

This Provisional PDF corresponds to the article as it appeared upon acceptance. Copyedited and fully formatted PDF and full text (HTML) versions will be made available soon.

NetPath: a public resource of curated signal transduction pathways

Genome Biology 2010, **11**:R3 doi:10.1186/gb-2010-11-1-r3

Kumaran Kandasamy (kumaran@ibioinformatics.org)
Sujatha Mohan (sujatha@rcai.riken.jp)
Rajesh Raju (rajesh@ibioinformatics.org)
Shivakumar Keerthikumar (shivakumar@ibioinformatics.org)
Ghantasala S SAMEER Kumar (sameer@ibioinformatics.org)
Abhilash K Venugopal (abhilash@ibioinformatics.org)
Deepthi Telikicherla (deepthi@ibioinformatics.org)
Daniel J Navarro (vimes656@gmail.com)
Suresh Mathivanan (suresh@ibioinformatics.org)
Christian Pecquet (christian.pecquet@bru.licr.org)
Sashi KANTH Gollapudi (sashi@ibioinformatics.org)
Sudhir GOPAL Tattikota (sudhir@ibioinformatics.org)
Shyam Mohan (shyam@ibioinformatics.org)
Hariprasad Padhukasahasram (hariprasad@ibioinformatics.org)
Yashwanth Subbannayya (yashwanth@ibioinformatics.org)
Renu Goel (renu@ibioinformatics.org)
Harrys K C Jacob (harrys@ibioinformatics.org)
Jun Zhong (jzhong@jhmi.edu)
Raja Sekhar (raja@ibioinformatics.org)
Vishalakshi Nanjappa (vishalakshi@ibioinformatics.org)
Lavanya Balakrishnan (lavanya@ibioinformatics.org)
Roopashree Subbaiah (roopa@ibioinformatics.org)
Y I Ramachandra (ylrpub@gmail.com)
Abdul Rahiman (rahimanbt@yahoo.com)
T s Keshava Prasad (keshav@ibioinformatics.org)
Jian-Xin Lin (linjx@nhlbi.nih.gov)
Jon C D Houtman (jon-houtman@uiowa.edu)
Stephen Desiderio (sdesider@jhmi.edu)
Jean-Christophe Renauld (Jean-Christophe.RENAULD@bru.licr.org)
Stefan Constantinescu (Stefan.Constantinescu@bru.licr.org)
Osamu Ohara (oosamu@rcai.riken.jp)
Toshio Hirano (hirano@molonc.med.osaka-u.ac.jp)
Masato Kubo (raysolfe@rs.noda.tus.ac.jp)
Sujay Singh (ssingh@imgenex.com)
Purvesh Khatri (purvesh@cs.wayne.edu)
Sorin Draghici (sod@cs.wayne.edu)
Gary D Bader (gary.bader@utoronto.ca)
Chris Sander (sander@mskcc.org)
Warren J Leonard (wjl@helix.nih.gov)
Akhilesh Pandey (pandey@jhmi.edu)

© 2010 Kandasamy *et al.*, licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ISSN 1465-6906
Article type Method
Submission date 21 April 2009
Acceptance date 12 January 2010
Publication date 12 January 2010
Article URL <http://genomebiology.com/2010/11/1/R3>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *Genome Biology* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *Genome Biology* go to

<http://genomebiology.com/info/instructions/>

NetPath: a public resource of curated signal transduction pathways

Kumaran Kandasamy^{1,2,*}, S Sujatha Mohan^{1,*,#}, Rajesh Raju^{1,4}, Shivakumar Keerthikumar¹, Ghantasala S Sameer Kumar¹, Abhilash K Venugopal¹, Deepthi Telikicherla¹, Daniel J Navarro¹, Suresh Mathivanan¹, Christian Pecquet³, Sashi Kanth Gollapudi¹, Sudhir Gopal Tattikota¹, Shyam Mohan¹, Hariprasad Padhukasahasram¹, Yashwanth Subbannayya¹, Renu Goel¹, Harrys K C Jacob^{1,2}, Jun Zhong², Raja Sekhar¹, Vishalakshi Nanjappa¹, Lavanya Balakrishnan¹, Roopashree Subbaiah¹, Y L Ramachandra⁴, B Abdul Rahiman⁴, T S Keshava Prasad¹, Jian-Xin Lin⁵, Jon C D Houtman⁶, Stephen Desiderio⁷, Jean-Christophe Renauld³, Stefan Constantinescu³, Osamu Ohara^{8,9}, Toshio Hirano^{10,11}, Masato Kubo^{12,13}, Sujay Singh¹⁴, Purvesh Khatri¹⁵, Sorin Draghici^{15,16}, Gary D Bader^{17,18}, Chris Sander¹⁸, Warren J Leonard⁵, Akhilesh Pandey^{2,19,§}

¹Institute of Bioinformatics, International Tech Park, Bangalore 560066, India

²McKusick-Nathans Institute of Genetic Medicine and the Department of Biological Chemistry, Johns Hopkins University, Baltimore, Maryland 21205, USA

³The Ludwig Institute for Cancer Research, Brussels Branch, and the Experimental Medicine Unit, Christian de Duve Institute of Cellular Pathology, Universite Catholique de Louvain, avenue Hippocrate 74, B-1200-Brussels, Belgium

⁴Department of Biotechnology and Bioinformatics, Kuvempu University, Jnanasahyadri, Shimoga 577451, India

⁵Laboratory of Molecular Immunology, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD 20892, USA

⁶Department of Microbiology, Carver College of Medicine, University of Iowa, Iowa City, Iowa 52242, USA.

⁷Department of Molecular Biology and Genetics, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

⁸Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, Kanagawa 230-0045, Japan

⁹Department of Human Genome Technology, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

¹⁰Laboratory for Cytokine Signaling, RIKEN Research Center for Allergy and Immunology, Yokohama, Kanagawa 230-0045, Japan

¹¹Laboratories of Developmental Immunology, Graduate School of Frontier Biosciences and Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan

¹²Research Institute for Biological Sciences, Tokyo University of Science, Yamazaki, Noda City, Chiba 278-0022, Japan

¹³Signal/Network Team, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, Suehiro-cho, Tsurumi, Yokohama, Kanagawa 230-0045, Japan

¹⁴IMGENEX India Pvt. Ltd., Bhubaneswar, Orissa 92121, India

¹⁵Department of Computer Science, Wayne State University, Detroit, Michigan 48202, USA

¹⁶Karmanos Cancer Institute, Wayne State University, Detroit, Michigan 48202, USA

¹⁷Banting and Best Department of Medical Research, Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College St, Toronto, Ontario M5S 3E1, Canada

¹⁸Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York 10021, USA

¹⁹Department of Oncology, Johns Hopkins University, Baltimore, Maryland 21205, USA

#Current Address: Research Unit for Immunoinformatics, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, Kanagawa, Japan

§Corresponding author

*These authors contributed equally

Corresponding author: Akhilesh Pandey (pandey@jhmi.edu).

Abstract

We have developed NetPath as a resource of curated human signaling pathways. As an initial step, NetPath provides detailed maps of a number of immune signaling pathways, which include ~1,600 reactions annotated from the literature and more than 2,800 instances of transcriptionally-regulated genes - all linked to over 5,500 published articles. We anticipate NetPath to become a consolidated resource for human signaling pathways that should enable systems biology approaches.

Background

Complex biological processes such as proliferation, migration and apoptosis are generally regulated through responses of cells to stimuli in their environment. Signal transduction pathways often involve binding of extracellular ligands to receptors, which trigger a sequence of biochemical reactions inside the cell. Generally, proteins are the effector molecules, which function as part of larger protein complexes in signaling cascades. Cellular signaling events are generally studied systematically through individual experiments that are widely scattered in the biomedical literature. Assembling these individual experiments and putting them in the context of a signaling pathway is difficult, time-consuming and cannot be automated.

The availability of detailed signal transduction pathways that can easily be understood by humans as well as be processed by computers is of great value to biologists trying to understand the working of cells, tissues and organ systems [1]. A systems-level understanding of any biological process requires, at the very least, a comprehensive map depicting the relationships among the various molecules involved [2]. For instance, these maps could be used to construct a complete network of protein-protein and transcriptional events, which would help in identifying novel transcriptional and other regulatory networks [3]. These can be extended to predict how the interactions, if perturbed singly or in combination, could affect individual biological processes. Additionally, they could be used to identify possible unintended effects of a candidate therapeutic agent on any clusters in a pathway [4]. We have developed a resource called NetPath, which allows biomedical scientists to visualize, process and manipulate data pertaining to signaling pathways in humans.

Results and Discussion

Development of ‘NetPath’ as a resource for signal transduction pathways

NetPath [5] is a resource for signaling pathways in humans. As an initial set, we have curated a list of 10 immune signaling pathways. The list of immune signaling pathways includes T and B cell receptor signaling pathways in addition to several interleukin signaling pathways as shown in Table 1. A query system facilitates searches based on protein/gene names or accession numbers to obtain the list of cellular signaling pathways involving the queried protein (Figure 1).

Signaling pathway annotation

To facilitate annotation of pathway data, we first developed a tool called ‘PathBuilder’ [6]. PathBuilder is a signal transduction pathway annotation tool that allows annotation of pathway information, storage of data, easy retrieval and export into community standardized data structures such as BioPAX [7], PSI-MI [8] and SBML [9] formats. PathBuilder facilitates the entry of information pertaining to protein interactions, enzyme–regulated reactions, intracellular translocation events and genes that are transcriptionally regulated.

Protein-protein interactions could be binary when two proteins directly interact with each other - ‘direct interaction,’ or, when the proteins are present in a complex of proteins - ‘complex interaction.’ Both types of protein interactions are comprehensively collected from the literature. We provide PubMed identifiers, experiment type and host organism in which the interaction has been detected.

Enzyme–regulated reactions such as post–translational modifications (e.g. phosphorylation, proteolytic cleavage, ubiquitination, prenylation or sulfation) are annotated as catalysis interactions. For each catalysis or modification event, the upstream enzyme, downstream targets and the site of the modification for a protein is annotated, if

it is available. Proteins that translocate from one compartment (e.g. cytoplasm) to another (e.g. nucleus) are represented as transport events. For all reactions, a brief comment describing the reaction is also provided.

Display of pathway information

The homepage of any given pathway contains a brief description of the pathway, a summary of the reaction statistics and a list of the molecules involved in the pathway. Reactions in a pathway are provided under three distinct categories including physical interactions, enzyme catalysis and transport. Furthermore, the pathway data is also provided in PSI-MI, BioPAX and SBML formats, which can also be visualized through other external network visualization software such as Cytoscape [10].

Cataloging transcriptionally regulated genes

In addition to the above pathway annotations, information on genes that are transcriptionally regulated is provided in NetPath. This is important because addition of most extracellular growth factors or ligands leads to an alteration in the transcriptome of the cell. Often, some of the transcriptionally regulated genes are used as ‘reporters’ in biological experiments where the pathway is being studied. We have cataloged a number of genes that are up or down regulated by the particular ligand involved in each pathway. These up/down regulated genes can be considered as ‘signatures’ for that particular pathway. We have incorporated both microarray and non-microarray (e.g. Northern blot, qRT-PCR, SAGE etc.) experiments for gene expression. In each case, the type of experiment (i.e. microarray, non-microarray or both) used to obtain the data is indicated. Additionally, we have also annotated the transcription factors that are responsible for transcriptional regulation of the downstream genes where such information is available. Given the large number of transcriptionally regulated genes for each pathway, we have also developed a query system that permits users to search such genes using gene symbol or accession numbers. This feature will be valuable for shortlisting genes that are common to several pathways or specific to any given pathway.

Pathway statistics

At present the ten annotated immune signaling pathways comprise 703 proteins and 1,572 reactions. The reactions can be grouped into 740 molecular association events, 727 enzyme catalysis and 105 translocation events. Our pathways provide list of ‘2,004’ and ‘889’ genes that are up or down regulated respectively at the level of mRNA expression. Including 10 cancer signaling pathways that are also available through Cancer Cell Map [11], NetPath now contains 1,682 proteins and 3,219 reactions which can be grouped into 1,800 molecular association events, 1,218 enzyme catalysis and 201 transport events. Table 1 shows the overall immune signaling pathways statistics as of November 1, 2009.

Comparison with other signaling databases

Although there are over 310 resources [12] that provide some form of pathway related information, many of these currently available resources belong to the category of protein-protein interaction databases, metabolic pathways, transcription factors/gene regulatory networks, genetic interaction networks. Some of these pathways include KEGG [13], BioCarta [14], Science’s STKE Connections Maps [15], Reactome [16], National Cancer Institute’s Pathway Interaction Database (PID) [17], Pathway database from Cell Signaling Technology [18], Integrating Network Objects with Hierarchies (INOH) [19], Signaling Pathway Database (SPAD) [20], GOLD.db [21], PATIKA [22], pSTIING [23], TRMP [24], WikiPathways [25] and PANTHER [26]. However, many of these pathway resources are not primary i.e. they combine data from many other sources. Thus, we have compared NetPath with eight other signaling pathways that contain manually curated human pathway data derived from experiments. Of all these pathways that are compared, NetPath stands out for 3 unique features. The first is that it includes annotation of transcriptionally regulated genes. Such a catalog of transcriptionally regulated genes pertaining to a given pathway should be highly useful in exploring pathway-specific expression signatures. The second unique feature is that NetPath provides manually curated textual description of each pathway reaction and such a

description of pathway reactions should facilitate an easier understanding of these pathways, aiding biomedical scientists to get an overview of the pathway reactions in a central repository. The third unique feature of NetPath is that this data can be searched using SPARQL – the recommended query language for the semantic web. Table 2 compares some of the salient features of NetPath with some of the other popular signaling pathway resources. In addition to the unique features, NetPath also provides a separate molecule page for every pathway component along with a brief textual description for each molecule. Overall, NetPath should be a useful pathway resource with unique features that should facilitate signaling research.

Interleukin-2 pathway as a prototype

One of the best studied immune signaling pathways is the interleukin-2 (IL-2) signaling pathway [27]. IL-2 is a multifunctional cytokine with pleiotropic effects on a number of cells of the immune system [27], [28]. IL-2 was originally discovered as a T-cell growth factor, TCGF [29], but it also was found to have actions related to B cell proliferation [30], and proliferation of and cytolytic activity of NK cells [31]. IL-2 also activates lymphokine activated killer cells [32]. In contrast to its proliferative effects, IL-2 also has potent activity in a process known as activation-induced cell death [33]. More recently, IL-2 was shown to promote tolerance through its effects on regulatory T cell (Treg) development [34]. IL-2 clinically has anti-cancer effects [35] as well as utility in supporting T cell numbers in HIV/AIDS [36].

There are three classes of IL-2 receptors, binding IL-2 with low, intermediate, or high-affinity [37]. The low affinity receptor (IL-2R α alone) is not functional, whereas signaling by IL-2 involves either the high affinity hetero-trimeric receptor containing IL-2R α , IL-2R β and the common cytokine receptor gamma chain (originally named IL-2R γ and now generally denoted as γ c) or the intermediate affinity hetero-dimeric receptor composed of IL-2R β and γ c [37] and [38]. Mutations in the *IL2RG* gene result in X-linked SCID (severe combined immunodeficiency disease) [39]. IL-2 stimulation induces the activation of Janus family tyrosine kinases JAK1 and JAK3, which associate with IL-

2R β and γ_c , respectively. These kinases in turn phosphorylate IL-2R β and induce tyrosine phosphorylation of STATs (signal transducers and activators of transcription) and various other downstream targets [40]. The downstream signaling pathway also involves MAPK and PI3-kinase signaling modules [41] leading to both mitogenic and anti-apoptotic signals [40], [41] and [42].

IL-2 signaling pathway currently comprises of 68 proteins, 155 reactions with 68 molecular association events, 76 enzymatic catalysis and 11 translocation events. Importantly, 840 transcriptionally regulated events, i.e. a list of IL-2 induced up or down regulated genes have been annotated from the published literature. In all, the reactions in the IL-2 pathway are supported by 1,289 links to research articles. Figure 2 shows the pathway page of the IL-2 pathway.

Integration of pathway information with other resources

The pathways developed by us have been integrated with the Human Protein Reference Database (HPRD) [43, 44]. The integration of pathways in HPRD helps identify each component of the pathway in the context of its detailed proteomic annotations [45]. As part of our community participation with other databases/resources, we hope to establish connections with other pathways databases such as KEGG [27] and Reactome [16] in the future.

Availability of pathway data

A digital representation of pathways is essential to be able to manipulate the large amount of available information [4]. The diversity among pathway databases is also reflected in differences in data models, data access methods and file formats. This leads to the incompatibility of data formats for the analysis of pathway data. To avoid this, data standards are adopted by many of the pathway databases [12, 46]. Data standards reduce the total number of translation operations needed to exchange data between multiple sources. To facilitate easy information retrieval from a wide variety of pathway

resources, a broad effort in the biological pathways community called BioPAX (Biological Pathways Exchange) was initiated. Since many less-detailed data types in a pathway database are difficult to represent in a very detailed format, BioPAX ontology uses hierarchical entity classes to present multiple levels of data resolution. All pathways in NetPath are available for download in BioPAX level 2, version 1.0. The Proteomics Standards Initiative's Molecular Interaction (PSI-MI) format was developed to exchange molecular interaction data between databases containing databases. PSI-MI data representation facilitates data comparison, exchange and verification [8]. The molecular interaction subset of NetPath pathways is also available in PSI-MI version 2.5. Systems Biology Markup Language (SBML) was developed as a medium for representation and exchange of biochemical network models [9]. NetPath provides all pathway data in SBML version 2.1 format. All data is made available under the Creative Commons license version 2.5 [47], which stipulates that the pathways may be freely used if adequate credit is given to the authors. Support for these data standards and free license enables the integration of knowledge from multiple sources in a coherent and reliable manner.

Enabling semantic web for NetPath

The semantic web envisions an Internet where specific information can be obtained from the web automatically using computers. Because providing computers with the intuitiveness of humans is nearly impossible as of now, creation of meta-data - data about data - can help computers identify what is being sought less ambiguously. However, annotating more data does not automatically imply that the data can be made easily accessible by the user. For instance, although many resources permit direct querying of individual molecules in the respective databases, queries based on 'relationships' between different entries in the databases cannot be handled. One possible solution to enable searching by such 'concepts' is to incorporate semantic web features that explicitly describe the inter-relationship between entries in the databases.

The W3C has established SPARQL as the standard semantic query language. Pathway data in BioPAX uses the web ontology language (OWL) format, which is highly descriptive in nature and can be used to make pathways semantically 'queryable.' In this regard, we have implemented an application programming interface (API) for NetPath that accepts SPARQL over HTTP to query the BioPAX files describing NetPath pathways. The return results are provided in SPARQL Query Results XML format. Although biologists cannot be expected to write SPARQL queries, the ability to send SPARQL queries over HTTP allows bioinformaticians to write client applications which can retrieve NetPath resources taking advantage of the descriptive richness of SPARQL and BioPAX.

Analyzing impact factor for pathways

It is becoming clear that pathway information can be used in the context of genome-scale gene expression experiments. A novel approach has been recently reported to measure the biological impact of perturbation of pathways in genomewide gene expression experiments [48]. This approach considers the topology of genes in a pathway in conjunction with classical statistics for microarray analysis. The impact factor is a statistical approach, which can capture the magnitude of the expression changes of each gene, the position of the differentially expressed genes on the given pathways, the topology of the pathway that describes how these genes interact, and the type of signaling interactions between them. Our previous results using KEGG pathways were found to correlate with known biological events, which were missed by other widely used classical analysis methods. However, this approach could not be applied to study immune responses because of limited availability of such pathways in humans.

As a proof of principle, we selected publicly available mRNA expression datasets from Gene Expression Omnibus (GEO), a repository for gene expression data [49]. Datasets which include expression analysis of immune cells under different experimental conditions were selected for this purpose.

One of the datasets used [GEO: GDS2214] (as described in ref [50]) was an experimental study of mRNA expression analysis of neutrophils isolated from blood of patients with sepsis-induced acute lung injury (ALI). The neutrophils were cultured with either lipopolysaccharide (LPS) or high mobility group box protein 1 (*HMGB1*), both of which are known to be mediators of inflammatory response. Gene expression analysis was carried out using the Affymetrix GeneChip Human Genome U133 Array Set HG-U133A oligonucleotide gene chip. The authors found enhancement of nuclear translocation activity of NF-kappaB and phosphorylation of Akt and p38 MAPK upon stimulation of LPS or HMGB1. We carried out impact factor analysis using this dataset on all 10 immune signaling pathways. The results corroborate with these findings since IL-1 and IL-6 pathways scores are highly affected while the rest of the NetPath pathways did not show significant scores.

Another dataset selected [GEO: GDS1407] (as described in ref [51]) was a part of the gene expression study that screened a cohort of 102 healthy individuals to investigate the distribution of inflammatory responses to LPS in the normal population in circulating leukocytes. Expression profiling with Affymetrix U95AV2 oligonucleotide microarray identified differentially regulated genes between two phenotypic subgroups which have been described as high LPS responders (lp_{high}) and low LPS responders (lp_{low}), based on the concentration cytokines produced in response to LPS. Gene expression analysis was done using the Affymetrix U95AV2 human oligonucleotide arrays. Impact factor analysis was carried out using this dataset on all the 10 immune signaling pathways. Impact factor scores for IL-1 and IL-6 NetPath pathways in lp_{high} group show high values whereas impact factor scores for lp_{low} do not show any significant perturbation of NetPath pathways. The scores are consistent with experimental results showing upregulation of IL-1 and IL-6 ligands in lp_{high} group. The impact factor gives an insight that not only the ligands are upregulated, but also the pathway seems to be highly affected. It should be noted that impact factor is not the only method and other methods will continue to be developed and could be applied to such pathway data.

Outlook

In addition to keeping these pathways updated on a regular basis, we will also add additional pathways to NetPath. We also hope to involve the biomedical community by allowing researchers to provide feedback as well as to volunteer to become ‘Pathway Authorities’ on specific pathways, similar to the successful contribution model of the BioCarta resource [14]. In this regard, we have already recruited several investigators to serve as pathways authorities in our initial effort. Multiple Pathway Authorities are possible for the same pathway if there are enough interested investigators with expertise who wish to contribute in this fashion. For instance, ten other signaling pathways pertaining to cancer signaling were developed for the Cancer Cell Map project [11], as a collaboration with Memorial Sloan-Kettering Cancer Center and this data is also available through Pathway Commons [52]. We also intend to map our human specific pathway data to corresponding mouse orthologs to create the mouse equivalent of our signaling pathways. Since large amounts of human signaling pathway data is modeled using the mouse, this will facilitate biological system modeling that relies on primary experimental data. We also intend to incorporate pathway visualization for all existing pathways in NetPath as well as those that will be added in the future using the PathVisio software [53]. PathVisio also supports visualization of gene expression data in the context of pathways, which will enable biologist to display a systems view of the signaling pathway.

Conclusions

We have developed a resource for integration of human cellular signaling events. These pathway specific protein-protein interaction data can be used to generate larger physical networks of protein-protein interactions which when coupled with data on genetic interactions could help in defining novel functional relationships among proteins. In addition, genetic interactions can functionally link proteins that belong to unconnected physical networks. These pathways could also be used to interrogate gene expression signatures in cancers and other human diseases to better understand the mechanisms or to

obtain profiles for diagnostic or therapeutic purposes. There is a large amount of known information about different cellular signaling pathways, which control a variety of cellular functions that is difficult to collect by one group. We support the vision of many data providers collecting data of interest and making it freely available in standard formats as a scalable way to represent all known pathway information in databases for a comprehensive analysis. Overall, we hope to engage the biomedical community in keeping the NetPath pathway resource up to date and as error-free as possible.

Materials and methods

The initial annotation process of any signaling pathway involves gathering and reading of review articles to achieve a brief overview of the pathway. This process is followed by listing all the molecules that were reported to be involved in the pathway under annotation. Information regarding potential pathway authorities were also gathered at this initial stage. Pathway experts were involved in initial screening of the molecules listed to check for any obvious omissions. In the second phase, annotators manually perform extensive literature searches using search keys which include all the alternate names of the molecules involved, name of the pathway, name of reactions, etc. In addition, iHOP [54] resource was also used to perform advanced PubMed based literature searches to collect the reactions that were reported to be implicated in a given pathway. The collected reactions were manually entered using the PathBuilder [6] annotation interface that is subjected to an internal review process, which involves Ph.D. level scientists with expertise in the areas of molecular biology, immunology and biochemistry. However, there are instances where although a molecule has been implicated in a pathway in a published report, the associated experimental evidence is either weak or differs from experiments carried out by other groups. For this purpose, we have recruited several investigators as ‘Pathway Authorities’ based on their expertise in individual signaling pathways. The review by ‘Pathway Authorities’ occasionally leads to correction of errors, or more commonly, to inclusion of additional information. Finally, the ‘Pathway Authorities’ help in assessing whether the work of all major laboratories has been incorporated for the given signaling pathway.

Abbreviations

KEGG: Kyoto Encyclopedia of Genes and Genomes, HPRD: Human Protein Reference Database, STKE: Signal Transduction Knowledge Environment, BioPAX : Biological Pathways Exchange, PSI-MI: Proteomics Standards Initiative - Molecular Interactions, SBML: Systems Biology Markup Language, GEO: Gene Expression Omnibus

Authors' contributions

SM1, RR, SK, GSSK, AKV, DT, DJN, SM2, CP, SKG, SGT, SM3, HP, YS, RG, HKCJ, JZ, RS1, VN, SB, RS2, YLR, BAR, TSKP and JL collected the data. JCDH, SD1, JR, SC, OO, TH, MK, SS, WJL and AP serve as 'Pathway Authorities.' KK, SM1 and AP wrote the manuscript. KK and SM2 developed the software. KK, AKV, DJN, SKG, PK and SD carried out the impact-factor analysis. KK, GDB, CS and AP participated in the study design. All authors read and approved the final manuscript.

Acknowledgements

Akhilesh Pandey is supported by grants from Johns Hopkins Breast Cancer SPORE (CA 88843) Career Development Award, Department of Defense Era of Hope Scholar (W81XWH-06-1-0428) and partly by National Institutes of Health grant U54 RR020839 (Roadmap Initiative for Technology Centers for Networks and Pathways).

References

1. Fukuda K, Takagi T: **Knowledge representation of signal transduction pathways.** *Bioinformatics* 2001, **17**:829-837.
2. Uetz P, Finley RL, Jr.: **From protein networks to biological systems.** *FEBS Lett* 2005, **579**:1821-1827.
3. Ideker T: **A systems approach to discovering signaling and regulatory pathways--or, how to digest large interaction networks into relevant pieces.** *Adv Exp Med Biol* 2004, **547**:21-30.
4. Schaefer CF: **Pathway databases.** *Ann N Y Acad Sci* 2004, **1020**:77-91.
5. **NetPath** [<http://www.netpath.org/>].
6. Kandasamy K, Keerthikumar S, Raju R, Keshava Prasad TS, Ramachandra YL, Mohan S, Pandey A: **PathBuilder--open source software for annotating and developing pathway resources.** *Bioinformatics* 2009, **25**:2860-2862.
7. **BioPAX: Biological Pathways Exchange** [<http://www.biopax.org/>].
8. Hermjakob H, Montecchi-Palazzi L, Bader G, Wojcik J, Salwinski L, Ceol A, Moore S, Orchard S, Sarkans U, von Mering C, Roechert B, Poux S, Jung E, Mersch H, Kersey P, Lappe M, Li Y, Zeng R, Rana D, Nikolski M, Husi H, Brun C, Shanker K, Grant SG, Sander C, Bork P, Zhu W, Pandey A, Brazma A, Jacq B *et al*: **The HUPO PSI's molecular interaction format--a community standard for the representation of protein interaction data.** *Nat Biotechnol* 2004, **22**:177-183.
9. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novere N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED *et al*: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19**:524-531.
10. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**:2498-2504.
11. **The Cancer Cell Map** [<http://cancer.cellmap.org/>].
12. Bader GD, Cary MP, Sander C: **Pathguide: a pathway resource list.** *Nucleic Acids Res* 2006, **34**:D504-506.
13. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**:27-30.
14. **BioCarta** [<http://www.biocarta.com/>].
15. **Connections Maps** [<http://stke.sciencemag.org/>].
16. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, Lewis S, Birney E, Stein L: **Reactome: a knowledgebase of biological pathways.** *Nucleic Acids Res* 2005, **33**:D428-432.
17. **NCI-Nature Pathway Interaction Database** [<http://pid.nci.nih.gov/>].

18. **Cell Signaling Technology** [<http://www.cellsignal.com/>].
19. **INOH Pathway Database** [<http://www.inoh.org/>].
20. **Signaling Pathway Database** [<http://www.grt.kyushu-u.ac.jp/spad/>].
21. Hackl H, Maurer M, Mlecnik B, Hartler J, Stocker G, Miranda-Saavedra D, Trajanoski Z: **GOLD.db: genomics of lipid-associated disorders database**. *BMC Genomics* 2004, **5**:93.
22. Demir E, Babur O, Dogrusoz U, Gursoy A, Nisanci G, Cetin-Atalay R, Ozturk M: **PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways**. *Bioinformatics* 2002, **18**:996-1003.
23. Ng A, Bursteinas B, Gao Q, Mollison E, Zvelebil M: **pSTIING: a 'systems' approach towards integrating signalling pathways, interaction and transcriptional regulatory networks in inflammation and cancer**. *Nucleic Acids Res* 2006, **34**:D527-534.
24. Zheng CJ, Zhou H, Xie B, Han LY, Yap CW, Chen YZ: **TRMP: a database of therapeutically relevant multiple pathways**. *Bioinformatics* 2004, **20**:2236-2241.
25. Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C: **WikiPathways: pathway editing for the people**. *PLoS Biol* 2008, **6**:e184.
26. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A: **PANTHER: a library of protein families and subfamilies indexed by function**. *Genome Res* 2003, **13**:2129-2141.
27. Nakamura M, Asao H, Takeshita T, Sugamura K: **Interleukin-2 receptor heterotrimer complex and intracellular signaling**. *Semin Immunol* 1993, **5**:309-317.
28. Morgan DA, Ruscetti FW, Gallo R: **Selective in vitro growth of T lymphocytes from normal human bone marrows**. *Science* 1976, **193**:1007-1008.
29. Mingari MC, Gerosa F, Carra G, Accolla RS, Moretta A, Zubler RH, Waldmann TA, Moretta L: **Human interleukin-2 promotes proliferation of activated B cells via surface receptors similar to those of activated T cells**. *Nature* 1984, **312**:641-643.
30. London L, Perussia B, Trinchieri G: **Induction of proliferation in vitro of resting human natural killer cells: IL 2 induces into cell cycle most peripheral blood NK cells, but only a minor subset of low density T cells**. *J Immunol* 1986, **137**:3845-3854.
31. Grimm EA, Mazumder A, Zhang HZ, Rosenberg SA: **Lymphokine-activated killer cell phenomenon. Lysis of natural killer-resistant fresh solid tumor cells by interleukin 2-activated autologous human peripheral blood lymphocytes**. *J Exp Med* 1982, **155**:1823-1841.
32. Greene WC, Leonard WJ: **The human interleukin-2 receptor**. *Annu Rev Immunol* 1986, **4**:69-95.
33. Green DR, Droin N, Pinkoski M: **Activation-induced cell death in T cells**. *Immunol Rev* 2003, **193**:70-81.
34. Malek TR, Bayer AL: **Tolerance, not immunity, crucially depends on IL-2**. *Nat Rev Immunol* 2004, **4**:665-674.
35. Rosenberg SA: **Progress in human tumour immunology and immunotherapy**. *Nature* 2001, **411**:380-384.

36. Paredes R, Lopez Benaldo de Quiros JC, Fernandez-Cruz E, Clotet B, Lane HC: **The potential role of interleukin-2 in patients with HIV infection.** *AIDS Rev* 2002, **4**:36-40.
37. Kim HP, Imbert J, Leonard WJ: **Both integrated and differential regulation of components of the IL-2/IL-2 receptor system.** *Cytokine Growth Factor Rev* 2006, **17**:349-366.
38. Leonard WJ: **Type I Cytokines and Interferons and Their Receptors**, Sixth edn. Philadelphia: Lippincott Williams & Wilkins.
39. Noguchi M, Yi H, Rosenblatt HM, Filipovich AH, Adelstein S, Modi WS, McBride OW, Leonard WJ: **Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans.** *Cell* 1993, **73**:147-157.
40. Lin JX, Leonard WJ: **Signaling from the IL-2 receptor to the nucleus.** *Cytokine Growth Factor Rev* 1997, **8**:313-332.
41. Ellery JM, Nicholls PJ: **Alternate signalling pathways from the interleukin-2 receptor.** *Cytokine Growth Factor Rev* 2002, **13**:27-40.
42. Ahmed NN, Grimes HL, Bellacosa A, Chan TO, Tsichlis PN: **Transduction of interleukin-2 antiapoptotic and proliferative signals via Akt protein kinase.** *Proc Natl Acad Sci U S A* 1997, **94**:3627-3632.
43. **HPRD: Human Protein Reference Database [<http://www.hprd.org/>].**
44. Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC, Yatish AJ, Kavitha MP, Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana R, Chandran S, Krishna S, Joy M *et al*: **Development of human protein reference database as an initial platform for approaching systems biology in humans.** *Genome Res* 2003, **13**:2363-2371.
45. Mishra GR, Suresh M, Kumaran K, Kannabiran N, Suresh S, Bala P, Shivakumar K, Anuradha N, Reddy R, Raghavan TM, Menon S, Hanumanthu G, Gupta M, Upendran S, Gupta S, Mahesh M, Jacob B, Mathew P, Chatterjee P, Arun KS, Sharma S, Chandrika KN, Deshpande N, Palvankar K, Raghavnath R, Krishnakanth R, Karathia H, Rekha B, Nayak R, Vishnupriya G *et al*: **Human protein reference database--2006 update.** *Nucleic Acids Res* 2006, **34**:D411-414.
46. Cary MP, Bader GD, Sander C: **Pathway information for systems biology.** *FEBS Lett* 2005, **579**:1815-1820.
47. **Creative Commons license version 2.5 [<http://creativecommons.org/licenses/by/2.5/>].**
48. Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R: **A systems biology approach for pathway level analysis.** *Genome Res* 2007, **17**:1537-1545.
49. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R: **NCBI GEO: mining tens of millions of expression profiles--database and tools update.** *Nucleic Acids Res* 2007, **35**:D760-765.

50. Eliezer Silva JA, Qianbin He, Daiva Svetkauskaite, Christopher Coldren, Jerry A. Nick, Katie Poch, Jong Sung Park, Anirban Banerjee, Edward Abraham: **HMGB1 and LPS induce distinct patterns of gene expression and activation in neutrophils from patients with sepsis-induced acute lung injury.** *Intensive Care Medicine* 2007, **33**:1829 - 1839.
51. Wurfel MM, Park WY, Radella F, Ruzinski J, Sandstrom A, Strout J, Bumgarner RE, Martin TR: **Identification of high and low responders to lipopolysaccharide in normal subjects: an unbiased approach to identify modulators of innate immunity.** *J Immunol* 2005, **175**:2570-2578.
52. **Pathway Commons** [<http://www.pathwaycommons.org/>].
53. van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, Evelo C: **Presenting and exploring biological pathways with PathVisio.** *BMC Bioinformatics* 2008, **9**:399.
54. Hoffmann R, Valencia A: **A gene network for navigating the literature.** *Nat Genet* 2004, **36**:664.

Figure legends:

Figure 1: The NetPath homepage

The search function allows users to query the database with multiple options including gene symbol, protein name, accession number and name of the pathway. The browse option links directly to a page listing all available pathways.

Figure 2: The IL-2 pathway page in NetPath

Hyperlinks to pathway specific information such as reactions, transcriptionally regulated genes, molecular associations, and catalysis events are listed. There is also an option to download pathway information in various data exchange formats from this page.

Table 1**Immune signaling pathway statistics**

	Pathway	Molecular association events	Catalysis events	Transport events	Total reactions	Number of upregulated genes annotated	Number of downregulated genes annotated	Number of PubMed links
1	T Cell Receptor	202	215	13	430	453	178	1153
2	B Cell Receptor	172	136	43	351	253	182	990
3	IL-1	55	44	9	108	161	79	461
4	IL-2	68	76	11	155	539	301	1289
5	IL-3	65	52	5	122	43	10	250
6	IL-4	59	47	5	111	222	90	519
7	IL-5	26	40	6	72	167	9	308
8	IL-6	65	58	7	130	84	25	332
9	IL-7	14	39	2	55	57	14	175
10	IL-9	14	20	4	38	25	1	103
Total	10	740	727	105	1572	2004	889	5580

Table 2**Comparison of salient features of NetPath with other popular curated signaling pathway resources**

	Pathway resource	Query option for pathway molecules	Whether genes transcriptionally regulated by the pathway are included ?	Whether pathways reviewed by pathway experts?	File formats available for download	Whether textual description of reactions provided ?	Other features or comments
1.	NetPath [5]	Yes	Yes	Yes	BioPAX, PSI-MI, SBML, Excel, Tab-delimited	Yes	Focus on human receptor mediated signaling. Also contains separate molecule pages with brief summary of the biology of the individual molecules.
2.	BioCarta [14]	Yes	No	Yes	No download option provided	No	BioCarta provides commercial links to antibody reagents
3.	Science's STKE [15]	No	No	Yes	SVG	No	Contains species specific and also cell type specific pathways
4.	KEGG [13]	Yes	No	No	KGML, BioPAX	No	Contains disease specific pathways

5.	Reactome [16]	Yes	No	Yes	BioPAX, SBML, PDF, SVG, Protégé, MySQL database dump	No	Also contains computationally inferred pathway reactions
6.	NCI-PID [17]	Yes	No	Yes	XML, BioPAX, SVG, JPG	No	Apart from NCI-Nature curated pathways, it also contains many pathways imported from BioCarta/Reactome
7.	CST [18]	Yes	No	Yes (in some cases)	PDF	No	Provides pathway information along with links to protein and commercial products available for that protein
8	WikiPathways [25]	Yes	No	No	GPML, GenMAP P, PDF, PNG, SVG	No	Any user can register and create a new pathway and also edit existing pathways
9	PANTHER [26]	Yes	No	Reviewed by Curation Coordinator	SBML, SBGN, PNG	No	Allows community pathway curation and also provides links to Applied Biosystems genomic products

NetPath

[Home](#)[Browse](#)[FAQs](#)[Immune Signaling Pathways](#)[Cancer Signaling Pathways](#)

Search

Gene Symbol e.g. [SHC1](#)
 Protein Name e.g. [EGF receptor](#)
 Accession Number e.g. [NP_002077.1](#)
 Pathway Name

Statistics

Curated Pathways	20
Molecules Involved	1682
Physical Interactions	1800
Genes Transcriptionally Regulated	6582
Transport	201
Enzyme Catalysis	1218
PubMed Citations	11739

'NetPath' is a curated resource of signal transduction pathways in humans. It is a collaboration between the [PandeyLab](#) at Johns Hopkins University and the [Institute of Bioinformatics](#). At this time, 10 immune and 10 cancer signaling pathways are available. The 10 cancer signaling pathways were developed in collaboration with the [Computational Biology Center](#) at [Memorial Sloan-Kettering Cancer Center](#) and with [Gary Bader's lab](#) at the University of Toronto for the '[Cancer Cell Map](#)'.

All 20 pathways are freely available in [BioPAX](#), [PSI-MI](#) and [SBML](#) formats at this website. [BioPAX](#) is an emerging standard for pathway data exchange. The 10 cancer signaling pathways are also available for download and browsing from the Cancer Cell Map project website. The pathways are freely available under an adaptive Creative Commons License which stipulates that the pathways may be used if adequate credit is given to the authors. These pathways are unpublished at this time, please cite the websites until the pathways are described in a publication.

Feedback on the NetPath website and its contents can be sent to info@netpath.org.

This is a joint project between:



IL-2 Signaling Pathway

Accession number : NetPath_14

Interleukin-2 belongs to a family of cytokines, which includes IL-4, IL-7, IL-9, IL-15 and IL-21. IL-2 signals through a receptor complex consisting of IL-2 specific IL-2 receptor alpha (CD25), IL-2 receptor beta (CD122) and a common gamma chain (γ_c), which is shared by all members of this family of cytokines. Binding of IL-2 activates the Ras/MAPK, JAK/Stat and PI 3-kinase/Akt signaling modules.

Pathway Statistics	
Molecules Involved	68
Physical Interactions	68
Enzyme Catalysis	76
Transport	11
Genes Transcriptionally Regulated	848

[Pathway Authority](#)
[Curators](#)
[Comments](#)

Molecules Involved in IL-2 signaling pathway (Total = 68)

[AKT1](#) [BCL2](#) [CBL](#) [CISH](#) [CREB1](#) [CRK](#) [CRKL](#) [EIF3S9](#) [EIF4E](#) [ETS1](#) [ETS2](#) [FOXO3A](#) [FRAP1](#) [FYN](#) [GAB2](#)
[GNB2L1](#) [GRB2](#) [HSP90AA1](#) [IL2](#) [IL2RA](#) [IL2RB](#) [IL2RG](#) [IRS1](#) [IRS2](#) [ITM2B](#) [JAK1](#) [JAK2](#) [JAK3](#) [LCK](#) [LYN](#)
[MAP2K1](#) [MAP2K2](#) [MAPK1](#) [MAPK3](#) [MKNK1](#) [NFKB1](#) [NMI](#) [NR3C1](#) [PIK3CA](#) [PIK3CB](#) [PIK3CD](#) [PIK3CG](#) [PIK3R1](#) [PIK3R2](#) [PIK3R3](#)
[PLCB1](#) [PRKCZ](#) [PTK2B](#) [PTPN11](#) [PTPN6](#) [RAF1](#) [RELA](#) [RPS6KB1](#) [SHB](#) [SHC1](#) [SOCS1](#) [SOCS3](#) [SOS1](#) [STAM](#) [STAM2](#)
[STAT1](#) [STAT3](#) [STAT5A](#) [STAT5B](#) [SYK](#) [TERT](#) [YAV1](#) [YBX1](#)

Reactions in IL-2 pathway (Total = 155)

[Physical Interactions](#)
[Enzyme Catalysis](#)
[Transport](#)

Genes Transcriptionally Regulated by IL-2 pathway (Total = 848)

[Upregulated genes](#)
[Downregulated genes](#)

Download Pathway

[BioPAX level 2.0](#)
[PSI-MI Version 2.5](#)
[SBML Version 2.1](#)

Download Transcriptionally Regulated Genes

[Excel Sheet](#)
[Tab Delimited](#)