



# iPTMnet: An Integrated Resource for Protein Post-Translational Modification Network Discovery

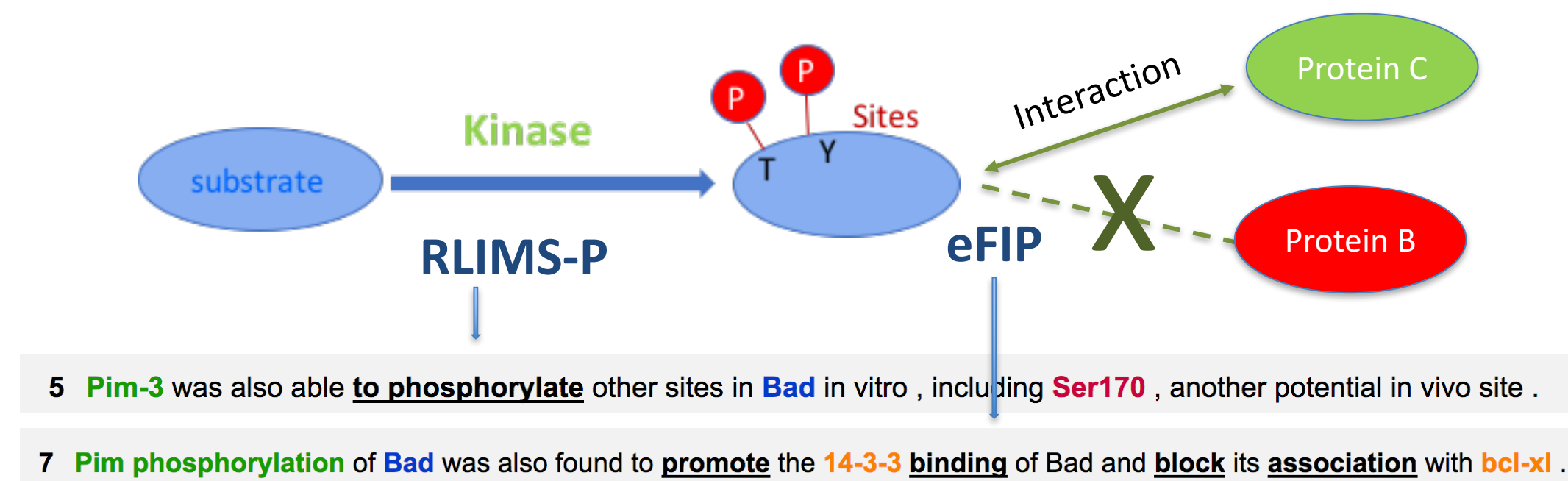
Cecilia Arighi<sup>1,2</sup>, Hongzhan Huang<sup>1,2</sup>, Karen E Ross<sup>3</sup>, Jia Ren<sup>1</sup>, Sachin Gavali<sup>1</sup>, Qinghua Wang<sup>1</sup>, Julie Cowart<sup>1</sup>, K Vijay-Shanker<sup>2</sup>, Cathy H Wu<sup>1,2,3</sup>

1 CBCB, University of Delaware; 2 CIS, University of Delaware; 3 PIR, Georgetown University Medical Center

iPTMnet for post-translational modification (PTM) knowledge discovery, employs an integrative bioinformatics approach—combining text and data mining, and ontological representation to capture rich PTM information, including PTM enzyme-substrate-site relationships, PTM-specific protein-protein interactions (PPIs), overlapping PTMs, conservation across species and sites affected in variants.

## Information Captured in iPTMnet

- Kinase-substrate-site and functional impact on protein-protein interaction from publications

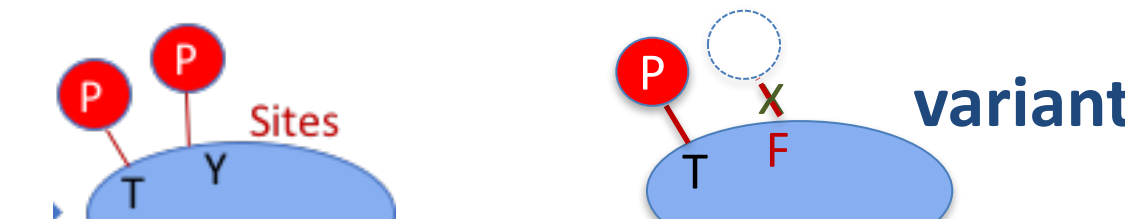


- PTM information from databases

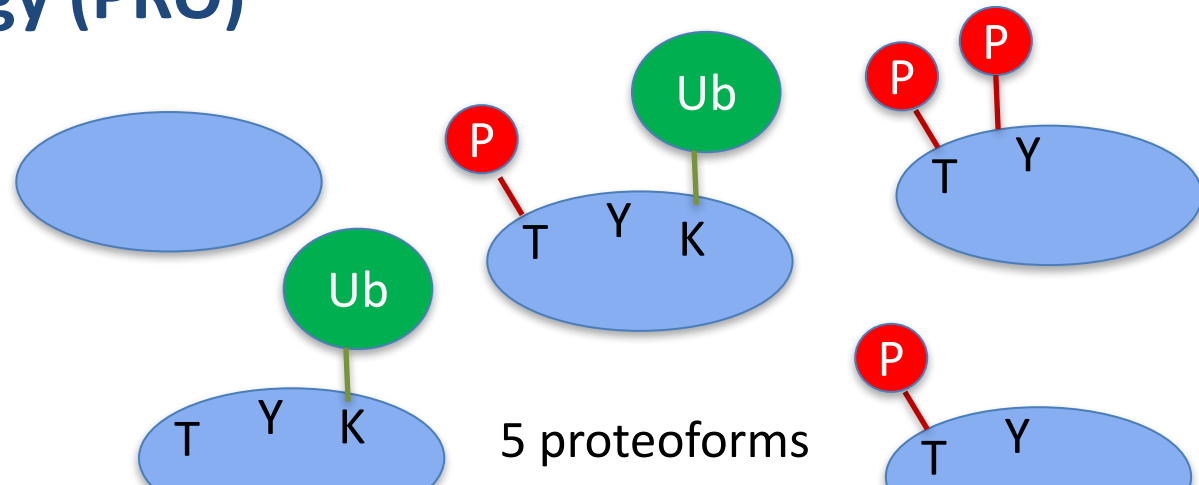
PTM-specific	Organism	Other DBs	Organism
phospho.ELM	any	HPRD	human
PhosphoSitePlus	mammals	PomBase	yeast
p3DB	plants	SGD	yeast
PhosphoGrid	yeast	neXtProt	human
RLIMS-P+	any	Signor	mammals
PhosPhAt	plants	UniProt	any
dbSNO	any	PRO	any

8 PTMs  
Phosphorylation  
Acetylation  
Methylation  
Glycosylation  
S-nitrosylation  
Sumoylation  
Ubiquitination  
Myristoylation

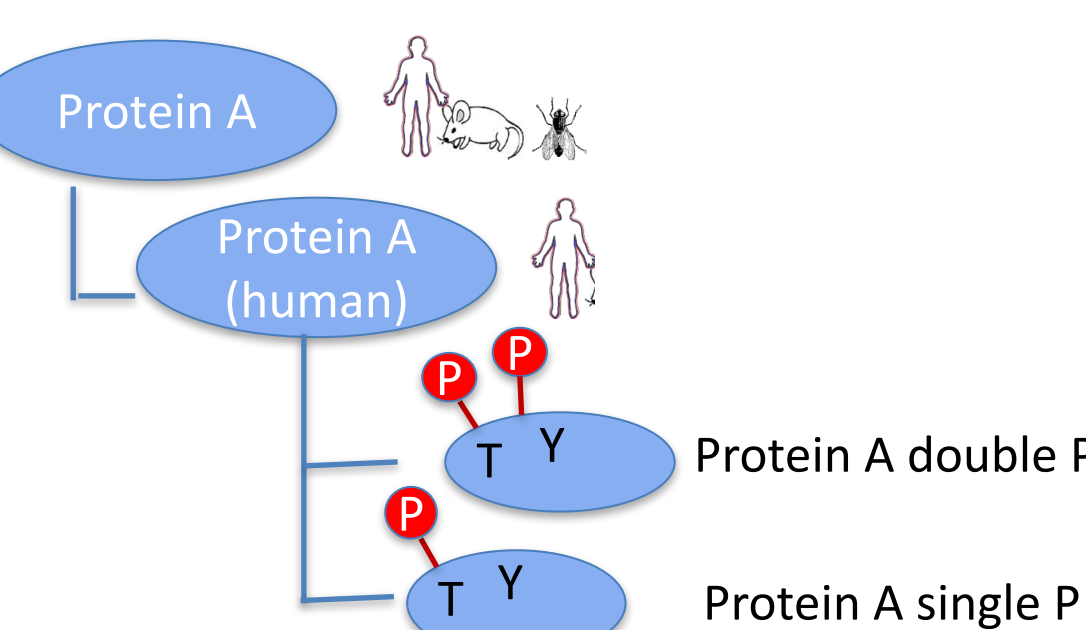
- Variant affecting PTM sites from BioMuta



- Proteoforms, combinations of PTMs observed, from Protein Ontology (PRO)



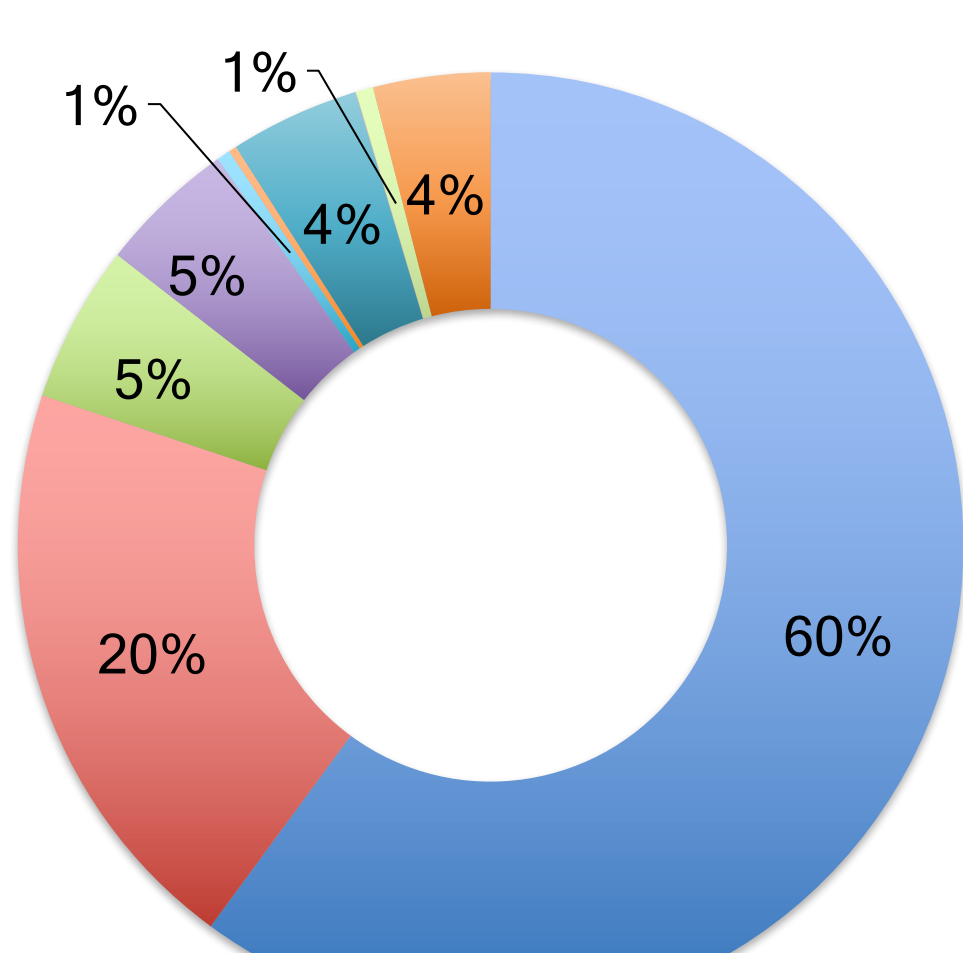
- Comparison across species from PRO



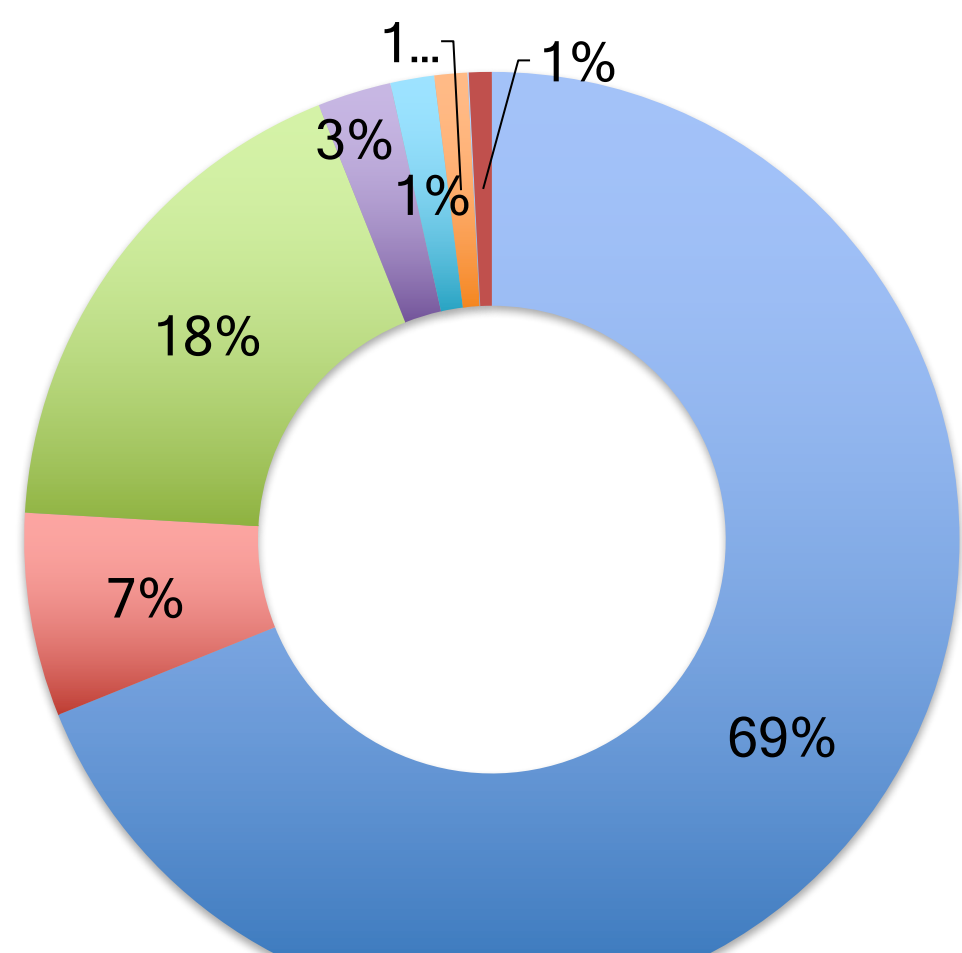
## Statistics release 5.0

Substrates (protein)	Substrates (proteoforms)	Sites	Enzymes	Enzyme-substrate pairs	Enzyme-substrate-site	PTM-dependent PPI	PMIDs	Variants
63475	12144	737070	1072	12106	23315	3031	32738	75130

Number of sites per organism



Number of sites per PTM type



## Data Confidence Score

The confidence score for each PTM information is calculated as:

$$\text{Score } S = S_q + S_n + S_p$$

**S<sub>q</sub> weights the quality of the underlying resource**

- S<sub>q</sub> = 2, data from curated resources supporting correct species assignment, 1, data from other curated resources, 0, data from automatic text mining;

**S<sub>n</sub> gives weight to multiple sources**

- S<sub>n</sub> = 1, data from multiple resources, 0, data from single resource;

**S<sub>p</sub> gives weight to the number of publications**

- S<sub>p</sub> = 1, data supported by multiple papers and at least one is not considered a large scale paper (LSP), 0, data supported by one non-LSP paper, -1, only LSP or no literature evidence.

Score

Evidence Source

Site	PTM Type	PTM Enzyme	Score	Source	PMID
S134	Phosphorylation	P08750 (Pim3), P18653 (Pdp4a1), P01532 (Pheaa), C08643 (Pka1), Q85838 (Pdp4b1), P31750 (Akt1), P06803 (Pim1)	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...
S136	Phosphorylation	P08750 (Pim3), P18653 (Pdp4a1), P01532 (Pheaa), C08643 (Pka1), Q85838 (Pdp4b1), P31750 (Akt1), P06803 (Pim1)	★★★★	phospho.ELM, PhosphoSitePlus, RLIMS-P, UniProt	10949026, 10611223, 12087097, ...

The Score ranges from 0 to 4 gold stars, with 4 being most confident

## Conclusion

iPTMnet connects PTM proteoforms with their modifying enzymes and multiple coordinated PTMs across taxa, thereby unifying fragmented PTM information into a biologically meaningful context for visual and systematic PTM knowledge discovery.

The work presented in this poster is supported by National Institutes of Health U01GM120953

## iPTMnet Website (<http://proteininformationresource.org/iPTMnet>)

### 1-Search

Use UniProtKB AC/ID, protein or gene name, PRO ID, and PMID. You can restrict search by PTM type (1), enzyme or substrate role (2), and/or organism (3)

Search for proteins in iPTMnet database

PTM Type: 1, Enzyme or Substrate: 2, Select organisms: 3

Search results for AKT1 in iPTMnet database

PTM ID	Protein Name	Gene Name	Enzyme	Substrate Site	Enzyme Site	PTM Enzyme	Site	Isotype
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1

The result page returns candidate entries matching the query ordered by the overall count of all PTM data in descending order

### 2-Batch Retrieval

Allows to obtain PTM enzyme or PTM dependent PPI information for up to 500 PTM sites at a time

iPTMnet Batch Retrieval

Search for proteins in iPTMnet database

Batch retrieval of PTM information for AKT1

PTM ID	Protein Name	Gene Name	Enzyme	Substrate Site	Enzyme Site	PTM Enzyme	Site	Isotype
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1
P01532	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1	AKT1

MYC PTM site table can be filtered for overlapping PTMs

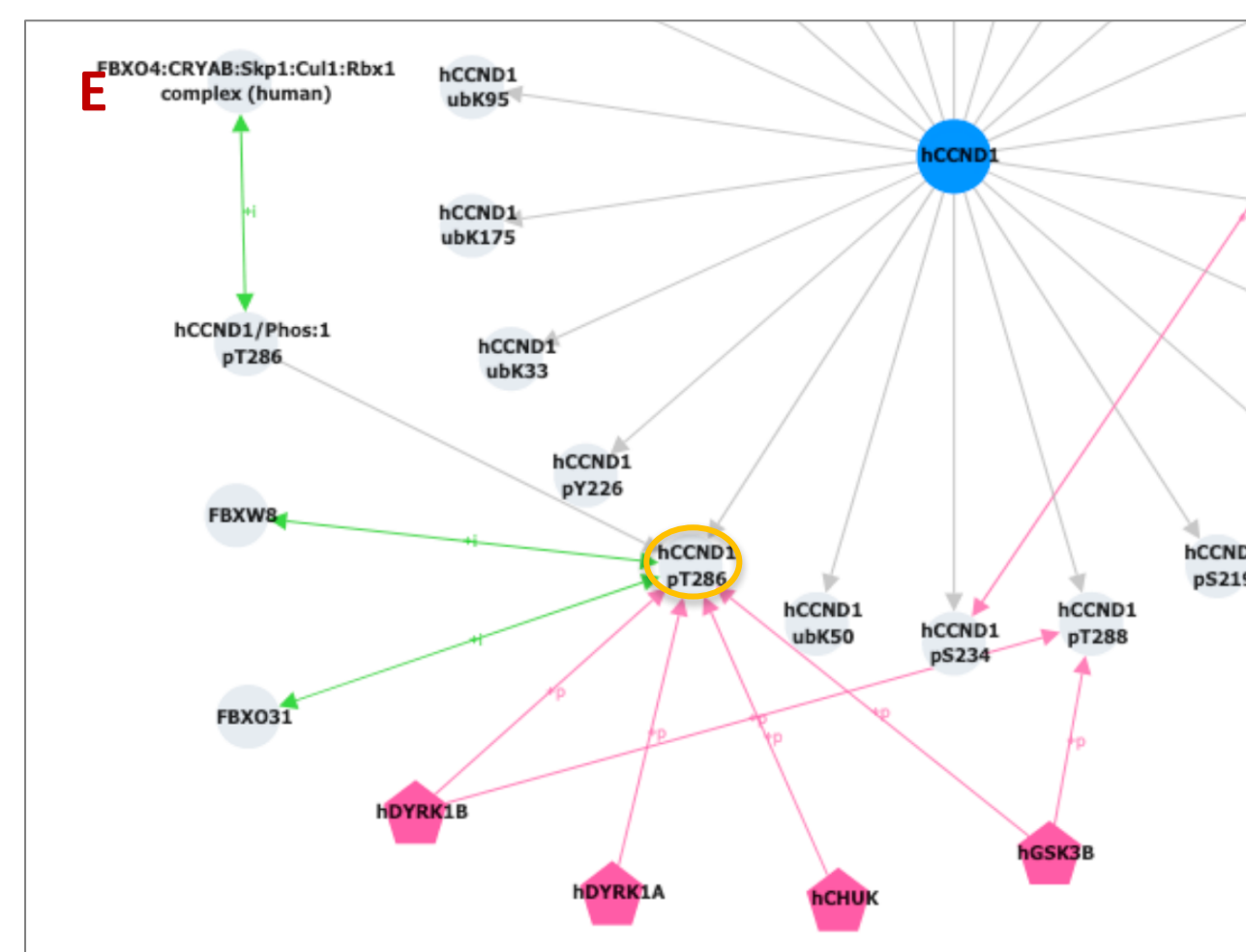
PTM ID	PTM Type	PTM Enzyme	Score	Source	PMID
P01106	Phosphorylation	P01106-1	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...
P01106-1	Phosphorylation	P01106-1	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...
P01106-2	Phosphorylation	P01106-2	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...

### 4b-Cross species comparison and functional interpretation of variants

Deregulated cyclin D1 degradation appears to be responsible for the increased levels of cyclin D1 in several cancers. iPTMnet can assist in connecting PTM sites affected in cancer-related variants and their functional consequence. T286 is conserved across vertebrates (A), and it is phosphorylated (B). This phosphorylation is necessary for subsequent degradation by the proteasome. The PPI section shows phospho-dependent interactions with proteasome proteins (C). The variants found in Biomuta for T286 are found in uterine cancer samples (D).

Variants at T286 would avoid proteasomal degradation, as lack of phosphorylation at position 286 would decrease or abolish interaction with proteasomal proteins.

The Cytoscape representation in (E) shows the phosphorylation site T286 along with kinases and PPI partners.



### 3-Entry

General protein information from UniProt and PRO

Sequence with color-coded sites based on PTM types

Site-centric list. PTM sites with confidence and evidence

PTM enzymes have an enzyme table with their substrates and sites

List of proteoforms for the entry as described in PRO.

Impact of phosphorylation on PPIs from text mining

PPIs for proteoforms from PRO

Variants affecting PTM sites

### 4-Use Cases

#### 4a-Finding overlapping PTM sites

Sites subjected to multiple PTMs are potential regulatory sites. Use Browse to find all proteins with overlapping PTM sites in an organism

Example: T-58 in human MYC is phosphorylated or O-glycosylated

<https://research.bioinformatics.udel.edu/iptmnet/entry/P01106>

Overlapping PTMs sites are shown in yellow in alignment

PTM ID	PTM Type	PTM Enzyme	Score	Source	PMID
P01106	Phosphorylation	P01106-1	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...
P01106-1	Phosphorylation	P01106-1	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...
P01106-2	Phosphorylation	P01106-2	★★★★	PhosphoSitePlus	10949026, 10611223, 12087097, ...

Multiple Sequence Alignment

Alignment of PTM sites across different species

Impact of phosphorylation on PPIs

PPIs for proteoforms from PRO

PTM sites affected in variants