Supplementary Information

Quantitative Maps of Protein Phosphorylation Sites across 14 different Rat Organs and Tissues

Alicia Lundby[1,2,*], Anna Secher[1,3,*], Kasper Lage[1,4,5,6], Nikolai B. Nordsborg[7], Anatoliy Dmytriyev[1], Carsten Lundby[8], and Jesper V. Olsen[1]

[1] NNF Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3b, DK-2200 Copenhagen, Denmark
[2] The Danish National Research Foundation Centre for Cardiac Arrhythmia, Copenhagen N, Denmark.
[3] Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv Denmark
[4] Pediatric Surgical Research Laboratory, Massachusetts General Hospital, 02114, Boston, MA, USA
[5] Harvard Medical School, 02115, Boston, MA, USA.
[6] Broad Institute, 02142, Cambridge, MA, USA.
[7] Department of Exercise and Sport Sciences, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen, Denmark.
[8] Zürich Centre for Integrative Human Physiology, Department of Physiology, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.

* These authors contributed equally to this work.

Correspondence: jesper.olsen@cpr.ku.dk

Supplementary Information contains:
Supplementary Figures S1-S12

All the Supplementary Data Tables are also available for download through CPR PTM Resource site:

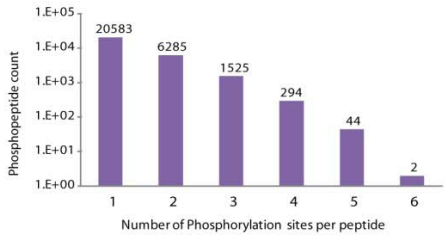
http://cpr1.sund.ku.dk/cgi-bin/PTM.pl

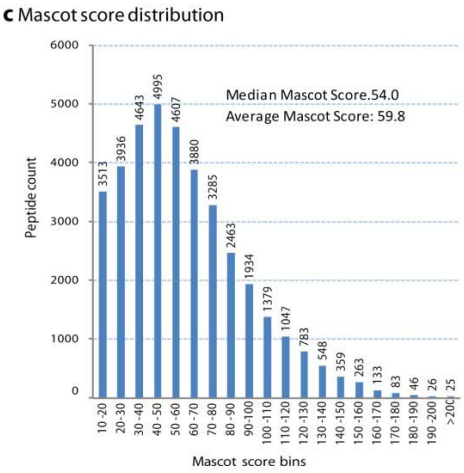
Supplementary Figure S1| Evaluation and assessment of MS data quality

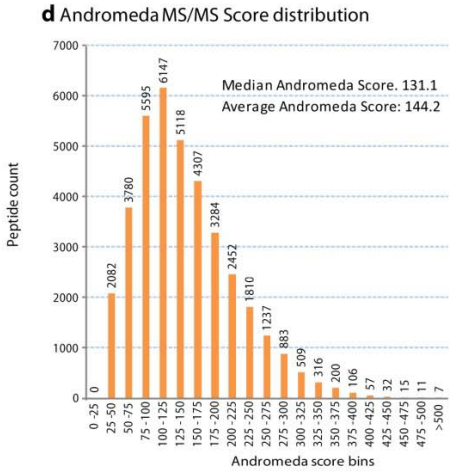
a MS/MS statistics

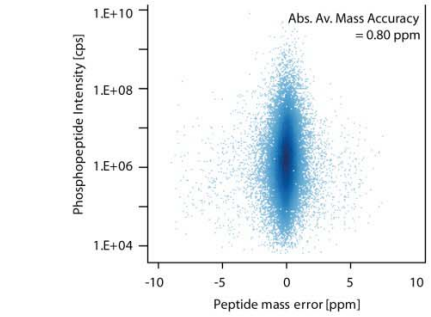
	Total	Peptide-Spectrum M	Fraction
HCD-MS/MSscans	876,203	374,191	0.43

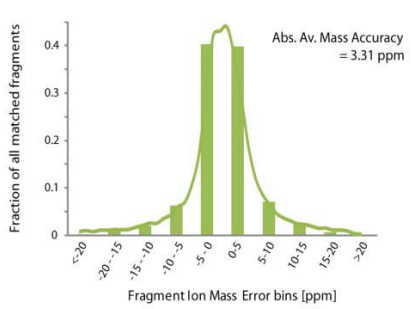
Peptide statistics	Unique Modified Peptides	Fraction
All	37,948	
Phosphopeptides	28,733	0.76
Non-Phosphopeptides	9,215	0.24

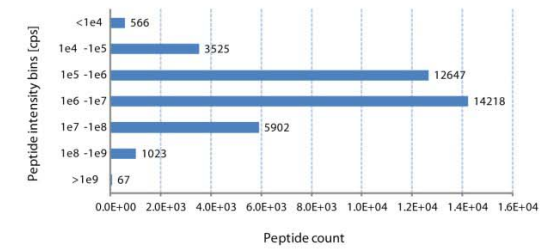


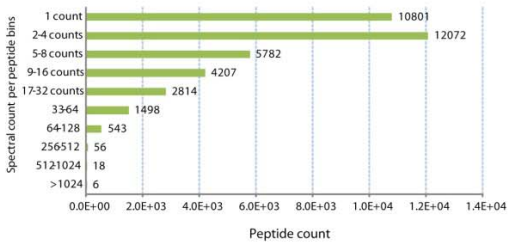






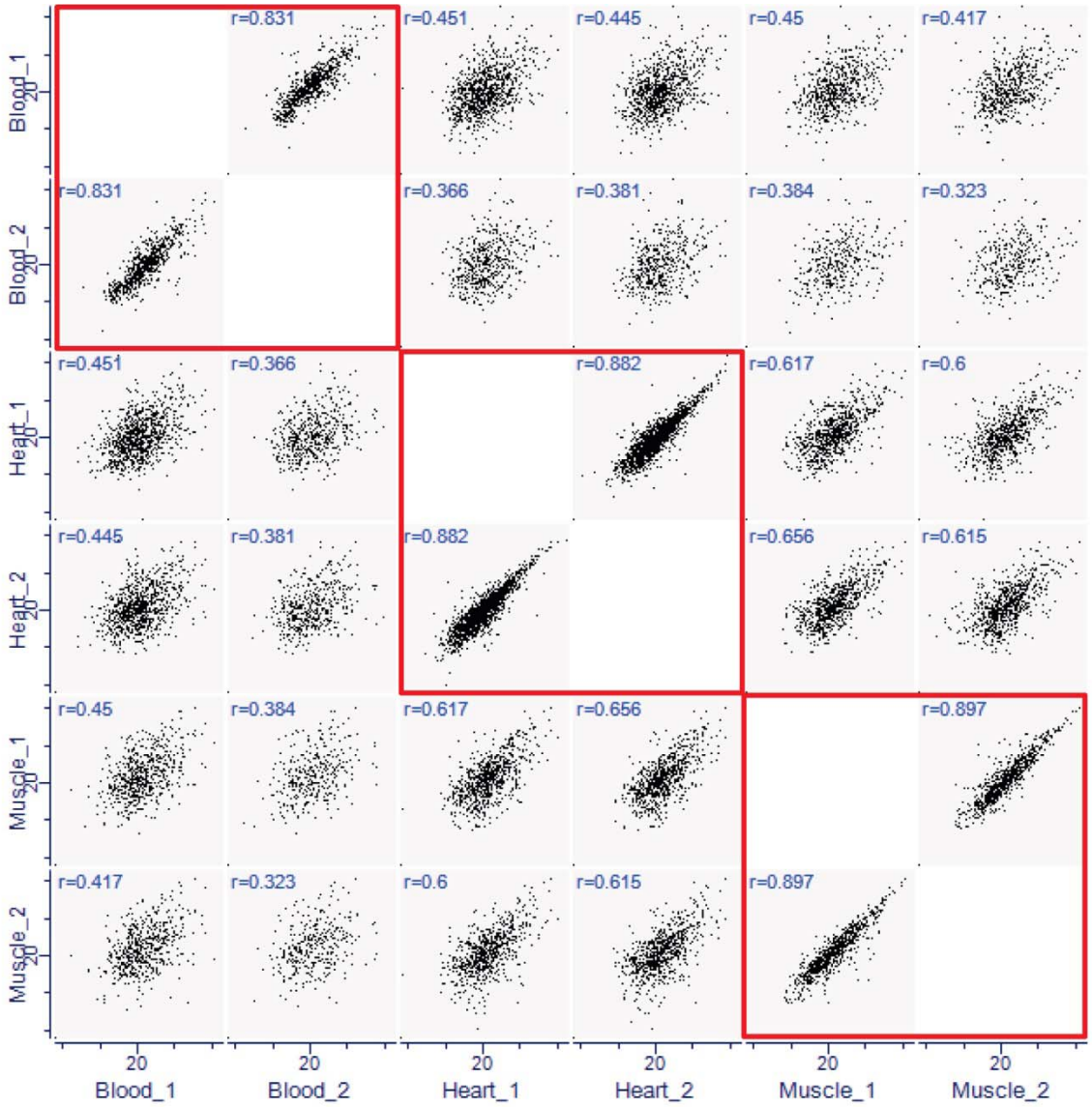






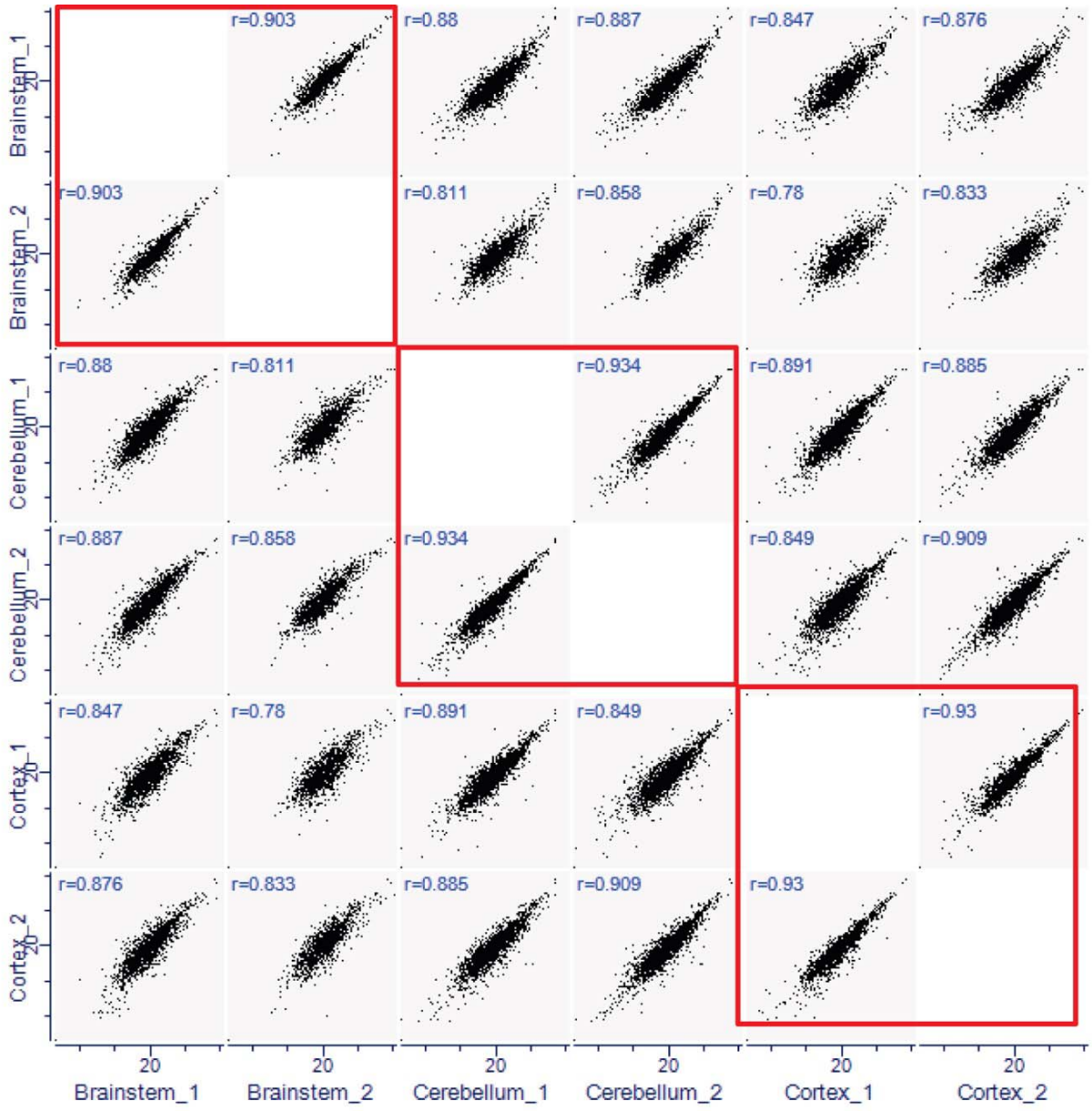
Supplementary Figure S1| Evaluation and assessment of MS data quality. a. Combined statistics for all 43 raw-files with fraction of identified MS/MS events (peptide-spectrum matches) and phosphopeptides versus unphosphorylated peptides. **b**. Phosphosite distribution of identified unique phosphopeptides. Histogram displays number of phosphosites per identified peptide. **c**. Histogram illustrating the Mascot score distribution of all modified peptides. Mascot scores are binned in 10-score units and each bin is displayed as a bar indicating the peptide count. **d**. Histogram of Andromeda score distribution of all modified peptides. Andromeda scores are binned in 25- score units. **e**. Peptide mass accuracy. Calibrated precursor mass errors measured for all peptides in parts-per-million (ppm). **f.** Fragment ion mass accuracy. Histogram of fragment ion errors binned in 5-ppm units. **g**. Distribution of peptide precursor intensities. Histogram of normalized and log10-based XIC-based intensities for all identified peptides. Peptide intensities are binned in log10-magnitudes. **h**. Spectral count distribution. Histogram of HCD-MS/MS spectra c ount per identified peptide.

Supplementary Figure S2| Reproducibility of phosphoprotein identification between first and second enrichment steps.



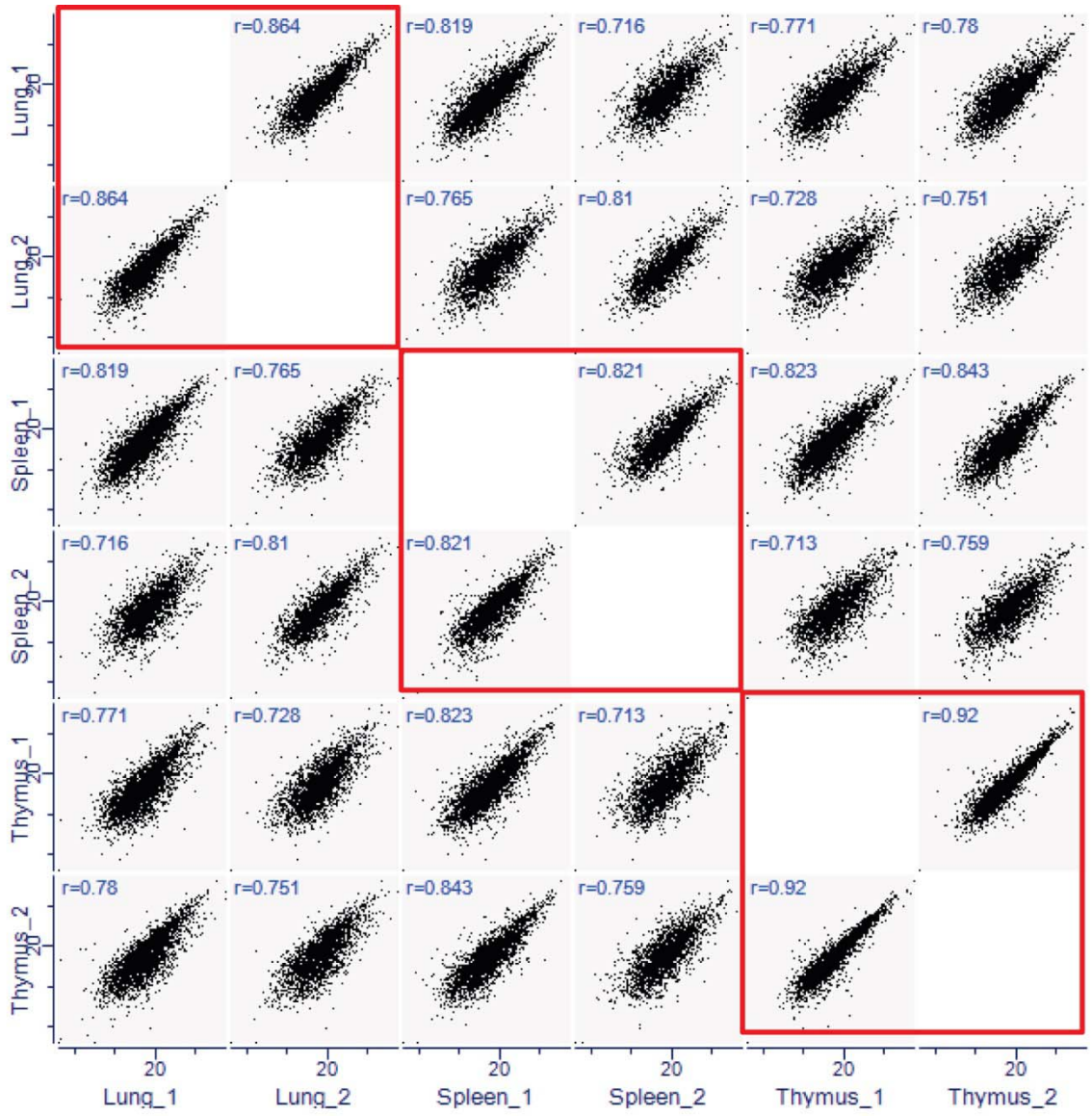
Supplementary Figure S2 | For each tissue two steps of enrichments was performed. The normalized phosphoprotein intensities calculated by summing all individual phosphopeptide intensities of the identified phosphoproteins were plotted against each other for the two enrichment steps for each tissue and the Pearson correlation coefficient was calculated for each plot, and is indicated in the upper left corner. One figure was made for each of the clusters found in the hierarchical clustering of phosphoproteins presented in Fig. 2a. This figure is for the blood, skeletal muscle and heart samples.

Supplementary Figure S3| Reproducibility of phosphoprotein identification between first and second enrichment steps.



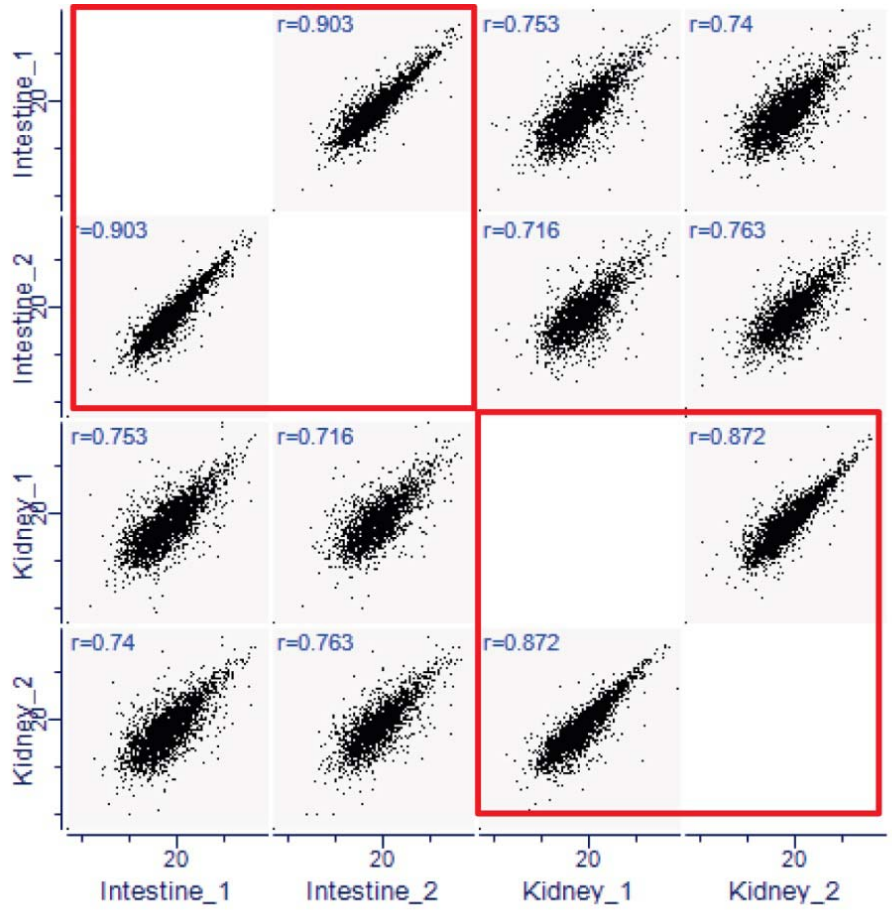
Supplementary Figure S3| For each tissue two steps of enrichments was performed. The normalized phosphoprotein intensities calculated by summing all individual phosphopeptide intensities of the identified phosphoproteins were plotted against each other for the two enrichment steps for each tissue and the Pearson correlation coefficient was calculated for each plot, and is indicated in the upper left corner. One figure was made for each of the clusters found in the hierarchical clustering of phosphoproteins presented in Fig. 2a. This figure is for the three brain parts investigated.

Supplementary Figure S4| Reproducibility of phosphoprotein identification between first and second enrichment steps.



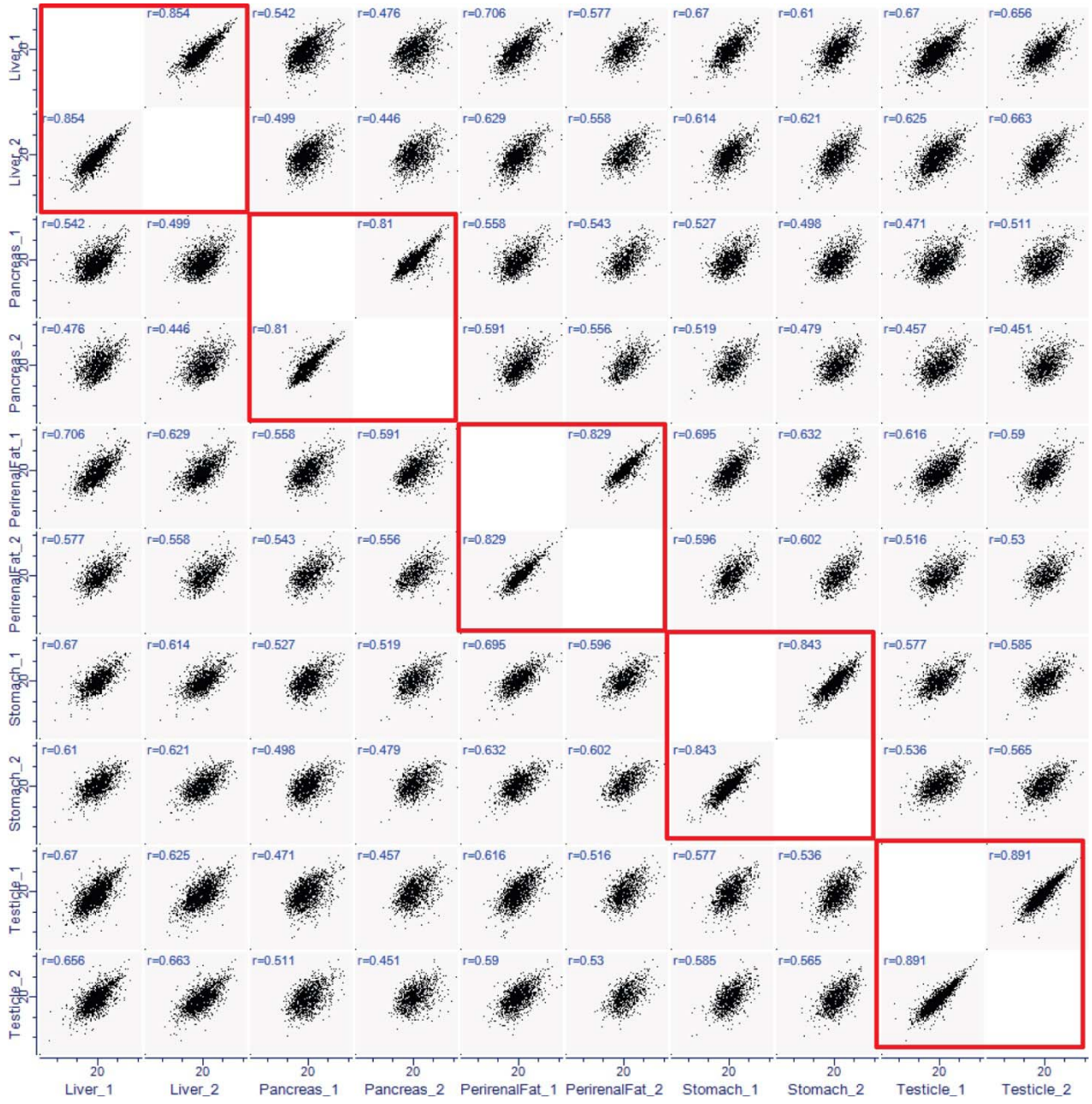
Supplementary Figure S4| For each tissue two steps of enrichments was performed. The normalized phosphoprotein intensities calculated by summing all individual phosphopeptide intensities of the identified phosphoproteins were plotted against each other for the two enrichment steps for each tissue and the Pearson correlation coefficient was calculated for each plot, and is indicated in the upper left corner. One figure was made for each of the clusters found in the hierarchical clustering of phosphoproteins presented in Fig. 2a. This figure is for the lung, spleen and thymus samples investigated.

Supplementary Figure S5| Reproducibility of phosphoprotein identification between first and second enrichment steps.



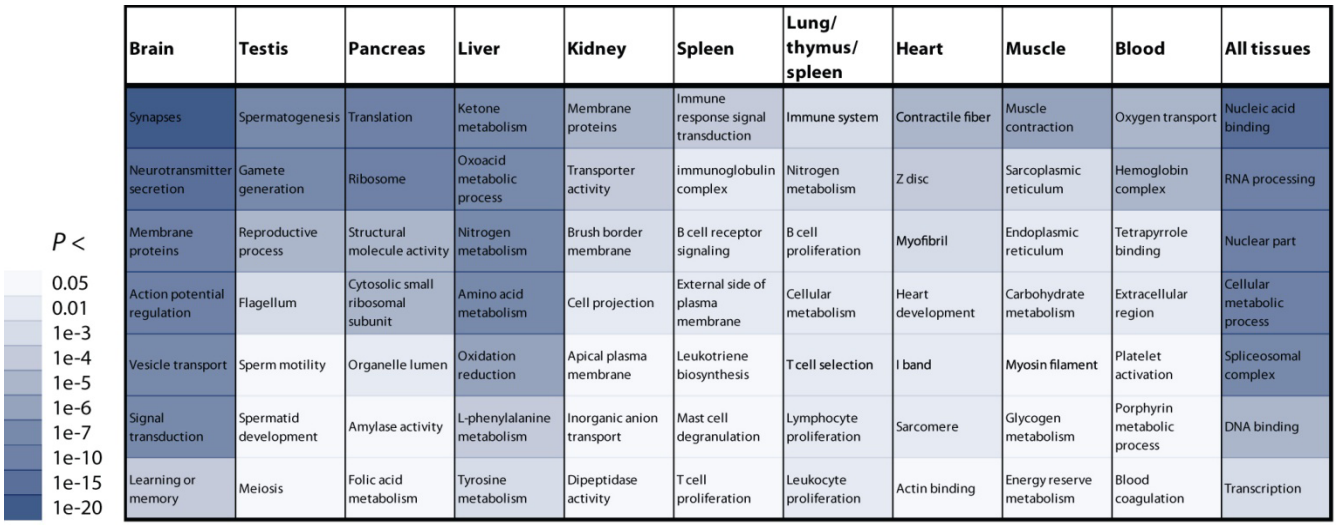
Supplementary Figure S5| For each tissue two steps of enrichments was performed. The normalized phosphoprotein intensities calculated by summing all individual phosphopeptide intensities of the identified phosphoproteins were plotted against each other for the two enrichment steps for each tissue and the Pearson correlation coefficient was calculated for each plot, and is indicated in the upper left corner. One figure was made for each of the clusters found in the hierarchical clustering of phosphoproteins presented in Fig. 2a. This figure is for the kidney and intestine samples investigated.

Supplementary Figure S6| Reproducibility of phosphoprotein identification between first and second enrichment steps.



Supplementary Figure S6| For each tissue two steps of enrichments was performed. The normalized phosphoprotein intensities calculated by summing all individual phosphopeptide intensities of the identified phosphoproteins were plotted against each other for the two enrichment steps for each tissue and the Pearson correlation coefficient was calculated for each plot, and is indicated in the upper left corner. One figure was made for each of the clusters found in the hierarchical clustering of phosphoproteins presented in Fig. 2a. This figure is for the liver, pancreas, perirenal fat, stomach and testis samples investigated.

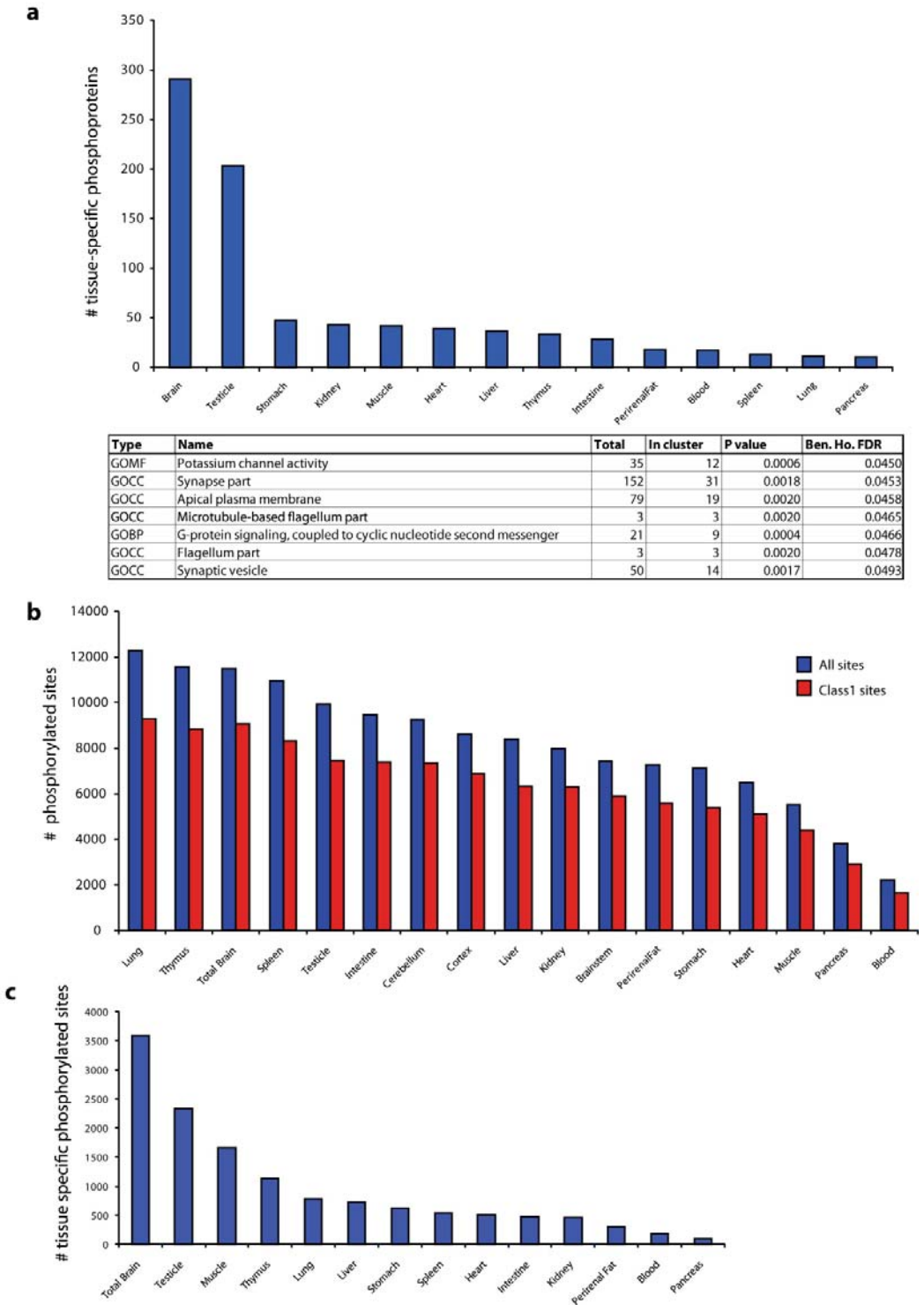
Supplementary Figure S7| Tissue-specific gene ontology analysis of phosphoproteins.



Brain	Testis	Pancreas	Liver	Kidney	Spleen	Lung/ thymus/ spleen	Heart	Muscle	Blood	All tissues
Synapses	Spermatogenesis	Translation	Ketone metabolism	Membrane proteins	Immune response signal transduction	Immune system	Contractile fiber	Muscle contraction	Oxygen transport	Nucleic acid binding
Neurotransmitter secretion	Gamete generation	Ribosome	Oxoacid metabolic process	Transporter activity	immunoglobulin complex	Nitrogen metabolism	Z disc	Sarcoplasmic reticulum	Hemoglobin complex	RNA processing
Membrane proteins	Reproductive process	Structural molecule activity	Nitrogen metabolism	Brush border membrane	B cell receptor signaling	B cell proliferation	Myofibril	Endoplasmic reticulum	Tetrapyrrole binding	Nuclear part
Action potential regulation	Flagellum	Cytosolic small ribosomal subunit	Amino acid metabolism	Cell projection	External side of plasma membrane	Cellular metabolism	Heart development	Carbohydrate metabolism	Extracellular region	Cellular metabolic process
Vesicle transport	Sperm motility	Organelle lumen	Oxidation reduction	Apical plasma membrane	Leukotriene biosynthesis	T cell selection	I band	Myosin filament	Platelet activation	Spliceosomal complex
Signal transduction	Spermatid development	Amylase activity	L-phenylalanine metabolism	Inorganic anion transport	Mast cell degranulation	Lymphocyte proliferation	Sarcomere	Glycogen metabolism	Porphyrin metabolic process	DNA binding
Learning or memory	Meiosis	Folic acid metabolism	Tyrosine metabolism	Dipeptidase activity	T cell proliferation	Leukocyte proliferation	Actin binding	Energy reserve metabolism	Blood coagulation	Transcription

Supplementary Figure S7| Gene ontology enrichment analysis was made for tissue-specific phosphoproteins by comparing the phosphoproteins encircled by the red boxes to all identified phosphoproteins. P-values for over-representation in a tissue-specific cluster was calculated with the Wilcoxon-Mann-Whitney test and corrected for multiple testing with a Benjamini Hochberg false discovery rate test. Enriched GO-terms are listed and their corresponding P-values are color coded according to the scale.

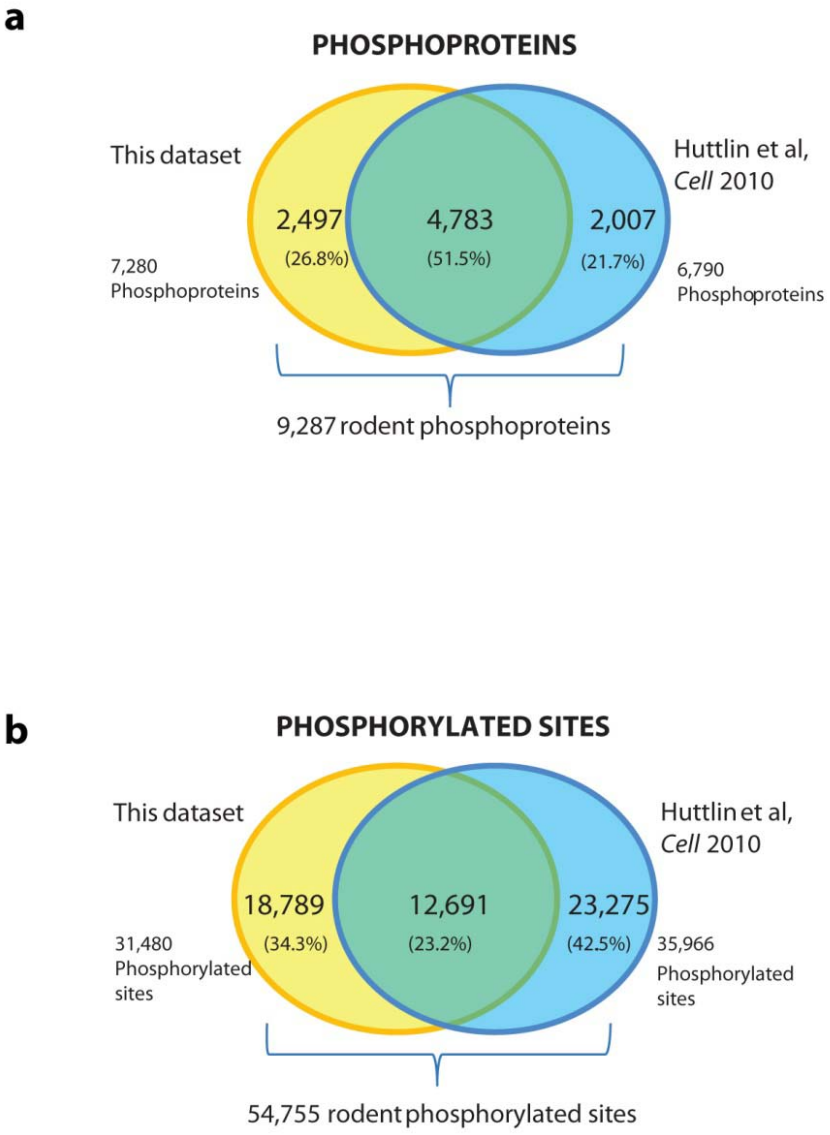
Supplementary Figure S8| Tissue distribution of phosphoproteins and phosphorylation sites.



Type	Name	Total	In cluster	P value	Ben. Ho. FDR
GOMF	Potassium channel activity	35	12	0.0006	0.0450
GOCC	Synapse part	152	31	0.0018	0.0453
GOCC	Apical plasma membrane	79	19	0.0020	0.0458
GOCC	Microtubule-based flagellum part	3	3	0.0020	0.0465
GOBP	G-protein signaling, coupled to cyclic nucleotide second messenger	21	9	0.0004	0.0466
GOCC	Flagellum part	3	3	0.0020	0.0478
GOCC	Synaptic vesicle	50	14	0.0017	0.0493

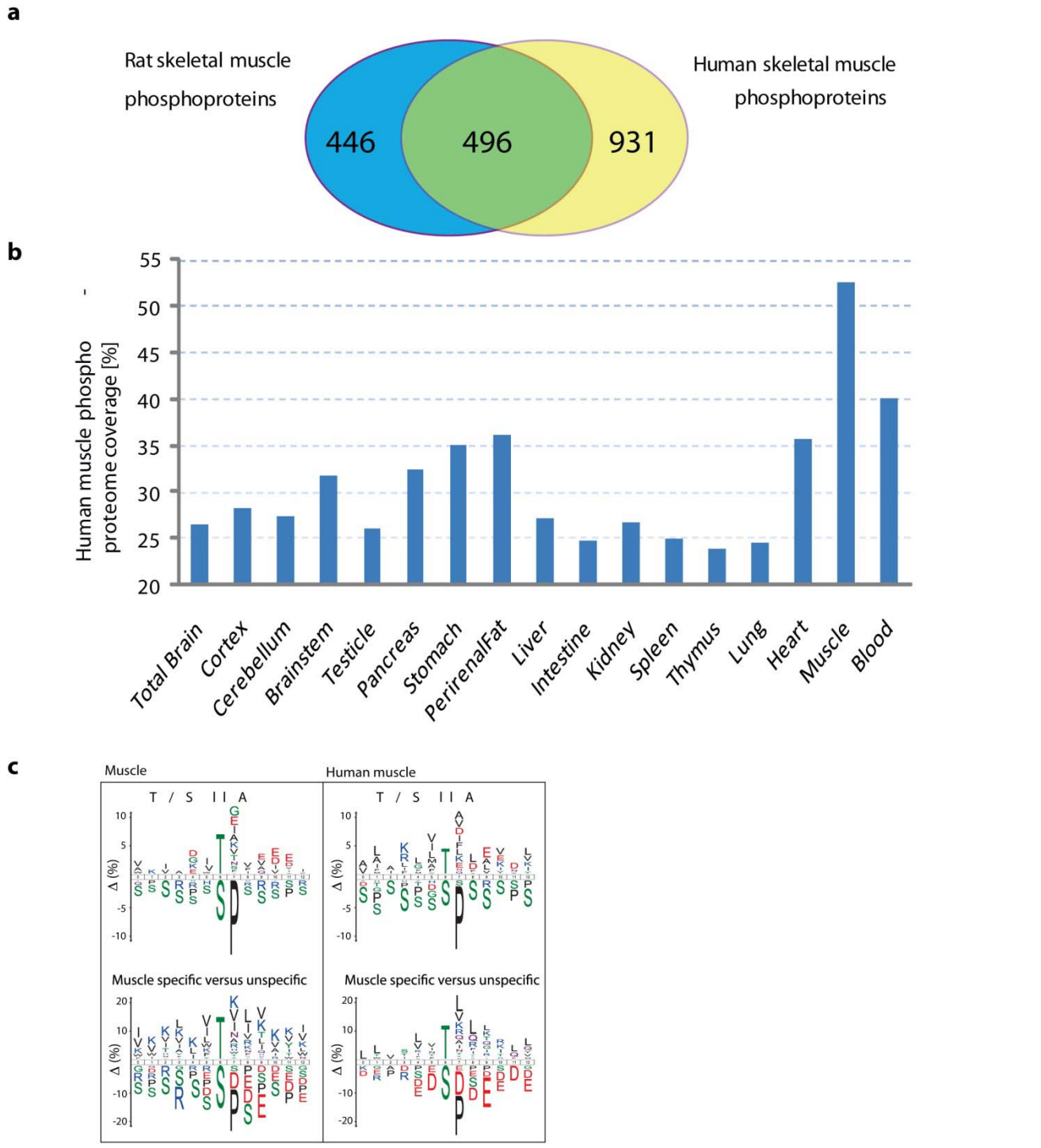
Supplementary Figure S8| a. The number of tissue-specific phosphoproteins are depicted in the histogram for all tissues investigated. Merging all the tissue-specific phosphoproteins and performing a GO term enrichment analysis relative to all other phosphoproteins in the dataset, reveal the terms indicated in the table to be enriched. The enriched terms are primarily associated with brain and testis proteins. b. Histogram depicting the number of phosphorylation sites (total number in blue and localized sites in red) identified in each tissue investigated. c. Histogram depicting the number of tissue-specific phosphorylation sites in each of the investigated tissues.

Supplementary Figure S9| Comparison of rat and mouse phosphoproteome atlases.



Supplementary Figure S9| We compared our combined rat phospho-proteome dataset from 14 tissues with the large-scale phosphoproteome dataset derived from nine mouse tissues (Huttlin et al, Cell 2010). **a**. Phosphoproteomes overlap between our rat dataset and the mouse phosphoproteome dataset based on Uniprot gene ortholog matching. **b**. Phosphorylation sites overlap between our rat dataset and the mouse phosphoproteome dataset based on sequence (sequence windows of ±6 residues adjacent to all phosphorylation sites) and tryptic peptide matching.

Supplementary Figure S10| Comparison of skeletal muscle phosphoproteome from rat and human samples.



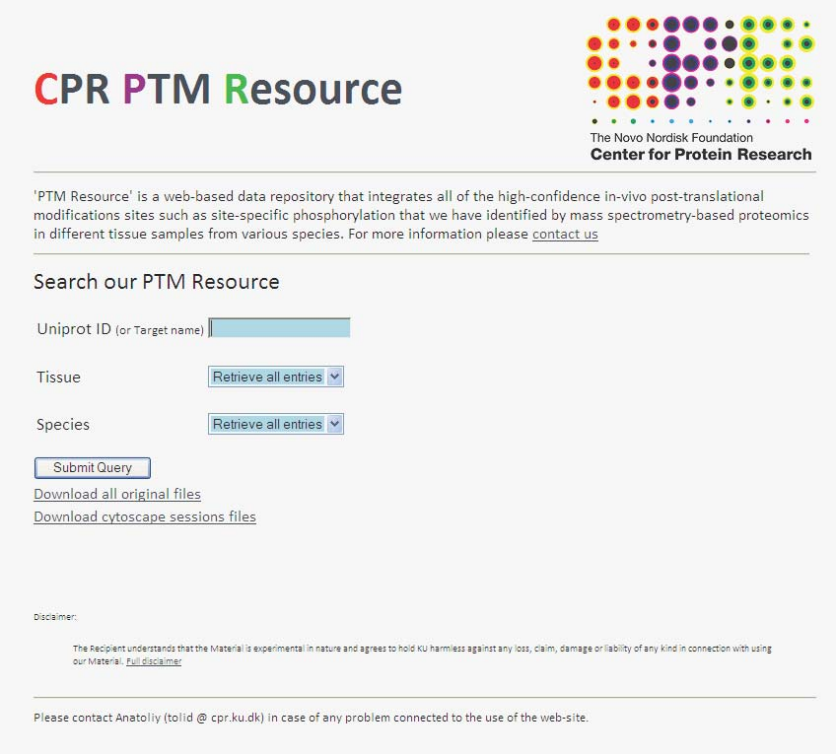
Supplementary Figure S10| a. Phosphoprotein overlap between rat and human skeletal muscle samples based on Uniprot gene ortholog matching. **b**. Percent overlap of human muscle phosphoproteins with all rat tissues tested. The greatest overlap is found for rat skeletal muscle samples. **c**. Sequence pattern analysis of rat and human S and T phosphorylation sites reveal similar amino acid sequence patterns. Top: Amino acid sequence pattern for S and T phosphorylation sites identified in rat or human muscle. Bottom: Amino acid sequence analysis of muscle specific S and T phosphorylation sites versus non-specific muscle phosphorylation sites for rat and human.

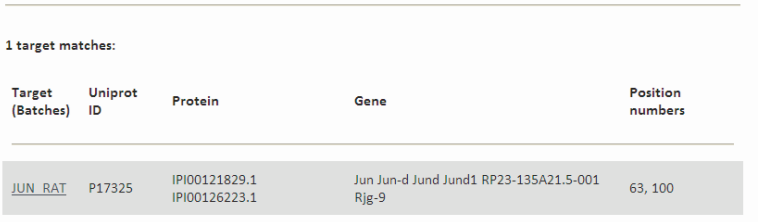
Supplementary Figure S11| Table of tissue-specific proline-directed phosphorylation sites of transcription factors.

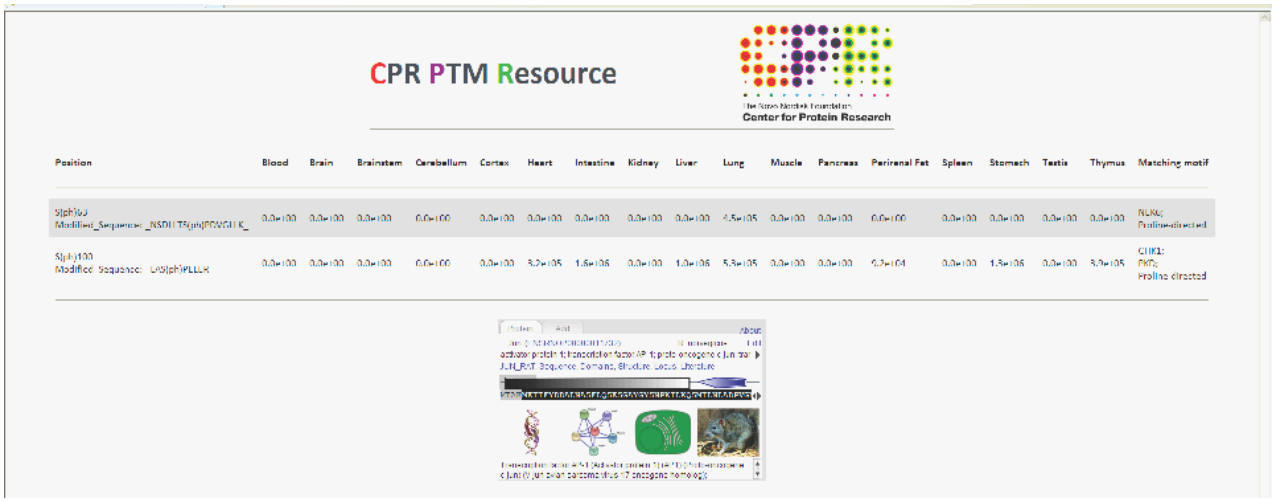
TF Gene Names	Phosphosite(s)	Phosphopeptide Sequence	Tissue(s) observed
Mlf1ip	T174	PAEPVT(ph)PR	Cortex
Tbr1	S646 & S657	RIS(ph)PADTPVSESSS(ph)PLK	Cortex
Tbr1	S84	SKLS(ph)PVLDGVSELR	Cortex
Fosl2	S200	S(ph)PPTSGLQSLR	Intestine & Stomach
Foxc2	S231 & S239	VETLS(ph)PEGALQAS(ph)PR	Kidney
Tcf20	S603	LSTS(ph)PATR	Kidney
Jun	S63	NSDLLTS(ph)PDVGLLK	Lung
Tead3	S572	FSPPS(ph)PLPQAVFSTSR	Lung
Hmga1	T42	EPSEVPT(ph)PK	Lung & Spleen
Tcf4	S393	NGGQASSS(ph)PNYEGPLHSLQSR	Lung & Thymus
Runx1	S14 & S21	RFT(ph)PPSTALS(ph)PGK	Lung, Spleen & Thymus
Mycn	S58 & S62	KFELLPT(ph)PPLS(ph)PSR	Spleen
Pou2f2	S26	(gl)QSLDS(ph)PSEHTDTER	Spleen
Pou2f2	S50 & S55	AS(ph)PFSVS(ph)PTGPSTK	Spleen
Foxn3	S85	SVS(ph)PVQDLDDDTPPSPAHSDMPYDAR	Spleen & Thymus
Tcf3	S140	DTSVGTLSQAGFLPGELGLSSPGPLS(ph)PSGVK	Spleen & Thymus
Aff1	S764 & S769	LLS(ph)PLRDS(ph)PPPTSLVVK	Thymus
Cbfa2t3	S297	RSCTLS(ph)PAQR	Thymus
Cbfa2t3	S39	AAAMPDS(ph)PAEVK	Thymus
E2f8	S71	MLISAVS(ph)PEIR	Thymus
Runx1	S212	VS(ph)PHHPAPTPNPR	Thymus
Runx1	S275	QIQPS(ph)PPWSYDQSYQYLGSITSSVHPATPIS(ph)PGR	Thymus
Mbd1	S440	QLQLSS(ph)PLK	Testis & Thymus
Mybl1	T441	FST(ph)PPTILR	Testis

Supplementary Figure S11| List of all tissue specific transcription factor phosphopeptides identified with information about gene name, phosphosites, and phosphopeptide sequence and in which tissue the phosphopeptide was identified.

Supplementary Figure S12| Screen shots from the web-based CPR PTM Resource database







Supplementary Figure S12| Screenshots taken from: http://cpr1.sund.ku.dk/cgi-bin/ptm.pl