

PROCEEDINGS

Open Access

# IIIDB: a database for isoform-isoform interactions and isoform network modules

Yu-Ting Tseng<sup>1</sup>, Wenyuan Li<sup>2</sup>, Ching-Hsien Chen<sup>3</sup>, Shihua Zhang<sup>4</sup>, Jeremy JW Chen<sup>5,6,7</sup>, Xianghong Jasmine Zhou<sup>2</sup>, Chun-Chi Liu<sup>1,5,7\*</sup>

From The Thirteenth Asia Pacific Bioinformatics Conference (APBC 2015)  
HsinChu, Taiwan. 21-23 January 2015

## Abstract

**Background:** Protein-protein interactions (PPIs) are key to understanding diverse cellular processes and disease mechanisms. However, current PPI databases only provide low-resolution knowledge of PPIs, in the sense that “proteins” of currently known PPIs generally refer to “genes.” It is known that alternative splicing often impacts PPI by either directly affecting protein interacting domains, or by indirectly impacting other domains, which, in turn, impacts the PPI binding. Thus, proteins translated from different isoforms of the same gene can have different interaction partners.

**Results:** Due to the limitations of current experimental capacities, little data is available for PPIs at the resolution of isoforms, although such high-resolution data is crucial to map pathways and to understand protein functions. In fact, alternative splicing can often change the internal structure of a pathway by rearranging specific PPIs. To fill the gap, we systematically predicted genome-wide isoform-isoform interactions (IIIs) using RNA-seq datasets, domain-domain interaction and PPIs. Furthermore, we constructed an III database (IIIDB) that is a resource for studying PPIs at isoform resolution. To discover functional modules in the III network, we performed III network clustering, and then obtained 1025 isoform modules. To evaluate the module functionality, we performed the GO/pathway enrichment analysis for each isoform module.

**Conclusions:** The IIIDB provides predictions of human protein-protein interactions at the high resolution of transcript isoforms that can facilitate detailed understanding of protein functions and biological pathways. The web interface allows users to search for IIIs or III network modules. The IIIDB is freely available at <http://syslab.nchu.edu.tw/IIIDB>.

## Background

Protein-protein interactions (PPIs) perform and regulate fundamental cellular processes. As a consequence, identifying interacting partners for a protein is essential to understand its functions. In recent years, remarkable progress has been made in the annotation of all functional interactions among proteins in the cell. However, in both experimentally derived and computationally predicted protein-protein interactions, a “protein” generally refers to “all isoforms of the respective gene.” Yet, it is known

that alternative splicing often impacts PPI by either directly affecting protein interacting domains, or by indirectly impacting other domains, which, in turn, impact the PPI binding [1]. That is, alternative splicing can modulate the PPIs by altering the protein structures and the domain compositions, leading to the gain or loss of specific molecular interactions that could be key links of pathways (reviewed in reference [2]). It is very likely that different isoforms of the same protein interact with different proteins, thus exerting different functional roles. For example, the protein BCL2L1 is alternatively spliced into two isoforms: Bcl-xL (long form) and Bcl-xS (short form) [3], in which Bcl-xL inhibits apoptosis whereas Bcl-xS promotes apoptosis [4]. Vogler et al. reported that

\* Correspondence: [jimliu@nchu.edu.tw](mailto:jimliu@nchu.edu.tw)

<sup>1</sup>Institute of Genomics and Bioinformatics, National Chung Hsing University, Taiwan

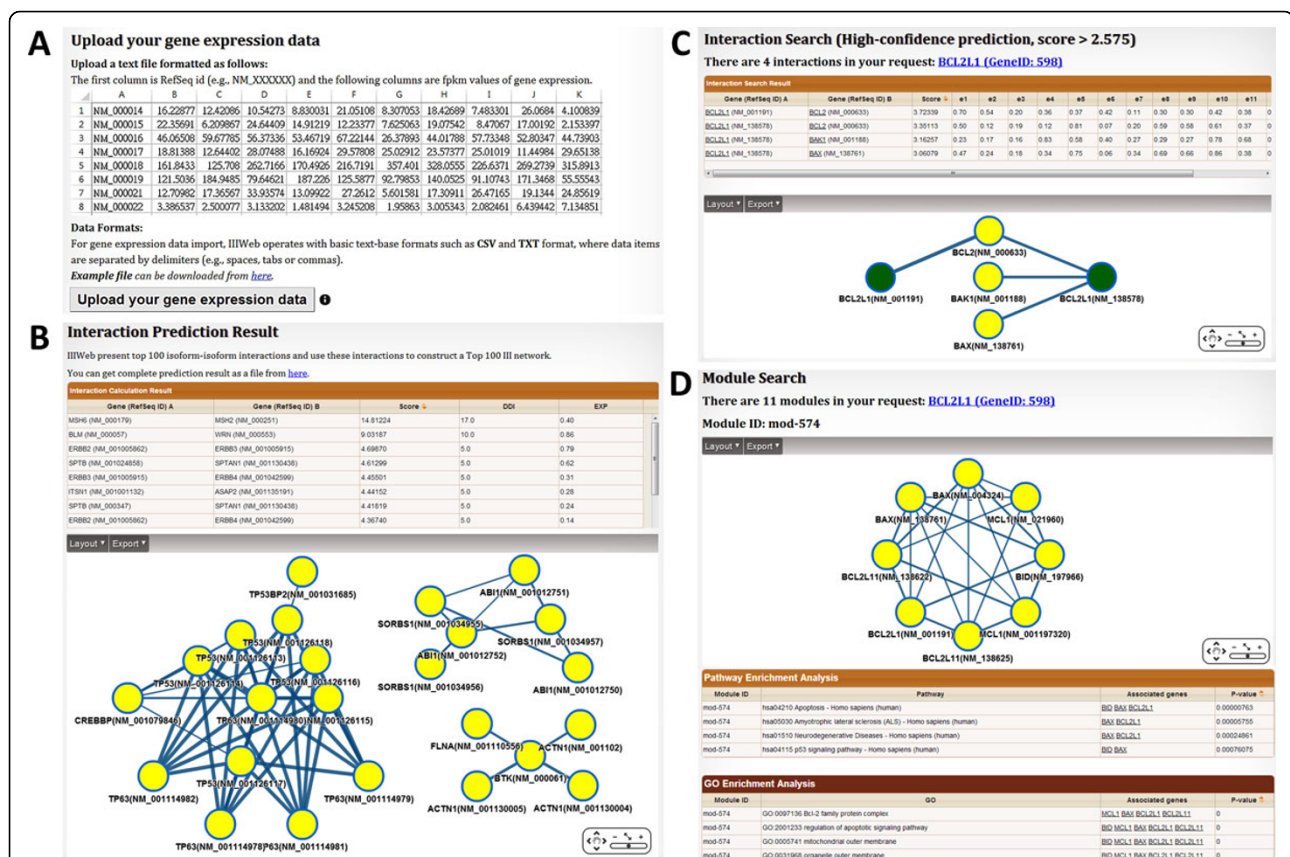
Full list of author information is available at the end of the article

the interaction of Bcl-xL and BAK1 in platelets ensures platelet survival [5]. Therefore, comprehensively identifying protein-protein interactions at the isoform level is important to systematically dissect cellular roles of proteins, to elucidate the exact composition of protein complexes, and to gain insights into metabolic pathways and a wide range of direct and indirect regulatory interactions.

Thus far, a series of studies have systematically predicted PPIs [6-10] and established PPI databases, e.g., OPHID [11], POINT [12], STRING [10] and PIPs [7]. With the exception that the IntAct database [13] contains 116 human PPIs with isoform specification, currently, none of those PPI databases has isoform-level PPI data. This is a huge knowledge gap yet to be filled. The rapid accumulation of RNA-seq data provides unprecedented opportunities to study the structures and topological dynamics of PPI networks at the isoform resolution. RNA-seq data provides two unique informative sources

for Isoform-Isoform Interaction (III) reconstruction: the absence or presence of specific isoforms under specific conditions, and the co-expression of two isoforms that may contribute to their interaction propensity. In this study, we seize this opportunity to comprehensively predict the possible interactions between splicing isoforms by integrating a series of RNA-seq data with domain-domain interaction data and PPI database. The resulting III network presents a high-resolution map of PPIs, which could be invaluable in studying biological processes and understanding cellular functions.

In this report, we described a database, IIIDB, for accessing and managing predicted human IIIs. In the IIIDB, users can differentiate access high-confidence and low-confidence predictions of human IIIs (see detailed description in Result section), and then see the full evidence values for each predicted III. Figure 1 shows the IIIDB web interface screen-shots of the III search and isoform module search function in the IIIDB. Users can



**Figure 1** The screenshots of the isoform interaction and module search function in the IIIDB. (A) Users can upload their own gene expression data for III prediction. (B) IIIDB shows top 100 IIIs and uses these interactions to construct an III network. Users can download complete prediction result as a file. (C) In interaction search function, users can input a gene symbol or gene ID to search associated IIIs and isoform modules. The interaction search section provides interface on searching high-confidence (score > 2.575) or low-confidence (score > 1.692) prediction. The resulting III table for interaction search shows full evidence values including Pearson correlations for 19 RNA-seq datasets and domain interaction score. (D) The resulting page for module search includes the user friendly network graph and pathway/GO enrichment analysis result.

upload their own gene expression data for III prediction, and then the users can download the predicted result. The searching function has three major parts: high-confidence interaction prediction search, low-confidence interaction prediction search, and isoform module search. The IIIDB provides auto-complete function in all search functions. Users can input a gene symbol or gene ID in the auto-complete field which provides an interface to quickly find and select matched values.

The IIIDB allows users to easily search IIIs and isoform modules, and then provides the evidence that led to each III prediction. To visualize the interactions with the isoforms of the input gene, we integrated CytoscapeWeb [14] to generate the interactive web-based III network (Figure 1B). Interestingly, the different isoforms within the same gene can be involved with different isoform modules that may open a new door to study differential functionality of isoforms of the gene. The IIIDB also provided GO/pathway enrichment analysis results for each isoform module, which helps biologists to study the biological insights of network modules at isoform level.

## Results

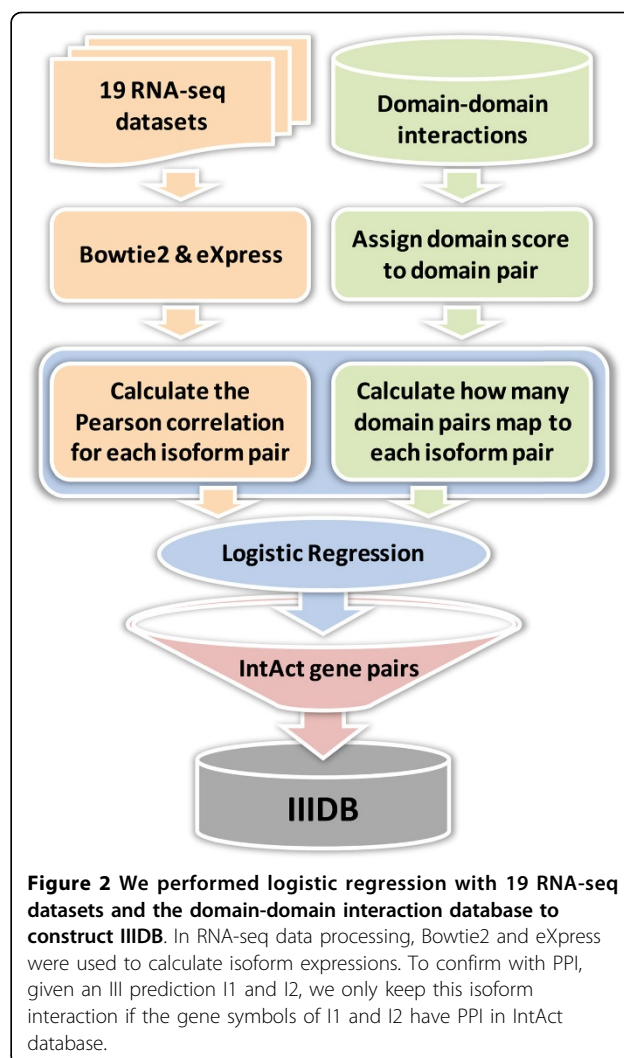
To obtain the isoform annotations in the human genome, we used the NCBI Reference Sequences (RefSeq) mRNAs as transcriptome annotation [15]. Figure 2 shows the framework of III prediction. We employed the logistic regression approach with 19 RNA-seq datasets (Table 1) and the domain-domain interaction database to infer IIIs. To confirm with PPI, given an III prediction  $I_1$  and  $I_2$ , we only keep this prediction if the gene symbols of  $I_1$  and  $I_2$  have PPI in IntAct database.

### High-confidence and low-confidence prediction of IIIs

In the IIIDB, we provided two III prediction sets using the logistic regression model: (a) high-confidence prediction: logit score  $> 2.575$  (precision 60% and recall 15%), it resulted in 4,476 IIIs; and (b) low-confidence prediction: logit score  $> 1.692$  (precision 40% and recall 68%). In addition, given a known PPI, the isoform pair with the best logit score among this PPI will be selected as low-confidence prediction. Thus, a PPI has at least one isoform-isoform interaction. It resulted in 54,605 IIIs (Figure 3).

### Isoform module discovery

To discover functional modules in the III network, we applied MODES network clustering method [16] on low-confidence III network to discover isoform modules with the given parameters (minimum module size 3, maximum module size 30, and density cutoff 0.7). An important feature of MODES is that it can discover overlapping dense isoform modules which allows one isoform to belong to multiple modules. We obtained



**Figure 2** We performed logistic regression with 19 RNA-seq datasets and the domain-domain interaction database to construct IIIDB. In RNA-seq data processing, Bowtie2 and eXpress were used to calculate isoform expressions. To confirm with PPI, given an III prediction  $I_1$  and  $I_2$ , we only keep this isoform interaction if the gene symbols of  $I_1$  and  $I_2$  have PPI in IntAct database.

1025 modules with size of 5.08 isoforms on average. To provide functional annotation and evaluate the module functional enrichment, we performed the enrichment analyses with GO [17] and KEGG pathway [18] databases. These databases are protein-level annotation which provides approximately functional annotation of isoforms.

To evaluate the significance, we randomly generated the same number and the same size of MODES isoform modules (i.e. 1025 randomized isoform modules). Table 2 shows the results of enrichment analyses for MODES and randomization modules for comparison. The MODES isoform modules have significantly higher functional enrichment rate than those of random cases, showing strong biological relevance of the predicted modules. The MODES isoform modules were further used to build the isoform module database in the IIIDB, and all isoform modules were listed in Additional file 1. We also stored the GO and KEGG enrichment results

**Table 1 19 RNA-seq datasets from SRA**

ID	SRA ID	# Exp	Title
d1	ERP000546	48	Illumina bodyMap2 transcriptome
d2	SRP005169	41	Widespread splicing changes in human brain development and aging
d3	SRP005408	31	Gene expression profile in postmortem hippocampus using RNAseq for addicted human samples
d4	SRP010280	31	Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins
d5	SRP002628	30	Comparative transcriptomic analysis of prostate cancer and matched normal tissue using RNA-seq
d6	ERP000550	29	Complete transcriptomic landscape of prostate cancer in the Chinese population using RNA-seq
d7	SRP005242	21	A Comparison of Single Molecule and Amplification Based Sequencing of Cancer Transcriptomes: RNA-Seq Comparison
d8	SRP002079	20	GSE20301: Dynamic transcriptomes during neural differentiation of human embryonic stem cells
d9	ERP000992	18	The effect of estrogen and progesterone and their antagonists in Ishikawa cell line compared to MCF7 and T47D cells
d10	SRP000727	16	Alternative Isoform Regulation in Human Tissue Transcriptomes
d11	SRP007338	16	GSE30017: Widespread regulated alternative splicing of single codons accelerates proteome evolution
d12	SRP010166	16	GSE34914: Deep Sequence Analysis of non-small cell lung cancer: Integrated analysis of gene expression, alternative splicing, and single nucleotide variations in lung adenocarcinomas with and without oncogenic KRAS mutations
d13	ERP000710	12	Transcriptome profiling of ovarian cancer cell lines
d14	SRP005411	11	RNA-Seq Quantification of the Complete Transcriptome of Genes Expressed in the Small Airway Epithelium of Nonsmokers and Smokers
d15	SRP006731	11	GSE29155: RNA-Seq analysis of prostate cancer cell lines using Next Generation Sequencing
d16	SRP013224	11	GSE38006: Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and the regulation of non-coding RNA expression in a CD4+ T cell line
d17	ERP000418	10	Gene expression profiles between normal and breast tumor genomes
d18	ERP000573	10	RNA and chromatin structure
d19	SRP010483	10	GSE35296: The human pancreatic islet transcriptome: impact of pro-inflammatory cytokines

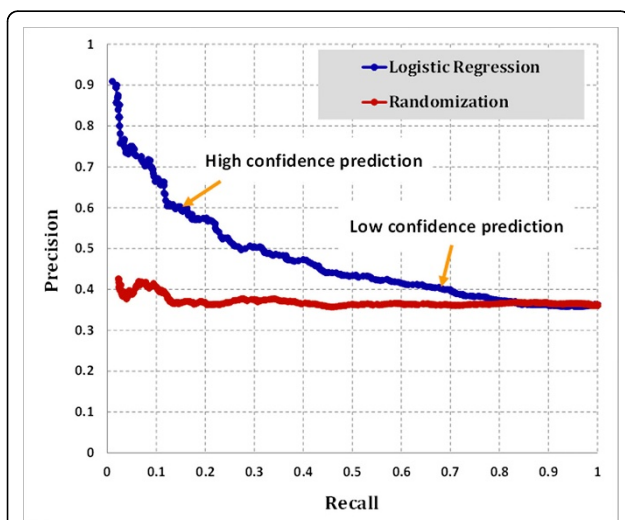
for all isoform modules in the IIIDB to provide potential functional annotation. In the IIIDB web interface, Figure 1C shows the result page of the isoform module search including the network graph and the enrichment analysis results. In the module result page, users can click any gene symbol

to iteratively search isoform modules, and sort the module result table by clicking the column header.

### Integrating diverse data sources for III prediction

Currently, most PPI databases do not provide information at the level of isoforms, which thus presents challenges for constructing a gold standard positive set (GSP) for the III prediction. Fortunately, in the June 2013 version of the IntAct database <http://www.ebi.ac.uk/intact/> [13], we identified 116 human PPIs with isoform specification (out of the total 43,508 distinct human PPIs). For example, IntAct has III between P29590-5 and P03243-1, which correspond to the 5th isoform of the protein P29590 and the 1st isoform of P03243. In addition, to obtain more IIIs for the GSP, we applied the following rule: given a PPI between protein P1 and P2, if both P1 and P2 only have single isoforms we also take it as the GSP. It resulted in 11,356 IIIs in the GSP set.

GSP covered 5,503 RefSeq IDs, and we used these RefSeq IDs to construct gold standard negative set (GSN). The GSN was defined as isoform pairs in which one isoform was assigned to the plasma membrane cellular component, and the other was assigned to the nuclear cellular component by the isoform-specific sub-cellular localization, in which we performed sequence-based predictions using the CELLO (subCELLular LOCALization predictor) [19]. To obtain the accurate isoform-specific annotation, we only used the cellular localization prediction results that consist of UniProt GO annotations.



**Figure 3 Precision and recall curve for logistic regression model.** At recall 15%, logistic regression model achieve 60% precision (High-confidence prediction); at recall 68%, logistic regression model achieve 40% precision (Low-confidence prediction).



**Table 2 The enrichment rate of isoform modules based on GO and pathway enrichments**

Isoform modules	# modules	% modules enriched with GO <sup>a</sup>	% modules enriched with pathway <sup>b</sup>
MODES modules	1025	88.7%	36.1%
Randomization	1025	49.7%	10.1%

<sup>a</sup> The modules enriched with GO term (P-value < 0.001).

<sup>b</sup> The modules enriched with pathway (P-value < 0.01).

It resulted in 36 RefSeq IDs for plasma membrane cellular component and 739 RefSeq IDs for nuclear cellular component. In addition, isoforms that are assigned to both the plasma membrane and the nuclear cellular component are excluded in GSN.

To calculate precision and recall, we used timestamp to divided GSP into training and test GSP sets, in which if an interaction with timestamp after 1st Jan 2012, it will be assigned to test GSP set (10,408 IIIs); otherwise, it will be assign to training GSP set (948 IIIs). When the GSP is decided, we used the RefSeq IDs covered in GSP to build GSN. Figure 3 shows the precision and recall curve for the logistic regression model.

#### Case studies

To demonstrate the biological importance of the IIIDB, we searched for isoform-associated reports in the literature. Although isoform-specific protein function studies are very rare, we found isoform-specific biological evidences with BCL2L1, which validated our III prediction of BCL2L1. In addition, we also found diverse biological functions with Ras association domain family in our isoform modules.

#### BCL2L1

BCL2L1 has two isoforms (Table 3), in which BCL2L1 isoform 1 is called Bcl-xL (long form) and BCL2L1 isoform 2 is called Bcl-xS (short form) [3]. These isoforms play important roles in apoptosis as follows: Bcl-xL inhibits apoptosis whereas Bcl-xS promotes apoptosis [4]. In our high-confidence prediction, BCL2L1 isoform 1 (Bcl-xL) interacts with BAK1, BAX and NLRP1 to inhibit apoptosis, but BCL2L1 isoform 2 (Bcl-xS) doesn't (Table 3). In previous reports, the fluorescence anisotropy, analytical ultracentrifugation, and NMR assays confirmed a direct interaction between Bcl-xL and BAK1 [5,20]. Vogler *et al.* also reported that the interaction of Bcl-xL and BAK1 in platelets ensures cell survival [5]. Edlich *et al.* reported that an interaction between Bcl-xL and BAX not only inhibits BAX activity but also maintains BAX in the cytosol [21]. On the other hand, Chang *et al.* demonstrated that Bcl-xL interacted with endogenous BAX in 293 cells.

However, no significant amount of BAX was detectable in the Bcl-xS immunoprecipitation [21], suggesting that Bcl-xS does not interact with BAX. In addition, Bruey *et al.* reported that Bcl-xL interacts with NALP1 to suppress apoptosis [22]. Thus, these previous biological studies validated our III prediction of BCL2L1.

#### RASSF1

RASSF1 is RAS association domain family member 1. RASSF1 has four isoforms, of which two isoforms belong to isoform modules (Mod-162 and Mod-384). The Mod-162 consists of RASSF1 isoform A, RASSF5 and STK4, whereas the Mod-384 consists of RASSF1 isoform C, RASSF3, RASSF4, SAV1, STK3 and STK4 (Figure 4). Interestingly, several studies reported that RASSF5 may interact with RASSF1 isoform A to suppress tumors [23-25], suggesting that the Mod-162 is validated. On the other hand, RASSF1 isoform C may play a completely different role as an oncogene in high-grade tumors [26]. Although the function of RASSF1 isoform C is still not clear, the RASSF1 isoform A and C should have distinct functions. The IIIDB assigned the RASSF1 isoform A and C into the Mod-162 and the Mod-384, respectively, suggesting a new hypothesis of the isoform-level modules.

## Methods

### Isoform coexpression network construction

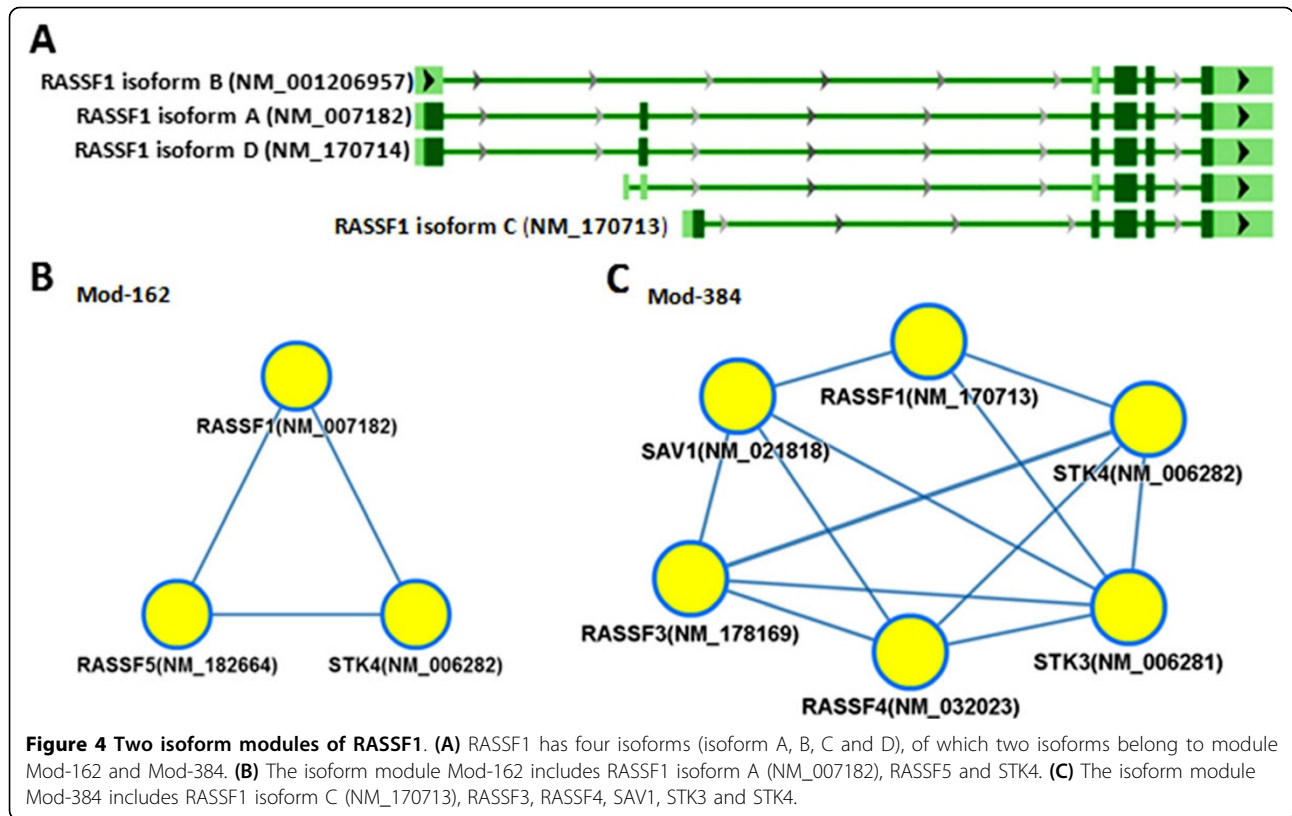
To construct the Isoform coexpression networks, we selected 19 RNA-seq datasets with at least 10 experiments from Sequence Read Archive (SRA) database (Table 1) [27]. We performed the eXpress [28] with Bowtie2 aligner [29] to obtain isoform expression values. NCBI Reference Sequences (RefSeq) mRNAs with protein sequences was used as transcriptome annotation, which included 31,454 RefSeq IDs (Jan 2013 version) [15].

### Logistic regression model

The logistic regression model has been applied to PPI prediction [30,31], as it is suitable to describe the relationship between a binary response variable and a set of

**Table 3 BCL2L1 isoform interaction partners.**

Isoform	RefSeq ID	mRNA length	Protein length	UniProt ID	Interaction partner (high confidence)
BCL2L1 isoform 1 (Bcl-xL)	NM_138578	2575	233	Q07817-1	BAK1 BAX NLRP1
BCL2L1 isoform 2 (Bcl-xS)	NM_001191	2386	170	Q07817-2	BCL2



explanatory variables. We used logistic regression approach to build the prediction model as follows:

$$\text{logit}(y_{ij}) = \alpha_0 + \alpha_1 E1_{ij} + \alpha_2 E2_{ij} + \dots + \alpha_{19} E19_{ij} + \alpha_{20} DDI_{ij}$$

where  $\alpha_0, \alpha_1, \dots, \alpha_{20}$  are regression coefficients and  $y_{ij}$  is the probability of the isoform interaction between isoform  $i$  and isoform  $j$ .  $DDI, E1, E2, \dots, E19$  are described as follows: (a) **Domain-domain interaction (DDI) score:** For all combinations of human isoform pairs, if a isoform pair has domain-domain interaction in the DOMINE database [32,33], then we assign the DDI score by the confidence level in the DOMINE database as follows: high-confidence prediction: 3, medium-confidence prediction: 2, and low-confidence prediction: 1. If the isoform pair has several DDI scores, we take the highest score. (b) **19 RNA-seq datasets (E1, E2, ..., E19):** the absolute values of Pearson correlations for all isoform pairs derived from RNA-seq datasets. Since each RNA-seq dataset may have different quality and data type, we used the logistic regression model to integrate 19 RNA-seq datasets, and then the coefficients of datasets reflected quality of the RNA-seq datasets.

## Additional material

**Additional file 1: The isoform modules.** There are 1025 isoform modules with gene symbols and RefSeq IDs.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements

The National Institutes of Health (NIH) [NHLBI MAPGEN U01HL108634 and NIGMS R01GM105431] and the National Science Foundation [0747475 to X.J.Z.]; the Taiwan Ministry of Science and Technology grant [MOST 103-2320-B-005-004] and the Ministry of Education, Taiwan, R.O.C. under the ATU plan (to C.C.L.). This article has been published as part of *BMC Genomics* Volume 16 Supplement 2, 2015: Selected articles from the Thirteenth Asia Pacific Bioinformatics Conference (APBC 2015): Genomics. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenomics/supplements/16/S2>

## Authors' details

<sup>1</sup>Institute of Genomics and Bioinformatics, National Chung Hsing University, Taiwan. <sup>2</sup>Molecular and Computational Biology, University of Southern California, USA. <sup>3</sup>Department of Internal Medicine, Division of Pulmonary and Critical Care Medicine and Center for Comparative Respiratory Biology and Medicine, University of California Davis, Davis, California, USA. <sup>4</sup>National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, CAS, Bei-jing, China. <sup>5</sup>Institute of Biomedical Sciences, National Chung Hsing University, Taiwan. <sup>6</sup>Institute of Molecular Biology, National Chung Hsing University, Taiwan. <sup>7</sup>Agricultural Biotechnology Center, National Chung Hsing University, Taiwan.

Published: 21 January 2015

## References

1. Ellis JD, Barrios-Rodiles M, Colak R, Irimia M, Kim T, Calarco JA, Wang X, Pan Q, O'Hanlon D, Kim PM, et al: Tissue-specific alternative splicing remodels protein-protein interaction networks. *Molecular cell* 2012, **46**(6):884-892.
2. Leeman JR, Gilmore TD: Alternative splicing in the NF-kappaB signaling pathway. *Gene* 2008, **423**(2):97-107.

3. Perumalsamy A, Fernandes R, Lai I, Detmar J, Varmuza S, Casper RF, Jurisicova A: **Developmental consequences of alternative Bcl-x splicing during preimplantation embryo development.** *FEBS J* 2010, **277**(5):1219-1233.
4. Rajan P, Elliott DJ, Robson CN, Leung HY: **Alternative splicing and biological heterogeneity in prostate cancer.** *Nat Rev Urol* 2009, **6**(8):454-460.
5. Vogler M, Hamali HA, Sun XM, Bampton ET, Dinsdale D, Snowden RT, Dyer MJ, Goodall AH, Cohen GM: **BCL2/BCL-XL inhibition induces apoptosis, disrupts cellular calcium homeostasis, and prevents platelet activation.** *Blood* 2011, **117**(26):7145-7154.
6. Cho YR, Shi L, Ramanathan M, Zhang A: **A probabilistic framework to predict protein function from interaction data integrated with semantic knowledge.** *BMC Bioinformatics* 2008, **9**:382.
7. McDowall MD, Scott MS, Barton GJ: **PIPs: human protein-protein interaction prediction database.** *Nucleic Acids Res* 2009, **37** Database: D651-656.
8. Rhodes DR, Tomlins SA, Varambally S, Mahavisno V, Barrette T, Kalyana-Sundaram S, Ghosh D, Pandey A, Chinnaiyan AM: **Probabilistic model of the human protein-protein interaction network.** *Nat Biotechnol* 2005, **23**(8):951-959.
9. Scott MS, Barton GJ: **Probabilistic prediction and ranking of human protein-protein interactions.** *BMC Bioinformatics* 2007, **8**:239.
10. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, Doerks T, Stark M, Muller J, Bork P, et al: **The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.** *Nucleic Acids Res* 2011, **39** Database: D561-568.
11. Brown KR, Jurisica I: **Online predicted human interaction database.** *Bioinformatics* 2005, **21**(9):2076-2082.
12. Huang TW, Tien AC, Huang WS, Lee YC, Peng CL, Tseng HH, Kao CY, Huang CY: **POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome.** *Bioinformatics* 2004, **20**(17):3273-3276.
13. Aranda B, Achuthan P, Alam-Faruque Y, Armean I, Bridge A, Derow C, Feuermann M, Ghanbarian AT, Kerrien S, Khadake J, et al: **The IntAct molecular interaction database in 2010.** *Nucleic Acids Res* 2010, **38** Database: D525-531.
14. Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD: **Cytoscape Web: an interactive web-based network browser.** *Bioinformatics* 2010, **26**(18):2347-2348.
15. Pruitt KD, Tatusova T, Brown GR, Maglott DR: **NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy.** *Nucleic Acids Res* 2012, **40** Database: D130-135.
16. Hu H, Yan X, Huang Y, Han J, Zhou XJ: **Mining coherent dense subgraphs across massive biological networks for functional discovery.** *Bioinformatics* 2005, **21**(Suppl 1):i213-221.
17. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene ontology: tool for the unification of biology.** The Gene Ontology Consortium. *Nat Genet* 2000, **25**(1):25-29.
18. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**(1):27-30.
19. Yu CS, Chen YC, Lu CH, Hwang JK: **Prediction of protein subcellular localization.** *Proteins* 2006, **64**(3):643-651.
20. Sot B, Freund SM, Fersht AR: **Comparative biophysical characterization of p53 with the pro-apoptotic BAK and the anti-apoptotic BCL-XL.** *J Biol Chem* 2007, **282**(40):29193-29200.
21. Edlich F, Banerjee S, Suzuki M, Cleland MM, Arnould D, Wang C, Neutzner A, Tjandra N, Youle RJ: **Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol.** *Cell* 2011, **145**(1):104-116.
22. Bruey JM, Bruey-Sedano N, Luciano F, Zhai D, Balpai R, Xu C, Kress CL, Bailly-Maitre B, Li X, Osterman A, et al: **Bcl-2 and Bcl-XL regulate proinflammatory caspase-1 activation by interaction with NALP1.** *Cell* 2007, **129**(1):45-56.
23. Park J, Kang SI, Lee SY, Zhang XF, Kim MS, Beers LF, Lim DS, Avruch J, Kim HS, Lee SB: **Tumor suppressor ras association domain family 5 (RASSF5/NORE1) mediates death receptor ligand-induced apoptosis.** *J Biol Chem* 2010, **285**(45):35029-35038.
24. Fernandes MS, Carneiro F, Oliveira C, Seruca R: **Colorectal cancer and RASSF family—a special emphasis on RASSF1A.** *Int J Cancer* 2013, **132**(2):251-258.
25. Richter AM, Pfeifer GP, Dammann RH: **The RASSF proteins in cancer; from epigenetic silencing to functional characterization.** *Biochim Biophys Acta* 2009, **1796**(2):114-128.
26. Pelosi G, Fumagalli C, Trubia M, Sonzogni A, Rektman N, Maisonneuve P, Galetta D, Spaggiari L, Veronesi G, Scarpa A, et al: **Dual role of RASSF1 as a tumor suppressor and an oncogene in neuroendocrine tumors of the lung.** *Anticancer Res* 2010, **30**(10):4269-4281.
27. Kodama Y, Shumway M, Leinonen R: **International Nucleotide Sequence Database C: The Sequence Read Archive: explosive growth of sequencing data.** *Nucleic Acids Res* 2012, **40** Database: D54-56.
28. Roberts A, Pachter L: **Streaming fragment assignment for real-time analysis of sequencing experiments.** *Nat Methods* 2013, **10**(1):71-73.
29. Liu Y, Schmidt B: **Long read alignment based on maximal exact match seeds.** *Bioinformatics* 2012, **28**(18):i318-i324.
30. Qi Y, Klein-Seetharaman J, Bar-Joseph Z: **A mixture of feature experts approach for protein-protein interaction prediction.** *BMC Bioinformatics* 2007, **8**(Suppl 10):S6.
31. Sprinzak E, Altuvia Y, Margalit H: **Characterization and prediction of protein-protein interactions within and between complexes.** *Proc Natl Acad Sci USA* 2006, **103**(40):14718-14723.
32. Raghavachari B, Tasneem A, Przytycka TM, Jothi R: **DOMINE: a database of protein domain interactions.** *Nucleic Acids Res* 2008, **36** Database: D656-661.
33. Yellaboina S, Tasneem A, Zaykin DV, Raghavachari B, Jothi R: **DOMINE: a comprehensive collection of known and predicted domain-domain interactions.** *Nucleic Acids Res* 2011, **39** Database: D730-735.

doi:10.1186/1471-2164-16-S2-S10

Cite this article as: Tseng et al.: **IIIDB: a database for isoform-isoform interactions and isoform network modules.** *BMC Genomics* 2015 **16**(Suppl 2):S10.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

