



PTM Bioinformatics

Yu Xue

Huazhong University of Science and Technology



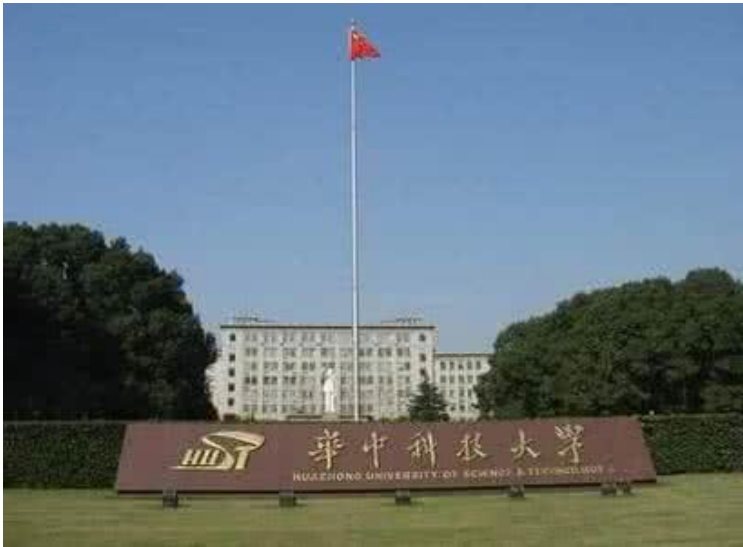


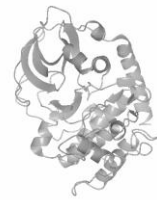
Wuhan, China





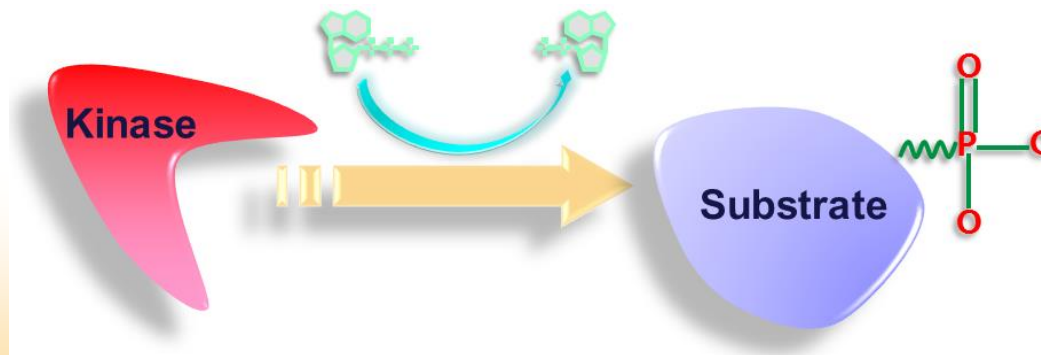
Huazhong University of Science and Technology





Post-translational modification

- **POST**: after RNA translated to protein
- **Covalent modification**: creation or disruption of covalent bond
- **Side or main chain of amino acid**
- **The PTM code**: writer-eraser-reader system
- **~610** type of PTMs
 - ◆ **Phosphorylation** (S/T/Y)
 - ◆ Acetylation, Ubiquitination, Propionylation (K)





PTM bioinformatics in China

- Methods for predicting PTM sites
- Databases of PTMs
- PTM proteome-based analysis
- Tools for PTM analysis
- **Future: PTM function prediction & validation**

遗传 Hereditas (Beijing) 2015 年 7 月, 37(7): 621—634
www.chinagene.cn

综述

Post-translational modification (PTM) bioinformatics in China: progresses and perspectives

Zexian Liu¹, Yudong Cai², Xuejiang Guo³, Ao Li⁴, Tingting Li⁵, Jianding Qiu⁶, Jian Ren⁷,
Shaoping Shi⁸, Jiangning Song⁹, Minghui Wang⁴, Lu Xie¹⁰, Yu Xue¹, Ziding Zhang¹¹,
Xingming Zhao¹²





Phosphorylation

- The most important and well-studied PTM
- Reversibility and regulation:
 - ◆ **Protein kinase**: Phosphorylation - *writer*
 - ◆ **Phosphatase**: De-phosphorylation - *eraser*
 - ◆ **Phospho-binding domain (PDB)**: interacts with phospho-sites - *reader*



The Nobel Prize in Physiology
or Medicine 1992

Edmond H. Fischer

Edwin G. Krebs





EKPD: Eukaryotic Kinase and Phosphatase Database



- Known data: 1855 kinases & 347 phosphatases
- EKPD: 50,433 kinases and 11,296 phosphatases in 84 eukaryotic species
- **iEKPD 2.0**: 148 species & phospho-binding proteins



The CUCKOO Workgroup

iEKPD - Integrated Eukaryotic Kinase & Phosphatase Database
Version 1.0



HOME BROWSE SEARCH DOCUMENTATION LINK USER GUIDE MAILING LIST

PRODUCTS OF CUCKOO

PTMs Predictor

[GPS \(Phosphorylation \)](#)
[iGPS \(Phosphorylation \)](#)
[CSS-Palm \(Palmitoylation \)](#)
[GPS-Lipid \(Lipid modifications \)](#)
[GPS-SUMO \(Sumoylation \)](#)
[GPS-SNO \(S-nitrosylation \)](#)
[GPS-YN02 \(Tyrosine Nitration \)](#)
[GPS-CCD \(Calpain Cleavage \)](#)
[GPS-Polo \(Polo-like Kinases \)](#)
[GPS-PUP \(Pupylation \)](#)

※ iEKPD database

In eukaryotes, phosphorylation-dependent signaling networks are, to a large extent, determined by the combined actions of protein **kinases**, protein **phosphatases** and **phosphoprotein-binding domains (PPBDs)** (Lim *et al.*, 2010). Protein kinase is a type of well understood enzyme which modifies other proteins by chemically adding phosphate groups to them (phosphorylation). Protein kinases have been found to be involved in varieties of cellular processes, including metabolism (Viollet and Andreelli 2011), transcription, cell cycle progression (Moniz *et al.*, 2011), cytoskeleton rearrangement and cell movement (Huang *et al.*, 2009), cell apoptosis (Wang, 2000), and differentiation (Taylor *et al.*, 2011). Contrary to phosphorylation, dephosphorylation is catalyzed by protein phosphatase through a way of removing a phosphate group from its substrate. Protein phosphatases contain two groups, PSP and PTP, which are responsible for the ser/thr dephosphorylation and tyr dephosphorylation respectively. Like protein kinases, protein phosphatases also play an important role in a lot of cellular processes, including proliferation (Zhang *et al.*, 2012), differentiation, cell adhesion (Besette *et al.*, 2008), motility and cell death (Gallego *et al.*, 2005). More over, recent studies discovered a few modular domains that particularly recognize pThr/pSer- or pThr-containing sequences, such as the breast-cancer-associated protein BRCA1 C-terminal (BRCT) repeats, SH2 domain and forkhead-associated (FHA) domain. These PPBD-containing proteins play a pivotal role in connecting the kinases and other effector molecules.

In this work, we have collected 1863 protein kinases, 383 protein phosphatases and 411 PPBDs-containing proteins from the scientific literature and various public databases. The data are further classified into 33 families for protein phosphatase, 148 families for protein kinase and 21 families for PPBD-containing proteins, respectively. To computationally detect more proteins in eukaryotes, we constructed hidden Markov model (HMM) profiles for these families. For families without it contains **156,006** unique prote

Phosphatase Database was lastly updated on June. **9, 2017**, which

<http://iekpd.biocuckoo.org/>





The classifications

- Kinase: 10 groups with 149 families
- Phosphatase: 10 groups with 33 families

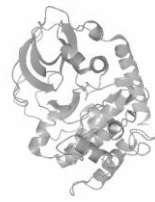
AGC		CAMK		CK1		TK		Other	
Akt	194.2	CAMK1	125.5	CK1	98.2	Abl	367.5	Aur	159.2
DMPK	81.2	CAMK2	116.4	Dual	407.8	Ack	97.8	BUB	87.8
GRK	188.5	CAMKL	89.9	TTBK	258.1	Alk	80.8	Bud32	93.1
MAST	86.3	CASK	183.8	TTBKL	164.1	Axl	159.8	CAMKK	138.3
NDR	113.1	DAPK	85.6	VRK	112.2	CCK4	196.2	CDC7	80.3
PDK1	98.5	DCAMKL	150.4	Worm10	434.6	Csk	144.6	HAL	185.7
PKA	153.2	MAPKAPK	112.7	Worm6	76.0	DDR	60.8	Haspin	124.7
PKG	164.6	MLCK	139.6	Worm7	203.5	EGFR	101.6	IKK	135.1
PKN	159.0	PHK	140.0	Worm8	524.7	Eph	120.1	IRE	86.2
RSK	101.2	PIM	149.7	Worm9	535.4	FAK	190.0	MOS	118.7
RSKL	81.5	PKD	101.6	Unclassified	63.4	FGFR	192.0	NAK	162.3
RSKR	139.7	PSK	284.6	TKL		NEK	81.4	NKF1	137.0
SGK	152.7	RAD53	106.7	IRAK	88.6	NKF2	125.0	NKF3	175.9
YANK	140.6	RSKb	78.9	LISK	75.8	NKF4	180.9	NKF5	92.3
Unclassified	38.9	TSSK	125.4	LRRK	118.2	NRBP	154.5	PEK	107.2
CMGC		Trbl	151.2	MLK	74.9	PLK	75.5	RAN	176.4
CDK	61.1	Trio	68.8	RAF	102.9	SCY	79.3	Slob	232.5
CDKL	110.8	Unclassified	102.8	RIPK	113.7	Ret	194.5	TBCK	58.8
CLK	91.2	STE		STKR	89.5	Ror	321.9	TLK	139.9
DYRK	97.7	STE11	113.4	Unclassified	136.5	Ryk	89.5	TOPK	149.2
GSK	124.1	STE20	42.2	Atypical		Sev	187.7	TTK	161.9
MAPK	61.5	STE7	75.6	ABC1	71.4	Src	115.7	ULK	64.1
RCK	92.0	Unclassified	117.2	Alpha	36.7	Syk	216.9	VPS15	220.9
SRPK	120.9	Unclassified		PDHK	120.3	Tec	116.1	WEE	63.6
Unclassified	109.9			PIKK	110.4	Tie	346.4	WNK	83.0
PK families				RIO	71.5	Trk	70.2	Worm1	581.4
				Unclassified		VEGFR	184.4	Worm2	492.7
						Unclassified	110.1	Worm3	639.1
						RGC		Worm4	413.5
						RGC	70.8	Worm5	436.4
						Unclassified		Unclassified	30.4

PPP		DSP	
PP1	169.7	aDSP	37.2
PP2A	120.6	MKP	69.0
PP2B	79.6	PRL	109.6
PP4	161.0	CDC14	60.3
PP5	120.9	SSH	72.7
PP6	168.1	Myotubularins	74.5
PP7	72.8	PTEN	65.0
Kelch	301.7	PPM	
SLP	67.4	PP2C_1	25.6
Unclassified	94.9	PP2C_2	38.7
Asp Based PTP		PP2C_3	24.7
FCP_SCP	30.4	PDP	171.0
EYA	49.7	LMWPTP	
MDP1	145.6	LMWPTP	77.2
Classical PTP		CDC25	
RPTP	62.1	CDC25	50.5
NRPTP	55.5	PTPLA	
		PTPLA	35.5

PP families	
-------------	--

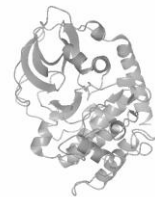
PK families

PP families



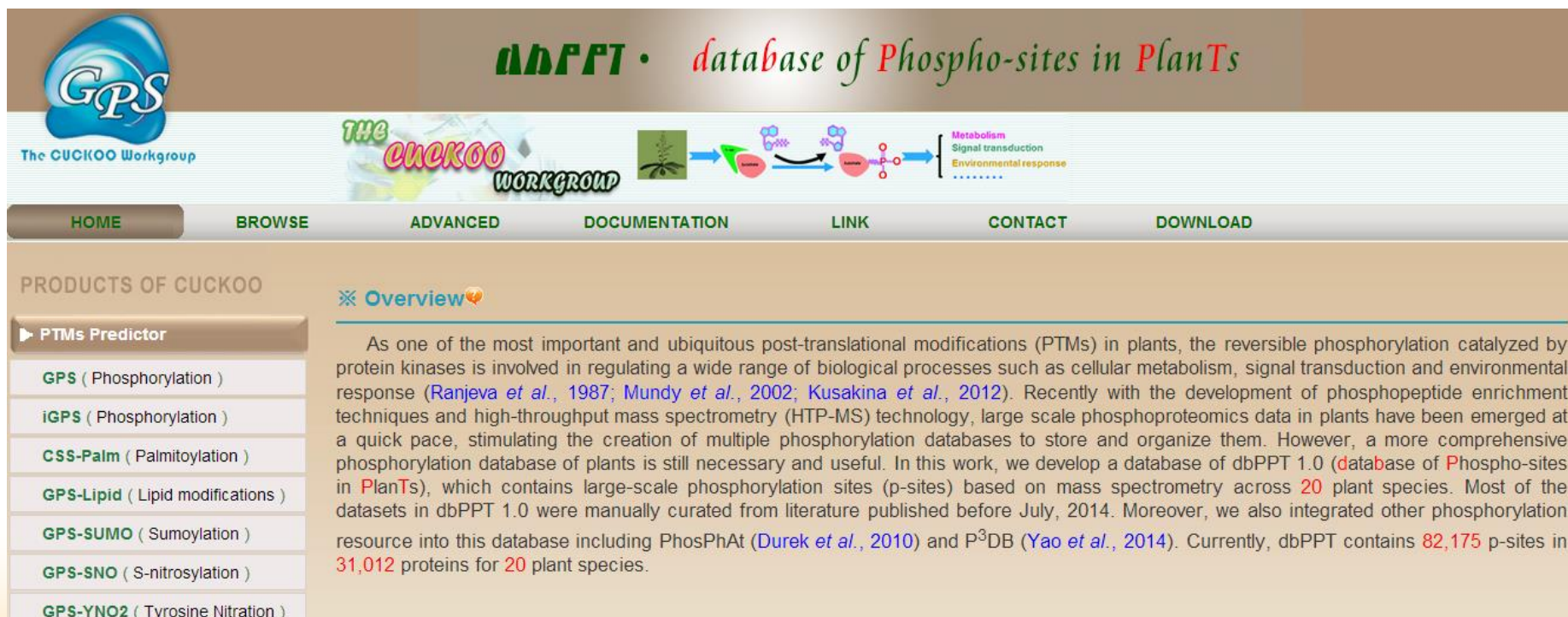
- **Kinase: 467 vs. 1,450**
- **Phosphatase: 144 vs. 192**





dbPPT

➤ Manual curation: 82,175 p-sites in 31,012 proteins for 20 plant species



The screenshot shows the dbPPT website. At the top, there's a header with the GPs logo (The CUCKOO Workgroup) on the left and the title "dbPPT • database of Phospho-sites in PlanTs" in the center. Below the header is a navigation bar with links: HOME, BROWSE, ADVANCED, DOCUMENTATION, LINK, CONTACT, and DOWNLOAD. The main content area is titled "PRODUCTS OF CUCKOO" and features a sidebar on the left with a list of PTMs Predictor tools: GPS (Phosphorylation), iGPS (Phosphorylation), CSS-Palm (Palmitoylation), GPS-Lipid (Lipid modifications), GPS-SUMO (Sumoylation), GPS-SNO (S-nitrosylation), and GPS-YNO2 (Tyrosine Nitration). The main text area is titled "Overview" and contains a paragraph describing the database. The paragraph states that phosphorylation is a key PTM in plants, involved in regulating various biological processes. It mentions the development of dbPPT 1.0, which contains large-scale phosphorylation sites (p-sites) based on mass spectrometry across 20 plant species. The database was manually curated from literature published before July 2014 and includes other phosphorylation resources like PhosPhAt and P³DB. Currently, dbPPT contains 82,175 p-sites in 31,012 proteins for 20 plant species.

dbPPT • database of Phospho-sites in PlanTs

THE CUCKOO WORKGROUP

HOME BROWSE ADVANCED DOCUMENTATION LINK CONTACT DOWNLOAD

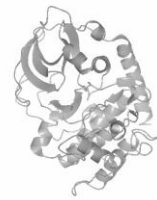
PRODUCTS OF CUCKOO

PTMs Predictor

- GPS (Phosphorylation)
- iGPS (Phosphorylation)
- CSS-Palm (Palmitoylation)
- GPS-Lipid (Lipid modifications)
- GPS-SUMO (Sumoylation)
- GPS-SNO (S-nitrosylation)
- GPS-YNO2 (Tyrosine Nitration)

Overview

As one of the most important and ubiquitous post-translational modifications (PTMs) in plants, the reversible phosphorylation catalyzed by protein kinases is involved in regulating a wide range of biological processes such as cellular metabolism, signal transduction and environmental response (Ranjeva *et al.*, 1987; Mundy *et al.*, 2002; Kusakina *et al.*, 2012). Recently with the development of phosphopeptide enrichment techniques and high-throughput mass spectrometry (HTP-MS) technology, large scale phosphoproteomics data in plants have been emerged at a quick pace, stimulating the creation of multiple phosphorylation databases to store and organize them. However, a more comprehensive phosphorylation database of plants is still necessary and useful. In this work, we develop a database of dbPPT 1.0 (database of Phospho-sites in PlanTs), which contains large-scale phosphorylation sites (p-sites) based on mass spectrometry across 20 plant species. Most of the datasets in dbPPT 1.0 were manually curated from literature published before July, 2014. Moreover, we also integrated other phosphorylation resource into this database including PhosPhAt (Durek *et al.*, 2010) and P³DB (Yao *et al.*, 2014). Currently, dbPPT contains 82,175 p-sites in 31,012 proteins for 20 plant species.



dbPSP

➤ 7,391 p-sites in 3,750 prokaryotic proteins



The screenshot displays the dbPSP website. The header features the 'GPS' logo with 'The CUCKOO Workgroup' text, the 'dbPSP Database of Phosphorylation Sites in Prokaryotes' title with 'Version 1.0', and a diagram of prokaryotic phosphorylation involving a kinase, substrate, and ATP/ADP cycle. A navigation bar includes links for HOME, BROWSE, ADVANCE, LINKS, USER GUIDE, CONTACT, and DOWNLOAD. The 'PRODUCTS OF CUCKOO' section lists various PTM predictors: GPS (Phosphorylation), iGPS (Phosphorylation), CSS-Palm (Palmitoylation), GPS-Lipid (Lipid modifications), GPS-SUMO (Sumoylation), GPS-SNO (S-nitrosylation), GPS-YN02 (Tyrosine Nitration), GPS-CCD (Calpain Cleavage), GPS-Polo (Polo-like Kinases), and GPS-PUP (Pupylation). The 'Overview' section provides a detailed description of the database's purpose and the challenges it addresses. The 'Search' section includes a search form with a dropdown menu, an input field, and buttons for 'Example', 'Clear Form', and 'Submit'.

GPS The CUCKOO Workgroup

dbPSP Database of Phosphorylation Sites in Prokaryotes Version 1.0

THE CUCKOO WORKGROUP Bacteria Archaea

Prokaryotic Phosphorylation

Kinase Substrate OH O-PI ATP ADP

HOME BROWSE ADVANCE LINKS USER GUIDE CONTACT DOWNLOAD

PRODUCTS OF CUCKOO

- PTMs Predictor
 - GPS (Phosphorylation)
 - iGPS (Phosphorylation)
 - CSS-Palm (Palmitoylation)
 - GPS-Lipid (Lipid modifications)
 - GPS-SUMO (Sumoylation)
 - GPS-SNO (S-nitrosylation)
 - GPS-YN02 (Tyrosine Nitration)
 - GPS-CCD (Calpain Cleavage)
 - GPS-Polo (Polo-like Kinases)
 - GPS-PUP (Pupylation)

※ **Overview**

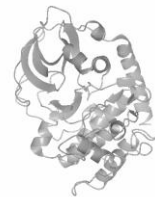
As one of the PTMs with tremendous studies, Phosphorylation regulates a wide variety of biological processes, including signal transduction events (Cohen, 1982). Whereas eukaryotic proteins of phosphorylation have been extensively studied, only limited information is available for phosphorylated proteins in prokaryotic organisms. Previous studies about constructing database of prokaryotic phosphorylation sites mainly focus on tyrosine-, serine, and threonine-phosphorylated proteins (Wurgler-Murphy, King and Kennelly, 2004). However, for the purpose of elucidating the mechanisms of phosphorylation in prokaryotic organisms, other residues which can also be phosphorylated should not be neglected. For example, protein histidine or aspartate phosphorylation plays important roles in two-signal-transduction events (Galperin, Nikolskaya, Koonin, 2001; Swanson, Alex and Simon, 1994).

To settle these challenges, we provide a comprehensive database of Prokaryotic Protein Phosphorylation Sites for 7 types of residues, including 7,391 phosphorylation sites in 3,750 proteins.

※ **Search** ?

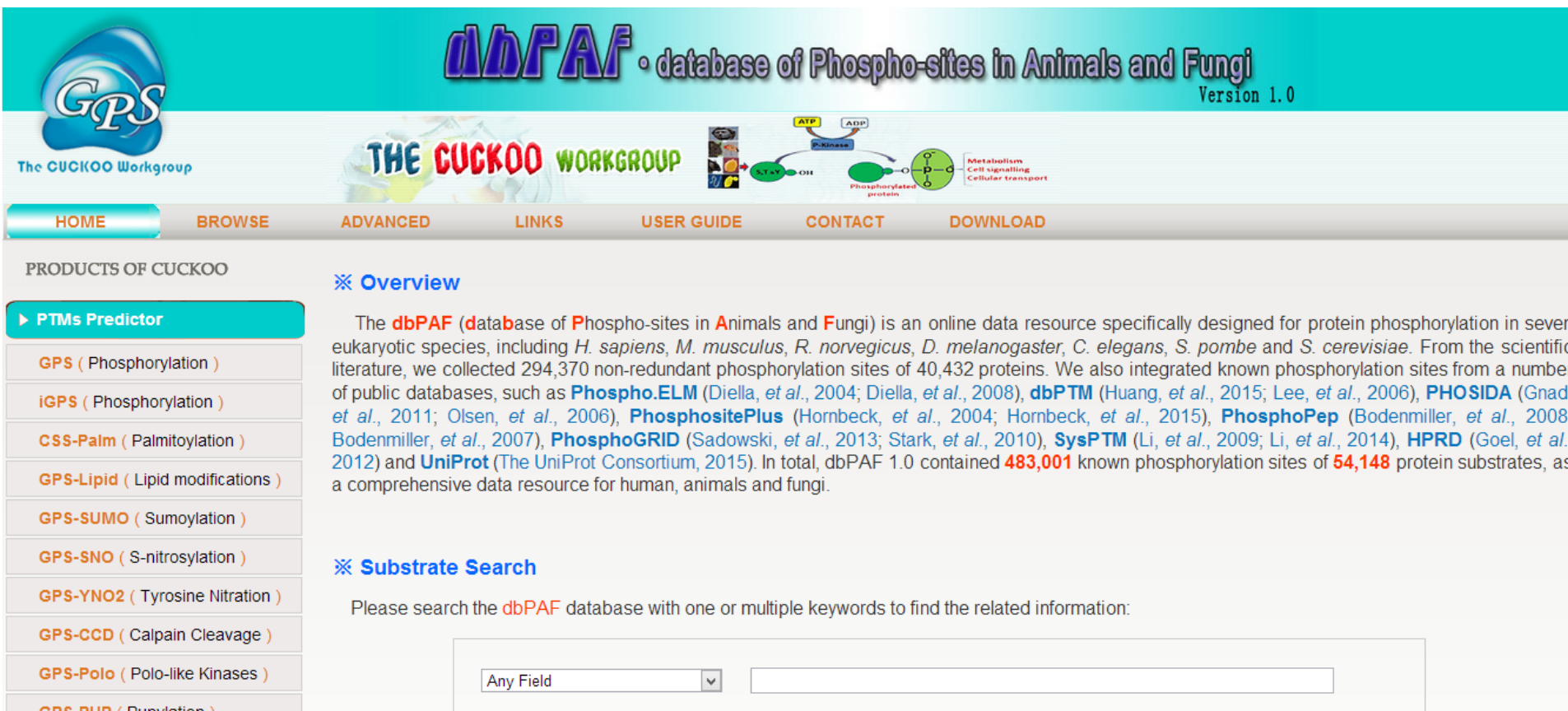
Any Field

Example Clear Form Submit



dbPAF

➤ 483,001 p-sites of 54,148 proteins for human, animals and fungi



The screenshot shows the dbPAF website interface. At the top, there is a header with the dbPAF logo and the text "database of Phospho-sites in Animals and Fungi" and "Version 1.0". Below the header is a navigation bar with links: HOME, BROWSE, ADVANCED, LINKS, USER GUIDE, CONTACT, and DOWNLOAD. The main content area is divided into two columns. The left column contains a sidebar with a "PRODUCTS OF CUCKOO" section, listing various PTMs Predictor tools: GPS (Phosphorylation), iGPS (Phosphorylation), CSS-Palm (Palmitoylation), GPS-Lipid (Lipid modifications), GPS-SUMO (Sumoylation), GPS-SNO (S-nitrosylation), GPS-YN02 (Tyrosine Nitration), GPS-CCD (Calpain Cleavage), GPS-Polo (Polo-like Kinases), and GPS-RIP (Ripulation). The right column features an "Overview" section with a paragraph describing the dbPAF database, followed by a "Substrate Search" section with a search bar and a dropdown menu.

dbPAF • database of Phospho-sites in Animals and Fungi
Version 1.0

THE CUCKOO WORKGROUP

ATP → ADP
P-Kinase
Phosphorylated protein
Metabolism
Cell signalling
Cellular transport

PRODUCTS OF CUCKOO

- PTMs Predictor
 - GPS (Phosphorylation)
 - iGPS (Phosphorylation)
 - CSS-Palm (Palmitoylation)
 - GPS-Lipid (Lipid modifications)
 - GPS-SUMO (Sumoylation)
 - GPS-SNO (S-nitrosylation)
 - GPS-YN02 (Tyrosine Nitration)
 - GPS-CCD (Calpain Cleavage)
 - GPS-Polo (Polo-like Kinases)
 - GPS-RIP (Ripulation)

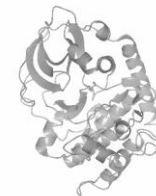
✧ **Overview**

The **dbPAF** (d**atabase of P**hospho-sites in **A**nimals and **F**ungi) is an online data resource specifically designed for protein phosphorylation in seven eukaryotic species, including *H. sapiens*, *M. musculus*, *R. norvegicus*, *D. melanogaster*, *C. elegans*, *S. pombe* and *S. cerevisiae*. From the scientific literature, we collected 294,370 non-redundant phosphorylation sites of 40,432 proteins. We also integrated known phosphorylation sites from a number of public databases, such as **Phospho.ELM** (Diella, *et al.*, 2004; Diella, *et al.*, 2008), **dbPTM** (Huang, *et al.*, 2015; Lee, *et al.*, 2006), **PHOSIDA** (Gnad, *et al.*, 2011; Olsen, *et al.*, 2006), **PhosphositePlus** (Hornbeck, *et al.*, 2004; Hornbeck, *et al.*, 2015), **PhosphoPep** (Bodenmiller, *et al.*, 2008; Bodenmiller, *et al.*, 2007), **PhosphoGRID** (Sadowski, *et al.*, 2013; Stark, *et al.*, 2010), **SysPTM** (Li, *et al.*, 2009; Li, *et al.*, 2014), **HPRD** (Goel, *et al.*, 2012) and **UniProt** (The UniProt Consortium, 2015). In total, dbPAF 1.0 contained **483,001** known phosphorylation sites of **54,148** protein substrates, as a comprehensive data resource for human, animals and fungi.

✧ **Substrate Search**

Please search the **dbPAF** database with one or multiple keywords to find the related information:

Any Field



CPLM & PLMD

- CPLM 2.0: 12 types of protein lysine modifications
- **PLMD 3.0: 20 lysine modifications**



CPLM • Compendium of Protein Lysine Modifications
Version 2.0



PLMD Protein Lysine Modification Database
version 3.0

THE CUCKOO WORKGROUP

metabolism protein synthesis sleep drugs abuse co

HOME BROWSE ADVANCED CITED BY LINKS USER GUIDE CONTACT

PRODUCTS OF CUCKOO

- + PTMs Predictor
- + Tools
- + Databases

0176140
Last update: Jun. 1st, 2017

※ **Overview**

PLMD (Protein Lysine Modifications Database) is an online data resource specifically designed for protein database (Liu *et al.*, 2011) and **CPLM 2.0** (Compendium of Protein Lysine Modifications) database (Liu *et al.*, 2007; Shahbazian *et al.*, 2007; Smith *et al.*, 2009), ubiquitination (Gao *et al.*, 2013), methylation (Cheng *et al.*, 2009; Zhang *et al.*, 2009), crotonylation (Tan *et al.*, 2011), malonylation (Xie *et al.*, 2012) (Rabut *et al.*, 2014; Soucy *et al.*, 2010), glutarylation (Tan *et al.*, 2014; Hirschey *et al.*, 2015), hydroxy (Posner *et al.*, 2013), formylation (Jiang *et al.*, 2007), carboxylation (Jimenez-Morales *et al.*, 2014), phospho

※ **Substrate Search**

Gene Name

Example Clear Form Submit

※ **Data Summary & Download:**

The data statistics of PLMD is shown below. The data for each type of PLMs and species or the total data set

Type	Protein	Site
Ubiquitination	25103	121742
Acetylation	33025	111253
Succinylation	6377	18593
Malonylation	3429	9584
Sumoylation	2801	8115
Glycation	3084	6591
Methylation	2819	6323
Glutarylation	211	715
Propionylation	192	413
Crotonylation	59	353
Pupylation	245	287
Formylation	97	188
Phosphoglycerlation	137	187
Hydroxylation	32	137
Butyrylation	28	96
2-hydroxyisobutyrylation	15	81
Neddylation	25	50
Carboxylation	36	36
Lipoylation	25	28
Biotinylation	8	8
Total	53501	284780

Species	Protein	Site
<i>Homo sapiens</i>	13378	131256
<i>Mus musculus</i>	7377	44065
<i>Saccharomyces cerevisiae</i>	3697	23561
<i>Rattus norvegicus</i>	5264	20928
<i>Escherichia coli</i>	1971	13600
<i>Emericella nidulans</i>	1918	4838
<i>Sulfolobus islandicus</i>	1158	3714
<i>Arabidopsis thaliana</i>	1834	3650
<i>Mycobacterium tuberculosis</i>	999	3161
<i>Plasmodium falciparum</i>	1214	3151
<i>Bacillus velezensis</i>	1146	2998
<i>Spiroplasma eriocheiris</i>	539	2494
<i>Schistosoma japonicum</i>	1251	2427
<i>Phytophthora sojae</i>	1177	2211
<i>Bacillus subtilis</i>	763	2039
<i>Drosophila melanogaster</i>	1051	2026
<i>Oryza sativa subsp</i>	913	2018
<i>Vibrio parahaemolyticus</i>	643	1929
<i>Corynebacterium glutamicum</i>	637	1897
Others	6571	12817
Total	53501	284780



Submit to JGG

- *Journal of Genetics and Genomics*
- Annual database/web server issue, IF: 4.051
- Article & Letter
- Guest Editors: Xiujie Wang & Yu Xue
- Submission: Early of December, 2017
- Publication: May, 2018
- E-mail: xueyu@hust.edu.cn



Journal of Genetics and Genomics
Volume 44, Issue 5, Pages 223-284 (20 May 2017)
Edited by Xiu-Jie Wang and Yu Xue



Journal of Genetics and Genomics

Volume 44, Issue 5, 20 May 2017, Pages 223–225



Editorial

Bioinformaticians wrestling with the big biomedical data

Yu Xue · Xiu-Jie Wang ·



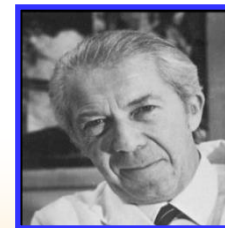
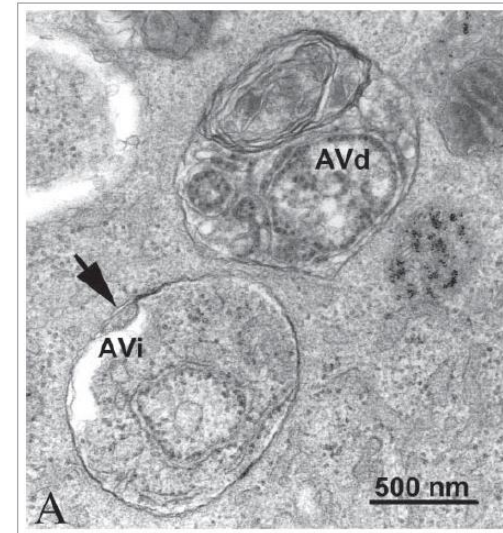
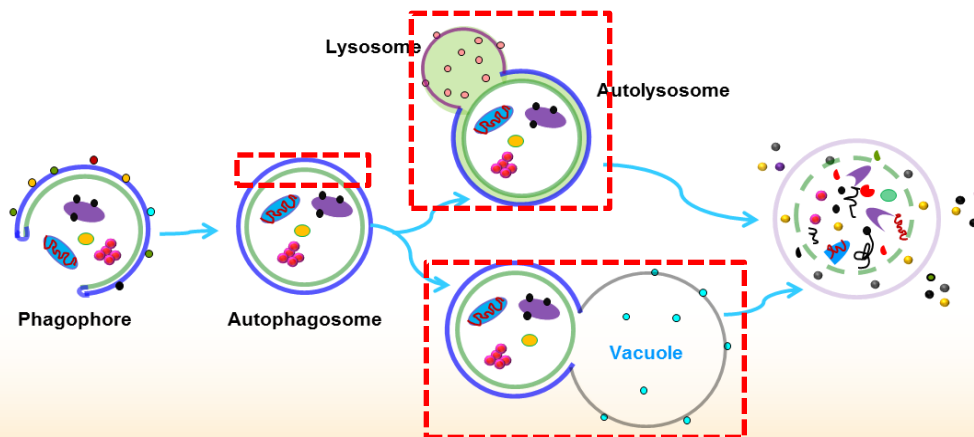
PTMomics

- Large-scale detection of *in vivo* PTM substrates, sites and motifs
 - ◆ Mass spectrometry (MS)
 - ◆ Protein, peptide, and PTM chips
- Current progress:
 - ◆ ~500,000 phosphorylation sites
 - ◆ ~140,000 ubiquitination sites
 - ◆ ~60,000 acetylation sites
- **Challenge:** What PTM Bioinformatics can do?



Autophagy

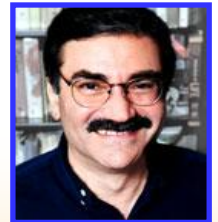
- **Macroautophagy**, microautophagy, chaperon-mediated autophagy (CMA)
- 1963, C de Duve, “self-eating” in Greek
- 1993, Atg1 in yeast
- 41 core ATG genes
- ~20 conserved in human



Christian de Duve



Yoshinori Ohsumi
(大隅良典)



Daniel J. Klionsky

Ohsumi Y, *Cell Res.*, 2014, 24, 9-23
Xie et al., *Autophagy*, 2015, 11, 28-45



Neuronal autophagy

- In Alzheimer's & Parkinson's diseases
 - ◆ Proteins accumulate in central nervous system
 - ◆ Defective autophagy in patient brains
- Enhanced autophagy
 - ◆ Neuroprotective by promoting the clearance of disease-associated aggregates
- Small-molecule autophagy enhancers
 - ◆ *Uncaria rhynchophylla* (Gouteng, 钩藤)

草部·钩藤

作者：李时珍

气味

甘、微寒、无毒。

主治

小儿惊热。用钩藤一两、硝石半两，甘草（炙）一分，共研为末。每服半钱，温水服，一天服三次。此方名“延龄散”。

斑疹。用钩藤的钩子、紫草茸，等分为末。每服三分或半钱，温酒送下。

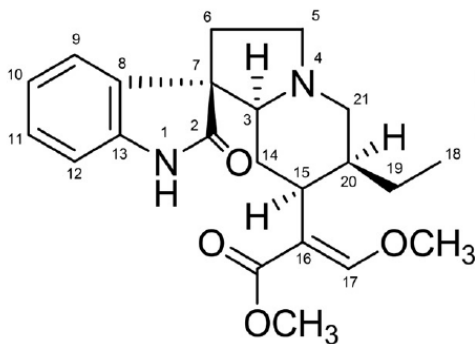




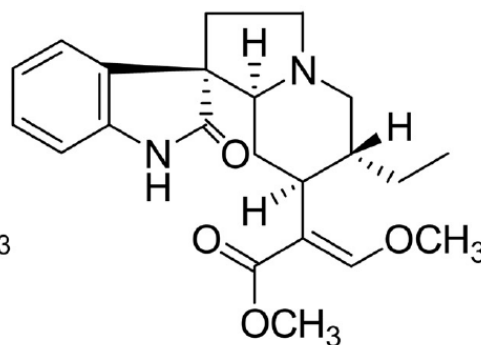
Neuroprotective alkaloids in Gouteng



- **Corynoxine (柯诺辛碱) & corynoxine B (柯诺辛B)**
 - ◆ Same molecular formula, different conformation
 - ◆ Induce autophagy in different way
- **Question:**
 - ◆ Find key regulators in neuronal autophagy
 - ◆ Distinguish two compounds



Corynoxine B

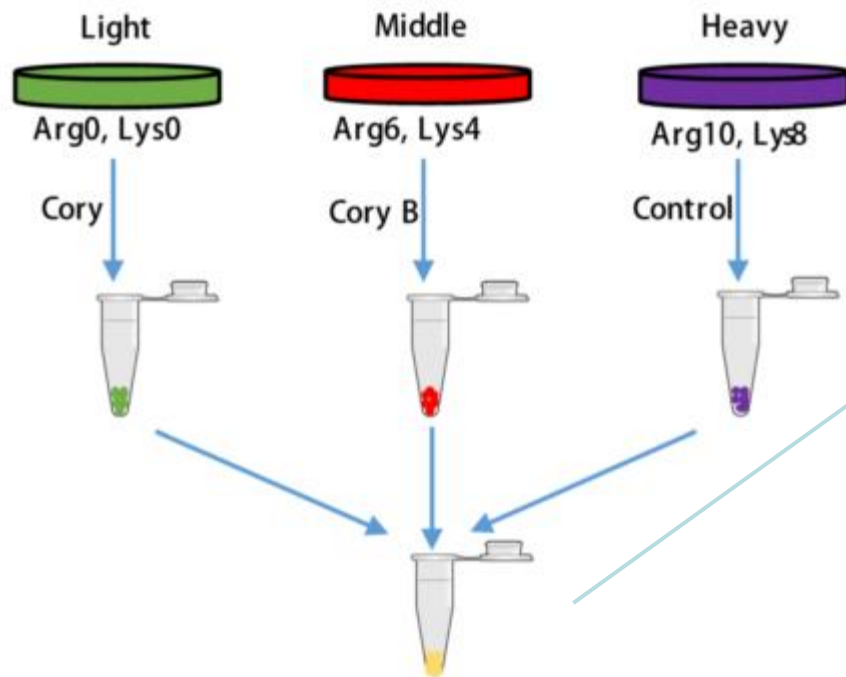


Corynoxine

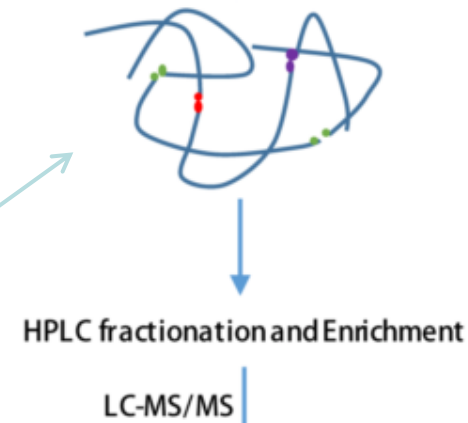


Experimental procedure

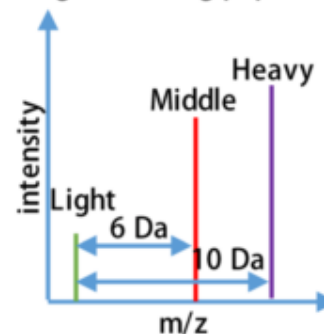
➤ N2a: a mouse neuroblastoma cell line



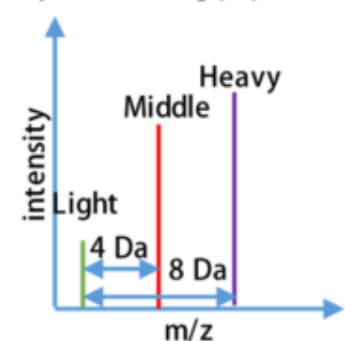
Tripsin digestion



Arg-containing peptide



Lys-containing peptide

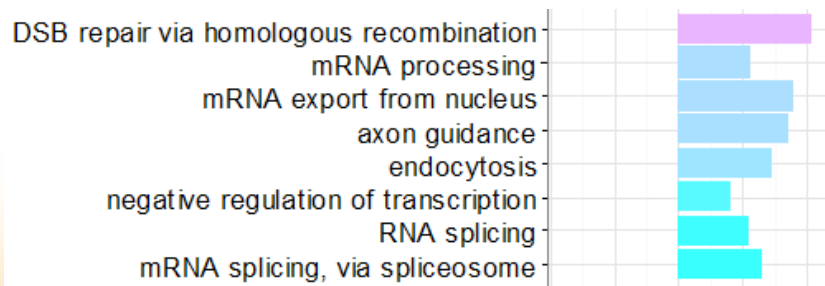
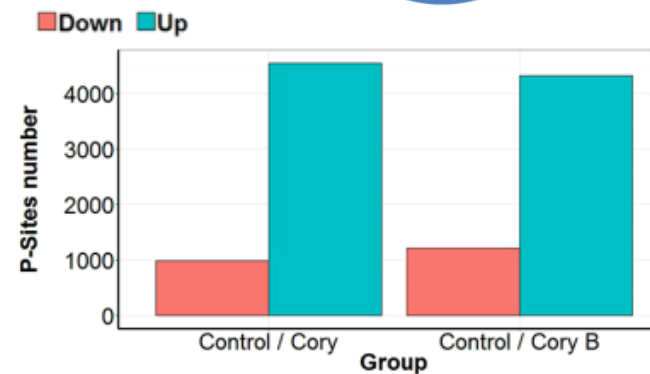
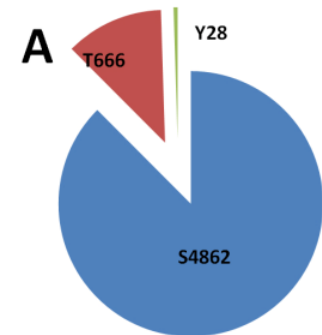


Treatment/Control ratio: >1 or <1

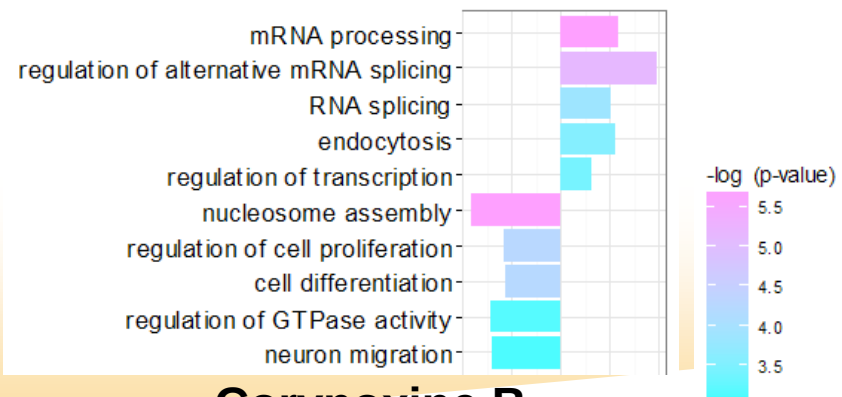


Phosphoproteomics profiling

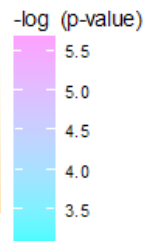
- Quantification: 2,317 proteins and 5,555 unique p-sites
- GSEA-based enrichment analysis
 - ◆ GO biological processes
- Limited difference in the distribution of p-sites



Corynoxine



Corynoxine B





'in vivo' GPS

➤ 'Kiss-then-farewell' model: the interaction of kinase-substrate

➤ iGPS:

- ◆ GPS algorithm
- ◆ Protein-protein interaction
- ◆ The phosphoproteomic data
- ◆ Much better than NetworkKIN

ICPS 1.0 - GPS algorithm with the interaction filter

File Tools Help

Phosphorylation

Protein Kinase

Site/Threonine Kinase

AKT

DMPK

GRK

PKA

PKB

PKC

PKG

RSK

SGK

CAMK

CK1

CMGC

STE

TKL

Atypical

Other

Tyrosine Kinase

TK

Predicted Site-specific Kinase-substrate Relations

Position	Code	Peptide	Matched ID	Gene Name	Kinase ID	Kinase Name	Interaction
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q2R1Y14	CRK7	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	P11862	CDK4	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q00537	PCTAIRE2	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q00538	PCTAIRE1	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q00535	CDK5	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q00524	CDK6	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q24821	PFTARE1	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q13523	PRF4	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04326	DYRK2	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04850	DYRK4	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04226	HPK2	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04225	HPK1	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04653	HPK4	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q43781	DYRK3	String
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q13627	DYRK1A	Exp/Strong
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q21453	DYRK1B	Exp/Strong
179	S	LVEDKPGRRRRSYT	ABK544	SFRS4	Q04452	HPK3	Exp/Strong

Enter the data in PhosPep/ELM/FASTA format

WNGLSPTSTDTTPKAPITLPSAREQMAPTLGER
KQAPPSVQAPGSPPTPAQOK
QLVAVYSQDSDNEELVER
KPGAGPYPSQAVLVD
pSAPPpSPPPGTR
DLDELLGNLpSETELK
SPFPKPPSPSWSSOR
KVGSPVK
EGMNPYSDEYADpSDEQHDAYLER

Options

Organism: H. sapiens

Format: PhosPep

Threshold: Low

Interaction: Exp/Strong

Console

PhosPep

Network

Clear

Submit

PK clusters	NetworkKIN				iGPS			
	Ac	Sn	Sp	MCC	Ac	Sn	Sp	MCC
NetworkKIN 1.1								
AGC/AKT	98.87%	58.89%	99.44%	0.5865	98.81%	52.22%	99.47%	0.5445
AGC/PKA	92.24%	32.20%	93.28%	0.1277	92.67%	57.56%	93.28%	0.2481
Atypical/PIKK/ATM	97.50%	77.42%	97.83%	0.5204	97.47%	75.27%	97.83%	0.5080
CAMK/CAMK2	98.43%	5.85%	99.91%	0.1671	98.51%	10.64%	99.91%	0.2580
CMGC/MAPK	93.45%	65.34%	94.06%	0.3312	93.60%	71.12%	94.09%	0.3614
CMGC/CDK/CDC2 ^a	94.58%	46.67%	95.69%	0.2825	94.75%	54.04%	95.69%	0.3268
Other/CK2	90.17%	55.49%	91.14%	0.2511	90.25%	54.85%	91.24%	0.2495
TK/EGFR	85.26%	12.16%	96.97%	0.1554	91.42%	56.76%	96.97%	0.6059
TK/Src	90.04%	23.97%	95.40%	0.2141	90.69%	32.64%	95.40%	0.2956

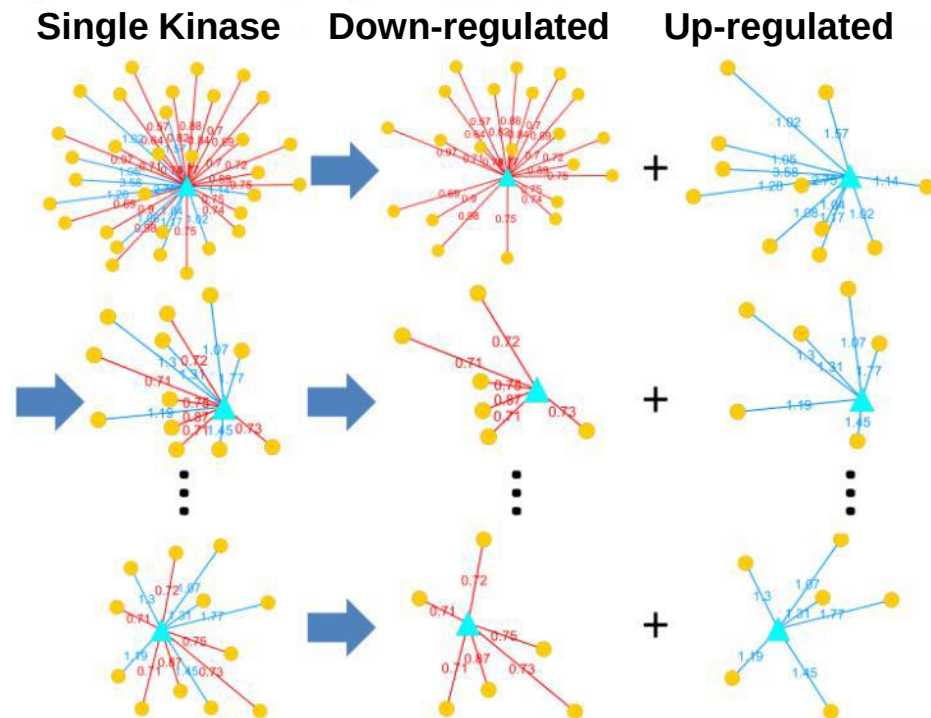
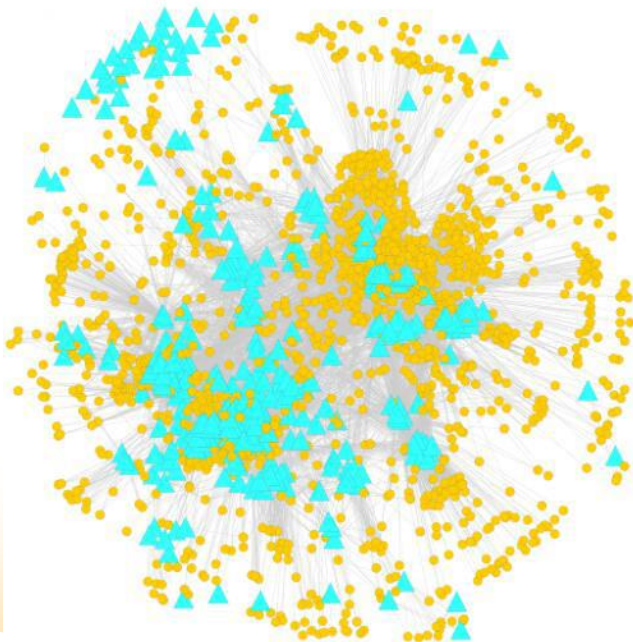
Song et al., MCP, 2012, 11, 1070-1083



Neuronal autophagy phosphorylation network



- iGPS: Cory- & Cory B-regulated phosphorylation networks
- Single kinase network
 - ◆ Down-regulated network & up-regulated network





iKAP algorithm

- For **Kinase i** , statistically test whether it prefer to be involved in up-regulated (high activity) or down-regulated (low activity) networks

- ◆ KA : kinase activity; KS : Treatment/Control (T/C) ratio

- **Up-regulated network**

- ◆ $T/C > 1$, $KA_{up}(i) = \sum_{j=1}^m int(KS_{ij})$

- **Down-regulated network**

- ◆ $T/C < 1$, $KA_{Down}(i) = \sum_{j=1}^n int(\frac{1}{KS_{ij}})$

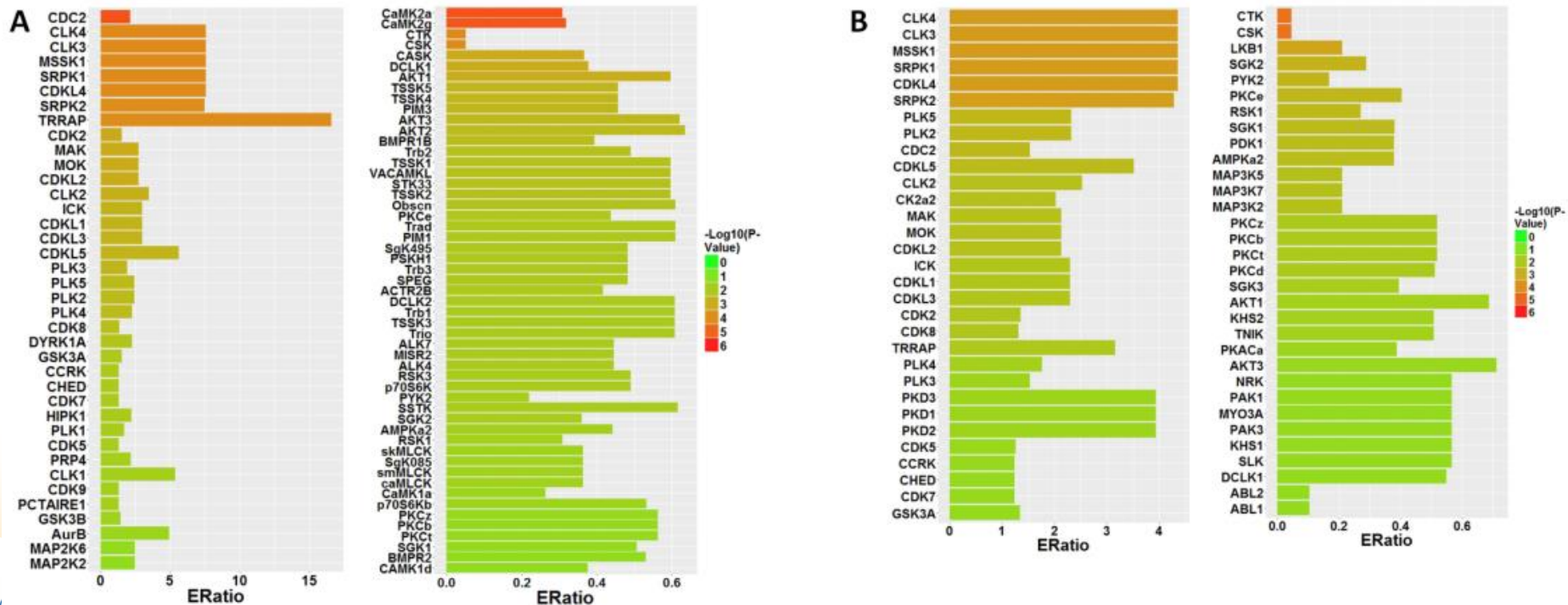
- **Yates' chi-squared test**

- ◆ $KA_{up} = \sum_{i=1}^k KA_{up}(i)$, $KA_{down} = \sum_{i=1}^l KA_{down}(i)$



Differentially activated kinases

- Up-regulated: 28 (Cory) & 28 (Cory B)
- Down-regulated: 51 (Cory) & 27 (Cory B)





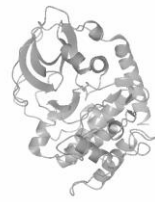
THANATOS database

- **THE Autophagy, Necrosis, Apoptosis OrchestratorS**
 - ◆ 144,153 proteins in 148 eukaryotes
 - ◆ Autophagy: 119 mouse kinases
- **THANATOS filter**
 - ◆ Up-regulated: 6 (Cory) & 4 (Cory B)
 - ◆ Down-regulated: 12 (Cory) & 7 (Cory B)

a. *The THANATOS database*

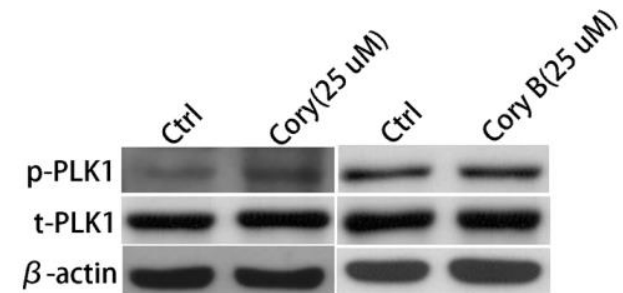
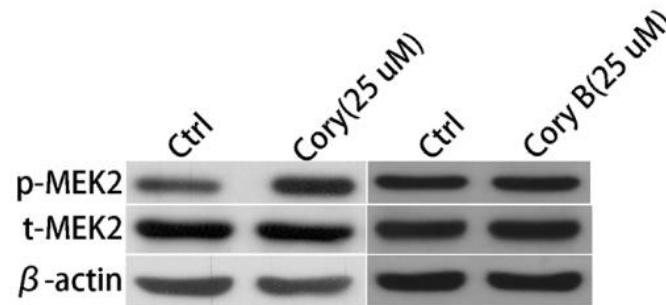
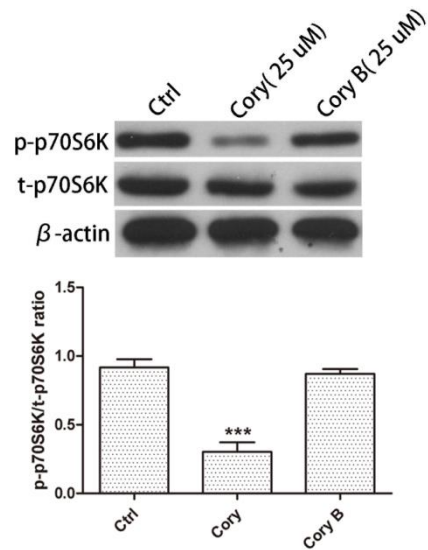
THANATOS (THE Apoptosis, Necrosis, AuTophagy OrchestratorS) is a resource being developed by the CUCKOO Workgroup at the Huazhong University of Science and Technology (Wuhan, Hubei, China). THANATOS is still under development (Y. Xue, personal communication) and it is focused on the integration of sequence data related to the main mechanisms leading to programmed cell death in eukaryotes. A simple web interface assists in data retrieval, using keyword searches, browsing by species and cell death type, performing BLAST searches with user-defined sequences, and by requesting the display of orthologs among predefined species. A Java application is also available to download for standalone usage of the THANATOS resource. The THANATOS database is publicly available online at the URL <http://thanatos.biocuckoo.org/>.

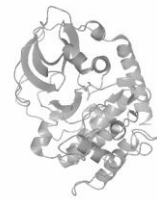




Validation

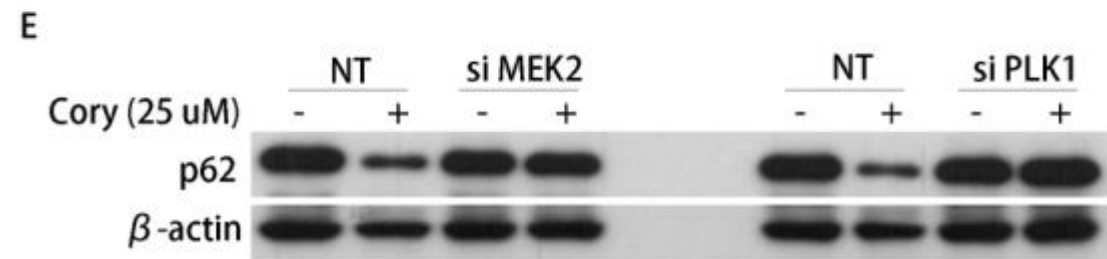
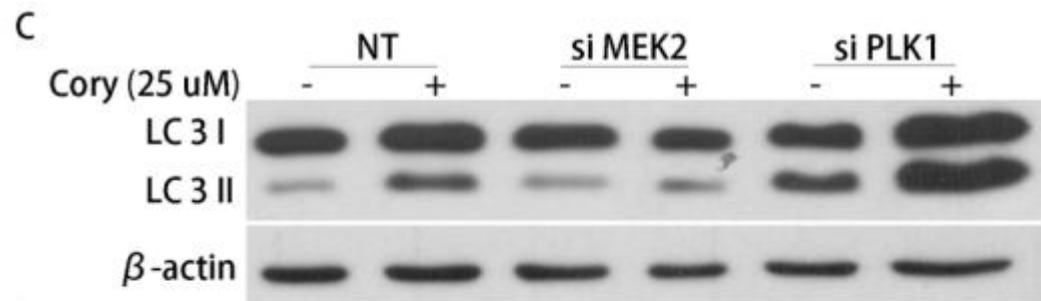
- **Cory: down-regulates p70S6K and up-regulates MEK2 & PLK1**





MEK2 & PLK1

- Silencing MEK2 but not PLK1 decreases LC3 II
- Silencing MEK2 & PLK1 both increase p62



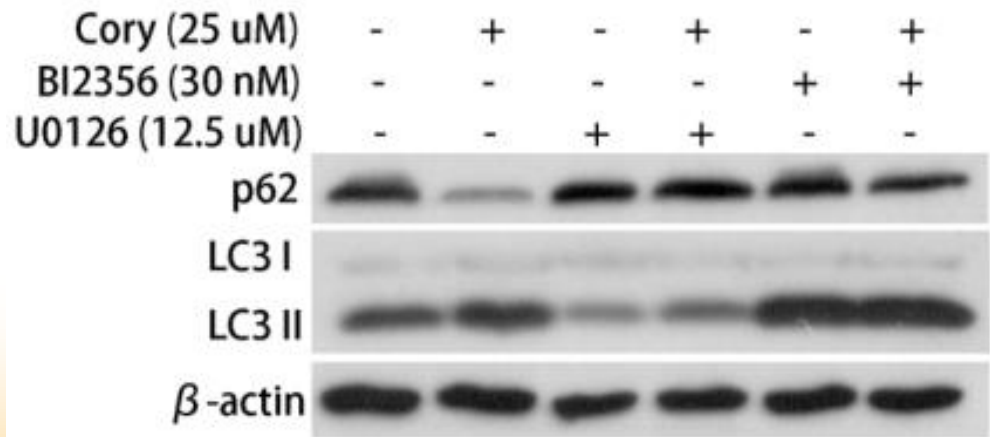
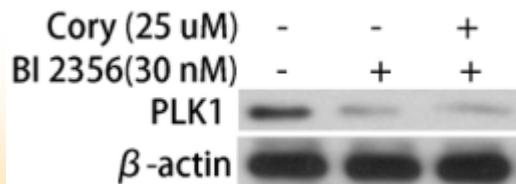
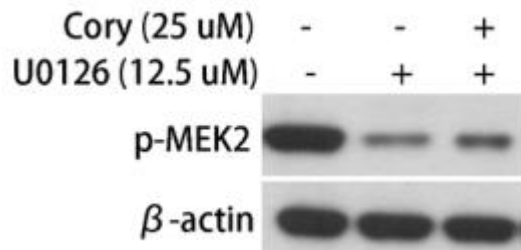
LC3 II $\uparrow$: autophagy
inhibition & activation

p62 $\downarrow$: autophagic flux $\uparrow$



MEK2 & PLK1 activation

- Inhibitors: U0126 (MEK2) & BI2356 (PLK1)
- The inhibitions of MEK2 and PLK1 both increase p62 and block autophagic flux

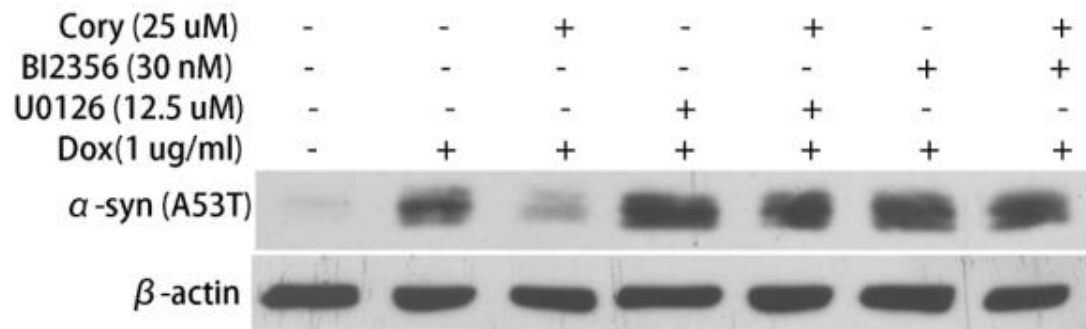
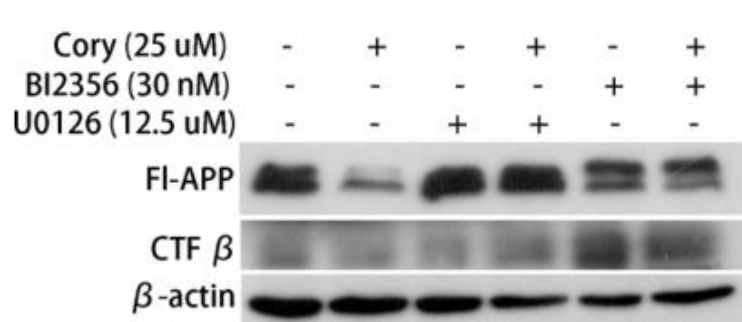




MEK2 & PLK1 in neuronal autophagy



- Alzheimer's disease: APP (β -amyloid precursor protein) & CTF β
- Parkinson's disease: α -synuclein (α -syn)
- The inhibition of MEK2 or PLK1 diminishes the clearance of disease-associated proteins by Cory
- The activation of MEK2 & PLK1: neuroprotective





Summarization

➤ PTM databases

- ◆ Phosphorylation: EKPD, dbPPT, dbPSP, dbPAF
- ◆ Lysine modifications: CPLM, PLMD

➤ Functional PTMs

- ◆ Functional protein kinases
- ◆ iKAP: a network-based algorithm
- ◆ Cory inhibits p70S6K and activates MEK2 & PLK1
- ◆ Inhibition of MEK2 & PLK1 block autophagic flux
- ◆ Activation of MEK2 & PLK1 is neuroprotective to clear disease-associated proteins



Acknowledgements

➤ Collaborators



Prof. Jian Ren,
SYSU



Prof. Li Yu, Tsinghua



Prof. Min Li, HKBU
Leilei Chen



Prof. Xiang-Jiao
Yang, McGill

➤ Students

◆ Zexian Liu,

◆ Tianshun Gao

◆ Han Cheng

◆ Yongbo Wang

◆ Wankun Deng

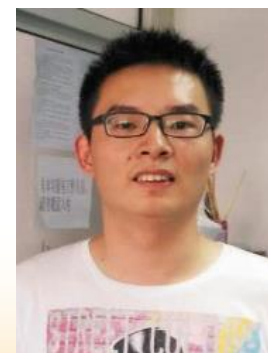
◆ Ying Zhang

◆ Jiaqi Zhou

◆ Shaofeng Lin

◆ Yang Xu

◆ Shuang Zhang



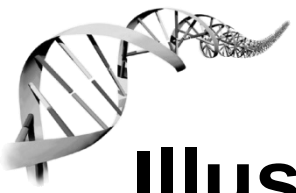
Prof. Zexian Liu,
SYSU



Thanks!



Any questions?



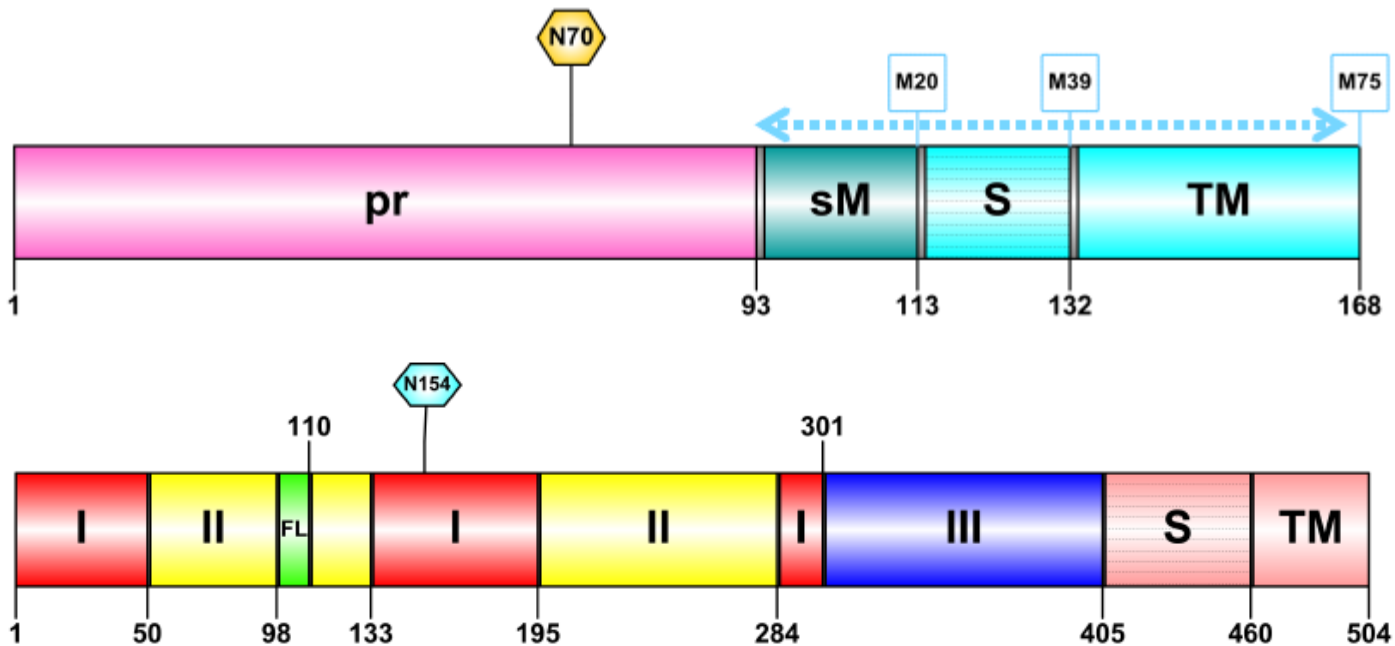
Illustrator for Biological Sequences (IBS)

Science

REPORTS

Cite as: Sirohi *et al.*, *Science*
10.1126/science.aaf5316 (2016).

The 3.8 Å resolution cryo-EM structure of Zika virus

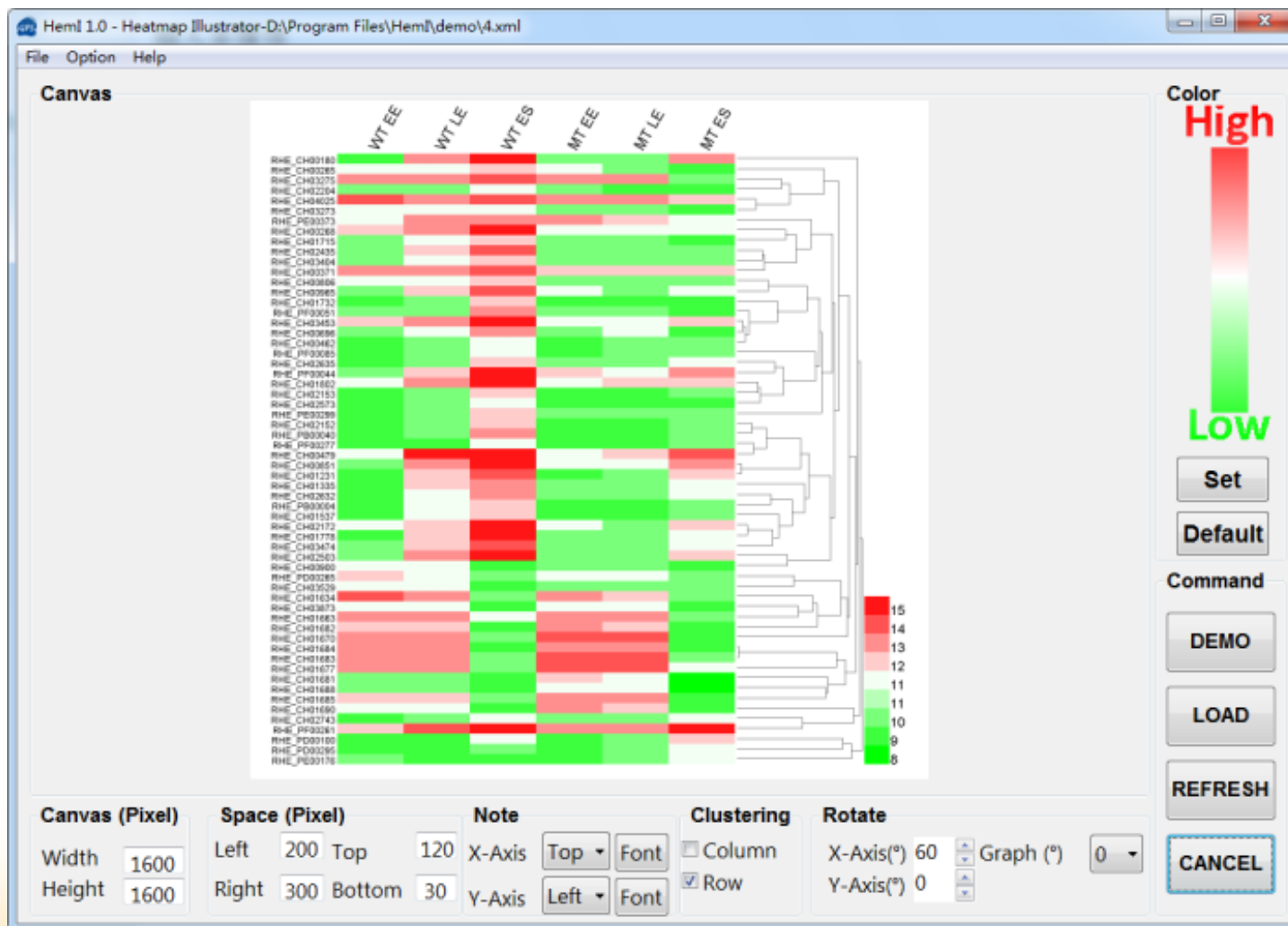


➤ <http://ibs.biocuckoo.org>





Heatmap Illustrator (Hemi)



➤ <http://hemi.biocuckoo.org>