INVESTIGATION OF NONCANONICAL WNT SIGNALING NETWORK IN *C. ELEGANS*

by

Feras Alazi

B.S., Chemical Engineering, Istanbul Technical University, 2007

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering
Boğaziçi University

2009

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and grateful thanks to my thesis advisors Prof. Kutlu Ö. Ülgen and Prof. Türkan Haliloğlu for their everlasting encouragement and support during my graduate studies. I also would like to thank my thesis jury; Prof. Zeynep İlsen Önsan, Assist. Prof. Elif Özkırımlı Ölmez and Assist. Prof. Nevra Özer for their critical comments on my thesis.

I gratefully acknowledge the financial support provided by TÜBİTAK-BİDEB graduate scholarship.

This thesis has been supported by BU-BAP through project 09A501P.

I am sincerely grateful to my friends in Biosystems Engineering Research Group, especially Duygu Dikicioğlu, Ayça Cankorur, Dicle Hasdemir, Güray Kuzu, Saliha Durmuş Tekir, Esra Börklü Yücel, Elif Dereli, Betül Kavun, Pelin Ümit, Yasemen Güngörmez, Ceyda Kasavi and K. Yalçın Arğa for their everlasting patience and support.

I would also like to thank my mother, Şirin Alazi, my brothers, Nizar and Aziz Alazi, my uncle Hikmet Hacımaraşlı and my grandparents Fatma and Ahmet Hacımaraşlı for their never ending love, support and encouragement through all my life.

Last but certainly not the least I would like to thank with all my heart to my love, my fiancée, Ebru Demirci for her everlasting love, continuous patience and support which makes everything bearable and beautiful.

ABSTRACT

INVESTIGATION OF NONCANONICAL WNT SIGNALING NETWORK IN *C. ELEGANS*

Wnt proteins, a well conserved family of signalling molecules, initiate signalling mechanisms that play important roles in embryonic development of animals. Wnt proteins activate three distinct signalling pathways. Wnt/β-catenin pathway is the canonical signalling pathway whereas Wnt/PCP and Wnt/Ca^{2+} are the β-catenin independent noncanonical pathways. Wnt-activated excessive β-catenin accumulation in nucleus is reported to cause tumour formation in different cells. On the other hand, one of the β-catenin independent signalling pathways; the Wnt/Ca^{2+} signalling pathway, is found to antagonize Wnt/β-catenin signalling pathway. Therefore, Wnt/Ca^{2+} cascade may have the ability to act as a tumour suppressor. *Caenorhabditis elegans* as an extremely suitable model to evaluate the functions of disease genes incorporates a well conserved set of Wnt signalling proteins. In the present computational study, Wnt signalling mechanisms that are involved in the developmental processes and tumorigenesis are investigated thoroughly in *C. elegans.* In order to better understand the molecular basis underlying the ability of Wnt proteins performing antagonistic or similar signalling activities, the protein-protein and domain-domain interaction networks of Wnt/Ca^{2+} signalling proteins were reconstructed using literature information and homology based modelling. The reconstructed network is further investigated at micro scale whether the majority of the genes in Wnt/Ca^{2+} signalling pathway are conserved during evolution and at macro scale whether the proposed Wnt/Ca^{2+} signalling network is biologically significant. The reconstructed network having a scale-free topology like most of the other biological networks successfully covers most of the protein and domain interactions of the Wnt/Ca^{2+} signalling mechanism in humans and hence can be used in studies for identifying drug targets.

ÖZET

KANONİK OLMAYAN WNT SİNYAL YOLİZLERİNİN *C. ELEGANS*'TA İNCELENMESİ

Wnt proteinleri, evrim sırasında korunmuş sinyal molleküllerı olup, hayvanların embriyonik gelişme evresinde önemli rol oynarlar. Wnt proteinleri üç farklı sinyal yolizini başlatmaktadırlar. Wnt/β-katenin sinyal yolizi kanonik sinyal yolizi iken Wnt/PCP and Wnt/Ca^{2+} sinyal yolizleri β-katenin'den bağımsız çalışan kanonik olmayan sinyal yolizleridir. Wnt aktivasyonu sonucu hücre çekirdeğinde β-katenin'in aşırı birikmesinin kansere sebep olduğu tespit edilmiştir. Öte yandan Wnt/Ca^{2+} sinyal yolizi ile Wnt/β-katenin sinyal yolizinin zıt çalıştığı tespit edilmiştir. Buna göre, Wnt/Ca^{2+} sinyal yolizinin tümör baskılayıcı işlevi olabileceği düşünülmektedir. *Caenorhabditis elegans,* hastalık genlerinin incelenmesi için çok uygun bir model organizmadır. *Caenorhabditis elegans,* Wnt proteinlerinin türdeşlerini de bulundurmaktadır. Yapılmış olan çalışmada, hem gelişim hem de hastalık süreçlerinde rol oynayan Wnt sinyal mekanizmaları *C. elegans*'ta hesapsal yöntemler kullanılarak derinlemesine incelenmiştir. Wnt proteinlerinin benzer ve zıt çalışma prensiplerinin daha iyi anlaşılabilmesi için Wnt/Ca^{2+} sinyal yolizinin protein etkileşim ile domain etkileşim ağyapıları literatür ve türler arası benzerlik bilgileri kullanılarak yeniden yapılandırılmıştır. Yeniden yapılandırılmış sinyal yolizi sonrasında mikro düzeyde ve makro düzeyde incelenmiştir. Mikro düzeyde, yeniden yapılandırılmış sinyal yolizinde bulunan proteinlerin türler arası benzerliği ve evrim içinde korunmuşluğu incelenmiştir. Makro düzeyde, yeniden yapılandırılmış sinyal yolizinin topolojik özelliklerinin literatürde mevcut olan biyolojik ağyapılarla tutarlılığı incelenmiştir. Yeniden yapılandırılmış sinyal yolizi, Wnt/Ca^{2+} sinyal yolizinde bulunan proteinlerin ve domainlerin çoğunluğunu başarıyla kapsamaktadır. Bu sayede yeniden yapılandırılmış sinyal yolizi tümör baskılayıcı ilaç araştırmaları için bir başlangıç noktası olarak kullanılabilir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

< l >	Mean path length
<k>	Mean degree
d	Network diameter
k	Degree
P(k)	Degree distribution
3D	Three dimensional
APC	Adenomatous polyposis coli
BiNGO	Biological networks gene ontology
C. elegans/CAEEL	*Caenorhabditis elegans*
Ca^{2+}	Calcium
CaCN	Calcineurin
CamKII	Ca2+-calmodulin-dependent protein kinase II
cGMP	Guanosine monophosphate
CRC	Colorectal cancer
D. melanogaster/DROME	*Drosophila melanogaster*
D. rerio/DANRE	*Danio rerio*
DAG	Diacylglycerol
DDI	Domain-domain interaction
Dsh/ DVL	Dishevelled
Fz2	Frizzled 2
Fzd	Frizzled
GO	Gene Ontology
GSK-3β	Glycogen synthase kinase 3
H. sapiens	*Homo Sapiens*
HTP	High throughput
JNK	c-Jun N terminal kinase
M. musculus	*Mus Musculus*
MCODE	Molecular complex detection

MSA	Multiple sequence alignment
NF-AT	Nuclear factor of activated T-cells
NLK	Nemo-like kinase
PCP	Planar cell polarity
PKC	Protein kinase C
PKG	Protein kinase G
PPI	Protein-protein interaction
PSI-BLAST	Position-Specific Iterated BLAST
S. cerevisiae	*Saccharomyces cerevisiae*
SPA	Selective Permissibility Algorithm
TAK1	TGF-β activated kinase 1
X. leavis/XENLA	*Xenopus leavis*

1. INTRODUCTION

Wnt proteins a well conserved family of signalling molecules, initiates signalling mechanism that plays important roles in embryonic development of animals. Furthermore, Wnt signalling pathway is known to branch between β-catenin dependent and β-catenin independent pathways.

The soil nematode *Caenorhabditis elegans* was the first multi-cellular organism of which the genome has been sequenced completely. Since December 1998 the *C. elegans* genome (100 Mega bases) is fully sequenced (The *C.elegans* Sequencing Consortium, 1998). Several studies in recent years have established *C. elegans* as a superb model to evaluate the function of disease genes (Yanagawa *et al.,* 1995; Han, 1997; Gleason *et al*., 2002; Schulz and Yutzey, 2004; Platzer and Meinzer, 2004; Korswagen, 2007). They suggest that indeed the large degree of sequence similarity found in genes in *C. elegans* and human genomes results in a significant functional similarity of the encoded proteins. Most importantly, the major signalling pathways of *C. elegans* and vertebrates, including Ras, Notch, TGFβ, Wnt/wingless and insulin signalling are remarkably conserved (Braungart *et al.,* 2004; Meissner *et al.,* 2004; Hertweck and Baumeister, 2005; Dirnberger *et al.,* 2008).

About 50-65 % of the currently known human (disease) genes have a homologue in *C. elegans*. Because its receptor pharmacology is remarkably similar to that of humans, drugs and side effects can be tested in large scale using *C. elegans* and its mutants as an *in vivo* model (whole animal testing). The mechanisms identified using *C. elegans* models have been shown to be remarkably similar in the mammalian organism, and have helped a lot to understand the biological function of the human gene that is dysfunctional in disease (Braungart *et al.,* 2004; Meissner *et al.,* 2004; Hertweck and Baumeister, 2005; Dirnberger *et al.,* 2008). Drug development using 'humanized worms' (e.g. *C. elegans* knockout strains expressing the homologous human disease mutant) in large scale is therefore feasible. And, perhaps the first effective medication against Alzheimer's or Parkinson's diseases or Duchenne Muscular Dystrophy will, one day, be the result of a high-throughput pharmacological screening procedure using *C. elegans* as a whole-animal model.

Recent works have demonstrated that Wnt-activated excessive β-catenin accumulation in nucleus cause tumour formation. On the other hand, one of the β-catenin independent signalling pathways; the Wnt/Ca^{2+} signalling pathway, is proposed to antagonize Wnt/β-catenin signalling pathway. Therefore, Wnt/Ca^{2+} cascade may have the ability to act as a tumour suppressor. However, still much more work has to be done on this subject as the mechanism of this antagonism is not well known. (Wang and Malbon, 2003; Giles *et al.*, 2003; Veeman *et al.,* 2003; Kohn and Moon, 2005; Slusarski and Pelegri, 2007; Maiese *et al.*, 2008). Although *C. elegans'* larva and embryo are simple models for these investigations and *C. elegans* is a powerful model organism for modelling Wnt signalling pathway (Hardin and King, 2008), the noncanonical Wnt signalling pathways particularly Wnt/Ca^{2+} pathway is less studied in *C. elegans* (Eisenmann, 2005). Especially, the most proximal steps of the Wnt signalling cascade beyond Wnt receptors are less well understood. Therefore, the aim of the present work is to reconstruct as well as to analyze the noncanonical Wnt signalling pathways in *C. elegans* and to better understand the molecular basis underlying the ability of Wnt proteins to perform antagonistic or similar signalling activities. That eventually would lead us to new ideas about how to suppress cancer cells in human metabolism.

This thesis comprises of 5 chapters. Chapter 2 is a summary of the available literature on Wnt signalling. Materials and methods utilized in the presented study are given in Chapter 3. Chapter 4 includes results of the study. Discussion of the results is also included in Chapter 4. Lastly conclusions and recommendations for future studies are given in Chapter 5. This section is followed by appendices composed of supplementary information about the chapters.

2. BACKGROUND ASPECTS

A few evolutionarily conserved signalling pathways control animal development**.** Family of Wnt proteins constitutes one of these conserved signal transduction pathways. Wnt proteins have been found in many organisms diversifying from humans to nematodes (Cadigan and Nusse, 1997, Wodarz and Nusse, 1998). Recently Wnt proteins have been also found in primitive animals such as cnidarians (Lee *et al.*, 2006). Wnt proteins are glycoproteins containing 350-400 amino acids with conserved cysteine residues (Weeraratna *et al.,* 2002; Maiese *et al.*, 2008). At least 19 Wnt members are known in human, 7 in fly and 5 in *C. elegans* (Giles *et al.*, 2003).

Wnt signalling has an important role in animal development. Wnt signalling is especially active during embryogenesis and controls multiple aspects such as cell fate specification, cell differentiation, cell proliferation, cell migration, and cell polarity (Cadigan and Nusse, 1997; Miller *et al*., 1999; Kuhl *et al*., 2000; Eisenmann, 2005; Maiese *et al.*, 2008). On the other hand, defects in the components of the Wnt signalling mechanism cause tumour formation in different cell types (Polakis, 2000; Peifer and Polakis, 2000; Maiese *et al.*, 2008). For instance, greater than 90% of all colorectal cancers (CRC) have been reported to have an activating mutation of the Wnt signalling pathway (Giles *et al.*, 2003). Although, initiated by Wnt proteins, mostly mutations in the proteins other than Wnt proteins in the signalling cascade cause tumour formation. The common result of all the mutations is the accumulation of β-catenin in the nucleus which is the results in tumour formation (Polakis, 2000; Giles *et al.*, 2003).

Work on vertebrates such as mouse (*Mus musculus),* frog (*Xenopus leavis*), zebrafish (*Danio rerio*), and invertebrates such as fruit fly (*Drosophila melanogaster*) and nematode (*Caenorhabditis elegans*) show that different Wnt molecules initiate different signalling mechanisms. These are; a canonical pathway that utilizes β-catenin and noncanonical pathways that function without utilizing β-catenin (Kuhl *et al*., 2000; Herman and Wu, 2004; Eisenmann, 2005; Bejsovec, 2005). The canonical pathway is well established and functions similarly in many organisms differing from human to nematode. On the other

hand, noncanonical Wnt signalling pathways show divergence between vertebrates and invertebrates (Eisenmann, 2005; Pandur, 2005; Korswagen, 2007).

A Wnt signalling pathway that increases intracellular Ca^{2+} levels has been proposed as Wnt/Ca^{2+} pathway in vertebrates (Kuhl *et al*., 2000). Different branches of this pathway utilizing different downstream components have been reported (Kuhl *et al*., 2001; Ishitani et al., 2003; Pandur 2005). There is also a Wnt/JNK (Wnt/c-Jun N terminal kinase) pathway in vertebrates that controls cell polarity (Veeman *et al.*, 2003; Pandur 2005). Recently a new Wnt signalling pathway called Wnt-5A/Ror2 signalling has also been proposed in *X. leavis* as a potential noncanonical Wnt signalling pathway (Schambony and Wedlich, 2007).

On the other hand, invertebrate noncanonical Wnt signalling pathways are less complicated but have similar functions with those in vertebrates. In flies, there is only Wnt/PCP pathway, which controls planar cell polarity, proposed as a noncanonical Wnt signalling pathway so far (Veeman *et al.*, 2003, Eisenmann, 2005).

Simplistic overview of known canonical and noncanonical Wnt signalling pathways in flies and vertebrates is illustrated in Figure 2.1. Wnt/β-catenin and Wnt/JNK (Wnt/PCP for fly) signalling pathways refer to both fly and vertebrate pathways.

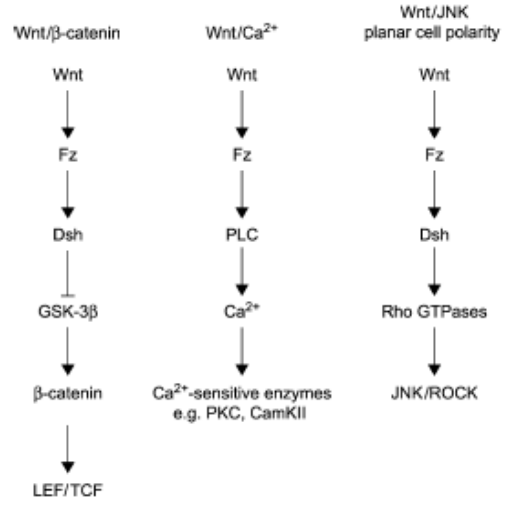

Figure 2.1. Simplistic overview of the canonical and noncanonical Wnt signalling pathways in flies and vertebrates (Pandur, 2005)

However, the divergence in noncanonical Wnt signalling increases even more in *C. elegans*. Processes such as P2/EMS signalling, T cell polarity, Z1/Z4 polarity and B cell polarity utilize noncanonical Wnt signalling. Interestingly there are three β-catenin homologues in *C. elegans* and the so called noncanonical Wnt pathways in *C. elegans* utilize one of these homologues while the other functions in the canonical pathway (Herman, 2002; Korswagen, 2002; Korswagen *et al.,* 2000; Herman and Wu, 2004; Eisenmann, 2005; Korswagen, 2007). Therefore, one of the differences between canonical and noncanonical Wnt pathways in *C. elegans* is the utilization of different β-catenin homologues. This is much more different than the β-catenin independent noncanonical pathways in fly and vertebrate. Consequently, the similarity of the noncanonical Wnt signalling pathways in *C. elegans* with those in vertebrates and fly is so far unclear (Eisenmann, 2005; Korswagen, 2007). However, a recent work suggests a novel β-catenin independent Wnt signalling pathway in *C. elegans* (Hingwing *et al*., 2009).

1.1. Wnt Signalling in Vertebrates and Flies (Experimental studies)

A body of work in vertebrates and flies has shown Wnt signals through a well established canonical or Wnt/β-catenin pathway and noncanonical pathway or pathways in a β-catenin independent manner. The number of noncanonical pathways is unclear (Veeman *et al.*, 2003; Eisenmann, 2005; Pandur, 2005).

2.1.1. Canonical Wnt Signalling

Briefly, Wnt/β-catenin signalling pathway starts when a Wnt ligand (e.g. Wnt1, Wnt3a and Wnt8) binds to Frizzled receptor at the cell surface (Figure 2.2). In addition, the co-receptor LRP-5/6 is required to activate this pathway. Wnt ligand-receptor complex leads to phosphorylation of Dishevelled (Dsh). Further, Dishevelled inhibits the formation of a complex of proteins including glycogen synthase kinase 3 (GSK-3β), Axin and the tumour suppressor protein adenomatous polyposis coli (APC) by binding to Axin. In the absence of the Wnt signal this complex destroys free β-catenin. Inhibition of this protein complex results in accumulation of β-catenin. Free β-catenin then activates Tcf/Lef transcription factors which lead to transcription and expression of different Wnt-responsive

genes. Some of these responsive genes are; c-Myc, cyclin D1, and Axin 2 (Miller *et al.*, 1999; Polakis, 2000; Herman, 2002; Bowerman, 2005; Maiese et al., 2008).

2.1.2. Noncanonical Wnt Signalling

Noncanonical Wnt pathways are less well understood but are under intense investigation. Wnt/PCP pathway which controls planar cell polarity is the noncanonical pathway well established in flies (Veeman *et al.*, 2003). Further, known pathways in vertebrates are Wnt/Ca^{2+} and Wnt/JNK pathways (Miller *et al.*, 1999; Kuhl *et al.*, 2000; Veeman *et al.*, 2003). Wnt/JNK pathway is thought to be analogous to Wnt/PCP pathway in *Drosophila*. (Veeman *et al.*, 2003; Pandur, 2005). On the other hand, to date Wnt/Ca^{2+} pathway has been found to function only in vertebrates (Miller *et al.*, 1999; Kuhl *et al.*, 2000; Kohn and Moon, 2005). Furthermore, it is unclear whether Wnt/JNK and Wnt/Ca^{2+} pathways are independent pathways functioning in vertebrates as some of the proteins function in both pathways (Kuhl *et al.*, 2001; Veeman *et al.*, 2003; Pandur, 2005).

Briefly, in Wnt/PCP pathway; Wnt proteins (Wnt5a and Wnt11) bind to Frizzled receptors which activate Dishevelled. Further, Dsh activates the small guanosine triphosphatase, Rho and Jun-N-terminal kinase (JNK). As a result, this signalling pathway helps in regulating cytoskeletal organization and coordinated polarization of cells (Huelsken and Birchmeier, 2001; Maiese *et al.*, 2008). This pathway regulates polarity of hairs, bristles and development of eye and wing in flies (Veeman *et al.*, 2003). Furthermore, regulation of gastrulation movements and convergent extension in the vertebrate embryo as a result of the polarization of migrating cells are examples Wnt/PCP pathway. These also show the analogy between fly Wnt/PCP and Wnt/JNK pathways (Veeman *et al.*, 2003; Pandur, 2005; Slusarski and Pelegri, 2007).

Wnt/Ca^{2+} pathway can be activated by Wnt5a. Binding of Wnt5a to Fz2 increases intracellular Ca^{2+} levels. Intracellular calcium levels are increased by Wnt5a-Fz2 in two different ways (Figure 2.3). First, Wnt/Fzd with G proteins activates PLC through Dishevelled. PLC activation generates diacylglycerol (DAG) and inositol-1,4,5-trisphosphate (IP_3) which eventually, increases Ca^{+2} concentration in the cell. Second, Wnt/Fzd activates cyclic guanosine monophosphate (cGMP) specific phosphodiesterase

PDE6 which decreases cellular cGMP and inactivates protein kinase G (PKG). This also leads to an increase in the cellular concentration of Ca^{2+}. Ca^{2+} increase triggers activation of calcium sensitive proteins; protein kinase C (PKC), Ca^{2+}-calmodulin-dependent protein kinase II (CamKII), and Calcineurin (CNA or CaCN) (Kuhl *et al*., 2000; Veeman *et al.,* 2003; Kohn and Moon, 2005; Pandur, 2005; Slusarski and Pelegri, 2007; Maiese *et al.*, 2008). After this point the pathway branches into three pathways.

CaCN activated by Wnt signalling, dephosphorylates NF-AT (nuclear factor of activated T-cells). NF-AT then translocates to nucleus to regulate gene expression. Studies in *Xenopus* show that this signalling is active in the development of ventral cell fate. This ventralizing activity inhibits Wnt/β-catenin pathway (Saneyoshi *et al.,* 2002; Veeman *et al.*, 2003; Pandur, 2005). No NF-AT homolog has been reported in invertebrates. The increased complexity of vertebrate development may have inserted this protein through evolution (Crabtree and Olson, 2002; Schulz and Yutzey, 2004).

CamKII further activates TGF-β activated kinase 1 (TAK1). TAK1 then stimulates nemo-like kinase (NLK). Afterwards NLK phosphorylates TCF. Consequently phosphorylation of TCF inhibits TCF/β-catenin complex (Kuhl *et al*., 2000). More studies have also showed that activation of NLK by CamKII requires the activation of TAK1 and CamKII by Wnt5a through increased calcium levels (Kuhl *et al*., 2001; Ishitani et al., 2003; Yasuda *et al.,* 2003).

The increase of Ca^{2+} levels also activates PKC (Sheldahl *et al*., 1999; Kuhl *et al*., 2000). PKC subsequently activates proteins such as Rho, JNK, cdc42 which are part of Wnt/JNK (or PCP) pathway. This common upstream regulation of small GTPases indicates overlapping between two noncanonical pathways in vertebrates (Veeman *et al.*, 2003; Pandur 2005).

As the accumulation of β-catenin is the common theme in tumour formation, it can be said that misexpression of Wnt/β-catenin signalling pathway would lead to cancer (Polakis, 2000; Giles *et al.*, 2003). On the other hand, Wnt/Ca^{2+} has been shown to inhibit Wnt/β-catenin pathway at different stages. Therefore, this antagonism might be used as a tumour suppressor in vertebrate cells (Kuhl *et al*., 2000b; Kuhl *et al*., 2001; Giles *et al.*,

2003; Veeman *et al.*, 2003). Recent work also supports this antagonism. It has been shown that Wnt5a can directly inhibit Wnt/β-catenin pathway (Mikels and Nusse, 2006).

An overview of the Wnt signalling pathways in vertebrates is shown in Figure 2.2 and Figure 2.3 in more detail. Left side of the figure represents canonical Wnt signalling pathway while the middle part represents Wnt/Ca^{2+} and the right part represents Wnt/PCP in these figures. Overlapping parts of the two noncanonical pathways and antagonism between canonical and noncanonical pathways can also be observed in these figures.

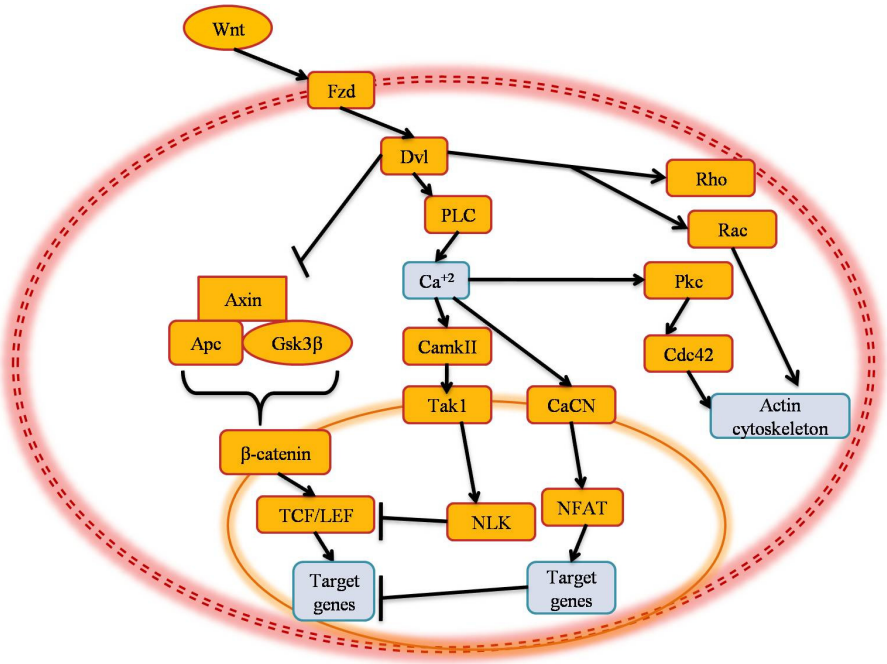

Figure 2.2. Overview of Wnt signalling pathways

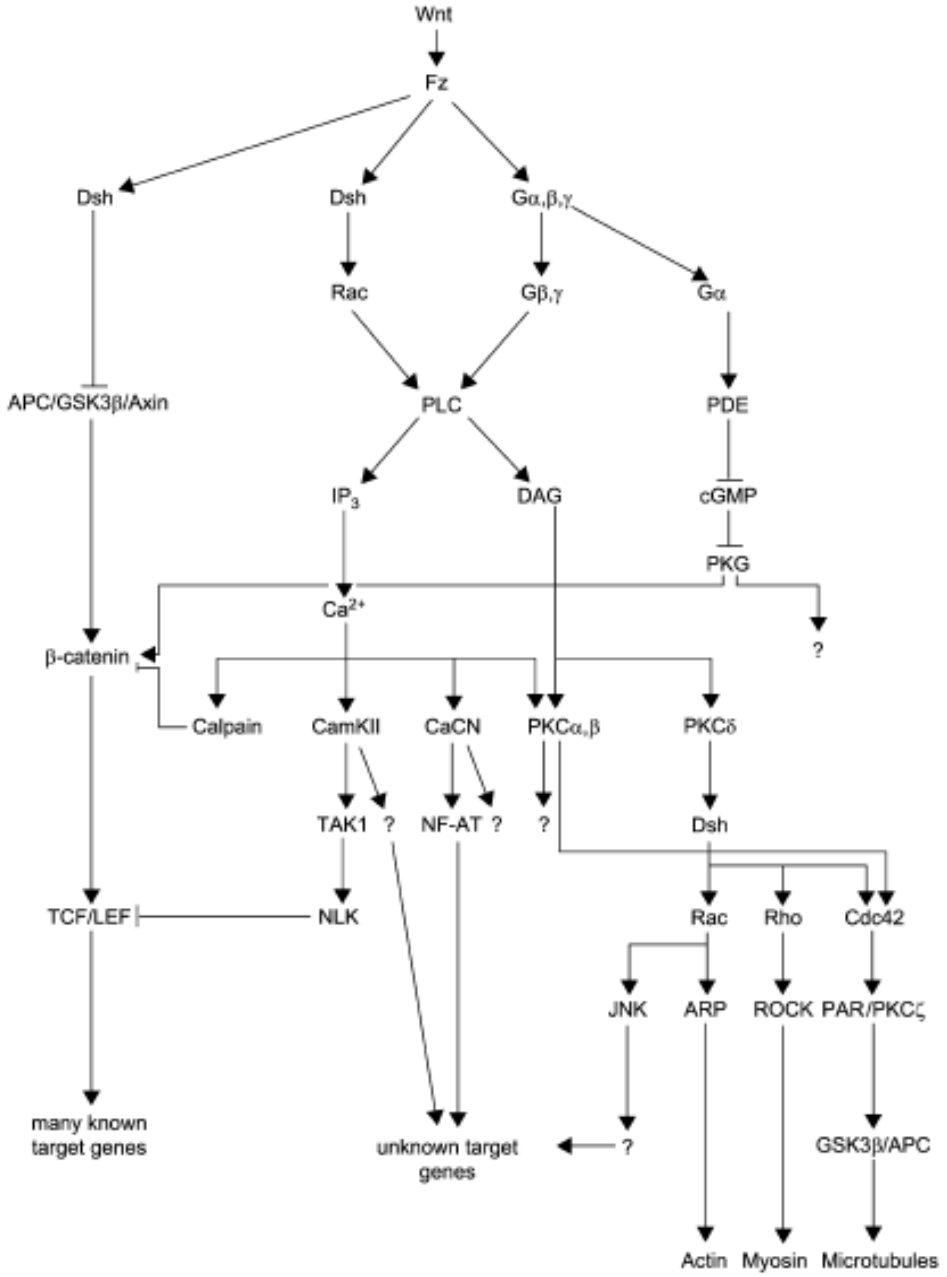

Figure 2.3. Overview of Wnt signalling pathways in vertebrates (Pandur, 2005)

2.2. Wnt Signalling Pathways in *C. Elegans*

Wnt signalling also plays an important role in *C. elegans* development. *C. elegans* has been used in many experimental researches on Wnt signalling as model organism. Advantages of *C. elegans* as model organism are; its transparent eggs which allows easy observation of embryonic development, its invariant cell lineage which allows easy study

on whole embryonic cell lineage, easiness in genetic modification, and finally its limited number of cells which is 558 at hatching and 959 in the adult hermaphrodite (Yanagawa *et al.,* 1995; Han, 1997; Gleason *et al*., 2002; Schulz and Yutzey, 2004; Platzer and Meinzer, 2004; Korswagen, 2007).

Considerable body of work suggests that important components of the canonical Wnt signalling pathway are present in *C. elegans* and the canonical pathway or pathways present are similar to those in flies and vertebrates. There is also considerable evidence suggesting the presence of noncanonical Wnt signalling pathways in *C. elegans* but with notable differences (Korswagen, 2002; Eisenmann, 2005). The similarity of noncanonical Wnt pathways in *C. elegans* with fly Wnt/PCP pathway and vertebrate Wnt/Ca^{2+} pathway also remains unclear (Herman and Wu, 2004; Eisenmann, 2005; Korswagen, 2007). However, MAPK part of noncanonical vertebrate and *C. elegans* pathways have similarity (Ishitani *et al.,*1999) as illustrated in Figure 2.3.

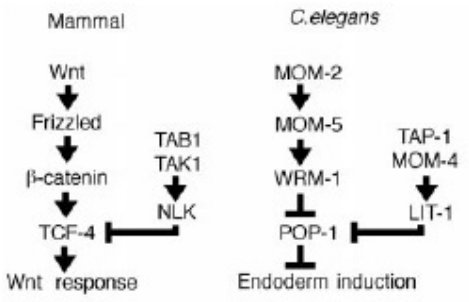

Figure 2.4. Comparison Wnt signalling between mammals and *C. elegans* (Ishitani *et al.,* 1999)

C. elegans has multiple Wnt ligands; lin-44, egl-20, mom-2, cwn-1 and cwn-2, multiple Frizzled family Wnt receptors; lin-17, mom-5, mig-1 and cfz-2, multiple Dishevelled proteins; mig-5, dsh-1, dsh-2, a GSK3β homolog; gsk-3, an Axin homolog; pry-1, an APC homolog; apr-1, a single TCF/LEF gene, pop-1, a TAK1 homolog; mom-4 and a NLK homolog; lit-1 like other species such as fly and vertebrates (Eisenmann, 2005). The main difference is the presence of three β-catenin homologues which is unique to *C. elegans.* These three β-catenin homologues are; WRM-1, BAR-1 and HMP-2. Interestingly, these β-catenin homologues show overall sequence identity less than 30%

(17%, 23%, 27% respectively) to β-catenins in fly and vertebrate while β-catenin homolog found in a simpler animal, *Hydra magnipapillata* member of cnidarians shows overall sequence identity over 60% to fly and vertebrate homologues (Natarajan *et al.,* 2001; Korswagen, 2002). Furthermore, canonical and noncanonical features of the Wnt signalling pathway are separated in terms of these three β-catenin homologues. Pathways utilizing BAR-1 are considered as canonical Wnt pathways, pathways utilizing WRM-1 are considered as noncanonical pathways and HMP-2 is thought to be utilized in cell adhesion only (Korswagen *et al*., 2000; Natarajan *et al*., 2001; Herman, 2002; Herman and Wu, 2004; Nelson and Nusse, 2004).

Canonical Wnt signalling pathway in *C. elegans* similar to those in flies and vertebrates is observed in two post-embryonic processes which are QL progeny migration and VPC fate specification (Korswagen *et al.,* 2002; Eisenmann, 2005; Wu and Herman, 2006). Components acting in QL progeny migration are EGL-20/Wnt, LIN-17/Fzd, MIG-1/Fzd, MIG-5/Dsh, BAR-1/b-cat, and POP-1/TCF (Eisenmann, 2005; Korswagen, 2007). Furthermore, known actors in VPC fate specification is *bar-1*/b-*cat*, *pry-1/axin*, *apr-1/APC* and *pop-1/TCF*. Wnt ligand and its receptor are unclear but there is high possibility of genetic redundancy (Eisenmann, 2005; Korswagen, 2007). In addition, canonical Wnt signalling is assumed to be active in P12 fate specification, postdereid formation from V5 and ray formation from V5 in the male tail (Eisenmann, 2005).

Wnt signalling pathways utilizing WRM-1 in *C. elegans* are considered to act in a noncanonical manner. Noncanonical Wnt signalling pathways (Figure 2.5) are known to act in P2/EMS signalling, T cell polarity, Z1/Z4 polarity (Herman, 2002; Korswagen, 2002). A novel noncanonical pathway that controls B cell polarity has also been proposed (Wu and Herman, 2006). Recently a noncanonical Wnt signalling pathway which is independent from β-catenin has also been proposed (Hingwing *et al*., 2009).

P2/EMS signalling is the best studied noncanonical Wnt signalling in *C. elegans*. Wnt signalling controls signalling from P2 to EMS cell that occurs at four cell-stage of embryonic development. EMS polarization and asymmetric cell division to MS and E daughter cell occurs in a Wnt mediated manner. This pathway utilizes mom-1/porcupine, mom-2/Wnt, mom-3/mig-14 (uncloned), mom-5/Fzd, dsh-2/Dsh, mig-5/Dsh, kin-19/CKI,

gsk-3/GSK3β, apr-1/APC, wrm-1/β-cat and pop-1/TCF and kinases MOM-4/Tak1 and LIT-1/Nlk and Pop-1/TCF which is downregulated. This pathway regulates spindle orientation and endoderm induction. (Korswagen, 2002; Eisenmann, 2005; Korswagen, 2007).

Secondly, T cell polarity is controlled by Wnt signals resulting in asymmetric division of TL and TR cells into anterior T.a cells and posterior T.p cells. T.a cells generate hypodermal cell fates while T.p cells generate neural cell fates (Eisenmann, 2005; Korswagen, 2007). The known Wnt pathway components are; lin-44/Wnt, lin-17/Fzd, lit-1/Nlk, wrm-1/β-cat and pop-1/TCF (reviewed in Eisenmann, 2005; Korswagen, 2007). In addition, work on egl-27/NURD suggests that egl-27 also plays role in Wnt-mediated T cell polarity (Herman *et al.,* 1999).

Thirdly, noncanonical Wnt pathway functions in establishing Z1/Z4 cells polarity. Somatic gonad develops from two precursor cells, Z1 and Z4. Asymmetric division of these cells establishes the axis of the somatic gonad. Wnt signalling pathway components involved are lin-17/Fzd, wrm-1/ β-cat, lit-1/Nlk and pop-1/TCF (Eisenmann, 2005; Korswagen, 2007). An overview of the important components of these pathways discussed so far is illustrated in Figure 2.4. Both EMS polarity and Spindle orientation parts refer to P2/EMS signalling.

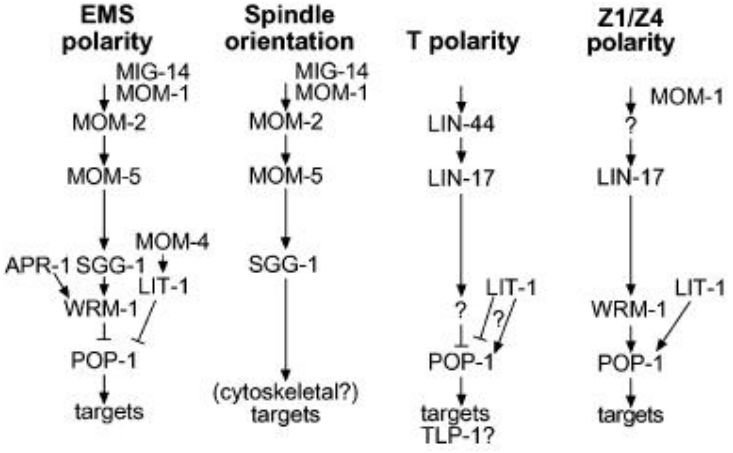

Figure 2.5. Noncanonical Wnt signalling pathways in *C. elegans* (Herman, 2002)

Furthermore, a noncanonical pathway has been proposed to control the polarity of B cells. It has been proposed to be a PCP like pathway (Wu and Herman, 2006). Recently a novel noncanonical Wnt pathway which is β-catenin independent like noncanonical pathways in vertebrate and fly has also been proposed. This pathway is thought to act in neuronal specification (Hingwing *et al.*, 2009).

2.3. Wnt Signalling in Human Diseases

Wnt signalling cascade, that controls many important developmental processes of cells and tissues involving cell differentiation and proliferation, leads to tumour formation when its components are altered. Especially, mutations in Wnt/β-catenin pathway are observed in different tumour types. Therefore, misregulated Wnt/β-catenin pathway is considered as a cancer cause (Miller *et al.,* 1999; Polakis, 2000; Peifer and Polakis, 2000; Giles *et al.*, 2003). Furthermore, there is strong evidence that Wnt5a is active in tumour formation although it is also considered as tumour suppressor (Weeraratna *et al.,* 2002).

The very first Wnt protein identified, Wnt1, was originally identified as an integration site for mouse mammary tumour virus (Nusse and Varmus, 1982). Furthermore, other Wnt ligands that are able to activate Wnt/β-catenin pathway are also observed in different tumour cells. However, direct relation between Wnt ligands and oncogenesis is unclear. Nevertheless, defects in Wnt ligands cause different diseases like schizophrenia and kidney damage (Miller *et al.,* 1999; Polakis, 2000; Giles *et al.*, 2003; Moon *et al.,* 2004).

Other than the Wnt proteins of wnt1, wnt-2, wnt3a, etc. that activate Wnt/β-catenin pathway, Wnt proteins such as Wnt5a are considered as tumour suppressors because of the antagonism between the noncanonical pathways they activate and the canonical Wnt/β-catenin pathway (Kuhl *et al.*, 2001; Veeman *et al.*, 2003; Moon *et al.,* 2004; Mikels and Nusse, 2006). However, Wnt5a is also reported to be active in oncogenesis. Increased Wnt5a expression is observed in human melanoma cells. Further observation states decrease in tumour grade when Wnt5a is inhibited. Therefore, Wnt5a is reported to increase cell motility and invasiveness of human melanoma cells (Weeraratna *et al.,* 2002).

Mostly, tumour formation related to Wnt signalling is the result of mutations in β-catenin and proteins related to β-catenin such as APC, Axin, Gsk3β, dishevelled in some cases and transcription factors that are activated by β-catenin. The common point of all the mutations causing tumour formation is the excessive accumulation of β-catenin in the nucleus. This causes overexpression of many target genes that are involved in cell growth (Miller *et al.,* 1999; Polakis, 2000; Giles *et al.*, 2003). Mutations in β-catenin gene (CTNNB1) are the most frequent defects observed in many different tumour cells. In addition, altered β-catenin not only overexpresses target genes with functions in cell growth but also provide cell survival signals to prevent apoptosis. High mutation frequency of β-catenin (CTNNB1) has been reported in colorectal cancer (CRC), liver cancer (Hepatocellular carcinoma; HCC) and thyroid cancer. Less mutation frequency has also been observed in some other tumour types such as ovarian, melanoma and prostate tumours (Miller *et al.,* 1999; Polakis, 2000; Giles *et al.*, 2003; Moon *et al.,* 2004).

Although, dishevelled upregulates β-catenin, overexpression or constitutive activation of dishevelled has not been reported to be involved in cancer. Its dual function in both canonical and noncanonical Wnt pathways might be the reason (Polakis, 2000). On the other hand, elevated expressions of dishevelled in lung cancer and promotion of the formation of lymphomas have also been reported (Polakis, 2000; Moon *et al.,* 2004).

Axin, Gsk3β and APC in spite of being downregulators of β-catenin are reported in tumour cells. Inactivation of Gsk3β in T-cell lymphomas has been reported in mice. Furthermore, loss of function mutations that eliminates β-catenin binding sites in Axin gene causes upregulation of β-catenin, which eventually leads to cancer (Polakis, 2000; Moon *et al.,* 2004). Third downregulator of the β-catenin which is reported to cause tumour formation is APC. Mutations in APC are the second most frequent mutations in Wnt signalling cascade. APC mutations are most frequently seen in colorectal cancer (CRC). Up to 85% of colorectal cancer cells have mutations in APC. Moreover, 300 different disease related mutations of APC have been also reported. Loss of function mutation in APC gene is the common reason of the oncogenesis and the most common mutation in APC is the deletion of AAAAG in codon 1309 (Polakis, 2000; Giles *et al.*, 2003; Moon *et al.,* 2004).

The summary of the Wnt signalling network genes that are involved in diseases are summarized in Table 2.1. It can be observed that except Wnt5a, genes causing diseases are all Wnt/β-catenin signalling pathway genes.

Table 2.1. Wnt signalling network genes that are involved in diseases (Moon *et al,* 2004)

Gene	Condition/disease
WNT1	Schizophrenia
WNT3	Tetra-amelia
WNT4	Intersex
WNT4	Kidney damage
WNT4	Polycystic kidney disease
WNT5a	Leukaemia
WNT5a	Metastasis
sFRP3	Osteoarthritis
DVL	Lung cancer
APC	Cancer
AXIN	Cancer
AXIN2	Cancer
β-catenin	Cancer

3. MATERIALS AND METHODS

3.1. Data

3.1.1. Protein-Protein Interaction Data Sources

Three public databases, BioGRID (Stark *et al.*, 2006), DIP (Salwinski *et al.*, 2004) and MINT (Chatr-aryamontri *et al.*, 2007) provide protein interactions for different organisms from experimentally determined interactions existing in literature and also from results of high throughput studies.

Biological General Repository for Interaction Datasets (BioGRID) is a freely accessible database available at http://www.thebiogrid.org, consisting of both physical and genetic interactions. Protein interactions included are collected from different high throughput studies (HTP). BioGRID release version 2.0 includes more than 116000 interactions for major model organisms such as *Saccharomyces cerevisiae, Caenorhabditis elegans, Drosophila melanogaster and Homo sapiens*. Furthermore, BioGRID includes large-scale HTP two-hybrid surveys for *C. elegans* consisting of 2801 proteins with physical 4453 interactions (Stark *et al.*, 2006).

The Database of Interacting Proteins (DIP) is another freely accessible database consisting of physical interactions for different organisms available at http://dip.doembi.ucla.edu. Interactions included in the database are collected both from high throughput experiments and small scale experiments existing in literature. DIP includes 20728 proteins with 57683 interactions for 274 different organisms. Furthermore, it includes 2650 proteins with 4043 interactions (3996 interactions from high throughput studies, 47 interactions from literature) for *C. elegans* (Salwinski *et al.*, 2004).

The Molecular Interaction Database (MINT) is a database of experimentally verified physical interactions freely accessible at http://mint.bio.uniroma2.it/mint/. MINT focuses only on physical interactions between proteins. MINT does not include genetic or

computationally inferred interactions. Interactions stored in the database are from large scale genome wide experiments (90 per cent of the interactions) and from small scale experiments reported in literature. As of September 2006, the database stores 27461 proteins with more than 95000 physical interactions from 325 different organisms. The database includes 3258 proteins with 7370 interactions for *C. elegans* (Chatr-aryamontri *et al.*, 2007).

3.1.2. Domain-Domain Interaction Data Sources

SUPERFAMILY freely available at http://supfam.cs.bris.ac.uk/SUPERFAMILY/ is a database consisting of structural and functional protein annotations for completely sequenced organisms. The database provides protein domain assignments at the Structural Classification of Proteins (SCOP), superfamily and family levels in over 1000 organisms. Domain assignments are generated using an expert curated set of profile hidden Markov models. The database includes 19945 domains in 905 unique superfamilies for *C. elegans* (Gough *et al.*, 2001).

3.1.3. Sequence Data Sources

Universal Protein Resource (UniProt) freely available at http://www.uniprot.org provides a stable and comprehensive resource on protein sequences and functional annotations. UniProt has four components for different uses. The UniProt Knowledgebase (UniProtKB) is an expertly curated database for protein information. The UniProt Archive (UniParc) is a comprehensive repository of the history of all protein sequences. UniProt Reference Clusters (UniRef) combines related sequences based on sequence identity for fast searches. Lastly, The UniProt Metagenomic and Environmental Sequences (UniMES) is a database of metagenomic and environmental data (The UniProt Consortium, 2008).

Both protein and domain sequences for many organisms are available at The UniProt Knowledgebase (UniProtKB). UniProtKB consists of two sections which are UniProtKB/Swiss-Prot and UniProtKB/TrEMBL. UniProtKB/Swiss-Prot contains manually annotated sequences, extracted from literature and curator-evaluated computational analysis. On the other hand, UniProtKB/TrEMBL contains only

computationally analyzed records. UniProtKB/Swiss-Prot release version 57.3 contains 468851 sequences for 250 organisms from 179873 references. It contains 3233 sequence entries for *C. elegans*. UniProtKB/TrEMBL release version 57.3 contains 7916844 sequences for 100 organisms. It contains 20678 sequence entries for *C. elegans* (The UniProt Consortium, 2008).

3.1.4. Gene Ontology Annotations

Results of genomic sequencing have shown that large division of genes specifying core biological roles are shared by all eukaryotes. Therefore, it is often possible to transfer knowledge of the shared biological roles from one organism to another. The Gene Ontology (GO) project aims to provide dynamic, controlled sets of descriptions that can be applied to all of the eukaryotes even if the knowledge of biological roles of genes are changing and accumulating. For this purpose, GO project has developed three structured controlled descriptions (ontologies) in a species-independent manner. These ontologies describe gene products in terms of their associated biological processes, cellular components and molecular functions. These ontologies are freely available at http://www.geneontology.org (The Gene Ontology Consortium, 2000).

Biological process refers to a biological objective that the gene or its product contributes. A process is accomplished via one or more molecular functions where a chemical or physical transformation occurs. ‘Cell growth and maintenance’ or ‘signal transduction’ are examples of broad biological processes whereas ‘pyrimidine metabolism’ or ‘cAMP biosynthesis is an example of more specific biological processes (The Gene Ontology Consortium, 2000).

Molecular function defines biochemical activity of gene product. This set of ontology defines what is done but does not define where or when. Broad examples for molecular functions are ‘enzyme’, ‘transporter’ or ‘ligand’ whereas adenylate cyclase’ or ‘Toll receptor ligand’ are examples of more specific functions (The Gene Ontology Consortium, 2000).

Lastly, cellular component refers to the place where a gene product is active. Cellular component terms help to understand structure of eukaryotic cells. It should be noted that not all the ontologies are applicable to all organisms. 'Ribosome', 'Golgi apparatus' or 'nuclear membrane' is example of cellular component descriptions (The Gene Ontology Consortium, 2000).

3.2. Data Visualization Tools

3.2.1. Cytoscape

Cytoscape is general purpose open source bioinformatics software capable of visualizing biomolecular interaction networks and integrating data such as annotations, gene expression profiles into visualized networks. It is available at http://www.cytoscape.org/. Core components of Cytoscape provide features for visualization and data integration. Features of Cytoscape can be extended via using plug-ins. Plug-ins can add features to core components such as, different analysis algorithms, new layouts and connection to databases. Although Cytoscape Core is open source, plug-ins are separate software which might be protected under any license (Shannon *et al.*, 2003).

3.2.2. Osprey

Osprey is freely available software capable of visualizing data-rich biological networks. Osprey utilizes BioGRID datasets as query. In addition, user-defined datasets can also be uploaded. Osprey visualizes genes/proteins by nodes and interactions between nodes by edges. The software embeds GO categories, experimental evidence and/or data source information in nodes and edges and also offers variety of graphical layouts. Osprey can be downloaded from http://biodata.mshri.on.ca/osprey (Stark *et al.*, 2006).

3.3. Methods

The methods this work utilizes can be reviewed in three main sections which are reconstruction of the signalling pathway, micro analysis and macro analysis of the reconstructed pathway. Reconstruction of the signalling pathway incorporates constitution

of core protein set, reconstruction of protein-protein interactions and domain-domain interactions of the signalling network. Micro scale analysis investigates homology and conservation of the protein sequences and the domains among different species. Macro scale analysis investigates the topological properties of the reconstructed signalling pathway by graph theoretical approaches.

3.3.1. Reconstruction of the Signalling Pathway

The core proteins that are known to have roles in noncanonical Wnt signalling were first identified by reviewing literature. Reviewed literature contains information of the proteins not only in *C. elegans* but also in different species such as *H. sapiens, M. musculus, X. leavis, D. rerio, D. melanogaster* (Wodarz and Nusse, 1998; Miller *et al.,* 1999; Kuhl *et al.,* 2000; Huelsken and Birchmeier, 2001; Herman, 2002; Korswagen, 2002; Giles *et al.,* 2003; Veeman *et al.,* 2003; Wang and Malbon, 2003; Herman and Wu, 2004; Eisenmann, 2005; Kohn and Moon, 2005; Pandur, 2005; Korswagen, 2007; Slusarski and Pelegri, 2007; Hingwing *et al.,* 2009). Proteins that are not reported in literature on *C. elegans* literature were found by sequence homology using proteins in different species. The methods and tools for finding homologues are described in detail in section 2.3.2 under 'Micro Analysis of the Reconstructed Pathway'. Consequently, a core set of proteins that are found to act in noncanonical Wnt signalling in *C. elegans* was formed using literature information and literature based homology search.

3.3.1.1. Reconstruction of PPI network. Protein-protein interaction (PPI) network was formed by using Selective Permissibility Algorithm, which was developed by Arga and colleagues (2007). Selective Permissibility Algorithm (SPA) designed for network reconstruction has two characteristic elements, the input proteins and selection criterion. The input proteins are set of proteins (core proteins) known to have a common role in a biological network such as signal transduction. The selection criterion is the collection of GO annotations (biological processes, cellular components and molecular functions) belonging to the core proteins (Arga *et al.,* 2007).

In the first step of the algorithm, all protein-protein interactions available for the core proteins are extracted from interactome data sources given in section 2.1.1. In the next

step, extracted interactions are evaluated according to the selection criterion. The interacting proteins are either accepted or rejected according to their GO terms. A candidate interacting protein is accepted into network if all of its three GO annotations match with those of the core protein set. Further, accepted proteins are used as input proteins and their interacting partners are evaluated using again the selection criterion. This iterative procedure is continued until no new interacting proteins can be added to the network (Arga *et al.,* 2007).

In the present work input proteins are the core protein set formed for "noncanonical Wnt signalling in *C. elegans"* (Table 4.1). The annotation collection set used for the evaluation incorporates GO annotations of these proteins. The annotations are retrieved from GO website (http://www.geneontology.org). The annotation collection table comprises of 160 unique annotations from 97320 total annotations of *C. elegans* present in GO website. The whole Annotation collection set used is presented in Appendix A.

3.3.1.2. Reconstruction of DDI network. The domain-domain interactions (DDIs) of the reconstructed signalling pathway are found and analyzed. Domains can be viewed as the basic components of proteins. They are defined as a minimal part of a protein that is able to perform a function; mostly experimentally assigned function. Most proteins consist of more than one domain. Further, domain combination of a protein can determine its function and its role in biological networks. Therefore it is logical to investigate protein interactions via domain combinations (Veretnik *et al.,* 2004; Santonico *et al.,* 2005; Bornberg-Bauer *et al.,* 2005). For instance, the diversification between canonical Wnt signalling pathway and noncanonical Wnt signalling pathways starts at Dishevelled protein. Dishevelled proteins are known to have three distinct domains which are DIX, PDZ and DEP. It has been reported that the utilization of DIX and PDZ domains activates canonical Wnt signalling pathway whereas DEP and PDZ domains are required for the activation of noncanonical Wnt signalling pathway (Sheldahl *et al.,* 2003; Wharton *et al.,* 2003; Pan *et al.,* 2004; Wallingford and Habas, 2005).

In the light of these information, domain-domain interactions of the reconstructed pathway were identified and analyzed as follows: Firstly, protein domains of the reconstructed pathway were extracted from the whole domain set of *C. elegans* available at

SUPERFAMILY website. Secondly, proteins were separated into domains and protein interactions were converted into domain interactions. As a consequence a protein-protein interaction network turns into domain-domain interaction network (Appendix E). Thirdly, every domain's probability of interaction was calculated by dividing domain's number of interaction to total number of domain interactions of the reconstructed network. Obtaining probabilities of the domain interactions enables one to comment about the most probable protein interactions and functions of the network. Furthermore, it is also possible to predict interactions for proteins pairs that have no interaction data in literature but have one or more domain pairs at high probability of interaction occurrence.

3.3.2. Micro Scale Analysis of the Reconstructed Pathway

The arrangement of amino acids provides proteins' functionality. Thus proteins in different species having sequence similarity are likely to have same functions. Furthermore, functionally important amino acids are often evolutionarily conserved from lower organisms to higher organisms. As domains are basic components of proteins performing a function, it is expected that, important domains might have similar sequences and be conserved through species. Therefore, to validate the presented network, the core proteins' similarity and conservation among different species have been investigated in detail at both protein and domain level. In other words, both overall and partial similarity and conservation of the proteins in noncanonical Wnt/Ca^{2+} signalling network were analyzed using bioinformatics tools.

The similarity of the proteins was analyzed on sequence level since the structures of the core proteins are unavailable. Sequence homology search was performed using Position-Specific Iterated BLAST algorithm (Altschul *et al.,* 1997). Position-Specific Iterated BLAST (PSI-BLAST) algorithm is available at http://blast.ncbi.nlm.nih.gov/Blast.cgi. PSI-BLAST algorithm is as fast as BLAST algorithm but it is more sensitive to weak but biologically relevant sequence similarities by building a position-specific scoring matrix to find distantly related sequences in latter iterations (Altschul *et al.,* 1997). The known noncanonical Wnt signalling pathway proteins in *H. sapiens* was used as initial query set. Protein and domain sequences of these initial proteins were obtained from UniProtKB (http://www.uniprot.org). A set of proteins

with similar sequences was obtained and used for the evaluation of similarity. The proteins used for comparison are from organisms of *H. sapiens, M. musculus, X. leavis, D. rerio, D. Melanogaster* and *C. elegans.* The reason of selecting these organisms is the higher occurrence of reports in literature on Wnt signalling compared to others. The same procedure was repeated for analyzing domain homology.

The conservation of these Wnt signalling proteins was analyzed by employing The Conseq Server (Berezin *et al.,* 2004) which is available at http://conseq.tau.ac.il/. This server first finds similar sequences available in UniProtKB, of the query sequence by utilizing PSI-BLAST algorithm. Further, The Conseq Server multiply (multiple sequence alignment-MSA) aligns the found sequences by utilizing ClustalW (Thompson *et al*., 1994). Finally the server estimates rate of amino acid substitutions in the multiple sequence alignment. Amino acids are scored and coloured according to their substitution rate. Consequently, less changing amino acids, thus the more conserved ones, have high scores while frequently changing amino acids, thus the less conserved ones, have low scores. This procedure is repeated for each protein in noncanonical Wnt signalling pathway and the subject proteins of the selected organisms are extracted from the whole BLAST results for evaluation. The same procedure is carried out for analyzing the domain conservation.

Both sections of UniProtKB, i.e. UniProtKB/Swiss-Prot and UniProtKB/TrEMBL, were scanned in order to find the maximum number of candidate proteins for the noncanonical Wnt signalling pathway.

3.3.3. Macro Scale Analysis of the Reconstructed Pathway

Topological analysis of complex biological networks enables better understanding of both the distribution of the components and their relations with each other either as group or as single. In order to validate the biological meaning of the reconstructed network, graph theory based analysis was performed. Further, the modules estimated to have important roles in the network were searched and analyzed.

3.3.3.1. Graph Theoretical Analysis. Graph theory in which proteins are illustrated as nodes and interactions between proteins as edges, explains several basic topological

properties that enables comparison and characterization of different complex biological networks. These properties are as follows:

- Degree (Connectivity): Degree, k, specifies number of links (edges) a node has to other nodes. In directed networks where edges have selected directions nodes have k_{in} degrees to specify edges pointing them and k_{out} degrees to specify outgoing edges from the nodes.
- Mean Degree (Connectivity): The average number of edges that a node has in a network defines the mean degree of a network. <k> denotes the mean degree. For an undirected network with N nodes and L edges, <k> is calculated by Equation 3.1:

$$< k > = \frac{2L}{N} \tag{3.1}$$

- Degree distribution: The probability of a selected node to have exactly k edges is defined as degree distribution, P(k). P(k) is calculated by dividing the number of nodes N(k), with k= 1,2,... edges by total number of nodes, N. Different classes of networks exist due to the degree distribution.
- Mean path length: Path length is the measurement of edges that are passed through to travel between two nodes. Shortest path is the path with the smallest number of edges among the alternative paths between two nodes. Furthermore, mean path length, < l >, is the average of the all shortest paths between all nodes in a network.
- Network diameter: The shortest path among all paths in the network is defined as the network diameter, d.

Degree distribution is the property of the network while connectivity is a property of nodes. Furthermore, mean path length and network diameter offer a measure of network's navigability (Jeong *et al.,* 2000; Barabási and Oltvai, 2004).

Until recently, biological networks have been modelled as random networks. In random networks each pair of nodes are connected with a probability, p that follows a Poisson distribution ($P(k) \sim e^{-k}$) peaking strongly at <k> and decaying exponentially for higher k values than <k>. Thus, most of the nodes have same number of edges and

probability of finding a highly connected node (node with k> <k> connectivity) is extremely rare (Barabási and Albert, 1999; Jeong *et al.,* 2000; Barabási and Oltvai, 2004).

On the other hand, empirical studies reported serious deviations from random network structure for not only world-wide web and social networks but also for biological networks. Therefore it is supposed that biological networks have a scale-free network structure instead of random network structure. Scale-free networks follow a power law distribution ($P(k) \sim k^{-\gamma}$, γ is degree exponent) and are extremely heterogeneous where a few highly connected nodes (hubs) are dominant and rest of the less connected nodes are linked to network by these hubs. Comparison between random networks and scale-free networks is illustrated in Figure 3.1. Furthermore, another feature of biological networks is their small-world character meaning that any two nodes in a network are connected by relatively shorter paths among existing paths. Moreover, a biological network can be characterized by its network diameter which is the shortest path in the network. Briefly, biological networks are thought to have scale-free topology with small-world character (Barabási and Albert, 1999; Jeong *et al.,* 2000; Wagner and Fell, 2001; Girvan and Newman, 2002; Barabási and Oltvai, 2004).

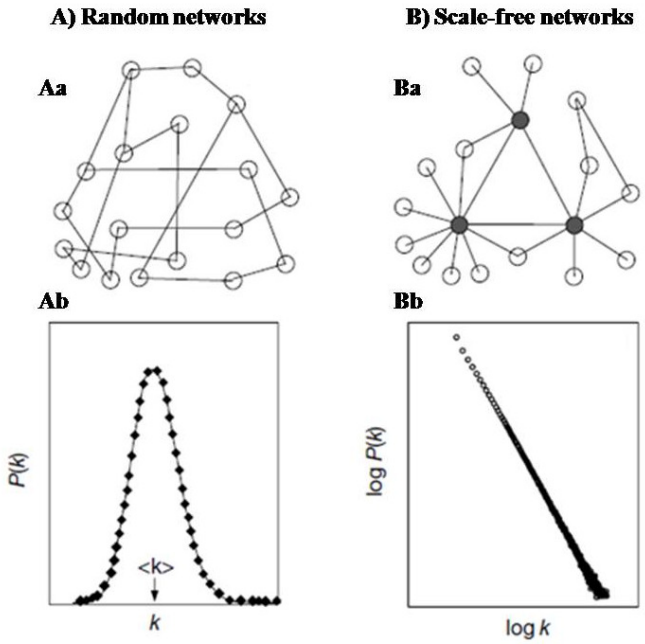

Figure 3.1. Representative structures of networks and their distributions (Jeong *et al.,* 2000)

On this basis reconstructed network's topology was investigated to validate its biological nature. For this purpose a novel algorithm developed by Arga and colleagues (2007) was used. The input of the algorithm is an adjacency matrix (S). Adjacency matrix is a binary square matrix representing edges that are protein-protein interactions. It is formed as follows:

- If (i,j) is an interacting protein pair than S(i,j)=1
- If (i,j) is not interacting protein than S(i,j)=0

The rows and columns of the adjacency matrix gives the ingoing and outgoing degrees of nodes (k_{in} and k_{out}) respectively. This enables the calculation of the degree distribution as it is dependent on the total number of edges. Furthermore, the shortest paths are calculated using the fact that $S^n(i,j)$ which is the element of the n^{th} power of the adjacency matrix, gives the number of paths of length *n* from node *i* to *j*. Finally, the mean path length and the network diameter are calculated from shortest path lengths (Arga *et al.,* 2007).

3.3.3.2. Module Detection. The modules in the reconstructed network were found and analyzed. Accumulating evidences suggest that protein-protein interaction networks are composed of subgraphs defined as modules or cliques. These modules are groups of proteins performing a certain biological function together. Therefore, detecting modules in networks enables better understanding of the cellular role of these networks. Furthermore, the prediction of a protein with unknown function as part of a module can give a clue about its function as proteins in a module have a common function. There are several algorithms developed in order to find modules in biological networks. While one approach forces all proteins of a network to be part of a module (complex, cluster, clique or sub-network), another approach suggests that only proteins which are more related (connected) with each other compared to other proteins should be considered as modules. Recent studies show that proteins sharing a common cellular function have more connections with each other than the other proteins (Bader and Hogue, 2003; Przulj *et al.,* 2004; Li *et al.,* 2008).

On this basis, the modules in the reconstructed network were detected and checked by using two Cytoscape plug-ins which are, MCODE (Bader and Hogue, 2003) and BiNGO (Maere *et al.,* 2005).

Molecular Complex Detection (MCODE) algorithm detects highly connected regions in networks in three stages which are vertex weighting, complex prediction and post-processing (to add or remove proteins from predicted complexes). The algorithm first weights vertices according to local network density. Then, it takes highest scoring vertex as seed and moves outward from the seed vertex including vertices above a given threshold. At the third stage, the proteins with degree lower than 2 are filtered. There are two additional options of the algorithm that can be utilized at this stage. First is "fluff" which increases the size of the complex for a given threshold. Proteins that are added with this parameter for one complex can also be added to another. As proteins are not strictly components of only one complex in cell, this option increases biological meaning of the results. The second option available is called "haircut", it filters the added proteins to the complex by fluff option that has lower connectivity to core complex. Optimal running parameters of the algorithm (Bader and Hogue, 2003) are;

- Fluff = true (on) with threshold = 0.2 (allows only 20 per cent difference)
- Haircut = true (on)
- Node score cut off = 0.

As explained above, proteins in a module share common cellular functions. Thus, estimated modules are checked whether its components share common function(s) or not. For this purpose another Cytoscape plug-in named, The Biological Networks Gene Ontology tool (BiNGO) was used (Maere *et al.,* 2005). For a given set of proteins (the detected modules, in our case), BiNGO retrieves GO annotations from GO website (http://www.geneontology.org) and evaluates their occurrence in a given network. Outputs are the most significant GO annotations (Maere *et al.,* 2005).

4. RESULTS AND DISCUSSION

4.1. Core Proteins

The core protein set of the noncanonical Wnt/Ca^{2+} signalling pathway was formed as explained at section 3.3.1. 64 core proteins were identified in *C.elegans* as a result of literature and homology search. The core proteins are presented in alphabetical order in Table 4.1.

Table 4.1. Core proteins

Name	ORF name	Name	ORF name	Name	ORF name
ama-1	F36A4.7	F59D6.7	F59D6.7	plc-1	F31B12.1
C09G4.2	C09G4.2	K06A9.1	K06A9.1	plc-2	Y75B12B.6
C24A1.3	C24A1.3	kin-1	ZK909.2	plc-3	T01E8.3
cdc-42	R07G3.1	kin-26	T06C10.6	plc-4	R05G6.8
ced-10	C09G12.8	lin-17	Y71F9B.5	pll-1	K10F12.3
cfz-2	F27E11.3	lin-44	E01A2.3	pph-4.2	Y49E10.3
chw-1	F22E12.2	lit-1	W06F12.1	pqn-15	C24A8.3
cmk-1	K07A9.2	mig-1	Y34D9B.1	pqn-20	C37A2.2
cnb-1	F55C10.1	mig-2	C35C5.4	rac-2	K03D3.10
crp-1	Y32F6B.3	mig-5	T05C12.6	rap-3	C08F8.7
cwn-1	K10B4.6	mom-2	F38E1.7	rho-1	Y51H4A.3
cwn-2	W01B6.1	mom-4	F52F12.3	sel-8	C32A3.1
dlk-1	F33E2.2	mom-5	T23D8.1	sma-9	T05A10.1
dpy-22	F47A4.2	mpk-1	F43C1.2	tax-6	C02F4.2
dsh-1	C34F11.9	mpz-1	C52A11.4	unc-43	K11E8.1
dsh-2	C27A2.6	pde-1	T04D3.3	Y105C5A.24	Y105C5A.24
egl-20	W08D2.1	pde-2	R08D7.6	Y39B6A.1	Y39B6A.1
egl-27	C04A2.3	pde-3	E01F3.1	Y71H2AL.1	Y71H2AL.1
egl-4	F55A8.2	pde-4	R153.1	ZC373.4	ZC373.4
egl-8	B0348.4	pde-6	Y95B8A.10	ZK856.8	ZK856.8
F22H10.2	F22H10.2	pkc-1	F57F5.5		
F30A10.1	F30A10.1	pkc-2	E01H11.1		

In literature, 10 proteins were reported to take place in the general structure of noncanonical Wnt/Ca^{2+} signalling pathway. Fly incorporates 16 core proteins and human incorporates 29 core proteins in noncanonical Wnt/Ca^{2+} signalling pathway (Table 4.2).

Table 4.2. Core proteins of different species.

General	WORM	FLY	HUMAN
WNT5A	cwn-2	N/A	WNT5A
FZD2	mom-5	fz2	FZD2
DVL	dsh-1 mig-5 dsh-2	dsh	DVL1 DVL2 DVL3
PLC	egl-8 plc-2 pll-1 plc-1	norpA Plc21C	PLCB1 PLCB2 PLCB3 PLCB4
CAMKII	unc-43 cmk-1 ZC373.4	CaMKII	CAMK2A CAMK2B CAMK2D CAMK2G
CACN	ZK856.8 Y71H2AL.1 cnb-1 F30A10.1 tax-6 pph-4.2	CACNB2 CACNA1 CG2185 CACNB CACNA-14F Pp2B-14D	CHP PPP3CA PPP3CB PPP3CC PPP3R1 PPP3R2
PKC	pkc-2 pkc-1	inaC Pkc53E	PRKCA PRKCB PRKCG
NFAT	C24A8.3 C37A2.2 F47A4.2 K06A9.1 T05A10.1	NFAT	NFAT5 NFATC1 NFATC2 NFATC3 NFATC4
TAK1	mom-4	Tak1	TAK1
NLK	lit-1	nmo	NLK

Although *C. elegans* is a simpler organism than both human and fly 64 core proteins were identified by homology search. As noncanonical Wnt signalling mechanism is less known in *C. elegans* a bigger set of core proteins was taken into consideration at the start of SPA in order to find more interactions related to Wnt/Ca^{2+} signalling. As, SPA rejects unrelated interactions, utilizing a large input set (core proteins) does not cause formation of an unrelated interaction network. The complete set of the core proteins justified according

to the general structure of noncanonical Wnt/Ca^{2+} signalling pathway is available in Table B.1 (Appendix B).

4.2. Protein-Protein Interaction Network

All protein interactions related to Wnt/Ca^{2+} signalling were first extracted from BioGRID database (release version 2.0.45; October 1 2008). Then DIP (release version Celeg20090126; January 26 2009) and MINT (release version 2009-04-14-mint-Caenorhabditis; April 14 2009) databases were scanned in order to increase the number of interacting proteins.

The reconstruction of the network started with 16 core proteins out of 64 proteins as the remaining 48 proteins does not have any interaction information either in databases scanned or in literature. Core proteins that have interactions are listed in Table 4.3.

Table 4.3. Core proteins with interactions

Name	ORF name	Name	ORF name
chw-1	F22E12.2	mom-4	F52F12.3
cmk-1	K07A9.2	mpk-1	F43C1.2
crp-1	Y32F6B.3	pkc-1	F57F5.5
dsh-1	C34F11.9	pkc-2	E01H11.1
dsh-2	C27A2.6	pqn-20	C37A2.2
egl-20	W08D2.1	sma-9	T05A10.1
lit-1	W06F12.1	unc-43	K11E8.1
mig-5	T05C12.6	Y39B6A.1	Y39B6A.1

Selective Permissibility Algorithm (SPA) initiated with 16 core proteins converged at the 9^{th} step (neighbour). The resulting Wnt/Ca^{2+} signalling network consists of 667 proteins (715 when non-interacting core proteins are added) with 922 interactions in total. An overview of the network is displayed in Figure 4.1.

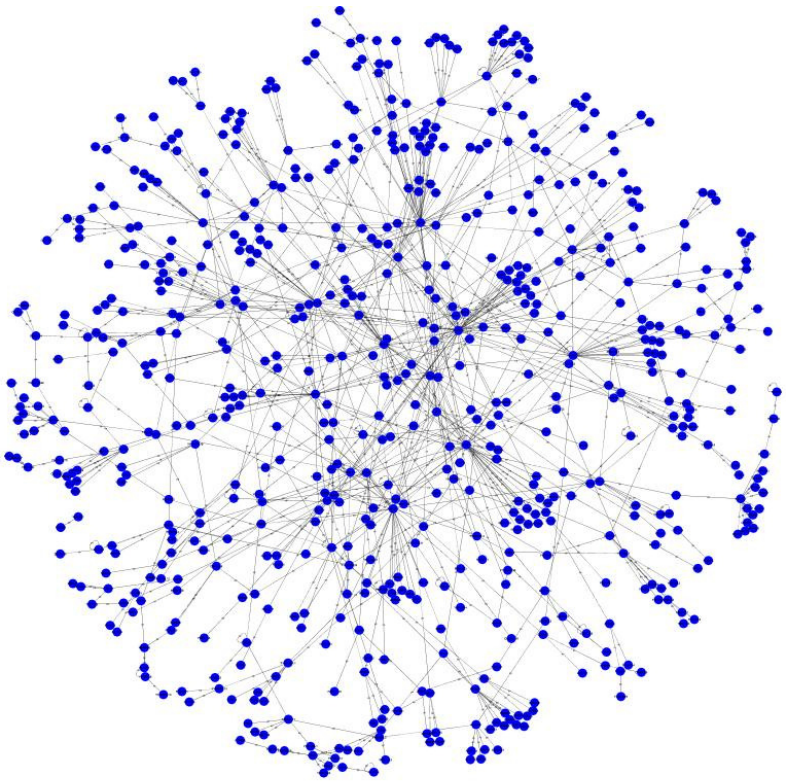

Figure 4.1. Representation of the reconstructed Wnt/Ca^{2+} signalling network

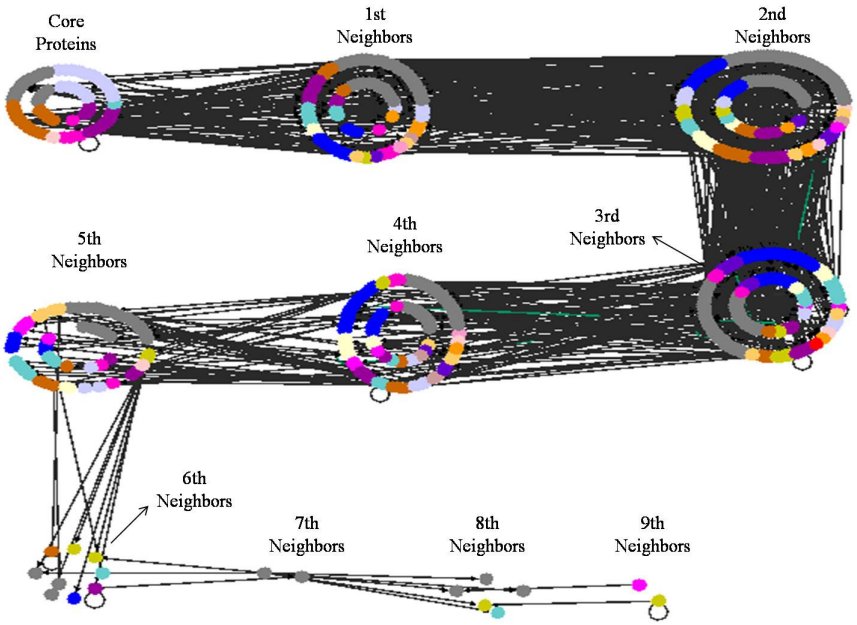

Figure 4.2. Representation of the reconstructed networks as neighbours

Furthermore, the interacting proteins are grouped as neighbours (Figure 4.2) which also display the steps of the algorithm. Every set of proteins in each neighbour is an input set of the algorithm to identify the next set of neighbours, eventually leading to the reconstruction of the whole network. Figure 4.1 was produced in Cytoscape and Figure 4.2 was produced in Osprey.

The statistical overview of the algorithm inputs for each neighbour is displayed in Figure 4.3a and b. The numbers represent the accepted or rejected proteins in each step,

Finally, all interactions are listed as binary interactions in Table B.2 (Appendix B). All proteins classified as neighbours are also listed in Appendix B (Tables B.3-B.9).

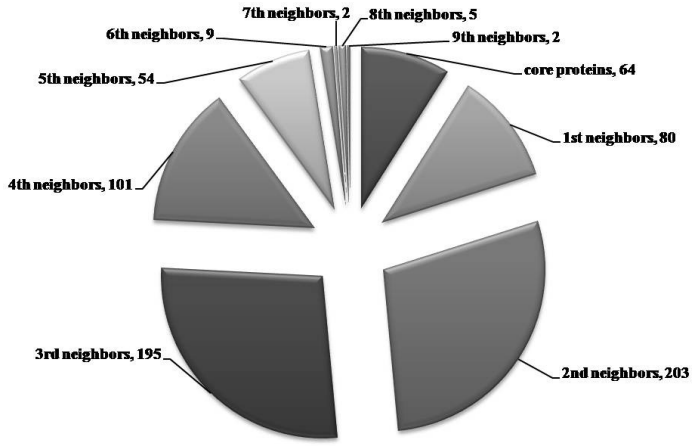

Figure 4.3. Accepted protein statistics

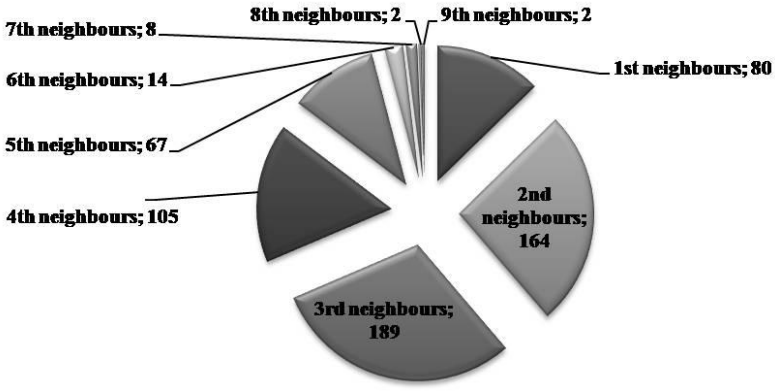

Figure 4.4 Rejected proteins statistics

Two important factors affect the efficiency of the reconstruction of the network: Reliability of the GO annotations and availability of protein interactions.

4.2.1. Reliability of the GO Annotations

Selection criterion set of the SPA is a collection of GO annotations of core proteins. Proteins are accepted to the network if all three of their GO annotations (Biological process, molecular function, cellular component) match the GO annotation collection set of core proteins. Otherwise, they are rejected, i.e. if one or more of their GO annotations do not match. Therefore, it is important to have a reliable set of GO annotations to be able to add the most significant proteins to the network. In present work, significance means the proteins' relationship to noncanonical Wnt/Ca^{2+} signalling pathway.

Sources of GO annotations are experimental, computational or other such as author statements or automatic electronic annotations. Experimental assays give results with high confidence while computational and other assay results might need more justification. Therefore, a reliable GO annotation collection set should include mostly, annotations with experimental evidence. Here, core proteins should have reliable GO annotations as the annotation collection table is formed from these annotations.

The annotation collection set utilized in the present work is composed of 160 unique annotations. These are extracted from the whole 97320 GO annotations of *C. elegans* available at GO website. The annotation collection set incorporates 13 cellular component annotations, 39 molecular function annotations and 108 biological process annotations. Sources of these annotations are summarized in Table 4.4. Other sources column refers to electronic annotations and traceable author statements. Experimental and computational sources are listed in table A.1 (see Appendix A). The complete list of annotations with their identity numbers and sources is given in Appendix A (Tables A.1-A.3).

Table 4.4. Sources of GO annotations

# of annotations	Sources of annotations			
	Experimental	Computational	Other	total
Cellular component	4	1	8	13
Molecular function	5	1	33	39
Biological process	92	1	15	108

Results indicate that only 30.8 per cent of cellular component annotations and 12.8 per cent of molecular function annotations have reliable sources (experimental) whereas 85.2 per cent of biological process annotations are reliable (Table 4.4). According to these results, cellular component and molecular function annotations are less reliable than biological process annotations.

However, the present annotation collection set consists of mostly reliable annotations (63 per cent) as biological process annotations dominate the annotation collection set. Furthermore, when the whole annotation set of *C. elegans* available at GO website (submission date; June 3, 2009) is investigated it can be seen that 47742 of total 97320 annotations (49 per cent) are rather reliable. Consequently the annotation collection set used can be considered as a reliable selection criterion set as it includes more reliable annotations than the complete *C. elegans* annotation set.

4.2.2. Availability of Protein Interactions

Reconstruction of the network also requires availability of protein interactions. Furthermore, protein interactions should have proteins with matching GO annotations in order to pass SPA's selection criterion. In other words, there should be enough protein interactions of good quality in order to achieve a successful evaluation. Therefore, three different databases (BioGRID, DIP and MINT) were scanned and a large set of core proteins (Table 4.1) was used in order to extract as much protein interactions as possible.

Although a large core protein set consisting of 64 proteins was utilized and three databases were scanned, only the interactions of 16 proteins could be obtained. In order to understand whether interaction information is missing only for *C. elegans* or for other organisms also, core protein sets for fly and human were formed and investigated in the

same methodology (see section 3.3.1). The results indicate that protein interactions of the core proteins in Wnt/Ca^{2+} signalling pathway are also missing in other species (Table 4.5). In Table 4.5, 16 interacting core proteins are listed. The complete set of interactions of 64 core proteins is given in Table C.1 (Appendix C).

Table 4.5. Core protein interactions of different species

CAEEL			FLY			HUMAN		
Orf name	Name	# of PPI	Name	Orf name	# of PPI	Name	Orf name	# of PPI
WNT			**WNT**			**WNT**		
W08D2.1	egl-20	1	N/A		0	WNT5A	EG7474	1
FZD			**FZD**			**FZD**		
T23D8.1	mom-5	0	fz2	CG9739	0	FZD2	EG2535	1
DVL			**DVL**			**DVL**		
C34F11.9	dsh-1	4	dsh	CG1836	1	DVL1	EG1855	12
T05C12.6	mig-5	33				DVL2	EG1856	17
C27A2.6	dsh-2	6				DVL3	EG1857	10
PLC			**PLC**			**PLC**		
B0348.4	egl-8	0	norpA	CG3620	3	PLCB1	RP4-654A7.1	3
CAMKII			**CAMKII**			**CAMKII**		
K11E8.1	unc-43	2	CaMKII	CG18069	2	CAMK2A	EG815	9
K07A9.2	cmk-1	1				CAMK2B	EG816	1
CACN			**CACN**			**CACN**		
C02F4.2	tax-6	0	CACNA1	CG1455	3	PPP3CA	EG5530	4
PKC			**PKC**			**PKC**		
E01H11.1	pkc-2	3	inaC	CG6518	3	PRKCA	EG5578	16
F57F5.5	pkc-1	1	Pkc53E	CG6622	0	PRKCB	EG5579	10
NFAT			**NFAT**			**NFAT**		
T05A10.1	sma-9	1	NFAT	CG11172	0	NFATC2	RP5-1009H6.1	8
C37A2.2	pqn-20	1				NFATC1	EG4772	3
Y39B6A.1	Y39B6A.1	11				NFAT5	EG10725	1
TAK1			**TAK1**			**TAK1**		
F52F12.3	mom-4	2	Tak1	CG18492	1	TAK1	EG6885	9
NLK			**NLK**			**NLK**		
W06F12.1	lit-1	10	nmo	CG7892	1	NLK	EG51701	3
F43C1.2	mpk-1	23						
CDC42			**CDC42**			**CDC42**		
F22E12.2	chw-1	1	cdc42	CG12530	1	cdc42	G25K	10
Y32F6B.3	crp-1	1						

It can be seen that core proteins that have no interactions in *C. elegans* such as mom-5 (FZD), egl-8 (PLC), tax-6 (CACN) have either no interactions or have limited

interactions in fly and human, i.e. homologues of mom-5 (FZD), egl-8 (PLC), tax-6 (CACN) have 0, 3, 3 interactions in fly and 1, 3, 4 interactions in human, respectively. On the other hand, core proteins that have higher number of interactions also have higher number of interactions in other species. Dishevelled homologues in *C. elegans*; dsh-1, mig-5 and dsh-2 have 4, 33, 6 interactions, respectively while human homologues; DVL1, DVL2 and DVL3 have 12, 17, 10 interactions, respectively. As a consequence, a general lack of protein interactions of the Wnt/Ca^{2+} signalling pathway core proteins can be seen. It can be suggested that this general lack of information may cause a decrease in the efficiency of the network reconstruction process.

In addition to the lack of protein interactions, lack of GO annotations decreased available interactions even further. Most of the protein interactions include proteins that have neither cellular component annotations nor molecular function annotations. In most cases, both of the cellular component and molecular function annotations are missing. Consequently, most of the interactions were rejected by the algorithm utilized (Figure 4.3).

In order to increase possible interactions, newer releases of the databases (see section 4.2 for dates) were scanned. 11 new interactions from MINT database, 7 new interactions from DIP database have been obtained for the core proteins. However none of these interactions could be added to the network. Due to lack of GO annotations of the interacting proteins SPA rejected all of these interactions. In the same manner, many of the interactions obtained were also rejected in latter stages of the algorithm. When the accepted and rejected proteins are further investigated and compared, it can be observed that, the lack of GO annotations affects the number of available protein interactions significantly. A total of 651 proteins were included to the network while 631 proteins were rejected (Figure 4.3 and Table 4.6). Consequently, it can be suggested that the general lack of GO annotations for Wnt/Ca^{2+} signalling pathway in *C. elegans* decreased the number of the possible proteins in the reconstructed network.

Table 4.6. Comparison of accepted and rejected proteins

	# of proteins	
	Accepted	Rejected
1st neighbours	80	80
2nd neighbours	203	164
3rd neighbours	195	189
4th neighbours	101	105
5th neighbours	54	67
6th neighbours	9	14
7th neighbours	2	8
8th neighbours	5	2
9th neighbours	2	2
Total	651	631

4.2.3. Overview of the Reconstructed Network

When the interactions in the Wnt/Ca^{2+} signalling network are further investigated, it can be seen that protein interactions in *C. elegans* have been obtained for most of the core proteins compared to the general Wnt/Ca^{2+} signaling pathway proteins given in literature (Figure 4.4). For seven core proteins, protein interactions have been found in databases. Five of these (DVL, CAMKII, TAK1, NLK, NFAT) belong to Wnt/Ca^{2+} pathway and two of these (PKC, CDC42) belong to another noncanonical pathway; Wnt/PCP pathway. Only interactions for WNT5A, FZD2, PLC and CACN proteins could not be obtained in *C. elegans* (Figure 4.4). However, the missing interactions of WNT5A and FZD2 proteins cause a significant weakness for the reconstructed network as WNT5A is the ligand that activates Wnt/Ca^{2+} signalling pathway and FZD2 is the receptor of WNT5A protein. In other words the interaction that activates the Wnt/Ca^{2+} signalling pathway is missing.

Furthermore, direct interactions that form the signalling pathway between core proteins are mostly missing. Only interactions between CAMKII, TAK1 and NLK could be obtained. Mom-4(TAK1)-lit-1(NLK) interaction that inhibits Wnt/β-catenin signalling pathway is also suggested in different studies (Korswagen, 2002; Eisenmann, 2005; Korswagen, 2007).

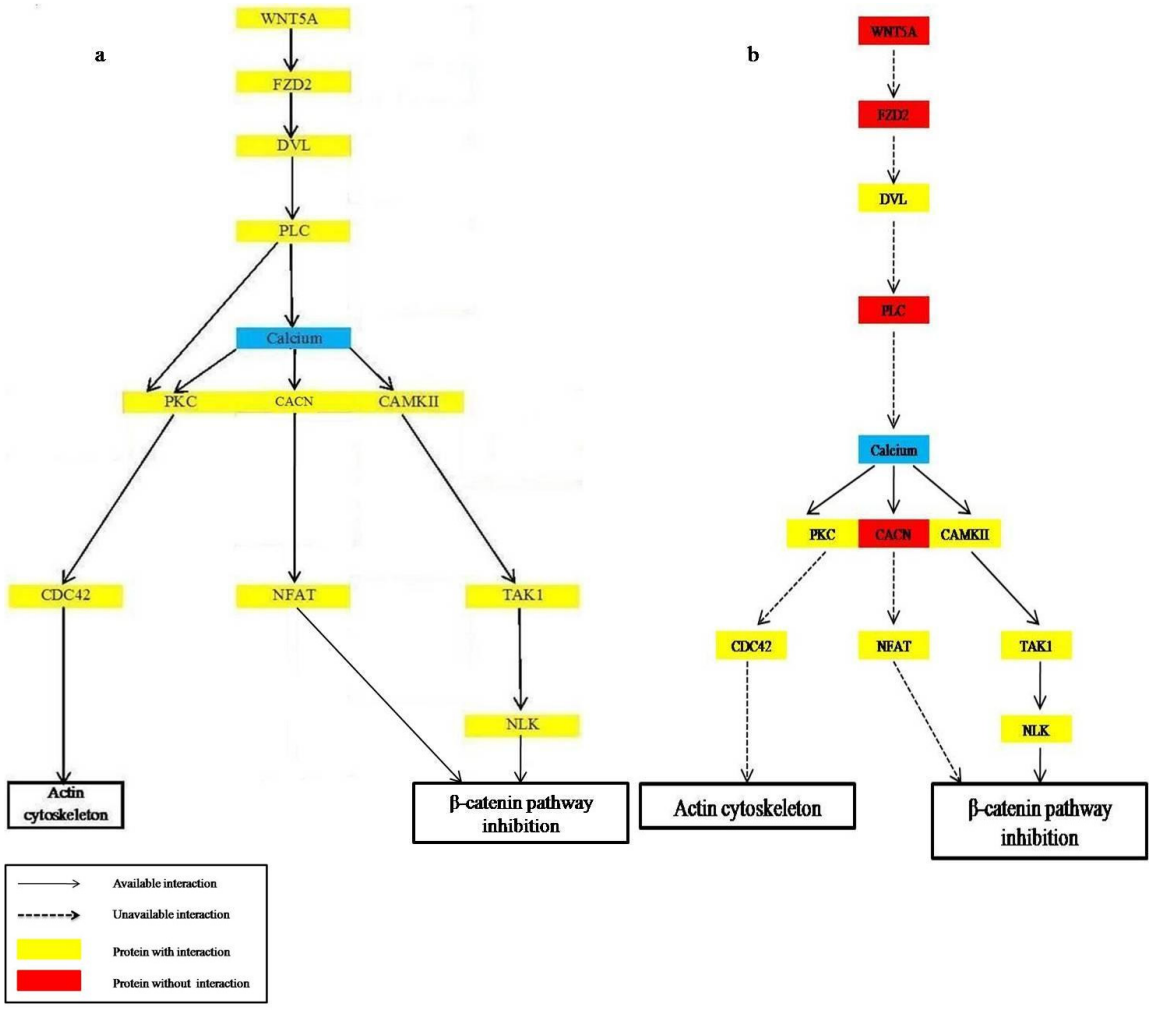

Figure 4.5. Wnt/Ca^{2+} signalling pathway a) General structure b) Reconstructed pathway in *C. elegans*

Nevertheless, the reconstructed network successfully covered most of the noncanonical Wnt signalling cascade in *C. elegans* in spite of missing direct interactions between core proteins.

In order to increase protein interactions in the reconstructed network, a large set of core proteins was utilized. To find possible core proteins in *C. elegans,* micro scale analysis was performed based on sequence homology and conservation. Results of micro analysis were used to form a large core protein set.

4.3. Micro Scale Analysis

Homologues of human Wnt/Ca^{2+} signalling pathway proteins were searched among different species. Analyses were carried on protein sequence level due to the lack of 3D structures of proteins. As functionally important amino acids are often evolutionarily conserved among species, analyzing conservation of proteins might give an opinion about functionally important parts of proteins.

The inputs of the micro analysis are the core proteins and domains of human Wnt/Ca^{2+} signalling pathway (see Table 4.7). Sequence information was retrieved from UniProt (http://www.uniprot.org). Methods and tools utilized are described in section 2.3.2 in detail. Micro scale analysis was performed among *H. sapiens* (Human), *M. musculus* (Mouse), *X. leavis* (Xenla), *D. rerio* (Danre), *D. melanogaster* (Drome), *C. elegans* (Caeel).

The output of the micro scale analysis at both protein and domain level is the percentage of identical amino acids. Conserved amino acid percentage is another output of the micro analysis. A summary of the micro scale analysis is displayed in Figure 4.6. In Figure 4.6, first row of a protein corresponds to the identity percentage and conservation percentage of amino acids, respectively. Second row of a protein corresponds to the identity percentage of the domains included. It can be observed that the similarity of Wnt/Ca^{2+} signalling pathway proteins among species is consistent with evolutional hierarchy, i.e. human proteins are more similar to mouse homologues than fly or worm homologues.

4.3.1. Identity and Conservation of *C. Elegans* Proteins

When the results are investigated, it is observed that the proteins in *C. elegans* have low identity percentages, i.e. they are less similar (see Figure 4.5). The identities of proteins in *C. elegans* vary between 25 to 50 per cent. Further investigation for WNT5A, FZD2, DVL1, DVL2 and DVL3 gives similar results. The identity percentages of WNT5A homologues vary between 29 to 43 per cent. The identity percentages of FZD2 homologues vary between 23 to 38 per cent. Lastly, the identity percentages of dishevelled

homologues vary between 32 to 43 per cent (see Table 4.8). Moreover, sequence homology and conservation of the domains in *C. elegans* are slightly higher than the overall protein identity and conservation (see Table 4.9).

Table 4.7. Core protein interactions of different species

Wnt/Ca^{2+} core proteins	Protein name	Domain name	Domain name	Domain name	Domain name
WNT	WNT5A	-			
FRIZZLED	FZD2	FZ			
DISHEVELLED	DVL1	DIX	PDZ	DEP	
	DVL2	DIX	PDZ	DEP	
	DVL3	DIX	PDZ	DEP	
PLC	PLCB1	PI-PLC X-box	PI-PLC Y-box	C2	
	PLCB2	PI-PLC X-box	PI-PLC Y-box	C2	
	PLCB3	PI-PLC X-box	PI-PLC Y-box	C2	
	PLCB4	PI-PLC X-box	PI-PLC Y-box	C2	
CAMKII	CAMK2A	Protein kinase			
	CAMK2B	Protein kinase			
	CAMK2G	Protein kinase			
CAN	CHP	EF-hand 1	EF-hand 2	EF-hand 3	EF-hand 4
	PPP3CA	-	-	-	-
	PPP3CB	-	-	-	-
	PPP3CC	-	-	-	-
	PPP3R1	EF-hand 1	EF-hand 2	EF-hand 3	EF-hand 4
	PPP3R2	EF-hand 1	EF-hand 2	EF-hand 3	EF-hand 4
PKC	PRKCA	C2	Protein kinase	AGC-kinase C-terminal	
	PRKCB	C2	Protein kinase	AGC-kinase C-terminal	
	PRKCG	C2	Protein kinase	AGC-kinase C-terminal	
NFAT	NFAT5	RHD			
	NFATC1	RHD			
	NFATC2	RHD			
	NFATC3	RHD			
	NFATC4	RHD			
TAK1	TAK1	Protein kinase			
NLK	NLK	Protein kinase			

Nine NFAT homologues were included in core protein data set (Table B.1, Appendix B). However, the results of micro scale analysis indicated a similarity less than 25 per cent for NFAT homologues in *C. elegans*. Moreover, no significant domain similarity could be found for these homologues (Figure 4.5). Furthermore, available literature also supports the results of micro scale analysis for NFAT proteins, i.e. no NFAT homolog has been reported in invertebrates so far (Crabtree and Olson, 2002; Schulz and Yutzey, 2004). Therefore, NFAT homologues in the core protein set (Table B.1) and their interactions (see Table 4.6) were removed from the reconstructed network.

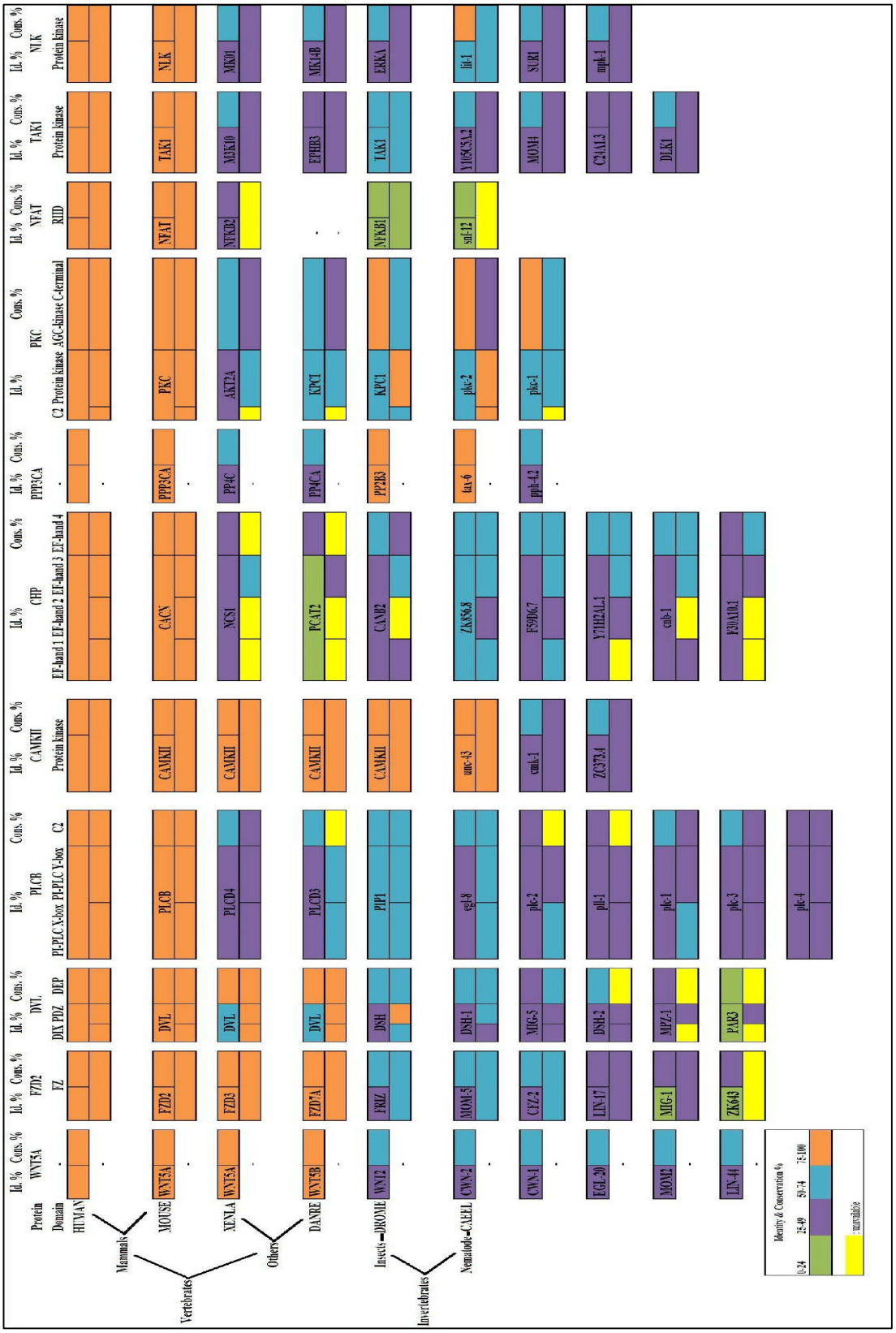

Figure 4.6. Summary of Micro scale analysis

Table 4.8. Overall protein identity and conservation

Name	Uni name	Organism	Identity %	Conservation %	aa length
WNT5A	**P41221**	**HUMAN**	**100**	**100**	380
WNT5A	P22725	MOUSE	98	98	380
WNT5A	P31286	XENLA	87	92	380
WNT5B	Q92050	DANRE	81	90	363
WNT2	P34889	CAEEL	43	64	360
WNT2	P28465	DROME	43	58	352
EGL-20	Q9TVJ1	CAEEL	35	52	393
WNT1	P34888	CAEEL	34	55	372
MOM2	Q10459	CAEEL	33	50	362
LIN-44	Q6F3D0	CAEEL	29	47	348
FZD2	**Q14332**	**HUMAN**	**100**	**100**	565
FZD2	Q9JIP6	MOUSE	99	99	570
FZD2	Q9PUU6	XENLA	88	92	551
FZD7A	Q7SZR7	DANRE	75	84	559
FRIZ	P18537	DROME	47	62	581
MOM-5	O16147	CAEEL	38	54	480
CFZ-2	Q9U8U6	CAEEL	37	51	585
LIN-17	Q94132	CAEEL	31	49	585
MIG-1	Q9N532	CAEEL	23	43	559
Y73B6BL.21	Q9GUF5	CAEEL	39	60	314
DVL1	**O14640**	**HUMAN**	**100**	**100**	695
DVL1	P51141	MOUSE	94	96	695
DVL3	Q6DKE2	XENLA	64	76	717
DVL-1	Q1LYR0	DANRE	64	74	707
DSH	P51140	DROME	49	65	623
DSH-1	B1Q238	CAEEL	43	60	570
MIG-5	O61720	CAEEL	35	49	666
DSH-2	Q18239	CAEEL	32	50	518
MPZ-1	Q7JMN5	CAEEL	32	47	838
DVL2	**O14641**	**HUMAN**	**100**	**100**	736
DVL2	Q60838	MOUSE	96	97	736
DVL2	P51142	XENLA	74	81	736
DVL2	Q803Q5	DANRE	68	77	747
DSH	P51140	DROME	58	75	623
DSH-1	B1Q238	CAEEL	41	56	570
MIG-5	O61720	CAEEL	33	50	666
DSH-2	Q18239	CAEEL	33	50	518
MPZ-1	Q7JMN5	CAEEL	35	50	838
DVL3	**Q92997**	**HUMAN**	**100**	**100**	716
DVL3	Q61062	MOUSE	98	99	716
DVL3	Q6DKE2	XENLA	84	91	717
DVL3	Q4KM11	DANRE	68	77	676
DSH	P51140	DROME	52	66	623
DSH-1	B1Q238	CAEEL	45	61	570
DSH-2	Q18239	CAEEL	32	49	666
MIG-5	O61720	CAEEL	34	34	518

Table 4.9 Overall domain identity and conservation (D. name corresponds to domain names, Id to identity and Con. to conservation)

Name	Uni name	Organism	D. Name	Id. %	Con.%	D. Name	Id. %	Con.%	D. Name	Id. %	Con.%
WNT5A	**P41221**	**HUMAN**	**-**	**-**							
WNT5A	P22725	MOUSE	-	-							
WNT5A	P31286	XENLA	-	-							
WNT5B	Q92050	DANRE	-	-							
WNT2	P34889	CAEEL	-	-							
WNT2	P28465	DROME	-	-							
EGL-20	Q9TVJ1	CAEEL									
WNT1	P34888	CAEEL	-	-							
MOM2	Q10459	CAEEL	-	-							
LIN-44	Q6F3D0	CAEEL	-	-							
FZD2	**Q14332**	**HUMAN**	**FZ**	**100**	**100**						
FZD2	Q9JIP6	MOUSE	FZ	100	100						
FZD2	Q9PUU6	XENLA	FZ	96	98						
FZD7A	Q7SZR7	DANRE	FZ	92	97						
FRIZ	P18537	DROME	FZ	61	75						
MOM-5	O16147	CAEEL	FZ	57	75						
CFZ-2	Q9U8U6	CAEEL	FZ	50	65						
LIN-17	Q94132	CAEEL	FZ	41	60						
MIG-1	Q9N532	CAEEL	FZ	38	60						
Y73B6BL.21	Q9GUF5	CAEEL	FZ	39	60						
DVL1	**O14640**	**HUMAN**	**DIX**	**100**	**100**	**PDZ**	**100**	**100**	**DEP**	**100**	**100**
DVL1	P51141	MOUSE	DIX	100	100	PDZ	100	100	DEP	98	100
DVL3	Q6DKE2	XENLA	DIX	77	85	PDZ	91	98	DEP	82	93
DVL-1	Q1LYR0	DANRE	DIX	91	97	PDZ	94	100	DEP	90	96
DSH	P51140	DROME	DIX	68	81	PDZ	83	93	DEP	64	82
DSH-1	B1Q238	CAEEL	DIX	39	63	PDZ	68	83	DEP	56	77
MIG-5	O61720	CAEEL	DIX	32	55	PDZ	43	58	DEP	53	73
DSH-2	Q18239	CAEEL	DIX	36	60	PDZ	49	71	DEP	0	0
MPZ-1	Q7JMN5	CAEEL	DIX	0	0	PDZ	39	60	DEP	0	0
DVL2	**O14641**	**HUMAN**	**DIX**	**100**	**100**	PDZ	**100**	**100**	DEP	**100**	**100**
DVL2	Q60838	MOUSE	DIX	100	100	PDZ	100	100	DEP	100	100
DVL2	P51142	XENLA	DIX	78	85	PDZ	98	98	DEP	92	96
DVL2	Q803Q5	DANRE	DIX	79	90	PDZ	97	100	DEP	88	94
DSH	P51140	DROME	DIX	71	84	PDZ	82	93	DEP	62	82
DSH-1	B1Q238	CAEEL	DIX	50	66	PDZ	68	82	DEP	61	77
MIG-5	O61720	CAEEL	DIX	39	56	PDZ	40	58	DEP	58	78
DSH-2	Q18239	CAEEL	DIX	45	63	PDZ	47	71	DEP	0	0
MPZ-1	Q7JMN5	CAEEL	DIX	0	0	PDZ	40	60	DEP	0	0
DVL3	**Q92997**	**HUMAN**	**DIX**	**100**	**100**	PDZ	**100**	**100**	DEP	**100**	**100**
DVL3	Q61062	MOUSE	DIX	100	100	PDZ	100	100	DEP	100	100
DVL3	Q6DKE2	XENLA	DIX	85	93	PDZ	95	98	DEP	97	100
DVL3	Q4KM11	DANRE	DIX	80	92	PDZ	97	98	DEP	94	98
DSH	P51140	DROME	DIX	70	85	PDZ	82	93	DEP	65	84
DSH-1	B1Q238	CAEEL	DIX	40	67	PDZ	71	82	DEP	56	80
DSH-2	Q18239	CAEEL	DIX	40	65	PDZ	52	71	DEP	0	0
MIG-5	O61720	CAEEL	DIX	31	54	PDZ	41	58	DEP	56	78

The results of micro analysis indicate that *C. elegans'* proteins have the lowest similarity values. Literature studies on *C. elegans* also suggested that Wnt signalling pathway proteins in *C. elegans* have lower sequence similarity than other organisms. For example, a body of work on β-catenin homologues in *C. elegans* stated a sequence similarity between 17 to 27 per cent. However, cnidarians which are evolutionarily simpler organisms than nematodes are reported to have 60 per cent of similarity to human homologues. In spite of low sequence similarity, *C. elegans* homologues studied are proved to function in Wnt signalling (Natarjan *et al.,* 2001; Korswagen, 2002). Therefore, it can be suggested that obtained *C. elegans* homologues from micro scale analysis except NFAT homologues were acceptable as proper homologues of Wnt/Ca^{2+} signalling pathway proteins. Thus, they were included as core proteins to the reconstructed network.

4.4. Macro Scale Analysis

4.4.1. Graph Theory Based Analysis

The results of the graph theory based topological analysis are illustrated in Figures 4.6, 4.7 and Table 4.10. Figure 4.6 shows the overall connectivity (degree) distribution of the proteins. It can be observed that a few proteins have high degrees while most of the proteins have one or two connections. The complete data of Figure 4.6 are included in Appendix D. Further, Figure 4.7 displays degree distribution of the network. The calculated properties based on degree distribution of the reconstructed network are listed in Table 4.10.

Table 4.10. Topological properties of the network

# of nodes	667
# edges	922
Mean degree	2.76
Mean path length	5.39
Network diameter	21

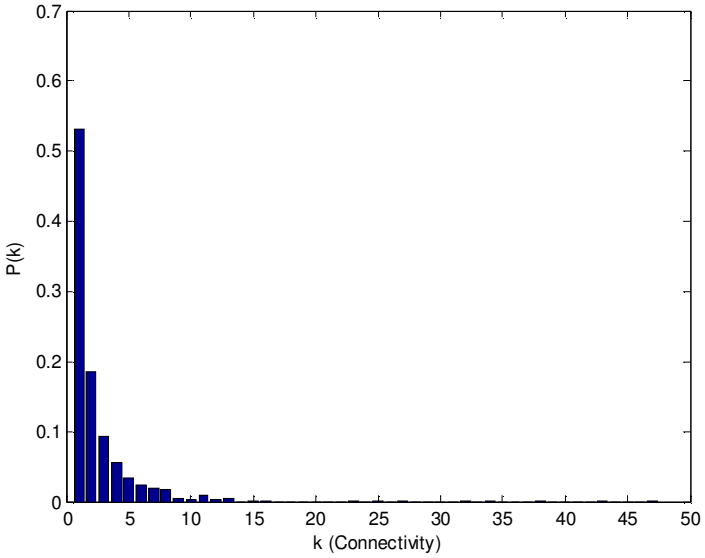

Figure 4.7. Degree distribution

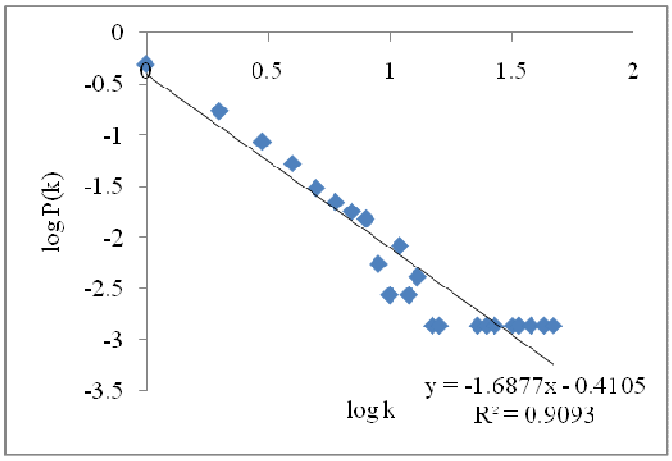

Figure 4.8. Log transform of degrees versus log transform of P(k)

4.4.2. Biological Significance of the Reconstructed Network

Biological networks are reported to have small world characteristics and a scale-free topology following power law distribution (Barabási and Albert, 1999; Jeong *et al.,* 2000; Wagner and Fell, 2001; Girvan and Newman, 2002; Barabási and Oltvai, 2004).

The present results suggest that the reconstructed network follows nearly a power law distribution with γ=1.69 and R^2=0.91 (Figure 4.7). Furthermore, the calculated

topological properties of the network (see Table 4.10) indicate that the network has small world characteristics as mean path length, mean degree and network diameter are smaller than the total number of proteins in orders of magnitude. In addition, these calculated topological properties of the reconstructed network show consistency with literature (see Table 4.11). Table 4.11 is a collection of protein-protein interaction networks of different species and includes graph theoretic properties of these networks (Tekir *et al.,* 2009).

When the calculated topological properties of the reconstructed network are compared to existing *C. elegans* PPI network, only 11 per cent and 12.3 per cent difference between mean degrees and mean path lengths and 50 per cent of differences between network diameters were observed. In addition, the reconstructed network shows only 6.2 per cent, 20.7 per cent and no difference to mean degree, mean path length and network diameter of *H. Sapiens* PPI network respectively (see Table 4.12). Thus it can be suggested that, reconstructed network's topological properties.are similar to existing biological networks.

Table 4.11. Graph theoretic properties of protein interaction networks (Tekir *et al.,* 2009)

Model	# of nodes	# of interactions	Mean Degree	Mean Path Length	Diameter
EGFR (Oda et al, 2005)	329	1795	10.91	4.7	11
S. cerevisiae	2115	2240	2.12	6.8	14
S. cerevisiae	5253	65673	25	5.21	11
S. cerevisiae (signaling)	1363	3649	5.35	6.81	9
DIP (*S. cerevisiae*-Core)	2640	6600	5	5	13
DIP (*S. cerevisiae*)	4773	15444	6.5	5.01	13
MIPS (*S. cerevisiae*)	2043	5434	5.32	7.71	16
DIP (*S. cerevisiae*)	5798	20098	6.93	4.9	12
Uetz (*S. cerevisiae*)	1870	4480	4.79	6.8	15
Ito (*S. cerevisiae*)	3280	8868	5.41	4.9	13
Ito (*S. cerevisiae*-Core)	797	1560	3.91	6.2	16
DIP (*D. Melanogaster*)	7451	22819	6.2	4.4	11
DIP (*C. elegans*)	2638	4030	3.1	4.8	14
DIP (*H. pylori*)	710	1420	4	4.1	9
DIP (*H. sapiens*)	1065	1369	2.6	6.8	21
DIP (*E. coli*)	553	761	2.8	5.5	16
DIP (*M. Musculus*)	329	286	1.7	3.6	9

Table 4.12. Comparison between reconstructed network and different networks.

	Difference (%)		
	Mean degree	Mean path length	Network diameter
DIP (*C. elegans*)	11.0	12.3	50
DIP (*H. sapiens*)	6.2	20.7	0

Furthermore, when highly connected proteins (hubs) in the reconstructed network are investigated, it is found that hubs are either core proteins or proteins that are connected to core proteins. It can be observed that dishevelled homolog mig-5 is one of the hubs and most of the other hubs are neighbours of mig-5 and other dishevelled homologues (dsh-1 and dsh-2) in *C. elegans* (Table 4.13). Moreover, dishevelled is also reported to be an important hub in the whole Wnt signalling network (Giles *et al.,* 2003; Wallingford and Habas, 2005). Therefore, hubs observed in the reconstructed Wnt/Ca^{2+} signalling network are found to be consistent with literature.

Table 4.13. Hubs of the reconstructed network.

Name	Orf name	Degree	Place	Important connections		
T04H1.2	T04H1.2	47	1st neighbour	dsh-2	mig-5	dsh-1
alp-1	T11B7.4	43	1st neighbour	mig-5	dsh-1	
pal-1	C38D4.6	38	1st neighbour	cmk-1		
mig-5	T05C12.6	34	core			
nhr-111	F44G3.9	32	2nd neighbour	mig-5	lit-1	
atn-1	W04D2.1	27	2nd neighbour	pkc-1	mig-5	dsh-1
ZK849.1	ZK849.1	25	1st neighbour	mig-5		
mpk-1	F43C1.2	23	core			

4.4.3. Module Detection and Analysis

Recent studies suggest that protein-protein interaction networks are composed of sub-graphs defined as modules or cliques. These modules are groups of proteins performing a certain biological function together. Therefore, detecting modules in networks enables better understanding of the cellular role of these networks (Bader and Hogue, 2003; Przulj *et al.,* 2004; Li *et al.,* 2008). On this basis, possible modules in the reconstructed network were searched and analyzed.

In order to detect possible modules, MCODE (version 1.31) was executed twice with the parameters stated below:

- Include Loops: true
- Degree Cutoff: 2
- Node Score Cutoff: 0.1
- Haircut: true
- Fluff: false (run1), Fluff: true with Fluff Density Cutoff: 0.2 (run2)

On the first run MCODE was executed without utilizing fluff option to find possible modules. On the second run fluff option was utilized to extend found modules. 4 modules were found (Figures 4.8-11 and Table 4.14).

The biological significance of the molecular functions of these modules was tested by utilizing BiNGO. The molecular functions of the proteins in each module were compared to those of the reconstructed Wnt/Ca^{2+} signalling network. Most significant molecular functions of each module are listed in Table 4.15.

Table 4.14. Modules detected.

Module	Score	Nodes	Edges	Node Ids (ORF)	Node Ids
1	1	3	3	F31E3.3	rfc-4
				C39E9.13	rfc-3
				C54G10.2	rfc-1
2	1	5	7	T11B7.4	alp-1
				C27B7.4	rad-26
				C50C3.8	bath-42
				K08F8.4	pah-1
				Y116A8C.26	Y116A8C.26
3	0.75	4	6	ZC239.15	ZC239.15
				F44G3.9	nhr-111
				ZK849.1	ZK849.1
				W04D2.1	atn-1
4	0.5	2	3	T01G9.5	mei-1
				ZK858.4	mel-26

Table 4.15. Most significant molecular functions of modules.

module id	fluff opt.	GO-ID	Description	p-value
1	off	33170	DNA-protein loading ATPase activity	3.64E-03
	on	33170	DNA-protein loading ATPase activity	5.46E-03
2	off	16597	amino acid binding	7.29E-03
	on	16597	amino acid binding	3.10E-02
3	off	3779	actin binding	7.29E-03
	on	16491	oxidoreductase activity	4.29E-02
4	off	17111	nucleoside-triphosphatase activity	1.51E-01
	on	4820	glycine-tRNA ligase activity	7.29E-03

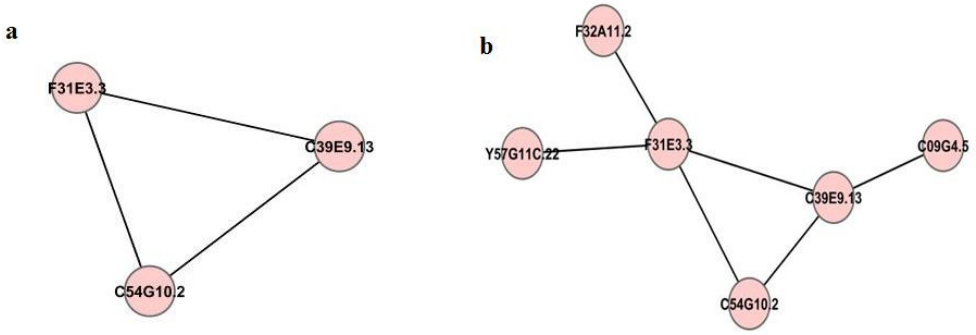

Figure 4.9. Representation of first Module a) fluff off b) fluff on

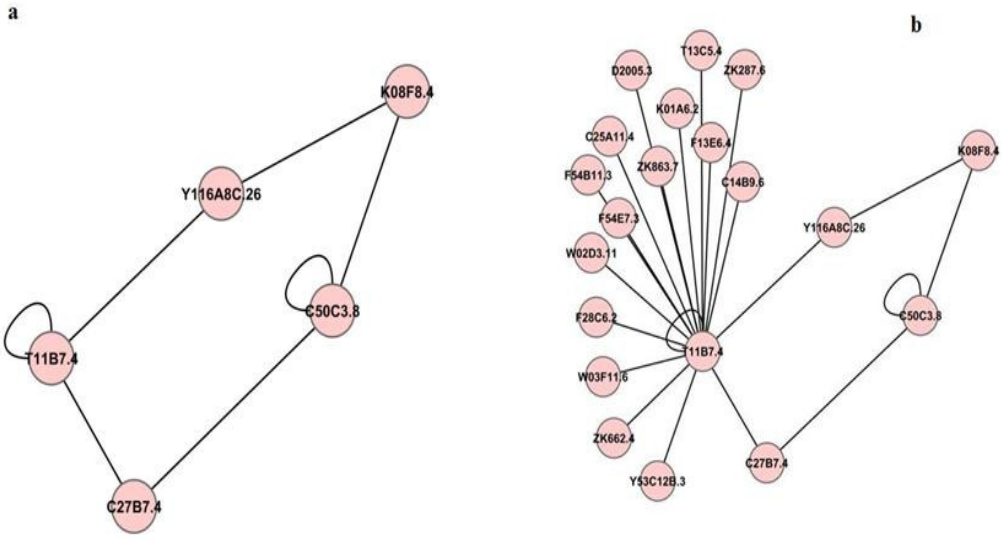

Figure 4.10. Representation of second Module a) fluff off b) fluff on

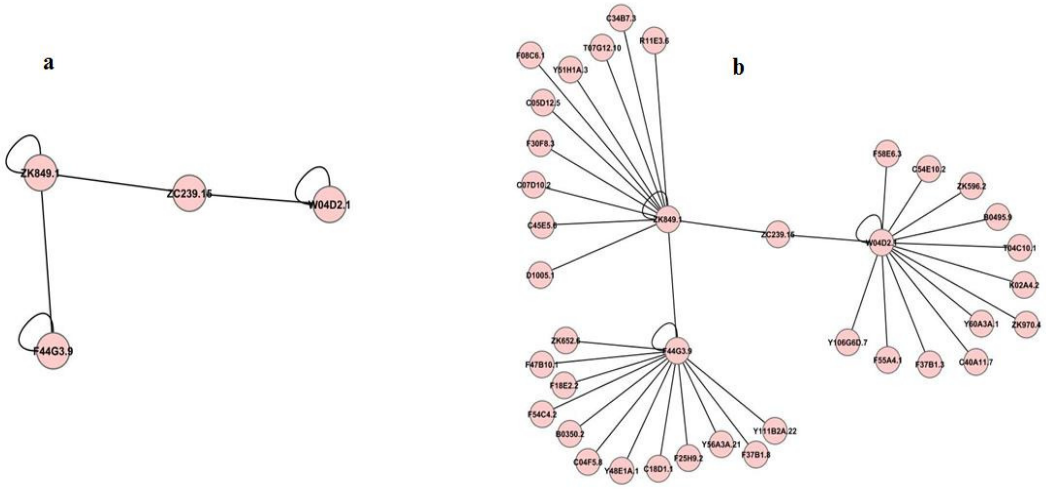

Figure 4.11. Representation of third Module a) fluff off b) fluff on

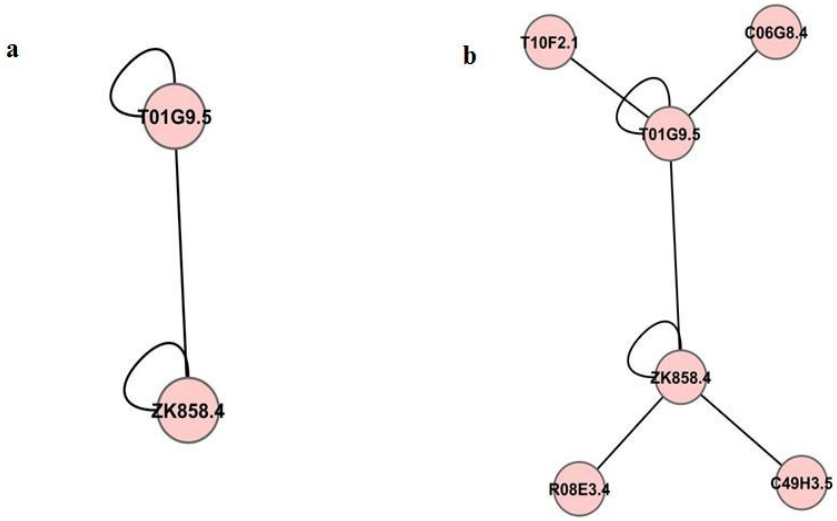

Figure 4.12. Representation of fourth Module a) fluff off b) fluff on

4.4.4. Efficiency of Module Detection

When modules detected are investigated, it is observed that second and third modules include hubs of the network. Second module contains alp-1 which is the second hub with most interactions. Third module contains nhr-111, ZK849.1 and atn-1 as hubs (see Table 4.13). On the other hand, first and fourth modules contain three and two proteins respectively (Figures 4.8-11 and Table 4.14).

The proteins of the first module are all members of the DNA replication factor family. Furthermore, the molecular function of this module is not a general GO annotation. It is on the eleventh step of the related GO annotation tree. Thus, the first module is likely to be a correct detection. On the other hand, molecular functions of both second and third modules are not significant. Their steps at the related GO annotation trees are three and four, respectively, indicating that the proteins in these modules have only general molecular functions in common. Nevertheless, the presence of hubs might have increased the probability of detection. Actually, the presence of hubs increases the number of connections which in turn increase the local density (algorithm parameter, see section *3.3.3.2* for details) resulting in high scores in the algorithm. Therefore, second and third modules containing hubs are also likely to be correct detections. Fourth module containing neither hubs nor significant molecular function has the lowest score. Therefore, fourth module is likely to be an insignificant detection.

Lastly, the general lack of interaction data decreased the efficiency of module detection, since module detection is based on interaction data. Modules are expected to have higher inner degrees than outer degrees (Bader and Hogue, 2003; Przulj *et al.,* 2004; Li *et al.,* 2008). The modules detected in the reconstructed network lack interconnectivity. When detected modules are extended, nodes (proteins) mostly assemble around one node, probably a hub, but they are not connected to each other. Therefore, extended versions of modules (fluff on option in Figures 4.9-12) might not have significant meaning. Furthermore, because of the incomplete GO annotations especially molecular functions, the biological significance of the modules' molecular functions could not be compared to either reconstructed network or whole *C. elegans'* annotation.

4.5. Domain-Domain Interaction Network

The domain assignments of the proteins in the Wnt/Ca^{2+} signalling network were obtained from the SUPERFAMILY website (release version WS147) at November 2nd, 2008. The results are summarized in Table 4.16. The most frequent 15 domain-domain interactions are listed in Table 4.17.

Table 4.16. Summary of DDI analysis

# of proteins with domains	611
# of protein interactions with domains	744
# of unique domains	332
# of domain-domain interactions	2526

Table 4.17. Most frequent domain-domain interactions

Domain 1	ID	Domain 2	ID	Freq.	Freq %
PDZ domain	50157	PDZ domain	50157	31	1.2272
LIM domain	57736	PDZ domain	50157	24	0.9501
Nuclear receptor	57721	Nuclear receptor	57721	20	0.7918
Spectrin repeat	46967	Canonical RBD	54929	20	0.7918
Nuclear receptor ligand-binding domain	48509	Nuclear receptor ligand-binding domain	48509	19	0.7522
Nuclear receptor ligand-binding domain	48509	Nuclear receptor	57721	19	0.7522
Protein kinases, catalytic subunit	88854	Homeodomain	46690	19	0.7522
Canonical RBD	54929	Canonical RBD	54929	18	0.7126
LASP-1	57743	PDZ domain	50157	16	0.6334
Spectrin repeat	46967	Spectrin repeat	46967	16	0.6334
Protein kinases, catalytic subunit	88854	Protein kinases, catalytic subunit	88854	16	0.6334
Eukaryotic type KH-domain (KH-domain type I)	54792	Canonical RBD	54929	12	0.4751
EF-hand modules in multidomain proteins	47547	Spectrin repeat	46967	12	0.4751
DNA-binding protein LAG-1 (CSL)	110080	EGF-type module	57197	12	0.4751

In addition, all unique domains of the reconstructed network and the whole set of domain-domain interaction data are listed in Table E.1 and Table E.2 (Appendix E) respectively.

4.5.1. Analysis of Domain Interactions with High Frequency

The results of DDI analysis indicate that hubs of the network possess domains with high frequency of interaction (Table 4.18). Hubs are important components of biological networks as they maintain biological networks' functionality by combining proteins with low connectivity into the network. Removal of hubs results in a malfunctioning network scattered in unrelated small sub-networks (Jeong *et al.,* 2000). Therefore, detailed knowledge about domains of hubs might support future studies on understanding the Wnt/Ca^{2+} signalling pathway.

Table 4.18. Most frequent domain-domain interactions

Name	Orf name	DDI Id	Frequency	Frequency %
T04H1.2	T04H1.2	52592-50157	8	0.0031671
alp-1	T11B7.4	57736-50157	24	0.0095012
pal-1	C38D4.6	88854-46690	18	0.0071259
mig-5	T05C12.6	57736-50157	24	0.0095012
nhr-111	F44G3.9	57721-57721	20	0.0079177
atn-1	W04D2.1	47577-54929	5	0.0019794
ZK849.1	ZK849.1	50157-50157	31	0.0122724
mpk-1	F43C1.2	88854-88854	12	0.0047506

The most frequent domain interaction observed utilizes PDZ domain (Tables 4.17 and 4.18). PDZ domain consisting of approximately 90 amino acids is the central domain of dishevelled proteins. Furthermore, PDZ domain is known to function in signal transduction and it has been proposed to function as a switch between Wnt signalling pathways depending on the binding partners it engages with (Sheldahl *et al.,* 2003; Wharton *et al.,* 2003; Pan *et al.,* 2004; Wallingford and Habas, 2005). Having the highest frequency of occurrence in the reconstructed network also indicates its role in signal transduction.

Furthermore, its key role in Wnt signalling makes PDZ domain a potential drug target. Inhibition of PDZ domain interactions either with upstream proteins such as Frizzled or downstream proteins might prevent β-catenin accumulation which leads to tumour formation. Recent studies show interaction between Frizzled FZ domain and Dishevelled PDZ domain might be a potential drug target in order to prevent tumour formation caused by β-catenin accumulation (Wong *et al.,* 2003; Fujii *et al*., 2007).

The micro scale analysis results also support domain-domain interaction analysis. Frizzled homologues observed in *C. elegans* have identical similarity to human Frizzled homologues between 31 to 38 per cent and protein sequences are conserved between 50 to 60 per cent. Although overall protein similarity is low, similarity on domain level increases. FZ domain is 57 per cent identical to the human FZ domain and 75 per cent of its sequence is conserved. Furthermore, dishevelled proteins in *C. elegans* that contain PDZ domain show similarity to human dishevelled proteins between 30 to 43 per cent while their sequence conservation is between 47 to 60 per cent. On the other hand, PDZ

domain similarity is higher than overall protein similarity. PDZ domains found in *C. elegans* show over 60 per cent similarity and over 80 per cent of their sequence is conserved (see Table 4.8 and 4.9). As a consequence, *C. elegans* contains conserved homologues of proteins with potential drug target domains with high similarity which enables further investigations to be carried on *C. elegans.*

4.5.2. Overview of the Domain-Domain Interaction Network

The presented domain-domain interaction analysis successfully covers PPI data. 85 per cent of the proteins in the network have a domain assignment and 81 per cent of the protein interactions incorporate interacting domain pairs (Table 4.16). These results enable to extend the investigation of the network in more detail.

In order to add new interactions to the network domains of core proteins without interaction were investigated further. Unfortunately, no significant domain data could be assigned to core proteins without interactions. 28 of them have no domain assignments. The remaining 20 core proteins have domain assignment but lack domain interactions in the reconstructed network. Four out of these 20 proteins are Frizzled homologues; mom-5, cfz-2, lin-17, mig-1. Dishevelled proteins' (dsh-1, mig-5, dsh-2) domains (DIX, PDZ, DEP) are available in the network. Although, no interaction between frizzled and dishevelled is available in the DDI analysis performed, FZ-DIX, FZ-PDZ, FZ-DEP domain interactions are reported in literature (Sheldahl *et al.,* 2003; Wharton *et al.,* 2003; Wong *et al.,* 2003; Pan *et al.,* 2004; Wallingford and Habas, 2005). Therefore, the interactions between frizzled and dishevelled proteins could be added to the reconstructed network.

4.6. Final Status of the Reconstructed PPI Network

PPI network reconstructed by SPA was modified by analyses performed. NFAT homologues and their interactions have been removed according to micro scale analysis' results. Furthermore, Frizzled-Dishevelled interaction has been added according to the domain analysis performed. Final status of the network is displayed in Figure 4.11. Wnt5a interaction could not be added to the reconstructed network even after all the analyses

performed. As Wnt5a is the activator of the Wnt/Ca^{2+} signalling pathway, lack of its interaction is the main shortage of the reconstructed network. Body of work suggest that Wnt homologues, cwn-2, egl-20, and lin-44 activate both canonical and noncanonical Wnt signalling pathways in *C. elegans.* However the activator of the noncanonical signalling pathway is not clear (Herman, 2002; Eisenmann, 2005; Korswagen, 2007).

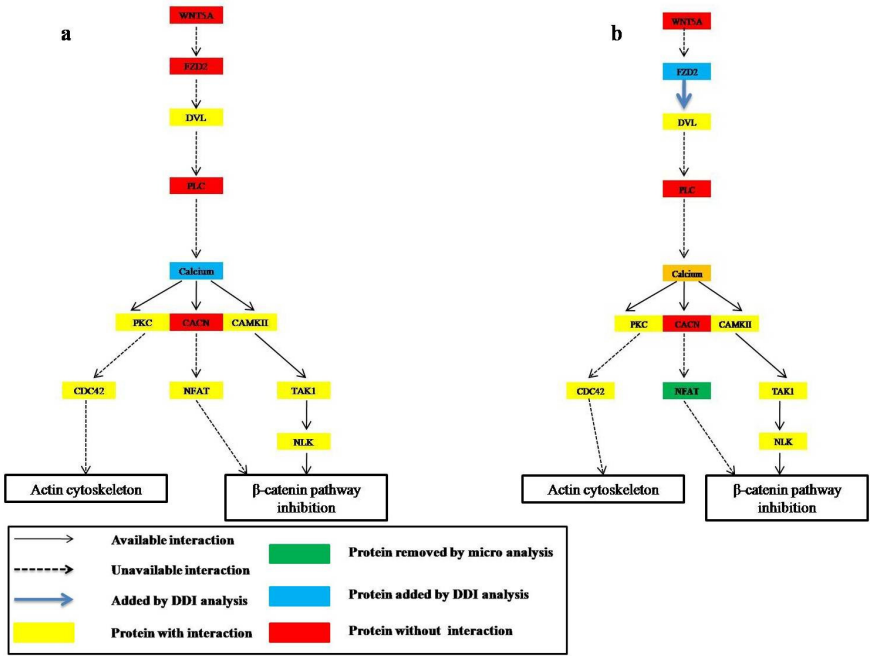

Figure 4.13. Wnt/Ca^{2+} signalling pathway in *C. elegans* a) reconstructed network b) reconstructed network after analyses

4.7. Wnt/Ca^{2+}-Wnt/β-catenin Antagonistic Mechanism in *C. Elegans*

Wnt/Ca^{2+} signalling pathway can inhibit Wnt/β-catenin pathway in two different ways (Figure 2.2). The reconstructed network successfully covers one of these inhibitory branches in Wnt/Ca^{2+} signalling pathway in *C. elegans.* CACN-NFAT branch that inhibits β-catenin function could not be obtained in the reconstructed network. CACN does not have any interactions and there no NFAT homologues in the reconstructed network. On the other hand the reconstructed network successfully covers CAMKII-TAK1-NLK branch which is able to inhibit Wnt/β-catenin signalling pathway. Consequently, antagonism between Wnt/Ca^{2+} and Wnt/β-catenin signalling pathways is partially covered in the reconstructed network.

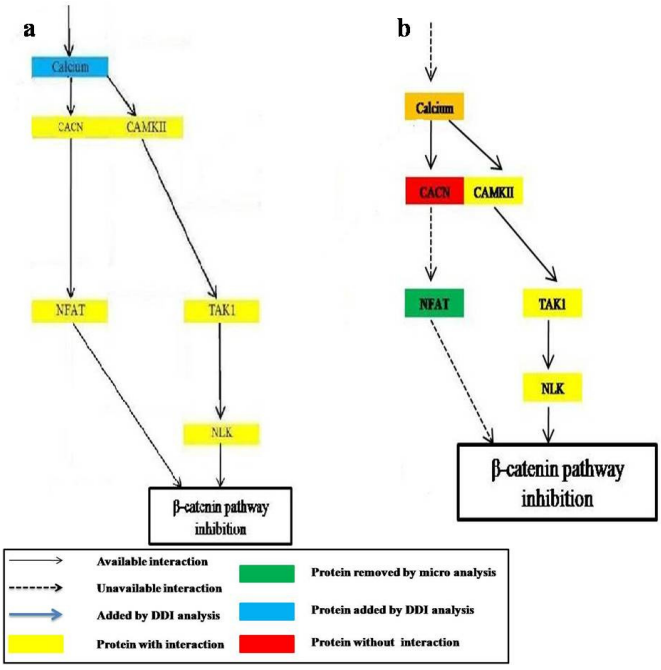

Figure 4.14. Inhibition part of the Wnt/Ca^{2+} signalling pathway a) in vertebrates b) in *C. elegans*

5. CONCLUSION and RECOMMENDATIONS

5.1. Conclusions

Wnt protein family is evolutionarily well conserved among animals. Wnt activated signalling pathways play important roles in embryonic development of many species ranging from human to worm. A very well established canonical pathway which activates β-catenin and two β-catenin independent pathways are known pathways that are activated by different Wnt proteins. Mutations in the canonical pathway have been reported to cause tumour formation. On the other hand, noncanonical Wnt/Ca^{2+} pathway is suggested to suppress tumour formation due to the antagonism between Wnt/Ca^{2+} pathway and canonical Wnt/β-catenin pathway. Therefore it is crucial to understand Wnt/Ca^{2+} pathway for developing proper disease treatments. As *C. elegans* is a simple model organism that possesses most of the Wnt signalling cascade proteins, it is convenient to study noncanonical Wnt signalling pathways in *C. elegans.* In order to find drug target candidates and to enlighten this antagonistic mechanism, in the present work the noncanonical Wnt signalling network was reconstructed in *C. elegans.* Both protein-protein interactions and domain-domain interactions were studied and analyzed. Further, the similarity and conservation of the proteins were investigated among different species. Finally, the topological properties of the network were analyzed to validate its biological significance.

Network reconstruction utilized 16 core proteins from 64 core proteins. Interaction data could not be obtained for the remaining 48 core proteins. Final network reconstructed incorporates 667 proteins with 922 interactions.

Sequence similarity of the *C. elegans* core proteins was found between 25 per cent and 50 per cent. Conservation percentage of *C. elegans* core proteins was found between 50 per cent and 75 per cent. Domains' similarity and conservation were found between 50 and 75 per cent in most cases. NFAT homologues and interactions were removed from the reconstructed network due to very low sequence similarity. Consequently, Wnt/Ca^{2+} components in *C. elegans* were identified, excluding NFAT, successfully.

Graph theory based topological analysis characterized the PPI network as scale-free with small world properties. The distribution of the network is close to power law model with γ=1.69 (R^2=0.91), mean path length=5.39 and network diameter=21. Obtained results are consistent with literature.

Four modules were detected in the network consisting of three, five, four and two proteins with three, seven, six and three interactions respectively. Extended version of modules associate 6, 20, 38 and 6 proteins with 6, 22, 40 and 7 interactions respectively.

Furthermore, 332 unique domains which cover 85 per cent of all proteins in the network and 2526 domain interactions which cover 81 per cent of interacting pairs of proteins were detected. Missing Frizzled-Dishevelled interaction was predicted from DDI analysis.

This study gives a general aspect of the noncanonical Wnt/Ca^{2+} signalling pathways in *C. elegans.* The network can be considered as a starting point for potential experimental studies to investigate candidate protein and domain interactions proposed.

5.2. Recommendations

Protein interaction data are missing for many components of the noncanonical signalling pathway not only in *C. elegans* but also in human and fly. Experimental studies validating suggested interactions could give insight to this signalling mechanism.

Reconstruction of the network depends on GO annotations. Existing GO annotations for *C. elegans* should be used with caution as nearly 50 per cent of the data is retrieved by computational analysis. *C. elegans* GO annotations quality should be improved by more experiments.

Micro scale analysis performed was based on sequence homology. Three dimensional structures of proteins could give better information about the conservation of noncanonical Wnt signalling components through different species.

APPENDIX A: GO ANNOTATION COLLECTION TABLES

This section gives the annotation collection tables utilized in SPA algorithm as selection criterion set. Evidence codes (Source) of the GO annotations listed are also included.

Table A.1. Evidence codes of the annotation collection set

Experimental Evidence Codes	
IDA	Inferred from Direct Assay
IGI	Inferred from Genetic Interaction
IMP	Inferred from Mutant Phenotype
IPI	Inferred from Physical Interaction
Computational Analysis Evidence Codes	
ISS	Inferred from Sequence or Structural Similarity
Author Statement Evidence Codes	
TAS	Traceable Author Statement
Automatically-assigned Evidence Codes	
IEA	Inferred from Electronic Annotation
ND	No biological Data available

Table A.2. Cellular component terms of the annotation collection set

GO ID	Cellular Component	Source
5954	calcium and calmodulin-dependent protein kinase complex	TAS
5952	cAMP-dependent protein kinase complex	IEA
5623	cell	IDA
5575	cellular component unknown	ND
5856	cytoskeleton	IEA
5665	DNA directed RNA polymerase II, core complex	IEA
5576	extracellular region	IEA
16021	integral to membrane	IDA
5622	intracellular	IEA
16020	membrane	IEA
5634	nucleus	IDA
16581	NuRD complex	ISS
5886	plasma membrane	IDA

Table A.3. Molecular function terms of the annotation collection set

GO ID	Molecular Function	Source
5524	ATP binding	IEA
5509	calcium ion binding	IEA
8603	cAMP-dependent protein kinase regulator activity	IEA
3824	catalytic activity	IEA
4692	cGMP-dependent protein kinase activity	IEA
19992	diacylglycerol binding	IEA
3677	DNA binding	IEA
3899	DNA-directed RNA polymerase activity	IEA
4930	G-protein coupled receptor activity	IEA
5525	GTP binding	IEA

Table A.3. Molecular function terms of the annotation collection set (continued)

GO ID	GO Function	Source
3924	GTPase activity	IDA
16787	hydrolase activity	IEA
16301	kinase activity	IMP
4707	MAP kinase activity	IEA
46872	metal ion binding	IEA
8168	methyltransferase activity	IEA
3674	molecular function unknown	ND
4926	nonG-protein coupled 7TM receptor activity	IEA
3676	nucleic acid binding	IEA
4435	phosphoinositide phospholipase C activity	IEA
4629	phospholipase C activity	IEA
4721	phosphoprotein phosphatase activity	IEA
5049	protein binding	IPI
6469	protein kinase activity	IEA
4697	protein kinase C activity	IEA
4674	protein serine/threonine kinase activity	IEA
4713	protein-tyrosine kinase activity	IEA
4872	receptor activity	IEA
5102	receptor binding	ISS
8757	S-adenosylmethionine-dependent methyltransferase activity	IEA
43565	sequence-specific DNA binding	IEA
4871	signal transducer activity	IEA

Table A.3. Molecular function terms of the annotation collection set (continued)

GO ID	GO Function	Source
16563	transcription activator activity	IMP
3700	transcription factor activity	IEA
16564	transcription repressor activity	IMP
16740	transferase activity	IEA
4888	transmembrane receptor activity	IEA
8270	zinc ion binding	IEA
4114	3',5'-cyclic-nucleotide phosphodiesterase activity	IEA

Table A.4. Biological process terms of the annotation collection set

GO ID	Biological Process	Source
186	activation of MAPKK activity	IGI
6915	apoptosis	IEA
8356	asymmetric cell division	IMP
45167	asymmetric protein localization involved in cell fate determination	IMP
48846	axon extension involved in axon guidance	IGI
8150	biological process unknown	IMP
10171	body morphogenesis	IMP
7155	cell adhesion	IMP
7049	cell cycle	IEA
51301	cell division	IEA
1708	cell fate specification	IMP
16477	cell migration	IMP
7166	cell surface receptor linked signal transduction	IMP
7635	chemosensory behavior	IMP
6935	chemotaxis	IMP
40004	collagen and cuticulin-based cuticle attachment to epithelium	IMP
40002	collagen and cuticulin-based cuticle development	IMP
40024	dauer larval development	IGI
30421	defecation	IMP
8340	determination of adult life span	IMP
7212	dopamine receptor signaling pathway	IMP
33563	dorsal/ventral axon guidance	IGI
40016	embryonic cleavage	IMP
9790	embryonic development	IMP
9792	embryonic development ending in birth or egg hatching	IMP
48598	embryonic morphogenesis	IDA

Table A.4. Biological process terms of the annotation collection set (continued)

GO ID	Biological Process	Source
7498	endoderm development	IMP
1706	endoderm formation	IMP
1714	endodermal cell fate specification	IMP
43652	engulfment of apoptotic cell	IMP
8544	epidermis development	IMP
7163	establishment and/or maintenance of cell polarity	IMP
30010	establishment of cell polarity	IMP
40001	establishment of mitotic spindle localization	IMP
16055	frizzled signaling pathway	IEA
7276	gamete generation	IMP
7369	gastrulation	IMP
51729	germline cell cycle switching, mitotic to meiotic cell cycle	IMP
40007	growth	IMP
40035	hermaphrodite genitalia development	IMP
6972	hyperosmotic response	IMP
6886	intracellular protein transport	IEA
7242	intracellular signaling cascade	IMP
6629	lipid metabolic process	IEA
40011	locomotion	IMP
7626	locomotory behavior	IMP
30539	male genitalia development	IMP
7498	mesoderm development	IMP
8078	mesodermal cell migration	IMP
7052	mitotic spindle organization and biogenesis	IEA
2009	morphogenesis of an epithelium	IMP
8045	motor axon guidance	IGI
7275	multicellular organismal development	IMP
30308	negative regulation of cell growth	IMP
40015	negative regulation of multicellular organism growth	IMP
46621	negative regulation of organ growth	IMP
40027	negative regulation of vulval development	IMP
2119	nematode larval development	IMP
48812	neurite morphogenesis	IMP
1764	neuron migration	IMP
7269	neurotransmitter secretion	IMP
6913	nucleocytoplasmic transport	IEA
42048	olfactory behavior	IMP

Table A.4. Biological process terms of the annotation collection set (continued)

GO ID	Biological Process	Source
18991	oviposition	IMP
6909	phagocytosis	IEA
6911	phagocytosis, engulfment	IMP
8150	physiological process unknown	IMP
9949	polarity specification of anterior/posterior axis	IMP
10085	polarity specification of proximal/distal axis	IMP
14057	positive regulation of acetylcholine secretion	IMP
40010	positive regulation of growth rate	IMP
40017	positive regulation of locomotion	IMP
40018	positive regulation of multicellular organism growth	IMP
45817	positive regulation of transcription from RNA polymerase II promoter	IMP
40032	post-embryonic body morphogenesis	IMP
12501	programmed cell death	IMP
35046	pronuclear migration	IMP
6468	protein amino acid phosphorylation	IEA
15031	protein transport	IEA
7265	Ras protein signal transduction	IGI
32970	regulation of actin filament-based process	IMP
9786	regulation of asymmetric cell division	IMP
42659	regulation of cell fate specification	IMP
30334	regulation of cell migration	IDA
7011	regulation of cytoskeleton	IMP
40012	regulation of locomotion	IMP
45664	regulation of neuron differentiation	IMP
46662	regulation of oviposition	IMP
50764	regulation of phagocytosis	IMP
43051	regulation of pharyngeal pumping	IMP
1932	regulation of protein amino acid phosphorylation	IEA
50807	regulation of synapse organization and biogenesis	IGI
43618	regulation of transcription from RNA polymerase II promoter	ISS
6355	regulation of transcription, DNA-dependent	IMP
40028	regulation of vulval development	IMP
3	reproduction	IMP
7165	signal transduction	IMP
7264	small GTPase mediated signal transduction	IEA
7051	spindle organization and biogenesis	IMP
45138	tail tip morphogenesis	IMP

Table A.4. Biological process terms of the annotation collection set (continued)

GO ID	Biological Process	Source
40040	thermosensory behavior	IMP
43052	thermotaxis	IMP
6350	transcription	IMP
6366	transcription from RNA polymerase II promoter	IEA
7179	transforming growth factor beta receptor signaling pathway	IMP
40025	vulval development	IMP
16055	Wnt receptor signaling pathway	IMP
7223	Wnt receptor signaling pathway, calcium modulating pathway	IEA

APPENDIX B: PROTEINS AND PROTEIN INTERACTIONS of NONCANONICAL WNT SIGNALING NETWORK in *C. ELEGANS*

Core proteins, protein interactions (as binary interaction) and all neighbours (steps of SPA algorithm also) are included in this section.

Table B.1. Core proteins justified according to general structure of the noncanonical Wnt signalling pathways.

General	Worm	
Name	Name	Orf name
WNT5A	cwn-2	W01B6.1
	cwn-1	K10B4.6
	egl-20	W08D2.1
	mom-2	F38E1.7
	lin-44	E01A2.3
FZD2	mom-5	T23D8.1
	cfz-2	F27E11.3
	lin-17	Y71F9B.5
	mig-1	Y34D9B.1
DVL1	dsh-1	C34F11.9
	mig-5	T05C12.6
	dsh-2	C27A2.6
	mpz-1	C52A11.4
DVL2	dsh-1	C34F11.9
	mig-5	T05C12.6
	dsh-2	C27A2.6
	mpz-1	C52A11.4
DVL3	dsh-1	C34F11.9
	mig-5	T05C12.6
	dsh-2	C27A2.6
	mpz-1	C52A11.4
PLCB1	egl-8	B0348.4
	plc-2	Y75B12B.6
	plc-4	R05G6.8
	pll-1	K10F12.3

Table B.1. Core proteins justified according to general structure of the noncanonical Wnt signalling pathways (continued)

General	Worm	
Name	Name	Orf name
PLCB2	egl-8	B0348.4
	plc-2	Y75B12B.6
	pll-1	K10F12.3
	plc-1	F31B12.1
	plc-3	T01E8.3
	plc-4	R05G6.8
PLCB3	egl-8	B0348.4
	plc-2	Y75B12B.6
	plc-3	T01E8.3
	plc-1	F31B12.1
	plc-4	R05G6.8
	pll-1	K10F12.3
PLCB4	egl-8	B0348.4
	plc-2	Y75B12B.6
	plc-1	F31B12.1
	plc-3	T01E8.3
	pll-1	K10F12.3
	plc-4	R05G6.8
CAMK2A	unc-43	K11E8.1
	cmk-1	K07A9.2
	ZC373.4	ZC373.4
CAMK2B	unc-43	K11E8.1
	cmk-1	K07A9.2
	ZC373.4	ZC373.4
CAMK2D	unc-43	K11E8.1
	cmk-1	K07A9.2
CAMK2G	unc-43	K11E8.1
	cmk-1	K07A9.2
	ZC373.4	ZC373.4
CHP	ZK856.8	ZK856.8
	Y71H2AL.1	Y71H2AL.1
	cnb-1	F55C10.1
	F30A10.1	F30A10.1
PPP3CA	tax-6	C02F4.2
	pph-4.2	Y49E10.3
PPP3CB	tax-6	C02F4.2
	pph-4.2	Y49E10.3

Table B.1. Core proteins justified according to general structure of the noncanonical Wnt signalling pathways (continued)

General	Worm	
Name	Name	Orf name
PPP3CC	tax-6	C02F4.2
	pph-4.2	Y49E10.3
PPP3R1	cnb-1	F55C10.1
	F59D6.7	F59D6.7
	ZK856.8	ZK856.8
	Y71H2AL.1	Y71H2AL.1
	F30A10.1	F30A10.1
PPP3R2	cnb-1	F55C10.1
	F59D6.7	F59D6.7
PRKCA	pkc-2	E01H11.1
	pkc-1	F57F5.5
PRKCB	pkc-2	E01H11.1
	pkc-1	F57F5.5
PRKCG	pkc-2	E01H11.1
	pkc-1	F57F5.5
NFAT5	pqn-15	C24A8.3
	pqn-20	C37A2.2
	dpy-22	F47A4.2
	K06A9.1	K06A9.1
	sma-9	T05A10.1
	egl-27	C04A2.3
	sel-8	C32A3.1
	F22H10.2	F22H10.2
	Y39B6A.1	Y39B6A.1
NFATC1	K06A9.1	K06A9.1
NFATC2	ama-1	F36A4.7
NFATC3	ama-1	F36A4.7
	K06A9.1	K06A9.1
TAK1	mom-4	F52F12.3
	Y105C5A.24	Y105C5A.24
	dlk-1	F33E2.2
	C24A1.3	C24A1.3
	kin-26	T06C10.6
NLK	lit-1	W06F12.1
	mpk-1	F43C1.2

Table B.1. Core proteins justified according to general structure of the noncanonical Wnt signalling pathways (continued)

General	Worm	
Name	Name	Orf name
PDE6	pde-2	R08D7.6
	pde-1	T04D3.3
	pde-4	R153.1
	pde-6	Y95B8A.10
	pde-3	E01F3.1
PKG	egl-4	F55A8.2
	C09G4.2	C09G4.2
	kin-1	ZK909.2
CDC42	cdc-42	R07G3.1
	ced-10	C09G12.8
	rac-2	K03D3.10
	mig-2	C35C5.4
	rho-1	Y51H4A.3
	chw-1	F22E12.2
	crp-1	Y32F6B.3
	rap-3	C08F8.7

Table B.2. Protein interactions

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
AC7.2	ZK792.6	F42G8.3	B0547.1	T05C12.6	T05C12.6
B0024.12	C38D4.6	F42H10.7	W05H7.4	T05C12.6	K04D7.2
B0024.12	B0365.1	F42H10.7	C27A2.6	T05C12.6	K03H6.2
B0024.12	C25F6.2	F42H10.7	ZK1098.4	T05C12.6	H15N14.2
B0207.4	C14F11.6	F42H10.7	T04D1.3	T05C12.6	F56D1.2
B0207.4	F41C3.5	F42H10.7	W04D2.1	T05C12.6	F55G1.8
B0218.3	C38D4.6	F43C1.2	Y17G7B.4	T05C12.6	F54F2.5
B0218.3	B0547.1	F43C1.2	W05H7.4	T05C12.6	F37A4.9
B0218.3	B0336.7	F43C1.2	T23H4.2	T05C12.6	F29G9.3
B0218.3	K11E8.1	F43C1.2	T02E1.3	T05C12.6	EEED8.7
B0218.3	T12G3.1	F43C1.2	T01H8.1	T05C12.6	C36E6.5
B0218.3	K04G7.3	F43C1.2	F32D1.1	T05C12.6	C34F11.9
B0218.3	C15C8.2	F43C1.2	C56G7.1	T05C12.6	C27A2.6
B0218.3	C44C8.6	F43C1.2	R07E5.8	T05C12.6	C23G10.3
B0218.3	C54G4.1	F43C1.2	F08C6.7	T05C12.6	C06G3.6
B0218.3	F32B6.1	F43C1.2	B0547.1	T05C12.6	C06A5.9
B0218.3	Y46G5A.31	F43C1.2	ZC477.9	T05C12.6	B0336.7
B0280.8	M03D4.1	F43C1.2	Y42H9AR.1	T05G5.3	C38D4.6
B0280.8	B0280.8	F43C1.2	F29G9.3	T07D4.2	C06A5.9
B0280.8	C27B7.4	F43C1.2	Y54E10BL.6	T07D4.2	F32G8.6
B0464.5	B0035.1	F43C1.2	C45E1.1	T07D4.2	T18D3.7

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
B0464.5	B0336.2	F43C1.2	T27F2.2	T09A5.2	R03D7.7
B0464.5	C47E8.5	F43C1.2	T27F2.1	T11B7.4	Y40C5A.1
B0464.5	F26H11.2	F43C1.2	F14F3.2	T11B7.4	W05H7.4
B0464.5	F46F11.2	F43C1.2	T08D10.1	T11B7.4	T11B7.4
B0496.7	B0547.1	F43C1.2	F42H10.7	T11B7.4	T27F2.2
B0496.7	R06F6.4	F43C1.2	T05C12.6	T11B7.4	F54F2.5
B0496.7	C09H6.2	F43C1.2	B0464.9	T11B7.4	C34F11.9
C01B10.8	C17E4.5	F43C1.2	K11E8.1	T11B7.4	T09A5.12
C01B10.8	C09H5.6	F43G9.11	C06A5.9	T11B7.4	T05C12.6
C01B10.8	C14B9.8	F43G9.11	ZK849.1	T11B7.4	K01A6.2
C01B10.8	Y113G7B.23	F43G9.11	W02D3.9	T11B7.4	Y53C12B.3
C01G6.4	Y104H12A.1	F43H9.2	T18D3.7	T11B7.4	AH6.5
C02F5.9	ZK1098.4	F44B9.6	T05C12.6	T11B7.4	T07F8.3
C02F5.9	C06G3.6	F44B9.6	R11A8.6	T11B7.4	C09G1.4
C02F5.9	D1005.1	F44B9.6	F45G2.3	T11B7.4	F59E12.9
C02F5.9	F44G3.9	F44B9.6	K12G11.3	T11B7.4	F54B11.3
C03C10.3	C03C10.3	F44B9.6	R11E3.6	T11B7.4	C14B9.6
C04F12.3	F37A4.9	F44B9.6	T27C4.4	T11B7.4	F13E6.4
C04F12.3	B0547.1	F44B9.6	W04D2.1	T11B7.4	C25A11.4
C04F12.3	F54G8.4	F44B9.9	C38D4.6	T11B7.4	F31E3.5
C04F12.3	ZK455.1	F44C4.3	T28F12.2	T11B7.4	F28F5.3
C04F12.3	F25B5.7	F44G3.9	ZK112.2	T11B7.4	T13C5.4
C04F12.3	T14F9.1	F44G3.9	Y119C1A.1	T11B7.4	F28C6.2
C04F12.3	C04F12.3	F44G3.9	C18D1.1	T11B7.4	W02D3.11
C04F12.3	Y39B6A.1	F44G3.9	K08E3.7	T11B7.4	F54E7.3
C05C10.5	C27A2.6	F44G3.9	B0464.5	T11B7.4	C47E8.5
C05C8.6	Y116F11B.12	F44G3.9	ZK652.6	T11B7.4	F46F11.2
C05C8.6	K04A8.6	F44G3.9	F47B10.1	T11B7.4	W04D2.1
C05C8.6	C05C8.6	F44G3.9	F25H9.2	T11B7.4	ZK863.7
C05C8.6	F37B1.1	F44G3.9	F18E2.2	T11B7.4	W03F11.6
C05C8.6	C27B7.4	F44G3.9	Y65B4BR.4	T11B7.4	F44D12.1
C05D11.10	F49H12.3	F44G3.9	B0350.2	T11B7.4	D1046.1
C05D2.1	C34E10.6	F44G3.9	F49H12.3	T11B7.4	C27B7.4
C05D2.1	C47E8.5	F44G3.9	F54C4.2	T11B7.4	D2005.3
C05D9.1	F45H10.4	F44G3.9	Y48E1A.1	T11B7.4	T10F2.1
C05D9.1	F25D7.1	F44G3.9	C04F5.8	T11B7.4	Y73B6BL.33
C05D9.1	Y75B8A.35	F44G3.9	F26D10.3	T11B7.4	ZK121.2
C05D9.1	Y59A8B.22	F44G3.9	Y111B2A.22	T11B7.4	ZK287.6
C07A12.3	Y38A10A.5	F44G3.9	F44G3.9	T12D8.7	T23H4.2
C07A12.3	F01F1.12	F44G4.4	C25A1.4	T12D8.7	Y119C1A.1
C07F11.1	C48D5.1	F44G4.4	F44G4.4	T12D8.7	C26B2.3
C07F11.1	T25C8.2	F46F2.2	T04A11.6	T12D8.7	T07C4.1
C07H6.5	W05H7.4	F46F2.2	F46F2.2	T12D8.7	Y104H12A.1
C07H6.5	C27A2.6	F46G11.1	C06G3.6	T12D8.7	Y46G5A.31
C07H6.5	C09F5.2	F46G11.1	Y119C1A.1	T12D8.7	Y56A3A.4
C07H6.5	F07C4.5	F46G11.1	F01F1.4	T14G12.4	Y73C8B.3
C07H6.5	F46A9.4	F46G11.1	F43G6.8	T14G12.4	Y75B8A.1

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
C07H6.5	F46H5.3	F46G11.3	C38D4.6	T20B12.8	C38D4.6
C07H6.5	W03C9.7	F46G11.3	F08F3.3	T20H4.5	C06A5.9
C07H6.5	ZK1025.9	F47D12.4	C37A2.2	T20H4.5	F53G12.5
C08B11.1	C38D4.6	F47D12.4	Y66D12A.5	T22B2.4	F46A9.6
C08B11.1	K04F10.6	F47D12.4	F18A1.3	T22B2.4	W02A11.3
C08F8.8	C38D4.6	F47D12.4	C27A12.2	T22B2.4	T01D1.2
C08F8.8	ZK1098.4	F47D12.4	F52B10.1	T22B2.4	T21G5.5
C08F8.8	C32F10.6	F47D12.4	T03G11.1	T22B2.4	R74.5
C08F8.8	C47E8.5	F47G6.1	C06G3.6	T22C8.3	Y49E10.23
C08F8.8	F49H12.3	F47G6.1	C34B2.4	T22C8.3	ZK909.4
C08F8.8	H19N07.1	F47G6.1	W07G1.3	T22D1.12	T18D3.7
C08F8.8	W02D3.9	F47G6.1	Y56A3A.13	T22D1.12	F32G8.6
C09G4.5	C39E9.13	F47G6.1	F57B10.3	T23C6.5	T06C10.3
C09H6.2	C09H6.2	F47G6.1	B0365.1	T23C6.5	C32F10.6
C09H6.2	F17E5.1	F47G6.1	W04D2.1	T23H4.2	T23H4.2
C12D8.1	ZK849.1	F47G6.1	F23F1.8	T23H4.2	C27B7.4
C12D8.1	AH6.5	F47G6.1	F44D12.1	T24D1.3	M7.5
C12D8.1	F59E12.9	F47G6.1	Y48C3A.17	T25C8.2	Y75B8A.2
C12D8.1	C27B7.4	F47G6.1	Y59A8B.22	T25C8.2	Y37D8A.10
C12D8.1	D1046.1	F48E8.5	C38D4.6	T25C8.2	B0207.4
C12D8.1	F32A11.6	F52F12.3	C48D5.1	T26E3.3	F53B3.1
C12D8.1	Y113G7B.23	F52F12.3	VW06B3R.1	T26E3.3	ZK849.1
C12D8.1	ZC513.6	F53F4.10	C38D4.6	T27F2.1	C27A2.6
C13B9.3	Y104H12A.1	F56A8.6	W05H7.4	T27F2.3	ZK1098.4
C13F10.7	C13F10.7	F56A8.6	F25B5.7	T27F2.3	B0547.1
C18D11.4	F46F11.2	F57B10.11	F26D10.3	T27F2.3	B0280.8
C18H9.7	Y119C1A.1	F57B10.11	C32F10.6	T27F2.3	Y119C1B.5
C18H9.7	Y40C5A.1	F57F5.5	C54E10.2	T27F2.3	F55A11.3
C18H9.7	Y42H9AR.1	F58A4.4	Y106G6H.2	T27F2.3	F23H11.1
C18H9.7	ZK1098.4	F58A4.8	C38D4.6	T27F2.3	F43C11.7
C18H9.7	E04F6.6	F58A4.8	C36E8.5	T27F2.3	F44G3.9
C18H9.7	C06A6.2	F58A4.8	C45G3.3	T27F2.3	H06H21.3
C18H9.7	C07A12.1	F58A4.8	F57B10.10	T27F2.3	T09A12.4
C18H9.7	C25B8.3	F58A4.8	F57B10.7	T27F2.3	Y37E11AR.2
C18H9.7	K08E5.3	F58A4.8	R06C7.10	T27F2.3	ZK1240.2
C18H9.7	R04B5.4	F59A2.4	B0464.5	T28A11.11	Y40D12A.2
C18H9.7	Y37E11AR.2	F59A2.4	C02F5.9	T28A11.11	F33C8.1
C18H9.7	Y45G5AM.1	F59A6.1	C38D4.6	T28A11.11	F23F12.6
C18H9.7	Y52D3.1	F59A6.1	F31E3.5	T28A11.11	Y53C12A.1
C23H3.3	F29G9.3	F59A6.1	H28O16.1	T28A11.11	F35G2.2
C23H3.3	C37C3.6	F59B2.3	T03F6.3	W01D2.2	F32B6.1
C27B7.1	B0336.7	F59B2.3	F25H5.3	W02B12.8	F18E2.1
C27B7.1	T05E7.1	F59E12.2	C38D4.6	W03D8.8	C27B7.4
C27B7.1	T04C12.5	F59E12.2	D1081.2	W03D8.8	D1046.1
C27B7.1	B0412.1	F59E12.2	H28O16.1	W04D2.1	C06G3.6
C28H8.6	T11B7.4	F59E12.2	Y106G6H.2	W04D2.1	C54E10.2
C28H8.6	Y105E8A.6	F59E12.2	ZC155.1	W04D2.1	F32D1.1

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
C28H8.9	C34F11.9	H42K12.1	C10F3.5	W04D2.1	C40A11.7
C28H8.9	T05C12.6	H42K12.1	Y65B4BR.4	W04D2.1	ZK970.4
C28H8.9	C44C10.4	K01G5.1	ZK1127.7	W04D2.1	ZK596.2
C28H8.9	W07B8.5	K01G5.1	R13H8.1	W04D2.1	F37B1.3
C28H8.9	Y39B6A.1	K01G5.1	R13D7.7	W04D2.1	B0495.9
C29F9.5	C38D4.6	K01G5.1	R11F4.1	W04D2.1	T04C10.1
C34B2.4	C38D4.6	K01G5.1	K12D12.1	W04D2.1	K02A4.2
C34B2.4	C08C3.3	K01G5.1	K09A11.3	W04D2.1	C27B7.4
C34B7.3	ZK849.1	K01G5.1	F46B6.7	W04D2.1	F55A4.1
C34E10.6	F13D12.6	K01G5.1	F11E6.5	W04D2.1	F58E6.3
C34E10.6	Y113G7A.6	K01G5.1	F07A5.1	W04D2.1	W04D2.1
C34E10.6	F38H4.8	K01G5.1	C08G9.2	W04D2.1	Y106G6D.7
C35D10.2	C32F10.6	K01G5.1	C03G6.3	W04D2.1	Y60A3A.1
C35D10.2	Y75B8A.1	K01G5.1	C32D5.12	W04D2.1	Y73B6BL.33
C36B1.4	C48D5.1	K01G5.1	T27C4.4	W04D2.1	ZC239.15
C36B1.4	H15N14.2	K01G5.4	F26E4.8	W04D2.1	ZC395.8
C36B1.4	D1054.2	K01G5.4	C39F7.4	W06D4.6	F37A4.1
C36B1.4	F23F1.8	K01G5.4	B0019.2	W06D4.6	Y43C5A.6
C37C3.6	C47E8.5	K01G5.4	Y24F12A.2	W06D4.6	F13D12.6
C37C3.6	F44G3.9	K01G5.4	Y38A10A.5	W06D4.6	F25H5.4
C39E9.13	C54G10.2	K05C4.1	W04D2.1	W06F12.1	ZK1290.4
C39E9.14	Y57G11C.22	K05C4.1	D1054.2	W06F12.1	ZC449.6
C39E9.14	F39B2.11	K06B4.1	C47E8.5	W06F12.1	Y119C1A.1
C39E9.14	C01B12.1	K06B4.11	Y39B6A.1	W06F12.1	W05H7.4
C39E9.14	F01F1.12	K07A9.2	C38D4.6	W06F12.1	T09F3.1
C40A11.2	F01F1.12	K07C11.2	F37A4.9	W06F12.1	T02E1.3
C40A11.2	Y39B6A.1	K07C11.2	C38D4.6	W06F12.1	F53B3.1
C44B7.1	H15N14.2	K07C11.2	R05F9.1	W06F12.1	F23B12.5
C44B7.1	F56H1.4	K07C11.2	C05E11.3	W06F12.1	F18A1.3
C44H4.5	C38D4.6	K08B12.5	K01G5.4	W06F12.1	C25G4.4
C44H4.5	C48D5.1	K08B4.1	B0547.1	W07B8.5	K08F11.3
C44H4.5	F46H5.3	K08B4.1	T05A10.1	W08D2.1	F57B9.6
C44H4.5	W05B10.1	K08B4.1	F02A9.6	Y105C5B.13	C31H1.5
C47B2.4	Y38A8.2	K08B4.1	C05C10.5	Y105C5B.13	F20B10.1
C47E8.5	C47E8.5	K08B4.1	C06A6.2	Y105C5B.13	Y104H12A.1
C49G7.1	Y57G11C.9	K08B4.1	K11D2.2	Y105E8B.5	Y105E8B.5
C49G7.1	C02C6.1	K08E3.7	Y53H1A.2	Y105E8B.5	W07B8.5
C49G7.1	H14N18.1	K08E3.7	C36E8.5	Y110A7A.10	C38D4.6
C49G7.1	C05C8.7	K08E3.7	Y39B6A.1	Y110A7A.10	EEED8.5
C49G7.1	F54B11.5	K08F11.3	C03C10.3	Y110A7A.10	T21E8.1
C50C3.8	C50C3.8	K08F4.7	T18D3.7	Y110A7A.10	C34E10.6
C50C3.8	K08F8.4	K08F8.2	Y49A3A.1	Y110A7A.10	F37H8.3
C50C3.8	C27B7.4	K09B11.1	C38D4.6	Y113G7B.23	Y113G7B.23
C50F4.13	C38D4.6	K09B11.1	Y75B8A.1	Y113G7B.23	C17G10.5
C52B11.2	Y42H9AR.1	K10C3.6	ZK697.2	Y116A8C.26	T11B7.4
C52B11.2	Y38A10A.5	K10C3.6	T26H2.9	Y116A8C.26	K08F8.4
C52B11.2	C52B11.2	K10C3.6	K10C3.6	Y15E3A.1	T09A5.2

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
C54G10.2	F31E3.3	K10C3.6	Y38E10A.18	Y15E3A.1	C07A12.1
C56C10.8	Y75B8A.2	K10C3.6	C05G6.1	Y15E3A.1	F55H12.6
C56C10.8	C18D11.4	K10C3.6	C06G3.1	Y15E3A.1	ZK867.1
CD4.6	ZK1098.4	K10C3.6	ZK1037.5	Y18D10A.5	F55H12.6
CD4.6	D1054.2	K10C3.6	C27C7.4	Y18D10A.5	C01G12.1
D1022.8	C31H1.6	K10C3.6	F54A5.1	Y18D10A.5	T12G3.1
D1046.1	D1046.1	K10C3.6	K11E4.5	Y18D10A.8	R186.5
D1053.1	D1053.1	K10C3.6	Y80D3A.4	Y18D10A.8	F44A2.5
D1054.2	H15N14.2	K10C3.6	Y5H2B.2	Y18D10A.8	F46G10.1
D1054.2	Y42H9AR.1	K10C3.6	T09A12.4	Y2H9A.1	F48E8.1
D1054.2	F56H1.4	K10C3.6	K06A1.4	Y2H9A.1	ZK632.6
D1054.2	D1054.2	K11D2.3	T05C12.6	Y2H9A.1	ZC513.6
E01H11.1	Y53F4B.33	K11D2.3	F23F12.9	Y2H9A.1	Y39B6A.1
E01H11.1	T26E3.3	K11D2.3	C43E11.4	Y37A1B.13	C38D4.6
E01H11.1	F54A3.4	K11D9.1	T11B7.4	Y37D8A.10	T07C4.1
E02H1.7	K10C3.6	K11D9.1	T01G1.1	Y38A8.2	C05D9.1
E02H1.7	K08H2.8	K11D9.1	W10D9.4	Y38A8.2	W04D2.1
E02H1.7	Y37E11AR.2	K11D9.1	Y37E11AR.2	Y39E4B.1	Y15E3A.1
F01D4.4	K08F11.3	M01E11.6	F58D5.1	Y43C5A.6	Y116A8C.13
F01F1.12	F01F1.12	M01E11.6	W07B8.5	Y43C5A.6	Y43C5A.6
F01F1.4	C06G3.6	M01H9.2	ZK637.5	Y43C5B.2	C32F10.6
F01F1.4	F11H8.1	M02A10.3	C38D4.6	Y47D7A.1	C35D10.2
F01F1.4	F44G3.9	M02A10.3	T23F6.4	Y48G1C.1	R151.2
F08C6.6	K08F11.3	M02A10.3	M7.5	Y48G1C.1	Y37E11AR.2
F08C6.6	ZK637.5	M02A10.3	F23C8.4	Y49E10.1	C48D5.1
F08F3.2	Y104H12A.1	M02A10.3	Y105E8B.5	Y49E10.1	Y113G7A.6
F09E5.1	T26E3.3	M03D4.1	K11D12.9	Y49F6C.3	F58A4.3
F10B5.5	C38D4.6	M03D4.1	C27A12.7	Y49F6C.3	B0272.4
F11G11.2	Y75B8A.1	M04F3.1	T20G5.1	Y49F6C.3	C45E5.6
F12F6.5	C46E10.9	M04F3.1	F18A1.5	Y51H4A.17	T11B7.4
F12F6.5	C26B2.3	M04F3.1	F44G3.9	Y51H4A.17	B0547.1
F14D12.2	Y105E8A.6	M18.5	F54F2.5	Y51H4A.17	C38D4.6
F14D12.2	Y39B6A.1	PAR2.1	C06G3.6	Y51H4A.17	F43E2.8
F14F3.1	C38D4.6	PAR2.1	K06A1.4	Y51H4A.17	T22G5.5
F14F3.1	F08C6.7	R03G5.2	ZK1098.4	Y51H4A.17	C06A6.2
F14F3.1	Y119C1A.1	R03G5.2	C38D4.6	Y51H4A.17	ZC155.1
F14F3.1	Y40C5A.1	R03G5.2	C48D5.1	Y51H4A.17	T08G11.4
F14F3.1	F38H4.9	R03G5.2	F25H2.9	Y51H4A.17	Y51H4A.17
F14F3.1	T07F8.3	R03G5.2	Y105C5B.28	Y53C12A.1	Y87G2A.14
F14F3.1	C09G1.4	R06B10.4	Y59A8B.22	Y53F4B.33	T23G5.1
F14F3.1	C07G1.5	R06B10.4	K02F2.2	Y53F4B.33	F54B11.6
F14F3.1	C27B7.4	R06B10.4	T20G5.1	Y54E10A.6	K02F2.2
F14F3.1	C52B11.2	R06B10.4	F38H4.9	Y54E10BL.6	T11B7.4
F14F3.1	D1046.1	R06B10.4	W04D2.1	Y54G11A.10	F39C12.2
F14F3.1	F26G5.9	R06B10.4	C06A6.2	Y55F3AM.10	T05G5.6
F14F3.1	F44G3.9	R07B1.4	Y42H9AR.1	Y55F3AM.10	T04D1.3
F14F3.1	Y5H2B.2	R07E4.6	C56E10.4	Y55F3AM.15	C38D4.6

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
F14F3.1	Y65B4BR.4	R07E4.6	F21F12.1	Y55F3AM.15	Y62E10A.16
F14F3.1	ZK121.2	R07E4.6	T04F3.1	Y55F3AM.15	K08F11.3
F17A2.5	Y113G7B.23	R07E4.6	Y40D12A.2	Y56A3A.20	C38D4.6
F17A2.5	T28F12.2	R07E4.6	ZK662.3	Y56A3A.20	T12D8.2
F17A2.5	C52B11.2	R07E4.6	F38E9.1	Y56A3A.21	C23G10.3
F17E5.1	C06G3.6	R07E4.6	K12G11.3	Y56A3A.21	F44G3.9
F17E5.1	M04B2.1	R07E5.8	Y47D3A.4	Y62E10A.16	Y62E10A.16
F17E5.1	C44F1.2	R07E5.8	K05G3.3	Y62E10A.16	T28F12.2
F17E5.1	F38A6.3	R07E5.8	C47E8.5	Y62E10A.16	T10C6.11
F17E5.1	R119.7	R107.7	T28A11.11	Y62E10A.16	F25H5.3
F17E5.1	Y54G11A.10	R10E11.2	ZK1058.2	Y62E10A.16	C27B7.4
F17E5.1	ZK652.9	R10E11.2	F31E3.5	Y73B6A.5	C38D4.6
F18C5.2	B0207.4	R11G1.3	C38D4.6	Y73B6A.5	H21P03.3
F19B6.1	F59E12.4	R12B2.1	Y48E1B.3	Y73B6A.5	K02F2.2
F22E12.2	C25G4.4	R12B2.1	R74.5	Y73B6A.5	F23H12.2
F23C8.4	T09A5.12	R12B2.1	F58G1.3	Y73B6A.5	F47F6.1
F23C8.4	C06A1.1	R12B2.1	D1037.4	Y73B6A.5	K07A12.2
F23C8.4	C41C4.8	R12B2.1	C54G7.4	Y73B6A.5	M01F1.2
F23C8.4	D1053.1	R12B2.1	C25F6.3	Y79H2A.11	B0336.7
F23C8.4	F44G3.9	R12B2.1	C04H5.6	Y79H2A.11	T05C12.6
F23C8.4	Y105E8B.8	R12B2.1	B0350.2	Y79H2A.11	C18D11.4
F23F1.8	F23F1.8	R12B2.1	K02B9.4	Y79H2A.11	C44B9.4
F23F1.8	F56H1.4	R12B2.1	B0281.5	Y79H2A.11	R05F9.1
F23F1.8	F31E3.5	R12B2.1	C47E8.5	Y79H2A.11	T02C5.1
F23F1.8	D1054.2	R13F6.9	C38D4.6	Y79H2A.11	W08E3.3
F25E2.5	Y113G7B.23	R13F6.9	R05F9.1	Y79H2A.11	Y24D9A.8
F25E2.5	F25H5.4	R13F6.9	D2096.8	Y79H2A.11	Y48B6A.14
F25E2.5	C25A1.5	R13F6.9	F01F1.12	Y81G3A.3	Y75B8A.1
F25E2.5	F09F3.9	R13F6.9	F25E2.5	Y92C3B.2	W05H7.4
F25E2.5	B0336.2	R13F6.9	Y116A8C.12	Y92C3B.2	T20G5.1
F25E2.5	ZK353.6	R144.1	C38D4.6	Y92C3B.2	C47E8.5
F25H2.9	H28O16.1	T01C8.1	C48D5.1	Y92C3B.2	Y116A8C.35
F25H2.9	C36B1.4	T01C8.1	C32F10.6	ZC434.8	C23G10.3
F25H5.3	F44G3.9	T01C8.1	C56C10.8	ZC434.8	ZK637.5
F26B1.2	T21G5.5	T01C8.1	F58E6.10	ZC434.8	C47E8.5
F26B1.2	F46A9.6	T01G9.5	T01G9.5	ZC434.8	F33D11.10
F26B1.2	Y49E10.14	T01G9.5	C06G8.4	ZC477.9	Y105E8A.6
F26B1.2	F59E12.4	T01G9.5	T10F2.1	ZC504.4	B0547.1
F26B1.2	C25A1.4	T01G9.5	ZK858.4	ZC504.4	C38D4.6
F26B1.2	Y48B6A.3	T03F6.1	C04C3.3	ZC504.4	F22B7.5
F26B1.2	Y59A8B.10	T04A11.6	K12D12.1	ZC504.4	C14F5.1
F26D11.11	Y113G7B.23	T04H1.2	Y40C5A.1	ZC504.4	F26D10.3
F26D11.11	C52B11.2	T04H1.2	T04H1.2	ZC504.4	F36A2.10
F26D11.11	F44G3.9	T04H1.2	K11D2.3	ZC504.4	K02F2.2
F26D11.11	Y37E11AR.2	T04H1.2	F53B3.1	ZC504.4	Y113G7A.6
F26F4.10	C38D4.6	T04H1.2	F32D1.1	ZC504.4	Y92C3B.2
F28F8.6	C41C4.8	T04H1.2	C27A2.6	ZC581.1	C26B2.3

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
F28H6.1	C38D4.6	T04H1.2	C06G3.6	ZK1307.8	Y39B6A.1
F28H6.1	W09C3.6	T04H1.2	C06A5.9	ZK1307.8	Y66A7A.6
F28H6.1	Y38F2AR.7	T04H1.2	ZK112.2	ZK353.6	ZK353.6
F29B9.4	Y75B8A.1	T04H1.2	F42H10.7	ZK353.6	ZK632.10
F29B9.4	C52B11.2	T04H1.2	ZK1098.4	ZK370.2	B0547.1
F29C4.1	C47E8.5	T04H1.2	C34F11.9	ZK370.2	R08E3.4
F29G9.5	C48D5.1	T04H1.2	T09A5.12	ZK632.12	Y67D2.6
F31C3.2	F23B12.5	T04H1.2	T05C12.6	ZK632.12	F54C1.7
F31C3.2	ZK849.1	T04H1.2	F33A8.1	ZK632.12	F37B1.1
F31C3.2	F42D1.2	T04H1.2	K08F8.2	ZK632.12	F42F12.3
F31C3.2	F32B4.4	T04H1.2	D2045.8	ZK632.12	D2045.8
F31C3.2	Y37E11AR.2	T04H1.2	T16H12.4	ZK632.12	C13F10.7
F31C3.2	F20G4.3	T04H1.2	R10E11.8	ZK632.12	R05F9.1
F31C3.2	ZK858.4	T04H1.2	B0464.5	ZK632.12	Y39B6A.1
F31E3.2	C06A5.9	T04H1.2	F45E6.2	ZK637.11	T23G7.1
F31E3.2	T19D12.5	T04H1.2	ZC504.4	ZK637.11	Y113G7A.6
F31E3.2	F59B1.7	T04H1.2	F31E3.5	ZK662.4	T11B7.4
F31E3.2	H14N18.1	T04H1.2	C47E8.4	ZK673.7	Y39B6A.1
F31E3.2	T28F12.2	T04H1.2	C27A12.2	ZK673.7	M03F4.2
F31E3.3	Y57G11C.22	T04H1.2	C06A6.2	ZK792.6	H09G03.2
F31E3.3	C39E9.13	T04H1.2	C13F10.7	ZK792.6	C56C10.8
F31E3.3	F32A11.2	T04H1.2	C30B5.4	ZK849.1	ZK849.1
F32A11.2	C47E8.5	T04H1.2	C31H1.6	ZK849.1	T26C12.3
F32D1.1	F32D1.1	T04H1.2	C43E11.6	ZK849.1	Y40C5A.1
F35C8.3	C38D4.6	T04H1.2	C52B11.2	ZK849.1	F18A1.3
F35C8.3	EEED8.16	T04H1.2	D1005.3	ZK849.1	F38H4.9
F35C8.3	Y51F10.2	T04H1.2	F23H11.5	ZK849.1	T07G12.10
F35C8.3	F43G6.8	T04H1.2	F26G5.9	ZK849.1	C07D10.2
F35G12.9	B0547.1	T04H1.2	F44D12.1	ZK849.1	F30F8.3
F35G12.9	F35G2.2	T04H1.2	F44G3.9	ZK849.1	F28F5.3
F35G12.9	D1046.1	T04H1.2	F46G10.1	ZK849.1	C45E5.6
F35G12.9	C52B11.2	T04H1.2	F47G6.1	ZK849.1	R11E3.6
F35G12.9	H10E21.3	T04H1.2	F49H12.6	ZK849.1	C05D12.5
F35G12.9	H10E21.4	T04H1.2	F57C9.4	ZK849.1	Y37E11AR.2
F35G2.2	H14A12.2	T04H1.2	H22K11.1	ZK849.1	F08C6.1
F35G2.2	F35G2.2	T04H1.2	W03F9.10	ZK849.1	D1005.1
F37B1.4	C52B11.2	T04H1.2	Y45G5AM.1	ZK849.1	F44G3.9
F37B1.8	F44G3.9	T04H1.2	ZC239.15	ZK849.1	T08G11.4
F38A6.1	Y113G7B.23	T04H1.2	ZC395.8	ZK849.1	Y51H1A.3
F38A6.1	T07D1.4	T04H1.2	ZK1240.9	ZK849.1	ZC239.15
F38A6.1	M01F1.2	T04H1.2	Y39B6A.1	ZK858.4	C38D4.6
F38A6.1	R74.5	T05C12.6	ZK849.1	ZK858.4	C27B7.4
F38A6.1	K08F8.2	T05C12.6	ZK112.2	ZK858.4	C49H3.5
F38A6.1	W04D2.1	T05C12.6	ZK1098.4	ZK858.4	R08E3.4
F38A6.1	D1046.1	T05C12.6	Y66D12A.5	ZK858.4	ZK858.4
F38A6.1	F26G5.9	T05C12.6	Y57G11C.9	ZK892.7	W02B12.8
F39C12.2	F08F3.4	T05C12.6	Y40C5A.1	ZK892.7	C04C3.3

Table B.2. Protein interactions (continued)

Interactor A	Interactor B	Interactor A	Interactor B	Interactor A	Interactor B
F39C12.2	R05D3.7	T05C12.6	Y32F6B.3	ZK892.7	C52B11.2
F42G10.2	T07A9.3	T05C12.6	T26C12.3	ZK892.7	Y46G5A.31
F42G10.2	B0478.1	T05C12.6	T09A5.2	ZK892.7	F44G3.9
F42G10.2	Y5H2B.2	T05C12.6	T09A5.12	ZK909.4	Y51H4A.4
				ZK909.4	K08F8.2

Table B.3. First neighbours

1st neighbours			
Name	**Name**	**Name**	**Name**
B0218.3	F02A9.6	K04D7.2	W05H7.4
B0336.7	F08C6.7	K06B4.11	Y119C1A.1
B0464.9	F14D12.2	K08B4.1	Y17G7B.4
B0547.1	F14F3.2	K08E3.7	Y2H9A.1
C04F12.3	F18A1.3	K11D2.3	Y40C5A.1
C05C10.5	F23B12.5	R07E5.8	Y42H9AR.1
C06A5.9	F29G9.3	T01H8.1	Y53F4B.33
C06G3.6	F32D1.1	T02E1.3	Y54E10BL.6
C07H6.5	F37A4.9	T04H1.2	Y57G11C.9
C23G10.3	F42H10.7	T08D10.1	Y66D12A.5
C25G4.4	F44B9.6	T09A5.12	Y79H2A.11
C28H8.9	F47D12.4	T09A5.2	ZC449.6
C36E6.5	F53B3.1	T09F3.1	ZC477.9
C38D4.6	F54A3.4	T11B7.4	ZK1098.4
C40A11.2	F54F2.5	T23H4.2	ZK112.2
C45E1.1	F55G1.8	T26C12.3	ZK1290.4
C48D5.1	F56D1.2	T26E3.3	ZK1307.8
C54E10.2	F57B9.6	T27F2.1	ZK632.12
C56G7.1	H15N14.2	T27F2.2	ZK673.7
EEED8.7	K03H6.2	VW06B3R.1	ZK849.1

Table B.4. Second neighbours

2nd neighbours					
Name	**Name**	**Name**	**Name**	**Name**	**Name**
AH6.5	C44H4.5	F32B6.1	K01A6.2	T20B12.8	ZC434.8
B0024.12	C45E5.6	F33A8.1	K04G7.3	T20H4.5	ZC504.4
B0464.5	C47E8.4	F35C8.3	K05G3.3	T23G5.1	ZC513.6
B0496.7	C47E8.5	F35G12.9	K07C11.2	T27C4.4	ZK1025.9
C02F5.9	C49G7.1	F37B1.1	K08F8.2	T27F2.3	ZK121.2
C05D12.5	C50F4.13	F38H4.9	K09B11.1	W02D3.11	ZK1240.9
C06A6.2	C52B11.2	F42F12.3	K11D2.2	W03C9.7	ZK287.6
C07D10.2	C54G4.1	F42G8.3	K11D9.1	W03F11.6	ZK370.2
C07F11.1	CD4.6	F43G9.11	K12G11.3	W03F9.10	ZK455.1
C08B11.1	D1005.1	F44B9.9	M02A10.3	W04D2.1	ZK632.6
C08F8.8	D1005.3	F44D12.1	M03F4.2	W07B8.5	ZK662.4
C09F5.2	D1046.1	F44G3.9	M18.5	W08E3.3	ZK858.4
C09G1.4	D1054.2	F45E6.2	PAR2.1	Y105E8A.6	ZK863.7
C12D8.1	D2005.3	F45G2.3	R03D7.7	Y110A7A.10	
C13F10.7	D2045.8	F46A9.4	R03G5.2	Y116A8C.26	
C14B9.6	F01F1.12	F46F11.2	R05F9.1	Y15E3A.1	
C15C8.2	F01F1.4	F46G10.1	R07B1.4	Y24D9A.8	
C18D11.4	F07C4.5	F46G11.1	R10E11.8	Y37A1B.13	
C18H9.7	F08C6.1	F46G11.3	R11A8.6	Y37E11AR.2	
C23H3.3	F09E5.1	F46H5.3	R11E3.6	Y45G5AM.1	
C25A11.4	F10B5.5	F47G6.1	R11G1.3	Y46G5A.31	
C27A12.2	F13E6.4	F48E8.1	R13F6.9	Y47D3A.4	
C27B7.1	F14F3.1	F48E8.5	R144.1	Y48B6A.14	
C27B7.4	F17E5.1	F49H12.6	T01C8.1	Y49E10.1	
C28H8.6	F23C8.4	F52B10.1	T02C5.1	Y51H1A.3	
C29F9.5	F23F12.9	F53F4.10	T03G11.1	Y51H4A.17	
C30B5.4	F23H11.5	F54B11.3	T04D1.3	Y53C12B.3	
C31H1.6	F25B5.7	F54B11.6	T05G5.3	Y53H1A.2	
C34B2.4	F26F4.10	F54C1.7	T07D4.2	Y55F3AM.15	
C34B7.3	F26G5.9	F54E7.3	T07F8.3	Y56A3A.20	
C36B1.4	F28C6.2	F54G8.4	T07G12.10	Y56A3A.21	
C36E8.5	F28F5.3	F56A8.6	T08G11.4	Y66A7A.6	
C43E11.4	F28H6.1	F57C9.4	T10F2.1	Y67D2.6	
C43E11.6	F29G9.5	F58A4.8	T12D8.7	Y73B6A.5	
C44B7.1	F30F8.3	F59A6.1	T12G3.1	Y73B6BL.33	
C44B9.4	F31C3.2	F59E12.2	T13C5.4	Y92C3B.2	
C44C10.4	F31E3.2	F59E12.9	T14F9.1	ZC239.15	
C44C8.6	F31E3.5	H22K11.1	T16H12.4	ZC395.8	

Table B.5. Third Neighbours

3rd neighbours				
Name	**Name**	**Name**	**Name**	**Name**
B0035.1	E04F6.6	F49H12.3	R06C7.10	Y116A8C.35
B0280.8	EEED8.16	F53G12.5	R06F6.4	Y119C1B.5
B0336.2	EEED8.5	F54B11.5	R07E4.6	Y18D10A.5
B0350.2	F08F3.3	F54C4.2	R08E3.4	Y18D10A.8
B0365.1	F11H8.1	F55A11.3	R10E11.2	Y38A10A.5
B0412.1	F17A2.5	F55A4.1	R119.7	Y38A8.2
B0495.9	F18E2.2	F55H12.6	R12B2.1	Y38F2AR.7
C02C6.1	F20G4.3	F56H1.4	T01G1.1	Y39E4B.1
C04F5.8	F22B7.5	F57B10.10	T01G9.5	Y48C3A.17
C05C8.6	F23F1.8	F57B10.3	T04C10.1	Y48E1A.1
C05C8.7	F23H11.1	F57B10.7	T04C12.5	Y48G1C.1
C05D2.1	F23H12.2	F58E6.10	T05E7.1	Y49A3A.1
C05E11.3	F25E2.5	F58E6.3	T07C4.1	Y49F6C.3
C06A1.1	F25H2.9	F59A2.4	T09A12.4	Y51F10.2
C07A12.1	F25H5.3	F59B1.7	T12D8.2	Y52D3.1
C07A12.3	F25H9.2	H06H21.3	T18D3.7	Y54G11A.10
C07G1.5	F26D10.3	H10E21.3	T19D12.5	Y55F3AM.10
C08C3.3	F26D11.11	H10E21.4	T20G5.1	Y56A3A.13
C09H6.2	F26H11.2	H14N18.1	T21E8.1	Y56A3A.4
C14F5.1	F29B9.4	H19N07.1	T22G5.5	Y59A8B.22
C18D1.1	F29C4.1	H21P03.3	T23F6.4	Y5H2B.2
C25B8.3	F32A11.2	H28O16.1	T25C8.2	Y60A3A.1
C25F6.2	F32A11.6	K01G5.1	T28F12.2	Y62E10A.16
C26B2.3	F32B4.4	K02A4.2	W01D2.2	Y65B4BR.4
C32F10.6	F32G8.6	K02F2.2	W02D3.9	Y75B8A.1
C34E10.6	F33D11.10	K04F10.6	W03D8.8	ZC155.1
C37C3.6	F35G2.2	K05C4.1	W05B10.1	ZK1240.2
C39E9.14	F36A2.10	K06A1.4	W07G1.3	ZK596.2
C40A11.7	F37B1.3	K06B4.1	W09C3.6	ZK637.5
C41C4.8	F37B1.4	K07A12.2	W10D9.4	ZK652.6
C44F1.2	F37B1.8	K08E5.3	Y104H12A.1	ZK652.9
C45G3.3	F37H8.3	K08F11.3	Y105C5B.28	ZK867.1
C49H3.5	F38A6.1	K08F8.4	Y105E8B.5	ZK892.7
C50C3.8	F38A6.3	M01E11.6	Y105E8B.8	ZK909.4
C56C10.8	F42D1.2	M01F1.2	Y106G6D.7	ZK970.4
D1022.8	F43C11.7	M04B2.1	Y106G6H.2	
D1053.1	F43E2.8	M04F3.1	Y111B2A.22	
D1081.2	F43G6.8	M7.5	Y113G7A.6	
D2096.8	F47B10.1	R04B5.4	Y113G7B.23	
E02H1.7	F47F6.1	R06B10.4	Y116A8C.12	

Table B.6. Fourth neighbours

4th neighbours				
Name	**Name**	**Name**	**Name**	**Name**
B0207.4	F01D4.4	F58A4.4	T10C6.11	ZK792.6
B0272.4	F07A5.1	F58D5.1	T14G12.4	
B0281.5	F08C6.6	F58G1.3	T22C8.3	
C01B10.8	F08F3.2	F59B2.3	T22D1.12	
C01B12.1	F09F3.9	H14A12.2	T23C6.5	
C01G12.1	F11E6.5	H42K12.1	T24D1.3	
C01G6.4	F11G11.2	K01G5.4	T28A11.11	
C03C10.3	F12F6.5	K02B9.4	W02B12.8	
C03G6.3	F13D12.6	K04A8.6	Y105C5B.13	
C04C3.3	F18A1.5	K08F4.7	Y116F11B.12	
C04H5.6	F21F12.1	K08H2.8	Y37D8A.10	
C05D11.10	F25H5.4	K09A11.3	Y40D12A.2	
C05D9.1	F28F8.6	K10C3.6	Y43C5B.2	
C06G8.4	F31E3.3	K12D12.1	Y48E1B.3	
C08G9.2	F38E9.1	M01H9.2	Y51H4A.4	
C13B9.3	F38H4.8	M03D4.1	Y54E10A.6	
C17G10.5	F39B2.11	R11F4.1	Y57G11C.22	
C25A1.5	F39C12.2	R13D7.7	Y75B8A.2	
C25F6.3	F42G10.2	R13H8.1	Y81G3A.3	
C32D5.12	F43H9.2	R151.2	ZC581.1	
C35D10.2	F44A2.5	R186.5	ZK1058.2	
C47B2.4	F44C4.3	R74.5	ZK1127.7	
C54G7.4	F46B6.7	T04F3.1	ZK353.6	
C56E10.4	F57B10.11	T05G5.6	ZK637.11	
D1037.4	F58A4.3	T07D1.4	ZK662.3	

Table B.7. Fifth neighbours

5th neighbours				
Name	**Name**	**Name**	**Name**	**Name**
AC7.2	C31H1.5	F33C8.1	T04A11.6	Y73C8B.3
B0019.2	C39E9.13	F41C3.5	T06C10.3	Y75B8A.35
B0478.1	C39F7.4	F45H10.4	T07A9.3	Y80D3A.4
C05G6.1	C46E10.9	F54A5.1	T22B2.4	ZK1037.5
C06G3.1	C54G10.2	H09G03.2	T23G7.1	ZK632.10
C09H5.6	F08F3.4	K08B12.5	T26H2.9	ZK697.2
C10F3.5	F18C5.2	K11D12.9	W06D4.6	
C14B9.8	F18E2.1	K11E4.5	Y24F12A.2	
C14F11.6	F20B10.1	R05D3.7	Y38E10A.18	
C17E4.5	F23F12.6	R107.7	Y47D7A.1	
C27A12.7	F25D7.1	T03F6.1	Y49E10.23	
C27C7.4	F26E4.8	T03F6.3	Y53C12A.1	

Table B.8. Sixth neighbours

6th neighbours	
Name	**Name**
C09G4.5	T21G5.5
F37A4.1	W02A11.3
F46A9.6	Y43C5A.6
F46F2.2	Y87G2A.14
T01D1.2	

Table B.9. Seventh, eighth and ninth neighbours

7th neighbours
Name
F26B1.2
Y116A8C.13
8th neighbours
Name
C25A1.4
F59E12.4
Y48B6A.3
Y49E10.14
Y59A8B.10
9th neighbours
Name
F19B6.1
F44G4.4

APPENDIX C: CORE PROTEIN INTERACTIONS OF DIFFERENT SPECIES

Table C.1. Comparison of core protein interaction between species

CAEEL			FLY			HUMAN		
Orf name	Name	# of PPI	Name	Orf name	# of PPI	name	Orf name	# of PPI
WNT			**WNT**			**WNT**		
W01B6.1	cwn-2	0	N/A		0	WNT5A	EG7474	1
K10B4.6	cwn-1	0						
W08D2.1	egl-20	1						
F38E1.7	mom-2	0						
E01A2.3	lin-44	0						
FZD			**FZD**			**FZD**		
T23D8.1	mom-5	0	fz2	CG9739	0	FZD2	EG2535	1
F27E11.3	cfz-2	0						
Y71F9B.5	lin-17	0						
Y34D9B.1	mig-1	0						
DVL			**DVL**			**DVL**		
C34F11.9	dsh-1	4	dsh	CG1836	1	DVL1	EG1855	12
T05C12.6	mig-5	34				DVL2	EG1856	17
C27A2.6	dsh-2	6				DVL3	EG1857	10
C52A11.4	mpz-1	0						
PLC			**PLC**			**PLC**		
B0348.4	egl-8	0	norpA	CG3620	3	PLCB1	RP4-654A7.1	3
Y75B12B.6	plc-2	0	Plc21C	CG4574	0	PLCB2	EG5330	3
K10F12.3	pll-1	0				PLCB3	EG5331	2
F31B12.1	plc-1	0				PLCB4	RP4-811H13.1	0
T01E8.3	plc-3	0						
R05G6.8	plc-4	0						
CAMKII			**CAMKII**			**CAMKII**		
K11E8.1	unc-43	2	CaMKII	CG18069	2	CAMK2A	EG815	9
K07A9.2	cmk-1	1				CAMK2B	EG816	1
ZC373.4	ZC373.4	0				CAMK2D	EG817	1
						CAMK2G	EG818	2
CACN			**CACN**			**CACN**		
ZK856.8	ZK856.8	0	CACNB2	CG11217	0	CHP	EG11261	0
Y71H2AL.1	Y71H2AL.1	0	CACNA1	CG1455	3	PPP3CA	EG5530	4
F55C10.1	cnb-1	0	CG2185	CG2185	0	PPP3CB	RP11-345K20.1	0
F30A10.1	F30A10.1	0	CACNB	CG4209	3	PPP3CC	EG5533	0
C02F4.2	tax-6	0	CACNA-14F	CG9819	0	PPP3R1	EG5534	0
Y49E10.3	pph-4.2	0	Pp2B-14D	CG9842	6	PPP3R2	EG5535	0
F59D6.7	F59D6.7	0						

Table C.1 Comparison of core protein interaction between species (continued)

CAEEL			FLY			HUMAN		
Orf name	Name	# of PPI	Name	Orf name	# of PPI	Name	Orf name	# of PPI
PKC			**PKC**			**PKC**		
E01H11.1	pkc-2	3	inaC	CG6518	3	PRKCA	EG5578	16
F57F5.5	pkc-1	1	Pkc53E	CG6622	0	PRKCB	EG5579	10
						PRKCG	EG5582	1
NFAT			**NFAT**			**NFAT**		
C24A8.3	pqn-15	0	NFAT	CG11172	0	NFAT5	EG10725	1
C37A2.2	pqn-20	1				NFATC1	EG4772	3
F47A4.2	dpy-22	0				NFATC2	RP5-1009H6.1	8
K06A9.1	K06A9.1	0				NFATC3	EG4775	0
T05A10.1	sma-9	1				NFATC4	EG4776	1
C04A2.3	egl-27	0						
C32A3.1	sel-8	0						
F22H10.2	F22H10.2	0						
Y39B6A.1	Y39B6A.1	11						
F36A4.7	ama-1	0						
TAK1			**TAK1**			**TAK1**		
F52F12.3	mom-4	2	Tak1	CG18492	1	TAK1	EG6885	9
Y105C5A.24	Y105C5A.24	0						
F33E2.2	dlk-1	0						
C24A1.3	C24A1.3	0						
T06C10.6	kin-26	0						
NLK			**NLK**			**NLK**		
W06F12.1	lit-1	10	nmo	CG7892	1	NLK	EG51701	3
F43C1.2	mpk-1	23						
CDC42			**CDC42**			**CDC42**		
R07G3.1	cdc-42	0	cdc42	CG12530	1	cdc42	G25K	10
C09G12.8	ced-10	0						
K03D3.10	rac-2	0						
C35C5.4	mig-2	0						
Y51H4A.3	rho-1	0						
F22E12.2	chw-1	1						
Y32F6B.3	crp-1	1						
C08F8.7	rap-3	0						

APPENDIX D: MACRO SCALE ANALYSIS RESULTS

In this section, connectivity distribution results of graph theory based analysis are included.

Table D.1 Connectivity distribution of the proteins

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
T04H1.2	47	F25E2.5	7	Y15E3A.1	5
T11B7.4	43	F26B1.2	7	C36B1.4	5
C38D4.6	38	F31C3.2	7	C02F5.9	5
T05C12.6	34	F42H10.7	7	ZK892.7	5
F44G3.9	32	F44B9.6	7	R03G5.2	5
W04D2.1	27	R07E4.6	7	M02A10.3	5
ZK849.1	25	Y73B6A.5	7	C05D9.1	5
F43C1.2	23	ZK858.4	7	T22B2.4	5
F14F3.1	16	C08F8.8	7	Y92C3B.2	5
K10C3.6	15	T12D8.7	7	Y42H9AR.1	5
C47E8.5	13	F23C8.4	7	F59E12.2	5
C18H9.7	13	W05H7.4	7	T01C8.1	4
K01G5.1	13	F35G12.9	6	T25C8.2	4
T27F2.3	12	C06A5.9	6	K07C11.2	4
F47G6.1	12	C27A2.6	6	B0207.4	4
B0218.3	11	F47D12.4	6	K08F8.2	4
B0547.1	11	F58A4.8	6	B0336.7	4
C27B7.4	11	Y62E10A.16	6	C01B10.8	4
R12B2.1	11	T28A11.11	6	C34F11.9	4
C52B11.2	11	K08B4.1	6	W07B8.5	4
Y39B6A.1	11	C32F10.6	6	C39E9.14	4
ZC504.4	10	Y75B8A.1	6	F32D1.1	4
W06F12.1	10	K01G5.4	6	F35G2.2	4
C06G3.6	9	F23F1.8	6	F46G11.1	4
Y51H4A.17	9	R13F6.9	6	H15N14.2	4
Y79H2A.11	9	R06B10.4	6	C56C10.8	4
ZK1098.4	9	Y119C1A.1	6	F35C8.3	4
C07H6.5	8	Y40C5A.1	6	K02F2.2	4
C12D8.1	8	Y110A7A.10	5	K11D2.3	4
C48D5.1	8	C34E10.6	5	K11D9.1	4
C04F12.3	8	C05C8.6	5	F26D11.11	4

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
F17E5.1	8	C06A6.2	5	T01G9.5	4
D1054.2	8	C28H8.9	5	Y2H9A.1	4
F38A6.1	8	C49G7.1	5	B0280.8	4
Y113G7B.23	8	F31E3.5	5	K08E3.7	4
B0464.5	8	F01F1.12	5	R05F9.1	4
Y37E11AR.2	8	F31E3.2	5	R07E5.8	4
ZK632.12	8	K08F11.3	5	F01F1.4	4
D1046.1	7	Y104H12A.1	5	W06D4.6	4
F31E3.3	4	H28O16.1	3	C09G1.4	2
C27B7.1	4	F26D10.3	3	C23H3.3	2
T09A5.12	4	ZK353.6	3	C25A1.4	2
T18D3.7	4	ZK792.6	3	C25G4.4	2
T23H4.2	4	F38H4.9	3	C27A12.2	2
T26E3.3	4	C09H6.2	3	C31H1.6	2
C44H4.5	4	M04F3.1	3	C40A11.2	2
Y113G7A.6	4	F42G10.2	3	C44B7.1	2
T28F12.2	4	Y5H2B.2	3	C54E10.2	2
ZC434.8	4	E02H1.7	3	F32G8.6	2
F39C12.2	3	C26B2.3	3	Y56A3A.20	2
F28H6.1	3	F59A6.1	3	ZK637.11	2
ZK637.5	3	F25H2.9	3	C41C4.8	2
F56H1.4	3	Y38A8.2	3	F56A8.6	2
B0496.7	3	C37C3.6	3	D1005.1	2
C50C3.8	3	R74.5	3	D2045.8	2
Y49F6C.3	3	Y43C5A.6	3	C05D2.1	2
C13F10.7	3	C39E9.13	3	F25H5.4	2
C23G10.3	3	C18D11.4	3	T27C4.4	2
C34B2.4	3	Y105C5B.13	3	R11E3.6	2
C35D10.2	3	T07D4.2	3	F08C6.6	2
F17A2.5	3	T09A5.2	3	F08C6.7	2
F43G9.11	3	F26G5.9	3	F13D12.6	2
ZK909.4	3	Y105E8A.6	3	F23B12.5	2
F46F11.2	3	Y65B4BR.4	3	F25B5.7	2
T20G5.1	3	Y105E8B.5	3	F43G6.8	2
Y38A10A.5	3	Y18D10A.5	3	F46G10.1	2
Y55F3AM.15	3	Y18D10A.8	3	F46H5.3	2
E01H11.1	3	Y53F4B.33	3	F52F12.3	2
F18A1.3	3	Y59A8B.22	3	F55H12.6	2
F25H5.3	3	ZC239.15	3	F59A2.4	2

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
F29G9.3	3	M03D4.1	3	F59B2.3	2
F37A4.9	3	ZK112.2	3	F59E12.4	2
F44D12.1	3	B0336.2	2	F59E12.9	2
F49H12.3	3	M7.5	2	T14G12.4	2
F53B3.1	3	B0365.1	2	T07F8.3	2
F54F2.5	3	F57B10.11	2	T10F2.1	2
B0024.12	3	C04C3.3	2	F37B1.1	2
Y46G5A.31	3	C05C10.5	2	D1053.1	2
C07A12.1	2	T12G3.1	2	C08B11.1	2
T04A11.6	2	T20H4.5	2	F49H12.6	1
F32A11.2	2	T22C8.3	2	F08F3.2	1
K11E8.1	2	T22D1.12	2	ZK455.1	1
K12D12.1	2	T23C6.5	2	M03F4.2	1
F46F2.2	2	T26C12.3	2	T04C12.5	1
M01E11.6	2	T27F2.1	2	F08C6.1	1
Y54G11A.10	2	T27F2.2	2	C25A11.4	1
M01F1.2	2	F12F6.5	2	Y116A8C.12	1
F46A9.6	2	F44G4.4	2	K11D2.2	1
AH6.5	2	F28F5.3	2	H22K11.1	1
ZC155.1	2	F46G11.3	2	F45E6.2	1
Y45G5AM.1	2	C28H8.6	2	C44B9.4	1
K06A1.4	2	C36E8.5	2	F29G9.5	1
C07A12.3	2	ZK673.7	2	F28F8.6	1
F32B6.1	2	C07F11.1	2	B0019.2	1
C45E5.6	2	H14N18.1	2	B0035.1	1
T09A12.4	2	W02D3.9	2	B0272.4	1
Y75B8A.2	2	B0350.2	2	B0281.5	1
Y106G6H.2	2	T04D1.3	2	B0464.9	1
K08F8.4	2	F14D12.2	2	B0495.9	1
PAR2.1	2	W03D8.8	2	F54B11.6	1
K05C4.1	2	Y53C12A.1	2	F23H11.1	1
H42K12.1	2	Y116A8C.26	2	C01G12.1	1
K09B11.1	2	Y37D8A.10	2	C01G6.4	1
CD4.6	2	Y40D12A.2	2	C04F5.8	1
F29B9.4	2	Y48G1C.1	2	C05C8.7	1
R08E3.4	2	Y54E10BL.6	2	C05D11.10	1
R10E11.2	2	Y55F3AM.10	2	C05D12.5	1
C54G10.2	2	Y56A3A.21	2	C05E11.3	1
W02B12.8	2	Y57G11C.22	2	C07D10.2	1

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
C03C10.3	2	Y57G11C.9	2	C08G9.2	1
Y49E10.1	2	Y66D12A.5	2	C13B9.3	1
ZK370.2	2	Y73B6BL.33	2	C14B9.8	1
K12G11.3	2	ZC477.9	2	C14F11.6	1
T21G5.5	2	ZC513.6	2	C14F5.1	1
T02E1.3	2	ZK121.2	2	C25A1.5	1
T07C4.1	2	ZK1307.8	2	C25F6.3	1
T08G11.4	2	ZC395.8	2	C27A12.7	1
C29F9.5	1	F45H10.4	1	F43C11.7	1
C30B5.4	1	T12D8.2	1	F44A2.5	1
C31H1.5	1	C02C6.1	1	F44B9.9	1
C32D5.12	1	E04F6.6	1	F45G2.3	1
C36E6.5	1	F38H4.8	1	F47B10.1	1
C37A2.2	1	T05G5.6	1	F47F6.1	1
C40A11.7	1	EEED8.16	1	F53F4.10	1
C44C10.4	1	EEED8.7	1	F54A3.4	1
C44F1.2	1	Y48C3A.17	1	F54A5.1	1
C45E1.1	1	F01D4.4	1	F54B11.5	1
C46E10.9	1	F11E6.5	1	F55A4.1	1
C47E8.4	1	K02B9.4	1	F55G1.8	1
C56G7.1	1	Y105E8B.8	1	F56D1.2	1
D1022.8	1	T01D1.2	1	F57B10.3	1
K05G3.3	1	F02A9.6	1	F57B9.6	1
ZK632.6	1	F08F3.4	1	F57C9.4	1
C06A1.1	1	F13E6.4	1	F57F5.5	1
T05G5.3	1	F14F3.2	1	F58E6.3	1
Y47D3A.4	1	F18A1.5	1	F58G1.3	1
C15C8.2	1	F18E2.1	1	T07D1.4	1
ZK652.9	1	F19B6.1	1	H09G03.2	1
F44C4.3	1	F21F12.1	1	H14A12.2	1
C25B8.3	1	F22E12.2	1	C14B9.6	1
F09F3.9	1	F23F12.9	1	C45G3.3	1
F25D7.1	1	F23H11.5	1	Y105C5B.28	1
K09A11.3	1	F23H12.2	1	Y116F11B.12	1
C34B7.3	1	F25H9.2	1	Y66A7A.6	1
D1005.3	1	F28C6.2	1	K02A4.2	1
D2005.3	1	F30F8.3	1	W09C3.6	1
D2096.8	1	F32B4.4	1	R107.7	1
B0412.1	1	F33D11.10	1	R11G1.3	1

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
F57B10.10	1	F36A2.10	1	F37B1.3	1
F29C4.1	1	F37A4.1	1	F37B1.4	1
R13H8.1	1	F37H8.3	1	F37B1.8	1
M18.5	1	F38E9.1	1	R07B1.4	1
C18D1.1	1	F39B2.11	1	K08F4.7	1
C25F6.2	1	F41C3.5	1	R13D7.7	1
F22B7.5	1	F42D1.2	1	F11G11.2	1
T23G7.1	1	F42F12.3	1	H06H21.3	1
H10E21.4	1	F32A11.6	1	F54C1.7	1
H19N07.1	1	C04H5.6	1	ZK1058.2	1
F58A4.3	1	EEED8.5	1	C47B2.4	1
C07G1.5	1	K08E5.3	1	F10B5.5	1
F38A6.3	1	K04F10.6	1	C10F3.5	1
C50F4.13	1	F52B10.1	1	T21E8.1	1
T10C6.11	1	C43E11.6	1	Y49E10.14	1
W05B10.1	1	Y87G2A.14	1	F09E5.1	1
Y48B6A.14	1	Y56A3A.13	1	F42G8.3	1
T20B12.8	1	F54G8.4	1	Y38F2AR.7	1
F55A11.3	1	C06G3.1	1	T03G11.1	1
F58D5.1	1	ZK1025.9	1	F58A4.4	1
F43E2.8	1	R04B5.4	1	T03F6.1	1
C54G7.4	1	C56E10.4	1	R06C7.10	1
F07A5.1	1	Y38E10A.18	1	R06F6.4	1
R11A8.6	1	Y80D3A.4	1	R119.7	1
B0478.1	1	ZK1037.5	1	R11F4.1	1
K03H6.2	1	ZK697.2	1	R151.2	1
K04A8.6	1	K08H2.8	1	C39F7.4	1
K04D7.2	1	ZK662.3	1	D1037.4	1
K07A12.2	1	K06B4.1	1	T23F6.4	1
K07A9.2	1	K06B4.11	1	F11H8.1	1
K11D12.9	1	W01D2.2	1	F08F3.3	1
T07A9.3	1	K11E4.5	1	ZK863.7	1
T01G1.1	1	C27C7.4	1	T23G5.1	1
R144.1	1	C05G6.1	1	F23F12.6	1
F33A8.1	1	T26H2.9	1	F26F4.10	1
ZK662.4	1	H10E21.3	1	C54G4.1	1
F48E8.1	1	F20B10.1	1	R186.5	1
C17G10.5	1	F20G4.3	1	F46A9.4	1
M01H9.2	1	R03D7.7	1	Y47D7A.1	1

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
C08C3.3	1	Y53C12B.3	1	H21P03.3	1
K01A6.2	1	C49H3.5	1	AC7.2	1
C44C8.6	1	F26H11.2	1	F54C4.2	1
T04C10.1	1	K04G7.3	1	F43H9.2	1
M04B2.1	1	C09F5.2	1	T22G5.5	1
C09G4.5	1	F48E8.5	1	C01B12.1	1
W03C9.7	1	C17E4.5	1	F07C4.5	1
F53G12.5	1	F54E7.3	1	C06G8.4	1
C03G6.3	1	W03F11.6	1	ZK652.6	1
F59B1.7	1	W03F9.10	1	ZK867.1	1
Y111B2A.22	1	W08D2.1	1	F46B6.7	1
C09H5.6	1	W10D9.4	1	ZK287.6	1
T01H8.1	1	F18C5.2	1	F31B12.1	0
T02C5.1	1	Y48B6A.3	1	T01E8.3	0
T03F6.3	1	Y106G6D.7	1	ZC373.4	0
T04F3.1	1	Y116A8C.13	1	B0348.4	0
T05A10.1	1	Y119C1B.5	1	C02F4.2	0
T05E7.1	1	Y17G7B.4	1	C04A2.3	0
T06C10.3	1	Y24D9A.8	1	C08F8.7	0
T08D10.1	1	Y24F12A.2	1	C09G12.8	0
T09F3.1	1	Y32F6B.3	1	C09G4.2	0
T13C5.4	1	Y43C5B.2	1	C24A1.3	0
T16H12.4	1	Y48E1A.1	1	C24A8.3	0
T19D12.5	1	Y48E1B.3	1	C32A3.1	0
T24D1.3	1	Y49A3A.1	1	C35C5.4	0
Y56A3A.4	1	Y49E10.23	1	C52A11.4	0
F18E2.2	1	Y51F10.2	1	E01A2.3	0
W08E3.3	1	Y51H1A.3	1	E01F3.1	0
Y39E4B.1	1	Y51H4A.4	1	F22H10.2	0
F33C8.1	1	Y52D3.1	1	F27E11.3	0
K08B12.5	1	Y53H1A.2	1	F30A10.1	0
F26E4.8	1	Y54E10A.6	1	F33E2.2	0
Y37A1B.13	1	Y59A8B.10	1	F36A4.7	0
F57B10.7	1	Y67D2.6	1	F38E1.7	0
C43E11.4	1	Y73C8B.3	1	F47A4.2	0
Y116A8C.35	1	Y81G3A.3	1	F55A8.2	0
R05D3.7	1	ZC449.6	1	F55C10.1	0
D1081.2	1	ZC581.1	1	F59D6.7	0
F58E6.10	1	T07G12.10	1	K03D3.10	0

Table D.1 Connectivity distribution of the proteins (continued)

ORF name	Connectivity	ORF name	Connectivity	ORF name	Connectivity
Y60A3A.1	1	ZK1240.9	1	T06C10.6	0
F54B11.3	1	ZK1290.4	1	T23D8.1	0
R10E11.8	1	ZK596.2	1	W01B6.1	0
T14F9.1	1	ZK632.10	1	Y105C5A.24	0
ZK970.4	1	K06A9.1	0	Y34D9B.1	0
VW06B3R.1	1	K10B4.6	0	Y49E10.3	0
W02A11.3	1	K10F12.3	0	Y51H4A.3	0
W02D3.11	1	R05G6.8	0	Y71F9B.5	0
Y75B8A.35	1	R07G3.1	0	Y71H2AL.1	0
W07G1.3	1	R08D7.6	0	Y75B12B.6	0
ZK1127.7	1	R153.1	0	Y95B8A.10	0
ZK1240.2	1	T04D3.3	0	ZK856.8	0
				ZK909.2	0

APPENDIX E: DOMAINS and DOMAIN INTERACTIONS of NONCANONICAL WNT SIGNALLING NETWORK in *C. ELEGANS*

Domains and domain interactions of the reconstructed network are included in this section.

Table E.1. Domain descriptions (Domain Id refers to Family Id and Domain description refers to Family description according to SCOP)

Domain ID	Domain Description
63912	1,4-beta-N-acetylmuraminidase
64162	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase, catalytic domain
64159	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase, substrate-binding domain
52048	28-residue LRR
53045	5' to 3' exonuclease catalytic domain
47808	5' to 3' exonuclease, C-terminal subdomain
53384	AAT-like
52686	ABC transporter ATPase domain-like
102620	Aclacinomycin methylesterase RdmC
52017	Aconitase
53733	Aconitase iron-sulfur domain
53068	Actin/HSP70
51736	Alcohol dehydrogenase-like, C-terminal domain
50136	Alcohol dehydrogenase-like, N-terminal domain
47221	alpha-catenin/vinculin
111353	Ammonium ransporter (Pfam 00909)
48404	Ankyrin repeat
47875	Annexin
47324	Anticodon-binding domain of a subclass of class I aminoacyl-tRNA synthetases
52955	Anticodon-binding domain of Class II aaRS
101888	Apyrase
53640	AraD-like aldolase/epimerase
101274	Archaeal tRNA CCA-adding enzyme substrate-binding domain
55191	Arginyl-tRNA synthetase (ArgRS), N-terminal 'additional' domain
48372	Armadillo repeat
56535	Aromatic aminoacid monoxygenases, catalytic and oligomerization domains
89851	Association domain of calcium/calmodulin-dependent protein kinase type II alpha subunit, CAMK2A

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
81850	Bacterial glucoamylase C-terminal domain-like
51699	Bacterial PLC
63492	BAG domain
48351	BCR-homology GTPase activation domain (BH-domain)
82509	Biotin biosynthesis protein BioH
51231	Biotinyl/lipoyl-carrier proteins and domains
52926	Branched-chain alpha-keto acid dehydrogenase beta-subunit, C-terminal-domain
88741	Branched-chain alpha-keto acid dehydrogenase Pyr module
63955	Breast cancer associated protein, BRCA1
47371	Bromodomain
54696	BTB/POZ domain
102492	CAF1-like ribonuclease
47502	Calmodulin-like
69236	Calnexin/calreticulin
47577	Calponin-homology domain, CH-domain
51210	cAMP-binding domain
54929	Canonical RBD
53522	Carbon-carbon bond hydrolase
51070	Carbonic anhydrase
49465	Carboxypeptidase regulatory domain
102932	Catalytic subunit of bi-partite nucleotidyltransferase
52778	CAT-like
90230	CCCH zinc finger
75285	Ccg1/TafII250-interacting factor B (Cib)
54586	Cdc48 domain 2-like
50708	Cdc48 N-terminal domain-like
52822	Cell cycle control phosphatase, catalytic domain
46847	Cell cycle transcription factor e2f-dp
46566	Chaperone J-domain
82424	Choline/Carnitine O-acyltransferase (Pfam 00755)
48257	Citrate synthase
51570	Class I aldolase
52375	Class I aminoacyl-tRNA synthetases (RS), catalytic domain
55682	Class II aminoacyl-tRNA synthetase (aaRS)-like, catalytic domain
57668	Classic zinc finger, C2H2
75521	Clathrin coat assembly domain
48390	Clathrin heavy chain proximal leg segment
48393	Clathrin heavy-chain linker domain
50990	Clathrin heavy-chain terminal domain

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
51900	CoA-binding domain
50282	Cold shock DNA-binding domain-like
52088	CRAL/TRIO domain
52103	Crotonase-like
47918	C-terminal domain of alpha and beta subunits of F1 ATP synthase
46844	C-terminal domain of RPA32
55857	Cytochrome b5
48265	Cytochrome P450
74693	Dachshund-homology domain (Pfam 02437)
81312	DEATH domain, DD
51375	Decarboxylase
52468	Deoxyhypusine synthase, DHS
63483	DEP domain
53518	Dienelactone hydrolase
46553	Dihydropyrimidine dehydrogenase, N-terminal domain
55879	DNA gyrase/MutL, N-terminal domain
54224	DNA gyrase/MutL, second domain
102396	DNA helicase UvsW
81300	DNA polymerase beta-like
48020	DNA polymerase III clamp loader subunits, C-terminal domain
56748	DNA primase
47795	DNA repair protein Rad51, N-terminal domain
110080	DNA-binding protein LAG-1 (CSL)
57939	DnaJ/Hsp40 cysteine-rich domain
89838	Doublecortin (DC)
51183	dTDP-sugar isomerase
54981	EF-G/eEF-2 domains III and V
47547	EF-hand modules in multidomain proteins
50466	EF-Tu/eEF-1alpha/eIF2-gamma C-terminal domain
57197	EGF-type module
57257	Elafin-like
50448	Elongation factors
56789	Epoxide hydrolase, N-terminal domain
110020	ERO1-like (Pfam 04137)
57717	Erythroid transcription factor GATA-1
111375	ETX/MTX2 (Pfam 03318)
54792	Eukaryotic type KH-domain (KH-domain type I)
81269	Extended AAA-ATPase domain
57354	Extracellular domain of cell surface receptors

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
81332	F1F0 ATP synthase subunit C
81381	F-box domain
54884	Ferredoxin domains from multidomain proteins
49885	FHA domain
51396	FMN-linked oxidoreductases
46832	Forkhead DNA-binding domain
51830	Formate/glycerate dehydrogenases, NAD-domain
63502	Frizzled cysteine-rich domain
53558	Fungal lipases
57904	FYVE, a phosphatidylinositol-3-phosphate binding domain
52592	G proteins
53417	GABA-aminotransferase-like
55782	GAF domain
50966	Galactose oxidase, central domain
49786	Galactose-binding domain
55932	Glutamine synthetase catalytic domain
54369	Glutamine synthetase, N-terminal domain
47617	Glutathione S-transferase (GST), C-terminal domain
52862	Glutathione S-transferase (GST), N-terminal domain
53089	Glycerol kinase
69594	Glycerol-3-phosphate (1)-acyltransferase
110734	Glycosyl transferases group 1 (Pfam 00534)
63588	Glycosyltransferase family 36 C-terminal domain
57185	Growth factor receptor domain
55621	GTP cyclohydrolase I
55935	Guanido kinase catalytic domain
48035	Guanido kinase N-terminal domain
48385	HEAT repeat
100935	Heat shock protein 70kD (HSP70), C-terminal subdomain
100921	Heat shock protein 70kD (HSP70), peptide-binding domain
55875	Heat shock protein 90, HSP90, N-terminal domain
56205	Hect, E3 ligase catalytic domain
69496	Helicase-like "domain" of reverse gyrase
55789	Heme-binding PAS domain
55519	Hemorrhagin
82200	Histone lysine methyltransferases
54198	HIT (HINT, histidine triad) family of protein kinase-interacting proteins
47460	HLH, helix-loop-helix DNA-binding domain
47096	HMG-box

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
46690	Homeodomain
47820	HRDC domain from RecQ helicase
49494	HSP40/DnaJ peptide-binding domain
110943	HSP90 C-terminal domain (C-terminal part of Pfam 00183)
102755	Hsp90 middle domain
74933	HtrA-like serine proteases
57263	Huristasin-like
46951	Hypothetical protein MTH1615
89751	Hypothetical protein Ta1320
82194	Hypoxia-inducible factor HIF ihhibitor (FIH1)
103184	Hypoxia-inducible factor Hif2a, C-terminal domain
49159	I set domains
110513	IF2B-like (Pfam 01008)
57925	Inhibitor of apoptosis (IAP) repeat
48620	Insect phospholipase A2
53301	Integrin A (or I) domain
69940	Integrin beta EGF-like domains
69688	Integrin beta tail domain
69180	Integrin domains
52059	Internalin LRR domain
100948	Isochorismatase-like hydrolases
102713	JAB1/MPN domain
47041	Kix domain of CBP (creb binding protein)
100940	Ku70 subunit middle domain
100959	Ku70 subunit N-terminal domain
100943	Ku80 subunit middle domain
100962	Ku80 subunit N-terminal domain
101289	L27 domain
49944	Laminin G-like module
57233	Laminin-type module
57743	LASP-1
48558	L-aspartase/fumarase
57425	LDL receptor-like module
53201	Leucine aminopeptidase, C-terminal domain
69433	Leucine rich effector protein YopM
57736	LIM domain
51696	Mammalian PLC
50912	Mannose 6-phosphate receptor domain
100908	MIF4G domain-like

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
47158	Mitochondrial import receptor subunit Tom20
69573	Molybdenum cofactor biosynthesis protein MoeB
52641	Motor proteins
63412	MPP-like
90124	Multidrug resistance ABC transporter MsbA, N-terminal domain
55812	MutT-like
46739	Myb/SANT domain
89743	N5-glutamine methyltransferase, HemK
55730	N-acetyl transferase, NAT
82227	N-acetylglucosamine-6-phosphate deacetylase, NagA
82261	N-acetylglucosamine-6-phosphate deacetylase, NagA, catalytic domain
52513	NagB-like
81279	NF-kappa-B/REL/DORSAL transcription factors, C-terminal domain
75142	Ngr ectodomain-like
101899	NHL repeat
56318	Nitrilase
52652	Nitrogenase iron protein-like
90194	Notch domain
51972	N-terminal domain of adrenodoxin reductase-like
50616	N-terminal domain of alpha and beta subunits of F1 ATP synthase
47669	N-terminal domain of cbl (N-cbl)
82764	N-terminal PAS domain of Pas kinase
57721	Nuclear receptor
48509	Nuclear receptor ligand-binding domain
47114	Nucleosome core histones
52541	Nucleotide and nucleoside kinases
90214	NZF domain
63571	PABC (PABP) domain
46748	Paired domain
53188	Pancreatic carboxypeptidases
54002	Papain-like
64225	PB1 domain
109671	PCI domain (PINT motif, Pfam 01399)
48548	PDEase
63888	P-domain of calnexin/calreticulin
50157	PDZ domain
51127	Pectate lyase
50646	Pepsin-like
47006	Peripheral subunit-binding domain of 2-oxo acid dehydrogenase complex

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
57911	PHD domain
55028	Phenylalanine hydroxylase N-terminal domain
53296	Phosphoribosylpyrophosphate synthetase
53272	Phosphoribosyltransferases (PRTases)
52584	Phosphoribulokinase/pantothenate kinase
50755	Phosphotyrosine-binding domain (PTB)
90198	Plant C2H2 finger (QALGGH zinc finger)
49563	PLC-like (P variant)
50730	Pleckstrin-homology domain (PH domain)
47770	Pointed domain
102856	Polo-box duplicated region
47414	POU-specific domain
55798	PR-1-like
82472	Precorrin-6Y methyltransferase (CbiT)
54815	Prokaryotic type KH domain (KH-domain type II)
56251	Proteasome subunits
57890	Protein kinase cysteine-rich domain (cys2, phorbol-binding domain)
88854	Protein kinases, catalytic subunit
48246	Protein prenyltransferases
56310	Protein serine/threonine phosphatase
81605	Protein serine/threonine phosphatase 2C, catalytic domain
53354	Protein-L-isoaspartyl O-methyltransferase
56301	Purple acid phosphatase
49364	Purple acid phosphatase, N-terminal domain
64269	PX domain
57864	Pyk2-associated protein beta ARF-GAP domain
51622	Pyruvate kinase
50801	Pyruvate kinase beta-barrel domain
52936	Pyruvate kinase, C-terminal domain
52002	R1 subunit of ribonucleotide reductase, C-terminal domain
48169	R1 subunit of ribonucleotide reductase, N-terminal domain
82709	R3H domain
111348	Rap/Ran-GAP (Pfam 02145)
48367	Ras GEF
54263	Ras-binding domain, RBD
52670	RecA protein-like (ATPase-domain)
48098	Regulator of G-protein signaling, RGS
63614	Regulatory subunit H of the V-type ATPase
81320	Rhodopsin-like

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
47253	Ribonucleotide reductase-like
110325	Ribosomal L27 protein
52162	Ribosomal protein L13
54822	Ribosomal protein S3 C-terminal domain
50371	Ricin B-like
57851	RING finger domain, C3HC4
52724	RNA helicase
64490	RNA-polymerase beta-prime
111454	RpoE2-like (Pfam 04035)
110843	RWD domain (Pfam 05773)
47478	S100 proteins
52300	S-adenosylhomocystein hydrolase
47773	SAM (sterile alpha motif) domain
63764	SAND domain
68907	SAP domain
49448	Second domain of Mu2 adaptin subunit (ap50) of ap2 adaptor
53499	Serine carboxypeptidase-like
55551	SH2 domain
50045	SH3-domain
54863	Short-chain ferredoxins
69198	SIAH, seven in absentia homolog
50263	Single strand DNA-binding domain, SSB
81380	Skp1 dimerisation domain-like
49880	SMAD domain
56367	SMAD MH1 domain
57363	Small Kunitz-type inhibitors & BPTI-like toxins
46967	Spectrin repeat
49855	Spermadhesin, CUB domain
69557	Spermidine synthase
64276	Splicing factor U2AF subunits
53449	Spore coat polysaccharide biosynthesis protein SpsA
55456	SRF-like
47656	STAT
81317	STAT DNA-binding domain
52211	Succinyl-CoA synthetase domains
56081	Succinyl-CoA synthetase, beta-chain, N-terminal domain
49575	Synaptotagmin-like (S variant)
64357	Synatpobrevin N-terminal domain
81268	Tandem AAA-ATPase domain

Table E.1. Domain descriptions (continued)

Domain ID	Domain Description
57934	TAZ domain
47134	TBP-associated factors, TAFs
54701	Tetramerization domain of potassium channels
48453	Tetratricopeptide repeat (TPR)
57899	TFIIH p44 subunit cysteine-rich domain
52918	Thioredoxin-like 2Fe-2S ferredoxin
57611	Thyroglobulin type-1 domain
57587	TNF receptor-like
52201	Toll/Interleukin receptor TIR domain
49600	TRAF domain
48671	Transducin (heterotrimeric G protein), gamma chain
54212	Translational machinery components
82458	Trunk domain of Sec23/24
53687	Tryptophan synthase beta subunit-like PLP-dependent enzymes
82896	TSP-1 type 1 repeat
55308	Tubulin, C-terminal domain
52491	Tubulin, GTPase domain
51191	Type I phosphomannose isomerase
56720	Type II DNA topoisomerase
51751	Tyrosine-dependent oxidoreductases
46935	UBA domain
110671	UbiE/COQ5-like (Pfam 01209)
89763	Ubiquitin activating enzymes (UBA)
54237	Ubiquitin-related
54250	UBX domain
50678	ValRS/IleRS/LeuRS editing domain
48468	VHS domain
49965	Vibrio cholerae sialidase, N-terminal and insertion domains
81323	Voltage-gated potassium channels
50979	WD40-repeat
51046	WW domain
82583	YchF GTP-binding protein, C-terminal domain

Table E.2. Domain-domain interactions of the reconstructed network

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
50157-50157	31	1.2272	54696-49786	1	0.0396	48404-47577	1	0.0396
57736-50157	24	0.9501	49786-54696	1	0.0396	48404-47547	1	0.0396
57721-57721	20	0.7918	54696-47617	1	0.0396	48404-47502	1	0.0396
46967-54929	20	0.7918	49786-52862	1	0.0396	100943-100940	1	0.0396
48509-48509	19	0.7522	54696-52862	1	0.0396	100943-100959	1	0.0396
48509-57721	19	0.7522	49786-47617	1	0.0396	100962-100940	1	0.0396
57721-48509	19	0.7522	49786-52724	1	0.0396	100962-100959	1	0.0396
88854-46690	18	0.7126	49786-81268	1	0.0396	100943-51070	1	0.0396
54929-54929	16	0.6334	57354-47918	1	0.0396	100962-51070	1	0.0396
57743-50157	16	0.6334	57354-50616	1	0.0396	100943-102755	1	0.0396
46967-46967	16	0.6334	57354-52670	1	0.0396	100943-55875	1	0.0396
88854-88854	12	0.4751	57354-55875	1	0.0396	100943-110943	1	0.0396
54792-54929	12	0.4751	57354-110943	1	0.0396	100962-102755	1	0.0396
47547-46967	12	0.4751	57354-102755	1	0.0396	100962-55875	1	0.0396
110080-57197	12	0.4751	64269-64269	1	0.0396	100962-110943	1	0.0396
46967-88854	12	0.4751	48509-69236	1	0.0396	81332-53301	1	0.0396
54792-57363	11	0.4355	57721-63888	1	0.0396	81332-69180	1	0.0396
57363-102755	11	0.4355	48509-63888	1	0.0396	81332-69688	1	0.0396
57363-55875	11	0.4355	57721-69236	1	0.0396	81332-52592	1	0.0396
57363-110943	11	0.4355	48509-51570	1	0.0396	81332-50466	1	0.0396
57363-57721	11	0.4355	57721-51570	1	0.0396	81332-50448	1	0.0396
57363-48509	11	0.4355	52201-57721	1	0.0396	49880-48246	1	0.0396
50157-54929	11	0.4355	69433-48509	1	0.0396	56367-48246	1	0.0396
88854-57721	10	0.3959	69433-57721	1	0.0396	49880-54929	1	0.0396
88854-54929	10	0.3959	52201-48509	1	0.0396	56367-54929	1	0.0396
88854-50157	9	0.3563	52048-46690	1	0.0396	49880-56310	1	0.0396
54792-54792	9	0.3563	48372-46690	1	0.0396	56367-56310	1	0.0396
47547-54929	9	0.3563	52048-101274	1	0.0396	49880-50979	1	0.0396
88854-48509	8	0.3167	48372-102932	1	0.0396	56367-50979	1	0.0396
81269-81269	8	0.3167	52048-102932	1	0.0396	49880-51396	1	0.0396
52592-50157	8	0.3167	48372-101274	1	0.0396	49880-46553	1	0.0396
56251-46967	8	0.3167	57721-110513	1	0.0396	49880-51972	1	0.0396
57736-54929	8	0.3167	57721-102755	1	0.0396	49880-54884	1	0.0396
57743-54929	8	0.3167	57721-55875	1	0.0396	56367-51396	1	0.0396
46967-47502	8	0.3167	57721-110943	1	0.0396	56367-46553	1	0.0396
46967-54701	8	0.3167	57721-52592	1	0.0396	56367-51972	1	0.0396
54929-48393	8	0.3167	57721-50466	1	0.0396	56367-54884	1	0.0396
88854-90230	7	0.2771	57721-50448	1	0.0396	49880-81268	1	0.0396
57851-57185	6	0.2375	57721-50979	1	0.0396	49880-52724	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
48453-57185	6	0.2375	50979-81269	1	0.0396	56367-81268	1	0.0396
54792-82896	6	0.2375	50979-48020	1	0.0396	56367-52724	1	0.0396
56251-81269	6	0.2375	50755-50755	1	0.0396	49880-81312	1	0.0396
56251-56251	6	0.2375	50755-88854	1	0.0396	56367-81312	1	0.0396
82896-102755	6	0.2375	50755-52541	1	0.0396	49880-57717	1	0.0396
82896-55875	6	0.2375	50755-50045	1	0.0396	56367-57717	1	0.0396
82896-110943	6	0.2375	50730-50157	1	0.0396	49880-102755	1	0.0396
82896-57721	6	0.2375	50730-52724	1	0.0396	49880-55875	1	0.0396
82896-48509	6	0.2375	88854-52724	1	0.0396	49880-110943	1	0.0396
88854-47221	6	0.2375	50730-54929	1	0.0396	56367-102755	1	0.0396
47096-57668	6	0.2375	88854-46739	1	0.0396	56367-55875	1	0.0396
57851-57611	6	0.2375	50730-46739	1	0.0396	56367-110943	1	0.0396
90230-57611	6	0.2375	49448-57721	1	0.0396	49880-46690	1	0.0396
110080-90194	6	0.2375	75521-48509	1	0.0396	56367-46690	1	0.0396
81279-57197	6	0.2375	49448-48509	1	0.0396	49880-51570	1	0.0396
50466-50157	6	0.2375	75521-57721	1	0.0396	56367-51570	1	0.0396
50448-50157	6	0.2375	54929-50282	1	0.0396	49880-49880	1	0.0396
50157-82896	6	0.2375	48453-110513	1	0.0396	49880-56367	1	0.0396
50730-90230	5	0.1979	57851-110513	1	0.0396	56367-49880	1	0.0396
57736-57736	5	0.1979	48453-90198	1	0.0396	56367-56367	1	0.0396
57736-57743	5	0.1979	57851-90198	1	0.0396	52641-46690	1	0.0396
50157-57851	5	0.1979	48453-54002	1	0.0396	49885-46690	1	0.0396
57851-54929	5	0.1979	57851-54002	1	0.0396	81269-81320	1	0.0396
57851-57257	5	0.1979	48453-53301	1	0.0396	81269-55682	1	0.0396
90230-57257	5	0.1979	57851-53301	1	0.0396	81269-52955	1	0.0396
52592-57668	5	0.1979	48453-69198	1	0.0396	81269-54696	1	0.0396
50466-57668	5	0.1979	57851-57851	1	0.0396	81269-49600	1	0.0396
50448-57668	5	0.1979	48453-57851	1	0.0396	51751-88741	1	0.0396
50157-63483	5	0.1979	57851-69198	1	0.0396	51751-52926	1	0.0396
57925-57851	5	0.1979	48453-88854	1	0.0396	52592-49448	1	0.0396
47577-54929	5	0.1979	57851-88854	1	0.0396	50466-49448	1	0.0396
47502-54929	5	0.1979	54792-75521	1	0.0396	50448-49448	1	0.0396
88854-102713	4	0.1584	54792-49159	1	0.0396	52592-75521	1	0.0396
88854-52592	4	0.1584	57911-52686	1	0.0396	50466-75521	1	0.0396
88854-57911	4	0.1584	57911-54002	1	0.0396	50448-75521	1	0.0396
57851-57721	4	0.1584	57911-47478	1	0.0396	52592-81269	1	0.0396
57851-48509	4	0.1584	47041-46690	1	0.0396	50466-81269	1	0.0396
54696-54696	4	0.1584	57934-46690	1	0.0396	50448-81269	1	0.0396
54696-81268	4	0.1584	48265-50157	1	0.0396	52592-101899	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
54696-52724	4	0.1584	52670-53499	1	0.0396	50466-101899	1	0.0396
52059-53068	4	0.1584	47918-53499	1	0.0396	50448-101899	1	0.0396
81268-90230	4	0.1584	50616-53499	1	0.0396	52592-110513	1	0.0396
50157-52541	4	0.1584	52670-46690	1	0.0396	50466-110513	1	0.0396
50157-57721	4	0.1584	47918-46690	1	0.0396	50448-110513	1	0.0396
57890-53687	4	0.1584	50616-46690	1	0.0396	52592-100908	1	0.0396
57736-47478	4	0.1584	52670-52103	1	0.0396	50466-100908	1	0.0396
46748-51046	4	0.1584	47918-52103	1	0.0396	50448-100908	1	0.0396
46690-51046	4	0.1584	50616-52103	1	0.0396	52592-82458	1	0.0396
49880-52592	4	0.1584	49159-102755	1	0.0396	50466-82458	1	0.0396
56367-52592	4	0.1584	49159-55875	1	0.0396	50448-82458	1	0.0396
88854-57851	4	0.1584	49159-110943	1	0.0396	52592-57899	1	0.0396
46832-46967	4	0.1584	49159-57721	1	0.0396	50466-57899	1	0.0396
57721-51046	4	0.1584	49159-48509	1	0.0396	50448-57899	1	0.0396
48509-51046	4	0.1584	48020-48020	1	0.0396	52592-81332	1	0.0396
47547-47502	4	0.1584	52592-47617	1	0.0396	50466-81332	1	0.0396
47547-50157	4	0.1584	52592-51570	1	0.0396	50448-81332	1	0.0396
88854-51046	4	0.1584	54701-51570	1	0.0396	52592-54929	1	0.0396
50730-51046	4	0.1584	54701-47478	1	0.0396	50466-54929	1	0.0396
52592-52592	4	0.1584	74933-54586	1	0.0396	50448-54929	1	0.0396
52641-57851	4	0.1584	74933-50708	1	0.0396	52592-111375	1	0.0396
55551-46690	4	0.1584	81605-46690	1	0.0396	50466-111375	1	0.0396
47669-54929	4	0.1584	81605-57721	1	0.0396	50448-111375	1	0.0396
55551-54929	4	0.1584	81605-48509	1	0.0396	52592-47773	1	0.0396
48404-46967	4	0.1584	81605-55935	1	0.0396	50466-47773	1	0.0396
52592-54701	4	0.1584	81605-48035	1	0.0396	50448-47773	1	0.0396
50466-54701	4	0.1584	81605-47114	1	0.0396	52592-52541	1	0.0396
50448-54701	4	0.1584	102755-102755	1	0.0396	50466-52541	1	0.0396
63483-50157	4	0.1584	55875-102755	1	0.0396	50448-52541	1	0.0396
50157-52592	4	0.1584	110943-102755	1	0.0396	52592-69594	1	0.0396
57736-46967	4	0.1584	102755-55875	1	0.0396	50466-69594	1	0.0396
57743-46967	4	0.1584	55875-55875	1	0.0396	50448-69594	1	0.0396
50157-46967	4	0.1584	110943-55875	1	0.0396	52592-50646	1	0.0396
46967-81269	4	0.1584	102755-110943	1	0.0396	50466-50646	1	0.0396
46967-47617	4	0.1584	55875-110943	1	0.0396	50448-50646	1	0.0396
46967-52862	4	0.1584	110943-110943	1	0.0396	52592-47478	1	0.0396
46967-48671	4	0.1584	63955-54929	1	0.0396	50466-47478	1	0.0396
46967-81268	4	0.1584	63955-52592	1	0.0396	50448-47478	1	0.0396
46967-52724	4	0.1584	48404-50730	1	0.0396	63483-101899	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
46967-64357	4	0.1584	63955-50730	1	0.0396	50157-101899	1	0.0396
46967-57185	4	0.1584	48404-52592	1	0.0396	63483-110513	1	0.0396
46967-57197	4	0.1584	63955-51191	1	0.0396	50157-110513	1	0.0396
46967-47577	4	0.1584	48404-51191	1	0.0396	63483-52641	1	0.0396
47577-46967	4	0.1584	63955-57851	1	0.0396	50157-52641	1	0.0396
47502-46967	4	0.1584	54696-56535	1	0.0396	63483-53558	1	0.0396
46967-47547	4	0.1584	49600-55028	1	0.0396	50157-53558	1	0.0396
55551-52686	4	0.1584	54696-55028	1	0.0396	63483-54586	1	0.0396
89838-57911	4	0.1584	49600-56535	1	0.0396	63483-50708	1	0.0396
54929-50990	4	0.1584	47114-46690	1	0.0396	50157-54586	1	0.0396
54929-48390	4	0.1584	54701-69236	1	0.0396	50157-50708	1	0.0396
54929-102755	4	0.1584	54701-63888	1	0.0396	63483-88854	1	0.0396
54929-55875	4	0.1584	54701-54701	1	0.0396	63483-102856	1	0.0396
54929-110943	4	0.1584	51070-111375	1	0.0396	50157-102856	1	0.0396
54929-64276	4	0.1584	88854-47617	1	0.0396	63483-49600	1	0.0396
54929-90230	4	0.1584	49575-47617	1	0.0396	63483-54696	1	0.0396
55730-50157	3	0.1188	88854-52862	1	0.0396	63483-75521	1	0.0396
88854-103184	3	0.1188	49575-52862	1	0.0396	50157-75521	1	0.0396
48509-81268	3	0.1188	49575-50157	1	0.0396	63483-47502	1	0.0396
57721-81268	3	0.1188	48509-69198	1	0.0396	63483-54822	1	0.0396
88854-57668	3	0.1188	57721-69198	1	0.0396	63483-54815	1	0.0396
88854-102755	3	0.1188	51570-51570	1	0.0396	50157-54822	1	0.0396
88854-55875	3	0.1188	57904-89763	1	0.0396	50157-54815	1	0.0396
88854-110943	3	0.1188	57904-57721	1	0.0396	56310-57851	1	0.0396
48404-54929	3	0.1188	57904-48509	1	0.0396	56310-55621	1	0.0396
88854-52670	3	0.1188	101888-52652	1	0.0396	57743-57736	1	0.0396
88854-50616	3	0.1188	69594-57721	1	0.0396	57743-57743	1	0.0396
88854-47918	3	0.1188	69594-48509	1	0.0396	50157-57736	1	0.0396
52059-57721	3	0.1188	64225-50157	1	0.0396	50157-57743	1	0.0396
52059-48509	3	0.1188	48351-57668	1	0.0396	57736-111348	1	0.0396
50755-50157	3	0.1188	48351-48509	1	0.0396	57743-111348	1	0.0396
50157-88854	3	0.1188	48351-57721	1	0.0396	50157-111348	1	0.0396
88854-81268	3	0.1188	46748-46690	1	0.0396	57736-90230	1	0.0396
50730-81268	3	0.1188	46748-56310	1	0.0396	57743-90230	1	0.0396
57743-46690	3	0.1188	46690-56310	1	0.0396	50157-90230	1	0.0396
74933-81269	3	0.1188	46748-50157	1	0.0396	57736-52592	1	0.0396
57890-50157	3	0.1188	46690-50157	1	0.0396	57736-50466	1	0.0396
81269-46690	3	0.1188	46748-48468	1	0.0396	57736-50448	1	0.0396
54792-90230	3	0.1188	46690-48468	1	0.0396	57743-52592	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
54792-90214	3	0.1188	46748-57904	1	0.0396	57743-50466	1	0.0396
50157-48509	3	0.1188	46690-57904	1	0.0396	57743-50448	1	0.0396
46832-54929	3	0.1188	46748-81268	1	0.0396	50157-50466	1	0.0396
47547-47577	3	0.1188	46690-81268	1	0.0396	50157-50448	1	0.0396
47547-47547	3	0.1188	46748-52724	1	0.0396	57736-46690	1	0.0396
47547-81269	3	0.1188	46690-52724	1	0.0396	57736-102755	1	0.0396
56748-54929	3	0.1188	46748-54701	1	0.0396	57736-55875	1	0.0396
81279-90194	3	0.1188	46748-54929	1	0.0396	57736-110943	1	0.0396
52641-54929	3	0.1188	46690-54929	1	0.0396	57743-102755	1	0.0396
50263-50263	3	0.1188	46748-57851	1	0.0396	57743-55875	1	0.0396
46844-50263	3	0.1188	46690-57851	1	0.0396	57743-110943	1	0.0396
52592-63483	3	0.1188	46748-56205	1	0.0396	50157-102755	1	0.0396
50466-63483	3	0.1188	46748-49563	1	0.0396	50157-55875	1	0.0396
50448-63483	3	0.1188	46690-56205	1	0.0396	50157-110943	1	0.0396
52592-57851	3	0.1188	46690-49563	1	0.0396	57736-50282	1	0.0396
50466-57851	3	0.1188	46748-82709	1	0.0396	57743-50282	1	0.0396
50448-57851	3	0.1188	46690-82709	1	0.0396	50157-50282	1	0.0396
63483-63483	3	0.1188	46690-46739	1	0.0396	57736-47577	1	0.0396
63483-81269	3	0.1188	52541-63764	1	0.0396	57736-47547	1	0.0396
50157-81269	3	0.1188	50045-63764	1	0.0396	57736-47502	1	0.0396
50157-54696	3	0.1188	50157-63764	1	0.0396	57743-47577	1	0.0396
57736-51046	3	0.1188	52541-103184	1	0.0396	57743-47547	1	0.0396
57743-51046	3	0.1188	50045-103184	1	0.0396	57743-47502	1	0.0396
50157-51046	3	0.1188	50157-103184	1	0.0396	50157-47577	1	0.0396
47134-48509	3	0.1188	88854-82764	1	0.0396	50157-47547	1	0.0396
47134-57721	3	0.1188	52541-82764	1	0.0396	57736-49885	1	0.0396
57925-48509	3	0.1188	50045-82764	1	0.0396	57736-54263	1	0.0396
57925-57721	3	0.1188	50157-82764	1	0.0396	57743-49885	1	0.0396
47577-88854	3	0.1188	52541-54929	1	0.0396	57743-54263	1	0.0396
47547-88854	3	0.1188	50045-54929	1	0.0396	50157-49885	1	0.0396
47502-88854	3	0.1188	52541-50157	1	0.0396	50157-54263	1	0.0396
88854-89851	2	0.0792	50045-50157	1	0.0396	57736-81268	1	0.0396
88854-46935	2	0.0792	88854-110325	1	0.0396	57736-52724	1	0.0396
88854-48453	2	0.0792	52541-110325	1	0.0396	57743-81268	1	0.0396
57721-52641	2	0.0792	50045-110325	1	0.0396	57743-52724	1	0.0396
57721-52724	2	0.0792	50157-110325	1	0.0396	50157-81268	1	0.0396
48509-52724	2	0.0792	88854-110671	1	0.0396	50157-52724	1	0.0396
57736-102713	2	0.0792	52541-110671	1	0.0396	57736-46951	1	0.0396
57736-50755	2	0.0792	50045-110671	1	0.0396	57743-46951	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
56251-110513	2	0.0792	50157-110671	1	0.0396	50157-46951	1	0.0396
56251-57721	2	0.0792	47820-88854	1	0.0396	57736-55682	1	0.0396
56251-48509	2	0.0792	52584-90214	1	0.0396	57743-55682	1	0.0396
48404-57851	2	0.0792	53272-90214	1	0.0396	50157-55682	1	0.0396
81312-54929	2	0.0792	52592-63764	1	0.0396	57736-52955	1	0.0396
52201-53068	2	0.0792	54250-47617	1	0.0396	57743-52955	1	0.0396
69433-53068	2	0.0792	54250-52862	1	0.0396	50157-52955	1	0.0396
81268-63483	2	0.0792	54250-57721	1	0.0396	57736-82709	1	0.0396
81268-50157	2	0.0792	54250-48509	1	0.0396	57743-82709	1	0.0396
81268-81320	2	0.0792	54250-110020	1	0.0396	50157-82709	1	0.0396
81268-81380	2	0.0792	81269-52592	1	0.0396	47134-51375	1	0.0396
81268-54696	2	0.0792	81269-50466	1	0.0396	47134-53272	1	0.0396
81268-55935	2	0.0792	81269-50448	1	0.0396	47134-110734	1	0.0396
81268-48035	2	0.0792	49880-46739	1	0.0396	47134-47134	1	0.0396
81268-48509	2	0.0792	56367-46739	1	0.0396	46832-53522	1	0.0396
81268-57721	2	0.0792	49880-54212	1	0.0396	46832-46690	1	0.0396
57721-46690	2	0.0792	49880-50448	1	0.0396	47096-46690	1	0.0396
57721-54701	2	0.0792	56367-54212	1	0.0396	54863-57851	1	0.0396
50157-50755	2	0.0792	56367-50448	1	0.0396	54929-57851	1	0.0396
50157-50045	2	0.0792	49880-55857	1	0.0396	54929-54792	1	0.0396
48453-50157	2	0.0792	56367-55857	1	0.0396	81320-55621	1	0.0396
57851-50157	2	0.0792	49880-53201	1	0.0396	81320-88854	1	0.0396
48453-57425	2	0.0792	56367-53201	1	0.0396	81320-55551	1	0.0396
57851-57425	2	0.0792	56251-52670	1	0.0396	81320-57721	1	0.0396
48453-48509	2	0.0792	56251-47918	1	0.0396	57925-110513	1	0.0396
48453-57721	2	0.0792	56251-50616	1	0.0396	57925-102713	1	0.0396
57911-50157	2	0.0792	51622-57721	1	0.0396	57925-50282	1	0.0396
57911-63483	2	0.0792	52936-57721	1	0.0396	57925-69198	1	0.0396
50157-46690	2	0.0792	50801-57721	1	0.0396	47617-53499	1	0.0396
56251-54586	2	0.0792	51622-48509	1	0.0396	52862-53499	1	0.0396
56251-50708	2	0.0792	52936-48509	1	0.0396	47617-50966	1	0.0396
81269-48020	2	0.0792	50801-48509	1	0.0396	47617-49855	1	0.0396
48020-81269	2	0.0792	52059-46739	1	0.0396	47617-57233	1	0.0396
49600-54696	2	0.0792	50157-46739	1	0.0396	52862-50966	1	0.0396
54696-49600	2	0.0792	52059-54701	1	0.0396	52862-49855	1	0.0396
49600-49600	2	0.0792	52059-69198	1	0.0396	52862-57233	1	0.0396
49600-81268	2	0.0792	52059-57851	1	0.0396	47617-81269	1	0.0396
49600-52724	2	0.0792	52375-46690	1	0.0396	52862-81269	1	0.0396
54701-50157	2	0.0792	47324-46690	1	0.0396	47617-88854	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
47617-47617	2	0.0792	55191-46690	1	0.0396	52862-88854	1	0.0396
52862-47617	2	0.0792	50730-46690	1	0.0396	47617-100948	1	0.0396
47617-52862	2	0.0792	88854-56310	1	0.0396	52862-100948	1	0.0396
52862-52862	2	0.0792	50730-56310	1	0.0396	48351-56301	1	0.0396
56251-50157	2	0.0792	50730-81269	1	0.0396	48351-49364	1	0.0396
57890-47617	2	0.0792	82194-46690	1	0.0396	52088-56301	1	0.0396
57890-52862	2	0.0792	82194-54701	1	0.0396	52088-49364	1	0.0396
88854-53687	2	0.0792	81300-52778	1	0.0396	102620-81268	1	0.0396
49575-53687	2	0.0792	81300-51231	1	0.0396	102620-52724	1	0.0396
48509-57851	2	0.0792	81300-47006	1	0.0396	102620-54929	1	0.0396
57721-57851	2	0.0792	101274-52778	1	0.0396	47577-81269	1	0.0396
47617-46690	2	0.0792	101274-51231	1	0.0396	47502-81269	1	0.0396
52862-46690	2	0.0792	101274-47006	1	0.0396	47577-47617	1	0.0396
46690-46690	2	0.0792	81300-50157	1	0.0396	47547-47617	1	0.0396
46748-57668	2	0.0792	101274-50157	1	0.0396	47502-47617	1	0.0396
46690-57668	2	0.0792	81300-53384	1	0.0396	47577-52862	1	0.0396
46690-54701	2	0.0792	101274-53384	1	0.0396	47547-52862	1	0.0396
46748-57721	2	0.0792	81300-54929	1	0.0396	47502-52862	1	0.0396
46690-57721	2	0.0792	101274-54929	1	0.0396	47577-48671	1	0.0396
46748-48509	2	0.0792	81300-69198	1	0.0396	47547-48671	1	0.0396
46690-48509	2	0.0792	101274-69198	1	0.0396	47502-48671	1	0.0396
88854-63764	2	0.0792	81300-57851	1	0.0396	47577-81268	1	0.0396
81268-88854	2	0.0792	101274-57851	1	0.0396	47547-81268	1	0.0396
54250-81269	2	0.0792	81300-52641	1	0.0396	47502-81268	1	0.0396
54250-54586	2	0.0792	101274-52641	1	0.0396	47577-52724	1	0.0396
54250-50708	2	0.0792	81300-54696	1	0.0396	47547-52724	1	0.0396
81269-56251	2	0.0792	101274-54696	1	0.0396	47502-52724	1	0.0396
49880-54981	2	0.0792	81300-49600	1	0.0396	47577-64357	1	0.0396
56367-54981	2	0.0792	101274-49600	1	0.0396	47547-64357	1	0.0396
49880-82424	2	0.0792	88854-81320	1	0.0396	47502-64357	1	0.0396
56367-82424	2	0.0792	81269-50157	1	0.0396	47577-57185	1	0.0396
50157-54701	2	0.0792	81269-102755	1	0.0396	47547-57185	1	0.0396
50157-69198	2	0.0792	81269-55875	1	0.0396	47502-57185	1	0.0396
88854-81269	2	0.0792	81269-110943	1	0.0396	47577-57197	1	0.0396
81269-57721	2	0.0792	57851-102713	1	0.0396	47547-57197	1	0.0396
81269-48509	2	0.0792	57851-100948	1	0.0396	47502-57197	1	0.0396
57851-47502	2	0.0792	57851-54701	1	0.0396	47577-47577	1	0.0396
88854-47502	2	0.0792	100948-48558	1	0.0396	47502-47577	1	0.0396
88854-63483	2	0.0792	100948-100948	1	0.0396	47577-47547	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
57721-52686	2	0.0792	47617-54701	1	0.0396	47502-47547	1	0.0396
48509-52686	2	0.0792	52862-54701	1	0.0396	52724-82509	1	0.0396
57721-48404	2	0.0792	47617-57721	1	0.0396	81268-82509	1	0.0396
48509-48404	2	0.0792	52862-57721	1	0.0396	52724-52670	1	0.0396
57721-53068	2	0.0792	47617-48509	1	0.0396	81268-52670	1	0.0396
48509-53068	2	0.0792	52862-48509	1	0.0396	52724-47795	1	0.0396
47096-52641	2	0.0792	46832-46739	1	0.0396	81268-47795	1	0.0396
47096-68907	2	0.0792	46832-52162	1	0.0396	52724-53499	1	0.0396
47547-57743	2	0.0792	46832-47577	1	0.0396	81268-53499	1	0.0396
47547-56318	2	0.0792	46832-47547	1	0.0396	52724-52592	1	0.0396
47547-54198	2	0.0792	46832-47502	1	0.0396	52724-54212	1	0.0396
47547-64159	2	0.0792	46832-57851	1	0.0396	52724-50448	1	0.0396
47547-64162	2	0.0792	53640-51751	1	0.0396	81268-52592	1	0.0396
47547-48257	2	0.0792	53640-52641	1	0.0396	81268-54212	1	0.0396
47547-56081	2	0.0792	88854-52468	1	0.0396	81268-50448	1	0.0396
47547-52211	2	0.0792	88854-100943	1	0.0396	88854-51751	1	0.0396
47547-51900	2	0.0792	88854-100962	1	0.0396	88854-52778	1	0.0396
47547-52541	2	0.0792	88854-75521	1	0.0396	88854-51231	1	0.0396
47547-46847	2	0.0792	88854-111348	1	0.0396	88854-47006	1	0.0396
47547-64269	2	0.0792	88854-57864	1	0.0396	81380-48620	1	0.0396
88854-63412	2	0.0792	88854-48404	1	0.0396	54696-48620	1	0.0396
63492-53068	2	0.0792	88854-53522	1	0.0396	81380-49965	1	0.0396
54237-53068	2	0.0792	57668-57851	1	0.0396	81380-57197	1	0.0396
57890-47502	2	0.0792	57668-50157	1	0.0396	54696-49965	1	0.0396
57851-56720	2	0.0792	57668-50979	1	0.0396	54696-57197	1	0.0396
90230-56720	2	0.0792	56310-46690	1	0.0396	81380-57721	1	0.0396
57851-53089	2	0.0792	54002-46690	1	0.0396	81380-48509	1	0.0396
90230-53089	2	0.0792	57721-101899	1	0.0396	53272-53272	1	0.0396
57851-57263	2	0.0792	48509-101899	1	0.0396	53272-54002	1	0.0396
90230-57263	2	0.0792	57721-54237	1	0.0396	46739-46739	1	0.0396
56251-47577	2	0.0792	48509-54237	1	0.0396	46739-63912	1	0.0396
56251-47547	2	0.0792	57721-88854	1	0.0396	48098-57736	1	0.0396
56251-47502	2	0.0792	48509-88854	1	0.0396	48098-57743	1	0.0396
88854-54701	2	0.0792	57721-56081	1	0.0396	48098-50157	1	0.0396
110080-102713	2	0.0792	48509-56081	1	0.0396	48098-56535	1	0.0396
110080-57668	2	0.0792	57721-52211	1	0.0396	48098-55028	1	0.0396
110080-48404	2	0.0792	48509-52211	1	0.0396	57721-90198	1	0.0396
81312-46690	2	0.0792	57721-54696	1	0.0396	48509-90198	1	0.0396
57851-69573	2	0.0792	48509-54696	1	0.0396	57721-57668	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
50263-48393	2	0.0792	57721-56205	1	0.0396	48509-57668	1	0.0396
46844-48393	2	0.0792	57721-49563	1	0.0396	54792-81323	1	0.0396
50263-57721	2	0.0792	48509-56205	1	0.0396	54792-100908	1	0.0396
50263-48509	2	0.0792	48509-49563	1	0.0396	82200-55798	1	0.0396
48404-48393	2	0.0792	57721-81312	1	0.0396	82200-69236	1	0.0396
47617-50157	2	0.0792	48509-81312	1	0.0396	82200-63888	1	0.0396
52862-50157	2	0.0792	48509-54701	1	0.0396	82200-47478	1	0.0396
51210-48509	2	0.0792	57721-111454	1	0.0396	56251-64269	1	0.0396
51210-53417	2	0.0792	48509-111454	1	0.0396	52686-57721	1	0.0396
51210-53499	2	0.0792	57721-64490	1	0.0396	54884-57721	1	0.0396
51210-57721	2	0.0792	48509-64490	1	0.0396	52686-48509	1	0.0396
51210-51699	2	0.0792	57721-89743	1	0.0396	54884-48509	1	0.0396
51210-50136	2	0.0792	48509-89743	1	0.0396	52670-81268	1	0.0396
51210-51736	2	0.0792	57721-100921	1	0.0396	47795-81268	1	0.0396
81332-69940	2	0.0792	57721-100935	1	0.0396	52670-102396	1	0.0396
49880-48404	2	0.0792	48509-100921	1	0.0396	47795-102396	1	0.0396
56367-48404	2	0.0792	48509-100935	1	0.0396	52670-52670	1	0.0396
49880-54701	2	0.0792	57721-69496	1	0.0396	47795-52670	1	0.0396
56367-54701	2	0.0792	48509-69496	1	0.0396	52670-47795	1	0.0396
81268-56720	2	0.0792	52670-57904	1	0.0396	47795-47795	1	0.0396
81268-55879	2	0.0792	52670-57851	1	0.0396	55551-57721	1	0.0396
81268-54224	2	0.0792	88854-111353	1	0.0396	81380-50157	1	0.0396
52592-50466	2	0.0792	48385-46690	1	0.0396	54696-50157	1	0.0396
52592-50448	2	0.0792	52918-46690	1	0.0396	49600-47114	1	0.0396
50448-52592	2	0.0792	63492-100921	1	0.0396	54696-52103	1	0.0396
50448-50466	2	0.0792	63492-100935	1	0.0396	49600-52103	1	0.0396
50448-50448	2	0.0792	54237-100921	1	0.0396	49600-48509	1	0.0396
50466-52592	2	0.0792	54237-100935	1	0.0396	49600-57721	1	0.0396
50466-50466	2	0.0792	63492-57721	1	0.0396	81317-57736	1	0.0396
50466-50448	2	0.0792	54237-57721	1	0.0396	81317-57743	1	0.0396
52592-88854	2	0.0792	49563-47502	1	0.0396	81317-50157	1	0.0396
50466-88854	2	0.0792	56748-63571	1	0.0396	47656-57736	1	0.0396
50448-88854	2	0.0792	52491-46690	1	0.0396	47656-57743	1	0.0396
52592-57721	2	0.0792	55308-46690	1	0.0396	47656-50157	1	0.0396
50466-57721	2	0.0792	52491-52491	1	0.0396	55551-57736	1	0.0396
50448-57721	2	0.0792	52491-55308	1	0.0396	55551-57743	1	0.0396
52592-48509	2	0.0792	55308-52491	1	0.0396	55551-50157	1	0.0396
50466-48509	2	0.0792	55308-55308	1	0.0396	81317-102713	1	0.0396
50448-48509	2	0.0792	52491-63588	1	0.0396	47656-102713	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
52592-47547	2	0.0792	55308-63588	1	0.0396	55551-102713	1	0.0396
50466-47547	2	0.0792	52491-52641	1	0.0396	81317-46690	1	0.0396
50448-47547	2	0.0792	55308-52641	1	0.0396	47656-46690	1	0.0396
63483-54929	2	0.0792	81269-88854	1	0.0396	81317-100921	1	0.0396
63483-52592	2	0.0792	47773-46690	1	0.0396	81317-100935	1	0.0396
63483-57851	2	0.0792	88854-50466	1	0.0396	47656-100921	1	0.0396
50157-49600	2	0.0792	88854-50448	1	0.0396	47656-100935	1	0.0396
50157-47502	2	0.0792	47773-52592	1	0.0396	55551-100921	1	0.0396
57736-63483	2	0.0792	47773-50466	1	0.0396	55551-100935	1	0.0396
57743-63483	2	0.0792	47773-50448	1	0.0396	81317-53417	1	0.0396
57736-52541	2	0.0792	47773-52670	1	0.0396	47656-53417	1	0.0396
57743-52541	2	0.0792	47773-47918	1	0.0396	55551-53417	1	0.0396
54863-54792	2	0.0792	47773-50616	1	0.0396	81317-47875	1	0.0396
53068-46690	2	0.0792	82261-52513	1	0.0396	47656-47875	1	0.0396
53068-88854	2	0.0792	82227-52513	1	0.0396	55551-47875	1	0.0396
47577-47502	2	0.0792	82261-51622	1	0.0396	81317-89751	1	0.0396
47502-47502	2	0.0792	82261-52936	1	0.0396	47656-89751	1	0.0396
47577-54701	2	0.0792	82261-50801	1	0.0396	55551-89751	1	0.0396
47547-54701	2	0.0792	82227-51622	1	0.0396	81317-81317	1	0.0396
47502-54701	2	0.0792	82227-52936	1	0.0396	47656-81317	1	0.0396
52724-54981	2	0.0792	82227-50801	1	0.0396	55551-81317	1	0.0396
81268-54981	2	0.0792	88854-55456	1	0.0396	81317-47656	1	0.0396
81380-49944	2	0.0792	88854-63571	1	0.0396	47656-47656	1	0.0396
54696-49944	2	0.0792	88854-47875	1	0.0396	55551-47656	1	0.0396
54696-57721	2	0.0792	88854-53354	1	0.0396	81317-55551	1	0.0396
54696-48509	2	0.0792	50730-53354	1	0.0396	47656-55551	1	0.0396
55551-81268	2	0.0792	88854-56205	1	0.0396	55551-55551	1	0.0396
55551-52724	2	0.0792	88854-49563	1	0.0396	88854-55812	1	0.0396
55551-50282	2	0.0792	50730-56205	1	0.0396	47617-52002	1	0.0396
55551-90124	2	0.0792	50730-49563	1	0.0396	52862-52002	1	0.0396
55551-52670	2	0.0792	57851-46832	1	0.0396	47617-48169	1	0.0396
55551-47918	2	0.0792	90230-46832	1	0.0396	52862-48169	1	0.0396
55551-50616	2	0.0792	57851-47617	1	0.0396	69433-52300	1	0.0396
55551-56789	2	0.0792	57851-52862	1	0.0396	69433-51830	1	0.0396
54792-54701	2	0.0792	90230-47617	1	0.0396	88854-57736	1	0.0396
82200-90230	2	0.0792	90230-52862	1	0.0396	88854-57743	1	0.0396
54696-47114	2	0.0792	57851-55879	1	0.0396	50157-53640	1	0.0396
81317-53068	2	0.0792	57851-54224	1	0.0396	110325-53640	1	0.0396
47656-53068	2	0.0792	90230-55879	1	0.0396	75285-52103	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
55551-53068	2	0.0792	90230-54224	1	0.0396	75285-50045	1	0.0396
54696-46690	2	0.0792	57851-48265	1	0.0396	109671-46690	1	0.0396
88854-52300	2	0.0792	90230-48265	1	0.0396	109671-54696	1	0.0396
88854-51830	2	0.0792	57851-57363	1	0.0396	102492-46690	1	0.0396
89838-63483	2	0.0792	90230-57363	1	0.0396	102492-54929	1	0.0396
89838-50157	2	0.0792	57851-81320	1	0.0396	54696-51622	1	0.0396
89838-54929	2	0.0792	90230-81320	1	0.0396	54696-52936	1	0.0396
89838-54701	2	0.0792	57851-51751	1	0.0396	54696-50801	1	0.0396
89838-57587	2	0.0792	90230-51751	1	0.0396	54263-46690	1	0.0396
89838-52592	2	0.0792	52592-52491	1	0.0396	57890-46690	1	0.0396
89838-82583	2	0.0792	52592-55308	1	0.0396	54263-52300	1	0.0396
89838-51570	2	0.0792	52592-46739	1	0.0396	57890-52300	1	0.0396
89838-47096	2	0.0792	52592-69236	1	0.0396	54263-51830	1	0.0396
89838-46690	2	0.0792	52592-63888	1	0.0396	57890-51830	1	0.0396
55935-81268	2	0.0792	48509-102755	1	0.0396	88854-47158	1	0.0396
48035-81268	2	0.0792	48509-55875	1	0.0396	54263-47158	1	0.0396
88854-49494	2	0.0792	48509-110943	1	0.0396	57890-47158	1	0.0396
88854-53068	2	0.0792	48509-47478	1	0.0396	54263-103184	1	0.0396
57904-81268	2	0.0792	57721-47478	1	0.0396	57890-103184	1	0.0396
50730-54701	2	0.0792	88854-49600	1	0.0396	88854-52059	1	0.0396
57904-54701	2	0.0792	88854-54696	1	0.0396	54263-52059	1	0.0396
47502-53068	2	0.0792	57890-52592	1	0.0396	57890-52059	1	0.0396
52059-52592	1	0.0396	81279-102713	1	0.0396	88854-52162	1	0.0396
75142-52592	1	0.0396	81279-57668	1	0.0396	54263-52162	1	0.0396
55730-46690	1	0.0396	81279-48404	1	0.0396	57890-52162	1	0.0396
55730-48257	1	0.0396	57851-52491	1	0.0396	88854-57587	1	0.0396
55730-56081	1	0.0396	57851-55308	1	0.0396	88854-82583	1	0.0396
55730-52211	1	0.0396	54237-52491	1	0.0396	88854-51570	1	0.0396
55730-51900	1	0.0396	54237-55308	1	0.0396	88854-47096	1	0.0396
55730-52541	1	0.0396	57851-47478	1	0.0396	55935-54822	1	0.0396
55730-50045	1	0.0396	54237-47478	1	0.0396	48035-54822	1	0.0396
55730-101289	1	0.0396	48509-47414	1	0.0396	55935-54815	1	0.0396
88854-51183	1	0.0396	48509-46690	1	0.0396	48035-54815	1	0.0396
88854-53499	1	0.0396	57721-47414	1	0.0396	55935-52652	1	0.0396
88854-47460	1	0.0396	49448-63483	1	0.0396	48035-52652	1	0.0396
88854-110734	1	0.0396	49448-50157	1	0.0396	55935-102755	1	0.0396
48509-52641	1	0.0396	75521-63483	1	0.0396	48035-102755	1	0.0396
88854-47371	1	0.0396	75521-50157	1	0.0396	55935-55875	1	0.0396
88854-50282	1	0.0396	49448-52592	1	0.0396	48035-55875	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
69557-54929	1	0.0396	49448-50448	1	0.0396	55935-110943	1	0.0396
82472-54929	1	0.0396	49448-50466	1	0.0396	48035-110943	1	0.0396
69557-81320	1	0.0396	75521-52592	1	0.0396	88854-46566	1	0.0396
82472-81320	1	0.0396	75521-50448	1	0.0396	88854-57939	1	0.0396
69557-81850	1	0.0396	75521-50466	1	0.0396	88854-100921	1	0.0396
82472-81850	1	0.0396	52641-57736	1	0.0396	88854-100935	1	0.0396
69557-46739	1	0.0396	52641-57743	1	0.0396	50912-47478	1	0.0396
82472-46739	1	0.0396	52641-50157	1	0.0396	50912-53449	1	0.0396
56251-48257	1	0.0396	52641-52641	1	0.0396	53201-53201	1	0.0396
56251-56081	1	0.0396	52641-50979	1	0.0396	49880-102713	1	0.0396
56251-52211	1	0.0396	52641-47134	1	0.0396	56367-102713	1	0.0396
56251-51900	1	0.0396	52641-69198	1	0.0396	49880-57668	1	0.0396
47253-47253	1	0.0396	52641-54002	1	0.0396	56367-57668	1	0.0396
48404-49600	1	0.0396	47669-46690	1	0.0396	50730-47502	1	0.0396
81312-54696	1	0.0396	47547-46690	1	0.0396	57904-47502	1	0.0396
48404-54696	1	0.0396	57851-46690	1	0.0396	50730-47617	1	0.0396
81312-49600	1	0.0396	47669-69573	1	0.0396	57904-47617	1	0.0396
48404-102713	1	0.0396	55551-69573	1	0.0396	50730-52862	1	0.0396
81312-102713	1	0.0396	47547-69573	1	0.0396	57904-52862	1	0.0396
48404-101899	1	0.0396	47669-54250	1	0.0396	50730-47478	1	0.0396
81312-57851	1	0.0396	55551-54250	1	0.0396	57904-47478	1	0.0396
81312-101899	1	0.0396	47547-54250	1	0.0396	52822-46847	1	0.0396
48404-53733	1	0.0396	57851-54250	1	0.0396	52822-46690	1	0.0396
81312-52017	1	0.0396	47669-53272	1	0.0396	47502-47478	1	0.0396
48404-52017	1	0.0396	55551-53272	1	0.0396	52592-51046	1	0.0396
81312-53733	1	0.0396	47547-53272	1	0.0396	50157-56310	1	0.0396
48404-63614	1	0.0396	57851-53272	1	0.0396	50157-57668	1	0.0396
81312-63614	1	0.0396	50263-50990	1	0.0396	50157-55519	1	0.0396
48404-48404	1	0.0396	50263-48390	1	0.0396	50157-48257	1	0.0396
81312-81312	1	0.0396	46844-50990	1	0.0396	50157-56081	1	0.0396
48404-81312	1	0.0396	46844-48390	1	0.0396	50157-52211	1	0.0396
81312-48404	1	0.0396	46844-57721	1	0.0396	50157-51900	1	0.0396
48404-47478	1	0.0396	46844-48509	1	0.0396	50157-89751	1	0.0396
81312-47478	1	0.0396	88854-110513	1	0.0396	49600-46690	1	0.0396
54696-53449	1	0.0396	88854-56251	1	0.0396	54696-57851	1	0.0396
49786-50371	1	0.0396	88854-55932	1	0.0396	49600-57851	1	0.0396
54696-50371	1	0.0396	88854-54369	1	0.0396	54696-54929	1	0.0396
49786-53449	1	0.0396	48404-64269	1	0.0396	49600-54929	1	0.0396
54696-51127	1	0.0396	48404-52300	1	0.0396	54696-57668	1	0.0396

Table E.2. Domain-domain interactions of the reconstructed network (continued)

DDIs	Freq.	Freq %	DDIs	Freq.	Freq %	DDIs	Freq.	Freq %
49786-81381	1	0.0396	48404-51830	1	0.0396	49600-57668	1	0.0396
54696-81381	1	0.0396	48404-50990	1	0.0396			
49786-51127	1	0.0396	48404-48390	1	0.0396			
49786-49786	1	0.0396	48404-56310	1	0.0396			

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J., 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res,* Vol. 25, pp. 3389-3402.

Arga, K. Y., Onsan, Z. I., Kirdar, B., Ulgen, K. O., & Nielsen, J. ,2007, "Understanding signaling in yeast: Insights from network analysis", *Biotechnology and Bioengineering,* Vol. 97, pp. 1246-1258.

Bader, G. D., and Hogue, C. W., 2003, "An automated method for finding molecular complexes in large protein interaction networks", *BMC Bioinformatics,* 4:2.

Barabasi, A. L., & Albert, R., 1999, "Emergence of scaling in random networks", *Science,* Vol. 286, pp. 509-512.

Barabasi, A. L., & Oltvai, Z. N., 2004, "Network biology: understanding the cell's functional organization", *Nature Reviews Genetics,* Vol. 5, pp. 101-113.

Bejsovec, A., 2005, "Wnt pathway activation: new relations and locations", *Cell,* Vol. 120, pp. 11-14.

Berezin, C., Glaser, F., Rosenberg, J., Paz, I., Pupko, T., Fariselli, P., Casadio, R., Ben-Tal, N., 2004, "ConSeq: the identification of functionally and structurally important residues in protein sequences", *Bioinformatics,* Vol. 20, pp. 1322-1324.

Bornberg-Bauer, E., Beaussart, F., Kummerfeld, S. K., Teichmann, S. A., & Weiner, J., 3rd., 2005, "The evolution of domain arrangements in proteins and interaction networks", *Cellular and Molecular Life Sciences,* Vol. 62, pp. 435-445.

Bowerman, B., 2005, "Advocating asymmetry and the POP-1 paradox: noncanonical Wnt signaling in C. Elegans", *Cell*, Vol. 121, pp. 662-664.

Braungart, E., Gerlach, M., Riederer, P., Baumeister, R., & Hoener, M. C., 2004, "Caenorhabditis elegans MPP+ model of Parkinson's disease for high-throughput drug screenings", *Neurodegenerative Diseases,* Vol. 1, pp.175-183.

Cadigan, K. M., and Nusse, R., 1997, "Wnt signaling: a common theme in animal development", *Genes and Development*, Vol. 11, pp. 3286-3305.

Chatr-aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L., et al., 2007, "MINT: the Molecular INTeraction database", *Nucleic Acids Res,* Vol. 35, pp. D572-574.

Crabtree, G. R., and Olson, E. N., 2002, "NFAT signaling: choreographing the social lives of cells", *Cell*, Vol. 109, pp. S67-S79.

Dirnberger, D., Messerschmid, M., & Baumeister, R., 2008, "An optimized split-ubiquitin cDNA-library screening system to identify novel interactors of the human Frizzled 1 receptor", *Nucleic Acids Res,* Vol. 36, pp. e37.

Durmus Tekir, S., Yalcin Arga, K., & Ulgen, K. O., 2009, "Drug targets for tumorigenesis: insights from structural analysis of EGFR signaling network", *J Biomed Inform,* Vol. 42, pp. 228-236.

Eisenmann, D. M., 2005, "Wnt signaling", *WormBook,* pp. 1-17.

Fujii, N., You, L., Xu, Z., Uematsu, K., Shan, J., He, B., Mikami, I., Edmondson, L. R., Neale, G., Zheng, J., Guy, R. K., Jablons, D. V., 2007, "An antagonist of dishevelled protein-protein interaction suppresses beta-catenin-dependent tumor cell growth", *Cancer Res,* Vol. 67, pp. 573-579.

Giles, R. H., van Es, J. H., and Clevers, H., 2003, "Caught up in a Wnt storm: Wnt signaling in cancer", *Biochim Biophys Acta*, Vol. 1653, pp. 1-24.

Girvan, M., and Newman, M. E., 2002, "Community structure in social and biological networks", *Proc Natl Acad Sci U S A,* Vol. 99, pp. 7821-7826.

Gleason, J. E., Korswagen, H. C., and Eisenmann, D. M., 2002, "Activation of Wnt signaling bypasses the requirement for RTK/Ras signaling during C. elegans vulval induction", *Genes Dev*, Vol. 16, pp. 1281-1290.

Gough, J., Karplus, K., Hughey, R., & Chothia, C., 2001, "Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure", *J Mol Biol,* Vol. 313, pp. 903-919.

Han, M., 1997, "Gut reaction to Wnt signaling in worms", *Cell*, Vol. 90, pp. 581-584.

Hardin, J., and King, R. S., 2008, "The long and the short of Wnt signaling in C. elegans", *Curr Opin Genet Dev,* Vol. 18, pp. 362-367.

Herman, M. A., 2002, "Control of cell polarity by noncanonical Wnt signaling in C. Elegans", *Semin Cell Dev Biol*, Vol. 13, pp. 233-241.

Herman, M. A., and Wu, M., 2004, "Noncanonical Wnt signaling pathways in C. elegans converge on POP-1/TCF and control cell polarity", *Front Biosci*, Vol. 9, pp. 1530-1539.

Hertweck, M., and Baumeister, R., 2005, "Automated assays to study longevity in C. elegans", *Mech Ageing Dev,* Vol. 126, pp. 139-145.

Hingwing, K., Lee, S., Nykilchuk, L., Walston, T., Hardin, J., Hawkins, N., 2009, "CWN-1 functions with DSH-2 to regulate C. elegans asymmetric neuroblast division in a β-catenin independent Wnt pathway", *Dev Biol,* Vol. 328, pp. 246-256.

Huelsken, J., and Birchmeier, W., 2001, "New aspects of Wnt signaling pathways in higher vertebrates", *Curr Opin Genet Dev,* Vol. 11, pp. 547-53.

Ishitani, T., Kishida, S., Hyodo-Miura, J., Ueno, N., Yasuda, J., Waterman, M., Shibuya, H., Moon, R. T., Ninomiya-Tsuji, J., and Matsumoto, K., 2003, "The TAK1-NLK mitogen-activated protein kinase cascade functions in the Wnt-5a/Ca($^{2+}$) pathway to antagonize Wnt/beta-catenin signaling", *Mol Cell Biol,* Vol. 23, pp. 131-139.

Ishitani, T., Ninomiya-Tsuji, J., Nagai, S., Nishita, M., Meneghini, M., Barker, N., Waterman, M., Bowerman, B., Clevers, H., Shibuya, H., and Matsumoto, K., 1999, "The TAK1-NLK-MAPK-related pathway antagonizes signalling between beta-catenin and transcription factor TCF", *Nature,* Vol. 399, pp. 798-802.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabasi, A. L., 2000, "The large-scale organization of metabolic networks", *Nature,* Vol. 407, pp. 651-654.

Kohn, A. D., and Moon, R. T., 2005, "Wnt and calcium signaling: beta-catenin-independent pathways", *Cell Calcium,* Vol. 38, pp. 439-446.

Korswagen, H. C., 2002, "Canonical and non-canonical Wnt signaling pathways in Caenorhabditis elegans: variations on a common signaling theme", *Bioessays,* Vol. 24, pp. 801-810.

Korswagen, H. C., 2007, "Wnt signaling in C. elegans: New insights into the regulation of POP-1/TCF-mediated activation and repression", *Advances in Dev Biol,* Vol. 17, pp. 95-110.

Korswagen, H. C., Herman, M. A., and Clevers, H. C., 2000, Distinct beta-catenins mediate adhesion and signalling functions in C. elegans, *Nature,* Vol. 406, pp. 527-532.

Kuhl, M., Geis, K., Sheldahl, L. C., Pukrop, T., Moon, R. T., and Wedlich, D., 2001, "Antagonistic regulation of convergent extension movements in Xenopus by Wnt/beta-catenin and Wnt/Ca^{2+} signaling", *Mech Dev,* Vol. 106, pp. 61-76.

Kuhl, M., Sheldahl, L. C., Malbon, C. C., and Moon, R. T., 2000, "Ca($^{2+}$)/calmodulin-dependent protein kinase II is stimulated by Wnt and Frizzled homologs and promotes ventral cell fates in Xenopus", *J Biol Chem,* Vol. 275, pp. 12701-12711.

Kuhl, M., Sheldahl, L. C., Park, M., Miller, J. R., and Moon, R. T., 2000, "The Wnt/Ca^{2+} pathway: a new vertebrate Wnt signaling pathway takes shape", *Trends Genet,* Vol. 16, pp. 279-283.

Lee, P. N., Pang, K., Matus, D. Q., and Martindale, M. Q., 2006, "A WNT of things to come: evolution of Wnt signaling and polarity in cnidarians", *Semin Cell Dev Biol,* Vol. 17, pp. 157-167.

Maere, S., Heymans, K., Kuiper, M., 2005, "BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks", *Bioinformatics,* Vol. 21, pp. 3448-3449.

Maiese, K., Li, F., Chong, Z. Z., and Shang, Y. C., 2008, "The Wnt signaling pathway: aging gracefully as a protectionist?", *Pharmacol Ther,* Vol. 118, pp. 58-81.

Meissner, B., Boll, M., Daniel, H., Baumeister, R., 2004, "Deletion of the intestinal peptide transporter affects insulin and TOR signaling in Caenorhabditis elegans", *J Biol Chem,* Vol. 279, pp. 36739-36745.

Mikels, A. J., and Nusse, R., 2006, "Purified Wnt5a protein activates or inhibits beta-catenin-TCF signaling depending on receptor context", *PLoS Biol,* Vol. 4, pp. e115.

Miller, J. R., Hocking, A. M., Brown, J. D., and Moon, R. T., 1999, "Mechanism and function of signal transduction by the Wnt/beta-catenin and Wnt/Ca^{2+} pathways", *Oncogene,* Vol. 18, pp. 7860-7872.

Moon, R. T., Kohn, A. D., De Ferrari, G. V., Kaykas, A. 2004, "WNT and beta-catenin signalling: diseases and therapies", Nat Rev Genet, Vol. 5, pp. 691-701.

Natarajan, L., Witwer, N. E., and Eisenmann, D. M., 2001, "The divergent Caenorhabditis elegans beta-catenin proteins BAR-1, WRM-1 and HMP-2 make distinct protein interactions but retain functional redundancy in vivo", *Genetics,* Vol. 159, pp. 159-172.

Nelson, W. J., and Nusse, R., 2004, "Convergence of Wnt, beta-catenin, and cadherin pathways", *Science,* Vol. 303, pp. 1483-1487.

Nusse, R., 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, Vol. 31, pp. 99-109.

Pan, W. J., Pang, S. Z., Huang, T., Guo, H. Y., Wu, D., Li, L., 2004, "Characterization of function of three domains in dishevelled-1: DEP domain is responsible for membrane translocation of dishevelled-1", *Cell Res,* Vol. 14, pp. 324-330.

Pandur, P., 2005, "Recent discoveries in vertebrate non-canonical Wnt signaling: Towards a Wnt signaling network", *Advances in Dev Biol,* Vol. 14, pp. 91-106.

Peifer, M., and Polakis, P., 2000, "Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus", *Science,* Vol. 287, pp. 1606-1609.

Platzer, U., and Meinzer, H. P., 2004, "Genetic networks in the early development of Caenorhabditis elegans", *Int Rev Cytol,* Vol. 234, pp. 47-100.

Polakis, P., 2000, "Wnt signaling and cancer", *Genes Dev,* Vol. 14, pp. 1837-1851.

Przulj, N., Wigle, D. A., Jurisica, I., 2004, "Functional topology in a network of protein interactions", *Bioinformatics,* Vol. 20, pp. 340-348.

Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., Eisenberg, D., 2004, "The Database of Interacting Proteins: 2004 update", *Nucleic Acids Res,* Vol. 32, pp. D449-451.

Saneyoshi, T., Kume, S., Amasaki, Y., and Mikoshiba, K., 2002, "The Wnt/calcium pathway activates NF-AT and promotes ventral cell fate in Xenopus embryos", *Nature,* Vol. 417, pp. 295-299.

Santonico, E., Castagnoli, L., Cesareni, G., 2005, "Methods to reveal domain networks", *Drug Discovery Today,* Vol. 10, pp. 1111-1117.

Schambony, A., and Wedlich, D., 2007, "Wnt-5A/Ror2 regulate expression of XPAPC through an alternative noncanonical signaling pathway", *Dev Cell,* Vol. 12, pp. 779-792.

Schulz, R. A., and Yutzey, K. E., 2004, "Calcineurin signaling and NFAT activation in cardiovascular and skeletal muscle development", *Dev Biol,* Vol. 266, pp. 1-16.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504.

Sheldahl, L. C., Park, M., Malbon, C. C., and Moon, R. T., 1999, "Protein kinase C is differentially stimulated by Wnt and Frizzled homologs in a G-protein-dependent manner", *Curr Biol,* Vol. 9, pp. 695-698.

Sheldahl, L. C., Slusarski, D. C., Pandur, P., Miller, J. R., Kuhl, M., Moon, R. T., 2003, "Dishevelled activates Ca^{2+} flux, PKC, and CamKII in vertebrate embryos", *J Cell Biol,* Vol. 161, pp. 769-777.

Slusarski, D. C., and Pelegri, F., 2007, "Calcium signaling in vertebrate embryonic patterning and morphogenesis", *Dev Biol,* Vol. 307, pp. 1-13.

Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A., Tyers, M., 2006, "BioGRID: a general repository for interaction datasets", *Nucleic Acids Res*, Vol. 34, pp. D535-539.

The *C. elegans* Sequencing Consortium, 1998, "Genome sequence of the nematode C. elegans: a platform for investigating biology", *Science,* Vol. 282, pp. 2012-2018.

The Gene Ontology Consortium, 2000, "Gene ontology: tool for the unification of biology", *Nature Genetics,* Vol. 25, pp. 25-29.

The UniProt Consortium, 2008, "The universal protein resource (UniProt)", *Nucleic Acids Res,* Vol. 36, pp. D190-195.

Thompson, J. D., Higgins, D. G., Gibson, T. J., 1994, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice", *Nucleic Acids Res,* Vol. 22, pp. 4673-4680.

Veeman, M. T., Axelrod, J. D., and Moon, R. T., 2003, "A second canon: Functions and mechanisms of beta-catenin-independent Wnt signaling", *Dev Cell,* Vol. 5, pp. 367-377.

Veretnik, S., Bourne, P. E., Alexandrov, N. N., Shindyalov, I. N., 2004, "Toward consistent assignment of structural domains in proteins", *J Mol Biol,* Vol. 339, pp. 647-678.

Wagner, A., Fell, D. A., 2001, "The small world inside large metabolic networks", *Proc Biol Sci,* Vol. 268, pp. 1803-1810.

Wallingford, J. B., Habas, R., 2005, "The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity", *Development,* Vol. 132, pp. 4421-4436.

Wang, H. Y., Malbon, C. C., 2003, "Wnt signaling, Ca^{2+}, and cyclic GMP: visualizing Frizzled functions", *Science,* Vol. 300, pp. 1529-1530.

Weeraratna, A. T., Jiang, Y., Hostetter, G., Rosenblatt, K., Duray, P., Bittner, M., and Trent, J. M., 2002, "Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma", *Cancer Cell,* Vol. 1, pp. 279-288.

Wharton, K. A., Jr., 2003, "Runnin' with the Dvl: proteins that associate with Dsh/Dvl and their significance to Wnt signal transduction", *Dev Biol,* Vol. 253, pp. 1-17.

Wodarz, A., and Nusse, R., 1998, "Mechanisms of Wnt signaling in development", *Annu Rev Cell Dev Biol,* Vol. 14, pp. 59-88.

Wong, H. C., Bourdelas, A., Krauss, A., Lee, H. J., Shao, Y., Wu, D., Mlodzik, M., Shi, D., Zheng, J., 2003, "Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled", *Mol Cell,* Vol. 12, pp. 1251-1260.

Wu, M., and Herman, M. A., 2006, "A novel noncanonical Wnt pathway is involved in the regulation of the asymmetric B cell division in C. Elegans", *Dev Biol,* Vol. 293, pp. 316-29.

Yanagawa, S., van Leeuwen, F., Wodarz, A., Klingensmith, J., and Nusse, R., 1995, "The dishevelled protein is modified by wingless signaling in Drosophila", *Genes Dev,* Vol. 9, pp. 1087-97.

Yasuda, J., Tsuchiya, A., Yamada, T., Sakamoto, M., Sekiya, T., and Hirohashi, S., 2003, "Nemo-like kinase induces apoptosis in DLD-1 human colon cancer cells", *Biochem Biophys Res Commun,* Vol. 308, pp. 227-33.