Supplementary figures

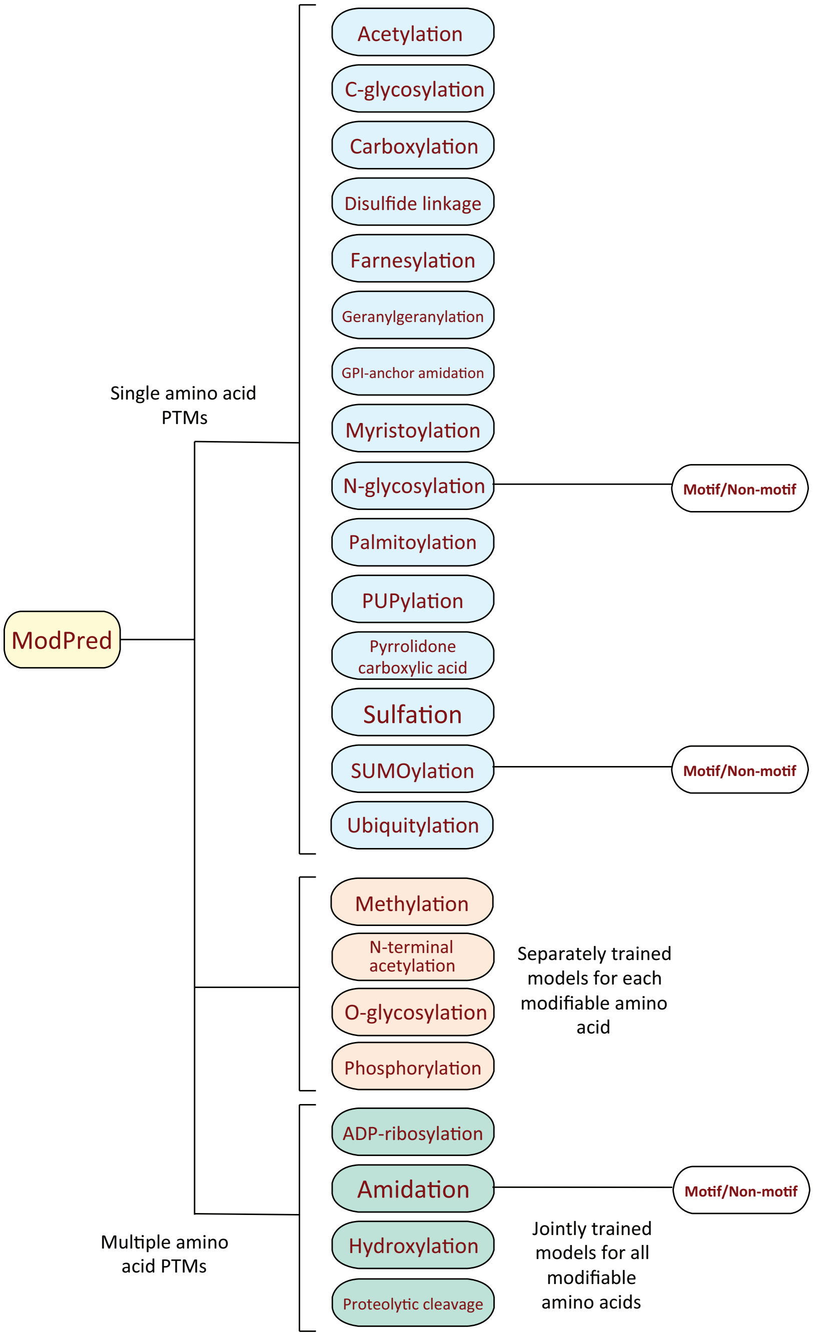
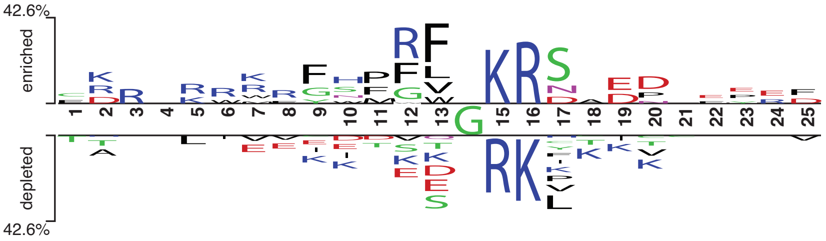
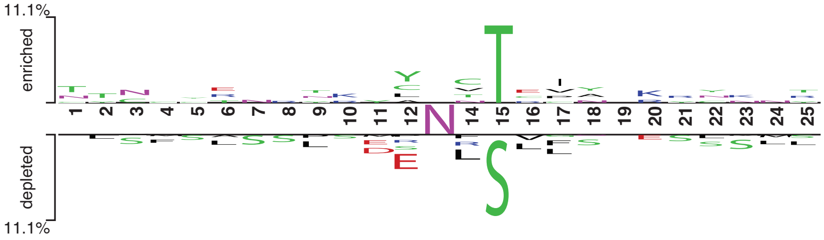


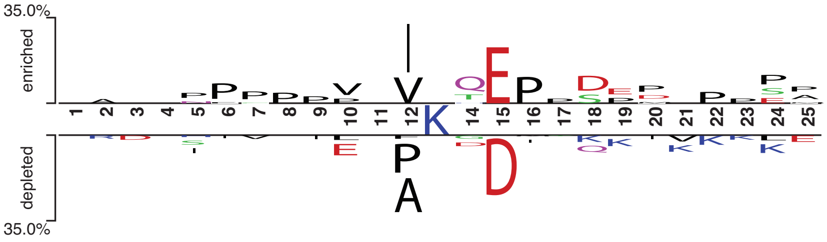
Figure S1. Schematic describing the training protocols for the 34 ensemble models in ModPred covering all 23 PTMs. Note that while certain PTMs are not exclusively restricted to one amino acid, we treat them as ‘single amino acid PTMs’ due to the lack of sufficient training data for the other amino acids.



(A) Amidation (Motif: **x** - G - [KR] - [KR])



(B) N-linked glycosylation (Motif: **N** - P - [ST] - P)



(c) SUMOylation (Motif: [AFILMPV] - **K** - x- [DE])

Figure S2. Two-sample logos generated for 25-amino acid fragments containing positive (modified) and negative (not known to be modified) sites for the three motif-associated modifications. In each case, the central position is occupied by the modifiable residue. Two-sample logos allow for the detection of enrichments and/or depletions of expected motifs when compared across two data sets. Here, modified sites in canonical motif regions were compared to motif-associated sites not known to be modified. In these cases, enrichments for motifs range from approximately 11% to 43%. For sizes of datasets, see Table S1.

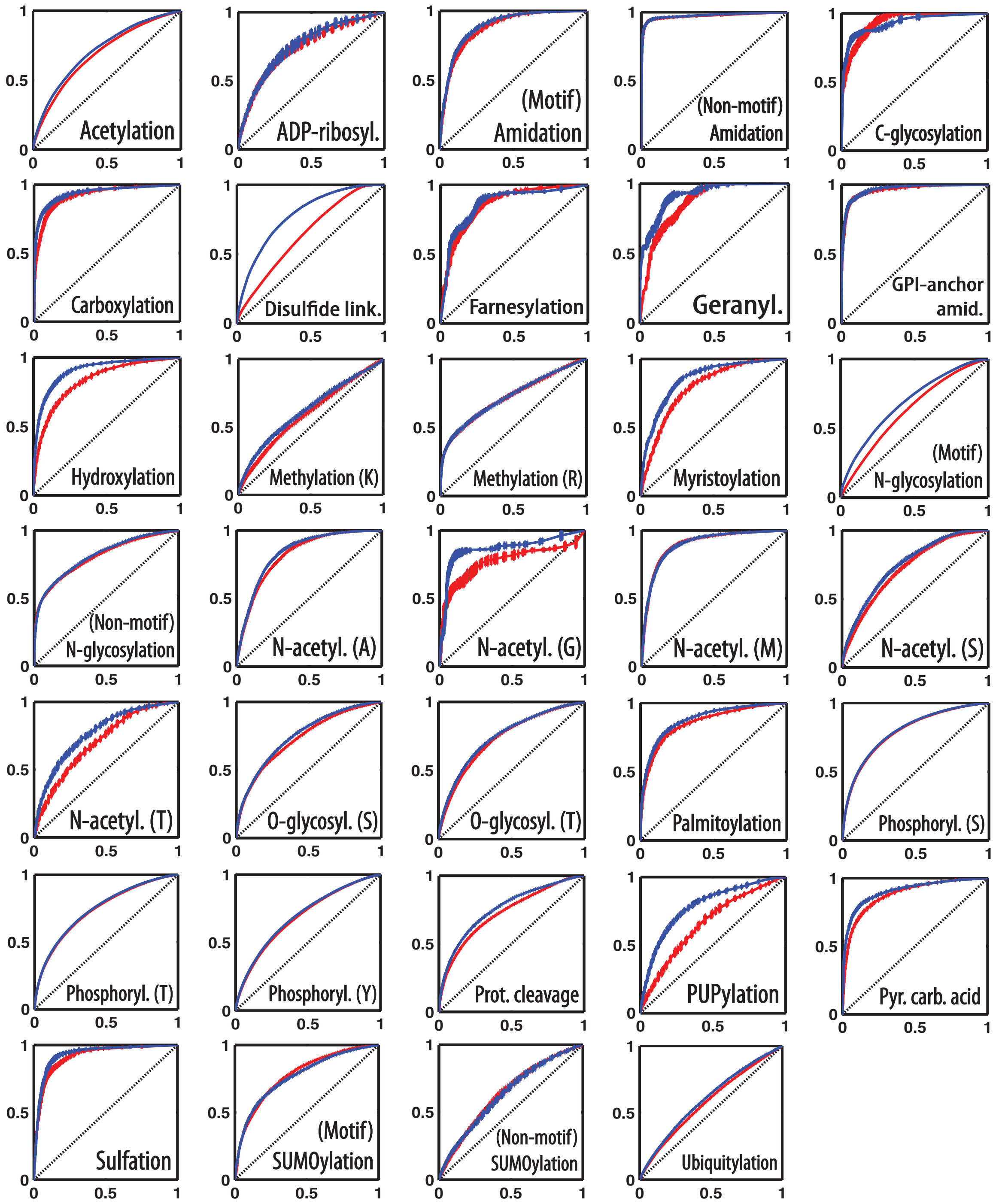

Acetylation
ADP-ribosyl.
(Motif) Amidation
(Non-motif) Amidation
C-glycosylation
Carboxylation
Disulfide link.
Farnesylation
Geranyl.
GPI-anchor amid.
Hydroxylation
Methylation (K)
Methylation (R)
Myristoylation
(Motif) N-glycosylation
(Non-motif) N-glycosylation
N-acetyl. (A)
N-acetyl. (G)
N-acetyl. (M)
N-acetyl. (S)
N-acetyl. (T)
O-glycosyl. (S)
O-glycosyl. (T)
Palmitoylation
Phosphoryl. (S)
Phosphoryl. (T)
Phosphoryl. (Y)
Prot. cleavage
PUPylation
Pyr. carb. acid
Sulfation
(Motif) SUMOylation
(Non-motif) SUMOylation
Ubiquitylation
0
0.5
1

Figure S3. ROC curves for the performance of individual bagged logistic regression models in ModPred, with (blue curve) and without (red curve) evolutionary features (PSSMs). Mean values from 30 repetitions of cross-validation are plotted and the error bars represent the standard deviation. The x-axis represents the false positive rate (1 − *sp*) and the y-axis represents the true positive rate (*sn*). The diagonal dashed line in each case represents the performance of a uniformly random predictor.

Supplementary tables

Modification type	**Residue / Type**	**Number of sites**		**AUC**		*sn* **at** *sp* = 0.90		*sn* **at** *sp* = 0.95		*sn* **at** *sp* = 0.99	
		Positive	Negative	No PSSM	PSSM	No PSSM	PSSM	No PSSM	PSSM	No PSSM	PSSM
Amidation	Motif	210	476*	0.877	**0.891**	0.624	**0.665**	0.422	**0.437**	**0.080**	0.067
	Non-motif	247	29,490	0.971	**0.973**	0.952	**0.956**	**0.935**	0.932	0.785	**0.808**
Methylation	K	303	10,087	0.594	**0.622**	0.194	**0.250**	0.113	**0.155**	0.027	**0.042**
	R	325	8,474	0.720	**0.722**	0.437	**0.445**	0.366	**0.368**	**0.211**	0.202
N-linked glycosylation	Motif	10,567	7,558	0.627	**0.686**	0.169	**0.277**	0.092	**0.175**	0.020	**0.066**
	Non-motif	719	70,492	0.802	**0.812**	0.548	**0.561**	**0.473**	**0.473**	**0.278**	0.255
N-terminal acetylation	A	527	583*	0.821	**0.838**	0.447	**0.464**	0.265	**0.285**	**0.069**	0.066
	G	21	89*	0.771	**0.863**	0.567	**0.789**	0.468	**0.576**	0.046	**0.056**
	M	334	616*	**0.898**	0.897	**0.713**	0.699	0.487	**0.501**	0.129	**0.145**
	S	345	525*	0.720	**0.751**	0.284	**0.337**	0.173	**0.216**	0.057	**0.074**
	T	83	189*	0.702	**0.769**	0.271	**0.392**	0.161	**0.253**	0.021	**0.027**
O-linked glycosylation	S	523	690	0.730	**0.751**	0.366	**0.379**	0.246	**0.263**	0.070	**0.089**
	T	221	16,849	0.733	**0.747**	0.333	**0.369**	0.200	**0.238**	0.052	**0.073**
Phosphorylation	S	59,458	201,815	0.794	**0.799**	0.473	**0.482**	0.330	**0.344**	**0.127**	0.116
	T	19,670	83,104	0.733	**0.747**	0.333	**0.369**	0.200	**0.238**	0.052	**0.073**
	Y	10,930	35,587	0.689	**0.698**	0.279	**0.297**	0.174	**0.183**	**0.052**	**0.052**
SUMOylation	Motif	523	690	**0.779**	0.770	0.433	**0.458**	0.298	**0.300**	0.048	**0.089**
	Non-motif	221	16,849	**0.639**	0.636	0.196	**0.210**	0.108	**0.126**	0.022	**0.027**

Table S1. A breakdown of performance for each amino acid in cases where the PTMs affect multiple residues or are motif-associated. The ‘No PSSM’ column represents the basic model and the ‘PSSM’ column represents the model with evolutionary features. *sn* and *sp* stand for sensitivity and specificity respectively. The closer each value in the AUC and sensitivity columns is to one, the better the performance. Values marked in bold indicate better show the better-performing model. The data sets marked with a ‘*’ indicate that the negatives were obtained from proteins different to those containing the positives through a random sampling procedure (See Materials and Methods).

Keyword	P-value			Frequencies (Main / Control)		
	All	Disordered	Ordered	All	Disordered	Ordered
4Fe-4S	-	2.51E-05	0.0016	-	0.0010 / 0.0034	0.0080 / 0.0027
ADP-ribosylation	2.70E-03	-	-	0.0028 / 0.0015	-	-
ANK repeat	5.90E-03	-	-	0.0114 / 0.0087	-	-
ATP-binding	2.23E-13	3.05E-32	-	0.0764 / 0.0992	0.0779 / 0.1260	-
Acetylation	1.81E-60	2.77E-32	5.00E-18	0.1600 / 0.0990	0.1571 / 0.1023	0.1771 / 0.0970
Actin-binding	6.02E-29	3.08E-20	-	0.0213 / 0.0089	0.0248 / 0.0107	-
Activator	3.32E-86	2.42E-14	-	0.0608 / 0.0237	0.0708 / 0.0480	-
Acyltransferase	8.47E-17	5.35E-05	0.0039	0.0040 / 0.0111	0.0037 / 0.0074	0.0056 / 0.0134
Adaptive immunity	4.00E-04	0.0021	-	0.0016 / 0.0034	0.0017 / 0.0037	-
Allosteric enzyme	1.00E-04	-	-	0.0012 / 0.0030	-	-
Alternative promoter usage	9.00E-04	-	-	0.0037 / 0.0020	-	-
Alternative splicing	2.55E-73	1.11E-05	3.93E-08	0.4845 / 0.3336	0.5551 / 0.4969	0.1759 / 0.2499
Amidation	5.76E-07	0.0002	3.40E-08	0.0045 / 0.0019	0.0033 / 0.0012	0.0117 / 0.0024
Amino-acid biosynthesis	5.25E-12	-	-	0.0014 / 0.0053	-	-
Aminoacyl-tRNA synthetase	2.49E-07	-	-	0.0008 / 0.0030	-	-
Angiogenesis	-	0.0038	-	-	0.0035 / 0.0060	-
Antibiotic	-	-	1.03E-06	-	-	0.0141 / 0.0041
Antimicrobial	5.10E-03	-	1.48E-08	0.0059 / 0.0082	-	0.0311 / 0.0125
Antiport	2.42E-05	0.002	-	0.0009 / 0.0027	0.0010 / 0.0028	-
Antiviral defense	9.60E-05	0.0019	-	0.0014 / 0.0034	0.0014 / 0.0034	-
Autocatalytic cleavage	1.30E-03	-	-	0.0008 / 0.0020	-	-
Autophagy	9.50E-03	0.0027	-	0.0029 / 0.0045	0.0027 / 0.0051	-
Auxin signaling pathway	8.90E-03	-	-	0.0008 / 0.0017	-	-
Biomineralization	2.00E-04	-	-	0.0020 / 0.0007	-	-
Bromodomain	1.83E-42	1.45E-21	-	0.0062 / 0.0003	0.0073 / 0.0009	-
Cadmium	2.00E-04	0.0047	-	0.0009 / 0.0001	0.0010 / 0.0002	-
Calcium	4.39E-15	5.65E-12	0.0001	0.0262 / 0.0406	0.0271 / 0.0428	0.0213 / 0.0392
Calcium transport	-	0.0061	-	-	0.0040 / 0.0065	-
Carbohydrate metabolism	1.69E-08	-	-	0.0028 / 0.0067	-	-
Cell adhesion	1.23E-06	3.41E-12	-	0.0116 / 0.0174	0.0129 / 0.0243	-
Cell cycle	5.42E-14	-	-	0.0419 / 0.0283	-	-
Cell division	2.00E-04	-	-	0.0228 / 0.0176	-	-
Cell junction	5.57E-10	-	0.0002	0.0318 / 0.0220	-	0.0056 / 0.0158
Cell membrane	2.07E-88	1.18E-47	7.93E-35	0.0723 / 0.1407	0.0771 / 0.1376	0.0457 / 0.1426
Cell projection	4.50E-07	-	-	0.0241 / 0.0171	-	-
Cell shape	1.20E-03	-	-	0.0023 / 0.0011	-	-
Cell wall	-	0.0029	-	-	0.0043 / 0.0022	-
Cell wall biogenesis/degradatio	7.07E-08	-	-	0.0029 / 0.0066	-	-

n						
Centromere	5.00E-04	-	-	0.0074 / 0.0048	-	-
Chaperone	6.53E-05	0.0025	-	0.0139 / 0.0097	0.0136 / 0.0096	-
Chemotaxis	2.16E-05	-	-	0.0013 / 0.0034	-	-
Chloroplast	2.09E-12	3.20E-20	-	0.0097 / 0.0178	0.0077 / 0.0207	-
Chromatin regulator	3.81E-46	2.70E-12	-	0.0278 / 0.0102	0.0328 / 0.0194	-
Chromosomal rearrangement	2.31E-17	6.60E-06	-	0.0095 / 0.0034	0.0106 / 0.0057	-
Chromosome	4.04E-78	2.28E-20	3.49E-25	0.0353 / 0.0104	0.0355 / 0.0179	0.0341 / 0.0059
Citrullination	2.45E-35	1.63E-15	2.03E-23	0.0057 / 0.0004	0.0044 / 0.0003	0.0137 / 0.0005
Cleavage on pair of basic residues	-	3.42E-08	0.0003	-	0.0074 / 0.0144	0.0218 / 0.0109
Coiled coil	8.74E-168	5.36E-11	-	0.1757 / 0.0774	0.2081 / 0.1686	-
Collagen	2.11E-79	7.11E-41	-	0.0123 / 0.0008	0.0145 / 0.0018	-
Cuticle	2.07E-06	5.30E-05	-	0.0019 / 0.0004	0.0021 / 0.0004	-
Cyclin	-	0.0002	-	-	0.0016 / 0.0041	-
Cytokine	1.57E-18	0.0003	-	0.0016 / 0.0073	0.0016 / 0.0040	-
Cytoplasm	5.01E-22	-	-	0.3007 / 0.2407	-	-
Cytoplasmic vesicle	3.70E-03	3.94E-06	-	0.0142 / 0.0179	0.0153 / 0.0232	-
Cytoskeleton	6.67E-44	0.0003	5.55E-08	0.0670 / 0.0368	0.0704 / 0.0585	0.0477 / 0.0239
DNA damage	-	5.83E-05	-	-	0.0186 / 0.0260	-
DNA recombination	-	1.67E-05	-	-	0.0035 / 0.0075	-
DNA repair	-	6.64E-06	-	-	0.0153 / 0.0230	-
DNA-binding	0.00E+00	2.38E-90	4.85E-92	0.2425 / 0.0659	0.2611 / 0.1400	0.1446 / 0.0251
DNA-directed RNA polymerase	6.50E-03	-	-	0.0013 / 0.0025	-	-
Defensin	-	-	2.29E-06	-	-	0.0108 / 0.0026
Developmental protein	1.67E-55	2.59E-06	-	0.0829 / 0.0447	0.0939 / 0.0762	-
Differentiation	6.31E-19	-	-	0.0386 / 0.0237	-	-
Dioxygenase	2.30E-03	-	-	0.0026 / 0.0045	-	-
Direct protein sequencing	-	-	1.70E-07	-	-	0.1862 / 0.1301
Disulfide bond	1.44E-255	1.91E-90	3.18E-32	0.0583 / 0.1831	0.0523 / 0.1272	0.0940 / 0.2202
EGF-like domain	8.00E-04	5.46E-06	-	0.0072 / 0.0104	0.0084 / 0.0144	-
ER-Golgi transport	8.67E-16	6.92E-06	-	0.0016 / 0.0068	0.0018 / 0.0050	-
Electron transport	5.94E-09	-	5.16E-05	0.0041 / 0.0088	-	0.0252 / 0.0123
Elongation factor	9.30E-03	-	1.03E-16	0.0021 / 0.0011	-	0.0122 / 0.0008
Endonuclease	7.47E-06	2.41E-08	-	0.0023 / 0.0050	0.0023 / 0.0069	-
Endoplasmic reticulum	2.16E-70	2.49E-22	1.32E-13	0.0212 / 0.0558	0.0205 / 0.0410	0.0252 / 0.0650
Endosome	8.32E-12	2.03E-14	0.0017	0.0102 / 0.0182	0.0108 / 0.0226	0.0065 / 0.0156
Exonuclease	1.70E-03	1.54E-05	-	0.0016 / 0.0031	0.0018 / 0.0048	-
Extracellular matrix	3.70E-05	0.007	-	0.0119 / 0.0080	0.0136 / 0.0100	-
FAD	3.32E-21	-	-	0.0025 / 0.0097	-	-
Fatty acid biosynthesis	1.85E-09	0.007	-	0.0010 / 0.0041	0.0008 / 0.0021	-

Fatty acid metabolism	1.19E-14	3.25E-05	-	0.0027 / 0.0083	0.0010 / 0.0035	-
Flavoprotein	5.85E-18	-	-	0.0028 / 0.0095	-	-
Fungicide	3.40E-03	-	5.11E-05	0.0031 / 0.0050	-	0.0184 / 0.0078
G-protein coupled receptor	7.39E-106	8.73E-22	3.01E-24	0.0047 / 0.0338	0.0045 / 0.0159	0.0056 / 0.0452
GTP-binding	7.21E-06	8.30E-08	5.67E-13	0.0136 / 0.0194	0.0068 / 0.0133	0.0549 / 0.0231
GTPase activation	8.44E-11	-	-	0.0111 / 0.0056	-	-
Glucose metabolism	-	-	0.0029	-	-	0.0056 / 0.0015
Glycolysis	-	-	5.00E-18	-	-	0.0228 / 0.0038
Glycoprotein	0.00E+00	4.37E-124	1.91E-103	0.0903 / 0.2775	0.0949 / 0.2175	0.0643 / 0.3173
Glycosidase	1.06E-32	-	-	0.0007 / 0.0083	-	-
Glycosyltransferase	1.71E-82	8.17E-28	-	0.0012 / 0.0195	0.0009 / 0.0099	-
Golgi apparatus	3.09E-62	2.08E-40	8.77E-18	0.0154 / 0.0436	0.0165 / 0.0439	0.0089 / 0.0434
Growth regulation	4.02E-06	0.0028	-	0.0046 / 0.0021	0.0053 / 0.0030	-
Helicase	2.18E-05	-	-	0.0113 / 0.0074	-	-
Heme	7.27E-28	0.0009	-	0.0031 / 0.0125	0.0008 / 0.0025	-
Homeobox	1.52E-53	1.48E-11	-	0.0241 / 0.0072	0.0278 / 0.0160	-
Hydrogen ion transport	3.11E-08	-	-	0.0014 / 0.0045	-	-
Hydrolase	1.10E-168	2.03E-70	2.71E-55	0.0405 / 0.1209	0.0435 / 0.1022	0.0233 / 0.1327
Hydroxylation	2.18E-33	2.64E-17	0.0003	0.0104 / 0.0023	0.0110 / 0.0031	0.0070 / 0.0018
Immunity	8.76E-11	1.19E-06	-	0.0091 / 0.0163	0.0080 / 0.0143	-
Immunoglobulin V region	8.28E-05	-	0.0054	0.0020 / 0.0042	-	0.0137 / 0.0068
Immunoglobulin domain	1.75E-41	1.53E-30	-	0.0074 / 0.0240	0.0059 / 0.0216	-
Inflammatory response	2.48E-10	2.46E-05	-	0.0014 / 0.0050	0.0015 / 0.0044	-
Innate immunity	1.15E-07	8.39E-05	-	0.0045 / 0.0089	0.0046 / 0.0085	-
Intermediate filament	1.46E-10	-	-	0.0057 / 0.0021	-	-
Ion channel	4.11E-13	8.91E-16	-	0.0093 / 0.0176	0.0103 / 0.0226	-
Ion transport	2.48E-53	2.18E-38	7.63E-10	0.0140 / 0.0386	0.0142 / 0.0392	0.0132 / 0.0382
Iron	4.65E-57	6.17E-14	0.0089	0.0086 / 0.0304	0.0055 / 0.0144	0.0272 / 0.0405
Iron-sulfur	6.28E-06	2.61E-07	-	0.0023 / 0.0051	0.0011 / 0.0044	-
Isomerase	2.47E-05	-	-	0.0054 / 0.0089	-	-
Isopeptide bond	2.70E-39	6.46E-10	1.40E-17	0.0349 / 0.0158	0.0338 / 0.0215	0.0416 / 0.0123
Kelch repeat	3.29E-16	0.0007	-	0.0011 / 0.0058	0.0013 / 0.0033	-
Keratin	2.47E-44	6.28E-15	9.77E-14	0.0108 / 0.0017	0.0110 / 0.0035	0.0098 / 0.0006
Keratinization	1.86E-20	3.71E-09	-	0.0035 / 0.0003	0.0040 / 0.0008	-
Kinase	9.12E-08	1.39E-15	0.0002	0.0327 / 0.0433	0.0351 / 0.0560	0.0194 / 0.0356
LIM domain	5.00E-04	-	-	0.0045 / 0.0025	-	-
Lectin	1.77E-13	0.0002	-	0.0033 / 0.0090	0.0037 / 0.0071	-
Leucine-rich repeat	4.02E-37	6.31E-22	-	0.0052 / 0.0189	0.0059 / 0.0184	-
Ligand-gated ion channel	1.44E-11	2.46E-10	-	0.0012 / 0.0050	0.0014 / 0.0059	-
Ligase	4.28E-07	1.90E-12	-	0.0166 / 0.0237	0.0151 / 0.0275	-

Lipid biosynthesis	1.29E-38	7.50E-10	6.70E-06	0.0020 / 0.0128	0.0015 / 0.0060	0.0047 / 0.0169
Lipid degradation	1.02E-23	3.17E-06	-	0.0011 / 0.0075	0.0011 / 0.0040	-
Lipid droplet	6.50E-03	-	-	0.0007 / 0.0017	-	-
Lipid metabolism	3.84E-76	3.69E-21	1.53E-07	0.0058 / 0.0296	0.0040 / 0.0146	0.0165 / 0.0390
Lipid transport	3.60E-03	0.0033	-	0.0016 / 0.0030	0.0015 / 0.0034	-
Lipid-binding	-	0.0014	-	-	0.0072 / 0.0110	-
Lipoprotein	1.25E-06	-	-	0.0277 / 0.0364	-	-
Lyase	8.74E-14	7.00E-06	-	0.0050 / 0.0118	0.0020 / 0.0054	-
Lysosome	3.24E-32	2.88E-09	-	0.0016 / 0.0106	0.0019 / 0.0066	-
MHC I	-	-	0.0027	-	-	0.0075 / 0.0025
Magnesium	2.49E-23	8.76E-19	-	0.0178 / 0.0337	0.0162 / 0.0327	-
Manganese	4.44E-22	3.01E-07	-	0.0055 / 0.0152	0.0046 / 0.0099	-
Membrane	0.00E+00	3.65E-186	9.23E-55	0.2312 / 0.5806	0.2209 / 0.4821	0.2946 / 0.6478
Metal-binding	8.18E-13	-	0.0011	0.2292 / 0.1915	-	0.2205 / 0.1773
Metal-thiolate cluster	1.84E-19	2.64E-11	-	0.0020 / 0.0000	0.0022 / 0.0000	-
Metalloprotease	1.73E-18	7.38E-13	-	0.0021 / 0.0082	0.0021 / 0.0082	-
Methylation	6.85E-52	4.26E-40	5.00E-18	0.0300 / 0.0107	0.0280 / 0.0082	0.0421 / 0.0123
Methyltransferase	9.00E-04	-	-	0.0076 / 0.0108	-	-
Microsome	1.92E-18	-	-	0.0019 / 0.0078	-	-
Microtubule	1.12E-14	-	1.89E-32	0.0195 / 0.0106	-	0.0366 / 0.0050
Mitochondrion	7.48E-74	6.77E-35	1.67E-05	0.0341 / 0.0774	0.0180 / 0.0438	0.1386 / 0.0991
Mitochondrion inner membrane	8.40E-32	1.07E-15	-	0.0056 / 0.0182	0.0014 / 0.0076	-
Mitochondrion outer membrane	1.71E-12	0.0009	-	0.0017 / 0.0061	0.0018 / 0.0040	-
Mitosis	4.53E-06	-	-	0.0163 / 0.0111	-	-
Monooxygenase	1.15E-24	-	-	0.0017 / 0.0090	-	-
Motor protein	1.45E-05	-	0.0005	0.0089 / 0.0054	-	0.0080 / 0.0024
Multifunctional enzyme	6.98E-22	-	-	0.0007 / 0.0061	-	-
Muscle protein	9.66E-20	1.48E-11	0.001	0.0068 / 0.0017	0.0069 / 0.0019	0.0061 / 0.0015
Myogenesis	2.09E-06	0.0012	-	0.0027 / 0.0009	0.0031 / 0.0012	-
Myosin	7.89E-11	0.0001	-	0.0051 / 0.0017	0.0052 / 0.0024	-
NAD	3.42E-14	8.39E-06	4.84E-05	0.0065 / 0.0142	0.0015 / 0.0045	0.0366 / 0.0202
NADP	1.81E-41	1.39E-09	-	0.0027 / 0.0151	0.0008 / 0.0045	-
Neurogenesis	4.64E-13	-	-	0.0154 / 0.0081	-	-
Neuropeptide	1.46E-05	-	2.09E-13	0.0031 / 0.0012	-	0.0127 / 0.0014
Nitration	-	-	4.48E-06	-	-	0.0089 / 0.0019
Nonsense-mediated mRNA decay	4.60E-06	-	-	0.0029 / 0.0010	-	-
Nuclease	1.86E-12	1.27E-16	-	0.0032 / 0.0086	0.0034 / 0.0119	-
Nucleosome core	2.30E-84	5.25E-30	7.76E-64	0.0107 / 0.0003	0.0076 / 0.0004	0.0292 / 0.0002
Nucleotide-binding	8.98E-22	1.94E-43	-	0.0932 / 0.1270	0.0870 / 0.1481	-
Nucleotidyltransferase	4.36E-08	4.00E-11	-	0.0025 / 0.0060	0.0025 / 0.0083	-

Nucleus	0.00E+00	1.51E-119	1.08E-14	0.6525 / 0.2349	0.7593 / 0.4375	0.2233 / 0.1367
Olfaction	1.30E-45	-	-	0.0009 / 0.0114	-	-
Oxidoreductase	2.30E-88	3.52E-17	-	0.0157 / 0.0519	0.0058 / 0.0163	-
Peroxisome	2.02E-14	0.0051	-	0.0023 / 0.0076	0.0014 / 0.0031	-
Phosphoprotein	0.00E+00	1.23E-111	0.0066	0.7811 / 0.2866	0.9462 / 0.5626	0.1948 / 0.1609
Plant defense	1.02E-06	-	-	0.0057 / 0.0101	-	-
Plastid	1.60E-12	3.20E-20	-	0.0097 / 0.0179	0.0077 / 0.0207	-
Polymorphism	6.33E-35	2.35E-07	-	0.2712 / 0.2021	0.2900 / 0.2495	-
Potassium	1.30E-05	2.21E-08	-	0.0044 / 0.0078	0.0040 / 0.0096	-
Potassium channel	-	0.0042	-	-	0.0029 / 0.0052	-
Potassium transport	3.07E-05	3.91E-08	-	0.0036 / 0.0066	0.0040 / 0.0094	-
Prenylation	3.90E-08	-	-	0.0043 / 0.0087	-	-
Protease	2.13E-67	7.80E-31	1.18E-18	0.0081 / 0.0321	0.0087 / 0.0267	0.0047 / 0.0354
Protease inhibitor	5.06E-12	-	-	0.0015 / 0.0056	-	-
Proteasome	3.95E-11	3.09E-06	-	0.0012 / 0.0049	0.0010 / 0.0037	-
Protein biosynthesis	-	0.0053	0.0006	-	0.0058 / 0.0088	0.0184 / 0.0089
Protein kinase inhibitor	2.00E-04	-	-	0.0016 / 0.0005	-	-
Protein phosphatase	3.25E-08	1.54E-05	-	0.0043 / 0.0088	0.0048 / 0.0093	-
Protein phosphatase inhibitor	6.75E-25	2.34E-13	-	0.0040 / 0.0003	0.0047 / 0.0006	-
Protein synthesis inhibitor	9.70E-03	-	-	0.0007 / 0.0002	-	-
Protein transport	3.61E-33	7.57E-25	1.35E-08	0.0181 / 0.0379	0.0188 / 0.0400	0.0137 / 0.0367
Proteoglycan	1.00E-10	0.0002	-	0.0055 / 0.0020	0.0057 / 0.0027	-
Proto-oncogene	3.25E-08	-	-	0.0088 / 0.0046	-	-
Pyridoxal phosphate	1.64E-19	-	-	0.0008 / 0.0059	-	-
Pyrrolidone carboxylic acid	-	0.0005	-	-	0.0034 / 0.0013	-
Pyruvate	-	-	2.16E-05	-	-	0.0070 / 0.0013
RNA-binding	1.02E-76	6.70E-30	-	0.0645 / 0.0276	0.0726 / 0.0403	-
Receptor	5.39E-164	8.62E-74	4.13E-42	0.0166 / 0.0739	0.0176 / 0.0597	0.0108 / 0.0829
Redox-active center	6.35E-10	0.0018	-	0.0014 / 0.0049	0.0009 / 0.0025	-
Repeat	1.88E-60	0.002	-	0.3435 / 0.2379	0.3771 / 0.3463	-
Repressor	1.93E-119	3.56E-33	-	0.0537 / 0.0155	0.0618 / 0.0313	-
Respiratory chain	-	-	4.05E-05	-	-	0.0156 / 0.0059
Ribonucleoprotein	7.22E-18	4.05E-12	5.30E-26	0.0329 / 0.0197	0.0265 / 0.0147	0.0717 / 0.0227
Ribosomal protein	2.36E-05	0.0021	9.01E-29	0.0207 / 0.0152	0.0130 / 0.0090	0.0674 / 0.0190
Ribosome biogenesis	6.68E-18	3.27E-07	-	0.0105 / 0.0039	0.0115 / 0.0058	-
S-adenosyl-L-methionine	5.00E-04	-	-	0.0064 / 0.0095	-	-
S-nitrosylation	-	-	0.0079	-	-	0.0051 / 0.0015
SH3 domain	8.11E-12	-	-	0.0132 / 0.0068	-	-
SH3-binding	2.01E-12	0.0012	-	0.0052 / 0.0016	0.0061 / 0.0033	-
Secreted	1.58E-42	-	0.0048	0.0593 / 0.0990	-	0.0968 / 0.1236

Sensory transduction	1.09E-37	-	3.30E-10	0.0049 / 0.0184	-	0.0056 / 0.0254
Serine protease	4.57E-29	1.30E-08	-	0.0009 / 0.0081	0.0010 / 0.0047	-
Serine protease inhibitor	1.94E-08	-	-	0.0010 / 0.0038	-	-
Serine/threonine-protein kinase	-	3.59E-06	-	-	0.0300 / 0.0409	-
Signal	4.33E-148	4.47E-39	2.66E-16	0.1022 / 0.2150	0.0941 / 0.1538	0.1513 / 0.2558
Signal-anchor	8.83E-47	1.36E-13	-	0.0069 / 0.0245	0.0074 / 0.0171	-
Sodium	8.49E-07	3.93E-05	-	0.0028 / 0.0061	0.0030 / 0.0065	-
Sodium transport	7.23E-07	2.87E-05	-	0.0027 / 0.0060	0.0029 / 0.0064	-
Spermatogenesis	6.80E-03	-	-	0.0081 / 0.0059	-	-
Spliceosome	1.21E-23	5.36E-15	-	0.0123 / 0.0042	0.0141 / 0.0053	-
Steroid metabolism	6.13E-14	-	-	0.0008 / 0.0046	-	-
Stress response	2.16E-05	-	6.09E-05	0.0114 / 0.0075	-	0.0170 / 0.0070
Sulfation	-	-	0.0065	-	-	0.0056 / 0.0018
Sushi	0.0013	-	-	0.0008 / 0.0021	-	-
Symport	1.18E-28	-	-	0.0007 / 0.0075	-	-
Synapse	0.0022	-	-	0.0161 / 0.0126	-	-
TPR repeat	6.71E-10	1.52E-11	-	0.0034 / 0.0080	0.0040 / 0.0109	-
Thiamine pyrophosphate	-	-	0.0054	-	-	0.0051 / 0.0015
Thick filament	8.17E-09	0.0009	-	0.0020 / 0.0003	0.0023 / 0.0008	-
Thiol protease	0.0004	0.0037	-	0.0043 / 0.0070	0.0051 / 0.0081	-
Thylakoid	2.47E-05	1.85E-05	-	0.0012 / 0.0033	0.0009 / 0.0033	-
Tight junction	1.65E-05	0.0062	-	0.0046 / 0.0022	0.0049 / 0.0028	-
Topoisomerase	0.0001	-	-	0.0012 / 0.0003	-	-
Transcription	0.00E+00	3.69E-89	6.14E-37	0.2691 / 0.0830	0.3005 / 0.1668	0.1138 / 0.0372
Transcription regulation	0.00E+00	3.48E-91	1.83E-40	0.2644 / 0.0789	0.2953 / 0.1621	0.1109 / 0.0335
Transducer	1.15E-106	6.85E-24	5.30E-26	0.0061 / 0.0375	0.0061 / 0.0196	0.0061 / 0.0487
Transferase	2.04E-105	8.38E-51	4.16E-28	0.0531 / 0.1192	0.0547 / 0.1079	0.0441 / 0.1263
Transit peptide	4.26E-55	4.60E-51	2.34E-06	0.0248 / 0.0562	0.0130 / 0.0424	0.0990 / 0.0649
Translation regulation	0.0034	-	-	0.0068 / 0.0047	-	-
Transmembrane	0.00E+00	1.28E-230	4.37E-58	0.1169 / 0.3961	0.1058 / 0.3029	0.1862 / 0.4602
Transmembrane helix	0.00E+00	2.09E-228	4.78E-57	0.1168 / 0.3933	0.1057 / 0.3012	0.1862 / 0.4565
Transport	1.07E-171	1.20E-96	7.68E-21	0.0516 / 0.1417	0.0484 / 0.1240	0.0706 / 0.1528
Tricarboxylic acid cycle	0.0001	-	6.07E-22	0.0037 / 0.0018	-	0.0218 / 0.0025
Triplet repeat expansion	1.68E-07	0.0012	-	0.0014 / 0.0002	0.0016 / 0.0004	-
Tumor antigen	3.87E-08	8.40E-05	-	0.0021 / 0.0004	0.0024 / 0.0006	-
Tumor suppressor	0.0003	-	-	0.0072 / 0.0045	-	-
Tyrosine-protein kinase	0.0017	1.89E-08	-	0.0034 / 0.0055	0.0038 / 0.0093	-
Ubiquinone	-	-	0.0013	-	-	0.0065 / 0.0018
Ubl conjugation	2.12E-56	1.77E-06	1.43E-14	0.0752 / 0.0391	0.0775 / 0.0615	0.0616 / 0.0258
Ubl conjugation pathway	8.74E-14	9.37E-21	3.05E-09	0.0239 / 0.0369	0.0265 / 0.0485	0.0089 / 0.0299

Vacuole	2.04E-13	4.65E-06	-	0.0018 / 0.0064	0.0016 / 0.0048	-
WD repeat	6.90E-18	6.59E-22	-	0.0063 / 0.0150	0.0074 / 0.0208	-
Wnt signaling pathway	1.86E-12	0.0003	-	0.0115 / 0.0055	0.0132 / 0.0085	-
Zinc	5.36E-112	4.08E-15	3.40E-15	0.1741 / 0.0906	0.1844 / 0.1416	0.1179 / 0.0614
Zinc-finger	1.72E-212	9.85E-34	4.34E-43	0.1523 / 0.0556	0.1615 / 0.1045	0.1018 / 0.0277
Zymogen	6.61E-34	6.29E-12	-	0.0018 / 0.0113	0.0016 / 0.0069	-
mRNA processing	1.22E-54	5.89E-31	-	0.0320 / 0.0115	0.0369 / 0.0154	-
mRNA splicing	5.71E-62	1.80E-38	-	0.0267 / 0.0077	0.0309 / 0.0100	-
rRNA processing	2.09E-06	-	-	0.0091 / 0.0053	-	-
rRNA-binding	0.004	-	4.03E-08	0.0034 / 0.0020	-	0.0117 / 0.0024
tRNA processing	2.24E-18	5.65E-12	-	0.0009 / 0.0058	0.0009 / 0.0054	-

Table S2. Complete list of significant keywords in the set of proteins enriched in multiple types of PTM sites (arranged in alphabetical order). All three main vs. control set analyses are shown in the table with P-values and frequencies. P-values are color-coded as follows: red – significantly depleted and blue – significantly enriched. Hyphens imply no significance.