We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

13

Data Mining Pubmed Using Natural Language Processing to Generate the β-Catenin Biological Association Network

Fengming Lan, Xiao Yue, Lei Han, Peiyu Pu and Chunsheng Kang
Department of Neurosurgery, Laboratory of Neuro-Oncology,
Tianjin Medical University General Hospital,
China

1. Introduction

β-catenin, originally identified in Drosophila as the segment polarity protein armadillo, is a multifunctional protein that is encoded in humans by the *CTNNB1* gene. β-catenin is found in at least three cellular pools: (i) at the adherens junctions, where β-catenin binds to the cytoplasmic domain of type I cadherins and modulates cadherin-dependent cell-cell adhesion by linking the cadherin/catenin complex to the cortical actin cytoskeleton through the binding of a-catenin; (ii) the cytoplasm, where β-catenin plays a critical role in the canonical Wnt signaling cascade by interacting with APC and GSK3β linked destruction complex, leading to its ubiquitination and subsequent degradation by the proteasome; and (iii) the nucleus, in association with other transcription factors. The crucial event in the canonical Wnt signalling cascade is the cytoplasmic stabilization of β–catenin, leading to its subsequent nuclear localization and gene transcription activity (Liu & Millar 2010). To date, numerous β-catenin target genes have been identified in diverse biological systems (http://www.stanford.edu/%7ernusse/wntwindow.html), yet little is understood about β–catenin outside of its roles in Wnt and cadherin signaling.

Biomedical literature is growing at a double-exponential pace, with more than 19,000,000 publications in MEDLINE of which more than three million were published in the last 5 years alone (Abrams et al., 1998). Over the last 10 years, the total size of MEDLINE (the database searched by PubMed) has grown at a ~4.1% compounded annual growth rate, and the number of new entries in MEDLINE each year has grown at a compounded annual growth rate of ~3.1% (Albert, 1999, 2002, 2005). Thus, a massive wealth of information is embedded in the literature and waiting to be discovered and extracted. Literature mining is a promising strategy to utilize this untapped information for knowledge discovery. Most mining is performed on the abstracts of biomedical articles, which represent a readily available resource of highly concentrated information and result in high quality extracted relations (Apte & Weiss, 1997; Card et al., 1996; Chien et al., 2007) . Text mining of biomedical literature has been applied successfully to various biological problems including the discovery and characterization of molecular interactions (protein-protein , gene-protein, gene-drug, protein sorting and molecular binding, consolidating information into a more accessible form (Babu et al., 2004; Balaji et al., 2006; Giot et al., 2003; Lee et al., 2006).

Recently, considerable interest and effort has been focused on the construction and analysis of genome-wide gene networks (McCraith et al., 2000). The task is complicated for heavily investigated transcription factors such as β-catenin due to the large volume of manuscripts published. As no searchable records are available to efficiently retrieve information relevant to the β-catenin gene network, we extracted gene/protein interactions by text mining Pubmed abstracts and constructed the β-catenin biologic association network. Our text-mining by natural language processing established an association of the β-catenin network with survival signaling, clarifying the fragmentary data that was previously available describing this relationship and confirming the crucial role of β-catenin in growth and development.

2. Methods

2.1 Natural Language Processing (NLP)

Medline/PubMed was used as the information source for bioinformatics text mining. Medline abstracts were retrieved using National Center for Biotechnology Information (NCBI) PubMed portal. We queried Pubmed with: (catenin OR CTNNB OR CTNNB1) AND ("1980/01/01"[PDAT]: "2009/05/24"[PDAT]). All abstracts were downloaded as HTML text without images and converted into XML documents. Sentence tokenization was performed with Lingpipe tools. Subsequent analysis was based on the sentence as the basic unit. Gene mentions (including the β-catenin gene) were tagged using ABNER (Egghe & Rousseau, 1990). To solve the matter of the plethora of gene aliases, all gene mentions were normalized to Entrez gene (http://www.ncbi.nlm.nih.gov/Entrez/) official gene symbols. A genetic interaction of the verb dictionary was established from BioNLP item (http://bionlp.sourceforge.net/), containing verbs such as repress, regulate, inhibit, interact, phosphorylate, downregulate, upregulate and all other verbs and their variants. Verbs in abstracts were tagged using Lingpipe and the interaction verb dictionary (Ghannad-Rezaie et al., 2006). Only sentences with the β-catenin gene, a proper interaction verb and another gene were selected. In order to test the null hypothesis 'the relationship between β-catenin and another gene is random', the hypergeometric distribution test was employed (Kim et al., 1997).

N represents the total number of PubMed abstracts and m and n represent the number gene mentions in PubMed for β-catenin and a related gene, respectively.

$$p=1-\sum_{i=0}^{k-1} p(i \mid n,m,N)$$

Where:

$$p(i \mid n,m,N)= \frac{n!(N-n)!m!(N-m)!}{(n-i)!i!(n-M)!(N-n-M+i)!N!}$$

The β-catenin-gene' relations with p-value<0.05 were then summarized and subjected to a relational database for further analysis. The flowchart of our NLP pipeline is shown as Figure 1.

2.2 Gene ontology analysis

Gene ontology analysis was performed using the GSEABase package of BioConductor (http://www.bioconductor.org/). A gene set enrichment analysis was performed on

the 543 β-catenin-related genes based on the gene ontology (GO) categories (Rual et al., 2005).

2.3 Pathway and gene network analysis

Expression Analysis Systematic Explorer (EASE) was used to analyze KEGG pathways. Over representation of genes in a KEGG pathway was present if a larger fraction of genes within that pathway was differentially expressed compared to all other genes in the genome. The 'β-catenin-verb-gene' relationships retrieved by our NLP system were filtered by pathway enrichment analysis. The links between β-catenin and related genes were visualized using Cytoscape software (Lopez & Blobel, 2008) (http://www.cytoscape.org/). Genes were grouped according to pathway. Genes that are involved in multiple pathways were assigned to a single pathway with the smallest enrichment p-value. Integrating PubMed text mining, homology prediction, gene neighbor, protein-protein interaction, gene fusion and other data sources through the Search Tool for the Retrieval of Interacting Genes/Proteins(STRING) , we created the β-catenin related genes knowledge-driven network (Sousa et al., 2004; Uetz et al., 2000).

3. Experimental results

3.1 Identification of β-catenin interaction genes

Query of β-catenin on Pubmed with: (catenin OR CTNNB OR CTNNB1) AND ("1980/01/01"[PDAT]: "2009/05/24"[PDAT]) led to the identification of 10018 articles describing putative interactions between β-catenin and other genes. Titles containing the β-catenin gene along with a proper interaction verb and another gene were selected for further analysis. A total of 543 genes with published interaction with β-catenin were identified (scheme describing the NLP process, Fig. 1; visualization of the 543 genes Fig. 2).

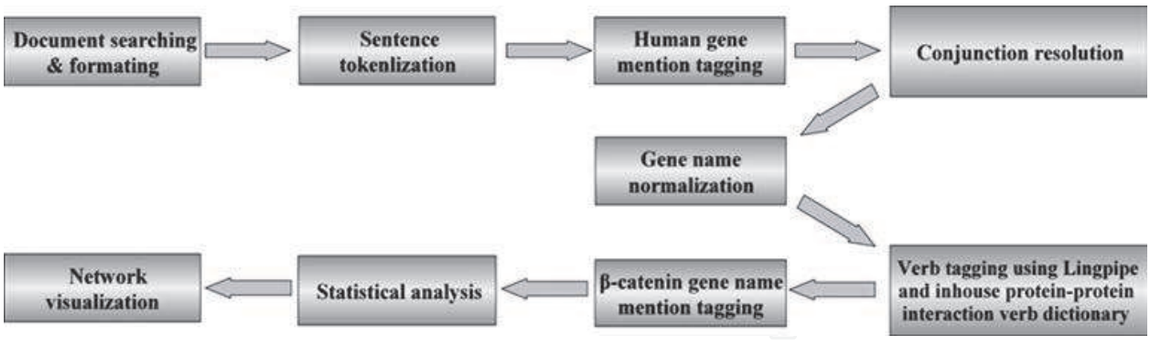


Fig. 1. Schematic representation of the literature-based gene network analysis. Literature mining by natural language processing was performed on all Pubmed abstracts available from 1/1/1980-5/24/2009. Mining identified 10018 articles describing putative interactions between β-catenin and another gene product, revealing 543 distinct β-catenin interacting proteins.

Hepatocyte nuclear factor 4 alpha (HNF4A) was the most prevalent gene identified, commonly referred to as a β-catenin "target" in the literature. Table 1 provides a list of the 10 most frequently published interactions with β-catenin, and their putative relationship.

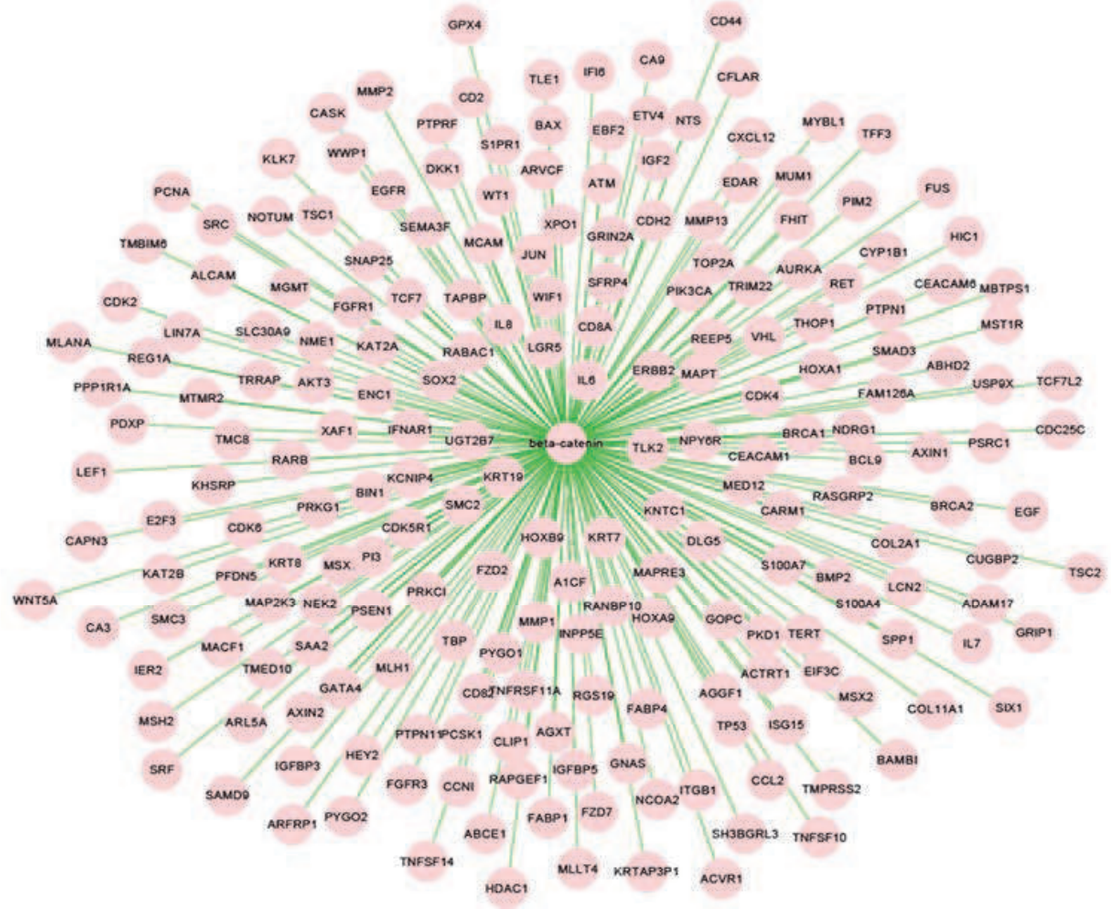


Fig. 2. Visualization of β-catenin interacting proteins. Data mining revealed 543 β-catenin interacting proteins, visualized using the Cytoscape software as described in Methods (http://www.cytoscape.org/).

Gene sym	Pubmed count	Putative interaction
HNF4A	121	target
APC	62	regulate
LEF1	39	activate
EGFR	30	associate
MAPK8	28	activate
LRP6	25	phosphorylate
MUC1	25	interact
IQGAP1	24	interact
CTNNBIP1	21	inhibit
DKK1	21	associate
CD44	20	associate

Table 1. Description of the 10 highest published interacting partners with β-catenin. The gene symbol, number of hits on Pubmed and putative interaction of the 10 highest frequency hits among the 543 interacting proteins identified by literature mining, as described in Methods.

Among the 543 gene products, 18 distinct putative protein-protein relationships were identified involving β-catenin, with the distribution of the seven most frequent relationships
included in Fig. 3. The most common relationship, complexing with β-catenin, was identified for 213 (39.2%) of the gene products (Fig. 3), including TCF4/TCF7L2 and LEF/LEF1. 54 (9.5%) gene products were identified as β-catenin targets, including cell cycle regulating proteins cyclinD1 and CDC42, proteins that influence cellular migratory behavior including uPA, Timp3, and CD44 and proteins that play a role in differentiation such as BMP-7, FGF8 and PPAR-d.

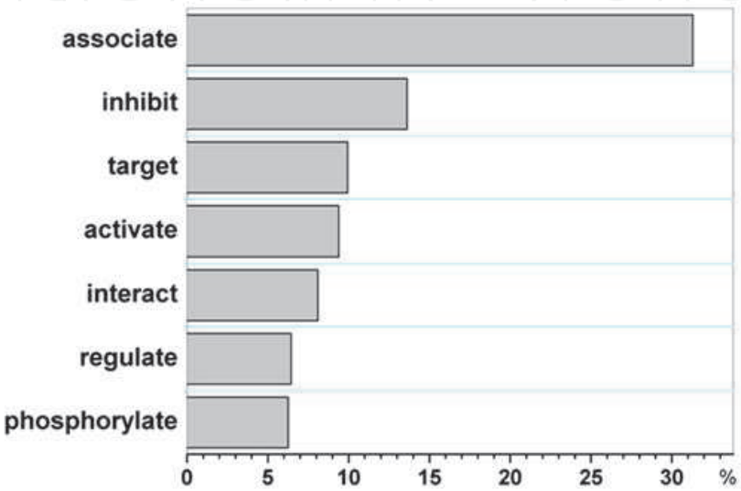


Fig. 3. Frequency distribution of protein 'relationships' with β-catenin. Literature mining revealed 18 types of relationships exhibited by β-catenin with the 543 identified gene products. Relationship type is expressed as the percent of total proteins analyzed. Analysis revealed that the most frequent relationship is "associate" (31.31%), while the least frequent relationship, occurring for only one associated protein, is "dephosphorylate".

3.2 β-catenin-related gene function

To better understand the biological role of the 543 β-catenin-related genes identified, and demonstrate the complexity of the β-catenin-related genes interaction network, we performed a Gene Ontology (GO) enrichment analysis.
GO provides structured, controlled ontologies for describing gene products in terms of their associated molecular function, biological process, or cellular compartment. Enrichment for molecular function revealed that most β-catenin-associated genes, including KIT, HSF1, XPO1, HSPA5, FGF8 and ALCAM, encode proteins that bind to β-catenin. Genes such as CAPN3, EPHA7, PTGER4, SRC, BMP4 composed the second largest category, encoding for proteins that act as signal transducers (Fig. 4, left panel). Enrichment for biological process revealed that the most common functions of gene products associated with β-catenin include developmental processes, cell communication and signaling transduction (Fig. 4, middle panel). Finally, enrichment for the cellular compartments where β-catenin associated gene products can be found primarily included the cytoplasm, nucleus and plasma membrane (Fig. 4, right panel).

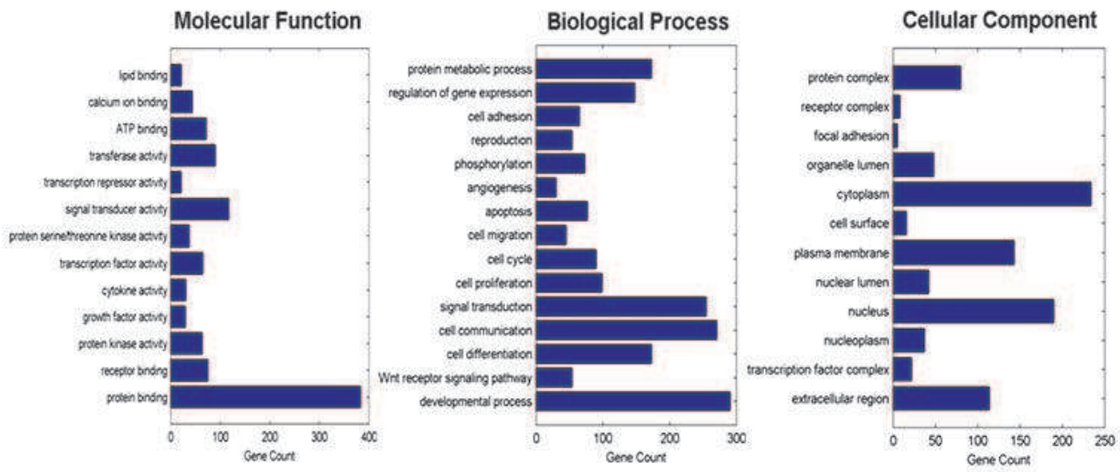


Fig. 4. Gene Ontology (GO) analysis of the β-catenin molecular network. GO enrichment sorted all data by Molecular Function, Biological Process and Cellular Compartment. GO analysis revealed that protein binding is the most prevalent molecular function of β-catenin interacting proteins, development, cell communication and signal transduction are the most common biological processes involved and the gene products are active primarily in cytoplasm, nucleus and plasma membrane.

3.3 Pathway and gene network analysis

Pathway information is required for understanding of gene function.For each of the 543 genes identified, we searched the Kyoto Encyclopedia of Genes and Genomes (KEGG)

Term	Count	%	P-Value
Wnt signaling pathway	54	10.06%	6.85E-26
Focal adhesion	36	6.70%	8.79E-08
MAPK signaling pathway	40	7.45%	1.07E-06
Adherens junction	21	3.91%	4.61E-08
p53 signaling pathway	18	3.35%	1.63E-06
ErbB signaling pathway	19	3.54%	1.03E-05
Apoptosis	17	3.17%	1.25E-04
Insulin signaling pathway	22	4.10%	2.18E-04
VEGF signaling pathway	15	2.79%	2.27E-04
Toll-like receptor signaling pathway	18	3.35%	4.18E-04
GnRH signaling pathway	17	3.17%	4.83E-04
Cell cycle	19	3.54%	5.61E-04
mTOR signaling pathway	11	2.05%	0.001869395
TGF-ß signaling pathway	14	2.61%	0.007074483

Table 2. Pathway analysis of the 543 β-catenin interacting proteins. β-catenin interacting proteins are involved in 14 different cell signaling pathways, identified using the Cytoscape software following NLP analysis.

database to identify their pathway information. 14 signaling pathways where identified whose corrected P-value was less than 0.01 (Table 2), and 321 of the genes belonged to these 14 pathways. Pathway analysis was visualized using the Cytoscape software (Fig. 5). The Using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), we created a β-catenin related genes network (Fig. 6). STRING incorporates known and predicted protein interaction information from HPRD, BioGrid, MINT, BIND, DIP, and imports known reactions from Reactome and KEGG pathways to generate a generalized source of protein interaction information. STRING analysis of β-catenin related genes revealed several hub genes, or genes in which high connection exists giving these genes an influential role in network stability. Hub genes identified include AKT1, CCND1, CTNNB1, JUN, TP53 and VEGFA (Fig. 7).

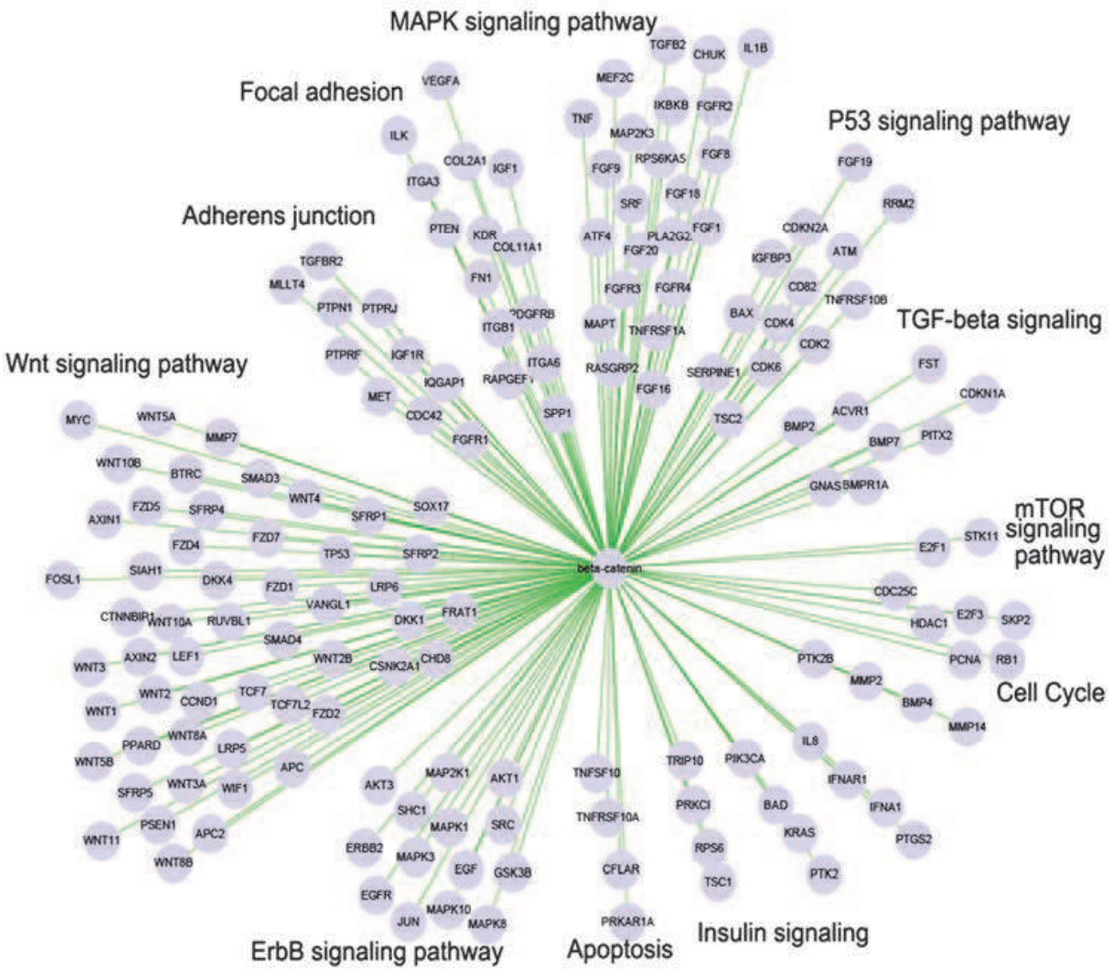


Fig. 5. Visualization of pathway distribution of the 543 β-catenin interacting proteins. . Pathway analysis of the 543 β-catenin interacting proteins was performed following NLP analysis and visualized using the Cytoscape software, as described in Methods. Interacting proteins fit into 14 different cell signaling pathways, including several pathways involved in cell survival signaling.

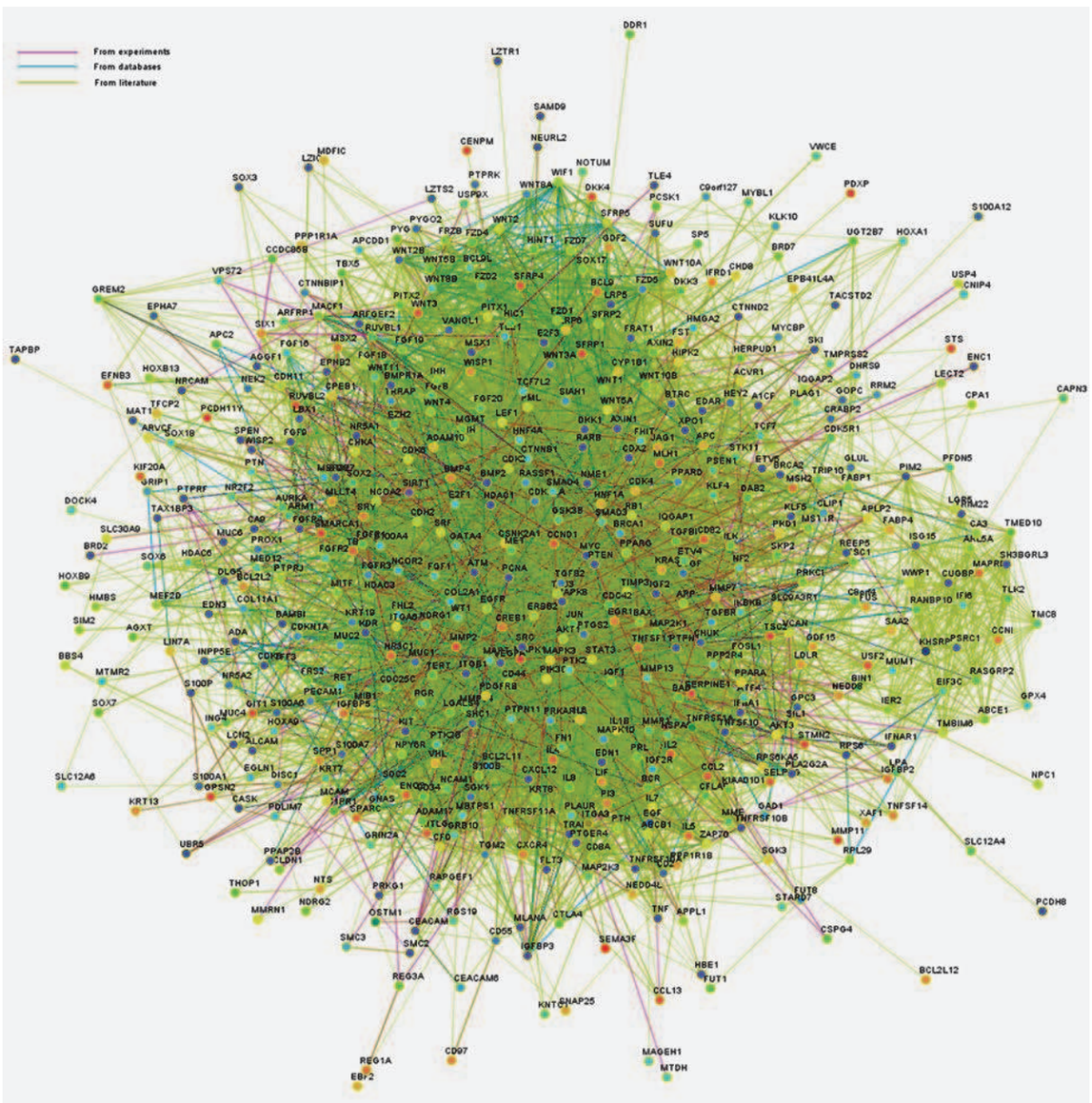


Fig. 6. Visualization of the β-catenin related genes network. Network analysis was performed using the Expression Analysis Systematic Explorer (EASE) to analyze KEGG pathways and Search Tool for the Retrieval of Interacting Genes/Proteins(STRING). Visualization was performed using Cytoscape. Pink lines indicate connections experimentally confirmed by other researches, Cyan lines indicate connections derived from databases (including the KEGG pathway and MIPS) and Green lines indicate connections compiled from co-citation data from literature mining PubMed abstracts.

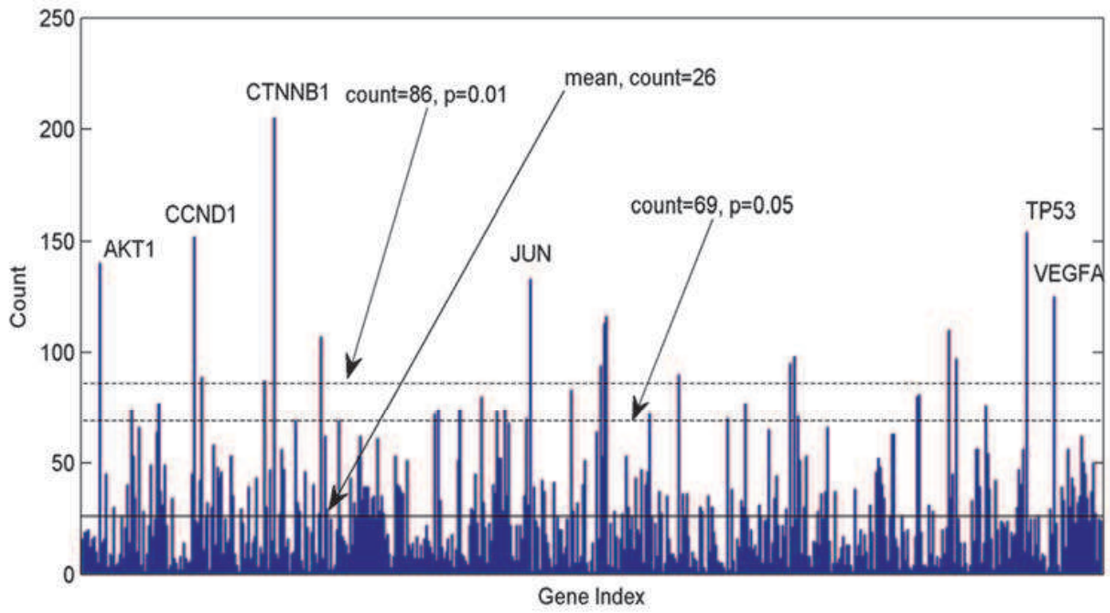


Fig. 7. Connectivity analysis of the β-catenin related genes network. Connectivity analysis was performed using the Search Tool for the Retrieval of Interacting Genes/Proteins(STRING) to generate the β-catenin related genes knowledge-driven network, as described in Methods. Analysis revealed AKT1, CCND1, CTNNB1, JUN, TP53 and VEGFA are important hub genes in the β-catenin network, with mean frequency counts >86 ($p < 0.001$).

4. Discussion

Gene/protein interaction networks provide critical information for a thorough understanding of cellular processes. Thus, detailed characterization of interactions between individual genes or proteins has become a primary focuses of biological research (Barabasi & Oltvai, 2004) . The complete biomedical literature database, containing a massive amount of information attained over a long period of time, is a largely untapped repository of information for study of gene/protein interaction networks (Chalmers et al., 1998; Gónzalez & Ochoa ,2008). Here, we generated a molecular network of β-catenin associated proteins. Analysis revealed that these proteins interacted with β-catenin via 18 different relationships, perform 13 biologic functions, take part in 15 cellular processes, localize to 12 cellular compartments, and signal in 14 different pathways. Significantly, this analysis identified a vast and mostly uncharacterized role for β-catenin in signal transduction pathways distinct from the Wnt and cadherin pathways. In particular, β-catenin may have a significant role in cell survival signaling.

Our analysis was performed using only abstracts instead of full text manuscripts. Our studies suggest that full-text articles contain too much detail for high-throughput analysis of biomedical research and development, while abstracts usually have higher information density and result in better quality relations extracted by text mining techniques. Further, abstracts are freely available through most public databases. Due to the large number of abstracts (10018) analyzed, we believe that the network data generated is both comprehensive and significant (Mackinnon et al., 2008; Ouzounis & Karp, 2000; Ptacek et al., 2005; Thieffry et al., 1998) .

Wnt/β-catenin signaling is involved in almost every aspect of embryonic development and controls homeostatic self-renewal in lots of adult tissues. Gene Otology identified three biological process (developmental process, cell communication and signaling transduction) that are significantly overrepresented within the β-catenin associated network. Validating these observations, several severe phenotypes in multiple tissues and organs can be observed in flies, frogs, fish, and mice following loss of Wnt/β-catenin signaling components. In adults, Wnt/β-catenin signaling remains essential throughout life for driving tissue renewal in rapidly self-renewing organs, including the intestine and skin. In addition, deregulation of Wnt/β-catenin signaling upsets the homeostatic balance in self-renewing tissues and leads to a variety of abnormalities and disease including bone defects and cancer. Pathway analysis of the β-catenin associated network revealed a close relationship between β-catenin and survival signaling (i.e. the Wnt, MAPK, insulin, and adhesion junction pathways), supporting an important role for β-catenin pathway in growth and development. Integration of PubMed text mining, homology prediction, gene neighbor, protein-protein interaction, gene fusion and other data sources identified AKT1, CCND1, CTNNB1, JUN, TP53 and VEGFA as hub genes in the β-catenin signaling network. Network analysis reveals an extremely high connectivity of these genes with other β-catenin associated genes. The involvement of these six genes in survival signaling, including anti-apoptosis, cell cycle and cell migration, provides further surport for a vital role for β-catenin in growth and development.

5. Conclusion

We performed natural language processing, a literature mining tool that can cluster a list of genes with keywords that are auto-extracted from their up-to-date related literature and then manually curated by the user, to establish the β-catenin biologic association network. Our results establish a significant association of this network with survival signaling. These data demonstrate the power of data-mining strategies as tools for biological discovery, suggesting that the use of similar strategies to consolidate all existing data for specific disease states, specifically cancer, may yield yield important discoveries in disease pathogenesis and identification of novel therapeutic targets. Further analysis of the β-catenin biologic association network may provide a deeper understanding of β-catenin signaling, particularly in relation to cell survival signaling.

6. Acknowledgements

This work was financially supported by the China National Natural Scientific Fund (81001128 and 30971136), the Tianjin Science and Technology Committee (09JCZDJC17600), and a Program for New Century Excellent Talents in University (NCET-07-0615), and Tianjin City High School Science & Technology Fund (2009CD01). We thank two anonymous referees for numerous excellent suggestions that led to substantial improvement of the manuscript.

7. References

Abrams, D., Baecker, R., and Chignell, M. Information Archiving with Bookmarks: PersonalWeb Space Construction and Organization, in *Proceedingsof CHI'98* (Los Angeles CA, April 1998), ACM Press, 41-48.

Albert, R.; Yeong H. & Barabasi, A. L. (1999). Diameter of the world-wide web. *Nature* 401 (September 1999) 130.

Albert, R. & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys,* 74, (January 2002) 47-97

Albert, R. (2005). Scale-free networks in cell biology. *Jcell Sci*, 118, (Pt 21) (October 2005)4947-57.

Apte C.; Weiss S. (1997) Data mining with decision trees and decision rules, *Future Generation Computer Systems*, Vol. 13, No 2-3, pp 197-210

Babu, M.M.; Luscombe N.M.; Aravind, L.; Gerstein, M. & Teichmann, S.A. (2004). Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* 14, (2004 Jun) 283-291.

Balaji, S.; Babu, M.M.; Iyer, L.M.; Luscombe, N.M. & Aravind, L (2006). Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. *J Mol Biol,* 360, (June 2006) 213-27.

Barabasi, A.L. & Oltvai, Z.N. (2004). Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet*, 5, (February 2004) 101-13.

Card, S.K., Robertson, G.C., and York, W. The WebBook and the Web Forager: An Information Workspace for the World-Wide Web, in *Proceedings of CHI'96*(Vancouver BC, April 1996), ACM Press, 111-117

Chalmers, M., Rodden, K., and Brodbeck, D. The Order of Things: Activity- Centred Information Access, In Proceedings of 7th International Conference on the WWW, 1998. (Brisbane Australia, April 1998), 359-367.

Chien C; Wanga W; Chenga J (2007) Data mining for yield enhancement in semiconductor manufacturing and an empirical study, Expert Systems with Applications, Vol 33, No 1,pp 192-198

Egghe, L and Rousseau, R. *Introduction to Informetrics: Quantitative methods in library, documentation, and information science*. Elsevier, New York, NY, 1990.

Ghannad-Rezaie M.; Soltanian-Zadeh H.; Siadat M.-R.; K.V. Elisevich (2006) Medical Data Mining using Particle Swarm Optimization for Temporal Lobe Epilepsy, *Proceedings of the IEEE World Congress on Computational Intelligence*, Vancouver, Canada, July 15-21

Giot, L.; Bader, J.S & Brouwer, C. et al. (2003). A protien interaction map of drosophila melanogaster. *Science*, 302, 5651, (December 2003) 1727-1736

Gónzalez S. & Ochoa A. Resolution of Japanese puzzles using Data Mining and Cultural Algorithms. (Accepted paper) In Proceedings of COMCEV'2008, México; 2008.

Kim, S. H.; Hyun Ju Noh (1997) Predictability of Interest Rates Using Data Mining Tools, EXPERT SYSTEMS WITH APPLICATIONS, Vol 13; No 2, pp 85-96

Lee, T. I.; Rinaldi, N. J & Robert, F. et al. (2002). Transcriptional regulatory networks in saccharomyces cerevisiae. *Sciencie*, 298, 5594, (October 2002) 799-804

Liu F. and Millar S.E. (2010).Wnt/beta-catenin signaling in oral tissue development and disease, In: Journal of Dental Research, Vol. 89, NO. 4, (January 8, 2010), PP. (318-330), ISSN (on-line) 1569-1551 (print) 1748-3050

Lopez DM, Blobel BG. (2008). A development framework for semantically interoperable health information systems. Int J Med Inform. 2008 Jul 11. [Epub ahead of print].

Mackinnon AD, Billington RA, Adam EJ, Dundas DD, Patel U. (2008). Picture archiving and communication systems lead to sustained improvements in reporting times and

productivity: results of a 5-year audit. Clin Radiol. 2008 Jul;63(7):796-804. Epub 2008 Mar 25.

McCraith, S.; Holtzman, T.; Moss, B. & Fields, S. (2000). Genome-wide analysis of vaccinia virus protein-protein interactions. *Proc. Natl. Acad. Sci. USA,* 97, 9, (April 2000) 4879-4884

Ouzounis, C. A. & Kar_finalp P.D. (2000). Global properties of the metabolic map of escherichia coli. *Genome Res.* 10, 4, (April 2000) 568-576

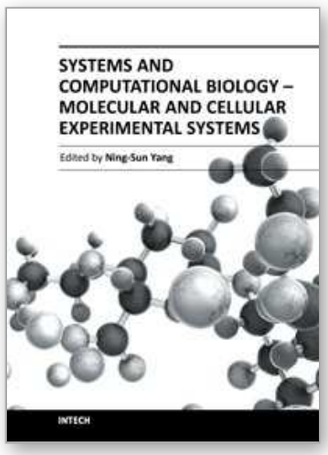
Ptacek, J.; Devegan, G.; Michaud,; G. Zhu,; H. Zhu,; X. Fasolo,; J. & et al. (2005). Global analysis of protein phosphorylation in yeast. *Nature*, 438, 7068, (December 2005) 679-84

Rual, J. F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Crico, A.; Li, N.; Berriz, G. F. & et al. (2005).Towards a proteome-scale map of the human protein-protein interaction network. *Nature,* 437, 7062, (October 2005) 1173-1178

Sousa T.; Silva A., Neves A. (2004) Particle Swarm based data mining algorithms for classification tasks, *Parallel Computing J.*, Vol. 30, pp. 767-783

Thieffry, D.; Huerta, A.M.; Pérez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. *Bioessays*, 20, 5, (1998 May) 433-40

Uetz, P.; Giot, L.; Cagney, G.; Mansfield,T.A.; Judson, R.S.; Knight, J.R.; & et al. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, 403, 6770, (February 2000) 623-627



Systems and Computational Biology - Molecular and Cellular Experimental Systems

Edited by Prof. Ning-Sun Yang

ISBN 978-953-307-280-7
Hard cover, 332 pages
Publisher InTech
Published online 15, September, 2011
Published in print edition September, 2011

Whereas some â€œmicroarrayâ€ or â€œbioinformaticsâ€ scientists among us may have been criticized as doing â€œcataloging researchâ€, the majority of us believe that we are sincerely exploring new scientific and technological systems to benefit human health, human food and animal feed production, and environmental protections. Indeed, we are humbled by the complexity, extent and beauty of cross-talks in various biological systems; on the other hand, we are becoming more educated and are able to start addressing honestly and skillfully the various important issues concerning translational medicine, global agriculture, and the environment. The two volumes of this book presents a series of high-quality research or review articles in a timely fashion to this emerging research field of our scientific community.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Fengming Lan, Xiao Yue, Lei Han, Peiyu Pu and Chunsheng Kang (2011). Data Mining Pubmed Using Natural Language Processing to Generate the β-Catenin Biological Association Network, Systems and Computational Biology - Molecular and Cellular Experimental Systems, Prof. Ning-Sun Yang (Ed.), ISBN: 978-953-307-280-7, InTech, Available from: http://www.intechopen.com/books/systems-and-computational-biology-molecular-and-cellular-experimental-systems/data-mining-pubmed-using-natural-language-processing-to-generate-the-catenin-biological-association-

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen