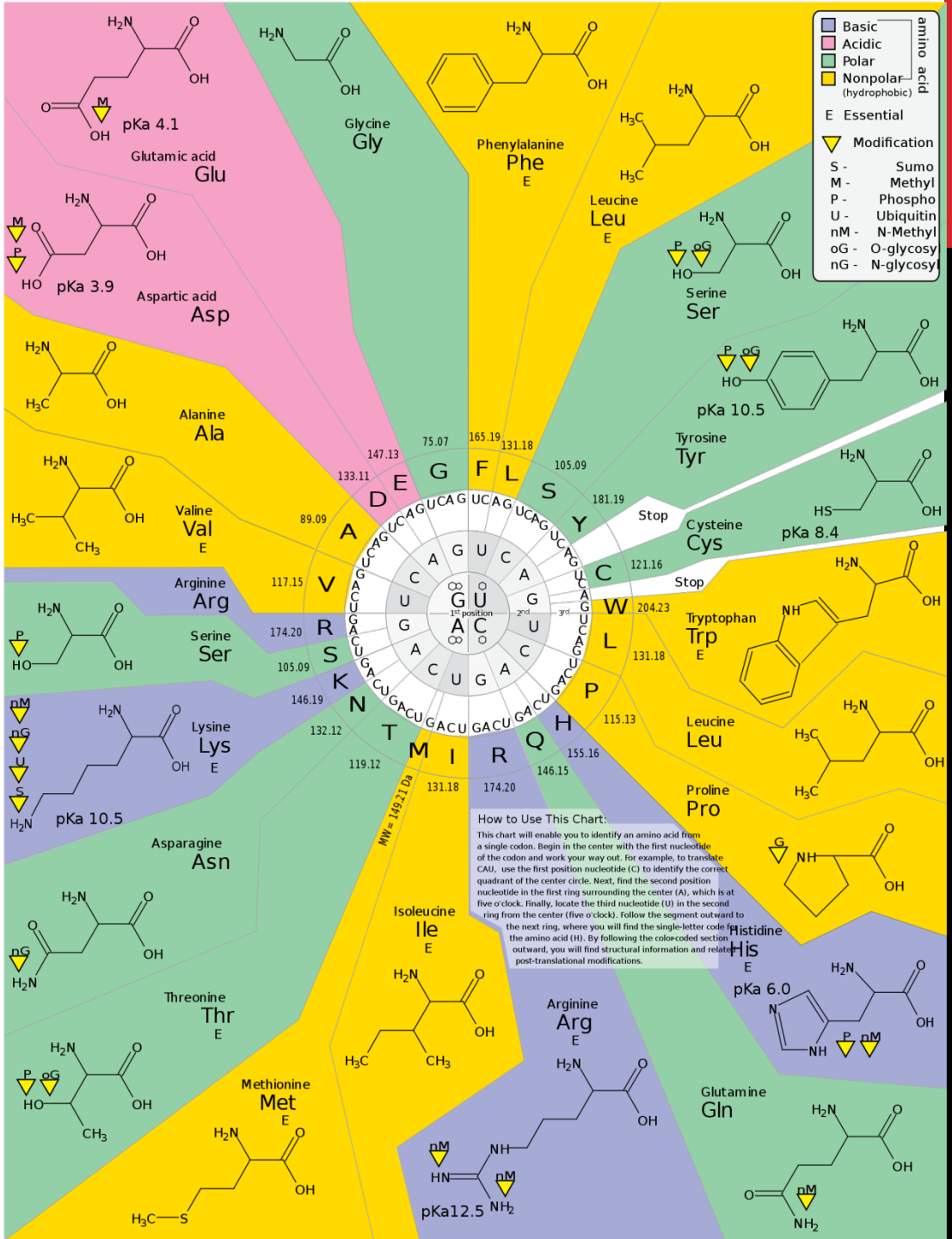
SYSTEMATIC FUNCTIONAL PRIORITIZATION OF PROTEIN POSTTRANSLATIONAL MODIFICATIONS

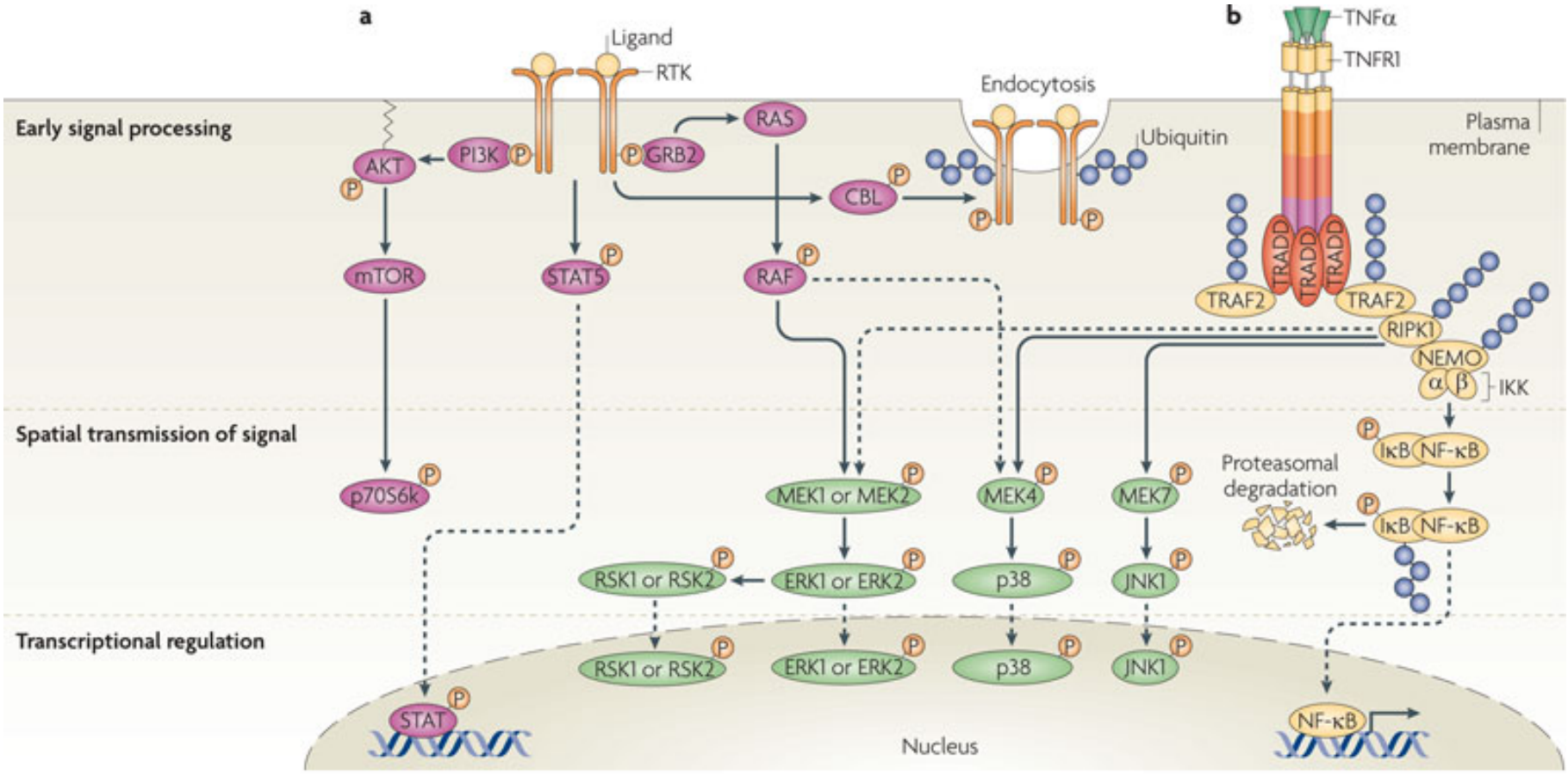
TECHNICAL JOURNAL CLUB – TUE, 23TH JULY 2013

KARL FRONTZEK, INSTITUTE OF NEUROPATHOLOGY

PROTEIN POSTTRANSLATIONAL MODIFICATIONS (PTM)

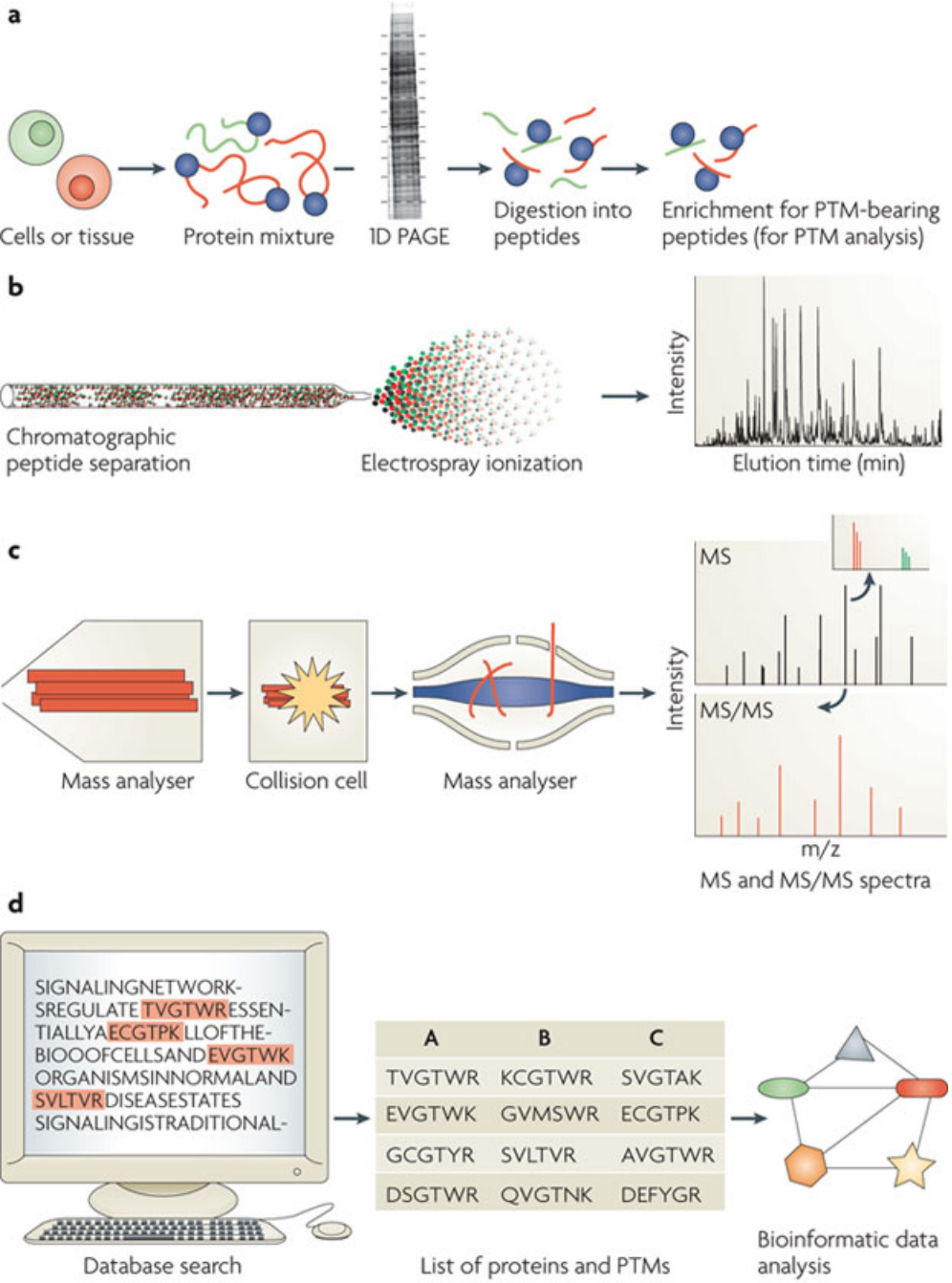


Gramatikoff K. in Abgent Catalog (2004-5) p.263



Nature Reviews | Molecular Cell Biology

Choudhary&Mann, Nat Rev Mol Cell Biol 2010



Nature Reviews | Molecular Cell Biology

PTM	Mass shift (Δm; Da)*	Enrichment methods	Largest MS study‡
Phosphorylation	79.96633	IMAC, TiO_2 and antibodies	20,443 sites[54]
Acetylation	42.01056	Pan anti-acetyl-Lys antibodies	3,600 sites[50]
Ubiquitylation (diGly tag)	114.04292	Tagged ubiquitin	110 sites[59]
Methylation	14.01565	Anti-methyl-Lys or anti-methyl-Arg antibodies	59 sites[55]
o-GlcNac	203.07937	Lectin	141 sites[124]

Choudhary&Mann, Nat Rev Mol Cell Biol 2010

EVOLUTION OF PHOSPHOSITES

- **phosphosites were shown to be under evolutionary constraint because they are having key roles in protein function**

  Boekhorst et al. Genome Biol 2008

  Gnad et al., Genome Biol 2007

- **but at the same time a lot of phosphosites were found with a more rapid evolutionary turnover or unknown function (i.e. unspecific phosphorylation <> non-functional phosphorylation)**

  Ubersax&Ferrell Jr, Nat Rev Mol Cell Biol 2007

  Lienhard, Trends Biochem Sci 2008

  Malik et al., Bioinformatics 2008

Genome Analysis

Weak functional constraints on phosphoproteomes

Christian R. Landry*, Emmanuel D. Levy* and Stephen W. Michnick

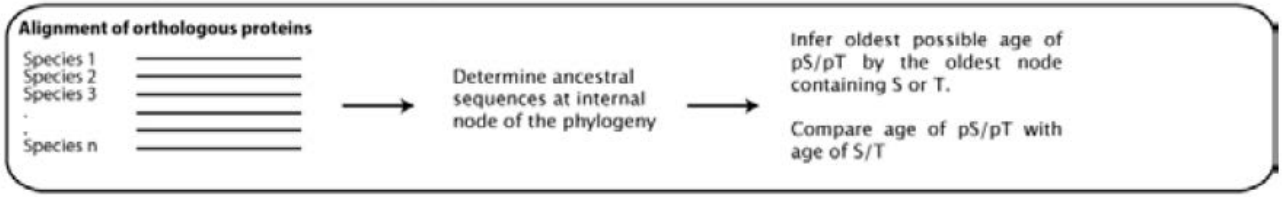
Trends in Genetics
Volume 25, Issue 5, May 2009, Pages 193–197

Scientific question:
Does the phosphoproteome (here pS/pT) have a slower evolutionary turnover than the non-phosphorylated proteome (here S/T)?

1. Mapping phosphosites on proteomes

Phosphosites were compiled from several phosphoproteomic experiments and databases (Table S1)

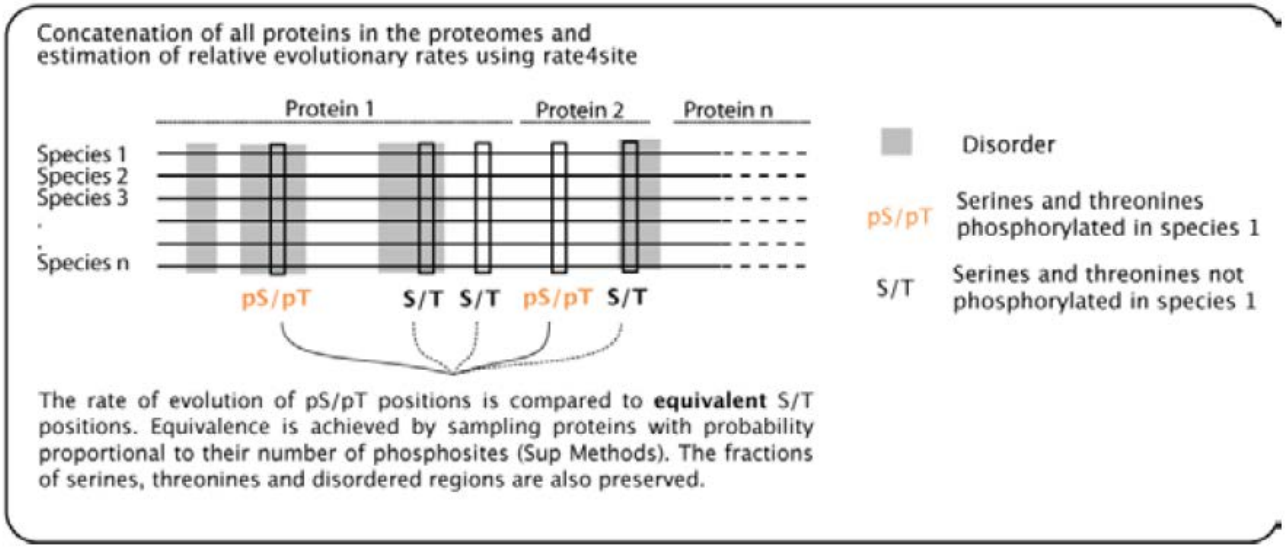
2. Reconstructing ancestral sequences by ML

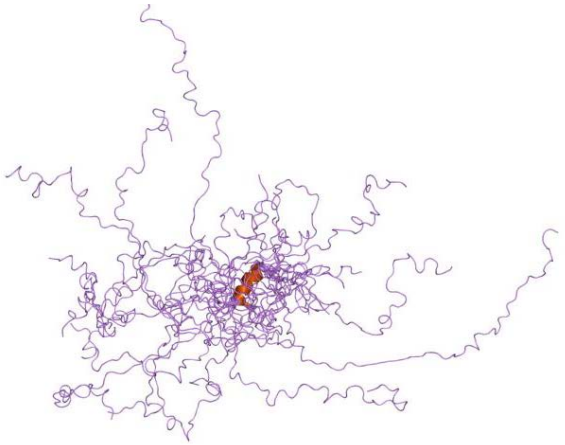


3. Mapping unstructured regions on proteomes

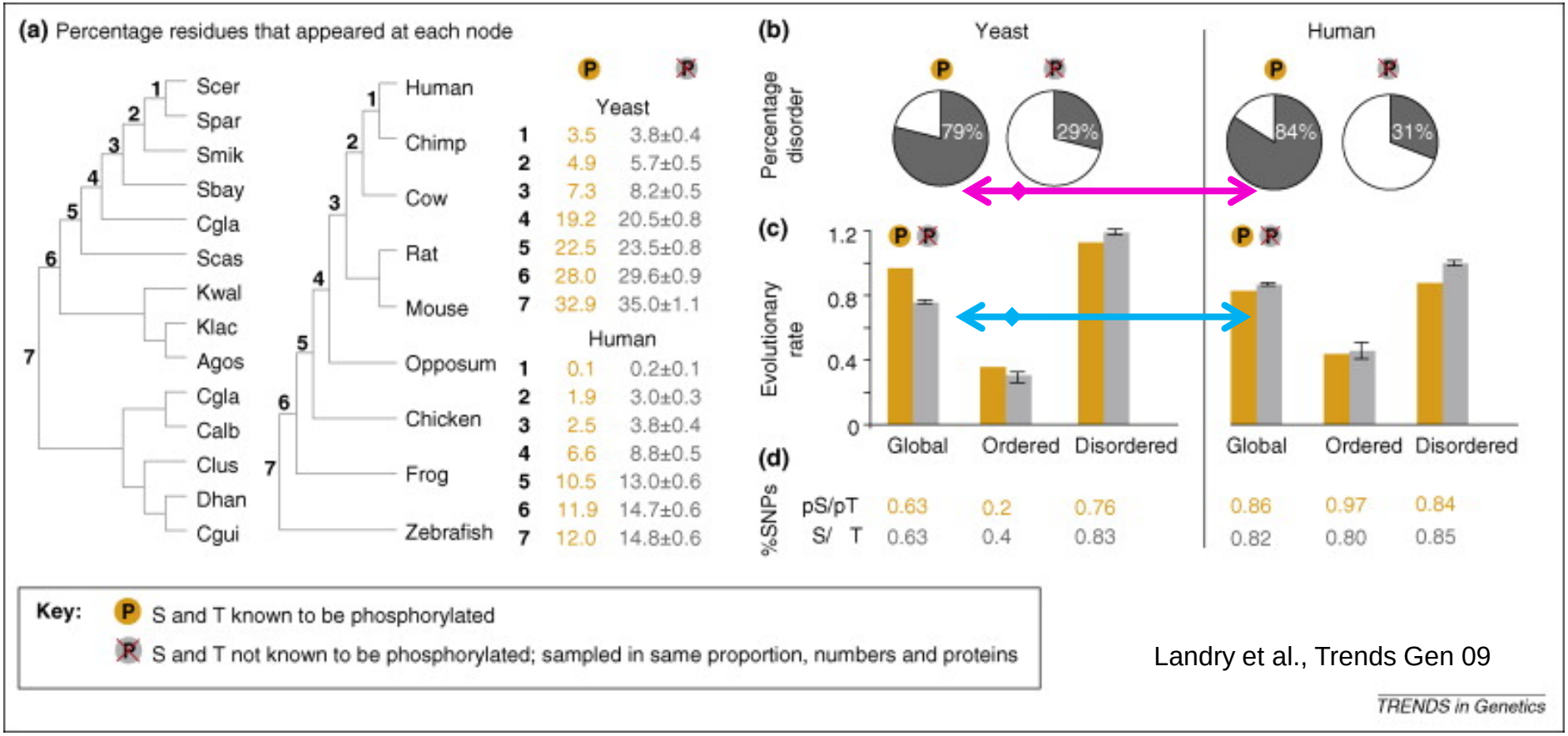
Disorder was predicted using DISOPRED

4. Compute the relative evolutionary rates of all residues in the proteomes





Song et al. Biochemistry 2006

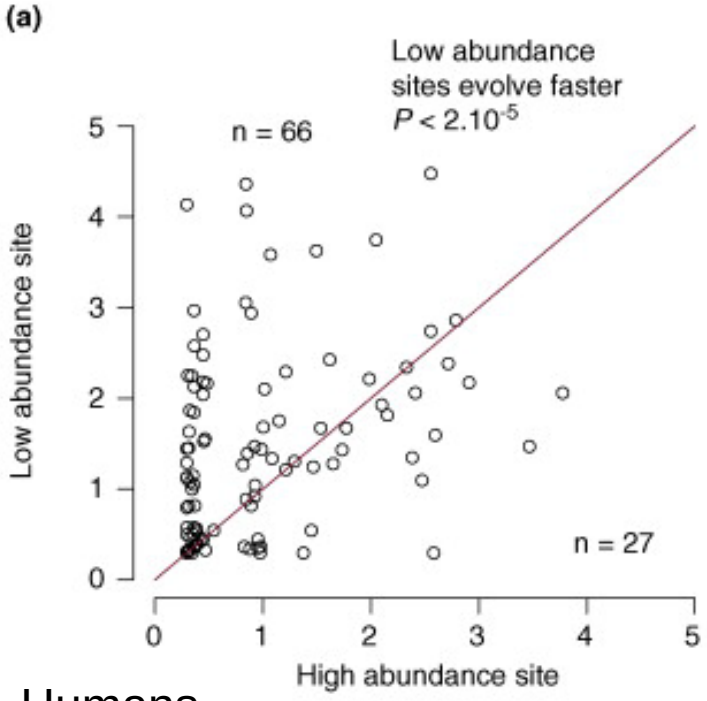


Landry et al., Trends Gen 09

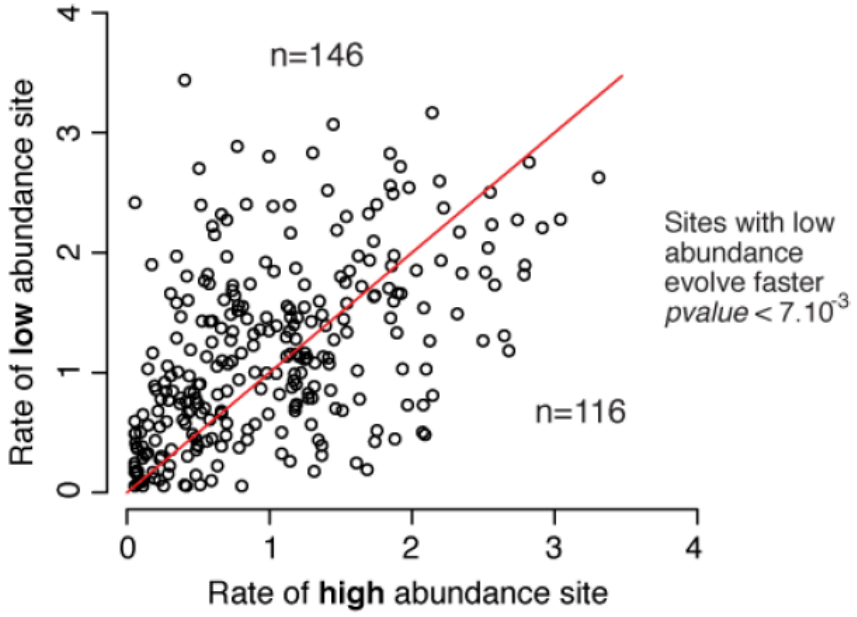
TRENDS in Genetics

HOW MANY NON-FUNCTIONAL PHOSPHOSITES ARE THERE? #1

1st assumption: non-functional phosphorylations are likely to represent off-target interactions >> rare molecular events >> should have lower abundance than (stochiometrically) higher phosphorylated sites



Humans



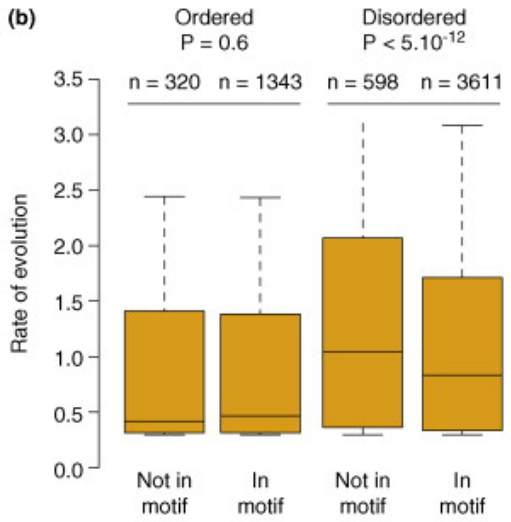
Yeast

Landry et al., Trends Gen 09

HOW MANY NON-FUNCTIONAL PHOSPHOSITES ARE THERE? #2

2nd assumption: if rapidly evolving sites results from non-functional (supposedly non-specific) phosphorylation events

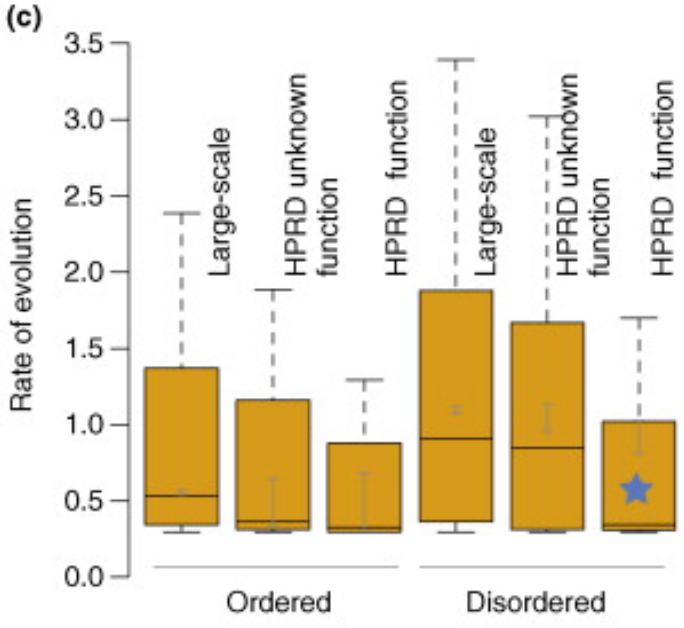
>> underrepresentation in common protein kinase recognition motifs



Humans

HOW MANY NON-FUNCTIONAL PHOSPHOSITES ARE THERE? #3

Last assumption: phosphosites with assigned function (e.g. site-directed mutagenesis and functional assay) evolve slower than those without assigned function



Humans

CONCLUSIONS

1) **Phosphoproteomes evolve at a similar rate to that of non-phosphorylated residues**

   Possible explanations:

   - Most phosphosites occur in disordered regions and these evolve rapidly
   - The experimental setups used to charactarized those site are highly sensitive and detect a fraction of non-functional sites

2) **For assessment of potentially meaningful (i.e. functional) phosphosites, more information about the protein should be taken into consideration, like**

   - Kinase recognition motifs
   - Abundance of phosphosites

COMPUTATIONAL BIOLOGY

Comparative Analysis Reveals Conserved Protein Phosphorylation Networks Implicated in Multiple Diseases

Science Signaling
AAAS

Chris Soon Heng Tan,[1,2*] Bernd Bodenmiller,[3*] Adrian Pasculescu,[1] Marko Jovanovic,[4] Michael O. Hengartner,[4] Claus Jørgensen,[1] Gary D. Bader,[1,2] Ruedi Aebersold,[3,5,6,7] Tony Pawson,[1,2] Rune Linding[8†]

(Published 28 July 2009; Volume 2 Issue 81 ra39)

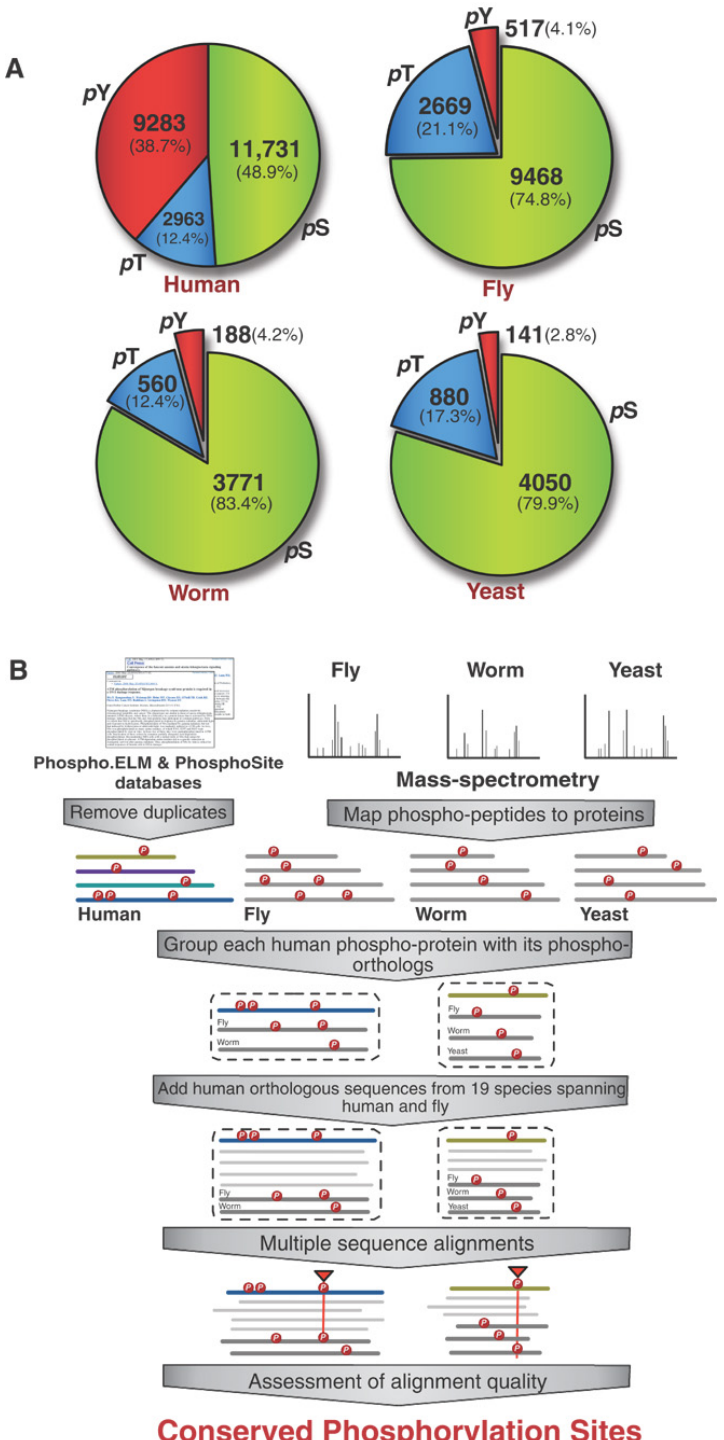
Analysis of conservation of phosphoproteomes from high-&low-throughput mass-spectrometry in yeast (S. cerevisiae), fly (D. melanogaster) and worm (C. elegans) with human reference set

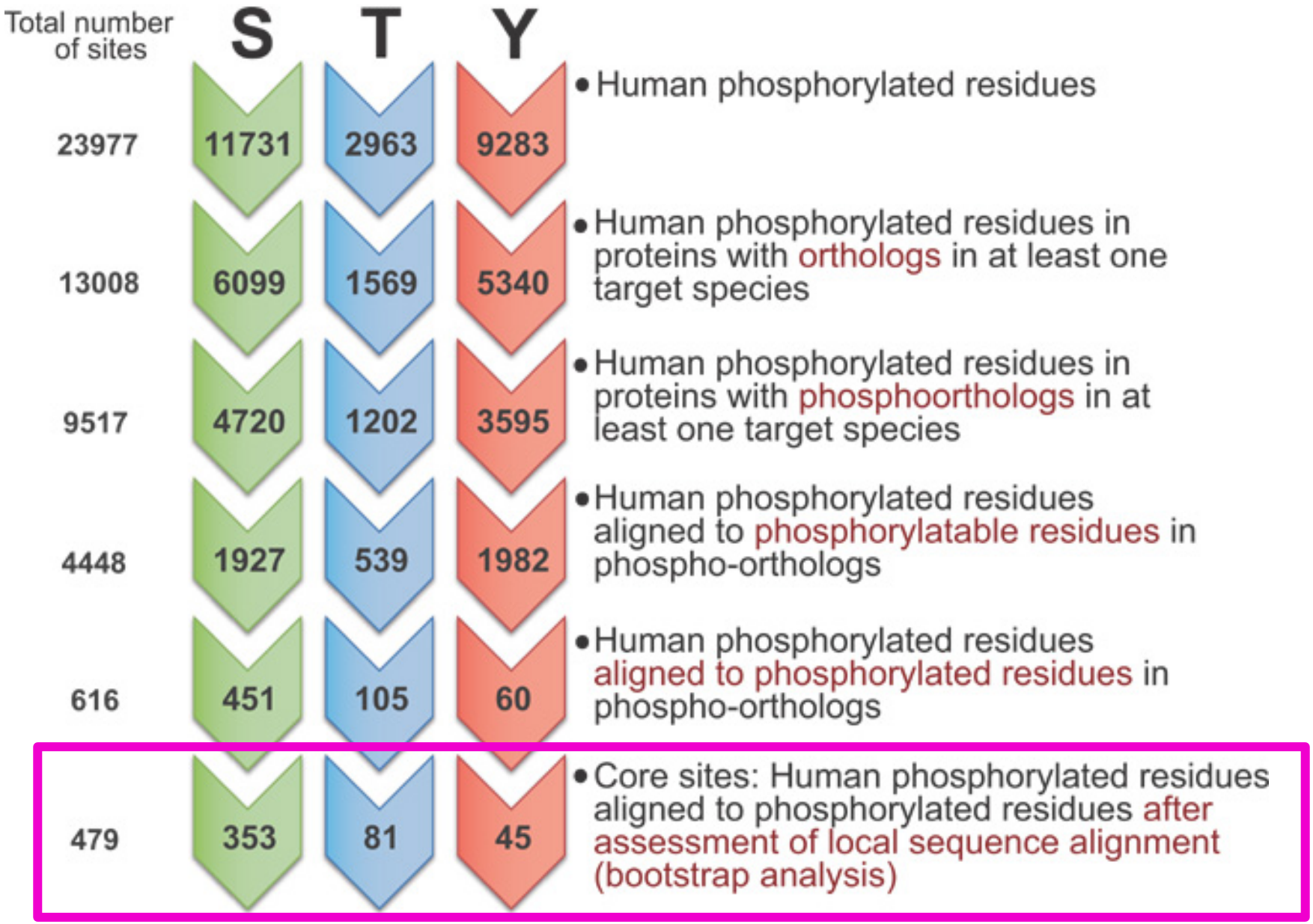
positionally conserved sites

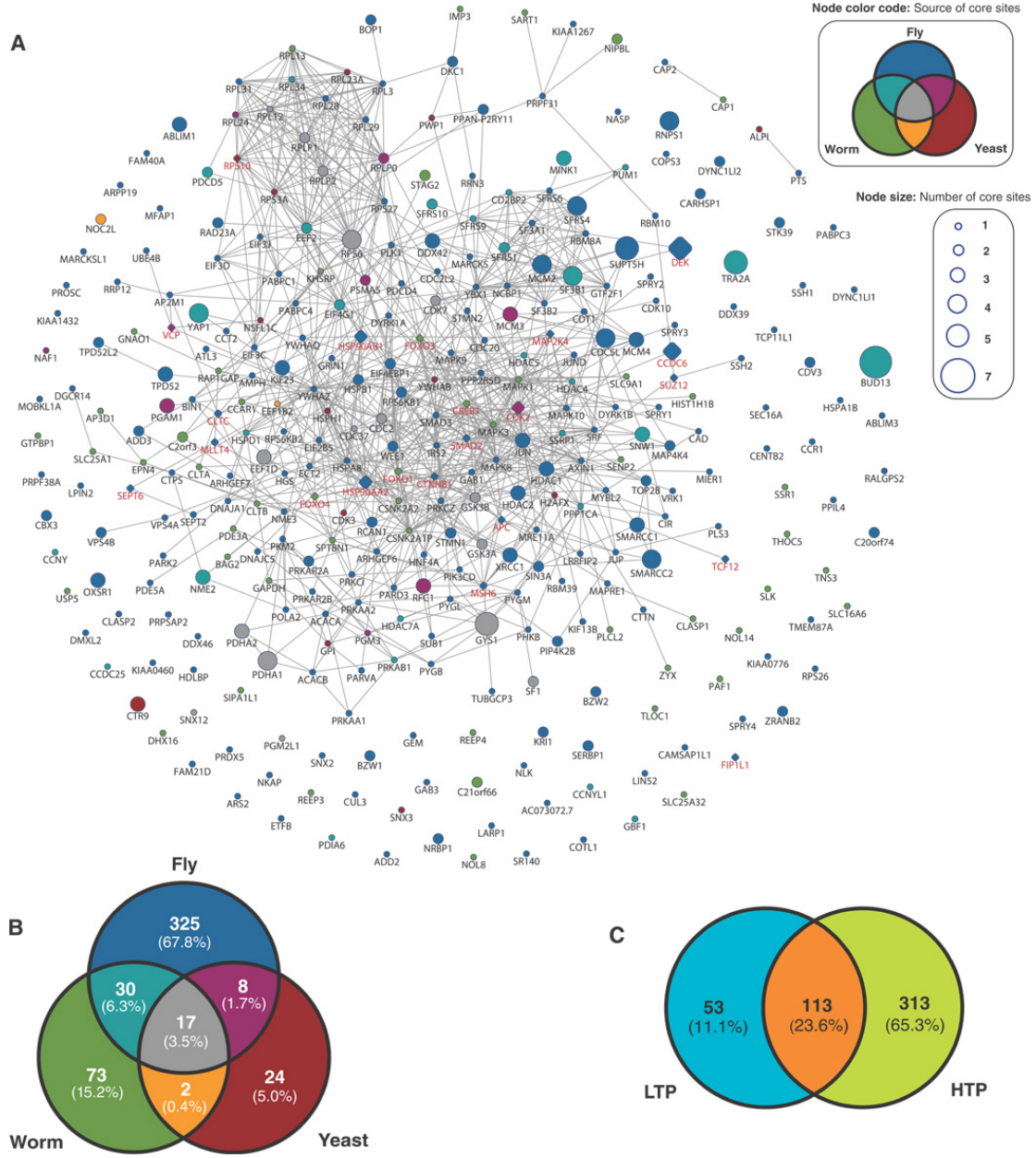
conserved kinase-substrate interactions (not necessarily positionally conserved)

Workflow of the study

(similar MS and computational alignment pipelines throughout the groups)




Total number of sites
S
T
Y
23977
11731
2963
9283
• Human phosphorylated residues
13008
6099
1569
5340
• Human phosphorylated residues in proteins with orthologs in at least one target species
9517
4720
1202
3595
• Human phosphorylated residues in proteins with phosphoorthologs in at least one target species
4448
1927
539
1982
• Human phosphorylated residues aligned to phosphorylatable residues in phospho-orthologs
616
451
105
60
• Human phosphorylated residues aligned to phosphorylated residues in phospho-orthologs
479
353
81
45
• Core sites: Human phosphorylated residues aligned to phosphorylated residues after assessment of local sequence alignment (bootstrap analysis)

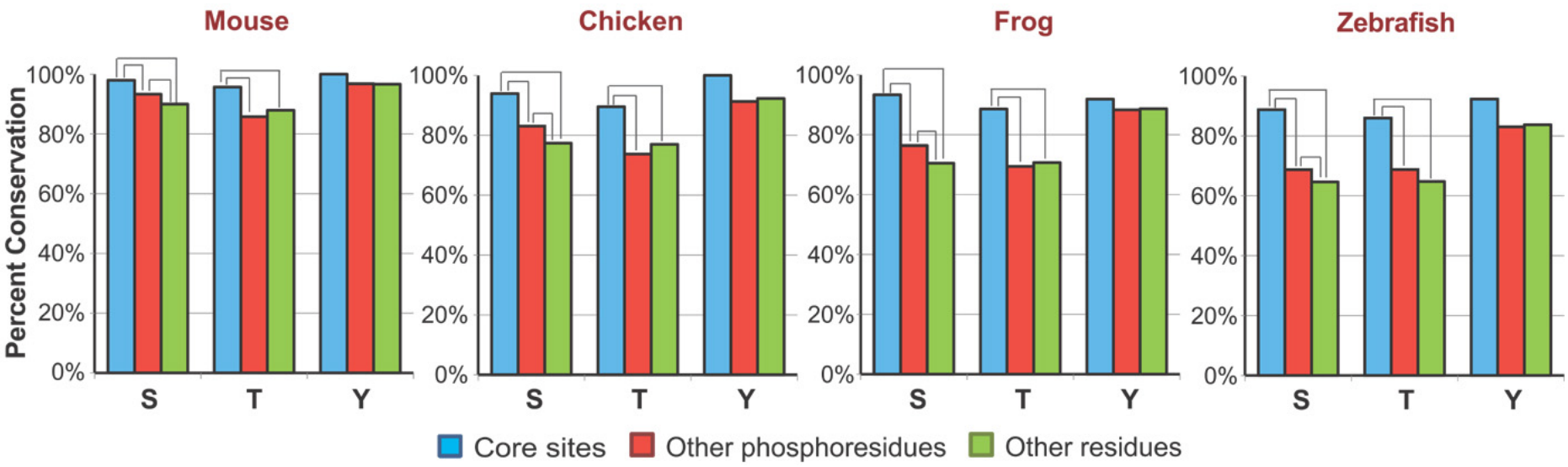


Conservation between compared species and human reference set, of all sites in 75 protein domain families:

- 57 at least in 2 target species
- 17 in all 3 target species

For kinase specificity, CDK1+CDK2 as well as casein kinase 2 were amongst most frequently predicted kinases

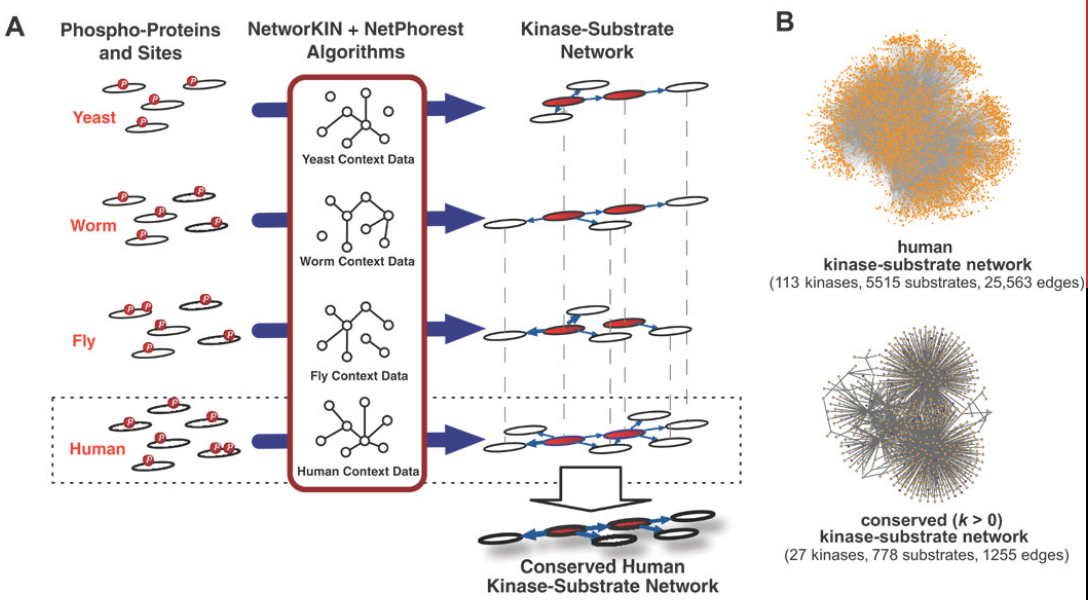
SEQUENCE HOMOLOGY IN OTHER SPECIES

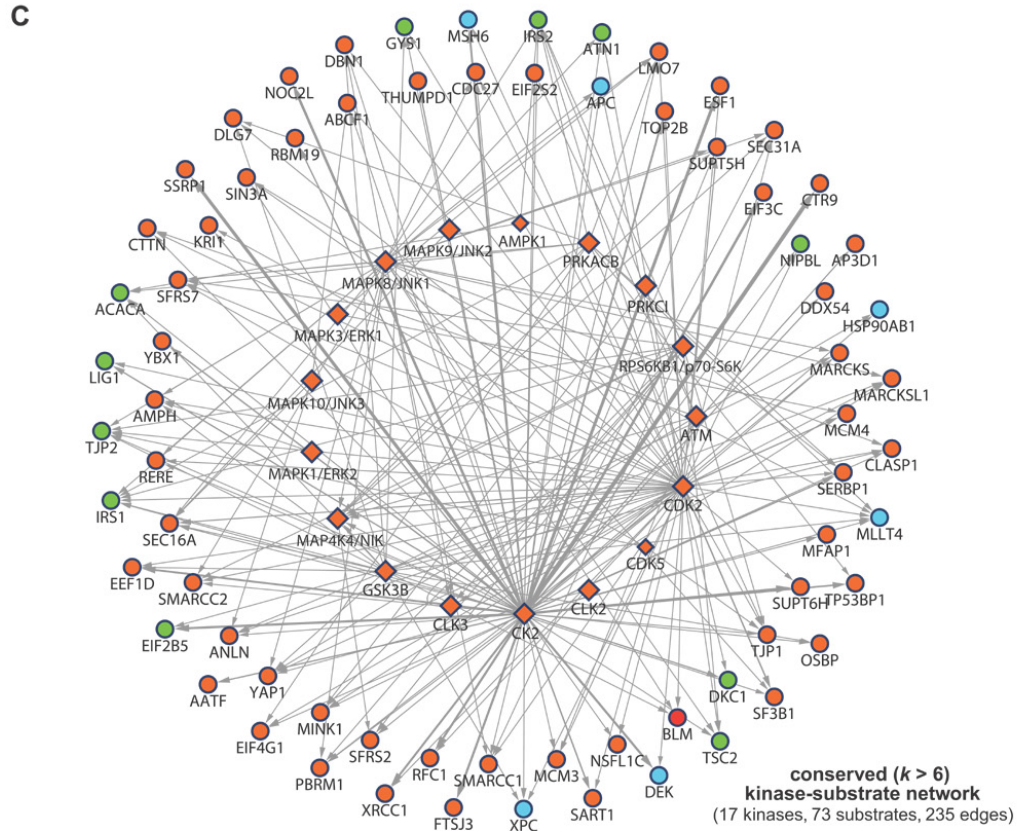


>> High rates of conservation concerning phosphorylated sequence motifs across species

CREATING A PHOSPHORYLATION-INTERACTION-NETWORK

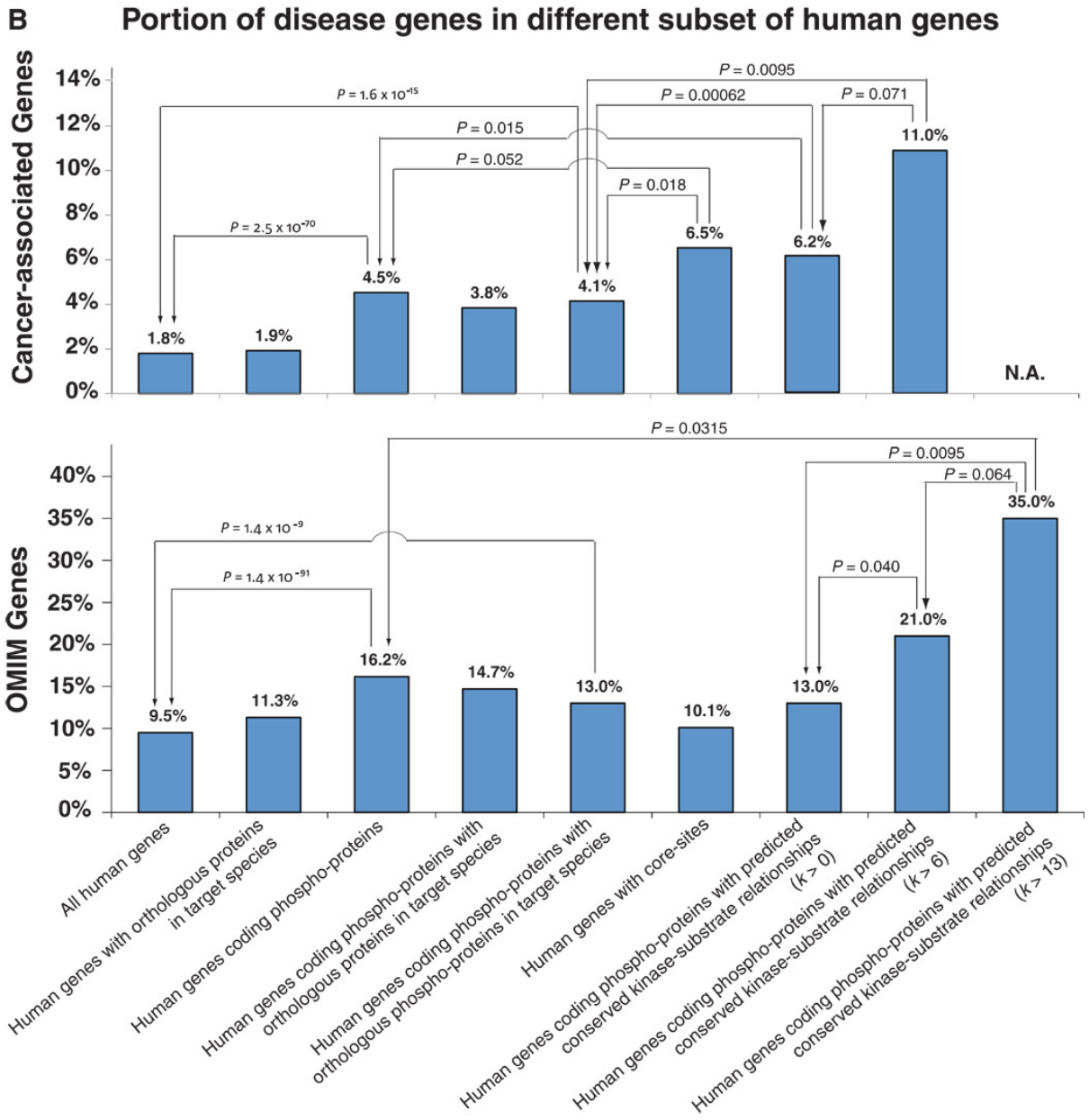
Randomized network analysis showed that those interactions were unlikely to have been occuring by chance





ENRICHED CORE-SITES/CONSERVED KINASE-SUBSTRATE RELATIONSHIPS IN CANCER GENES?

Enriched genes (amongst others)
FOXO1,3,4
CREB1
SMAD2
HSP90



CONCLUSIONS

- ***Pro*: Comparing phosphorylation patterns between humans (reference set) and multiple target species provides clues to disregulated phosphorylation hubs in cancer and other human diseases**
- ***Contra*: no functional validation provided**

Resource

Cell

Systematic Functional Prioritization of Protein Posttranslational Modifications

Pedro Beltrao,[1,3,*] Véronique Albanèse,[4] Lillian R. Kenner,[1,3] Danielle L. Swaney,[5] Alma Burlingame,[2,3] Judit Villén,[5] Wendell A. Lim,[1,3,6] James S. Fraser,[1,3] Judith Frydman,[4] and Nevan J. Krogan[1,3,7,*]

[1]Department of Cellular and Molecular Pharmacology

[2]Department of Pharmaceutical Chemistry

University of California, San Francisco, San Francisco, CA 94107, USA

[3]California Institute for Quantitative Biosciences, QB3, San Francisco, CA 94107, USA

[4]BioX Program, Biology Department, Clark Center, Stanford, CA 94305, USA

[5]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

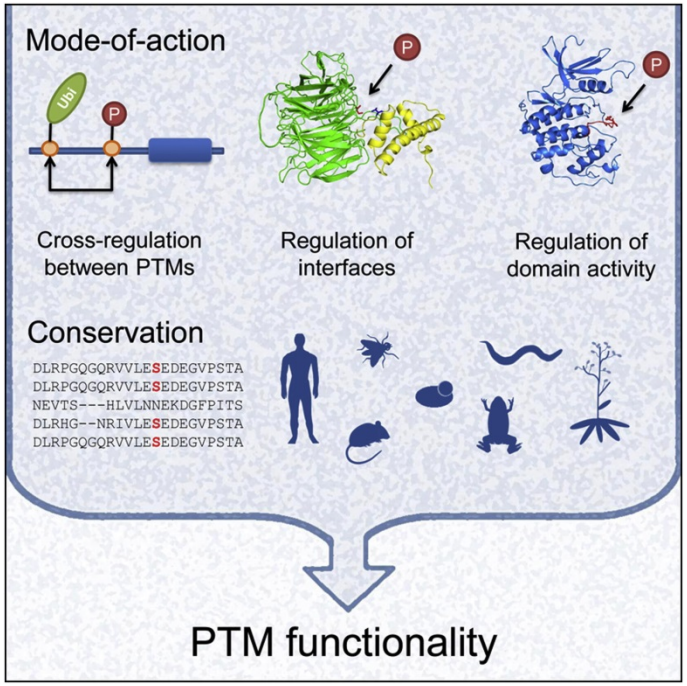
[6]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA

[7]J. David Gladstone Institute, San Francisco, CA 94158, USA

Scientific question: How to refine and identify functionally relevant protein posttranslational modifications (PTM) – «most significant bottleneck in proteomic studies of posttranslational modification»

SYSTEMATIC FUNCTIONAL PRIORITIZATION OF PROTEIN POSTTRANSLATIONAL MODIFICATIONS

- ***Compilation of ~ 200.000 phosphorylation, acetylation and ubiquitination sites from 11 eukaryotic species (incl. H. sapiens and M. musculus)***
- ***Experimental determination of ~ 2.500 ubiquitination sites for S. cerevisiae***



Beltrao et al., Cell 2012

DATA SETS

- Phosphorylation data set

3 fungi (*S. cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans*)

2 plant species (*Arabidopsis thaliana* and *Oryza sativa*)

3 mammals (*Homo sapiens*, *Mus musculus*, and *Rattus norvegicus*)

+ *Xenopus laevis*, *Drosophila melanogaster*, and *Caenorhabditis elegans*

- 13,133 lysine acetylation sites

H. sapiens, *M. musculus* and *Drosophila melanogaster*

- 22,000 human ubiquitylation sites

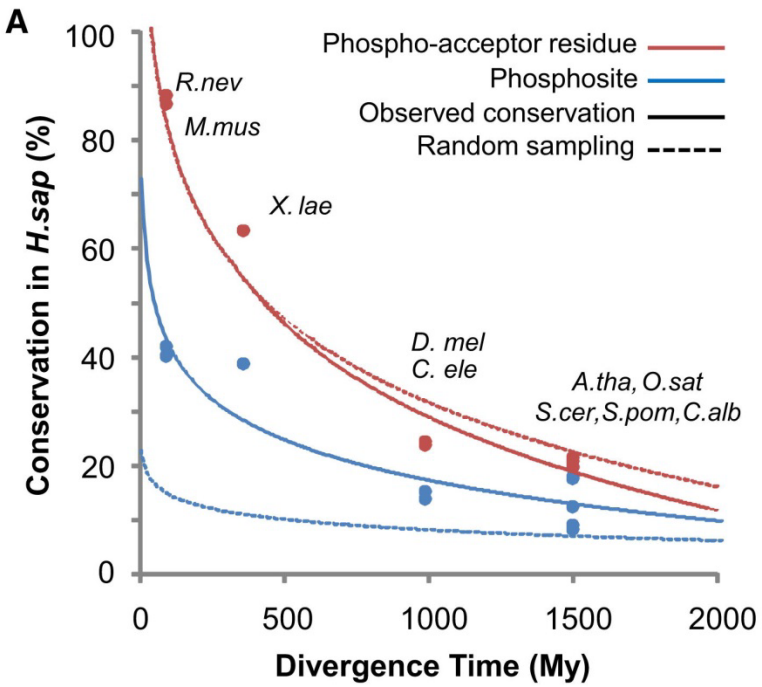
- MS approach to experimentally determine 2,500 ubiquitylation sites in S. cerevisiae to facilitate comparative studies.

Using a set of 12 different *S. cerevisiae* phosphoproteomics experiments, estimated false discovery rate < 4% of false-positive sites

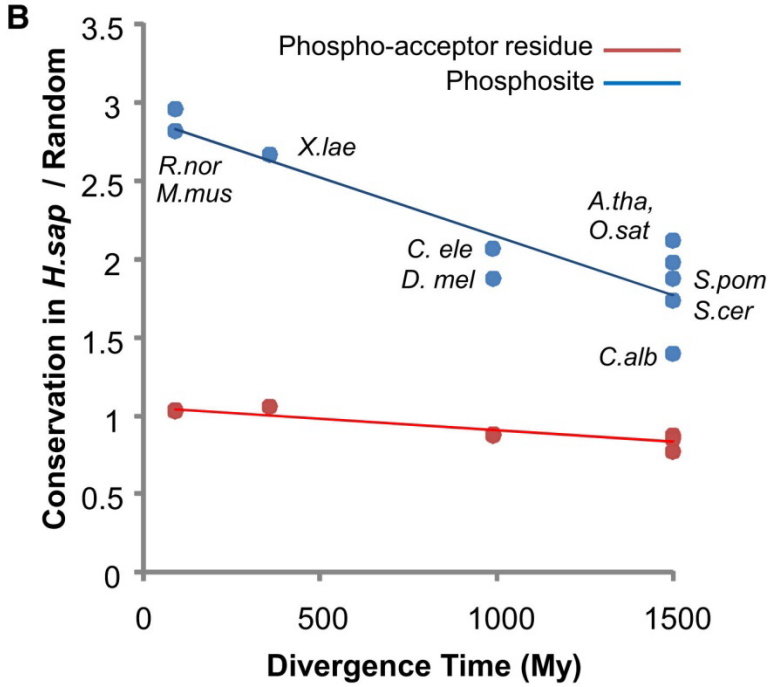
Human data set = reference data set

PHOSPHORYLATION DATA SET

observed conservation

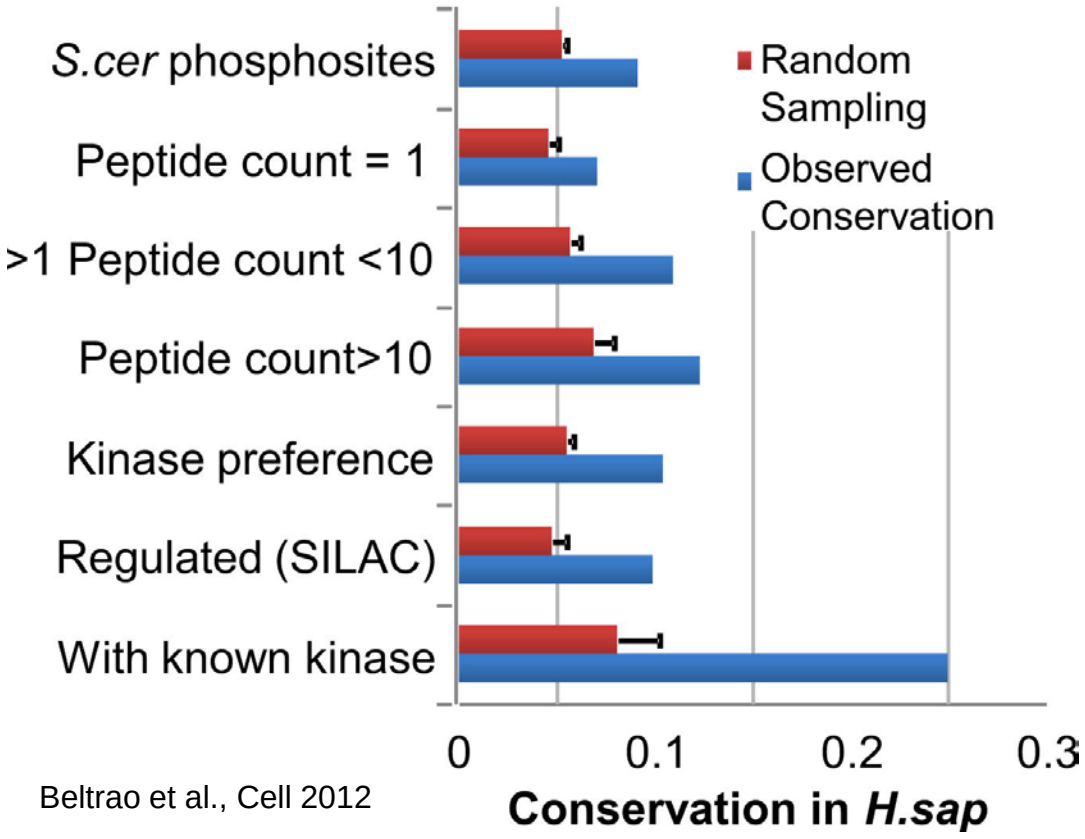


ratio observed/random conservation



Beltrao et al., Cell 2012

OBSERVED VS EXPECTED PHOSPHOSITES



Beltrao et al., Cell 2012

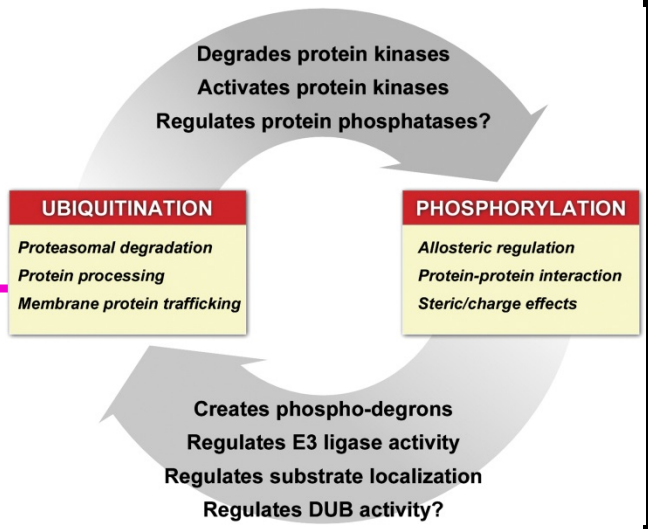
Conclusion: more functional assignments to PTMs are needed to improve data quality

error bars = 1 s.d.

FUNCTIONALITY OF PTMs IN UNSTRUCTURED DOMAINS

Species	PTM type	PTM total	Within PFAM domains	% Outside PFAM domain
H. sapiens	Phosphorylation	31165	11726	62.4
	Acetylation	8042	4604	42.8
	Ubiquitylation	22057	11079	49.7
M. musculus	Phosphorylation	24921	6825	72.6
	Acetylation	3384	2298	32.1
R. norvegicus	Phosphorylation	1885	913	51.6
X. laevis	Phosphorylation	470	149	68.3
C. elegans	Phosphorylation	6715	1074	84.0
D. melanogaster	Phosphorylation	17535	2081	88.1
	Acetylation	1707	858	49.7
S. pombe	Phosphorylation	2540	636	75.0
S. cerevisiae	Phosphorylation	15144	3747	75.3
	Acetylation	657	433	34.1
	Ubiquitylation	2499	1426	42.9
C. albicans	Phosphorylation	2910	532	81.7
A. thaliana	Phosphorylation	4527	648	85.7
O. sativa	Phosphorylation	3140	633	79.8

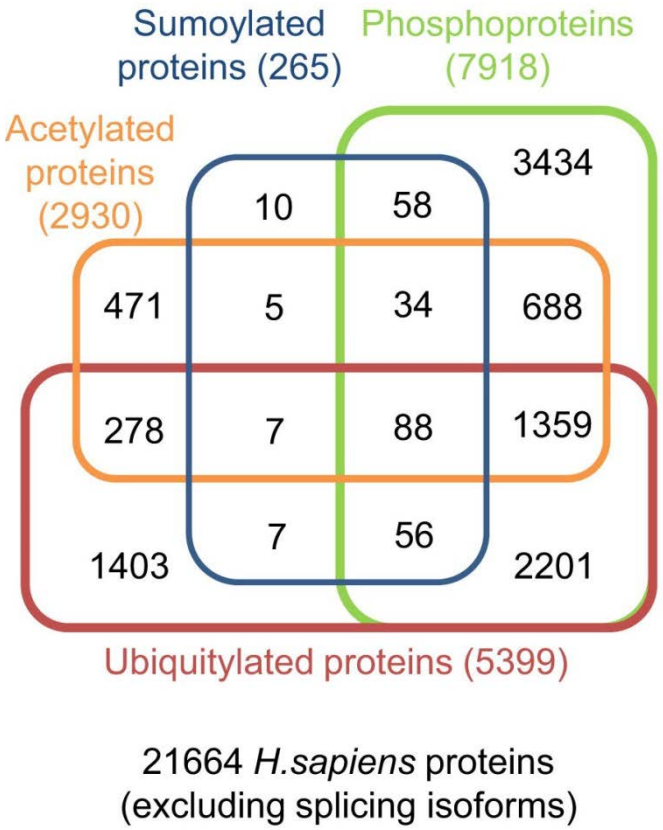
Beltrao et al., Cell 2012 (supplemental data)



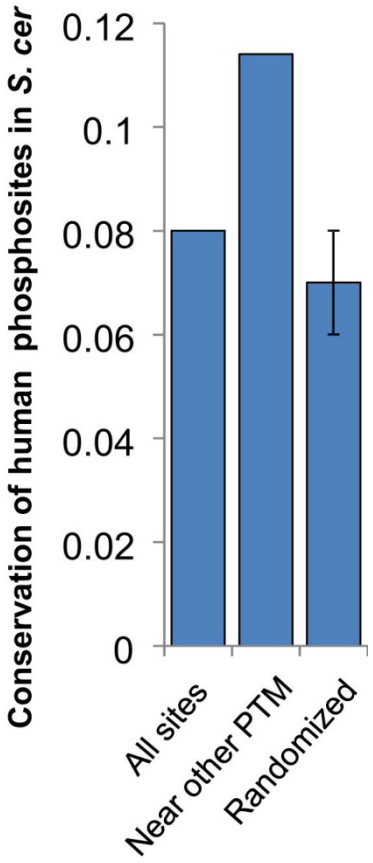
Hunter, Mol Cell 2007

Question: can functionality be assigned to PTM sites that are multiply posttranslationally modified?

ASSOCIATION OF PROTEIN PHOSPHORYLATION WITH LYSINE POSTTRANSLATIONAL MODIFICATIONS

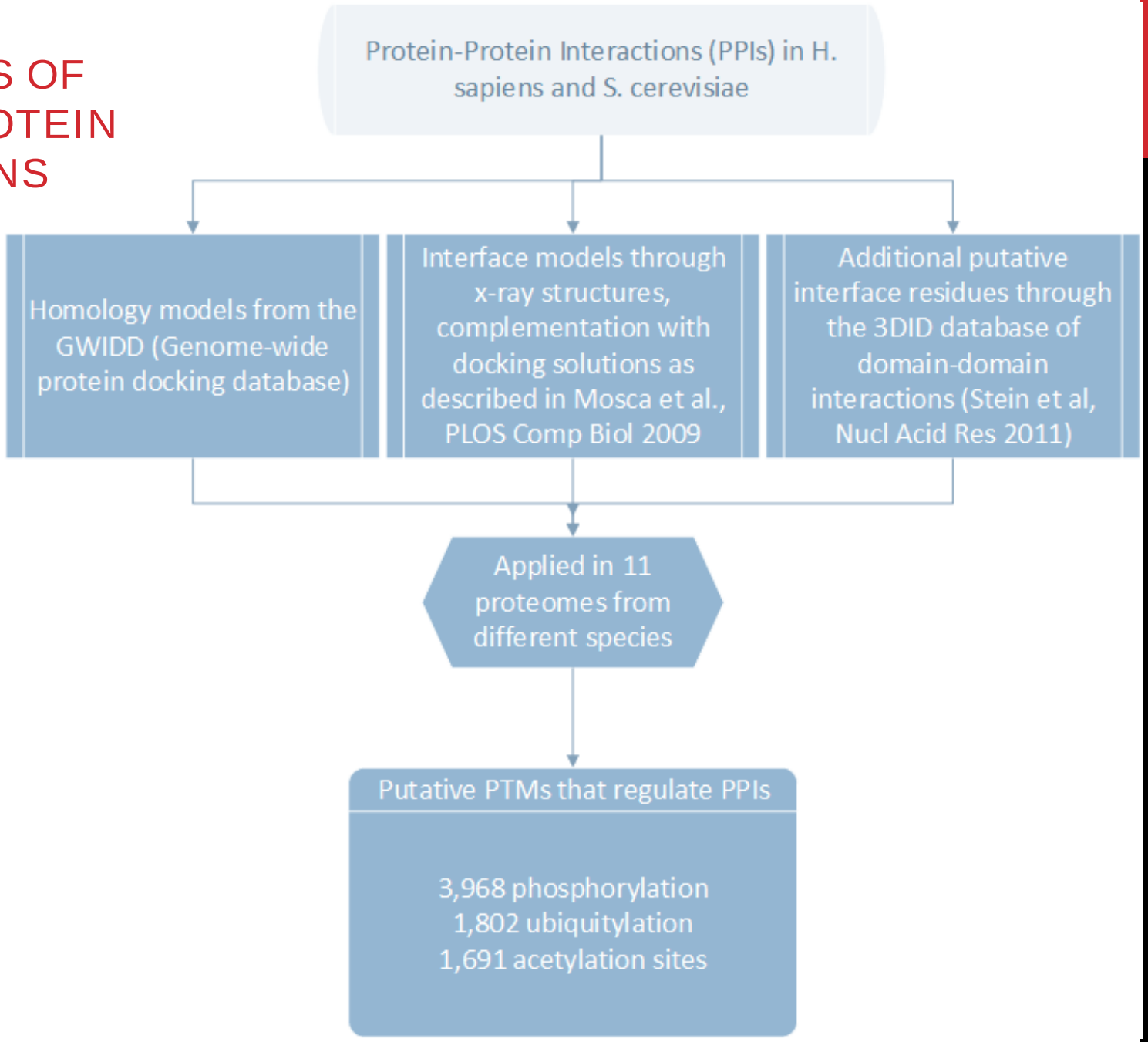


21664 *H.sapiens* proteins
(excluding splicing isoforms)

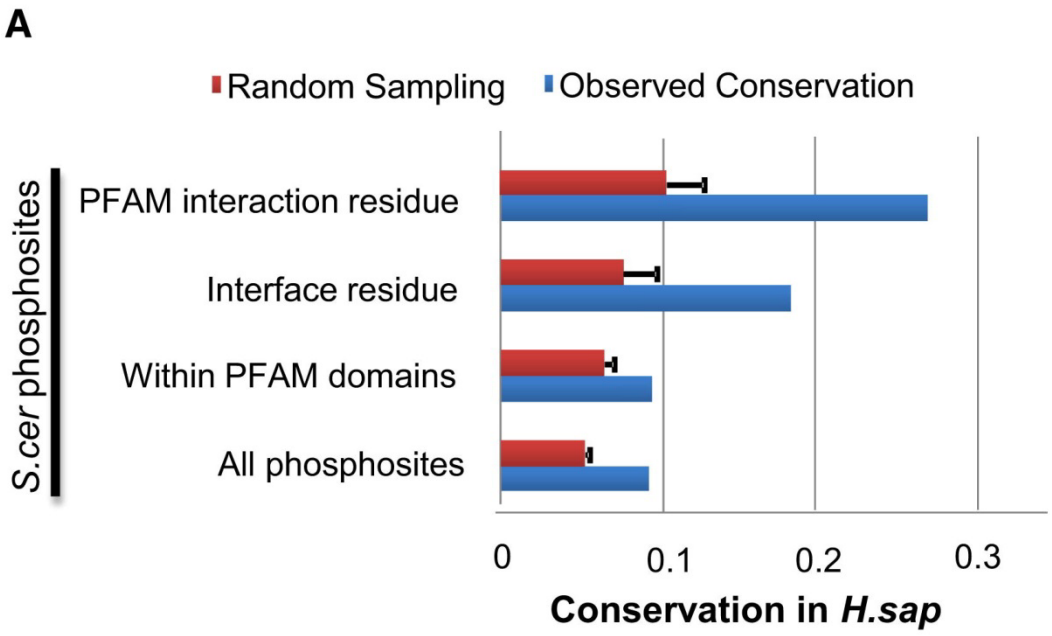


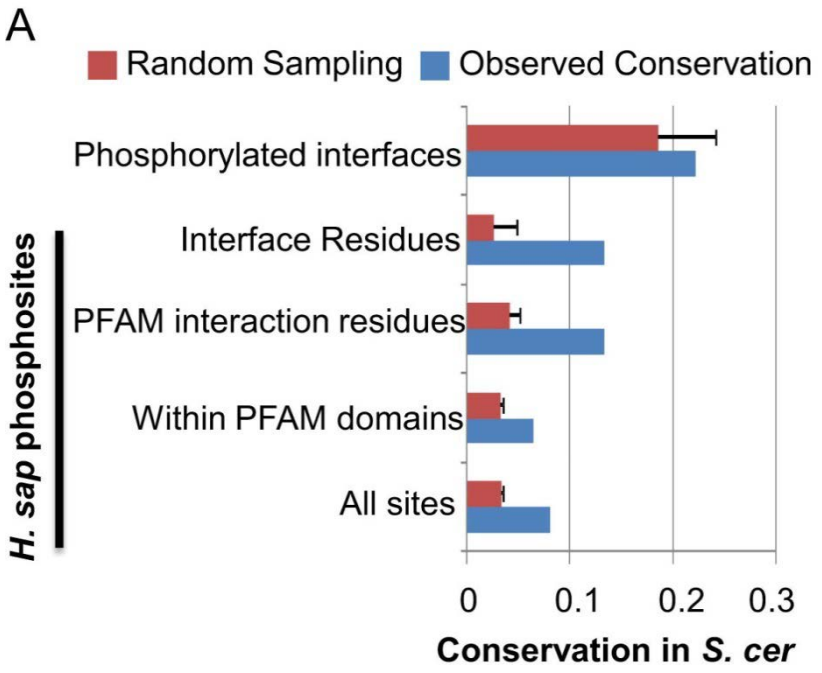
Beltrao et al., Cell 2012

PTMs AS REGULATORS OF PROTEIN-PROTEIN INTERACTIONS

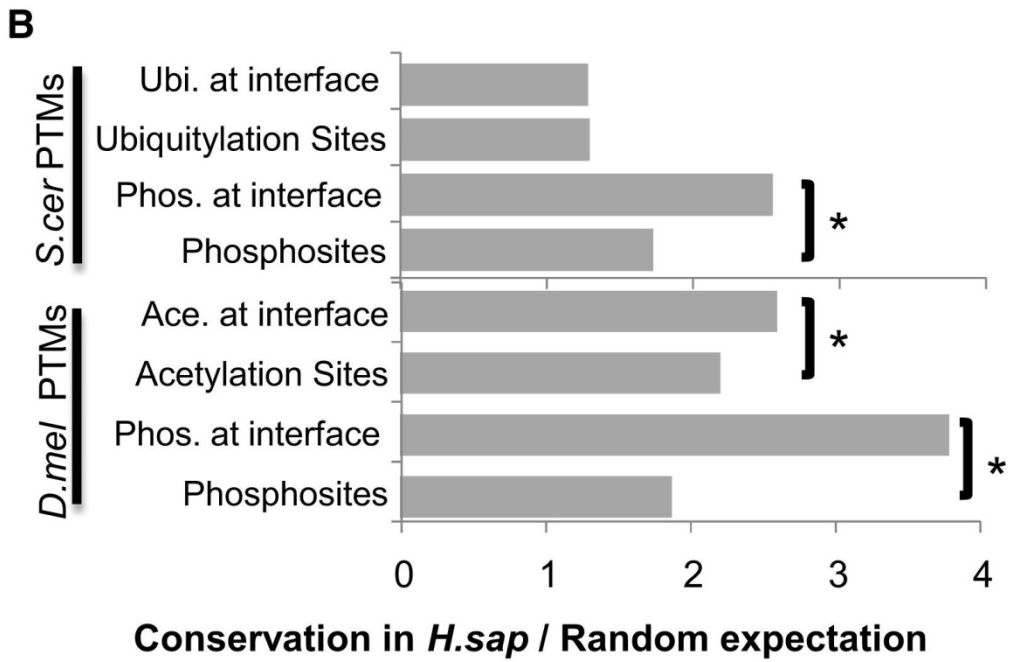


S. CEREVISIAE PHOSPHOSITES ARE MORE LIKELY TO BE CONSERVED AT INTERFACE RESIDUES THAN AVERAGE PHOSPHOSITES



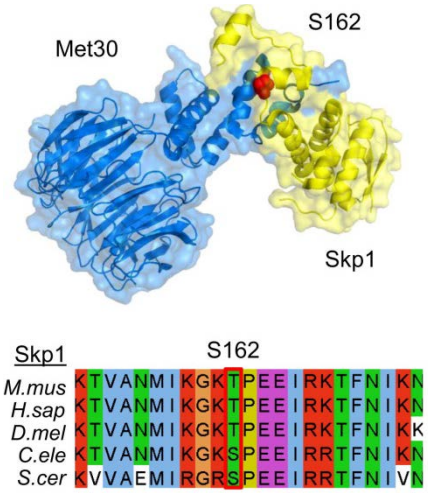


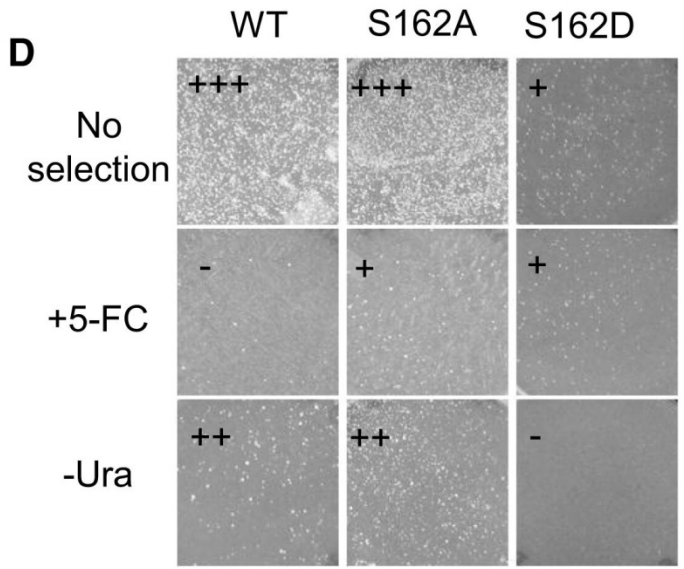
OTHER PTMS ARE ALSO MORE LIKELY TO BE CONSERVED WHEN RESIDING AT INTERFACE RESIDUES



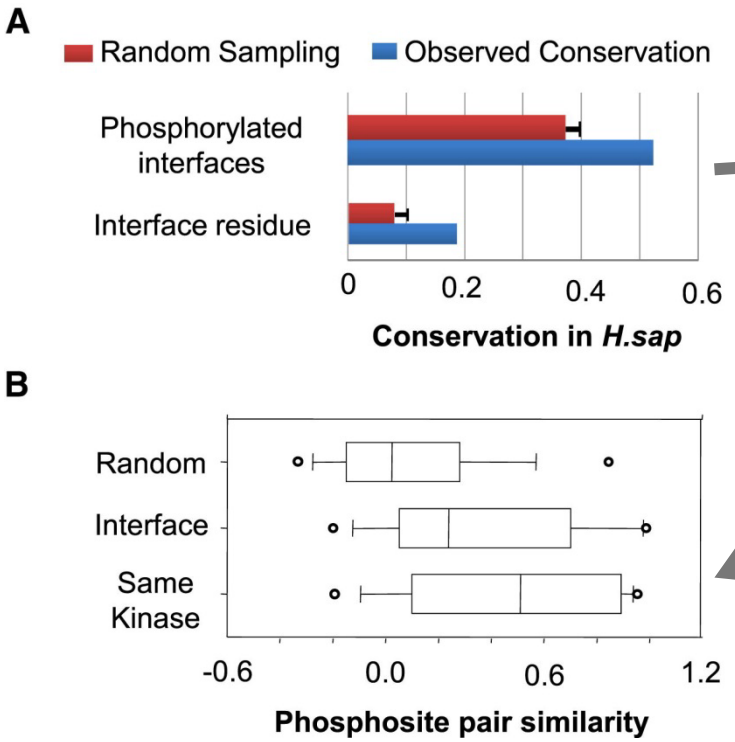
Conclusion: acetylation and phosphorylation, but not ubiquitylation are regulators of binding affinity of protein interactions

TESTING PROTEIN-PROTEIN INTERACTION DEPENDENCY ON PHOSPHORYLATION STATUS





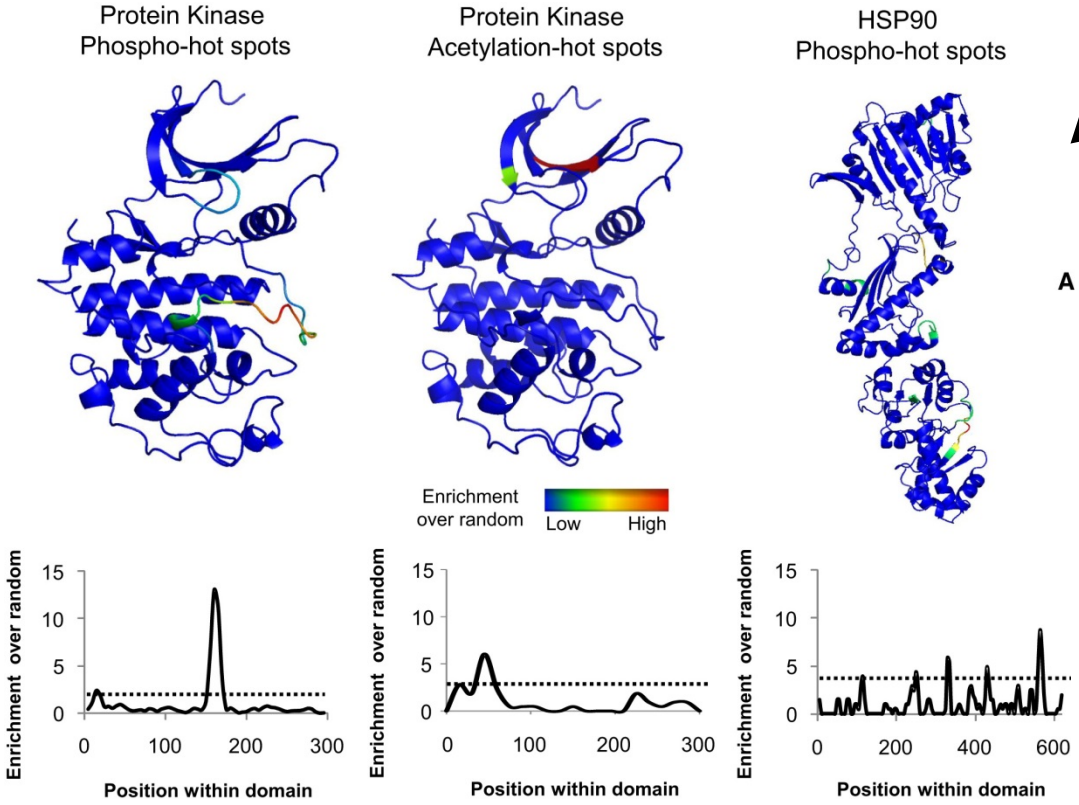
PHOSPHOSITE SIMILARITY



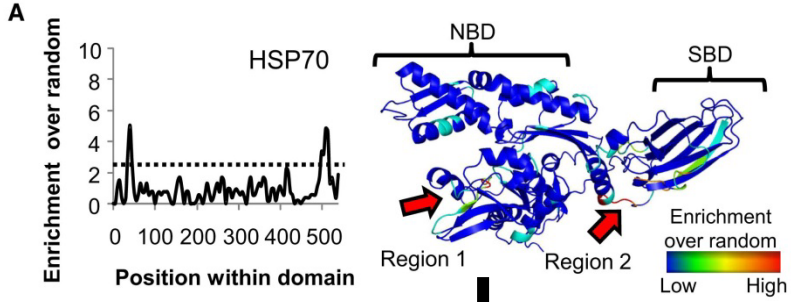
~ 50% of phosphorylated interfaces in S. cerevisiae are conserved in H. sapiens while only ~ 18% of the interface residues (i.e. AA) are conserved.
True observation? Lack of coverage?

POSTTRANSLATIONAL HOT SPOTS WITHIN DOMAIN FAMILIES

PFAM id	Domain name
Pkinase	Protein kinase
Pkinase_Tyr	Protein tyrosine kinase
HSP70	Heat shock proteins, Hsp70
RRM_1	RNA recognition motif
UCH	Ubiquitin carboxyl-terminal hydrolase
Ras	Ras domain
HSP90	Heat shock proteins, Hsp90
PH	Pleckstrin homology domain
MFS_1	Major Facilitator Superfamily (MFS) transporters
Mito_carr	Mitochondrial carrier

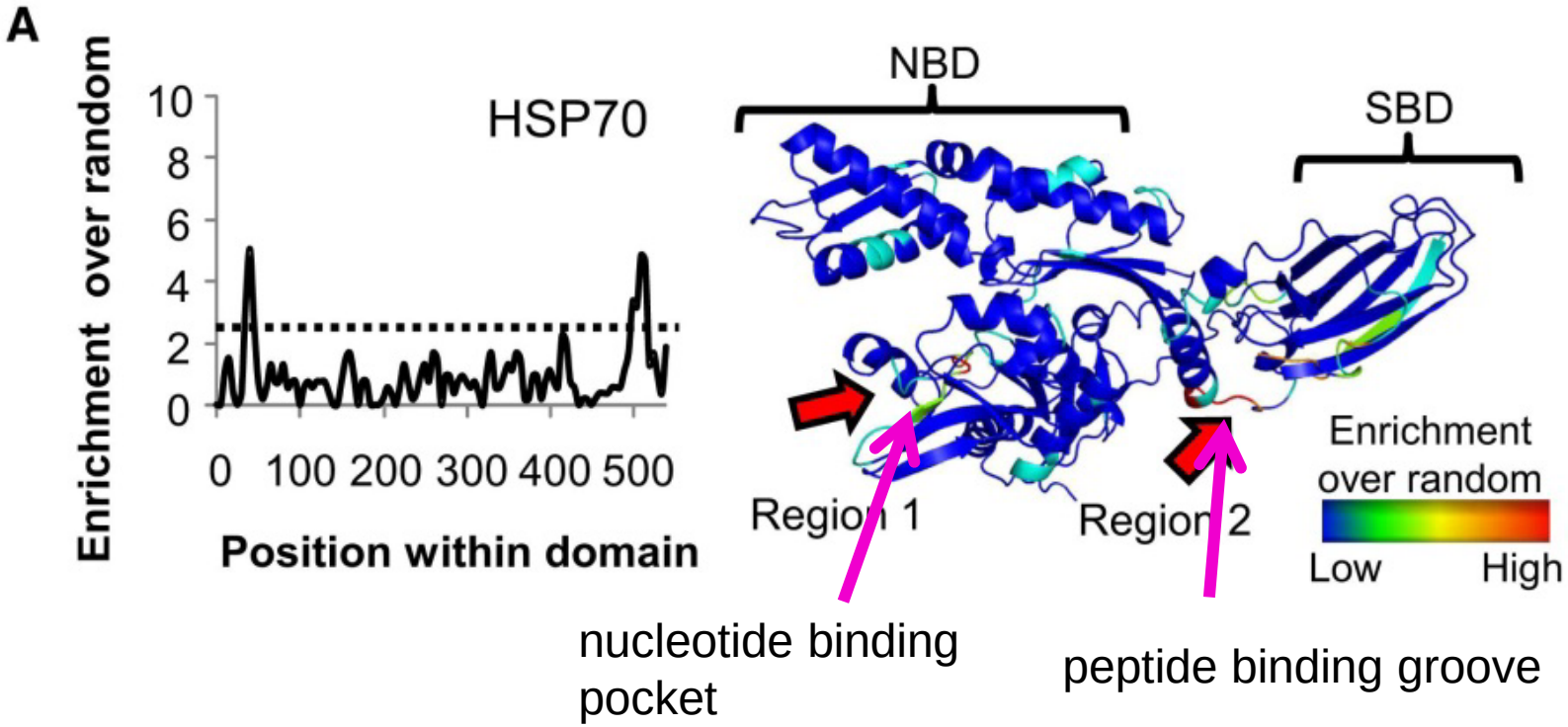


enrichment of PTM hot spots across 11 species above random expectation



HSP70 was taken for further experiments

HSP70

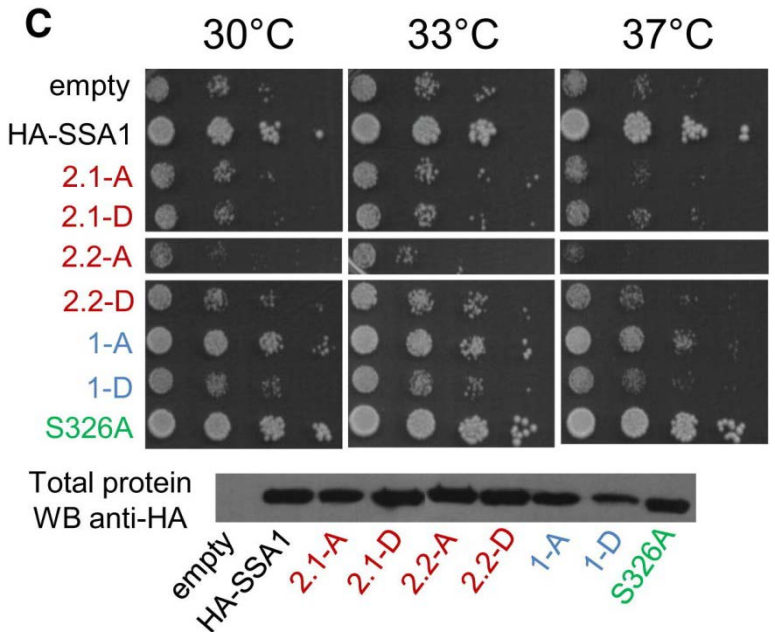


nucleotide binding pocket

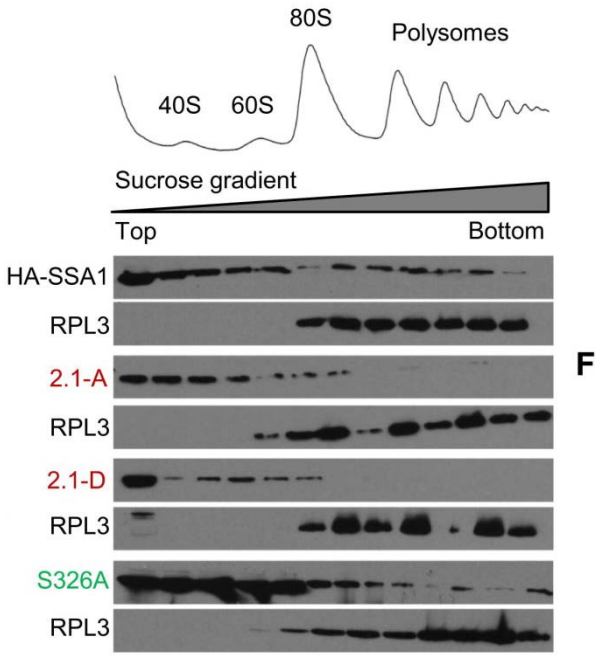
peptide binding groove

SSA1 is a cytosolic HSP70 in yeast >> D and A mutations in phosphosite hotspots

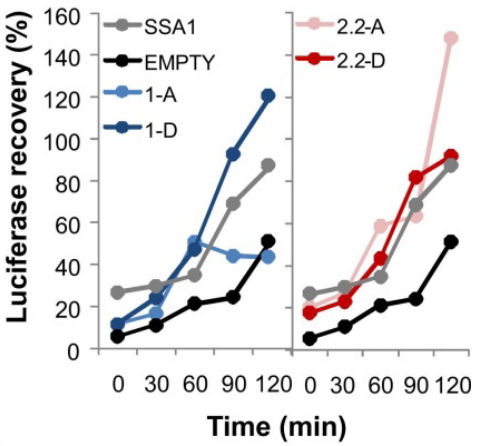
Yeast spotting assay

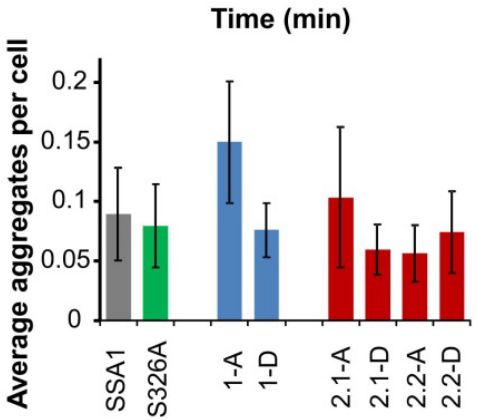


Binding of HSP70 to polysomes



HSP70 assist the refolding of denaturated proteins





CONCLUSIONS

- **Robust source of nearly 200.000 PTMs across 11 different species to investigate PPIs through protein-interfaces or domain activity**
- **Practical&theoritical example of functional priorization of PTMs**
- **Could also be used to study the evolution of posttranslational regulation**