

## **Appendix**

**Appendix Figure S1** – Decision tree for PSM searching, site localization, clustering, and search aggregation.

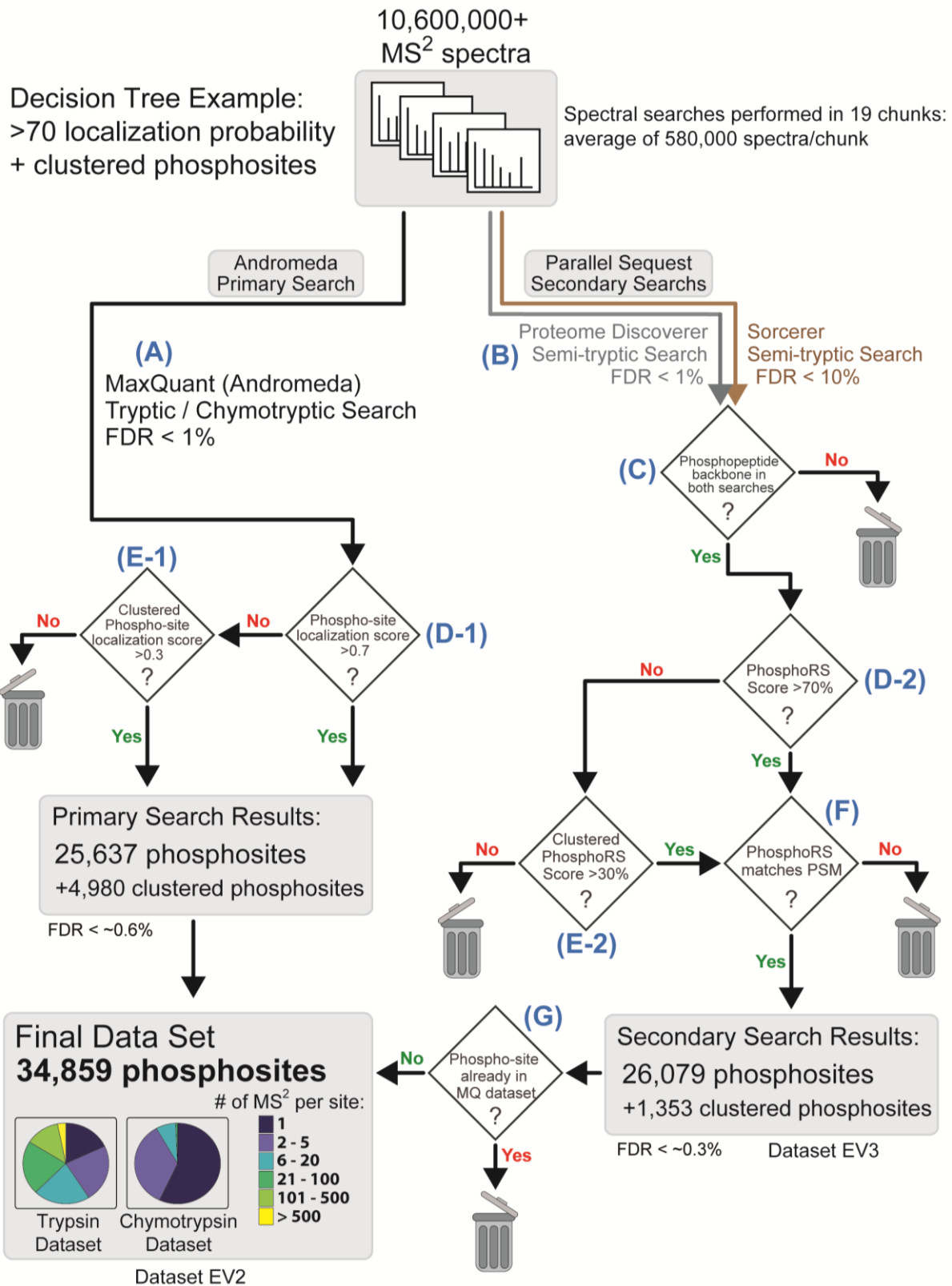
**Appendix Figure S2** – FDR analysis for primary, secondary, and combined searches.

**Appendix Figure S3** – Assessing phosphoproteome depth.

**Appendix Figure S4** – Scatterplots for two SILAC replicates and DNA damage in G2M.

**Appendix Figure S5** – Cell cycle analysis.

**Appendix Figure S6** – IP-MS of different Rad23 phosphomutants.



**Legend for Appendix Figure S1: Decision tree for PSM searching, site localization / clustering, and search aggregation.**

A decision tree for mapping the budding yeast phosphoproteome using mass spectrometry. Phosphopeptide fragments were captured as high resolution MS2 spectra using an orbitrap mass analyzer (Q-exactive). Three separate search engines were used to identify phosphopeptides from the fragmentation spectra.

(A) The primary search was performed using MaxQuant (Andromeda engine). Phosphosites from the primary search were extracted from MaxQuant's "Phospho STY" output table. A secondary search was performed using two separate Sequest-based engines, Proteome Discoverer (PD) and SORCERER.

(B) The secondary search utilized similar search parameters as the primary search, with the exception that tryptic enzyme digestion was set to semi-specific.

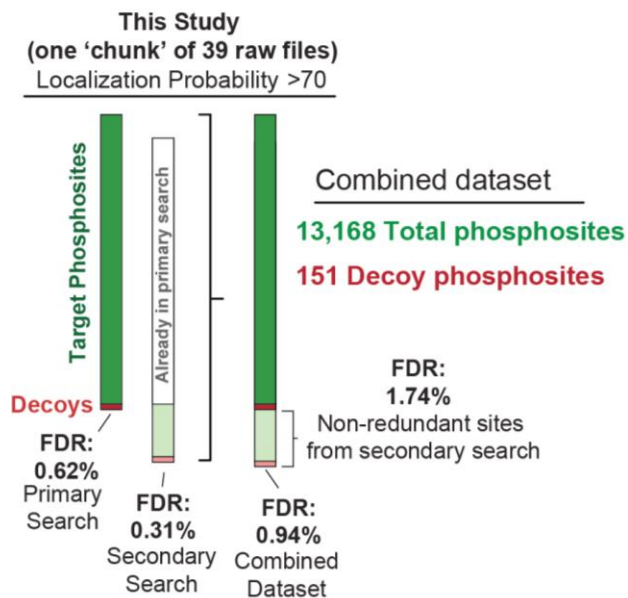
(C) To further increase the confidence in our Sequest searches, we only considered phosphopeptides whose backbone sequence appeared in both the PD and SORCERER searches.

(D) The phosphosite localization probabilities were determined using MaxQuant and the PhosphoRS node within Proteome Discoverer. Phosphosites with MQ localization / phosphoRS probabilities above 70 were considered to have "high-confidence" localization.

(E) A clustering algorithm was used to capture an additional class phosphosites that met our requirement for high-confidence localization. The logic used to cluster phosphosites is illustrated in Figure 1B.

(F) Because PhosphoRS runs independently of the PD PSM search, its phosphosite localization can sometimes conflict with the localization assigned in the rank 1 PSM. To further ensure data quality, we required agreement between the site localization determined by PD's PSM search and PhosphoRS node.

(G) PD and SORCERER searches are considered secondary because their PSM output is only included in the final data set (Dataset EV2) if a phosphosite was not already identified in the primary search. A version of the database sourced exclusively from the Sequest searches can be found in Dataset EV3.



**Legend for Appendix Figure S2: FDR analysis for primary, secondary, and combined searches**

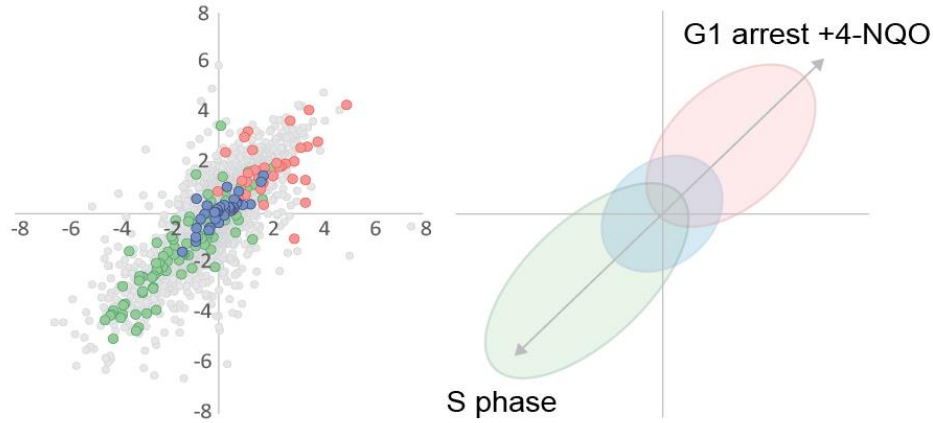
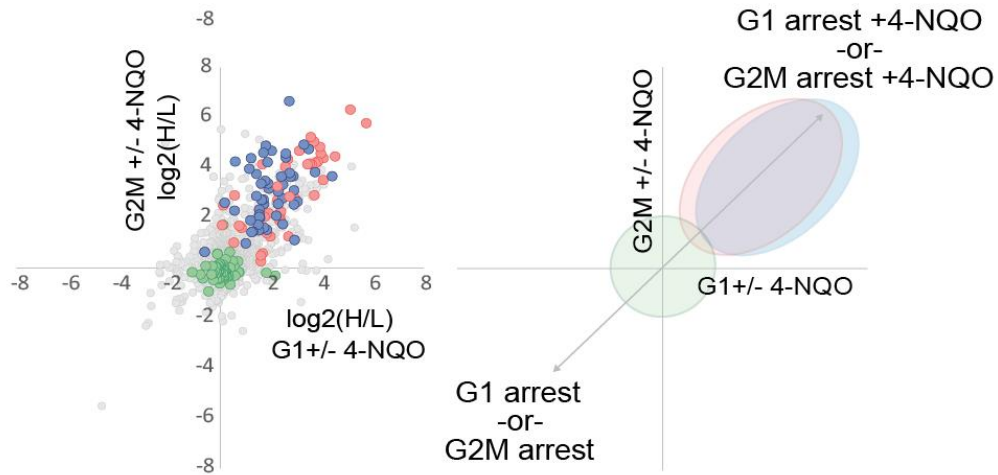
A schematic for the FDR evaluation of our dataset (not drawn to scale). By researching a chunk of our dataset (~40 raw files) with a manually embedded decoy database (Elias and Gygi, 2007), we were able to monitor FDR at each step of the search process and separately assign a FDR to the primary and secondary searches, both before and after combining them into the final dataset. This allowed us to monitor the FDR within the subgroup of phosphosites that were only identified after adding in the secondary search ("Non-redundant sites from the secondary search").

The overall FDR for the final dataset is less than 1%. Even when considering only the non-redundant phosphosites that are contributed from the secondary search, the FDR was less than 2%. We also note that the FDR for the secondary search is overestimated because decoys in Proteome Discoverer are not assigned a localization probability and thus cannot be filtered by localization confidence (the Target sites are filtered >70).



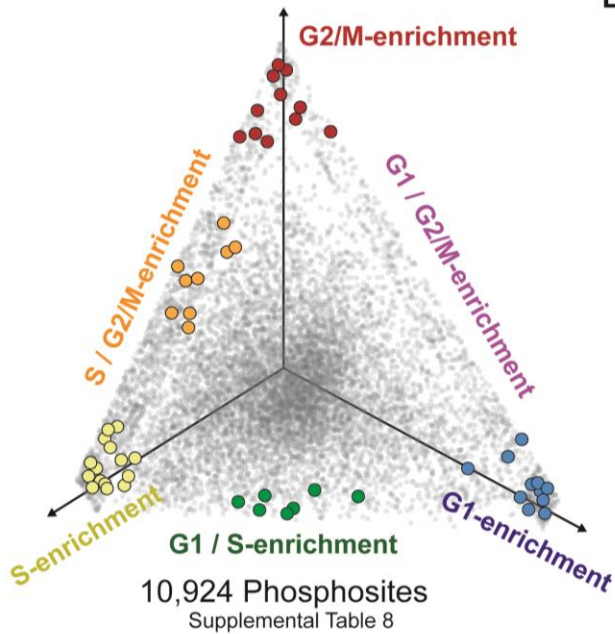
**Legend for Appendix Figure S3: Assessing phosphoproteome depth.**

Phosphosites identified in this study (exclusive to this study in blue). The bolded blue sites denoted by an asterisk represent putative phosphosites that were mutated and analyzed in previous studies. See main text for references.

**A****Light:** S phase synchrony**Heavy:** Alpha factor arrest + 4-NQO**B****Light:** G1 / G2M arrest**Heavy:** G1 / G2M arrest  
+ 4-NQO**Legend for Appendix Figure S4: Scatterplots for two SILAC replicates and DNA damage in G2M.**

Heavy / Light SILAC ratios (Log2) are plotted for phosphopeptides identified in two biological replicate experiments. (A) scatter plot representation of Figure 4B. (B) Alpha factor arrest was for 3hours. 0.4ug/ml 4-NQO was added 2hours and 20mins after G1 arrest. Cells were arrested in G2M with nocodazole for 2 hours and 40 mins. 0.4ug/ml 4-NQO was added 2hours after G2M arrest. Mec1/Tel1, Rad53, and CDK substrate phosphopeptides are highlighted in blue, red, and green, respectively. The right panel summarizes kinase activity based on the behavior of target phosphopeptides (see Dataset EV7).

A



B

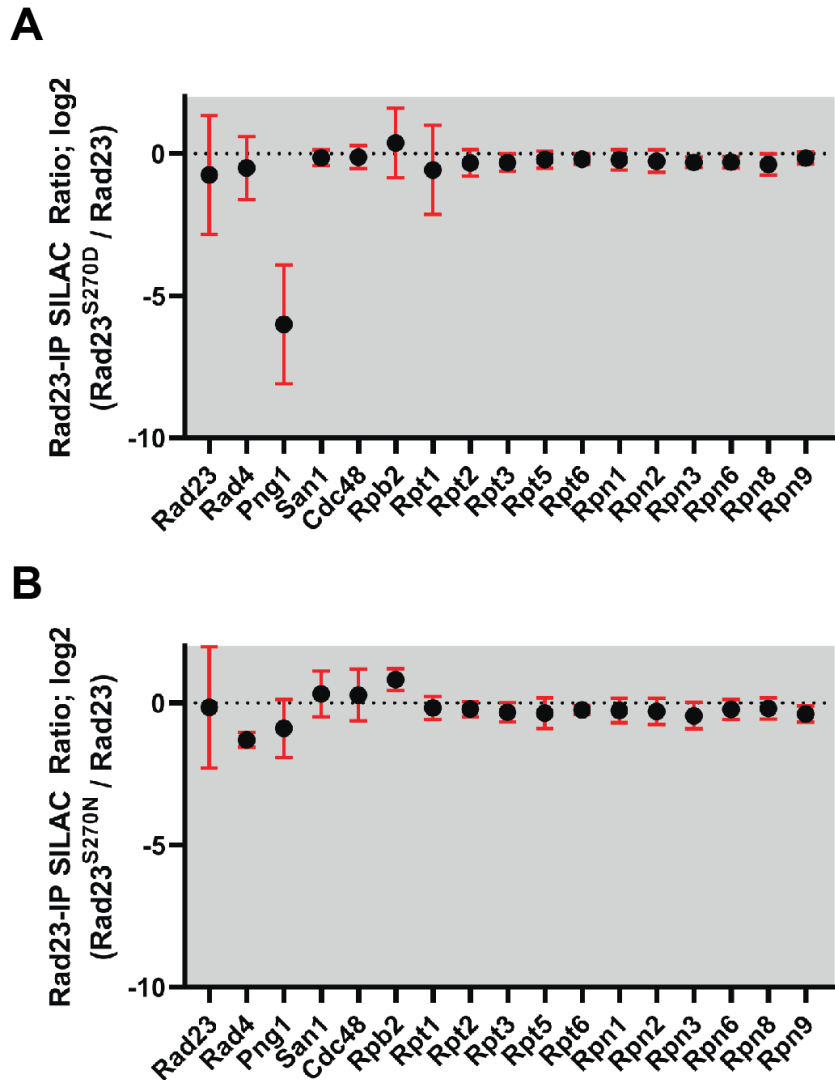
Phosphopeptide Identifications

	G1	S	G2/M	Gene_Site	Description	
●	22	3	2	STE12_29	G1-arrest	
●	19	1	0	STE2_310		
●	16	0	0	STE2_382		
●	16	0	2	STE4_335		
●	13	0	0	FAR1_306		
●	24	0	0	FAR1_43		
●	21	0	0	FAR1_8		
●	9	0	0	SIC1_5		CDK inhibitor
●	13	0	2	WHI4_205		Regulators of START
●	20	3	4	WHI4_236		
●	10	11	0	MBP1_133	G1/S Transcription Factors	
●	8	14	0	SWI6_547		
●	7	19	1	DAD1_89	Chromosome Segregation	
●	21	21	1	GIP4_478		
●	13	12	0	CHD1_1336	Chromatin remodeling	
●	6	9	0	RSC8_370		
●	20	9	1	CDC13_708	Telomere-associated	
●	0	13	0	RFA2_187	Replication Protein A Subunit	
●	0	12	1	RFA2_189		
●	0	10	0	SLD2_124	Replication Initiation	
●	0	7	0	SLD2_128		
●	0	10	0	SLD2_137		
●	0	5	0	SLD5_42		
●	1	21	1	CDC45_214		
●	0	8	0	DBF4_235	DDK	
●	0	19	2	DBF4_473		
●	0	9	0	DBF4_607		
●	1	11	0	RNR3_866	dNTP synthesis	
●	0	14	1	RNR3_868		
●	0	19	3	NRM1_145	G1/S Transcriptional Repressor	
●	0	12	0	NRM1_157		
●	2	11	14	FKH2_832	S/G2 Transcription Factor	
●	0	22	14	CDC5_238	Polo-like Kinase	
●	0	13	13	CDC5_70		
●	1	8	6	CDC14_429	Mitotic Phosphatase	
●	0	14	15	MAD1_502	Spindle Checkpoint	
●	4	38	25	CHS2_100		Chitin Synthesis
●	1	31	21	CHS2_133		
●	0	10	7	CHS2_60		
●	1	19	34	CHS2_69		
●	0	0	7	CDC5_419		Polo-like Kinase
●	0	1	6	CDC14_393	Mitotic Phosphatase	
●	0	0	6	CDC15_938		Mitotic Kinase
●	0	0	5	MAD1_233	Spindle Checkpoint	
●	0	1	5	BIR1_765		
●	1	0	8	MOB1_34	Mitotic Exit	
●	0	2	10	BNI1_1783	Formin	
●	0	0	5	BNI1_52		
●	0	0	6	BNI4_133	Glc7 Phosphatase regulator	
●	1	2	14	GLC8_184		
●	1	2	13	GLC8_224		
●	2	0	13	GLC8_224		

Legend for Appendix Figure S5: Cell cycle analysis.

(A) Ternary plot of the distribution phosphorylation events as a function of their detection in different stages of the cell cycle. A curated set of experiments from the larger dataset presented in Figure 1 was used for this analysis. Each gray dot represents a unique phosphosite. The position of each dot within the plot represents the fraction of times it was detected in either G1, S phase, or G2/M. A jitter was added to the plot to help visualize overlapping data points.

(B) Table that corresponds to the highlighted dots from (A) and the number of times they were detected in each cell cycle stage. See Dataset EV8 for the full dataset.



**Legend for Appendix Figure S6: IP-MS of different Rad23 phosphomutants.**

Quantitative mass spectrometry analysis of the Rad23 interaction network and the effect of a (A) Rad23 S270D and (B) Rad23 S270N. SILAC-labeled yeast cultures expressing Rad23-FLAG or an empty vector were subjected to anti-FLAG IP to pre-define the list of specific Rad23 interacting proteins shown in the graph. The average SILAC ratios represent changes for each of the Rad23 interactions in an IP using wild-type Rad23 versus Rad23 mutant as bait (Dataset EV12). Error bars represent the standard deviation of measurement made on multiple independent peptides. Each panel represents data from one biological replicate.

**Appendix References:**

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4, 207-214.