Supplementary Materials

Table S1. The Most Important Active Databases on the PPI

Acronym	Organisms	Number of Interaction	URL/FTP and References
3did/Interprets	global	367277	http://3did.irbbarcelona.org/ [1]
AtPIN	*A. thaliana*	28,062	http://www.megabionet.org/atpid [2, 3]
BindingDB	Global	793,068	http://www.bindingdb.org/bind [4, 5]
CBM	global	2784	ftp://ftp.ncbi.nlm.nih.gov/pub/cbm [6]
DIMA	Global	107436	http://webclu.bio.wzw.tum.de/dima/ [7]
DroID	Drosophila	520,978	http://www.droidb.org/ [8, 9]
GWIDD	Global	126 897	http://gwidd.bioinformatics.ku.edu/ [10]
HAPPI	Human	601,757	http://bio.informatics.iupui.edu/HAPPI/ [11]
HCPIN	Human Cancer associated Protein Interaction	9784	http://nesg.org:9090/HCPIN/ [12]
HiPredict	Model organisms	176983	http://hintdb.hgc.jp/htp/ [13, 14]
HIV PI Database	Human immunodeficiency virus type 1	2589	http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions [15-17]
HomoMINT	inferred human network from model organism	330377	http://mint.bio.uniroma2.it/HomoMINT [18]
hp-DPI	*Helicobacter Pylori*	?	http://dpi.nhri.org.tw/protein/hp [19]
HUGE ppi	interactions between large proteins of human	84	http://www.kazusa.or.jp/huge/ppi/ [20]
I2D (IMEx partner)	mammalian and eukaryotic PPI	681,404	http://ophid.utoronto.ca [21, 22]
IBIS	Global?	205090	http://www.ncbi.nlm.nih.gov/Structure/ibis/ibis.cgi [23]
InterDom	Global	5511	http://interdom.i2r.a-star.edu.sg/ [24]
iPfam	Global	2733	http://ipfam.sanger.ac.uk/ [25]
MPact/MIPS (IMEx partner)	S. cerevisiae	4500	http:// mips.gsf.de/genre/proj/mpact/ [26]
MPIDB (IMEx partner)	Microbial PPI	24,295	www.jcvi.org/mpidb [27]
MPPI	Mammals	about 1000	http://mips.helmholtz-muenchen.de/proj/ppi/ [28]
NetPro	Global	320,000	http://www.molecularconnections.com/home/en/home/products/NetPro
PepCyber :P~Pep	human PPI mediated by phosphoprotein binding domains	1858	http://www.pepcyber.org/PPEP/ [29]
PIPs	human	79441	http://www.compbio.dundee.ac.uk/www-pips/ [30]
POINT	Global	78125	http://point.bioinformatics.tw [31]
Prolinks	Global	34,924,6087	http://prl.mbi.ucla.edu/prlbeta/prolinks.jsp [32]

(Table 1) contd….

Acronym	Organisms	Number of Interaction	URL/FTP and References
SCOPPI	Global	105547	http://www.scoppi.org/ [33]
STRING	Global	>200,000,000 (predicted interactions) fa	http://string.embl.de [34-36]
iRefIndex	Global	544908	http://irefindex.org [37]
UniHI	Human, WORM, FLY, MOUSE, YEAST	36023	http://www.unihi.org/ [38]
Yeast Interacting Proteins Database	Yeast	4,549	http://itolab.cb.k.u-tokyo.ac.jp/Y2H/ [39]

REFERENCES

[1] Stein, A.; Ceol, A.; Aloy, P., 3did: identification and classification of domain-based interactions of known three-dimensional structure. *Nucleic acids research* **2011,** *39* (Database issue), D718-23.

[2] Li, P.; Zang, W.; Li, Y.; Xu, F.; Wang, J.; Shi, T., AtPID: the overall hierarchical functional protein interaction network interface and analytic platform for Arabidopsis. *Nucleic acids research* **2011,** *39* (Database issue), D1130-3.

[3] Cui, J.; Li, P.; Li, G.; Xu, F.; Zhao, C.; Li, Y.; Yang, Z.; Wang, G.; Yu, Q.; Shi, T., AtPID: Arabidopsis thaliana protein interactome database--an integrative platform for plant systems biology. *Nucleic acids research* **2008,** *36* (Database issue), D999-1008.

[4] Chen, X.; Lin, Y.; Liu, M.; Gilson, M. K., The Binding Database: data management and interface design. *Bioinformatics* **2002,** *18* (1), 130-9.

[5] Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K., BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic acids research* **2007,** *35* (Database issue), D198-201.

[6] Shoemaker, B. A.; Panchenko, A. R.; Bryant, S. H., Finding biologically relevant protein domain interactions: conserved binding mode analysis. *Protein Sci* **2006,** *15* (2), 352-61.

[7] Luo, Q.; Pagel, P.; Vilne, B.; Frishman, D., DIMA 3.0: Domain Interaction Map. *Nucleic acids research* **2011,** *39* (Database issue), D724-9.

[8] Murali, T.; Pacifico, S.; Yu, J.; Guest, S.; Roberts, G. G., 3rd; Finley, R. L., Jr., DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. *Nucleic acids research* **2011,** *39* (Database issue), D736-43.

[9] Yu, J.; Pacifico, S.; Liu, G.; Finley, R. L., Jr., DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. *BMC genomics* **2008,** *9*, 461.

[10] Kundrotas, P. J.; Zhu, Z.; Vakser, I. A., GWIDD: Genome-wide protein docking database. *Nucleic acids research* **2010,** *38* (Database issue), D513-7.

[11] Chen, J. Y.; Mamidipalli, S.; Huan, T., HAPPI: an online database of comprehensive human annotated and predicted protein interactions. *BMC genomics* **2009,** *10 Suppl 1*, S16.

[12] Huang, Y. J.; Hang, D.; Lu, L. J.; Tong, L.; Gerstein, M. B.; Montelione, G. T., Targeting the human cancer pathway protein interaction network by structural genomics. *Mol Cell Proteomics* **2008,** *7* (10), 2048-60.

[13] Patil, A.; Nakamura, H., Filtering high-throughput protein-protein interaction data using a combination of genomic features. *BMC bioinformatics* **2005,** *6*, 100.

[14] Patil, A.; Nakai, K.; Kinoshita, K., Assessing the utility of gene co-expression stability in combination with correlation in the analysis of protein-protein interaction networks. *BMC genomics* **2011,** *12 Suppl 3*, S19.

[15] Fu, W.; Sanders-Beer, B. E.; Katz, K. S.; Maglott, D. R.; Pruitt, K. D.; Ptak, R. G., Human immunodeficiency virus type 1, human protein interaction database at NCBI. *Nucleic acids research* **2009,** *37* (Database issue), D417-22.

[16] Ptak, R. G.; Fu, W.; Sanders-Beer, B. E.; Dickerson, J. E.; Pinney, J. W.; Robertson, D. L.; Rozanov, M. N.; Katz, K. S.; Maglott, D. R.; Pruitt, K. D.; Dieffenbach, C. W., Cataloguing the HIV type 1 human protein interaction network. *AIDS research and human retroviruses* **2008,** *24* (12), 1497-502.

[17] Pinney, J. W.; Dickerson, J. E.; Fu, W.; Sanders-Beer, B. E.; Ptak, R. G.; Robertson, D. L., HIV-host interactions: a map of viral perturbation of the host system. *AIDS (London, England)* **2009,** *23* (5), 549-54.

[18] Persico, M.; Ceol, A.; Gavrila, C.; Hoffmann, R.; Florio, A.; Cesareni, G., HomoMINT: an inferred human network based on orthology mapping of protein interactions discovered in model organisms. *BMC bioinformatics* **2005,** *1* (6).

[19] Lin, C. Y.; Chen, C. L.; Cho, C. S.; Wang, L. M.; Chang, C. M.; Chen, P. Y.; Lo, C. Z.; Hsiung, C. A., hp-DPI: Helicobacter pylori database of protein interactomes--embracing experimental and inferred interactions. *Bioinformatics* **2005,** *21* (7), 1288-90.

[20] Nakayama, M.; Kikuno, R.; Ohara, O., Protein-protein interactions between large proteins: two-hybrid screening using a functionally classified library composed of long cDNAs. *Genome Res* **2002,** *12* (11), 1773-84.

[21] Brown, K. R.; Jurisica, I., Online predicted human interaction database. *Bioinformatics* **2005,** *21* (9), 2076-82.

[22] Brown, K. R.; Jurisica, I., Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol* **2007,** *8* (5), R95.

[23] Shoemaker, B. A.; Zhang, D.; Thangudu, R. R.; Tyagi, M.; Fong, J. H.; Marchler-Bauer, A.; Bryant, S. H.; Madej, T.; Panchenko, A. R., Inferred Biomolecular Interaction Server--a web server to analyze and predict protein interacting partners and binding sites. *Nucleic acids research* **2010,** *38* (Database issue), D518-24.

[24] Ng, S. K.; Zhang, Z.; Tan, S. H.; Lin, K., InterDom: a database of putative interacting protein domains for validating predicted protein interactions and complexes. *Nucleic acids research* **2003,** *31* (1), 251-4.

[25] Finn, R. D.; Marshall, M.; Bateman, A., iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. *Bioinformatics* **2005,** *21* (3), 410-2.

[26] Guldener, U.; Munsterkotter, M.; Oesterheld, M.; Pagel, P.; Ruepp, A.; Mewes, H.; Stumpflen, V., MPact: the MIPS protein interaction resource on yeast. *Nucleic acids research* **2006,** *34*, D436 - D441.

[27] Goll, J.; Rajagopala, S. V.; Shiau, S. C.; Wu, H.; Lamb, B. T.; Uetz, P., MPIDB: the microbial protein interaction database. *Bioinformatics* **2008,** *24*

(15), 1743-4.
[28] Pagel, P.; Kovac, S.; Oesterheld, M.; Brauner, B.; Dunger-Kaltenbach, I.; Frishman, G.; Montrone, C.; Mark, P.; Stumpflen, V.; Mewes, H. W.; Ruepp, A.; Frishman, D., The MIPS mammalian protein-protein interaction database. *Bioinformatics* **2005,** *21* (6), 832-4.
[29] Gong, W.; Zhou, D.; Ren, Y.; Wang, Y.; Zuo, Z.; Shen, Y.; Xiao, F.; Zhu, Q.; Hong, A.; Zhou, X.; Gao, X.; Li, T., PepCyber:P~PEP: a database of human protein protein interactions mediated by phosphoprotein-binding domains. *Nucleic acids research* **2008,** *36* (Database issue), D679-83.
[30] McDowall, M. D.; Scott, M. S.; Barton, G. J., PIPs: human protein-protein interaction prediction database. *Nucleic acids research* **2009,** *37* (Database issue), D651-6.
[31] Huang, T. W.; Tien, A. C.; Huang, W. S.; Lee, Y. C.; Peng, C. L.; Tseng, H. H.; Kao, C. Y.; Huang, C. Y., POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome. *Bioinformatics* **2004,** *20* (17), 3273-6.
[32] Bowers, P. M.; Pellegrini, M.; Thompson, M. J.; Fierro, J.; Yeates, T. O.; Eisenberg, D., Prolinks: a database of protein functional linkages derived from coevolution. *Genome Biol* **2004,** *5* (5), R35.
[33] Winter, C.; Henschel, A.; Kim, W. K.; Schroeder, M., SCOPPI: a structural classification of protein-protein interfaces. *Nucleic acids research* **2006,** *34* (Database issue), D310-4.
[34] von Mering, C.; Jensen, L. J.; Snel, B.; Hooper, S. D.; Krupp, M.; Foglierini, M.; Jouffre, N.; Huynen, M. A.; Bork, P., STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic acids research* **2005,** *33* (Database issue), D433-7.
[35] von Mering, C.; Jensen, L. J.; Kuhn, M.; Chaffron, S.; Doerks, T.; Kruger, B.; Snel, B.; Bork, P., STRING 7--recent developments in the integration and prediction of protein interactions. *Nucleic acids research* **2007,** *35* (Database issue), D358-62.
[36] Szklarczyk, D.; Franceschini, A.; Kuhn, M.; Simonovic, M.; Roth, A.; Minguez, P.; Doerks, T.; Stark, M.; Muller, J.; Bork, P.; Jensen, L. J.; von Mering, C., The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research* **2011,** *39* (Database issue), D561-8.
[37] Razick, S.; Magklaras, G.; Donaldson, I. M., iRefIndex: a consolidated protein interaction database with provenance. *BMC bioinformatics* **2008,** *9* (1), 405.
[38] Chaurasia, G.; Malhotra, S.; Russ, J.; Schnoegl, S.; Hanig, C.; Wanker, E. E.; Futschik, M. E., UniHI 4: new tools for query, analysis and visualization of the human protein-protein interactome. *Nucleic acids research* **2009,** *37* (Database issue), D657-60.
[39] Ito, T.; Chiba, T.; Ozawa, R.; Yoshida, M.; Hattori, M.; Sakaki, Y., A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* **2001,** *98* (8), 4569 - 4574.

Table S2. Literature and Text Mining Tools for Protein-protein Interaction

Acronym	Description	URL/References
CBioC	Collaborative Bio Curation at the first step uses automatic text extraction from interaction database, then biologists contribute to eliminating inconsistencies. CBioC runs as a web browser extension and it can also be accessed directly (without having to install a plug-in).	http://cbioc.eas.asu.edu/
Chilibot	A tool for PubMed literature database to rapidly identify relationships between genes, proteins, or any keywords and the results are returned as a graph.	www.chilibot.net
GoPubMed	GoPubMed is a search engine for exploring PubMed search results.	www.gopubmed.org
iHOP	Information Hyperlinked over Proteins is a network of concurring genes and proteins extends through the scientific literature, using genes and proteins as hyperlinks between sentences and abstracts.	www.ihop-net.org/UniPub/iHOP [1, 2]
iProLINK	*i*ProLINK (*i*ntegrated Protein Literature, INformation andKnowledge) is a resource to facilitate text mining in the area of literature-based database curation, named entity recognition, and protein ontology development. The collection of data sources can be utilized by computational and biological scientists to explore literature information on proteins and their properties.	pir.georgetown.edu/iprolink [3]
PubGene	This tool gather information of genes and proteins such as co-occurrences in the abstracts of scientific papers, their sequence homology, and statistical probability of their co-occurrences, and then identify the relationships between genes and proteins, diseases, cell processes.	www.pubgene.org
Whatizit	Whatizit is a text mining tool and Medline abstracts retrieval/search engine. It can identifying molecular biology terms and linking them to publicly available databases. Identified terms are wrapped with XML tags that carry additional information, such as the primary keys to the databases where all the relevant information is kept. The wrapping XML is translated into HTML hypertext links.	www.ebi.ac.uk/webservices/whatizit/info.jsf

REFERENCES

[1] Fernandez, J.; Hoffmann, R.; Valencia, A., iHOP web services. *Nucleic acids research* **2007,** (35), W21 - W26.
[2] Hoffmann, R.; Valencia, A., A gene network for navigating the literature. *Nat Genet* **2004,** *36* (7), 664.
[3] Hu, Z. Z.; Mani, I.; Hermoso, V.; Liu, H.; Wu, C. H., iProLINK: an integrated protein resource for literature mining. *Computational biology and chemistry* **2004,** *28* (5-6), 409-16.

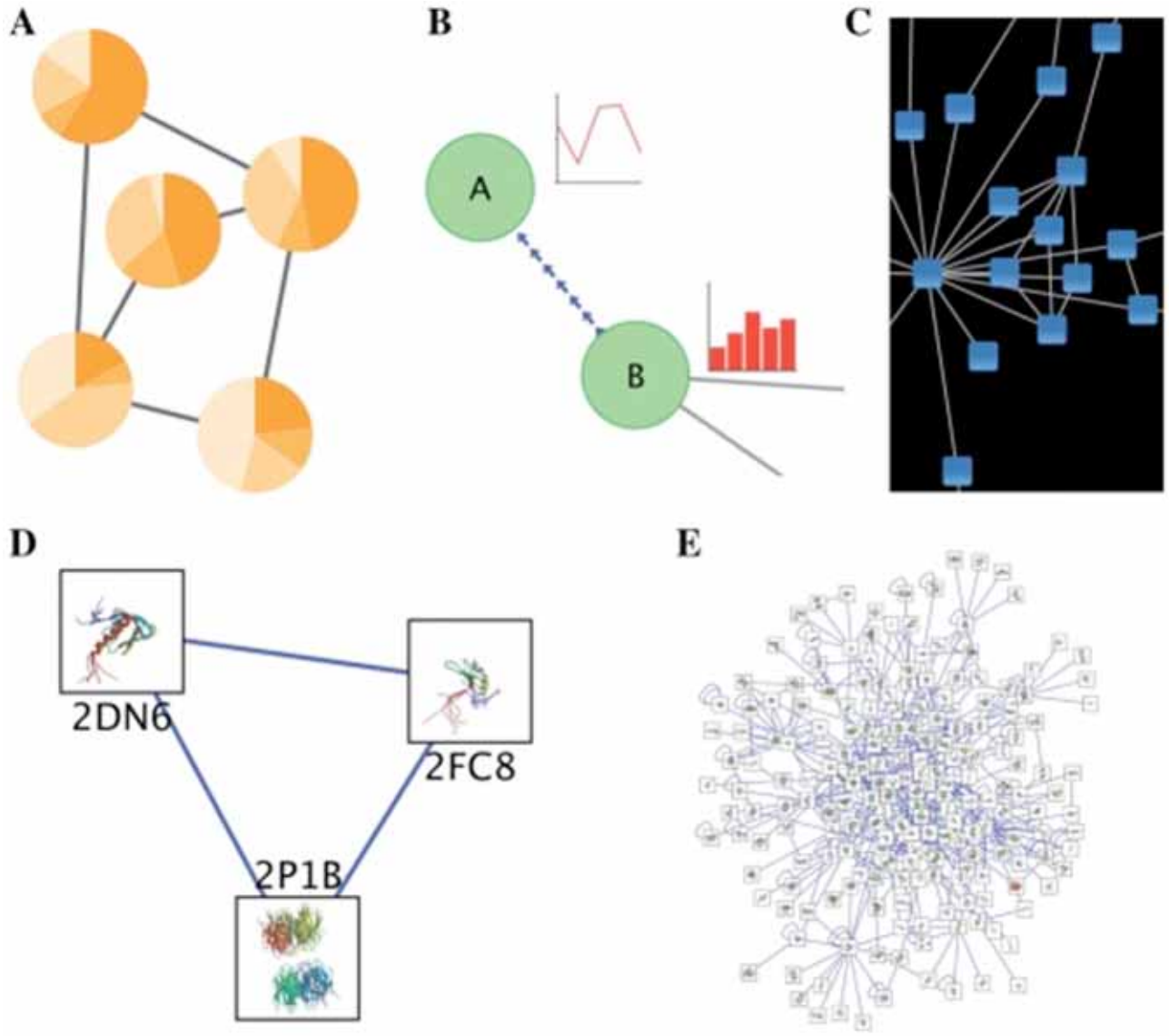

Fig. (S1). Rich network visualizations enabled by the Cytoscape features. Simple networks are shown with custom node images based on (**A**) pie chart displays or (**B**) line plots and bar charts generated using Google's Chart API. (**C**) Nodes have a transparent custom graphic to give the appearance of shading. (**D** and **E**) Protein–protein interaction networks in which each node contains a 3D image of the protein structure of the protein represented by the node [146].

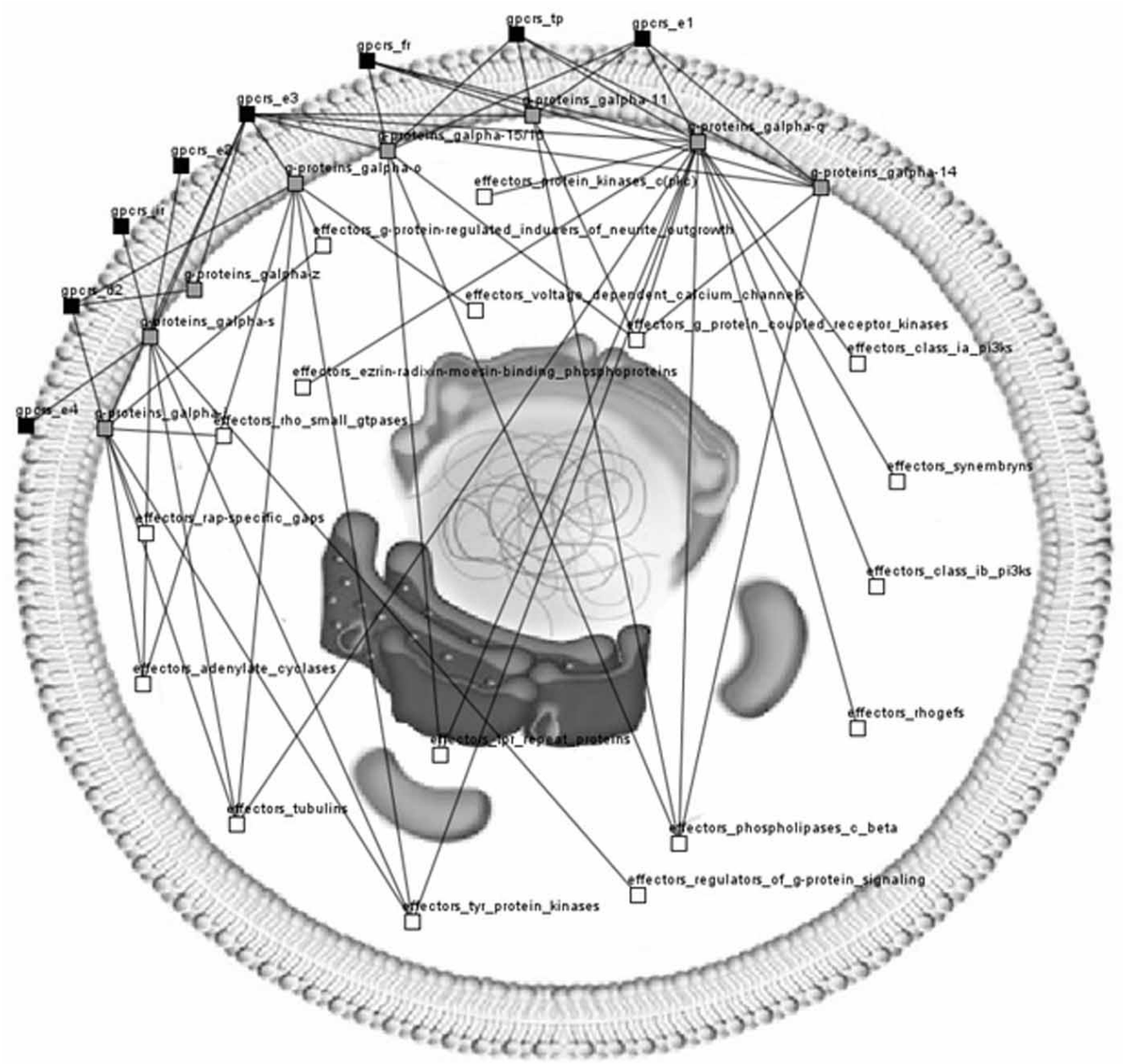

Fig. (S2). Visualization of human postanoid receptors and their interactions. GPCR transmembrane proteins (black) are classified in subfamilies whereas G-proteins (grey) and effectors (white) are classified in families according to the Human-gpDB database. A preloaded image of a cell shows the signal transduction from the outer to the inner part of the cell [156].

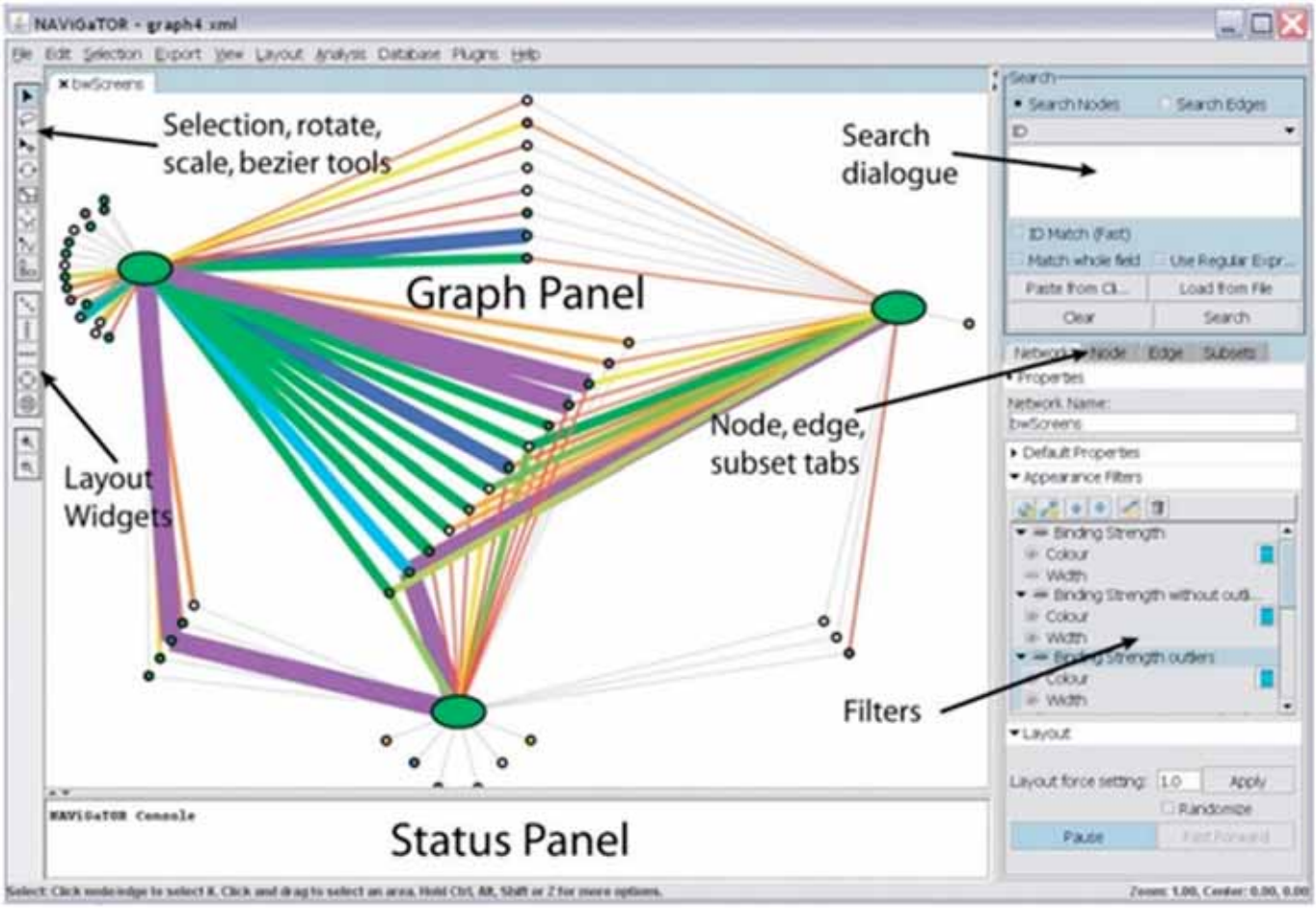

Fig. (S3). Various tools and options of the NAViGaTOR user interface. A graph is shown in the 'Graph Panel', with edges adjusted automatically by 'Edge Filters'. Filters can be used to automatically control visual attributes of both nodes and edges [158].