A sentinel protein assay for simultaneously quantifying cellular processes

Martin Soste[1], Rita Hrabakova[1,2], Stefanie Wanka[3], Andre Melnik[1], Paul Boersema[1], Alessio Maiolica[4], Timon Wernas[1], Marco Tognetti[1], Christian von Mering[3] & Paola Picotti[1]

We describe a proteomic screening approach based on the concept of sentinel proteins, biological markers whose change in abundance characterizes the activation state of a given cellular process. Our sentinel assay simultaneously probed 188 biological processes in *Saccharomyces cerevisiae* exposed to a set of environmental perturbations. The approach can be applied to analyze responses to large sets of uncharacterized perturbations in high throughput.

The measurement of proteins and their post-translational modifications (PTMs) can be exploited to reveal cellular processes altered in response to new conditions. Two general mass spectrometry (MS)-based proteomic approaches have been described for protein and PTM quantification. Shotgun proteomics enables measurement of thousands of proteins[1] and can be applied in discovery-based projects, but a limitation is its semi-stochasticity, which can lead to incomplete data sets[2]. Shotgun data are typically subjected to automated quantitative analyses using computational tools for global peptide ion detection, alignment and comparison across samples. The results of these steps differ depending on the software used[3] and produce lists of protein 'hits' (i.e., proteins that change their abundance across conditions) of varying confidence. Therefore, candidate proteins for subsequent biological follow-up are typically validated by more targeted approaches such as selected reaction monitoring (SRM)[4], the second general MS-based approach, which enables the quantification of proteins in complex samples with high precision and reproducibility. Selection of targets for SRM requires a priori knowledge of the system, thus far preventing its application in purely discovery-based projects. Methods that combine global fragment-ion maps and targeted, SRM-like data extraction—such as sequential window acquisition of all theoretical spectra (SWATH)-MS—overcome this limitation, but at the cost of lower sensitivity[5].

In molecular biology laboratories, a third, low-throughput approach is commonly applied. Typically, antibodies are used to quantify (by fluorescence imaging or western blotting) previously validated protein markers as representing the activity of biological processes. Examples of such markers include the yeast protein Atg8 (LC3 in mammals), which serves as an indicator for autophagy activation[6], and phosphorylation events in the activation loop of MAP kinases, indicating signaling along this pathway[7]. Limitations of this approach include a lack of antibodies, variable antibody quality and the limited number of markers that can be evaluated simultaneously. Antibody microarrays increase throughput but otherwise face the same challenges as other antibody-based methods[8].

We sought to devise a method that combines the strengths of each of the existing approaches to (i) generate precise, nonredundant quantitative data sets with no need for subsequent validation; (ii) provide an unbiased, system-wide view, suitable for application in discovery-based projects; and (iii) report on the activity status, not merely abundance, of multiple biological processes simultaneously, in high throughput. The approach we describe here is based on the concept of 'sentinel' proteins that report on the activation state of a specific biological process and are selected on the basis of literature evidence or computational prediction (**Fig. 1**). Our sentinel assay enables the simultaneous detection of the activation of many biological processes previously detected only through a large number of specific molecular biology experiments.

First, we identified 570 potentially suitable sentinels for *S. cerevisiae* from available biological data (**Supplementary Tables 1** and **2**). We chose specific proteins, phosphorylation sites or protein degradation products that report on four classes of biological relationships (**Fig. 1a**). To retrieve experimentally validated sentinels, we browsed the literature for proteins induced in expression, phosphorylated or cleaved under specific conditions. Our selection was guided by direct reading of the scientific literature, inspecting antibody databases, discussion with yeast biologists and browsing of manually curated protein databases[9]. We selected additional sentinels through computational prediction. We defined a set of criteria that would be desirable in a sentinel protein (for example, a high degree of functional characterization or specificity for one or few pathways) and then applied a weighted scoring scheme based on such criteria to automatically rank proteins from the entire yeast proteome (**Supplementary Note**). The 30 top-scoring proteins were added to the list of biologically validated sentinels. Biologically validated sentinels were deemed 'A-grade' sentinels (88% of the set); predicted sentinels and biologically validated phosphorylation-based sentinels with

[1]Institute of Biochemistry, Department of Biology, ETH Zurich, Zurich, Switzerland. [2]Laboratory of Applied Proteome Analyses, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Libechov, Czech Republic. [3]Institute of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, Zurich, Switzerland. [4]Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, Switzerland. Correspondence should be addressed to P.P. (paola.picotti@bc.biol.ethz.ch).

RECEIVED 24 JANUARY; ACCEPTED 30 JULY; PUBLISHED ONLINE 7 SEPTEMBER 2014; DOI:10.1038/NMETH.3101

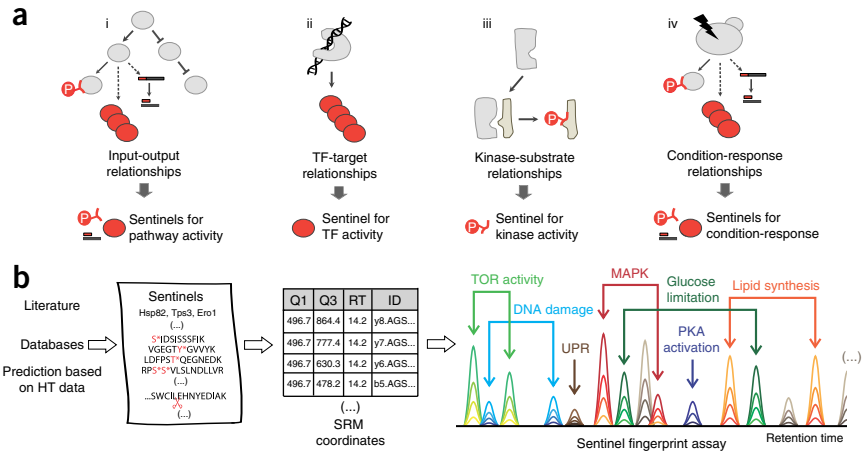

Figure 1 | Concept of sentinel proteins and the development of a sentinel fingerprint assay. (**a**) Sentinel proteins report the activation of specific cellular processes according to four general types of biological relationships. (i) If activation of a cellular pathway is known to require phosphorylation, cleavage or induction of a given pathway component, these products can be taken as sentinels for the activity of that pathway (input-output relationships). (ii) If expression of a gene is dependent on a known transcription factor (TF), the protein abundance of selected target genes can be used to hypothesize TF activation (TF-target relationships). (iii) If a protein is the known substrate of a kinase or phosphatase and the phosphorylation site is known, such a phosphorylation event can be used as a marker for the activity of that kinase and phosphatase system (kinase-substrate relationships). (iv) If cells activate a specific response upon a given environmental condition (for example, heat shock), proteins involved in such a response can be used as markers for the activation of that conditional response (condition-response relationships). P, phosphate group. (**b**) Sentinels were retrieved from databases on the basis of literature evidence or were predicted via mining of available high-throughput (HT) data sets. The optimal coordinates of specific targeted proteomic assays based on SRM were derived. Individual SRM assays were then assembled into two versions (protein- and phosphorylation-based sentinels) of a multiplexed fingerprint sentinel assay to report on the activation states of multiple cellular processes simultaneously. UPR, unfolded protein response.

an unclear phosphorylation site location were deemed 'B-grade' sentinels (12%).

Next, we developed quantitative SRM assays for 157 sentinel proteins and 152 sentinel phosphopeptides (**Fig. 1b** and **Supplementary Note**). The SRM assays (**Supplementary Table 3**) were assembled into two 30-min MS methods, enabling quantification of protein sentinels from whole proteome digests and phosphoprotein sentinels from digests subjected to a phosphopeptide-enrichment step. This 1-h multiplexed sentinel fingerprint assay reports on the activity of 188 unique biological processes.

We applied the sentinel assay to analyze yeast cells subjected to eight well-characterized perturbations: osmotic stress, osmotic stress adaptation, rapamycin treatment, amino acid and nitrogen starvation (–AA/N), entrance into stationary phase, 30 and 60 min of heat shock, and recovery from heat shock. Our unbiased SRM approach quantified 202 sentinels, allowing simultaneous monitoring of a variety of biological processes (**Fig. 2** and **Supplementary Tables 4** and **5**). Use of SWATH[5] as the targeted MS technique decreased the coverage to 66% of the protein sentinels detected by SRM (**Supplementary Table 6**), but SWATH could be advantageous for subsequent data mining and follow-up.

Our assay reported 74% of expected responses (**Supplementary Table 7**). For example, phosphorylation of T174 and Y176 on Hog1, a marker for the activation of the high-osmolarity branch of MAPK cascades[10], strongly increased only under osmotic stress conditions and returned to basal levels after adaptation to the stress (**Fig. 3a**). Tmt1, a sentinel for amino acid deprivation[11], was strongly upregulated in –AA/N (**Fig. 3b**). Several markers for the heat-shock response[12,13], including Hsp82, were induced under heat shock (**Fig. 3c**). In yeast treated with rapamycin, deactivation of TORC1 was detected by dephosphorylation of S232 and S233 on Rps6, a TORC1 target[14,15] (**Fig. 3d**). For the computationally predicted sentinels alone, we found expected responses in 73% of the cases evaluated. For both validated and predicted sentinels, response amplitudes were on average lower for the responses that did not agree with expectations.

Our sentinel screen also revealed new responses. For yeast grown to stationary phase, a condition that has been thoroughly characterized, we observed the known hallmarks of quiescence as well as novel events (**Fig. 3e** and **Supplementary Table 5**). Glucose limitation and Snf1 kinase activity were shown by increased phosphorylation of S14 on Hxk2, a known Snf1 target upon glucose deprivation[16]. Decreased phospholipid synthesis was demonstrated by dephosphorylation of S602 on the phosphatidate phosphatase, Pah1 (ref. 17). Increased phosphorylation of Bcy1 S129 indirectly reported that PKA and TORC1 activities were reduced in quiescent cells[18,19]. Novel responses included the massive induction of a marker for water deficiency, Gre1 (ref. 20), likely connected to the shrinkage of cells observed in quiescence and the resulting cytosol concentration. Our data also indicated a strong coordinated upregulation of a subset of heat-shock markers (Hsp12, Hsp26, Zim17, Pin3 and Sti1) during stationary phase, although Hsp78, Hsp82 and Hsp104 were slightly downregulated, indicating that different modules of the heat-shock response operate under specific stress types. The sentinel screen detected novel responses for several other perturbations and enabled the identification of different temporal patterns amongst sentinels, as well as separation of protein isoforms with parallel and divergent responses (**Supplementary Note**).

Our use of the term 'sentinel' represents an extension and proof of concept of what was previously proposed on a speculative basis for protein-protein interactions[21]. In principle, any protein feature that can be measured by MS at high throughput and serves as a marker for a specific biological process can be a sentinel. The sentinel assay is complementary to shotgun analyses, is fast, relies on high-confidence, information-rich markers, provides a readout for the activation status (rather than the abundance) of a multitude of biological processes, and is directly interpretable. It exploits the precision of SRM, so far applied only in hypothesis-driven projects[2,22–25], for discovery-based experiments. Possible applications include determining the network fingerprint of each drug in a collection or modeling pathway cross-talk in the context of systems biology projects. It could also be extended to mammalian systems, and most of our yeast sentinels have clear mammalian homologs. Hypotheses

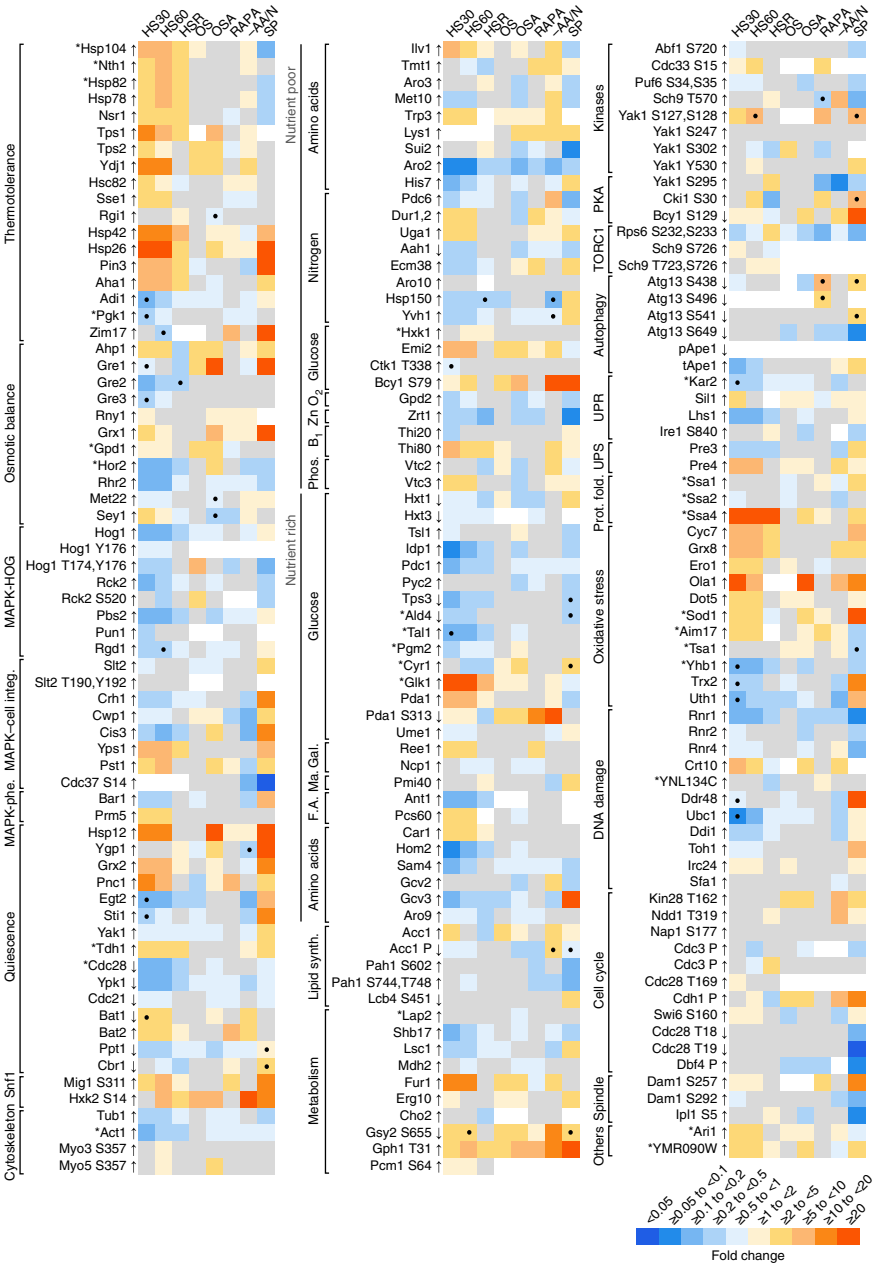

Figure 2 | Responses of yeast to eight environmental perturbations. Sentinel proteins are organized according to the cellular process they report on (left of each row) and grouped into broad categories, for space reasons. The detailed information about the specific biological process each sentinel reports on can be found in **Supplementary Tables 1** and **2**. The expected behavior of each sentinel upon activation of the corresponding biological process is indicated with upward (abundance increase) or downward (abundance decrease) arrows. The phosphorylation site targeted is indicated for phosphorylation-based sentinels. The average (triplicate cultures) abundance change of each sentinel upon a given perturbation is provided relative to the corresponding control. Significant changes (established with a linear mixed-effects model, false discovery rate–adjusted *P* value <0.05; **Supplementary Tables 4** and **5**) are depicted using a red (increased abundance) to blue (decreased abundance) color scale. Gray, no significant difference in abundance. White, not detected. "*" indicates that the sentinel was chosen by prediction. "•" indicates that the sentinel did not behave as expected from the literature (**Supplementary Table 7**). HS30, 30-min heat shock; HS60, 60-min heat shock; HSR, heat-shock recovery; OS, osmo-stress; OSA, osmo-stress adapted; RAPA, rapamycin; –AA/N, amino acid and nitrogen starvation; SP, stationary phase; MAPK–cell integ., MAPK cell integrity branch; MAPK-phe., MAPK pheromone branch; O_2, aerobic conditions; B_1, thiamine; Phos., phosphate; Gal., galactose; Ma., mannose; F.A., fatty acids; Lipid synth., lipid synthesis; UPR, unfolded protein response; UPS, ubiquitin proteosome system; Prot. fold., protein folding.

on a regulated process emerging from a sentinel screen can be followed up by targeted 'zoom-in' experiments, using biochemical tools or measuring all the components of the process that appeared regulated by SRM or SWATH after specific stimulation, as previously shown in targeted network analyses[2,24].

Our sentinel assay primarily includes validated biological markers (91% of the sentinels). Obviously, the identification of biological modules that are not covered by these markers is precluded. This could be overcome by extracting additional markers for an uncharacterized perturbation of interest from a shotgun proteomic analysis or from predictions (**Supplementary Table 8** and **Supplementary Note**). The resource of sentinel protein coordinates presented here could also be applied to directly mine shotgun proteomics data for sentinel behavior (though at lower quantitative precision and data completeness) while maintaining the possibility of discovering novel functional modules with the additional data.

A limitation of the approach is that the behavior of sentinels can be influenced by cross-talk between biological pathways. This concern applies equally to single-marker assays used in molecular biology laboratories. An advantage of the sentinel assay is that it contains multiple markers for several biological processes, and their simultaneous measurement allows assessment of their quality. The quality of our sentinels and of the associated results can also be evaluated using their underlying literature (**Supplementary Tables 1** and **2**). Some of the sentinels we could not detect are possibly associated with modules that were not activated under our conditions. Owing to the intrinsic definition of a sentinel, sentinel detectability must be considered strictly sample dependent, and the whole set of markers should always be quantified when analyzing an uncharacterized perturbation.

This straightforward approach should facilitate the application of proteomics tools in biology laboratories and enable high-throughput mechanistic analyses.

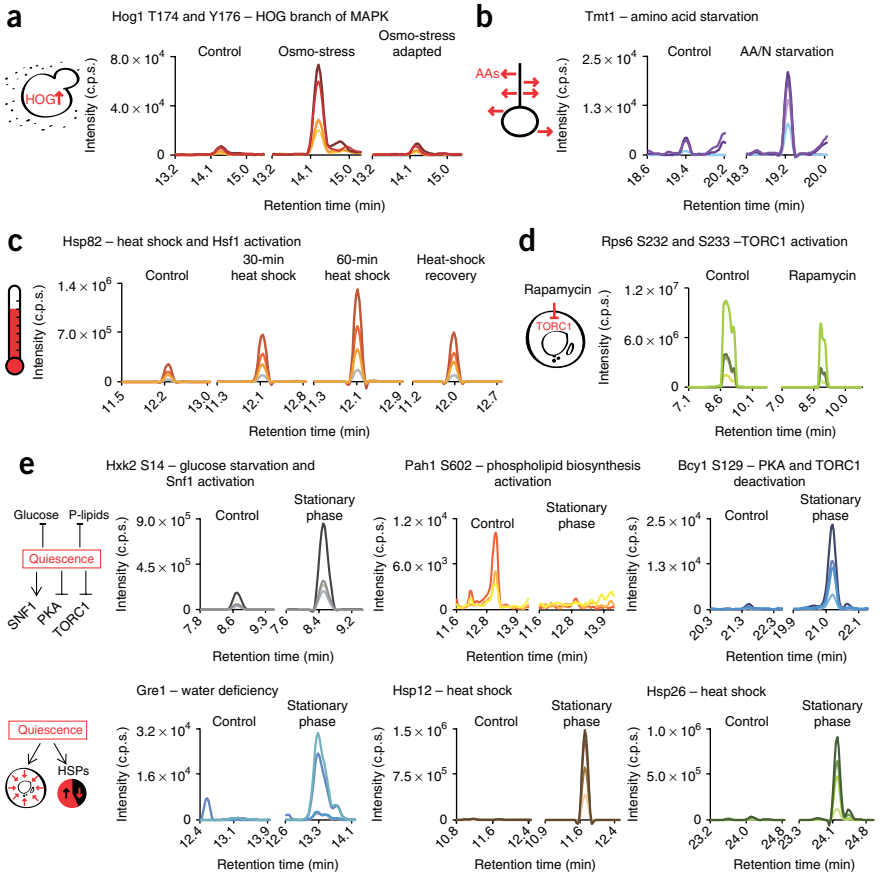

Figure 3 | Examples of confirmatory and novel sentinel behavior. Sentinel regulation is displayed as a chromatographic trace of peptide elution and detection by SRM. Colored traces represent the detection of different fragment ions from a common peptide precursor (SRM transitions). (**a**) T174 and Y176 phosphorylation on Hog1 mark activation of the HOG branch of MAPK signaling during osmotic stress. (**b**) Tmt1 abundance indicates amino acid limitation during AA/N starvation. (**c**) Hsp82 abundance suggests activation of the heat-shock response and Hsf1 during heat shock. (**d**) S232 and S233 phosphorylation on Rps6 indicate TORC1 deactivation during rapamycin treatment. (**e**) In stationary phase, phosphorylation of Hxk2 S14 indicates glucose starvation and Snf1 activation, Pah1 S602 dephosphorylation indicates decreased phospholipid biosynthesis, and phosphorylation of Bcy1 S129 indicates deactivation of PKA and TORC1. Gre1 abundance increase indicates water deficiency. A specific subset of heat-shock proteins, including Hsp12 and Hsp26, increased in abundance. P-lipids, phospholipid biosynthesis; c.p.s., counts per second.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. PeptideAtlas SRM Experiment Library (PASSEL): MS/MS data, PASS00396; SRM data, PASS00397; SWATH data, PASS00398.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

P.P. is supported by a European Research Council Starting Grant (ERC-2013-StG–337965), a 'Foerderungsprofessur' grant from the Swiss National Science Foundation (PP00P3_133670), an EU Seventh Framework Program Reintegration grant (FP7-PEOPLE-2010-RG-277147) and a Promedica Stiftung (2-70669-11). M.S. is supported by a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship D award (PGSD3-403808-2011). R.H. is supported by the Operational Program, Research and Development for Innovations (CZ.1.05/2.1.00/03.0124) and by a short-term EMBO fellowship (ASTF 475-20130). C.v.M. and S.W. acknowledge support by the SystemX.ch initiative. We thank our colleagues with experience in different aspects of yeast biology for their help: M. Peter, M. Kijanska, R. Deschant, S. Saad, A. Schreiber, C. Kraft and R. Loewith. We also thank M. Choi for assistance with MSstats and R. Shamir for insightful discussions. We are grateful to A. Bairoch and the UniProt team, as their excellent manual annotation of protein information in UniProt drastically accelerated our sentinel selection step. We thank the Peter laboratory (ETH Zurich) for the yeast strains used in this study and for sharing general lab reagents and equipment. We also thank Biognosys AG for assistance with scheduling SRM analyses using their iRT Kit.

AUTHOR CONTRIBUTIONS

M.S., R.H. and P.P. designed the experiments. M.S., R.H., A. Melnik., T.W. and M.T. performed experiments and analyzed the data. S.W. and C.v.M. designed and implemented the prediction strategy for computationally suggesting sentinel proteins. P.B. and A. Maiolica contributed to MS measurements. M.S., R.H. and P.P. wrote the manuscript. P.P. conceived of and supervised the project.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Aebersold, R. & Mann, M. *Nature* **422**, 198–207 (2003).
2. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D.A. & White, F.M. *Proc. Natl. Acad. Sci. USA* **104**, 5860–5865 (2007).
3. Nahnsen, S., Bielow, C., Reinert, K. & Kohlbacher, O. *Mol. Cell. Proteomics* **12**, 549–556 (2013).
4. Picotti, P. & Aebersold, R. *Nat. Methods* **9**, 555–566 (2012).
5. Gillet, L.C. *et al. Mol. Cell. Proteomics* **11**, 0111.016717 (2012).
6. Klionsky, D.J., Cuervo, A.M. & Seglen, P.O. *Autophagy* **3**, 181–206 (2007).
7. Cherkasova, V.A. *Methods* **40**, 234–242 (2006).
8. Stoevesandt, O. & Taussig, M.J. *Proteomics* **7**, 2738–2750 (2007).
9. Bairoch, A. *et al. Nucleic Acids Res.* **33**, D154–D159 (2005).
10. Saito, H. & Posas, F. *Genetics* **192**, 289–318 (2012).
11. Dumlao, D.S., Hertz, N. & Clarke, S. *Biochemistry* **47**, 698–709 (2008).
12. Morano, K.A., Grant, C.M. & Moye-Rowley, W.S. *Genetics* **190**, 1157–1195 (2012).
13. Erkine, A.M., Magrogan, S.F., Sekinger, E.A. & Gross, D.S. *Mol. Cell. Biol.* **19**, 1627–1639 (1999).
14. Nakashima, A., Sato, T. & Tamanoi, F. *J. Cell Sci.* **123**, 777–786 (2010).
15. Urban, J. *et al. Mol. Cell* **26**, 663–674 (2007).
16. Fernández-García, P., Peláez, R., Herrero, P. & Moreno, F. *J. Biol. Chem.* **287**, 42151–42164 (2012).
17. Homann, M.J., Poole, M.A., Gaynor, P.M., Ho, C.T. & Carman, G.M. *J. Bacteriol.* **169**, 533–539 (1987).
18. Soulard, A. *et al. Mol. Biol. Cell* **21**, 3475–3486 (2010).
19. Griffioen, G. *et al. Mol. Cell. Biol.* **21**, 511–523 (2001).
20. Garay-Arroyo, A., Colmenero-Flores, J.M., Garciarrubio, A. & Covarrubias, A.A. *J. Biol. Chem.* **275**, 5668–5674 (2000).
21. Michnick, S.W. *Curr. Opin. Biotechnol.* **14**, 610–617 (2003).
22. Altvater, M. *et al. Mol. Syst. Biol.* **8**, 628 (2012).
23. Bisson, N. *et al. Nat. Biotechnol.* **29**, 653–658 (2011).
24. Drabovich, A.P., Pavlou, M.P., Dimitromanolakis, A. & Diamandis, E.P. *Mol. Cell. Proteomics* **11**, 422–434 (2012).
25. Picotti, P., Bodenmiller, B., Mueller, L.N., Domon, B. & Aebersold, R. *Cell* **138**, 795–806 (2009).

ONLINE METHODS

Yeast culture. *Saccharomyces cerevisiae* cells—strain BY4741 for all experiments except the heat-shock experiment, where strain W303 was used (Peter laboratory, ETH Zurich)—were grown (3 independent biological replicates) at 30 °C in synthetic complete (SC) liquid medium[26] containing 2% glucose (w/v). Cultures for the heat-shock experiment were grown to an OD_{600} of ~0.7. Control cells were harvested, and cells for the heat-shock experiment were transferred to medium prewarmed to 42 °C and then harvested after 30 and 60 min. Remaining cells were transferred back to medium at 30 °C and then harvested after 2 h for the heat-shock recovery samples. To sample cells experiencing osmotic stress, cells in stationary phase, and their control, we split cultures for each different treatment. Control cells were harvested when the replicates had reached an OD_{600} of ~0.7. After the cultures were split, cells for osmotic stress experiments were treated with 0.4 M NaCl for 10 min and harvested. Cells adapted to osmotic stress were harvested 1.5 h after the salt stress was applied. Stationary-phase cells were grown for approximately 2 d to an OD_{600} of ~5.4. Rapamycin-treated samples were prepared by growing cells to an OD_{600} of ~0.3, at which point rapamycin in 90% ethanol/10% Tween-20 was added to the cultures to a final concentration of 220 nM, whereas control cultures received buffer alone. Rapamycin-treated cells and their controls were harvested at an OD_{600} of ~0.55 after 3 h. Amino acid and nitrogen (AA/N)-starvation samples were prepared by growing cultures to an OD_{600} of ~0.3 and then washing the cells and exchanging the medium to synthetic defined (SD) medium[26] containing 2% glucose without ammonium sulfate or amino acids added. The respective control cells were resuspended in the original growth medium. Cells for AA/N-starvation experiments and their controls were sampled at an OD_{600} of ~0.5–0.9 after 4 h. Cells were harvested by adding trichloroacetic acid (TCA) to a final concentration of 10% (w/v), cooling on ice for 10 min, centrifuging at 4 °C for 5 min and washing the pellets twice with 10 mL cold acetone. Cell pellets were stored at −80 °C until further processing. Cells were resuspended in a buffer containing 8 M urea, 100 mM ammonium bicarbonate (NH_4HCO_3) and 5 mM EDTA (pH 8.0) and disrupted by vortexing in the presence of acid-washed glass beads in four consecutive rounds of 10 min beating and 5 min centrifugation at 5,000*g* at 4 °C to remove cellular debris. Supernatants containing extracted proteins were centrifuged at 20,000*g* for 10 min in order remove any remaining debris. The protein concentration in the extracts was determined by the bicinchoninic acid assay, performed in triplicate (BCA Protein Assay Kit, Thermo Scientific).

Sample preparation for proteomic measurements. Proteins were reduced with 12 mM dithiothreitol or 10 mM TCEP for 30 min at 32 °C and alkylated with 40 mM iodoacetamide for 45 min at 25 °C, in the dark. Samples were diluted with 0.1 M NH_4HCO_3 to a final concentration of 1.5 M urea, and sequencing-grade porcine trypsin (Promega) was added to a final enzyme/substrate ratio of 1/100 (w/w). Tryptic digestion was conducted at 30 °C for ~13–15 h. The digestion was stopped by acidification with formic acid to a final pH <3. The peptide mixtures were loaded onto Sep-Pak tC18 cartridges (Waters) desalted according to the manufacturer instructions and eluted with 50% acetonitrile. Peptide samples were evaporated on a vacuum centrifuge to dryness and resolubilized in 0.1% formic acid for LC-MS analysis.

For phosphoproteomic measurements, the peptide mixtures were enriched for phosphopeptides by titanium dioxide (TiO_2) chromatography. Peptides (~3 mg) were reconstituted in a mixture of 80% ACN, 3.5% trifluoroacetic acid (TFA), saturated with phthalic acid. The peptide solution was added to TiO_2 resin (GL Science) at a ratio of 1 mg peptides:1 mg TiO_2 resin and was incubated for 1 h with end-over-end rotation. The resin was washed twice with the above described saturated phthalic acid solution, twice with a 80% ACN, 0.1% TFA solution, twice with a 40% ACN, 0.1% TFA solution and finally twice with 0.1% TFA. Phosphopeptides were eluted twice with 0.3 M NH_4OH. After elution, the pH was rapidly adjusted to 2.7 using 25% TFA, and phosphopeptides were purified using the C18 cartridge–based protocol described above. Dried phosphopeptide mixtures were solubilized in 0.1% formic acid and directly analyzed. All samples were processed in parallel with their respective controls.

Selection of biologically validated sentinels. To identify biologically relevant sentinels from existing biochemical data, we browsed the literature for proteins induced in expression, phosphorylated or cleaved under specific conditions. We mined databases of antibodies targeting *S. cerevisiae* proteins (Santa Cruz Biotechnology Inc., http://www.scbt.com/section/primary_antibodies_non-mammalian/yeast_proteins.html; Abcam, http://www.abcam.com/index.html?pageconfig=resource&rid=11996&pid=11320; and Cell Signaling Technology, http://www.cellsignal.com/catalog/model/s_cerevisiae.html) and manually extracted markers typically used in cell biology experiments to probe pathway activation. We also browsed the scientific literature, guided by information stored in databases on protein function, including the *Saccharomyces* Genome Database (SGD; http://www.yeastgenome.org/) and UniProt (http://www.uniprot.org/) and selected markers demonstrated to be regulated upon activation or deactivation of specific cellular processes and validated through biochemical or genetic experiments. We annotated the associated bibliographic citations. Additionally, we mined the "Induction" field of the UniProt complete data set in flat text format to identify proteins known to change their abundance when the responses to specific environmental stimuli are activated, and we extracted the associated literature. Further references to support the sentinel role of a protein were manually annotated. Additional markers were derived from discussions with yeast biologists. Information derived solely from high-throughput experiments was not considered sufficient to promote proteins to the status of sentinels.

For phosphorylation-based sentinels, tryptic peptides embedding the specific marker phosphorylation site of interest were derived from the sequence of the corresponding protein. To maximize the probability of detecting the phosphorylation site by MS, we included in the target phosphopeptide list also missed-cleaved or multiply phosphorylated versions of each tryptic peptide if these species had been previously reported in the database of mass spectrometric phosphopeptide measurements, PhosphoPep (http://www.phosphopep.org/). For comprehensiveness, phosphopeptide markers based on multiply phosphorylated tryptic peptides were also selected in their various singly phosphorylated forms. The resulting set of phosphorylated peptide sequences was synthesized on a small scale in a heavy-labeled, unpurified format (Thermo Scientific Biopolymers).

Sentinel prediction. The following criteria were established in order to best assess the ability of a protein to act as a sentinel. Candidate proteins should (i) have been detected in previous MS experiments, (ii) have shown abundance changes in response to perturbations on the (a) protein and/or (b) transcript level, (iii) have human orthologs, (iv) be highly connected in the yeast protein-protein interaction (PPI) network, (v) be well characterized, (vi) be associated to only one or few pathways, and (vii) have known chemical inhibitors. To inspect proteins for sentinel properties, we searched a number of published data sets and databases. MS-based proteomics data were compiled from refs. 27–34, and transcriptome data were from refs. 35,36. The STRING database[37] was used to determine the number of interaction partners (interaction confidence score >0.9) and to identify human ortholog(s) for each yeast protein. To estimate how well characterized a protein was, we counted the number of corresponding publications as provided by the SGD. Association of a protein to a pathway or a biological function was determined using UniProt 'keywords' from the controlled vocabulary category "Biological Process" (BP), considering only root terms, and the slim version of Gene Ontology (GO) BP[38]. Chemical inhibitor information was taken from UniProt[39].

All proteins in the yeast proteome, based on the SGD database, were evaluated for meeting sentinel properties. For each property a score was assigned, and we obtained a global 'sentinel score' by adding up all feature scores. Individual scores were awarded as follows. For properties (i) "MS detectability," (ii) "protein response to perturbation" and (iii) "transcript response to perturbation," the score was the relative frequency: how often a protein occurred in the respective data sets. Existence of human ortholog(s) (iv) was awarded with a basic score of 1. For properties (v) "connectivity in PPI," (vi) "degree of characterization" and (vii) "chemical inhibitors," the score was calculated on the basis of rank: feature values from all proteins were collected (for example, numbers of interaction partners) and unique values were sorted in descending order. The score for a specific feature value was then defined as its rank divided by the number of unique feature values. For evaluating (viii) "pathway membership," UniProt root terms and GO BP slim terms were weighted according to their number of proteins assigned. 'Rare' terms with few annotated proteins were assigned a higher score than those of more general terms comprising many proteins. The pathway score for a protein was the sum of weights of the term(s) the protein was assigned to, divided by the number of terms. Proteins that were not assigned to any term were awarded a basic score of 10. To address the significance of a property defining a sentinel, we weighted all sentinel properties according to the following scoring scheme: i, 12; ii, 14; iii, 10; iv, 5; v, 10; vi, 10; vii, 20, (unassigned, 10); and viii, 4. The results of the protein-sentinel prediction are shown in **Supplementary Table 8**.

SRM assay development. Two different approaches were used for SRM assay development, with the choice depending on whether the target was a protein- or a phosphorylation-based sentinel. Development and validation of SRM assays to measure abundances of yeast proteins was guided by a spectral library generated in-house through a LC-MS/MS analysis of the yeast samples analyzed in the present study (all conditions) on a 5600 TripleTOF mass spectrometer equipped with a nano-electrospray ion source (ABSciex). On-line chromatographic separation of the peptides was achieved with an Eksigent 1D-plus Nano liquid chromatography system (Eksigent/ABSciex) equipped with a 20-cm fused-silica column with 75-µm inner diameter (New Objective), packed in-house with Magic C18 AQ 3-µm beads (Michrom Bioresources). The peptide mixtures (~2 µg) were loaded from an AS-2 autosampler (Eksigent/ABSciex), cooled to 4 °C and separated with a linear gradient from 5% to 35% acetonitrile in 120 min. The instrument was operated as in ref. 40. The collected spectra were searched against the *S. cerevisiae* SGD protein database with Sorcerer-Sequest (Thermo Electron). Trypsin was set as the digesting protease with the tolerance of no missed cleavages and fully tryptic termini. The monoisotopic peptide and fragment mass tolerances were set to 50 p.p.m. and 0.8 Da, respectively. Carbamidomethylation of cysteines (+57.0214 Da) was defined as a fixed modification, and oxidation of methionines (+15.99492) as a variable modification. Protein identifications were filtered with a false discovery rate of <2%, calculated on the basis of a target-decoy approach[41]. SpectraST[42] was used to build a consensus spectral library. The raw data can be accessed at http://www.peptideatlas.org/PASS/PASS00396.

Proteotypic peptides matching to the spectral library were ranked by intensity using Skyline[43] (v2.0.9.4899, release date 29 August 2013, MacCoss Lab Software, University of Washington), and the ten most intense transitions consisting of doubly or triply charged precursor ions and singly or doubly charged fragment ions of the y- and b-ion series were experimentally tested in SRM mode to select the most suitable transitions for quantification experiments. Samples were measured on a triple-quadrupole/ion-trap mass spectrometer (5500 QTrap, ABSciex) equipped with a nano-electrospray ion source and operated in SRM mode. On-line chromatographic separation of the peptides was achieved with an Eksigent 1D-plus Nano liquid chromatography system equipped with a 20-cm fused-silica column with 75-µm inner diameter, packed in-house with Magic C18 AQ 5-µm beads (Michrom Bioresources). The peptide mixtures (~1 µg) were separated with a linear gradient from 5% to 35% acetonitrile in 30 min. SRM analysis was conducted with Q1 and Q3 operated at unit resolution (0.7 *m/z* half-maximum peak width) with a dwell time of 20 ms and a cycle time <3.0 s. Data were analyzed with Skyline, and for the final SRM assays, the top three or four transitions per peptide were retained. The identity of each SRM peak was confirmed against the spectral library by matching realigned retention times and relative fragment-ion intensities.

In total, 300 peptides matching to 156 proteins were selected for use during quantification experiments. For 139 proteins, the optimal SRM transitions of the two peptides with the highest signal-to-noise ratio in the fragment-ion spectrum were selected from the spectral library. If only one peptide was available in the library, either an additional proteotypic peptide from http://www.peptideatlas.org/ was selected (EKPLNQLEESSRPPLAK from Hsp42, STAAAVSQIGDGQVQATTK from Hsp150 and VNGLQDIFQIAK from Thi20) or one proteotypic peptide was used (14 proteins). For Atg8, an important autophagy marker, no fragment-ion spectrum was found in the library; therefore, to extract SRM coordinates, we used heavy-labeled unpurified synthetic peptides (Thermo Biopolymers) as described below. Indexed peptide retention times (iRTs, Biognosys AG) were annotated for time-scheduled SRM acquisition[44]. Overall,

doi:10.1038/nmeth.3101

1,174 transitions were selected to monitor protein sentinels (**Supplementary Table 3**).

For the protein Ape1, ten proteotypic peptides matching to the spectral library were tested in SRM mode, and assays mapping to both the mature (mApe1) and precursor (pApe1) proteins were developed (**Supplementary Fig. 1**). One additional assay, specific to the pApe1 form because it embeds the cleavage site, was developed by testing the most probable fragment ions. The two forms of Ape1 were counted as two distinct proteins but as only one sentinel because the ratio, and not the individual forms of the protein, indicates the activation of autophagy[6]. With up to six transitions per peptide, a total of 46 transitions for Ape1 peptides were selected (degradation sentinels) (**Supplementary Table 3**).

Development and validation of SRM assays for phosphopeptides was performed as previously described[45] using heavy-labeled, unpurified synthetic versions of each peptide. To extract their SRM coordinates, we first performed LC-MS/MS as described above with a 5600 TripleTOF mass spectrometer on mixtures of the synthetic peptides. The raw data can be accessed at http://www.peptideatlas.org/PASS/PASS00396. For each target peptide, the ten most intense SRM transitions consisting of doubly, triply or quadruply charged precursor ions and singly or doubly charged fragment ions of the b- and y-ion series and the associated neutral loss ions were selected using Skyline. For assay refinement, we experimentally tested the set of transitions in SRM mode to select the 4–6 most suitable transitions for quantification experiments. Where possible, we selected transitions that were diagnostic for phosphorylation on the specific targeted residue of the phosphopeptide. The identity of each SRM peak was confirmed against the spectral library by matching relative fragment-ion intensities. Synthetic phosphopeptides that could not be detected by automated LC-MS/MS were monitored by LC-SRM, testing doubly, triply or quadruply charged precursor ions and the most probable fragment ions from the b- and y-ion series and the corresponding neutral losses. Again, only the top 4–6 transitions per peptide with no obvious interferences were retained. A total of 807 transitions were selected to monitor phosphorylation-based sentinels (**Supplementary Table 3**). The indexed peptide retention times were annotated for scheduled SRM acquisition. The naked sequences of five targeted phosphosites (S[+80]TGGK from Hht1, PAS[+80]KAPAEK from Htb1, KGS[+80]MADVPK from Hxk2 and RAS[+80]S[+80]LKA from Rps6) were not strictly proteotypic (**Supplementary Table 3**). The peptide from Rps6 was mapped to YPL090C (Rps6a) and YBR181C (Rps6b), two protein identifiers referring to identical protein sequences. The same applied to the peptide from Hht1 mapping also to Hht2, a protein with an identical sequence and UniProt accession number. The peptide from Hxk2 was also mapped to Hxk1, a protein that shares 77% of its sequence with Hxk2. On the basis of the curated data in UniProt, there is evidence for phosphorylation of only S14 on Hxk2 (refs. 46–48). If an unknown phosphorylation occurred at Hxk1 S14, then this assay would not differentiate this from Hxk2 S14 and would report the sum abundance of the two sites. The target (phospho)peptide from Htb1 also maps to Htb2, a protein with 97% similarity in amino acid sequence. There is evidence for the same phosphosite on both histone proteins[49,50], and this SRM assay does not differentiate these events. If the protein of a phospho-sentinel was found in the spectral library generated from LC-MS/MS analyses (see above), then this protein was also included in the protein-sentinel fingerprint assay.

Sentinel quantification. Sentinels were measured in each perturbation condition by SRM or SWATH acquisition. Sentinel proteins and protein degradation products were measured from tryptic peptide mixtures. Sentinel phosphopeptides were measured from the same mixtures after phosphopeptide enrichment by TiO_2 chromatography.

In time-scheduled SRM experiments, iRTs extracted during the assay refinement were used to schedule acquisition of SRM traces within retention-time (RT) windows. One LC-SRM method was used for protein-sentinel analysis (total 1196 transitions, 3-min RT window, 2.0-s cycle time, ~2 µg peptides, 30-min gradient), and another LC-SRM method was used for phospho-sentinel analysis (total 825 transitions, 2.5- to 3-min RT window, 2.3-s cycle time, ~2–4 µg peptides, 30-min gradient). Corresponding conditions and controls were analyzed with the same RT window and amount of peptides injected. Both analyses used instrument settings as described above. The raw data can be accessed at http://www.peptideatlas.org/PASS/PASS00397.

The SWATH measurement was performed with a 5600 TripleTOF MS. The peptide mixture (~4 µg) was separated by a 120-min gradient from 1–35% acetonitrile, and the mass spectrometer was operated in a looped product-ion mode. Using an isolation width of 26 *m*/*z* (containing 1 *m*/*z* for the window overlap), a set of 32 overlapping windows was constructed covering a precursor mass range between 400 and 1,200 *m*/*z*. SWATH MS2 spectra were collected between 50 and 2,000 *m*/*z*. The collision energy for each window was determined according to the calculation for a charge 2+ ion centered upon the window with a spread of 15 V. For fragment-ion scans and for the high-resolution survey scan acquired at the beginning of each cycle, an accumulation time of 100 ms and the high-sensitivity mode were used, resulting in a total cycle time of ~3.3 s. The raw data can be accessed at http://www.peptideatlas.org/PASS/PASS00398.

Validation of protein sentinel identification using synthetic peptides. In order to validate the detection of protein-based sentinels in the sentinel fingerprint assay, we synthesized surrogate peptides containing a heavy-isotope label (Thermo Scientific Biopolymers). A mix of all synthetic peptides was spiked into a pooled sample of all conditions used in this study. Corresponding heavy and light transitions were targeted by scheduled SRM in order to monitor the coelution of endogenous (i.e., light) peptides and the spiked-in (heavy) surrogates. With the exception of the Pre3 peptide, WDGSSGGVIR, for which an assay based on the synthetic peptide could not be developed, all protein-based sentinel heavy peptides were also used to confirm the SRM coordinates presented in **Supplementary Table 3**.

Data analysis. SRM peaks were manually inspected using Skyline by checking for coelution, peak shape similarity, a match in relative intensities of fragment ions and retention times compared to the assay development phase (described above). Only SRM peaks appearing in either treatment or control with a signal-to-noise ratio of >3 for at least the top transition were considered. The top 3–6 transitions per peptide with no obvious interferences were retained for quantification. Raw SRM data (peak areas) were

exported from Skyline and further analyzed using the MSstats software package[51,52]. Briefly, the log-transformed values were normalized on the basis of endogenous signals of all peptides across MS runs (constant normalization) in order to remove technical variation. A linear mixed-effects model was applied for protein significance analysis to distinguish differential abundance from noise. Relative quantification of a given perturbation was performed with respect to its control. Significant abundance changes were reported as $\log_2$ fold changes with standard error, *T* value, degree of freedom and *P* value adjusted for multiple comparisons (**Supplementary Table 4**). Multiple SRM transitions per peptide (for protein- and phospho-sentinels) and multiple peptides per protein (for protein-sentinels) were used for the statistical analysis. An FDR-adjusted *P* value cutoff of 0.05 was used. The fold-change cutoff was calculated on the basis of the number of biological replicates per condition (3 in all experiments except for the AA/N starvation experiment, for which 1 of the 3 replicates was lost), peptides per protein and transitions per peptide while retaining a statistical power (the probability of detecting a true fold change) >0.8 (refs. 51,52).

To directly compare the detection rate of SWATH and SRM data acquisition for protein sentinels, we selected one replicate of the proteome digests from yeast subjected to osmotic stress. For the SWATH data, peaks were integrated in Skyline as described above for SRM data analysis. The number of successfully detected peptides and proteins in the SWATH and SRM data was then compared (**Supplementary Table 6**).

Success rate calculation. To calculate the success rate of the sentinel fingerprint assay, we referred to the literature to evaluate whether or not our data agreed with the expected responses. To determine the expected responses of our sentinels, we used the "Sentinel role" references in **Supplementary Tables 1** and **2** as well as published reviews covering the perturbations applied (see references in **Supplementary Table 7**). Only sentinels for which the connection between the perturbation and the biological process the sentinel reports on was clear were included in the calculation. The number of expected and unexpected responses was counted for both protein and phosphorylation-based sentinels. The average fold change was also calculated.

26. Sherman, F. *Methods Enzymol.* **350**, 3–41 (2002).
27. Wang, M. *et al. Mol. Cell. Proteomics* **11**, 492–500 (2012).
28. Costenoble, R. *et al. Mol. Syst. Biol.* **7**, 464 (2011).
29. Bodenmiller, B. *et al. Sci. Signal.* **3**, rs4 (2010).
30. Dephoure, N. & Gygi, S.P. *Sci. Signal.* **5**, rs2 (2012).
31. Nagaraj, N. *et al. Mol. Cell. Proteomics* **11**, M111.013722 (2012).
32. Poulsen, J.W. *et al. J. Proteomics* **75**, 3886–3897 (2012).
33. Soufi, B. *et al. Mol. Biosyst.* **5**, 1337–1346 (2009).
34. Vogel, C., Silva, G.M. & Marcotte, E.M. *Mol. Cell. Proteomics* **10**, M111.009217 (2011).
35. Gasch, A.P. *et al. Mol. Biol. Cell* **11**, 4241–4257 (2000).
36. Su, L.J. *et al. Dis. Model. Mech.* **3**, 194–208 (2010).
37. Franceschini, A. *et al. Nucleic Acids Res.* **41**, D808–D815 (2013).
38. Ashburner, M. *et al. Nat. Genet.* **25**, 25–29 (2000).
39. UniProt Consortium. *Nucleic Acids Res.* **41**, D43–D47 (2013).
40. Di Girolamo, F., Righetti, P.G., Soste, M., Feng, Y. & Picotti, P. *J. Proteomics* **89**, 215–226 (2013).
41. Elias, J.E. & Gygi, S.P. *Nat. Methods* **4**, 207–214 (2007).
42. Lam, H. *et al. Nat. Methods* **5**, 873–875 (2008).
43. MacLean, B. *et al. Bioinformatics* **26**, 966–968 (2010).
44. Escher, C. *et al. Proteomics* **12**, 1111–1121 (2012).
45. Picotti, P. *et al. Nat. Methods* **7**, 43–46 (2010).
46. Kriegel, T.M., Rush, J., Vojtek, A.B., Clifton, D. & Fraenkel, D.G. *Biochemistry* **33**, 148–152 (1994).
47. Behlke, J. *et al. Biochemistry* **37**, 11989–11995 (1998).
48. Holt, L.J. *et al. Science* **325**, 1682–1686 (2009).
49. Ahn, S.H. *et al. Cell* **120**, 25–36 (2005).
50. Ahn, S.H., Henderson, K.A., Keeney, S. & Allis, C.D. *Cell Cycle* **4**, 780–783 (2005).
51. Choi, M. *et al. Bioinformatics* doi:10.1093/bioinformatics/btu305 (2 May 2014).
52. Chang, C.Y. *et al. Mol. Cell. Proteomics* **11**, M111.014662 (2012).

doi:10.1038/nmeth.3101