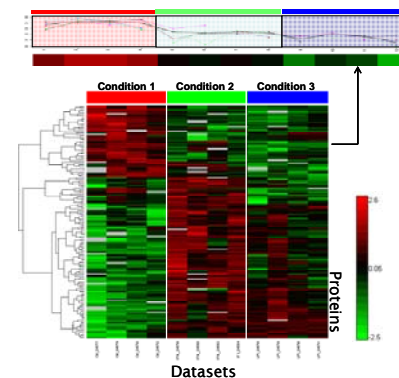


## EMSL helps build new tools to advance protein research

*Collaborative spirit provides no-cost resources to scientists around the world*

From developing bioenergy to detecting disease, proteins are crucial structural and functional elements of microbes, plants, animals, and humans. At the Department of Energy's EMSL, biological experts developed and deployed tools to further protein research. One of these tools, DAnTE (Data Analysis Tool Extension), was built by Pacific Northwest National Laboratory. DAnTE aids in performing downstream data analysis including normalization, missing value imputation algorithms, peptide to protein rollup methods, and hypothesis testing. Download the software at <http://omics.pnl.gov>.

Another tool, by University of Wisconsin-Madison, PNNL, and University of Texas-Houston collaborators, is a data analysis framework for temporal bottom-up proteomics data that can perform abundance normalization, missing-value imputation, and inference of peptide- and protein-level dynamic patterns. The framework was demonstrated on data from an *R. sphaeroides* time-course study that examined the transition between aerobic respiration and photosynthesis. The algorithms were implemented in MATLAB code to perform an analysis of temporal bottom-up proteomics data; the code is publicly available at [http://ober-proteomics.pnl.gov/software/Du\\_TimeCourseAnalysis\\_01072008.zip](http://ober-proteomics.pnl.gov/software/Du_TimeCourseAnalysis_01072008.zip).



*Plot from DAnTE shows a candidate set of proteins identified as significant biomarkers that differentiate three treatment conditions. Top bar shows the profile of one protein from the set.*

**Scientific impact:** Making new proteomics tools available at no cost to the scientific community allows more researchers to enter the proteomics field without investing in expensive tools. This work is just one example of EMSL's contributions to predicting biological functions from molecular and chemical data.

**Social impact:** Better tools for protein identification are vital to solving intractable problems such as converting agricultural waste into fuels, detecting bio-based threats, and quickly detecting and treating disease.

For more information, contact EMSL Communications Manager Mary Ann Showalter (509-371-6017).

**Citations:** Du X, SJ Callister, NP Manes, JN Adkins, RA Alexandridis, X Zeng, JH Roh, WE Smith, TJ Donohue, S Kaplan, RD Smith, and MS Lipton. 2008. "A Computational Strategy to Analyze Label-Free Temporal Bottom-Up Proteomics Data." *Journal of Proteome Research* doi:10.1021/pr0704837.

Polpitiya AD, WJ Qian, N Jaitly, VA Petyuk, JN Adkins, DG Camp II, GA Anderson, and RD Smith. 2008. "DAnTE: A Statistical Tool for Quantitative Analysis of -omics Data." *Bioinformatics* doi:10.1093/bioinformatics/btn217.

**Acknowledgments:** The DAnTE research was funded by the National Institutes of Health's (NIH's) National Institute of General Medical Sciences, National Center for Research Resources, and National Institute of Allergy and Infectious Diseases (NIAID) along with PNNL Laboratory Directed Research and Development program. The data analysis framework research was supported by the DOE Office of Biological and Environmental Research and NIH's NIAID and National Institute of General Medical Sciences.