Published in final edited form as:
Nat Biotechnol. 2022 March 01; 40(3): 382–390. doi:10.1038/s41587-021-01051-x.

High-throughput functional characterization of protein phosphorylation sites in yeast

Cristina Viéitez[1,2,*], **Bede P. Busby**[1,2,*], **David Ochoa**[1], **André Mateus**[2], **Danish Memon**[1], **Marco Galardini**[1], **Umut Yildiz**[2,3], **Matteo Trovato**[2,3], **Areeb Jawed**[2], **Alexander G. Geiger**[4], **Michaela Oborská-Oplová**[4,5], **Clement M. Potel**[2], **Sibylle C. Vonesch**[2], **Chelsea Szu Tu**[2], **Mohammed Shahraz**[1], **Frank Stein**[2], **Lars M. Steinmetz**[2,6,7], **Vikram G. Panse**[4,8], **Kyung-Min Noh**[2], **Mikhail M. Savitski**[2,#], **Athanasios Typas**[2,#], **Pedro Beltrao**[1,2,#]

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, CB10 1SD, Cambridge, UK

[2]European Molecular Biology Laboratory, Genome Biology Unit, 69117 Heidelberg, Germany

[3]Faculty of Biosciences, Heidelberg, Germany

[4]Institute of Medical Microbiology, University of Zurich, Zurich, Switzerland

[5]Institute of Biochemistry, ETH Zurich, Zurich, Switzerland

[6]Stanford Genome Technology Center, Stanford University, Palo Alto, California 94304, USA

[7]Department of Genetics, School of Medicine, Stanford University, Stanford, California 94305, USA

[8]Faculty of Science, University of Zurich, Zurich, Switzerland

Abstract

Phosphorylation is a critical post-translational modification involved in the regulation of almost all cellular processes. However, less than 5% of thousands of recently discovered phosphosites have been functionally annotated. Here, we devised a chemical genetic approach to study the functional relevance of phosphosites in S. cerevisiae. We generated 474 yeast strains with mutations in specific phosphosites that were screened for fitness in 102 conditions, along with a gene deletion

[#]Correspondence to: Mikhail M. Savitski (mikhail.savitski@embl.de), Athanasios Typas (typas@embl.de) and Pedro Beltrao (pbeltrao@ebi.ac.uk).
[*]Contributed equally

Reporting Summary
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Author Contributions
CV and BPB constructed the phospho-mutant library, performed the chemical genomics screen and spots assays with assistance from AJ and MS. DO, MG and DM performed computational analysis. CV and AM performed the 2D-TPP experiments with assistance from CMP. FS analyzed the proteomics data. BPB performed the lipidomics experiments. UY and MT generated the human phospho-mutant cell lines. CV, UY and MT performed viability assays on human cell line clones. AGG and MO-O generated the humanized phospho-mutant yeast strains and performed the FISH experiments. LS, VGP, K-MN, MMS, AT and PB supervised the project. CST and SV performed genome sequencing. PB and CV wrote the manuscript with assistance from all authors.

Competing interests
The authors declare no conflict of interests.

Europe PMC Funders Author Manuscripts

library. Of these phosphosites, 42% exhibited growth phenotypes, suggesting that these are more likely functional. We inferred their function based on the similarity of their growth profiles with that of gene deletions, and validated a subset by thermal proteome profiling and lipidomics. A high fraction exhibited phenotypes not seen in the corresponding gene deletion suggestive of a gain-of-function effect. For phosphosites conserved in humans, the severity of the yeast phenotypes is indicative of their human functional relevance. This high-throughput approach allows for functionally characterizing individual phosphosites at scale.

Introduction

Cells are constantly sensing and adapting to changes in environmental conditions. The relay of information from sensors to effector proteins often occurs via reversible protein post-translational modification (PTM), including protein phosphorylation. Regulation by phosphorylation allows cells to modulate protein activity, interactions and localization. In the past decade, mass spectrometry (MS)-based discovery of protein phosphorylation sites (phosphosites) has drastically changed our understanding of the extent of protein phosphorylation [1] . Recent estimates suggest that up to 75% of the yeast and human proteomes are phosphorylated, with over 10,000 and 150,000 non-redundant phosphosites described to date in the two organisms, respectively [2,3] . This vast extent of phosphorylation in eukaryotes has opened a debate on its physiological relevance [4–7] , fueled by the fact that even for yeast we currently know the functional roles for less than 5% of these sites [2] . It has been speculated that a significant fraction of phosphosites may have no function [4] , consistent with the fast rate of evolutionary change in phosphorylation detected across species [4,7–10] . Evolutionary studies have suggested that between 35% to 65% of sites are constrained and therefore functional [4,11] . The degree of functional relevance and the functional role of most phosphosites, thus remain open questions.

Computational approaches have been developed to try to bridge the gap of functional assignment [12–14] . For example, prediction of phosphosites that may regulate protein interactions or cross-regulatory interactions with other PTMs have been tried [7,15,16] . In addition, the conservation of phosphosites across species has been used to infer functionality, as phosphosites with well characterized functions are more likely to be conserved [9,17] . In contrast, experimental approaches that functionally characterize PTM sites in a comprehensive manner are still lacking or are geared to a small number of phosphosites [18,19] .

Reverse genetics of gene deletion strains have been instrumental for the systematic elucidation of gene function [20–22] . Similarly, reverse genetic screens of point mutants have shown promise for functional studies of protein regions [23] . Inspired by these studies, we set out to study the functional role of protein phosphorylation through a combination of chemical-genomics, molecular profiling and computational analysis. We created a yeast strain collection of phospho-deficient mutants that we screened for fitness phenotypes, together with the single-gene deletion library. In these assays 43% of the phospho-mutants displayed at least one significant growth phenotype when screened in 102 conditions, suggesting that these are likely functional. The correlation of the growth profiles across the

Europe PMC Funders Author Manuscripts

102 conditions, with those elicited by gene deletions was indicative of phosphosite function. From this large resource, we performed further characterization of the molecular changes elicited by some phospho-mutants by proteomics and lipidomics which further confirmed the inferences made from the genetic screen. Finally, the severity of the yeast phenotypes was indicative of the functional importance of orthologous human phosphosites.

Results

Phenotypic profiling of a phospho-deficient mutant library

In order to study the functional role of protein phosphorylation using chemical-genomic approaches, we generated a library of phospho-deficient mutant strains in *S. cerevisiae* (Fig. 1a and Supplementary Data 1). A compilation of experimentally determined *in vivo* phosphorylation sites was used to select 500 unique phosphosite positions including sites estimated to have different evolutionary ages [9] ; regulated *in-vivo* across different conditions; that were part of different structural elements; and within genes involved in a broad range of cellular processes (Supplementary Fig. 1). A total of 474 unique phospho-mutant strains, in 125 diverse genes, were generated with each position mutated to alanine at the genomic locus using a nearly scarless approach [24] (Supplementary Fig. 1).

The phospho-mutant library was arrayed as single colonies in a 1536 format and combined with the *S. cerevisiae* gene deletion library (4889 KOs) [25] . These were plated on a panel of 102 different stress conditions that were selected to span a wide range of diverse perturbations (Fig. 1b and Supplementary Data 2), including environmental, chemical or metabolic stresses (e.g. heat, DNA damage or nutrient starvation), drugs (e.g. antibiotics) and combination of stresses. Colony size was used as a proxy of fitness and a deviation from expected growth of each strain in each condition was computed using the S-score [22] (Methods and Fig. 1b). Positive and negative S-scores indicate that the mutant has increased or decreased resistance to stress, respectively. The resulting growth profiles (i.e. S-scores across all conditions) are shown as clustered heatmaps, separate for the phospho-mutants and KOs, indicating a diverse set of responses for both groups (Fig. 1c, available in Supplementary Data 3). The number of mutant phenotypes per condition is shown in Supplementary Fig. 2a.

The measured growth profiles were highly reproducible as shown by the correlation of S-scores across biological replicates for gene deletions (Supplementary Data 4 and Fig. 1d, r=0.775) and phospho-mutants (Supplementary Data 4 and Fig. 1d, r=0.698). In addition, the measurements recapitulate expected mutant phenotypes. For example, deletions of either of two key kinases of the high osmolarity glycerol pathway (Hog1 and Pbs2) and a mutation of a key regulatory phosphosite in Hog1 displayed severe growth defects under osmotic shock conditions (Fig. 1e and Fig.1g). In contrast, deleting the opposing phosphatase of the pathway (Ptc1) increased resistance in the same conditions. Similarly, we observed the expected sensitivity to grow under heat stress for deletion of Mck1 [26] and a mutant of the regulatory phosphosite of this kinase (Fig. 1e and Fig. 1g). To further validate the high-throughput growth measurements, we performed independent, albeit less sensitive spot assays to test 762 pairs of mutant-condition phenotypes (Methods, Fig. 1f and

Supplementary Data 5). The spot assays were analyzed and given a semi-quantitative score that was largely consistent with the high-throughput growth phenotypes (Fig. 1f, r=0.59).

Elucidating phosphosite function using growth profiles

Across all conditions, 43% of the phospho-mutants and 79% of the corresponding gene deletions showed at least one statistically significant phenotype (Fig. 2a). Across the different acceptor residues, 41.8 % (138/330) of serine, 41.4% (46/111) of threonine, and 60.6% (20/33) of tyrosine phospho-mutants exhibited phenotypes (Supplementary Fig. 2b). The number of phenotypes observed tended to be higher for phosphosites estimated to be older in evolutionary origin (Fig. 2a). Two observations suggest that more phosphosites may be functional than those captured by our assay: 5 out of 17 (29.4%) phosphosites reported to be functional did not show phenotypes (Fig. 2a); and rarefaction curve analysis indicated no saturation (Supplementary Fig. 2c), suggesting that additional conditions could yet reveal additional phenotypes. However, when manually re-curating literature we also noted that only 2 of the 17 phosphosites with described functions were actually tested as single point mutants for phenotypes. This suggests that for some of the 5 controls that did not show phenotypes in our screen, the literature annotated phenotypes may come from other mutations introduced in parallel.

Previous studies of gene deletion phenotypes have shown that the growth profiles across a large set of perturbations can be compared to assign functions to genes with unknown function based on a guilt-by-association principle [20–22,27,28] . The same principle can be applied to the study of single point mutants of phosphosites. For example, two point mutants in the Slt2 kinase (T190A and Y192A) show growth phenotypes that are highly similar to those elicited by the deletion of *SLT2* itself or the deletion of other members of the cell wall integrity pathway (Fig. 2b). These two positions are known to be critical for the full activation of the Slt2 kinase [29] . Another mutation (S414A) caused milder growth phenotypes in a subset of conditions, suggesting partial loss of function, whereas the remaining 5 point mutants did not show significant phenotypes. This result illustrates how the similarity of the growth profiles can be used for functional annotation of the phosphosites.

We classified each phospho-mutant according to the number of phenotypes and the extent by which the single point mutation shares the phenotypes (i.e. phenocopies) of the corresponding gene deletion (Fig. 2c). We identified a set of 36 phospho-mutants that phenocopy the corresponding gene deletion (“loss-of-function” mutants) and almost twice (68) the number of mutants that have a large number of significant growth phenotypes that are not mostly shared with the corresponding gene KO (“gain-of-function” mutants). The gain-of-function class also contains cases of partial-loss, where only some of the phenotypes of the gene KO are present. These gain-of-function mutants have growth profiles that can in turn correlate with other mutants which can be suggestive of their function. For each phosphosite mutated in our screen, we collected annotations related to conservation, condition specific regulation and structural properties (Supplementary Data 6). We then asked if any of these features are characteristic of either the gain- or loss-of-function group relative to a set of phosphosites without any significant phenotypes (Fig. 2d). Both

functional groups showed increased conservation relative to the no-phenotype group but the loss-of-function group was particularly more likely to have more conserved phosphosites. Relative to the no-phenotype group, the gain-of-function phosphosites are more likely to be regulated across a larger set of conditions when compared to the loss-of-function group.

In order to assign a putative functional role to each phosphosite, we computed all pairwise correlations of phenotypic profiles and built a network where two mutants are connected by an edge if they have significantly similar profiles (Fig. 2e). Based on this analysis, we provide a putative functional classification to phospho-mutants using the functional annotations of genes with similar growth profiles (Supplementary Data 7). As expected, the mutants tended to group in the network based on their functional characteristics (Fig. 2e) as identified by the Systematic Functional Annotation (SAFE) algorithm [30] . In this network we find cases such as the loss-of-function *VMA2-Y284A* with phenotypes that resemble the deletion of *VMA2* itself or other members of the vacuolar H+-ATPase complex. In contrast, the *INO1-S368A* mutant has phenotypes that resemble that of the deletion of several amino-acid biosynthesis genes instead of the deletion of *INO1* itself, due to *INO1-S368A* having only partially overlapping phenotypes with the *INO1* deletion.

We next investigated the incidence of additional mutations in the library and their potential phenotypic impact. We performed spot-tests on strains by PCR and whole ORF sequencing and carried out whole genome sequencing (Methods), generating high coverage sequence data for 244 phospho-mutant strains (Supplementary Data 8). We observed additional non-synonymous mutations in 132 strains (59 with phenotypes), with many of these mutations likely generated during amplification for sequencing. We observed 19 strains with an additional mutation in the phospho-mutant gene, 8 of these with phenotypes. Through detailed analysis, we defined 18 strains (3.8% of the library) with additional mutations that could be contributing to the phenotypes we observed as described in Supplementary Data 8. Furthermore, the added 18 base pairs (bp) at the 3'-UTR region of each mutant may, in some cases, perturb gene function. By measuring the average similarity of the mutant growth profiles with a control strain having only the 18 bp insertion (Supplementary Fig. 2d and Supplementary Data 8) we identified 3 genes (*CHD1*, *RPB2* and *SWI4*) where the insertion may account for phenotypes.

Discarding the 3 genes and the 18 strains flagged did not impact on the global quantitative results (Supplementary Fig. 3) with a similar fraction of phospho-mutants with phenotypes (38%), around twice the number of mutants classified as a gain vs loss of function (49 vs 23) with similar characteristics of conservation (Supplementary Fig. 3).

Phospho-mutant induced molecular changes

In order to associate the functional annotations derived from the growth profiles with molecular changes, we characterized 8 phospho-mutant strains, whose growth phenotypes were validated using two independently generated clones. Additionally, all KO strains used in the following experiments were validated via verification of their unique barcode sequence [31] . We used a combination of an in-depth proteomics methodology named two-dimensional thermal proteome profiling (2D-TPP) [32–34] , and lipidomics (Fig. 3a). In addition to the molecular data collected here (Supplementary Data 9), we analyzed

the 2D-TPP data collected previously on two of the mutants also included in this study (TDH3 S149A and T151A) [14]. The proteomics measurements allowed us to compare the thermal stability and abundance of thousands of proteins of each mutant strain relative to WT (Methods). For the lipidomics we quantified with LC-MS/MS the changes in 181 predicted lipid species (Supplementary Data 10 and 11), from which 19 were manually validated as described in Supplementary Data 12. We determined and compared the number of significant phenotypes and molecular changes in each mutant (Fig. 3b) noting a trend whereby mutants with the highest number of phenotypes had the largest number of protein and/or lipid changes. In particular, VMA2-Y284A and SIT4-S209A showed many molecular changes in these assays.

The in-depth characterization of protein thermal stability allowed us to ask if the phospho-mutant proteins tend to change their thermal stability due to the alanine mutation. None of them showed a significant effect in thermal stability relative to WT, however *SIT4-S209A* showed a measurable increase in thermal stability that did not pass the significance cut-off (Fig. 3c). We next compared the distribution of changes in protein abundance and thermal stability of protein interaction partners of the mutated proteins and overall did not detect a significant difference with the exception of *VMA2-Y284A* and *SIT4-S209A* (Supplementary Fig. 4), which were selected for further characterization. We then asked if the similarity of the phenotypic profiles for a pair of mutants is also reflected in a similar profile of molecular changes. We observe that this is true when looking at the protein abundance (Fig. 3d, r=0.43, p-value 0.022) or thermal stability changes (Fig. 3d, r=0.41, p-value 0.027) but not apparent when looking at the changes in lipid composition (not shown). The lack of correlation with the lipid measurements may be due to fewer molecular measurements and fewer strains analyzed when compared with the proteomics experiments. These results suggest that at least for the proteomics measurements there is an agreement between the phenotypic and molecular characterization that should allow for a more in-depth mechanistic explanation of the mutational effects.

Characterization of phosphosite mutants in Vma2 and Sit4

Vma2 is a subunit of the vacuolar H^+-ATPase (V-ATPase), a highly conserved proton pump critical for organelle pH homeostasis [35]. For Vma2 12 phospho-mutants were screened with 6 having phenotypes. One of these (Y284) resulted in phenotypes similar to the deletion of *VMA2* (r=0.70), and several complex members (Fig. 4a) representing an example of a loss-of-function mutant. A second mutant of *VMA2* (Y485), with a phenotypic profile that did not resemble the *VMA2* KO (r=0.20) was also selected for molecular characterization. Several growth phenotypes were validated by spot assays for these two mutants (Supplementary Fig. 5a).

In the proteomics data the Y284A mutant showed a strong reduction in protein levels of Vma2 itself and in 5 out of 8 subunits of the V_1 module (Fig. 4b), providing molecular evidence of the loss-of-function nature of this mutant. In contrast, Y485A showed no proteomic changes in *VMA2* or V-ATPase complex members, but it did show an increase in abundance of proteins enriched with annotations for mitochondrial localization (9/20) and transmembrane transport (GO:0055085, q-value<0.05; Supplementary Data 13). In

contrast, the Y284A mutant showed a decrease in abundance for proteins enriched in the same processes (GO:0055085,q-value<0.05; Supplementary Data 13), suggesting different functional roles for these two phosphosites in Vma2 (Supplementary Fig. 6). In the lipidomics data, *VMA2-Y284A* showed a significant reduction in the abundance of 12 lipids (Fig. 4c) including zymosterol, a precursor of ergosterol, previously described to be necessary for V-ATPase activity [36,37] . The reduction in zymosterol is consistent with the chemical genetic results, where we observed a high correlation between the *VMA2 KO*, *VMA2 Y284A* and 5 out of the 6 non-essential genes of the ergosterol biosynthesis pathway (Fig. 4a). In addition, the *VMA2-Y284A* mutant exhibited altered levels in 5 cardiolipin species, which have been previously suggested to play an important role in the maintenance, lubrication and/or activity the V-ATPase [38,39] . These changes were not shared by the *VMA2* KO (Supplementary Data 14) despite the fact that the KO and Y284A mutant shared most of the same growth phenotypes (Fig. 4a). Similar effects, albeit milder, were observed at the phenotypic and proteome level when we generated the Y284F phospho-deficient mutation, a more chemically conserved phospho-deficient mutation for tyrosine phosphosites (Supplementary Fig. 7). Overall, the molecular data further confirms the LoF effect of the Y284 mutant and the difference in growth phenotypes elicited by the two VMA2 Y284A and Y485A phospho-mutants.

Sit4 is a highly conserved PP2A-like phosphatase. Its activity and substrate specificity are regulated by the interaction with Sit4-associated proteins (SAPs) *SAP155*, *SAP185*, *SAP190* and *SAP4*. Deletion of all four SAPs together is equivalent to the deletion of *SIT4*, in terms of slower growth and cell cycle progression (as reviewed [40]). One phosphosite position in Sit4 (S209) has been reported [41] and mutated in our study. The S209A mutant shows 23 growth phenotypes with a growth profile that correlates strongly with only one of the SAP KOs (SAP155, r=0.72) (Fig. 4d). While S290A does not fully phenocopy *SIT4 KO*, it does share some of the same phenotypes such as reduced growth in heat stress and increased growth upon caffeine treatment. We also observed conditions where the S209A mutant had milder phenotypes than the *SIT4* KO such as reduced growth upon treatment with lipid biosynthesis inhibitors (ketoconazole and simvastatin). Several of these phenotypes were validated by spot assays (Supplementary Fig. 5b). In the 2D-TPP experiment, we observed a significant fraction of mitochondrial proteins changing in thermal stability, which agrees with the recently reported role of Sit4 regulating ATP synthase activity and mitochondrial function [42] . Sap155 was the only SAP protein showing a significant decrease in thermal stability in the S209A mutant (Fig. 4e). Together with the correlation of phenotypic profiles, this data suggests that S209A mutant has an impact on the Sit4-SAP155 phosphatase complex formation (Fig. 4f).

Functional annotation of conserved human phosphosites

We aligned the phospho-mutant genes to human orthologs identifying 72 yeast positions that aligned with 119 orthologous human phosphosites due to one-to-many gene orthology assignments. To query whether measuring the phenotypes of phospho-mutants in yeast could inform us on the functional relevance of human phosphosites, we used human information that we have recently integrated into a single score of phosphosite functional relevance [14] . We observed that human phosphosites that are predicted to be less tolerant to mutations

(SIFT p-value<0.05) or have a high predicted functional score are more likely to have extreme phenotypes when mutated in yeast (Fig. 5a). In addition, human phosphosites aligned with yeast phosphosites that have phenotypes when mutated are more likely to be regulated in cancer or under acute perturbations in human cells (fisher t-test p-value < 0.05, Supplementary Fig. 8). Overall, these results suggest that we can use the yeast fitness data to prioritize and annotate functionally important human phosphosites.

From the 72 yeast phospho-deficient mutants that align to a human phosphosite, we selected those with at least 2 significant phenotypes. For one-to-many cases we selected the human position with the highest functional score (Fig. 5b). We also obtained annotations on the number of conditions in which the human sites are regulated [43], the predicted impact of mutations [44] and the phosphosite functional score (Fig. 5b and Supplementary Data 15). This list contains human phosphosites with known functions such as the activation loop phosphosites of kinases (*MAPK14*-Y182/T180, *MAPK7*-Y221, *PRKX*-T203). To gain further insight into phosphosites conserved in human, two yeast positions (*GLO3*-S389 and *ELO2*-T334) and two human positions (RPS14 T133A and RPS14 S139A) from this list were included in the molecular characterization (Fig. 3 and Fig. 5). From the yeast phosphosites, we analyzed in more detail the T334A mutant of *ELO2* (a fatty acid elongase), which elicited a larger number of phenotypes (Fig. 5b) that were not strongly correlated with the corresponding gene deletion strain. A triple point mutant including this position had already been shown to impact on *ELO2* function [45] but the role of T334 has not yet been investigated on its own. Some of the growth differences between *ELO2-T334A* and *ELO2* KO were validated by spot assays. For example, the gene deletion shows resistance to an inhibitor of sterol biosynthesis (ketoconazole) while the T334A shows a modest increase in sensitivity (Supplementary Fig. 5c). We investigated the differences in lipid composition in these two mutants (Supplementary Data 14) finding a general trend of opposite impact on the phosphatidylinositol (PI) species, with an accumulation of PI (18:1/26:1) observed in the T334A and a depletion observed in the KO (Fig. 5c). Elo2 acts towards the start of the sphingolipid biosynthesis pathway, where it aids in the synthesis of very long-chain fatty acids, upstream of where ceramide is combined with PI (Fig. 5d). *ELO2* deletion has interactions with the enzymes involved in PI production and consumption - a negative genetic and physical interaction with *SAC1*, and is synthetic lethal with *STT4* [46]. These interactions are consistent with the observed perturbation in PI levels. The results indicate that the phospho-deficient mutant has an impact on the normal function of Elo2 and possibly acting in a gain-of-function fashion.

From the human phosphosites, RPS14 T133 and S139 were further characterized. For RPS14 T133, we generated a humanized-yeast strain (KO for both *RPS14A* and *RPS14B)* containing human RPS14-T133A in a plasmid. We observed that the T133A phospho-mutant exhibited cytoplasmic 20S pre-rRNA processing defect at low temperatures and rendered cells cold-sensitive (Fig. 5e and Supplementary Fig. 9), resembling the results observed for the yeast orthologue phosphosite (T119) (Supplementary Fig. 9). For RPS14 S139, we generated 3 independent homozygous human phospho-deficient cell lines using CRISPR (Methods). Mutation of RPS14 S139 phosphosite resulted in increased cellular viability after treatment with nickel sulfate (NiSO4) (Fig. 5f and Supplementary Fig. 10),

validating the nickel-resistant phenotype of its yeast orthologue (RPS14A S125A) observed in our phenotypic screen.

Discussion

We have developed here a chemical-genetic approach that can be used to systematically identify phosphosites that strongly contribute to fitness in a condition-specific manner. Conservation based studies have suggested that up to 65% of phosphosites may be non-functional [4]. While in our assay approximately 40% of the phospho-mutants displayed at least one significant phenotype, it is not straightforward to extrapolate this number to all phosphosites. The assays did not reach saturation as judged by the rarefaction curves or the lack of measured phenotypes for some phosphosites with known function. This could be due to a number of reasons - there could be additional relevant stress conditions that were not assayed; some phenotypes may only become relevant with phospho-mimetic mutations instead of alanine; and some phosphosites may work within clusters whereby one single mutation may be easily compensated by a nearby phosphosite. This would suggest that a larger fraction of phosphosites should be relevant for fitness. Nevertheless, we think this study may provide a lower bound estimate (i.e. ~40%) on the fraction of phosphosites that result in a growth phenotype when mutated. Notably, in addition to finding which phospho-mutants have the strongest phenotypes, the correlation of the growth profile of the phospho-mutants with those elicited by the knock-out strains generated a rich resource of functional annotations.

We observed a larger fraction of phenotypes for mutated tyrosines (60,6%) when compared to serine (41.8%) or threonine (41.4%). This difference could be explained, in part, by a higher impact of alanine mutations on tyrosine residues compared to replacing serine or threonine. We could show that at least for the case of VMA2 Y284A a more biochemically conservative Y284F mutation had similar growth and proteomic phenotypes but with smaller effect size. While tyrosine mutants appear to have a higher fraction of fitness phenotypes, the small absolute number of tyrosine mutants in our library and the differences in impact of alanine mutations on Y and S/T residues make it difficult to access if this difference is truly due to larger functional importance of phosphotyrosines.

It is possible that some phenotypes are induced by destabilization of the protein by mutations that would not be physiological. However, all mutated positions were selected based on mass spectrometry evidence for *in vivo* phosphorylation. Such positions should be able to accommodate a phosphate group and therefore be less likely to be destabilized by an alanine mutation. In addition, we were able to measure thermal stability changes for 8 phospho-mutants with one showing an increase and one showing a decrease that were not statistically significant. Although this is a small sample size it suggests that the mutations themselves are not often causing destabilization of the proteins. Therefore, the phenotypes measured should be, to a large extent, the result of the loss of function of that position (i.e. loss of phosphorylation). Some phenotypes may be induced by additional mutations, other than the phosphosite mutations, introduced during the generation of the library. We have identified additional mutations via whole genome sequencing and detailed analysis of these suggest they do not impact on the overall conclusions from our work

Europe PMC Funders Author Manuscripts

but may result in additional phenotypes for specific strains that have been flagged. Given the mutation rate of the yeast genome, we expect that many of the reported mutations could have been introduced in the amplification steps preceding the sequencing itself. The previously described mutation rate in *S. cerevisiae* is 3.3 +/− 0.1 x 10^{-3} per diploid genome per replication [47] . Prior to sequencing the strains were propagated for approximately 100 generations, this coupled with the ~0.25% /base error rate of Illumina sequencing [48] may account for the majority of the additional mutations observed in the sequencing carried out in this study.

Analysing different features of the phospho-mutants with phenotypes we observed that conservation was the best single predictor of functional relevance - in particular for the loss-of-function mutants - but conservation alone is not sufficient to identify most phosphosites with fitness phenotypes. This reinforces the need to develop machine learning based approaches to integrate different features [7,13,14] . Unbiased fitness measurements for a large number of phospho-mutants, as determined here, will be a crucial part of the further development of such predictors. Recent developments in CRISPR based targeted mutant libraries [49] will open the door to proteome wide interrogation of phosphosite function. While it is still not possible to perform large scale genome point mutant studies for human phosphosites there are several methods under development [50] that may reach a point where they can be employed for large scale phosphosite studies. Despite the overall poor conservation of phosphorylation between human and yeast, these studies in yeast were shown here to be useful to annotate orthologous human positions. We focused here on serine, threonine and tyrosine phosphorylation as these are readily accessible with MS approaches. However, there are other residues that can be phosphorylated including histidine and others which can in principle be studied via the same approaches. Finally, these methods have been used here to study protein phosphorylation but they can, in the future, be directed at other PTMs such as ubiquitination or acetylation.

Methods

Yeast strains and Media

The BY4741 Mat**a** haploid KO library (4889 KO's) and the Matα phospho-deficient mutant library (474 strains) were maintained on YPAD+G418 and YPAD respectively, prior to screening in 384 colony format. All strains used in this study are displayed in Supplementary Data 1 and were maintained on the appropriate selection media. Synthetic complete medium [51] plus indicated chemical or environmental stress was used for the chemical genomic screens.

Phospho-deficient mutant library construction

Single point mutations on phosphorylation sites (Serine/Threonine/Tyrosine to Alanine) were introduced using a modified homologous recombination method [52] in the strain Y8205 (*MAT*alpha *can1*Δ::*STE2*pr-Sp*HIS5 lyp1*Δ::*STE3*pr-*LEU2 LYS2 his3Δ1 leu2Δ0 MET15*+ *ura3Δ0*; a kind gift from Nevan Krogan). Before introducing the specific point mutation, the SceI endonuclease under the control of the galactose inducible promoter from plasmid pND32 [24] was inserted at the *LEU2* loci resulting in PBY5, the parental strain of our

phospho-deficient mutant library. All steps of the library construction are illustrated in Supplementary Fig. 1: **1**) The *URA3* selectable marker flanked by 18 bp recognition site (ATTACCCTGTTATCCCTA) for the SceI endonuclease was inserted after the stop codon of 125 genes and confirmed by PCR. **2**) The point mutation was introduced in a 55 bp primer (Fw1). The gene of interest was amplified in two overlapping products by PCR (Fw1/Rv2 and Fw3/Rv4). **3**) The two overlapping PCR products were simultaneously transformed into the PBY5 strain (WT Y8205 GALp-SceI). Point mutations were confirmed via sequence analysis of the point mutation containing region. **4**) The URA3 marker was removed via replica plating onto medium containing 2% galactose followed by a positive selection on 5-Fluoroorotic Acid [53] . All phospho-deficient mutants are listed in Supplementary Data 1. Phospho-mutant library can be obtained upon request.

High Throughput yeast transformation

Yeast transformations were performed as previously described [54] with some modifications. All steps were done in 96 well PCR plates and incubations were carried out in a thermocycler. Briefly, yeast cells were grown to exponential phase, harvested by centrifugation and washed with LiAc 100 mM, 1x TE pH 8.0 buffer, resulting in chemically competent yeast cells. Up to 30 uL of DNA (either PCR product or plasmid) was added to each well followed by transformation mix: 100 uL 50% PEG 3350, 15 uL 1M LiAc, 20 uL single-stranded carrier DNA from salmon sperm and 30 uL competent cells. The resulting mix was heat-shocked at 42°C for 40 min and plated onto the appropriate selection medium in individual agar plates, if selection was antibiotic resistance, an outgrowth of 4 hours at 30°C was carried out prior to selection.

Whole genome sequencing

Tagmentation adapters were annealed as previously described [55] and thawed on ice. For assembly, homemade Tn5R27S,E54K,L372P (0.5 mg/ml) was mixed with annealed adapters (70 uM stock) in 20 mM Tris pH 7.5 such that adapters were present in slight excess to Tn5 (2-3:1) and Tn5 was diluted 5.5X. The mixture was incubated at 23°C for 30-60 min at 300 rpm. For use in tagmentation, we further diluted the complex in dilution buffer (10 mM Tris pH 7.5, 150 mM NaCl) to a final dilution of Tn5 of 1:10.

For whole-genome sequencing, cells were inoculated into 100 ul YPAD from glycerol stocks and grown overnight at 800 rpm at 30°C. Saturated cultures were pelleted at 1000g for 3 minutes and pellets resuspended in 25 ul 300U/ml Zymolyase 100-T (AMSBIO) solution. The solution was incubated at 37°C for 30 minutes followed by 10 minutes at 95°C to inactivate Zymolyase. 1.25 ul of the solution was mixed with 1.25 ul 1:10 diluted, adapter-loaded Tn5R27S,E54K,L372P and 2.5 ul tagmentation buffer (8mM MgCl2, 20% DMF (v/v) and 16mM Tris-HCl (final concentrations), adjusted to pH 7.6 using acetic acid). Tagmentation buffer without DMF was prepared as a 2X solution and DMF was added fresh immediately before tagmentation for every experiment. Samples were incubated on a pre-heated thermocycler block for exactly 3 minutes at 55°C and 1.25 ul 0.2% SDS added immediately for neutralization. After a 5 minute incubation at room temperature, 1.75 ul 0.4mg/ml Proteinase K (diluted from 20mg/ml stock) was added and the reaction incubated for 30 min at 50°C followed by 15 min at 65°C. Samples were cleaned with 1.8 volumes

of AMPure XP beads (Beckman Coulter) and eluted in 6.25 ul nuclease-free water. For library amplification 1.25 ul each of indexed P5 and P7 primer (10uM), 6.75 ul of 2X KAPA HiFi Ready mix (KAPA Biosystems) and 0.75 ul of DMSO was added to the sample. A 12-cycle enrichment PCR was performed, samples were again purified using 1.8X volumes AMPure XP beads and eluted in 10 ul nuclease-free water. We quantified library yield using the Qubit high-sensitivity DNA kit and evaluated library quality on an Agilent Bioanalyzer (high-sensitivity DNA assay). All samples were pooled equimolar and sequenced on a Illumina NextSeq platform (150 PE, high output mode).

Raw reads were mapped against the S. cerevisiae S288c reference genome (Bioproject PRJNA128), using snippy v4.6.0 [https://github.com/tseemann/snippy] with default parameters. Reads coverage was computed from the aligned reads using mosdepth v0.2.9 [56] , with option "--fast-mode" and using a window size of 500 bases. Samples where less than 50% of the total genome length had coverage >= 10X where excluded. Copy Number Variants (CNVs) where called from the aligned reads using CNVkit v0.9.6 [57] using the aligned reads from the wild-type strain (PBY5); regions with absolute log2 of the mean coverage depth above 0.8 where considered CNVs. Shorter variants were extracted from the snippy output, discarding common variants (frequency across all samples > 10%), as those are likely to have been present in the parental strain. Synonymous variants were ignored as they are unlikely to have a significant effect. Raw reads have been uploaded to ENA under project number PRJEB38714.

Chemical genomics screen

Growth of the S288C (BY4741) KO and phospho-deficient strain collection was evaluated on concentrations of chemical and environmental stress conditions (Supplementary Data 2) that inhibit the growth of BY4741 by approximately 40%. The libraries were maintained and pinned with a Singer RoTor in 1536 colony format. Synthetic complete media (Kaiser et al 1994) was used with or without the stress condition, incubated at 30°C (unless temperature or medium was a stress, eg Nati medium (CSM35) was prepared as described previously [58]) for 48h or 72h. Plates were imaged using a SPIMAGER (S&P Robotics equipped with a Canon Rebel T3i digital camera.

The screen was carried out in two main batches, consisting of 70 conditions in batch 1 and 29 conditions in batch 2, all conditions and concentrations are listed in Supplementary Data 2. In order to account for variation of the method 13 conditions were tested in both batches (anaerobic growth, amphoteracin B, nystatin, DMSO, 2,4-D, glycerol, maltose, hepes buffered medium, caffeine, 6-AU, paraquat, 39°C and sorbitol).

Chemical genomics data analysis

Raw plate images were cropped using ImageMagick to exclude the plate plastic borders; raw colony sizes were extracted from the cropped images using gitter v1.1.1 [59] , using the "autorotate" and "noise removal" features on. Poor quality plates were flagged when no colony size could be reported for more than 5% of colonies, which indicates poor overall quality, or when no colony size could be reported for more than 90% of a whole row or column, which indicates a potential grid misalignment. Known empty spots in

each plate where used to flag incorrectly labelled plates. Conditions with less than three replicates across the two batches of the experiment were excluded from further processing. Raw colony sizes for the remaining conditions were used as an input for the S-score [22], with default parameters except the minimum colony size which was set to 5 pixels. The algorithm computes an S-score, which indicates whether the growth of each mutant/KO is deviating from the expected growth in each condition. The raw S-scores were further quantile normalized in each condition. Significant positive and negative phenotypes were highlighted by transforming the S-scores in Z-scores, given that the S-scores in each condition follow a normal distribution. P-values were derived using the survival function of the normal distribution and corrected using an FDR of 5% (false discovery rate). The whole dataset of mutant-condition scores is available as Supplementary Data 3.

The reproducibility of the chemical genomics screen was assessed by dividing the raw pictures according to the batch of origin; the EMAP algorithm (https://sourceforge.net/projects/emap-toolbox/) was then used to compute a set of S-scores for each batch. For the 13 conditions that were tested in both batches the S-score correlation was computed for the phospho-mutants and the KOs separately. We refer to this analysis as both technical and biological replicates because the inoculates are derived from the same source plate but at very different times (Fig. 1d).

To associate phosphosites with specific functions we correlated the growth profile, defined as the S-scores across conditions, of each phospho-mutant with the profile of all gene deletion strains. We defined the significant phospho-mutant to gene correlations (positive correlation with 5% FDR cutoff) and performed an enrichment test for the top most enriched GO terms of these associated genes using clusterProfiler package [60] in R. We excluded GO terms that were too unspecific (i.e. annotating more than 500 genes or linked to 10 or more of phospho-mutants). To reduce redundancy in the annotations we also measured GO-term similarity by calculating Jaccard Similarity Coefficient between all the GO gene sets. For each phospho-mutant we then selected the top 2 associated GO terms more specifically associated with each mutant from different GO term similarity clusters. Specificity of a GO term was defined by the number of phosphosites associated with it with broader terms presumed to be associated with more phosphosites.

Screen validation by serial dilution assay

Yeast cells were grown in 96 well microtiter plates in YPAD and incubated for 24 h until saturation. The strains were then serially diluted four times at a 1/20 dilution in fresh 96 well plates, the dilutions were performed using a liquid handler (Beckman Coulter BioMek FX^P). The diluted cells were spotted onto stress condition agar plates using the VP405 96 (V&P Scientific) format manual pinning tool. The agar plates were incubated for 48 and 72 hours at 30°C, unless other temperature is indicated. Agar plates were imaged using a SPIMAGER (S&P Robotics equipped with a Canon Rebel T3i digital camera). The resulting spot assays were scored 0-5 by eye-based on colony sizes compared to controls on each plate.

Two-dimensional thermal proteome profiling (2D-TPP)

Thermal proteome profiling [34] was performed as previously described [14] . Briefly, yeast cells were grown overnight at 30°C in YPAD and diluted to OD 0.1 in 50 ml fresh YPAD media. Cultures were collected by centrifugation at 4000x g for 5 min when OD_{600} reached ~0.7, washed with PBS, pelleted by centrifugation and resuspended in a volume of PBS that led to an equivalent of OD_{660} of 125. 20ul was aliquoted to 10 wells of a PCR plate and the plate was centrifuged at 4.000g for 5 min and was subjected to a temperature gradient (37, 40.4, 44, 46.9, 49.8, 52.9, 55.5, 58.6, 62, 66.3 °C) for 3 min in a thermocycler (Agilent SureCycler 8800) followed by 3 min at room temperature. Cells were lysed in 30 ul of cold lysis buffer (final concentration: Zymolyase (Amsbio) 0.5mg/ml, 1x protease inhibitor (Roche), 1x phosphatase inhibitor PhosSTOP (Roche), 250U/ml benzonase and 1mM MgCl2 in PBS) for 30 min shaking at 30°C, followed by five freeze-thaw cycles (freezing in liquid nitrogen, followed by 30s at 25°C in a thermocycler and vortexing. The remaining soluble protein at each temperature was collected by filtration as previously described [61] . Proteins were digested with trypsin and LysC according to a modified SP3 protocol [62] , as previously described [61] . Peptides were labelled with TMT11plex (Thermo Fisher Scientific), with the label used for each experiment as indicated in Supplementary Data 16. Mass spectrometry-based proteomics was performed using an UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) equipped with a trapping cartridge (Precolumn; C18 PepMap 100, 5 μm, 300 μm i.d. × 5 mm, 100 Å) and an analytical column (Waters nanoEase HSS C18 T3, 75 μm × 25 cm, 1.8 μm, 100 Å), coupled to to a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific), as previously described [61] . We then calculated abundance scores and thermal stability scores for every protein in every mutant as previously described [61] . Briefly, at each temperature, we calculated the ratio of the signal sum intensity of each mutant to the respective control (gene control with 18 bp scar at 3'UTR). The abundance score of each protein in each mutant was calculated as the average log2 fold change at the two lowest temperatures, and the thermal stability score of each protein in each mutant was then calculated by subtracting the abundance score from the log2 fold changes of all temperatures, and summing the resulting log2 fold changes. limma analysis was then used to assess significance [63] .

Lipid Extraction

Yeast cultures were grown in 5 mL synthetic complete medium to an OD_{600} of 1.0, harvested by centrifugation and washed with 1 mL dH_20, transferred to a 1.5 mL tube and the cells were pelleted by centrifugation. The following extraction process was carried out at 4 °C. Sterols were extracted from the cellular pellet by the addition of 150 μL acid washed glass beads along with 900 μL chloroform: methanol (1:2), vortexed vigorously for 10 mins and then incubated at 4 °C for 15 hours with shaking at 1100 rpm. After incubation a further 300 μL of chloroform and 300 μL of dH_20 was added, the mixture was vortexed for 30 sec, rested for 2 mins then centrifuged for 5 mins at 9,000 x g. The solvent fraction was removed and transferred into a new 1.5 mL tube. A second extraction was performed by the addition of 600 μL chloroform, vortexing vigorously for 30 sec followed by a final centrifugation for 5 mins at 9000 x g, the solvent fraction was removed and added to the previous solvent fraction. The combined solvent fractions were then dried over nitrogen and stored at −80 °C.

Lipidomics LC-MS/MS Analysis

Lipid extracts were separated on a Kinetex C18 2.1 x 100 mm, 2.6 μm column (Phenomonex, Aschaffenburg, De). Separation was achieved by gradient elution on a binary solvent Vanquish UHPLC (Thermo Fisher Scientific, Bremen, DE). Mobile Phase A consisted of ACN: H2O (60:40) while mobile phase B consisted of IPA: ACN (90:10). For positive ionization, the mobile phases were modified with 10 mM ammonium formate and 0.1% formic acid while for the negative ionization mode, the mobile phases were modified with 5 mM ammonium acetate and 0.1% acetic acid. A flow rate of 260 μL/min was used for the separation and the column and sample tray were held constant at 30°C and 4 °C respectively.

Lipidomics Mass Spectrometry Instrumentation

MS analysis was performed on a Q-Exactive plus Mass Spectrometer (Thermo Fisher Scientific, Bremen, DE) equipped with a heated electrospray ionization probe. In both the positive and negative ionization modes, the S-Lens RF level was set to 65, the capillary temperature was set to 320 °C, the sheath gas flow was set to 30 units and the auxiliary gas was set to 5 units. The spray voltage was set to 3.5 kV in the negative ionization mode and 4.5 kV in the positive ionization mode. In both modes, full scan mass spectra (Scan Range m/z 100-1500, R=35K) were acquired along with data-dependent (DDA) MS/MS spectra of the five most abundant ions. DDA MS/MS spectra were acquired using normalized collision energies of 30, 40, and 50 units (R= 17.5K and an isolation width = 1 m/z). The instrument was controlled using Xcalibur version 4.0.

Data analysis, lipid annotation and confirmation

Progenesis Q1, version 2.0 (Non-Linear Dynamics, A Waters Company, Newcastle upon Tyne, UK) was used for peak picking and for chromatographic alignment of all samples (with a pooled sample acquired in the middle of the LC-MS sequence used as a reference). Lipids were initially annotated from Progenesis Metascope and LipidBlast databases. Putative identification of lipids was done for ions that had MS/MS data. A pooled sample was used for the lipid identification of statistically significant lipids. An inclusion list was created to acquire MS/MS spectra of these lipids and spectra were acquired as described above. Full scan spectra of the pooled sample were also acquired by polarity switching between the positive and the negative ionization modes. Raw MS/MS spectra were log transformed and distribution of spectra were compared to remove any outlier samples. Differential analysis of the spectra were performed using Limma [63] for each phospho-mutant or KO with their respective controls.

Humanized yeast strains and plasmids

All recombinant DNA techniques were performed according to established procedures using *Escherichia coli* XL1 blue cells for cloning and plasmid propagation. Mutations in *hRPS14* were generated using the QuikChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA). All cloned DNA fragments and mutagenized plasmids were verified by sequencing. The RPS14 T133A phospho-mutant was generated by transforming JWY6851 strain [64] with plasmids listed in Supplementary Data 17. *pGal-*

yRPS14 was shuffled out on FOA plates. The plasmids used in this study are listed in Supplementary Data 17 and generated humanized yeast strains are listed in Supplementary Data 18.

Fluorescence *in situ* hybridization and microscopy

Cells were grown to mid-logarithmic phase and fixed with 4% formaldehyde in 0.1 M potassium phosphate buffer. Cells were then converted to spheroplasts using 0.1 M potassium phosphate buffer containing 1.2 M sorbitol and 500 µg of zymolyase. Spheroplasts were washed in 2× SSC buffer and incubated overnight at 37 °C with Cy3-labeled oligonucleotide probe (5′-Cy3-ATG CTC TTG CCA AAA CAA AAA AAT CCA TTT TCA AAA TTA TTA AAT TTC TT-3′) that is complementary to the 5′ portion of ITS1 (Faza et al., 2012). DNA was stained with 0.5 µg/ml DAPI. Cells were visualized using THUNDER Imager 3D Assay (Leica, Germany) equipped with a HCX PL-APO Fluotar 100×/1.44 NA oil immersion objective (Leica, Germany). Images were acquired with a fitted digital sCMOS camera and LAS X premium software (Leica, Germany).

Human RPS14 S139A cell lines generation

HEK293T cells were cultured in growth media composed of DMEM supplemented with 10 % FCS, 1% GlutaMax, 1% penicillin/streptomycin and 1 mM sodium pyruvate (all from Gibco). The HEK293T RPS14 S139A mutant lines were generated following the editing strategy previously described [65], with modifications. Briefly, the CRISPR-Cas9 ribonucleoprotein complex was formed by mixing 20 pmol of the Cas9 protein (produced in-house by the EMBL Protein Expression and Purification Facility) with 30 pmol of the guide RNA (sgRNA) and incubated at room temperature for 20 minutes. The sgRNA was obtained by mixing equimolar amount of the target-specific crRNA (Integrated DNA Technologies; standard desalting) with an ATTO-550 conjugated tracrRNA (Integrated DNA Technologies). Following incubation at 95°C for 5 minutes, the mixture was cooled down at room temperature to favour annealing. Prior to electroporation, 30 pmol of the single-stranded oligonucleotide (ssODN) repair template (82 bp, Integrated DNA Technologies Ultramers) was added to the mixture. One hundred thousand HEK293T cells were electroporated using the Neon® Transfection System 10 µl Kit (Thermo Fisher Scientific; settings: 1150v/20ms/2 pulses). After 48h, fluorescence-activated cell sorting was performed to select ATTO-550 positive cells. Clonal cell lines were established and genotyped. Sanger sequencing confirmed the correct insertion of the homozygous S139A mutation. For the successfully edited clones, the integrity of the top-three predicted off-target genes (according to [66]) was confirmed by Sanger sequencing.

HEK293T Cell viability assay

HEK 293T WT and RPS14 S139A cell lines were settled at a final volume of 0.2ml in 96-well flat-bottom plates (10.000 cells/well), grown for 24h and treated with 0mM (untreated), 1mM and 2mM of nickel sulphate ($NiSO_4$) for 24 hours. Fluorimetric resazurin reduction method (CellTiter Blue, Promega) was used to measure cell viability. Fluorescence was measured at 560/590nm 4 hours after addition of resazurin using Tecan Infinite M1000 plate reader. Fluorescence was determined for 6 technical replicates per condition and normalized

to the untreated control. The cell viability assay was performed for 3 independently generated phospho-mutant cell lines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Nevan Krogan and J Woolford Jr. for generously providing strains. Michael Knop. for generously providing plasmids. Magdalena Gierlach and Nadja Nepke from the EMBL Lab kitchen for their help making media and pouring screening plates. Natalia Gabrielli for providing nitrogen stress media. Rose Gathungu and Prasad Phapale from the EMBL Metabolomics Core Facility. Kevin Roy and Judit Villen for critical reading of the manuscript. This study has been funded by EMBL core funding and a Starting Grant Award from the European Research Council (ERC-2014-STG 638884 PhosFunc PB). L.M Steinmetz, Vonesch S.C and S.T.Chelsea were supported by the European Research Council Advanced Investigator Grant under the European Union's Horizon 2020 research and innovation programme (AdG-742804-SystGeneEdit). Vonesch S.C was supported by an Advanced Postdoc Mobility Fellowship from the Swiss National Science Foundation (grant number P300PA_177909). V.G. Panse is supported by grants from the Swiss National Science Foundation, NCCR RNA & Disease, Novartis Foundation and Olga Mayenfisch Stiftung. M. O.-O.was supported by Boehringer Ingelheim Fonds PhD fellowship. K.-M. N. was supported by the DFG fund (SPP 1738) and Fondazione Cariplo (2018-0525 DEPRECAD).

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD017929. Yeast mutant strain growth measurements are provided in Supplementary Data.

Code availability

Functional enrichment in the phosphorylation network was performed using SAFE (https://www.bioconductor.org/packages/release/bioc/html/safe.html). Image processing and S-score calculations were performed using EMAP (https://www.bioconductor.org/packages/release/bioc/html/safe.html). GO enrichment analysis to assign annotations to yeast phosphosites was performed using ClusterProfiler package (https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html).

References

1. Olsen JV, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell. 2006; 127: 635–648. [PubMed: 17081983]
2. Sadowski I, et al. The PhosphoGRID Saccharomyces cerevisiae protein phosphorylation site database: version 2.0 update. Database. 2013; 2013 bat026 [PubMed: 23674503]
3. Hornbeck PV, et al. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 2015; 43: D512–20. [PubMed: 25514926]
4. Landry CR, Levy ED, Michnick SW. Weak functional constraints on phosphoproteomes. Trends Genet. 2009; 25: 193–197. [PubMed: 19349092]
5. Lienhard GE. Non-functional phosphorylations? Trends Biochem Sci. 2008; 33: 351–352. [PubMed: 18603430]
6. Kanshin E, Bergeron-Sandoval L-P, Isik SS, Thibault P, Michnick SW. A cell-signaling network temporally resolves specific versus promiscuous phosphorylation. Cell Rep. 2015; 10: 1202–1214. [PubMed: 25704821]

7. Beltrao P, et al. Systematic functional prioritization of protein posttranslational modifications. Cell. 2012; 150: 413–425. [PubMed: 22817900]
8. Beltrao P, et al. Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. PLoS Biol. 2009; 7 e1000134 [PubMed: 19547744]
9. Studer RA, et al. Evolution of protein phosphorylation across 18 fungal species. Science. 2016; 354: 229–232. [PubMed: 27738172]
10. Freschi L, Osseni M, Landry CR. Functional divergence and evolutionary turnover in mammalian phosphoproteomes. PLoS Genet. 2014; 10 e1004062 [PubMed: 24465218]
11. Gray VE, Kumar S. Rampant purifying selection conserves positions with posttranslational modifications in human proteins. Mol Biol Evol. 2011; 28: 1565–1568. [PubMed: 21273632]
12. Tunc-Ozdemir M, et al. Predicted Functional Implications of Phosphorylation of Regulator of G Protein Signaling Protein in Plants. Front Plant Sci. 2017; 8 1456 [PubMed: 28890722]
13. Dewhurst HM, Choudhury S, Torres MP. Structural Analysis of PTM Hotspots (SAPH-ire)--A Quantitative Informatics Method Enabling the Discovery of Novel Regulatory Elements in Protein Families. Mol Cell Proteomics. 2015; 14: 2285–2297. [PubMed: 26070665]
14. Ochoa D, et al. The functional landscape of the human phosphoproteome. Nat Biotechnol. 2020; 38: 365–373. [PubMed: 31819260]
15. Nishi H, Hashimoto K, Panchenko AR. Phosphorylation in protein-protein binding: effect on stability and function. Structure. 2011; 19: 1807–1815. [PubMed: 22153503]
16. Šoštarić N, et al. Effects of Acetylation and Phosphorylation on Subunit Interactions in Three Large Eukaryotic Complexes. Mol Cell Proteomics. 2018; 17: 2387–2401. [PubMed: 30181345]
17. Strumillo MJ, et al. Conserved phosphorylation hotspots in eukaryotic protein domain families. Nat Commun. 2019; 10 1977 [PubMed: 31036831]
18. Nakic ZR, Seisenbacher G, Posas F, Sauer U. Untargeted metabolomics unravels functionalities of phosphorylation sites in Saccharomyces cerevisiae. BMC Systems Biology. 2016; 10
19. Oliveira AP, et al. Dynamic phosphoproteomics reveals TORC1-dependent regulation of yeast nucleotide and amino acid biosynthesis. Sci Signal. 2015; 8 rs4 [PubMed: 25921291]
20. Tong, aH; , et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 2001; 294: 2364–2368. [PubMed: 11743205]
21. Costanzo M, et al. The genetic landscape of a cell. Science. 2010; 327: 425–431. [PubMed: 20093466]
22. Collins SR, Schuldiner M, Krogan NJ, Weissman JS. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol. 2006; 7 R63 [PubMed: 16859555]
23. Braberg H, et al. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. Cell. 2013; 154: 775–788. [PubMed: 23932120]
24. Khmelinskii A, Meurer M, Duishoev N, Delhomme N, Knop M. Seamless gene tagging by endonuclease-driven homologous recombination. PLoS One. 2011; 6 e23794 [PubMed: 21915245]
25. Winzeler EA, et al. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science. 1999; 285: 901–906. [PubMed: 10436161]
26. Rayner TF, Gray JV, Thorner JW. Direct and novel regulation of cAMP-dependent protein kinase by Mck1p, a yeast glycogen synthase kinase-3. J Biol Chem. 2002; 277: 16814–16822. [PubMed: 11877433]
27. Beltrao P, Cagney G, Krogan NJ. Quantitative genetic interactions reveal biological modularity. Cell. 2010; 141: 739–745. [PubMed: 20510918]
28. Nichols RJ, et al. Phenotypic landscape of a bacterial cell. Cell. 2011; 144: 143–156. [PubMed: 21185072]
29. Martín H, Rodríguez-Pachón JM, Ruiz C, Nombela C, Molina M. Regulatory mechanisms for modulation of signaling through the cell integrity Slt2-mediated pathway in Saccharomyces cerevisiae. J Biol Chem. 2000; 275: 1511–1519. [PubMed: 10625705]
30. Baryshnikova A. Systematic Functional Annotation and Visualization of Biological Networks. Cell Syst. 2016; 2: 412–421. [PubMed: 27237738]
31. Giaever G, et al. Functional profiling of the Saccharomyces cerevisiae genome. Nature. 2002; 418: 387–391. [PubMed: 12140549]

32. Becher I, et al. Thermal profiling reveals phenylalanine hydroxylase as an off-target of panobinostat. Nat Chem Biol. 2016; 12: 908–910. [PubMed: 27669419]
33. Savitski MM, et al. Multiplexed Proteome Dynamics Profiling Reveals Mechanisms Controlling Protein Homeostasis. Cell. 2018; 173: 260–274. e25 [PubMed: 29551266]
34. Savitski MM, et al. Tracking cancer drugs in living cells by thermal profiling of the proteome. Science. 2014; 346 1255784 [PubMed: 25278616]
35. Kane PM. The where, when, and how of organelle acidification by the yeast vacuolar H+-ATPase. Microbiol Mol Biol Rev. 2006; 70: 177–191. [PubMed: 16524922]
36. Vasanthakumar T, et al. Structural comparison of the vacuolar and Golgi V-ATPases from Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2019; 116: 7272–7277. [PubMed: 30910982]
37. Zhang Y-Q, et al. Requirement for ergosterol in V-ATPase function underlies antifungal activity of azole drugs. PLoS Pathog. 2010; 6 e1000939 [PubMed: 20532216]
38. Duncan AL, Robinson AJ, Walker JE. Cardiolipin binds selectively but transiently to conserved lysine residues in the rotor of metazoan ATP synthases. Proc Natl Acad Sci U S A. 2016; 113: 8687–8692. [PubMed: 27382158]
39. Zhou M, et al. Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. Science. 2011; 334: 380–385. [PubMed: 22021858]
40. Ariño J, Velázquez D, Casamayor A. Ser/Thr protein phosphatases in fungi: structure, regulation and function. Microb Cell Fact. 2019; 6: 217–256.
41. Breitkreutz A, et al. A global protein kinase and phosphatase interaction network in yeast. Science. 2010; 328: 1043–1046. [PubMed: 20489023]
42. Pereira C, Pereira AT, Osório H, Moradas-Ferreira P, Costa V. Sit4p-mediated dephosphorylation of Atp2p regulates ATP synthase activity and mitochondrial function. Biochim Biophys Acta Bioenerg. 2018; 1859: 591–601. [PubMed: 29719209]
43. Ochoa D, et al. An atlas of human kinase regulation. Mol Syst Biol. 2016; 12: 888. [PubMed: 27909043]
44. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003; 31: 3812–3814. [PubMed: 12824425]
45. Zimmermann C, et al. TORC1 inhibits GSK3-mediated Elo2 phosphorylation to regulate very long chain fatty acid synthesis and autophagy. Cell Rep. 2013; 5: 1036–1046. [PubMed: 24239358]
46. Cherry JM, et al. Genetic and physical maps of Saccharomyces cerevisiae. Nature. 1997; 387: 67–73. [PubMed: 9169866]
47. Zhu YO, Siegal ML, Hall DW, Petrov DA. Precise estimates of mutation rate and spectrum in yeast. Proc Natl Acad Sci U S A. 2014; 111: E2310–8. [PubMed: 24847077]
48. Pfeiffer F, et al. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. Sci Rep. 2018; 8 10950 [PubMed: 30026539]
49. Roy KR, et al. Multiplexed precision genome editing with trackable genomic barcodes in yeast. Nat Biotechnol. 2018; 36: 512–520. [PubMed: 29734294]
50. Zafra MP, et al. Optimized base editors enable efficient editing in cells, organoids and mice. Nat Biotechnol. 2018; 36: 888–893. [PubMed: 29969439]
51. Kaiser, C, Michaelis, S, Mitchell, A. Cold Spring Harbor Laboratory. Methods in yeast genetics: a Cold Spring Harbor Laboratory course manual. Vol. 177. Plainview, NY: Cold Spring Harbor Laboratory Press; 1994. xiv
52. Toulmay A, Schneiter R. A two-step method for the introduction of single or multiple defined point mutations into the genome of Saccharomyces cerevisiae. Yeast. 2006; 23: 825–831. [PubMed: 16921548]
53. Boeke JD, Trueheart J, Natsoulis G, Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 1987; 154: 164–175. [PubMed: 3323810]
54. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007; 2: 31–34. [PubMed: 17401334]
55. Hennig BP, et al. Large-Scale Low-Cost NGS Library Preparation Using a Robust Tn5 Purification and Tagmentation Protocol. G3. 2018; 8: 79–89. [PubMed: 29118030]

56. Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics. 2018; 34: 867–868. [PubMed: 29096012]

57. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016; 12 e1004873 [PubMed: 27100738]

58. Ponomarova O, et al. Yeast Creates a Niche for Symbiotic Lactic Acid Bacteria through Nitrogen Overflow. Cell Syst. 2017; 5: 345–357. e6 [PubMed: 28964698]

59. Wagih O, Parts L. gitter: a robust and accurate method for quantification of colony sizes from plate images. G3. 2014; 4: 547–552. [PubMed: 24474170]

60. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16: 284–287. [PubMed: 22455463]

61. Mateus A, et al. Thermal proteome profiling in bacteria: probing protein state in vivo. Mol Syst Biol. 2018; 14 e8242 [PubMed: 29980614]

62. Hughes CS, et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nat Protoc. 2019; 14: 68–85. [PubMed: 30464214]

63. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43: e47. [PubMed: 25605792]

64. Jakovljevic J, et al. The carboxy-terminal extension of yeast ribosomal protein S14 is necessary for maturation of 43S preribosomes. Mol Cell. 2004; 14: 331–342. [PubMed: 15125836]

65. Gehre M, Buccitelli C, Diaz N, Korbel J, Noh K-M. Efficient strategies to detect genome editing and integrity in CRISPR-Cas9 engineered ESCs. bioRxiv. 2019; 635151 doi: 10.1101/635151

66. Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013; 31: 827–832. [PubMed: 23873081]

Europe PMC Funders Author Manuscripts

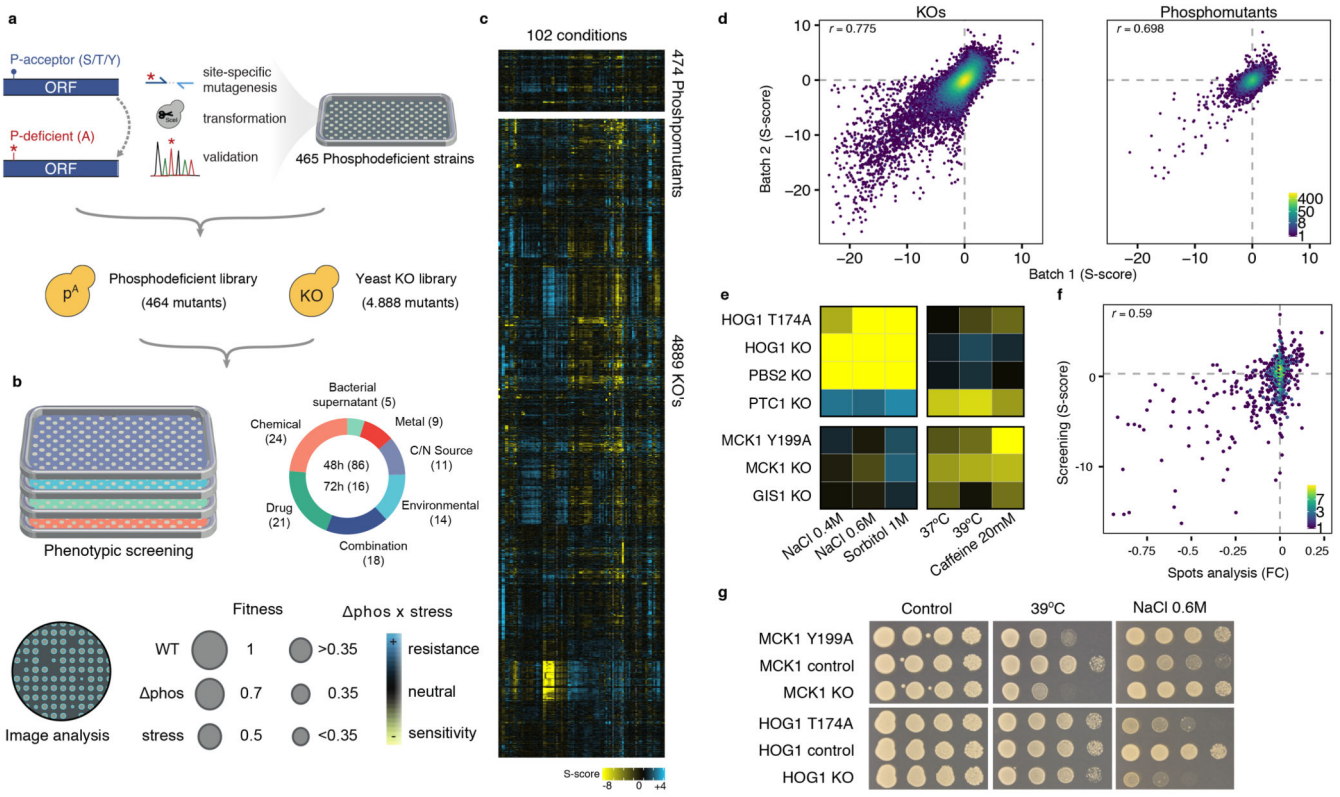

Figure 1. Phospho-mutant library construction, chemical genomics screen and quality control.

a, Phospho-mutant library construction (Supplementary Fig.1 and Supplementary Data 1) and arraying of mutant libraries. **b**, Chemical genomics screen against 102 conditions (Supplementary Data 2), summary of image analysis and phenotype calling (Methods). **c**, Overall heatmap of 474 phosphosites and 4889 KO's in 102 conditions (Supplementary Data 3). **d**, Correlation of S-scores for biological replicates of KO and phospho-mutants screened in 13 conditions (r=Pearson correlation coefficient) (Supplementary Data 3). **e**, Heatmap of known regulatory phospho-mutants HOG1-T174A and MCK1-Y199A, which phenocopy their corresponding KO. **f**, Correlation between screen S-scores and fold changes from spot analysis for 762 pairs of mutant-condition phenotypes (Supplementary Data 4). **g**, Example of serial dilution spot assays, recapitulating the S-scores of *MCK1* and *HOG1* Controls, KOs and phospho-mutants in three conditions: untreated, heat and osmotic stress.

Europe PMC Funders Author Manuscripts

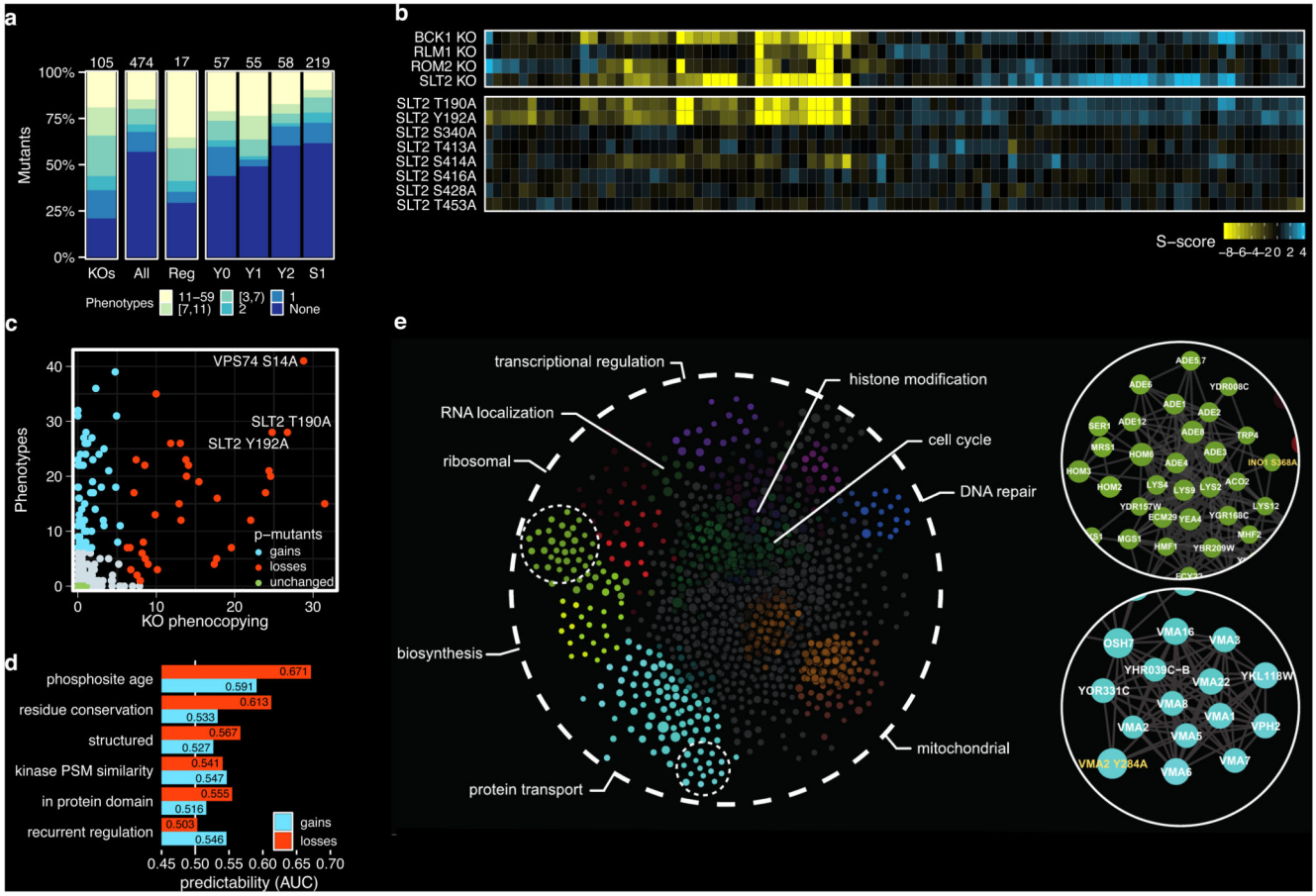

Figure 2. Chemical genomics screen functional analysis.

a, The percentage of mutants (KO or phospho-deficient) which have growth phenotypes (Supplementary Data 3). KO: 105 non-essential gene KO's for which we constructed phospho-mutants; All: 474 phospho-mutants in our library; Reg: 17 phospho-mutants previously described to be involved in protein function; Y0 to S1: phospho-mutants grouped based on their evolutionary age from more to less conserved: Y0= 731 Million years (My), Y1= 434 My, Y2 from 61 to 182 My and S1 <61 My. **b**, Growth fingerprints of *SLT2 KO* and eight phospho-mutants created in *SLT2*. Top: Heatmap depicting the phenotypic fingerprint of *SLT2 KO* which clusters with other related genes KO's (*BCK1*, *RML1* and *ROM2)*. Bottom: Unclustered phenotypic fingerprints of eight *SLT2* phospho-mutants. **c**, Phospho-mutants that phenocopy their corresponding gene KO are classified as loss of function (red) and phosphomutants that do not phenocopy their gene KO as gain of function (blue). KO phenocopying: log(*p*-value) of the correlation of the growth profiles between the phospho-mutant and corresponding gene-deletion. **d**, Area under the ROC curve (AUC) for the prediction of loss of function (red) or gain of function (blue) classification, relative to unchanged, based on the following categories: phosphosite age (evolutionary conservation); residue conservation; position on protein structure; kinase recognition (PSM) site similarity; recurrent regulation or conservation of protein domain. (Supplementary Data 6). **e**, SAFE analysis representation examining the functional organisation of phospho/gene KO network into functional gene ontology clusters. Highlighted in green and blue, are two examples

Europe PMC Funders Author Manuscripts

which show significant enrichment in the categories of ribosomal biosynthesis (INO1-S368) and protein transport (VMA2-Y284).

Europe PMC Funders Author Manuscripts

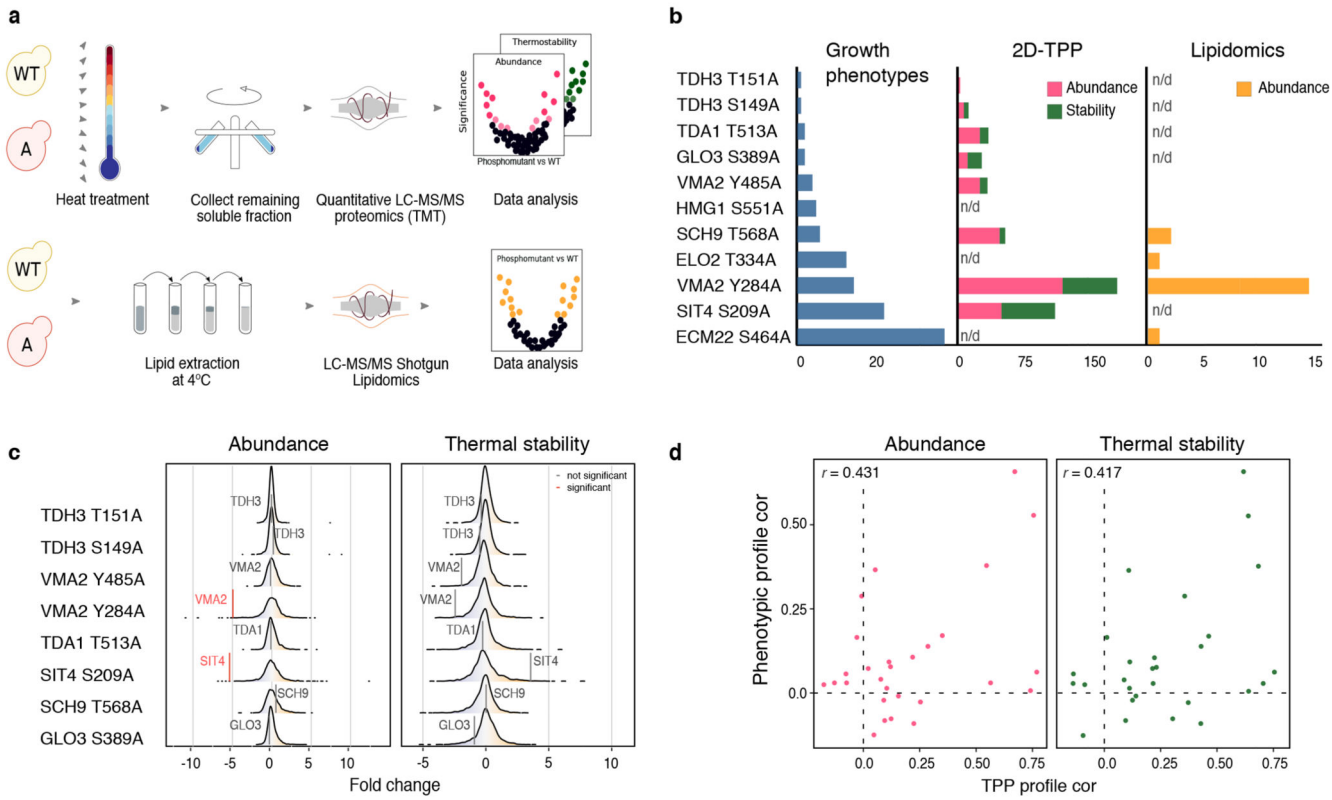

Figure 3. Molecular characterization of phospho-deficient mutants.

a, Graphical representation of the sample preparation, LC-MS/MS, and data analysis for 2D-TPP (top) and shotgun lipidomics (bottom). For both experiments, cells were growing in exponential phase in YPAD. **b**, Summary of the number of phenotypes observed in the genetic screen (blue), significant hits in 2D-TPP (green and pink) and lipidomics (orange) per phospho-mutant. **c**, Changes in protein abundance (left) and protein thermal stability (right) resulting from the phospho-mutation in their respective proteins. **d**, Pairwise correlations of the phenotypic profiles of the phospho-mutants correlated with pairwise correlations in protein abundance (left) or thermal stability (right) generated by 2D-TPP.

Europe PMC Funders Author Manuscripts

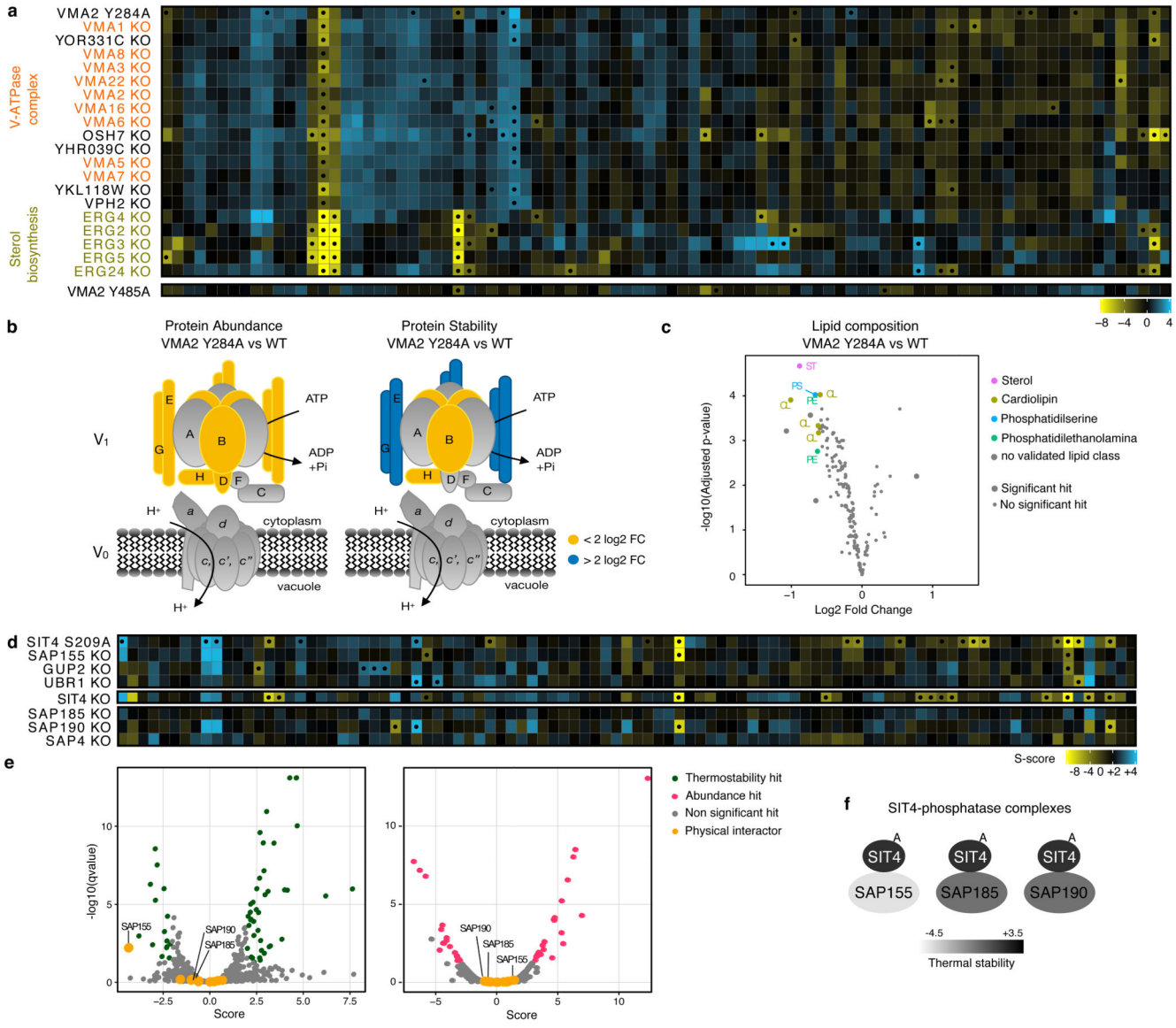

Figure 4. Molecular characterisation of Vma2 and Sit4 phospho-deficient mutants.

a, Top, growth profile of *VMA2-Y284A* which phenocopies *VMA2 KO* and clusters with members of the V-ATPase complex and members of the ergosterol biosynthesis pathway. Bottom, heat map (unclustered) of *VMA2-Y485A* which does not phenocopy *VMA2 KO*. **b**, Graphical representation of the cytoplasmic V_1 and membrane-bound V_0 subunits, colours represent complex members that significantly increase (blue) or decrease (yellow) in either protein abundance or protein thermal stability measured by 2D-TPP (Supplementary Data 9). **c**, Lipid composition of *VMA2-Y284A* measured by shotgun lipidomics, values are log2 fold change compared to a control strain (Supplementary Data 10, 11 and 12). Validated lipids (Supplementary Data 12) in the following classes showed significant decreases: sterol (ST), cardiolipin (CL), phosphatidylserine (PS), phosphatidylethanolamine (PE). **d**, Top, growth profile of *SIT4-S209A* in the same cluster with the deletion of *SAP155 KO*. Bottom, unclustered growth profiles of *SIT4 KO* and remaining phosphatase complex

Europe PMC Funders Author Manuscripts

members*: SAP4, SAP190* and *SAP185 KOs*. **e**, Volcano plots of *SIT-S209A* 2D-TPP data (Supplementary Data 9). Score is log2 fold change. Phosphatase complex members are labelled and physical interaction partners are shown in yellow. Significant changes in protein thermal stability (green) and significant changes in protein abundance (pink) are shown. **f**, Graphical summary of the phosphatase complexes that Sit4 forms with SAP155 (destabilized), SAP185 (non-destabilized) and SAP190 (non-destabilized), grayscale colors highlight changes in protein thermal stability in the Sit4 S209A mutant when compared to WT cells.

Europe PMC Funders Author Manuscripts

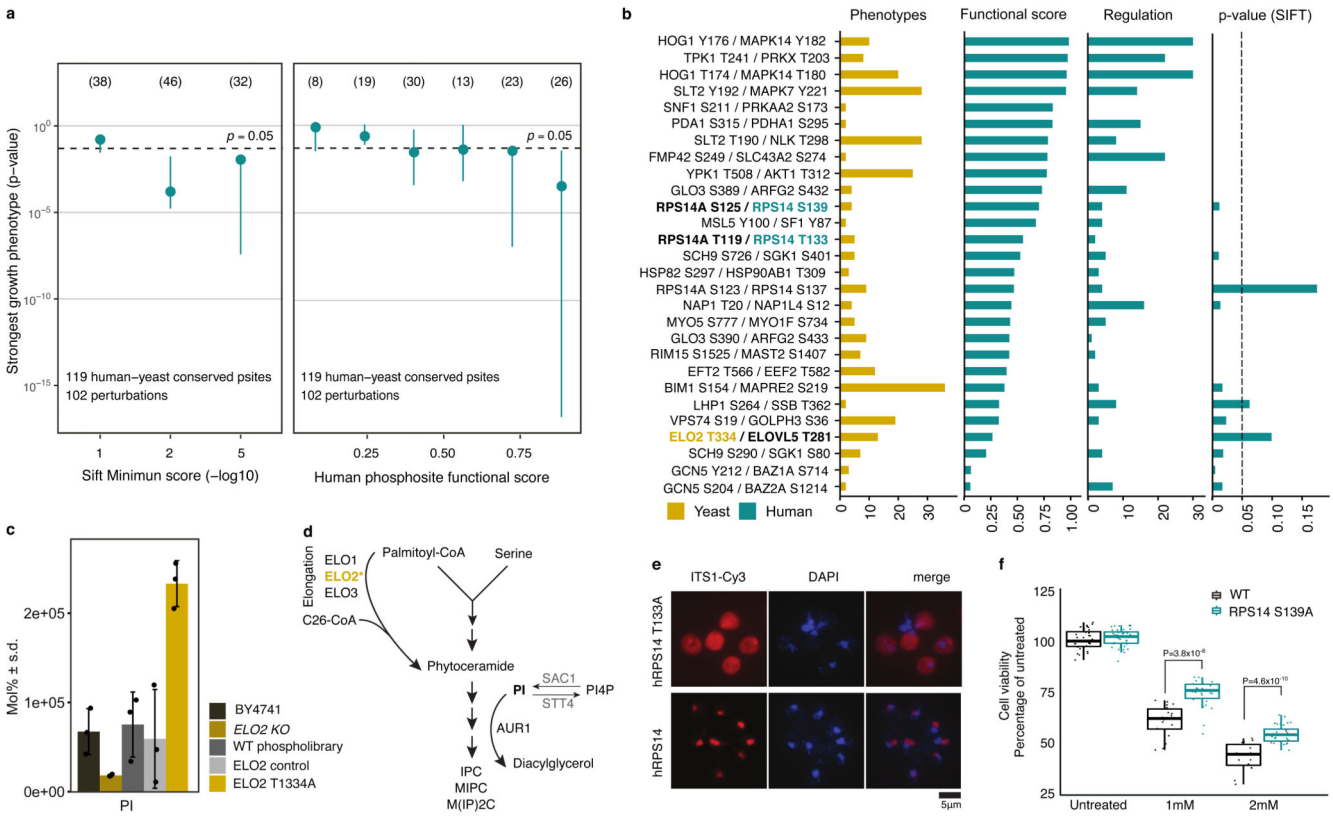

Figure 5. Yeast phospho-mutants with severe phenotypes can inform on the functional relevance of human phosphosites.

a, Human phosphosites orthologous to yeast phosphosites included in our screen were binned by SIFT and functional score and for each bin we represent the distribution of the impact of mutating the yeast phosphosites on yeast fitness (p-value of the most significant phenotype for each phospho-mutant). The p-value for a growth effect of each mutant in each condition was calculated as described in the Methods section. Data are presented as median +/- confidence intervals obtained from bootstrap (1000 resamples) of p-values. **b**, List of yeast (left) and conserved in human (right) phosphosites showing the number of yeast phenotypes in our screen (yellow) and information on the human phosphosites (green): Functional score; Regulation: phosphosite regulation in phosphoproteomic experiments in human cells and p-value; SIFT: the SIFT minimum score. **c**, Phosphatidylinositol (PI) (18:1/26:1) as molecular percentage and s.d in WT (WT phospholibrary and WT BY4741), *ELO2 KO, ELO2* control and *ELO2 T334A*. n=3 biologically independent samples and data are presented as mean values +/- SD. **d**, Simplified schematic of sphingolipid biosynthesis pathway highlighting *ELO2* (yellow); the incorporation of PI (bold); and known interactors Sac1 and Stt4 (grey). **e**, Humanized-yeast RPS14 T133A phospho-mutant exhibits cytoplasmic 20S pre-rRNA processing defect at 16°C. Localization of 20S pre-rRNA was analyzed by Fluorescence *in situ* hybridization using a Cy3-labeled oligonucleotide complementary to the 5’ sequence of ITS1 (red). Nuclear DNA was stained with DAPI (blue). Experiment was performed two times independently. **f**, Human HEK293T RPS14 S139A phospho-mutant shows increased cellular viability after treatment with

Europe PMC Funders Author Manuscripts

nickel sulfate ($NiSO_4$) 1mM (p-value=$3.86x10^{-8}$) and 2mM (p-value=$4.59x10^{-10}$). Cells were treated with 1mM or 2mM $NiSO_4$ and cell viability was measured using Cell Titer Blue assay 24h after. Cell viability is shown as percentage of the untreated control with respect to WT cells subjected to the same treatment. Data from 3 independently generated S139A phospho-mutant clones was compared to WT using a two-sided Student's t-test and P value<0.001 is displayed as ***. Data is shown for the 3 independently generated clones, 6 technical replicates and two independent experiments. Lower and upper bounds of boxplot specify 25th and 75th percentile, respectively, and lower and upper whiskers specify minimum and maximum, respectively, no further than 1.5× interquartile range. All data points are displayed.