Australian National University

RESEARCH SCHOOL OF BIOLOGY

Identification and Characterisation of Genes which are Critical for Murine Gastrulation

Alaa Alzahrani

A thesis submitted for the degree of Doctor of Philosophy of the Australian National University

June 2017

The results presented in this thesis are, except where otherwise acknowledged, my own original work.

Alaa Alzahrani

Work completed by others included in this thesis

Figure 1.1 adapted from Arkell *et al.,* (2012)

Figure 3.1 adapted from Ali *et al.,* (2012)

Figure 3.2, 3.3 adapted from Houtmeyers *et al.,* (2013)

Figure 3.5 reproduced from Elms *et al.,* (2004)

Figure 3.6, 3.8, 5.2 adapted from R.Arkell.

Figure 3.7 adapted from Koula Diamand

Panels (A-G) were adapted from Ahmed *et al.,* (2013). Panels (H-M) were adapted from Ware *et al.,* (2006b) in figure 3.10

Figure 3.11 adapted from Arkell *et al.,* (2013)

Panels (A-C) were adapted from Fossat *et al.,* (2011) in figure 3.12

SEM Experiments and taking photos performed with K. Barratt.

Figure 4.2 adapted from Petkov *et al.,* (2004)

The whole exome sequencing and the single nucleotide variants were performed at Australian Phenomics Genotyping Core Facility (APF, Australian National University).

Figure 5.1 adapted from iTranscription

Figure 5.3 reproduced from MacPhee *et al.,* (2000)

"Read in the name of your Lord who created"

"Created man, out of a (mere) clot of congealed blood"

"Read, and your Lord is the most Generous"

"Who taught by the pen"

"Taught man that which he knew not."

(Holy Qur'an 96:1-5)

Acknowledgments

I am eternally grateful to the Almighty for enabling to endure and complete this seemingly insurmountable PhD journey. Completing this is only made possible through the unwavering support and guidance of several individuals over the years.

First and foremost, I am grateful to my supervisor, Associate Professor Ruth Arkell for the opportunity to conduct research under her supervision and for providing invaluable advice over the last five and half years. Her assistance and feedback during the preparation of seminars and the writing of this thesis are highly appreciated. Under her supervision, I have gained invaluable skills which will be useful in my future undertakings as a research scientist.

My gratitude is also extended to the members of the Arkell lab past and present (Radiya Ali, Nicole Thomsen, Laura Woad, Jerry Ahmed, Koula Diamand, Kristen Barratt, and Hannah Jones); they have contributed immensely to my development as a research scientist. I would particularly like to thank Koula and Kristen for their assistance in the lab and friendship.

To the two most precious souls in the whole world, my Baba and Mama: thank you for allowing me exist in this world; and for providing me with your wonderful genetics. I owe to the both of you for me passing gastrulation safely. You have ensured I have everything I ever needed and for being there when I needed you. Moreover, you have taught me numerous life lessons you these made living a gratifying experience. I'm grateful to my Baba for teaching me the value of commitment, hard work and to my Mama for encouraging me to be persistent. To my siblings, a very special thank you for your constant love and encouragement.

This Ph.D. would not have been possible without my mentor, my bestie Rizsa Albarracin. I thank you for inspiring me always and pushing me to press harder.

To Soumya Joseph and Claudia Rodriguez-Delgado, thank you for proofreading my thesis.

A very special mention to the love of my life (Talal), thank you for just loving me purely even in time of crisis. Finally, I want to thank the wonderful 'chocolate' for keeping me happy and in high spirits. **I DID IT!**

Abstract

Gastrulation, a critical stage of embryogenesis, converts a group of relatively undifferentiated cells into the body plan of the growing foetus. Embryonic cells are allocated to one of three definitive germs layers and are organised with respect to three embryonic axes (anterior-posterior, dorsal-ventral and left-right). By the end of gastrulation the primordia of most organs and physiological systems have been established. In mammals, gastrulation occurs just after the embryo implants into the uterine wall. The small size of the mammalian gastrula and its inaccessibility means that, despite its central importance to embryonic development, gastrulation remains one of the most challenging and underexplored areas of mammalian biology. The study of other complex biological processes has been greatly facilitated by the use of genetic screens, which employ a powerful mutagen to introduce mutations into the germline of an organism and then examine descendants for altered biology (i.e. an abnormal phenotype) followed by post-hoc identification of the underlying genetic alteration. The utility of these screens is dependent upon quick, robust assays to identify pheno-deviants and it has long been thought that practical issues prevent this approach being applied to mammalian gastrulation. In this thesis, I examine this notion. In the first part of the thesis, I use phenotype analysis to infer the function of the gene altered in the katun mouse strain which was identified in a screen for mutations that alter embryonic development and was found to have a gastrulation phenotype. I show that the left-right axis defects in this strain are likely caused by dysregulated Wnt signalling at gastrulation which interferes with endoderm development. This confirms that mutants recovered from genetic screens can be used to infer gene function at gastrulation. In the second part of the thesis, I report results form a prototype, genome-wide, screen for recessive mutations that alter gastrulation. The screen was able to overcome the perceived difficulties of small litter sizes, of isolating a large number of gastrulation stage embryos, of distinguishing phenotypes from natural embryonic loss and of working with limiting amounts of genomic DNA. I demonstrate that mutant pedigrees can be identified, they can be maintained

in the absence of genetic markers via test breeding protocols and that low-resolution meiotic mapping can be combined with next generation sequencing technologies to identify the causative mutation. This confirms that gastrulation is amenable to genetic dissection by mutagenesis and I suggest a protocol that could be used to conduct a larger scale, efficient screen to identify the genes that direct mammalian gastrulation.

Table of Contents

List of Figures

List of Tables

Abbreviations

A-P	Posterior and anterior axis
APC	Adenomatous polyposis coli
APF	Australian Phenomics
ATP	Adenosine triphosphate
Bfc	BatFace allele
BI	Buffer one
BIV	Buffer four
BSA	Bovine serum albumin
Chr	Chromosome
CK1	Casein kinase 1
CNS	Central nervous system
Dkk1	Dickkopf1
DNA	Deoxyribonucleic acid
dNTP	di-Nucleotide Triphosphate
dpc	Dyas post coitum
DV	definitive endoderm
D-V	Dorsal and ventral axis
Dvl	Dishevelled protein
E	Glutamic Acid
EDTA	Ethylenediaminetraacetic acid
ENU	N-ethyl-N-nitrosourea
ES	Embryonic Stem Cells
F0	Founder
F1	First filial generation animal
FAM	4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein
FAP	Familial Adenomatous Polyposis
FGF	Fibroblast Growth Factor
FZ	Frizzled
G0	Generation 0

G1	First Generation
G2	Second Generation
G3	Third generation
GSK3	Glycogen synthase kinase 3
HRMA	High resolution melt analysis
K	Lysine protein
K^+	Potassium
Ka	Katun allele
KOMP	Knockout Mouse Project
LB	Bacteria bertani
LPM	Lateral plate mesoderm
L-R	Left and right axis
LRP	Lipoprotein receptor-related protein
L-V	Left ventricle
M6BNFQ	Minor groove binding non-fluorescent quencher
MGI	Mouse Genome Informatics Database
MIT	Microsatellite Marker
Mmu	Mouse (Mus Musculus) chromosome
n	Node
N1	First backcross generation, derived from the F1 animal
N2	Animals generated from specific intercross
Na	Sodium
NTC	No template control
Opa	odd-paired
PBS	Phosphate buffered saline
PCP	Planar cell polarity pathway
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PK	Protein K
PBS	Phosphate buffered saline
PTC	Premature termination codon

R-A	Right atrium
SACC	Subclass A C-terminal Conserved
SANC	Subclass A N-terminal Conserved
SEM	Scanning electronic microscopy
SNP	Single Nucleotide Polymorphism
SLT	Specific location sites
T	Thymine
TADs	Terminal transactivation domains
TCF	T-Cell Factor
TCF/LEF	Mediated transcription factor/ lymphoid enhancer binding factor
TD	Touchdown
TF	Transcription factors
Tm	Temperature
Tris	*Tris*(hydroxymethyl)aminomethane
UTR	Untranslated region
V	Valine
VIC	4,7,2'-trichloro-7
WMISH	Whole mount in situ hybridiSation
Y	Tyrosime
ZF	Zinc finger
ZFD	zinc finger domain
ZFNC	Zinc finger N-terminal conserved region
Zic	zinc finger of the cerebellum
ZOC	Zic-Opa conserved region

Chapter 1: Introduction

"It is not birth, marriage, or death, but gastrulation which is truly the most important time in your life." -Lewis Wolpert

1.1 Gastrulation and its importance in mammalian embryogenesis

One of the most profound biological questions is how an egg develops into a complete animal. This process is known as *embryogenesis* and it occurs when a fertilised egg, a single cell, eventually gives rise to a complex multicellular organism. Embryogenesis is orchestrated by genetics and cell communication via signalling pathways. Pioneering experiments demonstrated that this communication enables cells to differentiate into various cell types, which in turn will develop into specialised tissues and organs (Ingham, 1988, Harland *et al.*, 1997). For example, the initial diploid cell formed by fertilisation, the zygote, undergoes rapid cell division to produce a large number of totipotent cells. It is essential that all of these cells are programmed correctly during development to make the appropriate changes in their shape and position in order to form a specific part of the body. The core of this phenomenon, cell fate, is affected by the activities of genes encoding proteins and different RNAs that influence cell identity and control cell proliferation, migration, differentiation, and/or programmed cell death (Whitman *et al.*, 1989). Therefore, to ensure that each cell will fulfil its purpose, genes must be expressed at the correct time and place (McLaren, 1976, Hart *et al.*, 1985).

Abnormalities in the expression of genes that control embryonic development can lead to perturbations in critical processes, resulting in either severe developmental defects or embryo death (Kasarskis *et al.*, 1998, Garcia-Garcia *et al.*, 2005). Thus, it is critical to identify the genes and their genetic pathways that regulate embryogenesis to understand why these defects accrue. Importantly, every embryonic process is highly

dependent upon the success of the previous step. Consequently, minor aberrancies at an early stage are amplified as embryonic development progresses. This is illustrated by gastrulation.

1.1.1 Pre-gastrulation stage

During zygote formation, it takes multiple rounds of cleavages to establish two distinct cell lineages: the trophectoderm (TE) and the inner cell mass (ICM), which form the blastocyst between 3.5 and 4.5 days post coitum (dpc) (Figure 1.1). During implantation the TE gives rise to tissues in the placenta by invading the mother's uterine lining. After implantation, the ICM gives rise to the three main germ layers- the ectoderm, mesoderm and endoderm- ultimately generating all of the tissues in the animal body (Tam *et al.*, 2007). This is termed gastrulation. From this process a small number of cells can produce a whole animal within three days (in the mouse) (Tam *et al.*, 2006).

Positioning of the progenitor cell populations in the embryo is fundamental to establishing tissue lineage. These progenitors later delineate embryonic axes and coordinate morphogenetic cell and tissue movements (Tam *et al.*, 2007). The blastocyst is built as a vesicular structure (Figure 1.1), where the blastocyst epithelium encloses a cavity called the blastocoel. A cluster of cells attached to the wall on one side of the blastocoel constitutes the ICM, which further segregated into the epiblast (which gives rise to the entire embryo) and some components of the fetal extraembryonic membranes and the primitive endoderm (which lines the luminal surface of the cluster of epiblast cells).

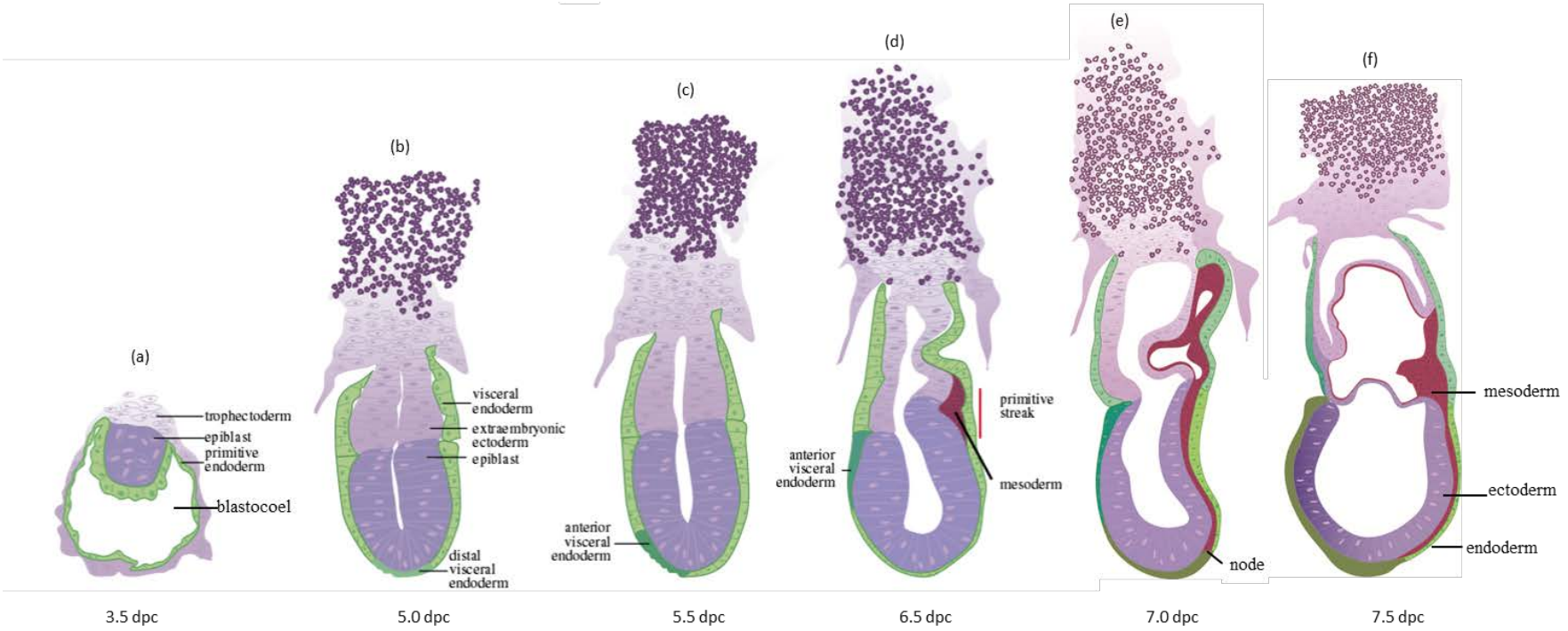

Figure 1.1: Formation of three primary germ layers in the gastrulating mouse embryo. (a) Blastocyst containing an inner cell mass comprising the epiblast and the primitive endoderm. (b, c) Egg cylinder embryo at 5.0 days post-coitum (dpc) with distal visceral endoderm, and 5.5 dpc with anterior visceral endoderm. (d) At 6.5 dpc a change in the shape and motility of the cells occurs at the posterior side of the embryo (cells in red). (e, f) At 7.0 dpc, these cells emerge distally, forming the node structure and the mesoderm (a new germ layer forms between the ectoderm and endoderm) by 7.5 dpc. Figure adapted from Arkell *et al.*, (2012).

The primitive endoderm forms by the time the embryo implants in the uterus at 4.5 dpc (Tam *et al.*, 2007). It is formed from the blastocoel surface of the ICM and gives rise to both mural trophectoderm and the visceral endoderm, which later forms the epithelium on the surface of the epiblast and extra embryonic ectoderm (EXE) (Tam *et al.*, 2007). Gene expression studies showed that GATA-binding protein 6 (GATA6) positive cells are the progenitors of the primitive endoderm and that *Nanog* expressing cells form the epiblast (Arman *et al.*, 1998, Mesnard *et al.*, 2004). In addition, it was found that growth factor receptor bound protein 2 (GRB2), a receptor-binding adaptor protein and member of mitogen-activated protein (MAP) kinase signalling pathways, is required for the expression of *Gata6* in the ICM and the formation of the primitive endoderm (Arman *et al.*, 1998, Chazaud *et al.*, 2006). This indicates a role for the receptor tyrosine of kinase signalling in the early stages of primitive endoderm formation. However, the precise mechanism by which cells of the primitive endoderm and other cell types resolve their fate remains unclear (Tam *et al.*, 2007).

1.1.2 Axis establishment

By 5.0 dpc the mouse embryo is called the egg cylinder (Figure 1.1 b). The prospective dorsal-ventral (D-V) axis becomes apparent, with the proamniotic cavity surface of the epiblast corresponding to the dorsal side of the embryo and the outer surface of the visceral endoderm corresponding to the ventral side of the embryo (Mohamed *et al.*, 2004). Asymmetric gene expression has been observed in the visceral endoderm of the mouse embryo at the distal visceral endoderm stage when *Nodal* antagonists *Cer1* (Cerberus 1 homologue) and *Lefty1* are highly expressed on the anterior side of the definitive endoderm (DVE) (Tam *et al.*, 2007). Similarly, *Wnt3* is expressed asymmetrically where the epiblast meets the extraembryonic endoderm. Recent studies showed that the distal visceral endoderm contributes mostly to the proximal portion of

anterior visceral endoderm as well as the visceral endoderm in lateral regions, but not to entire anterior visceral endoderm (AVE). The extensive movement of the visceral endoderm cells is initiated by the distal visceral endoderm at 5.5 dpc. Experiments showed that when the definitive endoderm is ablated by genetic manipulation, AVE cells formed at the distal end but fail to migrate (Takaoka *et al.*, 2011). Thus, the definitive endoderm guides the AVE to the anterior side of the embryo by initiating extensive movement of visceral endoderm cells (Takaoka *et al.*, 2011).

Before the formation of anterior-posterior (A-P) axis, the visceral endoderm becomes polarised in the proximal-distal direction. Between the pre-distal visceral endoderm and the definitive endoderm there is differentiation in gene expression. *Nodal* signalling, which originates in the epiblast and is transduced by SMAD2 in the extraembryonic tissues, is required for the formation of the DVE (Waldrip *et al.*, 1998, Brennan *et al.*, 2001). There is an inductive interaction between the extra-embryonic and embryonic tissues which is essential for lineage specification and embryonic patterning. These are mediated by both WNT and TGFβ signalling pathways (Tam *et al.*, 2006). In the mouse, the initial process to develop A-P axis is poorly understood, although several recent reports showed that early A-P molecular asymmetry could be tracked back to the pre-implantation stage (Stuckey *et al.*, 2011). Many genes determine the A-P axis formation in the pre-gastrulation stage. For example, endoderm-associated antigen (VE-1) is one of the earliest markers of the A-P axis at 5.5 dpc and can be detected on one side of the visceral endoderm. It is later localised to the anterior visceral endoderm (Rosenquist and Martin, 1995). Similarly, *Wnt3* expression, which will later lead to the formation of the primitive streak at gastrulation, is detected at the DVE (Tam *et al.*, 2006).

The expression pattern of genes such as *Evx1, Fgf8, Goosecoid (Gsc)* and *Nodal* are

posteriorly restricted in pre-gastrula mouse embryos. Prior to the primitive streak formation, both *Evx1* and *Fgf8* are expressed in the posterior epiblast and the extraembryonic ectoderm (Dush and Martin, 1992, Crossley and Martin, 1995). In addition, before gastrulation, *Gsc* expression is detected in the proximal-posterior epiblast, the prospective site of primitive streak formation (Faust *et al.,* 1995). Via a *lacZ* reporter, *Nodal* expression was identified in the presumptive posterior region of the epiblast prior to the formation of the primitive streak (Varlet *et al.*, 1997). As *Nodal* and *Evx1* are required for epiblast development/function, early embryonic lethality prior to gastrulation can occur as result of losing either proteins (Conlon *et al.,* 1991, Spyropoulos and Capecchi, 1994). Conlon *et al.* (1994) suggested that NODAL has a primary role in the induction and maintenance of the primitive streak. In contrest, *Gsc* is dispensable in mouse gastrulation (Rivera-Pérez *et al.,* 1995, Yamada *et al.,* 1995). Similarly, if a ligand or a receptor of the fibroblast growth factor (FGF) family (such as *Fgf8*) is absent the resulting mutant embryos exhibit gastrulation defects. The precise nature of these abnormalities, however are still under investigation (Crossley, 1995, Sun *et al.*, 1999). The observations that discrete A-P restricted expression domains in visceral endoderm and the epiblast exist in the pre-gastrula mouse embryo demonstrate that a molecular A-P asymmetry exists before the initiation of gastrulation (Tam *et al.*, 1997).

1.1.3 Gastrulation stage

Before primitive streak formation at 6.0 dpc, the newly established A-P axis changes its orientation and the polarity of bilateral symmetry as the embryo remodels its shape (Perea-Gomez *et al.*, 2004). Pre-gastrulation studies showed that the A-P axis of the embryo itself shifts to align with the long pass through axis at gastrulation (Mesnard *et al.*, 2004). The A-P axis becomes morphologically identifiable via primitive streak formation at the posterior side of the embryo at 6.5 dpc (Mohamed *et al.*, 2004) (Figure

1.1). Formation of the streak is the first overt morphologically distinct structure that breaks the bilateral symmetry of the embryo (Stuckey *et al.*, 2011). It forms within the epiblast at the extra-embryonic boundary diagonally to the anterior visceral endoderm and in doing this, it defines the future posterior of the embryo and the site of gastrulation at 6.5 dpc (Stuckey *et al.*, 2011). At the early primitive streak stage, Mohamed *et al.*, (2004) showed the Catenin beta-1 (*Ctnnb1*) expression is restricted to the primitive streak. The *Ctnnb1* gene encodes the β-CATENIN protein which is the central protein in the canonical Wnt signalling pathway. However, Maretto (2003) showed that, aside from being present in the primitive streak, *Ctnnb1* expression is also present in the proximal epiblast region. However, confirmation via sectioning is required. The transcriptional activity of β-CATENIN was detected in the primitive streak and the mature node of murine embryo, indicating that WNT signalling precedes formation of the epiblast cells that will go on to form primitive streak (Mohamed *et al.*, 2004). The transcriptional activity is sustained in the primitive streak of later staged embryo (Stuckey *et al.*, 2011).

Prior to gastrulation, *Wnt3* is expressed before gastrulation at the proximal epiblast at the junction between the embryonic and extra-embryonic ectoderm with higher levels detected posteriorly (Liu *et al.*, 1999). By 7.0 dpc the restricted expression of *Wnt3*, *Wnt2b* and *Ctnnb1* in the distal visceral endoderm and/or the posterior epiblast indicate the prospective site of the primitive streak. *Wnt3* plays a role in the formation and maintenance of the primitive streak, and the subsequent production of the mesoderm and definitive endoderm (Liu *et al.*, 1999). *Wnt3*$^{-/-}$ mice developed an egg cylinder shaped embryo but cannot form the primitive streak, mesoderm or node (*Nodal* was not expressed at 7.5 dpc) (Liu *et al.*, 1999). The epiblast continues to multiply in an

undifferentiated state that lacks anterior-posterior neural pattering, whilst the anterior visceral endoderm markers are expressed at the correct level and at the right position (Liu *et al.*, 1999). The expression of *Hex1*, which identifies the AVE as well as the earliest population of anterior definitive endoderm to emerge from the anterior primitive streak (Martinez Barbera *et al.,* 2000; Thomas *et al.,* 1998), is not detected in *Wnt3*$^{-/-}$ embryos at 7.5 dpc, reflecting an undifferentiated ectoderm rather than an anterior neural phenotype. These mutants also expressed *Pou5f1* in the ectoderm, which suggests that it has not differentiated beyond epiblast state (Liu *et al.*, 1999). These mutations illustrate how the regional pattering of the visceral endoderm is independent of the primitive streak, but the subsequent establishment of A-P neural patterning in the ectoderm is dependent on derivatives of the streak.

Studies have shown a role for the AVE in anterior neural pattering. The initial development of AVE occurs normally, but the overlying ectoderm does not acquire an anterior neural identity, which indicates that the AVE alone is not sufficient for anterior neural patterning (Thomas *et al.,* 1996). In the mouse, the requirement of the primitive streak derived anterior neural fate was suggested by experiments that explored germ layers. For example, *Lhx1* expression in AVE and primitive streak-derived anterior mesendoderm is essential for forebrain development. Similarly, β-CATENIN is required at the time of axis formation for correct positioning of the AVE. Although some AVE genes are expressed in embryos lacking β-CATENIN, the expressing cells remain at the distal tip of the embryo and fail to migrate to future anterior regions (Huelsken *et al.*, 2000). Similarly, a null allele of *Dickkopf1* (*Dkk1*), a WNT signalling inhibitor, results in anterior head truncations, suggesting it has a critical role in head formation (Mukhopadhyay *et al.*, 2001).

1.1.4 Node formation

At 7.0 dpc the mid-streak (MS) stage (Figure 1.1) a group of cells located near the anterior end of the primitive streak form the murine organiser structure, the node (Lee *et al.*, 2008). The formation of the node is the first sign of the embryonic left-right (L-R) midline axis establishment. At this stage, the node is invisible, with the outer surface of the embryo covered by squamous endoderm cells. However, morphogenesis of the mouse node becomes visible near the distal tip of the embryo at the late streak stage (Sulik *et al.,* 1994, Lee, 2008). Groups of cells at the anterior side of the embryonic midline with small apical surfaces begin to appear in the node at no bud (OB) stage (Poelmann, 1981, Sulik *et al.,* 1994). These cell clusters are separated by squamous endoderm cells with larger surfaces (Sulik *et al.,* 1994). These clustered cells develop cilia, which are approximately 2 μm at the OB stage and often appear to extend vertically from the surface of the node.

It is only recently that the transition mechanism between these two configurations - multiple small fields of constricted cells separated by endoderm cells and a single field of ventral node cells was explored. It is proposed that clusters of streak-derived, putative node cells insert between the endoderm cells (Poelmann, 1981). The arrangement of these node cells was documented via live imaging experiments following the expression of eGFP into the *Noto* locus (Yamanaka *et al.,* 2007). In the mouse embryo the expression of *Noto* is restricted in both the node and notochord, and the *Noto*-GFP allele recapitulates this pattern (Abdelkhalek *et al.,* 2004). Prior to node formation, *Noto*-GFP^{+} cells exist in the anterior primitive streak (Yamanaka *et al.,* 2007), indicating that these cells adopt a node or trunk notochord fate as they begin to divided from the streak (Yamanaka *et al.,* 2007). Presumably the *Noto*-expressing precursors of the node become mesenchymal as they exit from the primitive streak, but then reform an epithelium (with

the opposite polarity as the epiblast) beneath the endodermal layer, before the endoderm cells move away to reveal the node. Yamanaka *et al.* (2007) reported that prior to the node becoming visible on the ventral surface of the mouse embryo, *Noto*-GFP^{+} cells are already arranged in a sheet beneath the endodermal layer. Lee (2008) documented ciliated cells beneath the endoderm (via both scanning electron microscope (Semenov *et al.* 2000) and confocal microscopy) of the late streak (LS) stage in the mouse embryo. Therefore, the initial patchy appearance of ciliated, columnar cells may reflect the gradual removal of overlying endoderm cells, either by cell death or movement, although the organization of the node precursor cells prior to their complete emergence is unknown. Additionally, it is not known whether the field of node precursor cells change shape as the mature node forms or if the node progenitor cells are present in a single field or several clusters (Lee *et al.*, 2008). Importantly, these observations indicate that the ventral node and the endoderm are somehow coordinated. Consequently, the endoderm cells migrate off of the presumptive ventral node, at the same time that the ventral node cells converge toward the midline and assemble into the teardrop-shaped pit. The node epithelium eventually becomes connected with the adjacent epithelium of the definitive endoderm. The mechanism on what regulates epithelialisation of the ventral node and insertion into the endoderm is not explored and remains unclear (Lee *et al.*, 2008). Additionally, it is yet to be establishing how the signals between these two tissue layers might coordinate these behaviours.

By the late bud (LB) stage, the region of the node is flat and set below neighbouring crown cells. At the early head fold (EHF) stage (~7.75 dpc) the final shape of the node is formed, as well as the ciliated pit cells of the node. By this stage the cilia are 3–4 μm long and become posteriorly polarised over time. Signalling molecules, such as *Nodal* which

is the first molecule expressed at the start of the L-R axis, are expressed in the surrounding crown cells. If these signalling molecules are disrupted, congenital defects can form and be lethal for the embryo (Conlon *et al.*, 1994, Varlet *et al.*, 1997, Shen, 2007). The molecular mechanisms of L-R axis development is explained further in chapter 3.

It is important for ventral node cells, to have their apical surface facing ventrally (the opposite of the epiblast) to establish into a columnar epithelium. During gastrulation life is defined by the events of these morphological processes. By the end of gastrulation, the entire animal body is organised into its proper structure. The three major phases of body axis formation, anterior-posterior (A-P), dorsal-ventral (D-V), and left-right (L-R) are established. In addition, the formation of the three embryonic germ layers- the *endoderm*, *mesoderm*, and *ectoderm*- in response to gene expression organise the entire animal body into its proper structure (Figure 1.1) (Tam *et al.,* 1997). All adult tissues in the mouse are generated from these three germ layers; the ectoderm will eventually form the skin and neural tissues, the mesoderm forms muscle, blood, bone, and connective tissues, and the endoderm gives rise to the gut, intestinal tract, and organs such as the liver (Figure 1.2) (Arkell *et al.*, 2012). Thus, without correct gastrula progression, cells cannot assume their proper positions, leading to deformation and incorrect embryo development.

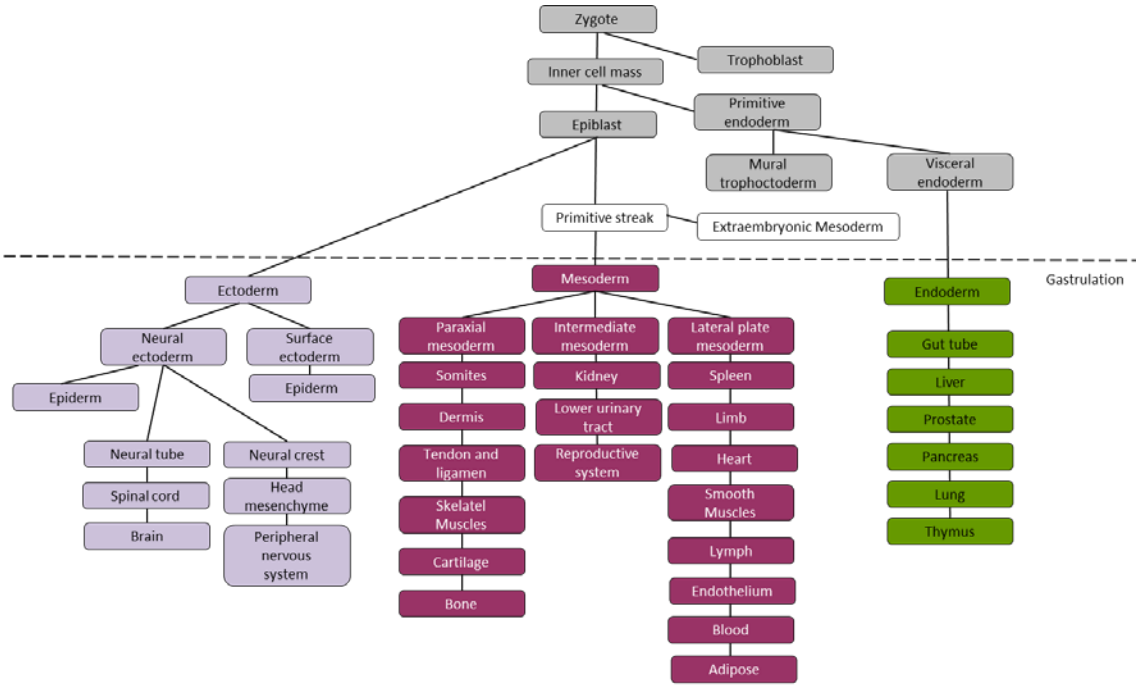

Figure 1.1: The mouse embryo development tree. This describes the mouse embryo developmental stages starting from the early stage (the zygote) through formation of the three germ layers (ectoderm, mesoderm and endoderm) at gastrulation to the ultimate adult tissues.

1.2 Identifying genes involved in mammalian embryogenesis

Genes required for normal gastrulation have previously been identified in studies of mouse genetics (Sun *et al.*, 1999). To generate the mutant phenotypes, researchers evaluated the roles of genes and signalling pathways in directing normal embryonic formation. The results showed that it is not individual gene activity, but rather complex genetic networking that has a crucial role in normal biological processes (Tam *et al.*, 2007). The process of gastrulation is difficult to study, however, because of the small size of the gastrula, making the mouse embryo inaccessible to manipulation during this stage. Additionally, rapid biological changes occur over a very short period time. Thus, our current understanding of the molecular mechanisms driving gastrulation is limited.

Mutagenesis is a powerful tool for determining the function of any gene in any organism: by inactivating or otherwise altering the function of a gene and observing the consequences to the organism, the normal function of the gene and gene product can be inferred. Numerous biological and biomedical studies have focussed on understanding gene function in mammalian organisms, and the mouse has become the most widely used mammalian organism for experimental genetics. Not only can the mouse genome be easily manipulated, but the biological and genetic similarities between mice and humans make it an excellent model for many aspects of human biology, including embryonic development (Denny *et al.*, 2000). Generally, efforts to understand gene function using mutagenesis fall into two categories: reverse genetics and forward genetics (Figure 1.3).

2.1.1 Reverse genetics

Reverse genetics (also known as a genotype-driven strategy), assesses gene function by modifying or disrupting a known gene of interest. The ability to use reverse genetics

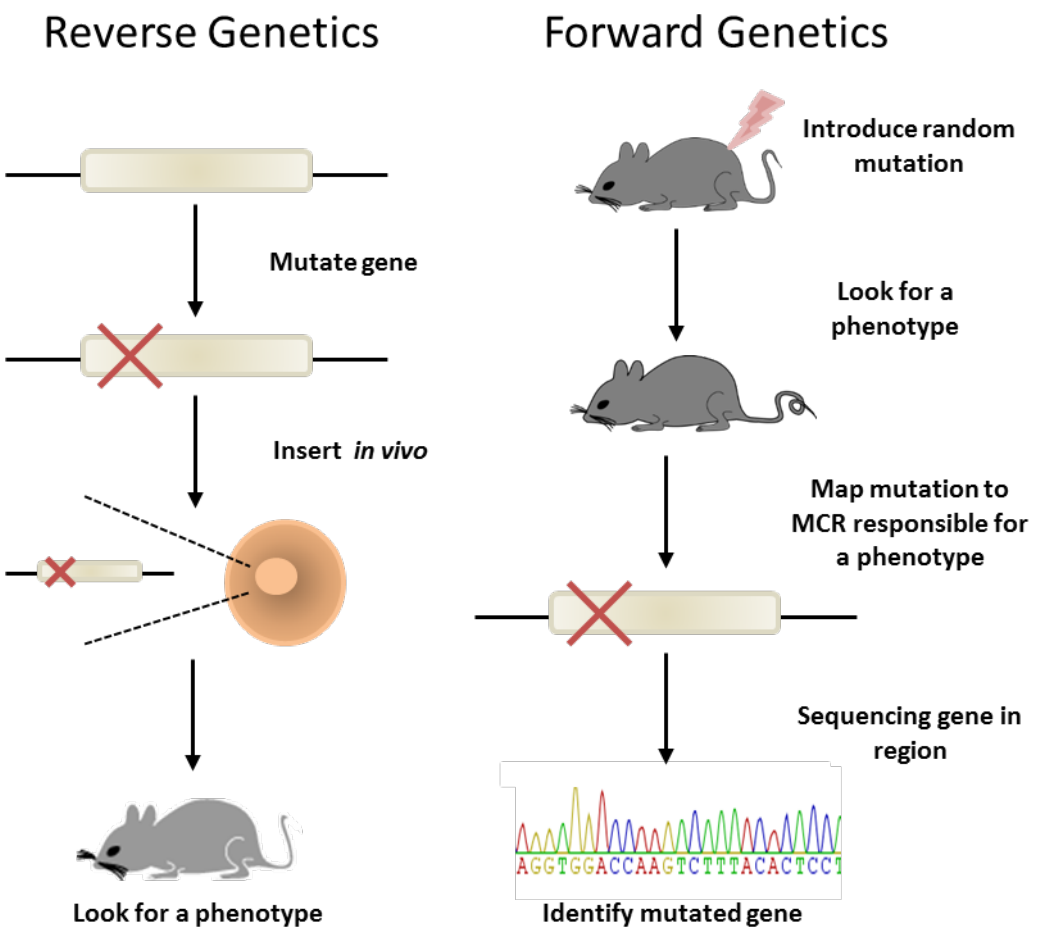

Figure 1.3: Common strategies for investigating gene function. Reverse genetics is used to evaluate the phenotypes resulting from mutation of a specific gene. It starts with selecting and mutating a gene of interest within the germline of an animal. The mutant animal is then assessed for phenotypes. In contrast, forward genetics is used to determine the genetic basis of a phenotype or trait. This method starts with generating random mutations by utilising a powerful mutagen. Next, screening is carried out to identify mutations that cause a phenotype. The mutated gene is then mapped to a minimal critical region (MCR) on a chromosome and the MCR sequenced to identify the gene responsible for the phenotype of interest.

in a mammalian organism was heralded by the discovery of murine embryonic stem (ES) cells (Evans *et al.*, 1981); subsequent studies demonstrated that genes could be specifically modified in ES cells via homologous recombination (Thomas *et al.*, 1987) before incorporation of ES cells back into a mouse embryo (Koller *et al.*, 1989). The ES cells are incorporated into an embryo via injection into the blastocoel cavity of a pre-implantation stage mouse embryo. The embryo is then transferred to the uterus of a pseudo-pregnant surrogate female mouse. The genetically modified ES cells mix with host embryonic cells to form a chimeric embryo. All tissues of the embryo can become chimeric, including the germ cells. When a chimeric mouse is mated with a wild-type animal, some progeny will be derived from the modified ES cells resulting in a mouse in which all cells carry the introduced mutation (Thompson *et al.*, 1989). This process is known as germline transmission. Gene function can ultimately be inferred by assessing the phenotypes of mice that are heterozygous or homozygous for the mutation.

The ability to target desired mutations to nearly any region of the mouse genome enabled reverse genetics to become the mainstay of mouse mutagenesis efforts for many years (Capecchi, 1989, Magin *et al.*, 1992, Bradley *et al.*, 1996, Muller, 1999, Capecchi, 2005, Bouabe *et al.*, 2013). Up until recently, all reverse genetic strategies in the mouse relied up the culture of, and genome manipulation in, ES cells. This powerful strategy does, however, have some drawbacks. Firstly, the germline transmission of ES cells through mice is labour-intensive and time-consuming. Moreover, extended culture of ES cells is associated with gross chromosomal rearrangements in ES cells. For example, compound trisomy 8 and 11 is often detected in ES cells, and the proportion of aneuploid cells increases with passage number. Aneuploidy does not appear to impact the ability of the cells to differentiate (Mallon *et al.,* 2004) but does, however,

affect the rate of germline transmission, with two recent studies demonstrating that the efficiency of germline transmission decreases markedly when ES cultures contain less than 50% euploid cells (Mallon *et al.,* 2004, Gaztelumend and Nogues, 2014)

More recently, the ability to directly modify genomic DNA within a zygote via CRISPR/Cas9-mediated genome editing has revolutionised reverse genetics. Circumventing the need for ES cells means that reverse genetics can now efficiently be applied to any organism for which the genome is well-characterised and a method exists for transduction of exogenous nucleic acids. Researchers have developed CRISPR/Cas9, an RNA-guided targeted genome technique, as a successful tool for introducing a wide range of genomic modifications in several cells and organisms (Doudna *et al.*, 2014). Moreover, Wang *et al.* (2013) showed that a one-cell mouse embryo (zygote) can be modified by CRISPR injection and this technique can be applied to produce mice carrying biallelic mutations in a specific gene, circumventing the need for homologous recombination in ES cells for targeting modification of the mouse genome.

CRISPR/Cas9 is a simple programming tool in which CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is associated with nuclease CRISPR-associated protein 9 (Cas9) to introduce site-specific DNA double-strand breaks, which are typically repaired by non-homologous end-joining or homology-directed repair with high accuracy and efficiency (Deriano *et al.*, 2013). Double-strand breaks occur via a synthetic single guide RNA (Jinek *et al.*, 2012), which guides Cas9 to a specific location in the genome and vie this method, the generation of insertion or deletion (indels) mutations can result in a loss-of-function allele (Jinek *et al.*, 2012). This technique has revealed the functional organisation of the genome at the systems level and established linkages between genetic variations and biological phenotypes. A major advantage of CRISPR is

the ability to design mutations by targeting/reporter specific loci to create knock-in of human single-nucleotide variants (SNV) (He *et al.*, 2016).

This method is only applicable when the function of the targeted gene is already known, however, this remains difficult to predict. For example, targets for reverse genetics are often selected because of previous work such as the knowledge of the gene expression pattern. Gene expression patterns, while suggestive, are not sufficient to distinguish sites of non-redundant expression. As such, reverse genetics often produces unexpected findings such as a different effect on the organism than expected from the genes function. Therefore, given that there are approximately 25,000 protein-coding genes in the mouse genome for which gene function must be described, reverse genetics needs to be complemented with other approaches (Rathkolb *et al.*, 2000).

2.1.2 Forward genetics

Forward genetics (also known as a phenotype-driven strategy) begins with an abnormal phenotype and subsequently identifies the genetic mutation causing the phenotype (Figure 1.3). Generally, a series of random mutations are introduced across the genome and the resulting animals are screened for a phenotype of interest (Rathkolb *et al.*, 2000). The major advantage of forward genetics is that it makes no *a priori* assumptions about gene function (Brown *et al.*, 2001). When applied systematically, forward genetics is theoretically capable of dissecting the genetic basis of any complex biological process in various organisms. This was dramatically demonstrated in simple organisms such as yeast and nematodes. For example, a forward genetic screen in the yeast *Schizosaccharomyces pombe* identified 14 unlinked genes that are essential for cell cycle progression (Nurse *et al.*, 1976). Similarly, large-scale screening of the nematode

Caenorhabditis elegans led to the characterisation of ~100 genes that generated viable, visible phenotypes (Brenner, 1974).

Despite the obvious utility of genetic screens for lethal mutations were, however, considered daunting because of the need to maintain each lethal mutation as a heterozygous strain. Additionally, a large proportion of genes and/or biological processes are essential for embryogenesis, suggesting that large numbers of mutations must be screened to recover multiple mutations giving rise to the same phenotype. Thus, it was several years before the Heidelberg mutagenesis screen generated thousands of lethal phenotypes in the fruit fly *Drosophila melanogaster* and revolutionised the understanding of how genes control developmental patterning (Nusslein-Volhard *et al.*, 1980, Wieschaus *et al.*, 1984). The success of this approach was increased by the availability of balancer chromosomes, which suppress recombination along the length of a chromosome and enable efficient maintenance of lethal mutations in the heterozygous state.

One of the most striking findings from these systematic forward genetics approaches was that different genes mutate to give the same or similar phenotypes. This resulted in the generation of 'phenotype classes' where several mutant strains were placed in the same group. For example, in the *Drosophila* screen, the segment-polarity mutants lack a particular region of each segment of the *Drosophila* embryo, whereas pair rule mutants all lack structures in alternate segments (Nusslein-Volhard *et al.*, 1980, Wieschaus *et al.*, 1984, St Johnston, 2002). Subsequent molecular identification of the genes altered in these mutants confirmed regulatory interactions between their gene products. Similarly, in *C. elegans* screens, 95 mutations affecting vulval development (either vulvaless or multivulva) were detected and were found to correspond to 22

genes. Subsequent molecular analysis showed that these mutants were involved in two signalling pathways (Ferguson *et al.*, 1985, Wang *et al.*, 2001). Thus, when forward genetics is applied on a large-scale, it is possible to group genes in pathways and networks. Furthermore, these screens demonstrated that in large collections of mutants the same gene can be mutated more than once, giving rise to an allelic series in which different mutant alleles of a gene cause a range of phenotypes. The power of genetic functional analysis is maximised when large collections of mutations are amassed such that both phenotype classes (same phenotype, multiple genes) and allelic series (same gene, multiple phenotypes) are generated.

The success of forward genetics in invertebrate organisms initiated the search of a vertebrate model system suitable for application of a large-scale forward genetics approach. This led to the development of the zebrafish, *Danio rerio*, as a model system (Streisinger *et al.*, 1981) and establishment of multiple large-scale mutagenesis screens (de Angelis *et al.*, 2000). The zebrafish is ideal for the phenotyping component of genetic screens because it produces large clutches of embryos (~100 per mating pair) and the externally fertilised eggs develop synchronously and much more rapidly than those of other related fish species studied in some laboratories (such as Medaka). Importantly, for the study of embryonic development, zebrafish embryos remain transparent through much of their embryonic development, permitting detailed microscopic observation throughout these stages.

However, there are limitations to using zebrafish for studying genetics to increase the understanding of human biology. One confounding factor is that zebrafish have maternal contribution to their embryonic development and patterning (Driever *et al.,* 1996, Heasman, J. 1997). At the early developmental stages, the normal genome of

fertilised zebrafish embryos is not transcriptionally active; thus, maternally derived products regulate early development (Jimenez *et al.,* 1982, Perrimon *et al.,* 1989). This maternal contribution (RNAs and proteins) can partially mask the effects of zygotic mutations in the zebrafish. To overcome this problem, Ciruna *et al.* (2002) developed a method for the production of maternal-zygotic zebrafish mutants by germline replacement. In contrast, zygotic transcription starts at the two-cell stage in mammals. The consequence for this on early development is shown by the recovery of cilia mutants in mouse genetic screens which firmly established a link between the formation and function of the primary cilium and Hh signalling (Huangfu and Anderson, 2005, Huangfu *et al.,* 2003, May *et al.,* 2005). The lack of cilia mutants from the zebrafish screens was initially interpreted to mean that this link was not conserved between fish and mammals, but maternal-zygotic mutation of a zebrafish cilia gene revealed this requirement (Huang and Schier, 2009).

Another minor point is because zebrafish are pseudo-tetraploid due to a partial genome duplication that occurred at the beginning of teleost radiation (Postlethwait *et al.*, 1998). Gene duplication facilitates redundancy and divergent function, meaning that functional relationships between sequence homologs cannot be presumed. Thus, the likelihood that a particular zebrafish gene mutates to give the same phenotype as the orthologous human gene is lower than that in diploid laboratory organisms such as mice. This is particularly important in the analysis of mammalian gene function during embryogenesis as the *in utero* development of the mammalian embryo precludes analysis of this period in human subjects. Despite the difficulties of working with mice (relative to vertebrate and invertebrate model organisms), efforts to understand mouse biology will remain central to the analysis of human clinical issues.

2.1.3 Forward genetic screens in mice

Phenotype-driven screens require that large numbers of the target organism be bred and analysed. The number of animals analysed in a genetic screen is inversely proportional to mutagen efficiency; for a more efficient mutagen, fewer animals are required. Any attempt at forward genetics in mice should therefore employ the most efficient mutagen possible. The effect of mutagens on the mouse genome can be quantified using specific test (*T*) stocks in which multiple recessive, visible alleles affecting easily recognisable traits are segregating. Two stocks have been generated for this purpose. One was developed at the Oak Ridge Laboratories (USA) in 1948 that harbours six alleles affecting coat colour: *a* (nonagouti, chromosome 2), *b* (brown, chromosome 4), c^{ch} (chinchilla at albino, chromosome 7), *d* (dilute, chromosome 9), *p* (pink-eyed dilution, chromosome 7), and *s* (piebald-spotting, chromosome 14) and one that controls ear morphology: *se* (short-ear, chromosome 9). The second was generated at The Harwell Laboratories (UK) in 1966 and harbours four alleles affecting the coat: *fz* (fuzzy, chromosome 1), *ln* (leaden, chromosome 1), *pa* (pallid, chromosome 2), and *pe* (pearl, chromosome 13), and one that alters the appendicular skeleton *bp* (brachypodism, chromosome 2). To use these stocks to calculate the rate at which a mutagen induces heritable gene mutations in mouse germ cells, a wild-type male is treated with the mutagen and crossed with *T*-stock females (N.B. the male germ cells are spermatogonial stem cells). Mutations in the *T*-stock are recessive, and thus (in the absence of a new mutation) all progeny appear to be wild-type, but induced, heritable mutations at any one of the test loci are immediately apparent. This specific locus test (SLT) enables calculation and comparison of the rate of mutation caused by different compounds.

Historically, mice were treated with X-rays to study the effects of radiation on the genome. In 1951, SLT was used to calculate the rate of radiation-induced mutation in mammals (Russell, 1951). Subsequently, SLT was used to assay for the harmful effects of chemicals, but the mouse was found to be surprisingly refractory to mutations in spermatogonial stem cells (Davis *et al.*, 1998). However, in 1979, a small pilot experiment showed that the rate of mutation induction by the chemical *N*-ethyl-*N*-nitrosourea (ENU) was 5-fold higher than the maximal rate obtained with X-rays (Russell *et al.*, 1979, Hitotsumachi *et al.*, 1985). Once ENU-treatment was optimised, it could induce mutations at a frequency of one per locus in every 700 gametes (12-fold more frequently than X-rays), (Russell *et al.*, 1982a, Russell *et al.*, 1982b, Hitotsumachi *et al.*, 1985) suggesting that ENU can be used in forward genetic approaches in mice. ENU creates lesions across the genome by transferring its ethyl group to oxygen or nitrogen radicals in DNA. The ethylated base is sometimes not properly recognised during the next round of DNA replication and can (after a second round of DNA replication) lead to the introduction of a base pair alteration that is invisible to cellular repair systems (Figure 1.4) (Noveroske *et al.*, 2000). Analysis of ENU-induced alterations in the mouse genome showed that AT to TA transversions (44%) and AT to GC transitions (38%) are the most common mutations induced by ENU (Justice *et al.*, 1999). Under typical conditions, treatment with ENU can produce approximately one mutation per gene within a cohort of 700 gametes (Favor *et al.*, 1988).

One feature of ENU that is both an advantage and disadvantage is that it introduces point mutations. Thus, ENU can accurately model subtle alleles (e.g., domain-specific or hypomorphic) associated with the single-nucleotide variants most often found in human disease (Justice, 2000). However, a significant limitation of using ENU as a mutagen is

that point mutations, which leave behind no molecular tag, make it difficult to efficiently identify the mutated gene responsible for a phenotype (Justice, 2000, Rathkolb *et al.*, 2000). To increase the efficiency of gene identification, many studies have adopted genetic strategies in ENU-mutagenesis experiments. One such strategy is to conduct saturation mutagenesis at a defined region of the genome using specialised genetic resources such balancer chromosomes (also known as inversions) or deficiencies (also known as deletions). In mice, these reagents do not cover the entire genome. Thus, region-specific screens, while valuable, do not allow investigation of the entire mouse genome. Another feature of mouse genetics that facilitates the localisation of ENU-induced alleles in genome-wide genetic screens is the availability of inbred strains of mice. Inbred mouse strains are the product of 20 generations of brother-sister mating (Potter *et al.*, 1995), which renders animals of the strain homozygous at all loci and enables the continued production of genetically identical mice.

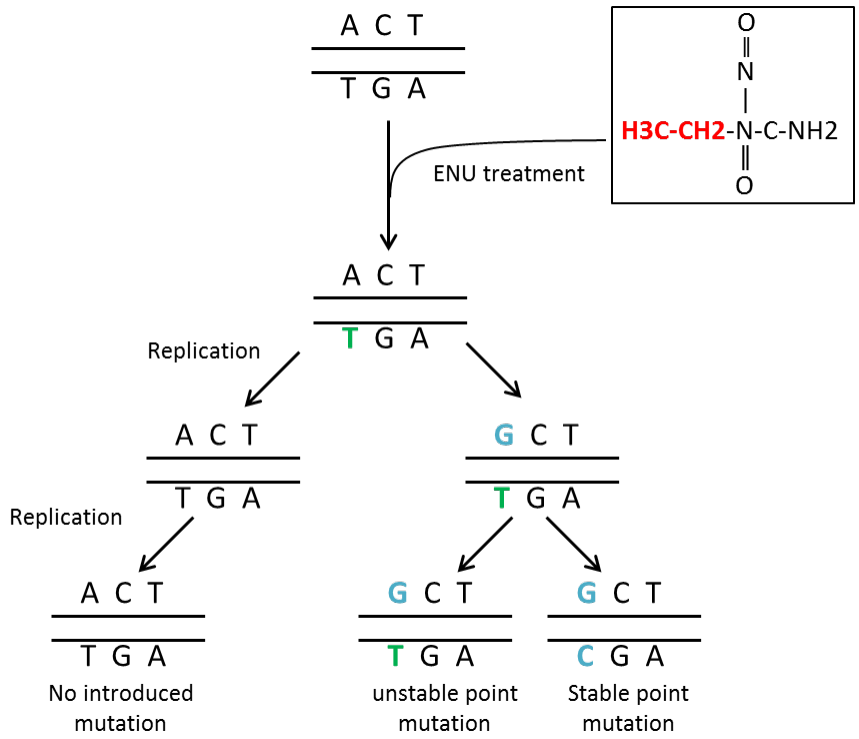

Figure 1.4: Mutagenic action of *N*-ethyl-*N*-nitrosourea (ENU). ENU can introduce point mutations in the mouse genome. The ethyl group of ENU (red) can be transferred to numerous reactive sites on any one of the four DNA nucleic acids. This generates a lesion that, if DNA repair mechanisms fail, can cause mispairing during DNA replication. This can result in the introduction of a heritable point mutation. Here, the ethyl group is transferred to a thymine (green). During the first round of DNA replication, this ethyl group may be repaired, preventing mutation. However, if repair fails, thymine will mispair with guanine (blue), creating a mutation and causing a transition from an A-T pair to a G-C pair (Harland *et al.*, 1997, Noveroske *et al.*, 2000).

Gene identification in an ENU-based genetic screen is facilitated by crossing two distinct inbred strains. Some regions of the genome are different (polymorphic) between the two inbred strains. These polymorphic regions are easy to track by recombination mapping (described below) to localise the induced mutation. The availability of different inbred strains also allows for selection of specific strains that are particularly suitable for the objectives of the screen. For example, when screening for embryonic phenotypes, a large litter size can be prioritised by using more fecund strains or a particular strain may be chosen because it is known to be permissive for the particular phenotype of interest.

1.3 Forward genetic strategies for isolating mutations affecting murine embryonic development

To initiate a genetic screen for ENU-induced mutations, male mice are injected intraperitoneally with ENU. ENU will act on all somatic and germ-line cells to introduce random point mutations (Hitotsumachi *et al.*, 1985, Justice *et al.*, 1999). Mutations in spermatogonial stem cells (i.e. the male germ-line) can be transmitted through the germ-line by mating an ENU-treated male mouse with a wild-type female animal. The progeny of these mice can be used in a variety of breeding and screening strategies depending upon the desired outcome of the genetic screen. To identify genes required for mammalian gastrulation, the Arkell laboratory has employed three different genetic screening strategies.

1.3.1 Region-specific screen for mutations affecting embryonic development

Region-specific screens employ either chromosome inversions or chromosome deletions to limit the region in which homozygous mutations are assayed. The use of inversions in mouse mutagenesis has been described previously (Justice, 2000). Its use in the mouse has largely been restricted to one screen across an engineered inversion

involving most of ***M**us **mu**sculus* chromosome 11 (Mmu 11) (Kile *et al.*, 2003). The utility of chromosome deletions in genetic screens was demonstrated by fine-scale functional analysis of genes at the *c* and *p* locus containing deletions of Mmu 7 (Rinchik *et al.*, 1990, Rinchik *et al.*, 1995). Chromosome deletions render one part of the genome hemizygous, causing the animal to be heterozygous null for each gene that normally lies within the deleted region. Mutation of the remaining allele behaves as a homozygous mutation. By conducting a two-generation screen, the effect of new, recessive alleles at genes within the deleted region can be observed. These screens are particularly efficient if the deletion removes one copy of a recessive allele for a visible trait such as for the *Del(13)Svea36H* deletion on Mmu 13. The deleted material includes a forkhead transcription factor, which is required for hair differentiation. Homozygous loss-of-function of this gene produces mice with a glossy coat referred to as satin (sa). This serves as a visible marker that enables the *Del(13)Svea36H* mouse, along with satin homozygous mice, to be used for two-generation genetic screening to identify viable and lethal recessive mutations within the segment of the mouse genome defined by the deletion (Figure 1.5).

In this type of screen, the absence of the coat colour class in a new pedigree of G2 animals indicates that the F0 male harbours a new mutation causing embryonic lethality. These types of screen rapidly identify genes essential for embryogenesis, but do not focus phenotype recovery efforts on a particular stage or biological process.

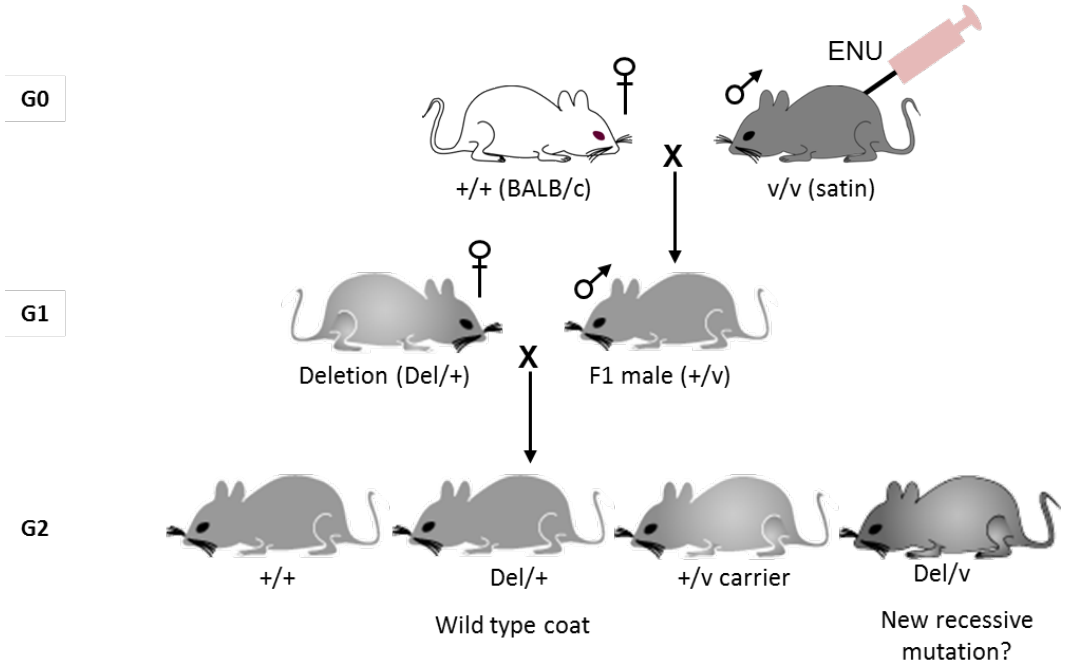

Figure 1.5: Isolation of ENU-induced mutations utilising a region-specific screen. First, the mutagenised male mouse (from the satin strain), with a visible phenotype such as coat colour, is crossed with a wild-type female (from the BALB/c strain). Each male progeny of this cross represents the F1 with the mutations inherited from their father. The F1 male is crossed with a female heterozygous for a deletion region to produce G2 animals. G2 animals fall into four genetic classes; animals with satin coat have inherited the deletion and the mutagenised v genome. These animals can be screened for phenotypes and their absence indicates new recessive lethal. v: visible coat colour, Del: deletion.

The Arkell laboratory conducted a two-generation screen across the *Del (13) Svea36H* deletion, which spans 12.7 Mb of Mmu 13 (Mallon *et al.*, 2004). From a sample of 1709 pedigrees, 13 strains carrying heritable recessive lethal mutations linked to *sa* were recovered. Most of these alleles led to phenotypes that were first evident at mid-gestation (Bogani *et al.*, 2005) and were not followed up as part of the doctoral studies reported here.

1.3.2 Screen for semi-dominant mutations altering embryonic patterning

Genetic screens for dominantly acting ENU-induced mutations require only one generation of breeding from the ENU-injected male mouse to reveal mutant phenotypes (Figure 1.6) and have been conducted on a large-scale (de Angelis *et al.*, 2000; Nolan, 2000). Dominant mutations may appear to be irrelevant to the study of embryonic lethality, but because of incomplete dominance, it is possible that these screens reveal some relevant phenotypes. In simple or complete dominance, the heterozygote and homozygote have the same phenotype. However, the phenotype of the heterozygote is often different from that of the homozygote. In mice, mutant strains displaying this type of incomplete dominance are referred to as semidominant. Two phenotype classes that are frequently associated with a semidominant mode of inheritance are tail phenotypes and coat pigmentation phenotypes. A survey of previously identified mouse mutants in the Mouse Locus Catalogue revealed 79 loci with phenotypic effects on either the tail or on coat spotting. Of these, three were modifier loci and 76 were mutant alleles with 37 of these being dominant. Of the dominant mutations, 35 exhibited semidominance (Blake *et al.*, 2003).

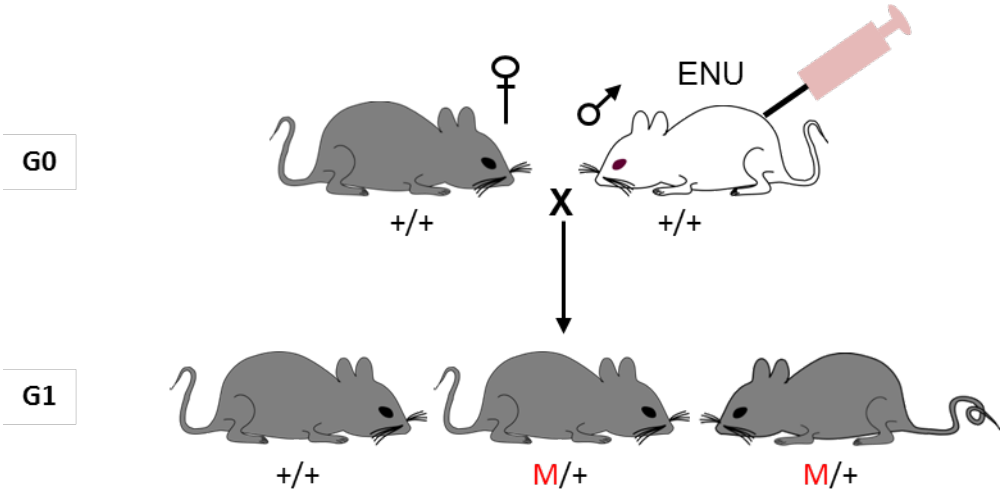

Figure 1.6: Isolation of ENU-induced mutations utilising a dominant screen. To identify the dominant mutation, the mutagenised male mouse is first crossed with a wild-type female from a different mouse strain. Each animal in G1 of this cross contains inherited mutations from the father. If any G1 animals show a visible phenotype (eg. curled tail), this indicates that a gene important for normal development has been mutated. Thus, the animal with a dominantly functioning mutation is selected for further breeding to determine if the mutation is semidominant and leads to aberrant gastrulation. M: mutation.

Although both tail and spotting phenotypes arise through patterning and/or morphogenesis defects at the end of gastrulation, because of the iterative nature of gene use during development, at least some of these mutants may reveal genes required for gastrulation. To test this hypothesis, a group of 13 newly recovered mouse mutants with dominantly inherited tail and/or spotting phenotypes were studied to determine what proportion were semidominant and whether loci required for gastrulation were uncovered by these mutations. All mutations were semidominant; for eight strains, the function of the mutated gene was traced to gastrulation (Bogani *et al.*, 2004, Fossat *et al.*, 2011). Two of these mutations, Katun (*Ka*) and Batface (*Bfc*), will be examined in this thesis and are further described in chapter 3.

1.3.3 Genome-wide screen for recessive mutations affecting gastrulation

Neither of the screens described above can recover all genes functioning in gastrulation and genome-wide screening is required to detect recessive mutations important in gastrulation. To identify recessive mutations amongst the progeny of ENU-treated animals, a three-generation breeding protocol is required (Figure 1.7). Individual male progeny of the ENU-injected animal are used as founder (F0) animals for a new pedigree. Each founder is crossed with a wild-type female to produce G2 animals. Each G2 animal has a 50% chance of inheriting any single mutation carried by the F0 male and will therefore be heterozygous for a subset (~50%) of the newly-induced mutations. These mutations can be brought to homozygosity using one of two approaches: backcross of G2 females to their F0 father (Figure 1.7) or intercross between G2 siblings (not shown). If the mutation segregation obeys Mendelian genetics, then using the backcross strategy, 25% of G3 progeny are expected to be homozygous for the mutation. Because only 50% of G2 animals are carriers for each mutation, the expected overall frequency

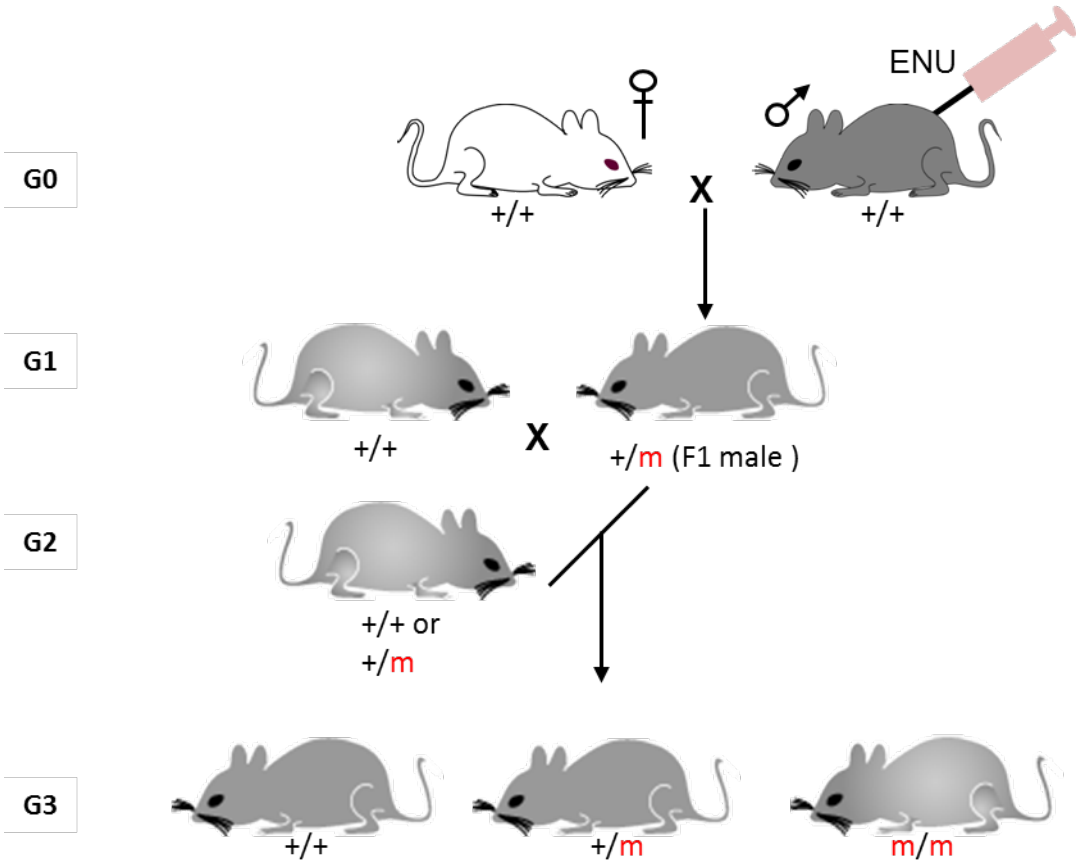

Figure 1.7: Isolation of ENU-induced mutations utilising a recessive screen. To identify recessive mutations, pedigrees must be established from individual F1 animals. First, the mutagenised male mouse is crossed with a wild-type female from a different strain. Each male progeny of this cross represents the F1. Each F1 male is crossed with a wild-type female to produce G2 animals. G2 females, which are 50% wild-type and 50% heterozygous for the mutations inherited from their father, are backcrossed to the F1 male. If both the F1 male and G2 female are heterozygous for the same recessive mutation, 25% of their G3 progeny will be homozygous for the mutation.

of homozygotes amongst all G3 progeny is 12.5%. Calculations have shown that the G2 female backcross protocol requires fewer mice to be bred than the G2 intercross strategy (Barbaric *et al.*, 2007) and the backcross protocol is most commonly employed for screening of embryonic phenotypes.

The utility and feasibility of genome-wide recessive screens for mouse developmental phenotypes was first demonstrated by Karsarkis *et al.,* (1998) who conducted a pilot screen of 86 pedigrees, in which each G2 female was backcrossed in timed mating to their F0 fathers and litters were dissected at early organogenesis (9.5 dpc). This procedure generated five pedigrees with new, heritable mutations, which caused defects observable by visual inspection under a dissecting microscope. Four pedigrees were readily mapped to a single locus by low-resolution recombination mapping (see section 1.4), whereas one pedigree was not, suggesting it had a more complex genetic basis. Since then, tens of genome-wide screens for recessive mutations affecting various aspects of mouse embryonic developments have been conducted (Table 1.1). On average, these screens identify one mutant pedigree among every ten screened.

None of these screens have directly sought to identify genes that function during gastrulation. Many pedigrees have been screened by dissection at 9.5 dpc because it is reasoned that mutations which affect post-implantation development can be identified by aberrant morphology at this stage, but will not have been completely resorbed (Kasarskis *et al.*, 1998). This is not entirely true, as mutations with a severe effect on the early stages of gastrulation (rather than on later patterning processes) are generally fully resorbed by 9.5 dpc. Therefore, the Arkell laboratory initiated a small screen aimed at recovering these early-acting mutations.

Table 1.1: Overview of mutations found in recessive screens.

Assay	Number of pedigrees	Number of heritable mutations/mapped	Reference
Visual inspections of G3 embryos at 9.5 dpc	86	4	Kasarskis *et al.*, 1998
Examining G3 embryos at 10.5 dpc	150	4	Hentges *et al.*, 1999
Examining G3 embryos at 18.5 dpc by necropsy using a hierarchical approach	21	9	Herron *et al.*, 2002
Morphology examination of G3 embryos at 9.5 dpc	380	42	García-García *et al.*, 2005
Examining G3 mice at 8 and 16 weeks of age for dilated cardiomyopathy	25	1	Fernandez *et al.*, 2009
Visual inspections of G3 embryos at 9.5 or 13.5 dpc for morphological L-R axis defects	135	10	Ermakov *et al.*, 2009
Visual inspections of G3 embryos at 10.5 dpc	19	8	Wansleeben *et al.*, 2011
Visual inspections of G3 embryos at 13.5 and 18.5 dpc	52	4	Caruana *et al.*, 2013
Examining G3 mice at 18.5 dpc or 21 weeks depending on the screening design	59	12	Ha *et al.*, 2015
Visual inspections of G3 mice from 12 till 80 weeks age	157	72	Potter *et al.*, 2016

The time of dissection was shifted 36 h earlier to 8.0 dpc. At this stage, embryos that have failed to initiate gastrulation as well as those which initiate but do not progress through gastrulation remain present. Additionally, some early patterning defects are already evident. The screen design and associated technical challenges are described in chapter 4 of this thesis, as are the experiments used to localise newly induced point mutations captured in the screen. In chapter 5, cloning of one mutant line is described. Together, these experiments demonstrated that a screen aimed directly at recovering mutations affecting the onset and progress of gastrulation is feasible in mice.

1.4 Identification of causal ENU-induced mutation

Positional candidate cloning refers to the process by which the gene responsible for a phenotype is identified (Hedrich, 2004). The steps involved in this process are (1) recombination mapping, (2) identification of pathogenic variants in the mapped interval, and (3) confirmation of the causal relationship between the gene mutation (i.e. the pathogenic variant) and mutant phenotype. The first step of recombination mapping (also known as linkage analysis) tracks meiotic events to identify recombination breakpoints that lie close to the mutation in a panel of progeny from a cross segregating the mutation and marker loci. The principle of recombination mapping is shown in Figure 1.8 (Avner *et al.*, 1988). In a cross designed for recombination mapping, the first generation requires that a mutant mouse is crossed with a wild-type animal from a different inbred strain (known as the mapping strain). The progeny of this cross are F1 hybrids, having inherited one complete set of chromosomes from each parent. In the next generation, F1 animals can be either intercrossed or backcrossed to the mapping strain. In all work presented in this thesis, the backcross strategy was used and will be described here. When backcrossed, meiotic recombination events events in germ cells

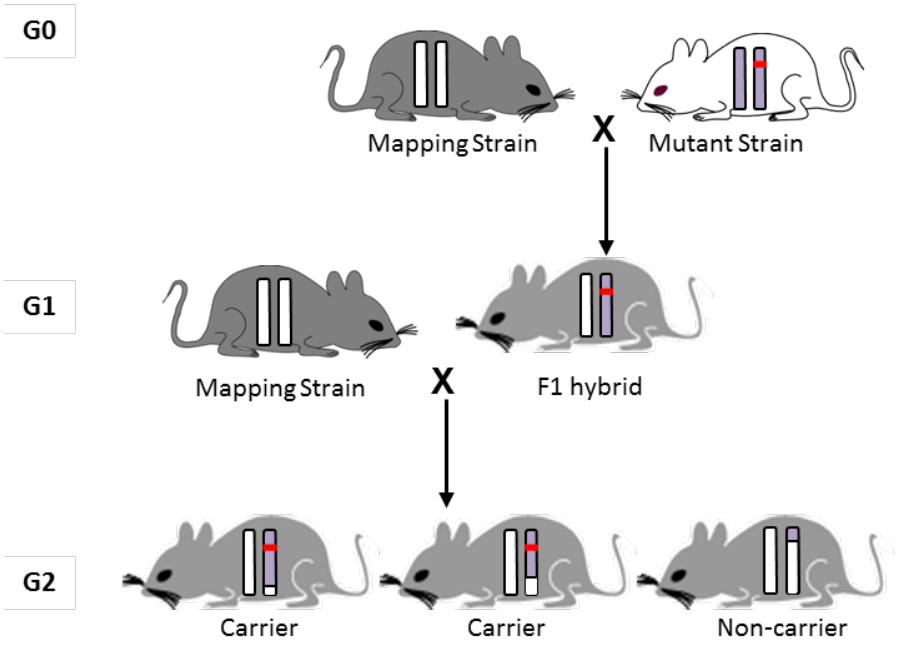

Figure 1.8: Principle of recombination mapping. In the first generation, a mutant strain is crossed with a wild-type animal from a different inbred strain (mapping strain). The F1 progeny inherit an entire chromosome from both the mapping (white) and mutant (purple) parental strains and may inherit the mutation (in red). In the next (G1) generation, an F1 animal is backcrossed to the mapping strain. Meiotic recombination events in germ cells of the F1 parent can be identified in G2 progeny using polymorphic markers. Animals carrying the mutation (red) regions close to the mutation (i.e. linked) will retain the mutant (purple) genome, whereas unlinked regions will generally show homozygosity for the mapping (white) genome. Adapted from (Avner *et al.*, 1988, Harland *et al.*, 1997). G: generation.

of the F1 parent can be identified using genomic loci (or markers), which are known to vary between the mapping and mutant strains (i.e. are polymorphic). This process can establish the combination of alleles along the length of a chromosome, which is known as a haplotype.

To identify recombination events, different types of polymorphic markers can be utilised. First, many simple sequence repeats occur across the genome of all eukaryotic organisms (Tautz, 1989). These repeat regions can, for example, be stretches of di-, tri- or tetranucleotide repeats known as microsatellites (MIT) (Love *et al.,* 1990). Expansions of these repeats occur at sufficient frequency that size differences exist between inbred strains of mice (Tautz, 1989). This variation in length, often referred to as a fragment length polymorphism, can be readily analysed by PCR across the repeats and therefore utilised to distinguish inbred strains of mice (Tautz, 1989). MIT markers have become widely used in genetic mapping (Love *et al.*, 1990) and more than 6000 known differences between inbred strains are curated on the Mouse Genome Informatics database, which serves as a reference for identifying MIT markers suitable for particular recombination mapping experiments. The mouse genome is approximately 300 Mb in size, leading to an estimate of the average spacing of microsatellite markers as one per 500 Kb. Not all markers will be polymorphic for the two strains used in any recombination mapping experiment and MIT markers are not evenly distributed across the genome. Thus, it may be difficult to identify a sufficient number of polymorphic MIT markers for high-resolution mapping (Love *et al.*, 1990). A second type of marker, which is more abundant across the genome, is the single-nucleotide polymorphism (SNP). However, SNPs are more difficult to detect and often require advanced technologies, which are more costly than simple PCR approaches that can distinguish MIT markers (Fairfield *et al.*, 2011). The most suitable panel of markers for recombination mapping

will be determined by a range of practical considerations, such as resolution of mapping required, local availability of specialised SNP detection techniques, and cost (Neuhaus *et al.*, 1998).

Once a genomic interval known to contain the mutation (i.e. the minimal critical region; MCR) has been established via recombination mapping, the next step is to identify pathogenic variants within the interval. This is generally accomplished by DNA sequencing of all exons and splice sites within the MCR. If recombination mapping has been carried out to high resolution, it may be feasible to use Sanger sequencing to scan all exons and exon/intron boundaries. The recent advent of low-cost, high-throughput sequencing (known as next-generation sequencing) has enabled a modified approach, in which low-resolution mapping is combined with whole exome sequencing (WES) (Arnold *et al.*, 2011). The exome is the set of all exons (or coding regions) in a genome and WES approaches cover not only coding regions, but also intron/exon boundaries (Fairfield *et al.*, 2011). WES data is then assessed for pathogenic variants that fall within the MCR. In both approaches, ENU-induced mutations in the non-coding regions will not be included in the analysis. Empirically, most ENU-induced mutations that cause abnormal phenotypes appear to lie in coding regions (Blake *et al.*, 2011, Fairfield *et al.*, 2011), with current data suggesting 2/3 of the mutations alter sequencing itself and 1/3 mRNA splicing. Reagents for WES therefor cover both exons and splice junction.

The final step in the positional candidate approach is to confirm the causal relationship between the pathogenic variant and mutant phenotype. The exact scope of this task varies with each project, but commonly involves compiling a variety of genetic and functional data that together provide evidence of causality. This may include evidence of tight linkage between the mutation and phenotype, such as non-segregation in a large

number (>1000) of meiosis. It could also include gene expression data, protein modelling approaches, and cell-based biochemical studies to compare functional attributes of the wild-type and mutant proteins. The strongest support, however, stems from additional mouse genetic experiments such as a complementation test with a second allele and/or transgenic rescue of the mutant phenotype.

1.5 Scope of this thesis

This thesis aims to contribute to the identification of genes that direct murine gastrulation by:

1. Employing a pre-existing genetic mutant, isolated from a screen aimed at ENU-induced semi-dominant alleles, to infer the function of the identified gene during gastrulation.

A genetic screen, aimed at identifying mutations that affect embryogenesis, isolated a mouse strain, katun (*Ka*), carrying a nonsense mutation in the zinc finger of the cerebellum 3 (*Zic3*) gene that generates a null allele (Ahmed *et al .,* 2013). This gene is a member of the zinc finger of the cerebellum (ZIC) family of transcriptional regulators; previous work has shown that ZIC proteins can inhibit Wnt signalling pathway *in vitro* (Aruga, 2004, Pourebrahim *et al.*, 2011). My research project is concerned with determining if *Zic3* inhibits canonical Wnt signalling *in vivo*. Toward that, katun mice were crossed with the batface (*Bfc*) activating allele of β-catenin. This increased the penetrance and severity of elevated Wnt-associated phenotypes, providing the first evidence that the ZIC proteins function as novel *in vivo* Wnt signalling inhibitors. In mouse and human, mutation of *Zic3*/*ZIC3* is associated with Heterotaxy, a disorder of left-right axis formation. My analysis of the *Ka* and *Bfc* strains will provide an insight to

the role, if any, of Wnt signalling in the development of Heterotaxy. These data are presented in chapter3.

2. Evaluating the feasibility of genetic screen aimed at ENU-induced recessive alleles that alter gastrulation.

Random point mutations were introduced into the genome of an inbred strain of mouse by the chemical mutagen ENU. The mutagenised mouse strain was crossed with the wild-type inbred strain to generate pedigrees with new mutations. Subsequent generations were examined at the end of gastrulation to determine if any of the denovo mutations altered gastrulation. The new mutant strains were selected by comparing the overt morphology of embryos, as viewed under a light microscope, at a stage close to the end of gastrulation. The mutant phenotypes selected, fell into different phenotypic classes. My project aims at taking each of the newly identified mutant strains through to the next stages of investigation: (i) to identify the location of the newly induced point mutation within the mouse genome of each of the eight strains and (ii) to determine the role of the mutated gene, during normal embryo development, by characterising the phenotype of mutant embryos. This study is likely to identify genes that have an important, but previously unrecognised, role in directing mammalian gastrulation. These data are presented in chapters 4 and 5.

Chapter 2: Materials and Methods

2.1 General molecular techniques

2.1.1 Bacterial transformation

Plasmid DNA was transformed into DH5α electro-competent Escherichia coli (*E. coli*) bacteria via electroporation. For electroporation ~20 ng of plasmid DNA was added to ice-cold *E. coli* cells. The mixture was transferred to an ice-cold cuvette and shocked at 2.5 kV using a BioRad *E. coli* Pulser (BioRad). Immediately after pulsing, 500 µL of Luria-Bertani (LB) Broth media (Bacto; Cat. No. 244620) was added to revive the bacteria and the solution was incubated for 1 hour on a rotating platform at 37°C. The bacterial solution was then plated on to LB Agar (Bacto; Cat. No. 244520) containing 0.1 mg/mL of the appropriate antibiotic (Ampicillin: Sigma-Aldrich; Cat. No. A9518-5G, or Kanamycin: Sigma-Aldrich; Cat. No. 60615-5G) and incubated for 16 hours at 37°C.

2.1.2 Plasmid isolation

Depending on the amount of DNA required, plasmid DNA was isolated using two methods:

1. For small scale isolation, single colonies from LB Agar plates were picked and each incubated in 4 mL of LB Broth (containing 0.1 mg/mL of the appropriate antibiotic) at 37°C for 16 hours on a rotating platform. 3 mL of this overnight culture was then used to isolate plasmid DNA via a High Pure Plasmid Isolation kit (Roche; Cat. No. 11754785001) (average yield: 7-10 µg).
2. To obtain a higher quantity and/or purity of plasmid DNA, single colonies from LB Agar plates were picked and each incubated in 2 mL of LB Broth (containing 0.1 mg/mL of the appropriate antibiotic) at 37°C for 6-8 hours on

a rotating platform. This culture was added to 100 mL of LB Broth (containing 0.1 mg/mL of the appropriate antibiotic) and incubated at 37°C for 16 hours on a rotating platform. Plasmid DNA was isolated using a NucleoBond® Xtra Midi kit (Macherey-Nagel; Cat. No. 740410.50).

In each case DNA was resuspended in Analar H_2O, instead of the supplied buffer.

2.1.3 Restriction enzyme digestion

DNA was digested with restriction enzymes for the preparation of riboprobes and analysis of plasmid inserts. For plasmids, 1 µg of DNA was digested using 5 U of the appropriate restriction enzyme (Table 2.1). Reactions were performed with the appropriate NEBuffer (10X stock), Bovine Serum Albumin (BSA; 10X stock) and enough Analar H_2O such that the final volume was 10-20 times greater than the volume of restriction enzyme added. Reactions were incubated as per the enzyme manufacture instruction and digestion was confirmed via agarose gel electrophoresis.

2.1.4 Agarose gel electrophoresis

After DNA was digested using restriction enzymes or amplified using PCR, the size of the resulting bands was visualised using agarose gel electrophoresis. Agarose gels were made by dissolving Ultrapure Agarose (Life Technologies) in TBE buffer (0.1 M Tris-HCl, 0.09 M Boric acid and 0.001 M EDTA; Amersco) to create 1-3% gels, dependant on the band size being resolved. Once dissolved and cooled slightly, RedSafeTM Nucleic Acid Staining Solution (Intron Biotechnology) was added at a 1:20,000 dilution to allow visualisation of the DNA and the gel was allowed to set within a mould. To prepare the DNA for gel electrophoresis, 1 µL of 4X loading dye was added for every 4 µL of sample with pipetting or vortexing to mix. The samples were then loaded into separate wells of the agarose gel. In an adjacent well, 0.6 µg of 1 Kb Plus DNA ladder (Invitrogen) was

loaded into the gel to indicate band sizes. Once all the samples and the ladder had been loaded, a voltage of 5-6 Vcm-1 was applied to the gel for 25-50 minutes to separate the DNA fragments. The gel was then photographed by a Gel Doc XR (Bio-Rad), which exposes the gel to a UV light source so that the DNA fragments are visible.

Table 2.1: Restriction enzymes used to digest plasmid DNA. All enzymes were purchased from New England BioLabs® / = denotes cut site.

Enzyme	Recognition sequence (5`-3`)	Cat. No.	NEBuffer
*EcoR*I	G/AATC	R0101	NEB1
*EcoR*V	GAT/ATC	R0195	NEB3
*Hind*III	A/AGCTT	R0104	NEB2
*Sac*I	GAGCT/C	R0156	NEB1
*Sac*II	CCGC/GG	R0157	NEB4
*Sa*lI	G/TCGAC	R0138	NEB3
*Xho*I	C/TCGAG	R0146	NEB4

2.1.5 Microsatellite marker selection

As oligonucleotide primer information was available for MIT markers in MGI; these sequences were utilised for primer design (Blake JA, 2011). For those MIT markers with no oligonucleotide sequence available, a pair of primers (forward and reverse) was designed by inspecting the genomic sequence surrounding the microsatellite. Markers within approximately 30 Mb from the most proximal and distal regions of each chromosome (MMu 1- MMu 19) were identified and selected for analysis along with additional medical markers for MMu 1, MMu 2 and MMu 3. Utilising the online resource *Ensembl,* the sequence of each MIT marker was obtained, with 200 bp of 5' (upstream) and 3' (downstream) flanking sequence in (See Table 1 and Table 2 in **Appendix A)**. PCR was used to reactions were optimised for each microsatellite (MIT) marker. Primer pairs for each MIT marker were tested in two different PCR buffers (ABgene Buffer I and ABgene Buffer IV), and at two different Touchdown PCR conditions, with initial annealing temperatures of 60°C and 65°C .

2.1.6 Primer design

17-mer forward and reverse primers were designed by visual inspection of this sequence according to the following guidelines: the base composition of each primer was either 9 G/C and 8 A/T bases or 8 G/C and 9 A/T (with each member of a primer pair of matching composition); each primer had a G/C clamp at the 3' end such that a minimum of three of the five 3' most bases was G or C; consideration of the melting temperatures (Tm) of primers which should be between 45-65°C for a successful PCR; and primer pairs, e.g. forward and revers primers used in the same PCR reaction, should have Tm within 55°C of each other. The primers used as MIT markers were designed to generate an amplicon of 200 – 300 bp in size.

2.1.7 Polymerase chain reaction (PCR)

DNA fragments for cloning or sequencing were amplified using PCR. PCR products were amplified from approximately 40 ng of plasmid DNA or 30 ng genomic DNA in the presence of 0.8 µM each oligonucleotide (Primers are listed in Table 2.2 and Table 1 and 2 in appendix A) and a commercial PCR buffer. Three commercial buffers were used: PCR Buffer I (BI) (AbGene) lead to final reaction conditions of 10 mM Tris-HCL (pH8.3 at 25^{o}C), 50 mM KCl and 1.5 mM $MgCl_2$; PCR Buffer IV (Kemp *et al.*) (AbGene) lead to final reaction conditions of 75 mM Tris-HCL (pH8.8 at 25^{o}C), 20 mM $(NH_4)_2SO_4$, 0.01 % (v/v) Tween 20 and 1.5 mM $MgCl_2$; and Immomix lead to final reaction conditions of 10x ImmoBuffer. 160 mM $(NH_4)_2$ SO_4, 1M Tris-HCI pH8.3 and enhancers. Thermal cycling was performed using a Touchdown (TD) PCR program (Don *et al.*, 1991) with an initial annealing temperature of either 60^{o}C (TD60) or 65^{o}C (TD65). Once prepared, the reaction mixture was cycled through the PCR amplification program (Table 2.3) using an Eppendorf Mastercycler PCR machine. Success of the PCR was analysed via gel electrophoresis.

2.1.8 Purification of DNA products

PCR fragments and digested plasmids were purified using the Nucleospin® Extract II purification kit (Scientifix) according to the manufacturer's instructions. PCR fragments were run directly through the kit for purification, whilst digested plasmid DNA was separated from contaminating DNA via gel electrophoresis and then purified. PCR fragments amplified for sequencing were purified using the Agencourt AMPure PCR purification system following the manufacturer's instructions. The concentration and purity of the eluted DNA were determined using a spectrophotometer (NanoDrop 1000 v8.1) and stored at -20°C until required.

Table 2.2: Primers used for genotyping and sequencing assays.

Primer	Sequence (5` - 3`)	Genotyping assay
Ark241_F	GCG CCT TCT TCC GTT ACA TG	Allelic discrimination
Ark242_R	AGC CTC CTC GAT CCA CTT ACA G	Allelic discrimination
Ark1002_Ubel_F	GAG GTC ATC ATG AAG GTC AG	HRMA
Ark1003_Ubel_R	CTG GAA AGT TTA TGC CC	HRMA
Ark1015_F	GAA AGC CGC TTC TTG TAA TC	HRMA
Ark1016_R	TGG CCT CAA GGG CTA GTT TCT	HRMA
Ark1085_F	CCT TCT TCC GTT ACA TGC G	Allelic discrimination
Ark1086_R	CTG AGC CTC CTC GAT CC	Allelic discrimination
Ark1345_Atp1a1_F	TGA GGT GG ATG AAA TGC	Sequencing
Ark1347_Atp1a1_R	CCA AGC ACA TTT TCT GG	Sequencing
Ark1393_Atp1a1_F	GG CAA TCC TGC CCA TCA C	Allelic discrimination
Ark1394_Atp1a1_R	GAT CCC CTC ACA CCT TCT CTT AAG	Allelic discrimination

F: forward primer; R: reverse primer.

Table 2.3: Thermal cycling conditions of PCR programs used for optimisation of genotyping assay.

Program:	TD 60		TD65	
Initial denaturation	94°C 4 min		94°C 4 min	
Denature	94°C 30 sec	X 29	94°C 30 sec	**X 19**
Anneal	60°C, decreasing 0.5°C per cycle 30 sec		65°C, decreasing 0.5°C per cycle 30 sec	
Extend	72°C 30 sec		72°C 30 sec	
Denature	94°C 30 sec	X 19	94°C 30 sec	**X 29**
Anneal	45°C 30 sec		55°C 30 sec	
Extend	72°C 30 sec		72°C 30 sec	
Polish	72°C 7 sec		72°C 7 sec	
Stop	**4°C**	**Hold**	**4°C**	**Hold**

2.2 Mouse colonies

2.2.1 Mouse husbandry

All mouse strains were maintained at the Australian Phenomics Facility at The Australian National University according to Australian Standards for Animal Care under ethics protocol A2015/10. Mice were maintained in a light cycle of 12 h light: 12 h dark, the midpoint of the dark cycle being 12 a.m. For the production of staged embryos, 12 p.m. on the day of the appearance of the vaginal plug is designated 0.5 dpc.

2.2.2 Low-quality genomic DNA preparation

Low-quality genomic DNA preparation was routinely used for genotyping of either ear biopsy tissue or embryonic tissue. Ear notches were collected in 50 µL of lysis solution [50 mM Tris(hydroxymethyl)aminomethane (Tris)-HCl (pH 8.5), 1 mM Ethylenediaminetetraacetic acid (EDTA), and 5% Tween 20] containing 2 µL of Proteinase K (10 µg/µL; Sigma Aldrich, Cat. No. P2308) and incubated at 55°C for 60 minutes, followed by 95°C for 10 minutes to inactivate Proteinase K. Tissue debris was pelleted by centrifuging at 2,000 *g* for 5 minutes and each sample was diluted 1:10 in Analar H_2O. Embryos were genotyped using a fragment of extra embryonic tissue/ectoplacental cone (7.5 dpc) or yolk sac (8.5 dpc and 9.5 dpc). Tissue was collected in 20 µL (7.5 dpc), 25 µL (8.5 dpc) or 35 µL (9.5 dpc) of lysis buffer (50 mM Tris-HCl, pH 8.5, 1 mM EDTA, 0.5% Tween 20) containing 1 µL of Proteinase K (10 µg/µL), followed by incubation at 55°C for 25 minutes and 95°C for 5 minutes to inactivate Proteinase K.

2.2.3 High-quality genomic DNA preparation

Approximately 0.5 cm of the tail or one ear were biopsied and homogenised in 1.5 mL of TE-SDS (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1% SDS) containing 0.5 µg/µL proteinase

K (PK) in a 12 mL glass corex tube on ice. The sample was sealed and incubated overnight at 55^{o}C. Cellular debris was precipitated by the addition 8M KOAc to a final concentration of 1.2 M, and the specimen was centrifuged in a Sorvall HB-4 rotor at (23,000 *g*, 4^{o}C) for 20 minutes to pellet debris. To precipitate the DNA 1.5 mL of the supernatant transferred to 3.75 mL of 100% EtOH in a 15 mL Falcon tube. The sample was inverted 10 times, and the DNA precipitate removed by spooling onto a glass Pasteur pipette with a flame sealed end. DNA was washed by immersion in 1 mL of 70% EtOH, allowed to air dry, after which it was resuspended in 500 µL of TE. DNA was quantified by measurement of the optical density at 260 nm via NanoDrop (NanoDrop 1000 v8.1) DNA preparations were diluted to 10 ng/µL for further PCR analysis.

2.2.4 Genotyping assays

Genotyping of adults and embryos from the ENU, *katun*, *batface* and *atp1a1* mouse strains was done using the following techniques:

2.2.4.1 High resolution melt analysis (HRMA)

Digested ear biopsies were genotyped using a previously described High Resolution Melt Analysis (HRMA) assay (Thomsen *et al.*, 2012) with no deviations. HRMA genotyping primers are shown in Table 2.2.

2.2.4.2 Allelic discrimination

PCR for allelic discrimination was done using TaqMan® Universal PCR Master Mix (Life Technologies; Cat. No. 4304437). All reactions were carried out in a final volume of 15 µL with ~30 ng of digested ear notch DNA in the presence of 0.9 µM of primers (Table 2.2). Additionally 0.25 µM of allele-specific probes (Applied Biosystems®) were used to identify the wild-type alleles and mutant allele (Figure 2.1, Table2.4). Reactions were set up in 96-well Half-Skirted PCR Microplates (Axygen®; Cat. No. PCR-96-LP-AB-C) covered

with Axygen Microplate Sealing Film (Fisher Scientific; Cat. No. UC500) and performed using the StepOnePlus™ Real-Time PCR System (Applied Biosystems®). The StepOne Software (version 2.2.2; Applied Biosystems®) was used to run the assay using the following conditions: an initial pre-PCR read at 60°C for 30 seconds to record background fluorescence, followed by 95°C for 10 minutes to denature the template and a cycling stage of 95°C for 15 seconds and of 60°C for 1 minute for 50 cycles. A post-PCR read was performed at 60°C for 30 seconds to collect data after completion of the PCR. Data was analysed using the same software that records the pre- and post-PCR reads and calculates normalised dye fluorescence (ΔRn) from the wild-type and mutant alleles as a function of cycle number. Based on this data the software called the sample as homozygous for either wild-type or mutant allele, or heterozygous with both alleles.

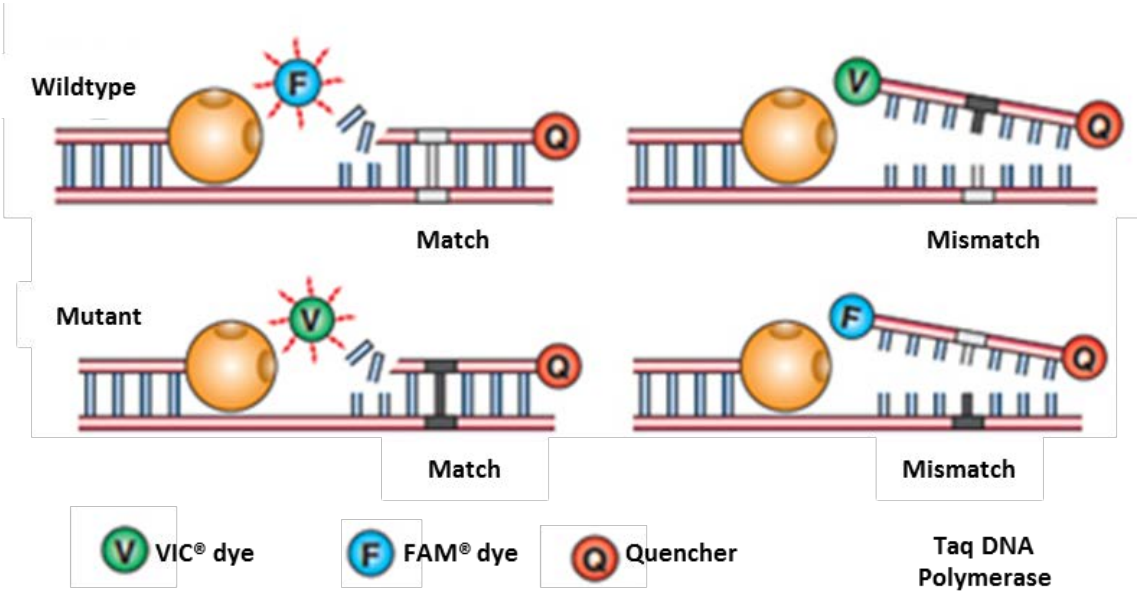

Figure 2.1: Allelic discrimination using fluorogenic probes. Wild-type and mutant alleles were distinguished using non-extendable oligonucleotide fluorogenic probes (Livak, 1999). The FAM™ (6-carboxyfluorescein) probe complemented the wild-type allele, while the VIC® (4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein) probe complemented the mutant allele. The probes differed by only one nucleotide (at the site of the mutation lucs), and at the 5' end were labelled with the reporter (VIC® or FAM™) dye and at the 3' end with the quencher (MGBNFQ: minor groove binding non-fluorescent quencher). The proximity of the reporter dye to the quencher reduced fluorescence via the reporter. During PCR as the forward primer extended the target specific probe was degraded by the 5'-3' exonuclease activity of the Taq DNA polymerase (Arya *et al.*, 2005), which released the reporter dye allowing it to fluoresce. Both probes were included in the same well for each PCR reaction thus fluorescence was only observed when target allele for the probe was present (McGuigan *et al.*, 2002). As PCR cycles increased, the fluorescence intensity increased that was proportional to the accumulation of the PCR product. The fluorescence was measured using a Real-Time PCR machine. This figure was adapted from Applied Biosystems' Allelic Discrimination Getting Started Guide.

Table 2.4: Probes used for allelic discrimination.

Probes name	Fluorogenic probes	Sequence (5` - 3`)	3‘ quencher
Ark_Atp1a1_wt	FAM	CCA GTG TAC ACG ACG	MGBNFQ
Ark_Atp1a1_mt	VIC	CAG TGT ACT CGA CGA TG	MGBNFQ
Ark_Bfc_wt	FAM	TTT ATT CCA GCA ACA TAC	MGBNFQ
Ark_Bfc_mt	VIC	TTT TAT TCC AGC AAA ATA	MGBNFQ
Ark_Zic3_wt	FAM	CAT CAA GCA GGA GCT	MGBNFQ
Ark_Zic3_mt	VIC	CAT CAA GCA GTA GCT TG	MGBNFQ

wt: wildt-ype, mt: mutant, FAM™: 6-carboxyfluorescein, VIC: 4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein and MGBNFQ: minor groove binding non-fluorescent quencher.

2.3 Carrier identification

Within each strain from the gastrulation screen, female animals were mated to an identified affected (carrier) male (F1 founder male or N2 male) and each litter was dissected and examined by the Arkell laboratory (Figure 1.3). Where mutant embryos were identified in a litter both adult parents were scored as affected.

2.4 Sequencing of Atp1a1 exon 8

The *Atp1a1* exon 8 amplicon was amplified from genomic DNA (30 ng) in a 15 µL reaction volume as described in section 2.2.4.2 with the following specific conditions (AbGene Buffer IV, TD60) and the primers 5'TGAGGTGGATGAAATGC 3' (forward) and 5'CCAAGCACATTTTCTGG 3' (reverse). PCR products were purified as per section 2.1.8. For sequence analysis, the purified products (3-10 ng) were amplified in 80 mM Tris-HCL (pH 9 at 25^{o}C) and 2 mM MgCl2 with 1 µL BigDye Terminator v3.0 Cycle Sequencing Ready Reaction (Applied Biosystems) and 3.2 pmol of the primer 5'CCAAGCACATTTTCTGG 3' in a total volume of 20 µL. Thermal cycling conditions were 94˚C for 5 minutes and 96˚C for 10 seconds, 50˚C for 5 seconds, 60˚C for 4 minutes for 30 cycles, 4˚C hold. The extension products were precipitated by the addition of EDTA (pH 8.0) to a final concentration of 10 mM and 3M NaOAc (pH 5.2) to a final concentration of 0.4 M together with 2 x the final volume of 100 % ethanol. Samples were incubated at room temperature for 15 minutes before centrifugation (3200 x *g*, 4°C) for 10 minutes. The supernatant was discarded and pellet washed with 250 µL of 70% EtOH before centrifugation (3200 *g*, 4°C) for 10 minutes. The supernatant was discarded and samples air dried to remove any additional EtOH. The reaction products were analysed using an ABI 3730 sequencer (Applied Biosystems) by the Biomolecular

Resource Facility (BRF; John Curtin School of Medical Research, Australian National University). The resulting sequence files were analysed using Geneious (Biomatters).

2.5 Embryo collection

All embryos were dissected from maternal tissue and Reichert's membrane removed in 10% (v/v) foetal bovine serum in phosphate buffered saline (PBS). The recovered embryos were staged according to Downs and Davies (1993). For ENU embryos, photos for analysis were taken after dissection, prior to fixing. Embryos were transferred to 4% paraformaldehyde (PFA) in PBS for overnight fixation at 4°C.

2.6 Whole Mount In Situ Hybridisation analysis

WMISH was performed as previously described (Elms *et al*., 2003) using riboprobes listed in Table 2.5. Template DNA was linearised and an antisense probe *in vitro* transcribed with the enzyme and RNA polymerase listed respectively. For each probe and stage of embryogenesis examined a minimum of four embryos were compared to precisely stage-matched littermates. Upon completion of the WMISH procedure, embryos were postfixed in 4% PFA and transferred via a glycerol series in PBS (50%, 80%, 100%) for 15 minutes each. For photography, embryos were flat-mounted under a glass coverslip or placed on 1.5% agarose plate in PBS and photographed by a Nikon SMZ 21500 Stereomicroscope and DS-Ri1 camera (Nikon).

Table 2.5: Probes used for WMISH.

Probe name	Plasmid	Enzyme (antisense)	RNA	Where it stained	Reference
Axin2	pBS-KS+	*EcoRI*	T3	At 8.5 dpc: in the midbrain and tailbud	Jho *et al.*, (2002)
Brachyury (*T*)	pBR331	*HindIII*	T7	At 9.5 dpc: tailbud and notochord	Herrmann *et al.*, (1990)
Dand5	pGEM-T easy	*SacII*	SP6	At 8.5 dpc: asymmetrically crown cells of the node	Barratt *et al.*, (2014)
Fgf8	pBS-KS+	*EcoRI*	T3	At 9.5 dpc: forebrain, midbrain-hindbrain junction and foregut cells	Mahmood *et al.*, (1995)
Fst	pBS-KS II	*EcoRV*	T3	At 9.5 dpc,: somites, in rhombomeres 2, 4 and 6 of the hindbrain	Albano *et al.*, (1994)
Lefty1/2	pBS-KS II	*XhoI*	T7	At 8.5 dpc: the midline (*Lefty1*) and left LPM (*Lefty2*)	Meno *et al.*, (1996)
Nodal	pBS-KS II	*EcoRI*	T3	At 8.5 dpc: asymmetrically at the node and LPM	Conlon *et al.*, (1994)
Noto	pGEMT easy	*SalI*	T7	At 7.5 dpc: pit cells of the node	Barratt *et al.*, (2014)
Nppa	pT7T3D	*EcoRI*	T3	At 9.5 dpc: the right atria and the left ventricle of the heart	Moorman *et al.*, (2003)
Ptix2	pBS-KS II	*SacI*	T3	At 8.5 dpc: left LPM At 9.5 dpc: left cardinal vein	Ryan *et al.*, (1998)
Pkd1l1	pBS-KS II	*EcoRI*	T7	At 7.5 dpc: at the node sensory cilia	Field *et al.*, (2011)
Shh	pBS-KS II	*HindIII*	T3	At 8.5 dpc: node, notochord and the floor plate	Echelard *et al.*, (1993)
Sp5	Blunt II-TOPO	*EcoRV*	SP6	At 8.5 dpc: in the midbrain, tailbud and neural plate	Weidinger *et al.*, (2005)

LPM: lateral plate mesoderm.

2.7 Scanning Electron Microscopy (SEM)

Mouse embryos at 8.0 dpc were dissected in 10% (v/v) fetal bovine serum in PBS and fixed overnight in fresh 2% paraformaldehyde/2.5% gluteraldehyde/0.1 M cacodylate buffer (pH 7.4) at 4°C. After rinsing with 0.1 M cacodylate buffer, embryos were postfixed in 1% OsO_4/0.1 M cacodylate for 20 minutes at room temperature. They were dehydrated through a graded EtOH series and dried at a critical point with a CPD010 (Balzers Union). Embryos were coated in platinum by an EMTECH K550X sputter coater. All imaging was performed on a Hitachi 4300SE/N FESEM at 3 kv.

2.7.1 SEM analysis

Node shape and border were outlined in Adobe Photoshop CS5 on a minimum of three embryos of *Ctnnb1*$^{+/+}$, *Ctnnb1*$^{Bfc/+}$ and *Ctnnb1*$^{Bfc/Bfc}$ at 8.5 dpc. Cilia length was measured in a minimum three embryos from each genotype (*Ctnnb1*$^{+/+}$, *Ctnnb1*$^{Bfc/+}$, and *Ctnnb1*$^{Bfc/Bfc}$) using SEM. For cilia length analysis, SEM images were recorded at 15,000× magnification. The file name of each image was altered to a number and the file order randomised by an independent worker so that the genotype of the embryo was unknown to the worker calculating cilia frequency and length. Cilia length was determined by measuring pixel length in Adobe Photoshop CS5 and conversion to µm (using a factor determined by the number of pixels per µm). In total, 10 node cilia were measured in a minimum of three embryos per genotype.

2.8 Statistical analysis

A two-way ANOVA with Fischers Unprotected LSD *post hoc* test was performed in Genstat in order to analyse cilia length. Statistical significance was determined when $p<0.05$. A Chi-square was performed in order to identify the linked regions on a

chromosome of the ENU induced mutation. Statistical significance was determined when $p<0.05$.

Chapter 3: The katun allele fails to inhibit canonical Wnt/β-catenin signalling *in vivo*

3.1 INTRODUCTION

3.1.1 The Zic gene family

The *Zic* (*z*inc finger of the cerebellum) gene family (named for the vertebrate tissue from which they were first isolated) has been found in a range of both vertebrate and invertebrate species (Aruga *et al.*, 1994). The initial homologous invertebrate gene was discovered in *Drosophil*a and named odd-paired (*opa*) (Benedyk *et al.*, 1994). This gene was isolated in a genetic screen for mutations that disrupt segmentation of the *Drosophila* embryo (Nusslein-Volhard *et al.*, 1980); it was proposed that it may encode a transcription factor, as it contains a zinc finger domain (ZFD) (Benedyk *et al.*, 1994). Flies with mutated *opa* failed to generate alternate segments during embryonic patterning and fell into the group of mutants known as pair-rule mutants. Further analysis of *opa* mutant flies revealed that the protein also regulates the expression of genes required for the development of the visceral mesoderm and midgut morphogenesis (Benedyk *et al.*, 1994).

Mice and humans have five paralogous *Zic* genes (named *Zic1-5*); each shows a high level of homology with the invertebrate *opa* gene and their respective *Zic* orthologues (Ali *et al.*, 2012). Mutational analyses in a variety of vertebrate and mammalian organisms have shown that *Zic* genes play important roles in the differentiation of ectoderm and mesoderm-derived structures during early embryogenesis (Houtmeyers *et al.*, 2013). However, much remains unknown about the molecular mechanisms of *Zic* gene function and regulation. Moreover, the long-standing uncertainty regarding whether ZIC proteins act as bona fide transcription factors, or as co-factors, is only now beginning to be

clarified with evidence that they can act as either class of molecule in a context-dependent manner (Figure 3.1) (Ali *et al.*, 2012, Ahmed *et al.,* 2016).

The first evidence that ZIC proteins may act as transcription factors or co-factors was found using a candidate approach, based on the fact that the ZIC ZFD shows high homology with GLI proteins (Aruga, 2004). GLI proteins are the vertebrate orthologues of the *cubitus interruptus* gene of *Drosophila*, which is a transcriptional mediator of the developmentally important Hedgehog signalling pathway (Koebernick *et al.*, 2002). ZIC proteins can bind GLI DNA binding sites (GLI-BS), albeit with much lower affinities, but do not stimulate transcription from GLI-BS (Mizugishi *et al.*, 2001). Furthermore, ZIC and GLI proteins can form heterodimers, where both proteins interact through their C2H2 zinc fingers (ZF) 3-5 (Koyabu *et al.*, 2001); however, no heterodimer formation has been observed between ZIC paralogues (Brown *et al.*, 2005). This interaction promotes the translocation of GLI proteins, which are predominantly localised in the cytoplasm, to the nucleus. Consequently, the ZIC-GLI heterodimers stimulate transcription at GLI-BS at higher levels than those observed in the presence of monomeric GLI proteins (Mizugishi *et al.*, 2001). Although GLI-ZIC heterodimers have never been purified from an *in vivo* model, *Zic3* loss-of-function rescues the digit phenotype of the *Gli3* mouse mutant known as ‘extra toes’, implying a functional relationship between these two molecules (Quinn *et al.*, 2012).

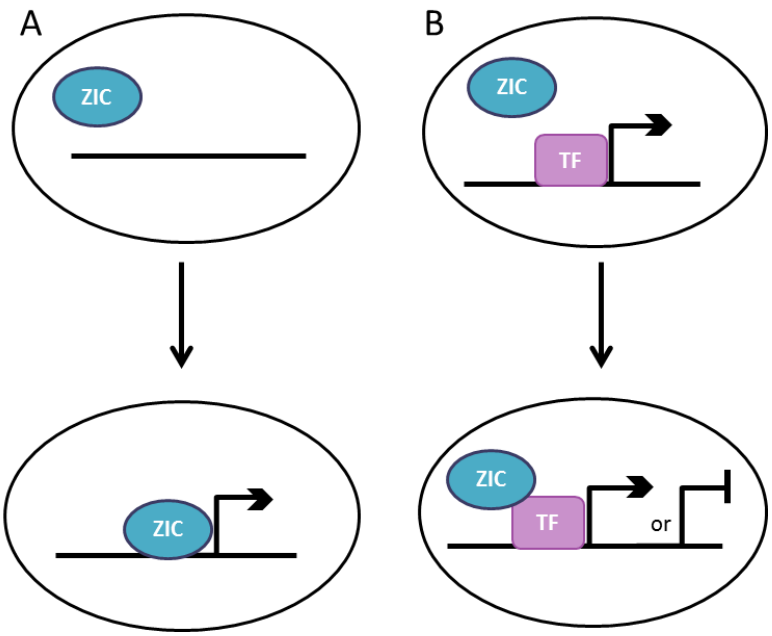

Figure 3.1: ZIC proteins can act as transcription factors and/or co-factors. (A) As transcription factors, ZIC proteins directly bind DNA. (B) As co-factors, ZIC proteins interact with and modulate the activity of other transcription factors (TF). Either scenario can result in enhancement or repression of target gene transcription, depending on the cellular context. Figure adapted from Ali *et al.* (2012).

Genomic arrangement

Phylogenetic sequence analysis has revealed that *Zic* genes have an unusual arrangement within the genomes of all vertebrates examined; the majority of genes are arranged as divergently transcribed gene pairs (Houtmeyers *et al.*, 2013). For example, in mammals, *Zic1* and *Zic4* exist as a divergently transcribed tandem gene pair (on human Chr3, mouse Chr9), as do *Zic2* and *Zic5* (on human Chr13, mouse Chr14), while *Zic3* is an X-linked singleton in both humans and mice (Figure 3.2) (Ali *et al.*, 2012, Houtmeyers *et al.*, 2013). Based on similarities in their nucleotide sequences, exon-intron boundaries, and protein sequences, the *Zic* genes are divided into two subclasses—subclass A and subclass B (Houtmeyers *et al.*, 2013). A hypothetical model of *Zic* gene evolution by Aruga *et al.* (2006) proposes that the ancestral single copy *Zic* gene contained one intron and a set of conserved structural domains. Tandem duplication of the ancestral gene and sequence modifications led to the generation of two isoforms: one that acquired an additional intron and another that underwent sequence divergence accompanied by the loss of certain structural domains. Additional genome duplication events presumably produced eight *Zic* genes organised into four gene clusters, with each cluster containing a *Zic* gene from subclass A (*Zic1*, *Zic2*, and *Zic3*) and subclass B (*Zic4* and *Zic5*). The maximum number of *Zic* genes identified in any species to date is seven in the zebrafish (Grinblat *et al.*, 1998, Toyama *et al.*, 2004). The variation in *Zic* gene copy numbers observed across vertebrate species is thought to have arisen through a combination of gene loss and *de novo* duplication events (Houtmeyers *et al.*, 2013).

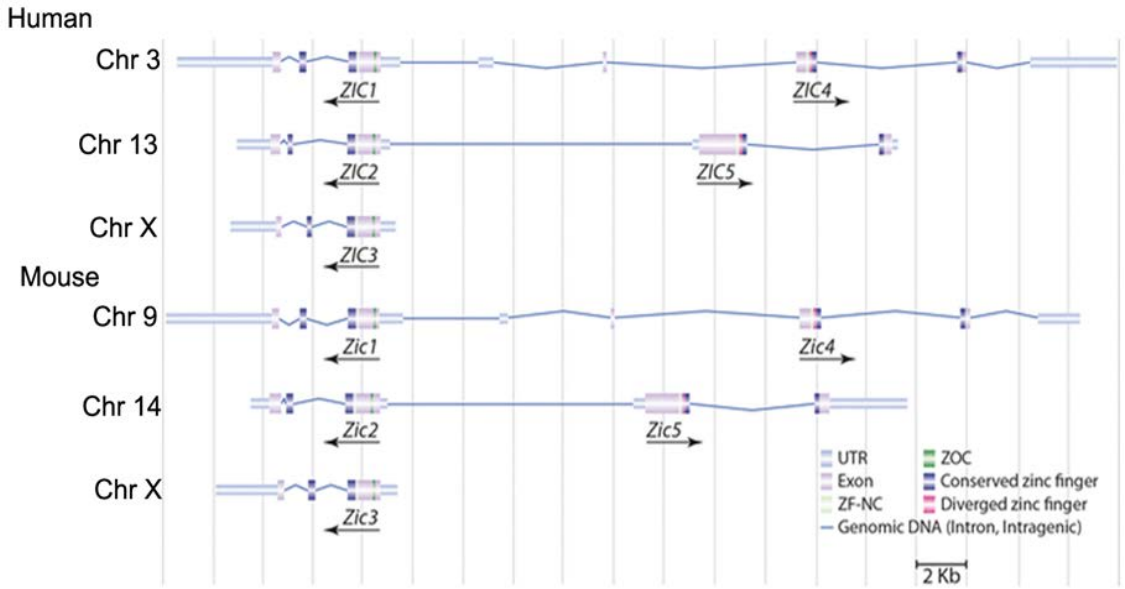

Figure 3.2: Genomic alignment of human and mouse *Zic* genes. The diagram illustrates the arrangement of *Zic* genes in both human and mouse genomes, as shown in the *Ensembl* genome browser. In humans, *Zic1* and *Zic4* are located on chromosome (Chr) 3, while *Zic2* and *Zic5* are located on Chr 13. In the mouse, *Zic1* and *Zic4* are located on Chr 9, while *Zic2* and *Zic5* are located on Chr 14. The *Zic3* gene is located on the X-chromosome in both human and mouse. UTR: untranslated region; ZF-NC: zinc finger N-terminal conserved region; ZOC: Zic-Opa conserved region; ZF: zinc finger. Figure adapted from Houtmeyers *et al.,* (2013).

The retention of some *Zic* genes as tandem pairs has implications for their function and study. This is because it is likely that the gene pairs share regulatory sequences and have overlapping expression patterns. For example, *Zic2* and *Zic5* show largely overlapping expression domains during murine gastrulation, with initial expression observed in the embryonic and extra-embryonic ectoderm (Furushima *et al.*, 2000, Elms *et al.*, 2004). This provides the potential for functional redundancy and/or other modes of paired activities, such as dominant-negative interference between co-expressed family members. Additionally, it is difficult to produce animal models in which more than one *Zic* gene is mutated; the common approach of crossing single mutant strains is impractical owing to the low probability of a crossover occurring between individual mutations.

The protein structure of the *Zic* gene family

The defining feature of ZIC proteins is the highly conserved ZFD which consists of five ZFs (Figure 3.3). This type of ZF contains two cysteine and two histidine residues that ligate the zinc ion (Zn^{2+}), enabling the protein to bind DNA (Iuchi, 2001). ZF1 is highly conserved in ZIC1, ZIC2, and ZIC3, but shows more divergence in ZIC4 and ZIC5 (Aruga *et al.*, 2006); this difference forms part of the distinction between subclass A and B (Table 3.1). It has been shown that the ZIC ZFD can bind DNA and participate in protein-protein interactions in a variety of contexts. It also appears to be responsible for localising ZICs to the nucleus, although this has only been demonstrated for ZIC3 (Bedard *et al.*, 2007, Houtmeyers *et al.*, 2013). Therefore, mutations within the ZFD can drastically affect the function of these proteins (Bedard *et al.*, 2007).

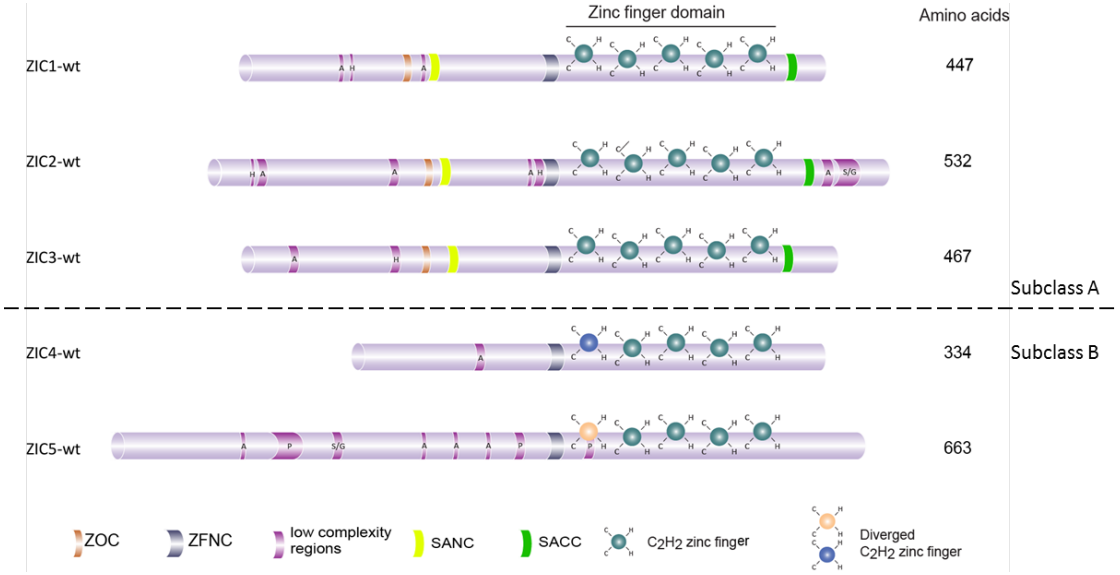

Figure 3.3: Schematic illustration of the protein structure of the human ZICs. All human ZICs contain a highly conserved zinc finger domain made of five C2H2-type zinc fingers. The *Zic* genes are divided into two subclasses (Subclass A and B). ZIC4 and ZIC5 exhibit a divergence in the first zinc finger. ZIC1, ZIC2, and ZIC3 share the ZOC, SANC, and SACC domains, but these are not found in ZIC4 or ZIC5. All five proteins contain a short highly conserved domain immediately upstream of the zinc fingers, called the ZFNC domain, as well as low complexity regions; the major amino acid found at each low complexity region is shown by its associated letter (A: alanine tract, H: histidine tract, P: proline tract, and S/G: serine/glycine tract, SANC: subclass A N-terminal conserved, SACC: subclass A C-terminal conserved). Figure adapted from Houtmeyers *et al.* (2013).

Table 3.1: Structural features of the two subclasses of ZIC proteins.

	ZF-CWCH2	ZFNC	ZOC	Low comp.regoins	ZF 1	ZF 2-5	SANC	SACC
Subclass A ZIC1-3	Conserved	Conserved	Conserved	Conserved	Conserved	Conserved	Conserved	Conserved
Subclass B ZIC4-5	Conserved	Conserved	Not-Consereved	Conserved	Diverged	Conserved	Not-Consereved	Not-Consereved

ZF: zinc fingers; ZFNC: zinc finger N-terminal conserved region; ZOC: Zic-Opa conserved region; low comp.regoins: low complexity regions; SANC: Subclass A N-terminal conserved; SACC: Subclass A C-terminal conserved.

Biochemical experiments have revealed some information regarding the DNA- and protein-binding activities of ZFs. Structural analysis of ZIC3, for instance, suggests that ZF1 and ZF2 may not be canonical (DNA-binding) C2H2-type ZFs. Both ZFs form a single structural unit called the tandem CWCH2 motif, which is characterised by the presence of a tryptophan (W) residue between the two canonical cysteines of each ZF (Aruga *et al.*, 2006). This motif is present in the first and second ZFs of ZIC3 in a wide range of metazoan species; this may have a biological impact, as a missense mutation of the tryptophan in ZF1 of ZIC3 is associated with congenital heart malformation (Chhin *et al.*, 2007). Although the molecular role of the CWCH2 motif is unknown, it is proposed that the motif modulates the DNA-binding capability of the other canonical ZFs, and/or participates in protein-protein interactions (Houtmeyers *et al.*, 2013). Furthermore, mutational analysis of *Zic2* has revealed that mutation of ZF4 ablates DNA-binding (Brown *et al.*, 2005) and causes *in vivo* loss-of-function (Elms *et al.*, 2003) without necessarily affecting protein-protein interactions (Pourebrahim *et al.*, 2011). Further work is required, however, to clearly delineate individual and cooperative ZF functions.

The N- and C-terminal sequences that flank the ZFs vary significantly amongst family members. There is one short region (14–21 amino acid) at the N-terminus of the ZFD—the zinc-finger N-terminal conserved region (ZFNC)—that is found in all family members across a wide range of species (Aruga *et al.*, 2006), but the function of this region is currently unknown. Additionally, a small (9-10 amino acids) N-terminally located domain is conserved in the subclass A ZIC proteins. The *Drosophila* orthologue, OPA, also contains this domain; hence, it was named the ZIC/Odd-paired conserved (ZOC) motif (Landy, 1989, Aruga *et al.*, 1996). This domain is required for protein-protein interactions (Mizugishi *et al.*, 2004, Himeda *et al.*, 2013). Aside from these conserved

domains, several low complexity domains, including poly-alanine, -histidine, -proline, and -serine/-glycine tracts have been found within individual ZICs (Figure 3.3). Although the function of these domains in ZICs is not well characterised, expansion mutations in the C-terminal alanine tract of ZIC2 (Brown *et al.*, 2005) and the N-terminal alanine tract of ZIC3 (Wessels *et al.*, 2010) are associated with pathological effects in humans. In addition, there is evidence that the number of protein binding partners is generally higher in proteins which contain low-complexity domains at the terminal regions (Coletta *et al.*, 2010), suggesting that these domains confer protein binding capabilities. Moreover, the Arkell laboratory has identified two previously unannotated regions that are only conserved in subclass A ZIC proteins. The first domain is a ten-amino acid sequence in the N-terminal region referred to as the subclass A N-terminal conserved (SANC) region. The second domain is a fourteen-amino acid sequence in the C-terminal region referred to as the subclass A C-terminal conserved (SACC) region. These two conserved domains contribute to the trans-activation of *Zic3* in a context-dependent manner (Ahmed *et al*., 2016).

3.1.2 The Zic3 gene

The *Zic3* gene is X-linked

The *Zic3* gene is the only mammalian *Zic* gene family member that is X-linked and unpaired in both human and mouse. X-linkage results in a specific pattern of inheritance, as X-linked genes have no allelic counterpart in males: a condition known as hemizygosity. Different progeny classes are generated depending on whether the mutation is passed through the female or the male (Figure 3.4). Both hemizygous males and heterozygous females inherit one copy of the mutant allele. However, in general, male carriers display more severe phenotypes than females. This is due to a

combination of two factors: firstly, male cells have only one copy of the X-chromosome, meaning that hemizygous mutations are functionally equivalent to a homozygous autosomal mutation. Secondly, female cells have only one active X-chromosome; owing to the random nature of X-inactivation, female tissues will contain mosaics of cells that are either wild-type or mutant.

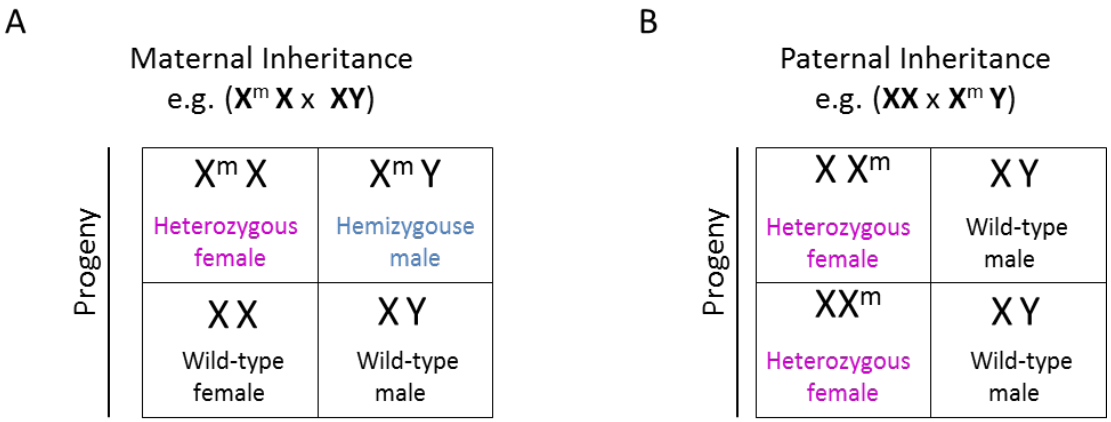

Figure 3.4: Schematic illustration of X-linked inheritance. A) When the *Zic3* mutation is passed through a carrier female ($X^{m}X$), the potential progeny are heterozygous female, hemizygous male, or wild-type male and female. B) When the *Zic3* mutation is passed through a hemizygous male ($X^{m}Y$), he produces only heterozygotic females and wild-type males who cannot transmit the mutation.

***Zic3* expression during murine gastrulation**

The expression pattern of murine *Zic3* has previously been examined by the Arkell laboratory and published in Elms *et al.* (2004). The information below has been reproduced in its entirerty from the publication and is included here for ease of reference. Their findings showed that expression of *Zic3* is initiated prior to gastrulation; transcripts were found throughout the extra-embryonic ectoderm of 5.0 dpc stage embryos and within some epiblast cells (Figure 3.5 A). In all pre-gastrulation stages examined, distal epiblast cells did not express *Zic3*, and there was no expression in the endoderm. This was confirmed through sectioning of pre-gastrulation stage embryos (Figure 3.5 B–D). This pattern of expression is maintained until the onset of gastrulation (Figure 3.5 E, F). Upon initiation of gastrulation, expression in the ectoderm recedes from the anterior half portion of the embryo. Instead, *Zic3* is expressed in the region of the embryo lined by the mesoderm, emerging from the primitive streak to encircle the embryonic portion of the egg cylinder (compare Figure 3.5 E, G). The mesoderm that moves into the extra-embryonic region does not contain *Zic3* transcripts, and expression at the posterior end of the embryo during this stage occurs in the ectoderm of the posterior amniotic fold (arrow in Figure 3.5 G). Whole-mount *in situ* hybridisation showed that mid- and late-streak stage embryos express *Zic3* at the furthest anterior region of the embryo: at the junction between the embryonic and extra-embryonic region. Sectioning of embryos revealed that this expression occurs in the anterior definitive endoderm (Figure 3.6 H, P). As the node forms at the anterior end of the primitive streak, it can be seen that the node cells at the external surface of the embryo do not express *Zic3*, although the ectoderm in this region contains transcripts (Figure 3.5 H). As gastrulation proceeds, the pattern of *Zic3* expression remains relatively constant; transcripts are located in the wings of the embryonic mesoderm, as well as

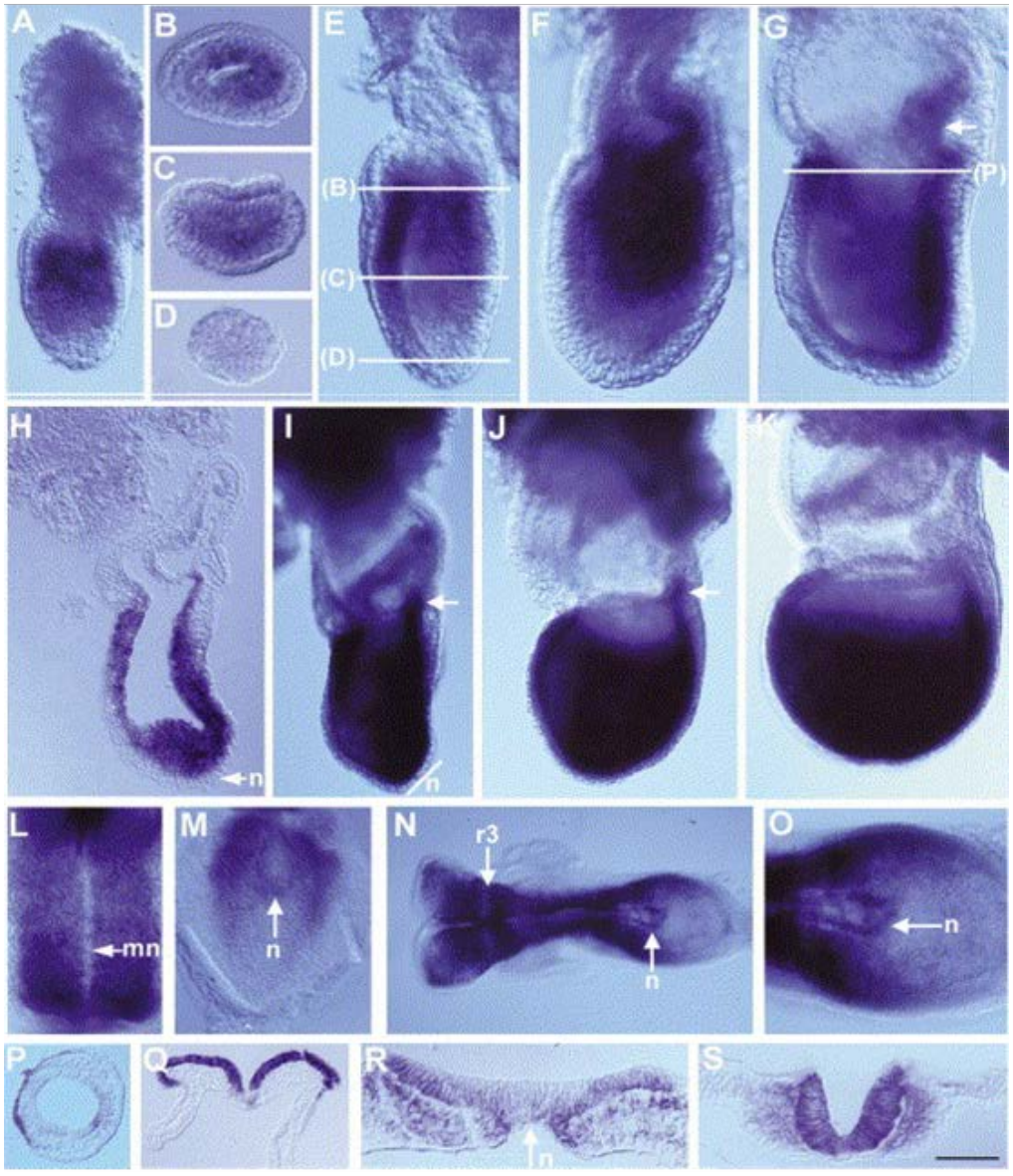
A
B
C
D
E
(B)
(C)
(D)
F
G
(P)
H
n
I
n
J
K
L
mn
M
n
N
r3
n
O
n
P
Q
R
n
S

Figure 3.5: *Zic3* expression during mouse gastrulation. In all pictures of intact embryos at stages where the anterior can be distinguished (after 6.0 dpc), anterior is on the left, except L and M, where anterior is at the top. Anterior is on the left in the sections shown in H and P. The intact embryos in A–K are shown from the lateral view; those in L–O are shown from the dorsal view. (A) A 5.0 dpc pre-streak stage embryo. (B) A transverse section through the extra-embryonic region of a pre-streak stage embryo at the level shown in E. (C) A transverse section through the proximal embryonic region of a pre-streak stage embryo at the level shown in E. (D) A transverse section through the distal embryonic region of a pre-streak stage embryo at the level shown in E. (E) A 6.0 dpc pre-streak stage embryo. (F) A 6.25 dpc embryo prior to the onset of gastrulation. (G) A 6.75 dpc mid-streak stage embryo, showing formation of the posterior amniotic fold. (H) A longitudinal section through the 7.0 dpc late-streak stage embryo shown in I. (I) A 7.0 dpc late-streak stage embryo. (J) A 7.25 dpc early allantoic bud stage embryo. (K) A 7.75 dpc early headfold stage embryo. (L) A flat-mount of the trunk neurectoderm of a headfold pre-somite stage embryo showing diminished transcription in the midline neurectoderm. (M) The posterior of a headfold pre-somite stage embryo with transcripts in the neurectoderm and at the edge of the node. (N) A 7-somite stage embryo; note regionalisation of expression in the cranial neurectoderm. (O) A higher power view of the posterior of a 7-somite stage embryo in which some transcripts are visible at the edge of the node and some signal is observed within the node itself. (P) A transverse section through the embryonic portion of a 6.75 dpc mid-streak stage embryo at the level shown in G. (Q) A transverse section through the trunk of a 1-somite stage embryo. (R) A transverse section through a 7-somite embryo at the level of the node. (S) A transverse section through the trunk of a 7-somite embryo anterior to the node. mn: midline neurectoderm, r3: prospective rhombomere 3, n: node. Scale bars: 50 µm (A), 90 µm (B–D), 100 µm (E–G), 150 µm (H–M, O), 300 µm (N), 170 µm (P), 120 µm (Q–S). Figure reproduced from Elms *et al.* (2004).

the primitive streak and ectoderm adjacent to the expressing mesoderm (Figure 3.5 I–K). In the posterior region of the embryo, expression can be seen in the ectoderm (arrow in Figure 3.5 I, J). The mesoderm in this region moves into the extra-embryonic portion of the embryo and does not express *Zic3*, as shown in Figure 3.5 H. *Zic3* expression was not observed in the ventral node at any of these primitive streak stages, or in the axial mesendoderm, which emerges from the node.

As the headfolds develop, *Zic3* expression is lost at the midline neurectoderm cells along the length of the axis, but continues throughout the rest of the neural plate (Figure 3.5 L, Q). In the cranial region, neurectoderm expression of *Zic3* begins to become regionalised, with areas of lower signal including the prospective midbrain and future hindbrain (Figure 3.5 N). In early somite embryos, the hindbrain contains one region with considerably lower expression levels. This region is immediately caudal of the periotic sulcus that is known to mark the junction of future rhombomere 2 and 3 (Trainor *et al.*, 1995); it has been identified as the prospective rhombomere 3. In later stage embryos, when rhombomeres can be clearly distinguished, it can be seen that *Zic3* transcripts are also down-regulated in rhombomere 5, which was identified by the location of the otic sulcus that marks the rhombomere 5/6 boundary (Trainor *et al.*, 1995). Additionally, during headfold stages, *Zic3* transcripts disappear from the primitive streak, but can be found in a group of cells anterior to the primitive streak that constitute the node (Figure 3.5 M). Node expression continues, becoming more prominent in embryos with seven somites (Figure 3.5 N, O, R) and is still present in embryos with twelve somites (data not shown). Inspection of node expression in both whole-mount specimens and sections failed to detect any evidence of asymmetric localisation of *Zic3* transcripts. During the somite stages of development, *Zic3* expression is also detected

in the paraxial mesoderm at the level of the node and, more rostrally, in the segmenting paraxial mesoderm (Figure 3.5 R, S).

Zic3 and *situs ambiguus* (heterotaxy)

Most organisms have several internal asymmetries, though they display outward symmetry. Humans, for example, have unpaired organs (such as the heart or liver) and asymmetrically paired organs (such as the lungs). These asymmetries derive from an initial symmetry-breaking event during early embryogenesis, which establishes the embryonic left-right (L–R) axis, distinguishing the left side of the embryo from the right side. Errors in establishing these differences can lead to laterality disorders known as heterotaxy (Figure 3.6), which is a Greek term meaning 'other arrangement' (Ware *et al.*, 2004). This condition is characterised by abnormally positioned thoracic and/or visceral organs, which can cause multiple congenital malformations (Sutherland *et al.*, 2009). The major cause of morbidity and mortality are complex cardiovascular malformations which are sometimes accompanied by defects in the pulmonary, gastrointestinal, genitourinary, immune and musculoskeletal systems (Kearns-Jonker, 2006). Epidemiological surveys show that 3% of congenital heart defects are caused by heterotaxy, which has an estimated prevalence of 1:10,000 live births (Lin *et al.*, 2000). The true prevalence may be underestimated because of the wide spectrum of possible phenotypic abnormalities, variable expressivity, and incomplete penetrance associated with laterality defects.

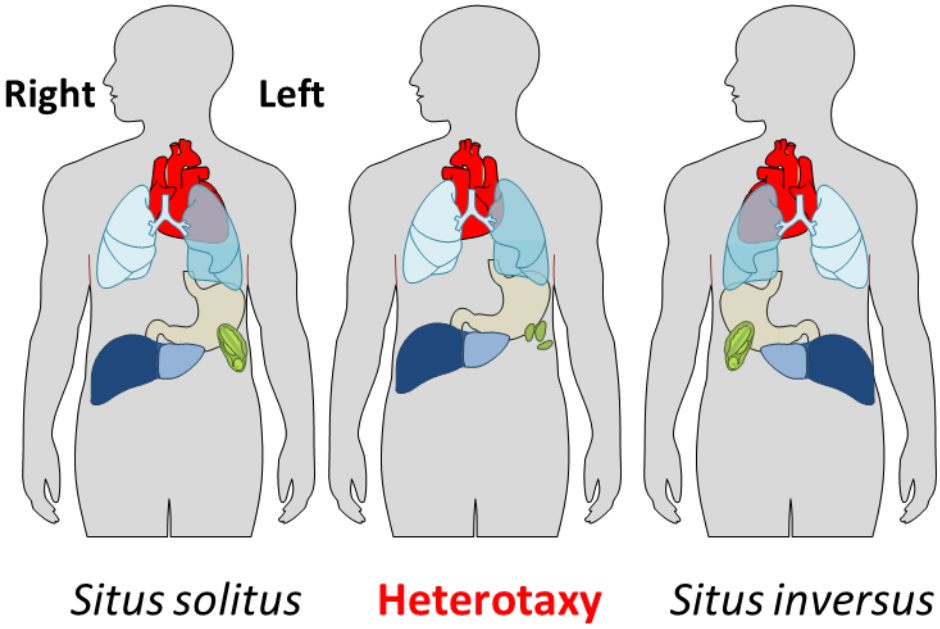

Figure 3.6: Schematic of *situs ambiguous*. Comparison between the normal arrangement of visceral organs (known as *situs solitus*—where the heart, stomach and spleen are on the left side and the liver on the right), heterotaxy (also known as *situs ambiguous*—where the heart and visceral organs are oriented independently of each other), and *situs inversus* (mirror image reversal of all organs). Figure adapted from R. Arkell.

The most common causes of heterotaxy are random genetic and environmental factors. Associated environmental factors include maternal diabetes, maternal cocaine use, and monozygotic twinning (Kuehl *et al.*, 2002). Mutational analysis in humans has identified heterozygous defects in a variety of autosomal genes (such as *Nodal*, *Cfc-1*, *ActrIIB*, *Foxh1* and *LeftyA*) in affected children and family members (Carmi *et al.*, 1992, Gebbia *et al.*, 1997). Despite the complications associated with studying these genes and their roles in L–R axis formation owing to species-specific differences in gene function; animal models have greatly advanced our understanding of this phenomenon (Kearns-Jonker, 2006). In addition, there are more than 100 genes in animal models that are likely to play a role in the development of human heterotaxy, as they have been shown to be required for the formation of the L–R axis in these models (Sutherland *et al.*, 2009).

Mutations in the ZIC3 transcription factor are most frequently associated with human heterotaxy (Gebbia *et al.*, 1997, Sutherland *et al.*, 2009), contributing to 5% of all familial cases of heterotaxy and 1% of cases of sporadic heterotaxy (Ware *et al.*, 2004). In humans, twenty-five different mutations in the *Zic3* gene have now been definitively linked to heterotaxy (Figure 3.7) (Gebbia *et al.*, 1997, Megarbane *et al.*, 2000, D'Alessandro *et al.*, 2013, Cowan *et al.*, 2014). As expected of an X-linked gene, male (hemizygous) carriers of *Zic3* mutations exhibit *situs ambiguus* heterotaxy with varying degrees of severity, and are affected more severely than heterozygous females (Lin *et al.*, 2000). However, some male carriers (from families with a history of ZIC3-related congenital defects) have been found to show normal phenotypes (Megarbane *et al.*, 2000, D'Alessandro *et al.*, 2013), perhaps indicating that other proteins can, in some cases, compensate for ZIC3 function. Females heterozygous for deleterious *Zic3* mutations are usually asymptomatic; however, one family has been found to contain

females with *situs inversus* who carry the heterozygous mutation (Gebbia *et al.*, 1997). The different phenotypes in females may be due to skewed X-inactivation, where one X-chromosome carrying normal *Zic3* is inactivated while the other X-chromosome carrying the mutated allele remains active (Chhin *et al.*, 2007). Notably, this pattern of inheritance exhibited by *Zic3,* along with the fluctuation of phenotypes observed, makes diagnosis difficult. It is difficult to pinpoint the specific tissues and the exact stage of development where ZIC3 function is required using human clinical data.

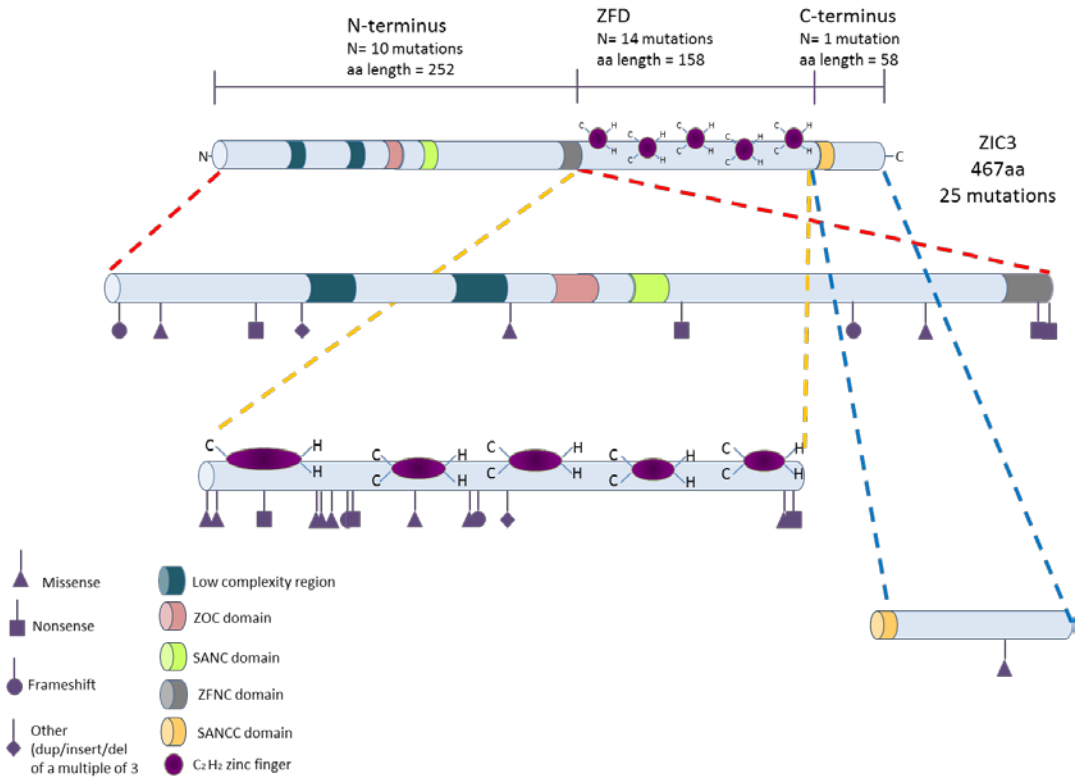

Figure 3.7: A schematic showing the location and class of all *Zic3* heterotaxy-associated human mutations. There are twenty-five known heterotaxy-associated human *Zic3* mutations. The N-terminus (defined as the first 252 amino acids before the zinc finger domain [ZFD]) contains ten mutations, the ZFD contains fourteen mutations, and the C-terminus (defined as the 58 amino acids following the ZFD) contains one mutation. There are notable shifts in the type of mutations that occur in the different regions of the protein. C: Cysteine, H: Histidine, aa: amino acid, Other: insertions, duplications, and deletions, N: number, SANC: subclass A N-terminal conserved, SACC: subclass A C-terminal conserved. Figure adapted from Koula Diamand.

The exact manner by which this occurs remains controversial. This asymmetric gene expression is thought to be induced by directional flow of extracellular fluid at the node (see 1.1.4 for a description of node formation), a process called nodal flow is shown in Figure 3.8. Two models have been proposed for how nodal flow is generated and perceived. First is the morphogen gradient model (McGrath and Brueckner, 2003). This model proposes the node cell monocilia moves a soluble morphogen towards the left side of the node. This establishes a concentration gradient of molecular signals, which then interact with their corresponding receptors on the left border of the node, resulting in signal transduction and the asymmetric expression of genes such as *Nodal* and *Pitx2*. Several secreted signals are produced in the node, such as Nodal (Brennan *et al.,* 2002), GDF-1 (Rankin *et al.,* 2000) and FGF-8 (Meyers and Martin 1999), however, no evidence for the differential distribution of these proteins in the node in response to nodal flow has been observed.

The second model, known as the two cilia model, proposes that the motile cilia of the pit cells direct fluid leftward which deflects left-side crown cell (primary, immotile) cilia. These cilia are thought to be mechanosensory. The bending of the node monocilia against their cell bodies triggers an increase in intracellular Ca^{2+} signalling in cells on the left margin of the node (McGrath *et al.,* 2013), which subsequently results in the asymmetric expression of genes such as Nodal on the left border of the node. This model is supported by the finding that the polycystic kidney disease 2 (*Pkd2*) gene and Pkd1-like 1 (*Pkd1l1*) gene are expressed in the ciliate cells of the node and are required for response to nodal flow and L-R aixs determination in the mouse (Pennekamp *et al.,* 2002, Field *et al.,* 2011). In the kidney, *PKD2* and the related molecule PKD1, form a mechanotransducing complex that detects urine flow and generates a Ca^{2+} signal in

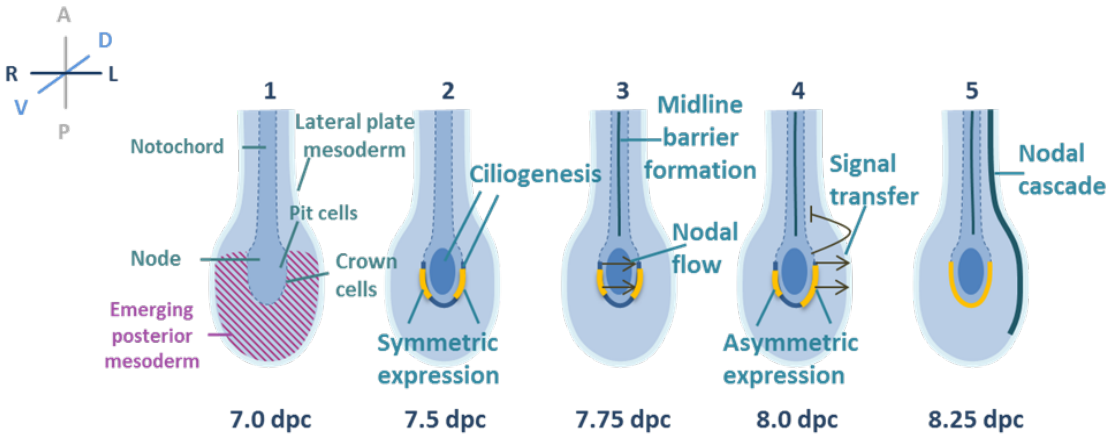

Figure 3.8: A schematic of L–R axis formation in the mouse embryo. 1: The node is induced to form at mid-gastrulation (7.0 dpc) and, in the next 24 hours, develops into a shallow, crescent-shaped depression on the ventral side of the embryo. The cells within the depression are termed pit cells, whereas those that form the raised surface that rings the pit are called crown cells (Norris, 2012). **2**: A monocilium, extending into the extracellular space, forms on the apical surface of each pit and crown cell; these become posteriorly polarised over time. Signalling molecules (such as *Nodal*) are expressed in the crown cells. **3**: The cilia of the pit cells rotate in a clockwise direction, directing first a disorganised, followed by a laminar leftward flow of extracellular fluid within the node. It is posited that this leftward nodal flow is sensed by crown cell cilia, prompting a Ca^{2+} flux in the left crown cells which modifies gene expression (Norris, 2012). By the end of this 24-hour period (at 8.0 dpc) the first known asymmetries in gene expression are detected within node crown cells (*Dand5,* and soon thereafter *Nodal*, become asymmetrically expressed in the node crown). A Wnt ligand is also asymmetrically expressed and canonical signalling then amplifies the initial *Dand5* asymmetry (Nakamura *et al.*, 2012). **4**: The asymmetric signal/s are propagated to the left lateral plate mesoderm (LPM) and prevented from spreading to the right LPM by the recently formed midline barrier. **5**: Nodal signalling in the left LPM controls its own expression, and that of other molecules which ultimately direct organ placement and other asymmetries (Norris, 2012). A: anterior, P: posterior, D: dorsal, V: ventral, R: right, L: left. This figure adapted from R. Arkell.

response (Nauli *et al.,* 2003). By analogy it is proposed that that PKD2 and PKD1L1 form a complex in the leftsided crown cell cilia and upon deflection by nodal flow generate at Ca^{2+} signal. Despite the observation of asymmetric calcium signalling at the node of dissected 7.75 dpc mouse embryos and its disruption in mutant embryos (McGrath *et al.,* 2003), however, Delling *et al*. (2016) reported that the crown cell cilia specific Ca^{2+} release was not observed at a physiological level of fluid flow when measured via a transgenic calcium indicator.

For both models, the outcome is an asymmetric (L greater than R) signal at the node which stimulates enhanced left-sided expression of the *Nodal* secreted growth factor and simultaneous decreased left-sided expression of the secreted *Dand5* nodal antagonist. The down-regulation of *Dand5* mRNA levels is accomplished by degradation of *Dand5* mRNA via the 3' UTR at the left side of the node in a calcium-dependent manner (Nakamura *et al.,* 2012). This suggests a potential role of the RNA-binding protein BICC1 because embryos null for this gene have defective L-R development (Maisonneuve *et al.,* 2009). Subsequently, an asymmetric Nodal signal is relayed from the node to the left lateral plate mesoderm (LPM) via an unresolved mechanism, resulting in the activation of the Nodal cascade in the left LPM and asymmetric development of organs along the L-R axis (Meno *et al.*, 1998, Nonaka *et al.*, 1998, Nonaka *et al.*, 2002, McGrath *et al.*, 2003, Shiratori *et al.*, 2006).

Despite the long association in both humans and mice between *ZIC3*/*Zic3* and heterotaxy, the precise role of this transcription factor in heterotaxy remains unknown. *Zic3*-null embryos are reported to exhibit random expression of *Nodal* at the early somite stage (Purandare *et al.*, 2002) and it has been proposed that *Zic3* directly up-regulates *Nodal* mRNA expression during the head-fold stages of embryogenesis (Ware

et al., 2006a). However, in recent studies, conditional inactivation of *Zic3* using a variety of Cre strains demonstrated that L–R axis defects results only when *Zic3* function is ablated at 7.0 dpc in the nascent mesoderm (using *T*Cre) (Sutherland *et al.*, 2009, Jiang *et al.*, 2013).

In addition, it is expressed in an array of developing embryonic tissues, such as those that form the central nervous system, muscles, and aspects of the visual system (Herman *et al.*, 2002). Overall, it seems that further analysis is needed of the *Zic3* mutant phenotype and of the molecular mechanisms of *Zic3*.

Mouse models of *Zic3* dysfunction

Multiple mouse strains with *Zic3* loss-of-function mutations are available (Table 3.2). In each case, mutant embryos displayed similar abnormalities to those observed in human patients with heterotaxy (Purandare *et al.*, 2002), indicating that the mouse is an excellent system for examining the function of ZIC3 in human heterotaxy. Precisely how ZIC3 acts to prevent heterotaxy and its relevance in developmental processes is yet to be investigated. Mouse embryos have become the mammalian model for investigating congenital defects in humans because of the practical and ethical constraints on studying human embryogenesis (Bedell *et al.*, 1997). There is a significant degree of sequence conservation between the human and mouse genomes. Thus, there are many mutations in the orthologous mouse genome and these mirrors the human phenotype. The human *ZIC3* gene, for instance, is 1,404 bp long with a protein product (ZIC3) of 467 amino acids; the mouse *Zic3* gene is 1,401 bp long with a protein product of 466 amino acids (Figure 3.9). The amino acid sequences of the human and mouse proteins show 98% identity, with 99.4% homology between their ZFDs. Therefore, due to genetic similarities, the mouse is a useful model organism for studying human genetic disorders and mammalian developmental genetics (Malakoff, 2000).

Table 3.2: Murine alleles for *Zic3* (downloaded from MGI: http://www.informatics.jax.org).

MGI Allele ID	Allele Symbol	Allele Name	Synonyms	Allele Type	Allele Attributes
MGI: 1856679	*Zic3^{Bn}*	Bent tail	*Bn*	spontaneous	Large deletion
MGI: 3043028	*Zic3^{Ka}*	Katun	*Ka*	spontaneous	Nonsense
MGI: 5476827	*Zic3$^{tm1.1Smwa}$*	Targeted mutation 1.1, Stephanie M. Ware	*Zic3^{flox}*	targeted	Conditional ready, no functional change
MGI: 2180720	*Zic3^{tm1Bca}*	Targeted mutation 1, John W. Belmont	*Zic3^{-}*, *Zic3^{null}*	targeted	Null/knockout, reporter
MGI: 3698161	*Zic3$^{tm1.Jwb}$*	Targeted mutation 1, John W. Belmont	*Zic3^{neo}*, *Zic3^{null}*	targeted	Null/knockout
MGI: 5470150	*Zic3$^{tm2.1.Jwb}$*	Targeted mutation 2.1, John W. Belmont	*Zic3^{flox}*	targeted	Conditional ready, no functional change

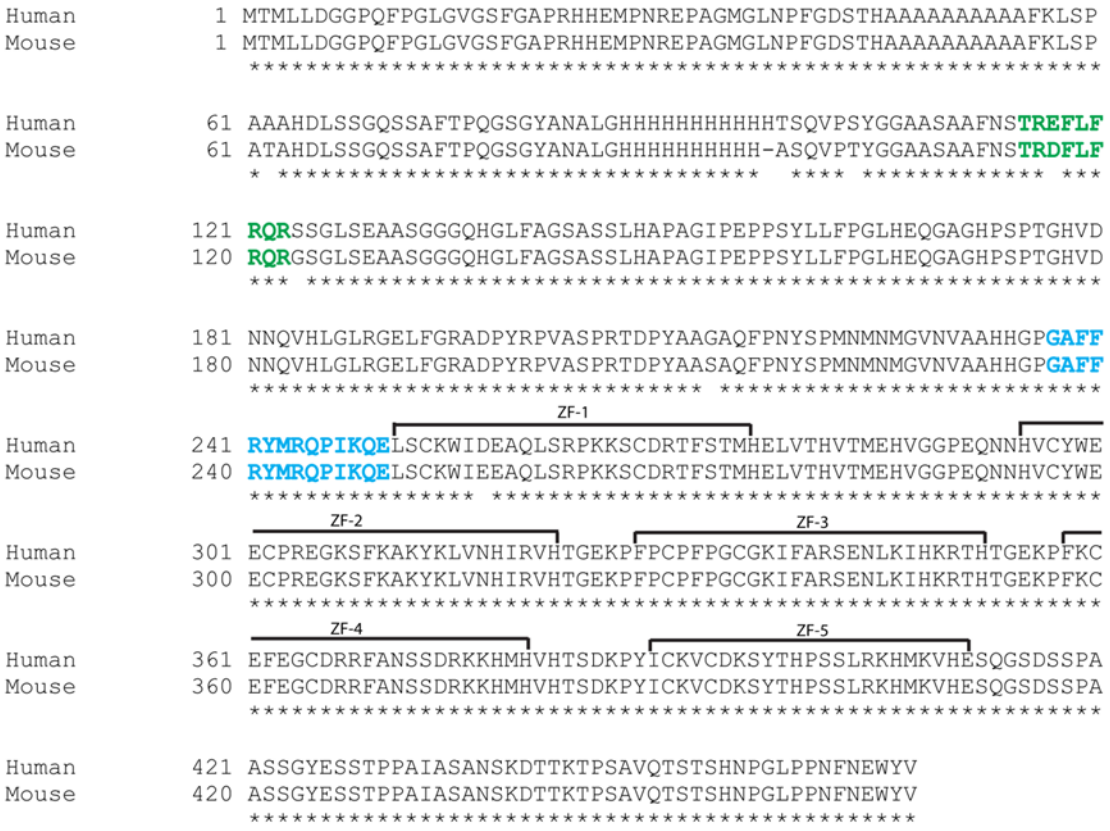

Figure 3.9: Sequence alignment of human and mouse ZIC3 proteins. The evolutionarily conserved five functional C2H2 zinc fingers (ZFs) are highly similar between human and mouse. Amino acids coloured in green and blue are the evolutionarily conserved ZIC/Odd-paired conserved (ZOC) and zinc finger N-terminal conserved (ZFNC) domains, respectively. The five C2H2 ZFs are indicated (ZF-1–5). Asterisk (*) = identical residues.

The katun mouse strain is a null allele of the *Zic3* gene

The katun (*Ka*) mouse mutant was isolated in an ENU mutagenesis screen designed to assess dominant mutations for semi-dominant phenotypes (Bogani *et al.*, 2004). In the pedigree generated within the dominant screen, it was observed that several G1 progenies exhibited a bent tail phenotype. This suggests that the mutation arose not as a consequence of ENU exposure in the male, as repeat mutations are not generally recovered with high frequency (Rinchik and Carpenter, 1999), but was instead a spontaneous, incompletely penetrant mutation in one of the parents (Bogani *et al.*, 2004). Inheritance of the mutation was consistent with X-linkage; the phenotypic similarities between *Ka* and deletion of the entire *Zic3* gene (known as the *Bn* allele) indicated that *Ka* was likely to be a new allele of *Zic3*. The coding region of *Zic3* was amplified by touchdown PCR from the genomic DNA of a heterozygous female, a hemizygous male, a C3H/HeH mouse, and a Balb/c mouse and sequenced. This revealed a single base change of G to T at nucleotide position 1,283 of accession no. NM_009575 in the affected mice (Figure 3.10 A-C). This base change converts codon 248 from a glutamate to a stop codon, generating a nonsense mutation that introduces a premature termination codon (PTC) (Figure 3.10 D).

Similar to the phenotypes caused by the *Bn* allele (Garber, 1952, Carrel *et al.*, 2000), the *katun* phenotype is incompletely penetrant. Females heterozygous for *katun* ($Zic3^{Ka/+}$) are genetically viable but may exhibit minor defects such as kinked tails (Figure 3.10 G). In contrast, hemizygote male mutants ($Zic3^{Ka/Y}$) are more severely affected; their body symmetry is altered and they exhibit heart defects in terms of the L–R axis (Bogani *et al.*, 2004). As the phenotype is incompletely penetrant, however, not all mutants display these defects to the same extent. Homozygous *katun* females

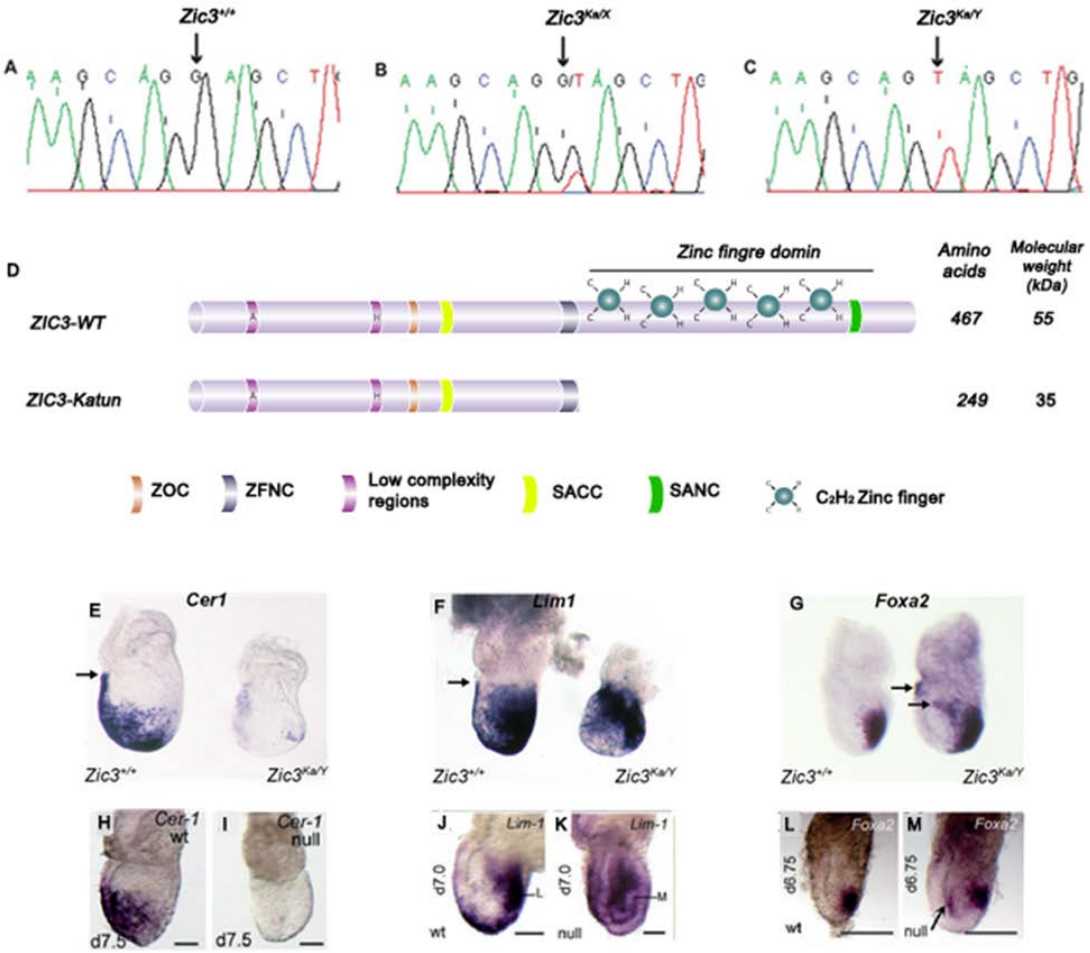

Figure 3.10: The *katun* mutation is a null allele for *Zic3*. This mutation converts codon 248 to a stop codon. (A-C) Sequence traces from exon 1 of murine *Zic3* gene of (A) wild-type (*Zic3*$^{+/+}$), (B) heterozygous female (*Zic3*$^{Ka/X}$), and (C) hemizygous male (*Zic3*$^{Ka/Y}$) mice with arrows indicating the altered base. (D) Protein structure of ZIC3-wt and ZIC3-katun. The katun protein lacks the entire zinc finger domain (ZFD) and C-terminus. (E-M) Stage-matched *Zic3*$^{Ka/Y}$ (E-G) and *Zic3*-null (H-M) embryos showing altered primitive streak, mesoderm, and endoderm formation. A lateral view of embryos is shown with anterior on the left, following whole-mount *in situ* hybridisation (WMISH) of the RNA named in each panel. (E-K) *Cer1* and *Lim1* are markers for endoderm and mesoderm formation. In *Zic3*$^{+/+}$ embryos, these markers are expressed in the anterior visceral endoderm (black arrow). In *Zic3*$^{Ka/Y}$ and *Zic3*-null embryos, *Cer1* expression is lacking, while *Lim1* expression is reduced. (G, L, M) *Foxa2* shows primitive streak formation. *Zic3*$^{Ka/Y}$ and *Zic3*-null embryos display ectopic expression of *Foxa2* going into the amniotic cavity (black arrows). Panels (A-G) were adapted from Ahmed *et al.* (2013). Panels (H-M) were adapted from Ware *et al.* (2006b).

(*Zic3*$^{Ka/Ka}$) exhibit similar defects to those seen in hemizygote males (Ahmed *et al.*, 2013).

Two sets of analyses conducted by the Arkell laboratory (Ahmed *et al.*, 2013) established that *katun* is a null allele. Molecular evidence indicated that the predicted protein (i.e. ZIC3 E249X) does not trans-activate and is unable to elicit transcription at an *Apoe* promoter in cell reporter assays. In addition, unlike the ZIC3 wild-type protein, the *katun* protein was unable to inhibit Wnt-dependent β-catenin-mediated transcription in a TopFlash assay. Moreover, phenotypic analysis of *katun* mutant mice revealed a reduction in the expression of markers for both endoderm and mesoderm formation (Ahmed *et al.*, 2013), similar to previous results in mice with other *Zic3*-null alleles (Figure 3.10 E-M) (Ware *et al.*, 2006b). As the *katun* allele behaves as a *Zic3*-null mutation in both human cells and mouse models, it can be used to elucidate the molecular role of ZIC3 during gastrulation and determine the mechanisms of ZIC3-associated heterotaxy.

3.1.3 The Wnt signalling pathway

Wnt signalling refers to several pathways activated by secreted Wnt ligands which have been separated into two classes: classical Wnts (Wnt1, -3a, -8, and -8b), which activate signalling through the canonical pathway; and non-classical Wnts (Wnt4, -5a, and -11), which activate signalling through non-canonical pathways such as the planar cell polarity (PCP) pathway and the Wnt/calcium signalling pathway (Rao *et al.* 2010). All three pathways require the binding of a Wnt ligand to a Frizzled (Fz) receptor, which sends a signal to Dishevelled (Dvl) in the cell cytoplasm. In the canonical pathway, the binding of Wnt to Fz results in the accumulation of β-catenin in the cytoplasm and its eventual translocation to the nucleus to activate the transcription of Wnt target genes. When

Wnt ligands are not bound, however, β-catenin is degraded via a destruction complex made up of Axin, glycogen synthase kinase 3 (GSK3), and adenomatous polyposis coli (APC). In contrast, the non-canonical pathways do not require β-catenin to function. The PCP pathway establishes cellular polarity by asymmetrical and polarised protein localisation and regulates polarised cell movements such as gastrulation and neural crest migration during embryonic development (Simons *et al.* 2008). The Wnt/calcium signalling pathway involves Wnt activating calcium-dependent pathways, which are critical to dorsal-ventral axis formation, intracellular calcium levels, and cardiac cell development (Purandare *et al.*, 2002).

Canonical Wnt/β-catenin signalling

Studies have shown that canonical Wnt/β-catenin signalling plays an essential role during all stages of life, from early embryogenesis to adult homoeostasis (Clevers *et al.*, 2012). In the early developing embryo, canonical Wnt/β-catenin signalling directs cell fate determination, migration, and proliferation in the coordination of tissues, as well as organ development during the establishment of the body axis (Hierholzer *et al.*). Mutations in this canonical Wnt pathway, therefore, have potentially lethal consequences for the developing embryo, including gastrulation defects (Morkel *et al.*, 2003), axis formation failure (Huelsken *et al.*, 2000), and embryonic lethality (Haegel *et al.*, 1995). In later life, canonical Wnt signalling also plays a critical role in tissue homoeostasis, cell renewal, regeneration, and the maintenance of bone health (Valenta *et al.*, 2012). Therefore, mutations in the canonical Wnt pathway during adulthood can contribute to a significant number of cancers and bone diseases (Nusse, 2005, Grigoryan *et al.*, 2008).

The central protein in canonical Wnt signalling— β-catenin—is encoded by the *Ctnnb1* gene, which consists of a central region of twelve 'armadillo' repeats flanked by N- and C-terminal trans-activation domains (TADs). These domains contain various phosphorylation and trans-activation sites that regulate β-catenin function (Valenta *et al.*, 2012). Because β-catenin does not contain a DNA-binding domain, the N- and C-terminal TADs of the armadillo repeats are required to support the various DNA-binding transcriptional co-activators that facilitate β-catenin-mediated transcription. This enables the cell-cell adhesion and signalling functions of β-catenin, such as transcription of Wnt target genes in the nucleus (Mosimann *et al.*, 2009).

As a cell adhesion molecule, β-catenin binds to the cytoplasmic domain of cadherin proteins involved in cell-cell adhesion. The β-catenin protein is essential for strong cadherin-mediated cellular adhesion, as it indirectly ties cadherin proteins on the membrane to the actin cytoskeleton to facilitate signalling events (Heuberger *et al.*, 2010). However, in the nucleus, β-catenin functions as a transcriptional co-activator by interacting with DNA-binding co-activator molecules to activate the transcription of Wnt target genes, owing to the lack of β-catenin DNA-binding sites (Henderson *et al.*, 2002, Bienz *et al.*, 2003). The primary nuclear partners of β-catenin that activate most canonical Wnt signalling pathways are the T-cell factor (TCF)/Lymphoid enhancer-binding factor 1 (LEF1) family. When, in the absence of a Wnt ligand, Wnt is not stimulated, TCF proteins act as transcriptional repressors by forming a complex with Groucho/TLE repressors (Figure 3.11 A) (Arce *et al.*, 2006). However, when Wnt is stimulated in response to Wnt ligand binding, the interaction of β-catenin with the N-terminus of TCF physically displaces the Groucho protein and, in doing so, transiently converts TCF into a transcriptional activator (Figure 3.11 B).

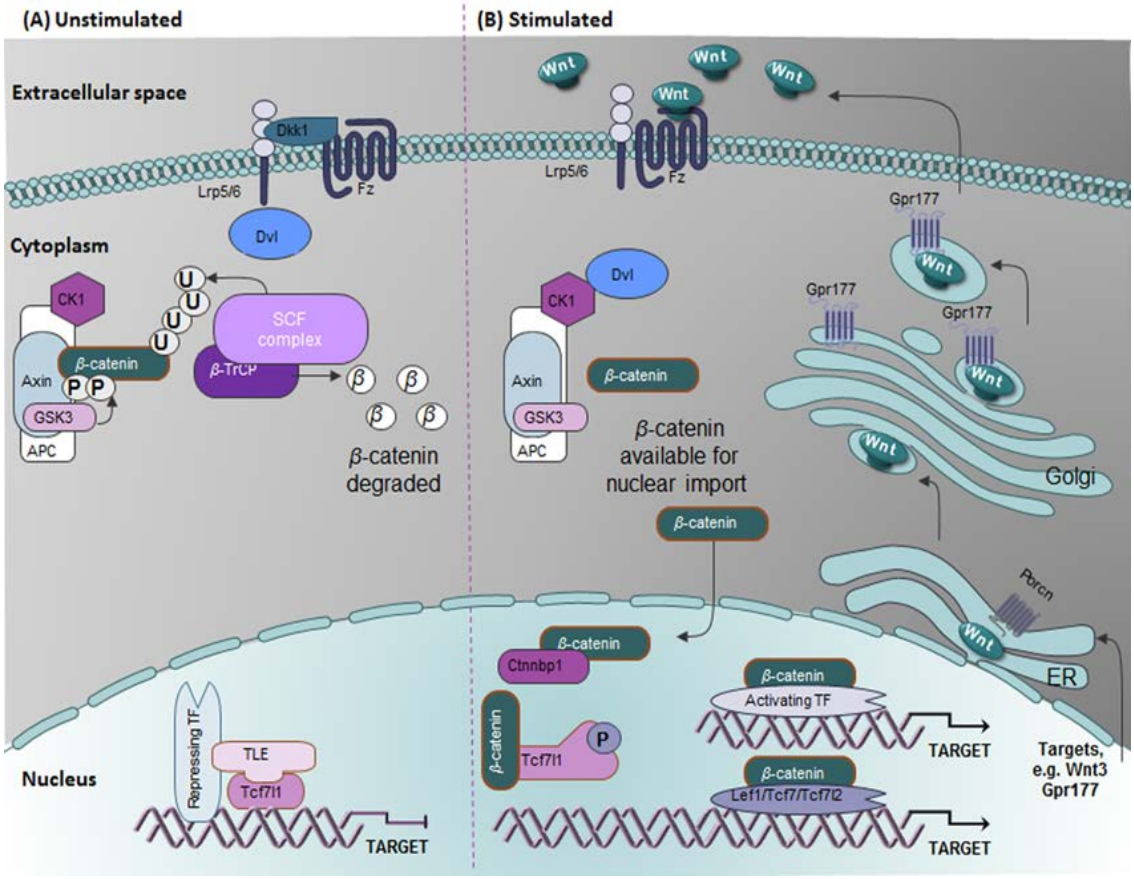

Figure 3.11: A schematic of canonical Wnt/β-catenin signalling. A) In the absence of a Wnt signal, β-catenin is targeted by a coordinated phosphorylation complex consisting of the scaffolding proteins AXIN and adenomatous polyposis coli (APC). AXIN and APC promote phosphorylation of β-CATENIN through the action of casein kinase 1 (CK1) and glycogen synthase kinase 3 (GSK3), which targets β-CATENIN for ubiquitin dependent degradation via the ubiquitin ligase β-TrCP (Cadigan and Peifer, 2009; MacDonald *et al.*, 2009). Without nuclear β-catenin, T-cell factor (TCF) proteins act as transcriptional repressors of target genes in the nucleus. B) When Wnt is stimulated by an extracellular Wnt protein binding to the Frizzled (Fz) receptor and the LRP5/6 co-receptor on the surface of cells, a Wnt/Fz/LRP receptor complex is formed, resulting in phosphorylation of the intracellular portion of LRP5/6; this, in turn, stops the β-CATENIN degradation complex in a Dishevelled (Dvl)-dependent manner (Cadigan and Peifer, 2009; MacDonald *et al.*, 2009; Kim *et al.*, 2013). Subsequently, β-CATENIN is transported into the nucleus, where it accumulates and binds to the TCF protein, subsequently acting as a transcription co-factor to activate target genes. Figure adapted from Arkell *et al.* (2013).

Elevation of Wnt/β-catenin signalling phenotypes

Wnt/β-catenin signalling activity must be tightly controlled during embryogenesis, as indicated by the dramatic phenotypes produced by aberrant Wnt signalling in mice. The absence of *Wnt3a* results in the loss of all caudal structures below the forelimbs (Takada *et al.*, 1994); embryos lacking the Wnt repressor *Dkk1* exhibit partial to complete head truncation (Mukhopadhyay *et al.*, 2001, Fossat *et al.*, 2011); and both *Dkk1* and *Axin* mutants exhibit duplicated axes (Zeng *et al.*, 1997) (Fossat *et al.*, 2011). Similarly, loss-of-function of the murine *TCF7L1* gene increases Wnt/β-catenin signalling during gastrulation, resulting in duplication of axis structures such as the node and notochord. This suggests that, in some contexts, the main function of the TCF/LEF is to repress *Wnt* target genes (Merrill *et al.*, 2004).

Wnt/β-catenin signalling can be inhibited at all levels of the signalling pathway. Proteins such as Wnt Inhibitory Factor (WIF) and secreted frizzled-related protein (SFRP), which are secreted extra- or intracellularly, contain regions mimicking Wnt receptors and can, therefore, bind and modulate the activity/stability of Wnt ligands (Bovolenta *et al.*, 2008). For instance, secreted proteins such as WISE and DKK inhibit Wnt/β-catenin signalling by interacting with the LRP5/6 receptor, thereby disrupting the formation of the Wnt/Fz/LRP receptor complex (Semenov *et al.*, 2001, Itasaki *et al.*, 2003). Wnt/β-catenin signalling antagonists have also been identified at the nuclear level. These proteins, which include CHIBBY, inhibitor of β-catenin (ICAT), and transcription factor 4 (TCF4), negatively regulate Wnt/β-catenin signalling in two ways: by disrupting the interaction of β-catenin with TCF, and by promoting the export of β-catenin from the nucleus (MacDonald *et al.*, 2009). Although β-catenin degradation usually occurs in the cytoplasm, it has also been suggested that SOX9 and OCT3/4 can downregulate Wnt/β-catenin signalling levels by causing degradation of β-catenin in the nucleus (Topol *et al.*,

2009, Abu-Remaileh *et al.*, 2010).

ZIC proteins can inhibit β-catenin-mediated transcription *in vitro*

Recently, two independent studies (using fish embryos, frog embryos, and cultured mammalian cells) demonstrated that ZIC proteins can inhibit TCF-dependent transcription, which is normally stimulated by Wnt signalling (Pourebrahim *et al.*, 2011, Fujimi *et al.*, 2012). Importantly, the study by Pourebrahim *et al.* demonstrated that this repression occurs because ZIC2 interacts with TCFs, rather than directly binding the DNA of Wnt targets; this represents a role for the ZIC proteins as co-factors in the Wnt/β-catenin pathway and provides further evidence that they are multifunctional proteins.

Both Wnt pathway components and ZICs show similar loss-of-function phenotypes. For instance, posterior axis duplications have been observed in *Zic3*-null mouse embryos (Ware *et al.*, 2006b), similar to defects previously associated with elevated/ectopic Wnt signalling (Popperl *et al.*, 1997, Zeng *et al.*, 1997). In addition, the kinked tail exhibited in adult *Zic3 katun* heterozygous and hemizygous mice (Bogani *et al.*, 2004, Ahmed *et al.*, 2013) was also seen in mice heterozygous for the Wnt-activating gain-of-function *Lrp6* allele *gwazi* (Bogani *et al.*, 2004, Fossat *et al.*, 2011). Moreover, it was found that ZIC3 prevented the formation of an additional axis by supressing β-catenin in *Xenopus* embryos (Fujimi *et al.*, 2012).

The Batface mouse strain

The batface (*Bfc*) mouse strain was first identified via an ENU recessive genetic screen (Nolan *et al.*, 2000); the causative mutation was found to be a C to A point mutation at nucleotide position 1,958 of the β-catenin gene, *Ctnnb1* (Figure 3.12 A) (Fossat *et al.*, 2011). The resulting amino acid substitution of threonine (T) to lysine (K) at residue 653 (T653K) falls within the last armadillo repeat and is a key trans-activation region for the

β-catenin protein (Figure 3.12 B) (Fossat *et al.*, 2011). In contrast to β-catenin-null mutant embryos, heterozygous *Bfc* mutant embryos survived beyond early embryogenesis and exhibited varying degrees of craniofacial and head formation defects (Figure 3.12) (Fossat *et al.*, 2011).

Despite the potential value of *Bfc* mutants, the effects of this mutation on β-catenin function, as well as the impact of the mutation on Wnt signalling, have not been quantitatively determined. Previous studies of early stage *Bfc* embryos crossed with the Tcf-Lef-LacZ reporter strain (BATGal) showed elevated levels of Wnt signalling activity during gastrulation compared to stage-matched wild-type embryos (Figure 3.12 C) (Fossat *et al.*, 2011). The elevated activity of the *Bfc* allele was confirmed and quantified using a TopFlash assay by the Arkell laboratory (Arkell lab, unpublished data); however, the underlying mechanism for this elevation is yet to be elucidated. It has been previously shown that replacement of a neutral tyrosine (Y) residue with a charged glutamic acid (E) residue at position 654 of β-catenin mimics phosphorylation of Y654 and results in increased Wnt-dependent transcription (Piedra *et al.*, 2001, van Veelen *et al.*, 2011). Because the *Bfc* mutation involves replacing a neutral threonine (T) residue with a charged lysine (K) residue at position 653 of the β-catenin protein, it is hypothesised that it acts by mimicking or enhancing phosphorylation of the neighbouring residue—Y654—by altering the structure of β-catenin to increase Wnt signalling activity.

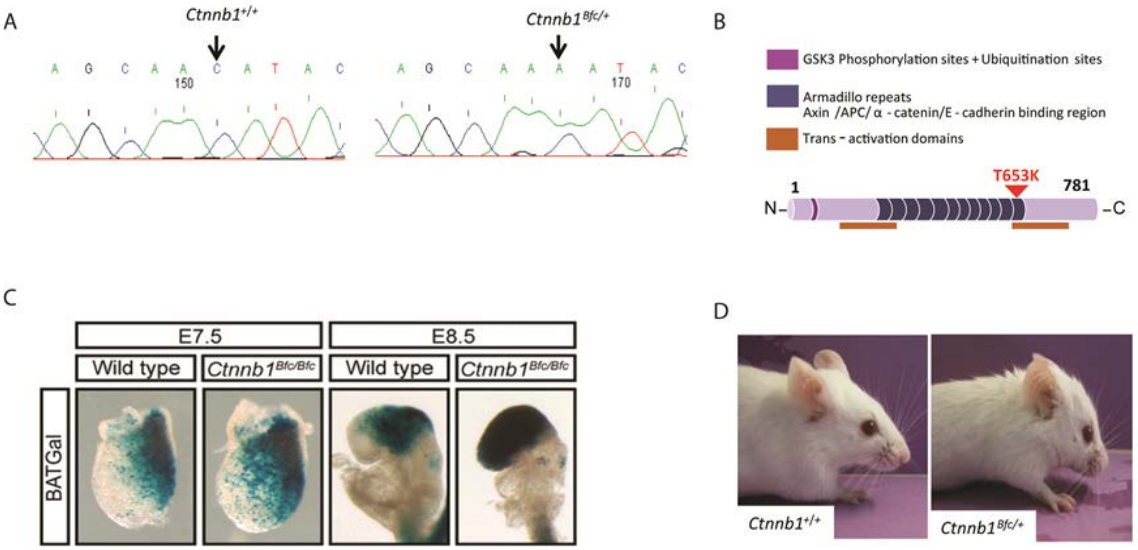

Figure 3.12: The Batface mouse. (A) A point mutation within the *Ctnnb1* gene (C to A) results in a threonine (T) to lysine (K) amino acid substitution. (B) The position of the *Bfc* mutation within the last armadillo repeat (red arrow) and C-terminal trans-activation domain of the β-catenin protein at position 653. (C) A later view of mouse embryos with anterior on the left. Stages are indicated, showing expression of the Tcf-Lef-lacZ reporter (BATGal) in both wild-type and *Ctnnb1*$^{Bfc/Bfc}$ mutant embryos. (D) Mice heterozygous for the *Bfc* mutation display craniofacial and head formation defects, including a squashed face, compared to the wild-type. (A-C) adapted from Fossat *et al.* (2011).

3.1.4 Aims

The overall aim is to characterise the phenotype of the *katun* mouse mutant that harbours a *Zic3*-null allele and to investigate the underlying molecular causes of the resultant phenotype. The possibility that the *katun* mutants fails to inhibit Wnt/β-catenin signalling during early embryogenesis will be investigated. Additionally the possibility that elevated Wnt/β-catenin signalling can cause L-R axis defects will be examined by phenotyping the *Bfc* allele of β-catenin known to have elevated canonical Wnt activity at gastrulation.

3.2 RESULTS

3.2.1 The *Zic3* loss-of-function phenotype is indicative of elevated canonical Wnt activity at gastrulation

As described in the introduction, several experimental systems demonstrate the ZIC proteins can inhibit Wnt/β-catenin dependent transcription via physical interaction with TCF proteins. Published data from the Arkell laboratory has shown the katun protein variant of murine ZIC3 is unable to inhibit Wnt/β-CATENIN dependent transcription in HEK293T cells (Ahmed *et al.*, 2013). The corresponding mouse strain (a null allele of *Zic3*) therefore provides the opportunity to test whether ZIC3 inhibits Wnt signalling *in vivo*, since loss-of-function of negative regulators of canonical Wnt signalling produces a consistent and well documented phenotype (Arkell *et al.*, 2013). Furthermore, previous analysis of the murine *Zic3* loss-of-function phenotype has already revealed features found in mouse strains harbouring activating canonical Wnt pathway alleles. For example, animals heterozygous for mutations in *Zic3* exhibit tail kinks; a phenotype repeatedly associated with elevated canonical Wnt signalling (Bogani *et al.*, 2004, MacDonald *et al.*, 2004, Ahmed *et al.*, 2013, Norum *et al.*, 2015, Doi *et al.*, 2016). Some subtle phenotypic features in embryos that lack *Zic3* are also shared with strains that have elevated canonical Wnt activity, such as the ectopic expression of *Foxa2* in the gastrula seen in embryos that lack *Zic3* (Ware *et al.*, 2006, Ahmed *et al.*, 2013) and in embryos that lack *Tcf7l1* (Merrill *et al.*, 2004).

To further assess whether elevated canonical Wnt signalling may contribute to the *Zic3* phenotype, embryos were specifically inspected for two features well known to result from elevated Wnt signalling at gastrulation: forebrain truncation and posterior midline defects. Such as seen in mutation in the Wnt signalling inhibitor *Fused* locus encoded

Axin (*Axin*^{Fu}) exhibited loss of its anterior structure and exhibiting posterior axis defects (Zeng *et al.,* 1997). WMISH of 9.5 dpc embryos to genes that delineate axial tissues (*T;* notochord Herrmann *et al.,* (1990), *Fst;* somites and brain structures, *Fgf8*; anterior neural ridge and midbrain/hindbrain junction [Mahmood *et al.,* 1995]) enabled scoring of these phenotypes. As shown in Table 3.3 and Figure 3.13, embryos heterozygous for the *Ka* allele of *Zic3* do not show these features, whereas a proportion of embryos hemizygous null for *Zic3* (*Zic3*$^{Ka/Y}$) for this allele exhibit mild anterior truncation (7%, N= 3/45). Additionally, *T* expression revealed that a subset of hemizygous null embryos (16%, N= 7/45) have bifurcations and/or small regions of duplicated trunk notochord. To further investigate the axis defects, embryos were harvested at 8.5 dpc (Figure 3.14) and WMISH was conducted for the secreted molecule sonic hedgehog (*Shh*), which is normally expressed in the node and notochord of the headfold stage embryo as well as in the floor plate and definitive endoderm during the early somites stages (Echelard *et al.*, 1993). This experiment revealed that the notochord bifurcations and small duplications observed following *T* WMISH were already present at earlier stages. Additionally, the node is irregularly shaped in hemizygous mutant embryos (data not shown), a feature consistent with previously published *Zic3* mutants (Sutherland *et al.*, 2013) and one that may underlie the axis defects seen in mutant embryos. Together the analysis implies that the *Zic3* phenotype, at least in part, can be attributed to elevated Wnt signalling at gastrulation and that one role of ZIC3 is to repress Wnt target gene expression.

Table 3.3: The frequency of phenotypes typically associated with elevated canonical Wnt signalling in 9.5 dpc *katun, Bfc* and compound mutant embryos.

Genotype	No.	No. with axis defects (%)	No. with anterior truncation (%)
$Zic3^{+/+}$	69	0 (0)	0 (0)
$Zic3^{Ka/X}$	36	0 (0)	0 (0)
$Zic3^{Ka/Y}$	45	7 (16)*	3 (7)
$Ctnnb1^{Bfc/+}$; $Zic3^{+/+}$	21	0 (0)	4 (19)*
$Ctnnb1^{Bfc/+}$; $Zic3^{Ka/X}$	15	4 (25)*	8 (50)*

* Significantly different from the wild-type at $p<0.001$ by $X^{2.}$

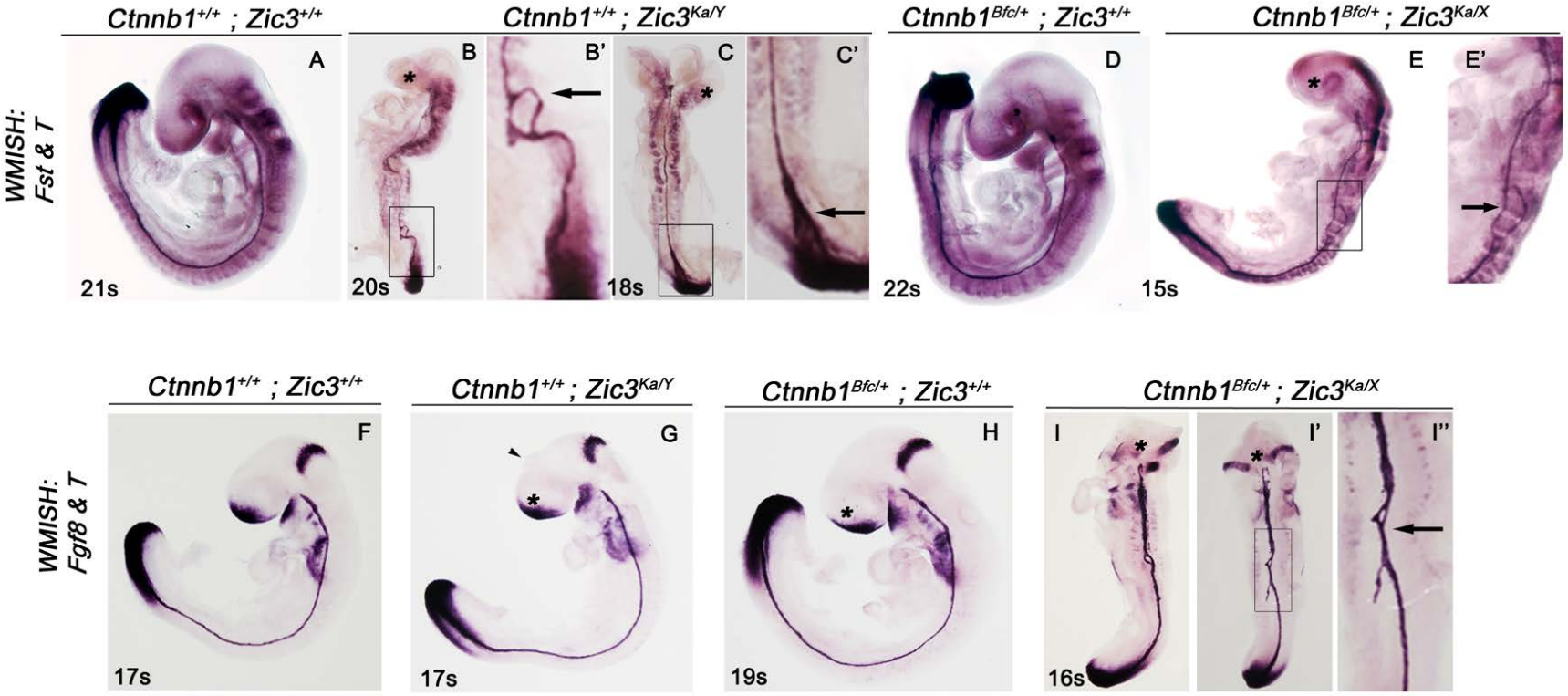
Ctnnb1+/+ ; Zic3+/+
Ctnnb1+/+ ; Zic3Ka/Y
Ctnnb1Bfc/+ ; Zic3+/+
Ctnnb1Bfc/+ ; Zic3Ka/X
WMISH: Fst & T
A
B
B'
C
C'
D
E
E'
21s
20s
18s
22s
15s
Ctnnb1+/+ ; Zic3+/+
Ctnnb1+/+ ; Zic3Ka/Y
Ctnnb1Bfc/+ ; Zic3+/+
Ctnnb1Bfc/+ ; Zic3Ka/X
WMISH: Fgf8 & T
F
G
H
I
I'
I''
17s
17s
19s
16s

Figure 3.13: Cross between *Ctnnb1*$^{Bfc/+}$ and *Zic3*$^{Ka/X}$ produces mutants consistent with an elevated canonical Wnt/β-catenin signalling phenotype. Mouse embryos following WMISH to *Brachyury* (*T*) and either *Follistatin* (*Fst*) (A-E') or *Fibroblast Growth Factor 8* (*Fgf8*) (F-I''). Stages are indicated on each panel. (A) Lateral view of a wild-type embryo with normal expression of *T* in the notochord and tail bud. (B) Lateral view of a *Ctnnb1*$^{+/+}$*; Zic3*$^{Ka/Y}$ embryo showing abnormal head development (asterisk) and axis duplication. The region within the rectangle in B is expanded in B' showing the axis duplication as indicated by black arrow. (C) Dorsal view of a *Ctnnb1*$^{+/+}$*; Zic3*$^{Ka/Y}$ embryo exhibiting a third head fold (asterisk) and an axis duplication close to the tailbud. The region within the rectangle in C is expanded in C' showing the axis duplication as indicated by a black arrow. (D) Lateral view of a *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{+/+}$ embryo showing a normal posterior axis. (E) Lateral view a *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{Ka/X}$ embryo showing a small head structure (asterisk) and an axis duplication. The region within the rectangle in E is expanded in E' showing the axis duplication as indicated by black arrow. (F) Lateral view of a wild-type embryo with normal expression of *Fgf8* in the forebrain. (G) Lateral view of a *Ctnnb1*$^{+/+}$*; Zic3*$^{Ka/Y}$ embryo showing reduced expression of *Fgf8* in the forebrain (asterisk) and a protrusion anteriorly to the head structure (black arrowhead). (H) Lateral view of a *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{+/+}$ embryo showing reduced expression of *Fgf8* in the forebrain (asterisk). Lateral (I) and dorsal (I') views of a *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{Ka/X}$ embryo exhibiting an absence of several head components (asterisk) and an axis duplication. The region within the rectangle in I' is expanded in I'' showing the axis duplication as indicated by black arrow. S: somites.

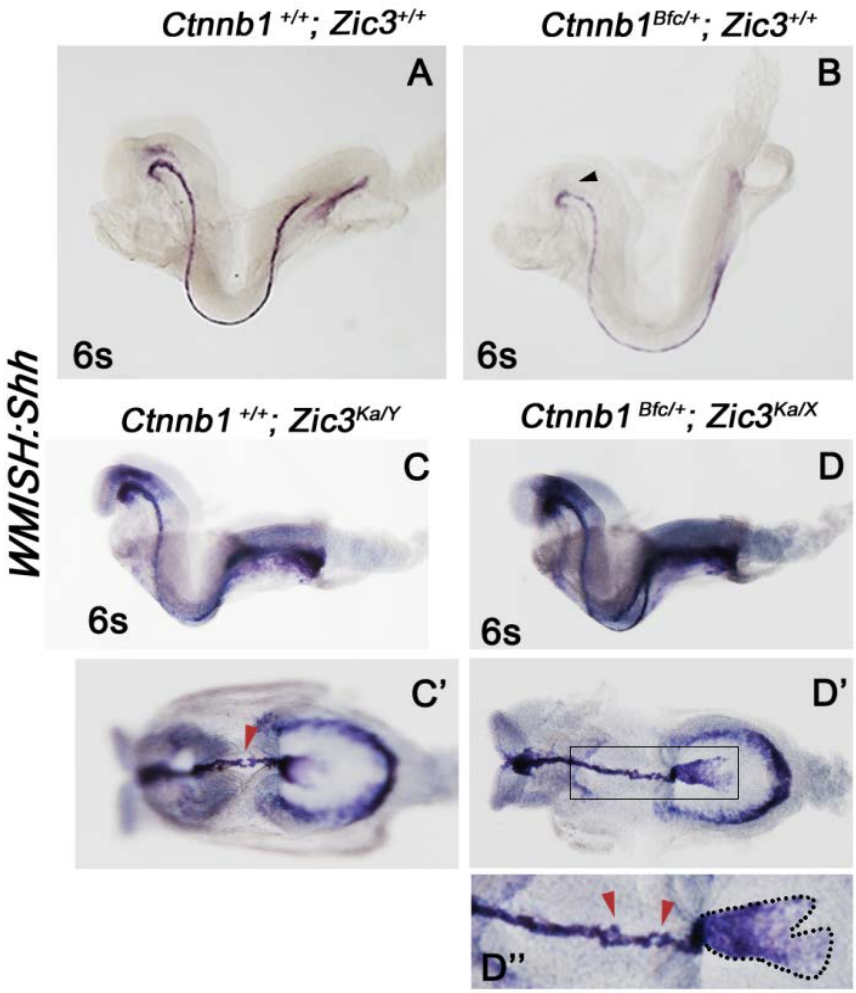

Figure 3.14: A partial axis duplication in the mutant embryos. Lateral view (A-D) and dorsal view (C'-D'') of mouse embryos at 6 somites following WMISH to *Sonic hedgehog* (*Shh*). Anterior is to the left side in all images. (A) A wild-type embryo with normal *Shh* expression in the notochord, in the floor plate and in the midbrain ventral neuroectoderm. (B) *Ctnnb1*$^{Bfc/+}$; *Zic3*$^{+/+}$ embryo with absence of *Shh* expression in the midbrain ventral neuroectoderm (black arrowhead). (C-C') *Ctnnb1*$^{+/+}$; *Zic3*$^{Ka/Y}$ embryo with reduced mid- and forebrain and axis duplication indicated by the red arrowhead. (D-D') *Ctnnb1*$^{Bfc/+}$; *Zic3*$^{Ka/X}$ embryo with reduced mid- and forebrain and an axis duplication. The region within the rectangle in D' is expanded in D'' showing two separate axis duplications as indicated by red arrowheads and an abnormal node shape (dotted line). S: somites.

3.2.2 Elevation of canonical Wnt signalling exacerbates the Zic3 loss-of-function phenotype

As described in the introduction, the *Ctnnb1* gene encodes β-CATENIN (the key transcriptional mediator for the activity of the Wnt/β-CATENIN signalling cascade) (Petersen *et al.*, 2009) and the batface (*Bfc*) mouse strain harbours a gain-of-function allele of *Ctnnb1* (Nolan *et al.*, 2000, Fossat *et al.*, 2011). The excess Wnt/β-CATENIN signalling activity at gastrulation manifests as anterior truncation in homozygous *Bfc* mutant embryos but not as axial defects (Fossat *et al.*, 2011). If ZIC3 does inhibit Wnt signalling at gastrulation then a positive interaction between the *Bfc* and *Ka* alleles would be expected, such as a more severe phenotype. Compound mutant embryos were produced using the strategy shown in Figure 3.15 and again analysed with markers of brain and axis development. As shown in Figure 3.13 and Table 3.3, 50% of $Ctnnb1^{Bfc/+}$; $Zic3^{Ka/X}$ embryos displayed various degrees of anterior truncation which is a significant increase compared to embryos heterozygous for one or other of these alleles (p<0.01). The embryos were also inspected for axial defects via WMISH to *T* and an increase in the penetrance of axis defects was observed amongst the compound embryos versus either individual heterozygote (Table 3.3). Likewise, the analysis of compound mutant embryos via WMISH to *Shh* revealed aberrant node morphology as found in the $Zic3^{Ka/Y}$ embryos (Figure 3.14).

Another indication of elevated Wnt signalling is the intrusion of Wnt target gene transcripts into the forebrain (the wild-type rostral limit of which is the midbrain). The distribution of transcripts from two such target genes (*Axin2* and *Sp5*) was therefore analysed in wild-type and mutant embryos at 8.5 dpc. As can be seen in Figure 3.16, both genes are expressed not only in the some posterior and/or lateral regions of the embryo but also in the anterior neurectoderm, with a rostral limit at the

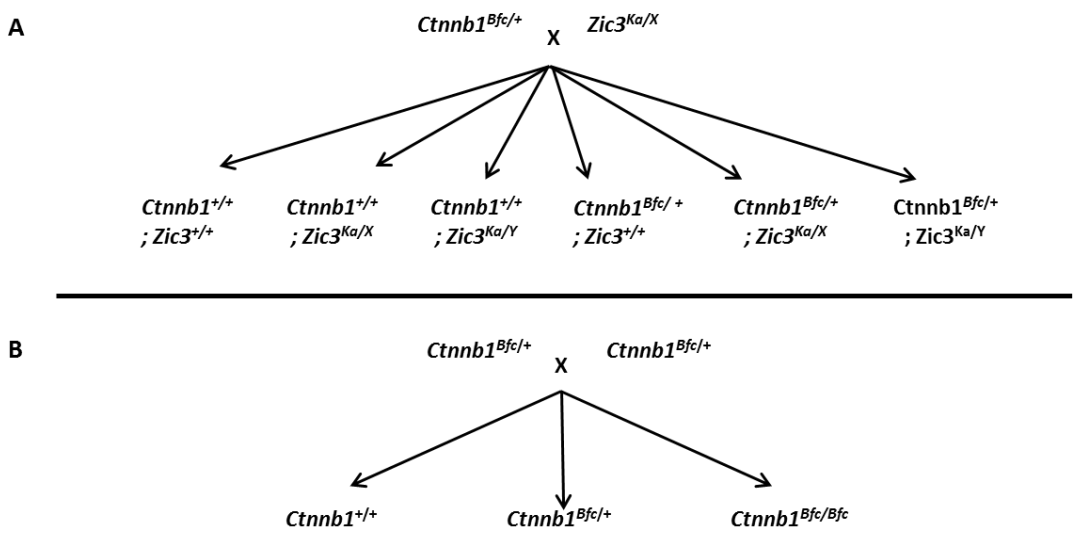

Figure 3.15: The experimental strategy. (A) A genetic cross was set up between a *Bfc* male and a *Ka* female to generate mutant embryos, genotypes are indicated on panel. Genotypes of the following embryos *Ctnnb1*$^{+/+}$*; Zic3*$^{+/+}$, *Ctnnb1*$^{+/+}$*; Zic3*$^{Ka/Y}$, *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{+/+}$ and *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{Ka/X}$ were used for downstream phenotypic analysis utilising WMISH. This analysis was used to examine phenotypes which are associated with elevated level of canonical Wnt/β-catenin signalling during early embryogenesis. (B) The genetic cross between a *Bfc* male and a *Bfc* female was set up to generate homozygous *Bfc* mutant embryos. These homozygous mutants were used for downstream phenotypic analysis utilising WMISH.

forebrain/midbrain boundary (Jho *et al.*, 2002). In a proportion of embryos null for *Zic3* (Ctnnb1$^{+/+}$; Zic3$^{Ka/Y}$) or trans-heterozygous for the *Ka* and *Bfc* alleles (*Ctnnb1$^{Bfc/+}$; Zic3$^{Ka/X}$*) the *Axin2 and Sp5* expression domains expand into the anterior forebrain of the mutants (16%, N=3/18) and *Ctnnb1$^{Bfc/+}$; Zic3$^{Ka/X}$* (20%, N= 2/10) (Figure 3.16 and Table 3.4) indicating that presumptive forebrain tissue was instead caudalised. Evidently, the mutant embryos decrease in forebrain tissue is a direct result of failing to inhibit the Wnt/β-catenin pathway during gastrulation (Fossat *et al.*, 2011). Overall the analysis supports the hypothesis that enhanced canonical Wnt signalling at gastrulation underpins the head truncation and axis defects associated with *Zic3* loss-of-function.

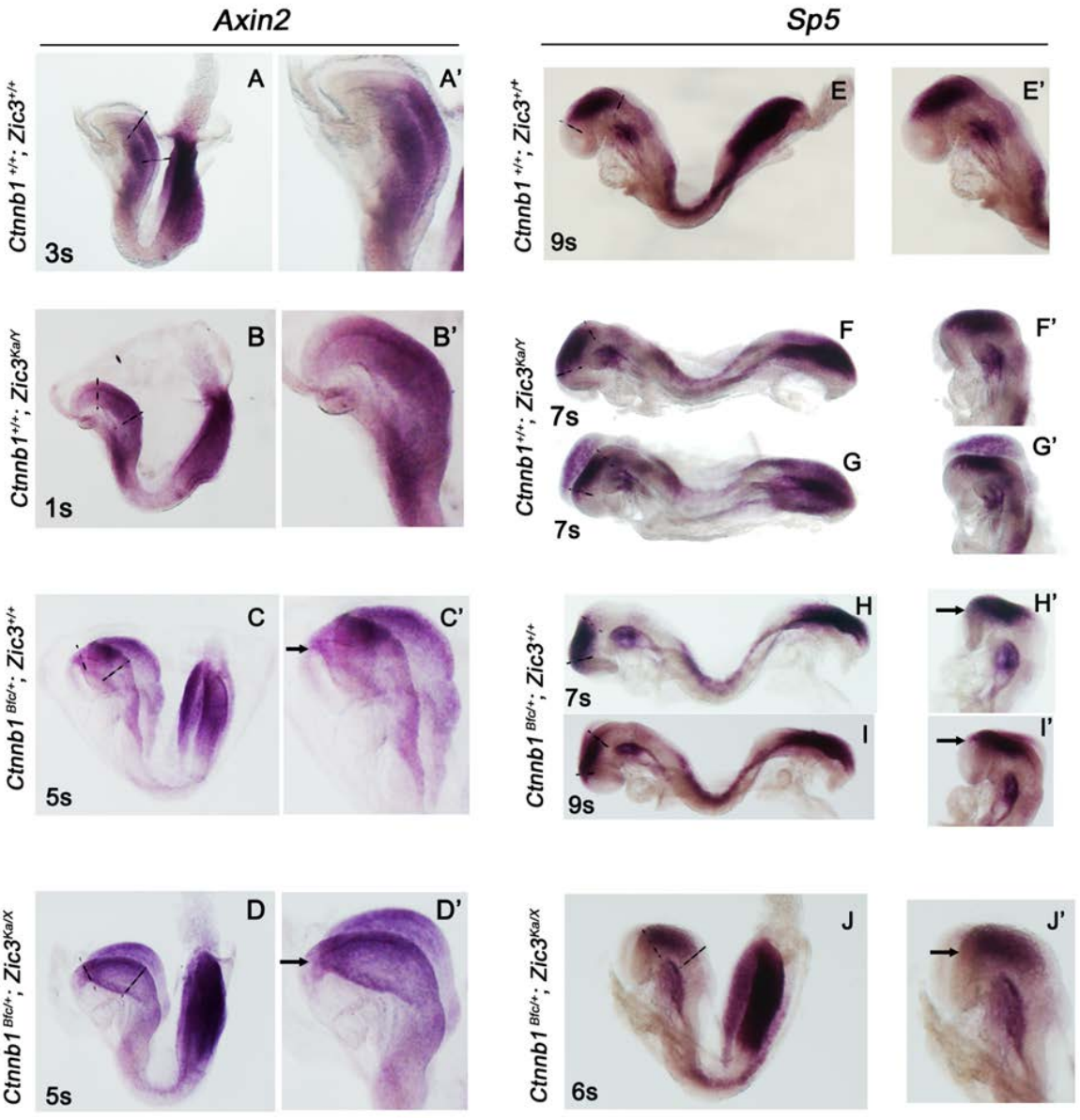
Axin2
Sp5
Ctnnb1 +/+; Zic3+/+
Ctnnb1+/+; Zic3Ka/Y
Ctnnb1 Bfc/+; Zic3+/+
Ctnnb1 Bfc/+; Zic3Ka/X
A
A'
B
B'
C
C'
D
D'
E
E'
F
F'
G
G'
H
H'
I
I'
J
J'
3s
1s
5s
5s
9s
7s
7s
7s
9s
6s

Figure 3.16: *Zic3* loss-of-function mutant embryos fail to inhibit the canonical Wnt signalling pathway at the anterior forebrain. (A-J') Lateral view of mouse embryos following WMISH to *Axin2* (A-D') and trans-acting transcription factor 5 (*Sp5)* (E-J'). Stages are indicated on the panel. Anterior is in the left side in all images. Wnt/β-catenin direct target genes *Sp5* and *Axin2* are not expressed in the anterior forebrain in the wild-type embryos. (A-A') The normal expression of *Axin2* in the midbrain. (B-B') An embryo with a slight ectopic expression extended in the forebrain as indicated by the black arrow in panel B'. The same observation seen in the mutants embryos *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{+/+}$ (C-C') and *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{Ka/X}$ (D-D') the expression of *Axin2* has shifted anteriorly as indicated by the black arrows in panels C' and D' respectively. (E) A wild-type embryo showing the normal expression of *Sp5* in the midbrain. (F-F') *Ctnnb1* $^{+/+}$*; Zic3*$^{Ka/Y}$embryo with normal expression of *Sp5*. (G-G') *Ctnnb1* $^{+/+}$*; Zic3*$^{Ka/Y}$ embryo showing anterior expression of *Sp5* in the forebrain as indicated by the black arrow in panel G'. The same observation seen in mutants embryos *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{+/+}$ (H-H' and I-I') and *Ctnnb1*$^{Bfc/+}$*; Zic3*$^{Ka/}$ (J-J') the expression of *Sp5* has shifted anteriorly in the forebrain as indicated by the black arrows in panels H', I' and J' respectively. The dashed lines are the boundaries of the midbrain.

Table 3.4: The frequency of phenotypes typically associated with elevated canonical Wnt signalling in 8.5 dpc *katun, Bfc* and compound mutant embryos.

Genotype	No.	No. with axis defects (%)	No. with anterior truncation (%)
$Zic3^{+/+}$	32	0 (0)	0 (0)
$Zic3^{Ka/X}$	21	0 (0)	0 (0)
$Zic3^{Ka/Y}$	18	5 (11)*	3 (16)*
$Ctnnb1^{Bfc/+}$; $Zic3^{+/+}$	19	0 (0)	5 (26)*
$Ctnnb1^{Bfc/+}$; $Zic3^{Ka/X}$	10	2 (20)*	2 (20)*

* Significantly different from the wild-type at $p<0.001$ by X^2

3.2.3 Elevation of canonical Wnt/β-catenin signalling during early embryogenesis effects cardiac situs

The experiments described above raised the possibility that elevated Wnt signalling may also underpin the left-right axis defects associated with *Zic3* loss-of-function in both mouse and man. To examine this possibility, I decided to investigate L-R axis formation in the *Bfc* mouse strain, since this strain affects the function of β-CATENIN; the non-redundant pivot of the canonical Wnt pathway. In contrast, ZIC proteins have been shown to interact with transcriptional mediators of the Wnt, Hedgehog and Nodal pathways (Cheng, 2006, Fernandes *et al.*, 2007, Pourebrahim *et al.*, 2011, Fujimi *et al.*, 2012, Houtmeyers *et al.*, 2016) complicating the attribution of phenotypic effects to a particular signalling pathway using *Zic* alleles. The first organ to exhibit L-R asymmetry is the heart (Ramsdell, 2005) and the expression of *Nppa* in the right atrium and left ventricle at 9.5 dpc can be used to examine cardiac situs (Moorman *et al.*, 2003). Embryos null for *Zic3*, and compound heterozygous for *Zic3* (*Ka*) and *Ctnnb1* (*Bfc*) alleles were recovered at 9.5 dpc and heart development assessed via WMISH to *Nppa*. While all wild-type embryos showed normal heart formation, 49% of the hemizygous *Ka* mutant embryos showed looping defects consistent with the findings of Ware *et al.,* (2006) and Ahmed *et al.,* (2013). In addition, 19% of the female trans-heterozygous mutant embryos ($Ctnnb1^{Bfc/+}$; $Zic3^{Ka/X}$) exhibited a leftward curve of the heart tube (sinistral looping heart, Figure 3.17 C). This result is statistically significant ($p<0.001$) compared to the wild-type embryos ($Ctnnb1^{+/+}$; $Zic3^{+/+}$) and indicates that enhanced Wnt/β-CATENIN signalling during early embryogenesis may contribute to cardiac defects. To further confirm this, homozygous *Bfc* ($Ctnnb1^{Bfc/Bfc}$) mutant embryos were harvested at 9.5 dpc and *Nppa* expression analysed (Figure 3.18). Interestingly, two of

the 17 embryos showed a completely reversed heart as seen in human cases of heterotaxy, indicating that elevated Wnt activity can disrupt L-R axis formation.

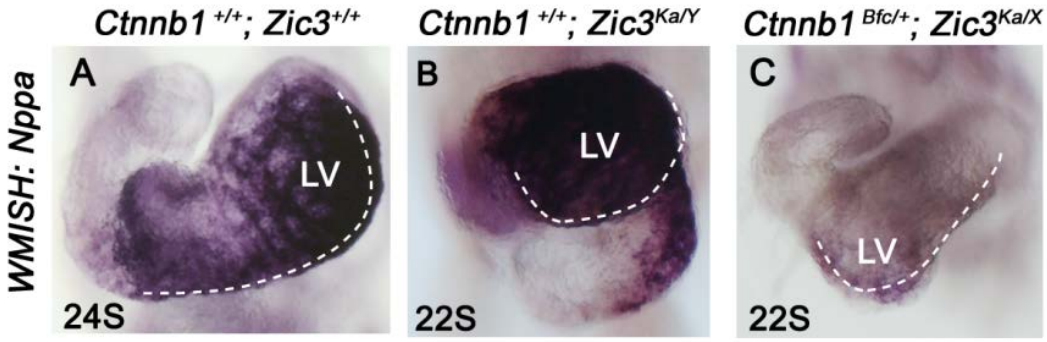

Figure 3.17: Heterotaxy might be associated with elevated canonical Wnt/β-catenin signalling. Ventral view (A-C) of mouse embryos following WMISH to *Nppa* to identify the heart chambers; the dotted line shows the extent of the heart tube. Genotypes and stages are indicated on the panel (A) A wild-type embryo with normal heart formation (dextral looping). (B) An embryo with abnormal heart formation (ventral looping). (C) An embryo exhibits a reversed heart formation (sinistral looping) comparing to the wild-type embryo. S: somites and LV: *left ventricle.*

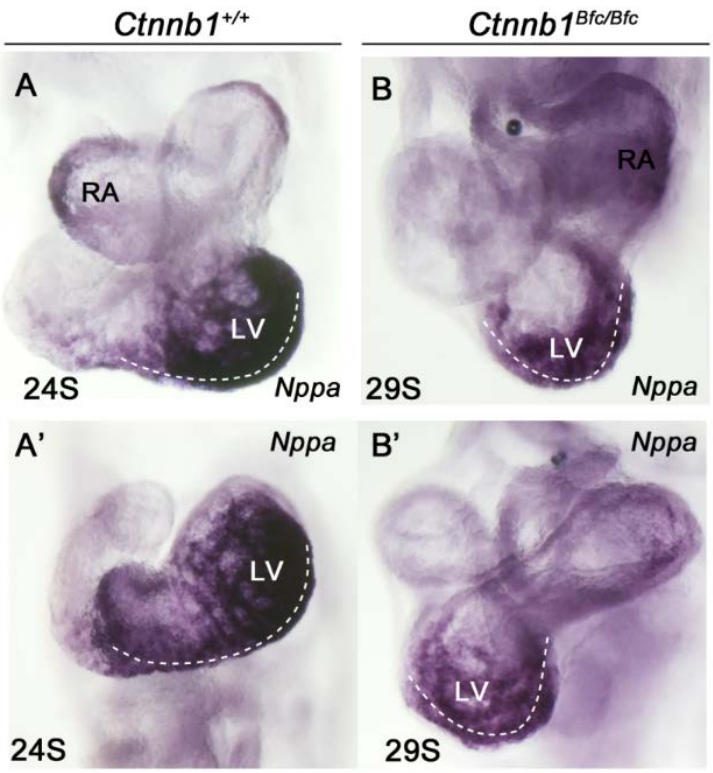

Figure 3.18: Elevation of canonical Wnt/β-catenin signalling contributes to heterotaxy. Images of mouse embryo hearts following WMISH to *Nppa* to identify the heart chambers. Genotype and stage are indicated on the panel. (A-B) Anterior view and (A'-B') ventral view of the heart: dotted line shows the extent of the heart tube. (A-A') A wild-type embryo showing the normal expression of *Nppa* in the RA and the *LV* of the heart. This embryo represents the normal heart formation (dextral looping). (B-B') Embryo showing the normal expression of *Nppa* in RA and LV. Mutant embryos exhibit complete reversal of the heart chambers (sinistral looping) similar to heterotaxy cases in humans. S: somites, RA: right atrium and LV: *left ventricle.*

Besides the two heterotaxic embryos, all of the remaining 15 homozygous *Bfc* embryos exhibited differing cardiac abnormalities as well as severe head truncations (Figure 3.19). The most severe defect was the occurrence of cardia bifida (Figure 3.19 C-C"); a failure of the bilateral myocardial cells to coalesce into a single central heart tube resulting in the presence of two independent hearts (Compernolle *et al.*, 2003). It has been reported that this is caused by a failure of the two lateral heart fields to fuse correctly. Qian *et al.* (2011) identified a hypomorphic mutation in *Axin2* (*Axin2*^{canp}) that resulted in the same phenotype and was thought to be due to a requirement for high levels of canonical Wnt signalling in cardiac progenitors (Qian *et al.*, 2011). Additionally, it could be induced by defective neural crest cells (NCC) (Compernolle *et al.*, 2003). Sinistral heart looping was present as well as an unusual looping that made it difficult to distinguish the heart chambers (Figure 3.19 A and D). Thus, dysregulating expression of Wnt/β-CATENIN signalling at early embryogenesis is essential for proper cardiac looping and ventricle formation.

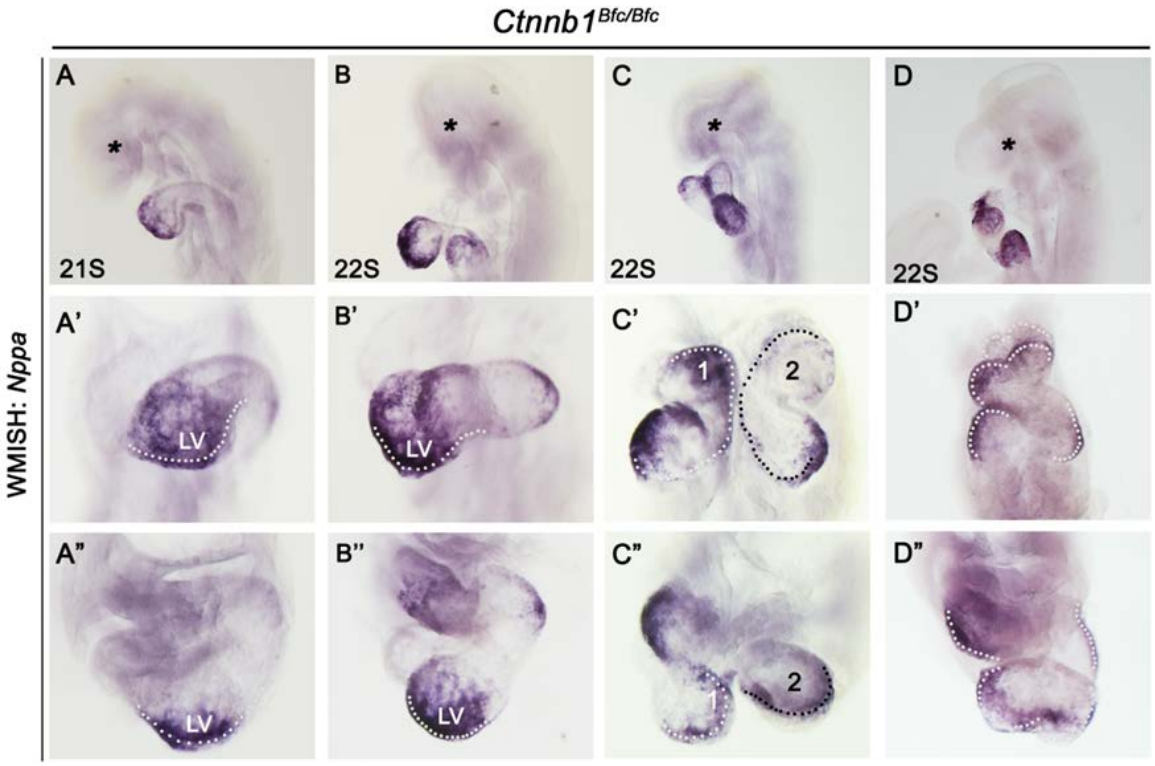

Figure 3.19: Elevation of canonical Wnt/β-catenin signalling in early embryogenesis can cause different cardiac situs defects. Images of *Ctnnb1*$^{Bfc/Bfc}$ mouse embryos following WMISH to *Nppa* to identify the heart chambers. Stages are indicated on panels. Lateral view (A-D), ventral view (A'-D') and anterior view (A"-D") of the heart: dotted line shows the extent of the heart tube. (A-D) All mutant embryos exhibited severe head truncations (asterisk). (A'-A") An embryo with an abnormal heart formation (ventral looping). (B'-B") An embryo with a completly reversed heart (sinistral looping), similar to a heterotaxy phenotype. (C'-C") An embryo with two separate hearts (numbers indicated on images), this phenotype is called cardia bifida. (D'-D'') An embryo with unidentified/unusual heart formation. S: somite and LV: *left ventricle.*

3.2.4 Left-Right axis formation is perturbed in a proportion of *Ctnnb1*$^{Bfc/Bfc}$ embryos

To determine whether the cardiac situs defects are part of a more generalised defect in L-R axis development, the formation of the cardinal vein was examined. Stereotypical cardinal vein morphology occurs in response to L-R axis establishment and can be independent of heart looping. For example, defects in cardinal vein morphology occur in conjunction with other organ asymmetries such as lung isomerisms and mislocation of the pancreas in embryos that lack *Pitx2* (Shiratori *et al.,* 2006). Normally, *Pitx2* is expressed in the left (but not right) cardinal vein at 9–9.5 dpc (Meno *et al.*, 1998). This marker was therefore used as a surrogate for the correct establishment of asymmetries other than heart looping. The expression of *Pitx2* in the left cardinal vein was detected in all five *Ctnnb1*$^{Bfc/Bfc}$ embryos examined (Figure 3.20), indicating that this element of left-right development is either unaltered or altered at low penetrance in *Ctnnb1*$^{Bfc/Bfc}$ embryos. At later stages of embryogenesis (10.5 dpc), *Pitx2* expression initiates in the ventral diencephalon (Martin *et al.*, 2002). Interestingly, ectopic expression of *Pitx2* was detected in the cranial region of all four *Ctnnb1*$^{Bfc/Bfc}$ embryos examined at 21 or 22 somites (9.5 dpc) but not in three embryos at 11 and 12 somites (Figure 3.20). A similar ectopic expression was noted in embryos that lack *Zic2* function (Barratt *et al.*, 2014), but in both cases sectioning is required to determine whether this represents early ventral diencephalon expression or an entirely ectopic domain.

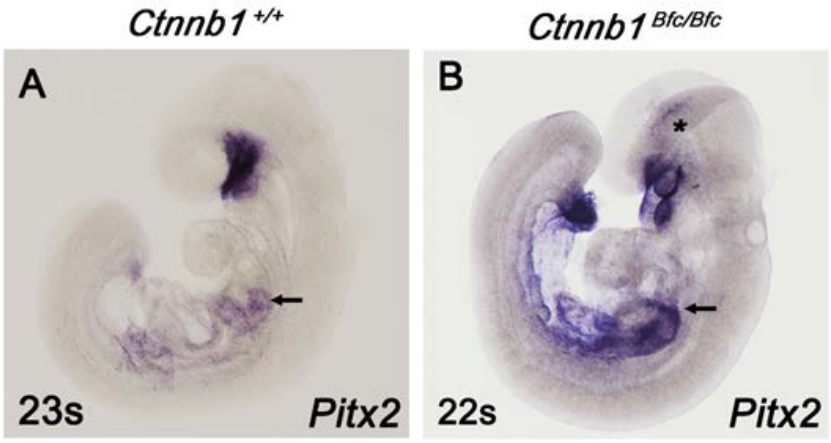

Figure 3.20: Ectopic expression of *Pitx2* in the midbrain of *Ctnnb1*$^{Bfc/Bfc}$ embryos. Lateral view of mouse embryos following WMISH to *Pitx2*. (A) A 23 somite *Ctnnb1*$^{+/+}$ embryo with normal expression of *Pitx2* in the left cardinal vein (black arrow). (B) A 22 somite *Ctnnb1*$^{Bfc/Bfc}$ embryo with normal *Pitx2* transcription in the left cardinal vein (black arrow) along with a severe head truncation and ectopic expression of *Pitx2* in midbrain (asterisk). s: somites.

As described in the chapter introduction, overt organ asymmetry is preceded by the asymmetric expression of certain genes in the lateral plate mesoderm (LPM) and the establishment of a midline barrier. I therefore determined whether these indicators of L-R axis establishment are seen in *Ctnnb1*$^{Bfc/Bfc}$ embryos using WMISH to known marker genes in carefully staged embryos. At the 3-6 somite stage of normal development genes encoding the secreted molecule NODAL, it's secreted antagonist LEFTY2 and the downstream transcription factor PITX2 are all expressed in the left LPM and the Nodal antagonist LEFTY1 is expressed at the axial midline. In 50% (N= 2/2) of homozygous *Bfc* embryos *Nodal* expression was present as usual in the LPM but transcripts were absent in the remaining 50% (N= 2/2) of homozygotes (Figure 3.21 and Table 3.5). The finding of some normal embryos is consistent with the incomplete penetrance of the cardiac situs defect documented in Table 3.6. Similarly, *Lefty2* expression was normal in 66% (N= 4/6) of examined homozygotes, but absent in the remaining 33%, whilst expression of the downstream transcription factor *Pitx2* was normal in 33% of embryos examined, but bilateral in the remaining 66% of embryos. Additionally, the expression of *Lefty1* which marks the midline barrier was normal in 66% (N= 4/6) of *Bfc* mutant embryos but absent in the remaining 33% (N= 2/3). Although the absence of the midline barrier in some cases is associated with the bilateral expression of *Nodal*, *Lefty2,* and *Pitx2 (*Meno *et al.,* 1998*,* Yamamoto *et al.,* 2003*),* there is precedent for the situation seen here (a reduction in *Nodal* and *Lefty1/2*) which most closely resembles the phenotype seen in the *Pkd2*$^{-/-}$ mutant (Pennekamp *et al.*, 2002). The ion channel polycystin-2 is required for L-R axis determination in mice (Pennekamp *et al.*, 2002). Overall the data confirms that the formation of the L-R axis is perturbed in a proportion of *Ctnnb1*$^{Bfc/Bfc}$ embryos consistent with the finding of incompletely penetrant cardiac situs abnormalities in embryos of this genotype.

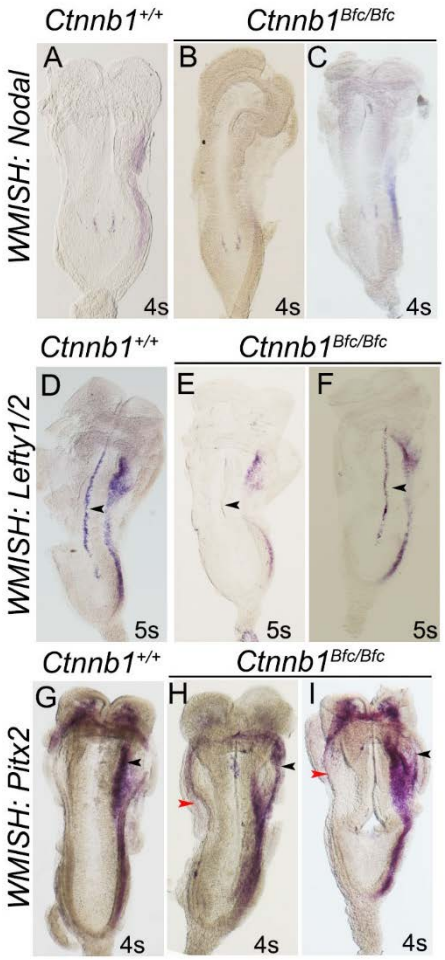

Figure 3.21: Defective formation of the L-R axis nodal cascade. (A-I) Ventral views of embryos following WMISH to the genes shown on embryos of the genotypes and stage shown. Anterior is to the top in all images. (A-C) The expression of *Nodal* is normal in the left LPM in wild-type embryos; however in *Ctnnb1*$^{Bfc/Bfc}$ (B-C) *Nodal* is reduced to varying degrees. (D) A wild-type embryo showing normal expression of *Lefty1* in the midline of the embryo and *Lefty2* in the left LPM. (E) *Lefty1* is absent (black arrowhead) and *Lefty2* expression is reduced in the left LPM in some *Ctnnb1*$^{Bfc/Bfc}$ embryos, but the expression for both genes (*Lefty1/2*) is normal in F. (G) The expression of *Pitx2* is normally in the cranial mesenchyme and in the LPM of wild-type embryos. (H-I) In *Ctnnb1*$^{Bfc/Bfc}$ embryos, *Pitx2* is present in the cranial mesenchyme and expressed in the outer layer of the left LPM (black arrowhead), along with a slight expression in the right LPM indicated (red arrowhead). LPM: lateral plate mesoderm, s: somites.

Table 3.5: Molecular phenotypes of *Ctnnb1*$^{Bfc/Bfc}$ embryos.

Lateral plate mesoderm		
Stages	L-LPM	R-LPM
3-6 somites		
LPM *Nodal* (N=4)	2 present 2 reduced	
LPM *Lefty2* (N=6)	4 present 2 reduced	
LPM *Pitx2* (N=3)	3 present	2 present
Midline		
Stages	Present	Absent
3-6 somites		
L*efty1* (N=6)	4	2

Shown is the number of embryos exhibiting the defect among the total examined. L: left, R: right, LPM: lateral plate mesoderm.

Table 3.6 : The frequency cardiac situs defects in 9.5 dpc *katun*, *Bfc* and compound mutant embryos.

Genotype	No.	No. with cardiac situs defect (%)
Zic3$^{+/+}$	69	0 (0)
Zic3$^{Ka/X}$	36	0 (0)
Zic3$^{Ka/Y}$	45	22(49)*
Ctnnb1$^{Bfc/+}$ ***; Zic3***$^{+/+}$	21	0 (0)
Ctnnb1$^{Bfc/+}$ ***; Zic3***$^{Ka/X}$	15	3 (18.7)*
Ctnnb1$^{Bfc/Bfc}$	17	9 (52.9)*

* Significantly different from the wild-type at $p<0.001$ by X^2

3.2.5 Gene expression at the node is perturbed in the *Ctnnb1*$^{Bfc/Bfc}$ embryos

At least two possible reasons for the near absence of both the midline barrier and the left LPM Nodal cascade exist; either the asymmetric signal from the node is not produced or it is not transferred to the LPM. To distinguish between these possibilities, gene expression at the node was first examined. The receipt of the leftward nodal flow is marked by the establishment of gene expression asymmetries at the node of early somite stage embryos. For example, the expression of *Noda*l itself becomes upregulated on the left margin of the node slightly before the initiation of *Nodal* left LPM expression. When analysed, *Bfc* embryos again show an incompletely penetrant phenotype, with 50% (N=2/4) of embryos showing normal expression of *Nodal* and the remaining 50% (N= 2/4) showing abnormal expression that fails to become asymmetric and in some cases is highly disorganised (Table 3.7 and Figure 3.22). Expression of another secreted molecule, the NODAL antagonist DAND5, also becomes asymmetric at the node of wild-type embryos prior to the initiation of *Nodal* LPM expression. In *Bfc* embryos, the expression of *Dand5* was either absent from the node or was present but failed to become elevated in the right margin of the node (Figure 3.22 and Table 3.7) and was again sometimes highly disorganised (Figure 3.22). Thus it appears that in *Bfc* embryos an asymmetrical node signal is aberrant in comparison with that seen in wild-type embryos.

The gene *Sox17* encodes a Sry-box containing protein and is expressed early in the differentiating definitive endoderm of the gastrula and is required for its normal formation (Kanai-Azuma M *et al.,* 2002). Hadjantonakis and colleagues (Viotti *et al*., 2012) reported that *Sox17*-null embryos exhibit defective L-R axis and an absence of the Nodal cascade in either the left or the right LPM. Gene expression at the node did however become asymmetric at the correct developmental stage. This suggests the

Table 3.7: *Nodal* and *Dand5* expression pattern in *Ctnnb1*$^{Bfc/Bfc}$ embryos.

Node				
	Stages	Normal	Absent	Not asymmetric
Nodal (N=4)	3-4 somites	1		1
	5-6 somites	1		1
Dand5 (N=4)	3-4 somites		2	2

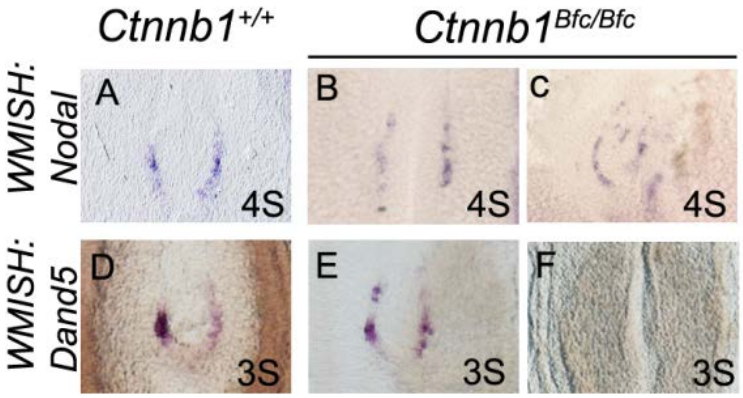

Figure 3.22: Aberrant expression of node markers (*Nodal* and *Dand5*) in *Ctnnb1*$^{Bfc/Bfc}$ embryos. (A-F) Ventral view of mouse embryos following WMISH to *Nodal* and *Dand5*. Genotypes and stages are indicated on each panel. Anterior is to the top in all images. (A) *Nodal* expression normally becomes asymmetric (leftward) in 4 somite stage wild-type embryos. (B-C) In the *Ctnnb1*$^{Bfc/Bfc}$ the perinodal expression failed to become asymmetric or was disorganised (D) *Dand5* is expressed asymmetrically (rightward) in the crown cells of the node in 3 somite stage wild-type embryos. (E-F) In *Ctnnb1*$^{Bfc/Bfc}$ embryos the perinodal expression failed to become asymmetric or was abcent. S: somites.

transmission of the asymmetric signals from the node to the left LPM requires *Sox17* expression to ensure formation of the definitive endoderm (Viotti *et al*., 2012).

To investigate whether the definitive endoderm of *Ctnnb1*$^{Bfc/Bfc}$ mutant embryos appears competent to transfer any signal, the expression of *Sox17* was investigated. As shown in Figure 3.23, all homozygous *Bfc* embryos appear to have normal expression of *Sox17* in the definitive endoderm suggesting that the signal is capable of being transferred to the left LPM, however a proportion of *Bfc* embryos (50%, N=4) showed ectopic expression of *Sox17* in the pit of the node (Figure 3.23 B'). Overall, the analysis of the *Ctnnb1*$^{Bfc/Bfc}$ embryos suggests that that nodal flow is not correctly interpreted in a proportion of *Ctnnb1*$^{Bfc/Bfc}$ embryos, as evidenced by the failure to establish asymmetrical *Nodal* and *Dand5* expression at the midline of early somite stage embryos. Additionally, it seems the pit of the node may be abnormal in some *Ctnnb1*$^{Bfc/Bfc}$ embryos, as judged by the ectopic expression of *Sox17*. This finding of ectopic expression of *Sox17* at the node of mutant embryos may be a possible cause of the failure of L-R establishment.

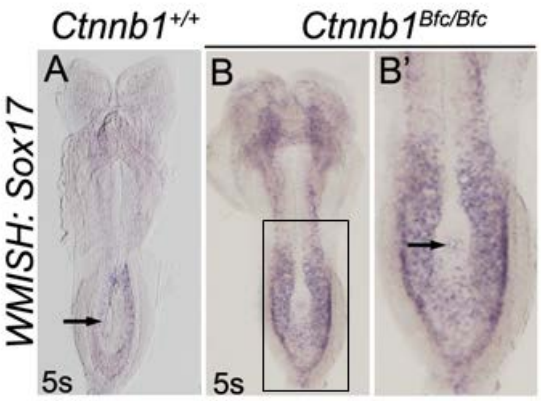

Figure 3.23: The expression of *Sox17* has expanded ectopically to the node in *Ctnnb1*$^{Bfc/Bfc}$ embryos. (A-B') Ventral view of embryos following WMISH to the gene shown on embryos of the genotypes and stages shown. Anterior is to the top in all images. (A-B) Normal expression of *Sox17* is in the DE; this expression is retained in the *Ctnnb1*$^{Bfc/Bfc}$ embryos as seen in B. However, the expression of *Sox17* has expanded ectopically into the node. The rectangle in B is expanded in B' showing DE cells within the node (black arrow). DE; definitive endoderm, s: somites.

3.2.6 The node of *Ctnnb1*$^{Bfc/Bfc}$ embryos is morphologically abnormal

Several assays indicate that the node is abnormal in *Bfc* homozygotes, therefore in collaboration with other lab members, scanning electron microscopy (SEM) (Semenov *et al.,* 2001) was used to visualise this structure in embryos at the 3–5 somite-stage. At this stage of development, the node pit cells can be recognised, even when not properly organised, since each cell has a long cilium and the apical surfaces of the cells are strongly constricted (Lee *et al.*, 2008, Lee *et al.*, 2010). The homozygous *Bfc* embryos (N=7) were harvested between 1-4 somite stage and compared with wild-type embryos (N=6) using SEM (Figure 3.24). Visual inspection of the overall morphology of the node (shape and size) suggested that *Ctnnb1*$^{Bfc/Bfc}$ nodes were strikingly abnormal (Figure 3.24 B-D). In the wild-type embryos, a teardrop-shaped depression at the distal tip of the embryo formed (Figure 3.24 A). In *Ctnnb1*$^{Bfc/Bfc}$ embryos, the node margin was irregular and non-ciliated, and flat cells were found within the pit. The morphology of these cells was the same as those of the surrounding endoderm and in some instances they were continuous with the endoderm and appeared to overlay the ciliated, rounded node cells (arrow in B'). The node cells appeared relatively normal in *Ctnnb1*$^{Bfc/Bfc}$ embryos. When quantified, it was found that the cilia of node pit cells occurred at the same frequency across all genotypes (i.e., ~1 cilium/cell) ($p > 0.05$) whereas the cilia length showed a small, but significant, decrease in *Ctnnb1*$^{Bfc/Bfc}$ embryos relative to their stage-matched wild-type counterparts (Figure 3.24 E'-F').

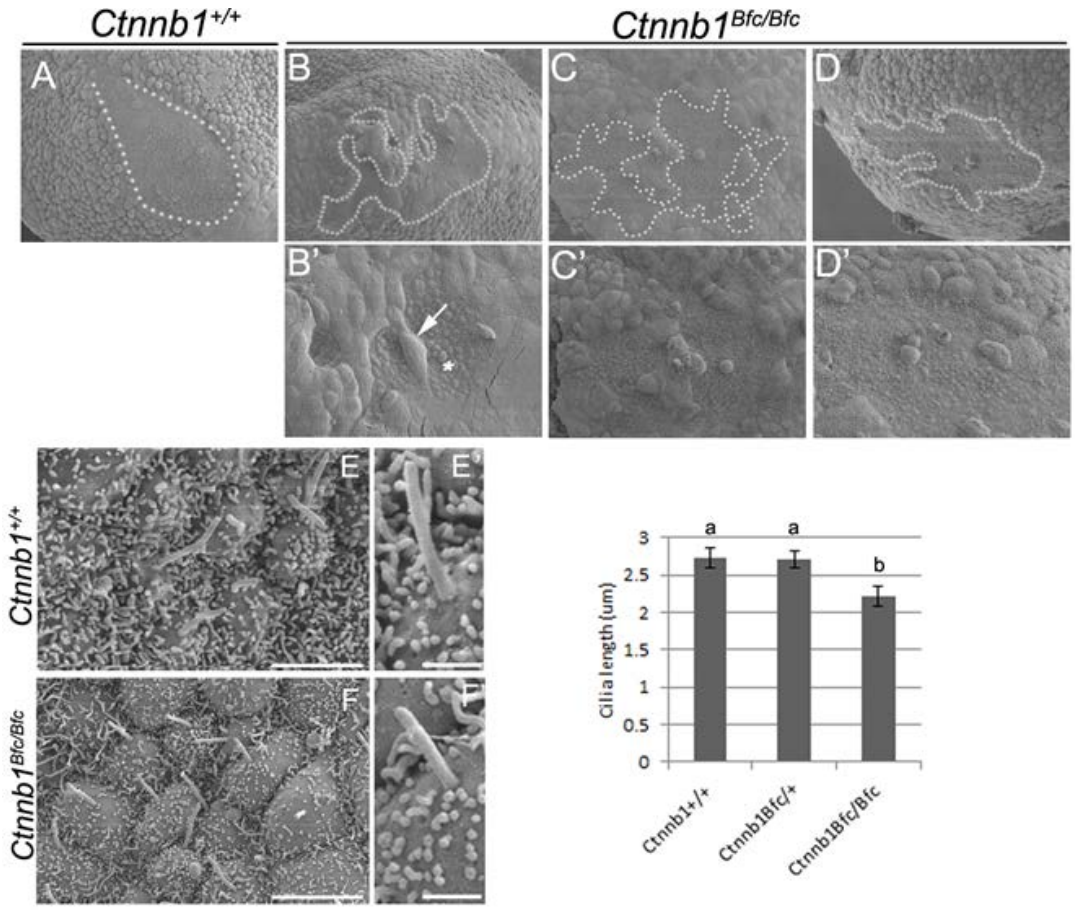

Figure 3.24: Elevation of canonical Wnt/β-catenin signalling in early embryogenesis disturbs node formation. SEM images of the ventral surface of the node ultrastructure in 0-2 somite wild-type embryos (A) and *Ctnnb1*$^{Bfc/Bfc}$ embryos (B-D'). The node in the wild-type embryo is a distinct tear-drop shaped depression with a clear edge. However, *Ctnnb1*$^{Bfc/Bfc}$ embryos have an abnormal node shape and a disorganised border (dotted line). Magnified images (B'-D') of the outlined areas of the irregular node shape in the *Ctnnb1*$^{Bfc/Bfc}$ embryos. Note the expansion of the endoderm cells (white arrow) within the node cells (white asterisk). (E-F') SEM images of the node cilia in 2-4 somite embryos of different genotypes shown on the panel. Scale bar: 5 μm (E-F), 1 μm (E'-F'). (G) A column graph depicting the mean cilia length in 1-4 somite embryos of different genotypes. Error bars: SEM, b: $P < 0.05$, Fisher's LSD ANOVA. Experiments and data analysis performed with K. Barratt.

3.2.7 Defective gene expression at the node reveals an early defect in node morphology

The expression of genes required for node ciliogenesis *Noto* (Beckers *et al.*, 2007) and sensory function *Pkd1l1* (Field *et al.*, 2011) were examined. To determine whether this early defect in mouse node formation may interfere with L-R axis establishment, embryos were harvested at two different stages during gastrulation: early bud stage (EB) and late head fold stage (LHF), and hybridised with *Noto* or *Pkd1l1*. Homozygous *Ctnnb1*$^{Bfc/Bfc}$ embryos (N=5/10, $P<0.05$) showed an expanded domain of *Noto* expression compared to wild-type embryos (Figure 3.25 A-B and Figure 3.26 A-B). The same was observed with *Pkd1l1* in homozygous *Bfc* embryos (N=3/10, $P<0.05$) (Figure 3.25 C-D and Figure 3.26 C-D). The node is normally tear-drop shaped in wild-type embryos however in the *Ctnnb1*$^{Bfc/Bfc}$ mutant embryos the node shape is abnormal and the expression of both genes is aberrant and completely irregular. These findings indicate that elevation of the Wnt/β-catenin pathway in the early stages of embryogenesis affects node structure formation.

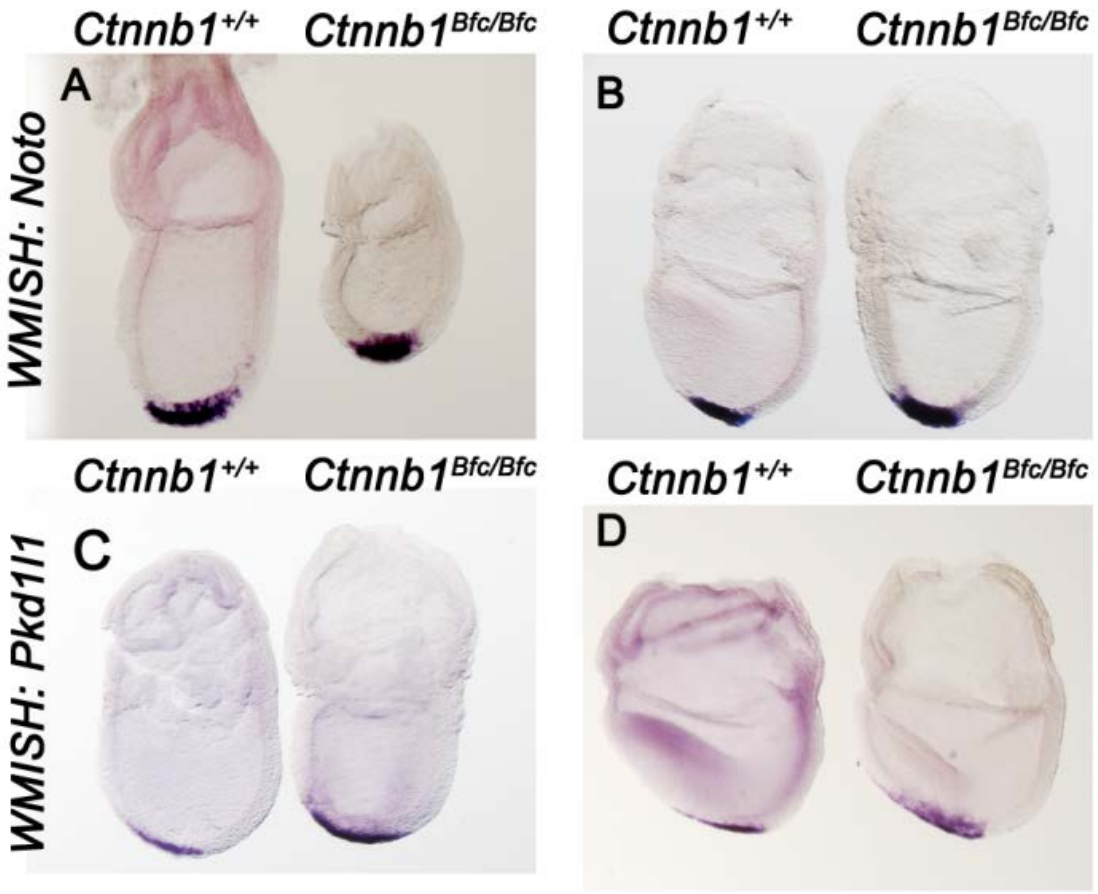

Figure 3.25: Ectopic expression of *Noto* and *Pkd1l1* genes at the gastrula node in *Ctnnb1*$^{Bfc/Bfc}$ embryos. Lateral view (A-D) of mouse embryos following WMISH to *Noto* and *Pkd1l1*. Anterior is to the left in all images. (A) No allantoic bud stage embryos. In wild-type embryos *Noto* expression is restricted to the node, but in *Ctnnb1*$^{Bfc/Bfc}$ embryos the expression is elevated and broad. (B) The same abnormality is observed in late allantoic bud stage in *Ctnnb1*$^{Bfc/Bfc}$ embryos. (C) Late allantoic bud stage embryos. In wild-type embryos *Pkd1l1* expression is restricted to the node, but in *Ctnnb1*$^{Bfc/Bfc}$ embryos the expression is elevated and broad. (D) The same abnormality is observed in late head fold stage *Ctnnb1*$^{Bfc/Bfc}$ embryos.

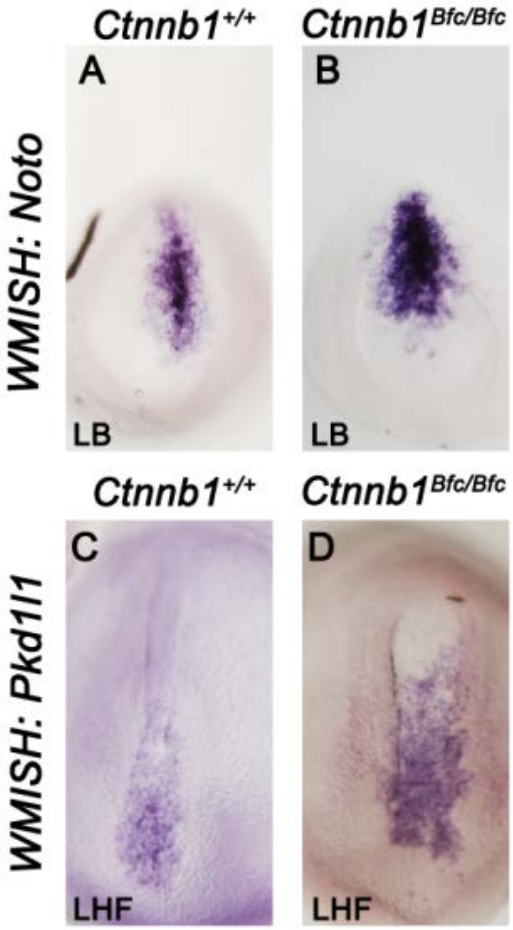

Figure 3.26: Irregular node shape in *Ctnnb1*$^{Bfc/Bfc}$ embryos with ectopic expression of *Noto* and *Pkd1l1* at the node. Ventral view (A-d) of mouse embryos following WMISH to *Noto* and *Pkd1l1*. Stages are indicated on the panel. Anterior is to the top in all images (A) A wild-type embryo showing the normal expression of *Noto* with a clear edge and normal node shape. (B) A mutant embryo exhibiting ectopic expression of *Noto* with a disorganised node shape in comparison to the wild-type. (C) A wild-type embryo showing normal expression of *Pkd1l1* with a clear edge and normal tear-drop node shape. (D) A mutant embryo exhibiting ectopic expression of *Pkd1l1* with an enlarged and disorganised node shape in comparison to the wild-type. LB: late allantoic bud stage, LHF: late head fold stage.

3.2.8 The Batface mutation on a C3H/HeH background

Originally, heterozygous *Bfc* animals were maintained on a Balb/b background. As difficulties were experienced timed-mating these animals, heterozygous *Bfc* Balb/c were backcrossed with C3H/HeH wild-type animals, resulting in C3H/HeH offspring with an even distribution between wild-type and heterozygous *Bfc* genotypes. In comparison to the Balb/c *Bfc* adults which displayed head truncations, additional phenotypes were observed in the C3H/HeH *Bfc* backcross. Namely, a kinked tail (which is commonly observed in *katun* animals) and a white ventral belly spot, indicative of a neural crest defect (Figure 3.27) (Silver *et al.*, 2006). Whether the additional phenotypes are due to increased penetrance of the *Bfc* mutation on the C3H/HeH background, or an increase in severity of preexisting phenotypes is yet to be established. As the phenotypic penetrance and expressivity of a mutation are known to vary due to the combinations of modifying alleles present in different genetic backgrounds (Montagutelli, 2000), it is possible that the C3H/HeH background exhibits higher levels of Wnt signalling than the Balb/c background.

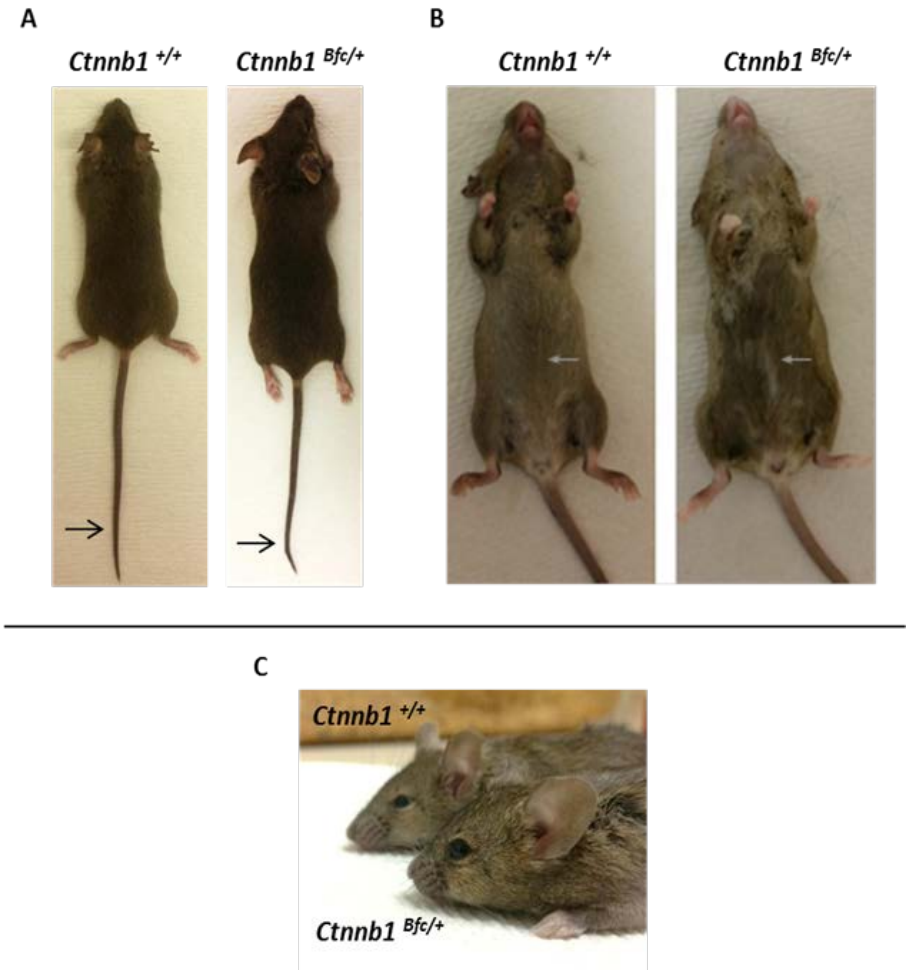

Figure 3.27: A C3H/HeH backcross produces additional Wnt phenotypes in *Bfc* adults. (A) In comparison to wild-type adults, some *Ctnnb1* $^{Bfc/+}$ adults exhibit a kinked tail (black arrow). (B) or ventral belly spot (grey arrow). (C) The expected phenotype associated with elevation of Wnt/β-catenin signalling, a head truncation, is exhibited in *Ctnnb1*$^{Bfc/+}$ animals regardless of their background.

3.2.9 Expression of transcription factor *Tcf7l1* during gastrulation

The findings that Zic3 loss-of-function mutation phenotypes were similar to the published *Tcf7l1* null allele phenotype in several respects. In addition, the other member of the ZIC family, ZIC2 can physically interact with TCF7L2 when it is overexpressed in cultured mammalian cells suggests that a physical interaction between Zic3 and Tcf7l1 may occur during embryogenesis. A pre-requisite for such a scenario is that the two proteins be expressed in the same cell. *Tcf7l1* protein accumulation has been examined in sectioned embryos, but whole-mount images of the *Tcf7l1* expression pattern across a wide range of embryonic stages relevant to axis formation has not been published. Therefore, embryos were harvested between 7.0 and 8.0 dpc (early and performed WMISH for *Tcf7l1*. From the onset of gastrulation until mid-gastrulation *Tcf7l1* is expressed in the ectoderm, and largely restricted to the anterior and extra-embryonic regions, but is not expressed in the extra-embryonic region of the mesoderm. *Tcf7l1* is not expressed in the node when it forms anterior to the primitive streak (Figure 3.28 B). During the early allantoic bud stage, *Tcf7l1* expression begins to recede gradually from the allantoic bud and embryonic ectoderm beginning from the posterior-most region (Figure 3.28 B, C and D). The extra-embryonic ectoderm continues to express *Tcf7l1* throughout the early allantoic bud stage. The expression in both the ectoderm and mesoderm continues until the head formation process begins, establishing the domain of the brain progenitor. However, to confirm the expression of *Tcf7l1* in ectoderm and mesoderm, sectioning of these embryos are required.

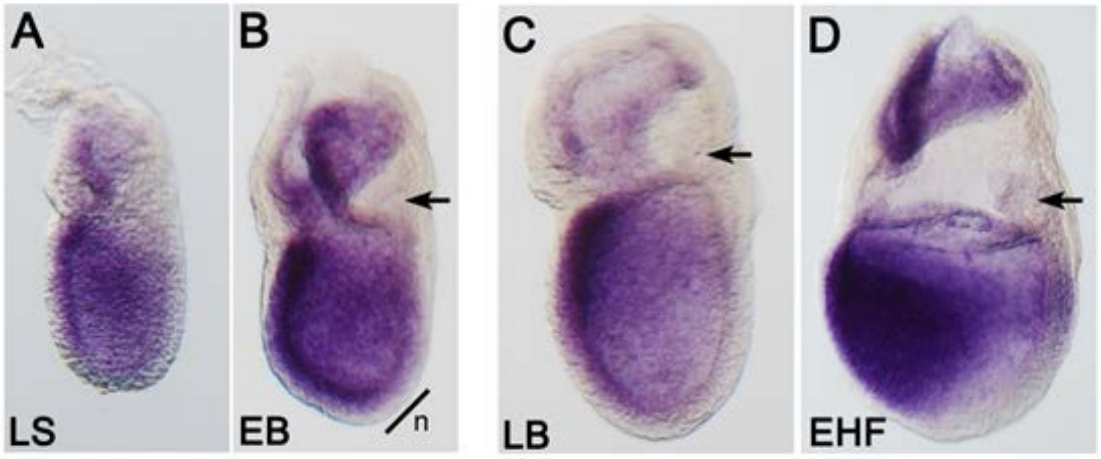

Figure 3.28: *Tcf7l1* is expressed anteriorly during mouse gastrulation. (A-D) Lateral view of wild-type embryos staged from early-gastrulation to early-head fold stage which have been hybridised with *Tcf7l1*. Anterior is to the left. The stages are indicated on the panel, the allontic bud is indicated by the black arrow. LS: late streak, EB: early bud, LB: late bud, EHF: early head fold and n: node.

3.3 DISCUSSION

3.3.1 Overview

The findings of this study suggest a novel basis for ZIC3-associated heterotaxy; namely that elevated canonical Wnt signalling perturbs endoderm development, compromising the establishment of the molecular asymmetries at the late gastrula node that drive organ pattern and position along the L-R axis. I show that during murine gastrulation *Zic3* loss-of-function precipitates a phenotype that shares features with mouse mutants that have elevated canonical Wnt activity. Additionally, a positive genetic interaction is seen between the *Zic3* loss-of-function allele and a known gain-of-function *Ctnnb1* allele (*Bfc*). This is consistent with the normal role of ZIC3 being an inhibitor of canonical Wnt signalling. The data raised the possibility that elevated Wnt activity may contribute to heterotaxy in the *Zic3* mutants. Reasoning that, if this is correct, then L-R axis defects should be found in other mutants with elevated Wnt activity such as evaluated embryos homozygous for the *Bfc* mutation. This demonstrated that they do have L-R axis defects and that elevated Wnt/β-catenin signalling at gastrulation is a previously unrecognised cause of aberrant L-R axis formation. Analysis of these defects indicates that elevated Wnt/β-catenin signalling is associated with abnormal node morphology and function. One feature of *Ctnnb1*$^{Bfc/Bfc}$ embryos is that endoderm cells are found in the pit of the node and it is possible that the L-R axis defects are secondary to aberrant endoderm differentiation. In this case one critical function of *Zic3* is to regulate endoderm formation early in gastrulation.

3.3.2 The Zic3 loss-of-function phenotype shows hallmarks of elevated Wnt activity at gastrulation

As described in the introduction, Wnt signalling must be tightly controlled during embryogenesis and germline mutations that lead to dysregulated Wnt signalling often

cause a phenotype first evident during gastrulation and patterning. When the pathway activity is lowered by loss-of-function mutations in activating components of the pathway (e.g. ligand, co-receptor, chaperone, signal transducing and trafficking molecules) primitive streak formation is blocked. Conversely, gain-of-function mutations in activators or loss-of-function mutations of negative regulators of canonical Wnt signalling (e.g. antagonist, components of the β-catenin destruction complex and transcriptional repressor) result in expansion of the primitive streak, anterior truncation and posterior axis duplication. Animals heterozygous for these mutations often have tail kinks, likely because Wnt signalling is required to regulate somitogenesis. Studies of mouse mutants with altered Wnt activity have led to a model where, at early stages of embryogenesis, the Wnt signal is distributed in a graded pattern increasing along the anterior-posterior axis of the gastrulation stage embryo (Reviewed in Arkell *et al.*, 2013). Previous analysis of embryos null for *Zic3* show that a proportion have an expanded primitive streak, as evidence by the ectopic expression of *Foxa2* and *Fgf8* (Ahmed *et al.*, 2013). In addition, I show here that *Zic3* loss-of -function is associated with anterior truncation and posterior axis duplication. Although not broadly described as *Zic3*-associated phenotypes, these features can be seen in other *Zic3* null alleles (Ware *et al.*, 2006a, Ware *et al.*, 2006b, Jiang *et al.,* 2013, Sutherland *et al.*, 2013). To confirm that these phenotypes are indicative of elevated Wnt activity, the consequence of *Zic3* loss-of-function was examined in a background of elevated Wnt activity via cross with a known gain of function *Ctnnb1* allele (*Bfc*). The penetrance and severity of anterior truncation and axis duplication phenotypes of the *Zic3* null embryos increased in this background confirming that *Zic3* intersects the Wnt pathway and functions to lower transcriptional output of the pathway. These genetic investigations were supplemented with molecular evidence of altered Wnt target gene expression in the developing

forebrain. During normal development, Wnt target genes are absent from the forebrain, but expressed in the midbrain. In embryos with elevated Wnt signalling, the anterior limit of Wnt target gene expression shifts rostrally. As seen in Figure 3.16 both *Axin2* and *Sp5* expression shifts rostrally when examined by wholemount in situ hybridisation. Ideally, this experiment should be augmented by a quantitative approach such as RT-qPCR from the brain of 9.5 dpc embryos as previously (Fossat *et al.,* 2011). Another method used to reveal Wnt pathway activity is via TCF reporter strains of mice. Members of the TCF family of transcription factors (*Tcf7*, *Tcf7l1*, *Tcf7l2* and *Lef1*) are transcriptional mediators of Wnt/β-catenin signalling that bind DNA at a consensus TCF binding site. Transgenic animals have been generated in which multiple copies of the consensus TCF binding site are placed upstream of a minimal promoter and the *β-galactosidase* cDNA to report sites of active Wnt signalling (reviewed in Barolo, 2006). The Arkell laboratory has previously used a cross with one such strain (BATGal) to demonstrate that Wnt signalling is elevated in *Bfc* homozygotes (Fossat *et al.*, 2011). Since then, a more sensitive Wnt reporter has been produced in which six multimerised TCF sites are placed upstream of the *hsp68* minimal promoter and a H2B-GFP reporter cassette, to produce the *TCF/Lef:H2B-GFP* mouse strain (Ferrer-Vaquer *et al.*, 2010). The Arkell laboratory has recently imported this mouse strain, but time constraints prevented me from crossing the *katun* mouse with this TCF reporter strain to visually validate elevated Wnt signalling in this strain. Overall, however, the available genetic and molecular evidence suggest that one role of the ZIC3 transcription factor during normal embryogenesis is to repress Wnt dependent transcription.

3.3.3 ZIC3 may repress Wnt dependent transcription by interaction with TCF7L1

The requirement for repression of Wnt target genes at gastrulation has previously been dramatically demonstrated by mutation of the *Tcf7l1* transcription factor. *Tcf7l1* is the only member of the Tcf/Lef family expressed during gastrulation, and loss-of-function of *Tcf7l1* results in embryos with an expanded primitive streak, anterior truncation, and duplicated posterior axis (Merrill *et al*., 2004). This identifies *Tcf7l1* as a negative regulator of Wnt activity and demonstrates that repression of TCF mediated transcription is critical for normal gastrulation. One interpretation of the similar phenotype between *Tcf7l1* and *Zic3* mutants is that these proteins co-operate to mediate repression of transcription at TCF sites. Along with data from others that indicates that the ZIC proteins can inhibit Wnt dependent transcription in *Xenopus* embryos and human cell lines via a physical interaction between ZIC and TCF proteins (Pourebrahim *et al.*, 2011, Fujimi *et al.*, 2012), I hypothesise that a direct interaction occurs between ZIC3 and TCF7L1 during murine gastrulation. Other members of the Arkell laboratory have shown (via bimolecular fluorescent complementation and proximity linked assays) that these two proteins can interact when overexpressed in cell lines (Koula Diamand, unpublished data). To assess whether this occurs *in vivo*, a proximity linked assay should be conducted on embryonic tissue, assuming antibodies compatible with this process can be identified. If not, CRISPR mutagenesis could be used to generate tagged alleles of ZIC3 and TCF7L1 and antibodies against the tags be used to assess *in vivo* interaction.

3.3.4 The L-R axis defects associated with elevated Wnt signalling may be secondary to aberrant node and/or endoderm differentiation

Bfc mutant embryos fail to form a tear-drop shaped, coalesced node and tightly packed midline. Instead, the node is irregular with clusters of endoderm cells present in the node and the midline is broad. Several studies have identified mutant mice with a similar node phenotype. In support of the idea that elevated Wnt signalling disrupts node formation, loss of function of the *Tcf7l1* Wnt transcriptional repressor leads to dysmorphic nodes with surface endoderm cells (Merrill *et al.,* 2004). If node dysmorphology is a general, yet previously unidentified, feature of elevated Wnt signalling and *Zic3*-null mutation causes elevated Wnt signalling, the mutant embryos should have irregularly shaped nodes and defective endoderm development. Sutherland *et al.* (2013) used immunohistochemistry (IHC) and SEM to examine the node of *Zic3*-null embryos and found that four of five *Zic3*-null embryos had an irregular shaped node with large cells partially covering the surface. Additionally, a single cilium was present on the small mesodermal cells of the node, although on average it was slightly shorter than that of wild-type embryos; similar to what we observed in the *Bfc* mutant embryos.

Moreover, analysis of *Zic3*$^{del/y}$*;T-Cre+* nodes demonstrated that five out of six embryos had dysmorphic nodes, unlike the Cre-negative controls. This suggests that ZIC3 function is required in the nascent mesoderm of the primitive streak-stage embryo for proper node morphology. However, the stage of embryos analysed was not documented in Sutherland et al and it is difficult to directly compare the *Zic3* null and *Zic3*$^{del/y}$*;T-Cre+* phenotype with the *Bfc* embryos analysed here. A detailed comparison of wild-type, *Bfc* and *Zic3* null nodes in a carefully staged series of embryos spanning the stages of initial node formation through to early somitogenesis would be useful. This may shed light on whether an abnormal amount of endoderm is produced by mutant embryos or whether

defective motility is responsible for the finding of endoderm cells in the pit of the overt node. An examination of early somite nodes in *Zic3*-null and *Bfc* mutant embryos should include the determination of the expression pattern of crown cell-expressed genes/proteins that do not become asymmetric (such as *Gdf1 and Pkd2*), to distinguish between the general effect on crown cells and that on symmetric gene expression.

Another group of molecules which, when mutated, produce dysmorphic, endoderm covered nodes are involved in extracellular matrix modification, cell migration and cytoskeletal dynamics, including the reorganisation of actin cytoskeleton (Migeotte *et al.,* 2001, Lee *et al.*, 2008; *Lee et al.,* 2010, Pulina *et al,* 2011). These mutants also have defective midline patterning and aberrant left-right asymmetry, suggesting that correct node morphology is required for left-right axis formation. The overlap in phenotype may also suggest that excess Wnt signalling is associated with dysregulation of cell-extracellular matrix interactions, but this has not yet been investigated.

The analysis of younger embryos presented in this chapter suggests that earlier events may alter node formation and potentially function in *Ctnnb1*$^{Bfc/Bfc}$ homozygotes. In particular, the node of these embryos is generally misshapen, endoderm cells persist in the pit of the node and the cilia of pit cells are significantly shorter than in wild-type, stage-matched embryos. Additionally, previous analysis of other *Zic3* null alleles (Sutherland *et al.,* 2013) and the *Tcf7l1* null allele (Merrill *et al.,* 2004) reveal similar perturbations in node and/or endoderm development suggesting that this may be a general yet previously unidentified feature of elevated Wnt signalling. The aberrant L-R axis formation in these strains may therefore be a secondary consequence of incorrect node and/or endoderm development during gastrulation.

3.3.5 Node development and endoderm cells

The node first becomes apparent at mid-gastrulation at the distal tip of late streak stage embryos. Prior to node emergence, the entire embryonic portion of the embryo is covered by a layer of visceral endoderm cells and the precursor cells of the node are found in at the anterior primitive streak. A subset of cells in the anterior primitive streak begin to express *Noto* and adopt a node or trunk notochord fate. These cells delaminate and exit from the streak to reform as an epithelium with small apical surfaces beneath the endoderm layer (Yamanaka *et al.*, 2007), where they begin to grow cilia (Sulik *et al.*, 1994). At the same time, some epiblast cells transit the primitive streak, begin to express *Foxa2* and adopt a definitive endoderm fate. Definitive endoderm cells then replace the visceral endoderm cells via intercalation and displacement. Between the mid- and late-streak stages of development, endoderm cells are gradually cleared from the site of the presumptive node and midline (Kwon *et al.*, 2008, Viotti *et al.*, 2014) revealing groups of columnar epithelial node cells. Consequently, after transitioning through a phase where node cells are still separated by squamous endoderm cells with lager surfaces, by the OB stage the node appears as a single field of cells with small apical surfaces and growing cilia. At the EHF stage, the node depresses to form the classical tear drop shaped pit of ciliated cells surrounded by ciliated crown cells.

The clearing of endoderm cells and the emergence of the node cells to the ventral surface are thought to be coordinated (Sulik *et al.*, 1994), however the molecular signals underlying the communication between these two layers is yet to be elucidated. The work in this chapter raises the possibility that Wnt signalling is involved either in endoderm differentiation, clearance from the node and/or communication between the node and endoderm cells. A survey of mouse mutants in which node shape and endoderm clearance are abnormal and associated with defects in L-R axis formation

provides clues to other molecules that may also be involved. Mutations of Noto (Beckers *et al.*, 2007) and Oct4 (Hasenöder, 2015) transcription factors, Notch ligand Dll1 (Przemeck *et al.*, 2003) as well as the fibronectin signalling molecule FN1 and it is ligand Integrin $\alpha5\beta1$ (Pulina *et al.*, 2011, Pulina *et al.*, 2014) all produce defective nodes similar to those seen when Wnt signalling is elevated. Evaluating the precise role and possible connections between these molecules is problematic because of inconsistent use of assays and stages analysed as well as inconsistent genetic backgrounds between the mutants. Moreover, in most cases the embryonic node is examined at stages beyond when the node is first emerging. It would be informative to carry out additional analyses on these mutant strains that compare the nodes (morphologically and via gene expression) at earlier time points, evaluate the level of Wnt signalling in each mutant, and track endoderm development, node formation and nodal flow in live embryos. This information could position mutants in a putative genetic pathway which could be tested by particular crosses.

The growing number of mutants with defective L-R axis formation in which abnormal node shape is accompanied by the presence of endoderm in the node suggests causation; it is unclear precisely where L-R axis formation fails in these mutants. When examined in these mutants, cilia formation on the pit cells of the node has been found to be apparently normal (Pulina *et al.*, 2014) or have relatively mild defects in cilia length (Beckers *et al.*, 2007) as in the *Bfc* mutant examined here. Multiple studies suggest the direction and angle that cilia protrude from cells and rotate (rather than cilia length) are the main limiting factors behind flow generation (Smith et al. 2010; Cartwright et al. 2004). Cilia direction and angle remain to be tested in both *Bfc* and *Zic3* mutant embryos and should in the future be examined via visualisation of basal bodies with a α-pericentrin antibody and cell membranes via F-actin staining (Mahaffey et al. 2013).

Another possibility that should be considered is that the presence of endoderm cells disrupts nodal flow and ideally nodal flow should be quantified via particle image velocimetry in the *Bfc* and *Zic3* mutants. However, in conjunction with the finding that even a weak nodal flow is sufficient to initiate the left sided Nodal cascade (Shinohara et al. 2012), it remains possible that reduced nodal flow is insufficient to explain the defective L-R axis formation that accompanies elevated canonical Wnt pathway activity. An alternative explanation is that the presence of endoderm cells and/or the irregular node shape may prevent or diminish receipt of the nodal flow signal. It would be interesting to use calcium sensing reporters in the crown cells of *Bfc* and *Zic3* mutants to evaluate flow reception.

If the presence of endoderm cells in the node is implicated, it will be important to consider how elevated Wnt levels prevent endoderm clearance from the node. One option is that Wnt signalling controls the number of cells that differentiate into endoderm. This possibility is raised by the observation that *Foxa2* is ectopically expressed in both *Tcf7l1* and *Zic3* null embryos (Merrill *et al.*, 2004, Ahmed *et al.*, 2013) and supported by the recent finding in the Arkell laboratory of a similar pattern of ectopic *Foxa2* expression in the *Bfc* homozygotes (data not shown). Consistent with this idea, TCF7L1 has been shown to bind downstream of the *Foxa*2 gene to repress its expression in ES cells (Morrison *et al.*, 2016). A working model is that ZIC3 and TCF7L1 co-operate at this element to repress *Foxa2* expression in the epiblast prior to gastrulation. When *Wnt3* expression begins at the site of primitive streak formation early in gastrulation, a limited de-repression of *Foxa2* expression occurs in cells close to this high Wnt activity, effectively setting the number of cells that differentiate into endoderm. When Wnt signalling is elevated early in gastrulation, *Foxa2* expression is de-repressed in additional cells and too much endoderm is formed. These cells then clog

the node, compromising nodal flow and/or crown cell receipt of the flow signal. To directly test this hypothesis, the relevant *Foxa2* element could be deleted in mouse embryos via CRISPR mutagenesis. If the theory is correct, this should lead to precocious endoderm formation and L-R axis defects.

3.3.6 Mutation of *Zic3* causes L-R axis defects by a mechanism distinct to that of *Zic2* mutation

Previously, the Arkell laboratory showed that mutation of a closely related transcription factor, *Zic2*, also causes cardiac situs defects (Barratt *et al.*, 2014). In the case of the *kumba* allele of *Zic2*, it appears the situs defects arise from a combination of perturbed nodal flow and reception, but which are not asscoaited with aberrant node shape nor with the presence of endoderm cells in the node pit (Barratt *et al.*, 2014) . Whilst it appears that mutation of either *Zic2* or *Zic3* can lead to similar cardiac situs defects the origin of these defects appears divergent. Interestingly, assays in HEK293T cells show that whilst ZIC2 and ZIC3 are each able to inhibit Wnt signalling as measured by a TOPflash assay, the *Zic3 katun* mutation removes this ability (Ahmed *et al.*, 2013) whereas the *Zic2 kumba* mutation (C370S) does not (Pourebrahim *et al.*, 2011). Thus, whilst the *Zic3* null phenotype may be attributed to a failure to inhibit canonical Wnt signalling and production of an elevated Wnt environment, this cannot be said for the phenotype of the *Zic2 kumba* mutants.

3.4 CONCLUSION

The level of activity along the canonical Wnt pathway is known to be important for a variety of processes during mammalian germ layer and axis formation. One unusual feature of the Wnt pathway is that mutations which generate a gain of pathway activity are relatively common. Evidently, negative regulation of this pathway is critical for

normal embryonic development and occurs at all levels of this signal transduction pathway (signal production, reception, transduction and transcription). The work in this chapter suggests that one previously unappreciated consequence of elevated Wnt signalling at gastrulation is a failure of L-R axis formation. It is possible that this defect can be traced to a role for Wnt signalling in endoderm differentiation and/or migration during gastrulation. One role of the ZIC3 transcription factor, mutation of which is known to be associated with defective L-R axis formation in both mouse and man, may be as a transcriptional co-repressor of Wnt/β-catenin mediated transcription. Additional experiments are warranted to confirm the data presented here and further explore the connections between *Zic3*, canonical Wnt signalling endoderm formation and L-R axis establishment. The work may ultimately provide the first molecular explanation of how ZIC3 loss-of-function causes heterotaxy in humans but also highlight a new set of candidate putative heterotaxy associated genes.

Chapter 4: The gastrulation screen; recombination mapping

4.1 INTRODUCTION

4.1.1 The Arkell laboratory gastrulation screen

A gastrulation screen (a genetic screen for *N*-ethyl-*N*-nitrosourea (ENU)-induced recessive mutations that alter gastrulation) conducted by the Arkell laboratory yielded eight pedigrees of interest (Table 4.1). Each pedigree exhibited one of three phenotypes: a failure to thrive, a failure to gastrulate, or defective patterning (Figure 4.1). The task now is to identify the causative mutations using the positional candidate cloning strategy described in Section 1.3.3 This process will be conducted in two separate stages.

Firstly, linkage analysis will be conducted; microsatellite (MIT) markers will be used to create a low-resolution recombination map and localise the mutation to a single chromosomal interval (known as the minimal critical region: MCR). MIT markers polymorphic for the strains used in the gastrulation screen were already available from the Arkell laboratory. Secondly, exome sequencing of carrier animals will be performed and the MCR analysed for heterozygous single nucleotide variants. Whole-exome sequencing and data analysis will be performed at the Australian Phenomics Genotyping Facility at ANU. Further carrier animals will then be selected for sequencing experiments to validate the inheritance of putative causative variants and a genotyping assay established for colony maintenance and downstream phenotype analysis.

Table 4.1: Novel mouse pedigrees with ENU-induced mutations that cause gastrulation defects. The chromosomal location indicates the information available at the project start.

Mutant pedigree	Mutagenised inbred strain	Mapping inbred strain	Chromosomal location
11-23*	C57Bl/6J	C3H/HeH	Mmu 3 V259E allele of *Atp1a1*
11-68*	C57Bl/6J	C3H/HeH	Mmu 6 (32.1-88 Mb)
16-5	BALB/cJ	FVB/N	Unknown
16-7	BALB/cJ	FVB/N	Unknown
18-5	FVB/N	BALB/cJ	Unknown
18-10	FVB/N	BALB/cJ	Unknown
18-19	FVB/N	BALB/cJ	Unknown
18-22	FVB/N	BALB/cJ	Unknown

*The ENU-11 pedigrees were obtained from preliminary ENU experiment animal at selecting the best strains. Recombination mapping for 11-23 and 11-68 was complete at the start of this project.

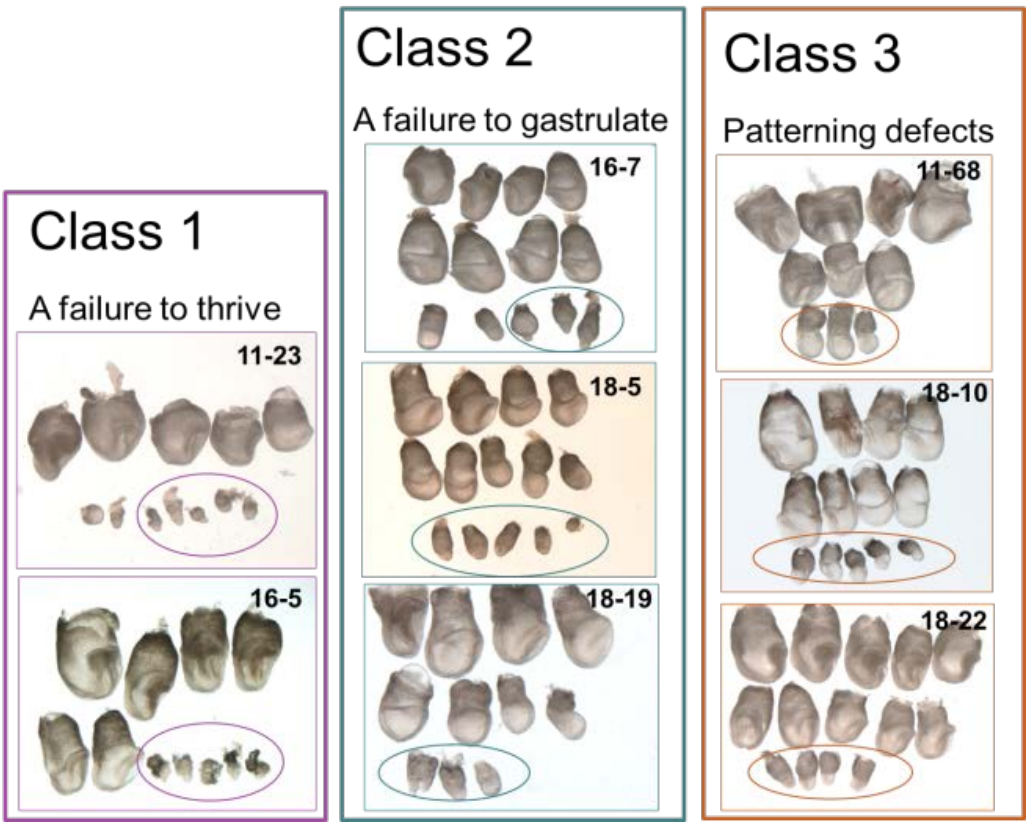

Figure 4.1: The gastrulation screen pedigrees. Lateral view of embryos dissected at 8.0 dpc. Each photo shows a single litter derived from one pedigree. The phenotypes can be divided into three classes. **Class 1** mutant embryos have a small embryonic and extra-embryonic region and do not initiate gastrulation. This phenotype is evident in pedigrees 11-23 and 16-5. **Class 2** mutant embryos have a small embryonic region relative to their extra-embryonic region and do not initiate gastrulation. This phenotype is evident in pedigrees 16-7, 18-5, and 18-19. **Class 3** mutant embryos initiate gastrulation that proceeds incorrectly; this is seen in pedigrees 11-68, 18-10, and 18-22. Mutant embryos are circled.

4.1.2 Breeding strategy for the gastrulation screen

As described in Section 1.3.3, ENU-induced recessive mutations can be isolated in a three-generation breeding protocol. The future identification of the causative mutation can be facilitated using different inbred strains within the breeding protocol. This prevents the need for a separate recombination mapping cross after mutant isolation. Additionally, the use of different strains introduces hybrid vigour; this generally results in an increased litter size, which is an important consideration for embryonic screens. Indeed, in the ENU-11 cross litter, the pregnancy rate was low; therefore, we looked for other strains. On the other hand, using two strains can confound the recovery of mutants because some phenotypes become incompletely penetrant or even lost owing to the modifying effects of the strain background on the phenotype (Montagutelli, 2000). The genetic relationship between inbred strains has been studied using a large number of single nucleotide variants (SNVs, Figure 4.2) (Petkov *et al.*, 2004). This study enabled us to identify fecund strains that are closely related (to minimise modifying effects on the phenotype), but still sufficiently distinct for high-resolution mapping. On this basis, the BALB/cJ and FVB/N strains were chosen. Additionally, published ENU mutagenesis protocols were already established for each strain (Nolan, 2000, Blewitt *et al.,* 2005) and each has an above average litter size relative to other inbred strains (Festing, 1998).

The breeding protocol used for the gastrulation screen is shown in Figure 4.3 A. Male mice of the mutagenised strain were treated with ENU and then bred with female mice of the mapping strain to produce G1 progeny, each of which carried a distinct set of newly induced mutations.

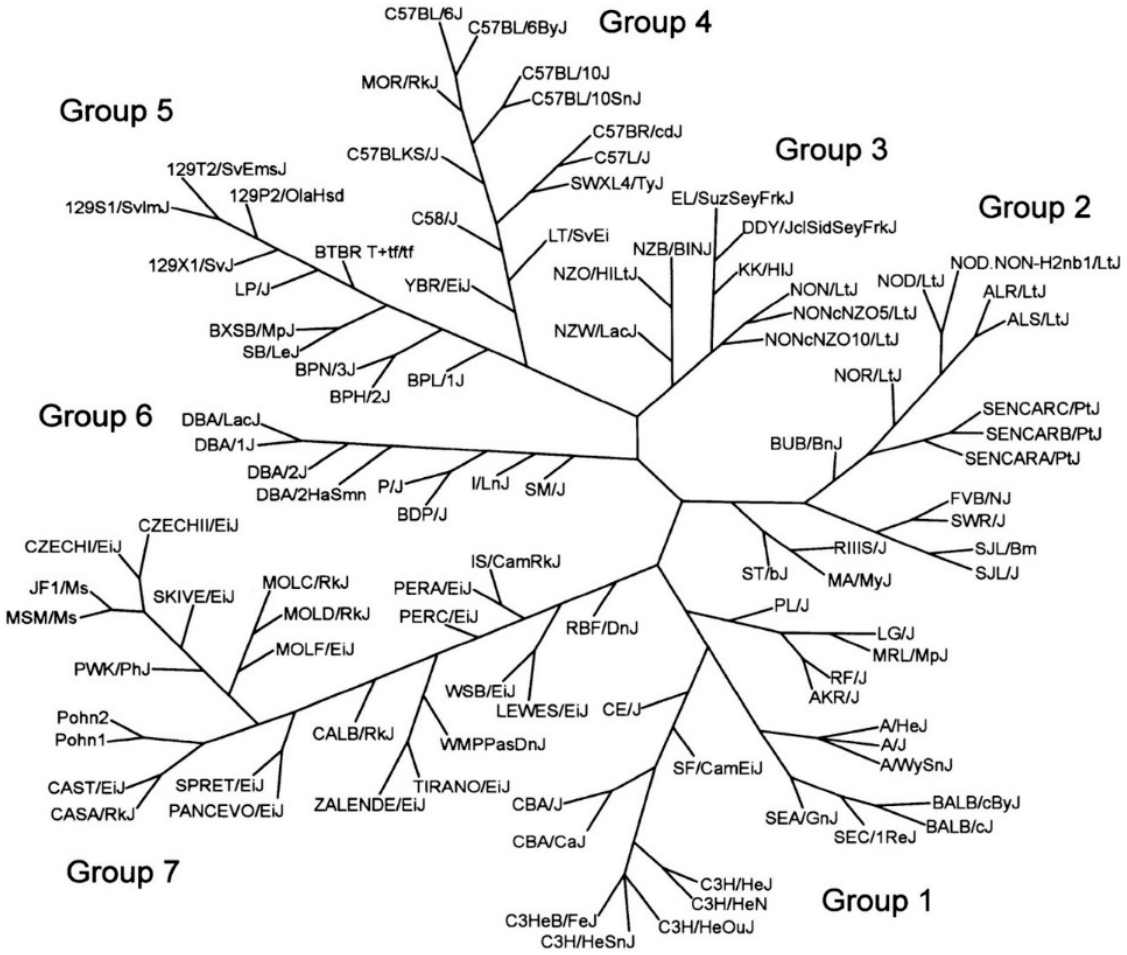

Figure 4.2: Mouse family tree. The 102 strains tested are organized into seven groups. Group 1, Bagg albino derivatives; Group 2, Swiss mice; Group 3, Japanese and New Zealand inbred strains; Group 4, C57/58 strains; Group 5, Castle's mice; Group 6, C. C. Little's DBA and related strains; Group 7, wild-derived strains. A detailed description of each group is included in the text. The length and angle of the branches have been optimized for printing and do not reflect the actual evolutionary distances between strains. Figure adapted from Petkov *et al.* (2004).

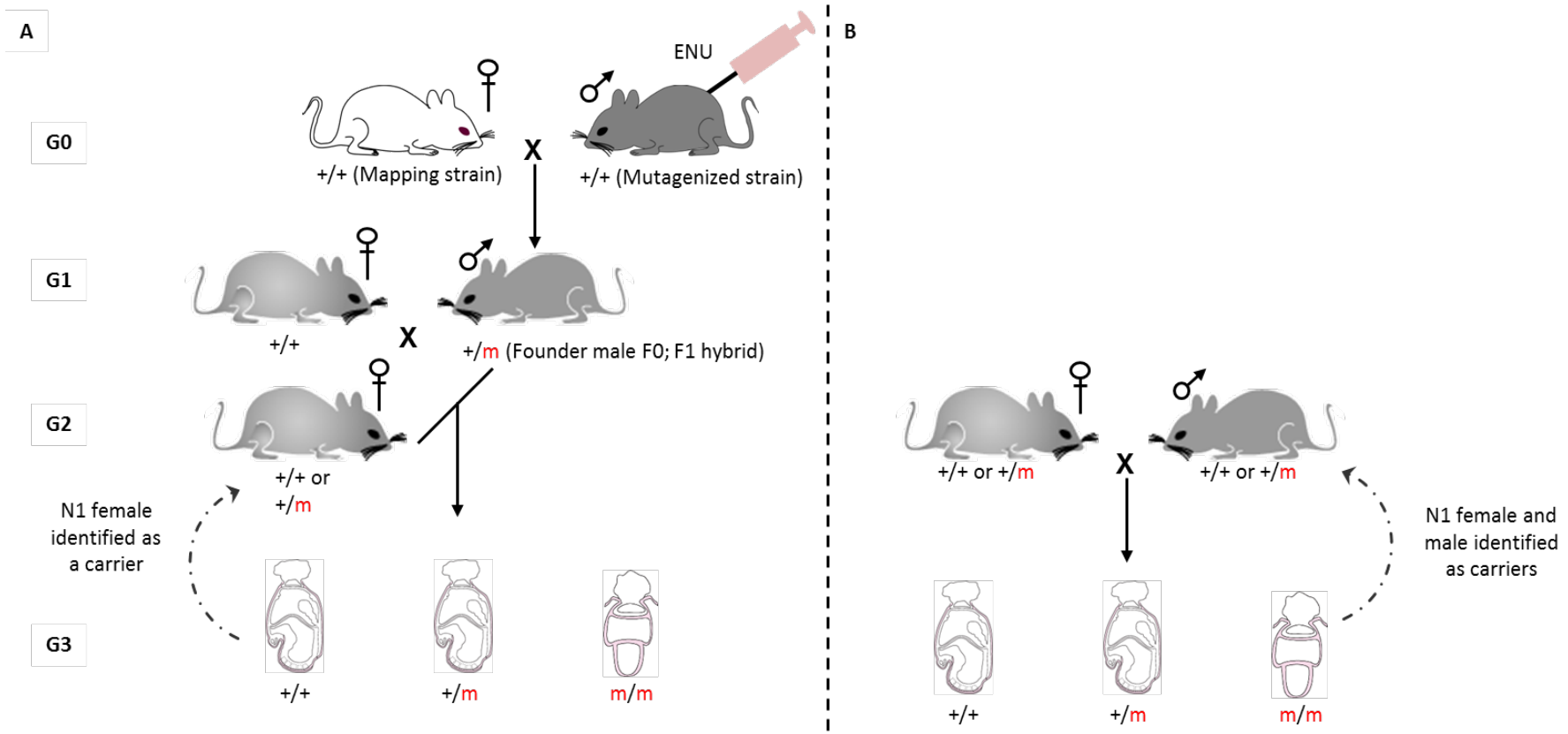
A
G0
G1
G2
G3
ENU
+/+ (Mapping strain)
X
+/+ (Mutagenized strain)
+/+
X
+/m (Founder male F0; F1 hybrid)
+/+ or
+/m
N1 female
identified as
a carrier
+/+
+/m
m/m
B
+/+ or +/m
X
+/+ or +/m
N1 female and
male identified
as carriers
+/+
+/m
m/m

Figure 4.3: The breeding and recombination mapping strategy for the gastrulation screen. (A) First, a mutagenised male mouse was crossed with a wild-type female from the mapping strain. Each male progeny from this cross represented the founder (F0) of a new pedigree. Each founder was crossed with a wild-type female from the mapping strain to produce G2 animals. G2 females (50% wild-type, 50% heterozygous for mutations inherited from the father) were backcrossed with the father (F0 male). If both the F0 male and G2 female were heterozygous for a newly induced mutation that impairs gastrulation, 25% of G3 progeny of the recovered embryos should exhibit a phenotype. These G3 embryos were recovered at the late gastrulation stage and visually inspected for defects. Overall, 12.5% of G3 embryos were expected to exhibit the same phenotype if the pedigree segregates a new, recessive mutation that impairs gastrulation. This would identify the backcrossed female as an N1 carrier, and her DNA would then be collected to conduct low-resolution recombination mapping. (B) G2 females were intercrossed (time-mated) with G2 males and their litters dissected at 8.0 dpc. If both parents were heterozygous for the mutation that impairs gastrulation, 25% of the recovered embryos should exhibit the known phenotype. This would identify the parents as N1 carriers, and the N1 male was selected for further breeding. G: generation, F0: founder, N: backcross animal.

Based on the average efficiency of ENU mutagenesis, we can estimate that each G1 progeny would be heterozygous for seventy new mutations (Noveroske *et al.*, 2000). To make these mutations, homozygous G1 males are mated with females of the mapping strain to produce G2 progeny, half of which are heterozygous for each particular mutation. G2 daughters are then intercrossed (time-mated) with the G1 father and the females are sacrificed at 8.0 dpc. The G3 embryos are recovered and screened for overt phenotypes by visual inspection. If the G2 female is heterozygous for the mutation, 25% of the G3 embryos will be homozygous for the mutation, as determined by their abnormal phenotypes. If the G2 female did not carry the mutation, all (or less than 25%) of the embryos within the litter would be normal. Overall, 1/8 (12.5%) of G3 embryos produced using this backcross breeding strategy should be homozygous. In the gastrulation screen, we aimed to examine at least thirty progenies per pedigree from a minimum of four pregnant G2 females. A pedigree was deemed worthy of follow-up when a similar mutant phenotype was found in at least two separate litters from the same pedigree. If homozygous mutants arise with a frequency of 1/8 and at least two mutants are required for follow up, then (using the binomial function) the probability of missing a mutation in thirty potential homozygous embryos is 0.096, which we considered an acceptable limit.

4.1.3 Low-resolution recombination mapping using interval haplotype analysis

The pedigree production protocol shown in Figure 4.3 A also served to generate recombinant DNAs for the first step in the positional candidate approach to identify the causative mutation in each gastrulation screen pedigree. One technical challenge posed by the gastrulation screen is that small mutant embryos (such as in Figure 4.1) do not contain sufficient genomic DNA to enable a low-resolution genome scan to be performed, as this is a procedure that would typically require sufficient DNA for 50-100

PCR reactions. To circumvent this difficulty, DNA from G2 daughters identified as carriers was used for recombination mapping. As shown in Figure 4.3 A, G2 daughters were backcrossed with their F1 fathers (F0 founder males) and, if heterozygous for a newly induced mutation that impairs gastrulation, 25% of the recovered embryos should exhibit a phenotype. Rather than using the mutant embryos for recombination mapping, DNA was collected from the G2 daughters. This G2 animal is designated the N1 animal (where N stands for backcross), as they are derived from the cross between an F1 hybrid and one of the parental strains (i.e. the mapping strain). For each N1 female from the gastrulation screen, meiotic recombination events in the germ cells of the F1 father can be identified using markers that are polymorphic between the BALB/cJ and FVB/N strains.

During the experiments designed to conduct low-resolution mapping of the causative variants in the gastrulation pedigrees, the F0 father would sometimes become too old for further breeding before sufficient N1 carrier females were identified for the mapping experiment. In these cases, the procedure shown in Figure 4.3 B was followed. G2 females were intercrossed (time-mated) with G2 males and their litters dissected at 8.0 dpc. If both were heterozygous for the mutation that impairs gastrulation, 25% of the recovered embryos should exhibit the (already known) phenotype. When this occurred, it retrospectively identified both the female as an N1 carrier suitable for the recombination mapping experiment and the male as a carrier suitable for breeding further progeny. As shown in Figure 4.4, this male was then mated with a wild-type female of the mapping strain and its female progeny assessed by backcross and dissection of litters at 8.0 dpc to identify N2 carrier females that could be used in the recombination mapping experiment.

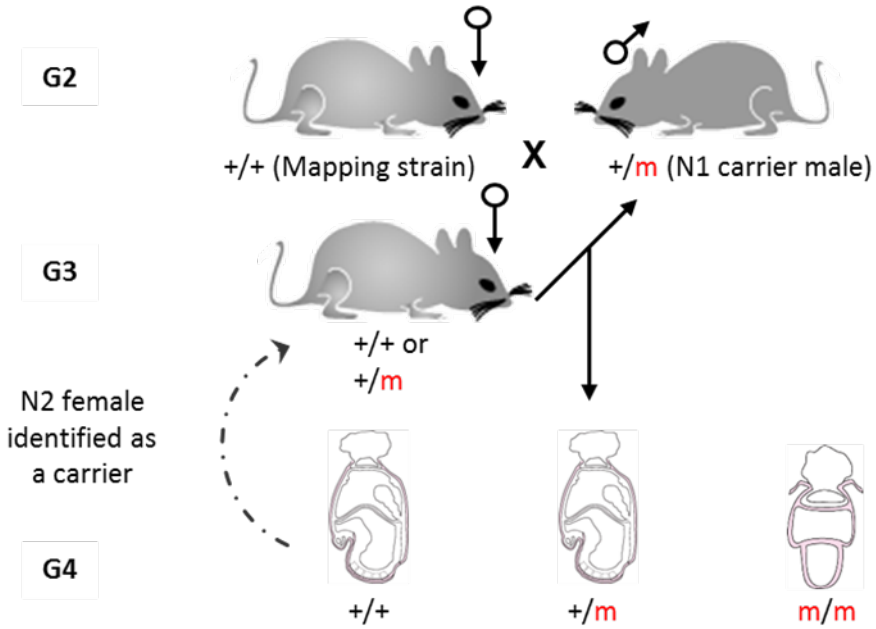

Figure 4.4: The breeding strategy used once F0 males were no longer available for breeding. The identified N1 carrier male is mated with a wild-type female from the mapping strain to produce G3 animals. Female progeny are assessed by backcross and dissection of litters at 8.0 dpc to identify N2 carrier females; their DNA can be collected for use in the recombination mapping experiment. G: generation.

One commonly used method for low-resolution mapping is to infer a haplotype based on the analysis of a small panel of polymorphic markers. A haplotype refers to the combination of alleles that are inherited together along the length of a chromosome. The interval haplotype mapping strategy (Neuhaus *et al.*, 1998) uses markers at the proximal and distal regions of all mouse autosomes (i.e. ***M****us* ***mu****sculus* [Mmu] chromosomes 1–19), with additional medial markers tested on three larger chromosomes (Mmu 1, Mmu 2, and Mmu 3). This divides the murine genome into approximately 22 intervals and enables segregation of the mutation to be followed in carrier animals. As shown in Figure 4.5, each haplotype indicates whether the chromosomal region has been inherited from the mutagenised or mapping strain (Potter *et al.*, 1995, Neuhaus *et al.*, 1998). In heterozygous carriers, the mutation must be in a region inherited from the mutagenised strain (i.e. a non-recombinant region); this region will be heterozygous for the mutagenised and mapping genomes. Linkage is therefore established when one chromosomal region is heterozygous among all confirmed carriers of the newly introduced mutation. In contrast, regions of the genome that are unlinked from the introduced mutation can undergo recombination without loss of the causative mutation. Such unlinked chromosomal regions will therefore tend towards homozygosity for the mapping strain. The interval haplotype analysis method requires only a small number of animals (approximately thirteen when empirically determined from other mapping experiments) to map a mutation to a specific chromosome. Often, by using the same DNA—or a small number of additional DNA samples—and a small number of additional polymorphic markers, a smaller MCR can be defined for each pedigree (Neuhaus *et al.*, 1998, Bogani *et al.*, 2004).

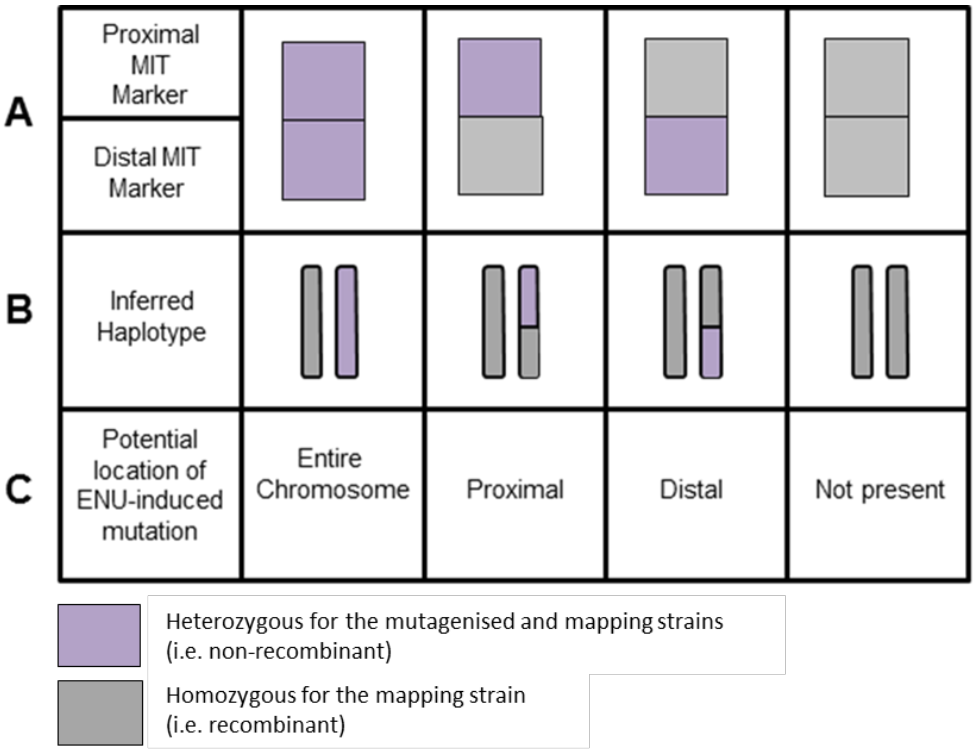

Figure 4.5: The principle of interval haplotype analysis. Analysis of proximal and distal MIT markers can determine whether recombination has occurred between these markers in parental meiosis. If no recombination occurred, the animal will remain heterozygous for the mapping and mutagenised strains (purple) at each marker. If recombination did occur, the animal will be homozygous for the mapping strain (grey) at one or both markers. (A) The genotype of an animal at each marker in the panel is tested. (B) The haplotype is inferred from the genotype results. (C) The potential location of the ENU-induced mutation can be inferred from the haplotype.

4.1.4 An MIT marker panel suitable for analysis of the BALB/cJ and FVB/N pedigrees

As previously stated, the gastrulation screen yielded eight novel mouse pedigrees that exhibited abnormal gastrulation phenotypes (Table 4.1). In the two pedigrees originating from the preliminary screen (11-23 and 11-68), the MCR had already been defined by other members of the Arkell laboratory. The location of the causative ENU-induced mutation for the remaining six pedigrees (16-5, 16-7, 18-5, 18-10, 18-19, and 18-22) was, however, unknown; linkage analysis was performed on each of these pedigrees. These pedigrees are referred to collectively as the BALB/cJ and FVB/N pedigrees.

The Arkell laboratory previously validated a panel of MIT markers suitable for distinguishing BALB/cJ and FVB/N genomes. The mouse genome informatics (MGI) database identified 125 MIT markers that may be suitable for use with the gastrulation screen (Table 1 and Table 2 in Appendix A). PCR amplicons from BALB/cJ and FVB/N DNA, as well as from an F1 hybrid animal (F0 founder male) produced by an outcross of these two strains, were analysed by agarose gel electrophoresis. Sixty-two polymorphic MIT markers were identified between the FVB/N and BALB/cJ inbred strains (see Figure 4.6 for a representative example of the analysis and Appendix B for optimised PCR conditions). To generate the MIT marker panel, one polymorphic marker at the proximal and distal regions of all autosomes (Mmu 1–19) was selected. A third intermediate marker was also selected for Mmu 1, Mmu 2, and Mmu 3 to give a panel of forty-one markers for the low-resolution mapping strategy (Figure 4.7). Eleven additional polymorphic markers were noted and subsequently used to increase the mapping resolution in chromosomal regions of interest.

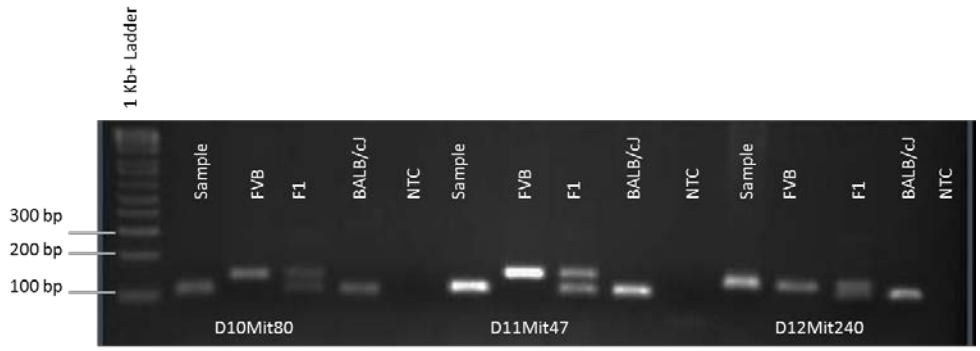

Figure 4.6: Representative example of agarose electrophoresis analysis of polymorphic MIT markers between the FVB/N and BALB/cJ inbred mouse strains. Genomic DNA samples from the FVB/N, BALB/cJ, and a F1 hybrid between the FVB/N and BALB/cJ inbred strains were amplified by PCR. The products were resolved on a 3% agarose gel. FVB: FVB/N, NTC: no template control.

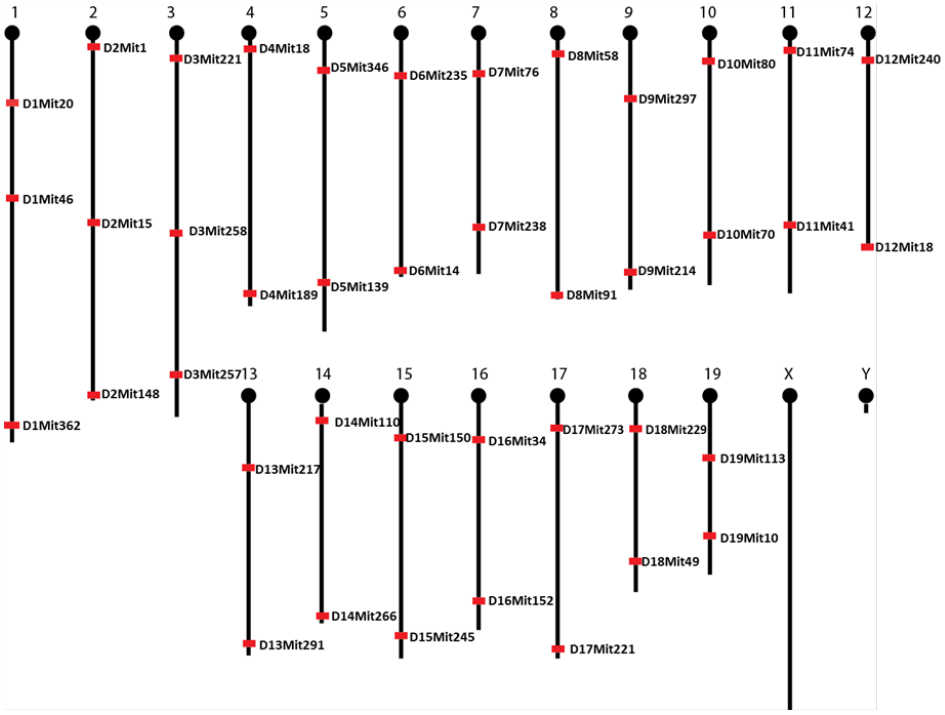

Figure 4.7: The FVB/N-BALB/cJ microsatellite marker panel. A total of forty-one markers were selected at the proximal and distal regions of each autosome. Additional medial markers on Mmu 1, Mmu 2, and Mmu 3 were also selected.

4.1.5 Aims

To identify genes that direct gastrulation, the overall aim of this chapter is to perform low-resolution recombination mapping on the six strains recovered from the BALB/cJ-FVB/N mutagenesis screen. This will identify an MCR for each strain that will limit the analysis required for future whole-exome sequencing experiments. It will also test the feasibility of a genome-wide screen for recessive mutations that affect gastrulation by determining whether the recovered phenotypes can be readily traced to one region of the genome.

4.2 RESULTS

4.2.1 Pedigree 16-5

Pedigree 16-5 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3). For this pedigree, the mutagenised strain was BALB/cJ and the mapping strain was FVB/N. Male G1 #9 was screened by backcross with five N1 females. Small embryos which failed to commence gastrulation were observed in three of the five litters examined, leading to an overall presumed homozygote frequency of 9% (N = 55). Thus, the male was considered to harbour a new recessive mutation that prevents gastrulation and was designated F0 #9. A further twenty-three N1 females were examined by backcross to F0 #9; the litters of twelve females were found to contain embryos with the expected phenotype at a frequency of 15% (N = 218). Ten N1 carrier females were selected for recombination mapping (shown in orange in Table 4.2). BALB/FVB marker panel analysis identified only one chromosome with strong linkage to the newly introduced mutation: Mmu 9. This was the only chromosome for which all animals were non-recombinant (i.e. heterozygous for the mutagenised and mapping strain) for at least one marker on the chromosome (Table 4.2). More carrier females were needed to generate additional genomic DNA samples for recombination mapping. The F0 #9 male was too old for further breeding and, using the strategy shown in Figure 4.3 B, additional male carriers (N1 #234, N1 #285, N1 #284, and N1 #287) were identified by intercross of the N1 progeny. The strategy shown in Figure 4.4 was then used to identify ten additional carrier N2 females from which genomic DNA could be isolated for further mapping.

Pedigree 16-5
Animal
(N1 #17 female x F0 # 9male)
(N1 #19 female x F0 #9 male)
(N1 #25 female x F0 #9 male)
(N1 #43 female x F0 #9 male)
(N1 #44 female x F0 #9 male)
(N1 #45 female x F0 #9 male)
(N1 #47 female x F0 #9 male)
(N1 #223 female x F0 #9 male)
(N1 #225 female x F0 #9 male)
(N1 #236 female x F0 #9 male)
(N2 #80 female x N1 #234 male)
(N2 #89 female x N1 #234 male)
(N2 #68 female x N1 #234 male)
(N2 #71 female x N1 #234 male)
(N2 #224 female x N1 #285 male)
(N2 #196 female x N1 #285 male)
(N2 #207 female x N1 #285 male)
(N2 #343 female x N1 #285 male)
(N2 #141 female x N1 #284 male)
(N1 #313 female x N1 #287 male)
Chr Marker Mb
Chr1 D1Mit20 30.9
Chr1 D1Mit46 75.6
Chr1 D1Mit362 190
Chr2 D2Mit1 3.8
Chr2 D2Mit15 91.9
Chr2 D2Mit148 179
Chr3 D3Mit221 7.92
Chr3 D3Mit258 82.1
Chr3 D3Mit257 142
Chr4 D4Mit18 13.9
Chr4 D4Mit189 146
Chr5 D5Mit346 4.8
Chr5 D5Mit139 125
Chr6 D6Mit235 23.1
Chr6 D6Mit14 146
Chr7 D7Mit76 19.6
Chr7 D7Mit238 107
Chr8 D8Mit58 7.1
Chr8 D8Mit91 127
Chr9 D9Mit297 34
Chr9 D9Mit214 115
Chr Marker Mb
Chr10 D10Mit80 11.4
Chr10 D10Mit70 104
Chr11 D11Mit74 5.2
Chr11 D11Mit41 89.9
Chr12 D12Mit240 11.3
Chr12 D12Mit18 114
Chr13 D13Mit217 29.2
Chr13 D13Mit76 111
Chr14 D14Mit110 9
Chr14 D14Mit266 112
Chr15 D15Mit150 14.1
Chr15 D15Mit245 94.1
Chr16 D16Mit34 15.9
Chr16 D16Mit152 86.3
Chr17 D17Mit273 9.1
Chr17 D17Mit221 90.5
Chr18 D18Mit229 12.3
Chr18 D18Mit49 76.2
Chr19 D19Mit113 19.4
Chr19 D19Mit10 47.2

Table 4.2: Genome-wide low-resolution mapping of N1 and N2 carrier females for pedigree 16-5. The genotype (homozygous FVB/N: grey, or heterozygous FVB/N; BALB/cJ: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Only one region (Mmu 9) showed strong linkage to the mutant phenotype exhibited in pedigree 16-5.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

These genomic DNA samples (shown in black in Table 4.2) were subjected to marker panel analysis; the results confirmed Mmu 9 as the most likely site of the mutation. Only one female (N2 #343) was recombinant at both the proximal and distal markers on Mmu 9 (Table 4.2).

Recombinant chromosomes can arise for two reasons: first, the N2 female may have been incorrectly assigned as a carrier, or second, a double recombination could have occurred. In N1 animals, in which the recombinant products from only one meiosis are examined, the frequency of double recombination should be low (<1%). In N2 animals, the result of two meioses is examined; even in linked regions, a double recombination could occur. To evaluate the possibility of double recombination in N2 female #343, additional Mmu 9 markers (D9Mit4, D9Mit12, and D9Mit347) were examined. As shown in Figure 4.8, this analysis indicated that a region of heterozygosity existed on Mmu 9 in this animal. In combination with the marker panel analysis of the additional females, the causative mutation was mapped to a 15.2 Mb interval of distal chromosome 9 between D9Mit12 and D9Mit214 (99.8–115 Mb).

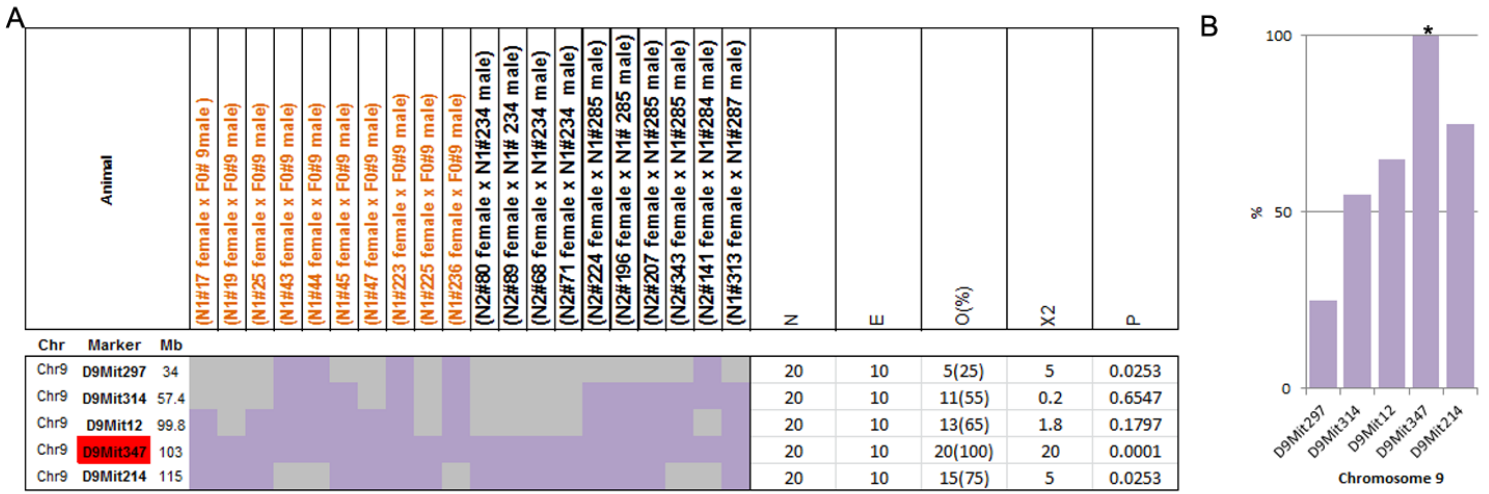

Figure 4.8: Increased resolution mapping on Mmu 9 using additional markers for pedigree 16-5. (A) The genotype (homozygous FVB: grey, or heterozygous FVB; BALB/cJ: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for Mmu 9. All carrier animals were heterozygous for marker D9Mit347 and the mutation is linked to the interval 99.8–115 Mb. (B) The percentage of 16-5 carrier animals (N = 20) that retained the mutagenised genome at each chromosome 9 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.001, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

4.2.2 Pedigree 16-7

Pedigree 16-7 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3) using BALB/cJ as the mutagenised strain and FVB/N as the mapping strain. Male G1 #11 was screened by backcross with four N1 females. Embryos with a small embryonic region and large extra-embryonic region were observed in two litters at an overall frequency of 15% (N = 40). Thus, the male (G1 #11) was considered to harbour a new recessive mutation that prevents the onset of gastrulation and was designated F0 #11. Genomic DNA was isolated from the four N1 carrier females for low-resolution recombination mapping. A further twenty-one N1 females were examined by backcross to F0 #11 and eleven females were found to contain embryos with the expected phenotype at a frequency of 18% of the collected embryos (N = 203). Recombination mapping with these thirteen N1 carrier females (shown in orange in Table 4.3) and the BALB/FVB marker panel analysis identified two regions (Mmu 2 and Mmu 7) in which only one animal was recombinant at both the proximal and distal markers, indicating strong linkage to the newly introduced mutation (Table 4.3).

Recombinant chromosomes may have arisen because the N1 female was incorrectly assigned as a carrier, or because a double recombination had occurred. To determine the cause of recombination, additional carrier females were identified. The F0 #11 male was too old for further breeding and additional carriers were identified by intercross of the N1 progeny (Figure 4.3 B). Each of four N1 males was test-mated with two N1 females. Through this process, one male (N1 #175) and his tested females were identified as carriers, with 23% of embryos displaying the phenotype of interest.

Pedigree 16-7
Animal
(N1 #23 female x F0 #11 male)
(N1 #27 female x F0 #11male)
(N1 #37 female x F0 #11 male)
(N1 #38 female x F0 #11 male)
(N1 #170 female x F0 #11 male)
(N1 #171 female x F0 #11 male)
(N1 #173 female x F0 #11 male)
(N1 #181 female x F0 #11 male)
(N1 #182 female x F0 #11 male)
(N1 #184 female x F0 #11 male)
(N1 #185 female x F0 #11 male)
(N1 #186 female x F0 #11 male)
(N1 #187 female x F0 #11 male)
(N1 #191 female x N1 #175 male)
(N1 #226 female x N1 #175 male)
Chr Marker Mb
Chr1 D1Mit20 31
Chr1 D1Mit46 76
Chr1 D1Mit362 190
Chr2 D2Mit1 3.8
Chr2 D2Mit15 92
Chr2 D2Mit148 179
Chr3 D3Mit221 7.9
Chr3 D3Mit258 82
Chr3 D3Mit257 142
Chr4 D4Mit18 14
Chr4 D4Mit189 146
Chr5 D5Mit346 4.8
Chr5 D5Mit139 125
Chr6 D6Mit235 23
Chr6 D6Mit14 146
Chr7 D7Mit76 20
Chr7 D7Mit238 107
Chr8 D8Mit58 7.1
Chr8 D8Mit91 127
Chr9 D9Mit297 34
Chr9 D9Mit214 115
Chr Marker Mb
Chr10 D10Mit80 11
Chr10 D10Mit70 104
Chr11 D11Mit74 5.2
Chr11 D11Mit41 90
Chr12 D12Mit240 11
Chr12 D12Mit18 114
Chr13 D13Mit217 29
Chr13 D13Mit76 111
Chr14 D14Mit110 9
Chr14 D14Mit266 112
Chr15 D15Mit150 14
Chr15 D15Mit245 94
Chr16 D16Mit34 16
Chr16 D16Mit152 86
Chr17 D17Mit273 9.1
Chr17 D17Mit221 91
Chr18 D18Mit229 12
Chr18 D18Mit49 76
Chr19 D19Mit113 19
Chr19 D19Mit10 47

Table 4.3: Genome-wide low-resolution mapping of N1 carrier females for pedigree 16-7. The genotype (homozygous FVB/N: grey, or heterozygous FVB/N; BALB/cJ: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Two regions (Mmu 2 and Mmu 7) showed strong linkage to the mutant phenotype exhibited in pedigree 16-7.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

Genomic DNA samples from the two females (shown in black in Table 4.3) were subjected to marker panel analysis; the results indicated that Mmu 7 was the most likely site of the mutation (Table 4.3). In parallel, additional markers

(D7Mit247, D7Mit69, and D7Mit276) were examined along Mmu 7 to evaluate the possibility of double recombination in N1 female #182. As shown in Figure 4.9, this analysis indicated that a region of heterozygosity existed on Mmu 7 (shown in green in Figure 4.9). In combination with the marker panel analysis of the additional females, this indicated a strong linkage to Mmu 7 (19.6–69.4 Mb, Figure 4.9).

To confirm linkage of the 16-7 causative mutation to Mmu 7, putative homozygous embryos were collected from further intercross matings between N1 females and the male carrier N1 #175. The embryos were genotyped along Mmu 7 and all mutant embryos were found to be homozygous for the mutagenised strain (BALB/cJ), confirming that the causative mutation was mapped between D7Mit247 and D7Mit69 (39–56.3 Mb, Figure 4.10). In contrast, normal embryos were homozygous for the mapping strain or heterozygous at the same tested markers. These findings indicate the mutation causing the abnormal gastrulation phenotype in pedigree 16-7 resides within the 49.8 Mb interval of Mmu 7.

A

Chr	Marker	Mb	(N1#23 female x F0# 11 male)	(N1#27 female x F0# 11male)	(N1#37 female x F0#11 male)	(N1#38 female x F0#11 male)	(N1#170 female x F0#11 male)	(N1#171 female x F0#11 male)	(N1#173 female x F0#11 male)	(N1#181 female x F0#11 male)	(N1#182 female x F0#11 male)	(N1#184 female x F0#11 male)	(N1#185 female x F0#11 male)	(N1#186 female x F0#11 male)	(N1#187 female x F0#11 male)	(N1#191 female x N1#175 male)	(N1#226 female x N1#175 male)	N	E	O(%)	X2	P
Chr7	D7Mit76	19.6																15	7.5	11(73.3)	3.267	0.0707
Chr7	D7Mit247	39											-					14	7	14(100)	14	0.0002
Chr7	D7Mit69	56.3																15	7.5	15(100)	15	0.0001
Chr7	D7Mit276	69.4																15	7.5	14(93.3)	11.267	0.0008
Chr7	D7Mit238	107.3																15	7.5	8(53.3)	0.067	0.7963

B

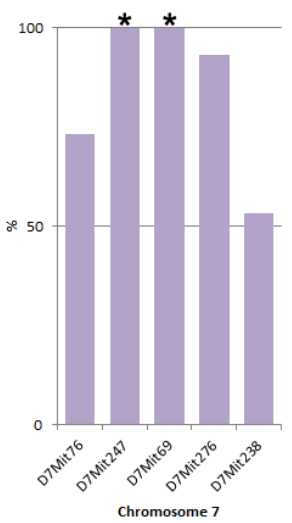

Figure 4.9: Increased resolution mapping on Mmu 7 using additional markers for pedigree 16-7. (A) The genotype (homozygous FVB/N: grey, or heterozygous FVB/N; BALB/cJ: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for Mmu 7. All carrier animals were heterozygous for markers D7Mit247 and D7Mit69; the mutation is linked to the interval 19.6–69.4 Mb. (B) The percentage of N1 16-7 carrier females (N = 14 or 15) that retained the mutagenised genome at each chromosome 7 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.001, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

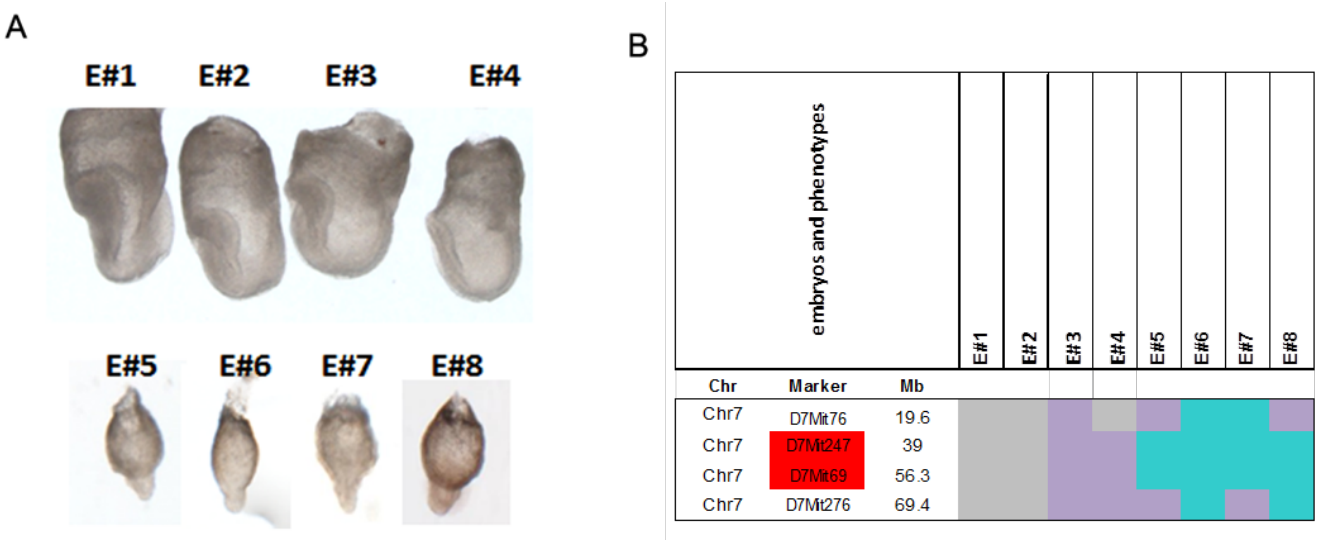

Figure 4.10: Linkage confirmation using presumed 16-7 homozygous embryos. (A) Lateral view of embryos dissected from one litter at 8.0 dpc. Embryos with wild-type phenotype (E1-4) are on the top row and embryos with the mutant phenotype (E5-8) are on the bottom row. (B) The genotype (homozygous FVB/N: grey, heterozygous FVB/N; BALB/cJ: purple or homozygous BALB/cJ: blue) of each embryo at each MIT marker in the linked Mmu 7 interval is shown. All embryos with an abnormal phenotype were homozygous for D7Mit247 and D7Mit69, confirming linkage to this interval.

	Heterozygous
	Homozygous (Mapping strain)
	Homozygous (Mutagenised strain)

4.2.3 Pedigree 18-5

Pedigree 18-5 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3). For this pedigree the mutagenised strain was FVB/N and the mapping strain was BALB/cJ. Male G1 #17 was screened by backcross with five N1 females. Embryos which displayed a small embryonic region and large extra-embryonic region were apparent in two litters at a frequency of 16% (N = 56). Thus, the male (G1 #17) was considered to harbour a new recessive mutation that prevents the onset of gastrulation and was designated F0 #17. An additional nine N1 females (from nineteen dissected) were found to contain embryos with the expected phenotype at an overall frequency of 16% (N = 197). Genomic DNA from the eleven N1 carrier females was used for low-resolution recombination mapping via the BALB/FVB marker panel (shown in orange in Table 4.4). Two chromosomes (Mmu 9 and Mmu 16) exhibited strong linkage, as all animals were non-recombinant for at least one marker on these chromosomes (Table 4.4).

The F0 #17 male was too old for further breeding, and additional male carriers were identified by intercross of N2 animals, as shown in Figure 4.3 B. This process identified one male (N1 #114) and one N2 female as carriers. Genomic DNA from the female (shown in black in Table 4.4) was analysed with the mapping panel; the female was found to be recombinant at a marker on Mmu 9, suggesting that Mmu 16 is the most likely site of the mutation (Table 4.4). To investigate this theory, five additional markers were analysed along Mmu 9, revealing weak linkage between D9Mit1001 and D9Mit347 from the genotyping of two identified N1 female carries (N1 #13 and N1 #17). Given the recombination in this region, as evidenced by the homozygosity, it is expected that the mutation is not linked to Mmu 9 (Figure 4.11). Chi square analysis of the markers on Mmu 9 (D9Mit1001 and D9Mit4) revealed a significant deviation from the frequency

expected in unlinked markers (Figure 4.11). However, given the haplotype analysis, it is expected that this significance was not due to linkage of these markers to the mutation of interest.

As shown in Figure 4.12, three additional markers were examined along Mmu 9 and 16; this analysis confirmed linkage to Mmu 16 (15.9–53.9 Mb). This result was statistically significant, confirming that the causative mutation was mapped between D16Mit34 and D16Mit152 (Figure 4.12). These findings indicate that the mutation causing the abnormal gastrulation phenotype in pedigree 18-5 resides within the 38 Mb interval of Mmu 16.

Animal
(N2 #168 female x N1 #114 male)
(N1 #232 female x F0 # 17 male)
(N1 #24 female x F0 #17 male)
(N1 #22 female x F0 #17 male)
(N1 #20 female x F0 #17 male)
(N1#17 female x F0#17 male)
(N1 #16 female x F0 #17 male)
(N1 #15 female x F0 #17 male)
(N1 #13 female x F0 #17 male)
(N1 #8 female x F0 #17 male)
(N1 #6 female x F0 #17 male)
(N1 #4 female x F0 #17 male)
Chr Marker Mb
Chr1 D1Mit20 31
Chr1 D1Mit46 76
Chr1 D1Mit362 190
Chr2 D2Mit1 3.8
Chr2 D2Mit15 92
Chr2 D2Mit148 179
Chr3 D3Mit221 7.9
Chr3 D3Mit258 82
Chr3 D3Mit257 142
Chr4 D4Mit18 14
Chr4 D4Mit189 146
Chr5 D5Mit346 4.8
Chr5 D5Mit139 125
Chr6 D6Mit235 23
Chr6 D6Mit14 146
Chr7 D7Mit76 20
Chr7 D7Mit238 107
Chr8 D8Mit58 7.1
Chr8 D8Mit91 127
Chr9 D9Mit297 34
Chr9 D9Mit214 115
Chr10 D10Mit80 11
Chr10 D10Mit70 104
Chr11 D11Mit74 5.2
Chr11 D11Mit41 90
Chr12 D12Mit240 11
Chr12 D12Mit18 114
Chr13 D13Mit217 29
Chr13 D13Mit76 111
Chr14 D14Mit110 9
Chr14 D14Mit266 112
Chr15 D15Mit150 14
Chr15 D15Mit245 94
Chr16 D16Mit34 16
Chr16 D16Mit152 86
Chr17 D17Mit273 9.1
Chr17 D17Mit221 91
Chr18 D18Mit229 12
Chr18 D18Mit49 76
Chr19 D19Mit113 19
Chr19 D19Mit10 47

Table 4.4: Genome-wide low-resolution mapping of N1 and N2 carrier females for pedigree 18-5. The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Two regions (Mmu 9 and Mmu 16) showed strong linkage to the mutant phenotype exhibited in pedigree 18-5.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

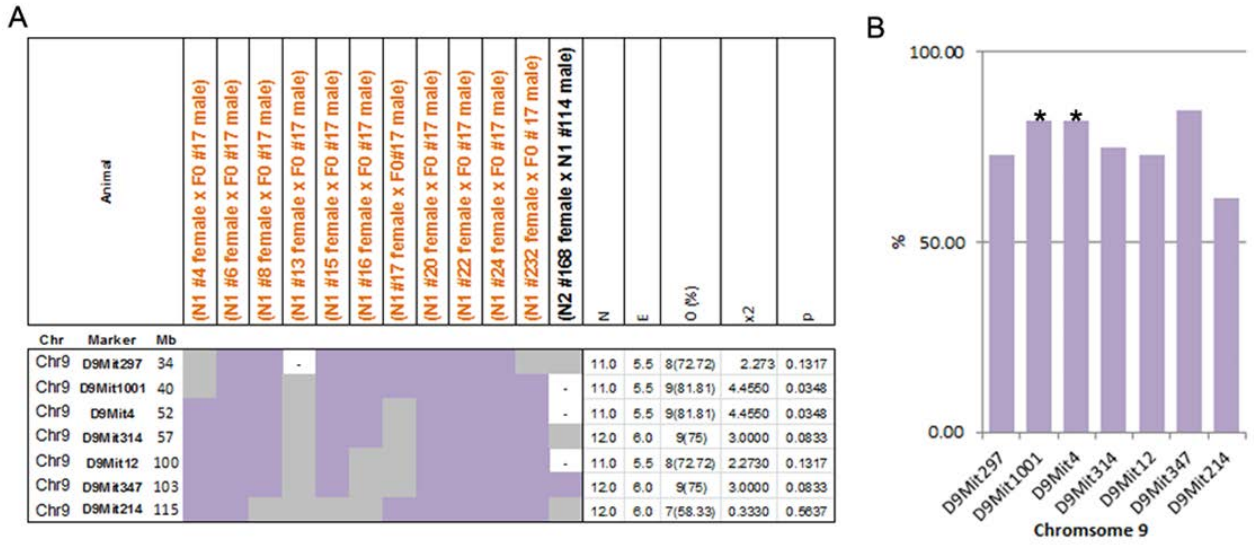

Figure 4.11: Increased resolution mapping on Mmu 9 using additional markers for pedigree 18-5. (A) The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for Mmu 9. (B) The percentage of carrier 18-5 animals (N = 11 or 12) that retained the mutagenised genome at each chromosome 9 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.05, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

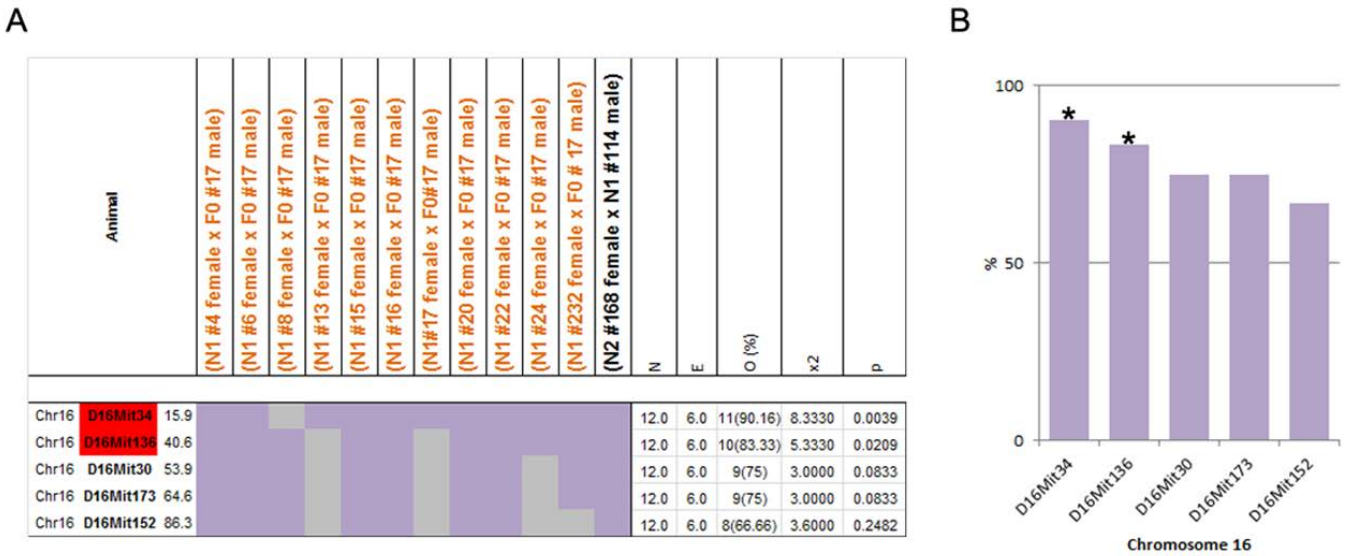

Figure 4.12: Increased resolution mapping on Mmu 16 using additional markers for pedigree 18-5. (A) The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for Mmu 16. Ten or eleven carrier females were heterozygous for markers D16Mit34 and D16Mit36; the mutation is linked to the interval 15.9–40.6 Mb. (B) The percentage of carrier 18-5 animals (N = 12) that retained the mutagenised genome at each chromosome 16 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.05, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

4.2.4 Pedigree 18-10

Pedigree 18-10 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3). For this pedigree the mutagenised strain was FVB/N and the mapping strain was BALB/cJ. Male G1 #32 was screened by backcross with six N1 females, two of which contained embryos with patterning errors at an overall frequency of 14% (N = 80). Thus, the male was considered to harbour a new recessive mutation that caused early patterning defects and was designated F0 #32. A further twenty-six litters from N1 females backcrossed to F0 #32 were examined and putative homozygous embryos were observed in six of these at an overall frequency of 8% (N = 236). This is lower than the 12.5% expected, and suggests that this mutation may not be fully penetrant. Genomic DNA samples were isolated from the eight N1 carrier females for low-resolution recombination mapping using the BALB/FVB marker panel analysis (shown in orange in Table 4.5). Two chromosomes (Mmu 5 and Mmu 12) exhibited strong linkage, since all animals were non-recombinant for at least one marker on these chromosomes (Table 4.5).

Additional carrier females were sought to refine the linkage position. The F0 #32 male was too old for further breeding and additional male carriers were identified by intercross of the N1 progeny (Figure 4.3 B). This identified three males (N1 #252, N1 #263, and N1 #265) and paired N1 females as carriers. When genomic DNA from the identified N1 carrier females (shown in black in Table 4.5) was subjected to the marker panel analysis, the results indicated that Mmu 12 was the most likely site of the mutation (Table 4.5). When one additional marker (D12Mit136) was examined, strong linkage to Mmu 12 was found, but did not further narrow the mapped interval (Figure 4.13).

Pedigree 18-10

Chr	Marker	Mb	(N1 #5 female x F0 #32 male)	(N1 #52 female x F0 #32 male)	(N1 #53 female x F0 #32 male)	(N1 #13 female x F0 #32 male)	(N1 #287 female x F0 #32 male)	(N1 #293 female x F0 #32 male)	(N1 #317female x F0 #32 male)	(N1 #314 female x F0 #32 male)	(N1 #217 female x N1 #252 male)	(N1 #233 female x N1 #252 male)	(N1 #192 female x N1 #263 male)	(N1 #334 female x N1 #263 male)	(N1 #189 female x N1 #265 male)	(N1 #269 female x N1 #265 male)
Chr1	D1Mit20	30.9														
Chr1	D1Mit46	75.6														
Chr1	D1Mit362	190												-		
Chr2	D2Mit1	3.8														
Chr2	D2Mit15	91.9												-		
Chr2	D2Mit148	179														
Chr3	D3Mit221	7.92														
Chr3	D3Mit258	82.1														
Chr3	D3Mit257	142														
Chr4	D4Mit18	13.9												-		
Chr4	D4Mit189	146														
Chr5	D5Mit346	4.8														
Chr5	D5Mit139	125														
Chr6	D6Mit235	23.1												-		
Chr6	D6Mit14	146												-		
Chr7	D7Mit76	19.6												-		
Chr7	D7Mit238	107												-		
Chr8	D8Mit58	7.1												-		
Chr8	D8Mit91	127												-		
Chr9	D9Mit297	34														
Chr9	D9Mit214	115														
Chr10	D10Mit80	11.4														
Chr10	D10Mit70	104														
Chr11	D11Mit74	5.2														
Chr11	D11Mit41	89.9														
Chr12	D12Mit240	11.3														
Chr12	D12Mit18	114												-		
Chr13	D13Mit217	29.2														
Chr13	D13Mit76	111														
Chr14	D14Mit110	9														
Chr14	D14Mit266	112														
Chr15	D15Mit150	14.1												-		
Chr15	D15Mit245	94.1												-		
Chr16	D16Mit34	15.9														
Chr16	D16Mit152	86.3												-		
Chr17	D17Mit273	9.1														
Chr17	D17Mit221	90.5														
Chr18	D18Mit229	12.3												-		
Chr18	D18Mit49	76.2														
Chr19	D19Mit113	19.4														
Chr19	D19Mit10	47.2														

Table 4.5: Genome-wide low-resolution mapping of N1 carrier females for pedigree 18-10. The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Only one region (Mmu 12) showed strong linkage to the mutant phenotype exhibited in pedigree 18-10.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

A

Chr	Marker	Mb	(N1 #5 female x F0 #32 male)	(N1 #52 female x F0 #32 male)	(N1 #53 female x F0 #32 male)	(N1 #13 female x F0 #32 male)	(N1 #287 female x F0 #32 male)	(N1 #293 female x F0 #32 male)	(N1 #317female x F0 #32 male)	(N1 #314 female x F0 #32 male)	(N1 #217 female x N1 #252 male)	(N1 #233 female x N1 #252 male)	(N1 #192 female x N1 #263 male)	(N1 #334 female x N1 #263 male)	(N1 #189 female x N1 #265 male)	(N1 #269 female x N1 #265 male)	N	E	O(%)	X2	P
Chr12	D12Mit240	11.3															14	7	7(50)	0	1
Chr12	D12Mit136	24.3				-		-	-								11	5.5	4(42.85)	0.818	0.3657
Chr12	D12Mit18	114												-			13	6.5	11(84.61)	6.231	0.0126

B

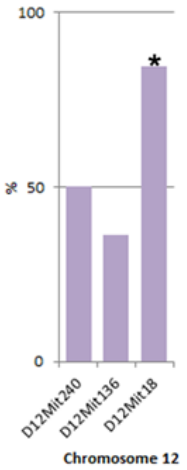

Figure 4.13: Increased resolution mapping on Mmu 12 using additional markers for pedigree 18-10. (A) The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each female at each MIT marker is shown. The results indicate the haplotype of each female for Mmu 12. Eleven out of thirteen carrier animals were heterozygous for marker D12Mit18. (B) The percentage of carrier 18-10 animals (N = 11 or 14) that retained the mutagenised genome at each chromosome 12 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.001, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

4.2.5 Pedigree 18-19

Pedigree 18-19 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3), using FVB/N as the mutagenised strain and BALB/cJ as the mapping strain. Male G1 #24 was screened by backcross with four N1 females. Embryos with a small embryonic region and large extra-embryonic region were apparent in two litters at an overall frequency of 24%. Thus, the male was considered to harbour a new recessive mutation that caused a failure to gastrulate and was designated F0 #24. Subsequently, a further four N1 carrier females (from eight tested) were identified by backcross with F0 #24; an overall frequency of 15% putative homozygous embryos was observed (N = 60). Recombination mapping with five N1 carrier females and the BALB/FVB marker panel identified four regions (Mmu 4, Mmu 17, Mmu 18, and Mmu 19) with strong linkage to the newly introduced mutation, since no animals were recombinant at both markers on these chromosomes (Table 4.6).

Additional carrier females were needed to confirm on linkage to one of these chromosomes. As the F0 #24 male was too old for further breeding, additional carriers were identified by intercross of the N1 progeny. Three males (N1 #208, N1 #209, and N1 #227) and paired females were identified as carriers. Additionally, the strategy shown in Figure 4.4 was used to generate N2 females for backcross with the newly identified N1 carrier males. An additional seven carrier females were identified (shown in black in Table 4.6) and their genomic DNA subjected to marker panel analysis. The data indicated that Mmu 4 was the most likely site of the mutation (Table 4.6). One additional marker (D4Mit26) was examined along Mmu 4 to evaluate the strength of the linkage. As shown in Figure 4.14, in combination with the marker panel analysis of the additional females, results suggested the mutation causing the abnormal gastrulation phenotype in pedigree 18-19 resides within a 70 Mb interval of Mmu 4.

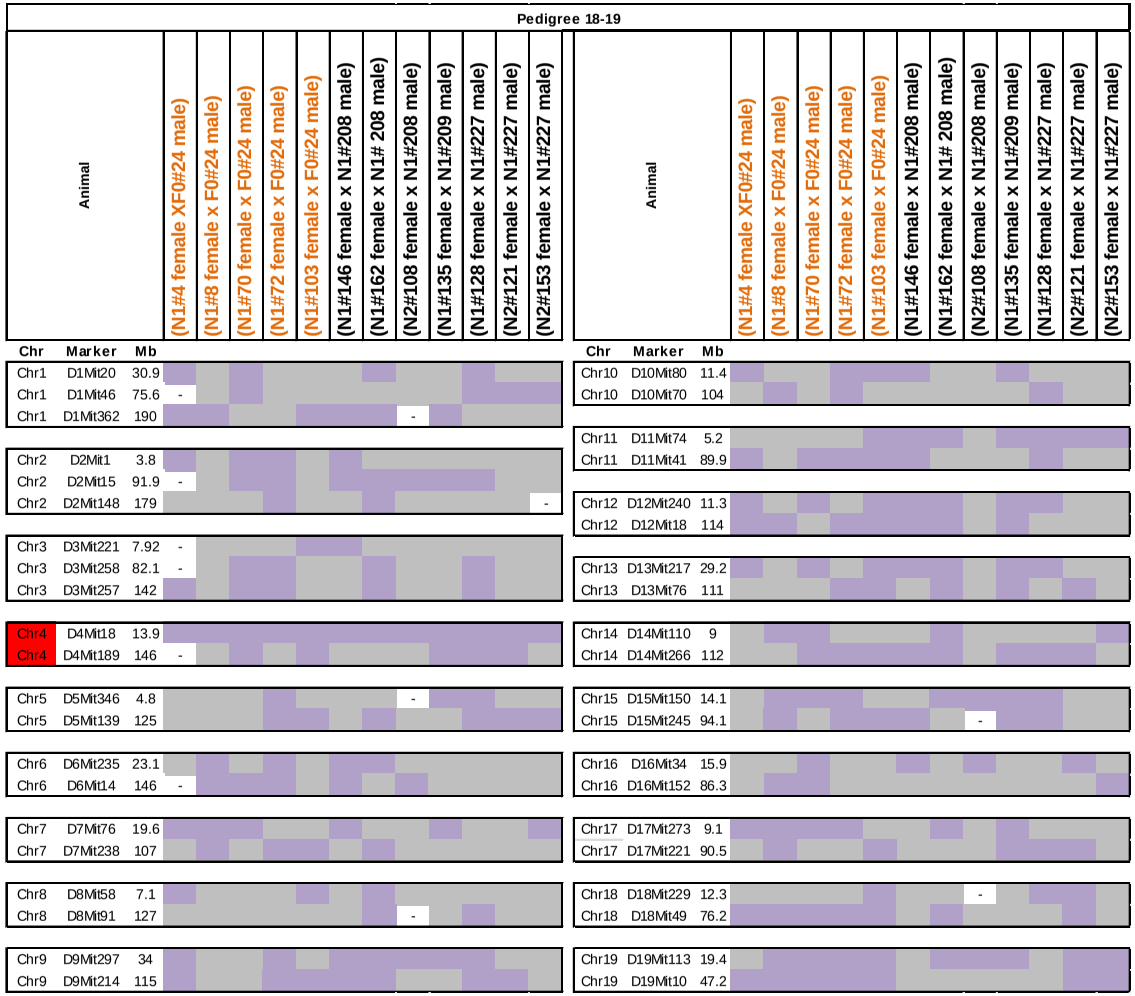
Pedigree 18-19
Animal
(N1#4 female XF0#24 male)
(N1#8 female x F0#24 male)
(N1#70 female x F0#24 male)
(N1#72 female x F0#24 male)
(N1#103 female x F0#24 male)
(N1#146 female x N1#208 male)
(N1#162 female x N1# 208 male)
(N2#108 female x N1#208 male)
(N1#135 female x N1#209 male)
(N1#128 female x N1#227 male)
(N2#121 female x N1#227 male)
(N2#153 female x N1#227 male)
Chr Marker Mb
Chr1 D1Mit20 30.9
Chr1 D1Mit46 75.6
Chr1 D1Mit362 190
Chr2 D2Mit1 3.8
Chr2 D2Mit15 91.9
Chr2 D2Mit148 179
Chr3 D3Mit221 7.92
Chr3 D3Mit258 82.1
Chr3 D3Mit257 142
Chr4 D4Mit18 13.9
Chr4 D4Mit189 146
Chr5 D5Mit346 4.8
Chr5 D5Mit139 125
Chr6 D6Mit235 23.1
Chr6 D6Mit14 146
Chr7 D7Mit76 19.6
Chr7 D7Mit238 107
Chr8 D8Mit58 7.1
Chr8 D8Mit91 127
Chr9 D9Mit297 34
Chr9 D9Mit214 115
Chr10 D10Mit80 11.4
Chr10 D10Mit70 104
Chr11 D11Mit74 5.2
Chr11 D11Mit41 89.9
Chr12 D12Mit240 11.3
Chr12 D12Mit18 114
Chr13 D13Mit217 29.2
Chr13 D13Mit76 111
Chr14 D14Mit110 9
Chr14 D14Mit266 112
Chr15 D15Mit150 14.1
Chr15 D15Mit245 94.1
Chr16 D16Mit34 15.9
Chr16 D16Mit152 86.3
Chr17 D17Mit273 9.1
Chr17 D17Mit221 90.5
Chr18 D18Mit229 12.3
Chr18 D18Mit49 76.2
Chr19 D19Mit113 19.4
Chr19 D19Mit10 47.2

Table 4.6: Genome-wide low-resolution mapping of N1 and N2 carrier females for pedigree 18-19. The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Only one region (Mmu 4) showed strong linkage to the mutant phenotype exhibited in pedigree 18-19.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

Figure 4.14: Increased resolution mapping on Mmu 4 using additional markers for pedigree 18-19. (A) The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each female at each MIT marker is shown. The results indicate the haplotype of each female for Mmu 4. All carrier females were heterozygous for markers D4Mit18 and the mutation is linked to the interval 13.9–84 Mb. (B) The percentage of carrier 18-19 females (N = 12 or 9) that retained the mutagenised genome at each chromosome 4 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (*p* = 0.001, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

4.2.6 Pedigree 18-22

Pedigree 18-22 was isolated from the gastrulation screen via the interspecific backcross strategy (Figure 1.3), using FVB/N as the mutagenised strain and BALB/cJ as the mapping strain. Male G1 #50 was screened by backcross with four N1 females. Embryos displaying gastrulation patterning errors were apparent in two litters at an overall frequency of 18% (N = 38). Thus, the male was considered to harbour a new recessive mutation that caused defective patterning and was designated F0 #50. A further twenty-nine litters were examined; nine were found to contain embryos with the putative homozygous mutants at an overall frequency of 11% (N = 234). Less than 50% of females produced embryos with this phenotype, indicating the mutation may be incompletely penetrant. Genomic DNA was isolated from the eleven N1 carrier females for low-resolution recombination mapping using the BALB/FVB marker panel analysis (shown in orange in Table 4.7). Of the eleven females, only one showed recombination at both markers on Mmu 9, whereas all other autosomes had multiple fully recombinant animals. Mmu 9 was considered to show the strongest linkage (Table 4.7).

The recombinant Mmu 9 in female N1 #157 could have arisen if she was incorrectly assigned as a carrier, or because of a double recombination event. To distinguish these possibilities, additional markers (D9Mit4, D9Mit12, and D9Mit347) were examined along Mmu 9. This showed that the N1 #157 female was homozygous along the length of Mmu 9, which led to its exclusion from the analysis (Figure 4.15 A). Additional carrier females were needed to complete the recombination mapping. As the F0 #50 male was too old for further breeding, intercross of N1 progeny identified two males (N1 #160 and N1 #169) and paired N1 females as carriers. Additionally, N2 females were backcrossed with male carrier N1 #160 to isolate extra genomic DNA samples for recombination mapping (shown in black in Table 4.7). This analysis confirmed Mmu 9 as the only site

with strong linkage. The genomic DNA samples were then subjected to additional Mmu 9 marker analysis (D9Mit4, D9Mit12, and D9Mit347), which indicated that the mutation causing the abnormal gastrulation phenotype in pedigree 18-22 resides within a 15 Mb interval of Mmu 9.

Pedigree 18-22
Animal
(N1#3 female x F0#50 male)
(N1#4 female x F0#50 male)
(N1#40 female x F0#50 male)
(N1#41 female x F0#50 male)
(N1# 42 female x F0#50 male)
(N1#147 female x F0#50 male)
(N1#157 female x F0#50 male)
(N1#185 female x F0#50 male)
(N1#187 female x F0#50 male)
(N1#249 female x F0#50 male)
(N1#251 female x F0#50 male)
(N2#37 female x N1# 160 male)
(N2# 38 female x N1#160 male)
(N2#39 female x N1#160 male)
(N2#55 female x N1#160 male)
(N1#109 female x N1#160 male)
(N2#120 female x N1#160 male)
(N1#146 female x N1#160 male)
(N1#108 female x N1# 169 male)
(N1#140 female x N1#169 male)
Chr Marker Mb
Chr1 D1Mit20 30.9
Chr1 D1Mit46 75.6
Chr1 D1Mit362 190
Chr2 D2Mit1 3.8
Chr2 D2Mit15 91.9
Chr2 D2Mit148 179
Chr3 D3Mit221 7.92
Chr3 D3Mit258 82.1
Chr3 D3Mit257 142
Chr4 D4Mit18 13.9
Chr4 D4Mit189 146
Chr5 D5Mit346 4.8
Chr5 D5Mit139 125
Chr6 D6Mit235 23.1
Chr6 D6Mit14 146
Chr7 D7Mit76 19.6
Chr7 D7Mit238 107
Chr8 D8Mit58 7.1
Chr8 D8Mit91 127
Chr9 D9Mit297 34
Chr9 D9Mit214 115
Animal
Chr Marker Mb
Chr10 D10Mit80 11.4
Chr10 D10Mit70 104
Chr11 D11Mit74 5.2
Chr11 D11Mit41 89.9
Chr12 D12Mit240 11.3
Chr12 D12Mit18 114
Chr13 D13Mit217 29.2
Chr13 D13Mit76 111
Chr14 D14Mit110 9
Chr14 D14Mit266 112
Chr15 D15Mit150 14.1
Chr15 D15Mit245 94.1
Chr16 D16Mit34 15.9
Chr16 D16Mit152 86.3
Chr17 D17Mit273 9.1
Chr17 D17Mit221 90.5
Chr18 D18Mit229 12.3
Chr18 D18Mit49 76.2
Chr19 D19Mit113 19.4
Chr19 D19Mit10 47.2

Table 4.7: Genome-wide low-resolution mapping of N1 and N2 carrier females for pedigree 18-22. The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each animal at each MIT marker is shown. The results indicate the haplotype of each animal for all autosomes (Mmu 1–19). Only one region (Mmu 9) showed strong linkage to the mutant phenotype exhibited in pedigree 18-22.

	Heterozygous
	Homozygous (Mapping strain)
-	No result

A

Chr	Marker	Mb	(N1#3 female x F0#50 male)	(N1#4 female x F0#50 male)	(N1#40 female x F0#50 male)	(N1#41 female x F0#50 male)	(N1# 42 female x F0#50 male)	(N1#147 female x F0#50 male)	(N1#157 female x F0#50 male)	(N1#85 female x F0#50 male)	(N1#87 female x F0#50 male)
Chr9	D9Mit297	34									
Chr9	D9Mit4	52									
Chr9	D9Mit12	100									
Chr 9	D9Mit347	103									
Chr9	D9Mit214	115									

B

Chr	Marker	Mb	(N1#3 female x F0#50 male)	(N1#4 female x F0#50 male)	(N1#40 female x F0#50 male)	(N1#41 female x F0#50 male)	(N1# 42 female x F0#50 male)	(N1#147 female x F0#50 male)	(N1#85 female x F0#50 male)	(N1#87 female x F0#50 male)	(N2#37 female x N1#160 male)	(N2# 38 female x N1#160 male)	(N2#39 female x N1#160 male)	(N2#55 female x N1#160 male)	(N1#109 female x N1#160 male)	(N1#146 female x N1#160 male)	(N1#108 female x N1# 169 male)	(N1#140 female x N1#169 male)	N	E	O(%)	X2	P
Chr9	D9Mit297	34																	16	8	13(81.25)	18.00	0.0001
Chr9	D9Mit4	52																	16	8	15(93.75)	14.22	0.0002
Chr9	D9Mit12	100																	16	8	15(93.75)	14.22	0.0002
Chr 9	D9Mit347	103																	16	8	16(100)	16.00	0.0001
Chr9	D9Mit214	115																	16	8	12(75)	3.56	0.0593

C

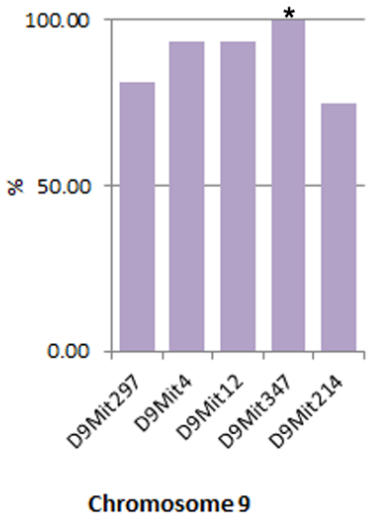

Figure 4.15: Increased resolution mapping on Mmu 9 using additional markers for pedigree 18-22. (A) The genotype (homozygous BALB/cJ: grey, or heterozygous BALB/cJ; FVB/N: purple) of each N1 female at each MIT marker is shown. The N1 #157 female (highlighted in green) was homozygous along Mmu 9. (B) The genotype of each N1 and N2 carrier female at each MIT marker is shown. The results indicate the haplotype of each female for Mmu 9. All carrier females were heterozygous for marker D9Mit347; the mutation is linked to the interval 100–115 Mb. (C) The percentage of carrier 18-22 females (N = 16) that retained the mutagenised genome at each chromosome 9 marker. *: markers for which the rate of heterozygosity was significantly above the 50% expected of an unlinked marker (p = 0.001, Chi squared test).

	Heterozygous
	Homozygous (Mapping strain)
-	No result

4.3 DISCUSSION

4.3.1 Overview

Despite the revolutionary use of ENU-based genetic screens in the mouse, a genome-wide screen that directly aims to identify genes that act in gastrulation has not previously been attempted. Several factors appear to mitigate the possibility of success for such a screen, including the difficulty of recovering gastrulation stage embryos, the difficulty of distinguishing mutant embryos from resorptions and the small embryo size that leads to insufficient recovery of DNA for extensive molecular analysis. In this chapter, I have contributed to a direct test of the feasibility of a genome-wide screen for ENU-induced alleles that alter gastrulation. I attempted to use a haplotype interval strategy to perform low-resolution recombination mapping for six newly isolated mouse strains. Each strain was thought to be segregating a new, recessive ENU-induced variant that prevented normal gastrulation. In each case, genomic DNA from carrier animals (rather than presumed homozygous animals) was analysed with a panel of polymorphic markers. For all six strains, the mutations were localised to a single chromosome and the MCR was narrowed to a sub-chromosomal interval for five of the strains. This indicates that the associated phenotypes likely arise from a single DNA variant, rather than a combination of mutations. This discovery will facilitate the cloning of the causative variant in future studies.

4.3.2 Phenotype classes at gastrulation

The gastrulation pedigrees discussed in this chapter come from a limited genetic screen in which only eight pedigrees were extensively investigated. Despite this small number, it appears that the strains can be grouped into distinct classes based on their phenotypes. The eight pedigrees can be broadly grouped into three classes:

Class 1: failure to thrive

In pedigrees 11-23 and 16-5, the mutant embryos exhibited a severe reduction in size compared to their presumed wild-type or heterozygous littermates, and both the embryonic and extra-embryonic regions were affected to a similar degree. These embryos mimicked those in which a gene required for embryo nutrition was mutated (EG); this class was designated 'failure to thrive'.

Class 2: failure to gastrulate

Pedigrees 16-7, 18-5, and 18-19 all contained presumed homozygotes that, like the class 1 embryos, were much smaller than their presumed wild-type or heterozygous littermates. In contrast to the class 1 mutants, the abnormal embryos in these pedigrees exhibited differential growth in the embryonic versus extra-embryonic regions. In all class 2 mutant pedigrees, the failure to grow was more severe in the embryonic region: the tissue from which all adult tissues are derived through gastrulation. Additionally, the mutant embryos lacked hallmarks of gastrulation such as an obvious primitive streak and evident anterior-posterior axis. Rather than a failure to thrive, these embryos appear to have not initiated gastrulation, and thus were designated 'failure to gastrulate'.

Class 3: patterning defects

Pedigrees 11-68, 18-10, and 18-22 contained presumed homozygous embryos in which the process of gastrulation was clearly initiated, as judged by the presence of the primitive streak. The embryos, however, failed to grow at the normal rate through gastrulation (i.e. they remained smaller than their presumed wild-type or heterozygous littermates) and lacked evidence of early patterning events such as the formation of the

neural plate. These three phenotypes were therefore considered to have arisen from 'patterning defects'.

Many patterning defects have been identified in other forward genetics experiments that examined later stages of embryogenesis (Kasarskis *et al.,* 1998 and Herron *et al.,* 2002). It is possible that the embryos of pedigrees 11-68, 18-10, and 18-22 have survived long enough to be recovered and screened for later phenotypes. The remaining five pedigrees, however, exhibited phenotypes that would not have been recovered in screens that examine embryos at 9.5 dpc.

4.3.3 Recombination mapping of the BALB/FVB pedigrees

Six mouse strains were selected for low-resolution recombination mapping and via a combination of strategies and an MCR was defined for each of these. As shown in Table 4.8, the size of the MCR varied from 15–103 Mb. In general, two factors influenced the size of the MCR: the density of polymorphic markers and the number of recombinant chromosomes. For any two inbred strains of mice, the density of polymorphic markers varied across the genome; this was evident in the present study. For example, in the case of pedigree 18-10, it was difficult to identify additional polymorphic markers at Mmu 12, which contributed to an inability to localise the mutation to a sub-chromosomal interval. Similarly, for pedigree 18-19, a scarcity of polymorphic markers on Mmu 4 contributed to the large—70 Mb—MCR on that chromosome. The density of markers affected all pedigrees equally, as all mapping experiments were conducted between the same two inbred strains.

The number of recombinant chromosomes also influenced the mapping outcomes, but this was a feature that affected the pedigrees to varying degrees; specifically, the ease

Table 4.8: Chromosomal location of the ENU-induced mutations in the gastrulation screens. In this study, a male mouse strain was mutagenised and crossed with a female mapping strain. The mutagenised and mapping strains are indicated for each of the eight isolated mouse pedigrees.

Mutant pedigree	Mutagenised inbred strain	Mapping inbred strain	Chromosomal location
11-23	C57Bl/6J	C3H/HeH	Mmu 3 V259E allele of *Atp1a1*
11-68	C57Bl/6J	C3H/HeH	Mmu 6 (32.1-88 Mb)
16-5	BALB/cJ	FVB/N	Mmu 9 (99.8-115 Mb)
16-7	BALB/cJ	FVB/N	Mmu 7 (19.6-69.5 Mb)
18-5	FVB/N	BALB/cJ	Mmu 16 (15.9-53.9 Mb)
18-10	FVB/N	BALB/cJ	Mmu 12 (11.3-114 Mb)
18-19	FVB/N	BALB/cJ	Mmu 4 (13.9-84 Mb)
18-22	FVB/N	BALB/cJ	Mmu 9 (100-115 Mb)

of correctly assigning carrier status varied across the strains. One feature that influences this is phenotype robustness. Phenotypes can vary with respect to penetrance and expressivity. The penetrance of a phenotype is defined as the proportion of individuals carrying a certain allele of a gene that also express an associated trait (the phenotype). When this proportion falls below 100%, the phenotype is said to have incomplete penetrance. The expressivity of a phenotype refers to the fact that not all phenotypes are manifested to the same extent in different individuals. Together, these two factors can lead to difficulty in inferring genotypes based on an observed phenotype; false negative conclusions can be reached (i.e. an individual that carries a mutant genotype could be assigned a non-carrier). In practical terms, the risk of false negative conclusions can be somewhat diminished by using mouse strains with large litter sizes. The cross between BALB/c and FVB/N animals routinely yielded litters with at least twelve embryos, in which the expected number of homozygotes was three. Even in pedigrees with incomplete penetrance and/or variable expressivity, more than one embryo with an overt phenotype would be expected.

In the gastrulation screen, this problem of incorrectly assigning carrier status was compounded by the fact that the litters were examined when resorption of wild-type embryos was still occurring. In class 3 embryos, this can lead to false positive conclusions where resorbing embryos are identified as putative homozygous embryos. Additionally, homozygous mutant embryos could wrongly be designated as a resorbing wild-type embryo (leading to false negative carrier identification).

Another aspect of robustness is how readily a phenotype can be shifted or modified. Modifying effects include environmental and genetic factors. The mouse genetic experiments in this chapter were conducted, as far as possible, under controlled and

consistent environmental conditions. The genetic background of each strain was, however, under constant flux because of continual backcrossing to the mapping strain. One recommendation, based on the recombination mapping described in this chapter, is that the number of pedigrees under analysis should be kept small to enable sufficient carrier females to be identified before the F0 male becomes too old for further mating (which requires carrier animals to be identified after another backcross).

Over the course of the recombination mapping experiments, it was noted that difficulties with phenotype loss during backcross generations were encountered most often with the patterning mutants. This was reflected in the fact that an increased number of potential carriers needed to be screened to identify sufficient animals for mapping; this produced an overall frequency of presumed homozygotes of less than the expected 12.5%. This indicator of incomplete penetrance was noted for both patterning defect pedigrees mapped in this chapter (18-10 and 18-22), but not for any of the three failure to gastrulate pedigrees (16-7, 18-5, and 18-19). Although the sample size was small, it is possible that this observation reflects a real difference in the ease with which the phenotypes themselves can be detected and/or adapted. These phenotypes differed in their time-of-onset relative to the screening time, i.e. the block to gastrulation occurred some 36 hours before the 8.0 dpc screen and was therefore strongly evident by the time at which the embryos were assessed. In contrast, when examined at 8.0 dpc, the patterning defects were relatively subtle when compared with the expected effects of the same mutation at the stages usually used to assess patterning (i.e. 9.5–10.5 dpc). Hence, a small change in the aberrant patterning phenotype could lead to it being overlooked in an 8.0 dpc screen. A second reason for the apparent difference in phenotype stability between these classes is that they essentially scored different

classes of variables. Patterning is a continuous variable, and a genetic modifier that ameliorates the phenotype does not need to cause a large phenotype shift to render the phenotype undetectable in a gross morphology screen. In contrast, a failure to gastrulate can be better considered a discrete variable (gastrulation versus no gastrulation), and overcoming a complete block to the onset of gastrulation may require more genetic variation than it is possible to obtain from different inbred strains of mice.

Overall, despite these difficulties, it was possible to map all pedigrees selected for this stage of the preliminary gastrulation screen. This indicates that, as in other reported ENU mutagenesis experiments, most phenotypes have a monogenic basis and should be amenable to the identification of the underlying genetic lesion. As shown in Figure 4.16, the recessive mutation that caused the abnormal phenotype in the eight pedigrees from the Arkell gastrulation screen localised to seven different chromosomal regions. The causative mutation in pedigrees 16-5 and 18-22 both mapped to distal chromosome 9; it is possible that these represent different alleles of the same gene. The 16-5 and 18-22 mutants fell into separate classes (16-5 was a class 1, failure to thrive mutant and 18-22 was a class 2, patterning defect mutant), which suggests they may have arisen from distinct mutations on Mmu 9. Moreover, there are more than 300 genes within this interval according to the MGI database (Blake *et al.,* 2011); it is entirely possible that these mutations affected different genes.

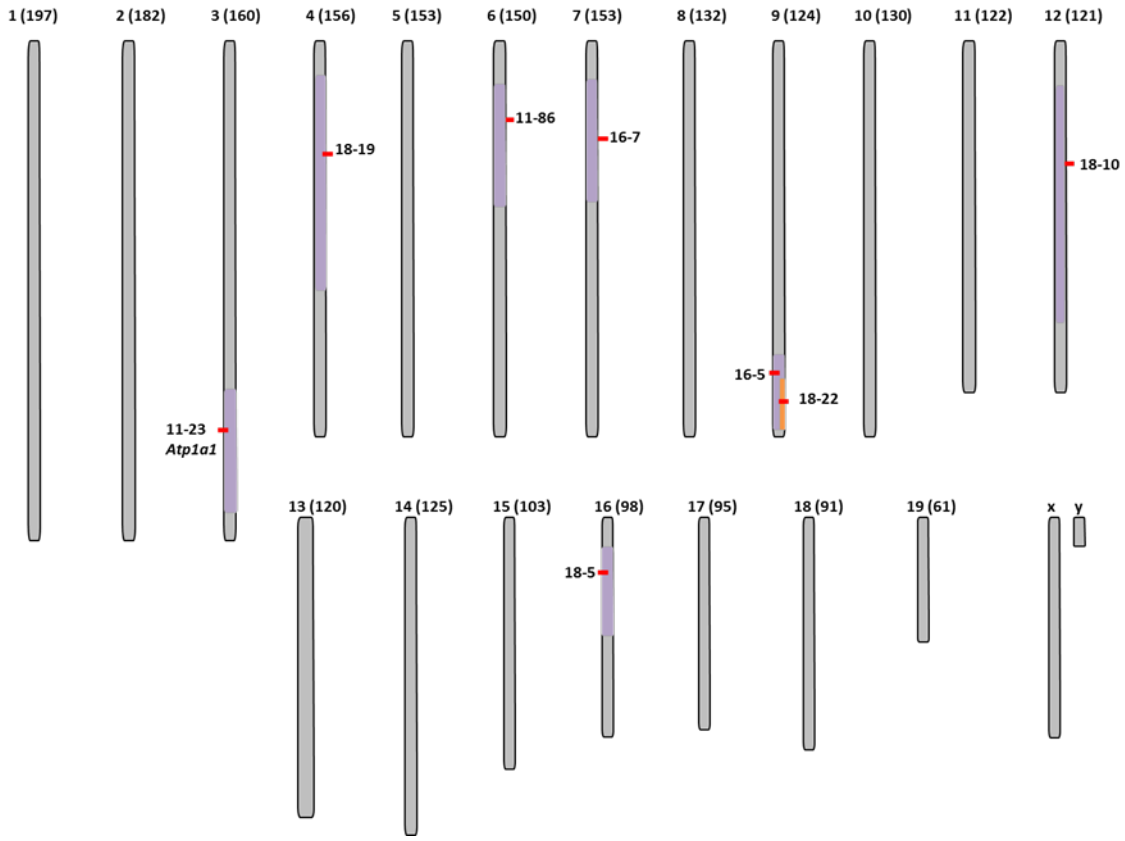

Figure 4.16: The chromosomal positions for each ENU-induced mutation of the gastrulation screen. Seven distinct chromosomal regions have been mutagenised. On chromosome 9, the map position for pedigree 16-5 is shown in purple, and the map position for 18-22 is shown in orange. The chromosome number is indicated at the top.

4.4 Future work

Each of the six pedigrees from the small-scale BALB/FVB gastrulation screen can now be subjected to the next phase of mutant identification, which aims to identify the relevant variant DNA sequence within the MCR. The work in this chapter has confirmed that low resolution mapping of strains obtained from a genetic screen for gastrulation mutants is feasible. It has also identified features of screening design and breeding strategies that should be incorporated into any larger-scale efforts to use ENU mutagenesis to identify genes required for mammalian gastrulation.

Chapter 5: The gastrulation screen: identifying the pathogenic variant

5.1 INTRODUCTION

As described in chapter 1 (section 1.3.3), following recombination mapping, the ENU-induced pathogenic variant underlying a particular phenotype needs to be identified within the minimal critical region (MCR), which is generally accomplished through DNA sequencing of all exons and splice sites across the MCR. As the pilot gastrulation screen (i.e. experiment ENU-11) was being conducted, however, the development of lower-cost whole exome sequencing (WES) technologies presented an opportunity to decrease the resolution of the recombination mapping step and increase the number of exons and splice sites scanned by sequencing. At the time, this approach had been successful in identifying homozygous mutations mapped with high resolution using single-nucleotide polymorphism (SNP) panels. For the gastrulation screen, however, it was necessary to identify the mutation in a carrier (heterozygous) animal, since the homozygous embryos did not contain sufficient genomic DNA for WES. Additionally, because of cost considerations, it was desirable to commence WES after low resolution mapping and to use a small number of carrier animals. To determine if such a strategy was possible, the 11-23 pedigree was selected for WES.

As described in section 4.1.2, the ENU-11 experiment was conducted as a pilot screen to determine the best strategy for pedigree production. In this experiment, the mutagenized strain was C57BL/6J and the mapping strain was C3H/HeH. The reference strain was also C57BL/6J, which aided the analysis. The 11-23 strain exhibited a failure to thrive (Figure 4.1) and had already undergone low-resolution mapping at the beginning of this project. Using procedures similar to those described in chapter 4, a

total of thirteen N1 and N2 carrier animals were identified and subjected to marker analysis using an interval haplotype strategy. This identified one region of strong linkage on MMu3, with a ~70 Mb MCR between 87,174,449 and 154,440,247 Mb.

5.2 RESULTS:

5.2.1 WES of the 11-23 MCR

The 11-23 MCR contains ~166 genes, and rather than using additional breeding to narrow the MCR, we applied next generation sequencing technologies. Based on the mutagenic frequency of ENU, the number of true exonic SNVs within a mutated genome is expected to be low (approximately 70 exonic SNVs in the F0 animal) (Keays *et al.*, 2006). Furthermore, this number should be reduced by ~50% in each generation during the initial mapping phase of our experiment. Typically, however, raw WES data included many more SNVs than the expected number. The additional variant calls can be attributed to multiple sources, including genetic drift of the C57BL/6J mouse strain versus the reference genome, known strain-specific polymorphisms for the mapping C3H strain, and the frequency of sequencing errors in massively parallel sequencing (Andrews *et al.*, 2012). To reduce the raw variant calls to a smaller number, highly enriched for ENU-induced mutations, a series of filters can be applied to remove known variants (present in dbSNP) and/or recurrent, platform–specific, false-positive variants. As the majority of causative ENU-induced alleles have been shown to be exonic, WES was selected as a suitable procedure for detecting the 11-23 causative mutation.

The whole exom sequencing and the single nucleotide variants were preformed at Australia Phenomic Genotyping Core Facility (APF, Australia National University). and The presumed homozygous embryos contained insufficient DNA for WES, and instead,

two confirmed carrier animals (G3 carrier males 18 and 54) were selected for the analysis. Genomic DNA was isolated from the kidneys, which were collected when the animals were euthanized at the end of the mapping experiment. The sequence reads were mapped to the NCBIM37 assembly of the reference mouse genome, and the identified SNVs were filtered to include those that occurred in exons and splice donor/acceptor sites and exclude known strain-specific and platform-specific variants. The resulting variant lists from each animal were constrasted, and a total of 33 heterozygous SNVs were common between the two datasets (Table 5.1). Three of the SNVs fell within the mapped interval (shown in blue). To determine which of the three variants within the MCR was likely to be causative, the functional annotations associated with each gene were investigated and the mutagenic potential of each identified SNV was assessed. The SNV identified at position 93,019,569 lies within the filaggrin family member 2 gene (*Flg2*; MGI: 3645678). It is one of five members of the S100 fused-type protein (SFTP) gene family,which also includes the protien trichohyalin (TCHH) (Lee *et al.*, 1993), hornerin (HRNR) (Makino *et al.*, 2001, Wu *et al.*, 2009b), repetin (RPTN) (Krieg *et al.*, 1997, Huber *et al.*, 2005), and cornulin (CRNN) (Xu *et al.*, 2000, Contzler *et al.*, 2005) in human and mice. The S100 genes encode a family of EF-hand calcium-binding proteins involved in the regulation of various cellular processes including cell growth, cell cycle regulation, differentiation, transcription, and motility (Zhao *et al.*, 1997, Marenholz *et al.*, 2004).

FLG2 is involved in epithelial homoeostasis and expressed in several tissues, including healthy human skin, where it is required for proper cornification (Smith *et al.*, 2006, Wu *et al.*, 2009a). In addition, the protein functions in skin barrier protection (Dale *et al.*,

1985, Brown *et al.*, 2012, Kezic *et al.*, 2014). Not surprisingly, defects in *FLG2* are associated with skin diseases (Smith *et al.*, 2006, Makino *et al.*, 2014).

Table 5.1: Validated potential SNVs detected in 11-23 pedigree exome sequencing.

chr	coord	ref_base	var_base	bases_and_counts	read_depth	median_base_quality	aa_change	gene_name
1	90108967	C	T	C:T:A:N:.:G_82:50:0:0:0:0	132	39	H->Y	0
1	90131829	G	A	G:A:T:N:C:._71:48:0:0:0:0	119	37	C->Y	Heatr7b1
1	158296543	G	T	T:G:A:N:C:._10:9:0:0:0:0	19	37	R->L	0
1	46086035	G	A	G:A:T:N:C:._8:5:0:0:0:0	13	37	P->S	Gm4776
1	11738127	A	C	A:C:T:N:.:G_8:4:0:0:0:0	12	38	H->P	A830018L16Rik
1	58703760	C	T	T:C:A:N:.:G_6:3:0:0:0:0	9	40	L->F	AC118698.1
2	14141002	T	A	A:T:N:C:.:G_39:23:0:0:0:0	62	39	I->N	Fam23a
2	32778609	A	G	A:G:T:N:C:._22:13:0:0:0:0	35	36	R->G	Fam129b
3	24232140	G	A	G:A:C:T:N:._34:31:1:0:0:0	66	38	G->S	AC139675.1
3	24232150	G	T	G:T:A:N:C:._34:31:0:0:0:0	65	38	G->V	AC139675.1
3	37613895	G	C	G:C:A:T:N:._8:5:0:0:0:0	13	40	P->A	0
3	93019569	A	T	A:T:N:C:.:G_22:18:0:0:0:0	40	37	Q->L	Flg2
3	101394617	A	T	T:A:N:C:.:G_17:15:0:0:0:0	32	38	V->E	Atp1a1
3	93979673	G	A	A:G:T:N:C:._13:13:0:0:0:0	26	38	A->V	Gm10696
4	145126763	A	G	A:G:C:T:N:._149:140:59:0:0:0	348	39.5	I->V	0
4	145126997	C	T	T:C:A:N:.:G_15:12:0:0:0:0	27	39	L->F	0
6	132157741	G	C	G:C:A:T:N:._77:40:0:0:0:0	117	38	P->A	Prb1
6	130072166	C	T	C:T:A:N:.:G_57:33:0:0:0:0	90	39	G->S	0
6	41564308	C	A	A:C:T:N:.:G_17:14:0:0:0:0	31	27	Y->*	Ephb6
6	132262491	T	G	T:G:A:N:C:._5:4:0:0:0:0	9	39	Q->H	Prpmp5
7	46663474	G	C	C:G:A:T:N:._45:38:0:0:0:0	83	38	P->R	Gm5114
7	24417780	T	G	T:G:A:N:C:._15:8:0:0:0:0	23	36	V->G	Vmn1r171
7	24417789	G	C	G:C:T:A:N:._11:8:2:1:1:0	23	36	R->P	Vmn1r171
8	26564987	C	T	C:T:A:N:.:G_48:25:0:0:0:0	73	38	D->N	0
8	41910115	G	A	G:A:T:N:C:._23:19:0:0:0:0	42	36.5	R->Q	Adam39
10	117642845	G	A	A:G:T:N:C:._21:12:0:0:0:0	33	38	V->I	Il22
12	24759537	C	A	C:A:T:N:.:G_25:14:0:0:0:0	39	37	S->I	0
14	51301112	C	A	A:C:T:N:.:G_7:4:0:0:0:0	11	40	A->S	Olfr747
14	54089653	C	T	C:T:A:N:.:G_4:2:0:0:0:0	6	40	S->L	TRAV8-1
15	76545859	A	C	A:C:T:N:.:G_22:21:0:0:0:0	43	38	D->E	0
15	91564504	G	A	G:A:T:N:C:._16:9:0:0:0:0	25	39	C->Y	Lrrk2
17	53647836	C	T	T:C:A:N:.:G_2:2:0:0:0:0	4	36.5	R->K	0
19	60543644	A	G	A:G:T:N:C:._30:13:0:0:0:0	43	36	F->S	Prlhr

Ch, chromosome; coord, coordinate; Ref_base, C57BL/6J reference genome base; var_base, most common variant base from sequencing; base read_depth, the absolute read depth; median_base_quality, median of the quality scores from the raw sequence data; aa, amino acid.

FLG2 expression is restricted to the epidermis (Wu, 2009a). *FLG2* mRNA was detected in thymus, stomach, tonsils, testis and placenta. The localisation of FLG2 in the epidermis and hair follicles is different from hornerin and trichohyalin, and similar to filaggrin, although filiggrin and FLG2 do not completely colocalise (Ishida-Yamamoto *et al.*, 1997, Wu *et al.*, 2009a, Wu *et al.*, 2009b). It has been suggested by Wu *et al.* (2009a) that FLG2 and filaggrin may play similar roles in skin biology, but may not directly cross-link with each other.

As there was no avaliable *in situ* data for this gene family in murine gastrulation, I used iTranscriptome, a database of genes expressed at the mid-gastrulation stage (7.0 dpc) designed by Chen *et al.,* (2017) to detect *Flg2* expression. Briefly, Chen *et al*., (2017) collected data from C57BL/6 embryos that were serially cyro-sectioned in the plane transverse to the proximal-distal axis. From a total of 22 sections, cells were sampled from 11 alternating sections. From each section, approximately 20 cells were isolated by laser capture microdissection (LCM) from epiblasts in each of the quadrants: anterior (A), posterior (P), left lateral (L), and right lateral (R), except for the distal-most section, where only anterior and posterior halves were sampled (Chen *et al.*, 2017). The methodology, which utilises geographical position sequencing (Geo-seq), combines LCM with single-cell RNA-seq technology. This allows study of the transcriptome of small samples, using as few as ten cells from defined locations in the embryo. Although Geo-seq cannot accurately report expression in individual cells, it can report gene expression in a specific region, indicating whether the gene is required for normal gastrulation or not. As shown in Figure 5.1, *Flg2* expression was barely detectable at mid-gastrulation, with only one cluster of cells in the L anterior epiblast (red) displaying high expression (Figure 5.1 A), and a group of cells in the distal epiblast

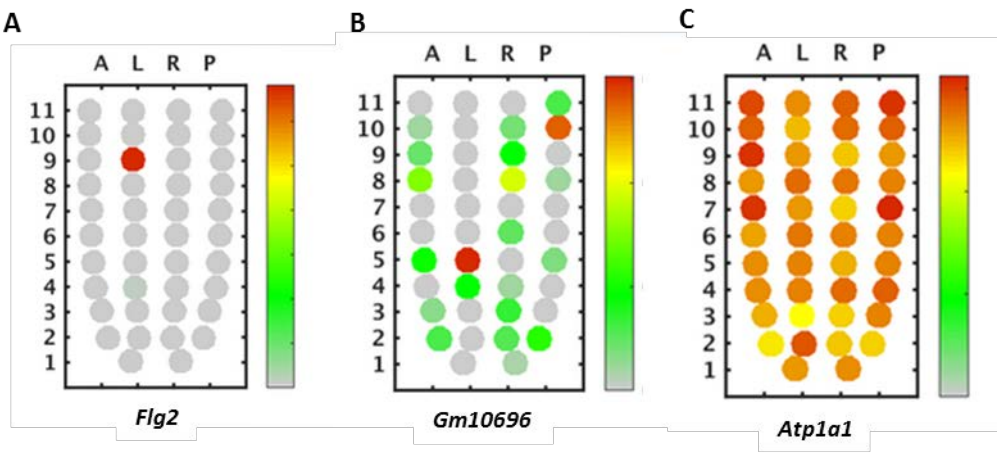

Figure 5.1: Predicted expression patterns of *Flg2*, *Gm10696*, and *Atp1a1* of mid-gastrulation. The corn plots represent the gene expression of the three potential causative mutation-carrying genes at mid-gastrulation in 2D format. (A) *Flg2* was highly expressed only in the proximal left lateral section (red) and very low expression was observed in the distal left lateral section (light green) of the embryo. (B) *Gm10696* was highly expressed only in the distal left lateral section (red) with low expression observed in the ventral left lateral section (light green) of the embryo. (C) The transcription of *Atp1a1* was observed throughout the anterior and posterior epiblast. A, anterior; P, posterior; L, left lateral; and R, right lateral. The colour scale indicates expression, with red indicating high expression and light green indicating low expression. Numbers indicate the 11 alternating sections. This figure was adapted from iTranscriptome.

(light green) displaying low expression. This suggests that Flg2 is not essential during gastrulation. In addition, the predictive online software Protein Variation Effect Analyser (PROVEAN) (Choi *et al.*, 2012, Choi *et al.*, 2015) suggests that this mutation has no effect on the function of the protien. Thus, it was excluded as the gene carrying the causative mutation for the gastrulation phenotype in the 11-23 pedigree.

The SNV identified at position 93,979,673 lies within predicted gene 10696 (*Gm10696,* MGI: 3702972). As shown in Figure 1.5 B, iTranscriptome predicted low expression of *Gm10696* throughout the embryo during gastrulation. One cluster of cells at the distal left lateral epiblast (indicated in red) displayed high expression, but the cells adjacent to it (indicated in light green) displayed very low expression. These observations were not specific to individual cells, making it difficult to interpret the data. Protein Variation Effect Analyser (PROVEAN) (Choi *et al.*, 2012, Choi *et al.*, 2015) suggests that this mutation has an effect on the function of the protien. However, because of the predicted lack of expression, this gene was excluded as a potential causative factor of the abnormal gastrulation phenotype in the 11-23 pedigree.

The SNV identified at position 101,394,617 lies within the ATPase, Na^+/K^+ transporting alpha 1 polypeptide gene (*Atp1a*, MGI: 88105). This protein is one of several isoforms of the α subunit of the Na^+/K^+-ATPase, an integral membrane protein complex that is responsible for establishing and maintaining Na^+ and K^+ ion gradients across the plasma membrane via active transport (Figure 5.2) . These gradients are essential for a variety of crucial cellular functions such as osmoregulation, sodium coupled transport of a variety of molecules, and the electrical excitability of nerve and muscle (Xie *et al.*, 1989, Levenson, 1994, James *et al.*, 1999, Chang *et al.*, 2012, Reinhard *et al.*, 2013).

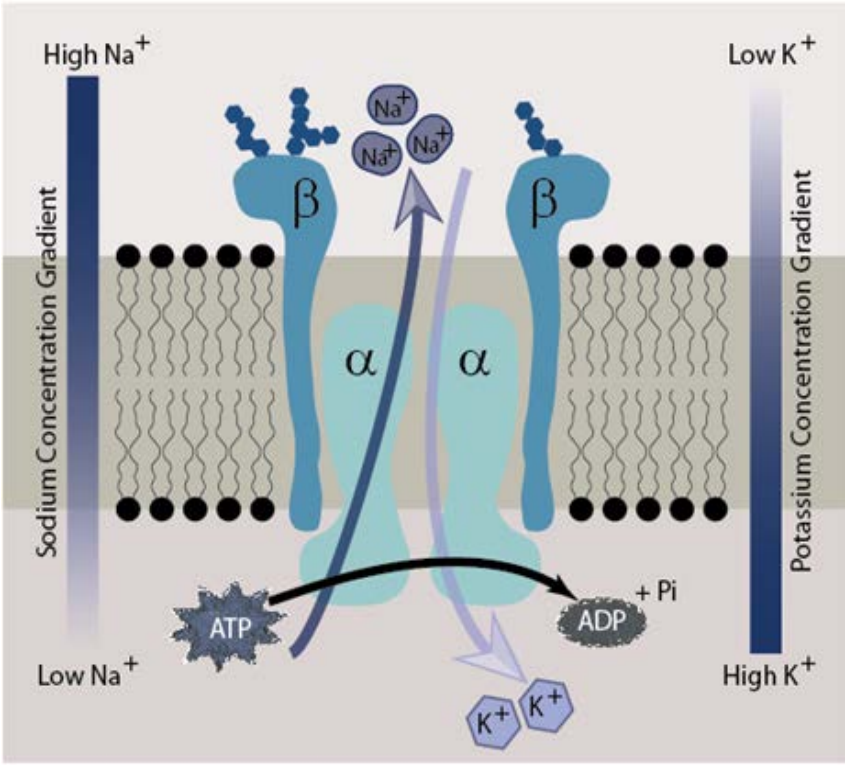

Figure 5.2: The Na^+/K^+ ATPase. This membrane pump uses the energy from ATP hydrolysis to actively pump sodium and potassium ions against their concentration gradients. The binding of three sodium ions to internal binding sites on the cytoplasmic side and phosphorylation of the pump induces a conformational change that releases Na^+ into the extracellular fluid. The subsequent binding of two extracellular potassium ions to the pump within the synapse, followed by dephosphorylation, causes it to return to its original confirmation and release the potassium ions into the cytoplasm. The Na^+/K^+ ATPase is expressed in almost all cell types, and the ionic gradient that it creates is used to carry out numerous important cellular processes, such as the transport of other molecules across membranes. Figure adapted from R. Arkell.

The enzyme is composed of two subunits; the large α subunit is the catalytic subunit, and the smaller β subunit is a glycoprotein that controls the assembly of the α/β heterodimer. Both subunits of the Na^+/K^+-ATPase are encoded by multiple genes. The critical role of Na^+/K^+-ATPase activity in embryogenesis has been previously demonstrated, as embryos that are homozygous for the null allele of *Atp1a1* die at the pre-implantation stage, before gastrulation (Barcroft *et al.*, 2004). *Atp1a1* expression was detected at the pre-implantation stage (3.5 dpc) in wild-type embryos, specifically in the basolateral (but not apical) membranes of the mural and polar trophectoderm, and the membranes of the inner cell mass (Figure 5.3) (MacPhee *et al.*, 2000). In addition, a significant focus has been placed upon exploring the activity of Na^+/K^+-ATPase in pre-implantation embryos. However, the role of Na^+/K^+-ATPase activity or *Atp1a1* during gastrulation is poorly defined (Baltz *et al.*, 1997). Therefore, it will be important to assess the expression and activity of the Na^+/K^+-ATPase α1 subunit in gastrulation stage embryos. It is recommended that whole-mount and sectioned gastrulation stage embryos be examined for Na^+/K^+-ATPase α1 subunit expression, through confocal immunofluorescence microscopy or immunoblot analysis (MacPhee *et al.*, 2000). Unlike *Flg2* and *Gm10696*, *Atp1a1* expression was predicted by iTranscriptome to be very high throughout the embryo at mid-gastrulation (Figure 5.1 C). The identified SNV (a T to A transition in exon 8) results in a missense mutation causing a valine to glutamate substitution at amino acid 259 (V259E). As shown in Figure 5.4, this amino acid lies within one of the cytoplasmic loop domains of the α subunit. The predictive program Protein Variation Effect Analyser (PROVEAN) (Choi *et al.*, 2012, Choi *et al.*, 2015) suggests that this mutation could interfere with ATP1A1 protien function. Therefore, we determined that *Atp1a1* is the likely mutation-carrying gene causing the gastrulation phenotype in the 11-23 pedigree.

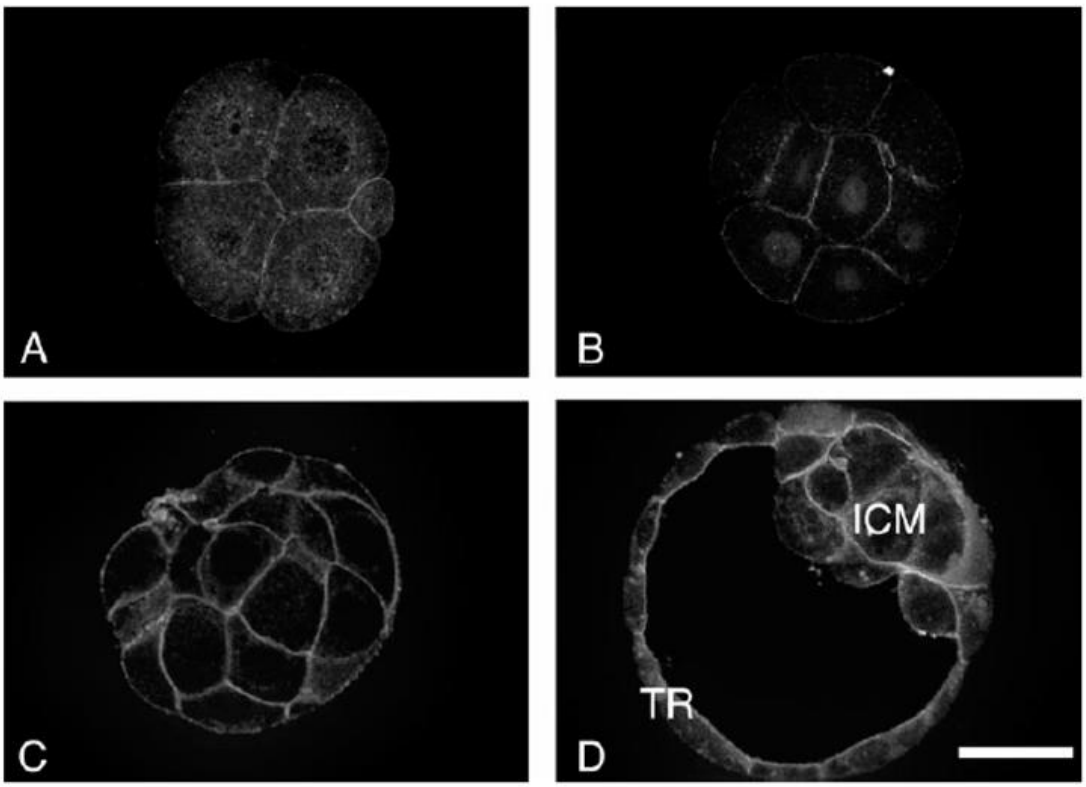

Figure 5.3: Na^+/K^+ ATPase expression during the pre-implantation stage. Localisation of the Na^+/K^+-ATPase a1 subunit in pre-implantation embryos by confocal immunofluorescence microscopy. Stages shown are (A) 4-cell, (B) 8-cell, (C) morula, and (D) blastocyst. The embryos were immunostained with an antibody. Nuclear staining was an artefact of the secondary antibody that was especially prominent in 8-cell embryos. The scale bar indicates 25 µm. Figure adapted from (MacPhee *et al.*, 2000). TR, trophectoderm; ICM, inner cell mass.

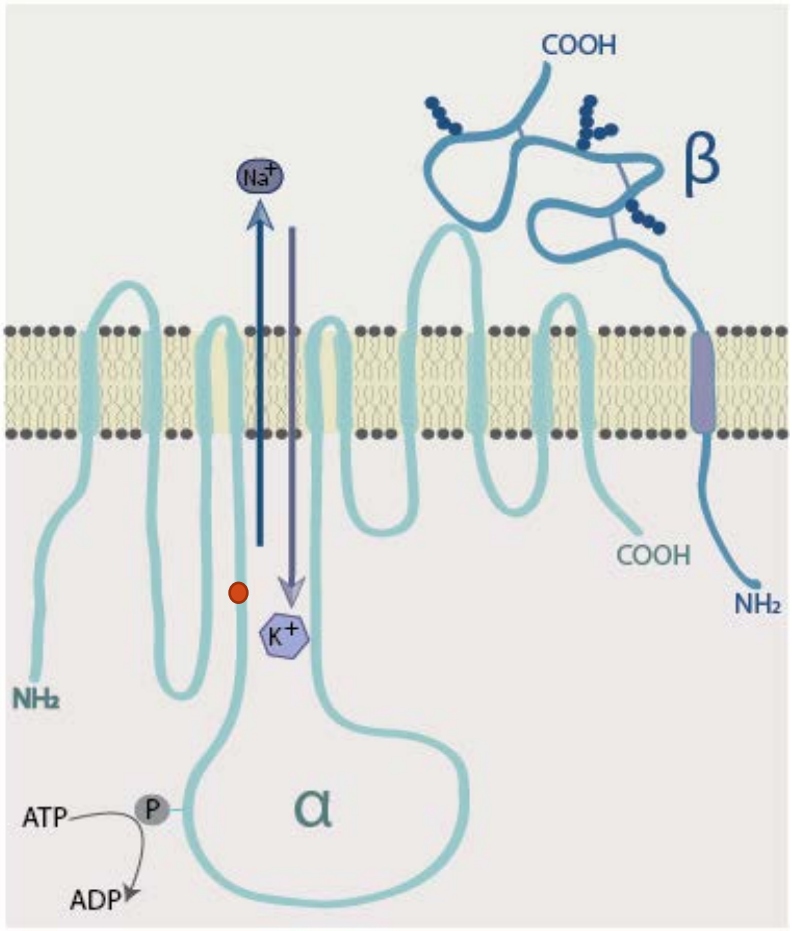

Figure 5.4: The α and β subunits of Na^+/K^+-ATPase. The approximate position of the V259E mutation, which may result in loss-of-function, is indicated by the red circle on one of the α loops.

5.2.2 *Atp1a1*^{V259E} allele validation via sequencing

Based on the known functions and expression patterns of the three candidate genes, the SNV in *Atp1a1*, at position 101,394,617 of MMu3, was considered the most likely causative variant. To determine whether this mutation segregated with the 11-23 phenotype, the relevant portion of *Atp1a1* was sequenced in two additional carrier animals (females 48 and 270) along with two non-carrier animals (females 4 and 14). DNA was extracted from a total of 27 carrier and non-carrier animals, but the data presented here was derived from carrier female #270 and non-carrier female #14. As shown in Figure 5.5, sequencing of female mouse #270, which was mated with a carrier, male #271, produced several dysmorphic embryos, confirming that she was heterozygous for the *Atp1a1*^{V259E} mutation. Additionally, DNA was extracted from the normal and dysmorphic embryos produced by carrier female #270 (Figure 5.5). The sequencing of these embryos, as expected, produced three clearly defined results: wild-type (*Atp1a1*$^{+/+}$), heterozygous (*Atp1a1*$^{V259E/+}$), and homozygous *Atp1a1*$^{V259E/V259E}$. The abnormal embryos were all *Atp1a1*$^{V259E/V259E}$ (Figure 5.5 B), confirming that these gastrulation mutant embryos were carriers of this SNP mutation, and suggesting causation. Conversely, the non-carrier female mouse #14, which was mated with carrier male #8, produced only normal embryos that did not carry the mutation (Figure 5.5). In all cases, the expected sequence was found. The data indicate that the *Atp1a1*^{V259E} mutation causes the abnormal phenotype in the 11-23 pedigree (Figure 5.5).

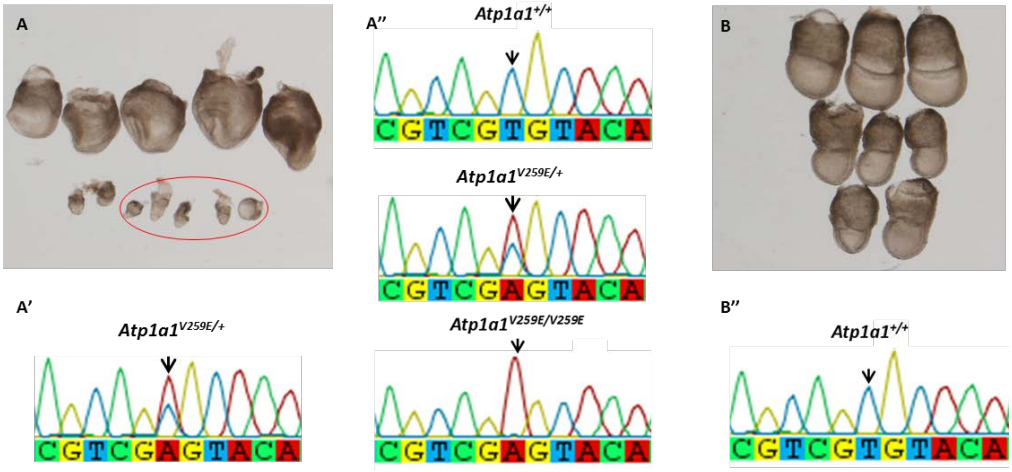

Figure 5.5: The identified *Atp1a1*^{V259E} mutation in the 11-23 pedigree. (A-B) Lateral views of dissected mouse embryos. Anterior view is shown on the left side in all images. (A) Dissected embryos at gastrulation from carrier female #270, including dysmorphic embryos (in red circle). (A') A sequence trace from exon 8 of the *Atp1a1* gene of female #270 displays both an A and a T at nucleotide position 775, confirming her heterozygosity. (A'') Sequence traces from the dissected embryos (in panel A). Normal gastrula embryos were wild-type or heterozygous for the *Atp1a1* gene, and abnormal embryos were homozygous for the *Atp1a1* mutant allele, indicating that abnormal gastrulation was due to the *Atp1a1*^{V259E} mutation. (B) Normal gastrulation stage embryos produced by a non-carrier female #14. (B'') A sequence trace from female #14 displayed a T at position 775 of the *Atp1a1* gene. In all traces, arrows indicate the affected base (T to A).

5.2.3 Genotyping of the *Atp1a1*^{V259E} mutation

An allelic discrimination assay was optimised using the TaqMan® SNP system to identify *Atp1a1*^{V259E}. All gastrulation stage embryos (a total of 240) were tested and genotyped by TaqMan® SNP genotyping. Primers that amplify the region of interest were used together with two different fluorescent probes: one that specifically binds to the wild-type allele (Caruana *et al.*) and another that binds to the mutant allele (VIC; see section 2.1.6 for primer design and section 2.2.4.2 for probes used for allelic discrimination). Changes in fluorescence associated with each probe were measured. Embryos at gastrulation did not contain sufficient genomic DNA, and there was slight maternal contamination; however, samples were still distinguishable. This assay allowed the identification of three distinct populations in the 11-23 pedigree: wild-type (*Atp1a1*$^{+/+}$), heterozygous (*Atp1a1*$^{V259E/+}$), and homozygous (*Atp1a1*$^{V259E/V259E}$) embryos (Figure 5.6 A-C). The FAM probe fluoresced when the wild-type allele was amplified, whereas only slight fluorescence was recorded from the VIC probe. When *Atp1a1*$^{V259E/+}$ embryonic DNA was present, both probes fluoresced at equal levels. When *Atp1a1*$^{V259E/V259}$ homozygous embryonic DNA was present, the VIC probe fluoresced while only slight fluorescence was recorded for the FAM probe. The allelic discrimination plot (Figure 5.6 D) displays the sorting of samples into one of three groups based on the fluorescence produced from each probe. Therefore, the allelic discrimination assay successfully differentiated homozygous and heterozygous carriers embryos of the *Atp1a1*^{V259E} mutation from the wild-type embryos.

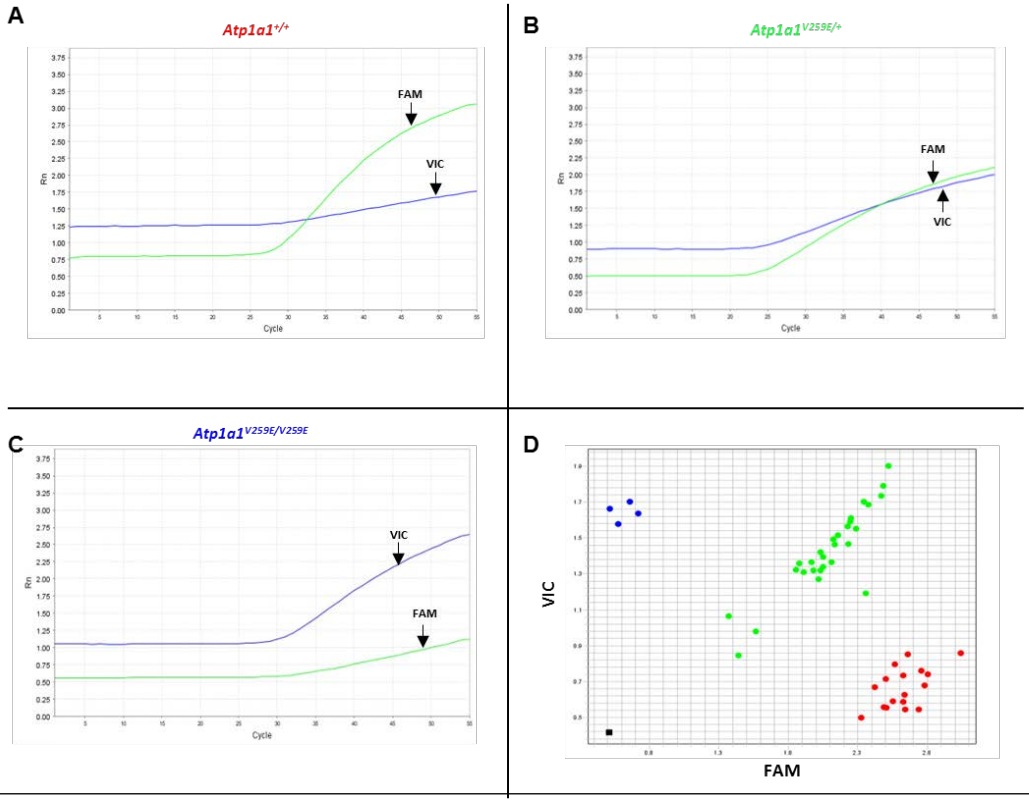

Figure 5.6: Genotyping of *Atp1a1*^{V259E} mutant embryos using allelic discrimination. Amplicons were generated using *Atp1a1*-specific primers. Wild-type and mutant alleles were distinguished using separate probes for each allele: FAM for wild-type (green traces) and VIC for *Atp1a1* mutants (blue traces). (A-C) The amplification plots displayed normalised fluorescence from each probe at each cycle. In wild-type (*Atp1a1*$^{+/+}$) embryos (A) fluorescence was detected from the FAM probe. In heterozygous (*Atp1a1*$^{V259E/+}$) embryos (B), fluorescence was detected from both probes. In homozygous (*Atp1a1*$^{V259E/V259E}$) embryos (C), fluorescence was detected from the VIC probe. (D) The allelic discrimination plot displaying normalised fluorescence from the VIC probe against the FAM probe. Genotyping results using unknown embryos from the 11-23 intercross specific mating produced three distinct populations: wild-type (red), heterozygous (green), and homozygous (blue). The black square at the bottom left of the graph represents a no-template control.

5.2.4 Protein modelling of the *Atp1a1*^{V259E} allele

The α subunit of ATP1A1 catalyses ATP hydrolysis, binds both sodium and potassium ions, and drives these ions across the plasma membrane (Levenson, 1994). Valine, a hydrophobic amino acid, is typically buried inside proteins while glutamate, which is hydrophilic, tends to be on protein surfaces (Figure 5.7). It is more energetically favourable for a hydrophobic residue, such as valine, to be buried and minimally exposed to solvent. Thus, the V to E substitution may alter the hydrophobicity of the mutation site. Therefore, this change may cause nearby residues to be arranged differently in the protein. Additionally, this change may make the protein unstable or mislocalised. Thus, a V to E mutation may perturb the backbone of the ATP1A1 protein, which alters the activity of the charged side chain or changes the folding of the protein. Further experimentation, however, is needed to investigate the effect of this mutation on Na^+/K^+-ATPase activity.

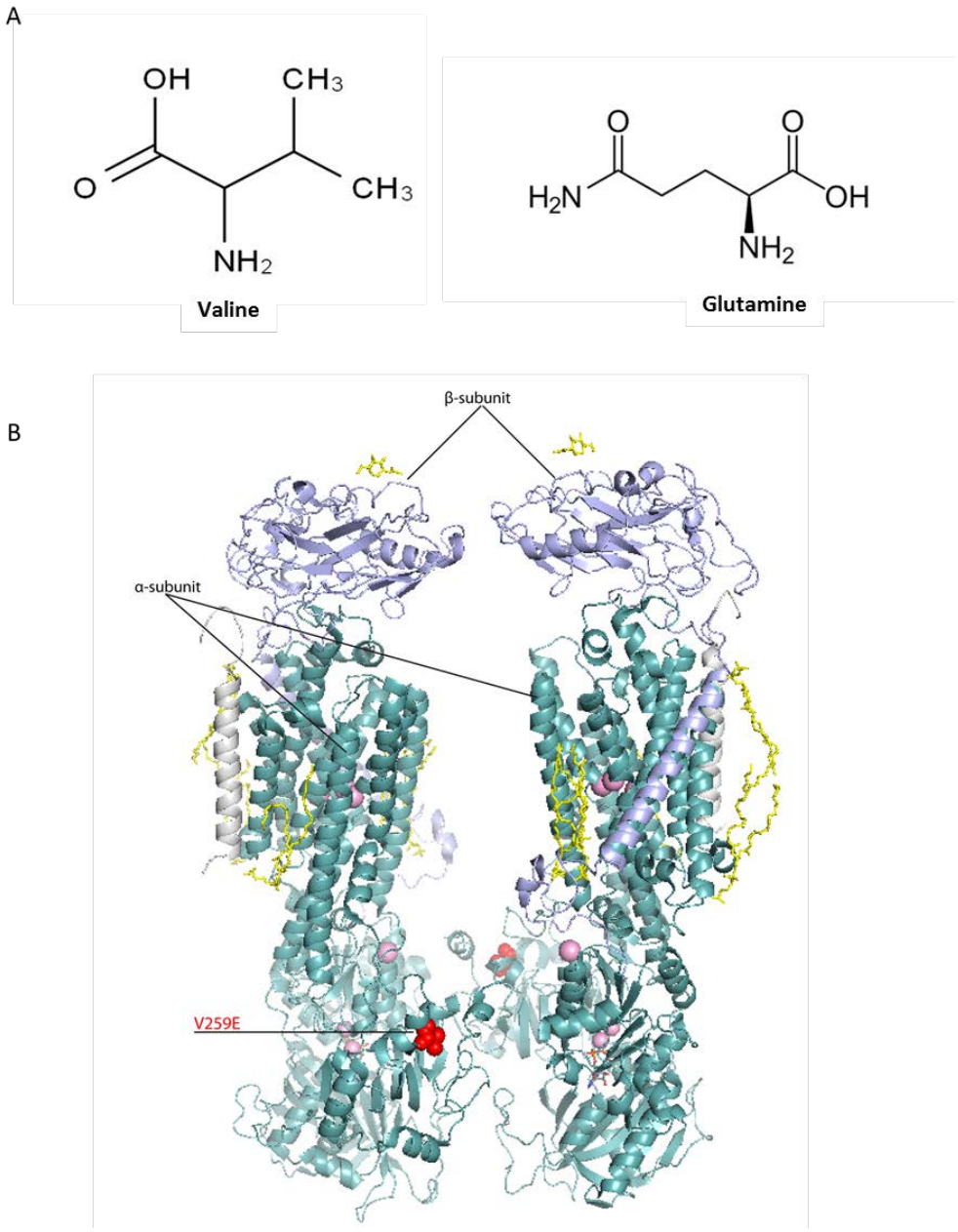

Figure 5.7: Crystal structure of porcine α-, β- and γ subunits of Na^+/K^+-ATPase (PDB ID: 3WGU). (A) The structures of valine and glutamate. (B) The Na^+/K^+-ATPase complex backbone is shown as a ribbon diagram. α subunit, teal; β subunit, lavender; V259 mutation, red; phosphatidylcholine, yellow sticks; Na^+/K^+ ATPase γ subunit transcript variant a, grey; sodium ions, pink spheres; N-acetyl-D-glucosamine, green; ADP, sticks with Corey, Pauling, Koltun colour scheme; magnesium ions, pink spheres. 3D image constructed in PyMOL Molecular Graphics System.

5.3 DISCUSSION

5.3.1 Overview

The advent of next-generation sequencing technologies has vastly altered the prospects for identifying causative point mutations in large genomes. In this chapter, I directly evaluate whether WES provides a useful method for identification of mutations in the gastrulation screen. I find that by conducting WES across the genome of two heterozygous carrier animals and applying filters for common and platform-specific variants, it is possible to readily identify the causative mutation, demonstrating the utility of this approach. The causative mutation lies in the *Atp1a1* gene, which encodes the α subunit of the Na^+/K^+-ATPase complex that energises the transport of ions and solutes across cell membranes. The *Atp1a1* gene has previously been shown to be essential for the peri-implantation stage of murine embryogenesis (Barcroft *et al.,* 2004). The data indicate that WES can be used to circumvent the need for fine recombination mapping of ENU-induced mutations that affect gastrulation. Simultaneously, by identifying all exonic mutations across the MCR, WES reduces the need for complementation studies traditionally used to ensure that the phenotype is not caused by a closely linked, alternative mutation.

5.3.2 Mutation of *Atp1a1* underlies the 11-23 phenotype

Several lines of evidence suggest that the *Atp1a1* mutation is causative. First, WES enabled the detection of all SNVs in the critical interval and revealed three putative variants within the MCR. Of the three altered genes, *Atp1a1* is the only one which is both expressed during gastrulation and predicted to cause a damaging protein alteration. Second, the *Atp1a1* gene has previously been disrupted via gene targeting in mice (James *et al.,* 1999), and embryos lacking ATP1A1 function die sometime between

3.5 and 8.5 dpc (Barcroft *et al.,* 2004), consistent with the 11-23 phenotype. Third, analysis using mutation prediction programs suggests that the non-conservative V295E substitution within the catalytic domain of the Na^+/K^+-ATPase complex is damaging. Finally, following the establishment of a genotyping assay, we found non-segregation of the phenotype and mutation in 843 meiosis samples.

Additional evidence that the *Atp1a1* V295E mutation underlies the phenotype can be obtained via complementation testing between 11-23 and the null *Atp1a1* allele. Complementation is the name given to the phenomenon whereby two strains of an organism, both of which have a homozygous recessive mutation that gives rise to the same phenotype, produce offspring with a wild-type phenotype when crossed. In general, complementation will occur if the two mutations are in separate genes. In contrast, if the two mutations are in the same gene, the homozygous mutant phenotype will generally be recapitulated, and the two strains are said to 'fail to complement' one another or be 'non-complementing'. In traditional positional cloning approaches, mutation scanning does not cover the entire MCR but generally is limited to candidate genes in the interval. In this case, complementation testing can be employed to rule out the possibility that a phenotype is caused by a closely linked, but unidentified mutation. The work in this chapter shows how WES can eliminate the need for complementation testing. By identifying and evaluating the likely pathogenicity of each mutation within the MCR, the possibility of an unrevealed, second-site causative mutation is minimised. In the case of *Atp1a1*, a search for available murine alleles showed that the *Atp1a1* targeted strain (*Atp1a1*^{tm1Ling}) was no longer available but that frozen stocks (embryos) of a gene trap allele (*Atp1a1*$^{Gt(OST31117)Lex}$) could be imported. Given the time and cost of importing, animating, and crossing this strain along with the strong evidence that the

other two exonic variants were unlikely to affect gastrulation, the decision was made not to proceed with an *Atp1a1* complementation test.

A second method that would provide additional support for the mutagenic potential of the V295E allele is the direct measurement of the Na^+/K^+-ATPase activity of this protein variant and the comparison to that of the wild-type protein. Na^+/K^+-ATPase activity can be measured by a variety of colourimetric assays (Sarkar, 2002) or by rubidium-86 influx assays, where K^+ influx is deduced by determining the rate of release of P_i from ATP (Xie *et al.*, 1989, Herrera *et al.,*1990, Liang *et al.*, 2006). These assays can be conducted in cell-free systems, an approach which would require the expression and purification of the V295E variant. Alternatively, they can be adapted for cell-based systems, which would require the generation of a cell line that either transiently or stably expresses wild-type or variant ATP1A1. *In vitro* systems for the measurement of Na^+/K^+-ATPase activity also exist; for example, activity can be measured in isolated intact skeletal muscles or in digests of these muscles (Clausen, 2013). This approach cannot be used with the homozygous embryos due to early lethality, but it may be possible to discern a difference between the muscles of heterozygous animals and those of wild-type animals. Given that the aim of the work in this chapter was to determine whether WES was suitable for identifying putative mutations in the MCR, these experiments were not deemed necessary to meet the project objectives. If, however, the 11-23 strain is used in the future to investigate the cellular and molecular roles of the ATP1A1 protein, these experiments should be performed.

One question raised by the work in this chapter is whether the phenotype of the 11-23 strain is different from that of the targeted null allele (*Atp1a1*^{tm1Ling}) (James *et al.,* 1999, Barcroft *et al.,* 2004). Previous characterisation of this allele spanned, but did not directly investigate, the immediate post-implantation period. Homozygous embryos

were recovered at the expected proportion at 3.5 dpc (Barcroft *et al.,* 2004), but none were recovered at 8.5 dpc (James *et al.,* 1999), illustrating that they died between these time points. Barcroft *et al.,* (2004) further investigated the viability of *Atp1a1*-null preimplantation embryos via culture of isolated eight-cell embryos or blastocysts. In each case, embryo culture adversely affected homozygous-null embryos (but not their wild-type or heterozygous littermates) as evidenced by premature cell dissociation and (in the case of the blastocysts) a failure to attach to cultureware and form trophoblast outgrowths. The researchers therefore concluded that *Atp1a1*-null embryos would not implant. The work presented in this chapter suggests that a severe mutation of *Atp1a1* may be compatible with implantation, but not with gastrulation. There are however two caveats to this inference: one is that the V295E allele may be hypomorphic, and the second is that a genetic background difference could lead to phenotype differences. There is some support for the notion that genetic background may influence the *Atp1a1* mutant phenotype since over the course of the experiments reported in this chapter, the phenotype of the homozygous embryos became more severe. After the 11-23 strain was backcrossed into the C3H strain, embryos recovered at 7.5 dpc showed little evidence of epiblast formation and were in the process of resorbing. In contrast, phenotype onset was delayed by outcrossing the C3H strain with C57BL/6 animals, suggesting that the C57BL/6 background moderates the phenotype.

5.3.3 WES is useful for the gastrulation screen

The WES analysis presented in this chapter indicates that it should be the method of choice for the gastrulation screens of the remaining pedigrees. Not only were a small number of exonic variants documented in the MCR by WES, but also resources such as the iTransriptome (Chen *et al.,* 2017) now enable the rapid assessment of the expression pattern of candidate genes. It seems likely that even when the MCR remains relatively

large, there will be only one or two realistic putative causative SNVs, and (as shown here with *Atp1a1*) these can be rapidly discriminated by Sanger sequencing across the variants in stored DNAs from the multi-generation pedigrees generated during the mapping process. Additionally, this is now the method of choice for many other ENU mutagenesis screens (Andrews *et al.,* 2012, Caruana *et al.,* 2013, Potter *et al.,* 2016). Many of the remaining pedigrees did not involve the C57BL/6 reference strain, but this should not confound their analysis because of mouse resequencing efforts (Keane *et al.*, 2011) and because the drop in the cost of genome sequencing means that (if required) additional carrier animals can be sequenced to validate putative de novo variants.

5.3.4 The gastrulation screen

One feature of the gastrulation screen is that even though only a small number of pedigrees were isolated (eight in total), these appeared to fall into three phenotype classes (Figure 4.1): Failure to thrive, Failure to gastrulate, and Patterning defects. Each class appears to correspond to failure at a particular time point (pre-gastrulation, onset of gastrulation, and during gastrulation, respectively). It is tempting to speculate that each class may be enriched for a particular type of molecule. For example, the mutation of housekeeping genes may be expected to lead to a failure to thrive, whereas the mutation of genes involved in signalling pathways that regulate the timing of gastrulation may lead to a failure to gastrulate. The finding that the *Atp1a1* gene is mutated in 11-23 (classified as a Failure to thrive) appears, on the surface, to support this hypothesis since *Atp1a1* is thought to be ubiquitously expressed and critical for the maintenance of basic cellular processes. Recently, however, the intracellular domains of the Na^+/K^+-ATPase have been shown to bind to an array of proteins involved in cell signalling pathways such as the PI_3K, Src, and MAPK pathways (Reinhard *et al.*, 2013). The V259E missense mutation occurs within the cytoplasmic domain in a region

implicated in these protein-protein interactions, and it is possible that the ENU-induced allele interferes with a more specific docking function of ATP1A1. Further understanding of the phenotype classes described here awaits the identification of the causative mutations in the other pedigrees.

Another noteworthy feature of the gastrulation screen is the high rate of recovery of pedigrees of interest. The eight lines reported in this thesis were recovered from a screen of forty-three pedigrees. This represents a recovery rate of 18%, which is very similar to the ~20% recovery rate reported in other genome-wide screens for recessive mutations that alter other stages of embryogenesis (see Caruana *et al.,* 2013 for discussion). This demonstrates that gastrulation, as with other stages of embryogenesis, represents a rich source of phenotypes that could reveal genes and protein domains critical to this process.

Chapter 6: Future Perspectives

The overall aim of this project was to determine if a genome-wide screen for genes required for gastrulation is possible and would likely contribute to our understanding of mammalian gastrulation. In chapter 3, I analysed the Katun mouse strain (a null allele of *Zic3*) and the batface mouse strain (a gain of function allele of *Ctnnb1*). The phenotypes of both these strains are first easily observed at 9.5 dpc, yet the likely cause of these phenotypes occurs during gastrulation. Moreover, other alleles of these genes have long been available (in the case of *Zic3*, the first allele called Bent tail (*Bn*) was published in 1952 (Garber, 1952), yet careful observation of the associated phenotypes has yielded novel information and hypotheses on the role of these genes during gastrulation. A search for additional ENU-induced mouse mutants with phenotypes that inform our understanding of gastrulation appears warranted.

The advent of CRISPR targeted mutagenesis has led to questions regarding the usefulness of ENU screens in the mouse. For many biological processes, there has been a shift in focus towards using next-generation sequencing technologies to identify variants in humans with certain diseases or disorders. The variants are then filtered with respect to functional annotations and likely disease involvement before CRISPR mutagenesis is used to introduce an analogous mutation in mice. The resulting animals are then analysed to determine whether the mutation appears pathogenic. Although useful for a wide range of diseases and disorders, the utility of this approach for the study of gastrulation is less clear. In humans, mutations that severely compromise gastrulation are lethal at ~ 7 weeks of gestation and may not be detected until routine scans that generally are scheduled for ~12 weeks of gestation – by which time the embryo has been resorbed. In contrast, the work presented in this thesis demonstrates

that a genome-wide ENU mutagenesis screen for phenotypes that affect gastrulation would be a sensible addition to efforts to understand the molecular basis of this process in mammals.

The progress in sequencing technologies and corresponding decrease in associated costs suggest that even the approach used in this thesis (low-resolution recombination mapping followed by WES sequencing of two carrier animals) has now been superseded. To carry out a further genome-wide screen for ENU-induced mutations that inform our study of mammalian gastrulation, the following procedure is recommended.

- Mutagenise males from either the BALB/c or FVB strain (since both have been well characterised for their response to ENU, are highly fecund, and have been re-sequenced).
- Cross mutagenised males with females of the same strain (it is no longer necessary to use a second mapping strain in the experiment) and designate male progeny as Founder animals for test mating.
- Cross a Founder male with a wild-type female (of the same inbred strain as the mutagenised male) until twenty G2 females have been produced.
- Screen four G2 females by backcrossing with the founder male, keeping tissues for future DNA extraction from each female.
- If two or more females produce embryos with the same phenotype, then the pedigree is ‘of interest’.
- Screen additional G2 females by backcrossing with the founder male until a total of four putative carrier females have been identified.
- Extract DNA from an ear biopsy of the founder male (estimated 70 exonic variants) and from the reserved tissue of each G2 carrier female (estimated 35 exonic variants) and perform WES on each of the five DNAs.

- Filter variants against known strain- and platform-specific variants.
- Intersect the variants across each G2 female and by comparison to the founder male (estimated 4 exonic variants common to all carriers).
- Sanger sequence putative variants in the remaining G2 females and formulate carrier hypotheses.
- Test the carrier hypotheses by backcrossing with the founder and examining the litters.
- Sanger sequence putative variants in affected (i.e. presumed homozygous) embryos (estimated 1 variant common between homozygous embryos and carrier females).
- Genotype G2 males to identify carriers for colony maintenance.

Supplementary Information

Appendix A

Table 1: The selected MIT markers at most the proximal, distal and the central regions of all autosomes. TD: touchdown PCR programme, B: buffer, D: DNA segment, Mit: Massachusetts Institute of Technology.

Proximal Marker					
MMu	Locus	Position (Mb)	Oligo Orientation	Oligonucleotide Sequence (5` - 3`)	Optimised conditions
1	D1Mit20	30.9	Forward	AGGTGAATGTACAAGTGTTTGGG	TD60, BI
			Reverse	TGAAATTCAGAAATCTTTCTGGG	
2	D2Mit1	3.8	Forward	CTTTTTCGTATGTGGTGGGG	TD65, BIV
			Reverse	AACATTGGGCCTCTATGCAC	
3	D3Mit221	7.92	Forward	ATATAAGGACATAAACAGGCATTTCT	TD65, BIV
			Reverse	AAAATGTAGATCCCTTCATACATGC	
4	D4Mit18	13.9	Forward	AATTAGCCCGGAGCTTGATT	TD65, BIV
			Reverse	GCTTCCATACATTTGCTTTTCC	
5	D5Mit346	4.8	Forward	TCAAACTCCTCTAATATGGAAGTGC	TD60, BIV
			Reverse	CTGTCTCATTAATCCATGGATCC	
6	D6Mit235	23.1	Forward	ATTTCTGTCGTGCCTGAGCT	TD65, BIV
			Reverse	TCATGTGTGCTTGCTTGTGA	
7	D7Mit76	19.6	Forward	CATGAGCACGTGGAGAAAGA	TD65, BIV
			Reverse	CGTGGAAACCTGATAAACTGA	
8	D8Mit58	7.1	Forward	TGTCTGCATGTAGAAGAGTGCA	TD65, BIV
			Reverse	CTGTGAACTGTGGCATGATAA	
9	D9Mit297	34	Forward	TGCCTCTCTCATCTCTTTAAACA	TD65, BIV
			Reverse	ATTTGTCTCTCTGTCTTTTGCTCA	
10	D10Mit80	11.4	Forward	CAAAAAAAACCCTGATTCTACCA	TD60, BI
			Reverse	GTGTGCATATGGCAGTAACTTTG	
11	D11Mit74	5.2	Forward	AAAACCTGAGTTCGACCCCT	TD65, BI
			Reverse	ATAAAGCCTCATCTACATGGGC	
12	D12Mit240	11.3	Forward	CTAGCAAAAGTAAAGGCTTTCAGG	TD65, BIV
			Reverse	GCAAGGAGTTAATGCTTTCCC	
13	D13Mit217	29.2	Forward	ACCTGCAAACTTGAGCAGAA	TD65, BIV
			Reverse	ACTAACTCATCTCTTGCTGCCC	
14	D14Mit110	9	Forward	GTAGGAGAGAACAACTGTCTTCTGC	TD65, BIV
			Reverse	CATGAATAAGAACGAAAAGGGC	
15	D15Mit150	14.1	Forward	TTGTGGGATGAAGTGTGCAT	TD60, BIV
			Reverse	TCAAATTTACAGGAAAAGAACAAGC	
16	D16Mit34	15.9	Forward	CCCAGTAGGAAGGAGACCCT	TD65, BI
			Reverse	TGGGAATATACATCCAAAAGCA	
17	D17Mit273	9.1	Forward	CAACACAAGGTTGTTTGATAAAGC	TD65, BIV
			Reverse	ATGGTTAGCCCAAGGGATTC	
18	D18Mit229	12.3	Forward	ATGTTGTGAGACCCTGTCCTG	TD65, BIV
			Reverse	TTCCTCATCCCCATTACTTCC	
19	D19Mit113	19.4	Forward	GGCTGGTTAATTCTTTGTGAGC	TD65, BIV
			Reverse	GAACTAGATTGTCAATATGAATGTGTG	

Central Marker					
MMu	Locus	Position (Mb)	Oligo Orientation	Oligonucleotide Sequence (5` - 3`)	Optimised conditions
1	D1Mit46	75.6	Forward	AGTCAGTCAGGGCTACATGATG	TD65, BIV
			Reverse	CACGGGTGCTCTATTTGGAA	
2	D2Mit15	91.9	Forward	ATGCCTTAGAAGAATTTGTTCCC	TD65, BIV
			Reverse	CTTGAAAAACACATCAAAATCTGC	
3	D3Mit258	142.1	Forward	ACACAGGAATATGGTGTGAGTCC	TD65, BI
			Reverse	AACAGTAGTAGAGGCAAGTAGTGTTTG	

Distal Marker					
MMu	**Locus**	**Position (Mb)**	**Oligo Orientation**	**Oligonucleotide Sequence (5` - 3`)**	**Optimised conditions**
1	D1Mit362	189.7	Forward	TGTGTGACTGCTTGGAAGATG	TD65, BIV
			Reverse	CTGAGTCCCTAAAGTTGTCCTTG	
2	D2Mit148	178.5	Forward	GTTCTCTGATCTACGGGCATG	TD65, BIV
			Reverse	TTCACTTCTACAAGTTCTACAAGTTCC	
3	D3Mit257	142.1	Forward	CCTAGCGCAGGAATAGTTAACC	TD65, BI
			Reverse	ACAAACAGAACAAAACAAAAAGTCC	
4	D4Mit189	145.7	Forward	CACTACACGGGGCTGAGATT	TD65, BIV
			Reverse	AAGATCCAACACCCTTTTTGG	
5	D5Mit139	125	Forward	ACCACCATGGAGAATGCG	TD60, BIV
			Reverse	GAAGTGGTAGGTGCAGTGCA	
6	D6Mit14	146	Forward	ATGCAGAAACATGAGTGGGG	TD65, BIV
			Reverse	CACAAGGCCTGATGACCTCT	
7	D7Mit238	107.3	Forward	GCATCTGCTTTTCTGCCTCT	TD65, BIV
			Reverse	AGGCACCTGACATTGACCTC	
8	D8Mit91	127	Forward	GCCTAACTTAGGGGAGGGAA	TD65, BIV
			Reverse	TATTGGGAATGAGGAGACATACA	
9	D9Mit214	115.4	Forward	AGCACAGGAAAAGGACGCTA	TD65, BIV
			Reverse	AACCTGTCTCTGTAAAACTATCTCCA	
10	D10Mit70	103.5	Forward	ACCTTTCTTCCTACCTACCTACCT	TD60, BI
			Reverse	TGGCACTTAGAAACTGATAAATGG	
11	D11Mit41	89.9	Forward	CGACTGAGCAAGTTGTATTTCTG	TD65, BI
			Reverse	ATAAAGCCTCATCTACATGGGC	
12	D12Mit18	114.4	Forward	CATGCATGTGAAGCTTAGAAGC	TD65, BI
			Reverse	TGGCTCTGTCTTCTGGTGC	
13	D13Mit291	111	Forward	AAAATTATGAATGCAAGTGCACA	TD65, BIV
			Reverse	GGGTATCTGATACCAGTGCCC	
14	D14Mit266	112.1	Forward	ATGCACAGGATTGATCTGCA	TD65, BIV
			Reverse	AGCATGACCTAAATAATGAGACCC	
15	D15Mit245	94.1	Forward	ACCAATACACTTCATGCAAACG	TD60, BIV
			Reverse	CAGTGACCGTAGGTTCATTAACC	
16	D16Mit152	86.3	Forward	AGAGACCTCTGGGGTGGG	TD65, BI
			Reverse	TTCAAGATAGACTATTCTGGAAAAAGC	
17	D17Mit221	90.5	Forward	AACCAGATCATTAACAGTAATAAAGCA	TD65, BIV
			Reverse	ATGGTTAGCCCAAGGGATTC	
18	D18Mit49	76.2	Forward	TTACTCACTTTTCCACTTGCTAGG	TD65, BIV
			Reverse	CTGCACACACCACTTGCC	
19	D19Mit10	47.2	Forward	GCCTTTAAGCCAGTCAAGACA	TD65, BIV
			Reverse	CCAGTCTGGACTTGTGAATGA	

Table2: The additional MIT markers used to narrow down the regions of interest.

Additional Marker					
MMu	**Locus**	**Position (Mb)**	**Oligo Orientation**	**Oligonucleotide Sequence (5` - 3`)**	**Optimised conditions**
4	D4Mit26	84	Forward	CGGTAGTTTTTCAAAGATCGTG	TD65, BIV
			Reverse	CTGATGGTAGCAGAATCAGGC	
7	D7Mit247	39	Forward	TGGAGGAGACATATCTTTGCG	TD65, IMMO
			Reverse	TCTTTTGACTTGATTTTGGCG	
	D7Mit69	56.3	Forward	CACAATGAAGGCTGAAAGCA	TD65, IMMO
			Reverse	CCCACCAGAGATCACCAAGT	
	D7Mit276	69.4	Forward	AATGCCCAGTGTAGAAGAAACC	TD65, IMMO
			Reverse	CTGGGAGGAATGTTCTCCAA	
9	D9Mit4	52	Forward	GACAGCCCATCACAGCTACA	TD65, BIV
			Reverse	TGCTGAGCAAGCTATGAGGA	
	D9Mit12	99.8	Forward	TGGTCCTGGTAAAACTGCCT	TD65, BIV
			Reverse	ATTCAAGGGGCAGTACACAT	
	D9Mit314	57.4	Forward	TCTCTCTAGTGGGATTATCAACACA	TD65, BIV
			Reverse	TCCTCCCTCAACCATCTCAG	
	D9Mit347	115	Forward	CTGTCCATCTATCATCTATCTGTCTG	TD65, BIV
			Reverse	CCTCCACATGTGCACTGCT	
	D9Mit1001	40	Forward	AGTCAGTCAGGGCTACATGATG	TD65, BIV
			Reverse	ATATCAGGCATGCATTATGATCC	
10	D10Mit298	8.5	Forward	AAACAACCAGGCTCCCAAG	TD65, BIV
			Reverse	TCCTGTTCACTGACTGTCTTCC	
12	D12Mit136	24.3	Forward	TTGCTACATGTACACTGATCTCCA	TD65, BIV
			Reverse	TTTAATTTTGAGTGGGTTTGGC	
16	D16Mit30	53.9	Forward	TCACTGCAGGGAGGTTCAG	TD65, BIV
			Reverse	GTGCACATACATACCACAGCG	
	D16Mit136	40.6	Forward	TTGAGAAGTTTGCCCTATAATGG	TD65, BIV
			Reverse	AGATAATTCCCGTGAGAATAAAACC	
	D16Mit173	64.6	Forward	TTCCATTCCATCTCTGATTTCA	TD65, BIV
			Reverse	TCCACCCCCATCACACTC	
17	D17Mit101	29	Forward	TTTCTCTCACAAATAGGGAGTGG	TD65, BIV
			Reverse	GTCCAGTTCCATGGGATCC	
	D17Mit175	32	Forward	TTGGAAAAGGTTGAGAGTAGATCA	TD65, BIV
			Reverse	TGGAAATCGGAGCCTCTG	
	D17Mit9	51.3	Forward	CCCCACCAACTGTCCTCTAA	TD65, BIV
			Reverse	TCAGCCCTTAAAAATTACTCTTGG	
18	D18Mit164	8.5	Forward	ACATGTGTCATAGCACATTCACA	TD65, IMMO
			Reverse	GGAAGAACAGTGGGAATTTATCC	
18	D18Mit119	28.9	Forward	GAGTGTATAGCGGACTTTTGGG	TD65, BIV
			Reverse	AGATGCTTGTGAAACATACATATGTG	

Reference

Ahmed, J. N., Ali, R. G., Warr, N., Wilson, H. M., Bellchambers, H. M., Barratt, K. S., Thompson, A. J. & Arkell, R. M. (2013). A murine Zic3 transcript with a premature termination codon evades nonsense-mediated decay during axis formation. *Disease models & mechanisms,* **6**(3): 755-767.

Ali, R. G., Bellchambers, H. M. & Arkell, R. M. (2012). Zinc fingers of the cerebellum (Zic): transcription factors and co-factors. *The international journal of biochemistry & cell biology,* **44**(11): 2065-2068.

Ahmed, J. N., Diamand, K. E., Arkell, R. M. (2016). Systematized reporter assays reveal ZIC protein regulatory abilities are subclass-specific and dependent upon transcription factor binding site context. Manuscript submitted for publication to *Scientific Report.*

alpha1-subunit gene (Atp1a1) does not prevent cavitation of the preimplantation mouse embryo. *Mechanisms of Development,* **121**(5): 417-426.

Andrews, T. D., Whittle, B., Field, M. A., Balakishnan, B., Zhang, Y., Shao, Y., Cho, V., Kirk, M., Singh, M., Xia, Y., Hager, J., Winslade, S., Sjollema, G., Beutler, B., Enders, A. & Goodnow, C. C. (2012). Massively parallel sequencing of the mouse exome to accurately identify rare, induced mutations: an immediate source for thousands of new mouse models. *Open Biology,* **2**.

Arce, L., Yokoyama, N. N. & Waterman, M. L. (2006). Diversity of LEF/TCF action in development and disease. *Oncogene,* **25**(57): 7492-7504.

Arkell, R. M. & Tam, P. P. (2012). Initiating head development in mouse embryos: integrating signalling and transcriptional activity. *Open Biology,* **2**(3): 120030.

Arkell, R. M., Fossat, N. & Tam, P. P. (2013). Wnt signalling in mouse gastrulation and anterior development: new players in the pathway and signal output. *Current opinion in genetics & development,* **23**(4): 454-460.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J. K. & Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. *Proceedings of the National Academy of Sciences of the United States of America,* **95**(9): 5082-5087.

Arnold, C. N., Xia, Y., Lin, P., Ross, C., Schwander, M., Smart, N. G., Muller, U. & Beutler, B. (2011). Rapid Identification of a Disease Allele in Mouse Through Whole Genome Sequencing and Bulk Segregation Analysis. *Genetics,* **187**(3): 633-641.

Aruga, J. (2004). The role of Zic genes in neural development. *Molecular and Cellular Neuroscience,* **26**(2): 205-221.

Aruga, J., Kamiya, A., Takahashi, H., Fujimi, T. J., Shimizu, Y., Ohkawa, K., Yazawa, S., Umesono, Y., Noguchi, H., ShirniZu, T., Saitou, N., Mikoshiba, K., Sakaki, Y., Agata, K. & Toyoda, A. (2006). A wide-range phylogenetic analysis of Zic proteins: Implications for correlations between protein structure conservation and body plan complexity. *Genomics,* **87**(6): 783-792.

Aruga, J., Yokota, N., Hashimoto, M., Furuichi, T., Fukuda, M. & Mikoshiba, K. (1994). A Novel Zinc-Finger Protein, Zic, Is Involved in Neurogenesis, Especially in the Cell Lineage of Cerebellar Granule Cells. *Journal of Neurochemistry,* **63**(5): 1880-1890.

Aruga, J., Yozu, A., Hayashizaki, Y., Okazaki, Y., Chapman, V. M. & Mikoshiba, K. (1996). Identification and characterization of Zic4, a new member of the mouse Zic gene family. *Gene,* **172**(2): 291-294.

Ashe, A., Butterfield, N. C., Town, L., Courtney, A. D., Cooper, A. N., Ferguson, C., Barry, R., Olsson, F., Liem, K. F., Jr., Parton, R. G., Wainwright, B. J., Anderson, K. V., Whitelaw, E. & Wicking, C. (2012). Mutations in mouse Ift144 model the craniofacial, limb and rib defects in skeletal ciliopathies. *Human molecular genetics,* **21**(8): 1808-1823.

Austin, C. P., Battey, J. F., Bradley, A., Bucan, M., Capecchi, M., Collins, F. S., Dove, W. F., Duyk, G., Dymecki, S., Eppig, J. T., Grieder, F. B., Heintz, N., Hicks, G., Insel, T. R., Joyner, A.,

Koller, B. H., Lloyd, K. C. K., Magnuson, T., Moore, M. W., Nagy, A., Pollock, J. D., Roses, A. D., Sands, A. T., Seed, B., Skarnes, W. C., Snoddy, J., Soriano, P., Stewart, D. J., Stewart, F., Stillman, B., Varmus, H., Varticovski, L., Verma, I. M., Vogt, T. F., von Melchner, H., Witkowski, J., Woychik, R. P., Wurst, W., Yancopoulos, G. D., Young, S. G., Zambrowicz, B. & Proje, C. K. M. (2004). The knockout mouse project. *Nature Genetics,* **36**(9): 921-924.

Avner, P., Amar, L., Dandolo, L. & Guenet, J. L. (1988). Genetic-analysis of the mouse using Iinterspecific crosses. *Trends in Genetics,* **4**(1): 18-23.

Baltz, J. M., Smith, S. S., Biggers, J. D. & Lechene, C. (1997). Intracellular ion concentrations and their maintenance by Na+/K+-ATPase in preimplantation mouse embryos. *Zygote,* **5**(1): 1-9.

Barbaric, I. & Dear, T. N. (2007). Optimizing screening and mating strategies for phenotype-driven recessive N-ethyl-N-nitrosourea screens in mice. *Journal of the American Association for Laboratory Animal Science : JAALAS,* **46**(6): 44-49.

Barratt, K. S., Glanville-Jones, H. C. & Arkell, R. M. (2014). The Zic2 gene directs the formation and function of node cilia to control cardiac situs. *Genesis,* **52**(6): 626-635.

Beckers, A., Alten, L., Viebahn, C., Andre, P. & Gossler, A. (2007). The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left-right patterning. *Proceedings of the National Academy of Sciences of the United States of America,* **104**(40): 15765-15770.

Bedard, J. E., Purnell, J. D. & Ware, S. M. (2007). Nuclear import and export signals are essential for proper cellular trafficking and function of ZIC3. *Hum Mol Genet,* **16**(2): 187-198.

Bedell, M. A., Jenkins, N. A. & Copeland, N. G. (1997). Mouse models of human disease. Part I: techniques and resources for genetic analysis in mice. *Genes & Development,* **11**(1): 1-10.

Benedyk, M. J., Mullen, J. R. & DiNardo, S. (1994). odd-paired: a zinc finger pair-rule protein required for the timely activation of engrailed and wingless in Drosophila embryos. *Genes & Development,* **8**(1): 105-117.

Bienz, M. & Clevers, H. (2003). Armadillo/beta-catenin signals in the nucleus - proof beyond a reasonable doubt? *Nature Cell Biology,* **5**(3): 179-182.

Blake, J. A., Bult, C. J., Kadin, J. A., Richardson, J. E., Eppig, J. T. & Grp, M. G. D. (2011). The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. *Nucleic Acids Research,* **39**: D842-D848.

Blake, J. A., Richardson, J. E., Bult, C. J., Kadin, J. A. & Eppig, J. T. (2003). MGD: the mouse genome database. *Nucleic Acids Research,* **31**(1): 193-195.

Blewitt, M. E., Vickaryous, N. K., Hemley, S. J., Ashe, A., Bruxner, T. J., Preis, J. I., Arkell, R. & Whitelaw, E. (2005). An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(21): 7629-7634.

Bogani, D., Warr, N., Elms, P., Davies, J., Tymowska-Lalanne, Z., Goldsworthy, M., Cox, R. D., Keays, D. A., Flint, J., Wilson, V., Nolan, P. & Arkell, R. (2004). New semidominant mutations that affect mouse development. *Genesis,* **40**(2): 109-117.

Bogani, D., Willoughby, C., Davies, J., Kaur, K., Mirza, G., Paudyal, A., Haines, H., McKeone, R., Cadman, M., Pieles, G., Schneider, J. E., Bhattacharya, S., Hardy, A., Nolan, P. M., Tripodis, N., Depew, M. J., Chandrasekara, R., Duncan, G., Sharpe, P. T., Greenfield, A., Denny, P., Brown, S. D. M., Ragoussis, J. & Arkell, R. M. (2005). Dissecting the genetic complexity of human 6p deletion syndromes by using a region-specific, phenotype-driven mouse screen. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(35): 12477-12482.

Bouabe, H. & Okkenhaug, K. (2013). A protocol for construction of gene targeting vectors and generation of homologous recombinant embryonic stem cells. *Methods in molecular biology,* **1064**: 337-354.

Bovolenta, P., Esteve, P., Ruiz, J. M., Cisneros, E. & Lopez-Rios, J. (2008). Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. *Journal of Cell Science,* **121**(6): 737-746.

Bradley, A. & Liu, P. (1996). Target practice in transgenics. *Nature Genetics,* **14**(2): 121-123.

Brennan, J., Lu, C. C., Norris, D. P., Rodriguez, T. A., Beddington, R. S. P. & Robertson, E. J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature,* **411**(6840): 965-969.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics,* **77**(1): 71-94.

Brown, L., Paraso, M., Arkell, R. & Brown, S. (2005). In vitro analysis of partial loss-of-function ZIC2 mutations in holoprosencephaly: alanine tract expansion modulates DNA binding and transactivation. *Human molecular genetics,* **14**(3): 411-420.

Brown, S. D. M. & Balling, R. (2001). Systematic approaches to mouse mutagenesis. *Current opinion in genetics & development,* **11**(3): 268-273.

Brown, S. J. & McLean, W. H. I. (2012). One Remarkable Molecule: Filaggrin. *Journal of Investigative Dermatology,* **132**(3): 751-762.

Capecchi, M. R. (1989). The new mouse genetics: altering the genome by gene targeting. *Trends in Genetics,* **5**(3): 70-76.

Capecchi, M. R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nature Reviews Genetics,* **6**(6): 507-512.

Carmi, R., Boughman, J. A. & Rosenbaum, K. R. (1992). Human situs determination is probably controlled by several different genes. *Am J Med Genet,* **44**(2): 246-249.

Carrel, T., Purandare, S. M., Harrison, W., Elder, F., Fox, T., Casey, B. & Herman, G. E. (2000). The X-linked mouse mutation Bent tail is associated with a deletion of the Zic3 locus. *Hum Mol Genet,* **9**(13): 1937-1942.

Caruana, G., Farlie, P. G., Hart, A. H., Bagheri-Fam, S., Wallace, M. J., Dobbie, M. S., Gordon, C. T., Miller, K. A., Whittle, B., Abud, H. E., Arkell, R. M., Cole, T. J., Harley, V. R., Smyth, I. M. & Bertram, J. F. (2013). Genome-wide ENU mutagenesis in combination with high density SNP analysis and exome sequencing provides rapid identification of novel mouse models of developmental disease. *PLoS One,* **8**(3): e55429.

Chang, J. T., Lowery, L. A. & Sive, H. (2012). Multiple roles for the Na,K-ATPase subunits, Atp1a1 and Fxyd1, during brain ventricle development. *Developmental biology,* **368**(2): 312-322.

Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Developmental Cell,* **10**(5): 615-624.

Chen, J., Suo, S. B., Tam, P. P. L., Han, J. D. J., Peng, G. D. & Jing, N. H. (2017). Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nature protocols,* **12**(3): 566-580.

Cheng, X., Hsu, C. M., Currle, D. S., Hu, J. S., Barkovich, A. J. & Monuki, E. S. (2006). Central roles of the roof plate in telencephalic development and holoprosencephaly. *Journal of Neuroscience,* **26**(29): 7640-7649.

Chhin, B., Hatayama, M., Bozon, D., Ogawa, M., Schon, P., Tohmonda, T., Sassolas, F., Aruga, J., Valard, A. G., Chen, S. C. & Bouvagnet, P. (2007). Elucidation of penetrance variability of a ZIC3 mutation in a family with complex heart defects and functional analysis of ZIC3 mutations in the first zinc finger domain. *Human Mutation,* **28**(6): 563-570.

Choi, Y. & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics,* **31**(16): 2745-2747.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. (2012). Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One,* **7**(10).

Clevers, H. & Nusse, R. (2012). Wnt/beta-catenin signaling and disease. *Cell,* **149**(6): 1192-1205.

Coletta, A., Pinney, J. W., Solis, D. Y. W., Marsh, J., Pettifer, S. R. & Attwood, T. K. (2010). Low-complexity regions within protein sequences have position-dependent roles. *Bmc Systems Biology,* **4**.

Compernolle, V., Brusselmans, K., Franco, D., Moorman, A., Dewerchin, M., Collen, D. & Carmeliet, P. (2003). Cardia bifida, defective heart development and abnormal neural crest migration in embryos lacking hypoxia-inducible factor-1alpha. *Cardiovascular research,* **60**(3): 569-579.

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. & Robertson, E. J. (1994). A Primary Requirement for Nodal in the Formation and Maintenance of the Primitive Streak in the Mouse. *Development,* **120**(7): 1919-1928.

Contzler, R., Favre, B., Huber, M. & Hohl, D. (2005). Cornulin, a new member of the "fused gene" family, is expressed during epidermal differentiation. *Journal of Investigative Dermatology,* **124**(5): 990-997.

Cowan, J., Tariq, M. & Ware, S. M. (2014). Genetic and Functional Analyses of ZIC3 Variants in Congenital Heart Disease. *Human Mutation,* **35**(1): 66-75.

Dale, B. A., Resing, K. A. & Lonsdaleeccles, J. D. (1985). Filaggrin - a Keratin Filament Associated Protein. *Annals of the New York Academy of Sciences,* **455**: 330-342.

D'Alessandro, L. C., Casey, B. & Siu, V. M. (2013). Situs inversus totalis and a novel ZIC3 mutation in a family with X-linked heterotaxy. *Congenit Heart Dis,* **8**(2): E36-40.

Davis, A. P. & Justice, M. J. (1998). An Oak Ridge legacy: the specific locus test and its role in mouse mutagenesis. *Genetics,* **148**(1): 7-12.

de Angelis, M. H., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M., Reis, A., Richter, T., Alessandrini, F., Jakob, T., Fuchs, E., Kolb, H., Kremmer, E., Schaeble, K., Rollinski, B., Roscher, A., Peters, C., Meitinger, T., Strom, T., Steckler, T., Holsboer, F., Klopstock, T., Gekeler, F., Schindewolf, C., Jung, T., Avraham, K., Behrendt, H., Ring, J., Zimmer, A., Schughart, K., Pfeffer, K., Wolf, E. & Balling, R. (2000). Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nature Genetics,* **25**(4): 444-447.

Delling, M., Indzhykulian, A. A., Liu, X., Li, Y., Xie, T., Corey, D. P. & Clapham, D. E. (2016). Primary cilia are not calcium-responsive mechanosensors. Nature, **531**(7596): 656-660.

Denny, P. & Justice, M. J. (2000). Mouse as the measure of man? *Trends in Genetics,* **16**(7): 283-287.

Deriano, L. & Roth, D. B. (2013). Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annual Review of Genetics, ,* **47**: 433-455.

Doi, A., Takagi, M., Fujimoto, K., Kakihara, J., Hayashi, Y., Tatsumi, H. & Yoshiyama, M. (2016). Long RP' Tachycardia With Unusual Entrainment Responses: What Is the Mechanism? Journal of Cardiovascular Electrophysiology, **27**(10): 1242-1244.

Doudna, J. A. & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science,* **346**(6213): 1077-+.

Echelard, Y., Epstein, D. J., Stjacques, B., Shen, L., Mohler, J., Mcmahon, J. A. & Mcmahon, A. P. (1993). Sonic-Hedgehog, a Member of a Family of Putative Signaling Molecules, Is Implicated in the Regulation of Cns Polarity. *Cell,* **75**(7): 1417-1430.

Elms, P., Scurry, A., Davies, J., Willoughby, C., Hacker, T., Bogani, D. & Arkell, R. (2004). Overlapping and distinct expression domains of Zic2 and Zic3 during mouse gastrulation. *Gene Expr Patterns,* **4**(5): 505-511.

Elms, P., Siggers, P., Napper, D., Greenfield, A. & Arkell, R. (2003). Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Developmental Biology,* **264**(2): 391-406.

Evans, M. J. & Kaufman, M. H. (1981). Establishment in Culture of Pluripotential Cells from Mouse Embryos. *Nature,* **292**(5819): 154-156.

Fairfield, H., Gilbert, G. J., Barter, M., Corrigan, R. R., Curtain, M., Ding, Y., D'Ascenzo, M., Gerhardt, D. J., He, C., Huang, W., Richmond, T., Rowe, L., Probst, F. J., Bergstrom, D. E., Murray, S. A., Bult, C., Richardson, J., Kile, B. T., Gut, I., Hager, J., Sigurdsson, S., Mauceli, E., Di Palma, F., Lindblad-Toh, K., Cunningham, M. L., Cox, T. C., Justice, M. J., Spector, M. S., Lowe, S. W., Albert, T., Donahue, L. R., Jeddeloh, J., Shendure, J. & Reinholdt, L. G.

(2011). Mutation discovery in mice by whole exome sequencing. *Genome biology,* **12**(9): R86.

Favor, J., Neuhauser-Klaus, A. & Ehling, U. H. (1988). The effect of dose fractionation on the frequency of ethylnitrosourea-induced dominant cataract and recessive specific locus mutations in germ cells of the mouse. *Mutation research,* **198**(2): 269-275.

Ferguson, E. L. & Horvitz, H. R. (1985). Identification and Characterization of 22 Genes That Affect the Vulval Cell Lineages of the Nematode Caenorhabditis-Elegans. *Genetics,* **110**(1): 17-72.

Fernandes, M., Gutin, G., Alcorn, H., McConnell, S. K. & Hebert, J. M. (2007). Mutations in the BMP pathway in mice support the existence of two molecular classes of holoprosencephaly. *Development,* **134**(21): 3789-3794.

Festing, Michael FW (1998). INBRED STRAINS OF MICE, Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: http://www.informatics.jax.org)

Field, S., Riley, K. L., Grimes, D. T., Hilton, H., Simon, M., Powles-Glover, N., Siggers, P., Bogani, D., Greenfield, A. & Norris, D. P. (2011). Pkd1l1 establishes left-right asymmetry and physically interacts with Pkd2. *Development,* **138**(6): 1131-1142.

Fossat, N., Jones, V., Khoo, P. L., Bogani, D., Hardy, A., Steiner, K., Mukhopadhyay, M., Westphal, H., Nolan, P. M., Arkell, R. & Tam, P. P. (2011). Stringent requirement of a proper level of canonical WNT signalling activity for head formation in mouse embryo. *Development,* **138**(4): 667-676.

Fossat, N., Jones, V., Khoo, P. L., Bogani, D., Hardy, A., Steiner, K., Mukhopadhyay, M., Westphal, H., Nolan, P. M., Arkell, R. & Tam, P. P. (2011). Stringent requirement of a proper level of canonical WNT signalling activity for head formation in mouse embryo. *Development,* **138**(4): 667-676.

Fujimi, T. J., Hatayama, M. & Aruga, J. (2012). Xenopus Zic3 controls notochord and organizer development through suppression of the Wnt/beta-catenin signaling pathway. *Dev Biol,* **361**(2): 220-231.

Furushima, K., Murata, T., Matsuo, I. & Aizawa, S. (2000). A new murine zinc finger gene, Opr. *Mechanisms of Development,* **98**(1-2): 161-164.

Garber, E. D. (1952). Bent-Tail, a Dominant, Sex-Linked Mutation in the Mouse. *Genetics,* **37**(5): 583-583.

Garcia-Garcia, M. J., Eggenschwiler, J. T., Caspary, T., Alcorn, H. L., Wyler, M. R., Huangfu, D., Rakeman, A. S., Lee, J. D., Feinberg, E. H., Timmer, J. R. & Anderson, K. V. (2005). Analysis of mouse embryonic patterning and morphogenesis by forward genetics. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(17): 5913-5919.

Gebbia, M., Ferrero, G. B., Pilia, G., Bassi, M. T., Aylsworth, A. S., PenmanSplitt, M., Bird, L. M., Bamforth, J. S., Burn, J., Schlessinger, D., Nelson, D. L. & Casey, B. (1997). X-linked situs abnormalities result from mutations in ZIC3. *Nature Genetics,* **17**(3): 305-308.

Grigoryan, T., Wend, P., Klaus, A. & Birchmeier, W. (2008). Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of alpha-catenin in mice. *Genes & Development,* **22**(17): 2308-2341.

Grinblat, Y., Gamse, J., Patel, M. & Sive, H. (1998). Determination of the zebrafish forebrain: induction and patterning. *Development,* **125**(22): 4403-4416.

Ha, S., Stottmann, R. W., Furley, A. J. & Beier, D. R. (2015). A Forward Genetic Screen in Mice Identifies Mutants with Abnormal Cortical Patterning. *Cerebral Cortex,* **25**(1): 167-179.

Haegel, H., Larue, L., Ohsugi, M., Fedorov, L., Herrenknecht, K. & Kemler, R. (1995). Lack of Beta-Catenin Affects Mouse Development at Gastrulation. *Development,* **121**(11): 3529-3537.

Hansmann, B., Schrder, J. M. & Gerstel, U. (2015). Skin-Derived C-Terminal Filaggrin-2 Fragments Are Pseudomonas aeruginosa-Directed Antimicrobials Targeting Bacterial Replication. *Plos Pathogens,* **11**(9).

Harland, R. & Gerhart, J. (1997). Formation and function of Spemann's organizer. *Annual Review of Cell and Developmental Biology,* **13**: 611-667.

Hart, C. P., Awgulewitsch, A., Fainsod, A., Mcginnis, W. & Ruddle, F. H. (1985). Homeo Box Gene-Complex on Mouse Chromosome-11 - Molecular-Cloning, Expression in Embryogenesis, and Homology to a Human Homeo Box Locus. *Cell,* **43**(1): 9-18.

He, X. J., Tan, C. L., Wang, F., Wang, Y. F., Zhou, R., Cui, D. X., You, W. X., Zhao, H., Ren, J. W. & Feng, B. (2016). Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Research,* **44**(9).

Hedrich, H. (2004). *The laboratory mouse*, Academic Press.

Henderson, B. R. & Fagotto, F. (2002). The ins and outs of APC and beta-catenin nuclear transport. *Embo Reports,* **3**(9): 834-839.

Hentges, K., Thompson, K. & Peterson, A. (1999). The flat-top gene is required for the expansion and regionalization of the telencephalic primordium. *Development,* **126**(8): 1601-1609.

Herman, G. E. & El-Hodiri, H. M. (2002). The role of ZIC3 in vertebrate development. *Cytogenet Genome Res,* **99**(1-4): 229-235.

Herrera, V. L. & Ruiz-Opazo, N. (1990). Alteration of alpha 1 Na+,K(+)-ATPase 86Rb+ influx by a single amino acid substitution. *Science,* **249**(4972): 1023-1026.

Herron, B. J., Lu, W., Rao, C., Liu, S., Peters, H., Bronson, R. T., Justice, M. J., McDonald, J. D. & Beier, D. R. (2002). Efficient generation and mapping of recessive developmental mutations using ENU mutagenesis. *Nat Genet,* **30**(2): 185-189.

Heuberger, J. & Birchmeier, W. (2010). Interplay of Cadherin-Mediated Cell Adhesion and Canonical Wnt Signaling. *Cold Spring Harbor Perspectives in Biology,* **2**(2).

Hierholzer, A. & Kemler, R. (2010). Beta-catenin-mediated signaling and cell adhesion in postgastrulation mouse embryos. *Developmental dynamics : an official publication of the American Association of Anatomists,* **239**(1): 191-199.

Himeda, C. L., Barro, M. V. & Emerson, C. P. (2013). Pax3 synergizes with Gli2 and Zic1 in transactivating the Myf5 epaxial somite enhancer. *Developmental Biology,* **383**(1): 7-14.

Hitotsumachi, S., Carpenter, D. A. & Russell, W. L. (1985). Dose-Repetition Increases the Mutagenic Effectiveness of N-Ethyl-N-Nitrosourea in Mouse Spermatogonia. *Proceedings of the National Academy of Sciences of the United States of America,* **82**(19): 6619-6621.

Houtmeyers, R., Souopgui, J., Tejpar, S. & Arkell, R. (2013). The ZIC gene family encodes multi-functional proteins essential for patterning and morphogenesis. *Cellular and molecular life sciences : CMLS,* **70**(20): 3791-3811.

Houtmeyers, R., Tchouate Gainkam, O., Glanville-Jones, H. A., Van den Bosch, B., Chappell, A., Barratt, K. S., Souopgui, J., Tejpar, S. & Arkell, R. M. (2016). Zic2 mutation causes holoprosencephaly via disruption of NODAL signalling. *Human molecular genetics*.

Huber, M., Siegenthaler, G., Mirancea, N., Marenholz, I., Nizetic, D., Breitkreutz, D., Mischke, D. & Hohl, D. (2005). Isolation and characterization of human repetin, a member of the fused gene family of the epidermal differentiation complex. *Journal of Investigative Dermatology,* **124**(5): 998-1007.

Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C. & Birchmeier, W. (2000). Requirement for beta-catenin in anterior-posterior axis formation in mice. *Journal of Cell Biology,* **148**(3): 567-578.

Ingham, P. W. (1988). The Molecular-Genetics of Embryonic Pattern-Formation in Drosophila. *Nature,* **335**(6185): 25-34.

Ishida-Yamamoto, A., Hashimoto, Y., Manabe, M., OGuin, W. M., Dale, B. A. & Iizuka, H. (1997). Distinctive expression of filaggrin and trichohyalin during various pathways of epithelial differentiation. *British Journal of Dermatology,* **137**(1): 9-16.

IshidaYamamoto, A., Hashimoto, Y., Manabe, M., OGuin, W. M., Dale, B. A. & Iizuka, H. (1997). Distinctive expression of filaggrin and trichohyalin during various pathways of epithelial differentiation. *British Journal of Dermatology,* **137**(1): 9-16.

Itasaki, N., Jones, C. M., Mercurio, S., Rowe, A., Domingos, P. M., Smith, J. C. & Krumlauf, R. (2003). Wise, a context-dependent activator and inhibitor of Wnt signalling. *Development,* **130**(18): 4295-4305.

Iuchi, S. (2001). Three classes of C2H2 zinc finger proteins. *Cellular and Molecular Life Sciences,* **58**(4): 625-635.

James, P. F., Grupp, I. L., Grupp, G., Woo, A. L., Askew, G. R., Croyle, M. L., Walsh, R. A. & Lingrel, J. B. (1999). Identification of a specific role for the Na,K-ATPase alpha 2 isoform as a regulator of calcium in the heart. *Molecular cell,* **3**(5): 555-563.

James, P. F., Grupp, I. L., Grupp, G., Woo, A. L., Askew, G. R., Croyle, M. L., Walsh, R. A. & Lingrel, J. B. (1999). Identification of a specific role for the Na,K-ATPase alpha 2 isoform as a regulator of calcium in the heart. *Molecular cell,* **3**(5): 555-563.

Jho, E. H., Zhang, T., Domon, C., Joo, C. K., Freund, J. N. & Costantini, F. (2002). Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Molecular and cellular biology,* **22**(4): 1172-1183.

Jiang, Z., Zhu, L., Hu, L., Slesnick, T. C., Pautler, R. G., Justice, M. J. & Belmont, J. W. (2013). Zic3 is required in the extra-cardiac perinodal region of the lateral plate mesoderm for left-right patterning and heart development. *Hum Mol Genet,* **22**(5): 879-889.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* **337**(6096): 816-821.

Justice, M. J. (2000). Capitalizing on large-scale mouse mutagenesis screens. *Nature Reviews Genetics,* **1**(2): 109-115.

Justice, M. J., Noveroske, J. K., Weber, J. S., Zheng, B. & Bradley, A. (1999). Mouse ENU mutagenesis. *Human molecular genetics,* **8**(10): 1955-1963.

Kasarskis, A., Manova, K. & Anderson, K. V. (1998). A phenotype-based screen for embryonic lethal mutations in the mouse. *Proceedings of the National Academy of Sciences of the United States of America,* **95**(13): 7485-7490.

Kasarskis, A., Manova, K. & Anderson, K. V. (1998). A phenotype-based screen for embryonic lethal mutations in the mouse. *Proceedings of the National Academy of Sciences of the United States of America,* **95**(13): 7485-7490.

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., Furlotte, N. A., Eskin, E., Nellaker, C., Whitley, H., Cleak, J., Janowitz, D., Hernandez-Pliego, P., Edwards, A., Belgard, T. G., Oliver, P. L., McIntyre, R. E., Bhomra, A., Nicod, J., Gan, X. C., Yuan, W., van der Weyden, L., Steward, C. A., Bala, S., Stalker, J., Mott, R., Durbin, R., Jackson, I. J., Czechanski, A., Guerra-Assuncao, J. A., Donahue, L. R., Reinholdt, L. G., Payseur, B. A., Ponting, C. P., Birney, E., Flint, J. & Adams, D. J. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature,* **477**(7364): 289-294.

Kearns-Jonker, M. (2006). *Congenital heart disease : molecular diagnostics*, Totowa, N.J. : Humana Press, c2006.

Keays, D. A., Clark, T. G. & Flint, J. (2006). Estimating the number of coding mutations in genotypic- and phenotypic-driven N-ethyl-N-nitrosourea (ENU) screens. *Mammalian Genome,* **17**(3): 230-238.

Kile, B. T., Hentges, K. E., Clark, A. T., Nakamura, H., Salinger, A. P., Liu, B., Box, N., Stockton, D. W., Johnson, R. L., Behringer, R. R., Bradley, A. & Justice, M. J. (2003). Functional genetic analysis of mouse chromosome 11. *Nature,* **425**(6953): 81-86.

Koebernick, K. & Pieler, T. (2002). Gli-type zinc finger proteins as bipotential transducers of Hedgehog signaling. *Differentiation,* **70**(2-3): 69-76.

Koller, B. H., Hagemann, L. J., Doetschman, T., Hagaman, J. R., Huang, S., Williams, P. J., First, N. L., Maeda, N. & Smithies, O. (1989). Germ-Line Transmission of a Planned Alteration Made in a Hypoxanthine Phosphoribosyltransferase Gene by Homologous Recombination in Embryonic Stem-Cells. *Proceedings of the National Academy of Sciences of the United States of America,* **86**(22): 8927-8931.

Koyabu, Y., Nakata, K., Mizugishi, K., Aruga, J. & Mikoshiba, K. (2001). Physical and functional interactions between Zic and Gli proteins. *J Biol Chem,* **276**(10): 6889-6892.

Krieg, P., Schuppler, M., Koesters, R., Mincheva, A., Lichter, P. & Marks, F. (1997). Repetin (Rptn), a new member of the ''fused gene'' subgroup within the S100 gene family encoding a murine epidermal differentiation protein. *Genomics,* **43**(3): 339-348.

Kuehl, K. S. & Loffredo, C. (2002). Risk factors for heart disease associated with abnormal sidedness. *Teratology,* **66**(5): 242-248.

Landy, A. (1989). Dynamic, Structural, and Regulatory Aspects of Lambda-Site-Specific Recombination. *Annual Review of Biochemistry,* **58**: 913-949.

Lee, J. D. & Anderson, K. V. (2008). Morphogenesis of the node and notochord: the cellular basis for the establishment and maintenance of left-right asymmetry in the mouse. *Developmental dynamics : an official publication of the American Association of Anatomists,* **237**(12): 3464-3476.

Lee, J. D., Migeotte, I. & Anderson, K. V. (2010). Left-right patterning in the mouse requires Epb4.1l5-dependent morphogenesis of the node and midline. *Dev Biol,* **346**(2): 237-246.

Lee, S. C., Wang, M., Mcbride, O. W., Okeefe, E. J., Kim, I. G. & Steinert, P. M. (1993). Human Trichohyalin Gene Is Clustered with the Genes for Other Epidermal Structural Proteins and Calcium-Binding Proteins at Chromosomal Locus 1q21. *Journal of Investigative Dermatology,* **100**(1): 65-68.

Levenson, R. (1994). Isoforms of the Na,K-ATPase: family members in search of function. *Reviews of physiology, biochemistry and pharmacology,* **123**: 1-45.

Lim, L. S., Loh, Y. H., Zhang, W. W., Li, Y. X., Chen, X., Wang, Y. N., Bakre, M., Ng, H. H. & Stanton, L. W. (2007). Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Molecular Biology of the Cell,* **18**(4): 1348-1358.

Lin, A. E., Ticho, B. S., Houde, K., Westgate, M. N. & Holmes, L. B. (2000). Heterotaxy: associated conditions and hospital-based prevalence in newborns. *Genetics in medicine : official journal of the American College of Medical Genetics,* **2**(3): 157-172.

Liu, P. T., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R. & Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nature Genetics,* **22**(4): 361-365.

Love, J. M., Knight, A. M., Mcaleer, M. A. & Todd, J. A. (1990). Towards construction of a high-resolution map of the mouse menome using pcr-analyzed microsatellites. *Nucleic Acids Research,* **18**(14): 4123-4130.

MacDonald, B. T., Adamska, M. & Meisler, M. H. (2004). Hypomorphic expression of Dkk1 in the doubleridge mouse: dose dependence and compensatory interactions with Lrp6. *Development,* **131**(11): 2543-2552.

MacDonald, B. T., Tamai, K. & He, X. (2009). Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Developmental cell,* **17**(1): 9-26.

MacPhee, D. J., Jones, D. H., Barr, K. J., Betts, D. H., Watson, A. J. & Kidder, G. M. (2000). Differential involvement of Na(+),K(+)-ATPase isozymes in preimplantation development of the mouse. *Dev Biol,* **222**(2): 486-498.

Magin, T. M., McWhir, J. & Melton, D. W. (1992). A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency. *Nucleic Acids Research,* **20**(14): 3795-3796.

Makino, T., Mizawa, M., Yamakoshi, T., Takaishi, M. & Shimizu, T. (2014). Expression of filaggrin-2 protein in the epidermis of human skin diseases: A comparative analysis with filaggrin. *Biochemical and Biophysical Research Communications,* **449**(1): 100-106.

Makino, T., Takaishi, M., Morohashi, M. & Huh, N. H. (2001). Hornerin, a novel profilaggrin-like protein and differentiation-specific marker isolated from mouse skin. *Journal of Biological Chemistry,* **276**(50): 47445-47452.

Malakoff, D. (2000). The rise of the mouse, biomedicine's model mammal. *Science,* **288**(5464): 248-253.

Mallon, A. M., Wilming, L., Weekes, J., Gilbert, J. G., Ashurst, J., Peyrefitte, S., Matthews, L., Cadman, M., McKeone, R., Sellick, C. A., Arkell, R., Botcherby, M. R., Strivens, M. A.,

Campbell, R. D., Gregory, S., Denny, P., Hancock, J. M., Rogers, J. & Brown, S. D. (2004). Organization and evolution of a gene-rich region of the mouse genome: a 12.7-Mb region deleted in the Del(13)Svea36H mouse. *Genome research,* **14**(10A): 1888-1901.

Marenholz, I., Heizmann, C. W. & Fritz, G. (2004). S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochemical and Biophysical Research Communications,* **322**(4): 1111-1122.

Marenholz, I., Heizmann, C. W. & Fritz, G. (2004). S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochemical and Biophysical Research Communications,* **322**(4): 1111-1122.

Martin, D. M., Skidmore, J. M., Fox, S. E., Gage, P. J. & Camper, S. A. (2002). Pitx2 distinguishes subtypes of terminally differentiated neurons in the developing mouse neuroepithelium. *Developmental biology,* **252**(1): 84-99.

McGrath, J., Somlo, S., Makova, S., Tian, X. & Brueckner, M. (2003). Two populations of node monocilia initiate left-right asymmetry in the mouse. *Cell,* **114**(1): 61-73.

McLaren, A. (1976). Genetics of the early mouse embryo. *Annual review of genetics,* **10**: 361-388.

Megarbane, A., Salem, N., Stephan, E., Ashoush, R., Lenoir, D., Delague, V., Kassab, R., Loiselet, J. & Bouvagnet, A. (2000). X-linked transposition of the great arteries and incomplete penetrance among males with a nonsense mutation in ZIC3. *European Journal of Human Genetics,* **8**(9): 704-708.

Meno, C., Shimono, A., Saijoh, Y., Yashiro, K., Mochida, K., Ohishi, S., Noji, S., Kondoh, H. & Hamada, H. (1998). lefty-1 is required for left-right determination as a regulator of lefty-2 and nodal. *Cell,* **94**(3): 287-297.

Merrill, B. J., Pasolli, H. A., Polak, L., Rendl, M., Garcia-Garcia, M. J., Anderson, K. V. & Fuchs, E. (2004). Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development,* **131**(2): 263-274.

Mesnard, D., Filipe, M., Belo, J. A. & Zernicka-Goetz, M. (2004). The anterior-posterior axis emerges respecting the morphology of the mouse embryo that changes and aligns with the uterus before gastrulation. *Current Biology,* **14**(3): 184-196.

Mizugishi, K., Aruga, J., Nakata, K. & Mikoshiba, K. (2001). Molecular properties of Zic proteins as transcriptional regulators and their relationship to GLI proteins. *Journal of Biological Chemistry,* **276**(3): 2180-2188.

Mizugishi, K., Hatayama, M., Tohmonda, T., Ogawa, M., Inoue, T., Mikoshiba, K. & Aruga, J. (2004). Myogenic repressor I-mfa interferes with the function of Zic family proteins. *Biochemical and Biophysical Research Communications,* **320**(1): 233-240.

Mohamed, O. A., Clarke, H. J. & Dufort, D. (2004). beta-catenin signaling marks the prospective site of primitive streak formation in the mouse embryo. *Developmental Dynamics,* **231**(2): 416-424.

Montagutelli, X. (2000). Effect of the genetic background on the phenotype of mouse mutations. *Journal of the American Society of Nephrology : JASN,* **11 Suppl 16**: S101-105.

Moorman, A. F. & Christoffels, V. M. (2003). Cardiac chamber formation: development, genes, and evolution. *Physiological reviews,* **83**(4): 1223-1267.

Morkel, M., Huelsken, J., Wakamiya, M., Ding, J., van de Wetering, M., Clevers, H., Taketo, M. M., Behringer, R. R., Shen, M. M. & Birchmeier, W. (2003). Beta-catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. *Development,* **130**(25): 6283-6294.

Mosimann, C., Hausmann, G. & Basler, K. (2009). Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nature reviews. Molecular cell biology,* **10**(4): 276-286.

Mukhopadhyay, M., Shtrom, S., Rodriguez-Esteban, C., Chen, L., Tsukui, T., Gomer, L., Dorward, D. W., Glinka, A., Grinberg, A., Huang, S. P., Niehrs, C., Belmonte, J. C. I. & Westphal, H. (2001). Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. *Developmental Cell,* **1**(3): 423-434.

Mukhopadhyay, M., Shtrom, S., Rodriguez-Esteban, C., Chen, L., Tsukui, T., Gomer, L., Dorward, D. W., Glinka, A., Grinberg, A., Huang, S. P., Niehrs, C., Belmonte, J. C. I. & Westphal, H. (2001). Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. *Developmental Cell,* **1**(3): 423-434. MacDonald, B. T., Adamska, M. & Meisler, M. H. (2004). Hypomorphic expression of Dkk1 in the doubleridge mouse: dose dependence and compensatory interactions with Lrp6. *Development,* **131**(11): 2543-2552.

Muller, U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mechanisms of Development,* **82**(1-2): 3-21.

Nakamura, T., Saito, D., Kawasumi, A., Shinohara, K., Asai, Y., Takaoka, K., Dong, F. L., Takamatsu, A., Belo, J. A., Mochizuki, A. & Hamada, H. (2012). Fluid flow and interlinked feedback loops establish left-right asymmetric decay of Cerl2 mRNA. *Nature Communications,* **3**.

Neuhaus, I. M. & Beier, D. R. (1998). Efficient localization of mutations by interval haplotype analysis. *Mammalian Genome,* **9**(2): 150-154.

Neuhaus, I. M. & Beier, D. R. (1998b). Efficient localization of mutations by interval haplotype analysis. *Mammalian genome : official journal of the International Mammalian Genome Society,* **9**(2): 150-154.

Nolan, P. M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I. C., Vizor, L., Brooker, D., Whitehill, E., Washbourne, R., Hough, T., Greenaway, S., Hewitt, M., Liu, X. H., McCormack, S., Pickford, K., Selley, R., Wells, C., Tymowska-Lalanne, Z., Roby, P., Glenister, P., Thornton, C., Thaung, C., Stevenson, J. A., Arkell, R., Mburu, P., Hardisty, R., Kiernan, A., Erven, H., Steel, K. P., Voegeling, S., Guenet, J. L., Nickols, C., Sadri, R., Naase, M., Isaacs, A., Davies, K., Browne, M., Fisher, E. M. C., Martin, J., Rastan, S., Brown, S. D. M. & Hunter, J. (2000). A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nature Genetics,* **25**(4): 440-443.

Nonaka, S., Shiratori, H., Saijoh, Y. & Hamada, H. (2002). Determination of left-right patterning of the mouse embryo by artificial nodal flow. *Nature,* **418**(6893): 96-99.

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., Kido, M. & Hirokawa, N. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell,* **95**(6): 829-837.

Norris, D. P. (2012). Cilia, calcium and the basis of left-right asymmetry. *BMC biology,* **10**: 102.

Norum, J. H., Bergstrom, A., Andersson, A. B., Kuiper, R. V., Hoelzl, M. A., Sorlie, T. & Toftgard, R. (2015). A conditional transgenic mouse line for targeted expression of the stem cell marker LGR5. *Developmental biology,* **404**(2): 35-48.

Noveroske, J. K., Weber, J. S. & Justice, M. J. (2000). The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mammalian genome : official journal of the International Mammalian Genome Society,* **11**(7): 478-483.

Nurse, P., Thuriaux, P. & Nasmyth, K. (1976). Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *Molecular & general genetics : MGG,* **146**(2): 167-178.

Nusse, R. (2005). Wnt signaling in disease and in development. *Cell Research,* **15**(1): 28-32.

Nusslein-Volhard, C. & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature,* **287**(5785): 795-801.

Pennekamp, P., Karcher, C., Fischer, A., Schweickert, A., Skryabin, B., Horst, J., Blum, M. & Dworniczak, B. (2002). The ion channel polycystin-2 is required for left-right axis determination in mice. *Current biology : CB,* **12**(11): 938-943.

Perea-Gomez, A., Camus, A., Moreau, A., Grieve, K., Moneron, G., Dubois, A., Cibert, C. & Collignon, J. (2004). Initiation of gastrulation in the mouse embryo is preceded by an apparent shift in the orientation of the anterior-posterior axis. *Current Biology,* **14**(3): 197-207.

Petersen, C. P. & Reddien, P. W. (2009). Wnt signaling and the polarity of the primary body axis. *Cell,* **139**(6): 1056-1068.

Petkov, P. M., Ding, Y., Cassell, M. A., Zhang, W., Wagner, G., Sargent, E. E., Asquith, S., Crew, V., Johnson, K. A., Robinson, P., Scott, V. E. & Wiles, M. V. (2004). An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome research,* **14**(9): 1806-1811.

Piedra, J., Martinez, D., Castano, J., Miravet, S., Dunach, M. & de Herreros, A. G. (2001). Regulation of beta-catenin structure and activity by tyrosine phosphorylation. *The Journal of biological chemistry,* **276**(23): 20436-20443.

Popperl, H., Schmidt, C., Wilson, V., Hume, C. R., Dodd, J., Krumlauf, R. & Beddington, R. S. (1997). Misexpression of Cwnt8C in the mouse induces an ectopic embryonic axis and causes a truncation of the anterior neuroectoderm. *Development,* **124**(15): 2997-3005.

Postlethwait, J. H., Yan, Y. L., Gates, M. A., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E. S., Force, A., Gong, Z. Y., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T. S., Yelick, P., Beier, D., Joly, J. S., Larhammar, D., Rosa, F., Westerfield, M., Zon, L. I., Johnson, S. L. & Talbot, W. S. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics,* **18**(4): 345-349.

Potter, M. & Mock, B. (1995). Mouse Genetics - Concepts and Applications - Silver,Lm. *Science,* **270**(5242): 1692-1693.

Potter, P. K., Bowl, M. R., Jeyarajan, P., Wisby, L., Blease, A., Goldsworthy, M. E., Simon, M. M., Greenaway, S., Michel, V., Barnard, A., Aguilar, C., Agnew, T., Banks, G., Blake, A., Chessum, L., Dorning, J., Falcone, S., Goosey, L., Harris, S., Haynes, A., Heise, I., Hillier, R., Hough, T., Hoslin, A., Hutchison, M., King, R., Kumar, S., Lad, H. V., Law, G., MacLaren, R. E., Morse, S., Nicol, T., Parker, A., Pickford, K., Sethi, S., Starbuck, B., Stelma, F., Cheeseman, M., Cross, S. H., Foster, R. G., Jackson, I. J., Peirson, S. N., Thakker, R. V., Vincent, T., Scudamore, C., Wells, S., El-Amraoui, A., Petit, C., Acevedo-Arozena, A., Nolan, P. M., Cox, R., Mallon, A. M. & Brown, S. D. M. (2016). Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nature Communications,* **7**.

Pourebrahim, R., Houtmeyers, R., Ghogomu, S., Janssens, S., Thelie, A., Tran, H. T., Langenberg, T., Vleminckx, K., Bellefroid, E., Cassiman, J. J. & Tejpar, S. (2011). Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling. *The Journal of biological chemistry,* **286**(43): 37732-37740.

Purandare, S. M., Ware, S. M., Kwan, K. M., Gebbia, M., Bassi, M. T., Deng, J. M., Vogel, H., Behringer, R. R., Belmont, J. W. & Casey, B. (2002). A complex syndrome of left-right axis, central nervous system and axial skeleton defects in Zic3 mutant mice. *Development,* **129**(9): 2293-2302.

Qian, L., Mahaffey, J. P., Alcorn, H. L. & Anderson, K. V. (2011). Tissue-specific roles of Axin2 in the inhibition and activation of Wnt signaling in the mouse embryo. *Proceedings of the National Academy of Sciences of the United States of America,* **108**(21): 8692-8697.

Quinn, M. E., Haaning, A. & Ware, S. M. (2012). Preaxial polydactyly caused by Gli3 haploinsufficiency is rescued by Zic3 loss of function in mice. *Hum Mol Genet,* **21**(8): 1888-1896.

Ramsdell, A. F. (2005). Left-right asymmetry and congenital cardiac defects: getting to the heart of the matter in vertebrate left-right axis determination. *Developmental biology,* **288**(1): 1-20.

Rao, T. P. & Kuhl, M. (2010). An updated overview on Wnt signaling pathways: a prelude for more. *Circulation research,* **106**(12): 1798-1806.

Rathkolb, B., Fuchs, E., Kolb, H. J., Renner-Muller, I., Krebs, O., Balling, R., de Angelis, M. H. & Wolf, E. (2000). Large-scale N-ethyl-N-nitrosourea mutagenesis of mice - from phenotypes to genes. *Experimental Physiology,* **85**(6): 635-644.

Reinhard, L., Tidow, H., Clausen, M. J. & Nissen, P. (2013). Na+,K+-ATPase as a docking station: protein-protein complexes of the Na+,K+-ATPase. *Cellular and Molecular Life Sciences,* **70**(2): 205-222.

Rinchik, E. M., Carpenter, D. A. & Handel, M. A. (1995). Pleiotropy in microdeletion syndromes: neurologic and spermatogenic abnormalities in mice homozygous for the p6H deletion are likely due to dysfunction of a single gene. *Proceedings of the National Academy of Sciences of the United States of America,* **92**(14): 6394-6398.

Rinchik, E. M., Carpenter, D. A. & Selby, P. B. (1990). A Strategy for Fine-Structure Functional-Analysis of a 6-Centimorgan to 11-Centimorgan Region of Mouse Chromosome-7 by High-Efficiency Mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America,* **87**(3): 896-900.

Russell, W. L. (1951). X-ray-induced mutations in mice. *Cold Spring Harbor Symposia on Quantitative Biology,* **16**: 327-336.

Russell, W. L., Hunsicker, P. R., Carpenter, D. A., Cornett, C. V. & Guinn, G. M. (1982a). Effect of Dose Fractionation on the Ethylnitrosourea Induction of Specific-Locus Mutations in Mouse Spermatogonia. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* **79**(11): 3592-3593.

Russell, W. L., Hunsicker, P. R., Raymer, G. D., Steele, M. H., Stelzner, K. F. & Thompson, H. M. (1982b). Dose-Response Curve for Ethylnitrosourea-Induced Specific-Locus Mutations in Mouse Spermatogonia. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* **79**(11): 3589-3591.

Russell, W. L., Kelly, E. M., Hunsicker, P. R., Bangham, J. W., Maddux, S. C. & Phipps, E. L. (1979). Specific-Locus Test Shows Ethylnitrosourea to Be the Most Potent Mutagen in the Mouse. *Proceedings of the National Academy of Sciences of the United States of America,* **76**(11): 5818-5819.

Sarkar, P. K. (2002). A quick assay for Na+-K+-ATPase specific activity. *Zeitschrift Fur Naturforschung C-a Journal of Biosciences,* **57**(5-6): 562-564.

Semenov, M. V., Tamai, K., Brott, B. K., Kuhl, M., Sokol, S. & He, X. (2001). Head inducer Dickkopf-1 is a ligand for Wnt coreceptor LRP6. *Current Biology,* **11**(12): 951-961.

Shen, M. M. (2007). Nodal signaling: developmental roles and regulation. *Development,* **134**(6): 1023-1034.

Shiratori, H. & Hamada, H. (2006). The left-right axis in the mouse: from origin to morphology. *Development,* **133**(11): 2095-2104.

Silver, D. L., Hou, L. & Pavan, W. J. (2006). The genetic regulation of pigment cell development. *Neural Crest Induction and Differentiation,* **589**: 155-169.

Simons, M. & Mlodzik, M. (2008). Planar Cell Polarity Signaling: From Fly Development to Human Disease. *Annual review of genetics,* **42**: 517-540.

Smith, F. J. D., Irvine, A. D., Terron-Kwiatkowski, A., Sandilands, A., Campbell, L. E., Zhao, Y. W., Liao, H. H., Evans, A. T., Goudie, D. R., Lewis-Jones, S., Arseculeratne, G., Munro, C. S., Sergeant, A., O'Regan, G., Bale, S. J., Compton, J. G., DiGiovanna, J. J., Presland, R. B., Fleckman, P. & McLean, W. H. I. (2006). Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nature Genetics,* **38**(3): 337-342.

Streisinger, G., Walker, C., Dower, N., Knauber, D. & Singer, F. (1981). Production of Clones of Homozygous Diploid Zebra Fish (Brachydanio-Rerio). *Nature,* **291**(5813): 293-296.

Stuckey, D. W., Clements, M., Di-Gregorio, A., Senner, C. E., Le Tissier, P., Srinivas, S. & Rodriguez, T. A. (2011). Coordination of cell proliferation and anterior-posterior axis establishment in the mouse embryo. *Development,* **138**(8): 1521-1530.

Sun, X., Meyers, E. N., Lewandoski, M. & Martin, G. R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes & Development,* **13**(14): 1834-1846.

Sutherland, M. J. & Ware, S. M. (2009). Disorders of left-right asymmetry: heterotaxy and situs inversus. *Am J Med Genet C Semin Med Genet,* **151C**(4): 307-317.

Sutherland, M. J., Wang, S. Y., Quinn, M. E., Haaning, A. & Ware, S. M. (2013). Zic3 is required in the migrating primitive streak for node morphogenesis and leftright patterning. *Human Molecular Genetics,* **22**(10): 1913-1923.

Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A. & McMahon, A. P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes & Development,* **8**(2): 174-189.

Takaoka, K., Yamamoto, M. & Hamada, H. (2011). Origin and role of distal visceral endoderm, a group of cells that determines anterior-posterior polarity of the mouse embryo. *Nature Cell Biology,* **13**(7): 743-U345.

Tam, P. P. & Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mechanisms of Development,* **68**(1-2): 3-25.

Tam, P. P. L. & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nature Reviews Genetics,* **8**(5): 368-381.

Tam, P. P. L., Loebel, D. A. F. & Tanaka, S. S. (2006). Building the mouse gastrula: signals, asymmetry and lineages. *Current opinion in genetics & development,* **16**(4): 419-425.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research,* **17**(16): 6463-6471.

Thomas, K. R. & Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem-cells. *Cell,* **51**(3): 503-512.

Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. & Melton, D. W. (1989). Germ Line Transmission and Expression of a Corrected Hprt Gene Produced by Gene Targeting in Embryonic Stem-Cells. *Cell,* **56**(2): 313-321.

Topol, L., Chen, W., Song, H., Day, T. F. & Yang, Y. (2009). Sox9 inhibits Wnt signaling by promoting beta-catenin phosphorylation in the nucleus. *The Journal of biological chemistry,* **284**(5): 3323-3333.

Toyama, R., Gomez, D. M., Mana, M. D. & Dawid, I. B. (2004). Sequence relationships and expression patterns of zebrafish zic2 and zic5 genes. *Gene Expression Patterns,* **4**(3): 345-350.

Trainor, P. A. & Tam, P. P. (1995). Cranial paraxial mesoderm and neural crest cells of the mouse embryo: co-distribution in the craniofacial mesenchyme but distinct segregation in branchial arches. *Development,* **121**(8): 2569-2582.

Valenta, T., Hausmann, G. & Basler, K. (2012). The many faces and functions of beta-catenin. *The EMBO journal,* **31**(12): 2714-2736.

van Veelen, W., Le, N. H., Helvensteijn, W., Blonden, L., Theeuwes, M., Bakker, E. R., Franken, P. F., van Gurp, L., Meijlink, F., van der Valk, M. A., Kuipers, E. J., Fodde, R. & Smits, R. (2011). beta-catenin tyrosine 654 phosphorylation increases Wnt signalling and intestinal tumorigenesis. *Gut,* **60**(9): 1204-1212.

van Veelen, W., Le, N. H., Helvensteijn, W., Blonden, L., Theeuwes, M., Bakker, E. R., Franken, P. F., van Gurp, L., Meijlink, F., van der Valk, M. A., Kuipers, E. J., Fodde, R. & Smits, R. (2011). beta-catenin tyrosine 654 phosphorylation increases Wnt signalling and intestinal tumorigenesis. *Gut,* **60**(9): 1204-1212.

Varlet, I., Collignon, J. & Robertson, E. J. (1997). Nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. *Development,* **124**(5): 1033-1044.

Viotti, M., Niu, L., Shi, S. H. & Hadjantonakis, A. K. (2012). Role of the gut endoderm in relaying left-right patterning in mice. *PLoS biology,* **10**(3): e1001276.

Waldrip, W. R., Bikoff, E. K., Hoodless, P. A., Wrana, J. L. & Robertson, E. J. (1998). Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. *Cell,* **92**(6): 797-808.

Wang, M. Q. & Sternberg, P. W. (2001). Pattern formation during C-elegans vulval induction. *Current Topics in Developmental Biology, Vol 51,* **51**: 189-220.

Wansleeben, C., van Gurp, L., Feitsma, H., Kroon, C., Rieter, E., Verberne, M., Guryev, V., Cuppen, E. & Meijlink, F. (2011). An ENU-mutagenesis screen in the mouse: identification of novel developmental gene functions. *PLoS One,* **6**(4).

Ware, S. M., Harutyunyan, K. G. & Belmont, J. W. (2006a). Heart defects in X-linked heterotaxy: evidence for a genetic interaction of Zic3 with the nodal signaling pathway. *Developmental dynamics : an official publication of the American Association of Anatomists,* **235**(6): 1631-1637.

Ware, S. M., Harutyunyan, K. G. & Belmont, J. W. (2006b). Zic3 is critical for early embryonic patterning during gastrulation. *Developmental Dynamics,* **235**(3): 776-785.

Ware, S. M., Peng, J., Zhu, L., Fernbach, S., Colicos, S., Casey, B., Towbin, J. & Belmont, J. W. (2004). Identification and functional analysis of ZIC3 mutations in heterotaxy and related congenital heart defects. *Am J Hum Genet,* **74**(1): 93-105.

Wessels, M. W., Kuchinka, B., Heydanus, R., Smit, B. J., Dooijes, D., de Krijger, R. R., Lequin, M. H., de Jong, E. M., Husen, M., Willems, P. J. & Casey, B. (2010). Polyalanine expansion in the ZIC3 gene leading to X-linked heterotaxy with VACTERL association: a new polyalanine disorder? *Journal of Medical Genetics,* **47**(5): 351-355.

Whitman, M. & Melton, D. A. (1989). Growth factors in early embryogenesis. *Annual review of cell biology,* **5**: 93-117.

Wieschaus, E., Nussleinvolhard, C. & Kluding, H. (1984). Kruppel, a gene whose activity is required early in the zygotic genome for normal embryonic segmentation. *Developmental Biology,* **104**(1): 172-186.

Wu, Z. H., Hansmann, B., Meyer-Hoffert, U., Glaser, R. & Schroder, J. M. (2009a). Molecular Identification and Expression Analysis of Filaggrin-2, a Member of the S100 Fused-Type Protein Family. *PLoS One,* **4**(4).

Wu, Z. H., Meyer-Hoffert, U., Reithmayer, K., Paus, R., Hansmann, B., He, Y. H., Bartels, J., Glaser, R., Harder, J. & Schroder, J. M. (2009b). Highly Complex Peptide Aggregates of the S100 Fused-Type Protein Hornerin Are Present in Human Skin. *Journal of Investigative Dermatology,* **129**(6): 1446-1458.

Xie, Z. J., Wang, Y. H., Ganjeizadeh, M., McGee, R., Jr. & Askari, A. (1989). Determination of total (Na+ + K+)-ATPase activity of isolated or cultured cells. *Analytical biochemistry,* **183**(2): 215-219.

Xu, Z. X., Wang, M. R., Xu, X., Cai, Y., Han, Y. L., Wu, K. M., Wang, J., Chen, B. S., Wang, X. Q. & Wu, M. (2000). Novel human esophagus-specific gene C1orf10: cDNA cloning, gene structure, and frequent loss of expression in esophageal cancer. *Genomics,* **69**(3): 322-330.

Yamamoto, M., Mine, N., Mochida, K., Sakai, Y., Saijoh, Y., Meno, C. & Hamada, H. (2003). Nodal signaling induces the midline barrier by activating Nodal expression in the lateral plate. *Development,* **130**(9): 1795-1804.

Zeng, L., Fagotto, F., Zhang, T., Hsu, W., Vasicek, T. J., Perry, W. L., Lee, J. J., Tilghman, S. M., Gumbiner, B. M. & Costantini, F. (1997). The mouse fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell,* **90**(1): 181-192.

Zhao, X. P. & Elder, J. T. (1997). Positional cloning of novel skin-specific genes from the human epidermal differentiation complex. *Genomics,* **45**(2): 250-258.