Database update

The human DEPhOsphorylation Database DEPOD: 2019 update

Nikhil P. Damle and Maja Köhn*

Signalling Research Centres BIOSS and CIBSS, Faculty of Biology, University of Freiburg, Schänzlestrasse 18, 79104 Freiburg, Germany

*Corresponding author: Tel.: +49 761 20367900; Fax: +49 761 20397419; Email: maja.koehn@bioss.uni-freiburg.de

Abstract

The human **Dep**h**o**sphorylation **D**atabase (DEPOD) is a manually curated resource that harbors human phosphatases, their protein and non-protein substrates, dephosphorylation sites and the associated signaling pathways. We report here an update to DEPOD by integrating and/or linking to annotations from 69 other open access databases including disease associations, phosphorylating kinases, protein interactions, and also genome browsers. We also provide tools to visualize protein interactions, protein structures, phosphorylation networks, evolutionary conservation of proteins, dephosphorylation sites, and short linear motifs within various proteins. The updated version of DEPOD contains 254 human phosphatases, 336 protein and 83 non-protein substrates, and 1215 manually curated phosphatase-substrate relationships. In addition, we have improved the data access as all the data in DEPOD can now be easily downloaded in a user-friendly format. With multiple significant improvements, DEPOD continues serving as a key resource for research on phosphatase-kinase networks.

Database URL: www.depod.org

Introduction

Phosphorylation is a reversible post-translational modification. Knowledge about protein kinases and phosphatases is critical to understand the phosphoregulation of a signaling event comprehensively. Advances in experimental and computational approaches led to the development of several kinase-centric resources and tools, whereas phosphatase-centric resources are very few. The human **DEP**h**O**sphorylation **D**atabase (DEPOD) was the first resource that curated information from literature about the phosphatase genes in humans along with their protein and non-protein substrates, the dephosphorylation sites, the protein interactions involving them and the signaling pathways regulated by them (1,2). Currently available other phosphatase-related resources either do not provide information about their substrates (PTP-central (3), Phosphatome. Net (4), EKPD (5) and iEKPD (6)) or provide information only about the protein substrates (HuPho (7)). Additionally, many of them rely on DEPOD for information on the phosphatase genes in humans and their substrates,

and have not been updated in years, and therefore do not reflect the current knowledge on phosphatases and their substrates.

We have now updated and also improved DEPOD significantly on the following three fronts: A. compilation, by adding new phosphatases and substrates; B. annotations, by integrating annotations from 69 other open access databases, and including two new modules on disease associations and protein kinases that are experimentally known to phosphorylate a given protein in DEPOD; and C. data access and visualization, by better organizing and interlinking the information and incorporating new visualization tools to facilitate data navigation.

DEPOD is now a unique resource that gives information about the bidirectional regulation of a phosphorylation event. New hypotheses may now be formulated based on this information and also based on the evolutionary conservation of proteins, the dephosphorylation sites and short linear motifs (SLiMs) within them. Disease associations in DEPOD may be of great value to increasing efforts to target phosphatases therapeutically (8). Overall, DEPOD will be a very valuable resource for the phosphorylation signaling community and also cater to the needs of wider research interests.

Organization of DEPOD

Figure 1 shows the overall organization of DEPOD consisting of the following three modules: A. compilation, B. annotations and C. access and visualization of data. Supplementary Figure S1 depicts the entity relationship diagram showing its architecture as a relational database.

A. Compilation of phosphatase genes and their substrates

Phosphatase genes in humans were fetched from Ensembl (GRCh38p12) using a keyword-based search. Candidates were filtered using 186 phosphatase-related gene ontology (GO) terms (Supplementary Table S1) to filter out non-phosphatase genes from the list. Remaining genes were compared with previous versions (1,2) of DEPOD and also with two other recently developed phosphatase-related resources—Phosphatome. Net (4) and iEKPD (6) to finally compile a list of 254 phosphatase genes in humans. These were then reclassified into superfamilies based on their CATH fold (9) and subsequently into families based on the sequences of their catalytic domains (Supplementary Table S2). This classification was carried out as reported previously (1, 2). Briefly, the catalytic domains of the phosphatases were fetched from their Uniprot FASTA sequence using the Pfam domain definitions

such as Y_phosphatase for Tyr-phosphatases and metallophosphatase domain for Ser/Thr-specific phosphatases. In cases where more than one phosphatase domain appeared in a single polypeptide, the domain having the catalytic motif signature as defined by Alonso *et al.* (10) was used for classification.

Unlike in DEPODv1.1, we have now renamed the superfamilies and families so that the nomenclature sufficiently reflects the substrate specificities of phosphatases and relates to the historic classes. For example, 'family 1' in DEPODv1.1 is now called 'Class I Cys-based PTPs' in accordance with previous literature (4, 11). Phylogenetic trees constructed using neighbor-joining algorithm in MEGA suite (12) and the catalytic domains of the members belonging to each of these superfamilies can be easily accessed using the 'view tree' links on the 'structure-based family' tab. The tree visualizations have been created using interactive tree of life (iTOL) (13). We now also provide in the download section the comparison between different databases and classifications, to provide clarity to this ongoing research. Of note, DEPOD contains more phosphatases than other recently developed phosphatase-related resources (4, 6), largely due to the inclusion of non-protein phosphatases and manual curation of the entries.

Fourteen different protein interaction databases were queried via PSICQUIC service (14) using the 'dephosphorylation' keyword. In-house scripts were developed to merge the outputs and create a list of unique dephosphorylation interactions. Interactions where no literature evidence was available (PubMed ID unassigned) were filtered out. Interactions where only fragments of the proteins (such as a peptide stretch or only a single domain in a multidomain protein) were used during dephosphorylation assay were also excluded. Remaining literature evidences published after January 2015 (DEPODv1.1) or not inspected during the development of DEPODv1.1 were inspected manually to compile a list of protein and non-protein substrates of the aforementioned phosphatase genes. A protein is regarded as a substrate only if the dephosphorylation of the full-length protein has been convincingly established using either or both of the *in vitro* and *in vivo* (in human cell lines) assays by single or multiple research groups. A complete list of PubMed IDs processed during this compilation along with the corresponding decisions and the underlying rationale can be found in Supplementary Table S3. Taken together, DEPOD now holds information about 336 protein substrates and 83 non-protein substrates of human phosphatases representing 1215 dephosphorylation interactions. The reliability scores of these phosphatase–substrate interactions were assigned as described in DEPODv1.1.
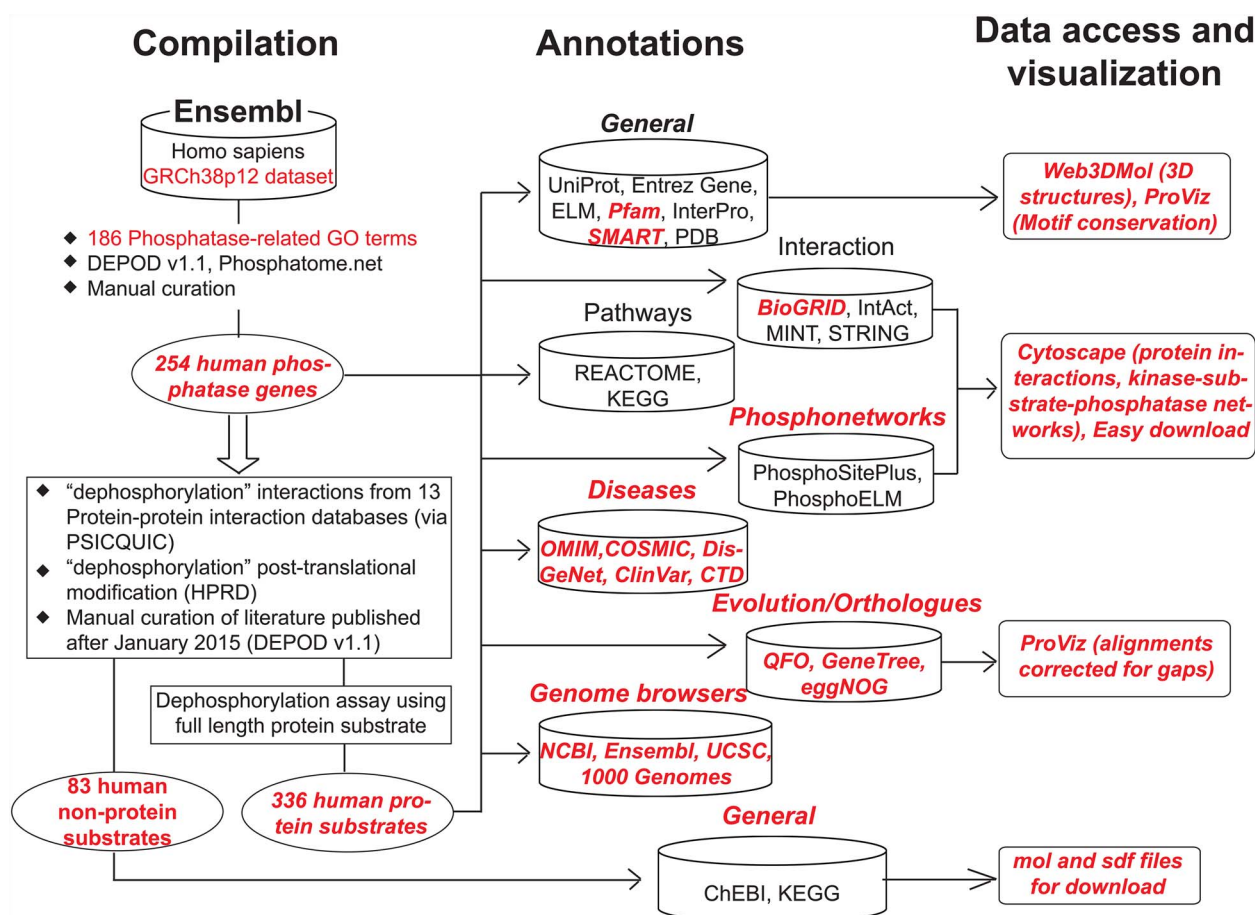
Figure 1. Schematic organization of DEPOD. DEPOD is organized in the following three modules: compilation, annotations and data access and visualization. Details on each are described in text. Red font indicates new additions or improvements in the current version over earlier versions.

B. New annotations in DEPOD

Figure 2 and Supplementary Figure S2 show a typical protein and a non-protein entry in DEPOD, respectively. DEPOD now provides several new and useful annotations for phosphatases and their substrates.

Disease associations. We have incorporated diseases associated with the phosphatases and their substrates from four different databases, namely COSMIC (15), DisGeNet (16,17), OMIM (https://omim.org/) and ClinVar (18) along with the corresponding reliability scores assigned by the parent databases. ('Domain organization, Expression, Diseases' subsection in the 'Basic Information' tab of each entry). The link to the PhosphoSitePlus database provided in this section may also be used to gain disease information on the gene/protein of interest and on involved or mutated phosphorylated sites.

Domain organization. Under the same section, we also provide links to the corresponding entry pages in Pfam (19), SMART (20) and InterPro (21) databases. For a quick overview, we now provide a visual schematic domain organization of a polypeptide with links to the corresponding domain families in Pfam.

Genomic annotations. Links to four genomic browsers (NCBI, UCSC, 1000Genomes and Ensembl) have been provided to explore genomic features of an entity. Comparative **T**oxicogenomics **D**atabase (CTD) (22) has also been linked to give a glimpse of the chemical–gene interactions and other manually curated information provided there.

Crosslinks. Each protein entry in DEPOD has now been linked with 69 other publicly accessible databases by creating a mapping between DEPOD entry names and other identifiers used in the corresponding databases. Some of these identifiers can be accessed using the links provided in DEPOD entries including the 'Query our ID-mapping table' link while others are provided as an excel sheet in the 'Download' section.

Catalytic signature motifs. For protein tyrosine phosphatases, DEPOD now provides their signature catalytic motifs incorporated from Alonso *et al*. (10).
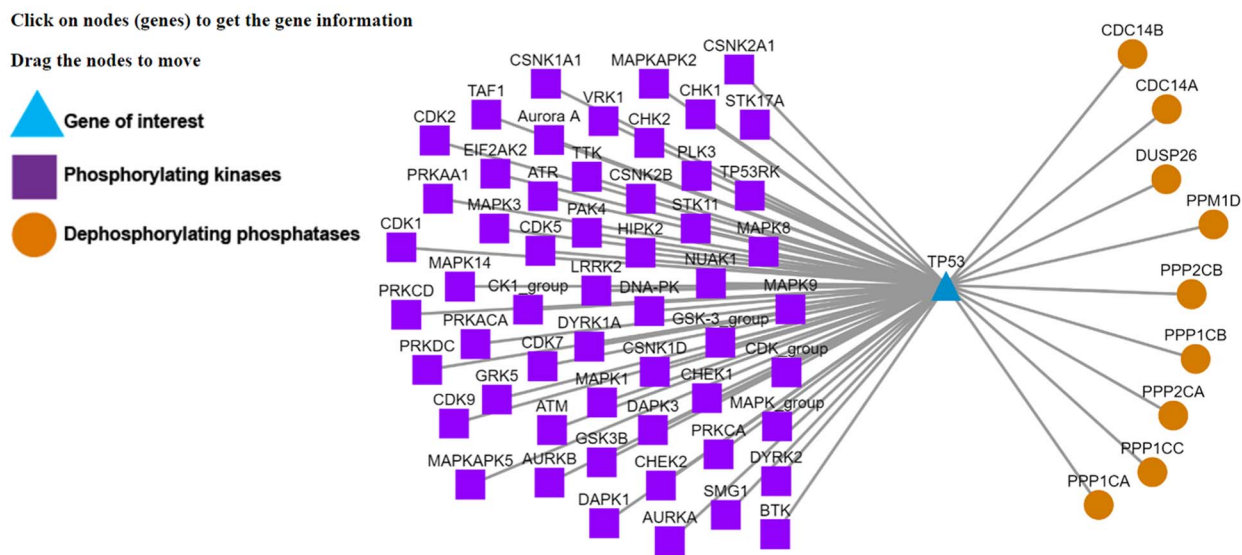
Figure 2. A typical protein entry in DEPOD. Various annotation fields are depicted and newly added annotations are highlighted separately. Phosphatase entries have a 'Substrates' tab that harbors information about its protein and non-protein substrates. Substrate entries have a 'Phosphatases' tab that includes information about phosphatases that dephosphorylate them.

Protein interactions. DEPOD provides protein interacting partners of phosphatases and their protein substrates under the 'Interacting Proteins' tab along with annotations about the detection methods, interaction types, interaction scores and links to original literature evidences supporting the interactions from BioGRID (23), IntAct (24) and MINT (25). The reliability scores for the protein interactions and the source publications have been directly imported from the corresponding databases.

Phosphorylating kinases. We have added a new module on protein kinases that are experimentally demonstrated to phosphorylate various protein entries in DEPOD by integrating information from PhosphoSitePlus (26), Phospho-ELM (27) and HPRD (28). Each phosphorylation site and the five amino acids on its either sides are annotated with the phosphorylation detection method and links to original supporting publications. In addition, evolutionary conservation of each site across metazoan and other clades may be visualized by clicking on each of the sites.

Evolutionary conservation. DEPOD provides information about the evolutionary conservation of entire proteins, the SLiMs and also the phospho and dephosphorylation sites within proteins. This information enables users to formulate hypotheses about the functional importance of a particular motif or site depending on whether or not it is conserved.

Annotations for non-protein substrates. Non-protein substrates of phosphatases is a unique feature of DEPOD. Substantially improving this area, non-protein substrates have now been annotated with their molecular formulae, SMILES and InChI notations and InChI keys. They have also been linked to KEGG (29), ChEBI (30), ChEMBL (31), PDB-CCD (32) and PubChem (33). Information about the 69 external databases may be found under the 'About DEPOD' section on the left-hand side panel.

C. Data access and visualization

DEPOD can be queried, accessed and navigated in a number of different ways. All protein entries in DEPOD can be queried using UniProt accessions, gene and protein names, NCBI gene identifiers and GenBank identifiers. Non-protein substrates may be queried using ChEBI or KEGG identifiers, compound names, molecular formulae and SMILES notations. Other query mechanisms such as quicksearch utility, KEGG or REACTOME pathway-based queries, dephosphorylation site-based queries and sequence-based queries have been retained from DEPODv1.1.

Data in DEPOD can now be easily accessed via various links provided with each entry. Users can download FASTA formatted sequences of protein substrates of phosphatases under the 'Protein substrates' sub-tab under the 'Substrates'

Figure 3. Kinase-substrate-phosphatase relationships (PSKRs) of TP53 in DEPOD. DEPOD imports the information about the protein kinases (purple squares) experimentally demonstrated to phosphorylate a gene (TP53 shown as blue triangle in this case) from PhosphoSitePlus, PhosphoELM and HPRD. The information about phosphatases (orange spheres) that dephosphorylate this gene is manually curated in DEPOD. PSKRs in DEPOD inform about bidirectional phosphoregulation of a gene of interest. At present, DEPOD distinguishes dephosphorylation of a substrate by different isoforms of a phosphatase only in cases where literature evidence specifically exists for that isoform. In the absence of any such evidence, all isoforms of that phosphatase are considered to dephosphorylate a given substrate (here for example PPP2CA and PPP2CB). The figure was created using Cytoscape (45).

tab in each phosphatase entry. The 'Interacting Proteins' tab provides a link to download excel sheets containing the protein interacting partners of protein entries in DEPOD from the three protein interaction databases. A non-redundant list created by merging all protein interacting partners together can also be downloaded under the 'All' sub-tab. The 'Basic Information' tab for non-protein entries provides links to download structures in .*mol* and .*sdf* formats.

The four major ways to navigate DEPOD have been retained from v1.1. Human phosphatases can be browsed using a link on the homepage. CATH (9) fold-based structural classification implemented in DEPOD is provided under the 'Structure-based family' tab. We renamed the phosphatase superfamilies to reflect as closely as possible their substrate preferences and to relate to their historical classification. Links associated with each superfamily, family, EC number and the historic class lead to a table describing the members belonging to that group along with their SCOP (34), CATH (9) and EC identifiers (35), DEPOD classification and the catalytic activity status. In addition, FASTA formatted polypeptide sequences of these members can be downloaded using a link provided with the table. We have now also added phylogenetic trees for each superfamily of phosphatases with color codes representing the member families within each of them. Pathways regulated by phosphatases or involving its substrates are listed under the 'Pathway' tab. Other phosphatases and substrates also

involved in the same pathway can be accessed by clicking on the corresponding pathway identifiers. DEPOD now stores information about 299 KEGG pathways (29) and 876 REACTOME pathways (36) along with the phosphatases and their substrates involved in them.

Under the 'Basic Information' tab, users can select and load PDB identifiers of their choice and interactively visualize the corresponding structure using Web3DMol (37). A link on how to use this tool has also been provided for an informed usage. Visualization of evolutionary conservation of the entire protein and the SLiMs and phospho and dephosphorylation sites within proteins is now possible through ProViz tool (38) using simple clickable links. We enabled interactive visualization of protein interaction networks and kinase-substrate-phosphatase relations (Supplementary Figure S3) by incorporating cytoscpe.js libraries (39) with each node linked to its NCBI gene annotations.

Exploring DEPOD

We now explain the use of some of the new and unique features of the database with TP53 as example, which users can search using the gene symbol or protein name under the 'Search Substrates' tab. Lack of phosphorylation at Thr55 of TP53 leads to its nuclear localization and eventually DNA damage-induced cell death (40). Manually curated information about the dephosphorylating phosphatases in DEPOD (PP2A—in the 'Phosphatases' tab

of the TP53 entry) along with the information about the phosphorylating kinases integrated from PhosphoSitePlus and Phopsho.ELM (GRK5, MAPK1 and TAF1—'Phosphorylating kinases' tab) and the neighboring sequence context (integrated from UniProt) will facilitate interpretations about phosphoregulation of this site without the need to visit different databases individually. A mouse click on this site will reveal the conservation of this entire stretch across different model organisms. This information may for example be used to design amino acid substitutions in a certain sequence neighborhood. Clicking on the phosphorylation network visualization link ('Phosphorylation Network' section on the 'Basic Information' tab or a separate link under the 'Phosphorylating Kinases' tab) will also provide information about other kinases and phosphatases that regulate TP53 (Figure 3). This information is valuable to understand alternate regulatory mechanisms.

Usage and Licensing

All data in DEPOD remains freely accessible and usable for academic and non-profitable usage and is provided under the 'Download' section on the homepage. Licensing information for non-academic and for-profit users should be inquired directly under depod@bioss.uni-freiburg.de. More information on how to use and navigate DEPOD along with a brief description of different fields in DEPOD entries can be found under the 'User Manual' and 'About DEPOD' links in the left-hand navigation panel.

Concluding Remarks

DEPOD has been used in many studies to formulate hypotheses or as a key resource for phosphatase substrates and networks (e.g. (2, 41, 42). With several significant and useful additions, DEPOD continues to serve as a curated and comprehensive resource for human phosphatases and their substrates. It is unique in harboring both the protein and the non-protein substrates of human phosphatases and similarly both active and inactive human phosphatases. Through annotations on protein kinases that phosphorylate human phosphatases and their substrates, DEPOD informs users about the bidirectional phosphoregulation of a signaling event. This knowledge may help in understanding individual signaling events comprehensively and also the signaling networks in general as specific targeted experiments can be designed based on it. This knowledge may also be useful in clinical studies. Phosphatases are widely pursued as potential drug targets (8). Owing to the conserved active sites within one phosphatase family, inhibition of a particular phosphatase may have several undesired effects (43,44). DEPOD now provides knowledge

about the diseases associated with human phosphatases and their substrates, which might help in proposing hypotheses about other potential signaling pathways that might also be inhibited by the drug candidates. This is key piece of information in designing selective inhibitors of phosphatases (44). Information about the extent of conservation of a particular motif or a dephosphorylation site provided by DEPOD may be useful in investigating its functional role in cellular signaling and may also be helpful in designing activity modulators of phosphatases, widely used tools to investigate phosphatase-signaling networks. In addition to encouraging the development of other phosphatase-centric resources, DEPOD also highlights the need to identify more substrates of phosphatases in humans.

Supplementary data

Supplementary data are available at *Database* Online.

References

1. Duan,G., Li,X. and Köhn,M. (2015) The human DEPhOsphorylation database DEPOD: a 2015 update. *Nucleic Acids Res.*, **43**, D531–D535.
2. Li,X., Wilmanns,M., Thornton,J. and Köhn,M. (2013) Elucidating human phosphatase-substrate networks. *Sci. Signal.*, **6**, rs10.
3. Hatzihristidis,T., Liu,S., Pryszcz,L. *et al.* (2014) PTP-central: a comprehensive resource of protein tyrosine phosphatases in eukaryotic genomes. *Methods*, **65**, 156–164.
4. Chen, M. J., Dixon, J. E., and Manning, G. (2017) Genomics and evolution of protein phosphatases. *Sci. Signal.* **10**, pii: eaag1796.
5. Wang,Y., Liu,Z., Cheng,H. *et al.* (2014) EKPD: a hierarchical database of eukaryotic protein kinases and protein phosphatases. *Nucleic Acids Res.*, **42**, D496–D502.
6. Guo,Y., Peng,D., Zhou,J. *et al.* (2019) iEKPD 2.0: an update with rich annotations for eukaryotic protein kinases, protein phosphatases and proteins containing phosphoprotein-binding domains. *Nucleic Acids Res.*, **47**, D344–D350.
7. Liberti,S., Sacco,F., Calderone,A. *et al.* (2013) HuPho: the human phosphatase portal. *FEBS J.*, **280**, 379–387.

8. Mullard,A. (2018) Phosphatases start shedding their stigma of undruggability. *Nat. Rev. Drug Discov.*, **17**, 847–849.

9. Dawson,N.L., Lewis,T.E., Das,S. *et al.* (2017) CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Res.*, **45**, D289–D295.

10. Alonso,A. and Pulido,R. (2016) The extended human PTPome: a growing tyrosine phosphatase family. *FEBS J.*, **283**, 1404–1429.

11. Alonso,A., Sasin,J., Bottini,N. *et al.* (2004) Protein tyrosine phosphatases in the human genome. *Cell*, **17**, 699–711.

12. Kumar,S., Stecher,G., Li,M. *et al.* (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, **35**, 1547–1549.

13. Letunic,I. and Bork,P. (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.*, **47**, W256–W259.

14. Aranda,B., Blankenburg,H., Kerrien,S. *et al.* (2011) PSICQUIC and PSICORE: accessing and scoring molecular interactions. *Nat. Methods*, **8**, 528–529.

15. Tate,J.G., Bamford,S., Jubb,H.C. *et al.* (2019) COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.*, **47**, D941–D947.

16. Pinero,J., Bravo,A., Queralt-Rosinach,N. *et al.* (2017) DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.*, **45**, D833–D839.

17. Pinero,J., Queralt-Rosinach,N., Bravo,A. *et al.* (2015) DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database (Oxford)*, bav028.

18. Landrum,M.J., Lee,J.M., Benson,M. *et al.* (2018) ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.*, **46**, D1062–D1067.

19. El-Gebali,S., Mistry,J., Bateman,A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.*, **47**, D427–D432.

20. Letunic,I. and Bork,P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.*, **46**, D493–D496.

21. Mitchell,A.L., Attwood,T.K., Babbitt,P.C. *et al.* (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.*, **47**, D351–D360.

22. Davis,A.P., Grondin,C.J., Johnson,R.J. *et al.* (2019) The comparative Toxicogenomics database: update 2019. *Nucleic Acids Res.*, **47**, D948–D954.

23. Stark,C., Breitkreutz,B.J., Reguly,T. *et al.* (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.*, **34**, D535–D539.

24. Orchard,S., Ammari,M., Aranda,B. *et al.* (2014) The MIntAct project–IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.*, **42**, D358–D363.

25. Licata,L., Briganti,L., Peluso,D. *et al.* (2012) MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.*, **40**, D857–D861.

26. Hornbeck,P.V., Zhang,B., Murray,B. *et al.* (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.*, **43**, D512–D520.

27. Dinkel,H., Chica,C., Via,A. *et al.* (2011) Phospho.ELM: a database of phosphorylation sites–update 2011. *Nucleic Acids Res.*, **39**, D261–D267.

28. Keshava Prasad,T.S., Goel,R., Kandasamy,K. *et al.* (2009) Human protein reference database–2009 update. *Nucleic Acids Res.*, **37**, D767–D772.

29. Kanehisa,M., Sato,Y., Furumichi,M. *et al.* (2019) New approach for understanding genome variations in KEGG. *Nucleic Acids Res.*, **47**, D590–D595.

30. Hastings,J., Owen,G., Dekker,A. *et al.* (2016) ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res.*, **44**, D1214–D1219.

31. Gaulton,A., Hersey,A., Nowotka,M. *et al.* (2017) The ChEMBL database in 2017. *Nucleic Acids Res.*, **45**, D945–D954.

32. Burley,S.K., Berman,H.M., Bhikadiya,C. *et al.* (2019) RCSB protein data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.*, **47**, D464–D474.

33. Kim,S., Chen,J., Cheng,T. *et al.* (2019) PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.*, **47**, D1102–D1109.

34. Chandonia,J.M., Fox,N.K. and Brenner,S.E. (2019) SCOPe: classification of large macromolecular structures in the structural classification of proteins-extended database. *Nucleic Acids Res.*, **47**, D475–D481.

35. Bairoch,A. (2000) The ENZYME database in 2000. *Nucleic Acids Res.*, **28**, 304–305.

36. Fabregat,A., Jupe,S., Matthews,L. *et al.* (2018) The Reactome pathway knowledgebase. *Nucleic Acids Res.*, **46**, D649–D655.

37. Shi,M., Gao,J. and Zhang,M.Q. (2017) Web3DMol: interactive protein structure visualization based on WebGL. *Nucleic Acids Res.*, **45**, W523–W527.

38. Jehl,P., Manguy,J., Shields,D.C. *et al.* (2016) ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Res.*, **44**, W11–W15.

39. Franz,M., Lopes,C.T., Huck,G. *et al.* (2016) Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*, **32**, 309–311.

40. Cai,X. and Liu,X. (2008) Inhibition of Thr-55 phosphorylation restores p53 nuclear localization and sensitizes cancer cells to DNA damage. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 16958–16963.

41. Goldman,A., Roy,J., Bodenmiller,B. *et al.* (2014) The calcineurin signaling network evolves via conserved kinase-phosphatase modules that transcend substrate identity. *Mol. Cell*, **55**, 422–435.

42. Fearnley, G. W., Young, K. A., Edgar, J. R., *et al.* (2019) The homophilic receptor PTPRK selectively dephosphorylates multiple junctional regulators to promote cell-cell adhesion. *Elife* **8**, pii: e44597.

43. Roy, J., and Cyert, M. S. (2019) Identifying new substrates and functions for an old enzyme: calcineurin. *Cold Spring Harb. Perspect. Biol.*, pii: a035436.

44. De,S., Köhn,M. and Bollen,M. (2013) Challenges and opportunities in the development of protein phosphatase-directed therapeutics. *ACS Chem. Biol.*, **8**, 36–45.

45. Shannon,P., Markiel,A., Ozier,O. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.