

P³DB 3.0: From plant phosphorylation sites to protein networks

Qiuming Yao^{1,2}, Huangyi Ge¹, Shangquan Wu¹, Ning Zhang^{1,2}, Wei Chen^{2,3},
Chunhui Xu^{2,4}, Jianjiong Gao⁵, Jay J. Thelen^{2,6} and Dong Xu^{1,2,*}

¹Department of Computer Science, University of Missouri, Columbia, MO 65211, USA, ²Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA, ³School of Communication and Information Engineering, Shanghai University, Shanghai 200444, People's Republic of China, ⁴Department of Biology, Brandeis University, MA 02453, USA, ⁵Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA and ⁶Department of Biochemistry, University of Missouri, Columbia, MO 65211, USA

Received September 15, 2013; Revised October 20, 2013; Accepted October 24, 2013

ABSTRACT

In the past few years, the Plant Protein Phosphorylation Database (P³DB, <http://p3db.org>) has become one of the most significant *in vivo* data resources for studying plant phosphoproteomics. We have substantially updated P³DB with respect to format, new datasets and analytic tools. In the P³DB 3.0, there are altogether 47 923 phosphosites in 16 477 phosphoproteins curated across nine plant organisms from 32 studies, which have met our multiple quality standards for acquisition of *in vivo* phosphorylation site data. Centralized by these phosphorylation data, multiple related data and annotations are provided, including protein–protein interaction (PPI), gene ontology, protein tertiary structures, orthologous sequences, kinase/phosphatase classification and Kinase Client Assay (KiC Assay) data—all of which provides context for the phosphorylation event. In addition, P³DB 3.0 incorporates multiple network viewers for the above features, such as PPI network, kinase-substrate network, phosphatase-substrate network, and domain co-occurrence network to help study phosphorylation from a systems point of view. Furthermore, the new P³DB reflects a community-based design through which users can share datasets and automate data depository processes for publication purposes. Each of these new features supports the goal of making P³DB a comprehensive, systematic and interactive platform for phosphoproteomics research.

INTRODUCTION

Phosphorylation is one of the most pervasive protein modification types in plants. Phosphorylation and dephosphorylation act as an important switch in signal transduction, chemical metabolism and other inter- or intra-cellular processes (1). In eukaryotes, *O*-phosphorylation (serine, threonine and tyrosine) predominates the landscape of protein phosphorylation. The burgeoning amount of experimental phosphorylation site data has necessitated the development of databases to warehouse these data and provide an essential infrastructure for the research community.

P³DB debuted in 2009 (2) when there was a need for depositing, requesting and sharing the wealth of experimental plant phosphorylation data beyond the reference plant *Arabidopsis*. Since then, P³DB has been actively developed and regularly updated with new datasets and features. Since the initial release of P³DB, high-quality phosphorylation sites in this database have accumulated at a rapid pace due to improvements in enrichment techniques and mass spectrometry [Supplementary Figure S1a and b]. Most of the datasets in the database came from large-scale experiments (MS/MS) (3), although several smaller datasets were also deposited. To help users analyze the proteome-wide phosphorylation data more systematically, the new P³DB 3.0 provides more information and annotations about phosphoproteins such as gene ontology, homolog, 3D structures, kinase and phosphatase families, protein–protein interactions (PPIs) and protein domains, together with protein–protein networks, kinase-substrate or phosphatase-substrate networks and domain co-occurrence networks (4).

Although plant phosphoproteomics has its origins in *Arabidopsis*, at present there are more experimentally-

*To whom correspondence should be addressed. Tel: +1 573 882 2299; Fax: +1 573 882 8318; Email: xudong@missouri.edu

mapped phosphorylation sites in nonmodel plants. As discovery plant phosphoproteomics extends beyond model organisms it is desirable to integrate and compare the diversity of phosphorylation events to fully interrogate the possibilities of regulation in *Viridiplantae*. In this regard P³DB aims to be a resource for the entire plant community. P³DB not only collects user-suggested datasets, but also allows research groups to directly deposit their data for the whole community or to share within a group. At the same time, they can collaborate and interact through this platform.

Besides P³DB, there are a number of other useful web resources for phosphorylation or other PTM data. HPRD, the Human Protein Reference Database (5,6), covers a wide range of PTM data including phosphorylation. However, it is restricted to human. PhosphoSitePlus (7) provides resources for integrating signaling pathways, but it only includes data for human, rat and mouse. PhosphoELM (8) contains many kinase-specific data, but it is not developed for plant phosphorylation data either. PhosphAT (9) is a rich resource for phosphorylation exclusively for Arabidopsis. However, it does not cover other plant species. MASCIP Gator (10) is also a comprehensive resource for proteomic data integration in Arabidopsis, but it does not host the data itself. Furthermore, none of the above web services provide any features for network-based analysis or community-based services. Thus, P³DB is a unique, complementary database to the current databases for its broad coverage of plant species, network-based data presentation and visualization, and community-based data services.

MATERIALS AND METHODS

Datasets

The datasets in P³DB 3.0 are curated from literature, online resources and in-house collaborations. With 32 studies (Supplementary Table S1) from nine plant species included, P³DB 3.0 now has 47923 nonredundant phosphorylation sites in 16477 phosphoproteins. *Arabidopsis thaliana* (contributing 30.15% to the phosphoproteins total data), *Medicago truncatula* (25.36%) and *Oryza sativa* (29.31%) are the three plants having the most phosphorylation data in P³DB (Supplementary Figure S1c).

P³DB also supports private datasets, which can be shared within research groups or with manuscript reviewers by a simple link that has password protection. These datasets can be merged to the public dataset after official acceptance of the publication or by user's authorized release.

Data quality criteria

Currently only *in vivo* experimental data are collected and archived in P³DB, except for the *in vitro* data obtained to identify kinase-client relationships, i.e. KiC assay (11,12) data. Most of the data are from high-throughput experiments from different laboratories using different instrumentation and data mining strategies; therefore, the quality of the data varies among different studies. P³DB

employs a strict data selection criterion: False Discovery Rate (FDR) <1% and <15 ppm precursor mass accuracy, to make sure phosphopeptide identification is of high quality.

Web services and interface

P³DB provides a user-friendly, interactive interface for data access. It is implemented by a back-end MySQL database, server-side PHP code and front-end Javascript and jQuery tools. The web services and interface follow the new definitions of HTML5 and CSS3, which ensure the smooth running on any html5-compliant systems including mobile devices. Cytoscape (13) JavaScript APIs are used in the new P³DB for displaying the network features.

Ontology

The gene ontology hierarchical architecture is retrieved from the ontology website (14,15) and integrated into P³DB. The annotations of plant proteins are from TAIR (16) and Phytozome (17). The ontology data can be used for searching and browsing phosphoproteins in different functional categories on the website.

Plant kinase and phosphatase family

The tree structure of the kinase and phosphatase families is obtained from PlantsP (18). The kinases or phosphatases from other data sources are mapped to the tree structure by their corresponding annotations. In the case that the kinase cannot be assigned to any of the subclasses, it will be assigned as an unknown class or general kinase class.

Kinase-Client assay

The KiC (Kinase-Client) assay is a technology to identify the phosphorylation sites associated with a kinase activity using a mass spectrometry-based *in vitro* screening process (11,12). A typical KiC assay requires purified kinase and a peptide library after which phosphorylation is quantified by mass spectrometry. After the peptide is identified as a target of a certain kinase, the whole-length protein can be used to further validate the kinase-substrate relationship (12). The KiC assay is a systematic way to screen the kinase-client relationship and help construct kinase-substrate networks.

Protein 3D structure

Protein structures are obtained from the Protein Data Bank (PDB) (19) with IDs mapped from the Uniprot (20,21). Some of the structures for Arabidopsis are predicted protein structures at Interactome 3D (22). Jmol (23) is used to visualize the tertiary structures.

PPI network

The PPI data are collected from four major databases: Biogrid (24), Intact (25), DIP (26) and Mint (27). The PPI data in plants other than Arabidopsis are still very limited. The interaction data are visualized in networks,

by calculating direct interactions or indirect interactions with intermediate nodes.

Protein domain network

Domains are structure- and function-independent units in proteins. The kinase domains, phosphatase domains, substrate domains and phosphorylation recognition domains are very important for understanding the functions of phosphorylation events. These domains may be fused into a single peptide to facilitate the phosphorylation activities, such as tyrosine kinase receptors. Such fusions can be represented by a domain co-occurrence network, where two domains form a link if they occur in the same protein. The domain library (Pfam A 27.0) is downloaded from the Pfam website (28), together with the proteomic annotations.

NEW AND UPDATED FEATURES

Protein chart

The protein chart provides information of local amino acid properties around phosphorylation sites, as shown in Figure 1. Phosphorylation sites are highlighted in green circles on the graph, and other related information

is aligned to the phosphorylation sites, including predicted hydrophobicity values, involved domains and predicted disorder scores (29). Phosphorylation is overrepresented in disordered regions, as shown in previous studies (30–32). It is also useful to present the substrate domain so that functional information may be revealed, since the substrate domain can be a regulatory or recognition domain in the downstream signaling cascade. The hydrophobicity often has a low value at the phosphorylation site, which indicates that phosphorylation sites are generally more hydrophilic than the background. This is not surprising as phosphorylation sites are usually exposed to the surface and are in disordered regions. Thus, the protein chart may be helpful to build hypotheses based on protein function and amino acid properties. The flexible architecture allows P³DB to display more potential position-specific factors in the future like protein-binding sites or polarity.

Orthologous sequence

Archiving the phosphoproteome of nonmodel plants in P³DB affords the plant biologist access to a larger, more complete resource of regulatory phosphorylation events. Questions about functional conservation can be directly queried through a new feature in version 3.0.

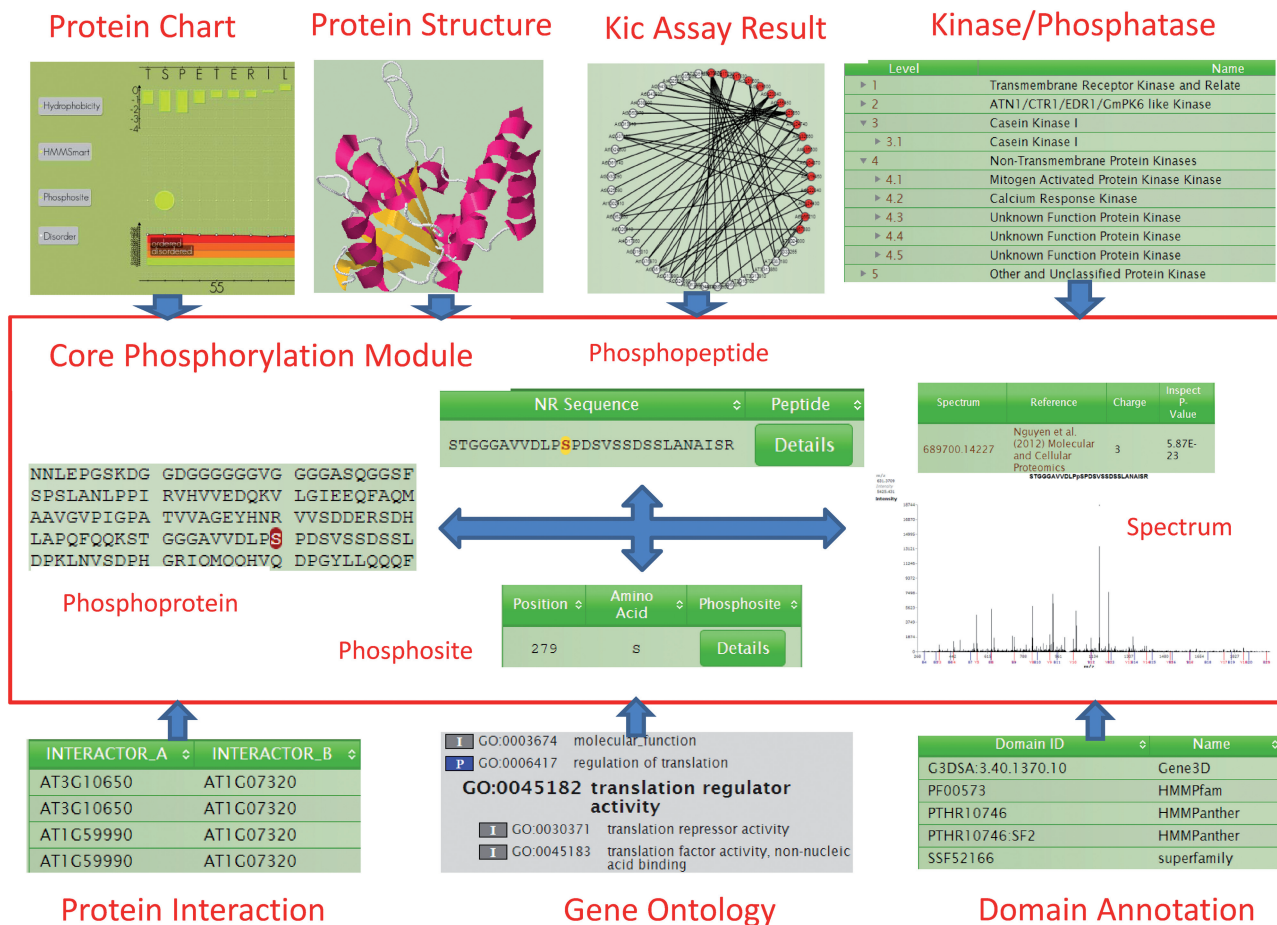


Figure 1. Module-based P³DB functionalities.

Phosphorylation sites may or may not be conserved among sequences in the same orthologous group in plants (31). Orthologous groups identified by MCL (33) are used in P³DB. Orthologous sequences are aligned based on positions; thus, it is easy to see whether the phosphorylation site is conserved. However, if the expected phosphorylation event in the ortholog is not observed, it may be due to lack of the experimental evidence in our database.

Gene ontology

The ontology terms can be browsed or searched in a hierarchical view. The tree view and the ontology are cross-linked between the protein page and the ontology page. On the protein page, the gene ontology terms are listed to help the user understand the functions of the phosphorylated protein. On the ontology page, the tree structure shows the parental and sibling terms, which help the user to navigate among related terms, and phosphorylated proteins under each ontology term are listed explicitly (Figure 2).

Kinase and phosphatase family

Kinases or phosphatases can be searched or browsed in a family tree view. This module is also cross-listed in the protein page. If a protein is annotated as a kinase, phosphatase or kinase substrate, the protein page will show this information. Notice that a protein can be both an enzyme and substrate. Every kinase or phosphatase is listed under the family tree view with a certain category. Some proteins are categorized to the unknown type or parent family due to lack of annotation.

Site prediction portal

Although we only curate experimental datasets, the green plant specific prediction of phosphorylation sites is available in P³DB through an embedded link to Musite (31) in the protein page.

Taxonomy browser

The taxonomy browser helps the user to explore the kingdom of *Viridiplantae* (green plants). The search page is also available for taxonomy information. The species page is cross-linked to phosphorylation datasets if available.

PPI network

Phosphoproteins in P³DB are also visualized in the context of PPI networks. Hypotheses in terms of the potential cause and effect can be constructed based on the PPI network and phosphorylation sites. For example, by searching interactions of AT4G26070, a MAPKK in Arabidopsis, the other two proteins AT4G01370 (MAPK4) and AT4G08500 (MAP3K) are found to have interactions with this MAPKK. The edge represents the data source validated or experimental verified relationships. In most cases, a single interaction is supported by multiple experiments or data resources. As seen from the interactions with these three MAP Kinases, there are 11 experimental evidences on them (34–37) (Figure 3a).

If two proteins do not interact with each other directly, P³DB will use the shortest path algorithm to find an interaction path that can connect these proteins. Based on the data and this algorithm, the pairwise linkage can be always found if it exists. This function is especially helpful to discover long-range relationships. The P³DB PPI network can also be expanded for each node with

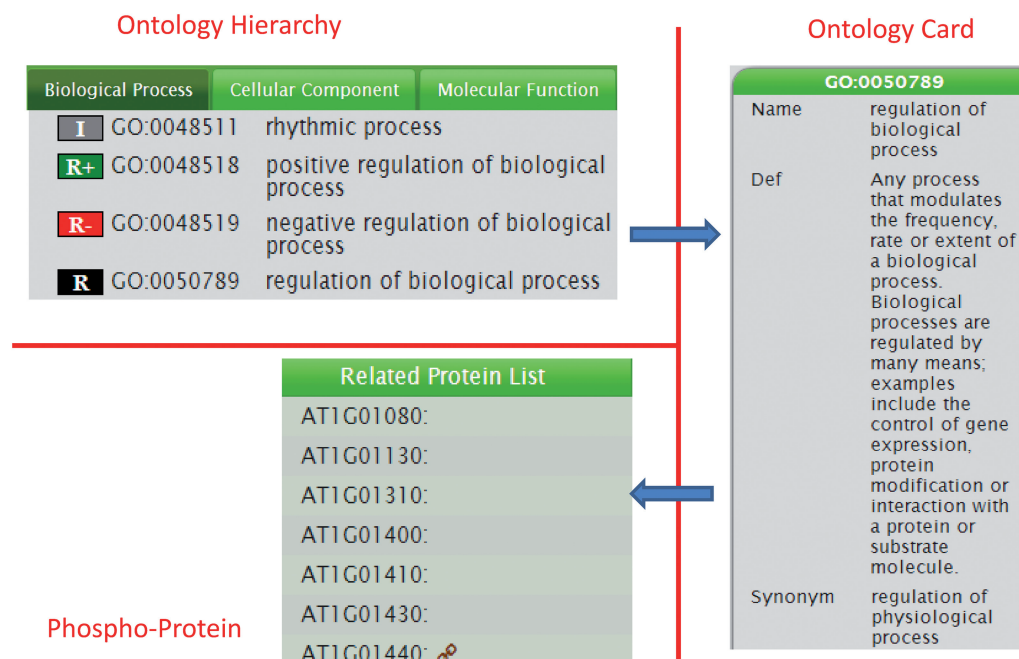


Figure 2. Ontology browser.

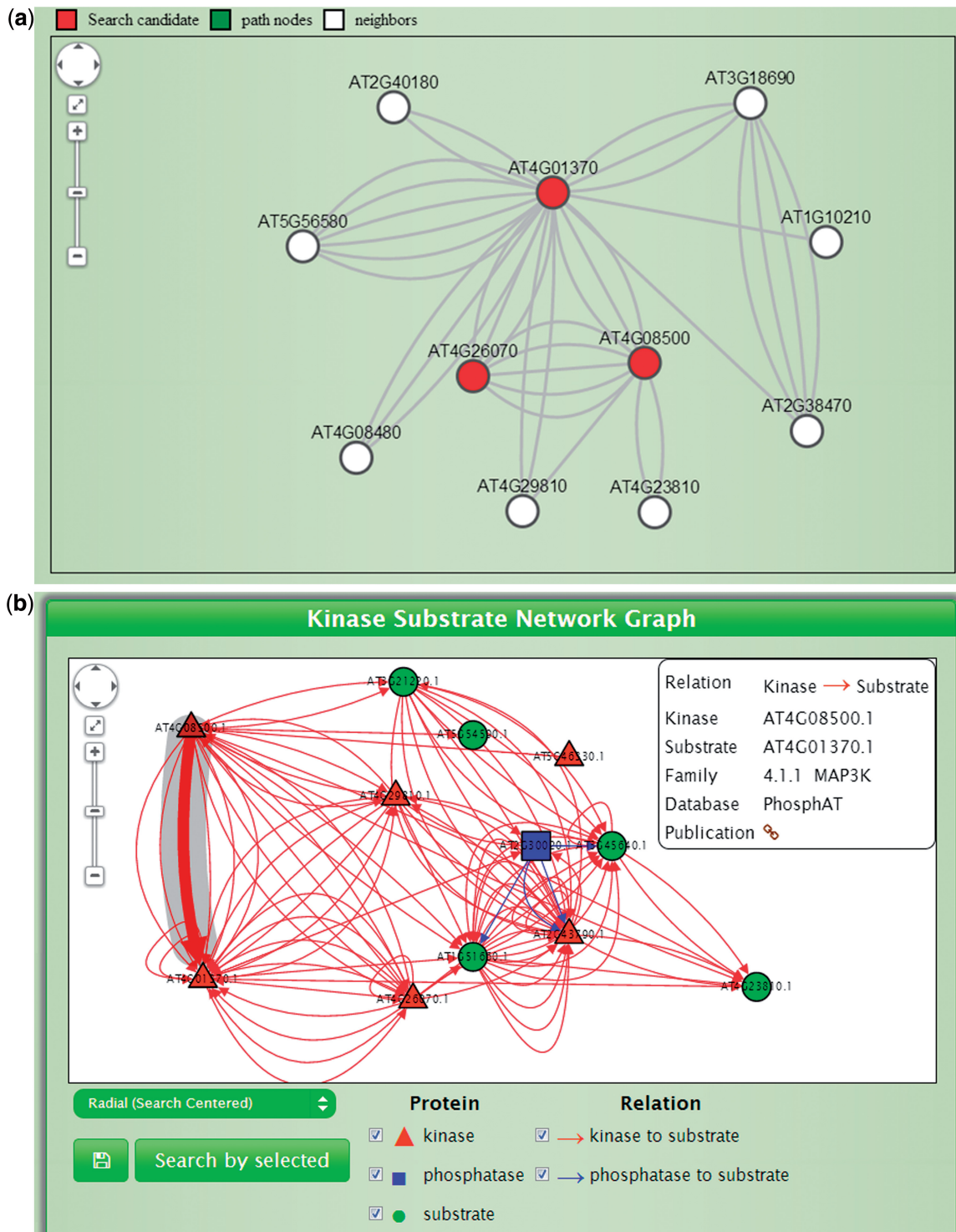


Figure 3. Cross-reference between PPI network and kinase/phosphatase networks. (a) PPI network; (b) kinase/phosphatase–substrate network.

its neighboring interactions by simply clicking the node. The white nodes in Figure 3a represent proteins that are directly interacting with the three red proteins.

Kinase and phosphatase-substrate network

An important data source for the kinase-substrate network is the emerging results from protein kinase and phosphatase

client screens. For example, P³DB displays KiC assay results separately from other enzyme target data, since they may provide further details of phosphorylation sites. Meanwhile, the KiC assay results are also merged into the pool of the kinase-substrate network.

The kinase-substrate network is overlaid with the phosphatase-substrate network. Different colors of the nodes

and edges are used to distinguish kinases and phosphatases, and phosphorylation and dephosphorylation. This network can provide complementary information to PPI networks. For example, in Figure 3b, the network obtained by searching from AT4G08500.1 and AT2G30020.1 contains kinases, a phosphatase and substrates with different colors and annotations. As a PPI partner seen in the previous example, AT4G01370.1 is also involved in this kinase-substrate network, which potentially reveals the MAP kinase cascade (38–41). Interestingly, AT2G30020.1 shows phosphatase activities in the network (42), which adds more information to what the PPI network can provide.

Domain network

Domain co-occurrence networks can help reveal kinase-domain interactions, regulatory domains and the recognition domains in phosphorylation-signaling pathways. For example, WW is a recognition domain for those phosphorylated proteins containing the pSer/pThr-Pro motif through the local conformational change of proline isomerization (43). In the WW neighboring network, the RNA capping methyltransferase domains, the RNA-binding domains and the helicase domains can be found. This may indicate that the related phosphorylation events for WW recognized proteins may eventually change the protein expression by affecting the mRNA metabolism, splicing, ribosome assembly and translation initiation (44). In the domain co-occurrence network, the protein that contributes to the domain network is cross-listed. The thickness of the edge between domains indicates the number of proteins that contribute to the domain linkage (Figure 4). The domain network

represents domain structures from different species with different colors, so that the conservation of the protein domains and functions can be viewed easily.

Community-based user experience

Automated data curating system

Users can now upload data on their own, and P³DB will automatically generate a customized data repository site for publication purposes. Users can also delete their own datasets easily if they wish (Figure 5).

Data sharing and security control

Users can decide the access level of their own data by selecting public, private or shared within a group. The public user data can be merged to the main depository pool for general P³DB display if the data meet the quality requirement.

Annotation by comments

The protein data, phosphosite data and phosphopeptide data can be annotated and commented by users. Users can also reply or follow other users' comments.

CONCLUSION

P³DB is a comprehensive, systematic and interactive plant protein phosphorylation data resource. It helps researchers to analyze protein phosphorylation events across the plant kingdom, providing homology-based evidence for function. P³DB 3.0 provides several network-based data representation and visualization tools to view the functions and context of phosphorylation sites in multiple aspects, by integrating all the related

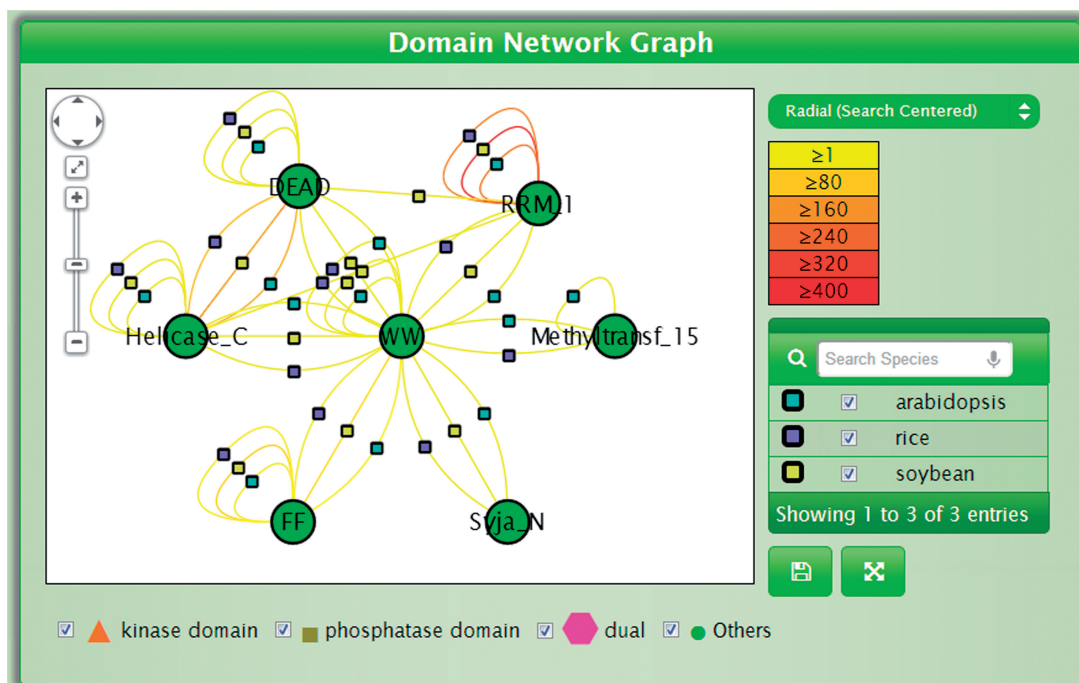


Figure 4. Domain co-occurrence network.

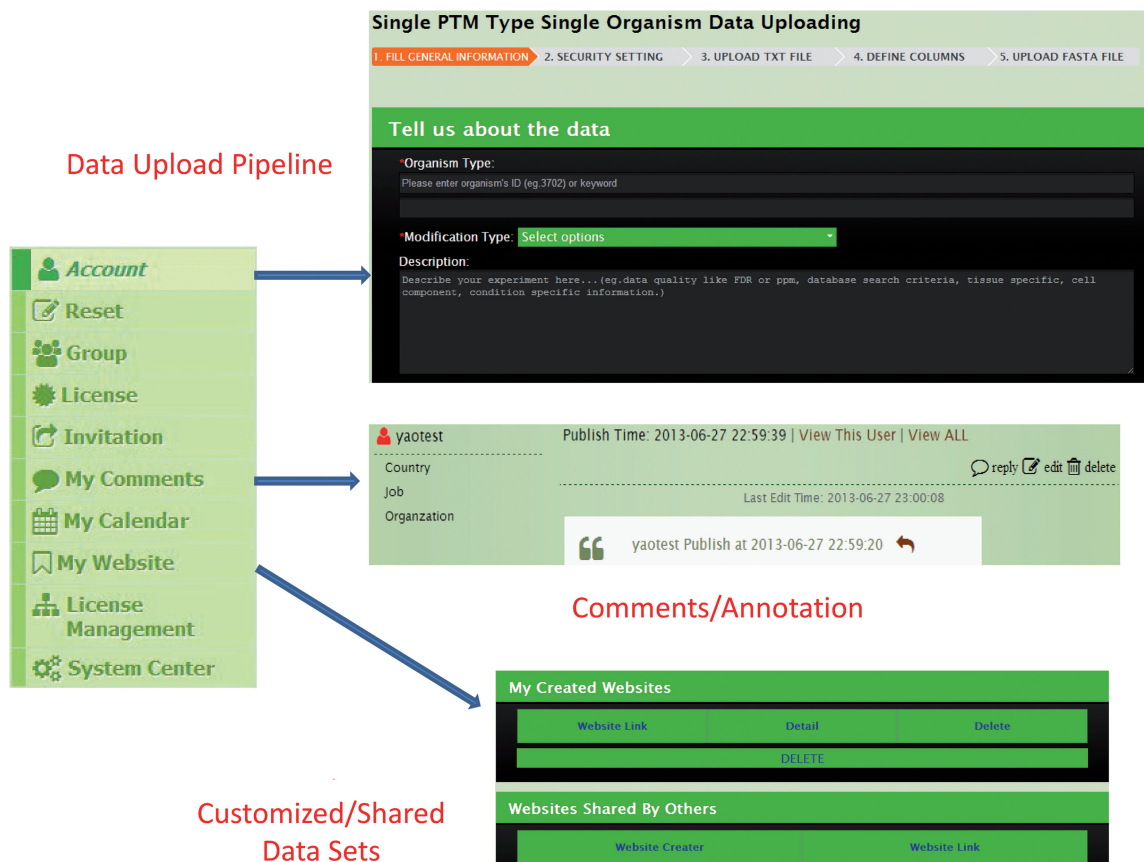


Figure 5. Community-based data curation.

information. The community-based design allows users to have better communication and control of their data in P³DB.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to acknowledge P³DB users for providing the suggestions and reporting some missing data from the literature.

FUNDING

The National Science Foundation [DBI-0604439 and 1126992] and National Institutes of Health [R01-GM100701]. Funding for open access charge: National Institutes of Health [R01-GM100701].

Conflict of interest statement. None declared.

REFERENCES

- Huang, P.H. and White, F.M. (2008) Phosphoproteomics: unraveling the signaling web. *Mol. Cell*, **31**, 777–781.
- Gao, J., Agrawal, G.K., Thelen, J.J. and Xu, D. (2009) P3DB: a plant protein phosphorylation database. *Nucleic Acids Res.*, **37**, D960–D962.
- Cantin, G.T., Yi, W., Lu, B., Park, S.K., Xu, T., Lee, J.D. and Yates, J.R. 3rd (2008) Combining protein-based IMAC, peptide-based IMAC, and MudPIT for efficient phosphoproteomic analysis. *J. Proteome Res.*, **7**, 1346–1351.
- Zhang, X.C., Wang, Z., Zhang, X., Le, M.H., Sun, J., Xu, D., Cheng, J. and Stacey, G. (2012) Evolutionary dynamics of protein domain architecture in plants. *BMC Evol. Biol.*, **12**, 6.
- Goel, R., Harsha, H.C., Pandey, A. and Prasad, T.S. (2012) Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. *Mol. bioSyst.*, **8**, 453–463.
- Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A. *et al.* (2009) Human Protein Reference Database–2009 update. *Nucleic Acids Res.*, **37**, D767–D772.
- Hornbeck, P.V., Kornhauser, J.M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V. and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.*, **40**, D261–D270.
- Dinkel, H., Chica, C., Via, A., Gould, C.M., Jensen, L.J., Gibson, T.J. and Diella, F. (2011) Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Res.*, **39**, D261–D267.
- Zulawski, M., Braginets, R. and Schulze, W.X. (2013) PhosPhAt goes kinases—searchable protein kinase target information in the plant phosphorylation site database PhosPhAt. *Nucleic Acids Res.*, **41**, D1176–D1184.
- Joshi, H.J., Hirsch-Hoffmann, M., Baerenfaller, K., Gruissem, W., Baginsky, S., Schmidt, R., Schulze, W.X., Sun, Q., van Wijk, K.J., Egelhofer, V. *et al.* (2011) MASCP Gator: an aggregation portal

- for the visualization of Arabidopsis proteomics data. *Plant Physiol.*, **155**, 259–270.
11. Huang, Y. and Thelen, J.J. (2012) KiC assay: a quantitative mass spectrometry-based approach for kinase client screening and activity analysis [corrected]. *Methods Mol. Biol.*, **893**, 359–370.
 12. Ahsan, N., Huang, Y., Tovar-Mendez, A., Swatek, K.N., Zhang, J., Miernyk, J.A., Xu, D. and Thelen, J.J. (2013) A versatile mass spectrometry-based method to both identify kinase client-relationships and characterize signaling network topology. *J. Proteome Res.*, **12**, 937–948.
 13. Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L. and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, **27**, 431–432.
 14. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. et al. (2000) Gene ontology: tool for the unification of biology. *Nat. Genet.*, **25**, 25–29.
 15. The Gene Ontology Consortium. (2013) Gene Ontology annotations and resources. *Nucleic Acids Res.*, **41**, D530–D535.
 16. Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M. et al. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.*, **40**, D1202–D1210.
 17. Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. et al. (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, **40**, D1178–D1186.
 18. Gribskov, M., Fana, F., Harper, J., Hope, D.A., Harmon, A.C., Smith, D.W., Tax, F.E. and Zhang, G. (2001) PlantsP: a functional genomics database for plant phosphorylation. *Nucleic Acids Res.*, **29**, 111–113.
 19. Rose, P.W., Beran, B., Bi, C., Bluhm, W.F., Dimitropoulos, D., Goodsell, D.S., Prlic, A., Quesada, M., Quinn, G.B., Westbrook, J.D. et al. (2011) The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res.*, **39**, D392–D401.
 20. Magrane, M. and Consortium, U. (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database J. Biol. Databases Curat.*, **2011**, bar009.
 21. (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.*, **41**, D43–D47.
 22. Mosca, R., Ceol, A. and Aloy, P. (2013) Interactome3D: adding structural details to protein networks. *Nat. Methods*, **10**, 47–53.
 23. Herraiz, A. (2006) Biomolecules in the computer: Jmol to the rescue. *Biochem. Mol. Biol. Educat. Bimonthly Publicat. Int. Union Biochem. Mol. Biol.*, **34**, 255–261.
 24. Stark, C., Breitkreutz, B.J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M.S., Nixon, J., Van Auken, K., Wang, X., Shi, X. et al. (2011) The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.*, **39**, D698–D704.
 25. Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., Duesbury, M., Dumousseau, M., Feuermann, M., Hinz, U. et al. (2012) The IntAct molecular interaction database in 2012. *Nucleic Acids Res.*, **40**, D841–D846.
 26. Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U. and Eisenberg, D. (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
 27. Licata, L., Briganti, L., Peluso, D., Perfetto, L., Iannuccelli, M., Galeota, E., Sacco, F., Palma, A., Nardoza, A.P., Santonico, E. et al. (2012) MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.*, **40**, D857–D861.
 28. Punta, M., Coghill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J. et al. (2012) The Pfam protein families database. *Nucleic Acids Res.*, **40**, D290–D301.
 29. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P. and Dunker, A.K. (2005) Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins*, **61(Suppl. 7)**, 176–182.
 30. Yao, Q., Bollinger, C., Gao, J., Xu, D. and Thelen, J.J. (2012) P(3)DB: an integrated database for plant protein phosphorylation. *Front. Plant Sci.*, **3**, 206.
 31. Yao, Q., Gao, J., Bollinger, C., Thelen, J.J. and Xu, D. (2012) Predicting and analyzing protein phosphorylation sites in plants using musite. *Front. Plant Sci.*, **3**, 186.
 32. Gao, J. and Xu, D. (2012) Correlation between posttranslational modification and intrinsic disorder in protein. *Pac. Symp. Biocomput.*, **94**, 103.
 33. Li, L., Stoeckert, C.J. Jr and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.*, **13**, 2178–2189.
 34. Zeng, Q., Chen, J.G. and Ellis, B.E. (2011) AtMPK4 is required for male-specific meiotic cytokinesis in Arabidopsis. *Plant J. Cell Mol. Biol.*, **67**, 895–906.
 35. Ichimura, K., Mizoguchi, T., Irie, K., Morris, P., Giraudat, J., Matsumoto, K. and Shinozaki, K. (1998) Isolation of ATMEKK1 (a MAP kinase kinase kinase)-interacting proteins and analysis of a MAP kinase cascade in Arabidopsis. *Biochem. Biophys. Res. Commun.*, **253**, 532–543.
 36. Mizoguchi, T., Ichimura, K., Irie, K., Morris, P., Giraudat, J., Matsumoto, K. and Shinozaki, K. (1998) Identification of a possible MAP kinase cascade in Arabidopsis thaliana based on pairwise yeast two-hybrid analysis and functional complementation tests of yeast mutants. *FEBS Lett.*, **437**, 56–60.
 37. Gao, M., Liu, J., Bi, D., Zhang, Z., Cheng, F., Chen, S. and Zhang, Y. (2008) MEKK1, MKK1/MKK2 and MPK4 function together in a mitogen-activated protein kinase cascade to regulate innate immunity in plants. *Cell Res.*, **18**, 1190–1198.
 38. Ichimura, K., Casais, C., Peck, S.C., Shinozaki, K. and Shirasu, K. (2006) MEKK1 is required for MPK4 activation and regulates tissue-specific and temperature-dependent cell death in Arabidopsis. *J. Biol. Chem.*, **281**, 36969–36976.
 39. Nakagami, H., Soukupova, H., Schikora, A., Zarsky, V. and Hirt, H. (2006) A Mitogen-activated protein kinase kinase kinase mediates reactive oxygen species homeostasis in Arabidopsis. *J. Biol. Chem.*, **281**, 38697–38704.
 40. Suarez-Rodriguez, M.C., Adams-Phillips, L., Liu, Y., Wang, H., Su, S.H., Jester, P.J., Zhang, S., Bent, A.F. and Krysan, P.J. (2007) MEKK1 is required for flg22-induced MPK4 activation in Arabidopsis plants. *Plant Physiol.*, **143**, 661–669.
 41. Berriri, S., Garcia, A.V., Dit Frey, N.F., Rozhon, W., Pateyron, S., Leonhardt, N., Montillet, J.L., Leung, J., Hirt, H. and Colcombet, J. (2012) Constitutively active mitogen-activated protein kinase versions reveal functions of Arabidopsis MPK4 in pathogen defense signaling. *Plant Cell*, **24**, 4281–4293.
 42. Schweighofer, A., Kazanaviciute, V., Scheikl, E., Teige, M., Doczi, R., Hirt, H., Schwanninger, M., Kant, M., Schuurink, R., Mauch, F. et al. (2007) The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in Arabidopsis. *Plant Cell*, **19**, 2213–2224.
 43. Yaffe, M.B. and Elia, A.E. (2001) Phosphoserine/threonine-binding domains. *Curr. Opin. Cell Biol.*, **13**, 131–138.
 44. Aubourg, S., Kreis, M. and Lecharny, A. (1999) The DEAD box RNA helicase family in Arabidopsis thaliana. *Nucleic Acids Res.*, **27**, 628–636.