

**Okinawa Institute of Science and Technology  
Graduate University**

Thesis submitted for the degree of  
Doctor of Philosophy

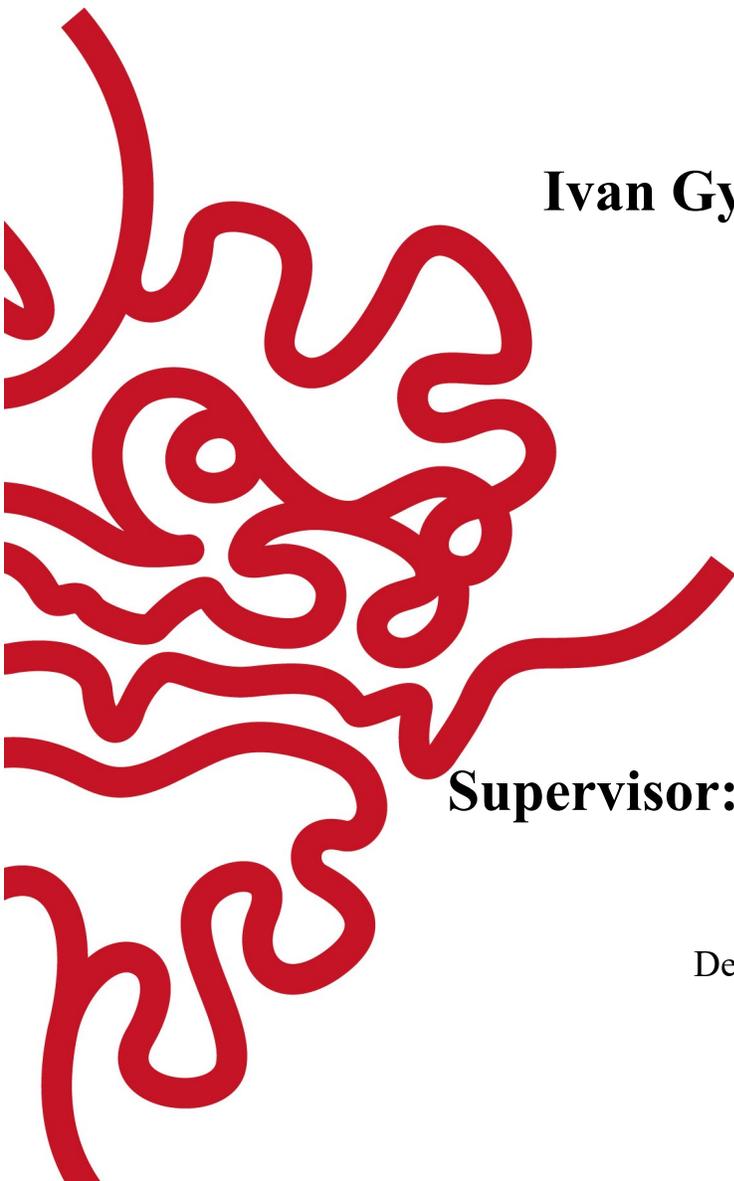
**The Evolution of Dual Functionality of  $\beta$ -catenin  
in Metazoans**

by

**Ivan Gyaviira Mbogo**

**Supervisor: Hiroshi Watanabe**

December 2021



# Declaration of Original and Sole Authorship

I, Ivan Gyaviira Mbogo, declare that this thesis entitled "**The Evolution of Dual Functionality of  $\beta$ -catenin in Metazoans**" and the data presented in it are original and my own work.

I confirm that:

- This work was done solely while a candidate for the research degree at the Okinawa Institute of Science and Technology Graduate University, Japan.
- No part of this work has previously been submitted for a degree at this or any other university.
- References to the work of others have been clearly attributed. Quotations from the work of others have been clearly indicated and attributed to them.
- In cases where others have contributed to part of this work, such contribution has been clearly acknowledged and distinguished from my own work.
- None of this work has been previously published elsewhere.

Signature: 

Date: 01/December/2021

## Abstract

The evolution of the multicellular body animals from unicellular organisms is still a significant and long-lasting subject of interest in biology. Acquisition of cell-cell adhesion with cadherin,  $\alpha$ - and  $\beta$ -catenin proteins is thought to be tightly coupled with the origin of animal epithelium and consequent evolutionary thrive of animals. On the other hand, much research has shown, in a wide range of animal lineages such as bilaterians and cnidarians, that  $\beta$ -catenin associates with diverse intracellular proteins involved in gene transcription/translation and plays an essential role in the induction of the signalling centre (organiser) during animal embryogenesis. The pleiotropic and evolutionary conserved functions of  $\beta$ -catenin suggest deep evolutionary roots of the  $\beta$ -catenin complexes and involvement in the emergence of basic animal body plan. Recent progress in genomics has identified genes of the cell-cell adhesion complex and signalling machinery of  $\beta$ -catenin in genomes of early-branching animals, including *Porifera* (sponges) and *Ctenophora* (comb jellies). However, due to difficulties in applying molecular and genetic technologies in these non-model animals, the ancestral functions of  $\beta$ -catenin complexes remain largely to be explored. In this study, I combined structural, proteomic, and functional approaches to understand evolutionarily conserved features of the  $\beta$ -catenin and its associated proteins. Structural analysis suggests a unicellular origin of the basic architecture of  $\beta$ -catenin protein, while amino acid residues critical in adhesive properties are conserved only within animals. To analyse evolutionarily conserved functional characteristics of basal animal  $\beta$ -catenins, I performed transphylectic studies where the basal animal  $\beta$ -catenins are expressed in *Xenopus* embryos. A series of proteomics analyses of  $\beta$ -catenin-associated proteins revealed the cadherin catenin complex's deep origin and evolutionary conservation. The transphylectic function studies and detailed sequence analysis also revealed the  $\beta$ -catenin's organiser-inducing function of *Cnidaria*, *Porifera*, but not *Ctenophora*. These data suggest that the primary function of ancestral  $\beta$ -catenin was to play adhesive roles, and its' signalling properties were equipped later during the evolution of basal animals.

## Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Hiroshi Watanabe, for the guidance and encouragement he has accorded me during this project. When I encountered the steepest learning curve in the early stages and felt that the project was big, his patience, understanding, and words of wisdom kept me going. Furthermore, I would like to thank Prof. Kentaro Tomii and Dr. Yuko Tsuchiya, both at the National Institute of Advanced Industrial Science, for their expertise in protein model prediction. I would also like to thank the members of the Evolutionary Neurobiology Unit for their support. In particular, Drs. Eisuke Hayakawa for introducing me to the world of proteomics, Chihiro Kawano for the technical assistance in mass spectrometry, Ryotaro Nakamura and Kurato Mohri for the thoughtful discussions. I would also like to thank Dr. Yuuri Yasuoka, at the RIKEN Center for Integrative Medical Sciences, for the training in carrying out functional studies using the *Xenopus* model, plus the thoughtful discussions and suggestions. I would like to thank Alejandro Villar-Briones for his patience when training me on using a mass spectrometer.

To my committee members, I would like to thank Profs. Nori Satoh and Ichiro Masai for their insights and suggestions, along with the training I got when I did rotations in their respective lab. To the Animal Resources Section, thank you for taking care of my frogs. I know it was not easy, and for this, I will always be grateful. To my friends, thanks for sharing the highs and lows of my projects. In particular, I would like to point out that Makoto Tokoro Schreiber, Kamila Mustafina, Larisa Sheloukhova have listened to every detail of my project since its inception. To my mind and stomach support team, I would like to thank Jalena Katic, Swathy Babu, Manana Kutsia and Mai Omar Ahmed for the advice, discussions and food whenever needed. To the Sanshin club, Matsushima-sensei, Tomomi Okubo, Hong Huat Hoh, and many others, thanks for moments that got me away from research and making my life in Okinawa to be more than just a PhD. To Alan and Kiyou Obace, thanks for welcoming me into your household and treating me like family.

I would like to thank the OIST Graduate School office for the immense work you have done during the ups and downs of this PhD and the financial support.

Finally, to my family, I would like to extend my appreciation to my father and siblings for being there whenever I needed you. To my dearest mother, words cannot express how I appreciate the support you have always forwarded me. This journey would have been impossible without you. All I can say to you is “*Webale nyo webalege. Nsimye nyo nsimidde dala*”.

## **Dedication**

This thesis is dedicated to the late Dr. Haruka Morikawa. I will forever be grateful for the constant encouragement during the difficult times of my research. Your presence made my PhD experience and life in Okinawa worthwhile and memorable. I will always appreciate your patience and understanding. Your memories, lessons and sense of humour will always stay with me. You are dearly missed.

## Abbreviations

ACN	Acetonitrile
AMBIC	Ammonium bicarbonate
ANOVA	Analysis of variance
AP	Anterior-Posterior
APC	Adenomatous Polyposis Coli
BLAST	Basic Local Alignment Search Tool
CBD	Catenin Binding Domain
CCC	Cadherin Catenin Complex
CTD	C-terminal domain
CTNND1	Catenin Delta 1
DTT	Dithiothreitol
DV	Dorsal-Ventral
FDR	False Discovery Rate
GBM	Groove Binding Motif
GSK3 $\beta$	Glycogen Synthase Kinase 3 Beta
HPLC	High-Pressure Liquid Chromatography
IAA	Indole-3-acetic acid
IG	In-gel trypsin digestion method
IP-MS	Immunoprecipitation with mass spectrometry
IS	In-solution digestion method
JUP	Junction plakoglobin
KPNB1	Karyopherin (Importin) Beta 1
LC-S/MS	Liquid chromatography with tandem mass spectrometry
MEME	Multiple Expectation maximizations for Motif Elicitation
NTD	N-terminal domain
PPIs	Protein-Protein Interactions
PSM	Peptide Spectrum Match
RPM	Rotations Per Minute
SDC	sodium deoxycholate
TCF	T-cell Factor
TFA	Trifluoroacetic acid

## TABLE OF CONTENTS

Abstract.....	iii
Acknowledgements.....	iv
Dedication.....	v
Abbreviations.....	vi
List of Figures.....	x
List of Tables.....	xi
<b>Chapter 1 Introduction.....</b>	<b>2</b>
1.1 On the origin of metazoans.....	2
1.1.1 <i>Bilateria</i> .....	3
1.1.2 <i>Cnidaria</i> .....	4
1.1.3 <i>Ctenophora</i> .....	4
1.1.4 <i>Placozoa</i> .....	4
1.1.5 <i>Porifera</i> .....	6
1.2 Canonical Wnt/ $\beta$ -catenin signalling.....	6
1.2.1 Wnt “OFF” state.....	6
1.2.2 Wnt “ON” state.....	7
1.2.3 Wnt signalling in the cytosol.....	7
1.2.4 Wnt signalling in the nucleus.....	7
1.3 $\beta$ -catenin.....	9
1.3.1 The bifunctional role of $\beta$ -catenin.....	10
1.3.2 Approach to solving remaining issues.....	12
1.3.3 Novelty of approach.....	12
<b>Chapter 2 Structural analysis of metazoan <math>\beta</math>-catenin evolution.....</b>	<b>15</b>
2.1 Summary.....	15
2.2 Background.....	15
2.3 Methods.....	16
2.3.1 Structural modelling of $\beta$ -catenin proteins.....	16
2.3.2 Conservation of crucial domains in metazoan $\beta$ -catenin.....	17
2.4 Results.....	17
2.4.1 Domain conservation.....	17
2.4.2 Homology models & structure validation.....	18
2.5 Discussion.....	20
2.5.1 Structural conservation of metazoan $\beta$ -catenin.....	20
2.5.2 Premetazoan origin of GSK3 $\beta$ phosphorylation site.....	21

2.5.3 Domains of the N- and C-terminal.....	22
2.6 Chapter conclusion.....	22
<b>Chapter 3 Proteomic approach to identify <math>\beta</math>-catenin protein interactions conserved in metazoans</b> .....	<b>24</b>
3.1 Summary.....	24
3.2 Background.....	24
3.3 Methods.....	25
3.3.1 Total RNA extraction.....	25
3.3.2 cDNA synthesis .....	26
3.3.3 Polymerase Chain Reaction (PCR).....	26
3.3.4 Restriction digestion .....	27
3.3.5 Synthesis of capped mRNA for microinjection.....	27
3.3.6 Animals.....	28
3.3.7 Ovulation induction in <i>X. laevis</i> females .....	28
3.3.8 Testes isolation from male <i>X. laevis</i> for <i>in vitro</i> fertilisation.....	28
3.3.9 Egg collection and <i>in vitro</i> fertilisation.....	28
3.3.10 $\beta$ -catenin mRNA microinjection.....	28
3.3.11 Protein extraction.....	29
3.3.12 Western blotting.....	29
3.3.13 Immunoprecipitation of flag-tagged $\beta$ -catenin.....	29
3.3.14 In-solution (IS) trypsin digestion .....	29
3.3.15 In-gel (IG) trypsin digestion .....	30
3.3.16 LC-MS/MS .....	30
3.3.17 Mass spectrometry data analysis.....	31
3.4 Results.....	31
3.4.1 Adjustment of expression level and immunoprecipitation of $\beta$ -catenin proteins.....	31
3.4.2 Proteomic characterisation of $\beta$ -catenin interactome.....	33
3.4.3 Evolutionarily conserved $\beta$ -catenin protein interactions in metazoans.....	34
3.4.4 $\beta$ -catenin protein interactions shared between <i>Bilateria</i> and <i>Cnidaria</i> .....	38
3.5 Discussion.....	40
3.5.1 The evolutionarily conserved metazoan $\beta$ -catenin machinery.....	40
3.5.2 Novel genes drove morphological and functional novelty in bilaterians.....	41
3.5.3 Challenges of study.....	41
3.6 Future improvements & chapter conclusion .....	41
<b>Chapter 4 Conservation of the cadherin-catenin interaction complex .....</b>	<b>44</b>

4.1 Summary .....	44
4.2 Background .....	44
4.3 Methods.....	45
4.4 Results.....	45
4.4.1 Evolutionary conservation of $\beta$ -catenin interaction with $\alpha$ E-catenin.....	45
4.4.2 Evolutionary conservation of $\beta$ -catenin interaction with E-cadherin in metazoans.....	47
4.4.3 Other adherens junctions associated proteins .....	48
4.5 Discussion .....	49
4.5.1 Premetazoan origin of $\beta$ -catenin's structural function .....	50
4.6 Chapter conclusion.....	52
<b>Chapter 5 Conservation of organiser inducing function by metazoan <math>\beta</math>-catenin.....</b>	<b>54</b>
5.1 Summary .....	54
5.2 Background .....	54
5.3 Methods.....	55
5.3.1 Secondary axis assay.....	55
5.3.2 X-gal fixing and staining.....	56
5.3.3 Luciferase reporter assay .....	56
5.4 Results.....	56
5.4.1 TCF- $\beta$ -catenin binding sites conserved in all metazoans but <i>Ctenophora</i> .....	56
5.4.2 Coprecipitation of canonical Wnt signalling elements with $\beta$ -catenin .....	57
5.4.3 Secondary axis induction assay.....	59
5.4.4 Metazoan canonical Wnt signalling activity induction .....	60
5.5 Discussion .....	60
5.5.1 Dorsal organizer inducing capacity.....	61
5.5.2 Formation of ectopic tail-like projections upon overexpression of ctenophore $\beta$ -catenin .....	63
5.6 Chapter conclusion.....	64
<b>Chapter 6 General discussion .....</b>	<b>66</b>
6.1 Conclusion .....	68
References.....	69
Appendices.....	85

## List of Figures

Figure 1.1 The origins of multicellularity.....	3
Figure 1.2. Cut through of placozoa showing the various cell types.....	5
Figure 1.3. Canonical Wnt signalling pathway.....	8
Figure 1.4. Structural domains of a typical bilaterian $\beta$ -catenin.....	9
Figure 1.5. $\beta$ -catenin's role in cell adhesion.....	11
Figure 1.6. Study design to understand the evolution of $\beta$ -catenin functionality.....	13
Figure 2.1 Conservation of crucial domains of $\beta$ -catenin.....	18
Figure 2.2. Ramachandran plot analysis.....	19
Figure 2.3. 3D architecture of metazoan $\beta$ -catenin proteins are highly conserved.....	20
Figure 2.4. Conservation of central region of metazoan $\beta$ -catenin.....	21
Figure 3.1. Identification of conserved $\beta$ -catenin PPIs.....	25
Figure 3.2. Processing and consensus workflows used in label-free quantification and identification of peptides and proteins.....	31
Figure 3.3. Optimisation of $\beta$ -catenin mRNA for microinjection..	32
Figure 3.4. Immunoprecipitation of flag-tagged metazoan $\beta$ -catenin..	33
Figure 3.5. Venn diagrams of common proteins interacting with metazoan $\beta$ -catenins.....	34
Figure 3.6. Evolutionarily conserved $\beta$ -catenin protein interactions.....	35
Figure 4.1. Validation of co-immunoprecipitation of $\alpha$ E-catenin with $\beta$ -catenin..	46
Figure 4.2. The $\alpha$ E-catenin binding region in $\beta$ -catenin.....	47
Figure 4.3. The two lysine residues that are critical in binding E-cadherin..	48
Figure 4.4. The metazoan groove binding motif (GBM) in E-cadherins.....	50
Figure 4.5. Evolution of $\beta$ -catenin role structural role.....	51
Figure 5.1. Secondary axis induction in <i>X. laevis</i> .....	54
Figure 5.2. Method to confirm conservation of metazoan $\beta$ -catenin's role in organizer induction.. .....	56
Figure 5.3. $\beta$ -catenin residues critical in binding TCF..	57
Figure 5.4. Secondary body axis induction assay.....	59
Figure 5.5. Canonical Wnt signalling activity for different metazoan $\beta$ -catenin.....	60
Figure 5.6. Evolution of canonical Wnt/ $\beta$ -catenin's signalling capacity. ....	64
Figure 6.1. The evolution of the dual functionality of $\beta$ -catenin.....	68

## List of Tables

Table 2.1 Percentage amino acid sequence similarity between different metazoan $\beta$ -catenins. ..	17
Table 2.2 Ramachandran plot calculation for 3D models of metazoan $\beta$ -catenin computed with PROCHECK. ....	19
Table 3.1. PCR reaction setup.....	26
Table 3.2. PCR conditions. ....	26
Table 3.3 Primers used in cloning $\beta$ -catenin.....	27
Table 3.4. Restriction digest enzymes used in plasmid insertion and corresponding length of the insert. ....	27
Table 3.5. IG method increased coverage and the number of unique peptides for better protein identification. ....	34
Table 3.6. Mitochondria localised $\beta$ -catenin protein interactions conserved in 4 different metazoans following $\beta$ -catenin mRNA expression in <i>X. laevis</i> embryos. ....	35
Table 3.7. Nucleus localised $\beta$ -catenin protein interactions conserved in 4 different metazoans following $\beta$ -catenin mRNA expression in <i>X. laevis</i> embryos. ....	36
Table 3.8. Plasma membrane localised $\beta$ -catenin protein interactions conserved in 4 different metazoans following $\beta$ -catenin mRNA expression in <i>X. laevis</i> embryos. ....	37
Table 3.9. Cytoskeleton associated $\beta$ -catenin protein interactions conserved in 4 different metazoans following $\beta$ -catenin mRNA expression in <i>X. laevis</i> embryos.. ....	38
Table 3.10. ER localised $\beta$ -catenin protein interactions conserved in 4 different metazoans following $\beta$ -catenin mRNA expression in <i>X. laevis</i> embryos. ....	38
Table 3.11. $\beta$ -catenin protein interactions that are shared between <i>Bilateria</i> and <i>Cnidaria</i> $\beta$ -catenin.....	39
Table 4.1. Coimmunoprecipitation of endogenous $\alpha$ E-catenin by metazoan $\beta$ -catenin. ....	45
Table 4.2. Coimmunoprecipitation of endogenous E-cadherin. ....	47
Table 4.3. Other proteins that are associated with the adherens junctions. ....	49
Table 5.1. Coimmunoprecipitation of endogenous TCF by metazoan $\beta$ -catenin.. ....	58
Table 5.2. Coimmunoprecipitation of endogenous APC by metazoan $\beta$ -catenin.....	58

# CHAPTER 1

---

“Nothing in biology makes sense except in the light of evolution”  
-Theodosius Dobhansky

---

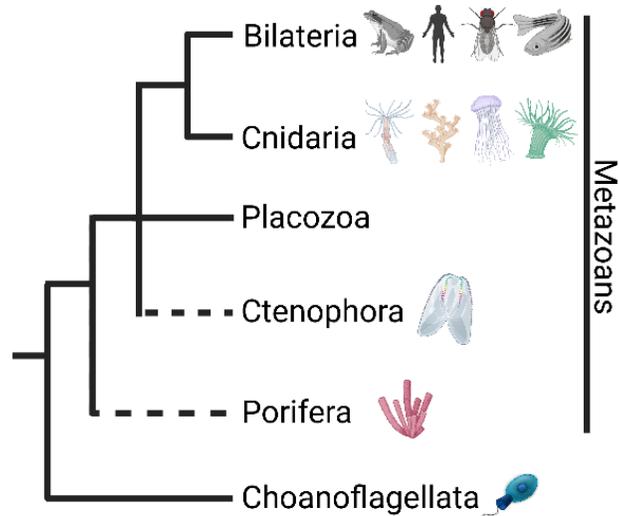
## Chapter 1 Introduction

In sexual organisms, embryonic development is initiated when the sperm fuses with the egg to form a zygote. From this moment onwards, several signalling cascades are required to finetune gene expression, allowing proper development. These signalling molecules include Fibroblast Growth Factor (Fgf) (Neugebauer *et al.*, 2009), Hedgehog (Briscoe and Thérond, 2013), Notch (Shi and Stanley, 2006), Bone Morphogenetic Protein (Bmp) (Wang *et al.*, 2014), Receptor Tyrosine Kinase (Rtk) (Sopko and Perrimon, 2013) and Wnt signalling (Clevers, 2006). These signalling genes are conserved in many animals, including basal metazoans, signifying their importance in organismal growth and development. Among these, I focused on the canonical Wnt signalling molecule,  $\beta$ -catenin, on which numerous research has shown its essential roles in establishing the primary body axis, induction of signalling centre, gastrulation and mediation of cell-cell adhesion (Haegel *et al.*, 1995; Röttinger *et al.*, 2012). In this thesis study, I investigated the conservation of structure, binding proteins, and developmental function of basal metazoan  $\beta$ -catenins.

### 1.1 On the origin of metazoans

Although numerous studies have been carried out on it for many years, understanding the origins and eventual diversification of multicellular organisms remains a subject of interest among evolutionary biologists. In understanding the origins of multicellularity, it is a widely accepted theory that choanoflagellates are the sister group of metazoans hence share a common ancestor. Indeed, choanoflagellates share a marked morphological similarity to poriferans' choanocytes; hence are believed to be the closest living relative to metazoans (King *et al.*, 2008). Choanoflagellates are unicellular eukaryotes with an ovoid shape surrounded by a fine collar (choano) and a flagellum used during locomotion and bacteria collection during feeding (Brunet and King, 2017). Although fundamentally unicellular, choanoflagellate, *Salpingoeca rosetta* has been shown to form multicellular-like colonies, suggesting that the last common ancestor of animals and choanoflagellates was possibly capable of “simple multicellularity” (Fairclough *et al.*, 2010; Dayel *et al.*, 2011). Furthermore, analysis of basal multicellular organisms' transcriptomes revealed that the evolution from unicellularity to multicellularity was matched with an increase in novel core genes like *TGF-beta* and *Wnt* (for further discussion, see below) (Paps and Holland, 2018).

Metazoans are multicellular organisms that, despite differences in body shape, are organised along the primary body axis, i.e., anterior-posterior (AP) axis in *Bilateria*, which is generally associated with locomotion with the mouth the anterior and the anus located at the posterior. In addition, the metazoan body plan frequently exhibits a dorsal-ventral (DV) axis, which is vital in the establishment of the organiser (a group of cells with the capacity to control the developmental trajectory of cells), which is the site of gastrulation. Using these axes, the five metazoan phyla, i.e., *Bilateria*, *Cnidaria*, *Placozoa*, *Ctenophora*, and *Porifera* (**Figure 1.1**), can be grouped into one of three symmetries, i.e., bilateral (*Bilateria*), radial (*Cnidaria*, *Ctenophora*, & *Porifera*), and asymmetrical (*Placozoa*).



**Figure 1.1 The origins of multicellularity.** The phylogenetic positions of metazoans with *Choanoflagellata* as the outgroup. The subject of the most basal metazoans is still a matter of debate, with some research suggesting *Porifera* while other *Ctenophora*. Created using Biorender.com.

### 1.1.1 Bilateria

This phylum encompasses the majority of metazoan organisms. Animals here are characterised by two orthogonal body axes (AP and DV axes), and in many cases, asymmetric patterning along the left-right axis. They also have three germ layers; hence are known as triploblastic. In the adult stage, the ectoderm gives rise to the nervous system and epidermal cells. The mesoderm gives rise to the muscle cells and connective tissue, while the internal lining of various organ systems such as digestive, skeletal, etc., is derived from the endoderm. Besides its role in axis polarity formation,  $\beta$ -catenin is also critical in germ layer formation in bilaterians. It was shown in the *Xenopus* embryos that  $\beta$ -catenin was crucial in early mesoderm induction (Schohl and Fagotto, 2003). Furthermore,  $\beta$ -catenin is also vital in forming the three germ layers in the ascidian embryo (Hudson *et al.*, 2013).

In its structural role,  $\beta$ -catenin is vital in cell adhesion. Cell-cell adhesion is crucial during embryo development and in the adult stage, where it facilitates tissue organisation. Various specialised adhesive functions in metazoan epithelial cells have been identified, and their roles are not limited to cell-cell adhesion but also intercellular communication and act as a barrier preventing molecules from passing between them (Farquhar and Palade, 1963; Miller *et al.*, 2013). Cell-cell adhesions that have been identified in bilaterians include adherens, tight, gap, septate junctions and, focal adhesions. Adherens junctions are localised in the apical membrane and composed of classical cadherins (particularly E-cadherin), which links cells through its extracellular cadherin (EC) domains,  $\beta$ -catenin which links with E-cadherin at the catenin binding domain (CBD) and,  $\alpha$ E-catenin which binds to  $\beta$ -catenin and links it to filamentous (F)-actin. Together these three protein components form the cadherin catenin complex (CCC). Other types of cell-cell adhesions are beyond the scope of this thesis but are critical in embryo development and structural integrity. For example, gap junctions permit the transport of small molecules between cells hence are vital in intercellular communication (Sáez *et al.*, 2003). Septate and tight junctions are categorised as occluding junctions because they are believed to serve similar roles, even though their molecular elements and evolutionary origins are different (Magie and Martindale, 2008). Focal adhesions are critical in mediating intracellular links between cells and the extracellular matrix.

### 1.1.2 *Cnidaria*

Cnidarians are the sister group to all bilaterians. They are marine animals with radial symmetry and present with two germ layers; hence, they are termed diploblastic. Most cnidarians belong to the class *Anthozoa*, which includes subclasses *Hexacorallia* (sea anemones and stony corals) and *Octocorallia* (sea pens). Other cnidarians fall under subphylum *Medusazoa*, which contains four classes, i.e., *Cubozoa* (box jellyfish), *Hydrozoa* (*Hydra*), *Scyphozoa* (moon jellyfish), *Staurozoa* (stalked jellyfish). Although cnidarians vary in shape, size, and lifestyle, they use stinging cells called cnidocytes to catch prey, which are characteristic of this phylum (Babonis and Martindale, 2014). Like in bilaterians, cnidarians bear distinct epithelial layers, ectoderm and endoderm-like inner layer. However, the nature of the inner layer is still a subject of debate. Unlike bilaterians, cnidarians do not exhibit a distinct mesoderm but rather a bifunctional layer referred to as endomesoderm or gastrodermis (Röttinger *et al.*, 2012; Salinas-Saavedra *et al.*, 2018). The endomesoderm forms the primitive digestive system and has markers characteristic of both the bilaterian meso- & endoderm (Technau and Scholz, 2003; Martindale *et al.*, 2004). As the sister group of bilaterians, researchers have recently begun relying on cnidarian animal models to understand why these two groups evolved to have different germ layers and what signals are critical in the specification of each layer.

Like in bilaterians, similar adhesive functions have been identified in *Cnidaria*, with a few exceptions. Sequence homology has confirmed the presence of elements of the adherens junctions, and functional analysis has confirmed their importance in cell-cell adhesion during embryonic development in *Nematostella* (Nathaniel Clarke *et al.*, 2019; Pukhlyakova *et al.*, 2019).

### 1.1.3 *Ctenophora*

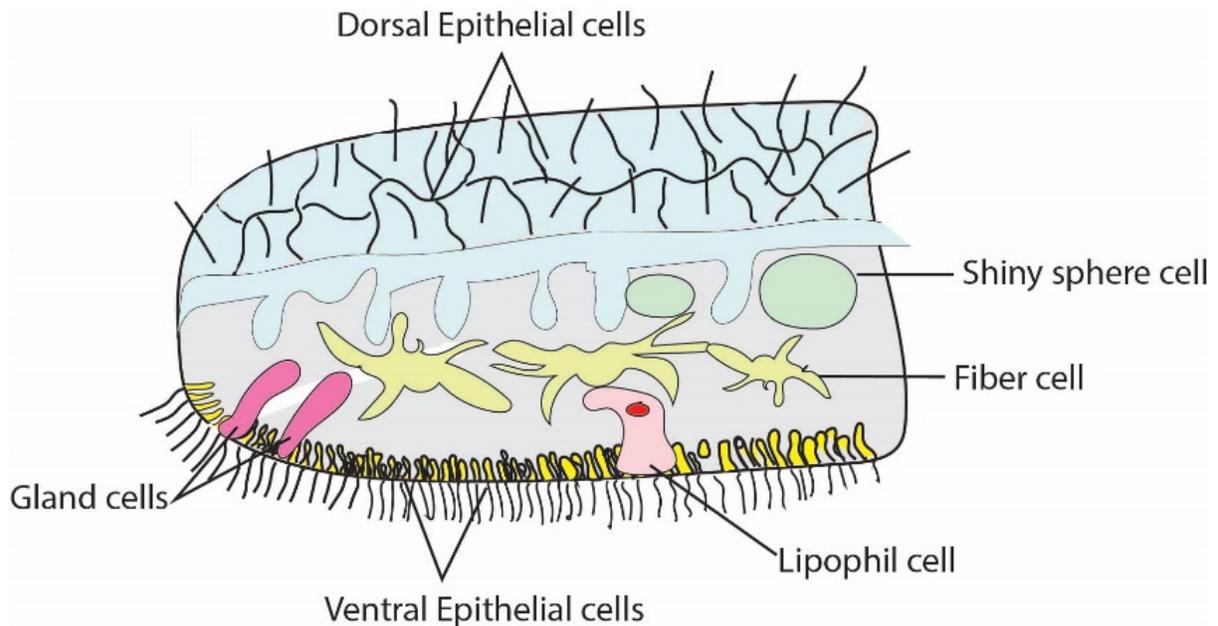
Also known as comb jellies (due to their jelly-like appearance), ctenophores are marine organisms with a modified biradial symmetry. Although there is still doubt about their place in the evolutionary tree of metazoans, some phylogenetic results have suggested that they are one of the earliest emerging metazoan lineages (Podar *et al.*, 2001; Dunn *et al.*, 2008; Li *et al.*, 2021). Their bodies have comb-like projections called ctenes used for locomotion, an apical sensory organ, and a pair of tentacles bearing colloblasts to catch prey (Pang and Martindale, 2008). Like cnidarians, ctenophores are diploblastic animals. However, compared to cnidarians, where striated muscles are found in a few animals, all ctenophores have fully differentiated muscles and mesoderm (Moroz, 2015). Thus, ctenophores would be an excellent model to study the origin of multicellularity and the evolution of germ layer formation.

Regarding cell-cell adhesion, in ctenophores, sequence homology has confirmed the presence of elements of the CCC, which form the adherens junctions (Ryan *et al.*, 2013; Moroz *et al.*, 2014). The presence of these homologs was also confirmed through expression studies (Pang *et al.*, 2010; Salinas-Saavedra *et al.*, 2019). However, there is still a lack of understanding if the adherens junctions elements in ctenophores function in cell-cell adhesion as they do in cnidarians and bilaterians. This is a big enigma since it was observed that *Mnemiopsis leidyi* E-cadherin lacks an evolutionary critical domain in the interaction with  $\beta$ -catenin (Belahbib *et al.*, 2018). Indeed a recent expression analysis found no evidence of  $\beta$ -catenin localisation in cell contacts in *M. leidyi* (Salinas-Saavedra *et al.*, 2019).

### 1.1.4 *Placozoa*

Placozoa is the simplest multicellular organism made up of a single epithelial layer hence have only three anatomical layers, i.e., upper, intermediate, & lower epithelium. Compared to other

metazoans, cell diversity in *Placozoa* is low, with only six cell types observed (Smith *et al.*, 2014a) (**Figure 1.2**). Although the *Placozoa* genome encodes genes involved in DV and AP axial patterning, they have only a DV axis comparable to the oral-aboral axis observed in cnidarians and lack the AP axis (DuBuc *et al.*, 2019). Compared to other metazoans, placozoans lack musculature (hence move by gliding using cilia on the ventral epithelium), a nervous, and digestive system (Schierwater *et al.*, 2009). Lack of a digestive system means that they can only carry out external digestion by secreting digestive enzymes in the proximity of algae at the ventral epithelium (Smith *et al.*, 2015). For a long time, the only known species in this phylum were *Trichoplax adhaerens* and other species, *Hoilungia hongkongensis* (Eitel *et al.*, 2018) & *Polyplacotoma mediterranea* (Osigus *et al.*, 2019) were recently discovered.



**Figure 1.2.** Cut through of placozoa showing the various cell types.

Interestingly, as simple an organism as it is, the *T. adhaerens* genome was found to contain 11,514 protein-coding genes, of which 87% were found in other metazoans, and 83% of genes conserved in sea anemones and bilaterians were found in *T. adhaerens* (Srivastava *et al.*, 2008). It remains unclear what this gene conservation imply for the evolution of cell diversity in metazoans. Furthermore, the phylogenetic position of placozoa with other metazoans is still debatable (Eitel *et al.*, 2013). Evaluation based on 18s rRNA of 52 taxa sequences placed placozoa near the base of bilaterians and maybe its sister group (Collins, 1998). However, further analysis of the mitochondrial genome of *T. adhaerens* placed it as the basal lower metazoan (Dellaporta *et al.*, 2006). Thus, several questions remain unanswered; hence more research needs to be carried out using placozoan models.

Like other metazoans, the *T. adhaerens* genome bears elements of the adherens junctions (Srivastava *et al.*, 2008). Recently, adherens junctions in *T. adhaerens* were visualised, and it was concluded that they are vital in modulating diffusion between epithelial cells (Smith and Reese, 2016). The presence or absence of other cell-cell adhesions is still a subject of study. So far, gap junctions have not been identified (Smith *et al.*, 2014b). Furthermore, both tight and septate junctions in *T. adhaerens* are absent (Ganot *et al.*, 2015; Smith and Reese, 2016).

### 1.1.5 *Porifera*

By virtue of the presence of feeding cells called choanocytes that share a morphological similarity to choanoflagellates, poriferans (sponges) are considered the most basal metazoan phylum (King *et al.*, 2008). Poriferans are aquatic benthic multicellular animals that are still a significant subject of study among evolutionary developmental biologists. Poriferans undergo gastrulation forming a diploblastic ciliated embryo (Nakanishi *et al.*, 2014). Although the homology of germ layers between sponges and eumetazoans is still debatable, a finding of transcription factor, *GATA*, a highly conserved eumetazoan endomesodermal marker in the inner cell layer of *Amphimedon queenslandica*, suggests that the germ layers observed in sponges may be similar to those in other metazoans (Nakanishi *et al.*, 2014).

The nature of cell-cell adhesions in poriferan epithelial cells is still a subject of study. Homologous genes of the adherens junction have been observed in various poriferan genomes (Srivastava *et al.*, 2010; Nichols *et al.*, 2012; Alié *et al.*, 2015). Although the function of the adherens junctions in cell adhesion is yet to be examined, a recent study confirmed the precipitation of adherens junction components with endogenous  $\beta$ -catenin in *Ephydatia muelleri* (Schippers and Nichols, 2018).

From the above, we can observe a large diversity in anatomical and even cellular features among metazoans. Therefore, it is essential to identify factors that brought about the differences between metazoan phyla, i.e., what made each phylum unique. However, it is also essential to identify and study protein functions and interactions conserved throughout evolution. Comparative genome analyses have revealed an unexpected level of conservation of genes involved in embryonic development (Kusserow *et al.*, 2005), despite anatomical diversities of metazoans. One example of such are Wnt genes, which are metazoan specific and are effectors of the Wnt signalling pathway (Holstein *et al.*, 2011).

## 1.2 Canonical Wnt/ $\beta$ -catenin signalling

In addition to an essential role in adherens junctions,  $\beta$ -catenin is known to play crucial regulatory roles in gene expression under the Wnt signalling pathway. Initial discoveries into the Wnt signalling pathway arose from cloning a novel mice proto-oncogene, integration-1 (Int-1) (Nusse and Varmus, 1982). Later studies found that Int-1 was a homolog of wingless (Wg) found in *Drosophila* and is critical during embryonic development, particularly in segmentation (Nüsslein-Volhard and Wieschaus, 1980). Wnt signalling can be divided into three signalling pathways, i.e.,  $\beta$ -catenin dependent (canonical) and  $\beta$ -catenin independent (non-canonical) pathways. The  $\beta$ -catenin dependent pathway is the focus of the current study and will be looked at in greater detail.

### 1.2.1 Wnt “OFF” state

In the absence of Wnt ligands,  $\beta$ -catenin that is not bound to E-cadherin is phosphorylated through the activity of a multiprotein destruction complex resulting in its' subsequent degradation. The destruction complex is composed of Axin, adenoma polyposis coli (APC), glycogen synthase kinase (GSK)  $3\beta$ , casein kinase (CK)  $1\alpha$  and protein phosphatase 2A (PP2A). Axin plays a critical role in the serial phosphorylation of  $\beta$ -catenin at serine 45 by CK $1\alpha$  followed by phosphorylation at threonine 41, serine 37 and 33 by GSK $3\beta$  (Kimelman and Xu, 2006). Axin, therefore, plays a role of the rate limiting factor of the destruction complex (Lee *et al.*, 2003). Phosphorylation at S33 and 37 results in creating an E3 ubiquitin ligase  $\beta$ -TrCP binding site, which ubiquitinates  $\beta$ -catenin and targets it for destruction by 26S proteasome (Hart *et al.*, 1999). The phosphorylation of  $\beta$ -catenin has positive feedback through the interaction of GSK $3\beta$  and CK1 with Axin and APC. It was found that GSK $3\beta$  and CK $1\alpha$  can also phosphorylate Axin and APC, leading to increased

phosphorylation of  $\beta$ -catenin by the two proteins (Kimelman and Xu, 2006; Huang and He, 2008). In the protein complex, PP2A seems to counteract the effects of GSK3 $\beta$  and CK1 $\alpha$  in that if the phosphorylated  $\beta$ -catenin is not associated with APC when it leaves the destruction complex, then PP2A dephosphorylates hence reducing  $\beta$ -catenin levels that are targeted for destruction (Su *et al.*, 2008).

### 1.2.2 Wnt “ON” state

Wnt ligands are a large family of highly hydrophobic cysteine-rich glycoproteins that are highly conserved in metazoans. The initiation of Wnt signalling requires two types of cell membrane receptors, i.e., Frizzled (Fzd) and low-density lipoprotein (LDL) receptor-related protein 5 and 6 (LRP5 and 6). Fzds are seven-transmembrane proteins characterised with a cysteine-rich domain at the N-terminal (Hsieh *et al.*, 1999; Dann *et al.*, 2001). The LRP receptor family was initially discovered in the *Drosophila* homolog arrow (Nüsslein-Volhard and Wieschaus, 1980). Later work in mice and *Xenopus* mutants established arrow and LRP5 and 6 to be homologs (Hey *et al.*, 1998; DiNardo *et al.*, 2000). The LRP receptor family act as co-receptors in Wnt signalling. On the initiation of the signal, dimerisation occurs through the interaction of Fzd with LRP5/6 (Janda *et al.*, 2017). The Wnt-Fzd-LRP5 complex recruits the scaffolding protein Dishevelled (Dvl) via its' PDZ domain (Wong *et al.*, 2003) and oligomerises via its' DIX domain (Schwarz-Romond *et al.*, 2007), resulting in the phosphorylation and activation of LRP6 and the recruitment of the Axin protein complex to the receptors (MacDonald *et al.*, 2009).

### 1.2.3 Wnt signalling in the cytosol

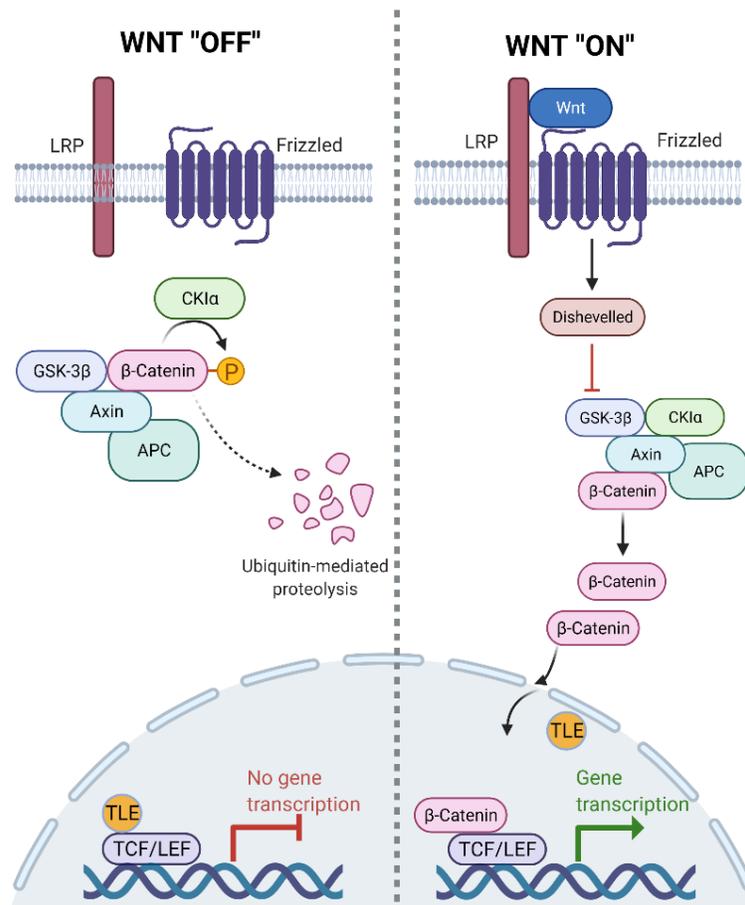
It is understood that the inhibition of GSK3 $\beta$  activity is critical for canonical Wnt signalling to proceed. However, the mechanisms of how this occurs are still not well understood. Two different models have been proposed for how this occurs. In the first model, the binding of Wnt ligands to the co-receptors results in the phosphorylation of the PPPSPxS motifs in the cytosolic region of LRP5/6 hence inhibiting the interaction of GSK3 $\beta$  to Axin (Wu *et al.*, 2009). This leads to a buildup in  $\beta$ -catenin levels as phosphorylation is reduced. However, an alternative model suggests that multivesicular endosomes into which GSK3 $\beta$  is sequestered are key in Wnt signal-transduction (Taelman *et al.*, 2010). Following accumulation in the cytosol,  $\beta$ -catenin is shuttled to the nucleus.

### 1.2.4 Wnt signalling in the nucleus

When canonical Wnt signalling is activated,  $\beta$ -catenin accumulates in the cytoplasm and translocates to the nucleus. The method of translocation is still a subject of research.  $\beta$ -catenin controls the transcription activity of canonical Wnt target genes, hence necessitating its' binding to DNA. However,  $\beta$ -catenin lacks a DNA binding domain; therefore, it requires binding partners to bring it in contact with the promoters of target genes (Huber *et al.*, 1997). The transcription factors T-cell factor (TCF) and lymphoid enhancer-binding protein (LEF) (from here on referred to as TCF) are critical in the repression and activation of gene targets for canonical Wnt signalling. Invertebrates carry a single TCF ortholog, while vertebrates have four TCF paralogs (five in Zebrafish) (Arce *et al.*, 2006). TCF homologs have also been identified in non-bilaterian metazoans but not in choanoflagellates suggesting that the signalling role of  $\beta$ -catenin was inherently a metazoan innovation.

Knowledge of TCF's role in canonical Wnt signalling arose from yeast two-hybrid screens that showed  $\beta$ -catenin binds strongly to TCF with a conserved motif in the N-terminal of TCF being

observed in both invertebrates and vertebrates (Behrens *et al.*, 1996; Huber *et al.*, 1996). Deletion of this conserved motif inhibited the induction of secondary body axis in *Xenopus* embryos with overexpressed  $\beta$ -catenin (Molenaar *et al.*, 1996). TCFs utilise their high mobility group (HMG) domain to bind to a specific consensus sequence called the Wnt responsive element (WRE) CCTTGATWW (W is either T or A) in the DNA's minor groove (Valenta *et al.*, 2012). In the absence of  $\beta$ -catenin in the nucleus, TCFs repress transcription by forming a complex with Groucho/TLE resulting in histone deacetylation & chromatin stabilisation. The repressive role of TCFs is reversed to activation in the presence of  $\beta$ -catenin, which displaces Groucho (Daniels and Weis, 2005) (**Figure 1.3**). However, the binding of  $\beta$ -catenin to TCF can also repress transcription. Three mechanisms have been suggested about how this possibly happens. The first mechanism described as competitive repression involves TCF/ $\beta$ -catenin displacing or inhibiting other DNA binding transcription activators (Piepenburg *et al.*, 2000; Daniels and Weis, 2005). The second mechanism is direct repression, where the TCF/ $\beta$ -catenin complex binds to the WREs and recruits corepressors (Jamora *et al.*, 2003; Theisen *et al.*, 2007). The final mechanism involves an AGAWAW consensus TCF binding sequence which in *Drosophila* has been shown to induce TCF/ $\beta$ -catenin mediated repression (Blauwkamp *et al.*, 2008).



**Figure 1.3. Canonical Wnt signalling pathway.** During Wnt “off”,  $\beta$ -catenin phosphorylated in the cytoplasm and targeted for degradation hence gene transcription is inhibited by TLE/ Groucho’s interaction with TCF. On the contrary, during Wnt “on”, the  $\beta$ -catenin degradation complex is inhibited, resulting in  $\beta$ -catenin accumulation in the cytoplasm and translocation to the nucleus. Here it displaces TLE resulting in gene transcription of canonical Wnt signalling target genes (Created with Biorender.com).

The failure to degrade  $\beta$ -catenin resulting in its' accumulation in the nucleus and uncontrolled gene activation is a characteristic of many cancers; hence a finetuned control of the Wnt signalling pathway is crucial (Zhan *et al.*, 2017). Therefore, the list of genes that are targeted by the Wnt/ $\beta$ -catenin signalling pathway is ever-growing. However, these genes require context-dependent regulation depending on spatial-temporal localisation.

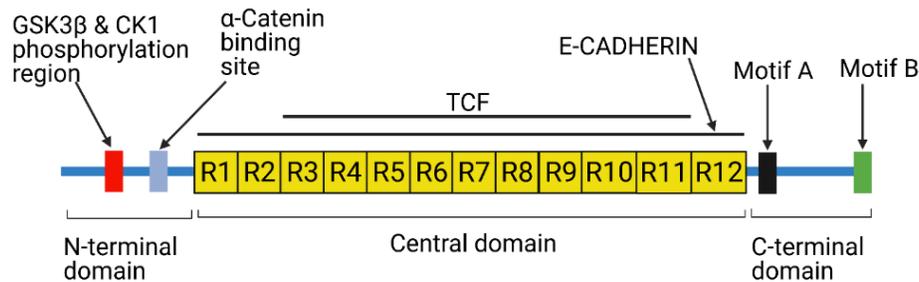
It is essential to understand whether  $\beta$ -catenin's transcription regulatory role was present at the emergence of metazoans. For now, most answers emanate from the use of bilaterian models. By a transphyletic approach, this research sought to understand how and if  $\beta$ -catenin's transcription regulatory role was present in other non-metazoans.

### 1.3 $\beta$ -catenin

The catenin family consists of three groups, i.e., p120, alpha and beta subfamily.  $\beta$ -catenin is the founding member of the beta subfamily (Zhao *et al.*, 2011). Its presence in all metazoans is an indicator of its importance in the emergence of multicellularity. Although there are lineage-specific domains, it is primarily composed of three main domains, i.e., the central domain that is characterised by 12 armadillo repeats (R1-R12) and the flanking N- and C-terminal domains (NTD and CTD, respectively) (**Figure 1.4**).

In the central domain, each armadillo repeat comprises approximately 42 amino acids that form three alpha-helices arranged together in a compact superhelix (Huber *et al.*, 1997). The central domain is the leading interaction site with several key transcription factors in Wnt signalling, e.g. TCF interacts via R3-R10 (Graham *et al.*, 2000).

The binding sites for GSK3 $\beta$  and CK1 $\alpha$  are located within the NTD where mutations in this region, specifically S45 & T41, often appear in malignant tumours and S37 & S33, widespread in benign entities (Provost *et al.*, 2003). These sites are crucial in  $\beta$ -catenin's phosphorylation and subsequent degradation. In addition, the  $\alpha$ E-catenin binding domain is also localised in the NTD, and as will be explained later, this site is crucial in the cell-cell adhesion function of  $\beta$ -catenin.



**Figure 1.4. Structural domains of a typical bilaterian  $\beta$ -catenin.** R1-R12, armadillo repeats. Binding sites of proteins involved cell-cell adhesion and transcriptional regulation. Created using Biorender.com.

The CTD can further be divided into motif A and B, which are highly conserved in bilaterians apart from *Caenorhabditis elegans* (Schneider *et al.*, 2003). Just adjacent to R12 is motif A, which is involved in chromatin remodelling and gene transcription as it is a binding site for protein p300/CBP (Hecht *et al.*, 2000) and transcription factor Teashirt (Gallet *et al.*, 1999). Further evidence for the importance of motif A was proven in armadillo mutants (point and deletion mutations), which resulted in interference of cell adhesion function but not Wnt signalling and vice versa (Orsulic and Peifer, 1996). Motif B is localised at the tip of the C-terminal and contains a characteristic PDZ-like consensus sequence (DTDLD). Several researchers have indicated the possible role of this DTDLD sequence motif in binding other proteins to  $\beta$ -catenin, e.g. LIN7 (Perego *et al.*, 2000) and TIP1 (Kanamori *et al.*, 2003).

Identifying domain similarities and differences between different phyla is essential as it could lay a foundation for understanding the functional evolution of  $\beta$ -catenin. To this end, our research carried out an in-depth metazoan comprehensive identification of domain conservations and differences between phyla—the effect of these domain differences reflected in our proteomic and functional analysis results.

### 1.3.1 The bifunctional role of $\beta$ -catenin

Through sequence homology, it can be concluded that  $\beta$ -catenin is an evolutionarily conserved protein present in all metazoans. In *Ctenophora*, *Placozoa*, *Porifera*, and *Cnidaria*,  $\beta$ -catenin protein homologs have been identified, and these bear similarities to those of *Bilateria*. Its high conservation between different phyla is an indication of its importance in the emergence of metazoans.  $\beta$ -catenin functions are categorised into either structural or signalling, i.e., cell adhesion at the adherens junction and transcription regulation in the nucleus. There is still debate on which of the two functions emerged first, but a recent study suggested that both roles were present in poriferan, *Ephydatia muelleri* (Schippers and Nichols, 2018). In contrast, a different study suggests that only the transcription regulatory role and not cell adhesion was present in *Ctenophora* (Salinas-Saavedra *et al.*, 2019).

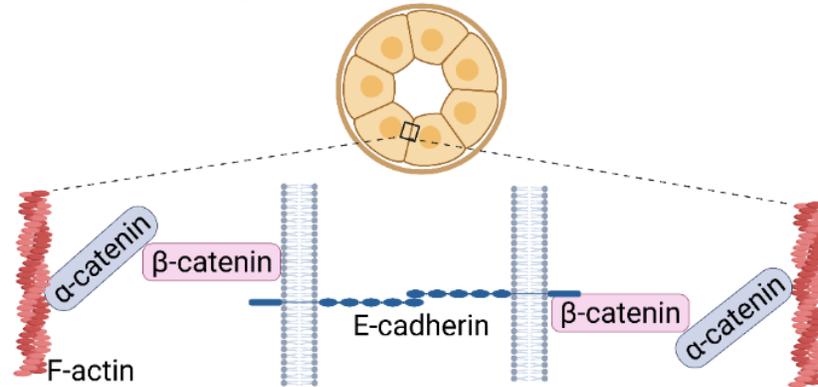
#### Role of $\beta$ -catenin in the cell-cell adhesion

A simple epithelium is characteristic of metazoans and is critical for coordinated multicellularity, i.e., division & migration. The emergence of multicellularity required the innovation of cell-cell adhesion machinery that could maintain the epithelia's structural integrity (Abedin and King, 2010). The importance of cell-cell adhesion was demonstrated in knockdown experiments, which resulted in developmental defects & or arrests and cancer (Halbleib and Nelson, 2006; Michalina *et al.*, 2020). Research has shown that classical cadherins, particularly transmembrane E-cadherins, are vital in developing epithelial cell-cell adhesion in metazoans. Cell-cell adhesion is mediated at the adherens junction where the E-cadherin extracellular domains of two cells link. The E-cadherin cytoplasmic domain through a conserved domain binds  $\beta$ -catenin, which in turn binds alpha ( $\alpha$ )-E-catenin, forming the cadherin catenin complex (CCC) (**Figure 1.5**).  $\alpha$ E-catenin, in turn, binds filamentous actin, linking the adherens junction complex to the cytoskeleton resulting in maintenance of cell shape, strengthening of cell-cell contact and polarised vesicle trafficking (Adams and Nelson, 1998; Mellman and Nelson, 2008).

In bilaterians, extensive research using different models has been carried to understand the role of  $\beta$ -catenin and other adherens junctions' elements in cell-cell adhesion with mutations leading to failure in gastrulation and developmental arrests. Early research demonstrated that E-cadherin null mutant mouse embryos could not form a trophectodermal epithelium, highlighting E-cadherin's vital role in early morphogenetic events (Larue *et al.*, 1994). Furthermore,  $\beta$ -catenin knockout has been demonstrated to affect mouse development at the gastrulation stage (Haegel *et al.*, 1995). Meanwhile, the depletion of  $\alpha$ E-catenin in *Xenopus* embryos has been shown to affect intercellular adhesion at the blastula stage (Kofron *et al.*, 1997). In human disorders, disruption of normal expression of adherens junctions elements has been associated with epithelial cancers (Benjamin and Nelson, 2008; Jeanes *et al.*, 2008). All this evidence suggests the general importance of the CCC in cell-cell adhesion among bilaterian animals. However, it remains obscure when and in what form of CCC became functional during metazoan evolution. Recently, researchers have started to investigate this question by using basal metazoans.

In basal metazoans, little functional analysis has been carried to investigate the role of the CCC in

cell-cell adhesion. In cnidarians, a recent study using *Nematostella* revealed that the  $\alpha$ E-catenin is vital for tissue cohesion and embryonic development (Nathaniel Clarke *et al.*, 2019). As stated earlier, CCC homologs have been identified in *Placozoa*, *Porifera* and, *Ctenophora*. Although there is a lack of functional analysis, in *Porifera*, proteomic and expression evidence suggests that  $\beta$ -catenin and other components of the CCC may be functional at cell contacts (Schippers and Nichols, 2018). In *Ctenophora*, it is not known whether  $\beta$ -catenin is vital in cell-cell adhesion as it does not localise at cell-cell contacts in *M. leidyi* (Salinas-Saavedra *et al.*, 2019) and its E-cadherin homolog has a divergent catenin binding motif that would not permit that binding of  $\beta$ -catenin (Belahbib *et al.*, 2018). All of the above highlights the necessity to investigate the nature of cell-cell adhesion as mediated by the CCC.



**Figure 1.5.  $\beta$ -catenin's role in cell adhesion.** A simple epithelium (top image) is composed of cells linked to each other at the adherens junction localised in the apical membrane. Cell adhesion is mediated at the CCC by E-cadherin, which links two cells,  $\beta$ -catenin, and  $\alpha$ E-catenin, which interacts with F-actin. (Created with Biorender.com).

### Transcriptional regulatory role of $\beta$ -catenin

The majority of knowledge on the role of  $\beta$ -catenin in canonical Wnt signalling has been derived from the use of bilaterian models where it has been shown to play developmental roles in gastrulation, axial patterning, stem cell maintenance (Schohl and Fagotto, 2002; Clevers, 2006; Hudson *et al.*, 2013). However, there is a considerable lack of understanding of  $\beta$ -catenin's transcription regulatory role in basal metazoans. The origin of  $\beta$ -catenin is intimately linked to the emergence and evolution of animal body organization (Holstein *et al.*, 2011). The evolutionarily conserved  $\beta$ -catenin functions in embryonic organizer induction, patterning, gastrulation and neural development have also been shown in cnidarians (Hobmayer *et al.*, 2000; Röttinger *et al.*, 2012; Watanabe *et al.*, 2014).

In basal metazoans, functional analysis of canonical Wnt signalling is limited. Early research using *Hydra* suggested that  $\beta$ -catenin might be functional in axis formation as its expression in *Xenopus* embryo resulted in induction of a secondary axis (Hobmayer *et al.*, 2000). Later, when using *Nematostella*, it was shown that  $\beta$ -catenin accumulates in nuclei of cells at the blastopore (Wikramanayake *et al.*, 2003) and treatment of GSK3 $\beta$  inhibitor results in exogastrulation (Röttinger *et al.*, 2012). These findings indicate that the  $\beta$ -catenin signalling has deep evolutionary roots and may have played essential roles for early animals in establishing the embryonic organizer. Even though genomic studies have shown the conservation of structural components of  $\beta$ -catenin proteins in basal metazoans, some upstream and downstream genes associated with its subcellular localization (cytoplasm or nucleus) do not seem to exist in the genomes of basal metazoans such as poriferans and ctenophores.

In the ctenophore, *Mnemiopsis leidyi*, the main components of canonical Wnt signalling have been identified (Pang *et al.*, 2010). Interestingly, unlike in bilaterians & cnidarians,  $\beta$ -catenin expression is observed after gastrulation and inhibition of GSK3 $\beta$  produced no phenotype in *M. leidyi* embryos during early development (Pang *et al.*, 2010). This, therefore, creates a question of whether canonical Wnt signalling activity is functional in ctenophores. In poriferans *Amphimedon queenslandica* & *Sycon ciliatum*,  $\beta$ -catenin signalling elements were found to be localised posteriorly, suggesting that they may be involved in axial patterning of the primary axis (Adamska *et al.*, 2007; Leininger *et al.*, 2014). Indeed of evidence of a functional canonical Wnt signalling pathway in poriferans stems from evidence in *Oscarella*, where inhibition of GSK3 $\beta$  results in a significant increase in the number of ostia (Lapébie *et al.*, 2009).

### 1.3.2 Approach to solving remaining issues.

A functional analysis across all metazoan phyla would be the most suitable way to investigate the evolutionary role of  $\beta$ -catenin. However, understanding the role of  $\beta$ -catenin in cell-cell adhesion and body plan development has primarily been carried out using bilaterian models because there are challenges associated with carrying out gene function experiments in non-bilaterian models, including basal metazoans, e.g., difficulties in keeping the laboratory culture and obtaining their eggs/embryos constantly. In addition to such technical issues, there are difficulties also in analysing the evolutionarily conserved functional feature(s) of  $\beta$ -catenin because of co-option and re-wiring of the  $\beta$ -catenin functional network during a long evolutionary time, e.g., Wnt/ $\beta$ -catenin signalling in *Hydrozoa* (*Cnidaria*) has been demonstrated to be involved in cell differentiation of nematocytes (sting cells) that are a unique characteristic of *Cnidaria* (Teo *et al.*, 2006; Khalturin *et al.*, 2007). The lineage-specific modification of  $\beta$ -catenin function could be achieved via changes in the gene regulatory network, the protein complex's composition, and the employment of lineage-specific gene products. Therefore, to illustrate the “core” molecular and functional feature of the  $\beta$ -catenin protein, it is necessary to compare the function of basal metazoan  $\beta$ -catenin proteins in a similar biological context.

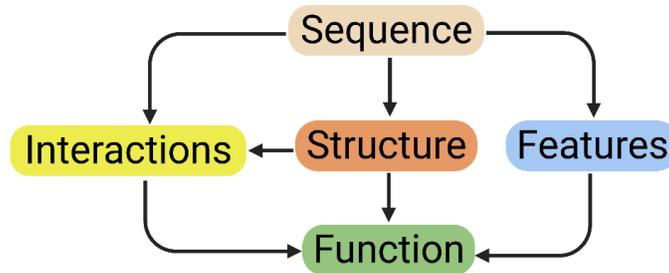
In the current study, I designed an experiment that involved the expression of metazoan  $\beta$ -catenin in the *Xenopus* embryo. This approach could allow the investigation of;

- a)  $\beta$ -catenin protein-protein interactions systematically through the comparative level of expression of FLAG-tagged  $\beta$ -catenin protein at the same developmental stage, purification of the complex using an anti-FLAG antibody, and proteomic analysis by LC-MS/MS.
- b) Conservation of  $\beta$ -catenin signalling function in the organiser (embryonic signalling centre) inducing activity using the conventional secondary axis induction and TOPflash (Wnt activity) assays.
- c)

### 1.3.3 Novelty of approach

With the advancement of sequencing technologies, the genomes/ transcriptomes of several organisms have been deciphered. From this, assumptions on protein functions have been made based on either the presence or absence of a protein of interest. However, this alone is not enough because protein function is based not only on the presence of a gene in the genome, but also on multiple factors, including protein-protein interactions, protein structure, sequence, and domains/ motifs (**Figure 1.6**). To this end, this study combined an integrative method involving functional, domain, structural, and proteomic analysis by expressing non-bilaterian  $\beta$ -catenin in a *Xenopus*

embryo. The *Xenopus* embryo is an advantageous model for this kind of analysis because the embryos can tolerate manipulations (e.g., tissue manipulations and germ layer dissections). Injection of mRNA in a specific region (blastomere) of the embryos is also possible.



**Figure 1.6. Study design to understand the evolution of  $\beta$ -catenin functionality.** A protein's functionality in different species depends on multiple factors, including structure, interactions, sequence, and features (domains).

Most studies about  $\beta$ -catenin protein complexes were done using cell lines (Tian *et al.*, 2004; Semaan *et al.*, 2019). These results added knowledge to our understanding of  $\beta$ -catenin interactions and function, but they have one limitation. Since the  $\beta$ -catenin functionality is context-dependent (Nakamura *et al.*, 2016), there will be differences in interactions between different cell types & tissues and at different developmental stages. Using a gastrula stage *Xenopus* embryo (as per this study) would uncover  $\beta$ -catenin interacting proteins from specified cell types within the three germ layers resulting in a global interactome, i.e.,  $\beta$ -catenin protein complexes specific to the ecto-, meso-, and endoderm cells. This can be significant as  $\beta$ -catenin interacting proteins could function as coactivators when bound to  $\beta$ -catenin resulting in a cell/tissue-specific transcription response. Coactivators could work in such a way that  $\beta$ -catenin interaction with TCF would not be sufficient to elicit gene transcription but would require the binding of a cell/ tissue-specific coactivator for a gene-targeted transcription.

Furthermore, apart from bilaterians and cnidarians, the role of  $\beta$ -catenin in axis specification in *Ctenophora* and *Porifera* is not known. By carrying out a  $\beta$ -catenin transphylectic secondary axis induction assay, we can check with these non-bilaterian  $\beta$ -catenins play a role for axis specification in a similar biological context. Confirmation of this would be a piece of crucial information to examine the endogenous function of  $\beta$ -catenin in their respective species. This study will provide novel insights into some fundamental questions regarding  $\beta$ -catenin signalling in basal metazoans.

# CHAPTER 2

## Chapter 2 Structural analysis of metazoan $\beta$ -catenin evolution

### 2.1 Summary

Proteins fold into specific structures according to their amino acid sequences. The resulting 3D structure can affect the protein's function, permitting or inhibiting vital interacting proteins. Using template-based protein structure modelling, this study revealed that regardless of evolutionary amino acid substitutions since metazoan emergence, the  $\beta$ -catenin structure has not changed. Surprisingly, unicellular choanoflagellate " $\beta$ -catenin-like" protein bears some structural similarity to that of metazoans suggesting that the basic architecture of  $\beta$ -catenin has its origins in unicellular organisms. Sequence alignment also reflected the conservation of critical domains across metazoans with some crucial residues conserved in choanoflagellate " $\beta$ -catenin-like" protein, indicating a possible pre-metazoan innovation of this crucial protein.

### 2.2 Background

Advancement in sequencing technology has led to an increase in the availability of sequence data of various organisms over the years. As a result, several evolutionary studies have been sequence centred. However, compared to the sequence data, the amount of structural information is still low. The sequence conservation in protein homologs has led to assumptions that protein function across different species is also conserved. This may indeed be true in closely related organisms, i.e., organisms within the same phylum. However, the functionality of a protein is also dependent on its structure. This is because a slight change in structure can allow or inhibit the binding of an essential interacting protein, possibly altering the functionality of that protein (**Figure 1.6**). These changes in structures are usually down to mutations in DNA. Several mutations in DNA are silent; hence they do not affect the resulting amino acid. However, in some cases, amino acid substitutions occur under missense, deletion, or insertion mutations. In such circumstances, a single amino acid substitution can lead to a change in the backbone conformation and hydrogen-bonding pattern, thereby constraining evolutionary choices by affecting which proteins can bind (Worth *et al.*, 2009). Researchers are now emphasising incorporating protein structures into evolutionary models to understand the effect of amino acid substitution on protein stability (Jaimes *et al.*, 2020; Di Nardo *et al.*, 2021).

X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are the most commonly used techniques in protein molecular structure determination (Yee *et al.*, 2005). X-ray crystallography requires first the protein to be crystallised and the resulting highly ordered crystals diffracted at high resolution. In NMR spectroscopy, purified protein molecules in solution are placed in a magnetic field resulting in the alignment of atoms' nuclei with the magnetic field. Information is then attained on the spatial arrangement of the protein's atoms when the sample is exposed to radio frequency electromagnetic radiation (Howard, 1998). However, as good as X-ray crystallography and NMR spectroscopy are, they are very costly to implement, require a high level of expertise in the respective field to get an excellent 3D structure, and need a high amount of purified protein. Recently computational biologists have developed tools that allow the prediction of protein structure. These are divided into two categories:

- a) Template-based protein structure modelling: Also known as comparative modelling, it depends on the principles regarding the divergence and/or relationship between structure and sequence in proteins (Chothia and Lesk, 1986). It primarily consists of four steps.
  - i. Identification of a suitable structural template. This can be obtained from the protein databank (PDB) database.
  - ii. Alignment of the template to the target sequence.

- iii. Refining model to account for amino acid substitutions, insertions, and deletions.
- iv. A final step may involve checking the quality of the model against various parameters. If the model does not meet the required parameters, further alignment may be required.

Template-based protein modelling is perfect for closely related species, but it can be optimised for distantly related organisms with good alignment.

- b) *De novo* protein structure modelling: This has no limitation observed in comparative modelling where the molecular structure of the template against which the unknown protein structure will be constructed ought to be known. It can therefore be used to predict the 3D structure of any protein from a given sequence. It works on the notion that the native state of a protein, when functional, is at the global free energy minimum and computationally carries out an extensive search of conformational space for tertiary protein structures that are predominantly low in free energy for the given amino acids (Baker and Sali, 2001).

In this chapter, using comparative modelling, our study presents the predicted structures of  $\beta$ -catenin from across five representative metazoans and one non-metazoan. An in-depth analysis of critical regions critical in cell-cell adhesion and transcriptional regulation will be described in later chapters.

## 2.3 Methods

### 2.3.1 Structural modelling of $\beta$ -catenin proteins

Bidirectional BLAST searches using *Mus musculus*  $\beta$ -catenin were carried out against databases chosen representative organisms of *Bilateria* (*Xenopus laevis*), *Cnidaria* (*Nematostella vectensis*), *Ctenophora* (*Bolinopsis mikado*), & *Porifera* (*Ephydatia fluviatilis*) and as an outgroup, *Choanoflagellata* (*Salpingoecca rosetta*) (see appendix 1). Single homologs were identified from each organism, while in *S. rosetta*, the best hit was denoted “ $\beta$ -catenin-like” protein.

With technical assistance from collaborators, the amino acid sequences of *E. fluviatilis* & *B. Mikado* were aligned with MIQS, a scoring matrix optimised to detect distant homologs (Yamada and Tomii, 2014). The putative  $\beta$ -catenin of *S. rosetta* was aligned with structurally known  $\beta$ -catenin using HHpred (Zimmermann *et al.*, 2018). Due to a high substitution rate in the N- and C-termini, alignment was focused on the region starting the  $\alpha$ E-catenin binding site to the 12<sup>th</sup> armadillo repeat. Multiple sequence alignments of  $\beta$ -catenin were calculated using FAMSA (Deorowicz *et al.*, 2016).  $\beta$ -catenin 3D models were built by using MODELLER v9.20 (Webb and Sali, 2016), based on alignments of target  $\beta$ -catenins for structurally known  $\beta$ -catenins, calculated by FORTE/ DELTA-FORTE, which are profile-profile alignment methods (Tomii *et al.*, 2005).

I then validated the predicted structural models through PROCHECK analysis (Laskowski *et al.*, 1993) which generated Ramachandran plots that showed the distribution of the combinations of the backbone dihedral angles Phi & Psi (Ramachandran *et al.*, 1963). Finally, visualisation of the final model structure was carried out using UCSF Chimera (Version 1.15) (Pettersen *et al.*, 2004). Template comparison was carried out by superimposing the 3D structures to the template *M. musculus*  $\beta$ -catenin structure using the default settings of the Matchmaker tool of UCSF Chimera. This was followed by an analysis of repeats in the central region using armadillo repeat boundaries defined by Huber *et al.* (1997).

### 2.3.2 Conservation of crucial domains in metazoan $\beta$ -catenin

$\beta$ -catenin sequences from various taxa were collected using NCBI protein databases by bidirectional BLAST as described in section 2.3.1.

This dataset extended to major taxa of bilaterians (chordates, ambulacrarians, lophotrochozoans and ecdysozoans), cnidarians, placozoan, poriferans, ctenophores and non-metazoan choanoflagellate. In total, 20 species were used in the analysis. Sequence alignments were performed using MUSCLE (Edgar, 2004). Manual curation was then undertaken to match the current alignment with the alignment used for structural analysis, resulting in a final high quality alignment.

Much research points to  $\beta$ -catenin C-terminal functioning as a transactivation domain (Gallet *et al.*, 1999; Vleminckx *et al.*, 1999; DuChez *et al.*, 2019). Therefore, to get an in-depth understanding of this region, Multiple Expectation Maximisation for Motif Elicitation (MEME) analysis was carried out on the C-terminals of unaligned protein sequences. MEME Suite 5.3.3 was used in motif identification using default settings (Bailey *et al.*, 2009).

## 2.4 Results

### 2.4.1 Domain conservation

Using the complete sequence of *Mus musculus* (mouse)  $\beta$ -catenin and aligning it with other metazoans, we found that the percentage sequence identity decreased towards basal metazoans (Table 2.1). The alignment of various metazoans indicated that most of the critical domains required for  $\beta$ -catenin functionality are conserved. In the N-terminal, the GSK3 $\beta$  and CK1 $\alpha$  phosphorylation sites critical in  $\beta$ -catenin degradation were conserved in all metazoans except for *H. vulgaris* which lacked the site corresponding to S37 (Figure 2.1B). Interestingly in unicellular *S. rosetta*, two of these sites corresponding to S33 & T41 were also conserved (See appendix 2 for complete alignment).

	<i>M.musculus</i>	<i>X.laevis</i>	<i>N.vectensis</i>	<i>E.fluviatilis</i>	<i>B.mikado</i>	<i>S.rosetta</i>
<i>M.musculus</i>	100	97.2	65.5	41.5	27	11.9
<i>X.laevis</i>	97.2	100	65.9	41.2	27.2	11.8
<i>N.vectensis</i>	65.5	65.9	100	40.7	27.1	11.6
<i>E.fluviatilis</i>	41.5	41.2	40.7	100	26.4	11.4
<i>B.mikado</i>	27	27.2	27.1	26.4	100	12.4
<i>S.rosetta</i>	11.9	11.8	11.6	11.4	12.4	100

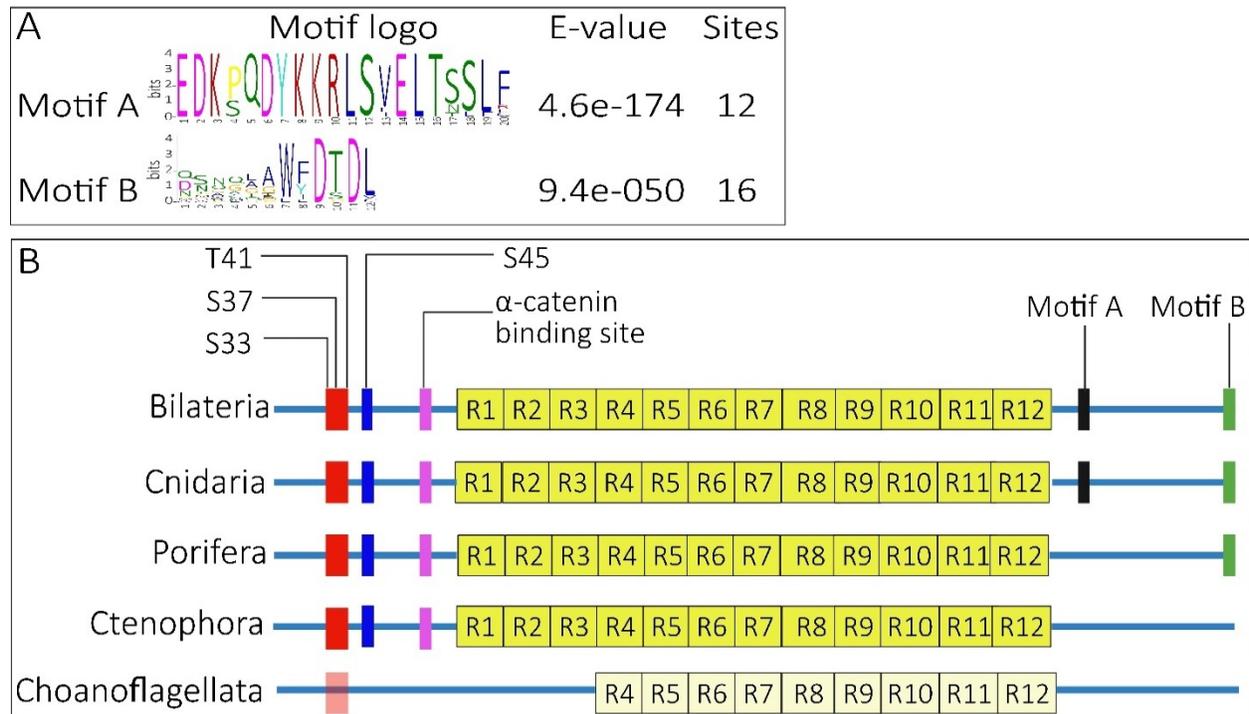


**Table 2.1 Percentage amino acid sequence similarity between different metazoan  $\beta$ -catenins.** Sequence similarity was in *Bilateria* and *Cnidaria* (>50%) & low in *Porifera* and *Ctenophora*. Non-metazoan *S. rosetta*  $\beta$ -catenin-like protein displayed the lowest similarity to mouse  $\beta$ -catenin.

The central region comprised of 12 armadillo repeats was highly conserved between bilaterians and cnidarians (Figure 2.1B and Figure 2.4). Nevertheless, in *Porifera* and *Ctenophora*, 12 armadillo repeats were observed following structural alignment but with variable amino acid substitutions within some helices. Similarly, *S. rosetta*  $\beta$ -catenin-like protein displayed some conservation within the central region, particularly between the 4<sup>th</sup> and 12<sup>th</sup> armadillo repeats. For understanding the functional role of  $\beta$ -catenin in cell-to-cell adhesion, it was found that the two

lysine residues key in the interaction with E-cadherin were conserved in all metazoans. These will be looked at in greater detail in a later chapter. Similarly, the three arginine residues critical in TCF interaction were conserved in all metazoans, but ctenophores expressed conservation of the first residue corresponding to mouse  $\beta$ -catenin R474 but had substituted the other crucial residues. These will also be looked at in greater detail in a later chapter.

In the C-terminal, the degree of conservation was low. However, MEME analysis for conserved motifs identified two statistically significant motifs in the C-terminal (Figure 2.1A). These were previously denoted motifs A and B in a study of bilaterians and cnidarians (Schneider *et al.*, 2003). MEME analysis confirmed that motif A is present in only bilaterians and cnidarians. Motif B, which forms a PDZ (DTDLD) domain, was conserved among bilaterians, cnidarians, poriferans but not in ctenophores and choanoflagellate.



**Figure 2.1 Conservation of crucial domains of  $\beta$ -catenin.** (A) MEME analysis identified two statistically significant motifs in the C-terminal. The sequence of each motif is indicated with large letters indicating greater significance and vice versa. (B) Analysis of alignment revealed a high degree of conservation of the central region, the GSK3 $\beta$  (red) & CK1 $\alpha$  (blue) phosphorylation sites and  $\alpha$ -catenin binding site across metazoans. In the C-terminal, motif A was conserved in only bilaterians & cnidarians, while motif B was conserved in bilaterians, cnidarians & poriferans only. In choanoflagellate, *S. rosetta*, the possible GSK3 $\beta$  phosphorylation site presented 2 out of the three sites (hence light shade of red). Similarly, although conservation was low in the central region, the structural analysis revealed conservation of repeats with those of metazoans.

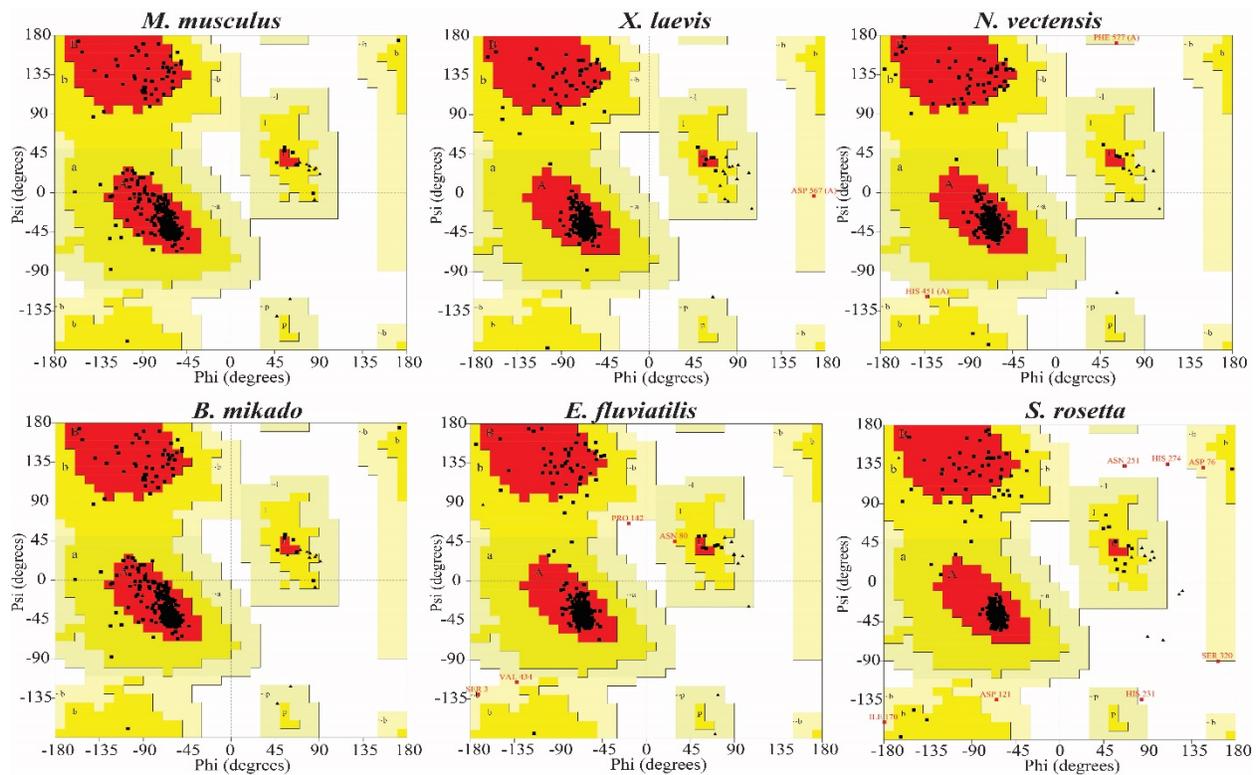
#### 2.4.2 Homology models & structure validation

The Phi & Psi distribution of the Ramachandran plots generated by PROCHECK indicated that over 90% of amino acid residues of modelled structures were in favoured positions (Table 2.2 & Figure 2.2). Despite the amino acid variation observed across the metazoan phyla, the subsequent predicted structure of  $\beta$ -catenin has not evolved much since metazoan emergence (Figure 2.3). Interestingly, despite its low sequence similarity to metazoans (as observed in Table 2.1), between residues 452-878, the *S. rosetta*  $\beta$ -catenin-like protein possesses a structure that bears considerable

similarity to metazoan  $\beta$ -catenins. However, *S. rosetta* “ $\beta$ -catenin like” presented a short loop, whereas the metazoan  $\beta$ -catenins presented a long loop.

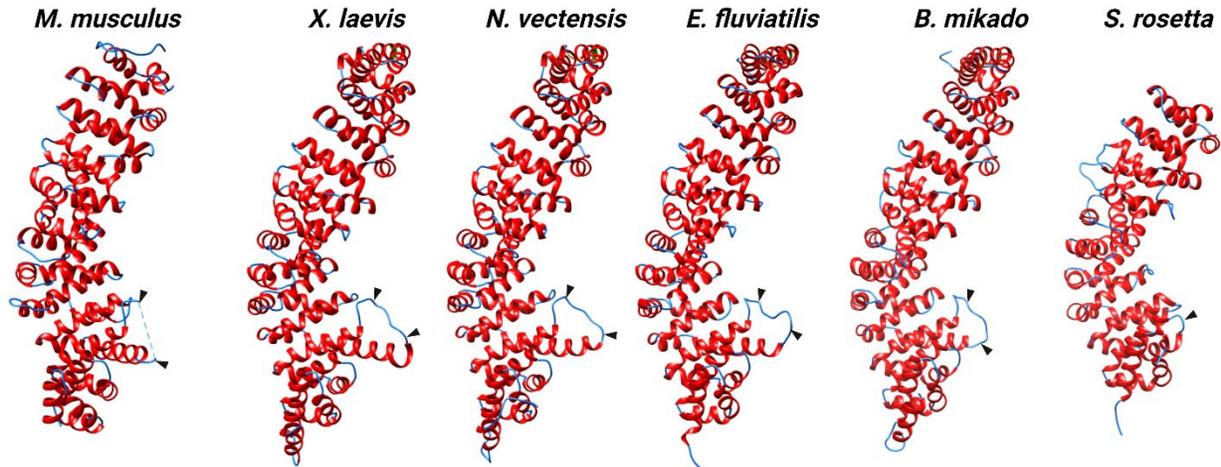
	<i>M. musculus</i>	<i>X. laevis</i>	<i>N. vectensis</i>	<i>E. fluviatilis</i>	<i>B. mikado</i>	<i>S. rosetta</i>
Residues in the most favored region (%)	95.4	96.4	96.2	95.5	94.4	90.1
Residues in additionally allowed region (%)	4.6	3.4	3.4	3.9	4.8	8.1
Residues in generously allowed region (%)	0	0.2	0.4	0.6	0.4	1.0
Residues in disallowed region (%)	0	0	0	0	0.4	0.8

**Table 2.2** Ramachandran plot calculation for 3D models of metazoan  $\beta$ -catenin computed with PROCHECK. Based on non-glycine, non-proline residues, most of the residues were in the most favoured regions.



**Figure 2.2.** Ramachandran plot analysis. The majority of the residues (black squares) were localised in favoured positions (red regions), followed by those in allowed positions (yellow regions). Few residues were localised in generously allowed positions (light yellow). Disallowed regions are indicated in white.

Superimpositions of the 3D structures revealed that all metazoans had 12 armadillo repeats that aligned with those of the known mouse  $\beta$ -catenin structure (**Figure 2.4A**). Superimposing of the *S. rosetta* “ $\beta$ -catenin like” protein onto that of mouse  $\beta$ -catenin showed that although the degree of amino acid conservation in the central region was low, *S. rosetta* “ $\beta$ -catenin like” protein bore armadillo repeats that aligned with those of mouse  $\beta$ -catenin. However, this was only observed in the 4<sup>th</sup> to 12<sup>th</sup> armadillo repeat (**Figure 2.4B**).

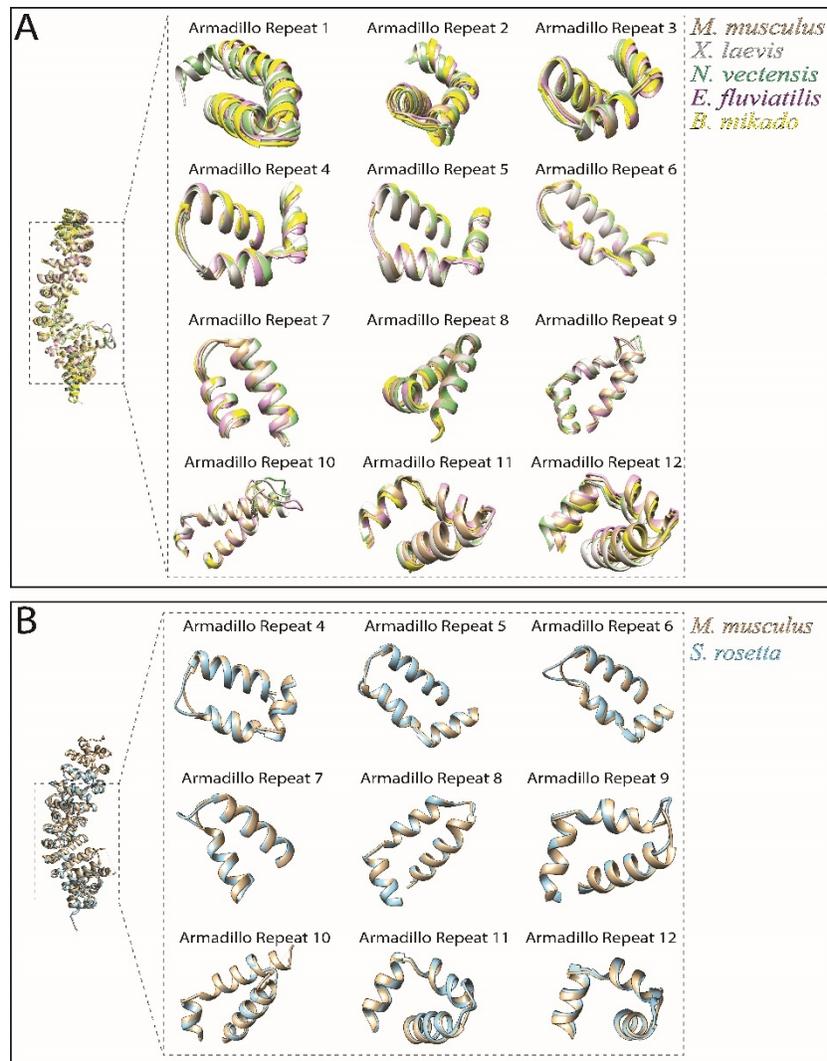


**Figure 2.3. The 3D architecture of metazoan  $\beta$ -catenin proteins is highly conserved.** Compared to the significant level of divergence of their amino acid sequences, no noticeable difference was observed in their predicted 3D structures, although some considerable differences were observed in pre-metazoan *S. rosetta*  $\beta$ -catenin-like protein. For example, a small loop between repeat 10 and 11 (black arrow).

## 2.5 Discussion

### 2.5.1 Structural conservation of metazoan $\beta$ -catenin

The fact that the structure of  $\beta$ -catenin has not evolved much since the protein's innovation was not so much of a surprise. It has been suggested that a 30% similarity threshold between protein sequences is required for proteins to adopt the same structure (Gilson *et al.*, 2017). The sequence identities of *B. mikado* & *S. rosetta* were under this threshold (**Table 2.1**). In *B. mikado*, with a 27% sequence similarity to mouse  $\beta$ -catenin, the resulting structure was not different from that of metazoan  $\beta$ -catenins. Similarly, the structure of *S. rosetta* “ $\beta$ -catenin-like” protein, even with an 11.88% sequence identity to mouse  $\beta$ -catenin, showed a structure that was not significantly different from that of mouse  $\beta$ -catenin. This could imply that the previously suggested 30% similarity threshold required for two proteins to adopt the same structure could be extended to 10%. Furthermore, the conservation of the alpha-helices forming armadillo repeats in *S. rosetta* “ $\beta$ -catenin-like” protein suggests that this feature was not necessarily a metazoan invention. With the structure seemingly conserved in all metazoans, the interaction of  $\beta$ -catenin with other proteins will be a function of the amino acid residues and the strength of the bonds they form with other proteins. This may mainly hold true in the central region where several crystal structures of  $\beta$ -catenin interact with other proteins, e.g. APC,  $\alpha$ -catenin, ICAT & TCF (Daniels and Weis, 2002; Xing *et al.*, 2003; Ha *et al.*, 2004; Sampietro *et al.*, 2006). However, this means that the binding sites of several  $\beta$ -catenin interacting proteins possibly overlap, as has already been found in TCF and E-cadherin, which share critical binding sites at lysines 312 & 435 (Graham *et al.*, 2000; Huber and Weis, 2001).



**Figure 2.4. Conservation of central region of metazoan  $\beta$ -catenin.** Superimposition of predicted  $\beta$ -catenin structures onto template *M. musculus*  $\beta$ -catenin showed that model metazoan  $\beta$ -catenin had twelve conserved armadillo repeats while (A) while *S. rosetta*  $\beta$ -catenin like protein had nine conserved armadillo repeats (B).

### 2.5.2 Premetazoan origin of GSK3 $\beta$ phosphorylation site

Although there was an extensive amino acid substitution in  $\beta$ -catenin of different phyla, most key regulatory and interaction domains were conserved in metazoans. The GSK3 $\beta$  phosphorylation sites (S33, S37 & T41) are critical in the control of  $\beta$ -catenin destruction and have been previously shown to be conserved in most cnidarians & poriferans (Schneider *et al.*, 2003; Schippers and Nichols, 2018). Conservation of these sites in ctenophores almost led to the conclusion that this control system was a metazoan innovation. However, surprisingly, in unicellular *S. rosetta*, sites corresponding to S33 & T41 were conserved. Moreover, a checking of databases revealed the presence of GSK3 $\beta$  homologs in *S. rosetta* (XP\_004990991.1) and the other choanoflagellate, *M. brevicollis* (XP\_001745587.1). It is unknown whether 2 out of 3 phosphorylation sites are enough to elicit degradation, and most mutation targeting these sites has been carried out on all three residues (Yost *et al.*, 1996; Liu *et al.*, 1999). Also, lack of conservation at the potential CK1 $\alpha$  phosphorylation site in *S. rosetta*  $\beta$ -catenin-like protein could affect degradation as this site has been shown that its (S45) mutation prevents the phosphorylation of T41 in *Xenopus* embryos (Liu

*et al.*, 2002). Nevertheless, the presence of GSK3 $\beta$  homologs in *S. rosetta* and 2 out of 3 sites crucial for phosphorylation might indicate that this was a premetazoan innovation.

### 2.5.3 Domains of the N- and C-terminal

The highest rate of amino acid substitutions is usually observed in the flanking N- & C-terminals of proteins, which was the same in metazoan  $\beta$ -catenins. However, within these regions, there lie some essential sites. The role of the C-terminal is still not very clear. However, previously, it was shown that a TCF (LEF1)- $\beta$ -catenin fusion lacking the C-terminal domain is impaired in signalling while the fusion of just the  $\beta$ -catenin C-transactivational domain to the DNA-binding domain of LEF1 was sufficient for axis-induction (Vleminckx *et al.*, 1999). This suggests that the C-terminal could be vital in the signalling role of  $\beta$ -catenin.

Two motifs, A & B, were found in the C-terminal. Of these, motif A located adjacent to the 12<sup>th</sup> armadillo repeat based on its ability to bind transcription factor *Teashirt* to modulate Wingless signalling has previously been suggested to be responsible for the transactivation role in the C-terminal (Gallet *et al.*, 1999). However, this claim would require further experimental validation. Also, as motif A was conserved in only bilaterians and cnidarians, what impact (if any) could it have on the dual functionality of  $\beta$ -catenin. Motif B, which forms a PDZ domain (DTDLD), was present in all metazoans but ctenophores. Although the role of this domain is yet to be clearly understood, a recent work indicated that expression of a mutant  $\beta$ -catenin lacking the DTDLD domain resulted in a decrease in the transcriptional activity of  $\beta$ -catenin (DuChez *et al.*, 2019). Future analysis of the role of endogenous  $\beta$ -catenin in ctenophore is therefore essential.

In the N-terminal, a binding domain, the  $\alpha$ E-catenin binding domain was present in all metazoans. This will be discussed in greater detail in a later chapter. Similarly, the central regions where the majority of interactions occur was highly conserved. The regions of interest here are those involved in the interaction with TCF and E-cadherin. These will be discussed in detail in later chapters, where the results observed in the structural evolution could explain results obtained from other experiments.

## 2.6 Chapter conclusion

Structural analysis suggests the unicellular origin of the basic 3D architecture of  $\beta$ -catenin. Furthermore, the presence of some GSK3 $\beta$  phosphorylation sites and GSK3 $\beta$  homologs in *S. rosetta* suggests that a  $\beta$ -catenin-like protein and control mechanism were present in premetazoan eukaryotes. Clear-cut domain differences were observed in the C-terminal. In reconstructing the nature of ancestral  $\beta$ -catenin, it is vital to understand how these differences affect  $\beta$ -catenin functionality.

The next chapter will introduce our novel approach to identify the conserved  $\beta$ -catenin protein interactions in metazoans.

# CHAPTER 3

## Chapter 3 Proteomic approach to identify $\beta$ -catenin protein interactions conserved in metazoans

### 3.1 Summary

The ability of a protein to interact with other critical proteins can tremendously affect its function. Using a transphylectic and proteomic analysis, this study found conservation of the adhesion complex in metazoans. Furthermore, novel putative  $\beta$ -catenin interacting proteins were identified, suggesting possible new evolutionary roles of  $\beta$ -catenin in the mitochondrial pathways. These results may present new knowledge regarding the roles of  $\beta$ -catenin in different biological and molecular processes in metazoans.

### 3.2 Background

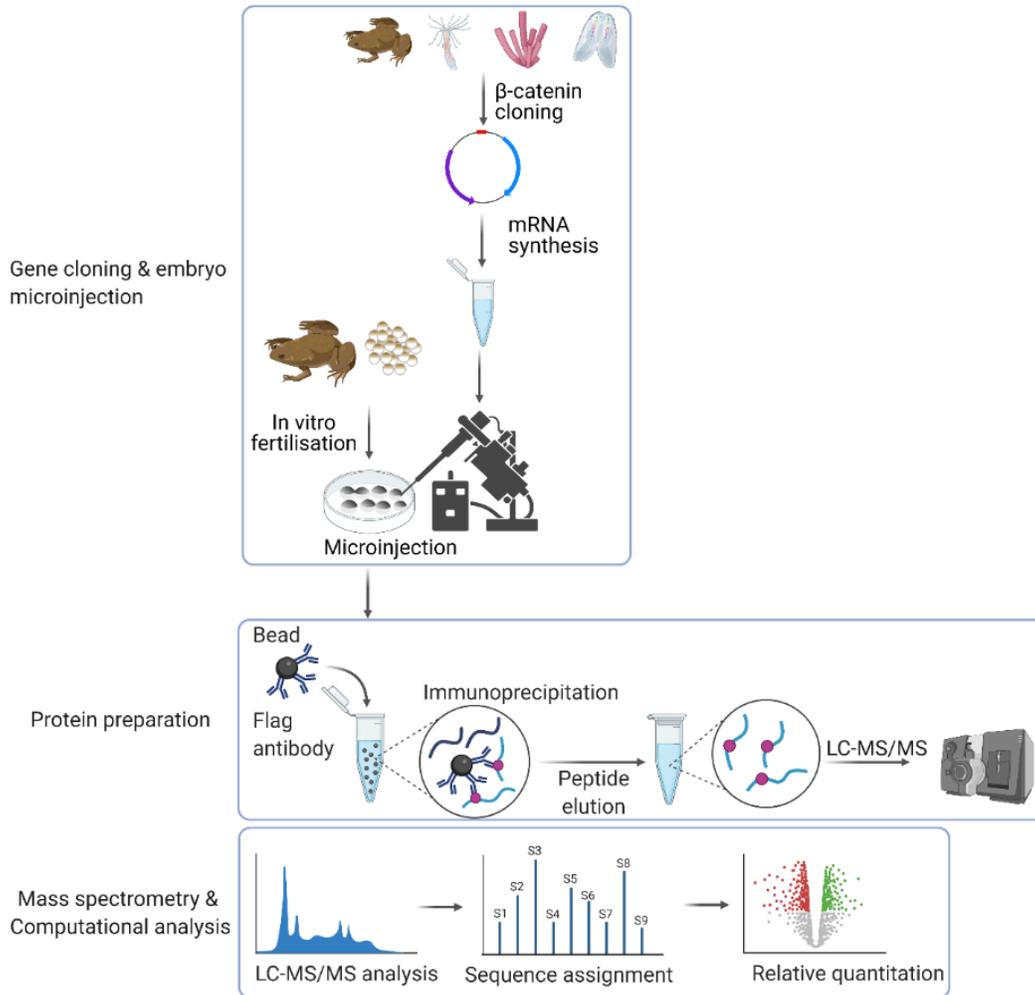
In animal cells, proteins do not work individually but interact with each other in evolutionarily well-engineered protein-protein interaction (PPI) complexes to allow signalling cascades (Alberts, 1998). This can easily be observed in critical cellular processes like transcription, mRNA translation, and DNA replication which call for a carefully organised number of proteins. Similarly, the evolution from unicellular organisms to multicellularity required the innovation of coordinated pieces of machinery that could play either structural or signalling roles, contributing to cell & tissue diversity and body plan development.

$\beta$ -catenin was important in multicellularity emergence as it plays both structural and signalling roles and is highly conserved in metazoans. In its structural roles at the cell membrane, it has been shown to interact with classical cadherins and  $\alpha$ -catenin, whereas, in its signalling role in the nucleus, it has been shown to interact with TCF and other transcription factors. Considerable information about  $\beta$ -catenin PPIs using immunoprecipitation with mass spectrometry (IP-MS) has primarily relied on cell cultures (Tian *et al.*, 2004; Semaan *et al.*, 2019). Cell cultures are used because they have a high growth rate and are easy to manipulate with various molecules, allowing the study of interacting partners under different conditions. However, recently in the study of  $\beta$ -catenin PPIs, researchers have begun to move away from cell cultures onto cultured tissues, allowing identification of tissue-specific PPIs (Hwang *et al.*, 2017; Amit *et al.*, 2020). Much knowledge on  $\beta$ -catenin interaction partners is a result of work on bilaterian models. There is, therefore, a need to understand the ability of the early metazoans  $\beta$ -catenin proteins to interact with these proteins. However, owing to difficulties in applying molecular and genetic technologies in these non-model basal metazoans, the ancestral functions and nature of  $\beta$ -catenin protein complexes remain largely unexplored.

In this study, our transphylectic expression of basal metazoan  $\beta$ -catenins in *Xenopus* embryos and proteomic analysis sought to provide the first experimental evidence about evolutionarily conserved protein components that can interact with basal metazoan  $\beta$ -catenins. To maximise the advantage of this approach, I decided to express N-terminally tagged  $\beta$ -catenin proteins. The lack of a pan-metazoan  $\beta$ -catenin antibody necessitated the use of a tagged protein. Furthermore, it would be possible to compare protein levels between metazoans by using western blot analysis. This tag was placed at the N-terminal to minimise the effect of protein interaction on the C-terminal as the C-terminal is a critical transactivation domain and site for the interaction of various proteins (Shibata *et al.*, 2003; Mosimann *et al.*, 2009; DuChez *et al.*, 2019).

### 3.3 Methods

A pictorial summary of the sample processing and analysis followed to identify proteins interacting with different metazoan  $\beta$ -catenin following expression in *Xenopus* embryo (**Figure 3.1**).



**Figure 3.1. Identification of conserved  $\beta$ -catenin PPIs.** The method followed was divided into three stages, i.e., cloning & embryo manipulation, protein preparation and mass spectrometry & computational analysis. Different techniques were used in each category, and these will be explained in detail. Created using Biorender.com

#### 3.3.1 Total RNA extraction

Total RNA was extracted from planula stage *N. vectensis*, larval stage *B. mikado*, and juvenile *E. fluviatilis* using Qiagen RNeasy mini kit (Qiagen, Hilden, Germany) following manufacturer's guidelines. In summary, animals were suspended in 350  $\mu$ l RLT buffer containing 1% (v/v)  $\beta$ -mercaptoethanol and homogenised. The lysate was then vortexed for 30 seconds, and 350  $\mu$ l of 70% ethanol was added & vortexed. The lysate was applied to an RNeasy column and centrifuged at 10000 RPM for 30 seconds. 250  $\mu$ l buffer RW1 was added to the column and centrifuged at 10000 RPM for 30 seconds. 80  $\mu$ l DNase1 solution (5  $\mu$ l DNaseI + 75  $\mu$ l RDD buffer) was applied to the column and incubated at room temperature for 15 minutes. Buffer RPE was applied to the column, and centrifugation was carried out as before. Final centrifugation was carried to remove any remaining wash buffer. The RNeasy column was placed into a new collection tube, 30  $\mu$ l of

RNase-free water was applied to the column and incubated at room temperature for 5 minutes. Elution was carried out by centrifugation at 10050 RPM for 1 minute.

### 3.3.2 cDNA synthesis

To 1  $\mu$ g of total RNA, 1  $\mu$ l 50 mM OligoDT, 1  $\mu$ l 10 mM dNTP mix and RNase-free water was added to 13  $\mu$ l. Solutions were mixed by pipetting and incubated for 5 minutes at 60°C. Next, an enzyme mix containing 4  $\mu$ l 5x First-strand buffer, 0.5  $\mu$ l Superscript IV RT enzyme, 2  $\mu$ l 0.1M DTT and 0.5  $\mu$ l RNaseOUT was added, gently mixed by pipetting. Initial incubation for 10 minutes at 25°C followed by a 2-hour incubation at 50°C. The reaction was stopped by incubating for 15 minutes at 70°C. Stock cDNA was stored at -80°C.

### 3.3.3 Polymerase Chain Reaction (PCR)

50  $\mu$ l PCR reactions were set up using Q5 High-Fidelity DNA Polymerase (M0491) (NEB, Ipswich, USA) (Table 3.1 & Table 3.2). The primer annealing temperature was calculated using the NEB Tm Calculator (<https://tmcalculator.neb.com/>) (Table 3.3). The reaction was run on a Mastercycler Nexus GX2 (Eppendorf). PCR products were run on a 1% agarose gel to confirm product size. In cases where multiply bands were observed, DNA gel extraction was carried out using a QIAquick gel extraction kit following the manufacturer's guideline; otherwise, a QIAquick PCR purification kit was used. PCR products were then cloned into pCS2+ plasmids using restriction digest enzyme in Table 3.4. Bacterial transformation with plasmids was then carried out using DH5 $\alpha$  competent cells (Takara Bio, Japan) following the manufacturer's guidelines. Individual white colonies were picked and inoculated in 5 ml LB medium containing ampicillin (100 mg/ml) and incubated at 37°C overnight at 200 RPM. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's guidelines and eluted in TE buffer and insertion presence confirmed by double digestion & sanger sequencing. Eluted DNA was stored at -20°C.

Reagents	Volume ( $\mu$ l)	Final Concentration
Nuclease free water	Variable	
5X Q5 Reaction buffer	10	1X
10 $\mu$ M Forward Primer	2.5	0.5 $\mu$ M
10 $\mu$ M Reverse Primer	2.5	0.5 $\mu$ M
10 mM dNTP	1	200 $\mu$ M
Template cDNA (1:100)	Variable	< 1000 ng
Q5 high-fidelity polymerase	0.5	0.02 U/ $\mu$ l
Total Volume	50	

Table 3.1. PCR reaction setup.

Step	Temperature	Time
Initial denaturation	98°C	30 seconds
35 Cycles	98°C	10 seconds
	*Variable	30 seconds
	72°C	30 seconds/kb
Final Extension	72°C	2 minutes
Hold	4°C	

Table 3.2. PCR conditions. Asterisk is Tm which varied with size & GC content of primer.

Gene	Primers	Primer sequence (5'-----3')	Tm/°C	Length
<i>N. vectensis</i> $\beta$ -catenin	Forward primer	AGA <b>GGATCC</b> GCCACCATG <b>GACTACAAAGACGATGACG</b> <b>ACAAG</b> ATGGAGACACACGGTATGGG	61	2256
	Reverse primer	TTCTC <b>TCTAGA</b> TCAGAGGTCTGTGT		
<i>E. fluviatilis</i> $\beta$ -catenin	Forward primer	ATAGA <b>GGATCC</b> GCCACCATG <b>GACTACAAAGACGATGA</b> <b>CGACAAG</b> ATGGAAACGCCAGTATATCA	61	2732
	Reverse primer	TCTC <b>TCTAGA</b> GTGTCATAGATCTGAGTCAAACCA		
<i>B. mikado</i> $\beta$ -catenin	Forward primer	A <b>GGATCC</b> GCCACCATG <b>GACTACAAAGACGATGACGAC</b> <b>AAG</b> ATGGAAACGCCAGTATATCA	60	2697
	Reverse primer	GA <b>CTCGAG</b> TCAGATGGCCGCGTAC		

**Table 3.3 Primers used in cloning  $\beta$ -catenin.** Restriction digest sites (highlighted blue (BamH1), yellow (Xba1) and green (Xho1)). A flag-tag coding sequence (highlighted pink) was also included in the forward primer. Green and red nucleotides represent binding sites for forward and reverse primers, respectively. Bold "ATG" represents the start codon. A Kozak sequence (underlined) was also included before the start codon to improve translation initiation. *X. laevis*  $\beta$ -catenin and *S. rosetta*  $\beta$ -catenin-like were generated by FASMAC, Japan.

### 3.3.4 Restriction digestion

Digestion was conducted at 37°C for 2 hours using commercially supplied enzymes and buffers (New England Biolabs). The restriction enzymes were deactivated as per the supplier's guidelines, usually by heat inactivation. Double digest products were purified by gel extraction, while single digest products were purified using a QIAquick Gel Extraction Kit and QIAquick PCR purification, respectively, following manufacturer's guidelines.

Species	Restriction enzymes	Length/ base pairs
<i>X. laevis</i>	BamH1 & EcoR1	2386
<i>N. vectensis</i>	BamH1 & Xba1	2256
<i>E. fluviatilis</i>	BamH1 & Xba1	2732
<i>B. mikado</i>	BamH1 & Xho1	2697
<i>S. rosetta</i>	BamH1 & Xba1	2685

**Table 3.4. Restriction digest enzymes used in plasmid insertion and corresponding length of the insert.**

### 3.3.5 Synthesis of capped mRNA for microinjection

Initially, plasmid DNA was linearised, as detailed above in section 3.3.4, using the Not1 restriction enzyme. Then, 1  $\mu$ g of linearized DNA was used for mRNA synthesis using mMACHINE SP6 Transcription Kit (Ambion) following manufacturer's guidelines with a few changes. Briefly, the components of the transcription kit were thawed at room temperature and then kept on ice apart from the 10x Reaction buffer. Next, the reaction mix was set up by adding together 10  $\mu$ l 2X NTP/CAP, 2  $\mu$ l 10x Reaction buffer, 1  $\mu$ g linear template DNA, 2  $\mu$ l enzyme mix and an appropriated amount of nuclease-free water added up to 20  $\mu$ l. The reaction was mixed by pipetting and incubated for 4 hours at 37°C. 1  $\mu$ l TURBO DNase was then added to the reaction mix and incubated at 37°C for 20 minutes. Finally, MEGAclean Transcription Clean-Up Kit (Ambion) was used to purify the synthesised mRNA following the manufacturer's instructions. Concentration was measured using a Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Inc), and its quality was confirmed using Agilent 4200 TapeStation (Agilent). The mRNA was then aliquoted and stored at -80°C.

### 3.3.6 Animals

Adult males and females of *X. laevis* were purchased from the *Xenopus* breeder in Ibaraki Prefecture and maintained in our frog vivarium. All experiments with *X. laevis* were approved by the Animal Care and Use Committees in the Okinawa Institute of Science and Technology Graduate University.

### 3.3.7 Ovulation induction in *X. laevis* females

A week before an experiment, *X. laevis* females were induced for ovulation by priming them with 100 IU Pregnant Mare Serum Gonadotropin (PMSG) (Aska Pharmaceutical, Japan) injected in the dorsal lymph sac. Then, 12-16 hours before the experiment, primed females were injected with 300 IU of human Chorionic Gonadotropin (hCG) (Aska Pharmaceutical, Japan) into the dorsal lymph sac.

### 3.3.8 Testes isolation from male *X. laevis* for *in vitro* fertilisation

One frog was immersed in 0.01% (w/v) tricaine methanesulfonate (MS-222) for 20 minutes at room temperature. Before dissection, the frog was placed on its' back to check if the frog was utterly limp. Then, using sharp forceps to pick loose skin and sharp scissors, the frog's ventral wall and musculature were cut through, exposing the underlying viscera. The testes lie at the base of the fat bodies on either side of the midline; hence testes were carefully dissected away. Dissected testes were further dissected into small pieces and stored in 1x De Boer's solution at 4°C. Stored testes remain viable for up to 2 weeks. Finally, the frog carcass was wrap in a plastic film and frozen at -20°C until disposal in biohazard materials.

### 3.3.9 Egg collection and *in vitro* fertilisation

A primed female frog that had already started to spawn was held by its hind legs and gently pushed her body by fingers to collect eggs from the cloaca on to 90 mm petri dish. The maximum holding time is 1 minute. A small piece of the isolated testis was homogenised in 600 ul 1x De Boer's solution, and the homogenate was mixed with the collected eggs, which were then spread out onto the petri dish using a Pasteur pipette. The eggs were then left for 5 minutes. Following this, the eggs were flooded with 0.1x Steinberg solution and incubated for 20 minutes at room temperature to allow fertilisation to proceed. Next, dejellying of fertilised eggs was carried out using a 2% L-Cysteine/ 0.1x Steinberg solution for up to 5 minutes. Dejellied eggs were then washed five times with 0.1x Steinberg solution and incubated at 20°C until the appropriate stage for experimental analysis.

### 3.3.10 $\beta$ -catenin mRNA microinjection

Transcribed mRNA was injected into fertilised *X. laevis* embryos at 1-cell stage. Due to differences in codon translation efficiency, the amount of mRNA injected for each metazoan was optimised such that flag-tagged protein expression was roughly the same across all organisms. Therefore, 3 ng of *B. mikado* flag-tagged  $\beta$ -catenin mRNA while 1 ng of *X. laevis*, *E. fluviatilis* and *N. vectensis*  $\beta$ -catenins were injected. The embryos were then cultured in 5% Ficoll in 1x Steinberg Solution at 20°C until gastrula (10-12) stage. They were then washed three times in 0.1x Steinberg Solution, snap-frozen until ready for protein extraction.

### 3.3.11 Protein extraction

Embryos were initially thawed on ice for 10 minutes and then lysed by pipetting in 1 ml cold lysis buffer (20 mM Tris-HCl, pH 8, 70 mM KCl, 1 mM EDTA, 1% NP-40, 10% glycerol and 1x cOmplete proteinase inhibitor (Roche)). This was followed by sonication in a water bath at 4°C for 60 seconds. The lysate was then clarified by centrifugation for 15 mins at 14,000 rpm at 4°C to remove cell debris, pigment, and yolk. The supernatant was recovered, and subsequently, protein concentration was measured using a Direct Detect infrared spectrometer (Merck, Darmstadt, Germany). Aliquots of the supernatant were then snap-frozen in liquid nitrogen and stored at -80°C.

### 3.3.12 Western blotting

Samples were made up to 1  $\mu\text{g}/\mu\text{l}$  by adding 2x sample buffer and a calculated amount of lysis buffer and boiled at 95°C for 5 minutes. Except for immunoprecipitation eluates where 8% of elution volume was loaded, 10  $\mu\text{g}$  per lane were loaded. SDS-PAGE was carried out using 7.5% polyacrylamide gels (BioRad, Hercules, Ca). The gel was initially run at 50 V for 30 minutes and then 100 V for 1 hour or until the leading marker was 1 cm from the base. Subsequently, protein extracts were transferred onto a methanol activated polyvinylidene difluoride (PVDF) membrane. The PVDF membrane and gel were sandwiched between two filter papers soaked in 1x transfer blot buffer. A Trans-Blot Turbo System (BioRad, Hercules, Ca) was then used to transfer the proteins at 25 V for 7 minutes. After transfer, the PVDF membrane was blocked with 5% skimmed milk in 0.1% Tween 20 in PBS (PBST) for 1 hour at room temperature on a shaker. Then, the blocking buffer was removed, and an overnight incubation at 4°C of the membrane with primary antibody diluted in PBST was carried out. Following overnight incubation, the membranes were washed 3 x 5 minutes in PBST and then incubated for 1 hour at room temperature in horseradish peroxidase (HRP) linked secondary antibody in 5% skimmed milk in PBST. Finally, the membrane was rewashed 3 x 5 minutes in PBST and chemiluminescence development by treating the membrane with ImmunoStar Zeta (Wako Pure Chemical Industries, Osaka).

### 3.3.13 Immunoprecipitation of flag-tagged $\beta$ -catenin

This was carried out on stage 10-12 (gastrula) embryos. Protein extraction was carried out as described in section 3.3.11. Initially, anti-flag M2 magnetic beads (Sigma Aldrich) were equilibrated by washing them in 1 ml of lysis buffer for 10 minutes. 40  $\mu\text{l}$  of the magnetic beads were then added to 3 mg of protein lysate made up to 1 ml by adding cold lysis buffer. Overnight incubation with rotation was carried at 4°C. The beads were then pulled down, washed 3 x 5 minutes in cold lysis buffer and eluted in two ways, as detailed in sections 3.3.14 & 3.3.15.

### 3.3.14 In-solution (IS) trypsin digestion

75  $\mu\text{l}$  elution buffer (2% SDC, 100 mM Tris-HCl (pH 8), 10 mM DTT) was added to the beads and subsequently boiled at 60°C for 30 minutes. Alkylation by adding 1 M IAA (4.2  $\mu\text{l}$ ) to 55 mM and incubating at room temperature in the dark for 30 minutes. To 50  $\mu\text{l}$  of supernatant for mass spectrometric analysis, 2 mM  $\text{CaCl}_2$  (400  $\mu\text{l}$ ) and vortexed. 150 ng of trypsin was then added, vortexed and initially incubated in a water bath at 37°C at 400 watts for 1 hour. Subsequent overnight incubation was then carried out in an incubator at 37°C. Trypsin inactivation was carried out by adding & vortexing with 400  $\mu\text{l}$  Ethyl Acetate. 40  $\mu\text{l}$  10% Trifluoroacetic acid (TFA) were added, vortexed and centrifuged at 4,200 RPM resulting in separation into two phases. The top phase was discarded. An Empore SDB-RPS stage tip (3M, Minneapolis) was conditioned by

centrifuging at 4,200 RPM with 50  $\mu$ l Ethyl Acetate/ 0.1 % TFA in water. The sample was loaded onto the tip and centrifuged for 4 minutes at 3,000 RPM, following which the tip was washed with 150  $\mu$ l Ethyl Acetate/ 0.1 % TFA in water. Two final washes were carried out with 200  $\mu$ l 0.1% TFA in water. Peptide elution was carried out with 200  $\mu$ l 5%  $\text{NH}_3\text{OH}$  with 80% Acetonitrile (ACN) and then dried at 42°C for 50 minutes using the HPLC method of a Genevac EZ-2 vacuum evaporator.

### 3.3.15 In-gel (IG) trypsin digestion

Elution was carried out in 65  $\mu$ l 1x sample buffer (Nacalai Tesque, Inc) by boiling at 95°C for 5 minutes. 40  $\mu$ l of the eluate was loaded onto a 7.5% polyacrylamide gel, and SDS-PAGE carried out as described before. The gel was then stained using SimplyBlue SafeStain (Thermo Fisher Scientific), following the manufacturer's guidelines. Individual gel lanes were separated into six fractions from high to low molecular weight using a razor blade. The fractions were further diced into small pieces (2 x 2 mm) and transferred to a low-bind 96-well Eppendorf plate. The gels were washed with 100  $\mu$ l MilliQ water by gentle vortexing and incubating for 50 minutes. They were then dehydrated by removing all the water and replacing it with 100 $\mu$ l 50% ACN in 50 mM Ammonium bicarbonate (AMBIC) (pH 7.8), vortexing and incubating for 10 minutes. Finally, all this solution was removed and replaced with 100% ACN and incubated for 10 minutes before removing all ACN. Dehydrated gel cubes were then subjected to reduction with 80  $\mu$ l 10 mM DTT in 50 mM AMBIC (pH 7.8) for 30 minutes at 56°C. This was followed by 30 minutes of alkylation in the dark with 80  $\mu$ l of 55 mM IAA in 50 mM AMBIC (pH 7.8). Sequential washing of the gel pieces was carried with 100  $\mu$ l MilliQ water, 10 minutes incubation in 100 $\mu$ l 50% ACN in 50 mM AMBIC (pH 7.8) and finally dehydrated in 100% ACN. MS-grade trypsin was made up to 10 ng/ $\mu$ l by suspending it 50 mM AMBIC in 10% ACN. Gel rehydration was carried out by adding 15  $\mu$ l (150 ng) of trypsin enzyme solution and incubating on ice for 15 minutes. Finally, 50  $\mu$ l AMBIC in 10% ACN was added to the gels and incubated overnight at 37°C. 25  $\mu$ l formic acid was added to each well, and the plate gently vortexed and incubated for 1 minute. The supernatant was collected in a new low bind well Eppendorf tube. 100  $\mu$ l 50% ACN/ 5% formic acid was then added to each well, vortexed and incubated at 37°C for 45 minutes. Samples were then sonicated in a water bath for 5 minutes and the supernatant pooled with the previously collected supernatant. 100  $\mu$ l 100% ACN was added to the gels, vortexed and incubated for 5 minutes and the supernatant pooled with the previously collected supernatant. The supernatant was dried at 42°C for 50 minutes using the HPLC method of a Genevac EZ-2 vacuum evaporator.

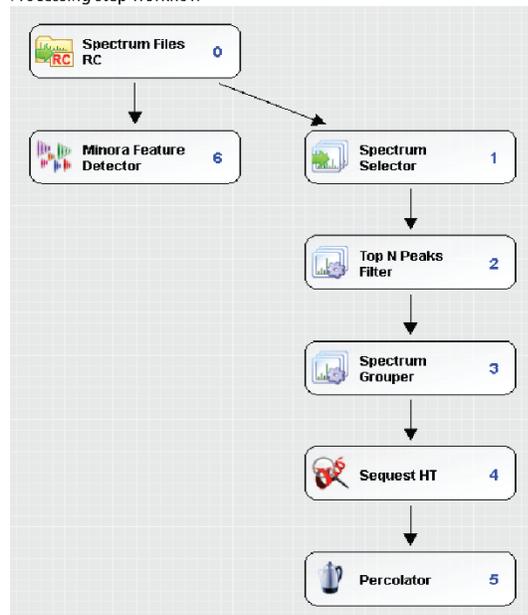
### 3.3.16 LC-MS/MS

100  $\mu$ l resuspension buffer (1% acetic acid, 0.5% formic acid) was added to the dried peptides, vortexed and desalted using C18 filter membranes fitted in a 200  $\mu$ l pipette tip. The C18 filter membranes were initially wetted with 100  $\mu$ l methanol and centrifuged at 3,200 RPM for 1 minute. This was followed by sequential washing and centrifugation with 200  $\mu$ l elution buffer (80% ACN, 0.1% formic acid) and 200  $\mu$ l resuspension buffer. The resuspended peptides were loaded and centrifuged at 3,200 RPM. The membranes were washed with 200  $\mu$ l 0.1% formic acid, and elution carried out with 100  $\mu$ l elution buffer. The eluted peptides were then dried at 42°C for 50 minutes using the HPLC method of a Genevac EZ-2 vacuum evaporator. Finally, they were resuspended in 20  $\mu$ l of 0.1% formic acid. Peptides were analysed by liquid chromatography-nanoelectrospray ionisation -tandem mass spectrometry using an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, Inc).

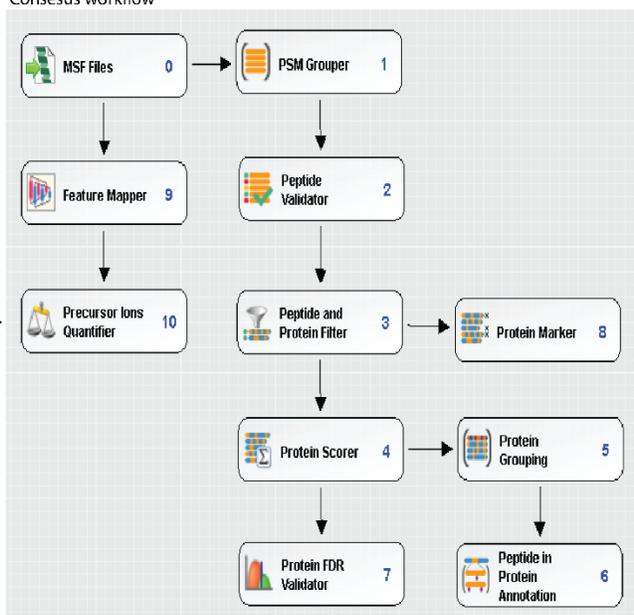
### 3.3.17 Mass spectrometry data analysis

Data-dependent analysis was carried out using Proteome Discoverer software V2.2 (Thermo Fisher Scientific, Inc) with the SEQUEST HT search engine (**Figure 3.2**). Database searches were performed against the Uniprot *X. laevis* database (11/Dec/2019; 57070 entries) and the common Repository of Adventitious Proteins (cRAP; <http://www.thegpm.org/crap/>). Search parameters were; trypsin enzyme, allowing up to 2 missed cleavages, with precursor and fragment mass tolerance set to 20 ppm & 0.8 Da, respectively. Carboxyamidomethylation of cysteine was set as a fixed modification, while methionine oxidation, asparagine, and glutamine deamidation, and N-terminal acetylation were set as variable modifications. The results were filtered using a False Discovery Rate (FDR) of <1% as a cutoff threshold, determined by the Percolator algorithm in Proteome Discoverer software. Functional annotation was based on results from Genecards (<https://www.genecards.org/>). Protein localisation was derived from the Human Protein Atlas (HPA) (<http://www.proteinatlas.org>) (Thul *et al.*, 2017). In addition, BioGRID (<https://thebiogrid.org/>) database was searched to understand any previously identified interactions of the protein with  $\beta$ -catenin (Stark *et al.*, 2006). Manual searches were also conducted using Google Scholar (<https://scholar.google.com/>) to identify any possible interaction with  $\beta$ -catenin for every identified protein.

Processing step workflow



Consensus workflow



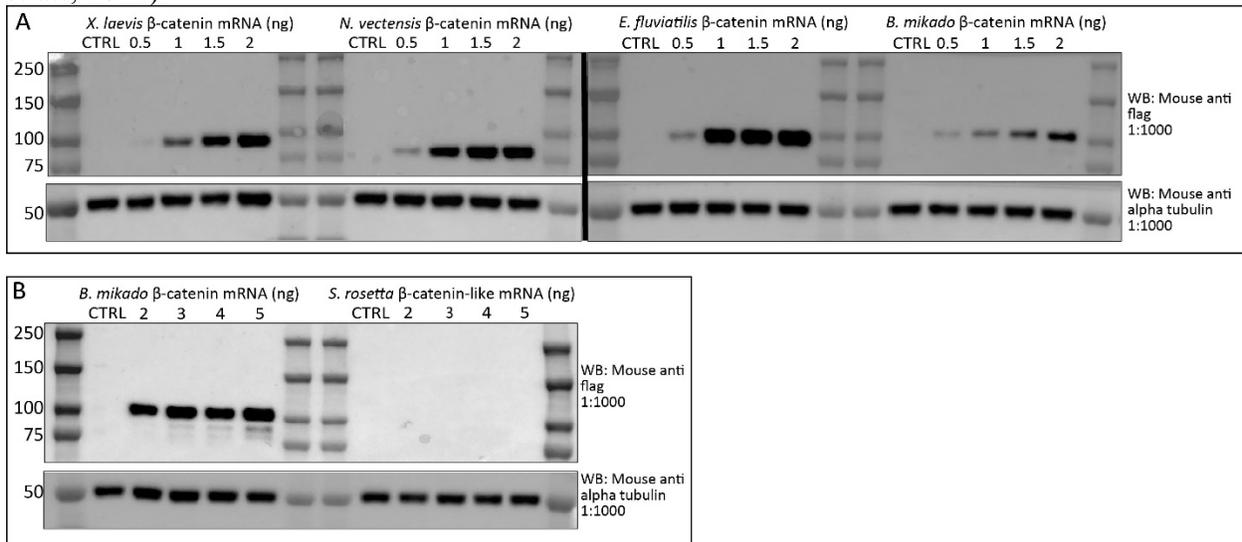
**Figure 3.2. Processing and consensus workflows used in label-free quantification and identification of peptides and proteins.** RAW files obtained from LC-MS/MS analysis were initially analysed in the processing step using the criteria described above and producing compiled MSF files that were passed onto the consensus analysis step. In the consensus step, peptides were validated and matched to respective proteins.

## 3.4 Results

### 3.4.1 Adjustment of expression level and immunoprecipitation of $\beta$ -catenin proteins

Codon differences in mRNA from different organisms can impact the translation rate when expressed in another organism. Therefore, I started by conducting a series of microinjections to optimize and try to attain relatively similar flag-tagged protein expression (**Figure 3.3**). Flag-tagged  $\beta$ -catenin protein was observed at the expected band sizes in all metazoans. For all metazoan  $\beta$ -catenins, protein expression increased with the amount of mRNA injected. However,

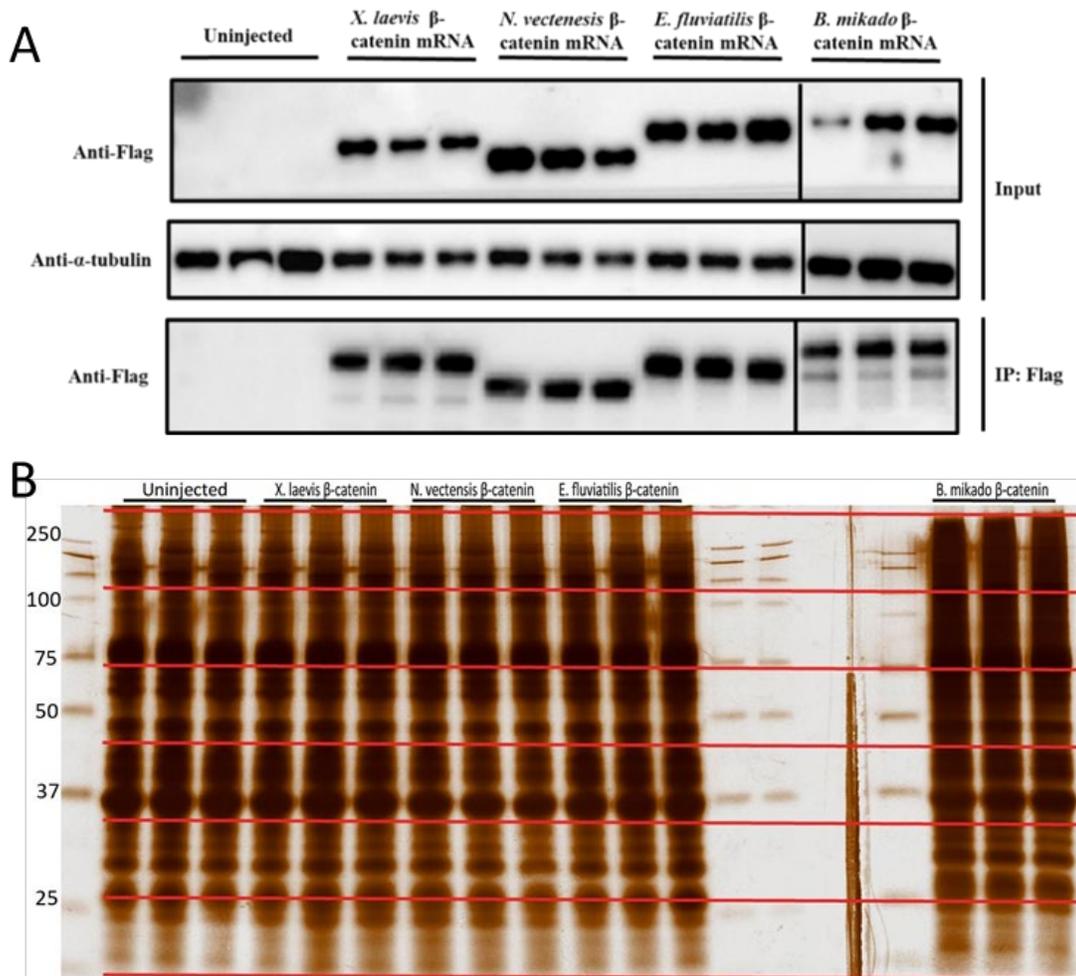
*B. mikado*  $\beta$ -catenin was translated more slowly than the rest of the metazoans; hence would require more mRNA to be injected to make protein expression comparable to that of other metazoans. Unfortunately, no protein expression was observed when *S. rosetta* “ $\beta$ -catenin-like” mRNA was injected, even though codon optimization for expression in the *Xenopus* system had been carried out. It might be that *S. rosetta*  $\beta$ -catenin-like mRNA is translated, but the protein is rapidly degraded, possibly due to its inability to bind E-cadherin, which offers protection from degradation by shielding  $\beta$ -catenin from the destruction complex (Huber and Weis, 2001; Valenta *et al.*, 2012).



**Figure 3.3. Optimisation of  $\beta$ -catenin mRNA for microinjection.** (A) Although the flag-tagged  $\beta$ -catenin expression increased with mRNA concentration, expression (hence translation) was lower for *B. mikado*  $\beta$ -catenin compared to other metazoans. (B) Although codon optimisation had been carried out for *S. rosetta*  $\beta$ -catenin-like mRNA, no bands were observed at all injected mRNA concentrations. Therefore, I did not continue with its analysis.

To execute proteomic characterisation of  $\beta$ -catenin interaction proteins, isolation of exogenous flag-tagged  $\beta$ -catenin was performed. Gastrula stage *Xenopus* embryos were homogenised in a non-reducing buffer to preserve interactions. For immunoprecipitation, anti-flag M2 magnetic beads were used to target the flag-tag added to the N-terminal side of each  $\beta$ -catenin sequence. For negative control, uninjected embryos were incubated with anti-flag M2 beads.

Immunoblotting of elutes revealed efficient precipitation of flag-tagged  $\beta$ -catenin from the different metazoan species, whereas no precipitation was observed in the uninjected control (Figure 3.4). This confirmed that our approach to isolate exogenous  $\beta$ -catenin interactome was successful. However, silver staining of retrieved proteins (elutes) showed a high number of proteins present in control (uninjected) in our metazoan samples. These are unspecific interactions that are either bound to the antibody or the magnetic beads. Therefore, this necessitated rigorous statistical analysis to reduce the presence of false positives (see section 3.3.17 for details).



**Figure 3.4. Immunoprecipitation of flag-tagged metazoan  $\beta$ -catenin.** Gastrula stage *Xenopus* embryos previously injected with  $\beta$ -catenin mRNA were homogenised. **(A)** Expression of flag-tagged  $\beta$ -catenin was confirmed using whole cell lysates (input) with an anti-flag M2 antibody. Similarly, the success of immunoprecipitation using anti-flag M2 magnetic beads was confirmed using the same antibody (IP). No bands were observed in the uninjected embryos. **(B)** Retrieved elutes were then visualised by silver staining of SDS gels. No noticeable difference was observed between the uninjected and the study samples. This means that the antibody and probably the magnetic beads bound many unspecific proteins. Therefore, this necessitated rigorous statistical analysis to reduce the presence of false positives.

### 3.4.2 Proteomic characterisation of $\beta$ -catenin interactome

To identify the proteins interacting with different metazoan  $\beta$ -catenin following expression in *X. laevis* embryos, a proteomic analysis of the immunoprecipitation samples was performed. The proteins were digested using trypsin enzyme and then subjected to LC-MS/MS followed by label-free quantification using Proteome Discoverer. Initially, conventional protein digestion combined with mass spectrometry was carried out. This was designated as the “in-solution” (IS) method. The IS method identified 676 proteins at a 1% false discovery rate (FDR). To increase protein identification, a workflow was designed which increased coverage of unique peptides for proteins. This workflow involved separating proteins by SDS-PAGE and fragmenting each lane into six

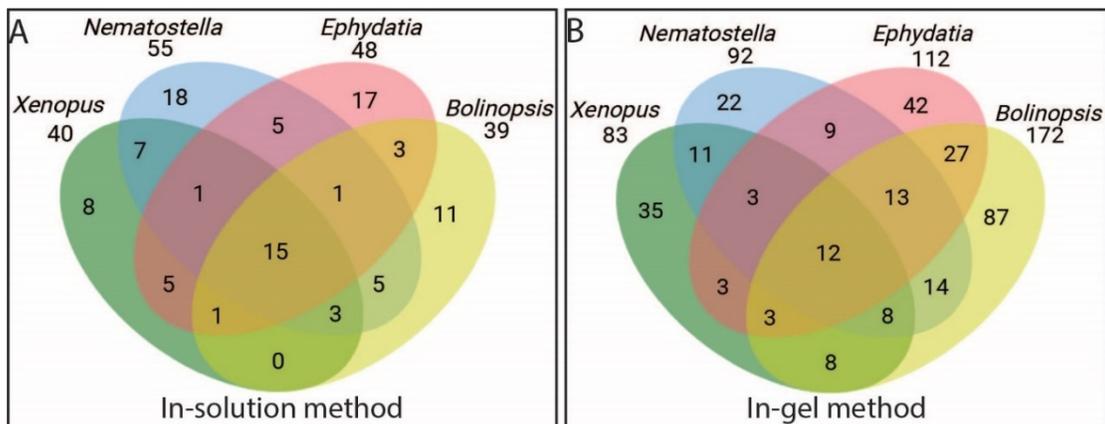
fractions independently digested in trypsin and prepared for LC-MS/MS. This approach was designated the “in-gel” (IG) method. The IG method revealed 2106 proteins at 1% FDR. This was thrice as many as those obtained when using the IS method. Both methods confirmed significant enrichment of exogenous  $\beta$ -catenin from each metazoan that had been expressed (**Table 3.5**).

Protein	IS trypsin digestion method				IG trypsin digestion method			
	PSMs	Uniq Peps	AR	P-value	PSMs	Uniq Peps	AR	P-value
<i>X.lae</i> $\beta$ -cat	262	26	26.58	4.4E-16	1956	9	100	1E-17
<i>N. vec</i> $\beta$ -cat	305	23	100	1E-17	1982	47	100	1E-17
<i>E. flu</i> $\beta$ -cat	179	24	97.67	1E-17	1199	41	35.5	3.5E-12
<i>B. mik</i> $\beta$ -cat	195	27	100	1E-17	2006	59	72.53	1.9E-09

**Table 3.5. Confirmation of the enrichment of exogenous  $\beta$ -catenin.** Both methods confirmed the presence of exogenous  $\beta$ -catenin protein. However, better identification was obtained when using the IG method as evidenced from the higher peptide spectrum matches (PSMs), abundance ratios (AR) and unique peptide. The low unique peptide observed in the for *X. laevis*  $\beta$ -catenin could be due to a high similarity with the endogenous  $\beta$ -catenin. Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), *B. mik* (*B. mikado*), and Uniq Peps (Unique peptides)

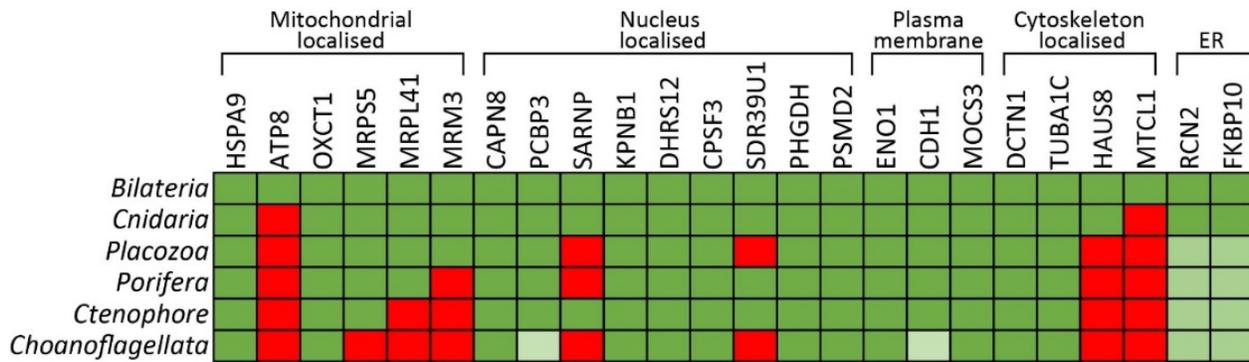
### 3.4.3 Evolutionarily conserved $\beta$ -catenin protein interactions in metazoans

To reduce false positives in identified proteins from the IS and IG methods, only proteins with an abundance ratio greater than or equal to 2 ( $p$ -value < 0.05) are generally used as a threshold for “true” interactions (**Figure 3.5**). First, I analysed the *X. laevis*  $\beta$ -catenin and confirmed the presence of known  $\beta$ -catenin interacting proteins, e.g., ARVCF, APC, JUP, FMR1, CTNND1, CTNNA1, CDH1 etc. This further assured me that most of the identified proteins across the mass spectrometry analysis were “true” interactions.



**Figure 3.5. Venn diagrams of common proteins interacting with  $\beta$ -catenin from four different metazoans.** At a  $p$ -value < 0.05 and a fold ratio to control greater or equal to 2, the number of high confidence (1% FDR) protein interactions (16) was greater in the IS method (A) than in the in-gel method (12) (B) even though the in-gel method had resulted in greater protein identification. Details of the proteins in each quadrant are in appendices 5 and 6.

From the IS (15) & IG (12) methods, the 27 UniProt IDs equated to 24  $\beta$ -catenin protein interactions were conserved in the four metazoan models used in this study (**Figure 3.6**). Except for ATP8 and MTCL1, which were *Bilateria* specific, Blast searches confirmed the presence of homologs in other metazoan phyla. Here, I will present, based on their localization as obtained from the Human Protein Atlas, known & previously unknown interactions.



**Figure 3.6. Potential evolutionarily conserved  $\beta$ -catenin protein interactions.** From the IS and IG trypsin digestion methods, the mass spectrometric analysis identified 24 proteins interacting with the injected metazoan  $\beta$ -catenin. Blast database searches were then carried to check for the presence of homologs in the respective phylum. Green indicates the homolog is present. Lighter shades of green indicate reduced conservation, while red indicates the absence of homolog.

### Mitochondria localized proteins

The majority of these proteins were identified using the IS method (Table 3.6). Only HSPA9 has been previously shown to interact with  $\beta$ -catenin (Tian *et al.*, 2004). Also, recently it has been shown that HSPA9 suppression inhibits  $\beta$ -catenin signalling in ovarian carcinoma (Xu *et al.*, 2019). As homologs of HSPA9 were found in other metazoans, a similar mechanism of  $\beta$ -catenin regulation may be present in basal metazoans but under a different context. Other proteins were novel interactions with little knowledge about their function and possible relationship with  $\beta$ -catenin. Future research could focus on understanding these proteins' function and how they relate or interact with  $\beta$ -catenin.

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
Q7ZX34	Heat Shock Protein Family A Member 9	3.47	3.64	3.11	4.11	1.59E-03	4.09E-03	1.13E-02	1.64E-02
F5KDD9	ATP synthase protein 8	3.07	3.16	3.98	4.49	4.26E-03	1.07E-02	2.01E-03	1.05E-02
A0A1L8I381	Succinyl-CoA:3-ketoacid coenzyme A transferase 1	4.26	4.09	3.96	4.76	2.54E-04	1.74E-03	2.10E-03	7.84E-03
A0A1L8G7F0	Mitochondrial Ribosomal Protein S5	2.92	4.16	3.47	3.52	6.12E-03	1.52E-03	5.50E-03	3.32E-02
Q6DJI4	Mitochondrial Ribosomal Protein L41	5.87	7.88	5.54	7.67	3.93E-05	3.21E-05	1.28E-04	4.59E-04
A0A1L8HG50	RNA Methyltransferase Like 1	27.21	16.80	28.38	13.65	4.69E-12	1.12E-07	2.77E-05	9.27E-04

**Table 3.6. Mitochondria localised  $\beta$ -catenin protein interactions conserved in 4 different metazoans following  $\beta$ -catenin mRNA expression in *X. laevis* embryos.** Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

Nucleus localized proteins

$\beta$ -catenin's role in transcription regulatory role in the nucleus has been highly researched over many years. Here, I found that nine protein interactions were conserved in the metazoans used in this study (**Table 3.7**). Except for SARNP and SDR39U1, all proteins had homologs present in all metazoans (**Figure 3.6**). Among these proteins, only KPNB1 has previously been indicated to interact with  $\beta$ -catenin. A recent study confirmed by immunoprecipitation the interaction of KPNB1 with  $\beta$ -catenin and showed that knockdown of KPNB1 resulted in decreased  $\beta$ -catenin expression in the nucleus (Lu *et al.*, 2016). KPNB1 could therefore be essential in the transportation of  $\beta$ -catenin into the nucleus, and as its homologs were found in basal metazoans, this relationship with  $\beta$ -catenin could be the ancient mechanism through which  $\beta$ -catenin transport to the nucleus occurs. Although there is little knowledge about most of these novel interactions, a recent showed that the overexpression of DHSR12 results in the inactivation of the  $\beta$ -catenin signalling pathway, thereby inhibiting proliferation in osteosarcoma (Xu *et al.*, 2020). Therefore, it would be interesting to investigate whether this function is conserved in basal metazoans and under what context.

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
A0A1L8G6E2	Calpain 8	4.39	5.00	4.05	5.53	1.89E-04	3.34E-04	1.77E-03	3.45E-03
A0A1L8HA56	Poly(RC) Binding Protein 3	7.00	3.85	7.25	11.27	1.13E-06	2.72E-03	9.15E-06	2.89E-05
A0A1L8HHY7	SAP Domain Containing Ribonucleoprotein	9.35	12.83	14.63	22.66	2.54E-08	1.08E-08	1.76E-09	6.09E-08
P52297	Importin subunit beta-1	2.45	2.84	2.77	5.53	2.06E-02	2.10E-02	2.20E-02	3.43E-03
A0A1L8HIR4	Dehydrogenase /Reductase 12	24.55	16.03	27.99	32.94	1.05E-09	5.87E-06	7.47E-05	5.91E-06
A0A1L8G578	Cleavage And Polyadenylation Specific Factor 3	5.91	8.58	8.46	21.72	5.50E-03	8.22E-03	3.97E-02	1.90E-04
A0A1L8HPP5	Short Chain Dehydrogenase/Reductase Family 39U Member 1	3.01	6.04	14.19	19.63	2.65E-02	4.23E-04	4.06E-06	1.37E-04
A0A1L8ER90	D-3-phosphoglycerate dehydrogenase	23.06	3.82	8.90	8.01	3.27E-13	2.09E-02	7.53E-03	1.00E-02
A0A1L8G3M7	26S proteasome non-ATPase regulatory subunit 2	6.91	7.02	12.86	18.84	2.70E-04	1.12E-03	1.38E-03	1.72E-04

**Table 3.7. Nucleus localised  $\beta$ -catenin protein interactions conserved in 4 different metazoans following  $\beta$ -catenin mRNA expression in *X. laevis* embryos.** Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

Plasma membrane localised proteins

In the plasma membrane,  $\beta$ -catenin is a cadherin catenin complex (CCC) component, essential in cell adhesion. One component of the CCC, E-cadherin, was identified to interact with all metazoan  $\beta$ -catenins (**Table 3.8**). This could imply that this interaction was present in the last common of metazoans. However, it does not indicate whether the interaction is functional and necessary in basal metazoans as ctenophore E-cadherin has a divergent catenin binding domain that may impede interaction with  $\beta$ -catenin (Belahbib *et al.*, 2018). Surprisingly,  $\alpha$ E-catenin (another vital component of the CCC) was only precipitated with *X. laevis*, *N. vectensis* & *B. mikado*  $\beta$ -catenin. Detailed analysis of these protein interactions will be described in the next chapter.

Two novel interactions, MOCS3 and ENO1, were also identified. Although little is known about their relationship with  $\beta$ -catenin, recent research has shown ENO1 regulates  $\beta$ -catenin-driven trans-differentiation of murine alveolar epithelial cells (Mutze *et al.*, 2015). Therefore, as a pan-metazoan protein, ENO1 could probably be functional in basal metazoans, working with  $\beta$ -catenin in a cell-specific context.

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
A0A1L8FLY7	Enolase 1	32.37	30.59	2.47	49.84	1.00E-17	1.55E-14	4.43E-02	9.85E-12
A0A1L8GET9	E-cadherin	46.95	54.23	11.48	41.68	1.00E-17	1.00E-17	2.52E-03	1.20E-06
P33152		58.90	79.19	12.36	42.48	1.00E-17	1.00E-17	1.98E-05	1.06E-06
Q58E95	Molybdenum Cofactor Synthesis 3	8.08	8.03	16.10	37.86	2.05E-04	8.61E-04	1.49E-03	2.34E-06

**Table 3.8.** Plasma membrane localised  $\beta$ -catenin protein interactions conserved in 4 different metazoans following  $\beta$ -catenin mRNA expression in *X. laevis* embryos. Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

Cytoskeleton associated proteins.

Although  $\beta$ -catenin is linked to the cytoskeleton through  $\alpha$ E-catenin, none of the proteins identified have been previously identified to interact with  $\beta$ -catenin (**Table 3.9**). That way, it is not easy to speculate the possible function of the interaction. Furthermore, like any other interaction, this could be an indirect interaction through an intermediate protein binding to  $\beta$ -catenin. Nevertheless, as these results were obtained from strict statistical filtration, these interactions could be investigated in the future.

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
Q6PCJ1	Dynactin Subunit 1	8.25	4.05	3.74	7.48	6.91E-07	3.97E-03	3.21E-03	5.42E-04
Q6NRV3	Tubulin Alpha 1C	3.50	5.59	3.35	6.93	1.47E-03	1.23E-04	6.91E-03	8.92E-04
Q5U4V6		3.01	6.81	2.50	4.46	4.91E-03	1.83E-05	4.17E-02	1.09E-02
A0A1L8HXE0	HAUS Augmin Like Complex Subunit 8	6.14	4.28	5.31	8.00	6.63E-04	1.73E-02	4.42E-02	1.01E-02
A0A1L8FYI1	Microtubule Crosslinking Factor 1	3.14	7.13	22.02	10.28	1.57E-02	1.18E-04	1.71E-06	3.51E-03

**Table 3.9. Cytoskeleton associated  $\beta$ -catenin protein interactions conserved in 4 different metazoans following  $\beta$ -catenin mRNA expression in *X. laevis* embryos *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).**

#### Endoplasmic reticulum (ER) localized proteins

The ER is involved in various cellular functions, including protein synthesis, folding & -transportation, lipid & steroid synthesis, and calcium storage. HPA database search showed that only two proteins, RCN2 and FKBP10 were localised in the ER, and they had no known prior interaction with  $\beta$ -catenin (**Table 3.10**).

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
Q6IP80	Reticulocalbin 2	6.79	6.62	6.73	75.59	4.55E-14	1.97E-05	2.42E-05	1.65E-06
A9JS50	FKBP Prolyl Isomerase 10	3.82	3.30	7.10	13.24	1.28E-02	4.34E-02	1.44E-02	1.08E-03

**Table 3.10. ER localised  $\beta$ -catenin protein interactions conserved in 4 different metazoans following  $\beta$ -catenin mRNA expression in *X. laevis* embryos.** Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

#### 3.4.4 $\beta$ -catenin protein interactions shared between *Bilateria* and *Cnidaria*

Due to a high percentage amino acid sequence identity between *X. laevis* and *N. vectensis*  $\beta$ -catenin, combined with conservation of critical domains and structures (see section 2.4 for details), I expected several previously known interactions to be shared between these two metazoans. From both methods, 13 protein interactions were common in *X. laevis* and *N. vectensis*  $\beta$ -catenin (**Table 3.11**). The majority of these were novel protein interactions. Five (APC, CTNND1, CTNNBIP1/ICAT, and JUP) proteins have previously been shown to interact with  $\beta$ -catenin. APC is a vital component of the destruction complex with homologs in cnidarians and poriferans (Schenkelaars *et al.*, 2017). However, in ctenophore, *M. leidy*, *in silico* searches showed a partial APC lacking some domains (Ryan *et al.*, 2013). An in-depth analysis of APC and its binding region in  $\beta$ -catenin will be discussed in a later chapter.

Uniprot ID	Protein	Abundance ratio				P-value			
		<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat	<i>X. lae</i> $\beta$ -cat	<i>N. vec</i> $\beta$ -cat	<i>E. flu</i> $\beta$ -cat	<i>B. mik</i> $\beta$ -cat
G5E9G0	Ribosomal protein L3	9.17	7.67	0.77	0.29	2.00E-09	1.90E-05	6.23E-01	4.54E-02
P70039	Adenomatous polyposis coli	13.52	10.65	1.02	1.26	6.34E-09	8.66E-06	9.73E-01	9.19E-01
A0A1L8I1P0		14.28	10.57	0.98	2.57	3.05E-08	5.82E-06	9.39E-01	3.84E-01
Q5MNU4	Staufen Double-Stranded RNA Binding Protein 1	3.16	3.45	2.88	4.21	8.13E-03	1.46E-02	1.53E-01	9.22E-02
Q2TAT7	Cytochrome b-c1 complex subunit 8	3.06	3.05	0.07	0.05	1.58E-02	3.08E-02	3.85E-06	4.69E-06
A0A1L8FB29	Nucleolar and spindle-associated protein 1	5.48	3.22	1.19	1.07	3.48E-04	3.12E-02	8.57E-01	8.33E-01
Q8AXM9	Catenin delta-1	5.15	3.98	1.25	3.41	8.57E-05	2.99E-03	8.52E-01	1.63E-01
A0A1L8EKG6	Junction plakoglobin	3.07	5.24	2.87	4.44	3.15E-02	1.78E-03	2.47E-01	7.85E-02
A0A1L8F6I3	Heat Shock Protein Family A (Hsp70) Member 5	3.73	3.64	1.95	2.88	3.04E-04	1.29E-03	2.69E-01	2.46E-01
P49739	Minichromosome Maintenance Complex Component 3	2.67	3.18	0.81	2.46	1.17E-02	1.03E-02	6.26E-01	1.36E-01
A0A1L8GNF1	RAB11 Family Interacting Protein 4	2.86	3.08	2.56	1.86	1.89E-02	4.93E-02	N/A	N/A
A0A1L8HTV8	Pericentriolar Material 1	2.49	3.36	2.31	2.45	1.85E-02	7.21E-03	6.26E-02	1.38E-01
Q6DED3	Catenin Beta Interacting Protein 1	22.48	20.39	0.73	2.38	1.33E-14	1.32E-11	N/A	1.51E-01

**Table 3.11.  $\beta$ -catenin protein interactions that are shared between *Bilateria* and *Cnidaria*  $\beta$ -catenin.** Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

Although precipitated by both *X. laevis* and *N. vectensis*  $\beta$ -catenin, the proteins JUP and CTNND1 (p120) are catenin family members specific to bilaterians, particularly in chordates (Zhao *et al.*, 2011). They are both essential in cell-cell adhesion, where CTNND1 stabilises E-cadherin at the adherens junction (Xiao *et al.*, 2007), while JUP is a critical protein in desmosomal junctions (Green *et al.*, 2020). CTNNBIP1 is a negative regulator of  $\beta$ -catenin signalling as it blocks the interactions of  $\beta$ -catenin to TCF (Qi *et al.*, 2015). However, I carried out a database search and found that it is a bilaterian specific protein. These results imply that the  $\beta$ -catenin protein in the last common ancestor of bilaterians and cnidarians had already established the machinery to bind these proteins. However, as they are present only in bilaterians, it implies they probably emerged (evolved) to drive morphological or another functional novelty.

The rest of the proteins were novel interactions. Of these, NUSAP1, through its upregulation, was shown to stimulate metastasis of cervical carcinoma cells by activating  $\beta$ -catenin signalling (Li *et al.*, 2019). I carried out an *in silico* search of NUSAP1, and it revealed a homolog is present in cnidarians hence may be functional.

### 3.5 Discussion

#### 3.5.1 The evolutionarily conserved metazoan $\beta$ -catenin machinery

As detailed in section 2.4, although there is varying amino acid conservation within some regions, the basic structure of  $\beta$ -catenin in metazoans has been conserved. Therefore, this implies that the ability of exogenous  $\beta$ -catenin to bind endogenous *Xenopus* protein will mainly function the essential interacting residues (**Figure 1.6**). In this study, I designed a transphylectic approach combined with proteomic analysis to identify conserved  $\beta$ -catenin protein interactions across metazoans. Precipitation of elements of adherens junction & other known  $\beta$ -catenin interacting proteins, e.g., APC, JUP (Hwang *et al.*, 2017), FMR1 (Ehyai *et al.*, 2018) with *Xenopus*  $\beta$ -catenin, indicates that the results obtained could be representative of the actual interactions at the gastrula stage. Therefore, this increases the probability that the precipitated  $\beta$ -catenin interacting proteins are “true” interactions. Other than the adherens junction elements, most of these were novel interactions, and it is not easy to understand their roles in basal metazoan embryo development. Therefore, investigations into some of these protein interactions could set a new direction for understanding the new roles of  $\beta$ -catenin in development. Of these, the presence of mitochondria localised proteins was the most surprising.

Several studies have shown that mitochondria play vital roles in various cellular processes like apoptosis, fatty acid & nucleotide biosynthesis etc. Their importance in the course of evolution is that all eukaryotes have them, and the parasites that lack mitochondria or mitochondria-like organelles possibly lost them as they evolved (Karnkowska *et al.*, 2016; Zachar and Szathmáry, 2017). However, regardless of their importance, there are still some questions, including how the regulatory factors of mitochondria acquired their mitochondrial role & localization. Furthermore, it remains to be investigated how some regulatory factors are found in both the nucleus and mitochondria function (Medini *et al.*, 2020). The results of mitochondria localized  $\beta$ -catenin interacting proteins might begin to shed some light on these questions.

$\beta$ -catenin lacks a nuclear localization signal (NLS); hence its translocation to the nucleus has remained enigmatic. However, I found KPNB1 (Karyopherin (Importin) Beta 1) association with  $\beta$ -catenin to be conserved across the four metazoans. Previous work noted that  $\beta$ -catenin’s armadillo repeats share striking structural similarities to the HEAT repeats of importin beta proteins, suggesting that it may interact with the Fg-repeats of nucleoporins (Xu and Massagué, 2004; Fagotto, 2013). Furthermore, a recent study found decreased  $\beta$ -catenin expression in the nucleus of cells depleted of KPNB1 (Lu *et al.*, 2016). Therefore, this suggests that KPNB1 might be involved in the nuclear localization of  $\beta$ -catenin. Database searches were carried and found that KPNB1 was evolutionarily conserved in metazoans. Therefore, KPNB1 could be the conventional shuttle system used to translocate  $\beta$ -catenin to the nucleus. However, confirmation of this would need functional analysis across different metazoans.

In conclusion, the evolutionarily conserved  $\beta$ -catenin protein machinery in metazoans comprises a diverse range of proteins. How these proteins function might vary between phyla. Future research could focus on performing analysis on some of these proteins in both bilaterians and basal

metazoans to try to establish functional roles and how they might relate to  $\beta$ -catenin's dual functions.

### 3.5.2 Novel genes drove morphological and functional novelty in bilaterians

In section 3.4.4, we saw that some critical proteins, e.g., JUP, CTNND1 and CTNNBIP1, interacted with *X. laevis* and *N. vectensis*  $\beta$ -catenin. Previous research has shown that JUP and CTNND1 are specific to chordates only (Zhao *et al.*, 2011). Furthermore, Blast searches confirmed that CTNNBIP1 is bilaterian specific. These create two implications, i.e., 1) architectural design (structure and amino acid sequence) of  $\beta$ -catenin in the last common ancestor of bilaterians and cnidarians was complete. As a result of this, cnidarian  $\beta$ -catenin can bind these bilaterian specific proteins. 2) Novel genes evolved in bilaterians to enhance or finetune context-specific roles while interacting with  $\beta$ -catenin. For example, CTNNBIP1 knockdown significantly inhibits cell apoptosis resulting in the over-proliferation of keratinocytes (Wang *et al.*, 2021). CTNND1 (also known as p120) is a multifunctional protein that stabilises E-cadherin in the adherens junctions (Ishiyama *et al.*, 2010), and interacts with ZBTB33 (Kaiso), a transcriptional repressor to facilitate Wnt signal transduction (Park *et al.*, 2005; del Valle-Pérez *et al.*, 2011). In a recent study, CTNND1 knockdown in *Xenopus* resulted in craniofacial and heart defects, thereby illustrating its importance in embryo development in chordates (Alharatani *et al.*, 2020). JUP (plakoglobin) is highly homologous to  $\beta$ -catenin. Its structural role is well established as a component of the desmosomal junction (Green *et al.*, 2020). There is, however, a significant subject of study on whether JUP is involved in transcription regulation. Early work using NCL-H28 cells showed that JUP could elicit Wnt signalling independently without interacting with  $\beta$ -catenin (Maeda *et al.*, 2003). However, a later study demonstrated that JUP was insufficient for TCF-dependent transcriptional activation by Wnt-3a in F9 cells (Shimizu *et al.*, 2008). To understand whether these novel genes are only functional in bilaterians, future studies could perform knockin studies in basal metazoans, particularly cnidarians and study the effect on embryonic development.

### 3.5.3 Challenges of study

The novelty & attraction of the approach used in this study (using whole embryos) also resulted in some difficulties, making the use of cell lines in interactomics attractive. One challenge faced was the variance between biological replicates. Although some of this variance could have been brought about by the preparation technique (difficult to homogenize and prepare samples precisely the same way), the leading cause was down to the embryos themselves. As a result, there were differences in abundances between biological replicates, which affected the p-value. Therefore, although several identified proteins had a qualifying abundance ratio at 1% FDR, they were filtered out due to their insignificant p-value. There is, therefore, a high possibility that some true interactions were negated.

On the other hand, strict statistical filtration reduced the presence of possible false positives. To further reduce the presence of false positives, crosslinking was not carried out. Chemical crosslinking allows the strengthening of transient or weak protein interactions. However, some proteins end up being crosslinked just by being nearby, resulting in many false positives. Nevertheless, the precipitation of known interactions gives us hope that several interactions were not false positives.

### 3.6 Future improvements & chapter conclusion

In summary, our study has uncovered the evolutionarily conserved  $\beta$ -catenin machinery made up of a wide range of proteins. As most of these proteins are conserved in metazoans, there is a

possibility that they were crucial in  $\beta$ -catenin's dual functions leading up to multicellularity. As most of the proteins we uncovered were novel interactions, future works could identify their respective roles in embryo development.

Although this study was conducted on gastrula stage embryos, this method could be used on later stages, e.g., neurula stage; hence allowing identification of more stage-specific interactions. For the novel interactions identified in this work, it would be necessary to validate them. Validation of interaction with a protein of interest is crucial before concluding that an observed interaction is actual. However, one problem when using *X. laevis*, is that not all commercial antibodies target proteins perfectly. This was one of the main difficulties that we faced. Furthermore, although the IG method was more superior to the IS method, there is still room for improvement on the IG method. A recent study found that introducing Lys-C treatments before & during tryptic digestion along with improvement of peptide separation using electrostatic repulsion hydrophilic interaction chromatography significantly increased unique peptides and protein identification (Taoufiq *et al.*, 2021).

The subsequent chapters of this thesis will focus on three proteins, i.e., TCF, E-cadherin and  $\alpha$ E-catenin, which are vital in  $\beta$ -catenin's dual functionality and were identified in the mass spectrometry analysis.

# CHAPTER 4

## Chapter 4 Conservation of the cadherin-catenin interaction complex

### 4.1 Summary

Here, my results revealed domain conservation of E-cadherin binding site in  $\beta$ -catenin, and equally, E-cadherin was precipitated with all metazoan  $\beta$ -catenin.  $\alpha$ E-catenin was not precipitated by sponge  $\beta$ -catenin, possibly due to domain differences identified around a critical binding site. Nevertheless, our results suggest that the adhesive role of  $\beta$ -catenin was present in the last common ancestor of metazoans.

### 4.2 Background

Cell-cell adhesion has been intimately linked with the emergence of multicellularity (Abedin and King, 2010; Knoll, 2011). It is, therefore, crucial to unravel how cell-cell adhesion emerged and its role in metazoan diversity. Early research into cell-cell adhesion showed that it was calcium-dependent in *Porifera* (Maas, 1906). In metazoans, cell-cell contact between epithelial cells is essential in the organization and formation of tissues and is mainly mediated by tight junctions, adherens junctions, and desmosomes. Mediation through the adherens junction requires the interaction of E-cadherin,  $\beta$ -catenin, and  $\alpha$ E-catenin (see **Figure 1.5**). Together, these form the cadherin catenin complex (CCC). In the presence of calcium, opposing cells' E-cadherin's calcium-dependent domains (CADs) homophilically interact in the extracellular space while the cytoplasmic cadherin domain interacts with  $\beta$ -catenin.  $\alpha$ E-catenin is the adapter protein that interacts with  $\beta$ -catenin hence facilitating interaction with the actin cytoskeleton. CCC protein homologs have been identified in several metazoans whose genomes or transcriptomes have been sequenced. This could imply that the CCC adhesion complex was already present in the last common ancestor of all metazoans. Indeed, many cadherin-like homologs have been identified in unicellular choanoflagellates, *S. rosetta* (Nichols *et al.*, 2012) & *M. brevicolis* (Abedin and King, 2008). However, although they exhibit the CAD domains, they lack cytoplasmic domains that facilitate binding with  $\beta$ -catenin.

Until recently, functional knowledge about the CCC adhesion complex has come from bilaterian models. Early research found that an E-cadherin mutant lacking the  $\beta$ -catenin binding domain (CBD) cannot interact with other catenins resulting in failed cell adhesion (Nagafuchi and Takeichi, 1988). Similarly,  $\alpha$ E-catenin mutations in bilaterian embryos have resulted in cell adhesion defects and developmental arrest (Kofron *et al.*, 1997; Costa *et al.*, 1998; Sarpal *et al.*, 2012).  $\beta$ -catenin knockdown and mutations have also illustrated its importance in cell adhesion, embryo development and tissue maintenance (Haegel *et al.*, 1995; Messerschmidt *et al.*, 2016). Aside from the functional analysis, several studies have proved the interaction and colocalization of the CCC adhesion complex components (Tian *et al.*, 2004; Guo *et al.*, 2014; Semaan *et al.*, 2019). However, there is a still lack of understanding about the nature of the CCC adhesion complex at the early phases of animal evolution, especially the role of  $\beta$ -catenin.

To this end, recently, researchers have begun investigating the nature of the epithelium and CCC adhesion complex in basal animals. It was recently demonstrated in *Nematostella* that  $\alpha$ E-catenin plays a role in embryo development (Clarke *et al.*, 2019) and that cadherins are also crucial in germ layer formation (Pukhlyakova *et al.*, 2019). In *Porifera*, much research has been sequence-based. However, a recent study using *E. muelleri* showed that the interaction of  $\beta$ -catenin with CDH2 (N-cadherin) &  $\alpha$ E-catenin is conserved (Schippers and Nichols, 2018). Furthermore, it showed that  $\beta$ -catenin is localised at cell contact points, suggesting that its structural role is probably conserved in poriferans (Schippers and Nichols, 2018). Nevertheless, it is still unknown

whether the CCC adhesion complex in basal metazoans, particularly poriferans and ctenophores, is functional.

Through a comprehensive domain and proteomic analysis, I attempted to answer when the structural role of  $\beta$ -catenin in the CCC emerged.

### 4.3 Methods

As described in sections 2.3 and 3.3.

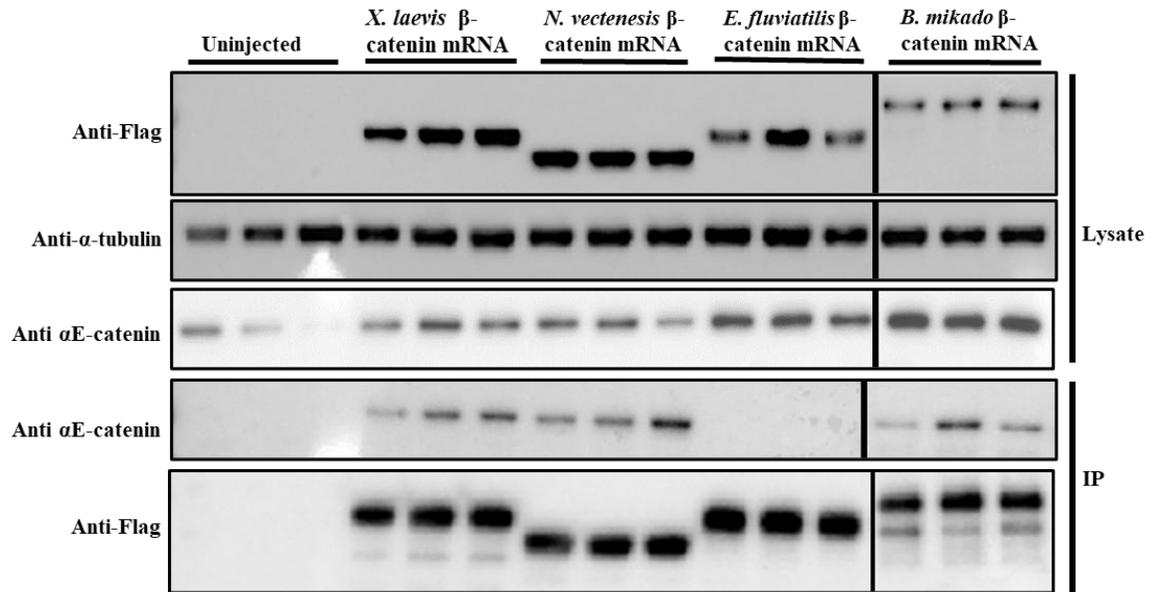
### 4.4 Results

#### 4.4.1 Evolutionary conservation of $\beta$ -catenin interaction with $\alpha$ E-catenin.

The  $\alpha$ -catenin family consists of four members, i.e., alpha E, N, T and catulin, which play crucial and specific functions in development and differentiation (Kobielak and Fuchs, 2004). Previous research using mass spectrometry confirmed the interaction of  $\beta$ -catenin to  $\alpha$ E-catenin in mammalian cell culture (Tian *et al.*, 2004). Our mass spectrometry analysis of affinity-purified metazoan flag-tagged  $\beta$ -catenin was able to detect  $\alpha$ E-catenin but none of the other  $\alpha$ -catenin family members. However, the detection was confined only to *X. laevis*, *N. vectensis* & *B. mikado*  $\beta$ -catenins and not *E. fluviatilis* (**Table 4.1**). A validation western blot on the elute confirmed the mass spectrometry results as no bands was observed in elutes of *E. fluviatilis* and the uninjected (IgG) control. In contrast, the other lanes had a band corresponding to the size of  $\alpha$ E-catenin (**Figure 4.1**).

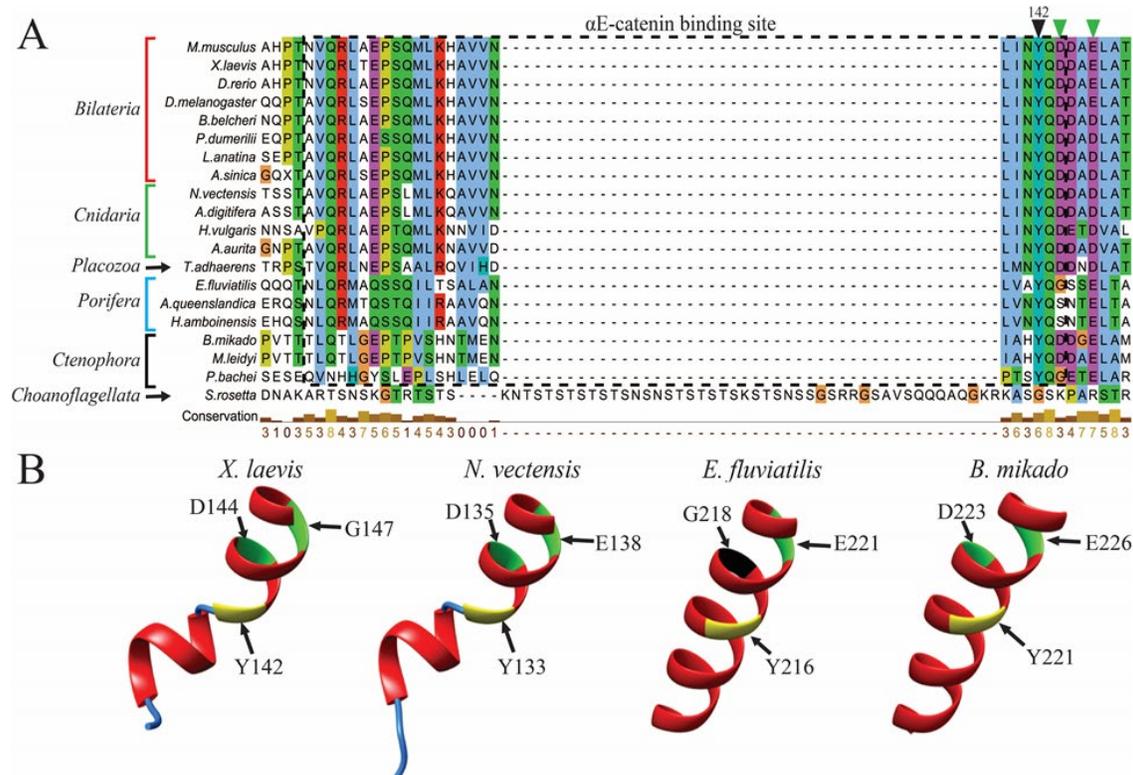
		$\alpha$ E-catenin								
		Unique peptides	Abundance ratio				P-value			
Uniprot ID	PSMs		<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin	<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin
Q91682	826	22	30.20	27.65	1.23	12.97	1.00E-17	1.33E-15	8.11E-01	1.19E-03
A0A1L8GWT0	611	17	12.64	8.54	0.83	5.30	2.24E-10	1.04E-05	8.20E-01	4.50E-02

**Table 4.1. Coimmunoprecipitation of endogenous  $\alpha$ E-catenin by metazoan  $\beta$ -catenin.** Transphylectic expression of metazoan  $\beta$ -catenin followed by affinity purification and proteomic analysis detected an abundant presence of  $\alpha$ E-catenin peptides in all metazoan  $\beta$ -catenins except for *E. fluviatilis*  $\beta$ -catenin. Results are from three biological replicates.



**Figure 4.1. Validation of co-immunoprecipitation of  $\alpha$ E-catenin with  $\beta$ -catenin.** Flag-tagged  $\beta$ -catenin mRNA was expressed in 1-cell stage *Xenopus* embryos. At gastrula stage, protein lysates were immunoprecipitated with anti-flag M2 magnetic beads and eluate subjected to western blot with the anti- $\alpha$ E-catenin & anti-flag M2 antibodies. The  $\alpha$ -tubulin was used as a loading control from the lysates. Although present in all lysates,  $\alpha$ E-catenin was not immunoprecipitated in the case of flag-tagged *E. fluviatilis*  $\beta$ -catenin.

To understand why *E. fluviatilis*  $\beta$ -catenin failed to precipitate endogenous *Xenopus*  $\alpha$ E-catenin, I analysed the metazoan  $\beta$ -catenin alignment focusing on the  $\alpha$ E-catenin binding site. The  $\alpha$ E-catenin binding site in  $\beta$ -catenin was previously narrowed down to amino acids 118-146 (in mouse) (Aberle *et al.*, 1996). Metazoan-wide conservation across this region was not high except for a few residues (**Figure 4.2**). In mouse  $\beta$ -catenin, Tyrosine-142 (Y142) is vital for  $\alpha$ E-catenin binding as when mutated to alanine,  $\beta$ -catenin- $\alpha$ E-catenin interaction is obliterated (Aberle *et al.*, 1996). The site corresponding to Y142 was conserved in all metazoans, including *E. fluviatilis* (**Figure 4.2**). However, two acidic residues, Asp-144 (D144) and Glu-147 (E147), which are close to the hydroxyl group of Y142, have previously been shown to affect  $\alpha$ E-catenin interaction (Piedra *et al.*, 2003). I found that the residue corresponding to D144 was not conserved among poriferans (**Figure 4.2**). Furthermore, structural analysis around the region surrounding Y142 ruled structural differences as a reason for failure to precipitate  $\alpha$ E-catenin because it showed that *E. fluviatilis* and *B. mikado*  $\beta$ -catenin had similar structures. Therefore, the failure of *E. fluviatilis*  $\beta$ -catenin to bind endogenous *Xenopus*  $\alpha$ E-catenin could be due to a lack of conservation at the residue corresponding to D144.



**Figure 4.2. The  $\alpha$ E-catenin binding region in  $\beta$ -catenin.** **A**, Alignment of  $\beta$ -catenin indicating the suggested region where  $\alpha$ -catenin binds (dotted black box). The Y142 residue (black triangle) critical in this interaction was conserved in all metazoans, including placozoan, *T. adhaerens*. Green arrows represent acidic residues that can inhibit  $\alpha$ E-catenin binding by affecting charge distribution (Piedra *et al.*, 2003). **B**, predicted structure of  $\beta$ -catenin focusing on the residues around Y142, key in binding  $\alpha$ E-catenin. Yellow residues are key in binding  $\alpha$ E-catenin. Green residues are acidic aspartate & glutamate, which can affect charge distribution hence affect  $\alpha$ E-catenin binding. In poriferans only, the aspartic acid residue was substituted for a neutral glycine residue.

#### 4.4.2 Evolutionary conservation of $\beta$ -catenin interaction with E-cadherin in metazoans.

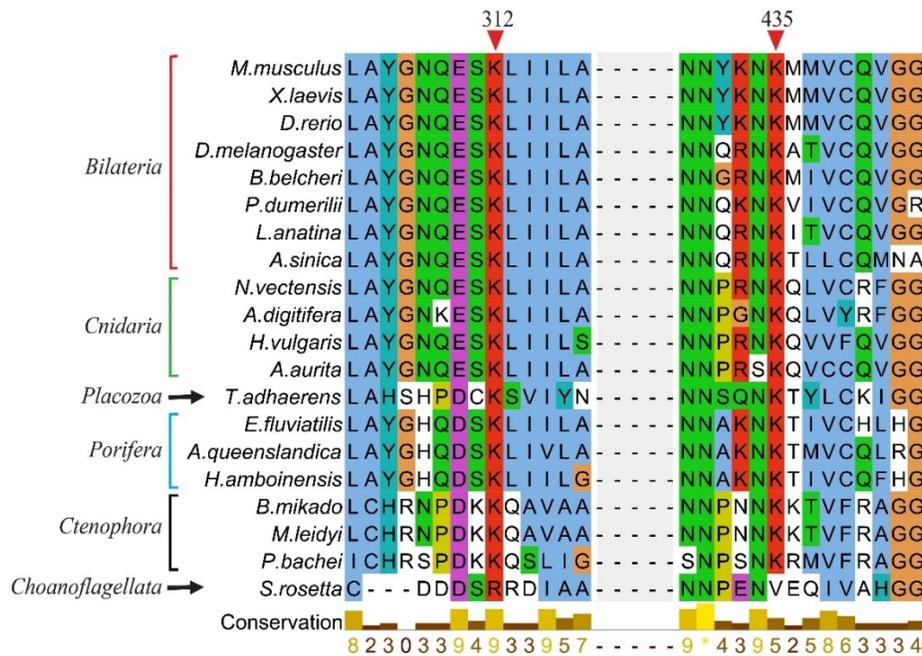
Cadherins are a family of cell surface proteins including N (neural)-, E (epithelial)-, VE (vascular-endothelial)-, P (placental), K (kidney)- and R (retinal)-cadherins (Hulpiau and van Roy, 2009). Mass spectrometry analysis detected E-cadherin in all metazoan  $\beta$ -catenin elutes (Table 4.2). The validation by immunoblot analysis was not possible due to the unavailability of a suitable antibody. Although there have been studies that have indicated the binding of  $\beta$ -catenin to VE-cadherin (Margariti *et al.*, 2010), K-cadherin (Shimoyama *et al.*, 1999), N-cadherin (Straub *et al.*, 2003), and P-cadherin (Ribeiro *et al.*, 2013), none of these were found to interact with any of the purified exogenous metazoan  $\beta$ -catenin. This could probably be due to the early developmental stage at which the analysis was done, i.e., they are likely not expressed at the gastrula stage.

		E-cadherin								
		Abundance ratio				P-value				
Uniprot ID	PSMs	Unique peptides	<i>X.lae</i>	<i>N.vec</i>	<i>E.flu</i>	<i>B.mik</i>	<i>X.lae</i>	<i>N.vec</i>	<i>E.flu</i>	<i>B.mik</i>
			$\beta$ -catenin							
P33152	262	12	58.90	79.19	12.36	42.48	1.00E-17	1.00E-17	1.98E-05	1.06E-06
A0A1L8GET9	194	10	46.95	54.23	11.48	41.68	1.00E-17	1.00E-17	2.52E-03	1.20E-06

**Table 4.2. Coimmunoprecipitation of endogenous E-cadherin.** Transphylectic expression of metazoan  $\beta$ -catenin followed by affinity purification and proteomic analysis found that all metazoan  $\beta$ -catenins can bind endogenous E-cadherin.

Next, an investigation was carried out to ascertain whether metazoan  $\beta$ -catenins have the critical amino acids to interact with endogenous *X. laevis* E-cadherin. All metazoans  $\beta$ -catenins displayed conservation of the two critical lysine amino acids (K312 & K435 in mouse  $\beta$ -catenin) that are critical in the interaction binding E-cadherin to  $\beta$ -catenin (Graham *et al.*, 2000; Huber and Weis, 2001) (**Figure 4.3**). Interestingly, in *S. rosetta*, the residue corresponding to K312 was substituted for an arginine that displays similar chemical properties to lysine. Interestingly (as we will see in the next chapter), this residue (K312) is also essential in binding TCF to  $\beta$ -catenin (Graham *et al.*, 2000; Von Kries *et al.*, 2000). This could indicate that the  $\beta$ -catenin to enable structural and transcriptional function began as early as at the emergence of choanoflagellates.

Other residues, Y331, K335, D390 & R582 in mouse  $\beta$ -catenin, have been suggested to be crucial in binding E-cadherin at the CBD intracellular domain (Belahbib *et al.*, 2018). These were conserved in all metazoans but ctenophores. However, precipitation of *Xenopus* E-cadherin by *B. mikado*  $\beta$ -catenin implies that these residues may not be evolutionarily vital in binding E-cadherin as previously suggested (Belahbib *et al.*, 2018).



**Figure 4.3. The two lysine residues that are critical in binding E-cadherin.** Residues K312 & K435 (in mouse  $\beta$ -catenin) are critical to the interaction with E-cadherin. These were conserved in all metazoans. Choanoflagellate, *S. rosetta* had the residue corresponding to K312 substituted for an arginine amino acid.

#### 4.4.3 Other adherens junctions associated proteins

Apart from E-cadherin and  $\alpha$ E-catenin, mass spectrometry analysis detected other proteins that are associated with the adherens junctions (**Table 4.3**). The association of JUP and CTNND1 with *X. laevis* and *N. vectensis*  $\beta$ -catenin was earlier highlighted and explained (see details in section 3.4.4). Another protein that I found was ARVCF. It has been shown to localise both in the nucleus and the adherens junction (Mariner *et al.*, 2000). Like  $\alpha$ E-catenin, ARVCF was significantly abundant in all metazoan  $\beta$ -catenin but *E. fluviatilis*  $\beta$ -catenin. Unlike other p120 subfamily proteins, ARVCF is present in non-vertebrates, including non-bilaterians (Zhao *et al.*, 2011). However, it is difficult to conclude whether this interaction is direct with  $\beta$ -catenin or indirectly through another protein.

Uniprot ID	Protein	PSMs	Unique peptides	Abundance ratio				P-value			
				<i>X. lae</i> β-catenin	<i>N. vec</i> β-catenin	<i>E. flu</i> β-catenin	<i>B. mik</i> β-catenin	<i>X. lae</i> β-catenin	<i>N. vec</i> β-catenin	<i>E. flu</i> β-catenin	<i>B. mik</i> β-catenin
A0A1L8HQG5	ARVCF	189	3	7.47	5.26	2.78	11.09	2.69E-05	1.32E-03	2.53E-01	2.49E-03
Q6INZ6		199	5	8.37	6.54	1.19	6.69	1.10E-07	6.56E-05	8.79E-01	2.00E-02
Q8AXM9	CTNND1	103	20	5.15	3.98	1.25	3.41	8.57E-05	2.99E-03	8.52E-01	1.64E-01
A0A1L8EKG6	JUP	144	1	3.07	5.24	2.87	4.44	3.15E-02	1.78E-03	2.47E-01	7.85E-02

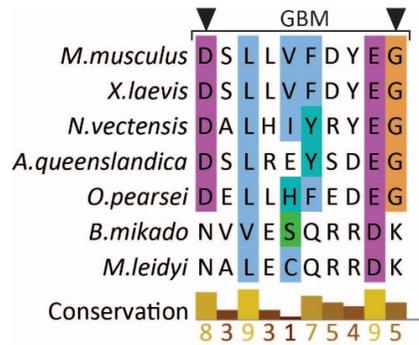
**Table 4.3. Other proteins that are associated with the adherens junctions.** JUP (plakoglobin) was significantly abundant in *X. laevis* and *N. vectensis* β-catenin. In contrast, the p120 subfamily proteins, CTNND1 and ARVCF, were found in all metazoan β-catenins but *E. fluviatilis*. *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*).

#### 4.5 Discussion

Much research focuses on the presence or absence of a gene or domain between different species to investigate the evolution of molecular pathways or protein functions. However, although a domain deletion or insertion can drastically affect a protein's function, a single nucleotide change can have just the same effect (Anderson *et al.*, 2016). Although CCC homologs are present in all metazoans, molecular interactions between E-cadherin, β-catenin, and αE-catenin have not been thoroughly investigated. The results from this study suggest that the capability of *Xenopus* E-cadherin to interact with metazoan β-catenin indicates that this invention was probably already present in the last common ancestor of metazoans (**Figure 4.5**). This was further supported by the observation that residues in β-catenin critical in this interaction were conserved throughout metazoans but not fully conserved in *S. rosetta* seems to support the hypothesis that cell-cell adhesion as mediated by the CCC is ancestral to unikonts (Miller *et al.*, 2013).

Furthermore, the conservation of E-cadherin-β-catenin interaction serves a protective role for both proteins. Interaction between these proteins starts in the endoplasmic reticulum, and the two proteins translocate together to the plasma membrane (Valenta *et al.*, 2012). During translocation, β-catenin inhibits E-cadherin degradation by shielding E-cadherin's PEST motif associated with rapid protein turnover (Rechsteiner and Rogers, 1996) and when recognized by ubiquitin ligase marks E-cadherin for degradation (Hinck *et al.*, 1994; Huber *et al.*, 2001). Mutually, E-cadherin prevents the binding of destruction complex proteins onto β-catenin (Huber and Weis, 2001). This protective relationship between E-cadherin and β-catenin might explain why I got no protein expression when I expressed *S. rosetta* β-catenin-like mRNA in the *Xenopus* embryo. Theoretically, the lack of conservation of one of the critical lysine residues means that *S. rosetta* β-catenin-like protein cannot bind endogenous *Xenopus* E-cadherin; hence it is subject to rapid degradation. The results suggest that E-cadherin & β-catenin most likely co-evolved to facilitate cell-cell adhesion at the adherens junction.

However, it is essential to note that this IP-MS data is from a transphylectic expression analysis. This is particularly crucial when interpreting the results of *B. mikado* β-catenin interaction. Although it precipitated endogenous *X. laevis* E-cadherin, it is likely that it cannot precipitate endogenous *B. mikado* E-cadherin. Ten amino acids in E-cadherin (DXXXXϕXXEG, ϕ = aromatic residue) forming the groove binding motif (GBM) in the catenin binding domain (CBD) are vital in the interaction with β-catenin (Nichols *et al.*, 2012; Murray and Zaidel-Bar, 2014; Belahbib *et al.*, 2018). In *B. mikado* (and *M. leidy*) E-cadherins, the D is replaced with an N while the G was replaced with a K, possibly resulting in poor interaction between endogenous E-cadherin and β-catenin (**Figure 4.4**). Combined with evidence in the next chapter (for details, see **section 5.4**) (absence of crucial TCF binding sites, lack Wnt activity by ctenophore β-catenin), I believe that ctenophore features are ancestral.



**Figure 4.4. The metazoan groove binding motif (GBM) in E-cadherins.** Ctenophores exhibit a lack of conservation in the GBM where the crucial aspartate (D) and glycine (G) residues (black arrows) are not conserved. This may impede the interaction with ctenophore  $\beta$ -catenin.

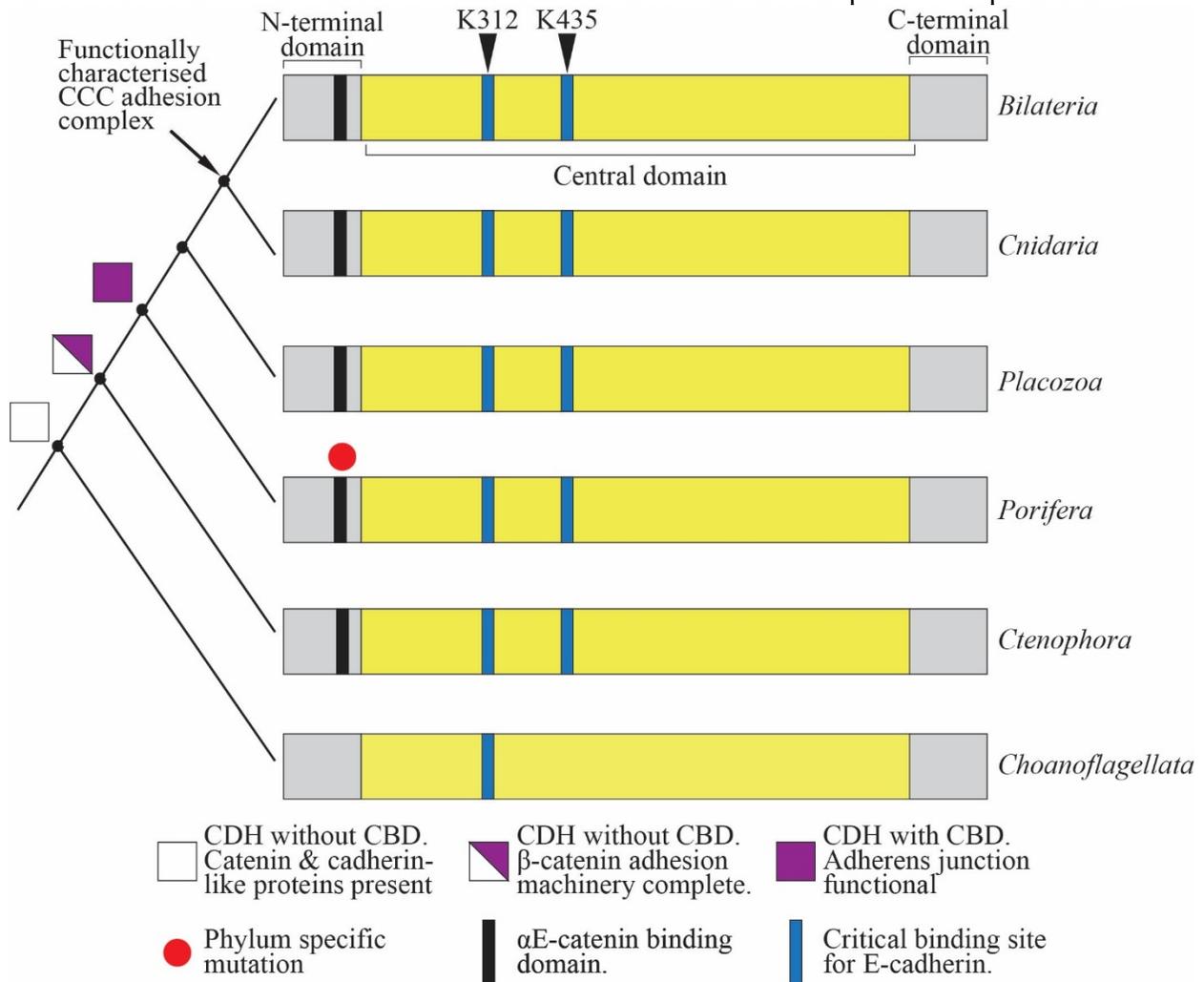
Regarding the conservation of  $\beta$ -catenin interaction with  $\alpha$ E-catenin, it was surprising to find that *E. fluviatilis*  $\beta$ -catenin was not interacting with *Xenopus*  $\alpha$ E-catenin. It was fascinating because recent research had confirmed by affinity purification of endogenous *E. muelleri*  $\beta$ -catenin interaction  $\alpha$ E-catenin (Schippers and Nichols, 2018). The possible cause was narrowed down to charge distribution near the critical tyrosine residue in the  $\alpha$ E-catenin binding site. In mouse  $\beta$ -catenin, the position of the Y142's hydroxyl group is in close vicinity to two acidic amino acids, Asp-144 & Glu-147. Substituting these residues or introducing a negative charge on Y142 could alter charge distribution, thereby preventing  $\alpha$ E-catenin binding (Piedra *et al.*, 2003) (**Figure 4.2**). Among poriferans, acidic residues were conserved at the sites corresponding to Glu-147. In contrast, conservation was not observed at the sites corresponding to Asp-144. The residue corresponding to Asp-144 in *E. fluviatilis*  $\beta$ -catenin was substituted for neutral amino acid glycine, possibly impeding or weakening  $\alpha$ E-catenin binding. As a similar substitution is observed in *E. muelleri*, it could imply that poriferan  $\beta$ -catenin and  $\alpha$ E-catenin co-evolved to either allow or strengthen interaction.

As no functional experiments have been carried out in poriferans to understand the function of  $\alpha$ E-catenin in cell-cell adhesion, it is impossible to conclude its role in cell-cell adhesion in the epithelium. However, it raises a question about the nature of the poriferan epithelia. Although there are still some unanswered questions, sponges are considered mainly to be epithelial organisms. As observed in bilaterians, cell-cell adhesion in the epithelium requires the complete CCC adhesion complex (Gumbiner, 2005; Nelson, 2008; Miller *et al.*, 2013). The results here would therefore support the idea that cnidarians but poriferans possess a true epithelium. In the case of *Ephydatia* and probably other poriferans, different adhesion receptors may be essential. This would support the claim that the proteoglycan Aggregation Factor (AF) is vital for cell adhesion in demosponge (Fernández-Busquets and Burger, 1997; Grice *et al.*, 2017). However, new evidence found that in *E. muelleri*, the adherens junction & focal adhesion proteins were present and functional, meaning that the Aggregation Factor might not be as critical in cell adhesion as was previously thought (Mitchell and Nichols, 2019). The role of adherens junction in the epithelia of *Porifera* would be challenging to understand without knockdown experiments. At this juncture, most of our conclusions have been drawn from expression analyses.

#### 4.5.1 Premetazoan origin of $\beta$ -catenin's structural function

From the sequence alignment and the general structure prediction, we can see that *S. rosetta*  $\beta$ -catenin bears some similarities to metazoan  $\beta$ -catenins. Previously, we saw conservation of some GSK3 $\beta$  phosphorylation sites. Regarding the  $\beta$ -catenin's adhesive role, one residue corresponding

to mouse  $\beta$ -catenin K312 was conserved with an arginine which bears similar chemical properties. On the other hand, in the  $\alpha$ -catenin binding site, *S. rosetta*  $\beta$ -catenin-like protein had a 45 amino acid insertion which was not the case in other metazoans (**Figure 4.2**). Therefore, this region may have been deleted & or reorganised during the emergence of metazoans, thereby creating a binding site for  $\alpha$ E-catenin. If this is the case, this would suggest that the interaction between E-cadherin and  $\beta$ -catenin was of more vital importance in the rise of multicellularity (**Figure 4.5**). The interaction with the actin cytoskeleton through  $\alpha$ E-catenin could have served to improve the strength of the cell-cell contact and give the cells more structure in metazoans. This hypothesis could be supported by the fact that  $\alpha$ E-catenin siRNA injection into early *N. vectensis* embryos results in not so severe gastrulation defects that only emerge following gastrulation initiation (Clarke *et al.*, 2019). On the other hand, knockdown of *N. vectensis* E-cadherin results in total inhibition of growth and gastrulation (Pukhlyakova *et al.*, 2019). Future work should therefore endeavour to understand the role of cadherin and  $\alpha$ E-catenin in ctenophores and poriferans.



**Figure 4.5. Evolution of  $\beta$ -catenin role structural role.** Pre-metazoan existence of cadherin like homologs (CDH) in choanoflagellates lacks cytoplasmic catenin binding domains (CBD). However, a single site critical for binding E-cadherin was already established in a  $\beta$ -catenin-like homolog in *S. rosetta*. By the emergence of ctenophores, domains critical for binding both  $\alpha$ E-catenin & E-cadherin had been formed in  $\beta$ -catenin, but cadherins still lacked a CBD. The evolution of CDH1 was completed around the emergence of poriferans, allowing interaction with  $\beta$ -catenin and functioning in the adherens junction.

#### 4.6 Chapter conclusion

We began our investigation by asking when  $\beta$ -catenin evolved its structural role leading up to the emergence of multicellular animals. A challenge for understanding the role of  $\beta$ -catenin in the CCC during the evolution of multicellularity is carrying out experiments in basal metazoans. Using a transphyletic approach, I have shown that all metazoan  $\beta$ -catenins can bind E-cadherin through evolutionarily conserved domains. Therefore, it implies that the  $\beta$ -catenin of the last common ancestor of metazoan was fully capable of binding E-cadherin. However, there remain some questions about the role of  $\alpha$ E-catenin in this complex. There is a possibility that compared to E-cadherin and  $\beta$ -catenin, its role in cell-cell adhesion in basal animals might vary from one phylum to another. Therefore, this might mean its current role as the adapter protein that links  $\beta$ -catenin to the actin cytoskeleton might have been a later innovation. Sequence homology alone cannot provide answers to such fundamental questions. Therefore, it is necessary to carry out functional and proteomic analysis in basal metazoans mainly, ctenophores and poriferans.

# CHAPTER 5

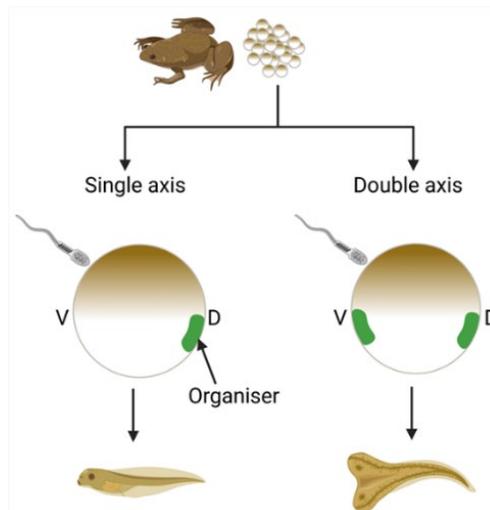
## Chapter 5 Conservation of organiser inducing function by metazoan $\beta$ -catenin

### 5.1 Summary

The evolutionary roots of  $\beta$ -catenin in organiser induction have been widely studied in bilaterians but remain poorly understood in basal metazoans. Using a transphylectic approach combined with detailed sequence analysis, I examined the organizer induction activity of different metazoan  $\beta$ -catenins. Results revealed that  $\beta$ -catenin's organiser-inducing function is conserved in *Cnidaria* and *Porifera* but not *Ctenophora*. Combined with the absence of transcription driving domains in ctenophores and lack of critical TCF binding residues, these results imply that the organiser inducing function of  $\beta$ -catenin was not present in the last common ancestor of metazoans but at around the time of poriferans.

### 5.2 Background

Dorsoventral axis specification in vertebrate embryos happens through the formation of the Spemann/ dorsal organizer and has been shown to involve various signalling mechanisms, including canonical Wnt signalling. The dorsal organizer was first described in a newt gastrula stage embryo as a group of cells capable of inducing a secondary embryonic axis when transplanted to the ventral side of the embryo (Spemann and Mangold, 1924) (**Figure 5.1**). Research using various vertebrate models later showed that nuclear accumulation of  $\beta$ -catenin in the dorsal region before mid-blastula transition was crucial in establishing dorsal cell fate (Kelly *et al.*, 2000; Bellipanni *et al.*, 2006). The importance of canonical Wnt signalling was first proven in *Xenopus* experiments where exogenous expression of Wnt mRNA or proteins generated a secondary dorsal organizer (Smith and Harland, 1991; Sokol *et al.*, 1991). Recently it was demonstrated that  $\beta$ -catenin signalling is an evolutionarily conserved determinant of the chordate dorsal organizer (Kozmikova and Kozmik, 2020). From all the above, it is clear that most of the research has utilized bilaterian models to understand the role of canonical Wnt signalling during embryonic development. Recently, the investigation into the role canonical Wnt signalling pathway in development in non-bilaterians has picked up, but there remain difficulties in using these models.



**Figure 5.1. Secondary axis induction in *X. laevis*.** In normal development, the presence of a single dorsal organizer results in the formation of a single axis. However, the expression of a canonical Wnt signalling agent on the ventral side induces a dorsal organizer resulting in secondary body axis formation. V, ventral and D, dorsal. Created using Biorender.com.

The oral-aboral axis is the main body axis in cnidarians, and considerable evidence from research on *Nematostella*, *Clytia* & *Hydra* suggests that canonical Wnt signalling controls the oral-aboral axis formation (Momose and Houliston, 2007; Lapébie *et al.*, 2014; Kraus *et al.*, 2016). Like in vertebrate embryos, it has been shown  $\beta$ -catenin accumulate in the nuclei of cells at the blastopore (Wikramanayake *et al.*, 2003). Similar to the Spemann & Mangold experiment, in *Nematostella*, transplantation of a fragment from the blastopore lip to the aboral ectoderm induces a secondary body axis (Kraus *et al.*, 2007). Furthermore, pharmaceutical targeting of canonical Wnt signalling using lithium chloride and 1-azakenpaullone hinders GSK3 $\beta$ -facilitated degradation of  $\beta$ -catenin, leading to gastrulation defects (Röttinger *et al.*, 2012). In *Hydra*, transplantation of the hypostome (comparable to the cnidarian blastopore) to a different site induces a secondary body axis (MacWilliams, 1983; Broun *et al.*, 2005). Furthermore, it was shown that treatment with alsterpaullone, a GSK3 $\beta$  inhibitor, resulted in the body column gaining markers of the head organizer (Broun *et al.*, 2005). These results obtained from cnidarian models suggest that the role of canonical Wnt signalling in axis specification is conserved.

In ctenophores & poriferans,  $\beta$ -catenin signalling is yet to be functionally assessed, but elements of  $\beta$ -catenin signalling have been found localised at particular positions along the developing embryo's axis (Adamska *et al.*, 2007; Pang *et al.*, 2010; Jager *et al.*, 2013; Leininger *et al.*, 2014). This may imply that canonical Wnt signalling may play an active role in body axis formation.

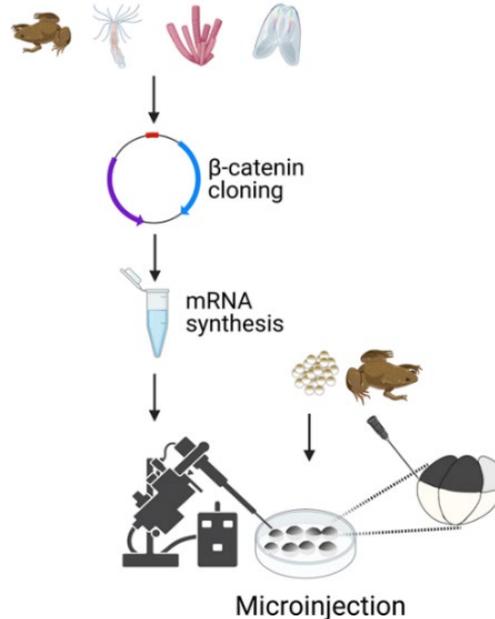
This chapter will focus on the results obtained from domain and structural analysis of the TCF binding domain in metazoan  $\beta$ -catenin and how they complement & explain the results of the secondary axis induction experiment. Alongside this, results of TCF interaction with  $\beta$ -catenin obtained from IP-MS analysis will be presented.

### 5.3 Methods

As detailed in sections 2.3 and 3.3.

#### 5.3.1 Secondary axis assay

Flag-tagged  $\beta$ -catenin mRNA (100, 250, 500 & 1000 pg) was coinjected with tracer LacZ mRNA (encoding  $\beta$ -galactosidase) (25 pg) into one blastomere in the ventral equatorial zone of 4-cell stage *X. laevis* embryos (**Figure 5.2**). Injected embryos were then incubated at 20°C in 5% Ficoll in 1x Steinberg solution until stage 33-37. At this stage, they were prepared for X-gal staining as described below in section 5.3.2. Scoring of axis duplication success was classed as complete if the second axis had eyes and cement gland, incomplete in the head features of the secondary axis were not fully developed and normal if there was only one axis. For each experiment, uninjected and LacZ mRNA injected embryos were used as negative controls.



**Figure 5.2. Method to confirm conservation of metazoan  $\beta$ -catenin's role in organizer induction.**  $\beta$ -catenin mRNA was injected in one-ventral blastomere at the 4-cell stage. Injected embryos were cultured until 33-37.

### 5.3.2 X-gal fixing and staining

Embryos were collected then fixed for 1 hour on ice in gal fix solution (37% formaldehyde, 25% glutaraldehyde, 10% TritonX-100, 10%, 0.7% PBS). Following fixation, embryos were washed in 0.7x PBS then prepared for gal staining in fresh gal solution (0.1 M  $K_3Fe(CN)_6$ , 0.1M  $K_4Fe(CN)_6$ , 10% X-gal, 1 M  $MgCl_2$ , 0.7x PBS) at 30°C for 3 hours. Following staining, embryos were rewashed in 0.7x PBS and refixed in MEMFA (10% MEM salts, 10% formaldehyde, MilliQ water) for 2 hours at room temperature. This was followed by a wash in 0.7x TBS, 50% methanol, and three 5 minute washes in 100% methanol. Finally, the embryos could either be imaged or stored in methanol at -20°C.

### 5.3.3 Luciferase reporter assay

A TOPflash luciferase assay analysis was carried out to measure and compare the activity of metazoan  $\beta$ -catenin. Analysis was carried out in triplicates. First, 100 pg  $\beta$ -catenin mRNA together was coinjected with either 50 pg M50 Super 8X TOPflash or 50 pg M51 Super 8X FOPflash (TOPflash mutant) (Veeman *et al.*, 2003) at 1-cell stage in the animal pole, and the embryos were incubated at 20°C until gastrula stage (11-12). Then, embryos were washed in a 0.1x Steinberg solution and prepared for analysis following the protocol detailed in Yasuoka and Taira (2019). The luciferase assay system (E1500, Promega) was used in homogenisation following the manufacturer's guidelines. Finally, lysates were transferred to individual wells in a 96-well plate, and luciferase activity was measured using a Centro LB960 microplate reader (Berthold Technologies). Statistical analysis was carried out using GraphPad Prism 9 software (San Diego, California USA, [www.graphpad.com](http://www.graphpad.com)), and two-way ANOVA analysed statistical significance.

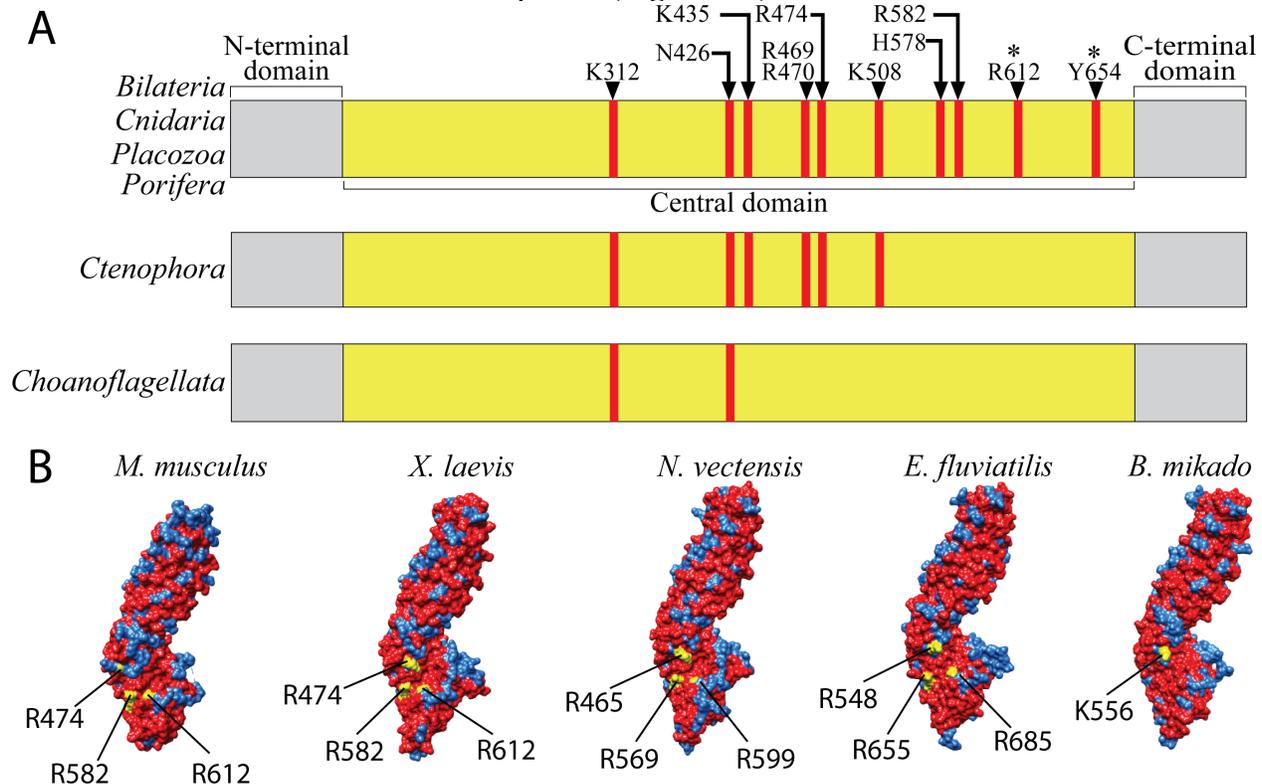
## 5.4 Results

### 5.4.1 TCF- $\beta$ -catenin binding sites conserved in all metazoans but *Ctenophora*

Previous studies have shown that  $\beta$ -catenin binds to TCF between armadillo repeat 3 & 10 (Graham *et al.*, 2000; Von Kries *et al.*, 2000) (Figure 1.4). As stated in section 2.4, the degree of conservation within the central region was high. Therefore, particular attention was initially placed

on three arginines (R474, R582 & R612 in mouse  $\beta$ -catenin ) that have been indicated to be critical in this interaction (Graham *et al.*, 2000) (Figure 5.3). In metazoans, the residues corresponding to R474 were highly conserved, although, in ctenophores and placozoan, they had been substituted with a lysine residue. R582 & R612 showed high conservation in all metazoans but ctenophores (Figure 5.3). On the other hand, *S. rosetta*  $\beta$ -catenin-like protein showed no conservation in these critical regions. Comparison of structural models revealed that the three critical residues were also localised in the same region as observed in the target model (mouse  $\beta$ -catenin).

Other high positive electrostatic potential sites have been suggested to be essential in  $\beta$ -catenin-TCF interaction, and these are K435, K508, K312, R469, H470 & N426 (Graham *et al.*, 2000; Von Kries *et al.*, 2000). Sequence alignment and comparison showed that these sites were highly conserved in all metazoans. A site-directed mutagenesis study also found that H578 & Y654 in *Xenopus*  $\beta$ -catenin are also vital in binding TCF (Fasolini *et al.*, 2003). These sites were found to be conserved in all metazoans but ctenophores (Figure 5.3).



**Figure 5.3.  $\beta$ -catenin residues critical in binding TCF.** (A) Amino acid residues that essential in binding TCF. Apart from ctenophores, the rest of the metazoans displayed conservation of residues key in TCF interaction. Ctenophores had only 7 of the 11 residues conserved. Asterisk sign indicates lack of conservation in *T. adhaerens*. Similarly, ctenophore, *P. bachei* had a conserved arginine residue at the site corresponding to R612. (B) Surface depiction of  $\beta$ -catenin homology structural models were compared with that mouse  $\beta$ -catenin. This confirmed that the three crucial arginine residues (yellow spots) had a localization similar to that observed in mouse  $\beta$ -catenin. *B. mikado*  $\beta$ -catenin's K556 corresponded in placement to mouse  $\beta$ -catenin R474. Structure visualized using UCSF Chimera software.

#### 5.4.2 Coprecipitation of canonical Wnt signalling elements with $\beta$ -catenin

The interaction of  $\beta$ -catenin and TCF has been previously proven (Hanson *et al.*, 2012; Satow *et al.*, 2012). Vertebrates have a family of four TCF members compared to the single ortholog found in invertebrates (Arce *et al.*, 2006). To understand how the differences in the  $\beta$ -catenin TCF

binding domain may affect the interaction, immunoprecipitation with mass spectrometric analysis was carried out as described previously in Chapter 3. In the current study, the TCF7L1 ortholog was found to interact with only the *X. laevis*  $\beta$ -catenin. With other metazoan  $\beta$ -catenins, although mass spectrometry results indicated the presence of some TCF7L1 peptides, their abundance was not statistically significant (**Table 5.1**). Since cnidarian and poriferan  $\beta$ -catenins exhibited conserved residues to bind TCF, I found their inability to precipitate TCF surprising. Although it remains to be clarified why TCF was not precipitated at significant levels, several recent works on proteomics of  $\beta$ -catenin-associated proteins were also unsuccessful in detecting  $\beta$ -catenin and TCF interaction (Hwang *et al.*, 2017; Ehyai *et al.*, 2018; Semaan *et al.*, 2019). One possible reason is that TCF- $\beta$ -catenin interaction is highly regulated; hence little  $\beta$ -catenin is required to drive transcription of target genes. The other reason could be a methodology and technology limitation. As little TCF interacts with  $\beta$ -catenin, the method needs thorough optimisation to detect the few available TCF peptides.

Uniprot ID	Protein	PSMs	Unique peptides	Abundance ratio				P-value			
				<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin	<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin
B7ZQ53	TCF7L1	10	3	7.19	1.25	1.17	1.38	8.19E-07	7.57E-01	N/A	N/A

**Table 5.1. Coimmunoprecipitation of endogenous TCF by metazoan  $\beta$ -catenin.** From three replicates, immunoaffinity purification of metazoan flag-tagged  $\beta$ -catenin followed by mass spectrometry analysis identified three unique peptides of TCF7L1. Its abundance in the *X. laevis* immunoprecipitation elute was statistically significant compared to that observed in other metazoans. AR, abundance ratio.

Another canonical Wnt signalling element detected was APC, a critical component of the destruction complex. As with E-cadherin and TCF, APC also binds to K312 & K435 of  $\beta$ -catenin. These were highly conserved in metazoans, as previously indicated (see section 4.4.2). Other interaction sites with  $\beta$ -catenin are F253, F293, A295, and I296 (Graham *et al.*, 2000). Except for F253, which was substituted for a tyrosine, in ctenophores, conservation was observed in all metazoans and *S. rosetta* for all residues. Both large and small homologs of APC were detected at significant levels in  $\beta$ -catenin elutes of *X. laevis* and *N. vectensis*, but not of *E. fluviatilis* and *B. mikado* (**Table 5.2**). However, the inability of *E. fluviatilis* and *B. mikado*  $\beta$ -catenin to bind APC should not affect  $\beta$ -catenin's degradation as it has been shown that GSK3 $\beta$  can target  $\beta$ -catenin for phosphorylation without requiring other protein members of the destruction complex (Yost *et al.*, 1996). Unfortunately, no other canonical Wnt pathway components, e.g., Axin, GSK3 $\beta$  or CK1 $\alpha$ , were identified.

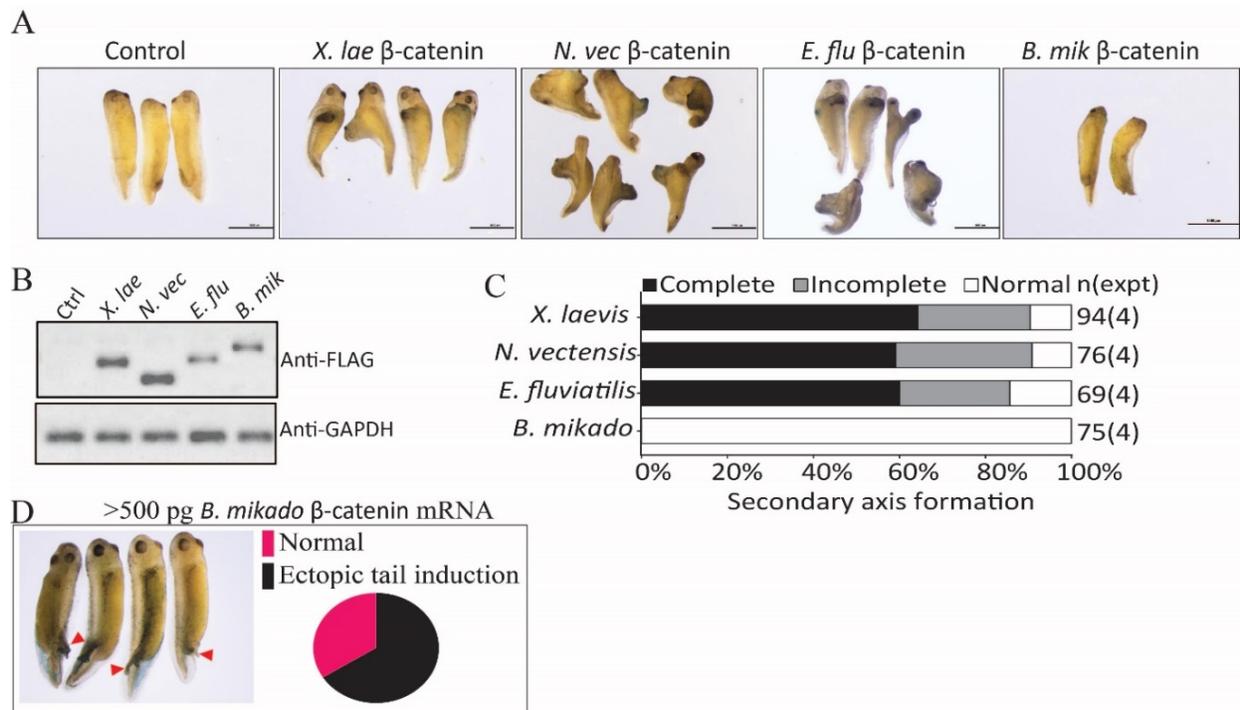
Uniprot ID	Protein	PSMs	Unique peptides	Abundance ratio				P-value			
				<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin	<i>X.lae</i> $\beta$ -catenin	<i>N. vec</i> $\beta$ -catenin	<i>E. flu</i> $\beta$ -catenin	<i>B. mik</i> $\beta$ -catenin
P70039	APC	75	1	13.52	10.65	1.02	1.26	6.34E-09	8.66E-06	9.73E-01	9.19E-01
A0A1L8I1P0		69	5	14.28	10.57	0.98	2.57	3.05E-08	5.82E-06	9.39E-01	3.84E-01

**Table 5.2. Coimmunoprecipitation of endogenous APC by metazoan  $\beta$ -catenin.** From three replicates, immunoaffinity purification of flag-tagged  $\beta$ -catenin followed by mass spectrometry analysis identified APC. Its abundance in the *X. laevis* & *N. vectensis* immunoprecipitation elute was statistically significant compared to that observed in other metazoans. AR, abundance ratio.

In order to examine whether the basal metazoan  $\beta$ -catenin can activate the signalling pathway in the *Xenopus* context, I carried out a functional analysis.

### 5.4.3 Secondary axis induction assay

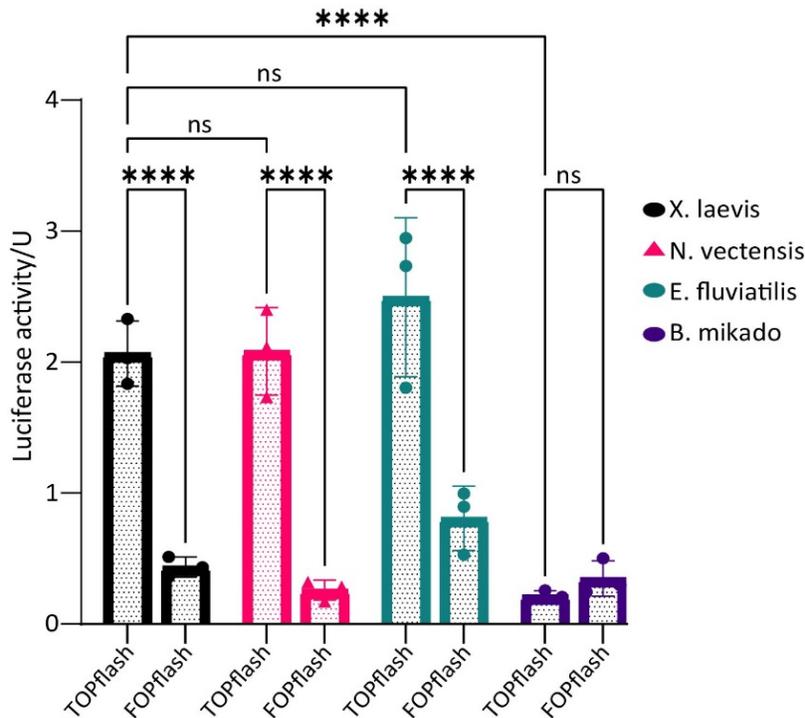
Unfortunately, my IP-MS analysis was not successful in examining TCF interaction with  $\beta$ -catenin proteins of *N. vectensis*, *E. fluviatilis*, and *B. mikado*, whereas the lack of conservation of crucial residues that bind TCF suggests that only ctenophore  $\beta$ -catenin may not have the capacity to bind TCF (for detail, see section 5.4.1). To test whether basal metazoans'  $\beta$ -catenins can functionally link to TCF, metazoan  $\beta$ -catenin was injected at the 4-cell stage into the ventral marginal zone of *X. laevis* embryos, and protein expression was verified by western blotting. The secondary axis induced bodies were classified into three categories, i.e., complete (cement gland present), incomplete (cement gland absent), and normal (single body axis). The results showed that *X. laevis*  $\beta$ -catenin, *N. vectensis* & *E. fluviatilis*  $\beta$ -catenins could induce a secondary body axis up from 100 pg. The majority of these were in the complete secondary body axes (**Figure 5.4**). In contrast, *B. mikado*  $\beta$ -catenin did not induce a secondary axis when expressed in the ventral marginal zone. These results suggest that  $\beta$ -catenin's organiser inductive capacity during embryo development is conserved in all metazoans. However, at high concentrations (>500 pg), induction of ectopic tail-like structures was observed in some embryos (**Figure 5.4D**). Normal axis development was observed in embryos injected with LacZ mRNA. The inability of *B. mikado*  $\beta$ -catenin to induce a secondary axis body could imply that *Ctenophora*  $\beta$ -catenin is probably not functional in the canonical Wnt signalling pathway.



**Figure 5.4. Secondary body axis induction assay.** Embryos were injected in the ventral marginal zone of one blastomere at the 4-cell stage with *X. laevis*, *N. vectensis*, *E. fluviatilis* & *B. mikado*  $\beta$ -catenin as described in Methods (section 5.3.1). (A) Only *X. laevis*, *N. vectensis* & *E. fluviatilis*  $\beta$ -catenin induced axis duplication while a single axis was observed in embryos injected with *B. mikado*  $\beta$ -catenin. (B) Western blotting confirmed the presence of flag-tagged  $\beta$ -catenin from each of the metazoans at the expected band size. (C) The results from four different concentrations (100, 250, 500 & 1000 pg) combined to represent the efficiency of secondary axis induction. The majority of embryos with duplicated axes presented complete duplication with two pairs of eyes. (D) At high concentrations (500 & 1000 pg), the injection of *B. mikado*  $\beta$ -catenin resulted in an ectopic tail-like structure (red arrows) in some embryos (60%,  $n = 30$ ). Abbreviations: *X. lae* (*X. laevis*), *N. vec* (*N. vectensis*), *E. flu* (*E. fluviatilis*), and *B. mik* (*B. mikado*). Ctrl (control (uninjected))

#### 5.4.4 Metazoan canonical Wnt signalling activity induction

To further confirm that the basal metazoan  $\beta$ -catenins actually drive gene expression via TCF, I carried out a TOPflash luciferase assay using *Xenopus* embryos. The TOPflash plasmid utilizes luciferase expression under the control of TCF interaction sites (Molenaar *et al.*, 1996; Griffin *et al.*, 2018). FOPflash is a negative control plasmid that contains mutated TCF binding sites. Consistent with the secondary axis induction experiments, there was a significant increase in canonical Wnt signalling activity following expression of *X. laevis*, *N. vectensis* & *E. fluviatilis*  $\beta$ -catenin. On the other hand, the expression of *B. mikado*  $\beta$ -catenin did not increase the luciferase activity (Figure 5.5). These results complemented the secondary axis induction assay results and confirmed that, indeed, the signalling activity of  $\beta$ -catenin is not conserved in all metazoans.



**Figure 5.5. Canonical Wnt signalling activity for different metazoan  $\beta$ -catenin.** Expression of *X. laevis*, *N. vectensis* & *E. fluviatilis* flag-tagged  $\beta$ -catenin mRNA (100 pg) in *Xenopus* embryos resulted in a significant increase in Wnt activity which was not observed in the case *B. mikado* expression. Each bar is a representation of the mean of three biological replicates. ns, not significant ( $p$ -value>0.5); \*\*\*\* $p$ -value<0.0001 (Two way ANOVA).

#### 5.5 Discussion

Secondary axis induction assays in *Xenopus* embryos have been used in several studies to ascertain different protein functions and how they might relate or affect canonical Wnt/  $\beta$ -catenin signalling. This chapter presented for the first the origins and evolutionary roots of  $\beta$ -catenin with regards to its role as an organiser inducer during early embryo development. Canonical Wnt signalling occurs on the future dorsal side of the developing *Xenopus* embryo with overexpression of Wnt signalling components on the ventral side resulting in induction of ectopic dorsal development. However, for dorsal development to proceed, the interaction of  $\beta$ -catenin with TCF is paramount.

In the current study, analysis of the known sites in  $\beta$ -catenin where TCF binds were largely conserved across several metazoan phyla apart from ctenophores. Three arginine residues in  $\beta$ -catenin form critical interactions with TCF. It was shown that R474 & R612 form two salt bridges

with the corresponding sites on *Xenopus* TCF3 at aspartate-10 & glutamate-11, respectively, while R582 also interacts with *Xenopus* TCF3 at glutamate-11 (Graham *et al.*, 2000). In ctenophores, the site corresponding to R474, the charge had been conserved by mutation arginine (R) to a lysine (K). Furthermore, the structural analysis revealed that this residue was also conserved in its localization, indicating that endogenous *Xenopus* TCF would interact with it. However, the sites corresponding to R582 and R612 in *B. mikado*  $\beta$ -catenin had been substituted with a glutamate and glycine residue, resulting in a change in charge and probably affecting TCF interaction. The importance of each of these residues in binding TCF is still unknown as site-directed mutagenesis studies are limited and produce varying results. However, a previous study using *Xenopus*  $\beta$ -catenin showed that mutations R612A & Y654A (both of which are not conserved in *B. mikado*) do not significantly affect TCF interaction (Fasolini *et al.*, 2003). These may therefore not be evolutionarily critical in binding TCF. However, TCF interaction with  $\beta$ -catenin was found to decrease (about 30 fold) in  $\beta$ -catenin mutations R582A and H578A (Fasolini *et al.*, 2003). The lack of conservations at these sites in *B. mikado*  $\beta$ -catenin could contribute to weak interaction with endogenous *Xenopus*  $\beta$ -catenin and subsequent lack of dorsal organiser induction capability (**Figure 5.4**). Therefore, this could partly explain the results obtained from the luciferase assay of canonical Wnt signalling activity (**Figure 5.5**).

We employed affinity purification with mass spectrometry to determine if the interaction of metazoan  $\beta$ -catenin with TCF has been conserved since the emergence of metazoans. TCF7L1 (TCF3) peptides were detected to be significantly abundant in the elute *X. laevis*  $\beta$ -catenin. It has been shown that the four TCF proteins in vertebrates function differently, with some acting as repressors and other activators of gene transcription. LEF1 has been suggested to be a transcriptional activator as its overexpression on the ventral side induces axis duplication in *Xenopus* embryos (Behrens *et al.*, 1996). *Xenopus* TCF1 was found to exhibit site-specific repression and activation of Wnt target genes ventrally and dorsally, respectively (Standley *et al.*, 2006). The same study found that TCF4 is an activator of a nodal related gene, NR3 and chordin, in the organizer cells at the gastrula stage (Standley *et al.*, 2006). The most abundant of TCF proteins is TCF7L1 (Hrckulak *et al.*, 2016; Liang and Liu, 2018). Indeed, proteomic analysis of  $\beta$ -catenin interactome in this study managed to identify only TCF7L1. Relying on the conserved sites where TCF binds in *N. vectensis* & *E. fluviatilis*  $\beta$ -catenin, it could be hypothesised that endogenous TCF7L1 should have the capacity to interact with these  $\beta$ -catenins. This is supported by *in situ* hybridization evidence that  $\beta$ -catenin and TCF genes are coexpressed in the developing oocyte of *Sycon ciliatum* (Leininger *et al.*, 2014). Similarly, in cnidarians, it was shown that TCF and  $\beta$ -catenin colocalize in the head and the budding zones (Hobmayer *et al.*, 2000). However, in the ctenophore *M. leidyi*, TCF and  $\beta$ -catenin expression are not so apparent as TCF is diffusely expressed in the ectoderm and later in the blastopore while  $\beta$ -catenin is primarily expressed in the blastopore (Pang *et al.*, 2010). The failure to detect significant levels of TCF7L1 peptides in *N. vectensis* and *E. fluviatilis*  $\beta$ -catenins could be due to a limitation of the mass spectrometry analysis. It could also be the reason why none of the other TCF orthologs were detected.

### 5.5.1 Dorsal organizer inducing capacity

Presently few studies have investigated whether protein homology as reflected in sequence similarity is conserved in function. The current study utilized a classical approach, secondary axis induction assay, to understand whether  $\beta$ -catenin's capacity in organiser induction had been conserved in all metazoans. Apart from *B. mikado*  $\beta$ -catenin, which induced ectopic tail-like features, overexpression of metazoan  $\beta$ -catenin induced a secondary body axis. The reason for the

failure of *B. mikado*  $\beta$ -catenin to induce a secondary body axis could be down to several factors. First, as observed in the sequence alignments, *B. mikado*  $\beta$ -catenin, compared to other metazoans, lacks critical binding sites for TCF interaction. TCF binding sites are localised in the central domain, which by itself is sufficient to induce secondary body axis in *Xenopus* embryos (Funayama *et al.*, 1995). However, is the lack of conservation of TCF binds sites enough to elicit failure in organiser induction by *B. mikado*  $\beta$ -catenin?

To answer the question above, we would have to start by analysing domain differences in metazoan  $\beta$ -catenin, particularly in the C-terminal (**Figure 2.1**). Previous studies have shown that  $\beta$ -catenin C-termini can function as transcriptional activators when fused to TCF (Van de Wetering *et al.*, 1997; Hsu *et al.*, 1998). Furthermore, it was shown that  $\beta$ -catenin C-terminus-LEF1 fusion is sufficient for secondary body axis induction in *Xenopus* (Vleminckx *et al.*, 1999). So, which region of the C-terminal could be vital in  $\beta$ -catenin's organiser induction capability?

First, we observed in section 2.4.1 that motif A in the C-terminal is present in only bilaterians & cnidarians. This motif is localised within the Helix-C region, a key player in Wnt-stimulated transcription activity (Xing *et al.*, 2008). Mutations targeting Helix-C have been shown to affect the signalling function of  $\beta$ -catenin without interfering with its cell adhesion role (Orsulic and Peifer, 1996). Furthermore, the importance of this region was highlighted in a review which indicated that the interaction of several  $\beta$ -catenin transcriptional coactivators requires an intact Helix-C (Mosimann *et al.*, 2009). However, motif A is probably not crucial in organiser induction as it is absent in *E. fluviatilis*  $\beta$ -catenin, which managed to induce a secondary body axis.

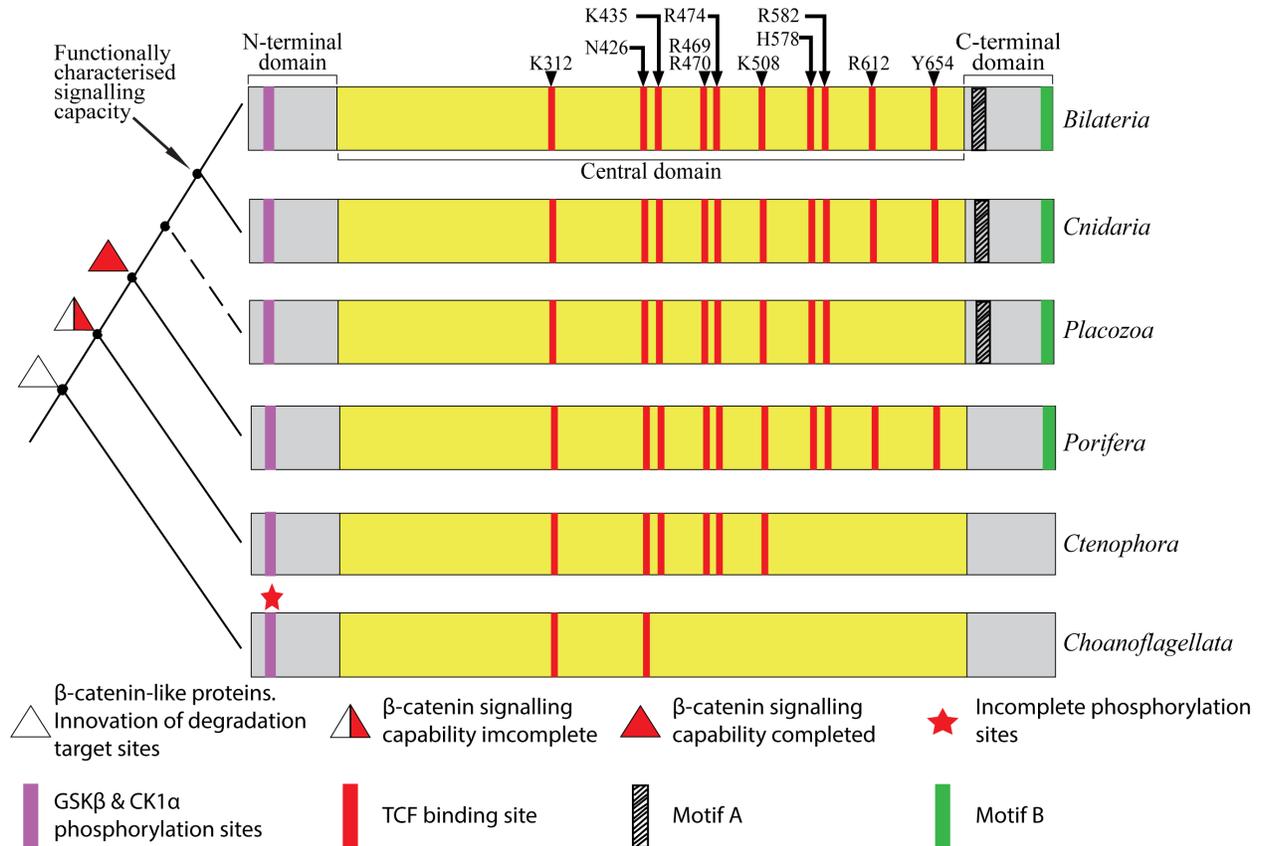
Therefore, this leaves us with motif B, which forms a PDZ binding domain (DTDLD) present in all metazoans but ctenophores. A recent study showed that mutation of this domain resulted in a decrease in TCF dependent transcriptional activity (DuChez *et al.*, 2019). Furthermore, some studies have shown this domain binds some proteins that may affect cell adhesion or signalling, e.g. LIN-7 (Perego *et al.*, 2000), EBP50 (Shibata *et al.*, 2003), SCRIB (Sun *et al.*, 2009). The organiser induction capability could depend on the interaction at  $\beta$ -catenin's PDZ domain with another protein. Therefore, *B. mikado*  $\beta$ -catenin's inability to induce a secondary body axis could be because it cannot fully activate canonical Wnt signalling, possibly due to insufficient interaction at the PDZ domain with endogenous proteins. This is supported by low canonical Wnt signalling activity in embryos overexpressed with *B. mikado*  $\beta$ -catenin, as determined by the TOPflash reporter assay. Integrating these results can explain why the inactivation of GSK3 $\beta$  in ctenophore, *M. leidy* by treating it with Lithium Chloride & alsterpaullone produced no phenotype (Pang *et al.*, 2010). Inhibition of GSK3 $\beta$  activates canonical Wnt signalling, and in *N. vectensis*, it has been shown to result in exogastrulation (Röttinger *et al.*, 2012). The failure to produce any visible phenotype following GSK3 $\beta$  inactivation in *M. leidy* is because although  $\beta$ -catenin accumulated, it had no downstream effect as it cannot interact with TCF due to lack of crucial residues and it lacks crucial transactivation motifs in the C-terminal.

These results combined create a story of when  $\beta$ -catenin's role as a transcriptional regulator most likely evolved. Secondary axis induction capability by *E. fluviatilis*  $\beta$ -catenin indicates that canonical Wnt signalling activity was established at *Porifera* emergence. This was maintained in bilaterians, cnidarians (and probably in placozoans). However, the missing TCF binding sites along with the lack of motifs A & B can lead to two hypotheses, i.e., 1) Ctenophores are the most basal of metazoans; hence the signalling capacity of  $\beta$ -catenin was still "under construction", 2) Poriferans are the most basal and  $\beta$ -catenin's signalling capacity was already established in

poriferans and was only lost in ctenophores due to truncation of sites critical in signalling. The former seems more parsimonious from these results as it is parallel with the functionality results obtained from this study (**Figure 5.6**). Also, combined with the knowledge that ctenophore E-cadherin has a divergent cytoplasmic domain that may impede  $\beta$ -catenin interaction garners more support that ctenophores are the earliest emerging metazoans (Belahbib *et al.*, 2018). Further support for this conclusion comes from a recent study that performed systematic and standardized analysis on 136 past phylogenomic studies and concluded that most of them placed ctenophores as the most basal animals (Li *et al.*, 2021). Therefore, it can be concluded that the  $\beta$ -catenin in the last common metazoan ancestor could not activate canonical Wnt signalling activity.

### **5.5.2 Formation of ectopic tail-like projections upon overexpression of ctenophore $\beta$ -catenin**

Although it could not induce a secondary axis, the most surprising result of this study was forming ectopic tail-like projections following overexpression of ctenophore  $\beta$ -catenin at the 4-cell stage. The same phenotype was recently observed following *CLK2* & *EFGF* mRNA expression in *Xenopus* (Virginia *et al.*, 2019). Multiple studies suggest that ectopic tail-like projections result from an effect on the FGF pathway (Chen *et al.*, 1999; Böttcher *et al.*, 2004). One intriguing possibility is that  $\beta$ -catenin's role as a dorsal organizer emerged after ctenophore. However, in ctenophore, although  $\beta$ -catenin was present, it might have primarily enhanced (either directly or indirectly) the FGF pathway in anterior-posterior axis formation. To investigate this, direct targets genes of ctenophore  $\beta$ -catenin need to be identified, and the role of endogenous  $\beta$ -catenin in ctenophore development also needs to be investigated.



**Figure 5.6. Evolution of canonical Wnt/  $\beta$ -catenin's signalling capacity.**  $\beta$ -catenin-like protein with incomplete phosphorylation/ degradation and TCF binding sites were present in the sister group of metazoans, choanoflagellates. More TCF binding sites were innovated around the emergence of ctenophores, but these were possibly insufficient to activate  $\beta$ -catenin signalling.  $\beta$ -catenin signalling, as seen in bilaterians, was initiated around the emergence of poriferans and was marked with new critical domain formation and TCF interacting sites. Later, motif A was created in other metazoans, which served to support  $\beta$ -catenin signalling.

## 5.6 Chapter conclusion

The functional capacity of both poriferan and ctenophore  $\beta$ -catenin has not been investigated until now. This study has shown that the signalling role of  $\beta$ -catenin is conserved in cnidarians and poriferans but not in ctenophores. An interesting phenotype induced by ctenophore  $\beta$ -catenin may indicate that it works through a different pathway (possibly the earliest function of  $\beta$ -catenin). Confirmation of this would, however, require analysis in ctenophores. The IP-MS results obtained do leave for improvement, resulting in identifying individual transcription pieces of machinery each  $\beta$ -catenin functions through. With this in hand, we shall begin to get a clearer view of how signalling activity evolved.

# CHAPTER 6

“Science works on the frontier between knowledge and ignorance. We're not afraid to admit what we don't know—there's no shame in that. The only shame is to pretend that we have all the answers”.

-Neil deGrasse Tyson

---

## Chapter 6 General discussion

There is no doubt that  $\beta$ -catenin was involved in the rise of multicellularity. However, little is known about its evolutionary roots and its role in basal metazoans, particularly poriferans and ctenophores. The current study investigated this by employing structural, proteomic, functional approaches combined with detailed sequence analysis of metazoan  $\beta$ -catenin.

Here we found that the structure of  $\beta$ -catenin is conserved in all metazoans indicates this architecture was already present in the last common metazoan ancestor. Interestingly, part of the metazoan  $\beta$ -catenin structure was similar to that unicellular *S. rosetta*  $\beta$ -catenin-like protein, meaning that the basic architecture of  $\beta$ -catenin has its origins in unicellular organisms.

There is still a lack of understanding of the role of  $\beta$ -catenin in basal animals particularly, poriferans and ctenophores. A recent study concluded that it was involved in cell adhesion and signalling in sponge *E. muelleri* (Schippers and Nichols, 2018). Our study supports this view as *E. fluviatilis*  $\beta$ -catenin exhibited canonical Wnt signalling activity and induced a secondary body axis when it was expressed in *Xenopus* embryos. Sequence analysis also confirmed the lysine residues in  $\beta$ -catenin that bind E-cadherin across all metazoans, and this was complemented with detection E-cadherin in all metazoan  $\beta$ -catenin elutes. The failure to detect  $\alpha$ E-catenin *E. fluviatilis*  $\beta$ -catenin could be due to a co-evolution of  $\beta$ -catenin and  $\alpha$ E-catenin in poriferans (**Figure 4.2**). All in all, our results suggest that poriferan  $\beta$ -catenin is functionally similar to bilaterian  $\beta$ -catenin.

Meanwhile, recent research indicated that in ctenophore *M. leidy*,  $\beta$ -catenin was involved in cell fate specification (signalling) but not in cell adhesion (Salinas-Saavedra *et al.*, 2019). However, this conclusion might be premature to an extent, particularly regarding the cell specification conclusion. It was found that ctenophore E-cadherin has a divergent cytoplasmic domain at the groove binding motif (GBM), which may impede the interaction with  $\beta$ -catenin (**Figure 4.4**) (Belahbib *et al.*, 2018). Indeed, *M. leidy*  $\beta$ -catenin was not visualized at cell-cell contacts in other metazoans but in the nucleus (Salinas-Saavedra *et al.*, 2019). This can confirm that it is not involved in cell adhesion. However, visualisation of  $\beta$ -catenin in the nucleus does not necessarily imply that it is functional as this comes down to several factors, as summarised in **Figure 1.6**. Indeed, earlier work from the same group has previously observed that the inactivation of GSK3 $\beta$  in developing *M. leidy* embryos produces no obvious phenotype (Pang *et al.*, 2010). Furthermore, during *M. leidy* development, all four *Wnt* genes, *Tcf* and  $\beta$ -catenin are detected after gastrulation (Pang *et al.*, 2010). In conjunction with our results, i.e., domain differences, weak canonical Wnt signalling activity & lack of dorsal organizer induction capability, it can be inferred that in ctenophores, Wnt/  $\beta$ -catenin signalling is not functionally conserved as it is in other metazoans. Nevertheless, ctenophore  $\beta$ -catenin does seem to target through an unknown pathway (possibly FGF (Chen *et al.*, 1999; Böttcher *et al.*, 2004), leading to the development of ectopic tail-like phenotype in *X. laevis* embryos that we observed. This could perhaps be the most ancestral

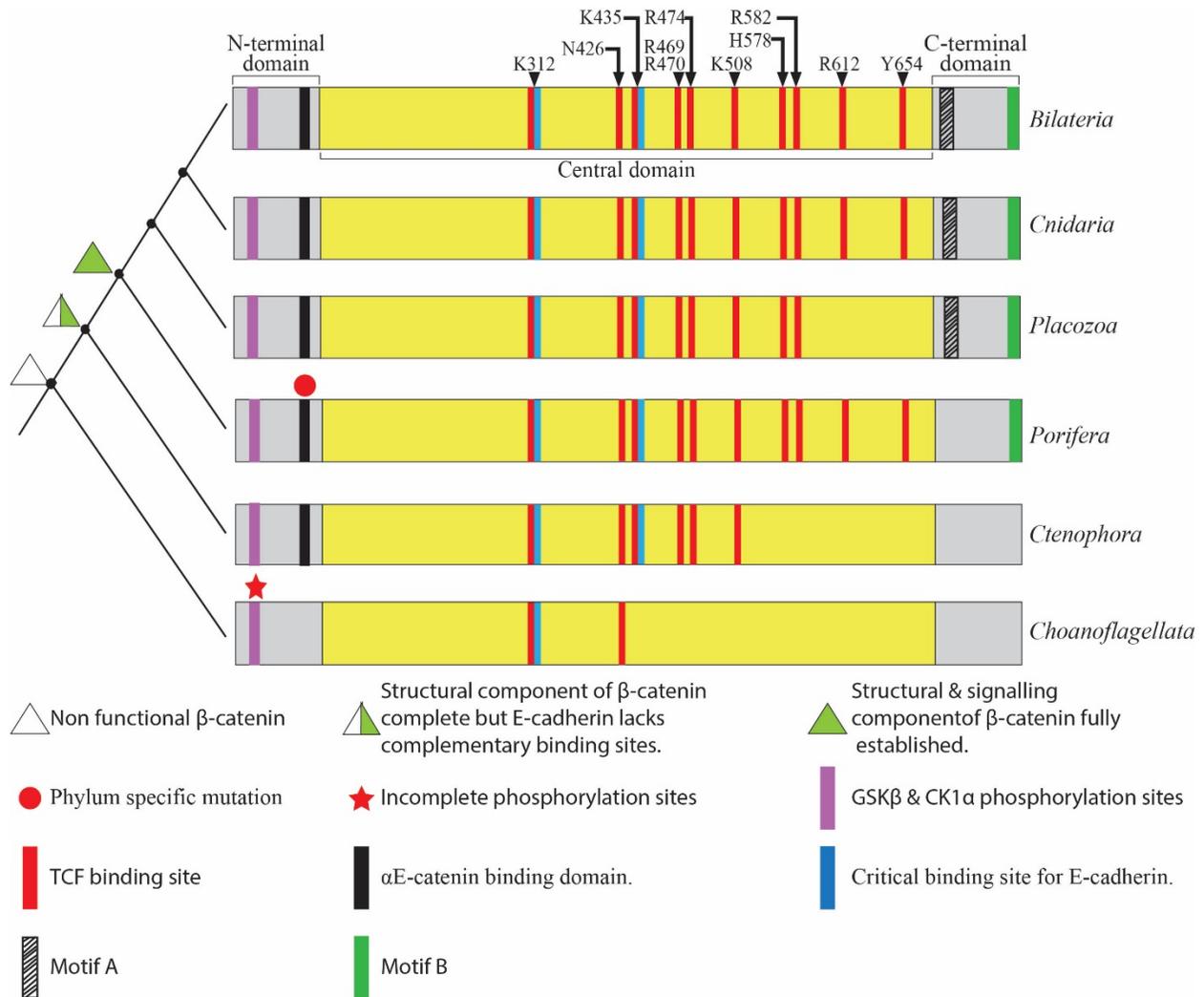
pathway of  $\beta$ -catenin functionality before it was modified to allow canonical Wnt signalling pathway.

In light of evidence from this research, I suggest a hypothesis that places ctenophores as the most basal metazoans (**Figure 6.1**). Support for this hypothesis stems from the following facts;

- 1) Ctenophores and choanoflagellates exhibit cadherins that lack a crucial  $\beta$ -catenin binding domain, hence not interacting with  $\beta$ -catenin. This could suggest that the CCC may not be structurally involved in cell-cell adhesion in either of these organisms.
- 2) Some residues critical in binding TCF are missing in ctenophore  $\beta$ -catenin, and two key motifs, A & B, in the C-terminal are absent. These absences could explain the observed low canonical Wnt signalling activity and the inability to induce a secondary body axis.

Therefore, the evolution of dual functionality of  $\beta$ -catenin can be described as follows:

- i) A  $\beta$ -catenin-like homolog was present in a choanoflagellate-like ancestor of metazoans. This homolog had conservation of some GSK3 $\beta$  phosphorylation sites to allow protein degradation and a single site critical for E-cadherin interaction. Reorganisation in the N-terminal of this protein probably played a part in the innovation of the  $\alpha$ E-catenin binding site in metazoans.
- ii) The emergence of multicellularity necessitated creating a new E-cadherin binding site along with sites to bind TCF, hence allowing signalling. However, signalling was likely weak in ctenophores, as seen in **Figure 5.5**, resulting in more TCF binding sites and another domain transactivation domain, motif B. Key sites to bind E-cadherin and  $\alpha$ E-catenin had been formed in ctenophores. However, interaction with E-cadherin was impossible due to a GBM in the catenin binding domain.
- iii) By the emergence of poriferans, a dual functional  $\beta$ -catenin had been established due to the creation of sites detailed above.
- iv) Finally, a new signalling domain, motif A, was established in the last common ancestor to bilaterian and cnidarians.



**Figure 6.1. The evolution of the dual functionality of  $\beta$ -catenin.** The architecture of  $\beta$ -catenin began as early as *Choanoflagellata* existence, and its dual functionality was established around the emergence of poriferans. *Ctenophora*  $\beta$ -catenin probably lacks structural and canonical Wnt signalling capabilities.

## 6.1 Conclusion

In summary, detailed sequence analysis combined with transphylectic and proteomic analysis has led us to conclude that ctenophores are most likely the most basal metazoans, and the  $\beta$ -catenin of the last common metazoan ancestor was equipped for adhesive roles, and the signalling roles evolved later. Our IP-MS data uncovered several novel  $\beta$ -catenin interacting proteins that are conserved in all metazoans. The presence of homologs in non-bilaterian phyla indicates that they may have played an essential role during the emergence of metazoans. Future studies could look into some of these proteins to understand their roles in development. Finally, although ctenophore  $\beta$ -catenin exhibited weak canonical Wnt signalling activity and lacked dorsal organizer induction capacity, it induced ectopic tail-like phenotypes. The pathway through which this phenotype is created could be the earliest function of  $\beta$ -catenin before committing to canonical Wnt signalling through new domain creation. Therefore, future research into understanding the earliest roles and targets of  $\beta$ -catenin could involve conducting functional experiments in ctenophores.

**References**

- Abedin, M. and King, N. (2008) 'The premetazoan ancestry of cadherins', *Science*, 319(5865), pp. 946–948.
- Abedin, M. and King, N. (2010) 'Diverse evolutionary paths to cell adhesion', *Trends in Cell Biology*. Elsevier Current Trends, pp. 734–742.
- Aberle, H., Schwartz, H., Hoschuetzky, H. and Kemler, R. (1996) 'Single amino acid substitutions in proteins of the armadillo gene family abolish their binding to  $\alpha$ -catenin', *Journal of Biological Chemistry*, 271(3), pp. 1520–1526.
- Adams, C. L. and Nelson, W. J. (1998) 'Cytomechanics of cadherin-mediated cell-cell adhesion', *Current Opinion in Cell Biology*, 10(5), pp. 572–577.
- Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., *et al.* (2007) 'Wnt and TGF- $\beta$  expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning', *PLoS ONE*, 2(10), p. e1031.
- Alberts, B. (1998) 'The cell as a collection of protein machines: Preparing the next generation of molecular biologists', *Cell*. Cell Press, pp. 291–294.
- Alharatani, R., Ververi, A., Belez-Meireles, A., Ji, W., *et al.* (2020) 'Novel truncating mutations in CTNND1 cause a dominant craniofacial and cardiac syndrome', *Human Molecular Genetics*, 29(11), p. 1900.
- Alié, A., Hayashi, T., Sugimura, I., Manuel, M., *et al.* (2015) 'The ancestral gene repertoire of animal stem cells.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(51), pp. E7093-100.
- Amit, C., Padmanabhan, P. and Narayanan, J. (2020) 'Deciphering the mechanoresponsive role of  $\beta$ -catenin in keratoconus epithelium', *Scientific Reports*, 10(1), pp. 1–16.
- Anderson, D. P., Whitney, D. S., Hanson-Smith, V., Woznica, A., *et al.* (2016) 'Evolution of an ancient protein function involved in organized multicellularity in animals', *eLife*, 5(JANUARY2016).
- Arce, L., Yokoyama, N. N. and Waterman, M. L. (2006) 'Diversity of LEF/TCF action in development and disease', *Oncogene*, 25(57), pp. 7492–7504.
- Babonis, L. S. and Martindale, M. Q. (2014) 'Old cell, new trick? Cnidocytes as a model for the evolution of novelty', in *Integrative and Comparative Biology*. Oxford University Press, pp. 714–722.
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., *et al.* (2009) 'MEME SUITE: tools for motif discovery and searching.', *Nucleic acids research*, 37(Web Server issue), pp. W202-8.
- Baker, D. and Sali, A. (2001) 'Protein structure prediction and structural genomics', *Science*. American Association for the Advancement of Science, pp. 93–96.
- Behrens, J., von Kries, J. P., Kühl, M., Bruhn, L., *et al.* (1996) 'Functional interaction of  $\beta$ -catenin with the transcription factor LEF-1', *Nature*, 382(6592), pp. 638–642.

- Belahbib, H., Renard, E., Santini, S., Jourda, C., *et al.* (2018) 'New genomic data and analyses challenge the traditional vision of animal epithelium evolution', *BMC Genomics*, 19(1), pp. 1–15.
- Bellipanni, G., Varga, M., Maegawa, S., Imai, Y., *et al.* (2006) 'Essential and opposing roles of zebrafish  $\beta$ -catenins in the formation of dorsal axial structures and neurectoderm', *Development*, 133(7), pp. 1299–1309.
- Benjamin, J. M. and Nelson, W. J. (2008) 'Bench to bedside and back again: Molecular mechanisms of  $\alpha$ -catenin function and roles in tumorigenesis', *Seminars in Cancer Biology*. *Semin Cancer Biol*, pp. 53–64.
- Blauwkamp, T. A., Chang, M. V and Cadigan, K. M. (2008) 'Novel TCF-binding sites specify transcriptional repression by Wnt signalling.', *The EMBO journal*, 27(10), pp. 1436–46.
- Böttcher, R. T., Pollet, N., Delius, H. and Niehrs, C. (2004) 'The transmembrane protein XFLRT3 forms a complex with FGF receptors and promotes FGF signalling', *Nature Cell Biology*, 6(1), pp. 38–44.
- Briscoe, J. and Thérond, P. P. (2013) 'Hedgehog (Hh) was first identified by genetic screens in *Drosophila melanogaster*', *Nature Publishing Group*, 14.
- Broun, M., Gee, L., Reinhardt, B. and Bode, H. R. (2005) 'Formation of the head organizer in hydra involves the canonical Wnt pathway', *Development*, 132(12), pp. 2907–2916.
- Brunet, T. and King, N. (2017) 'The Origin of Animal Multicellularity and Cell Differentiation', *Developmental Cell*. Cell Press, pp. 124–140.
- Chen, Y., Hollemann, T., Grunz, H. and Pieler, T. (1999) 'Characterization of the Ets-type protein ER81 in *Xenopus* embryos', *Mechanisms of Development*, 80(1), pp. 67–76.
- Chothia, C. and Lesk, A. M. (1986) 'The relation between the divergence of sequence and structure in proteins.', *The EMBO Journal*, 5(4), pp. 823–826.
- Clevers, H. (2006) 'Wnt/ $\beta$ -Catenin Signaling in Development and Disease', *Cell*, 127(3), pp. 469–480.
- Collins, A. G. (1998) 'Evaluating multiple alternative hypotheses for the origin of Bilateria: An analysis of 18S rRNA molecular evidence', *Proceedings of the National Academy of Sciences of the United States of America*, 95(26), pp. 15458–15463.
- Costa, M., Raich, W., Agbunag, C., Leung, B., *et al.* (1998) 'A putative catenin-cadherin system mediates morphogenesis of the caenorhabditis elegans embryo', *Journal of Cell Biology*, 141(1), pp. 297–308.
- Daniels, D. L. and Weis, W. I. (2002) 'ICAT inhibits  $\beta$ -catenin binding to tcf/lef-family transcription factors and the general coactivator p300 using independent structural modules', *Molecular Cell*, 10(3), pp. 573–584.
- Daniels, D. L. and Weis, W. I. (2005) ' $\beta$ -catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation', *NATURE STRUCTURAL & MOLECULAR BIOLOGY*, 12(4).

- Dann, C. E., Hsieh, J.-C., Rattner, A., Sharma, D., *et al.* (2001) 'Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains', *Nature*, 412(6842), pp. 86–90.
- Dayel, M. J., Alegado, R. A., Fairclough, S. R., Levin, T. C., *et al.* (2011) 'Cell differentiation and morphogenesis in the colony-forming choanoflagellate *Salpingoeca rosetta*', *Developmental Biology*, 357(1), pp. 73–82.
- Dellaporta, S. L., Xu, A., Sagasser, S., Jakob, W., *et al.* (2006) 'Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum', *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), pp. 8751–8756.
- Deorowicz, S., Debudaj-Grabysz, A. and Gudys, A. (2016) 'FAMSA: Fast and accurate multiple sequence alignment of huge protein families', *Scientific Reports*, 6(1), pp. 1–13.
- DiNardo, S., Wehrli, M., Dougan, S. T., Caldwell, K., *et al.* (2000) 'arrow encodes an LDL-receptor-related protein essential for Wingless signalling', *Nature*, 407(6803), pp. 527–530.
- DuBuc, T. Q., Ryan, J. F. and Martindale, M. Q. (2019) "'Dorsal–Ventral" Genes Are Part of an Ancient Axial Patterning System: Evidence from *Trichoplax adhaerens* (Placozoa)', *Molecular Biology and Evolution*. Edited by J. True, 36(5), pp. 966–973.
- DuChez, B. J., Hueschen, C. L., Zimmerman, S. P., Baumer, Y., *et al.* (2019) 'Characterization of the interaction between  $\beta$ -catenin and sorting nexin 27: Contribution of the type I PDZ-binding motif to Wnt signaling', *Bioscience Reports*, 39(11).
- Dunn, C. W., Hejnal, A., Matus, D. Q., Pang, K., *et al.* (2008) 'Broad phylogenomic sampling improves resolution of the animal tree of life', *Nature*, 452(7188), pp. 745–749.
- Edgar, R. C. (2004) 'MUSCLE: Multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Research*, 32(5), pp. 1792–1797.
- Ehyai, S., Miyake, T., Williams, D., Vinayak, J., *et al.* (2018) 'FMRP recruitment of  $\beta$ -catenin to the translation pre-initiation complex represses translation', *EMBO reports*, 19(12), p. e45536.
- Eitel, M., Osigus, H.-J., DeSalle, R. and Schierwater, B. (2013) 'Global Diversity of the Placozoa', *PLoS ONE*. Edited by H. Browman, 8(4), p. e57131.
- Fagotto, F. (2013) 'Looking beyond the Wnt pathway for the deep nature of  $\beta$ -catenin', *EMBO Reports*, pp. 422–433.
- Fairclough, S. R., Dayel, M. J. and King, N. (2010) 'Multicellular development in a choanoflagellate', *Current Biology*. Cell Press, pp. R875–R876.
- Farquhar, M. G. and Palade, G. E. (1963) 'Junctional complexes in various epithelia.', *The Journal of cell biology*, 17(2), pp. 375–412.
- Fasolini, M., Wu, X., Flocco, M., Trosset, J. Y., *et al.* (2003) 'Hot spots in Tcf4 for the interaction with  $\beta$ -catenin', *Journal of Biological Chemistry*, 278(23), pp. 21092–21098.
- Fernàndez-Busquets, X. and Burger, M. M. (1997) 'The main protein of the aggregation factor responsible for species-specific cell adhesion in the marine sponge *Microciona prolifera* is highly polymorphic', *Journal of Biological Chemistry*, 272(44), pp. 27839–27847.

- Funayama, N., Fagotto, F., McCrea, P. and Gumbiner, B. M. (1995) 'Embryonic axis induction by the armadillo repeat domain of beta-catenin: evidence for intracellular signaling', *The Journal of Cell Biology*, 128(5), p. 959.
- Gallet, A., Angelats, C., Erkner, A., Charroux, B., *et al.* (1999) 'The C-terminal domain of Armadillo binds to hypophosphorylated Teashirt to modulate Wingless signalling in *Drosophila*', *The EMBO Journal*, 18(8), pp. 2208–2217. Available at: <http://emboj.embopress.org/content/embojnl/18/8/2208.full.pdf> (Accessed: 30 August 2017).
- Ganot, P., Zoccola, D., Tambutté, E., Voolstra, C. R., *et al.* (2015) 'Structural Molecular Components of Septate Junctions in Cnidarians Point to the Origin of Epithelial Junctions in Eukaryotes', *Molecular Biology and Evolution*, 32(1), pp. 44–62.
- Gilson, A. I., Marshall-Christensen, A., Choi, J. M. and Shakhnovich, E. I. (2017) 'The Role of Evolutionary Selection in the Dynamics of Protein Structure Evolution', *Biophysical Journal*, 112(7), pp. 1350–1365.
- Graham, T. A., Weaver, C., Mao, F., Kimelman, D., *et al.* (2000) 'Crystal structure of a  $\beta$ -catenin/Tcf complex', *Cell*, 103(6), pp. 885–896.
- Green, K. J., Roth-Carter, Q., Niessen, C. M. and Nichols, S. A. (2020) 'Tracing the Origins of the Desmosome: a Vertebrate Innovation', *Current biology : CB*, 30(10), p. R535.
- Grice, L. F., Gauthier, M. E. A., Roper, K. E., Fernández-Busquets, X., *et al.* (2017) 'Origin and Evolution of the Sponge Aggregation Factor Gene Family', *Molecular Biology and Evolution*, 34(5), pp. 1083–1099.
- Griffin, J. N., del Viso, F., Duncan, A. R., Robson, A., *et al.* (2018) 'RAPGEF5 Regulates Nuclear Translocation of  $\beta$ -Catenin', *Developmental Cell*, 44(2), pp. 248-260.e4.
- Gumbiner, B. M. (2005) 'Regulation of cadherin-mediated adhesion in morphogenesis', *Nature Reviews Molecular Cell Biology* 2005 6:8, 6(8), pp. 622–634.
- Guo, Z., Neilson, L. J., Zhong, H., Murray, P. S., *et al.* (2014) 'E-cadherin interactome complexity and robustness resolved by quantitative proteomics', *Science Signaling*, 7(354), p. rs7.
- Ha, N. C., Tonzuka, T., Stamos, J. L., Choi, H. J., *et al.* (2004) 'Mechanism of phosphorylation-dependent binding of APC to  $\beta$ -catenin and its role in  $\beta$ -catenin degradation', *Molecular Cell*, 15(4), pp. 511–521.
- Haegel, H., Larue, L., Ohsugi, M., Fedorov, L., *et al.* (1995) 'Lack of beta-catenin affects mouse development at gastrulation', *Development*, 121(11). Available at: <http://dev.biologists.org/content/121/11/3529.long> (Accessed: 12 July 2017).
- Halbleib, J. M. and Nelson, W. J. (2006) 'Cadherins in development: Cell adhesion, sorting, and tissue morphogenesis', *Genes and Development*. Cold Spring Harbor Laboratory Press, pp. 3199–3214.
- Hanson, A. J., Wallace, H. A., Freeman, T. J., Beauchamp, R. D., *et al.* (2012) 'XIAP Monoubiquitylates Groucho/TLE to Promote Canonical Wnt Signaling', *Molecular Cell*, 45(5), pp. 619–628.

- Hart, M., Concordet, J. P., Lassot, I., Albert, I., *et al.* (1999) 'The F-box protein beta-TrCP associates with phosphorylated beta-catenin and regulates its activity in the cell.', *Current biology : CB*, 9(4), pp. 207–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10074433> (Accessed: 28 August 2017).
- Hecht, A., Vleminckx, K., Stemmler, M. P., van Roy, F., *et al.* (2000) 'The p300/CBP acetyltransferases function as transcriptional coactivators of beta-catenin in vertebrates.', *The EMBO journal*, 19(8), pp. 1839–50.
- Hey, P. J., Twells, R. C. J., Phillips, M. S., Yusuke Nakagawa, *et al.* (1998) 'Cloning of a novel member of the low-density lipoprotein receptor family', *Gene*, 216(1), pp. 103–111.
- Hinck, L., Näthke, I. S., Papkoff, J. and Nelson, W. J. (1994) 'Dynamics of cadherin/catenin complex formation: Novel protein interactions and pathways of complex assembly', *Journal of Cell Biology*, 125(6), pp. 1327–1340.
- Hobmayer, B., Rentzsch, F., Kuhn, K., Happel, C. M., *et al.* (2000) 'WNT signalling molecules act in axis formation in the diploblastic metazoan Hydra', *Nature 2000 407:6801*, 407(6801), pp. 186–189.
- Holstein, T. W., Watanabe, H. and Özbek, S. (2011) 'Signaling Pathways and Axis Formation in the Lower Metazoa', *Current Topics in Developmental Biology*, 97, pp. 137–177.
- Howard, M. J. (1998) 'Protein NMR spectroscopy.', *Current biology : CB*. Elsevier, pp. R331–R333.
- Hrckulak, D., Kolar, M., Strnad, H. and Korinek, V. (2016) 'TCF/LEF transcription factors: An update from the internet resources', *Cancers*. MDPI AG, pp. 330–337.
- Hsieh, J.-C., Rattner, A., Smallwood, P. M. and Nathans, J. (1999) 'Biochemical characterization of Wnt-Frizzled interactions using a soluble, biologically active vertebrate Wnt protein', *Biochemistry*, 96, pp. 3546–3551. Available at: <http://www.pnas.org/content/96/7/3546.full.pdf> (Accessed: 29 August 2017).
- Hsu, S.-C., Galceran, J. and Grosschedl, R. (1998) 'Modulation of Transcriptional Regulation by LEF-1 in Response to Wnt-1 Signaling and Association with  $\beta$ -Catenin', *Molecular and Cellular Biology*, 18(8), pp. 4807–4818.
- Huang, H. and He, X. (2008) 'Wnt/beta-catenin signaling: new (and old) players and new insights.', *Current opinion in cell biology*, 20(2), pp. 119–25.
- Huber, A. H., Stewart, D. B., Laurents, D. V., Nelson, W. J., *et al.* (2001) 'The cadherin cytoplasmic domain is unstructured in the absence of  $\beta$ -catenin. A possible mechanism for regulating cadherin turnover', *Journal of Biological Chemistry*, 276(15), pp. 12301–12309.
- Huber, A. H., Nelson, W. J. and Weis, W. I. (1997) 'Three-Dimensional Structure of the Armadillo Repeat Region of  $\beta$ -Catenin', *Cell*, 90(5), pp. 871–882.
- Huber, A. H. and Weis, W. I. (2001) 'The Structure of the  $\beta$ -Catenin/E-Cadherin Complex and the Molecular Basis of Diverse Ligand Recognition by  $\beta$ -Catenin', *Cell*, 105(3), pp. 391–402.
- Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., *et al.* (1996) 'Nuclear localization of  $\beta$ -catenin

- by interaction with transcription factor LEF-1', *Mechanisms of Development*, 59(1), pp. 3–10.
- Hudson, C., Kawai, N., Negishi, T. and Yasuo, H. (2013) 'β-catenin-driven binary fate specification segregates germ layers in ascidian embryos', *Current Biology*, 23(6), pp. 491–495.
- Hulpiau, P. and van Roy, F. (2009) 'Molecular evolution of the cadherin superfamily', *International Journal of Biochemistry and Cell Biology*. Pergamon, pp. 349–369.
- Hwang, J. R., Chou, C. L., Medvar, B., Knepper, M. A., *et al.* (2017) 'Identification of β-catenin-interacting proteins in nuclear fractions of native rat collecting duct cells', *American Journal of Physiology - Renal Physiology*, 313(1), pp. F30–F46.
- Ishiyama, N., Lee, S. H., Liu, S., Li, G. Y., *et al.* (2010) 'Dynamic and Static Interactions between p120 Catenin and E-Cadherin Regulate the Stability of Cell-Cell Adhesion', *Cell*, 141(1), pp. 117–128.
- Jager, M., Dayraud, C., Mialot, A., Quéinnec, E., *et al.* (2013) 'Evidence for involvement of Wnt signalling in body polarities, cell proliferation, and the neuro-sensory system in an adult ctenophore', *PLoS ONE*, 8(12).
- Jaimes, J. A., André, N. M., Chappie, J. S., Millet, J. K., *et al.* (2020) 'Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop', *Journal of Molecular Biology*, 432(10), pp. 3309–3325.
- Jamora, C., DasGupta, R., Kocieniewski, P. and Fuchs, E. (2003) 'Links between signal transduction, transcription and adhesion in epithelial bud development', *Nature*, 422(6929), pp. 317–322.
- Janda, C. Y., Dang, L. T., You, C., Chang, J., *et al.* (2017) 'Surrogate Wnt agonists that phenocopy canonical Wnt and β-catenin signalling', *Nature*, 545(7653), pp. 234–237.
- Jeanes, A., Gottardi, C. J. and Yap, A. S. (2008) 'Cadherins and cancer: How does cadherin dysfunction promote tumor progression?', *Oncogene*. *Oncogene*, pp. 6920–6929.
- Kanamori, M., Sandy, P., Marzinotto, S., Benetti, R., *et al.* (2003) 'The PDZ protein tax-interacting protein-1 inhibits beta-catenin transcriptional activity and growth of colorectal cancer cells.', *The Journal of biological chemistry*, 278(40), pp. 38758–64.
- Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S. C., *et al.* (2016) 'A eukaryote without a mitochondrial organelle', *Current Biology*, 26(10), pp. 1274–1284.
- Kelly, C., Chin, A. J., Leatherman, J. L., Kozlowski, D. J., *et al.* (2000) 'Maternally controlled (beta)-catenin-mediated signaling is required for organizer formation in the zebrafish', *Development*, 127(18).
- Khalturin, K., Anton-Erxleben, F., Milde, S., Plötz, C., *et al.* (2007) 'Transgenic stem cells in Hydra reveal an early evolutionary origin for key elements controlling self-renewal and differentiation', *Developmental Biology*, 309(1), pp. 32–44.
- Kimelman, D. and Xu, W. (2006) 'β-Catenin destruction complex: insights and questions from a structural perspective', *Oncogene*, 25(57), pp. 7482–7491.

- King, N., Westbrook, M. J., Young, S. L., Kuo, A., *et al.* (2008) ‘The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans’, *Nature*, 451(7180), pp. 783–788.
- Knoll, A. H. (2011) ‘The multiple origins of complex multicellularity’, *Annual Review of Earth and Planetary Sciences*, 39, pp. 217–239.
- Kobielak, A. and Fuchs, E. (2004) ‘ $\alpha$ -catenin: At the junction of intercellular adhesion and actin dynamics’, *Nature Reviews Molecular Cell Biology*, pp. 614–625.
- Kofron, M., Spagnuolo, A., Klymkowsky, M., Wylie, C., *et al.* (1997) ‘The roles of maternal  $\alpha$ -catenin and plakoglobin in the early *Xenopus* embryo’, *Development*, 124(8), pp. 1553–1560.
- Kozmikova, I. and Kozmik, Z. (2020) ‘Wnt/ $\beta$ -catenin signaling is an evolutionarily conserved determinant of chordate dorsal organizer’, *eLife*, 9.
- Kraus, Y., Fritzenwanker, J. H., Genikhovich, G. and Technau, U. (2007) ‘The blastoporal organiser of a sea anemone’, *Current Biology*.
- Kraus, Y., Aman, A., Technau, U. and Genikhovich, G. (2016) ‘Pre-bilaterian origin of the blastoporal axial organizer’, *Nature Communications*, 7(1), pp. 1–9.
- Von Kries, J. P., Winbeck, G., Asbrand, C., Schwarz-Romond, T., *et al.* (2000) ‘Hot spots in  $\beta$ -catenin for interactions with LEF-1, conduction and APC’, *Nature Structural Biology*, 7(9), pp. 800–807.
- Kusserow, A., Pang, K., Sturm, C., Hroudá, M., *et al.* (2005) ‘Unexpected complexity of the Wnt gene family in a sea anemone’, *Nature*, 433(7022), pp. 156–160.
- Lapébie, P., Gazave, E., Ereskovsky, A., Derelle, R., *et al.* (2009) ‘WNT/ $\beta$ -Catenin Signalling and Epithelial Patterning in the Homoscleromorph Sponge *Oscarella*’, *PLoS ONE*, 4(6).
- Lapébie, P., Ruggiero, A., Barreau, C., Chevalier, S., *et al.* (2014) ‘Differential Responses to Wnt and PCP Disruption Predict Expression and Developmental Function of Conserved and Novel Genes in a Cnidarian’, *PLoS Genetics*, 10(9), p. e1004590.
- Larue, L., Ohsugi, M., Hirchenhain, J. and Kemler, R. (1994) ‘E-cadherin null mutant embryos fail to form a trophectoderm epithelium’, *Proceedings of the National Academy of Sciences of the United States of America*, 91(17), pp. 8263–8267.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) ‘PROCHECK: a program to check the stereochemical quality of protein structures’, *Journal of Applied Crystallography*, 26(2), pp. 283–291.
- Lee, E., Salic, A., Krüger, R., Heinrich, R., *et al.* (2003) ‘The Roles of APC and Axin Derived from Experimental and Theoretical Analysis of the Wnt Pathway’, *PLoS Biology*. Edited by Roel Nusse, 1(1), p. e10.
- Leininger, S., Adamski, M., Bergum, B., Guder, C., *et al.* (2014) ‘Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans’, *Nature Communications*, 5(1), pp. 1–15.
- Li, H., Zhang, W., Yan, M., Qiu, J., *et al.* (2019) ‘Nucleolar and spindle associated protein 1

promotes metastasis of cervical carcinoma cells by activating Wnt/ $\beta$ -catenin signaling', *Journal of Experimental & Clinical Cancer Research* 2019 38:1, 38(1), pp. 1–18.

Li, Y., Shen, X.-X., Evans, B., Dunn, C. W., *et al.* (2021) 'Rooting the Animal Tree of Life', *Molecular Biology and Evolution*.

Liang, R. and Liu, Y. (2018) 'Tcf711 directly regulates cardiomyocyte differentiation in embryonic stem cells 06 Biological Sciences 0604 Genetics 06 Biological Sciences 0601 Biochemistry and Cell Biology', *Stem Cell Research and Therapy*, 9(1), p. 267.

Liu, C., Kato, Y., Zhang, Z., Do, V. M., *et al.* (1999) ' $\beta$ -Trcp couples  $\beta$ -catenin phosphorylation-degradation and regulates *Xenopus* axis formation', *Proceedings of the National Academy of Sciences of the United States of America*, 96(11), pp. 6273–6278.

Liu, C., Li, Y., Semenov, M., Han, C., *et al.* (2002) 'Control of  $\beta$ -catenin phosphorylation/degradation by a dual-kinase mechanism', *Cell*, 108(6), pp. 837–847.

Lu, T., Bao, Z., Wang, Y., Yang, L., *et al.* (2016) 'Karyopherin $\beta$ 1 regulates proliferation of human glioma cells via Wnt/ $\beta$ -catenin pathway', *Biochemical and Biophysical Research Communications*, 478(3), pp. 1189–1197.

Maas, O. (1906) 'Über die Einwirkung karbonatfreier und kalkfreier Salzlösungen auf erwachsene Kalkschwämme und auf Entwicklungsstadien derselben', *Archiv für Entwicklungsmechanik der Organismen*, 22(4), pp. 581–599.

MacDonald, B. T., Tamai, K. and He, X. (2009) 'Wnt/ $\beta$ -Catenin Signaling: Components, Mechanisms, and Diseases', *Developmental Cell*, 17(1), pp. 9–26.

MacWilliams, H. K. (1983) 'Hydra transplantation phenomena and the mechanism of Hydra head regeneration. II. Properties of the head activation', *Developmental Biology*, 96(1), pp. 239–257.

Maeda, O., Usami, N., Kondo, M., Takahashi, M., *et al.* (2003) 'Plakoglobin ( $\gamma$ -catenin) has TCF/LEF family-dependent transcriptional activity in  $\beta$ -catenin-deficient cell line', *Oncogene* 2004 23:4, 23(4), pp. 964–972.

Magie, C. R. and Martindale, M. Q. (2008) 'Cell-cell adhesion in the cnidaria: Insights into the evolution of tissue morphogenesis', *Biological Bulletin*. Marine Biological Laboratory, pp. 218–232.

Margariti, A., Zampetaki, A., Xiao, Q., Zhou, B., *et al.* (2010) 'Molecular Medicine Histone Deacetylase 7 Controls Endothelial Cell Growth Through Modulation of-Catenin'.

Mariner, D. J., Wang, J. and Reynolds, A. B. (2000) 'ARVCF localizes to the nucleus and adherens junction and is mutually exclusive with p120(ctn) in E-cadherin complexes', *Journal of Cell Science*, 113(8), pp. 1481–1490.

Martindale, M. Q., Pang, K. and Finnerty, J. R. (2004) 'Investigating the origins of triplosblasty: "Mesodermal" gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa)', *Development*, 131(10), pp. 2463–2474.

Medini, H., Cohen, T. and Mishmar, D. (2020) 'Annual Review of Genetics Mitochondria Are

- Fundamental for the Emergence of Metazoans: On Metabolism, Genomic Regulation, and the Birth of Complex Organisms’.
- Mellman, I. and Nelson, W. J. (2008) ‘Coordinated protein sorting, targeting and distribution in polarized cells’, *Nature Reviews Molecular Cell Biology*, 9(11), pp. 833–845.
- Messerschmidt, D., De Vries, W. N., Lorthongpanich, C., Balu, S., *et al.* (2016) ‘ $\beta$ -catenin-mediated adhesion is required for successful preimplantation mouse embryo development’, *Development (Cambridge)*, 143(11), pp. 1993–1999.
- Michalina, X., †1, J., Candido Primi, M. and Izard, T. (2020) ‘Cell adhesion in cancer: Beyond the migration of single cells’.
- Miller, P. W., Clarke, D. N., Weis, W. I., Lowe, C. J., *et al.* (2013) ‘The Evolutionary Origin of Epithelial Cell–Cell Adhesion Mechanisms’, in *Current Topics in Membranes*. Academic Press Inc., pp. 267–311.
- Mitchell, J. M. and Nichols, S. A. (2019) ‘Diverse cell junctions with unique molecular composition in tissues of a sponge (Porifera)’, *EvoDevo*, 10(1), pp. 1–16.
- Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., *et al.* (1996) ‘XTcf-3 Transcription Factor Mediates  $\beta$ -Catenin-Induced Axis Formation in Xenopus Embryos’, *Cell*, 86(3), pp. 391–399.
- Momose, T. and Houlston, E. (2007) ‘Two oppositely localised frizzled RNAs as axis determinants in a cnidarian embryo’, *PLoS Biology*, 5(4), pp. 889–899.
- Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., *et al.* (2014) ‘The ctenophore genome and the evolutionary origins of neural systems’, *Nature*, 510(7503), pp. 109–114.
- Moroz, L. L. (2015) ‘Convergent evolution of neural systems in ctenophores’, *Journal of Experimental Biology*. Company of Biologists Ltd, pp. 598–611.
- Mosimann, C., Hausmann, G. and Basler, K. (2009) ‘ $\beta$ -Catenin hits chromatin: regulation of Wnt target gene activation’, *Nature Reviews Molecular Cell Biology*, 10(4), pp. 276–286.
- Murray, P. S. and Zaidel-Bar, R. (2014) ‘Pre-metazoan origins and evolution of the cadherin adhesome’, *Biology Open*, 3(12), pp. 1183–1195.
- Mutze, K., Vierkotten, S., Milosevic, J., Eickelberg, O., *et al.* (2015) ‘Enolase 1 (ENO1) and protein disulfide-isomerase associated 3 (PDIA3) regulate Wnt/ $\beta$ -catenin-driven trans-differentiation of murine alveolar epithelial cells’, *DMM Disease Models and Mechanisms*, 8(8), pp. 877–890.
- Nagafuchi, A. and Takeichi, M. (1988) ‘Cell binding function of E-cadherin is regulated by the cytoplasmic domain.’, *The EMBO journal*, 7(12), pp. 3679–3684.
- Nakamura, Y., de Paiva Alves, E., Veenstra, G. J. C. and Hoppler, S. (2016) ‘Tissue- and stage-specific Wnt target gene expression is controlled subsequent to  $\beta$ -catenin recruitment to cis-regulatory modules.’, *Development (Cambridge, England)*, 143(11), pp. 1914–25.
- Nakanishi, N., Sogabe, S. and Degnan, B. M. (2014) ‘Evolutionary origin of gastrulation: insights from sponge development’, *BMC Biology 2014 12:1*, 12(1), pp. 1–9.

- Di Nardo, G., Zhang, C., Marcelli, A. G. and Gilardi, G. (2021) 'Molecular and structural evolution of cytochrome p450 aromatase', *International Journal of Molecular Sciences*, 22(2), pp. 1–16.
- Nathaniel Clarke, D., Lowe, C. J. and James Nelson, W. (2019) 'The cadherin-catenin complex is necessary for cell adhesion and embryogenesis in *Nematostella vectensis*', *Developmental Biology*, 447(2), pp. 170–181.
- Nelson, W. J. (2008) 'Regulation of cell–cell adhesion by the cadherin–catenin complex', *Biochemical Society Transactions*, 36(Pt 2), p. 149.
- Neugebauer, J. M., Amack, J. D., Peterson, A. G., Bisgrove, B. W., *et al.* (2009) 'FGF signalling during embryo development regulates cilia length in diverse epithelia', *Nature*, 458(7238), pp. 651–654.
- Nichols, S. A., Roberts, B. W., Richter, D. J., Fairclough, S. R., *et al.* (2012) 'Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/ -catenin complex', *Proceedings of the National Academy of Sciences*, 109(32), pp. 13046–13051.
- Nusse, R. and Varmus, H. E. (1982) 'Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome', *Cell*, 31(1), pp. 99–109.
- Nüsslein-Volhard, C. and Wieschaus, E. (1980) 'Mutations affecting segment number and polarity in *Drosophila*', *Nature*, 287(5785), pp. 795–801.
- Orsulic, S. and Peifer, M. (1996) 'An in vivo structure-function study of armadillo, the beta-catenin homologue, reveals both separate and overlapping regions of the protein required for cell adhesion and for wingless signaling.', *The Journal of cell biology*, 134(5), pp. 1283–300. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8794868> (Accessed: 30 August 2017).
- Osigus, H. J., Rolfes, S., Herzog, R., Kamm, K., *et al.* (2019) 'Polyplacotoma mediterranea is a new ramified placozoan species', *Current Biology*. Cell Press, pp. R148–R149.
- Pang, K., Ryan, J. F., Mullikin, J. C., Baxevanis, A. D., *et al.* (2010) 'Genomic insights into Wnt signaling in an early diverging metazoan, the ctenophore *Mnemiopsis leidyi*', *EvoDevo*, 1(1), pp. 1–15.
- Pang, K. and Martindale, M. Q. (2008) 'Comb jellies (Ctenophora): A model for basal metazoan evolution and development', *Cold Spring Harbor Protocols*, 3(11), p. pdb.emo106.
- Paps, J. and Holland, P. W. H. (2018) 'Reconstruction of the ancestral metazoan genome reveals an increase in genomic novelty', *Nature Communications*, 9(1).
- Park, J. Il, Kim, S. W., Lyons, J. P., Ji, H., *et al.* (2005) 'Kaiso/p120-Catenin and TCF/ $\beta$ -Catenin Complexes Coordinately Regulate Canonical Wnt Gene Targets', *Developmental Cell*, 8(6), pp. 843–854.
- Perego, C., Vanoni, C., Massari, S., Longhi, R., *et al.* (2000) 'Mammalian LIN-7 PDZ proteins associate with beta-catenin at the cell-cell junctions of epithelia and neurons.', *The EMBO journal*, 19(15), pp. 3978–89.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., *et al.* (2004) 'UCSF Chimera - A

- visualization system for exploratory research and analysis’, *Journal of Computational Chemistry*, 25(13), pp. 1605–1612.
- Piedra, J., Miravet, S., Castaño, J., Pálmer, H. G., *et al.* (2003) ‘p120 Catenin-Associated Fer and Fyn Tyrosine Kinases Regulate  $\beta$ -Catenin Tyr-142 Phosphorylation and  $\beta$ -Catenin- $\alpha$ -Catenin Interaction’, *Molecular and Cellular Biology*, 23(7), pp. 2287–2297.
- Piepenburg, O., Vorbrüggen, G. and Jäckle, H. (2000) ‘Drosophila Segment Borders Result from Unilateral Repression of Hedgehog Activity by WINGLESS Signaling’, *Molecular Cell*, 6(1), pp. 203–209.
- Podar, M., Haddock, S. H. D., Sogin, M. L. and Harbison, G. R. (2001) ‘A molecular phylogenetic framework for the phylum Ctenophora using 18S rRNA genes’, *Molecular Phylogenetics and Evolution*, 21(2), pp. 218–230.
- Provost, E., Yamamoto, Y., Lizardi, I., Stern, J., *et al.* (2003) ‘Functional correlates of mutations in beta-catenin exon 3 phosphorylation sites.’, *The Journal of biological chemistry*, 278(34), pp. 31781–9.
- Pukhlyakova, E. A., Kirillova, A. O., Kraus, Y. A., Zimmermann, B., *et al.* (2019) ‘A cadherin switch marks germ layer formation in the diploblastic sea anemone *Nematostella vectensis*’, *Development (Cambridge)*, 146(20).
- Qi, W., Chen, J., Cheng, X., Huang, J., *et al.* (2015) ‘Targeting the Wnt-Regulatory Protein CTNNBIP1 by microRNA-214 Enhances the Stemness and Self-Renewal of Cancer Stem-Like Cells in Lung Adenocarcinomas’, *STEM CELLS*, 33(12), pp. 3423–3436.
- Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. (1963) ‘Stereochemistry of polypeptide chain configurations’, *Journal of Molecular Biology*. J Mol Biol, pp. 95–99.
- Rechsteiner, M. and Rogers, S. W. (1996) ‘PEST sequences and regulation by proteolysis’, *Trends in Biochemical Sciences*. Elsevier Ltd, pp. 267–271.
- Ribeiro, A. S., Sousa, B., Carreto, L., Mendes, N., *et al.* (2013) ‘P-cadherin functional role is dependent on E-cadherin cellular context: a proof of concept using the breast cancer model’, *The Journal of Pathology*, 229(5), pp. 705–718.
- Röttinger, E., Dahlin, P. and Martindale, M. Q. (2012) ‘A framework for the establishment of a cnidarian gene regulatory network for “endomesoderm” specification: the inputs of  $\beta$ -catenin/TCF signaling.’, *PLoS genetics*, 8(12), p. e1003164.
- Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A. D., *et al.* (2013) ‘The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution’, *Science*, 342(6164).
- Sáez, J. C., Berthoud, V. M., Brañes, M. C., Martínez, A. D., *et al.* (2003) ‘Plasma membrane channels formed by connexins: Their regulation and functions’, *Physiological Reviews*. American Physiological Society, pp. 1359–1400.
- Salinas-Saavedra, M., Rock, A. Q. and Martindale, M. Q. (2018) ‘Germ layer-specific regulation of cell polarity and adhesion gives insight into the evolution of mesoderm’, *eLife*, 7.
- Salinas-Saavedra, M., Wikramanayake, A. and Martindale, M. (2019) ‘ $\beta$ -catenin has an ancestral

- role in cell fate specification but not cell adhesion’, *bioRxiv*, p. 520957.
- Sampietro, J., Dahlberg, C. L., Cho, U. S., Hinds, T. R., *et al.* (2006) ‘Crystal Structure of a  $\beta$ -Catenin/BCL9/Tcf4 Complex’, *Molecular Cell*, 24(2), pp. 293–300.
- Sarpal, R., Pellikka, M., Patel, R. R., Hui, F. Y. W., *et al.* (2012) ‘Mutational analysis supports a core role for Drosophila  $\alpha$ -Catenin in adherens junction function’, *Development*, 139(5).
- Satow, R., Shitashige, M., Jigami, T., Fukami, K., *et al.* (2012) ‘ $\beta$ -catenin inhibits promyelocytic leukemia protein tumor suppressor function in colorectal cancer cells’, *Gastroenterology*, 142(3), pp. 572–581.
- Schenkelaars, Q., Pralong, M., Kodjabachian, L., Fierro-Constain, L., *et al.* (2017) ‘Animal multicellularity and polarity without Wnt signaling’, *Scientific Reports 2017 7:1*, 7(1), pp. 1–11.
- Schierwater, B., Eitel, M., Jakob, W., Osigus, H.-J., *et al.* (2009) ‘Concatenated Analysis Sheds Light on Early Metazoan Evolution and Fuels a Modern “Urmotazoon” Hypothesis’, *PLoS Biology*. Edited by D. Penny, 7(1), p. e1000020.
- Schippers, K. J. and Nichols, S. A. (2018) ‘Evidence of Signaling and Adhesion Roles for  $\beta$ -Catenin in the Sponge *Ephydatia muelleri*’, *Molecular Biology and Evolution*. Edited by P. Wittkopp, 35(6), pp. 1407–1421.
- Schneider, S. Q., Finnerty, J. R. and Martindale, M. Q. (2003) ‘Protein evolution: structure-function relationships of the oncogene beta-catenin in the evolution of multicellular animals’, *Journal of Experimental Zoology*, 295B(1), pp. 25–44.
- Schohl, A. and Fagotto, F. (2002) ‘ $\beta$ -catenin, MAPK and Smad signaling during early *Xenopus* development’, *Development*, 129(1), pp. 37–52.
- Schohl, A. and Fagotto, F. (2003) ‘A role for maternal  $\beta$ -catenin in early mesoderm induction in *Xenopus*’, *EMBO Journal*, 22(13), pp. 3303–3313.
- Schwarz-Romond, T., Fiedler, M., Shibata, N., Butler, P. J. G., *et al.* (2007) ‘The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization’, *Nature Structural & Molecular Biology*, 14(6), pp. 484–492.
- Semaan, C., Henderson, B. R. and Molloy, M. P. (2019) ‘Proteomic screen with the proto-oncogene beta-catenin identifies interaction with Golgi coatamer complex I’, *Biochemistry and Biophysics Reports*, 19.
- Shi, S. and Stanley, P. (2006) ‘Evolutionary Origins of Notch Signaling in Early Development’, *Cell Cycle Landes Bioscience Cell Cycle*, 53(3), pp. 274–278. Available at: <http://www.landesbioscience.com/journals/cc/abstract.php?id=2396> (Accessed: 11 July 2017).
- Shibata, T., Chuma, M., Kokubu, A., Sakamoto, M., *et al.* (2003) ‘EBP50, a  $\beta$ -catenin-associating protein, enhances Wnt signaling and is over-expressed in hepatocellular carcinoma’, *Hepatology*, 38(1), pp. 178–186.
- Shimizu, M., Fukunaga, Y., Ikenouchi, J. and Nagafuchi, A. (2008) ‘Defining the Roles of  $\beta$ -Catenin and Plakoglobin in LEF/T-Cell Factor-Dependent Transcription Using  $\beta$ -Catenin/Plakoglobin-Null F9 Cells’, *Molecular and Cellular Biology*, 28(2), pp. 825–835.

- Shimoyama, Y., Takeda, H., Yoshihara, S., Kitajima, M., *et al.* (1999) 'Biochemical characterization and functional analysis of two type II classic cadherins, cadherin-6 and -14, and comparison with E-cadherin', *Journal of Biological Chemistry*, 274(17), pp. 11987–11994.
- Smith, C. L., Varoqueaux, F., Kittelmann, M., Azzam, R. N., *et al.* (2014a) 'Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*', *Current Biology*, 24(14), pp. 1565–1572.
- Smith, C. L., Varoqueaux, F., Kittelmann, M., Azzam, R. N., *et al.* (2014b) 'Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*', *Current Biology*, 24(14), pp. 1565–1572.
- Smith, C. L., Pivovarov, N. and Reese, T. S. (2015) 'Coordinated Feeding Behavior in *Trichoplax*, an Animal without Synapses', *PLOS ONE*. Edited by R. E. Steele, 10(9), p. e0136098.
- Smith, C. L. and Reese, T. S. (2016) 'Adherens junctions modulate diffusion between epithelial cells in *Trichoplax adhaerens*', *Biological Bulletin*, 231(3), pp. 216–224.
- Smith, W. C. and Harland, R. M. (1991) 'Injected Xwnt-8 RNA acts early in *Xenopus* embryos to promote formation of a vegetal dorsalizing center', *Cell*, 67(4), pp. 753–765.
- Sokol, S., Christian, J. L., Moon, R. T. and Melton, D. A. (1991) 'Injected Wnt RNA induces a complete body axis in *Xenopus* embryos', *Cell*, 67(4), pp. 741–752.
- Sopko, R. and Perrimon, N. (2013) 'Receptor tyrosine kinases in *Drosophila* development', *Cold Spring Harbor Perspectives in Biology*, 5(6), p. a009050.
- Spemann, H. and Mangold, H. (1924) 'über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren', *Archiv für Mikroskopische Anatomie und Entwicklungsmechanik*, 100(3–4), pp. 599–638.
- Srivastava, M., Begovic, E., Chapman, J., Putnam, N. H., *et al.* (2008) 'The *Trichoplax* genome and the nature of placozoans', *Nature*, 454(7207), pp. 955–960.
- Srivastava, M., Simakov, O., Chapman, J., Fahey, B., *et al.* (2010) 'The *Amphimedon queenslandica* genome and the evolution of animal complexity', *Nature*, 466(7307), pp. 720–726.
- Standley, H. J., Destrée, O., Kofron, M., Wylie, C., *et al.* (2006) 'Maternal XTcf1 and XTcf4 have distinct roles in regulating Wnt target genes', *Developmental Biology*, 289(2), pp. 318–328.
- Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., *et al.* (2006) 'BioGRID: a general repository for interaction datasets.', *Nucleic acids research*, 34(Database issue).
- Straub, B. K., Boda, J., Kuhn, C., Schnoelzer, M., *et al.* (2003) 'A novel cell-cell junction system: The cortex *adhaerens* mosaic of lens fiber cells', *Journal of Cell Science*. The Company of Biologists Ltd, pp. 4985–4995.
- Su, Y., Fu, C., Ishikawa, S., Stella, A., *et al.* (2008) 'APC Is Essential for Targeting Phosphorylated  $\beta$ -Catenin to the SCF $\beta$ -TrCP Ubiquitin Ligase', *Molecular Cell*, 32(5), pp. 652–661.

- Sun, Y., Aiga, M., Yoshida, E., Humbert, P. O., *et al.* (2009) 'Scribble Interacts with  $\beta$ -Catenin to Localize Synaptic Vesicles to Synapses', *Molecular Biology of the Cell*. Edited by T. F. J. Martin, 20(14), pp. 3390–3400.
- Taelman, V. F., Dobrowolski, R., Plouhinec, J.-L., Fuentealba, L. C., *et al.* (2010) 'Wnt Signaling Requires Sequestration of Glycogen Synthase Kinase 3 inside Multivesicular Endosomes', *Cell*, 143(7), pp. 1136–1148.
- Taoufiq, Z., Ninov, M., Villar-Briones, A., Wang, H. Y., *et al.* (2021) 'Hidden proteome of synaptic vesicles in the mammalian brain', *Proceedings of the National Academy of Sciences of the United States of America*, 117(52), pp. 33586–33596.
- Technau, U. and Scholz, C. B. (2003) 'Origin and evolution of endoderm and mesoderm', *International Journal of Developmental Biology*. UPV/EHU Press, pp. 531–539.
- Teo, R., Möhrle, F., Plickert, G., Müller, W. A., *et al.* (2006) 'An evolutionary conserved role of Wnt signaling in stem cell fate decision', *Developmental Biology*, 289(1), pp. 91–99.
- Theisen, H., Syed, A., Nguyen, B. T., Lukacsovich, T., *et al.* (2007) 'Wingless Directly Represses DPP Morphogen Expression via an Armadillo/TCF/Brinker Complex', *PLoS ONE*. Edited by C.-P. Heisenberg, 2(1), p. e142.
- Thul, P. J., Akesson, L., Wiking, M., Mahdessian, D., *et al.* (2017) 'A subcellular map of the human proteome', *Science*, 356(6340).
- Tian, Q., Feetham, M. C., Tao, W. A., He, X. C., *et al.* (2004) *Proteomic analysis identifies that 14-3-3 interacts with-catenin and facilitates its activation by Akt*. Available at: [www.pnas.org/cgi/doi/10.1073/pnas.0406499101](http://www.pnas.org/cgi/doi/10.1073/pnas.0406499101).
- Tomii, K., Hirokawa, T. and Motoso, C. (2005) 'Protein structure prediction using a variety of profile libraries and 3D verification', *Proteins: Structure, Function, and Bioinformatics*, 61(S7), pp. 114–121.
- Valenta, T., Hausmann, G. and Basler, K. (2012) 'The many faces and functions of  $\beta$ -catenin.', *The EMBO journal*, 31(12), pp. 2714–36.
- del Valle-Pérez, B., Casagolda, D., Lugilde, E., Valls, G., *et al.* (2011) 'Wnt controls the transcriptional activity of Kaiso through CK1 $\epsilon$ -dependent phosphorylation of p120-catenin', *Journal of Cell Science*, 124(13), pp. 2298–2309.
- Veeman, M. T., Slusarski, D. C., Kaykas, A., Louie, S. H., *et al.* (2003) 'Zebrafish prickle, a modulator of noncanonical Wnt/Fz signaling, regulates gastrulation movements', *Current Biology*, 13(8), pp. 680–685.
- Virginia, R. P., Jahan, N., Okada, M., Takebayashi-Suzuki, K., *et al.* (2019) 'Cdc2-like kinase 2 (Clk2) promotes early neural development in *Xenopus* embryos', *Development, Growth & Differentiation*, 61(6), pp. 365–377.
- Vleminckx, K., Kemler, R. and Hecht, A. (1999) 'The C-terminal transactivation domain of  $\beta$ -catenin is necessary and sufficient for signaling by the LEF-1/ $\beta$ -catenin complex in *Xenopus laevis*', *Mechanisms of Development*, 81(1–2), pp. 65–74.

- Wang, C., Wang, H., Peng, Y., Zeng, B., *et al.* (2021) 'CTNNBIP1 modulates keratinocyte proliferation through promoting the transcription of  $\beta$ -catenin/TCF complex downstream genes', *Journal of the European Academy of Dermatology and Venereology*, 35(2), pp. 368–379.
- Wang, R. N., Green, J., Wang, Z., Deng, Y., *et al.* (2014) 'Bone Morphogenetic Protein (BMP) signaling in development and human diseases', *Genes & Diseases*, 1(1), pp. 87–105.
- Watanabe, H., Kuhn, A., Fushiki, M., Agata, K., *et al.* (2014) 'Sequential actions of  $\beta$ -catenin and Bmp pattern the oral nerve net in *Nematostella vectensis*', *Nature Communications*, 5.
- Webb, B. and Sali, A. (2016) 'Comparative protein structure modeling using MODELLER', *Current Protocols in Bioinformatics*, 2016, pp. 5.6.1-5.6.37.
- Van de Wetering, M., Cavallo, R., Dooijes, D., Van Beest, M., *et al.* (1997) 'Armadillo coactivates transcription driven by the product of the *Drosophila* segment polarity gene dTCF', *Cell*, 88(6), pp. 789–799.
- Wikramanayake, A. H., Hong, M., Lee, P. N., Pang, K., *et al.* (2003) 'An ancient role for nuclear  $\beta$ -catenin in the evolution of axial polarity and germ layer segregation', *Nature*, 426(6965), pp. 446–450.
- Wong, H.-C., Bourdelas, A., Krauss, A., Lee, H.-J., *et al.* (2003) 'Direct Binding of the PDZ Domain of Dishevelled to a Conserved Internal Sequence in the C-Terminal Region of Frizzled', *Molecular cell*, 12(5), p. 1251. Available at: [/pmc/articles/PMC4381837/](https://pubmed.ncbi.nlm.nih.gov/12510000/) (Accessed: 17 August 2021).
- Worth, C. L., Gong, S. and Blundell, T. L. (2009) 'Structural and functional constraints in the evolution of protein families', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, pp. 709–720.
- Wu, G., Huang, H., Abreu, J. G., He, X., *et al.* (2009) 'Inhibition of GSK3 Phosphorylation of  $\beta$ -Catenin via Phosphorylated PPPSPXS Motifs of Wnt Coreceptor LRP6', *PLoS ONE*. Edited by D.-Y. Jin, 4(3), p. e4926.
- Xiao, K., Oas, R. G., Chiasson, C. M. and Kowalczyk, A. P. (2007) 'Role of p120-catenin in cadherin trafficking', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1773(1), pp. 8–16.
- Xing, Y., Clements, W. K., Kimelman, D. and Xu, W. (2003) 'Crystal structure of a  $\beta$ -catenin/Axin complex suggests a mechanism for the  $\beta$ -catenin destruction complex', *Genes and Development*, 17(22), pp. 2753–2764.
- Xing, Y., Takemaru, K. I., Liu, J., Berndt, J. D., *et al.* (2008) 'Crystal Structure of a Full-Length  $\beta$ -Catenin', *Structure*, 16(3), pp. 478–487.
- Xu, L. and Massagué, J. (2004) 'Nucleocytoplasmic shuttling of signal transducers', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, pp. 209–219.
- Xu, M., Jin, T., Chen, L., Zhang, X., *et al.* (2019) 'Mortalin is a distinct bio-marker and prognostic factor in serous ovarian carcinoma', *Gene*, 696, pp. 63–71.
- Xu, Z., He, W., Ke, T., Zhang, Y., *et al.* (2020) 'DHRS12 inhibits the proliferation and

- metastasis of osteosarcoma via Wnt3a/ $\beta$ -catenin pathway', *Future oncology*, 16(11), pp. 665–674.
- Yamada, K. and Tomii, K. (2014) 'Revisiting amino acid substitution matrices for identifying distantly related proteins', *Bioinformatics*, 30(3), pp. 317–325.
- Yost, C., Torres, M., Miller, J. R., Huang, E., *et al.* (1996) 'The axis-inducing activity, stability, and subcellular distribution of  $\beta$ -catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3', *Genes and Development*, 10(12), pp. 1443–1454.
- Zachar, I. and Szathmáry, E. (2017) 'Breath-giving cooperation: Critical review of origin of mitochondria hypotheses', *Biology Direct*. BioMed Central Ltd., pp. 1–26.
- Zhan, T., Rindtorff, N. and Boutros, M. (2017) 'Wnt signaling in cancer', *Oncogene*, 36(11), pp. 1461–1473.
- Zhao, Z. M., Reynolds, A. B. and Gaucher, E. A. (2011) 'The evolutionary history of the catenin gene family during metazoan evolution', *BMC Evolutionary Biology*, 11(1), pp. 1–13.
- Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., *et al.* (2018) 'A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core', *Journal of Molecular Biology*, 430(15), pp. 2237–2243.

## Appendices

### Appendix 1

#### Species used in $\beta$ -catenin domain analysis.

Accession IDs or scaffold references where candidate genes were sourced from different databases. Accession IDs in red were translated to respective protein sequences using ExpASy ([web.expasy.org/translate/](http://web.expasy.org/translate/)). Asterisks, transcriptomic analysis carried out by our group.

Taxa	Species	Accession ID	Source
Bilateria	<i>Mus musculus</i>	NP_001159374.1	GenBank
	<i>Xenopus laevis</i>	NP_001080749.1	GenBank
	<i>Danio rerio</i>	NP_571134.2	GenBank
	<i>Drosophila melanogaster</i>	NP_476666.1	GenBank
	<i>Branchiostoma belcheri</i>	XP_019614271.1	GenBank
	<i>Platynereis dumerilii</i>	ABQ85061.1	GenBank
	<i>Artemia sinica</i>	ADC35062.1	GenBank
	<i>Lingula anatina</i>	XP_013379954.1	GenBank
Cnidaria	<i>Nematostella vectensis</i>	XP_001647517.2	GenBank
	<i>Aurelia aurita</i>	scaffold25.g74.t1	Ref 1
	<i>Hydra vulgaris</i>	NP_001267780.1	GenBank
	<i>Acropora digitifera</i>	XP_015769412.1	GenBank
Porifera	<i>Ephydatia fluviatilis</i>	m.31095	Ref 2
	<i>Amphimedon queenslandica</i>	NP_001266234.1	GenBank
	<i>Haliclona amboinensis</i>	c56049_g2_i2 mm.6942	<a href="http://compagen.org">compagen.org</a>
Ctenophora	<i>Bolinopsis mikado</i>	c56689_g1_i1	*
	<i>Mnemiopsis leidyi</i>	ADO34159.1	GenBank
	<i>Pleurobachia bachei</i>	sb 2661557	<a href="http://neurobase.rc.ufl.edu">neurobase.rc.ufl.edu</a>
Placozoa	<i>Trichoplax adherens</i>	DN14072_c0_g3_i1	*
Choanoflagellata	<i>Salpingoecca rosetta</i>	XP_004991097.1	GenBank

Asterisk, sourced from transcriptomes generated by our research group.

Ref 1: Khalturin, K., *et al.* Medusozoan genomes inform the evolution of the jellyfish body plan. *Nat Ecol Evol* 3, 811–822 (2019). doi.org/10.1038/s41559-019-0853-y

Ref 2: Alié, A. *et al.* ‘The ancestral gene repertoire of animal stem cells.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(51), pp. E7093-100 (2015). doi: 10.1073/pnas.1514789112.

\* Transcriptome sequencing was carried out in our lab.

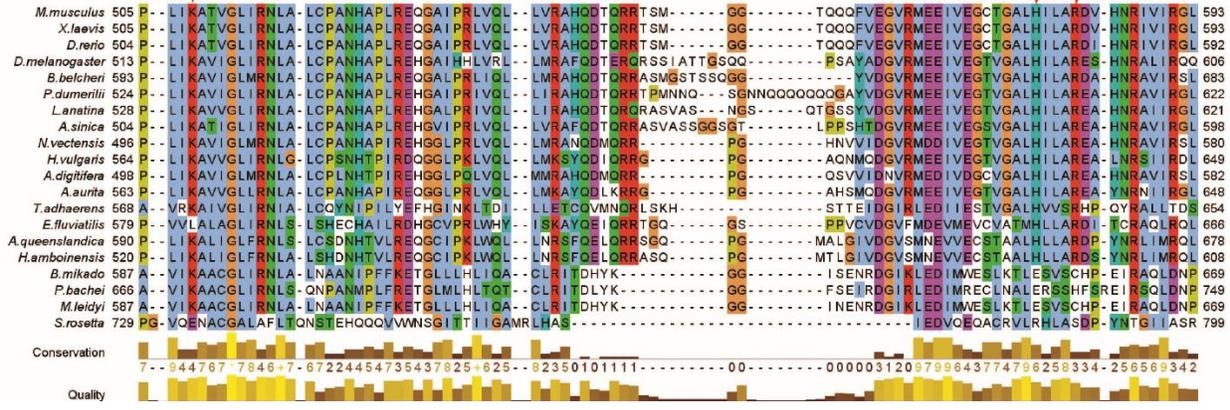
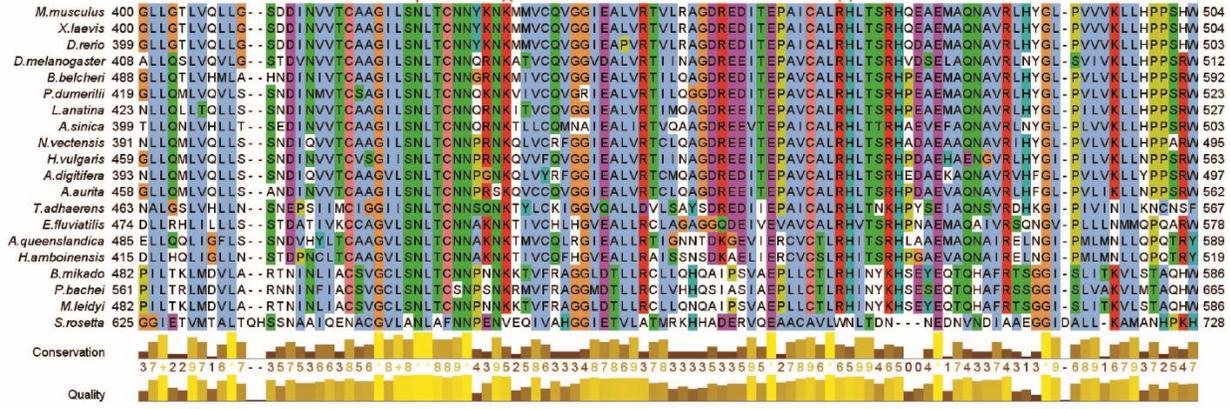
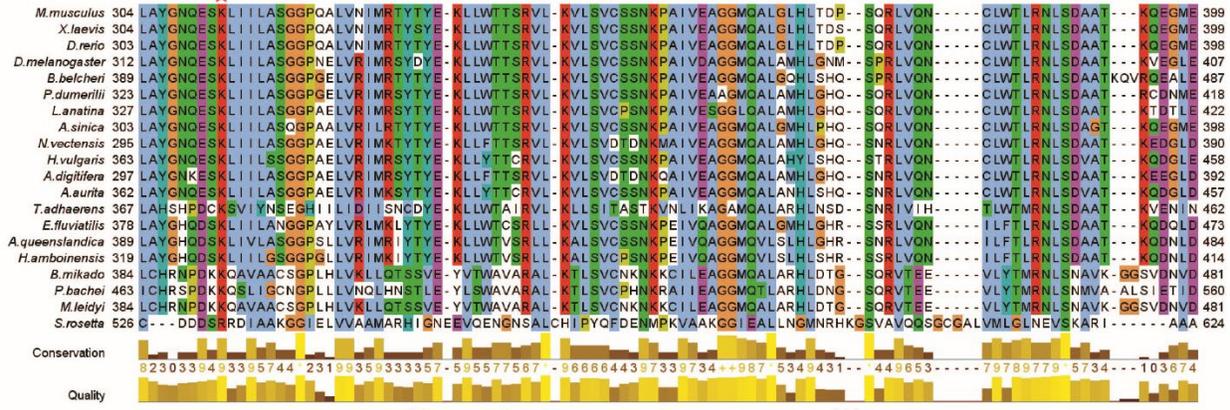
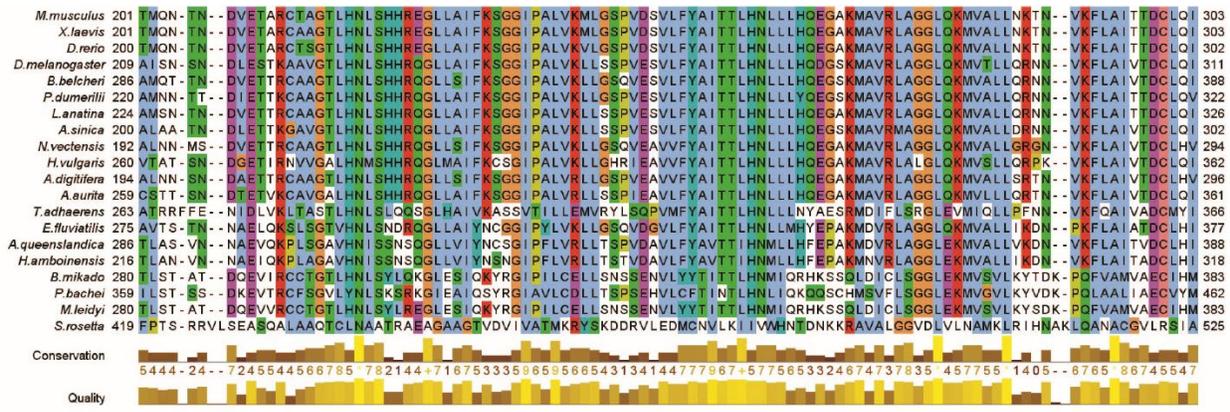
## Appendix 2

### Alignment of metazoan $\beta$ -catenins & *S. rosetta* $\beta$ -catenin-like protein.

Red and black circles represent GSK3 $\beta$  and CK1 $\alpha$  phosphorylation sites, respectively. Black bold line represent  $\alpha$ E-catenin binding motif. Black arrow represents tyrosine residue critical to binding  $\alpha$ E-catenin while green arrows represents acidic residues that promote or inhibit this interaction. Red asterisk represents lysine residues vital in the binding of both TCF and E-cadherin. Red arrows represent residues key in binding TCF. Motif A and B in the C-transactivation domain are labelled with dotted boxes.



Appendices





### Appendix 3

#### Antibodies and plasmids used

Primary antibodies	Manufacturer (Cat #)	Source	Dilution
$\alpha$ -tubulin	Sigma-Aldrich (T6074)	Mouse polyclonal	1:1000
Anti-flag M2	Sigma-Aldrich (F1804)	Mouse monoclonal	1:1000
$\alpha$ E-catenin	Santacruz (SC9988)	Mouse monoclonal	1:500
GAPDH	Santacruz (SC-47724)	Mouse monoclonal	1:1000
Secondary antibodies			
Anti-mouse	JacksonImmuno Research (115035003)	Goat polyclonal	1:10000

Plasmid	Description	Source
pCS2+ Flag-Nvec- $\beta$ -cat	Flag-tagged <i>N. vectensis</i> $\beta$ -catenin	This study
pCS2+ Flag-Xlae- $\beta$ -cat	Flag-tagged <i>X. laevis</i> $\beta$ -catenin	This study
pCS2+ Flag-Eflu- $\beta$ -cat	Flag-tagged <i>E. fluviatilis</i> $\beta$ -catenin	This study
pCS2+ Flag-Bmik- $\beta$ -cat	Flag-tagged <i>B. mikado</i> $\beta$ -catenin	This study
pCS2+ $\beta$ -gal	$\beta$ -galactosidase	Yuuri Yasuoka
*pCSf107_8x Super TOPFLASH	Beta-catenin reporter. TCF/LEF sites upstream of a luciferase reporter.	Veeman et al. (2003) (From Yuuri Yasuoka)
*pCSf107_8x Super FOPFLASH	Contains mutated TCF binding sites upstream of a luciferase reporter.	Veeman et al. (2003) (From Yuuri Yasuoka)

\*pCSf107 is vector that contains SP6/T7 terminator sequences (Mii and Taira, 2009).

#### References:

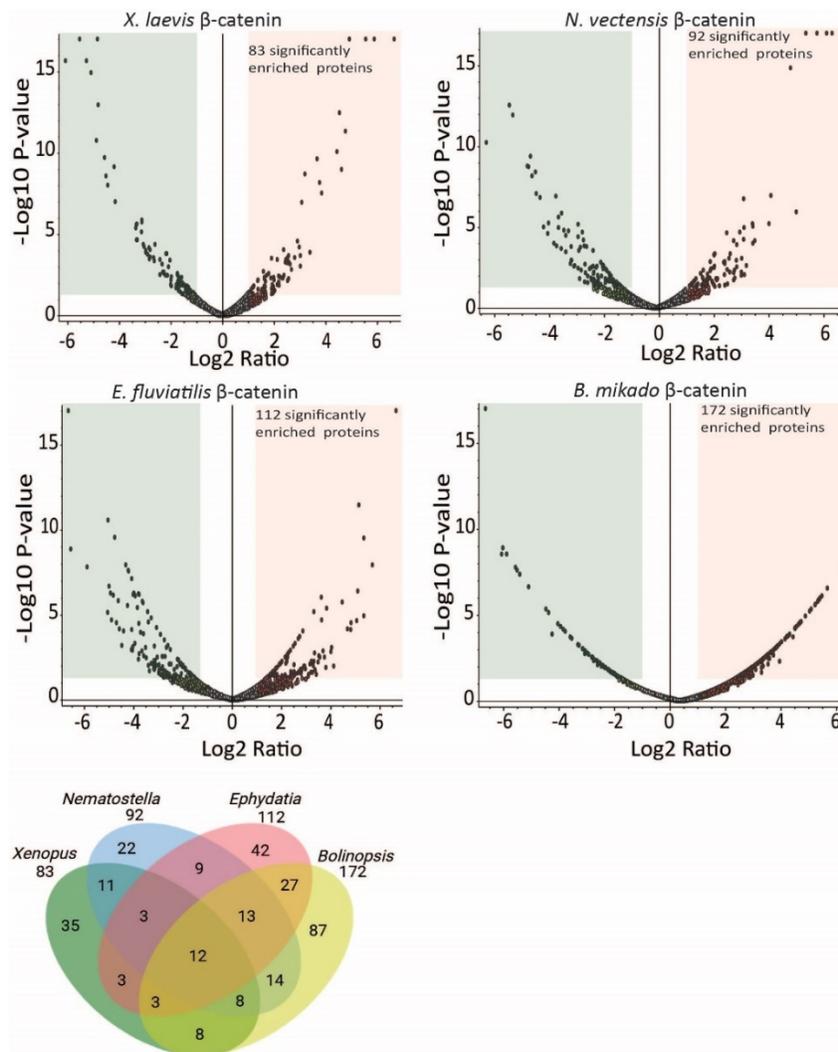
Mii Y, & Taira, M. (2009). Secreted Frizzled-related proteins enhance the diffusion of wnt ligands and expand their signalling range. *Development* 136:4083–4088.

Veeman et al. (2003). Zebrafish prickles, a modulator of noncanonical Wnt/Fz signaling, regulates gastrulation movements. *Current Biology*, 13(8), 680–685.

## Appendix 5

### Proteins identified using the in-gel (IG) method

High confidence proteins (1% FDR), with a fold abundance ratio  $\geq 2$  and a p-value  $< 0.05$ , were considered to be “true” interactions. The volcano plots below show identified proteins. Proteins that met the criteria are in the light red quadrant. Common interactions were then identified, as shown in the venn diagram and the following tables. Blast analysis was conducted against the human proteome to identify and confirm the closest homolog for each protein. Function annotation was conducted using data from Genecards ([genecards.org/](http://genecards.org/)) while localisation was carried by using the Human Protein Atlas ([proteatlas.org/](http://proteatlas.org/)). Any known interactions with  $\beta$ -catenin were carried out using Google scholar (<https://scholar.google.com/>) and BioGrid ([thebiogrid.org/](http://thebiogrid.org/)).



Appendices

**Protein interactions conserved in *X. laevis*, *N. vectensis*, *E. fluviatilis*, and *B. mikado*  $\beta$ -catenin**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
A0A1L8HIR4	Dehydrogenase/Reductase 12	DHRS12	Inhibits the proliferation and metastasis of osteosarcoma via $\beta$ -catenin pathway	Nucleoplasm	No
A0A1L8G578	Cleavage And Polyadenylation Specific Factor 3	CPSF3	Critical in RNA polyadenylation	Nucleoplasm	No
A0A1L8HG50	RNA Methyltransferase Like 1	MRM3	Catalyzes the creation of 2'-O-methylguanosine	Mitochondria	No
A0A1L8GET9, P33152	E-cadherin	CDH1	Critical in cell-cell adhesion	Plasma membrane	Several
A0A1L8HXE0	HAUS Augmin Like Complex Subunit 8	HAUS8	Essential for maintaining spindle integrity and chromosomal stability	Cytoskeleton	No
A0A1L8HPP5	Short Chain Dehydrogenase/Reductase Family 39U Member 1	SDR39U1	Metabolises steroid hormones and prostaglandins	Nucleoplasm & cytosol	No
A0A1L8ER90	D-3-phosphoglycerate dehydrogenase	PHGDH	Catalyzes the reversible oxidation of 3-phospho-D-glycerate to 3-phosphonoxypropionate, the first step of the phosphorylated L-serine biosynthesis pathway.	Nucleoplasm & cytosol	No
A9JS50	FKBP Prolyl Isomerase 10	FKBP10	Accelerates the folding of proteins during protein synthesis	Vesicles & endoplasmic reticulum	No
Q58E95	Adenylyltransferase and sulfurtransferase MOCS3	MOCS3	Essential during biosynthesis of the molybdenum cofactor.	Cytosol	No
A0A1L8G3M7	26S proteasome non-ATPase regulatory subunit 2	PSMD2	Involved in ATP dependent degradation of ubiquitinated proteins.	Nucleoplasm & cytosol	No
A0A1L8FYI7	Microtubule Crosslinking Factor 1	MTCL1	Development and maintenance of non-centrosomal microtubule bundles at the lateral membrane in polarized epithelial cells.	Cytoskeleton & plasma membrane	No

**Protein interactions with *X. laevis*  $\beta$ -catenin only**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Flag_Bcat_Xenopus laevis	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
Q6DEE7	Family With Sequence Similarity 76 Member B	FAM76B	Unknown	Nucleoplasm	No
A0A1L8F1E9	Proline Rich Coiled-Coil 2B	PRRC2B	Unknown	Cytosol	No
A2RV58	Acylglycerol kinase	AGK	Phosphorylates monoacylglycerol and diacylglycerol to form lysophosphatidic acid and phosphatidic acid, respectively	Mitochondria	No
Q3KPL2	Nucleolin	NCL	Transcriptional elongation	Nucleoplasm	
A0A1L8EMX9	Signal Recognition Particle 68	SRP68	Targets secretory proteins to the rough endoplasmic reticulum membrane.	Cytosol	No
Q3B8M1	FKBP Prolyl Isomerase 3	FKBP3	Inhibits T-cell proliferation by arresting two distinct cytoplasmic signal transmission pathways.	Nucleoplasm	No
Q7ZWY9	Lysine--tRNA ligase	KARS1	Catalyzes the specific attachment of an amino acid to its cognate tRNA in a 2-step reaction	Cytosol & nucleoplasm	No
A0A1L8GX20	IK Cytokine	IK	Involved in pre-mRNA splicing as a component of the spliceosome	Nucleoplasm	Binds to CTNNB1
Q800A4	DNA primase small subunit	PRIM1	Plays an essential role in the initiation of DNA synthesis	Nucleoplasm	No
A0A1L8GZU4	Filamin-C	FLNC	Plays a central role in muscle cells, probably by functioning as a large actin-cross-linking protein.	Cytosol, cytoskeleton & plasma membrane	No
Q7ZYR7	Protein phosphatase 1G	PPM1G		Nucleoplasm	No
A2VDC8	Eukaryotic translation initiation factor 3 subunit H	EIF3H	Required for in the initiation of protein synthesis.	Cytosol	No
A0A1L8FQW5	Glycylpeptide N-Tetradecanoyltransferase 2	NMT2	Adds a myristoyl group to the N-terminal glycine residue of certain cellular and viral proteins.	Golgi apparatus, cytosol & plasma membrane	No
A0A1L8GDF1	Cleavage stimulation factor subunit 3	CSTF3	Required for polyadenylation and 3'-end cleavage of mammalian pre-mRNAs.	Nucleoplasm	No
Q7ZYL5	Kinesin Family Member 22	KIF22	Involved in spindle formation and the movements of chromosomes during mitosis and meiosis.	Nucleoplasm	No
Q5XH65	LIM Domain And Actin Binding 1	LIMA1	Involved in actin cytoskeleton regulation and dynamics.	Cytosol	Binds to junction plakoglobin
O93355	Geminin	GMNN	Critical role in cell cycle regulation.	Cytosol and nucleoplasm	

## Appendices

A0A1L8F746	Cullin-4B	CUL4B	Mediates the ubiquitination and subsequent proteasomal degradation of target proteins.	Nucleoplasm	No
P09560	RAF proto-oncogene serine/threonine-protein kinase	RAF1	Acts as a regulatory link between the membrane-associated Ras GTPases and the MAPK/ERK cascade.	Nucleoplasm	No
Q52QH5	Actin Alpha 2, Smooth Muscle	ACTA2	Essential in cell motility.	Cytosol & cytoskeleton	No
Q6DDV8	TPX2 Microtubule Nucleation Factor	TPX2	Necessary for normal assembly of mitotic spindles.	Cytoskeleton & nucleoplasm	No
A0A1L8EKJ1	Cell division control protein 6	CDC6	Involved in the initiation of DNA replication.	Cytosol & nucleoplasm	No
A0A1L8HPJ8	Neural precursor cell expressed, developmentally down-regulated 8	NEDD8	Ubiquitin-like protein which plays an important role in cell cycle control and embryogenesis.	Cytosol & nucleoplasm	No
A0A1L8EZH4	Actin Alpha Cardiac Muscle 1	ACTC1	Essential in cell motility.	Cytosol & cytoskeleton	No
A0A1L8F2W4	Fragile X mental retardation 1	FMR1	Vital in neuronal development and synaptic plasticity through the regulation of alternative mRNA splicing, mRNA stability.	Cytosol & nucleoplasm	Ehyai et al (2018)
Q6GNQ0	Cortactin	CTTN	Contributes to the organization of the actin cytoskeleton and cell shape.	Plasma membrane & cytosol	Binds to alpha & delta catenin
Q5XHJ8	Gelsolin	GSN	Calcium-regulated, actin-modulating protein that binds to the plus (or barbed) ends of actin monomers or filaments, preventing monomer exchange.	Plasma membrane & cytoskeleton	No
A0A1L8FQC1	Microtubule-associated protein 4	MAP4	Promotes microtubule assembly.	Plasma membrane & cytosol	No
A0A1L8H338	KH-Type Splicing Regulatory Protein	KHSRP	Binds to the dendritic targeting element and may play a role in mRNA trafficking.	Nucleoplasm	No
Q7SYT0	Elongation factor 1-alpha 2	EEF1A2	Promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis.	Cytosol	No
A0A1L8G7Q1	Spectrin Repeat Containing Nuclear Envelope Protein 1	SYNE1	Forms a linking network between organelles and the actin cytoskeleton to maintain the subcellular spatial organization.	Plasma membrane, nucleoplasm & nuclear membrane	No
Q7ZWS3	Junction plakoglobin	JUP	Necessary in mediating cell adhesion.	Cytosol & plasma membrane	Several evidence
A0A1L8GSW1	Tropomodulin-3	TMOD3	Development of short actin protofilament.	Cytosol	No
Q4QR01	WD repeat-containing protein 18	WDR18	Functions as a component of the Five Friends of Methylated CHTOP (5FMC) complex; the 5FMC complex is recruited to ZNF148 by methylated	Nucleoplasm	No

Appendices

			CHTOP, leading to desumoylation of ZNF148 and subsequent transactivation of ZNF148 target genes		
--	--	--	---	--	--

**Protein interactions with *N. vectensis*  $\beta$ -catenin only.**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Flag_Bcat_Nematostella	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
A0A1L8EVQ5	G-rich sequence factor-1	GRSF1	Regulator of post-transcriptional mitochondrial gene expression, required for assembly of the mitochondrial ribosome and for recruitment of mRNA and lncRNA.	Nucleoplasm & cytosol	No
A0A1L8EPA4	Heterogeneous Nuclear Ribonucleoprotein A3	HNRNPA3	Plays a role in cytoplasmic trafficking of RNA. Binds to the cis-acting response element, A2RE.	Nucleoplasm	No
A0A1L8FD07	Protoporphyrinogen Oxidase	PPOX	Catalyzes the 6-electron oxidation of protoporphyrinogen-IX to form protoporphyrin-IX	Mitochondria & cytosol	No
A0A1L8G5U4	WAP Four-Disulfide Core Domain 5	WFDC5	A proteinase inhibitor.	Extracellular	No
P19011	Vitellogenin-B2	VITB2	Xenopus specific precursor of major egg-yolk proteins		No
A0A1L8FQP1, Q6DD69	ATP-dependent 6-phosphofructokinase, muscle type	PFKM	Catalyzes the phosphorylation of D-fructose 6-phosphate to fructose 1,6-bisphosphate by ATP, the first committing step of glycolysis.	Endoplasmic reticulum	No
Q64116	Perilipin 2	PLIN2	May be involved in development and maintenance of adipose tissue.	Lipid droplets	No
Q5EAV7	OCIA Domain Containing 2	OCIAD2	Unknown	Mitochondria	No
A0A1L8ERQ6	Keratin 19	KRT19	Involved in the organization of myofibers.	Cytoskeleton	(Hwang <i>et al.</i> , 2017)
Q6NRY2	Nucleolar Protein 6	NOL6	RNA processing	Nucleoplasm	No
A0A1L8IOJ5	Hypermethylated in cancer 2	HIC2	Transcriptional repressor.	Nucleoplasm	No
Q2VPF9	DEAD-Box Helicase 54	DDX54	Represses the transcriptional activity of nuclear receptors.	Nucleoplasm	No
Q7ZZY7	Ribonucleic Acid Export 1	RAE1	Plays a role in mitotic bipolar spindle formation	Nucleoplasm	No
Q52L44	Propionyl-CoA Carboxylase Subunit Beta	PCCB	Catalyzes the carboxylation of propionyl-CoA/propanoyl-CoA to D-methylmalonyl-CoA/(S)-methylmalonyl-CoA	Mitochondria	No

## Appendices

A0A1L8HH71	Propionyl-CoA Carboxylase Subunit Alpha	PCCA	Involved in the catabolism of odd chain fatty acids, branched-chain amino acids isoleucine, threonine, methionine, and valine	Mitochondria	No
A0A1L8EXW2	Sad1 And UNC84 Domain Containing 1	SUN1	Involved in telomere attachment to nuclear envelope in the prophase of meiosis	Nuclear membrane	No
Q6DJE0	SEC23 Homolog B, COPII Coat Complex Component	SEC23B	Promotes the formation of transport vesicles from the endoplasmic reticulum	Endoplasmic reticulum	No
A0A1L8HLC5	Factor Interacting With PAPOLA And CPSF1	FIP1L1	Contributes to poly(A) site recognition and stimulates poly(A) addition.	Nucleoplasm	No
A0A1L8HYG8	Phosphoserine aminotransferase	PSAT1	Catalyzes the reversible conversion of 3-phosphohydroxypyruvate to phosphoserine and of 3-hydroxy-2-oxo-4-phosphonooxybutanoate to phosphohydroxythreonine.	Cytosol	No
Q8AX85	KH-Type Splicing Regulatory Protein	KHSRP	Binds to the dendritic targeting element and may play a role in mRNA trafficking.	Nucleoplasm	No

### Protein interactions with *E. fluviatilis* $\beta$ -catenin only.

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Flag_Bcat_Ephydatia	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
A0A1L8HXW7	Focadhesin	FOCAD	Potential tumor suppressor in gliomas.	Mitochondria	No
Q5U4V6, A0A1L8HEI6, Q7ZWQ5	Tubulin alpha-1C chain	TUBA1C	Major constituent of microtubules.	Cytoskeleton	No
A0A1L8HRX7	Lectin, Mannose Binding 1	LMAN1	Forms a specific cargo receptor for the ER-to-Golgi transport of selected proteins.	Cytosol	No
A0A1L8HPG1	Phosphoenolpyruvate Carboxykinase 1	PCK1	Catalyzes the reversible decarboxylation and phosphorylation of oxaloacetate.	Cytosol	No
Q7ZY50	Tubulin Beta-4B Chain	TUBB4B	Major constituent of microtubules.	Cytoskeleton	No
A0A1L8H328	Coiled-coil and C2 domain-containing protein 1A	CC2D1A	Transcription factor that binds specifically to the dual repressor element and represses HTR1A gene transcription in neuronal cells.	Nucleoplasm & cytosol	No
A0A1L8F028, A0A1L8F9Y5	Glycogen phosphorylase, liver form	PYGL	Vital in carbohydrate metabolism.	Cytosol	No
F6KDD8	Mitochondrially Encoded Cytochrome C Oxidase II	COX2	Last enzyme in the mitochondrial electron transport chain which drives oxidative phosphorylation.	Mitochondria	Moutaoufik (2019)

## Appendices

A0A1L8HHY6	ATP Synthase F1 Subunit Beta	ATP5F1B	Production of ATP from ADP during the respiratory chain.	Mitochondria	No
A0A1L8FCH5	Skint1 Like	BTN12	Unknown	Unknown	Unknown
A0A1L8GA84	Lipoma Preferred Partner	LPP	May play a structural role at sites of cell adhesion in maintaining cell shape and motility.	Focal adhesion sites & cytosol	No
P30883	Tubulin beta-4B	TUBB4B	Major constituent of microtubules.	Cytoskeleton	No
B1H1T9	Kinesin Family Member 20A	KIF20A	Required for chromosome passenger complex (CPC)-mediated cytokinesis.	Nucleoplasm	<a href="#">Capalbo et al (2019)</a>
A0A1L8H3C0	Minichromosome Maintenance Complex Component 7	MCM7	Putative replicative helicase essential for DNA replication initiation & elongation in eukaryotic cells.	Nucleoplasm	No
Q5XGR6, Q6NTQ9	NADH:Ubiquinone Oxidoreductase Subunit A10	NDUFA10	Functions in the transfer of electrons from NADH to the respiratory chain.	Mitochondria	No
Q0IHH7	Acyl-CoA Dehydrogenase Long Chain	ACADL	Catalyzes the initial step of mitochondrial fatty acid beta-oxidation	Mitochondria	No
A0A1L8GA35	26S proteasome non-ATPase regulatory subunit 2	PSMD2	Involved in the ATP-dependent degradation of ubiquitinated proteins.	Nucleoplasm & cytosol	No
Q7ZXE9	Testin LIM Domain Protein	TES	Scaffold protein that may play a role in cell adhesion, cell spreading and in the reorganization of the actin cytoskeleton.	Focal adhesion sites & cytosol	No
A0A1L8EU56	Adenosylhomocysteinase	AHCY	May play a key role in the control of methylations via regulation of the intracellular concentration of adenosylhomocysteine.	Cytosol	No
Q9W603	FACT Complex Subunit SPT16	SUPT16H	Acts as a histone chaperone that both destabilizes and restores nucleosomal structure.	Nucleoplasm	No
Q6NRV3	Tubulin Alpha Like 3	TUBAL3	Major constituent of microtubules.	Cytoskeleton	No
Q7ZY98	Staphylococcal Nuclease And Tudor Domain Containing 1	SND1	Endonuclease that mediates miRNA decay of both protein-free and AGO2-loaded miRNAs.	Cytosol	<a href="#">Moutaoufik MT (2019)</a>
O12975	Aldolase, Fructose-Bisphosphate A	ALDOA	Plays a key role in glycolysis and gluconeogenesis.	Cytosol	No
A0A1L8FXX3	Thioredoxin Domain Containing 5	TXNDC5	Possesses thioredoxin activity. Has been shown to reduce insulin disulfide bonds.	Endoplasmic reticulum	No
A0A1L8H4A9	Acyl-CoA Dehydrogenase Very Long Chain	VLCAD	Catalyzes the initial step of mitochondrial fatty acid beta-oxidation.	Nucleoplasm & mitochondria	No

## Appendices

Q6TEC1	Nudix Hydrolase 16	NUDT16	RNA-binding and decapping enzyme that catalyzes the cleavage of the cap structure of snoRNAs and mRNAs in a metal-dependent manner.	Nucleoplasm	No
A0A1L8GZZ1	Transient Receptor Potential P8	TRPM8	Positively regulates the plasma membrane cation channel TRPM8 activity.	Plasma membrane	No
Q6PAB3	Malate Dehydrogenase 1	MDH1	Catalyzes the reduction of aromatic alpha-keto acids in the presence of NADH	Cytosol	No
A0A1L8HPL6	Zinc Finger Homeobox 2	ZFHX2	Transcriptional regulator that is critical for the regulation of pain perception and processing of noxious stimuli.	Nucleoplasm	No
A0A090AZM4	Minichromosome Maintenance Complex Component 6	MCM6	Putative replicative helicase necessary for DNA replication initiation & elongation in eukaryotic cells.	Nucleoplasm	No
A0A1L8GPN4	Minichromosome Maintenance Complex Component 2	MCM2	Putative replicative helicase necessary for DNA replication initiation & elongation in eukaryotic cells.	Nucleoplasm	Drissi <i>et al</i> (2015)
A0A1L8HX68	60S Ribosomal Protein L40	RPL40	Influences the ribosomal biogenesis as a ribosomal protein	Nucleoplasm & cytosol	No
Q91855	G1 To S Phase Transition 1	GSPT1	Involved in translation termination in response to the termination codons UAA, UAG and UGA	Cytosol	No
Q7ZYH7	Phosphoglycerate Kinase 1	PGK1	Catalyzes one of the two ATP producing reactions in the glycolytic pathway via the reversible conversion of 1,3-diphosphoglycerate to 3-phosphoglycerate	Cytosol	No
A0A1L8ELW9	Cytochrome C Oxidase Subunit 4I1	COX4I1	Final enzyme in the mitochondrial electron transport chain which drives oxidative phosphorylation.	Mitochondria	Moutaoufik et al (2019)
Q9I8J7	Transketolase	TKT	Catalyses the transfer of a two-carbon ketol group from a ketose donor to an aldose acceptor, via a covalent intermediate with the cofactor thiamine pyrophosphate.	Nucleoplasm	No
A0A1L8FJ72	B-TFIID TATA-Box Binding Protein Associated Factor 1	BTAF1	Regulates transcription in association with TATA binding protein (TBP).	Nucleoplasm	No

### Protein interactions with *B. mikado* $\beta$ -catenin only.

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Flag_Bcat_Bolinop	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various

## Appendices

A0A1L8H7Y9, A0A1L8I0K6	Clathrin Heavy Chain	CLTC	Contribute to stabilization of kinetochore fibers of the mitotic spindle by acting as inter-microtubule bridge	Cytosol & mitotic spindles	No
A0A1L8GM46	Nuclear Autoantigenic Sperm Protein	NASP	Required for DNA replication, normal cell cycle progression and cell proliferation.	Nucleoplasm	No
A0A1L8HKM5	Transforming Acidic Coiled-Coil Containing Protein 3	TACC3	Plays a role in the microtubule-dependent coupling of the nucleus and the centrosome.	Mitotic spindle & Cytosol	No
O93398, P09010	Lamin B1	LMNB1	Provides a framework for the nuclear envelope and may also interact with chromatin.	Nuclear membrane	<a href="#">Chen et al (2016)</a>
A0A1L8GW77	Uncharacterized protein				
Q6IRQ1	ATP Binding Cassette Subfamily D Member 3	ABCD3	Probable transporter involved in the transport of branched-chain fatty acids and C27 bile acids into the peroxisome.	Peroxisomes	No
A0A1L8HR04	Sorting Nexin 2	SNX2	Involved in several stages of intracellular trafficking. Interacts with membranes containing phosphatidylinositol 3-phosphate	Endosomes	No
A0A1L8FAS3	Exonuclease 3'-5' Domain Containing 2	EXD2	Acts as an exoribonuclease in mitochondrion, possibly by regulating ATP production and mitochondrial translation	Mitochondria	No
A0A1L8GNZ8	Phospholipase A2 Group VI	PLA2G6	Involved in phospholipid remodeling with implications in cellular membrane homeostasis, mitochondrial integrity and signal transduction.	Cytosol & mitotic spindles	No
A0A1L8EUC8	Nucleoporin 85	NUP85	Required for NPC assembly and maintenance	Nuclear membrane	No
A0A1L8I190	Kinetochore Associated 1	KNTC1	Prevents cells from prematurely exiting mitosis.	Cytosol & plasma membrane	No
Q4KLQ6	Nucleoporin 88	NUP88	Required for NPC assembly and maintenance	Nuclear membrane	No
Q6GNW7	Acetylserotonin O-Methyltransferase Like	ASMTL	Nucleoside triphosphate pyrophosphatase that hydrolyzes dTTP and UTP.	Cytosol	No
B4F6R0	Zona Pellucida Glycoprotein 2	ZP2	Facilitates sperm binding, induction of the acrosome reaction and prevents post-fertilization polyspermy	Plasma membrane	No
Q6AX37	Phosphatidylinositol-Glycan Biosynthesis Class S Protein	PIGS	Essential for transfer of GPI to proteins, particularly for formation of carbonyl intermediates.	Endoplasmic reticulum	No
Q91349	Nucleoporin 62	NUP62	Required for NPC assembly and maintenance	Nuclear membrane	No
Q91673	Zona Pellucida Glycoprotein 4	ZP4	Facilitates sperm binding, induction of the acrosome reaction and prevents post-fertilization polyspermy	Plasma membrane	No

## Appendices

A0A1L8EPP4	Ras-associated and pleckstrin homology domains-containing protein 1	RAPH1	lamellipodial dynamics regulation and negatively regulates cell adhesion.	Nucleoplasm, plasma membrane & cytosol	No
Q6GPS9	Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Subunit Alpha	HADHA	Mitochondrial trifunctional enzyme catalyzes the last three of the four reactions of the mitochondrial beta-oxidation pathway.	Mitochondria	No
Q7ZTL9	Ubiquitin Protein Ligase E3A	UBE3A	E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and transfers it to its substrates	Nucleoplasm & cytosol	No
Q694W8	Myosin X	MYO10	Regulates cell shape, cell spreading and cell adhesion.	Cytosol & plasma membrane	No
A0A1L8H8C4	Rabankyrin 5	RANK5	Proposed effector of Rab5.	Endosomes	No
Q641D5	Histone Acetyltransferase 1	HAT1	Has intrinsic substrate specificity that modifies lysine in recognition sequence GXGKXG.	Nucleoplasm	No
A0A1L8ELC5	Exportin-2	XPO2	Mediates importin-alpha re-export from the nucleus to the cytoplasm after import substrates have been released into the nucleoplasm.	Nucleoplasm & cytosol	No
A0A1L8FWJ0	DnaJ Heat Shock Protein Family (Hsp40) Member C13	DNAJC13	Involved in membrane trafficking through early endosomes.	Cytosol	No
Q0IHC3	Sideroflexin 1	SFXN1	Mediates transport of serine into mitochondria, an important step of the one-carbon metabolism pathway	Mitochondria	No

### Protein interactions conserved in *N. vectensis* and *X. laevis* $\beta$ -catenin.

Entry	Protein Description	Protein symbol	Functional Annotation	Localization	Publication
G5E9G0	Ribosomal protein L3	RPL3	Component of the large subunit of cytoplasmic ribosomes.	Nucleoplasm & cytosol	No
P70039, A0A1L8I1P0	Adenomatous polyposis coli	APC	Degradation of CTNNB1 and participates in Wnt signaling as a negative regulator.	Cytosol, nucleoplasm & plasma membrane	Several evidence
Q5MNU4	Staufen Double-Stranded RNA Binding Protein 1	STAU1	Role in specific positioning of mRNAs at given sites in the cell by cross-linking cytoskeletal and RNA	Cytosol	No

## Appendices

			components, and in stimulating their translation at the site.		
Q2TAT7	Cytochrome b-c1 complex subunit 8	UQCRCQ	Part of the mitochondrial electron transport chain which drives oxidative phosphorylation.	Mitochondria	No
A0A1L8FB29	Nucleolar and spindle-associated protein 1	NUSAP1	Plays a role in spindle microtubule organization during mitosis	Nucleoplasm	No
Q8AXM9	Catenin delta-1	CTNND1	Stabilises E-cadherin in the adherens junction	Plasma membrane	(Hwang <i>et al.</i> , 2017)
A0A1L8HDN2	Pyruvate Dehydrogenase E1 Subunit Alpha 1	PDHA1	Catalyzes the overall conversion of pyruvate to acetyl-CoA and CO <sub>2</sub> , and thereby links the glycolytic pathway to the tricarboxylic cycle.	Mitochondria	No
A0A1L8EKG6	Junction plakoglobin	JUP		Cytosol & plasma membrane	Several evidence
A0A1L8F6I3	Heat Shock Protein Family A (Hsp70) Member 5	HSPA5	Endoplasmic reticulum chaperone that plays a key role in protein folding and quality control in the endoplasmic reticulum lumen	Cytosol	No

### Protein interactions conserved in *N. vectensis* and *E. fluviatilis* $\beta$ -catenin.

Entry	Protein Description	Protein symbol	Functional Annotation	Localization	Publication
A0A1L8GCZ4	Pyruvate carboxylase	PC	Catalyzes a 2-step reaction involving the ATP-dependent carboxylation of the covalently attached biotin in the first step and the transfer of the carboxyl group to pyruvate in the second.	Mitochondria	No
A2BD87, Q9PSN9, A0A1L8GLT4	Vitellogenin-A2	VTGA2	Precursor of the major egg-yolk proteins that are sources of nutrients during early development of oviparous organisms.		No
A8E5X7	Tubulin beta 2B	TUBB2B	Plays a critical role in proper axon guidance in both central and peripheral axon tracts	Cytoskeleton	No
Q5U4L0	intelectin-2	ITLN2	May play a role in the defense system against pathogens.	Extracellular matrix	No
Q6GN63	Isocitrate dehydrogenase [NAD] subunit alpha	IDH3A	Catalyzes the decarboxylation of isocitrate into alpha-ketoglutarate.	Mitochondria & nucleus	No

## Appendices

E9LYZ5	Prominin 1	PROM1	May play a role in cell differentiation, proliferation, and apoptosis	Plasma membrane & endoplasmic reticulum	No
--------	------------	-------	---	---	----

### Protein interactions conserved in *N. vectensis*, *B. mikado* and *E. fluviatilis* $\beta$ -catenin.

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Q5XHK2	14-3-3 protein beta/alpha	YWHAB	Adapter protein implicated in the regulation of a large spectrum of both general and specialized signaling pathways.	Cytosol & nucleoplasm	No
A0A1L8HXD2	Maternal Embryonic Leucine Zipper Kinase	MELK	Serine/threonine-protein kinase involved in various processes such as cell cycle regulation, self-renewal of stem cells, apoptosis and splicing regulation.	Plasma membrane & nucleoplasm	No
A0A1L8H4P4	RNA Polymerase II Subunit A	RPB1	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.	Nucleus	No
Q3KQG0	Electron transfer flavoprotein subunit alpha	ETFA	Transfers the electrons to the mitochondrial respiratory chain via ETF-ubiquinone oxidoreductase	Mitochondria	No
Q6U7I9	Ubiquitin Specific Peptidase 5	USP5	Cleaves linear and branched multiubiquitin polymers.	Cytosol & nucleoplasm	No
Q6INR2	Pyrroline-5-Carboxylate Reductase 2	PYCR2	Housekeeping enzyme that catalyzes the last step in proline biosynthesis.	Mitochondria	No
Q6DFJ1	1,4-alpha-glucan-branching enzyme	GBE1	Required for normal glycogen accumulation	Cytosol & nucleoplasm	No
Q6GMB7	Methylcrotonoyl-CoA Carboxylase 1	MCCC1	Essential in leucine and isovaleric acid catabolism.	Mitochondria	No
A0A1L8FZP0	14-3-3 protein zeta/delta	YWHAZ	Induces ARHGEF7 activity on RAC1 as well as lamellipodia and membrane ruffle formation	Cytosol & nucleoplasm	Several
A9UM19	Glutamine amidotransferase-like class 1 domain-containing protein 3A	GATD3A			No
Q6PAF2	Calpain 1	CAPN1	Involved in cytoskeletal remodeling and signal transduction.	Cytosol	No
A0A1L8G1U5	Protein Transport Protein Sec23B	SEC23B	Stimulates the creation of transport vesicles from the endoplasmic reticulum.	Endoplasmic reticulum	No
A0A1L8FIF9	Prohibitin-2	PHB2	Mediator of transcriptional repression by nuclear hormone receptors via recruitment of histone deacetylases	Mitochondria & nucleoplasm	<a href="#">Ewing et al (2007)</a>

**Protein interactions conserved in *N. vectensis* and *B. mikado*  $\beta$ -catenin.**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Q6GMF0	Delta-1-pyrroline-5-carboxylate synthase	ALDH18A1	Converts glutamate to glutamate 5-semialdehyde, during the biosynthesis of proline, ornithine and arginine.	Mitochondria	No
A0A1L8GVZ4	Lymphoid-Restricted Membrane Protein	LRMP	Delivery of peptides to MHC1 molecules.	Nucleoplasm & cytosol	No
A0A1L8EQC6	Sad1 And UNC84 Domain Containing 1	SUN1	Involved in the connection between the nuclear lamina and the cytoskeleton	Nucleoplasm	No
A0A1L8HBE3	Nuclear pore complex protein Nup98-Nup96	NUP98	Involved in the bidirectional transport across the NPC.	Nucleoplasm	No
Q6IP85	Vacuolar protein sorting-associated protein VTA1	VTA1	Involved in the endosomal multivesicular bodies (MVB) pathway.	Cytosol & nucleoplasm	No
Q5XGL4	Proteasome subunit alpha type-6	PSMA6	Involved in the proteolytic degradation of intracellular proteins.	Cytosol & nucleoplasm	No
A0A1L8HUR0	Protein Transport Protein Sec24B	SEC24B	Promotes the creation of transport vesicles from the endoplasmic reticulum	Cytosol, nucleoplasm & endoplasmic reticulum	No
A0A1L8HJ31	Nucleoporin p58/p45	NUP58	Involved in the bidirectional transport across the NPC.	Nucleoplasm	No
A0A1L8GD17	CAD Protein	CAD	Necessary protein in the pyrimidine pathway	Cytosol & nucleoplasm	No
Q6PKY5	Metaxin-1	MTX1	Involved in transport of proteins into the mitochondrion.	Mitochondria	No
Q6NTZ7	Nuclear receptor-binding protein	NRBP1	May play a role in subcellular trafficking between the endoplasmic reticulum and Golgi apparatus through interactions with the Rho-type GTPases.	Cytosol & nucleoplasm	No
A0A1L8GKJ9	Nucleoporin 93	NUP93	Involved in the bidirectional transport across the NPC.	Nucleoplasm	No
A0A1L8F9A2	NIMA Related Kinase 9	NEK9	Regulator of mitotic progression, participating in the control of spindle dynamics and chromosome separation.	Cytosol, nucleus & mitochondria	No
A0A1L8I066	eIF-2-alpha kinase activator GCN1	GCN1	Acts as a positive activator of the EIF2AK4/GCN2 protein kinase activity in response to amino acid starvation.	Cytosol	No

Appendices

**Protein interactions conserved in *X. laevis* and *E. fluviatilis*  $\beta$ -catenin.**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Q6IND5	von Willebrand factor A domain-containing protein 5A	VWA5A	May play a role in tumorigenesis as a tumor suppressor.	Nucleoplasm	No
A1L2N5	Actin, cytoplasmic 1	ACTB	Actin is a highly conserved protein that polymerizes to produce filaments that form cross-linked networks in the cytoplasm of cells	Cytoskeleton	<a href="#">Chang et al (2012)</a>
Q7ZY48	60S ribosomal protein L13a	RPL13A	Component of the GAIT (gamma interferon-activated inhibitor of translation) complex which mediates interferon-gamma-induced transcript-selective translation inhibition in inflammation processes.	Cytoplasm & nucleoplasm	No

**Protein interactions conserved in *X. laevis*, *N. vectensis* & *B. mikado*  $\beta$ -catenin.**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
A0A1L8HQG5, Q6INZ6	Armadillo repeat protein deleted in velo-cardio-facial syndrome	ARVCF	Involved in protein-protein interactions at adherens junctions.	Plasma membrane & nucleoplasm	Various
Q6DE12	Perilipin 3	PLIN3	Transport of mannose 6-phosphate receptors from endosomes to the trans-Golgi network.	Cytosol	No
Q91682, A0A1L8GWT0	Catenin Alpha 1	CTNNA1	Essential in cell-cell adhesion	Plasma membrane & cytoskeleton	Various
A0A1L8GG92	Formin Binding Protein 1 Like	FNBP1L	Directs membrane tubulation with reorganization of the actin cytoskeleton during endocytosis.	Cytosol	No
A0A1L8GKU1	IST1 Factor Associated With ESCRT-III	IST1	Recruits VPS4A and VPS4B to the midbody of dividing cells	Cytosol	No
A0A1L8G704	Cullin 9	CUL9	Mediates ubiquitination and subsequent degradation of BIRC5 to maintain microtubule dynamics and genome integrity.	Cytosol & nucleoplasm	No

**Protein interactions conserved in *X. laevis*, *E. fluviatilis* and *B. mikado*  $\beta$ -catenin.**

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
-------	---------------------	----------------	-----------------------	--------------	-------------

## Appendices

A0A1L8I396	Myosin-IIIb	MYO3B	Probable actin-based motor with a protein kinase activity.	Cytoskeleton	No
A0A1L8F071	26S proteasome regulatory subunit 10B	PSMC6	Involved in the ATP-dependent degradation of ubiquitinated proteins.	Cytosol, nucleoplasm & plasma membrane	No
Q7ZYE8	5,6-dihydroxyindole-2-carboxylic acid oxidase	TYRP1	Catalyzes the oxidation of 5,6-dihydroxyindole-2-carboxylic acid into indole-5,6-quinone-2-carboxylic acid.	Vesicles	No

### Protein interactions conserved in *X. laevis*, *N. vectensis* and *E. fluviatilis* $\beta$ -catenin.

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
A0A1L8GLM1, A0A1L8GFH3	Phosvitin	VTG1	Xenopus specific protein		No
Q52KX4	Smoothelin	SMTN	Structural protein of the cytoskeleton.	Nucleoplasm & cytoskeleton	No

### Protein interactions conserved in *E. fluviatilis* and *B. mikado* $\beta$ -catenin.

Entry	Protein Description	Protein symbol	Functional Annotation	Localisation	Publication
Q7ZWP4, A0A1L8H339	Tubulin beta-4A	TUBB4A	Major constituent of microtubules.	Cytoskeleton	No
A0A1L8HXI3	Armadillo Repeat Containing 6	ARMC6	Unknown	Cytosol	No
A0A1L8F6S2	TRNA Methyltransferase 12 Homolog	TRMT12	Component of the wybutosine biosynthesis pathway.	Nucleoplasm & mitochondria	No
A0A1L8F8L5	Tubulin beta chain	TUBB	Major constituent of microtubules	Cytoskeleton	No
A0A1L8G0I0	5-Oxoprolinase	OPLAH	Catalyzes the cleavage of 5-oxo-L-proline to form L-glutamate coupled.	Cytosol	No
Q91728	Zona Pellucida Glycoprotein 3	ZP3	Essential for sperm binding and zona matrix formation.		No
A0A1L8GJR2	SCY1 Like Pseudokinase 1	SCYL1	Regulates COPI-mediated retrograde protein traffic at the interface between the Golgi apparatus and the ER.	Cytosol	No

## Appendices

Q52L34	Methenyltetrahydrofolate Synthetase Domain Containing	MTHFSD	Novel RNA binding protein that is a component of stress granules.	Nucleoplasm	No
Q642R6	Nucleoporin 205	NUP205	Vital in the nuclear pore complex assembly and/or maintenance.	Nucleoplasm	No
P13602	Tubulin beta 2A	TUBB2A	Major constituent of microtubules	Cytoskeleton	No
Q6NRZ4	Ceramide Transporter 1	CERT	Facilitates the intracellular trafficking of ceramides and diacylglycerol lipids.	Nucleoplasm	No
A1L3K7	Enolase 1	ENO1	Glycolytic enzyme the catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate	Plasma membrane & cytosol	No
Q4QR22	Electron Transfer Flavoprotein Subunit Alpha	ETFA	Electron transfer to the mitochondrial respiratory chain via ETF-ubiquinone oxidoreductase	Mitochondria	No
A0A1L8FMQ6	Tripartite Motif Containing 28	TRIM28	Corepressor for KRAB domain-containing zinc finger proteins	Nucleoplasm	No
A0A1L8FF52	SEC31 Homolog B, COPII Coat Complex Component	SEC31B	May function in vesicle budding and cargo export from the endoplasmic reticulum.	Nucleoplasm & cytosol	No
A0A1L8F8T2	SUMO1 Activating Enzyme Subunit 1	SAE1	Mediates ATP-dependent activation of SUMO proteins followed by formation of a thioester bond between a SUMO protein and a conserved active site cysteine residue on UBA2/SAE2.	Nucleoplasm & cytosol	No
Q6DCF9	Annexin A7	ANXA7	Calcium/phospholipid-binding protein which promotes membrane fusion and is involved in exocytosis.	Nucleoplasm & cytosol	No
A0A1L8H6E2	Autophagy Related 3	ATG3	Enzyme required for the cytoplasm to vacuole transport ,autophagy, and mitochondrial homeostasis.	Cytosol	No
Q641I4	O-Linked N-Acetylglucosamine (GlcNAc) Transferase	OGT1	Catalyzes the transfer of a single N-acetylglucosamine from UDP-GlcNAc to a serine or threonine residue in cytoplasmic and nuclear proteins.	Nucleoplasm & plasma membrane	No
A0A1L8F992	Eukaryotic Translation Initiation Factor 2B Subunit Beta	EIF2B2	Catalyzes the exchange of eukaryotic initiation factor 2-bound GDP for GTP.	Nucleoplasm & plasma membrane	No
A0A1L8GPQ3	Ubiquinol-Cytochrome C Reductase Core Protein 1	UQCRC1	Part of the mitochondrial electron transport chain which drives oxidative phosphorylation.	Mitochondria	Moutaoufik (2019)
Q7ZXB0	Survival motor neuron protein	SNM1	Plays an important role in the splicing of cellular pre-mRNAs.	Nucleoplasm & cytosol	No
Q6INS0	Glutathione S-transferase theta-1	GSTT1	Conjugation of reduced glutathione to exogenous and endogenous hydrophobic electrophiles.	Unknown	No
A0A310UH28	Lysophospholipase II	LYPLA2	Acts as a acyl-protein thioesterase hydrolyzing fatty acids from S-acylated cysteine residues.	Nucleoplasm & cytosol	No

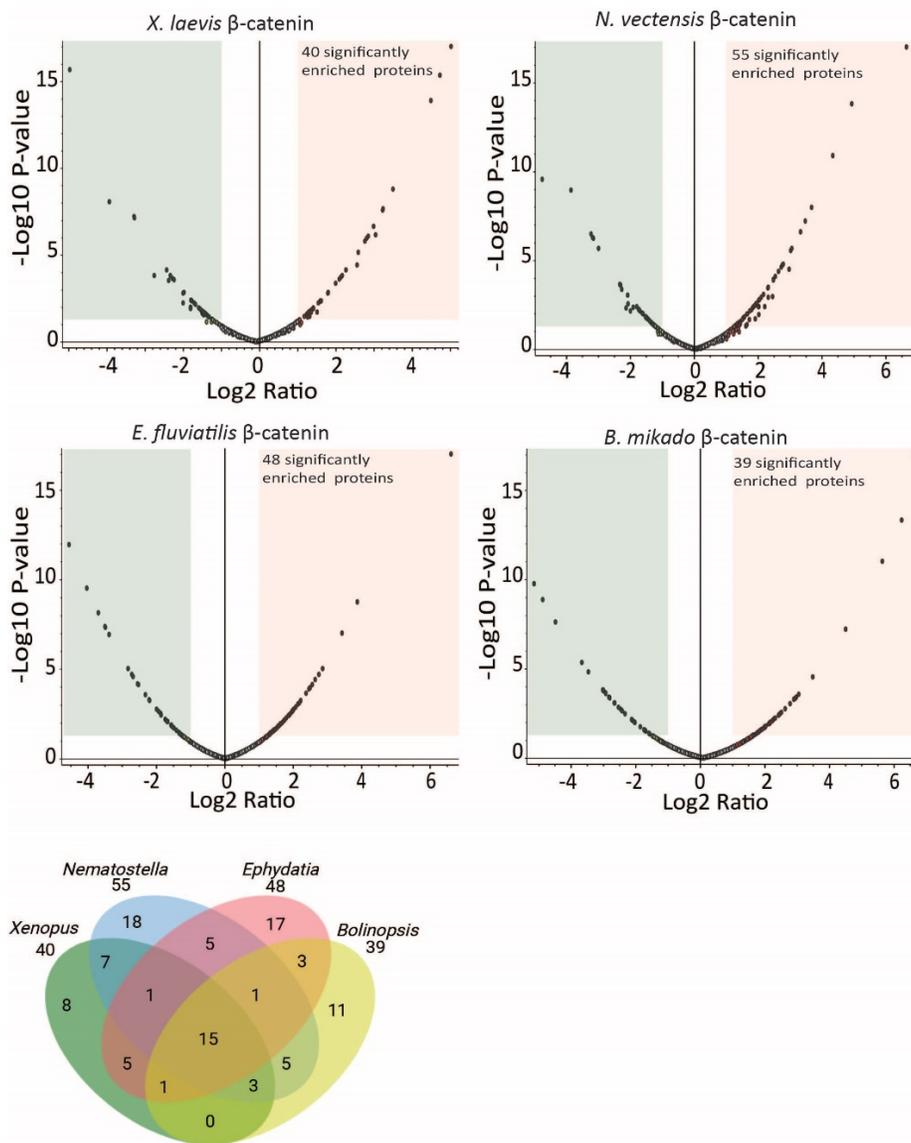
Appendices

Q4V7H4	Uncharacterised protein	Unknown	Unknown	Unknown	Unknown
Q05AT7	Zona Pellucida Glycoprotein 4	ZP4	Component of the zona pellucida.	Extracellular matrix & plasma membrane	No

Appendix 6

Proteins identified using the in-solution (IS) method

High confidence proteins (1% FDR), with a fold abundance ratio  $\geq 2$  and a p-value  $< 0.05$ , were considered to be “true” interactions. The volcano plots below show identified proteins. Proteins that met the criteria are in the light red quadrant. Common interactions were then identified, as shown in the venn diagram and the following tables. Blast analysis was conducted against the human proteome to identify and confirm the closest homolog for each protein. Function annotation was conducted using data from Genecards (<https://www.genecards.org/>) while localisation was carried by using the Human Protein Atlas (<https://www.proteinatlas.org/>). Any known interactions with  $\beta$ -catenin were carried out using Google scholar (<https://scholar.google.com/>).



Appendices

**Protein interactions conserved in *X. laevis*, *N. vectensis*, *E. fluviatilis*, and *B. mikado*  $\beta$ -catenin**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Q7ZX34	Heat Shock Protein Family A (Hsp70) Member 9	HSPA9	Important role in mitochondrial iron-sulfur cluster biogenesis.	Mitochondria	No
Q6AZV1	Heat Shock Protein 90 Alpha Family Class B Member 1	HSP90AB1	Promotes the maturation, structural maintenance and proper regulation of specific target proteins.	Cytosol	<a href="#">Tian et al (2004)</a> ,
F6KDD9	ATP synthase protein 8	ATP8	Produces ATP from ADP during the respiratory chain.	Mitochondria	No
Q6PCJ1	Dynactin Subunit 1	DCTN1	Involved in ER-to-Golgi transport, the centripetal movement of lysosomes	Microtubules	No
Q6NRV3, Q5U4V6	Tubulin Alpha 1C	TUBA1C	It binds two moles of GTP, one at an exchangeable site on the beta chain and one at a non-exchangeable site on the alpha chain.	Microtubules	No
A0A1L8FLY7	Enolase 1	ENO1	In addition to glycolysis, involved in various processes such as growth control, hypoxia tolerance and allergic responses	Plasma membrane & cytosol	No
A0A1L8G6E2	Calpain 8	CAPN8	Involved in membrane trafficking in the gastric surface mucus cells.	Nucleoplasm	No
Q6IP80	Reticulocalbin 2	RCN2	Binds calcium but function in unknown	Endoplasmic reticulum	No
A0A1L8I381	Succinyl-CoA:3-ketoacid coenzyme A transferase 1	OXCT1	Key enzyme for ketone body catabolism. Transfers the CoA moiety from succinate to acetoacetate.	Mitochondria	No
A0A1L8G7F0	Mitochondrial Ribosomal Protein S5	MRPS5	Affects mitoribosome accuracy and confers stress-related behavioural alterations.	Mitochondria	No
A0A1L8HA56	Poly(RC) Binding Protein 3	PCBP3	Single-stranded nucleic acid binding protein that binds preferentially to oligo dC.	Nucleoplasm & cytosol	No
A0A1L8HHY7	SAP Domain Containing Ribonucleoprotein	SARNP	Binds both single-stranded and double-stranded DNA with higher affinity for the single-stranded form.	Nucleoplasm	No
Q6DJ14	Mitochondrial Ribosomal Protein L41	MRPL41	Enhances p53, thereby contributing to p53-induced apoptosis.	Mitochondria	No
P52297	Importin subunit beta-1	KPNB1	Possibly essential in nuclear protein importation	Nucleoplasm	<a href="#">Lu et al (2016)</a>

Appendices

**Proteins interactions with *X. laevis*  $\beta$ -catenin only**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Flag_Bcat_Xenopus_laevis	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
A0A1L8G9H1	Synaptotagmin Binding Cytoplasmic RNA Interacting Protein	SYNCRIP	Component of the CRD-mediated complex that promotes MYC mRNA stability.	Nucleoplasm	No
Q6NRW7	Chloride Nucleotide-Sensitive Channel 1A	CLNS1A	Involved in both the assembly of spliceosomal snRNPs and the methylation of Sm proteins	Nucleoplasm & cytosol	No
A0A1L8H3Q4	Zona pellucida sperm-binding protein 3	ZP3	Component of the zona pellucida.		
A0A1L8HGS2	Mitochondrial Ribosomal Protein S9	MRPS9		Mitochondria	No
A0A1L8G769	Malate Dehydrogenase 1	MDH1	Catalyzes the reduction of aromatic alpha-keto acids in the presence of NADH	Cytosol	No
Q6Q2J3	Oviduct protein p20		Xenopus specific protein	Extracellular	No
B7ZQ53	Transcription factor 7-like 1	TCF7L1	Participates in the canonical Wnt signaling pathway.	Nucleoplasm	<a href="#">Graham et al (2000)</a>
A0A1L8F5F1	Ribosomal Protein L10	RPL10	Plays a role in the formation of actively translating ribosomes.	Cytosol & endoplasmic reticulum	No

**Protein interactions with *N. vectensis*  $\beta$ -catenin only.**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Flag_Bcat_Nematostella	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
A0A1L8F9Z9	WD Repeat And HMG-Box DNA Binding Protein 1	WDHD1	Acts as a DNA replication initiation factor that brings together the MCM2-7 helicase and the DNA polymerase.	Nucleoplasm	No
A0A1L8GWT0	Catenin alpha-1	CTNNA1	Essential in cell-cell adhesion	Plasma membrane	<a href="#">Tian et al (2004)</a> ,
A0A1L8GM46	Nuclear Autoantigenic Sperm Protein	NASP	Required for DNA replication, normal cell cycle progression and cell proliferation.	Nucleoplasm	No
A0A1L8HEI6	Tubulin Alpha 1c	TUBA1C			
A0A1L8F6I3	Heat Shock Protein Family A (Hsp70) Member 5	HSPA5	Plays a key role in protein folding and quality control in the endoplasmic reticulum lumen	Cytosol	No

Appendices

A0A1L8HGG4	Complement C1q Binding Protein	C1QBP	involved in various function including inflammation and infection processes, ribosome biogenesis, protein synthesis in mitochondria, regulation of apoptosis,	Plasma membrane, cytosol & mitochondria	
A0A1L8ERQ6	Keratin 19	KRT19	Involved in the organization of myofibers.	Intermediate filaments	No
Q6NS00	Importin 7	IPO7	Possibly essential in nuclear protein importation	Nucleoplasm	No
A0A1L8HJ81	Ferredoxin 1	FDX1	Essential for the synthesis of various steroid hormones	Mitochondria	No
A1A5J2	Coenzyme Q8	COQ8	Atypical kinase involved in the biosynthesis of coenzyme Q, also named ubiquinone, an essential lipid-soluble electron transporter for aerobic cellular respiration	Mitochondria	No
A0A1L8GD17	Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase, And Dihydroorotase	CAD	This protein is a 'fusion' protein encoding four enzymatic activities of the pyrimidine pathway	Nucleoplasm	No
A0A1L8H9Z8	Phosphofructokinase, muscle	PFKM	Catalyzes the phosphorylation of D-fructose 6-phosphate to fructose 1,6-bisphosphate by ATP, the first committing step of glycolysis.	Cytosol	No
A0A1L8I066	GCN1 Activator Of EIF2AK4	GCN1	Acts as a positive activator of the EIF2AK4/GCN2 protein kinase activity in response to amino acid starvation.	Cytosol	No
Q9PVY6	Proteasome 20S Subunit Alpha 7	PSMA7	Component of the 20S core proteasome complex involved in the proteolytic degradation of most intracellular proteins.	Nucleoplasm & cytosol	No
A0A1L8FLJ6	ATP synthase subunit g, mitochondrial	ATP5MG	Produces ATP from ADP in the presence of a proton gradient across the membrane which is generated by electron transport complexes of the respiratory chain.	Mitochondria	No
A0A1L8G704	Cullin 9	CUL9	Core component of a Cul9-RING ubiquitin-protein ligase complex, a complex that mediates ubiquitination and subsequent degradation of BIRC5 and is required to maintain microtubule dynamics and genome integrity.	Cytosol	No
Q6P699	Ribosomal Protein Lateral Stalk Subunit P1	RPLP1	Plays an important role in the elongation step of protein synthesis.	Cytosol	No

Appendices

**Protein interactions with *E. fluviatilis*  $\beta$ -catenin only.**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Flag_Bcat_Ephydatia	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
Q6PGS0	Signal Sequence Receptor Subunit 1	SSR1	Binds calcium to the ER membrane and thereby regulate the retention of ER resident proteins.	Endoplasmic reticulum	No
A0A1L8F833	Optic Atrophy 3	OPA3	May play some role in mitochondrial processes.	Nucleoplasm & cytosol	No
A0A1L8GQG8	Single Stranded DNA Binding Protein 1	SSBP1	Binds preferentially and cooperatively to pyrimidine rich single-stranded DNA	Mitochondria	No
Q8AVJ2	LSM14A MRNA Processing Body Assembly Factor	LSM14A	Formation of P-bodies, cytoplasmic structures that provide storage sites for translationally inactive mRNAs and protect them from degradation	Cytosol	No
A0A1L8GPS7	Ribophorin I	RPN1	Essential in the first step of protein N-glycosylation.	Cytosol & endoplasmic reticulum	No
Q66KT2	NADH:Ubiquinone Oxidoreductase Core Subunit S8	NDUFS8	Essential for the catalytic activity and assembly of complex I	Mitochondria	No
A0A1L8GTA1	Cytochrome C Oxidase Subunit 5A	COX5A	Component of enzyme complex vital for electron transfer & oxidative phosphorylation	Mitochondria	No
A0A1L8G711	Mitochondrial Ribosomal Protein L14	MRPL14	Forms part of the two bridges in the assembled ribosome.	Mitochondria	No
A0A1L8HIK9	Eukaryotic Translation Initiation Factor 4B	EIF4B	Required for the binding of mRNA to ribosomes.	Nucleoplasm & cytosol	<a href="#">Ehyai et al., 2018</a>
A0A1L8GU71	NADH:Ubiquinone Oxidoreductase Subunit A5	NDUFA5	Vital in electron transfer.	Mitochondria	No
A0A1L8H1I9	Nucleoporin 133	NUP133	Involved in poly(A)+ RNA transport.	Nuclear membrane	No
A0A1L8HIG7	Protein D7	D7	Involved in oocyte maturation.	Cytoplasm	No
P28049	Sjogren syndrome antigen B	SSB	Binds to the 3' poly(U) terminus of nascent RNA polymerase III transcripts, protecting them from exonuclease digestion and facilitating their folding and maturation	Nucleoplasm	No
Q7ZY03	Microfibril Associated Protein 1	MFAP1	Component of the mRNA spliceosome.	Nucleoplasm	No
Q7ZWU3	Protein Disulfide Isomerase Family A Member 3	PDIA3	Catalyzes the rearrangement of -S-S- bonds in proteins.	Endoplasmic reticulum	No

## Appendices

A0A1L8GEB3	Protein Arginine Methyltransferase 3	PRMT3	Methylates the guanidino nitrogens of arginyl residues in some proteins.	Cytosol	No
------------	--------------------------------------	-------	--	---------	----

### Protein interactions with *B. mikado* $\beta$ -catenin only.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Flag_Bcat_Bolinop	Flag tagged Beta catenin	CTNNB1	Cell adhesion and transcriptional regulations	Nucleoplasm & cell membrane	Various
A0A1L8I2H7	Complement Component 9	C9	Vital in the innate and adaptive immune response by forming pores in the plasma membrane of target cells.	Plasma membrane & extracellular	No
A0A1L8G1F3	Spectrin Beta, Erythrocytic	SPTB	Major constituent of the cytoskeletal network underlying the erythrocyte plasma membrane.	Cytosol	No
Q4V7T5	Mitochondrial Fission Regulator 1 Like	MTFR1L	Unknown	Mitochondria & cell junctions	No
A0A1L8F045	Kinesin Light Chain 1	KLC1	Kinesin is a microtubule-associated force-producing protein that may play a role in organelle transport.	Nucleoplasm, plasma membrane & cytosol	No
A0A1L8GUC2	Nucleosome Assembly Protein 1 Like 1	NAP1L1	Histone chaperone that plays a role in the nuclear import of H2A-H2B and nucleosome assembly.	Nucleoplasm	No
A0A1L8F6P6	Apoptosis Inducing Factor Mitochondria Associated 1	AIFM1	Functions both as NADH oxidoreductase and as regulator of apoptosis.	Mitochondria	No
A0A1L8HFS1	Retinoblastoma-Binding Protein 4	RBBP4	Core histone-binding subunit that targets chromatin assembly factors, chromatin remodeling factors.	Nucleoplasm	No
A0A1L8ELM2	VAMP Associated Protein A	VAPA	Binds to OSBPL3, which mediates recruitment of VAPA to plasma membrane sites	Endoplasmic reticulum	No
A0A1L8GYM8	Cullin Associated And Neddylation Dissociated 1	CAND1	Acts as a F-box protein exchange factor.	Nucleoplasm	No
Q7ZYR7	Protein Phosphatase, Mg <sup>2+</sup> /Mn <sup>2+</sup> Dependent 1G	PPM1G	Responsible for the dephosphorylation of Pre-mRNA splicing factors, which is important for the formation of functional spliceosome.	Nucleoplasm	No

## Appendices

### Protein interactions conserved in *X. laevis* and *N. vectensis* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
A0A1L8GET9, P33152	E-cadherin	CDH1	Essential in cell-cell adhesion	Plasma membrane	Several
P49739	Minichromosome Maintenance Complex Component 3	MCM3	Essential for 'once per cell cycle' DNA replication initiation and elongation in eukaryotic cells.	Nucleoplasm	<a href="#">Ehyai et al (2018)</a>
A0A1L8GNF1	RAB11 Family Interacting Protein 4	RAB11FIP4	Acts as a regulator of endocytic traffic by participating in membrane delivery.	Plasma membrane, endosome	No
A0A1L8HTV8	Pericentriolar Material 1	PCM1	Required for centrosome assembly and function.	Nucleoplasm & cytosol	No
Q6DED3	Catenin Beta Interacting Protein 1	CTNNBIP1	Prevents the interaction between $\beta$ -catenin and TCF and acts as negative regulator of the Wnt signaling pathway.	Nucleoplasm, cytosol & plasma membrane	Several

### Protein interactions conserved in *X. laevis* and *E. fluviatilis* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
O93400, P15475	Actin Beta	ACTB	Polymerizes to produce filaments that form cross-linked networks in the cytoplasm of cells	Cytoskeleton, cytosol	<a href="#">Chang et al. (2012)</a>
A0A1L8FGR3	Ribosomal Protein L18	RPL18	Component of the large ribosomal subunit.	Nucleoplasm & cytosol	No
A0A1L8H0F9	Tropomodulin 3	TMOD3	Blocks the elongation and depolymerization of the actin filaments at the pointed end.	Cytosol & actin filaments	No
A0A1L8HW57	RNA Polymerase II Subunit B	POLR2B	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.	Nucleoplasm	No

Appendices

**Protein interactions conserved in *N. vectensis* and *E. fluviatilis*  $\beta$ -catenin.**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Q505N1	Acidic Nuclear Phosphoprotein 32 Family Member E	ANP32E	Histone chaperone that specifically mediates the genome-wide removal of histone H2A.Z/H2AZ1 from the nucleosome	Nucleoplasm	No
Q6GPQ9	NADH:Ubiquinone Oxidoreductase Subunit S6	NDUFS6	Functions in the transfer of electrons from NADH to the respiratory chain.	Mitochondria	No
A0A1L8F508	Tropomyosin 3	TPM3	Binds to actin filaments in muscle and non-muscle cells.	Cytosol & cytoskeleton	No
A0A1L8HEP1	Arginyl Aminopeptidase	RNPEP	Exopeptidase which selectively removes arginine and/or lysine residues from the N-terminus of several peptide substrates	Golgi apparatus	No
Q7T105	Mesoderm Induction Early Response Protein 1	MIER1	Transcriptional repressor regulating the expression of a number of genes including SP1 target genes.	Nucleoplasm	No

**Protein interactions conserved in *N. vectensis* and *B. mikado*  $\beta$ -catenin.**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
A0A1L8EXB4	Hemoglobin Subunit Zeta	HBZ	An alpha-type chain of mammalian embryonic hemoglobin.	Cytosol	No
Q7ZY50	Tubulin Beta-4B Chain	TUBB4B	Major constituent of microtubules.	Cytoskeleton	No
Q6DD08	Reticulocalbin 3	RCN3	Probable molecular chaperone assisting protein biosynthesis and transport in the endoplasmic reticulum	Endoplasmic reticulum	No
Q4V7H4	LOC733268 protein	LOC733268	Uncharacterised protein		
A0A1L8HNA1	Ribosomal Protein S15	RPS15	Essential in peptide elongation.	Cytosol	No

**Protein interactions conserved in *E. fluviatilis* and *B. mikado*  $\beta$ -catenin.**

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Q05AX6	Keratin 19	KRT19	Involved in the organization of myofibers.	Intermediate filaments	No
A0A1L8FGS8	Pre-mRNA Processing Factor 31	PRPF31	Involved in pre-mRNA splicing as component of the spliceosome	Nucleus	Interacts with CTNBL1 <a href="#">Ganesh et al (2011)</a>

## Appendices

A0A1L8GCM7	Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Subunit Alpha	HADHA	Mitochondrial trifunctional enzyme catalyzes the last three of the four reactions of the mitochondrial beta-oxidation pathway	Mitochondria	No
------------	---	-------	---	--------------	----

### Protein interactions conserved in *X. laevis*, *N. vectensis* and *E. fluviatilis* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
A0A090AZM4	Minichromosome Maintenance Complex Component 6	MCM6	Essential for 'once per cell cycle' DNA replication initiation and elongation in eukaryotic cells.	Nucleoplasm	No

### Protein interactions conserved in *X. laevis*, *N. vectensis*, and *B. mikado* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
Q8AVE2	Heat Shock Protein Family A (Hsp70) Member 8	HSPA8	Essential in various processes, including protection of the proteome from stress, folding and transfer of newly synthesized polypeptides.	Cytosol & nucleoplasm	No
Q0IH35	Solute Carrier Family 25 Member 11	SLC25A11	Catalyzes the transport of 2-oxoglutarate across the inner mitochondrial membrane in an electroneutral exchange for malate or other dicarboxylic acids	Mitochondria	No
Q6DD17	Coiled-Coil Domain Containing 124	CCDC124	Required for proper progression of late cytokinetic stages.	Plasma membrane and cytosol	No

### Protein interactions conserved in *X. laevis*, *E. fluviatilis* and *B. mikado* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
A0A1L8EUE2	Potassium Channel Tetramerization Domain Containing 2	KCTD2	Control of gliomagenesis	Nucleoplasm	No

### Protein interactions conserved in *N. vectensis*, *E. fluviatilis* and *B. mikado* $\beta$ -catenin.

Entry	Protein Description	Gene symbol	Functional Annotation	Localization	Publication
A0A1L8FRY6	Zinc Finger Protein 830	ZNF830	May play a role in pre-mRNA splicing as component of the spliceosome	Nucleoplasm	No