WikiPathways: building research communities on biological pathways

Thomas Kelder[1,*], Martijn P. van Iersel[1], Kristina Hanspers[2], Martina Kutmon[1], Bruce R. Conklin[2], Chris T. Evelo[1,3] and Alexander R. Pico[2]

[1]Department of Bioinformatics—BiGCaT, Maastricht University, Maastricht, The Netherlands, [2]Gladstone Institute of Cardiovascular Disease, San Francisco, California, USA and [3]Netherlands Consortium for Systems Biology (NCSB), The Netherlands

ABSTRACT

Here, we describe the development of WikiPathways (http://www.wikipathways.org), a public wiki for pathway curation, since it was first published in 2008. New features are discussed, as well as developments in the community of contributors. New features include a zoomable pathway viewer, support for pathway ontology annotations, the ability to mark pathways as private for a limited time and the availability of stable hyperlinks to pathways and the elements therein. WikiPathways content is freely available in a variety of formats such as the BioPAX standard, and the content is increasingly adopted by external databases and tools, including Wikipedia. A recent development is the use of WikiPathways as a staging ground for centrally curated databases such as Reactome. WikiPathways is seeing steady growth in the number of users, page views and edits for each pathway. To assess whether the community curation experiment can be considered successful, here we analyze the relation between use and contribution, which gives results in line with other wiki projects. The novel use of pathway pages as supplementary material to publications, as well as the addition of tailored content for research domains, is expected to stimulate growth further.

INTRODUCTION

WikiPathways (http://www.wikipathways.org) is a resource for biological pathways in the form of a wiki. It serves as a repository for biological knowledge in the form of pathway diagrams and as platform for curating, sharing and publishing pathways. We launched WikiPathways in 2008 as an experiment in community-based curation of biological pathways (1). WikiPathways has continued to develop and has been adopted by the research community in several ways.

Here, we present the latest developments for WikiPathways focusing on two specific aspects. In the first part, we highlight new features developed in recent years, and show how they fit with our general philosophy of community curation and collaboration. In the second half, we analyze the size and activity of the community to assess the success of WikiPathways so far.

New features of WikiPathways

Pathway diagrams are found everywhere: in textbooks, in review articles, on posters and on whiteboards. Their utility is to turn abstract knowledge into an understandable visualization. WikiPathways enables biologists to capture their rich, intuitive mental models of biological pathways and avoid the incomprehensible 'hairball' state of networks derived from big data.

From this, understanding of how pathways could be used, we derived a number of assumptions that have guided the addition of new features. The first assumption is that manual curation leads to higher quality pathways in the long term. Consequently, we developed features that make manual interaction with the site easier. Our second assumption is that the success of WikiPathways is dependent on the community of curators, and thus we have added features to stimulate growth of the community and avoided features that increase the barrier to entrance for newcomers. Finally, it is our goal for WikiPathways to be a public resource for biological research. The content should be easily accessible for a wide range of

*To whom correspondence should be addressed. Tel: +31 88 866 4270; Email: thomaskelder@gmail.com
Correspondence may also be addressed to Alexander R. Pico. Tel: +1 415 734 2741; Fax: +1 415 355 0960; Email: apico@gladstone.ucsf.edu
Present addresses:
Thomas Kelder, TNO, Research Group Microbiology and Systems Biology, Zeist, The Netherlands.
Martijn P. van Iersel, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, UK.

applications, including those that we support directly, as well as those that we have not envisioned ourselves.

New developments in the pathway page

The most visible aspect of WikiPathways is the pathway page. Each pathway has a dedicated page that provides information on a specific biological mechanism, including the pathway diagram, description, hyperlinks to detailed information about genes, proteins and metabolites, and relevant literature references. Entities such as genes, proteins or metabolites in a pathway can be annotated with many different identifier systems, such as Ensembl (2), Entrez Gene (3) or ChEBI (4). Hyperlinks are provided to many different external sources of information, such as genome browsers, experimental platforms, protein databases, Gene Ontology and Wikipedia. These are directly available by clicking on a gene or protein box on the pathway page and allow the researcher to browse to detailed information related to components of the pathway. This way, the pathway page provides an organized summary of information related to a biological mechanism, and provides a starting point for researchers to browse through more specific and detailed resources.

Earlier versions of the pathway page contained a static image of the pathway, without interactive access to linked databases. Zooming in was possible only after opening the editor applet. Aiming to increase the manual interaction with the pathway page, we replaced the static pathway image with an interactive pathway viewer (Figure 1). This interactive viewer makes it possible to zoom and pan the diagram and click on elements in the pathway to view detailed information, similar to the popular Google maps interface. Genes, proteins and metabolites can be clicked for direct access to external databases. The interactive viewer also contains a search function that makes it easier to locate elements on the pathway diagram.

Organizing pathway information

Currently, WikiPathways contain over 1600 pathways and supports 21 species, including vertebrates, several plant species, bacteria and model organisms such as worm, yeast and fruit fly. To support manual interaction with this growing amount of knowledge, we attempted to make it easier to organize and browse. Therefore, we recently implemented an ontology annotation feature that allows each pathway to be annotated with terms from several ontologies covering different topics [these currently include the Pathway, Human disease and Cell type ontologies from the BioPortal collection (5)]. This allows users to exactly specify the context that applies to the pathway, e.g. a specific disease, organ or cell type. By using the pathway ontology, a hierarchical organization of the pathways can be created that groups smaller, more specific pathways (e.g. oxidative phosphorylation) into larger superpathways (e.g. energy metabolism). These annotations make it easier to find pathways related to a specific topic, by allowing users to search pathways by ontology terms. It will also facilitate data analysis and visualization methods, for example, a pathway interaction network could be built from the hierarchical organization in the pathway ontology to analyze biological mechanisms at different levels of detail.

Pathways as a public resource

WikiPathways is a public resource intended to benefit biological research worldwide. All content is freely available under the Creative Commons attribution
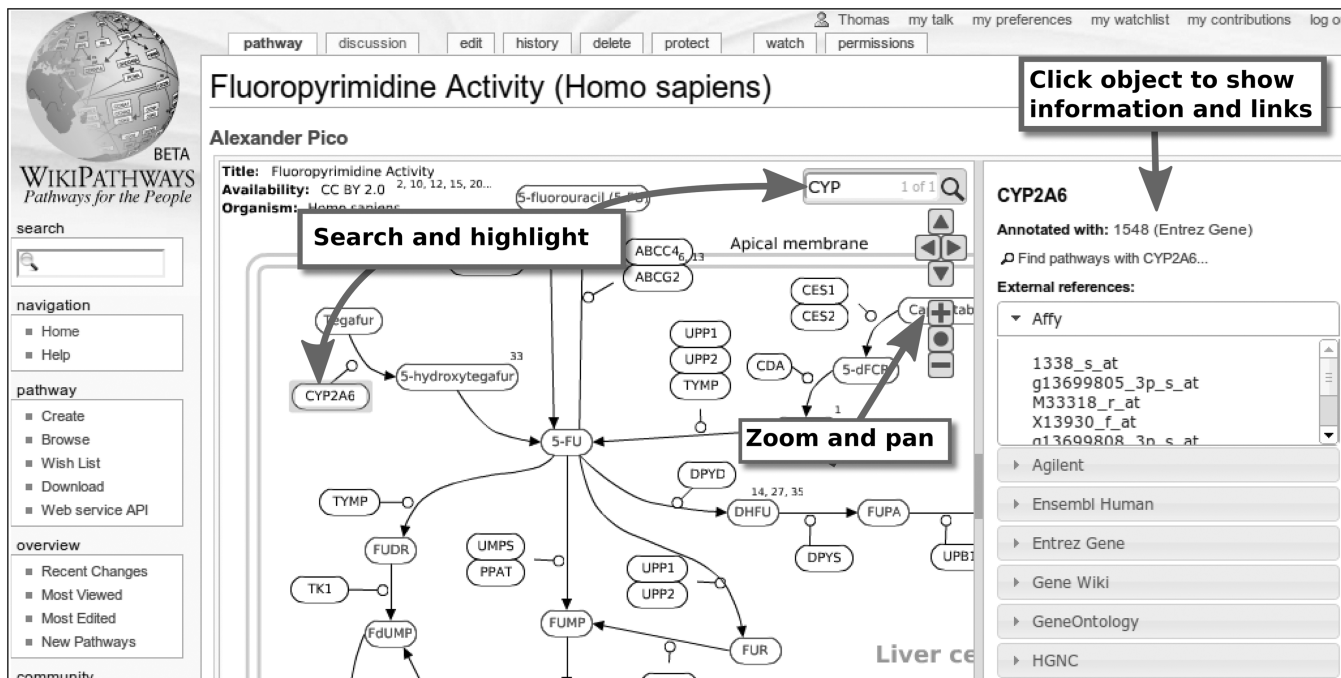

Figure 1. The new interactive pathway viewer.

license (6). Because of this license, there is no limit to re-distribution of pathway content. The WikiPathways content is already accessible from several external resources, such as NCBI BioSystems (7) and BioGPS (8). We are continually looking for new ways to make pathway information available to a wider public. For example, we recently added Interactive Pathway Maps into human gene and pathway-related articles at Wikipedia (9). Implemented as reusable templates, these interactive pathway maps link related Gene Wiki (10) articles. Each of these pathways at Wikipedia is viewed tens of thousands of times per day. We were careful not to add flagrant links to WikiPathways. Clicking on pathway at Wikipedia, for example, keeps you at Wikipedia. Nevertheless, we witnessed a 38% increase in new users and Wikipedia is consistently one of our top referring sites. Our pathway template passed through the Molecular and Cellular Biology Wikiproject proposal process and was vetted in accordance to rigorous guidelines by a number of 'Wikipedians' who specialize in biology-related content. WikiPathways content can now be deployed as 'stub' articles to initiate new articles around important biology topics.

Each pathway in WikiPathways carries an identifier of the form 'WP1234' (where 1234 can be a number of any size), which enables stable URLs to be formed to link directly to a pathway. This simple yet effective feature is important for online collaboration, as a URL can be shared via email. We recently added support for linking directly to genes, proteins or metabolites in the pathway. The linked element is opened in the new interactive pathway viewer and highlighted. This way users can be directly pointed to an element of interest, for example, from a search result or external resource. The pathway viewer can also be included in any website as a widget, for example to share a pathway of interest via a blog post, or to publish a pathway on the project page of a research group. These new ways of linking to pathway content serve to increase the cohesion of WikiPathways with other online biology and bioinformatics resources.

In some use cases, it is not desirable that a new pathway is immediately publicly available. For example, when the pathway is used as supplementary material to a still unpublished manuscript, it should only be visible to a limited number of users. Therefore, we implemented the possibility to create a pathway, but postpone its publication by temporarily marking it as private and thereby hiding it from public view. The pathway author can then set permissions for specific user accounts, for example, to allow only collaborators to view and edit the pathway. This way authors can retrieve a stable identifier that can be used to send as URL to collaborators or added as reference in a manuscript and allow referees to access the pathway during the peer review process. By default, the pathway will automatically become public after 1 month, but the author can actively postpone this deadline each month. By requiring a periodic action to prevent the pathway from becoming public, we expect that all private pathway information will eventually become publicly available to the WikiPathways community.

Pathways in bioinformatics applications

Pathways produced at WikiPathways are in a format that can be directly used in downstream data analysis by a number of software tools. Thus, we complete a cycle starting with researcher knowledge that when synthesized with standardized data, leads to novel pathway models that can be used to visualize and analyze other data sets, leading to new insights, experiments and knowledge.

WikiPathways content is distributed through numerous online resources and bioinformatics software packages. We provide pathways in an open, XML standard format, called GPML, which is explicitly compatible with a handful of analysis tools, such as GenMAPP (11), PathVisio (12), Cytoscape (13) and GO-Elite (14). These tools support various workflows involving visualization and analysis of experimental data. The GPML format can be made compatible with any tool that chooses to use it since it is cross-platform, open and actively supported. For an even broader audience, we provide our pathways in BioPAX (15) format as well. This for example allows integration of the content into pathway unification efforts such as Pathway Commons (16).

Pathway information is also available through our open web service API, providing access to WikiPathways content to a broad spectrum of software developers (17). This web service processed over 45,000 requests by external scripts per month over 2010. It can be used to integrate pathway information directly from WikiPathways into scripts, data analysis workflows or external tools. An example of a web application that uses the WikiPathways web service for pathway analysis is WebGestalt (18), which allows researchers to find over-represented pathways from a user-specified input list. An example of a locally installed tool that integrates WikiPathways content via the web service is DomainGraph (19), a plugin for the network analysis tool Cytoscape that can be used to visualize alternative-splicing data.

Growth and activity of the WikiPathways community

The goal of WikiPathways is to capture knowledge about biological pathways (the elements, their interactions and layout) in a form that is both human readable and amenable to computational analysis. Curating and maintaining a public collection of pathways is a large and never-ending task, as a continuous stream of new knowledge is being generated. Given the current fast growth of knowledge, centralized curation as employed in existing pathway resources may not scale well (20). Wikis provide an effective platform for community-based curation that may be better scalable since users can directly contribute, update and expand the content. Given a large and active enough group of users, this improves comprehensiveness and quality of the content in the long term. The mechanism behind this can be described as a positive feedback loop (21).
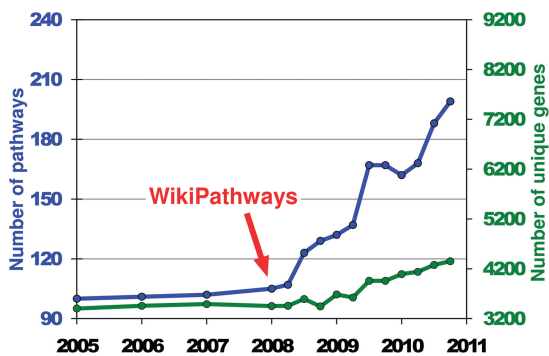
Figure 2. Size and growth of the human pathway archive before and after launching WikiPathways. The numbers for years 2005–2008 are derived from the size of the GenMAPP archive, which served as the initial seed content for WikiPathways.

Figure 3. Relation between views and edits for the pathways at WikiPathways. Each dot in the graph represents a group of pathways that each has been edited by the same number of unique authors (indicated by color). The size of the dot represents the number of pathways in that group. The x-axis represents the number of times the pathway has been viewed (averaged over each group) and the y-axis represents the number of times the pathway has been edited (averaged over each group).

The wiki starts with some initial content which attracts users. A small fraction of these users will also correct or extend the content, which on average leads to an improvement of the quality and usefulness of the content. This will attract more users, thereby also increasing the number of potential editors, which improves the content even more. Based on the statistics gathered during the short history of WikiPathways, it might be possible to obtain more insight into how well the wiki approach has worked so far for biological pathways.

If a relation would exist between usage and contribution in WikiPathways, the number of edits to a pathway is expected to increase with the number of views. Indeed, pathways that have many views are also among the ones with the most edits and have also been edited by a larger group of authors (Figure 3). Following the positive feedback mechanism, usage would also result in a growth of content. Since the number of site visits has increased from over 1100 per month over the 3 months prior to publication to almost 5400 per month over the last 3 months of 2010, a growth and improvement of content would also be expected. Indeed, compared to the initial content of WikiPathways, the number of human pathways has grown by 128% and the number of annotated human genes in these pathways has increased with 30% (Figure 2). To put this number in perspective, the pathway collection for GenMAPP, on which the initial WikiPathways content was based, grew with only 1% in number of pathways and 5% in number of genes in the 3 years preceding WikiPathways.

In addition to the usage and content growth of WikiPathways, the size and activity of its community have also grown. Since January 2008, WikiPathways went from 100 to over 1800 registered users with an increasing percentage of members creating and editing pathways. In 2008, on average 10 users per month made one or more edits to a pathway and 87 edits were made per month. These numbers have grown to on average 16 editing users and 261 edits per month over 2010, a growth of 56 and 200%, respectively. The barrier to contribute still seems fairly high, since on average only 0.36% of the website visitors actually edited a pathway
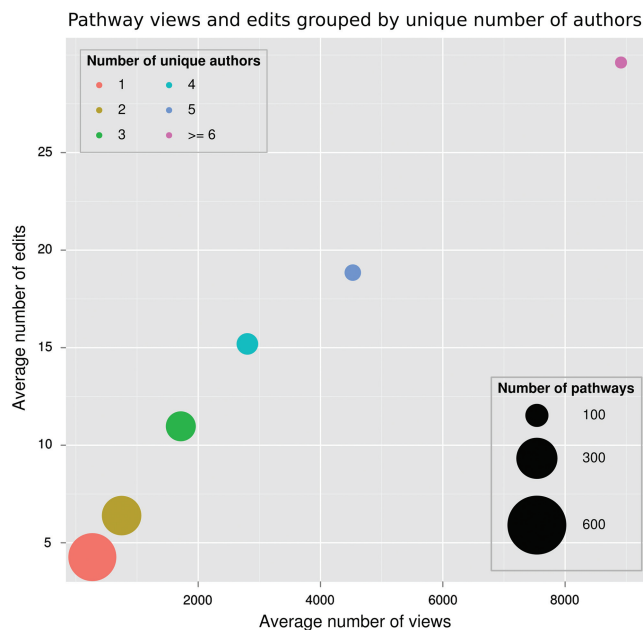
one or more times. However, when compared to Wikipedia, a wiki with a very active community, these numbers seem more reasonable. For the English Wikipedia, only 0.02–0.03% of the visitors are active contributors [defined as at least five edits in a given month (22)] and for WikiPathways this translates to almost 0.19% averaged over 2010. However, in contrast to Wikipedia, the content at WikiPathways is focused on a smaller domain and a large part of the target audience is expert in this domain. Therefore, it is worth trying to lower this barrier even further.

Although the active community shows growth, it remains relatively small, probably too small to effectively keep up with the growth of biological knowledge that can be captured in pathways. The size and activity of the community can be improved in two main ways that reinforce each other. First, the portion of users that become contributors can be improved to increase the activity of the community. Second, improving usability of the pathways will increase the size of the community, since the number of users is proportional to the utility of the content. The following sections will highlight several use cases that we are supporting and actively stimulating to grow the active community and improve usability of content.

Pathway publishing

Pathway diagrams are widely used in scientific publications as figures accompanying the text. These figures are being translated into annotated pathways in digital

form; however, this is a rather tedious and often error prone task. Using WikiPathways, authors can directly create a fully annotated pathway and either save it supplementary material. Several authors have already taken this approach (23,24), which offers additional advantages from the author's point of view. First, the pathway page at WikiPathways improves the reader experience, by providing links to additional detailed information and unambiguously annotated genes, proteins and metabolites. In contrast, when using a static image, manual searching is required to find more information about a protein on the diagram and ambiguous protein names may lead to incorrect interpretations. Secondly, the online pathway improves visibility and usability of the author's results, by making it searchable via internet search engines such as Google and by redistributing it with several bioinformatics tools for use in data analyses of other researchers. When such analyses lead to new publications, the original source of the pathway can be tracked and cited, thereby increasing its measurable impact. Once at WikiPathways, the pathway can be updated by both the original author and the community based on new research results that may become available over time to provide a better representation of the biological mechanism. The new private pathways feature has increased the usability of WikiPathways as publishing tool and we hope this encourages more authors to publish their pathway diagram as supplementary material at WikiPathways.

Curated pathway diagrams describing novel findings and perspectives can also be published as posters, presentations and initial research reports at *Nature Precedings*. We are establishing a WikiPathways Collection at *Nature Precedings* through which we will encourage, collect and promote publications from the community. This is an attractive and innovative publishing route for pathway curation efforts that are not yet associated with traditionally defined 'publishable' results.

Building communities

The explosion of wikis and social curation tools in the biological sciences is a testament to the demand for such involvement across a wide array of subdisciplines (e.g. model organisms, genes, SNPs, structures). These communities were not *created* by wiki tools (anyone who has tried to start a wiki knows that the statement 'build it and they will come' does not apply). Rather, a wiki simply enables and gives coherence to a community that already existed, but had yet to find each other. Thus, the real innovation of WikiPathways is not necessarily the wiki technology, but rather the fact that we revealed this potential community of pathway curators. This innovation is changing the definition of what a pathway resource is and does. A main focus for pathway databases has been to collect and curate a set of canonical pathways. But, more recently, we are seeing many WikiPathways contributors focus on content that is tailored to a particular research perspective. Thus, we are breaking the 'canonical' mold. For example, pathways related to topics such as

pluripotency, heart development, miRNA, addiction, SIDS, ossification, aflatoxin were contributed to WikiPathways, which are typically not considered among canonical pathways. In the next phase of growth, our innovation will be to enable not merely a curation community, but an expanding collection of curation communities, each with their own research interests. If we properly capitalize on this innovation, our distributed model could experience exponential growth, not possible by a traditional resource, while simultaneously increasing quality.

WikiPathways has already been used by specific curation communities to build and maintain a set of focused pathways supporting ongoing projects or research collaborations. For example, the Micronutrient Genomics project (25) aims to provide a resource for knowledge on the biological context of micronutrients and uses WikiPathways to collaboratively edit a subset of pathways related to these topics. A core team of experts builds pathways and streamlines contributions from the community. Along the same lines, the California Institute for Regenerative Medicine (CIRM) has adopted WikiPathways to highlight a subset of pathways contributed by the stem cell research community. For initiatives like these, WikiPathways provides the option to create a portal page, which provides an access point to the subset of pathways and can be customized with the project logo and announcements. Such portals make the content more attractive for a specific group of users because it provides a more convenient entry point that is focused on their research subject.

Supporting centrally curated databases

WikiPathways also provides a framework for collecting community contributions for centrally curated databases. In this setup, the content of the database is mirrored on WikiPathways which provides a medium for its users to contribute corrections or additions. This way, WikiPathways complements centrally curated databases by providing a staging ground for new content that can then be reviewed by appointed curators for inclusion in the database. The Reactome database is currently in the process of setting up this workflow using WikiPathways to improve their ability to collect community contributions (26). A similar approach at smaller scale is taken by the maintainers of the PluriNetwork (27), an electronic resource for curated protein interactions relevant to pluripotency, for which a version is maintained at WikiPathways that allows users to contribute new interactions. As another example, the NetPath cancer and immune signaling pathways are also maintained at WikiPathways to complement their focused, more stringent, system of curation (28). Partnering with other pathway or interaction resources directly increases the community and contribution rate of WikiPathways. In addition, it increases database inter-compatibility and makes each contribution from the community more valuable, because it will eventually be distributed over different pathway resources.

Future plans

Improving the curation experience. We aim to improve the curator tools by bringing relevant data to the curator. For example, if the pathway already contains a given gene, we could provide the curator with known interactions n-degrees from that gene, or other genes related by functional annotation. The benefits to the curator are 2-fold. First, they have access to relevant snippets of data that are otherwise buried in various databases. And, second, they can copy the snippet into the pathway they are editing and maintain the relationships and annotations from the source database, including evidence codes and literature references. This way both ease and quality of user contributions can be improved.

Structured pathway content. The anticipated growth of WikiPathways necessitates a forward-looking approach to data storage and representation. There are three major considerations: scalability, accommodating constant and dynamic data changes, and the ability to support complex queries. Complex queries are an essential consideration because WikiPathways users will need more than basic 'select' and 'join' query access. They will need to be able to query across multiple resources and levels, essentially accessing various, dynamically generated 'super pathways'. This goes well beyond the capabilities of our current web service API and MySQL database solutions. Therefore, we aim to make WikiPathways content more accessible and connected through semantic technologies. We will extend WikiPathways with customized semantic components and derive triples from our structured GPML content, inferred pathway information, pathway metadata and selected external content. By periodically synchronizing our semantic data with major biological data repositories, our content can be effectively connected with these massive and growing collections. This way, our data will be accessible to growing numbers of semantic tools for advanced search, data integration and bioinformatics analysis.

Public data integration. To enhance the usability of the pathway information on the WikiPathways website, we aim to support direct integration of publicly available data and allow the user to customize the information content displayed on a given pathway page. For example, we could directly map reference gene expression data, for example to visualize the expression level of the genes in a pathway in a given tissue. Or by querying gene–disease associations, we could display a ranked list of potentially relevant disease terms per pathway. This way, WikiPathways may actually be serving as a high-level knowledge management tool, providing researchers and domain experts access to related snippets of data from disparate resources, enabling them to annotate and qualify new connections in the context of biological pathways, and ultimately producing novel snippets of data for future reuse.

CONCLUSION

We presented new features of WikiPathways that improve both usability and curation of pathway content. The growth in content active community indicates that WikiPathways is being adopted by the research community as pathway resource as well as a framework for publishing and curating biological knowledge. Eventually, both usage and contribution need to reach a critical mass to establish a stable active community that can keep up with the continuously growing amount of knowledge and publications. The open nature of the WikiPathways project (both in content, code and collaborations) allows different user groups to adapt and implement features to support a specific use case, and stimulates new communities to adapt WikiPathways for their research. In the end, we hope this will contribute to a better representation of our knowledge as biological pathways, and contribute to improving exploratory pathway analysis.

REFERENCES

1. Pico,A.R., Kelder,T., van Iersel,M.P., Hanspers,K., Conklin,B.R. and Evelo,C. (2008) WikiPathways: pathway editing for the people. *PLoS Biol.*, **6**, e184.
2. Flicek,P., Amode,M.R., Barrell,D., Beal,K., Brent,S., Chen,Y., Clapham,P., Coates,G., Fairley,S., Fitzgerald,S. *et al.* (2011) Ensembl 2011. *Nucleic Acids Res.*, **39**, D800–D806.
3. Maglott,D., Ostell,J., Pruitt,K.D. and Tatusova,T. (2011) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **39**, D52–D57.
4. de Matos,P., Alcántara,R., Dekker,A., Ennis,M., Hastings,J., Haug,K., Spiteri,I., Turner,S. and Steinbeck,C. (2010) Chemical entities of biological interest: an update. *Nucleic Acids Res.*, **38**, D249–D254.
5. Noy,N.F., Shah,N.H., Whetzel,P.L., Dai,B., Dorf,M., Griffith,N., Jonquet,C., Rubin,D.L., Storey,M.-A., Chute,C.G. *et al.* (2009) BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.*, **37**, W170–W173.
6. Creative Commons Attribution 3.0 Unported. Available at: http:// creativecommons.org/licenses/by/3.0/ (Accessed 6 August 2011).
7. Geer,L.Y., Marchler-Bauer,A., Geer,R.C., Han,L., He,J., He,S., Liu,C., Shi,W. and Bryant,S.H. (2010) The NCBI BioSystems database. *Nucleic Acids Res.*, **38**, D492–D496.
8. Wu,C., Orozco,C., Boyer,J., Leglise,M., Goodale,J., Batalov,S., Hodge,C., Haase,J., Janes,J., Huss,J. *et al.* (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.*, **10**, R130.
9. Interactive Pathway Maps at Wikipedia. Available at: http://en .wikipedia.org/wiki/Template:Interactive_pathway_maps/doc (Accessed 6 August 2011).
10. Huss,J.W., Lindenbaum,P., Martone,M., Roberts,D., Pizarro,A., Valafar,F., Hogenesch,J.B. and Su,A.I. (2010) The Gene Wiki: community intelligence applied to human gene annotation. *Nucleic Acids Res.*, **38**, D633–D639.
11. Salomonis,N., Hanspers,K., Zambon,A., Vranizan,K., Lawlor,S., Dahlquist,K., Doniger,S., Stuart,J., Conklin,B. and Pico,A. (2007) GenMAPP 2: new features and resources for pathway analysis. *BMC Bioinform.*, **8**, 217.

12. Iersel,M., van, Kelder,T., Pico,A., Hanspers,K., Coort,S., Conklin,B. and Evelo,C. (2008) Presenting and exploring biological pathways with PathVisio. *BMC Bioinform.*, **9**, 399.
13. Smoot,M.E., Ono,K., Ruscheinski,J., Wang,P.-liang and Ideker,T. (2010) Cytoscape 2.8: new features for data integration and network visualization two illustrative examples: integration of custom graphics with equations. *Bioinformatics*, **27**, 431–432.
14. GO-Elite, Software for Extended Pathway Analysis. Available at: http://www.genmapp.org/go_elite/ (Accessed 6 August 2011).
15. Demir,E., Cary,M.P., Paley,S., Fukuda,K., Lemer,C., Vastrik,I., Wu,G., D'Eustachio,P., Schaefer,C., Luciano,J. *et al.* (2010) The BioPAX community standard for pathway data sharing. *Nat. Biotechnol.*, **28**, 935–942.
16. Cerami,E.G., Gross,B.E., Demir,E., Rodchenkov,I., Babur,O., Anwar,N., Schultz,N., Bader,G.D. and Sander,C. (2011) Pathway commons, a web resource for biological pathway data. *Nucleic Acids Res.*, **39**, D685–D690.
17. Kelder,T., Pico,A.R., Hanspers,K., van Iersel,M.P., Evelo,C. and Conklin,B.R. (2009) Mining biological pathways using WikiPathways web services. *PloS One*, **4**, e6447.
18. Duncan,D., Prodduturi,N. and Zhang,B. (2010) WebGestalt2: an updated and expanded version of the Web-based Gene set analysis toolkit. *BMC Bioinform.*, **11**, P10.
19. Emig,D., Salomonis,N., Baumbach,J., Lengauer,T., Conklin,B.R. and Albrecht,M. (2010) AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res.*, **38**, W755–W762.
20. Howe,D., Costanzo,M., Fey,P., Gojobori,T., Hannick,L., Hide,W., Hill,D.P., Kania,R., Schaeffer,M., St Pierre,S. *et al.* (2008) Big data: the future of biocuration. *Nature*, **455**, 47–50.
21. Huss,J.W., Orozco,C., Goodale,J., Wu,C., Batalov,S., Vickers,T.J., Valafar,F. and Su,A.I. (2008) A gene wiki for community annotation of gene function. *PLoS Biol.*, **6**, e175.
22. Wikimedia Users—Strategic Planning. Available at: http://strategy.wikimedia.org/wiki/Wikimedia_users#Contributors_as_a_percentage_of_all_visitors (Accessed 6 August 2011).
23. Rückert,F., Dawelbait,G., Winter,C., Hartmann,A., Denz,A., Ammerpohl,O., Schroeder,M., Schackert,H.K., Sipos,B., Klöppel,G. *et al.* (2010) Examination of apoptosis signaling in pancreatic cancer by computational signal transduction analysis. *PLoS ONE*, **5**, e12243.
24. Jennen,D.G.J., Gaj,S., Giesbertz,P.J., van Delft,J.H.M., Evelo,C.T. and Kleinjans,J.C.S. (2010) Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes. *Drug Discov. Today*, **15**, 851–858.
25. Ommen,B., El-Sohemy,A., Hesketh,J., Kaput,J., Fenech,M., Evelo,C.T., McArdle,H.J., Bouwman,J., Lietz,G., Mathers,J.C. *et al.* (2010) The micronutrient genomics project: a community-driven knowledge base for micronutrient research. *Genes Nutr.*, **5**, 285–296.
26. Croft,D., O'Kelly,G., Wu,G., Haw,R., Gillespie,M., Matthews,L., Caudy,M., Garapati,P., Gopinath,G., Jassal,B. *et al.* (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.*, **39**, D691–D697.
27. Som,A., Harder,C., Greber,B., Siatkowski,M., Paudel,Y., Warsow,G., Cap,C., Schöler,H. and Fuellen,G. (2010) The PluriNetWork: an electronic representation of the network underlying pluripotency in mouse, and its applications. *PloS One*, **5**, e15165.
28. NetSlim Curation Criteria. Available at: http://www.netpath.org/netslim/criteria.html (Accessed 2 October 2011).