

# ResponseNet2.0: revealing signaling and regulatory pathways connecting your proteins and genes—now with human data

Omer Basha<sup>1</sup>, Shoval Tirman<sup>1</sup>, Amir Eluk<sup>1</sup> and Esti Yeger-Lotem<sup>1,2,\*</sup>

<sup>1</sup>Department of Clinical Biochemistry & Pharmacology, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel and <sup>2</sup>National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

Received February 26, 2013; Revised May 8, 2013; Accepted May 22, 2013

## ABSTRACT

Genome sequencing and transcriptomic profiling are two widely used approaches for the identification of human disease pathways. However, each approach typically provides a limited view of disease pathways: Genome sequencing can identify disease-related mutations but rarely reveals their mode-of-action, while transcriptomic assays do not reveal the series of events that lead to the transcriptomic change. ResponseNet is an integrative network-optimization approach that we developed to fill these gaps by highlighting major signaling and regulatory molecular interaction paths that connect disease-related mutations and genes. The ResponseNet web-server provides a user-friendly interface to ResponseNet. Specifically, users can upload weighted lists of proteins and genes and obtain a sparse, weighted, molecular interaction subnetwork connecting them, that is biased toward regulatory and signaling pathways. ResponseNet2.0 enhances the functionality of the ResponseNet web-server in two important ways. First, it supports analysis of human data by offering a human interactome composed of proteins, genes and micro-RNAs. Second, it offers a new informative view of the output, including a randomization analysis, to help users assess the biological relevance of the output subnetwork. ResponseNet2.0 is available at <http://netbio.bgu.ac.il/respnet>.

## INTRODUCTION

Massive efforts are being invested worldwide in cataloguing the mutations and transcriptomic changes characterizing a large variety of human diseases to

identify the cellular pathways involved in each disease (1–4). These wealth of data provide a promising starting point for unraveling disease pathways. However, the functions of many of the identified mutations and the signaling pathways that lead to altered transcriptional regulation often remain elusive.

Molecular interaction networks (interactomes), where nodes represent molecules such as genes and proteins and edges represent their various inter-relationships, offer a powerful framework for enhancing our understanding of protein functions and the cellular processes underlying diseases (5). First, molecular interactions govern biological processes. Being the union of these interactions, an interactome provides a skeleton from which the functions of proteins and the organization of pathways can be inferred (6–9). Second, because edges represent molecular relationships, the road from an interactome-based hypothesis to experimental testing is short (10–14). These observations motivated a wide variety of interactome-based approaches for shedding light on disease genes and pathways (15–19).

ResponseNet is an integrative interactome-based approach that uses known molecular interactions to bridge the gap between condition-specific mutations and transcriptomic changes, revealing a broader view of the underlying cellular processes (12). Specifically, given weighted lists of proteins and genes related to a specific condition, ResponseNet identifies a sparse high-probability molecular interaction subnetwork by which the input proteins may lead to the altered transcription of input target genes. This is achieved by formulating a minimum-cost flow optimization problem that is solved by linear programming. By applying ResponseNet to data of large-scale genetic and transcriptomic screens of a yeast disease model, we successfully mapped recognized disease pathways and exposed previously hidden pathways that we validated experimentally (12).

The ResponseNet web-server that we reported previously enabled users to meaningfully integrate their data

\*To whom correspondence should be addressed. Tel/Fax: +972 8 642 8675; Email: [estiyl@bgu.ac.il](mailto:estiyl@bgu.ac.il)

and to substantially expand their understanding of the cellular conditions they study (20). Specifically, users could upload weighted lists of proteins and genes and obtain the connecting output subnetwork. Here we present ResponseNet2.0, a new version of the ResponseNet web-server that features enhanced functionality. We first describe the extension of ResponseNet toward the analysis of human pathways, including evaluation of its performance over manually curated human pathways. We then describe new features of ResponseNet2.0 that help users assess the biological relevance of ResponseNet results.

## ResponseNet2.0: ANALYSIS OF HUMAN PATHWAYS

ResponseNet was originally developed to analyze data gathered from budding yeast (12). The ResponseNet web-server supported analysis of yeast data by providing a weighted model of the yeast interactome, which consisted of physical and regulatory interactions among yeast proteins and genes. Analysis of data from other organisms was also supported, given that users upload their corresponding interactomes. ResponseNet2.0 extends ResponseNet by offering, in addition, a weighted model of the human interactome. Similarly to the interactome of budding yeast, the human interactome contains physical and regulatory interactions among human proteins and genes. Yet unlike yeast, it also contains interactions involving micro-RNAs (miRs), in accordance with their significant roles in regulating a large variety of cellular processes in health and disease (21).

### Construction of a weighted model of the human interactome

We gathered experimentally identified interactions from several databases. Protein–protein interactions (PPIs) were gathered from BioGRID (22), DIP (23), MINT (24) and IntAct (25) using the PSICQUIC interface (26). These interactions were then weighted using a weighting scheme that works in two steps, as previously described (12). First, given a specific subset of biological processes, in this case signaling pathways, the scheme calculates a score for each interaction-detection method, which reflects the ability of the method to identify process-related interactions (see Methods). Second, the scheme assigns a weight for every interaction according to the interaction-detection methods that identified it, using a Bayesian computation. The scheme therefore assigns higher weights to interactions identified by multiple interaction-detection methods with a bias toward certain methods, and is not limited to known pathways. A histogram describing the distribution of PPI weights is provided in Supplementary Figure S1. In addition to PPI, transcription regulation interactions between transcription factors and their target genes were gathered from TRANSFAC (27). Lastly, miR interactions were gathered from TarBase (28), miRecords (29) and TransmiR (30). All regulatory interactions were associated with uniform weights. The resulting weighted human

interactome contained 88 016 experimentally identified interactions among 11 347 proteins, 426 miRs and 10 014 target genes.

### ResponseNet2.0 evaluation over human cellular pathways

To assess the validity and performance of ResponseNet2.0, we applied it to data from manually curated human pathways. To this end, we used the SPIKE database, which contained 24 maps of human cellular pathways (31). The analysis was performed as follows: From each pathway map, we extracted an input source set consisting of upstream pathway proteins and an input target set consisting of downstream pathway genes. We applied ResponseNet2.0 to these inputs and then compared between the output subnetwork predicted by ResponseNet2.0 and the original pathway map. In addition to this ‘full’ map analysis, we repeated the analysis for different path lengths per map. For every length  $k = 1, \dots, 12$ , the input target set was left intact and the input source set consisted of proteins with a shortest distance of  $k$  from the target set. The median input and output sizes per analysis are presented in Table 1. Figure 1 shows the performance of ResponseNet2.0 as measured by its sensitivity and precision over nodes and interactions (specificity was always 0.99 owing to the large number of true negatives relative to false positives). In general, ResponseNet’s performance was highest for shortest paths and decreased with path length. The sensitivity and precision were much higher for nodes than for interactions, implying that ResponseNet often revealed the relevant proteins but not necessarily their pathway connectivity. Evaluation of ResponseNet analysis per map and upon using different values of gamma and capping are presented in Supplementary Tables S1–S3. Notably, ResponseNet can be applied to source and target sets of various sizes and to networks containing  $>10^4$  nodes and  $>10^5$  edges. A typical run of ResponseNet2.0 over source and target sets containing 20 protein and 500 genes took about 3 minutes when executed locally.

An example of ResponseNet2.0 output is shown in Figure 2. In this case, we used ResponseNet to identify the regulatory subnetwork connecting the melanoma-related protein ‘BRCA1 associated protein-1’ (BAP1), which when mutated leads to uveal melanoma, to 693 target genes that were found to be down-regulated in melanoma cell lines (32). ResponseNet2.0 correctly predicted the main interactor of BAP1, the HCFC1 protein, and highlighted two connecting transcription factors, GABPA and SP1, the latter previously linked to melanoma (33).

### ResponseNet 2.0: NEW FEATURES

ResponseNet2.0 contains several features that we added to help users assess the biological relevance of ResponseNet results.

#### Informative presentation of ResponseNet results

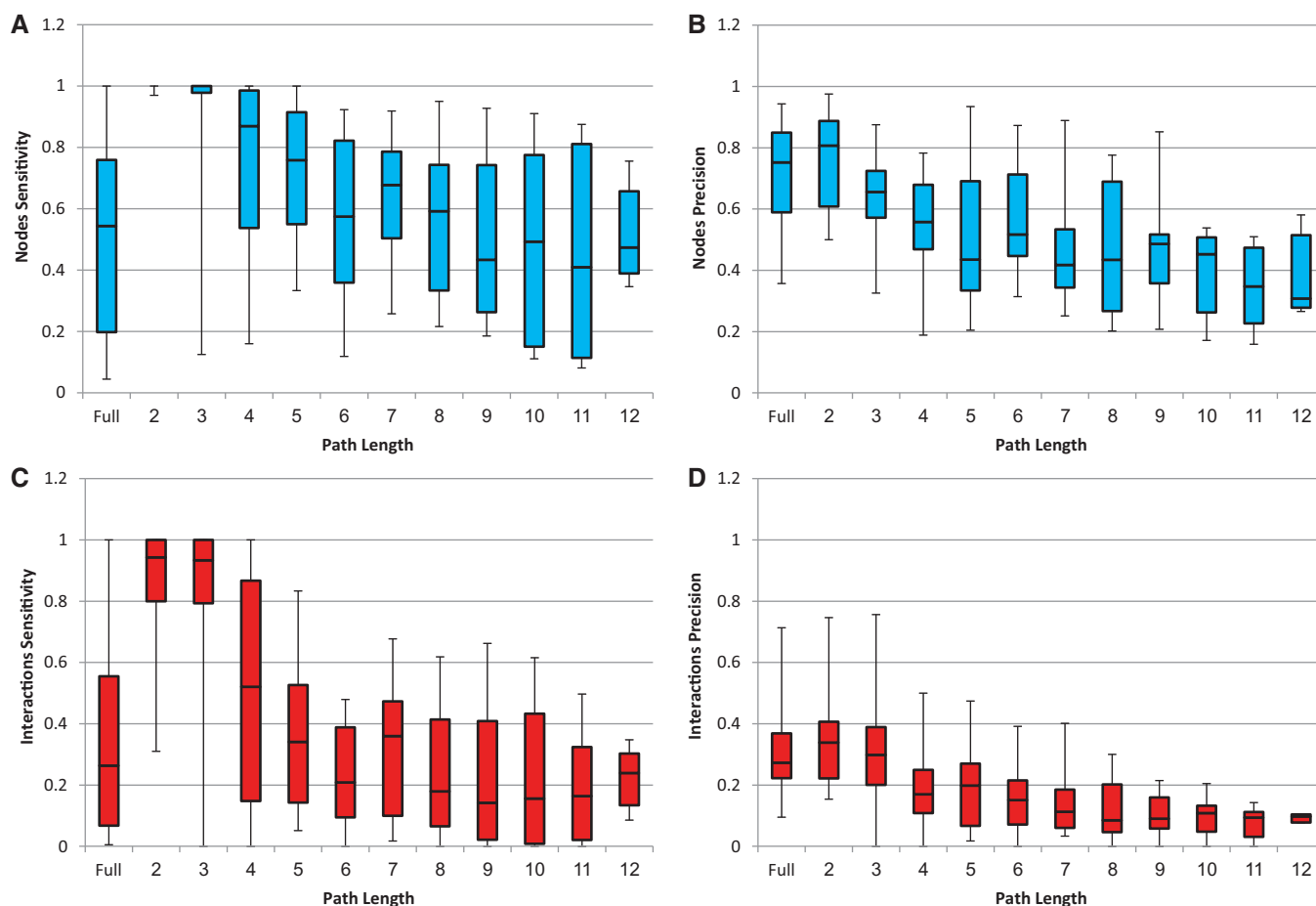
The main output of ResponseNet is a subnetwork that is presented graphically using Cytoscape-web (34).

**Table 1.** The scope of ResponseNet analysis of known human pathways by using SPIKE maps

| Paths length | Number of maps tested | Input size <sup>a</sup> (source, target) | Map size <sup>a</sup> (nodes, edges) | ResponseNet output size <sup>a</sup> (nodes, edges) | Overlap between map and output <sup>a</sup> (nodes, edges) |
|--------------|-----------------------|--|--------------------------------------|---|--|
| Full         | 24                    | 6, 13                                    | 73.5, 127.5                          | 31, 51.5  | 23, 19   |
| 2            | 24                    | 9.5, 13                                  | 27.5, 22                             | 31.5, 64.5  | 27.5, 19   |
| 3            | 22                    | 9, 17.5                                  | 27, 33                               | 43.5, 98  | 27, 27   |
| 4            | 20                    | 12.5, 17.5                               | 41, 64                               | 54, 134.50  | 31, 19   |
| 5            | 17                    | 19, 22                                   | 62, 97                               | 61, 128   | 31, 40   |
| 6            | 14                    | 23, 13                                   | 69, 124                              | 47, 104.50  | 40, 15   |
| 7            | 14                    | 13, 13                                   | 64.5, 88.5                           | 57.5, 117   | 39.5, 21.5   |
| 8            | 11                    | 12, 13                                   | 70, 99                               | 74, 150   | 45, 30   |
| 9            | 11                    | 16, 13                                   | 66, 93                               | 68, 169   | 37, 27   |
| 10           | 9                     | 10, 22                                   | 52, 116                              | 68, 130   | 35, 20   |
| 11           | 6                     | 4, 27.5                                  | 45.5, 83.5                           | 57.5, 127   | 14, 8.5  |
| 12           | 5                     | 6, 33                                    | 38, 70                               | 52, 107   | 18, 11   |
| 13           | 4                     | 8, 48.5                                  | 53.5, 108.5                          | 46.5, 99  | 13, 4.5  |
| 14           | 3                     | 4, 33                                    | 39, 73                               | 14, 18  | 7, 1   |
| 15           | 2                     | 3, 17.5                                  | 33, 52                               | 27.5, 44.5  | 7.5, 2.5   |

<sup>a</sup>Median values are shown.

ResponseNet was executed with gamma of 2.5 and capping of 0.8. The sensitivity and precision for each type of run are shown in Figure 1.

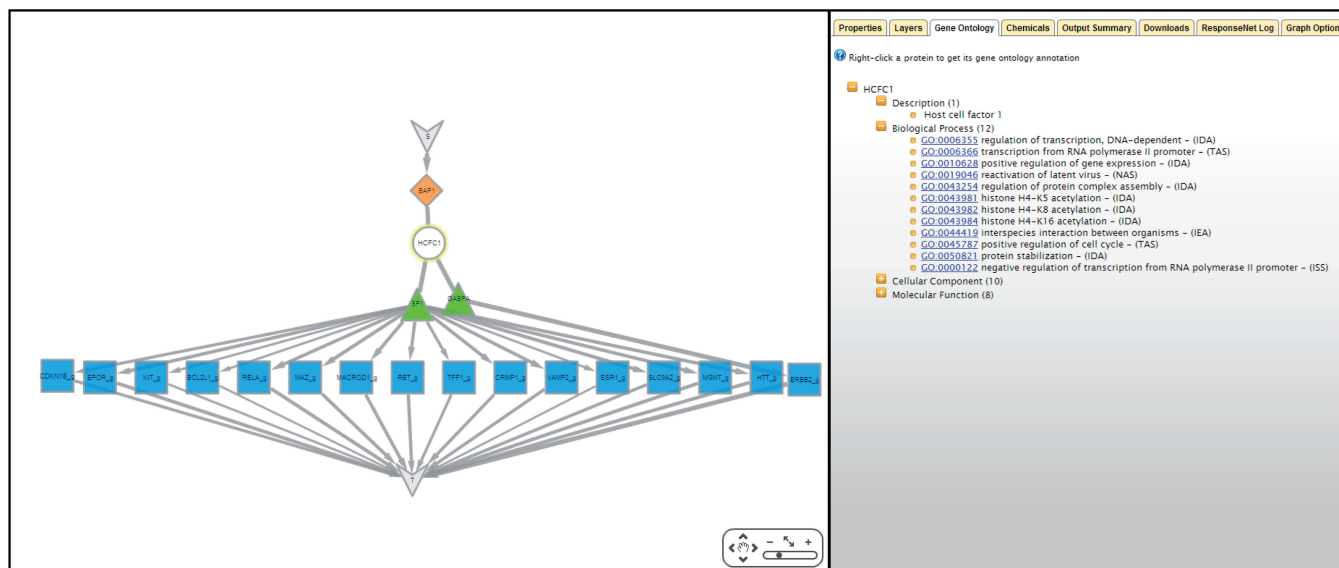


**Figure 1.** ResponseNet sensitivity and precision upon analysis of known human pathways. Results are shown for full maps (denoted 'Full') and for different path lengths (lengths >12 involved <5 maps and were therefore ignored). Each box-plot diagram shows the quartile values (25, 50 and 75%). (A) Nodes sensitivity. (B) Nodes precision. (C) Interactions sensitivity. (D) Interactions precision. Specificity was 99% in every case and is therefore not shown.

## ResponseNet 2.0



Now offering human interactome



**Figure 2.** An example of ResponseNet2.0 output. The human protein BAP1 is a deubiquitinating enzyme that acts as a regulator of cell growth by mediating deubiquitination of HCFC1, and is frequently mutated in uveal melanomas. We used ResponseNet2.0 to identify a regulatory subnetwork that connects BAP1 to human genes that were found to be down-regulated in melanoma cell lines. ResponseNet correctly predicted the BAP1-interacting protein HCFC1, and connected HCFC1 to two transcription factors, GABPA and SP1, that regulate the transcription of 16 target genes. Notably, SP1 was previously linked to melanoma (33). S and T are auxiliary nodes that are part of ResponseNet formulation.

ResponseNet2.0 enhances this output by providing a new tabs menu, where each tab provides additional information regarding the output (Figure 2): The ‘Gene Ontology’ (GO) tab provides the GO annotations and evidence codes of subnetwork proteins and transcripts (35). The ‘Chemicals’ tab provides data of small molecules known to bind subnetwork proteins, which may be used in the design of future experiments (36). The ‘Properties’ tab provides information regarding subnetwork interactions, including the method(s) by which they were detected and the database that reported them. Other tabs enable users to download data, remove subnetwork nodes, change the network layout, and more.

### Randomization analysis

This novel feature in ResponseNet2.0 enables users to estimate the probability of observing subnetwork components by chance. Upon executing ResponseNet, users can choose whether to randomize the source set, the target set or both. In each randomization run, an input set of size equal to the original input set is selected randomly from the interactome, and ResponseNet is executed with all other data and parameters as in the original run. The results of the randomized runs are presented as a table in the ‘Properties’ tab. The table lists for each node in the original output the number of times that this node was predicted in the randomized runs. A lower number of appearances thus indicates higher statistical significance. Owing to the computational complexity of

ResponseNet, we currently support up to 10 randomized runs per input set.

### Session maintenance and subnetwork comparisons

ResponseNet users typically execute ResponseNet analyses anonymously. In ResponseNet2.0, we also provide users with an option to login to ResponseNet, store their sessions and load previous sessions. Notably, users can analyze and compare between their sessions using a layers tabs. A new layer can be created by uploading another network in GraphML format, or by creating a new layer from the existing network. Once a layer was created, the layers tab allows users to unite layers, intersect them to identify common components, or hide common components and show network differences. The layers interface can be used to evaluate runs with different parameters, and to identify common versus unique subnetwork components.

### SUMMARY

The post-genomic era is marked by massive mapping of the genome, transcriptome, methylome, microbiome and other ‘omes’ of healthy and sick individuals (19). Owing to initiatives such as The Cancer Genome Atlas and the 1000 Genomes Project, these data are accumulating at a high pace (37). However, the next important step in the analysis, namely deciphering these large-scale data, often becomes the limiting step. This calls for the development



of web-servers that offer users meaningful analysis of large-scale data. By taking into account known molecular interactions, tools such as ResponseNet (20) and SteinerNet (38) shed light on the roles of mutations and uncover additional components of the underlying pathways.

ResponseNet2.0 is a new version of the ResponseNet web-server that features richer user interface and offers analysis of human pathways. ResponseNet users can now upload lists of human proteins, genes and miRs, and obtain a high-probability subnetwork of the human interactome that connects them (Figure 2). Our evaluation of ResponseNet using manually curated human pathways revealed that ResponseNet recovered pathway nodes at a median sensitivity, specificity and precision of 54%, 99% and 75%, respectively. Notably, these results are almost as high as ResponseNet performance over yeast pathways (12). As interactions data improve, we anticipate that tools such as ResponseNet will become standard in the analysis of the accumulating state-of-the-art disease data, and provide the much needed views into the currently hidden mechanisms underlying human diseases.

## METHODS

### PPI weighting scheme

The first step in the scheme is the scoring of interaction-detection methods based on their ability to detect interactions within signaling and response pathways. To this end, we created positive and negative sets of PPIs as follows. The positive set included PPIs among members of a common signaling and regulatory process, where processes were extracted from Gene Ontology (35). The negative set included PPIs between proteins that were not annotated to any of these processes. The score of each interaction-detection method was then computed based on the ratio of positive to negative interactions that the detection method identified. In the second step of the scheme, the weight of each PPI was calculated according to the interaction-detection methods that identified it using a Bayesian computation. More details about the scheme can be found elsewhere (12).

### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3 and Supplementary Figure 1.

### ACKNOWLEDGEMENTS

We thank Ilan Y. Smoly and Martin Ramzi Kahil for help in integrating protein interactions data. We are grateful to all members of the Yeger-Lotem lab for helpful comments on the design of ResponseNet2.0.

### FUNDING

European Union Seventh Programme under the FP7-PEOPLE-MCA-IRG Funding scheme [256360 to E.Y.-L.], and the Israel Cancer Institute (to E.Y.-L.).

Funding for open access charge: Israel Cancer Institute (to E.Y.-L.).

*Conflict of interest statement.* None declared.

## REFERENCES

- Hudson,T.J., Anderson,W., Artez,A., Barker,A.D., Bell,C., Bernabe,R.R., Bhan,M.K., Calvo,F., Eerola,I., Gerhard,D.S. *et al.* (2010) International network of cancer genome projects. *Nature*, **464**, 993–998.
- Pankratz,N., Wilk,J.B., Latourelle,J.C., DeStefano,A.L., Halter,C., Pugh,E.W., Doheny,K.F., Gusella,J.F., Nichols,W.C., Foroud,T. *et al.* (2009) Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *Hum. Genet.*, **124**, 593–605.
- Olgiati,P., Politis,A.M., Papadimitriou,G.N., De Ronchi,D. and Serretti,A. (2011) Genetics of late-onset Alzheimer's disease: update from the alzgene database and analysis of shared pathways. *Int. J. Alzheimers Dis.*, **2011**, 832379.
- Pleasance,E.D., Cheetham,R.K., Stephens,P.J., McBride,D.J., Humphray,S.J., Greenman,C.D., Varela,I., Lin,M.L., Odonez,G.R., Bignell,G.R. *et al.* (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*, **463**, 191–196.
- Vidal,M., Cusick,M.E. and Barabasi,A.L. (2011) Interactome networks and human disease. *Cell*, **144**, 986–998.
- Schwikowski,B., Uetz,P. and Fields,S. (2000) A network of protein-protein interactions in yeast. *Nat. Biotechnol.*, **18**, 1257–1261.
- Kelley,R. and Ideker,T. (2005) Systematic interpretation of genetic interactions using protein networks. *Nat. Biotechnol.*, **23**, 561–566.
- Sharan,R., Ulitsky,I. and Shamir,R. (2007) Network-based prediction of protein function. *Mol. Syst. Biol.*, **3**, 88.
- Huttenhower,C., Haley,E.M., Hibbs,M.A., Dumeaux,V., Barrett,D.R., Collier,H.A. and Troyanskaya,O.G. (2009) Exploring the human genome with functional maps. *Genome Res.*, **19**, 1093–1106.
- Pujana,M.A., Han,J.D., Starita,L.M., Stevens,K.N., Tewari,M., Ahn,J.S., Rennert,G., Moreno,V., Kirchhoff,T., Gold,B. *et al.* (2007) Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat. Genet.*, **39**, 1338–1349.
- Schadt,E.E., Friend,S.H. and Shaywitz,D.A. (2009) A network view of disease and compound screening. *Nat. Rev. Drug Discov.*, **8**, 286–295.
- Yeger-Lotem,E., Riva,L., Su,L.J., Gitler,A.D., Cashikar,A.G., King,O.D., Auluck,P.K., Geddie,M.L., Valastyan,J.S., Karger,D.R. *et al.* (2009) Bridging high-throughput genetic and transcriptional data reveals cellular responses to alpha-synuclein toxicity. *Nat. Genet.*, **41**, 316–323.
- Shapira,S.D., Gat-Viks,I., Shum,B.O., Dricot,A., de Grace,M.M., Wu,L., Gupta,P.B., Hao,T., Silver,S.J., Root,D.E. *et al.* (2009) A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell*, **139**, 1255–1267.
- Frezza,C., Zheng,L., Folger,O., Rajagopalan,K.N., MacKenzie,E.D., Jerby,L., Micaroni,M., Chaneton,B., Adam,J., Hedley,A. *et al.* (2011) Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature*, **477**, 225–228.
- Wang,Y. and Xia,Y. (2008) *The Second International Symposium on Optimization and Systems Biology (OSB'08)*. Lijiang, China, pp. 333–340.
- Ideker,T. and Sharan,R. (2008) Protein networks in disease. *Genome Res.*, **18**, 644–652.
- Barabasi,A.L., Gulbahce,N. and Loscalzo,J. (2011) Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.*, **12**, 56–68.
- Ma,H., Schadt,E.E., Kaplan,L.M. and Zhao,H. (2011) COSINE: COndition-Specific sub-NEtwork identification using a global optimization method. *Bioinformatics*, **27**, 1290–1298.
- Huang,S.S. and Fraenkel,E. (2012) Swimming upstream: identifying proteomic signals that drive transcriptional changes

- using the interactome and multiple “-omics” datasets. *Methods Cell Biol.*, **110**, 57–80.
20. Lan, A., Smoly, I.Y., Rapaport, G., Lindquist, S., Fraenkel, E. and Yeger-Lotem, E. (2011) ResponseNet: revealing signaling and regulatory networks linking genetic and transcriptomic screening data. *Nucleic Acids Res.*, **39**, W424–W429.
  21. Mendell, J.T. and Olson, E.N. (2012) MicroRNAs in stress signaling and human disease. *Cell*, **148**, 1172–1187.
  22. Stark, C., Breitkreutz, B.J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M.S., Nixon, J., Van Auken, K., Wang, X., Shi, X. *et al.* (2011) The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.*, **39**, D698–D704.
  23. Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U. and Eisenberg, D. (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
  24. Ceol, A., Chatr-Aryamontri, A., Licata, L., Peluso, D., Briganti, L., Perfetto, L., Castagnoli, L. and Cesareni, G. (2010) MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res.*, **38**, D532–D539.
  25. Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A.T., Kerrien, S., Khadake, J. *et al.* (2010) The IntAct molecular interaction database in 2010. *Nucleic Acids Res.*, **38**, D525–D531.
  26. Aranda, B., Blankenburg, H., Kerrien, S., Brinkman, F.S., Ceol, A., Chautard, E., Dana, J.M., De Las Rivas, J., Dumousseau, M., Galeota, E. *et al.* (2011) PSICQUIC and PSISCORE: accessing and scoring molecular interactions. *Nat. Methods*, **8**, 528–529.
  27. Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K. *et al.* (2006) TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.*, **34**, D108–D110.
  28. Vergoulis, T., Vlachos, I.S., Alexiou, P., Georgakilas, G., Maragkakis, M., Reczko, M., Gerangelos, S., Koziris, N., Dalamagas, T. and Hatzigeorgiou, A.G. (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res.*, **40**, D222–D229.
  29. Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X. and Li, T. (2009) miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.*, **37**, D105–D110.
  30. Wang, J., Lu, M., Qiu, C. and Cui, Q. (2010) TransmiR: a transcription factor-microRNA regulation database. *Nucleic Acids Res.*, **38**, D119–D122.
  31. Paz, A., Brownstein, Z., Ber, Y., Bialik, S., David, E., Sagir, D., Ulitsky, I., Elkou, R., Kimchi, A., Avraham, K.B. *et al.* (2011) SPIKE: a database of highly curated human signaling pathways. *Nucleic Acids Res.*, **39**, D793–D799.
  32. Broad Melanoma Portal Release 1.0 <http://www.broadinstitute.org/melanoma/pages/publicPortalHome.jsf>.
  33. Karlen, S. and Braathen, L.R. (1999) Regulation of the melanoma cell adhesion molecule gene in melanoma: modulation of mRNA synthesis by cyclic adenosine monophosphate, phorbol ester, and stem cell factor/c-kit signaling. *J. Invest. Dermatol.*, **113**, 711–719.
  34. Lopes, C.T., Franz, M., Kazi, F., Donaldson, S.L., Morris, Q. and Bader, G.D. (2010) Cytoscape Web: an interactive web-based network browser. *Bioinformatics*, **26**, 2347–2348.
  35. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
  36. Kuhn, M., Szklarczyk, D., Franceschini, A., von Mering, C., Jensen, L.J. and Bork, P. (2012) STITCH 3: zooming in on protein-chemical interactions. *Nucleic Acids Res.*, **40**, D876–D880.
  37. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E. and McVean, G.A. (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.
  38. Tuncbag, N., McCallum, S., Huang, S.S. and Fraenkel, E. (2012) SteinerNet: a web server for integrating ‘omic’ data to discover hidden components of response pathways. *Nucleic Acids Res.*, **40**, W505–W509.