scientific reports

OPEN

Maximized quantitative phosphoproteomics allows high confidence dissection of the DNA damage signaling network

Vitor Marcel Faca[1,2,4], Ethan J. Sanford[1,4], Jennifer Tieu[1], William Comstock[1], Shagun Gupta[3], Shannon Marshall[1], Haiyuan Yu[3] & Marcus B. Smolka[1✉]

The maintenance of genomic stability relies on DNA damage sensor kinases that detect DNA lesions and phosphorylate an extensive network of substrates. The Mec1/ATR kinase is one of the primary sensor kinases responsible for orchestrating DNA damage responses. Despite the importance of Mec1/ATR, the current network of its identified substrates remains incomplete due, in part, to limitations in mass spectrometry-based quantitative phosphoproteomics. Phosphoproteomics suffers from lack of redundancy and statistical power for generating high confidence datasets, since information about phosphopeptide identity, site-localization, and quantitation must often be gleaned from a single peptide-spectrum match (PSM). Here we carefully analyzed the isotope label swapping strategy for phosphoproteomics, using data consistency among reciprocal labeling experiments as a central filtering rule for maximizing phosphopeptide identification and quantitation. We demonstrate that the approach allows drastic reduction of false positive quantitations and identifications even from phosphopeptides with a low number of spectral matches. Application of this approach identifies new Mec1/ATR-dependent signaling events, expanding our understanding of the DNA damage signaling network. Overall, the proposed quantitative phosphoproteomic approach should be generally applicable for investigating kinase signaling networks with high confidence and depth.

Protein phosphorylation is of central importance in both normal physiology and pathological conditions. Phosphorylation-mediated switches regulated by protein kinases and protein phosphatases can affect protein structure and function, with consequences in enzymatic activity, protein localization, protein interactions and turnover[1–3]. The control circuits of the DNA damage response (DDR) are extensively regulated by phosphorylation, with the kinase Mec1 (human ATR) playing a major role in both activation of the DNA damage checkpoint as well as phosphorylation of substrates involved in a range of nuclear processes including DNA repair, DNA replication, and transcription[4–10], To date a number of Mec1 substrates have been mapped by phosphoproteomics[11–13]. However, the current network of identified Mec1 substrates remains incomplete. Many DNA repair proteins are not highly expressed[14], and represent a challenge for phosphoproteomic analyses of DNA damage signaling to achieve proper depth with high quality quantitative data. Improvements in global quantitative phosphoproteomic analyses are therefore necessary to comprehensively map the Mec1-dependent signaling network. Similar challenges exist for the study of other kinases and represent important barriers for progress in understanding kinase action in general.

Phosphoproteomics, the systematic and unbiased mapping of phosphorylation events, is achieved mainly using mass spectrometry (MS)-based approaches. Both instrumentation and bioinformatic tools applied for phosphopeptide identification have been continuously evolving[15–17], culminating in large phosphoproteomic datasets in recent years[18–24]. In addition to in depth coverage of the phosphoproteome, comprehensive mapping of kinase-mediated signaling also requires quantitative analysis of each phosphopeptide or phosphorylation site to monitor its abundance in conditions of active kinase compared to conditions in which the kinase of interest is chemically and/or genetically ablated[11,25,26]. Various quantitative mass spectrometric approaches have been

[1]Department of Molecular Biology and Genetics, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, USA. [2]Department of Biochemistry and Immunology and Cell-Based Therapy Center, Ribeirao Preto Medical School, University of Sao Paulo, Ribeirao Preto, SP 14049-900, Brazil. [3]Department of Computational Biology, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, USA. [4]These authors contributed equally: Vitor Marcel Faca and Ethan J. Sanford. ✉email: mbs266@cornell.edu

applied for the mapping of kinase signaling, including stable isotope labeling in cell culture (SILAC)[26–28] and isobaric labelling strategies such as tandem mass tag (TMT)[29,30]. In a recent systematic comparison of quantitative phosphoproteomic strategies, SILAC was considered the most accurate, although TMT-based analyses yielded better coverage of the phosphoproteome[17]. SILAC is based on peptide precursor ion quantification to detect and quantify, in relative terms, the ratio between "heavy" and "light" isotopologues of amino acids (most commonly lysine and arginine) incorporated metabolically into cells[31,32]. Such an approach allows early mixing of labeled protein extracts in phosphoproteomic workflows to minimize technical variation.

Phosphoproteomics faces inherent issues for achieving identification and quantitation of phosphopeptides with high confidence. Different than proteomics, where the analysis of proteins is based on identification and quantification of multiple redundant representative peptides for a given protein, phosphoproteomics relies on phosphopeptides that are often unique (non-redundant) species represented by one or a few peptide spectral matches (PSMs) in the dataset. The lack of multiple redundant events for informing identification, quantification and phospho-site localization hobbles the acquisition of high-quality data due to the low numbers of PSMs per phosphopeptide[33]. The ability of acquiring high quality identification and quantification data is further compli-cated by the fact that many key phosphopeptides of biological interest are present at very low levels in the pool of phosphopeptides enriched from whole cell lysates. Even in cases when identification of a phosphopeptide based on one or two PSMs is successful, the associated quantitative information can suffer from signal interference derived from sample complexity and other intrinsic technical noise[34–37]. As a result, a significant part of the gener-ated phosphoproteomic data is not suited for reliable quantitative analysis and biological inference, representing one of the major bottlenecks in large-scale quantitative phosphoproteomic analysis of kinase-mediated signaling.

Here we report a phosphoproteomic approach for increasing reliability in phosphopeptide identifi-cation and quantification, while minimizing loss of data from phosphopeptides with low PSM counts. The approach builds on the established concept of SILAC labeling swap, relying on quantitation consistency among reversed isotopically labeled samples as a central filtering step for removing false positive identifications and erroneous quantifications[31,38,39]. While isotopic label swapping has been a common practice in SILAC-based experiments[38–42], its contribution to the reduction of false positive identifications and quantitations has not been systematically characterized, especially for cases of phosphopeptides with low PSM counts. By performing an in-depth analysis of label swap phosphoproteomics we monitor experimental error or biological variation in phosphopeptide quantitation and propose an approach for drastic reduction of false positive identifications and quantitation. The reported approach balances both sensitivity and specificity to detect phosphorylation changes with high confidence, even in the case of phosphopeptides with low PSM counts. Overall, the simple approach presented here enhances the reliability of quantitative phosphoproteomics in biological interrogations of kinase-mediated signaling networks.

Results

Error and variation in SILAC-based phosphopeptide quantitation is unidirectionally biased.

We set out to develop an approach to maximize confidence in quantitative data from phosphoprot-eomic experiments. We postulated that SILAC-based quantitation might be particularly well suited for separat-ing meaningful biological changes from: (1) aberrant quantitation during data processing (herein referred as "Error"), and/or (2) changes in phosphopeptide abundance unintentionally introduced during sample handling (herein referred as "Variation"). If both Error and/or Variation (EV) are mostly associated with artifacts that are independent of true biological differences in the cell lines or drug treatment conditions being compared, phos-phoproteomic analysis should reveal a unidirectional bias in the generated ratios of data points reflecting EVs (Fig. 1A–C). We further reasoned that a strong bias in EVs would enable their systematic exclusion from large-scale phosphoproteomic datasets and, in principle, enable the generation of high confidence quantitative data even from phosphopeptides with only one PSM detected in each reciprocal, labeling swap SILAC experiment.

To test this idea, we mixed equal amounts of protein extracts from budding yeast grown in light ($^{12}C^{14}N$ arginine and lysine) or heavy ($^{13}C^{15}N$ arginine and lysine) SILAC media and subjected lysates to a quantitative phosphoproteomic and data analysis pipeline outlined in Fig. 1 and detailed in Supplemental Figure S1. An independent biological replicate was performed to mimic a reciprocal, labeling swap experiment. As shown in Fig. 1A, data points with a SILAC ratio not reflecting the expected 1:1 ratio (simulated within a 33% coefficient of variation, or approximately a twofold change), were considered to reflect methodological error and/or variation. Comparison of experiments A and B (control and reciprocal label swap) should reveal if the error and/or vari-ation exhibit any biased distribution in a quantitative plot (Fig. 1B,C). As shown in Fig. 2A (see Supplementary Table S1 for detailed dataset), separate experiments revealed thousands of data points outside a simulated range of 33% coefficient of variation (indicated in yellow). We reasoned that these points reflect EVs in the experiment. Notably, comparison of the ratio of each phosphopeptide in experiments A and B revealed a clear bias in EV dis-tribution toward quadrants Q2 and Q4 (Fig. 2B,C) such that 92% of all EVs fell within these quadrants. Notably, EVs accounted for 17% of all phosphopeptides present in our dataset when considering phosphopeptides with 1 PSM in each experiment, underscoring the importance of their exclusion. Data points in Q2 and Q4 represent phosphopeptides whose SILAC ratios did not revert in the reciprocal experiment. Overall, these results reveal that the use of a SILAC labeling swap in phosphoproteomic experiments allows efficient detection of intrinsic EV in the dataset, which may be used for achieving high confidence quantitative analysis, even for phosphopeptides represented by a low number of PSMs. This ability to filter signal from noise, even when PSM numbers are low, is crucial for phosphoproteomic experiments which often rely on difficult-to-detect phosphopeptides. In fact, approximately a third of the data points in the correlation plot shown in Fig. 2B reflect phosphopeptides with only one PSM in one of the experiments (see Supplementary Table S1).
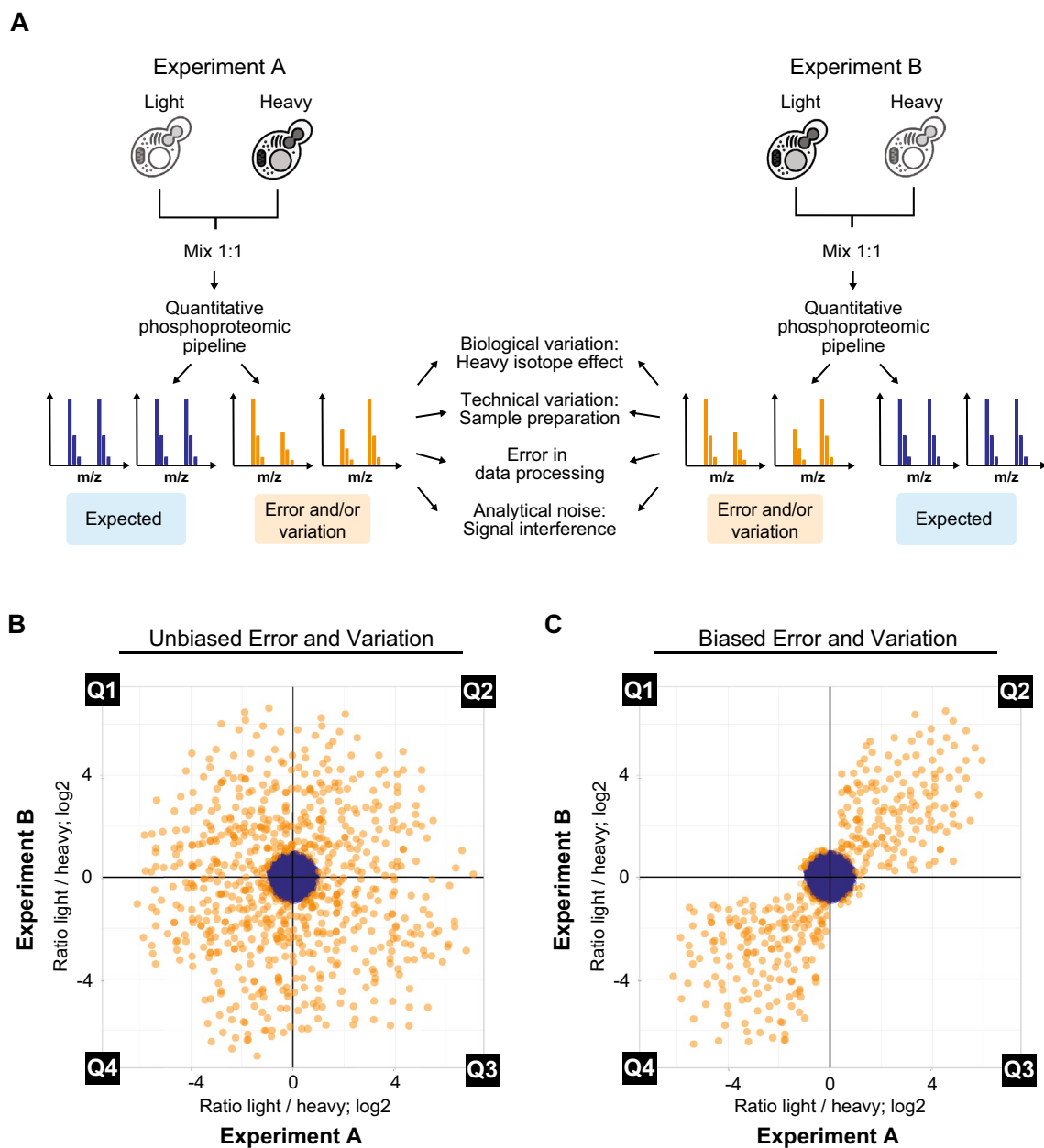
Figure 1. Modeling outcomes of SILAC reciprocal labeling as a means to reduce technical error and/or variation (EV). (**A**) Workflow showing reciprocal labeling scheme with a forward experiment (Experiment A, left) and a reverse experiment (Experiment B, right) with anticipated outcomes and proposed causes of EVs (middle). (**B**) Anticipated distribution of false positives in a comparison of two identical samples if error and variation occurred randomly and independent of isotopic labeling. (**C**) Anticipated distribution of false positives in a comparison of two identical samples if error and variation were unidirectionally biased (i.e. similar ratio in both a forward and reverse experiment).

Data filtering approaches for reducing error and variation. To apply data filtering approaches for efficiently eliminating EVs while minimizing loss of data, we evaluated the effects of imposing thresholds on the minimal number of observations (PSMs) required for each phosphorylation site identified. While each phosphosite requires at least 2 observations (1 in each of the reciprocal experiments) to be shown in the correlation plot, increasing the requirement for 2 or more observations in each experiment decreased the proportion of EVs in relation to the entire dataset (Fig. 3A). Considering specifically data points present in Q1 and Q3, where inverse correlation is expected between phosphopeptide ratios in reciprocal experiments, we find that the proportion of EVs is about 1.5% when considering 1 or more PSM in each experiment. This EV proportion is reduced by approximately half, to 0.8% of the data points, when a minimum of 2 observations is required in each experiment (Fig. 3B). However, this additional requirement also decreased sensitivity, reducing the total number of data points from 15,062 to 10,439 (Supplementary Table S1).
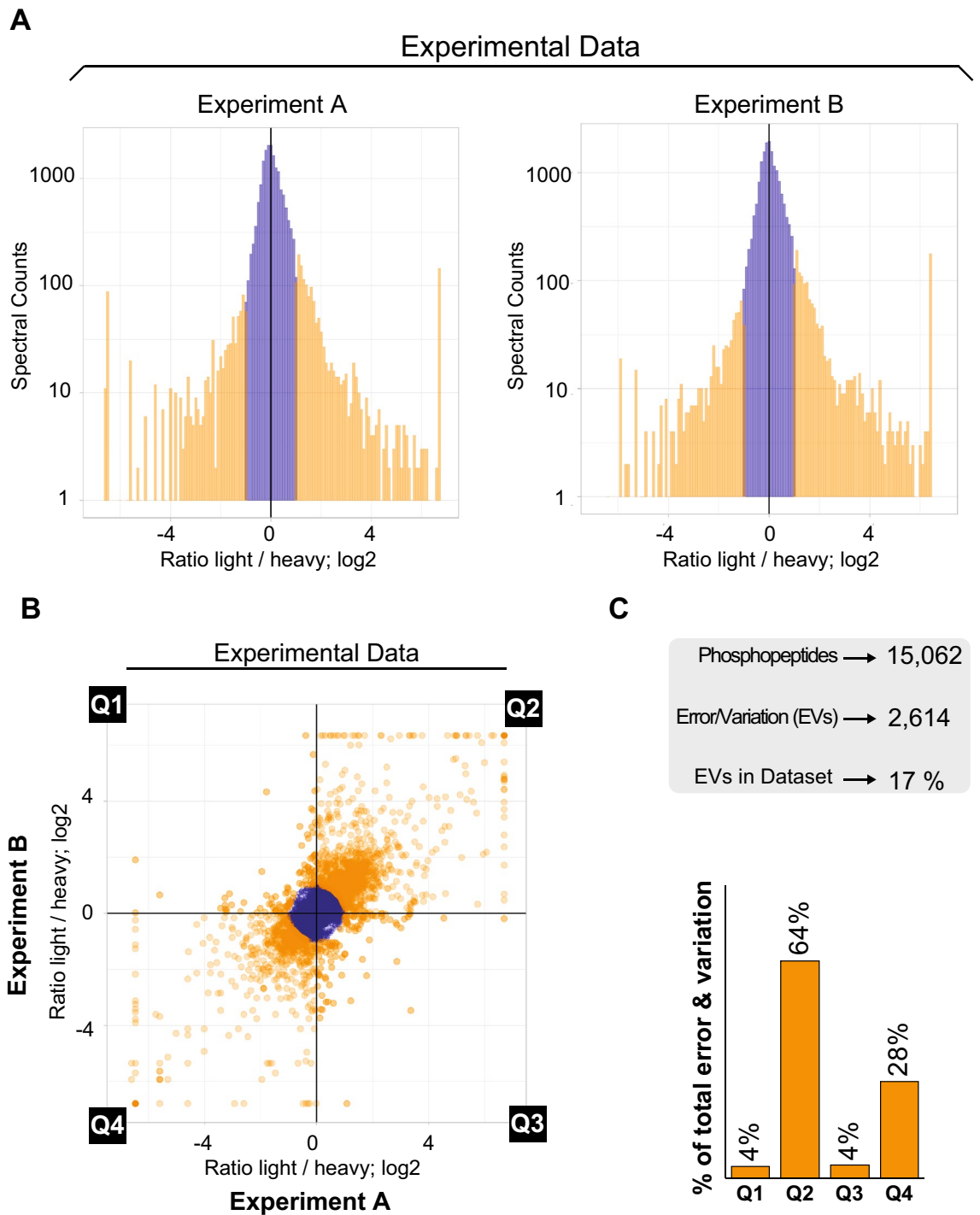
Figure 2. Reciprocal labeling in an isogenic yeast cell line reveals extensive error and variation that is unidirectionally biased. (**A**) Histograms for SILAC ratios of two independent phosphoproteome experiments comparing isogenic wild-type *S. cerevisiae* as depicted in Fig. 1A. EVs are colored in orange. (**B**) Scatterplot comparing experimental data from the two SILAC experiments shown in (**A**). EVs are colored in orange. (**C**) Histogram showing unidirectional bias of error and variation toward quadrants 2 and 4 in plot from (**B**).

During our EV analyses, we noticed a clear prevalence of data points close to the X-axis and Y-axis in Q1 and Q3 (Fig. 3A,C), revealing data points with a deviated ratio in only one of the experiments. By employing a simple "quadrant filtering" approach, whereby points in Q2 and Q4 are excluded, and points in Q1 and Q3 are kept, we cannot exclude highly variable phosphopeptide measurements that are also likely the result of error and/ or variation (Fig. 3C). To circumvent this issue and more efficiently remove EVs for improved data quality, we designed an alternative filtering approach where data points in Q1 and Q3 were required to be within an interval of correlation correspondent to fourfold of the log2 scale (hereafter referred to as "Bow-tie filtering") (Fig. 3C).
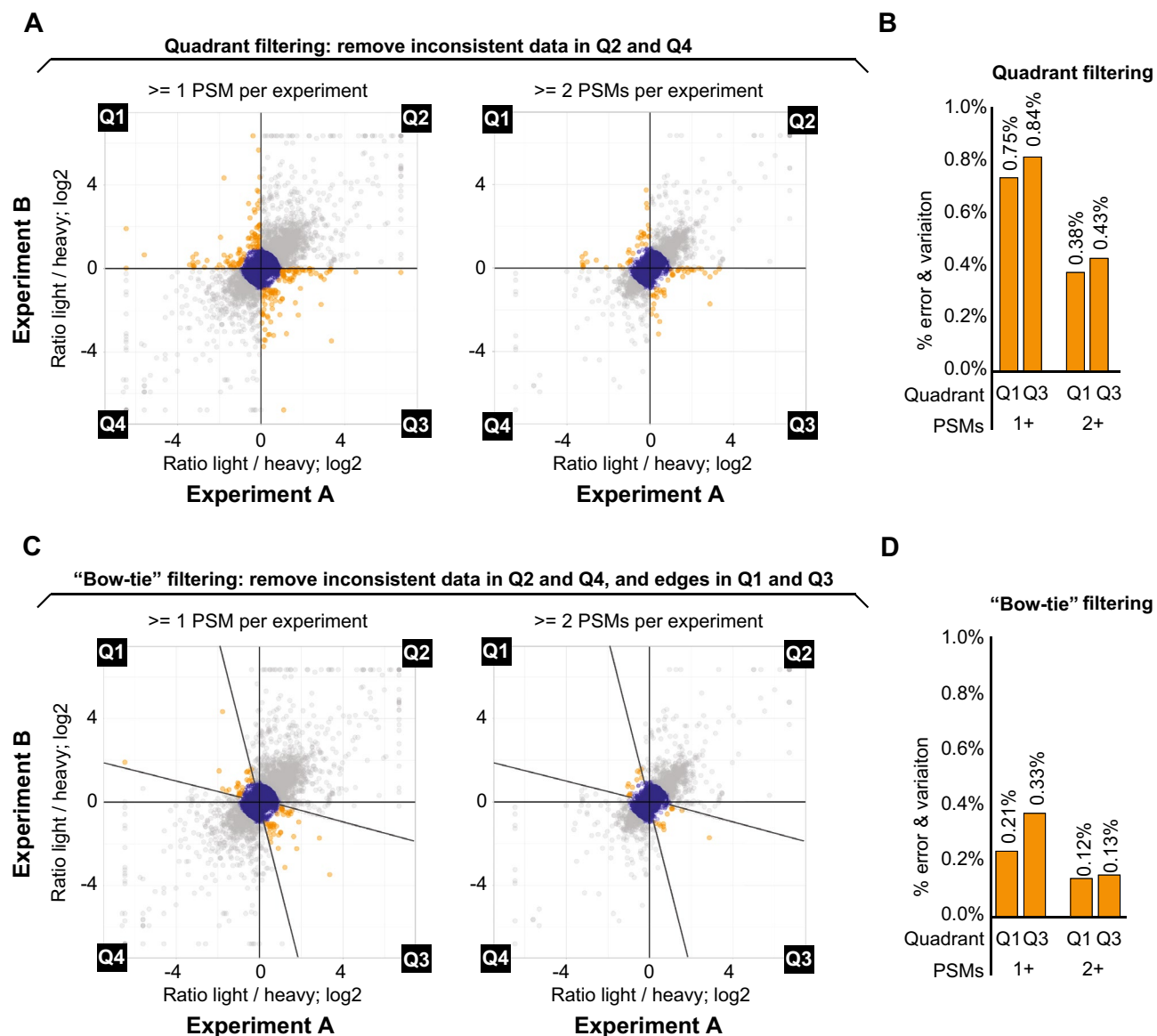
Figure 3. Data filtering based on quantitation consistency drastically reduces error and variation. (**A**) Scatterplot from Fig. 2B. indicating the "Quadrant" filtering scheme (gray data points removed in Q2 and Q4) and additional filtering based on the requirement for at least 2 PSMs per experiment for each data point (plot on the right). (**B**) Histogram showing EVs in Q1 and Q3 (orange data points) from A as a percentage of total dataset using either 1 PSM or 2 PSM filtering. (**C**) Scatterplot from Fig. 2B. indicating the "Bow-tie" filtering scheme (gray data points removed) and additional filtering based on the requirement for at least 2 PSMs per experiment for each data point (plot on the right). For Bow-tie filtering, in addition to removing EVs in Q2 and Q4, data points in Q1 and Q3 were required to be within an interval of correlation correspondent to fourfold of the log2 scale. (**D**) Histogram showing highlighted points in quadrants 1 and 3 from (**C**) as a percentage of total dataset using either 1 PSM or 2 PSM filtering.

As shown in Figs. 3C,D, the use of Bow-tie filtering, even where peptides with 1 PSM in each experiment were included, reduced the proportion of EVs to 0.54% of the dataset. When Bow-tie filtering was combined with the threshold of at least 2 PSMs per experiment, the proportion of EVs again dropped by approximately half to 0.25% of the dataset. These results reveal that the ability to identify EVs in SILAC-based phosphoproteomic experiments allows the utilization of filtering strategies that drastically reduce error and variation in the dataset, therefore increasing the confidence in the data even when considering phosphopeptides represented by a single PSM per experiment.

Eliminating error and variation in quantitation reduces decoy identifications. SILAC labeling with stable isotopes shifts the mass of parent ions and their fragments in both MS1 and MS2, respectively. We reasoned that this mass shift should enable more efficient exclusion of false positive identifications in the dataset, since a misidentification would need to occur in both reciprocal experiments and be consistent between two
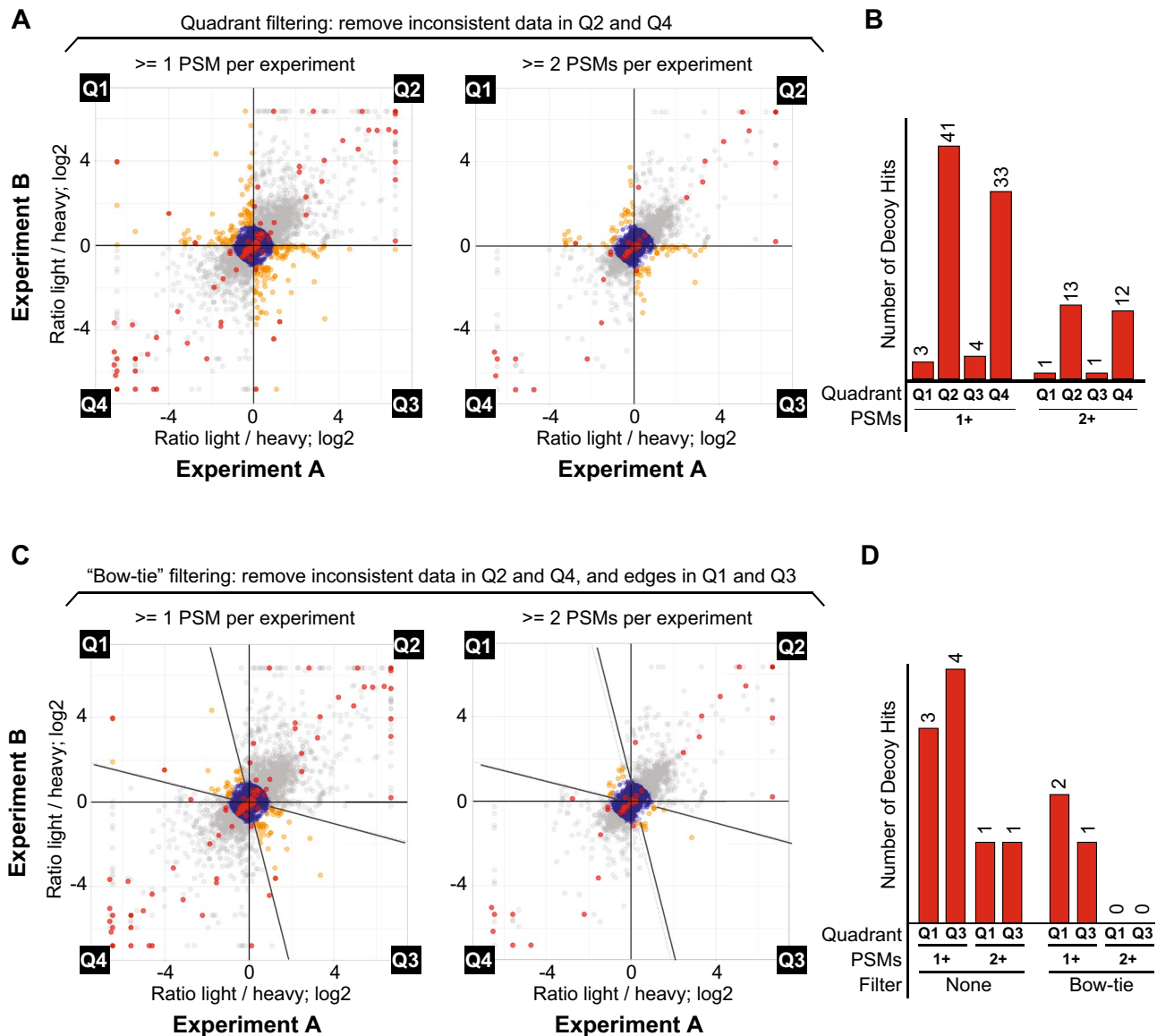
5

Figure 4. "Bow-Tie" approach efficiently reduces decoy peptide identifications. (**A**) Scatterplot from Fig. 2B. indicating hits from a decoy database. Decoy peptides are displayed in red. As in Fig. 3, unfiltered EVs in Q1 and Q3 are displayed in orange. (**B**) Histogram displaying number of decoy peptide hits in Q1 and Q3 with quadrant filtering applied. (**C**) Scatterplot from Fig. 2B. indicating hits from a decoy database and employment of Bow-tie filtering. Decoy peptides are displayed in red. As in Fig. 3, unfiltered EVs in Q1 and Q3 after Bow-tie filtering are displayed in orange. (**D**) Histogram displaying number of decoy peptide hits in Q1 and Q3 with Bow-tie filtering applied.

parental ions with different m/z. To give a more detailed example, a false identification in a $^{12}C^{14}N$ (light) sample with a high light/heavy ratio, should not be reciprocally identified in the $^{13}C^{15}N$ (heavy) form, or if identified in the light form in the reciprocal experiment, it should not display an inverted low light/heavy ratio. If most of these cases reflect intrinsic experimental artefacts consistently present over biological replicates, independently of SILAC labeling swap, these false identifications should be prevalent in quadrants Q2 and Q4, because the peptide in question would be very unlikely misidentified in a reciprocal experiment due to its having a different m/z and/or display an inverted ratio. In such a context, consistency in quantitation over two or more biological replicates of label swapped experiments could be used as a parameter for efficiently excluding false identifications from final datasets, especially in the region of data points with high fold changes containing most of the key data that would be used for biological inference, such as for the identification of kinase substrates.

To test if performing a reciprocal labeling experiment indeed reduces false-positive identification and quantification, we estimated the error rate of phosphopeptide identifications by monitoring the distribution of reversed decoy hits from the list of phosphopeptide identifications that passed our basal quality criteria (Peptide-Prophet > 0.9 and < 20 ppm precursor ion error). As shown in Fig. 4A,B, decoy hits display a clear distribution bias towards quadrants Q2 and Q4, congruent with our rationale that false identifications are mostly unidirectional

Minimal # of PSMs per experiment	Total phospho sites matched*	Identification		Quantitation		Data filter	Phospho-sites with consistent quantitation (center)	Identification		Quantitation	
		Total # of decoys hits	% of decoy hits	Total # of error and variation	% of error and variations			Persisting decoy hits in Q1+Q3	% of decoy hits post-filtering	Persisting error and variation in Q1+Q3	% of error and variation post-filtering
1	15,062	130	0.86	2614	17.35	Quadrant	12,448	7	0.06	198	1.59
1	15,062	130	0.86	2614	17.35	Bow-tie	12,448	3	0.02	67	0.54
2	10,439	46	0.44	1263	12.10	Quadrant	9176	2	0.02	75	0.82
2	10,439	46	0.44	1263	12.10	Bow-tie	9176	0	0.00	23	0.25
Data filter: PTMPPROPHET score ≥ 0.9											
1	13,391	118	0.88	2232	16.67	Quadrant	11,159	8	0.07	143	1.28
1	13,391	118	0.88	2232	16.67	Bow-tie	11,159	4	0.04	40	0.36
2	9771	41	0.42	1149	11.76	Quadrant	8622	2	0.02	61	0.71
2	9771	41	0.42	1149	11.76	Bow-tie	8622	0	0.00	16	0.19

Table 1. Summary of phosphoproteomic data from Figs. 2, 3 and 4, comparing Quadrant to Bow-tie filtering, 1 PSM cutoff to 2 PSM cutoff, and filter for PTMPROPHET score ≥ 0.9.

in quantitation and likely reflect artefacts that are extremely unlikely to occur in two reciprocal experiments, independently. Of all decoy hits in the unfiltered dataset, more than half (81 out of 130) were found to display ratios outside the twofold change range (Fig. 4A,B; Table 1). Notably, we were able to remove all decoy hits from Q1 and Q3 (regions expected to contain key data for biological inference of true changes in phosphorylation events) using the Bow-tie filtering strategy in combination with a threshold of at least 2 PSMs per experiment (Fig. 4C,D; Table 1). Even when phosphopeptides reflected by 1 PSM per experiment were allowed in the dataset, the number of decoy hits in Q1 or Q3 remained low (2 hits) (Fig. 4D). We also tested a stringent filter for phosphorylation site localization (PTMProphet score equal or above 0.9), which further reduced EVs in Q1 and Q3 to 0.36% without drastically reducing the overall coverage of the dataset (Table 1). These findings highlight the usefulness of our approach, which hinges on conducting a reciprocal SILAC experiment to improve confidence in both identification and quantitation in phosphoproteomic studies. Importantly, the described approach results in minor loss of valuable data content from low abundance phosphopeptides represented by only one PSM in each of the two reciprocal experiments.

High confidence dissection of the Mec1-dependent signaling network.

Mec1, the *Saccharomyces cerevisiae* ortholog of mammalian ATR, is a phosphoinositide 3-kinase-related kinase (PIKK) kinase that is a key mediator of DNA damage responses[43–45]. We have previously used quantitative phosphoproteomics comparing WT and *mec1Δ* cells to uncover phosphorylation events dependent on Mec1[11]. Here we applied our optimized quantitative phosphoproteomic approach to the study of Mec1 in order to benchmark our bow-tie approach and expand the Mec1-dependent signaling network. We carried out the experiments in cells treated with the DNA alkylating agent MMS (methyl methanesulfonate) and lacking the checkpoint adaptor Rad9 to minimize indirect downstream phosphorylation and preferentially reveal direct Mec1 substrates[6,46]. Overall phosphoproteome coverage was similar to the control experiment, with approximately 20,000 phosphopeptide identifications for each SILAC reciprocal experiment. Upon application of our most relaxed filtering scheme, which considers phosphopeptides with 1 or more PSM in each experiment and a PeptideProphet score of 0.9 or greater, a total of 13,456 unique phosphosites from 2778 different proteins were identified. In order to ensure confidence in phosphopeptide localization, we applied a PTMProphet score filter of greater than or equal to 0.9, somewhat reducing the total number of unique phosphopeptides identified in the Mec1 experiment to 11,950. The list of all phosphosites identified and quantified in our experiment is supplied in Supplemental Table S2. As shown in Fig. 5A, Q1 after Bow-tie filtering contained a large number of phosphosites consistently downregulated in *rad9Δ* cells lacking Mec1 in both reciprocal SILAC experiments. The number of phosphosites in Q1 was approximately equal to the number of EVs in Q2 and Q4 (Supplemental Table S3), indicating that if experiments we performed using only one labeling scheme, many of these EVs excluded in our Bow-tie approach would have been erroneously called Mec1-dependent sites, obfuscating true biological effects of *MEC1* loss. Reassuringly, phosphopeptides in Q1 or Q3 (representing phosphorylation events lost or induced upon deletion of *MEC1*) were approximately eight times more prevalent than in the WT (1:1) control experiment (Fig. 5B). Our filtering strategy allows minimal loss of data while increasing stringency for identification and exclusion of false positives through SILAC label swapping. To systematically and quantitatively demonstrate that the set of Mec1-dependent phosphorylation events had a low rate of EVs, we sought to stratify the bow-tie filter into several bins of increasing fold-change and calculate the false discovery rate (FDR) for each bin. Mathematically, the FDR is equal to the number of points in a given bin in the control-experiment divided by the number of points in a given bin in the Mec1 experiment. For the purposes of this calculation, points in Q1 and Q3 were considered together (Supplemental Figures S2A,C). Expectedly, FDR decreased with increasing distance from the center and was further reduced depending on how close points fell to the line of symmetry (Supplemental Figures S2B,D). The majority of the data points in Q1 and Q3 encompassed by the bow-tie filter have a p value less than 0.05 (Fig. 5C), thus validating our bow-tie filtering approach as a means to improve data quality while allowing the inclusion of difficult-to-detect phosphorylation events.
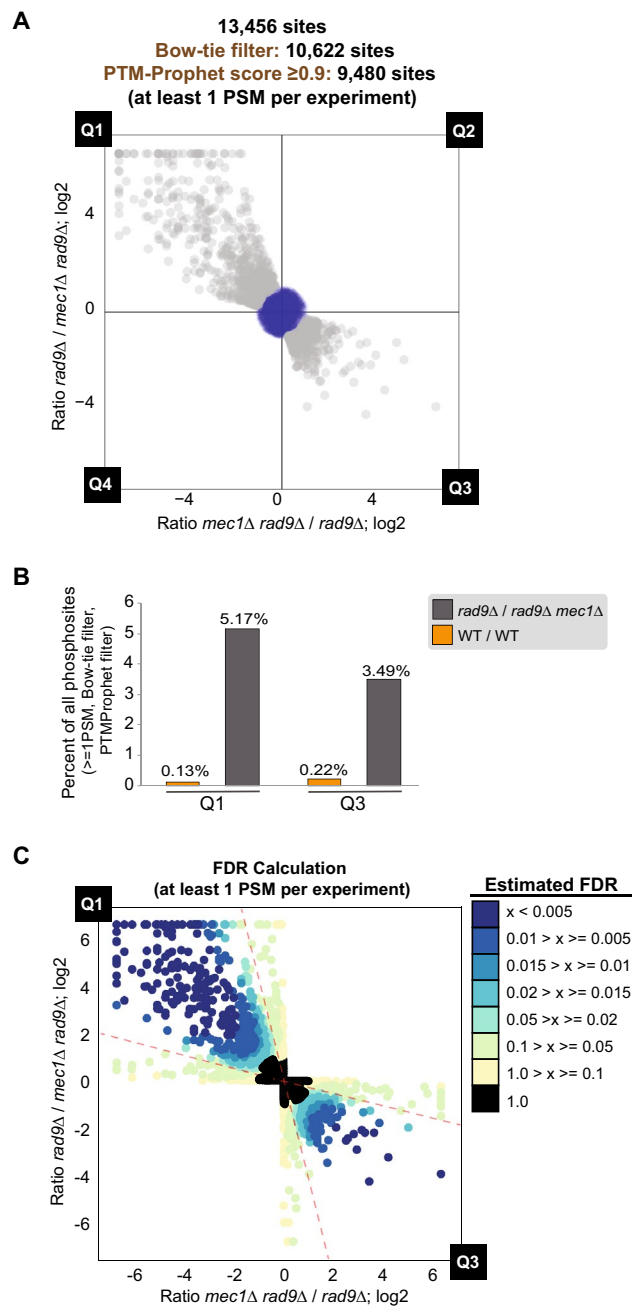
Figure 5. Quantitative phosphoproteomic analysis of Mec1-dependent signaling. (**A**) Scatterplot (with Bow-tie filter applied and PTMProphet score ≥ 0.9; only data points within Bow-tie filter displayed) of forward and reciprocal SILAC experiment comparing phosphoproteome of *rad9Δ* cells to phosphoproteome of *rad9Δ mec1Δ* cells. Cells were treated with 0.02% MMS for 2hrs. (**B**) Histogram depicting distribution of phosphorylation sites in Q1 and Q3 compared to control experiments. (**C**) Estimation of false discovery rate (FDR) in quantitative analysis for experiment in 5A. FDR for quadrants 1 and 3 is estimated based on error and variation in wild-type control experiment (Fig. 2). See "Materials and methods" for more details.

The results of our experiment revealed an extensive network of Mec1-dependent phosphorylation events, many not published before and mostly phosphorylated at the preferential S/T-Q motif (Fig. 6A, green dots), which was overrepresented in Q1 (Fig. 6B). Whereas the S/T-Q motif represents only about 3% of the phospho-sites in the entire dataset, it represents 33% of the Q1 sites, and 49% of the group of highly Mec1-dependent sites (over twofold depletion in *rad9Δmec1Δ* cells). Besides S/T-Q sites, Q1 also contained a number of sites with the S/T-ψ (where ψ denotes the bulky hydrophobic residues F, I, L and V) phospho motif (Supplemental Table S3), which is associated with the downstream checkpoint kinase Rad53 that is activated by Mec1[47]. The occurrence of S/T-ψ phosphorylation in the absence of the major *RAD9*-dependent pathway of Rad53 activation likely reflects Rad53 activation via the Mrc1 adaptor[48]. Indeed, Mec1-dependent phosphorylation sites were detected
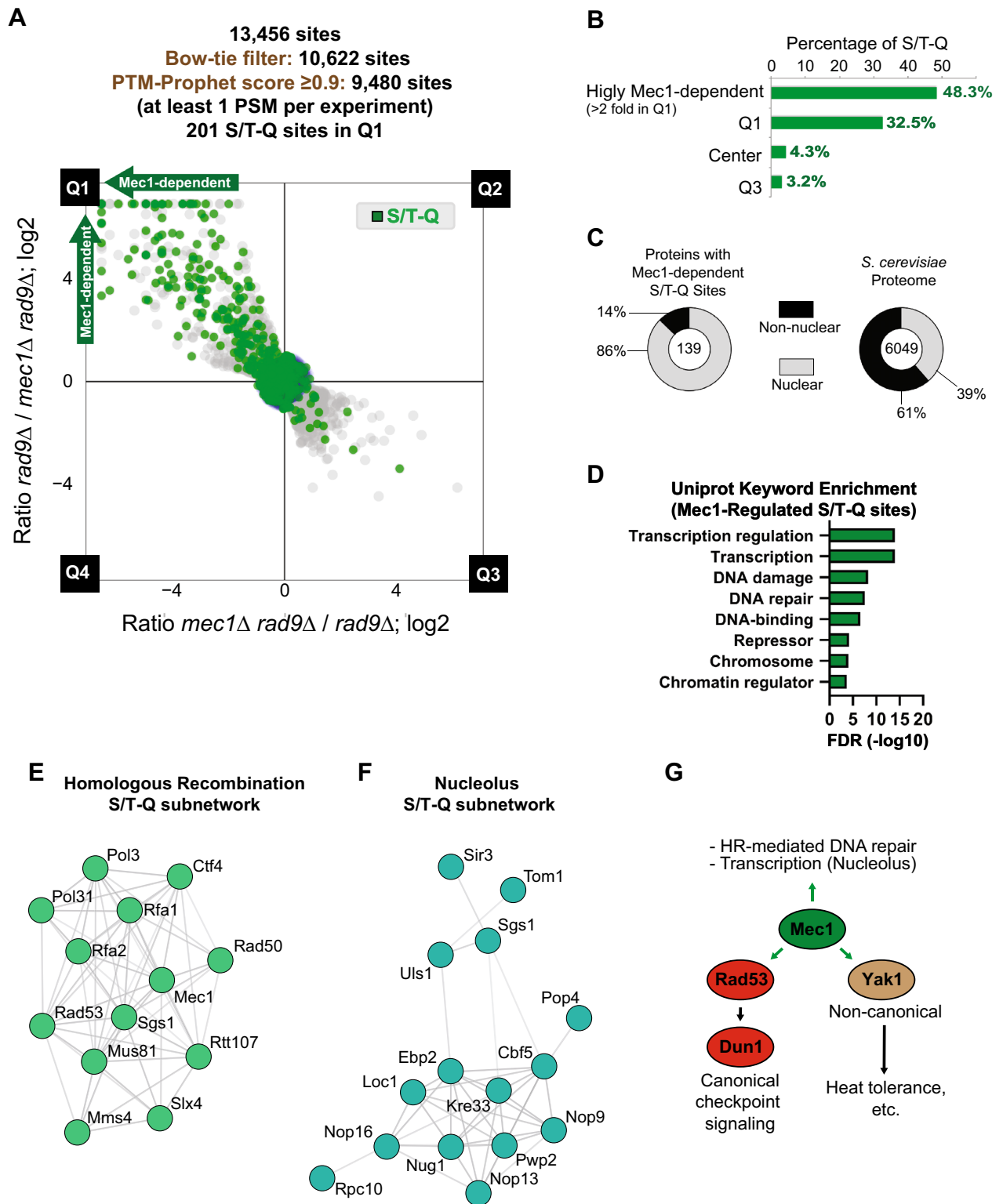
Figure 6. Expanding the Mec1 signaling network. (**A**) Scatterplot (with Bow-tie filter applied and PTMProphet score ≥ 0.9; only data points within Bow-tie filter displayed) of forward and reciprocal SILAC experiment comparing phosphoproteome of *rad9Δ* cells to phosphoproteome of *rad9Δ mec1Δ* cells. S/T-Q consensus motif is highlighted in green. Cells were treated with 0.02% MMS for 2 h. (**B**) Histogram of proportion of S/T-Q phospho-motif by quadrant in the Mec1 experiment from (**A**). (**C**) Pie chart showing proportion of nuclear proteins (GO Cellular Location[79,80]) in Mec1-dependent (log2 ratio > 1.0) S/T-Q sites from the Mec1 experiment in (**A**). (**D**) Uniprot keyword enrichment analysis performed on proteins containing Mec1-dependent S/T-Q phosphorylation from (**A**). (**E**) String analysis of proteins with Mec1-dependent phosphorylation in the S/T-Q consensus revealed a sub-network of proteins involved in DNA repair via homologous recombination (HR). Image adapted from https://string-db.org/. (**F**) String analysis of proteins with Mec1-dependent phosphorylation in the S/T-Q consensus revealed a sub-network of proteins related to the nucleolus. Image adapted from https://string-db.org/. (**G**) Simplified model depicting major processes directly regulated (green arrows) by Mec1. Results suggest that the Yak1 represents a novel kinase under direct control by Mec1.

in Rad53, several of which are known Rad53 autophosphorylation sites[49] and indicate that this kinase is activated in *rad9Δ* cells expressing Mec1.

In total, our quantitative phosphoproteomic approach using Bow-tie filtering of inconsistent ratios resulted in the identification of 201 S/T-Q Mec1-dependent phosphosites, which at least triples the number of Mec1 targets identified compared to our previous screen[11]. Consistent with Mec1 being a nuclear kinase, these sites identified in our screen occurred largely on nuclear proteins (Fig. 6C). Gene enrichment analysis of all Mec1-regulated S/T-Q sites in Q1 (with a log2 ratio > 1 in *rad9Δ* cells relative to *mec1Δ rad9Δ*) was consistent with our previous study showing that the substrate repertoire of this kinase was enriched for nuclear proteins involved in DNA repair, chromatin dynamics, and transcription (Fig. 6D; Supplemental Table S4). String network analysis[50] of the proteins with regulated S/T-Q sites revealed extensive Mec1-dependent phosphorylation of components of the homologous recombination machinery (Fig. 6E), including proteins such as Rad50 that act early in HR during the resection step[51,52], as well as proteins that act later during HR to regulate the processing of joint molecules, such as Sgs1 and Mus81-Mms4[53–55]. Additionally, we found extensive Mec1-dependent phosphorylation of nucleolar proteins at the S/T-Q consensus, suggesting direct control of nucleolar processes by Mec1 (Fig. 6F). Analysis of Mec1-regulated sites containing a consensus motif that was not S/T-Q revealed that the scope of Mec1's downstream signaling also largely encompassed proteins related to DNA damage, repair, and transcription, while also showing prevalence of cell-cycle, DNA replication and cytoplasmic proteins (Supplemental Figure S3A; Supplemental Table S5). Similar to the SQ consensus sites, the majority of the non-SQ sites were in nuclear proteins (Supplemental Figure S3B). String analysis of non-S/T-Q signaling events in the "cell cycle" node revealed non-canonical Mec1-dependent phosphorylation of the spindle assembly protein Mad3 and the condensin subunit Smc4 (Supplemental Figure S3C).

We also identified Mec1-dependent phosphorylation sites in the Dun1 kinase, which is known to function downstream of Mec1 and Rad53 in the canonical DNA damage checkpoint signaling pathway[56–58]. Interestingly, our SILAC-based filtering approach revealed a number of Mec1-dependent sites that did not contain the S/T-Q or S/T-ψ consensus, raising the possibility that Mec1 regulates the action of other kinases in addition to Rad53 and Dun1 in response to DNA damage. An example of a potentially new kinase targeted by Mec1 in our data is the DYRK-family kinase Yak1, which was phosphorylated in a Mec1-dependent manner in response to DNA damage on serine 663 (Supplemental Table S2; Fig. 6G). Both Yak1 and Mec1 have been reported to be important for acute heat shock resistance[59,60], raising the possibility that Mec1 and Yak1 may be acting in the same stress-response pathway. Lastly, analysis of phosphorylation sites in Q3 revealed a likely up-regulation of the Tel1 kinase, a Mec1-related PI3K-like Kinase (PIKK) with roles in DNA double strand break (DSB) repair and telomere maintenance[61–63]. Q3 included phosphorylation of the telomere maintenance protein Rif1 at serine 1308, which was previously shown to be dependent on Tel1[64]. In fact, ATM/Tel1 signaling has been reported to be up-regulated in the absence of ATR/Mec1 in mammals[65–67]. Q3 also contained additional phosphorylation sites in proteins related to DNA double strand break (DSB) repair and telomere maintenance (Supplemental Figure S4). Notably, some of these phosphorylation sites were not present in the canonical S/T-Q motif (Supplemental Table S2), suggesting additional non-canonical Tel1-dependent phosphorylation and/or the involvement of additional kinases. Taken together, these findings highlight the efficacy of our optimized quantitative SILAC-based phosphoproteomic approach and Bow-tie filtering method in identifying and quantifying kinase-dependent signaling events at high depth and specificity, while minimizing false positives.

Discussion

The field of phosphoproteomics has made significant strides toward improved phosphopeptide detection and quantitation since the seminal paper by Fenselau et al. which described the first application of FAB mass spectrometry for phosphopeptide characterization[68]. Throughput as well as robustness has increased, and modern instruments and workflows can routinely detect and quantitate thousands of phosphopeptides in a single run. Still, the intrinsic issue of lack of redundancy in data representation for each phosphopeptide remains, leading to lack of statistical power for generating high confidence quantitation and identification for large portions of the dataset, especially for low abundance phosphopeptides that often rely on single PSMs with noisy signals. This issue has been tackled predominantly by requiring higher numbers of PSMs per phosphopeptide, with the consequent trade-off of eliminating a substantial fraction of the dataset that may contain most of the biologically meaningful regulatory, and low abundant, events. This problem is especially salient for nuclear proteins involved in the DNA damage response which often exist at low levels in the cell[14]. In this work we present a workaround that allows for the efficient exclusion of technical noise and variation through the use of a reciprocal SILAC experiment, while allowing for the identification and quantitation of low abundance phosphopeptides. We leveraged the sensitivity of this pipeline by combining the proposed Bow-tie analysis with samples that had been pre-fractionated using HILIC chromatography. The result is a drastic expansion in coverage with concomitant reduction in error and technical variation in the overall quantitative data. This combination of high specificity, low-PSM phosphoproteomics with HILIC, which is particularly suited to phosphopeptide fractionation due to the hydrophilicity of the phosphate group[69] revealed ~ 15,000 unique phosphopeptides in a short fractionation schema (15 fractions). Importantly, we demonstrate the utility of this approach by identifying new Mec1-dependent signaling events in *S. cerevisiae*.

Central to the Bow-tie filtering strategy presented in this study is the use of metabolic labelling with stable isotopes (SILAC) and the consequent shift in mass of parent and fragment ions of phosphopeptides. Such a large delta mass between phosphopeptides in reciprocal experiments forces a stringent requirement in which phosphopeptide identification with inverted fold change in each experiment should also exhibit proper delta mass shift. In addition to allowing efficient detection of EVs, this approach also led to a dramatic reduction in the number of decoy database peptide identifications in quadrants 1 and 3. This serves as definitive proof that

reciprocal labeling reduces false-positive identification and associated quantitation. False-positives identifications are proposed to represent either artefacts, exogenous sample contaminants not represented in the searched database, or containing other types of modifications not considered in our search as variable modifications[70]. The Bow-tie approach applied to the mock dataset reduced false-positive hits to virtually zero, when Q1 and Q3 were considered and at a PeptideProphet score minimum score of 0.9, satisfying our needs for highly sensitive and comprehensive strategy to uncover phosphopeptides of low abundance and low PSM counts. A near-zero frequency of false-positive identifications appearing in Q1 and Q3 is essential to our SILAC-based approach, because peptide identification essentially serves as the most important gatekeeper of filtering meaningful biological data from technical noise.

The mass spectrometric data processing pipeline employed in this study relied on the trans-proteomic pipeline (TPP) suit of proteomic tools, including updated tools for peptide identification with Comet[71] and scoring with PeptideProphet[72]. For SILAC quantitation, we used the Xpress precursor ion intensity quantitation tool[73], and for phosphosite localization and scoring we used the newly described PTMProphet tool[74]. PTMProphet models the potential sites of phosphorylation independently of the spectrum identification provided by the search engine and calculates probabilities for each potential modification site. This feature allowed us to design additional steps in our R-based scripts for handling clustered S/T/Y residues, which is a common occurrence in phosphopeptides. Unambiguous, high-confidence phosphorylated S/T/Y residues with neighboring S/T/Y residues were kept separate; medium or low-confidence phosphorylated S/T/Y residues with adjacent S/T/Y residues were combined with their neighbors and considered in our subsequent analyses as a "cluster".

The ability of our Bow-tie approach to separate biologically meaningful phosphorylation from technical noise is exemplified by the observed regulation in our Mec1 phospho-mapping dataset. In contrast to the control dataset, in which there were a small number of points in quadrants 1 and 3, there were a number of regulated sites in Q1 in our Mec1 dataset (and much less in Q3). S/T-Q consensus motif sites were overrepresented in Q1, indicating primary Mec1-dependent phosphorylation in response to DNA damage that was ablated in the absence of the *MEC1* gene. In addition to revealing many known Mec1 targets identified in other studies, which was our intention as a validation of our method, we revealed a number of previously unreported proteins with Mec1-dependent phosphorylation events, including a subset of nucleolar proteins. For example, we identified phosphorylation on serine 1007 (an S/T-Q site) of Kre33, a relatively understudied protein that promotes maturation of 18S rRNA[75,76]. Future work should be targeted toward understanding how Mec1 signaling contributes to nuclear homeostasis independently of its established roles in activation of the DNA damage checkpoint. One new kinase target of Mec1 present in our dataset is Yak1, which we found to be phosphorylated on Serine S663, near Yak1's kinase domain. Yak1 is a member of the family of Ser/Thr protein kinases known as dual-specificity Tyr phosphorylation-regulated kinases (DYRKs). Yak1 has been described as a growth antagonist downstream of Ras/PKA pathway, phosphorylated by PKA and translocated to the nucleus upon nutrient deprivation[77]. Indeed, cells lacking *YAK1* are sensitive to acute heat stress[60]. Intriguingly, cells lacking *MEC1* are sensitive to proteotoxic and heat stress[59]. No previous reports have linked Yak1 to the DNA damage response or to Mec1, and we speculate that this could be a new point of crosstalk between DNA damage signaling and cellular stress responses.

In summary, here we report a simple, robust SILAC-based phosphoproteomic data analysis pipeline that allows for identification and quantitation of phosphopeptides with high confidence and coverage. The depth of the analyses allowed identification of a range of novel Mec1-dependent signaling events, including a potentially new mode of Mec1 signaling targeting the nucleolus. While this work highlights the utility of SILAC for high confidence and in depth quantitative phosphoproteomics, the same rationale could be applied to improve the quantitative analysis of other low-abundance post-translational modifications such as sumoylation, ubiquitylation, and acetylation.

Materials and methods

Yeast cell culture and manipulation.
A list of yeast strains used in this study is found in Supplemental Table S6. The strain background for all yeast used was S288C. We performed whole ORF deletions of *MEC1* and *RAD9* using established PCR-based methods for amplifying resistance cassettes containing homology to the target gene. Gene manipulations were verified by PCR. Primers used for gene deletions are available upon request. Yeast were grown at 30 °C in synthetic SILAC media lacking arginine and lysine and supplemented with "light" lysine and arginine (^{12}C and ^{14}N) or supplemented with "heavy" lysine and arginine (L-lysine ^{13}C$_6$,^{15}N$_2$·HCl and L-arginine ^{13}C$_6$,^{15}N$_4$·HCl). Media was also supplemented with excess L-proline to prevent conversion of arginine to proline.

Sample preparation for phosphoproteomic analysis.
200–300 mL of yeast was grown in either "heavy" or "light" SILAC media to mid-log phase and treated as described in the figure legend and the text, depending on the experiment. Cells were pelleted at 1000×g and washed once with TE (10 mM Tris pH 8.0, 5 mM EDTA) buffer containing 1 mM PMSF. Cells were lysed by bead beating with 0.5 mm glass beads for 3 cycles of 10 min with 1-min rest time between cycles at 4 °C in lysis buffer (150 mM NaCl, 50 mM Tris pH 8.0, 5 mM EDTA, 0.2% Tergitol type NP40) supplemented with protease inhibitor cocktail (Pierce), 5 mM sodium fluoride and 10 mM β-glycerophosphate. 5–7 mg of each light and heavy labeled protein lysate was denatured and reduced with 1% SDS and 5 mM DTT at 42 °C, then alkylated with 25 mM iodoacetamide. Lysates (light and heavy) were mixed and precipitated with a cold solution of 50% acetone, 49.9% ethanol, 0.1% acetic acid. Post-precipitation protein pellet was then resuspended in 2 M urea and subsequently digested with TPCK-treated trypsin overnight at 37 °C. Phosphoenrichment was performed using a High-Select Fe-NTA phosphopeptide enrichment kit (ThermoFisher Scientific, cat# A32992) as described in the manufacturer's instructions. Purified phosphopeptides were then dried in a SpeedVac and fractionated via HILIC chromatography as described below.

HILIC fractionation.   Dried phosphopeptide samples were reconstituted in 15 μL H₂O, 10 μL 10% formic acid (v/v), and 60 μL HPLC-grade acetonitrile. 80 μL of the reconstituted sample was injected and fractionated by hydrophilic interaction liquid chromatography (HILIC) using a TSK gel Amide-80 column (2 mm × 150 mm, 5 μm; Tosoh Bioscience). Three solvents were used for the gradient: buffer A (90% acetonitrile), buffer B (75% acetonitrile and 0.005% trifluoroacetic acid), and buffer C (0.025% trifluoroacetic acid). A short gradient was used for the mock control and Mec1 experiments and consisted of 100% buffer A at time = 0 min; 88% of buffer B and 12% of buffer C at time = 5 min; 60% of buffer B and 40% of buffer C at time = 30 min; and 5% of buffer B and 95% of buffer C from time = 35 to 45 min in a flow of 150 μl/min. 30-s fractions were collected between 9 and 18 min. Individual fractions were dried in speedvac and submitted to LC–MS/MS analysis.

Phosphoproteomics data acquisition.   Individual phosphopeptide fractions were resuspended in 0.1% trifluoroacetic acid and subjected to LC–MS/MS analysis in an UltiMate 3000 RSLC nano chromatographic system coupled to a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific). The chromatographic separation was carried out in 35-cm-long 100-μm inner diameter column packed in-house with 3 μm C₁₈ reversed-phase resin (Reprosil Pur C18AQ 3 μm). Q-Exactive HF was operated in data-dependent mode with survey scans acquired in the Orbitrap mass analyzer over the range of 380–1800 m/z with a mass resolution of 60,000 (at m/z 200). MS/MS spectra was performed selecting the top 15 most abundant + 2, + 3 or + 4 ions and a with an precursor isolation window of 2.0 m/z. Selected ions were fragmented by Higher-energy Collisional Dissociation (HCD) with normalized collision energies of 28 and the mass spectra acquired in the Orbitrap mass analyzer with a mass resolution of 15,000 (at m/z 200), AGC target set to 1e⁵ and max injection time set to 120 ms. A dynamic exclusion window was set for 30 s.

Phosphopeptide and phosphosite identification.   The peptide identification and quantification pipeline relied on TPP tools[78]. The search engine used was Comet (v. 2019.01.1)[71]. Search parameters included semi-tryptic requirement, 20 ppm for the precursor match tolerance, differential mass modification of 8.0142 for lysine, 10.00827 for arginine, 79.966331 for phosphorylation of serine, threonine and tyrosine, 97.976896 for phosphorylation dehydration, and static mass modification of 57.021465 for alkylated cysteine residues. The protein sequence database was the SGD yeast supplemented with the decoy reversed sequences and common contaminants (downloaded in Aug 2019, 11,968 entries). Original ThermoScientific .raw files were converted to mzXML before the search with Comet. After searches, peptides were filtered and scored by the PeptideProphet algorithm[72] using the following parameters: minimum probability of 0.9, minimum peptide length of 7 amino acid residues, accurate mass binning, restriction to + 2, + 3 and + 4 ion charge states and Phospho-Information enabled. After scoring and filtering, relative quantitation based on SILAC were obtained using Xpress and specific parameters were: mass tolerance of 0.005 daltons; minimum number of chromatogram points needed for quantitation = 1; number of isotopic peaks = 0. Phosphopeptides were then evaluated by PTMProphet[74] in order to obtain accurate phosphosite localization score. The complete lists of identified, quantified, scored, and filtered phosphopeptides were further processed using a R-script developed in-house. The script separates phosphosites with high PTMProphet probability (> 0.9) from those with ambiguous localization containing 2 or more adjacent potentially phosphorylated residues, here denominated "clusters". Separately, high confidence phosphosites and clustered phosphosites had their SILAC quantitation median calculated and additional R-scripts were used for combining, correlating, and plotting the data.

Estimation of false discovery rate (FDR) in quantitative analysis.   All points (from the mock and Mec1 experiment) belonging to quadrant 2 and 4 are removed along with all the points that have fold change (FC) of less than or equal to 2 and hence would lie in a circle with radius 1 (since the scale is log transformed FC). The points in quadrant 1 and quadrant 3 are combined in order to get symmetric parabolic bins. A choice of aperture is made by sampling this space using multiple parabolas rotated at 45° (~ 0.78 radians) with their vertex on the circle with radius 1. This is done to ensure accordance with the underlying assumption that the highest confidence points would lie far away from origin along the line of symmetry y = − x. False discovery rate (FDR) is calculated as the percentage of false positives given by the mock experiment to the false positives and true positives given by the Mec1 experiment that lie within each parabola. The false positives are indicated with red color and the true positives are indicated with green color in Supplementary Figure S1. The parabolic bins that gave FDR values closest to commonly used FDR values (5% and 2%) were retained and the bin aperture that gave a 2% FDR is then used to further sample the space by varying the vertex of the parabolas. The vertices were chosen so that the obtained FDR would be the first local minimum within an FDR range. This was done to ensure that the number of false positives are minimized, and the number of true positives are maximized.

String analysis of S/T-Q and non-S/T-Q motif Mec1-dependent sites.   A subset of phosphorylation sites (e.g. all S/T-Q sites in the experiment from Fig. 5 with a log2 ratio > 1) was selected and the list of gene names uploaded to https://string-db.org/. In cases where there were multiple sites under the same gene name entry, the gene name was used only once. Interaction networks were generated considering only high confidence interactions (score > 0.700). Next, the genes in the list corresponding to a specific biological process or pathway (e.g. nucleolus) were again uploaded to https://string-db.org/.

Uniprot keyword enrichment analysis of S/T-Q and non-S/T-Q motif Mec1-dependent sites.   A subset of phosphorylation sites (e.g. all S/T-Q sites in the experiment from Fig. 5 with a log2 ratio > 1) was selected and the list of gene names uploaded to https://string-db.org/. In cases where there were multiple

sites under the same gene name entry, the gene name was used only once. Interaction networks were generated considering only high confidence interactions (score > 0.700). Next, the top 8-ranked Uniprot Keyword enrichment terms were exported along with the associated false discovery rate (FDR). For visualization, the FDR was log10-transformed. The terms "Nucleus" and "Phosphoprotein" were manually excluded from the figure because they represent processes that are too general to be informative.

Data availability

References

1. Day, E. K., Sosale, N. G. & Lazzara, M. J. Cell signaling regulation by protein phosphorylation: A multivariate, heterogeneous, and context-dependent process. *Curr. Opin. Biotechnol.* **40**, 185–192 (2016).
2. Krebs, E. G. & Fischer, E. H. The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *BBA Gen. Subj.* **20**, 150–157 (1956).
3. Taylor, S. S., Keshwani, M. M., Steichen, J. M. & Kornev, A. P. Evolution of the eukaryotic protein kinases as dynamic molecular switches. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2517–2528 (2012).
4. Flott, S. *et al.* Regulation of Rad51 function by phosphorylation. *EMBO Rep.* **12**, 833–839 (2011).
5. Osborn, A. J. *et al.* Checking on the fork: The DNA-replication stress–response pathway. *Trends Cell Biol.* **12**, 509–516 (2002).
6. Schwartz, M. F. *et al.* Rad9 phosphorylation sites couple Rad53 to the *Saccharomyces cerevisiae* DNA damage checkpoint. *Mol. Cell* **9**, 1055–1065 (2002).
7. Memisoglu, G. *et al.* Mec1ATR autophosphorylation and Ddc2ATRIP phosphorylation regulates DNA damage checkpoint signaling. *Cell Rep.* **28**, 1090–1102 (2019).
8. Ohouo, P. Y., Bastos de Oliveira, F. M., Almeida, B. S. & Smolka, M. B. DNA damage signaling recruits the Rtt107-Slx4 scaffolds via Dpb11 to mediate replication stress response. *Mol. Cell* **39**, 300–306 (2010).
9. Toh, G.W.-L. *et al.* Mec1/Tel1-dependent phosphorylation of Slx4 stimulates Rad1–Rad10-dependent cleavage of non-homologous DNA tails. *DNA Repair Amst.* **9**, 718–726 (2010).
10. Weinert, T. A., Kiser, G. L. & Hartwell, L. H. Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair. *Genes Dev.* **8**, 652–665 (1994).
11. BastosdeOliveira, F. M. *et al.* Phosphoproteomics reveals distinct modes of Mec1/ATR signaling during DNA replication. *Mol. Cell* **57**, 1124–1132 (2015).
12. Lanz, M. C. *et al.* Separable roles for Mec1/ATR in genome maintenance, DNA replication, and checkpoint signaling. *Genes Dev.* **32**, 822–835 (2018).
13. Chen, S. H., Albuquerque, C. P., Liang, J., Suhandynata, R. T. & Zhou, H. A proteome-wide analysis of kinase-substrate network in the DNA damage response. *J. Biol. Chem.* **285**, 12803–12812 (2010).
14. Ho, B., Baryshnikova, A. & Brown, G. W. Unification of protein abundance datasets yields a quantitative *Saccharomyces cerevisiae* proteome. *Cell Syst.* **6**, 192–205 (2018).
15. Kelstrup, C. D. *et al.* Performance evaluation of the Q exactive HF-X for shotgun proteomics. *J. Proteome Res.* **17**, 727–738 (2018).
16. Deutsch, E. W. *et al.* Trans-proteomic pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteom. Clin. Appl.* **9**, 745–754 (2015).
17. Hogrebe, A. *et al.* Benchmarking common quantification strategies for large-scale phosphoproteomics. *Nat. Commun.* **9**, 1045. https://doi.org/10.1038/s41467-018-03309-6 (2018).
18. Li, J., Paulo, J. A., Nusinow, D. P., Huttlin, E. L. & Gygi, S. P. Investigation of proteomic and phosphoproteomic responses to signaling network perturbations reveals functional pathway organizations in yeast. *Cell Rep.* **29**, 2092–2104 (2019).
19. Sharma, K. *et al.* Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep.* **8**, 1583–1594 (2014).
20. Humphrey, S. J., Karayel, O., James, D. E. & Mann, M. High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nat. Protoc.* **13**, 1897–1916 (2018).
21. Ochoa, D. *et al.* The functional landscape of the human phosphoproteome. *Nat. Biotechnol.* **38**, 365–373 (2019).
22. Balakrishnan, R. *et al.* YeastMine—an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. *Database* https://doi.org/10.1093/database/bar062 (2012).
23. Hu, Z. *et al.* Multilayered control of protein turnover by TORC1 and Atg1. *Cell Rep.* **28**, 3486–3496 (2019).
24. Lanz, M. C., Yugandhar, K., Gupta, S., Sanford, E. & Faça, V. In-depth and 3-dimensional exploration of the budding yeast phosphoproteome. *bioRxiv* https://doi.org/10.1101/700070 (2019).
25. Bastos de Oliveira, F. M., Kim, D., Lanz, M. & Smolka, M. B. Quantitative analysis of DNA damage signaling responses to chemical and genetic perturbations. *Methods Mol. Biol.* **1672**, 645–660 (2018).
26. Hertz, N. T. *et al.* Chemical genetic approach for kinase-substrate mapping by covalent capture of thiophosphopeptides and analysis by mass spectrometry. *Curr. Protoc. Chem. Biol.* **2**, 15–36 (2010).
27. Shinde, M. Y. *et al.* Phosphoproteomics reveals that glycogen synthase kinase-3 phosphorylates multiple splicing factors and is associated with alternative splicing. *J. Biol. Chem.* **292**, 18240–18255 (2017).
28. Amanchy, R. *et al.* Identification of c-Src tyrosine kinase substrates using mass spectrometry and peptide microarrays. *J. Proteome Res.* **7**, 3900–3910 (2008).
29. Schwill, M. *et al.* Systemic analysis of tyrosine kinase signaling reveals a common adaptive response program in a HER2-positive breast cancer. *Sci. Signal.* **12**, eaau2875. https://doi.org/10.1126/scisignal.aau2875 (2019).
30. Pease, B. N. *et al.* Characterization of plasmodium falciparum atypical kinase PfPK7-dependent phosphoproteome. *J. Proteome Res.* **17**, 2112–2123 (2018).
31. Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteom.* **1**, 376–386 (2002).
32. Mann, M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
33. Casado, P. & Cutillas, P. R. A self-validating quantitative mass spectrometry method for assessing the accuracy of high-content phosphoproteomic experiments. *Mol. Cell. Proteom.* https://doi.org/10.1074/mcp.M110.003079 (2011).

34. Chen, X., Wei, S., Ji, Y., Guo, X. & Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **15**, 3175–3192 (2015).

35. Sandberg, A. S., Branca, R. M. M., Lehtiö, J. & Forshed, J. Quantitative accuracy in mass spectrometry based proteomics of complex samples: The impact of labeling and precursor interference. *J. Proteom.* **96**, 133–144 (2014).

36. Li, Z. *et al.* Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ orbitrap velos. *J. Proteome Res.* **11**, 1582–1590 (2012).

37. Wong, C. C. L., Cociorva, D., Venable, J. D., Xu, T. & Yates, J. R. Comparison of different signal thresholds on data dependent sampling in orbitrap and LTQ mass spectrometry for the identification of peptides and proteins in complex mixtures. *J. Am. Soc. Mass Spectrom.* **20**, 1405–1414 (2009).

38. Ong, S. E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **1**, 2650–2660 (2006).

39. Francavilla, C., Hekmat, O., Blagoev, B. & Olsen, J. V. SILAC-based temporal phosphoproteomics. *Methods Mol. Biol.* **1188**, 125–148 (2014).

40. Aggelis, V. *et al.* Proteomic identification of differentially expressed plasma membrane proteins in renal cell carcinoma by stable isotope labelling of a von Hippel-Lindau transfectant cell line model. *Proteomics* **9**, 2118–2130 (2009).

41. Alli-Shaik, A., Wee, S., Lim, L. H. K. & Gunaratne, J. Phosphoproteomics reveals network rewiring to a pro-adhesion state in annexin-1-deficient mammary epithelial cells. *Breast Cancer Res.* **19**, 132. https://doi.org/10.1186/s13058-017-0924-4 (2017).

42. Park, S. S. *et al.* Effective correction of experimental errors in quantitative proteomics using stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteom.* **75**, 3720–3732 (2012).

43. Friedel, A. M., Pike, B. L. & Gasser, S. M. ATR/Mec1: Coordinating fork stability and repair. *Curr. Opin. Cell Biol.* **21**, 237–244 (2009).

44. Lanz, M. C., Dibitetto, D. & Smolka, M. B. DNA damage kinase signaling: Checkpoint and repair at 30 years. *EMBO J.* **38**, 101801. https://doi.org/10.15252/embj.2019101801 (2019).

45. Pardo, B., Crabbé, L. & Pasero, P. Signaling pathways of replication stress in yeast. *FEMS Yeast Res.* https://doi.org/10.1093/femsyr/fow101 (2017).

46. Toh, G. W. L. & Lowndes, N. F. Role of the *Saccharomyces cerevisiae* Rad9 protein in sensing and responding to DNA damage. *Biochem. Soc. Trans.* **31**, 242–246 (2003).

47. Smolka, M. B., Albuquerque, C. P., Chen, S. H. & Zhou, H. Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. *Proc. Natl. Acad. Sci. USA* **104**, 10364–10369 (2007).

48. Alcasabas, A. A. *et al.* Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nat. Cell Biol.* **3**, 958–965 (2001).

49. Smolka, M. B. *et al.* Dynamic changes in protein-protein interaction and protein phosphorylation probed with amine-reactive isotope tag. *Mol. Cell. Proteom.* **4**, 1358–1369 (2005).

50. Szklarczyk, D. *et al.* STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

51. Mimitou, E. P. & Symington, L. S. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**, 770–774 (2008).

52. Cannavo, E. & Cejka, P. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* **514**, 122–125 (2014).

53. West, S. C. *et al.* Resolution of recombination intermediates: Mechanisms and regulation. *Cold Spring Harb. Symp. Quant. Biol.* **80**, 1–7 (2016).

54. Hickson, I. D. & Mankouri, H. W. Processing of homologous recombination repair intermediates by the Sgs1–Top3-Rmi1 and Mus81–Mms4 complexes. *Cell Cycle* **10**, 3078–3085 (2011).

55. Bermúdez-López, M. *et al.* Sgs1's roles in DNA end resection, HJ dissolution, and crossover suppression require a two-step SUMO regulation dependent on Smc5/6. *Genes Dev.* **30**, 1339–1356 (2016).

56. Chen, S. H., Smolka, M. B. & Zhou, H. Mechanism of Dun1 activation by Rad53 phosphorylation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **282**, 986–995 (2007).

57. Zhao, X. & Rothstein, R. The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. *Proc. Natl. Acad. Sci. USA* **99**, 3746–3751 (2002).

58. Andreson, B. L., Gupta, A., Georgieva, B. P. & Rothstein, R. The ribonucleotide reductase inhibitor, Sml1, is sequentially phosphorylated, ubiquitylated and degraded in response to DNA damage. *Nucleic Acids Res.* **38**, 6490–6501 (2010).

59. Corcoles-Saez, I. *et al.* Essential function of Mec1, the budding yeast ATM/ATR checkpoint-response kinase, protein homeostasis. *Dev. Cell* **46**, 495–503 (2018).

60. Hartley, A. D., Ward, M. P. & Garrett, S. The Yak1 protein kinase of *Saccharomyces cerevisiae* moderates thermotolerance and inhibits growth by an Sch9 protein kinase-independent mechanism. *Genetics* **136**, 465–474 (1994).

61. Lee, K., Zhang, Y. & Lee, S. E. *Saccharomyces cerevisiae* ATM orthologue suppresses break-induced chromosome translocations. *Nature* **454**, 543–546 (2008).

62. Mallory, J. C. & Petes, T. D. Protein kinase activity of Tel1p and Mec1p, two *Saccharomyces cerevisiae* proteins related to the human ATM protein kinase. *Proc. Natl. Acad. Sci. USA* **97**, 13749–13754 (2000).

63. Morrow, D. M., Tagle, D. A., Shiloh, Y., Collins, F. S. & Hieter, P. TEL1, an *S. cerevisiae* homolog of the human gene mutated in ataxia telangiectasia, is functionally related to the yeast checkpoint gene MEC1. *Cell* **82**, 831–840 (1995).

64. Sridhar, A., Kedziora, S. & Donaldson, A. D. At short telomeres Tel1 directs early replication and phosphorylates Rif1. *PLoS Genet.* **10**, 1004691. https://doi.org/10.1371/journal.pgen.1004691 (2014).

65. Myung, K., Datta, A. & Kolodner, R. D. Suppression of spontaneous chromosomal rearrangements by S phase checkpoint functions in *Saccharomyces cerevisiae*. *Cell* **104**, 397–408 (2001).

66. Gobbini, E., Cesena, D., Galbiati, A., Lockhart, A. & Longhese, M. P. Interplays between ATM/Tel1 and ATR/Mec1 in sensing and signaling DNA double-strand breaks. *DNA Repair (Amst).* **12**, 791–799 (2013).

67. Ozeri-Galai, E., Schwartz, M., Rahat, A. & Kerem, B. Interplay between ATM and ATR in the regulation of common fragile site stability. *Oncogene* **20**, 20 (2008).

68. Fenselau, C., Heller, D. N., Miller, M. S. & White, H. B. Phosphorylation sites in riboflavin-binding protein characterized by fast atom bombardment mass spectrometry. *Anal. Biochem.* **150**, 309–314 (1985).

69. McNulty, D. E. & Annan, R. S. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol. Cell. Proteom.* **7**, 971–980 (2008).

70. Chick, J. M. *et al.* A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nat. Biotechnol.* **33**, 743–749 (2015).

71. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: An open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).

72. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).

73. Han, D. K., Eng, J., Zhou, H. & Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19**, 946–951 (2001).

74. Shteynberg, D. D. *et al.* PTMProphet: Fast and accurate mass modification localization for the trans-proteomic pipeline. *J. Proteome Res.* **18**, 4262–4272 (2019).
75. Sharma, S. *et al.* Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. *Nucleic Acids Res.* **43**, 2242–2258 (2015).
76. Pagé, N. *et al.* A *Saccharomyces cerevisiae* genome-wide mutant screen for altered sensitivity to K1 killer toxin. *Genetics* **163**, 875–894 (2003).
77. Lee, P., Paik, S. M., Shin, C. S., Huh, W. K. & Hahn, J. S. Regulation of yeast Yak1 kinase by PKA and autophosphorylation-dependent 14-3-3 binding. *Mol. Microbiol.* **79**, 633–646 (2011).
78. Deutsch, E. W. *et al.* A guided tour of the trans-proteomic pipeline. *Proteomics* **10**, 1150–1159 (2010).
79. Carbon, S. *et al.* The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
80. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

Acknowledgements

Author contributions

V.M.F. and M.B.S. conceptualized the project and the pipeline. E.J.S. and W.J.C. performed experiments. All authors contributed to data analysis. J.T. and V.M.F. wrote the data analysis script. E.J.S., V.M.F., and M.B.S. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-74939-4.

Correspondence and requests for materials should be addressed to M.B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.