

ES-related protein interaction network. D. PTMPathway results showing all phosphoproteins in mouse focal adhesion pathway. E. PTMCluster results for the protein CTNA1_MOUSE. The yellow bar is a PTM position cluster predicted by PTMCluster. F. PTMPhylog results for the protein CTNA1_MOUSE. Residues shown in red are recorded in SysPTM-A; residues in blue are uniquely from SysPTM-B. S641 and S652 are identified phosphosites in the mES case.

SUPPLEMENTARY MATERIAL:

(Supplementary material is freely available at: <http://www.biosino.org/papers/SysPTM>)

Supplementary Figure 1: A simplified entity-relationship (ER) diagram for the SysPTM database.

Supplementary Figure 2: A neighborhood model to find PTM position clusters.

Supplementary Figure 3: Amino acid bias of different modification types. For a given modification type, the percentage of experimentally identified modifications that occur at each type of amino acid is shown by heatmap.

Supplementary Figure 4: Sample pages showing SysPTM database interface. A. Search interface. Users can query a protein by sequence or keyword. B. Browse interface. Users can browse proteins by modification type, data source, or KEGG pathway. C. Browse/search results for the example protein entry M17929. Each section of the protein entry can be opened by clicking to reveal detailed information (Supplementary Fig. 5).

Supplementary Figure 5: Detailed view of example SysPTM entry result, showing eight expanded sections: Summary, PTMsite-Statistics, PTMsite-Map, PTMsite-Table, PTMsite-Source, PTMsite-Cluster, PTMprotein-Sequence, and PTMprotein-Annotation.

Supplementary Table 1: Quality control and summary of MS/MS literature providing data for SysPTM-B.

Supplementary Table 2: Versions and statistics for public resources providing data for SysPTM-A.

Supplementary Table 3: Statistics for MS/MS literature providing data for SysPTM-B.

Supplementary Table 4: Number of PTM sites and proteins for each PTM type

covered in SysPTM.

Supplementary Table 5: List of phosphopeptides identified by SEQUEST in mouse embryonic stem cells. The symbols * and # indicate phosphorylation and oxidation, respectively. Clicking the hyperlink on the peptide sequence will link automatically to the mass spectrum.

Supplementary Table 6: Significantly enriched ($P \leq 0.01$) functional annotations of six frequent modifications (acetylation, glycosylation, methylation, phosphorylation, sumoylation and ubiquitination), in the categories of GO cellular component, GO biological process, GO molecular function, KEGG pathway, and Pfam domain. P-value was calculated by the chi-square test.