Analysis of yeast protein kinases using protein chips

Heng Zhu[1], James F. Klemic[2,3], Swan Chang[2], Paul Bertone[1], Antonio Casamayor[1], Kathryn G. Klemic[4], David Smith[1], Mark Gerstein[5], Mark A. Reed[2,3] & Michael Snyder[1,5]

We have developed a novel protein chip technology that allows the high-throughput analysis of biochemical activities, and used this approach to analyse nearly all of the protein kinases from *Saccharomyces cerevisiae*. Protein chips are disposable arrays of microwells in silicone elastomer sheets placed on top of microscope slides. The high density and small size of the wells allows for high-throughput batch processing and simultaneous analysis of many individual samples. Only small amounts of protein are required. Of 122 known and predicted yeast protein kinases, 119 were overexpressed and analysed using 17 different substrates and protein chips. We found many novel activities and that a large number of protein kinases are capable of phosphorylating tyrosine. The tyrosine phosphorylating enzymes often share common amino acid residues that lie near the catalytic region. Thus, our study identified a number of novel features of protein kinases and demonstrates that protein chip technology is useful for high-throughput screening of protein biochemical activity.

Introduction

The sequencing of entire genomes has resulted in the identification of large numbers of novel ORFs. The challenge ahead is to gain information about the function of identified genes[1,2]. Currently, significant effort is devoted to understanding gene function by mRNA expression patterns and by gene disruption phenotypes[3,4]. Important advances in this effort have been possible, in part, by the ability to analyse thousands of gene sequences in a single experiment using gene chip technology. Much information about gene function comes from the analysis of the biochemical activities of the encoded protein. Currently, these types of analyses are done by individual investigators studying a single protein at a time. This can be time consuming because it can take years to purify and identify a protein on the basis of its biochemical activity. The availability of an entire genome sequence makes it possible to perform biochemical assays on every protein encoded by the genome. As such, it would be extremely powerful to analyse hundreds or thousands of protein samples using a single protein chip. Such approaches lend themselves well to high-throughput experiments in which large amounts of data can be generated and analysed.

Several groups have devised methods for expressing large numbers of proteins with potential utility for biochemical genomics in *S. cerevisiae*. InVitrogen has cloned ORFs into an expression vector that uses the *GAL* promotor and fuses the protein to a HISX6 tag; thus far they have prepared and confirmed expression of approximately 2,000 yeast protein fusions[5]. Using a recombination strategy, Eric Phizicky's group has cloned approximately 85% of the yeast ORFs into a vector that produces GST fusion proteins inder the control of the *CUP1* promotor (inducible by copper[6]). Using a pooling strategy, they identified the gene encoding several important biochemical activities (for example, phosphodiesterase and Appr-1′′-P-processing activities). Strategies to analyse large numbers of individual protein samples have not been described.

We have also overproduced yeast proteins as GST fusions and developed a protein chip technology suitable for rapidly analysing large numbers of samples; this approach was applied to the analysis of nearly all yeast protein kinases. The yeast genome has been sequenced and contains approximately 6,200 ORFs greater than 100 codons in length. Of these, 122 are predicted to encode protein kinases, and 24 of these protein kinase genes have not been studied previously[7]. Except for two histidine protein kinases, all of the yeast protein kinases are members of the Ser/Thr family; tyrosine kinase family members do not exist, although seven protein kinases that phosphorylate serine/threonine and tyrosine have been reported[7].

Here we overexpress nearly all (119) of the yeast protein kinases and used a novel protein chip technology to analyse their specificity using 17 different substrates. We find that 32 kinases preferentially phosphorylate one or two substrates, and 27 kinases readily phosphorylate poly(Tyr-Glu), suggesting that there are many more potential tyrosine kinases than were known previously. Correlation of functional specificity with amino acid sequence information reveals that the kinases that use poly(Tyr-Glu) as a substrate contain amino acids near the catalytic region that are distinct from those that do not. We expect this technology to be valuable for the analysis of entire proteomes and the information to be very valuable to researchers studying kinase-substrate reactions.

Results

Yeast kinase cloning and protein purification

Using a recombination-directed cloning strategy[8], we cloned the entire coding regions of 122 yeast protein kinase genes in a high-copy expression vector (pEG(KG)) that produces GST fusion proteins under the control of the galactose-inducible *GAL1* promoter[9] (Fig. 1*a*). GST::kinase constructs were rescued into *Escherichia coli*, and sequences at the 5′ end of each construct were determined. We successfully cloned 119 of the protein kinase genes in-frame. The three kinase genes that we did not clone were very large (4.5–8.4 kb).

Departments of [1]Molecular, Cellular, and Developmental Biology, [2]Electrical Engineering, [3]Applied Physics, [4]Cellular and Molecular Physiology, and [5]Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, USA. Correspondence should be addressed to M.S. (e-mail: michael.snyder@yale.edu).

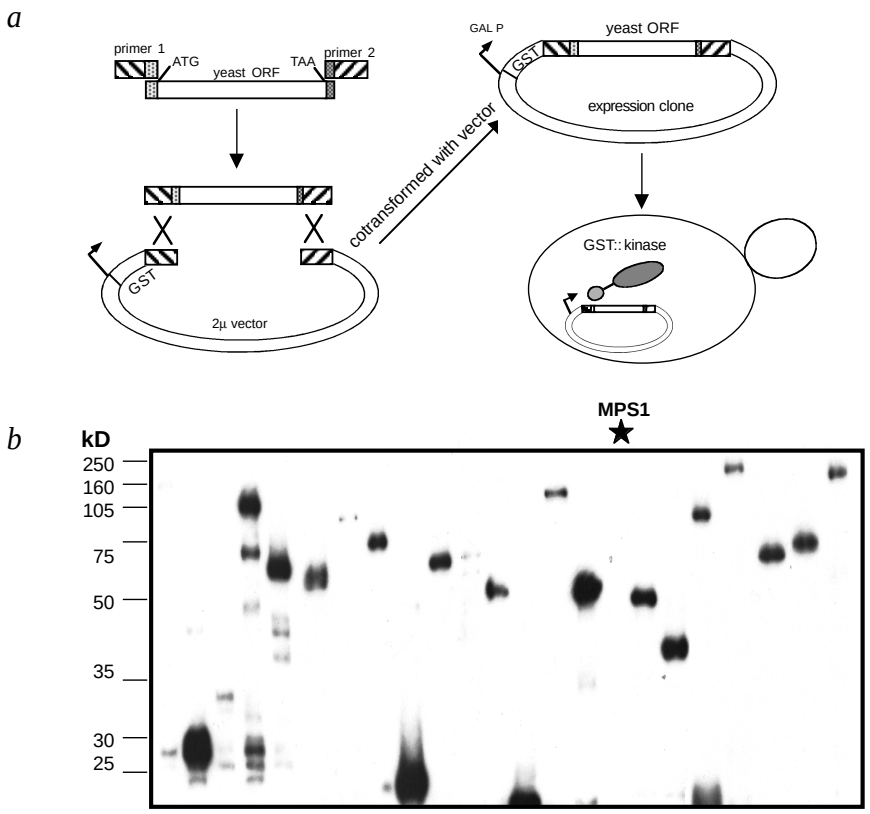


Fig. 1 Strategy to overproduce yeast protein kinases. ***a***, Using the recombination strategy[8], 119 yeast protein kinases were cloned in a high-copy *URA3* expression vector (pEG(KG)) that produces GST fusion proteins under the control of the galactose-inducible *GAL1* promoter[9]. GST:kinase constructs were rescued into *E. coli*, and sequences at the 5′ of each construct were determined. The whole procedure was repeated when mutations were discovered. ***b***, Immunoblots of GST:kinase fusion proteins purified as described. From 3 attempts we purified 105 kinase proteins. In spite of repeated attempts, we were unable to detect 14 of 119 GST fusions by immunoblotting analysis, for example, Mps1p in the lane labelled with a star.

The GST::kinase fusion proteins were overproduced in yeast and purified from 50-ml cultures using glutathione beads and standard protocols[10]. For the case of Hog1p, in the last five minutes of induction the yeast cells were treated with high salt to activate the enzyme; for the rest of the kinases, synthetic media (URA$^-$/raffinose) was used. Immunoblot analysis of all 119 fusions using anti-GST antibodies revealed that 105 of the yeast strains produced detectable GST::fusion proteins; in most cases the fusions were full length. Up to 1 μg of fusion protein per millilitre of starting culture was obtained (Fig. 1*b*), but we failed to detect 14 of 119 GST::kinase samples by immunoblotting analysis, despite repeated attempts. Presumably, these proteins are not stably overproduced in the *pep4* protease-deficient strain used, or these proteins may form insoluble aggregates that do not purify using our procedures. Although this procedure was successful, purification of GST fusion proteins using 50-ml cultures is time consuming and is not applicable for preparing thousands of samples. Therefore, we have developed a procedure for purifying proteins in a 96-well format. Using this procedure, we prepared and purified 119 GST fusions in 6 hours with approximately twofold higher yields per millilitre of starting culture relative to the 50-ml method.

Protein chip design

We developed protein chips to conduct high-throughput biochemical assays of these 119 protein kinases (Fig. 2). These chips consist of an array of microwells in a disposable silicone elastomer, poly(dimethylsiloxane) (PDMS; ref. 10). Microwell arrays allow small volumes of different analytes to be densely packed on a single chip, yet remain physically segregated during subsequent batch processing. Proteins were covalently attached to the wells using a crosslinker 3-glycidoxypropyltrimethoxysilane[11] (GPTS). Up to 8×10^{-9} μg/μm^2 of protein can be attached to the surface.

For the purposes of the protein kinase assays described here, we configured the protein chip technology to be compatible with standard sample handling and recording equipment. Using radioisotope labelling (^{33}P), the kinase assays described below and manual loading, we tested a variety of microarray configurations and found that the following chips produced the best results: round wells 1.4 mm in diameter and 300 μm deep (approximately 300 nl), in a 10×14 rectangular array configuration with a 1.8 mm pitch. We then made a master mold of 12 of these arrays and repeatedly cast microarrays for the protein kinase analysis. Chips were placed atop microscope slides for handling purposes (Fig. 2*a*); the arrays covered slightly more than one-third of a standard microscope slide and we typically used two arrays per slide (Fig. 2*b*). Although we used a manual pipette method to place proteins in each well, automated techniques may also be used. In addition, this protein chip configuration may also be used with other tagging methods such as fluorescent antibodies.

Large-scale kinase assays using protein chips

All 119 GST:protein kinases were tested for *in vitro* kinase activity[12] in 17 different assays using ^{33}Pγ-ATP and the following 17 substrates: (i) the kinases themselves (autophosphorylation); (ii) bovine histone H1 (a common kinase substrate); (iii) bovine casein (a common substrate); (iv) myelin basic protein (a common substrate); (v) Axl2 carboxy terminus-GST (Axl2 is a transmembrane phosphoprotein involved in budding[13]); (vi) Rad9 (a phosphoprotein involved in the DNA damage checkpoint[14]); (vii) Gic2 (a phosphoprotein involved in budding[15]); (viii) Red1 (a meiotic phosphoprotein important for chromosome synapsis[16]); (ix) Mek1 (a meiotic protein kinase important for chromosome synapsis[17]); (x) Poly(tyrosine-glutamate 1:4) (poly (Tyr-Glu); a tyrosine kinase substrate[18]); (xi) Ptk2 (a small-molecule transport protein[19]); (xii) Hsl1 (a protein kinase involved in cell cycle regulation[20]); (xiii) Swi6 (a phosphotranscription factor involved in G1/S control[21]); (xiv) Tub4 (a protein involved in microtubule nucleation[22]); (xv) Hog1 (a protein kinase involved in osmoregulation[23]); (xvi) Hog1 (an inactive form of the kinase); and (xvii) GST (a control). For the autophosphorylation assay, the kinases were directly adhered to the treated PDMS wells and ^{33}Pγ-ATP was added; for substrate reactions, the substrates were bound to the wells, and then kinases and ^{33}Pγ–ATP were added. After the reactions were completed, the slides were washed and the phosphorylation signals were acquired and quantified using a high-resolution phosphoimager (Fig. 3). To identify kinase activities, the quantified signals were converted into

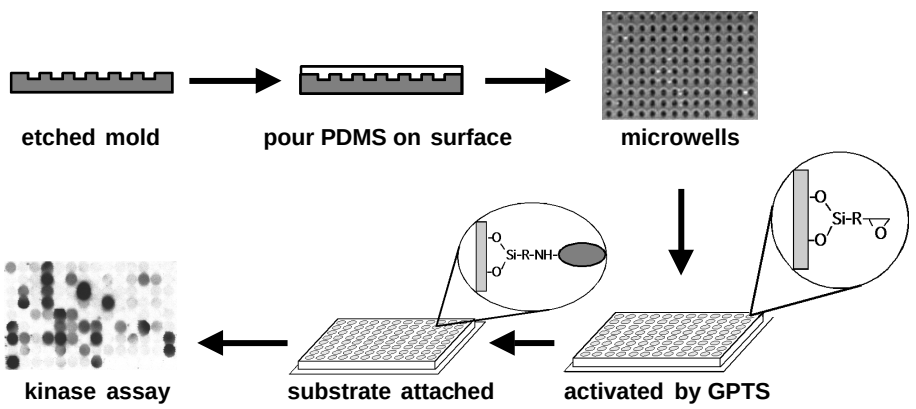
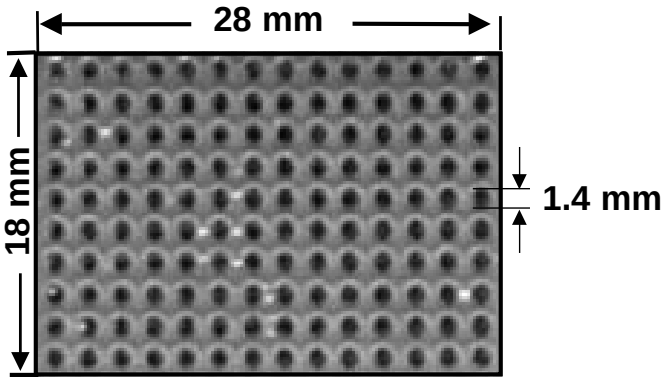


Fig. 2 Protein chip fabrication and kinase assays. ***a***, Kinase activities were detected using protein chips. PDMS was poured over the acrylic mold. After curing, the chip containing the wells was peeled away and mounted on a glass slide. The next step included modification of the surface and then attachment of proteins to the wells. Wells were blocked with 1% BSA before kinase, $^{33}P\gamma$-ATP and buffer were added. After incubation for 30 min at 30 °C, the chips were washed extensively and exposed to both X-ray film and a phosphoimager, which has a resolution of 50 μm and is quantitative. For 12 substrates each kinase assay was repeated at least twice; for the remaining 5 the assays were performed once. ***b***, An enlarged picture of the protein chip.



fold increases relative to GST controls and plotted for further analysis (Fig. 4*a*).

Most (112/119; 94%) kinases exhibited activity fivefold or greater over background for at least one substrate (Fig. 4*a*). As expected, Hrr25p, Pbs2p and Mek1p phosphorylated their known substrates[24–26], Swi6p (400-fold higher than the GST control), Hog1p (10-fold higher) and Red1p (10-fold higher), respectively. Using this assay, we found that 18 of 24 predicted protein kinases that have not been previously studied phosphorylate one or more substrates. Several unconventional kinases[7], including the histidine kinase YIL042c and phospholipid kinase Mec1p, phosphorylate protein substrates in *trans*.

To determine substrate specificity, the activity of a particular kinase was further normalized against the average of its activity against all substrates (Fig. 4*b*; all data are available at http://bioinfo.mbb.yale.edu/genome/yeast/chip). We found that 32 kinases had substrate specificity on a particular substrate with specificity index (SI) equal or higher than 2, and, reciprocally, most substrates are preferentially phosphorylated by a particular protein kinase or set of kinases. For example, the preferred substrates for YIL042C and Mec1p were Swi6p and Axl2p. The C terminus of Axl2, a protein involved in yeast cell budding, is also preferentially phosphorylated by Dbf20p, Kin2p, Yak1p and Ste20p relative to other proteins. Previous studies found that Ste20p was localized at the tip of emerging buds similar to Axl2p, and a *ste20Δcla4*^{ts} mutant is unable to bud or form fully polarized actin patches or cables[27]. Another example is the phosphoprotein Gic2, which is also involved in budding[15]. Ste20p and Skm1p strongly phosphorylate Gic2p (Fig. 4*b*). Previous studies suggested that Cdc42p interacts with Gic2p, Cla4p (ref. 28), Ste20p and Skm1p. Our results raise the possibility that Cdc42p may function to promote the phosphorylation of Gic2p by recruiting Ste20p and/or Skm1p.

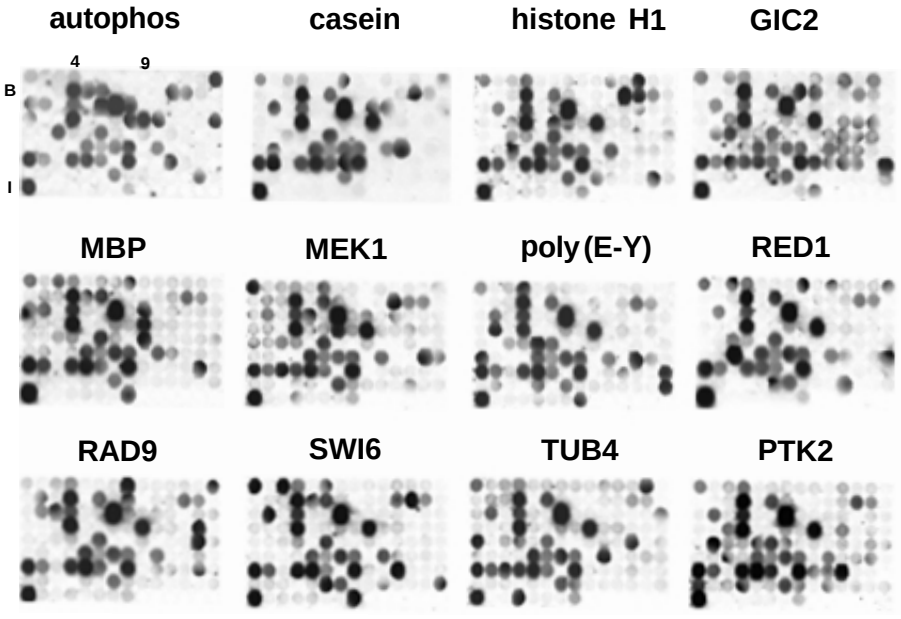


Fig. 3 The protein chip and kinase assays. Position I9 on every chip indicates the signal of negative GST control. Mps1p at position B4 exhibited strong kinase activities in all 12 kinase reactions, although no visible signal was detected by immunoblot analysis (Fig. 1*b*).

Many yeast kinases phosphorylate poly(Tyr-Glu)

On the basis of sequence analysis, all but two yeast protein kinases belong to the Ser/Thr family of protein kinases; the two exceptions are members of the histidine kinase family. Proteins of the conventional tyrosine kinase sequence family are lacking. At the time we started our study, however, seven protein kinases (Mps1, Rad53, Swe1, Ime2, Ste7, Hrr25 and Mck1) were reported to phosphorylate tyrosine[18]. We confirmed that Swe1p, Mps1p, Ime2p and Hrr25p readily phosphorylate poly(Tyr-Glu), but we did not detect any tyrosine kinase activity for Ste7p, Rad53p or Mck1p. Mck1p did not show strong activity in any of our assays, but Ste7p and Rad53p are very active in other assays. Thus, their inability to phosphorylate poly(Tyr-Glu) indicates that they either are very weak tyrosine kinases in general or are at least weak with the poly(Tyr-Glu) substrate. Consistent with the latter possibility, others have found that poly(Tyr-Glu) is a poor substrate for Rad53p (ref. 19; D. Stern, pers. comm.). We found that 23 other kinases also efficiently use poly(Tyr-Glu) as a substrate, indicating that there are at least 27 kinases in yeast that are capable of acting *in vitro* as tyrosine kinases. One of these, Rim11p, was recently shown to phosphorylate a Tyr residue on its *in vivo* substrate, Ime1p, indicating that it is a *bona fide* tyrosine kinase[29]. Thus, our experiment roughly tripled the number of kinases capable of phosphorylating tyrosine, and has raised questions about some of those classified as such kinases.

Correlation between functional specificity and amino sequences of the poly(Tyr-Glu) kinases

The large-scale analysis of yeast protein kinases allowed us to compare the functional relationship of the protein kinases with one another. We found that many of the kinases that phosphorylate poly(Tyr-Glu) are related to one another in their amino acid sequences: 70% of the poly(Tyr-Glu) kinases cluster into a distinct four groups on a dendrogram in which the kinases are organized relative to one another based on sequence similarity of their conserved protein kinase domains (Fig. 5*a*). Further examination of the amino acid sequence revealed four types of amino acids that are preferentially found in the poly(Tyr-Glu) class of kinases relative to the kinases that do not use poly(Tyr-Glu) as a substrate (three are lysines and one is a methionine); one residue (an asparagine) was preferentially located in the kinases that do not readily use poly(Tyr-Glu) as a substrate (Fig. 5*b*). Most of the residues lie near the catalytic portion of the molecule[30] (Fig. 5*b*), suggesting that they may have a role in substrate recognition.

Discussion

Large-scale analysis of protein kinases. We used a novel protein chip technology to characterize the activities of 119 protein kinases for 17 different substrates. We found that particular proteins are preferred substrates for particular protein kinases and that, vice versa, many protein kinases prefer particular substrates.

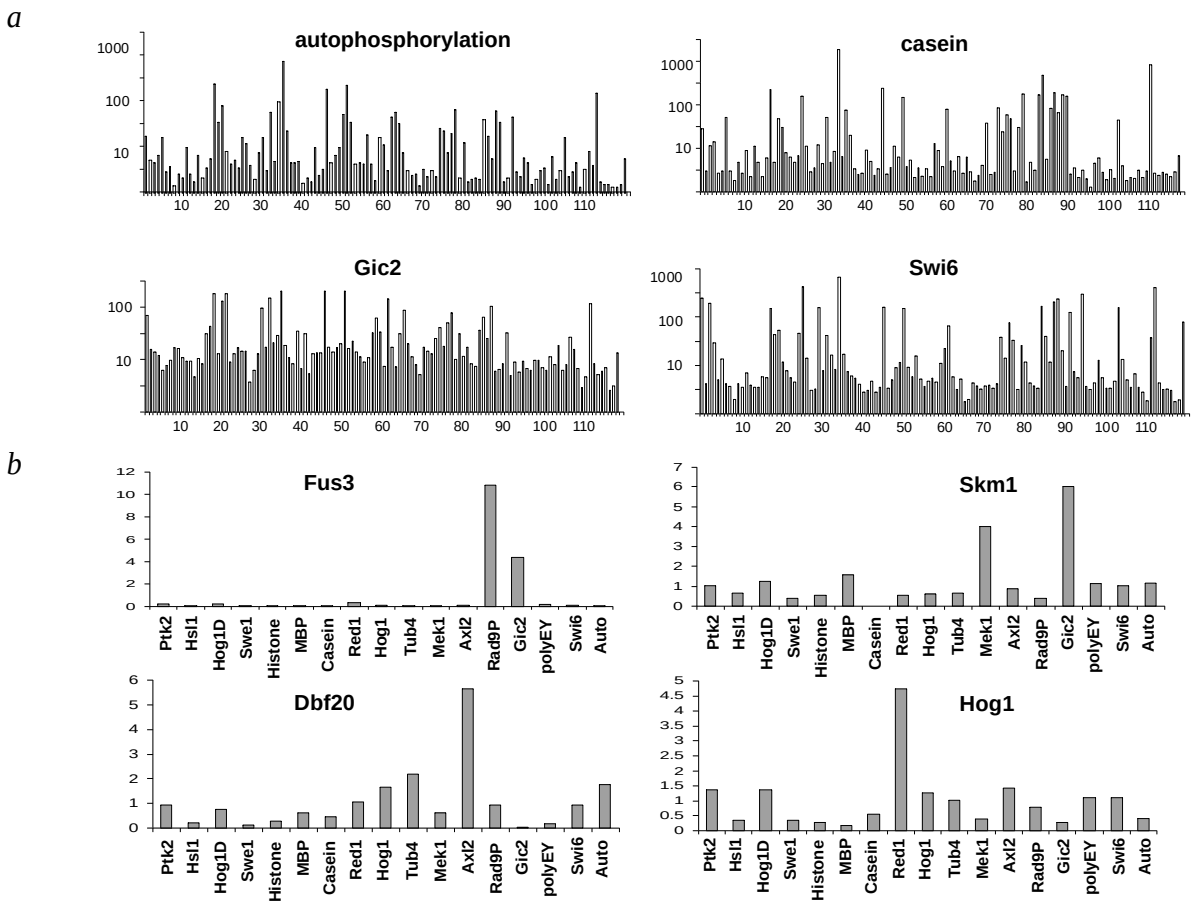


Fig. 4 Quantitative analysis of protein kinase reactions. Kinase activities were determined using a phosphoimager. The kinase signals were then transformed into fold increases by normalizing the data against negative control. ***a***, Signals of 119 kinases in 4 reactions were shown in log scale. The fold increases ranged from 1 to 1000-fold. The numbers on the axis indicate the particular kinase that was analysed (for reference numbers, see Table A, http://genetics.nature.com/supplementary_info/). ***b***, To determine substrate specificity, specificity index (SI) was calculated using the following formula: $SI_{ir}=F_{ir}/[F_{il}+F_{i2}+\cdots\cdots+F_{ir})/r]$ where *i* represents the ID of a kinase used, *r* represents the ID of a substrate, and F_{ir} represents the fold increase of a kinase *i* on substrate r compared with GST alone. Several examples of kinase specificity are shown when SI is greater than 3.The entire set of fold increase data can be retrieved from our web site (http://bioinfo.mbb.yale.edu/genome/yeast/chip).

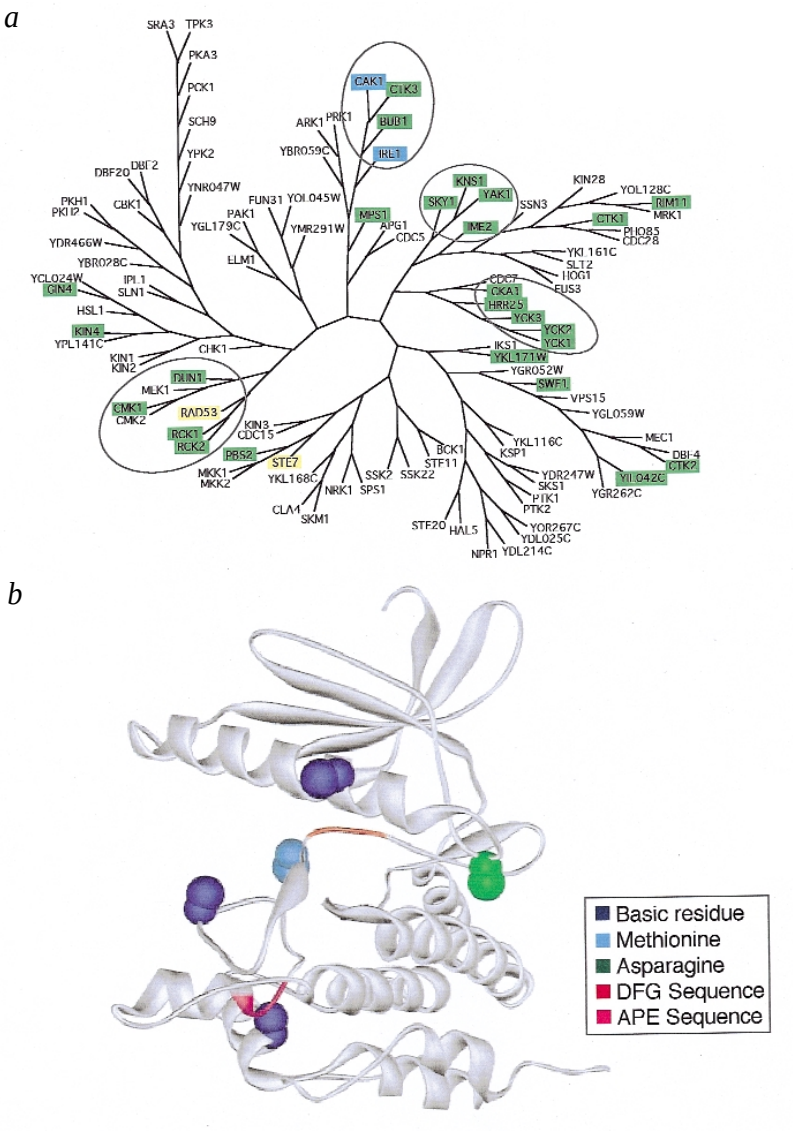


Fig. 5 Phylogenetic tree derived from the kinase core domain multiple sequence alignment, illustrating the correlation between functional specificity and amino sequences of the poly(Tyr-Glu) kinases. ***a***, Kinases that can use poly(Tyr-Glu) as a substrate often map to specific regions on a sequence comparison dendrogram. The kinases that efficiently phosphorylate poly(Tyr-Glu) are indicated in green; two kinases that weakly use this substrate are in indicated in blue. Rad53p and Ste7p, which could not phosphorylate poly(Tyr-Glu), are indicated in yellow. As shown, 70% of these kinases lie in four sequence groups (circled). ***b***, Structure of the rabbit muscle phosphorylase kinase[28] (PHK). The positions of residues preferentially found in kinases that can use poly(Tyr-Glu) as a substrate are indicated in blue (dark blue indicates a basic residue; light blue indicates a methionine); the asparagine residue that is usually found in kinases that do not use poly(Tyr-Glu) is indicated in green. The conserved DFG region that is implicated in catalysis is indicated in red, whereas the conserved APE region of the substrate binding domain is indicated in purple.

One concern with these studies is that it is possible that kinases other than the desired enzyme are contaminating our preparations. Although this cannot be rigorously ruled out, analysis of five of our samples by Coomassie staining and immunoblot staining with anti-GST antibodies does not reveal any detectable bands in our preparation that are not GST fusions.

It is important to note that *in vitro* assays do not ensure that a substrate for a particular kinase *in vitro* is phosphorylated by the same kinase *in vivo*. Other factors might restrict kinase-substrate recognition *in vivo* such as the presence of additional regulatory factors and subcellular localization. Nevertheless, these experiments indicate that certain proteins are capable of serving as substrates for specific kinases, thereby allowing further analysis. In this respect, these assays are analogous to two-hybrid studies in which candidate interactions are detected. Further experimentation is necessary to determine if the processes normally occur *in vivo*.

Consistent with the idea that many of the substrates are likely to be *bona fide* substrates *in vivo* is the observation that three kinases, Hrr25p, Pbs2p and Mek1p, phosphorylate their known substrates in our assays. Moreover, many of the kinases (for example, Ste20p) co-localize with their *in vitro* substrates (for example, Axl2p). Thus, we expect many of the kinases that phosphorylate substrates in our *in vitro* assays are likely to also do so *in vivo*.

Although most of the kinases were active in our assays, several were not. Presumably, these latter kinase preparations either lack sufficient quantities of an activator or were not purified under activating conditions. For example, Cdc28p, which was not active in our assays, might be lacking its activating cyclins. For the case of Hog1p, we treated cells with high salt to activate the enzyme. As nearly all of our kinase preparations showed activity, we presume that at least some of the enzyme in the preparation has been properly activated and/or contains the necessary cofactors. It is likely that the overexpression of these enzymes in their native organism contributes to the high success of obtaining active enzymes. It is also possible that the use of GST fusions that are capable of dimerization might augment activation of some kinases through *trans* phosphorylation. This is not the case for Hog1, which is not activated unless high salt is added to the medium.

Our assays identified many kinases that use poly(Tyr-Glu) as substrate. The large-scale analysis of many kinases allowed the novel approach of correlating functional specificity of poly(Tyr-Glu) kinases with specific amino acid sequences. Many of the residues of the kinases that phosphorylate poly(Tyr-Glu) contain basic residues. This might be expected if there were electrostatic interactions between the kinases residues and the Glu residues. The roles of some of the other residues, however, are not obvious, such as the Met residues on the kinases that phosphorylate poly(Tyr-Glu) and the Asn on those that do not. These kinase residues may confer substrate specificity by other mechanisms. Regardless, analysis of additional substrates should allow a further correlation of functional specificity with protein kinase sequence for all protein kinases.

Protein chip technology. In addition to the rapid analysis of large number of samples, the protein chip technology described here has substantial advantages over conventional methods. First, the chip-based assays have very high signal-to-noise ratios. We found that the signal-to-noise ratio exhibited using the microwell chips is much better (>10-fold) than that observed for traditional microtitre dish assays (data not shown). Presumably this is due to the fact that ^{33}Pγ-ATP does not bind the PDMS as much as microtitre dishes. Second, the amount of material needed is very small. Reaction volumes are 1/20–1/40 the amount used in the 384-well microtitre dishes; less than 20 ng of protein kinase was used in each reaction. Third, the enzymatic assays using protein chips are extremely sensitive. Even though only 105 fusions were detectable by immunoblot analysis, 112 had enzymatic activity greater than fivefold over background for

at least 1 substrate. For example, Mps1p consistently exhibited the strongest activity in many of the kinase assays, even though we have never been able to detect this fusion protein by immunoblot analysis (Figs 1*b* and 3*a*). Fourth, the chips are inexpensive; the material costs less than eight cents for each array. The microfabricated molds are also easy to make and inexpensive.

In addition to the analysis of protein kinases, this protein chip technology is also applicable to a wide variety of additional assays, such as ATP and GTP binding assays, nuclease assays, helicase assays and protein-protein interaction assays. In an independent study, yeast proteins were expressed as GST fusions under the much weaker *CUP1* promotor[6]. Although the quality of these clones has not been established, biochemical activities were identified using pools of yeast strains containing the fusion proteins. The advantage of our protein chip approach is that all samples can be analysed in a single experiment. The fact that many protein kinases are active in the autophosphorylation assay indicates that at least some of the attached protein kinases retain enzymatic activity.

We used microwells that have the advantage of reducing evaporation and segregating samples, which is particularly useful for solution-based reactions. Flat PDMS chips and glass slides, however, can also be used for different assays at high density (H.Z. and M.S., unpublished data); these have the advantage that they can be used with standard pinning tool microarrayers. This technology can also be applied to facilitate high-throughput drug screening in which one can screen for compounds that inhibit or activate enzymatic activities of any gene products of interest. Because these assays will be carried out at the protein level, the results will be more direct and meaningful to the molecular function of the protein.

We configured the protein chip technology for a specific protein kinase assay using commonly available sample handling and recording equipment. For this purpose, array dimensions remained relatively large compared with dimensions readily available with micromolded silicone elastomer structures[10,31]. Thus, it should be possible to make micromolded protein chips with microwell densities increased by several orders of magnitude and carry out high-throughput biochemical assays using arrays of 10,000 to 1,000,000 microwells using automatic sample handling and measurement techniques.

We have developed an inexpensive, disposable protein chip technology for high-throughput screening of protein biochemical activity. Its usefulness was demonstrated through the analysis of 119 protein kinases from *S. cerevisiae* assayed for phosphorylation of 17 different substrates. These protein chips permit the simultaneous measurement of hundreds of protein samples. The use of micromolded microwell arrays as the basis of the chip technology allows array densities to be increased by several orders of magnitude. With the development of appropriate sample handling and measurement techniques, these protein chips may be adapted for the simultaneous assay of several thousand to millions of samples.

Methods

Cell culture, constructs and protein purification. Using a published recombination strategy[8], we cloned 119 of 122 yeast protein kinase genes in a high-copy *URA3* expression vector (pEG(KG)) that produces GST fusion proteins under the control of the galactose-inducible *GAL1* promoter[32]. Briefly, primers complementary to the end of each ORF were purchased (Research Genetics). The ends of these primers contain a common 20-bp sequence. In a second round of PCR, we modified the ends of these products by adding sequences that are homologous to the vector. The PCR products containing the vector sequences at their ends were transformed along with the vector into a *pep4* yeast strain (which lacks several yeast proteases[9]), and Ura$^+$ colonies were selected. Plasmids were rescued into *E. coli*, verified by restriction endonuclease digestion and the DNA sequence spanning the vector-insert junction was determined using a primer complementary to the vector. For the GST::Cla4 construct, a frameshift mutation was found in a poly(A) stretch in the amino-terminal coding region. Three independent clones were required to find the correct one that maintained reading frame. For eight kinase genes we were unable to obtain a PCR product, presumably because the genes were large. For five of these genes two overlapping PCR products were obtained and introduced into yeast cells. Confirmed plasmids were reintroduced into the *pep4* yeast strain for kinase protein purification.

For preparing samples using the 96-well format, we grew cells (0.75 ml) in medium containing raffinose to O.D.(600) ~0.5 in boxes containing 2 ml wells; two wells were used for each strain. Galactose was added to a final concentration of 4% to induce protein expression, and the cells were incubated for 4 h. The cultures of the same strain were combined, washed once with 500 µl lysis buffer, resuspended in 200 µl lysis buffer and transferred into a 96×0.5 ml plate (Dot Scientific) containing 100 µl chilled glass beads. Cells were lysed in the box by repeated vortexing at 4 °C and the GST fusion proteins were purified from these strains using glutathione beads and standard protocols[19] in a 96-well format. The purity of five purified GST::kinase proteins (Swe1, Ptk2, Pkh1, Hog1, Pbs2) was determined by comparing the Coomassie staining patterns of the purified proteins with the patterns obtained by immunoblot analysis using anti-GST antibodies. The results indicated that the purified proteins are more than 90% pure. To purify the activated form of Hog1p, cells were challenged with NaCl (0.4 M) in the last 5 min of the induction. Protein kinase activity was stable for at least 2 months at –70 °C with little or no loss of kinase activity.

Chips fabrication and protein attachment. Chips were made from the silicone elastomer PDMS (Dow Chemical) cast over micromachined molds. Liquid PDMS was poured over the molds and, after curing (at least 4 h at 65 °C), flexible silicone elastomer array sheets were peeled from the reusable molds. Although PDMS may be readily cast over microlithographically fabricated structures, for the purposes of the kinase assay described herein, molds made from sheets of acrylic patterned with a computer-controlled laser milling tool (Universal Laser Systems) sufficed.

We tested over 30 different arrays. The variables tested were width and depth of the wells (widths ranging from 100 µm to 2.5 mm, depths from 100 µm to 1 mm), spacing between wells (100 µm to 1 mm), configuration (either rectangular arrays or closest packed) and microwell shape (square versus round). The use of laser-milled acrylic molds offered a fast and inexpensive method to realize a large number of prototype molds of varying parameters.

To determine the conditions that maximize protein attachment to the wells, we treated PDMS with H_2SO_4 (5 M), NaOH (10 M), hydrogen peroxide or a crosslinker GPTS (Aldrich; ref. 11). We have found that GPTS treatment resulted in the greatest absorption of protein to the microwells relative to untreated PDMS or PDMS treated other ways. Briefly, after washing with 100% ethanol three times at RT, the chips were immersed in 1% GPST solution (95% ethanol, 16 mM HOAc) with shaking for 1 h at RT. After 3 washes with 95% ethanol, the chips were cured at 135 °C for 2 h under vacuum. Cured chips can be stored in dry argon for months[11]. To attach proteins to the chips, protein solutions were added to the wells and incubated on ice for 1–2 h. After rinsing with cold HEPES buffer (10 mM HEPES, 100 mM NaCl, pH 7.0) three times, the wells were blocked with 1% BSA in PBS (Sigma) on ice for >1 h. Because of the use of GPTS, any reagent containing primary amine groups was avoided.

To determine the concentration of proteins that can be crosslinked to the treated PDMS, HRP anti-mouse Ig (Amersham) was attached to the chip using serial dilutions of the enzyme. After extensive washing with PBS, the bound antibodies were detected using an ECL kit (Amersham). We found that up to 8×10^{-9} µg/µm^2 of protein can be attached to the surface; a minimum 8×10^{-13} µg/µm^2 is required for detection by our immunostaining methods[33].

Immunoblotting, kinase assay and data acquisition. GST::protein kinases were tested for *in vitro* kinase activity[12] using ^{33}Pγ-ATP. In the autophosphorylation assay, the GST:kinases were directly adhered to GPTS-treated PDMS and the *in vitro* reactions carried out with ^{33}Pγ-ATP in appropriate buffer. In the substrate reactions, the substrate was adhered to the wells, and the wells were washed with HEPES buffer and blocked with 1% BSA before kinase, ^{33}Pγ-ATP and buffer were added. The total reaction volume was kept below 0.5 µl per reaction. After incubation for 30 min at 30 °C, the chips were washed extensively, and exposed to both X-ray film and a Molecular Dynamics phosphoimager, which has a resolution of 50 µm and is quantitative. For 12 substrates each kinase assay was repeated at least twice; for the remaining 5 the assays were performed once.

Kinase sequence alignments and phylogenetic trees. Multiple sequence alignments based on the core kinase catalytic domain subsequences of the 107 protein kinases were generated with the CLUSTAL W algorithm[33], using the Gonnet 250 scoring matrix[34]. Kinase catalytic domain sequences were obtained from the SWISS-PROT (ref. 35), PIR (ref. 36) and GenBank (ref. 37) databases. For those kinases whose catalytic domains are not yet annotated (DBF4/YDR052C and SLN1/YIL147C), probable kinase subsequences were inferred from alignments with other kinase subsequences in the data set with the FASTA algorithm[38,39] using the BLOSUM 50 scoring matrix[40]. Protein subsequences corresponding to the 11 core catalytic subdomains[41] were extracted from the alignments, and the phylogenetic trees were computed with the PROTPARS (ref. 42) program (Fig. 5*a*).

Functional grouping of protein chip data. To visualize the approximate functional relationships between protein kinases relative to the experimental data, kinases were hierarchically ordered based on their ability to phosphorylate the 12 different substrates (data available on web site). A profile corresponding to the positive or negative activity of the 107 protein kinases to each of the substrates was recorded, with discretized values in [0,1]. Matrices were derived from the pairwise Hamming distances between experimental profiles, and unrooted phylogenies were computed using the Fitch-Margoliash least-squares estimation method[43] as implemented in the FITCH program[34] of the PHYLIP software package[42]. In each case, the input order of taxa was randomized to negate any inherent bias in the organization of the data set, and optimal hierarchies were obtained through global rearrangements of the tree structures.

Acknowledgements

We thank M. Schwartz, D. Stern, J. Bailus, G. Michaud, M. Jaquenoud and M. Peter for substrates; G. Michaud for devising methods for preparing GST:fusions; G. Michaud, B. Manning, C. Horak and S. Bidlingmaier for critical comments on the manuscript; E. Skoufas for the list of protein kinases; and F.J. Sigworth for the use of his laboratory facilities to cast silicone elastomer microwells. This research was supported by grants from the National Institutes of Health, Defense Research Project Agency and the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation.

Received 9 May; accepted 26 September 2000.

1. Fields, S., Kohara, Y. & Lockhart, D.J. Functional genomics. *Proc. Natl Acad. Sci. USA* **96**, 8825–8826 (1999).
2. Goffeau, A. *et al*. Life with 6000 genes. *Science* **274**, 563–567 (1996).
3 DeRisi, J.L., Iyer, V.R. & Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680–686 (1997).
4. Winzeler, E.A. *et al*. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
5. Heyman, J.A. *et al*. Genome-scale cloning and expression of individual open reading frames using topoisomerase I-mediated ligation. *Genome Res*. **9**, 383–392 (1999).
6. Martzen, M.R. *et al*. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153–1155 (1999).
7. Hunter, T. & Plowman, G.D. The protein kinases of budding yeast: six score and more. *Trends Biol. Sci.* **22**, 18–22 (1997).
8. Hudson, J.R. *et al*. The complete set of predicted genes from *Saccharomyces cerevisiae* in a readily usable form. *Genome Res*. **7**, 1169–1173 (1997).
9. Mitchell, D.A., Marshall, T.K. & Deschenes, R.J. Vector for the inducible overexpression of glutathione S-transferase fusion protein in yeast. *Yeast* **9**, 715–723 (1993).
10. Xia, Y. *et al*. Complex optical surfaces formed by replica molding against elastomeric masters. *Science* **273**, 347–349 (1996).
11. Rogers, Y.-H. *et al*. Immobulization of oligonucleotides onto a glass support via disulfide bonds: a method for preparation of DNA microarrays. *Anal. Biochem.* **266**, 23–30 (1999).
12. Hunter, T. & Sefton, B.M. Protein phosphorylation. *Methods Enzymol*. **200**, 35–83 (1991).
13. Roemer, T.K. *et al*. Selection of axial growth sites in yeast requires Axl2p, a novel plasma membrane glycoprotein. *Genes Dev*. **10**, 777–793 (1996).
14. Weinert, T.A. & Hartwell, L.H. Cell cycle arrest of cdc mutants and specificity of the RAD9 checkpoint. *Genetics* **134**, 63–80 (1993).
15 Jaquenoud, M., Gulli, M.P., Peter, K. & Peter, M. The Cdc42p effector Gic2p is targeted for ubiquitin-dependent degradation by the SCFGrr1 complex. *EMBO J.* **17**, 5360–5373 (1998).
16. Menees, T.M., Ross-MacDonald, P.B. & Roeder, G.S. MEI4, a meiosis-specific yeast gene required for chromosome synapsis. *Mol. Cell. Biol.* **12**, 1340–1351 (1992).
17. Bailis, J.M. & Roeder, G.S. Synaptonemal complex morphogenesis and sister-chromatid cohesion require Mek1-dependent phosphorylation of a meiotic chromosomal protein. *Genes Dev.* **12**, 3551–3563 (1998).
18. Stern, D.F., Zheng, P., Beidler, D.R. & Zerillo, C. Spk1, a new kinase from *Saccharomyces cerevisiae* phosphorylates proteins on serine, threonine, and tyrosine. *Mol. Cell. Biol*. **11**, 987–1001 (1991).
19. Kaouass, M. *et al*. The STK2 gene, which encodes a putative Ser/Thr protein kinase, is required for high-affinity spermidine transport in *Saccharomyces cerevisiae*. *Mol. Cell. Biol*. **17**, 2994–3004 (1997).
20. Barral, Y., Parra, M., Bidlingmaier, S. & Snyder, M. Nim1-related kinases coordinate cell cycle progression with the organization of the peripheral cytoskeleton in yeast. *Genes Dev.* **13**, 176–187 (1999).
21. Madden, K., Sheu, Y.-J., Baetz, K., Andrews, B. & Snyder, M. SBF cell cycle regulator as a target of the yeast PKC-MAP kinase pathway. *Science* **275**, 1781–1784 (1997).
22. Sobel, S.G. & Snyder, M. A highly divergent gamma-tubulin gene is essential for cell growth and proper microtubule organization in *Saccharomyces cerevisiae*. *J. Cell. Biol*. **131**, 1775–1788 (1995).
23. Ferrigno, P., Posas, F., Koepp, D., Saito, H. & Silver, P.A. Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin β homologs NMD5 and XPO1. *EMBO J.* **17**, 5606–5614 (1998).
24. Ho, U., Mason, S., Kobayashi, R., Heokstra, M. & Andrew, B. Role of the casein kinase I isoform, Hrr25, and the cell cycle-regulatory transcription factor, SBF, in the transcriptional response to DNA damage in *Saccharomyces cerevisiae. Proc. Natl Acad. Sci. USA* **94**, 581–586 (1997).
25. Wurgler-Murphy, S.M., Maeda, T., Witten, E.A. & Saito, H. Regulation of the *Saccharomyces cerevisiae* HOG1 mitogen-activated protein kinase by the PTP2 and PTP3 protein tyrosine phosphatases. *Mol. Cell. Biol.* **17**, 1289–1297 (1997).
26. Santos, T. & Hollingsworth, N.M. Red1p, a MEK1-dependent phosphoprotein that physically interacts with Hop1p during meiosis in yeast. *J. Biol. Chem.* **274**, 1783–1790 (1999).
27. Holly, S.P. & Blumer, K.J. PAK-family kinases regulate cell and actin polarization throughout the cell cycle of *Saccharomyces cerevisiae*. *J. Cell Biol.* **147**, 845–856 (1999).
28. Richman, T.J., Sawyer, M.M. & Johnson, D.I. The Cdc42p GTPase is involved in a G2/M morphogenetic checkpoint regulating the apical-isotropic switch and nuclear division in yeast. *J. Biol. Chem.* **274**, 16861–16870 (1999).
29. Malathi, K., Xiao, Y. & Mitchell, A.P. Catalytic roles of yeast GSK3β/shaggy homolog Rim11p in meiotic activation. *Genetics* **153**, 1145–1152 (1999).
30. Owen, D.J., Noble, M.E., Garman, E.F., Papageorgiou, A.C. & Johnson, L.N. Two structures of the catalytic domain of phosphorylase kinase: an active protein kinase complexed with substrate analogue and product. *Structure* **3**, 467–474 (1995).
31. Jackman, R.J., Duffy, D.C., Cherniavskaya, O. & Whitesides, G.M. Using elastomeric membranes as dry resists and for dry lift-off. *Langmuir* **15**, 2973–2984 (1999).
32. Mylin, L.M., Hofmann, K.J., Schultz, L.D. & Hopper, J.E. Regulated GAL4 expression cassette providing controllable and high-level output from high-copy galactose promoters in yeast. *Methods Enzymol.* **185**, 297–308 (1990).
33. Higgins, D.G., Thompson, J.D. & Gibson, T.J. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**, 383–402 (1996).
34. Gonnet, G.H., Cohen, M.A. & Benner, S.A. Exhaustive matching of the entire protein sequence database. *Science* **256**, 1443–1445 (1992).
35. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL. *Nucleic Acids Res.* **27**, 49–54 (1999).
36. Barker, W.C. *et al*. The PIR-International Protein Sequence Database. *Nucleic Acids Res.* **27**, 39–43 (1999).
37. Benson, D.A. *et al*. GenBank. *Nucleic Acids Res.* **27**, 12–17 (1999).
38. Lipman, D.J. & Pearson, W.R. Rapid and sensitive protein similarity searches. *Science* **277**, 1435–1441 (1985).
39. Pearson, W.R. & Lipman, D.J. Improved tools for biological sequence comparison. *Proc. Natl Acad. Sci. USA* **85**, 2444–2448 (1988).
40. Dayhoff, M.O., Schwartz, R.M. & Orcutt, B.C. A model of evolutionary change in proteins. in *Atlas of Protein Sequence and Structure* (ed. Dayhoff, M.O.) 345–352 (National Biomedical Research Foundation, Washington DC, 1978).
41. Hanks, S.K. & Hunter, T. Protein Kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596 (1995).
42. Felsenstein, J. PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166 (1989).
43. Fitch, W.M. & Margoliash, E. Construction of phylogenetic trees. *Science* **155**, 279–284 (1967).