
Supporting Information:

Mouse primary T cell phosphotyrosine proteomics

enabled by BOOST

Xien Yu Chua¹, Kenneth P. Callahan², Alijah A. Griffith², Tobias Hildebrandt²,
Guoping Fu³, Mengzhou Hu¹, Renren Wen³, Arthur R. Salomon^{2,*}

1 Department of Molecular Pharmacology, Physiology & Biotechnology, Brown University, Providence, RI, 02912

2 Department of Molecular Biology, Cell Biology & Biochemistry, Brown University, Providence, RI, 02912

3 Blood Research Institute, Blood Center of Wisconsin, Milwaukee, WI, 53226
** Corresponding Author**

E-mail: art@drsalomon.com

Table Of Contents

Supporting Table 1:	(.XLSX) All unique peptides observed in the BOOST experiment without ΦSDM
Supporting Table 2:	(.XLSX) All unique peptides observed in the BOOST and 1.0 mg Control experiments without ΦSDM
Supporting Table 3:	(.XLSX) All unique peptides observed in the 1.0 mg Control experiment without ΦSDM
Supporting Table 4:	(.XLSX) All unique peptides observed in the BOOST experiment with ΦSDM
Supporting Table 5:	(.XLSX) All unique peptides observed in the BOOST and 1.0 mg Control experiments with ΦSDM
Supporting Table 6:	(.XLSX) All unique peptides observed in the 1.0 mg Control experiment with ΦSDM
Supporting Folder 2	(.ZIP) All Python3 code used for data analysis
Supporting Folder 1	(.ZIP) All tables output from MaxQuant
Supporting Figure 1:	Histogram of all PSMs containing a phosphorylated amino acid in all conditions binned by localization probability
Supporting Figure 2:	Missing value bar charts for each experiment
Supporting Figure 3:	\log_{10} transformed reporter intensity box-and-whisker plots
Supporting Figure 4:	Pairwise replicate comparisons of unique peptides identified in BOOST when ΦSDM is disabled
Supporting Figure 5:	Pairwise replicate comparisons of unique peptides identified in BOOST when ΦSDM is enabled
Supporting Figure 6:	Pairwise replicate comparisons of unique peptides identified in 1.0 mg Control when ΦSDM is disabled
Supporting Figure 7:	Pairwise replicate comparisons of unique peptides identified in 1.0 mg Control when ΦSDM is enabled
Supporting Figure 8:	Venn Diagram and Volcano Plots for BOOST and 1.0 mg Control when ΦSDM is disabled
Supporting Figure 9:	Venn Diagram and Volcano Plots for BOOST and 1.0 mg Control when ΦSDM is enabled
Supporting Figure 10:	Venn Diagrams for BOOST conditions (with and without ΦSDM) and 1.0 mg Control (with and without ΦSDM)
Supporting Figure 11:	Cumulative distributions of unique pTyr peptides from BOOST experiments (with and without ΦSDM)
Supporting Figure 12:	Overlap between Mouse-BOOST and Jurkat-BOOST experiments

Supporting Table 1: All unique peptides observed exclusively in the BOOST experiment with ΦSDM disabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values, BOOST factors) and WikiPathways¹ Annotations for each unique peptide.

Supporting Table 2: All unique peptides observed in both the BOOST experiment and the 1.0 mg Control experiment with ΦSDM disabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values, BOOST factors) and WikiPathways¹ Annotations for each unique peptide.

Supporting Table 3: All unique peptides observed exclusively in the 1.0 mg Control experiment with ΦSDM disabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values) and WikiPathways¹ Annotations for each unique peptide.

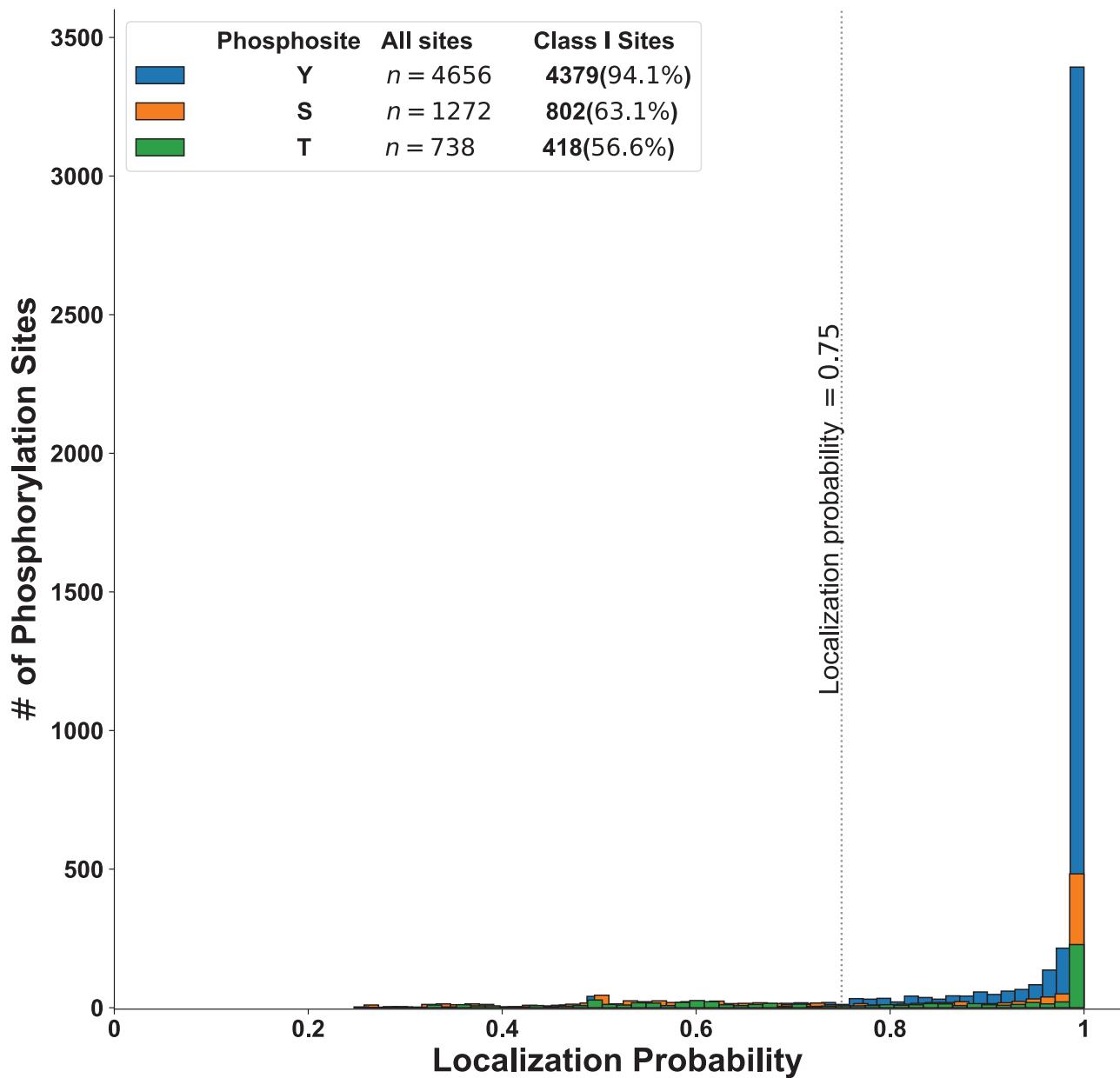
Supporting Table 4: All unique peptides observed exclusively in the BOOST experiment with ΦSDM enabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values, BOOST factors) and WikiPathways¹ Annotations for each unique peptide.

Supporting Table 5: All unique peptides observed in both the BOOST experiment and the 1.0 mg Control experiment with ΦSDM enabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values, BOOST factors) and WikiPathways¹ Annotations for each unique peptide.

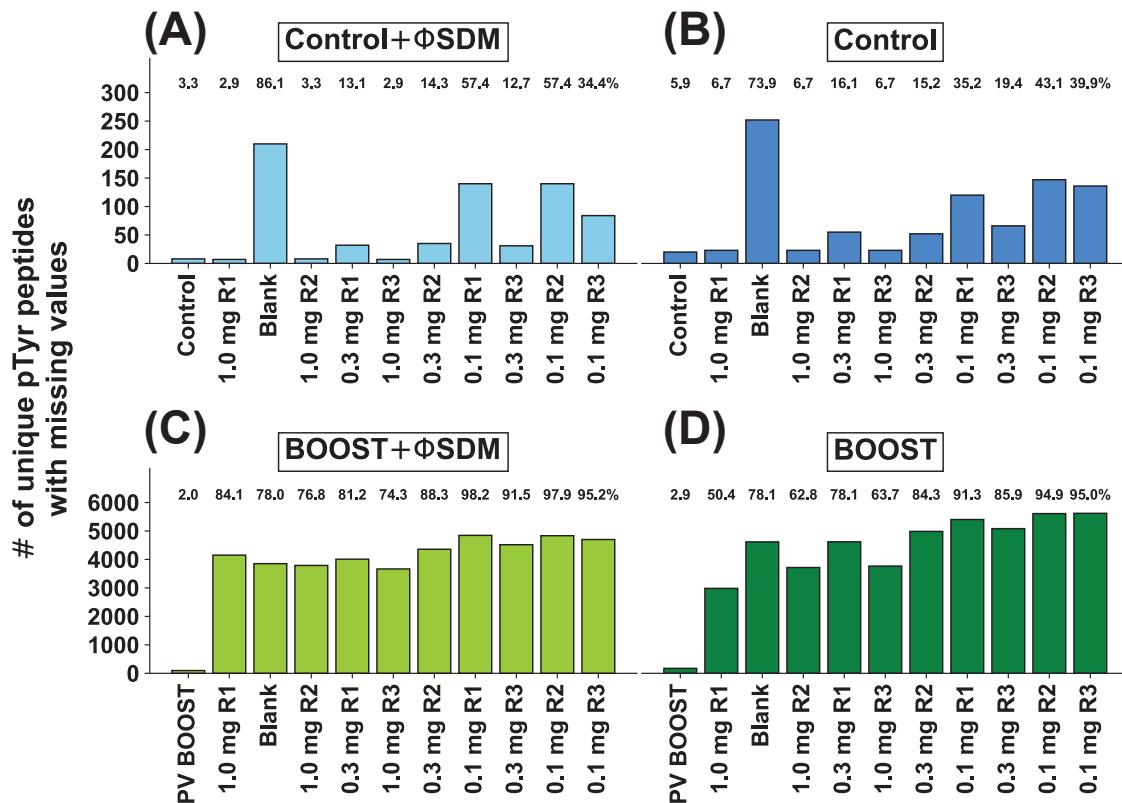
Supporting Table 6: All unique peptides observed exclusively in the 1.0 mg Control experiment with ΦSDM enabled and reporter ions quantified in at least one experimental channel. Included are MaxQuant outputs such as the gene name(s), protein name(s), modified sequence, modified sequence with phosphorylation site localization probabilities, sequence windows, and corrected TMT reporter ion intensities, as well as PhosphoSitePlus® site annotations, calculations used in the manuscript (mean intensity values, standard deviations, CV percentages, *p*-values, *q*-values) and WikiPathways¹ Annotations for each unique peptide.

Supporting Folder 1: All Python3 code used to analyze the MaxQuant output files and databases referenced. These include “data_analysis.py” (script used to generate plots), “helpers/” (Python3 files used to assist in data analysis), “database/” (all external databases used in analysis), and “maxquant_results” (the “evidence.txt” and “Phospho (STY)Sites.txt” files from Supporting Folder 2), as well as the output folders “figures/” (all figures generated by data_analysis.py) and “curated_results/” (all .txt output files from Python3 analysis, which are aggregated and formatted in Supporting Tables 1-6).

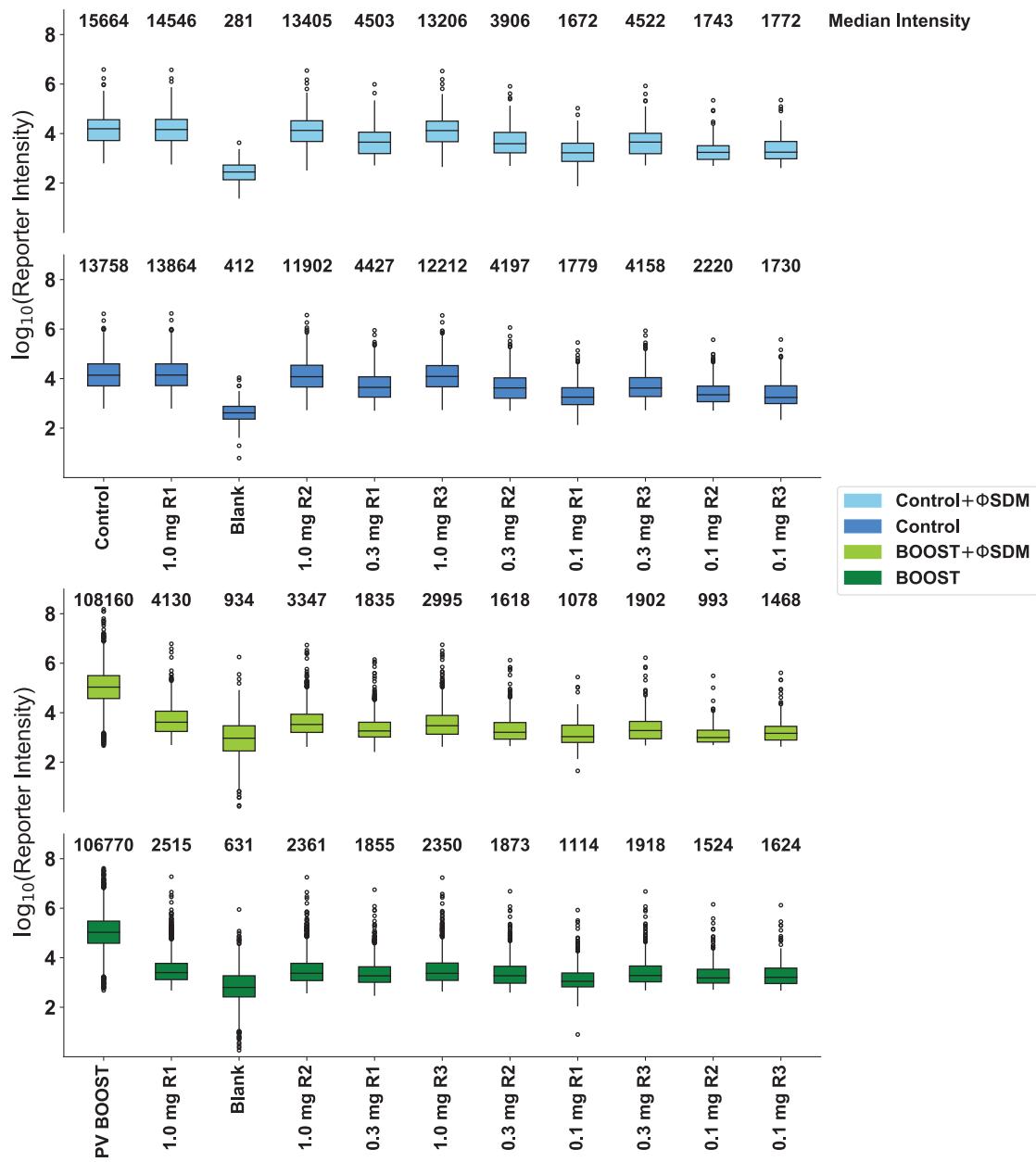
Supporting Folder 2: All tables generated by MaxQuant as text files. These include “summary.txt” (a summary of parameters, information, .raw files, and statistics used for peak detection), “evidence.txt” (all information about unique peptides quantified from .raw files), “peptides.txt” (information about the peptides identified from .raw files), “modification-SpecificPeptides.txt” (information about posttranslational modifications to the peptides), “Oxidation (M)Sites.txt” (information about oxidized peptides), “Phospho (STY)Sites.txt” (information about phosphorylated peptides), “proteinGroups.txt” (information about estimated protein abundance from the .raw files), “allPeptides.txt” (all information for each unique peptide identified in each .raw file), “msScans.txt” (information about the scans observed on the mass spectrometer), “mzRange.txt”, “msmsScans.txt” (information about the MS/MS scans for each .raw file), and “msms.txt” (information about the MS/MS spectra for each peptide identified in each .raw file).



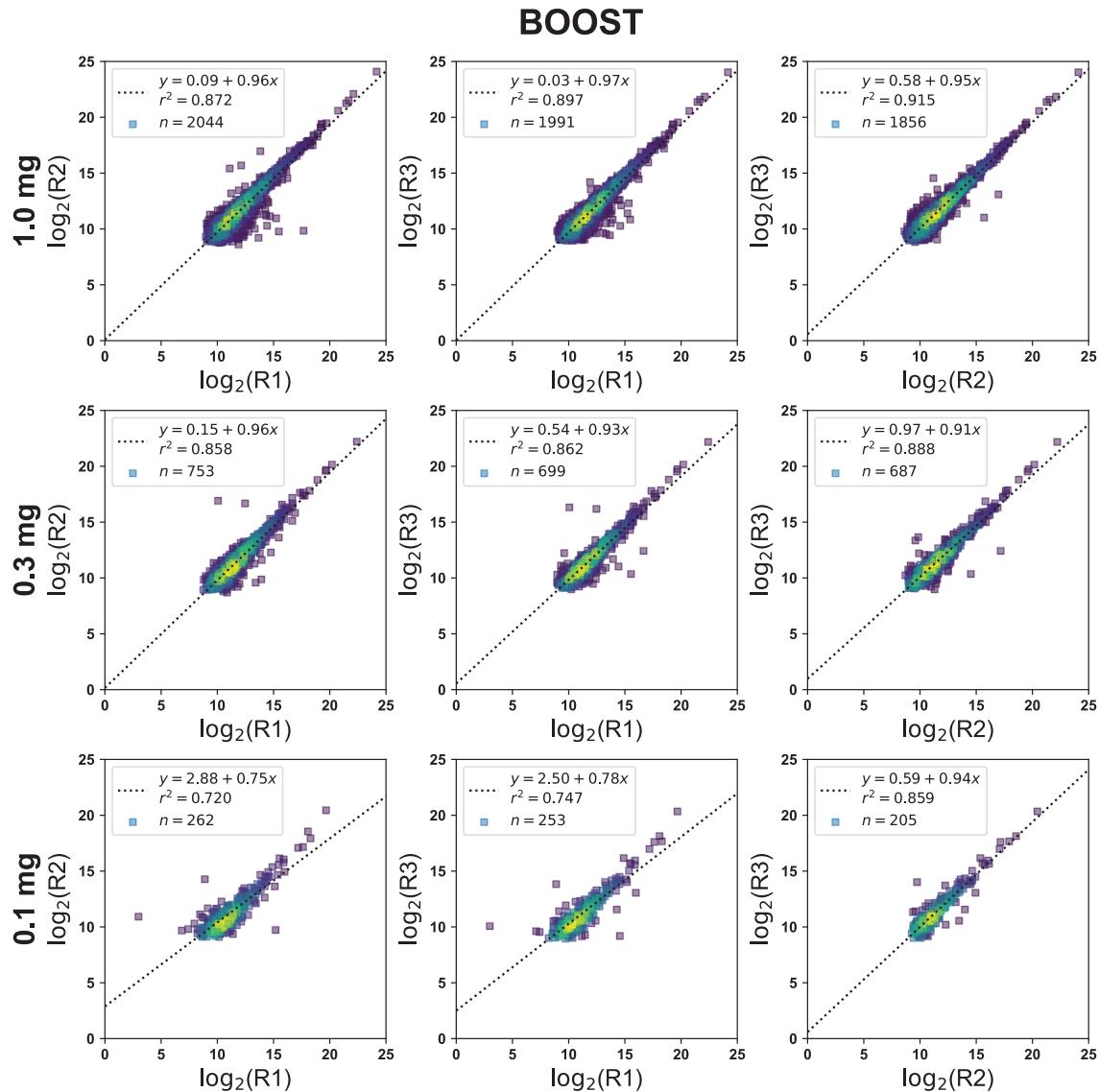
Supporting Figure 1: A histogram with depicting all PSMs from all experiments containing at least one phosphorylated serine (S), threonine (T), or tyrosine (Y) amino acid as a function of localization probability ($n_{bins} = 75$). The total number and number of Class I (localization probability > 0.75) phosphorylation sites for each amino acid are noted in the Figure Legend.



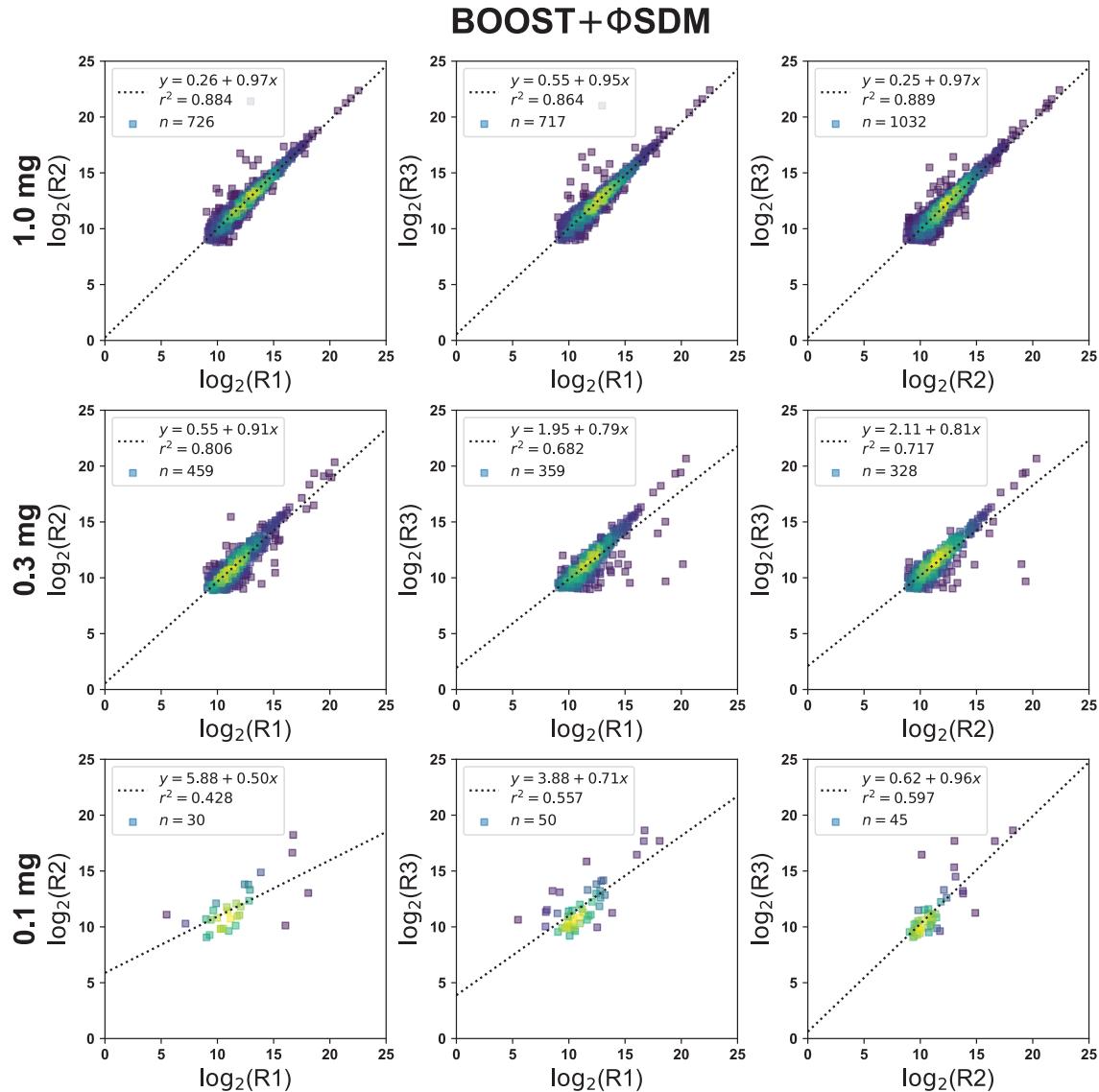
Supporting Figure 2: Disabling ΦSDM reduces the proportion of missing values in BOOST and 1.0 mg Control experiments. The number of missing values in each TMT channel for the (A) 1.0 mg Control experiment with ΦSDM enabled, (B) 1.0 mg Control experiment with ΦSDM disabled, (C) BOOST experiment with ΦSDM enabled, (D) BOOST experiment with ΦSDM disabled. The percentage of missing values in each TMT channel is indicated above each bar.



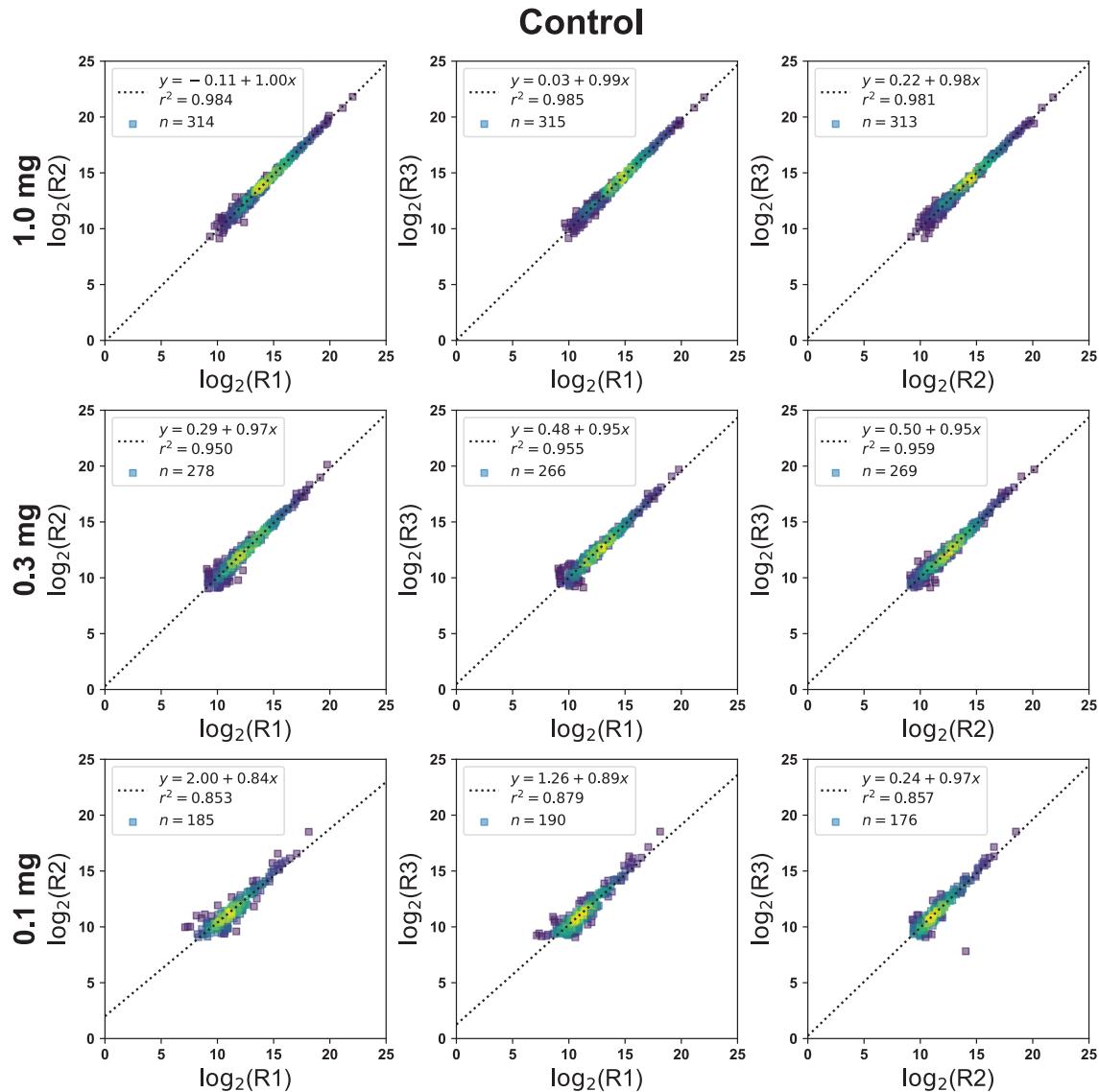
Supporting Figure 3: Box-and-whisker plots showing the \log_{10} transformed reporter intensities for each TMT mix and each condition. Non-transformed, median intensities are displayed above each box-and-whisker plot.



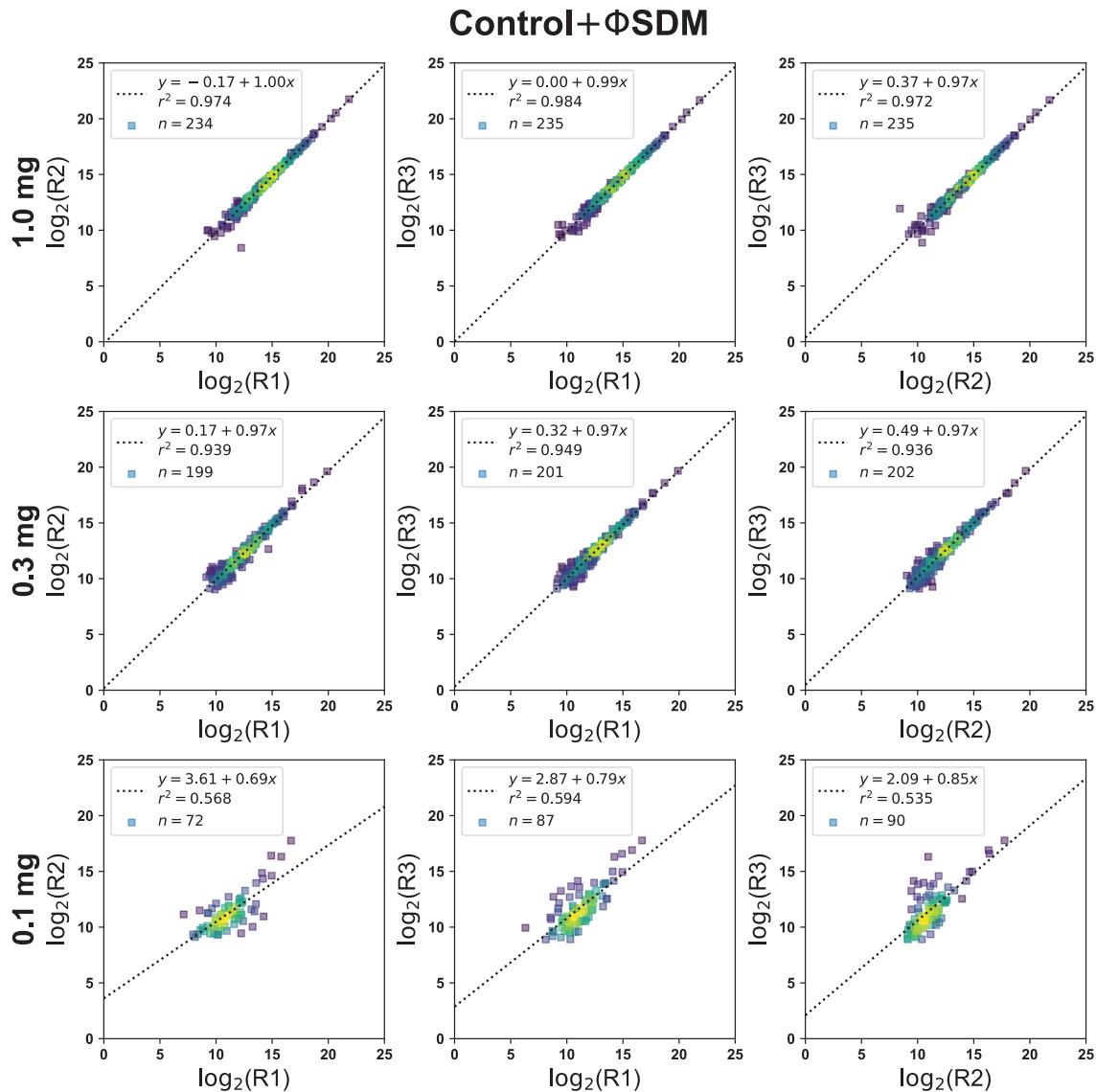
Supporting Figure 4: Replicate reproducibility is stable when ΦSDM is disabled for low protein input samples in the pervanadate BOOST condition. Evaluation of replicate reproducibility in the BOOST experiment (with ΦSDM disabled) using pairwise comparisons of \log_2 transformed abundances for phosphopeptides with the same charge state and assigned sequence. For each comparison, we show the line of best fit as determined by simple least-squares regression, the r^2 value as an estimate of the quality of the fitted line, and the total number of points (n) in each comparison.



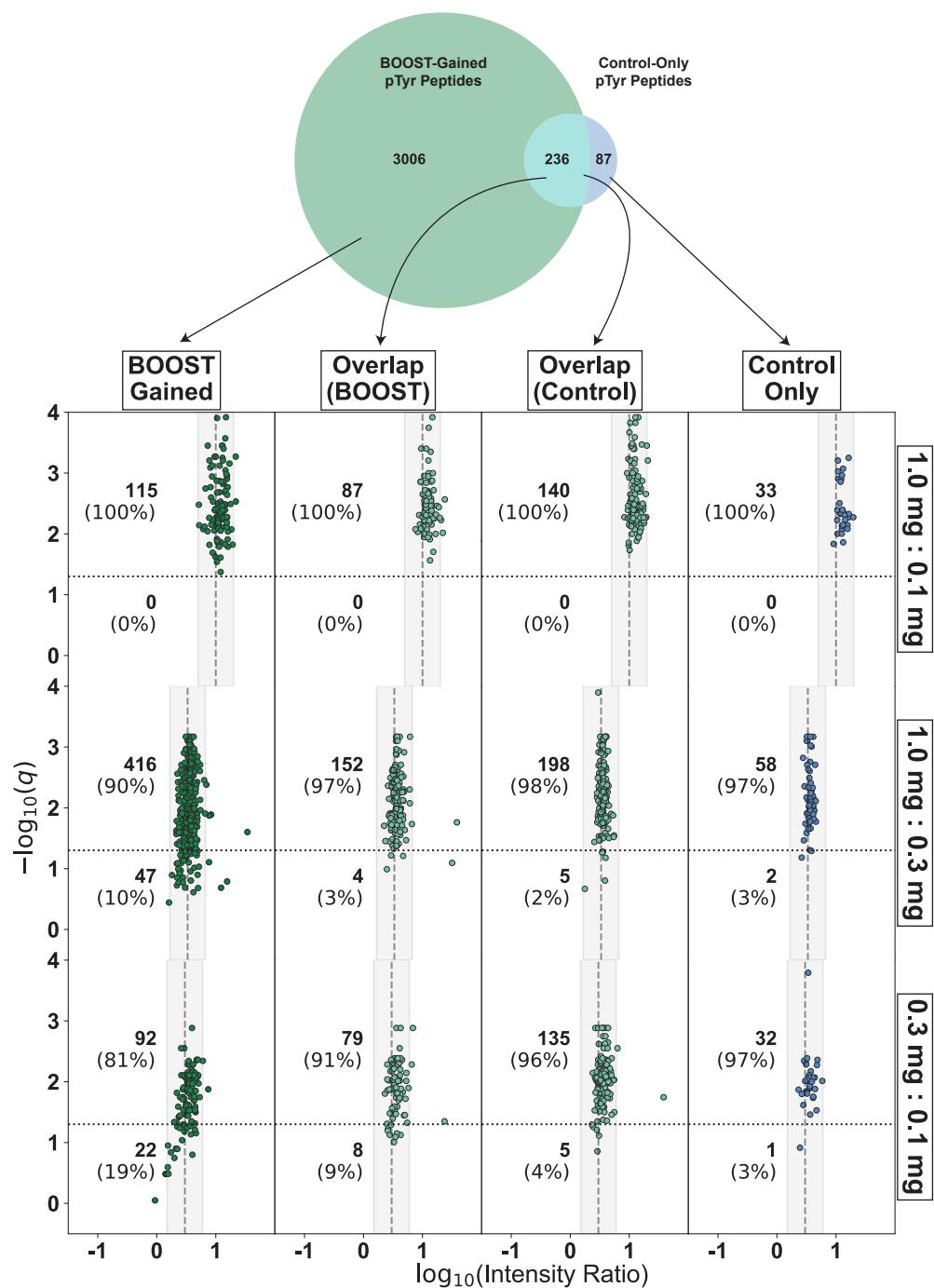
Supporting Figure 5: Replicate reproducibility is severely degraded when ΦSDM is enabled for low protein input samples in the pervanadate BOOST condition. Evaluation of replicate reproducibility in the BOOST experiment (with ΦSDM enabled) using pairwise comparisons of \log_2 transformed abundances for phosphopeptides with the same charge state and assigned sequence. For each comparison, we show the line of best fit as determined by simple least-squares regression, the r^2 value as an estimate of the quality of the fitted line, and the total number of points (n) in each comparison.



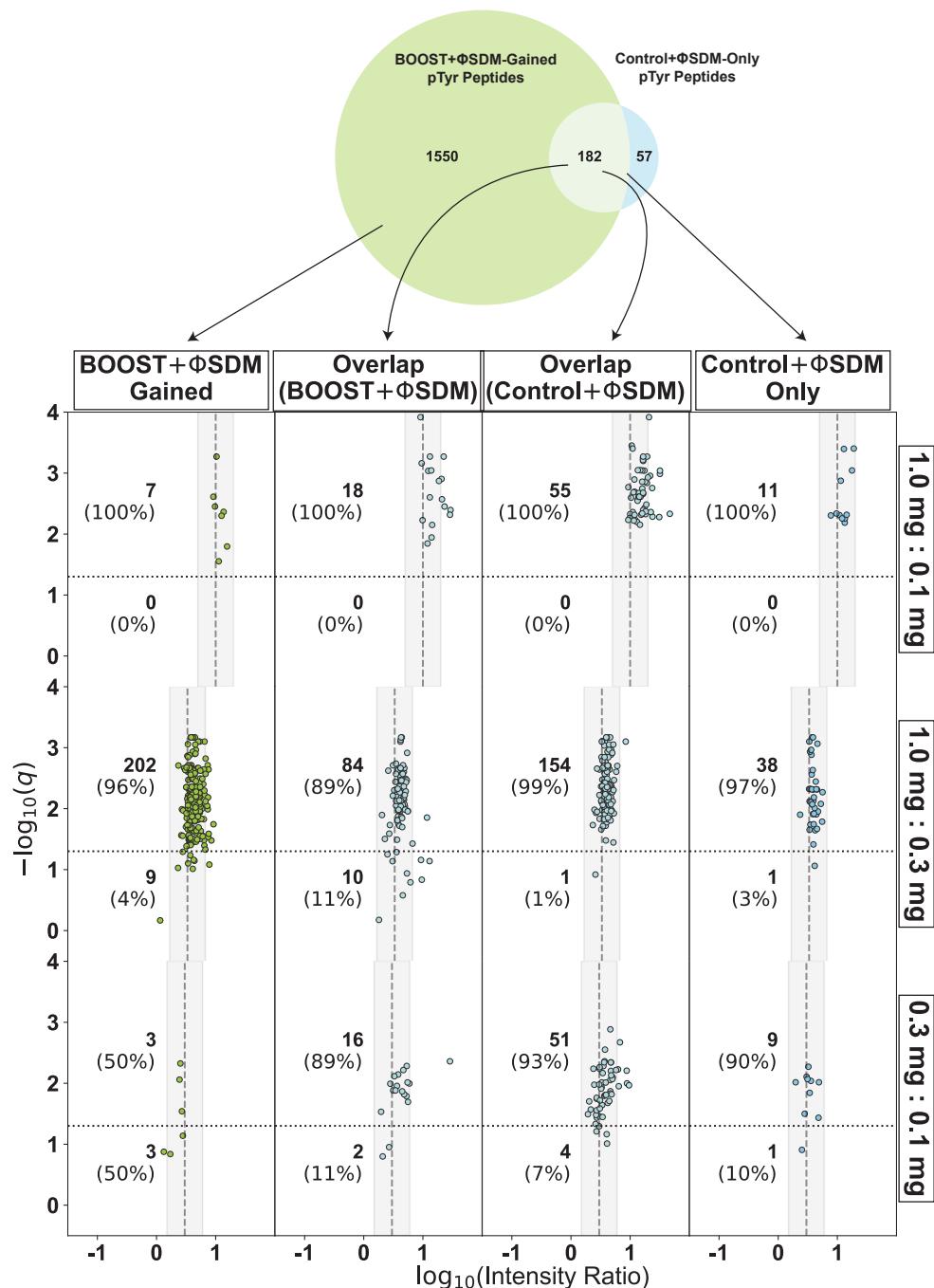
Supporting Figure 6: Replicate reproducibility is stable when ΦSDM is disabled for low protein input samples in the 1.0 mg Control condition. Evaluation of replicate reproducibility in the 1.0 mg Control experiment (with ΦSDM disabled) using pairwise comparisons of \log_2 transformed abundances for phosphopeptides with the same charge state and assigned sequence. For each comparison, we show the line of best fit as determined by simple least-squares regression, the r^2 value as an estimate of the quality of the fitted line, and the total number of points (n) in each comparison.



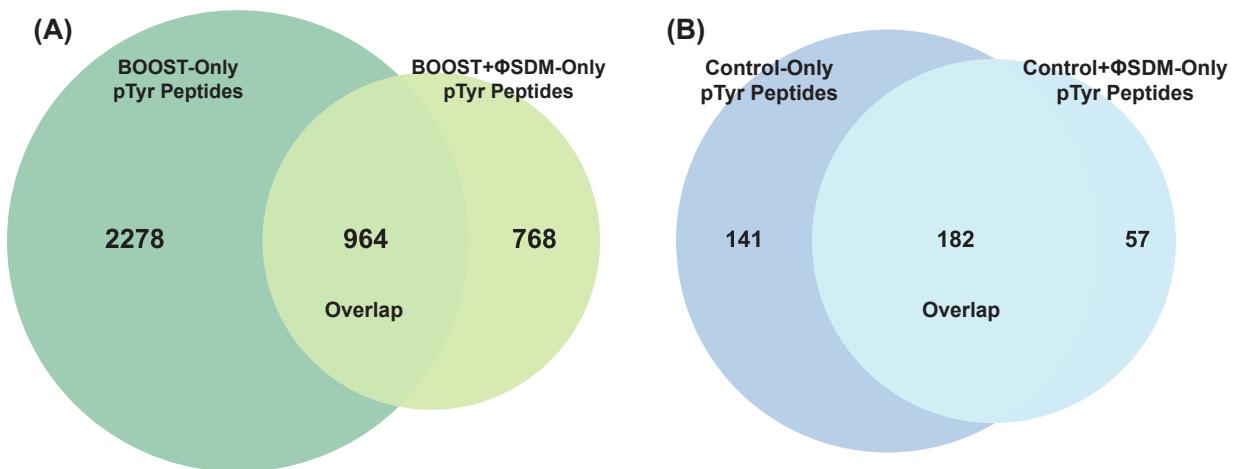
Supporting Figure 7: Replicate reproducibility is degraded when ΦSDM is enabled for low protein input samples in the 1.0 mg Control condition. Evaluation of replicate reproducibility in the 1.0 mg experiment (with ΦSDM enabled) using pairwise comparisons of \log_2 transformed abundances for phosphopeptides with the same charge state and assigned sequence. For each comparison, we show the line of best fit as determined by simple least-squares regression, the r^2 value as an estimate of the quality of the fitted line, and the total number of points (n) in each comparison.



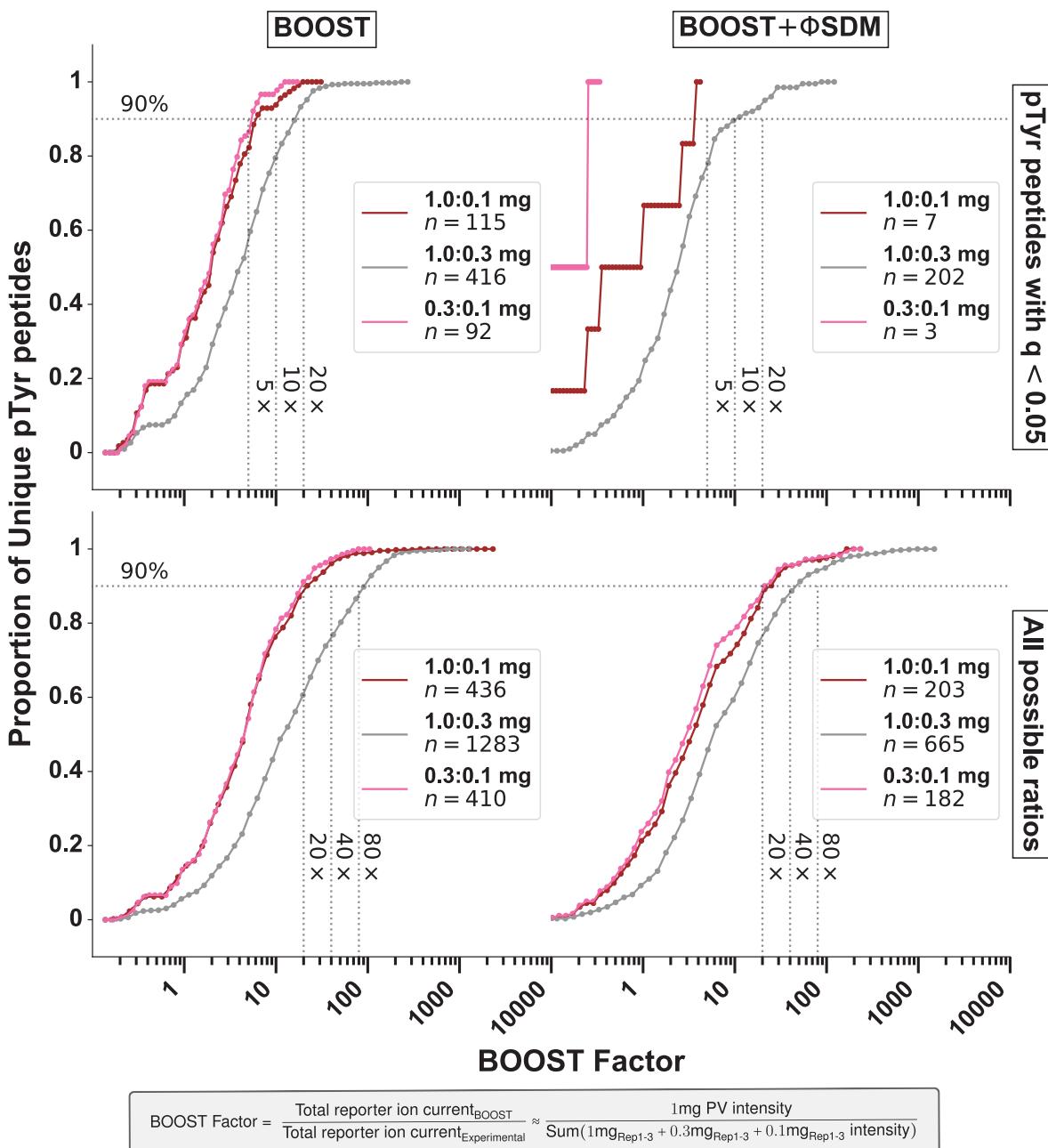
Supporting Figure 8: With Φ SDM disabled, the pervanadate BOOST channel dramatically increases the number of unique pTyr peptides observed as compared to a 1.0 mg Control channel. A Venn diagram showing the overlap of unique pTyr peptides between the BOOST and 1.0 mg Control experiments (with Φ SDM disabled). Volcano plots show $-\log_{10}(q\text{-value})$ as a function of $\log_{10}(\text{Intensity Ratio})$ for unique pTyr peptides from groups shown in the Venn diagram. For the overlapping section, volcano plots were created using data from both the BOOST experiment and the control experiment acquired with Φ SDM disabled.



Supporting Figure 9: The pervanadate BOOST channel increases the number of unique pTyr peptides observed when ΦSDM is enabled, although few peptides are observed in low abundance samples. A Venn diagram showing the overlap of unique pTyr peptides between the BOOST+ΦSDM and 1.0 mg Control+ΦSDM experiments. Volcano plots show $-\log_{10}(q\text{-value})$ as a function of $\log_{10}(\text{Intensity Ratio})$ for unique pTyr peptides from groups shown in the Venn diagram. For the overlapping section, volcano plots were created using data from both the BOOST experiment and the control experiment acquired with ΦSDM enabled.



Supporting Figure 10: Enabling ΦSDM results in lower yield in both pervanadate BOOST and 1.0 mg Control conditions. Venn diagrams showing the number of unique pTyr peptides observed when ΦSDM is enabled or disabled using (A) pervanadate BOOST samples, and (B) 1.0 mg Control samples.



Supporting Figure 11: Enabling ΦSDM decreases quantitation depth, particularly in low abundance samples. Cumulative distribution of BOOST factors for unique pTyr peptides identified in the pervanadate BOOST experiments with ΦSDM disabled or with ΦSDM enabled for pTyr peptides with a statistically significant ratio ($q < 0.05$) or for all calculable ratios. For each cumulative distribution, the range of BOOST factors are split into 50 bins of equal size on a \log_{10} scale.

Gene	Mouse	Human	Gene	Mouse	Human	Gene	Mouse	Human	Gene	Mouse	Human
AKT2	Y122	Y122	ITK	Y243	Y237	PIK3CD	Y523	Y524	PTPRC	Y852	Y861
CARD11	Y489	Y489		Y517	Y512	PIK3R1	Y76	Y76		Y871	Y880
CBLB	Y363	Y363	LAT	Y46	Y45		Y416	Y416		Y937	F946
	Y763	Y763		Y195	Y220		Y452	Y452		Y969	Y978
CD247	N64	Y64	LCK	Y192	Y192		Y467	Y467		F971	Y980
	Y72	Y72		Y394	Y394		Y470	Y470	RHOA	Y66	Y66
	Y83	Y83		Y414	Y414		Y688	Y688	TEC	Y205	Y206
	Y111	Y111		Y470	Y470	PIK3R2	Y458	Y464		Y227	Y228
	Y123	Y123		Y505	Y505	PIK3R3	Y202	Y202		Y280	Y281
	Y142	Y142	LCP2	Y173	Y173	PLCG1	Y210	Y210		Y518	Y519
	Y153	Y153		Y483	Y483		Y472	Y472	VAV1	Y110	Y110
CD28	Y189	Y191	MAP3K7	Y558	Y585		Y771	Y771		Y192	Y192
	Y204	Y206	MAPK1	Y185	Y187		Y775	Y775		Y541	Y541
	Y207	Y209	MAPK10	Y223	Y223		Y783	Y783		Y791	Y791
CD3D	Y149	Y149	MAPK11	Y182	Y182		Y1003	Y1003	VAV2	Y142	Y142
	Y160	Y160	MAPK12	Y185	Y185		Y1253	Y1253	VAV3	Y141	Y141
CD3E	Y170	Y188	MAPK14	Y182	Y182	PRKCQ	Y28	Y28		Y217	Y217
	Y181	Y199	MAPK3	Y205	Y204		Y545	Y545	ZAP70	Y69	Y69
CD3G	Y160	Y160	MAPK8	Y185	Y185	PTPN6	Y61	Y61		Y87	Y87
	Y171	Y171	MAPK9	Y185	Y185		Y64	Y64		Y164	Y164
CDC42	Y64	Y64	NCK1	Y13	Y13		Y213	Y213		Y178	Y178
CDK4	Y17	Y17		Y55	Y55		Y214	Y214		Y198	Y198
CTLA4	Y201	Y201		Y105	Y105		Y276	Y276		Y209	Y209
DLG1	Y399	Y399	NCK2	Y110	Y110		Y301	Y301		Y211	Y211
	Y761	Y760	NFATC2	Y754	Y752		Y306	Y306		Y221	Y221
	Y785	Y784	NFATC3	Y86	Y86		Y374	Y374		Y248	Y248
FYN	Y28	Y28		Y150	Y150		Y377	Y377		Y290	Y292
	Y214	Y214	NFKB1	Y238	Y240		Y536	Y536		Y314	Y315
	Y420	Y420	PAK1	Y142	Y142		Y541	Y541		Y396	Y397
	Y440	Y440		Y153	Y153		Y564	Y564		Y491	Y492
GRAP2	Y45	Y45		Y474	Y474	PTPRC	Y631	Y640		Y492	Y493
	Y218	Y222	PAK2	Y130	Y130		Y672	Y681		Y505	Y506
GRB2	Y209	Y209		Y139	Y139		Y678	Y687		Y596	Y597
GSK3B	Y216	Y216		Y453	Y453		Y680	Y689		Y597	Y598
ITK	Y40	Y40	PAK6	Y366	Y365		Y711	Y720			
	Y126	Y120	PDCD1	Y225	Y223		Y754	Y763			
	Y226	Y220	PIK3CA	Y317	Y317		Y781	Y790			

Supporting Figure 12: Comparison of pTyr sites identified using BOOST in primary T cells from mice and pTyr sites identified using BOOST in Jukat T cells. Our mouse data acquired without the ΦSDM enabled were compared to the data from Chua et al. collected in Jurkat T cells and acquired on an Orbitrap Fusion Lumos Tribrid mass spectrometer without the ΦSDM enabled. Flanking sequences (phosphorylation site ±7 amino acids) and phosphorylation sites for each protein in the Kyoto Encyclopedia of Genes and Genomes T cell receptor signaling pathway were manually curated from PhosphoSitePlus®³ for humans and mice. Flanking sequences for each peptide in the Mouse-BOOST and Jurkat-BOOST datasets were compared with the manually curated KEGG TCR/PhosphoSitePlus flanking sequences and filtered for unique sites. Gene names are colored purple. Phosphotyrosine sites identified in mice are colored red. Phosphotyrosine sites identified in Jurkat T cells are colored green. Sites that were not identified either mice or Jurkat T cells are colored grey.

References

- (1) Martens, M.; Ammar, A.; Riutta, A.; Waagmeester, A.; Slenter, D. N.; Hanspers, K.; A. Miller, R.; Digles, D.; Lopes, E. N.; Ehrhart, F., et al. WikiPathways: connecting communities. *Nucleic acids research* **2021**, *49*, D613–D621.
- (2) Chua, X. Y.; Mensah, T.; Aballo, T.; Mackintosh, S. G.; Edmondson, R. D.; Salomon, A. R. Tandem mass tag approach utilizing pervanadate BOOST channels delivers deeper quantitative characterization of the tyrosine phosphoproteome. *Molecular & Cellular Proteomics* **2020**, *19*, 730–743.
- (3) Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic acids research* **2015**, *43*, D512–D520.