FUNGAL BIOLOGY XXX (2017) 1–7

British Mycological Society promoting fungal science

journal homepage: www.elsevier.com/locate/funbio

FPD: A comprehensive phosphorylation database in fungi☆

Youhuang BAI[a,b,1], Bin CHEN[b,1], Mingzhu LI[a,1], Yincong ZHOU[c], Silin REN[a], Qin XU[b], Ming CHEN[c,**], Shihua WANG[a,*]

[a]Key Laboratory of Pathogenic Fungi and Mycotoxins of Fujian Province and School of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, 350002, People's Republic of China
[b]Department of Bioinformatics, Fujian Agriculture and Forestry University, Fuzhou, 350002, People's Republic of China
[c]Department of Bioinformatics, State Key Laboratory of Plant Physiology and Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou, 310058, People's Republic of China

ARTICLE INFO

Article history:
Received 13 March 2017
Received in revised form
9 June 2017
Accepted 18 June 2017
Corresponding Editor:
Gustavo H. Henrique Goldman

Keywords:
Fungal
Motif
Phosphosites
Protein database

ABSTRACT

Protein phosphorylation, one of the most classic post-translational modification, plays a critical role in diverse cellular processes including cell cycle, growth, and signal transduction pathways. However, the available information about phosphorylation in fungi is limited. Here, we provided a Fungi Phosphorylation Database (FPD) that comprises high-confidence *in vivo* phosphosites identified by MS-based proteomics in various fungal species. This comprehensive phosphorylation database contains 62 272 non-redundant phosphorylation sites in 11 222 proteins across eight organisms, including *Aspergillus flavus*, *Aspergillus nidulans*, *Fusarium graminearum*, *Magnaporthe oryzae*, *Neurospora crassa*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Cryptococcus neoformans*. A fungi-specific phosphothreonine motif and several conserved phosphorylation motifs were discovered by comparatively analysing the pattern of phosphorylation sites in plants, animals, and fungi.

Introduction

Post-translational modification (PTM) plays a significant role in a wide range of cellular processes, including cell cycle, growth, and signal transduction pathways. The event of protein PTM provides an additional layer to determine the kinetics and cellular plasticity through the dynamical cell signalling networks. Phosphorylation is a very common mechanism for regulating the activity of enzymes to monitor events and initiate appropriate responses throughout eukaryotes. It mainly occurs on three types of amino acids (serine, threonine, and tyrosine). In fungi, phosphorylation is known to regulate several aspects of fungal biology, including central metabolism, cell cycle, transcription, maintenance of cellular integrity, and fungal pathogenicity (Manning *et al.* 2002; Oliveira *et al.* 2012; Albataineh & Kadosh 2016).

☆ Database URL: http://bis.zju.edu.cn/FPD/ or; http://funbiotox.fafu.edu.cn/bioinformatics/fpd/current/.

* *Corresponding author*. Tel./fax: +86 (0)591 87984471.
** *Corresponding author*. Tel./fax: +86 (0)571 88206612.
E-mail addresses: mchen@zju.edu.cn (M. Chen), wshyyl@sina.com (S. Wang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.funbio.2017.06.004

Please cite this article in press as: Bai Y, et al., FPD: A comprehensive phosphorylation database in fungi, Fungal Biology (2017), http://dx.doi.org/10.1016/j.funbio.2017.06.004

As the most prominent model organism for fungal research, yeast attracts special attention from the fungi community and it takes continuous effort to improve the quality of connection between the dynamic phosphorylation event and function of proteins (Amoutzias *et al.* 2012; Oliveira *et al.* 2012; Schulz *et al.* 2014; Kanshin *et al.* 2015). Although the phosphoproteomics techniques in yeast developed quickly, phosphoproteome analysis in other fungi is still limited. Xiong *et al.* identified and quantified 5882 phosphorylation sites in 1942 proteins with high confidence in *Neurospora crassa* (Xiong *et al.* 2014). Selvan *et al.* identified 1089 phosphopeptides derived from 648 proteins including 45 kinases in *Cryptococcus neoformans* (Selvan *et al.* 2014). Recently, our group analysed the *Aspergillus flavus* phosphoproteome and identified 598 high-confidence phosphorylation sites in 278 phosphoproteins (Ren *et al.* 2016). And a great deal of effort has been made to identify phosphorylation events in other fungi, such as *Aspergillus nidulans* (Ramsubramaniam *et al.* 2014), *Alternaria brassicicola* (Davanture *et al.* 2014), *Botrytis cinerea* (Davanture *et al.* 2014; Lineiro *et al.* 2016), *Fusarium graminearum* (Rampitsch *et al.* 2012), *Magnaporthe oryzae* (Franck *et al.* 2015), and *Candida albicans* (Beltrao *et al.* 2009; Willger *et al.* 2015). It was revealed that changes in phosphorylation status of the fork-head family transcription factor Fkh2 drive the pathogenic switch in the opportunistic human fungal pathogen *C. albicans* (Greig *et al.* 2015). Very recently, Studer *et al.* examined the regulation and evolution of protein phosphorylation across 18 different fungal species (Studer *et al.* 2016). With the accumulation of phosphoproteome studies in fungi, a large number of phosphorylation events and sites were identified, while data maintenance and sharing became increasingly challenging. These numerous phosphorylation data in fungi may contribute to expand the understanding of molecular mechanisms and functional roles of phosphorylation in fungal communities.

Until now, several online databases collected and integrated numerous phosphorylated substrates with their sites from different species, such as PhosphositePlus (Hornbeck *et al.* 2015), Phospho.ELM (Dinkel *et al.* 2011), PHOSIDA (Gnad *et al.* 2011), PhosphoPep (Bodenmiller *et al.* 2008), and P3DB (Yao *et al.* 2014). Recently, Xue's laboratory built three databases of protein phosphorylation sites in prokaryotes (Pan *et al.* 2015), plants (Cheng *et al.* 2014), animals and fungi (Ullah *et al.* 2016). However, the dbFAF database had only collected the phosphorylation sites in two fungi *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. Besides, several databases mainly focused on phosphorylation in specific species, for example, PhosphoGRID (*S. cerevisiae*) (Sadowski *et al.* 2013), PhosPhAt (*Arabidopsis thaliana*) (Zulawski *et al.* 2013), and HPRD (Human) (Goel *et al.* 2012). To our knowledge, only a limited proportion of the identified phosphoproteins and phosphosites in *S. cerevisiae* and *S. pombe* were covered by PhosphoGRID and dbFAF, while a growing number of phosphoproteome data in fungi need to be integrated for further analysis.

In this study, we provide a comprehensive collection of 62 272 non-redundant phosphorylation sites from 11 222 proteins that were collected from literature and database. The phosphoproteins were from eight fungal organisms, including *A. flavus*, *A. nidulans*, *F. graminearum*, *M. oryzae*, *N. crassa*, *S. cerevisiae*, *S. pombe*, and *C. neoformans*. These data were stored in the Fungal Phosphorylation Database (FPD). In addition, we identified several conserved phosphorylation motifs and a fungi-specific phosphothreonine motif by comparative analysis of phosphosite motifs in plants, animals, and fungi. Taken together, the FPD database could serve as a comprehensive protein phosphorylation data resource for further studies in fungi.

Methods

The FPD is a relational database built on MySQL server. The web application runs on an Apache version 2.4.7 server; in-house developed PHP scripts provide data retrieval. The web interface is based on PHP and CSS scripts. The FPD is publicly available at http://bis.zju.edu.cn/FPD/index.php.

We conducted a literature search using PubMed with multiple keywords: [('phosphoproteomics' OR 'phosphoproteomic' OR 'phosphoproteome') and ('fungal' OR 'fungi')]. All retrieved articles were carefully curated. Fungal protein sequences were collected from Uniprot database. The identified phosphorylated proteins, peptides, and sites in each fungus were downloaded (Supplementary Table S1). Although the phosphoproteome data of these fungi was analysed with different methods, a stringent filtering criteria of 1 % false discovery rate with the decoy fusion method was applied to identify the phosphopeptides in *Aspergillus flavus*, *Aspergillus nidulans*, *Cryptococcus neoformans*, *Fusarium graminearum*, *Magnaporthe oryzae*, and *Neurospora crassa*. For each specie, we mapped corresponding phosphorylated peptides to the protein sequences, and the phosphorylation sites were exactly pinpointed by using in-house Perl scripts. Besides curation from literature, phosphorylation sites of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* in public database dbPAF were integrated into our database. The detailed annotations of phosphoproteins were retrieved from Uniprot. The protein ID conversion was implemented using custom perl script.

To identify the conserved phosphosite motifs, we analysed the motifs of all phosphorylated sites in plants, animals, and fungi. All phosphorylated sites in four representative plants (*Arabidopsis thaliana*, *Medicago truncatula*, *Oryza sativa*, and *Zea mays*) were downloaded from dbPPT. The data in five animals (*Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Drosophila melanogaster*, and *Caenorhabditis elegans*) were downloaded from dbPAF. Phosphorylated peptides with a length of 13 amino acids with central characters of S/T/Y residues were prepared as the foreground data set, while non-phosphorylated peptides in the same proteins were regarded as the background data set. The phosphorylation motifs were calculated for three types of residues, respectively. By submitting phosphorylation peptides and non-phosphorylation peptides to Motif-X (http://motif-x.med.harvard.edu/motif-x.html), all significant motifs for each specie were gathered. The data of *H. sapiens* and *M. musculus* were excluded in phosphoserine motif analysis because the data size is too large to get result under the default parameter setting of Motif-X.

Database usage

In order to provide a convenient and efficient service for the dissemination of information about phosphorylation to the

Please cite this article in press as: Bai Y, et al., FPD: A comprehensive phosphorylation database in fungi, Fungal Biology (2017), http://dx.doi.org/10.1016/j.funbio.2017.06.004