



Toward Personalized Network Biomarkers in Alzheimer's Disease: Computing Individualized Genomic and Protein Crosstalk Maps

Kanchana Padmanabhan^{1,2}, Katie Shpanskaya³, Gonzalo Bello¹, P. Murali Doraiswamy^{4,5} and Nagiza F. Samatova^{1,2*}

¹ Department of Computer Science, North Carolina State University, Raleigh, NC, United States, ² Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, ³ Stanford University School of Medicine, Stanford, CA, United States, ⁴ Department of Psychiatry, Duke University, Durham, NC, United States, ⁵ Duke Institute for Brain Sciences, Duke University, Durham, NC, United States

Keywords: Alzheimer's disease, pathways, crosstalk, SNP data, bioinformatics, biomarker

INTRODUCTION

Alzheimer's disease (AD) affects about 10% of the population over 65 years old (Brookmeyer et al., 2011) and evidence suggests it may have an extended preclinical phase during which treatments are likely to be most effective. Thus, it is important to discover and develop accurate biomarkers that reflect the complexity of the disease at an individual level. Fluid (e.g., blood and spinal fluid markers) and imaging (e.g., MRI or PET imaging) biomarkers remain important for diagnosis and prognosis but in their current state do not capture the full underlying heterogeneity.

Over 20 years of genomic and proteomic studies have yielded a rich array of information on the molecular cascade of AD suggesting a role for a diverse array of underlying molecular mechanisms (Juhász et al., 2011). Indeed, an estimated 70% of AD risk is attributed to genetics; however, the currently recognized genetic mutations linked to AD, including amyloid precursor protein (APP) and presenilins (PSEN) 1 and 2, only account for 5% of AD cases (Ballard et al., 2011). This discrepancy has progressed genomic studies of AD to consider that complex interactions across numerous molecular pathways likely contribute to AD initiation and progression.

Characterization of such complex crosstalks (i.e., interactions) across multiple molecular pathways is a non-trivial endeavor. One approach to identify whether crosstalk exists between two pathways is to determine if both pathways work together to perform a biological function such as the Toll-like receptor and complement pathways interacting to reinforce innate immunity (Hajishengallis and Lambris, 2010). Crosstalks can also occur between signal transduction pathways, usually taking the form of direct protein or transmembrane interactions. For example, interactions between the major regulatory NFκB pathway and multiple oncogenic signaling pathways, such as Ras and p53, are key to the development of carcinogenesis (Oeckinghaus et al., 2011). In AD, several potential crosstalks have been noted *in vitro*, such as those between amyloid and tau pathways and inflammation (Selkoe, 2001; Ballatore et al., 2007; Lanni et al., 2007).

From the computational methodology standpoint, the study of predicting crosstalks is still in its infancy. Existing methods predict crosstalks between known metabolic pathways using physical protein interaction networks (Myers et al., 2005; Li et al., 2008; Liu et al., 2010; Xu et al., 2011; Mukherjee et al., 2014). However, these computational methods do not take advantage of the different physical evidences available such as direct protein binding, biochemical evidences such as phosphorylation, and functional evidences such as transcriptional regulation. Moreover, discovery, characterization, and utilization of pathway crosstalks as biomarkers for disease prognosis have not been investigated.

OPEN ACCESS

Edited by:

Catarina Oliveira,
University of Coimbra, Portugal

Reviewed by:

Franc Llorens,
Centro de Investigación Biomédica en
Red Sobre Enfermedades
Neurodegenerativas, Spain
Charlotte Elisabeth Teunissen,
VU University Amsterdam,
Netherlands

*Correspondence:

Nagiza F. Samatova
samatova@csc.ncsu.edu

Received: 14 May 2017

Accepted: 15 September 2017

Published: 26 September 2017

Citation:

Padmanabhan K, Shpanskaya K,
Bello G, Doraiswamy PM and
Samatova NF (2017) Toward
Personalized Network Biomarkers in
Alzheimer's Disease: Computing
Individualized Genomic and Protein
Crosstalk Maps.
Front. Aging Neurosci. 9:315.
doi: 10.3389/fnagi.2017.00315

It is our opinion that the next step in biomarker discovery would be to go beyond discrete biomarkers to complex personalized network biomarkers. A variety of combinatorial approaches have been proposed and applied previously for fluid, tissue, and imaging markers (reviewed in Atluri et al., 2013). However, the development of dynamic network biomarkers using the rich array of available genomic AD data has yet to be realized. To construct such biomarkers, we envision the creation of generic pathway crosstalk maps that can be enriched with patient-specific genomic data (e.g., SNPs) to generate personalized genetic risk profiles. Specifically, we propose the following schematic steps: (A) identify *potential* pathway crosstalks using existing gene/protein/pathway-level data (**Figure 1A**), (B) identify *patient-specific* pathway crosstalks, for example by using SNP information (**Figure 1B**), and (C) utilize clinical datasets to identify *significant* pathway crosstalks as biomarkers for AD prediction.

Creation of a Generic Pathway Crosstalk Map

Traditional statistical approaches for characterizing interactions have only a limited ability to capture the heterogeneity and complexity of pathway crosstalk and consequently a number of novel informatics approaches have been proposed (Myers et al., 2005; Li et al., 2008; Liu et al., 2010; Xu et al., 2011; Hsu and Yang, 2012; Atluri et al., 2013; Diao et al., 2013; Wang et al., 2013; Tegge et al., 2015). One could quantify via scores the likelihood that a pair of pathways will crosstalk based on existing biological datasets that provide evidence for possible crosstalks (including physical interaction, genetic interaction, and transcription factors). To have a more robust pathway crosstalk map, one could incorporate a wide array of evidences. Scores from each of these evidences can then be combined to build a generic pathway crosstalk reference map analogous to the “Kyoto Encyclopedia of Genes and Genomes” (KEGG) pathway reference map. The likelihood of two pathways crosstalking can be scored utilizing one of several different methods. One method could be based on the presence of “common elements,” such as shared enzymes and metabolites. For example, the regulatory signaling protein kinase A (PKA) and protein kinase C (PKC) pathways converge onto the MAPK kinase system using the small G protein, Ras, a common element of both the PKA and PKC pathway (Franco et al., 2017). Another method could rely on the presence of “interacting elements,” such as physical protein-protein interactions. An example of such an interaction in humans is the ANP32A protein binding to Axin-1, another protein, to enhance its suppression of the Wnt pathway, a key pathway in adult tumor formation (Stelzl et al., 2005). Using such methods, one could build a network from a generic pathway crosstalk reference map where the nodes represent pathways and the edges represent a statistically significant *p*-value for crosstalk likelihood between a pathway pair.

Characterization of Patient-Specific Pathway Crosstalks

To determine which of the pathway crosstalks in the generic reference map may be utilized as a biomarker for AD, one could

then go further to characterize *patient-specific* pathway crosstalks. For this purpose, one could make use of single-nucleotide polymorphism (SNP) and gene expression (transcriptomics) data collected in large naturalistic studies such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Dominant Inherited Alzheimer's Network (DIAN) or large clinical trials such as the Alzheimer's Disease Genetic Consortium (ADGC). Using SNP-data as an illustrative example, characterization of patient-specific pathway crosstalks could be broken down into four steps schematically as shown in **Figure 1B**:

1. Obtain a mapping of SNPs to pathways using genetic information.
2. Identify the list of SNPs that are present in a patient.
3. Use the mapping obtained in Step 1 and the patient-specific SNP list in Step 2 to obtain the pathways that are “SNP-enriched” in the patient.
4. Use the “SNP-enriched” pathways from Step 3 to obtain patient-specific pathway crosstalks.

If a different “-omics” data source is used such as proteomics or metabolomics, then a simple substitution of SNPs for the desired data source in the above schema is needed to obtain “patient specific data-enriched” pathways.

A number of other informatics approaches that attempt to predict crosstalk are possible and have been reviewed (Myers et al., 2005; Dotan-Cohen et al., 2009; Liu et al., 2010; Xu et al., 2011; Atluri et al., 2013). These approaches are limited to physical protein interaction alone and do not capitalize upon additional physical and functional crosstalk evidences available (Li et al., 2008). Moreover, personalization of pathway crosstalks and subsequent utilization of such crosstalks as prognostic biomarkers has not been broached.

Genetic Crosstalk Findings in Alzheimer's Disease

A recent systems biology analysis conducted by Hu et al. (2017) constructed an AD pathway crosstalk map from 430 human genes associated with AD. A total of 68 biological pathways were found to be enriched by these AD-related genes. Pathway crosstalks were determined by the proportion of overlapping genes with a minimum of two shared genes between pathways as a requirement for inclusion in the pathway crosstalk map. Their crosstalk network revealed three core interacting modules, consisting of immune modulation-related pathways, cell growth/survival and neuroendocrine-related pathways, and neuronal and drug-metabolism pathways (**Figure 1C**). Such an AD immune-endocrine-neuronal regulatory network is consistent with our current understanding of AD pathogenesis. These promising findings highlight the need for moving beyond the traditional single-gene based studies to network and pathway-based methodologies. Furthermore, we believe enriching sophisticated pathway crosstalk analyses with patient-specific data could yield powerful personalized biomarkers that could advance our understanding of disease mechanisms and potential therapeutic targets.

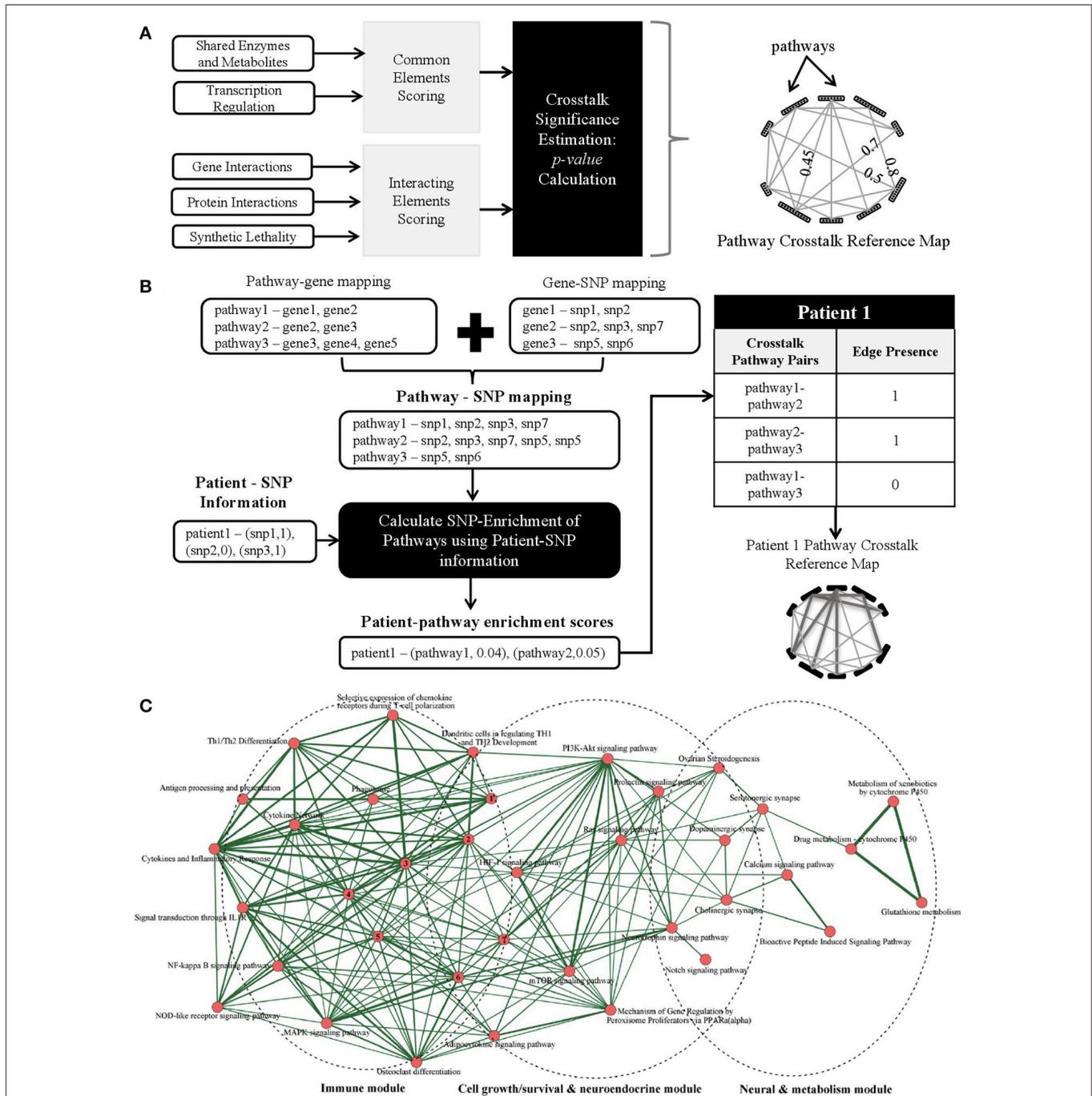


FIGURE 1 | (A) Proposed methods to identify *potential* pathway crosstalks. The methodology has three steps: (1) quantify crosstalk likelihood using multiple individual evidences to score each pathway pair, (2) obtain a combined score from a variety of evidence for possible crosstalks, and (3) build the generic crosstalk reference map. **(B)** Schematic methods to identify enriched *patient-specific* pathways and pathway crosstalks using SNP data as an example with three steps: (1) map SNPs to genes and, in turn, to pathways using SNP and gene location information, (2) choose a genetic model and calculate a patient-specific SNP enrichment score for each pathway using the patient's allele information, and (3) overlaying the patient-specific pathway enrichment scores onto the reference crosstalk map to build patient-specific pathway crosstalk maps. **(C)** Crosstalk network amid Alzset-overrepresented pathways. Vertices, biological pathways; lines, crosstalks among pathways. Width of one line (edge) shows direct proportion with the crosstalk level of a given pathway pair. Nodes tagged with numbers represent the following corresponding pathways: 1, intestinal immune network for IgA production; 2, toll-like receptor signaling pathway; 3, cytokine–cytokine receptor interaction; 4, hematopoietic cell lineage; 5, TNF signaling pathway; 6, apoptosis; 7, Fcε RI signaling pathway. Panel C is reproduced with permission from Hu et al. (2017) courtesy of Ju Wang Ph.D., Tianjin Medical University.

Validation of Such Approaches

Emerging robust bioinformatics models will ultimately need to be validated using existing datasets and biological models of disease. Some recent studies have begun to utilize such methods. Mukherjee et al. (2014) performed a network analyses incorporating a human protein-protein interaction database mined from 12 different sites including BIND, BioGRID, Intct to the HapMap2-imputed combined ADGC data set from 15 studies. They identified a set of significant modules and candidate genes and then demonstrated an initial functional validation of some of these candidate markers as modifiers of amyloid-beta toxicity *in vivo* using a transgenic *C. elegans* model. Liu et al. (2016) analyzed 528 biomarkers in ADNI data (proteomics, MRI, cognitive tests) to examine the sequence of network changes related to the risk for progression from mild cognitive impairment (MCI) to AD. A semi-mechanism-based Bayesian network with 26 nodes and 43 arcs was generated, which achieved a high 10-fold cross-validated prediction performance with 95% sensitivity and 65% specificity. The network analyses identified several markers of relevance such as fibrin clot formation and hyperinsulinemia. This is consistent with our own preliminary work using longitudinal ADNI data from 91 MCI subjects which suggests that constructing patient-specific SNP crosstalk maps may enhance the predictive accuracy above and beyond the traditional approach of combining discrete MRI and cognitive test markers. Clearly, further experimental validation will be critical and readers are referred elsewhere for a more comprehensive review of validation methods (Chen et al., 2015).

CONCLUSIONS

In summary, we call for the AD field to move beyond discrete biomarkers and utilize the full power of informatics and big data approaches to build and test personalized markers at a pathway and network level. While we have chosen AD as an example, the issues we propose are also highly relevant to other neurodegenerative disorders, such as vascular dementia

REFERENCES

- Atluri, G., Padmanabhan, K., Fang, G., Steinbach, M., Petrella, J. R., Lim, K., et al. (2013). Complex biomarker discovery in neuroimaging data: finding a needle in a haystack. *NeuroImage Clin.* 3, 123–131. doi: 10.1016/j.nicl.2013.07.004
- Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., and Jones, E. (2011). Alzheimer's disease. *Lancet* 377, 1019–1031. doi: 10.1016/S0140-6736(10)61349-9
- Ballatore, C., Lee, V. M.-Y., and Trojanowski, J. Q. (2007). Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nat. Rev. Neurosci.* 8, 663–672. doi: 10.1038/nrn2194
- Brookmeyer, R., Evans, D. A., Hebert, L., Langa, K. M., Heeringa, S. G., Plassman, B. L., et al. (2011). National estimates of the prevalence of Alzheimer's disease in the United States. *Alzheimer's Dement.* 7, 61–73. doi: 10.1016/j.jalz.2010.11.007
- Chen, H., Zhu, Z., Zhu, Y., Wang, J., Mei, Y., and Cheng, Y. (2015). Pathway mapping and development of disease-specific biomarkers: protein-based network biomarkers. *J. Cell. and Mol. Med.* 19, 297–314. doi: 10.1111/jcmm.12447

or dementia with Lewy bodies, where even less is known about how various biomarkers interact. Indeed, the availability of rich biomarker data across a range of neurodegenerative disorders would enable more accurate pathology-based classification of such conditions (as opposed to the current predominantly clinical classifications). We further hypothesize that building dynamic network biomarkers and pathway crosstalk reference maps using the combined power of several protein/gene-level knowledge priors could accelerate discovery of disease-specific mechanisms and novel drug targets by enrichment with patient-specific genetic information. Application of this methodology to large public AD datasets is needed to test our hypotheses and refine the methods. Subsequent replication in independent datasets and population studies as well as functional validation of mechanisms in laboratory models will be the next steps. Ultimately, it is our hope that such novel methods may yield further insights into both disease mechanisms as well as novel targets for biomarker development and drug discovery.

AUTHOR CONTRIBUTIONS

PD and NS designed and provided conceptual guidance for all aspects of this work. KP, KS, and GB executed the work, conducted a thorough review of literature, and collaborated with PD and NS to draft, critically revise, and approve the final opinion piece.

FUNDING

We wish to thank the Karen L. Wrenn Trust and Cure Alzheimer's Fund for their funding to apply informatics methods to AD discovery. PD has also received research grants and/or advisory fees from several government agencies, advocacy groups, and pharmaceutical/imaging companies for other studies, and owns stock in several companies whose products are not discussed here. PD and NS are applicants on a preliminary patent for discovering crosstalk pathways which is not licensed and has not been granted yet.

- Diao, H., Li, X., and Hu, S. (2013). The identification of dysfunctional crosstalk of pathways in Parkinson disease. *Gene* 515, 159–162. doi: 10.1016/j.gene.2012.11.003
- Dotan-Cohen, D., Letovsky, S., Melkman, A. A., and Kasif, S. (2009). Biological Process Linkage Networks. *PLoS ONE* 4:e5313. doi: 10.1371/journal.pone.0005313
- Franco, R., Martínez-Pinilla, E., Navarro, G., and Zamarbide, M. (2017). Potential of GPCRs to modulate MAPK and mTOR pathways in Alzheimer's disease. *Prog. Neurobiol.* 149–150, 21–38. doi: 10.1016/j.pneurobio.2017.01.004
- Hajishengallis, G., and Lambris, J. D. (2010). Crosstalk pathways between Toll-like receptors and the complement system. *Trends Immunol.* 31, 154–163. doi: 10.1016/j.it.2010.01.002
- Hsu, C.-L., and Yang, U.-C. (2012). Discovering pathway cross-talks based on functional relations between pathways. *BMC Genomics* 13(Suppl. 7), S25. doi: 10.1186/1471-2164-13-S7-S25
- Hu, Y.-S., Xin, J., Hu, Y., Zhang, L., and Wang, J. (2017). Analyzing the genes related to Alzheimer's disease via a network and pathway-based approach. *Alzheimer's Res. Ther.* 9:29. doi: 10.1186/s13195-017-0252-z

- Juhász, G., Földi, I., and Penke, B. (2011). Systems biology of Alzheimer's disease: how diverse molecular changes result in memory impairment in AD. *Neurochem. Int.* 58, 739–750. doi: 10.1016/j.neuint.2011.02.008
- Lanni, C., Uberti, D., Racchi, M., Govoni, S., and Memo, M. (2007). Unfolded p53: A potential biomarker for Alzheimer's disease. *J. Alzheimer's Dis.* 12, 93–99. doi: 10.3233/JAD-2007-12109
- Li, Y., Agarwal, P., and Rajagopalan, D. (2008). A global pathway crosstalk network. *Bioinformatics* 24, 1442–1447. doi: 10.1093/bioinformatics/btn200
- Liu, H., Zhou, X., Jiang, H., He, H., Liu, X., Weiner, M. W., et al. (2016). A semi-mechanism approach based on MRI and proteomics for prediction of conversion from mild cognitive impairment to Alzheimer's disease. *Sci. Rep.* 6:26712. doi: 10.1038/srep26712
- Liu, Z.-P., Wang, Y., Zhang, X.-S., and Chen, L. (2010). Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst. Biol.* 4(Suppl. 2), S11. doi: 10.1186/1752-0509-4-S2-S11
- Mukherjee, S., Kaerberlein, M., Kauwe, J., Naj, A. C., and Crane, P. (2014). A systems-biology approach to identify candidate genes for Alzheimer's disease by integrating protein-protein interaction network and subsequent *in vivo* validation of candidate genes using a *C. elegans* model of ab toxicity. *Alzheimer's Dement.* 10, P298–P299. doi: 10.1016/j.jalz.2014.04.499
- Myers, C. L., Robson, D., Wible, A., Hibbs, M., Chiriack, C., Theesfeld, C. L., et al. (2005). Discovery of biological networks from diverse functional genomic data. *Genome Biol.* 6:R114. doi: 10.1186/gb-2005-6-13-r114
- Oeckinghaus, A., Hayden, M. S., and Ghosh, S. (2011). Crosstalk in NF- κ B signaling pathways. *Nat. Immunol.* 12, 695–708. doi: 10.1038/ni.2065
- Selkoe, D. J. (2001). Alzheimer's disease: genes, proteins, and therapy. *Physiol. Rev.* 81, 741–766.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122, 957–968. doi: 10.1016/j.cell.2005.08.029
- Tegge, A. N., Sharp, N., and Murali, T. M. (2015). Xtalk: a path-based approach for identifying crosstalk between signaling pathways. *Bioinformatics* 32, 242–251. doi: 10.1093/bioinformatics/btv549
- Wang, T., Gu, J., Yuan, J., Tao, R., Li, Y., and Li, S. (2013). Inferring pathway crosstalk networks using gene set co-expression signatures. *Mol. Biosyst.* 9, 1822–1828. doi: 10.1039/c3mb25506a
- Xu, Y., Hu, W., Chang, Z., DuanMu, H., Zhang, S., Li, Z., et al. (2011). Prediction of human protein-protein interaction by a mixed Bayesian model and its application to exploring underlying cancer-related pathway crosstalk. *J. R. Soc. Interface* 8, 555–567. doi: 10.1098/rsif.2010.0384

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Padmanabhan, Shpanskaya, Bello, Doraiswamy and Samatova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.