YENEPOYA
(DEEMED TO BE UNIVERSITY)
Recognized under Sec 3(A) of the UGC Act 1956
Accredited by NAAC with 'A' Grade

Details of the Collaborative Activity

2019-20

Name of the Collaborating Institute: Division of Biochemistry and Informatics Institute, University of Missouri, Columbia, Missouri 65211, United States

Name of the Collaborating Department: Yenepoya Research Centre

Activities:

Dr. Shyam Prasad Rao collaborated with Faculty from Division of Biochemistry, and Informatics Institute, University of Missouri, Columbia, for proteomic based research studies and resulted in the publications of the following article.

- Ahsan N, Wilson RS, **Rao RSP**, Salvato F, Sabila M, Ullah H, Miernyk JA. "Mass spectrometry-based identification of phospho-Tyr in plant proteomics. *Journal of Proteome Research*. 2020; 19:561-571.

ATTESTED

Dr. Gangadhara Somayaji K S
Registrar
Yenepoya (Deemed to be University)
University Road, Deralakatte
Mangalore 575 018, Karnataka.

Journal of proteome research

pubs.acs.org/jpr

Reviews

Mass Spectrometry-Based Identification of Phospho-Tyr in Plant Proteomics

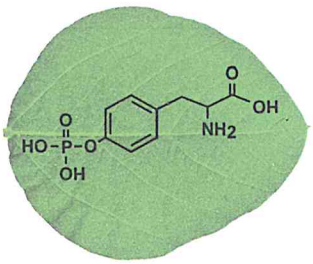
Nagib Ahsan,* Rashaun S. Wilson, R. Shyama Prasad Rao, Fernanda Salvato, Mercy Sabila, Hemayet Ullah, and Ján A. Miernyk*

Cite This: https://dx.doi.org/10.1021/acs.jproteome.9b00550

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: O-Phosphorylation (phosphorylation of the hydroxyl-group of S, T, and Y residues) is among the first described and most thoroughly studied posttranslational modification (PTM). Y-Phosphorylation, catalyzed by Y-kinases, is a key step in both signal transduction and regulation of enzymatic activity in mammalian systems. Canonical Y-kinase sequences are absent from plant genomes/kinomes, often leading to the assumption that plant cells lack *O*-phospho-L-tyrosine (pY). However, recent improvements in sample preparation, coupled with advances in instrument sensitivity and accessibility, have led to results that unequivocally disproved this assumption. Identification of hundreds of pY-peptides/proteins, followed by validation using genomic, molecular, and biochemical approaches, implies previously unappreciated roles for this "animal PTM" in plants. Herein, we review extant results from studies of pY in plants and propose a strategy for preparation and analysis of pY-peptides that will allow a depth of coverage of the plant pY-proteome comparable to that achieved in mammalian systems.



KEYWORDS: *mass spectrometry, plant proteomics, regulation, signaling, tyrosine phosphorylation*

INTRODUCTION

The total number of proteins comprising the proteome is many-fold larger than the translated genome.[1] A substantial component of this difference is the result of posttranslational modifications. O-Phosphorylation (phosphorylation of the hydroxyl-group of Ser (S), Thr (T), and Tyr (Y) residues) is among the first described and most thoroughly studied posttranslational modification.[2] O-Phosphorylation has been shown to be critically important in regulation of function at the levels of proteins, organelles, cells, tissues, organs, and whole organisms.[3,4] It has been suggested that the human P-proteome includes more than 100 000 O-phosphorylation sites.[5]

Prior to the development of contemporary mass spectrometers,[6,7] analysis of protein O-phosphorylation typically involved incubation of cells with ^{32}Pi, separation of proteins by two-dimensional (2-D) electrophoresis, excision of radiolabeled spots, total chemical hydrolysis of the proteins, separation of (P)-amino acids, and identification by either 2-D thin-layer electrophoresis[8−11] or HPLC-retention times. Potential P-proteins were characterized by M_r and pI values, although it is necessary to keep in mind that 2-D gel spots seldom contain a single protein.[12] In some instances, physiological manipulations were used to provide insight into protein IDs from 2-D gel spots.[13] Additional protein ID information could be obtained from immunoblots,[14] or, if the sample protein was sufficiently abundant, from N-terminal sequence analysis.[15] These techniques allow the detection of P-proteins but do not indicate the position of the phosphorylation sites. In addition, sample recoveries are less than quantitative at each step, and final recoveries were often quite low, which is problematic when the phosphate/protein stoichiometry is low within the cell, all of which hamper studies of the P-proteome.[16] To circumvent these difficulties, the development of enrichment methods combined with sensitive mass spectrometry (MS)-based approaches paved the way for high-throughput identification of pY-sites/pY-peptides.

Improvements in recovery, methods, and instrument sensitivity led to simultaneous identification of P-proteins and analysis of P-sites, and the near universal adoption of an MS-based strategy for analysis of the P-proteome.[6,17,18] As our knowledge base increases in size and scope, it has been possible to quantify the occurrence of O-phosphorylation in each branch of the tree of life.

O-Phosphorylation is a simple, reversible post-translational modification (PTM) capable of regulating protein functions.[19−22] It has been estimated that O-phosphorylation affects as many as one-third of all proteins.[3] Phosphorylation of protein Y-residues is critical in the regulation of growth, differentiation, and in cell-signaling in animal systems.[23,24] However, lack of a bona fide tyrosine protein kinase (YPK) outside of metazoa led to the view that YPK-group kinases are

Received: August 18, 2019

frontiers in Plant Science

METHODS
published: 23 May 2018
doi: 10.3389/fpls.2018.00634

MU-LOC: A Machine-Learning Method for Predicting Mitochondrially Localized Proteins in Plants

Ning Zhang[1,2], R. S. P. Rao[2,3†], Fernanda Salvato[3†], Jesper F. Havelund[4], Ian M. Møller[4], Jay J. Thelen[2,3] and Dong Xu[1,2,5]*

[1] Informatics Institute, University of Missouri, Columbia, MO, United States, [2] Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, United States, [3] Department of Biochemistry, University of Missouri, Columbia, MO, United States, [4] Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark, [5] Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, United States

OPEN ACCESS

Edited by:
Chuang Ma,
Northwest A&F University, China

Reviewed by:
Shihua Zhang,
Academy of Mathematics and Systems Science (CAS), China
Fengfeng Zhou,
Jilin University, China

****Correspondence:***
Dong Xu
xudong@missouri.edu

†Present address:
R. S. P. Rao,
Biostatistics and Bioinformatics Division, Yenepoya Research Center, Yenepoya University, Mangalore, India
Fernanda Salvato,
Institute of Biology, University of Campinas, Campinas, Brazil

Specialty section:
This article was submitted to Plant Systems and Synthetic Biology, a section of the journal Frontiers in Plant Science

Received: *19 February 2018*
Accepted: *23 April 2018*
Published: *23 May 2018*

Citation:
Zhang N, Rao RSP, Salvato F, Havelund JF, Møller IM, Thelen JJ and Xu D (2018) MU-LOC: A Machine-Learning Method for Predicting Mitochondrially Localized Proteins in Plants. Front. Plant Sci. 9:634. doi: 10.3389/fpls.2018.00634

Targeting and translocation of proteins to the appropriate subcellular compartments are crucial for cell organization and function. Newly synthesized proteins are transported to mitochondria with the assistance of complex targeting sequences containing either an N-terminal pre-sequence or a multitude of internal signals. Compared with experimental approaches, computational predictions provide an efficient way to infer subcellular localization of a protein. However, it is still challenging to predict plant mitochondrially localized proteins accurately due to various limitations. Consequently, the performance of current tools can be improved with new data and new machine-learning methods. We present MU-LOC, a novel computational approach for large-scale prediction of plant mitochondrial proteins. We collected a comprehensive dataset of plant subcellular localization, extracted features including amino acid composition, protein position weight matrix, and gene co-expression information, and trained predictors using deep neural network and support vector machine. Benchmarked on two independent datasets, MU-LOC achieved substantial improvements over six state-of-the-art tools for plant mitochondrial targeting prediction. In addition, MU-LOC has the advantage of predicting plant mitochondrial proteins either possessing or lacking N-terminal pre-sequences. We applied MU-LOC to predict candidate mitochondrial proteins for the whole proteome of Arabidopsis and potato. MU-LOC is publicly available at http://mu-loc.org.

Keywords: machine learning, mitochondrial targeting, deep neural network, support vector machine, position weight matrix, gene co-expression

INTRODUCTION

Mitochondria play an essential role in plant cells. They are responsible for a diversity of biological processes, such as energy production, biosynthesis of several co-factors and vitamins, photorespiration, and programmed cell death (Millar et al., 2011; Welchen et al., 2014). It has been estimated that more than 95% of mitochondrial proteins in plants are encoded by nuclear genes (Millar et al., 2005). Two basic mechanisms exist for mitochondrial targeting. One group of proteins has cleavable N-terminal targeting peptides (also called pre-sequences), and the other group does not possess pre-sequences and instead has short, internal signal sequences that are