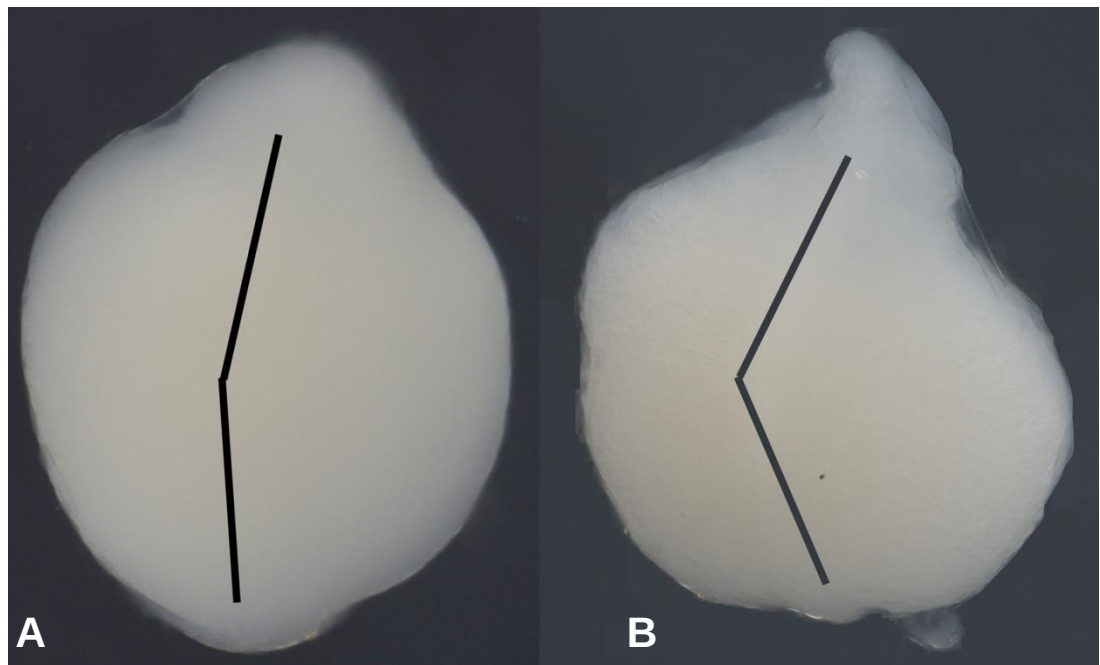


Figure 19: Comparison Between Wild type Embryo and Claudin 3B Morphant at Stage E6. (A) Posterior view of a wild type showing normal shape of the neural tube. (B) Cldn 3b morphant has an abnormal body morphology and irregular neural tube shape.

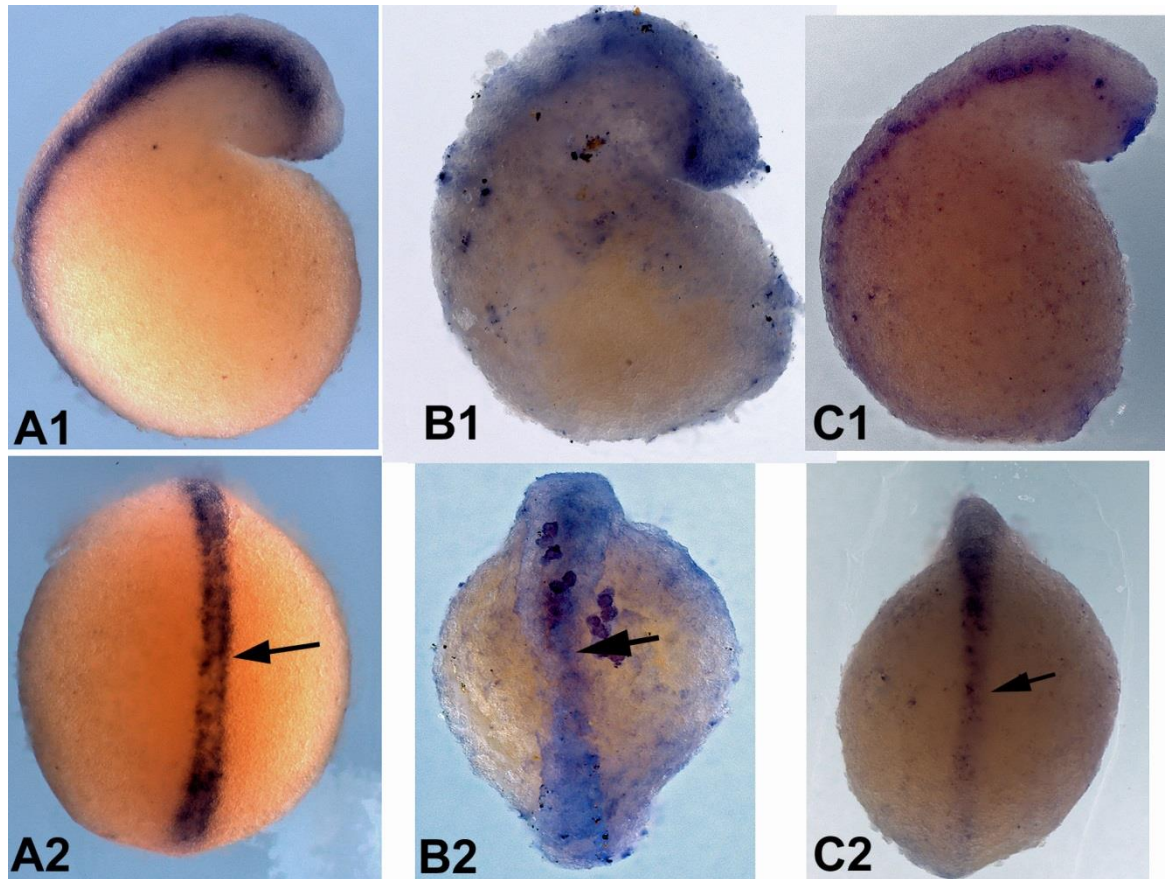


Figure 20: Claudin 3B Knockout Morphants at stage E6. The first image is the lateral view of the embryo and the second image below is the posterior view of the same embryo. (A) Control wild type (B) Cldn 3b morphant (C) Cldn 19f morphant; post *in situ* hybridization using Sox-2 probe as a marker. The arrows are indicating the Neural tube.

3.5.Claudin 8 is important for ear placode development

The ear placode is an important trait since it is unique to the vertebrate lineage. Since this gene showed strong expression in most stages of embryonic development these structures were made the focus of the experiment. Knockout experiments were performed using Cldn8bMO.

Embryos that were injected with Cldn8bMO showed a high rate of mortality at E6. This data, along with the expression profile data, suggested that the deficiency of this gene in early embryos leads to a fragile ectoderm, implicating the importance of the gene in the ectoderm stability. Reducing the concentration of the morpholino injected into the embryos increase the ration of embryos surviving past this stage.

Fig. 21 shows morphants at stage E8 post *in situ* hybridization using *Tbx-1* as the marker. These morphants were injected at the two cell stage of development. The morphological defects seen in theses morphants were different for the side of the embryo injected with Cldn8bMO versus the other side not injected with the morpholino. The side that was not injected with the morpholino showed a range of phenotype in the ear placode. The morphants which showed a strong phenotype, had no placode on either side possibly due to leakage of the morpholino into the side that was not directly injected. Embryos which showed weak phenotype, were the ones with reduced or absent ear placodes only on the side directly injected with Cldn8bMO (Fig. 21.B-D). Most of the morphants (31 of 64) showed the weak phenotype with 8 out of 64 morphants showing the strong phenotype (Fig. 22). The remaining 25 embryos showed normal development of the ear placode on both sides of the embryo.

The standard control morpholino was injected at the whole cell stage of injection. Only a few embryos were viable for this particular stage and less than 16% (2 / 13) of these embryos were successfully stained however those that were stained showed normal placode development (Fig. 23). This data suggest that there was a lot of degradation of the RNA during the storage period, however the control appears to not affect the development of the ear placode at this stage.

At stage E10, *cldn 8b* is expressed strongly in the ear placode (Fig. 9.G). Embryos injected with Cldn8bMO at the whole cell stage of development showed strict phenotypes at this later stage of embryogenesis.

Few morphants collected at stage E10 completely lacked placodes. This phenotype was seen in 2 of 6 embryos. The remaining 4 embryos showed normal development of the ear placode (Fig. 25). This result is expected at an older stage when the morpholino gets diluted.

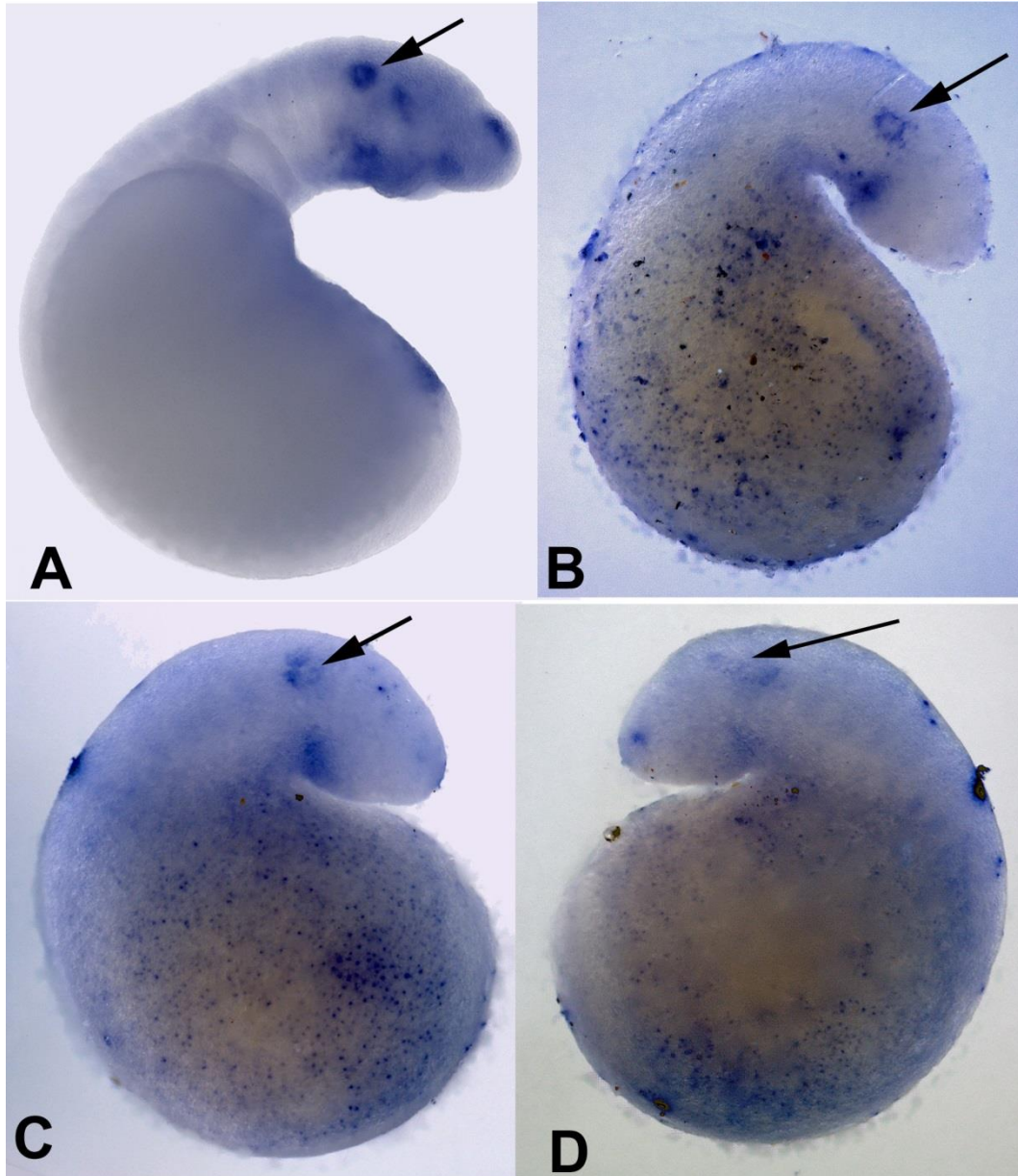


Figure 21: Stage E8 Claudin 8B Knockout Morphants Injected at Two Cell Stage. Embryos were stained using the Tbx1 probe. Arrows indicate the ear placode. (A) Wild type. (B – C) One side of the embryos showed strong or weak expression however there was normal development of the ear placode. (D) Same embryo as (C), left side. Side of the embryo which was injected with the morpholino resulted in a lack of ear placode.

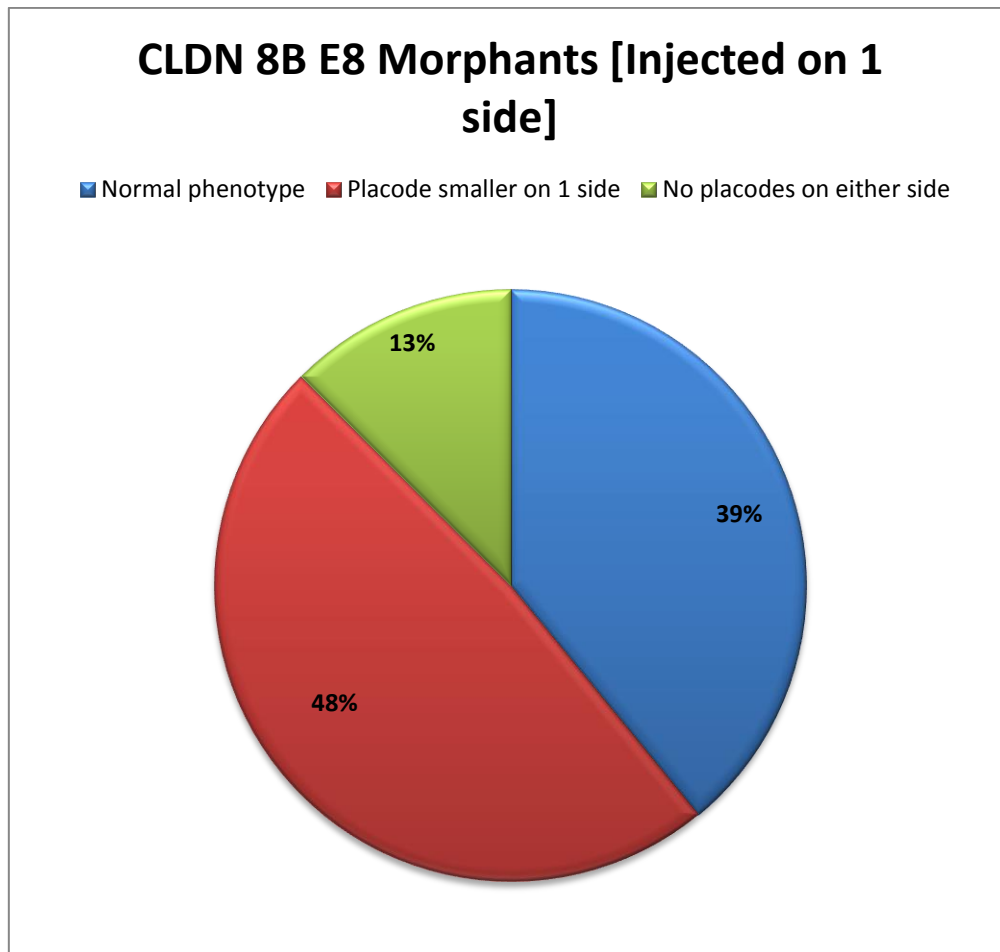


Figure 22: Pie Chart Showing Results From the Claudin 8B Knockouts on One side of the Embryos. One side showed weak ear placode development while the other showed normal developments. 39% of the embryos showed wild type phenotype. 13% showed no placodes on either side of the embryo. 48% were affected on the side injected with the morpholino, showing reduced placodes on 1 side.

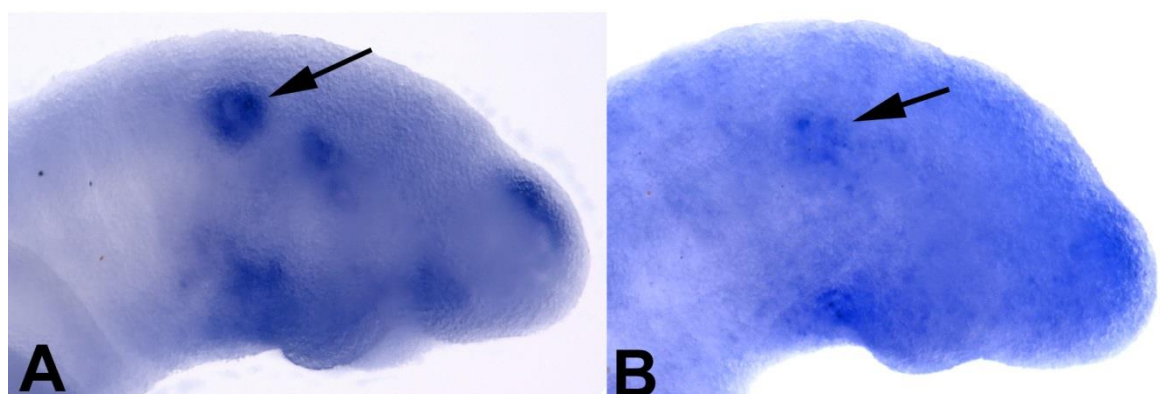


Figure 23: Whole Mount *in situ* Hybridization Using Tbx1 on Claudin 8b Control Morpholino Injected Embryos at Stage E8. Arrows indicate the ear placode. (A) Wild type (B) Standard control morpholino injected embryo showing normal ear placode. Ear placodes are indicated by the black arrows

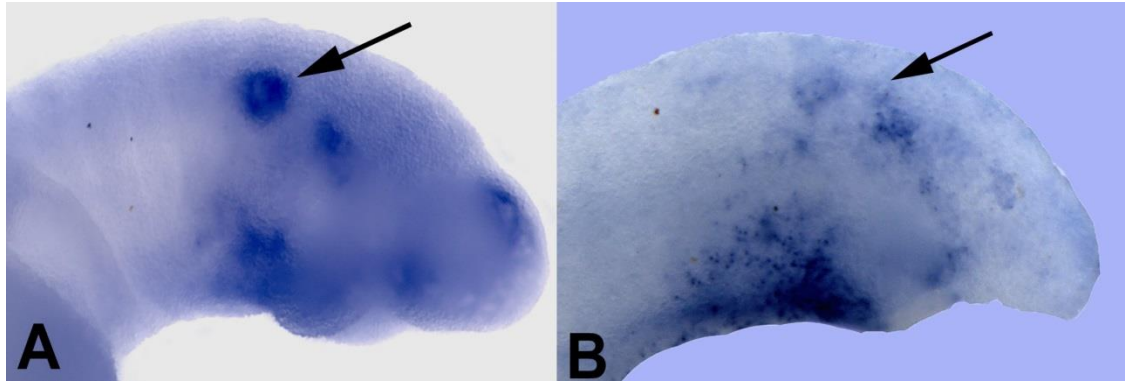


Figure 24: Whole Mount *in situ* Hybridization Using Tbx1 on Claudin 8b Knockout Morphants at Stage E10. Arrows indicate the ear placode. (A) Wild type (B) Cldn 8B knockout morphant showing lack of ear placode.

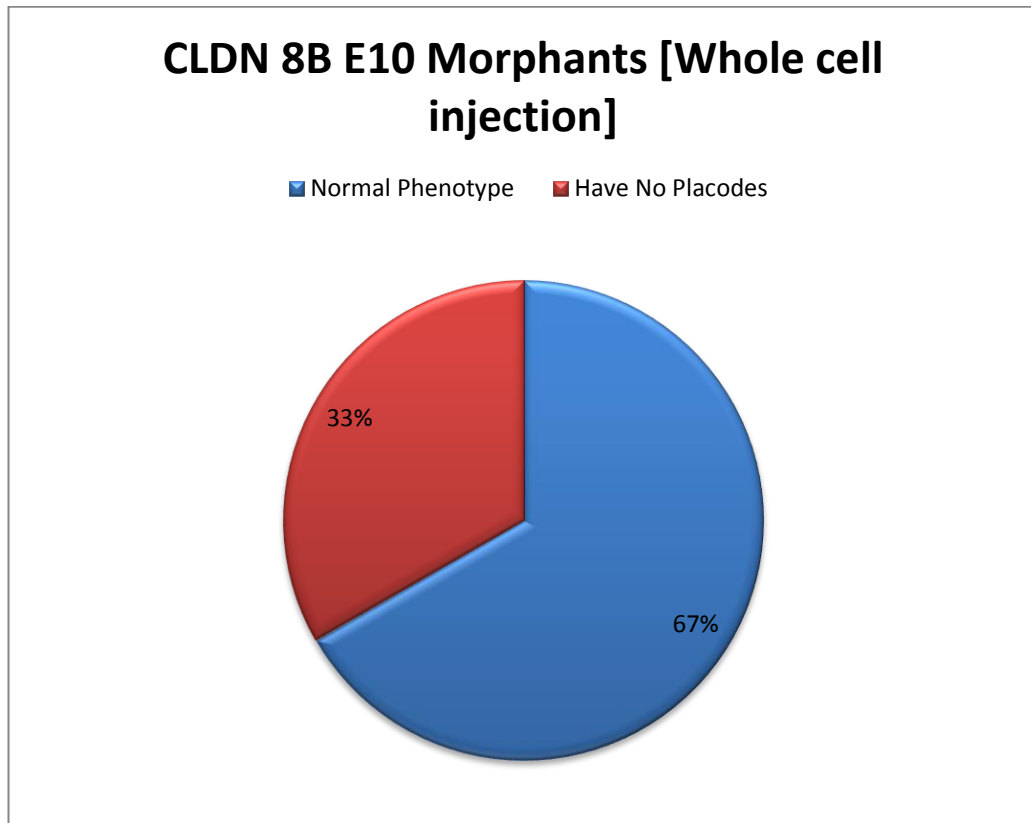


Figure 25: Pie Chart Showing the Ratio of Embryos at Stage E10 Affected By Cldn 8b Knockout. 67% showed wild type phenotypes and the rest were severely affected and developed no placodes.

CHAPTER FOUR – DISCUSSION AND CONCLUSION

4.1. Insight from the phylogenetic tree of evolution of claudins in vertebrates

4.1.1. Lamprey Claudin Gene Family

The initial six fold screening of the lamprey embryonic cDNA library yielded six possible lamprey embryonic CLDNs, initially labelled *cldn* A – F. Subsequent BLAST searches, using the NCBI database confirmed the identities of three, as actual lamprey CLDNs; *cldn* A, B and F (see Appendix B). This not only confirmed the identities of the genes but also aligned DNA sequence with protein sequences thereby the output similarities was in part based on functional relationship. The DNA STAR lasergene software was used to identify the ORF of each of these genes so that the amino acid sequence could also be determined using the online tool ExPASy Translation tool (Results not shown) and thus could be used in the phylogenetic tree.

RNA sequencing data used by Jeramiah Smith, to build up the lamprey genome was obtained and used to find more lamprey *cldn* sequences. The coding sequence (CDS) of the genes were determined for each gene and where necessary the sequences were reverse-complemented so that the CDS were in the correct direction.

Pooling together the two dataset revealed that *cldn* A was the same as the designated *cldn* 3b found in the RNA sequence of adult lamprey, this suggest that the gene is not only expressed during embryonic development but as previously stated with CLDNs, this gene is expressed also in adult lamprey. Interestingly *cldn* B and *cldn* F were identified only in the screening suggesting that these genes may be more specific for developmental purposes and is mostly expressed during embryonic development.

This goes to show how even in one of the earliest vertebrates, *cldns* played a role in the embryonic morphogenesis based on the *cldn* expression profile of the organism and the reconstructive limitations on the epithelia covering the surfaces of the body. This was made possible by the two whole genome duplication events which allowed for the initial expansion on the genes, including the CLDN gene family consisting of 20 members.

As previously mentioned, claudins predate vertebrates, evident in the expression of *cldns* and CLDN-like genes in the invertebrates *C. elegans*; *D. melanogaster* and *C. intestinalis*. The *C.elegans*, belong to the Nematoda phylum and have been reported to express 5 CLDN-related genes. Three CLDN-like genes have been identified in the

D.melanogaster, which is part of the Arthropoda phylum which in itself is more closely related to the vertebrate phylum, Chordata. The invertebrate chordate, *C.intestinalis* has 11 true *cldn* members which showed great similarities in their amino acid sequences to CLDNs in the vertebrate lineage. These results support the theory of an expansion of the claudin gene family with the rise of more complex organisms, with 20 genes identified in the sea lamprey, of which *CLDN 4* is not present. As previously mentioned Claudin gene family size in different organisms vary and table 10 has integrated our findings of Lamprey gene family size with that from other organism data in the NCBI database.

Table 10: The Size of the Claudin Gene Family in Different Organisms, Including *P.marinus*

Organism	Claudin Gene Family Size
<i>H.sapiens</i>	27
<i>M. musculus</i>	23
<i>G.gallus</i>	19
<i>X.laevis</i>	20
<i>D.rerio</i>	33
<i>P.marinus</i>	20
<i>C.intestinalis</i>	11
<i>D.melanogaster</i>	3

Out of the 20 genes identified as claudins in the lamprey genome, 15 of them belong to the Classic type of claudin. By definition they show high levels of conservation with *cldn 1*. This is signifying an early expansion with the variety between the duplicates remaining at a low level as most of the protein domains remain very similar.

The lamprey *cldn* mRNA sequences were translated to the amino acid sequence so that they may be used in the phylogenetic tree construction.

4.1.2. Phylogenetic Tree

To understand the evolutionary relationships between CLDNs in all vertebrates, a phylogenetic tree was constructed. Only the amino acid sequences were used as to account for the difference in evolution between the introns and exons. The tree included CLDNs from vertebrate model organisms, *Mus musculus*, *Gallus gallus*, *Xenopus laevis* and *Danio rerio*. Additionally the sea lamprey's 20 members identified in this project, were also included into the tree in order to determine their homologs, genes related by

descent from a common ancestral DNA sequence, and orthologs, genes in different species that evolved from a common ancestral gene by speciation, in higher vertebrates. Additionally *Ciona intestinalis* (11 members) were included as the closest relatives to vertebrates. However since CLDNs predate vertebrates, the tree was further extended and rooted with the invertebrate *Drosophila melanogaster* (3 members).

The phylogenetic tree was constructed on the online program phylogeny.fr. this implements two methods, the non-parametric bootstrap and the approximate likelihood ratio test, to evaluate branch supports in a statistical framework.

Figure 6 shows the final phylogenetic tree. Many of the lampreys *cldns* are closely related to their orthologs in higher vertebrates. The results shown in this tree correlate closely with the tree constructed by Loh et al., 2004, although this tree does not include the Fugu fish. However region 3 on this phylogenetic tree shows that even though the zebrafish lineage does not contain the CLDN, other genes such as zebrafish *cldn a, b, d* and *f* are part of the same evolutionary branch as CLDN 3. This is also seen in the tree constructed Loh et al., 2004.

Focusing on the lamprey, the tree shows that lamprey claudin identified in our initial screening, Cldn A, was grouped with CLDN 3. For this reason the gene was subsequently renamed *cldn 3B*. This was very interesting since generally CLDN 3 is known to be expressed in previously mentioned vertebrate-specific traits, the ear placode and the cranial ganglion.

Paralogs, genes related by a duplication event within a genome, have played a vital role in evolution of many gene families and the claudin gene family is no different. Paralogs can already be seen in the Sea lamprey suggesting early expansion of the CLDNs in the vertebrate lineage. Unlike orthologs, paralogs generally evolve new functions even if the function is related to the original one. Based on the importance of claudins during compartmentalization, this expansion in the gene family members would make it possible for the development of novel traits. The lamprey paralogs can be seen on the tree with CLDN 8 and Cldn B. *Cldn B* was identified in the screening of the lamprey cDNA and BLAST analysis showed this gene to be homologues to CLDN 8 in higher vertebrates (see Appendix B). The pairing of CLDN 8 and Cldn B was expected and thus it was renamed *cldn 8b* as the two are paralogs shown on the tree.

However partial sequences have a huge effect on alignment and therefore pairing on the tree. This could explain why the Cldn 3a and Cldn 3b do not appear as paralogs on this

tree since *cldn 3a* obtained is a partial sequence. This sequence is missing a few sequences at the 5' end which is why initially it was excluded from the dataset however it was included because it formed part of the novel lamprey claudin raw dataset and initial BLAST analysis of *cldn A* showed homologues belonging to *cldn 3* in higher vertebrates (see Appendix B). This means that it is very likely that *cldn A* is a paralog of *cldn 3a*.

The red brackets on the tree mark off regions with no lamprey claudins, signifying certain members of this gene family in higher vertebrates. The lamprey claudin gene family does not include a *cldn 4*. On our tree, Region 3 shows the CLDN 4 group, which is closely related to the CLDN 5 and CLDN 6 groups. All of these groups stream from the same lineage as the CLDN 3 group which, as previously mentioned, includes a claudin from the lamprey. *Cldn 4* and *cldn 3* arose by duplication in tetrapod lineage. In humans these genes are located 60kb apart on chromosome 7q11.23. The same region shows CLDN 14 group which has no members from the lamprey however this group is closely related to group CLDN 2.

Region 2 includes the CLDN 7 which does not include a member from the lamprey. However once again this group is closely related to groups that does include members from the lamprey, CLDN 19 group. Region 1 has CLDNs from the group 20 and above, these members are not found in the lamprey. All of these are examples of the gene family expanding and obtaining new members, hence way to date over 30 diverse members of this gene family have been identified so far in different animals and this phylogenetic tree show how closely related the orthologs are to each other and other groups. In this way novel function develop through duplications in individual lineages and is passed on through evolution. Hereby, leading to the rise of vertebrate specific traits but also novel characteristics, in this case it would be a preserved trait such as the ear placode development and the diverse trait such as the auditory organ development.

4.2. Claudin 3B

4.2.1. Expression

Whole mount RNA in situ hybridization was done on lamprey embryos between the stages E4 – E31. *Cldn 3b* probes were used in order to determine the expression of *cldn 3b* during embryogenesis. The results showed that *cldn 3b* is strongly expressed in the ectoderm, the neural fold at stage E5. At this stage the ectoderm cells opposite the primitive streak, begin to thicken and flatten to become the neural plate. This formation plays an important role in the first three steps of primary neurulation. The neural plate

develops into the elevated neural folds and then by stage E6 the neural folds fuse to form the neural tube, which also expresses *cldn 3b* (Fig. 7.C). Other areas which show expression of this gene include the stomodeum, the ear placode and forebrain. At later stages expression is localised in the pharyngeal arches, the pericardial cavity which contains the pericardial fluid.

Table 11: Comparison of *P.marinus* claudin 3 expression to that of higher vertebrates

	Species			
	Regions of Expression	<i>P.marinus</i>	<i>G.gallus</i>	<i>M.musculus</i>
Adult Tissues	Bladder	---	No	Yes
	Digestive Tract	---	Yes	Yes
	Gills	Yes	No	No
	Kidney	---	No	Yes
	Liver	---	No	Yes
	Lung	---	Yes	Yes
	Mouth	Yes	No	Yes
	Nasal Cavity	---	No	Yes
	Digestive tract	Yes	Yes	Yes
Embryonic Tissues	Dorsal Fin	Yes	No	No
	Ear Placode	Yes	Yes	Yes
	Ectoderm	Yes	Yes	Yes
	Endoderm	No	Yes	No
	Eye	Yes	No	No
	Forebrain	Yes	Yes	No
	Mesoderm	No	Yes	No
	Nasal Placode	No	Yes	No
	Neural Crest	Yes	No	No
	Neural Sheath	No	Yes	No
	Neural Tube	No	Yes	No
	Somites	No	Yes	No
	Stomodeum	Yes	No	Yes
	Pericardial Cavity	No	Yes	No
	Pharyngeal arches	Yes	Yes	
	Primitive Streak	No	Yes	---

For the sake of similarities only comparable body structures between the *P.marinus* and higher vertebrates has been included. More work has been done on expression of this gene in adults of higher vertebrates. Homologous structures have been grouped together on the bases of the body system in which the gene is expressing. This was done to get a more comparative data, as the gene may express “similar functions” in the individual regions of expression relating to the body system. The symbol “---“ marks region with unknown expression profile of claudin 3, either due to lack of published data or lack of comparative features. The figure includes data previously published by Chihara et al., 2013; Ohazama et al., 2007; Ohta et al., 2006; Haddad et al., 2011; Haworth et al., 2005; Ozden et al., 2010; as well as data from this study.

The expression profile in the lamprey shows similarities with expression of CLDN 3 in higher vertebrates, shown in Table 11.

The expression of *cldn 3b* is very strong in the ear placode starting from stage E8 and persisting through to stage E23. In comparison to the most well-known expression profile, the chick, the same pattern can be seen. Starting from the HH stage 11, the gene is strongly expressed in the Ear placode. This data suggest a preserved function of the gene in different vertebrates. Especially, since CLDN 3 has also been identified in the ear placode of other vertebrates, such as mouse which are closely related to humans.

It should also be noted that in both profile, gene expression starts very early in the ectoderm and also the head process. Although expression in the somites were seen in the developing chick embryo, this expression pattern was not seen in the lamprey suggesting a gain in function of the gene in higher vertebrates more closely related to humans. Alternatively, expression is seen in the pericardial cavity of the developing lamprey embryo but no such expression is seen in the chick or the mouse. However expression is seen in the gut, bladder and pancreas epithelium.

The uniqueness of the pharyngeal arches in vertebrates compared to the pharyngeal slit in invertebrates, the outpocketing of the endoderm, this gene expression site suggests that *cldn 3b* may play a role in the this development by sealing off areas which will not expand moving further away from the ectoderm. This will lead to the formation of the outpocketing shape and allowing the neural crest cells and mesoderm to fill the space. In invertebrates the pharyngeal region and the pharyngeal slits are relatively simple and do not have the same shape. This could be an example of a CLDN in vertebrate specific trait, beginning in one of the early members of the lineage. Throughout the course of evolution, paralogs could emerge in any of the claudins thereby further modifying the different traits within the different vertebrate. In the chick expression was also seen in the pharyngeal arches in later stages of embryo development. This suggest that although in both case the CLDN 3 is important for development of these structures, however in the sea lamprey, *cldn 3B* is required at an earlier stage to set up the bases for the development of the arches hence the expression of the gene in the neural crest cells located at the regions where the arches will develop. In chick embryo, keeping in mind the expansion of the gene family, possible another claudin is recruited to set up the bases for CLDN 3 to aid in the formation of the complex vertebrate specific arches.

Expression was seen earlier in the cranial neural crest cells. This is fascinating as these cells migrate dorsolaterally to produce the craniofacial mesenchyme that differentiates into bone, connective tissue, cartilage and cranial neuron which themselves originate from the vertebrate specific cluster of cell known as the ganglion. Section through the body of E31 embryos revealed expression to be in the dorsal fin and the coelom, along with these regions expression continues in the neural tube and the eye spot.

These expression profile compared to that of the chick embryo show some similarities especially in the vertebrate specific traits suggesting a preserved function of this gene in the lineage. However different expression profiles in the species specific traits support the theory that this gene family plays a role in morphogenesis and as such a mutation in the functions of these genes would contribute to the morphological diversification between species of the subphylum vertebrata.

4.2.2. Function

Knockout experiments, using morpholinos, were performed at the California Institute of Technology, Bronner Lab. The embryos were fixed in MEMFA and stored in 100% EtOH to be transported to Wits, School of Molecular and Cell Biology, Dr Nikitina Lab, for analysis. The morpholinos used in the experiments had a FITC-tag which was useful in allowing the visualization of the morpholino in the embryos under a fluorescence microscope.

The morpholinos were introduced into one side of the embryo at two cell stage of development. This was done so that one side of the embryo would be affected and the other side would develop normally thus the phenotypic mutation can be seen on the same embryo as a control. The embryos were also injected at 1 cell stage so that the mutations could be seen over the entire embryos. Control oligos with no target sequence in the lamprey embryos were used as negative controls. These morpholinos should produce no abnormalities in the embryonic phenotype; supporting the specificity of the target morpholino.

In order to visualize the structures of interest, *Sox-2*; *Sox1* and *Tbx1* were selected to be used as markers as these genes showed expression in the neural tube, otic placode and pharyngeal arch respectively..

Early analysis of the *cldn 3b* morphant embryos at stage E6 showed a mortality of 33%. This highlights the importance of the gene very early during development. Analysis of the

embryos seen in Fig. 15 & 20 revealed that these embryos had fragile ectoderms. This trait was evident during in situ hybridization, leading to the embryos falling apart during the procedure. This is seen clearly in Fig. 20.B.1 where the embryo shows ectodermal damage on the lateral side. Both Fig. 19 & 20.B. shows that without the *cldn 3b*, the shape of the embryo is disrupted along with the shape of the neural tube.

Expression profile of *sox-2* at the early stage of development in wild type is very strong in the neural tube. This was ideal to observe the abnormalities in the neural tube at this stage. The results from the *cldn 3b* morphants at stage E6 suggests that the function of this gene in this region is to aid with the fusion of the neural folds in order to form the neural tube and complete the neutralization process. The hybridization also revealed that the ectoderm is hugely affected by the lack of *cldn 3b*.

Reducing the concentration of morpholino injected into the embryos allowed more embryos to develop to a later stage for observation.

At stage E8, morphants that were injected at two cell stage showed two types of abnormalities in the ear placode development on the side which incorporated the morpholino, this is shown in Fig. 17.B & C. Over 62% of the embryos showed reduced or absent ear placode phenotype at this stage. Most of the morphants were severely affected and showed no placode formation. 12% of the morphants showed ear placode formation however the structure was abnormally small. It is possible that these embryos did not incorporate as much morpholino as the others, or the individual embryos expressed a higher level of the gene allowing them to develop the reduced ear placode. These results show the importance of *cldn 3b* in otic placode development and the absence of this gene would result in no placode or reduced placode. This is linked to the data shown in Fig. 16, which shows that half of the embryos were affected and the rest showed wild type placode development.

It should be noted that since *cldn 3b* and *8b* show similar expression in this region it is possible that in some embryos, *cldn 8b* could be recruited to compensate for the decreased expression of *cldn 3b* this way allowing for normal development at this early stage since the genes are very similar in their coding regions and expression site at this stage. Maybe recruitment is possible at this early stage.

The fragile ectoderm at stage E8 is still evident in the affected embryos in an embryo that was injected at the zygote stage. The results suggest that *cldn 3b* is important in the normal development of vertebrates' traits, the ear placode and the neural tube as well as

the ectoderm by stage E8. The high mortality in early development of the morphants can be due to the failure to develop normal ectodermal protective layer and likewise, the neural tube.

The probe selected to use during *in situ* hybridization for these older morphants was *SoxE1* since it was strongly expressed in the ear placode as well as the pharyngeal arches.

Fig. 11 shows the results of the whole mount *in situ* hybridization on the morphants. Fig. 11.A is the wild type expression test, done to ensure the expression of *SoxE1* in the areas of interest. This would also be used as the template for normal development of the pharyngeal arches and the ear placode. Fig. 11B is the embryo which was injected with the non-specific control morpholino. The phenotype of the embryo is normal in the development of the pharyngeal arches as well as the ear placode. The same phenotype is seen in the embryos injected with *cldn 3b* control morpholino. The controls therefore show that any abnormalities seen in the experiments could not be attributed to the effects of morpholinos introduced into the embryo or the injection technique.

The bar groups seen in Fig. 13 and 14 show the percentage of abnormal embryos in each group injected with a morpholino. The group injected with the non-specific control morpholino, shows approximately 83% of them had normal phenotypes and the remaining 17%, had abnormal body morphology which could be due to physical damage to the embryos during injection. Weaker embryos do not recover from the injection process at a very early stage. The same phenomenon can be seen with the embryos injected with *cldn 3b* control morpholino.

However due to the effect of the storage of the morphants during the transport from the USA to RSA, the level of RNA were compromised thus affecting the strength of the staining during *in situ* hybridization. Expression ranged from those with strong expression in the ear placode and pharyngeal arches, some had strong expression in the arches and weak in the ear placode and vice versa. These variations in RNA levels did not affect our ability to distinguish between the different phenotypes.

The *cldn 3b* morpholino injected embryos showed that just over 61% of the embryos had defects in the ear placode and/or the pharyngeal arches. Thus these morphants were further divided into three groups. 15.6% (13/83) of the morphants showed defects in both the ear placode and the pharyngeal arches. This can be seen in Fig. 11F. The morphants showed defects in the sealing off of the pharyngeal arches, specifically in the first two arches. This data suggests that *cldn 3b* functions in sealing off the compartments destined

to develop into the arches. Additionally the other morphants only showed phenotypic defects in the ear placode, where in the growth of these structures were retarded and their location was in closer proximity to that of the arches. This morphant can be seen in Fig. 11E. The class of embryos with defects only in the pharyngeal arches can be seen in Fig. 11D.

The arches are much smaller as the expansion is retarded. Fig. 12 shows the section through the morphants that belonging to the two classes with defects in the pharyngeal arches. Fig. 12.B' is a section through an embryo that had defects in the ear placode and the arches. The arches do not appear to have sealed off endorsing the phenomenon seen in from the whole mount *in situ*. The abnormality in the arches was again evident in Fig. 12.C'. The reduced development of the arches seen in the whole mounts *in situ*, Fig. 12.A.

This signifies how different phenotypes may arise from the decreased expression of the same gene. This could be as a result of the different levels of morpholino may have been incorporated. However the results were the same, abnormal developments of the ear placode and pharyngeal arches, which supports the theory that this gene function in normal compartmentalization during development of these two characteristics.

The implications of this feature, specifically in relation to the size and shape, suggest that a mutation in this gene could lead to a different morphology allowing for the development of novel characteristics. Since spatial restriction has a huge influence on body morphology, differential expression pattern of claudins would result in variations in the development of unique structures resulting in diverse organizations at the adult stage of different vertebrates.

4.3.Claudin 8B

4.3.1. Expression

Since this gene was identified from the screening of the embryonic cDNA library it was concluded that this gene is expressed during early development. RNA isolation identified this gene in embryos from stage E4 – E31, therefore these stages were again the focus of the *in situ* hybridization using a DIG-labelled *cldn 8b* RNA probe. The *cldns* used in *in situ* hybridization showed great similarities in their coding region. This was a huge factor which affected the first runs of *in situ* hybridization, which resulted in the sum expression of the claudins during development instead of the isolated individuals. Sequence analysis

revealed that the similarities were not as high in the untranslated regions therefore these regions and therefore could be isolated and used for the individual genes.

Fig. 9 shows the expression profile of lamprey *cldn 8b*. Fig. 9.A-C, stage E4 – E6, shows expression in the ectoderm much like *cldn 3b* expression at the same stages. However unlike *cldn 3b* E6, expression is not shown in the developing neural plate and the neural tube. This feature is not seen in the development of the mouse embryo, with expression of *Cldn 8* located to the Endoderm of the early stages of development. This suggests a gain in function of the gene as previously mentioned, with the emergence of new duplicated genes would be recruited for novel function often similar to the original one. With the expansion of the gene family over the evolutionary distance between the lamprey and the mouse, a lot of mutation could have occurred. The rise of new claudin gene family members has led to many novel claudins which developed new functions. This plays a role in increasing the complexity of the lineage and the rise of novel traits. Those traits could then be maintained by genes that are better suited to develop them. This is seen in the development of the swim bladder (which expresses *cldn 4, 5, 6 & 7*) and the development of the modified lung (which in itself expresses *cldn 1, 3, 4, 5 & 7*). Like *cldn 6* being no longer involved in the development of the modified structure, *Cldn 8* may evolve to be more specifically suited to the development of a different germ layer. However other expression profiles are more strongly maintained over the course of evolution.

This is seen in the expression profile of the developing ear placode from stage E8, as seen in Fig. 9.E. This expression pattern is maintained throughout the development of the embryo until stage E31. This feature is maintained in the mouse embryo which solidifies the importance of this gene in the development of the vertebrate specific trait.

Sections through the head show strong expression in the frontal regions of the head ectoderm. A variance between the two expression profiles can also be seen in the area between the developing stomodeum and the rest of the body, the region which will develop into the pharyngeal region.

At stage E9 expression begins in the stomodeum, similarly to that of *cldn 3b*. At E9, Fig. 9.F, expression begins in the neural tube however unlike the expression of *cldn 3b*, which is not only expressed in the neural tube but also the region destined to become the pericardial cavity. Sections through the head show strong expression in the developing placode and the frontal region of the head. Between stages E10 – E14, Fig. 9.G – 9.J,

expression is not only in the frontal regions of the head but also beings in the pharyngeal regions, much like *cldn 3b*. In mouse embryo development, CLDN 8 is seen in the forebrain and hindbrain. Again showing how, this gene family is important in development and thus it is preserved. Even in later stages of *Xenopus* embryonic development, Cldn 8 is seen expressing in the forebrain, midbrain and hindbrain.

The similar expression profile between *cldn 3b* and *8b* is very evident in the whole mount *in situ* hybridization, hence the sections through the embryos which reveal the expression profile in greater detail. Fig. 9.K – 9.M shows strong expression of *cldn 8b* in the pharyngeal arches just as it was for *cldn 3b*, minus the expression in the pericardial cavity and the fin fold. Fig. 10.D' is the section through the pharyngeal arches, specifically in the endoderm still it is different from that of *cldn 3b*. The difference in the expression profile can be seen in the schematic representation the genes in Fig. 26.

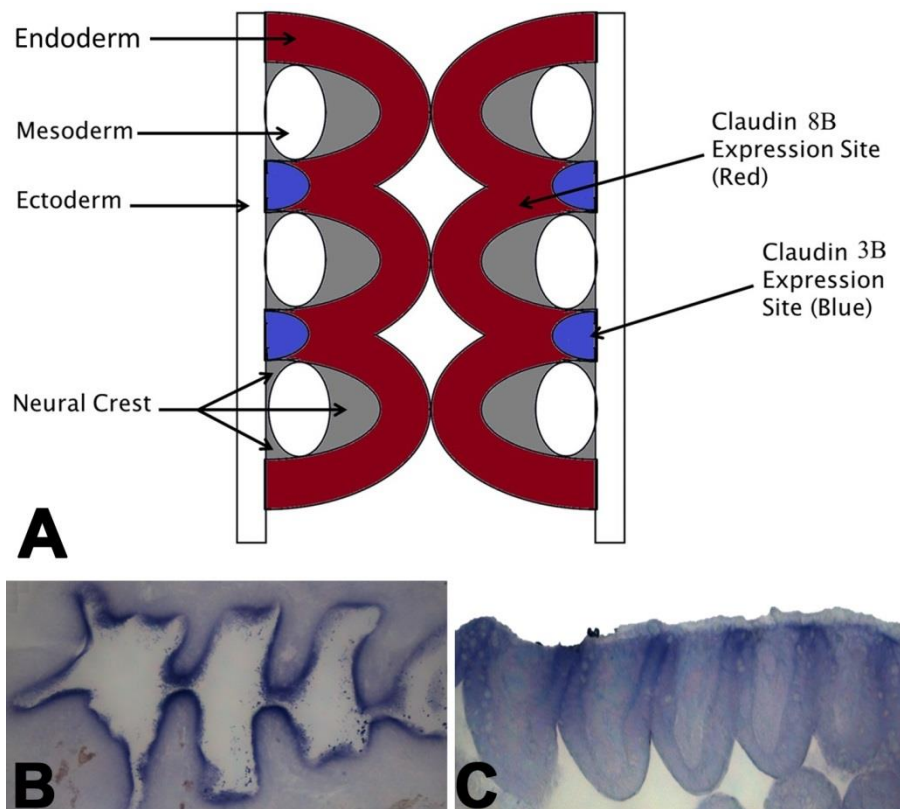


Figure 26: Expression profile of Cldn 3B and 8B in the pharyngeal arches of the lamprey. (A) Schematic representation of the Claudin 3B and 8B expression profile in the pharyngeal arch of the lamprey. (B) Image of lamprey pharyngeal arches after *in situ* hybridization using *cldn 8B* probe. Original magnification, 10X. (C) Image of lamprey pharyngeal arches after *in situ* hybridization using *cldn 3B* probe. Original magnification, 20X.

As previously stated, *cldn 3b* expression is only in certain regions of the endoderm which suggests its importance in the formation of the outpocketing of the endoderm. However *cldn 8b* expression throughout the endoderm, suggests that this gene is required in the development of the compartment of this region initially even before the overall pharyngeal arch structural formation. Expression in the endoderm at this stage is interesting, keeping in mind that in the mouse, *cldn 8* expression in the endoderm begins at a younger stage.

Fig. 9.M.2 shows some strong expression not only in the pharyngeal arches and the ear placode but also in the developing gut at stage E17, this was not seen in the *cldn 3b* expression profile. Additionally expression in the mouse is seen in the development of the urinal system as oppose to the digestive tract again supporting the theory of recruitment. Likewise the *Xenopus* embryo also expresses *Cldn 8* in the development urinal system showing conservation which evolved after the divergence of the lamprey from the higher vertebrate lineage.

During the later stages of development expression of the gene was also more specific, at E23 the expression is only in the ear placode and the first pharyngeal arch. The other regions of the body do not show clear expression but a little bit of some background staining except for the head region. This suggests that this gene is important for the development of the compartments in the body then the gene expression is decreased and only maintained in a few areas seen in Fig. 9.N. At stage E31, Fig. 9.O.1 shows some expression in the mouth region which was not seen at the earlier stage. At this stage the oral cirri begin to develop, thus the gene may be reactivated to play an important role in the structural changes. Fig. 10.E' shows expression of the gene in the placode and also the connective tissue of the mouth, it is also expressed in the eye spot at E31. Mouse embryo also showed expression of *cldn 8b* in the developing eye. This expression is not like the expression of *cldn 3b*, which was more widely spread in the cartilage of the head.

Table 12: Comparison of *P. marinus* claudin 8 expression to that of higher vertebrates

	Regions of Expression	Species		
		<i>P.marinus</i>	<i>D.rerio</i>	<i>M.musculus</i>
Adult Tissues	Brain	---	Yes	No
	Ear	---	No	Yes
	Gut	---	No	Yes
	Heart	---	No	Yes
	Kidney	---	Yes	No
	Pelvic urothelium	---	No	Yes
	Skin	---	No	Yes
	Ear Placode	Yes	No	Yes
Embryonic Tissues	Ectoderm	Yes	Yes	Yes
	Endoderm	Yes	No	Yes
	Eye	Yes	No	Yes
	Heart	No	No	Yes
	Mouth Structures	Yes	No	Yes
	Nervous System/Brain	Yes	Yes	Yes
	Neural Crest	Yes	No	No
	Neural Tube	Yes	No	No
	Pharyngeal Structures	Yes	Yes	Yes
	Urinary System/Kidney	No	Yes	Yes

Only comparative body structures between the *P.marinus* and higher vertebrates have been included. Expression is seen at an embryonic stage as well as the embryonic stage of development. Homologous structures have been classified together. Certain features have been grouped based on the system in which the gene is expressing, seen with the urinary system. This was done to get a more comparative data, as the gene may express “similar functions” in the individual regions of expression relating to the body system. “---“, Unknown” expression. The figure includes data previously published by Raciti et al., 2008; Clellan & Kelly, 2010; Thisse et al., 2004; Hou et al., 2010; Ohazama et al., 2007; Ohta et al., 2006; Fujita et al., 2006; as well as data from this study.

The similarities of the coding regions allow the genes to have very similar functions and the expression profile has shown a lot of similarities between *cldn 3b* and *cldn 8b*, however they showed differentiations in expression sites showing how each one of the genes have been recruited for specific function during development of the embryo. Even in higher vertebrates expression of both genes can be seen in the ear of the mouse nonetheless expression in other regions of the body is different, like for example CLDN 8 expression in the pharynx during mouse development. A summary of known CLDN 8 expression in vertebrates can be seen in Table 12.

The expression of *cldn 8* is preserved in mouse, chick and the lamprey, indicating the importance of *cldn 8* in vertebrate development. The expression of this gene is preserved especially in vertebrate specific traits such as the pharyngeal structures. Surprisingly the zebrafish does not show expression of *cldn 8* in the ear placode which may be attributed to the extra round of duplication which may have led to another claudin fulfilling the role of *cldn 8* during development of this structure. Supporting the hypothesis of how expansion of this gene family could lead to the rise of novel genes which are vital to the morphology of vertebrate lineages.

4.3.2. Function

Knockout experiment using Cldn8bMO were performed as described for Cldn3bMO knockout experiments.

Early analysis of the *cldn 8b* morphant embryos saw a high mortality of 66%, higher than that seen with the *cldn 3b* morphants. It was necessary to reduce the concentration of the morpholino injected into the embryos. In some cases the volume of the morpholino introduced into the embryos was decreased to try and get the embryos to develop to the required stage. Although this analysis shows that both genes are important in the formation of the ectoderm, the different mortality rate suggests that *cldn 8b* is more important in the formation of a stable ectoderm. Therefore, incorporation of the morpholino proved fatal to most embryos.

For this reason, as with *cldn 3b*, in some cases the concentration of morpholino injected into the embryos had to be reduced in order to allow them to develop to the desired stage. However in some cases, good quality embryos were able to develop even after injection with a smaller volume of 1mM morpholino solution.

Tbx1 was used as the marker for ear placodes at later stage of developmental stages to better visualise the defects in areas which normally express *cldn b*. Since the expression profile of this gene is also in the ear placode, this was made the focus of the experiments and the morphants were expected to show a reduced development in this region as it was seen in *cldn 3b* morphants.

Embryos injected on one side at the two cell stage and the development can be seen in Fig. 21. Fig. 21.A shows the normal expression of *tbx1* at stage E8. Expression is strong in the ear placode and the ectoderm is properly sealed thus it did not capture any of the

excess staining. Fig. 21.B - C show the side of the embryo which was not injected with the morpholino.

Usually when injecting a cell at this stage minimal amount of morpholinos still manages to drift into the other side of the developing embryo. As a result, the side which wasn't injected with morpholino showed two phenotypes in different embryos. The first class includes the embryos which had normal ear placode development on the side that was not injected with the morpholino (Fig. 21.B). The ectoderm still showed irregularities.

The second class consisted of embryos which had weak placode on one side and no placode on the injected side. This could be attributed to the drifting morpholino into the other cell at the two cell stage of development. Together these two classes were grouped under the embryos that showed no phenotype. Fig. 22 shows that 39% of the embryos developed normal placodes on both side.

The other 61%, showed severe irregularities in the placode, Fig. 21.D shows the side of the embryo which was injected with the morpholino, that side was severely affected and completely lacked the ear placode. However some of the embryos were not just affected on the side with the morpholino, 13%, also showed strong phenotypic abnormalities on the other side. The embryos had no ear placodes on either side. The remaining 48% of embryos developed no ear placode on the side that had integrated morpholino.

These results suggest that *cldn 8b* expression in the ear placode may be to initiate the development of the structure, unlike in the *cldn 3b* morphants, which showed retarded and slight spatial abnormalities of the structure. Together the two gene may play a role in determining whether the placode will form or not and also where and the shape of the placode during early development.

Fig. 23 shows stage E8 embryos injected with the standard control at the whole cell stage of development. As previously mentioned, during the long storage stage of these embryos as they were being delivered from the Bronner lab to Nikitina Lab, the RNA started getting compromised and started to degrade even though the embryos were fixed in MEMFA before being stored in MeOH for transport. This has to do with the storage conditions during transport. Generally the embryos are stored at -20°C to preserve the RNA for long periods of time. Unfortunately these embryos were kept in storage for months before they were delivered thus during that time some of the RNA had already degraded.

This was seen in all of the embryos but more so in the younger embryos at stage E8 and younger. The effects of this storage was seen here with more than 84% of the embryos not staining the placodes adequately enough even after a long staining period. Re-antibody and re-staining followed in order to try getting a better result but the RNA in these embryos was degraded and the results could not be maximised. Future experiments would require a better transport method in order to preserve the RNA. Alternatively, if time and funding permits, the *in situ* hybridization experiments would need to be done at the Bronner lab.

Whole mount *in situ* hybridization on *cldn 8b* knockout morphants at stage E10 lacked ear placode is shown in Fig. 24. Since these embryos were injected at the whole cell stage of development the resulting abnormality was seen on both sides at this stage, Fig. 24.B. Showing how important the gene is in developing a stable ear placode. Again supporting the hypothesis that *cldn 8b* is involved in whether this placode will form or not. Interestingly most of the embryos showed normal phenotype at this stage. When comparing the data shown in Fig. 25 against other data at a younger stage of development the results suggests that there may be some kind of compensation mechanism which helps with the normal development of ear placodes at a later stage. For instance maybe *cldn 3b*, or another lamprey claudin, could be recruited to fulfil the role of the down regulated *cldn 8b* since both genes are strongly expressed in this area during normal development and, as members of the same gene family they share biochemical functions also seen with other members of the lamprey claudin gene family.

Keeping in mind that claudins are indeed cis and trans-acting but also homophilic and heterophilic. Thus future studies would have to focus on the type of interactions formed by *cldn 8b* in the ear placode and how these interactions is affected by the down regulation of this claudin. It is possible that there is another claudin that interacts with *cldn 8b* in a heterophilic trans-interaction and in the absence of *cldn 8B* both cells produce this claudin so that the interactions become homophilic, this would depend on the conformation of the extracellular loops creating the required type of TJ.

Nonetheless the current data shows the importance of *cldn 8b* in ear placode formation and since orthologs generally retain the same function, this gene may also be thought of as vital in the development of the ear placode in mice and other higher vertebrates which have a preserved expression profile for this gene.

4.4. Conclusion

Claudins have had an important role in development since even before the evolution of vertebrates seen with discovery of claudins not only in all vertebrates but also in invertebrates such as the fruit fly, *Drosophila melanogaster*, which belong to a different phylum. The protein synthesised by these genes not only share similar amino acid sequences with some vertebrate claudin but also play a vital role in organization of the septate junction. On a larger scale these proteins participate in morphogenesis by way of their role in regulating the paracellular transport through the septate junction. These junctions are analogous to the vertebrate specific tight junction. Early expansion is seen in the *Ciona intestinalis*, with 11 members identified, showing evolution in the chordate lineage.

Since their discovery in 1998, about 30 homologues have been identified in all vertebrates. The phylogenetic tree showed a lot of homologues and orthologs in vertebrates that aren't seen in early vertebrates such as the sea lamprey suggesting great expansion in individual lineages. Most of the variations are in the size and the ECL1 domain; which is involved in regulating the pore size properties at the TJ but it is the highly conserved ECL2 which governs the size of the paracellular cleft. Generally while CLDNS are located at the apical region, novel claudins have also been identified also in the basolateral membrane, which is oriented away from the lumen, thus showing recruitment of claudins for new functions. CLDNS also maintain the apical-basolateral gradient of epithelial cells, thus insuring that other integral proteins maintain their position.

Vertebrates have become one of the most diverse and complex subphylum of animals. Novel characteristics emerging along with other traits which were maintained, and at times modified, has been one of the driving forces behind the success of this lineage in different environments. Advanced sense organs, brain encased in bone, jaws and lungs are examples of features which have been elaborated throughout the course of evolution, from the simple body plan. The phenomenon seen here also allowed for the rise of unique vertebrate specific traits; such as neurogenic placode derived from the ectoderm, important for the development of the cranial sensory system, amongst these placodes is the ear placode. Another trait is the emergence of the pharyngeal pouches in vertebrates. Novel trait would lead to the evolution of new compartments and ultimately leading to the appearance of novel types of vertebrates.

The results shown in this experiment show that even in the earliest vertebrates, this gene family is involved in the embryogenesis, specifically in certain vertebrate specific traits. Mutations in these genes affect the presence of certain structures, their shape and size are also affected. Thus non-fatal mutations would change the entire morphology in adults, possibly leading the development of novel characteristics allowing for diversification.

The early vertebrate, *Petromyzon marinus*, showed 20 members of the gene family. Humans have 27 identified members while over 25 members have been seen in mammals. So far 16 claudins are identified in chick. *Xenopus* have 20 claudins identified and zebrafish has about 33 members of the claudin gene family. The overall pattern suggest a correlation between the claudin gene family size and complex vertebrate evolution.

Compartmentalization is an important stage of morphogenesis and it forms the basis of variations in individual lineages during vertebrate evolution. Claudins play an important role in determining the boundaries which separate the compartments, as seen with the lamprey *cldn 3b* expressed in the fold endoderm of the pharyngeal arches. The separation of the compartments is done by the sealing off of the claudins which assists in creating the hydrostatic pressure. Together with the extracellular materials, the hydrostatic pressures form the spatial restriction for the epithelia which covers all structures in the body. As mentioned in the text, compartmentalization controls the division of functions within the different cell lines and the supply of nutrients and information over great distances, stressing the importance of the claudins, as the core unit of the TJ complex in the epithelial sheet. The role of CLDNs in development of vertebrate specific traits, such as *cldn 3* and *8* in the lamprey and mice, further supports the importance of these genes in morphogenesis of vertebrates. Also other structures, example *cldn 5a* in zebrafish development of the embryonic cerebral-ventricular barrier system; mice *Cldn 4* and *6* important in the formation of the blastocyst and even *Cldn 1* important for the epithelia formation in mice. Additionally CLDNs continued expression in adults where they continue to evolve and develop new function such as the mice *Cldn 13* involved in the stress induced erythropoiesis pathway. In humans, 7 CLDNs are expressed during organogenesis of the complex lung and only 5 are expressed in development of the swim bladder the evolutionary precursor of lungs.

Mutation such as gene duplication, would lead to complex modification, as the genes are recruited for novel traits during morphogenesis, in this way the expansion of the claudin gene family may lead to divergence of that organism possibly into a novel more complex

vertebrate species as further mutations occur. Further experiments would need to be done on the other lamprey claudins in order to obtain a complete lamprey claudin expression profile which could give a more complete story with the functional evolution of other genes such as *CLDN 16* or *CLDN 7*. However current data already supports the hypothesis of claudins being involved in the rise of complex vertebrates.

REFERENCES

1. Abdelilah-Seyfreid, S. (2010). Barriers : Another brick in the wall, 768–776. doi:10.1002/bies.201000045
2. Antin, P.B., Kaur, S. Stanislaw, S., Davey, S., Konieczka, J.H., Yatskievych, T.A. & Darnell, D.K. (2007) Gallus expression in situ hybridization analysis: a chicken embryo gene expression database. *Poultry Science*, 86 (7): 1472-1477.
3. Ashida, H., Ogawa, M., Kim, M., Mimuro, H. & Sasakawa, C. (2012) Bacteria and host interactions in the gut epithelial barrier. *Nature Chemical Biology*, 8 (1) 36-45. doi:10.1038/nchembio.741
4. Ballabh, P., Braun, A., & Nedergaard, M. (2004). The blood-brain barrier: an overview: structure, regulation, and clinical implications. *Neurobiology of disease*, 16(1), 1–13. doi:10.1016/j.nbd.2003.12.016
5. Chapman, A. D., (2009) Numbr of Living Species in Australia and the World. 2nd edition. Australian Government, Department of the Environment, Water, Heritage and the Arts. Canberra, Australia.
6. Chihara, M., Ikebuchi, R., Otsuka, S., Ichii, O., Hashimoto, Y., Suzuki, A., Saga, Y. & Kon, Y. (2013) Mice stage-specific claudin 3 expression regulates progression of meiosis in early stage spermatocytes. *Biology of Reproduction*, (86 [1]) 3
7. Clelland, E. S. & Kelly, S. P. (2010) Tight junction proteins in zebrafish ovarian follicles: stage specific mRNA abundance and reponse to 17beta-estradiol, human chorionic gonadotropin, and maturation inducing hormone. *General and Comparative. Endocrinology*, 168 (3) 388-400
8. Collins, M. M., Baumholtz, A. I. & Ryan, A. K. (2013) Claudin family members exhibit unique temporal and spatial expression boundaries in the chick embryo. *Tissue Barriers* (1) 3. DOI:104161/tisb.24517
9. Coulombe, P.A., & Lee, L. (2001). Defining Keratin Protein Function in Skin Epithelia: Epidermolysis Bullosa Simplex and Its Aftermath. *Journal of Investigative Dermatology* 132: 763–775
10. Crosfill, B. M. L., Widdicombe, J. G., & Bartholomew, S. (1961). IN DIFFERENT MAMMALIAN SPECIES From the Department of Physiology ,

Medical College of mammalian species is that small animals breathe faster than large ones ., 1–14.

11. Crow, K. D. & Wager, G. P. (2006) What is the role of genome duplication in the evolution of complexity and vertebrate evolution. *Trends Ecol. Evol.* (20) 312-319
12. Cybulski, L. E., & Mendoza, D. De. (2011). Playing with transmembrane signals, *Communicative & Integrative Biology*, 4(1), 69–71. doi:10.4161/cib.4.1.13778
13. Dörfel, M. J., & Huber, O. (2012). Modulation of Tight Junction Structure and Function by Kinases and Phosphatases Targeting Occludin, *Journal of Biomedicine and Biotechnology*, 2012. doi:10.1155/2012/807356
14. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. Vol 32 (5) 1792-1797
15. Fall, B. (2008). BIOL 101L: Principles of Biology Fall 2008.
16. Fujita, H., Chiba, H., Yokozaki, H., Sakai, N., Sugimoto, K., Wada, T., Kojima, T., Yamashita, T & Sawada, N. (2006) Differential expression and subcellular localization of claudin-7, -8, -12, -13 and -15 along the mouse intestine. *J. Histochem Cytochem.* 54 (8) 933-44
17. Gilbert, S. F. (2000) *Developmental Biology* 6th Edition: Neural Crest. Sunderland (MA): Sinauer Associates
18. Graham, A., & Begbie, J. (2000) Neurogenic Placodes: A Common Front. *Trends Neurosci.* 2000: 313-316
19. Gumbiner, B.M 1993 Breaking Through the Tight Junction Barrier. *The Journal of Cell Biology*, Vol 123, 1993
20. Günzel, D., & Yu, A. S. L. (2013). Claudins and The Modulation of Tight Junction Permeability. *Physiological Reviews*, 1–45. doi:10.1152/physrev.00019.2012
21. Gupta, I.R. and Ryan, A.K. (2010). Claudins: Unlocking the code to tight junction function during embryogenesis and in disease. *Clinical Genetics* Vol. 77 (4): 314–325

22. Haddad, N., El Andalous, J., Khairallah, H., Yu, M., Ryan, A. K & Gupta, I. R. (2011) The Tight Junction Protein Claudin-3 Shows Conserved Expression in the Nephric Duct and Uretric Dub and Promotes Tubulogenesis in vitro. *American Journal of Renal Physiology*, 301(5): 1057-65
23. Haworth, K. E., El-hanfy, A., Prayag, S., Healy, C., Dietrich, S. & Sharpe, P. (2005) Expression of Claudin-3 during chick development. *gene Expression patterns* (6) 40-44. doi:10.1016/j.modgep.2005.05.002
24. Hayashi, D., Tamura, A., Tanaka, H., Yamazaki, Y., Watanabe, S., Suzuki, K., Suzuki, K., Sentani, K., Yasui, W., Rakugi, H., Isaka, Y. & Tsukita, S. (2012). Deficiency of claudin-18 causes paracellular H⁺ leakage, up-regulation of interleukin-1 β , and atrophic gastritis in mice. (2012), *Gastroenterology*, 142(2), p292. doi: 10.1053/j.gastro.2011.10.040.
25. Hiroshi, A., Michinaga, O., Minsoo K., Hitomi M., & Chihiro S., (2012). "Bacteria and host interactions in the gut epithelial barrier." *Nature Chemical Biology* 8(1): 36–45. <http://dx.doi.org/10.1038/nchembio.741>.
26. Hou, J., Renigunta A., Yang. J. & Waldegger, S (2010) Claudin-4 forms paracellular chloride channel in the kidney and requires claudin-8 for tight junction localization. *Proceedings of the National Academy of Sciences*. 107 (42) 18010-5
27. Kaarteenaho, R., Merikallio, H., Lehtonen, S., Harju, T., & Soini, Y. (2010). Divergent expression of claudin -1, -3, -4, -5 and -7 in developing human lung. *Respiratory*, 1–10.
28. Krause, G., Winkler, L., Mueller, S. L., Haseloff, R. F., Piontek, J. & Blasig, I.E. (2008) Structure and function of claudins. 1778, 631-645. doi:10.1016/j.bbamem.2007.10.018
29. Krause, G., Winkler, L., Piehl, C., Blasig, I., Piontek, J., & Müller, S. L. (2009). Structure and function of extracellular claudin domains. *Annals of the New York Academy of Sciences*, 1165, 34–43. doi:10.1111/j.1749-6632.2009.04057.x
30. Loh, Y. H., Christoffels, A., Brenner, S., Hunziker, W., & Venkatesh, B. (2004). Extensive expansion of the claudin gene family in the teleost fish, *Fugu rubripes*. *Genome Research*, 14(7), 1248–57. doi:10.1101/gr.2400004

31. McCauley, D. W. & Bronner-Fraser, M. (2003) Neural crest contributions to the lamprey head. *Development*, 130, 2317-2327
32. Marchio, C., Weigelt, B., & Reis-filho J.S. (2010) Adenoid cystic carcinomas of the breast and salivary glands (or ‘ The strange case of Dr Jekyll and Mr Hyde ’ of exocrine gland carcinomas). *Journal of Clinical Pathology*, 63, 220-228
33. Nag, M. L., & Morin, P. J. (2009). Protein family review The claudins. *Genome Biology*, (10) 235.1-235.7
34. Nikitina, N., Bronner-Fraser, M. & Sauka-Spengler, T. (2009) The Sea Lamprey *Petromyzon marinus*: A Model for Evolution and Developmental Biology. *Emerging Model Organisms: A Laboratory Manual*, Vol. 1.
35. Nikitina, N., Sauka-Spengler, T & Bronner-Fraser, M (2008) Dissecting early regulatory relationships in the lamprey neural crest gene network. *PNAS* 2008 105 (51) 20083-20088; published ahead of print December 22, 2008, doi:10.1073/pnas.0806009105
36. Ohazama, A. & Sharpe, P.T. (2007) Expression of claudins in murine tooth development. *Developmental Dynamics*, 236 (1) 290-4
37. Ohta, H., Adachi, H. & Inaba, M. (2006) Developmental changes in the tight junction protein claudins in murine metanephroi and embryonic kidneys. *Journal of Veterinary Medical Science*, 68 (2) 149-55
38. Ozden, O., Black, B. L., Ashwell, C. M., Tipsmark, C. K., Borski, R. J. & Grubb, B. J. (2010) Developmental profile of claudin-3,-5 and -16 protein in the epithelium of chick intestine. *The Anatomical Record*, (7) 1175-83
39. Putman, N. H., Butts, T., Ferrier, D. E. K., Furlong, R.F., Hellstem, U., Kawashima, T. & Rokhsa, D. S. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature*, 453 (7198), 1064-71. Doi:1038/nature06967
40. Raciti, D., Reggiani, L., Geffers, L., Jiang, Q., Bacchion, F., Subrizi, A.E., Clements, D., Tindal, C., Davidson, D.R., Kaissling, B. & Brandli, A.W. (2008). Organization of the pronephric kidney revealed by large-scale gene expression mapping. *Genome Biol.* 9 (5):R84

41. Rossa, J., Ploeger C., Vorreiter F., Saleh T., Protze J., Günzel D., Wolburg H., Krause G. & Piontek J. J. (2014) Claudin-3 and claudin-5 protein folding and assembly into the tight junction are controlled by non-conserved residues in the transmembrane 3 (TM3) and extracellular loop 2 (ECL2) segments. *The Journal of Biological Chemistry*, 289, 7641-7653 (2014)
42. Saldaña, C., Alta, R. De, Carlos, S., Investigación, D. De, & Facultad, B. (n.d.). Claudins in Normal and Lung Cancer State.
43. Sasakura, Y., Shoguchi, E., Takatori, N., Wada, S., Meinertzhagen, I. A., Satou, Y. & Satoh, N. (2003) A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. *Dev. Genes Evol.* 213: 303 - 313
44. Sasom, R.S. (2009). Phylogeny, classification and character polarity of the Osteostrace (Vertebrata). *Journal of Systematic Palaeontology*, Vol. 7 (01); 95-115. Doi: <http://dx.doi.org/10.1017/S1477201908002551>
45. Sauka-Spengler, T. & Bronner-Fraser, M (2008) A gene regulatory network orchestrates neural crest formation. *Nature reviews Molecular Cell Biology* (9) 7: 557-568
46. Shang X., Lin X., Manorek G., Howell S.B. (2013) Claudin-3 and claudin-4 regulate sensitivity to cisplatin by controlling expression of the copper and cisplatin influx transporter CTR1. *Mol. Pharmacol.* **83**, 85-94
47. Shen, L., Weber, C.R., Raleigh, D.R., Yu, D. & Turner, J.T. (2011) Tight Junction Pore and Leak Pathways: A Dynamic Duo. *Annual Reviews of Physiology*. Vol. 73. 283-309
48. Smith, J. J., Antonacci, F., Eichler, E. E., & Amemiya, C. T. (2009). Programmed loss of millions of base pairs from a vertebrate genome, *Proceeding of the National Academy of Science*, 106(27), 11 212 - 11 217. doi:10.1073/pnas.0902358106.
49. Smith, J.J., Kurank, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., Yandell, M. D., Manousaki, T., Meyer, A., Bloom, O. E., Morgan, J. R., Buxbaum, J. D., Sachidanandam, R., Sims, C., Garruss, A. S., Cook, M., Krumlauf, R., Wiedemann, L. M., Sower, S. A., Decatur, W. A., Hall, J. A., Amemiya, C. T., Saha, N. R., Buckley, K. M. & Rast, J. P. (2012) Sequencing of the Lamprey (*Petromyzon marinus*) genome provides insight into vertebrate evolution. *Nature Genetics* 45, 415 - 421

50. Soini, Y. (2011). Claudins in lung diseases. *Respiratory Research*, 12(1), 70. doi:10.1186/1465-9921-12-70
51. Takezaki, N. Figueroa, F. Zaleska-Rutezynska, Z. & Klein, J. 2003. Molecular Phylogeny of Early Vertebrates: Monophyly of the Aganthans as Revealed by Sequences of 35 Genes. *Molecular Biology and Evolution*. 20 (2), 287-292.
52. Thisse, B. & Thisse, C. (2004) Fast Release Clones: A high throughput expression analysis. ZFIN direc data submission.
53. Thompson, P. D., Tipney, H., Brass, A., Noyes, H., Kemp, S., Naessens, J., & Tassabehji, M. (2010). Claudin 13, a member of the claudin family regulated in mouse stress induced erythropoiesis. *PloS one*, 5(9). doi:10.1371/journal.pone.0012667
54. Tipsmark, C. K., Baltzegar, D. a, Ozden, O., Grubb, B. J., & Borski, R. J. (2008). Salinity regulates claudin mRNA and protein expression in the teleost gill. *American journal of physiology. Regulatory, integrative and comparative physiology*, 294(3), R1004–14. doi:10.1152/ajpregu.00112.2007
55. Tsukita, S., & Furuse, M. (2000). Pores in the wall: claudins constitute tight junction strands containing aqueous pores. *The Journal of cell biology*, 149(1), 13–6. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2175101&tool=pmcentrez&rendertype=abstract>
56. Wen-jin, Z., & Min, Z. (2007). Diversification and Faunal shift of Siluro-Devonian vertebrates of China. *Geological Journal*, 42; 351-369.
57. The World Conservation Union. 2010. IUCN Red List of Threatened Species. Summary Statistics for Globally Threatened Species. Table 1: Numbers of threatened species by major groups of organisms (1996–2010). http://www.iucnredlist.org/documents/summarystatistics/2010_1RL_Stats_Table_1.pdf
58. Xu R., Li Q., Zhou J., Zhou X.D., Perelman J.M. & Kolosov V.P. (2013) The degradation of airway tight junction protein under acidic conditions is probably mediated by transient receptor potential vanilloid 1 receptor. *Bioscience Reports*, 33

59. Yanai, I., Camacho, C. J., & Delisi, C. (2000). Predictions of Gene Family Distributions in Microbial Genomes: Evolution by Gene Duplication and Modification, 2641–2644.
60. Yanai, I., Peshkin, L., Jorgensen, P. & Kirschner, M.W. (2011) Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Developmental Cell*, 20 (4): 483-96
61. Van de Peer, Y., Maere, S. & Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature Reviews. Genetic*, 10(10), 725-32
62. Van Itallie, C. M. & Anderson, J. M. (2006) Claudin and Epithelial paracellular transport. *Annu. Re. Physiol.* (68) 403-29
63. Zhang, J., Piontek, J., Wolburg, H., Piehl, C., Liss, M., Otten, C., Christ, A., Willnow, T.E., Blasig, I.E., & Abdelilah-Seyfried. S. (2010). Establishment of a neuroepithelial barrier by Claudin5a is essential for zebrafish brain ventricular lumen expansion. *Proceedings of the National Academy of Sciences*, 107 (4), 1425-1430.

APPENDIX

6.1. APPENDIX A

6.1.1 Preparation of MEMFA

10 x MEM Salts was prepared as follows in order to be used to make up the final MEMFA solution:

<u>Reagents</u>	<u>Quantity</u>
MOPS (pH7.4)	104.65ml
EGTA	3.804ml
MgSO ₄	0.602ml
H ₂ O, distilled	390.944ml

The solution was stored in the dark at RT

The final MEMFA solution was subsequently prepared as follows:

<u>Reagents</u>	<u>Quantity</u>
Formaldehyde	2.5ml
10x MEM salts	1ml
H ₂ O, distilled	6.5ml

Stored at 4°C for up to 12 hours protected from light

6.1.2 Preparation of LB Media and LB Agar Ampicillin Plates

LB media where prepared, in a 1 L Erlenmeyer flask, as shown below:

<u>Reagents</u>	<u>Quantity</u>
Tryptone	5g
Yeast extracts (Provides the minerals for bacterial growth)	2.5g
NaCl	5g
dH ₂ O	500 mL

The solution was mixed on a hot plate for approximately 30 minutes then autoclaved for approximately 1 hour.

For LB that would be used for bacterial strain consisting of the plasmid and the insert; 500ul of ampicillin was added to the solution after it had cooled to 55°C or lower.

In order to make the LA plates; 7.5 grams of Agar was added to the previously mentioned recipe for LB media.

The solution was sterilized in the same manner as LB media. For plates that would be used for screening after transformation; 500 µl of ampicillin was added to the solution after it had cooled below 55°C at after which the agar was poured out into sterile petri dish. The plates were left to cool until the agar had solidified then the plates were left upside down at 37°C overnight. The plates were stored at 4°C until they were needed.

6.1.3 Preparation of SOC Media

SOB Medium:

<u>Reagents</u>	<u>Quantity</u>
Tryptone	20g
Yeast Extract	5g
NaCl	0.5g
ddH ₂ O	1 L

Autoclave

SOC Medium:

<u>Reagents</u>	<u>Quantity</u>
SOB Medium	10 ml
1M MgCl ₂	0.1 ml
1M MgSO ₄	0.1 ml

Filter Sterilize

Add 0.1ml of 2M filter sterilized glucose solution OR 0.2ml of 20% (w/v) glucose. Filter Sterilize

6.1.4 Preparation of Transformation Buffer

<u>Reagents</u>	<u>Quantity</u>
0.5 M CaCl ₂	20 ml
1M Tris.HCl pH 8)	2 ml
ddH ₂ O	178 ml
Total Volume	200

6.1.5 DNA Miniprep Protocol

- Resuspend Cells, Lyse and Neutralize
 - Add to pelleted cells:
 - 250µl of Resuspended Solution (with RNase A) and vortex
 - 250µl of Lysis Solution and gently invert the tube 4-6 times
 - 350µl of Neutralization Solution and invert the tube 4-6 times
 - Centrifuge at 12 000 RCF for 5min
- Bind DNA
 - Load the supernatant into a GeneJET spin column
 - Centrifuge at 12 00 RCF for 1min
- Wash the Column
 - Add 500µl of Wash Solution and centrifuge for 30-60sec
 - Discard the flow-through
 - Repeat
 - Centrifuge the column one more time for 1min
- Elute Purified DNA
 - Add 500µl of Elution Buffer to the column and incubate for 2 min
 - Centrifuge for 2 min and collect the flow-through containing the DNA

6.1.6 Insert Digest

<u>Reagents</u>	<u>Quantity</u>
Plasmid DNA	4 – 5 µl
ddH ₂ O	18.5 – 19.5 µl
FastDigest Buffer	2.5 µl
Fast Digest SalI	0.5 µl
Fast Digest NotI / XhoI	0.5 µl
Total Volume	25 µl

Incubate in 37°C waterbath for 30 min

Increase Temperature to 65°C for 1min to inactivate enzyme

6.1.7 Preparation of Ladder Mix and Agarose Gel

The ladder mixture was prepared as follows (In accordance with the Thermo Scientific GeneRuler protocol):

<u>Reagents</u>	<u>Quantity</u>
DNA Ladder	30 µl
6X DNA Loading Dye	30 µl
Deionized Water	<u>120 µl</u>
	180 µl

Aliquoted in 60 ul stocks

6.1.8 Preparation of HybeMix

Reagents	Quantity
Formamide	250ml
20X SSC _{DEPC} (pH 5w/citric acid)	32.5ml
0.5M EDTA (pH 8)	5ml
tRNA (20mg/ml in H ₂ O _{DEPC})	25ml (or 2.5g)
Heparin	50mg
Fill to final volume with ddH ₂ O	500ml

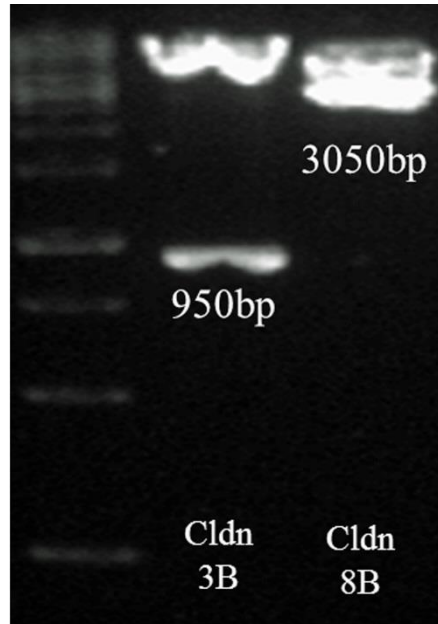
6.1.9 Preparation of NTMT

NTMT	Quantity
5M NaCl	1ml
1M Tris.Hcl (pH9.5)	5ml
2M MgCl ₂	1.25ml
10% Tween-20	0.5ml
ddH ₂ O	42.25ml
Total	50ml

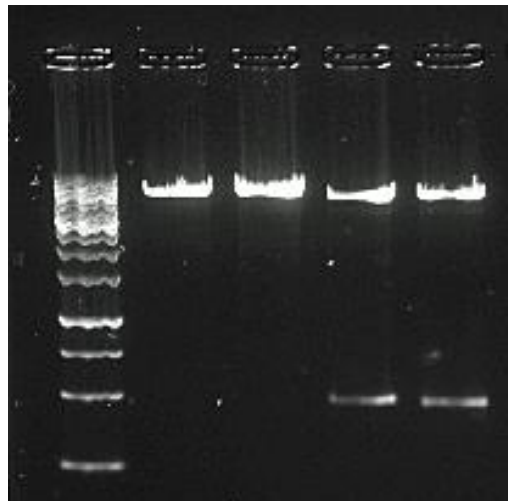
6.2. APPENDIX B

6.2.1. Electrophoresis

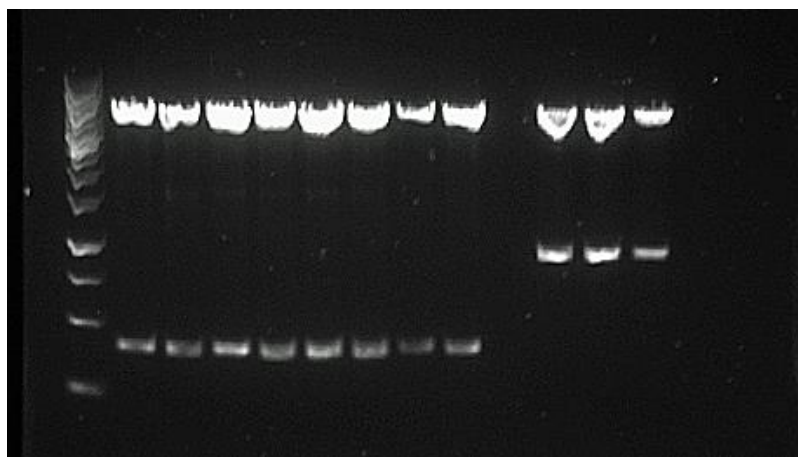
6.2.1.1. Double Digest



NotI and *SalI* double digest of different pCMV-Sport 6 to cut out the lamprey Cldn 3B & Cldn8B inserts respectively. Lane 1 is the 1 Kb DNA ladder.

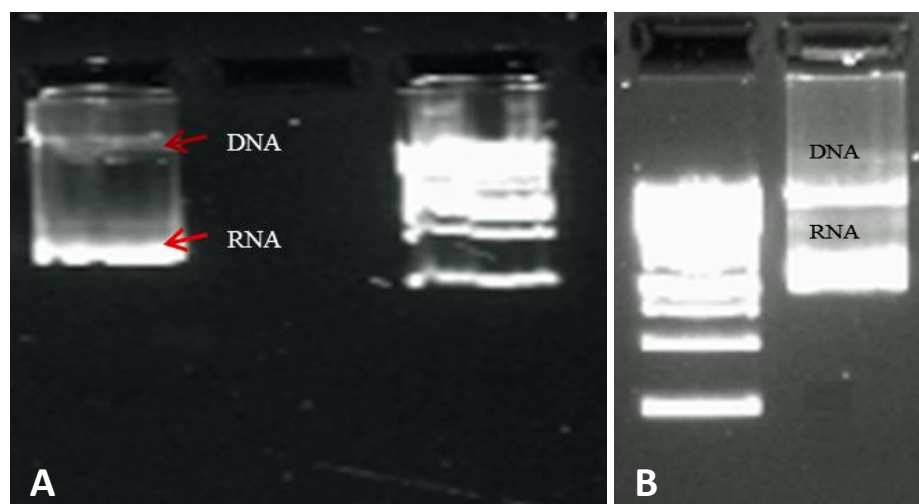


NotI restriction of pCMV-Sport 6 containing Cldn 3B insert versus *NotI* and *PvuII* double digest of different pCMV-Sport 6 to cut out the coding region of lamprey Cldn 3B. Lane 1 is the 1 Kb DNA ladder. Lane 2 – 3 is the *NotI* restriction samples. Lane 4 – 5 is the *NotI* & *PvuII* double restriction samples.



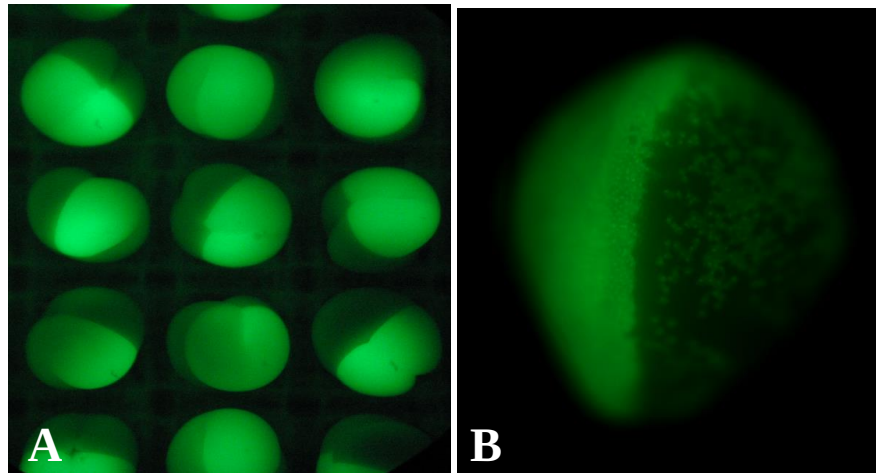
NotI and *SalI* double digest of pCMV-Sport 6 containing Cldn 3B insert versus *NotI* and *PvuII* double digest of different pCMV-Sport 6 to cut out the coding region of lamprey Cldn 3B. Lane 1 is the 1 Kb DNA ladder. Lane 2 – 9 is the *NotI* and *PvuII* double restriction samples. Lane 11 – 13 is the *NotI* and *SalI* double restriction samples.

6.2.1.2. Probe Synthesis



Newly synthesised DIG-Labelled RNA probes of Cldn 3B (Image A) and Cldn 8B (Image B) before purification of the probe. The Top band is the DNA template fragment and the lower band is the RNA fragment.

6.2.2. Morpholino Injections



Images from Nikitina et al. (2008) In *Emerging Model Organisms, a laboratory manual*. Vol. 1, Chapter 16, p 405-429. (A) MO injected embryos at 4 cell stage. MO is incorporated into one side of the embryo. (B) MO injected lamprey embryo at E4 stage.

6.2.3. Lamprey Claudin Family

6.2.1.1 Six fold screening of lamprey embryonic cDNA library

Results from sequence analysis comparing the six possible lamprey claudin clones (A-F) obtained from the embryonic cDNA library against the NCBI database

<u>Lamprey Claudin</u>	<u>Confirmed Claudin Family Member</u>	<u>Homolog</u>	<u>Species</u>
A	Yes	Claudin 3; Claudin 6	<i>Xenopus (Silurana) tropicalis</i>
		Claudin 4	<i>Xenopus laevis</i>
		Claudin 4	<i>Bufo japonicus</i>
B	Yes	Claudin 8 isoform 1	<i>Pongo Abellii</i>
		Claudin 8	<i>Homo sapiens</i>
		Claudin 3	<i>Anolis carolinensis</i>
C	No	SP4 Transcription Factor	<i>Homo Sapiens; Bos Taurus</i>
D	No	Glycosyltransferase	<i>Triticum aestivum</i>
F	Yes	Claudin 19	<i>Sarcophilus harrisii</i>
E (2011)	Yes	Claudin 9	<i>Danio rerio</i>
		Claudin 3	<i>Anolis carolinensis</i>

6.2.1.2 Claudin Gene Family Members

>PmCLDN1a :

mhvvglfst lgfgltmvi ylnwsiysh gpgtymqsgl wnicvmast ergevcqfy

talsqprffq vvrallmav avgaialfvs tlgtacarrf sehastktkl tcaggallll

tgclvmlavs lyanfvvsql ynplepygvr fsfgaalytg wlagafivig galmvavffr

knslrtyrf vitdivasp

>PmCLDN1b:

miggvngtrv pvssliegta epdadenavm malcalgggt glwgrclnlp shrgelcltp

gflqewaprf sepgnhnsqf dpvrtalcra qlllplvalt lllltlalal igcasnssla
lipcailtll aglctlttas svslslpcv dassspgwst flalasacqd iaatglliwa
raaslrhyrs vl

>PmCLDN2:

mvstglqiva lclavlgllg tlaatilphw qvtahvgtsi itatgqmrgl wmecawmstg
ffqcrpfysi lamnpalkaa qammviscvl taagigaatv gmkctlllgp ssrrskslva
vsgglclivs alctlvpvaw nthntvrqfy dpwypspvky elgaaiylgy aaailtmaag
vmmalscpek rragpprhl rprdgpprg vesagtgtpe vhtgmswv

>PmCLDN3a

mqqckvydsml algqdlqaar altvvavvaa aaallvaiaag akcttcieeee etkarvsiaa
gagfvlagll elvpvswsan svvrdfynpl vpdstkreig aamfigwaaa tllvlggall
ccscppkqrk yaappkggggg ggsgggggggg ddgddrkkgk

>PmCLDN3b_CLDNA

mqlglalci vgwlgaitvc alpmwkv saf ignnivvaqi vwegiwmncv vqstgqmck
mydsmlalgq dlqaaralsv iaivltvlgf lvsivgakct tciedettk a ritvgsgif
ilsgymelip vsaantiir dfynpivpda mkreigasmy igwaaaalli lggallccqc
prkqekymgp rsvapprsva tarsnyv

>PmCLDN5a

mgsatvegla valcsigwlc diiacaipmw rvtafignni vtaqgiwegl wmscvfqstg
qvqckvydsv lalgpelqaa rslsvlsall almalplgla gapctgcarg darskarvsa
lagtlaalag alqlappswt ahavirdfhd pripaskre lgpcvyaawl gagllllggs
llcggscgac gahtestrge wrdpspawdr ppqqpl

>PmCLDN8b_CLDNB

manaalqiag ialgilgwig siavtvpqw rvtalpnas nivvgqvfw giwmqcvsq
tgqqqckvyd slaltpdmq aarglmcsv algtlamia tlgmqcthc enerakaya
msggflfiia alivlipvcw tanaiirdfy npsvpealkr elggalygw isafmlvagg
allacscppk depgyrpaks ayavsaprtv qdthqpram sttysmkefv

>PmCLDN8

manaalqiag ialgilgwig siavtvpqw rvtalpnas nivvgqvfw giwmqcvsq
tgqqqckvyd slaltpdmq aarglmcsv algtlamia tlgmqcthc enerakaya
msggflfiia alivlipvcw tanaiirdfy n

>PmCLDN9

mqqkvhdsm angelqaar amtvlanlla llavgvaatg mrcttcgded gsakgrltat
ggvgfalagl lmlapvswa hnivrdfydv gvpvltkrel gsalyvgwga atllflgggl
lccswprpep grprpfyglq pgglamhgtg pnapgaplpr mhsaaaaaa

>PmCLDN10

msgtclqvlg tiiaclgwvg atvatamnew rvtssrsvi tatwvfqglw mncagnalga
vhrphltvf klenyqvacr afmitsvflg clamicspfg macirisfvs dkakrnitfi
tgficalqgk prrlmgtys gnsvtveyd ptiigakyl gtalfigwas alltltggti
ivcsgvgrsy kqprpwrall pspwrft

>PmCLDN11

mgascfqlg vsllglgwva lvvatatndw vllcsyfta cvrfedmgsk glwaqcisnn
gidhckpfae ilalpvyiqv cralmvvacv lglpaillas tglkcirlge etasekaqr
vaggvmlilv glctlvattw fpvgthkesl kgvvnfgysl fagwvgtala ivsgallccc
sgskeeryya ypatasaaaa

>PmCLDN12

msgcgflvl wslcgaaalc llvaaalmps wrdlilvdpt krsivvrhgl wvrcagyegd

vpggvswaa awfgpdsgke gvdgggspgc vvrdsawpia tdsldrlrq lgvptaclva
lvalalnvsg avvlcrsgrr sgrlygeded stslstsss ssefsskrk rkkkkaqqcc
qfnpcgcqv asvlyllasa lvlppalwal lsvdelnrah vsgdarwrpg eavpvclaac
agllaaagl amwyctcrpl ppaplpapl algsllgmea lglsgragpa eacgrstrqs
hrttgsfpag ptsvssrpms rateldipm yelh

>PmCLDN15

maldvgtll ciagwtlvgt iitnhywks tvagsvittn tiyenlwqac atdsmgvsnc
rdfsmlnlp ghvqacralv itaivlgffg tllaflgmrc tnivssdaev kgrivfsagi
iyileglsai iavswyaaqv vaefdpfya gkyelgpal ymgwagsgla ilgggllccv
cctgarsstq rdhtfkysaa raatpistat prtildmegq tyarqgyv

>PmCLDN15LB

msgtclqvlg tiiaclgwg atvatamnew rvtsrassvi tatwvfqglw mncagnalga
vhcrphltvf klenyvqacr afmitsvflg clamicspfg macirisfvs dkakrnitfi
tgfcilqgk prlmgltys gnsvtveyyd ptiigakyl gtalfigwas alltltggti
ivcsgvgrsy kqprpwrall pspwrft

>PmCLDN16

mitalqllaf fflvsasfl vvatwtcdwm vnaddslevs qkcrglwwec vtnafdgirt
cdqfdisilae hplkivlsrs liltadilac faciilgl dcvrvlsnep riktricyta
gfiyllagip gmvgsvwyav dvyverstlv qqnvflgiqy efgwscwlgm agsmacilag
ivltccvylf rdanaarlhr spyvaprtsl gkmyamashv

>PmCLDN17

mantrqllla malallgwc mlvscmlpmw kvsgmssssa flaqqvfegl wvhcvlqgtg

dmqcrthgsm mslgtsmqit ralvvcattml gyvaffvala gakcttwlde satktrlats
aavcfavvql lelvaavscma hhiisdhnp lispdmrrei glslylgfga aallimgssf
mlcasrgsdr lqfl

>PmCLDN18

mhvvqfalst lgfgltvmvi ylnwsiysh gpgtymqsgl wnicvmast ergevcrqfy
talswprffq vvrallmav avgaialfvs tlgtacarrf sehastktkl tcaggtlwl
tvlsrvvcs sfvgslvtef pgwlrsvpts pggvggllss fgelqtlytr fsfgaalytg
wlagafivig galm

>PmCLDN19a

mansglqmlg yalgmcgwia itaatalpqw ktsayagevi itavsiyegl fmascasqstg
qiqckvfds lalptrvqia ralmicsivv glfalgvsav gmkctrmgad nkarnriai
iggvvflvag llcliatsyy asdiarefys pntpvqarye fgqalfvgwa gacvlmggs
flccscnsks sgktsrprgp prgppsstg saqkeyv

>PmCLDN19b_CLDNF

mansglqmlg yalgmcgwia iiaatglpqw rtsayagevi itavsiyegl fmascasqstg
qiqckifds lalptrvqit ralmicsivv glfalgvsav gmkctrmgad nktrknrfai
lggvvflvag llclsgtsyy aaeiarefys pltpvqarye fgqalfvgwv gacvlmgga
flccscsskp sgkksrprgp prhgpprskg sgrtdyv

>PmCLDN20

mvstglqiva lclaivgllg tlaatilphw qvtahvgtsi itatgqmrgl wmecawmstg
ffqcrpfysi lamnpalkaa qammviscvl taagigaatv gmkctllgp ssrrskslva
vsqglclivs alctlvpaw nthntvrqfy dpwypspvky elgaaiylgy aaailtmaag

v

6.3. APPENDIX C

6.3.1 Construction of the phylogenetic tree

The Sequences used to construct the phylogenetic tree along with the accession number

Claudin	Sequences	Accession
CiCLDN1Like	avsyglinis avlgvlsslm mliililpew rinsvvasei nmsmvrareg iwircvssyn aywncdtfdv sliglpayln afralmiigf iaslggvvcs vlgmrcasvv deeskakrwi viiggllchia agvltgicvs wyaaevvwey yspyyqtqkl kytlgaclyv gwiamlfalv agclmtccac cnkpadstnt ysynpaklqa pnhsyppgnsf aflyi	Ensembl: ENSCINP00000025043
CiCLDN1Like2	mskswihidg fwkecigytt gewhcdgfgd flrlfstyli lgrtlmivti gcaslavfls lagmncstml pedssikhrv yklsvlflfl saisatvitl wsairtssef yacvqqqrsq qkvnk	Ensembl: ENSCINP00000025340
CiCLDN1Like3	mvngglqigg flgligwil givtcalpdw rkndlqgevi esvirttqlw vrcttqatgl wtcdydsff lglpvalqaa racvclsiam gflscicsvf glectnisag npkmmkaritl vagllnfigg vsvgiavswf aanvlqeywn plgaatsntr yvygsalfvg wismavsvia glvqccsswn vddedngrpy tynpskpkas saeyi	Ensembl: ENSCINP00000029937
CiCLDN7Like	maisscqivg ylvgtlwg allvvtlpdw khsyvsersf sgvlvvrqlw ttctvkesgs vecinisepv falptikgc rtsvlgtvl afcavaippl gmrctriggd aktklriarc cavlfglai ligiavswya edvakdfhnp lsrlklppvs dlkyvygnal yigwatmatg aisstfl	Ensembl: ENSCINP00000016825
CiCLDN7Like2	massalqvfg fdlglawvg migtcaspew rknsghgtli dntlkqeglw vscfafvtgq tnckdyaeff vnlpqvlqic ralmissigt gfvaaiitvl gmkcvglgss nhivkarial lglllfigis vltgaavsvy aqgvlqdyyn psfaqgltnv p	Ensembl: ENSCINP00000033582
CiCLDN17Like	madyatdsas sviaplckfa gfifgltaas cvfvvtvtpc wyviyaesi spyencyglw kfcstglcg kwlsqllgs hgaqrvtcrg lmitacllg falilnvmgm rcillfrllp rtkekltra avvwivagtf vllstwygl evmysvwdeg kaselsdgif vgwatafay aaagllwiga tvsrserkea ekrrdsilst vaaayaasrs gsgasttccc ylaprlgag hecghsntg ddkirtlepi tecqvtshar ttym	Ensembl: ENSCINP00000009847
CiCLDN18Like	masgllhiavc ilyviafile vvtvflpywk rndledtld sivrheglwi rcqtfatgnw dcdfrnfrll glptelqaar gflsllsfis figiclavaa tvlsvkslqn ktiasggas mlaglliaa vswyandiri qheyiltrll tgsgtrytf gyalflgwvg galqvvgism alciscgnsd eedipptiry hptkstlaps qeyv	Ensembl: ENSCINP00000035894
CiCLDN19LikeV1	mavkpalrtv fflgmlifv gvfttcvlpv wkisswqtss alygrmptyd gvwvqcmtlk aggfcsnyg kpsiflsadi qamralmcia tiasflamvc slsldctta leantkpkni itmvsygfhi vaglmcgav swyawrvtye fyfpyqere fyeyfsgcly igwvssaial laglvlicga cknkdedeh ssypytyqpr yeiyv	Ensembl: ENSCINP00000019620
CiCLDN19LikeV2	msksvlqavq wvflgcca tigvtlshkw kntvegihh svldsvsnva glwdfcqvfp sgqtqcpwia atsvpallf crilmifgl lgcavagfli csfkcttlg earharekls lycglmnlls gicvvatss yiyetlntpp svdprgqsvm nevseaaqis lafsiiwcla svvmiagstt taepkvkyhy ptkpimqey v	Ensembl: ENSCINP00000029858
CiCLDN19LikeV3	nlgmygqili fvlfavstsl itttvwnqw avngptsssv ltaiwawrgi wgdcmqmygg qyqceemqsm lnekgyilai rammcials svigcvtli gmecsklldg nlsakrnlmr isgvllfvag vltasvswy tgvvvedffs vftqdnmiy emgaalylgf vssalallta ilcfcmgpgk pededspyk ytpgtvire pyetkvqipk skdkysevv	Ensembl: ENSCINP00000000852
CiCLDN19Like2	mvngglqigg flgligwil givtcalpdw rkndlqgevi esvirttqlw vrcttqatgl wtcdydsff lglpvalqaa racvclsiam gflscicsvf glectnisag npkmmkaritl vagllnfigg vsvgiavswf aanvlqeywn plgaatsntr yvygsalfvg wismavsvia glvqccsswn vddedngrpy tynpskpkas saeyi	Ensembl: ENSCINP00000029937
DmCLDNLSINU	mrelnkqqsq dtdsgvekq dypratngvv fgaivtfasf fvlmmsfscsp ywiesyeetr asfknmgllwq ycfkdvfyypk yafkqftgc hnfisheyyv ireyllpgwl mavqgfvtns fiivflval lsliiripl kavlyewll vrlsymgtai sslfmflavc ifggcayrrd wmmypkfnvl gwsyalavvt fmlglalaali lqrearqayd argeqknlvn qmemeqpgyq pprhhsqsr slqgyi	NCBI: NP_647971.3
DmCLDNLKUNE	mgasnrilqv alcavsfsls cfviafstpy wlvtdgrlqn prftnlglwe vcfnnfqdih rffdsnfgc lwvfeeyyi ihdfllpgfy isvqlfatlc fvmclvvipl tvaflrtsrd	NCBI: NP_610179.2

	ddrymvllla igscqvvgsv fgfiavvifg akgsrdwmp gwqnnmgws falgvgavl llpsgvlylv earrerykrl neisnreise ygdeyyqnqg psavltaps pyqsyfasep srprpqqs asnsavqggi qtdi	
DmCLDNLEGA	mqqkshfare dsqehlptsh tqnshtvnmk rrtlsqscgv gvfvafafi viafatpswl vsdyritgak ldriglwhc frslpdvndd sqrrffvgr wvydpftgy deirgflpa fmiatqffyt lafigmlvsa igvlvfilca gpdqkhfitl ikslgyvllg agvsaaiavi vfagfgnmng wmpenhannwf gwsfilacvg tvltlvastl flseahvqhk kriqfkesqt rfelvrg	NCBI: NP_726742.1
DrCLDN1	mahagqlmgl yclgflgllg liastamaew kmssyagdni itaqaqyegl wqscvsqstg qlqckkydsi lklpgeiqga rglmltgifl cglstlvsvf gmkcttclse apqvkksval aggvfitggl lfaliatswy gekirqkffd pftptnarye fgkalyvgwg ssalsiiggs llccicgsea sekpsyppar aagrgpdrv	NCBI: NP_571845.1
DrCLDN1.1	mvdnryatal vigsvlslla tvylsvamgt qhwfyhsqp asndinngsd lqnlrldfeg edvdemsysk alfqlngtlg lwwrcirvhq dshwfkapdp vmvtkcqaft lqqqweskyy spgnintged llrtylwkcq flpllisvgl vflaaligvc aclcrsitpt lfvglhlma glcslgtvcc fltghllhd iselpgmdy spgwsflal vssplqvmaa alfiwaarsh rqhytrmtay rva	NCBI: NP_001003751.1
DrCLDN1.2	mvdnrhatal viaavlslla avyvsvisgs phwrycspp vraepnatel ralqeefldg dfdektasda lfrnngtvgl wwrvcvqtpd ahwyrepdpq mvmevcsfsl sqqltpkyre pgnhnsgegm irtylwrcql llplaalllv lfsallgfca cvcgsvtpal figllhlag lcslatvccf lagtdllhrv svlpdrvdaa lgwalylmml aapllmmaaa lmvwaarshg qnyarmtayr va	NCBI: NP_956255.1
DrCLDN2	mailalelmgl fffgligmlg tlvatllpyw atsahvgpni vtavvsmkgl wmcvqvstg afqcetyntl lgltdlqaa rammvmssif svmacavstv gmqctvcmdg ssvktkvagv ggsmfllagl lslipvawkt hevvtqfym pmpaskfci gdclyvglas slsmlgggl lsaccddid gnrgsrhpy ypemalrnp shsmtyqpaa lhstnpnit nktqtlnsht ssgthsiqag dsrkmtrqnt aagydvtygy	NCBI: NP_001004559.2
DrCLDN4	mvsmcreilg mclaiigflg aiialcpmw kvtafigani vtaqtiwegl wmcvcmqstg qmckkiydsi lalpqdlqaa ralvviaiiv sffalilgia gkctnfvcr edakakvsia sgvifiagv lslvpvcwsa ntiirdfynp lltadqrrem gaslyigwgv aallilgggi lccscppkde kymmkysqpr statsrayv	NCBI: NP_571840.1
DrCLDN5A	masaalellg lilvcgtilg emvacglptw kvtafieani vvaqtiwdgl wmscavqstg qmckkvhdsm lalshdlqaa raltviissv slvalmvvia gaqctncike dnvkarvvnv ggviyilsai fvlvplcwma nniisdynp qvlpaqkrei gaalyigwaa sallllggsi lccscpasgs sgysvkapt kratsngeyd kmnyv	NCBI: NP_998439
DrCLDN5B	manmisacle ivglcltvgt tllvmvacgl pmwkvafie gnivvaqniw dglwmscvvq stgqmckkmh dsvalttdl qtaraltvis avlgvlalmi tiagaqctnc intesikskv vnaggvmyiv agflvliplc wmanniisdf ydpqvyakk reigaalyig waasamllvg gailcfsrph dekstypky ippptkgti ngdydkmyv	NCBI: NP_001006044.2
DrCLDN7A	mansgvqlg flslgllg liigtimpqw kmsayvgdni itavatyqgl wmscafqstg qlckkiydsi lqlsdldqat ralmivgiiv siaglgvasi gmkcttcgad dkvrkrtam tggiiilvga lcavvacswf ahnvirafyn pftpvntkfe fgaaifiawg gsflldvlgga mlaascprsk qvskypksns trsangsne yv	NCBI: NP_001002340.1
DrCLDN7B	mahkgllqllg flslgllg liigtimpqw kmsayvgdni itaiamyqgl wmscayqstg qqqckvydsv lqlsalqat ralmvvaill tvaglgvasm gmkctncggd dkvkskriam tggiiilvga lcsivacgwf tsqiirdfyn pftpvntkfe fgaaifiawa gafldimggg mlasscsksg sspnypkssr pvkssrppss skevy	NCBI: NP_571712.1
DrCLDN9	masfglelv vtlsvlgwil nivccalpmw rvtafigtni vtaqvywegi wmscvqvstg qmckkvysdm lalpadlqaa ralvvaaiiv gvalfvaiv gakctnciee eaakarvmis sgaafitasv lqlipvcwsa htvielfysp lypeaqkmei gaslylgwaa samllvggsi lccscppkde tryppqsria ysahhsvaps tynkrdyv	NCBI: NP_571839.1
DrCLDN10	mgsmateiva fltisgwil vstlptdyw kvssvdgtvi ttatfwsnlw ktcvtdstgv snckdfpsml aldayiqvcr glmisavclg ffgasalvg mkctriggse tkaritclc glhfilsgic sitacslyah ritseffdl fvkqkfelga alfigwagsv lsilgglic fsmtegfisr eysyngatsv istrkitkt nkstksqkdp sdqmfsgrrf grmayv	NCBI: NP_001007038.1
DrCLDN11A	mantclqfqt fvmsflgwil livatatnew vftckyqmnt crkmdleak glwadcvist alyhchitq ilelpaiqt sralmvthi lglpavvmvl mmscinlgg epesaknkr vlggilillt afcgivstvw fpigahhekg lmsfgslsly gwvgtalcll ggciiscsv dspatyenn rfyyskqgpa htgptstnha ksahv	NCBI: NP_001002624.1
DrCLDN11B	msnsrllcg flmsfvwig iiaitstndw vlsctyghs crnmddletk glwtecvist alyhchitq vripayia crvlmvsas lglpalall lampcvkvsq etegtkhrra vqggliilvl slcgmvstvw fpigkldglm sfgslslyagv vgsalcfag svmvccskdh spenpesrh yysnpdgvtv tpsenpetr yysnhdgvt ssgtsenqet rfyysnhdgv tsygpaknsh aksehv	NCBI: NP_571847.2
DrCLDN12	mscrlihatn afafviamls vggltvatli pqwrtmrllt fnrnaknvtv ydglwtckvr	NCBI: NP_571848.1

	rdgssgcyyf dsdwyakvdq ldlrlqfcl ptglfssla llcltgmck taccsktpdd iknsrclvns sgchlvagml lflggaiaip psvwflfthr hlnerydhlf avefgyvvai gsaggilaa lfmfwycmc kklpspflw lpevptmtqs lsaqlmting lptpvyapq sfppavldap afvpmqgypq pastqpmpph vmpqmsvdp gygsevgasq vysyapsqsyapsqgyapsq syapsqrvag hrysrtsts gieidipvl d	
DrCLDN15A	mdpiieaval flgfvswtmv gitipnrywk vssldgvtit tslyenlwm scatdstgvh qcrefplla lsgyiqasra lmiaavvsqt fgvvatlign qcskaggeny vlkgriagt gvffllqglc tmsvswyaa nitqeffnpl ypgqkyeige alyigwasav laicggvclm fscldgtek taysyhtpre tvysasasrr etqstygnkna yv	NCBI: NP_956698.1
DrCLDN15LA	mstalevtgy fmccligwvlt glavandywk issiggnv snrlyenlwh acgedstgka ncqdfqsmia lphviqacra lvialllgl vglvltmgl ncikigske eskgnmfig giylggic tmgvswyaa rvvqefndp ygvrfelgs glyigwagaa lclmgggfc sayqrfsksk ekgayypaqk pqtiytaqs naetskayv	NCBI: NP_571846.1
DrCLDN15B	mqnllqifal llglgcsa vlslynhywk vstddsvii tsnifenlwm scaddssgtf scrdfqslla lpgyiqacra liitsvltg flvatlvgm qcfkigggny imksriagag gvffllqglc tmtavswyaf nitqeffdpl ypgikyeye glyigwssat lalcsgscll cscglctpmg kksfrqshlq dfrsshqpqs inivptvst shylsrqyv	NCBI: NP_001035404.1
DrCLDN15LB	mavmvlqvl flglgvcwcl essinsvsw kssshgeavv tassqfeglw fscasnlga ihcqrkttl glpgyiqacr almiialilg llsvlasmg lkctklgsts eakgkislt agimfilsil cvivavswya arvvqefndp fyggtykelg aglylgwaaa alcllggtl ctsfkgsspa qtrgpgynys aapqkiyrs apsdnsitka yv	NCBI: NP_001002446.2
DrCLDN17	mvqgpcdvla lclsliglv vatvtalpmw rvtafigeni ivmetrwegl wmnryqani rmqckvydsl lyllpdlqaa rglctclsal cggiivslf glrcmsclsd qpmksilml itgiieilas fcivpvswt ahtiirdfyn pllldaqre lgealyigwv tsafllasgv iflcrvdsn eqpfyarprr nkqvpnhfqt fnsirrpql qnnsfstpqh ssaafsavt pssgqysqhn vqpvvngsli yspnmippqr vnyngqnvml sqilasnsqh pvsdprdsfn tgsafhpvqq nplyvgynys rvqsgssns tglri	NCBI: NP_001104000.1
DrCLDN18	msapalqitg fvsqvigtg vfaatlmvsw cfrnqqddp qrvssiytyk glwkdcmesg tvfpecpily shpnysgilg aaralmiaai lvaviavfig fclclclmkm nmqlstrakl ilssaiiffi agicgiaaas vyadqlvpsf mmpqfnqkqg ekggvqcann vaamdpasr ytfgpalyia wvgaallilg gilmsiaykr mqcktetreg yiynaaqra aeescqctq gqslsqcqvq chsqvqri	NCBI: XP_002660862.1
DrCLDN19	mansgfllg yflalggwig iistvlpqw kqssyagda imavglyegl wmscasqstg qvqckifdsl lsdvniqtc ralmvisvml gfmgiivsv gmcktkvgdn npstkskiai sgslfllsg vcltvavswy aaqvsahffd pntptnakyd fgtalfvgwa asvlmmlgga flccscined rrgqfyqrq qpsttrezv	NCBI: NP_001017736.1
DrCLDN20	mmpsstmqmf afilallgv satvatllpn wkvsadvgsn imtaisqmqg lwmdctwyst gvfscslkys vlapaylqt arttmvlscm matmgclaa lglkctrwgg syrskghtai sagacfiag ilclvpaswf tnevittfmd skvpksskye pgaavyvafv sagflaagv ifclscpprr ggaldsgssh pekpkqlklr qkqptgnqkk qhreqqsqt etpeqeesp etlhqekppr qgekssvdr lnlekekqr spsrtrmqda kavyslhdyv	NCBI: XP_005173866.1
DrCLDN22L	mepgscalel lgvffslac lclslstmmt rwtlstell presfelglw mtcvvelg tecrpysllg lppdirlari mmctsavaag salvfaipgi nlvnscknra dsieakrtlk ifggilslss gvlgiqvsvy vahltvrrff desvpsvpr wefgdalflg wtagclqvva glitscffl qdktrglges ihmdrvntr sprmrtenv	NCBI: XM_001342099.1
DrCLDN23	mhtpasmimg vvfaplgvl vfaatpqw regqarlgaa gkggatell lrsdglwesc lqvvhseve cwpvsgsyqr dsrvrmvqgm vlsslflga givlasigvr cwtidplrsv aaagglval agvlsaalg vytchlsklg idvdsnlset qglnahkppr lthpagsly fgwvgaqvq lggiallygf kdmkcsfcrk qlrkdstemy avnc	NCBI: XP_001333025.2
DrCLDNA	mvsaglmvg ialavigwig vivvcilpmw qvtafigqni vtaqvtwegi wmscvvqstg qmckvydsm lalssdlqaa ralciisilv gvggillaaa ggkcttike erakakvci sgaifivagv lclipvcwta nniirdfyn mtnsaqkrel gaslfigwga salliggs lcancppqdg ykatytarsg gkgyv	NCBI: NP_571837.2
DrCLDNB	mastglqmlg ialaifgwig vivlcalpmw kvtafigani vtsqtswegi wmscvvqstg qmckvydsm lalssdlqaa ralvisivi gvmgimlsma ggkctnciee esskakvgit agvifiisgv lclipvcwta naiiqdfyn lvvqaqkre gaslyigwga salliggs lcchcpeksd sgkytakyna tprseasaps gknfv	NCBI: NP_571838.1
DrCLDND	masvqlqla tvlaiigwlg eivicalpmw kvtafignni vtaqifwegl wmnvcvqstg qmckvydsm lalqdlqaa ralvisiiv tfmgvflta ggkctncied qdakavvva agvfflvggi lclipvcwta nsvikdfyn tldaqlrel gaslfigwga sglllglg lcccpkneg raysvkysap rsapgayv	NCBI: NP_851295.1
DrCLDNF	mgriakevag qlclfigvg iciccgipmw rvtyigani vtgqivwdgl wmnvcvqstg qmckiqdsi mnltdlqva ralvisili gfmgmlltfi ggqcsclkn essmakvli ggilcivagv lllipvcwta aytisdyvsv ttiqtqrel gasiyvgwga saflffggii lctscppren ippypgmyy yqgpyvghqy apsavysgin kpyhppvayt paprqyl	NCBI: NP_571841.1
DrCLDNG	mstglqlgt tglglwlg iiscaiqlwr vtafignni vtaqtmwegl wmscvvqstg qmckvydsm lalqdlqas ailvisaigv liamfasag gkctncladn sakalvattg gvafiiagil glvpswtan tiirdfynl vaeaqkrefg aaificwgaa vllviggl cssypkgrts srgrypasq ngrerseyv	ENSARP00000015898

DrCLDNH	msmgleggi algiigwiis ivacalpmwr vsafvganiv taqviweglw mncvqvstgq mqckvydsml algqdlqasr amtvaiila vlgymismvg akctnciede gakakvmivsgimfiagil dlipsawvan qiirdfynpl lpgaqqrelg asiyigfaaa alliigamlcctcppkek ykparmgysa prsasagydk kdyv	NCBI: NM_131767.1
DrCLDNI	mgsagvqivc valgilglia tvvtiaipqw ktsafiqnii itaqvseegl wmqcvqvstg qqqcksydsi lilssdlqaa ramtiiscml svlsllilcg gadfttcien edvkpvtlv saigililagl lviipvwaa nnvvrdfnnp mvpeaqkrel gpiciygwas gvllilaggl lccfsrprss gssgaakyys nsasapskny v	NCBI: NP_571843.1
DrCLDNJ	malqvlgitl smigfagti icalpmwkv afignivva qvfweglwmt cyerigqmq cklydalldl dpflqasrgl ivttmalasl aflilglad ctncnpra kgrivvsvgi tfmlsglttv vpvswtadi irdfhnppvh ealkremgaa lyvgwtagf lfvggailct scpperdnyl prytltksgt hsgyavknyv	NCBI: NP_571844.1
DrCLDNK	matgmqllg lvlsiilvg gflvctlpmw rvtafignni vtaqitwegl wmniciwstg eiqckgydsi lalpsdmkaa rglvtlaili cslstlgl gikctecvlg pslkarlarv sgvlfviagf lilvpvcwta hsiirdfydp yvaaphkrel gpalylgwga salllignsl lyagsnppgi pssptfssde ssprraggss qvkgvy	NCBI: NP_001003464.1
GgCLDN1	masgglqllg fvlaflgwmg iistampqw kmasyagdni vtaqalyegl wmscamqstg qiqckvydsi lklegslqat ralmvaail glvgvfvavt gmkcmkmed dqvkkmmav fggvifiag lsalvatswy gnrvarafyd pftpvntfr fgsalfigwa aasallgga lccscprse tsypsrsgyp knapstgkdy v	NCBI: NP_001013629.1
GgCLDN2	mvsnglqlvg yivaflyig tltilpnw kissyigssi vtavstkgi wmeatystg itqcdiysl lnppdiqaa qalmvsscav sslaciavv gmrctvfnqg spakdrvava ggvvfllgl lcfiplvwni hvvldfhnpl lpdstkfem gealyigsi sltliggf lcascpprdp sgypsprlla srspqpsikq mqpksefss ynltyv	NCBI: NP_001264551.1
GgCLDN3	msmgleggv alsvlglwls iiccalpmwr vtafigniv taqiiweglw mncvqvstgq mqckvydsml alpdlqaar allvvaiva vlglmvaivg aqctrcvede ttakativs gvifllsgim tlipvswan tiirdfynpl vidaqkrel tslyvgwaa alllfggall ccscppkder yapskvaysa prsavtsydk rnyv	NCBI: NP_989533.1
GgCLDN4	masmgqlvg ialsvigwla svlcalpmw retafvgnni vvaqiiwegl wmncvqvstg qmqrvydsi lalpdqlqaa rallvvaiv avlgtllaia ggkcttcved etakakmil sgitfivagv lilipicwsa naiirdfydp lvadsqkrel gsslyvgwaa salllfggal lccscpprge kpysakysaa rslpasnyv	NCBI: XP_003642430.1
GgCLDN5	masaaveilg lglgilwvg vilacglpmw qvsafidvni vvaqtiwegl wmcvqvstg mqckvydsi lalrevqag raltvival glvalmvttv gaqctncirp gmkksrivia ggtyilcgv lylvplcwfa nivisdfydp svppsqkrei gaalyigwaa talllfggcl icccscqlrd etsfpvkysa prptssgey ddknyv	NCBI: NP_989532.1
GgCLDN8	misgalqiag lllglgltmgt viltacglpmw kvssidgtvi ttatfwanlv ktcvtstgv rmqckvydsi lalpdqlqas rglmcagsvl sflafvatv gmkctrcrqs gwptkgyilv aagllfilsq ttalpicwv antiirdfyn pvvnvaqkre lgealylgwv aafcllaaga ifccfrcce ktrsynysmp pqhlplnqhl hgatesaysk sqv	NCBI: XP_004938436.1
GgCLDN10	mastsaeiva flitigwvl vssltptydvw kvssidgtvi ttatfwanlv ktcvtstgv snckdfpsml aldyiqacr glmisavclg ffgsvfvlvg mkctkiggsd qnkariacla glifilcglc smtgcslyah ritseffdp fvaqkyelga alfigwagas lciiggsifc fsiaenskp rrayayngaa svmsrtrkhi nsvpdktspk hfdknayv	NCBI: NP_001264696.1
GgCLDN11	mvatclhlag fvcfigwig vvvatatndw vvtcgytitt crkmdelgsk glwadcvmat glyhckplvd ililpgyvqa cralmiasv lglpaiflli tvlpcirmgh epgaakyrss qlggiliill amcgvvatiw fpvcahreth imsfygslyt gwigsalclf ggcvivccsg daqtfgnrf yyagsspspt haksahv	NCBI: XP_422797.1
GgCLDN12	mgcrdvhaat vflafsgtas vagilaavll pnwrqmrlt fnknemvttv ytlgwiccar fdgsrdcviy dpqwytavdq ldlrvlqfal plsmftavs lflcmigmcn tafvssvpnv klakclvnsa gchlvagllf llacaiclt siwvifynny lnryepvfv fdisvfiia sagglfttsi llflwycack slpspfwqpl yshapsmhsy asqpysarsr lsameidiv vths	NCBI: XP_004939337.1
GgCLDN14	mastavqlg fslglglg tiatilphw wrtahvgti itavaymkgi wmcvwhstg vyqcqvhrrs lalpdqlqaa rammvscil svlacviavi gmkctrcakg tsakasiavs ggivfilagl tclvpvswtt ndvvtdfynp llpqgmkyei gqalylgfvs asltlfggal lctsgqcgtd qgprqprcs trtapscrpp taykgnhass ltsashgyr lndyv	NCBI: XP_004934594.1
GgCLDN15	massslqics lllalagfti llltmsnrw kisdttavlv psewiseglw mdcgttdfga iqcmkflhml ssethvqacr tlmiiisllg fkaavsvllg lkctnvglsd edekvkvavt ggflfllgl csmvavswya amitapfnp lyagtkykg dalylgwags vlcmllggifl tcpckgkekq eysaskyrss aggasgqqqi ytedaetvss pkeyi	NCBI: XP_422589.1
GgCLDN16	mpldvtpag mrrflqyagc ffaftstgl iastwtcdwm vnaddslevs tkcrglwec vtvfdgiqt cdeydsifae hpvklvltra mmitadilag fgflfvlgl dcvkflpdep liklrcivs gvtliaglp gitgsvyai dvtyersslv fhniifgiy kfgwscwlgm agslglfsg sltccmylf retssgrlhs ayslrkyss agtvtvnhl pssqtatak yavdtrv	NCBI: XP_426702.1
GgCLDN18	msttlcqvmg flvsslgvvg cimstamdww stqldyndp tavfvyqglw	NCBI: XP_426691.2

	rsccvqssgf tecrpyftil glpamfqavr almivgivlg ilgfvfifa mkcirigtme dsakanmtlt sgimfiagl caitgvsvfa nmlvtnfwms tnmymqgmqn vqnrytfgsa lfvgwvagg alvggimmcl ackgmvpes rykavsynas grnysyrtgp ykagsgyepe tktrkgpye eaprsedagr sypskydyv	
GgCLDN19	mgggarelag ylaalggwva alaaalpqw rqsyyagdai itavglhegl wmscaaqtg qvqcrldsl lsldvhiqts ralmvislll gffgiivsvv gmkctkvgee dptksriav aggvlfilcg lctlaavslv atqvtyeffs astpinarye fgsalfvgwg aasltvlggs llccspake rggqgyyrs qpstarephv kissaireq cl	NCBI: XP_003642589.2
GgCLDN20	mhanailyek grshsretlf fedidseenn sdtstsys lgyeatnqlh afrekradk rgqqitmasa glqffafmla lfvgfvdiaa tllpnwkvna dvgsniitai tqmqglwmdc twystgmfc tkysvslp vyiqtartm vmcilsafg icittvgmkc tklggdpdsk shctfaggvc filagifglv ptswytrei snfldtipe sgkhepggav yigfisagfl lvagvifcts cfkkgqgawi yppkqqlps tmqessagyn lkdyv	NCBI: XP_004935700.1
GgCLDN25	mawscnarlq vgamlallfg wvsscvtfv plwknfnldl neleiwtmgl wqvcimqeeq ateckahdsf lalplelrvs rvlmclsngl gllalfasm glcdwktceg kpdakkqlll aggaafgaag ittlvpvswv aytvlefw d eavpdivprw efgeamflgw fagsflaagg lllystcls rtgtpaappa trlprvths patmspshh shpkndlvi	NCBI: XP_004948118.1
MmCLDN1	managlllg filaslwig sivistalpaw kiysyagdni vtaqaiyegl wmscvstg qiqckvfdsl lnlnstqat ralmvigill gliaifvsti gmkcmrcled devqkmwma iggiiiflis latlvatawy gnrvqefyd pltpinarye fgqaltgwa aaslclggv llscscprkt tsyptpryp kptpsgkdy v	NCBI: NP_057883.1
MmCLDN2	mgvkslqtg gnllnllsi ltlvstttny wtrqggghsg lwqecthgkc snipcqtva vsaacmvlaa tfsivalgig iriqcreas rrsqntivll flsgllllia lavytsknaw kpevfswsy ffgwlpfl fiagfcfla dmlqsteai sgfpvcl	NCBI: NP_083125.1
MmCLDN2.1	maslgvqlvg yilglllg tsiamllpnw rtssyvgsi vtagvfkgl wmcathstg itqcdiyst lglpadiqaa qammvtssam sslaciisv gmrcvfcqd srakdravv ggvfilggi lglpivawnl hgildrfysp lvpds mkfei gealygiis alflsvagvi lfcscpqgn rtmydygyqa gplatrssr saqqpkakse fnsysltgyv	NCBI: NP_057884.1
MmCLDN3	msmgletgt slavlgwlt ivccalpmw rvtafignsi vtaqtwegl wmcvqvstg mqckmydsl alpdlqaa alivvsilla afgllvalvg aqctncvqde takakitiva gvflaall tlvpvswan tiirdfynpl vpeaqkrem aglyvgwaaa alqllggall ccscpprdky aptkilysap rstgpgtggt taydrkdyv	NCBI: NP_034032.1
MmCLDN4	masmgqlvg islavlgwlg iilscalpmw vtafldhni vtaqtwegl wmcvqvstg qmckmydsm lalpqdlqaa ralmvisiiv galgmllsvv gkctncmed etvakimi agavfivasm limvpvswa hvirdfynp mvasgqkrem gaslyvgwaa sgllllggllccscpprsn dkpysakysa arsvpasnyv	NCBI: NP_034033.1
MmCLDN5	mgaaileil lclclvgwlg iilscalpmw vtafldhni vtaqtwegl wmcvqvstg hmckkyesv lalsaeqvaa raltgvall alvalvtlt gaqctcvap gpvkarvalt ggalyavcgl lalvplcwa nivvrefydp tvpvsqkyel gaalyigwaa sallmcgggl vccgawvctg rpefsvpky saprrtang dydkknyv	NCBI: NP_038833.2
MmCLDN6	mastglqilg ivtlglwn alvscalpmw kvtafignsi vvaqmvwgl wmcvqvstg qmckvydsl lalpqdlqaa ralcvvlll vllglvyla gaktctved mksrlvli sgiiifvisg ltlpvcwta hsiidfynd lvaadaqkrel gaslylgwaa sgllllgggl lccacssggg qgprhymacy stsvphsrgp seypknyv	NCBI: NP_061247.1
MmCLDN7	mansglqllg fsmamlgwg liastaipqw qmssyagdni itaqamykgl wmcvqtstgmmsckmyds lalpgalqat ralmvslvl glamfvatm gmktctregd dkakkariam tggivfivag laalvacswi ghqivtdfyn pltpmnvkey fgpaifigwa gsalvllgga llscscpgse skaayraps ypknskey v	NCBI: NP_058583.1
MmCLDN8	matyalqmaa lvlgvgmvv tvavtimpw rvsafiesni vvfennwgl wmcmrhani rmqckvydsl lalspdlqas rglmcaasvl aflafmtail gmktctregd denyksrill tagiiffitg lvlpvswv ansiirdfyn plvdvalkre lgealyigwt talvliagg lfccvfccte rnsyrysvp shrtqrsfh aekrspsys ksqyv	NCBI: NP_061248.1
MmCLDN9	mastglelg mtlavlgwlg tlvsclplw kvtafignsi vvaqmvwgl wmcvqvstg qmckvydsl lalpqdlqaa ralcvvlll allglvait gaqctctved egakarivlt agvlllsgl lvlpvcwta haiidfynd lvaalkrel gaslylgwaa aallmlgggl lccctppshf erprglrgy siprsrgasg ldkrdyv	NCBI: NP_064689.2
MmCLDN10	mastaleiva fvvisgwl vstlptdyw kvstidgtv ttatyfanlw kicvtdstg anckefpsml aldgyiqacr glmiaavslg ffsifalgl mkctkvgsd qakakiacla givfilsclg smtgcslyan kitteffdl ymeqkyelga alfigwagas lciiggvifc fsisdnnkt rmgtyngpt samssrtkyq ggegdfktg pskqfdknay v	NCBI: NP_067361.2
MmCLDN11	mvatclqvvg fvtsvfwg ivtvtstndw vvtcsytip crkmdelgsk glwadcvmat glyhckplvd ililpgyvqa cralmiaasv lglpailll tvlpcirmgh epgvakyr qlagvlll alcaivatiw fpvcahreit ivsfygslya gwigavmclv ggcvivccsg daqsfgnrf yyssgsspt haksahv	NCBI: NP_032796.1
MmCLDN12	mgcrdvhaat vlsflcgias vaglfgatll pnwrklrlit fnmeknlti ytlvkwcar	NCBI: EDL14638.1

	ydgsdclmy drtwylsvdq ldlrvlqfal plsiwiamga llcligmcn tafnssvpni klakclvnsa gchlvaqlf flagtvslsp siwaifynsh lnrkfevpt fdyavfvta ssglftmlt llfwycack slspsfwqpl yshapgmhty sqpyssrsrl saiedipvv shst	
MmCLDN13	mvsvkqeaissvtslgwvg aivscvlpvw rvtppddetd pdatiweglw hicqvrenrw iqctlydtri lvaqdikvsr vfmvictigt wlgllcvlg dwrincfmnf tieenllkva ggmlfsvgl mlvplswvth niihgfnpl lgfskkvqmg ssllawtss lllllggill cvnipvcrdf prcietspar psganndtld v	NCBI: NP_065250.1
MmCLDN14	mastavqllg flsflgmvg tlittlphw rrtahvgtni ltavsyklg wmecvwhstg iyqccqyrsi lalprdlqaa ralmviscll sgmacacavv gmkctrcakg tpakttfavl ggalfilag lcmvavswtt ndvvqnfynp llpsgmkfci gqalylgfis sslliggtl lclscqdeap yrpyppqsra gattatapa yrppaaykdn rapsvtsaah sgylndyv msvavetfgf fmsalglml gltlnsywr vstvhgnvit tntifenlwy scatdsgvs ncwdfpsmla lsgyvqgcr lmitailgl lglflgmvg rctnvgnmdl skkakllaia gtlhilagac gmvaiswyav nittdffnpl yagtkyelgp alylgwsasi lsilggicvf stccsskee patraglpyk pstvviptat sdesdisfgk ygknayv	NCBI: NP_062373.3
MmCLDN15	msvavetfgf fmsalglml gltlnsywr vstvhgnvit tntifenlwy scatdsgvs ncwdfpsmla lsgyvqgcr lmitailgl lglflgmvg rctnvgnmdl skkakllaia gtlhilagac gmvaiswyav nittdffnpl yagtkyelgp alylgwsasi lsilggicvf stccsskee patraglpyk pstvviptat sdesdisfgk ygknayv	NCBI: NP_068365.1
MmCLDN16	mkdllyaac flafstgfl ivatwtdcwm vnaddslvs tkcrglwwec vtnafdgirt cdeydsiay hplklvtra lmitadilag fgfitllgl dcvkflpddp qikvrlcva gttliagtp giigsvwyav dvyversslv lhnifligy kfgwscwlgm agslgcflag alltcclylf kdvgpennyp yamrkpysta gysmaksyka prtetakmya vdrtrv	NCBI: NP_444471.1
MmCLDN17	mafypqlqag lvgffglvg tigtllpqw rvsafignsi iiferiwegl wmcicqqamv tlqckfyysi lalppvlea ralmcvaval alvaliigic gmkqlqctgs servkayllg tsgvlfiltg ifvlpvswt aniiirdfyd pthvagqkre lggalfgwa taavfiggg llcgycconr kerwhryppv ayrvpqkdnq mvtvprkss tsyv	NCBI: NP_852467.1
MmCLDN18	mattecvqvg lllsllglag ciaatgmdmw stqldydnvp tavfyqeglw rscvqqssgf tecrpyftil glpamlqavr almivgvlvlg vigilvsifa lkcirigsm dskakmtlt sgilfiisgi caiigvsvfa nmlvtnfwms tanmysgmvg mvgmvqvtvt rytfgaalfv gwwaggltil ggvmnciacr gltpddsnfk avsyhasqgn vayrpggkfa stfgsntrn kkiydgart edeqshptk ydyv	NCBI: NP_062789.1
MmCLDN19	mansglqllg yfalggwvg iastalpqw kqssyagda itavglyegl wmscasqstg qvqcklydsi laldghiqsa ralmvavil gfvamvsvv gmkctrvgds nptaksvai sggalfilag lcltavswy atlvteffn pstpvnyare fgpalvfgwa saglamlggs flcctcepe ransipqpyr sgpstaarey v	NCBI: NP_694745.1
MmCLDN20	masagllqlla filavsgvg vlaatlpnw kvnayagpni vtavvqvqgl vwdctwystg mfsctlkysi lslpvhvqta ratmvlacvl sawgietsia gmncthlggd thtkiskisa ggvcfitagi sgliptvwyet keianfldl tipeshkyep ggalyigfis amllisgvi fctsyiqknq epwiypkqk lttwqpknr rahnlkdym	NCBI: NP_001095030.1
MmCLDN22	mglvfrtatq aaallslg wvlscnllyl phwknlnlel nemenwtmgl wkscviqueev grqckdfdsf lalpaqlvs rvlmslcnlg lglgllasgc gldclrlget qeglkrrlt lggllwtsg vmvlpvswv ahktvrefwd etmpeivprw efgealfgw fagfclvlgg cvlhcaacs papaashya vagprdhqgh lelqanpei	NCBI: NP_083659.1
MmCLDN23	mrtpvmtlg mvltpcglll nlvstlapgw rlvkgfldqp vdvvlyqglw dicreqssre recgqpdown yfqtqpqvva rglmitslat talgllas gvrwcqdeph yglaglsgrv ffvagfslsi pvswnhfls dpdvlaapss pvtvqvsyl vlgylgscil lggfslals fapwceercr rcrkappagp rrsistvyv dwpealtpa ikyysdgqhr ppptaehrtd sklkvgfmp rpppkstytp mdvlegeekk tatsqggss rstrpcqnsi pcddsl	NCBI: NP_082274.1
MmCLDN24	mafifrtamq svglslslg wvlaityl phwknlnlel nemenwtmgl wkscviqueev grqckdfdsf lalpaqlvs rvlmslcnlg lglgllasgc gldclrlget qeglkrrlt lggllwtsg vmvlpvswv ahktvrefwd etmpeivprw efgealfgw fagfclvlgg cvlhcaacs papaashya gagprdhgsy lengtvqpkv	NCBI: NP_001104788.1
MmCLDN25	mdnrfatav iacvlslist iymaasigt fwyeyrsipq enssdsnkia wedflgdead ektyndvler yngslglwrr citipknthw yappertesf dvvtkcmsft lneqfmekyv dpghnsgid llrtylwrcq flpfvslg mcfgaliglc acicrslypt latgilhlla glctlgsvsc yvagiellhq kvelpkdvsq efgwsfclac vsaplqfmaa alfiwaahtn rkeytlmkay rva	NCBI: NM_001252450.1
MmCLDN26	mrvtlqalag aalsgalsf vllaaigt fwyiiderl erssqrmrdq gpanrsqqep lsshsqglwrt crvqssctpl mnpfwqenvt vsdssrqlt mhgtfvillp lslivmvfgg mtgflsflr ahlllltgi lflfgamvtl tgsiayis avafreavcl leeralldqv dirfgwslal gwisfvsell tgvvflaaar alsisqrqda ai	NCBI: NM_029070.2
XICLDND1	mdnrfatalv igsvslslsv iylsaavgtv swyhyftssa qanisytpe dftsdaekmd ekaytdalfh fngslglwrr citvpqdhs qqlsaltsp tlfscvflsi sdqflekyqe pgnhnseidi vrytlwrcq flplvalgli ffgalvlag cvcqslpal gtgilhlag vctlgsvlcf ssgihmleqr lplptgirge ygwsfclac csplqvmaga lflwasrstr reyslmkayr va	NCBI: NP_001016259.1
XICLDN1	managllqllg falacglwig fivciaipqw kmssfagda itaqityegl wmscvmqstg qmqcktydsi lkldtmqat ralmicgilv gffamciaav gmkcltclqd devkkakvgv vggalfivag lcvliatawy gnkiakdfyn vftptnskye fgpalvfgwa gaalailgga llccsppre tsyppprgyn ksappagkdy v	NCBI: NP_001015704.1
XICLDN2	mtsvglqmmg yfmgilglvg svvatlmpnw kvssyvgasi vtavgfsgl wmecasytg	NCBI: XP_004916813.1

	itqcdiynsm lglpsdtqaa qalmitscvl sslaglsif gmktctfsqd spgkakaiv ggvlfilggv lclvpicwnl hslrldfynp iadsmryei gpalylgims sivsylvgsi lcascppree ghafysryqr kplvankgan atvltetskr eqsgynltg v	
XICLDN3	msmgleilgv alsivgwlg tviccalpmwr vtafignniv vaqtiweglw mncvqvstgq mqckmydsll alpdlqaar almvisivia vlglvisiig akctncvqde sakakimivs gvifilsglm tlipvswsan tiirdfynpl vvdaqkrelg ssmyigwaas allmlggaml ccscppkeky ptsrvaysaa rstnpgydrk dyv	NCBI: NP_001005709.1
XICLDN4	masmgqlvlg ialsvigwlg tviccalpmw rvtafignni vvaqiiwegl wmcvqvstg qmckkvydsl lalpqdlqaa ralvvaiav avlallmai ggkctncved esskakvmiv agvvfivagv ltlipvswsa nniirdfynp lvvdaqkrel gaslyigwaa aallmlggam lccncpprdq kpyasakytas rsgatsnyv	NCBI: NP_001016663.1
XICLDN5	masaaieilg lslsilgwvg vilacglpmw qvsafienni vvaqiiwegl wmcvqvstg qmckkvydsi lalsqelqag raltvmasiv gligllvtiv gaktncldg nsakgrvlla ggviyilcgi lvliplcwia nniitefydp rvpsqkrem gaalyvgwaa tallllggs lccsfamkdg isnlpvkysa prmpstngdy dkknvy	NCBI: NP_001006707.1
XICLDN6	mastglqilg malaligwvg siitcalpmw rvtafignni vvaqiiwegl wmcvqvstg qmckkvydsm lalsqdlqaa raltviciav allailigv gaktncied entkakvsmi sgivfvvsg lmlipvcwsa nsiirdfynp lvveaqkrel gaalyigwaa sgllmlggal lccscpkdd nhysaryta tsnprsdyps knvy	NCBI: NP_988873.1
XICLDN7	mansglqllg fvlallgwia liaatimpqw kmssyagdq itavaiyqgl wmscatqstg qiqckvydsl lqldaslat ralmvsviil gvfglaistm gmktncggd dkvksriam tggfvlilgg laaliacsww gnqiirdfid evckghwggf	NCBI: NP_989077.17
XICLDN8	malqlvlgvl ggigligta vtgmppqrwt afidnnivvf eaqweglwmm cvrqanirmq ckvydslal tpdldqagral mcvavcltfl sfmiaiigmk ctvcvgdnar tkgiillvag itfilsgivv lipvswtngq iirdfynplv lssqkrelgd alyigwtal vliagglilc ctfrsgekev ryslppksvt sappksais vpirkpsly sksqv	NCBI: NP_001120297.1
XICLDN10	mahmaseiia filavsgwvl vssltptdyw kvssidgtvi ttatfwsnlw ktcvtdstg snckdfpsml aleghiqvcr gliaavsfv fgsifalig mkctkiggsd ntakitcfs gllyilsgic amagvslyah kiitseffdt fleqkyelga alfigwggss lcviggavlc fslndikap ritysykgvt svrsarsniq nsgngqnaqr gvprnfdkna yv	NCBI: XP_002939572.1
XICLDN11	mggtclqffg fissvigwig vvvattndw vltcdysint crkldegsr glwadcvsv glyhckplad ilalpahiqa aralmitssf lglpavfill sampfirmgh dpgtekhkrf rlggilvvl alctiatiw fpvsthrett ivtfgysly gwigaaclf ggsvivccag eshafmenrf ysssqgsgsp thaksanv	NCBI: XP_002931594.1
XICLDN12	mgcqvahavt lfafvcgtas isglfaatl pqwrqmklyt ynmeknlv siglwkctr lewsrecmmy dmewyasvdq ldirvlqfal pfsiltaasa lilcltgin tftsgvynl klvklvnss gchlvagllv ilagimsimp siwsvfynsv lntkygpyft ydiavfaig ssgmfftav llflwyack slpsfwqpl yshvpsmhsy vppsyrqr msaiedipv vtha	NCBI: NP_001016851.1
XICLDN14	masmalqllg fsvaligfig tivatlphw wrtahvgnt itavaymkg lwmecvwhstg iyqcqvhqsq lalprldqva rammvascvl svlasvsvf gmktqcakg ssskrmiaaf ggftfalagl mclipawst ndvvqdfynp glpygmkyei gqalyvgfis gglsviggim ilstscqrnd tlpypqpqr yprkapsrs qpvknsnhvp swssasrhgy hlndfv	NCBI: XP_004912176.1
XICLDN15	msviletigf ffgvigstml gidfianwr vssvdgnvit tsllfenlwm scatdstgy ncrefslia lsgylqasa lmitlsvfvg fscllgmigl qctkvaednp vvkglataa gaisifggc agiavswyaf nitkdffnpl ypgtkeyigp alylgwsgal ldivggailc csyktqvsaa knkytynra pmtvssakaa ppatlpyes snsickeygn ayv	NCBI: NP_001096379.1
XICLDN16	mrtleytgf flafvsagfl iaatwtcdwm enaddslvs tkcrglwec vtnvfdgirt cdeydsiae hpmklvatrg lmitadilag fgfiillgl dcvkftdes dvklkicyva gitlligvp giigsvwyai dvyverstv ihnvflgiy kfgwscwlgm igslgclfcg gililchylf kengygrnhf ssrkytllrg gsaakfdshp rtetakmyav dtiv	NCBI: XP_002934087.1
XICLDN18	msvtmctmg flvsalgfag iivataldpw stqldynpv tavfyqqlw kscvqessgf tecrpyytil glpamfqavr almivgvlv gigllvaifs mkcirigsme dsakanitlt sgimfilagl csiigsvvfa nmlvtnfwm tanmytgga sgmggmglg tlqtrytfga alfvgwvagg ltmiggvmmc iacrglmpee snykavsyhv stktpaykts aydekskksi ynesrsedg ksypskydyv	NCBI: XP_002938162.1
XICLDN19	mansglqllg yfalaggwig iistalpqw kqssyagda itavglyegl wmscasqstg qvqckvydsl lslevhiqt ralmvvaill gfvgiisv gmktkvgd nptksriav sgvlfilag lctliavswy atqvthdffn pntpvarye fgsalvgwa sasltmlggs flccscpted rrqqyyrqs qpsattreit kmpakkeet s	NCBI: XP_002937080.1
XICLDN19L	maaaieigf lmslagnlit gvtlssnrwk tstvhgsvit tsnyvenlw scaedstgis nckkfdsls lpvhiqvcra lmiislvl mscivsigl kctkfgtsne qtkgkialsg glifilagll tltavswyaa mitaqffdl yltkyelgs alyigwaas lsilggsflc csfkkkkpat kagaykyny dpdsdfkqk erketsvask ayv	NCBI: XP_002933422.2
XICLDN23	mrtpvvmiig mvlapcglvl nlctvapnw rqlsqtgkp qdvvehqgiw dicdetvtnq nkvcgisnte yglkvvmva ralmtslv talgiivaal gvrwqdeph fvlsgigglv ifisvlsli pvswnngmy slstdssi tiqvdyalvl gylgscfei ggflslsfv rgckdcigr rksptstlyy kqnpmhsgn psvysirt rdnlgyvsgn dsvyqndrr yapevsvpk sysnprvid kehsrgprp asqlslpcd sdll	NCBI: XP_004911118.1
XICLDN22	malvkrtnvq fggillsitg wilscvtyi pvwknlnel nelenwtmgl wqtcvqvdeg pmqckdfdsf lalselkis rilmlsngs giigllisf gldecrigkg nqaqknrlil	NCBI: XP_002934304.1

	fggimlwtsg lsalvpvswv ayetvqefwd etipdivprw efgealfmgw fggffliilgg sllaslycfa radhapvyyt ptqqheklh ldpkypdlti	
--	--	--