

CoralP: Flexible visualization of the human phosphatome

Amit Min¹, Erika Deoudes¹, Marielle L. Bond¹, Eric S. Davis², Douglas H. Phanstiel^{1,2,3,4,5,6}

¹Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC 27599, USA

²Curriculum in Bioinformatics & Computational Biology, University of North Carolina, Chapel Hill, NC 27599, USA

³Curriculum in Genetics & Molecular Biology, University of North Carolina, Chapel Hill, NC 27599, USA

⁴Department of Cell Biology & Physiology, University of North Carolina, Chapel Hill, NC 27599, USA

⁵Lineberger Comprehensive Cancer Research Center, University of North Carolina, Chapel Hill, NC 27599, USA

⁶Correspondence: douglas_phanstiel@med.unc.edu

Protein phosphatases and kinases play critical roles in a host of biological processes and diseases via the removal and addition of phosphoryl groups. While kinases have been extensively studied for decades, recent findings regarding the specificity and activities of phosphatases have generated an increased interest in targeting phosphatases for pharmaceutical development. This increased focus has created a need for methods to visualize this important class of proteins within the context of the entire phosphatase protein family. Here, we present CoralP, an interactive web application for the generation of customizable, publication-quality representations of human phosphatome data. Phosphatase attributes can be encoded through edge colors, node colors, and node sizes. CoralP is the first and currently the only tool designed for phosphatome visualization and should be of great use to the signaling community. The source code and web application are available at <https://github.com/PhanstielLab/coralp> and <http://phanstiel-lab.med.unc.edu/coralp> respectively.

INTRODUCTION

Protein phosphorylation is an important post-translational modification that plays a major role in regulating protein activity and function. Misregulation of phosphorylation has been implicated in a host of human diseases including cancer (Ostman et al., 2006; Tonks, 2006), rheumatoid arthritis (Begovich et al., 2004; Hendriks et al., 2013), and various neurological disorders (Pulido and Hooft van Huijsduijnen, 2008; Robinson and Dixon, 2005). As such, altering protein phosphorylation levels has become a major focus of pharmaceutical development (Cohen, 2002). Protein phosphorylation is regulated primarily by protein kinases and protein phosphatases which add and remove phosphoryl groups respectively. While protein kinases have been the primary focus of research for the past several decades, recent findings regarding the activities and specificities of their counterparts have generated an increased interest in the role of phosphatases in human disease research (Andersen et al., 2004; Sacco et al., 2012; Tiganis and Bennett, 2007; Tonks, 2006). Because of this interest and the fact that there are over 180 human

phosphatases, there exists a great need for methods to analyze, interpret, and communicate experimental results within the context of the entire protein family (Chen et al., 2017).

While numerous methods have been developed to visualize the human kinome (Chartier et al., 2013; Eid et al., 2017; Metz et al., 2018), no such software exists for the human phosphatome. Existing methods to visualize kinases make use of a phylogenetic kinase tree that was constructed shortly after the sequencing of the human genome and later modified by Cell Signaling (www.cellsignal.com). These software packages allow for the encoding of multiple kinase attributes and data types through colors, sizes, and shapes laid out in a kinase tree organized by sequence and functional similarity. Last year, we developed Coral, an interactive web application that produces publication-quality visualizations of the human kinome (Metz et al., 2018). Coral allows for the simultaneous encoding of three attributes per kinase, supports three modes of visualization, and produces high-resolution vector images. Recently, Chen et al. published a phylogenetic phosphatase tree (Chen et al., 2017); however, to the best of our knowledge, there are currently no available tools to visualize phosphatase data within the context of this or any other format.

Here, we describe CoralP, a web application for the generation of customizable, publication-quality visualizations of the human phosphatome. Qualitative and quantitative features can be represented in branch colors, node colors, and node sizes. Phosphatases can be organized using the published phosphatome tree or as radial or force-directed networks. CoralP is simple to use, well documented, and freely available. It is the first and only dedicated tool for phosphatome visualization and is widely applicable to a variety of data types including those generated from proteomic, genomic, epidemiological, and high-throughput screening experiments.

RESULTS

CoralP is built using the same underlying framework as Coral and therefore makes use of similar strategies for data input, setting selection, and data download. Phosphatase attributes can be entered using pull-down

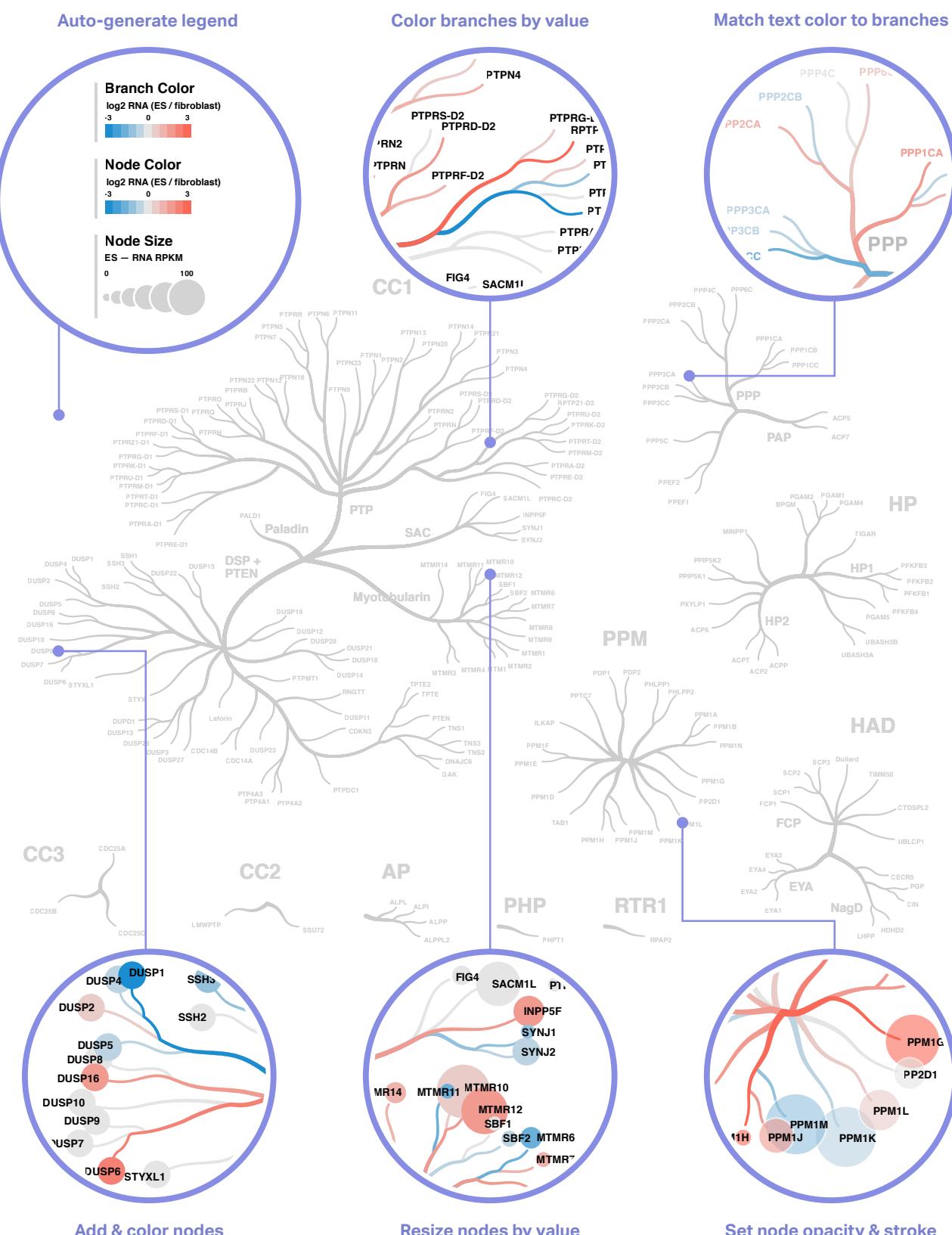


Figure 1. CoralP tree plot highlighting some of its features. The default view of a CoralP tree plot is shown in grey. Circular inserts reveal some of the formatting options. (Top left) Legends are automatically generated as users configure settings. (Top middle) Branch color can be used to encode quantitative or qualitative information. (Top right) Users can adjust the font type and color of phosphatase labels and even match the label color to the branch color. (Bottom left) Nodes can be added and recolored to depict quantitative or qualitative information. (Bottom middle) Node size can be used to encode quantitative information. (Bottom right) Users can control many aesthetic options including node opacity and stroke color.

menus or pasted into text boxes. CoralP supports multiple identifiers including UniProt, ENSEMBL, Entrez, and HGNC. Visualization options including color palettes, data scales, fonts, and protein identifiers can all be edited via an intuitive graphical user interface. CoralP was written using the reactive programming package Shiny (Chang et al., 2018), so CoralP plots update in real-time as users adjust settings, making it easy to generate a plot perfectly customized to the user's preferences.

CoralP offers highly flexible methods for visualizing qualitative and quantitative phosphatase attributes within the context of the entire protein family (Figure 1). Data can be encoded via multiple modalities and displayed via either tree or network views. Qualitative and quantitative phosphatase attributes can be encoded through three separate modalities: branch color, node color, and node size (for quantitative data only). Users can employ individual modalities or any combination thereof. The resulting data can be displayed either as a phosphatase tree (modified from Chen et al.) or as a network organized in a radial or force-directed fashion.

For ease of use, CoralP is highly documented and freely available. An 'Info' page describes every feature of CoralP and includes screenshots demonstrating how different settings impact the appearance of the final plots. For maximal reproducibility and stability, the code is open-source, version controlled, and publicly available at <https://github.com/PhanstielLab/coralp>. The web application is available at <http://phanstiel-lab.med.unc.edu/coralp>.

DISCUSSION

To the best of our knowledge, CoralP is the first and only dedicated tool for visualizing data within the context of the human phosphatome. It offers a rich selection of highly customizable features and produces high-resolution publication-quality figures in seconds. It is available online, employs a simple graphical user interface, and includes detailed documentation and examples. As such, CoralP is highly accessible to users, independent of operating system or computational expertise. Given its ease of use, the extensive adoption of the kinase tree to visualize kinase data, and the growing importance of phosphatases as an area of research and drug development, we anticipate that CoralP will be of great use to the signaling community. CoralP will expedite phosphatase research by enhancing our ability to interpret and communicate studies focused on the human phosphatome.

METHODS

CoralP is adapted from Coral and is written using R and Javascript. The R package shiny (Chang et al., 2018) and extensions shiny-

dashboard (Chang and Borges Ribeiro, 2018), shinyBS (Bailey, 2015) and shinyWidgets (Perrier et al., 2019) were used for the web framework.

The R packages readr (Wickham et al., 2018) and rsvg (Ooms, 2018) were used for data manipulation and rendering the SVG elements.

RColorBrewer (Neuwirth, 2014) was utilized for color palettes.

The Circle and Force layouts were written using the D3.js library (Bostock et al., 2011).

The phosphatome tree was adapted from Chen et al. and manually redrawn using vector graphics in Adobe Illustrator (Chen et al., 2017).

ACKNOWLEDGMENTS

We thank Bulent Arman Aksoy, Jeff Hammerbacher, Matthew E. Berginski, and Shawn M. Gomez for their help with the infrastructure required to build and host Coral and CoralP.

FUNDING

D.H.P. is supported by the National Institutes of Health (NIH), National Human Genome Research Institute (NHGRI) grant R00HG008662 and National Institutes of Health (NIH), National Institute of General Medical Sciences (NIGMS) grant R35GM128645.

AUTHOR CONTRIBUTIONS

Conceptualization: DHP
Software: AM, EMD, ESD, DHP
Writing: AM, MLB, DHP
Visualization: EMD
Supervision: DHP
Resources: DHP

DECLARATION OF INTERESTS

The authors declare no competing interests.

WEB RESOURCES

CoralP Source Code, <https://github.com/PhanstielLab/coralp>
CoralP, <http://phanstiel-lab.med.unc.edu/coralp>

REFERENCES

Begovich, Ann B., Victoria E. H. Carlton, Lee A. Honigberg, Steven J. Schrodin, Anand P. Chokkalingam, Heather C. Alexander, Kristin G. Ardlie, et al. 2004. "A Missense Single-Nucleotide Polymorphism in a Gene Encoding a Protein Tyrosine Phosphatase (PTPN22) Is Associated with Rheumatoid Arthritis." *American Journal of Human Genetics* 75 (2): 330–37.

Chen, Mark J., Jack E. Dixon, and Gerard Manning. 2017. "Genomics and Evolution of Protein Phosphatases." *Science Signaling* 10 (471): eaal1910.

(474). <https://doi.org/10.1126/scisignal.aag1796>.

Cohen, Philip. 2002. "Protein Kinases--the Major Drug Targets of the Twenty-First Century?" *Nature Reviews. Drug Discovery* 1 (4): 309–15.

Hendriks, Wiljan J. A. J., Ari Elson, Sheila Harroch, Rafael Pulido, Andrew Stoker, and Jeroen den Hertog. 2013. "Protein Tyrosine Phosphatases in Health and Disease." *The FEBS Journal* 280 (2): 708–30.

Metz, Kathleen S., Erika M. Deoudes, Matthew E. Berginski, Ivan Jimenez-Ruiz, Bulent Arman Aksoy, Jeff Hammerbacher, Shawn M. Gomez, and Douglas H. Phanstiel. 2018. "Coral: Clear and Customizable Visualization of Human Kinome Data." *Cell Systems* 7 (3): 347–50.e1.

Ostman, Arne, Carina Hellberg, and Frank D. Böhmer. 2006. "Protein-Tyrosine Phosphatases and Cancer." *Nature Reviews. Cancer* 6 (4): 307–20.

Pulido, Rafael, and Rob Hooft van Huijsdijnen. 2008. "Protein Tyrosine Phosphatases: Dual-Specificity Phosphatases in Health and Disease." *The FEBS Journal* 275 (5): 848–66.

Robinson, Fred L., and Jack E. Dixon. 2005. "The Phosphoinositide-3-Phosphatase MTMR2 Associates with MTMR13, a Membrane-Associated Pseudophosphatase Also Mutated in Type 4B Charcot-Marie-Tooth Disease." *The Journal of Biological Chemistry* 280 (36): 31699–707.

Sacco, Francesca, Livia Perfetto, Luisa Castagnoli, and Gianni Cesareni. 2012. "The Human Phosphatase Interactome: An Intricate Family Portrait." *FEBS Letters* 586 (17): 2732–39.

Tonks, Nicholas K. 2006. "Protein Tyrosine Phosphatases: From Genes, to Function, to Disease." *Nature Reviews. Molecular Cell Biology* 7 (11): 833–46.

Andersen, J.N., Jansen, P.G., Echwald, S.M., Mortensen, O.H., Fukada, T., Del Vecchio, R., Tonks, N.K., and Møller, N.P.H. (2004). A genomic perspective on protein tyrosine phosphatases: gene structure, pseudogenes, and genetic disease linkage. *FASEB J.* 18, 8–30.

Bailey, E. (2015). shinyBS: Twitter Bootstrap Components for Shiny.

Begovich, A.B., Carlton, V.E.H., Honigberg, L.A., Schrodin, S.J., Chokkalingam, A.P., Alexander, H.C., Ardlie, K.G., Huang, Q., Smith, A.M., Spoerke, J.M., et al. (2004). A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* 75, 330–337.

Bostock, M., Ogievetsky, V., and Heer, J. (2011). D3: Data-Driven Documents. *IEEE Trans. Vis. Comput. Graph.* 17, 2301–2309.

Chang, W., and Borges Ribeiro, B. (2018). shinydashboard: Create Dashboards with "Shiny."

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and McPherson, J. (2018). shiny: Web Application Framework for R.

Chartier, M., Chénard, T., Barker, J., and Najmanovich, R. (2013). Kinome Render: a stand-alone and web-accessible tool to annotate the human protein kinome tree. *PeerJ* 1, e126.

Chen, M.J., Dixon, J.E., and Manning, G. (2017). Genomics and evolution of protein phosphatases. *Sci. Signal.* 10.

Cohen, P. (2002). Protein kinases--the major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.* 1, 309–315.

Eid, S., Turk, S., Volkamer, A., Rippmann, F., and Fulle, S. (2017). KinMap: a web-based tool for interactive navigation through human kinome data. *BMC Bioinformatics* 18, 16.

Hendriks, W.J.A.J., Elson, A., Harroch, S., Pulido, R., Stoker, A., and den Hertog, J. (2013). Protein tyrosine phosphatases in health and disease. *FEBS J.* 280, 708–730.

Metz, K.S., Deoudes, E.M., Berginski, M.E., Jimenez-Ruiz, I., Aksoy, B.A., Hammerbacher, J., Gomez, S.M., and Phanstiel, D.H. (2018). Coral: Clear and Customizable Visualization of Human Kinome Data. *Cell Syst* 7, 347–350.e1.

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes.

Ooms, J. (2018). rsvg: Render SVG Images into PDF, PNG, PostScript, or Bitmap Arrays.

Ostman, A., Hellberg, C., and Böhmer, F.D. (2006). Protein-tyrosine phosphatases and cancer. *Nat. Rev. Cancer* 6, 307–320.

Perrier, V., Meyer, F., and Granjon, D. (2019). shinyWidgets: Custom Inputs Widgets for Shiny.

Pulido, R., and Hooft van Huijsdijnen, R. (2008). Protein tyrosine phosphatases: dual-specificity phosphatases in health and disease. *FEBS J.* 275, 848–866.

Robinson, F.L., and Dixon, J.E. (2005). The phosphoinositide-3-phosphatase MTMR2 associates with MTMR13, a membrane-associated pseudophosphatase also mutated in type 4B Charcot-Marie-Tooth disease. *J. Biol. Chem.* 280, 31699–31707.

Sacco, F., Perfetto, L., Castagnoli, L., and Cesareni, G. (2012). The human phosphatase interactome: An intricate family portrait. *FEBS Lett.* 586, 2732–2739.

Tiganis, T., and Bennett, A.M. (2007). Protein tyrosine phosphatase function: the substrate perspective. *Biochem. J* 402, 1–15.

Tonks, N.K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* 7, 833–846.

Wickham, H., Hester, J., and Francois, R. (2018). readr: Read Rectangular Text Data.