Characterisation of transmembrane protein 114 (TMEM114), a protein associated with juvenile onset cataract.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2011

Geoffrey Joseph Maher

School of Medicine

Table of Contents

Final word count = 59,077

List of Figures

List of tables

Abstract

Characterisation of transmembrane protein 114 (TMEM114), a protein associated with juvenile onset cataract.

Geoffrey Joseph Maher, The University of Manchester, PhD in Genetic Medicine, final version submitted January 2011.

Transmembrane protein 114 (TMEM114) is an uncharacterised predicted transmembrane protein expressed in the lens epithelium. A balanced translocation that transects the putative promoter of *TMEM114* is associated with autosomal dominant congenital cataract (ADCC), however, coding sequence variants in *TMEM114* were not identified in a panel of ADCC patients. Subsequent to the identification of TMEM114 a similar novel transmembrane protein named TMEM114-like protein 1 (TMLP1) was identified. This study aimed to functionally characterise the two novel proteins TMEM114 and TMLP1.

TMEM114 and TMLP1 showed homology to voltage dependent calcium channel gamma (γ) subunits, but TMEM114 and TMLP1 lacked some of the key domains present in these proteins. Expression of *TMEM114* and *TMLP1* in the developing human eye was identified. TMLP1 was also expressed in developing neural tissue. To aid functional characterisation, the murine orthologues of Tmem114 and Tmlp1 were cloned and polyclonal antibodies were generated. Bioinformatic tools predicted co-and post-translational modifications. The predicted plasma membrane localisation of Tmem114 was confirmed *in vitro* in polarised MDCK II cells and the membrane localisation was shown to be dependent on the presence of N-linked oligosaccharides. Murine Tmlp1 was localised in the endoplasmic reticulum (ER) in MDCK II cells, possibly due to the fact that murine Tmlp1 lacks an N-glycosylation site present in other species.

The knockdown of Tmem114 in *Xenopus tropicalis* using antisense morpholinos results in microphthalmia (small eye) confirming the protein's role in eye development and growth. Sequencing of patients with microphthalmia and anophthalmia identified a novel heterozygous missense mutation (p.R2Q) not present in controls. The p.R2Q variant was not found to affect the expression or localisation of the protein. The p.A147V variant, previously reported as a SNP, mis-localises to the endoplasmic reticulum.

In conclusion, this study identifies that TMEM114 is a transmembrane glycoprotein with an important role in ocular development in vertebrates, although its precise functional role remains to be elucidated. Knockdown of Tmem114 in *X. tropicalis* results in microphthalmia, suggesting that loss of function mutations in *TMEM114* may be associated with human ocular disease. Further characterisation of TMLP1 is required to identify its role in human biology.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on presentation of Theses

Acknowledgments

I would first of all like to thank my supervisors Dr Forbes Manson and Prof Graeme Black for giving me the opportunity to undertake this project. Thank you for your unrelenting guidance, support and encouragement. Thanks also to my advisor Nicoletta Bobola for her guidance over the years. Many thanks also to Fight for Sight who funded this project.

Many thanks to present and past members of the Manson/Black group, in particular Dr Jill Urquhart and Dr Emma Hilton. Jill undertook the initial functional work on TMEM114 and passed on constructs and her knowledge which made my work that bit easier. I would like to thank Emma for her general help and advice aswell as *Xenopus* expertise. Your help with morpholino injections and assessment of phenotypes is greatly appreciated. Thanks also to Dr Jim Bellingham for his advice and general enthusiasm for science, to Rob Lea for teaching me the technique of *in situ* hybridisation and to Dr Kathy Williamson and Dr David Fitzpatrick for providing the patient DNA panel.

A big thank you to my parents who have supported me in everything I've ever done and to my brothers for their interest and support over the years. Thank you also to other relatives, particularly Aunt Margaret and Aunt Rita, for their continued interest and support. To my friends Neil, Brendan, Miles and John and all others who listened to my rants about failed experiments and to other friends who I haven't seen enough of the last few years. Thanks also to Julian Reindorp for his proof-reading skills. And finally, to Alice who has helped me more than she'll ever know.

Abbreviations

ADCC	autosomal dominant congenital cataract
AMPAR	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor
BAC	bacterial artificial chromosome
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BMP	bone morphogenic protein
bp	base pairs
c.	coding DNA
CDK	cyclin dependent kinase
cDNA	complementary DNA
CLDN	claudin
CNS	central nervous system
CX	connexin
Da	Dalton
DAPI	4,6-diamidine-2-phenylindoledihydochloride
dH_20	distilled H_20
DIG	digoxigenin
DLAD	DNase II-like acid DNase
DLGA	Drosophila disc large tumor suppressor
dNTP	deoxyribonucleoside triphosphate
DMEM	dulbecco's modified Eagle's medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ECL	extracellular loop
EDTA	ethylene diamine tetraacetic acid
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EGTA	ethylene glycol tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmic reticulum
ERAD	ER-associated degradation
EST	expressed sequence tag
EtOH	ethanol
F	forward
FBS	foetal bovine serum
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
g	gram
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein

GluR	glutamate receptor
hCG	human chorionic gonadotropin
HEK293	human embryonic kidney 293
HEPES	N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
Het.	heterozygous
HMM	Hidden Markov Model
Homo.	homozygous
HPRT	hypoxanthine phosphoribosyl transferase
hr	hour
HRP	horseradish peroxidise
IF	immunofluorescence
IGF	insulin-like growth factor
IGEPAL	tert-Octylphenoxy poly(oxyethylene)ethanol
IHC	immunohistochemistry
IMS	industrial methylated spirit
IOL	intraocular lens
IPTG	isopropyl-β-D-thiogalactoside
KLH	keyhole limpet haemocyanin
l	litre
LB	Luria Bertani
LEDGF	lens epithelium-derived growth factor
M	molar
m	meter
MAB	maleic acid buffer
MAC	microphthalmia/anophthalmia/coloboma
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation-Time of Flight
MDCK	Madin-Darby canine kidney
MEMFA	Minimum Essential Medium with formaldehyde
MetOH	Methanol
min	minute
MLPA	multiplex ligation-dependant probe amplification
MMR	Marc's Modified Ringers
MO	morpholino
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger RNA
MS222	Tricaine mesylate
NBT	Nitro blue tetrazolium
NEP	New England Peptide
NHS	Nance-Horan syndrome
nLC-MS/MS	Nano-scale liquid chromatographic tandem mass spectrometry
NN	neural network

NT	non-transfected
OCT	optimal cutting temperature
OFZ	organelle free zone
OD	optical density
OMIM	Online Mendelian Inheritance in Man
O/N	overnight
p.	protein sequence
PBS	phosphate buffer saline
PBS-Tw	phosphate buffer saline-Tween20
PCR	polymerase chain reaction
PDGF	platelet derived growth factor
PDZ	PSD95/DLGA/ZO-1
PLE	presumptive lens ectoderm
PMSG	pregnant mare serum gonadotropin
PSD-95	post synaptic density protein
R	reverse
RCF	relative centrifugal force
RIPA	radioimmunoprecipitation assay
RNA	ribonucleic acid
RNase	ribonucelase
RPE	retinal pigment epithelium
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcription-PCR
s	second(s)
SDM	site directed mutagenesis
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SFK	Src family tyrosine kinases
SNP	single nucleotide polymorphism
SPM	signal peptide mutant
SRP	signal recognition particle
SSC	saline-sodium citrate
TAE	Tris-acetate-EDTA
TARP `	Transmembrane AMPA receptor protein
TBST	Tris-Buffered Saline Tween20
TE	Tris EDTA
TEMED	N,N,N',N'-Tetramethylethyenediamine
TER	transepithelial resistance
Tm	annealing temperature
TMD	transmembrane domain

TPM	transcripts per million
WB	western blot
WHD	WAVE homology domain
U	unit of enzyme
UPP	ubiquitin proteasome pathway
UPR	unfolded protein response
UTR	untranslated region
UV	ultraviolet
v/v	volume in volume
WHO	world health organisation
WT	wildtype
w/v	weight in volume
X-gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside
Y2H	Yeast two-hybrid
ZO	zonula occluden

Magnitudes

p pico

n nano

µ micro

m milli

k kilo

Chapter 1: Introduction

1.1 Epidemiology of Cataract

In 2002, 37 million people worldwide were affected by blindness, with another 125 million visually impaired (Pascolini *et al*. 2004). In developed countries age-related macular degeneration is the most frequent cause of blindness but worldwide cataract is the leading cause, responsible for almost half of all cases (Pascolini *et al.* 2004). Cataract is an opacification of the lens resulting in visual impairment. The formation of cataracts is generally age-related but congenital and juvenile onset forms also occur, albeit less frequently. In developed countries it is estimated that 3-13.6 children per 10,000 live births are diagnosed with cataract by the age of 15, the majority of which are diagnosed in early childhood (Rahi and Dezateux 2001; SanGiovanni *et al.* 2002; Holmes *et al.* 2003; Graw 2004).

In developing countries incidence rates of juvenile cataract are higher due to environmental causes such as rubella infections during pregnancy and due to recessive traits exposed by consanguinity. Inherited forms of congenital cataract are estimated to account for 50 % of cases (Francis and Moore 2004). Congenital cataracts may develop as part of a syndrome such as Lowe (OMIM 309000), Martsolf (OMIM 212720) and Nance-Horan (OMIM 302350) (see section 1.6.3), and are also often seen in association with other anterior ocular diseases such as microphthalmia and microcornea (Abouzeid *et al.* 2009). Cataract is more frequently inherited independently, most commonly as an autosomal dominant trait and occasionally in an autosomal recessive or X-linked manner (see section 1.6) (Francis *et al.* 1999).

1.2 Role and structure of the lens

1.2.1 Role of the lens

The lens is a highly specialised ovoid tissue in the anterior segment of the eye, suspended behind the pupil by zonular fibres from the ciliary body (Figure 1.1A). In conjunction with the cornea, the lens functions to transmit and focus light onto the retina, which transmits nerve impulses to the brain, being interpreted as vision. Thus, transparency and refraction of the lens are crucial to visual acuity. These optical properties are achieved by the highly organised structure and protein content of the lens. Disruption of the refractive properties of the lens by altering the protein or cellular organisation results in cataract.

1.2.2 Structure of the mammalian lens

The human lens is biconvex with the posterior surface more curved than the anterior surface. The lens is surrounded by the lens capsule, a thick layer of basement membrane (secreted by the lens epithelial cells), which is bathed by the aqueous humour on the anterior side and the vitreous humour on the posterior side (Wederell and de Iongh 2006). The majority of the avascular lens is composed of differentiated lens fibre cells which are surrounded by a layer of cuboidal epithelial cells on the anterior side (Figure 1.1B). The centre of the lens, the lens nucleus, is formed during lens development from differentiated posterior epithelial cells (Section 1.3.1). The surrounding cortical fibres are formed from anterior epithelial cells which divide and differentiate to form fibre cells, a process which continues throughout life in

vertebrates (Section 1.4) (Gordon and Donzis 1985). As there is no cell turnover in the lens, the lens is ever-growing with newly differentiating cells surrounding the older cells at the centre (Gordon and Donzis 1985).

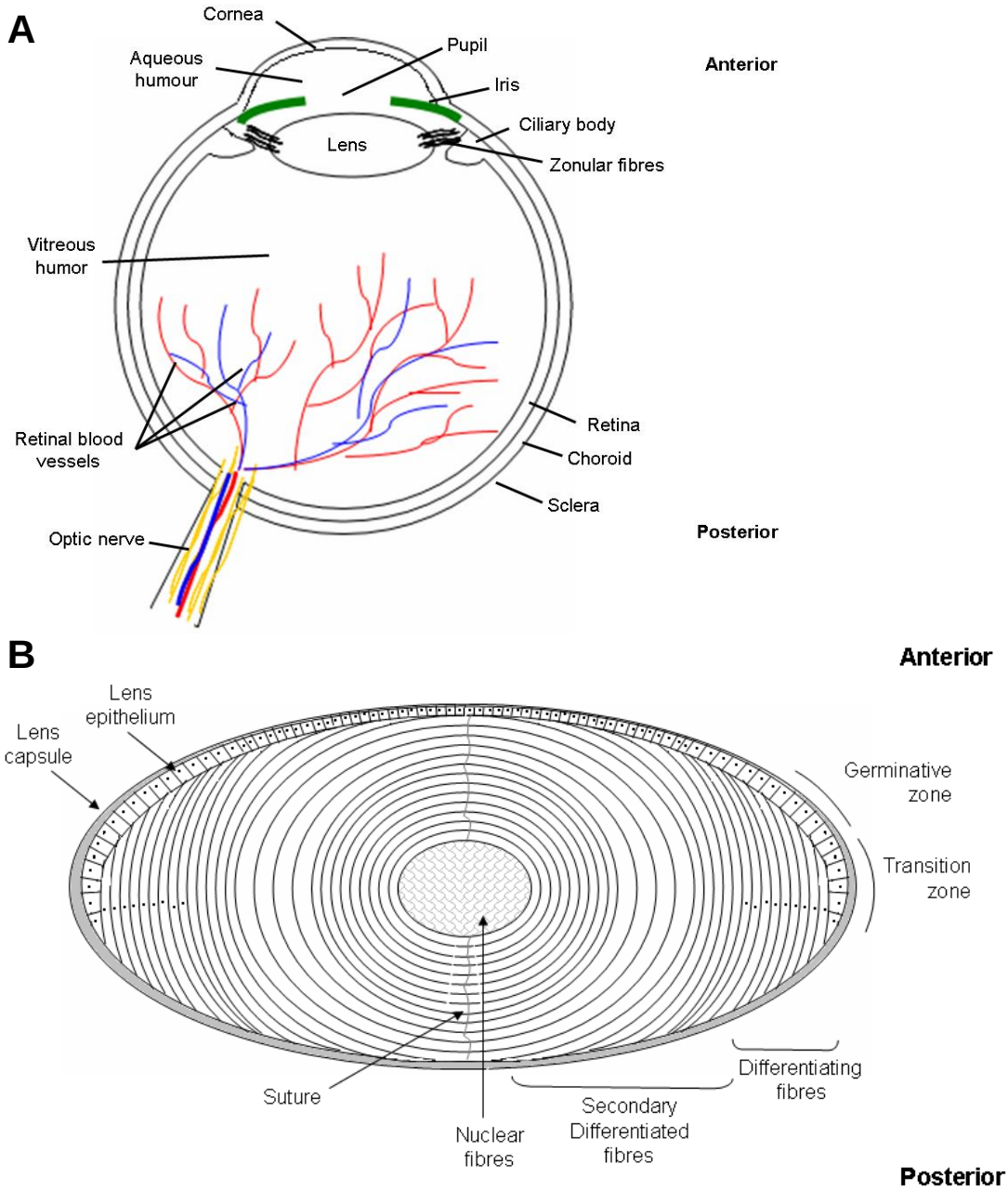


Figure 1.1. Anatomy of the mature human eye and lens. A) Cross-section of an adult human eye. The lens is located in the anterior segment, attached to the ciliary body by zonular fibres. B) Cross-section of an adult lens. The lens is comprised of two main cell types: epithelia and fibre cells. Central epithelial cells at the anterior of the lens are mitotically quiescent, have a flat morphology and become more cuboidal towards the germinative zone. In the germinative zone the cells divide, and in the transitional zone the post-mitotic cells begin to elongate. The apical ends of the differentiating cells then turn and point towards the anterior. The elongating differentiating cells lose their cellular organelles and become internalised as more cells differentiate. Elongation continues until opposing elongating cells meet and form sutures. Differentiating cells are continually concentrically added throughout life to the inner nuclear fibres which are laid down early in development.

1.2.3 Refractive properties of the lens

Most living tissues appear non-transparent as the difference in refractive index between and within cells (i.e. between cytoplasm and membrane-bound organelles) results in light scatter. In order to remain transparent, the lens has a number of adaptations to minimise the absorption and scatter of light. Cells in the equatorial region of the lens display little adaptation as these cells are located behind the iris so are not on the visual axis. The lens epithelium is composed of a thin single layer of cells at the anterior and as this layer of cells is so thin, the light scattering caused by it is minimal (Bassnett 2009).

Differentiating and mature fibre cells use different mechanisms to prevent light scatter. Differentiating fibre cells at the periphery are uniform in size and hexagonal in cross-section (Bassnett and Winzenburger 2003). The hexagonal structure allows close packing of the fibre cells to such an extent that the space between the cells is smaller than the wavelength of light, minimising light scatter (Michael *et al.* 2003). The differentiating cells also elongate drastically (Section 1.4.4) and elongation increases the cytoplasm to nucleus ratio, reducing the contribution of nuclei to light scatter (Coulombre and Coulombre 1963). As fibre cells further differentiate they lose their organelles (Section 1.4.6), reducing light scatter associated with membrane-bound organelles (Bassnett 1997).

The central fibre cells are much more irregular in morphology and in order than the peripheral fibre cells (Bassnett and Winzenburger 2003). These cells are connected by numerous tongue and groove inter-digitations (Okinami 1978). This irregular order should increase the refractive index but light scatter from the irregular-shaped cells is prevented by equilibrating the refractive indexes of the cytoplasm and the cell membrane (Michael *et al*. 2003). The lens nucleus contains a high crystallin concentration (90 % of all water soluble protein) which increases the refractive index of the cytoplasm (Keenan *et al.* 2008). The refractive index of the fibre membranes is reduced to match that of the cytoplasm by an increase in sphingomyelin content and cholesterol to phospholipids ratio of the membrane (Li *et al.* 1985; Greiner *et al.* 1994; Tholozan and Quinlan 2007).

The centre of the lens has a uniform refractive index (~1.41) which is higher than that of the periphery (~1.37) (Jones *et al.* 2007; Kasthurirangan *et al.* 2008), and a gradient of refractive indexes occurs from the outside of the nucleus to the lens periphery (Jones and Pope 2004; Augusteyn *et al.* 2008)). This gradient reflects the protein concentration in the lens (Fagerholm *et al.* 1981) and is necessary to correct the longitudinal spherical aberration of the lens (Fernald and Wright 1983). Syncytia form between the lateral membranes of differentiating cells of the same age and remain in the mature fibres (i.e. from the same stratum) (Shestopalov and Bassnett 2003; Shi *et al.* 2009). Absence of syncytia is associated with altered refractive properties of the lens and it has been proposed that the presence of syncytia make a more even distribution of protein between cells helping to create a more uniform refractive index in cells of the same stratum (Shi *et al*. 2009).

The lens can alter its refractive properties by the process of accommodation. In order to focus light from near objects on to the retina, the lens increases refractive power by changing its shape to become more globular (Jones *et al*. 2007) and by reducing the distance between it and the cornea (Koretz, Cook, & Kaufman, 2002). This movement is controlled by the ciliary body and the zonular fibres that attach to the lens.

1.3 Mammalian eye and lens development

Development of the eye is a complex process involving the interaction of neuroectoderm, surface ectoderm, and mesoderm of the developing embryo. Late in gastrulation (mouse embryonic day 7 (E7); day 17 human gestation) precursors of the eyes, the eye fields, develop in the anterior neural plate. Growth of the diencephalon causes the eye fields to move forward (E8/day 21) where they evaginate forming optic pits. The optic pit evaginates further forming the optic vesicle which is joined to the neural tube by the optic stalk.

The optic vesicle extends, displacing the mesenchyme between it and the overlying surface ectoderm. The resulting interaction between the neural ectoderm of the optic vesicle and the surface ectoderm causes the surface ectoderm to thicken, forming the lens placode (E9, day 27) (Figure 1.2A) (Pei and Rhodin 1970). This is the first step in lens formation and is a key stage in eye development. The hyaloid artery, which later in development regresses, supplies the developing optic vesicle with nutrients (Beebe 2008). This artery causes differential growth in the lens placode, with the cells near the nutrient supply multiplying much quicker, causing the lens placode to invaginate forming the lens cup (E10.5, day 29) (Figure 1.2B,C) (Graw 2003). Filopodia connect the lens placode to the presumptive retina and so invagination of both tissues is co-ordinated (Chauhan *et al.* 2009).

The almost spherical lens vesicle is formed when the lens cup closes, creating a large central lumen (E11.5; day 31-35) (Graw 2003). The lens vesicle transiently remains attached to the surface ectoderm via the lens stalk. This surface ectoderm, with neural crest-derived mesenchyme, will later develop into the cornea (Hay 1979). Mice lacking lenses undergo skin differentiation in presumptive corneal epithelial tissue suggesting that signalling from the lens is required for corneal development (Zhang *et al.* 2008). During the formation of the vesicle, the neuroectoderm surrounding the invaginating lens epithelium folds over forming two layers (Figure 1.2C). The inner part, surrounding the lens vesicle, will later develop into the neural retina, whilst the outer layer will give rise to the retinal pigment epithelium (RPE). Contrasting with results in chicken which suggested that the presence of the lens is necessary for correct retinal development (reviewed in (Chow and Lang 2001), differentiation of the major retinal cell types was not affected in mice in which the lens placode was ablated, although ectopic RPE formation occured (Zhang *et al*. 2008).

.

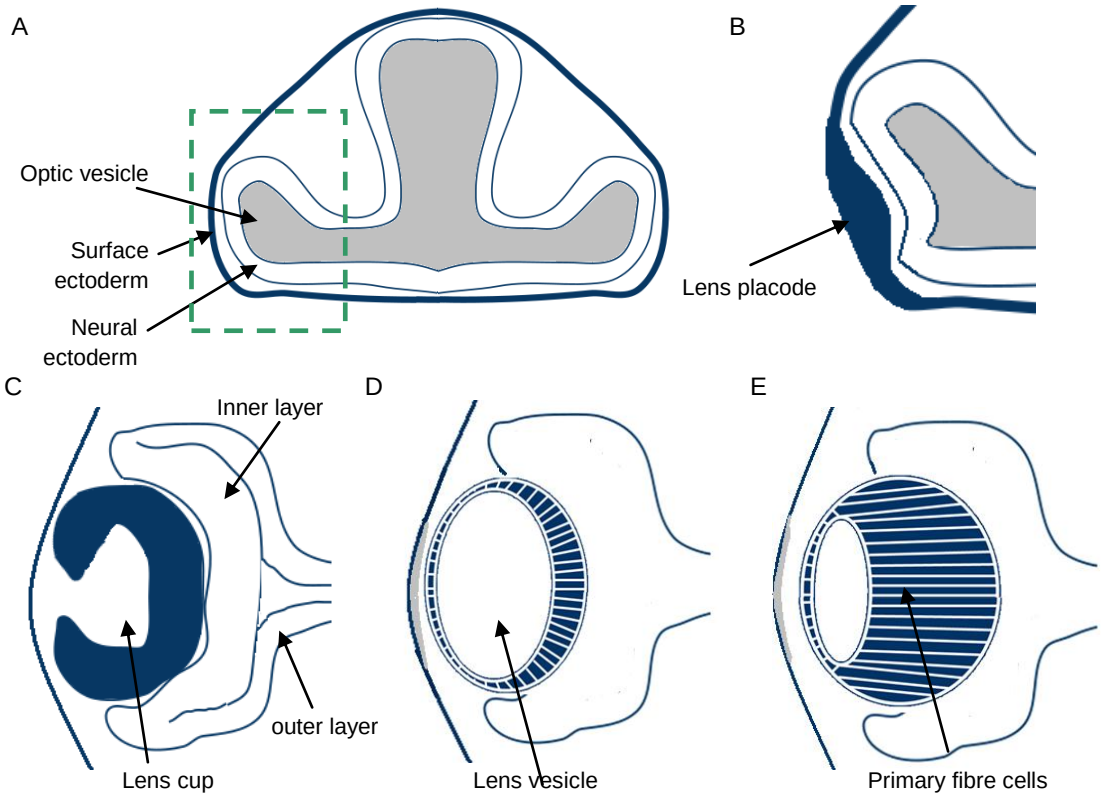


Figure 1.2. Development of the mouse lens. A) The lens placode is induced by the interaction of the surface ectoderm and the neural ectoderm of the optic vesicle (E9). B) - E) enlargement of area contained in green box in A). B) The surface ectoderm thickens forming the lens placode. C) The lens placode invaginates forming the lens cup and the neuroectoderm surrounding the invaginating lens epithelium folds over forming two layers the inner nuclear layer and the outer nuclear layer. D) The lens cup closes to form the lens vesicle. The retina begins to differentiate and the future cornea (grey) begins to appear. E) The cavity of the vesicle is filled by the elongating posterior lens fibre cells.

The lumen of the lens vesicle is then filled by elongating primary lens fibre cells, which develop from the posterior wall of the vesicle as soon as it is formed (E12.5-14) (Figure 1.2D,E) (Zhou *et al.* 2002). However, the epithelial cells at the anterior wall remain as a cuboidal epithelial monolayer, establishing polarity in the lens. Once the vesicle is filled, mitotic cells from the central epithelial region divide and move to the equatorial region. Here at the lens bow, the cells elongate and differentiate into secondary fibres. The growing secondary cells surround the central primary cells, forming concentric layers (E15.5). The elongated mature fibre cells span the length of the lens, and sutures form where the apical ends of two cells are in contact in the anterior of the lens and where the basal ends of two cells are in contact in the posterior of the lens (Figure 1.1B).

Rapid growth of the lens occurs during embryogenesis and the first two years of life with continual addition of secondary fibre cells (~90 % of adult size) (Gordon and Donzis 1985). Throughout the rest of life, secondary fibres are continually added, albeit at a much slower

rate, resulting in an ever-growing lens (Gordon and Donzis 1985).

1.4 Secondary lens fibre differentiation

All lens fibre cells originate from lens epithelial cells which have undergone differentiation. This co-ordinated process involves proliferation of the epithelial cells, withdrawal from the cell cycle, elongation of the cells, expression of fibre-specific genes and degradation of organelles. Growth factors from the aqueous, the vitreous and the ciliary body all contribute to lens fibre differentiation.

1.4.1 The central lens epithelium remains undifferentiated

The cells in the central lens epithelium remain in an undifferentiated state. These cells are relatively quiescent with a small number of proliferating cells (Zhou *et al.* 2006). The central epithelium cells divide very infrequently leading to the hypothesis that these cells represent a stem cell population (Zhou *et al*. 2006). To inhibit differentiation, the central lens epithelial cells express a number of transcription factors. Pitx3, expressed in the central lens epithelium and at the lens equator, limits the expression of the cyclin dependent kinase (CDK) inhibitors $p27^{Kip1}$ and $p57^{Kip2}$ which are required for cell cycle withdrawal (Zhang *et al.* 1998; Ho *et al.* 2009). AP-2α regulates expression of the cell-adhesion molecule E-cadherin in the epithelium and inhibits expression of the fibre specific aquaporin0 (Ohtaka-Maruyama *et al.* 1998; West-Mays *et al.* 2002). Pax6, necessary for lens placode formation, inhibits expression of the lens fibre specific βB1-crystallin and the pro-fibre transcription factor c-Maf in the lens epithelium (Duncan *et al.* 2004). The transcription factor lens epithelium-derived growth factor (LEDGF) enhances growth and survival of the all lens epithelial cells (Singh *et al.* 2000; Kubo *et al.* 2003).

1.4.2 Cell proliferation

At the peripheral region of the lens epithelium, the germinative zone, the epithelial cells proliferate and move to the transition zone (Zhou *et al*. 2006). Low concentrations of FGF, found in the aqueous humor that baths the central epithelium, are associated with proliferation of lens epithelial cells (Chamberlain and McAvoy 1989). The lens expresses four tyrosine kinase receptors (FGFR1 to 4) (Robinson 2006) and FGFR-1 and -2 are highly expressed in the equatorial region (Garcia *et al.* 2005). Mice in which FGFR-1, -2 and -3 have been knocked out in combination have abnormal proliferation in prospective lens fibre cells as well as other anomalies (Zhao *et al.* 2008). *In vitro* FGF-1 and -2, insulin-like growth factor (IGF), epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and aqueous humour are all associated with lens cell proliferation (Iyengar *et al.* 2006). The Wnt/β-catenin pathway is associated with proliferation as loss of β-catenin results in decreased cell proliferation (Cain *et al.* 2008) and overactivating the Wnt pathway increases epithelial cell proliferation (Martinez *et al.* 2009).

At the molecular level, proliferation is promoted by the transcription factors Pitx3 and FoxE3

(Medina-Martinez *et al.* 2005; Ho *et al*. 2009). Loss of the gap junction protein connexin 50 (Cx50) is associated with reduced proliferation of the equatorial and central epithelium (White *et al.* 2007). Activity of Src family tyrosine kinases (SFKs) is associated with inhibiting assembly of N-cadherin junctions which is thought to be necessary to maintain cell proliferation (Griep 2006).

1.4.3 Withdrawal from the cell-cycle

In the transition zone cells undergo the first step in differentiation: withdrawal from the cell cycle. The cell cycle inhibitor $p57^{Kip2}$ is highly expressed in cells at the transition zone (Lovicu and McAvoy 1999) and combined loss of $p57^{Kip2}$ and $p27^{kip1}$ results in hyperproliferation and apoptosis of fibre cells (Zhang *et al*. 1998). Upregulation of $p57^{Kip2}$ and $p27^{kip1}$ is dependent on expression of the homeobox transcription factor Prox1 (Wigle *et al.* 1999), the presence of FGF and the inhibition of SFK activity (Walker *et al.* 2002).

As the cells move more posteriorly during differentiation they are exposed to higher concentrations of FGF from the vitreous humour (Lovicu and McAvoy 2005). This high concentration of FGF increases Erk1/2 activation leading to cell cycle withdrawal and differentiation (Lovicu and McAvoy 2001; Iyengar *et al*. 2006; Iyengar *et al.* 2007; Zhao *et al*. 2008). Of the four FGF receptors expressed in the lens FGFR-2 has been shown to be directly involved in cell cycle withdrawal (Garcia *et al*. 2005).

The retinoblastoma protein (pRb) is reported to be abundant at the lens equator (Griep 2006) and epithelial cells lacking pRb expression fail to withdraw from the cell cycle (Fromm *et al.* 1994; Morgenbesser *et al.* 1994). Active pRb binds to and represses E2F transcription factors, leading to down-regulation of the E2F targets involved in cell cycle progression (Dyson 1998). The transcription factor Pax6 also controls cell cycle withdrawal, as Pax6-deficient cells at the lens equator fail to exit the cell cycle (Shaham *et al.* 2009).

1.4.4 Cell migration and elongation

Past the transition zone, the apical ends on the differentiating cells turn through 90° and point towards the anterior (Figure 1.3) (Grove *et al.* 2004). The apical ends of the migrating differentiating cells contact the apical membranes of the lens epithelium at the epithelial cell-fibre interface (Grove *et al*. 2004). The basal membrane remains in contact with the lens capsule. As the cell migrates to the interior of the lens it elongates several hundred fold as the depth of the lens increases (Coulombre and Coulombre 1963). This movement is dependent on the reorganisation of cell-cell contacts between the apical membrane of the epithelium and the anterior end of the fibre cells, and between the posterior end of fibre cells and the posterior lens capsule. The cells finally detach from the lens epithelium and lens capsule and interlock with the opposing cell forming a suture, with posterior sutures forming before anterior (Bassnett and Winzenburger 2003; Kuszak *et al.* 2004). Although sutures are present in vertebrate lenses, the arrangement of sutures is not conserved, with four distinct types (Kuszak *et al*. 2004).

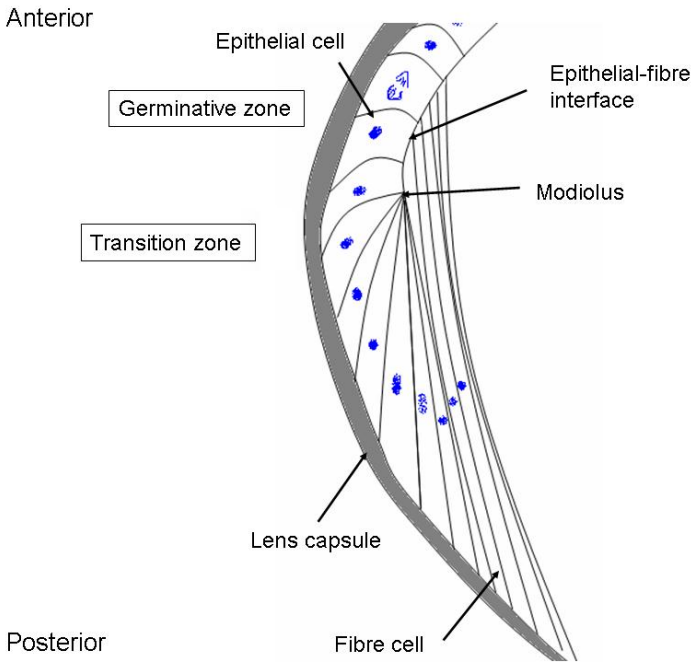


Figure 1.3. Cell-cell contacts at the lens equator. The epithelial cells in the germinative zone divide and at the transition zone the cells elongate. Past the transition zone, the differentiating cells turn towards the anterior. The apical ends of the elongating cells form a suture known as the modiolus or lens fulcrum. The basal membranes of the newly differentiated fibre cells remain in contact with the lens capsule and the apical membranes of the cells are in contact with the apical membranes of the lens epithelial cells at the epithelial-fibre interface.

Signalling factors from the ciliary body and vitreous humour promote migration of the differentiating cells (Zelenka 2004). As cells migrate and elongate the adhesive and junctional properties of the cells change. The transcription factor AP2-α is expressed in lens epithelia and expression is lost at the lens equator (West-Mays *et al.* 1999). AP2-α down-regulates expression of the adherens junction protein E-cadherin, which, in many tissues, is associated with the onset of cell migration (West-Mays *et al*. 2002). The adaptor protein Abi-2 is expressed at the ends of migrating fibre cells in the transition zone and is present at the epithelial-fibre interface. Loss of Abi-2 expression results in defects in orientation, migration and elongation of differentiating cells (Grove *et al*. 2004). Phosphorylated regulatory myosin light chain-2 (phospho-MLC), which colocalises with actin filament in elongating and differentiating primary and secondary lens fibres, is also thought to have a role (Maddala *et al.* 2007).

When lens differentiation is initiated, actin, present in the form of stress fibres in the epithelium, dissociates and is reorganised into membrane-attached cortical bundles (Lo *et al.* 1997; Weber and Menko 2006). This is in part dependent on actin-binding proteins (Zelenka 2004) and on the formation of N-cadherin cell-cell junctions (Ferreira-Cornwell *et al.* 2000; Weber and Menko 2006). Thus, a switch from E-cadherin junctions in the epithelia to N-cadherin junctions in the differentiating fibres occurs (Pontoriero *et al.* 2009).

In both lens epithelial and fibre cells, N-cadherin forms a complex with β-catenin (Leonard *et al.* 2008). Localisation of N-cadherin in the lens fibre cell membrane, where it is mainly in the short sides of the hexagons and colocalises with β-catenin (Straub *et al.* 2003) and F-actin (Leonard *et al*. 2008), is dependent upon the activation of ephrin-A2 (Cooper *et al.* 2008). Activation of ephrin-A2 is dependent on expression of the axon guidance molecule ephrin-A5. In ephrin-A5 knockout lenses, which develop cataract, N-cadherin fails to bind β-catenin and remains in the cytoplasm (Cooper *et al*. 2008). In the differentiating and mature fibres there is an increase in N-cadherin binding to γ-catenin and they colocalise at the cell vertices (Leonard *et al*. 2008).

Movement along the posterior capsule is mediated by the basal membrane of the fibre cell, which is rich in β1 and α6 integrins and associated cytoskeletal elements, interacting with the extracellular matrix (Bassnett *et al.* 1999; Walker and Menko 1999). The serine/threonine kinase Cdk5 localises to the tips of elongating fibre cells, where it is present in high concentrations (Gao *et al.* 1997; Negash *et al.* 2002). Substrates of Cdk5 have roles in cell migration, and *in vitro* overexpression of Cdk5 enhances adhesion of lens epithelial cells to extracellular matrix (Negash *et al*. 2002).

1.4.5 Expression of fibre-specific proteins

Differentiation is associated with large changes in gene expression, with downregulation of epithelium-specific genes and up regulation of a number of fibre-specific genes including β-crystallin, aquaporin0, MP20, Cx46 and beaded filament proteins. The effects of FGF on expression of lens fibre specific proteins are dependent on the activity of members of the bone morphogenetic protein (BMP) family of proteins (Boswell *et al.* 2008). *In vitro*, EGF and PDGF-A, present in the vitreous as well as the more potent IGF-1 upregulate the expression of fibre-specific genes such as β-crystallin (Wang *et al*. 2010a). Post cell-cycle withdrawal FGF is also involved in the accumulation of the transcription factor c-Maf, induction of crystallin expression and gap junction formation (Le and Musil 2001a; Lovicu and McAvoy 2001; Zhao *et al*. 2008).

1.4.5.1 Expression of β- and γ-crystallins

High concentrations of the structural crystallin proteins exist in the lens contributing to refractive power. Expression of β-crystallins serve as a marker of lens fibre cells, and expression coincides with the loss of Pitx3 (Ho *et al*. 2009). c-Maf is upregulated during fibre cell differentiation (Duncan *et al*. 2004) and enhances expression of multiple crystallins, particularly the *γ*-crystallins (Kim *et al.* 1999). Similarly, the transcription factors Sox1 and Prox1 up-regulate expression of multiple *γ*-crystallins (Nishiguchi *et al.* 1998; Wigle *et al*. 1999). Loss of the epithelial transcription factor Pitx3 also associates with increased *γ*-crystallin expression (Ho *et al*. 2009).

1.4.5.2 Change in aquaporin expression

Aquaporins are integral membrane proteins that facilitate the passive movement of water between cells (Agre *et al.* 2002). Aquaporin1 (AQP1), expressed in the lens epithelium is 40

times more efficient at facilitating water movement than AQP0, which is expressed only in terminally differentiated fibre cells (Varadaraj *et al.* 2007). AQP0 is the most abundant membrane protein in lens fibre cells (Bassnett *et al.* 2009) and, in addition to its water function, it is postulated to have a role in cell-cell adhesion (Kumari and Varadaraj 2009).

1.4.5.3 Change in connexin expression

The lens expresses three connexins (Cx), the subunits of gap junctions. Gap junctions between cells allow the passage of small molecules and ions, and are important in maintaining physiology in lens fibre cells (Patterson and Delamere 1992). Cx43 is restricted to epithelial cells where it is expressed at low levels and is required for maintaining lens transparency (Beyer *et al.* 1989; Gong *et al.* 1997). Upon differentiation Cx46 expression is initiated and high levels are found in lens fibres (Paul *et al.* 1991). In the differentiating fibres Cx46 forms gap junctions with Cx50 (Konig and Zampighi 1995), a connexin that is expressed in both epithelial and fibre cells (White *et al.* 1992; White *et al*. 2007). The gap junction coupling conductance in *in vitro* differentiating fibres is increased by FGF at concentrations expected in the lens equator (Le and Musil 2001b). Mice lacking Cx46 display normal lens growth and fibre differentiation and develop cataract postnatally (Gong *et al*. 1997). Whilst Cx50 is necessary for lens growth and development, Cx46 is necessary for the maintenance of lens transparency (Gong *et al*. 1997; Rong *et al.* 2002).

1.4.5.4 Expression of beaded filament proteins

The expression of beaded filament proteins occurs after elongation of fibre cells, and have a role in stabilisation after differentiation (Blankenship *et al.* 2001). These lens specific intermediate filament proteins, filensin and phakinin/CP49 (BFSP1 and BFSP2, respectively), are highly expressed and localise at the cell membrane of peripheral cortical fibres and the cytoplasm in deeper cortical fibres (Oka *et al.* 2008). The beaded filament proteins act to minimise light scatter by their involvement in ensuring that cortical fibre cells adopt the correct spatial organisation (Sandilands *et al.* 2003; Song *et al.* 2009).

1.4.6 Degradation of organelles

To maintain transparency, mature lens fibres lack organelles (Bassnett and Beebe 1992). All cytoplasmic organelles are degraded in a coordinated manner in primary lens fibre cells and in the final stage of secondary lens fibre differentiation (Bassnett 2002). In mice the loss of organelles begins at E17/18 and is complete when the eyelids are first opened, 14 days after birth (Vrensen *et al.* 1991). In the chicken lens, organelles in the lens core are degraded at E12 forming an organelle free zone (OFZ) (Bassnett and Beebe 1992). At first the OFZ only includes the primary cells but this region expands at a rate of 20 cell widths per day to include the secondary fibres. The mitochondria in cells bordering the OFZ swell and suddenly fragment. This process is rapid with the loss of numerous mitochondrial markers within 2-4 hours (Bassnett and Beebe 1992; Dahm *et al.* 1998; Zandy and Bassnett 2007). Degradation of the endoplasmic reticulum also occurs within a short time period (Bassnett 1995).

The Golgi apparatus fragment into small structures apical to the cell nuclei which later disintegrate (Bassnett 1995). Degradation of the nucleus is less rapid and the process varies among species. In primates, the nucleus disintegrates and residual nuclear particles persist indefinitely in the cytoplasm (Bassnett 1997). The nucleus initially becomes more spherical and the chromatin condenses and fragments. DNase II-like acid DNase (DLAD) is expressed only in the lens and is believed to digest the chromosomal DNA (Nakahara *et al.* 2007). This degradation of nuclear material results in mature fibre cells losing the capacity for *de novo* protein synthesis. Thus, proteins in the core of the lens nucleus have been present since pre-natal and early post-natal development.

The mechanism behind organelle degradation remains to be resolved. Mice lacking the effector caspases known to be expressed in the lens undergo normal organelle breakdown, suggesting that apoptosis is not the mechanism (Zandy *et al.* 2005). Mice lacking the autophagy gene *Atg5* are also unaffected (Matsui *et al.* 2006). Recent evidence suggests that organelle degradation may be dependent upon the ubiquitin proteasome pathway (UPP) (Zandy and Bassnett 2007). Upon differentiation, members of the UPP degradation system redistribute from the cytoplasm to the nucleus (Girao *et al.* 2005). Inhibition of the UPP results in the elimination of the mitochondrial marker succinate-ubiquinone oxidoreductase from the differentiating central fibre cells (Zandy and Bassnett 2007).

In summary, the adult lens is continually growing, forming from the proliferation and differentiation of epithelial cells at the lens periphery. Differentiating cells increase in length several hundred fold as they migrate, and lose the capacity for gene expression as the nucleus and other cytoplasmic organelles are degraded. These changes and the expression of crystallins, AQP0, connexins and beaded filaments all contribute to the transparency of the lens.

1.5 Diagnosis, classification, and treatment of congenital cataract

1.5.1 Diagnosis and classifications of congenital cataract

Cataract is readily identifiable as light scattering regions by slit lamp examination (oblique illumination) or retroillumination. Congenital cataract is genetically and phenotypically heterogeneous and types are classified according to position and type of opacity (Table 1.1) (Reddy *et al.* 2004). Lens opacities are commonly found in the nuclear and cortical fibres and may also be found in both (“total cataract”) (Amaya *et al.* 2003). Nuclear cataracts are restricted to the embryonic nuclear fibres (Figure 1.4A). Opacities in the central nuclear fibres may be solid or dust-like (pulverulent) (Reddy *et al*. 2004). Visual acuity is greatly affected by the presence of solid and dust like opacities, but cerulean cataracts have little effect on visual acuity. Isolated cortical opacities are rare and as they typically lie outside of the visual axis they therefore have little effect on visual acuity (Ionides *et al.* 1999). Cerulean cataracts,

which typically have a juvenile rather congenital onset, appear as blue and white dots throughout the lens and don't require surgery until adulthood (Bodker *et al.* 1990).

Sutural cataracts uniquely display a Y-shaped opacity and can be present at the anterior or posterior (Figure 1.4B) (Zhang *et al.* 2006). Sutural cataracts are distinct form anterior polar and posterior polar cataracts which are more common (Ionides *et al*. 1999). In lamellar cataracts, the opacities are restricted to a layer, or layers, of cells which represent a specific generation of secondary lens fibres. A defect in the development of this specific generation of cells when they were at their most metabolically active results in restricted opacification (Brown and Bron 1996). Although a clear separation from the cortex is usually seen, the nuclear fibre may also show some opacification (Figure 1.4C) (Amaya *et al*. 2003). Acueliform cataracts are rare and present as unique fibreglass–like projections (Figure 1.4D) (Zenteno *et al.* 2005).

1.5.2 Why study a treatable disease?

Cataract is a treatable disease but the only effective medical treatment is surgery to remove the cataractous lens. Treatment of congenital or juvenile cataract is expensive with surgery and post-operative care recently estimated to cost an average of almost $20,000 per patient (Stager *et al.* 2009). The loss of the lens may be corrected by the insertion of an artificial intraocular lens (IOL) that is positioned in the place of the original lens, although this may be preceded by a period when the refractive error caused by loss of the lens is corrected using contact lenses. There are a number of risks associated with cataract surgery in the young. Aphakic glaucoma, strabismus and amblyopia are common risks of lens removal procedures, with aphakic glaucoma particularly prevalent in the very young (Allen *et al.* 2009; Khan and Al-Dahmesh 2009; Yu and Dahan 2009). Other rare complications include retinal detachment and endophthalmitis (inflammation of the intraocular cavities) (Stager *et al*. 2009).

Early vision loss due to congenital cataracts affects visual development at the time of vision loss and has also been shown to affect aspects of vision that are acquired later on in visual development (Lewis and Maurer 2009). Therefore, early diagnosis and treatment of congenital and early onset forms of cataract is necessary to maximise the visual capabilities of patients later on in life. Understanding the genetic causes of cataract may benefit future treatment by identifying the molecular mechanisms behind cataract formation. Cataract is the most common cause of blindness in adults worldwide (Pascolini *et al*. 2004), and identifying genes involved in juvenile onset cataract will aid understanding of the causes of age-related cataract. Identifying mutations which cause congenital cataract can also help in genetic counselling. Parents at risk of having a child with congenital cataract can be counselled and made aware of the potential for cataract formation in their future offspring. This can help with early diagnosis and treatment which may aid their prospects of achieving maximum visual acuity.

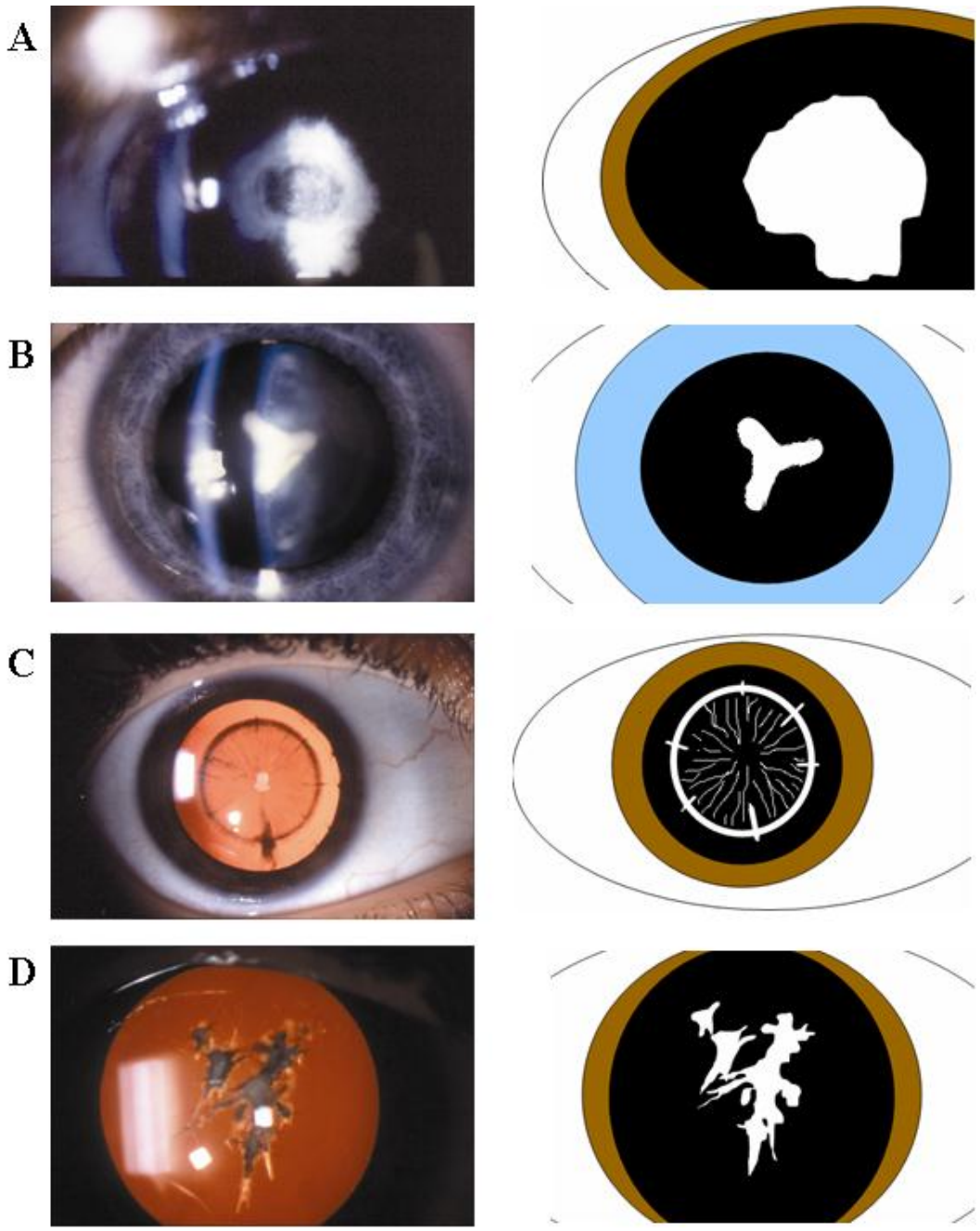


Figure 1.4. Examples of congenital cataract. Lenses of cataract patients visualised using direct photogrpahy (A,B) or retroillumination (C,D). Cataracts are visualised as silhouettes when light is reflected from the retina using retroillimuniation. To the right hand side are schematics representing the images on the left. (A) Nuclear cataract. B) Sutural cataract has a typical Y shaped appearance. C) Lamellar cataract visualised by retroillumination. D) Acueliform cataract visualised by retroillumination. Schematic legend: White shape on black background = cataract. Brown/blue = iris. White = sclera. Cataract images were supplied by Professor Graeme Black DPhil FRCOphth, Central Manchester University Hospitals NHS Foundation Trust.

1.6 Mutations associated with congenital cataract

Mutations in more than 25 genes have been associated with non-syndromic congenital or early-onset cataract (Table 1.1). Approximately half of cases which present have mutations

in genes encoding the crystallins, and about a quarter are associated with connexin mutations (Hejtmancik 2008). Mutations in numerous transcription factors result in congenital cataract, although these are usually seen with other ocular abnormalities.

Table 1.1. Human genes with mutations associated with non-syndromic congenital or juvenile cataract. Accession numbers are from the GenBank database. Chrom: chromosomal location. Inh: inheritance pattern. AD: autosomal dominant. AR: autosomal recessive. XL: X-linked. OMIM: Online Mendelian Inheritance in Man reference numbers. Where OMIM accessions were unavailable, the original paper reporting the mutation is referenced.

Gene	Accession	Chrom	Inh	Morphology	OMIM
Transcription Factors					
FOXE3	NM_012186	1p32	AD	No lens pathology details	107250, 601094
HSF4	NM_001538	16q21-q22.1	AD	Lamellar, nuclear, total	602438
			AR	Total	(Smaoui *et al.* 2004)
MAF	NM_001031804	16q23	AD	Lamellar pulverulent, cerulean	177074
PITX3	NM_005029	10q25	AD	Posterior polar	602669
VSX2/CHX10	NM_182894	14q24.3	**AR**	No lens pathology details	142993
PAX6	NM_000280,	11p13	AD	No lens pathology details	607108
Structural Proteins					
CRYAA	NM_000394	21q22.3	AD	Nuclear, lamellar, total	123580
			AR	No lens pathology details	
CRYAB	NM_001885	11q23.3−24.2	AD	Posterior polar, lamellar	123590
CRYBA3/BA1	NM_005208	17q11-q12	AD	Nuclear, sutural	600881
CRYBA4	NM_001886	22q11.2	AD	Lamellar	123631
CRYBB1	NM_001887	22q11.2	AD	Pulverulent	600929
			AR	Nuclear	
CRYBB2	NM_00496	22q11.2	AD	Variable	123620
CRYBB3	NM_004076	22q11.2	**AR**	Nuclear	123630
CRYGC	NM_020989	2q33−q35	AD	Nuclear lamellar, lamellar, zonular perverulent	123660, 123680, 601286
CRYGD	NM_006891	2q33−q35	AD	Variable	115700, 123690
CRYGS	NM_017541	3q26.3-qter	AD	Polymorphic cortical	123730
BFSP1	NM_001195	20p11.23-p12.1	**AR**	Zonular nuclear, pulverulent	603307
BFSP2	NM_003571	3q21-q25	AD	Nuclear, sutural, lamellar	603212
Vimentin	NM_003380	10p13	AD	Pulverulant	193060
Membrane Associated Proteins					
GJA3/CX46	NM_021954	13q11−13	AD	Variable	601885
GJA8/CX50	NM_005267	1q21-q25	AD	Nuclear, zonular pulverulent, total, posterior subcapsular	116200
			AR	No lens pathology details	(Ponnam *et al.* 2007)
AQP0	NM_012064	12q12−14.1	AD	Variable	154050
LIM2 (MP20)	NM_002316	19q	**AR**	Total	154045
NHS	NM_198270	Xp22.13	XL	Total	300457
Other					
GCNT2	NM_001491	6p24-p23	**AR**	No lens pathology details	110800
EYA1	NM_172060	8q13.3	AD	Nuclear	601653
CHMP4B	NM_176812	20q11.22	AD	Posterior	610897
EPHA2	NM_004431.2	1p36	AD	posterior polar	176946
			AR	Nuclear	(Kaul *et al.* 2010)

1.6.1 Transcription factors

The transcription factor PAX6 is one of the key proteins in the induction of eye formation. Pax6 is expressed in the presumptive lens ectoderm, and expression is required for formation of the lens placode and subsequently the lens. Mutations in murine *Pax6* result in the failure of the optic vesicle and overlying ectoderm to interact, compromising one of the key inductive events in eye development (Hogan *et al.* 1986). When Pax6-deficient surface ectoderm is transplanted onto a wildtype optic vesicle, lens formation fails to occur (Fujiwara *et al.* 1994). The inductive properties of Pax6 are conserved in vertebrates and non-vertebrates as ectopic expression of murine Pax6 induces functional ommatidal eyes in the legs or antennae of Drosophila (Halder *et al.* 1995). Pax6 is a master regulator of other transcription factors important for eye development such as Foxe3, Pitx3, Maf, Mitf and Vsx2 (Cvekl *et al.* 2004) and mutations in these genes are associated with congenital cataract (Table 1). Mutations in *PAX6 are* commonly associated with aniridia (absence of the iris) (OMIM 106210) (Tzoulaki *et al.* 2005) but a number of mutations in *PAX6* have been associated with congenital cataract (Glaser *et al.* 1994; Azuma *et al.* 1999; Hanson *et al.* 1999).

1.6.2 Structural proteins

Mutations in lens structural proteins are most commonly associated with congenital cataract (Hejtmancik 2008). The α-crystallins, large complexes of two subunits, αA- and αB- crystallins, represent ~30 % of the soluble protein in the lens (Bloemendal and Groenewoud 1981). These proteins have a number of roles in the lens other than their contribution to refractive index. Due to the lack of organelles, there is no protein turnover in the mature lens fibres. The α-crystallins act as chaperones preventing the aggregation and degradation of the life-long expressed proteins in the lens, including other crystallins (Horwitz 1992; Acosta-Sampson and King 2010). Both α-crystallins are necessary for microtubule assembly in lens epithelia (Xi *et al.* 2006). Other roles specific to αA-crystallin in the lens include protection against apoptosis (Andley *et al.* 2002; Morozov and Wawrousek 2006; Xi *et al.* 2008) and assembly of beaded filaments (Carter *et al.* 1995). αB-crystallin is predominantly expressed in the epithelium and may have a role in genomic stability (Andley *et al.* 2001).

Mutations in αA-crystallin have been associated with autosomal dominant and recessive forms of congenital cataract. Heterozygous mutations in the conserved core α-crystallin domain cause dominant cataract by altering protein interactions (Fu and Liang 2003). The dominat cataract associated heterozygous mutation R49C, which lies outside the core α-crystallin domain, loses the ability to protect against apoptosis (Mackay *et al.* 2003), and activates the unfolded protein response (Watson and Andley 2010). Cell expressing the R12C, present in the heterozygous state in cases of autosomal dominant cataract, exhibit an altered heat shock response (Zhang *et al.* 2009). A homozygous nonsense mutation in the N-terminus of CRYAA results in recessive congenital cataract (Pras *et al.* 2000), suggesting that lack of CRYAA function is associated with recessive cataract, but gain of function mutations are associated with dominant cases.

The β- and γ-crystallins are genetically and functionally distinct from the α-crystallins (Graw 2009). These crystallins are expressed in the lens fibres and contribute to the refractive power of the lens. Mutations in many of these genes, as well as genes that regulate β- and γ-crystallin expression (e.g. MAF and HSF4) are associated with congenital cataract. The lens fibres express two lens-specific intermediate filament proteins filensin and phakinin/CP49 (BFSP1 and BFSP2, respectively) which form the beaded filament in differentiated lens fibre cells (Maisel and Perry 1972). Mutations in both proteins are associated with juvenile-onset cataract (Conley *et al.* 2000; Ramachandran *et al.* 2007) and a role in myopia has also been hypothesised for BSFP2 (Zhang *et al.* 2004). Recently, a dominant missense mutation in the widely expressed intermediate filament protein vimentin was associated with pulverant congenital cataract (Sandilands *et al.* 1995; Muller *et al.* 2009). Although vimentin knockout mice do not display a lens phenotype (Colucci-Guyon *et al.* 1994), missense mutations in (Bornheim *et al.* 2008), and overexpression of (Capetanaki *et al.* 1989) vimentin in mice results in cataract.

1.6.3 Membrane associated proteins

The lens expresses a number of unique membrane proteins such as the gap junction proteins CX46 and CX50, and the highly prevalent AQP0 and MP20. Mutations in each of these proteins have been associated with congenital cataract, underlying their importance in lens homeostasis. As the lens is avascular, gap junctions and water channels are necessary for the delivery of nutrients to and the removal of waste from the central fibre cells (Gong *et al.* 2007). As stated previously in section 1.4.5.3, connexins play an important role in fibre cell physiology and mutations in both of the fibre cell-expressed connexin genes (CX46 and CX50) result in congenital cataract (Addison *et al.* 2006; Vanita *et al.* 2006; Gao *et al.* 2010).

AQP0 is the most abundant membrane protein in lens fibre cells (Bassnett *et al*. 2009) and *Aqp0* knockout mice display congenital cataract (Shiels *et al.* 2001; Al-Ghoul *et al.* 2003). AQP0 is a member of the aquaporin family of proteins which facilitate the passive movement of water between cells (Agre *et al*. 2002). However, AQP0 is much less efficient at facilitating water movement across the membrane than AQP1 (Chandy *et al.* 1997), which is expressed in the lens epithelium (Varadaraj *et al*. 2007). An *in vitro* role in cell-cell adhesion has been demonstrated for AQP0, but not AQP1, suggesting that AQP0 has functional roles beyond water transport (Kumari and Varadaraj 2009; Varadaraj *et al.* 2010). Deletions and missense mutations affecting both the transmembrane and extracellular loops have been associated with congenital lamellar cataract (Francis *et al*. 2000; Geyer *et al*. 2006; Gu *et al*. 2007; Wang *et al*. 2010c).

A single missense mutation in the fourth transmembrane domain of lens intrinsic membrane protein 2 (LIM2) has been associated with recessive congenital cataract (Ponnam *et al.* 2008). A phenotypically milder form of recessive cataract with a later, but pre-senile, onset is associated with a missense mutation in the third transmembrane domain (Pras *et al.* 2002).

This protein, the second most abundant membrane protein in lens fibres, will be discussed in more detail below (Section 1.7.4).

Null mutations in the Nance-Horan Syndrome protein (NHS) cause Nance-Horan Syndrome, an X-linked disorder typified by anterior ocular anomalies, dental anomalies, dysmorphic facial features and mental retardation (Ding *et al.* 2009). Copy number variants of the NHS gene are associated with congenital cataract with a possible link with heart defects (Coccia *et al.* 2009). NHS is expressed in at least 5 isoforms (Brooks *et al.* 2010). According to the nomenclature used by Brooks *et al*., isoforms NHS-A and NHS-1A contain exon 1 and are expressed in epithelia (Sharma *et al.* 2006; Sharma *et al.* 2009). The *Xcat* mouse which displays bilateral total cataract as well as other NHS phenotypes has reduced expression of exon 1 isoforms (Huang 2006). NHS-1A interacts with the tight junction protein ZO-1 and colocalises with ZO-1 in developing epithelial tissues (Sharma *et al*. 2009) and *in vitro* (Sharma *et al*. 2006).

Exon 1 of *NHS* encodes a WAVE homology domain (WHD) and NHS isoforms with this domain have roles in maintenance of cell morphology and size (Brooks *et al*. 2010). These WHD containing isoforms interact with the Abl-interactor (Abi) adaptor proteins associated with actin polymerisation (Stradal *et al.* 2001) and the NHS isoforms maintain the contractile actin ring at the cell circumference (Brooks *et al*. 2010). This actin ring acts to provide cell shape and strength for cell adhesion. Abi2 is expressed at the tips of migrating secondary fibre cells at the epithelial fibre interface and knockout mice display defects in migration and orientation of secondary fibres and fail to form sutures (Grove *et al*. 2004).

1.7 TMEM114

1.7.1 Discovery of *TMEM114*

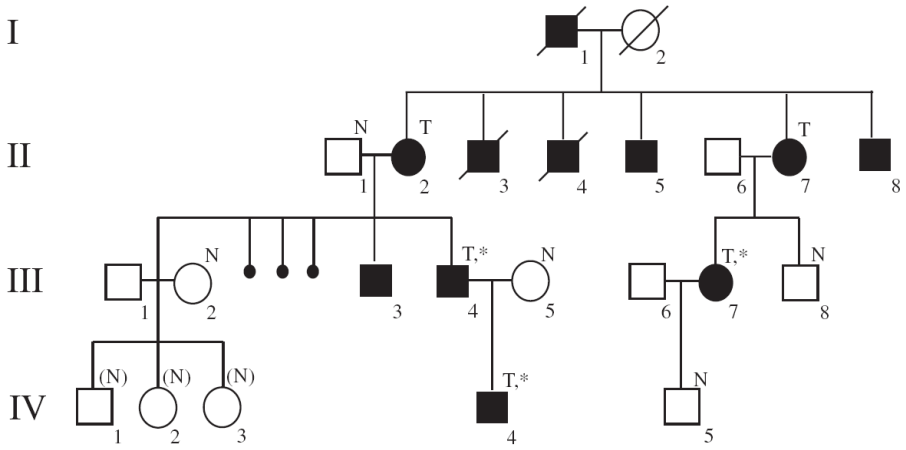
Transmembrane protein 114 (*TMEM114*) is associated with dominant bilateral congenital or juvenile cataract through its proximity to the breakpoint in a balanced translocation between 16p13.3 and 22q11.2 (Jamieson *et al*. 2007). A four generation family presented with congenital or juvenile cataract inherited in a dominant pattern (Figure 1.5A). Some of the affected also presented with developmental kidney abnormalities (Figure 1.5B). A balanced translocation t(16;22)(p13.3;q11.2) segregated with those affected. Limited clinical details and no photographs were available for the affected members of the family.

The chromosomal region 16p.13.3 had previously been identified as a locus associated with congenital cataract and microphthalmia (OMIM 156850) (Yokoyama *et al.* 1992). Analysis of the breakpoint regions using somatic cell hybrids revealed the translocation transected a region of 16p13.3 not sequenced by the human genome project. The BAC AC074052, located telomeric to the breakpoint, contained an EST (BF727259) from a human lens library. A homologous sequence encoded by four exons was identified in mouse, FANTOM clone

4930511J11 (Kawai *et al.* 2001). The predicted third and fourth exons of the human EST (BF727259) are present in BAC AC074052 with the predicted fourth exon furthest away from the breakpoint. Blast analysis against the human genome with the predicted human and murine exons 1 and 2 failed to identify any sequence similarity. This suggested that exons 1 and 2 were located in a gap in the sequence of the genome. Exons 1 and 2 were later identified in a short fragment of sequence by Celera.

The breakpoint lay in the putative promoter region of *TMEM114*, 673 nucleotides upstream of the predicted start codon and 151 nucleotides from the predicted transcription start site. The translocation was predicted to result in the separation of the coding sequence from the upstream regulatory regions of *TMEM114*, suggesting dysregulation as the likely disease mechanism (Jamieson *et al*. 2007).

A



B

Patient	Lens opacities	Age at operation	Other phenotypes
I.1	Juvenile onset	Childhood to early adulthood	Not known
II.2-5, II.8	Juvenile onset	Childhood to early adulthood	Not known
III.7	Juvenile onset	Childhood to early adulthood	Partial right duplex collecting system
III.4	Juvenile onset	Right eye: 7 years. Left eye: 23 years	Partial right duplex collecting system
III.3	Juvenile onset	Childhood to early adulthood	Not known
IV.4	Congenital: lamellar, central core, posterior sutural Later: anterior polar	Left eye: 6 years Right eye: 7 years	Absent Left kidney

Figure 1.5. Pedigree and phenotypes of the family with the balanced translocation t(16;22)(p13.3;q11.2). A). Pedigree of the family displaying juvenile onset cataract. Affected individuals are highlighted in black, unaffected in white. Individuals tested and found to have the balanced translocation 46,XX,t(16;22)(p13.3;q11.2) or 46,XY,t(16;22)(p13.3;q11.2) are indicated by a T. Individuals tested and found to have a normal karyotype are indicated by a N. Individuals not tested but inferred to have a normal karyotype are indicated by a (N). An asterisk denotes those tested and found to have a renal abnormality. B) Clinical details of the available patients.

1.7.2 TMEM114 bioinformatics

The predicted *TMEM114* gene encodes a 223 amino acid protein from four exons. Orthologues are found in mammals, chicken, frog and zebrafish (Jamieson *et al*. 2007). TMEM114 belongs to the PMP-22/EMP/MP20/Claudin family of transmembrane proteins (Pfam PF00822) (Jamieson *et al*. 2007). Proteins in this family are small (~18-25 kDa) and have four transmembrane domains with two extracellular loops. Members of this family include the epithelial tight junction proteins claudins and the more distantly related voltage dependent calcium channel gamma (γ) subunits.

Murine Tmem114 was first identified as a possible voltage dependent calcium channel γ subunit in the RIKEN collection of brain cDNA clones (4930511J11) (Kawai *et al*. 2001; Gustincich *et al.* 2003). Cataract has been associated with increased intracellular Ca^{2+} levels in lens epithelial cells resulting from the loss of Ca^{2+} regulation (Gupta *et al*. 2004a). Reduction of intracellular calcium levels has been shown to inhibit proliferation of epithelial cells in a human lens epithelial cell line (Meissner and Noack 2008). To date, mutations in voltage dependent calcium channel γ subunits have not been associated with congenital cataract.

1.7.3 Claudins and tight junctions

Claudins are transmembrane proteins that form the major component of epithelial tight junctions (Simon *et al.* 1999; Van Itallie *et al.* 2001). Head-to-head binding of claudins at the apical end of the lateral border between adjacent epithelial cells forms close connections, producing a selective barrier to diffusion which regulates the flow of ions, solutes and water across epithelia (paracellular transport). This barrier is measured as transepithelial resistance (TER) and the properties of barriers are determined by the claudins present which can vary from cell to cell (Gow *et al.* 1999; Simon *et al*. 1999), or in a developmental stage-specific manner (Simon *et al*. 1999; Troy *et al.* 2007). There are 23 claudins in the human genome, but the properties and function of a number of claudins (13, 17, 20, 21-24) have yet to be elucidated. A number of claudins are expressed in the kidney, and claudin expression varies in distinct regions of the kidney (reviewed in (Angelow *et al.* 2008). Mutations in two claudins have been associated with human diseases, both genes associated with a renal tube disorder familial hypercalciuric hypomagnesemia with nephrocalcinosis (FHHNC) (Konrad *et al.* 2006; Hampson *et al.* 2008).

Tight junctions divide polarised epithelial cells into distinct apical and basolateral membrane domains, although they are not involved in sorting proteins to specific domains (Umeda *et al.* 2006). Localisation of claudins at the tight junction is dependent on the scaffold proteins zonula occludens-1 (ZO-1) and ZO-2 (Umeda *et al*. 2006). An isoform of the Nance-Horan syndrome protein, a protein associated with cataract and Nance-Horan syndrome, localises to the tight junction, although its role there remains to be elucidated (Sharma *et al*. 2009).

Tight junctions have been observed by electron microscopy in human (Lo and Harding 1983) and rat lenses (Zampighi *et al.* 2000), but not in the mouse lens (Goodenough *et al.* 1980). The presence of tight junctions in the lenses of frog (Lo and Harding 1983) and chicken (Goodenough *et al*. 1980), however, suggests that tight junctions are conserved and are likely to be present in mouse lens epithelia. Further indirect support for tight junctions in mouse lens epithelia is the presence of relatively high levels of ZO-1 compared to those found in the lens fibres (Nielsen *et al.* 2003). Tight junctions in bovine lens epithelia have low transepithelial resistance (TER) but are functional (Zhang and Jacob 1994). In bovine lens epithelia some colocalisation of the adherens junction protein E-cadherin and ZO-1 is observed (Sugiyama *et al.* 2008). In the rat lens, tight junctions, which have the appearance of having low TER, are present in the central epithelia where they colocalise with gap junctions (Zampighi *et al*. 2000). The epithelial cells close to, and at, the equator lack tight junctions, but gap junctions are increased and are found connecting the neighbouring fibre cells (Zampighi *et al*. 2000).

1.7.4 Members of the Pfam00822 family in the lens

In the lens claudin-1, a ubiquitously expressed claudin (Furuse *et al.* 2002), localises to the tight junctions of the epithelium (Sugiyama *et al*. 2008). The presence of other claudins in the lens has not yet been reported. Peripheral myelin protein 22 (Pmp22), highly expressed in Schwann cells, has recently been detected in lens fibre cells (Ivanov *et al.* 2005; Bassnett *et al*. 2009). Mutations and copy number variants of *PMP22* are associated with multiple neuropathies including Charcot-Marie-Tooth type 1A disease (Nelis *et al.* 1996; Jen *et al.* 2005).

Lens intrinsic membrane protein, LIM2 (MP20), is the second most prevalent membrane protein in lens fibre cells and is highly conserved in vertebrates (Rao *et al.* 1989; Steele *et al.* 2000). In peripheral fibre cells Mp20 is located in the cytoplasm but in deeper fibre cells Mp20 relocates to the plasma membrane (Tenbroek *et al.* 1992; Grey *et al.* 2003). This relocation coincides with the formation of a diffusion barrier between the fibre cells and suggests Mp20 has a role in cell adhesion (Grey *et al*. 2003). Further support for Mp20's role in adhesion lies with its apparent interaction with galectin-3, a galactosidase-binding lectin which modulates cell-to-cell adhesion (Gonen *et al.* 2001). Mp20 also has a role in the fusion of cells in the formation of the lens syncytium (Shi *et al*. 2009).

Homozygous missense mutations in human *LIM2* are associated with recessive congenital (Ponnam *et al*. 2008) and presenile (Pras *et al*. 2002) cataracts. The To3 (total opacity of lens #3) mouse displays autosomal semi-dominant congenital cataracts caused by a missense mutation (p.G15V) in *Lim2* (Kerscher *et al.* 1996; Steele *et al*. 2000). In homozygous mutants the primary fibres are disorganised, as are the secondary fibres, some of which still retain their nuclei (Steele *et al.* 1997). This results in abnormal refractive properties in the peripheral cortex (Shiels *et al.* 2007). The mutated Mp20 is restricted to a compartment of the Golgi apparatus and fails to insert in the plasma membrane (Zhou *et al*. 2002). Lim2 knockout mice

fail to develop syncytia and present with pulverulent nuclear opacities and have altered refractive properties in the peripheral fibres (Shiels *et al*. 2007; Shi *et al*. 2009).

Peptide sequences from two other distantly related members of the PF00822 family, TMEM47 and PERP were also identified in a mass spectrometry analysis of lens fibres (Bassnett *et al*. 2009). TMEM47, also known as Brain Cell Membrane Protein 1 (BCMP1) or transmembrane 4 superfamily member 10 (TM4SF10), is widely expressed, with high levels in the brain (Christophe-Hobertus *et al.* 2001). TMEM47 is the vertebrate homologue of the *C. elegans* adherens junction protein VAB-9 (Christophe-Hobertus *et al*. 2001; Simske *et al.* 2003). A study focusing on TMEM47 expression in the kidney identified that Tmem47 was temporarily expressed in the membrane of developing podocytes and renal tubules (in the basal part of the lateral membranes in renal tubules) (Bruggeman *et al.* 2007). PERP (p53 effector related to PMP22) is ubiquitously expressed, with highest expression in stratified epithelia where it localises to desmosomes (Hildebrandt *et al.* 2000; Ihrie *et al.* 2005). Downregulation of PERP is associated with numerous aspects of cancer such as cell cycle control (Attardi *et al.* 2000), loss of cell-cell adhesion (Ihrie *et al*. 2005), upregulation of inflammation-related genes (Beaudry *et al*. 2010b) and has been associated with squamous cell carcinoma in human skin (Beaudry *et al*. 2010b) and uveal melanoma (Davies *et al.* 2009). PERP is also involved in wound healing, promoting cell-adhesion in wound closure (Beaudry *et al*. 2010a). The function of either protein in the lens has not been identified, although PERP is upregulated in p53 knockout lenses (Wang *et al*. 2010d).

1.7.5 Expression of *TMEM114*

In the adult mouse, *Tmem114* is expressed in the eye, cerebellum and testis (Figure 1.6A) (Jamieson *et al*. 2007). In the developing eye *Tmem114* is expressed as early as E13.5 and is upregulated post-natally (Figure 1.6B). Expression is significantly higher in the epithelial cells than the fibre cells of the lens. *Tmem114* expression was detected in the central lens epithelium by *in situ* hybridisation (Figure 1.7A). The strongest signal was obtained in the cuboidal cells close to the periphery above the lens equator (Figure 1.7B). Expression was also detected in the transitional zone where epithelial cells differentiate into elongated fibre cells, but expression is down-regulated upon differentiation of lens epithelial cells into lens fibre cells (Figure 1.7B) (Jamieson *et al*. 2007).

1.7.6 Sequence variants in *TMEM114*

The breakpoint in the translocation that led to the identification of *TMEM114* lies in the predicted promoter region, 151 nucleotides upstream of the predicted transcription start site (Jamieson *et al*. 2007). Promoter sequence analysis suggests that it contains binding sites for differentiation- and lens-specific transcription factors including Pax6 and AP2. This suggests that the translocation results in dysregulation of expression of the gene (Jamieson *et al*. 2007).

After identifying *TMEM114*'s association with autosomal dominant juvenile cataracts, a panel of 130 autosomal dominant congenital cataract (ADCC) patients was screened for mutations in *TMEM114*. Three heterozygous sequence variants which result in amino acid substitutions were identified (Jamieson *et al*. 2007). The sequence variants c.104T>C (p.I35T) and c.318T>G (p.F106L) were each found in one patient, and were not detected in 200 normal control chromosomes. Apparently unaffected family members of these patients had the same heterozygous changes. The missense mutation c.440C>T (p.A147V) was identified three times in the cohort, but when family members of one of the affected were screened, it was found that one affected family member did not harbour this sequence variant. Therefore, this variant was deemed a polymorphism (Jamieson *et al*. 2007).

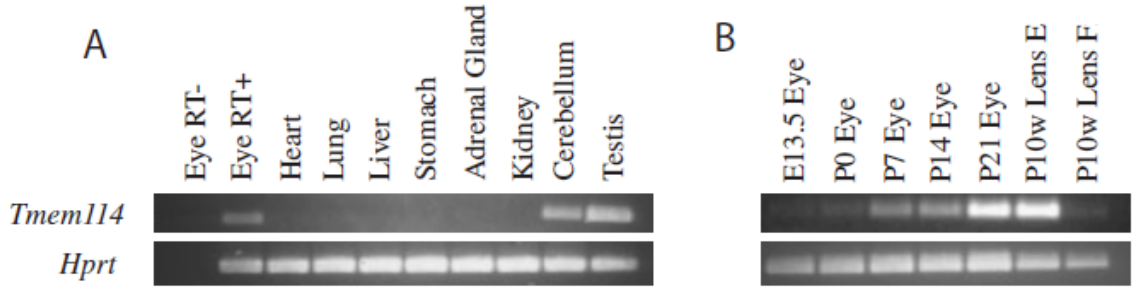


Figure 1.6. Expression of *Tmem114* in embryonic and adult mouse. A) RT-PCR from a panel of adult mouse tissues reveals expression of Tmem114 in the eye, cerebellum and testis. B) RT-PCR in the developing eye reveals weak expression from E13.5. Expression is upregulated postnatally. In the lens, expression is mainly restricted to the lens epithelium (E), with little expression in the lens fibres (F). Figure adapted from (Jamieson *et al*. 2007).

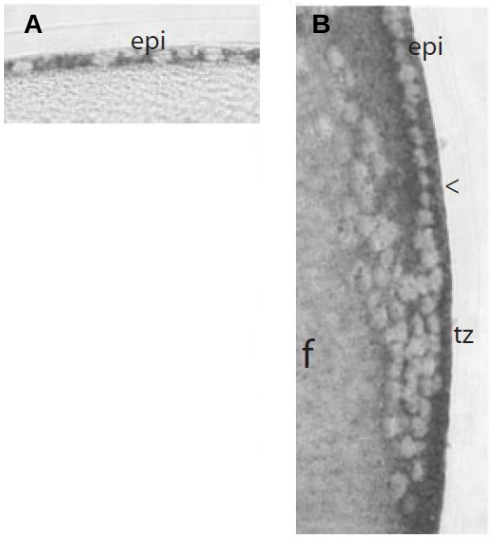


Figure 1.7. *Tmem114 in situ* hybridisation in P21 mouse lens. Staining was observed in the central epithelium (epi) (A) and in the epithelial cells at the lens equator including the at the transition zone (tz) (B). Expression was highest above the lens equator (arrowhead). Staining was not detected in the lens fibres (f). Figure adapted from (Jamieson *et al.* 2007).

Chapter 2: Materials and Methods

2.1 Suppliers

All chemicals used were purchased from Sigma-Aldrich Company Ltd or Fisher Scientific and were of molecular biology grade, unless otherwise indicated in the text. Nuclease-free water (Applied Biosystems) was used for all RNA work and PCR-based techniques. Millipore filtered water (Milli-Q® Integral) was used for all other techniques. All DNA primers used were synthesised by Eurofins MWG.

2.2 Buffers and Solutions

Alkaline phosphatase buffer

100 mM Tris pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, 0.1 % (v/v) Tween-20

Bleaching solution

MetOH/H_20_2/formamide (14:6:1)

Blocking solution for western blots

5 % (w/v) skimmed milk powder (Marvel) in TBST (1 x).

Blotting buffer for western blots

25 mM Tris-HCl, 150 mM glycine, 20 % (v/v) methanol, 1 % (w/v) SDS.

Bouin's fixative

10 % formaldehyde, 5 % (v/v) acetic acid

Glycosylation lysis buffer

0.1 % SDS, 0.1 % β-mercaptoethanol, 0.5 % IGEPAL in PBS

Hybridisation buffer

50 % (v/v) formamide, 5 x SSC buffer (Sigma), 0.1 % (w/v) Torula RNA, 0.01 % (w/v) heparin, 1 x Denhart's solution (Sigma), 0.1 % (v/v) Tween-20, 0.1 % (w/v) CHAPS, 10 mM EDTA.

***In situ* hybridisation blocking reagent (5 x)**

10 % (w/v) DIG blocking reagent (Roche) in MAB. Autoclaved and stored at 2-8°C.

Maleic Acid Buffer (MAB)

100 mM Maleic acid, 150 mM NaCl. Adjusted to pH 7.5

MEMFA Salts (10 X)

1 M MOPS (pH 7.4), 20 mM EGTA, 10 mM $MgSO_4$. Autoclaved

MEMFA Fixative

1 X MEMFA salts, 3.7 % formaldehyde (Baker)

Marc's Modified Ringers (MMR) (10 x)

1 M NaCl, 20 mM KCl, 10 mM $MgSO_4$, 20 mM $CaCl_2$, 50 mM HEPES (pH 7.8). Sterilised by autoclaving.

PBS-Tw

0.1 % Tween-20 in sterile PBS (PAA)

RIPA Buffer

50 mM Tris-HCl (pH 7.4), 1 % (v/v) NP-40, 0.25 % (w/v) sodium-deoxychloate, 1 mM NaCl, complete EDTA-free protease inhibitor cocktail (Roche).

SDS-PAGE running buffer

25 mM Tris-HCl, 20 mM glycine, 5 % (w/v) SDS

SDS-PAGE loading buffer (5 x)

60 mM Tris-HCl (pH6.8), 25 % (v/v) glycerol, 2 % (w/v) SDS, 5 % (w/v) β-mercaptoethanol, 0.025 % (w/v) bromophenol blue.

SSC buffer (20 X)

0.3 M sodium citrate, pH approx. 7.0, containing 3 M NaCl

Sorenson's phosphate buffer

0.2 M NaH_2PO_4 was added to 0.2 M Na_2HPO_4 until the pH reached 7.3

TAE (10 x)

400 mM Tris-acetate, 10 mM EDTA, pH 8

TBST

10 mM Tris-HCl (pH7.5), 15 mM NaCL, 0.1 % (v/v) Tween 20

TE Buffer

10 mM Tris-HCl, 1 mM EDTA

2.3 Bacteriological procedures

2.3.1 Genotypes of bacterial strains used in this study.

Table 2.1. Genotypes of bacteria used in this study.

Bacterial Strain	Genotype
XL-1 blue competent cells	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*q*ZΔM15* Tn*10* (Tetr)].
XL10-Gold® ultracompetent cells	TetrΔ (mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac The [F' proAB lacIqZΔM15 Tn10 (Tetr) Amy Camr]

2.3.2 Preparation of Luria-Bertani (LB) broth or LB agar

LB broth and LB agar was prepared according to the manufacturer's instructions. Antibiotics were used in the following concentrations; ampicillin 0.1 mg/ml, kanamycin 0.025 mg/ml. All LB broth and agar preparations were autoclaved and cooled before antibiotics were added.

2.3.3 Bacterial Transformation using XL-1 blue competent cells

1.7 µl of β-mercaptoethanol was added to 100 µl of chemically competent XL-1 blue cells and incubated for 10 min on ice. Approximately 200 ng of DNA was then added and incubated on ice for 30 mins. The cells were heat shocked at 42 ºC for 45 s and cooled on ice for a further 2 min. 900 µl of LB broth was added to cells prior to incubation period of 1 hr at 37 ºC and 225 rpm. 200 µl of the transformation reaction were plated onto LB agar containing antibiotic and incubated at 37 ºC for a minimum of 16 hr.

2.3.4 Blue/white selection of bacteria

To determine if DNA had inserted into the TA vector pGEM T easy, bacteria were screened for the production of β-galactosidase. Screening of bacteria was performed by plating cells onto LB ampicillin plates containing 0.1 nM IPTG and 40 µg/ml X-gal. Blue colonies indicated lack of insert in the plasmid and were disregarded; the white colonies were selected for further culture.

2.3.5 Bacterial Transformation using XL-10 ultracompetent cells

4 µl of β-mercaptoethanol mix was added to 100 µl of chemically competent XL-10 cells and incubated for 10 min on ice. Approximately 200 ng of DNA was then added and incubated on ice for 30 mins. The cells were heat shocked at 42 ºC for 30 s and cooled on ice for a further 2 min. 900 µl of LB medium was added to cells prior to incubation period of 1 hr at 37 ºC and 225 rpm. 200 µl of the transformation reaction were plated onto LB plates containing antibiotic and incubated at 37 ºC for a minimum of 16 hr.

2.3.6 Bacterial Culture Growth

Single colonies were isolated from an LB plates containing antibiotic and resuspended in 2-5 ml (small scale) or 100 ml (medium scale) of LB broth containing antibiotic. The cultures were incubated for 16 hrs at 37 ºC and with shaking at 225 rpm.

2.4 Cell Culture procedures

2.4.1 Cell Culture conditions

All cells were incubated in their recommended media (stated below) in 75 cm^2 sterile, non-pyrogenic, polystyrene flasks at 37 ºC and 5 % CO_2. When the cells approached confluency they were removed from the flask using trypsin-EDTA (1 x) (PAA) and split at an appropriate level into fresh 75 cm^2 flasks.

2.4.2 Cell Culture Media

MDCK II cells were incubated in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen), supplemented with 10 % (v/v) FBS (PAA), 100 units/ml penicillin/streptomycin (PAA) and 1 % non-essential amino acids (PAA).

HEK293 cells were incubated in Minimum Essential Eagle Medium with L-glutamine (Sigma), supplemented with 10 % (v/v) FBS (PAA), 1 % non-essential amino acids (PAA), 0.25 mg/ml gentamycin (Sigma) and 100 units/ml penicillin/streptomycin (PAA).

FHL124 cells were incubated in DMEM (Invitrogen), supplemented with 5 % (v/v) FBS (PAA) and 100 units/ml penicillin/streptomycin (PAA).

2.4.3 Transient transfections

MDCK II cells were split 24 hrs prior to transfection and were seeded at a density of 1 x 10^6 cells per well of a 6 well plate and transfected with 4 µg of plasmid DNA, or seeded at a density of 5 x 10^5 cells per Transwell® polyester membrane insert (pore size 0.4; Corning) and transfected with 1 µg of plasmid. DNA was transfected with Lipofectamine™2000 or Lipofectamine™LTX (Invitrogen) in accordance with the manufacturer's protocol. Medium was changed after 6 hrs.

HEK293 cells were split 24 hrs prior to transfection and 2.5 x 10^5 cells were seeded per well in a 6 well plate. Cells were transfected with 0.4 µg of plasmid DNA using Effectene® transfection reagent (Qiagen) in accordance with the manufacturer's protocol.

FHL124 cells were split 24 hrs prior to transfection and were seeded at a density of 2.5 X 10^5 cells per Transwell® polyester membrane insert (pore size 0.4; Corning) and transfected with 0.4 µg of plasmid. DNA was transfected with Effectene® transfection reagent (Qiagen) in accordance with the manufacturer's protocol.

2.4.4 Stable cell generation

The expression vector pcDNA3.1(-) contains a Neomycin resistance gene (Figure 9.1). Stable cell lines expressing pcDNA3.1(-) were created by selecting for resistance to G418/Geneticin (PAA). Cloned pcDNA3.1(-) constructs were linearised with the restriction enzyme *Bgl* II, run on an agarose gel and purified by gel extraction (Section 2.5.1.13). MDCK II cells were transfected with the linearised plasmid DNA (Section 2.4.3). 48 hrs post-transfection, cells were split 1:60 in normal maintenance medium supplemented with 1.5 mg/ml G418. Media containing 1.5 mg/ml G418 was changed every 24 hrs for 5 days. Resistant colonies were then isolated and cultured further in 24 well plates. Stable transfected cells were identified by western blot.

2.5 Nucleic acid procedures

2.5.1 DNA methods

2.5.1.1 Polymerase chain reaction (PCR)

Routinely, PCRs were performed using Reddymix™ Custom PCR master mix (Abgene) with extension at 72 °C. When sequence fidelity was essential or for the amplification of long templates, Reddymix™ HiFidelity PCR master mix (Abgene) was used with extension at 68 °C. Reactions were composed of 1 x Reddymix™ master mix, 0.5 µM forward and reverse primers and 20 ng of DNA. 10 % DMSO (v/v) was additionally added for the amplification of GC rich templates. PCR reactions were prepared in 0.2 ml thin-walled PCR tubes or plates (Thermo Scientific) and cycled on an MJ Research Peltier Thermal Cycler PTC-225 or a G-Storm GS482 Thermal cycler**.** Unless otherwise indicated, standard PCR cycling parameters used for all reactions, stated in Table 2.2.

Table 2.2. Cycling parameters for PCR

Cycling stage	Temperature (°C)	Time	Cycles
Initial denaturation	94 °C	5 min	X 1
Denaturation	94 °C	30 s	X 30
Annealing	Y °C	30 s	
Extension	72 / 68 °C	1 min kb^{-1}	
Final Extension	72 / 68 °C	10 min	X 1
Hold	4 °C	Hold	X 1

Y = optimal annealing temperature for each primer pair. Optimal annealing temperatures were identified by running gradient temperatures. Primer sequences and annealing temperatures are available in section 9.3.

For RT-PCR reactions, the expression of a housekeeping gene was used as a control for equal cDNA. The endpoint of all RT-PCRs was determined to be within the exponential phase by performing linearity reactions. Controls for the reverse transcriptase reaction (no reverse transcriptase added to the cDNA reaction – Section 2.5.13) and controls for the PCR reaction (no template control) were also performed. All primers were designed to amplify across at

least one intron and all RT-PCR products were sequenced to confirm indentity (Section 2.5.1.8).

2.5.1.2 Agarose gel electrophoresis

DNA was analysed by gel electrophoresis typically with 2 % (w/v) (for fragments up to 1 kb) and 1 % (w/v) (for fragments over 1 kb) agarose-TAE gels. DNA was mixed with 1 X DNA loading dye (Bioline) and loaded into wells next to a standard size marker (Bioline) and run at typically 90 V in 1 X TAE buffer until the desired separation was achieved. Loading dye was not added to Reddymix samples. Prior to casting 0.01 % (v/v) µl of SafeView Nucleic Acid Stain (NBS Biologicals) was added to all gels to enable the DNA to be visualised on a transilluminator (UV light of wavelength 205 nm).

2.5.1.3 Measurement of DNA concentration

DNA concentration was measured by resolving a known volume on an agarose gel and comparing band intensity with standardised DNA molecular weight ladders (Hyperladder I and IV; Bioline). DNA concentration was also measured with a photometer at 260 nm (BioPhotometer plus, Eppendorf).

2.5.1.4 Column purification of PCR products

PCR products were purified using either Montage Millipore PCR clean up kit or Geneclean Turbo kit in accordance with the manufacturers' instructions and eluted in dH_20.

2.5.1.5 Enzymatic purification of PCR products

When large numbers of PCR products were required to be purified the enzyme ExoSAP-IT® (USB) was used accordance with the manufacturer's instructions.

2.5.1.6 Ethanol precipitation of DNA

To each DNA sample 2.5 X the sample volume of 95 % (v/v) ethanol and 0.1 X the sample volume of 3 M sodium acetate (pH 5.3) was added. When purifying sequencing reactions EDTA was added to a final concentration of 10 mM. Samples were then incubated at room temperature for 15 min. The DNA was pelleted at 16000 RCF for 15 mins and washed in 70 % (v/v) ethanol. The pellet was air-dryed and resuspended in nuclease-free H_20.

2.5.1.7 Phenol/chloroform extraction of DNA

DNA was purified using a phenol/chloroform extraction method when the resultant DNA was required to be of particularly high quality. Samples were made up to 200 µl with TE buffer and an equal volume of phenol:chloroform (1:1) was added. The samples were vortexed for 15 s, and centrifuged at 16,000 RCF for 45 s. The upper phase was removed and transferred to a new 1.5 ml tube. The phenol phase was re-extracted by adding 100 µl of TE buffer, vortexing for 15 s, and centrifuging at 16,000 RCF for 45 s. The upper phase was added to the supernatant from the previous extraction. Next 30µl of 3 M Sodium Acetate pH 5.2, 3 µl of 1 M

$MgCl_2$ and 800 µl 100 % EtOH were added and the sample was vortexed and incubated on dry ice for 1 hr. The sample was pelleted at 14,000 rpm at 4 °C for 15 mins. The supernatant was discarded and the pellet was washed with 70 % EtOH, pelleted for 5 mins and the 70 % EtOH was removed. The pellet was left to air dry and resuspended in nuclease-free dH_20

2.5.1.8 BigDye™ Terminator Cycle Sequencing

Sequencing analysis of plasmid DNA and PCR products was performed using BigDye™ Terminator version 3.1 (Applied Biosystems). The 10 µl sequencing reactions consisted of the following components: 10-250 ng of plasmid DNA or purified PCR product, 0.5 µM sequencing primer, 0.875 x BigDye Sequencing buffer and 1 µl of BigDye™ Terminator v3.1 and dH_20 up to 10. Sequencing reactions were prepared in 0.2 ml thin-walled PCR tubes and cycled on an MJ Research Peltier Thermal Cycler PTC-225. The following cycling parameters were used: 94 °C for 4 mins followed by 25 cycles of 94 °C for 20 s, 55 °C for 20 s and 60 °C for 4 mins.

2.5.1.9 Purification of sequencing products

Ordinarily, DNA was then precipitated in ethanol (Section 2.5.1.6) and the resultant pellet was sequenced by the core facility using (Faculty of Life Sciences, The University of Manchester) an ABI Prism 3100 Genetic Analyser (Applied Biosystems).

For large-scale sequencing in 96 well plates, DNA was purified using Sephadex plates. Sephadex plates were prepared by filling wells of Millipore multiscreen 96 well filter plates with Sephadex and adding 300 µl of dH_2O. Catch plates were placed under the Sephadex plates and the plates were centrifuged at 910 RCF for 5 mins. The catch plates were then emptied, 150 µl of dH_2O was added to each Sephadex plate and the plates were centrifuged at 910 RCF for 5 mins with catch plates underneath. The catch plates were then emptied. A 10 ul volume of Hidi Formamide (Applied Biosystems) was placed into each well of a non-skirted 96 well plate. The sequencing reaction product was placed into the wells of the Sephadex plate, and then centrifuged at 910 RCF for 5 mins into the plate containing Hidi Formamide. Sequencing was then performed on an ABI Prism 3100 (Applied Biosystems) at the core facility (Faculty of Life Sciences, The University of Manchester).

2.5.1.10 Sequence trace analysis

Patient sequences were compared to a wildtype reference sequence using the Staden package Pregap4 1.4b1. All other sequencing reactions were analysed using BioEdit version 7.0.

2.5.1.11 Plasmid DNA Preparation

Plasmid DNA was extracted from overnight bacterial cultures (LB containing the appropriate antibiotic) using QIAprep mini kit® (2-5 ml) and the QIAprep midi kit® (100 ml) (Qiagen) plasmid preparation kits according to manufacture's protocol. DNA samples extracted using

QIAprep mini kit® and the QIAprep midi kit® were eluted with 50 µl and 500 µl of dH_20, respectively.

2.5.1.12 Restriction enzyme digests

Restriction enzyme digests were performed using restriction enzymes (NEB or Roche) and their corresponding buffers according to the manufacturers' protocols.

2.5.1.13 Gel extraction of DNA

PCR fragments or digested DNA fragments were run on agarose gels (Section 2.5.1.2). DNA was visualised on a transilluminator (UV light of wavelength 205 nm) and the desired bands were excised from the gel using a clean scalpel blade. DNA was extracted from the gel fragments using QIAquick Gel® extraction kit (Qiagen) in accordance with manufacturer's protocol. DNA was eluted in nuclease-free H_20.

2.5.1.14 Dephosphorylation of DNA

Alkaline phosphatase was used to prevent the re-ligation of linearised vector DNA with compatible ends. Dephosphorylation of DNA was achieved by adding 2 U of alkaline phosphatase (Roche) to approximately 1 µg of DNA and alkaline phosphatase buffer (1 x) (Roche). Reactions were incubated at 37 ºC for 1 hr. The reaction was terminated by adding one tenth of the volume of 0.2 M EGTA and heating 65 ºC for 10 mins.

2.5.1.15 Ligation of DNA

Following restriction enzyme digests ligation reactions were performed with 50-100 ng vector DNA, 1 U T4 DNA ligase and 1 x buffer (Promega). Vector and insert ratios were determined using the following equation:

$$\frac{\text{vector (ng) x insert (kb)}}{\text{vector (kb)}} \text{ x ratio of insert: vector} = \text{insert (ng)}$$

Generally, ratios of 1:1 and 3:1 insert:vector were used for ligation reactions. Reactions were incubated at room temperature for 4 hrs or 4 ºC overnight. A vector only ligation was used as a negative control.

2.5.1.16 TA cloning

PCR products were cloned using pGEM-T Easy kit (Promega) (Figure 9.2) in accordance with the manufacturer's protocol.

2.5.1.17 Site-directed mutagenesis

Site-directed mutagenesis was performed using the Quikchange Lightning kit (Stratagene) in accordance with the manufacturer's instructions. *Dpn* I digestion was extended to 1 hr.

2.5.2 RNA methods

2.5.2.1 RNA extraction from cells and tissues

RNA was extracted and purified from mammalian cells using the QIAshredder® kit (Qiagen) and the RNeasy® mini kit (Qiagen) according to the manufacturer's instructions. Tissue was snap-frozen, crushed by RNase-free pestle and mortar and RNA from 30 mg of tissue was purified using the QIAshredder® kit (Qiagen) and the RNeasy® mini kit (Qiagen) according to the manufacturer's instructions. RNA was eluted in 30 µl RNase-free dH_20. The integrity of each RNA extraction was confirmed by loading 5 µl of each sample into 2 % agarose gel and confirming the presence of the 28S and 18S rRNA bands. RNA extracted from human eyes at Carnegie stage 21 and Foetal stage 2 was obtained from the MRC-Wellcome Trust Human Developmental Biology Resource at Newcastle University, UK.

2.5.2.2 DNase treatment of RNA

30 µl RNA sample was DNase treated with 2U RNase-free DNase I (Ambion) and 3.5 µl DNase I Buffer at 37 °C for 1 hr. The sample was heated to 70 °C for 10 min to denature the enzyme.

2.5.2.3 Reverse transcription

Approximately 1 µg of RNA, 0.2 µg of random hexamers (Roche) and RNase-free H_20 up to total volume of 12 µl was mixed and heated at 70 ºC for 10 mins and incubated on ice for 5 mins. Next, 10 U of RNase inhibitor (Roche), 1 mM dNTPs, 1 x Bioscript reaction buffer and 50 units of Bioscript reverse transciptase (Bioline) were then added to the reaction and incubated at 37 ºC for 1 h. Reactions were terminated by incubating at 70 ºC for 10 min to denature the reverse transcription enzyme. Control reactions replacing the enzyme with RNase-free H_20 were performed for each sample. All resultant cDNA was stored at 4 ºC and used within 1 week of synthesis.

2.5.2.4 *In vitro* transcription of DIG-labelled probes

Linearised DNA fragments used for *in situ* hybridisation probe transcription were purified by Phenol/Chloroform extraction followed by ethanol precipitation (Section 2.5.1.7). 2.5 µg linearised template DNA was incubated in 1 X transcription buffer (Roche), 2.5 µl DIG RNA labelling mix, 90 U RNA polymerase (T7 or SP6), 20 U RNase inhibitor (Roche) and made up to a volume of 50 µl with nuclease-free H_20 at 37°C for 2 hrs. 1 µg RNAse-free DNase was added and incubated for 10 mins at 37°C. RNA was then purified on a Micro bio-spin P6 column (Biorad) in accordance with the manufacturer's instructions, ethanol precipitated (Section 2.5.1.6) on dry ice and resuspended in nuclease-free H_20. The integrity and concentration of the RNA was identified by heating a sample to 65 °C for 5 mins and running on an agarose gel against a marker of known size and concentration. The probe was then diluted in hybridisation buffer at a concentration of 1 µg/mL and stored at -20 °C.

2.5.2.5 *In situ* hybridisation

The following steps were performed at room temperature unless otherwise stated. Dehydrated *Xenopus tropicalis* embryos (Section 2.7.5) were placed in glass vials and

washed in 75 % MetOH in H_2O, 50 % MetOH in H_2O, 25 % MetOH in PBS-Tw for 5 mins each, followed by two washes in PBS-Tw. The embryos were then incubated in 5 µg/ml Proteinase K for 10 mins, followed by two 5 min washes in 0.1 M triethanolamine (pH 7.5). 2.5 µL of acetic anhydride was added to the 0.1 M triethanolamine sample and incubated for 5 mins. A further 2.5 µL of acetic anhydride was then added and incubated for 5 mins, followed by two washes in PBS-Tw. Embryos were then fixed in 4 % formaldehyde in PBS-Tw, followed by two 5 min washes PBS-Tw. Samples were then incubated in hybridisation buffer for 3 hrs at 60 ºC followed by an over-night incubation at 60 ºC in hybridisation buffer containing a DIG-labelled RNA probe (1 µg/mL).

The probe was removed and stored at –20 ºC for future use. Probes were re-used no more than three times. Embryos were washed in hybridisation buffer for 5 mins followed by four 20 min washes in 2 X SSC buffer. Embryos were RNase-treated with 10 µL RNase cocktail (Roche) per ml of 2 X SSC for 30 mins at 37ºC, and the RNase was removed by washing with 2 X SSC at room temperature for 10 mins. Samples were subsequently washed twice with 0.2 X SSC at 60ºC for 20 mins each. SCC buffer was removed by washing the samples twice for 5 mins in MAB. Samples were blocked for 1 hr in 1 x *in situ* blocking solution in MAB, and incubated with anti-DIG alkaline phosphatase Fab fragments in blocking solution containing 20 % (v/v) heat-treated Lamb serum (HTLS) for 5 hrs. Excess antibody was removed by washing three times in MAB for 5 mins followed by an overnight wash in MAB at 4ºC. Samples were then washed two times in alkaline phosphatase buffer and incubated in the dark in alkaline phosphatase buffer containing 4.5 µL/mL NBT (Roche) and 3.5 µL/mL BCIP (Roche). When a satisfactory signal was obtained, samples were washed two times in MAB for 5 mins and fixed in Bouin's fixative for 1 hr. Embryos were then rinsed in 70 % EtOH and washed for 5 mins in MetOH and bleached on a light box in bleaching solution. Bleaching solution was gradually replaced with PBS and samples were stored in PBS at 4ºC.

2.5.2.6 *In vitro* transcription of capped mRNA for injection

Template DNA was linearised downstream of the termination codon and purified by phenol/chloroform extraction. Capped mRNA was transcribed using the mMESSAGE mMACHINE® Kit (Ambion). The following reaction was assembled at room temperature: 10 µl NTP/CAP, 2 µl 10 x transcription buffer, 2 µl T7 polymerase, 1 µg template DNA and H_20 up to 20 µl. The reaction was performed at 37 °C for 2 hrs. Template DNA was digested by the addition of 1 U RNase-free DNase I for 10 mins at 37 °C. RNA was purified on a Micro bio-spin column (Section 2.5.2.4) and ethanol precipitated (Section 2.5.1.6) on dry ice and resuspended in nuclease-free H_20. The integrity and concentration of the RNA was identified by heating a sample to 65°C for 5 mins and running on an agarose gel against a marker of known size and concentration.

2.6 Protein Procedures

2.6.1 Generation of custom polyclonal antibodies

2.6.1.1 Sigma-Genosys custom polyclonal antibodies

Synthetic peptides and antibody production was performed by Sigma-Genosys. KLH-conjugated peptides were synthesised and purified on a Discovery HS C18 HPLC Column (3 µm) by Sigma-Genosys. New Zealand White Rabbits were selected as the host animal. Single peptides were used to immunise two rabbits. Peptides and their final purity are as follows:

Protein	Rabbit Numbers	Peptide	Purity
Tmem114	4671 & 4672	KLH-[C]-ARALSLSQRQDQAI	59 %
Tmlp1	4673 & 4674	KLH-[C]-SLSQRPGVPHSVIL	66 %

Purified antibodies were supplied as solutions. 1 mg of dehydrated peptide was resuspended in 50 µl acetic acid and made up to 1 ml with sterile PBS as according to manufacturer's instructions.

Table 2.3. Sigma-Genosys antibody production schedule. *estimated. Complete Freund's Adjuvant (CFA); Incomplete Freund's Adjuvant (IFA).

Time	Procedure	Additional information
Day 0	Pre-immune bleed Immunisation #1	 200 µg peptide in CFA
Day 14	Immunisation #2	100 µg peptide in IFA
Day 28	Immunisation #3	100 µg peptide in IFA
Day 42	Immunisation #4	100 µg peptide in IFA
Day 49	Test bleed #1	5 ml
Day 56	Immunisation #5	100 µg peptide in IFA
Day 63	Test bleed #2 ELISA	5 ml
Day 70	Immunisation #6	100 µg peptide in IFA
Day 77	Test bleed #3	5 ml
Day 155	Immunisation #7	100 µg peptide in IFA
Day 162	Test bleed #4	5 ml
Day 179	Immunisation #8	100 µg peptide in IFA
Day 186	Final Bleed	Tmem114: 56 ml, 52 ml Tmlp1: 61 ml 68 ml
Day 214*	Purification ELISA	

2.6.1.2 New England Peptide custom polyclonal antibodies

Synthetic peptides and antibody production was performed by New England Peptide. Peptides and their final purity are as follows:

Protein	Rabbit Numbers	Peptide	Purity
Tmem114	F4140 & F4141	KLH-[C]-ARALSLSQRQDQAI	>85 %
Tmlp1	F4142 & F4144	KLH-[C]-SLSQRPGVPHSVIL	>85 %

1 mg of dehydrated purified antibody was resuspended in 1 ml distilled H_20 as according to manufacturer's instructions. Resuspended purified antibodies were aliquoted and stored at -80°C.

Table 2.4. New England Peptide antibody production schedule

Time	Procedure	Additional information
Day 0	Pre-immune bleed Immunisation #1	5 mL per rabbit peptide in CFA
Day 14	Immunisation #2	peptide in IFA
Day 28	Immunisation #3	peptide in IFA
Day 42	Immunisation #4	No details available
Day 49	Production bleed #1 ELISA	40-50 mL
Day 56	Immunisation #5	No details available
Day 63	Production bleed #2 ELISA	40-50 mL
Day 70	Immunisation #6	No details available
Day 77	Production bleed #3	40-50 mL
Day 84	Final bleed Purification	 40-50 mL antiserum per column

2.6.2 Antibodies used in this study

Table 2.5. Primary antibodies used in this study. Abbreviations: Western blot (WB), immunofluorescence (IF), immunohistochemistry (IHC), overnight (O/N).

Primary Antibody	Technique	Supplier	Temp	Time	Working dilution
Custom polyclonal anti-mTmem114 final bleed	WB, IHC IF	Sigma	4 °C RT	O/N 1 hr	1:1000
Custom polyclonal anti-mTmem114 purified	WB, IHC IF	Sigma	4 °C RT	O/N 1 hr	1:20
Custom polyclonal anti-mTmlp1 final bleed	WB IF	Sigma	4 °C RT	O/N 1 hr	1:1000
Custom polyclonal anti-mTmlp1 final bleed	WB IF	NEP	4 °C RT	O/N 1 hr	1:1000
Custom polyclonal anti-mTmlp1 purified	WB IF	NEP	4 °C	O/N 1 hr	1:200
Mouse monoclonal anti-V5 tag IgG_{2a}	WB, IF	AbD Serotec	RT	1 hr	1/2,000
Rabbit polyclonal anti-ZO-1 (61-7300)	IF	Invitrogen	RT	1 hr	1/1,000
Rat monoclonal anti-ZO-1 (R26.4C)	IF	Developmental Studies Hybridoma Bank	RT	1 hr	1:200
FITC conjugated mouse monoclonal anti-E-Cadherin IgG_{2a} (612130)	IF	BD Biosciences	RT	45 mins	1/1,000
Mouse monoclonal anti–GM130 IgG_1	IF	BD Biosciences	RT	1 hr	1/1,000
Goat polyclonal anti-GRP94	IF	Santa Cruz Biotechnology	RT	1 hr	1/200-1/500
Mouse monoclonal anti–GAPDH $1gG_1$	WB	Santa Cruz Biotechnology	RT	1 hr	1/20,000
Goat polyclonal anti-β-actin	WB	Santa Cruz Biotechnology	RT	1 hr	1/20,000
Rabbit primary antibody isotype control (08-6199)	IHC	Invitrogen	4 °C	O/N	Neat

Table 2.6. Secondary antibodies used in this study. Abbreviations: Western blot (WB), immunofluorescence (IF), immunohistochemistry (IHC).

Secondary Antibody	Technique	Supplier	Incubation temperature	Incubation Time	Working dilution
Goat anti-mouse HRP	WB	Pierce	RT	45 mins	1/2,000
Goat anti-rabbit HRP	WB	Pierce	RT	45 mins	1/2,000
Rabbit anti-goat HRP	WB	Pierce	RT	45 mins	1/2,000
AlexaFluor 488 chicken anti-mouse IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000
AlexaFluor® 488 goat anti-mouse IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000
AlexaFluor® 594 goat anti-mouse IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000
AlexaFluor® 488 goat anti-rabbit IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000
AlexaFluor® 594 goat anti-rabbit IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000
AlexaFluor® 488 goat anti-rat IgG (H+L)	IF	Invitrogen	RT	45 mins	1/2,000

Table 2.7 Labelled stains used for immunofluorescence and immunohistochemistry

Name	Target	Supplier	Incubation temperature	Incubation Time	Working Conc.
Phalloidin-Tetramethylrhodamine B isothiocyanate (P1951)	Filamentous actin	Sigma-Aldrich	RT	45 mins	2.5µg/mL
VECTASHIELD® Hard Set Mounting Medium with DAPI (H-1500)	DNA	Vector	RT	-	1.5µg/mL

2.6.3 Sample lysis and preparation

Cells or tissues homogenised by pestle and mortar were lysed in ice cold RIPA buffer containing complete protease inhibitor cocktail (Roche), rotated for 30 mins at 4 °C, centrifuged at 4 °C for 15 mins. The supernatant was taken to a new tube to which 5 X SDS loading buffer was added.

2.6.4 *In vitro* transcription and translation

Proteins were *in vitro* transcribed and *in vitro* translated from plasmid DNA using the Human *In Vitro* Glycoprotein Expression Kit for DNA Templates (Pierce) in accordance with the manufacturer's instructions.

2.6.5 De-glycosylation assays

2.6.5.1 N-glycosidase F (PNGase F) de-glycosylation assay

For N-glycosylation assays, cells were lysed in glycosylation lysis buffer and EDTA was added to a final concentration of 50 mM. 1 U PNGase F was then added to 25 µl samples and incubated at 37 °C overnight. 5 µl of 5 X SDS loading buffer was added to each sample to stop the reaction.

2.6.5.2 Endo H de-glycosylation assay

2 µl of 10 X denaturation buffer was added to 18 µl of RIPA lysates and the samples were incubated at room temperature for 10 mins. 2.5 µl of G5 reaction buffer and 2 µl of Endo H (Roche) were added and incubated at 37 °C for 2 hrs. The reaction was stopped by the addition of 6 µl of 5 X SDS loading buffer.

2.6.6 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was used for the detection and analysis of protein. Prior to SDS-PAGE, samples denatured at 37 °C for 20 mins. The denatured proteins were then run on an SDS denaturing gel at 120-150 volts in 1 X running buffer next to an appropriately sized molecular weight marker (Bio-rad). Components of the gels were as follows:

Stacking Gel

625 mM Tris HCl pH 6.8, 0.1 % (w/v) SDS, 4 % (v/v) acrylamide/bis, 0.075 % (w/v) ammonium persulphate, 0.2 % (v/v) Tetramethylethylenediamine (TEMED).

Resolving Gel

375 mM Tris HCl pH 8.8, 0.1 % (w/v) SDS, 12-15 % (v/v) acrylamide/bis, 0.125 % (w/v) ammonium persulphate, 0.1 % (v/v) TEMED

2.6.7 Coomassie stain

SDS-PAGE gels washed three times in dH_20 for 5 mins, then incubated in EZBlue^{TM} gel staining reagent (Sigma) for 1 hr at room temperature and then washed in dH_20 until bands appeared visible and background stain was removed. The whole protocol was performed on a rocking platform. Gels were stored in dH_20 at 4 °C.

2.6.8 Western Blot Analysis

Following SDS-PAGE gels were washed in transfer buffer and layered between a pre-soaked nitrocellulose membrane and three layers of filter paper in a semi-dry blotter (Bio-Rad). Protein was transferred onto the membrane at 15 volts for 1 hr. The membrane was then incubated with the following solutions on a rocking or rotating platform. Membranes were blocked in 5 % (w/v) skimmed milk powder (Marvel) in TBST for 2 hrs at room temperature. Membranes were then incubated in primary antibody diluted in 5 % skimmed milk powder/TBST as according to Table 2.5. The membranes were then washed 3 times for 10 mins in TBST. Next membranes were incubated in secondary antibody in skimmed milk powder in TBST for 45 mins at room temperature (Table 2.6). The membranes were then washed 3 times for 10 mins in TBST. Finally, the membranes were incubated with enhanced chemiluminescence Western Blotting Detection Reagents (Amersham Bioscience) for 1 min, and then exposed onto Hyperfilm enhanced chemiluminescence (Amersham Bioscience) until bands were visible. Films were developed in GBX Developer (Kodak), fixed in GBX fixer (Kodak) and washed in dH_20.

2.6.9 Western blot stripping

When re-probing of a membrane was required, the nitrocellulose membrane was stripped by incubating in Restore Western Blot stripping buffer (Pierce) for 30 mins at 37 °C. The membrane was then washed three times in TBST for 10 mins and then blocked and probed with antibodies as described in Table 2.5 and Table 2.6.

2.6.10 Peptide competition assay

Antibodies were incubated with an equal volume of peptides at a concentration of 1 mg/ml at room temperature for 2 hrs. The antibody was then diluted to the optimised working dilution (Table 2.5).

2.6.11 Mass spectrometry

MALDI (Matrix Assisted Laser Desorption/Ionisation) TOF-TOF (Time Of Flight) mass spectrometry was performed by the Bimolecular analysis core facility (Faculty of Life Sciences, The University of Manchester) using an Ultraflex II (Bruker). Bands of interest were excised from EZBlue-stained SDS-polyacrylamide gels using sterile razor blades. Gel pieces were incubated in acetonitrile for 5 mins and then dried by vacuum centrifugation. The protein within each sample was then reduced by incubating in 10 mM dithiothreitol in 25 mM ammonium hydrogen carbonate at 56 °C for 1 hr. The samples were incubated in alkylating agent, 55 mM Iodoacetamide, in 25 mM ammonium hydrogen carbonate at 22 °C for 45 mins in a dark room. The samples were washed with 25 mM ammonium hydrogen carbonate, acetonitrile, 25 mM ammonium hydrogen carbonate and acetonitrile for 5 mins each respectively. The gel pieces were dried by vacuum centrifugation. Samples were incubated in 10 % (v/v) trypsin solution (12.5 ng/µl) in 25 mM ammonium hydrogen carbonate at 4 °C for 45 mins and then at 37 °C for 16 hrs. Peptides were extracted from the samples with 20 mM ammonium hydrogen carbonate for 20 mins followed by two further extractions with 5 % formic acid in 50 % acetonitrile. Supernatant samples were pooled and concentrated. Peptide samples were stored at -20 °C until mass spectrometry analysis was performed.

Nano-scale liquid chromatographic tandem mass spectrometry (nLC-MS/MS) was performed by the Fingerprints Proteomics Facility at the University of Dundee. Bands of interest were excised from EZBlue-stained SDS-polyacrylamide gels using sterile razor blades, and shipped to the Fingerprints Proteomics Facility at ambient temperature in 1.5 ml sterile tubes. Gel bands were further sectioned and the proteins in the gel were digested by and in-gel digestion procedure in a laminar flow cabinet. The digests were then fractioned by Strong Cation eXchange into 10 or more fractions prior to 1D nLC-MS-MS. Obitrap protein identification analysis was performed on a 4000 QTRAP (Applied Biosystems) MS-MS system by personnel at the Fingerprints Proteomics Facility. Identified protein fragments were compared to the International Protein Index (IPI) and Genbank databases.

2.6.12 Immunofluorecent staining of cell lines

Non-polarised cells were grown on chamber slides (Scientific Lab Supplies Ltd). In order to polarise MDCK II cells the cells were grown on Transwell® polyester membrane inserts (pore size 0.4; Corning) for 48 hrs. Cells were fixed 48 hrs post-transfection. Cells were then rinsed 3 times with PBS on ice and cells were fixed in a methanol:acetone mix (1:1) for 10 mins and rehydrated with PBS for 10 mins. Alternatively, cells were rinsed 3 times with PBS on ice and cells were fixed with 4 % paraformaldehyde in PBS for 20 min. Cells were then washed for 2 x 5 min with PBS. Autofluorescence of the sample was quenched by 3 x 5 min washes with 0.25 % NH_4Cl. Cells were subsequently permeabilised with 0.1 % Triton X-100 for 10 min and washed for 2 x 5 min with PBS.

Cells were blocked in 5 % FBS in PBS for 30 min at room temperature and incubated in primary antibody in block solution at room temperature for 1 h. Cells were then washed with 3 x 5 min PBS plus 0.1 % Tween-20 (PBS-Tw) and incubated for a further 20 mins in block solution. Cells were then incubated with secondary antibody(s) in block solution at room temperature for 1 h and subsequently washed for 3 x 5 min with PBS-Tw. Transwell filters were then cut from the inserts and mounted onto microscope slides. The chambers on chamber slides were removed. Cover slips were finally mounted onto a microscope slide with VECTASHIELD® Hard Set Mounting Medium with DAPI (H-1500) (Vector Laboratories) and sealed with nail varnish. Stained cells were visualised on a Leica DM 5000 microscope and for the majority of experiments, images were taken on a Nikon upright confocal microscope (Section 2.6.14).

2.6.13 Immunohistochemistry

For staining, OCT embedded sections were washed four times for 5 mins in PBS, and blocked for 1 hr at room temperature in 5 % goat serum (Abcam). Slides were then incubated in primary antibody in blocking solution or a rabbit isotype control at 4 °C overnight in a humid container. Slides were washed two times for 10 mins in PBS followed by 10 mins in blocking solution. Blocked slides were incubated with fluorescent secondary antibodies at appropriate concentrations in a dark chamber for 45 mins at room temperature and then washed three times in PBS for 10 mins. Cover slips were finally mounted with VECTASHIELD® Hard Set Mounting Medium with DAPI (H-1500) (Vector Laboratories) and sealed with nail varnish. Stained cells were visualised on a Leica DM 5000 microscope and in some cases images were taken on a Nikon upright confocal microscope (Section 2.6.14).

2.6.14 Confocal microscopy

Images were collected using a Nikon C1 confocal on an upright 90i microscope with a 60 x 1.40 Plan Apo objective and 3 x confocal zoom. The confocal settings were as follows; pinhole 30 µm, scan speed 400 Hz unidirectional, format 1024 x 1024. Images for DAPI, Alexa Fluor 488 and Alexa Fluor 594 were excited with the 405 nm, 488 nm and 594 nm laser lines respectively.

2.6.15 Transmission electron microscopy

Stable MDCK II cells expressing mTmem114 were cultured on Transwell® polyester membrane inserts (pore size 0.4; Corning) for 5 days. Samples were fixed with 4 % PFA in 0.1 M phosphate buffer overnight. The fixed cells were then prepared for imaging by Dr. Aleksandr Mironov, an experimental officer in the Electron Microscopy Facility, Faculty of Life Sciences University of Manchester. The samples were dehydrated through the series of ethanol and embedded in LR White resin that was polymerized over 48 hrs at 4 °C. Sections of 70 nm thickness were labelled with primary and subsequently secondary antibodies conjugated with 10 nm colloidal gold. After labelling sections were contrasted with 1 % uranyl acetate in water for 10 min. Images were acquired with FEI Tecnai 12 Biotwin microscope at 80 kV accelaration voltage by Dr. Aleksandr Mironov.

2.7 *Xenopus tropicalis* procedures

2.7.1 Induction of ovulation and obtaining embryos

To prime adult females for ovulation, 15 U of pregnant mare serum gonadotropin (PMSG) was injected into the dorsal lymph sac 72 hrs before full induction. To induce ovulation, females were injected with 150 U of human chorionic gonadotropin (hCG). The male was also injected with 150 U of hCG. The oocytes were collected 3-4 hrs post-induction. Oocytes were collected manually by squeezing the abdomen of the females and then placed in a Petri dish coated with L-15 media supplemented with 10 % FBS.

2.7.2 Fertilisation of oocytes

Adult males were anaesthetised in 0.1 % MS222 for 25 mins. A reflex test was performed to confirm anaesthetisation and the male was sacrificed via pithing methods. Testes were removed and homogenised in L-15 containing 10 % FBS using mini-homogeniser/1.5 ml tube immediately before use. Homogenised testes in L-15 containing 10 % FBS were pipetted on to unfertilised oocytes and incubated for 2 mins. The fertilising oocytes were then covered with 0.1 x MMR for 20 mins, removed, and replaced with 2 % free-base cysteine in 0.1 x MMR at pH 7.8-8.0 for 20 mins to remove the jelly from the embryos. Cysteine was removed by washing three times in 0.1 x MMR, with a final wash into 0.1 x MMR + 3 % Ficoll.

2.7.3 Injection of embryos

Embryos in 3 % Ficoll in 0.1 x MMR were placed on 800 µm meshes in petri dishes. The solution was removed until the surface tension holds the embryos in place. Typically, 1 nl of diluted morpholino, DNA or capped mRNA was injected at the 1-2 cell stage for bilateral injections, or injected into one cell at the 4 cell stage for unilateral injections using a Picospritzer microinjector (Nanotech). Injected and uninjected embryos were then placed in 3 % Ficoll in 0.1 x MMR on agarose (1 % (w/v) in 0.01 x MMR). Dividing embryos were counted and sorted into 6-well dishes coated with 2.5 % agarose in 0.1 x MMR and incubated in 0.1 x MMR with 1 x gentamycin at 22-25 °C. Embryos were removed to normal culture at stage 4.

2.7.4 Culture and phenotyping of embryos

Embryos were routinely cultured to stage 45 in 6-well plates (approximately 50 embryos per well) in 0.1 x MMR containing 1 x gentamycin. Developmental stages were assessed in accordance with Nieuwkoop and Faber (1967). Phenotyping of morpholino-injected embryos was performed at stages 41-43. Phenotyping of EGFP-injected embryos was performed at stages 33-35. Phenotyping of capped mRNA-injected embryos was performed 16 hrs post-injection.

2.7.5 Fixing embryos for *in situ* hybridisation

Membranes were removed from embryos manually using forceps. The embryos were incubated in MEMFA for 2 hrs, washed in 30 % MetOH: 70 % MEMFA, 50 % MetOH/:50 % MEMFA, 70 % MetOH: 30 % MEMFA, for 10 mins each, washed in 100 % MetOH and stored in fresh 100 % MetOH at -20 ºC.

2.8 Histology

2.8.1 OCT embedding and sectioning

Lenses from sacrificed adult mice were washed in PBS on ice, fixed in 4 % paraformaldehyde at room temperature for 10 mins, and washed three times for 10 mins in PBS. Lenses were mounted, and sectioned by the University of Manchester core histology facility. Fixed lenses were mounted in OCT compound on pre-chilled chucks. 10 mm thick sections (equatorial plane) were cryosectioned at -18 °C on a cryostat (Shandon AS260) and placed on poly-L-lysine-coated slides. Slides were stored at – 80 °C until use.

2.8.2 Parrafin embedding and sectioning

Whole sacrificed *Xenopus tropicalis* embryos were washed in sterile PBS on ice, and fixed in MEMFA overnight at 4°C. Samples were then dehydrated in 30 %, 50 % and 70 % EtOH-sterile PBS for 1 hr each. Samples were then processed for paraffin was embedding on a Shandon Citadel 2000 tissue processor using the following protocol. Samples were immersed in 70 % IMS three times for 30 mins, followed by three immersions in 90 % IMS for 30 mins each. Samples were finally dehydrated in 100 % IMS three times for 30 mins. Dehydrated tissues were then immersed in xylene three times for 30 mins each at room temperature. Finally samples were immersed in paraffin at 58 °C two times for 1 hr each. Samples were then manually orientated and embedded in paraffin at 58 °C. Once set, the paraffin block was sectioned at 5 µm using a rotary microtome. Sections were placed on poly-lysine-coated slides and dried overnight at 37 °C.

2.8.3 Haematoxylin and eosin staining of sections

Sections were stained with hematoxylin and eosin by immersing in the following reagents consecutively at room temperature (Table 2.8). Slides were then mounted in Histomount mouting medium (National diagnostics) and covered with a glass cover slip.

Table 2.8 Haematoxylin and eosin staining solutions

Wash	Reagent	Time (mins:ss)	Wash	Reagent	Time (mins:ss)
1	Xylene	5:00	13	Tap H_20	2:00
2	Xylene	3:00	14	Tap H_20	5:00
3	Xylene	3:00	15	Eosin	0:30
4	100 % IMS	3:00	16	Distilled H_20	1:00
5	100 % IMS	2:00	17	70 % IMS	1:00
6	100 % IMS	2:00	18	95 % IMS	1:00
7	95 % IMS	2:00	19	100 % IMS	2:00
8	70 % IMS	2:00	20	100 % IMS	2:00
9	Tap H_20	2:00	21	100 % IMS	2:00
10	Haematoxylin	4:00	22	Xylene	2:00
11	Tap H_20	3:00	23	Xylene	3:00
12	Acid alcohol	0:04	24	Xylene	5:00

2.9 Bioinformatics

The URLs for software used are available in Section 9.1. Accession numbers for sequences used are available in Section 9.2.

2.9.1 Topology prediction

Topologies were predicted using default settings of the prediction tools HMMTOP, TMMOD, TMPRED, TOPPRED, TMHMM and Philius.

2.9.2 Signal peptide prediction

Signal peptides were predicted using both neural network (NN) and Hidden Markov Model methods in the SignalP3.0 software. Signal peptides were also predicted using default settings of Philius.

2.9.3 Glycosylation prediction

N- and O-glycosylation sites were predicted using default settings of the NetNGlyc1.0 Server and NetOGlyc3.1 Server, respectively.

2.9.4 Phosphorylation prediction

Phosphorylation sites were predicted at tyrosine, threonine and serine residues using the NetPhos2.0 software and the kinases predicted to act at the identified sites were predicted using NetPhosK1.0 software.

2.9.5 Palmitoylation prediction

Palmitoylation sites were predicted using the CSS-Palm2.0 software with the high threshold setting.

2.9.6 Sulphation prediction

Sulphation sites were predicted using the Sulfinator prediction tool.

2.9.7 Molecular mass prediction

The molecular mass of proteins was predicted using the average setting on the Compute pI/MW tool.

2.9.8 Multiple alignments

Typically multiple alignments were created using ClustalW with default settings at the EBI server or using default settings of ClustalW embedding in the BioEdit software. On some occasions manual readjustments to the alignment were made. For phylogenetic tree construction sequences were aligned using default settings of the ClustalW embedded in the Seaview software.

2.9.9 Pairwise alignment

Percentage pairwise identities were determined using default settings of the ALIGN software.

2.9.10 Phylogenetic tree construction

Phylogenetic trees were constructed using Seaview software. Trees were constructed from aligned sequences (Section 2.9.8) using the neighbour-joining method based on observed distances with gap sites ignored. Bootstrap proportions (Felsenstein 1985) were used to assess the robustness of tree with 10,000 bootstrap replications. Bootstrap values of less than 50 were not considered significant. The protein Clarin-1 was used as an outgroup as recommended by (Adato *et al.* 2002; Price *et al.* 2005).

2.9.11 Gene structure identification

Where available the gene structure was retrieved from the Ensembl database. The gene structure of TMEM114 was previously reported (Jamieson *et al.* 2007). The gene structure of TMLP1 was determined by comparing the cDNA sequences to genomic DNA sequences.

Chapter 3: In silico analysis of TMEM114 and TMLP1

3.1 Introduction

As is evident from Chapter 1, there is little reported information about *TMEM114* and the protein it encodes. Jamieson *et al*., who first identified *TMEM114,* determined its exomic structure and the 223 amino acid encoded protein. The protein was found to have a predicted transmembrane topology that was similar to that of the PMP-22/EMP/MP20 family of proteins. However, the relationship of TMEM114 to these other proteins was not investigated. Some TMEM114 orthologues were identified such as in mouse, chicken and zebrafish, but information regarding some intermediate species and lower species was lacking. Also, the presence or absence of functional motifs and post-translational modifications was not reported. Therefore, in this chapter, the aim was to expand upon the available information about TMEM114 using in silico techniques. This information would then a) lead to greater understanding of TMEM114's relationships and functionality and b) be used to devise appropriate experiments to investigate its functional role. Prior to the undertaking of this project, as part of a Blastp (http://blast.ncbi.nlm.nih.gov/) database search using the TMEM114 protein sequence to identify putative orthologues, a novel putative human protein was identified. This chapter thus encompasses the analysis of TMEM114 and this novel putative protein of which no analysis had previously been performed.

3.2 Identification of TMLP1

As stated in the introduction, a novel putative protein (XP_946151.1) was identified which is encoded by a gene with five exons located at 17q25.3 (LOC283999) (Urquhart, unpublished) (see Section 3.8 for exon structure). The encoded protein is 223 amino acids in length, the same length as TMEM114. When both proteins are compared by pairwise alignment, they share 43 % amino acid identity and 61 % similarity (Figure 3.1). Due to its similarity to TMEM114, this protein has been named TMEM114-like protein 1 (TMLP1).

Figure 3.1. Amino acid sequence of TMLP1 aligned with TMEM114. Alignment of human TMEM114 with its paralog TMLP1 showing their similarity. Black highlight = identical amino acid. Grey highlight = similar amino acids. Sequences were aligned with ClustalW.

3.3 Predicted topologies of TMEM114 and TMLP1

3.3.1 Topology prediction

The transmembrane topology of TMEM114, predicted using the transmembrane topology prediction tool TMHMM (Sonnhammer *et al.* 1998), has previously been reported (Jamieson *et al*. 2007). TMEM114 has four predicted transmembrane domains (TMDs) with two

extracellular loops (ECLs) and short intracellular amino- (N-) and carobxy- (C-) termini. To support this predicted topology, four further transmembrane topology prediction tools were used. Although there are slight differences in the positioning of the TMDs the overall topology of TMEM114 as having four TMDs with a short intracellular C-terminus is retained (Table 3.1).

Table 3.1. Topology predictions for TMEM114 and TMLP1. Numbers refer to amino acid numbering from N- to C-termini. TMD = transmembrane domain

TMEM114

	HMMTOP	TMMOD	TMPRED	TOPPRED	TMHMM
Inside	1-5	1-11	1-7	1-7	1-4
TM1	**6-29**	**12-35**	**8-33**	**8-28**	**5-27**
Outside	30-105	36-105	34-108	29-103	28-106
TM2	**106-129**	**106-130**	**109-135**	**104-124**	**107-129**
Inside	130-135	131-133	136-136	125-132	130-133
TM3	**136-159**	**134-158**	**137-161**	**133-153**	**134-156**
Outside	160-185	159-188	162-193	154-182	157-186
TM4	**186-210**	**189-209**	**194-213**	**183-203**	**187-209**
Inside	211-223	210-223	214-223	204-223	210-223

TMLP1

	HMMTOP	TMMOD	TMPRED	TOPPRED	TMHMM
Inside	1-3	1-3	1-3	1-3	1-4
TM1	**4-27**	**4-27**	**4-27**	**4-24**	**5-27**
Outside	28-95	28-95	28-95	25-94	28-96
TM2	**96-120**	**96-106**	**95-120**	**95-115**	**97-119**
Inside	121-124	117-122	121-124	116-121	120-125
TM3	**125-149**	**123-148**	**125-147**	**122-142**	**126-148**
Outside	150-173	149-171	148-188	143-188	149-209
TM4	**174-198**	**172-196**	**189-209**	**189-209**	
Inside	199-203	197-223	210-223	210-223	

To determine if TMLP1 also has a transmembrane topology, all five transmembrane topology prediction tools were used (Table 3.1). All five topology prediction tools predicted TMLP1 to be a transmembrane protein, four of which predicted it to have four transmembrane domains (i.e a tetraspan topology). The predicted topology consists of a short intracellular N-terminus, two extracellular loops, the first (ECL1) of which is larger than the second (ECL2), a short intracellular loop between TMD2 and TMD3 and an intracellular C-terminus. Of the four, two tools [HMMTOP (Tusnady and Simon 2001) and TMMOD (Kahsay *et al.* 2005)] predicted an ECL2 of <24 amino acids resulting in an internal C-terminus of ~25 amino acids. TMPRED (Hofmann and W. 1993) and TOPPRED (von Heijne 1992) predicted ECL2 to be ~40 amino acids, shortening the intracellular C-terminus to 13 amino acids. This is a similar topology to that predicted for TMEM114. The prediction tool TMHMM, which predicts TMEM114 to have

four transmembrane domains (Jamieson *et al*. 2007), predicted TMLP1 to have just the first three TMD domains with a long extracellular C-terminus .

3.3.2 Signal peptide prediction

Eukaryotic membrane proteins contain signal peptides which are necessary for the insertion of the proteins into the endoplasmic reticulum (ER) membrane. The prediction tool SignalP (Bendtsen *et al.* 2004) predicts the probability of a protein containing a signal peptide and the probability that the predicted peptide is cleaved or uncleaved after transport. Predictions are made using one of two methods: neural networks (SignalP-NN) or Hidden Markov Models (SignalP-HMM). Signal peptides were predicted in both TMEM114 and TMLP1. For TMEM14 SignalP-NN predicted a cleavage site of the predicted signal peptide between residues p.A26 and p.I27 and SignalP-HMM between residues p.G28 and p.T29. This predicted signal peptide encompasses the first transmembrane domain (TMD1), which is predicted to end at p.I27 (Jamieson *et al*. 2007). Cleavage sites for signal peptide in TMLP1 were predicted between residues p.S28 and p.29D (SignalP-NN) or between p.A27 and p.S28 (SignalP-HMM) which also encompass TMD1 (Table 3.1).

3.3.3 Combined signal peptide and topology prediction

The predicted signal peptides for TMEM114 and TMLP1 consist of the intracellular N-terminus and TMD1. Cleavage at these sites would result in loss of TMD1, making the N-terminus in the mature protein extracellular. However, in the presence of downstream TMDs, predicted signal peptides may remain uncleaved (known as signal anchors) acting as the first transmembrane domain of the protein (Schatz and Dobberstein 1996). There are some examples of proteins with multiple transmembrane proteins with signal peptides, e.g. calcitonin receptor (Uniprot P30988). The combined signal peptide and topology prediction tool Philius (Reynolds *et al.* 2008) is designed specifically for transmembrane proteins, was used to test prior predictions made (Section 3.3.1 and 3.3.2). TMEM114 is predicted to have a cleaved signal peptide with cleavage occurring between amino acids p.I27 and p.G28 (Figure 3.2). Confidence in the signal peptide prediction was 63 %. The TMD and ECL predictions were similar to previously described (Jamieson *et al*. 2007). Confidence in this prediction was 63 % for the topology and 66 % for the overall prediction. TMLP1 was predicted to have a similar structure, with cleavage occurring between residues p.S28 and p.D29, as predicted by SignalP-NN (Figure 3.2). Confidence in the signal peptide prediction was high (97 %). Philius predicted an intercellular C-terminus of 23 amino acids, similar to that predicted by HMMTOP and TMMOD (Table 3.1). Confidence in this prediction was 83 % for the topology and 99 % for the overall prediction.

Figure 3.2. Signal peptide and membrane topology for TMEM114 and TMLP1 as predicted by Philius. TMEM114 contains a signal peptide (red) of 27 amino acids and three transmembrane domains (yellow). Between the transmembrane domains lie large extracellular domains (green) and short intracellular regions (blue). TMLP1 contains a signal peptide (red) of 28 amino acids and three transmembrane domains (yellow). A long intracellular C-terminus of 23 amino acids is predicted.

3.3.4 Extracellular loop 1 (ECL1)

3.3.4.1 Prosite motif PDOC01045

As discussed in section 1.7.2, TMEM114 shows homology to the Pfam00822 family of transmembrane proteins. This family of small (~18-25 kDa) integral membrane proteins encompasses the claudins, MP20, PMP22 and epithelial membrane proteins-1, -2 and -3 (EMP-1, -2, -3). These proteins contain four transmembrane domains with two extracellular loops, the first of which is larger than the second. The first extracellular loop (ECL1) contains a conserved W-GLW-C-C motif (Prosite motif PDOC01045) (Figure 3.3). This transmembrane topology and the W-GLW-C-C motif are also features of the voltage dependent calcium channel γ subunits, a more distally related group of proteins. ECL1 is functionally important for both claudins and the γ subunits (Tomita *et al.* 2004; Wen *et al.* 2004). To begin to elucidate the function of TMEM114 and TMLP1, both proteins were compared to Pfam00822 family members and γ subunits to identify shared properties and domains.

To determine if TMLP1 also contains a W-GLW-C-C motif in ECL1, an alignment of ECL1 from Pfam00822 family members and the voltage dependent calcium channel γ subunits was performed. ECL1 in TMEM114 and TMLP1 are larger than those in claudins but of similar size to that of the γ subunits (Figure 3.3). TMEM114 and TMLP1 also contain the PDOC01045 motif in ECL1. The complete W-GLW-C-C motif is not present in EMP2, EMP3, MP20, TMEM47, PERP, claudin-10b and claudin-15 as these proteins lack the conserved glycine (G) residue. In the Prosite annotation of the PDOC01045 motif, the final cysteine (C) is immediately preceded by either S, T, D, E, N, Q, or H, but alternative residues are found in TMEM114 and TMLP1. The spacing between the two cysteines is generally less than nine amino acids in the majority of family members, but is greater than twenty amino acids in γ1, γ6 and claudin-like protein 24 (CLP24). The majority of the claudins and the γ subunits contain a glycine residue (G) four residues before the final cysteine, but this glycine is not present in TMEM114, or the EMPs, CLP24, TMEM47 and PERP. The majority of family members also contain an arginine-alanine (RA) sequence 3-4 amino acids from TMD2 which

is present in TMLP1 but not in TMEM114.

3.3.4.2 Glycosylation

Glycosylation is one of the most abundant protein modifications in eukaryotes (Turnbull and Field 2007). Glycosylation is the addition of oligosaccharides to specific amino acids and is a common feature in secreted proteins and membrane proteins with extracellular regions (Cao *et al.* 2009). Addition of oligosaccharides to asparagine residues, known as N-linked glycosylation, occurs in a co-translational manner in the ER. N-glycosylation may promote protein folding by interacting with chaperones (Tokhtaeva *et al.* 2010) and may affect sorting to specific membrane domains (Vagin *et al.* 2009). O-linked glycosylation is the post-translation addition of oligosaccharides to serine or threonine residues and may also have a role in protein trafficking (Yeaman *et al.* 1997). All eight γ subunits contain predicted N-glycosylation sites in ECL1 as do EMP-1,-2,-3, PMP22 and MP20 (Price *et al.* 2005). Only two claudins contain predicted N-glycosylation sites in ECL1 (CLDN-1 and -12) (Price *et al*. 2005).

The NetNGlyc 1.0 Server (Gupta *et al*. 2004b) was used to identify putative N-glycosylation sites in the ECLs of TMEM114 and TMLP1. TMEM114 contains 2 predicted sites in ECL1 at residues p.N55 and p.N89 (Figure 3.3). These predicted glycosylation sites are conserved across species (Figure 3.4A). TMLP1 contains a single predicted site at residue p.N41 (Figure 3.3). However, both murine and rat Tmlp1 lack the conserved asparagine-linked site (Figure 3.4B). No O-linked glycosylation sites were predicted for TMEM114 or TMLP1 by NetOGlyc 3.1 (Julenius *et al.* 2005).

3.3.5 Extracellular loop 2 (ECL2)

In TMEM114, TMLP1 and other Pfam00822 family members the second extracellular loop (ECL2) is shorter than ECL1. No functional motifs have been identified to date in Pfam00822 family members or in the calcium channel γ subunits. To identify any common residues or motifs, ECL2 of TMEM114, TMLP1, Pfam00822 family members and the voltage dependent calcium channel γ subunits were aligned (Figure 3.5). The regions of ECL2 adjoining the transmembrane domains in TMEM114 and TMLP1 were similar, beginning with the residues YIxY (where x is any amino acid) and ending with the GWS sequence of amino acids. These motifs were similar to those of γ2, γ3, γ4, γ5, γ7, γ8 and claudin-16. Almost all members of the Pfam00822 family had a tyrosine (Y) or phenylalanine (F) close to the end of ECL2. The majority of claudins had a FY or FF motif in the centre of ECL2 and some claudins had conserved proline (P), lysine (K) and glutamic acid (E) amino acids. ECL2 of the EMPs, PMP22, MP20, CLP24, PERP and TMEM47 were shorter and lacked the conserved motifs present in the other family members.

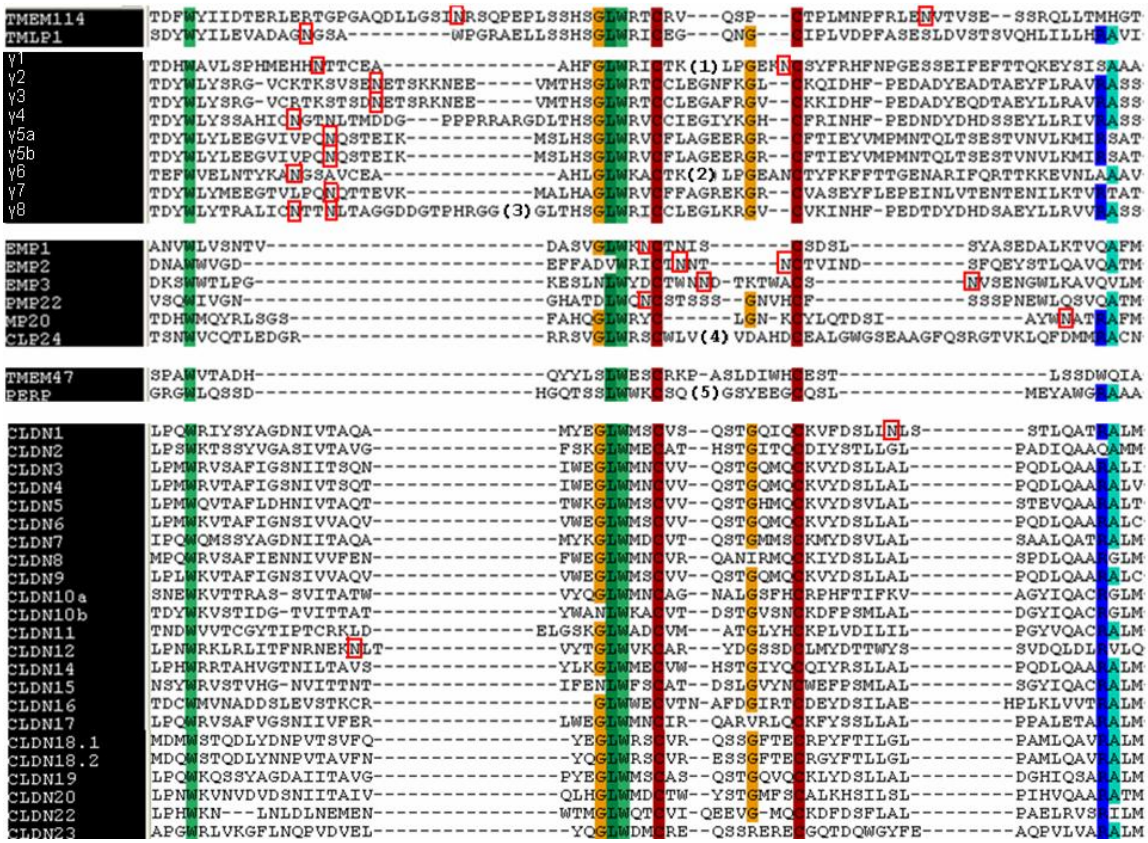


Figure 3.3. Alignment of the first extracellular loop (ECL1) of TMEM114, TMLP1 and Pfam00822 proteins. Identical residues present in more than 70 % of sequences are highlighted. Predicted N-linked glycosylation sites (NetNGlyc) are boxed in red. To facilitate alignment, gaps were introduced (-) and some amino acids were removed: (1) = RIPMDDSKTCGPIT. (2) = RLWQADVPVDRDTCGPAE. (3) = GGASEKKDPG. (4) = DRTRGGPSPGARAGQ. (5) = EGGGS. ECL1 sequences were based on TMPRED toplogy predictions.

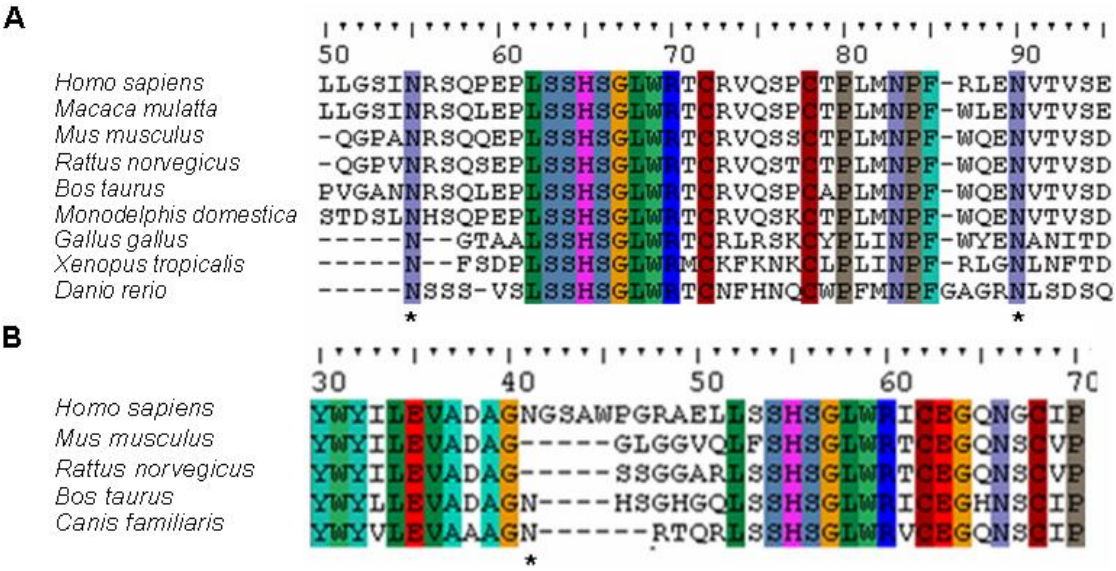


Figure 3.4. Conservation of N-glycosylation sites in TMEM114 and TMLP1. A) Two predicted N-glycosylation sites (*) are present in the 1st extracellular loop of TMEM114 at residues 55 and 89. B) One predicted N-glycosylation site (*) is present in the 1st extracellular loop of TMLP1 at residue 41. This site is not present in mouse and rat. Numbers refer to human amino acid sequence.

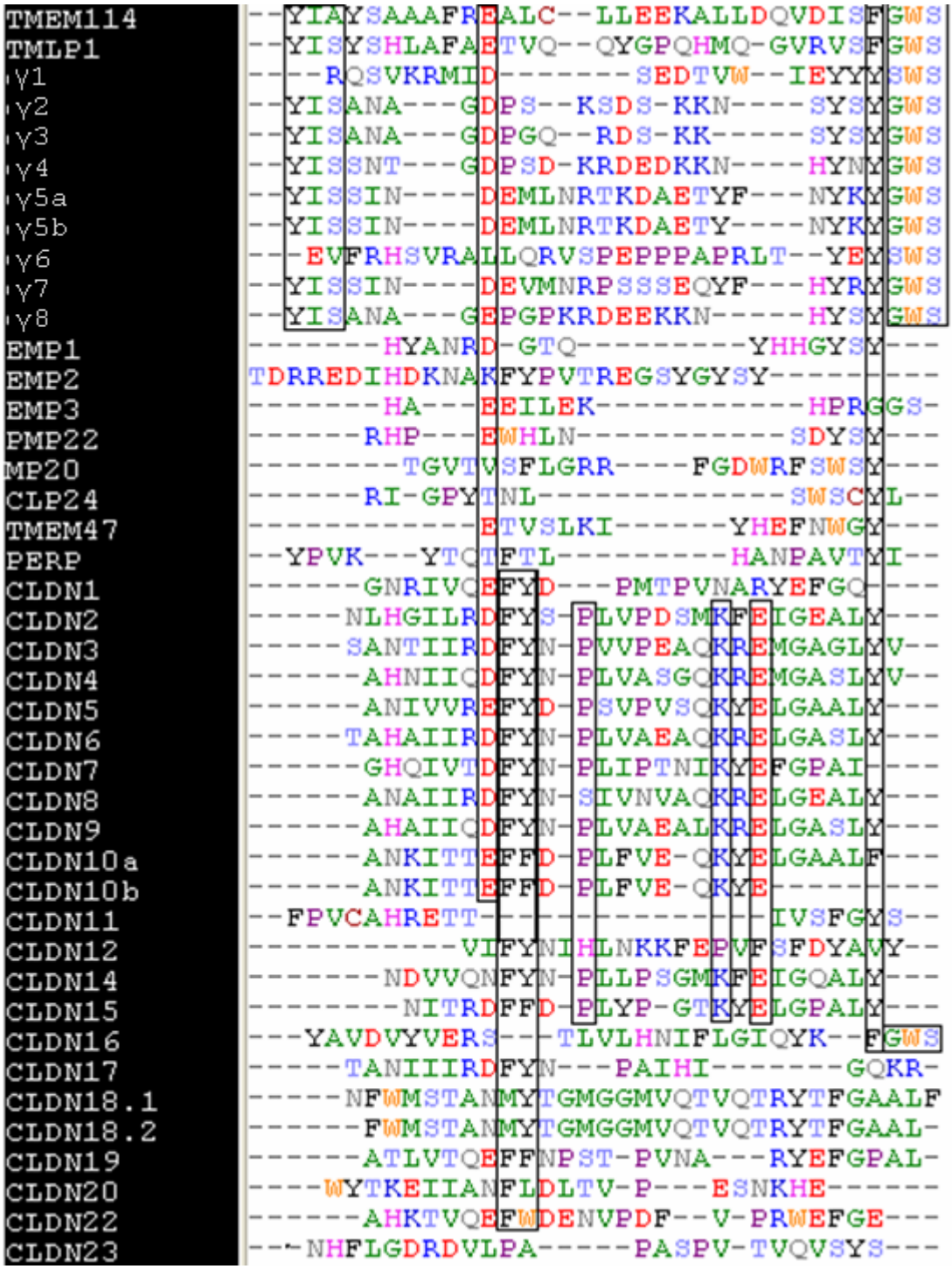


Figure 3.5. Alignment of the second extracellular loops of TMEM114, TMLP1, Pfam00822 family members and the voltage dependent calcium channel γ subunits. Conserved amino acids are boxed. Amino acids with similar properties are grouped into colour-coded groups. Dark green = A/G/I/M/L/V; Red = D/E; Blue = K/R; Light blue = S/T; Black = F/Y; Grey = N/Q; Wine = C; Purple = P; Pink = H; Orange = W

3.3.6 Carboxy termini

To determine if the C-termini of TMEM114 or TMLP1 contain conserved motifs the C-termini of TMEM114 and TMLP1 orthologues were aligned (Figure 3.6). The C-terminus of TMEM114 is conserved in mammals, but shows little conservation in chicken (*Gallus gallus*), African clawed frog (*Xenopus tropicalis*) or zebrafish (*Danio rerio*) (Figure 3.6A) (Jamieson *et al*. 2007). This C-terminus is rich in arginine, serine and leucine residues and terminates with two hydrophobic residues (alanine and isoleucine). The C-terminus of TMLP1 in primates is distinct from that of non-primates (Figure 3.6B). The final 3 amino acids in primates are encoded by a fifth exon not present in non-primates (Section 3.8). In non-primate mammals TMLP1 terminates in three hydrophobic residues, but in primates the C-terminus is highly polar.

The short intracellular C-terminus of claudins is required for correct localisation and the majority of claudins end in tyrosine-valine (YV) PDZ binding motifs which interact with adaptor proteins (Figure 3.7) (Itoh *et al.* 1999; Jeansonne *et al.* 2003). The C-termini of the γ subunits are highly variable in length and four of the subunits (γ2, γ3, γ4, γ8) also contain PDZ binding motifs (TTPV) for protein interaction (Milstein and Nicoll 2009). The C-termini of TMEM114 and primate and non-primate TMLP1 do not contain PDZ-binding motifs and are much shorter than the C-termini of claudins and the majority of the γ subunits (Figure 3.7). Although similar in size to some of the other family members, the C-termini TMEM114 and TMLP1 show little similarity to those proteins in terms of amino acid sequence.

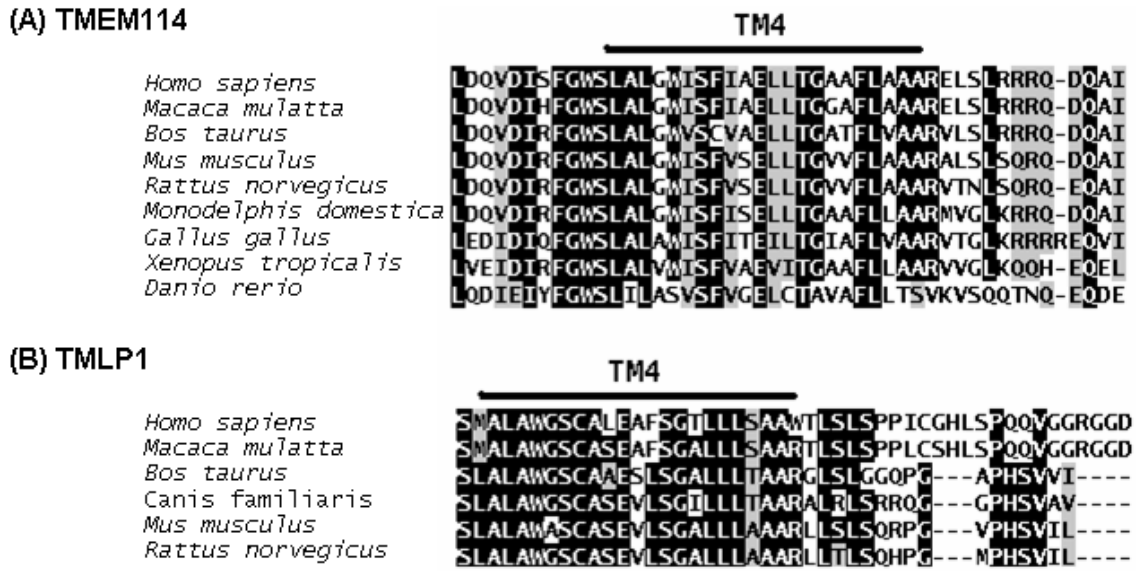


Figure 3.6. Alignment of the carboxy termini from TMEM114 and TMLP1 orthologues. Carboxy termini and the 4th transmembrane domain (TMD4) as predicted by Philius. Identical residues are highlighted in black. Grey = similar residues. Threshold levels for similar residues were 80 % for TMEM114 and 66 % for TMLP1.

Figure 3.7. Alignment of the carboxy termini of Pfam00822 family members. Depending on prediction tools used, TMLP1 has a relatively short (14 residue) or long (24 residue) (brackets) C-terminus. PDZ binding motifs (blue box) are present in the voltage dependent calcium channel γ2, γ3, γ4 and γ8 subunits and in the majority of claudins. * = end of protein sequence. To aid alignment the following amino acids were deleted from the γ8 subunit: (1) = LAGAGGGGGGAVGAFGGAAGGAGGGGGGGGGAGAE,
(2) = TVTVTGPPAPPAPAPPAPSAPAPGTLAKEAAASNT

3.4 Post-translational modifications

3.4.1 Phosphorylation

The function of proteins may be regulated by the addition and removal of phosphates. To assess whether TMEM114 and TMLP1 are potentially regulated by phosphorylation, phosphorylation sites were predicted using NetPhos2.0 (Blom *et al.* 1999) and the kinases

predicted to act at the identified sites were determined using NetPhosK 1.0 (Blom *et al.* 2004). A number of phosphorylation sites were predicted for TMEM114, all of which were in non-transmembrane regions (Figure 3.8). One serine site was conserved down to *Xenopus tropicalis,* but the confidence level for this prediction was moderate. A number of serine sites and a single threonine site were conserved in human and mouse. Predictions for TMLP1 phosphorylation sites show little conservation (Figure 3.9).

A

```
hTMEM114        MRVHLGGLAGAAALTGALSFVLLAAAIGTDFWYIIDTERLERTGPGAQDLLGSINRSQPE 60
mTmem114        MRVRLGALAGAAALSGALSFVLLAAAIGTDFWYIIDTERLERSSQRMRDQ-GPANRSQQE 59
xTmem114        MKLKLSMLSVFVAVVGILSFISLVVAIGTDFWYIIDASRLEKITN-----------FSD 48
                *:::*. *:  .*: * ***: *..***********:.***:               :

hTMEM114        PLSSHSGLWRTCRVQSPCTPLMNPFRLENVTVSESSRQLLTMHGTFVILLPLSLILMVFG 120
mTmem114        PLSSHSGLWRTCRVQSSCTPLMNPFWQENVTVSDSSRQLLTMHGTFVILLPLSLIVMVFG 119
xTmem114        PLSSHSGLWRMCKFKNKCLPLINPFRLGNLNFTDSQKQLLSMHGTLVILLPLSLILMIFG 108
                ********** *:.:. * **:***   *:..::*.:***:****:*********:*:**

hTMEM114        GMTGFLSFLLQAYLLLLLTGILFLFGAMVTLAGISVYIAYSAAAFREALCLLEEKALLDQ 180
mTmem114        GMTGFLSFLLRAHLLLLLTGILFLFGAMVTLTGISIYIAYSAVAFREAVCLLEERALLDQ 179
xTmem114        GMTGFVSILARAYLLLLLTGMLFLFGALVTLTGISIYIAYSAAAFKDAVCILGN-KILED 167
                *****:*:* :*:*******:******:***:***:******.**::*:*:* :  :*::

hTMEM114        VDISFGWSLALGWISFIAELLTGAAFLAAARELSLRRRQDQAI 223
mTmem114        VDIRFGWSLALGWISFVSELLTGVVFLAAARALSLSQRQDQAI 222
xTmem114        IDIQFGWSLALAWISFITEILTGIAFLVAARVTGLKRRRREQV 210
                :** *******.****::*:*** .**.***  .* :*: : :
```

B

Homo sapiens			*Mus musculus*			*Xenopus tropicalis*		
AA	**Con**	**Kinase**	**AA**	**Con**	**Kinase**	**AA**	**Con**	**Kinase**
T37	0.97	CKII	T37	0.97	CKII			
			S44	1.00	DNAPK			
S57	0.93	DNAPK, ATM, cdc2	S56	0.73	DNAPK, cdc2			
			S62	0.85		S51	0.93	
S76	0.98	P38mapk, PKA, cdk5	S75	0.72	PKA			
S95	0.71		S94	0.83		S83	0.85	DNAPK
S184	0.77							
S214	0.93	PKC	S213	0.64	RSK, PKC, PKA			
			S215	0.91	DNAPK, ATM, PKC			

Figure 3.8. Predicted phosphorylation sites and the associated kinases for TMEM114. A) Alignment of TMEM114 from *Homo sapiens* (h), *Mus musculus* (m) and *Xenopus tropicalis* (x) with phosphorylation sites predicted by NetPhos 2.0. Blue = Serine site. Yellow = Threonine site. Bold font = predicted transmembrane domain. B) Amino acids (AA) predicted to be phosphorylated with the confidence (Con) level of the prediction according to NetPhos 2.0. Predicted kinases which act at the site, predicted using NetPhosK 1.0 are also included. Abbreviations: CKII = casein kinase II; DNAPK = DNA-dependent protein *kinase; ATM* = Ataxia telangiectasia mutated; cdc2 = cell division control protein 2 (or Cyclin dependent kinase 1); P38mapk = p38 mitogen-activated protein kinase; PKA = Protein kinase A, cdk5 = Cyclin-Dependent Kinase5; PKC = Protein kinase C*;* RSK = ribosomal s6 kinase

A

```
hTMLP1          MARLGALLLAAALGALLSFALLAAAVASDYWYILEVADAGNGSAWPGRAELLSSHSGLWR 60
mTmlp1          MALLATLLLSAALGALLSFALLAAAVASDYWYILEVADAGG-----LGGVQLFSHSGLWR 55
cTmlp1          MARLGALLLAAALGALLSFALLVAAVASDYWYVLEVAAAGN------RTQRLSSHSGLWR 54
                ** *.:***:************.*********:**** **.          * *******

hTMLP1          ICEGQNGCIPLVDPFASESLDVSTSVQHLILLHRAVIVVLPLSLVLLVCGWICGLLSSLA 120
mTmlp1          TCEGQNSCVPLIDPFASAGLEVSPSVQHLLSLHRTVMVVLPLSLVLIVCGWVCGLLSSLS 115
cTmlp1          VCEGQNSCIPLIDPFASENLDVPTSVRHLISLHRAIMVVLPLSLVLVVCGWICGLFSSLA 114
                 *****.*:**:***** .*:*..**:**: ***:::*********:****:***:***:

hTMLP1          QSVSLLLFTGCYFLLGSVLTLAGVSIYISYSHLAFAETVQQYGPQHMQGVRVSFGWSMAL 180
mTmlp1          QSVPLLLATGCYFLLGGALTLAGLSIYISYSHLAFVEAARTYGVTHVQNVHISFGWSLAL 175
cTmlp1          QSVLLLLFTGCYFLLGGALTLVGISVYISYSHLAFAETARQFGPRHVQHVRVGFGWSLAL 174
                *** *** ********..***.*:*:*********.*:.: :*  *:* *:::****:**

hTMLP1          AWGSCALEAFSGTLLLSAAWTLSLSPPICGHLSPQQVGGRGGD 223
mTmlp1          AWASCASEVLSGALLLAAARLLSLSQRPGV---PHSVIL---- 211
cTmlp1          AWGSCASEVLSGILLLTAARALRLSRRQGG---PHSVAV---- 210
                **.*** *.:** ***:**  * **        *:.*
```

B

Homo sapiens			*Mus musculus*			*Canis familiaris*		
AA	**Con**	**Kinase**	**AA**	**Con**	**Kinase**	**AA**	**Con**	**Kinase**
S53	0.59					S47	0.99	
						S48	0.99	RSK, PKA
S83	0.72							
						S79	0.82	
			S86	0.80	PKC			
						S143	0.77	
						T152	0.77	
			T156	0.57	PKC			
S173	0.72	PKC, PKA	S198	0.59	RSK, PKC, PKA			
			S200	0.98	DNAPK, ATM	S199	1.00	PKC

Figure 3.9. Predicted phosphorylation sites and the associated kinases for TMLP1. A) Alignment of Tmlp1 from *Homo sapiens* (h), *Mus musculus* (m) and *Canis familiaris* (c) with phosphorylation sites predicted by NetPhos 2.0. Blue = Serine site. Yellow = Threonine site. Bold font = predicted transmembrane domain. B) Amino acids (AA) predicted to be phosphorylated with the confidence (Con) level of the prediction according to NetPhos 2.0. Predicted kinases which act at the site, predicted using NetPhosK 1.0 are also included. Abbreviations: RSK = ribosomal s6 kinase; PKA = Protein kinase A; PKC = Protein kinase C; DNAPK = DNA-dependent protein kinase; ATM = Ataxia telangiectasia mutated.

3.4.2 Palmitoylation

Palmitoylation of membrane proteins increases the hydrophobicity of transmembrane regions to promote membrane association. Palmitoylation involves the addition of the lipid palmitate to cysteine residues. Palmitoylation sites were predicted using CSS-Palm 2.0 (Ren *et al.* 2008). TMLP1 contains three predicted palmitoylation sites which were conserved across species. Scores were highest for the two sites present in TMD2, p.C109 and p.C113 (Figure 3.10). A weak, but significant score was obtained for p.C131. TMEM114 contains a single predicted site at p.C170 in ECL2 but the score was not significant (data not shown).

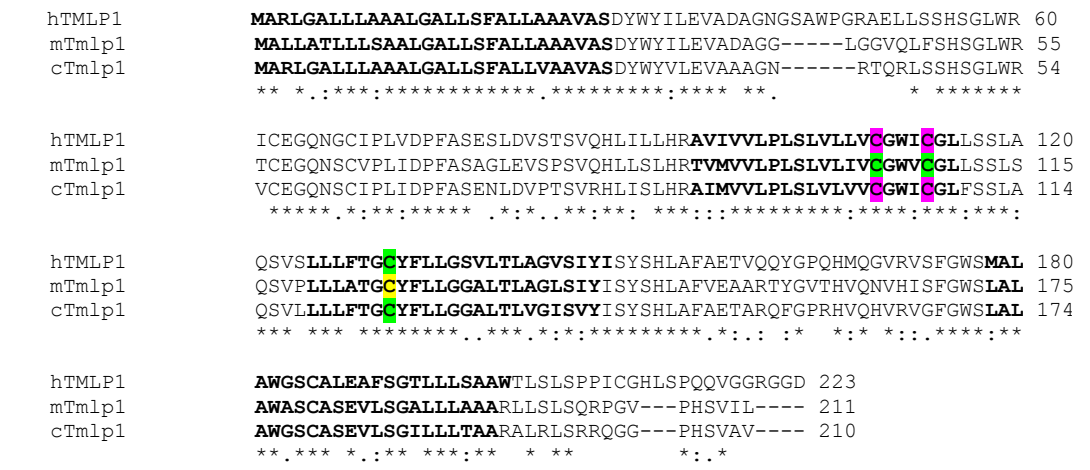


Figure 3.10. Predicted palmitoylation sites for TMLP1. TMLP1 contains three predicted palmitoylation sites which are conserved in *Homo sapiens* (h), *Mus musculus* (m) and *Canis familiaris* (c). Two sites with the highest prediction scores are present in TMD2. Colours refer to prediction scores: Yellow = <1, Green = 1-1.49, Pink 1.5-2. Sites were predicted using CSS-Palm 2.0.

3.5 Phylogenetic Analysis of TMEM114 and TMLP1 with Pfam00822

Both TMEM114 and TMLP1 have similar features to members of the Pfam00822 family of claudins, EMPs and voltage dependent calcium channel γ subunits as they have a predicted tetraspan topology and contain the W-GLW-C-C motif in ECL1. To determine the relationship of TMEM114 and TMLP1 to members of this family a phylogenetic tree was constructed by a neighbour-joining method based on ClustalW multiple alignments using the software Seaview (Gouy *et al.* 2010). Bootstrap values are an indicator of how robust branches of a tree are (scored as percentages with 100 indicating maximum robustness). Only branches with bootstrap values of >50 were considered significant. The claudins form distinct branches from the γ subunits and the EMP/PMP/MP20 family members (Figure 3.11), as previously described (Price *et al*. 2005). TMEM114 and TMLP1 show the greatest similarity to each other and they cluster with the γ subunits on distinct branches from γ1 and γ6 and from the other γ subunits (Figure 3.11). Members of the Pfam00822 family of proteins vary greatly in their intracellular domains, particularly the C-termini (Figure 3.5). Therefore, a phylogenetic tree was also constructed using the most conserved regions of the proteins in this family, i.e. from the beginning of TMD1 to the end of TMD 4. A similar tree was constructed with these sequences, with TMEM114 and TMLP1 branching with the γ subunits but distinct from γ1 and γ6 and the other γ subunits (Figure 3.12).

Performing a pairwise alignment with of TMEM114 and TMLP1 with each other and other members of the Pfam00822 family of proteins confirmed that TMEM114 and TMLP1 show the greatest similarity to each other (Table 3.2). TMLP1 shows greatest amino acid identity with the γ6 subunit when full-length sequences and TMD1-TMD4 sequences are compared. The TMD1-TMD4 region of TMEM114 shows greatest similarity to the γ2, γ3, γ4 and γ8 subunits (Table 3.2).

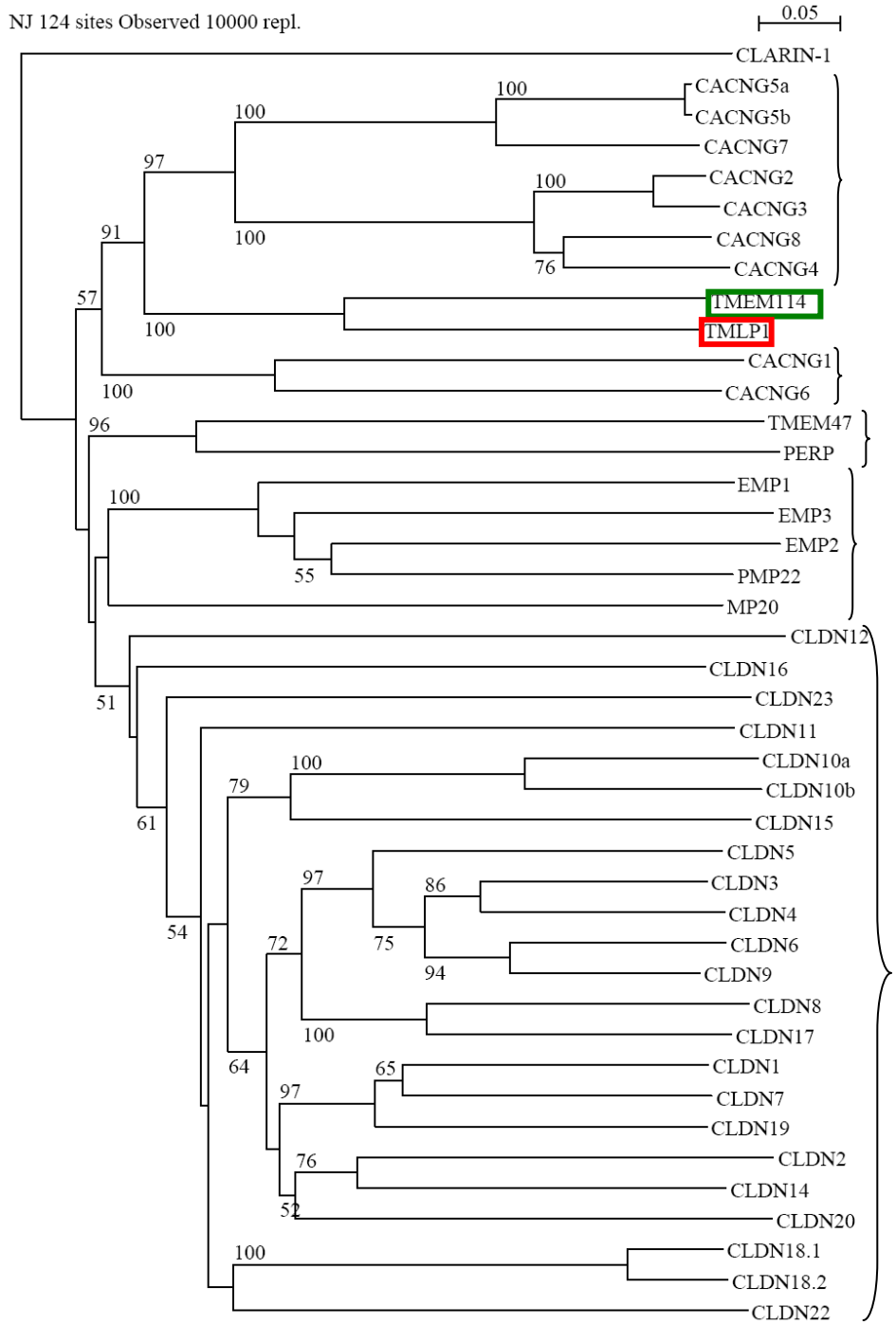


Figure 3.11. Phylogenetic tree of human TMEM114, TMLP1 and Pfam00822 proteins. A phylogram of full-length protein sequences was constructed using ClustalW and the neighbour-joining method. The protein Clarin-1 was used an outgroup. TMEM114 (green box) and TMLP1 (red box) cluster with the voltage dependent calcium channel γ subunits (CACNGs) on distinct branches from the claudins and EMP/PMP/MP20 family members. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position. Numbers indicate bootstrap values as the percentage of 10,000 replicates for each branch. Bootstrap values of less than 50 were considered non-significant and are not displayed.

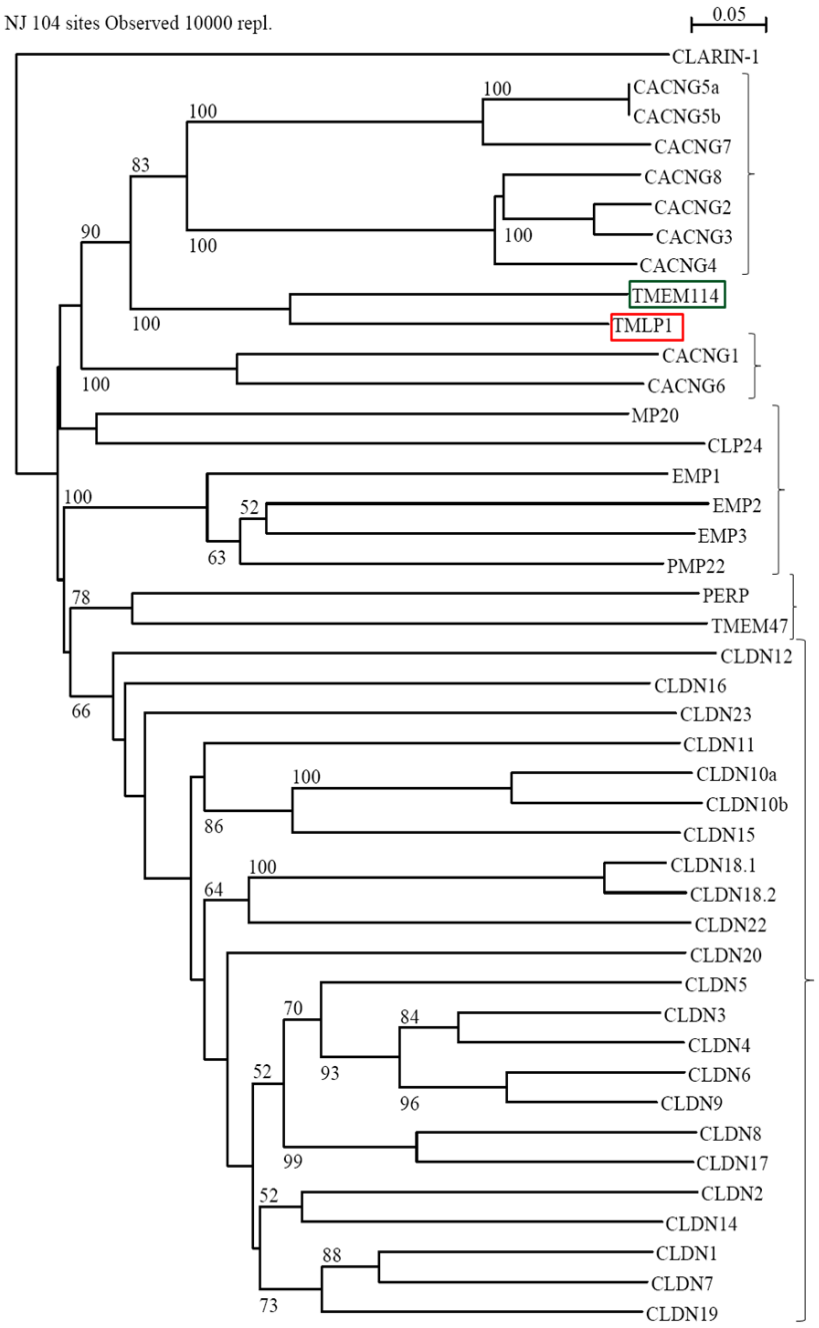


Figure 3.12. Phylogenetic tree of TMD1-TMD4 of human TMEM114, TMLP1 and Pfam00822 proteins. A phylogram of an amino acid alignment of the 1st transmembrane domain through to the 4th transmembrane domain was constructed using ClustalW and the neighbour-joining method. The protein Clarin-1 was used an outgroup. TMEM114 (green box) and TMLP1 (red box) cluster with the calcium channel γ subunits (CACNGs) on distinct branches from the claudins and EMP/PMP/MP20 family members. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position. Numbers indicate bootstrap values as the percentage of 10,000 replicates for each branch. Bootstrap values of less than 50 were considered non-significant and are not displayed.

Table 3.2. Pairwise alignment of TMEM114 and TMLP1 with voltage dependent calcium channel γ subunits and EMPs, MP20, PMP22 and CLP24. Percentage pairwise amino acid identities were determined using the ALIGN program (Genestream).

	TMLP1	MP20	EMP-1	EMP-2	EMP-3	PMP22	CLP24	PERP	TMEM47	
TMEM114	42.7	19.8	21.1	23.2	22.0	23.2	20.7	20.3	19.7	**full-length**
	42.7	20.9	21.2	21.2	21.4	23.2	22.6	19.9	19.4	**TM1-TM4**

	γ1	γ2	γ3	γ4	γ5	γ6	γ7	γ8	
TMEM114	20.6	18.1	19.3	21.1	18.1	18.9	16.7	17.6	**full-length**
	22.3	25.4	25.6	28.9	23.2	19.4	19.5	26.6	**TM1-TM4**

	TMEM114	MP20	EMP-1	EMP-2	EMP-3	PMP22	CLP24	PERP	TMEM47	
TMLP1	42.7	22.2	19.3	22.0	23.7	21.4	20.6	21.9	21.3	**full-length**
	42.7	22.1	20.4	23.3	24.2	21.3	21.0	21.1	22.9	**TM1-TM4**

	γ1	γ2	γ3	γ4	γ5	γ6	γ7	γ8	
TMLP1	22.9	19.8	22.0	21.0	23.5	25.0	21.0	17.4	**full-length**
	22.2	23.9	25.4	25.7	26.3	28.7	23.3	23.1	**TM1-TM4**

3.6 Identifying orthologues of TMLP1

In order to identify orthologues of TMLP1 a Blastp search was performed using the predicted protein sequence of TMLP1. A number of TMEM114 orthologues were identified, confirming the similarity of TMLP1 to TMEM114. Other hypothetical proteins were also identified in animals as evolutionary distant as zebrafish. In zebrafish two distinct protein sequences with moderate levels of similarity to TMLP1 were identified (XP_001919554; XP_001340845) both annotated as TMEM114-like proteins (Figure 3.13).

Figure 3.13. Alignment of putative zebrafish Tmlp1 orthologues XP_001919554 and XP_001340845 with human and murine Tmlp1. Black highlight = identical residue in three or more proteins.

To determine if either of the zebrafish TMLP1-like proteins is an orthologue of TMLP1, the genes surrounding *TMLP1* were compared to those surrounding both zebrafish genes. In humans, *TMLP1* is located on chromosome 17q25.3 and is flanked by *threonine aldolase pseudogene 1* (*THA1P*), *suppressor of cytokine signalling 3* (*SOCS3*), *cytohesin 1*, *ubiquitin specific peptidase 36* (*USP36*) and *tissue inhibitor of metalloproteinase 2* (*TIMP2*) on the telomeric side. To the centromic side *TMLP1* is flanked by a *baculoviral IAP repeat* (*BIRC5*), *arylformamidase* (*AFMID*), *thymidine kinase* (*TK1*) and *synaptogyrin* (*SYNGR2*) (Figure

3.15). The zebrafish gene encoding the protein XP_001919554 (XM_001919519 / LOC100150722) is located on chromosome 12 immediately next to *threonine aldolase 1* and *Socs3b*. On the other flanking side of LOC100150722 is a *baculoviral IAP repeat*, *cytohesin 1-like*, and *Timp2*. The zebrafish gene encoding the protein XP_001340845 (XM_001340809 / LOC100000694), located on chromosome 3 is in proximity to *Socs3a*, *cytohesin 1*, *Usp36* and *Timp2b*. On the other flanking side is a *Timp2b*-like gene and *galectin-3*.

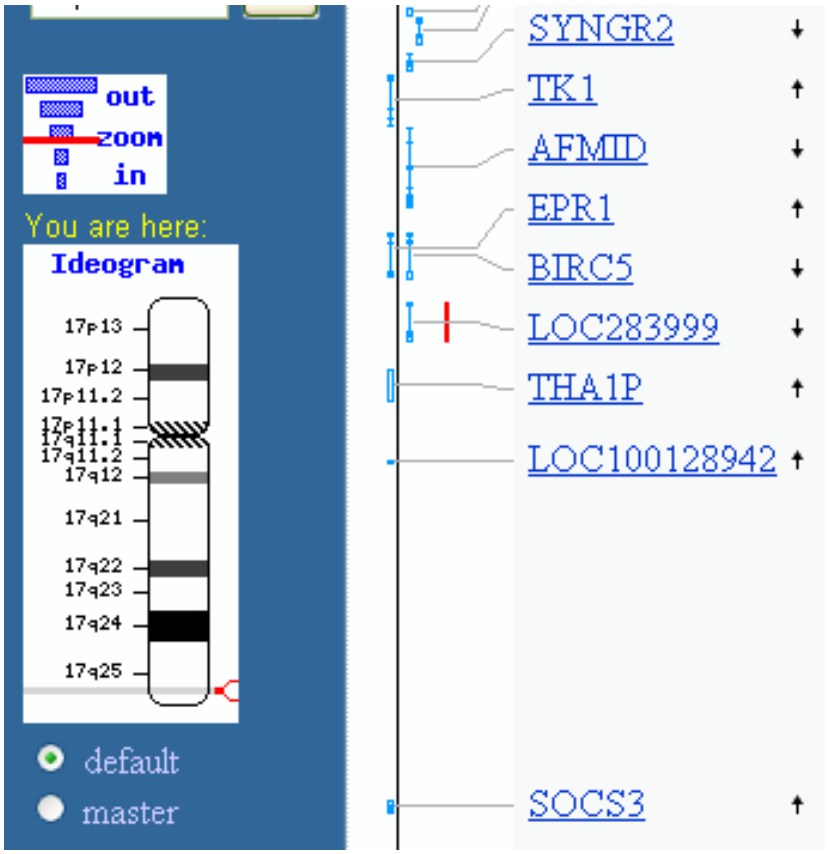


Figure 3.14. *TMLP1* and its surrounding genes on human chromosome 17q25.3. Human *TMLP1* (LOC283999) is flanked by *baculoviral IAP repeat* (*BIRC5*), *arylformamidase (AFMID)*, *thymidine kinase (TK1)* and *synaptogyrin (SYNGR2)* on the centromeric side and *threonine aldolase pseudogene 1 (THA1P)* and more distantly, suppressor of cytokine signalling 3 (*SOCS3)*, on the telomeric side. Arrows represent direction of transcription. Image taken from NCBI Mapviewer human (*Homo sapiens*) Build 37.1.

Both putative *Tmlp1* orthologues in zebrafish are in close proximity to *Socs3* genes. The genomes of zebrafish and other fish contain two copies of *Socs3,* but a single copy of *Socs3* is present in the genome of animals that evolved after fish (Wang *et al*. 2010b). Therefore, if one can determine if the *SOCS3* present in the human genome is descended from either *Socs3a* or *Socs3b*, then one may infer that human *TMLP1* is descended from the putative *Tmlp1* close to the ancestor of *SOCS3*. To identify which zebrafish copy of *Socs3* mammals show most similarity to a multiple alignment was created using protein sequences (Figure 3.15). Both Socs3a and Socs3b show significant similarity to murine and human Socs3. It is therefore not obvious which zebrafish gene the mammalian genes are descended from.

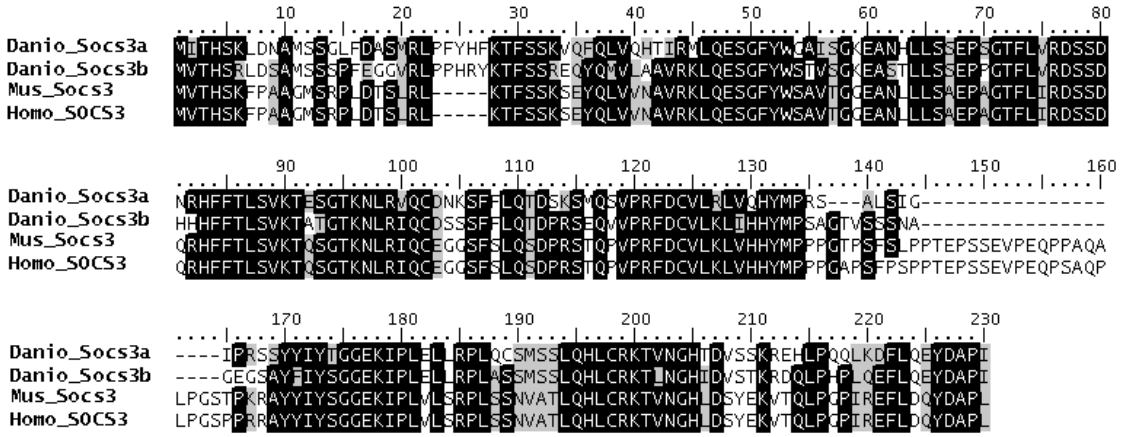


Figure 3.15. Alignment of zebrafish suppressor of cytokine signalling 3 genes Socs3a and Socs3b with murine and human Socs3. Sequences were aligned with Clustalw. Numbering refers to human SOCS3 amino acids. Black highlight = identical residues

In the next attempt to elucidate the conservation of Tmlp1, the aim was to identify Tmlp1 in species intermediate to mammals and fish; chicken (*Gallus gallus*) and *X. tropicalis*. Using Blastp, tBlastn and Blastn searches against nucleotide and EST databases, no sequences of reasonable similarity to TMLP1 were identified. Performing the same searches with both zebrafish sequences also failed to identify similar proteins. The regions syntenic to 17q25.3 in humans were then identified in the chicken genome. In chicken, *threonine aldolase pseudogene 1* and *arylformamidase* are present at the same locus but Tmlp1 is not present between these genes on MapViewer (Figure 3.16). When the sequence between the two genes was downloaded it revealed that there was a gap in the sequence in contig NW_001471506.1. Similarly, there are also gaps in the sequence in this region in the opossum *Monodelphis domestica* (NW_001581875.1).

Creating a phylogenetic tree with the TMEM114 orthologues, the TMLP1 orthologues and putative zebrafish orthologues and human and zebrafish Pfam00822 family members revealed that the putative zebrafish orthologues of Tmlp1 are closer in relationship to TMLP1 than to other zebrafish proteins (Figure 3.17).

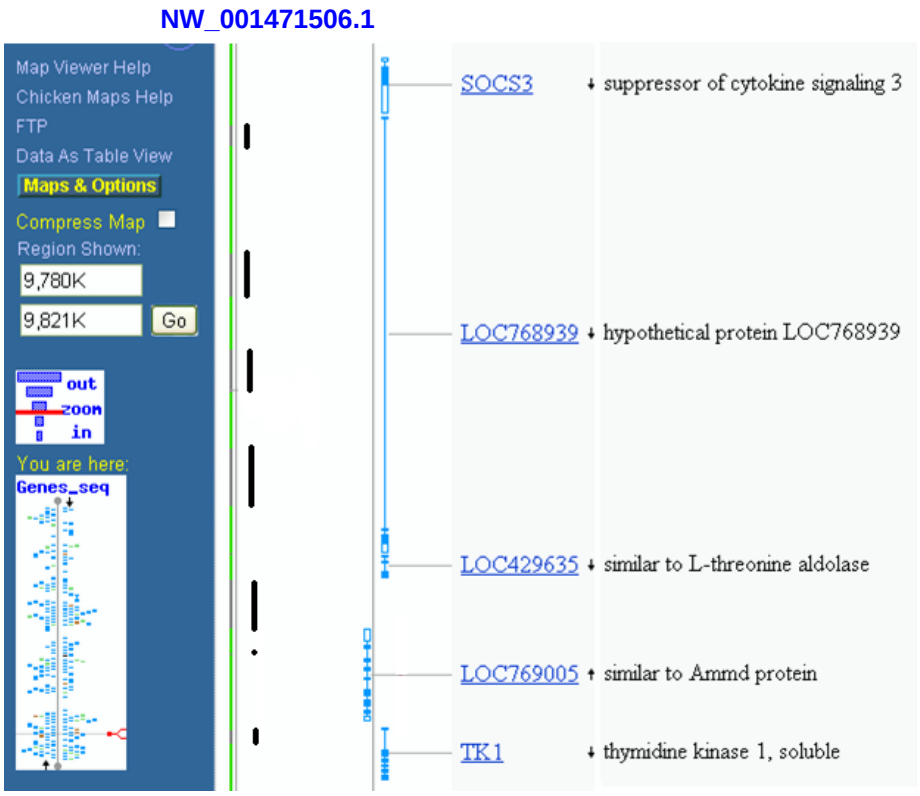


Figure 3.16. Chicken contig NW_001471506.1 on chromosome 18. No gene is annotated between *arylformamidase* (*Ammd*) and *L-threonine aldolase*. Sequence gaps (black bars) in the contig (green bar), show that the sequence between the two genes has not been identified. Arrows represent direction of transcription. Image from NCBI Mapviewer chicken (*Gallus gallus*) Build 2.1.

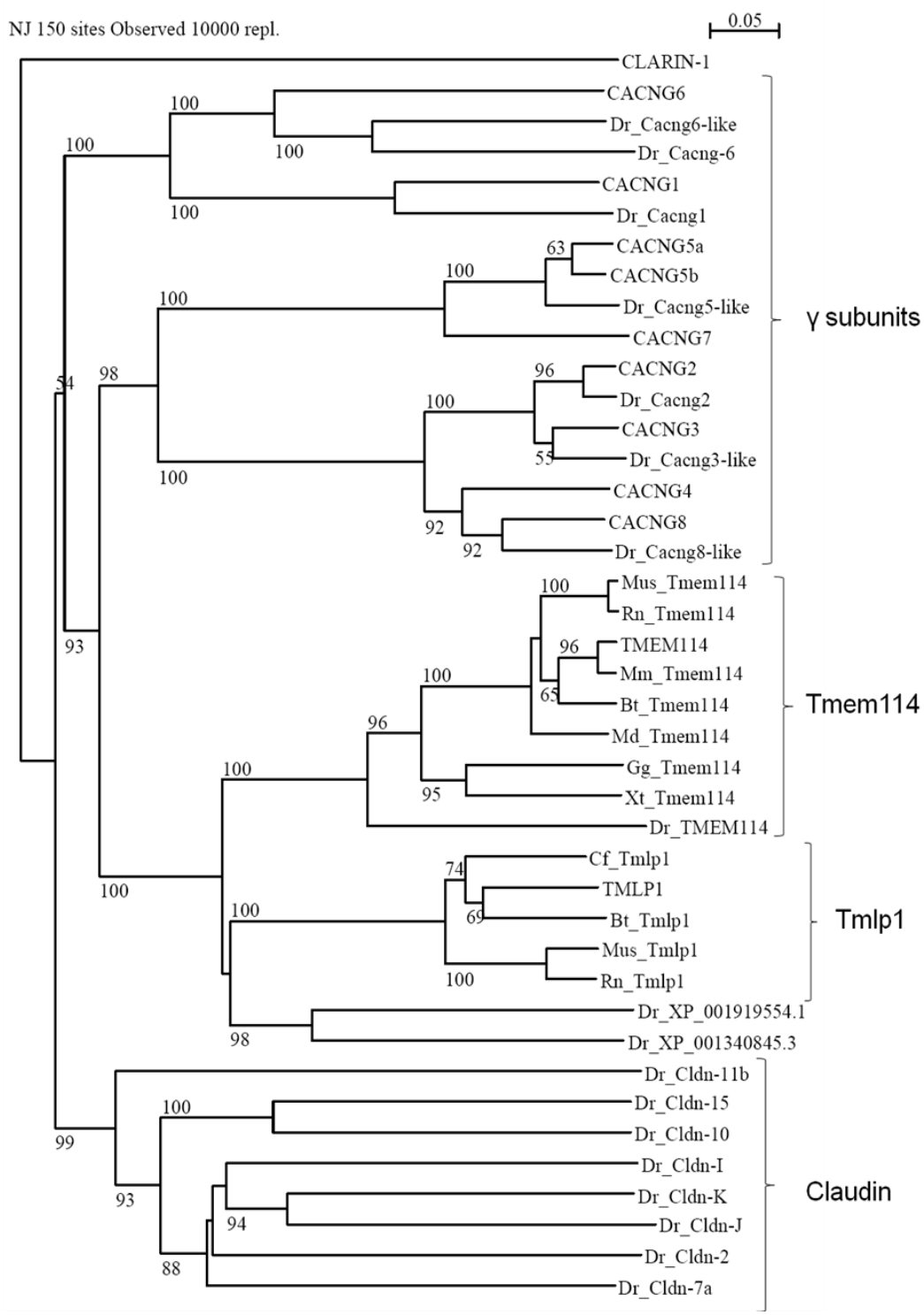


Figure 3.17. Phylogenetic tree of human and zebrafish orthologues and putative orthologues of TMEM114, TMLP1 and Pfam00822 proteins. A phylogram of full-length protein sequences aligned using ClustalW was constructed using the neighbour-joining method. Cldn = claudin; CACNG = voltage dependent calcium channel gamma subunit; Dr = *Danio rerio* (Zebrafish); Mm = *Macaca mulatta*; Mus = *Mus musculus*; Rn = *Rattus norvegicus*; Bt = *Bos Taurus*; Cf = *Canis familiaris*; Md = *Monodelphis domestica*; Gg = *Gallus gallus*; Xt = *Xenopus tropicalis*. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position. Numbers indicate bootstrap values as the percentage of 10,000 replicates for each branch. Bootstrap values of less than 50 were considered non-significant and are not displayed.

3.7 Expression of *TMEM114* and *TMLP1*

ESTs encoding the third and fourth exons of *TMEM114* belong to the Unigene cluster Hs.150849. ESTs are present in the foetus and more prominently in the adult. Expression is restricted to three tissues: prostate (5 transcripts per million (TPM)), lens (47 TPM) and lung (94 TPM) (NCBI aceview LOC283953). All lung samples are from carcinoids (rare slow-growing neuroendocrine tumours) (Dahabreh *et al.* 2009). Single ESTs for the murine orthologue were cloned from the retina, brain, testis, oocyte and stomach (Unigene Mm.159318). All murine ESTs were from adult tissue.

ESTs from TMLP1 have been detected in adult brain (hippocampus and hypothalamus) (11 TPM) and in embryonic stem cells (Unigene Hs.632228). Available cDNA clones from brain tissue reveal that it is alternatively spliced (Figure 3.18). When exon 2 is spliced out, 28 amino acids of ECL1 are omitted. When exons 2 and 3 are spliced out, 73 amino acids of ECL1, TMD2 and TMD3 are omitted. This results in a shift of the TMDs producing a three-TMD protein with an extracellular C-terminus (Figure 3.18).

To identify the 5'UTR of *TMLP1* the EST BC042066.1 was compared to the genomic sequence. The predicted 5'UTR was 624 nucleotides in length and omitted 90 nucleotides from the genomic sequence. This short sequence begins and ends in intron splice donor and acceptor sites (GTXXX and XXXAG, respectively) suggesting that transcription occurs from another exon upstream (Exon 1a) of what was previously considered to be exon1 (Figure 3.19).

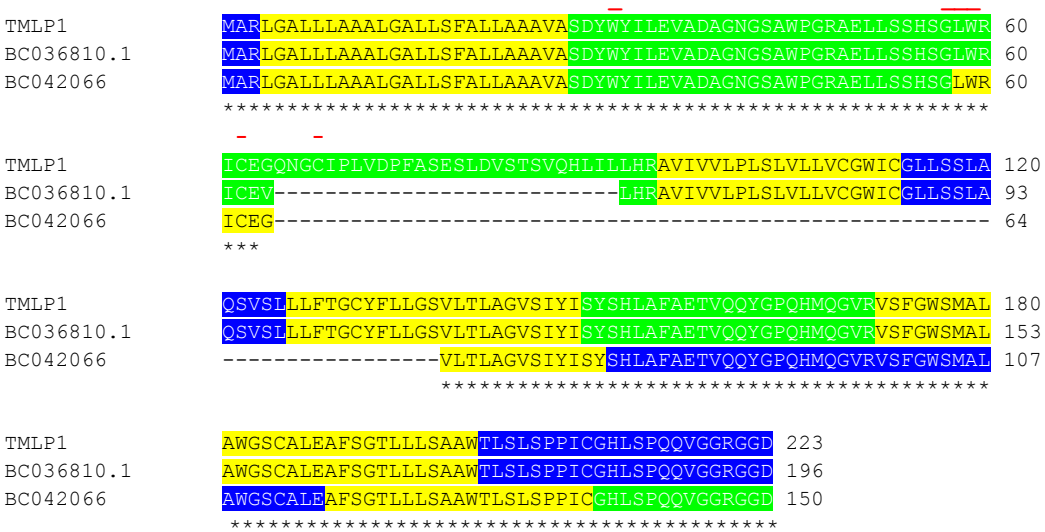


Figure 3.18. Amino acid sequences encoded by alternatively spliced cDNAs from *TMLP1*. The protein encoded by cDNA BC036810, which lacks exon 2, maintains the same topology as full-length TMLP1, but the altered splice site results in the incorporation of a valine (V) in place of glycine (G) at p.64. cDNA BC042066.1 lacks both exons 2 and 3 and encodes a protein with three predicted transmembrane domains and an extracellular C-terminus. Yellow = TMD, Green = extracellular domain, Blue = intracellular domain. The W-GLW-C-C motif is marked by the red bar above the sequence. Topology predicted by TMPRED.

```
Genomic      GTCCTAGAAGGTTCTTTCTGCCTGTTTTCTCCgtgagtgactggacaggcagaggccggccttgctggagggg
BC042066.1   GTCCTAGAAGGTTCTTTCTGCCTGTTTTCTCC----------------------------------------

Genomic      gcatttgtaattatgagtgaatccaaaacaaggttttttccttccgcagCCCCCCGCCCGGCTGTGGGGCCCA
BC042066.1   ----------------------------------------------CCCCCCGCCCGGCTGTGGGGCCCA

Genomic      GCCACTGCACTTCACCGGATGCCGTCTGGTTGGTCCTCAGGACTGATACAGACCAGGACCCCAGGGCCAGCCC
BC042066.1   GCCACTGCACTTCACCGGATGCCGTCTGGTTGGTCCTCAGGACTGATACAGACCAGGACCCCAGGGCCAGCCC

Genomic      GTGCCAGGCTCCTATGCTTCCAGGAGCACGGGTGGGTGGTCCTGCTGCCTGGCCGGCCATCCTCCTGGGGTCG
BC042066.1   GTGCCAGGCTCCTATGCTTCCAGGAGCACGGGTGGGTGGTCCTGCTGCCTGGCCGGCCATCCTCCTGGGGTCG

Genomic      GTCTCTGGCCGATCCTCCCTCCTCCTCTCAAGCCCTGCACAGCCCGGCCAGGCAGGTGCATCTTGTTTGGCTG
BC042066.1   GTCTCTGGCCGATCCTCCCTCCTCCTCTCAAGCCCTGCACAGCCCGGCCAGGCAGGTGCATCTTGTTTGGCTG

Genomic      CTGAGGAGCCGGGGGTTCAGGGAAATTAAGGAACGTGCCCAGGGACCCGGGGCCAGCCCGTGGGGACGCTGGG
BC042066.1   CTGAGGAGCCGGGGGTTCAGGGAAATTAAGGAACGTGCCCAGGGACCCGGGGCCAGCCCGTGGGGACGCTGGG

Genomic      ATTGGAGCCCAAGCCCCAGGTTCGCCGCGCGGCTCTCGACTTCCTCTCCTTTCCCCCAGGGGCGAGCTCAGCG
BC042066.1   ATTGGAGCCCAAGCCCCAGGTTCGCCGCGCGGCTCTCGACTTCCTCTCCTTTCCCCCAGGGGCGAGCTCAGCG

Genomic      ACCGCAGAGAGGTGGGGTCGATCTCCCTGCGACCCCAGGGGGCCCGCGAGGCCAGTGCGCGGGCAGGAGCGGG
BC042066.1   ACCGCAGAGAGGTGGGGTCGATCTCCCTGCGACCCCAGGGGGCCCGCGAGGCCAGTGCGCGGGCAGGAGCGGG

Genomic      GACGTGCTCAGAAGAGCCGGGCGCCGCCGCGCCCGCCCGCCCCCCGTCCCCCGGCTCCCGGCTCCGCGCGCCC
BC042066.1   GACGTGCTCAGAAGAGCCGGGCGCCGCCGCGCCCGCCCGCCCCCCGTCCCCCGGCTCCCGGCTCCGCGCGCCC

Genomic      CCCGCCGCCCCCGGGGCCCTGCTACCCCCGACCCGTCCCCTCCCGCCGGCCGCCCCCATGGCCCGGCTGGGCG
BC042066.1   CCCGCCGCCCCCGGGGCCCTGCTACCCCCGACCCGTCCCCTCCCGCCGGCCGCCCCCATGGCCCGGCTGGGCG
```

Figure 3.19. Identification of the 5'UTR of *TMLP1*. Comparison of EST BC042066.1 to the genomic sequence revealed a gap in the alignment of 90 nucleotides (lower case). This gap begins with a GT and ends in AG suggesting it is an intron. The translation start site (ATG) is underlined and the coding sequence is in bold.

3.8 Gene structure of *TMEM114* and *TMLP1*

To look further at the conservation of TMEM114 and TMLP1, the structure of both genes was compared in a number of species. When the structure of *TMEM114* was being investigated a Blastn search with the available sequence for the sequence upstream of *TMEM114* revealed a gap in the sequence assembly in the human genome project (Figure 3.19A). To identify the size and sequence of the gap between the two adjoining contigs, primers were designed to cross the gap. PCR amplification with the primers resulted in a 3,360 bp amplicon, representing a gap of 1,577 bp (Figure 3.19B). The sequence of this 3.4 kb amplicon was then determined by primer walking and submitted to the GenBank database (accession number GQ267544.1) (Sequence in Section 9.6).

The exon structure of *TMEM114* is largely conserved (Figure 3.21). Exons 2, 3 and 4 do not differ in length in different species. However, these exons have different levels of conservation of encoded amino acids. The amino acids encoded by exon 3 which form TMD2 and half of TMD3 are highly conserved. Exon 4 shows much higher conservation in its encoded transmembrane domains regions than in ECL2 or the intracellular C-terminus. Exon 1 shows relatively low levels of conservation and differs in length in different species. Two regions containing the residues W-GLW-C of the W-GLW-C-C motif are conserved. The residues in TMD1 are conserved in mammals (human and mouse) but are less so in chicken and *Xenopus*. Exon 2 is also highly similar in humans and mouse, but is less conserved in chicken and *Xenopus*.

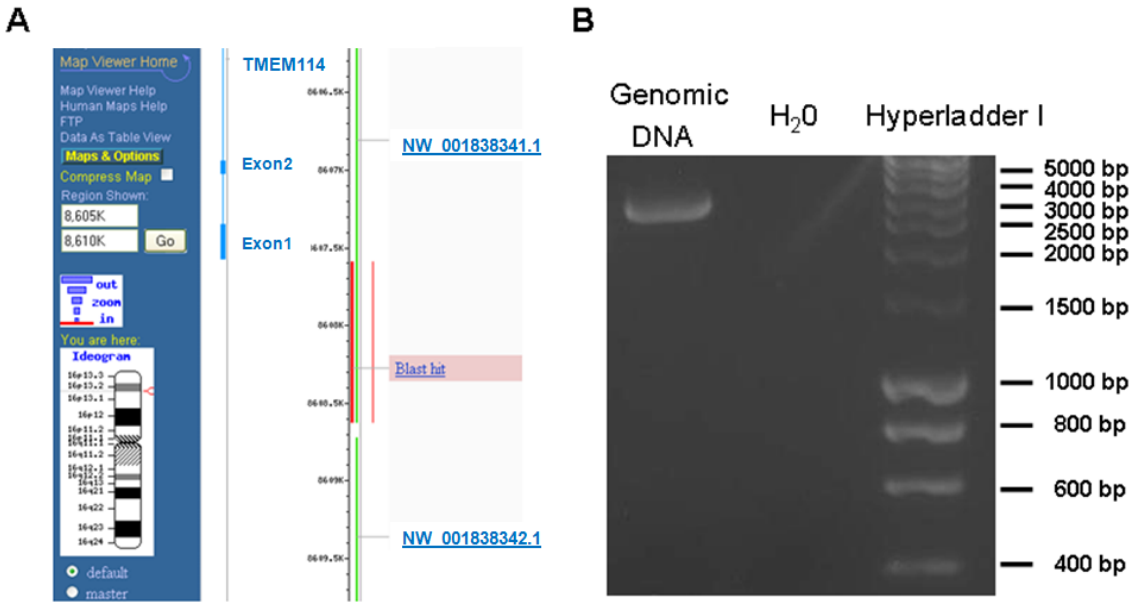


Figure 3.20. Identification of the gap between contigs NW_001838341 and NW_001838342. A) A Blastn search with the available DNA sequence upstream of exon 1 of TMEM114 (red bar) aligned to the contig NW_001838341 (green bar). A gap in sequence between NW_001838341 and NW_001838342 (green bar) was identified. Image taken from NCBI Mapviewer human (*Homo sapiens*) Build 37.1. B) Amplification across the gap region between the two contigs produced a 3.360 bp amplicon, representing a gap of 1,577 bp.

Human *TMLP1* contains six exons and encodes TMLP1 from the latter 5 exons, in comparison to being encoded by four exons in non-primates. The final exon in TMLP1 encodes just three amino acids (Figure 3.22). Exon 3 is highly conserved and encodes parts of TMD2 and TMD3. The regions of exon 4 that encode TMD3 and TMD4 are also conserved. Exon 2 shows some degree of conservation. Exon 1 is highly conserved except for a short region in ECL1 that is only present in humans.

Comparison of the structure of *TMEM114* and *TMLP1* revealed that exons 2 and 3 are the same size in both genes, but the intron between them is vastly different in size (Table 3.3). The intron/exon boundaries surrounding exon 2 appear similar between both genes, but other intron/exon boundaries show little similarity. Exon 3 of *TMEM114* and *TMLP1* is the same size as exon 3 in γ1 and γ6 and is similar to that of the other γ subunits (Table 3.4). Exon 2 of *TMEM114* and *TMLP1* is similar in size to the γ subunits, differing by 1-2 codons. The structure of claudins varies with some encoded by single exons (e.g. *CLDN2*) and others by four (e.g. *CLDN1*) or five (e.g. *CLDN16*) exons. The sizes of the exons of the latter are distinct from those of the γ subunits. The *EMPs* and *PMP22*, although similar in size in terms of exon 3, show little similarity with the γ subunits or *TMEM114* and *TMLP1* in terms of other exons. The structures of *CLP-24*, *TMEM47* and *PERP* do not match those of the γ subunits or *TMEM114* and *TMLP1*.

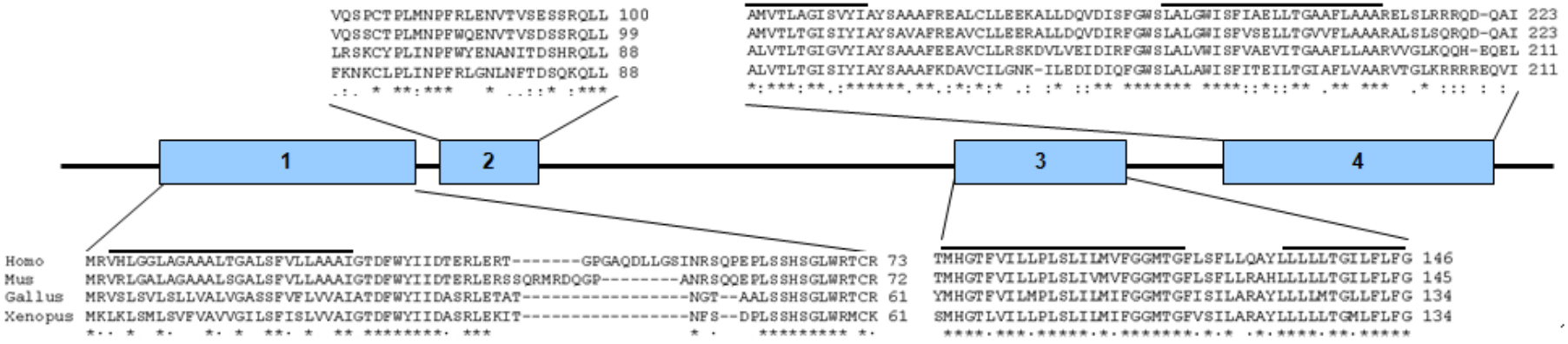


Figure 3.21. Gene structure of *TMEM114*. *TMEM114* encodes a 223 amino acids protein from four exons. Encoded amino acids are shown above or below the exons. Amino acid sequences from human (Homo), mouse (Mus), chicken (Gallus) and Xenopus were aligned with Clustalw. Transmembrane regions are marked by a black bar.

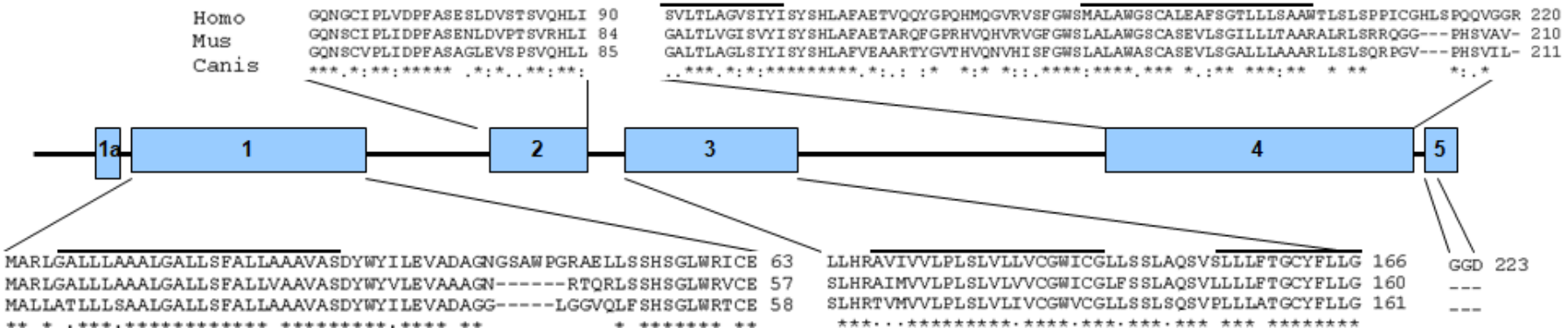


Figure 3.22. Exon structure of *TMLP1*. Human TMLP1 encodes a 223 amino acids protein from 5 exons. Murine (Mus) and canine (Canis) encode 210 and 211 amino acids proteins, respectively, from 4 exons. Encoded amino acids are shown above or below the exons. Amino acid sequences were aligned with Clustalw. Transmembrane regions are marked by a black bar.

Table 3.3. Intron/exon boundaries of TMEM114 and TMLP1. Intronic splice donor (GT) and splice acceptor (AG) sites are in bold. Exon size refers to nucleotides encoding amino acids. Stop codons are coloured in red.

Exon	Gene	5' Sequence	Exon sequence	3' sequence	Exon size	Intron size
1a	*TMLP1*		GTCCTAGAAG.....TGTTTTCTCC	**gt**aaccggcc.....	-	90bp
1	*TMEM114*	.....gggctcggcg	ATGCGGGTGC.....ACCTGCCGGG	**gt**aagggtgc.....	220bp	325bp
1	*TMLP1*	.....ggccgccccc	ATGGCCCGGC.....ATCTGCGAAG	**gt**aaccggcc.....	190bp	1,681bp
2	*TMEM114*	.....tcctatcc**ag**	TGCAGAGCCC.....CAACTTCTCA	**gt**gagtccgg.....	81bp	16,973bp
2	*TMLP1*	.....tgtttccc**ag**	GGCAGAACGG.....CACCTCATCT	**gt**gagtcctg.....	81bp	617bp
3	*TMEM114*	.....ctcttcta**ag**	CCATGCATGG.....CTCTTTGGAG	**gt**aggtgccc.....	138bp	2,081bp
3	*TMLP1*	.....tttcctgc**ag**	TGCTGCACCG.....CTGCTGGGG**A**	**gt**gagtctgg.....	138bp	4,293bp
4	*TMEM114*	.....ctccctgc**ag**	CCATGGTGAC.....GGCCATATGA		233bp	
4	*TMLP1*	.....tctccccc**ag**	GTGTCCTGAC.....TGGGAGGG**AG**	**gt**aagagggg.....	250bp	506bp
5	*TMLP1*	.....cattttac**ag**	AGGGGGAGACTGA		13bp	

Table 3.4. Gene structure of TMEM114 and TMLP1 compared to members of Pfam00822. *: exon size refers to nucleotides encoding amino acids. - : exon does not encode amino acids

	TMEM114	TMLP1	γ1	γ6	γ5	γ7	γ2	γ3	γ4	γ8
Exon 1*	220	190	289	331	196	196	211	211	220	181
Intron 1	325	1,681	9,106	5,030	1,443	1,472	114,815	89,768	53,057	14,720
Exon 2	81	81	75	75	87	87	84	84	84	84
Intron 2	16,973	617	1,032	1,320	1,497	778	20,971	80,145	6,587	1,637
Exon 3	138	138	138	138	141	141	141	141	141	141
Intron 3	2,081	4,293	804	12,179	3,818	25,964	1,466	6,378	5,465	2,072
Exon 4*	233	250	227	239	146	146	536	512	539	770
Intron 4		506			321	420				
Exon 5*		13			258	258				

	MP20	EMP1	EMP2	EMP3	PMP22	CLP24	PERP	TMEM47	CLDN1	CLDN2	CLDN16
Exon 1*	301	78	78	78	78	280	214	253	233	-	324
Intron 1	4,575	1,890	3,879	598	1,456	7,365	10,632	30,694	8,947	7,499	13,892
Exon 2	150	97	91	103	100	156	141	180	165	693	103
Intron 2	1,778	105	5,495	1,728	19,482	12,705	4,085	13,380	2,618		2,322
Exon 3	135	141	147	141	141	245	227	266	85		165
Intron 3	242	712	4,835	808	8,390				1,729		3,387
Exon 4*	62	158	188	170	164				163		192
Intron 4											1,397
Exon 5*											134

3.9 Clarification of the database entries for TMLP1

Note: The 223 amino sequence used for TMLP1 was based on the GenBank accession XP_946151.1 which was identified through the blast search with the TMEM114 protein sequence. However, since this sequence was first identified, the GenBank annotation has changed. The accession XP_946151.1 has been replaced by NP_001139001.1 which describes a 394 amino acid protein. This 394 amino acid protein does not contain the W-GLW-C-C motif and is predicted to be a globular protein and not a transmembrane protein (Philius). This predicted protein sequence is a translation of the predicted mRNA sequence NM_001145529.1 which is based in part on the EST BC042066.1 (Figure 3.18, 3.19). Translation of this EST in the +3 frame produces the 394 amino acid protein NP_001139001. Translation of this EST is the +1 frame produces the splice variant of TMLP1 encoded by exons 1, 4 and 5, based on the exonic structure devised in this chapter (Figure 3.22).

Thus, both proteins are predicted to be encoded from the same transcript. The inclusion of exon 3, as found in the EST BC036810 (Figure 3.18), does not disrupt the correct frame of either of the predicted proteins. However, inclusion of exon 2 which encodes the latter part of ECL1 in TMLP1, would result in a premature stop codon being encoded in NP_001139001. NP_001139001 shows no similarity to any protein in the human genome and no functional motifs were identified in the predicted protein. Thus it is much more likely that the protein encoded by locus 283999 is TMLP1 and not NP_001139001. This will be confirmed if transcripts including exon 2 can be detected.

On the Ensembl database, both the predicted proteins have been annotated from Loc283999, The 196 amino acid isoform of TMLP1 is annotated ENSP00000364084, and the 394 amino acid protein encoded from the same locus is annotated ENSP00000402790. However, this conflicts with the annotation on the GenBank database as ENSP00000364084 refers to NP_001139001. To clarify, for future reference, the Ensembl annotations will be used (i.e. ENSP00000364084 = TMLP1 and ENSP00000402790 = the 394 amino acid protein NP_001139001) (Table 3.5).

Table 3.5. Accession numbers for predicted proteins encoded by loc283999. The Ensembl accession numbers and names are used in this study.

Name	Protein type	AA	Frame	Genbank Accession	Ensembl Accession
TMLP1	Transmembrane	196	+1	-	ENSP00000364084
?	Globular	394	+3	NP_001139001/ ENSP00000364084	ENSP00000402790

3.10 Discussion

3.10.1 Identification of a novel gene TMLP1

Through a Blastp search a novel hypothetical protein similar to TMEM114 was identified. This protein, named TMEM114-like protein 1 (TMLP1), shares 43 % amino acid identity with TMEM114 (Figure 3.1, Table 3.2). This level of amino acid identity is above the "twilight zone" of 20-30 % identity at which level it is difficult to assess homology (Doolittle 1987; Pevsner 2009). This suggested that the proteins may be related. Both proteins are the same size (223 amino acids), have the same predicted topology and both contain a conserved motif in ECL1.

In order to learn more about the conservation of *TMLP1* and to possibly identify a shared ancestor of *TMEM114* and *TMLP1,* orthologues of *TMLP1* were searched for. Orthologues were identified in placental mammals, but not in marsupials, chicken or *X. tropicalis*. However, the contig which contains the homologous chromosomal region in chicken has gaps in the sequence where one may have expected to find Tmlp1 (Figure 3.16). Two potential orthologues were identified in zebrafish. Both genes, although more dissimilar than to be expected from true orthologues, were in close proximity to genes that human TMLP1 is linked with. Animals that evolved after fish contain only one *Socs3* gene, but zebrafish and other fish species have *Socs3a* and *Socs3b* genes (Wang *et al*. 2010b). The putative zebrafish *Tmlp1* orthologue XP_00130845 is linked with *Socs3a*, *cytohesin 1*, *Usp36* and *Timp2* on chromosome 3, whilst XP_001919554 is linked with *threonine aldolase* and *Socs3b* on chromosome 12. Although both genes are in the correct loci to be orthologues of *TMLP1*, and show greater similarity to *TMLP1* than to other zebrafish genes (Figure 3.18), their encoded proteins show one distinct difference from TMLP1. Both proteins lack the two cysteine residues in the conserved motif W-GLW-C-C (Figure 3.13). The full motif is present in all zebrafish claudins, γ subunits and zebrafish Tmem114. Thus these similar proteins may not be true functional orthologues of *TMLP1*.

If *Tmlp1* is not present in zebrafish, the question of when it arose and what its closest ancestor is remains unanswered. This information is difficult to ascertain due to gaps in the available genomes for intermediate species. In mammals, the transmembrane regions of TMLP1 are largely conserved, as is ECL1. The intracellular C-terminus is encoded by two exons in primates, but only one in non-primates and shows little conservation between the two groups of species. This may represent the adoption of a new functional domain.

3.10.2 TMEM114, TMLP1 and the Pfam00822 family

3.10.2.1 Topology of TMEM114 and TMLP1

The predicted topology of TMEM114 and TMLP1 is suggestive of a relationship with the Pfam00822 family of proteins. Both proteins have four transmembrane domains, short intracellular N- and C-termini, two extracellular loops, the first of which being larger than the second and containing a conserved motif, and a very short intracellular loop between TMD2

and TMD3. As described in section 1.7.2, the Pfam00822 family of proteins contains claudins, the main constituent of tight junctions and the more distally related EMPs, PMP22, MP20 and CLP24. Although they show little similarity at the sequence level, this family also contains the voltage dependent calcium channel γ subunits as these proteins share the same topology and conserved motif.

3.10.2.2 Voltage dependent calcium channel γ subunits

There are eight voltage dependent calcium channel γ subunits in the human genome, but despite their name, the main function of the majority of this group of proteins does not involve calcium channel activity. Phylogenetically, and functionally, the γ subunits are divided into three clusters: (γ1, γ6), (γ5, γ7) and (γ2, γ3, γ4, γ8) (Burgess *et al.* 2001; Chu *et al.* 2001; Price *et al.* 2005). The first subunit described, γ1, is a functional calcium channel subunit expressed in skeletal muscle which decreases expression and activity of the pore forming alpha subunit ($Ca_v1.1$) (Sandoval *et al.* 2007). This is mediated by a direct interaction between the two subunits (Arikkath *et al.* 2003). The γ6 subunit is expressed in skeletal and cardiac muscle, with limited neural expression (Burgess *et al.* 2001; Chu *et al.* 2001; Fukaya *et al.* 2005). This subunit interacts with and negatively regulates cardiac $Ca_v3.1$ (Hansen *et al.* 2004; Lin *et al.* 2008).

The stargazer mouse, which displays absence epilepsy and cerebellar ataxia, was found to have a disruption in the gene expressing stargazin (Letts *et al.* 1998). This gene shared exonic and protein structure with the γ1 subunit (Letts *et al*. 1998). As the phenotype of the stargazer mouse also showed similarities to mice with mutations in other voltage dependent calcium channel subunits (Khosravani and Zamponi 2006), the gene encoding stargazin was annotated the γ2 subunit. Although no change was observed in calcium current in neurons of stargazer mice (Chen *et al.* 2000) there was a notable absence of AMPA receptors (AMPARs) at the synapses of cerebellar mossy fibre-granule cells (Hashimoto *et al.* 1999). Transfection of γ2 into the stargazin knockout rescued AMPAR activity by mediating trafficking of AMPARs to the cell surface (Chen *et al*. 2000). The ability of the subsequently identified γ3, γ4 and γ8 to traffic AMPARs (Glutamate receptors 1-4) to the cell membrane gave rise to these subunits being named as TARPs (transmembrane AMPA receptor regulating proteins) (Tomita *et al.* 2003). As well as modulating AMPAR activity by trafficking, TARPs also regulate the gating of AMPARs (Milstein *et al.* 2007; Sager *et al.* 2009).

There is little experimental evidence on the ability of γ5 and γ7 subunits to alter calcium current, although γ7 can regulate stability of *$Ca_v2.2$* mRNA (Ferron *et al.* 2008). Even though both proteins can alter AMPAR functionality, the function of γ5 and γ7 is distinct from the TARPs, and they have been recently annotated as Type II TARPS (Kato *et al.* 2008). Trafficking of AMPARs by γ7 is specific to GluR1 and GluR2 (Kato *et al.* 2007), and although it does not regulate AMPAR trafficking, γ5 can regulate GluR2 gating, a by regulating specific GluR2 transcripts (Kato *et al*. 2008).

3.10.2.3 Relationship of TMEM114 and TMLP1 to Pfam00822 family members

A phylogenetic tree was constructed to identify the relationship of TMEM114 and TMLP1 to other Pfam00822 family members. The claudins form a distinct clade from the calcium channel γ subunits and the EMPs, PMP22, MP20 and CLP24. The claudin family separated into the classic highly conserved claudins (1-10, 14, 15, 17, 19) and the non-classic non-conserved claudins (11-13, 16, 18, 20-23) (Krause *et al.* 2009).

The γ subunits clustered as expected; γ1 and γ6, γ5 and γ7, and γ2, γ3, γ4 and γ8 (Figure 3.11). TMEM114 and TMLP1 cluster together, in the same clade as the calcium channel γ subunits and the EMPs, PMP22, MP20 and CLP24. The bootstrap values for the division of the calcium channel γ subunits and the EMPs, PMP22, MP20 and CLP24 were low which prevents certainty in the distinction of the relationship between the γ subunits and the EMPs/PMP22/MP20. However, it is clear from the phylogenetic tree that TMEM114 and TMLP1 are closer in relationship to the calcium channel γ subunits and the EMPs, PMP22, MP20, CLP24, PERP and TMEM47 than to any claudin family members. .

The C-termini of proteins in the Pfam00822 superfamily are variable and are much longer in γ subunits than the other proteins (Figure 3.7) (Price *et al*. 2005). To remove this potential confounding factor, a phylogenetic tree based on the amino acids from TMD1-TMD4 was constructed as recommended by Price *et al.*, 2005. The tree produced was similar with TMEM114 and TMLP1 clustering with the γ subunits. However the bootstrap values were slightly higher in the TMD1-TMD4 tree. TMLP1 and γ6 had 25 % amino acid identity over the full-length of the proteins, and 29 % from TMD1-TMD4, which was more than with any other Pfam00822 family member (Table 3.2). TMEM114 showed greatest amino acid identity with γ4 (29 %). This level of amino acid identity was lower than that for TMLP1 and γ6 which explains why TMEM114 and TMLP1 branched separately from γ4 and the other TARPs.

Pairwise amino acid identity between γ1 and γ6 is 33 %, between γ5 and γ7 is 47 %, and between each of γ2, γ3, γ4 and γ8 are greater than 50 % (Burgess *et al*. 2001). Thus, the 43 % identity between TMEM114 and TMLP1 is likely to be suggestive of a relationship between the two genes. Pairwise identity between non-TARP and TARP γ subunits was typically <25 % (Burgess *et al*. 2001), similar to the level of identity between TMEM114 and the other γ subunits suggesting a possible relationship.

3.10.2.4 Gene structure

The gene structure of *TMEM114* and *TMLP1* was found to be much closer to the non-claudin members of the Pfam00822 family. Claudins vary in their exon structure with some encoded by a single exon (e.g., *claudin-2*), others by four (e.g., *claudin-1*) or five exons (e.g., *claudin-16*). The genes encoding the γ subunits (*CACNGs*) can be divided into specific subgroups by their exon structure (Table 3.4). Structurally, *TMEM114* is closest to *CACNG1* (encoding

γ1) and *CACNG6* (encoding γ6). TMLP1 in humans is encoded by 5 exons. The murine orthologue is encoded by four exons, with a similar structure to that of *TMEM114*, *CACNG1* and *CACNG6*. Exon 3 in *TMEM114, TMLP1, CACNG1* and *CACNG6* is the same size but the encoded amino acids show little identity. Exon 2 is similar in size to the *CACNGs*, but again they show little similarity. In *TMEM114*, *TMLP1* and the *CACNGs* exon 1 encodes TMD1 and ECL1 but in the *EMP-1, -2, -3* and *PMP22*, an intron divides TMD1 and ECL1 (Table 3.4) (Jetten and Suter 2000), separating these from the other family members. The structure of *CLP-24, TMEM47 and PERP* revealed that they are distinct from the *EMP*s, *PMP22* and *MP20*, despite clustering with them in the phylogenetic tree.

3.10.2.5 Carboxy-termini

Calcium channel γ subunits vary in length from 222 amino acids (γ1) to 425 amino acids (γ8), but typically TMD4 ends at approximately 200 amino acids. Thus, the variety in length is due to the length of the intracellular C-terminus. The topology predictions for TMEM114 and TMLP1 also predicted TMD4 to finish at approximately 200 amino acids, with C-termini of 13 and 23 amino acids, respectively (Figure 3.3).

The C-terminus of claudins can dictate the half life of the protein (Van Itallie *et al.* 2004) and is required for targeting to the tight junction (Ruffer and Gerke 2004). The majority of claudins contain PDZ binding domains in the final residues of their C-termini. These residues, which generally are hydrophobic, are necessary for the interaction of claudins with ZO proteins (Itoh *et al*. 1999) and other PDZ proteins (Roh *et al.* 2002; Jeansonne *et al*. 2003). The TARP proteins also terminate in PDZ-binding domains (Milstein and Nicoll 2009). Although the majority of claudins have tyrosine-valine (YV) type of PDZ binding motifs, the TARPs end in a TTPV motif (Figure 3.7). This motif is required for interaction with PSD-95 involved in targeting AMPARs at the cell membrane to the synapse (Chen *et al*. 2000; Schnell *et al.* 2002). This synaptic targeting is also dependent on a nPIST binding domain in the C-terminus (Cuadra *et al.* 2004). Although γ7 (ending in -TTPC) does not end in a typical PDZ-binding domain it has the ability to bind PSD-95 and traffic AMPARs (Kato *et al*. 2007). γ5, which ends in –STPC does not bind PSD-95 and does not traffic AMPARs to the synapse (Kato *et al*. 2007).

Thus it appears that PDZ-binding motifs facilitate the trafficking of AMPARs to the synapse via interaction of TARPs with PSD-95, whilst in epithelial cells claudins with PDZ-binding motifs bind the tight junction protein ZO-1. The γ subunits which lack PDZ binding domains (γ1, γ5, γ6) also lack the ability to bind AMPARs (Chen *et al*., 2007). Consequently it is therefore unlikely that TMEM114 and TMLP1 interact with ZO-1 or other PDZ proteins involved in trafficking of AMPARs. γ1 and γ6 also have short C-termini (22 and 16 amino acids, respectively) which lack known functional motifs. The short C-termini of TMEM114 and TMLP1 show no similarity to γ1, γ5, γ6 or to any of the claudins. The C-terminus of TMEM114 is moderately conserved (Figure 3.6) suggesting it may be functionally less important than the C-termini of TARPs or claudins. The C-terminus of TMLP1 shows little conservation (Figure

3.6). Primates have a longer C-terminus with the final three amino acids encoded by an additional exon (Figure 3.22) which may represent the addition of a functional domain.

3.10.2.6 Extracellular loop 1 (ECL1)

Claudins form apical tight adhesions between the lateral membranes of neighbouring cells; the tight junctions. The first extracellular loop (ECL1) of claudins is believed to bridge the intercellular space, with opposing loops interacting across the space (Heiskala *et al.* 2001). This head-head *trans*-interaction can be homo- or hetero-philic (Coyne *et al.* 2003; Daugherty *et al.* 2007; Krause *et al*. 2009) and forms charge- (Colegio *et al.* 2002) and size- (Van Itallie *et al.* 2008) selective pores. The charge of amino acids in ECL1 determines the specificity of claudin permeability (Colegio *et al*. 2002; Wen *et al*. 2004).

Non-claudin members of the Pfam00822 family are also associated with intercellular adhesion. PMP22, expressed at high levels in myelin-forming Schwann cells, is also expressed in epithelial cells of the lens, lung and intestine (Baechner *et al.* 1995; Taylor *et al.* 1995). *In vitro* and *in vivo*, Pmp22 colocalises with tight junction markers (Notterpek *et al.* 2001). EMP-1 localises to tight junctions in the endothelial cells of the blood brain barrier (Bangsow *et al.* 2008). Indirect evidence from a variety of sources also suggests a role in adhesion for the lens fibre specific protein MP20. In differentiating lens fibre cells, MP20 inserts into the plasma membrane at a point which correlates with the formation of a barrier to extracellular diffusion (Grey *et al*. 2003) and lens fibres of MP20-/- mice are much more easily separated from each other than those of wildtype mice (Shiels *et al*. 2007). MP20 expression is necessary for formation of the lens syncytium, a process which begins with cell adhesion (Shi *et al*. 2009).

The adhesive capacities of claudins has been demonstrated by stable expression of the protein of interest in mouse fibroblast L-cells (Kubota *et al.* 1999). These cells have little endogenous cell-cell adhesion capabilities (Nagafuchi *et al.* 1987) but cells transfected with claudins aggregate (Kubota *et al*. 1999). Surprisingly, the voltage dependent calcium channel γ2 subunit was also demonstrated to have adhesive properties in L-cells (Price 2005). This suggests there are common properties between claudins and the γ subunits. It has been postulated that the claudins and the γ subunits diverged from a common ancestor and that this common ancestor had an adhesive function (Price *et al*. 2005). Over time the proteins diverged and acquired new functions. Both groups of proteins contain the W-GLW-C-C motif in ECL1 (Figure 3.4). Mutating these sites in claudins affects homotypic *trans*-interactions suggesting that these residues are important for cell-cell-adhesion (Cukierman *et al.* 2009). TMEM114 and TMLP1 contain the same W-GLW-C-C motif in the first extracellular loop and thus may have adhesive functions.

ECL1 of γ2 mediates binding of AMPAR GluR1 but when ECL1 of γ2 is replaced with the ECL1 of γ5 binding to GluR1 does not occur (Tomita *et al*. 2004). This suggests that

although ECL1 is involved in binding of GluR1, the W-GLW-C-C motif is not involved in this binding. This suggests that the function of ECL1 is not limited to the W-GLW-C-C motif.

3.10.2.7 Extracellular loop 2 (ECL2)

In claudins-3 and -4, ECL2 acts as a receptor for *Clostridium perfringens* enterotoxin (Katahira *et al.* 1997; Fujita *et al.* 2000). This causes structural changes in the tight junction resulting in endocytosis of the claudins in the intestine (Katahira *et al*. 1997). Like ECL1, ECL2 forms *trans*-interactions with claudins of opposing cells (Piontek *et al.* 2008) thereby also affecting the transepithelial resistance of the tight junctions (Piehl *et al.* 2010). The *trans*-interactions are thought to be mediated by an interface of aromatic amino acids at p.F147, p.Y148 and p.Y158 (Piontek *et al*. 2008). In TMEM114 and TMLP1 there are phenylalanine residues (F) in ECL2 that are not immediately followed by tyrosine (Y) (Figure 3.5). TMEM114, TMLP1 and the TARPs all end in an aromatic amino acids followed by GWS. TMLP1 also contains two other aromatic amino acids within its ECL2, p.Y162 and p.H166. Most claudins also contain proline residues in ECL2 (Krause *et al*. 2009), as does TMLP1 at p.P164. TMLP1 shows a high degree of similarity to γ6, but their ECL2 show little similarity. Both TMEM114 and TMLP1 show conservation of the beginning and end of ECL2 with the TARPs. As the functional importance of ECL2 in TARPs has yet to be elucidated the significance of this similarity remains to be identified.

3.10.3 Co- and post-translational modifications

Glycosylation of proteins can serve many purposes including protein folding and maintenance of correct conformation, intracellular trafficking and cell surface expression (Helenius 1994; Gahmberg and Tolvanen 1996). The calcium channel γ subunits and EMPs/PMP22/MP20 all contain predicted or confirmed N-linked glycosylation sites, whereas the majority of claudins do not (Figure 3.3) (Pareek *et al.* 1993; Sharp *et al.* 2001; Ervin *et al.* 2005). The N-linked glycosylation site in γ2 has been demonstrated to be functionally important as mutants at this site fail to localise to the cell membrane (Price *et al.* 2005).

The γ1 and γ6 subunits contain two predicted N-linked glycosylation sites in ECL1 at either side of the GLW of the W-GLW-CC motif. TMEM114 also contains sites at either side of the GLW (Figure 3.3). These sites are conserved to zebrafish (Figure 3.4). In γ1 transfected cells treated with the antibiotic tunicamycin, which inhibits N-linked glycosylation, γ1 retains its ability to interact with $Ca_V1.1$ (Arikkath *et al*. 2003). This suggested that γ1 is not glycosylated as it maintains its functional interactions. Treating cells with tunicamycin prevents glycosylation of all N-linked glycoproteins and disrupts cellular function. Therefore, as the localisation of the unglycosylated γ1 was not determined, it is possible that γ1 and $Ca_V1.1$ were not localised at the plasma membrane and that the interaction between γ1 and $Ca_V1.1$ may have occurred in the ER, as loss of glycosylation sites in some membrane proteins is associated with misfolding and retention in the ER (Kelley and Kinsella 2003).

TMLP1 has a single predicted N-linked glycosylation site. This site, although present in canines and bovines, is absent in rodents. This suggests that the prediction may be a false positive or that the N-linked glycosylation at this residue may not be functionally significant. Notably although all of the non-claudins proteins mentioned contain predicted sites, *in vivo* glycosylation has not been demonstrated for all. MP20 contains a predicted N-glycosylation site and interacts with galectin-3 (Gonen *et al*. 2001), which binds oligosaccharides (Fradin *et al.* 2000; Hirabayashi *et al.* 2002). However, glycosylation of the site has not been demonstrated (Ervin *et al.* 2005). N-linked glycosylation predictions are based on the presence of the N-X-S/T consensus motif (where X is any amino acid except proline). The presence of the N-X-S/T site in a protein does not guarantee the addition of N-linked oligosaccharides as the conformation of the protein may prevent access of the glycosylation machinery to the site (Apweiler *et al.* 1999). Thus, the glycosylation of TMEM114 and TMLP1 must be confirmed *in vivo*.

The intracellular C-termini of claudins contains phosphorylation sites for protein kinase C (PKC), casein kinase II and c-amp dependent kinase (Van Itallie and Anderson 2006). The effect of the phosphorylation state of claudins is dependent of the specific kinases and phosphatases, as dephosphorylation has been shown to both enhance (Nunbhakdi-Craig *et al.* 2002) and disrupt (Banan *et al.* 2005) epithelial barrier function. Similarly, phosphorylation can enhance (Fujibe *et al.* 2004) or decrease (Kahle *et al.* 2004; Yamauchi *et al.* 2004) barrier function. The PDZ-binding domain (TTPV) of TARPs contains multiple phosphorylation sites (Chen *et al.* 2007) and phosphorylation of γ2 by mitogen-activated protein kinase (MAPK) and Protein Kinase A (PKA) modulate the synaptic clustering of γ2 (Stein and Chetkovich 2010).

The intracellular C-terminus of TMEM114 contains a single PKC site (Figure 3.8). ECL1 also contains DNA-dependent protein kinase (DNAPK) and PKA sites. Therefore TMEM114 may be regulated at the protein level by phosphorylation. The predicted phosphorylation sites for TMLP1 show little conservation suggesting that it is less likely to be regulated by phosphorylation. Similarly, there are few predicted phosphorylation sites in γ1 and γ6 and those that are present are not conserved (Chen *et al*. 2007).

Only γ5 and γ7 are predicted to be phosphorylated at tyrosine residues (Chen *et al*. 2007). These sites are present in the intracellular C-terminus. Human and murine TMEM114 contain a predicted tyrosine phosphorylation site in ECL1. Phosphorylation of tyrosine residues is associated with signal transduction (Medvedev *et al.* 2007; Metere *et al.* 2009). The TARPs and γ7 contain predicted tyrosine sulphation sites in ECL1 (Chen *et al*. 2007). Sulphation of tyrosines occurs in the Golgi apparatus lumen and strengthen protein-protein interactions (Moore 2003). No sulphation sites were predicted for TMEM114 or TMLP1.

Palmitoylation is a reversible post-translational modification that involves the attachment of the lipid palmitate to cysteine residues of membrane proteins (Wolven *et al.* 1997). This

typically occurs in the border region between the transmembrane domain and the cytoplasmic domain. Palmitoylation of membrane proteins has roles in trafficking (Singaraja *et al.* 2009), signalling (Macdonald-Obermann and Pike 2009), protein-protein interactions (Yu *et al.* 2006), internalisation (Ohno *et al.* 2009) and degradation (Percherancier *et al.* 2001). Palmitoylation of claudin-14 near the cytoplasmic end of TMD2 and TMD4 enhances membrane localisation (Van Itallie *et al.* 2005). The other claudin family members also contain predicted palmitoylation signals at this site (Van Itallie *et al*. 2005). The calcium channel γ subunits also contain predicted palmitoylation sites, γ6 at the cytoplasmic end of TMD2, the TARPs at their N-termini, and γ5 and γ7 at the N- and C-termini (Chen *et al*. 2007).

The predicted palmitoylation site for TMEM114 was likely to be a false positive as the prediction score was weak and the site is present in the extracellular domain. TMLP1 contains two predicted sites at the cytoplasmic end of TMD2, similar to γ6 subunit, although the surrounding amino acids show little similarity. Thus, palmitoylation may have a role in TMLP1 function.

Signal sequences are N-terminal sequences that direct transport of nascent peptides of membrane proteins to the ER. Signal sequences are recognised by the signal recognition particle (SRP) which bind to SRP receptors in the ER membrane (Janda *et al.* 2010). Typically, signal sequences are cleaved after transport. Cleaved signal sequences are known as signal peptides, non-cleaved signal sequences are known as signal anchors. Both TMEM114 and TMLP1 contain predict signal peptides, with <2 % probability of being signal anchors (SignalP). The predicted signal peptides include the first transmembrane domain. However, it is well documented that misclassification of signal peptides in transmembrane proteins with transmembrane domains near the N-terminus can occur (Kall *et al.* 2007; Reynolds *et al*. 2008). The prediction tool Philius is a combined transmembrane topology and signal peptide program designed to overcome the confusion in signal peptide prediction (Reynolds *et al*. 2008). Philius confirmed the presence of signal peptides in TMEM114 and TMLP1, although the confidence level for TMEM114 was only 63 %. However, the likelihood that TMEM114 and TMLP1 contain signal peptides is strengthened by the observation that all members of the claudins, γ subunits and EMPs/PMP/MP20 were tested and not predicted to contain signal peptides. Cleavage of signal peptides which are predicted as transmembrane domains is thought to occur for both the calcitonin receptor (Uniprot P30988) and calcitonin receptor-like receptor (CRLR) (Uniprot Q16602) (Walker *et al.* 2010). Cleavage of these TMD peptides results in the N-terminus of the mature protein being extracellular. The alteration of the structure of the protein is ultimately likely to affect the function of the protein, but without functional evidence for other proteins, it is difficult to predict the functional effect on TMEM114 or TMLP1. In the uncleaved proteins, the beginning of ECL1 is tethered to the transmembrane domain. Cleavage of the signal peptide would release this domain from the plasma membrane making it accessible to extracellular molecules, suggesting this region may contain interacting domains.

3.10.4 Expression of *TMEM114* and *TMLP1*

Human *TMEM114* ESTs were identified in the lens, as was found experimentally in mouse (Jamieson *et al*. 2007). The highest expression levels were found in lung carcinoids, but there were no ESTs from wildtype lung tissue. Claudin dysregulation is associated with lung cancers (Chao *et al.* 2009) and multiple other forms of cancer (Jung *et al.* 2010; Kuo *et al.* 2010). The type of claudin being misexpressed and the direction of misexpression can be used to differentiate between carcinomas (Moldvay *et al.* 2007; Paschoud *et al.* 2007).

Expression of Tmem114 is present in the cerebellum of mouse (Jamieson *et al*. 2007), but ESTs of human TMEM114 were not present in the cerebellum. It is not known if the expression in the cerebellum is present in the endothelium like claudins (Nitta *et al.* 2003; Shimizu *et al.* 2008), or if there is neuronal expression like the γ subunits (Burgess *et al*. 2001). Like PMP22, TMEM114 potentially may be expressed in epithelial (Baechner *et al*. 1995), endothelial (Roux *et al.* 2004) and neuronal cells (Parmantier *et al.* 1995).

ESTs for *TMLP1* were only available for the human orthologue where it was found in the hippocampus and hypothalamus. Human and murine *Cacng2*, *3*, and *4* are expressed in the hippocampus and multiple other regions of the brain (Sharp *et al.* 2001; Moss *et al.* 2003). The stargazing-encoding *Cacng2* is particularly highly expressed in the hippocampus (Sharp *et al.* 2001). TMLP1 shows greatest similarity to γ6. However, γ6 is predominantly expressed in skeletal and cardiac muscle, with limited expression in the brain (Burgess *et al*. 2001; Chu *et al*. 2001; Fukaya *et al*. 2005). The *TMLP1* ESTs were detected in two splice isoforms which lack exon 2 or exons 2 and 3. Loss of these exons does not alter the frame of the encoded protein. Loss of exon 2 shortens ECL1 and results in the loss of the second cysteine of the W-GLW-C-C motif. It has been postulated, but never demonstrated that the two cysteines of this motif form disulphide bonds (Van Itallie and Anderson 2006). If disulphide bond formation does occur, then omission of exon 2 may be functionally significant. Exons 2 and 3 encode the latter part of ECL1, TMD2 and half of TMD3, thus omission of this exon alters the topology of the protein. The topology prediction tools predict that this isoform has three TMDs with an extracellular C-terminus. *CACNG6* is also detected in alternate isoforms which lacks exon 3 or exons 2 and 3 (Burgess *et al*. 2001). This is predicted to result in the in-frame loss of TMD2 and TMD3 (Burgess *et al*. 2001). The functional significance of this isoform has yet to be elucidated.

3.11 Conclusion

As well as showing more similarity to each other than to any other protein, TMEM114 and TMLP1 are distinct from the other Pfam00822 proteins due to the presence of a predicted signal peptide (Figure 3.23). TMEM114 and TMLP1 belong to the same clade as the voltage dependent calcium channel γ subunits and the EMPs, PMP22 and MP20. The pairwise amino acid alignments revealed that TMLP1 shows greatest (25 %) identity with γ6. This level of identity was greater than the level of identity between γ6 and all of the subunits other than

γ1 (Burgess *et al*. 2001). The exonic structure of TMLP1 in non-primates also resembled that of γ6. Similar to TMLP1, γ6 is short with a short intracellular C-terminus which lacks PDZ binding domains. The multiple splice isoforms of TMLP1 also reflected those of γ6. However, there were notable differences between TMLP1 and γ6.

TMLP1 is expressed in the brain, but γ6 is mainly expressed in skeletal and cardiac muscle, with much lower levels of expression in the brain (Burgess *et al*. 2001; Chu *et al*. 2001). γ6 decreases activity of the cardiac low voltage activated $Ca_v3.1$ *in vitro* (Hansen *et al*. 2004). This modulation is dependent on a GxxxA motif in TMD1 (Lin *et al*. 2008). This motif is not present in TMLP1 although there are GxxxxA and GxxxxxA motifs in TMD1. γ6 null mice are phenotypically normal (http://www.informatics.jax.org/external/ko/lexicon/2332.html) although information regarding voltage dependent calcium channel activity is not available for these mice. Unlike the TARPs, the ability of γ6 to traffic proteins has not been demonstrated.

The γ subunit family are thought to have evolved from a common ancestor through a combination of tandem repeats and chromosomal duplications (Burgess *et al*. 2001; Chu *et al*. 2001). These genes are thought to have evolved through duplication with the ancestor of the PKC genes, as three of the *CACNG* loci are in linkage with *PKCA*, *PKCB* and *PKCG* (Burgess *et al*. 2001). In humans *CACNG6*, *CACNG*7 and *CACNG*8 are located at 19q13.4, *CACNG1*, *CACNG4* and *CACNG5* are located at 17q24, and *CACNG2* and *CACNG3* are found individually at 22q13.1 and 16p12, respectively. With the limited information regarding TMLP1 orthologues being currently available it is difficult to ascertain if TMLP1 could have arisen through duplication of γ6 or the ancestor of γ6.

TMEM114 was previously reported as being related to the claudins, EMPs, PMP22 and MP20 (Jamieson *et al*. 2007). The analysis performed here suggests that in terms of pairwise alignment and amino acid identity TMEM114 displays most similarity to the TARPs. However, there are clear differences between TMEM114 and the TARPs. The TARPs have long intracellular C-termini which contain interact with nPIST for trafficking to the plasma membrane where the PDZ-binding motif interacts with PSD-95 to mediate trafficking to the synapse (Cuadra *et al*. 2004). TMEM114 lacks both of these interacting domains. However, EMP2, which also lacks nPIST and PDZ binding domains, modulates the trafficking of various proteins to the cell surface (Wadehra *et al.* 2004; Wadehra *et al.* 2005). The TARPs are predominantly expressed in neurons (Letts *et al*. 1998) but Tmem114 is expressed in the epithelium of the lens (Jamieson *et al*. 2007). The amino acid identity of TMEM114 with PMP22, a protein which localises to tight junctions in epithelia (Notterpek *et al.* 2001), was also high (Table 3.2). Like TMEM114, PMP22 does not appear to contain a PDZ-binding domain. Thus TMEM114 may localise to tight junctions in epithelia.

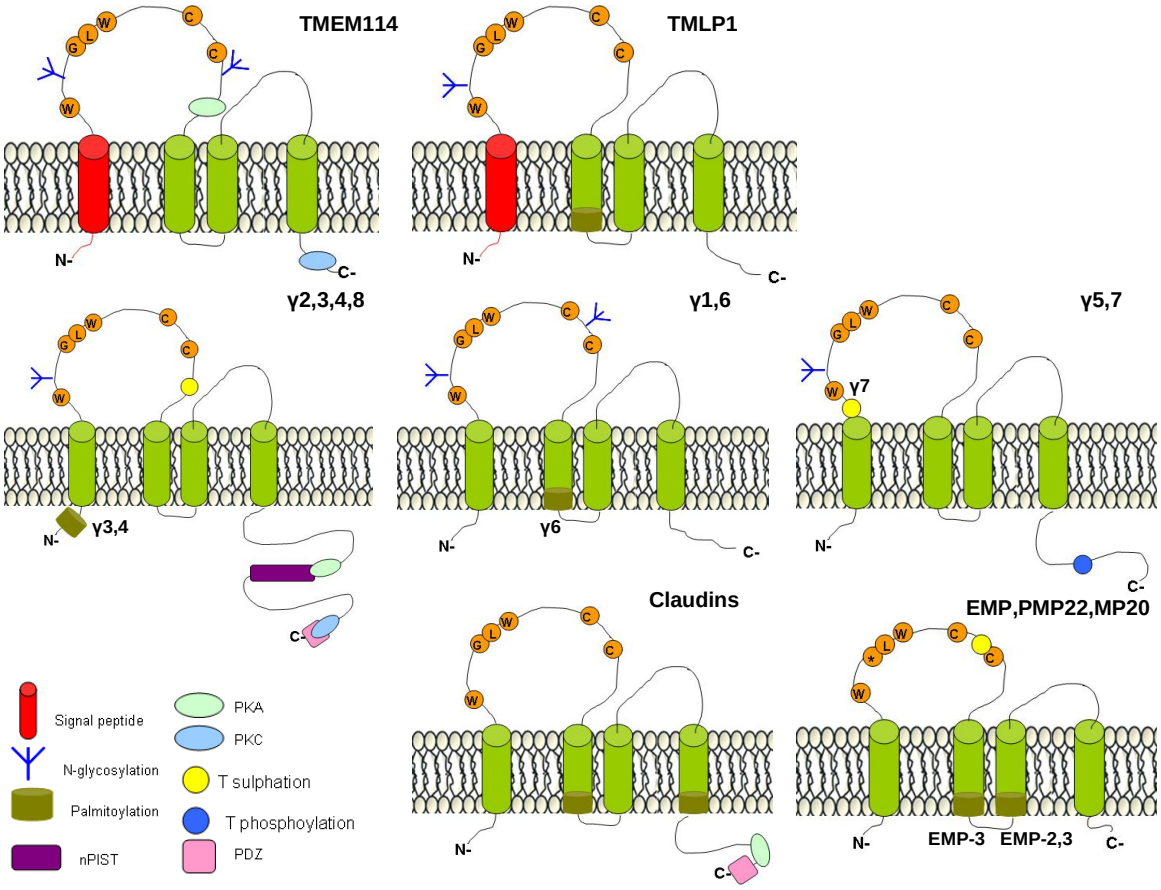


Figure 3.23. Schematic displaying the predicted topologies and predicted co- or post-translational modifications of TMEM114, TMLP1, voltage dependent calcium channel gamma subunits, claudins and EMPs, PMP22 and MP20. * = G is not present in EMP2, EMP3, PMP22. Protein names close to modifications indicate that only these specific proteins contain the modification.

Chapter 4: Expression and localisation of Tmem114

4.1 Introduction

The developmental expression of *Tmem114* was previously described in mouse (Jamieson *et al*. 2007). In this chapter the aim was to further characterise the developmental expression of *TMEM114* in humans, and by generating polyclonal antibodies to Tmem114, determine the expression of the Tmem114 protein in mouse. As TMEM114 is a predicted transmembrane protein the localisation of Tmem114 is investigated in a model epithelial cell system. The presence of co- and post-translational modifications is also investigated.

4.2 Expression of *TMEM114*

4.2.1 Expression of *TMEM114* in human embryonic development

Tmem114 has previously been shown to be expressed in the eye, cerebellum and testes of adult mice and expressed as early as E13.5 in embryonic mice (Section 1.7.5) (Jamieson *et al*. 2007). Using a panel of human embryonic cDNA from eyes and neural tissue (forebain, midbrain, hindbrain, spinal cord) from three developmental stages (Carnegie stage 16/days 37-42; Carnegie stage 21/days 53-54; Foetal stage 2/day 70-77) the expression of *TMEM114* in the developing human was determined. Expression was not detected in any of the tested tissue at CS16 (data not shown). At stages CS21 and F2 expression was restricted to the developing eye where expression increased with development (Figure 4.1).

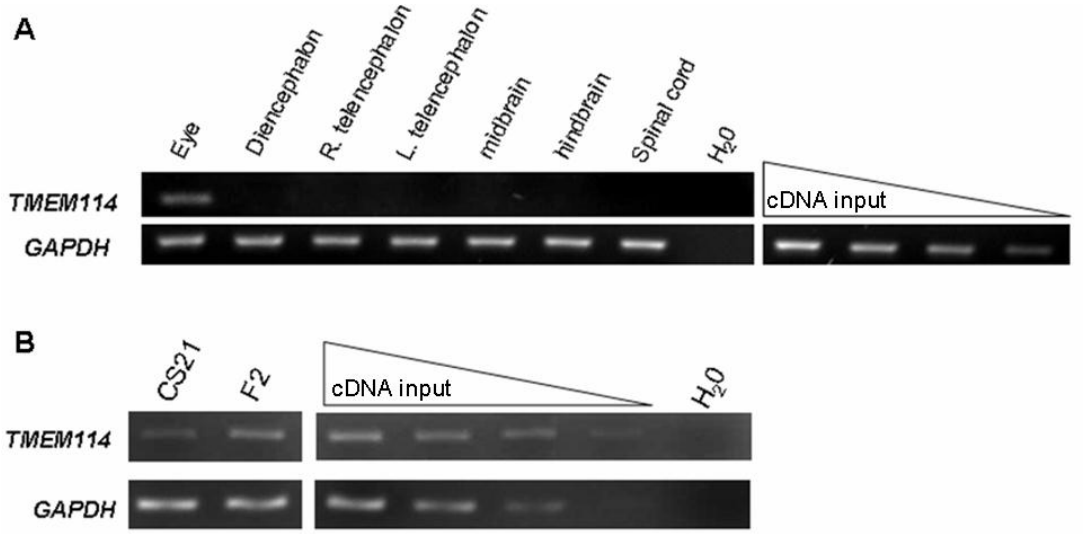


Figure 4.1. Embryonic expression of *TMEM114* determined by RT-PCR. Expression of *TMEM114* was detected by amplifying cDNA with primers amplifying exons 1 and 2. A) Expression in a panel of eye and CNS tissue at Carnegie Stage 21 (CS21). At CS21 expression of *TMEM114* is detected in the developing eye. B) Expression in the developing eye. Expression of *TMEM114* increases between developmental stages CS21 and Foetal stage 2 (F2). Detection of the housekeeping gene *GAPDH* was used as a control for equal cDNA loading in the RT-PCR reactions. The endpoint of all PCRs was determined to be within the exponential phase as demonstrated by the linearities for each reaction (triangle). H_20 represents a no template control.

4.2.2 Expression of human and canine *TMEM114* in cell lines

Expression of *TMEM114* was not detected in the human lens epithelial cell line FHL124 or in Human Embryonic Kidney (HEK) 293 cells (Figure 4.2A). MDCK II cells are of canine origin. As no canine positive control was available, primers were designed to amplify canine *Tmem114* as well as human *TMEM114* to act as a positive control. Canine *Tmem114* expression was not detected in MDCK II cells (Figure 4.2B).

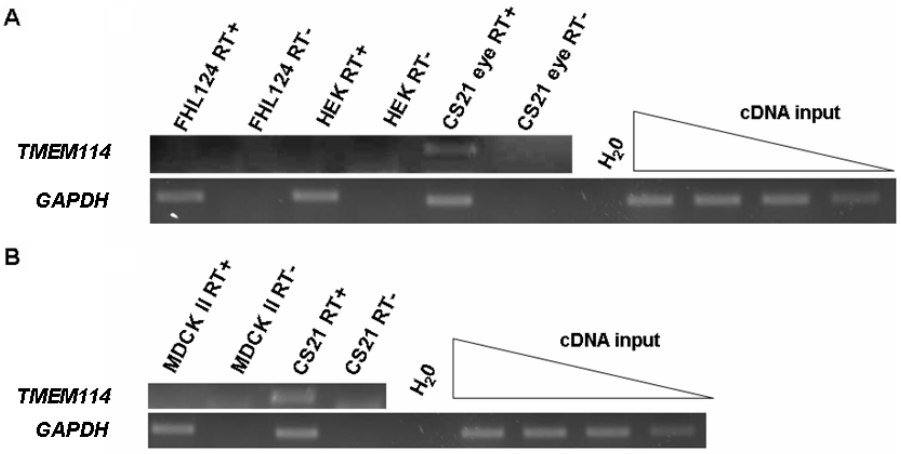


Figure 4.2. Expression of human *TMEM114* and canine *Tmem114* in cell lines. A) *TMEM114* expression (exon 1 – exon 2) was not detected in the human cell lines FHL124 and HEK293. cDNA from a CS21 eye was used as a positive control. B) Tmem114 is not expressed in the canine kidney epithelial cell line MDCK II. Primers amplifying exons 3 and 4 of canine Tmem114 were designed to also amplify human cDNA. cDNA from a CS21 eye was used as a positive control. Detection of the housekeeping gene GAPDH was used as a control for equal cDNA loading in the RT-PCR reactions. The endpoint of all PCRs was determined to be within the exponential phase as demonstrated by the linearities for each reaction (triangle). 'RT-' are samples to which no reverse transcriptase was added. H_20 represents a no template control.

4.3 Generation of a polyclonal antibody to murine Tmem114

As an image clone for human *TMEM114* was not available, our studies were performed using the highly similar murine orthologue for which a FANTOM clone was available (RIKEN cDNA 4930511J11). C-terminal V5-tagged *Tmem114* (Tmem114-V5) had previously been cloned from this FANTOM clone (Urquhart, unpublished). From this V5-tagged clone, untagged *Tmem114* was cloned into the pGEM-T easy vector (Figure 9.2) and subsequently into the expression vector pcDNA3.1(-) (Figure 9.1). Constructs were bi-directionally sequenced to confirm that the *mTmem114* sequence was correct and in-frame.

4.3.1 Epitope design

The final C-terminal 14 residues of mTmem114 are predicted to localise intracellularly (Table 3.1). This intracellular C-terminus is conserved in Tmem114 and showed no significant similarity to any known protein. This region is highly conserved in human and bovine Tmem114 which would potentially enable the use of the anti-Tmem114 antibody in multiple species (Figure 4.3).

```
Homo            VDISFGWSLALGWISFIAELLTGAAFLAAARELSLRRRQDQAI 223
Bos             VDIRFGWSLALGWVSCVAELLTGATFLVAARVLSLRRRQDQAI 223
Mus             VDIRFGWSLALGWISFVSELLTGVVFLAAARALSLSQRQDQAI 222
                *** *********:* ::*****..**.*** *** :****** 
```

Figure 4.3. Epitope used for polyclonal antibody generation. Protein sequences were aligned with ClustalW2. The 14 amino acid epitope used for polyclonal antibody production is based on murine Tmem114 and is boxed in red. This C-terminal region is conserved in humans and bovines. The predicted fourth transmembrane domain of murine Tmem114 is in bold.

4.3.2 Initial test bleeds

Keyhole limpet haemocyanin (KLH) conjugated mTmem114 peptides were synthesised, purified and used to immunise two New Zealand white rabbits (#4671 and #4672) (Section 2.6.1.1). The initial test bleeds were then compared to the pre-immune serum for their reactivity to non-transfected MDCK II cells and MDCK II cells transfected mTmem114 by western blot. MDCK II cells expressing V5-tagged mTmem114 were used as a control to determine if the bands detected were specific to mTmem114. Four main bands ranging from approximately 45 kDa to 23 kDa were detected by the V5 antibody (Figure 4.4). The 23 kDa band is close to the predicted molecular mass of 25.7 kDa for mTmem114 with the V5 epitope. Bleed 3 from rabbit #4671 detected bands which were not present in the non-transfected cells and were not detected by the pre-immune serum (Figure 4.4A). The lowest molecular weight band at 22 kDa is close to the predicted molecular mass of 24.4 kDa for mTmem114. Additionally, the bands were comparable to those detected with an anti-V5 antibody against lysates from mTmem114-V5 transfected cells. These data suggest that the detected bands are various forms of mTmem114. Non-specific bands were detected at approxiamtely 36 kDa in the transfected and non-transfected samples (Figure 4.4A). The anti-mTmem114 antisera did not detect these 4 bands in cells transfected with V5-tagged mTmem114. As the V5-tag was immediately C-terminal to the epitope, it is possible that the V5-tag was blocking access of the antibody to the epitope. Rabbit #4672 failed to detect mTmem114 by western blot (Figure 4.1B).

4.3.3 Final test bleeds and purified antibodies

Rabbits #4671 and #4672 were immunised twice more (Section 2.6.1.1) and the final bleeds were tested. The final bleed of rabbit #4672 failed to detect transfected mTmem114 (data not shown). The final bleed of rabbit #4671, although functional at 1:1000 dilution (Figure 4.5), had significantly lower titres than the initial bleeds, as shown by the ELISA results (Figure 4.6). The specificity of the final bleed was confirmed using peptide incubation. When the antisera was incubated with the peptide antigen prior to use, the antibody failed to detect mTmem114, confirming that the antibody is specific against the mTmem114 antigen (Figure 4.7). The final bleed of rabbit #4671 was chosen for purification on an affinity column containing the purified peptide used as the antigen. The purified Tmem114 antibody identified Tmem114 although a 1:100 dilution with a long exposure time was required (Figure 4.5 B).

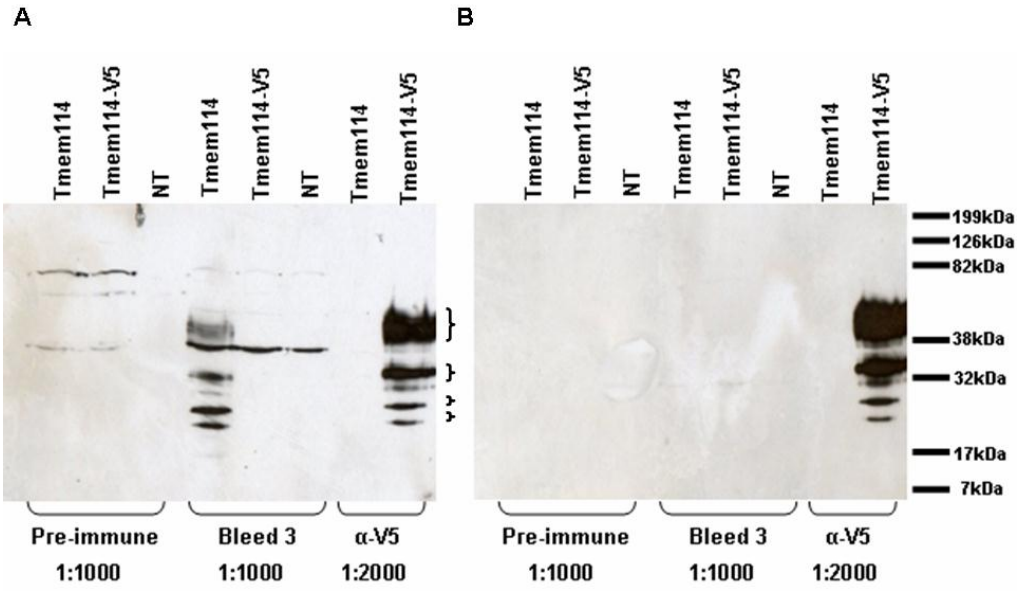


Figure 4.4. Testing of anti-mTmem114 antisera from rabbits #4671 (A) and #4672 (B). Sera were tested on non-transfected MDCK II cells (NT) and MDCK II cells transfected with untagged Tmem114 (Tmem114) or V5-tagged Tmem114 (Tmem114-V5). V5-tagged mTmem114 detected by western blot displays 4 bands ranging from approximately 23-45 kDa (braces). A) Bleed 3 from rabbit #4671 recognised these same bands when probed against untagged mTmem114, but not when probed against V5-tagged mTmem114. One non-specific band of approximately 36 kDa is present in all samples detected by bleed 3 and these bands were also present in the pre-bleed. B) Bleeds from rabbit #4672 failed to detect mTmem114.

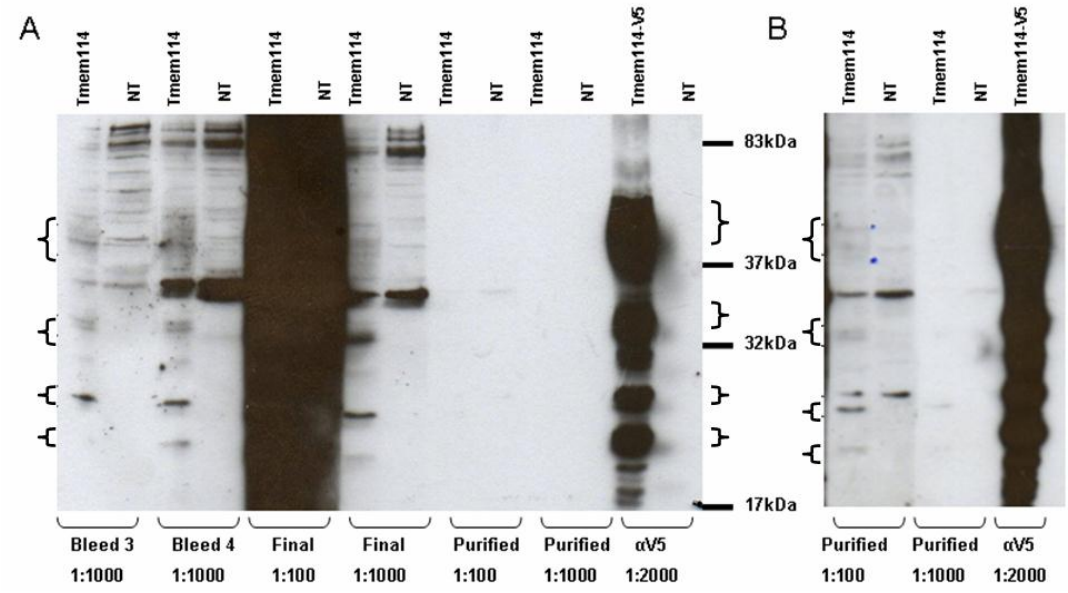


Figure 4.5. Testing of final bleed antisera and purified anti-mTmem114 antibody from rabbit #4671. Sera were tested on non-transfected MDCK II cells (NT) and MDCK II cells transfected with untagged Tmem114 (Tmem114) or V5-tagged Tmem114 (Tmem114-V5). A) 10 minute exposure. B) 1hr exposure. Bands specific (marked by braces) to the MDCK II transfected lysates (Tmem114) that were not present in the non-transfected cells (NT) were detected by bleeds 3, 4 and the final bleed. The purified antibody failed to detect the bands of interest when exposed for 10 mins but faint expression was detected at 1:100 dilution when over-exposed.

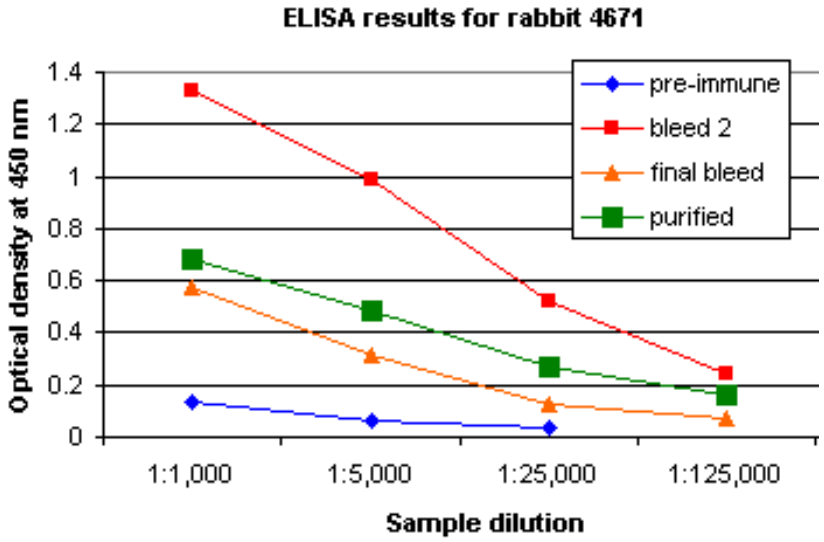


Figure 4.6. ELISA results for rabbit #4671 bleeds and purified antibody against mTmem114. At a dilution of 1:1000, reads for all tested post-immunised sera were greater than the pre-immunised serum (blue). The final bleed (orange) and purified antibody (green) were much weaker than Bleed 2 (red).

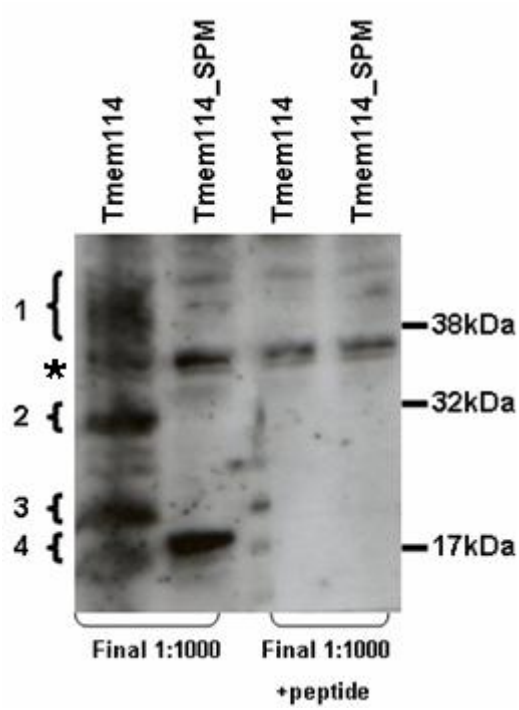


Figure 4.7. Peptide incubation. To confirm the specificity of the #4671 bleed, the antibody was pre-incubated with the peptide antigen before probing the transfected MDCK II lysates by western blot. Pre-incubation ablated the detection of the 3 bands (braces 1-3), confirming they represent mTmem114. Tmem114_SPM represents mTmem114 with its signal peptide (amino acids 2-28) deleted. This was detected at a lower molecular weight (brace 4) than WT Tmem114. Pre-incubation with the peptide ablated the detection of this solitary band confirming it represents mTmem114. Non-specific bands are detected at approxiamtely 36 kDa (asterisk).

4.4 Immunofluorescent staining of mTmem114

4.4.1 Localisation of V5-tagged mTmem114 in FHL124 cells

TMEM114 is a predicted transmembrane protein expressed in the lens epithelium (Jamieson *et al*. 2007). To identify the localisation of mTmem114 in an appropriate cell line, mTmem114-V5 was transfected in the lens epithelial cell line FHL124. In FHL124 cells mTmem114-V5 localised in the cytoplasm and showed strong perinuclear staining (Figure 4.8).

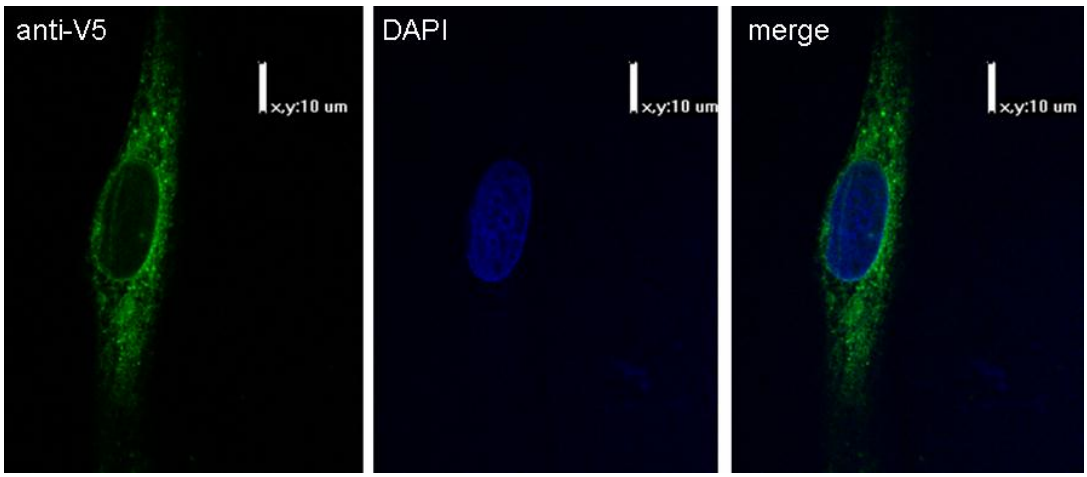


Figure 4.8. Localisation of V5-tagged mTmem114 in FHL124 cells. mTmem114-V5 (green) was localised in the cytoplasm and showed strong perinuclear staining. No staining was observed in the nucleus (labelled blue with DAPI). Scale bar = 10 µm

4.4.2 Localisation of V5-tagged mTmem114 in MDCK II cells

4.4.2.1 Transient expression in MDCK II cells

MDCK II cells are routinely used to study the localisation of claudins. When mTmem114-V5 was transiently expressed in MDCK II cells, Tmem114-V5 localised at the cell membrane (Figure 4.9). Transient over-expression also resulted in the detection of cytoplasmic mTmem114-V5. The stable, moderate expression of a gene is likely to form a truer representation of protein localisation than seen with the high levels of transient expression.

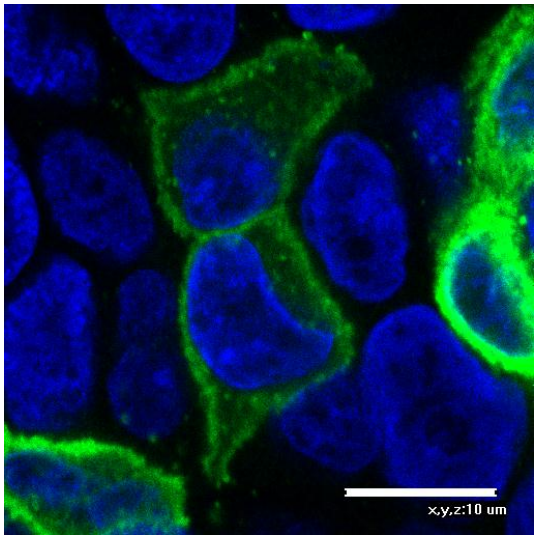


Figure 4.9. Membrane localisation of V5-tagged mTmem114 in transiently transfected MDCK II cells. Tmem114-V5 is detected at the cell membrane and in the cytoplasm. The cells on the right are over-expressing Tmem114-V5. Cells were fixed with 4 % PFA, permeabilised with 0.2 % Triton X-100, and visualised with Leica DM 5000 fluorescent microscope. Bar = 10 µm

4.4.2.2 Stable expression of mTmem114 in MDCK II cells

To establish MDCK II cell lines which stably express either untagged mTmem114 or V5-tagged mTmem114, cells were transfected with a linearised expression vector containing a neomycin resistance gene. Colonies that continued to grow in G-418 selection media were isolated and cultured further. Colonies expressing the protein of interest were then identified by western blot (Figure 4.10A). When these cells were visualised by immunofluorescence, it was found that the cells were not clonal, as some cells did not express the protein of interest (Figure 4.10B). In the cells expressing the V5-tagged (Figure 4.10Bi) or untagged mTmem114 (Figure 4.10Bii), the majority of the protein localised to the cell membrane, with a small pool of protein in the cytoplasm. Cells that failed to express mTmem114 displayed little background staining, confirming the specificity of the antibody.

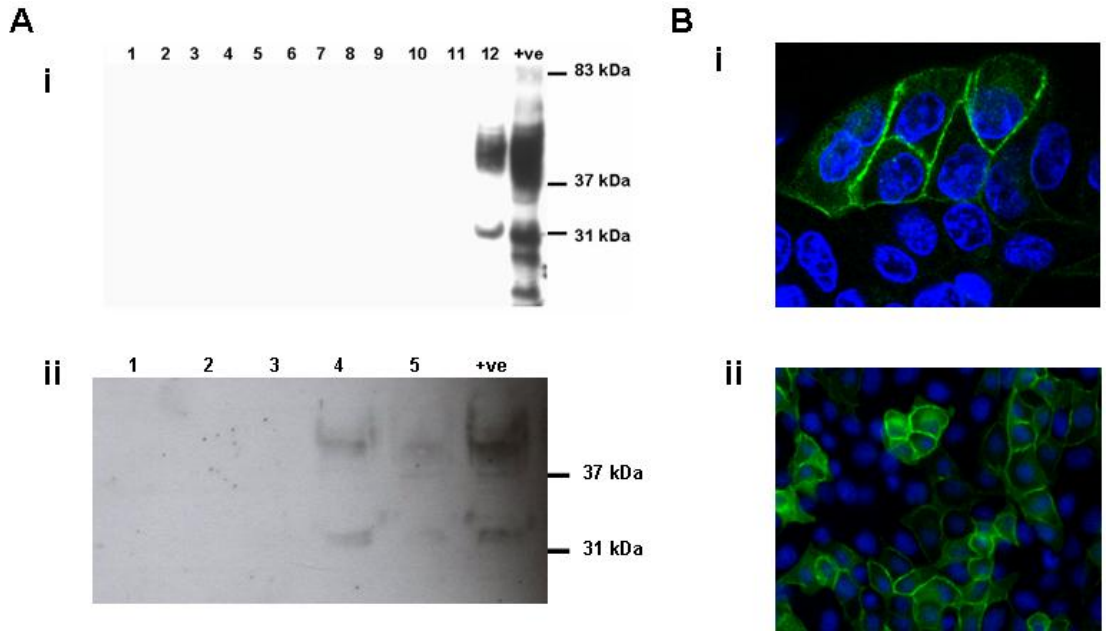


Figure 4.10. Establishment of MDCK II cell lines which stably express V5-tagged or untagged mTmem114. A) Western blot of cell lysates from G-418 resistant colonies. i) Detection of V5-tagged mTmem114 using anti-V5 antibody (1:5000) demonstrating positive expression from colony #12. ii) Detection of untagged mTmem114 using anti-mTmem114 antibody (1:1000) demonstrating positive expression from colonies # 4 and 5. B) Stable Tmem114-expressing cells visualised by immunofluorescence using anti-V5 (1:1000) (i) and anti-mTmem114 (1:1000) (ii) antibodies. Both the tagged (i) and untagged (ii) colonies were not clonal as not all cells expressed mTmem114-V5 (i) or mTmem114 (ii).

4.4.2.3 Localisation of Tmem114 in polarised MDCK II cells

To determine if mTmem114 is present at tight junctions, like claudins and other members of Pfam00822 family (Section 3.10.2.6), confluent MDCK II cells stably expressing untagged and V5-tagged mTmem114 were grown on transwell inserts until polarised. MDCK II cells grown in this manner were shown to be polarised by labelling with the basolateral marker E-cadherin and the tight junction marker ZO-1 (Figure 4.11). The polarised Tmem114 stable cells were

then labelled with an anti-ZO-1 antibody and anti-Tmem114 antibodies and visualised using a confocal microscope. Tmem114 localised to the apical surface and at lateral membranes but did not colocalise with tight junction marker ZO-1 (Figure 4.12A). The lateral membrane localisation of Tmem114 resembled that of the adherens junction protein E-cadherin, so further colocalisation studies with adherens junction proteins and other junctional proteins were attempted. Tmem114 colocalised with E-cadherin at the lateral membrane (Figure 4.12B). A similar localisation was detected with V5-tagged Tmem114-V5 (Figure 4.12C).

To identify the precise localisation in polarised MDCK II cells, Tmem114 was labelled with immuno-nanogold and viewed under an electron microscope. Tmem114 localised to the microvilli of the apical cell membrane and to the apical regions of the cytoplasm (Figure 4.13). Low levels of staining were also observed in non-transfected cells (data not shown).

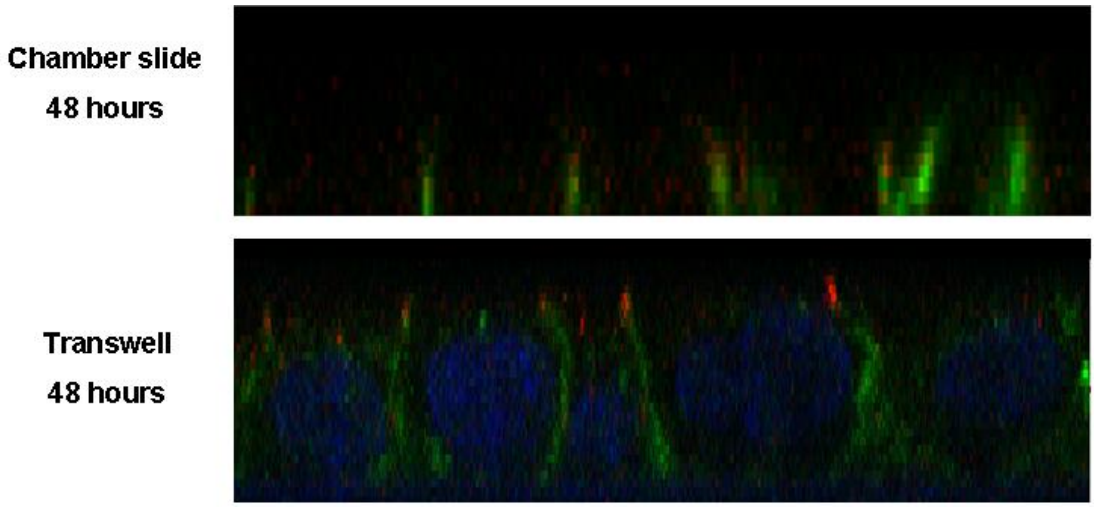


Figure 4.11. Establishment of a polarised epithelial cell line. MDCK II cells were seeded at the same density on chamber slides or Transwell filters and labelled 48 h post-seeding. In the cells grown on the chamber slide (top), the adherens junction marker E-cadherin (green) and the tight junction marker ZO-1 (red) colocalised. When cells were grown on transwell inserts (bottom), ZO-1 localised to the tight junction above E-cadherin. Cells were fixed with 1:1 methanol:acetone, labelled, and viewed by confocal microscopy.

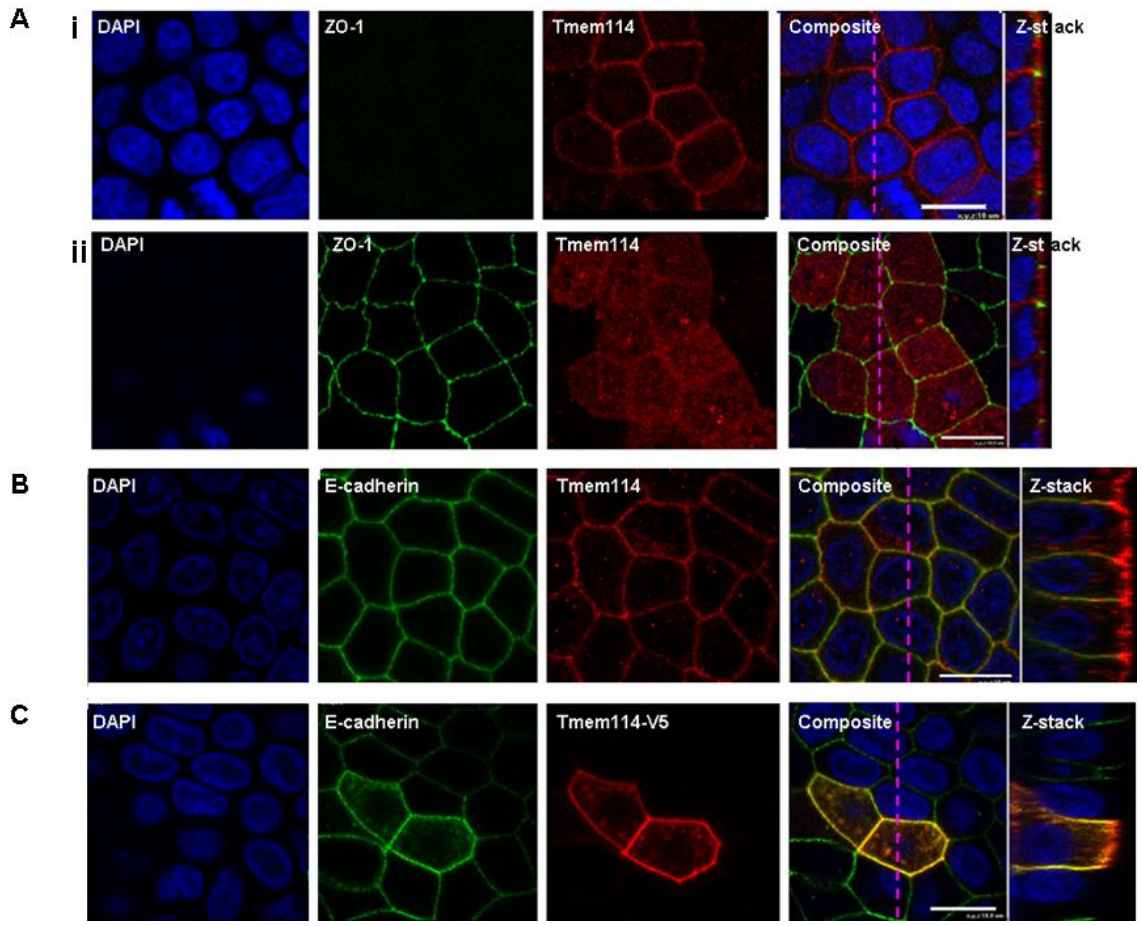


Figure 4.12. Localisation of stably-expressed mTmem114 and mTmem114-V5 in polarised MDCK II cells. Cells were labelled with anti-mTmem114 (red), anti-V5 (red), anti-ZO-1 (green), or anti-E-cadherin (green) and stained with DAPI (blue) as indicated. Colocalisation of red and green labels shows as yellow in the composite image. A) mTmem114 localises to the lateral cell membrane (Ai) and apical cell membrane (Aii) but does not colocalise with the tight junction marker ZO-1 (Aii). B) mTmem114 is predominantly expressed at the apical cell membrane and colocalises with E-cadherin at the lateral cell membrane. C) mTmem114-V5 localises to the apical cell membrane and colocalises with E-cadherin at the lateral cell membrane. Z-stack images to the right of each panel show Y-Z cross-sections through the dashed pink line in the X-Y composite images. The left and right sides show the basal and apical membranes, respectively. Bar = 10 µm. Cells were fixed with 1:1 methanol:acetone and viewed by confocal microscopy.

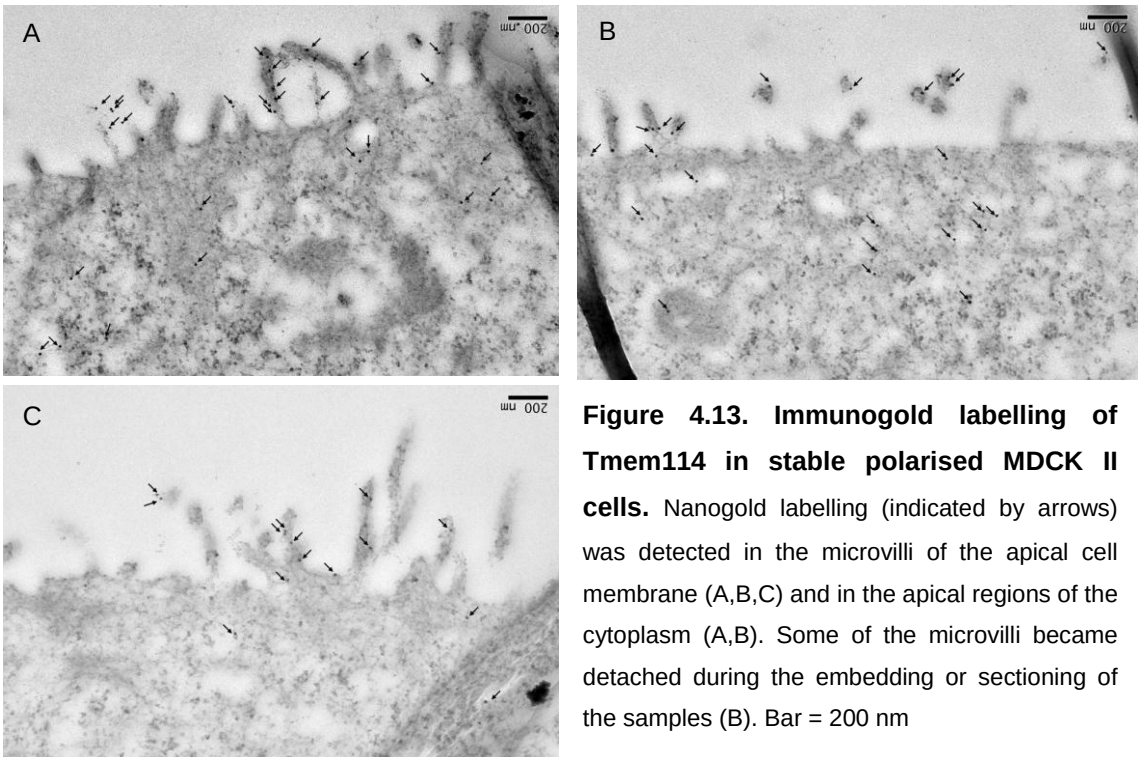


Figure 4.13. Immunogold labelling of Tmem114 in stable polarised MDCK II cells. Nanogold labelling (indicated by arrows) was detected in the microvilli of the apical cell membrane (A,B,C) and in the apical regions of the cytoplasm (A,B). Some of the microvilli became detached during the embedding or sectioning of the samples (B). Bar = 200 nm

4.5 Glycosylation of TMEM114

4.5.1 Establishing the glycosylation status of mTmem114

Murine Tmem114 contains two putative N-glycosylation sites at p.N54 and p.N88 (Figure 3.4) and when detected by western blot it has a greater molecular mass than predicted (Section 4.3.2). This suggested post-translational modification of the protein. N-Glycosidase F (PNGase F) is an amidase that removes all N-linked oligosaccharides by cleaving the innermost N-acetylglucosamine (GlcNAc) from the asparagine residue. When mTmem114-V5 is treated with PNGase F it is detected at 23 kDa, which approximately corresponds to the predicted mass of mTmem114, confirming that Tmem114 is N-glycosylated (Figure 4.14A). The antibiotic tunicamycin prevents N-glycosylation by inhibiting *N*-acetylglucosamine transferases. When mTmem114-V5 transfected cells were treated with varying amounts on tunicamycin, mTmem114-V5 was detected at 23 kDa in the treated cells, but at 40-45 kDa and 30 kDa in the mock treated (Figure 4.14B). This confirms, by an independent method, that mTmem114 is N-glycosylated.

The enzyme Endoglycosidase H (Endo H) cleaves N-linked mannose-rich oligosaccharides, some N-linked hybrid oligsaccharides but not N-linked complex oligosaccharides. Treating mTmem114-V5 lysates with Endo H resulted in a shift in isoform (Figure 4.14C). The 40-45 kDa isoform was present in both untreated and Endo H treated samples but there was a shift of isoform from 30 kDa in the untreated sample to 23 kDa in the treated sample. This confirms that the 40-45 kDa isoform contains complex oligosaccharides, the 30 kDa isoform contains mannose-rich oligosaccharides and the 23 kDa isoform is unglycosylated.

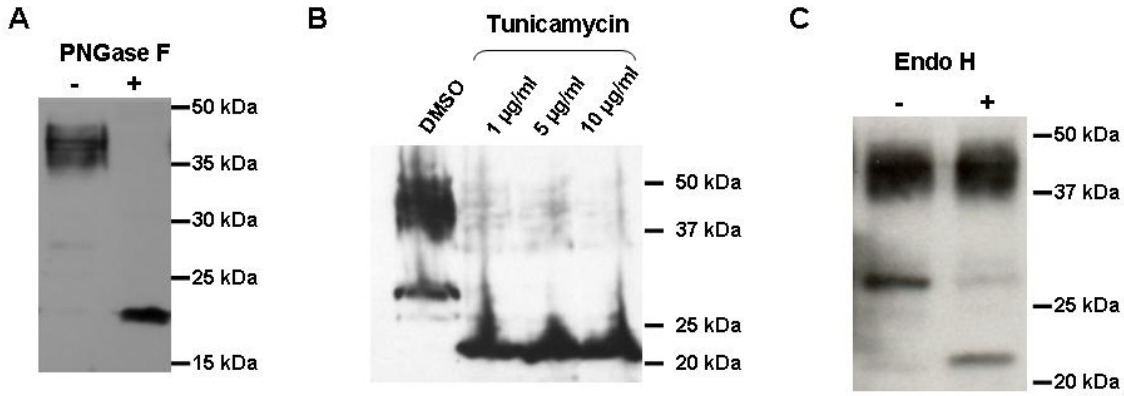


Figure 4.14. Determining the glycosylation state of mTmem114. MDCK II cells were transfected with V5-tagged Tmem114 were treated during culture (B) or post-lysis (A, C) to determine the presence of N-linked oligosaccharides. A) Enzymatic removal of all N-linked oligosaccharides using PNGase F. In untreated lysates (-) mTmem114-V5 is predominantly detected at 40-45 kDa, but in PNGase F treated lysates (+) mTmem114-V5 is detected at 23 kDa. B) Preventing the addition of N-linked oligosaccharides by treating cells with tunicamycin. In mock treated cells (DMSO) mTmem114-V5 was detected predominantly at 40-45 kDa, and to a lesser extent at 30 kDa. In cells treated with varying concentrations of tunicamycin mTmem114-V5 was detected at 23 kDa. C) Enzymatic removal of N-linked high mannose oligosaccharides using Endo H. In untreated lysates (-) mTmem114-V5 is predominantly detected at 40-45 kDa and at 30 kDa. In Endo H treated lysates, mTmem114-V5 is detected at 40-45 kDa and at 23 kDa.

4.5.2 Establishing the N-glycosylation sites of Tmem114

To confirm that the predicted N-glycosylation sites, p.N54 and p.N88, are those glycosylated, mutants were created by site-directed mutagenesis (SDM) to change the two asparagine residues to lysine residues. Both single mutants and a double mutant were created. When detected by western blot, the single mutants ran at a smaller mass than the wildtype: p.N54K at ~32 kDa and p.N88K at ~30 kDa (Figure 4.15A). The double mutant was detected at 23 kDa, the same mass as de-glycosylated Tmem114, confirming that it is unglycosylated. When all forms were treated with N-Glycosidase F they were detected at 23 kDa. This confirms that the increase in mass of wildtype mTmem114 is due to glycosylation at p.N54 and p.N88.

The single glycosylation mutants were detected in a number of isoforms (Figure 4.15A). To determine if the added oligosaccharides were mannose-rich or complex, lystates were digested with Endo H (Figure 4.15B). As in Figure 4.14C, the intermediate isoform of wildtype Tmem114 is sensitive to Endo H digestion. The intermediate isoforms of both p.N54K and p.N88K were sensitive to Endo H digestion, but the highest molecular mass isoforms were not. Compared to the wildtype, p.N54K and p.N88K had a greater proportion of Endo H sensitive to insensitive protein. A higher molecular mass isoform (40 kDa) was detected in the unglycosylated p.N54K N88K sample.

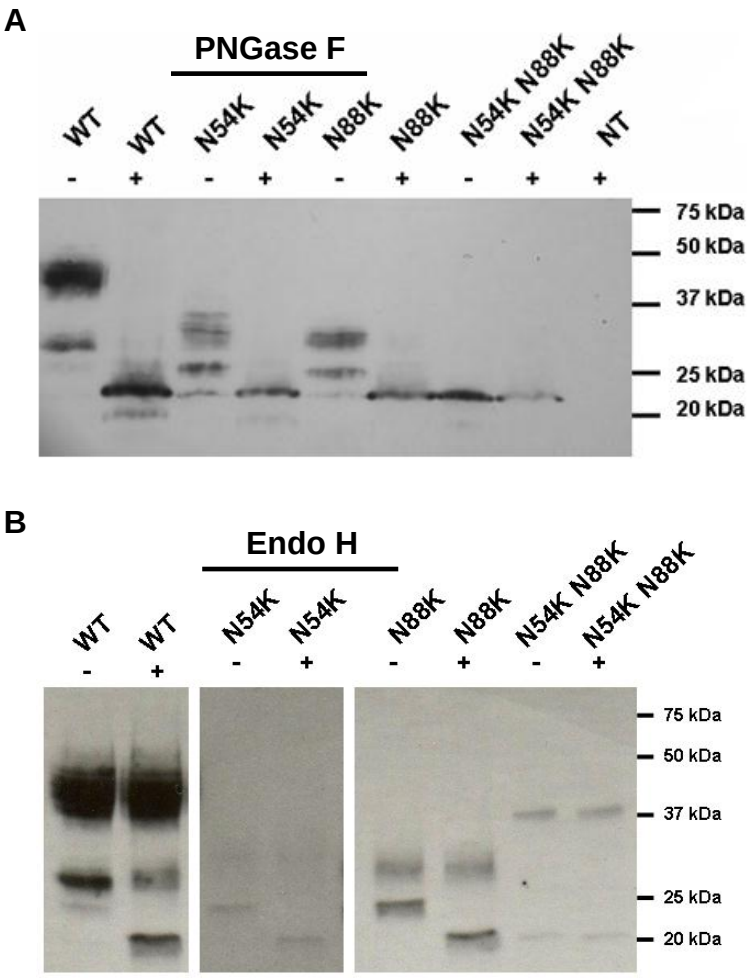


Figure 4.15. Identification of the glycosylated residues of mTmem114. MDCK II cells were transfected with wildtype or mutant V5-tagged Tmem114 and detected by western blot. A) Detection of V5-tagged glycosylation mutants by western blot and treatment with PNGase F. Single glycosylation mutants (p.N54K and p.N88K) were detected at a lower molecular mass than the wildtype (WT). Treating the wildtype and single glycosylation mutants with PNGase F (+) resulted in a shift to the 23 kDa isoform confirming the presence of N-linked oligosaccharides. Untreated (-) and PNGase F treated (+) double mutant (p.N54K N88K) protein were also detected at 23 kDa indicating it is not glycosylated. B) Sensitivity of the glycosylation mutants to Endo H. Treatment with Endo H resulted in a shift from the intermediate isoforms of wildtype, p.N54K and p.N88K to the unglycosylated isoform at 23 kDa. The highest molecular mass isoform was resistant to Endo H digestion.

4.5.3 Effect of glycosylation on the cellular localisation of mTmem114

As N-linked glycosylation may have a role in protein stability and transport the effect of the mutated sites on cellular localisation was investigated. V5-tagged WT and mutant constructs were expressed transiently for 48 hours in MDCK II cells on transwell inserts. Wildtype Tmem114-V5 localised to the lateral and apical cell membranes (Figure 4.16 Ai, Bi). The single glycosylation mutants (N54K and N88K) localised to the lateral cell membrane (Figure 4.16 Aii,iii) but not at the apical cell membrane (Figure 4.16 Bii,iii). Some cytoplasmic localisation was also detected (Figure 4.16 Aii,iii,Bii,iii). The double mutant N54K N88K failed to localise to the cell membrane and had a perinuclear appearance (Figure 4.16 Aiv,Biv).

The perinuclear localisation of the double mutant was suggestive of ER localisation. A number of attempts were made to immuno-label the ER but none of the commercially available antibodies tested labelled the ER in MDCK II cells, which are of canine origin.

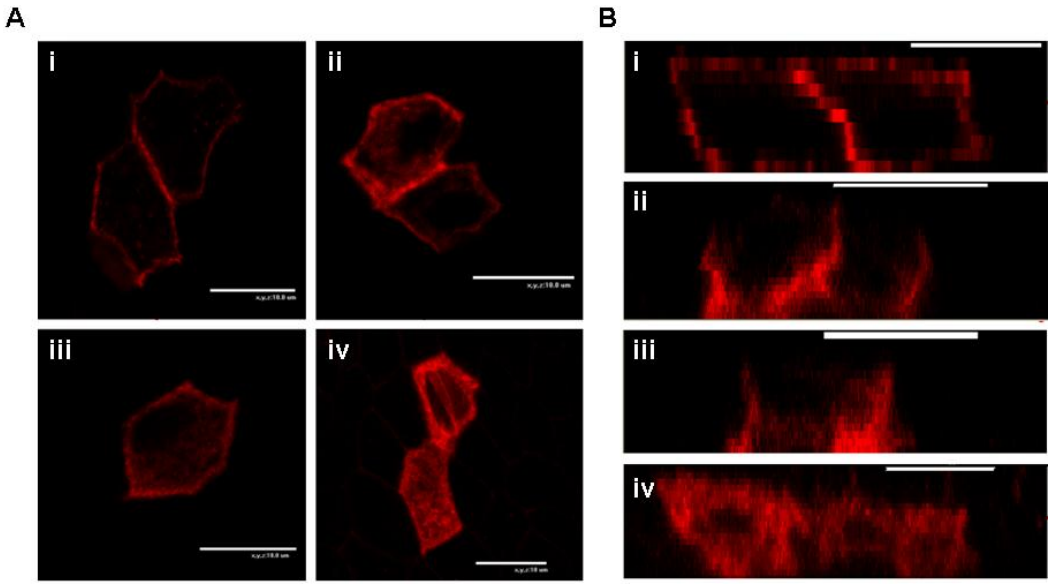


Figure 4.16. Localisation of V5-tagged wildtype and glycosylation mutants of mTmem114 in MDCK II cells. A) Planar view of labelled cells. B) Z-stack images displaying cross-section of labelled cells. WT Tmem114 localised at the lateral (Ai) and apical (Bi) cell membrane. The single mutants N54K (Aii, Bii) and N88K (Aiii, Biii) were detected at the lateral cell surface (A ii, iii) but not at the apical cell membrane (B ii, iii). The unglycosylated double mutant N54K N88K localised to the cytosol (Aiv, Biv). Bar = 10 µm. Cells were fixed with MetOH:Acetone, labelled and imaged by confocal microscopy.

4.6 Signal peptide

The mTmem114 protein contains a predicted signal peptide of 28 amino acids (Section 3.3.3). Signal peptides are typically cleaved after transport of the protein. To identify if the predicted signal peptide of mTmem114 is cleaved, a construct of mTmem114 was made in which the codons encoding the predicted signal peptide (amino acids 2-28) were deleted. This signal peptide mutant (SPM) was detected at a lower molecular mass than wildtype mTmem114 (Figure 4.17).

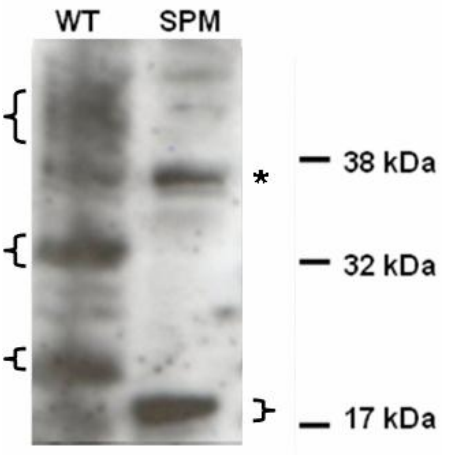


Figure 4.17. Detection of untagged wildtype and signal peptide deleted Tmem114 in MDCK II cells. Wildtype mTmem114 (WT) is detected as 3 bands that vary from 45 kDa to 23 kDa (braces at left hand side). Mutant mTmem114 lacking its signal peptide (SPM) is detected at a lower molecular mass than the WT (brace on right hand side). A non-specific band is detected at 36 kDa (asterisk).

4.7 Tissue expression of mTmem114

4.7.1 Mouse lens

After confirming that the antibody was specific to mTmem114 (Section 4.3.3), the presence of endogenous mTmem114 in adult mouse tissue was investigated. When the antibody was used to probe an adult mouse lens lysate, a single band of ~30 kDa was detected (Figure 4.18A). This band was not present in the blot incubated with the pre-immune bleed. The band ran at a different mass from exogenously expressed mTmem114, but was similar in size to the mannose-rich isoform of Tmem114.

To determine if the lens band represented a N-glycosylated form of mTmem114 that was modified with different oligosaccharides to those when the protein is processed by MDCK II cells, the adult lens fibre lysate was treated with PNGase F and probed with the anti-mTmem114 antibody by western blot. Although the exogenously expressed mTmem114 in MDCK II cells showed a reduction in molecular mass when treated, the molecular mass of the lens sample remained constant (Figure 4.18B), confirming that this band was not a glycosylated form of mTmem114.

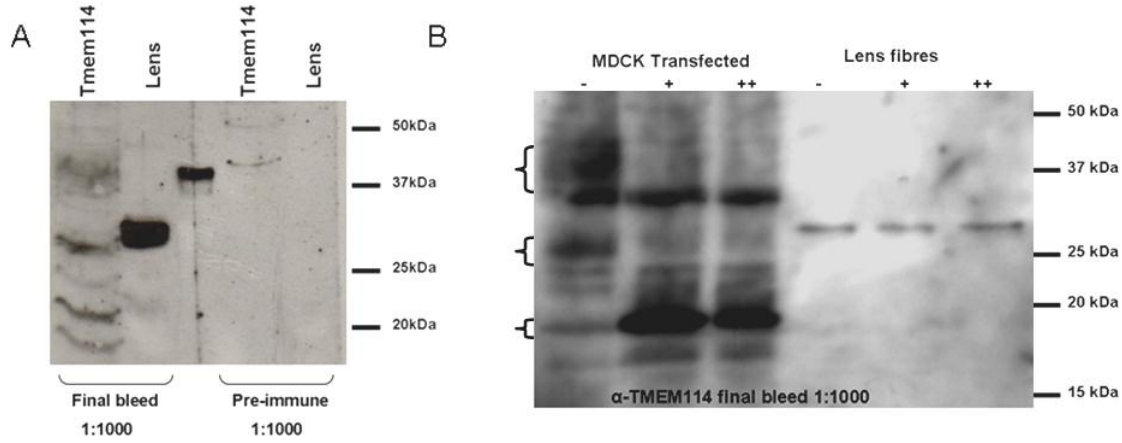


Figure 4.18. Expression of mTmem114 in mouse lens. A) MDCK II cells transfected with mTmem114 and whole mouse lens lysate were probed with the final bleed and pre-immune bleeds. A single band of approx 32 kDa was identified in the lens sample which was not present in the pre-immune. The band in the central lane is the 37 kDa band from the molecular marker. B) De-glycosylation assay. Lysates from transfected samples and mouse lens fibres were incubated with (+) and without (-) PNGase F. ++ represents the addition of double the recommended volume of PNGase F. Untreated transfected Tmem114 is detected in three isoforms (braces) at 40-45 kDa, 30 kDa and 23 kDa. PNGAse F treatment results in a shift to the 23 kDa isoform. No shift is detected for the band from the lens fibre lysate.

To determine if the lens fibre band was specific, the mTmem114 antisera was pre-incubated with the peptide antigen prior to using it to probe a western blot. Blocking the mTmem114 antibody only caused the bands from the transfected samples to disappear; the band in the lens sample remained at approximately the same intensity (Figure 4.19A). To determine if the

continued detection of a band in the lens sample was due to the high abundance of mTmem114 in this tissue, or too a high concentration of the antibody, a higher dilution of antibody and a greater amount of peptide was used (Figure 4.19B). Increasing the amount of the peptide ten-fold reduced, but did not ablate the signal from the 1:1000 antibody dilution. When the antibody was used at a dilution of 1:10,000, the band was still detected but when the peptide was added, the band was no longer detected. This suggested that the lens fibre band is specific for mTmem114.

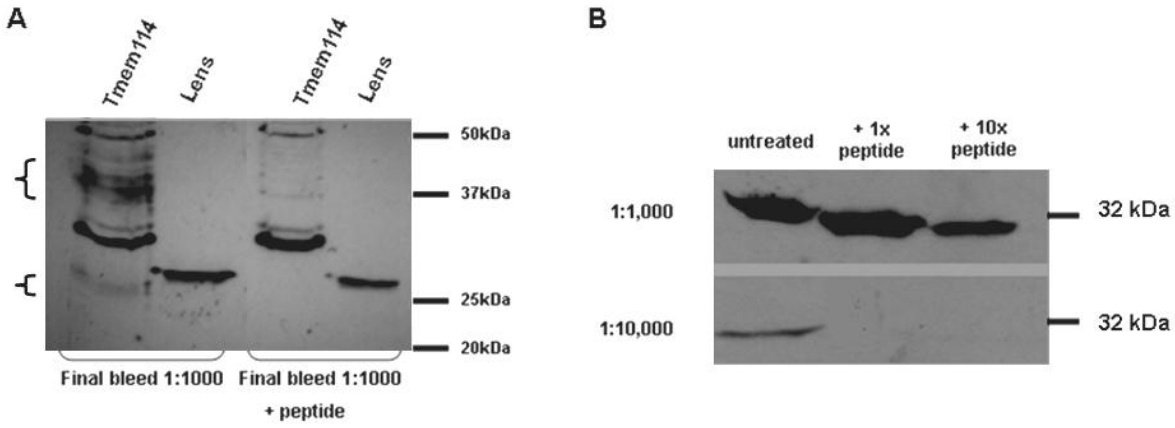


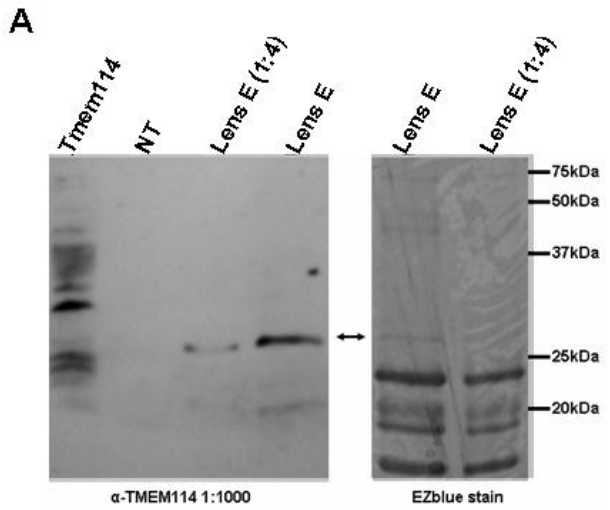
Figure 4.19. Pre-incubation of mTmem114 antibody against its antigen prior to western blot of mouse lens lysate. A) Pre-incubation of the mTmem114 antibody with its peptide antigen prevented the detection of mTmem114-specific bands (braces) in the MDCK II transfected sample, but the band in the lens sample remained. B) Increasing the peptide concentration to 10 x the suggested blocking concentration reduced the band intensity. Reducing the working antibody concentration to 1:10,000 resulted in the band still being detected. When the antibody at this dilution was pre-incubated with the peptide the band was not detected.

4.7.2 Mass spectrometry

In order to confirm the band detected in the lens sample was mTmem114 mass spectrometry analysis was performed. Two lens epithelium lysates were run on SDS-PAGE gels and the gel was divided in two: one half was stained with EZBlue (Coomassie Brilliant Blue G-250) and the other was western blotted with the anti-mTmem114 antibody (Figure 4.20A). The band of interest was identified in the EZBlue gel by aligning with the western blot, excised, and analysed by MALDI-TOF by the University of Manchester's core facility. mTmem114 was not detected in the MALDI-TOF analysis (Figure 4.20B).

To reduce the risk of contamination from the highly soluble crystalline proteins, purified membrane samples from bovine lens fibres (a gift from Professor Roy Quinlan, University of Durham) were used to perform the next mass spectrometry analysis. The anti-mTmem114 antibody detected a band of similar molecular mass to the mouse sample in three bovine lens purified samples; lens fibre cytoskeletal fraction, purified membranes from the lens cortex and purified membranes from the lens nucleus. The putative mTmem114 band was not detected

by the pre-immune serum. A band of greater molecular mass was also identified in the cortex membrane and cytoskeleton samples but as this band was also detected by the pre-immune serum it suggests it is non-specific. (Figure 4.21A). The two bands of approximately 33 kDa and 34 kDa detected in the cortex fibre sample were excised from the EZBlue gel and subjected to MALDI-TOF analysis by the University of Manchester's core facility (Figure 4.21B). Both bands were identified as Beta crystallin B1.



B

Identified Proteins	Uniprot Accession	Molecular Weight	Number of peptides
Beta-crystallin B1	CRBB1_MOUSE	28 kDa	13
Gamma-crystallin B	CRGB_MOUSE	21 kDa	4
Beta-crystallin B2	CRBB2_MOUSE	23 kDa	2

Figure 4.20. Mass spectrometry of putative mTmem114 band from mouse lens. A) Lysates from mouse lens epithelium were run on SDS-PAGE gels and EZBlue stained (right hand side) or detected by western blot (left hand side) The corresponding band detected by the anti-mTmem114 antibody (arrow) by western blot was extracted from the EZBlue gel and MALDI-TOF analysis was performed. B) MALDI-TOF results. The resulting peptides were matched against the Uniprot mouse database. Peptides from three crystalline proteins were identified.

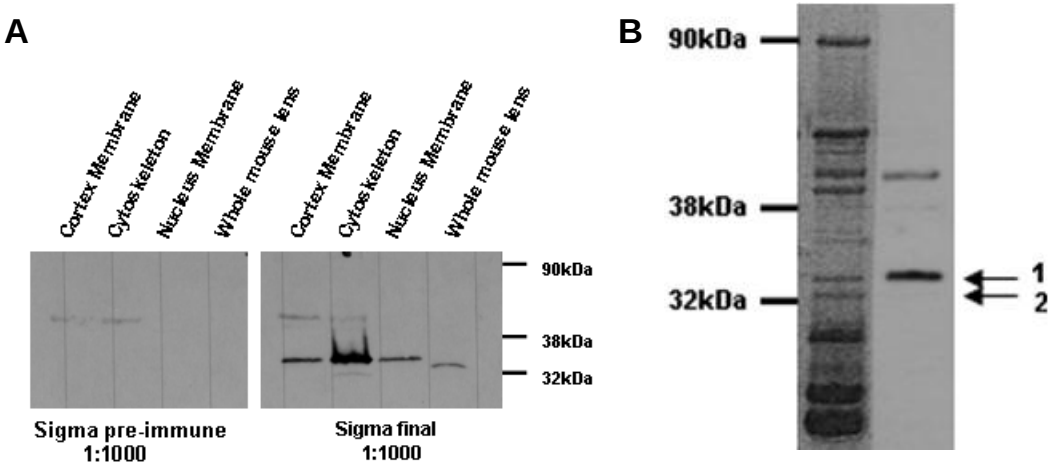


Figure 4.21. Western blot for Tmem114 in purified bovine lens fibre samples. A) A band of similar MW to that found in the mouse lens was identified in all three purified bovine lens fibre samples. A band of greater MW was also identified in the cortex membrane and cytoskeleton sample, but this band was also detected by the pre-immune bleed suggesting it is non-specific. B) MALDI-TOF analysis for two bands detected in the lens cortical fibre membrane sample. Bands were compared to a EZBlue stained gel of the same lysate. Both bands were identified as Beta crystallin B1.

A targeted approach was then performed using an orbitrap mass spectrometer at the Fingerprints Proteomics Facility at the University of Dundee. The sequence of mTmem114 was built into an inclusion list and this sequence was searched for in the list of all peptides detected in the submitted band from a mouse lens epithelium lysate as in Figure 4.20. The predominant results were for crystallins. mTmem114 was not detected in the sample (Figure 4.22).

Database	**:**	**IPI-mouse (56692 sequences; 25503219 residues)**
Timestamp	**:**	**18 Jun 2009 at 13:57:45 GMT**
Protein hits	**:**	IPI00126051.7 Tax_Id=10090 Gene_Symbol=Crybb1 Beta-crystallin B1
		IPI00222211.5 Tax_Id=10090 Gene_Symbol=Crybb2 Beta-crystallin B2
		IPI00109729.1 Tax_Id=10090 Gene_Symbol=Cryaa Isoform 2 of Alpha-crystallin A chain
		IPI00116498.1 Tax_Id=10090 Gene_Symbol=Ywhaz 14-3-3 protein zeta/delta
		IPI00134845.1 Tax_Id=10090 Gene_Symbol=Cryba1 Crystallin, beta A1
		IPI00122270.3 Tax_Id=10090 Gene_Symbol=Crybb3 Beta-crystallin B3
		IPI00138274.1 Tax_Id=10090 Gene_Symbol=Cryab Alpha-crystallin B chain
		IPI00227782.5 Tax_Id=10090 Gene_Symbol=Cryba4 Beta-crystallin A4
		IPI00230707.6 Tax_Id=10090 Gene_Symbol=Ywhag 14-3-3 protein gamma
		IPI00408378.4 Tax_Id=10090 Gene_Symbol=Ywhaq Isoform 1 of 14-3-3 protein theta
		IPI00130922.3 Tax_Id=10090 Gene_Symbol=Crygs Beta-crystallin S
		IPI00230682.6 Tax_Id=10090 Gene_Symbol=Ywhab Isoform Long of 14-3-3 protein beta/alpha
		IPI00755181.1 Tax_Id=10090 Gene_Symbol=Krt10 keratin complex 1, acidic, gene 10
		IPI00139301.3 Tax_Id=10090 Gene_Symbol=Krt5 Keratin, type II cytoskeletal 5
		IPI00347110.2 Tax_Id=10090 Gene_Symbol=Krt73 Keratin, type II cytoskeletal 73
		IPI00622240.4 Tax_Id=10090 Gene_Symbol=Krt2 Keratin, type II cytoskeletal 2 epidermal
		IPI00625729.2 Tax_Id=10090 Gene_Symbol=Krt1 Keratin, type II cytoskeletal 1
		IPI00227849.5 Tax_Id=10090 Gene_Symbol=Cryba2 Beta-crystallin A2
		IPI00132096.3 Tax_Id=10090 Gene_Symbol=Mtap S-methyl-5'-thioadenosine phosphorylase
		IPI00331092.7 Tax_Id=10090 Gene_Symbol=Rps4x 40S ribosomal protein S4, X isoform
		IPI00314736.3 Tax_Id=10090 Gene_Symbol=Anp32a Acidic leucine-rich nuclear phosphoprotein 32 family member A
		IPI00110827.1 Tax_Id=10090 Gene_Symbol=Acta1 Actin, alpha skeletal muscle
		IPI00121534.11 Tax_Id=10090 Gene_Symbol=Car2 Carbonic anhydrase 2
		IPI00346834.1 Tax_Id=10090 Gene_Symbol=Krt76 Keratin, type II cytoskeletal 2 oral
		IPI00277001.4 Tax_Id=10090 Gene_Symbol=Psma4 Proteasome subunit alpha type-4
		IPI00114285.1 Tax_Id=10090 Gene_Symbol=Gsto1 Glutathione S-transferase omega-1
		IPI00118821.2 Tax_Id=10090 Gene_Symbol=Pafah1b2 Platelet-activating factor acetylhydrolase IB subunit beta
		IPI00110864.3 Tax_Id=10090 Gene_Symbol=Crygf Gamma-crystallin F
		IPI00336881.1 Tax_Id=10090 Gene_Symbol=Ddah2 N(G),N(G)-dimethylarginine dimethylaminohydrolase 2
		IPI00115650.4 Tax_Id=10090 Gene_Symbol=Cacybp Calcyclin-binding protein
		IPI00124223.3 Tax_Id=10090 Gene_Symbol=Psme1 Proteasome activator complex subunit 1
		IPI00227522.2 Tax_Id=10090 Gene_Symbol=Fcgbp Putative uncharacterized protein
		IPI00348328.4 Tax_Id=10090 Gene_Symbol=Krt78 keratin Kb40
		IPI00109918.1 Tax_Id=10090 Gene_Symbol=Olfr703 T4 olfactory receptor
		IPI00331644.5 Tax_Id=10090 Gene_Symbol=Psma3 Proteasome subunit alpha type-3

Figure 4.22. Mass spectrometry of putative mTmem114 band isolated from mouse lens epithelium sample using an Obitrap mass spectrometer. Orbitrap results from submitted band using the International Protein Index (IPI) database

4.7.3 Endogenous localisation of mTmem114

The use of the anti-mTmem114 antibody in western blotting against mouse tissue did not identify mTmem114, however this may have been due to low levels of expression and/or expression in a subset of cells. To test this hypothesis the antibody was used for immunohistochemistry (IHC) on sections of paraffin-embedded mouse eyes and on mouse lens cryosections. Labelling of the lens epithelium was predominantly in the apical membrane, with some lateral staining, where it colocalised with F-actin (Figure 4.23Ai). Labelling was also observed in the both the cortical and nuclear lens fibres (Figure 4.23Aii). However, similar labelling patterns were observed with the pre-immune bleed (Figure 4.23B).

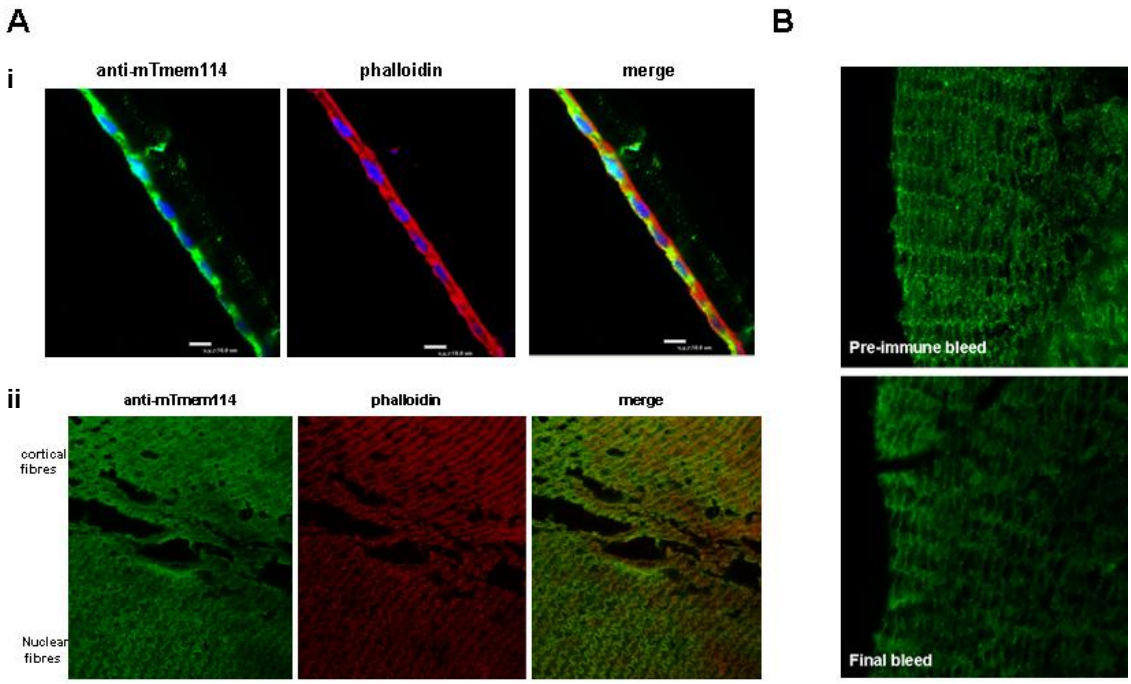


Figure 4.23. Anti-mTmem114 immunohistochemical labelling of adult mouse lens. A) labelling of adult mouse lens with mTmem114 antibody and rhodamine-labelled phalloidin. i) In an isolated lens epithelium sample, F-actin (red) was localised at all cell membranes. Tmem114 labelling (green) was present at the apical and lateral membrane, but was not present at the basal membrane. ii) In both cortical and nuclear fibres, labelling for Tmem114 (green) was similar to that of F-actin (red). B. Labelling of cortical lens fibres with pre-immune bleed and anti-mTmem114 final bleed.

4.8 Discussion

4.8.1 mRNA expression

Expression of *TMEM114* was not detected in the human embryonic kidney (HEK) cell line or in the canine MDCK II cell line which are of kidney epithelium origin. Expression of *TMEM114* was also not detected in the adult lens epithelial cell line FHL124. This was unexpected as the properties of FHL124 cells are considered very similar to those of freshly obtained lens epithelial cells (Lauf *et al.* 2008) and *Tmem114* is expressed in the mouse lens epithelium (Jamieson *et al*. 2007). The morphology of this cell line does not reflect that of *in vivo* lens epithelium, suggesting that some changes have occurred during establishment of the cell line. Thus, *TMEM114* expression may have been lost during establishment of the FHL124 cell line. Alternatively, the specific cells from which the cell line developed did not express *TMEM114*. *TMEM114* expression was detected in the developing human eye (Figure 4.1) and the expression was similar to that of murine *Tmem114* (Figure 1.6) (Jamieson *et al*. 2007). Expression of *TMEM114* was detected at CS21 (equivalent to murine stage E14), which correlates with onset of expression at E13.5 in mouse. As was identified in mouse, expression of *TMEM114* increased with development. In mouse there is a marked increase in *Tmem114* expression postnatally (Jamieson *et al*. 2007). Postnatal human samples were not available for testing so expression in the adult has not been confirmed.

4.8.2 Custom antibody testing

Two custom polyclonal antibody projects were undertaken using the 14 amino acids of the C-terminus of murine Tmem114 as the epitope. Although one antibody (#4672) failed to detect transfected mTmem114, the other (#4671) recognised transfected mTmem114 (Figures 4.4, 4.5). The bands detected by the antibody were equivalent to those detected by an anti-V5 antibody used against cells transfected with mTmem114-V5. The specificity of the antibody was shown by pre-incubating the anti-sera against the peptide antigen, and when this was done the antibody failed to detect the specific mTmem114 bands on the western blot (Figure 4.7). The non-specific bands remained, confirming that the peptide specifically inhibited detection of mTmem114.

The purified antibody, although functional, required high concentrations (1:50 for western blot and 1:20 for immunofluorescence). The antibody production schedule was protracted (Section 2.6.1.1) and the titres of the final serum and purified antibody were significantly lower than the initial sera according to the ELISA results (Figure 4.6). The final bleed was taken 7 days post-immunisation, short of the recommended 10-14 days (Harlow and Lane 1988). This sera however still proved functional at a dilution of 1:1000. Thus it was unexpected that the purified antibody was very weak. The use of an elution buffer that was too severe may have been responsible for this, or the fact that the purity of the peptide used was poor (Section 2.6.1.1) may have affected the affinity purification.

In the V5-tagged construct of mTmem114 the V5-epitope was positioned at the C-terminus of Tmem114, immediately downstream of the antibody epitope. The Tmem114 antibody failed to detect mTmem114-V5 suggesting that the V5 epitope blocked access of the antibody to the mTmem114 epitope. This further suggests that the antibody is specific to the C-terminus of mTmem114. This specificity should allow differential labeling of tagged and untagged Tmem114. However, although the antibody could detect mTmem114 when heterologously overexpressed, it failed to detect mTmem114 in the adult lens. Non-specific labeling was observed when the antibody was used against adult mouse tissue. This non-specificity may have masked the detection of low expression levels of mTmem114.

4.8.3 Glycosylation of mTmem114

N-linked glycosylation of proteins can serve many purposes including protein folding and maintenance of correct conformation, intracellular trafficking and cell surface expression (Helenius 1994; Gahmberg and Tolvanen 1996). N-linked glycosylation of extracellular sites of membrane proteins occurs in a co-translational fashion in the lumen of the ER (Weerapana and Imperiali 2006). In the ER lumen, the core oligosaccharide, composed of glucose (Glc), mannose (Man) and N-acetyl glucosamine (GlcNAc) ($Glc_3Man_9GlcNAc_2$) is added to asparagine residues in the consenus motif N-X-S/T (Helenius and Aebi 2004). Removal of two of the three glucoses by glucosidases I and II occurs almost immediately, leaving a single glucose on the glycoprotein (Trombetta 2003). The ER resident chaperones calnexin and

calreticulin interact only with the monoglucosylated form of the glycoprotein (Zapun *et al.* 1997; Vassilakos *et al.* 1998). The interaction with calnexin and calreticuin is terminated by cleavage of the remaining glucose by glucosidase II. If the released glycoprotein is correctly folded, it exits the ER. However, if after release from the chaperones, the glycoprotein has not adopted its native conformation, a single glucose is added back to the oligosaccharide by UDP-glucose:glycoprotein glucosyltransferase (UGGT1), facilitating binding of calnexin and calreticulin (Ritter *et al.* 2005). This cycle of binding, release and re-binding continues and when the protein is correctly folded it exits the ER (Ellgaard and Helenius 2003). These proteins are then exported to the Golgi apparatus where the majority of glycoproteins undergo further processing of their sugar residues (Vagin *et al*. 2009). The mannose-based oligosaccharides, which are sensitive to Endo H digestion, are removed and replaced with more complex oligosaccharides which are Endo H insensitive (Helenius and Aebi 2004). However, if after multiple cycles the glycoprotein remains misfolded, a single mannose is cleaved by ER α1,2-mannosidase I which promotes association with ER degradation-enhancing 1,2-mannosidase-like protein (EDEM) (Hosokawa *et al.* 2001), triggering ER-associated degradation (ERAD) of the misfolded protein (Molinari *et al.* 2003; Oda *et al.* 2003).

When mTmem114 was detected by western blot three main isoforms were identified. A diffuse band at 40-45 kDa was sensitive to PNGase F but not Endo H digestion (Figure 4.13A,C) suggesting it represented the mature isoform with complex carbohydrates. An isoform at ~30 kDa, which when resolved further appeared as two bands, was sensitive to Endo H digestion suggesting it contained mannose oligosaccharides (Figure 4.13C). This represents an intermediate isoform in the endoplasmic reticulum before export. A third isoform of ~23 kDa was approximately the same size as the predicted molecular mass of mTmem114. Treating mTmem114 with de-glycosylating enzymes increased the intensity of this band suggesting it represents the unglycosylated form of the protein. As the addition of oligosaccharides is a co-translational event, the detection of Tmem114 in an unglycosylated state suggests that the protein was not being processed correctly by the MDCK II cells. This is most likely due to the high levels of overexpression in the transiently transfected cells, as the unglycosylated isoform was not present in the stable cell lines which had moderate levels of expression of mTmem114 (Figure 4.10A).

The glycosylated sites in Tmem114 were identified by mutating the predicted glycosylation sites by SDM. The single mutants (N54K and N88K) were detected in three isoforms but had a reduced molecular mass compared to the wildtype (Figure 4.15). The double mutant was detected at ~23 kDa, the same mass as the de-glycosylated wildtype protein, confirming that this mutant is unglycosylated.

4.8.4 Localisation of mTmem114

The localisation of the predicted transmembrane protein mTmem114 at the plasma membrane was confirmed by immunofluorescence *in vitro*. Membrane localisation was found to be cell-type dependent. When heterologously expressed in the lens epithelial cell line FHL124, mTmem114-V5 was detected in the cytoplasm, with strong perinuclear staining (Figure 4.8). No membrane localisation was observed. However, as previously mentioned, the morphology of this cell line does not reflect that of *in vivo* lens epithelium, and this may have affected localisation. Other lens epithelial cell lines could have been used to detect localisation. HLE-B3 cells (ATCC CRL-11421) were derived from an infant's lens and although they display an epithelial morphology, they fail to polarise (Andley *et al.* 1994). A physiologically relevant cell model for studying mTmem114 localisation requires the cell line to be polarised. Using rat lens epithelial explants, Stump and colleagues (2006) showed that treating non-polarised lens epithelial cells with lithium chloride (LiCl) promoted polarization. However, the ability to transfect these cells was not shown.

Polarised cells such as MDCK II cells are recommended for studying the localisation of membrane proteins (van Beest *et al.* 2006). Polarised cells have distinct apical and basolateral domains, demarcated by tight junctions. During polarisation, adherens junctions form between opposing cells (Yonemura *et al.* 1995). ZO-1 is recruited to this contact site where it dimerises and functions to polymerise claudins (Itoh *et al*. 1999; Umeda *et al*. 2006). This polymerisation is necessary for tight junction formation (Umeda *et al*. 2006). The polymerised claudins and ZO proteins then form a distinct junctional complex, the tight junction, apical to the cadherin-based adherens junction. Once tight junctions are formed, ZO-1 localises specifically to the junction (Miyoshi and Takai 2005). Thus, the localisation of ZO-1 distinct and apical to that of E-cadherin can serve as a marker for tight junctions. When MDCK II cells were cultured on plastic surfaces, ZO-1 colocalised with E-cadherin at the lateral border, indicating that tight junctions had not formed (Figure 4.11). However, by culturing the cells on the transwell porous filters, which substitute the extracellular matrix on which polarised epithelial cells lie *in vivo* (Mostov *et al.* 2003), the MDCK II cells polarise and ZO-1 localises to the tight junction, apical to the adherens junction (Figure 4.11).

TMEM114 shows homology to proteins which localise to tight junctions in MDCK II cells and *in situ* (Section 3.10.2.6) (Notterpek *et al.* 2001; Ruffer and Gerke 2004; Sugiyama *et al.* 2008). Therefore, localisation of Tmem114 to the tight junction may have been predicted. However mTmem114 localised to the lateral and apical cell membranes and did not colocalise with the tight junction marker ZO-1 (Figure 4.12A). Both tagged and untagged mTmem114 displayed the same localisation, confirming both the specificity of the custom antibody and that the V5 epitope does not affect localisation of the protein (Figure 4.12B,C). The cytoplasmic C-termini of claudins, which includes the PDZ binding motif which interacts with ZO-1, target and stabilize localization at the tight junction (Itoh *et al*. 1999; Ruffer and Gerke 2004). Loss of the C-terminus results in intracellular retention (Ruffer and Gerke 2004). The C-terminus of

TMEM114 is shorter than those of claudins and shows little similarity to them (Figure 3.7). The V5-tag was attached to the final C-terminal amino acid and the western blot data suggests that the V5 epitope blocks the anti-mTmem114 antibody epitope of Tmem114 which also located at the C-terminal of the protein. As both the tagged and untagged proteins have the same localisation this suggests that the amino acids in the C-terminal epitope are not involved in membrane localisation.

Although lateral membrane localisation of mTmem114 was detected, expression was predominantly localised to the apical cell membrane (Figure 4.12B). The apical membrane of epithelia is made up of a base membrane and protruding microvilli (Roper *et al.* 2000; Giepmans and van Ijzendoorn 2009). To determine if mTmem114 specifically localised to either membrane segment, immunoelectron microscopy using nano-gold labelled anti-mTmem114 antibody was performed in MDCK II cells stably expressing mTmem114. Although labelling was observed in the microvilli, with some apical cytoplasmic labelling, background levels of labelling were too high for the results to be unambiguous.

Apical localisation of proteins is associated with a number of motifs which typically lie in the extracellular domains or in transmembrane domains (Rodriguez-Boulan and Musch 2005; Cullinane *et al.* 2010). The addition of N-linked (Scheiffele *et al.* 1995) and O-linked oligosaccharides (Alfalah *et al.* 1999) and GPI anchors (Sharom and Lehto 2002) are associated with apical trafficking. mTmem114 contains two N-linked glycosylation sites (Section 4.4.2) and the absence of N-linked oligosaccharides at both sites resulted in intracellular localisation, possibly due to misfolding and retention in the ER. A blot of the unglycosylated mutant identified a second band at 40 kDa which may represent a dimer of the misfolded protein (Figure 4.15B).

The presence of oligosaccharides at a single site resulted in lateral membrane localisation, but the presence of oligosaccharides at both sites was necessary for apical membrane localisation. The single glycosylation mutant proteins were also detected in the cytoplasm (Figure 4.16) and the single mutants had a higher proportion of Endo H sensitive isoforms than wildtype Tmem114 (Figure 4.15B). Endo H sensitivity is an indicator of ER localisation. Thus, although some of the single glycosylated proteins were reaching the plasma membrane, a significant proportion were present in the ER. This implies that both glycosylation sites are necessary for efficient membrane trafficking of Tmem114, supported by the fact that both glycosylation sites are evolutionarily conserved. Only when both sites are glycosylated does Tmem114 localise to the apical cell membrane suggesting that the apical localisation is functionally significant.

The NHS-A protein is currently the only protein localised at tight junctions that is associated with congenital cataract (Sharma *et al*. 2006). Although the localisation of mTmem114 to tight junctions was not detected, its localisation suggests it may be part of the apical junctional

complex. *mTmem114* is expressed in cells of the central lens epithelium and epithelial cells at the lens equator (Jamieson *et al*. 2007). Due to the apical localisation of mTmem114 in the polarised MDCK II cells a similar localisation of mTmem114 in the lens epithelium was postulated. This was investigated using the custom polyclonal antibody to probe sectioned adult mouse lenses. However, although labelling was observed with the antibody, similar labelling was also observed with the pre-immune serum (Figure 4.22B). To determine if this was due to non-specific binding of the fluorescently labelled secondary antibody, or autofluorescence, a rabbit antibody isotype control was used in place of the primary antibody. No staining was observed with this isotype control confirming that the staining was not due to the secondary antibody or due to autofluorescence. Thus the observed labelling was likely to have been due to non-specific binding of an antibody present in the serum.

4.8.5 Signal peptide of Tmem114

Signal peptides are typically 20-30 amino acids in length and show little sequence conservation (Hegde and Bernstein 2006). There is a conserved structure to signal peptides, with a basic N-terminal domain (N region), followed by a hydrophobic domain (H region) and a slightly polar C region which contains the cleavage site (Nielsen and Krogh 1998). Signal peptides of nascent polypeptides emerging from the ribosome are recognised by a signal recognition particle (SRP) and the rate of translation is slowed until the nascent polypeptide/ribosome complex is targeted to the ER (Jiang *et al.* 2008). The growing polypeptide chain enters the lipid bilayer of the ER membrane via the translocon and the signal peptide is cleaved.

Transmembrane domains are hydrophobic regions, which when present near the N-terminus of a protein, may be mis-annotated as signal peptides (Reynolds *et al*. 2008). These domains are recongnised by SRP for targeting to the ER (Keenan *et al.* 2001). In silico predictions using Philius (Section 3.3.3), which is designed specifically to overcome such mis-annotation of signal peptides in transmembrane proteins, suggested that Tmem114 has a signal peptide, although confidence in the prediction was only 66 %. To determine if the signal peptide of Tmem114 is cleaved, Tmem114 lacking its signal peptide was cloned. When both wildtype and signal peptide mutant were detected by western blot, the signal peptide mutant was detected at a lower molecular mass than the wildtype confirming that the signal peptide is not cleaved. Thus, Tmem114 retains TMD1 and has a tetraspan topology similar to the other Pfam00822 proteins.

4.9 Conclusion

Although this study has confirmed that Tmem114 is a transmembrane protein and showed that N-linked glycosylation is required for localisation at the membrane, the functional role of Tmem114 at the cell membrane remains to be elucidated. It is clear that Tmem114 is not a typical claudin as in polarised MDCK II cells it does not localise to the tight junction. This

was not unexpected as the C-terminus of Tmem114 shows no similarity to the C-termini of claudins, which are necessary for tight junction localisation (Ruffer and Gerke 2004; Umeda *et al*. 2006)

Although Tmem114 did not localise to the tight junction, by localising to the lateral cell membrane it may have a role in barrier formation between cells. The effect of a claudin on barrier function can be determined by overexpressing or knocking down the expression of the claudin of interest in MDCK II cells on transwell membranes. By measuring the transepithelial resistance of the cell monolayer one can determine the effect of expression of the claudin on paracellular permeability (Furuse *et al*. 2001). The effect of mTmem114 on transepithelial resistance was not measured as the stable cell lines that were created were not clonal. This would have affected the integrity of the barrier as the monolayer would be constituted of Tmem114 expressing and non-expressing cells.

What is possible to conclude is that Tmem114 and other members of Pfam00822 are expressed in polarised tissues. Although distinct in function, neurons (Hung *et al.* 2007) and epithelial cells (Horne-Badovinac and Bilder 2008) are both polarised. The lens develops from a single layer of polarised epithelial cells – the lens vesicle. In the lens the epithelial cells are polarised and the differentiated fibres retain this polarity (Lo *et al.* 2000; Zampighi *et al*. 2000). Expression of *Tmem114* occurs in both the central and equatorial epithelium of the lens (Jamieson *et al*. 2007). Epithelial cells at above the equator, where *Tmem114* expression is highest (Jamieson *et al*. 2007) lack tight junctions and are in contact with newly differentiating fibre cells at the epithelial-fibre interface (Zampighi *et al*. 2000).

At the epithelial-fibre interface, the apical membranes of the differentiating fibres are in contact with the apical membranes of the lens epithelial cells (Figure 1.3) (Kuszak *et al.* 1995). The apical and lateral localisation of Tmem114 in a model polarised epithelial cell line and the expression of *Tmem114* in the epithelial cells of this region is suggestive of a role in the epithelial fibre interface. Posterior to the epithelial-fibre interface lies the modiolus or lens fulcrum, a structure formed from the joining of the apical ends of newly differentiating epithelial cells which have begun to elongate (Figure 1.3) (Zampighi *et al*. 2000; Bassnett 2005). The elongating cells also express *Tmem114* (Jamieson *et al*. 2007) suggesting that Tmem114 may have a role in the formation of the lens fulcrum. These hypotheses could not be tested as the antibody generated did not function *in situ*.

If Tmem114 functions in the epithelial-fibre interface and/or the lens fulcrum, it may require the adhesive properties of other members of the Pfam00822 family. Claudins, EMP-1, MP20 and the more distantly related calcium channel γ subunits contain the W-GLW-C-C motif and have the ability to act as adhesion molecules in mouse fibroblast L cells (Section 3.10.2.6). Unfortunately, due to time constraints, the adhesive capability of Tmem114 in L cells was not tested. If adhesive properties are shown it would suggest a possible role in adhesion

between lens epithelial and fibre cells. The localisation of Tmem114 in the correct positions of the lens would support this. Confirmation of the localisation would require the production of an antibody that functioned *in situ.* The functional role of Tmem114 *in vivo* may be identified by creating an animal model. As of October 2010, no knockout mouse is available, but Tmem114 knockout embryonic stem cell colonies have been created (http://www.knockoutmouse.org/genedetails/MGI:1921970).

Tmem114 is also expressed in the cerebellum (Jamieson *et al*. 2007), where the calcium channel γ2 subunit stargazin is expressed (Tomita *et al*. 2003). Stargazin acts as a TARP, mediating the membrane delivery of AMPA receptors (Chen *et al*. 2000). Delivery of AMPA receptors to the synapse is dependent of PDZ binding domains (Tomita *et al.* 2005). Although Tmem114 lacks the PDZ binding domains found in claudins and TARPs (Section 3.10.2.5), as it has relatively high levels of amino acid identity with the TARPs (Table 3.2), it may have the ability to traffic proteins to the membrane. To determine if Tmem114 functions to traffic proteins, pull-down experiments using lysates from lenses or cerebella could be performed.

Chapter 5: Expression and localisation of Tmlp1

5.1 Introduction

The in silico analysis performed in chapter 3 suggests that TMLP1 is a transmembrane protein expressed in the brain. However, to date, no experimental analysis has been performed on this novel gene. There were two main aims in this chapter: to determine the expression pattern of murine *Tmlp1* at the mRNA level by RT-PCR, and to determine the tissue specific expression as well as the cellular localisation of the predicted transmembrane protein Tmlp1 using custom anti-Tmlp1 polyclonal antibodies.

5.2 Expression of Tmlp1

5.2.1 *Tmlp1* expression in mouse

RT-PCR was used to determine the expression pattern of *Tmlp1* in the adult mouse. Expression of *Tmlp1* was detected in the brain, eye, lung, stomach and kidney, but was not detected in the liver or intestine (Figure 5.1A). *Tmlp1* is also expressed in the developing brain and developing eye (Figure 5.1B). Expression in the eye was found to increase at later developmental stages, reaching a peak during adulthood. In the adult lens, expression is highest in the lens epithelium, with lower levels detected in the lens fibres.

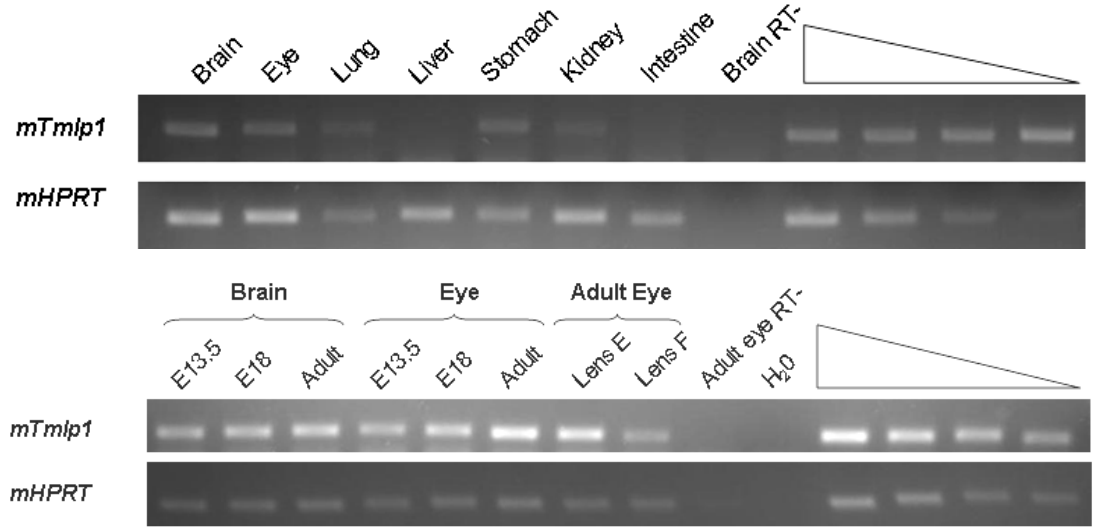


Figure 5.1. Expression profile of murine Tmlp1 detected by RT-PCR. Expression of *mTmlp1* was detected by amplifying cDNA with primers amplying from exon 1 to exon 4.A) Expression in the adult mouse. B) Expression pattern in the developing and adult brain and eye. Detection of the housekeeping gene *Hprt* was used as a control for equal cDNA loading in the RT-PCR reactions. The endpoint of all PCRs was determined to be within the exponential phase as demonstrated by the linearities for each reaction (triangle). 'RT-' are samples to which no reverse transcriptase was added. H_20 is a no template control.

5.2.2 *TMLP1* expression in human development

The expression pattern of *TMLP1* in a select panel of human embryonic eye and neural tissue was determined by RT-PCR. *TMLP1* expression was detected in the hindbrain at CS21

(Figure 5.2). *TMLP1* expression was also detected in the eye at foetal stage 2 (F2), but not at the earlier stage of Carnegie Stage 21 (CS21). In the spinal cord, expression was detected at F2 but not at the earlier stages of CS21 and F1. The primers used to amplify *TMLP1* encompass exon 1 to exon 4. Three different isoforms were detected in the different tissues. Sequencing of the amplicons confirmed that they were alternatively spliced isoforms of *TMLP1*. In the eye, the full-length isoform and an isoform which lacked exons 2 and 3 were expressed. This shorter isoform was also expressed in the hindbrain. The F2 spinal cord had weak expression of an isoform that lacked exon 2 only. The alternatively spliced isoforms produce in-frame shortened isoforms of TMLP1 with alternative topologies (Figure 5.3).

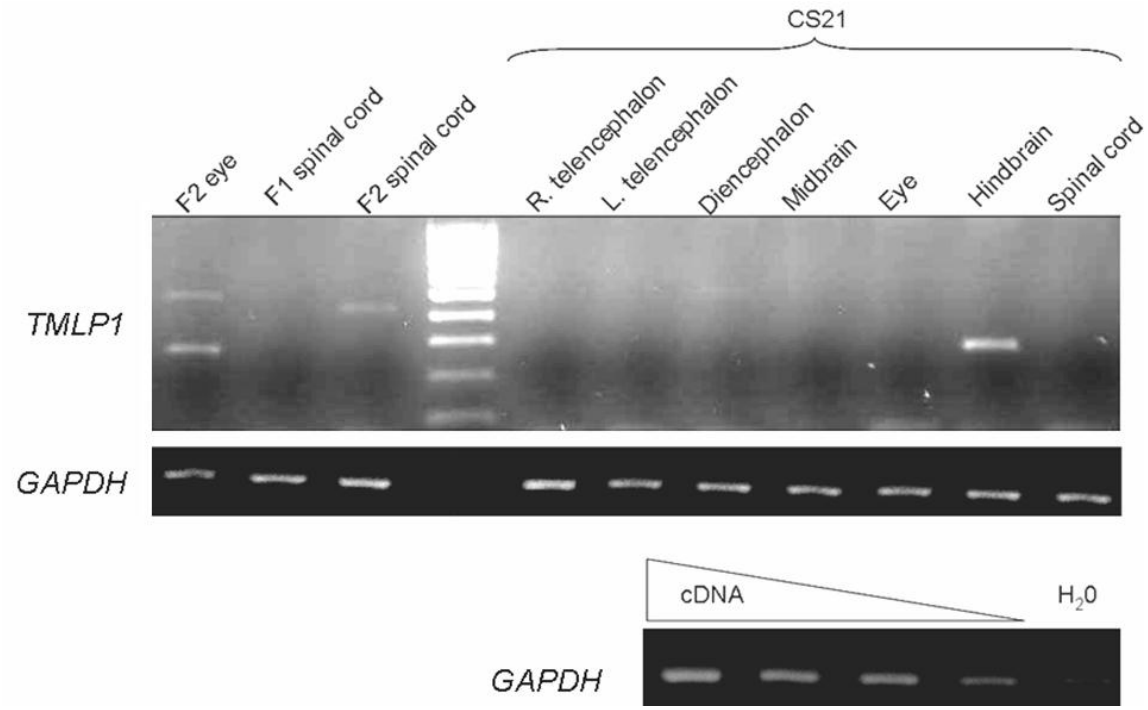


Figure 5.2. Expression profile of TMLP1 in developing eye and neural tissues. Expression of *TMLP1* was detected by amplifying cDNA with primers amplifying from exon 1 to exon 4. TMLP1 was expressed in the hindbrain at Carnegie stage 21 (CS21) in a splice isoform which lacks exons two and three. At Foetal stage two (F2) *TMLP1* was expressed in the eye and spinal cord. Expression in the spinal cord was weak and was restricted to a splice isoform which lacks exon 2. In the eye *TMLP1* was in two isoforms: full length *TMLP1* and *TMLP1* which lacked exons 2 and 3. Detection of the housekeeping gene *GAPDH* was used as a control for equal cDNA loading in the RT-PCR reactions. The endpoint of all PCRs was determined to be within the exponential phase as demonstrated by the linearities for each reaction (triangle). H_20 is a no template control. Lane 4 is the DNA marker Hyperladder IV. Bands are of 100 bp increments (100-1000 bp).

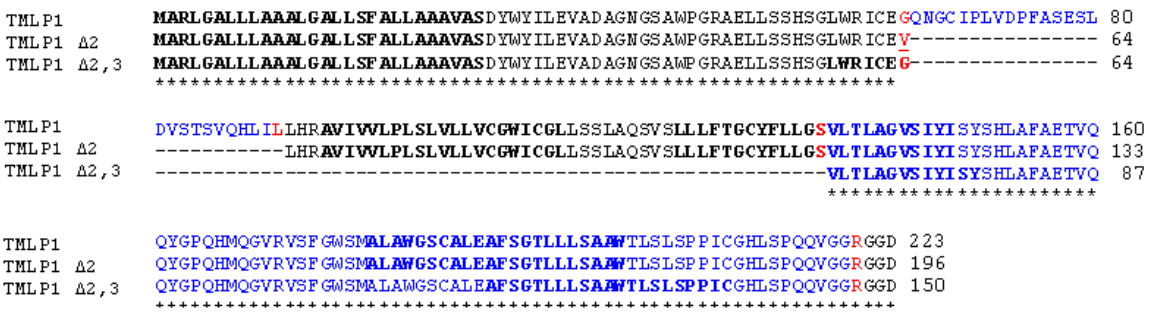


Figure 5.3. Predicted topologies of full-length TMLP1 and the splice variants lacking exon two or exons two and three. The splice isoform lacking exon 2 (TMLP1 Δ2) maintains the same topology as full-length TMLP1 but the altered splice site results in the incorporation of a valine (V) (underlined) in place of glycine (G) at p.64. The isoform lacking exons 2 and 3 (TMLP1 Δ2,3) encodes a protein 73 amino acids shorter than the predicted full length protein. This isoform is predicted to have three transmembrane domains (bold). Alternate exons are in black and blue. Amino acids encoded by overlapping splice are coloured red.

5.2.3 *TMLP1* expression in cell lines

Expression of *TMLP1* was not detected in the human lens cell line FHL124 or HEK293 cells by RT-PCR (data not shown). Canine *Tmlp1* (XM_548670) was not detected in MDCK II cells (data not shown).

5.2.4 Cloning of wildtype and mutant isoforms of murine *Tmlp1*

The coding sequence of *Tmlp1* was amplified from murine cDNA and cloned into the pGEM-T easy vector. The full-length wildtype *Tmlp1* clone was subsequently used as a template to generate a mutant isoform lacking the predicted signal peptide (amino acids 2-29) (Tmlp1-SPM). Tmlp1-SPM was cloned by amplifying the coding sequence downstream of the predicted cleavage site. C-terminal V5-tagged constructs were created by amplifying the template DNA with a reverse primer which contained the sequence encoding the V5 epitope followed by a TGA termination codon. All four coding sequences were subsequently cloned from pGEM-T easy into the expression vector pcDNA3.1(-) and bi-directionally sequenced to confirm that the sequence was correct and in-frame.

5.3 Detection by western blot

5.3.1 V5-tagged Tmlp1 and Tmlp1-SPM

MDCK II cells transfected with wildtype V5-tagged mTmlp1 (mTmlp1-V5) and the V5-tagged signal peptide mutant (mTmlp1-SPM-V5) were lysed, run on an SDS-PAGE gel and detected by western blot (Figure 5.4). mTmlp1-V5, which has a predicted molecular mass of 23.3 kDa, was detected in four isoforms at approximately 19, 22, 23, and 27 kDa. The most prominent isoform was at ~19 kDa. mTmlp1-SPM-V5, which has a predicted molecular mass of 20.9 kDa was detected at ~21 kDa.

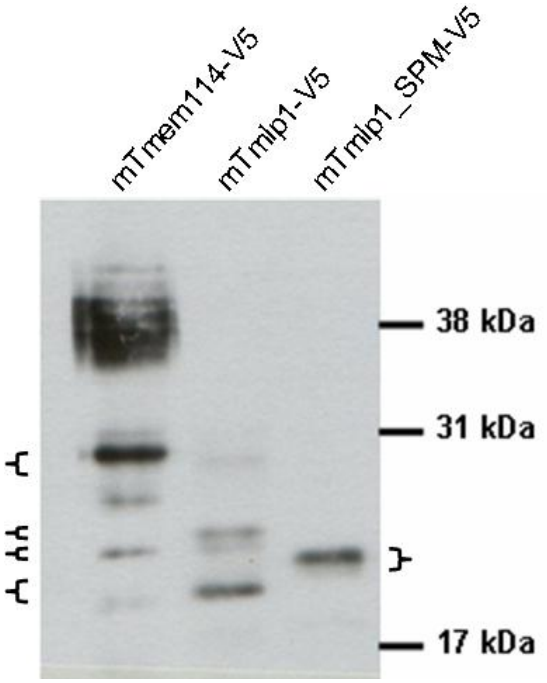


Figure 5.3. Detection of MDCK II cells transfected with V5-tagged wildtype (mTmlp1-V5) and signal peptide mutant Tmlp1 (mTmlp1-SPM-V5) by western blot. Tmlp1-V5 was detected in four isoforms (left braces). Tmlp1_SPM-V5 was detected at 21 kDa (right brace). V5-tagged Tmem114 was used as a positive control.

The detection of multiple isoforms of Tmlp1 by western blot (Figure 5.4) was suggestive of co-translation or post-translation modifications. Although mTmlp1 does not contain predicted N-glycosylation sites, the banding pattern of Tmlp1 was similar to that of the glycosylated protein Tmem114-V5 (Figure 5.4). Therefore, the presence of N-linked oligosaccharides was tested using PNGase F and tunicamycin. Treating transfected cells with the glycosylation inhibitor tunicamycin resulted in a shift to the 19 kDa isoform suggesting that mTmlp1 is glycosylated (Figure 5.5A). Treatment of mTmlp1 lysates with an enzyme that cleaved N-linked oligosaccharides (PNGase F) also resulted in a shift to the 19 kDa isoform. To determine if the N-linked oligosaccharides were mature complex oligosaccharides or immature mannose-rich oligosaccharides mTmlp1 lysate was treated with the enzyme Endo H which resulted in a shift to the 19 kDa isoform (Figure 5.5C). Thus the fastest migrating isoform (19 kDa) represents the unglycosylated form of mTmlp1-V5. Unexpectedly, this unglycosylated isoform migrated faster than Tmlp1-V5 lacking its signal peptide (21 kDa). However, on prolonged exposure of Tmlp1-SPM-V5, a faster migrating isoform was identified at ~17 kDa (Figure 5.5B). The slower 21 kDa migrating isoform was not formed in the presence of tunicamycin and Tmlp1-SPM-V5 was sensitive to both PNGase F (Figure 5.5B) and Endo H (Figure 5.5C) digestion indicating it contained N-linked mannose-rich oligosaccharides.

If the predicted signal peptide of Tmlp1 is cleaved, the de-glycosylated isoforms of Tmlp1-V5 and Tmlp1-SPM-V5 should migrate at the same speed on an SDS-PAGE gel. However, the deglycosylated isoform of Tmlp1-SPM-V5 migrated faster than the deglyosylated isoform of the full length Tmlp1-V5 (Figure 5.5C) indicating that the signal peptide of Tmlp1 is not cleaved.

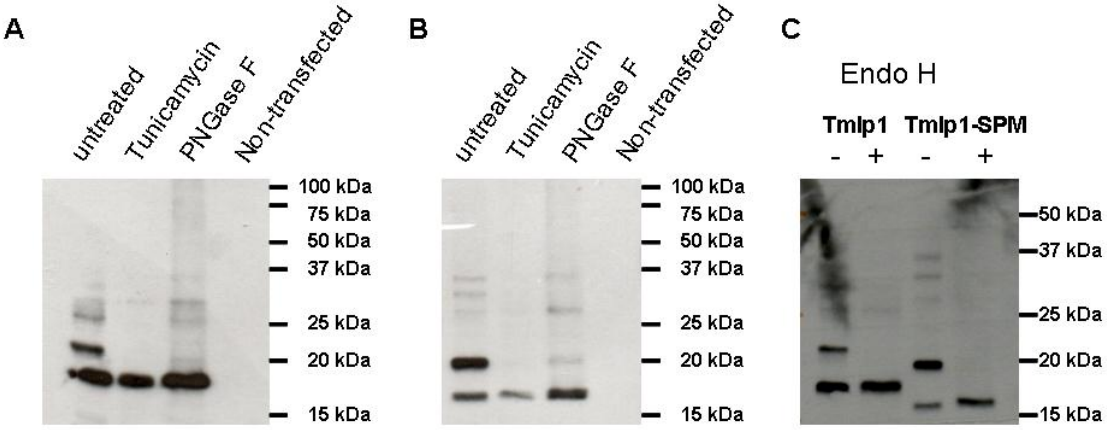


Figure 5.5. Testing of Tmlp1 and Tmlp1-SPM for the presence of N-linked oligosaccharides. MDCK II cells were transfected with V5-tagged Tmlp1 (A, C) or V5-tagged Tmlp1 signal peptide mutant (SPM) (B, C) were treated during culture or post-lysis to determine the presence of N-linked oligosaccharides. Tunicamycin prevents the addition of N-linked oligosaccharides. PNGase F cleaves all N-linked oligosaccharides from proteins. Endo H cleaves N-linked mannose-rich oligosaccharides or hybrid oligosaccharides, but not complex oligosaccharides processed in the Golgi Apparatus. A) N-linked glycosylation of Tmlp1-V5. Untreated Tmlp1-V5 is detected in three isoforms from 19 kDa to ~28 kDa. When transfected cells are grown in the presence of tunicamycin Tmlp1-V5 is detected in the 19 kDa isoform only. Tmlp1 lysates are sensitive to PNGase F, being detected in the 19 kDa isoform only. B) N-linked glycosylation of mTmlp1-SPM-V5. Untreated Tmlp1-SPM-V5 is detected in two isoforms at 17 kDa and 20 kDa. When transfected cells are grown in the presence of tunicamycin Tmlp1-V5 is detected in the 17 kDa isoform only. Tmlp1 lysates are sensitive to PNGase F, being detected in the 17 kDa isoform only. C) Endo H treatment of Tmlp1-V5 and Tmlp1-SPM-V5. Tmlp1-V5 and Tmlp1-SPM-V5 are sensitive to Endo H digestion as seen by the shift to the fastest migrating isoform.

5.3.2 Testing of custom polyclonal antibodies to Tmlp1

5.3.2.1 Epitope design

The final C-terminal 14 residues of mTmlp1 had previously been selected as the epitope for polyclonal custom antibody generation. This epitope is unique to Tmlp1 and it is predicted to be accessible as it does not lie in transmembrane domain (Figure 5.6). As discussed in section 3.3.6 this intracellular C-terminus is poorly conserved in mammals. The C-terminus of rat Tmlp1 is similar to that of mouse Tmlp1, possibly facilitating use of the antibody to detect rat Tmlp1 (Figure 5.6).

```
Homo sapiens        AWGSCALEAFSGTLLLSAAWTLSLSPPICGHLSPQQVGGRGGD 223
Macaca mulatta      AWGSCASEAFSGALLLSAARTLSLSPPLCSHLSPQQVGGRGGD 206
Mus musculus        AWASCASEVLSGALLLAAARLLSLSQRPG---VPHSVIL---- 211
Rattus norvegicus   AWGSCASEVLSGALLLAAARLLTLSQHPG---MPHSVIL---- 211
Bos taurus          AWGSCAAESLSGALLLTAARGLSLGGQPG---APHSVVI---- 212
Canis familiaris    AWGSCASEVLSGILLLTAARALRLSRRQG---GPHSVAV---- 210
                    **.*** * :** ***:**. * *.        *:.*
```

Figure 5.6. Multiple alignment of the C-terminus of Tmlp1 orthologues. The epitope used for custom antibody production is highlighted in the red box. This epitope is conserved in rat (*Rattus norvegicus*), but is clearly distinct to the C-terminus of primates (*Homo sapiens*, *Macaca mulatta*).

5.3.2.2 Sigma custom polyclonal antibody generation

Sigma-Aldrich synthesised and purified the designed KLH-conjugated peptides (purity 66 %) which were then used to immunise two New Zealand white rabbits (#4673 and #4674). The initial test sera were then compared to the pre-immune serum for their reactivity to transfected mTmlp1 by western blot. At the time of testing, a V5-tagged positive control had not been cloned, to aid in the identification of specific bands between the transfected to non-transfected cells.

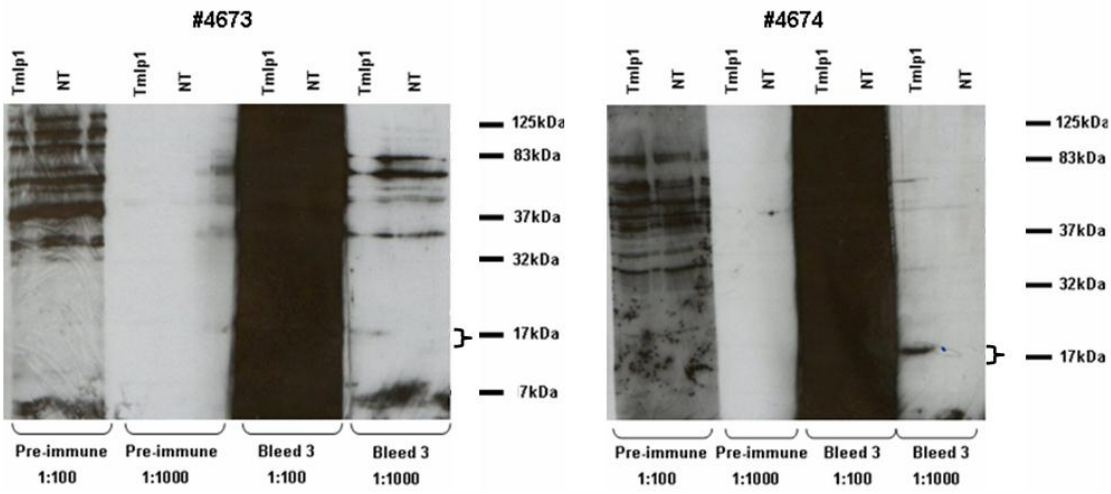


Figure 5.7. Testing of the #4673 and #4674 anti-mTmlp1 sera. Sera were tested on non-transfected MDCK II cells (NT) and MDCK II cells transfected with untagged Tmlp1 (Tmlp1). A single band of ~19 kDa (brace) is detected in the transfected cells by sera from both rabbits. This band was not detected in the non-transfected cells (NT) and was not detected by the pre-immune bleed.

Initial test sera from both rabbits detected a specific band at ~19 kDa in Tmlp1-transfected cells that was not detected with the sera of the pre-immunised rabbits (Figure 5.7). This was similar in mass to the predicted molecular mass of mTmlp1 (22 kDa). As the signal was stronger with the test sera from rabbit #4674, sera from this rabbit was chosen for antibody purification. The specific band at 19 kDa detected by the initial sera was also detected by the fourth test sera and final sera (Figure 5.8A). However this band was not detected by the purified antibody at concentrations as high as 1:100 (Figure 5.8A). The specificity of the final test sera was confirmed using peptide incubation. The final test sera detected full-length and the signal peptide mutant isoform of Tmlp1 (Figure 5.8B). When peptide pre-incubation was performed, the sera failed to detect two transfected isoforms of mTmlp1 (Figure 5.8B), confirming that the antibody is specific. Non-specific bands at 35 kDa and above 40 kDa remained in the pre-incubated samples.

The purified antibody failed to detect the 19 kDa band at dilutions as high as 1:50 dilution, and the ELISA results show that the purified antibody titre was significantly lower than that of the initial test sera (Figure 5.9).

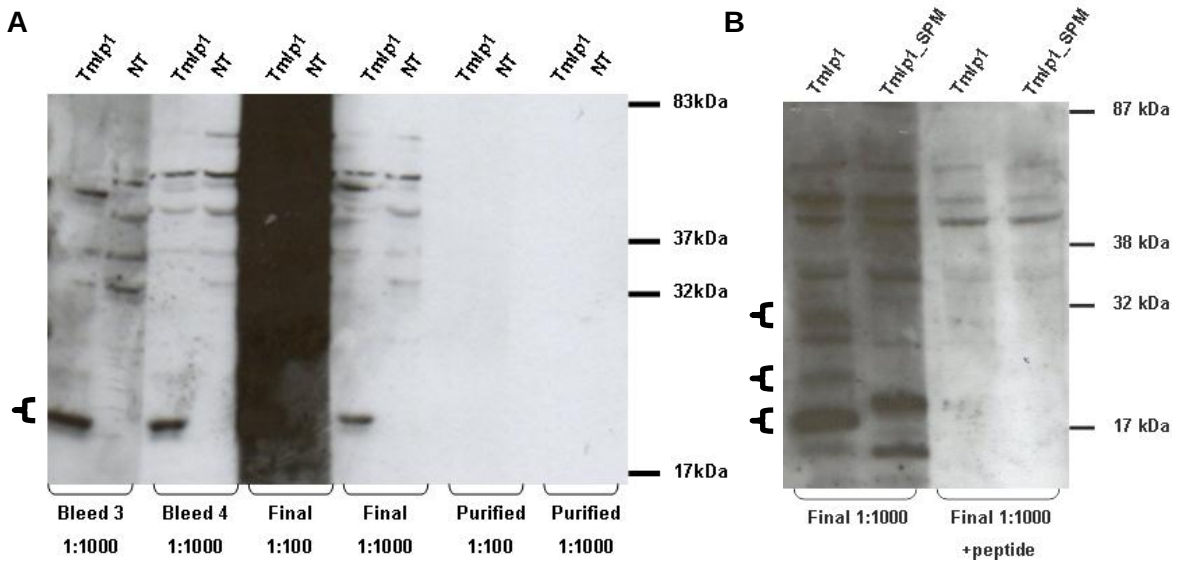


Figure 5.8. Testing of #4674 bleeds and purified antibody in transfected MDCK II cells. A) Comparison on Tmlp1-transfected to non-transfected cells. A single band of ~19 kDa (brace) is detected by all bleeds in the Tmlp1 transfected cells, but is not detected by the affinity purified antibody. B) Peptide incubation. Pre-incubation of the final bleed with the peptide epitope ablated the detection of the Tmlp1 bands (braces), confirming they represent mTmlp1. Detection of the Tmlp1 signal peptide mutant (Tmlp1-SPM) bands was also prevented by pre-incubation with the peptide epitope.

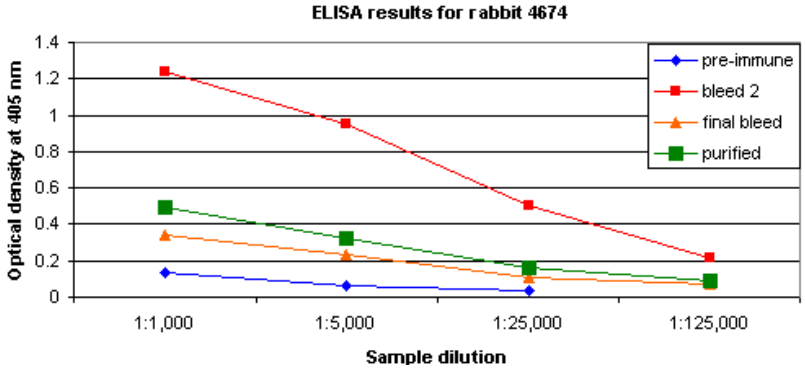


Figure 5.9. ELISA results for rabbit #4674 sera and purified antibody against mTmlp1. At a dilution of 1:1000, reads for all tested post-immunised sera were greater than pre-immunised serum (blue). The final bleed (orange) and purified antibody (green) were much weaker than Bleed 2 (red).

5.3.2.3 NEP custom polyclonal antibody generation

Due to the failure of the antibody purification process, custom polyclonal antibody generation was initiated with another company, New England Peptide (NEP) (Section 2.6.1.2). The same C-terminal 14 residues of mTmlp1 were used as the epitope. The synthesised KLH-conjugated peptide was >85 % pure and was used to immunise two New Zealand white rabbits (#4142 and #4144). Test sera were used to probe MDCK II cells transfected with Tmlp1. The Sigma-Aldrich final serum was used as a positive control for the presence of Tmlp1. Test sera from rabbit #4142 failed to detect Tmlp1 (Figure 5.10A). The sera from rabbit #4144 detected a single band at 19 kDa comparable to that detected by the Sigma-Aldrich antibody at dilutions of 1:1000 (Figure 5.10B).

Pre-incubating the final test serum from #4144 with the peptide epitope resulted in ablation of the band at 19 kDa indicating the band was specific to Tmlp1 (Figure 5.11A). This final serum from #4144 was affinity purified using the peptide epitope. The purified antibody was functional at a 1:200 dilution (Figure 5.11B). Figure 5.11B shows that the final sera from Sigma, the purified NEP antibody detect bands which match those of V5-tagged Tmlp1, although a background band at ~34 kDa was observed.

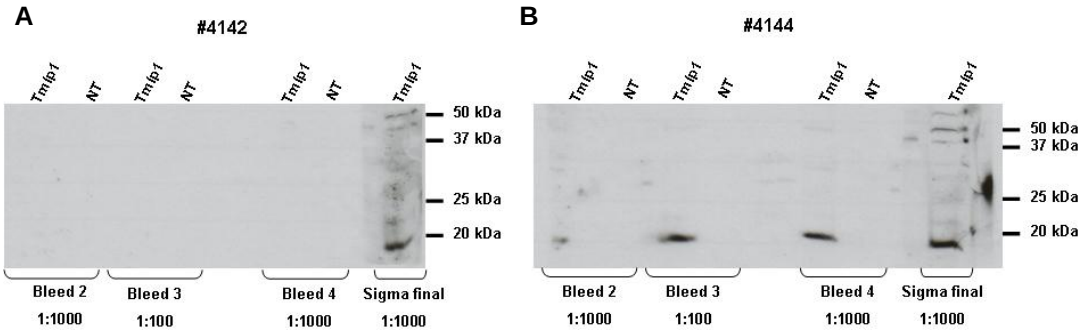


Figure 5.10. Testing of the #4142 and #4144 anti-mTmlp1 sera. Non-transfected MDCK II cells or MDCK II cells were transfected with mTmlp1, were lysed, run on SDS-PAGE gels and detected by western blot. A) Sera from rabbit #4142 failed to detect transfected Tmlp1, which was detected by the Sigma antibody at a dilution of 1:1000. B) Sera from rabbit #4144 detected a single band of ~19 kDa which was also detected by the Sigma antibody. This band was not detected in the non-transfected (NT) cells.

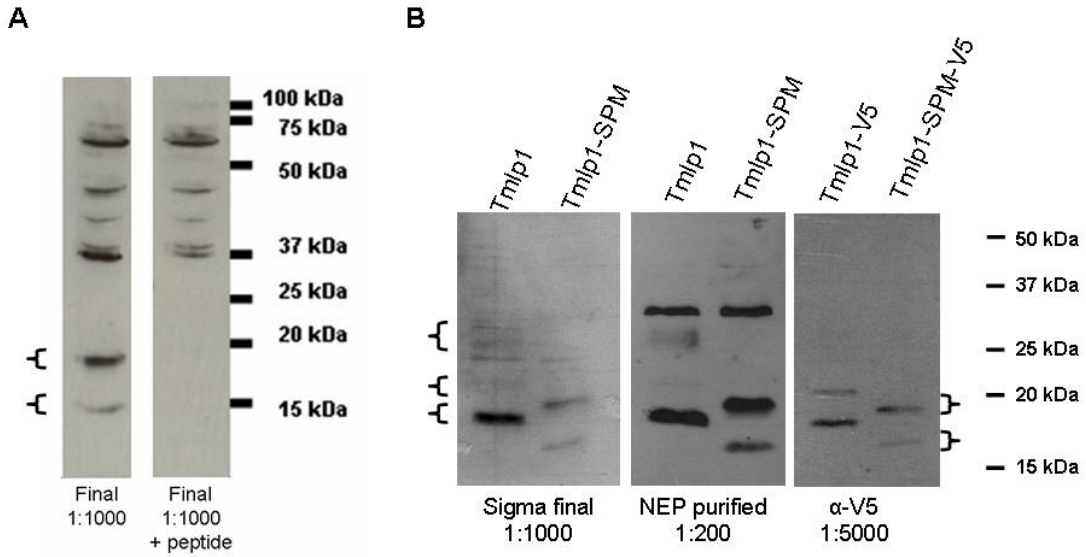


Figure 5.11. Confirming the specificity of the NEP anti-Tmlp1 antibody against transfected MDCK II cells. A) Peptide incubation. The final sera from rabbit 4144 detected two isoforms at 17 kDa and 19 kDa (braces) as well as some other slower migrating bands in Tmlp1 transfected MDCK II cells. Pre-incubation of the antibody with the peptide antigen blocked detection of the 17 kDa and the 19 kDa bands, but not the other bands, confirming that the 17 kDa and 19 kDa bands are specific. B) Comparison of isoforms of Tmlp1 detected by the Sigma antibody and NEP purified antibody and the V5-tagged Tmlp1 detected with the V5 antibody. Braces on left hand side: Tmlp1 isoforms. Braces on right hand side: Tmlp1-SPM isoforms.

5.2.4 Use of custom polyclonal antibodies to detect endogenous Tmlp1

After confirming the specificity of the purified antibody to the C-terminal epitope and optimising the concentration required for detection by western blot, the anti-Tmlp1 NEP purified antibody was used to probe lysates from a variety of mouse tissues. Bands were identified in a number of tissues, with most tissues displaying more than one band (Figure 5.12). The lens tissue contained a single band of ~20 kDa, close in molecular mass to the transfected Tmlp1 band. This 20 kDa lens band was not present in the whole eye sample suggesting it was enriched in the lens. However, pre-incubation with the peptide epitope did not prevent detection of the band in the lens suggesting it was non-specific (Figure 5.13B). No mouse brain lysate was available for testing, but a rat brain homogenate was available. Therefore, as the region of mouse Tmlp1 which was used to generate the custom antibodies was similar to that of rat Tmlp1 (Figure 5.6), a rat brain homogenate was probed with the anti-Tmlp1 antisera. Although the single band in the mouse lens sample was detected, no bands of the correct size were identified in rat brain homogenate (Figure 5.13A).

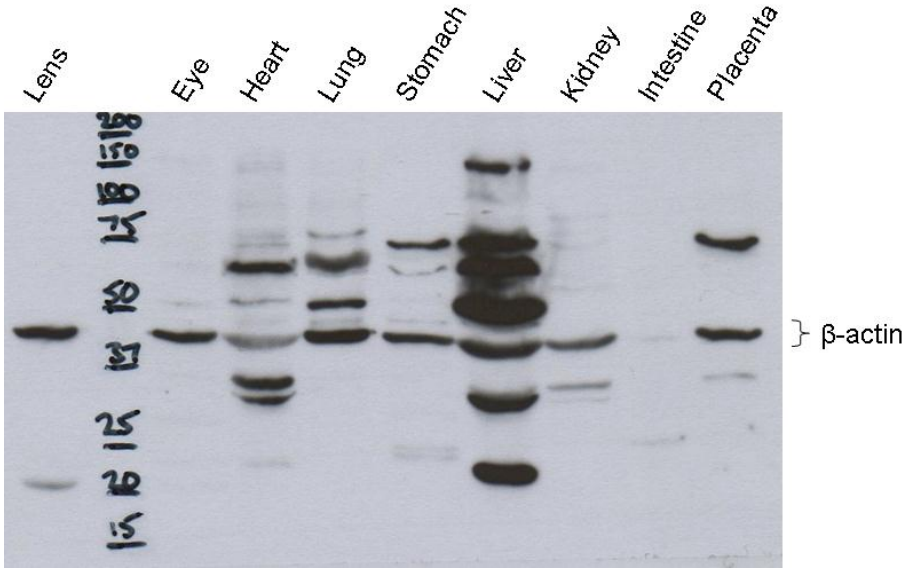


Figure 5.12. Western blot using the NEP anti-Tmlp1 purified antibody against a panel of adult mouse tissues. This blot was dual-probed for Tmlp1 and β-actin labelling. The housekeeping protein β-actin, detected at 44 kDa, was used as a loading control. All other bands are attributed to the Tmlp1 purified antisera.

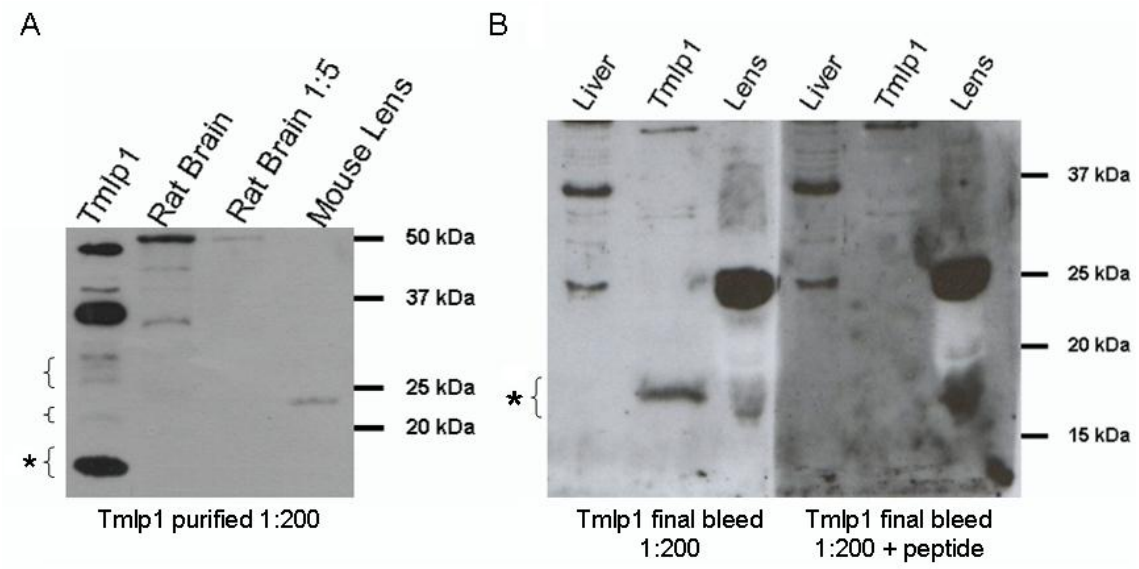


Figure 5.13. Western blots using anti-Tmlp1 antibody against rat brain, mouse lens and mouse liver. A) Tmlp1 transfected in MDCK II cells is detected in three isoforms (brace) and is mainly detected in the 17 kDa isoform (asterisk). Other bands represent background labelling. No similar bands were detected in rat brain lysates. A solitary band is detected in the mouse lens lysate at 24 kDa. B) Peptide pre-incubation with anti-Tmlp1 antibody. Tmlp1 transfected in MDCK II cells is detected at 17 kDa (asterisked brace) but is not detected when the antibody is pre-incubated with the peptide antigen. The bands detected in the liver (25 kDa, 35 kDa) and lens lysates (24 kDa) were present when the antibody was pre-incubated with the peptide antigen suggesting the bands were non-specific.

5.4 Localisation of mTmlp1-V5 and untagged mTmlp1 in MDCK II cells

The in silico analysis performed in chapter 3 suggested that Tmlp1 is a transmembrane protein (Table 3.1). Multiple unsuccessful attempts were made at creating stable MDCK II cells expressing V5-tagged or untagged Tmlp1. Therefore, transient transfections of Tmlp1 in MDCK II cells were used to determine the cellular localisation of the protein. V5-tagged Tmlp1 displayed an intracellular localisation pattern in transiently transfected polarised MDCK II cells (Figure 5.14Ai, ii). Untagged Tmlp1 detected with NEP purified anti-Tmlp1 antibody had a similar distribution to that of V5-tagged Tmlp1 (Figure 5.14B). The intracellular staining pattern resembled that of the ER, but as previously mentioned (Section 4.5.3) an antibody that detected the canine ER was not available to test for colocalisation of Tmlp1 with the ER. As Tmlp1 had a similar intracellular localisation when detected with both the anti-V5 antibody and the NEP purified anti-Tmlp1 antibody it was hypothesised that a) the localisation of V5-tagged Tmlp1 was not affected by the presence of the V5-tag, and b) the NEP purified anti-Tmlp1 antibody specifically detected Tmlp1 by immunofluorescence. Consistently, the number of cells expressing Tmlp1-V5 or Tmlp1 was found to be lower than for other constructs transfected at the same concentration.

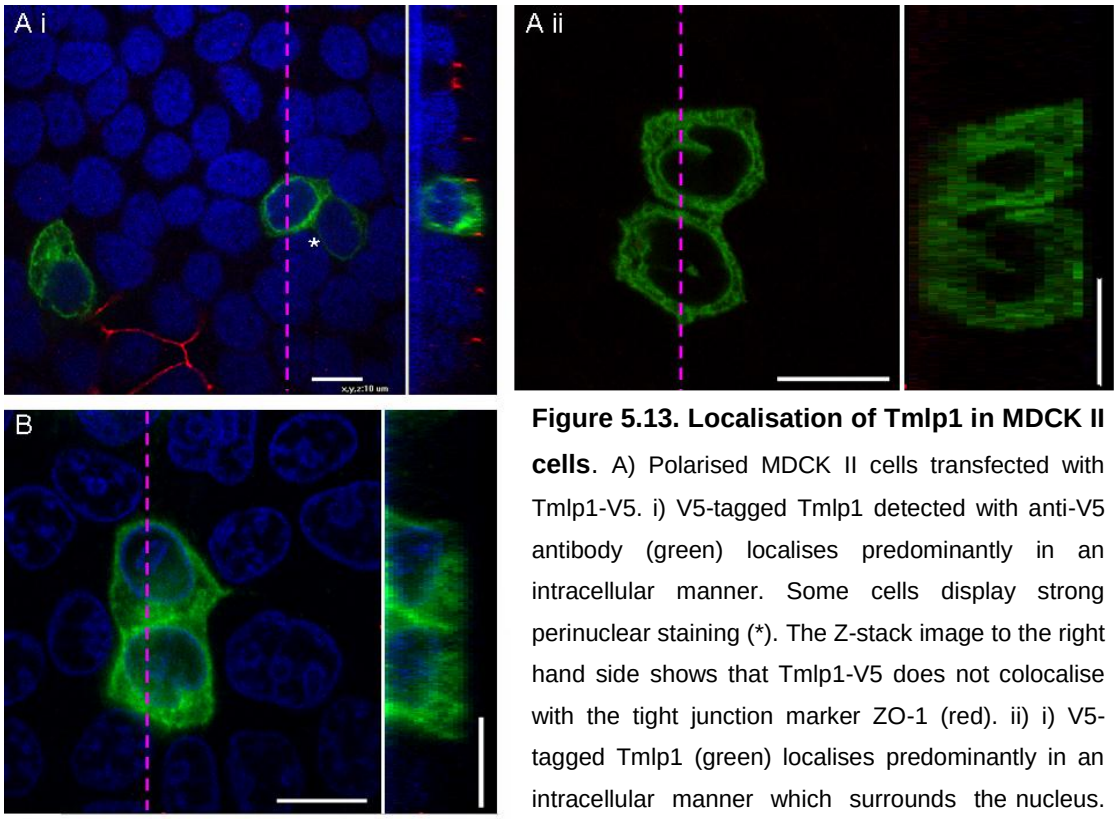


Figure 5.13. Localisation of Tmlp1 in MDCK II cells. A) Polarised MDCK II cells transfected with Tmlp1-V5. i) V5-tagged Tmlp1 detected with anti-V5 antibody (green) localises predominantly in an intracellular manner. Some cells display strong perinuclear staining (*). The Z-stack image to the right hand side shows that Tmlp1-V5 does not colocalise with the tight junction marker ZO-1 (red). ii) i) V5-tagged Tmlp1 (green) localises predominantly in an intracellular manner which surrounds the nucleus. Polarised MDCK II cells transfected with untagged Tmlp1. Untagged Tmlp1 detected with anti-Tmlp1 purified antibody (green) localises predominantly in an intracellular manner with some localisation at the plasma membrane. Z-stack images to the right of each panel show Y-Z cross-sections through the dashed pink line in the X-Y composite images. The left and right sides show the basal and apical membranes, respectively. Bar = 10 μm. Cells were fixed with 1:1 methanol:acetone and viewed by confocal microscopy.

5.5 Discussion

5.5.1 Expression of *TMLP1*

Expressed sequence tags (ESTs) for *TMLP1* have only been reported in adult human tissue (hippocampus and hypothalamus) and in embryonic stem cells (Unigene Hs.632228) (Section 3.7). To determine if *TMLP1* is also expressed in neural tissue during embryonic development, a panel of human cDNA from various central nervous system (CNS) tissues and eyes of different developmental stages was analysed for expression of *TMLP1* by RT-PCR. Interestingly, multiple splice isoforms were identified which matched those previously reported in EST databases. Although a *TMLP1* EST is reported from the adult hypothalamus (DB496856.1), expression was not detected in the developing diencephalons (which contain the hypothalamus). This may be due to the low levels of expression in the hypothalamus being diluted by the other tissues of the diencephalon or that expression in the hypothalamus is restricted to adult stages.

The earliest expression of *TMLP1* was detected in the hindbrain at CS21. This isoform lacked

exons two and three. An EST of the same splice form is present in the adult hippocampus (BC042066). Exons two and three encode the latter part of ECL1, TMD2 and part of TMD3 (Figure 5.3). The omission of exons 2 and 3 does not alter the frame of the protein, but results in the loss of domains found in the full-length protein. Thus, the resulting protein has three instead of four predicted TMDs; TMD2 is lost whilst TMD3 extends to include what was part of ECL2 (Figure 5.3). Additionally the smaller isoform is predicted to have an extracellular instead of intracellular C-terminus. Although many of the claudins and the TARPs contain protein interacting domains in their intracellular C-termini (Itoh *et al*. 1999; Schnell *et al*. 2002), these motifs are not present in the intracellular C-terminus of Tmlp1 (Section 3.3.6). This suggests two possibilities: 1) the C-terminus has an important function which can be performed inside or outside the plasma membrane; 2) TMLP1 does not contain interacting domains in its C-terminus and it becoming extracellular has no effect on protein function.

Any functional difference between the wildtype and the alternatively spliced isoform may be due to loss of latter half of ECL1. Loss of this region results in the loss of the second cysteine of the W-GLW-C-C motif. Also the alternatively spliced isoform has an altered membrane topology which results in the LW-C residues forming part of TMD2. Thus, this isoform lacks the signature motif (W-GLW-C-C) of the Pfam00822 family. A study which used peptides of the first half of ECL1 (includes W-GLW) and the second half of ECL1 (includes C-C) of claudin-1 showed that the latter peptide interacts with full-length claudin-1 and disrupts tight junction assembly (Mrsny *et al.* 2008). Thus, the latter half of ECL1 in claudins is functionally important for oligomerisation. Introducing hemagglutinin tags into the end, but not the middle, of ECL1 of the γ2 subunit prevents its interaction with AMPARs, suggesting that these residues mediate the interaction with AMPARs (Tomita *et al*. 2004). Omission of this domain in TMLP1 due to alternative splicing is likely to be functionally significant, possibly affecting its ability to interact with itself or other proteins with the W-GLW-C-C motif.

The short isoform lacking exons 2 and 3 was also detected in the foetal stage 2 eye where it was co-expressed with the full-length *TMLP1*. Expression of *TMLP1* was not detected at the earlier stages of CS16 and CS21. Expression of the full length isoform had not previously been reported and its detection confirms the predicted exonic structure in Figures 3.19 and 3.22 (based on the now obsolete Genbank accession XM_211287.4). As described in section 3.9, the TMLP1 locus loc283999 is predicted to encode two distinct proteins, encoded by exons 1, 3 ,4 and 5 (Figure 3.21); ENSP00000364084 is a 194 amino acid splice isoform of TMLP1, ENSP00000402790 is a 394 amino acid protein encoded by the same transcript translated in a different frame. ENSP00000402790 is a predicted globular protein (Philius) that has no known homologous proteins. Section 3.9 describes how the retention of exon 2 does not alter the frame of *TMLP1* but results in a premature stop codon in ENSP00000402790. Therefore the detection of the expression of an amplicon containing exons 1, 2, 3 and 4 by RT-PCR confirms that TMLP1 is the protein encoded at this locus.

The spinal cord at developmental stage foetal stage 2 showed weak expression of the isoform that lacked exon 2 (the 194 amino acid isoform of TMLP1 ENSP00000364084). Omission of exon 2 results in the loss of the latter half of ECL1, a region which includes the second cysteine of the W-GLW-C-C motif. As discussed previously this may affect the protein's ability to interact with itself and other proteins. An investigation into the prevalence of alternative splice variants in transmembrane proteins in the human genome found that splice variants were common (Cline *et al.* 2004). Of the 15 proteins with four transmembrane domains tested, 13 showed alternative splice variants with differing transmembrane architectures (Cline *et al*. 2004). Thus the identification of TMLP1 isoforms with different transmembrane topologies is not unusual. The γ6 calcium channel subunit is also expressed in three isoforms – full length, without exon 2, and without exons 2 and 3 with the consequent loss of predicted transmembrane domains (Burgess *et al.*, 2001). The full-length isoform of *Cacng6* is expressed in rat skeletal muscle (Chu *et al.*, 2001). Expression of *CACNG6* in the human brain is present in all three isoforms, predominantly in the isoform lacking only exon 2 (Burgess *et al.*, 2001). However, the functional significance of the alternative isoforms of γ6 remains to be elucidated.

5.5.2 Expression of murine *Tmlp1*

Expression of murine *Tmlp1* had not previously been reported in any database (Section 3.7). Full-length *Tmlp1* was detected by RT-PCR but no additional alternatively spliced variants were identified. *Tmlp1* was detected in the brain, eye, lung, stomach and, at very low levels, in the kidney (Figure 5.1A). The concentration of RNA extracted from heart and muscle repeatedly was weak so these tissues were not tested. Expression of *Tmlp1* was also detected in the developing brain and eye (Figure 5.1B), similar to the expression of *TMLP1* in humans (Figure 5.2). In the lens, *Tmlp1* was expressed in the epithelium and to a lesser extent in the fibres.

Translation of the full-length murine Tmlp1 sequence in other frames does not produce a protein similar to the globular protein ENSP00000402790 (data not shown) further confirming that Tmlp1 is the protein encoded by this locus.

To determine the expression pattern of Tmlp1, custom rabbit polyclonal antibodies were generated against the final 14 amino acids of the intracellular C-terminus. The sera generated by Sigma-Aldrich detected transfected Tmlp1 however the purified antibody did not (Figure 5.8A). Therefore a second rabbit polyclonal antibody was generated by New England Peptide (NEP). The purified antibody detected transfected Tmlp1 when used at dilutions of 1:200 (Figure 5.11B). A panel of adult mouse tissues were probed with the anti-Tmlp1 antibodies to investigate the expression pattern of Tmlp1. Bands were present in the majority of tissues tested, including some that were negative for expression of *Tmlp1* at the mRNA level (Figure 5.1A). However, the molecular mass of the bands detected was much larger than expected and no two tissues had the same bands, suggesting that the bands were non-specific (Figure

5.12). A single band was present in the lens sample that was not present in the whole eye sample suggesting the protein was enriched in the lens. However the lens band could not be competed out by peptide pre-incubation, suggesting the band may be non-specific (Figure 5.13B).

Thus, although the anti-Tmlp1 antibodies generated were able to detect Tmlp1 when it was over-expressed in cell lines, Tmlp1 was not detected in tissue. Expression at the mRNA level is low as detection by RT-PCR required a large number of PCR cycles. Assuming this correlated with low levels of protein expression this may explain why the anti-Tmlp1 antibodies were unable to detect Tmlp1 *in situ*. An alternative explanation is that *Tmlp1* is unstable and is rapidly degraded before detection. However, unstable proteins tend to be rich in amino acids that can be phosphorylated (Yen *et al.* 2008) and Tmlp1 has few predicted phosphorylation sites (Figure 3.9).

5.5.3 *In vitro* detection of Tmlp1

Tmlp1 has a predicted molecular mass of 22 kDa, and with cleavage of its predicted signal peptide it has a predicted molecular mass of 19.1 kDa. When Tmlp1 was detected by western blot it was detected at approximately 19 kDa, suggesting the signal peptide was cleaved. To determine if this hypothesis was correct, Tmlp1 was cloned without its signal peptide. When the signal peptide mutant (Tmlp1-SPM) was detected by western blot it was detected at a lower molecular mass than the wildtype Tmlp1. This suggests that the signal peptide is not cleaved (Figure 5.5C). Thus, similar to the claudins and voltage dependent calcium channel γ subunits, the N-terminus of Tmlp1 is a signal anchor and not a signal peptide.

Although Tmlp1 was detected at lower molecular mass (19 kDa) than predicted (23.3 kDa), Tmem114-V5 was also detected at a smaller molecular mass (22 kDa) than expected (25.7 kDa) (Figure 5.4). The protein markers used are predicted to be 95% accurate (http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_4110025.pdf) which does not account for the 4 kDa difference between the expected and the observed molecular masses. Therefore the difference may be due to the properties of the amino acids in the protein affecting its gel resolution. Proteins with hydrophobic domains, such as transmembrane proteins, generally bind more SDS than hydrophilic proteins, which speeds up migration (Rath *et al.* 2009).

Several isoforms of Tmlp1 were detected from cells transiently transfected with Tmlp1 (Figure 5.5). The higher molecular mass isoforms were sensitive to digestion by enzymes that cleave N-linked oligosaccharides, suggesting that the protein is N-glycosylated. This finding was unexpected as no N-linked glycosylation sites were predicted in murine Tmlp1 (Figure 3.4). The majority of the protein detected was in the non-glycosylated form suggesting that the N-linked glycosylation process was inefficient. Typically oligosaccharides are added to asparagine residues (N) that are followed by any amino acid (except proline) and then a

serine or threonie (N-X-S/T) (Gavel and von Heijne 1990). Oligosaccharides may also be added to asparagines in the motif N-X-C but this is a rare occurrence (Giuffrida *et al.* 1997) and occurs less efficiently than at N-X-S/T sites (Breuer *et al.* 2001). Tmlp1 contains a potential N-X-C site in ECL1 at p.N61 (NSCV). This N-X-C site is also present in rat, which also lacks N-X-S/T sites (Figure 3.4). The asparagine at p.N61 is located four amino acids downstream, and two amino acids upstream, of the cysteine residues in the W-GLW-C-C motif (Figure 3.4). As discussed in section 3.10.4, it has been postulated, but never demonstrated, that the cysteine residues of this motif in claudins and Pfam00822 members form disulphide bonds.

Both N-linked glycosylation (Kaplan *et al.* 1987) and disulphide bond formation (Chakravarthi *et al.*, 2007) occur in a co-translational fashion in the ER lumen. The enzyme oligosaccharyl transferase, which adds the initial oligosaccharide (tetradecasaccharide) to the asparagine residue, is thought to interact with the hydroxyl group of the serine/threonine of the N-X-S/T motif (Bause and Legler 1981). It has been postulated that the sulphydryl group (SH) of the cysteine in N-X-C can act in place of the hydroxyl group (OH) on the serine or threonine residues for hydrogen bonding with oligosaccharyl transferase, albeit less efficiently (Breuer *et al*. 2001). If oligosaccharyl transferase binds the SH group before disulphide bonds form, disulphide bond formation may be prevented. Inversely, formation of the disulphide bond may prevent glycosylation of the asparagine residue. Prevention of disulphide formation in the haemagglutinin-neuraminidase (HN) of Newcastle disease virus results in glycosylation of an asparagine site between the cysteines that is not normally glycosylated (McGinnes and Morrison 1997). A similar effect is thought to occur with the human receptor activity modifying protein-3 (RAMP3), as blocking of disulphide bond formation increases N-linked glycosylation (Flahaut *et al.*, 2003). Mutating the first of the two cysteine residues may establish if the cysteine residues form a disulphide bridge and if so, determine if prevention of disulphide bride formation promotes glycosylation at the p.N61 site.

The significance of the N-X-C site in Tmlp1 is unknown. There is an equivalent N-X-C site in the human orthologue at p.N66 that encompasses the second cysteine of the W-GLW-C-C motif (Figure 3.3). However, human TMLP1 also contains a N-X-S/T site at p.N41 and so if human TMLP1 is glycosylated *in vivo* it is likely to be at this site and not at the N-X-C site as glycosylation is more efficient at N-X-S/T sites than N-X-C sites (Breuer *et al*. 2001).

5.5.4 Localisation of Tmlp1

N-linked oligosaccharides covalently linked to Tmlp1 were sensitive to Endo H. Endo H cleaves N-linked oligosaccharides that have not undergone further processing to complex oligosaccharides in the Golgi apparatus, inferring that the protein has not been processed beyond the ER. When the localisation of Tmlp1 was investigated in MDCK II cells, Tmlp1 was localised in the cytoplasm in a reticulate pattern indicative of the ER (Figure 5.14). As previously mentioned in section 4.5.3 the ER antibodies tested did not label the ER in the

canine-derived MDCK II cells, preventing colocalisation experiments being performed. However, the sensitivity to Endo H digestion confirms the ER localisation of Tmlp1. Integral membrane proteins of the ER may also be glycosylated (Kornmann *et al.*, 2009). Human pancreas-specific protein disulfide-isomerase (PDIp), an ER chaperone, is N-glycosylated (Desilva *et al.* 1997). Maintenance of mitochondrial morphology 1 (MMM1), previously thought to be a mitochondrial protein, was re-annotated as an integral ER membrane protein due to its localisation and sensitivity to Endo H (Kornmann *et al.*, 2009). Thus, Tmlp1 may be an N-linked glycosylated integral ER membrane protein.

However, misfolded plasma membrane proteins may also be retained in the ER (Hwa *et al.* 1997; Choi *et al.* 2005; Gallagher *et al.* 2007). As discussed in section 4.8.3 glycosylation of proteins may be required for correct folding and trafficking of membrane proteins and Tmlp1 was predominantly detected in the non-glycosylated isoform. If Tmlp1 is being misfolded due to a lack of glycosylation it may be retained in the ER. Misfolded proteins retained in the ER are detected and degraded by ER-associated degradation (ERAD), a proteasome dependent degradation pathway (Gelman *et al.* 2002; Meusser *et al.* 2005). There were a number of indicators that Tmlp1 was being degraded. Repeatedly, when detected by immunofluorescence, very few cells were positive for Tmlp1 expression. Also, numerous unsuccessful attempts were made to create stable cell lines expressing Tmlp1 or Tmlp1-V5. In contrast, transiently transfected Tmlp1 and Tmlp1-V5 were detected at high levels of expression. With the more moderate levels of expression associated with stable transfections, the cellular machinery may efficiently degrade the misfolded protein. The vector backbone, cells and transfection procedure used to express Tmlp1 were the same as used for Tmem114, suggesting that the low levels of protein detected was due to something inherent in Tmlp1 and not some other factor.

When studying the cellular localisation of a protein the choice of an appropriate cell line is important (Davidson *et al.* 2009). As *Tmlp1* is expressed in epithelium of the lens and in other epithelial tissues (lung, stomach, and kidney) MDCK II cells were chosen to study its localisation as these cells are a model epithelial cell system. As Tmlp1 was also expressed in regions of the brain perhaps a neuronal cell line such as HT-22 cells (Sarkar *et al.* 2008) would correctly process Tmlp1. Or perhaps a murine cell line such as L cell fibroblasts (Kubota *et al.* 1999) would correctly process murine Tmlp1.

5.5.5 Limitations and future directions

The study of murine Tmlp1 has revealed little in terms of functional aspects of the protein. However, there were a number of reasons for studying murine Tmlp1 and not the human orthologue. Firstly, the antibodies had previously been designed against murine Tmlp1 to facilitate their use for western blotting and immunohistochemistry with mouse tissue. While the C-terminus of Tmlp1 was a good choice of epitope in respect of the fact that because it would detect all splice variants, it limited the study to murine tissue as the C-terminus is not

conserved (Figure 5.6). However, there are no regions other than the transmembrane domains which show enough conservation between mouse and human for the antibody to detect both proteins. The region surrounding the G-LW-C-C is the most similar but using this region as the epitope may have resulted in non-specificity as it may detect other Pfam00822 proteins. Also, the predicted presence of glycosylation sites and disulphide bond sites may have resulted in a tertiary structure which would block access to the epitope. The second reason for studying murine Tmlp1 was that murine cDNA was widely available, allowing Tmlp1 to be cloned. The panel of human cDNAs only became available half-way through this study, by which point the antibodies directed against the C-terminus of murine Tmlp1 had been generated and the murine orthologues had been cloned. No full-length IMAGE clone for human TMLP1 exists.

With more time and resources, and the data generated from studying Tmlp1, it would be desirable to characterise the human orthologue. All three isoforms of TMLP1 detected by RT-PCR could be cloned, initially with a V5 tag. The presence of N-linked oligosaccharides could be detected using de-glycosylating enzymes and if glycosylation was confirmed, the glycosylated site (predicted to be p.N41) could be identified by creating mutant protein. The localisation of each of the three TMLP1 isoforms could then be investigated, initially in MDCK II cells, and then in neuronal cells. If TMLP1 is glycosylated at the predicted site and correctly processed (i.e. the majority of the protein in the glycosylated form) then one could be confident in the localisation of the protein. As the C-terminus of the TMLP1 isoform lacking exons 2 and 3 is predicted to be extracellular, positioning the V5 tag at the C-terminus could determine if the C-terminus is extracellular or not by permeabilising the cells before labelling.

With sufficient resources the expression profile of *TMLP1* could also be investigated more extensively in developing eye and CNS tissues. The genes encoding the γ1 and γ6 subunits are expressed in skeletal muscle (Sandoval *et al*. 2007) and skeletal and cardiac muscle (Chu *et al*. 2001), respectively. *In vitro* γ1 and γ6 interact with voltage dependent calcium channels expressed in their representative tissues (Arikkath *et al*. 2003; Hansen *et al*. 2004). These tissues were not tested for *Tmlp1* expression as the extracted RNA was of poor quality. A wide range of adult and developing human cDNA samples could be used to create a more complete expression profile, which may show tissue specificity of the different isoforms. Once the tissues expressing *TMLP1* are identified, *in situ* hybridisation could be performed to determine a more precise spatial expression pattern.

To determine a functional role for *TMLP1* the knowledge of which tissues it is expressed in could be utilised to determine the localisation and post-translational modifications in an appropriate cell line (as previously discussed) and to identify interaction partners. The latter could be performed using a targeted approach such as co-immunoprecipitation with calcium channel subunits expressed in the tissues identified, or using a non-targeted approach such as pull-downs. Human lens epithelial cells express a number of α-subunits ($Ca_v1.2$, $Ca_v3.1$,

$Ca_v3.2$, and $Ca_v3.3$), and *in vitro* blocking of voltage dependent calcium channel activity decreases the rate of proliferation of cultured HLE-B3 cells (Meissner and Noack 2008). *Tmlp1* is expressed in the lens epithelium (Figure 5.1). Future co-immunoprecipitation experiments using lysates of HLE-B3 cells transfected with TMLP1 may identify an interaction between TMLP1 and the endogenously expressed voltage dependent calcium channels.

Bioinformatic analysis (chapter 3) showed that TMLP1 is similar to the calcium channel γ6 subunit, which modulates $Ca_v3.1$ activity *in vitro* (Hansen *et al*. 2004). The ability of Tmlp1 to modulate voltage dependent calcium channel activity was not tested for two main reasons. Firstly, Tmlp1 localised to the ER whereas voltage dependent calcium channel subunits localise to the plasma membrane (Brice and Dolphin 1999). Secondly, to test the ability of Tmlp1 to modulate calcium channel activity would have required the creation of cell lines expressing the α-, β-, and α2δ-subunits. There are a number of α-subunits which can be tissue-specific, requiring the creation of multiple cell lines. After consulting a researcher of voltage dependent calcium channels at the University of Manchester (Dr Liz Fitzgerald) it was decided that pursuing this area of research would be unwise considering the time required and the number of α-subunits that would need to be tested.

5.6 Conclusion

This study has identified that *TMLP1* is temporally and developmentally expressed in a wide variety of tissues. Such a wide expanse of tissue expression has also been established for the voltage dependent calcium channel γ subunits (Burgess *et al*. 2001) and for some claudins (Hewitt *et al.* 2006). Interestingly, three different splice isoforms were identified, and the presumptive full-length isoform, that was not annotated in any current database, was detected. Expression of the full-length isoform confirms that TMLP1 is expressed from loc283999. Although the custom antibodies generated detected murine Tmlp1 *in vitro*, Tmlp1 was not detected *in vivo*. The failure to detect murine Tmlp1, coupled with the unclear state of glycosylation of murine Tmlp1 limited the study. However, creating an antibody specific to human TMLP1, which would enable the study of the different isoforms, would have limited the work to *in vitro* analysis.

TMLP1 is a novel gene expressed in a number of tissues in the developing embryo and in the adult and its functional characterisation may reveal insights into eye and brain development. To proceed with the characterisation of TMLP1, V5-tagged constructs of all three splice isoforms could be cloned and expressed in epithelial and neuronal cell lines to study their localisation and possible interacting partners. With further time and resources the effects of TMLP1 expression on voltage dependent calcium channel activity would also merit investigation.

The knockout mouse project (http://www.knockoutmouse.org) currently (as of October 2010)

has knockout stem cells for Tmlp1. A knockout mouse may be a useful resource, but as discussed previously, murine Tmlp1 may not be representative of human TMLP1. However, if the N-X-S/T glycosylation site present in primates and ungulates is functionally significant it is surprising that it is not present in rodents. Due to the fact that a Tmlp1 orthologue in *Xenopus* was not identified and the presence of a Tmlp1 orthologue in zebrafish is unknown, animal models for Tmlp1 do not provide a good avenue for investigating Tmlp1 function.

Chapter 6: Knockdown of Tmem114 in *Xenopus tropicalis*

6.1 Introduction

To investigate the functional role of Tmem114 *in vivo*, a knockdown model was created. The African clawed frog *Xenopus tropicalis* was chosen to create the model for a number of reasons. Offspring are produced in large numbers and are amenable to knockdown of protein expression using antisense morpholinos. The embryos have a short developmental time and their development is external facilitating monitoring of phenotypes at all developmental stages. Morpholinos are modified DNA oligonucleotides that bind to complementary sequences in the target mRNA. Binding to the mRNA blocks translation by the ribosome by steric hindrance (Heasman *et al.* 2000). This results in the knockdown of target protein expression. Morpholinos are easily injected into *Xenopus* zygotes allowing the knockdown of a specific protein in the developing embryo.

6.2 Expression of *xtTmem114* in *Xenopus tropicalis* development

Before attempting to create an animal model for Tmem114, it was necessary to confirm that Tmem114 was expressed in developing *X. tropicalis* using RT-PCR and *in situ* hybridisation experiments were performed.

6.2.1 RT-PCR in stage and tissue series

RT-PCR using primers that crossed the first intron of *Tmem114* in *X. tropicalis* (*xtTmem114*) were used to detect expression in a stage series of developing embryos (Figure 6.1 A). The forward primer was designed to the 5'UTR and encompassed the morphlino binding sites (Sections 6.3.1 and 6.3.2). Maternal expression of *xtTmem114* was detected at stages 1-7, but transcription by the developing embryo does not occur until stage 23, when it is weakly expressed. Expression is upregulated in later stages. To determine which tissues *xtTmem114* is expressed, a panel of cDNA from various adult tissues was used for RT-PCR. Expression was observed in the eye, lung and intestine (Figure 6.1B). Expression in the developing eye was also detected at similar levels to that in the adult eye (Figure 6.1C).

6.2.2 *In situ* hybridisation

To determine the expression of *Tmem114* in different tissues in the developing embryo, *in situ* hybridisation was performed. Two probes were used on separate occasions. Initially, probes transcribed from full-length *xtTmem114* were used but there appeared to be no difference in the staining pattern between the antisense and the sense (negative control) probes. Staining was observed but this appeared to be due to probe trapping. To improve tissue penetration of the probe, a shorter probe was then used. This probe was based on the only *X. tropicalis Tmem114* EST (Genbank accession CX373776, IMAGE: 7632995), which includes exons 3 and 4. However, no staining difference was seen between the antisense and sense probes (Figure 6.2A). The lens specific gene *Lim2* was used as a marker for the lens and as a positive control for the *in situ* hybridisation experiments. Figure 6.2B clearly demonstrates that *xtLim2* is expressed in the lens and the staining was specific to the antisense probe, confirming that the protocol used could detect expression of genes in the lens.

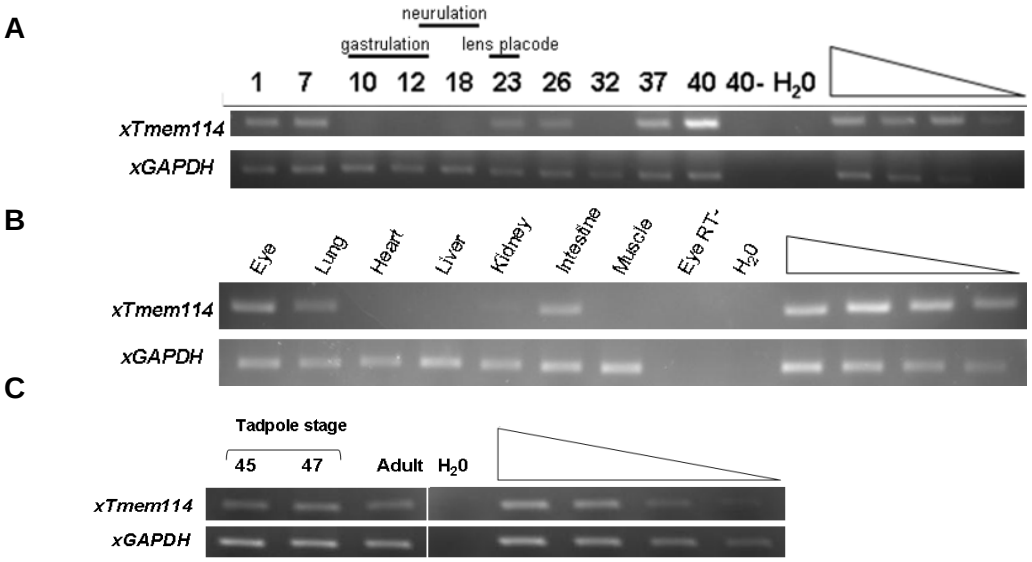


Figure 6.1. Expression of xtTmem114 in developing and adult X. tropicalis. A) Expression in the developing whole embryo. Maternal expression was detected (stages 1-7). Expression in the developing embryo is restricted to later developmental stages. B) Expression in adult tissues. Expression was detected in the adult eye, lung and intestine. C) Expression in the developing and adult eye. Detection of the housekeeping gene *Gapdh* was used as a control for equal cDNA loading in the RT-PCR reactions. The linearity for each reaction is indicated (triangle). 40- and Eye RT- represent samples to which no reverse transcriptase was added. H_20 represents a no template control.

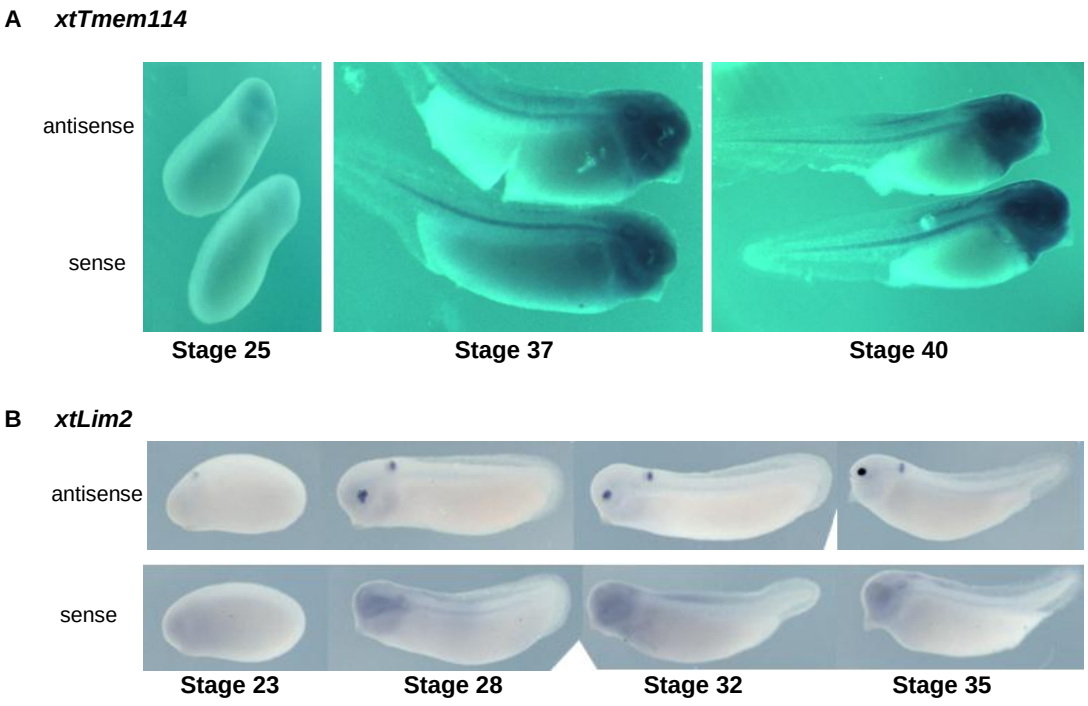


Figure 6.2. Developmental expression analysis by *in situ* hybridisation. A) xtTmem114. No difference in staining was observed between the antisense and sense (negative control) probes based on the xtTmem114 EST CX373776. B) xtLim2. Staining with the antisense probe was detected in the lens and a single rhombencephalon. Diffuse staining was detected with the sense probe.

6.3 Morpholino knockdown of xtTmem114

To knock down expression of Tmem114, morpholinos were designed to block translation of *Tmem114* mRNA. Searches were performed with morpholino sequences against sequence databases to identify possible cross-reaction before being synthesised.

6.3.1 Morpholino 1

6.3.1.1 Design and optimisation

Morpholino 1 (MO1) was 25 nucleotides in length and was designed to bind across the translation start site, binding from 3 nucleotides upstream of the A of the translation start site (AUG) to 21 nucleotides downstream (Figure 6.3A). Initially, a range of concentrations up to 20 ng of MO1 were injected in order to ascertain what levels were toxic to the embryo. Doses up to 10 ng had normal levels of embryo survival (Figure 6.3C). When the developing embryos reached at least stage 40, phenotypes were observed and scored. Bilateral and unilateral microphthalmia were observed in a dose dependent manner (Figure 6.3B,D). Other ocular anomalies such as coloboma occured at much lower frequencies (Figure 6.4A iv).

6.3.1.2 Unilateral injections

To increase the sensitivity of scoring for microphthalmia, unilateral injections were performed. Unilateral injections of morpholino create a unilateral knockdown allowing comparison with the uninjected side as an internal control. As survival rates were greatest for 10 ng bilateral injections of MO1 (Figure 6.3C), 5ng of MO1 was chosen as the optimal dose for unilateral injections. 5 ng of MO1 when injected unilaterally resulted in 39 % of embryos displaying an observable difference in eye size (Figure 6.4).

6.3.1.3 EGFP injections

To confirm the observed phenotype developed on the injected half of the embryo, MO1 was co-injected unilaterally with a plasmid expressing enhanced green fluorescent protein (EGFP). Embryos were visualised using a fluorescent microscope at stage 30-33. At later stages EGFP expression was lost. Figure 6.5 is a representative image displaying that the microphthalmia associated with the EGFP injected side of the embryo.

6.3.1.4 Increased defects in gut development in MO-injected embryos

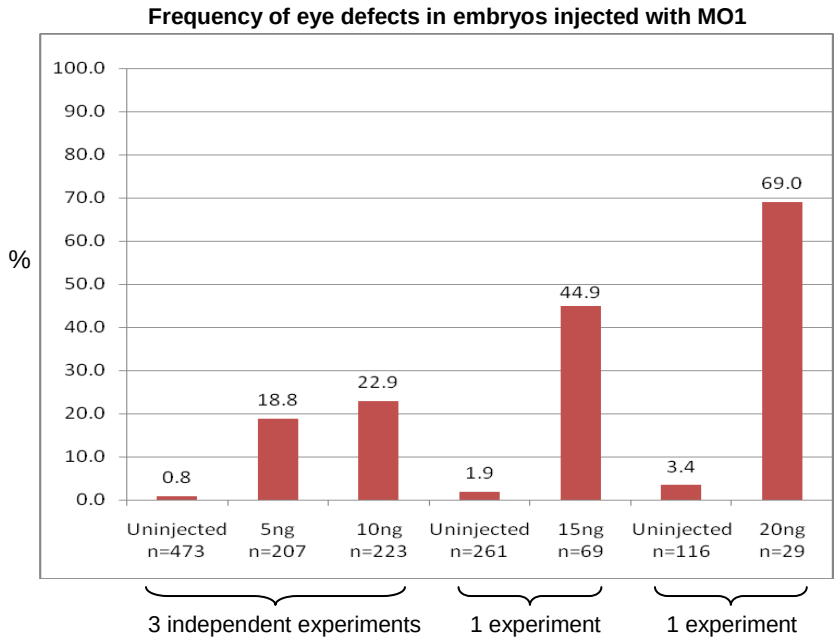
When injections were performed with higher doses of MO1 an increase in gut developmental defects was observed (Figure 6.6). The developing gut typically forms a coil structure formed from two external anticlockwise loops with a small inner clockwise loop (Chalmers and Slack 1998) (Figure 6.6 A). Embryos injected with 15 ng of MO1 bilaterally or 7.5 ng of MO1 unilaterally displayed defects in gut development at a frequency of 51 % and 24 % respectively, compared to 7 % of uninjected embryos (Figure 6.6B). The majority of the morphants had uncoiled guts which developed as a single curved tube whilst others formed more bulbous structures (Figure 6.6A).

A

aacacaggaacaaguggauuccucuucuguaucugca**AUGAAGCUGAAACUGAGUAUGUUAUCUGUCUU**

CGTTACTTCGACTTTGACTCATACA

B



C

	Final n=	Original n=	% survival
Uninjected	473	665	71.1
5ng	207	424	48.8
10ng	223	385	57.9
Uninjected	261	600	43.5
15ng	69	462	14.9
Uninjected	116	219	53.0
20ng	29	162	17.9

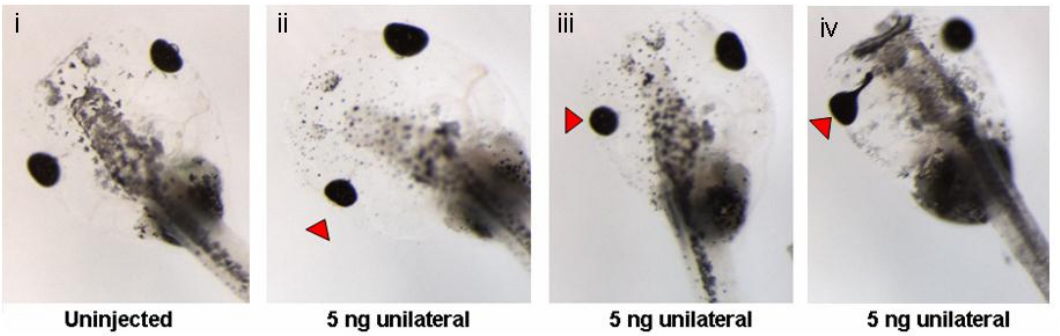
D

	Uninjected		
	n=	eye defect	%
Run1	206	2	0.97
Run2	116	1	0.86
Run3	151	1	0.66
Combined	**473**	**4**	**0.85**

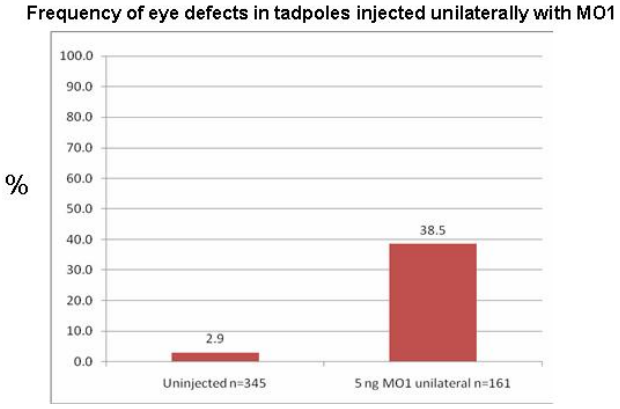
	5ng			10ng		
	n=	eye defect	%	n=	eye defect	%
Run1	80	16	20	107	26	24.3
Run2	53	13	24.53	24	5	20.83
Run3	74	10	13.51	92	20	21.74
Combined	**207**	**39**	**18.84**	**223**	**51**	**22.87**

Figure 6.3. Bilateral injections of MO1. A) Binding site for MO1. MO1 (red) binds three bases of the 5' UTR (lower case) and the first 22 nucleotides of the coding sequence of *xtTmem114* (upper case, bold). B) Frequency of eye defects (%) observed in uninjected embryos and embryos injected with MO1. Frequency of eye defects was dose dependent C) Survival rates of embryos for each dose. D) Data for the 3 independent experiments for 5ng and 10 ng MO1 injections

A



B



C

	Uninjected		
	n=	eye defect	%
Run1	116	4	3.4
Run2	111	5	4.5
Run3	118	1	0.8
Combined	345	10	2.9

	5ng unilateral		
	n=	eye defect	%
Run1	77	26	33.8
Run2	47	20	42.6
Run3	37	16	43.2
Combined	161	62	38.5

Figure 6.4. Unilateral injection of 5 ng MO1. A) Injected embryos displaying unilateral microphthalmic eyes (ii, iii) (arrowhead). On rare occasions embryos displayed unilateral retinal coloboma (iv) (arrowhead). B) Frequency of unilateral microphthalmia observed uninjected embryos and in embryos unilaterally injected with 5 ng MO1. C) Data for the 3 independent experiments for 5ng unilateral injections

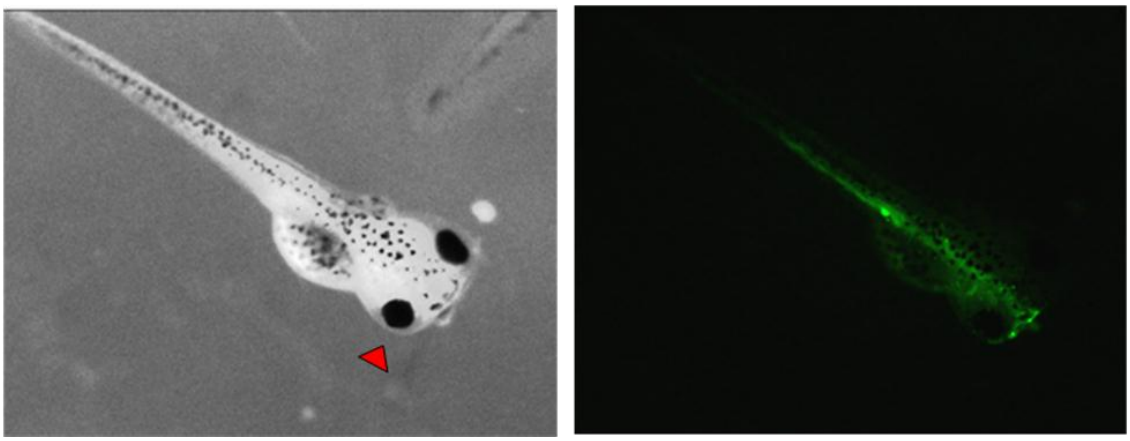
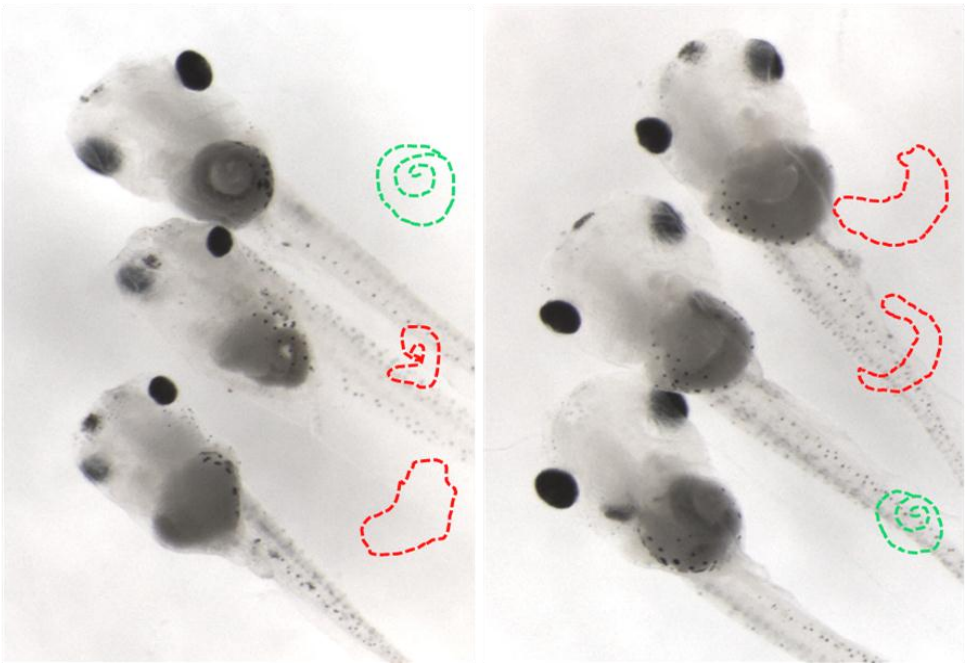


Figure 6.5. Unilateral co-injection of EGFP with 5 ng MO1. An MO1/EGFP injected embryo visualised under white light (left) and GFP filter (right) confirms the phenotype is observed on the MO1 injected side.

A



B Eye and gut defects in uninjected embryos and embryos injected with MO1

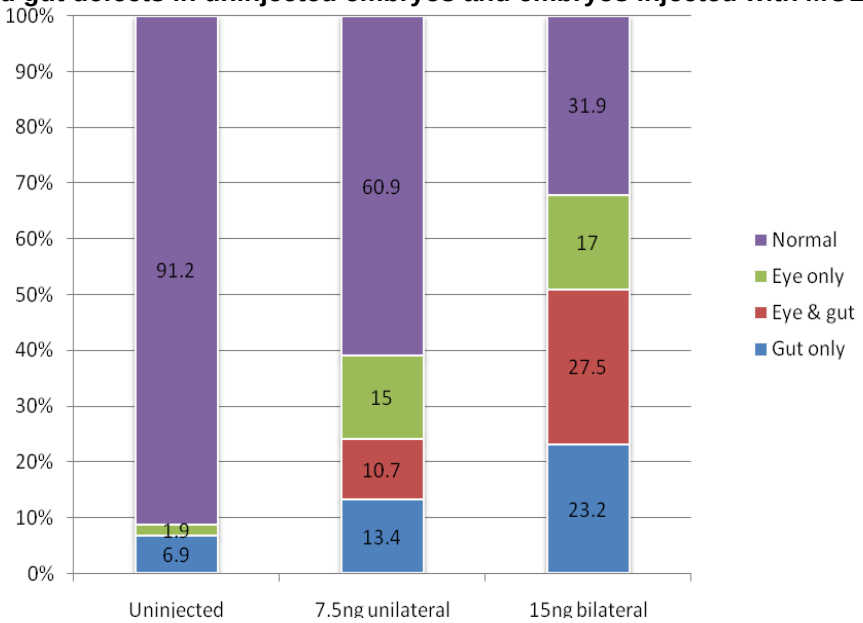


Figure 6.6. Gut defects in MO1-injected embryos. A) Embryos injected with 15 ng MO1 displaying defects in gut development. Gut morphology is illustrated to the right of the embryos which display their ventral surface. The correctly developed guts (green illustrations) of uninjected embryos form a coiled structure of antisense loops. The morphants displayed failure of the coil to form correctly, formation of a singular tubular gut or formation of a bulbous structure (red illustrations). B) Frequency of eye and gut defects observed in embryos injected unilaterally with 7.5 ng MO1 and bilaterally with 15 ng MO1. Data presented is from a single experiment. Eye only = unilateral eye defects with no other abnormalities. Gut only = Abnormal gut development with no other abnormalities. Eye and gut = Unilateral microphthalmia with gut abnormalities.

6.3.1.5 Knockdown of xtTmem114 with morpholino *in vitro*

The ability of MO1 to prevent translation of *xtTmem114* was tested *in vitro*, as the custom polyclonal antibody to mTmem114 does not detect xtTmem114 (data not shown). Full-length *xtTmem114* including 23 nucleotides of the upstream 5'UTR was cloned with a C-terminal V5 tag. When transcribed and translated *in vitro* multiple bands between 10 and 18 kDa were detected by western blot (Figure 6.7). These bands were not present when no *in vitro* transcribed RNA was added to the *in vitro* translation reaction (Figure 6.7, negative control). To test if MO1 blocks translation of these xtTmem114 products, MO1 was added to the *in vitro* translation reaction at a final concentration similar to that of the final concentration of MO1 in the injected embryo (Hilton *et al.* 2007). A separate reaction using a standard control morpholino at the same concentration was also performed. Detection of the protein bands was not prevented by addition of MO1 (Figure 6.7).

The V5-tagged xtTmem114 has a predicted molecular mass of 24.8 kDa, but the bands detected were of a lower molecular mass (12-18 kDa). When the coding sequence of this construct is translated in silico seven isoforms which end in an in-frame V5-tag are produced ranging from full-length (24.8 kDa) to 10.5 kDa (Figure 6.8). The bands detected by western blot are similar in molecular mass to the predicted molecular mass of the four putative isoforms (10.5 kDa, 12.6 kDa, 14.7 kDa, 18.3 kDa). This suggests that the translational machinery was utilising alternative start codons producing shorter isoforms, that were not blocked by MO1 which targets only the true start codon.

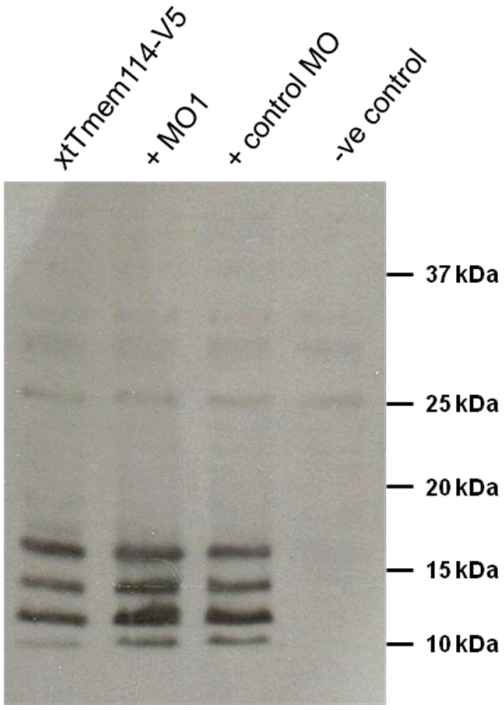


Figure 6.7. Detection of in vitro translated products by western blot. Non-specific bands at 25 kDa, 30 kDa and 34 kDa were detected in all samples. Three major bands between 12 and 18 kDa and one minor band of 10.5 kDa were detected. These bands were not detected when no RNA was added to the *in vitro* translation reaction (-ve control). Expression of these isoforms was not prevented by addition of xtTmem114 MO1 or a standard control MO.

A

>pGEM_xtTmem114-V5

taatacgactcactatagggcgaattgggcccgacgtcgcatgctcccggccgccatggcggccgcgggaattcgat
tTGGATTCCTCTTCTGTATCTGCAATGAAGCTGAAACTGAGTATGTTATCTGTCTTTGTTGCTGTGGTTGGCATATT
AAGTTTCATATCATTAGTTGTTGCTATTGGAACTGATTTTTGGTATATTATTGATGCGTCCAGACTGGAAAAGATCA
CCAATTTCTCTGATCCATTAAGTTCCCATTCAGGACTCTGGAGGATGTGTAAATTTAAGAACAAATGTCTCCCTCTG
ATAAATCCCTTTCGGCTTGGGAACTTAAACTTCACTGATTCACAGAAACAGCTTTTGAGTATGCATGGAACACTAGT
GATTTTACTACCACTTAGCCTAATCTTGATGATATTTGGAGGAATGACTGGGTTTGTTAGTATTCTTGCAAGAGCTT
ATCTACTTCTGCTTTTGACCGGAATGCTTTTCCTGTTTGGCGCACTAGTAACATTAACGGGAATCAGCATCTACATT
GCCTACTCTGCTGCTGCATTTAAAGATGCCGTTTGCATCTTAGGGAACAAAATTCTGGAAGATATTGACATACAATT
TGGCTGGTCCTTGGCTCTGGCATGGATATCATTTATTACAGAGATTCTCACAGGGATAGCCTTCCTGGTGGCTGCCA
GAGTAACTGGGTTAAAAAGAAGAAGACGGGAGCAGGTTATAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGAT
TCTACG**TGA**

B

24.8 kDa
MKLKLS**M**LSVFVAVVGILSFISLVVAIGTDFWYIIDASRLEKITNFSDPLSSHSGLWR**M**CKFKNKC
LPLINPFRLGNLNFTDSQKQLLS**M**HGTLVILLPLSLIL**M**IFGG**M**TGFVSILARAYLLLLLTG**M**LFL
FGALVTLTGISIYIAYSAAAFKDAVCILGNKILEDIDIQFGWSLALAWISFITEILTGIAFLVAAR
VTGLKRRRREQVIGKPIPNPLLGLDST

24 kDa
MLSVFVAVVGILSFISLVVAIGTDFWYIIDASRLEKITNFSDPLSSHSGLWR**M**CKFKNKCLPLINP
FRLGNLNFTDSQKQLLS**M**HGTLVILLPLSLIL**M**IFGG**M**TGFVSILARAYLLLLLTG**M**LFLFGALVT
LTGISIYIAYSAAAFKDAVCILGNKILEDIDIQFGWSLALAWISFITEILTGIAFLVAARVTGLKR
RRREQVIGKPIPNPLLGLDST

*18.3kDa
MCKFKNKCLPLINPFRLGNLNFTDSQKQLLS**M**HGTLVILLPLSLIL**M**IFGG**M**TGFVSILARAYLLL
LLTG**M**LFLFGALVTLTGISIYIAYSAAAFKDAVCILGNKILEDIDIQFGWSLALAWISFITEILTG
IAFLVAARVTGLKRRRREQVIGKPIPNPLLGLDST

*14.7 kDa
MHGTLVILLPLSLIL**M**IFGG**M**TGFVSILARAYLLLLLTG**M**LFLFGALVTLTGISIYIAYSAAAFKD
AVCILGNKILEDIDIQFGWSLALAWISFITEILTGIAFLVAARVTGLKRRRREQVIGKPIPNPLLG
LDST

13.1 kDa
MIFGG**M**TGFVSILARAYLLLLLTG**M**LFLFGALVTLTGISIYIAYSAAAFKDAVCILGNKILEDIDI
QFGWSLALAWISFITEILTGIAFLVAARVTGLKRRRREQVIGKPIPNPLLGLDST

*12.6 kDa
MTGFVSILARAYLLLLLTG**M**LFLFGALVTLTGISIYIAYSAAAFKDAVCILGNKILEDIDIQFGWS
LALAWISFITEILTGIAFLVAARVTGLKRRRREQVIGKPIPNPLLGLDST

*10.5 kDa
MLFLFGALVTLTGISIYIAYSAAAFKDAVCILGNKILEDIDIQFGWSLALAWISFITEILTGIAFL
VAARVTGLKRRRREQVIGKPIPNPLLGLDST

Figure 6.8. Possible translation start sites in pGEM-xtTmem114-V5 and the predicted protein products which may produce in-frame V5-tagged proteins. A) xtTmem114 (upper case) was *in vitro* transcribed from the T7 promoter (bold) of pGEM T easy (lower case). Blue = 5'UTR sequence. Green highlight = Translation start site for xtTmem114. Yellow highlight = possible alternative start sites that encode in-frame products. Red highlight = possible alternative start site that encodes out-of-frame products. TGA = stop codon. Underline = V5 tag. B) Translated sequences from putative transcripts which encode V5 tags and their predicted molecular weight. * Predicted to represent the bands detected by western blot. Underline = V5 tag. Sequences translated using Expasy translate tool (http://www.expasy.ch/tools/dna.html). Molecular weight predicted by the Expasy pI/Mw tool (http://www.expasy.ch/tools/pi_tool.html).

6.3.2 Morpholino 2

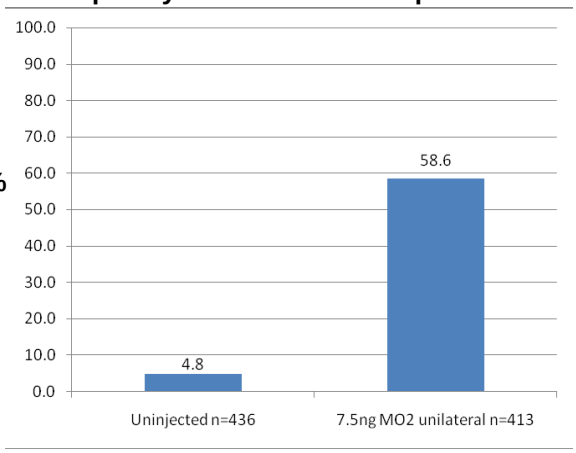
A second morpholino (MO2) designed to bind immediately upstream of the binding site for MO1 was generated to further confirm the phenotype observed (Figure 6.9A). Embryos unilaterally injected with 7.5 ng MO2 displayed unilateral microphthalmia at a frequency of 59 % compared to 4.8 % in uninjected embryos (Figure 6.9 B,C,D). Gut and/or eye abnormalities were observed in more than 81 % of MO2 injected in comparison with 19 % of uninjected (Figure 6.10). The level of gut abnormalities (13.8 %) was consistently higher than the levels of unilateral microphthalmia observed (4.8 %) in uninjected embryos. This level was relatively low when compared to the levels in the MO2 injected embryos (51 %).

A

aacacaggaacaaguggauuccucuucuguaucugca**AUGAAGCUGAAACUGAGUAUGUUAUCUGUCUUU**
TGTTCACCTAAGGAGAAGACATAGA

B

C

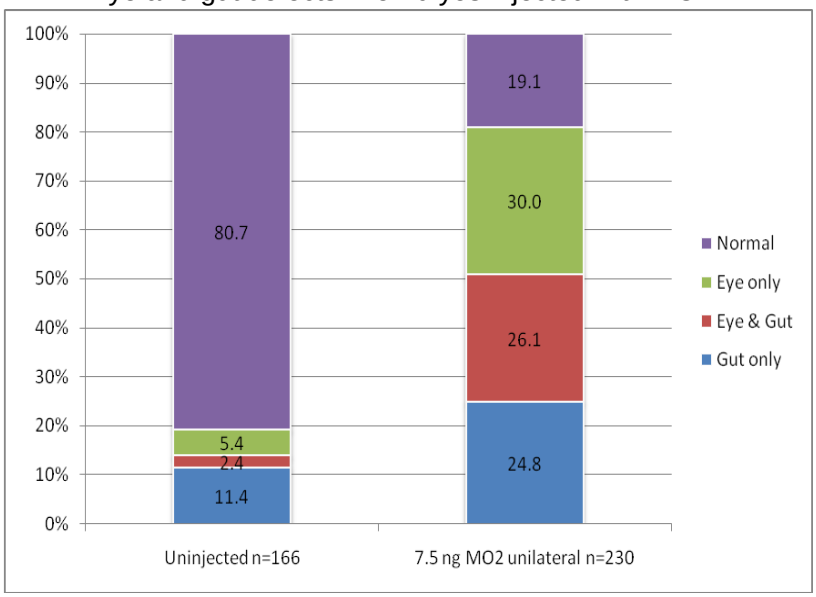


D

	Uninjected			7.5ng unilateral		
	n=	eye defect	%	n=	eye defect	%
Run1	54	3	5.5	163	91	55.8
Run2	158	5	3.2	93	53	57.0
Run3	112	5	4.5	90	58	64.4
Run4	51	2	3.9	20	12	60.0
Run5	61	6	9.8	47	28	59.6
Combined	**436**	**21**	**4.8**	**413**	**242**	**58.6**

Figure 6.9. Second morpholino (MO2) directed against the 5'UTR of *Tmem114*. A) Binding sites of MO2. MO2 (blue) binds the 5'UTR (lower case) of *xtTmem114* four bases upstream of the translation start site (AUG). Coding sequence of *xtTmem114* is in upper case and bold. **B)** Injected Embryo injected unilaterally with 7.5 ng MO2 displaying microphthalmic eye (arrowhead). **C)** Frequency of unilateral microphthalmia observed in embryos unilaterally injected with 7.5 ng MO2. **D)** Data for the 5 independent experiments.

A



B

Uninjected

	n=	eye	gut	eye & gut	normal
Run1	54	3.7%	5.5%	1.9%	88.9%
Run4	51	3.9%	19.6%	3.9%	72.6%
Run5	61	8.1%	9.8%	1.6%	80.3%
Combined	**166**	**5.4%**	**11.4%**	**2.4%**	**80.7%**

7.5ng unilateral

	n=	eye	gut	eye & gut	normal
Run1	163	28.2%	27.6%	27.6%	16.6%
Run4	20	15.0%	15.0%	35.0%	35.0%
Run5	47	42.6%	19.1%	17.0%	21.3%
Combined	**230**	**30.0%**	**24.8%**	**26.1%**	**19.1%**

Figure 6.10. Gut defects in MO2-injected embryos. A) Frequency of eye and gut defects observed in embryos unilaterally injected with 7.5 ng MO2. B) Data for the 3 independent experiments. Eye only = unilateral eye defects with no other abnormalities. Gut only = Abnormal gut development with no other abnormalities. Eye and gut = Unilateral microphthalmia with gut abnormalities.

6.3.3 Structure of the microphthalmic eyes

To investigate the effect of knockdown of xtTmem114 expression on the structure of the eye, embryos which were unilaterally injected with 7.5 ng of MO2 were sorted into those displaying unilateral microphthalmia of the right eye and those displaying unilateral microphthalmia of the left eye. The sorted embryos were embedded in paraffin, sectioned and stained with haematoxylin and eosin (H&E). Comparing both eyes of the microphthalmic embryos revealed no observable difference between the uninjected and injected eye other than overall eye size (Figure 6.11). The lenses were clear and comparable in size although the finer cellular

structure of the lens could not be determined in the sections obtained. In the retina, there was no difference in the structure of the ganglion cell layer, inner and outer plexiform layer, inner and outer nuclear layer, the outer segment of the photoreceptors or in the retinal pigmented epithelium (RPE). No differences were observed in the cornea.

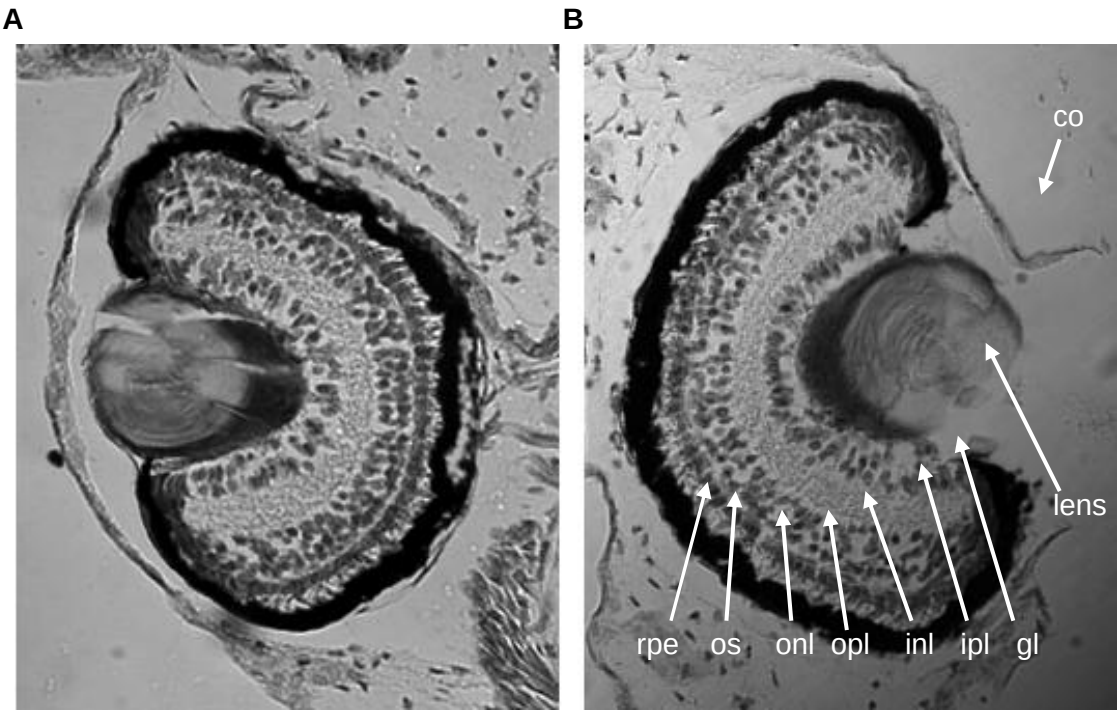


Figure 6.11. Sectioned microphthalmic (A) and normal (B) eye of a stage 43 embryo. No structural differences are detectable between the two eyes. Co = cornea, le = lens, gl = ganglion layer, ipl = inner plexiform layer, inl = inner nuclear layer, opl = outer plexiform layer, onl = outer nuclear layer, os = outersegments of photoreceptors, rpe = retinal pigmented epithelium. Images were taken at 20 x magnification.

6.3.4 Attempt at rescue of phenotype

Murine *Tmem114* (*mTmem114*) differs from xtTmem114 in the binding regions of the *xtTmem114* MOs (Figure 6.12A). Therefore, the morpholinos should have no effect on translation of *mTmem114*. If mTmem114 and xtTmem114 are functionally conserved then heterologous expression of mTmem114 should compensate for the loss of endogenous xtTmem114 expression caused by the morpholinos and rescue the phenotype observed in the MO injected embryos. Before attempting a phenotypic rescue *mTmem114-V5* was overexpressed in embryos to determine if such expression was tolerated by the embryos. Varying amounts of V5-tagged *mTmem114* were injected into embryos to determine at what dose no observable phenotypes were detected. A marked increase in the number of embryos failing to develop past gastrulation was observed. At doses of 100 pg and 200 pg, 37 % and 61 % of embryos, respectively, displayed gastrulation-specific defects such as exogastrulation (Figure 6.12B). Higher frequencies of grossly abnormal or dead embryos were also observed.

A

```
xtTmem114 aacaaguggauuccucuucuguaucugcaAUGAAGCUGAAACUGAGUAUGUUAUCUGUC
          TGTTCACCTAAGGAGAAGACATAGACGTTACTTCGACTTTGACTCATACA
mTmem114  gccggcgcccgggggcccaacgcugggagAUGCGGGUGCGCCUGGGCGCUCUGGCGGGA
            *           *       *  *  ***  * **   *** *     *  * *
```

B

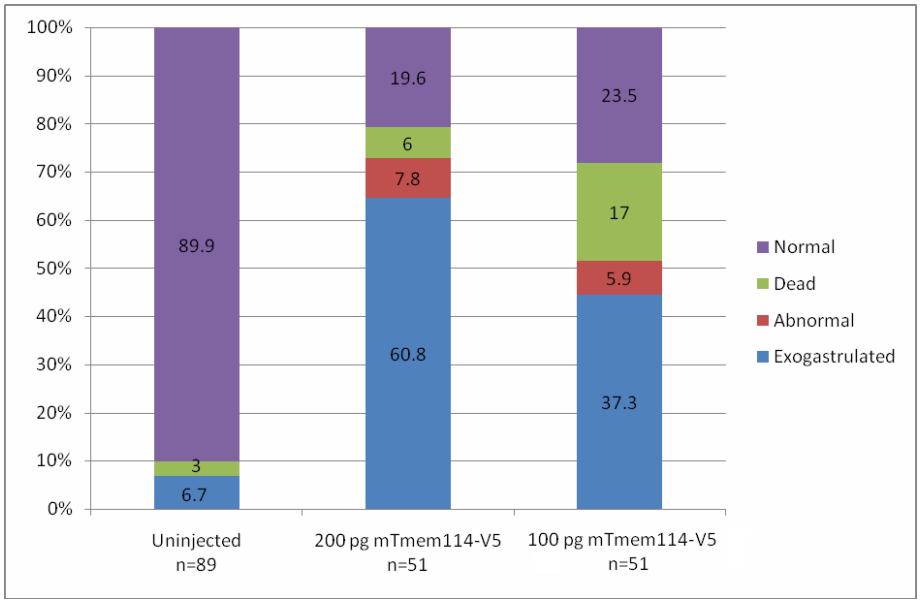


Figure 6.12. Overexpression of mTmem114-V5. A) Comparison of the region surrounding the translation start sites of *xtTmem114* and *mTmem114*. Identical residues are marked by an asterisk. The binding site for MO1 (red) in xtTmem114 has 11 identical nucleotides of 25 when compared to mTmem114. The binding site for MO2 (blue) in the 5'UTR (lower case) shows little conservation. B) Phenotypes observed with injection of mTmem114-V5. A marked increase in the number of abnormal and exogastrulated embryos was observed with injection of *mTmem114-V5*. Data presented is from a single experiment.

To determine if the observed phenotypes correlated with *mTmem114-V5* expression, uninjected and *mTmem114-V5* injected embryos were lysed and protein expression was detected by western blot. Prior to western blotting, the membrane was stained with Ponceau S confirming protein transfer to the membrane (Figure 6.13A). Subsequently, V5 expression was analysed by western blot. All samples tested were negative for V5 indicating the injected mRNA was not being translated (Figure 6.13B).

To check the integrity of the mRNA injected, total RNA was extracted from injected embryos and cDNA was made. The presence of *mTmem114-V5* was detected by RT-PCR using primers specific for *mTmem114* (Figure 6.13C). *mTmem114* was detected in the cDNA from injected samples, but not in the uninjected samples or in negative control samples to which RNA was added but reverse transcriptase was not.

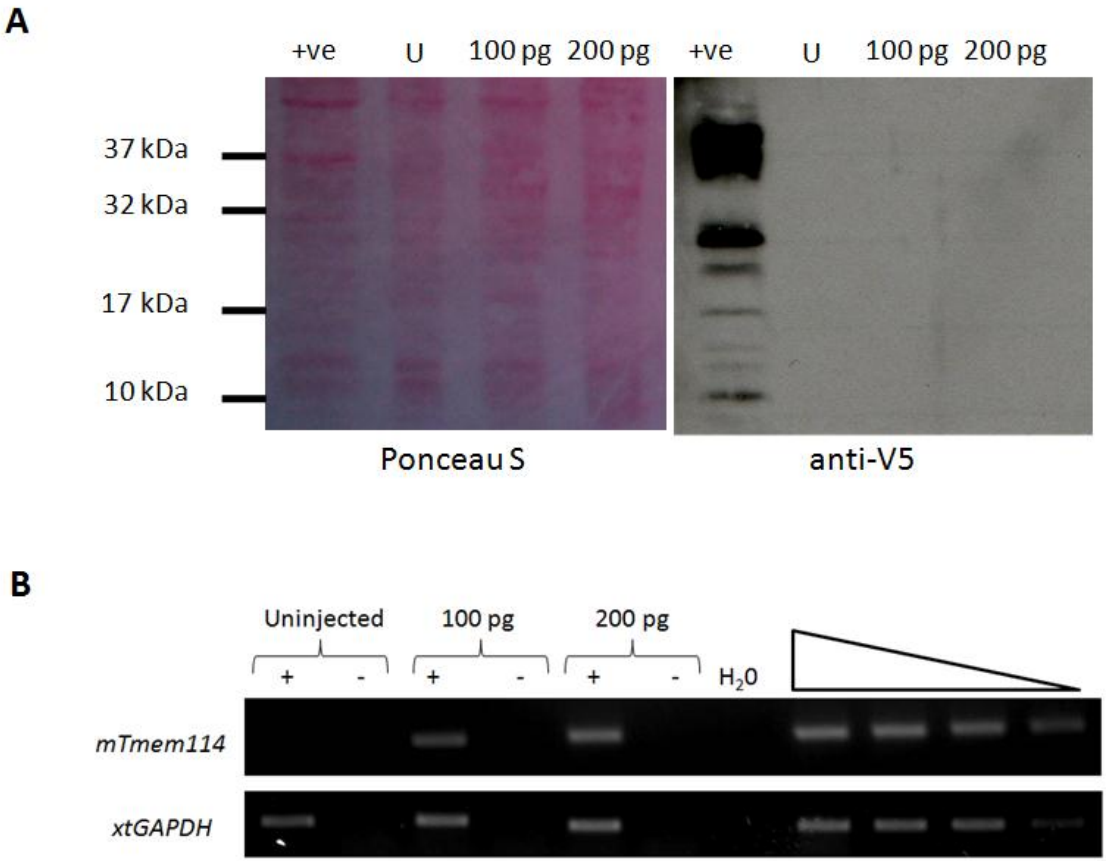


Figure 6.13. Detection of the overexpression of *mTmem114-V5* mRNA. A) Detection of mTmem114-V5 by western blot. mTmem114-V5 was not detected in the injected (100 pg or 200 pg RNA) or uninjected (U) embryos, but was detected in the lysate of mTmem114-V5 transfected cells. The Ponceau S stained membrane demonstrates that similar levels of protein were present in all samples. B) Detection of *mTmem114-V5* by RT-PCR with primers amplifying exons 1 and 2. Expression of *mTmem114-V5* was detected in the embryos injected with 100 pg *mTmem114-V5* mRNA and higher levels of expression were detected in the embryos injected with 200 pg *mTmem114-V5* mRNA. Expression was not detected in the uninjected embryos. Detection of the housekeeping gene *GAPDH* was used as a control for equal cDNA loading in the RT-PCR reactions. The endpoint of all PCRs was determined to be within the exponential phase as demonstrated by the linearities for each reaction (triangle). H_20 represents a no template control.

6.4 Discussion

6.4.1 Early developmental expression of *xtTmem114*

Expression of *xtTmem114* was detected at specific stages of development (Figure 6.1). In early *Xenopus* development RNA synthesis does not occur until the midblastula stage (stage 8). Until this point, protein expression is dependent on maternally inherited mRNA and protein (Newport and Kirschner 1982). The maternally expressed mRNAs are translated at specific times which are thought to be associated with the adenylation state of the mRNAs (Graindorge *et al.* 2006). The maternal expression of *Tmem114* detected in *X. tropicalis* had not previously been described in other species. This early expression in *X. tropicalis* is

suggestive of a functional role prior to gastrulation in which case the injection of morpholinos may disrupt blastula development. A reduction in embryo survival was observed with the increase in MO dose (Figure 6.3C), which may suggest that knockdown of Tmem114 was lethal to the embryos, but may have been due to non-specific effects caused by the injection or by non-specific toxicity associated with the high concentrations of morpholino, or by poor quality of the embryos (discussed in Section 6.4.4).

When *mTmem114-V5* was heterologously expressed in developing embryos gastrulation defects were observed. This suggested that expression of Tmem114 may inhibit gastrulation. Gastrulation involves the organisation of cells of the blastula into the three germ layers: endoderm, mesoderm and ectoderm. This involves highly co-ordinated cell movement which is dependent on cell-cell and cell-extracellular matrix adhesion (Steinberg 1996). Improper expression of a protein with adhesive properties may disrupt the co-ordinated cell movement involved in gastrulation. Overexpression of claudins in *Xenopus* increases cell-cell adhesion and impairs gastrulation (Brizuela *et al.* 2001; Chang *et al.* 2010). Tmem114 is similar to a number of proteins with adhesive properties (Section 3.10.2.6) and therefore, expression of Tmem114 at gastrulation, when it is not normally expressed, may alter the cell-cell adhesions, affecting the movement of cells necessary for successful gastrulation. However, the link between mTmem114-V5 overexpression and exogastrulation could not be proved. Although the injected *mTmem114-V5* mRNA appeared stable (as it was detected by RT-PCR) translation to Tmem114-V5 could not be detected by western blot. This indicated that the phenotype observed may have been due to toxicity of the injected mRNA or due to the injection procedure itself, and not due to expression of mTmem114. Attempts were also made to inject plasmid DNA encoding *mTmem114-V5* but this also was not detected by western blot (data not shown).

With further time and resources the overexpression of mTmem114-V5 could have been optimised. Once overexpression had been successfully achieved, a rescue of the phenotype could be attempted. Rescue of a phenotype by expression of the target mRNA is the most stringent test of specificity of the phenotype observed in knockdown models. A more extensive investigation of the effect of xtTmem114 MOs on early development may identify the early functional role. The dissection of the early embryos into animal and vegetal hemispheres may also detect if expression of xtTmem114 is localised to specific poles.

6.4.2 Zygotic expression of *xtTmem114*

Weak zygotic expression of *xtTmem114* began at stage 23, approximately the time of interaction between the presumptive lens ectoderm (PLE) and the optic vesicle (Ishibashi and Yasuda 2001; Henry *et al.* 2008). Expression in the eye at later developmental stages (45, 47) was shown by RT-PCR, but expression at earlier stages could not be tested by RT-PCR as the tissues were too small to dissect. Attempts were made to detect expression by *in situ* hybridisation, but although the protocol worked, as shown by the staining for *xtLim2*, specific

staining for *xtTmem114* could not be detected. The background staining seen with both antisense and sense probes was suggestive of trapping of reagents in cavities ("probe trapping") which may have been due to low levels of expression of *xtTmem114* or due to the chemistry of the probes used. A similar staining pattern was observed with the sense *xtLim2* probe confirming that the staining observed was non-specific.

Expression of *xtTmem114* restarts at stage 37 at higher levels than earlier developmental stages. This may represent upregulation of expression in the eye, as seen in mouse (Jamieson *et al*. 2007) or may represent a more widespread expression of the gene. The increase in expression coincides with two major developments in the eye: at stage 35/36, the cavity of the lens vesicle begins to be filled by the elongating primary fibre cells and the retina begins to differentiate into its specific layers (Henry *et al*. 2008). By stage 41 the primary lens cells have filled the lumen of the lens vesicle (Henry *et al*. 2008), the equivalent of E13.5 in mouse (Faber *et al.* 2002). Expression of murine *Tmem114* begins at E13.5 and increases thereafter (Jamieson *et al*. 2007). Expression in *Xenopus* therefore occurs slightly earlier than in mouse, but expression is increased at stage 40, approximately the equivalent stage of onset of murine expression.

The earliest detection of *mTmem114* expression in the developing mouse eye is detected at E13.5 (Jamieson *et al*. 2007), the equivalent stage of development as stage 41 in *Xenopus*. In *Xenopus*, the expression at stages 23-26 (the time of PLE/optic vesicle contact) was much lower than at 37-40. Therefore, expression in the developing PLE in mouse may also be occurring, but was not detected due to the low levels of expression.

Expression of *xtTmem114* was also detected in the adult intestine, but developing tissues were not tested. The major development of the digestive system takes place form stage 36 onwards (Estabel *et al.* 2003) and at stage 40 the gut is formed of a simple tube, with the formation of coiled loops occurring at approximately stage 43 (Chalmers and Slack 1998). In the adult *xtTmem114* is also expressed in the lung. In the developing embryo external gills begin to develop at stage 36 (Segerdell *et al.* 2008) and at stage 46 the external gills are replaced by internal gills which later in development are replaced by lungs (Estabel *et al*. 2003).

Developmental, but not adult, expression of murine Tmem114 has been reported in kidney and developmental kidney defects were identified in patients with the balanced translocation through the putative promoter of *TMEM114* (Jamieson *et al*. 2007). Expression in the kidney of adult *X. tropicalis* was not detected. During kidney development in *X. tropicalis*, the pronephros begins to function at stage 36 (Chan and Asashima 2006) and this is followed by a wave of gene expression (Brandli 1999). This coincides with the onset of *xtTmem114* expression, suggesting a possible role in kidney development. The failure of the *in situ* hybridisation means that this hypothesis could not be tested.

Thus, the onset of *xtTmem114* expression at stage 37 may represent expression in the eye, kidney, gill or intestine or may represent expression in a number of these developing tissues. Determination of spatial expression of *xtTmem114* at this stage will require *in situ* hybridisation.

6.4.3 Knockdown of xtTmem114 results in microphthalmia

The xtTmem114 MO injected embryos displayed the developmental ocular defect microphthalmia. Microphthalmia was observed in 39 % of MO1 unilateral injected embryos and 59 % of MO2 unilateral injected embryos, with background (uninjected) frequencies of 2.9-4.8 %. Although background frequencies were higher than from other published work from the laboratory (1.8-2.4 %) (Hilton *et al*. 2007), frequencies were in line with other published work. Other published work reporting eye defects in the range of 65-77 % reported frequencies of 3-9 % in control MO injected embryos (Elkins and Henry 2006; Hanel and Hensey 2006; Wang *et al*. 2009b).

Although phenotypes were scored after stage 40 when the eyes of the embryos were relatively large, presentation of the phenotype preceded stage 37, as shown in the EGFP/MO1 injected embryos (Figure 6.5). This suggests that the phenotype was due to knockdown of expression at the earlier stages of 23-26, rather than when expression restarts at stage 37. This implies that *xtTmem114* is expressed in the eye at the stages of 23-26. In *Xenopus* the PLE is in contact with the optic vesicle at these stages and its thickening to form the lens placode occurs at stage 27 (Ishibashi and Yasuda 2001). Thus *xtTmem114* expression precedes lens placode formation but occurs at the time of its precursor tissue.

Impaired lens formation is a major cause of microphthalmia (Landel *et al.* 1988; van Raamsdonk and Tilghman 2000) and in humans seven of the eight genes which are known to cause isolated microphthalmia/anophthalmia encode transcription factors required for lens formation (reviewed in Becker *et al.,* 2010). Reduction of the size of lens placode caused by haploinsufficiency for *Pax6,* a gene essential for lens placode formation (Grindley *et al.* 1995) results in the formation of a small lens causing microphthalmia (van Raamsdonk and Tilghman 2000). Cases of microphthalmia in which lens development is defective at later stages, such as the disrupted elongation of lens fibre cells (Oda *et al.* 1980), or a decrease in the number of fibre cells (White *et al.* 1998) also decreases the size of the lens relative to the remainder of eye.

As the lens placode is key to development of the whole eye, disrupting development of the PLE may have resulted in the overall reduction in eye size. Disruption of PLE development by a conditional knockout of the transcription factor *Six3* prevents development of the lens placode and lens (Liu *et al*. 2006b). The lenses of the xtTmem114 MO microphthalmic eyes were of normal size relative to the remainder of the eye, i.e., the affected eye showed a

reduction in size in all tissues and not specifically the lens. This suggests that the defect in development caused by the MOs delayed growth of the whole eye rather than inhibiting development of the lens.

TMEM114 is associated with early onset dominant cataract through its proximity to a balanced translocation (Jamieson *et al*. 2007). The breakpoint of the translocation transects the predicted promoter of *TMEM114*. However, no association was identified between coding sequence variants in *TMEM114* and autosomal dominant congenital cataract (Jamieson *et al*. 2007). The microphthalmic lenses of Tmem114 MO injected embryos showed no sign of opacification and cataract was not observed in any of the other morphants. Cataract, although rarely reported, is detectable in MO-injected *Xenopus* (Dirscherl *et al.* 2005). Thus, coding sequence variants in *TMEM114* and knockdown of Tmem114 have not been shown to result in cataract. However, the effect of the balanced translocation on *TMEM114* expression is unknown. If the effect was to inhibit expression of *TMEM114*, one may have expected to observed cataract in the morphants. As this did not occur it is possible that the translocation has the effect of dysregulating expression of *TMEM114*, with expression or over-expression occurring at the wrong time or place. The effect of mutations on promoter activity can be tested with luciferase assays, with luciferase expression controlled by the promoter of interest. Although both wildtype and mutant promoter fragments could have been cloned, the lack of a suitable cell system in which to perform the luciferase assays prevented the experiment from being pursued. As described in section 4.2.2, endogenous *TMEM114* expression was not detected in any of the cell lines tested.

The knockdown/knockout of a number of genes in which missense mutations in humans result in cataracts causes microphthalmia. Missense mutations in the gap junction protein connexin 50 (Cx50) are associated with a variety of forms of dominant congenital cataract (Schmidt *et al*. 2008; Vanita *et al*. 2008; Wang *et al*. 2009a; Gao *et al*. 2010). Cx50 knockout mice display congenital cataract and microphthalmia (White *et al*. 1998; Rong *et al*. 2002). Microphthalmia also presents in some mice and rats homozygous for missense mutations in Cx50 (Liska *et al.* 2008; Bakthavachalu *et al.* 2010). Knockout of *Pitx3* results in microphthalmia due to failure of the lens to develop (aphakia) (Rieger *et al.* 2001) and homozygous frameshift mutations in *PITX3* cause microphthalmia and neurological defects (Bidinost *et al.* 2006). Knockout of αA-crystallin in mouse results in cataract and microphthalmia (Brady *et al.* 1997). Interestingly homozygous missense mutations affecting residue p.R54 of αA-crystallin in mouse have different effects. Both p.R54C and p.R54H (Chang *et al.* 1999) mice display cataract but the p.R54C mouse also displays microphthalmia (Xia *et al.* 2006). The substitution of histidine for arginine replaces a positively charged amino acid with another positively charged amino acid. But, altering the positively charged arginine (R) to the non polar cysteine results in the more severe phenotype microphthalmia. This suggests that loss of protein or proteins with severe amino acid alterations are associated with microphthalmia.

In conclusion, knockdown of xtTmem114 in *Xenopus* results in microphthalmia and biallelic missense or non-sense mutations in a number of genes expressed in the lens are associated with microphthalmia. Therefore one may expect to find biallelic missense or non-sense mutations in *TMEM114* associated with microphthalmia in humans. This hypothesis is tested in Chapter 7.

6.4.4 Knockdown of xtTmem114 is associated with an increase in gut defects

An increased number of embryos displaying abnormal guts were observed with injection of MO1. This was followed up with three independent experiments with injection of MO2 also resulting in an increased frequency of abnormal gut development. Expression of *xtTmem114* was detected in the adult intestine supporting a role for xtTmem114 in this tissue. As discussed above there was a marked increase in expression of xtTmem114 from stage 37. At this stage the gut is simple but is beginning to develop. By stage 43 the gut has lengthened, has begun to forms its cavity, and has formed its coiled structure which further develops with time (Chalmers and Slack 1998; Chalmers and Slack 2000). Although the increase in expression of *xtTmem114* coincides with this stage of gut development, the failure of the guts to coil in the MO injected embryos is not likely to be due to inhibition of expression at such late a stage of development. MOs become more dilute as the embryo grows and so phenotypic effects are usually due to inhibition at earlier stages. Gut coiling has also been shown to sensitive at earlier stages of development, particularly at stage 25 (Zeynali and Dixon 1998) which coincides with the earlier expression of *xtTmem114*.

Even though gut development is well documented and used as a marker for embryo staging (Nieuwkoop and Faber 1967) there was consistently a high background rate of gut defects in the uninjected embryos (7.4 %-23.5 %). Previous work published by this laboratory had background gut defect frequencies of 5-6 %, compared to background levels of eye defects of <1 % (Hilton *et al*. 2007). This suggests that gut development was particularly sensitive to perturbation. To determine if the observed increase in frequency of gut abnormalities was due specifically to the xtTmem114 MOs a number of control experiments could have been performed. Embryos could have been injected with the same concentration of control morpholino to determine if the effect was due to injection of oligonucleotides or injected with water to determine if the effect was due to the injection procedure itself.

Although the expression in the adult intestine is suggestive of expression in the developing intestine, this developmental expression has not been determined. Also, the effect of morpholinos decreases with the number of cell divisions occurring as the morpholino becomes more dilute with each cell division. Thus, when xtTmem114 expression is at its peak (~stage 40), the morpholino is likely to be nearing its most dilute state. Also, *Tmem114* expression has only been detected in epithelial tissues, but although gut development begins earlier, the intestinal epithelium is not formed until stage 45 (Chalmers and Slack 2000). Thus there are a number of questions that need answering before the association between knockdown of xtTmem114 and defects in gut development is confirmed.

It is common that the quality of embryos varies from batch to batch, especially if the females used are old or not ovulating regularly (Tomlinson *et al.* 2005). This can affect the numbers of embryos surviving through gastrulation and those that do may have an increased rate of abnormalities or be more susceptible to perturbation. For this reason data from three or more independent experiments using different females are performed. The initial dose optimisation experiments for MO1 using 15 ng and 20 ng injections had low survival rates of the uninjected (~50 %) and very low survival rates for the injected (~ 15 %). But these doses were only tested once. Based on this data 10 ng was chosen as the optimal dose (5 ng for unilateral injections). However, repetition of the higher doses may have identified that these doses are tolerable to the embryos and an increased dose may have increased the numbers of embryos displaying microphthalmia from the moderate but significant level of 39 %.

Using MO1, 39 % of embryos displayed microphthalmia and using a higher dose of MO2, 59 % of embryos displayed microphthalmia. Thus, complete penetrance of the phenotype was not observed. However, as the majority (80 %) of MOs from Gene Tools are reported to achieve 70-98 % knockdown of their specific targets (Summerton 2007), one cannot expect to achieve phenotypes in 100 % of MO injected embryos. As discussed above, the lower rate of microphthalmia with MO1 may be due to using a sub-optimal dose. The co-injection of EGFP with the MOs routinely would have allowed selection of embryos specifically with EGFP/MO located in the early developing eye which may have increased the frequency of microphthalmia detected. However, the co-expression of EGFP may have been a confounder so this was not performed. Injections can also be performed at later stages, e.g. injection into the animal region of dorsal blastomeres at the 16 cell stage specifically targets the brain, eyes and olfactory tissues (Moody 1987; Ciesiolka *et al.* 2004). But this strategy requires more time and expertise and reduces the number of injections being performed.

6.4.5 Limitations of this study

There were two major limitations to this study: the failure of the *in situ* hybridisation for *xtTmem114* and the inability to confirm that MO1 and MO2 block translation of their targets. Although the temporal expression of *xtTmem114* was successfully determined in this study, there was only limited spatial expression determined by RT-PCR. A successful *in situ* hybridisation would have aided this study in numerous ways. The expression at stages 23-26 is likely to occur in the PLE due to the expression of *mTmem114* in the lens and that microphthalmia is observed before the later onset of expression at stage 37. Wholemount *in situ* hybridisations at a range of stages from 20-30 would have confirmed this hypothesis as true or false. There is large increase in expression of *xtTmem114* at stage 37 which coincides with developmental processes in the eye, kidney and intestine. A functional *in situ* hybridisation at this stage would have determined the tissue(s) in which it is expressed at this stage.

Due to the lack of an antibody that detected xtTmem114 the knockdown of xtTmem114 could not be shown *in vivo*. Therefore an alternative approach based on the *in vitro* translation of V5-tagged xtTmem114 was used. However, *xtTmem114* contains a number of potential ATG initiation codons, and when detected by western blot, it was apparent that translation was being initiated downstream of the "true" initiation codon. Therefore, the addition of MO1 had no effect as the ribosome was binding at an alternative site. If time permitted an alternative approach could have been taken to prove that the MOs bind their targets and knockdown expression. The 5'UTR of *xtTmem114* could have been cloned upstream of a protein such as GFP. As shown in figure 6.5, *X. tropicalis* embryos readily transcribe and translate GFP from injected circular plasmids. The ability of MO2, which binds the 5'UTR of *xtTmem114* to knockdown expression of its target could be tested by co-injecting the xtTmem114-5'UTR/GFP construct with and without MO2. As a further control, MO2 should not prevent translation of GFP which lacked the 5'UTR of *xtTmem114*. This method has previously been used to prove efficacy of morpholinos *in vivo* (Jia *et al.* 2008). However, although the knockdown of xtTmem114 was not demonstrated, the observation of similar phenotypes with MO1 and MO2 which have independent targets was suggestive of specificity.

Another limitation of the study was that a control MO was not used. Injection of scrambled MOs or standard control MOs available commercially act as a better control than using uninjected embyros as it controls for the injection procedure itself and it acts as a control for the injection of the 25bp oligonucleotide. However, injecting control MOs in the same experiments as the test MOs halves the number of test embryos injected. Previous MO work in the laboratory showed little difference in the rates of defects between MO, water and uninjected embryos (Hilton *et al*. 2007). For this reason, control MOs were not employed.

6.5 Conclusion

The use of *X. tropicalis* as an animal model has confirmed the role of Tmem114 in eye development but has also uncovered questions that remain to be answered. Maternal expression of *xtTmem114* was detected but the functional role for this is yet to be determined. The conservation of this maternal expression also remains to be identified. The data generated suggests that xtTmem114 is expressed in the early stages of eye development, earlier than reported in mouse (Jamieson *et al*. 2007). This has led to the hypothesis that TMEM114 may be associated with microphthalmia in humans, a hypothesis which is tested in chapter 7. The detection of expression in the intestine and the lung suggests a more widespread expression pattern than previously detected. In the lens *Tmem114* is expressed in the epithelial cell layer (Jamieson *et al*. 2007). The lung and the intestine are also epithelial tissues. If the hypothesis that TMEM114 is associated with microphthalmia in humans is correct, one may expect to find the microphthalmia occurring in association with other non-ocular anomalies.

Chapter 7: Mutational screening of *TMEM114*

7.1 Introduction

Knockdown of Tmem114 in *Xenopus tropicalis* resulted in microphthalmia (Chapter 6) and the expression of *TMEM114* in the developing human eye (Figure 4.1) suggested a role in eye development. *TMEM114* is located at a locus associated with isolated microphthalmia with cataract (MCOPCT1, OMIM 156850). Therefore mutations were sought in a panel of patients with microphthalmia to determine if mutations in *TMEM114* are associated with microphthalmia in humans.

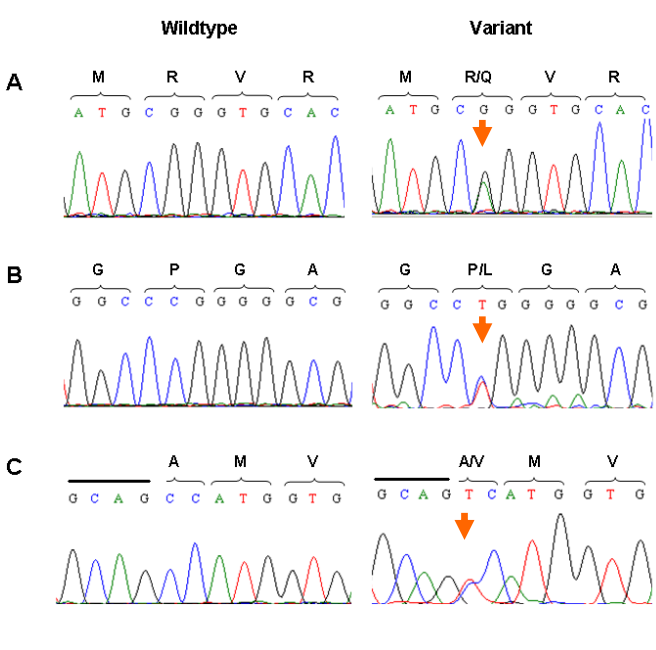
7.2 Screening of microphthalmia cohort for mutations in *TMEM114*

This panel consisted of 77 patients with the congenital structural eye defects microphthalmia (small eye), anophthalmia (absence of ocular tissue in the orbit) and coloboma (failure of part of the fetal fissure to close) (Table 7.1). Some of these phenotypes were in isolation, but the majority of patients had other ocular and non-ocular phenotypes (Full list of clinical phenotypes available in Table 9.14). The microphthalmia/anophthalmia/coloboma (MAC) patients were screened for mutations in the coding sequence and flanking intronic sequences of *TMEM114*. Three patients were identified to have heterozygous missense variants (Figure 7.1) and a number of SNPs were identified in patients and in controls (Table 7.2).

Table 7.1. Clinical phenotypes of 77 MAC patients screened for mutations in *TMEM114.*

n = number of patients. NS = number of patients with non-syndromic (eye only) phenotypes. Total = combined numbers of bilateral cases (unilateral anophthalmia with contralateral anophthalmia are considered bilateral microphthalmia).

Phenotype	n	NS	Total
Bilateral anophthalmia	13	6	13
Unilateral anophthalmia with contralateral microphthalmia	12	7	48
Unilateral anophthalmia with contralateral microphthalmia & coloboma	2	1	
Bilateral microphthalmia	19	8	
Bilateral microphthalmia with bilateral colobomata	11	9	
Bilateral microphthalmia with unilateral coloboma	4	3	
Unilateral microphthalmia	3	1	
Unilateral microphthalmia with contralateral coloboma	4	1	
Bilateral colobomata	3	2	8
Bilateral colobomata with unilateral microphthalmia	5	5	
Unilateral coloboma	1	0	



Phenotype

c.5g>a p.R2Q

> bilateral microphthalmia, retinal coloboma
> absent right kidney, cystic dysplastic left kidney, bilateral cryptorchidism, adducted left thumb, pseudoarthrosis of right clavicle
> Ethnicity: Vietnamese

c.134c>t p.P45L

> bilateral microphthalmia, colobomas, retinal dystrophy.
> Ethnicity: unknown

c.440c>t p.A147V

> bilateral anophthalmia, optic nerve hypoplasia, atrophic chiasm
> interhemispheric cyst, corpus callosum agenesis, Dandy Walker malformation, seizures, dysplastic cup shaped ears, extended abdomen.

Figure 7.1. Missense variants identified in patients with microphthalmia. A) Heterozygous G>A change at position 5 (arrow) resulting in substitution of glutamine for arginine B) Heterozygous C>T change at position 134 (arrow) resulting in substitution of lysine for proline. C) Heterozygous C>T change at position 440 (arrow) resulting in substitution of valine for alanine. Bar indicates intronic sequence. Phenotypes and ethnicity of the affected individuals are presented on the right.

The variant c.134C>T, resulting in the substitution of non-polar proline with the non-polar leucine (p.P45L), was identified in a patient with bilateral microphthalmia, coloboma and retinal dystrophy (Figure 7.1 B). The affected proline residue, located in ECL1, is conserved in mammals, although the surrounding amino acids show little conservation (Figure 7.2). Parental DNA was unavailable for testing. This variant was identified in one of 224 white European control chromosomes tested, which suggests that it is a rare non-pathogenic variant (Table 7.2).

A c.5G>A transition resulting in an arginine to glutamine substitution (p.R2Q) was identified in a patient with bilateral microphthalmia, retinal coloboma and congenital kidney defects as well as other anomalies (Figure 7.1A). Parental DNA was unavailable for testing as was information on their affected status. This variant was absent from 224 white European control chromosomes screened by direct sequencing. The arginine at position 2 in TMEM114 is conserved down to chick and is located in the predicted short intracellular N-terminus of TMEM114 (Figure 7.2). The potential functional effect of p.R2Q was investigated using tools which predict the effect of non-synonymous substitutions on protein function. Supportive of the mutations pathogenicity *in vivo* PolyPhen (Ramensky *et al.* 2002) predicted the substitution as “probably damaging” and SIFT (Kumar *et al.* 2009) predicted that the change would be “not tolerated”.

The third variant identified, c.440C>T changes the first nucleotide in exon 4 resulting in the substitution of the highly conserved non-polar alanine with the non-polar valine (p.A147V) in TMD3 (Figure 7.1C). This patient had bilateral anophthalmia, hypoplasia of both globes, optic nerve hypoplasia and multiple non-ocular anomalies (atrophic chiasm, interhemispheric cyst, corpus callosum agenesis, Dandy Walker malformation, seizures, dysplastic cup shaped ears and extended abdomen). Sequencing of the unaffected parents of this individual identified the same c.440C>T variant in the patient's unaffected mother confirming that the variant is not the causative mutation in this patient. This variant disrupts an *Nco I* site (CCATGG). The *Nco I* digested exon four fragment from the patient was detected as two bands of 372 bp (undigested) and 301 bp (digested), indicating heterozygosity at the *Nco I* site (Figure 7.3). However, all of the 160 white European control chromosomes screened were sensitive to *Nco I* digestion indicating they were wildtype for the *Nco I* site.

Two non-coding variants and one synonymous substitution were also identified in the panel. The c.13C>T variant does not alter the encoded amino acid suggesting it is a SNP. A C>T variant in the 5'UTR, 24 nucleotides upstream of the A of the initiation codon was present in patients in heterozygous and homozygous states but was also identified in controls heterozygous and homozygous states suggesting it was a SNP. The c.440-7C>G change in at the end of intron three has previously been reported as a SNP in the database dbSNP (rs62018861).

Table 7.2. Variants identified in the coding and non-coding regions of *TMEM114* in MAC patients and controls. H/h = heterozygous. h/h homozygous. * = tested by *Nco I* digestion. ^ = rs62018861

			Patients		Controls	
	Variants		Prevalence in H/h state	Prevalence in h/h state	Prevalence in H/h state	Prevalence in h/h state
	Nucleotide	Protein				
Exon1	-24C>T	-	20/77	6/77	36/112	1/112
Exon1	c.5G>A	p.R2Q	1/77	0/77	0/112	0/112
Exon1	c.13C>T	p.L5L	1/77	0/77	0/112	0/112
Exon1	c.134C>T	p.P45L	1/77	0/77	1/112	0/112
Intron3	c.440-7C>G^	-	16/77	3/77	not tested	not tested
Exon4	c.440C>T	p.A147V	1/77	0/77	0/80*	0/80*

A

```
H. sapiens          MRVHLGGLAGAAALTGALSFVLLAAAIGTDFWYIIDTERLERT--------GPGAQDLLGS 53
M. mulatta          MRVHLGGLAGAAALTGALSFVLLAAAIGTDFWYIIDTERLERS--------GPGAQDLLGS 53
M. musculus         MRVRLGALAGAAALSGALSFVLLAAAIGTDFWYIIDTERLERS-SQRMRDQGP-------- 52
R. norvegicus       MRVRLGALAGAAALSGALSFVLLAAAIGTDFWYIIDTERLEKS-SPRMRDQGP-------- 52
B. taurus           MRVHLGALAGAAALTGALSFVLLAAAIGTDFWYIIDTERLERG--------GPGARGPVGA 53
M. domestica        MRVRLGALAGAAALTGALSFVLLAAAIGTDFWYIIDTARLEQRRPGASGARDPGGSSTDS- 60
G. gallus           MRVSLSVLSLLVALVGASSFVFLVVAIATDFWYIIDASRLETATN---------------- 45
X. tropicalis       MKLKLSMLSVFVAVVGILSFISLVVAIGTDFWYIIDASRLEKITN---------------- 45
D. rerio            MKVTFTGLAGFVAVFGVLSFIGLVLAIGTDFWYIIDTSKRENSSS---------------- 45
                    *:: :  *:  .*: *  **: *. **.********: : *

H. sapiens          INRSQPEPLSSHSGLWRTCRVQSPCTPLMNPFRLENVTVSESSRQLLTMHGTFVILLPLS 113
M. mulatta          INRSQLEPLSSHSGLWRTCRVQSPCTPLMNPFWLENVTVSESSRQLLTMHGTFVILLPLS 113
M. musculus         ANRSQQEPLSSHSGLWRTCRVQSSCTPLMNPFWQENVTVSDSSRQLLTMHGTFVILLPLS 112
R. norvegicus       VNRSQSEPLSSHSGLWRTCRVQSTCTPLMNPFWQENVTVSDSSKQLLTMHGTFVILLPLS 112
B. taurus           NNRSQLEPLSSHSGLWRTCRVQSPCAPLMNPFWQENVTVSDSSRQLLTMHGTFVILLPLS 113
M. domestic         LNHSQPEPLSSHSGLWRTCRVQSKCTPLMNPFWQENVTVSDSDRQLLTMHGTFVILLPLS 120
G. gallus           ----GTAALSSHSGLWRTCRLRSKCYPLINPFWYENANITDSHRQLLYMHGTFVILMPLS 101
X. tropicalis       ----FSDPLSSHSGLWRMCKFKNKCLPLINPFRLGNLNFTDSQKQLLSMHGTLVILLPLS 101
D. rerio            ------VSLSSHSGLWRTCNFHNQCWPFMNPFGAG-RNLSDSQRQILNMQGTFIILLPLS 98
                           .********* *..:. * *::***     ..::* :*:* *:**::**:***

H. sapiens          LILMVFGGMTGFLSFLLQAYLLLLLTGILFLFGAMVTLAGISVYIAYSAAAFREALCLLE 173
M. mulatta          LILMVFGGMTGFLSFLLRAYLLLLLTGTLFLFGAMVTLAGISVYIAYSAAAFREALCLLE 173
M. musculus         LIVMVFGGMTGFLSFLLRAHLLLLLTGILFLFGAMVTLTGISIYIAYSAVAFREAVCLLE 172
R. norvegicus       LIVMVFGGMTGFLSFLLRAHLLLLLTGTLFLFGAMVTLTGISIYIAYSAVAFREAMCLLE 172
B. taurus           LILMVFGGMTGFLSFLLRASCLLLLTGTLFLFGALVTLAGISVYIAYSAAAFQEALCLLQ 173
M. domestica        LILMVFGGMTGFISILARAYLLLLMTGMLFLFGALVTLAGISVYIAYSAAAFQEAVCLLE 180
G. gallus           LILMIFGGMTGFISILARAYLLLLMTGLLFLFGALVTLTGIGVYIAYSAAAFEEAVCLLR 161
X. tropicalis       LILMIFGGMTGFVSILARAYLLLLLTGMLFLFGALVTLTGISIYIAYSAAAFKDAVCILG 161
D. rerio            VILLFVGGMLGFISLLARAYVLLLLTGVLLLFGAVLSLAGICVYMAYSAAAFREAVDISG 158
                    :*::..*** **:*:* :*  ***:** *:****:::*:** :*:****.**.:*: ::

H. sapiens          EKALLDQVDISFGWSLALGWISFIAELLTGAAFLAAARELSLRRRQDQAI- 223
M. mulatta          EKALLDQVDIHFGWSLALGWISFIAELLTGGAFLAAARELSLRRRQDQAI- 223
M. musculus         ERALLDQVDIRFGWSLALGWISFVSELLTGVVFLAAARALSLSQRQDQAI- 222
R. norvegicus       ERALLDQVDIRFGWSLALGWISFVSELLTGVVFLAAARVTNLSQRQEQAI- 222
B. taurus           EKTLLDQVDIRFGWSLALGWVSCVAELLTGATFLVAARVLSLRRRQDQAI- 223
M. domestica        EKDLLDQVDIRFGWSLALGWISFISELLTGAAFLLAARMVGLKRRQDQAI- 230
G. gallus           SKDVLVEIDIRFGWSLALVWISFVAEVITGAAFLLAARVVGLKQQHEQEL- 211
X. tropicalis       NK-ILEDIDIQFGWSLALAWISFITEILTGIAFLVAARVTGLKRRRREQVI 211
D. rerio            HK-TLQDIEIYFGWSLILASVSFVGELCTAVAFLLTSVKVSQQTNQEQDE- 207
                      :  * :::* ***** *  :* : *: *. .** ::   .   .: :
```

B

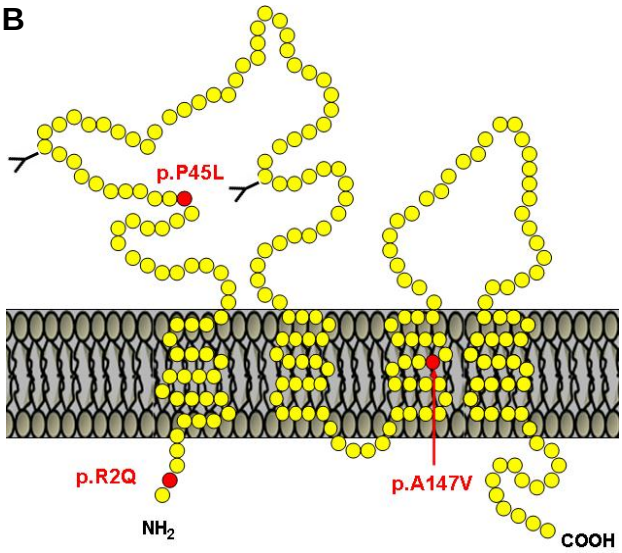


Figure 7.2. Conservation and location of the affected residues. A) Multiple alignment of TMEM114 orthologues. p.R2 is conserved to chick. p.P45 is conserved in mammals but is located in a non-conserved region. p.A147 is conserved in all species and is the first amino acid encoded by exon 4. Variants are highlighted in yellow. Alternate exons are coloured in black and blue font. Amino acids bridging two exons are in red font. B) Location of affected residues. p.R2 is located in the intracellular N-terminus. p.P45 is located in the 1st extracellular loop. p.A147 is located in the 3rd transmembrane domain. Topology predicted by TMHMM.

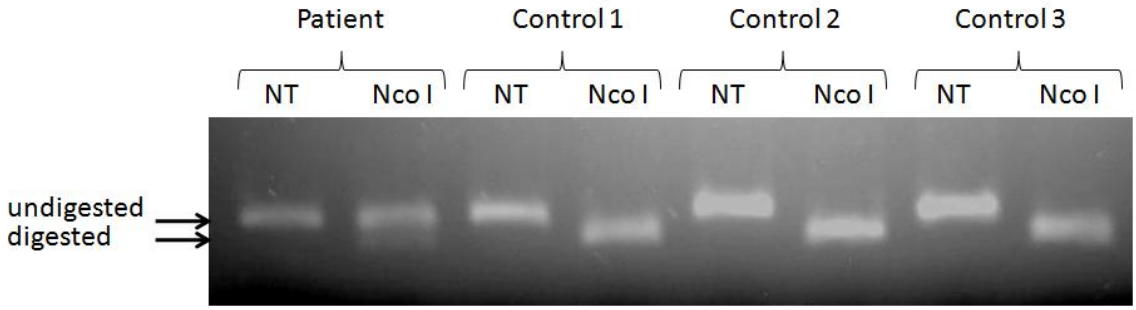


Figure 7.3 Screening for the c.440C>T variant in patients and controls by *Nco* I digestion. The c.440C>T variant disrupts and *Nco* I restriction site in exon 4 of TMEM114. In control DNA *Nco* I treatment digested the DNA forming a 301 bp DNA fragment, compared to the 372 bp fragment in the non-treated sample (NT). The amplicon from the patient identified to be heterozygous for the c.440C>T variant produced two bands of 372 bp and 301 bp when digested with *Nco* I indicating that the *Nco* I site was disrupted on one allele.

7.3 Functional analysis

The panel screened was from an external source and there was a delay in acquiring information regarding the patients and acquiring parental DNA samples. Therefore, before the appropriate controls had been screened, *in vitro* analysis was undertaken to determine if the mutations identified had functional effects on the TMEM114. To assess the potential pathogenic consequences of the human TMEM114 variants identified, mutations were introduced at the equivalent site in mTmem114-V5 using SDM (Figure 7.4). The murine equivalent of p.A147V, p.A146V, had previously been cloned (Urquhart, unpublished) due to the presence of this variant in patients with autosomal dominant congenital cataract (ADCC) (Jamieson *et al*. 2007).

7.3.1 Expression levels

Expression levels of the variants p.R2Q, p.P52L and p.A146V were compared by transfecting into MDCK II cells and detecting by western blot. Forty eight hours post-transfection the p.R2Q variant and the p.P52L (murine equivalent of p.P45L) had similar expression levels and pattern to the wildtype (Figure 7.5A). Unexpectedly, the p.A146V isoform (murine equivalent of p.A147V) was not detected in the 40-45 kDa isoform, but was detected predominantly in the intermediate isoform at ~32 kDa (Figure 7.5A). This variant was sensitive to digestion by Endo H which cleaves mannose-rich oligosaccharides, which are typical of glycoproteins that have not been trafficked beyond the ER (Figure 7.5B).

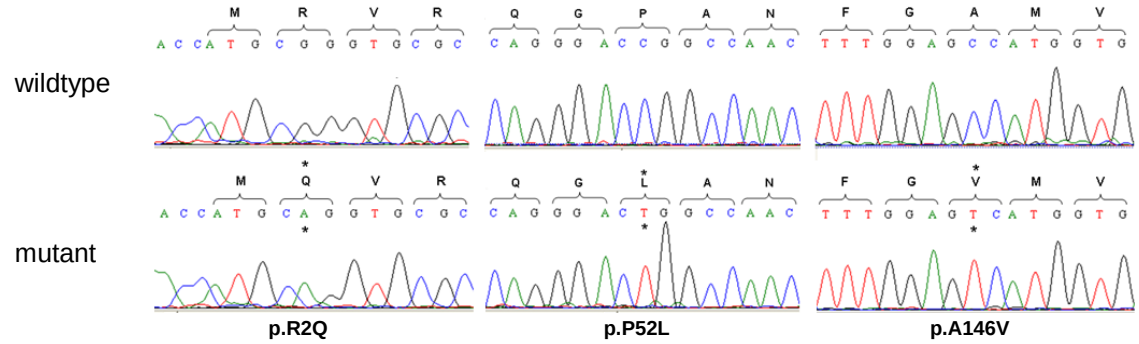


Figure 7.4. Sequence chromatograms of the mTmem114-V5 mutants created by site-directed mutagenesis. Mutations (asterisks) were created at the equivalent sites of murine Tmem114

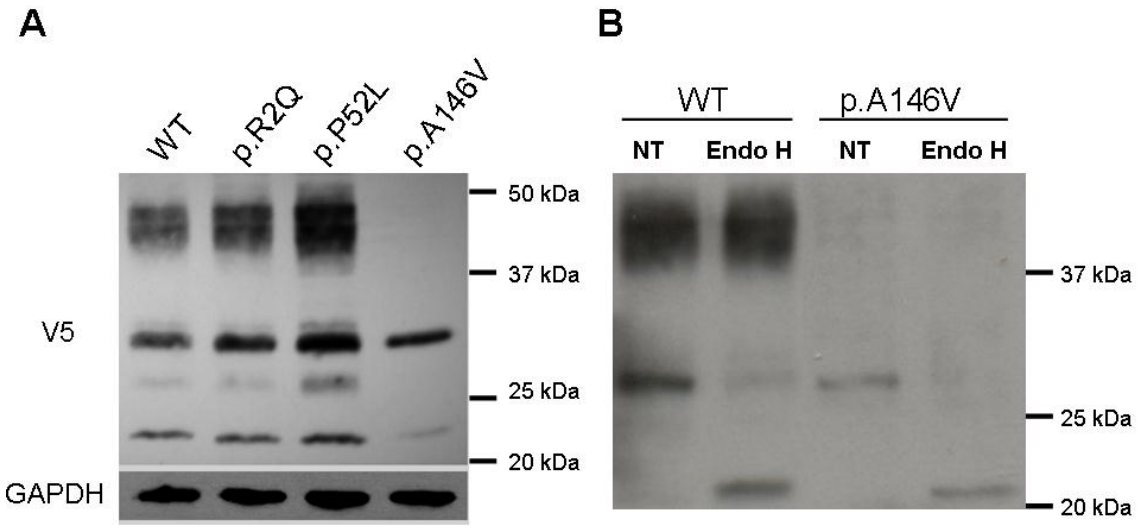


Figure 7.5. Detection of wildtype and missense isoforms of C-terminal V5 tagged Tmem114 by western blot. A) Expression of wildtype and missense isoforms in MDCK II cells. Tmem114 is detected in three main isoforms; unglycosylated (23 kDa), with covalently linked mannose-rich oligosaccharides (32 kDa), and with covalently linked complex oligosaccharides (40-45 kDa). p.R2Q and p.P52L are expressed at similar levels to wildtype Tmem114 (when normalised to the housekeeping protein GAPDH). p.A146V is expressed predominantly in the mannose-rich 32 kDa isoform. B) Endo H digestion of wildtype and p.A146V. Non-treated (NT) p.A146V is detected at 32 kDa. Endo H treatment of p.A146V results in a shift to the unglycosylated form (23 kDa) of Tmem114. Endo H treatment of wildtype Tmem114 results in a shift of the mannose-rich oligosaccharide isoform, but not the complex oligosaccharide isoform.

7.3.2 Localisation of variants in polarised MDCK II cells

To determine if the missense mutations had effects on the localisation of Tmem114, the localisation of the V5-tagged variants was sought in polarised MDCK II cells. The p.R2Q and p.P45L variants localised to the lateral and apical cell membrane, comparable to that of

wildtype Tmem114 (Figure 7.6). When V5-tagged p.A146V was expressed in polarised MDCK II cells it displayed a predominantly perinuclear localisation. Co-staining with the membrane protein E-cadherin confirmed p.A146V does not localise at the cell membrane (Figure 7.7A). The staining pattern of p.A146V was indicative of ER localisation. This could not be directly demonstrated because, as mentioned in section 4.5.3, it was not possible to optimise an ER antibody for MDCK II cells. Using the Golgi apparatus marker GM-130 revealed that p.A146V-V5 is not localised to the Golgi apparatus (Figure 7.7B). This lack of colocalisation with the Golgi apparatus and the plasma membrane supports the conclusion that p.A146V-V5 does not undergo further processing of its N-linked oligosaccharides, as shown by its sensitivity to Endo H digestion (Section 7.2.1).

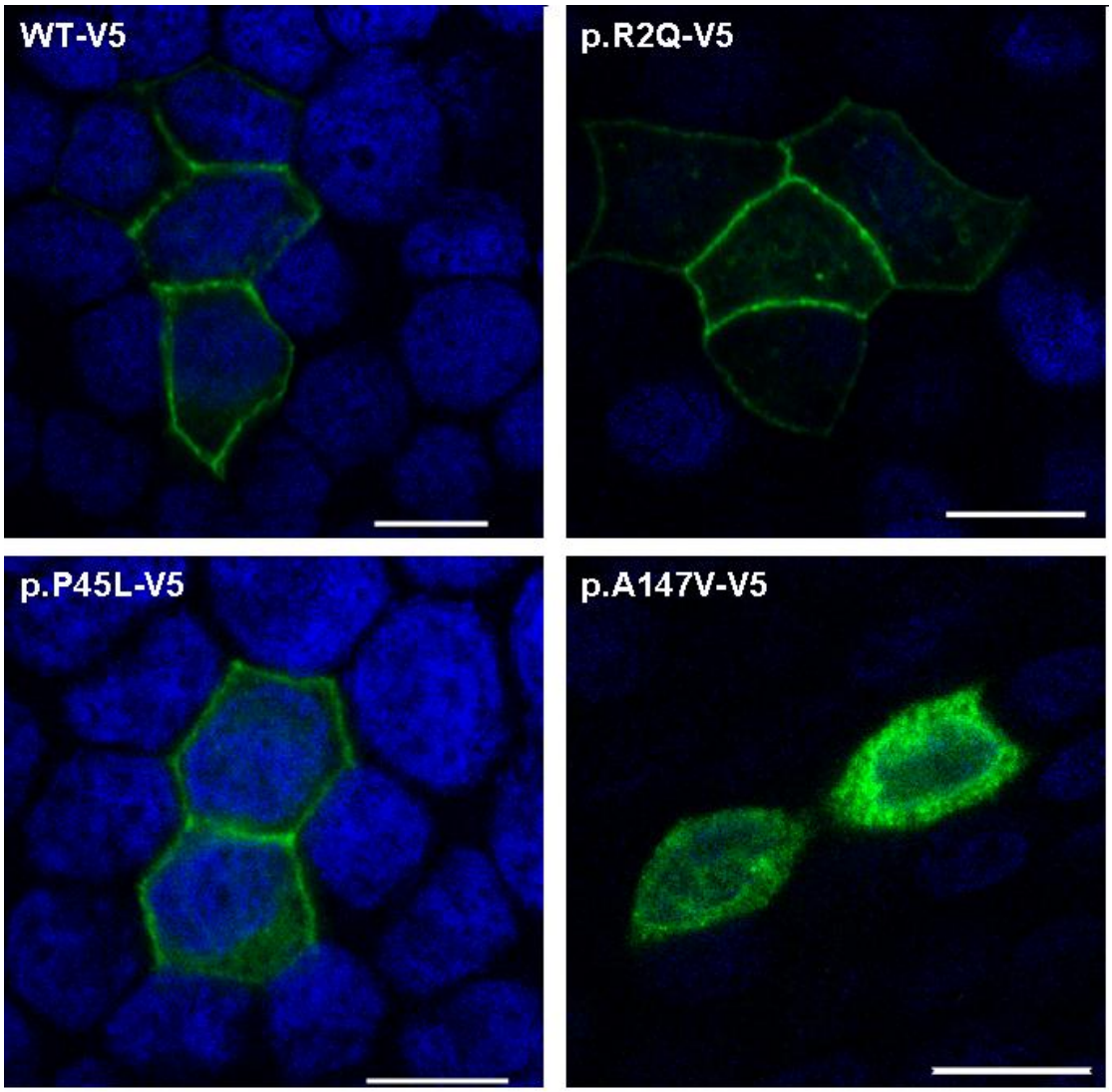


Figure 7.6. Localisation of wildtype and missense isoforms of V5-tagged mTmem114 in polarised MDCK II cells. Transfected cells were polarized on transwell filters and labeled with anti-V5 antibody (green) 48 hours post-transfection. Wildtype mTmem114-V5 localises to the plasma membrane. The variants p.R2Q and p.P52L also localise to the plasma membrane. The missense isoform p.A146V does not localise to the plasma membrane and is retained in the cytoplasm. Nuclei are stained with DAPI (blue). Bar = 10 µm. Cells were fixed with 1:1 methanol:acetone and viewed by confocal microscopy.

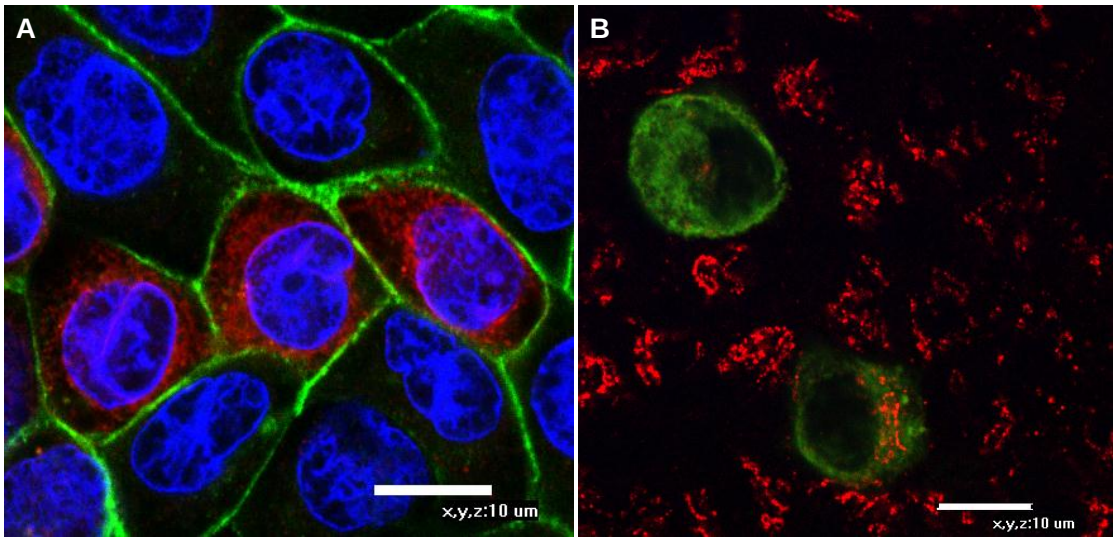


Figure 7.7. Localisation of mTmem114 p.A146V-V5 in MDCK II cells. A) mTmem114 p.A146V-V5 (red) fails to localise to the cell membrane labelled with E-cadherin (green). Nuclei are stained with DAPI (blue). B) mTmem114 p.A147V-V5 (green) does not colocalise with the Golgi apparatus marker GM130 (red). Bar = 10 µm. Cells were fixed with 1:1 methanol:acetone and viewed by confocal microscopy.

7.3.3 Effects of p.A146V on wildtype Tmem114

Missense mutations which result in retention of the mutant protein in intraceullular organelles have the potential to act in a dominant-negative manner by inhibiting the trafficking of wildtype protein (Saliba *et al.* 2002). To determine if expression of p.A146V-V5 influences the localisation of wildtype mTmem114, untagged wildtype mTmem114 was co-transfected with V5-tagged p.A146V into polarised MDCK II cells and visualised by immunofluorescence. As previously demonstrated in figure 4.4, the V5-epitope prevents detection by the custom polyclonal antibody. Cells co-transfected with untagged wildtype mTmem114 and p.A146V-V5 and labelled with anti-Tmem114 antisera (i.e. detecting the wildtype isoform only) displayed plasma membrane staining (Figure 7.8Ai). Co-transfected cells labelled with anti-V5 antibody (detecting the p.A146V isoform only) displayed cytoplasmic staining (Figure 7.8Aii). Cells transfected with p.A146-V5 alone and labelled with anti-Tmem114 antisera showed very low levels of background staining (Figure 7.8Aiii). This confirms that the anti-Tmem114 antisera was specifically detecting the untagged protein.

In co-transfected cells labelled with both anti-Tmem114 anti-sera and anti-V5 antisera, the wildtype untagged protein was detected at the cell membrane whilst the mutant protein was retained in the cytoplasm (Figure 7.8B). Therefore, co-transfection of p.A146V Tmem114 had no effect on the trafficking of the wildtype protein.

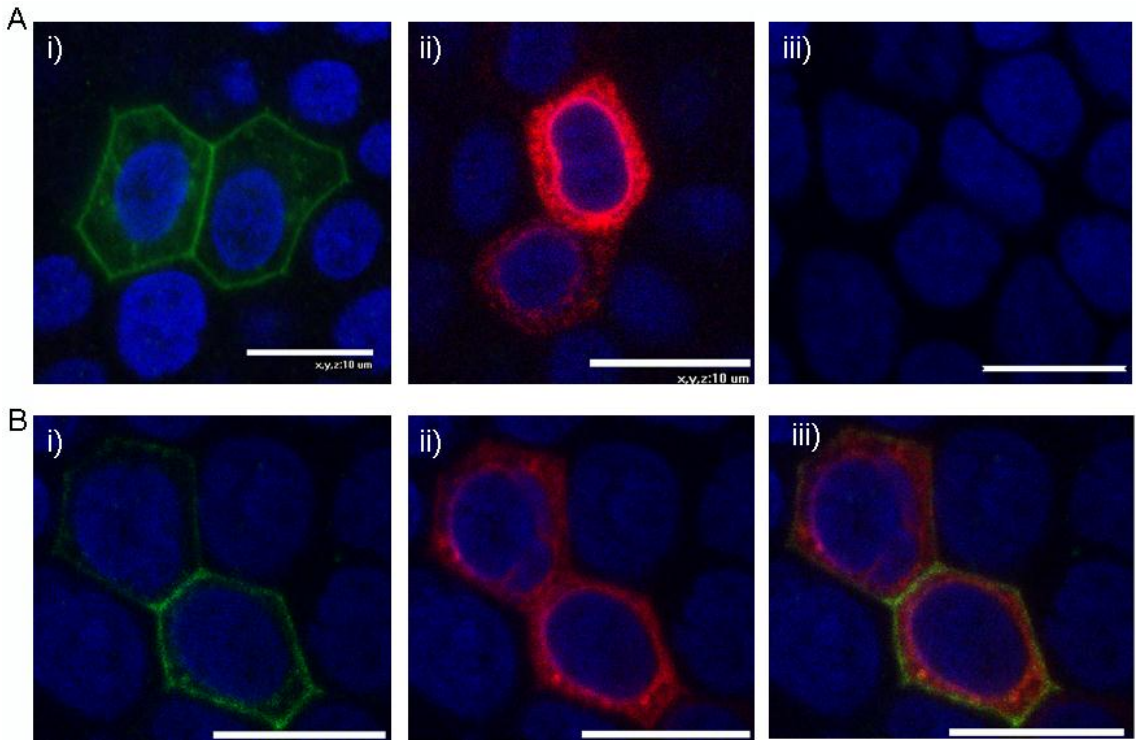


Figure 7.8. Co-transfection of wildtype mTmem114 and p.A146V-V5 in MDCK II cells. A. i) In co-transfected cells, anti-Tmem114 antisera (green) detects only the untagged wildtype isoform and shows only plasma membrane staining. ii) In co-transfected cells, anti-V5 antibody (red) detects only the p.A146V isofrom and shows cytoplasmic but not plasma membrane staining. iii) In cells only transfected with the p.A146V-V5 isofrom, anti-Tmem114 antisera (green) only detects very low levels of background staining due to epitope masking. B) Dual staining with anti-mTmem114 and V5 antibodies in co-transfected cells. Dual-labelled cells for anti-Tmem114 (green) (i), anti-V5 (red) (ii), and the merged image (iii). In co-transfected cells, wildtype Tmem114 localises to the plasma membrane, but p.A146V-V5 is retained in the cytoplasm. Nuclei are stained with DAPI (blue). Bar = 10 µm. Cells were fixed with 1:1 methanol:acetone and viewed by confocal microscopy.

7.4 Possible effect of c.440C>T on splicing

The nucleotide c.440C is the first nucleotide of exon 4 of *TMEM114* and therefore at the intron-exon boundary. Mutations at intron/exon boundaries can affect the splicing of transcripts and are a common source of mutation in human disease (Ars *et al.* 2000; Schmid *et al.* 2010). To investigate if the sequence variant c.440C>T has an effect on the splicing of *TMEM114*, bioinformatic splicing prediction programs were used to predict possible splicing differences between the wildtype and c.440C>T isoforms. Of the 4 programs used, two (PESX and FAS-ESS) did not detect any difference between the 2 sequences. ESEfinder predicted the loss of a SRp40 site and the gain of a SRp55 site in the variant sequence and the program Rescue-ESE predicted the loss of an exon splice-enhancing site in the variant sequence (Figure 7.9).

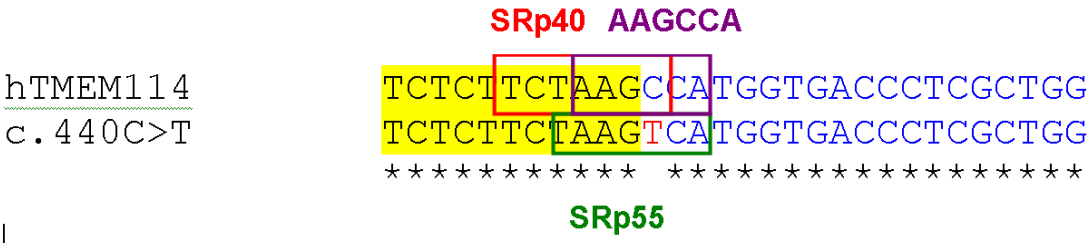


Figure 7.9. Predicted splice factor binding sites in wildtype and c.440C>T *TMEM114*. Wiltype *TMEM114* (top) contains an SRp40 binding site and an AAGCCA binding motif that are lost by the alteration of C>T at nucleotide c.440 (bottom, in red). The mutant sequence contains a predicted SRp55 site not present in the wildtype. Intron 3 (highlighted in yellow) and exon 4 (blue font). Variant nucleotide is highlighted in red font.

7.5 Discussion

7.5.1 Prevalence and etiology of microphthalmia

As discussed in section 6.4.3, a number of genes associated with congenital cataract are also associated with microphthalmia. The *To3* mouse is homozygous for missense mutations in *Lim2*, a gene homologous to *Tmem114*, displays microphthalmia (Steele *et al*. 1997). *TMEM114* is located at a locus associated with isolated microphthalmia with cataract (MCOPCT1). This data, taken in conjunction with what has been identified in this study (that *TMEM114* is expressed in the developing human eye (Figure 4.1) and knockdown of xtTmem114 results in microphthalmia (Chapter 6) makes *TMEM114* a good candidate for association with microphthalmia.

Microphthalmia is a congenital ocular developmental defect responsible for 3.2-11.2 % of blindness in children (Fujiki *et al.* 1982; Fraunfelder *et al.* 1985; Traboulsi 1999), with an estimated prevalence of 5-14 per 100,000 live births (Morrison *et al.* 2002; Kallen and Tornqvist 2005; Shah *et al.* 2010). Microphthalmia is often associated with other ocular defects such as cataract (Billingsley *et al.* 2006) and coloboma (Aijaz *et al.* 2004), as well as with non-ocular defects such as mental retardation and cardiac anomalies in syndromes such as CHARGE syndrome (OMIM 214800) and Lenz microphthalmia syndrome (OMIM 309800).

Although a number of causative genes have been identified for microphthalmia (e.g. *SOX2*, *PAX6, VSX2*) the genetic etiology of the majority of cases remains unknown (Verma and Fitzpatrick 2007; Chao *et al.* 2010). The majority of the genes currently associated with microphthalmia and anophthalmia encode transcription factors, which typically lead to more systemic phenotypes than microphthalmia and anophthalmia alone. Isolated cases of microphthalmia and anophthalmia are in the minority (Morrison *et al*. 2002) and relatively few genes have been identified. Homozygous mutations in the homeobox gene *VSX2/CHX10* (Burkitt Wright *et al.* 2010; Iseri *et al.* 2010) and compound heterozygous mutations in another homeobox gene *RAX* (Voronina *et al.* 2004; Lequeux *et al.* 2008) have been associated with recessive forms of isolated microphthalmia. Biallelic mutations in *FOXE3*, a transcription

factor with a role in the proliferation and differentiation of lens cells (Medina-Martinez *et al*. 2005), are associated with recessive microphthalmia in multiple ethnic groups (Reis *et al*. 2010b). The homeobox gene *SIX6* is located in a locus associated with isolated microphthalmia (14q32; OMIM 251600). Although a sequence variant in *SIX6* has been identified in a case of microphthalmia, pathogenicity of the variant has not been demonstrated (Gallardo *et al.* 2004).

7.5.2 Variants identified in *TMEM114*

The panel of 77 patients screened here contained 48 cases of bilateral isolated or syndromic microphthalmia and 13 cases of bilateral isolated or syndromic anophthalmia. A c.5G>A transition resulting in an arginine to glutamine substitution (p.R2Q) was identified in a patient with bilateral microphthalmia, retinal coloboma, kidney defects and other non-ocular anomalies such as bilateral cryptorchidism, adducted left thumb, and pseudoarthrosis of the right clavicle. This variant was found to be absent from 224 control chromosomes tested. Unfortunately, parental DNA and clinical information was unavailable making the significance of their affected status uncertain. The p.R2Q variant identified is predicted to be damaging according to the prediction tools PolyPhen2 and SIFT. As discussed in section 4.8.5 the N-terminal region (the “N domain”) of signal sequences are basic. Substitution of the basic arginine residue at p.R2 with a polar glutamine residue may have altered the N domain, possibly reducing the efficiency of insertion into the ER membrane. However, in polarized MDCK II cells the expression and localisation of p.R2Q variant was similar to that of wildtype Tmem114 (Figure 7.6).

A heterozygous arginine to glutamine missense mutation in the intracellular N-terminus of bestrophin-1, also close to the first predicted transmembrane domain (TMD), is associated with the autosomal dominant retinal disorder Best disease (Marquardt *et al.* 1998), supporting the likely pathogenicity of the p.R2Q variant in TMEM114. The function of the intracellular N-terminus of TMEM114 is unknown. Although short, the N-terminus may contain protein interaction domains which are disrupted by the arginine to glutamine substitution.

In addition to the ocular defects, the patient is also reported to display kidney defects, with an absent right kidney and a cystic dysplastic left kidney. Developmental kidney defects were also observed in individuals with cataracts carrying the balanced translocation upstream of *TMEM114* (Jamieson *et al*. 2007). These kidney phenotypes and the weak expression of *Tmem114* in the developing kidney (Jamieson *et al*. 2007) suggest a putative role for TMEM114 in kidney development. Biallelic mutations in *STRA6,* involved in vitamin A metabolism, are associated with microphthalmia with a duplicated kidney collecting system, although parental DNA was available to confirm the inheritance pattern (White *et al.* 2008). Expression of *Tmem114* was also detected in the mouse testis (Jamieson *et al*. 2007) and the p.R2Q patient also presented with bilateral cryptorchidism (failure of the testes to descend into the scrotal sac). However, bilateral cryptorchidism is the most frequent congential

defect in males, with an estimated prevalence of 2-4 % of live male births (Barthold and Gonzalez 2003).

The absence of the p.R2Q variant in 224 control chromosomes is suggestive of pathogenicity, but the limited scope of the functional assays available means that the functional effect of the mutation remains to be elucidated. To strengthen confidence in this finding, it would be helpful to screen the proband's DNA for mutations in *STRA6* in order to eliminate this as a potential cause of their condition.

7.5.3 Identification of the missense variant p.A147V

The missense variant c.440C>T (p.A147V) was identified in a patient with bilateral anophthalmia, hypoplasia of both globes, optic nerve hypoplasia and multiple non-ocular defects. The severe systemic phenotype suggested that the missense mutation was unlikely to be the causative mutation and this was confirmed by finding the same variant in the unaffected mother. Interestingly, the p.A147V variant was identified three times in a heterozygous state in a panel of 130 ADCC patients (Jamieson *et al*. 2007). However, in one of the families it was absent from an affected family member, suggesting it is a polymorphism (Jamieson *et al*. 2007). However, although previously reported as a polymorphism, this variant was found absent from 160 control chromosomes tested.

The mislocalisation of the murine equivalent of p.A147V, p.A146V, was not expected as the variant had been reported as a SNP (Jamieson *et al*. 2007). When expressed in MDCK II cells p.A146V displayed an intracellular localisation indicative of ER localisation. This ER localisation was confirmed by the fact that p.A146V was sensitive to Endo H digestion. Integral membrane glycoproteins are translated in the ER and exit the ER when they have been folded into the correct conformation by ER resident chaperones (Dejgaard *et al.* 2004). Mutations in TMDs of membrane proteins may disrupt their native conformation, leading to incomplete folding or misfolding (Hwa *et al*. 1997; Choi *et al*. 2005; Gallagher *et al*. 2007). Proteins which remain misfolded after repeated cycles of chaperone binding are targeted for ER-associated degradation (ERAD). ERAD is a quality control system which targets misfolded or unassembled proteins within the ER lumen for degradation via the cytoplasmic ubiquitin-proteasome system (Gelman *et al*. 2002; Meusser *et al*. 2005). However, severely misfolded proteins may aggregate in the cell triggering the unfolded protein response (UPR) (Roboti *et al.* 2009). The UPR is a stress response that induces expression of chaperones to promote folding (Friedlander *et al.* 2000; Okada *et al.* 2002) and halts protein translation to prevent protein build up in the ER (Okada *et al*. 2002; Ron 2002).

The 40-45 kDa Endo H insensitive isoform of Tmem114, which localises to the cell membrane, was not detected in the p.A146V lysate. This suggested that p.A146V was either being trafficked through the Golgi apparatus to the plasma membrane but was being rapidly degraded, or it was not being trafficked and was being retained in the ER. Some misfolded

proteins escape the ER quality control system and are trafficked to the plasma membrane. However, the lysosomal degradation system detects and degrades misfolded proteins that have been transported to the Golgi apparatus (Arvan *et al.* 2002) or plasma membrane (Luzio *et al.* 2007). Lysosomes are membrane-bound organelles that contain multiple hydrolytic enzymes to degrade misfolded proteins.

Different mutations in the same protein which affect protein folding may be processed in very different ways and yet still result in the same phenotype as is the case for mutations in claudin-16 which cause familial hypomagnesemia with hypercalciuria and nephrocalcinosis (FHHNC) (Kausalya *et al.* 2006). If p.A146V was being retained in the ER it may be degraded via the proteasome or the misfolded protein may have formed multimers and aggregated. Such protein aggregates fail to resolve on SDS-PAGE gels and are retained in the stacking gel (Goldberg *et al.* 2001; Mishra *et al.* 2008). No such aggregates were identified in the p.A146V-V5 lysates.

The majority of cellular proteins are routinely being turned over at rates which are protein dependent. Protein turnover involves degradation of the cellular protein and its replacement by newly synthesised protein. Misfolded proteins degraded by the cellular machinery are degraded much more rapidly than the correctly folded proteins. The alanine to valine substitution at p.A147 may disrupt the conformation of TMEM114 resulting in its misfolding leading it to be more rapidly degraded than the wildtype. A heterozygous alanine to valine mutation in a TMD of rhodopsin which results in misfolding (Hwa *et al*. 1997) is associated with the retinal disease retinitis pigmentosa (RP) (Fuchs *et al.* 1994). Although alanine and valine are both hydrophobic, valine has a larger sidechain which results in the disruption of residues in another TMD (Stojanovic *et al.* 2003).

To determine if p.A147V is being degraded at a faster rate than the wildtype, pulse-chase experiments could be performed. Pulse-chase experiments involve the cellular proteins being radiolabeled for a set period of time (the pulse) followed by a period of time when protein synthesis occurs without radiolabeled amino acids (the chase). The degradation of the radiolabeled protein can then be monitored over time. Pulse-chase analysis, would determine the rate of turnover of wildtype Tmem114 and then the rate of degradation of p.A146V could then be determined and compared to the wildtype. If a more rapid rate of degradation was confirmed, the pathway by which p.A146V is being degraded could be determined using proteasome and lysosome inhibitors and detecting changes in expression levels (Roboti *et al*. 2009).

The identification of the molecular effect of the alanine to valine substitution on the folding of TMEM114, i.e., whether it is rapidly degraded by the proteasome or lysosome or if it is not degraded efficiently leading to a cytotoxic build up of misfolded protein, is important in terms of disease.

7.5.4 The p.A147V variant and dominant disease

The identification of the p.A147V variant in three patients with ADCC suggests that, 1) it is a rare polymorphism and has no association with congenital cataract or, 2) it is has a dominant effect and the family member with ADCC in which this variant was not detected was either a phenocopy or was heterozygous for c.440C>T but the variant was not detected due to allelic dropout. Although both latter options are unlikely, such occurrences are not without precedence (Bidinost *et al*. 2006; Tester *et al.* 2006).

ER-retained misfolded proteins exert a dominant negative effect via two main mechanisms. Firstly, by preventing trafficking of the wildtype protein to its correct destination, or secondly, by creating a build-up of misfolded protein in the ER that triggers the unfolded protein response (UPR). To investigate the first possibility wildtype and p.A146V Tmem114 were co-transfected in polarised MDCK II cells. Co-transfection of the p.A146V mutant isoform had no effect on the localisation of wildtype Tmem114. The possibly that p.A146V forms stable aggregates in the ER and triggers UPR has yet to be determined. However, as previously discussed the lack of aggregates resolving on SDS-PAGE gels suggests this phenomenon is not occurring. Future pulse chase experiments investigating the degradation rates of wildtype and p.A146V Tmem114 would more comprehensively address this possibility.

Heterozygous mutations may also have a dominant effect as a result of haploinsufficiency. The appearance of the p.A147V variant in the heterozygous state in the unaffected mother of a patient with microphthalmia suggests that individuals who are haploinsufficient for wildtype TMEM114 are asymptomatic. This also implies that the p.A147V variant in a heterozygous state is not associated with pre-senile cataract or other ocular disease.

7.5.5 Association of p.A147V with age-related cataract?

Age-related cataract occurs due to the misfolding and aggregation of protein which accumulates over the lifetime of an individual (Andley 2009). Although heritability of age-related cataract is well established (Hammond *et al.* 2001) there are few genetic variations associated with age-related cataract, although the number of associations are increasing (Jun *et al.* 2009; Zhou *et al.* 2010; Zuercher *et al.* 2010). Lifelong expression of misfolded p.A147V in combination with other misfolded proteins that naturally accumulate over time may result in age-related cataract. Therefore, mutations in *TMEM114* may be a contributing factor in age-related cataract.

7.5.6 Association of p.A147V with recessive disease?

Although the p.A147V variant does not appear to act in a dominant manner, its mislocalisation and possible degradation would suggest it is non-functional. Co-transfection with wildtype Tmem114 did not rescue the localisation of p.A146V, suggesting that this variant acts in a recessive manner. In the previous chapter it was demonstrated that knockdown of

xtTmem114 *in vivo* results in microphthalmia, suggesting that a lack of functional Tmem114 causes ocular disease. Therefore it is possible that recessive *TMEM114* mutations are associated with ocular phenotypes, as is the case for missense mutations in the TMD of LIM2 (Pras *et al*. 2002; Ponnam *et al*. 2008). Unfortunately, no panels of recessive cataract or microphthalmia were available to test this hypothesis.

7.5.7 Splicing effect of the c.440 C>T variant

The alanine to valine substitution at p.A147 is caused by the c.440C>T variant at the first nucleotide of exon three. The functional investigation into the c.440C>T variant in this chapter was based on the amino acid substitution. However, the altered nucleotide's location at an intron/exon boundary suggested that the DNA variation it may affect splicing. Splice site prediction tools identified that the mutation results in the loss of some splicing factor binding sites that may cause a failure of splicing or aberrant splicing at a cryptic splice site. This could have been tested using an *ex vivo* splice assay but due to time restrictions this was not possible.

Missense mutations in the coding sequence of the RPE gene *BEST1* which cause autosomal dominant vitreoretinalchoroidopathy (ADVIRC) have been demonstrated *in vitro* to cause misplicing resulting in the produciton of deletion isoforms (Yardley *et al.* 2004). The missense mutations are predicted to result in the production of two isoforms from a single allele *in vivo*: the missense isoform and a deletion isoform caused by exon skipping (Yardley *et al*. 2004). The phenotype of ADVIRC patients with missense splice site mutations is distinct to those with the heterozygous missense mutations in *BEST1* that cause Best disease. Aberrant splicing of *TMEM114* due to the c.440C>T variant may result in the production of a mutant protein with more severe functional consequenses than caused by the straightforward p.A147V missense isoform.

Alternatively the c.440C>T variant may prevent splicing of intron 3 resulting in a premature stop codon, which may lead to nonsense mediated decay. Thus, if the c.440C>T variant has an effect on splicing, the pathogenic mechanism may result from loss- or gain- of function. If the c.440C>T change does alter splicing, the mis-spliced isoform or the missense mutation isoform may not be produced exclusively and there may be a mixed population of protein products. It is likely that the proportions would be determined by the splicing machinery of the individual and so the c.440C>T variant may show altered penetrance depending on the genetic background of the individual.

7.5.8 Limitations of the *TMEM114* sequencing strategy

As with the majority of sequencing strategies used to detect mutations there were a number of limitations. Only the coding exonic sequences and the surrounding intron/exon boundaries and 70 bp of 5'UTR were sequenced. The primers used previously to detect mutations in *TMEM114* (Jamieson *et al*. 2007) had binding sites close to intron/exon boundaries.

Sequences close to the primer binding site (within 50 bases) are often not read as the short amplicons produced may be removed during the purification process. Therefore, if sequences were read in only one direction, the use of these primers (positioned close to the intron/exon boundaries) may have resulted in splice site altering variants in the intronic regions being missed. To overcome this potential problem, new primers were designed to amplify from deeper within the introns. This had the advantage of producing a clearer sequence trace for the exonic sequences and their proximal intronic sequences. The possible disadvantage of this strategy is that the intronic regions chosen may harbour SNPs which might result in allelic dropout. This would result in the sequencing of a single allele for which the patient would appear homozygous for the wildtype sequence.

The breakpoint on chromosome 16 in the balanced translocation t(16;22)(p13.3;q11.2) associated with juvenile onset cataract is located 673 bp upstream of the predicted transcription start site, and 151 bp upstream of the translation start site of *TMEM114* (Jamieson *et al*. 2007). Extending the upstream region of *TMEM114* that was sequenced may have identified sequence variants in the 5'UTR or in the putative promoter that alter regulatory elements. However, although regulatory elements may be close to the coding sequence of genes, e.g. in hereditary hyperferritinemia cataract syndrome (HHCS), the causative mutations are located ~160 bp upstream in the 5'UTR (Rochow *et al.* 2009), or much further away, e.g. regulatory elements in the transcription factor *PAX6* have been identified up to 150 kb away (Kleinjan *et al.* 2001).

Deletions in the genes *SOX2* (Reis *et al*. 2010a) and *BCOR* (Ng *et al.* 2004), and a large genomic deletion encompassing *TMX3* (Chao *et al*. 2010) are associated with syndromic and non-syndromic microphthalmia, respectively. It is possible that single allele deletions of coding sequence regions that presented as homozygous traces were missed. Also, as the focus of this sequencing strategy was on coding sequence variants, deletions of regulatory elements in the promoter, UTRs or introns would have been missed. To detect copy number variants (loss or gain of DNA) techniques such as multiplex ligation-dependent probe amplification (MLPA) could have been utilised. Due to time limitations, the strategy used focused only on PCR based detection methods.

Once variants in patients are identified it is important to determine if the same variants are present in unaffected individuals of the same ethnicity. Whilst 224 control chromsomes were screened (typically giving over 80 % power to detect a variant with a frequency of 0.01 (Collins and Schwartz 2002)) the controls were not ethnically matched. The c.5g>a (p.R2Q) variant was identified in a patient of Vietnamese origin (Figure 7.1). As Vietnamese or southeast Asian controls were not available, controls of white European descent were screened. Screening of southeast Asian controls may have identified the c.5g>a (p.R2Q) variant, suggesting that it is a non-pathogenic variant found in this population. Thus, future

work should involve the acquisition and screeening of southeast Asian controls and the access of the data from the 1000 genomes project (www.1000genomes.org).

7.6 Conclusion

This study and a previous study (Jamieson *et al*. 2007) have failed to identify a direct link between mutations in *TMEM114* and over 200 cases of human ocular disease. The *in vivo* in *X. tropicalis* data supports the hypothesis that the human orthologue of *TMEM114* is a strong candidate for recessive microphthalmia. The inheritance pattern of the microphthalmia cases in the panel of patients screened for mutations in *TMEM114* cases was unknown, but recessive cases are likely to represent a small fraction of the total number of the 48 cases (Reis *et al*. 2010b). To date, four loci (14q32, *VSX2*, *RAX*, *FOXE3*) have been associated with isolated autosomal recessive microphthalmia/anophthalmia (Bar-Yosef *et al*. 2004; Gallardo *et al*. 2004; Lequeux *et al*. 2008; Reis *et al*. 2010a), although more have been associated with syndromic forms. These four loci are estimated to account for less than 20 % of all cases (Reis *et al*. 2010b) suggesting there are more genes yet to be identified. Subsequent to its identification as a cause of autosomal recessive microphthalmia, screening of a panel of 198 microphthalmia and anophthalmia patients failed to identify mutations in VSX2 (Morrison *et al*. 2002). This suggests that even if mutations are not identified in one panel of patients, screening further panels may identify pathogenic mutations. Further support for this comes from other examples of ocular disease in which candidate genes were chosen due to the phenotype of mouse models. The gene *PCDH21* was chosen for screening of patients with retinal dystrophies as Pcdh21 knockout mice display retinal degeneration (Rattner *et al.* 2001). Screening of a panel of 279 patients with autosomal recessive retinal disease failed to identify mutations in *PCDH21* (Bolz *et al.* 2005), but five years later mutations were identified in a homozygous or biallelic state in families with autosomal recessive retinal disease (Henderson *et al.* 2010; Ostergaard *et al.* 2010). Thus, *TMEM114* remains a good candidate gene for microphthalmia.

Chapter 8: Discussion

8.1 Introduction

The human genome has an estimated 21,257 protein-coding genes, the majority of which encode proteins of unknown function (http://www.ensembl.org/Homo_sapiens/Info/StatsTable) (October 2010). Characterising the function of these proteins is of vital importance to our understanding of biological processes at the molecular level. With the advancement of high resolution SNP arrays and next generation sequencing, genes are readily being associated with human diseases (Daly *et al.* 2010; Johnston *et al.* 2010; Ng *et al.* 2010; Nishimura *et al.* 2010). In this functional genomics era the major challenge is to understand the biological role of proteins and the functional significance of disease associated variants. This is paramount to the identification of disease pathways and it is only when the pathogenic mechanisms are understood can one begin to identify possible therapeutics.

There are a number of ways to identify the function of a protein, and typically functional characterisation depends upon information gathered from multiple sources. The association with a disease may give clues as to the function, e.g. hereditary hyperferritinemia cataract syndrome patients have high serum ferritin levels, suggesting the gene associated with the disease is involved in ferritin production or regulation, as was the case (Beaumont *et al.* 1995). Often, however, there is little specific information that can be inferred from the disease phenotype, e.g. an association with cataract may reveal a protein's role in the lens, but provides scant information regarding the molecular function of the protein. The existence of characterised homologues or orthologues can also give clues to the function of the protein of interest and suggest experiments which can be utilised to test the function of the new protein. The creation, or existence of, an animal model for an orthologous protein can greatly aid the understanding of a protein's function and the identification of interacting protein partners can illuminate molecular pathways.

Many large publicly available databases containing information regarding DNA and protein expression and function are accessible online. Data from microarray-based expression studies can be useful to identify if the gene or interest is up- or down-regulated at different developmental time points or under different conditions such as after pharmaceutical treatment, or in diseased tissue (e.g. cancers). To date, major changes in the regulation of TMEM114 or TMLP1 have not been reported in microarray studies. Genome-wide association studies (GWAS), which compare the frequency of SNPs in disease and control populations, have been a useful tool to identify the genetic contributors to multigenic diseases (Rioux *et al.* 2007; Sladek *et al.* 2007). GWAS may also identify previously undiscovered functions of a protein. The role of the oncogene *BCL11A* in foetal haemoglobin production was unknown until a GWAS study mapped a new F cell quantitative trait locus to *BCL11A* (Menzel *et al.* 2007), even though a mouse model had previously been created (Liu *et al.* 2003) and the gene was well characterised *in vitro* (Liu *et al*. 2006a).

This discussion aims to give a brief overview of what has been achieved in this study in terms of the characterisation of TMEM114 and TMLP1 and discuss areas that merit further investigation.

8.2 Homology to proteins of known function

An initial approach to potentially identify the function of a protein is to identify homologous proteins of known function. This can subsequently be used to aid experimental design to test the function of the protein. The search for proteins homologous to TMEM114 identified a similar protein of unknown function, TMLP1. The predicted topologies and the presence of conserved motifs placed TMEM114 and TMLP1 in the Pfam00822 family of transmembrane proteins. The relationship of TMEM114 and the newly identified TMLP1 to members of the Pfam00822 family was investigated in silico and *in vitro* in this study. TMEM114 and TMLP1, although similar in size to the tight junction claudin proteins, lack the PDZ-binding domains through which claudins interact with ZO proteins (Itoh *et al*. 1999). In polarised MDCK II cells, Tmem114 and Tmlp1 failed to colocalise with ZO-1 indicating they are likely to have distinct functions to claudins.

In chapter 3 both TMEM114 and TMLP1 were shown to be more similar to voltage dependent calcium channel γ subunits, a family of eight proteins divided into two functionally distinct groups. γ1 and γ6 interact with and modulate the activity of voltage dependent calcium channel α-subunits (Ca_v) expressed in cardiac and skeletal muscle (Arikkath *et al*. 2003; Hansen *et al*. 2004; Lin *et al*. 2008). The determination of the ability of a protein to modulate calcium channel activity required resources which were not available, and so was not investigated in this study. If more evidence was accrued to strengthen the possibility that TMEM114 or TMLP1 interact with Ca_vs, then performing patch-clamp experiments to identify a modulating role in calcium channel activity would be an interesting area to pursue. The other voltage dependent calcium channel γ subunits, despite their name, have debatable effects on calcium channel activity. These γ subunits are expressed in the brain where they regulate trafficking and modify activity of AMPA receptors, and hence are known as transmembrane AMPA receptor regulatory proteins (TARPs) (Tomita *et al*. 2005). The TARPs contain key protein interaction domains in their C-termini required for AMPA receptor trafficking that are not present in the C-termini of TMEM114 or TMLP1.

TMEM114 and TMLP1 contain the W-GLW-C-C motif in their extracellular loop 1, and claudins and other Pfam00822 proteins with this motif have roles in cell adhesion (Kubota *et al*. 1999; Grey *et al*. 2003). The voltage dependent calcium channel γ subunits have cell-cell adhesion capabilities *in vitro*, although this has yet to be demonstrated *in vivo* (Price *et al*. 2005). The testing of this adhesion function, initially *in vitro* using L-cell fibroblasts, would determine if TMEM114 and TMLP1 have adhesive capacities. A positive result may suggest, although not necessarily infer, an adhesion role *in vivo*.

In mouse *Tmem114* expression was detected at E13.5 (Jamieson *et al*. 2007). At this point in lens development the primary fibres have elongated and filled the lens vesicle. Expression of *Tmem114* was upregulated postnatally (Jamieson *et al*. 2007) and although the formation of secondary lens fibres occurs throughout life, secondary fibre formation is much more active in the early postnatal period (White *et al*. 1998). Therefore the onset of *Tmem114* expression coincides with when the elongating primary fibre cells reach the anterior lens epithelial cells, and the postnatal upregulation of Tmem114 coincides with the increased levels of secondary lens fibre cell differentiation (White *et al*. 1998; Jamieson *et al*. 2007). Tmem114 localises predominantly to the apical cell membrane in polarised epithelial cells *in vitro* and in the lens expression was detected in the epithelium (Jamieson *et al*. 2007). The conserved adhesion properties in Pfam00822 proteins has lead to the hypothesis that TMEM114 acts as an adhesion molecule between the apical membrane of epithelial cells and fibre cells in the lens. An alternative hypothesis is that TMEM114 may be involved in secondary fibre cell proliferation or differentiation, as high levels of expression occur in the germinative region of the lens (Jamieson *et al*. 2007). The development of an antibody which detects endogenous Tmem114 would facilitate the precise localisation of Tmem114 *in vivo*. Using such an antibody, immunogold labelling and visualisation by electron microscopy may determine if Tmem114 is present in junctions between the cells. The generation of a knockout mouse may be the best means of testing these hypotheses (Section 8.3).

8.3 Animal models

The creation of animal models and the study of spontaneous animal mutants have greatly facilitated the identification of protein function and the aetiology of many diseases. Typically in animal models the loss of the particular protein is investigated. Thus, the functional role of a protein may be identified by determining the effect of loss of the protein on the animal's phenotype. An orthologue of TMLP1 was not identified in *X. tropicalis* and so the *in vivo* analysis in this study was restricted to TMEM114. Knockdown of Tmem114 in *X. tropicalis* resulted in a small eye phenotype suggesting TMEM114 has a role in eye growth and development. Although *X. tropicalis* is a useful tool for investigating protein function (Section 6.1), in evolutionary terms *Xenopus* is much more distant to humans than other animal models such as mouse. Also, morpholinos cause only a temporary and incomplete knockdown of the target protein. Knockout, i.e. the complete ablation, of a protein can be achieved in mouse by targeted recombination. Additionally, much more sophisticated phenotypic and behavioural analysis can also be performed in mouse than the simpler analysis performed with *Xenopus* (Boothby and Roberts 1995; Lambert *et al.* 2004). A murine model would be useful to test the role of Tmem114, particularly in the eye and brain, as subtle changes may not have been detected in *Xenopus,* and as the knockout is permanent non-developmental phenotypes could also be detected.

Developmental, but not adult, expression of Tmem114 was also detected in the kidney and all of the ADCC patients with the t(16;22)(p13.3;q11.2) balanced translocation tested have developmental kidney defects (Jamieson *et al*. 2007). Developmental kidney defects were also reported in the TMEM114 p.R2Q patient. A Tmem114 knockout mouse would provide a useful tool to study the possible role of Tmem114 in kidney development.

A number of tissues express *Tmlp1* in mouse (Figure 5.1). However, as previously discussed mouse Tmlp1 lacks the N-X-S/T glycosylation site which is present in human TMLP1 (Figure 3.4), and this appeared to affect trafficking of Tmlp1. This may complicate any analysis of a Tmlp1 knockout mouse as the extrapolation of data may be limited. It may be, however, that mice can efficiently process Tmlp1 without the need for N-linked glycans. Changes in the amino acid composition of Tmlp1 throughout evolution may have led to the need for a N-linked glycan to stabilise the structure. Thus, although a Tmlp1 knockout may be a useful tool, the costs involved and the doubts over the translation of the findings from mice to humans would suggest that a Tmlp1 knockout would not be cost-effective at this time. Therefore, the focus would be to find interacting partners for human TMLP1 using the targeted or non-targeted approaches discussed in section 8.5. Knockout mouse stem cells are available for both Tmem114 (Project: VG15800) and Tmlp1 (Project: 66468) (http://www.knockoutmouse.org/).

8.4 Association with disease

The association of sequence variants with disease can also aid in understanding protein function. The *TMEM114* gene was identified through its proximity to a balanced translocation which was associated with ADCC (Jamieson *et al*. 2007), although a direct association between mutations in the coding sequence of *TMEM114* and ADCC was not identified (Jamieson *et al*. 2007). The functional data generated in this study suggested that *TMEM114* represented a strong candidate for association with microphthalmia in humans. One heterozygous variant that was not present in controls was identified in a panel of MAC patients, but pathogenicity was not confirmed. Thus the association between *TMEM114* and microphthalmia has yet to be demonstrated, as discussed in section 8.6.

It is important to fully investigate the functional effects of mutations on protein function. For example heterozygous mutations may have a pathogenic effect by reducing the level of wildtype protein expressed (haploinsufficiency) or by having a dominant negative effect on the wildtype protein. Although in terms of genetic counselling how a dominant mutation exerts its effects is of little importance, in terms of possible future treatment it is imperative to distinguish between these effects as dominant negative effects may not be corrected simply by increasing gene expression by gene therapy or pharmacological means. The p.A147V variant, identified in ADCC patients (Jamieson *et al*. 2007) and in a patient with anophthalmia (Chapter 7) was also present in an unaffected relative implying it is a SNP. However, *in vitro*

this variant was retained in the ER possibly due to misfolding. The functional consequences of this ER retention remain to be investigated. Pulse-chase analysis could be utilised to determine if this p.A147V variant is rapidly degraded (and by which pathway) or if it is retained in the ER and causes ER stress.

Although knockout mice can prove valuable tools, genetically they represent the null phenotype. Typically diseases are more complex than simply representing the null phenotype, often being caused by gain-of-function missense mutations (Botstein and Risch 2003). The *in vitro* data generated in this study revealed that the p.A147V variant does not appear to function like wildtype TMEM114. Creation of a knock-in mouse expressing p.A146V Tmem114 (equivalent of human p.A147V) may answer a number of questions: does p.A146V in the heterozygous state cause congenital cataract or age-related cataract?; do mice homozygous for p.A146V present with microphthalmia or other congenital phenotypes?; does the C>T DNA change at the start of exon four, which causes the missense substitution, alter the splicing of Tmem114? If phenotypes were observed they could be correlated with characterisation of the mutant protein (i.e. pulse-chase analysis).

An attempt to associate *TMLP1* with human disease was not made as there was limited functional data to base a hypothesis on. However, its expression in the developing eye and brain suggest it may be a good candidate for screening if chromosome 17q25 is identified in a linkage study for association with developmental ocular or neurological disorders.

8.5 Identification of interacting protein partners

Another way to identify a possible function for a protein is to identify interacting protein partners to determine the pathway in which the protein functions. This area of research was not investigated in this study due to time restrictions. As the expression profile of Tmem114 and Tmlp1 have been identified in this study and previously (Jamieson *et al*. 2007), lysates from these specific tissues could be used as a source of potentially interacting proteins. The TMEM114 and TMLP1 proteins could be tagged with glutathione S-transferase (GST) (Roh *et al*. 2002) to probe for interacting proteins in lysates of particular tissues. The Split-Ubiquitin Membrane Yeast Two-Hybrid system was successfully employed to identify interacting proteins for claudin-16 and claudin-19 (Hou *et al.* 2008).

As transmembrane proteins expressed in the lens epithelium there is potential that TMEM114 and TMLP1 may interact. An interaction between the two proteins could be investigated using a targeted approach such as co-immunoprecipitation or fluorescence resonance energy transfer (Blasig *et al.* 2006). Other possible interaction targets are the voltage dependent calcium channel α-subunits, as TMLP1 shares a number of common features with the γ6 subunit which modulates $Ca_v3.1$ channel activity *in vitro* (Lin *et al*. 2008). Lens epithelial cells

express a number of $Ca_v3.1$ and other voltage dependent calcium channels (Meissner and Noack 2008) and these may represent good targets for interaction experiments.

8.6 TMEM114 and congenital ocular disease

Over 200 cases of congenital ocular disease have been screened for mutations in the coding sequence of *TMEM114* without a clear association. The initial screening of 130 ADCC patients was based upon the fact that *TMEM114* lies in close proximity to a balanced translocation t(16;22)(p13.3;q11.2) that associated with juvenile onset cataract (Jamieson *et al*. 2007). It was postulated that the translocation resulted in misexpression of *TMEM114* (Jamieson *et al*. 2007), but the molecular effect of the balanced translocation is unknown. To identify a possible effect of the balanced translocation on *TMEM114* expression, some of the sequence at 16p13.3 upstream of *TMEM114* from wildtype and translocated chromosomes could be cloned into luciferase reporter constructs. However, as the cell lines tested did not endogenously express *TMEM114* (Section 4.2.2), an appropriate cell line will need to be identified. This experiment would also assume that the regulatory regions of *TMEM114* are in close proximity to the transcription start site, which may not the case. This experiment also assumes that it is the dysregulation of *TMEM114* expression that is associated with disease, and not that the translocated putative *TMEM114* promoter is causing misexpression of other genes on chromosome 22q11.2.

As discussed in section 7.6, mutations in *TMEM114* appear not to be a common cause of congenital ocular disease and screening more patients, and perhaps a specific subset of patients (e.g. recessive cases), may identify causative mutations. It is also possible that the phenotype observed in *X. tropicalis* knocked down for xtTmem114 is not representative of the human phenotype. The onset of expression of *xtTmem114* occurrs at a stage of eye development earlier than was detected in mouse or humans, and hence the phenotype in *X. tropicalis* may be more severe. Although unlikely, *TMEM114* mutations may contribute to digenic disease, in which mutations in two genes are required for pathogenicity. For example, both the *peripherin/RDS* and *ROM1* genes encode transmembrane proteins expressed in the same retinal cells (Bascom *et al.* 1993) and the proteins form a multimeric complex with each other (Goldberg *et al.* 1995). Mutations in *ROM1* alone have not been associated with retinal disease but a combination of heterozygous mutations in both *peripherin/RDS* and *ROM1* are associated with retinitis pigmentosa (Kajiwara *et al.* 1994).

Finally, for reasons such as compensation, mutations in *TMEM114* may not be associated with congenital ocular disease at all. Developmental expression of *Tmem114* is relatively weak when compared to its postnatal expression (Jamieson *et al*. 2007), suggesting that perhaps it is more likely that disease phenotypes associated with Tmem114 would be of later onset (e.g. age-related cataract). It may also be that a disease associated with mutations in either TMEM114 or TMLP1 are non-ocular as *Tmem114* is also expressed in the

cerebellum, testis and developing kidney (Jamieson *et al*. 2007). In terms of cloned ESTs (Section 3.7), *TMEM114* expression is highest in carcinoids of the lung (rare neuroendocrine tumors that represent 1-2 % of lung cancers (Hage *et al.* 2003)). Altered expression of claudins and other Pfam00822 proteins are associated with cancers (Beaudry *et al*. 2010b; Jung *et al*. 2010). Alteration of claudin expression is associated with lung carcinoids (Moldvay *et al*. 2007), although unlike TMEM114 expression, claudin expression was downregulated (Moldvay *et al*. 2007). The upregulation of TMEM114 as a cause or effect of carcinoid formation may be an interesting area to investigate.

8.7 Conclusion

The overall aim of this study was to characterise the function of the novel protein TMEM114. While progress was made and a role in eye development confirmed, the precise functional role remains to be elucidated. This study has predominantly focused on the ocular aspects of TMEM114 due to its association with ADCC (Jamieson *et al*. 2007) and the small eye phenotype observed in the xtTmem114 knockdown. The expression of *Tmem114* and *Tmlp1* was also detected in other tissues and the role of these proteins in non-ocular tissues represent interesting areas to be pursued. The distinctions between TMEM114, TMLP1 and each of the other groups of proteins of known function in the Pfam00822 family made it difficult to identify a functional role for the proteins in such a short time-frame. With further *in vitro* and *in vivo* characterisation, combined with information accumulating from microarray, GWAS and massively parallel sequencing, the biological roles of TMEM114 and TMLP1 will be revealed.

Chapter 9: Appendix

9.1 Uniform Resource Locators (URL) for online resources

9.1.1 Databases

Ensembl	http://www.ensembl.org/
Genbank	http://www.ncbi.nlm.nih.gov/genbank/
IPI	http://www.ebi.ac.uk/IPI/
OMIM	http://www.ncbi.nlm.nih.gov/omim/
Pfam	http://pfam.sanger.ac.uk/
Prosite	http://expasy.org/prosite/
Unigene	http://www.ncbi.nlm.nih.gov/unigene/
Uniprot	http://www.uniprot.org/

9.1.2 Prediction tools

Compute pI/MW	http://expasy.org/tools/pi_tool.html
CSS-Palm 2.0	http://csspalm.biocuckoo.org/
HMMTOP	http://www.enzim.hu/hmmtop/
NetNGlyc 1.0	http://www.cbs.dtu.dk/services/NetNGlyc/
NetOGlyc 3.1	http://www.cbs.dtu.dk/services/NetOGlyc/
NetPhos2.0	http://www.cbs.dtu.dk/services/NetPhos/
NetPhosK 1.0	http://www.cbs.dtu.dk/services/NetPhosK/
Philius	http://www.yeastrc.org/philius/pages/philius/runPhilius.jsp
PolyPhen	http://genetics.bwh.harvard.edu/pph/
SIFT	http://sift.jcvi.org/
SignalP	http://www.cbs.dtu.dk/services/SignalP/
Sulfinator	http://expasy.org/tools/sulfinator/
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/
TMMOD	http://liao.cis.udel.edu/website/servers/TMMOD/
TMPRED	http://www.ch.embnet.org/software/TMPRED_form.html
TOPPRED	http://www.sbc.su.se/~erikw/toppred2/

9.1.3 Other tools

ALIGN	http://xylian.igh.cnrs.fr/bin/align-guess.cgi
Blast	http://blast.ncbi.nlm.nih.gov/
ClustalW	http://www.ebi.ac.uk/Tools/clustalw2/index.html
Translate	http://expasy.org/tools/dna.html

9.1.4 Downloaded Software

BioEdit	http://www.mbio.ncsu.edu/bioedit/bioedit.html
EZ-C1	http://www.nikon-instruments.jp/eng/service/download/DLList.aspx?CID=3
Seaview	http://pbil.univ-lyon1.fr/software/seaview.html

9.2 Database accession numbers for DNA and protein sequences

Table 9.1 *Homo sapiens* DNA and protein sequences

Name	DNA	Protein sequences for alignments	Ensembl accession numbers for gene structure
TMEM114	NM_001146336.1	NP_001139808.1	-
TMLP1	XM_941058.2	XP_946151.1	-
CLARIN-1	-	NP_777367.1	-
CLAUDIN-1	-	NP_066924.1	ENSP00000295522
CLAUDIN-2	-	NP_065117.1	ENSP00000336571
CLAUDIN-3	-	NP_001297.1	-
CLAUDIN-4	-	NP_001296.1	-
CLAUDIN-5	-	NP_003268.2	-
CLAUDIN-6	-	NP_067018.2	-
CLAUDIN-7	-	NP_001298.3	-
CLAUDIN-8	-	NP_955360.1	-
CLAUDIN-9	-	NP_066192.1	-
CLAUDIN-10a	-	NP_878268.1	-
CLAUDIN-10b	-	NP_008915.1	-
CLAUDIN-11	-	NP_005593.2	-
CLAUDIN-12	-	NP_036261.1	-
CLAUDIN-14	-	NP_036262.1	-
CLAUDIN-15	-	NP_055158.1	-
CLAUDIN-16	-	NP_006571.1	ENSP00000264734
CLAUDIN-17	-	NP_036263.1	-
CLAUDIN-18.1	-	NP_057453.1	-
CLAUDIN-18.2	-	NP_001002026.1	-
CLAUDIN-19	-	NP_683763.2	-
CLAUDIN-20	-	NP_001001346.1	-
CLAUDIN-22	-	NP_001104789.1	-
CLAUDIN-23	-	NP_919260.2	-
EMP-1	-	NP_001414.1	ENSP00000410755
EMP-2	-	NP_001415.1	ENSP00000352540
EMP-3	-	NP_001416.1	ENSP00000270221
PMP22	-	NP_000295.1	ENSP00000409824
LIM2/MP20	-	NP_001155220.1	ENSP00000221973
CLP24	-	NP_078876.2	ENSP00000253934
PERP	-	NP_071404.2	ENSP00000397157
TMEM47	-	NP_113630.1	ENSP00000275954
CACNG1	-	NP_000718.1	ENSP00000226021
CACNG2	-	NP_006069.1	ENSP00000300105
CACNG3	-	NP_006530.1	ENSP00000005284
CACNG4	-	NP_055220.1	ENSP00000262138
CACNG5a	-	NP_665810.1	ENSG00000075429
CACNG5b	-	NP_055219.1	-
CACNG6	-	NP_665813.1	ENSP00000252729
CACNG7	-	NP_114102.2	ENSP00000222212
CACNG8	-	NP_114101.4	ENSP00000270458
GAPDH	NM_002046.3	-	-
SOCS3	-	NP_003946.3	-

Table 9.2 *Mus musculus* DNA and protein sequences

Name	DNA	Protein
Tmem114	NM_029070.2	NP_083346.1
Tmlp1	NM_001085535.1|	NP_001079004.1
Hprt	NM_013556.2	-
Socs3	-	NP_031733.1

Table 9.3 *Xenopus tropicalis* DNA and protein sequences

Name	DNA	Protein
Tmem114	XM_002939438.1	XP_002939484.1
Lim2	NM_001102779.1	-
Gapdh	NM_001004949.1	-

Table 9.4 *Danio rerio* DNA and protein sequences

Name	DNA	Protein
Claudin-2	-	NP_001004559.2
Claudin-7	-	NP_001002340.1
Claudin-10	-	NP_001007038.1
Claudin-11	-	NP_571847.2
Claudin-15	-	NP_956698.1
Claudin-I	-	NP_571843.1
Claudin-J	-	NP_571844.1
Claudin-K	-	NP_001003464.1
Cacng1	-	NP_999849.1
Cacng2	-	NP_956935.1
Cacng3-like	-	XP_693299.1
Cacng5-like	-	XP_001337872.1
Cacng6	-	NP_001076560.1
Cacng6-like	-	XP_001342833.1
Cacng8-like	-	XP_001342922.2
Socs3a	-	NP_956244.1
Socs3b	-	NP_998469.1

Table 9.5 TMEM114 orthologues DNA and Protein Sequences

Name	DNA	Protein
Homo sapiens	NM_001146336.1	NP_001139808.1
Macaca mulatta	-	NP_001181105.1
Mus musculus	NM_029070.2	NP_083346.1
Rattus norvegicus	-	NP_001128110.1
Bos taurus	-	NP_001178169.1
Canis familiaris	XM_547134.2	XP_547134.2
Gallus gallus	-	XP_001235110.1
Monodelphis domestica	-	XP_001367110.1
Xenopus tropicalis	XM_002939438.1	XP_002939484.1
Danio rerio	-	XP_001922308.1

Table 9.6 TMLP1 orthologues DNA and Protein Sequences

Name	DNA	Protein
Homo sapiens	XM_941058.2	XP_946151.1
Mus musculus	NM_001085535.1	NP_001079004.1
Rattus norvegicus	-	XP_001081744.1
Bos taurus	-	XP_616745.3
Canis familiaris	XM_548670.1	XP_548670.1

9.3 Primer Information

9.3.1 Sequencing Primers

Table 9.7 Primers used to sequence coding regions of *TMEM114*
Tm = Annealing temperature

Primer	Sequence 5'-3'	Tm	Additional information
Exon 1F	GACGGGGGCAGGAGCAGAG	62.5 °C	5 % DMSO
Exon 1R	GAGGAGTGCGCGCCAAGC	62.5 °C	5 % DMSO
Exon 2F	CAGACGAGGGTGCTGCACAGA	62.5 °C	5 % DMSO
Exon 2R	CCCGGCTGTTTTGGGAAAGG	62.5 °C	5 % DMSO
Exon 3F	CGTAGGGGCAGTGACAGTGTGG	62.5 °C	5 % DMSO
Exon 3R	GGATCCCCCTCGTTTGCTGA	62.5 °C	5 % DMSO
Exon 4F	GCGAGGAGGAGGAGGGGAGT	62.5 °C	5 % DMSO
Exon 4R	CGATGGAGATCGGTCGGTGA	62.5 °C	5 % DMSO

Table 9.8 Primers used to sequence plasmids

Primer	Sequence 5'-3'	Tm	Additional information
M13 F	GTAAAACGACGGCCAG	45 °C	-
M13 R	CAGGAAACAGCTATGAC	45 °C	-
T7 F	TAATACGACTCACTATAGGG	50 °C	-

9.3.2 RT-PCR primers

Table 9.9 Primers used to detect *Homo sapiens* cDNA expression. h = *Homo sapiens*. m = *Mus musculus*. xt = *Xenopus tropicalis*. cf = *Canis familiaris*

Primer	Sequence 5'-3'	Tm	Additional information
hTMEM114 exon 1F	GCTCAGCTTTGTGCTCCTG	58.8 °C	-
hTMEM114 exon 2R	GCTGACTGTCACGTTCTCCA	58.8 °C	-
hTMLP1 exon 1F	GCCAGCGACTACTGGTACATC	55 °C	-
hTMLP1 exon 4R	CTGAGAGCAGGAGGGTTCC	55 °C	-
hGAPDH F	GAAGGTGAAGGTCGGAGTC	60 °C	-
hGAPDH R	GAAGATGGTGATGGGATTTC	60 °C	-
mTmem114 exon 1F	AGAGCCTCTGAGCTCCCACT	58.5 °C	-
mTmem114 exon 2R	GCTGACCGTCACATTCTCCT	58.5 °C	-
mTmlp1 exon 1F	CCAGCGACTACTGGTACATCC	55 °C	-
mTmlp1 exon 4R	GAAGCTGATGTGCACGTTCT	55 °C	-
mHprt F	GCTGGTGAAAAGGACCTCT	52.5 °C	-
mHprt R	CACAGGACTAGAACACCTGC	52.5 °C	-
xtTmem114 5'UTR F	TGGATTCCTCTTCTGTATCTGC	55.3 °C	-
xtTmem114 exon 2R	CCAAGCCGAAAGGGATTTAT	55.3 °C	-
xtGapdh F	GCAGTGTATGTGGTGGAATCT	45 °C	-
xtGapdh R	AAGTTGTCATTGATGACCTTTGC	45 °C	-
cfTmem114 exon 3F	CAGCCTGATCCTGATGGTTT	52 °C	-
cfTmem114 exon 4R	GGTCCAGGAGGGCTTTCTC	52 °C	-
cfTmlp1 exon 1 F	GGTGGCCAGCGACTACTGGT	53.5 °C	-
cfTmlp1 exon 4 R	GGAACTGAAGAGGCCGCAGA	53.5 °C	-
cfGapdh F	GAAGGTGAAGGTCGGAGTC	55 °C	-
cfGapdh R	GAAGATGGTGATGGGATTTC	55 °C	-

9.3.3 Primers used for cloning

Table 9.10. Primers used for cloning. Restriction enzyme sites are coloured in blue. Termination codons are coloured in red. V5 tags are coloured in green.

Insert	Primer	Sequence 5'-3'	Tm
mTmem114	mTmem114 Xba I F	GGCTATCTAGAACCATGCGGGTGCGCCTG GGCGCTCTG	60 °C
	mTmem114 Eco RI R	GGCTAGAATTCTCAGATGGCTTGGTCCTG CCTCTGGCTCAG	60 °C
mTmem114_SPM	mTmem114_SPM XbaI F	GGCTATCTAGAACCATGACCGACTTCTGGT ACATCATCG	60 °C
	mTmem114 Eco RI R	GGCTAGAATTCTCAGATGGCTTGGTCCTG CCTCTGGCTCAG	60 °C
mTmem114_SPM-V5	mTmem114_SPM XbaI F	GGCTATCTAGAACCATGACCGACTTCTGGT ACATCATCG	60 °C
	mTmem114 V5 R	CCTCACGTAGAATCGAGACCGAGGAGAGG GTTAGGGATAGGCTTACCGATGGCTTGGT CCTGCCTCTGGCTCAG	60 °C
mTmlp1	mTmlp1 Xba I F	GGCTATCTAGAACCATGCGGGTGCGCCTG GGCGCTCTG	64 °C
	mTmlp1 Eco RI R	GGCTAGAATTCTCAGAGGATCACAGAGTG GGGCACTC	64 °C
mTmlp1-V5	mTmlp1 Xba I F	GGCTATCTAGAACCATGCGGGTGCGCCTG GGCGCTCTG	64 °C
	mTmlp1 V5 Eco RI R	GGCTAGAATTCTCACGTAGAATCGAGACC GAGGAGAGGGTTAGGGATAGGCTTACCGA GGATCACAGAGTGGGGCACTC	64 °C
mTmlp1_SPM	mTmlp1_SPM XbaI F	GGCTATCTAGAACCATGTGGTACATCCTG GAGGTGGCG	64 °C
	mTmlp1 Eco RI R	GGCTAGAATTCTCAGAGGATCACAGAGTG GGGCACTC	64 °C
mTmlp1_SPM-V5	mTmlp1_SPM XbaI F	GGCTATCTAGAACCATGTGGTACATCCTG GAGGTGGCG	64 °C
	mTmlp1 V5 Eco RI R	GGCTAGAATTCTCACGTAGAATCGAGACC GAGGAGAGGGTTAGGGATAGGCTTACCGA GGATCACAGAGTGGGGCACTC	64 °C
xtTmem114-5'UTR-V5	xtTmem114 5'UTR F	TGGATTCCTCTTCTGTATCTGC	60 °C
	xtTmem114-V5clone_R	CCCTCACGTAGAATCGAGACCGAGGAGAG GGTTAGGGATAGGCTTACCTATAACCTGCT CCCGTCTTCTT	60 °C
xtTmem114-EST	xtTmem114_in situ F	GGAATGCTTTTCCTGTTTGG	60 °C
	xtTmem114_in situ R	GCTTTCCGTTTTCAGAAGCA	60 °C
xtLim2	xtLim2 F	GGATGATGTACAGCTTTATGGGAGGAGG	60 °C
	xtLim2 R	GGTTAACGGGGTCCTGACCCTCG	60 °C

Table 9.11 Site-directed mutagenesis primers

Clone	Primer	Sequence 5'-3'	Tm
mTmem114-V5 p.R2Q	R2Q F	TCTAGAACCATGCAGGTGCGCCTGGGC	68 °C
	R2Q R	GCCCAGGCGCACCTGCATGGTTCTAGA	68 °C
mTmem114-V5 p.P52L	P52L F	GGGACCAGGGACTGGCCAACCGCAG	68 °C
	P52L R	CTGCGGTTGGCCAGTCCCTGGTCCC	68 °C
mTmem114-V5 p.N54K	N54K F	GGGACCGGCCAAACGCAGCCAGCAAGAGCC	68 °C
	N54K R	GGCTCTTGCTGGCTGCGTTTGGCCGGTCCC	68 °C
mTmem114-V5 p.N88K	N88K F	TTCTGGCAGGAGAAAGTGACGGTCAGCGAC	68 °C
	N88K R	GTCGCTGACCGTCACTTTCTCCTGCCAGAA	68 °C

Table 9.12 TMEM114 promoter amplification and sequencing primers

Primer Use	Primer name	Primer sequence	Tm
Amplification	Amplification_F	CCACCTTCAGGTAGGCCCCAGTGTTTGTTGC	68 °C
Amplification	Amplification_R	TCTGGATTTCCGCCCCAGCCCTTTTTCTTGGC	68 °C
Sequencing	Amplification_F	CCACCTTCAGGTAGGCCCCAGTGTTTGTTGC	68 °C
Sequencing	WTPromSeqStep1F	TGCTGATGCCCTTGATAATG	60 °C
Sequencing	WTPromSeqStep2F	CCTTACGAATGGATCACAGGA	60 °C
Sequencing	WTPromSeqStep3F	ATGGTAATGACCGTGGTGGT	60 °C
Sequencing	WTPromSeqStep4F	TGGTGGTGGTAATGGTTGG	60 °C
Sequencing	PromGap_R	CATCACCCACGGACAACAC	60 °C
Sequencing	Gap259bp_F	GGCAATGGTGGTGATAGTGA	60 °C

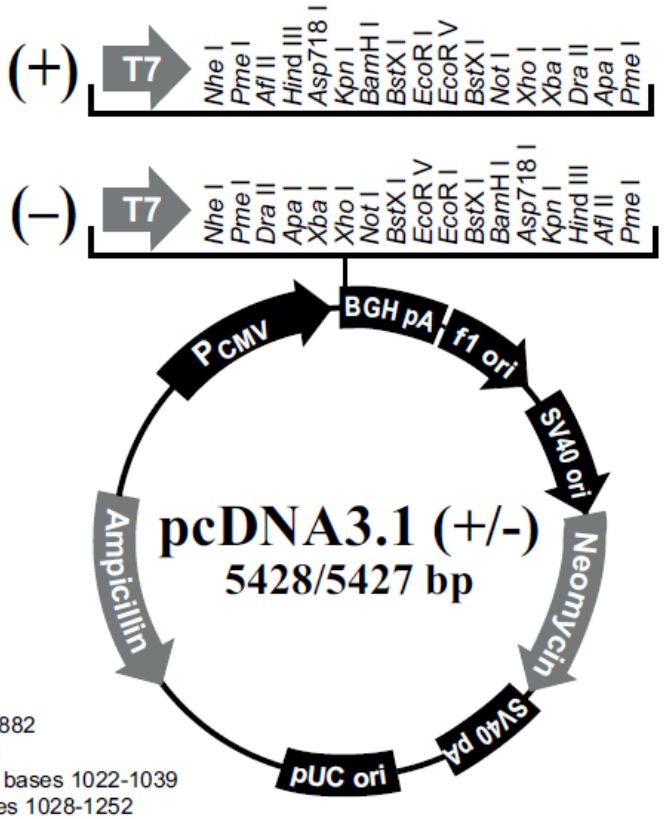
9.4 Constructs used in this study

Table 9.13 Constructs used in this study

Construct Name	Template	Cloned by
pcDNA3.1(-) mTmem114	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-SPM	pcDNA3.1(-) mTmem114	GJ Maher
pcDNA3.1(-) mTmem114-V5	RIKEN 4930511J11 cDNA IMAGE 5357758	Dr. J Urquhart
pcDNA3.1(-) mTmem114-V5-SPM	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-V5 p.R2Q	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-V5 p.P52L	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-V5 p.A146V	pcDNA3.1(-) mTmem114-V5	Dr. J Urquhart
pcDNA3.1(-) mTmem114-V5 p.N54K	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-V5 p.N88K	pcDNA3.1(-) mTmem114-V5	GJ Maher
pcDNA3.1(-) mTmem114-V5 p.N54K N88K	pcDNA3.1(-) mTmem114-V5 p.N88K	GJ Maher
pcDNA3.1(-) mTmlp1	*Mus musculus* testis cDNA	GJ Maher
pcDNA3.1(-) mTmlp1-SPM	pcDNA3.1(-) mTmlp1	GJ Maher
pcDNA3.1(-) mTmlp1-V5	pcDNA3.1(-) mTmlp1	GJ Maher
pcDNA3.1(-) mTmlp1-V5-SPM	pcDNA3.1(-) mTmlp1-V5-SPM	GJ Maher
pGEM T easy xtTmem114-5'UTR-V5	*X. tropicalis* stage 45 eye cDNA	GJ Maher
pGEM T easy xtTmem114-EST	*X. tropicalis* stage 45 eye cDNA	GJ Maher
pGEM T easy xtLim2	*X. tropicalis* stage 45 eye cDNA	GJ Maher

9.5 Vector maps

Figure 9.1 pcDNA3.1(-) vector map and sequence reference points. (Invitrogen catalogue number V795-20)



Comments for pcDNA3.1 (+)
5428 nucleotides

CMV promoter: bases 232-819
T7 promoter/priming site: bases 863-882
Multiple cloning site: bases 895-1010
pcDNA3.1/BGH reverse priming site: bases 1022-1039
BGH polyadenylation sequence: bases 1028-1252
f1 origin: bases 1298-1726
SV40 early promoter and origin: bases 1731-2074
Neomycin resistance gene (ORF): bases 2136-2930
SV40 early polyadenylation signal: bases 3104-3234
pUC origin: bases 3617-4287 (complementary strand)
Ampicillin resistance gene (*bla*): bases 4432-5428 (complementary strand)
ORF: bases 4432-5292 (complementary strand)
Ribosome binding site: bases 5300-5304 (complementary strand)
bla promoter (P3): bases 5327-5333 (complementary strand)

Figure 9.2 pGEM T easy vector map, sequence reference points and sequence of the promoter and multiple cloning site (Promega catalogue number A1360)

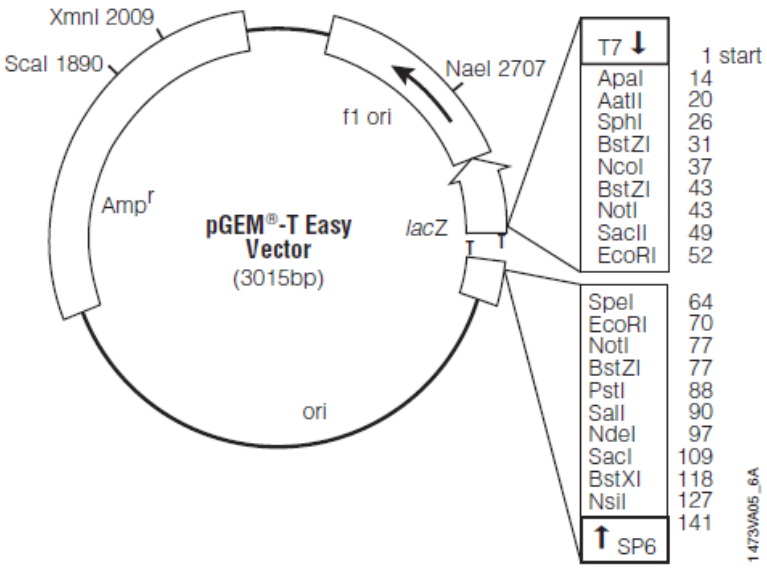


Figure 3. pGEM®-T Easy Vector circle map and sequence reference points.

pGEM®-T Easy Vector sequence reference points:

T7 RNA polymerase transcription initiation site	1
multiple cloning region	10–128
SP6 RNA polymerase promoter (–17 to +3)	139–158
SP6 RNA polymerase transcription initiation site	141
pUC/M13 Reverse Sequencing Primer binding site	176–197
*lac*Z start codon	180
lac operator	200–216
β-lactamase coding region	1337–2197
phage f1 region	2380–2835
lac operon sequences	2836–2996, 166–395
pUC/M13 Forward Sequencing Primer binding site	2949–2972
T7 RNA polymerase promoter (–17 to +3)	2999–3

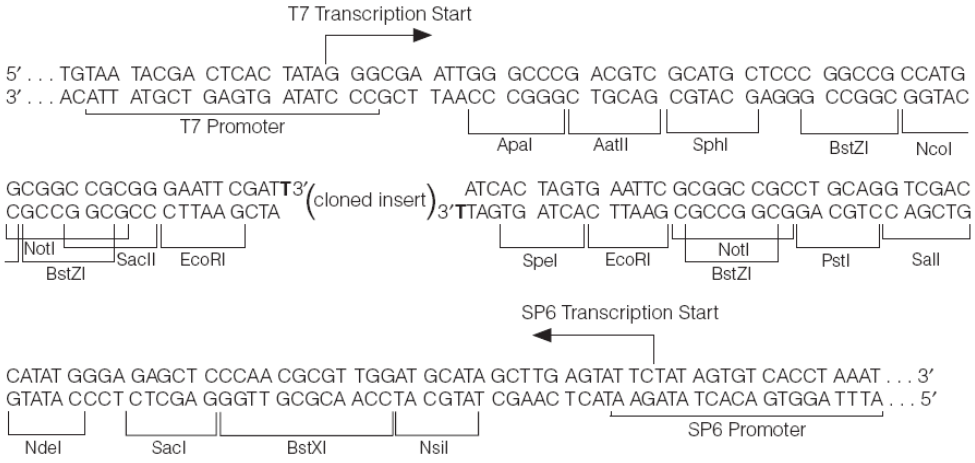
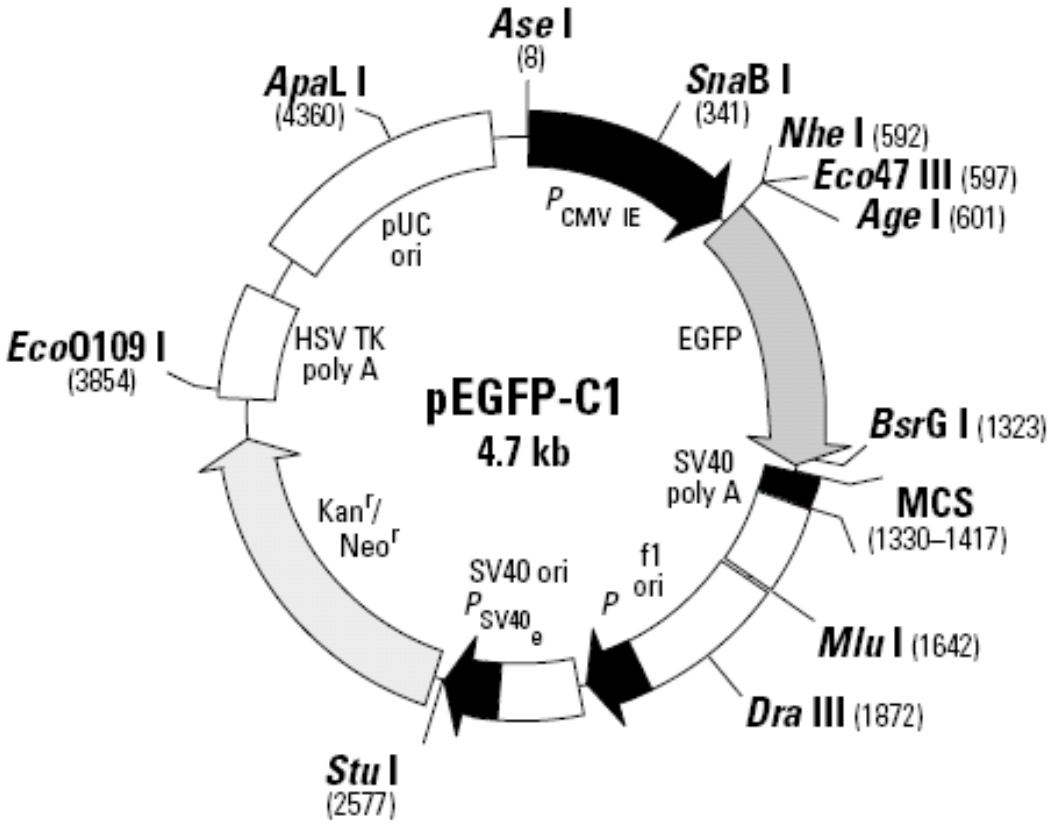


Figure 9.3. pEGFP-C1 vector map and sequence reference points (Clontech catalogue number 6084-1)



9.6 Sequence of the gap in the upstream sequence of *TMEM114.*

Bold = forward primer. Italic = reverse primer. Colour represents sequence identified by that specific primer. Underlined = new sequence. Grey font = 5'UTR. Black font = coding sequence. The highlighted sequence was submitted to Genbank (accession GQ267544.1)

NW_001838342

ccaccttcaggtaggccccagtgtttgttgctcccctctatgtgtccatgtgttcttatcatttagctccc
acttgtaagtgaaaacatgcaatatttggttttctgttccggtgttaatttgctaaggataatggcctcca
gctccatccgtgttcctgcaaaggacataatcttgttctttttatggcagcatagtattccatggtatag
cctaccacattttctttatccaatctaccattgatgggcatttaggttgattctatatctttgctatcata
gttggactcaaatatttagggtgggtctgacttctcccactctggctccaagagtagacgggtgactgagg
cttggccaattagcacatctggtccctttgaatgatagagattgattcagagacgggcacatgaccaaaaa
caggccaatgagatttaattctggggatctggcactattgagagagagattatttcttttttcttggagctg
ctaagctgtgaaactggctcctatgattgaacaaatgagagcctgcttgaaaaatgaagtcaagcaagagg
aaaacaactgagaggtggcaagaagaagaggaagagggaagagattgaatctcaaagacattgtaagggct
cctgggatcagcctcaactgaattttacctacatctctcaagtcgacggtgaggactattttaatgatggt
ggaagtaaagataacagtagtagtgtgatggaggtgacagcaaaattatcagtgatgatggttttggaaaa
gatctatcttcatcaagatagtggtggcggggtgatgaaacagaatggtctgcagaatcaagggagagtgt
attcttgctggtgatagctctgccaggg**tgctgatgcccttgataatg**aagatgaggatgacagtgatgta
gattaaggtggtattgggaaggaagaataaagggaaaggataactctcaaccccagcctatcaccgagccc
taacaactctactccctcaacttttaaaaactctatccacttcactctctgctcgctgctgcccccaccct
agcccaagtcagcaggatcttgcttcttcttctacgatgtcctcctaagtggccttcttgttccctttctt
gtctctctatattctatattccacactacagcaagcatagtgagatttttttaatacgtaagttagaacat
tttacttttgtgctgaaaacggttcagtggtttggactaaaattgaactctttatcatggatctcaagagc
tgcatgacctgacaaattattataccctcaatctcatttccctctggttgctcttcccttactacattcta
gcctgatggcccttctttcgatttctctagaatggaccaaaccccctctataagccattctctagacacgga
tacttttataaccttcatgtctcagcatcaatgtcacttcctcagagaagcccagtgagatctcccctctg
atactttctctcatagcaccctgtcgtaagcacgctttgatgattctcttccaacctgaccagcacttacc
aacttttggaccttgatacaatggtc**ccttacgaatggatcacagga**gtgccaagattttgattcctcagc
cataaggtcagcattgaggggctttagtagctgtaactcctgaggattggtcatccccagtatctggaagc
ctcttctgttaatgagccatgcaaatataggtattttctctgtgtcccctactttgtaaaagtctgggaag
cctgaatcctatatttcttaagggcaggacttgtttttggattgttcatgcctatattcttaacatagact
atagcatcagacacatagaaggtaccggaagtatttcttgaaggagtaaaaggttttatggtggtgaagaa
gatggtggaattgataagagtgatggtgatgataaagaggagagtggtgctggtgctggtggtggtagtga
tgatgatggtgatggtcatggtggtggtg**atggtaatgaccgtggtggt**gatagtggtaatggtggtgatg
atggtggtgatgatgatggtggtgatagtggtaatggtggtgatgatggtggtgatgatgatggtgatggt
ggtgatga**tggtggtggtaatggttgg**tggtgatggtggtggttggtggtgatggtggtggtgatgatgat
ggtgatgacgat**ggcaatgatggtggtagtaa**tgatggtggtggtggtgatggtggtgatggtggtgatgg
tggtggtggtggtgatggtggtgatggtggtgttgtctgtggtgatggtgatggtggtggtgatggtggtg
gtgatggtgatggtggtgatggtggtgttgtccatggtgatggtgatggtggtggttgtggtgatggtggt
ggtgatgatgatgg*tgatgacgatggcaatgg*tggtgatagcgatgatggtggtggtggtgatggtggtga
tggtggtggtggtgggtggtgatggtggtgaagatgatgtggtgatggtagtgatgatggcggtgttggtg
gtgatgatggtgatggtggtgatggtg*gtgttgtccgtgggtgatg*gtgatggtggtggtggtgatgatga
tgatggtgatggtggtggtgttggtgatgctgctgctggtgatgatggtggtgatgatgatgatggtgatg
gtagtgatgacggtgttgttgttggtgatgatgatgatggtgatggtggtgctgttggtgatgctgctgct
ggtgatggtggtgatgatgatggtgatggtagtgatgattgtgttgttgttggtggtgatgttgatgatgg
tgatggtggtggtggtgatgatggtggcaatgctggtggctatgataatgaagaaagactttgggggtaga
aatgatagattctgaggacctgaactacctcagtgctttagagaacttttatgaaaaattaatgtgccttt
tcacattctagtcagcccagactacactcgaaggtggattttggaggaagggtacttggagctggaggggc
tcttggagatgatttgcaaatcctggcgggactgtgaggcccgccacacccctccttctttctgccccagg
tggtcccaaggccgcctgcccttccacaccggcagcctccagcgctccagggctggggaccgcccccgtcc
atgctgcgtgcgcactccgagcttttcctcggcgaggaggagggctcagagggtcagctgcgccccggcca
gagcaggagctgcgggaagaccctggagagcagggctgggagtggggatctggcagtgtgggg*gccaagaaa*
*aagggctggggcggaaatccaga*GACAGAGGCAGAGAGGGGAGCATCTGGGATCGTGGGGAGGGTTCGTGG
AGAGTGAGGCAGGGGAATGGAGGGAGTAGGAGTCTGGGGAAAGTGAGGATTTTGAAGTGAGAGAGTCCGCT
TGGGCTGGGCGTGGAGGGCCGAGGGTGGAGTGGGGTTGAGGGACGAGGTGGGTTAAGGTGCTGGTGGGGAA
GGACTAGAGCTTCCAGGAGGGTCTGAAGGCGGAGGTGGAGGTCGGACAGCACAAGTGAGAGTGAGGGGGTC
TGTAGGTAGAGCTGGCATTGGAGGGGTGGGAACGTGGACGGCGGCTGAGCGGGGGGCGCGTGCAGGTCGGA
GCCGGCAGGTGAGGAGCCAGGGTCTAAGCTAGGGCCCGGGGTCCAAAGGAAAGGCTTTCAGATCTGAGCTC
CGCGGCTCTGGCGGCGGTGGAGCTGGGAGCCCGCGGTGGCCGGCTGGGGACGGGGGCAGGAGCAGAGGGGT
CGCGCCGCGGCCACTGGAACCCGCGGCTGCTGGTCCTGGGCTCGGCG**ATG**CGGGTGCACCTGGGCGGGCTG
M R V H L G G L

9.7 Clinical phenotypes of patients screened for mutations in *TMEM114*

Table 9.14. Full clinical phenotypes and DNA changes of patients screened for mutations in *TMEM114.* Patients are categorised according to clinical phenotype.

Code	Ocular phenotype	Non-ocular phenotype	DNA change	Protein change
B7	bilateral anophthalmia	-	-	-
B10	bilateral anophthalmia	-	-	-
C7	bilateral anophthalmia (familial)	-	-24C>T (homo); c.440 -7C>G (homo)	-
D9	bilateral anophthalmia	-	'-24C>T (homo)	-
E7	bilateral anophthalmia	-	-24C>T; c.440 -7C>G	-
H9	bilateral anophthalmia	-	-	-
B1	bilateral anophthalmia	bilat sensory neural deafness, axial hypotonia and developmental delay.	-	-
D10	bilateral anophthalmia, optic nerve hypoplasia, atrophic chiasm	interhemispheric cyst, seizures, cup shaped ears, agenesis of corpus callosum	-24C>T; c.440C>T	p.A147V
E1	bilateral anophthalmia	microcephaly, spastic dystonic quadriplegia, ASD, MR, laryngeal cleft, diaphragmatic eventration, 2-3 toe syndactyly, normal chromosomes	-	-
E8	bilateral A	saddle nose, dysmorphic ears and navel, micrognathia, hypospadias, aortic coartation	-	-
F8	bilat A, ON agenesis	rhizomelia, contractures	-24C>T; c.440 -7C>G	-
G2	bilateral anophthalmia	cleft palate, cutaneous syndactyly of left III-IV fingers and of bilateral II-III toes, microcephaly, severe mental delay. Brain MRI: hydrocephalus, vermian hypoplasia and corpus callosum hypogenesia.	-	-
H3	bilateral anophthalmia	right microtia, profound developmental delay and mental retardation, no speech, difficulty standing and walking, hypotonic, absent left kidney, MRI scan pending	-24C>T; c.440 -7C>G	-
A6	M, A	-	-24C>T; c.13C>T; c.440 -7C>G	p.L5L
A10	anophthalmia (right) and severe microphthalmia (left)	-	-	-
B6	anopthalmia, microphthalmia, arhinia	-	-	-
C6	M, A	-	-	-
F5	Anophthalmia Unilateral & Microphthalmia Unilateral	-	-24C>T	-
G5	Micropthalmia and anopthalmia.	-	-24C>T	-
H7	A and severe M,	-	-	-
C10	Anophthalmia (left), microphthalmia (right), blind	bilateral congenital choanal stenosis/atresia, facial hemimicrosomia	-	-
D2	anophthalmia (right) and microphthalmia	bilateral cleft lip and choanal atresia, micropenis, cryptorchidism, developmental delay, increased liver dimension, unilateral (right) kidney ectopia with rotation defect, severe anomalies on X-ray	-	-

D6	L anophthalmia, R microphthalmia	pituitary hypoplasia, hypoplastic optic chiasm, hypothalamic abnormality per AJB, hypoplastic CC, CLP.	-	-
H1	left sided anophthalmia & right sided severe microphthalmia	mental retardation, bilateral cleft lip & palate and a scoliosis, aplasia of septum pellucidum	c.440 -7C>G	-
H2	Septo-optic dysplasia with anophthalmia OD/microphthalmia OS,	severe developmental delay	c.440 -7C>G	-
C11	A, M with C, optic disc hypoplasia	-	-24C>T	-
D7	left A, right severe M with coloboma	repetitive hand movements, rocking and echolalia. Normal chrs	-	-
A4	Bilateral microphthalmia & congenital cataracts, secondary glaucoma, nystagmus and global developmental delay	-	-	-
A9	bilateral microphthalmia and cataracts	-	-	-
B8	bilat M and sclerocornea	-	-24C>T	-
B9	bilateral microphthalmia and corneal opacities	-	-24C>T; c.440 -7C>G	-
C5	Microphthalmia & Cataracts Bilateral.	-	-	-
D8	bilateral microphthalmia, cataracts	-		
F6	bilateral microphthalmia	-	-	-
G7	bilat severe M, corneal opacification, iris abnormalities	-	-	-
B3	Bilateral microphthalmia & schleroderma of cornea bilaterally	facial dysmorphism, microcephaly, short stature & developmental delay	-	-
D1	bilat microphthalmia	renal reflux, cryptorchidism, delayed speech and walking, PDA that resolved	-24C>T	-
E6	bilateral microphthalmia and sclerocornea	L undescended testes, chordee with penile torsion, very severe microcephaly, complete agenesis of the corpus callosum	-24C>T (homo); c.440 -7C>G (homo)	-
E9	bilateral microphthalmia and cataract, iris hypoplasia, corneal haze, ectopia uvea	infantile spasms	-	-
F1	microphthalmia, aniridia, totally blind	developmental delay, pervasive disorder?	-	-
F9	Bilateral microphthalmia, left cornea transplant (unsuccessful), rudimentary lens & possible iris, right sclerocornea & glaucoma, senses light.	Respiratory tract infection, cerebellar hypoplasia (vermis), streak gonads?, empty scrotum, small penis & corpus callosum agenesis	-	-
G6	bilateral aniridia and microphthalmia	limbal stem cell deficiency with early pannus, keratopathy, corneal oedema, cataract, uveal coloboma.	c.440 -7C>G	-
G8	severe bilateral microphthalmia/anophthalmia	diaphragmatic hernia	-24C>T; c.440 -7C>G	-
H4	Severe bilateral microphthalmia	Severe undergrowth, unilateral cleft lip/palate repaired, hypoplastic midface, clipped ears, anterior encephalocele repaired, cryptorchidism, severe to profound mental retardation, no speech		

H5	Bilateral anopthalmia/micropthalmia	congenital abnormalities	-	-
H8	bilat M, sclerocornea	Pierre Robin Syndrome	-	-
A7	bilateral microphthalmia, colobomas, retinal dystrophy		c.134C>T	p.P45L
A11	bilat M and C	-	-	-
B12	bilat M, C	-	-	-
D5	Microphthalmia & Coloboma Bilateral.	-		
D11	bilat M, C, heterochromia, cataract	-	-	-
E10	bilateral microphthalmia, coloboma, acentric pupils	-	-	-
E11	bilat M and C	-	-	-
G11	bilat M, C, heterochromia	-	-24C>T	-
H12	bilat M and C and cataracts, abnormal anterior segment	-	-24C>T (homo)	-
G3	Bilateral severe microphthalmia and iris, retinal and optic cup coloboma	cryptorchidism, hypotonia, microcephaly, severe mental retardation	-	-
G10	bilateral microphthalmia, retinal coloboma	absent right kidney, cystic dyspastic left kidney, bilateral cryptorchidism, adducted left thumb, pseudoarthrosis of right clavicle	-24C>T c.5G>A	p.R2Q
B11	bilat M, unilat C	-	-24C>T	-
C9	congenital bilateral microphthalmia, translucency of cornea, coloboma on MRI	-	-	-
C12	severe bilat M, left iris C, retinal defect	-	-	-
C2	bilateral microphthalmia and cataracts, unilateral coloboma	microcephaly, seizures, self-aggressive and autistic behaviour, severe mental retardation	-	-
F10	unilateral (left) microphthalmia with cataract,	-	-	-
A2	left microphthalmia	v high palate, joint laxity, scoliosis, left accessory nipple, learning problems, chr15q12.2 del	-24C>T; c.440 -7C>G	-
B4	Unilateral severe microphthalmia	severe short stature, hypospadias, 2/3 sundactyly and mild developmental delay	-24C>T; c.440 -7C>G	-
F11	left M, right C, abnormal retina	-	-24C>T; c.440 -7C>G	-
C3	Microphthalmia left and coloboma right,	developmental delay, some dysmorphic features including a single palmar crease in right hand and umbilical hernia	-	-
D3	unilateral microphthalmia (left), retinal coloboma	bilateral sensorineural deafness, developmental delay, ankyloglossia, dysplastic ear, micropenis	-24C>T	-
H6	left microphthalmia, right coloboma	ambiguous genetalia and cleft right hand, complex phenotype	-	-
E5	Optic Fissure Closure Defect Bilat. Possible Retinal Degeneration		-	-

A1	bilateral coloboma/mild aniridia	epilepsy, pachygyria, hypotonia, developmental delay	-	-
A3	bilateral iris and choroiretinal coloboma	developmental delay, dilated renal pelvis, laryngomalacia, undescended testis.	-	-
D12	unilat M, bilat C and cataracts		-	-
E12	unilat M, bilat C and cataracts		-	-
F12	unilat M, bilat C, iris pupillary membrane remnant, cataract		-	-
A12	left M, bilat C, heterochromia		-	-
B2	unilateral (R) microphthalmia, bilateral coloboma		-24C>T (homo); c.440 -7C>G (homo)	-
F3	Right iris coloboma	macrocephaly, learning difficulties.	-	-

References

Abouzeid, H., F. M. Meire, I. Osman, N. ElShakankiri, S. Bolay, F. L. Munier and D. F. Schorderet (2009). "A new locus for congenital cataract, microcornea, microphthalmia, and atypical iris coloboma maps to chromosome 2." Ophthalmology **116**(1): 154-162 e1.

Acosta-Sampson, L. and J. King (2010). "Partially folded aggregation intermediates of human gammaD-, gammaC-, and gammaS-crystallin are recognized and bound by human alphaB-crystallin chaperone." J Mol Biol **401**(1): 134-52.

Adato, A., S. Vreugde, T. Joensuu, N. Avidan, R. Hamalainen, O. Belenkiy, T. Olender, B. Bonne-Tamir, E. Ben-Asher, C. Espinos, J. M. Millan, A. E. Lehesjoki, J. G. Flannery, K. B. Avraham, S. Pietrokovski, E. M. Sankila, J. S. Beckmann and D. Lancet (2002). "USH3A transcripts encode clarin-1, a four-transmembrane-domain protein with a possible role in sensory synapses." Eur J Hum Genet **10**(6): 339-50.

Addison, P. K., V. Berry, K. R. Holden, D. Espinal, B. Rivera, H. Su, A. K. Srivastava and S. S. Bhattacharya (2006). "A novel mutation in the connexin 46 gene (GJA3) causes autosomal dominant zonular pulverulent cataract in a Hispanic family." Mol Vis **12**: 791-5.

Agre, P., L. S. King, M. Yasui, W. B. Guggino, O. P. Ottersen, Y. Fujiyoshi, A. Engel and S. Nielsen (2002). "Aquaporin water channels--from atomic structure to clinical medicine." J Physiol **542**(Pt 1): 3-16.

Aijaz, S., B. J. Clark, K. Williamson, V. van Heyningen, D. Morrison, D. Fitzpatrick, R. Collin, N. Ragge, A. Christoforou, A. Brown and I. Hanson (2004). "Absence of SIX6 mutations in microphthalmia, anophthalmia, and coloboma." Invest Ophthalmol Vis Sci **45**(11): 3871-6.

Al-Ghoul, K. J., T. Kirk, A. J. Kuszak, R. K. Zoltoski, A. Shiels and J. R. Kuszak (2003). "Lens structure in MIP-deficient mice." Anat Rec A Discov Mol Cell Evol Biol **273**(2): 714-30.

Alfalah, M., R. Jacob, U. Preuss, K. P. Zimmer, H. Naim and H. Y. Naim (1999). "O-linked glycans mediate apical sorting of human intestinal sucrase-isomaltase through association with lipid rafts." Curr Biol **9**(11): 593-6.

Allen, R. J., L. Speedwell and I. Russell-Eggitt (2009). "Long-term visual outcome after extraction of unilateral congenital cataracts." Eye (Lond).

Amaya, L., D. Taylor, I. Russell-Eggitt, K. K. Nischal and D. Lengyel (2003). "The morphology and natural history of childhood cataracts." Surv Ophthalmol **48**(2): 125-44.

Andley, U. P. (2009). "Effects of alpha-crystallin on lens cell function and cataract pathology." Curr Mol Med **9**(7): 887-92.

Andley, U. P., H. C. Patel and J. H. Xi (2002). "The R116C mutation in alpha A-crystallin diminishes its protective ability against stress-induced lens epithelial cell apoptosis." J Biol Chem **277**(12): 10178-86.

Andley, U. P., J. S. Rhim, L. T. Chylack, Jr. and T. P. Fleming (1994). "Propagation and immortalization of human lens epithelial cells in culture." Invest Ophthalmol Vis Sci **35**(7): 3094-102.

Andley, U. P., Z. Song, E. F. Wawrousek, J. P. Brady, S. Bassnett and T. P. Fleming (2001). "Lens epithelial cells derived from alphaB-crystallin knockout mice demonstrate hyperproliferation and genomic instability." Faseb J **15**(1): 221-229.

Angelow, S., R. Ahlstrom and A. S. Yu (2008). "Biology of claudins." Am J Physiol Renal Physiol **295**(4): F867-76.

Apweiler, R., H. Hermjakob and N. Sharon (1999). "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database." Biochim Biophys Acta **1473**(1): 4-8.

Arikkath, J., C. C. Chen, C. Ahern, V. Allamand, J. D. Flanagan, R. Coronado, R. G. Gregg and K. P. Campbell (2003). "Gamma 1 subunit interactions within the skeletal muscle L-type voltage-gated calcium channels." J Biol Chem **278**(2): 1212-9.

Ars, E., E. Serra, J. Garcia, H. Kruyer, A. Gaona, C. Lazaro and X. Estivill (2000). "Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1." Hum Mol Genet **9**(2): 237-47.

Arvan, P., X. Zhao, J. Ramos-Castaneda and A. Chang (2002). "Secretory pathway quality control operating in Golgi, plasmalemmal, and endosomal systems." Traffic **3**(11): 771-80.

Attardi, L. D., E. E. Reczek, C. Cosmas, E. G. Demicco, M. E. McCurrach, S. W. Lowe and T. Jacks (2000). "PERP, an apoptosis-associated target of p53, is a novel member of the PMP-22/gas3 family." Genes Dev **14**(6): 704-18.

Augusteyn, R. C., C. E. Jones and J. M. Pope (2008). "Age-related development of a refractive index plateau in the human lens: evidence for a distinct nucleus." Clin Exp Optom **91**(3): 296-301.

Azuma, N., Y. Yamaguchi, H. Handa, M. Hayakawa, A. Kanai and M. Yamada (1999). "Missense mutation in the alternative splice region of the PAX6 gene in eye anomalies." Am J Hum Genet **65**(3): 656-63.

Baechner, D., T. Liehr, H. Hameister, H. Altenberger, H. Grehl, U. Suter and B. Rautenstrauss (1995). "Widespread expression of the peripheral myelin protein-22 gene (PMP22) in neural and non-neural tissues during murine development." J Neurosci Res **42**(6): 733-41.

Bakthavachalu, B., S. Kalanke, S. Galande, B. Ramanamurthy, P. Parab, K. N. Kohale and V. Seshadri (2010). "Dense cataract and microphthalmia (dcm) in BALB/c mice is caused by mutations in the GJA8 locus." J Genet **89**(2): 147-54.

Banan, A., L. J. Zhang, M. Shaikh, J. Z. Fields, S. Choudhary, C. B. Forsyth, A. Farhadi and A. Keshavarzian (2005). "theta Isoform of protein kinase C alters barrier function in intestinal epithelium through modulation of distinct claudin isotypes: a novel mechanism for regulation of permeability." J Pharmacol Exp Ther **313**(3): 962-82.

Bangsow, T., E. Baumann, C. Bangsow, M. H. Jaeger, B. Pelzer, P. Gruhn, S. Wolf, H. von Melchner and D. B. Stanimirovic (2008). "The epithelial membrane protein 1 is a novel tight junction protein of the blood-brain barrier." J Cereb Blood Flow Metab **28**(6): 1249-60.

Bar-Yosef, U., I. Abuelaish, T. Harel, N. Hendler, R. Ofir and O. S. Birk (2004). "CHX10 mutations cause non-syndromic microphthalmia/ anophthalmia in Arab and Jewish kindreds." Hum Genet **115**(4): 302-9.

Barthold, J. S. and R. Gonzalez (2003). "The epidemiology of congenital cryptorchidism, testicular ascent and orchiopexy." J Urol **170**(6 Pt 1): 2396-401.

Bascom, R. A., K. Schappert and R. R. McInnes (1993). "Cloning of the human and murine ROM1 genes: genomic organization and sequence conservation." Hum Mol Genet **2**(4): 385-91.

Bassnett, S. (1995). "The fate of the Golgi apparatus and the endoplasmic reticulum during lens fiber cell differentiation." Invest Ophthalmol Vis Sci **36**(9): 1793-803.

Bassnett, S. (1997). "Fiber cell denucleation in the primate lens." Invest Ophthalmol Vis Sci **38**(9): 1678-87.

Bassnett, S. (2002). "Lens organelle degradation." Exp Eye Res **74**(1): 1-6.

Bassnett, S. (2005). "Three-dimensional reconstruction of cells in the living lens: the relationship between cell length and volume." Exp Eye Res **81**(6): 716-23.

Bassnett, S. (2009). "On the mechanism of organelle degradation in the vertebrate lens." Exp Eye Res **88**(2): 133-9.

Bassnett, S. and D. C. Beebe (1992). "Coincident loss of mitochondria and nuclei during lens fiber cell differentiation." Dev Dyn **194**(2): 85-93.

Bassnett, S., H. Missey and I. Vucemilo (1999). "Molecular architecture of the lens fiber cell basal membrane complex." J Cell Sci **112 (Pt 13)**: 2155-65.

Bassnett, S., P. A. Wilmarth and L. L. David (2009). "The membrane proteome of the mouse lens fiber cell." Mol Vis **15**: 2448-63.

Bassnett, S. and P. A. Winzenburger (2003). "Morphometric analysis of fibre cell growth in the developing chicken lens." Exp Eye Res **76**(3): 291-302.

Bause, E. and G. Legler (1981). "The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis." Biochem J **195**(3): 639-44.

Beaudry, V. G., R. A. Ihrie, S. B. Jacobs, B. Nguyen, N. Pathak, E. Park and L. D. Attardi (2010a). "Loss of the desmosomal component perp impairs wound healing in vivo." Dermatol Res Pract **2010**: 759731.

Beaudry, V. G., D. Jiang, R. L. Dusek, E. J. Park, S. Knezevich, K. Ridd, H. Vogel, B. C. Bastian and L. D. Attardi (2010b). "Loss of the p53/p63 Regulated Desmosomal Protein Perp Promotes Tumorigenesis." PLoS Genet **6**(10): e1001168.

Beaumont, C., P. Leneuve, I. Devaux, J. Y. Scoazec, M. Berthier, M. N. Loiseau, B. Grandchamp and D. Bonneau (1995). "Mutation in the iron responsive element of the L ferritin mRNA in a family with dominant hyperferritinaemia and cataract." Nat Genet **11**(4): 444-6.

Beebe, D. C. (2008). "Maintaining transparency: a review of the developmental physiology and pathophysiology of two avascular tissues." Semin Cell Dev Biol **19**(2): 125-33.

Bendtsen, J. D., H. Nielsen, G. von Heijne and S. Brunak (2004). "Improved prediction of signal peptides: SignalP 3.0." J Mol Biol **340**(4): 783-95.

Beyer, E. C., J. Kistler, D. L. Paul and D. A. Goodenough (1989). "Antisera directed against connexin43 peptides react with a 43-kD protein localized to gap junctions in myocardium and other tissues." J Cell Biol **108**(2): 595-605.

Bidinost, C., M. Matsumoto, D. Chung, N. Salem, K. Zhang, D. W. Stockton, A. Khoury, A. Megarbane, B. A. Bejjani and E. I. Traboulsi (2006). "Heterozygous and homozygous mutations in PITX3 in a large Lebanese family with posterior polar cataracts and neurodevelopmental abnormalities." Invest Ophthalmol Vis Sci **47**(4): 1274-80.

Billingsley, G., S. T. Santhiya, A. D. Paterson, K. Ogata, S. Wodak, S. M. Hosseini, S. M. Manisastry, P. Vijayalakshmi, P. M. Gopinath, J. Graw and E. Heon (2006). "CRYBA4, a novel human cataract gene, is also involved in microphthalmia." Am J Hum Genet **79**(4): 702-9.

Blankenship, T. N., J. F. Hess and P. G. FitzGerald (2001). "Development- and differentiation-dependent reorganization of intermediate filaments in fiber cells." Invest Ophthalmol Vis Sci **42**(3): 735-42.

Blasig, I. E., L. Winkler, B. Lassowski, S. L. Mueller, N. Zuleger, E. Krause, G. Krause, K. Gast, M. Kolbe and J. Piontek (2006). "On the self-association potential of transmembrane tight junction proteins." Cell Mol Life Sci **63**(4): 505-14.

Bloemendal, H. and G. Groenewoud (1981). "One-step separation of the subunits of alpha-crystallin by chromatofocusing in 6 M urea." Anal Biochem **117**(2): 327-9.

Blom, N., S. Gammeltoft and S. Brunak (1999). "Sequence and structure-based prediction of eukaryotic protein phosphorylation sites." J Mol Biol **294**(5): 1351-62.

Blom, N., T. Sicheritz-Ponten, R. Gupta, S. Gammeltoft and S. Brunak (2004). "Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence." Proteomics **4**(6): 1633-49.

Bodker, F. S., M. A. Lavery, T. N. Mitchell, E. W. Lovrien and I. H. Maumenee (1990). "Microphthalmos in the presumed homozygous offspring of a first cousin marriage and linkage analysis of a locus in a family with autosomal dominant cerulean congenital cataracts." Am J Med Genet **37**(1): 54-9.

Bolz, H., I. Ebermann and A. Gal (2005). "Protocadherin-21 (PCDH21), a candidate gene for human retinal dystrophies." Mol Vis **11**: 929-33.

Boothby, K. M. and A. Roberts (1995). "Effects of site and strength of tactile stimulation on the swimming responses of Xenopus laevis embryos." J. Zool. **235**: 113-125.

Bornheim, R., M. Muller, U. Reuter, H. Herrmann, H. Bussow and T. M. Magin (2008). "A dominant vimentin mutant upregulates Hsp70 and the activity of the ubiquitin-proteasome system, and causes posterior cataracts in transgenic mice." J Cell Sci **121**(Pt 22): 3737-46.

Boswell, B. A., P. A. Overbeek and L. S. Musil (2008). "Essential role of BMPs in FGF-induced secondary lens fiber differentiation." Dev Biol **324**(2): 202-12.

Botstein, D. and N. Risch (2003). "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease." Nat Genet **33 Suppl**: 228-37.

Brady, J. P., D. Garland, Y. Duglas-Tabor, W. G. Robison, Jr., A. Groome and E. F. Wawrousek (1997). "Targeted disruption of the mouse alpha A-crystallin gene induces cataract and cytoplasmic inclusion bodies containing the small heat shock protein alpha B-crystallin." Proc Natl Acad Sci U S A **94**(3): 884-9.

Brandli, A. W. (1999). "Towards a molecular anatomy of the Xenopus pronephric kidney." Int J Dev Biol **43**(5): 381-95.

Breuer, W., R. A. Klein, B. Hardt, A. Bartoschek and E. Bause (2001). "Oligosaccharyltransferase is highly specific for the hydroxy amino acid in Asn-Xaa-Thr/Ser." FEBS Lett **501**(2-3): 106-10.

Brice, N. L. and A. C. Dolphin (1999). "Differential plasma membrane targeting of voltage-dependent calcium channel subunits expressed in a polarized epithelial cell line." J Physiol **515 (Pt 3)**: 685-94.

Brizuela, B. J., O. Wessely and E. M. De Robertis (2001). "Overexpression of the Xenopus tight-junction protein claudin causes randomization of the left-right body axis." Dev Biol **230**(2): 217-29.

Brooks, S. P., M. Coccia, H. R. Tang, N. Kanuga, L. M. Machesky, M. Bailly, M. E. Cheetham and A. J. Hardcastle (2010). "The Nance-Horan syndrome protein encodes a functional WAVE homology domain (WHD) and is important for co-ordinating actin remodelling and maintaining cell morphology." Hum Mol Genet **19**(12): 2421-32.

Brown, N. and A. J. Bron (1996). Lens disorders: a clinical manual of cataract diagnosis. Oxford, Butterworth-Heinemann Ltd.

Bruggeman, L. A., S. Martinka and J. S. Simske (2007). "Expression of TM4SF10, a Claudin/EMP/PMP22 family cell junction protein, during mouse kidney development and podocyte differentiation." Dev Dyn **236**(2): 596-605.

Burgess, D. L., L. A. Gefrides, P. J. Foreman and J. L. Noebels (2001). "A cluster of three novel Ca2+ channel gamma subunit genes on chromosome 19q13.4: evolution and expression profile of the gamma subunit gene family." Genomics **71**(3): 339-50.

Burkitt Wright, E. M., R. Perveen, N. Bowers, S. Ramsden, E. McCann, M. O'Driscoll, I. C. Lloyd, J. Clayton-Smith and G. C. Black (2010). "VSX2 in microphthalmia: a novel splice site mutation producing a severe microphthalmia phenotype." Br J Ophthalmol **94**(3): 386-8.

Cain, S., G. Martinez, M. I. Kokkinos, K. Turner, R. J. Richardson, H. E. Abud, J. Huelsken, M. L. Robinson and R. U. de Iongh (2008). "Differential requirement for beta-catenin in epithelial and fiber cells during lens development." Dev Biol **321**(2): 420-33.

Cao, J., C. Shen, H. Wang, H. Shen, Y. Chen, A. Nie, G. Yan, H. Lu, Y. Liu and P. Yang (2009). "Identification of N-glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach." J Proteome Res **8**(2): 662-72.

Capetanaki, Y., S. Starnes and S. Smith (1989). "Expression of the chicken vimentin gene in transgenic mice: efficient assembly of the avian protein into the cytoskeleton." Proc Natl Acad Sci U S A **86**(13): 4882-6.

Carter, J. M., A. M. Hutcheson and R. A. Quinlan (1995). "In vitro studies on the assembly properties of the lens proteins CP49, CP115: coassembly with alpha-crystallin but not with vimentin." Exp Eye Res **60**(2): 181-92.

Chalmers, A. D. and J. M. Slack (1998). "Development of the gut in Xenopus laevis." Dev Dyn **212**(4): 509-21.

Chalmers, A. D. and J. M. Slack (2000). "The Xenopus tadpole gut: fate maps and morphogenetic movements." Development **127**(2): 381-92.

Chamberlain, C. G. and J. W. McAvoy (1989). "Induction of lens fibre differentiation by acidic and basic fibroblast growth factor (FGF)." Growth Factors **1**(2): 125-34.

Chan, T. and M. Asashima (2006). "Growing kidney in the frog." Nephron Exp Nephrol **103**(3): e81-5.

Chandy, G., G. A. Zampighi, M. Kreman and J. E. Hall (1997). "Comparison of the water transporting properties of MIP and AQP1." J Membr Biol **159**(1): 29-39.

Chang, B., N. L. Hawes, T. H. Roderick, R. S. Smith, J. R. Heckenlively, J. Horwitz and M. T. Davisson (1999). "Identification of a missense mutation in the alphaA-crystallin gene of the lop18 mouse." Mol Vis **5**: 21.

Chang, D. J., Y. S. Hwang, S. W. Cha, J. P. Chae, S. H. Hwang, J. H. Hahn, Y. C. Bae, H. S. Lee and M. J. Park (2010). "Xclaudin 1 is required for the proper gastrulation in Xenopus laevis." Biochem Biophys Res Commun **397**(1): 75-81.

Chao, R., L. Nevin, P. Agarwal, J. Riemer, X. Bai, A. Delaney, M. Akana, N. JimenezLopez, T. Bardakjian, A. Schneider, N. Chassaing, D. F. Schorderet, D. FitzPatrick, P. Y. Kwok, L. Ellgaard, D. B. Gould, Y. Zhang, J. Malicki, H. Baier and A. Slavotinek (2010). "A male with unilateral microphthalmia reveals a role for TMX3 in eye development." PLoS One **5**(5): e10565.

Chao, Y. C., S. H. Pan, S. C. Yang, S. L. Yu, T. F. Che, C. W. Lin, M. S. Tsai, G. C. Chang, C. H. Wu, Y. Y. Wu, Y. C. Lee, T. M. Hong and P. C. Yang (2009). "Claudin-1 is a metastasis suppressor and correlates with clinical outcome in lung adenocarcinoma." Am J Respir Crit Care Med **179**(2): 123-33.

Chauhan, B. K., A. Disanza, S. Y. Choi, S. C. Faber, M. Lou, H. E. Beggs, G. Scita, Y. Zheng and R. A. Lang (2009). "Cdc42- and IRSp53-dependent contractile filopodia tether presumptive lens and retina to coordinate epithelial invagination." Development **136**(21): 3657-67.

Chen, L., D. M. Chetkovich, R. S. Petralia, N. T. Sweeney, Y. Kawasaki, R. J. Wenthold,

D. S. Bredt and R. A. Nicoll (2000). "Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms." Nature **408**(6815): 936-43.

Chen, R. S., T. C. Deng, T. Garcia, Z. M. Sellers and P. M. Best (2007). "Calcium channel gamma subunits: a functionally diverse protein family." Cell Biochem Biophys **47**(2): 178-86.

Choi, M. Y., A. W. Partridge, C. Daniels, K. Du, G. L. Lukacs and C. M. Deber (2005). "Destabilization of the transmembrane domain induces misfolding in a phenotypic mutant of cystic fibrosis transmembrane conductance regulator." J Biol Chem **280**(6): 4968-74.

Chow, R. L. and R. A. Lang (2001). "Early eye development in vertebrates." Annu Rev Cell Dev Biol **17**: 255-96.

Christophe-Hobertus, C., C. Szpirer, R. Guyon and D. Christophe (2001). "Identification of the gene encoding Brain Cell Membrane Protein 1 (BCMP1), a putative four-transmembrane protein distantly related to the Peripheral Myelin Protein 22 / Epithelial Membrane Proteins and the Claudins." BMC Genomics **2**(1): 3.

Chu, P. J., H. M. Robertson and P. M. Best (2001). "Calcium channel gamma subunits provide insights into the evolution of this gene family." Gene **280**(1-2): 37-48.

Ciesiolka, M., M. Delvaeye, G. Van Imschoot, V. Verschuere, P. McCrea, F. van Roy and K. Vleminckx (2004). "p120 catenin is required for morphogenetic movements involved in the formation of the eyes and the craniofacial skeleton in Xenopus." J Cell Sci **117**(Pt 18): 4325-39.

Cline, M. S., R. Shigeta, R. L. Wheeler, M. A. Siani-Rose, D. Kulp and A. E. Loraine (2004). "The effects of alternative splicing on transmembrane proteins in the mouse genome." Pac Symp Biocomput: 17-28.

Coccia, M., S. P. Brooks, T. R. Webb, K. Christodoulou, I. O. Wozniak, V. Murday, M. Balicki, H. A. Yee, T. Wangensteen, R. Riise, A. K. Saggar, S. M. Park, N. Kanuga, P. J. Francis, E. R. Maher, A. T. Moore, I. M. Russell-Eggitt and A. J. Hardcastle (2009). "X-linked cataract and Nance-Horan syndrome are allelic disorders." Hum Mol Genet **18**(14): 2643-55.

Colegio, O. R., C. M. Van Itallie, H. J. McCrea, C. Rahner and J. M. Anderson (2002). "Claudins create charge-selective channels in the paracellular pathway between epithelial cells." Am J Physiol Cell Physiol **283**(1): C142-7.

Collins, J. S. and C. E. Schwartz (2002). "Detecting polymorphisms and mutations in candidate genes." Am J Hum Genet **71**(5): 1251-2.

Colucci-Guyon, E., M. M. Portier, I. Dunia, D. Paulin, S. Pournin and C. Babinet (1994). "Mice lacking vimentin develop and reproduce without an obvious phenotype." Cell **79**(4): 679-94.

Conley, Y. P., D. Erturk, A. Keverline, T. S. Mah, A. Keravala, L. R. Barnes, A. Bruchis, J. F. Hess, P. G. FitzGerald, D. E. Weeks, R. E. Ferrell and M. B. Gorin (2000). "A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense mutation in the beaded filament structural protein-2." Am J Hum Genet **66**(4): 1426-31.

Cooper, M. A., A. I. Son, D. Komlos, Y. Sun, N. J. Kleiman and R. Zhou (2008). "Loss of ephrin-A5 function disrupts lens fiber cell packing and leads to cataract." Proc Natl Acad Sci U S A **105**(43): 16620-5.

Coulombre, J. L. and A. J. Coulombre (1963). "Lens Development: Fiber Elongation and Lens Orientation." Science **142**: 1489-90.

Coyne, C. B., T. M. Gambling, R. C. Boucher, J. L. Carson and L. G. Johnson (2003). "Role of claudin interactions in airway tight junctional permeability." Am J Physiol Lung Cell Mol Physiol **285**(5): L1166-78.

Cuadra, A. E., S. H. Kuo, Y. Kawasaki, D. S. Bredt and D. M. Chetkovich (2004). "AMPA receptor synaptic targeting regulated by stargazin interactions with the Golgi-resident PDZ protein nPIST." J Neurosci **24**(34): 7491-502.

Cukierman, L., L. Meertens, C. Bertaux, F. Kajumo and T. Dragic (2009). "Residues in a highly conserved claudin-1 motif are required for hepatitis C virus entry and mediate the formation of cell-cell contacts." J Virol **83**(11): 5477-84.

Cullinane, A. R., A. Straatman-Iwanowska, A. Zaucker, Y. Wakabayashi, C. K. Bruce, G. Luo, F. Rahman, F. Gurakan, E. Utine, T. B. Ozkan, J. Denecke, J. Vukovic, M. Di Rocco, H. Mandel, H. Cangul, R. P. Matthews, S. G. Thomas, J. Z. Rappoport, I. M. Arias, H. Wolburg, A. S. Knisely, D. A. Kelly, F. Muller, E. R. Maher and P. Gissen (2010). "Mutations in VIPAR cause an arthrogryposis, renal dysfunction and

cholestasis syndrome phenotype with defects in epithelial polarization." Nat Genet **42**(4): 303-12.

Cvekl, A., Y. Yang, B. K. Chauhan and K. Cveklova (2004). "Regulation of gene expression by Pax6 in ocular cells: a case of tissue-preferred expression of crystallins in lens." Int J Dev Biol **48**(8-9): 829-44.

Dahabreh, J., G. P. Stathopoulos, J. Koutantos and S. Rigatos (2009). "Lung carcinoid tumor biology: treatment and survival." Oncol Rep **21**(3): 757-60.

Dahm, R., C. Gribbon, R. A. Quinlan and A. R. Prescott (1998). "Changes in the nucleolar and coiled body compartments precede lamina and chromatin reorganization during fibre cell denucleation in the bovine lens." Eur J Cell Biol **75**(3): 237-46.

Daly, S. B., J. E. Urquhart, E. Hilton, E. A. McKenzie, R. A. Kammerer, M. Lewis, B. Kerr, H. Stuart, D. Donnai, D. A. Long, B. Burgu, O. Aydogdu, M. Derbent, S. Garcia-Minaur, W. Reardon, B. Gener, S. Shalev, R. Smith, A. S. Woolf, G. C. Black and W. G. Newman (2010). "Mutations in HPSE2 Cause Urofacial Syndrome." Am J Hum Genet.

Daugherty, B. L., C. Ward, T. Smith, J. D. Ritzenthaler and M. Koval (2007). "Regulation of heterotypic claudin compatibility." J Biol Chem **282**(41): 30005-13.

Davidson, A. E., I. D. Millar, J. E. Urquhart, R. Burgess-Mullan, Y. Shweikh, N. Parry, J. O'Sullivan, G. J. Maher, M. McKibbin, S. M. Downes, A. J. Lotery, S. G. Jacobson, P. D. Brown, G. C. Black and F. D. Manson (2009). "Missense mutations in a retinal pigment epithelium protein, bestrophin-1, cause retinitis pigmentosa." Am J Hum Genet **85**(5): 581-92.

Davies, L., D. Gray, D. Spiller, M. R. White, B. Damato, I. Grierson and L. Paraoan (2009). "P53 apoptosis mediator PERP: localization, function and caspase activation in uveal melanoma." J Cell Mol Med **13**(8B): 1995-2007.

Dejgaard, S., J. Nicolay, M. Taheri, D. Y. Thomas and J. J. Bergeron (2004). "The ER glycoprotein quality control system." Curr Issues Mol Biol **6**(1): 29-42.

Desilva, M. G., A. L. Notkins and M. S. Lan (1997). "Molecular characterization of a pancreas-specific protein disulfide isomerase, PDIp." DNA Cell Biol **16**(3): 269-74.

Ding, X., M. Patel, A. A. Herzlich, P. C. Sieving and C. C. Chan (2009). "Ophthalmic pathology of Nance-Horan syndrome: case report and review of the literature." Ophthalmic Genet **30**(3): 127-35.

Dirscherl, S. S., J. J. Henry and J. E. Krebs (2005). "Neural and eye-specific defects associated with loss of the imitation switch (ISWI) chromatin remodeler in Xenopus laevis." Mech Dev **122**(11): 1157-70.

Doolittle, R. F. (1987). Of URFs and ORFs: a primer on how to analyze derived amino acid sequences. Mill Valley California, University Science Books.

Duncan, M. K., L. Xie, L. L. David, M. L. Robinson, J. R. Taube, W. Cui and L. W. Reneker (2004). "Ectopic Pax6 expression disturbs lens fiber cell differentiation." Invest Ophthalmol Vis Sci **45**(10): 3589-98.

Dyson, N. (1998). "The regulation of E2F by pRB-family proteins." Genes Dev **12**(15): 2245-62.

Elkins, M. B. and J. J. Henry (2006). "Isolation and characterization of a novel gene, xMADML, involved in Xenopus laevis eye development." Dev Dyn **235**(7): 1845-57.

Ellgaard, L. and A. Helenius (2003). "Quality control in the endoplasmic reticulum." Nat Rev Mol Cell Biol **4**(3): 181-91.

Estabel, J., A. Mercer, N. Konig and J. M. Exbrayat (2003). "Programmed cell death in Xenopus laevis spinal cord, tail and other tissues, prior to, and during, metamorphosis." Life Sci **73**(25): 3297-306.

Faber, S. C., M. L. Robinson, H. P. Makarenkova and R. A. Lang (2002). "Bmp signaling is required for development of primary lens fiber cells." Development **129**(15): 3727-37.

Fagerholm, P. P., B. T. Philipson and B. Lindstrom (1981). "Normal human lens - the distribution of protein." Exp Eye Res **33**(6): 615-20.

Felsenstein, J. (1985). "Confidence limits on phylogenies: an approach using the bootstrap." Evolution **39**: 783-791.

Fernald, R. D. and S. E. Wright (1983). "Maintenance of optical quality during crystalline lens growth." Nature **301**(5901): 618-20.

Ferreira-Cornwell, M. C., R. W. Veneziale, G. B. Grunwald and A. S. Menko (2000). "N-cadherin function is required for differentiation-dependent cytoskeletal reorganization in lens cells in vitro." Exp Cell Res **256**(1): 237-47.

Ferron, L., A. Davies, K. M. Page, D. J. Cox, J. Leroy, D. Waithe, A. J. Butcher, P. Sellaturay, S. Bolsover, W. S. Pratt, F. J. Moss and A. C. Dolphin (2008). "The

stargazin-related protein gamma 7 interacts with the mRNA-binding protein heterogeneous nuclear ribonucleoprotein A2 and regulates the stability of specific mRNAs, including CaV2.2." J Neurosci **28**(42): 10604-17.

Francis, P., J. J. Chung, M. Yasui, V. Berry, A. Moore, M. K. Wyatt, G. Wistow, S. S. Bhattacharya and P. Agre (2000). "Functional impairment of lens aquaporin in two families with dominantly inherited cataracts." Hum Mol Genet **9**(15): 2329-34.

Francis, P. J., V. Berry, A. T. Moore and S. Bhattacharya (1999). "Lens biology: development and human cataractogenesis." Trends Genet **15**(5): 191-6.

Francis, P. J. and A. T. Moore (2004). "Genetics of childhood cataract." Curr Opin Ophthalmol **15**(1): 10-5.

Fraunfelder, F. T., J. M. LaBraico and S. M. Meyer (1985). "Adverse ocular reactions possibly associated with isotretinoin." Am J Ophthalmol **100**(4): 534-7.

Friedlander, R., E. Jarosch, J. Urban, C. Volkwein and T. Sommer (2000). "A regulatory link between ER-associated protein degradation and the unfolded-protein response." Nat Cell Biol **2**(7): 379-84.

Fromm, L., W. Shawlot, K. Gunning, J. S. Butel and P. A. Overbeek (1994). "The retinoblastoma protein-binding region of simian virus 40 large T antigen alters cell cycle regulation in lenses of transgenic mice." Mol Cell Biol **14**(10): 6743-54.

Fu, L. and J. J. Liang (2003). "Alteration of protein-protein interactions of congenital cataract crystallin mutants." Invest Ophthalmol Vis Sci **44**(3): 1155-9.

Fuchs, S., H. Kranich, M. J. Denton, E. Zrenner, S. S. Bhattacharya, P. Humphries and A. Gal (1994). "Three novel rhodopsin mutations (C110F, L131P, A164V) in patients with autosomal dominant retinitis pigmentosa." Hum Mol Genet **3**(7): 1203.

Fujibe, M., H. Chiba, T. Kojima, T. Soma, T. Wada, T. Yamashita and N. Sawada (2004). "Thr203 of claudin-1, a putative phosphorylation site for MAP kinase, is required to promote the barrier function of tight junctions." Exp Cell Res **295**(1): 36-47.

Fujiki, K., A. Nakajima, N. Yasuda, N. Tanabe and K. Kabasawa (1982). "Genetic analysis of microphthalmos." Ophthal Paediat Genet(1): 139-149.

Fujita, K., J. Katahira, Y. Horiguchi, N. Sonoda, M. Furuse and S. Tsukita (2000). "Clostridium perfringens enterotoxin binds to the second extracellular loop of claudin-3, a tight junction integral membrane protein." FEBS Letters **476**(3): 258-261.

Fujiwara, M., T. Uchida, N. Osumi-Yamashita and K. Eto (1994). "Uchida rat (rSey): a new mutant rat with craniofacial abnormalities resembling those of the mouse Sey mutant." Differentiation **57**(1): 31-8.

Fukaya, M., M. Yamazaki, K. Sakimura and M. Watanabe (2005). "Spatial diversity in gene expression for VDCCgamma subunit family in developing and adult mouse brains." Neurosci Res **53**(4): 376-83.

Furuse, M., K. Furuse, H. Sasaki and S. Tsukita (2001). "Conversion of zonulae occludentes from tight to leaky strand type by introducing claudin-2 into Madin-Darby canine kidney I cells." J Cell Biol **153**(2): 263-72.

Furuse, M., M. Hata, K. Furuse, Y. Yoshida, A. Haratake, Y. Sugitani, T. Noda, A. Kubo and S. Tsukita (2002). Claudin-based tight junctions are crucial for the mammalian epidermal barrier: a lesson from claudin-1-deficient mice. **156:** 1099-1111.

Gahmberg, C. G. and M. Tolvanen (1996). "Why mammalian cell surface proteins are glycoproteins." Trends Biochem Sci **21**(8): 308-11.

Gallagher, M. J., L. Ding, A. Maheshwari and R. L. Macdonald (2007). "The GABAA receptor alpha1 subunit epilepsy mutation A322D inhibits transmembrane helix formation and causes proteasomal degradation." Proc Natl Acad Sci U S A **104**(32): 12999-3004.

Gallardo, M. E., S. Rodriguez De Cordoba, A. S. Schneider, M. A. Dwyer, C. Ayuso and P. Bovolenta (2004). "Analysis of the developmental SIX6 homeobox gene in patients with anophthalmia/microphthalmia." Am J Med Genet A **129A**(1): 92-4.

Gao, C. Y., Z. Zakeri, Y. Zhu, H. He and P. S. Zelenka (1997). "Expression of Cdk5, p35, and Cdk5-associated kinase activity in the developing rat lens." Dev Genet **20**(3): 267-75.

Gao, X., J. Cheng, C. Lu, X. Li, F. Li, C. Liu, M. Zhang, S. Zhu and X. Ma (2010). "A novel mutation in the connexin 50 gene (GJA8) associated with autosomal dominant congenital nuclear cataract in a Chinese family." Curr Eye Res **35**(7): 597-604.

Garcia, C. M., K. Yu, H. Zhao, R. Ashery-Padan, D. M. Ornitz, M. L. Robinson and D. C. Beebe (2005). "Signaling through FGF receptor-2 is required for lens cell survival and for withdrawal from the cell cycle during lens fiber cell differentiation." Dev Dyn **233**(2): 516-27.

Gavel, Y. and G. von Heijne (1990). "Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering." Protein Eng **3**(5): 433-42.

Gelman, M. S., E. S. Kannegaard and R. R. Kopito (2002). "A principal role for the proteasome in endoplasmic reticulum-associated degradation of misfolded intracellular cystic fibrosis transmembrane conductance regulator." J Biol Chem **277**(14): 11709-14.

Geyer, D. D., M. A. Spence, M. Johannes, P. Flodman, K. P. Clancy, R. Berry, R. S. Sparkes, M. D. Jonsen, S. J. Isenberg and J. B. Bateman (2006). "Novel single-base deletional mutation in major intrinsic protein (MIP) in autosomal dominant cataract." Am J Ophthalmol **141**(4): 761-3.

Giepmans, B. N. and S. C. van Ijzendoorn (2009). "Epithelial cell-cell junctions and plasma membrane domains." Biochim Biophys Acta **1788**(4): 820-31.

Girao, H., P. Pereira, A. Taylor and F. Shang (2005). "Subcellular redistribution of components of the ubiquitin-proteasome pathway during lens differentiation and maturation." Invest Ophthalmol Vis Sci **46**(4): 1386-92.

Giuffrida, M. G., M. Cavaletto, C. Giunta, B. Neuteboom, A. Cantisani, L. Napolitano, V. Calderone, J. Godovac-Zimmermann and A. Conti (1997). "The unusual amino acid triplet Asn-Ile-Cys is a glycosylation consensus site in human alpha-lactalbumin." J Protein Chem **16**(8): 747-53.

Glaser, T., L. Jepeal, J. G. Edwards, S. R. Young, J. Favor and R. L. Maas (1994). "PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects." Nat Genet **7**(4): 463-71.

Goldberg, A. F., L. M. Fales, J. B. Hurley and N. Khattree (2001). "Folding and subunit assembly of photoreceptor peripherin/rds is mediated by determinants within the extracellular/intradiskal EC2 domain: implications for heterogeneous molecular pathologies." J Biol Chem **276**(46): 42700-6.

Goldberg, A. F., O. L. Moritz and R. S. Molday (1995). "Heterologous expression of photoreceptor peripherin/rds and Rom-1 in COS-1 cells: assembly, interactions, and localization of multisubunit complexes." Biochemistry **34**(43): 14213-9.

Gonen, T., A. C. Grey, M. D. Jacobs, P. J. Donaldson and J. Kistler (2001). "MP20, the second most abundant lens membrane protein and member of the tetraspanin superfamily, joins the list of ligands of galectin-3." BMC Cell Biol **2**: 17.

Gong, X., C. Cheng and C. H. Xia (2007). "Connexins in lens development and cataractogenesis." J Membr Biol **218**(1-3): 9-12.

Gong, X., E. Li, G. Klier, Q. Huang, Y. Wu, H. Lei, N. M. Kumar, J. Horwitz and N. B. Gilula (1997). "Disruption of alpha3 connexin gene leads to proteolysis and cataractogenesis in mice." Cell **91**(6): 833-43.

Goodenough, D. A., J. S. Dick, 2nd and J. E. Lyons (1980). "Lens metabolic cooperation: a study of mouse lens transport and permeability visualized with freeze-substitution autoradiography and electron microscopy." J Cell Biol **86**(2): 576-89.

Gordon, R. A. and P. B. Donzis (1985). "Refractive development of the human eye." Arch Ophthalmol **103**(6): 785-9.

Gouy, M., S. Guindon and O. Gascuel (2010). "SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building." Mol Biol Evol **27**(2): 221-4.

Gow, A., C. M. Southwood, J. S. Li, M. Pariali, G. P. Riordan, S. E. Brodie, J. Danias, J. M. Bronstein, B. Kachar and R. A. Lazzarini (1999). "CNS myelin and sertoli cell tight junction strands are absent in Osp/claudin-11 null mice." Cell **99**(6): 649-59.

Graindorge, A., R. Thuret, N. Pollet, H. B. Osborne and Y. Audic (2006). "Identification of post-transcriptionally regulated Xenopus tropicalis maternal mRNAs by microarray." Nucleic Acids Res **34**(3): 986-95.

Graw, J. (2003). "The genetic and molecular basis of congenital eye defects." Nat Rev Genet **4**(11): 876-88.

Graw, J. (2004). "Congenital hereditary cataracts." Int J Dev Biol **48**(8-9): 1031-44.

Graw, J. (2009). "Genetics of crystallins: cataract and beyond." Exp Eye Res **88**(2): 173-89.

Greiner, J. V., D. B. Auerbach, C. D. Leahy and T. Glonek (1994). "Distribution of membrane phospholipids in the crystalline lens." Invest Ophthalmol Vis Sci **35**(10): 3739-46.

Grey, A. C., M. D. Jacobs, T. Gonen, J. Kistler and P. J. Donaldson (2003). "Insertion of MP20 into lens fibre cell plasma membranes correlates with the formation of an extracellular diffusion barrier." Experimental Eye Research **77**(5): 567-574.

Griep, A. E. (2006). "Cell cycle regulation in the developing lens." Semin Cell Dev Biol **17**(6): 686-97.

Grindley, J. C., D. R. Davidson and R. E. Hill (1995). "The role of Pax-6 in eye and nasal development." Development **121**(5): 1433-42.

Grove, M., G. Demyanenko, A. Echarri, P. A. Zipfel, M. E. Quiroz, R. M. Rodriguiz, M. Playford, S. A. Martensen, M. R. Robinson, W. C. Wetsel, P. F. Maness and A. M. Pendergast (2004). "ABI2-deficient mice exhibit defective cell migration, aberrant dendritic spine morphogenesis, and deficits in learning and memory." Mol Cell Biol **24**(24): 10905-22.

Gu, F., H. Zhai, D. Li, L. Zhao, C. Li, S. Huang and X. Ma (2007). "A novel mutation in major intrinsic protein of the lens gene (MIP) underlies autosomal dominant cataract in a Chinese family." Mol Vis **13**: 1651-6.

Gupta, P. D., K. Johar and A. Vasavada (2004a). "Causative and preventive action of calcium in cataracto-genesis." Acta Pharmacol Sin **25**(10): 1250-6.

Gupta, R., E. Jung and S. Brunak (2004b). "Prediction of N-glycosylation sites in human proteins." In preparation.

Gustincich, S., S. Batalov, K. W. Beisel, H. Bono, P. Carninci, C. F. Fletcher, S. Grimmond, N. Hirokawa, E. D. Jarvis, T. Jegla, Y. Kawasawa, J. LeMieux, H. Miki, E. Raviola, R. D. Teasdale, N. Tominaga, K. Yagi, A. Zimmer, Y. Hayashizaki and Y. Okazaki (2003). "Analysis of the mouse transcriptome for genes involved in the function of the nervous system." Genome Res **13**(6B): 1395-401.

Hage, R., A. B. de la Riviere, C. A. Seldenrijk and J. M. van den Bosch (2003). "Update in pulmonary carcinoid tumors: a review article." Ann Surg Oncol **10**(6): 697-704.

Halder, G., P. Callaerts and W. J. Gehring (1995). "Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila." Science **267**(5205): 1788-92.

Hammond, C. J., D. D. Duncan, H. Snieder, M. de Lange, S. K. West, T. D. Spector and C. E. Gilbert (2001). "The heritability of age-related cortical cataract: the twin eye study." Invest Ophthalmol Vis Sci **42**(3): 601-5.

Hampson, G., M. A. Konrad and J. Scoble (2008). "Familial hypomagnesaemia with hypercalciuria and nephrocalcinosis (FHHNC): compound heterozygous mutation in the claudin 16 (CLDN16) gene." BMC Nephrol **9**: 12.

Hanel, M. L. and C. Hensey (2006). "Eye and neural defects associated with loss of GDF6." BMC Dev Biol **6**: 43.

Hansen, J. P., R. S. Chen, J. K. Larsen, P. J. Chu, D. M. Janes, K. E. Weis and P. M. Best (2004). "Calcium channel gamma6 subunits are unique modulators of low voltage-activated (Cav3.1) calcium current." J Mol Cell Cardiol **37**(6): 1147-58.

Hanson, I., A. Churchill, J. Love, R. Axton, T. Moore, M. Clarke, F. Meire and V. van Heyningen (1999). "Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations." Hum Mol Genet **8**(2): 165-72.

Harlow, E. and D. Lane (1988). Antibodies: A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.

Hashimoto, K., M. Fukaya, X. Qiao, K. Sakimura, M. Watanabe and M. Kano (1999). "Impairment of AMPA receptor function in cerebellar granule cells of ataxic mutant mouse stargazer." J Neurosci **19**(14): 6027-36.

Hay, E. D. (1979). "Development of the vertebrate cornea." Int Rev Cytol **63**: 263-322.

Heasman, J., M. Kofron and C. Wylie (2000). "Beta-catenin signaling activity dissected in the early Xenopus embryo: a novel antisense approach." Dev Biol **222**(1): 124-34.

Hegde, R. S. and H. D. Bernstein (2006). "The surprising complexity of signal sequences." Trends Biochem Sci **31**(10): 563-71.

Heiskala, M., P. A. Peterson and Y. Yang (2001). "The roles of claudin superfamily proteins in paracellular transport." Traffic **2**(2): 93-8.

Hejtmancik, J. F. (2008). "Congenital cataracts and their molecular genetics." Semin Cell Dev Biol **19**(2): 134-49.

Helenius, A. (1994). "How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum." Mol Biol Cell **5**(3): 253-65.

Helenius, A. and M. Aebi (2004). "Roles of N-linked glycans in the endoplasmic reticulum." Annu Rev Biochem **73**: 1019-49.

Henderson, R. H., Z. Li, M. M. Abd El Aziz, D. S. Mackay, M. A. Eljinini, M. Zeidan, A. T. Moore, S. S. Bhattacharya and A. R. Webster (2010). "Biallelic mutation of protocadherin-21 (PCDH21) causes retinal degeneration in humans." Mol Vis **16**: 46-52.

Henry, J. J., J. M. Wever, M. N. Vergara and L. Fukui (2008). Xenopus, an Ideal Vertebrate System for Studies of Eye Development and Regeneration. Animal Models in Eye Research P. A. Tsonis, Academic Press**:** 57-72.

Hewitt, K. J., R. Agarwal and P. J. Morin (2006). "The claudin gene family: expression in normal and neoplastic tissues." BMC Cancer **6**: 186.

Hildebrandt, T., J. Preiherr, N. Tarbe, S. Klostermann, G. N. Van Muijen and U. H. Weidle (2000). "Identification of THW, a putative new tumor suppressor gene." Anticancer Res **20**(5A): 2801-9.

Hilton, E. N., F. D. Manson, J. E. Urquhart, J. J. Johnston, A. M. Slavotinek, P. Hedera, E. L. Stattin, A. Nordgren, L. G. Biesecker and G. C. Black (2007). "Left-sided embryonic expression of the BCL-6 corepressor, BCOR, is required for vertebrate laterality determination." Hum Mol Genet **16**(14): 1773-82.

Ho, H. Y., K. H. Chang, J. Nichols and M. Li (2009). "Homeodomain protein Pitx3 maintains the mitotic activity of lens epithelial cells." Mech Dev **126**(1-2): 18-29.

Hofmann, K. and S. W. (1993). "Tmbase - A database of membrane spanning proteins segments." Biol. Chem. Hoppe-Seyler **347** 166.

Hogan, B. L., G. Horsburgh, J. Cohen, C. M. Hetherington, G. Fisher and M. F. Lyon (1986). "Small eyes (Sey): a homozygous lethal mutation on chromosome 2 which affects the differentiation of both lens and nasal placodes in the mouse." J Embryol Exp Morphol **97**: 95-110.

Holmes, J. M., D. A. Leske, J. P. Burke and D. O. Hodge (2003). "Birth prevalence of visually significant infantile cataract in a defined U.S. population." Ophthalmic Epidemiol **10**(2): 67-74.

Horne-Badovinac, S. and D. Bilder (2008). "Dynein regulates epithelial polarity and the apical localization of stardust A mRNA." PLoS Genet **4**(1): e8.

Horwitz, J. (1992). "Alpha-crystallin can function as a molecular chaperone." Proc Natl Acad Sci U S A **89**(21): 10449-53.

Hosokawa, N., I. Wada, K. Hasegawa, T. Yorihuzi, L. O. Tremblay, A. Herscovics and K. Nagata (2001). "A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation." EMBO Rep **2**(5): 415-22.

Hou, J., A. Renigunta, M. Konrad, A. S. Gomes, E. E. Schneeberger, D. L. Paul, S. Waldegger and D. A. Goodenough (2008). "Claudin-16 and claudin-19 interact and form a cation-selective tight junction complex." J Clin Invest **118**(2): 619-28.

Hung, W., C. Hwang, M. D. Po and M. Zhen (2007). "Neuronal polarity is regulated by a direct interaction between a scaffolding protein, Neurabin, and a presynaptic SAD-1 kinase in Caenorhabditis elegans." Development **134**(2): 237-49.

Hwa, J., P. Garriga, X. Liu and H. G. Khorana (1997). "Structure and function in rhodopsin: packing of the helices in the transmembrane domain and folding to a tertiary structure in the intradiscal domain are coupled." Proc Natl Acad Sci U S A **94**(20): 10571-6.

Ihrie, R. A., M. R. Marques, B. T. Nguyen, J. S. Horner, C. Papazoglu, R. T. Bronson, A. A. Mills and L. D. Attardi (2005). "Perp is a p63-regulated gene essential for epithelial integrity." Cell **120**(6): 843-56.

Ionides, A., P. Francis, V. Berry, D. Mackay, S. Bhattacharya, A. Shiels and A. Moore (1999). "Clinical and genetic heterogeneity in autosomal dominant cataract." Br J Ophthalmol **83**(7): 802-8.

Iseri, S. U., A. W. Wyatt, G. Nurnberg, C. Kluck, P. Nurnberg, G. E. Holder, E. Blair, A. Salt and N. K. Ragge (2010). "Use of genome-wide SNP homozygosity mapping in small pedigrees to identify new mutations in VSX2 causing recessive microphthalmia and a semidominant inner retinal dystrophy." Hum Genet **128**(1): 51-60.

Ishibashi, S. and K. Yasuda (2001). "Distinct roles of maf genes during Xenopus lens development." Mech Dev **101**(1-2): 155-66.

Itoh, M., M. Furuse, K. Morita, K. Kubota, M. Saitou and S. Tsukita (1999). Direct Binding of Three Tight Junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH Termini of Claudins. **147:** 1351-1363.

Ivanov, D., G. Dvoriantchikova, A. Pestova, L. Nathanson and V. I. Shestopalov (2005). "Microarray analysis of fiber cell maturation in the lens." FEBS Lett **579**(5): 1213-9.

Iyengar, L., B. Patkunanathan, O. T. Lynch, J. W. McAvoy, J. E. Rasko and F. J. Lovicu (2006). "Aqueous humour- and growth factor-induced lens cell proliferation is dependent on MAPK/ERK1/2 and Akt/PI3-K signalling." Exp Eye Res **83**(3): 667-78.

Iyengar, L., Q. Wang, J. E. Rasko, J. W. McAvoy and F. J. Lovicu (2007). "Duration of ERK1/2 phosphorylation induced by FGF or ocular media determines lens cell fate." Differentiation **75**(7): 662-8.

Jamieson, R. V., N. Farrar, K. Stewart, R. Perveen, M. Mihelec, M. Carette, J. R. Grigg, J. W. McAvoy, F. J. Lovicu, P. P. Tam, P. Scambler, I. C. Lloyd, D. Donnai and G. C. Black (2007). "Characterization of a familial t(16;22) balanced translocation associated with congenital cataract leads to identification of a novel gene, TMEM114, expressed in the lens and disrupted by the translocation." Hum Mutat **28**(10): 968-77.

Janda, C. Y., J. Li, C. Oubridge, H. Hernandez, C. V. Robinson and K. Nagai (2010). "Recognition of a signal peptide by the signal recognition particle." Nature **465**(7297): 507-10.

Jeansonne, B., Q. Lu, D. A. Goodenough and Y. H. Chen (2003). "Claudin-8 interacts with multi-PDZ domain protein 1 (MUPP1) and reduces paracellular conductance in epithelial cells." Cell Mol Biol (Noisy-le-grand) **49**(1): 13-21.

Jen, J., R. H. Baloh, A. Ishiyama and R. W. Baloh (2005). "Dejerine-Sottas syndrome and vestibular loss due to a point mutation in the PMP22 gene." J Neurol Sci **237**(1-2): 21-4.

Jetten, A. M. and U. Suter (2000). "The peripheral myelin protein 22 and epithelial membrane protein family." Prog Nucleic Acid Res Mol Biol **64**: 97-129.

Jia, S., Z. Ren, X. Li, Y. Zheng and A. Meng (2008). "smad2 and smad3 are required for mesendoderm induction by transforming growth factor-beta/nodal signals in zebrafish." J Biol Chem **283**(4): 2418-26.

Jiang, Y., Z. Cheng, E. C. Mandon and R. Gilmore (2008). "An interaction between the SRP receptor and the translocon is critical during cotranslational protein translocation." J Cell Biol **180**(6): 1149-61.

Johnston, J. J., J. K. Teer, P. F. Cherukuri, N. F. Hansen, S. K. Loftus, K. Chong, J. C. Mullikin and L. G. Biesecker (2010). "Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate." Am J Hum Genet **86**(5): 743-8.

Jones, C. E., D. A. Atchison and J. M. Pope (2007). "Changes in lens dimensions and refractive index with age and accommodation." Optom Vis Sci **84**(10): 990-5.

Jones, C. E. and J. M. Pope (2004). "Measuring optical properties of an eye lens using magnetic resonance imaging." Magn Reson Imaging **22**(2): 211-20.

Julenius, K., A. Mølgaard, R. Gupta and S. Brunak (2005). "Prediction, conservation analysis and structural characterization of mammalian mucin-type O-glycosylation sites." Glycobiology **15**: 153-164.

Jun, G., H. Guo, B. E. Klein, R. Klein, J. J. Wang, P. Mitchell, H. Miao, K. E. Lee, T. Joshi, M. Buck, P. Chugha, D. Bardenstein, A. P. Klein, J. E. Bailey-Wilson, X. Gong, T. D. Spector, T. Andrew, C. J. Hammond, R. C. Elston, S. K. Iyengar and B. Wang (2009). "EPHA2 is associated with age-related cortical cataract in mice and humans." PLoS Genet **5**(7): e1000584.

Jung, H., K. H. Jun, J. H. Jung, H. M. Chin and W. B. Park (2010). "The Expression of Claudin-1, Claudin-2, Claudin-3, and Claudin-4 in Gastric Cancer Tissue." J Surg Res.

Kahle, K. T., G. G. Macgregor, F. H. Wilson, A. N. Van Hoek, D. Brown, T. Ardito, M. Kashgarian, G. Giebisch, S. C. Hebert, E. L. Boulpaep and R. P. Lifton (2004). "Paracellular Cl- permeability is regulated by WNK4 kinase: insight into normal physiology and hypertension." Proc Natl Acad Sci U S A **101**(41): 14877-82.

Kahsay, R. Y., G. Gao and L. Liao (2005). "An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes." Bioinformatics **21**(9): 1853-8.

Kajiwara, K., E. L. Berson and T. P. Dryja (1994). "Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci." Science **264**(5165): 1604-8.

Kall, L., A. Krogh and E. L. Sonnhammer (2007). "Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server." Nucleic Acids Res **35**(Web Server issue): W429-32.

Kallen, B. and K. Tornqvist (2005). "The epidemiology of anophthalmia and microphthalmia in Sweden." Eur J Epidemiol **20**(4): 345-50.

Kaplan, H. A., J. K. Welply and W. J. Lennarz (1987). "Oligosaccharyl transferase: the central enzyme in the pathway of glycoprotein assembly." Biochim Biophys Acta **906**(2): 161-73.

Kasthurirangan, S., E. L. Markwell, D. A. Atchison and J. M. Pope (2008). "In vivo study of changes in refractive index distribution in the human crystalline lens with age and accommodation." Invest Ophthalmol Vis Sci **49**(6): 2531-40.

Katahira, J., N. Inoue, Y. Horiguchi, M. Matsuda and N. Sugimoto (1997). "Molecular cloning and functional characterization of the receptor for Clostridium perfringens enterotoxin." J Cell Biol **136**(6): 1239-47.

Kato, A. S., E. R. Siuda, E. S. Nisenbaum and D. S. Bredt (2008). "AMPA receptor subunit-specific regulation by a distinct family of type II TARPs." Neuron **59**(6): 986-96.

Kato, A. S., W. Zhou, A. D. Milstein, M. D. Knierman, E. R. Siuda, J. E. Dotzlaf, H. Yu, J. E. Hale, E. S. Nisenbaum, R. A. Nicoll and D. S. Bredt (2007). "New transmembrane AMPA receptor regulatory protein isoform, gamma-7, differentially regulates AMPA receptors." J Neurosci **27**(18): 4969-77.

Kaul, H., S. A. Riazuddin, M. Shahid, S. Kousar, N. H. Butt, A. U. Zafar, S. N. Khan, T. Husnain, J. Akram, J. F. Hejtmancik and S. Riazuddin (2010). "Autosomal recessive congenital cataract linked to EPHA2 in a consanguineous Pakistani family." Mol Vis **16**: 511-7.

Kausalya, P. J., S. Amasheh, D. Gunzel, H. Wurps, D. Muller, M. Fromm and W. Hunziker (2006). "Disease-associated mutations affect intracellular traffic and paracellular Mg2+ transport function of Claudin-16." J Clin Invest **116**(4): 878-91.

Kawai, J., A. Shinagawa, K. Shibata, M. Yoshino, M. Itoh, Y. Ishii, T. Arakawa, A. Hara, Y. Fukunishi, H. Konno, J. Adachi, S. Fukuda, K. Aizawa, M. Izawa, K. Nishi, H. Kiyosawa, S. Kondo, I. Yamanaka, T. Saito, Y. Okazaki, T. Gojobori, H. Bono, T. Kasukawa, R. Saito, K. Kadota, H. Matsuda, M. Ashburner, S. Batalov, T. Casavant, W. Fleischmann, T. Gaasterland, C. Gissi, B. King, H. Kochiwa, P. Kuehl, S. Lewis, Y. Matsuo, I. Nikaido, G. Pesole, J. Quackenbush, L. M. Schriml, F. Staubli, R. Suzuki, M. Tomita, L. Wagner, T. Washio, K. Sakai, T. Okido, M. Furuno, H. Aono, R. Baldarelli, G. Barsh, J. Blake, D. Boffelli, N. Bojunga, P. Carninci, M. F. de Bonaldo, M. J. Brownstein, C. Bult, C. Fletcher, M. Fujita, M. Gariboldi, S. Gustincich, D. Hill, M. Hofmann, D. A. Hume, M. Kamiya, N. H. Lee, P. Lyons, L. Marchionni, J. Mashima, J. Mazzarelli, P. Mombaerts, P. Nordone, B. Ring, M. Ringwald, I. Rodriguez, N. Sakamoto, H. Sasaki, K. Sato, C. Schonbach, T. Seya, Y. Shibata, K. F. Storch, H. Suzuki, K. Toyo-oka, K. H. Wang, C. Weitz, C. Whittaker, L. Wilming, A. Wynshaw-Boris, K. Yoshida, Y. Hasegawa, H. Kawaji, S. Kohtsuki and Y. Hayashizaki (2001). "Functional annotation of a full-length mouse cDNA collection." Nature **409**(6821): 685-90.

Keenan, J., D. F. Orr and B. K. Pierscionek (2008). "Patterns of crystallin distribution in porcine eye lenses." Mol Vis **14**: 1245-53.

Keenan, R. J., D. M. Freymann, R. M. Stroud and P. Walter (2001). "The signal recognition particle." Annu Rev Biochem **70**: 755-75.

Kelley, L. P. and B. T. Kinsella (2003). "The role of N-linked glycosylation in determining the surface expression, G protein interaction and effector coupling of the alpha (alpha) isoform of the human thromboxane A(2) receptor." Biochim Biophys Acta **1621**(2): 192-203.

Kerscher, S., P. H. Glenister, J. Favor and M. F. Lyon (1996). "Two new cataract loci, Ccw and To3, and further mapping of the Npp and Opj cataracts in the mouse." Genomics **36**(1): 17-21.

Khan, A. O. and S. Al-Dahmesh (2009). "Age at the time of cataract surgery and relative risk for aphakic glaucoma in nontraumatic infantile cataract." J Aapos **13**(2): 166-9.

Khosravani, H. and G. W. Zamponi (2006). "Voltage-gated calcium channels and idiopathic generalized epilepsies." Physiol Rev **86**(3): 941-66.

Kim, J. I., T. Li, I. C. Ho, M. J. Grusby and L. H. Glimcher (1999). "Requirement for the c-Maf transcription factor in crystallin gene regulation and lens development." Proc Natl Acad Sci U S A **96**(7): 3781-5.

Kleinjan, D. A., A. Seawright, A. Schedl, R. A. Quinlan, S. Danes and V. van Heyningen (2001). "Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6." Hum Mol Genet **10**(19): 2049-59.

Konig, N. and G. A. Zampighi (1995). "Purification of bovine lens cell-to-cell channels composed of connexin44 and connexin50." J Cell Sci **108 (Pt 9)**: 3091-8.

Konrad, M., A. Schaller, D. Seelow, A. V. Pandey, S. Waldegger, A. Lesslauer, H. Vitzthum, Y. Suzuki, J. M. Luk, C. Becker, K. P. Schlingmann, M. Schmid, J. Rodriguez-Soriano, G. Ariceta, F. Cano, R. Enriquez, H. Juppner, S. A. Bakkaloglu, M. A. Hediger, S. Gallati, S. C. Neuhauss, P. Nurnberg and S. Weber (2006). "Mutations in the tight-junction gene claudin 19 (CLDN19) are associated with

renal magnesium wasting, renal failure, and severe ocular involvement." Am J Hum Genet **79**(5): 949-57.

Krause, G., L. Winkler, C. Piehl, I. Blasig, J. Piontek and S. L. Muller (2009). "Structure and function of extracellular claudin domains." Ann N Y Acad Sci **1165**: 34-43.

Kubo, E., D. P. Singh, N. Fatma, T. Shinohara, P. Zelenka, V. N. Reddy and L. T. Chylack (2003). "Cellular distribution of lens epithelium-derived growth factor (LEDGF) in the rat eye: loss of LEDGF from nuclei of differentiating cells." Histochem Cell Biol **119**(4): 289-99.

Kubota, K., M. Furuse, H. Sasaki, N. Sonoda, K. Fujita, A. Nagafuchi and S. Tsukita (1999). "Ca(2+)-independent cell-adhesion activity of claudins, a family of integral membrane proteins localized at tight junctions." Curr Biol **9**(18): 1035-8.

Kumar, P., S. Henikoff and P. C. Ng (2009). "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm." Nat Protoc **4**(7): 1073-81.

Kumari, S. S. and K. Varadaraj (2009). "Intact AQP0 performs cell-to-cell adhesion." Biochem Biophys Res Commun **390**(3): 1034-9.

Kuo, S. J., S. Y. Chien, C. Lin, S. E. Chan, H. T. Tsai and D. R. Chen (2010). "Significant elevation of CLDN16 and HAPLN3 gene expression in human breast cancer." Oncol Rep **24**(3): 759-66.

Kuszak, J. R., L. A. Novak and H. G. Brown (1995). "An ultrastructural analysis of the epithelial-fiber interface (EFI) in primate lenses." Exp Eye Res **61**(5): 579-97.

Kuszak, J. R., R. K. Zoltoski and C. Sivertson (2004). "Fibre cell organization in crystalline lenses." Exp Eye Res **78**(3): 673-87.

Lambert, T. D., J. Howard, A. Plant, S. Soffe and A. Roberts (2004). "Mechanisms and significance of reduced activity and responsiveness in resting frog tadpoles." J Exp Biol **207**(Pt 7): 1113-25.

Landel, C. P., J. Zhao, D. Bok and G. A. Evans (1988). "Lens-specific expression of recombinant ricin induces developmental defects in the eyes of transgenic mice." Genes Dev **2**(9): 1168-78.

Lauf, P. K., S. Misri, A. A. Chimote and N. C. Adragna (2008). "Apparent intermediate K conductance channel hyposmotic activation in human lens epithelial cells." Am J Physiol Cell Physiol **294**(3): C820-32.

Le, A.-C. N. and L. S. Musil (2001a). "FGF Signaling in Chick Lens Development." Developmental Biology **233**(2): 394-411.

Le, A. C. and L. S. Musil (2001b). "A novel role for FGF and extracellular signal-regulated kinase in gap junction-mediated intercellular communication in the lens." J Cell Biol **154**(1): 197-216.

Leonard, M., Y. Chan and A. S. Menko (2008). "Identification of a novel intermediate filament-linked N-cadherin/gamma-catenin complex involved in the establishment of the cytoarchitecture of differentiated lens fiber cells." Dev Biol.

Lequeux, L., M. Rio, A. Vigouroux, M. Titeux, H. Etchevers, F. Malecaze, N. Chassaing and P. Calvas (2008). "Confirmation of RAX gene involvement in human anophthalmia." Clin Genet **74**(4): 392-5.

Letts, V. A., R. Felix, G. H. Biddlecome, J. Arikkath, C. L. Mahaffey, A. Valenzuela, F. S. Bartlett, 2nd, Y. Mori, K. P. Campbell and W. N. Frankel (1998). "The mouse stargazer gene encodes a neuronal Ca2+-channel gamma subunit." Nat Genet **19**(4): 340-7.

Lewis, T. L. and D. Maurer (2009). "Effects of early pattern deprivation on visual development." Optom Vis Sci **86**(6): 640-6.

Li, L. K., L. So and A. Spector (1985). "Membrane cholesterol and phospholipid in consecutive concentric sections of human lenses." J Lipid Res **26**(5): 600-9.

Lin, Z., K. Witschas, T. Garcia, R. S. Chen, J. P. Hansen, Z. M. Sellers, E. Kuzmenkina, S. Herzig and P. M. Best (2008). "A critical GxxxA motif in the gamma6 calcium channel subunit mediates its inhibitory effect on Cav3.1 calcium current." J Physiol **586**(Pt 22): 5349-66.

Liska, F., B. Chylikova, J. Martinek and V. Kren (2008). "Microphthalmia and cataract in rats with a novel point mutation in connexin 50 - L7Q." Mol Vis **14**: 823-8.

Liu, H., G. C. Ippolito, J. K. Wall, T. Niu, L. Probst, B. S. Lee, K. Pulford, A. H. Banham, L. Stockwin, A. L. Shaffer, L. M. Staudt, C. Das, M. J. Dyer and P. W. Tucker (2006a). "Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells." Mol Cancer **5**: 18.

Liu, P., J. R. Keller, M. Ortiz, L. Tessarollo, R. A. Rachel, T. Nakamura, N. A. Jenkins and N. G. Copeland (2003). "Bcl11a is essential for normal lymphoid development." Nat Immunol **4**(6): 525-32.

Liu, W., O. V. Lagutin, M. Mende, A. Streit and G. Oliver (2006b). "Six3 activation of Pax6 expression is essential for mammalian lens induction and specification." Embo J **25**(22): 5383-95.

Lo, W. K. and C. V. Harding (1983). "Tight junctions in the lens epithelia of human and frog: freeze-fracture and protein tracer studies." Invest Ophthalmol Vis Sci **24**(4): 396-402.

Lo, W. K., A. P. Shaw, D. F. Paulsen and A. Mills (2000). "Spatiotemporal distribution of zonulae adherens and associated actin bundles in both epithelium and fiber cells during chicken lens development." Exp Eye Res **71**(1): 45-55.

Lo, W. K., A. P. Shaw and X. J. Wen (1997). "Actin filament bundles in cortical fiber cells of the rat lens." Exp Eye Res **65**(5): 691-701.

Lovicu, F. J. and J. W. McAvoy (1999). "Spatial and temporal expression of p57(KIP2) during murine lens development." Mech Dev **86**(1-2): 165-9.

Lovicu, F. J. and J. W. McAvoy (2001). FGF-induced lens cell proliferation and differentiation is dependent on MAPK (ERK1/2) signalling. **128:** 5075-5084.

Lovicu, F. J. and J. W. McAvoy (2005). "Growth factor regulation of lens development." Dev Biol **280**(1): 1-14.

Luzio, J. P., P. R. Pryor and N. A. Bright (2007). "Lysosomes: fusion and function." Nat Rev Mol Cell Biol **8**(8): 622-32.

Macdonald-Obermann, J. L. and L. J. Pike (2009). "Palmitoylation of the EGF receptor impairs signal transduction and abolishes high-affinity ligand binding." Biochemistry **48**(11): 2505-13.

Mackay, D. S., U. P. Andley and A. Shiels (2003). "Cell death triggered by a novel mutation in the alphaA-crystallin gene underlies autosomal dominant cataract linked to chromosome 21q." Eur J Hum Genet **11**(10): 784-93.

Maddala, R., N. Skiba and P. Vasantha Rao (2007). "Lens fiber cell elongation and differentiation is associated with a robust increase in myosin light chain phosphorylation in the developing mouse." Differentiation **75**(8): 713-25.

Maisel, H. and M. M. Perry (1972). "Electron microscope observations on some structural proteins of the chick lens." Exp Eye Res **14**(1): 7-12.

Marquardt, A., H. Stohr, L. A. Passmore, F. Kramer, A. Rivera and B. H. Weber (1998). "Mutations in a novel gene, VMD2, encoding a protein of unknown properties cause juvenile-onset vitelliform macular dystrophy (Best's disease)." Hum Mol Genet **7**(9): 1517-25.

Martinez, G., M. Wijesinghe, K. Turner, H. E. Abud, M. M. Taketo, T. Noda, M. L. Robinson and R. U. de Iongh (2009). "Conditional mutations of beta-catenin and APC reveal roles for canonical Wnt signaling in lens differentiation." Invest Ophthalmol Vis Sci **50**(10): 4794-806.

Matsui, M., A. Yamamoto, A. Kuma, Y. Ohsumi and N. Mizushima (2006). "Organelle degradation during the lens and erythroid differentiation is independent of autophagy." Biochem Biophys Res Commun **339**(2): 485-9.

McGinnes, L. W. and T. G. Morrison (1997). "Disulfide bond formation is a determinant of glycosylation site usage in the hemagglutinin-neuraminidase glycoprotein of Newcastle disease virus." J Virol **71**(4): 3083-9.

Medina-Martinez, O., I. Brownell, F. Amaya-Manzanares, Q. Hu, R. R. Behringer and M. Jamrich (2005). "Severe defects in proliferation and differentiation of lens cells in Foxe3 null mice." Mol Cell Biol **25**(20): 8854-63.

Medvedev, A. E., W. Piao, J. Shoenfelt, S. H. Rhee, H. Chen, S. Basu, L. M. Wahl, M. J. Fenton and S. N. Vogel (2007). "Role of TLR4 tyrosine phosphorylation in signal transduction and endotoxin tolerance." J Biol Chem **282**(22): 16042-53.

Meissner, A. and T. Noack (2008). "Proliferation of human lens epithelial cells (HLE-B3) is inhibited by blocking of voltage-gated calcium channels." Pflugers Arch.

Menzel, S., C. Garner, I. Gut, F. Matsuda, M. Yamaguchi, S. Heath, M. Foglio, D. Zelenika, A. Boland, H. Rooks, S. Best, T. D. Spector, M. Farrall, M. Lathrop and S. L. Thein (2007). "A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15." Nat Genet **39**(10): 1197-9.

Metere, A., E. Iorio, D. Pietraforte, F. Podo and M. Minetti (2009). "Peroxynitrite signaling in human erythrocytes: synergistic role of hemoglobin oxidation and band 3 tyrosine phosphorylation." Arch Biochem Biophys **484**(2): 173-82.

Meusser, B., C. Hirsch, E. Jarosch and T. Sommer (2005). "ERAD: the long road to destruction." Nat Cell Biol **7**(8): 766-72.

Michael, R., J. van Marle, G. F. Vrensen and T. J. van den Berg (2003). "Changes in the refractive index of lens fibre membranes during maturation--impact on lens transparency." Exp Eye Res **77**(1): 93-9.

Milstein, A. D. and R. A. Nicoll (2009). "TARP modulation of synaptic AMPA receptor trafficking and gating depends on multiple intracellular domains." Proc Natl Acad Sci U S A **106**(27): 11348-51.

Milstein, A. D., W. Zhou, S. Karimzadegan, D. S. Bredt and R. A. Nicoll (2007). "TARP subtypes differentially and dose-dependently control synaptic AMPA receptor gating." Neuron **55**(6): 905-18.

Mishra, A., P. Dikshit, S. Purkayastha, J. Sharma, N. Nukina and N. R. Jana (2008). "E6-AP promotes misfolded polyglutamine proteins for proteasomal degradation and suppresses polyglutamine protein aggregation and toxicity." J Biol Chem **283**(12): 7648-56.

Miyoshi, J. and Y. Takai (2005). "Molecular perspective on tight-junction assembly and epithelial polarity." Adv Drug Deliv Rev **57**(6): 815-55.

Moldvay, J., M. Jackel, C. Paska, I. Soltesz, Z. Schaff and A. Kiss (2007). "Distinct claudin expression profile in histologic subtypes of lung cancer." Lung Cancer **57**(2): 159-67.

Molinari, M., V. Calanca, C. Galli, P. Lucca and P. Paganetti (2003). "Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle." Science **299**(5611): 1397-400.

Moody, S. A. (1987). "Fates of the blastomeres of the 16-cell stage Xenopus embryo." Dev Biol **119**(2): 560-78.

Moore, K. L. (2003). "The biology and enzymology of protein tyrosine O-sulfation." J Biol Chem **278**(27): 24243-6.

Morgenbesser, S. D., B. O. Williams, T. Jacks and R. A. DePinho (1994). "p53-dependent apoptosis produced by Rb-deficiency in the developing mouse lens." Nature **371**(6492): 72-4.

Morozov, V. and E. F. Wawrousek (2006). "Caspase-dependent secondary lens fiber cell disintegration in alphaA-/alphaB-crystallin double-knockout mice." Development **133**(5): 813-21.

Morrison, D., D. FitzPatrick, I. Hanson, K. Williamson, V. van Heyningen, B. Fleck, I. Jones, J. Chalmers and H. Campbell (2002). "National study of microphthalmia, anophthalmia, and coloboma (MAC) in Scotland: investigation of genetic aetiology." J Med Genet **39**(1): 16-22.

Mostov, K., T. Su and M. ter Beest (2003). "Polarized epithelial membrane traffic: conservation and plasticity." Nat Cell Biol **5**(4): 287-93.

Mrsny, R. J., G. T. Brown, K. Gerner-Smidt, A. G. Buret, J. B. Meddings, C. Quan, M. Koval and A. Nusrat (2008). "A key claudin extracellular loop domain is critical for epithelial barrier integrity." Am J Pathol **172**(4): 905-15.

Muller, M., S. S. Bhattacharya, T. Moore, Q. Prescott, T. Wedig, H. Herrmann and T. M. Magin (2009). "Dominant cataract formation in association with a vimentin assembly disrupting mutation." Hum Mol Genet **18**(6): 1052-7.

Nagafuchi, A., Y. Shirayoshi, K. Okazaki, K. Yasuda and M. Takeichi (1987). "Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA." Nature **329**(6137): 341-3.

Nakahara, M., A. Nagasaka, M. Koike, K. Uchida, K. Kawane, Y. Uchiyama and S. Nagata (2007). "Degradation of nuclear DNA by DNase II-like acid DNase in cortical fiber cells of mouse eye lens." Febs J **274**(12): 3055-64.

Negash, S., H. S. Wang, C. Gao, D. Ledee and P. Zelenka (2002). "Cdk5 regulates cell-matrix and cell-cell adhesion in lens epithelial cells." J Cell Sci **115**(Pt 10): 2109-17.

Nelis, E., C. Van Broeckhoven, P. De Jonghe, A. Lofgren, A. Vandenberghe, P. Latour, E. Le Guern, A. Brice, M. L. Mostacciuolo, F. Schiavon, F. Palau, S. Bort, M. Upadhyaya, M. Rocchi, N. Archidiacono, P. Mandich, E. Bellone, K. Silander, M. L. Savontaus, R. Navon, H. Goldberg-Stern, X. Estivill, V. Volpini, W. Friedl, A. Gal and et al. (1996). "Estimation of the mutation frequencies in Charcot-Marie-Tooth disease type 1 and hereditary neuropathy with liability to pressure palsies: a European collaborative study." Eur J Hum Genet **4**(1): 25-33.

Newport, J. and M. Kirschner (1982). "A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription." Cell **30**(3): 687-96.

Ng, D., N. Thakker, C. M. Corcoran, D. Donnai, R. Perveen, A. Schneider, D. W. Hadley, C. Tifft, L. Zhang, A. O. Wilkie, J. J. van der Smagt, R. J. Gorlin, S. M. Burgess, V. J. Bardwell, G. C. Black and L. G. Biesecker (2004). "Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR." Nat Genet **36**(4): 411-6.

Ng, S. B., K. J. Buckingham, C. Lee, A. W. Bigham, H. K. Tabor, K. M. Dent, C. D. Huff, P. T. Shannon, E. W. Jabs, D. A. Nickerson, J. Shendure and M. J. Bamshad (2010). "Exome sequencing identifies the cause of a mendelian disorder." Nat Genet **42**(1): 30-5.

Nielsen, H. and A. Krogh (1998). "Prediction of signal peptides and signal anchors by a hidden Markov model." Proc Int Conf Intell Syst Mol Biol **6**: 122-30.

Nielsen, P. A., A. Baruch, V. I. Shestopalov, B. N. Giepmans, I. Dunia, E. L. Benedetti and N. M. Kumar (2003). "Lens connexins alpha3Cx46 and alpha8Cx50 interact with zonula occludens protein-1 (ZO-1)." Mol Biol Cell **14**(6): 2470-81.

Nieuwkoop, P. D. and J. Faber (1967). Normal table of Xenopus Laevis. Amsterdam, The Netherlands, North Holland Publishing Co.

Nishiguchi, S., H. Wood, H. Kondoh, R. Lovell-Badge and V. Episkopou (1998). "Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice." Genes Dev **12**(6): 776-81.

Nishimura, D. Y., L. M. Baye, R. Perveen, C. C. Searby, A. Avila-Fernandez, I. Pereiro, C. Ayuso, D. Valverde, P. N. Bishop, F. D. Manson, J. Urquhart, E. M. Stone, D. C. Slusarski, G. C. Black and V. C. Sheffield (2010). "Discovery and functional analysis of a retinitis pigmentosa gene, C2ORF71." Am J Hum Genet **86**(5): 686-95.

Nitta, T., M. Hata, S. Gotoh, Y. Seo, H. Sasaki, N. Hashimoto, M. Furuse and S. Tsukita (2003). "Size-selective loosening of the blood-brain barrier in claudin-5-deficient mice." J Cell Biol **161**(3): 653-60.

Notterpek, L., K. J. Roux, S. A. Amici, A. Yazdanpour, C. Rahner and B. S. Fletcher (2001). "Peripheral myelin protein 22 is a constituent of intercellular junctions in epithelia." Proc Natl Acad Sci U S A **98**(25): 14404-9.

Nunbhakdi-Craig, V., T. Machleidt, E. Ogris, D. Bellotto, C. L. White, 3rd and E. Sontag (2002). "Protein phosphatase 2A associates with and regulates atypical PKC and the epithelial tight junction complex." J Cell Biol **158**(5): 967-78.

Oda, S., K. Watanabe, H. Fujisawa and Y. Kameyama (1980). "Impaired development of lens fibers in genetic microphthalmia, eye lens obsolescence, Elo, of the mouse." Exp Eye Res **31**(6): 673-81.

Oda, Y., N. Hosokawa, I. Wada and K. Nagata (2003). "EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin." Science **299**(5611): 1394-7.

Ohno, Y., A. Ito, R. Ogata, Y. Hiraga, Y. Igarashi and A. Kihara (2009). "Palmitoylation of the sphingosine 1-phosphate receptor S1P is involved in its signaling functions and internalization." Genes Cells **14**(8): 911-23.

Ohtaka-Maruyama, C., F. Hanaoka and A. B. Chepelinsky (1998). "A novel alternative spliced variant of the transcription factor AP2alpha is expressed in the murine ocular lens." Dev Biol **202**(1): 125-35.

Oka, M., H. Kudo, N. Sugama, Y. Asami and M. Takehana (2008). "The function of filensin and phakinin in lens transparency." Mol Vis **14**: 815-22.

Okada, T., H. Yoshida, R. Akazawa, M. Negishi and K. Mori (2002). "Distinct roles of activating transcription factor 6 (ATF6) and double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK) in transcription during the mammalian unfolded protein response." Biochem J **366**(Pt 2): 585-94.

Okinami, S. (1978). "Freeze-fracture replica of the primate lens fibers." Albrecht Von Graefes Arch Klin Exp Ophthalmol **209**(1): 51-8.

Ostergaard, E., M. Batbayli, M. Duno, K. Vilhelmsen and T. Rosenberg (2010). "Mutations in PCDH21 cause autosomal recessive cone-rod dystrophy." J Med Genet **47**(10): 665-9.

Parmantier, E., F. Cabon, C. Braun, D. D'Urso, H. W. Muller and B. Zalc (1995). "Peripheral myelin protein-22 is expressed in rat and mouse brain and spinal cord motoneurons." Eur J Neurosci **7**(5): 1080-8.

Paschoud, S., M. Bongiovanni, J. C. Pache and S. Citi (2007). "Claudin-1 and claudin-5 expression patterns differentiate lung squamous cell carcinomas from adenocarcinomas." Mod Pathol **20**(9): 947-54.

Pascolini, D., S. P. Mariotti, G. P. Pokharel, R. Pararajasegaram, D. Etya'ale, A. D. Negrel and S. Resnikoff (2004). "2002 global update of available data on visual

impairment: a compilation of population-based prevalence studies." Ophthalmic Epidemiol **11**(2): 67-115.

Patterson, C. and N. Delamere (1992). The lens. Adler's Physiology of the Eye. W. Hart. St Louis, CV Mosby**:** 348-390.

Paul, D. L., L. Ebihara, L. J. Takemoto, K. I. Swenson and D. A. Goodenough (1991). "Connexin46, a novel lens gap junction protein, induces voltage-gated currents in nonjunctional plasma membrane of Xenopus oocytes." J Cell Biol **115**(4): 1077-89.

Pei, Y. F. and J. A. Rhodin (1970). "The prenatal development of the mouse eye." Anat Rec **168**(1): 105-25.

Percherancier, Y., T. Planchenault, A. Valenzuela-Fernandez, J. L. Virelizier, F. Arenzana-Seisdedos and F. Bachelerie (2001). "Palmitoylation-dependent control of degradation, life span, and membrane expression of the CCR5 receptor." J Biol Chem **276**(34): 31936-44.

Pevsner, J. (2009). Pairwise Sequence Alignment. Bioinformatics and Functional Genomics Wiley-Blackwell**:** 74-75.

Piehl, C., J. Piontek, J. Cording, H. Wolburg and I. E. Blasig (2010). "Participation of the second extracellular loop of claudin-5 in paracellular tightening against ions, small and large molecules." Cell Mol Life Sci **67**(12): 2131-40.

Piontek, J., L. Winkler, H. Wolburg, S. L. Muller, N. Zuleger, C. Piehl, B. Wiesner, G. Krause and I. E. Blasig (2008). "Formation of tight junction: determinants of homophilic interaction between classic claudins." Faseb J **22**(1): 146-58.

Ponnam, S. P., K. Ramesha, S. Tejwani, J. Matalia and C. Kannabiran (2008). "A missense mutation in LIM2 causes autosomal recessive congenital cataract." Mol Vis **14**: 1204-8.

Ponnam, S. P., K. Ramesha, S. Tejwani, B. Ramamurthy and C. Kannabiran (2007). "Mutation of the gap junction protein alpha 8 (GJA8) gene causes autosomal recessive cataract." J Med Genet **44**(7): e85.

Pontoriero, G. F., A. N. Smith, L. A. Miller, G. L. Radice, J. A. West-Mays and R. A. Lang (2009). "Co-operative roles for E-cadherin and N-cadherin during lens vesicle separation and lens epithelial cell survival." Dev Biol **326**(2): 403-17.

Pras, E., M. Frydman, E. Levy-Nissenbaum, T. Bakhan, J. Raz, E. I. Assia, B. Goldman and E. Pras (2000). "A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family." Invest Ophthalmol Vis Sci **41**(11): 3511-5.

Pras, E., E. Levy-Nissenbaum, T. Bakhan, H. Lahat, E. Assia, N. Geffen-Carmi, M. Frydman, B. Goldman and E. Pras (2002). "A missense mutation in the LIM2 gene is associated with autosomal recessive presenile cataract in an inbred Iraqi Jewish family." Am J Hum Genet **70**(5): 1363-7.

Price, M. G., C. F. Davis, F. Deng and D. L. Burgess (2005). "The alpha-amino-3-hydroxyl-5-methyl-4-isoxazolepropionate receptor trafficking regulator "stargazin" is related to the claudin family of proteins by Its ability to mediate cell-cell adhesion." J Biol Chem **280**(20): 19711-20.

Rahi, J. S. and C. Dezateux (2001). "Measuring and interpreting the incidence of congenital ocular anomalies: lessons from a national study of congenital cataract in the UK." Invest Ophthalmol Vis Sci **42**(7): 1444-8.

Ramachandran, R. D., V. Perumalsamy and J. F. Hejtmancik (2007). "Autosomal recessive juvenile onset cataract associated with mutation in BFSP1." Hum Genet **121**(3-4): 475-82.

Ramensky, V., P. Bork and S. Sunyaev (2002). "Human non-synonymous SNPs: server and survey." Nucleic Acids Res **30**(17): 3894-900.

Rao, G. N., K. A. Gutekunst and R. L. Church (1989). "Bovine lens 23, 21 and 19 kDa intrinsic membrane proteins have an identical amino-terminal amino acid sequence." FEBS Lett **250**(2): 483-6.

Rath, A., M. Glibowicka, V. G. Nadeau, G. Chen and C. M. Deber (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins." Proc Natl Acad Sci U S A **106**(6): 1760-5.

Rattner, A., P. M. Smallwood, J. Williams, C. Cooke, A. Savchenko, A. Lyubarsky, E. N. Pugh and J. Nathans (2001). "A photoreceptor-specific cadherin is essential for the structural integrity of the outer segment and for photoreceptor survival." Neuron **32**(5): 775-86.

Reddy, M. A., P. J. Francis, V. Berry, S. S. Bhattacharya and A. T. Moore (2004).

"Molecular genetic basis of inherited cataract and associated phenotypes." Surv Ophthalmol **49**(3): 300-15.

Reis, L. M., R. C. Tyler, A. Schneider, T. Bardakjian and E. V. Semina (2010a). "Examination of SOX2 in variable ocular conditions identifies a recurrent deletion in microphthalmia and lack of mutations in other phenotypes." Mol Vis **16**: 768-73.

Reis, L. M., R. C. Tyler, A. Schneider, T. Bardakjian, J. M. Stoler, S. B. Melancon and E. V. Semina (2010b). "FOXE3 plays a significant role in autosomal recessive microphthalmia." Am J Med Genet A **152A**(3): 582-90.

Ren, J., L. Wen, X. Gao, C. Jin, Y. Xue and X. Yao (2008). "CSS-Palm 2.0: an updated software for palmitoylation sites prediction." Protein Eng Des Sel **21**(11): 639-44.

Reynolds, S. M., L. Kall, M. E. Riffle, J. A. Bilmes and W. S. Noble (2008). "Transmembrane topology and signal peptide prediction using dynamic bayesian networks." PLoS Comput Biol **4**(11): e1000213.

Rieger, D. K., E. Reichenberger, W. McLean, A. Sidow and B. R. Olsen (2001). "A double-deletion mutation in the Pitx3 gene causes arrested lens development in aphakia mice." Genomics **72**(1): 61-72.

Rioux, J. D., R. J. Xavier, K. D. Taylor, M. S. Silverberg, P. Goyette, A. Huett, T. Green, P. Kuballa, M. M. Barmada, L. W. Datta, Y. Y. Shugart, A. M. Griffiths, S. R. Targan, A. F. Ippoliti, E. J. Bernard, L. Mei, D. L. Nicolae, M. Regueiro, L. P. Schumm, A. H. Steinhart, J. I. Rotter, R. H. Duerr, J. H. Cho, M. J. Daly and S. R. Brant (2007). "Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis." Nat Genet **39**(5): 596-604.

Ritter, C., K. Quirin, M. Kowarik and A. Helenius (2005). "Minor folding defects trigger local modification of glycoproteins by the ER folding sensor GT." Embo J **24**(9): 1730-8.

Robinson, M. L. (2006). "An essential role for FGF receptor signaling in lens development." Semin Cell Dev Biol **17**(6): 726-40.

Roboti, P., E. Swanton and S. High (2009). "Differences in endoplasmic-reticulum quality control determine the cellular response to disease-associated mutants of proteolipid protein." J Cell Sci **122**(Pt 21): 3942-53.

Rochow, N., N. Bachmaier, F. Tost, J. F. Beck and T. Bernig (2009). "The case of a 1-year-old girl with hereditary hyperferritinemia cataract syndrome." Pediatr Hematol Oncol **26**(3): 136-41.

Rodriguez-Boulan, E. and A. Musch (2005). "Protein sorting in the Golgi complex: shifting paradigms." Biochim Biophys Acta **1744**(3): 455-64.

Roh, M. H., C. J. Liu, S. Laurinec and B. Margolis (2002). "The carboxyl terminus of zona occludens-3 binds and recruits a mammalian homologue of discs lost to tight junctions." J Biol Chem **277**(30): 27501-9.

Ron, D. (2002). "Translational control in the endoplasmic reticulum stress response." J Clin Invest **110**(10): 1383-8.

Rong, P., X. Wang, I. Niesman, Y. Wu, L. E. Benedetti, I. Dunia, E. Levy and X. Gong (2002). "Disruption of Gja8 (alpha8 connexin) in mice leads to microphthalmia associated with retardation of lens growth and lens fiber maturation." Development **129**(1): 167-74.

Roper, K., D. Corbeil and W. B. Huttner (2000). "Retention of prominin in microvilli reveals distinct cholesterol-based lipid micro-domains in the apical plasma membrane." Nat Cell Biol **2**(9): 582-92.

Roux, K. J., S. A. Amici and L. Notterpek (2004). "The temporospatial expression of peripheral myelin protein 22 at the developing blood-nerve and blood-brain barriers." J Comp Neurol **474**(4): 578-88.

Ruffer, C. and V. Gerke (2004). "The C-terminal cytoplasmic tail of claudins 1 and 5 but not its PDZ-binding motif is required for apical localization at epithelial and endothelial tight junctions." Eur J Cell Biol **83**(4): 135-44.

Sager, C., J. Terhag, S. Kott and M. Hollmann (2009). "C-terminal domains of transmembrane alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) receptor regulatory proteins not only facilitate trafficking but are major modulators of AMPA receptor function." J Biol Chem **284**(47): 32413-24.

Saliba, R. S., P. M. Munro, P. J. Luthert and M. E. Cheetham (2002). "The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation." J Cell Sci **115**(Pt 14): 2907-18.

Sandilands, A., A. R. Prescott, J. M. Carter, A. M. Hutcheson, R. A. Quinlan, J. Richards and P. G. FitzGerald (1995). "Vimentin and CP49/filensin form distinct networks in

the lens which are independently modulated during lens fibre cell differentiation." J Cell Sci **108 (Pt 4)**: 1397-406.

Sandilands, A., A. R. Prescott, A. Wegener, R. K. Zoltoski, A. M. Hutcheson, S. Masaki, J. R. Kuszak and R. A. Quinlan (2003). "Knockout of the intermediate filament protein CP49 destabilises the lens fibre cell cytoskeleton and decreases lens optical quality, but does not induce cataract." Exp Eye Res **76**(3): 385-91.

Sandoval, A., J. Arikkath, E. Monjaraz, K. P. Campbell and R. Felix (2007). "gamma1-dependent down-regulation of recombinant voltage-gated Ca2+ channels." Cell Mol Neurobiol **27**(7): 901-8.

SanGiovanni, J. P., E. Y. Chew, G. F. Reed, N. A. Remaley, J. B. Bateman, T. A. Sugimoto and M. A. Klebanoff (2002). "Infantile cataract in the collaborative perinatal project: prevalence and risk factors." Arch Ophthalmol **120**(11): 1559-65.

Sarkar, S. N., R. Q. Huang, S. M. Logan, K. D. Yi, G. H. Dillon and J. W. Simpkins (2008). "Estrogens directly potentiate neuronal L-type Ca2+ channels." Proc Natl Acad Sci U S A **105**(39): 15148-53.

Schatz, G. and B. Dobberstein (1996). "Common principles of protein translocation across membranes." Science **271**(5255): 1519-26.

Scheiffele, P., J. Peranen and K. Simons (1995). "N-glycans as apical sorting signals in epithelial cells." Nature **378**(6552): 96-8.

Schmid, F., E. Glaus, F. P. Cremers, B. Kloeckener-Gruissem, W. Berger and J. Neidhardt (2010). "Mutation- and tissue-specific alterations of RPGR transcripts." Invest Ophthalmol Vis Sci **51**(3): 1628-35.

Schmidt, W., N. Klopp, T. Illig and J. Graw (2008). "A novel GJA8 mutation causing a recessive triangular cataract." Mol Vis **14**: 851-6.

Schnell, E., M. Sizemore, S. Karimzadegan, L. Chen, D. S. Bredt and R. A. Nicoll (2002). "Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number." Proc Natl Acad Sci U S A **99**(21): 13902-7.

Segerdell, E., J. B. Bowes, N. Pollet and P. D. Vize (2008). "An ontology for Xenopus anatomy and development." BMC Dev Biol **8**: 92.

Shah, S. P., A. Taylor, J. C. Sowden, N. K. Ragge, I. Russell-Eggitt, J. S. Rahi and C. Gilbert (2010). "Anophthalmos, microphthalmos and typical coloboma in the UK: a prospective study of incidence and risk." Invest Ophthalmol Vis Sci.

Shaham, O., A. N. Smith, M. L. Robinson, M. M. Taketo, R. A. Lang and R. Ashery-Padan (2009). "Pax6 is essential for lens fiber cell differentiation." Development **136**(15): 2567-78.

Sharma, S., S. L. Ang, M. Shaw, D. A. Mackey, J. Gecz, J. W. McAvoy and J. E. Craig (2006). "Nance-Horan syndrome protein, NHS, associates with epithelial cell junctions." Hum Mol Genet **15**(12): 1972-83.

Sharma, S., K. S. Koh, C. Collin, A. Dave, A. McMellon, Y. Sugiyama, J. W. McAvoy, A. K. Voss, J. Gecz and J. E. Craig (2009). "NHS-A isoform of the NHS gene is a novel interactor of ZO-1." Exp Cell Res **315**(14): 2358-72.

Sharom, F. J. and M. T. Lehto (2002). "Glycosylphosphatidylinositol-anchored proteins: structure, function, and cleavage by phosphatidylinositol-specific phospholipase C." Biochem Cell Biol **80**(5): 535-49.

Sharp, A. H., J. L. Black, 3rd, S. J. Dubel, S. Sundarraj, J. P. Shen, A. M. Yunker, T. D. Copeland and M. W. McEnery (2001). "Biochemical and anatomical evidence for specialized voltage-dependent calcium channel gamma isoform expression in the epileptic and ataxic mouse, stargazer." Neuroscience **105**(3): 599-617.

Shestopalov, V. I. and S. Bassnett (2003). "Development of a macromolecular diffusion pathway in the lens." J Cell Sci **116**(Pt 20): 4191-9.

Shi, Y., K. Barton, A. De Maria, J. M. Petrash, A. Shiels and S. Bassnett (2009). "The stratified syncytium of the vertebrate lens." J Cell Sci **122**(Pt 10): 1607-15.

Shiels, A., S. Bassnett, K. Varadaraj, R. Mathias, K. Al-Ghoul, J. Kuszak, D. Donoviel, S. Lilleberg, G. Friedrich and B. Zambrowicz (2001). "Optical dysfunction of the crystalline lens in aquaporin-0-deficient mice." Physiol Genomics **7**(2): 179-86.

Shiels, A., J. M. King, D. S. Mackay and S. Bassnett (2007). "Refractive defects and cataracts in mice lacking lens intrinsic membrane protein-2." Invest Ophthalmol Vis Sci **48**(2): 500-8.

Shimizu, F., Y. Sano, T. Maeda, M. A. Abe, H. Nakayama, R. Takahashi, M. Ueda, S. Ohtsuki, T. Terasaki, M. Obinata and T. Kanda (2008). "Peripheral nerve pericytes originating from the blood-nerve barrier expresses tight junctional molecules and transporters as barrier-forming cells." J Cell Physiol **217**(2): 388-99.

Simon, D. B., Y. Lu, K. A. Choate, H. Velazquez, E. Al-Sabban, M. Praga, G. Casari, A. Bettinelli, G. Colussi, J. Rodriguez-Soriano, D. McCredie, D. Milford, S. Sanjad and R. P. Lifton (1999). "Paracellin-1, a renal tight junction protein required for paracellular Mg2+ resorption." Science **285**(5424): 103-6.

Simske, J. S., M. Koppen, P. Sims, J. Hodgkin, A. Yonkof and J. Hardin (2003). "The cell junction protein VAB-9 regulates adhesion and epidermal morphology in C. elegans." Nat Cell Biol **5**(7): 619-25.

Singaraja, R. R., M. H. Kang, K. Vaid, S. S. Sanders, G. L. Vilas, P. Arstikaitis, J. Coutinho, R. C. Drisdel, D. El-Husseini Ael, W. N. Green, L. Berthiaume and M. R. Hayden (2009). "Palmitoylation of ATP-binding cassette transporter A1 is essential for its trafficking and function." Circ Res **105**(2): 138-47.

Singh, D. P., N. Ohguro, T. Kikuchi, T. Sueno, V. N. Reddy, K. Yuge, L. T. Chylack, Jr. and T. Shinohara (2000). "Lens epithelium-derived growth factor: effects on growth and survival of lens epithelial cells, keratinocytes, and fibroblasts." Biochem Biophys Res Commun **267**(1): 373-81.

Sladek, R., G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, B. Balkau, B. Heude, G. Charpentier, T. J. Hudson, A. Montpetit, A. V. Pshezhetsky, M. Prentki, B. I. Posner, D. J. Balding, D. Meyre, C. Polychronakos and P. Froguel (2007). "A genome-wide association study identifies novel risk loci for type 2 diabetes." Nature **445**(7130): 881-5.

Smaoui, N., O. Beltaief, S. BenHamed, R. M'Rad, F. Maazoul, A. Ouertani, H. Chaabouni and J. F. Hejtmancik (2004). "A homozygous splice mutation in the HSF4 gene is associated with an autosomal recessive congenital cataract." Invest Ophthalmol Vis Sci **45**(8): 2716-21.

Song, S., A. Landsbury, R. Dahm, Y. Liu, Q. Zhang and R. A. Quinlan (2009). "Functions of the intermediate filament cytoskeleton in the eye lens." J Clin Invest **119**(7): 1837-48.

Sonnhammer, E. L., G. von Heijne and A. Krogh (1998). "A hidden Markov model for predicting transmembrane helices in protein sequences." Proc Int Conf Intell Syst Mol Biol **6**: 175-82.

Stager, D. R., Jr., J. Felius and G. R. Beauchamp (2009). "Congenital cataract cost." Ophthalmology **116**(12): 2484 e1-2.

Steele, E. C., Jr., S. Kerscher, M. F. Lyon, P. H. Glenister, J. Favor, J. Wang and R. L. Church (1997). "Identification of a mutation in the MP19 gene, Lim2, in the cataractous mouse mutant To3." Mol Vis **3**: 5.

Steele, E. C., Jr., J. H. Wang, W. K. Lo, D. A. Saperstein, X. Li and R. L. Church (2000). "Lim2(To3) transgenic mice establish a causative relationship between the mutation identified in the lim2 gene and cataractogenesis in the To3 mouse mutant." Mol Vis **6**: 85-94.

Stein, E. L. and D. M. Chetkovich (2010). "Regulation of stargazin synaptic trafficking by C-terminal PDZ ligand phosphorylation in bidirectional synaptic plasticity." J Neurochem **113**(1): 42-53.

Steinberg, M. S. (1996). "Adhesion in development: an historical overview." Dev Biol **180**(2): 377-88.

Stojanovic, A., I. Hwang, H. G. Khorana and J. Hwa (2003). "Retinitis pigmentosa rhodopsin mutations L125R and A164V perturb critical interhelical interactions: new insights through compensatory mutations and crystal structure analysis." J Biol Chem **278**(40): 39020-8.

Stradal, T., K. D. Courtney, K. Rottner, P. Hahne, J. V. Small and A. M. Pendergast (2001). "The Abl interactor proteins localize to sites of actin polymerization at the tips of lamellipodia and filopodia." Curr Biol **11**(11): 891-5.

Straub, B. K., J. Boda, C. Kuhn, M. Schnoelzer, U. Korf, T. Kempf, H. Spring, M. Hatzfeld and W. W. Franke (2003). "A novel cell-cell junction system: the cortex adhaerens mosaic of lens fiber cells." J Cell Sci **116**(Pt 24): 4985-95.

Sugiyama, Y., A. R. Prescott, F. M. Tholozan, S. Ohno and R. A. Quinlan (2008). "Expression and localisation of apical junctional complex proteins in lens epithelial cells." Exp Eye Res.

Summerton, J. E. (2007). "Morpholino, siRNA, and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity." Curr Top Med Chem **7**(7): 651-60.

Taylor, V., A. A. Welcher, A. E. Program and U. Suter (1995). "Epithelial membrane

protein-1, peripheral myelin protein 22, and lens membrane protein 20 define a novel gene family." J Biol Chem **270**(48): 28824-33.

Tenbroek, E., M. Arneson, L. Jarvis and C. Louis (1992). The distribution of the fiber cell intrinsic membrane proteins MP20 and connexin46 in the bovine lens. **103:** 245-257.

Tester, D. J., L. B. Cronk, J. L. Carr, V. Schulz, B. A. Salisbury, R. S. Judson and M. J. Ackerman (2006). "Allelic dropout in long QT syndrome genetic testing: a possible mechanism underlying false-negative results." Heart Rhythm **3**(7): 815-21.

Tholozan, F. M. and R. A. Quinlan (2007). "Lens cells: More than meets the eye." Int J Biochem Cell Biol **39**(10): 1754-9.

Tokhtaeva, E., K. Munson, G. Sachs and O. Vagin (2010). "N-glycan-dependent quality control of the Na,K-ATPase beta(2) subunit." Biochemistry **49**(14): 3116-28.

Tomita, S., H. Adesnik, M. Sekiguchi, W. Zhang, K. Wada, J. R. Howe, R. A. Nicoll and D. S. Bredt (2005). "Stargazin modulates AMPA receptor gating and trafficking by distinct domains." Nature **435**(7045): 1052-8.

Tomita, S., L. Chen, Y. Kawasaki, R. S. Petralia, R. J. Wenthold, R. A. Nicoll and D. S. Bredt (2003). "Functional studies and distribution define a family of transmembrane AMPA receptor regulatory proteins." J Cell Biol **161**(4): 805-16.

Tomita, S., M. Fukata, R. A. Nicoll and D. S. Bredt (2004). "Dynamic interaction of stargazin-like TARPs with cycling AMPA receptors at synapses." Science **303**(5663): 1508-11.

Tomlinson, M. L., R. A. Field and G. N. Wheeler (2005). "Xenopus as a model organism in developmental chemical genetic screens." Mol Biosyst **1**(3): 223-8.

Traboulsi, E. (1999). Genetic Diseases of the Eye. New York, Oxford University Press.

Trombetta, E. S. (2003). "The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis." Glycobiology **13**(9): 77R-91R.

Troy, T.-C., Y. Li, L. O'Malley and K. Turksen (2007). "The temporal and spatial expression of Claudins in epidermal development and the accelerated program of epidermal differentiation in K14-CaSR transgenic mice." Gene Expression Patterns **7**(4): 423-430.

Turnbull, J. E. and R. A. Field (2007). "Emerging glycomics technologies." Nat Chem Biol **3**(2): 74-7.

Tusnady, G. E. and I. Simon (2001). "The HMMTOP transmembrane topology prediction server." Bioinformatics **17**(9): 849-50.

Tzoulaki, I., I. M. White and I. M. Hanson (2005). "PAX6 mutations: genotype-phenotype correlations." BMC Genet **6**(1): 27.

Umeda, K., J. Ikenouchi, S. Katahira-Tayama, K. Furuse, H. Sasaki, M. Nakayama, T. Matsui, S. Tsukita, M. Furuse and S. Tsukita (2006). "ZO-1 and ZO-2 independently determine where claudins are polymerized in tight-junction strand formation." Cell **126**(4): 741-54.

Vagin, O., J. A. Kraut and G. Sachs (2009). "Role of N-glycosylation in trafficking of apical membrane proteins in epithelia." Am J Physiol Renal Physiol **296**(3): F459-69.

van Beest, M., J. H. Robben, P. J. Savelkoul, G. Hendriks, M. A. Devonald, I. B. Konings, A. K. Lagendijk, F. Karet and P. M. Deen (2006). "Polarisation, key to good localisation." Biochim Biophys Acta **1758**(8): 1126-33.

Van Itallie, C., C. Rahner and J. M. Anderson (2001). "Regulated expression of claudin-4 decreases paracellular conductance through a selective decrease in sodium permeability." J Clin Invest **107**(10): 1319-27.

Van Itallie, C. M. and J. M. Anderson (2006). "Claudins and epithelial paracellular transport." Annu Rev Physiol **68**: 403-29.

Van Itallie, C. M., O. R. Colegio and J. M. Anderson (2004). "The cytoplasmic tails of claudins can influence tight junction barrier properties through effects on protein stability." J Membr Biol **199**(1): 29-38.

Van Itallie, C. M., T. M. Gambling, J. L. Carson and J. M. Anderson (2005). "Palmitoylation of claudins is required for efficient tight-junction localization." J Cell Sci **118**(Pt 7): 1427-36.

Van Itallie, C. M., J. Holmes, A. Bridges, J. L. Gookin, M. R. Coccaro, W. Proctor, O. R. Colegio and J. M. Anderson (2008). "The density of small tight junction pores varies among cell types and is increased by expression of claudin-2." J Cell Sci **121**(Pt 3): 298-305.

van Raamsdonk, C. D. and S. M. Tilghman (2000). "Dosage requirement and allelic expression of PAX6 during lens placode formation." Development **127**(24): 5439-48.

Vanita, V., H. C. Hennies, D. Singh, P. Nurnberg, K. Sperling and J. R. Singh (2006). "A novel mutation in GJA8 associated with autosomal dominant congenital cataract in a family of Indian origin." Mol Vis **12**: 1217-22.

Vanita, V., J. R. Singh, D. Singh, R. Varon and K. Sperling (2008). "A mutation in GJA8 (p.P88Q) is associated with "balloon-like" cataract with Y-sutural opacities in a family of Indian origin." Mol Vis **14**: 1171-5.

Varadaraj, K., S. S. Kumari and R. T. Mathias (2007). "Functional expression of aquaporins in embryonic, postnatal, and adult mouse lenses." Dev Dyn **236**(5): 1319-28.

Varadaraj, K., S. S. Kumari and R. T. Mathias (2010). "Transgenic expression of AQP1 in the fiber cells of AQP0 knockout mouse: effects on lens transparency." Exp Eye Res **91**(3): 393-404.

Vassilakos, A., M. Michalak, M. A. Lehrman and D. B. Williams (1998). "Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin." Biochemistry **37**(10): 3480-90.

Verma, A. S. and D. R. Fitzpatrick (2007). "Anophthalmia and microphthalmia." Orphanet J Rare Dis **2**: 47.

von Heijne, G. (1992). "Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule." J Mol Biol **225**(2): 487-94.

Voronina, V. A., E. A. Kozhemyakina, C. M. O'Kernick, N. D. Kahn, S. L. Wenger, J. V. Linberg, A. S. Schneider and P. H. Mathers (2004). "Mutations in the human RAX homeobox gene in a patient with anophthalmia and sclerocornea." Hum Mol Genet **13**(3): 315-22.

Vrensen, G. F., J. Graw and A. De Wolf (1991). "Nuclear breakdown during terminal differentiation of primary lens fibres in mice: a transmission electron microscopic study." Exp Eye Res **52**(6): 647-59.

Wadehra, M., A. Forbes, N. Pushkarna, L. Goodglick, L. K. Gordon, C. J. Williams and J. Braun (2005). "Epithelial membrane protein-2 regulates surface expression of alphavbeta3 integrin in the endometrium." Dev Biol **287**(2): 336-45.

Wadehra, M., L. Goodglick and J. Braun (2004). "The tetraspan protein EMP2 modulates the surface expression of caveolins and glycosylphosphatidyl inositol-linked proteins." Mol Biol Cell **15**(5): 2073-83.

Walker, C. S., A. C. Conner, D. R. Poyner and D. L. Hay (2010). "Regulation of signal transduction by calcitonin gene-related peptide receptors." Trends Pharmacol Sci.

Walker, J. L. and A. S. Menko (1999). "alpha6 Integrin is regulated with lens cell differentiation by linkage to the cytoskeleton and isoform switching." Dev Biol **210**(2): 497-511.

Walker, J. L., L. Zhang, J. Zhou, M. J. Woolkalis and A. S. Menko (2002). "Role for alpha 6 integrin during lens development: Evidence for signaling through IGF-1R and ERK." Dev Dyn **223**(2): 273-84.

Wang, K., B. Wang, J. Wang, S. Zhou, B. Yun, P. Suo, J. Cheng, X. Ma and S. Zhu (2009a). "A novel GJA8 mutation (p.I31T) causing autosomal dominant congenital cataract in a Chinese family." Mol Vis **15**: 2813-20.

Wang, Q., J. W. McAvoy and F. J. Lovicu (2010a). "Growth factor signaling in vitreous humour-induced lens fiber differentiation." Invest Ophthalmol Vis Sci.

Wang, T., Q. Gao, P. Nie and C. J. Secombes (2010b). "Identification of suppressor of cytokine signalling (SOCS) 6, 7, 9 and CISH in rainbow trout Oncorhynchus mykiss and analysis of their expression in relation to other known trout SOCS." Fish Shellfish Immunol **29**(4): 656-67.

Wang, W., J. Jiang, Y. Zhu, J. Li, C. Jin, X. Shentu and K. Yao (2010c). "A novel mutation in the major intrinsic protein (MIP) associated with autosomal dominant congenital cataracts in a Chinese family." Mol Vis **16**: 534-9.

Wang, W. L., Q. Li, J. Xu and A. Cvekl (2010d). "Lens fiber cell differentiation and denucleation are disrupted through expression of the N-terminal nuclear receptor box of NCOA6 and result in p53-dependent and p53-independent apoptosis." Mol Biol Cell **21**(14): 2453-68.

Wang, Y., D. Mukhopadhyay, S. Mathew, T. Hasebe, R. A. Heimeier, Y. Azuma, N. Kolli, Y. B. Shi, K. D. Wilkinson and M. Dasso (2009b). "Identification and developmental expression of Xenopus laevis SUMO proteases." PLoS One **4**(12): e8462.

Watson, G. W. and U. P. Andley (2010). "Activation of the unfolded protein response by a cataract-associated alphaA-crystallin mutation." Biochem Biophys Res Commun **401**(2): 192-6.

Weber, G. F. and A. S. Menko (2006). "Actin filament organization regulates the induction of lens cell differentiation and survival." Dev Biol **295**(2): 714-29.

Wederell, E. D. and R. U. de Iongh (2006). "Extracellular matrix and integrin signaling in lens development and cataract." Semin Cell Dev Biol **17**(6): 759-76.

Weerapana, E. and B. Imperiali (2006). "Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems." Glycobiology **16**(6): 91R-101R.

Wen, H., D. D. Watry, M. C. Marcondes and H. S. Fox (2004). "Selective decrease in paracellular conductance of tight junctions: role of the first extracellular domain of claudin-5." Mol Cell Biol **24**(19): 8408-17.

West-Mays, J. A., B. M. Coyle, J. Piatigorsky, S. Papagiotas and D. Libby (2002). "Ectopic expression of AP-2alpha transcription factor in the lens disrupts fiber cell differentiation." Dev Biol **245**(1): 13-27.

West-Mays, J. A., J. Zhang, T. Nottoli, S. Hagopian-Donaldson, D. Libby, K. J. Strissel and T. Williams (1999). "AP-2alpha transcription factor is required for early morphogenesis of the lens vesicle." Dev Biol **206**(1): 46-62.

White, T., T. Lu, R. Metlapally, J. Katowitz, F. Kherani, T. Y. Wang, K. N. Tran-Viet and T. L. Young (2008). "Identification of STRA6 and SKI sequence variants in patients with anophthalmia/microphthalmia." Mol Vis **14**: 2458-65.

White, T. W., R. Bruzzone, D. A. Goodenough and D. L. Paul (1992). "Mouse Cx50, a functional member of the connexin family of gap junction proteins, is the lens fiber protein MP70." Mol Biol Cell **3**(7): 711-20.

White, T. W., Y. Gao, L. Li, C. Sellitto and M. Srinivas (2007). "Optimal lens epithelial cell proliferation is dependent on the connexin isoform providing gap junctional coupling." Invest Ophthalmol Vis Sci **48**(12): 5630-7.

White, T. W., D. A. Goodenough and D. L. Paul (1998). "Targeted ablation of connexin50 in mice results in microphthalmia and zonular pulverulent cataracts." J Cell Biol **143**(3): 815-25.

Wigle, J. T., K. Chowdhury, P. Gruss and G. Oliver (1999). "Prox1 function is crucial for mouse lens-fibre elongation." Nat Genet **21**(3): 318-22.

Wolven, A., H. Okamura, Y. Rosenblatt and M. D. Resh (1997). "Palmitoylation of p59fyn is reversible and sufficient for plasma membrane association." Mol Biol Cell **8**(6): 1159-73.

Xi, J. H., F. Bai, J. Gross, R. R. Townsend, A. S. Menko and U. P. Andley (2008). "Mechanism of small heat shock protein function in vivo: a knock-in mouse model demonstrates that the R49C mutation in alpha A-crystallin enhances protein insolubility and cell death." J Biol Chem **283**(9): 5801-14.

Xi, J. H., F. Bai, R. McGaha and U. P. Andley (2006). "Alpha-crystallin expression affects microtubule assembly and prevents their aggregation." Faseb J **20**(7): 846-57.

Xia, C. H., H. Liu, B. Chang, C. Cheng, D. Cheung, M. Wang, Q. Huang, J. Horwitz and X. Gong (2006). "Arginine 54 and Tyrosine 118 residues of {alpha}A-crystallin are crucial for lens formation and transparency." Invest Ophthalmol Vis Sci **47**(7): 3004-10.

Yamauchi, K., T. Rai, K. Kobayashi, E. Sohara, T. Suzuki, T. Itoh, S. Suda, A. Hayama, S. Sasaki and S. Uchida (2004). "Disease-causing mutant WNK4 increases paracellular chloride permeability and phosphorylates claudins." Proc Natl Acad Sci U S A **101**(13): 4690-4.

Yardley, J., B. P. Leroy, N. Hart-Holden, B. A. Lafaut, B. Loeys, L. M. Messiaen, R. Perveen, M. A. Reddy, S. S. Bhattacharya, E. Traboulsi, D. Baralle, J. J. De Laey, B. Puech, P. Kestelyn, A. T. Moore, F. D. Manson and G. C. Black (2004). "Mutations of VMD2 splicing regulators cause nanophthalmos and autosomal dominant vitreoretinochoroidopathy (ADVIRC)." Invest Ophthalmol Vis Sci **45**(10): 3683-9.

Yeaman, C., A. H. Le Gall, A. N. Baldwin, L. Monlauzeur, A. Le Bivic and E. Rodriguez-Boulan (1997). "The O-glycosylated stalk domain is required for apical sorting of neurotrophin receptors in polarized MDCK cells." J Cell Biol **139**(4): 929-40.

Yen, H. C., Q. Xu, D. M. Chou, Z. Zhao and S. J. Elledge (2008). "Global protein stability profiling in mammalian cells." Science **322**(5903): 918-23.

Yokoyama, Y., K. Narahara, K. Tsuji, S. Ninomiya and Y. Seino (1992). "Autosomal dominant congenital cataract and microphthalmia associated with a familial t(2;16) translocation." Hum Genet **90**(1-2): 177-8.

Yonemura, S., M. Itoh, A. Nagafuchi and S. Tsukita (1995). "Cell-to-cell adherens junction

formation and actin filament organization: similarities and differences between non-polarized fibroblasts and polarized epithelial cells." J Cell Sci **108 (Pt 1)**: 127-42.

Yu, G. Y., K. J. Lee, L. Gao and M. M. Lai (2006). "Palmitoylation and polymerization of hepatitis C virus NS4B protein." J Virol **80**(12): 6013-23.

Yu, T. and E. Dahan (2009). "Bilateral uneven cataracts in children: amblyopia management by sequential intraocular lens implantation." Eye (Lond) **23**(6): 1451-5.

Zampighi, G. A., S. Eskandari and M. Kreman (2000). "Epithelial organization of the mammalian lens." Exp Eye Res **71**(4): 415-35.

Zandy, A. J. and S. Bassnett (2007). "Proteolytic mechanisms underlying mitochondrial degradation in the ocular lens." Invest Ophthalmol Vis Sci **48**(1): 293-302.

Zandy, A. J., S. Lakhani, T. Zheng, R. A. Flavell and S. Bassnett (2005). "Role of the executioner caspases during lens development." J Biol Chem **280**(34): 30263-72.

Zapun, A., S. M. Petrescu, P. M. Rudd, R. A. Dwek, D. Y. Thomas and J. J. Bergeron (1997). "Conformation-independent binding of monoglucosylated ribonuclease B to calnexin." Cell **88**(1): 29-38.

Zelenka, P. S. (2004). "Regulation of cell adhesion and migration in lens development." Int J Dev Biol **48**(8-9): 857-65.

Zenteno, J. C., M. E. Morales, V. Moran-Barroso and A. Sanchez-Navarro (2005). "CRYGD gene analysis in a family with autosomal dominant congenital cataract: evidence for molecular homogeneity and intrafamilial clinical heterogeneity in aculeiform cataract." Mol Vis **11**: 438-42.

Zeynali, B. and K. E. Dixon (1998). "Effects of retinoic acid on the endoderm in Xenopus embryos." Dev Genes Evol **208**(6): 318-26.

Zhang, J. J. and T. J. Jacob (1994). "A new approach to measuring transepithelial potentials in the bovine lens reveals a chloride-dependent component." Exp Physiol **79**(5): 741-53.

Zhang, L., L. Gao, Z. Li, W. Qin, W. Gao, X. Cui, G. Feng, S. Fu, L. He and P. Liu (2006). "Progressive sutural cataract associated with a BFSP2 mutation in a Chinese family." Mol Vis **12**: 1626-31.

Zhang, L. Y., G. H. Yam, P. O. Tam, R. Y. Lai, D. S. Lam, C. P. Pang and D. S. Fan (2009). "An alphaA-crystallin gene mutation, Arg12Cys, causing inherited cataract-microcornea exhibits an altered heat-shock response." Mol Vis **15**: 1127-38.

Zhang, P., C. Wong, R. A. DePinho, J. W. Harper and S. J. Elledge (1998). "Cooperation between the Cdk inhibitors p27(KIP1) and p57(KIP2) in the control of tissue growth and development." Genes Dev **12**(20): 3162-7.

Zhang, Q., X. Guo, X. Xiao, J. Yi, X. Jia and J. F. Hejtmancik (2004). "Clinical description and genome wide linkage study of Y-sutural cataract and myopia in a Chinese family." Mol Vis **10**: 890-900.

Zhang, Y., D. Burgess, P. A. Overbeek and V. Govindarajan (2008). "Dominant inhibition of lens placode formation in mice." Dev Biol **323**(1): 53-63.

Zhao, H., T. Yang, B. P. Madakashira, C. A. Thiels, C. A. Bechtle, C. M. Garcia, H. Zhang, K. Yu, D. M. Ornitz, D. C. Beebe and M. L. Robinson (2008). "Fibroblast growth factor receptor signaling is essential for lens fiber cell differentiation." Dev Biol **318**(2): 276-88.

Zhou, J., J. Hu and H. Guan (2010). "The association between copy number variations in glutathione S-transferase M1 and T1 and age-related cataract in a Han Chinese population." Invest Ophthalmol Vis Sci **51**(8): 3924-8.

Zhou, L., T. Chen and R. L. Church (2002). "Temporal expression of three mouse lens fiber cell membrane protein genes during early development." Mol Vis **8**: 143-8.

Zhou, M., J. Leiberman, J. Xu and R. M. Lavker (2006). "A hierarchy of proliferative cells exists in mouse lens epithelium: implications for lens maintenance." Invest Ophthalmol Vis Sci **47**(7): 2997-3003.

Zuercher, J., J. Neidhardt, I. Magyar, S. Labs, A. T. Moore, F. C. Tanner, N. Waseem, D. F. Schorderet, F. L. Munier, S. Bhattacharya, W. Berger and B. Kloeckener-Gruissem (2010). "Alterations of the 5'untranslated region of SLC16A12 lead to age-related cataract." Invest Ophthalmol Vis Sci **51**(7): 3354-61.