INVESTIGATION OF LINEAR PATHS IN HUMAN WNT SIGNALING NETWORK

by

Bahar Nalbantoğlu

B.S., Chemical Engineering, Boğaziçi University, 2009

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

2011

dedicated to my beloved parents

ACKNOWLEDGEMENT

I would like to offer my sincere and grateful thanks to my thesis advisor Prof. Kutlu Özergin Ülgen for her support during my graduate studies and counsel throughout this study.

I also would like to thank my thesis jury; Assist. Prof. Elif Özkırımlı Ölmez and Prof. Belma Kın Özbek for their critical comments and feedbacks on my thesis.

I also thank my laboratory friends; Ceyda, Elif, Esra and Saliha for their suggestion on this study and answering my never ending questions.

I offer my thanks to my friends Merve, Ünal, Çağlar, Duygu, Şefik, Tarık, Murat, Nilay, Vasfiye, Ali, Burcu, Selcen, İpek, Aybüke, Caner, Gülsüm, İrfan, Aslıgül, Ömür, Simay, Ahmet, Seval, Dilek, Gözde, Okan for their love and support.

I would also like to thank my mother, Hazan Nalbantoğlu, my father, Şahabettin Nalbantoğlu, my sister Özlem Nalbantoğlu Elmas and my brother, Ufuk Nalbantoğlu for their never ending love, support and encouragement through all my life.

ABSTRACT

INVESTIGATION OF LINEAR PATHS IN HUMAN WNT SIGNALING NETWORK

Recently, the evolutionarily conserved signaling pathways which are involved in embryonic development are on the march for many researchs since the deregulations seen in the mechanism of these pathways results in several diseases, especially in cancer. Hence, interaction networks have begun to be appreciated because it may be useful to understand the general principles of biological systems by means of systems biology. This study is concerned with the Wnt signaling pathway that are known as Wnt/β-catenin, planar cell polarity (PCP) and Wnt-calcium (Wnt/Ca^{2+}) pathways and presents potential targeting points for cancer drug progression. Here, the reconstruction of Wnt signaling network was performed for *Homo Sapiens* via integration of interactome data and Gene Ontology annotations since the whole Wnt mechanism is not well understood in human due to limited experimental findings. In addition to the Wnt sub-networks, the whole Wnt signaling network was also investigated to see the global picture. All the reconstructed networks showed scale-free topologies. Moreover, the linear paths in which the signal is transferred from ligands (input) to transcription factors (output) were identified. The graph theoretical analysis was performed for analyzing the toplogical properties of the network proteins. Then, the crosstalk analysis was applied in order to detect the significant bridging proteins in all sub-networks. Finally, the size reduction of the network was done; however, it caused a significant loss in core proteins, and did not result in a small network for convenient usage. The proteins Beta-catenin, LEF1, GSK3b, APC, AXIN and DKK1, which are reported to be potential drug targets in literature, are found as the essential proteins of the Wnt network. Besides, LEF1, WNT7A and FZD9 proteins which are the participants with the highest percentages in the linear path analysis and CNKD3, TGFR2 and MYC which are the specific proteins in linear paths are also proposed as potential drug targets.

ÖZET

İNSANDAKİ WNT SİNYAL AĞININ DOĞRUSAL YOLİZLERİNİN İNCELENMESİ

Son yıllarda, evrimsel açıdan korunmuş ve embriyonik gelişimde yer alan sinyal ağları, düzenlemelerinde (regülasyon) görülen değişiklerin başta kanser olmak üzere çeşitli hastalıklara sebep olması dolayısıyla birçok araştırmacının gözdesi olmuştur. Sistem biyolojisi aracılığıyla biyolojik sistemlerin genel ilkelerini anlamak mümkün olabildiğinden, protein ve etki alanı etkileşim ağları önem kazanmıştır. Bu çalışma Wnt/beta-katenin, Wnt/düzlemsel hücre kutuplaşma ve Wnt/kalsiyum izyolları olarak bilinen Wnt sinyal ağının hesapsal yöntemlerle oluşturulması ve analizini içermekte, ve kanser ilaçlarının gelişimi için muhtemel ilaç hedeflerini sunmaktadır. İnsan ele alındığında yapılabilecek araştırmaların yasal boyutları dolayısı ile kısıtlı deneysel kapasite ile çalışılabilindiğinden Wnt sinyal ağı bütünüyle anlaşılamamıştır. Burada, protein etkileşim datası ve Gen Ontolojisi ifadelerinin bütünleştirilmesi yoluyla insandaki Wnt sinyal ağı yeniden kurulmuştur. Wnt sinyal sistemini oluşturan izyolları (Wnt/beta-katenin, Wnt/düzlemsel hücre kutuplaşma ve Wnt/kalsiyum) hem ayrı ayrı hemde genel resmi görmek amacıyla Wnt ağı bir bütün olarak incelenmiştir. Oluşturulan tüm ağyapıları küçük dünya özellikleri göstermiştir. Buna ek olarak, sinyalin ligandlardan (girdi) transkripsiyon faktörlerine (çıktı) aktarıldığı doğrusal ileti yolları incelenmiştir. Ayrıca, ağ proteinlerinin topolojik özelliklerinin incelenmesi için grafik kuramsal analizleri uygulanmıştır. Sonrasında ortak olan ve köprü niteliği taşıyan proteinleri bulmak amacıyla etkileşim incelemesi uygulanmıştır. Son olarak ağ küçültme işlemi yapılmış; fakat bu uygulama esas proteinlerin önemli miktarda azalmasına sebep olmuş ve küçük bir ağ elde edilememiştir. Literatürde ilaç hedefi olarak rapor edilen Beta-katenin, LEF1, GSK3b, APC, AXIN ve DKK1 proteinleri doğrusal ileti yollarının önemli proteinleri olarak tespit edilmişlerdir. Ayrıca doğrusal yol analizinde en çok yer alan proteinler olan LEF1, WNT7A ve FZD9 proteinleri ile doğrusal yollarda özgül olarak bulunan CNKD3, TGFR2 ve MYC proteinleri de bu çalışma çerçevesinde muhtemel ilaç hedefleri olarak önerilmektedirler.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Signaling is a part of system communication among living cells by processing biological information that governs basic cellular activities and coordinates cell actions. The signals are transmitted by means of signaling molecules to the surface of the target cell. Then ligands bind to receptor that activates the signaling mechanism and signaling proteins transfer the information and this process ends up with altering of gene transcription in the nucleus resultsing in many cellular processes such as differentiation, proliferation etc. The deviations, abnormalities and errors in signal transduction are responsible for diseases such as cancer, autoimmunity, and diabetes. Recently, interaction networks have begun to be appreciated because it may be useful to understand the general principles of biological systems by means of systems biology.

Protein-protein interaction can be defined as the delivery of the biological information when the proteins are in contact. Identification of protein-protein interactions is an useful starting point for the representation of physiological cell signaling pathways. Regulated interactions between proteins govern signaling pathways within and between cells. Cells that are defined as the functional basic unit are ready to respond to essential signals in their environment. Cell signaling is part of a complex system of communication that governs basic cellular activities and coordinates cell actions. The deviations, abnormalities and errors in signal transduction are responsible for diseases such as cancer, autoimmunity, and diabetes. Systems biology research helps us understand the underlying structure of cell signaling networks and how changes in these networks may affect the transmission and flow of information. By the help of systems biology, depending on the information obtained by cell signaling networks, diseases may be treated effectively. Due to reasoning that disease states often result from genetic alterations in physiologically important biological signaling pathways, disease-derived protein variants have frequently been used to provide a biomedical reference for the study of protein-protein interactions (Golemis and Adams, 2005).

Reconstruction of protein-protein interactions is based on the recent developments of new high throughput technologies. There are several methods available to determine PPIs

such as Yeast Two Hybrid, Tandem Affinity Purification and computational methods like Phylogenetic profile, Correlated Domain Signature Method and several integrative methods. Most of these approaches cover only a sub-set of possible interactions (Nandy *et al*., 2010). In order to reconstruct causal protein networks and predict complex system behaviors, a combination of multiple types of data, including genotypic, expression, transcription factor binding site (TFBS), and protein–protein interaction (PPI) data previously generated from a number of experiments is generally used. Some previous examples of reconstructed networks performed through thesis work in our department are as follows: insuling signaling in human (Ümit *et al*., 2009), hedgehog signaling in fly (Eren A, 2010) and Ca signaling in multiple species (Tiveci S, 2011).

Within the framework of the present thesis study, one of the major signaling systems, Wnt signaling, which is known to have role in regulating the direct processes such as embryonic development and growth, cell polarity and morphology, is investigated thoroughly. The deregulations and mutations in this signaling pathway cause several human diseases including lung, breast, colon and colorectal cancers. Investigating the Wnt signaling pathway is therefore attractive for determination of the strategies to identify the suitable drug targets for therapeutic intervention in cancer treatment (Cadigan and Liu, 2006; Cong *et al.*, 2004; Widelitz, 2005; Nusse, 2005). Many researchers studied and analyzed this signaling pathway via experimental and computational works in order to enucleate the governing mechanism through essential components (DasGupta, 2005; Kuhl *et al.*, 2000; Kohn and Moon, 2005). Comprising the small pieces is more effective in order to understand the mechanism and whereas generating the whole picture is more informative for drug target studies. Hence, identification of protein-protein interactions (PPIs) is necessary for the representation of physiological cell signaling pathways. In this study, reconstruction of Wnt signaling network was performed for *Homo Sapiens*. However, the whole Wnt mechanism is not well understood in human due to restricted experimental capacity/allowances and more attention should be devoted to this organism.

Following the brief information and aims of this study, the second chapter gives information about the theoretical background, introduces canonical pathways with several examples and the insights of the Wnt signaling pathway and its impact on diseases. This chapter also includes information about the pathway reconstruction methods and servers.

The third chapter concentrates on the methods of reconstruction and computational analysis of the Wnt signaling network. Domain–domain interaction network was also studied in order to fill the gaps and hence to improve the protein interaction network. The canonical and non-canonical Wnt sub-networks and the whole Wnt network were decomposed into linear paths from inputs (i.e. ligands) to outputs (i.e. transcription factors) to provide a better understanding of the signaling mechanism. The fourth chapter gives the results of the reconstruction and analyses of these networks. Finally, in the fifth chapter, the conclusions which give the summary of the main points and the recommendations are presented for future investigations.

2. BACKGROUND ASPECTS

2.1. Signaling Networks

Signaling proteins transfer the information that ends up with altering of gene transcription in the nucleus which results in many cellular processes. The improvement of high-throughput technologies enables the reconstruction of signaling networks and nowadays, interaction networks have begun to be appreciated because it may be useful to understand the general principles of biological systems by means of systems biology. Investigation of signaling networks are crucial for studying diseases especially cancer (Finkel and Gutkind, 2003). The Notch, Wnt, Ras-MAPK and JAK-STAT signaling pathways, which are mentioned here, are the examples of the most familiar signaling networks that are known to be related to cancer. Table 2.1 gives information about the major canonical signaling pathways.

Table 2.1. Some examples of major canonical signaling pathways.

Canonical pathway	Function	Reference
Cytokinin Signaling	multistep phosphorelay system	Müller and Sheen, 2007
Notch signaling	binary cell-fate decisions	Ehebauer *et al.*, 2006
Wnt signaling	diverse processes during development	Moon, 2005
JAK-STAT signaling	hematopoiesis, sex determination, migration processes	Johansen *et al.*, 2002
Ras-MAPK signaling	cell survival promotion and apoptosis inhibition	Bonni *et al.*, 1999
FasL-Fas signaling	diverse processes such as activation-induced cell death (AICD) in T cells, cytotoxic T cell (CTL) function, immune privilege, tumor surveillance, proliferation, differentiation,	Wajant, 2003
Hedgehog signaling	development and tissue homeostasis	Jacob and Lum, 2007

2.1.1. Notch Signaling

The Notch signaling pathway has an important role in cell-cell communication including gene regulation mechanisms that control multiple cell differentiation processes during embryonic and adult life. In addition to that, Notch signaling also participates in the many processes such as neuronal function and development, regulation of crucial cell

communication events between endocardium and myocardium during both the formation of the valve primordial and ventricular development and differentiation, expansion of the hematopoietic stem cell compartment during bone development. Notch signaling also takes part in commitment to the osteoblastic lineage, bone regeneration and osteoporosis, regulation of cell-fate decision in mammary glands at several distinct development stages and control of the actin cytoskeleton through the tyrosine kinase Abl. Notch signaling is modulated by fringe following the binding of the DSL ligand, and serial cleavage events provide Notch receptors activation within the transmembrane domain mediated by γ-secretase (presenilin) activity that results in translocation of intracellular domain of Notch (NICD) into the nucleus where it can regulate gene expression by activating the transcription factor CSL, C promoter binding factor (Liu *et al., 2005).* Deregulations in Notch signaling result in many disease states which involve cancer, T-cell acute lymphoblastic leukema (T-ALL) and Multiple Sclerosis (MS). Recent studies showed that Notch signaling crosstalks with Hedgehog and Wnt signaling (Gaiano and Fishell, 2002; Bolós , Grego-Bessa and de la Pompa, 2007; Aguirre , Rubio and Gallo, 2010; Hitoshi *et al.*,2002; Grego-Bessa *et al.*, 2007; Nobta *et al.*, 2005; Dontu *et al.*, 2004).

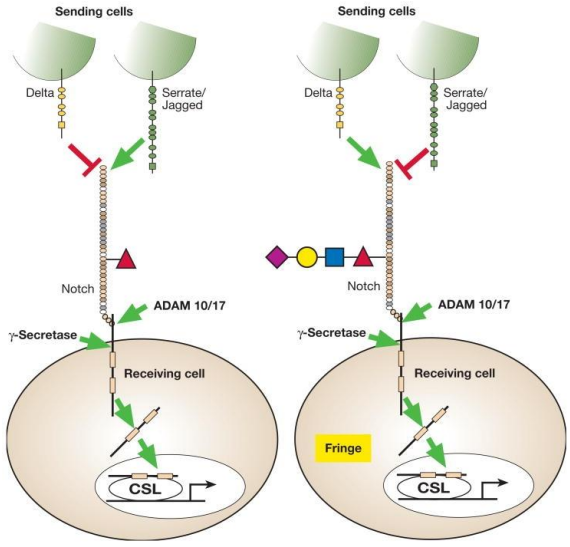

Figure 2.1. An overview of Notch signaling pathway (Varki *et al.*, 2009).

2.1.2. JAK-STAT Signaling Pathway

JAK-STAT (Janus kinase- signal transducer and activator of transcription) signaling pathway is one of the major signaling pathways in which three main components are

involved: a receptor, JAK and STAT. There are four JAK proteins in mammals which are JAK1, JAK 2, JAK 3 and TYK2. In this pathway, the information received from chemical signals of extracellular polypeptides is transmitted into the gene promoters in the nucleus through transmembrane receptors. A signal activates the receptor, then this results in the activation of JAK which autophosphorylates itself. After that stimulation, STAT protein binds to phosphorylated receptor through specific binding between STAT SH2 domain and receptor phosphotyrosine residues. There are seven STAT genes (STAT1, STAT2, STAT 3, STAT 4, STAT5A, STAT5B and STAT6) in mammals and their gene products are found in the cytoplasm of unstimulated target cells. After binding the stimulated receptor, STAT dimers are transported to nucleus from cytoplasm, and in nucleus they bind to DNA and trancription starts. This transcription occurred in the nucleus affects basic cell functions like cell growth, differentiation and death. Mutations and deregulations seen in JAK-STAT functionality can result in immune deficiency syndromes and cancers. Figure 2.2 illustrates the three specific JAK-STAT pathways in which STAT1 and STAT2 are used in type I interferon (IFNγ/β), STAT1 is used in type II interferon IFNγ and STAT3 is used in the third one (Aaronson and Horvath, 2002).

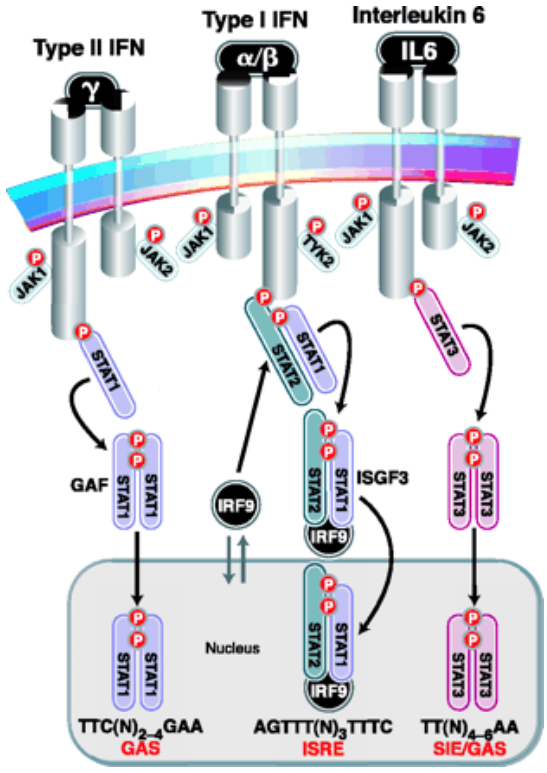

Figure 2.2. Three examples of signaling in JAK-STAT pathway (Aaronson and Horvath, 2002).

2.1.3. Ras-MAPK Signal Transduction Pathway

Ras-MAPK pathway is also a major pathway which starts from receptors and ends with protein expression in nucleus. Since this pathway regulates many cellular processes such as blood glucose concentration, gene expression etc., it is necessary for growth and differentiation. When a mutation is seen in the proteins of this pathway, it gets "on/off" position which is important for the following steps in cancer treatment. In this pathway, adaptor protein binds to phosphorylated receptors and guanine nucleotide exchange factors (GEFs) join into the complex. The stimulation of GDP to GTP exchange by GEFs on RAS results in interaction of RAS with RAF reducing its inhibition by 14-3-3 proteins. RAF activates MEK, which then activates final MAPK in the cascade, and stream substrates, which are transcription factors, are then phosphorylated by final MAPK (Thatcher, 2010). Figure 2.3 illustrates the Ras-MAPK pathway.

For the controlling of malignant diseases, the drugs that are selectively down regulating MAPK cascades are appreciable therapeutic agents. It was also shown that the mutations in components of the RAS-MAPK pathway cause clinically distinct human "neuro-cardio-facial-cutaneous" (NCFC) syndromes (Alj *et al.*, 2006).

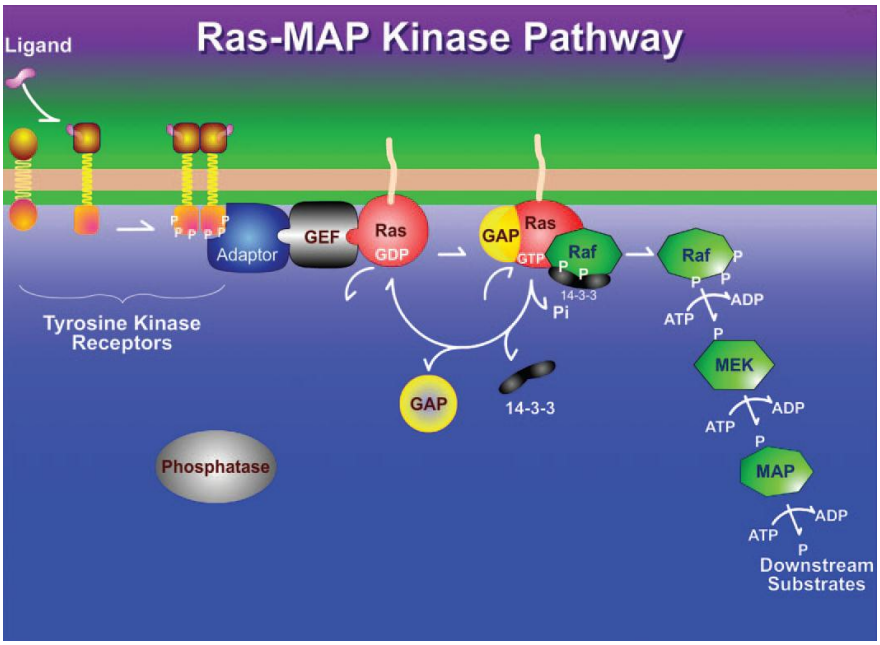

Figure 2.3.An overview of Ras-MAPK signaling pathway (Thatcher, 2010).

2.2 Wnt Signaling Pathway

The Wnt signaling network regulates diverse processes during development such as cell fate determination, structural remodeling, cell polarity and morphology, cell adhesion and growth. Some of the Wnt signaling components are obviously conserved in a variety of organisms, from *Caenorhabditis elegans* to *Homo sapiens* although the strength of any signaling effect changes depending on the types of wnt ligand, cell and organism. There are three major types of Wnt signaling which can be classified as Wnt/β-catenin, the planar cell polarity (PCP) and the Wnt-calcium (Wnt/Ca^{2+}) pathways (Figure 2.4). The Wnt/β-catenin pathway is known as the canonical pathway whereas the others are called as non-canonical pathways. Defects in the components of the Wnt signaling may cause tumor formation in different cell types (Katoh, 2005; Nusse, 2005; DasGupta, 2005).

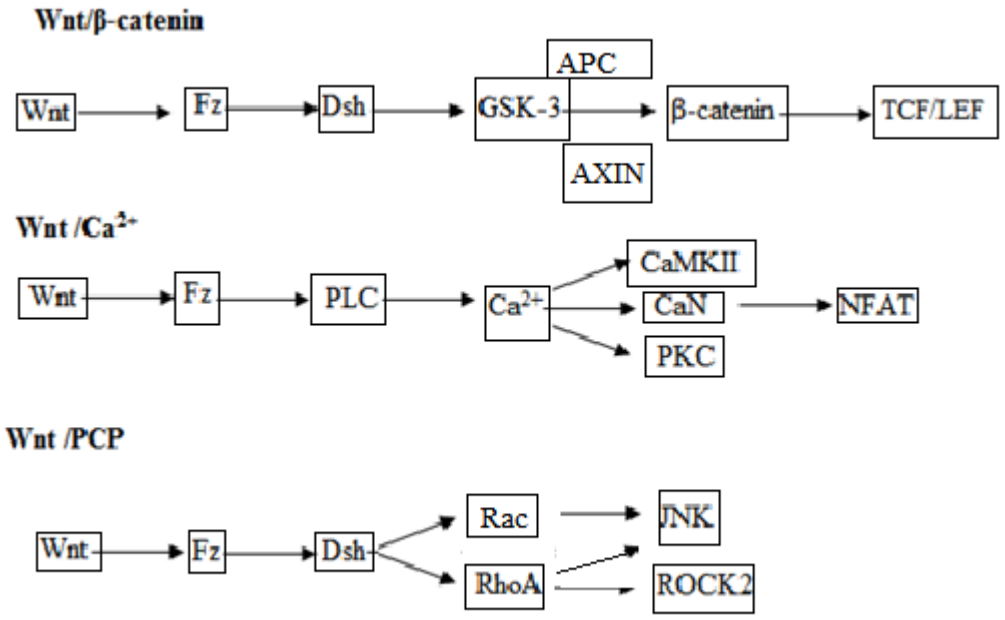

Figure 2.4. An overview of canonical and non-canonical Wnt signaling pathways in human (KEGG).

Wnt signaling is regulated by the presence or absence of the intracellular protein β -catenin. A large multiprotein machine that includes proteins of the APC and Axin families normally facilitate the addition of phosphate groups to β-catenin by glycogen synthase kinase-3b (GSK3b). Phosphorylated β -catenin binds to a protein called bTrCP, and it is then modified by the covalent addition of a small protein called ubiquitin. Proteins tagged with ubiquitin are degraded by the proteosome, the cell's protein-recycling center. When Wnt

signal is absent, the signal transduction pathway is OFF because β -catenin is rapidly destroyed. When cells are exposed to Wnt, it binds to cell surface receptors of the Frizzled family. Receptor activation antagonizes the APC-Axin "destruction complex" by an unknown mechanism that requires Dishvelled protein. This blocks b-catenin phosphorylation and its subsequent ubiquitination. β -Catenin is thus diverted from the proteosome, and it accumulates and enters the nucleus, where it finds a partner, a DNA binding protein of the TCF/ LEF family. Together, they activate new gene expression programs (Kadigan, 2008).

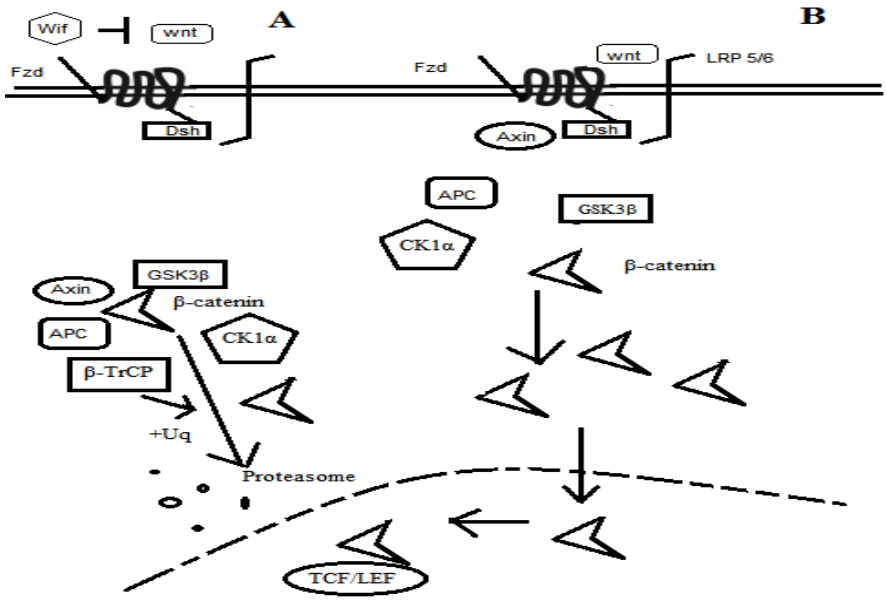

Figure 2.5. Regulation of β-catenin stability by Wnt signaling.

The illustration (Figure 2.5) depicts a cell in the absence (A) or presence (B) of Wnt protein. In unstimulated cells, β-catenin not complexed with the cadherin adhesion complex is phosphorylated by CKI and GSK3b, part of the destruction complex, leading to β-TrCP dependent ubiquitination and proteosomal degradation. The presence of Wnt, on the hand, promotes LRP and Fz association, leading to recruitment of Dvl to the complex and GSK3 and CKI phosphorylation of LRP. This stabilizes recruitment of Axin to the receptor which in turn may disrupt the activity of the destruction complex, and hence allows accumulation of β-catenin and nuclear translocation (Kadigan, 2008).

Wnt signaling can also activate an alternative signaling pathway involved in planar cell polarity (PCP) that may lead to protein kinase C (PKC) and Jun kinase (JNK) activation, resulting in calcium release and cytoskeletal rearrangements. Calcium has been implicated as

an important messenger in wnt pathway and recent studies showed that higher frequencies of calcium transients were associated with faster rates of outgrowth (Kadigan, 2008). Wnt5a activates the Wnt/Ca2+ signaling. The binding of Wnts to Frizzled receptors leads to an activation of heterotrimeric G-proteins and subsequent activation of phospholipase C by the G-protein beta/gamma dimer. This enzyme cleaves phosphatidylinositol-4, 5-bisphosphate (PIP2) into inositol-1,4,5-trisphosphate (IP3) and diacylglycerol (DAG). IP3 is released from the membrane, and binds to the IP3 receptor which subsequently releases calcium ions from intracellular stores. This calcium release is locally restricted, and calcium ions can subsequently activate calcium-sensitve proteins like protein kinase C (PKC), calcium-calmodulin dependent kinase II (CamKII), and/or calcineurin (CaCN). IP3 is degraded to inositol by specific phosphatases and recycled to PIP2 (i.e., the PI cycle). The activation of the beta2- adrenergic/Rfz-2 chimeric receptor stimulates Wnt signaling and is sufficient to trigger intracellular calcium release within minutes (Kuhl *et al.*, 2000; Kohn and Moon, 2005; Kuhl, 2004; Slusarskia *et al.*, 1997).

Since the Wnt signaling pathway has a complex structure with important consequences that result in a variety of diseases, it is needed to study the canonical and non-canonical Wnt signaling and understand the mechanism for both healthy and defective situations.

2.2.1. Wnt signaling in Diseases

The malfunctions and defects in wnt signaling are generally caused by activity changes occurring in the cascade. Recent studies show that mutations or deregulations in Wnt signaling can cause different types of responses. Wnt pathway has been implicated in several diseases such as cancer, heart disease, Alzheimer disease, osteoarthritis and diseases of bone, eye and heart and that’s why there is a remarkable interest in this pathway for drug discovery researches (Cadigan, 2008; Klaus and Birchmeier, 2008;Paul and Dey, 2008). The Wnt signaling components and related diseases are given in Table 2.2 below.

In the strategies for drug targeting, the first step is to identify the essential regulatory proteins and their functions in the signaling cascade. In this section, present drug design strategies for this pathway are explained. Since there is no selective inhibitor for Wnt signaling, the key components are usually considered as potential drug targets. The strategy

follows the main steps in the pathway which are receptor-ligand interaction, β-catenin complex in the cytoplasm and β-catenin-TCF/LEF complex (Yanaga and Sasaguri, 2007; Luo *et al.*, 2007).

Table 2.2. Wnt signaling pathway components involved in diseases.

Gene	**Disease**	**Mutation or Activity change**	**References**
Beta catenin	Carcinogenesis, hepatocellular carcinomasWilms' tumors	Gain of function	Klaus and Birchmeier, 2008; Maiti *et al.*,2000
DVL	Myocardial infarction	Loss of function	Luo *et al.*, 2007
FZDs	Gastric cancer,colorectal cancer& carcinogenesis	Loss of function	Kirikoshi, Sekihara and Katoh, 2001;Ueno *et al.*, 2008
APC	Colorectal cancer, carcinogenesis	Loss of function	Klaus and Birchmeier, 2008; Ueno *et al.*, 2008
Axin2	Familial tooth agenesis; colorectal cancer	Loss of function	Luo *et al.*, 2007
Dkk1	Increased osteolytic metastasis of multiple myeloma	Overexpression	Tian *et al.*, 2003
sFRP(s)	colon cancer, mesothelioma, bladder cancer	Reduced	Tan and Kelsey, 2009; Paul and Dey, 2008; Gehrke, Gandhirajan and Kreuzer, 2009.
GSK-3β	colorectal cancer	Deregulation	Ge and Wang, 2010
LRP5	A high bone mass phenotype	Gain of function	Luo *et al.*, 2007
Wnt 1	Schizophrenia	Increased expression	Miyaoka *et al.*, 1999
Wnt 5B	type II diabetes, tumor suppressor	Overexpression	Luo *et al.*, 2007

The first step in wnt signaling is the interaction between wnt ligands and frizzled receptors. Although the information about ligand-receptor is limited, the frizzled proteins can be thought as drug targets due to their similar structure to the G protein-coupled receptors (GPCRs) which are preferred drug targets. In addition to that, the overexpression of natural

inhibitors such as WIF-1, sFRPs, DKK etc. has been considered as another logical approach (Luo *et al.*, 2007; Janssens, 2006).

The second strategy for inhibition of wnt signaling is targeting the multiprotein β-catenin destruction complex. Gsk-3b and APC are the validated drug targets in cancer treatment since the differentiation-inducing factors (DIFs) activates Gsk-3b and inhibits wnt signaling. Also, AXIN mutation is known to induce apoptosis in colorectal cancers. In addition to that, targeting β-catenin is a rational approach since abnormous wnt signaling results in increased β-catenin levels (Satoh *et al.*, 2000; Yanaga and Sasaguri, 2007; Luo *et al.*, 2007).

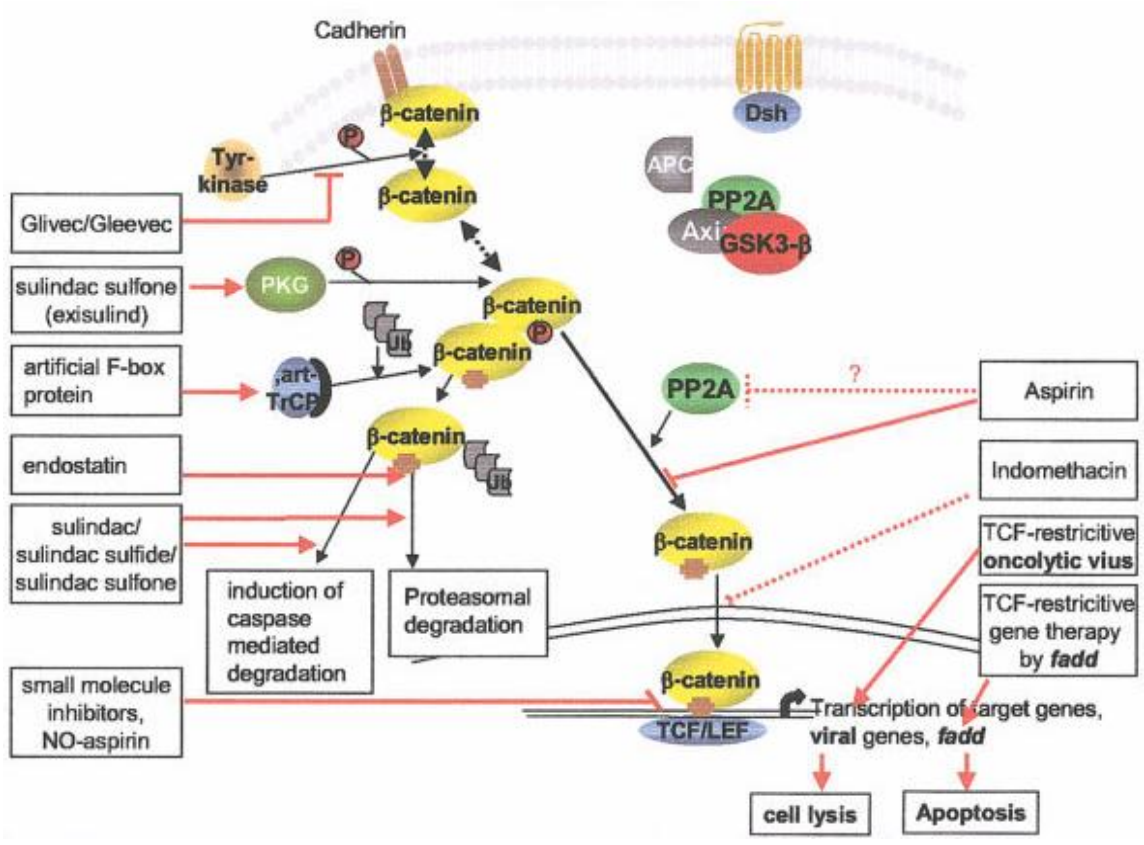

Figure 2.6. Schematic representation of inhibitors of wnt signaling used as anti-cancer drugs (Dihlmann *et al.*, 2005).

Furthermore, targeting another β-catenin complex, β-catenin-TCF/LEF, is an alternative approach used in inhibition of wnt signaling. Recent studies showed that some of the

important downstream mediators of β-catenin/Tcf signaling, such as c-Myc, cyclinD1, PPARd, COX-2, CD44 and MMP7, can also be targeted. However, it is important to design a selective inhibitor to impede the complex without disturbing other interactions of β-catenin (Luu *et al.*, 2004; Satoh *et al.*, 2000; Janssens, 2006).

Table 2.3. Inhibitors used as anti-cancer drugs in wnt signaling.

Agent	**Type**	**Mechanism**	**Developmental stage**	**References**
Indomethacin	NSAID	Repression of beta-catenin expression	Different types of in phase II/III	http://www.cancer.gov/clinicaltrials
Celecoxib	Selective COX-2 inhibitor	beta-catenin relocalization	FDA-approved for FAP Bladder cancer; Aktinic keratosis, breast cancer in phase III clinical trial HNPCC; SCLC; HCC in phase I/II clinical trial	http://www.cancer.gov/clinicaltrials
Sulindac, sulindac sulfide	NSAID, Sulindac metabolite	Induction of beta-catenin degradation by the proteasome	Different types of in phase II	http://www.cancer.gov/clinicaltrials Boon *et al.*, 2004,Rice *et al.*, 2003
Aspirin	NSAID	Stabilization of S/T-phosphorylated, inactive beta-catenin	CRC in Phase II clinical trial (119,120)	Baron *et al.*, 2003; Snadler *et al.*, 2003; Dihlmann *et al.*, 2003
Glivec (Gleevec/STI-571)	inhibitor for Tyrosine kinase	Inhibition of tyrosine phosphorylation; Relocalization of beta-catenin to the plasma membrane	Gastrointestinal stromal tumors (GIST) and CML: clinical Trials	Zhou *et al.* 2003, http://www.cancer.gov/clinicaltrials

The COX inhibitors, also known as anti-inflammatory drugs (NSAIDs), which are aspirin, indomethacin and sulindac, are known as the widely used drugs for diminishing the

wnt signaling. In addition to that, at present, there is only one drug, celecoxib, which is allowed by Food and Drug Administration (FDA) of the USA (Yanaga and Sasaguri, 2007). Table 2.2 gives the inhibitors used in wnt signaling as potential drugs, and Figure 2.6 illustrates schematic overview of these drugs.

2.3. Pathway reconstruction servers

Since the biological information is growing rapidly, automated computational methods are required for functional annotations and for deciphering biological roles of these genomes. There are several servers and methods that can be used in the reconstruction of biological pathways. In this part, some important servers that are recently introduced are explained (Table 2.4).

Table 2.4. The methods and availability of the common servers for network reconstruction.

Server	Availability	URL	Ref
KAAS	Web	http://www. Genome.jp/keg/kaas	Moriya *et al.*, 2007
FMM	Web	http://FMM.mbc.nctu.edu.tw/.	Chou *et al.*, 2009
RAST	Web	Rast.nmpdr.org	Aziz *et al.*, 2008
CARMEN	Web	http://carmen.cebitec.uni-bielefeld.de/cgi-bin/index.cgi	Schneider *et al.*, 2010
PathPred	Web	http://www.genome.jp/tools/pathpred/	Moriya *et al.*, 2010

2.3.1. KAAS: an automatic genome annotation and pathway reconstruction server

Due to increasing dynamic information there is a need to use automated servers in biological approaches. Functional annotation depends on the application of sequence similarity with well-annotated sequences. Recent works showed that for enzymes, 40 to 70% sequence identity is needed for acceptable function prediction in which 90% accuracy is obtained (Rost, 2002; Tian and Skolnick, 2003). GO annotations for different species are difficult to integrate due to the usage of different databases. KEGG GENES database supplies a source in which cross-species annotations of all available genomes are reached by a standard mechanism called KEGG ortholog (KO) system. KAAS uses this integrated

method with Smith-Waterman algorithm. First, Blast scores between the query set and reference set taken from KEGG GENES are computed and homologs are found. Then these homologs are filtered using a cut off value by bi-directional hit rate (BHR). Later ortholog candidates are classified by KO groups according to the annotations. Finally, each KO group is scored by probability and heuristics. The user gives the FASTA format of the query gene and obtains an annotation result which is highlighted in the map. This server is available at http://www. Genome.jp/keg/kaas (Moriya *et al.*, 2007).

2.3.2. RAST: Rapid Annotation using Subsystem Technology

RAST is a fully automated server, which is designed to annotate bacterial and archaeal genomes. This server is used to reconstruct the metabolic networks via uploading a genome as a set of contigs in Fasta format. The main idea of the system is based on the reference pathways which are called as the subsystems. The subsystems assign a collection of functionally connected genes and the system maintains a number of organism-specific variants. First the protein-encoding, rRNA and tRNA genes are identified and in the following step, function assignments are performed in order to predict which subsystems are represented in the genome. In the light of this information the metabolic network is reconstructed and the output, which involves relevant information and annotations, are ready to download.

The RAST server also supplies the user an annotated genome-browsing environment to compare it to different genomes within the SEED environment. The output of the annotated genome can be attained within 12-24 hours after submission. This server is available at http://rast.nmpdr.org/ and can be utilized after registration (Aziz *et al.*, 2007).

2.3.3. CARMEN: Comparative Analysis and Reconstruction of Metabolic Networks

The software CARMEN is a tool, which is designed to support functional and comprehensive analysis. It is used for the reconstruction of organism-specific metabolic networks and based on two main applications, which are pathway information from KEGG database (KGML models) and the generation of system biology Markup Language templates (SBML templates). SBML is an open source XML based file format that expedites the

description and the analysis of the model via tools such as CellDesigner (Funahashi *et al.*, 2003; Kitano *et al.*, 2005).

CARMEN is an effective tool in understanding the insights of metabolic networks of published or unpublished genome sequences. The automatically generated metabolic networks which are stored in standardized SBML format enables user to visualize the network and to detect the metabolic features by means of the combination of biochemical reactions and genome annotation data. The server is available at http://carmen.cebitec.uni-bielefeld.de/cgi-bin/index.cgi and free for academic use (Schneider *et al.*, 2010).

2.3.4. PathPred : Pathway Prediction server

PathPred is another pathway prediction server, which is used for the enzyme-catalyzed metabolic pathways from a query compound. The starting point of this server is the global similarity search of the query compound in the KEGG compound database by the help of SIMCOMP program. In the following step, RDM pattern is used for determining the local similarity of the matched compounds. The last step focuses on the transformation and generation of the query compound after the matched patterns. The Xenobiotics biodegradation (Bacteria) and biosynthesis of secondary metabolites (plants) are the two reference pathways provided by the server for RDM pattern libraries.

The server gives the results as a tree-shaped graph for all predicted multi-step reaction pathways. This server is available at http://www.genome.jp/tools/pathpred/ (Moriya *et al.*, 2010).

2.3.5. FMM: a metabolic pathway reconstruction server

Although there are many servers, which reconstruct metabolic pathways from metabolic maps, it needs a great effort to reconstruct metabolic pathways from two interesting metabolites. For this purpose, a new server, FMM (From metabolite to Metabolite) is introduced in order to reconstruct metabolic pathways from one metabolite to another among different species based on KEGG database and other integrated databases. This server is important and effective for detecting the genes from which species should be cloned into those microorganisms based on FMM pathway comparative analysis. FMM uses a reaction matrix which is developed for the identification of numerous reaction processes.

The data are collected from KEGG/ ligand, KEGG PATHWAY for reaction maps and enzyme lists, from UniProtKB/Swiss-Prot for enzyme annotations and from dbPTM for protein post-translational modifications. After the formation of a reaction matrix is formed by reactions, the pathways from a metabolite to the other one are searched and best maps are selected. Then enzymes are identified in different species and the results show the pathways and maps in the output. FMM is an effective server and it can be used in metabolic engineering and synthetic biology (Chou *et al.*, 2009).

2.3.6. SPA: Selective Permissibility Algorithm for Network Reconstruction

Although the above servers are very useful tools in system biology approaches, they are not applicable in the reconstruction of human wnt signaling network since it is not a metabolic network and the reconstruction has be performed for only one organism. Unfortunately, the studies on the reconstruction and analysis of signaling networks are limited when compared with those of the metabolic and regulatory networks.

Another approach is SPA, in which protein-protein interaction data are integrated with gene ontology data. In SPA method there are two distinctive elements which are input proteins that are known to have a certain function in the pathway and selection criterion. First, all the interactions of the set proteins are extracted from the databases and in the following step, the interactions are filtered according to the selection criterion based on GO annotations. In the next step, the accepted proteins are used as input proteins and this procedure continues until no interaction can be added and all the data are scanned. The input introduced to the algorithm is the adjacency matrix (*S*), which is a binary, square matrix that shows the presence of the interactions between nodes by $S(i,j)=1$ for interacting proteins and $S(i,j)=0$ for non-interacting pairs (Arga *et al.*, 2007).

In this study, this SPA method is modified as the attainable information for human is limited. The new algorithm and the improvements are explained in detail in Materials and Methods section 3.3.1.

3. MATERIALS AND METHODS

The reconstruction of Wnt signaling network together with its sub-networks is performed using the system biology approach in which many types of biological data and analysis techniques are integrated.

3.1. Data

3.1.1. Protein-Protein Interaction(Interactome) Data Sources

MINT (Molecular INTeraction database) is a protein interaction database which is mined from the scientific literature by expert curators.The current version of MINT contains 90467 interactions, 31923 proteins and 4552 pmids (pubmed citations). All data in MINT can be accessed and downloaded freely from MINT home page (http://mint.bio.uniroma2.it/mint/Welcome.do).

HomoMINT, Domino and VirusMINT are the three databases related to MINT. Domino is a domain peptide interaction database, VirusMINT is a virus protein interactions database, and HomoMINT is a database of human interactions (Persico *et al.*, 2005). These interactions are either experimentally verified or inferred from model organisms.The interactions between proteins of model organisms, especially rat, mouse, yeast, worm and fly, are imported into HomoMINT after mapping to the human orthologs. The current version of HomoMINT contains 26913 interactions, 8438 proteins and 2271 pmids. All data in HomoMINT can be accessed and downloaded freely from MINT home page (http://mint.bio.uniroma2.it/HomoMINT/Welcome.do).

BioGRID (Biological General Repository for Interaction Datasets) is another online interaction repository in which data are compiled through comprehensive curation efforts (Stark *et al.,* 2006).The protein interactions are both physical and genetic, and they are collected from different high throughput studies (HTP). BioGRID release version 3.1.75 includes protein interactions for major model organisms such as *Saccharomyces cerevisiae, Caenorhabditis elegans, Drosophila melanogaster* and *Homo sapiens*. Furthermore, BioGRID includes 10226 proteins with 59404 physical interactions for *Homo Sapiens* (Stark *et al.*, 2006).

Human Protein Reference Database (HPRD) is another freely accessible database available at http://www.hprd.org (Prasad *et al.*, 2009). In addition to protein-protein interactions, HPRD also reports interactions of proteins with nucleic acids and small molecules. The PPI data in HPRD are classified as binary or complex interactions based on topology and number of participants. The current version of HPRD includes 30047 unique proteins and 39194 protein interactions.

3.1.2. Domain-Domain Interaction Data Sources

InterPro is a database of protein families, domains, regions and sites in which different databases using different methodologies are combined. It includes the integration of 453 new methods from the HAMAP, PIRSF, SUPERFAMILY, PROSITE, GENE3D, PANTHER and Pfam databases. The regular expressions and Hidden Markov models are the examples of the models used in this database (Hunter*et al.*, 2009). The last release (31.0) includes 5936 domain information and it is accessible at http://www.ebi.ac.uk/interpro/.

PFAM is a database of protein families that are represented by multiple sequence alignment generated by Hidden Markov Model. Pfam has two components: Pfam-A and Pfam-B. The entries of Pfam-A are of high quality, i.e. manually curated families, whereas the entries of Pfam- B are of low quality. However, Pfam-B entries can be used in identifying the functionally conserved regions, if no entry in Pfam-A is obtained (Finn *et al.*, 2008; Finn *et al.*, 2006). The latest release (25.0) contains a total of 12273 families, with 384 new families and it is available at http://pfam.sanger.ac.uk/.

3D Interacting Domains (3did) is another database of domain-domain interactions in proteins for which high-resolution three-dimensional structures are known, and it also contains Gene Ontology-based functional annotations. It uses structural details at molecular level in order to understand how interactions occur. In the last release, there are 159 557 3D structures of domain–domain interactions in 3did, involving 161 996 proteins and the website is available at http://3did.irbbarcelona.org/ (Stein, Céol and Aloy, 2011).

3.1.3. Gene Ontology Annotations

The Gene Ontology (GO) project supplies vocabulary (ontology) for describing the roles of genes and gene products and it is used in addressing the need for consistent descriptions of gene products in different databases. Thus, it is possible to transfer knowledge of the shared biological roles from one organism to another by means of the dynamic set of descriptions used in Gene Ontology project. The three ontologies defined by Gene Ontology Consortium are cellular component, biological process and molecular function and they are available at http://www.geneontology.org.

Cellular component can be defined as the location of the protein in the cell where it is active.A gene product may be located in a particular cellular component or it may be in a subcomponent of the cellular component. Cytosol, nucleus and nuclear membrane are the examples of the cellular components. Biological process refers to a biological objective in which a series of events are accomplished by the genes or gene products. Signal transduction, cell maintenance and multi-cellular organism development are the basic examples of the biological processes. Molecular function refers to the activities of the gene product such as catalytic activity, transporter activity and protein binding that occurs at the molecular level (The Gene Ontology Consortium, 2000).

3.2.Data Visualization

In construction of biological networks, softwares employed by computational methods are very purposive, and they are helpful for systems biology approach. Cytoscape is free open source bioinformatics software (http://www.cytoscape.org) for visualizing and analyzing biomolecular interaction networks. Cytoscape is enriched with many plug-ins which have different goals. Network analyzer (Assenov *et al.*, 2008), MCODE (Bader *et al.*, 2003) and BINGO (Maere *et al.*, 2005) are the freely available plugins also used in this study and they are described in detail in Methods in section 3.3.

3.3.Methods

3.3.1. Network Reconstruction by Interactome Data (PPI) and GO Annotations

The reconstruction of Wnt signaling network in human was performed by integrating Gene Ontology (GO) annotations and PPI data. First of all, the core proteins that are known to have roles in Wnt signaling in *Homo Sapiens* were obtained from literature. Then an annotation collection table was prepared by using GO annotations (biological process, cellular component and molecular function) of the core proteins. All the GO annotations of all human proteins were downloaded from the file *gene_association.goa_human.gz* from GO website (http://www.geneontology.org/GO.current.annotations.shtml).

Next, the human proteins whose all three GO terms showed exact match with the annotation collection table were accepted to the network. Thereafter, the physical protein-protein interaction data among these human proteins were extracted from MINT (Molecular Interactions Database), homo_MINT (an inferred human network), HPRD (Human Protein Reference Database) and BioGRID (The Biological General Repository for Interaction Datasets). Finally, the interaction partners of the proteins that passed the GO annotation filter were obtained and these protein pairs were used in the reconstruction of the network. The algorithm used in the network reconstruction is illustrated in Figure 3.1.

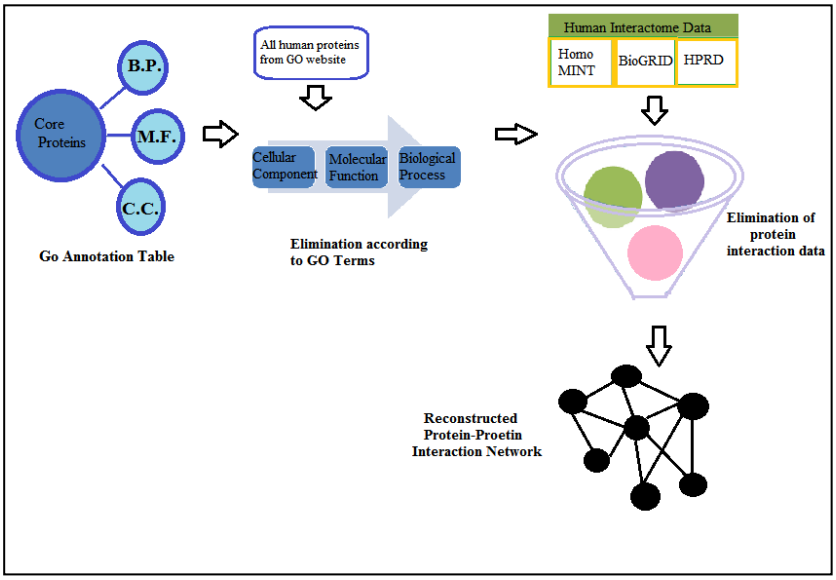

Figure 3.1. The schematic representation of the algorithm used for the network reconstruction.

3.3.2. Reconstruction of Domain Domain Interaction

The protein-protein interactions have an important place in systems biology due to their role in cellular processes especially in identifying the protein functions. However, the protein interaction data are not enough for determining how these interactions occur. The domain-domain interactions may be helpful in order to get a deep understanding of these mechanisms, which then may lead to deciphering the interaction mechanisms underlying the diseases and consequently designing the relevant drugs.

Domains are the basic parts of the proteins, and are responsible for performing specific function. In addition to that, domain interactions are also used in the prediction of protein-protein interactions (Spinzak and Margalit, 2001; Nguyen and Ho, 2006). For this purpose the domain-domain interactions of the wnt signaling network were studied and analyzed here. First of all, the domains of the human proteins were extracted from Interpro (protein2ipr.dat.gz) and the protein-protein interactions of the reconstructed wnt sub-networks were converted into domain-domain interactions via a code written in Matlab.

In the computational prediction of domain interactions, the association method was used to determine the over-represented domain pairs in interacting protein pairs. In association method, the log ratio of the frequency of occurrences of domains in interacting proteins to the frequency of independent occurrences of those domains are calculated by a statistical scoring system that is called PID, Potentially Interacting Domain pairs (Sprinzak and Margalit, 2001). The proteins that have no domain information were ignored in calculating PID (Potentially Interacting Domain) scores.

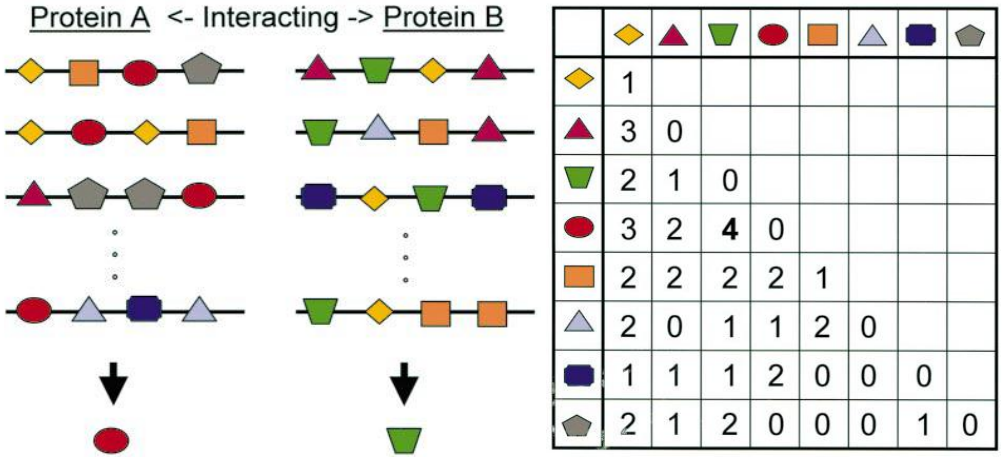

Figure 3.2. A schematic representation of analysis for the study of correlated sequence-signatures as markers of protein-protein interaction (Sprinzak and Margalit, 2001).

In this study, it is assumed that if a domain pair (a, b) are found in an interacting protein pair (A, B), the domain pair (a, b) is considered as 'Potentially Interacting Domain pair (PID)'. If a PID (a, b) is found more frequently than expected by random probability, this domain pair can be thought as a truly interacting domain pair. The domain set of protein A is defined as A = [a_1 , a_2 , . . ., a_i , . . ., a_m] and that of protein B is B= [b_1 , b_2 , . . ., b_j , . . ., b_n]. Then the related frequencies of the domains a_i and b_j are given as E(a_i) and E(b_j), respectively. Then PID score is calculated as follows;

$$PID\ score(a_i, b_j) = \sum_{AB} \log(F(a_i, b_j)/E(a_i, b_j)) - treshold \qquad (3.1)$$

where m, n : numbers of domain species in protein A, B

- E(a_i) , E(b_j): frequencies of domain a_i and b_j
- F(a_i, b_j) = 1/(n × m) : weighted frequency of PID (a_i, b_j)
- E(a_i, b_j) = E(a_i)E(b_j) × 2 expected frequency of PID (a_i, b_j) in a randomly chosen protein pair (reciprocal domain pairs of (a_i, b_j) and (b_j, a_i) are merged to one PID)
- T_{int}: number of total interaction in training set
- D_{avg}: average number of domain species per protein
- Threshold = max {log ((1/E (ai, bj)),T_{int}×E (a_i, b_j) ×log ($(1/D_{avg})^2$/E (ai, bj))}

In summary, firstly, the domains of the proteins were extracted from the whole domain set of *Homo sapiens* using InterPro database. Then, the protein-protein interactions are converted into domain-domain interactions. The frequency values of each domain pair were obtained and used in PID score calculations. In our calculations, the random

probability value was taken as the average probability of the entire reconstructed ddi network. The most frequent domain pairs were examined since this enables us to comment about the most probable protein interactions and to predict potential interactions for proteins pairs that have no interaction data in literature but have one or more domain pairs at high probability of interaction occurrence. Finally, the possible domain pairs obtained by this method were verified with the experimental domain-domain interaction data taken from Pfam and 3did databases.

3.3.3. Graph Theoretical Analysis

Graph theoretical analysis enables better understanding of the structure of the complex networks and the distribution of the components via analyzing topological properties. The topology of the reconstructed network was determined by Network Analyzer Plug-in of Cytoscape (ver. 2.7.0) which computes several topological properties of both directed and undirected networks given. In the network analysis, the proteins are called nodes whereas the interactions between the nodes are called edges. The input is the binary protein interactions and output is the topological parameters such as the degree distribution of the nodes, the number of highly connected nodes (hubs), network diameter, mean path length, clustering coefficient, number of shortest paths and average shortest path length.

Degree (Connectivity): Degree (k) is used to determine the number of the edges of a node with the others.

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.

Network Diameter: The network diameter is the maximum non-infinite length of a shortest path from a node to another node in the network. In other words, the network diameter can be defined as the maximum length of shortest paths between two nodes.

The shortest path length: The path length is the total number of edges that is needed to be crossed over from one node to another. The shortest path length is the smallest number of

edges between two nodes whereas the average shortest path length (or the characteristic path length) is the average expected distance between two connected nodes.

Clustering coefficient: The clustering coefficient of a node *N* can be defined as the ratio of the number of edges between the neighbors of *N* to the maximum number of edges that could possibly exist between the neighbors of *N*. The network clustering coefficient is the average of the clustering coefficients for all nodes in the network. This value is always between 0 and 1 (Alterovitz *et al*., 2005).

Betweenness centrality: It gives the total number of non-redundant shortest paths going through a certain node or edge. In other words, it can be explained by the measure of the centrality of the nodes in a network and most of the shortest paths in a network go through the nodes with high betweenness centrality and these nodes are called as the bottleneck proteins (Yu *et al.*, 2007; Freeman, 1977; Girvan and Newman; 2002).

Until recently, biological networks have been modeled as random networks. In random networks, the connection of the nodes are related to a probability, p that follows a Poisson distribution ($P(k) < e^{-k}$) in which most of the edges have the same number of connections and nearly approximately equal to the mean degree of the network. Depending on the recent studies, many social and biological networks cannot be regarded as random networks. Scale-free networks are found to be more suitable to describe these systems, in which P (k) follows a power-law distribution, $P(k) \approx k^{-\gamma}$, where γ is defined as degree exponent. In these networks, there are a few highly connected nodes (hubs) and the other nodes, which have smaller number of edges, are linked to network via these hubs. Bader and Houge (2003) showed that protein interaction networks and metabolic networks both have scale-free topology (Jeong *et al.*, 2000; Barabási and Oltvai, 2004). A comparison of random networks and scale-free networks is illustrated in Figure 3.3.

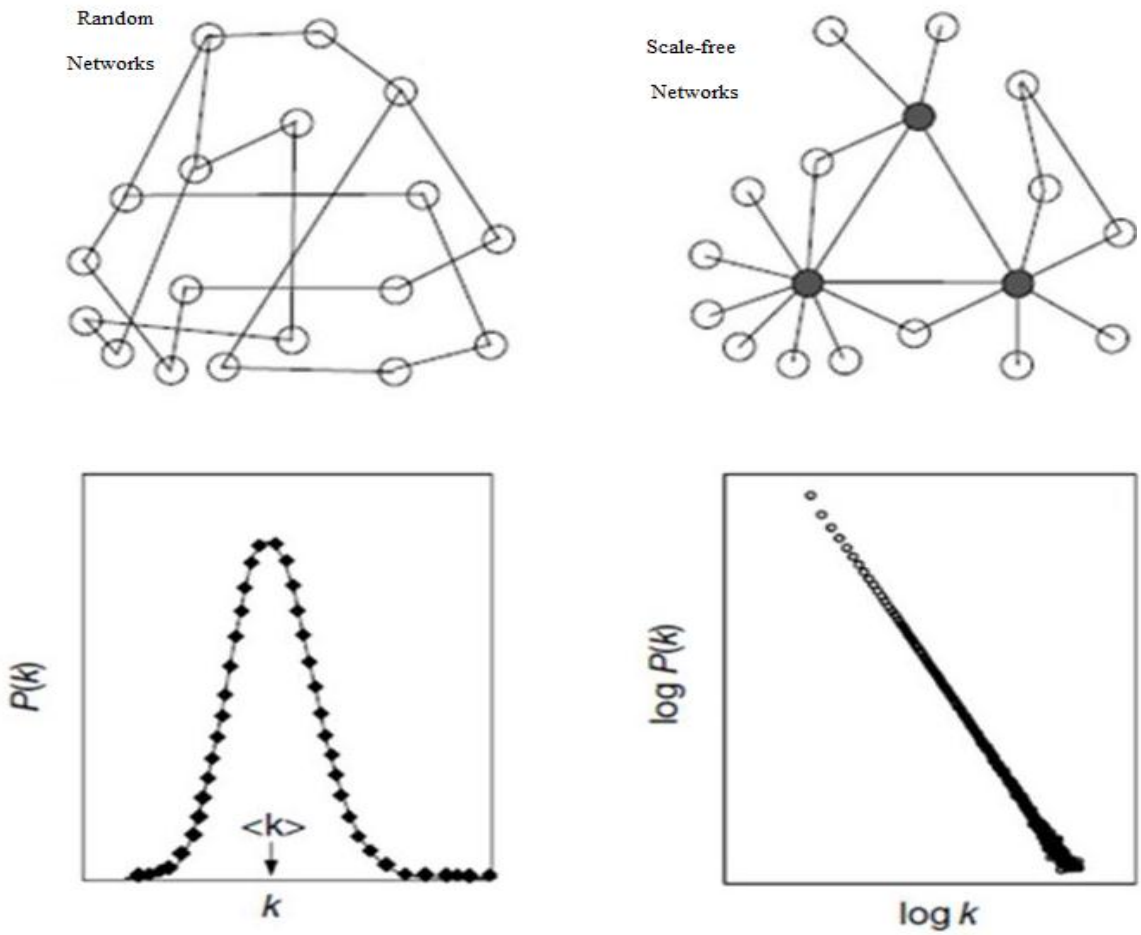

Figure 3.3. Representative structures of Networks and their distributions (Jeong*et al.*, 2000)

3.3.4. Module Detection and Analysis

Scale-free networks are known to be composed of clustered regions and in biological networks these clustered regions correspond to molecular complexes named as modules. In other words, these modules are the sub-graphs of the network that performs a molecular function together. In order to get an insight on the cellular role of the network, a Cytoscape plug-in, MCODE is designed to identify these highly interconnected regions in the present protein-protein interaction network.

The MCODE algorithm depends on three stages which are vertex weighting, complex prediction and optionally post-processing for filtering or adding proteins in the resulting complexes by certain connectivity criteria.In the first step, MCODE weights all vertices

according to the local network density by using the minimum degree (k-score). The density which ranges between 0.0 and 1.0 is calculated by dividing the number of connections to the theoretical maximum number of connections. A scale-free network performs a vertex connectivity distribution that follows a power law which has a few highly connected regions whereas the rest of the network has smaller number of degrees. In the second step, the highest vertex is taken as seed and MCODE moves outward from the seed vertex including the vertices above a given threshold. In the last step, the vertices which have lower degree than threshold, 2, are eliminated (Bader and Houge, 2003).

After the detection of the modules, the statistically overrepresented Gene Ontology functions are obtained by the help of the Biological Networks Gene Ontology tool (BINGO). BiNGO combines the relevant GO annotations and relates them upwards through the GO hierarchy. The output of this analysis gives the most significant GO annotations with the lowest P-value. The hypergeometric test and the binomial test are the two statistical tests that are used by BiNGO. In hypergeometric test, probabilities of a set of genes (detected modules) in a given reference set (all GO annotations) are given in p-value form whereas in binominal test an approximate P-value is obtained in a shorter time (Maere *et al*., 2005). In the new version of Cytoscape (2.7.0), these two plug-ins are replaced by ClusterViz plug-in.

3.3.5. Network Decomposition Analysis

Network decomposition analysis allows us to decompose the protein interaction networks into linear paths in which ligands are taken as inputs and transcription factors are taken as outputs. This analysis enables us to investigate the network communication route via linear paths. In this study, the linear paths of the reconstructed network were found by the NetSearch program (Steffen *et al.*, 2002). At a specified length, all the linear paths are obtained using the interaction list from input protein to output protein. An important approach for determining the essentiality of the proteins in the network is to calculate the participation frequencies of the proteins. Therefore, the participation percentage of each protein in the reconstructed network is calculated for a better understanding of its contribution to the signaling network through linear paths.

3.3.6. Crosstalk Analysis

Signaling networks are communicating systems and they form complex signaling networks due to the interaction with each other instead of behaving in isolation. Therefore, transmittal of knowledge occurs through the node interactions that are related to a particular cellular process. Crosstalk analysis is a way to analyze i) the components that connect the pathways in signal transduction and ii) the responses that result in different conditions in signaling, in order to design suitable therapeutic agents against a disease caused by the disturbance of these components.

Crosstalk analysis is mainly done by two ways, i.e., network crosstalk and path crosstalk. The network crosstalk of a node can be defined as the difference in degree of the node in all considered networks and the maximum degree of this node in any individual pathway. If a node has a high network crosstalk value then it means that this component is a branch node connecting two or more pathways. The second type of crosstalk is the path crosstalk. In this type, signal-flow centrality, which means the shortest paths from ligands to transcription factors, is considered. The path crosstalk of a node can be defined as the difference between its signal-flow centrality for the entire network and its maximum signal-flow centrality in any individual pathway. A high path crosstalk value implies that a node is more important in the combined network than it was in the individual pathways (Zielinski *et al.*, 2009; Tekir *et al.*, 2009).

In the crosstalk analysis of this study, common nodes in three wnt signaling subnetworks were found, and the network crosstalk values were calculated since higher network crosstalk value of a node increases its importance in the pathways for taking a branch role between these pathways. The nodes with high crosstalk value were accepted as the bridging nodes.

3.3.7. Reduction of the Network

When studying the molecular processes in the cell, the network studies come into prominence due to being effective tools. However, analyzing the data becomes more complicated when the size of the considered networks is too large. Network reduction is a

common process used for simplicity. Hence, in this study, the reduction of the network for Wnt signaling in human was performed for a better analysis of the whole network.

For this purpose, a matlab code written by Dr. Çakır for the reduction of networks was used. The MatlabBGL package which is designed by David Gleich to work with large sparse graphs with thousands of nodes was used in here. The code first constituted the adjacency matrix and then the graph theory aspects such as betweenness centrality, network diameter etc. were calculated without disturbing the degree values. In the second step, hundred random networks were constructed and analyzed. Following the random network analyses, the application of t-test analysis enabled the elimination of relatively trivial nodes. The t-test, which determines whether the means of two groups are statistically different from each other, is a statistical hypothesis test that can be explained as the ratio of the differences between two group means to the variability of the groups.

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}} \tag{3.2}$$

Where $\overline{x}$: the mean of the betweenness centrality of the random networks,

μ_0 : The betweenness centrality value of the nodes obtained from the first step,

S: Standart deviation of the samples in random networks,

n: number of random networks.

After the t- value is determined, a p- value is then obtained by using distribution tables. If the calculated p- value is found to be below the threshold, which is taken as 0.01 for statistical significance of this case, then the null hypothesis is rejected. In this study, the nodes which have negative values after the probability calculations are eliminated from the network since the biological networks are modeled as scale-free networks instead of random networks.

4. RESULTS AND DISCUSSION

In the present study, the protein-protein interaction sub-networks of Wnt signaling were reconstructed in *Homo sapiens* and the structural analyses of these sub-networks were then performed. The reconstruction process of protein interaction networks may undergo some difficulties due to the false positives, i.e. interactions determined in vivo but actually not identified in the cell, and false negatives, i.e. interactions are not experimentally detected but identified in the cell. Besides that, the metabolic and regulatory networks are more prevail in the reconstruction and analysis of network studies than signaling networks. Recently, there are several attempts on this topic which try to integrate different biological features (Lu *et al*., 2005; Patil and Namakura, 2005). In order to handle this problem and to improve the quality of the data, a computational approach was here used by integrating Gene Ontology (GO) annotations with interactome data.

4.1. Reconstruction of Wnt Signaling Networks

In the first step, the proteins that are known to be involved in the Wnt signaling in *Homo Sapiens* were identified for through literature search. The NetPath and KEGG databases were also used as good starting points. These selected Wnt proteins that belong to both canonical and non-canonical Wnt pathways in human were taken as the core proteins of the sub-networks to be constructed. The list of the core proteins of Wnt/β-catenin (canonical), Wnt/PCP and Wnt/Ca^{2+} (non-canonical) pathways are listed in Tables 4.1 and 4.2, respectively.

For canonical Wnt pathway 68 core proteins were identified. For PCP and Ca++ pathways these numbers were 33 and 32, respectively. For canonical pathway, 10592 physical protein interactions, which were obtained from MINT, BioGRID and HPRD databases, were accepted to the network. For PCP network 5928 protein-protein interactions

Table 4.1. Core proteins of canonical Wnt signaling pathway.

General Name(Canonical)-Uniprot ID			General Name(Canonical)-Uniprot ID		
WNT	WNT1	P04628	AXAM	SENP2	Q9HC62
	WNT2	P09544	DKK	DKK1	O94907
	WNT2B	Q93097	NKD	NKD2	Q969F2
	WNT3	P56703		NKD1	Q969G9
	WNT3A	P56704	IDAX	CXXC4	Q9H2H0
	WNT4	P56705	SKP1	SKP1	P63208
	WNT7A	O00755	CUL1	CUL1	Q13616
	WNT10B	O00744	NLK	NLK	Q9UBE8
FZD	FZD1	Q9UP38	PROTEIN52	RUVBL1	Q9Y265
	FZD2	Q14332	SMAD4	SMAD4	Q13485
	FZD4	Q9ULV1	SMAD3	SMAD3	P84022
	FZD5	Q13467	CtBP	CTBP1	Q13363
	FZD7	O75084		CTBP2	P56545
LRP	LRP5	O75197	TAK1	MAP3K7	O43318
	LRP6	O75581	LEF	LEF1	Q9UJU2
DVL	DVL1	O14640	TCF	TCF7	P36402
	DVL2	O14641		TCF7L1	Q9HCS4
	DVL3	Q92997	BETA-TRCP	BTRC	Q9Y297
GBP	FRAT1	Q92837	SIAH1	SIAH1	Q8IUQ4
	FRAT2	O75474	CBP	EP300	Q09472
GSK3B	GSK3B	P49841		FBXW11	Q9UKB1
AXIN	AXIN1	O15169	PS1	PSEN1	P49768
	AXIN2	Q9Y2T1	WIF1	WIF1	Q9Y5W5
APC	APC2	O95996	PROC	PORCN	Q9H237
	APC	P25054	CER1	CER1	O95813
	PPP2CA	P67775	FRP	SFRP1	Q8N474
CKIALPHA	CSNK1A1	P48729		SFRP2	Q96HF1
	CSNK1A1L	Q8N752		SFRP4	Q6FHJ7
	CSNK1D	P48730		SFRP5	Q5T4F7
	CSNK1E	P49674	XSOX17	SOX17	Q9H6I2
CK2	CSNK2A2	P19784	DUPLIN	CHD8	Q9HCK8
	CSNK2B	P67870	EBI1	TBL1X	O60907
BETA-CAT	CTNNB1	P35222		TBL1XR1	Q9BZK7
ICAT	CTNNBIP1	Q9NSA3	SIP	CACYBP	Q9HB71

Table 4.2. Core proteins of non-canonical Wnt signaling pathway.

General Name (PCP)-Uniprot ID			General Name (Ca++)-Uniprot ID		
WNT	WNT5A	P41221	WNT	WNT5A	P41221
	WNT5B	Q9H1J7		WNT11	O96014
	WNT11	O96014		WNT1	P04628
FZD	FZD3	Q9NPG1	PLC	PLCB1	Q9NQ66
	FZD2	Q14332		PLCB2	Q00722
	FZD6	O60353		PLCB3	Q01970
	MAGI3	Q5TCQ9		PLCB4	Q15147
	ROR1	Q01973	CaMKII	CAMK2A	Q9UQM7
	ROR2	Q01974		CAMK2B	Q13554
	PTK7	Q13308		CAMK2D	Q13557
Stbm	VANGL1	Q8TAA9		CAMK2G	Q13555
	VANGL2	Q9ULK5	CaN	CHP	Q99653
Knypek	CELSR1	Q9NYQ6		PPP3CA	Q08209
	CELSR2	Q9HCU4		PPP3CB	Q8N1F0
	CELSR3	Q9NYQ7		PPP3CC	P48454
DVL	DVL1	O14640		PPP3R1	P63098
	DVL2	O14641		PPP3R2	Q96LZ3
	DVL3	Q92997		CHP2	O43745
PRINCKLE	PRINCKLE1	Q96MT3	PKC	PRKCA	P17252
	PRINCKLE2	Q7Z3G6		PRKCB	P05771
Nkd	NKD1	Q969G9		PRKCG	P05129
	NKD2	Q969F2	NFAT	NFAT5	O94916
	ANKRD6	Q9Y2G4		NFATC1	O95644
DAAM1	DAAM1	Q9Y4D1		NFATC2	Q13469
	DAAM2	Q86T65		NFATC3	Q12968
RhoA	RHOA	P61586		NFATC4	Q14934
ROCK2	ROCK1	Q13464	Frizzled	FZD2	Q14332
	ROCK2	O75116		FZD3	Q9NPG1
RAC	RAC1	P63000		FZD4	Q9ULV1
	RAC2	P15153		FZD6	O60353
JNK	MAPK8	P45983	NLK	NLK	Q9UBE8
	MAPK9	P45984			
	MAPK10	P53779			

and for Ca++ network 6080 protein-protein interactions were accepted to these sub-networks. The number of the core proteins, nodes and interactions of the sub-networks are given in Table 4.3.

Table 4.3. The quantitative data of the Wnt signaling sub-networks.

Sub-Network	Number of Cores	Number of Interactions	Number of Nodes	Number of Edges
Wnt/β-catenin (Canonical)	68	10592	3251	9304
Wnt/PCP	33	5928	1952	5001
Wnt/Ca^{2+}	32	6080	2112	5293

However, 6 core proteins could not be included into the whole Wnt network as they do not have physical interactions. The proteins with missing interactions in canonical pathway are Wnt10B, SFRP4 and SFRP5. Wnt 11 and DAAM2 are the unwired core proteins of PCP pathway, and Wnt 11 and PPP3R2 are the unwired ones for calcium pathway.

After the reconstruction process was completed, the isolated smaller parts were removed and the sub-networks were further analyzed. For Wnt/β-catenin pathway a network of 3251 nodes and 9304 edges was obtained. For PCP pathway 1952 nodes and 5001 edges were obtained. Finally, for Calcium pathway 2112 nodes and 5293 edges were obtained. Table 4.3 shows the overview of the reconstructed networks..

Finally, the sub-networks were integrated and the whole Wnt network was analyzed. For this network, 3489 nodes and 10092 edges was obtained. The common proteins of each sub-network were identified and the numbers and the percentages of the common nodes are given in Table 4.4 and Figure 4.1. There are 1474 proteins which correspond to 42 % of all proteins, obtained in all Wnt sub-networks.

Table 4.4 Quantitative information of common nodes.

Networks	Number of common nodes	Total number of nodes	Percentages
Canonical-Ca^{2+}-PCP Wnt (human)	1474	3484	42.31 %
Canonical-PCP Wnt (human)	1837	3366	54.57 %
Canonical-Ca^{2+} Wnt (human)	1951	3412	57.21 %
PCP-Ca^{2+} Wnt (human)	1517	2547	59.56 %

When the common nodes are examined, it is seen that dishvelled (DVL1, DVL2, DVL3), axin (AXIN1, AXIN2), Casein Kinase (CSNK1A1, CSNK1D, CSNK1E, CSNK2A1, CSNK2B), beta-catenin (CTNB1), Frizzled (FZD1, FZD2, FZD4, FZD7, FZD8, FZD9), Adenomatous polyposis coli protein (APC), Glycogen synthase kinase-3 beta (GSK3β), Mitogen-activated protein kinase (MAP3K2, MAP3K4, MAP3K7, MAP3K8), Nuclear factor of activated T-cells (NFATC1, NFATC2), Smad family (SMAD 1, SMAD 2, SMAD 3, SMAD 4, SMAD 5, SMAD 6, SMAD 7, SMAD 9), Transcription factor 7-like 2 (TCF7L2), Cellular tumor antigen p53(TP53), Protein kinase C inhibitor protein 1 (YWHAZ) proteins exist in all there sub-networks. The total list of these proteins is given in Appendix B. The relationship of these common proteins to diseases is given in Table 4.5. A dysfunction of these proteins is reported to cause mainly cancer.

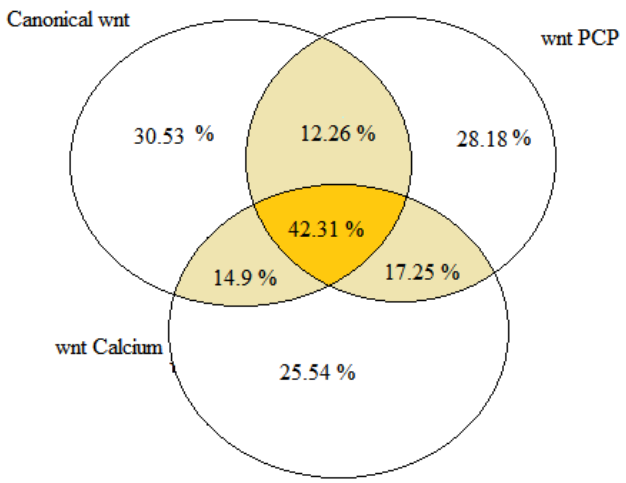

Figure 4.1. Venn Schema Representation of Common Nodes.

Table 4.5. The common proteins found to be related to diseases.

Protein	***Disease***	***References***
Beta catenin	Carcinogenesis, hepatocellular carcinomasWilms' tumors	Klaus and Birchmeier,2008; Maiti *et al.*,2000
DVL	Lung cancer	Yang *et al.*,2010
FZDs	Gastric cancer,colorectal cancer& carcinogenesis	Kirikoshi, Sekihara and Katoh, 2001;Ueno *et al.*,2008
APC	Colorectal cancer, carcinogenesis	Klaus and Birchmeier, 2008; Ueno *et al.*, 2008
KC1AL	Alzheimer Disease	Li,Yin and Kuret, 2004
YWHAZ	Breast cancer Obesity, Diabetes	Peng, Wang and Shan, 2009
sFRP(s)	colon cancer, mesothelioma, bladder cancer	Tan and Kelsey, 2009; Paul and Dey, 2008; Gehrke, Gandhirajan and Kreuzer, 2009.
GSK-3β	colorectal cancer	Ge and X. Wang, 2010
Smad3	Osteoarthritis	Valdes*et al*, 2010

4.2. Reconstruction of Domain Domain Interaction Networks for Wnt Signaling

Domains are the structural/functional units of the proteins which contain specific signature sequences that conserved through evolution to perform the function of the protein. Since not only the investigation of protein interactions but also that of domain interactions is important for understanding the roles of proteins in biological networks, the domain-domain interactions (DDIs) of the reconstructed signaling sub-networks were found and analyzed for this purpose. It was also aimed to computationally predict the missing protein-protein interactions through *in silico* domain-domain interactions. The domains of the human proteins were extracted from Interpro (protein2ipr.dat.gz) and the protein interactions of the reconstructed sub-networks were converted into domain interactions via a code written in Matlab (Appendix A).

In the computational prediction of domain interactions, the *association method* was used to determine the over-represented domain pairs in interacting protein pairs. The log ratio of the frequency of occurrences of domains in interacting proteins to the frequency of independent occurrences of those domains was used in the calculation of PID, Potentially Interacting Domain pairs (Sprinzak and Margalit, 2001). The domain pairs which have higher scores than the average score were taken as the potentially interacting domain pairs and they were compared with the experimental domain interaction data downloaded from Domine Database. In order to improve the accuracy, both Pfam and 3did domain databases were also taken into account.

For canonical Wnt signaling sub-network, the obtained numbers of unique domains and domain interactions are 1784 and 35569, respectively. For non-canonical sub-networks 1211 unique domains and 44492 domain interactions are obtained in case of PCP signaling, whereas these numbers are 1204 and 23465 in case of Wnt/Calcium signaling, respectively. These results are given in Table 4.6.

Table 4.6. Summary of DDI analysis.

Model	**# of unique proteins**	**# of protein interactions**	**# of unique domains**	**# of domain interactions**
Wnt/β-catenin (*H. Sapiens*)	3201	9304	1784	35569
Wnt /PCP Signaling (*H. Sapiens*)	1987	5001	1211	44492
Wnt/Ca^{+2} (*H. Sapiens*)	2095	5293	1204	23465

When the frequency values of these domain pairs are examined, it is seen that the self interaction of protein kinase domain (PF00069-PF00069) is the most frequent one. In addition to that, the domain pair, SH3 domain (PF00018) and protein kinase domain (PF00069), shows remarkable frequency. SH3 (Src homology 3) domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization which also implicates Wnt signaling (http://pfam.sanger.ac.uk). Another significant

interaction is between the protein kinase domain (PF00069) and PDZ domain (PF00595) which are found in diverse signaling proteins. The domain pairs and related frequency values are given in Table 4.7.

Table 4.7. The highest frequency values of the domain pairs of canonical Wnt signaling.

Domain_ID	Domain_Name	Domain_ID	Domain_Name	Frequency
PF00069	Protein kinase domain	PF00069	Protein kinase domain	2,45E-03
PF00018	SH3 domain	PF00069	Protein kinase domain	1,69E-03
PF00069	Protein kinase domain	PF00018	SH3 domain	1,69E-03
PF00017	SH2 domain	PF00069	Protein kinase domain	1,63E-03
PF00069	Protein kinase domain	PF00017	SH2 domain	1,63E-03
PF00069	Protein kinase domain	PF07714	Protein tyrosine kinase	1,54E-03
PF07714	Protein tyrosine kinase	PF00069	Protein kinase domain	1,54E-03
PF00069	Protein kinase domain	PF00595	PDZ domain	1,45E-03
PF00595	PDZ domain	PF00069	Protein kinase domain	1,45E-03
PF00018	SH3 domain	PF00018	SH3 domain	1,17E-03

The highest PID (potentially interacting domain pairs) score obtained for canonical Wnt DDI sub-network is 18.80 for PF04857-PF07742 domain pair which are members of CAF1 family and BTG family, respectively. The domains of BTG family are important for having 3 conserved cysteine residues and they are involved in negative regulation of cell proliferation. In addition to that, BTG2 seems to have a signal sequence (Pfam-database, http://pfam.sanger.ac.uk). The *in silico* interaction of this domain pair was checked in DOMINE database (Database of Protein Domain interactions-http://domine.utdallas.edu) and it is seen that BTG (PF07742) forms interactions with CAF1 family ribonuclease (PF04857). These interactions were verified experimentally and reported in 3did database. In this ddi network, the number of domain pairs which have higher PID scores than average is 17690 (49.7 %). The experimental data of Pfam and 3did databases were used for validation of predictions and it is seen that 1196 computationally predicted domain pairs are in agreement with the experimental data and that corresponds to nearly 7%

accuracy. The accuracy increases up to 26 % when the highest top fifty PID scores are taken into account.

When the frequency values of Wnt/PCP signaling domain pairs are examined, it is again seen that the protein kinase domain pair (PF00069-PF00069) has the highest frequency value. In addition to that, other interaction partners of kinase domain such as SH3 domain (PF00018), PDZ domain (PF00595) and protein tyrosine kinase domain (PF07714) show remarkable significance in these putative interactions. The domain pairs and related frequency values are given in Table 4.8. The number of domain pairs which have higher PID scores than average is 11159 (49.65 %). The highest PID score obtained for this DDI sub-network is 17.99 again for PF04857-PF07742 domain pair which are members of CAF1 family and BTG family, respectively. It is seen that only 878 computationally predicted domain pairs are consistent with the experimental data and that corresponds to nearly 8 % accuracy. The highest accuracy percentage obtained is 16 when the top fifty PID scores are taken into account.

Table 4.8. The highest frequency values of the domain pairs of Wnt/PCP signaling.

Domain_ID	**Domain_Name**	**Domain_ID**	**Domain_Name**	**Frequency**
PF00069	Protein kinase domain	PF00069	Protein kinase domain	2,63E-03
PF00018	SH3 domain	PF00069	Protein kinase domain	2,33E-03
PF00069	Protein kinase domain	PF00018	SH3 domain	2,33E-03
PF00017	SH2 domain	PF00069	Protein kinase domain	2,32E-03
PF00069	Protein kinase domain	PF00017	SH2 domain	2,32E-03
PF00069	Protein kinase domain	PF07714	Protein tyrosine kinase	2,20E-03
PF07714	Protein tyrosine kinase	PF00069	Protein kinase domain	2,20E-03
PF00018	SH3 domain	PF00018	SH3 domain	2,06E-03
PF00017	SH2 domain	PF00018	SH3 domain	2,05E-03
PF00018	SH3 domain	PF00017	SH2 domain	2,05E-03
PF00017	SH2 domain	PF00017	SH2 domain	2,04E-03
PF00018	SH3 domain	PF07714	Protein tyrosine kinase	1,95E-03
PF07714	Protein tyrosine kinase	PF00018	SH3 domain	1,95E-03
PF00069	Protein kinase domain	PF00595	PDZ domain	1,94E-03
PF00595	PDZ domain	PF00069	Protein kinase domain	1,94E-03

When the frequency values of Wnt/Ca^{+2}signaling domain pairs are examined, it is seen that the self interaction of protein kinase domain (PF00069) is the most frequent one. In addition to that, again, SH3 domain, PDZ domain and protein tyrosine kinase domain show remarkable frequency values in their interactions with this protein kinase domain (PF00069). In addition to that, another specific domain C2 domain (PF00168) shows high frequency values when paired with protein kinase domain. The domain pairs and related frequency values are given in Table 4.9.

Table 4.9. The highest frequency values of the domain pairs of Wnt/ Ca^{+2}signaling.

Domain_ID	Domain_Name	Domain_ID	Domain_Name	Frequency
PF00069	Protein kinase domain	PF00069	Protein kinase domain	5,55E-03
PF00069	Protein kinase domain	PF07714	Protein tyrosine kinase	3,35E-03
PF07714	Protein tyrosine kinase	PF00069	Protein kinase domain	3,35E-03
PF00018	SH3 domain	PF00069	Protein kinase domain	3,22E-03
PF00069	Protein kinase domain	PF00018	SH3 domain	3,22E-03
PF00017	SH2 domain	PF00069	Protein kinase domain	3,17E-03
PF00069	Protein kinase domain	PF00017	SH2 domain	3,17E-03
PF00168	C2 domain	PF00069	Protein kinase domain	2,16E-03
PF00069	Protein kinase domain	PF00168	C2 domain	2,16E-03
PF07714	Protein tyrosine kinase	PF07714	Protein tyrosine kinase	2,02E-03
PF00069	Protein kinase domain	PF00069	Protein kinase domain	5,55E-03
PF00069	Protein kinase domain	PF07714	Protein tyrosine kinase	3,35E-03
PF07714	Protein tyrosine kinase	PF00069	Protein kinase domain	3,35E-03
PF00018	SH3 domain	PF00069	Protein kinase domain	3,22E-03
PF00069	Protein kinase domain	PF00018	SH3 domain	3,22E-03

The number of domain pairs which have higher PID scores than average is 11750 (50.07 %). The highest PID score obtained for Wnt/Ca^{2+}DDI sub-network is 18.08 for again PF04857-PF07742 domain pair which are members of CAF1 family and BTG family, respectively. This domain pair always has the highest score for all three DDI sub-networks. It is seen that 729 computationally predicted domain pairs are in agreement with the experimental data of Pfam and 3did and that corresponds to nearly 6 % accuracy. 30 percent accuracy is obtained for this ddi network when the highest fifty PID scores are taken into account.

The highest PID score is obtained for PF04857-PF07742 domain pair and this domain pair is seen in CAF2 (Q9UFF2) and BTG1 (P62324) protein pair. CAF2 has a role in negative regulation of cell proliferation and sequence-specific DNA binding transcription factor activity whereas BTG1 is involved in negative regulation of cell growth, negative regulation of cell proliferation, positive regulation of endothelial cell differentiation, regulation of apoptosis and transcription cofactor activity.

Depending on these results it can be stated that higher accuracy percentages are obtained when top 25 PIDs are taken into account. The accuracy percentages are increasing up to top 25 PIDs and then it decreases with increasing number of domain pairs taken among the top scored group. The accuracy values are shown in Figure 4.2 and the experimental verifications are listed in Tables 4.10. The list of accuracy values for different number of domain pairs with highest PIDs are given in Appendix C.

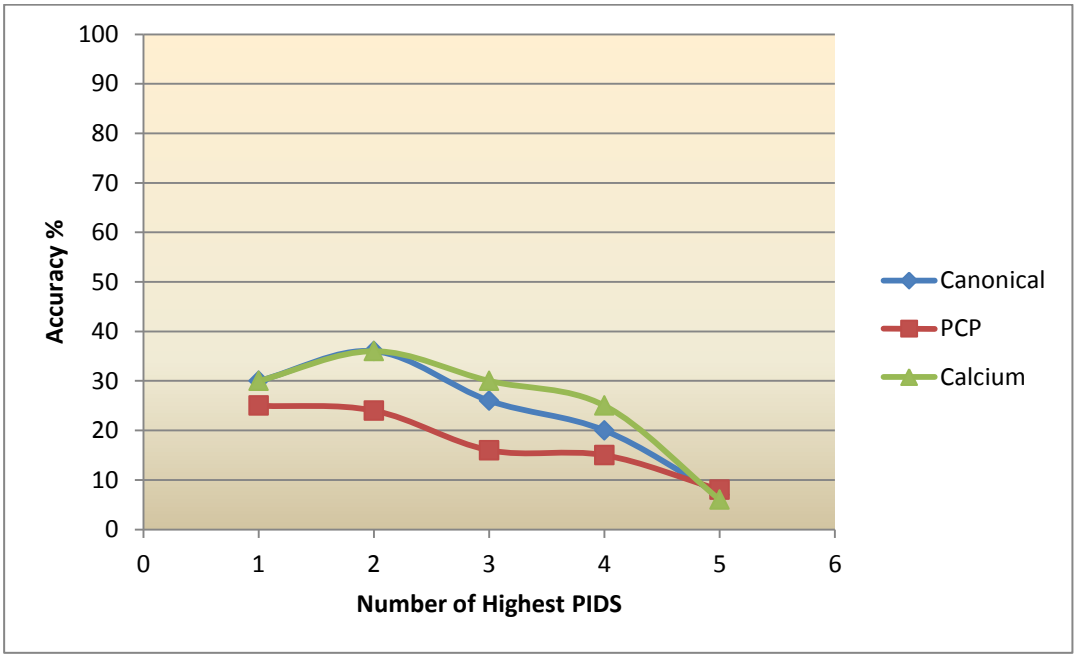

Figure 4.2. Accuracy percentage vs. number domain pairs with highest PIDs.
(1) Top 20 PID pairs , (2) Top 25 PID pairs, (3) Top 50 PID pairs, (4) Top 100 PID pairs, (5)All PID Pairs.

Table 4.10. Top scored 25PIDs and experimental verification of these DDI pairs.

Canonical DDI Pair	Experimental Verification	PCP DDI Pair	Experimental Verification	Calcium DDI Pair	Experimental Verification
PF04857-PF07742	√	PF04857-PF07742	√	PF04857-PF07742	√
PF09753-PF06775	-	PF06292-PF01504	-	PF00876-PF00864	-
PF00876-PF00864	-	PF00876-PF00864	-	PF09238-PF00727	-
PF09238-PF00727	-	PF09238-PF00727	-	PF00146-PF00329	-
PF08612-PF11568	-	PF00147-PF00354	-	PF00068-PF00305	-
PF11107-PF11510	-	PF01290-PF00193	-	PF09766-PF11957	-
PF08318-PF01105	-	PF01290-PF02469	-	PF00147-PF00354	-
PF11315-PF11594	-	PF00326-PF00048	-	PF10018-PF05527	-
PF11594-PF11315	-	PF00930-PF00048	-	PF11315-PF11594	-
PF09738-PF09738	-	PF10018-PF05527	-	PF11594-PF11315	-
PF03467-PF04113	-	PF00146-PF05368	-	PF04722-PF04722	√
PF00574-PF04106	-	PF04699-PF02947	-	PF01997-PF01997	√
PF03807-PF03807	√	PF00146-PF01058	-	PF00542-PF00542	√
PF03301-PF03301	√	PF01733-PF10494	-	PF00459-PF00459	√
PF04733-PF03208	-	PF08612-PF09637	√	PF00243-PF00243	√
PF03045-PF03045	-	PF09637-PF08612	-	PF01287-PF01916	-
PF02223-PF02223	√	PF03467-PF04113	-	PF08612-PF11568	-
PF02127-PF02127	√	PF03807-PF03807	√	PF01596-PF10601	-
PF00810-PF01650	-	PF00542-PF00542	√	PF05450-PF06105	-
PF00883-PF00883	√	PF00459-PF00459	√	PF00383-PF03870	-
PF00542-PF00542	√	PF00210-PF00210	√	PF00059-PF00357	-
PF00266-PF00266	√	PF00899-PF09029	-	PF05450-PF10251	-
PF00243-PF00243	√	PF01033-PF00079	-	PF03153-PF03153	√
PF01733-PF10494	-	PF00110-PF03062	-	PF01704-PF01704	√
PF06292-PF01504	-	PF01596-PF10601	-	PF00383-PF00383	√

All these three Wnt sub-networks resemble to each other when the most frequent domain pairs are taken into account. In Table 4.11, the most frequent domain pairs are shown.

Table 4.11. The most frequent domain pairs in all Wnt sub-networks.

PF00069	Protein kinase domain	PF00069	Protein kinase domain
PF00069	Protein kinase domain	PF07714	Protein tyrosine kinase
PF00069	Protein kinase domain	PF00018	SH3 domain
PF00069	Protein kinase domain	PF00017	SH2 domain
PF00069	Protein kinase domain	PF00595	PDZ domain

The protein kinase domain pair (PF00069-PF00069) is the most frequent one. Protein kinases are known to be enzymes that have a catalytic subunit which transfers phosphate and affects the function of protein because of the conformational change (Hanks, Quinn and Hunter, 1988).Protein kinase function has been evolutionarily conserved from Escherichia coli to human since they play a role in a many cellular processes, including division, proliferation, apoptosis, and differentiation. It is known that the catalytic subunits of protein kinases are highly conserved and there are kinase-specific inhibitors for the treatments of many diseases (Manning *et al.*, 2002; Li *et al.*, 2004).

Other interactions showing high frequencies are the interactions between protein kinase domain (PF00069) and Src family domains which are SH3 (PF00018) and SH2 (PF00017). The SH3 (Srchomology 3) domains are often indicative of a protein involved in signal transduction related to cytoskeletal organisation. The surface of the SH3-domain bears a flat, hydrophobic ligand-binding pocket which consists of three shallow grooves defined by conservative aromatic residues in which the ligand adopts an extended left-handed helical arrangement (Pfam). SH2 domains function as regulatory modules of intracellular signalling cascades by interacting with high affinity to phosphotyrosine-containing target peptides in a sequence-specific fashion which differs from one SH2 domain to another and strictly phosphorylation-dependent manner (Marengere and Pawson, 1994; Pawson and Schlessingert, 1993; Mayer and Baltimore, 1993; Pawson, 1995).

In addition to that the PDZ domain (PF00595) is found to be frequent in reconstructed canonical Wnt-beta catenin and non-canonical Wnt-PCP sub-networks. PDZ domains are found in diverse signalling proteins in many organisms including yeasts, plants and vertebrates. Dishvelled proteins, which are fundamental proteins of Wnt signaling, also have PDZ domains.

Domain-domain interactions can be used in prediction of protein-protein interactions. In the present sub-networks, there are unwired proteins (WNT10B, WNT11, SFRP4, SFRP5, PPP3R2, and DAAM 2) that have no physical interaction. That's why domain interactions of these proteins were examined. Unfortunately, the domains of Wnt proteins (WNT10B, WNT11), SFRP5 and PPP3R2 are not existing in Interpro database. On the other hand, the Formin homology 2 domain (PF02181) of DAAM2 protein has experimentally verified interactions with Actin domain (PF00022). As a result of that, DAAM2 may have interaction with the proteins ACL6A (O96019), ARP3 (P61158), ACTC (P68032) and ACTA1 (P68133) which contain actin domain. In addition to that, the SFRP4 protein has Fz domain (PF01392) and Fz domain has interaction with itself. So it can be stated that SFRP4 protein can interact with proteins that have FZ domain (especially Frizzled proteins) in this network. As a result it can be considered that SFRP4 protein can interact with FZD9 (O00144), MUSK (O15146), FZD7 (O75084), FZD5 (Q13467), FZD2 (Q14332), FZD4 (Q9ULV1) and SFRP1 (Q8N474). These estimations are listed in Table 4.12 below.

Table 4.12. Protein interaction assumptions based on ddi predicted by association method.

Unwired Proteins	**Domain**	**Predicted interaction domain**	**Possible Protein Interaction Pairs**
WNT10B	None	-	-
WNT11	None	-	-
PPP3R2	None	-	-
DAAM2	Formin homology 2 domain (PF02181)	Actin domain (PF00022)	ACL6A , ARP3, ACTC and ACTA1
SFRP4	Fz domain (PF01392)	Fz domain (PF01392)	FZD9, MUSK, FZD7, FZD5, FZD2, FZD4 and SFRP1
SFRP5	None	-	-

4.3. Graph Theoretical Analysis

The topological analyses of the reconstructed networks were performed using Network Analyzer plug-in of Cytoscape (ver. 2.7.0).The mean degree (average number of interactions per protein), clustering coefficients (normalized number of interactions between neighbors of each protein), mean path lengths, network diameters (longest path between any two nodes), power-law distribution exponents (γ), and centrality values were obtained to gain insight on the characteristics of each sub-network. The degree distribution of each sub-network approximates a power law model ($P(k) \cong k^{-\gamma}$) with few nodes having high degree (hub proteins) and the others having low degree. The degree distribution of the Wnt/β-catenin signaling network is shown in Figure 4.3(a), where the exponent is equal to 1.78 ($P(k) \cong k^{-\gamma}, \gamma = 1.78, R^2 = 0.89$). The degree distributions of the non-canonical Wnt/PCP ($P(k) \cong k^{-\gamma}, \gamma = 1.80, R^2 = 0.93$) and Wnt/Calcium ($P(k) \cong k^{-\gamma}, \gamma = 1.68, R^2 = 0.88$) sub-networks are given in Figure 4.3b, and 4.3c, respectively. The degree distributions of the reconstructed sub-networks show consistency with most of biological networks having a small-world topology. The power law exponent,$\gamma$and R^2 values used in the analyses of the degree distributions of different reconstructed sub-networks are given in Table 4.13.

Table 4.13. Power law exponents of different reconstructed sub-networks.

Model	**Power law exponent (γ)**	**R^2**	**Reference**
Wnt/β-catenin (H. Sapiens)	1.78	0.89	Present study
Wnt/PCP (H. Sapiens)	1.80	0.93	Present study
Wnt/Ca^{2+} (H. Sapiens)	1.68	0.88	Present study
Wnt/β-catenin (D.Melanogaster)	1.78	0.93	Toku, 2010
Hedgehog (D. Melanogaster)	1.75	0.93	Toku, 2010
Sphingolipid (H. Sapiens)	1.68	0.90	Özbayraktar, 2011
EGFR (H. Sapiens)	1.86	0.84	Tekir *et al.*, 2009
Insulin (H. Sapiens)	1.53	0.83	Tekir *et al.*, 2010
ca-signaling (H. Sapiens)	1.64	0.90	Tiveci, 2011

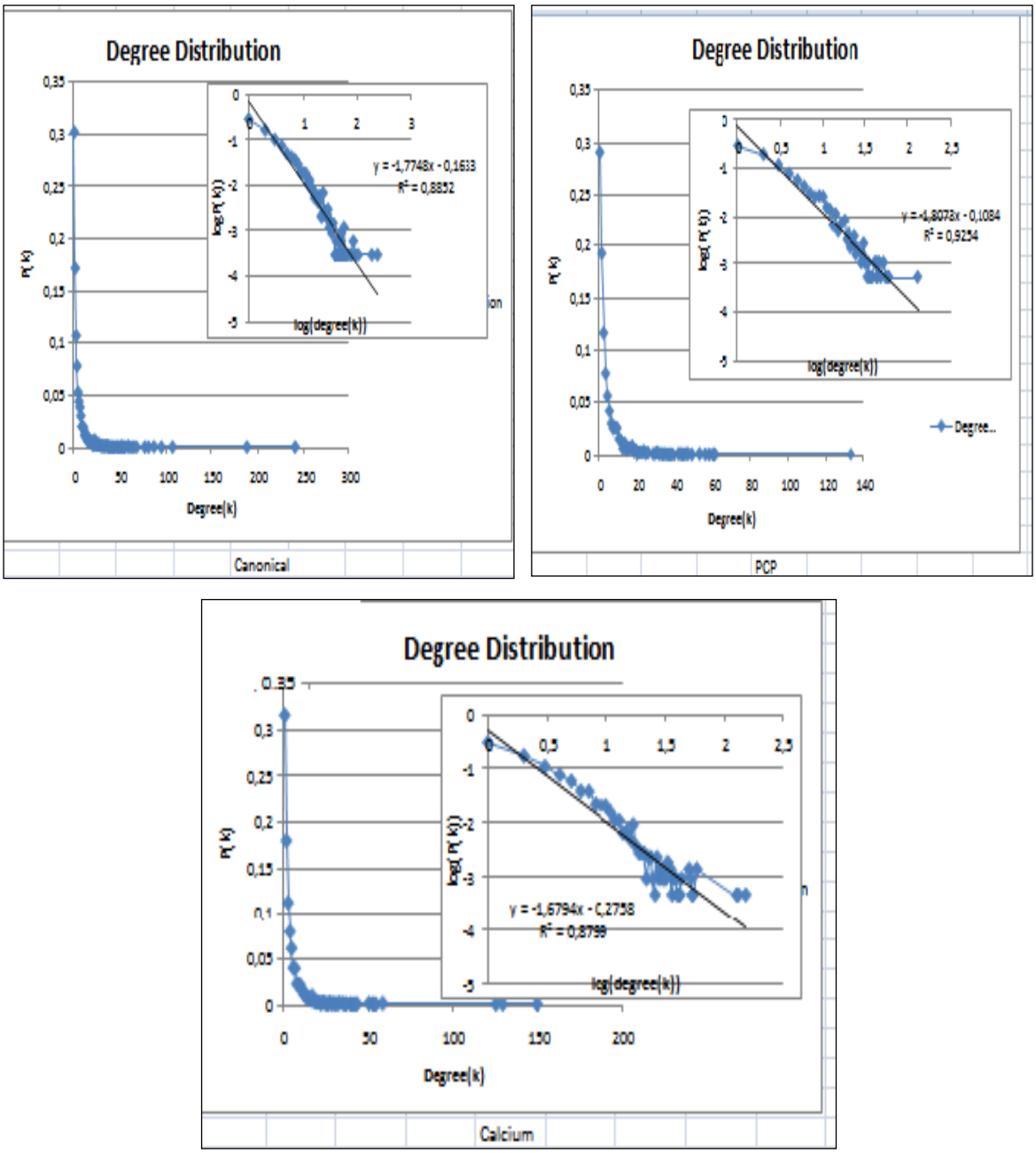

Figure 4.3. Degree distribution of (a)The canonical Wnt network (b)Degree distribution of the PCP Wnt network (c)Degree distribution of the calcium Wnt network.

The graph theoretic properties of the protein interaction sub-networks of Wnt signaling are given in Table 4.14. The topological properties of several networks reported in literature are given in Table 4.15 for comparison.

The network diameter is the maximum node eccentricity. The eccentricity here can be defined as the maximum non-infinite length of a shortest path from a node to another node in

the network. In other words, the network diameter can be defined as the maximum length of shortest paths between two nodes. The network diameter value indicates the speed of signal flow. As the complexity of the organism increases, the time needed for baud rate (information transmission rate) increases.

The diameters are 14, 13, and 15 for Wnt canonical, PCP and calcium signaling networks, respectively. The network diameter of the whole Wnt signaling network in which these three sub-networks are integrated, is found to be 15 (Table 4.14). When the diameters of the reconstructed human Wnt sub-networks are examined, it is seen that they are in agreement with the values reported in literature for fly and worm. Also, the diameter of the reconstructed human sphingolipid protein interaction network (Özbayraktar, 2011), which is 13, shows similarity with these Wnt networks. On the other hand, the diameters of the reconstructed Wnt sub-networks are found to be higher than that of insulin signalling network which has a diameter of 5. This implies that the information transfer rate is slower in Wnt signalling networks (See Table 4.15).

Table 4.14. Graph theoretical properties of PPI Wnt networks.

Model	**Number of Nodes**	**Number of Interactions**	**Mean Path Length**	**Diameter**	**Mean Connectivity**	**Clustering Coefficient**
Wnt/β-catenin	3251	9304	4.46	14	5.72	0.114
Wnt/PCP	1952	5001	4.61	13	5.12	0.112
Wnt/Ca^{+2}	2112	5293	4.56	15	5.01	0.105
Wnt (integrated)	3489	10092	4.40	15	5.78	0.113

The average shortest path length (or the characteristic path length) is the expected distance between two connected nodes. In reconstructed human Wnt sub-networks, the average path length values are found to be around 4.5. This value is found to be 4.40 in whole Wnt signalling. The network diameter and the shortest path length distribution may indicate small-world properties of the analyzed network.As it will be discussed in section

4.5, the shortest path length is also found to be 5 by the linear path analysis. The proteins that participate in these shortest paths are mentioned in detail in section 4.5.

Table 4.15. Graph theoretical properties of the protein interaction networks.

Model	Number of Nodes	Number of Interactions	Mean Path Length	Diameter	Model
Wnt/β-catenin (H. Sapiens)	3251	9304	4.46	14	Present work
Wnt/PCP (H. Sapiens)	1952	5001	4.61	13	Present work
Wnt/Ca^{+2} (H. Sapiens)	2112	5293	4.56	15	Present work
Wnt(whole) (H. Sapiens)	3489	10092	4.40	15	Present work
Wnt/β-catenin (D.melanogaster)	656	1253	4.80	13	Toku *et al.*, 2010
Hedgehog (D.melanogaster)	568	975	4.80	14	Toku *et al.*, 2010
EGFR (H. Sapiens)	329	1795	4.70	11	Tekir *et al.*, 2009
glucosesignaling (S. cerevisiae)	1363	3649	6.81	9	Arga *et al.*, 2007
DIP (C.elegans)	2638	4030	4.80	14	Wu *et al.*, 2005
Sphingolipid (H. Sapiens)	3097	11064	4.10	13	Özbayraktar, 2011
Insulin (H. Sapiens)	498	2887	2.9	5	Tekir *et al.*, 2010
ca-signaling- (H. Sapiens)	2593	9890	3.87	9	Tiveci *et al.*, 2011

Another parameter, clustering coefficient of a node *N* can be defined as the ratio of the number of edges between the neighbors of *N* to the maximum number of edges that could possibly exist between the neighbors of *N*. The network clustering coefficient is the average

of the clustering coefficients for all nodes in the network. This value is always between 0 and 1. The clustering coefficient values found in this research are about 0.1 which means that the numbers of edges of the nodes are nearly 10 percent of the maximum number of the edges that could be exist. When the values of each Wnt sub-network are compared, it is seen that the canonical Wnt signaling network has the highest value (0.114) whereas Wnt/Ca^{+2}signaling network has the lowest value (0.105).

In addition to that, the average (mean) connectivity values are 5.72, 5.12 and 5.01 for β-catenin, PCP and Ca^{+2} pathways, respectively (Table 4.14). For insulin network, this value is obtained as 2.9 which is smaller than the values obtained in Wnt networks. The clustering coefficient values found in this research are in agreement with literature values (Alterovitz *et al.*, 2005).

Another significant topologic aspect is the degree values of the nodes. In this study, the degree distributions of the reconstructed sub-networks are showing a small-world topology which indicates that there are few nodes which have many interaction partners than the other nodes. These highly connected proteins called hubs play an important role in biological processes of the cell (Hsing *et al.*, 2008). In order to analyze the high degree proteins (hubs), the connectivity values (=degree) of all proteins were calculated. The hubs of the Wnt signaling and relevant information are summarized in Tables 4.16 and 4.17.

The hubs of the canonical pathway are obtained as KC1AL (Casein kinase I isoform alpha-like), YWHAZ (Protein kinase C inhibitor protein 1) and TBL1XR1 (F-box-like/WD repeat-containing protein). Casein kinases preferentially use acidic proteins (especially caseins) as substrates. Therefore, they can phosphorylate many proteins. APC tumour suppressor protein has roles in the regulation of Wnt signaling and cytoskeletal dynamics and it is known that casein kinase-1-alpha forms beta-catenin destruction complex when connected to the proteins APC, beta-catenin and glycogen synthase kinase-3-beta (GSK3-beta) (Faux *et al.*, 2008).Besides that, KC1AL has interactions with the core proteins,AXIN1,AXIN2, CSNK1A1, CSNK1D and CSNK1E (String database).

TBL1XR1 is a core protein of canonical Wnt signaling which has an interaction with TBL1X. This protein is involved in signal transduction and cytoskeletal assembly and plays

an essential role in transcription activation mediated by nuclear receptors and has effects on cytotypic differentiation. It is also reported that low level of TBL1XR1 gene expression causes breast cancer (Kadota *et al.*, 2009).

Table 4.16. GO terms of the hub proteins.

Protein	**Biological Process**	**Cellular Component**	**Molecular Function**
KC1AL	Wnt receptor signaling pathway	cytoplasm	ATP binding protein serine/threonine kinase activity
TBL1XR1	Negative regulation of transcription,DNA dependent Canonical Wnt signaling	Cytoplasm nucleus	DNA binding Beta-catenin binding Transcription activator activity
YWHAZ	Signal transduction protein targeting to mitochondiron responsetodrug	Cytoplasm cytosol nucleus	protein domain specific binding transcription factor binding protein binding
PRKCA	intracellular signal transduction cellular calcium ion homeostasis	Cytoplasm Cytosol nucleus	metal ion binding protein binding calciumdependent protein kinase C activity transferaseactivity
PRKCB	signal transduction	cytosol nucleus	protein kinase C activity zincionbinding ATP binding

YWHAZ (14-3-3 protein zeta/delta /Protein kinase C inhibitor protein 1) is a key component in both canonical and non-canonical Wnt signaling and has an interaction with the core protein CSNK1A1 (CSKI alpha).14.3.3 proteins are a group of highly conserved proteins that are involved in many vital cellular processes such as metabolism, protein trafficking, signal transduction, apoptosis and cell cycle regulation. Because many interactions of 14.3.3 proteins are phosphorylation dependent, 14.3.3 proteins have been integrated into the core regulatory pathways that are crucial for normal growth and development such as Wnt signaling pathway. It is an adapter protein implicated in the

regulation of a large spectrum of both general and specialized signaling pathways. In addition to its interaction with canonical pathway core protein of CSNK1A1, YWHAZ also has interactions with core proteins of NFATC2, NFATC4 and MAPK8 of non-canonical Wnt signaling. Its binding to a large number of partners via recognition of a phosphoserine or phosphothreonine motif results in the modulation of the activity of the binding partner (Olsen *et al.*, 2006). YWHAZ protein is the common hub and also bottleneck protein in all reconstructed Wnt signaling sub-networks.

Table 4.17. General information about hub proteins.

Hub proteins	**Conservation**	**Degree**	**Core / member**	**Related Disease**	**Related Pathways**
KC1AL (Wnt/β-catenin)	87.93%	241	Core	Alzheimer Disease(Li,Yin and Kuret,2004)	Axin regulated JNK activity(Zhang *et al.*, 2002) Hedgehog signalingpathway(kegg)
YWHAZ (Wnt/β-catenin)	66.54%	189	Member	Breast cancer Obesity Diabetes	Cell cycle Oocyte meiosis N urotrophinsig aling pathway Pathogenic Escherichia coli Infection(Kegg)
TBL1XR1 (Wnt/β-catenin)	54.09%	107	Core	Breast cancer Acute lymphoblastic leukemia(Parker*et al.*,2008)	Wnt, Notch, NF-κB, and nuclear receptor pathways (Kadota *et al.*, 2009)
YWHAZ (Wnt/PCP)	66.54%	133	Member	Breast cancer	Cell cycle Oocyte meiosis Neurotrophinsignaling pathway Pathogenic Escherichia coli Infection(Kegg)
PRKCB (Wnt/Ca^{2+})	53.55%	149	Core	Gastric cancer (Schwartz *et al.*,1993)	NF-kappa-B signalling MAPK signaling pathway ErbBsignaling pathway Calcium signaling pathway
PRKCA (Wnt/Ca^{2+})	44.09%	129	Core	MS(Multiple sclerosis) (Saarela *et al.*,2006) colon cancer(Chen *et al.*,1999)	Insulin receptor signaling pathway (GO Annotation) Nerve growth factor receptor signaling
YWHAZ (Wnt/Ca^{2+})	66.54%	125	Member	Breast cancer	Cell cycle Oocyte meiosis Neurotrophinsignaling pathway Pathogenic Escherichia coli infection (Kegg)

Analyzing the interaction network of proteins in the cell is important in understanding how complex processes lead to diseases. It is therefore needed to identify key mediator proteins, which possibly maintain direct relationships among proteins causing diseases, by studying the relationships between topology and functionality of the protein-protein interaction networks. YWHAZ is also found to be a key mediator protein in various diseases involving various types of cancers, heart diseases, obesity, diabetes and autism (Nguyen and Jordán, 2010). YWHAZ also contributes to chemotherapy resistance and recurrence of breast cancer (Ralhan *et al.*, 2008). Moreover, TP53 **(**Cellular tumor antigen p53), which acts as a tumor suppressor in many tumor types and induces growth arrest or apoptosis depending on the physiological circumstances and cell type, has an interaction with YWHAZ; this interaction enhances TP53 transcriptional activity (http://www.pathwaycommons.org/pc/record2.do?id=92132) .

The hubs of the Wnt/Ca^{2+} pathway are PRKCB (Protein kinase C beta type), PRKCA (Protein kinase C alpha type) and also YWHAZ (Protein kinase C inhibitor protein 1). Protein kinase C (PKC) is a family of serine- and threonine-specific protein kinases that can be activated by calcium and second messenger diacylglycerol. PKC family members phosphorylate a wide variety of protein targets and are known to be involved in diverse cellular signaling pathways. PKC phosphorylation and subsequent activation of multiple target proteins are involved in signal transduction pathways such as receptor desensitization, modulation of membrane structure events, regulation of transcription, regulation of cell growth, immune responses and also in learning and memory (http://www.cellapplications.com/product_desc.php?id=1020). PRKCA kinase was reported to play roles in many different cellular processes, such as cell adhesion, cell transformation, cell cycle checkpoint, and cell volume control (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?db=human&c=Gene&l=PRKCA). Knockout studies in mice suggest that this kinase may be a fundamental regulator of cardiac contractility and Ca^{2+}handling in myocytes (http://www.ncbi.nlm.nih.gov). PRKCA also binds to RHOA which is another core protein in Wnt/PCP signaling. PRKCB, calcium-activated and phospholipid-dependent serine/threonine-protein kinase, is involved in various processes such as regulation of the B-cell receptor (BCR) signalosome, apoptosis and transcription regulation and it has an interation with the core protein, dishevelled 2 (DVL2)

and the common hub protein YWHAZ. The obtained hub proteins except YWHAZ are core proteins of the Wnt signaling network (Table 4.17).

Recently, in addition to degree, the proteins are also classified by using another topological aspect which is betweenness centrality. Betweenness centrality value gives the total number of non-redundant shortest paths going through a certain node or edge. In other words, it can be explained by the measure of the centrality of the nodes in a network and most of the shortest paths in a network go through the nodes with high betweenness centrality and these nodes are called as the bottleneck proteins (Yu *et al.*, 2007; Freeman, 1977; Girvan and Newman, 2002). In these reconstructed sub-networks, the proteins with highest degrees also have the highest betweenness centrality values which imply that the hub proteins are also the bottleneck proteins of the networks. Another node centrality measure in networks is closeness centrality which is defined as the inverse of the sum of distances to all other nodes. In other words, it is the inverse of the sum of all shortest paths to others or the smallest number of ties to go through to reach all others individually (Freeman, 1978; Opsahl *et al.*, 2010; Wasserman and Faust, 1994). The betweenness and closeness centrality values obtained for the bottleneck proteins are listed in Table 4.18. The topological properties of the top 350 nodes are given in Appendix D.

Table 4.18. Topological properties of bottleneck proteins in human Wnt signaling.

Model	**Uniprot ID_Name**	**Betweenness Centrality**	**Closeness Centrality**	**Clustering Coefficient**	**Degree**	**Average Shortest Path Length**
Wnt/Canonical	Q8N752 (KC1AL)	0.168	0.356	0.0060	241	2.817
	P63104 (YWHAZ)	0.124	0.350	0.0071	189	2.855
	Q9BZK7 (TBL1XR1)	0.052	0.289	0.0071	107	3.464
Wnt/PCP	P63104 (YWHAZ)	0.182	0.351	0.0094	133	2.850
Wnt/Ca^{2+}	P17252 (PRKCA)	0.160	0.353	0.0122	129	2.830
	P63104 (YWHAZ)	0.136	0.343	0.0099	125	2.917
	P05771 (PRKCB)	0.135	0.334	0.0074	149	2.997

In identifying the drug target, some topological aspects such as betweenness centrality, closeness centrality and degree are important. A protein can be a suitable drug target if its betweenness centrality and degree values are lower whereas closeness centrality is higher than the average. A comparison of our network proteins with commercial drug targets under clinical trial will be given in section 4.6.

4.4. Module Detection and Analysis

Modules in PPI networks are groups of proteins performing a certain biological function together. The statistically overrepresented molecular function terms were obtained by GO enrichment analysis. The canonical pathway was clustered into 75 complexes. The PCP and Ca^{+2} networks were clustered into 41 and 57 complexes, respectively. Table 4.19 shows the significant modules detected in the reconstructed sub-networks, i.e., the modules either having score higher than 2 or having number of nodes greater than 10.

Many of the proteins in the modules have roles in binding, catalytic activity and transcriptional regulation. The top two scored modules in canonical Wnt signaling is related to eukaryotic cell surface binding (1), threonine endopeptidase activity (2), respectively, and the third one is related to DNA-directed RNA polymerase activity. The proteins in the first module of PCP signaling have role in hydrogen ion transporting ATPase activity and those in these module have role in RNA polymerase activity. The first module of Wnt-Calcium signaling is related to dependent nuclear receptor transcription coactivator activity whereas the second module is related to DNA-directed RNA polymerase activity.

In addition to that, the modules with significant molecular functions directly related to Wnt signaling were detected by GO enrichment analysis (Table 4.20). The proteins in module 22 of Wnt/β-catenin (canonical) pathway are enriched in Wnt protein binding. In module 26, NADH dehydrogenase (ubiquinone) activity is dominant. NADH Dehydrogenase is the first enzyme (Complex I) of the mitochondrial electron transport chain. Complex I may have a role in triggering apoptosis. In fact, it has been shown that there is a correlation between mitochondrial activities and programmed cell death (PCD) during somatic embryo development (Chomova and Racay, 2010; Petrussa *et al.*, 2009).

Table 4.19. The significant modules in Wnt networks.

Network	Cluster	Score	Nodes	Edges	Network	Cluster	Score	Nodes	Edges
Canonical	1	11.5	24	277	PCP	1	7	15	106
	2	8.9	19	179		2	4.5	10	49
	3	4.5	10	55		3	4	9	45
	4	4.5	10	45		4	4	9	36
	5	4.0	25	120		5	3	7	28
	6	3.2	13	50		6	3	7	26
	7	3.0	13	45		7	2	27	73
	8	2.5	6	20		8	2	5	15
	9	2.2	30	81		9	2	5	11
	10	2.0	5	15		10	2	39	100
	11	2.0	14	37		32	1	10	13
	14	1.6	17	37	Ca++	1	4.5	10	49
	17	1.5	46	78		2	4.5	10	45
	18	1.4	10	15		3	4	9	45
	21	1.3	12	23		4	4	9	36
	23	1.2	22	33		5	3	7	28
	25	1.2	13	16		6	2	22	58
	26	1.1	8	15		7	2	5	15
	27	1.1	10	18		8	1.8	10	21
	32	1.0	16	29		9	2	48	119

The corepressor CtBP (C-terminal Binding Protein) has roles in transcriptional pathways important for development, cell cycle regulation and transformation, and embryonic development is important in Wnt signaling. Zhang *et al*. (2002) showed that NAD+/NADH regulate CtBP binding.

When the modules in Planar Cell Polarity (PCP) sub-network are investigated, the proteins of module 18 show potassium channel activity which is important in polarity. In module 15, Rho GDP-dissociation inhibitor activity is significant. Rho GTPase family has a role in breast cancer which is a Wnt signaling related disease. RhoBTB2 is identified independently as a gene deleted in breast tumors and shows growth inhibitory activity when reintroduced into Deleted Breast Cancer2-deficient cells (Hamaguchi *et al.*, 2002; Karnoub *et al.*, 2004).

The proteins of module 6 in Wnt/Ca^{2+}subnetwork are enriched in calcium ion binding as expected.

Table 4.20.Most significant molecular functions of the modules in Wnt network.

Network	**Module**	**GO-ID**	**corr p-value**	**Description**	**Genes in test set**
Canonical	22	17147	2.00E-02	Wnt-protein binding	P34925 (RYK)
	26	8137	4.99E-07	NADH dehydrogenase (ubiquinone) activity	P03886(NU1M), O75251(NDUS7), Q16795(NDUA9)
PCP	15	5094	1.92E-04	Rho GDP-dissociation inhibitor activity	P52565(GDIR1), P52566(GDIR2)
	18	5267	5.35E-03	potassium channel activity	Q693B1(KCD11), Q719H9(KCTD1)
Ca++	6	5509	1.54E-05	calcium ion binding	P62166(NCS1), Q9NZI2(KCIP1), P43080(GUC1A),

The information obtained by module analysis enabled us to confirm the present Wnt signaling network reconstructed using an integrated approach of interactomics and GO annotations, and thus it can be used in further analysis for drug target identification.

4.5. Network Decomposition Analysis

The linear paths in the reconstructed whole Wnt signaling network and those in each sub-network of Wnt signaling were determined via NetSearch algorithm of Steffen *et al.* (2002) in order to examine the signal transmittal steps. In this algorithm, the membrane (ligand) proteins were used as input whereas the transcription factors were used as output (Table 4.21) of Wnt signaling network in *Homo Sapiens.*

Table 4.21. Input and output proteins of the linear paths.

Input Protein (Uniprot_ID)	**Protein Name**	**Output Protein (Uniprot_ID)**	**Protein Name**
P04628	WNT1	O94916	NFAT5
P09544	WNT2	O95644	NFATC1
Q93097	WNT2B	Q13469	NFATC2
P56703	WNT3	Q12968	NFATC3
P56704	WNT3A	Q14934	NFATC4
P56705	WNT4	Q9UJU2	LEF1
P41221	WNT5A	P36402	TCF7
Q9H1J7	WNT5B	Q9HCS4	TF7L1
O00755	WNT7A		
O00744	WN10B		
O96014	WNT11		

The shortest path length is found to be 5, which includes 5 proteins connected by 4 linear interactions from WNT3A to LEF1. There are 2 linear paths at path length of 5. The proteins that participate in linear paths of length 5 and 6 are shown in Tables 4.22 and 4.23.

WNT3A and LEF1 are the core proteins of canonical Wnt/beta-catenin signaling network. WNT3A is a ligand for members of the frizzled family of seven transmembrane receptors and plays distinct roles in cell-cell signaling during morphogenesis of the developing neural tube. It has an interaction with PORCN protein. In addition to that it is a component of Wnt-Fzd-LRP5-LRP6 signaling complex and has direct interactions with FZD1 and LRP6 (Bourhis *et al.*, 2010).

LEF1 activates transcription of target genes in the presence of CTNNB1 and EP300 and it may play a role in hair cell differentiation and follicle morphogenesis. Besides that, it regulates T-cell receptor alpha enhancer function and binds to DNA in a sequence-specific manner (Jesse *et al.*, 2010). It has a strong role in detection of cancer. It is not detected in normal colon, but highly expressed in colon cancer biopsies and colon cancer cell lines, also expressed in several pancreatic tumors and weakly expressed in normal pancreatic tissue (www.uniprot.org).

Table 4.22. The linear paths at path length of 5.

Path Length	**Input Protein**				**Output Protein**
5	*P56704* (Wnt 3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q13485* (SMAD4)	*Q9UJU2* (LEF1)
	P56704 (Wnt 3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q15796* (SMAD2)	*Q9UJU2* (LEF1)

Table 4.23. The linear paths at path length of 6.

Path Length	**Input Protein**					**Output Protein**
6	*P04628* (WNT1)	*P56704* (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q13485* (SMAD4)	*Q9UJU2* (LEF1)
	P04628 (WNT1)	*P56704* (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q15796* (SMAD2)	*Q9UJU2* (LEF1)
	P04628 (WNT1)	*Q9H461* (FZD8)	*Q9HD26* (GOPC)	*Q13255* (GRM1)	*P52294* (IMA1)	*Q9UJU2* (LEF1)
	P09544 (WNT2)	*P56704* (WNT3A)	*Q07954* (LRP1)	P12757 (SKIL)	Q13485 (SMAD4)	*Q9UJU2* (LEF1)
	P09544 (WNT2)	*P56704* (WNT3A)	*Q07954* (LRP1)	P12757 (SKIL)	Q15796 (SMAD2)	*Q9UJU2* (LEF1)
	P56704 (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q13485* (SMAD4)	*P84022* (SMAD3)	*Q9UJU2* (LEF1)
	P56704 (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q13485* (SMAD4)	*Q15796* (SMAD2)	*Q9UJU2* (LEF1)
	P56704 (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q13485* (SMAD4)	*Q15797* (SMAD1)	*Q9UJU2* (LEF1)
	P56704 (WNT3A)	*Q07954* (LRP1)	*P12757* (SKIL)	*Q15796* (SMAD2)	*Q13485* (SMAD4)	*Q9UJU2* (LEF1)

By graph theoretical analysis (section 4.3), the diameters were found as 14, 13, 15 and 15 in for Wnt/beta-catenin, Wnt/PCP, Wnt/Calcium sub-networks and whole Wnt network, respectively. The path length is therefore increased in order to cover all the proteins in the network. However, a maximum number of 13 steps can be achieved due to the computer capacity. The numbers of linear paths with the number of steps and covered protein percentages are shown in Table 4.24. The number of linear paths obtained at 13 steps is 1 086 956, in which only 59 (50%) of 118 core proteins and 1244 (34%) of 3676 proteins are covered.

Table 4.24. Results of the network decomposition analysis of the Wnt signaling network.

Number of steps	Number of linear paths	Number of core proteins exist in the paths	Number of nodes exist in the paths
5	2	3 (3%)	6 (0.16%)
6	9	6 (6%)	14 (0.3%)
7	24	8 (7%)	29 (0.8%)
8	72	16 (14%)	70 (2%)
9	376	27 (23%)	205 (6%)
10	2182	33 (28%)	449 (12%)
11	16298	41 (35%)	762 (21%)
12	128207	52 (44%)	1044 (29%)
13	1086956	59 (50%)	1244 (34%)

4.5.1. Network decomposition analysis for canonical and non-canonical wnt pathways

Network decomposition analysis was then performed for canonical and non-canonical wnt pathways separately. A maximum number of 13 steps can be achieved for canonical pathway and 15 steps for non-canonical pathway. For canonical pathway the number of linear paths obtained at 13 steps is 815627, in which only 33 (42 %) of 68 core proteins and 1115 (32%) of 3251 proteins are covered. For non-canonical pathway the number of linear paths obtained at 15 steps is 1098373, in which only 27 (48 %) of 60 core proteins and 817

(34%) of 2547 proteins are covered (Table 4.25). For a consistent comparison the linear path analysis is also performed at 13 steps for canonical and non-canonical Wnt pathways. Unfortunately, for non-canonical pathway the number of linear paths obtained at 13 steps is found to be 29082, in which 546 of 2547 nodes and only one core protein of 60 core proteins is covered. Hence, the number of steps is increased for the purpose of high coverage of the nodes. Also, this result seems to be logical since the diameter of the non-canonical pathway is found to be larger than that of canonical pathway, which implies that the signal transfer is faster in canonical pathway.

A minimum number of 7 steps is obtained in non-canonical pathway whereas this is 5 in canonical pathway. In canonical pathways, the information flow prefers short routes, i.e. it goes through 5 proteins at fastest whereas it has to pass 7 proteins in case of non-canonical pathways such as PCP or Wnt/Ca^{2+} signaling.

Table 4.25 Results of network decomposition analysis of both canonical and non-canonical Wnt networks.

Model	**Maximum Number of steps**	**Number of linear paths**	**Number of core proteins exist in the paths**	**Number of nodes exist in the paths**
Canonical Wnt Pathway	13	815627	33 (42 %)	1115 (32%)
Non-canonical Wnt Pathways	13	29082	1 (2%)	546 (21%)
Non-canonical Wnt Pathways	15	1098373	27(48 %)	817 (34%)

4.5.2. Participation of Proteins in Linear Paths

The participation values of the core proteins and sub-network proteins (nodes) are low at the maximum of the shortest path lengths (i.e. network diameter) obtained. However, this result is expected because the directions of the interactions are taken into account in the linear path analysis whereas it is not in graph theoretic analysis.

Table 4.26. Proteins with the highest participation percentages in Wnt signaling pathway.

Uniprot ID	**Participation in linear paths(%)**	**Protein Name**	**Recommended Name**	**Canonical/ Noncanonical**	**Degree**
Q9UJU2	56.19	LEF1	T cell-specific transcription factor 1-alpha	Canonical	17
O00755	51.91	WNT7A	Protein Wnt-7A	Canonical	2
O00144	51.91	FZD9	Frizzled-9	Canonical/PCP/ Ca^{2+}	4
Q99750	50.94	MDFI	MyoDfamilyinhibitor	Canonical/PCP/ Ca^{2+}	50
P04628	47.20	WNT1	Proto-oncogene Wnt-1	Canonical/Ca^{2+}	10
Q9HD26	46.89	GOPC	Golgi-associated PDZ andcoiled-coil motif-containing protein	Canonical/PCP	18
Q9H461	46.87	FZD8	Frizzled-8	Canonical/PCP/ Ca^{2+}	4
P33992	42.94	MCM5	DNA replicationlicensingfactor MCM5	Canonical/Ca^{2+}	6
Q14566	38.83	MCM6	DNA replicationlicensingfactor MCM6	Canonical/Ca^{2+}	28
Q15797	29.23	SMAD1	SMAD familymember 1	Canonical/PCP/ Ca^{2+}	60
P28070	28.75	PSB4	Proteasomesubunit beta type-4	Canonical	19

The percentages of each protein contributing to all 1 086 956 linear paths in the whole Wnt network were found. The proteins having participation percentages higher than 20 are listed in Table 4.26. T cell-specific transcription factor 1-alpha (LEF1) has the highest percentage since it is one of the output proteins. Wnt7A and Wnt1 are the input proteins. These three proteins (Wnt7A, Wnt1 and Lef1) out of listed 11 proteins are also the core proteins of the canonical Wnt signaling sub-network. The other highly participating proteins are fzd9, mdfi, gopc, fzd8, mcm5, mcm6, smad1 and psb4 (Table 4.26). These proteins have

high percentages in the linear paths since they bind to essential proteins, especially to the input/output proteins or core proteins which are common to many paths in the network.

Frizzled 9 (fzd9), which is first defined as Fzd3 in literature, has a high participation percentage of 52, and it is common to all three sub-networks of Wnt signaling. It is also a receptor for Wnt proteins and leads to the activation of disheveled proteins, inhibition of GSK-3 kinase, nuclear accumulation of beta-catenin and activation of Wnt target genes. It is also hypothesized that fzd9 may be involved in transduction and intercellular transmission of polarity information during tissue morphogenesis and/or in differentiated tissues (www.uniprot.org).

Another protein common to all three Wnt sub-networks is MyoD family inhibitor protein (MDFI) which inhibits the transactivation activity of the Myod family of myogenic factors and represses myogenesis. It also regulates the transcriptional activity of TCF7L1/TCF3 by direct interaction to it and prevents TCF7L1/TCF3 from binding to DNA. It affects the axin regulation and therefore, it has an important role in Wnt signaling pathway since its binding to axin complex results in an increase in the level of free beta-catenin. The DNA replication licensing factor proteins (MCM5 and MCM6) have interaction with each other and MCM5 also binds to MDFI and beta-catenin, which is an essential protein for Wnt signaling pathway. Besides that, SMAD1-OAZ1-PSMB4 complex mediates the degradation of the CREBBP/EP300 repressor SNIP1. None of these proteins are bottleneck proteins, i.e., may have low effect in signaling.

The following 12 proteins (Nup62, Ga45a, Nu160, Lrsm1, Terf2, Mltk, Mark1, Chm4c, Stag1, Alk, Rbp2 and Myh9) participate in only one linear path within the Wnt network which means that the absence of these proteins does not affect the signaling mechanism thoroughly. The information related to these proteins is given in Table 4.27. When the proteins in Table 4.27 that have lowest participation percentages in linear paths are evaluated according to the criteria of low betweenness and high closeness centrality values, four proteins seem to be important and need further examination. These proteins are LRSAM1, MLTK, MARK1 and miyosin 9.

Table 4.27. Proteins with the lowest percentages in Wnt signaling pathway.

Uniprot ID	**Participation in linear paths (%)**	**Protein Name**	**Recommended Name**	**Canonical/ noncanonical**	**Degree**
P37198	8.1E-05	*NUP62*	Nuclearporeglycoprotein p62	Canonical / Ca^{2+} /PCP	10
P24522	8.1E-05	*GA45A*	Growtharrestand DNA damage-inducible protein GADD45 alpha	Canonical / Ca^{2+} /PCP	3
Q12769	8.1E-05	*NU160*	Nuclearporecomplex protein Nup160	Canonical	6
Q6UWE0	8.1E-05	*LRSAM1*	E3 ubiquitin-protein ligase LRSAM1	Canonical	5
Q15554	8.1E-05	*TERF2*	Telomericrepeat-bindingfactor 2	Canonical / Ca^{2+}	3
Q9NYL2	8.1E-05	*MLTK*	Mitogen-activated protein kinasekinasekinase MLT (ZAK)	Canonical / Ca^{2+} /PCP	3
Q9P0L2	8.1E-05	*MARK1*	Serine/threonine-protein kinase MARK1	Canonical / Ca^{2+} /PCP	2
Q96CF2	8.1E-05	*CHM4C*	Chargedmultivesicular body protein 4c	Canonical	6
Q8WVM7	8.1E-05	*STAG1*	Cohesinsubunit SA-1	Canonical	5
Q9UM73	8.1E-05	*ALK*	ALK tyrosinekinasereceptor	Canonical / Ca^{2+} /PCP	2
P49792	8.1E-05	*RBP2*	E3 SUMO-protein ligase RanBP2	Canonical / Ca^{2+} /PCP	9
P35579	8.1E-05	*MYH9*	miyosin9	Canonical /Ca^{2+}/PCP	10

LRSAM1 (leucine rich repeat and sterile alpha motif containing1), also called as RIFLE and TAL (TSG101-associated ligase), is an E3 type ubiquitin ligase. TSG101 itself is a tumor suppressor gene, which has a role in maturation of human immunodeficiency virus, and LRSAM1 is implicated in its metabolism directly by polyubiquitination (Guernsey *et al.*, 2010). Tsg101 is a gene whose functional disruption led both to cellular transformation and to tumors that metastasized spontaneously in nude mice (Li and Cohen,

1996). In addition to that, although genomic alterations in tsg101 are rare in human cancer, functional inactivation of the gene enhances metastatic growth of murine fibroblasts (Li and Cohen 1996).

Another protein is ZAK (Q9NYL2) which inhibits human lung cancer cell growth via ERK and JNK activation in an AP-1-dependent manner (Yang *et al.*, 2010). Also, over expression of ZAK result in apoptosis (OMIM).

Another protein is Serine/threonine-protein kinase MARK1. Cellular studies showed that over expression of MARK1 resulted in shorter dendrite length and decreased transport speed. MARK1 over-expression in individuals with autism may underlie subtle changes in synaptic plasticity linked to dendritic trafficking (Maussion *et al.*, 2008; OMIM).

The last protein is miyosin9. Fechtner syndrome, which is an autosomal dominant disorder characterized by the triad of thrombocytopenia, giant platelets, and Dohle body-like inclusions in peripheral blood leukocytes, with the additional features of nephritis, hearing loss, and eye abnormalities, mostly cataracts, is caused by heterozygous mutation in the gene encoding nonmuscle myosin heavy chain-9 (MYH9; 160775) on chromosome 22q11 (Peterson *et al.*, 1985; OMIM).

The linear paths, that these 4 proteins participate in, are shown in Table 4.28. As it can be seen from Table 4.28, Zak and MARK1 both bind to SFN which has interaction with YWHAZ. YWHAZ is found to be hub and bottleneck protein in these reconstructed sub-networks due to its high degree and betweenness centrality value, respectively. YWHAZ has a participation percentage of 0.95 in linear path analysis. YWHAZ is also found to be a key mediator protein in various diseases involving various types of cancers, heart diseases, obesity, diabetes and autism (Nguyen and Jordán, 2010) as mentioned in section 4.3. Key mediators are proteins that bind to significant proteins (mostly hubs) and so they can be chosen as the drug targets.

Table 4.28. Linear paths of lowest participant proteins.

	Protein _ID	**Name**	**Protein _ID**	**Name**	**Protein _ID**	**Name**	**Protein _ID**	**Name**
	Q6UWE0	**LRSAM1**	*Q9P0L2*	**MARK1**	*P35579*	**MHY9**	*Q9NYL2*	**ZAK**
Input Protein	O00755	WNT7A	P04628	WNT1	P04628	WNT1	P04628	WNT1
	O00144	FZD9	Q9H461	FZD8	Q9H461	FZD8	Q9H461	FZD8
	Q99750	MDFI	Q9HD26	GOPC	Q9HD26	GOPC	Q9HD26	GOPC
	Q12906	ILF3	P13569	CFTR	P13569	CFTR	P13569	CFTR
	Q8N752	KC1AL	P08670	VIME	P08670	VIME	P08670	VIME
	Q9UQM7	CAMK2A	O43353	RIPK2	Q12873	CHD3	O43353	RIPK2
	Q13554	CAMK2B	P05771	PRKCB	Q14974	IMB1	P05771	PRKCB
	P48443	RXRG	Q9P0L2	**MARK1**	Q00722	PLCB2	Q9NYL2	**ZAK**
	Q6UWE0	**LRSAM1**	P31947	SFN	Q96QT4	TRPM7	P31947	SFN
	Q99816	TS101	P63104	**YWHAZ**	P35579	**MHY9**	P63104	**YWHAZ**
	Q13464	ROCK1	P30291	WEE1	P19838	NFKB1	P30291	WEE1
	Q15796	**SMAD2**	P84022	**SMAD3**	P17252	**PRKCA**	P84022	**SMAD3**
Output Protein	Q9UJU2	LEF1	Q9UJU2	LEF1	O95644	NFAC1	Q9UJU2	LEF1
Path Length	13		13		13		13	

Besides that, another hub protein TBL1XR1 also has a low participation percentage (0.16). The betweenness centrality value of this protein is found to be 0.046, which is higher than the average of all nodes (=0.01). The closeness centrality value of this protein is obtained as 0.288, which is higher than the average; whereas the clustering coefficient value is obtained as 0.007, which is lower than the average value.

4.5.3. Specific Proteins in Linear Paths

The proteins in the linear paths ending at transcription factors specific to canonical and noncanonical pathways were examined in detail. The linear paths were found to reach to LEF1 (Q9UJU2) in canonical subnetwork and NFATC1 (O95644), NFATC2 (Q13469), NFATC3 (Q12968) in noncanonical subnetwork. The proteins, which participate in the linear paths leading to one transcription factor only, are called specific proteins of that particular pathway.

287 specific proteins were obtained which were then investigated according to their topological properties such as lower betweennness centrality, higher clşoseness centrality and higher clustering coefficient than the average for drug target identification. As a result, 51 proteins meet these criteria, and out of these 51 proteins 4 proteins seem to be important since they are either related to important diseases or connected to significant proteins in the network. These proteins are Myc proto-oncogene protein (MYC), TGF-beta receptor type-2 (TGFR2), Cyclin-dependent kinase inhibitor 3 (CDKN3) and F-box-like/WD repeat-containing protein TBL1X (canonical).

MYC is a protein that participates in the regulation of gene transcription. The mutations and overexpressions seen in MYC resulted in cell proliferation and consequently formation of cancer. The translocations such as t (8:14) are the reasons of the development of Burkitt's lymphoma. Soucek *et al.,* 2008 demostrated that the temporary inhibition of Myc selectively killed mouse lung cancer cells, making it a potential drug target in cancer (Gearhart *et al.*, 2007; Soucek *et al*., 2008).

TGFR2 is receptor protein of TGF-beta and also known to be involved in tumor suppression..It forms receptor complexes with serine/threonine protein kinases and has role in activation of SMAD transcriptional regulators. The mutations and defects seen in this protein are associated with Lynch sendrome, Loeys-Deitz aortic aneurysm syndrome, Osler-Weber-Rendu syndrome, hereditary non-polyposis colorectal cancer type 6 (HNPCC6) and esophageal cancer (Tanaka *et al*. , 2000; Lu *et al*., 1998).

TBL1X is a protein that plays an essential role in transcription activation mediated by nuclear receptors. Besides, it is a component of E3 ubiquitin ligase complex containing UBE2D1, SIAH1, CACYBP/SIP, SKP1, APC and TBL1X proteins. It has interactions with essential proteins of Wnt signaling such as APC and beta-catenin and it is also a core protein of reconstructed canonical Wnt signaling pathway (Matsuzawa and Reed, 2001).

CDKN3 is a member of cyclin-dependent kinases (CDKs) which have roles in regulating the cell cycle, regulating transcription, mRNA processing, and the differentiation of nerve cells (Gyuris *et al*., 1993). The overexpression and defects seen in

this protein leads to prostate cancer and hepatocellular carcinoma (HCC) (Yeh *et al*., 2003; Lee *et al*., 2000).

These specific proteins except TBL1X are related to cancer and they are suitable for drug target applications according to their topological properties. Hence, they need further investigation.

4.5.4. Crosstalk of Proteins in Wnt Sub-networks

Signaling networks are communicating systems and they interact with each other rather than behaving in isolation. Crosstalk analysis is branched in two ways that are network crosstalk and path crosstalk. The network crosstalk of a node can be defined as the difference in degree of the node in all considered networks and the maximum degree of this node in any individual pathway. If a node has a high network crosstalk value then it means that this component is a branch node connecting two or more pathways. For analyzing the components that connect the pathways in signal transduction, the network crosstalk analysis was performed and 239 proteins are found to be common among Wnt sub-networks which are β-catenin sub-network (3251 nodes), PCP sub-network (1952 nodes) and Calcium sub-network (2112 nodes). The list of these common proteins is given in Appendix E. In addition to that, the known signaling proteins and hubs of the reconstructed Wnt sub-networks are listed in Table 4.29.

One of the highest crosstalk value belongs to YWHAZ protein. This situation is rational since this protein was obtained as the hub and bottleneck protein of all Wnt sub-networks. Besides, DVL2 has a significant crosstalk value. Dishevelled proteins have high participation in the subnetworks since they interact with the core proteins such as frizzled receptors and GSK3B in Wnt/beta-catenin sub-network and with frizzled receptors and DAAM1 in Wnt/PCP sub-network. Smad proteins also have considerable crosstalk value since they have interactions with axin, beta-catenin and Lef1 proteins. PRKCA, which was found as hub and core protein in Wnt/calcium sub-network, also has a non-zero crosstalk value.

Table 4.29. Proteins and network crosstalk values.

Proteins		**Network crosstalk values**
YWHAZ	Core protein (all sub-networks)	11
DVL2	Core protein (beta-catenin and Wnt/ PCP sub-networks)	11
CAMK2A	Core protein (Wnt/Ca^{2+} sub-network)	4
SMAD3-4	Core proteins (beta-catenin subnetwork)	4
GSK3B	Core protein (beta-catenin sub-network)	2
PRKCA	Hub-Core protein (Wnt/ Ca^{2+} subnetwork)	2
RAC1	Core protein (Wnt/PCP sub-network)	2
NFATC2	Core protein (Wnt/Ca^{2+} subnetwork)	1
AXIN1	Core protein (beta-catenin sub-network)	1

The understanding of the crosstalk mechanism is difficult due to the fact that the reconstructed networks are undirected. On the other hand, detecting these connector proteins by network crosstalk analysis is a promoter step for further experimental work in cancer treatment researches.

4.6. Potential Drug Targets in the Reconstructed Wnt Signaling Networks

Wnt signaling pathways regulate many cellular processes such as proliferation, migration and differentiation in embryonic development and maintenance of homeostasis in matured tissues. The deregulations and mutations in Wnt signaling pathway result in cancer. Unfortunately, there is no selective inhibitor for Wnt signaling. That is why targeting key components, such as SFRPs, WIF-1, DKK-1, APC, AXIN2, ICAT, LEF1 and beta-catenin, of

the Wnt signaling seems to be applicable in cancer treatment (Aguilera *et al.,* 2007). The topological values of these target proteins are shown in Table 4.30.

Luu *et al.* (2004) suggested that targeting beta-catenin could be a rational approach in cancer treatment. In our reconstructed networks, beta catenin has a connectivity value of 40 and participation percentage of 4.36 %. In addition to that, Frizzled 9, Wnt7A and Lef 1 proteins are found to be essential due to their high participation percentages in linear path analysis. Albers *et al.* (2011) shows that the Wnt receptor Frizzled-9 (Fzd-9) can be a new potential target for the treatment of osteoporosis by promoting bone formation. Also, it is known that the re-expression of Wnt7A and signaling through Fzd-9 are associated with increased differentiation and used in the lung cancer treatment (Winn *et al.*, 2005). Frizzled proteins are the receptors for Wnt ligand, and they are structurally similar to G protein-coupled receptors (GPCRs) which are the targets of and more than 50% of chemically applicable drugs (Yanaga and Sasaguri, 2007). So targeting frizzled proteins seems to be logical in cancer treatment.

The clustering coefficient is another topological aspect that is used in determining the essential proteins in a network. The clustering coefficient measures the degree of cliquishness of a node in the network. Cliquishness refers to the graph theory term “cliquish” which is used to designate complete graphs, i.e., those in which all nodes are connected to each other, and it is known that within the interaction network, essential proteins tend to be more cliquish (Yu *et al.,* 2004; Estrada, 2006). In the present Wnt network, there are two essential proteins (AXIN2 and APC) that have higher clustering coefficient values than the average (Table 4.29).

Ranking proteins according to their centrality measures can additionally be useful in selecting the possible drug targets (Estrada, 2006). Closeness Centrality, which is defined as the shortest paths to other nodes, is another aspect used in network analysis. The list of the putative target proteins that have higher closeness centrality values are listed in Table 4.30. From Table 4.30, it is seen that GSK3β and APC can be potential drug targets in Wnt signaling. This result is rational since Gsk3β and APC are essential proteins of Wnt signaling. In the absence of the ligands (Wnt proteins), the complex of Axin-beta catenin-APC and GSK3b result in beta catenin degradation, but when Wnt binds to frizzled proteins,

then beta catenin is stabilized and activate transcription. In addition to that, APC is related with colorectal cancer and APC-activating mutations are very common in colorectal cancer (Garber, 2009; Yanaga and Sasaguri, 2007).

Table 4. 30. The topological values of the target proteins.

Uniprot ID	Name	Average Shortest Path Length	Betwenness Centrality	Closeness Centrality	Clustering Coefficient	Degree	Participation Percentage
P35222	Beta-cat	**3.42**	1.43E-02	**0.292**	1.92E-02	40	**4.36E+00**
Q9UJU2	LEF1	**3.58**	4.52E-03	**0.279**	4.41E-02	17	**5.62E+01**
P49841	GSK3b	**3.42**	2.97E-03	**0.292**	8.57E-02	15	2.08E-01
P25054	APC	**3.66**	**7.45E-04**	**0.273**	**1.67E-01**	4	5.47E-02
Q8N474	SFRP1	6.49	**6.93E-04**	0.154	0.00E+00	4	-
O94907	DKK-1	**4.22**	**5.94E-04**	0.237	0.00E+00	2	-
Q9Y2T1	AXIN2	4.63	**5.73E-04**	0.216	**3.33E-01**	3	1.54E-03
Q14332	fzd2	4.57	**2.42E-04**	0.219	0.00E+00	2	-
Q9NSA3	ICAT	4.41	**3.36E-05**	0.227	0.00E+00	2	1.30E-03
Q9Y5W5	WIF-1	6.49	0.00E+00	0.154	0.00E+00	1	-
Q6UWE0	LRSAM1	**4.18**	8.73E-04	**0.239**	0.00E+00	5	8.10E-05
Q9P0L2	MARK1	**3.90**	4.78E-05	**0.256**	0.00E+00	2	8.10E-05
Q9NYL2	ZAK	**3.90**	6.21E-04	**0.256**	0.00E+00	3	8.10E-05
P35579	MHY9	**3.58**	5.61E-04	**0.280**	1.11E-01	10	8.10E-05
O00144	FZD9	4.80	4.71E-03	0.208	0.00E+00	4	**5.19E+01**
O00755	WNT7A	5.57	**4.27E-05**	0.179	0.00E+00	2	**5.19E+01**
Q8N752	KC1AL	**2.83**	1.40E-01	0.354	6.09E-03	241	**4.81E+00**
P63104	YWHAZ	**2.83**	1.18E-01	**0.353**	6.97E-03	200	9.46E-01
Q9BZK7	TBL1XR1	**3.46**	4.62E-02	**0.289**	7.14E-03	107	1.60E-01
Q16667	CNKD3	**4.31**	0.00E+00	**0.232**	**1.00E+00**	2	8.10E-05
P37173	TGFR2	**4.31**	**1.57E-04**	**0.232**	**6.00E-01**	5	4.69E-02
P01106	MYC	**4.14**	**1.53E-04**	**0.242**	**2.00E-01**	5	2.00E-03
Average		4.40	9.76E-04	0.232	1.13E-01	5.8	7.78E-01

Moreover, recent researches show that the betweenness centrality (number of shortest paths going through a node) and bridging centrality (nodes between and connecting clusters defined by the ratio of the number of interactions of a neighboring node over the number of remaining edges) are also effective in identifying the drug targets due to their position in communication (Hopkins, 2008, Hwang *et al.*, 2008). In order to prevent side effects and high lethality, the essential nodes with lower betweenness centrality values are chosen as drug targets on the purpose of not affecting the neighbors of the targeted protein which results in side effects. The list of the target proteins which have lower betweenness centrality values are given in Table 4.30. It is seen that APC, DKK1, SFRP1, AXIN2, FZD2, WNT7A,

ICAT and WIF1 are consistent with this fact. SFRP1 protein also needs more attention since its loss causes breast cancer (Klopocki *et al.*, 2004).

In this study, the topological properties of the important proteins are examined in addition to those of existing drug targets in literature. For this purpose, the average shortest path length, clustering coefficient, closeness centrality, betweenness centrality, degree and participation values are taken into consideration. In this step, the nodes that have lower average shortest path length, higher clustering coefficient, higher closeness centrality, lower betweenness centrality and higher participation percentages than the average values are obtained and written in bold (see Table 4.30). In this table the green color is used for existing drug targets in literature, blue color is used for the nodes which are obtained from linear path analysis with low participation percentage, brown color is used for the nodes which are obtained from linear path analysis for specific proteins whereas the pink color is used for the essential proteins obtained from linear path analysis and hubs of the network. It is seen that, the nodes which are specific to canonical and noncanonical pathways (CNKD3, TGFR2 and MYC) and the nodes which are detected as hub proteins (YWHAZ, TBL1XR1) have the quality of conformance since they have lower average shortest path length and higher closeness centrality values than the average. Besides that, Wnt7A has a lower betweenness centrality value and KC1AL has lower average shortest path length value than the average.

The Wnt signaling pathway is important for being related to several diseases including cancer when a deregulation or mutation occurs. That’s why it is one of the most attractive targets in cancer treatment studies. However, there is no selective inhibitor available for this pathway. The COX inhibitors, also known as anti-inflammatory drugs (NSAIDs), which are aspirin, indomethacin and sulindac, are known as the widely used drugs for diminishing the Wnt signaling. In addition to that, at present, there is only one drug, celecoxib, that is allowed by Food and Drug Administration (FDA) of the USA (Yanaga and Sasaguri, 2007).

4.7. Reduction of Reconstructed Signaling Network

The networks are important tools for studying the molecular processes in the cell. When the size of the considered networks is too large, then analyzing the data becomes more complicated. Network reduction is a common process used for simplicity. Hence, in this study, the reduction of the network for Wnt signaling in human was performed for a better understanding of the whole network. For this purpose, the betweenness centrality values were calculated for hundred random networks and the nodes are eliminated after the application of t-test with a threshold value of 0.01.

Only 479 nodes that correspond to 14 % of all the nodes in the Wnt network (3490 nodes) are eliminated by network reduction process. As a result, the number of accepted nodes is 3011 after this process. When the presence of the core proteins is examined, it is seen that 26 cores including the unwired ones are eliminated during the process and that corresponds to 22 % of the cores.

Table 4.31 gives the list of core proteins that are eliminated after reduction process. In Table 4.31, some information about eliminated proteins is missing since these proteins have no interaction and removed from the network as isolated nodes. The proteins Wnt10B, SFRP4, SFRP5, Wnt11 and DAAM2 are the unwired core proteins in the network. Consequently, the network reduction attempt caused a significant loss in core proteins, but did not result in a small network that can be used for further analysis.

Table 4.31. Eliminated core proteins after reduction.

Sub-network	Name	Uniprot ID	Degree	Participation in	Clustering
WNT/Beta-Catenin	WNT2B	Q93097	-	-	-
	WNT10B	O00744	-	-	-
	FRAT2	O75474	1	-	0.00E+00
	CSNK2A2	P19784	1	-	0.00E+00
	CTNNBIP1	Q9NSA3	2	1.30E-03	0.00E+00
	MAP3K7	O43318	9	4.30E-02	2.50E-01
	TCF7	P36402	2	-	0.00E+00
	TCF7L1	Q9HCS4	-	-	-
	WIF1	Q9Y5W5	1	-	0.00E+00
	CER1	O95813	-	-	-
	SFRP4	Q6FHJ7	-	-	-
	SFRP5	Q5T4F7	-	-	-
	CACYBP	Q9HB71	-	-	-
WNT/PCP	WNT11	O96014	-	-	-
	MAGI3	Q5TCQ9	5	1.11E-02	0.00E+00
	ROR2	Q01974	3	-	0.00E+00
	PTK7	Q13308	-	-	-
	CELSR1	Q9NYQ6	-	-	-
	CELSR3	Q9NYQ7	-	-	-
	PRINCKLE2	Q7Z3G6	1	-	0.00E+00
	DAAM2	Q86T65	-	-	-
	MAPK10	P53779	3	0.00E+00	0.00E+00
WNT/Calcium	PLCB3	Q01970	6	5.81E-01	1.33E-01
	PLCB4	Q15147	3	1.94E-03	3.33E-01
	PPP3R2	Q96LZ3	-	-	-
	NFATC4	Q14934	2	-	0.00E+00

5. CONCLUSIONS AND RECOMMENDATIONS

Wnt signaling is a major signaling pathway which has important roles in embryonic development of many species, and it is known that Wnt protein family is evolutionarily conserved among animals. The Wnt signaling is branched as canonical (Wnt/β-catenin) and non-canonical pathways (the planar cell polarity (PCP) and the Wnt-calcium (Wnt/Ca^{2+})). The deregulations and mutations in this signaling pathway cause several human diseases involving lung, breast, colon and colorectal cancers. Investigating the Wnt signaling pathway is therefore attractive for the determination of strategies to identify the suitable drug targets for therapeutic intervention in cancer treatment.

In this study, the reconstruction of Wnt signaling sub-networks and consequently the whole Wnt signaling network was performed in *H. Sapiens*. Furthermore, in order to provide an insight into the governing mechanisms of canonical and noncanonical wnt signaling, the following analyses were carried out. The graph theoretical analysis was enabled us to characterize the structure of the network via topological properties. The module analysis was done to understand the cellular role of the network and hence to validate the present reconstruction. The network decomposition analysis and crosstalk analysis allowed us to obtain detailed information about signaling mechanism such as bottleneck and bridging proteins, specific proteins, proteins with high and low participation in the linear paths, identification of putative drug target proteins, etc. Finally network reduction analysis is performed to see whether size reduction by applying statistical tests is rational or not for providing convenient usage of the reconstructed wnt signaling network for further analyses.

5.1. Conclusion

68, 33 and 32 core proteins were identified for canonical Wnt/beta-catenin and noncanonical Wnt/PCP and Wnt/Ca++ pathways respectively. The physical protein interaction data was obtained from MINT, BioGRID and HPRD databases, and 10592, 5928 and 6080 physical protein interactions were obtained from MINT, BioGRID and HPRD databases for beta-catenin, PCP and Ca++ pathways respectively.

6 core proteins could not be included into the whole Wnt network as they do not have any physical interactions. The proteins with missing interactions of the canonical network are Wnt10B, SFRP4 and SFRP5. Wnt 11 and DAAM2 are the unwired core proteins of PCP pathway, and Wnt 11 and PPP3R2 are the unwired ones for calcium pathway.

The reconstructed Wnt/β-catenin pathway is consisted of 3251 nodes and 9304 edges. For Wnt/PCP pathway 1952 nodes and 5001 edges were obtained. And the Wnt/ Calcium pathway had 2112 nodes and 5293 edges. For the whole Wnt network, 3489 nodes and 10092 edges were obtained. The network reduction attempt caused a significant loss in core proteins, and did not result in a small network that can be used for further analysis.

The network of domain-domain interactions were next reconstructed by association method. The numbers of unique domains were obtained as 1784, 1211 and 1204 for canonical, PCP and Ca++ pathways, respectively. The domain interaction numbers were found as 35569, 44492 and 23465 for canonical, PCP and Ca++ pathways, respectively.

The highest PID (potentially interacting domain pairs) score obtained for all Wnt DDI sub-networks was calculated as 18.80 for PF04857-PF07742 domain pair which are members of CAF1 family and BTG family, respectively. Depending on the frequency of domain-domain interactions, some protein interaction predictions were done. It is concluded that DAAM2 having formin homology2 domain may have interactions with the proteins ACL6A, ARP3, ACTC and ACTA1 which have actin domain. In addition to that, the SFRP4 protein can interact with FZD9, MUSK, FZD7, FZD5, FZD2, FZD4 and SFRP1 due to their Fz domains.

The degree distribution of the reconstructed Wnt networks followed a power-law model showing a scale-free network rather than a random one. The network diameters were found as 14, 13, 15 and 15 for Wnt/beta-catenin, PCP, calcium and whole Wnt signaling networks, respectively. The average shortest path lengths (or the characteristic path length) were found to be 4.46, 4.61, 4.56 and 4.40 in the same order as above. The clustering coefficients were obtained as 0.114, 0.112, 0.105 and 0.113 for Wnt/beta-catenin, PCP, calcium and whole Wnt signaling networks, respectively.

The hub proteins of the Wnt networks were also determined since these proteins are important for being dominant in the network and less connected nodes are linked to the network via these proteins. The KCLA1 (CSNK1A1L), YWHAZ and TBL1XR1 were detected as the hub proteins of the canonical Wnt network. PRKCA and PRCKB were obtained as the hub proteins of Wnt/Ca^{2+} network. In addition to that, YWHAZ protein was found to be the hub protein of all Wnt sub-networks. Moreover, the hub proteins KCLA1 (CSNK1A1L), YWHAZ, TBL1XR1, PRKCA and PRKCB were also detected as the bottleneck proteins of Wnt networks. All these proteins have certain roles in diseases like cancer, Alzheimer's, diabetes and Multiple sclerosis.

The module analysis was performed for understanding the cellular role of the network. Many of the proteins in the modules have roles in binding, catalytic activity and transcriptional regulation. Wnt protein binding and NADH dehydrogenase (ubiquinone) activity modules were detected for Wnt/beta-catenin sub-network. Rho GDP-dissociation inhibitor activity and potassium channel activity modules were detected in Planar Cell Polarity (PCP) sub-network and the proteins of module 6 were enriched in calcium ion binding in Wnt calcium sub-network.

The linear paths starting from Wnt ligand to the transcription factor proteins (Table 4.21) at 13 steps included 50% of the nodes in the whole Wnt network.

LEF1, WNT7A, FZD9, MDFI, WNT1, GOPC, FZD8, MCM5, MCM6, SMAD1 and PSB4 proteins were found to be involved in linear paths with high participation percentages (>25%). High participation percentages are naturally expected for WNT, FZD and LEF proteins since Wnts were the input and LEF1 was the output proteins of the linear path analysis.

LRSAM1, ZAK, MARK1 and MHY9 proteins are involved in only one linear path, and they have roles in certain diseases like cancer, autism and Fechtner syndrome.

The crosstalk analysis was performed and 239 nodes were obtained as common proteins in all Wnt sub-networks. YWHAZ and DVL2 proteins were detected as the major

crosstalk players due to their highest crosstalk values. This result is expectable since they are the essential proteins of the Wnt networks.

The key components of the Wnt signaling such as beta-catenin, GSK3B, APC, AXIN, DKK1, LEF1 and ICAT which are reported as potential drug targets in literature were also found to be drug targets due to their topological properties. Besides, CNKD3, TGFR2 and MYC proteins were here proposed as drug targets for Wnt signaling. In addition to that, targeting FZD9 and WNT7A seems rational since they have high participation percentages in Wnt linear paths and hence their inhibition will cause the whole wnt signaling down. The specific proteins obtained in linear path analysis are also proposed as drug target proteins due to their suitable topological properties. Finally, the hub proteins YWHAZ and TBL1XR1 may also be drug targets due to their effective roles in Wnt signaling pathway.

5.2. Recommendations

In the reconstruction process, a comprehensive database is needed since the data found in different databases are inconsistent in terms of protein IDs, not only in human but also in many species.

The interactome data are missing for many components of human signaling pathways. Experimental studies validating suggested interactions could give insight to the signaling mechanism in question.

The non-canonical Wnt signaling networks are not well understood and more investigations both experimental and computational should be done.

The conserved modules of the Wnt signaling network could be investigated among different species via various servers in order to provide a better understanding into the functionality of the proteins detected in the modules.

The reconstructed protein interaction networks are not directed and this situation constitutes a problem in understanding the signal transmittal mechanism of the networks, especially in crosstalk analysis of the different signaling networks. Nevertheless, the major

crosstalk players found in the investigations could be used in further experimental studies. The servers developed for the prediction of direction of interactions can be employed for obtaining a directed network.

The hubs and essential proteins (of the reconstructed Wnt networks), that are proposed to be putative drug targets as the result of graph theoretical and network decomposition analyses, could be used in further experimental investigations for drug design.

APPENDIX A: MATLAB CODES USED IN RECONSTRUCTION OF DOMAIN-DOMAIN INTERACTION NETWORKS AND NETWORK REDUCTION

```
clear all;
[num, txt]=xlsread('can-ddi-1.xlsx', 'Sheet4');
a=txt(:,3:18);
delete bahar_CAN_ddi_1.xls;
% clear c
b=length(a);
k=0;
hh = waitbar(0,'Please wait...');
for i=1:b
   waitbar(i/b);
   for j=1:8
      for n=9:16
         if length(a{i,j})>0
            if length(a{i,n})>0
               k=k+1;
               c(k,:)=[i, a(i,j), a(i,n)];
            end
         end
      end
   end
end
close(hh);
xlswrite('bahar_CAN_ddi_1.xls', c);
```

Figure A.1. Matlab code used in DDI in Wnt/β-catenin signaling network.

```
clear all;
[num, txt]=xlsread('pcpddi1.xlsx', 'Sheet4');
a=txt(:,3:16);
delete bahar_ppi_ddi_1.xls;
% clear c
b=length(a);
k=0;
hh = waitbar(0,'Please wait... ');
for i=1:b
   waitbar(i/b);
   for j=1:7
      for n=8:14
         if length(a{i,j})>0
            if length(a{i,n})>0
               k=k+1;
               c(k,:)=[i, a(i,j), a(i,n)];
            end
         end
      end
   end
end
close(hh);
xlswrite('bahar_pcp_ddi_1.xls', c);
```

Figure A.2. Matlab code used in DDI in Wnt/PCP signaling network.

```
clear all;
[num, txt]=xlsread('Caddi1.xlsx', 'Sheet4');
a=txt(:,3:16);
delete bahar_ca1.xls;
% clear c
b=length(a);
k=0;
hh = waitbar(0,'Please wait... ');
for i=1:b
   waitbar(i/b);
   for j=1:7
      for n=8:14
         if length(a{i,j})>0
            if length(a{i,n})>0
               k=k+1;
               c(k,:)=[i, a(i,j), a(i,n)];
            end
         end
      end
   end
end
close(hh);
xlswrite('bahar_ca1.xls', c);
```

Figure A.3. Matlab code used in DDI in Wnt/Ca^{2+} signaling network.

```
%% 1. READING INTERACTION FILE & FORMING ADJACENCY MATRIX
clear all
format compact
tic
%AGENET.TXT is the interaction file, supplied by Esra in xls format.
[node1, node2] = textread('int.txt','%s %s');
nInt= length(node1); % number of interactions
nodecomb = [node1; node2];
[nodelist,ii,ind]= unique(nodecomb);  %nodelist = nodecomb(ii);
ind1 = ind(1:nInt);
ind2 = ind(nInt+1:end);
% a number was assigned to each node.
% ind1 and ind2 correspond to the two columns in the text file in "numberized" format.
% nodelist: names of all nodes in the system.

num_nodes = length(nodelist); % number of nodes
AdjM = sparse(num_nodes, num_nodes);
for i = 1:length(node1)
   AdjM(ind1(i), ind2(i)) = 1;
   AdjM(ind2(i), ind1(i)) = 1;
end % Adjacency Matrix, AdjM was formed, by considering the symmetry of interactions.
nodelist;



%% 2. PARAMETERS  (using MATLAB BGL package)
bc   = betweenness_centrality(AdjM);
cf   = clustering_coefficients(AdjM);
spl  = mean(mean(all_shortest_paths(AdjM))); % mean shortest path length
```

Figure A.4. The matlab code used in the reduction of Wnt signaling network.

```
dmtr = max(max(all_shortest_paths(AdjM))); % diameter

AdjM_orig = AdjM; % back-up for the original adjacency matrix

%% 3. SWITCH RANDOMIZATION
num_rand = 100; % number of randomized networks to construct
for k = 1:num_rand
AdjM = AdjM_orig;
% AdjM is symmetric; so focus on the upper diagonal part.
[nh(:,1), nh(:,2)] = find(triu(AdjM)); % finds nonzero elements- i.e. true edges in upper diagonal
%nh is similar to [ind1 ind2] above; "numberized" versions of true interactions

num_edges = length(nh); %number of edges (interactions)
nh_orig = nh; % back-up

   for i = 1:num_edges % repeat a randomization for number-of-interaction times
      rind = ceil(rand(1,2)*num_edges); % not round but ceil to prevent zero index
      a = rind(1);  b = rind(2); % two edges are selected for switching

      edge1 = nh(a,:);
      edge2 = nh(b,:);
      new_edge1 = [edge1(1) edge2(2)];
      new_edge2 = [edge2(1) edge1(2)]; % the edges are switched
      if ((ismember(new_edge1, nh,'rows'))==0) & ...
        ((ismember(new_edge2, nh,'rows'))==0 )
         nh(a,2) = edge2(2);
         nh(b,2) = edge1(2);
```

Figure A.4. The matlab code used in the reduction of Wnt signaling network (Continued).

```
end % the edge switching is performed only if the "candidate edges" are not available in the network.
   end

   for j = 1:length(node1) % construct randomized adjacency matrix
      AdjM(nh(j,1), nh(j,2)) = 1;
      AdjM(nh(j,2), nh(j,1)) = 1;
   end
   bc_r(:,k)= betweenness_centrality(AdjM); %each column of bc_r corresponds to betweenness centrality
                                            % vector for one random network
end
toc

% this is a 3-way matrix: third index shows the index for randomized
0% network
```

Table A.4. The matlab code used in the reduction of Wnt signaling network (Continued).

APPENDIX B: COMMON NODES IN WNT SIGNALING NETWORKS

Table B.1. The common nodes detected in all Wnt sub-networks

P63104	P63165	P30153	P68133	Q9Y4J8	P38919	Q14738	Q86X55	Q9UMX6
Q9Y6K9	Q15843	P60953	Q9UQF2	O15151	Q06124	Q13177	Q8NI35	P11802
P61981	Q09472	Q9Y265	P55196	Q13362	Q15717	P55209	O14497	Q9Y6W5
P31947	Q15628	P36897	P61247	Q9H0H5	P17844	P23443	O43318	O75925
P54253	Q96GM5	P16885	P62875	O75496	Q99962	O00429	Q03135	P46940
P04637	P12236	P15531	Q14232	P17542	Q99963	Q8TAQ2	O43251	O15360
P31946	P67870	P62258	P05107	P24928	O60861	P51693	P09471	Q96LA8
P84022	Q92925	P20226	P05106	O00459	P15121	P68032	Q9NVC6	P06730
Q00653	P29353	Q92769	Q9UMX0	Q13616	Q8TEW0	P48729	Q9NPC8	Q9P2R6
Q15797	P27348	O43353	P19388	P06400	Q9UQM7	P48730	P18206	Q12906
P00533	Q6STE5	P27708	Q06187	Q14118	P63272	P67809	P10644	O14640
O15198	O15105	P68400	P19174	P21817	O95751	P53602	Q04864	Q693B1
P20333	Q13547	Q15645	P61978	Q6SJ96	Q00535	P48552	P07858	Q9UER7
P11021	P11142	P00519	Q13191	P31150	P04049	P43405	O96017	P68402
P62993	P31749	Q04206	P63000	Q15811	P10275	P62244	P51149	Q04771
P18031	O14920	Q9Y3R0	Q15172	Q9BZL6	P61218	P37198	P08069	O75791
Q15796	Q9UHD2	Q13546	Q96EY1	Q969Q1	Q15303	Q9UNH7	Q93074	Q9H9Y6
Q13233	Q13077	Q9Y297	O15530	O15379	Q05397	P02511	Q13330	Q9UDY8
P19838	Q99717	Q05516	Q92796	Q92830	P62913	Q96ST3	Q9Y4H2	Q07812
P04626	Q05586	P50570	P08107	P62487	P62942	Q6A1A2	P17275	Q8IXK0
Q9Y572	Q8IZP0	O14964	Q96B97	P53350	Q96FC9	Q16576	P43080	P42773
Q99750	P06241	Q15102	P55072	O43639	Q96RN5	O43541	P05412	P05783
Q99558	Q9NRD5	P62195	O15169	P30876	P18074	P17028	P49792	P56524
Q15834	P21796	P38936	Q14974	P42345	Q8WV28	Q92922	P10276	Q14005
Q01201	Q8N2W9	Q16537	P52435	Q7Z6J2	P15056	Q15654	O14776	O75569
Q13485	P24941	P17252	Q13232	Q13363	Q12929	O60880	Q9Y2X0	Q00534
P19438	P67775	P25963	P62714	Q9UHX1	P46108	P16220	Q9Y295	Q07889
P10809	P27986	P42858	P14672	Q9NZI2	P45983	Q9UHV7	P51532	Q9UKG1
Q12873	P12004	Q92844	Q14192	P46087	P56945	O95630	Q9NQ87	P49815
P35222	P30101	P03372	Q96PU8	Q9NS61	P08047	P06213	Q13163	P62820
P54259	Q93009	P07948	O14908	P78527	Q07817	P50395	O75376	O00418
P12757	P28482	Q16665	O14936	O75533	P41240	P42574	P04792	P31751
P62158	Q9NYJ8	Q92993	Q07666	P49768	Q96J02	P13569	O00308	P63098
Q5VTD9	P08842	Q9NZC7	Q99728	Q9Y2W7	P68106	Q58A45	P10415	Q719H9
Q9HAU4	P12956	O15111	Q8WUM4	Q9UKB1	P09874	P05556	Q9BXS5	Q9NPI1
Q12933	P07437	P27361	P62380	Q14790	P49674	A0JLT2	P54868	Q9Y5K6
P07900	P06493	O14745	O60674	Q13148	Q13224	P33151	Q96SB3	P98170
P78352	Q9NVJ2	O95376	P06748	P49841	Q12824	Q13153	Q96RT1	P06396

Table B.1. The common nodes detected in all Wnt sub-networks(Continued)

Q96F07	P81274	P15884	P19235	P11488	Q15311	Q53H80	P49639	Q13627	Q02447
P00492	Q9Y2A7	O15392	P21580	Q92831	Q96GD4	P05129	Q96AE4	Q9NPJ6	Q9HBW0
P00441	Q12879	Q63HR2	P61201	P16591	Q15369	P48637	P54764	O95390	Q14432
O75116	O75528	O43521	P25791	Q9NY61	P23467	P24385	P40189	O75369	P11717
Q9NP86	O75525	Q53EL6	P49023	Q08380	O00444	P21675	O00257	Q02078	P30556
P27797	P49757	O43561	P17535	Q8IVT5	Q14289	Q8N726	Q9NS23	O95999	Q9UNL4
Q9HCE7	O14757	P20936	P19793	P30305	O00499	Q9H461	Q9NS37	P20340	P37023
Q9NPB3	Q9Y243	P62491	P46527	P02751	P35226	P14373	O00233	P20337	Q00978
O00560	O00167	Q12905	Q96GD3	P45973	Q6P1K2	P51946	Q13554	P50552	Q02363
P05067	Q6QNY1	Q8N5S9	Q15349	P02768	P35240	O15399	P60763	P36405	P62330
P12830	Q13671	Q92904	P23458	P15172	O95373	Q6P5Z2	Q9BUZ4	P61106	P08758
Q86YR5	P49810	P45984	P04150	P15923	Q9ULH1	P29597	Q8WZ42	Q9Y616	O75928
Q9BQA1	Q9Y6Q9	P30874	P60484	P15924	O75312	P40394	Q86U70	Q13043	P29972
Q9BZE0	Q13114	P42338	P08581	Q8IXI2	P36406	Q9UNS2	Q13263	P11047	O75899
Q9BZE4	P57796	P10912	Q05513	Q07820	P09619	O43463	O14786	P55957	O15357
P46934	P02489	Q16594	O75381	Q13315	P48023	O14654	Q8WYH8	P98172	Q66K89
P48357	P16989	P30291	P48058	P01106	P51617	P51114	Q9UBF8	Q06587	P60174
P32320	Q09028	Q8IX12	P98161	P01116	Q9UKV3	P21980	P11217	P46695	Q16513
P29590	Q9UJU6	P16615	P18848	Q13426	Q15208	Q08345	P09769	Q9BXL7	P06733
Q9NXV2	O00401	Q99417	Q14134	Q9UQL6	P61088	P84077	P11230	Q14160	P17948
Q9H2X6	Q13625	Q8IXI1	P63167	O60315	O00311	P51153	O14775	O43157	Q5VSY0
O96013	Q8N163	Q13469	Q5TCQ9	P35900	O75081	P51159	P05549	O75096	Q6UVK1
P61586	Q8TAI7	P52566	P13861	P32121	P61457	Q495M9	P25054	O75084	O60934
P42336	Q9Y5P4	Q9UQB8	Q08881	O00254	Q9C0K0	P56915	Q9UBP9	P14921	P10071
P29350	Q9NR12	Q9HB75	Q96RU7	O00238	P42261	Q9H0Z9	O00144	P53667	P35711
Q99961	P63146	Q9HB90	P50613	P14635	Q15118	O43597	Q9Y239	Q99814	P08134
P10911	P14923	P36575	P37173	O75563	P42229	Q15633	Q9Y2U5	Q96S99	Q15759
O43586	Q99816	P42771	P17676	Q99504	P04632	P01243	O14733	Q99828	Q96KQ4
O60884	Q15119	O60313	P40763	Q99502	Q96EB6	O14543	Q05655	Q9H4B6	O96018
Q8IXH7	P13010	Q9NZN5	Q7LFL8	Q12852	Q13867	P01111	P32780	P28749	Q92934
Q9Y4C8	P43403	P46781	Q06830	O60260	Q01814	Q13402	P08631	P42224	P24522
O15117	Q99741	P20749	O43424	Q13255	O43255	Q9UQC2	P61244	P13807	P07949
P11387	P05023	Q9UHF1	Q99708	P04040	P20290	P36507	Q9Y6R0	P22460	O15273
Q93034	P29323	P40337	P25445	P51813	P20248	P50281	P61962	O60711	Q16643
P42684	Q96EK5	P39687	P50750	P24864	P10599	O60381	Q02750	O60716	P33402
O00213	Q08117	P20591	Q92759	Q96GX9	P10586	P17301	P61224	Q15078	Q9NVW2
O00291	Q9UM54	Q13227	P03886	P00966	O95163	Q5VZM2	P49840	O00716	Q92997
P56545	O60760	O15085	P32745	Q13136	Q96CV9	O15146	O60504	Q8N8B7	P17813

Table B.1. The common nodes detected in all Wnt sub-networks(Continued)

P30279	P49759	P26358	Q8IUD2	Q92928	P21860	Q00987	Q15173	P26010
P30281	Q15418	Q9UKW4	Q8WV24	P21964	P25705	Q92905	P35568	P49407
Q14526	P51812	Q9BXP8	P22392	P31994	P12931	O14773	P19387	Q04724
Q9H1R3	P12314	Q9UKL3	Q9HCK4	P14416	O14641	P06239	Q16539	Q15648
O60870	Q07890	Q8N108	Q16890	P51178	Q9Y4K3	O60563	Q15653	P38398
Q08209	Q9NZ42	Q92574	Q9H3D4	P19429	Q6VMQ6	Q14164	P21333	P01112
O14578	P25025	Q92542	Q92754	Q16610	P01100	P15692	Q9Y219	O94992
Q13467	P25098	P98194	Q8IUX8	Q8ND76	Q9HB21	P55042	Q99608	Q15375
Q13322	P78364	P41134	O60603	O43609	O15197	P55040	Q9Y2T1	P43250
Q99496	P78368	Q96SB4	Q92786	P51148	Q16288	P52434	Q9Y2W1	P05230
Q01094	Q13642	P47900	O95232	Q04759	P61020	Q13557	Q9Y371	O00463
Q12772	Q13639	P12755	Q00975	Q16620	Q16236	O95644	Q9Y2U8	P07384
Q12778	O75489	P21802	Q13936	P11532	P61006	P16144	Q8N2S1	P26441
Q8IXJ6	P52757	O14492	P24394	Q9BT81	Q99471	Q96P48	Q9UBS5	Q13618
Q13489	Q96PU5	P53671	P08754	O15294	Q9NSA3	P48201	P46821	O94907
Q15596	Q9NQC7	Q9HBA0	P07315	Q16667	Q13393	O75582	Q9UBS0	P15407
Q16342	P52701	P18075	Q9NP31	P30307	P28331	Q07954	P32239	Q9BZ71
Q15583	Q13155	Q15109	Q86YT6	P10909	P42680	Q13535	O00139	Q13794
Q15572	Q9Y6R4	Q15915	Q92800	Q99967	P42681	Q9H190	P32242	Q9ULH7
Q9NRW1	P49848	Q86SG6	P21127	P48443	Q13490	Q9NYB0	Q9Y2Y9	Q8TB22
P41279	P0C7P4	P27695	Q9H2G4	Q99933	P28300	O15031	P27037	Q8N196
Q9NZQ3	Q8IZX4	P48551	Q03701	P60033	Q9Y4K0	Q13291	P01730	P36402
Q8N488	Q9NQ66	P14174	P10826	Q7Z434	O00206	Q8WYP3	P08648	P04198
P42704	Q9Y5V3	Q92633	P52333	Q53ET0	Q9NS56	Q9NYA1	P61204	P53778
P56539	Q8IZW8	Q13882	P06746	O43572	P41235	Q9NYF8	P51572	Q13009
O00273	O00422	P16284	Q16584	P29466	P49683	Q9Y2H1	Q9NQB0	P17482
Q9NRM7	P43246	O43294	Q8WWW0	P20827	O60333	Q15382	Q68EM7	P50591
O14827	Q9UJM3	Q9ULL4	P29692	O76075	P11309	Q9UBC3	P04899	Q13049
P12429	Q92481	Q9ULJ8	P43694	Q9NX02	Q15573	Q8NEU8	P47736	P62070
Q9BX63	Q92499	Q9HBM6	Q969G3	Q96BI3	Q9UHR5	P22736	Q14185	Q06547
P46060	P07332	Q99732	Q14686	Q9UP38	P11362	Q9UBE8	P51531	P48788
Q13564	P35269	P29317	P08100	O14514	Q86VP6	Q9NYL2	P04083	Q14814
Q13523	P07550	Q15942	P17096	P15153	Q9NRZ9	P78317	Q14781	Q9UKV0
Q8IYF1	P51692	Q9ULV1	P19419	Q96BK5	Q15528	P49795	P23497	Q9UL26
P49715	O95409	Q9UM47	P26006	P15173	P54753	Q8N302	Q9Y6K1	P04271
Q9H0T7	P14859	Q99081	P08151	P01258	O00268	P11234	P29083	P19022
Q9H160	O75340	Q9HBZ2	Q9H307	Q96BR1	Q02040	P49760	P11137	O14907
Q12888	P20339	P09429	Q96KS0	O14544	Q7KZ85	O14763	Q15326	Q6W2J9

Table B.1. The common nodes detected in all Wnt sub-networks(Continued)

Q14155	O60244	P53041	Q14457	Q02156	P23634	Q92783	Q01196	P37840	Q7L576	Q8IYM9
Q9ULK4	Q15475	Q15269	Q9NVP2	Q9Y4W6	P55060	P49366	O60264	Q01970	P11831	Q9UBG7
Q92793	Q04917	P62826	Q99704	P55345	P35968	P49336	Q06330	Q00403	P49842	Q04900
P36954	Q13501	Q9Y5X4	P17612	P98082	Q13573	Q99697	P62745	Q05086	P52735	Q9UBL3
Q8IUQ4	O75448	P15498	P10636	Q99836	Q07912	Q9BUB5	Q9UIS9	O15397	P55211	P11233
Q9H422	Q13158	P23528	P15311	P07196	Q9H0M0	O15350	O14727	Q92879	P46531	P56693
O14939	P26583	Q9GZV4	P53355	P24588	Q9Y3M8	P51843	Q13127	Q13017	Q53QZ3	P53779
P35398	Q8IUC4	Q9H3U1	P01344	P58753	P43003	Q9Y2H0	Q9BYF1	Q13033	Q92556	P68036
O14976	Q9H3M7	P08709	O14682	O14508	P41231	Q13253	P61296	P58340	Q9P1Z2	P48736
Q9UGP4	P30542	P31260	Q8IVS8	Q8TF72	P78504	P09848	Q96H20	P31785	Q92736	Q6R327
Q8TDC3	P18146	P51948	O96019	O75691	Q9NS86	O60229	O95398	Q13045	P30626	Q7KZI7
Q9H4A3	Q92734	P13631	P01374	Q9UPZ9	Q8WZA2	Q9H7Z6	Q9H6Z9	Q03167	Q9P286	Q06546
Q9H4B4	Q8N9M5	O43513	P27540	Q99426	Q9UHB7	Q9UBI6	P55210	P37275	Q8IV90	Q96TC7
Q03112	Q6IE81	P12270	O15228	P01133	P25942	P49765	P55212	Q14872	Q13972	P51398
Q15121	P82094	P06703	Q9NWB7	O95886	P26232	P30086	Q8TBA6	Q86VZ5	P28347	Q07343
O00625	Q8N9B5	Q92870	P61587	Q13445	O14807	Q9Y2R2	O60519	Q03181	P31483	Q9BZW8
A6NI72	Q14332	O43524	Q96B36	O95838	P32119	Q8WYB5	Q9NQ69	Q9NR80	Q9UPX8	Q8IUC6
P04629	Q9GZX5	P10827	Q96AQ6	O95835	P55055	P24855	P62834	P61077	Q9UPY3	O15455
Q13873	O43374	P30988	Q04721	Q9UQB3	Q9UHF7	Q15406	Q13085	O43186	P78536	O14686
O43281	O60602	Q16512	O15264	Q96IZ0	O00220	P78395	P11926	P98174	Q14957	Q8N5U6
Q9P0L2	Q13952	Q92835	P51168	Q13404	Q7Z6Z7	O43815	Q9NPP4	P26367	Q6ZNA4	P01375
Q9HC29	Q13951	P29474	P84085	P28358	P54577	Q9Y266	P09629	Q9P035	P62873	Q96KR1
Q9UM73	Q00994	P31273	P51170	P36508	Q99570	P16066	O00468	Q9NR46	Q5T5U3	Q8NEC5
O15520	P09382	P31274	P31321	O75628	Q13219	P32238	q92908	Q9NQS1	P08575	O14672
O95197	Q9P212	P31270	P30872	O75626	Q8NER5	O00141	Q86Y01	O60447	P35227	Q14511
P43487	Q92785	P31268	O43657	O15156	O75593	P32249	O95379	Q9NR48	P32856	Q93084
P08887	Q13976	P20963	Q92974	Q9Y4G8	Q07960	P78347	Q9BYV6	Q9UL25	Q15036	Q13470
P10589	P36956	O43464	P31391	P17302	O00192	P22694	Q14209	Q9NQR1	Q08116	Q13464
Q9ULZ3	Q9BQI3	P43699	P29376	P21359	O15054	Q9Y6J9	Q9H5H4	Q9BXL6	O15524	Q13352
O95155	Q9UN19	Q99996	O43683	Q9UPV9	O15021	Q7L523	P20309	O14901	O15516	P01135
P19634	P05112	P20020	Q16595	Q99466	Q9H832	Q9UKD1	Q8TB24	Q96RJ6	Q6NYC1	
P24468	Q99623	O60921	Q9NWQ8	Q9H8V3	O60271	P53805	P63211	P35368	Q96EV8	
O95171	Q9Y6U3	P36896	P09172	P42685	Q86U86	P16035	Q08722	Q9UKN5	P06858	
O75962	Q13905	P08123	Q9C0C4	Q99490	Q8IZ40	O75460	P20336	Q96RL1	O43679	
Q16828	Q8TCU6	Q9NXR7	Q66K74	P42679	P53999	Q96PU4	P49137	Q13705	Q9BT49	
P34741	Q9GZU2	Q9P2T0	P15976	Q9UHL9	Q8N3E9	Q96Q42	O60481	P82279	O43680	
Q08050	Q9GZT9	Q9H2X0	Q32P28	Q6DT37	Q9UIW2	Q13683	P80723	Q9BY67	P25208	
Q9BPU6	Q9BZC1	O14649	Q15669	Q6PGN9	Q06413	Q15276	Q96QZ7	P23508	P39880	

APPENDIX C: TOP SCORED PIDS AND THEIR EXPERIMENTAL VERIFICATION

Table C.1. Top scored 25 and 50 PIDS and their verification of Wnt/Ca^{2+} signalling

Top 20 PIDS		**TOP 25 PIDS**	
PF04857-PF07742	√	PF04857-PF07742	√
PF00876-PF00864	-	PF00876-PF00864	-
PF09238-PF00727	-	PF09238-PF00727	-
PF00146-PF00329	-	PF00146-PF00329	-
PF00068-PF00305	-	PF00068-PF00305	-
PF09766-PF11957	-	PF09766-PF11957	-
PF00147-PF00354	-	PF00147-PF00354	-
PF10018-PF05527	-	PF10018-PF05527	-
PF11315-PF11594	-	PF11315-PF11594	-
PF11594-PF11315	-	PF11594-PF11315	-
PF04722-PF04722	√	PF04722-PF04722	√
PF01997-PF01997	√	PF01997-PF01997	√
PF00542-PF00542	√	PF00542-PF00542	√
PF00459-PF00459	√	PF00459-PF00459	√
PF00243-PF00243	√	PF00243-PF00243	√
PF01287-PF01916	-	PF01287-PF01916	-
PF08612-PF11568	-	PF08612-PF11568	-
PF01596-PF10601	-	PF01596-PF10601	-
PF05450-PF06105	-	PF05450-PF06105	-
PF00383-PF03870	-	PF00383-PF03870	-
		PF00059-PF00357	-
		PF05450-PF10251	-
		PF03153-PF03153	√
		PF01704-PF01704	√
		PF00383-PF00383	√

Table C.2. Top scored 50 PIDS and their verification of Wnt/Ca^{2+} signaling

TOP 50 PIDS-I		**TOP 50 PIDS-II**	
PF04857-PF07742	√	PF04503-PF01803	-
PF00876-PF00864	-	PF08920-PF01125	-
PF09238-PF00727	-	PF00086-PF00049	-
PF00146-PF00329	-	PF00219-PF00049	-
PF00068-PF00305	-	PF08612-PF09637	√
PF09766-PF11957	-	PF09637-PF08612	-
PF00147-PF00354	-	PF01023-PF08205	-
PF10018-PF05527	-	PF01248-PF00210	-
PF11315-PF11594	-	PF02269-PF04719	√
PF11594-PF11315	-	PF04719-PF02269	-
PF04722-PF04722	√	PF04968-PF07093	-
PF01997-PF01997	√	PF06083-PF00113	-
PF00542-PF00542	√	PF06083-PF03952	-
PF00459-PF00459	√	PF07524-PF03540	-
PF00243-PF00243	√	PF10406-PF03540	-
PF01287-PF01916	-	PF04695-PF04695	√
PF08612-PF11568	-	PF04078-PF04078	√
PF01596-PF10601	-	PF03367-PF03367	-
PF05450-PF06105	-	PF01387-PF01387	-
PF00383-PF03870	-	PF01221-PF01221	√
PF00059-PF00357	-	PF00596-PF00596	√
PF05450-PF10251	-	PF00751-PF08776	-
PF03153-PF03153	√	PF01391-PF00354	-
PF01704-PF01704	√	PF12489-PF06741	-
PF00383-PF00383	√	PF12489-PF07145	-

Table C.3. Top scored 100 PIDS and their verification of Wnt/Ca^{2+} signalling

TOP 100 PIDS-I		TOP 100 PIDS-II		TOP 100 PIDS-III		TOP 100 PIDS-IV	
PF04857-PF07742	√	PF04503-PF01803	-	PF00110-PF03062	-	PF05181-PF01286	-
PF00876-PF00864	-	PF08920-PF01125	-	PF06920-PF04727	-	PF05181-PF05181	-
PF09238-PF00727	-	PF00086-PF00049	-	PF02919-PF03367	-	PF00479-PF00479	√
PF00146-PF00329	-	PF00219-PF00049	-	PF01028-PF03367	-	PF00479-PF02781	√
PF00068-PF00305	-	PF08612-PF09637	√	PF00467-PF01916	-	PF02781-PF00479	-
PF09766-PF11957	-	PF09637-PF08612	-	PF00068-PF01477	-	PF02781-PF02781	√
PF00147-PF00354	-	PF01023-PF08205	-	PF04727-PF06920	-	PF01187-PF01398	-
PF10018-PF05527	-	PF01248-PF00210	-	PF05669-PF03980	-	PF10357-PF00118	-
PF11315-PF11594	-	PF02269-PF04719	√	PF05129-PF01214	-	PF05669-PF12052	-
PF11594-PF11315	-	PF04719-PF02269	-	PF01585-PF00657	-	PF02186-PF11521	-
PF04722-PF04722	√	PF04968-PF07093	-	PF12045-PF07412	-	PF11521-PF02186	-
PF01997-PF01997	√	PF06083-PF00113	-	PF05186-PF05186	√	PF01166-PF02798	-
PF00542-PF00542	√	PF06083-PF03952	-	PF01285-PF01285	√	PF01166-PF00043	-
PF00459-PF00459	√	PF07524-PF03540	-	PF00083-PF08232	-	PF06487-PF09743	-
PF00243-PF00243	√	PF10406-PF03540	-	PF04969-PF07093	-	PF06553-PF06553	√
PF01287-PF01916	-	PF04695-PF04695	√	PF05529-PF04194	-	PF06553-PF06553	√
PF08612-PF11568	-	PF04078-PF04078	√	PF01110-PF09240	-	PF06553-PF06553	√
PF01596-PF10601	-	PF03367-PF03367	-	PF00079-PF00262	-	PF01329-PF01329	√
PF05450-PF06105	-	PF01387-PF01387	-	PF05641-PF05994	-	PF00111-PF01044	-
PF00383-PF03870	-	PF01221-PF01221	√	PF12235-PF05994	-	PF00384-PF01044	-
PF00059-PF00357	-	PF00596-PF00596	√	PF00110-PF02019	-	PF09326-PF01044	-
PF05450-PF10251	-	PF00751-PF08776	-	PF00057-PF00110	-	PF10588-PF01044	-
PF03153-PF03153	√	PF01391-PF00354	-	PF06529-PF06529	-	PF07535-PF11600	-
PF01704-PF01704	√	PF12489-PF06741	-	PF01286-PF01286	-	PF07535-PF12253	-
PF00383-PF00383	√	PF12489-PF07145	-	PF01286-PF05181	√	PF00288-PF00579	-

Table C.4. Top scored 20 and 25 PIDS and their verification of Wnt/PCP signalling

TOP 20 PIDS		**TOP 25 PIDS**	
PF04857-PF07742	√	PF04857-PF07742	√
PF06292-PF01504	-	PF06292-PF01504	-
PF00876-PF00864	-	PF00876-PF00864	-
PF09238-PF00727	-	PF09238-PF00727	-
PF00147-PF00354	-	PF00147-PF00354	-
PF01290-PF00193	-	PF01290-PF00193	-
PF01290-PF02469	-	PF01290-PF02469	-
PF00326-PF00048	-	PF00326-PF00048	-
PF00930-PF00048	-	PF00930-PF00048	-
PF10018-PF05527	-	PF10018-PF05527	-
PF00146-PF05368	-	PF00146-PF05368	-
PF04699-PF02947	-	PF04699-PF02947	-
PF00146-PF01058	-	PF00146-PF01058	-
PF01733-PF10494	-	PF01733-PF10494	-
PF08612-PF09637	√	PF08612-PF09637	√
PF09637-PF08612	-	PF09637-PF08612	-
PF03467-PF04113	-	PF03467-PF04113	-
PF03807-PF03807	√	PF03807-PF03807	√
PF00542-PF00542	√	PF00542-PF00542	√
PF00459-PF00459	√	PF00459-PF00459	√
		PF00210-PF00210	√
		PF00899-PF09029	-
		PF01033-PF00079	-
		PF00110-PF03062	-
		PF01596-PF10601	-

Table C.5. Top scored 50 PIDS and their verification of Wnt/PCP signalling

TOP 50 PIDS I		**TOP 50 PIDS II**	
PF04857-PF07742	√	PF05450-PF06105	-
PF06292-PF01504	-	PF00146-PF00329	-
PF00876-PF00864	-	PF02374-PF06079	-
PF09238-PF00727	-	PF07763-PF03823	-
PF00147-PF00354	-	PF00111-PF02374	-
PF01290-PF00193	-	PF03172-PF01257	-
PF01290-PF02469	-	PF00899-PF00155	-
PF00326-PF00048	-	PF05450-PF10251	-
PF00930-PF00048	-	PF01248-PF10147	-
PF10018-PF05527	-	PF00735-PF00735	√
PF00146-PF05368	-	PF04045-PF00262	-
PF04699-PF02947	-	PF00080-PF04777	-
PF00146-PF01058	-	PF00068-PF00305	-
PF01733-PF10494	-	PF09766-PF11957	-
PF08612-PF09637	√	PF00219-PF00049	-
PF09637-PF08612	-	PF06632-PF09302	-
PF03467-PF04113	-	PF01585-PF00657	-
PF03807-PF03807	√	PF10357-PF00755	-
PF00542-PF00542	√	PF00467-PF01916	-
PF00459-PF00459	√	PF01287-PF01916	-
PF00210-PF00210	√	PF06083-PF00113	-
PF00899-PF09029	-	PF06083-PF03952	-
PF01033-PF00079	-	PF05110-PF03145	-
PF00110-PF03062	-	PF03849-PF02792	-
PF01596-PF10601	-	PF04695-PF04695	√

Table C.6. Top scored 100 PIDS and their verification of Wnt/PCP signalling

TOP 100 PIDS I		**TOP 100 PIDS II**		**TOP 100 PIDS III**		**TOP 100 PIDS IV**	
PF04857-PF07742	√	PF05450-PF06105	-	PF03367-PF03367	-	PF05485-PF10226	-
PF06292-PF01504	-	PF00146-PF00329	-	PF02374-PF02374	√	PF04538-PF10226	-
PF00876-PF00864	-	PF02374-PF06079	-	PF00808-PF00808	√	PF05510-PF01485	-
PF09238-PF00727	-	PF07763-PF03823	-	PF00596-PF00596	√	PF00313-PF00029	-
PF00147-PF00354	-	PF00111-PF02374	-	PF00248-PF00248	√	PF00313-PF03508	-
PF01290-PF00193	-	PF03172-PF01257	-	PF01342-PF01257	-	PF00313-PF10582	-
PF01290-PF02469	-	PF00899-PF00155	-	PF06487-PF10198	-	PF05693-PF04212	-
PF00326-PF00048	-	PF05450-PF10251	-	PF00657-PF04969	-	PF00878-PF10254	-
PF00930-PF00048	-	PF01248-PF10147	-	PF00383-PF03870	-	PF02919-PF03367	-
PF10018-PF05527	-	PF00735-PF00735	√	PF01242-PF01553	-	PF01028-PF03367	-
PF00146-PF05368	-	PF04045-PF00262	-	PF04969-PF00657	-	PF00571-PF04739	√
PF04699-PF02947	-	PF00080-PF04777	-	PF03792-PF04617	-	PF04739-PF00571	-
PF00146-PF01058	-	PF00068-PF00305	-	PF00343-PF01115	-	PF00288-PF00579	-
PF01733-PF10494	-	PF09766-PF11957	-	PF00235-PF00021	-	PF00288-PF01588	-
PF08612-PF09637	√	PF00219-PF00049	-	PF04057-PF00246	-	PF00579-PF00288	-
PF09637-PF08612	-	PF06632-PF09302	-	PF08646-PF00246	-	PF01588-PF00288	-
PF03467-PF04113	-	PF01585-PF00657	-	PF01433-PF04062	-	PF07654-PF00129	-
PF03807-PF03807	√	PF10357-PF00755	-	PF09127-PF04062	-	PF07654-PF06623	-
PF00542-PF00542	√	PF00467-PF01916	-	PF06467-PF07412	-	PF09377-PF09377	-
PF00459-PF00459	√	PF01287-PF01916	-	PF12045-PF07412	-	PF09377-PF01172	-
PF00210-PF00210	√	PF06083-PF00113	-	PF07763-PF05793	-	PF01172-PF09377	√
PF00899-PF09029	-	PF06083-PF03952	-	PF05615-PF00155	-	PF01172-PF01172	-
PF01033-PF00079	-	PF05110-PF03145	-	PF01285-PF01285	√	PF06529-PF06529	-
PF00110-PF03062	-	PF03849-PF02792	-	PF05529-PF04194	-	PF05187-PF05187	-
PF01596-PF10601	-	PF04695-PF04695	√	PF03145-PF08658	-	PF05187-PF01946	-

Table C.7. Top scored 20 and 25 PIDS and their verification of Wnt/β-catenin signalling

TOP 20 PIDS		**TOP 25 PIDS**	
PF04857-PF07742	√	PF04857-PF07742	√
PF09753-PF06775	-	PF09753-PF06775	-
PF00876-PF00864	-	PF00876-PF00864	-
PF09238-PF00727	-	PF09238-PF00727	-
PF08612-PF11568	-	PF08612-PF11568	-
PF11107-PF11510	-	PF11107-PF11510	-
PF08318-PF01105	-	PF08318-PF01105	-
PF11315-PF11594	-	PF11315-PF11594	-
PF11594-PF11315	-	PF11594-PF11315	-
PF09738-PF09738	-	PF09738-PF09738	-
PF03467-PF04113	-	PF03467-PF04113	-
PF00574-PF04106	-	PF00574-PF04106	-
PF03807-PF03807	√	PF03807-PF03807	√
PF03301-PF03301	√	PF03301-PF03301	√
PF04733-PF03208	-	PF04733-PF03208	-
PF03045-PF03045	-	PF03045-PF03045	-
PF02223-PF02223	√	PF02223-PF02223	√
PF02127-PF02127	√	PF02127-PF02127	√
PF00810-PF01650	-	PF00810-PF01650	-
PF00883-PF00883	√	PF00883-PF00883	√
		PF00542-PF00542	√
		PF00266-PF00266	√
		PF00243-PF00243	√
		PF01733-PF10494	-
		PF06292-PF01504	-

Table C.8. Top scored 50 PIDS and their verification of Wnt/β-catenin signalling

TOP 50 PIDS I		**TOP 50 PIDS II**	
PF04857-PF07742	√	PF02516-PF10384	-
PF09753-PF06775	-	PF02516-PF12346	-
PF00876-PF00864	-	PF08612-PF09637	√
PF09238-PF00727	-	PF09637-PF08612	-
PF08612-PF11568	-	PF05450-PF06105	-
PF11107-PF11510	-	PF10018-PF05527	-
PF08318-PF01105	-	PF01088-PF02953	-
PF11315-PF11594	-	PF01058-PF08547	-
PF11594-PF11315	-	PF02374-PF06079	-
PF09738-PF09738	-	PF07084-PF05076	-
PF03467-PF04113	-	PF07690-PF00209	-
PF00574-PF04106	-	PF05450-PF10251	-
PF03807-PF03807	√	PF04889-PF08216	-
PF03301-PF03301	√	PF05557-PF05557	√
PF04733-PF03208	-	PF03153-PF03153	√
PF03045-PF03045	-	PF00318-PF00293	-
PF02223-PF02223	√	PF12309-PF01244	-
PF02127-PF02127	√	PF00080-PF04777	-
PF00810-PF01650	-	PF03172-PF01257	-
PF00883-PF00883	√	PF09766-PF11957	-
PF00542-PF00542	√	PF06372-PF11095	√
PF00266-PF00266	√	PF05368-PF08547	-
PF00243-PF00243	√	PF00329-PF08547	-
PF01733-PF10494	-	PF08645-PF06632	-
PF06292-PF01504	-	PF11095-PF06372	-

Table C.9. Top scored 100 PIDS and their verification of Wnt/β-catenin signalling

TOP 100 PIDS I		TOP 100 PIDS II		TOP 100 PIDS III		TOP 100 PIDS IV	
PF04857-PF07742	√	PF02516-PF10384	-	PF11894-PF04097	-	PF09787-PF00995	-
PF09753-PF06775	-	PF02516-PF12346	-	PF00146-PF01058	-	PF00146-PF00329	-
PF00876-PF00864	-	PF08612-PF09637	√	PF12196-PF08065	-	PF05110-PF03145	-
PF09238-PF00727	-	PF09637-PF08612	-	PF04695-PF04088	-	PF06083-PF00113	-
PF08612-PF11568	-	PF05450-PF06105	-	PF02269-PF04719	√	PF06083-PF03952	-
PF11107-PF11510	-	PF10018-PF05527	-	PF06553-PF04437	-	PF01596-PF10601	-
PF08318-PF01105	-	PF01088-PF02953	-	PF04719-PF02269	-	PF07303-PF12063	-
PF11315-PF11594	-	PF01058-PF08547	-	PF04968-PF07093	-	PF00450-PF08658	-
PF11594-PF11315	-	PF02374-PF06079	-	PF01287-PF01916	-	PF04938-PF06003	-
PF09738-PF09738	-	PF07084-PF05076	-	PF03849-PF02792	-	PF05517-PF01387	-
PF03467-PF04113	-	PF07690-PF00209	-	PF10406-PF03540	-	PF10195-PF09324	-
PF00574-PF04106	-	PF05450-PF10251	-	PF03801-PF07160	-	PF02991-PF02144	-
PF03807-PF03807	√	PF04889-PF08216	-	PF03367-PF03367	-	PF06920-PF04727	-
PF03301-PF03301	√	PF05557-PF05557	√	PF02374-PF02374	√	PF03299-PF00637	-
PF04733-PF03208	-	PF03153-PF03153	√	PF01997-PF01997	√	PF03299-PF12451	-
PF03045-PF03045	-	PF00318-PF00293	-	PF01997-PF01997	√	PF04727-PF06920	-
PF02223-PF02223	√	PF12309-PF01244	-	PF04503-PF01803	-	PF12443-PF02213	-
PF02127-PF02127	√	PF00080-PF04777	-	PF01650-PF01650	-	PF06632-PF09302	-
PF00810-PF01650	-	PF03172-PF01257	-	PF00579-PF00579	√	PF00146-PF08547	-
PF00883-PF00883	√	PF09766-PF11957	-	PF00459-PF00459	√	PF00146-PF04095	-
PF00542-PF00542	√	PF06372-PF11095	√	PF09416-PF00735	-	PF04095-PF00146	-
PF00266-PF00266	√	PF05368-PF08547	-	PF00146-PF05368	-	PF01284-PF12063	-
PF00243-PF00243	√	PF00329-PF08547	-	PF02466-PF04495	-	PF04116-PF04116	-
PF01733-PF10494	-	PF08645-PF06632	-	PF00383-PF03870	-	PF01285-PF01285	√
PF06292-PF01504	-	PF11095-PF06372	-	PF11510-PF02106	-	PF05529-PF04194	-

Table C.10. The percentages of accuracies for different number of domain pairs with highest PIDs

Model	**Number of Domain Pairs**	**Accuracy (%)**
Wnt Canonical	Top 20 domain Pairs(2)	30
	Top 25 domain Pairs(3)	36
	Top 50 domain Pairs(4)	26
	Top 100 domain Pairs(5)	20
	All(6)	7
Wnt PCP DDI	Top 20 domain Pairs(2)	25
	Top 25 domain Pairs(3)	24
	Top 50 domain Pairs(4)	16
	Top 100 domain Pairs(5)	15
	All(6)	8
Wnt Calcium DDI	Top 20 domain Pairs(2)	30
	Top 25 domain Pairs(3)	36
	Top 50 domain Pairs(4)	30
	Top 100 domain Pairs(5)	25
	All(6)	6

APPENDIX D: THE TOPLOGICAL PROPERTIES OF THE TOP 350 NODES IN WNT NETWORK

Table D.1. Topological Properties of the nodes in Wnt signaling network

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
Q8N752	2.827E+00	1.400E-01	3.538E-01	6.086E-03	241
P63104	2.830E+00	1.180E-01	3.534E-01	6.973E-03	200
P17252	2.946E+00	8.022E-02	3.395E-01	1.272E-02	131
P05771	3.061E+00	6.267E-02	3.267E-01	7.641E-03	149
P04637	3.028E+00	4.813E-02	3.303E-01	1.739E-02	72
Q9BZK7	3.463E+00	4.619E-02	2.887E-01	7.143E-03	107
P54253	3.248E+00	3.726E-02	3.079E-01	4.785E-03	79
P61981	3.075E+00	3.254E-02	3.252E-01	1.465E-02	91
Q15428	3.333E+00	3.144E-02	3.000E-01	1.660E-02	95
Q9Y6K9	3.200E+00	2.693E-02	3.125E-01	4.009E-02	91
P84022	3.270E+00	2.617E-02	3.058E-01	1.058E-02	67
O15198	3.362E+00	2.570E-02	2.975E-01	7.650E-03	61
O14641	3.171E+00	2.419E-02	3.153E-01	1.883E-02	45
Q15796	3.281E+00	2.239E-02	3.048E-01	1.629E-02	57
Q99750	3.832E+00	2.181E-02	2.610E-01	0.000E+00	50
P00533	3.186E+00	2.106E-02	3.139E-01	5.424E-02	62
P31947	3.286E+00	2.058E-02	3.044E-01	1.793E-02	85
Q15797	3.439E+00	1.955E-02	2.908E-01	2.057E-02	60
Q15834	3.540E+00	1.933E-02	2.825E-01	1.064E-02	48
Q13526	3.439E+00	1.920E-02	2.908E-01	6.970E-02	40
Q5VTD9	3.515E+00	1.726E-02	2.845E-01	1.422E-03	38
P68400	3.206E+00	1.714E-02	3.119E-01	2.597E-02	22
P12757	3.553E+00	1.681E-02	2.814E-01	1.991E-02	38
P60709	3.170E+00	1.676E-02	3.155E-01	2.068E-01	56
P61586	3.510E+00	1.607E-02	2.849E-01	1.414E-02	47
Q12873	3.427E+00	1.595E-02	2.918E-01	3.414E-02	40
P12956	3.235E+00	1.531E-02	3.092E-01	1.333E-02	25
P08670	3.163E+00	1.466E-02	3.162E-01	3.175E-02	30
Q96KM6	3.532E+00	1.435E-02	2.831E-01	2.699E-03	39
P35222	3.423E+00	1.432E-02	2.921E-01	1.923E-02	40
Q13485	3.366E+00	1.421E-02	2.971E-01	3.656E-02	49
P62993	3.501E+00	1.382E-02	2.856E-01	4.644E-02	55
P18031	3.345E+00	1.362E-02	2.990E-01	2.413E-02	54
P20333	3.366E+00	1.357E-02	2.971E-01	1.475E-02	63
P62158	3.253E+00	1.349E-02	3.074E-01	4.319E-02	43
Q15843	3.427E+00	1.279E-02	2.918E-01	1.411E-02	32
P16885	3.475E+00	1.251E-02	2.878E-01	2.767E-02	23
Q6VMQ6	3.643E+00	1.228E-02	2.745E-01	5.682E-03	33

Tabl D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
Q09472	3.501E+00	1.216E-02	2.857E-01	1.515E-02	33
P12931	3.224E+00	1.181E-02	3.102E-01	9.269E-02	36
P07900	3.281E+00	1.171E-02	3.048E-01	8.333E-02	42
Q00610	3.259E+00	1.170E-02	3.069E-01	5.923E-02	44
P19838	3.196E+00	1.152E-02	3.129E-01	7.294E-02	53
Q9HAU4	3.496E+00	1.152E-02	2.860E-01	1.619E-02	41
P67870	3.389E+00	1.137E-02	2.951E-01	1.839E-02	32
Q9UQM7	3.312E+00	1.128E-02	3.019E-01	3.598E-02	35
Q15393	3.228E+00	1.122E-02	3.098E-01	5.714E-02	21
P31946	3.328E+00	1.092E-02	3.005E-01	7.925E-03	66
P11021	3.169E+00	1.076E-02	3.155E-01	2.240E-01	56
O75400	3.433E+00	1.051E-02	2.913E-01	3.175E-02	28
P54259	3.549E+00	1.032E-02	2.818E-01	1.652E-02	39
P31749	3.395E+00	1.029E-02	2.945E-01	3.941E-02	29
P21246	3.745E+00	1.011E-02	2.671E-01	1.846E-02	28
Q9Y572	3.427E+00	1.008E-02	2.918E-01	6.061E-03	57
Q9UKY1	3.438E+00	1.003E-02	2.909E-01	7.143E-02	21
P78352	3.586E+00	9.717E-03	2.789E-01	5.079E-02	36
P62136	3.414E+00	9.566E-03	2.929E-01	1.994E-02	27
P10809	3.352E+00	9.444E-03	2.983E-01	3.124E-01	42
P24941	3.407E+00	9.306E-03	2.936E-01	7.333E-02	27
Q13233	3.238E+00	9.001E-03	3.088E-01	1.388E-02	52
Q00653	3.313E+00	8.829E-03	3.019E-01	6.594E-02	60
P38646	3.315E+00	8.766E-03	3.016E-01	7.958E-02	37
P52294	3.385E+00	8.606E-03	2.954E-01	1.070E-02	34
P36897	3.605E+00	8.548E-03	2.774E-01	3.692E-02	26
Q93009	3.526E+00	8.538E-03	2.836E-01	3.986E-02	26
P06241	3.347E+00	8.498E-03	2.988E-01	7.672E-02	28
P06748	3.331E+00	8.441E-03	3.002E-01	7.692E-02	14
Q92905	3.519E+00	8.355E-03	2.842E-01	1.107E-01	23
Q01201	3.388E+00	8.340E-03	2.952E-01	6.364E-02	47
P67775	3.536E+00	8.255E-03	2.828E-01	2.279E-02	27
Q9NRD5	3.593E+00	8.079E-03	2.783E-01	2.769E-02	28
P23396	3.336E+00	8.042E-03	2.998E-01	3.509E-02	19
P21333	3.295E+00	7.889E-03	3.035E-01	2.924E-02	19
Q7L5N1	3.692E+00	7.879E-03	2.708E-01	6.154E-03	26
Q8N2W9	3.417E+00	7.778E-03	2.926E-01	8.000E-02	26
P22681	3.479E+00	7.506E-03	2.874E-01	7.460E-02	36

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
P30101	4.047E+00	7.474E-03	2.471E-01	1.000E-02	25
P63165	3.509E+00	7.464E-03	2.850E-01	6.629E-02	35
Q14974	3.261E+00	7.338E-03	3.067E-01	5.000E-02	16
P04626	3.569E+00	7.291E-03	2.802E-01	3.695E-02	52
Q9HD26	4.096E+00	7.276E-03	2.441E-01	8.333E-03	18
O95376	3.690E+00	7.249E-03	2.710E-01	5.848E-03	19
Q9Y3C7	3.747E+00	7.233E-03	2.669E-01	0.000E+00	25
O43707	3.543E+00	7.206E-03	2.822E-01	1.169E-01	24
P11142	3.267E+00	7.177E-03	3.060E-01	9.113E-02	29
P27348	3.417E+00	7.170E-03	2.927E-01	1.075E-02	31
P04406	3.250E+00	7.130E-03	3.077E-01	3.366E-01	41
O60563	3.750E+00	7.071E-03	2.667E-01	4.667E-02	25
Q8IZP0	3.651E+00	7.054E-03	2.739E-01	2.462E-02	28
Q13469	3.563E+00	6.988E-03	2.807E-01	0.000E+00	20
P12004	3.957E+00	6.912E-03	2.527E-01	7.905E-03	25
P15531	3.701E+00	6.910E-03	2.702E-01	6.316E-02	22
P08107	3.276E+00	6.893E-03	3.053E-01	5.833E-02	16
Q15645	3.940E+00	6.892E-03	2.538E-01	5.848E-03	21
P04628	5.489E+00	6.779E-03	1.822E-01	6.667E-02	10
Q99558	3.305E+00	6.725E-03	3.026E-01	5.762E-02	48
Q05516	3.926E+00	6.686E-03	2.547E-01	5.263E-03	20
Q96GM5	3.458E+00	6.630E-03	2.892E-01	4.234E-02	32
P11940	3.439E+00	6.584E-03	2.908E-01	1.732E-02	22
Q9Y4K3	3.599E+00	6.564E-03	2.778E-01	4.991E-02	34
Q06609	3.489E+00	6.354E-03	2.866E-01	3.557E-02	25
P00519	3.436E+00	6.292E-03	2.911E-01	9.048E-02	21
Q15047	3.671E+00	6.245E-03	2.724E-01	1.994E-02	27
Q96RN5	3.729E+00	6.192E-03	2.682E-01	2.727E-01	12
Q9Y230	3.511E+00	6.075E-03	2.848E-01	4.737E-02	20
P61978	3.305E+00	6.035E-03	3.025E-01	1.500E-01	16
O43395	3.632E+00	6.027E-03	2.753E-01	3.297E-02	16
Q13547	3.573E+00	6.008E-03	2.798E-01	8.201E-02	30
Q96EY1	3.521E+00	6.007E-03	2.840E-01	1.667E-02	16
O15105	3.629E+00	5.924E-03	2.756E-01	7.389E-03	29
P27986	3.541E+00	5.903E-03	2.824E-01	7.692E-02	26
P50570	3.712E+00	5.841E-03	2.694E-01	8.497E-02	20
P29353	3.474E+00	5.830E-03	2.879E-01	1.089E-01	32
Q06187	3.575E+00	5.749E-03	2.798E-01	2.941E-02	19

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
Q14194	3.696E+00	5.736E-03	2.706E-01	0.000E+00	21
Q12933	3.414E+00	5.724E-03	2.929E-01	1.569E-01	36
P06493	3.435E+00	5.635E-03	2.911E-01	9.957E-02	24
P62805	3.560E+00	5.596E-03	2.809E-01	3.810E-02	15
O95405	3.627E+00	5.591E-03	2.757E-01	1.732E-02	22
Q9UMX0	3.836E+00	5.544E-03	2.607E-01	0.000E+00	18
P40818	3.528E+00	5.540E-03	2.835E-01	3.333E-02	16
P42858	3.688E+00	5.474E-03	2.711E-01	2.632E-02	20
Q9Y3C5	3.649E+00	5.463E-03	2.740E-01	1.581E-02	23
Q12906	3.581E+00	5.428E-03	2.793E-01	3.571E-02	8
P55196	3.321E+00	5.352E-03	3.011E-01	5.147E-02	17
O14964	3.686E+00	5.310E-03	2.713E-01	4.762E-02	21
P38936	3.763E+00	5.306E-03	2.657E-01	4.678E-02	19
P03372	3.509E+00	5.283E-03	2.850E-01	5.882E-02	18
Q05586	3.690E+00	5.265E-03	2.710E-01	6.462E-02	28
P49407	3.521E+00	5.256E-03	2.840E-01	1.099E-02	14
P63261	3.448E+00	5.245E-03	2.900E-01	9.247E-02	33
P30153	3.355E+00	5.154E-03	2.980E-01	9.474E-02	22
Q16539	3.502E+00	5.098E-03	2.856E-01	3.509E-02	19
P61247	3.461E+00	5.094E-03	2.889E-01	5.882E-02	17
Q16665	3.642E+00	5.088E-03	2.745E-01	1.307E-02	18
O75554	3.579E+00	5.088E-03	2.794E-01	4.762E-02	28
P27797	3.718E+00	5.062E-03	2.689E-01	2.778E-02	9
P05556	3.658E+00	5.000E-03	2.734E-01	3.636E-02	11
P19438	3.509E+00	4.917E-03	2.850E-01	4.343E-02	47
P63279	3.529E+00	4.913E-03	2.833E-01	7.619E-02	17
P06239	3.603E+00	4.862E-03	2.776E-01	3.986E-02	24
P28482	3.371E+00	4.817E-03	2.966E-01	7.971E-02	26
P25705	3.306E+00	4.813E-03	3.025E-01	4.397E-01	36
O00144	4.802E+00	4.708E-03	2.083E-01	0.000E+00	4
Q9Y3A3	4.053E+00	4.603E-03	2.467E-01	0.000E+00	12
A0JLT2	4.113E+00	4.585E-03	2.432E-01	0.000E+00	14
Q9UKV8	3.685E+00	4.576E-03	2.714E-01	1.705E-02	33
Q93062	3.790E+00	4.571E-03	2.639E-01	3.268E-02	20
P14672	4.025E+00	4.562E-03	2.485E-01	0.000E+00	15
O75808	3.551E+00	4.526E-03	2.816E-01	3.846E-02	13
P78527	3.392E+00	4.523E-03	2.948E-01	9.890E-02	14
Q9UJU2	3.585E+00	4.522E-03	2.790E-01	4.412E-02	17

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
O95067	3.737E+00	4.488E-03	2.676E-01	2.198E-02	14
P60953	3.678E+00	4.430E-03	2.719E-01	4.743E-02	23
Q9Y4J8	3.987E+00	4.399E-03	2.508E-01	9.524E-02	15
P07948	3.469E+00	4.391E-03	2.882E-01	4.678E-02	19
P62258	3.325E+00	4.391E-03	3.007E-01	3.810E-02	21
O75496	4.434E+00	4.354E-03	2.255E-01	0.000E+00	14
P62714	3.600E+00	4.349E-03	2.778E-01	8.333E-03	16
O14745	4.041E+00	4.343E-03	2.474E-01	1.170E-02	19
P21817	4.108E+00	4.334E-03	2.435E-01	0.000E+00	13
Q13432	4.002E+00	4.310E-03	2.499E-01	0.000E+00	23
Q14192	3.755E+00	4.291E-03	2.663E-01	0.000E+00	17
Q99717	3.572E+00	4.277E-03	2.800E-01	5.128E-02	27
P62316	3.554E+00	4.266E-03	2.814E-01	2.745E-01	18
Q9Y297	3.672E+00	4.265E-03	2.724E-01	1.053E-01	22
P35609	3.693E+00	4.251E-03	2.708E-01	1.028E-01	25
P40425	4.344E+00	4.219E-03	2.302E-01	0.000E+00	10
Q92993	3.741E+00	4.186E-03	2.673E-01	0.000E+00	18
O15530	3.620E+00	4.162E-03	2.763E-01	3.333E-02	16
P63172	3.885E+00	4.127E-03	2.574E-01	1.099E-02	14
Q15102	3.558E+00	4.107E-03	2.810E-01	7.353E-03	19
P55072	3.452E+00	4.049E-03	2.897E-01	2.198E-02	16
P08621	3.566E+00	4.028E-03	2.805E-01	1.978E-01	14
Q8TBB1	3.743E+00	4.023E-03	2.672E-01	7.353E-03	17
Q9Y6I3	3.690E+00	4.005E-03	2.710E-01	4.396E-02	14
O43639	3.691E+00	3.947E-03	2.709E-01	1.099E-02	14
Q92925	3.501E+00	3.918E-03	2.857E-01	4.368E-02	30
P56524	3.517E+00	3.892E-03	2.843E-01	7.143E-02	8
Q15365	3.507E+00	3.883E-03	2.852E-01	1.000E-01	16
P63000	3.668E+00	3.882E-03	2.726E-01	3.922E-02	18
Q14118	4.271E+00	3.874E-03	2.341E-01	1.061E-01	14
Q14232	3.816E+00	3.866E-03	2.620E-01	1.667E-02	18
P38398	3.658E+00	3.853E-03	2.734E-01	2.198E-02	14
Q07954	4.530E+00	3.831E-03	2.208E-01	0.000E+00	2
P05783	3.409E+00	3.830E-03	2.934E-01	1.556E-01	10
Q9H1D0	3.770E+00	3.804E-03	2.652E-01	0.000E+00	4
P05549	4.249E+00	3.773E-03	2.354E-01	4.762E-01	7
O14908	3.879E+00	3.771E-03	2.578E-01	2.198E-02	16
P25963	3.603E+00	3.755E-03	2.775E-01	1.930E-01	19

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
P21860	3.751E+00	3.736E-03	2.666E-01	2.857E-02	35
P48730	3.489E+00	3.735E-03	2.866E-01	1.556E-01	12
Q99873	3.501E+00	3.651E-03	2.856E-01	1.324E-01	19
O14936	3.985E+00	3.643E-03	2.510E-01	0.000E+00	15
P49674	3.489E+00	3.623E-03	2.866E-01	1.455E-01	13
Q92769	3.597E+00	3.610E-03	2.780E-01	1.111E-01	21
P31150	3.647E+00	3.585E-03	2.742E-01	2.121E-01	14
P54274	4.145E+00	3.520E-03	2.412E-01	4.762E-02	9
Q04206	3.614E+00	3.507E-03	2.767E-01	1.316E-01	20
O15169	3.598E+00	3.505E-03	2.779E-01	6.593E-02	16
P07910	3.384E+00	3.502E-03	2.955E-01	0.000E+00	10
Q8IUQ4	3.556E+00	3.500E-03	2.812E-01	1.515E-02	12
Q6STE5	3.533E+00	3.459E-03	2.831E-01	3.448E-02	29
O14543	4.396E+00	3.436E-03	2.275E-01	1.667E-01	4
P55198	4.725E+00	3.435E-03	2.116E-01	0.000E+00	2
O43765	4.217E+00	3.422E-03	2.371E-01	0.000E+00	10
P20226	4.145E+00	3.421E-03	2.412E-01	1.503E-01	20
P56704	5.441E+00	3.393E-03	1.838E-01	1.429E-01	9
P63208	3.732E+00	3.385E-03	2.679E-01	2.092E-01	20
Q9H461	4.902E+00	3.385E-03	2.040E-01	0.000E+00	4
O94973	3.622E+00	3.363E-03	2.761E-01	1.944E-01	9
P25788	3.755E+00	3.321E-03	2.663E-01	4.800E-01	28
P52292	3.631E+00	3.294E-03	2.754E-01	4.301E-03	31
O15085	3.702E+00	3.291E-03	2.701E-01	0.000E+00	7
P02792	4.213E+00	3.290E-03	2.374E-01	0.000E+00	9
Q9UHD2	3.349E+00	3.281E-03	2.986E-01	1.396E-01	29
P04049	3.536E+00	3.231E-03	2.828E-01	6.593E-02	14
P49720	3.585E+00	3.231E-03	2.789E-01	5.833E-01	24
O00505	3.620E+00	3.227E-03	2.762E-01	2.646E-03	28
P17542	3.974E+00	3.227E-03	2.516E-01	4.396E-02	14
O15111	3.370E+00	3.213E-03	2.968E-01	2.579E-01	20
Q13363	3.877E+00	3.201E-03	2.579E-01	1.282E-02	13
P13569	3.627E+00	3.196E-03	2.757E-01	2.222E-02	12
Q13257	3.799E+00	3.192E-03	2.632E-01	3.810E-02	17
P63146	4.247E+00	3.188E-03	2.355E-01	0.000E+00	7
Q3ZCQ8	3.622E+00	3.187E-03	2.761E-01	7.368E-02	20
Q9UL18	3.753E+00	3.183E-03	2.665E-01	1.994E-02	29
P06576	3.405E+00	3.183E-03	2.937E-01	5.057E-01	33

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
P17980	3.795E+00	3.165E-03	2.635E-01	2.583E-01	16
P05107	3.622E+00	3.135E-03	2.761E-01	4.575E-02	18
P53350	3.581E+00	3.118E-03	2.793E-01	7.273E-02	13
Q00987	3.609E+00	3.118E-03	2.771E-01	7.619E-02	23
Q9Y333	3.936E+00	3.112E-03	2.541E-01	7.353E-02	17
Q07666	3.616E+00	3.083E-03	2.766E-01	1.905E-02	15
P13861	4.142E+00	3.076E-03	2.415E-01	0.000E+00	5
Q13077	3.534E+00	3.052E-03	2.830E-01	1.200E-01	28
Q9BQQ3	4.085E+00	3.050E-03	2.448E-01	0.000E+00	12
P10276	3.802E+00	3.048E-03	2.630E-01	2.222E-02	10
Q13554	3.574E+00	3.045E-03	2.798E-01	1.282E-01	15
P16220	3.805E+00	3.030E-03	2.628E-01	1.818E-02	11
Q92793	3.696E+00	3.023E-03	2.706E-01	0.000E+00	14
Q9BZL6	3.613E+00	3.003E-03	2.768E-01	9.524E-03	15
P49841	3.423E+00	2.973E-03	2.921E-01	8.571E-02	15
P26368	3.563E+00	2.965E-03	2.807E-01	7.692E-02	13
Q9NQ87	3.700E+00	2.956E-03	2.703E-01	2.778E-02	9
Q92997	3.520E+00	2.956E-03	2.841E-01	5.147E-02	19
O43353	3.470E+00	2.950E-03	2.881E-01	1.170E-01	21
P04156	3.809E+00	2.920E-03	2.625E-01	0.000E+00	12
P62244	3.745E+00	2.892E-03	2.670E-01	1.026E-01	13
Q13402	5.243E+00	2.871E-03	1.907E-01	0.000E+00	4
Q16181	4.747E+00	2.865E-03	2.107E-01	0.000E+00	6
P53567	5.722E+00	2.864E-03	1.748E-01	0.000E+00	2
Q9Y478	4.058E+00	2.864E-03	2.464E-01	5.000E-01	4
P62195	3.860E+00	2.856E-03	2.591E-01	2.279E-01	19
Q8NI35	3.816E+00	2.845E-03	2.621E-01	3.030E-02	12
P62333	3.754E+00	2.826E-03	2.664E-01	2.288E-01	20
Q9H0H5	3.635E+00	2.818E-03	2.751E-01	8.791E-02	14
P62140	3.507E+00	2.800E-03	2.851E-01	2.924E-02	19
P08047	3.710E+00	2.790E-03	2.695E-01	1.818E-02	11
Q99963	3.727E+00	2.765E-03	2.683E-01	1.515E-02	12
P17844	3.415E+00	2.763E-03	2.928E-01	7.273E-02	13
P42330	4.283E+00	2.760E-03	2.335E-01	6.545E-01	13
P48729	3.665E+00	2.748E-03	2.729E-01	5.455E-02	11
O75586	3.870E+00	2.713E-03	2.584E-01	2.564E-02	13
P27361	3.427E+00	2.704E-03	2.918E-01	1.397E-01	19
O75376	3.906E+00	2.690E-03	2.560E-01	1.944E-01	9

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
Q14566	3.968E+00	2.689E-03	2.520E-01	4.708E-01	28
Q16623	3.662E+00	2.665E-03	2.731E-01	5.556E-02	9
P10911	4.245E+00	2.641E-03	2.356E-01	0.000E+00	10
Q9H422	4.021E+00	2.625E-03	2.487E-01	0.000E+00	12
Q14155	3.868E+00	2.623E-03	2.585E-01	3.846E-02	13
P31689	3.594E+00	2.621E-03	2.782E-01	1.225E-01	25
P54257	4.004E+00	2.620E-03	2.497E-01	8.333E-03	16
P68133	3.517E+00	2.618E-03	2.843E-01	1.917E-01	18
Q9NZC7	3.652E+00	2.600E-03	2.738E-01	3.922E-02	18
P25786	3.819E+00	2.594E-03	2.618E-01	4.206E-01	28
P53602	4.548E+00	2.587E-03	2.199E-01	0.000E+00	10
P51532	4.010E+00	2.565E-03	2.494E-01	2.778E-02	9
Q15811	3.962E+00	2.556E-03	2.524E-01	1.282E-02	13
P29350	3.836E+00	2.544E-03	2.607E-01	0.000E+00	8
P19174	3.802E+00	2.541E-03	2.630E-01	1.500E-01	16
Q92796	3.733E+00	2.534E-03	2.679E-01	1.667E-01	16
P50453	3.560E+00	2.531E-03	2.809E-01	1.905E-02	15
Q9Y3R0	3.811E+00	2.511E-03	2.624E-01	4.737E-02	20
Q9UNE7	3.768E+00	2.505E-03	2.654E-01	0.000E+00	10
P49736	4.051E+00	2.502E-03	2.469E-01	1.515E-02	14
Q96PU8	3.976E+00	2.489E-03	2.515E-01	3.810E-02	15
P61968	3.905E+00	2.472E-03	2.561E-01	2.222E-02	10
P32780	3.995E+00	2.470E-03	2.503E-01	0.000E+00	4
P14373	4.025E+00	2.467E-03	2.485E-01	0.000E+00	5
P16104	3.716E+00	2.456E-03	2.691E-01	7.273E-02	11
P48552	3.883E+00	2.452E-03	2.575E-01	4.444E-02	10
P27708	3.318E+00	2.419E-03	3.014E-01	1.000E-01	22
P09874	3.634E+00	2.417E-03	2.752E-01	5.455E-02	11
Q9UQB8	3.762E+00	2.410E-03	2.658E-01	0.000E+00	6
Q13625	3.547E+00	2.403E-03	2.819E-01	9.524E-02	7
P07947	3.544E+00	2.393E-03	2.822E-01	0.000E+00	8
Q96KQ7	4.145E+00	2.384E-03	2.412E-01	1.515E-02	12
P42345	3.380E+00	2.372E-03	2.959E-01	3.030E-02	14
Q9UEF7	6.486E+00	2.370E-03	1.542E-01	0.000E+00	4
O75533	3.556E+00	2.369E-03	2.812E-01	1.455E-01	13
Q13153	4.195E+00	2.358E-03	2.384E-01	4.444E-02	10
P35568	3.643E+00	2.355E-03	2.745E-01	1.250E-01	17
Q15459	3.677E+00	2.351E-03	2.720E-01	7.692E-02	13

Table D.1. Topological Properties of the nodes in Wnt signaling network (Continued)

Nodes	Ave. Shortest Path Length	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Degree
P46108	3.653E+00	2.350E-03	2.738E-01	1.455E-01	11
Q13163	4.002E+00	2.344E-03	2.499E-01	2.778E-02	9
Q96RK0	3.711E+00	2.339E-03	2.694E-01	9.524E-02	7
Q7Z6J2	4.011E+00	2.325E-03	2.493E-01	1.818E-02	13
P02751	4.433E+00	2.323E-03	2.256E-01	6.667E-02	6
P10415	3.724E+00	2.322E-03	2.685E-01	3.636E-02	11
O96013	3.418E+00	2.321E-03	2.926E-01	1.111E-01	9
O95751	3.832E+00	2.316E-03	2.610E-01	8.974E-02	13
Q15717	3.939E+00	2.311E-03	2.539E-01	3.030E-02	12
Q15475	3.785E+00	2.306E-03	2.642E-01	1.818E-02	11
P35638	6.719E+00	2.292E-03	1.488E-01	0.000E+00	3
P40189	5.394E+00	2.292E-03	1.854E-01	0.000E+00	4
O95409	4.390E+00	2.292E-03	2.278E-01	0.000E+00	3
P06400	3.680E+00	2.271E-03	2.717E-01	6.410E-02	15
P35080	3.882E+00	2.259E-03	2.576E-01	2.778E-02	9
P49768	3.897E+00	2.246E-03	2.566E-01	2.564E-02	13
Q13616	3.820E+00	2.225E-03	2.618E-01	1.319E-01	14
Q15653	3.691E+00	2.212E-03	2.709E-01	1.699E-01	18
O14920	3.384E+00	2.210E-03	2.955E-01	1.552E-01	31
Q93034	3.598E+00	2.208E-03	2.780E-01	9.524E-02	7
Q9BXS5	3.801E+00	2.193E-03	2.631E-01	0.000E+00	11
Q96S59	3.797E+00	2.192E-03	2.634E-01	3.030E-02	12
Q15628	3.515E+00	2.187E-03	2.845E-01	1.075E-01	33
P67809	3.408E+00	2.172E-03	2.934E-01	5.455E-02	11
P11388	3.592E+00	2.172E-03	2.784E-01	1.389E-01	11
Q15427	3.955E+00	2.171E-03	2.528E-01	1.071E-01	8
Q13158	3.983E+00	2.166E-03	2.511E-01	5.455E-02	11
Q02575	3.669E+00	2.163E-03	2.726E-01	0.000E+00	16
Q02577	3.669E+00	2.163E-03	2.726E-01	0.000E+00	16
P09471	4.076E+00	2.161E-03	2.454E-01	0.000E+00	10
Q9NR12	3.603E+00	2.160E-03	2.776E-01	5.556E-02	9
Q13224	3.650E+00	2.153E-03	2.740E-01	1.923E-01	13
P45983	3.555E+00	2.150E-03	2.813E-01	1.818E-02	11
Q15648	3.861E+00	2.149E-03	2.590E-01	5.055E-01	14
P36873	3.498E+00	2.141E-03	2.859E-01	4.412E-02	17
O95400	3.875E+00	2.134E-03	2.581E-01	5.000E-02	16
Q15654	3.765E+00	2.130E-03	2.656E-01	2.222E-02	10

APPENDIX E: CROSSTALK VALUES OF PROTEINS IN WNT SIGNALING

Table E.1. The signaling proteins that have non-zero network crosstalk values

Protein_ID	Name	Crosstalk Values	Protein_ID	Name	Crosstalk Values
P63104	YWHAZ	11	Q14192	FHL2	2
O14641	DVL2	11	Q9Y297	BTRC	2
P62158	CALM3	6	P63000	RAC1	2
P07900	HSP90AA1	6	P49674	CSNK1E	2
Q9Y572	RIPK3	6	P04049	RAF1	2
Q15796	SMAD2	5	O15111	CHUK	2
P36897	TGFBR1	5	P13569	CFTR	2
P04637	TP53	4	P05107	ITGB2	2
P61981	YWHAG	4	Q9BZL6	PRKD2	2
Q9Y6K9	IKBKG	4	P49841	GSK3B	2
P84022	SMAD3	4	P62244	RPS15A	2
P31947	SFN	4	Q8NI35	INADL	2
Q13485	SMAD4	4	P27708	PYR1	2
P20333	TNFRSF1B	4	P10415	BCL2	2
Q9HAU4	SMURF2	4	O96013	PAK4	2
Q9UQM7	CAMK2A	4	O14920	IKBKB	2
P00533	EGFR	3	Q9BXS5	AP1M1	2
Q06187	BTK	3	Q9NR12	PDLIM7	2
P19438	TNFRSF1A	3	Q13224	GRIN2B	2
Q9NZI2	KCNIP1	3	P45984	MAPK9	2
Q9NS61	KCNIP2	3	O14640	DVL1	2
Q9Y2W7	KCNIP3	3	O60674	JAK2	2
P43080	GUCA1A	3	P98082	DAB2	2
Q9UMX6	GUCA1B	3	Q04724	TLE1	2
P17252	PRKCA	2	Q05513	PRKCZ	2
P54253	ATXN1	2	Q04771	ACVR1	2
O15198	SMAD9	2	P06730	EIF4E	2
P68400	CSNK2A1	2	Q9UKB1	FBXW11	2
P16885	PLCG2	2	Q9Y6W5	WASF2	2
P12931	SRC	2	Q13148	TARDBP	2
P21333	FLNA	2	Q8IV90	WASF4	2
P63165	SUMO1	2	O00238	BMPR1B	2
O60563	CCNT1	2	Q15797	SMAD1	1
P29353	SHC1	2	Q15834	CCDC85B	1
Q9UMX0	UBQLN1	2	Q5VTD9	GFI1B	1
Q16539	MAPK14	2	P12757	SKIL	1
P27797	CALR	2	P12956	XRCC6	1
A0JLT2	MED19	2	P35222	CTNNB1	1

Table E.1. The signaling proteins that have non-zero network crosstalk values (Continued)

Protein_ID	Name	Crosstalk Values	Protein_ID	Name	Crosstalk Values
Q09472	EP300	1	O43639	NCK2	1
P19838	NFKB1	1	Q14232	EIF2B1	1
P67870	CSNK2B	1	P05783	KRT18	1
P31946	YWHAB	1	O14908	GIPC1	1
P54259	ATN1	1	P48730	CSNK1D	1
P24941	CDK2	1	Q92769	HDAC2	1
Q93009	USP7	1	P31150	GDI1	1
P06241	FYN	1	O15169	AXIN1	1
P67775	PPP2CA	1	O15085	ARHGEF11	1
Q9NRD5	PICK1	1	Q9UHD2	TBK1	1
Q14974	KPNB1	1	Q13554	CAMK2B	1
P04626	ERBB2	1	P16220	CREB1	1
O95376	ARIH2	1	Q92793	CREBBP	1
P27348	YWHAQ	1	Q92997	DVL3	1
Q8IZP0	ABI1	1	O43353	RIPK2	1
Q13469	NFATC2	1	P17844	DDX5	1
P15531	NME1	1	P27361	MAPK3	1
P08107	HSPA1A	1	P10911	MCF2	1
Q15645	TRIP13	1	P68133	ACTA1	1
P00519	ABL1	1	P29350	PTPN6	1
Q13547	HDAC1	1	Q92796	DLG3	1
P06493	CDK1	1	P42345	MTOR	1
P42858	HTT	1	P02751	FN1	1
O14964	HGS	1	O95751	LDOC1	1
Q05586	GRIN1	1	Q15628	TRADD	1
P05556	ITGB1	1	P09471	GNAO1	1
P06239	LCK	1	P46934	NEDD4	1
P28482	MAPK1	1	P30291	WEE1	1
P25705	ATP5A1	1	P18206	VCL	1
P78527	PRKDC	1	P10644	PRKAR1A	1
P60953	CDC42	1	Q00535	CDK5	1
Q9Y4J8	DTNA	1	P15498	VAV1	1
P07948	LYN	1	Q96EK5	KIAA1279	1
P62258	YWHAE	1	Q14134	TRIM29	1
P62714	PPP2CB	1	Q06124	PTPN11	1
O14745	SLC9A3R1	1	Q9NZN5	ARHGEF12	1
O15530	PDPK1	1	Q13255	GRM1	1
P55072	VCP	1	Q92844	TANK	1

Table E.1. The signaling proteins that have non-zero network crosstalk values (Continued)

Protein_ID	Name	Crosstalk Values	Protein_ID	Name	Crosstalk Values
P17612	PRKACA	1	P52735	VAV2	1
Q16537	PPP2R5E	1	Q92481	TFAP2B	1
Q9HCE7	SMURF1	1	P16591	FER	1
P41240	CSK	1	Q9UBL3	ASH2L	1
P26010	ITGB7	1	Q14289	PTK2B	1
Q96GD4	AURKB	1	Q9P286	PAK7	1
Q13232	NME3	1	Q9Y5V3	MAGED1	1
P06213	INSR	1	P49840	GSK3A	1
Q12929	EPS8	1	P42684	ABL2	1
P50750	CDK9	1	Q8N302	AGGF1	1
Q16594	TAF9	1	Q13557	CAMK2D	1
Q16512	PKN1	1	O15357	INPPL1	1
Q14164	IKBKE	1	Q15172	PPP2R5A	1
Q9UPX8	SHANK2	1	Q13467	FZD5	1
Q12879	GRIN2A	1	P51812	RPS6KA3	1
P05129	PRKCG	1	P51948	MNAT1	1
Q08345	DDR1	1	P49757	NUMB	1
P29323	EPHB2	1	Q96FC9	DDX11	1
P39687	ANP32A	1	Q15349	RPS6KA2	1
Q16513	PKN2	1	P32320	CDA	1
Q07817	BCL2L1	1	Q99704	DOK1	1
O60861	GAS7	1	P36954	POLR2I	1
Q969Q1	TRIM63	1	Q7KZI7	MARK2	1
P29597	TYK2	1	Q86U86	PBRM1	1
Q9BZE4	GTPBP4	1	P08631	HCK	1
Q9UBE8	NLK	1	Q9NQ66	PLCB1	1
Q15173	PPP2R5B	1	Q9NP86	CABP5	1
P17676	CEBPB	1	Q9NPB3	CABP2	1
O14976	GAK	1	P57796	CABP4	1
P50613	CDK7	1	P11217	PYGM	1
P21980	TGM2	1	P00966	ASS1	1
Q06330	RBPJ	1	Q13352	ITGB3BP	1
O60760	HPGDS	1	Q9NXV2	KCTD5	1
P31751	AKT2	1	Q9Y2U5	MAP3K2	1
P49366	DHPS	1	O00429	DNM1L	1
P42261	GRIA1	1	P04040	CAT	1
Q8WZ42	TTN	1	Q01970	PLCB3	1
Q01814	ATP2B2	1	Q13470	TNK1	1

Table E.1. The signaling proteins that have non-zero network crosstalk values (Continued)

Protein_ID	Name	Crosstalk Values
Q9UNS2	COPS3	1
P48551	IFNAR2	1
P62491	RAB11A	1
Q16584	MAP3K11	1
P42680	TEC	1
P51946	CCNH	1
P48443	RXRG	1
Q15078	CDK5R1	1
Q13393	PLD1	1
Q9UBG7	RBPJL	1
P20020	ATP2B1	1

REFERENCES

Aaronson, D. S. and C. M. Horvath, 2002, "A Road Map for Those who don't Know JAK-STAT", *Science*, Vol. 296, Issue 5573, pp. 1653–5.

Aguilera, O., A. Muñoz, M. Esteller and M. F. Fraga, 2007, "Epigenetic Alterations of The Wnt/beta-catenin Pathway in Human Disease", *Endocr Metab Immune Disord Drug Targets*, Vol. 7(1), pp. 13-21.

Aguirre, A., M. E. Rubio and V. Gallo, 2010, "Notch and EGFR Pathway Interaction Regulates Neural Stem Cell Number and Self-renewal", *Nature*, Vol. 467 (7313), pp. 323–327.

Albers, J., J. Schulze, F. T. Beil, M. Gebauer, A. Baranowsky *et al.*, 2011, "Control of Bone Formation by The Serpentine Receptor Frizzled-9", *J. Cell Biol*., March, pp. 1-16.

Alj, M. B., M. I. Kontaridis and B. G. Neel, 2006, "Stops Along The RAS Pathway in Human Genetic Disease", *Nature Medicine*, Vol. 12, pp. 283 – 285.

Alterovitz G., M. Xiang, M. F. Ramoni and I. S. Kohane, 2005, "Protein Network Topology Metric Conservation: From Yeast to Human," in *RECOMB*, Cambridge, MA, (Invited) .

Arga, K. Y., Z. İ. Önsan, B. Kırdar, K. Ö. Ülgen and J. Nielsen, 2007, "Understanding Signaling in Yeast: Insights From Network Analysis", *Biotechnology and Bioengineering*, Vol. 97, No. 5, pp. 1246-1258.

Assenov, Y., F. Ramirez, S. E. Schelhorn, T. Lengaur and M. Albrecht, 2008, "Compuing Topological Parameters of Biological Networks", *Bioinformatics*,Vol. 24, pp. 282-284.

Aziz, R. K., D. Bartels, A. A. Best, M. DeJongh, T. Disz *et al*., 2008, "The RAST Server: Rapid Annotations Using Subsystems Technology", *BMC Genomics*, Vol. 9, Issue 75.

Bader, G. D. and C. W. Hogue, 2003, "An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks", *BMC Bioinformatics*, Vol. 4., No. 2, pp. 1471-2105.

Barabasi, A. L. and Z. N. Oltvai, 2004, "Network Biology: Understanding The Cell's Functional Organization", *Nature Reviews Genetics*, Vol. 5, pp. 101-113.

Baron, J. A., R. S. Sandler, S. Halabi, S. Budinger, E. Paskett *et al.*, 2003, "A Randomized Trial of Aspirin to Prevent Colorectal Adenomas", *N Engl J Med*, Vol. 348, pp. 891–9.

Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich *et al*., 2004, "The Pfam Protein Families Database", *Nucleic Acids Research*, Vol. 32, pp. D138–D141.

Bolós, V, J. Grego-Bessa and J. L. de la Pompa, 2007, "Notch Signaling in Development and Cancer", *Endocr. Rev*., Vol.28, Issue 3, pp. 339–63.

Bonni, A., A. Brunet, A. E. West, S. R. Datta, M. A. Takasu and M. E. Greenberg, 1999, "Cell Survival Promoted by the Ras-MAPK Signaling Pathway by Transcription-Dependent and -Independent Mechanisms", *Science*, Vol. 286, No. 5443, pp. 1358-1362.

Boon, E. M., J. J. Keller, T. A. Wormhoudt, F. M. Giardiello, G. J. Offerhaus, R. van der Neut, S. T. Pals, 2004, "Sulindac Targets Nuclear Beta-catenin Accumulation and Wnt Signaling in Adenomas of Patients with Familial Adenomatous Polyposis and in Human Colorectal Cancer Cell Lines", *Br J Cancer*, Vol. 90, pp. 224 –9.

Bourhis, E., C. Tam, Y. Franke, J. F. Bazan, J. Ernst, J. Hwang, M. Costa, A. G. Cochran and R. N. Hannoush, 2010, "Reconstitution of a Frizzled8.Wnt3a.LRP6 Signaling

Complex Reveals Multiple Wnt and Dkk1 Binding Sites on LRP6", *J Biol Chem*., Vol. 285, Issue 12, pp. 9172-9.

Cadigan, K. M., 2008, "Wnt–β-catenin Signaling", *Curr. Biol.*, Vol.18, pp. 943-947.

Cadigan, K. M. and Y. I. Liu, 2006, "Wnt Signaling: Complexity at The Surface", *J. Cell. Sci*., Vol. 119, pp. 395-402.

"Cell Applications, Inc.", http://www.cellapplications.com/product_desc.php?id=1020, May 2011.

Chen, A., B. H. Davis, M. Bissonnette, B. Scaglione-Sewell, T. A. Brasitus, 1999, "1,25-Dihydroxyvitamin D(3) Stimulates Activator Protein-1-dependent Caco-2 Cell Differentiation", *J. Biol. Chem.,* Vol. 274, Issue 50, pp. 35505-13.

Chomova, M. and P. Racay, 2010, "Mitochondrial Complex I in The Network of Known and Unknown Facts", *General Physiology and Biophysics*, Vol. 29, Issue 1, pp. 3–11.

Chou, C. H., W. C. Chang, C. M. Chiu, C. C. Huang and H. D. Huang, 2009, " FMM: A Web Server for Metabolic Pathway Reconstruction and Comparative Analysis", *Nucleic Acids Research,* Vol. 37, W129-134.

Cong, F., L. Schweizer and H. Varmus, 2004, "Wnt Signals Across The Plasma Membrane to Activate The Beta-catenin Pathway by Forming Oligomers Containing Its Receptors, Frizzled and LRP", *Development*, Vol. 131, pp. 5103-5115.

DasGupta, R., 2005, "Functional Genomic Analysis of the Wnt-Wingless Signaling Pathway", *Science*, Vol. 308, Issue 826.

Dihlmann, S., S. Klein and M. von Knebel Doeberitz, 2003, "Reduction of Beta-catenin/t-cell Transcription Factor Signaling by Aspirin and Indomethacin is Caused by An Increased Stabilization of Phosphorylated Beta-catenin", *Mol Cancer Ther*, Vol. 2, pp. 509 –16.

Dontu G, K. W. Jackson, E. McNicholas, M. J. Kawamura, W. M. Abdallah and M. S. Wicha, 2004, "Role of Notch Signaling in Cell-fate Determination of Human Mammary Stem/progenitor Cells", *Breast Cancer Res*., Vol. 6 , Issue 6, R605–15.

Ehebauer, M., P. Hayward, and A. Martinez-Aria, 2006, "Notch Signaling Pathway", *Sci. STKE*, Vol. 2006, Issue 364, pp. cm7.

Estrada, E., 2006, "Virtual Identification of Essential Proteins within The Protein Interaction Network of Yeast", *Proteomics*, Vol. 6, pp. 35-40.

Faux, M. C., J. L. Coates, B. Catimel, S. Cody, A. H. A. Clayton, M. J. Layton and A. W. Burgess, 2008, "Recruitment of Adenomatous Polyposis Coli and β-catenin to Axin-puncta", *Oncogene*, Vol. 27, pp. 5808–5820.

Finkel, T. and J. S. Gutkind, 2003, "Signal Transduction and Human Disease", *Wiley-Liss*, Hoboken, New Jersey, USA.

Finn, D., J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich *et al*., 2006, "Pfam: Clans, Web Tools and Services", *Nucleic Acids Research*, Vol. 34, pp. D247–D251.

Freeman, L. C., 1977, "Set of Measures of Centrality Based on Betweenness", *Sociometry,* Vol. 40, pp. 35–41.

Freeman, L. C., 1978, "Centrality in Social Networks: Conceptual Clarification", *Social Networks*, Vol. 1, pp. 215-239.

Funahashi, A., N. Tanimura, M. Morohashi and H. Kitano, 2003, "CellDesigner: Process Diagram Editor for Gene-regulatory and Biochemical Networks", *BIOSILICO*, Vol. 1, pp. 159-162.

Gaiano, N. and G. Fishell, 2002, "The Role of Notch in Promoting Glial and Neural Stem Cell Fates", *Annu. Rev. Neurosci*, Vol. 25, Issue 1, pp. 471–90.

Garber, K., 2009, “Drugging the Wnt Pathway: Problems and Progress”, *Journal of the National Cancer Institute*, Vol. 101, Issue 8, pp. 548-550.

Ge, X. and X. Wang, 2010, “Role of Wnt Canonical Pathway in Hematological Malignancies”, *Journal of Hematology & Oncology*, Vol. 3, Issue 33.

Gearhart, J., E. E. Pashos, M. K. Prasad, 2007, “Pluripotency Redeux -- Advances in Stem-cell Research”, N Engl J Med, Vol. 357(15), pp. 1469-72.

Gehrke, I., R. K. Gandhirajan and K. Kreuzer, 2009, “Targeting The Wnt/β-catenin/TCF/LEF1 Axis in Solid and Haematological Cancers: Multiplicity of Therapeutic Options”, *Eur J Cancer*, Vol. 45, Issue 16, pp. 2759-2767.

Girvan, M. and M. E. Newman, 2002, “Community Structure in Social and Biological Networks”, *Proc Natl Acad Sci USA*, Vol. 99, pp. 7821–7826.

Golemis, E. A. and P. D. Adams, 2005, “Protein-Protein Interactions: A Molecular Cloning Manual”, Second Edition, ix-xi.

Grego-Bessa, J., L. Luna-Zurita, G. del Monte, V. Bolós, P. Melgar *et al.*, 2007, "Notch Signaling is Essential for Ventricular Chamber Development", *Dev. Cell*, Vol. 12, Issue 3, pp. 415–29.

Gyuris, J., E. Golemis, H. Chertkov, R. Brent, 1993, “Cdi1, A Human G1 and S Phase Protein Phosphatase that Associates with Cdk2”, *Cell*, Vol. 75, pp. 791-803.

Hanks, S. K., A. M. Quinn and T. Hunter, 1988, “The Protein Kinase Family: Conserved Features and Deduced Phylogeny of The Catalytic Domains”, *Science*, Vol. 241, No. 4861, pp. 42-52.

Hitoshi, S., T. Alexson, V. Tropepe, D. Donoviel, A. J. Elia, J. S. Nye, R. A. Conlon, T. W. Mak, A. Bernstein, D. van der Kooy, 2002, "Notch Pathway Molecules are Essential

for The Maintenance, but not The Generation, of Mammalian Neural Stem Cells", *Genes Dev.,*Vol. 16, Issue 7, pp. 846–858.

Hopkins A. L., 2008, "Network Pharmacology: The Next Paradigm in Drug Discovery", *Nature Chemical Biology, Vol.* **4,** pp. 682 – 690.

Hsing M., G. B. Kendall and A. Cherksov, 2008, "The Use of Go Terms for Predicting Highly-connected 'Hub' Nodes in Protein-Protein Interaction Networks", *BMC Systems Biology*,Vol. 2,No. 80.

Hunter S., R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman *et al* , 2009, "InterPro: The Integrative Protein Signature Database", *Nucleic Acids Research*, Vol. 37, Issue 1, pp. D211-D215.

Hwang, W. C., A. Zhang and M. Ramanathan, 2008, "Identification of Information Flow Modulating Drug Targets: A Novel Bridging Paradigm for Drug Discovery", *Clinical Pharmacology & Therapeutics*, Vol. 84, pp. 563-572.

Jacob, L. and L. Lum, 2007, "Hedgehog Signaling Pathway", *Sci. STKE*, Vol. 2007, Issue 407, pp. cm6.

Janssens, N., M. Janicon and T. Perera, 2006, "The Wnt-dependent Signalling Pathways as Target in Oncology Drug Discovery", *Investigational New Drugs*, Vol. 24, pp. 263-280.

Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabasi, 2000, "The Large-scale Organization of Metabolic Networks", *Nature*, Vol. 407, pp. 651-654.

Jesse, S., A. Koenig, V. Ellenrieder and A. Menke, 2010, "Lef-1 Isoforms Regulate Different Target Genes and Reduce Cellular Adhesion", *Int J Cancer.*, Vol. 126, Issue 5, pp. 1109-20.

Johansen, K. A., D. D. Iwaki and J. A. Lengye, 2002, "Localized JAK/STAT Signaling is Required for Oriented Cell Rearrangement in A Tubular Epithelium", *Development*, Vol. 130, No. 1, pp. 135-145.

Kadota, M., M. Sato, B. Duncan, A. Ooshima, H. H. Yang *et al*., 2009, "Identification of Novel Gene Amplifications in Breast Cancer and Coexistence of Gene Amplification with An Activating Mutation of PIK3CA", *Cancer Res.,* Vol. 69, Issue 18, pp. 7357–7365.

Katoh, M., 2005, "WNT/PCP Signaling Pathway and Human Cancer", *Oncol Rep.,* Vol. 14, Issue 6, pp. 1583-8.

Keshava, P. T. S., R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar *et al*., 2009, "Human Protein Reference Database - 2009 Update", *Nucleic Acids Research*, Vol. 37, D767-72.

Kim, W. K., J. Park and J. K. Suh, 2002, "Large Scale Statistical Prediction of Protein-Protein Interaction by Potentially Interacting Domain (PID) Pair", *Genome Informatics*, Vol. 13, pp. 42-50.

Kirikoshi, H., H. Sekihara and M. Katoh, 2001, "Expression Profiles of 10 Members of Frizzled Gene Family in Human Gastric Cancer", *Int J Oncol.,* Vol. 19, Issue 4 pp. 767-71.

Kitano, H., A. Funahashi, Y. Matsuoka and K. Oda, 2005, "Using Process Diagrams for The Graphical Representation of Biological Networks", *Nat. Biotechnol*., Vol. 23, pp. 961-966.

Klaus, A. and W. Birchmeier, 2008, "Wnt Signaling and Its Impact on Development and Cancer", *Nat Rev Cancer*, Vol. 8, pp. 378-398.

Klopocki, E., G. Kristiansen, P. J. Wild, I. Klaman, E. Castanos-Velez *et al.*, 2004, "Loss of SFRP1 is Associated with Breast Cancer Progression and Poor Prognosis in Early Stage Tumors", *Int J Oncol.* , Vol. 25(3), pp. 641-9.

Kohn, A. D. and R. T. Moon, 2005, "Wnt and Calcium Signaling: Beta-catenin-independent Pathways", *Cell Calcium*, Vol. 38(3-4), pp. 439-46.

Kuhl, M., 2004, "The Wnt/Calcium Pathway: Biochemical Mediators, Tools and Future Requirements", *Frontiers in Bioscience*, Vol. 9, pp. 967-974.

Kuhl, M., L. C. Sheldahl, M. Park, J. R. Miller and R. T. Moon, 2000, "The Wnt/Ca^{2+} Pathway: A New Vertebrate Wnt Signaling Pathway Takes Shape", *Trends Genet*, Vol. 16, pp. 279-283.

Lee, S.W., C. L. Reimer, L. Fang, M. L. Iruela-Arispe., S. A. Aaronson, 2000, "Overexpression of Kinase-associated Phosphatase (KAP) in Breast and Prostate Cancer and Inhibition of The Transformed Phenotype by Antisense KAP Expression", Mol. Cell. Biol., Vol. 20, pp. 1723-1732.

Li, B., Y. Liu, T. Uno and N. Gray, 2004, "Creating Chemical Diversity to Target Protein Kinases", *Comb Chem High Throughput Screen*, Vol. 7, Issue 5, pp. 453-72.

Li, G., H. Yin and J. Kuret, 2004, "Casein Kinase 1 Delta Phosphorylates Tau and Disrupts Its Binding to Microtubules", *Journal of Biological Chemistry*, Vol. 279, pp. 15938-15945.

Liu, S., G. Dontu and M. S. Wicha, 2005, "Mammary Stem Cells, Self-renewal Pathways, and Carcinogenesis", *Breast Cancer Research*, Vol. 7, pp. 86-95.

Lu, S.-L., M. Kawabata, T. Imamura, Y. Akiyama, T. Nomizu, K. Miyazono, Y. Yuasa, 1998, "HNPCC Associated with Germline Mutation in The TGF-beta Type II Receptor Gene", *Nature Genet*,, Vol. 19, pp. 17-18.

Lu, L. J., Y. Xia, A. Paccanaro, H. Yu and M. Gerstein, 2005, "Assessing The Limits of Genomic Data Integration for Predicting Protein Networks", *Genome Res.*, Vol. 15, pp. 945-953.

Luo, J., J. Chen, Z. L. Deng, X. Luo, W. X. Song, K. A. Sharff, N. Tang, R. C. Haydon, H. H. Luu and T. C. He, 2007, "Wnt Signaling and Human Diseases: What Are The Therapeutic Implications? ", *Laboratory Investigation*, Vol. 87, pp. 97-103.

Luu, H. H., R. Zhang, R.C. Haydon, E. Rayburn, Q. Kang *et al*., 2004, "Wnt/beta-catenin Signalling Pathway as A Novel Cancer Drug Target", Curr Cancer Drug Targets, Vol. 4, pp. 653–671.

Maiti, S., R. Alam, C. I. Amos, V. Huff, 2000, "Frequent Association of beta-Catenin and WNT1 Mutations in Wilms Tumors", *Cancer Res*, Vol. 60, pp. 6288-6292.

Manning, G., G. D. Plowman, T. Hunter and S. Sudarsanam, 2002, "Evolution of Protein Kinase Signaling From Yeast to Man", Trends Biochem Sci.,Vol. 27, Issue 10, pp. 514-20.

Marengere, L. E. and T. Pawson, 1994, "Structure and Function of SH2 Domains", *J Cell Sci Suppl.*, Vol. 18, pp. 97-104.

Matsuzawa, S. I., J. C. Reed, 2001, "Siah-1, SIP, and Ebi Collaborate in A Novel Pathway for Beta-catenin Degradation Linked to p53 Responses", *Mol Cell*., Vol. 7(5), pp. 915-26.

Mayer, B. J. and D. Baltimore, 1993, "Signalling Through SH2 and SH3 Domains", *Trends Cell Biol.*, Vol. 3, Issue 1, pp. 8-13.

Miyaoka, T., H. Seno and H. Ishino, 1999, "Increased Expression of Wnt-1 in Schizophrenic Brains", *Schizophr Res*, Vol. 38, pp.1–6.

Moon, R. T., 2005, "Wnt/beta-Catenin Pathway", *Sci. STKE*, Vol. 2005, Issue 271, pp. cm1.

Moriya, Y., M. Itoh, S. Okuda, A. C. Yoshizawa and M. Kanehisa, 2007, "KAAS: An Automated Genome Annotation and Pathway Reconstruction Server", *Nucleic Acids Research*, Vol. 35, W 182-185.

Moriya, Y., D. Shigemizu, M. Hattori, T. Tokimatsu, M. Kotera, S. Goto, M. Kanehisa M, 2010, "PathPred: An Enzyme-catalyzed Metabolic Pathway Prediction Server", Nucleic Acids Research, Vol. 38, W138-143.

Müller, B. and J. Sheen, 2007, "Advances in Cytokinin Signaling", *Science*, Vol. 318, No. 5847, pp. 68-69.

Nandy, S. K., J. Nielsen and P. Jouhten, 2010, "Reconstruction of The Yeast Protein-Protein Interaction Network Involved in Nutrient Sensing and Global Metabolic Regulation", *BMC Systems Biology*, Vol. 4, Issue 68.

"National Center for Biotechnology Information", http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?db=human&c=Gene&l=PRKCA, April 2011.

"National Center for Biotechnology Information", http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=5578, May 2011.

"National Cancer Institute", http://www.cancer.gov/clinicaltrials, May 2011.

Nguyen, T. P. and F. Jordán, 2010, "A Quantitative Approach to Study Indirect Effects Among Disease Proteins in The Human Protein Interaction Network", *BMC Syst Biol*., Vol. 4, Issue 103.

Nguyen, T. P. and T. B. Ho, 2006, "Discovering Signal Transduction Networks Using Signaling Domain-Domain Interactions", *Genome Informatics*, Vol. 17, Issue 2, pp. 35-45.

Nobta, M., T. Tsukazaki, Y. Shibata, C. Xin, T. Moriishi, S. Sakano, H. Shindo and A. Yamaguchi, 2005, "Critical Regulation of Bone Morphogenetic Protein-Induced

Osteoblastic Differentiation by Delta1/Jagged1-activated Notch1 Signaling", *J. Biol. Chem.*,Vol. 280, Issue 16, pp. 15842–8.

Nusse, R., 2005, "Wnt Signaling in Disease and in Development", *Cell Research*, Vol. 15, pp. 28-32.

Olsen, J. V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen and M. Mann, 2006, "Global, in Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks", *Cell*, Vol. 127, Issue 3, pp. 635-48.

"OMIM, Online Mendelian Inheritance in Man", http://www.ncbi.nlm.nih.gov/omim, April 2011.

Opsahl, T., F. Agneessens and J. Skvoretz, 2010, "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths", *Social Networks*, Vol. 32, pp. 245-251.

Ozbayraktar, B. K., 2011, "From Yeast to Human: Unraveling Sphingolipid Metabolism through Macroscopic and Microscopic Analyses", PhD Thesis, Boğaziçi Üniversitesi, İstanbul.

Parker, H., Q. An, K. Barber, M. Case, T. Davies *et al*., 2008, "The Complex Genomic Profile of ETV6-RUNX1 Positive Acute Lymphoblastic Leukemia Highlights A Recurrent Deletion of TBL1XR1", *Genes, Chromosomes and Cancer*, Vol. 47, Issue 12, pp. 1118–1125.

Patil, A. and H. Namakura, 2005, "Filtering High-throughput Protein-Protein Interaction Data Using A Combination of Genomic Features", *BMC Bioinformatics*, Vol. 6, Issue. 100.

"Pathway Commons", http://www.pathwaycommons.org/pc/record2.do?id=92132, April 2011.

Paul S. and A. Dey, 2008, "Wnt Signaling and Cancer Development: Therapeutic Implication", *Neoplasma*, Vol. 55, pp. 165-176.

Pawson, T. and J. Schlessingert, 1993, "SH2 and SH3 domains", *Curr Biol*., Vol. 3, Issue 7, pp. 434-42.

Pawson, T., 1995, "Protein Modules and Signalling Networks", *Nature*, Vol. 373, No. 6515, pp. 573-80.

Peng, Z., H. Wang and C. Shan, 2009, "Expression of Ubiquitin and Cullin-1 and Its Clinicopathological Significance in Benign and Malignant Lesions of The Lung*", Zhong Nan Da Xue Xue Bao Yi Xue Ban*., Vol. 34, Issue 3, pp. 204-9.

Persico, M., A. Ceol, C. Gavrila, R. Hoffmann, A. Florio and G. Cesareni, 2005, "HomoMINT: An Inferred Human Network Based on Orthology Mapping of Protein Interactions Discovered in Model Organisms", *BMC Bioinformatics*,Vol. 6, (Suppl. 4): S21.

Petrussa, E., A. Bertolini, V. Casolo, J. Krajnáková, F. Macrì, A. Vianello A, 2009, "Mitochondrial Bioenergetics Linked to The Manifestation of Programmed Cell Death During Somatic Embryogenesis of Abies Alba", *Planta*, Vol. 231, Issue 1, pp. 93–107.

"PFAM-Database of Protein Families", http://pfam.sanger.ac.uk/, March 2011.

Ralhan, R., L. V. De Souza, A. Matta, S. C. Tripathi, S. Ghanny, S. D. Gupta, S. Bahadur and K. W. M. Siu, 2008, "Discovery and Verification of Head-and-neck Cancer Biomarkers by Differential Protein Expression Analysis Using iTRAQ Labeling, Multidimensional Liquid Chromatography, and Tandem Mass Spectrometry", *Molecular and Cellular Proteomics*, Vol. 7, pp. 1162-1173.

Rice, P. L., J. Kelloff, H. Sullivan, L. J. Driggers, K. S. Beard, S. Kuwada, G. Piazza, D. J. Ahnen, 2003, "Sulindac Metabolites Induce Caspase- and Proteasome-dependent Degradation of Beta-catenin Protein in Human Colon Cancer Cells", *Mol Cancer Ther*, Vol. 2, pp. 885–92.

Rost, B., 2002, "Enzyme function Less Conserved than Anticipated", *J. Mol. Biol*., Vol. 318, pp. 595-608.

Saarela, J., S. P. Kallio, D. Chen, A. Montpetit, A. Jokiaho *et al*., 2006, "PRKCA and Multiple Sclerosis: Association in Two Independent Populations", *PLoS Genetics,* Vol. 2, Issue 3, e 42.

Sandler, R. S., S. Halabi, J. A. Baron, S. Budinger, E. Paskett *et al*., 2003, "A Randomized Trial of Aspirin to Prevent Colorectal Adenomas in Patients with Previous Colorectal Cancer", *N Engl J Med*, Vol 348, pp. 883–90.

Satoh, S., Y. Daigo, Y. Furukawa, T. Kato, N. Miwa *et al*., 2000, "AXIN1 Mutations in Hepatocellular Carcinomas, and Growth Suppression in Cancer Cells by Virus-Mediated Transfer of AXIN1", *Nat Genet*, Vol. 24, pp. 245–250.

Schneider, J., F. J. Vorhölter, E. Trost, J. Blom, Y. R. Musa, H. Neuweger, K. Niehaus, S. Schatschneider, A. Tauch, A. Goesmann, 2010, "CARMEN- Comparative Analysis and *in silico* Reconstruction of Organism-specific Metabolic Networks", *Genet. Mol. Res*., Vol. 9, Issue 3, pp. 1660-1672.

Schwartz, G. K., J. Jiang, D. Kelsen and A. P. Albino, 1993, "Protein kinase C: A Novel Target for Inhibiting Gastric Cancer Cell Invasion", *J. Natl. Cancer Inst.,* Vol. 85, Issue 5, pp. 402-7.

Slusarskia, D. C., J. Yang-Snyderb, W. B. Busaa and R. T. Moon, 1997, "Modulation of Embryonic Intracellular Ca^{2+} Signaling by Wnt-5A", *Developmental Biology*, vol. 182, pp. 114-120.

Soucek, L., J. Whitfield, C. P. Martins, A. J. Finch, D. J. Murphy, N. M. Sodir, A. N. Karnezis, L. B. Swigart, S. Nasi, G. I. Evan, 2008, "Modelling Myc Inhibition As A Cancer Therapy", *Nature,*Vol. 455 (7213), pp. 679–83.

Stark, C., B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz and M. Tyers, 2006, "BioGRID: A General Repository for Interaction Datasets", *Nucleic Acids Res*, Vol.34, pp. D535-539.

Steffen, M., A. Petti, J. Aach, P. D'haeseleer and G. Church, 2002, "Automated Modelling of Signal Transduction Networks", *BMC Bioinformatics*, Vol. 3, No. 34, pp. 1471-2105.

Stein A., A. Céol and P. Aloy, 2011, "3did: Identification and Classification of Domain-based Interactions of Known Three-dimensional Structure", *Nucleic Acids Res*., Vol. 39, pp. D718-723.

Tan, S. G. and K. T. Kelsey, 2009, "Epigenetic Inactivation of SFRP1 and SFRP2 Genes as Biomarkers of Invasive Bladder Cancer", *Honors Thesis*, Brown University.

Tanaka, S., M. Mori, K. Mafune, S. Ohno, K. Sugimachi, 2000, "A Dominant Negative Mutation of Transforming Growth Factor-beta Receptor Type II Gene in Microsatellite Stable Oesophageal Carcinoma", *Brit. J. Cancer*, Vol. 82, pp. 1557-1560.

Tekir S. D., K. Y. Arga and K. Ö. Ülgen, 2009, "Drug Targets for Tumorigenesis: Insights from Structural Analysis of EGFR Signaling Network", *Journal of Biomedical Informatics*, Vol. 42, pp. 228–236.

Thatcher, J. D., 2010, "The Ras-MAPK Signal Transduction Pathway", *Sci. Signal.*, Vol. 3, Issue 119, tr1.

Tian, E., F. Zhan, R. Walker, E. Rasmussen, Y. Ma, B. Barlogie, J. D. Shaughnessy, 2003, "The Role of The Wnt-signaling Antagonist DKK1 in The Development of Osteolytic Lesions in Multiple Myeloma", *N Engl J Med*, Vol. 349, pp. 2483–2494.

Tian, W. and J. Skolnick, 2003, "How well is Enzyme Fuction and Conserved of Pairwise Sequence Identity? ", *J. Mol. Biol*,Vol. 333, pp. 863-882.

Tiveci, S., 2011, "Ca Signaling in Multiple Species", PhD Thesis, Boğaziçi Üniversitesi, İstanbul.

Toku, A. E., 2010, "Investigation of Crosstalk between Wnt and Hedgehog Signaling Pathways in Fly", MSc Thesis, Boğaziçi Üniversitesi, İstanbul.

Ueno, K., M. Hiura, Y. Suehiro, S. Hazama, H. Hirata, M. Oka, K. Imai, R. Dahiya, Y. Hinoda, 2008, "Frizzled-7 as a Potential Therapeutic Target in Colorectal Cancer", *Neoplasia*, Vol. 10, pp. 697–705.

"UNIPROT (Universal Protein Resource)", http://www.uniprot.org/, May 2010.

Ümit, P. and K. Ö. Ülgen, 2009, "Reconstruction of Insulin Signaling Network", ICSB09, San Francisco, USA.

Valdes, A. M., T. D. Spector, A. Tamm, K. Kisand, S. A. Doherty *et al.*, 2010, "Genetic Variation in The SMAD3 Gene is Associated with Hip and Knee Osteoarthritis", *Arthritis Rheum.*, Vol. 66 Issue 8, pp. 2347-52.

Varki, A., R. D. Cummings, J. D. Esko, H. H. Freeze HH, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler (editors), 2009, *Source Essentials of Glycobiology,* Chapter 12 Second edition, Cold Spring Harbor Laboratory Press, New York, USA.

Wajant, H., 2003, "The FasL-Fas Signaling Nework", *Sci. STKE*, Vol. 2003, Issue 194, pp. cm5.

Wasserman, S. and K. Faust, 1994, "Social Network Analysis: Methods and Applications", Cambridge University Press, New York, NY.

Watts, D. J. and S. H. Strogatz, 1998, "Collective Dynamics of 'Small-world' Networks", Nature, Vol. 393, pp. 440-2.

Widelitz, R., 2005, "Wnt Signaling Through Canonical and Non-canonical Pathways: Recent Progress", *Growth Factors*, Vol. 23, pp. 111-116.

Winn, R. A., L. Marek and S. Y. Han, 2005, "Restoration of Wnt-7a Expression Reverses Non-small Cell Lung Cancer Cellular Transformation through Frizzled-9-mediated

Growth Inhibition and Promotion of Cell Differentiation", *The Journal of Biological Chemistry*, Vol. 280, No. 20, Issue of May 20, pp. 19625–19634.

Wu, X. R., Y. Zhu, Y. Li, 2005, "Analyzing Protein Interaction Networks via Random Graph Model", *International Journal of Information technology*, Vol. 12, pp. 1251-1260.

Yanaga F. T. and T. Sasaguri, 2007, "The Wnt/β-Catenin Signaling Pathway as a Target in Drug Discovery", *J Pharmacol Sci*., Vol. 104, pp. 293 – 302.

Yang, Z. Q., Y. Zhao, Y. Liu, J. Y. Zhang, S. Zhang *et al*., 2010, "Downregulation of HDPR1 is Associated with Poor Prognosis and Affects Expression Levels of p120-catenin and β-catenin in Nonsmall Cell Lung Cancer", *Molecular Carcinogenesis*, Vol. 49, Issue 5, pp. 508–519.

Yeh, C. T., S. C. Lu, T. C. Chen, C. Y. Peng, Y. F. Liaw, 2003, "Aberrant Transcripts of The Cyclin-dependent Kinase-associated Protein Phosphatase in Hepatocellular Carcinoma", Cancer Res., Vol. 60, pp. 4697-4700.

Yu, H., D. Greenbaum, H. X. Lu, X. Zhu and M. Gerstein, 2004, "Genomic Analysis of Essentiality within Protein Networks", *Trends Genet*.,Vol. 20,Issue 6, pp. 227–231.

Yu, H., P. M. Kim, E. Sprecher, V. Trifonov and M. Gerstein, 2007, "The Importance of Bottlenecks in Protein Networks: Correlation with Gene Essentiality and Expression Dynamics", PLoS Computational Biology, Vol. 3, Issue 4, e 59.

Zhang Q., R. H. Goodman and D.W. Piston, 2002, "Regulation of Corepressor Function by Nuclear NADH", *Science* , Vol. 295 no. 5561 pp. 1895-1897.

Zhou, L., N. An, R. C. Haydon, Q. Zhou, H. Cheng *et al*., 2003, "Tyrosine Kinase Inhibitor STI-571/Gleevec Down-regulates The Beta-catenin Signalling Activity", Cancer Lett., Vol. 193, pp. 161–70.

Zielinski R., P. F. Przytycki, J. Zheng, D. Zhang, T. M. Przytycka and J. Capala, 2009, “The Crosstalk Between EGF, IGF, and Insulin Cell Signaling Pathways-Computational and Experimental Analysis”, BMC Bioinformatics, Vol. 3, No. 88.