19 Using In Vitro Expression Cloning to Identify Interacting Proteins

Fumihiko Kanai,[1] Michael B. Yaffe,[1] and P. Todd Stukenberg[2]

[1]Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [2]Department of Biochemistry and Molecular Genetics, University of Virginia Medical School, Charlottesville, Virginia 22908-0733

INTRODUCTION

An important step in the characterization of any protein is to determine whether it exists in a complex with other proteins and, if it does, to identify its partners. This chapter describes a new technique to identify interacting proteins. It is a based on in vitro expression cloning (IVEC), which was recently developed to screen cDNA libraries rapidly and systematically according to the function of the protein encoded on each gene. This variation on traditional expression cloning has been successfully used to identify the substrates of kinases (Stukenberg et al. 1997), substrates of proteases (Cryns et al. 1997; Kothakota et al. 1997; Li 1998; McGarry and Kirschner 1998), DNA-binding activities (Mead et al. 1998), and components of signaling pathways (Andresson and Ruderman 1998). To date, IVEC has not been used to identify proteins that interact in complexes. This chapter describes a modification of the technique that successfully identified five binding partners of the specific phosphoserine/phosphothreonine-binding protein 14-3-3 (Kanai et al. 2000). Thus, IVEC can also complement the two-hybrid technique and biochemical purification to identify interacting proteins systematically.

IVEC was originally reported in 1997 (Lustig et al. 1997; Stukenberg et al. 1997). The technique combines the power of expression cloning with the ability to manipulate a reaction exper-

Protein–Protein Interactions: A Molecular Cloning *Manual*, © 2002 by Cold Spring Harbor Laboratory Press, Chapter 19.

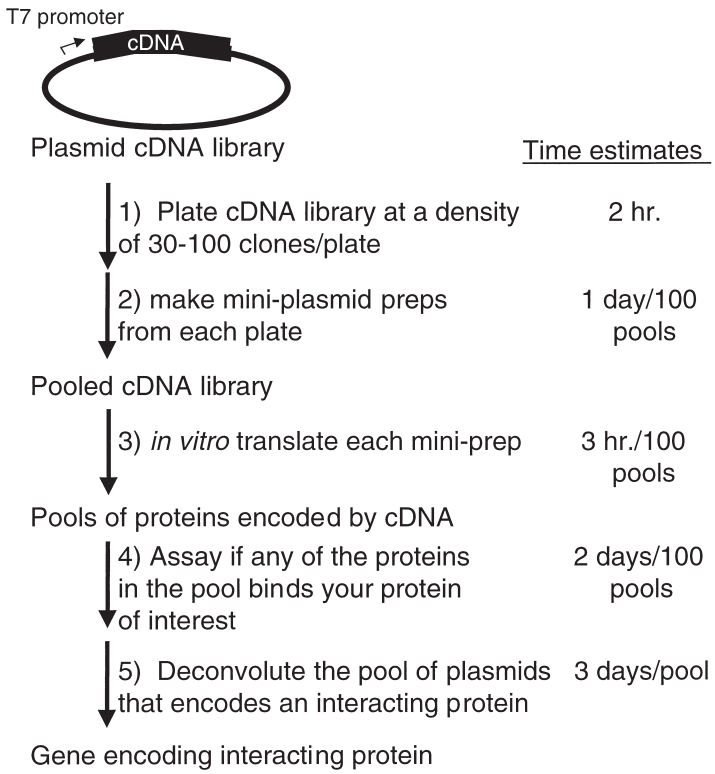

FIGURE 1. The strategy of in vitro expression cloning. See text for details.

imentally in vitro. An overview of this strategy is shown in Figure 1. The first step involves constructing or purchasing a cDNA plasmid library in which the cDNAs have been directionally cloned downstream from a T7, SP6, or T3 promoter and transformed into bacteria. (This chapter focuses on the T7 promoter, although IVEC has also been succesfully employed using SP6-based vectors.) Bacteria containing the library plasmids are plated at a density of about 30–100 clones per plate, and plasmid DNA is pooled by scraping colonies from each plate and performing small-scale plasmid purification ("minipreps"). Each plasmid pool, containing 30–100 distinct cDNAs, is then transcribed and translated in vitro in the presence of [^{35}S]methionine. The resulting labeled pools of proteins are assayed for a particular activity such as protein or ligand binding, phosphorylation by a particular kinase, or cleavage by a particular protease. Once a pool that contains the gene expressing the candidate activity has been identified, the plasmids encoding the cDNAs within the pool are subdivided and individually retested until the single cDNA encoding the protein of interest is isolated.

We describe here the general methods used to prepare library pools of cDNAs and protein, and the sib-selection techniques used to subdivide a pool once it is found to contain a candidate activity. These methods were developed in traditional academic research settings, but they could be easily modified to allow high-throughput using robotics. As an example, we describe how protein pools can be screened in vitro for binding to the specific phosphoserine/phosphothreonine-binding molecule, 14-3-3. We conclude with a discussion about the practicalities of performing IVEC and the situations where IVEC may be better than two-hybrid techniques to screen for interacting proteins.

Protocol 1

In Vitro Expression Cloning

In this protocol, we use the example of phosphoserine/phosphothreonine-binding molecule 14.3.3 to describe how protein pools can be screened for binding using IVEC.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!>.

Buffers and Reagents

E. coli, DH5α or XL2-Blue Strain
Glutathione-agarose (Sigma)
Glycerol (40%)
GST–14-3-3 fusion protein
LB-agar plates (10 cm) containing antibiotic
LB media
[^{35}S]Methionine
Plasmid cDNA library (Discovery Line Premade cDNA Library, Invitrogen)
Recombinant RNasin (ribonuclease inhibitor) (40 units/μl) (Promega)
TNT T7 coupled reticulocyte lysate system (Promega)
Tris-HCl (10 mM, pH 8.0)
Unamplified library and bacteria

Special Equipment

Eppendorf tubes (1.5 ml)
Microfuge
Molecular Imager System (Model GX-525; Bio-Rad)
Pipettor
Plates (96-well)
QIAprep Spin Miniprep Kit (Qiagen)
Rubber policeman
SDS-PAGE equipment

METHOD

Stage 1: Preparation of cDNA Pools

To obtain detectable amounts of protein by in vitro translation, it is critical to produce plasmid pools that encode a limited number of individual cDNAs (30–100 clones/pool). Although this step constitutes the major time and expense of IVEC, this single investment can be used for ~50 IVEC screens. In fact, at least three laboratories that constructed pooled libraries have used them to screen for multiple activities and have given the libraries out as collaborations with other laboratories as well.

Preparation of cDNA pools and in vitro translation reactions are performed essentially as described previously (Lustig et al. 1997). We used a commercially available plasmid cDNA library

of a human HeLa cell line from Invitrogen. This cDNA library is primed with the oligo(dT) (*Not*I) primer, and double-stranded cDNA is subcloned unidirectionally downstream from the CMV/T7 promoter of pcDNA3.1 (Invitrogen). Although this cDNA library was originally made for in vivo mammalian expression cloning, it functions well for in vitro expression screening. Furthermore, after identifying the cDNA having the desired interaction, the same construct can then be used for gene expression in mammalian cells without the necessity of recloning. For other applications, it may be advantageous to purchase or prepare cDNA libraries from a particular tissue source or organism. In each case, however, it is necessary to ensure that the library is of good quality by examining the average insert size and the percentage of empty vectors. In the human HeLa library available from Invitrogen, we found that 11 of 12 randomly chosen colonies contained plasmids with cDNA inserts, and the average insert size was ~1.2 kb.

1. To prepare pools of library cDNA that contain ~100 cDNA clones per pool, thaw an aliquot of the frozen unamplified library in bacteria, dilute with LB medium, and plate on LB-agar plates (10-cm diameter) containing appropriate antibiotics (ampicillin in the case of the Invitrogen HeLa cell library).

   It is crucial to know the titer of the original library before plating. For example, if the titer of the original library is 1×10^7 colonies/µl, then a $1:10^6$ dilution is performed and 10 µl of the final dilution is plated to get ~100 colonies/plate. The number of plates that must be prepared is calculated based on the expected number of distinct cDNAs expected to be present in the library and the depth of coverage desired during the screening. For example, we prepared 680 plates of cDNA pools that we estimated contained more than 20,000 total proteins (a density of 100 clones per pool gives ~30 proteins when translated.

In the literature, most cases have used pool sizes of ~100 clones/pool. However, we have found that translation of these pools only gives about 30 bands, suggesting that there is competition among the plasmids, and only the transcripts with the best Kozak sequences are translated in sufficient quantities to be detected. A titration of pool size suggested that the ideal pool size is closer to 30 proteins (P.T. Stukenberg and M. Kirschner, data not shown). At this lowered pool size, there are about 25 bands/pool, suggesting that most of the cDNAs in the pool are translated to a high enough concentration to be detected (Stukenberg 1997). Thus, by lowering the pool size, one can increase the coverage of the cDNA library almost threefold. To assay more than 30 proteins per assay, one can easily combine pools after the in vitro translation step (P.T. Stukenberg and M. Kirschner, data not shown).

2. Incubate the plates at 37°C until the size of bacterial colonies reaches ~1.5 mm in diameter (this typically takes overnight to 24 hours). Wrap the plates with Parafilm and store at 4°C.

   To keep the number of plates and plasmid pools manageable, the following procedures are performed using only 100 plates at a time, and subsequently repeated until all the plates have been processed. In our experience, the procedure seems to work best if the work is shared by two people working simultaneously.

3. Add 1 ml of LB medium to each of the 100 plates, and collect the bacterial colonies into microcentrifuge tubes by scraping and resuspending them using a rubber policeman. Store a small aliquot of the pooled bacteria as a 20% (v/v) glycerol stock at –80°C in either 1.5-ml Eppendorf tubes or 96-well plates.

4. Centrifuge the remainder of the pooled bacteria briefly in a microfuge, and isolate the plasmid DNA from the 100 bacterial pellets using commercially available miniprep kits. Typically, we use the miniprep kit from Qiagen (QIAprep Spin) and have consistently obtained good results for in vitro translation. The plasmid yield from one 10-cm plate with the Qiagen miniprep kit is ~5–30 µg, representing 300 ng/starting colony. The purified plasmid DNA is stored at –20°C.

Some commercially available kits for plasmid purification do not provide sufficiently good quality plasmid DNA for in vitro translation. Best results were obtained with Qiagen QIAprep Spin, and both Bio101 and Promega Wizard preps have been used successfully.

Stage 2: Preparation of In-vitro-labeled Protein Pools

1. Prepare the protein pools directly from cDNA pools using a coupled transcription/ translation system.

   We prepare protein pools using the TNT-coupled reticulocyte lysate system (Promega) according to the manufacturer's instructions, although the conditions are rescaled to use 0.1–0.6 µg of pooled cDNA per reaction in a 10-µl reaction volume. To prepare radioactively labeled protein pools, each 10-µl translation reaction contains 8 µCi of [^{35}S]methionine (10 mCi/ml) (New England Nuclear) and a mixture of amino acids lacking methionine that is supplied with the TNT kit.

   Amounts of each reagent for a single translation reaction are tabulated below. It is generally advantageous to prepare a master mix containing enough of each reagent EXCEPT THE DNA for 100 reactions.

	Per reaction	Master mix -100 reactions
TNT lysate	5.0 µl	500 µl
TNT reaction buffer	0.4 µl	40 µl
T7 RNA polymerase	0.2 µl	20 µl
Amino acid mix (–methionine)	0.2 µl	20 µl
[^{35}S]Methionine (10 mCi/ml)	0.8 µl	80 µl
Recombinant RNasin (40 units/µl)	0.2 µl	20 µl
ddH_2O	2.2 µl	220 µl
DNA (0.1~0.6 µg/µl)	1.0 µl	—
Total	10.0 µl	900 µl

   Perform the in-vitro-coupled transcription/translation reactions either in 1.5-ml microcentrifuge tubes or in 96-well plates. If the master mix is prepared, place 1 µl of DNA from each pool into each tube or well on the 96-well plate; then add 9 µl of the master mix to each tube/well. Incubate the reactions for 90 minutes at 30°C either in a water bath or hot block

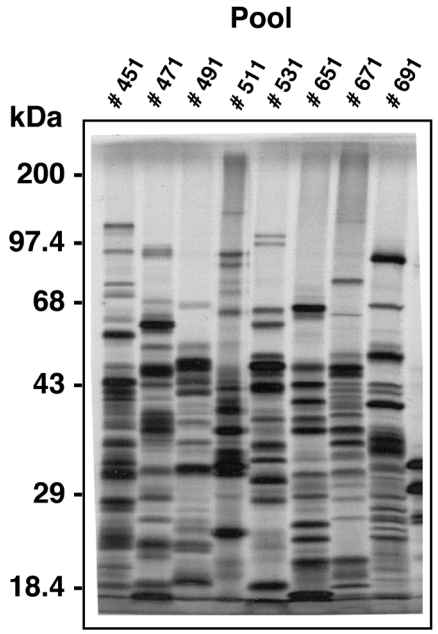

FIGURE 2. Example of in-vitro-translated pools. A total of 2 µl of each in-vitro-translated pool was analyzed by 10% SDS-PAGE and autoradiography. The mobility of molecular mass standards in kilodaltons is indicated to the left; pool numbers are shown above lanes.

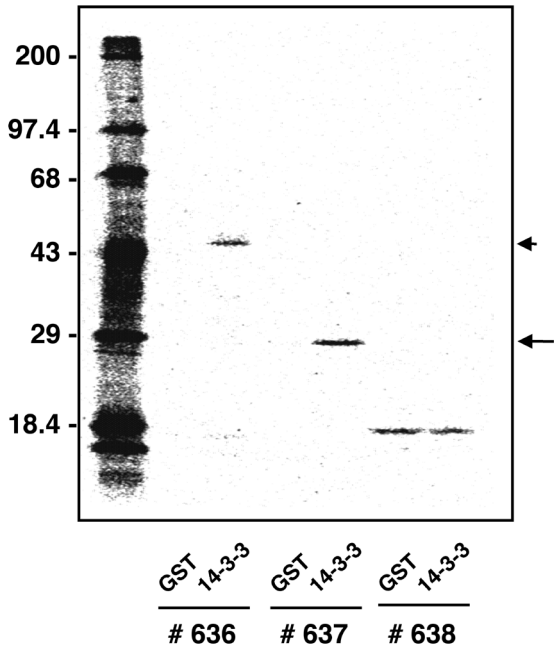

FIGURE 3. Identification of 14-3-3-binding proteins. ^{35}S-Labeled proteins from cDNA pools #636–638 were incubated with Sepharose beads containing GST or GST-14-3-3 (indicated as 14-3-3) for 2 hours at 4°C. Bound proteins were analyzed by 10% SDS-PAGE and autoradiography. The arrow and arrowhead indicate specific 14-3-3-interacting proteins later identified as a partial fragment of human TAZ (Pool #636, *arrow*) and 14-3-3 beta (Pool #637, *arrowhead*). (14-3-3 proteins are known to form homo- and heterodimers.) There is also a single band in Pool #638 that binds to both GST and GST-14-3-3, suggesting an interaction with GST rather than 14-3-3. The mobility of molecular mass standards in kilodaltons is indicated to the left.

(for microcentrifuge tubes) or in a 30°C hybridization oven (for 96-well plates). Store the radiolabeled protein pools on ice for up to 2 hours or at –80°C for 1 week.

2. Check the efficiency of translation of each pool by SDS-PAGE using 10% polyacrylamide gels followed by standard autoradiography using an overnight exposure on Model GS-525 Molecular Imager System (Bio-Rad) or a 2- to 5-day exposure on conventional X-ray film (Fig. 2).

An Example of a Binding Reaction: Using IVEC to Identify Novel 14-3-3-binding Proteins

To search for novel 14-3-3 binding proteins:

1. Preclear each ^{35}S-labeled pool by incubation with 10 pmole of GST beads for 2 hours at 4°C to reduce background binding. Incubate with 10 pmole of bead-immobilized 14-3-3–GST fusion proteins for 2 hours at 4°C with gentle agitation.
2. Centrifuge briefly, then wash the beads four times. Resuspend in Laemmli sample buffer.
3. Analyze the aliquots by SDS-PAGE and detect the radioactive protein bands using a Model GS-525 Molecular Imager System.

   In addition, it is also worthwhile to expose the gels on X-ray film because long exposures (6 days) can often uncover bands that are not detectable by phosphorimager. Using this approach, five distinct 14-3-3-binding proteins were identified (Fig. 3) (Kanai et al. 2000).

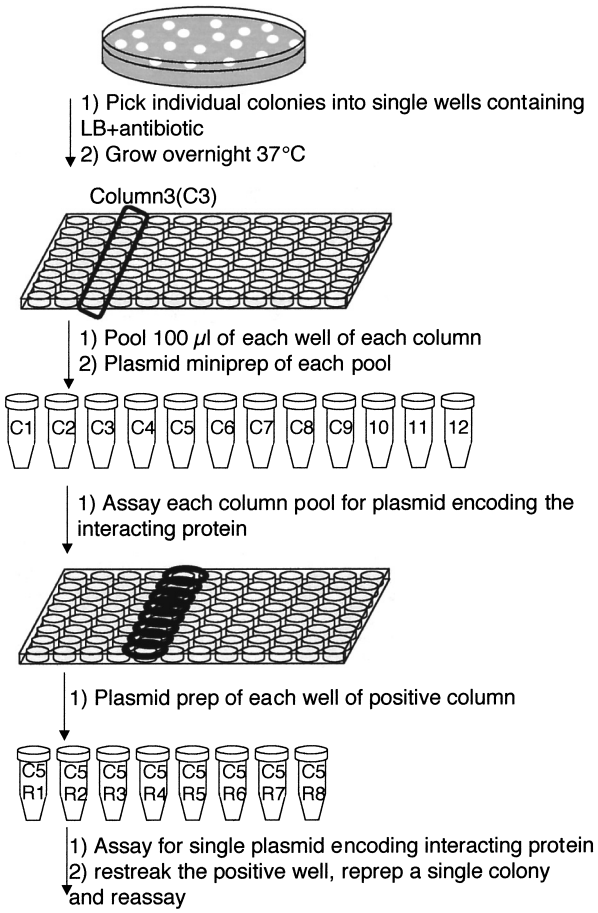

FIGURE 4. Strategy to isolate a single plasmid encoding the 14-3-3 interacting proteins from a positive pool of 100 plasmids. See text for details.

Stage 3: Sib Selection of Positive Pools

Once a cDNA pool containing a desired activity has been identified, it is progressively subdivided until a single cDNA clone containing that activity has been isolated (Fig. 4). The initial step is to retransform or replate the bacteria containing the cDNA pool to provide >100 colonies. Individual bacteria are grown in a 96-well plate format, and wells from individual columns and rows in the plate are pooled and re-screened to identify the isolated cDNA clone unambiguously.

1. To subdivide the pooled cDNAs, perform a 1:100–1:1000 dilution of the original miniprep DNA into sterile H_2O or a 10 mM Tris-HCl (pH 8.0) solution. Use 1 μl of this dilution to transform 50 μl of competent DH5α or XL2-Blue strains of *E. coli* by electroporation. The exact dilution of the library that is used depends on the transformation efficiency of the bacterial strain used.

2. Following electroporation, add 1 ml of LB medium and allow the transformed bacteria to recover at 37°C for 1 hour with shaking. Pellet the bacteria briefly in a microfuge, resuspend in ~100 μl of LB, and plate the entire suspension onto a 10-cm LB-agar plate with the appropriate antibiotic. In our experience, this typically yields 200–1000 colonies on the plate.

   Alternatively, it is acceptable to plate a small aliquot of the corresponding glycerol bacterial stock onto LB-agar plates containing the appropriate antibiotic.

3. After the plates are incubated overnight at 37°C, use an autoclaved toothpick to transfer a small sample of each colony to a single well of a 96-well flat-bottom plate containing 150 µl of LB medium with antibiotic. Avoid cross-contamination between wells.
4. Incubate the 96-well plate overnight at 37°C without shaking. Mix the cultures by repipetting, and remove 100-µl samples from each well in a column of the 96-well plate and pool in a separate bacterial culture tube. This yields 12 samples corresponding to each of the 12 columns on the 96-well plate, and each sample contains bacteria from eight different wells. Add 2 ml of LB containing antibiotic to each culture tube and shake the culture for several hours to get enough bacteria for plasmid purification. Then supplement the original cultures remaining in the 96-well plate (50 µl should be left in each well) with 50 µl of LB medium containing 40% glycerol and the appropriate antibiotic, mix by pipetting, and freeze at –80°C to provide glycerol stocks of the individual cDNA clones. The final glycerol concentration is 20%.
5. Isolate plasmid DNA from each of the 12 column pools, synthesize the proteins, and assay as described above. If no positives are detected, pick 96 more colonies and repeat the same procedure.
6. Further subdivide the positive column pools by individually preparing and testing the cDNA in each of the eight wells using the glycerol stock of the 96-well plate described above. Cross-contamination between wells can occur; therefore, it is critical to verify that the final observed activity is a consequence of the action of a single cDNA by restreaking the final clone on a fresh LB plate, picking a single clone, and verifying that it maintains the desired activity. Sequence the isolated cDNA using a T7 oligonucleotide primer to identify the protein of interest. Alternatively, a two-dimensional matrix can also be used to find a single active cDNA in one step (e.g., 12 column pools by 8 row pools to make 20 pools). However, because one can pick a positive clone more than one time (which confuses interpretation of a two-dimensional matrix), the outlined two-step procedure is often less work.

When to Use IVEC

IVEC has been most successful when used to isolate cDNAs for simple biochemical activities that were not possible by traditional methods. For example, before IVEC, effective genetic screens to identify kinase or protease substrates were elusive. A few practical considerations may help you decide whether IVEC is the correct approach for your application.

1. Because small pools of cDNAs are being expressed, it is unlikely that IVEC could identify activities that require more than one subunit being encoded by the library. To circumvent this limitation, other components or crude extracts to complement missing activities can be added.
2. IVEC may not be the best technique to screen for rare messages. In most of the screens performed to date, the rate-limiting step has been the running of SDS-PAGE gels, and it is difficult to electrophorese more than 50–100 pools of clones (1500–3000 proteins) per day. Thus, until good techniques are developed to normalize libraries, saturating screens are not practical. For smaller genomes (i.e., viruses, yeast, and worms) it would be ideal to prearray clones where each plasmid encoding cDNA is located in individual wells, which are then later pooled to simplify the translation. Once such a library is made, it allows rapid saturating IVEC screens.
3. There are presently at least three complementary techniques to identify interacting proteins, and all of them have distinct advantages and disadvantages. Protein purification is a powerful tool for identifying interacting proteins. Some advantages of protein purification are that it makes no assumptions about the binding conditions, and a whole native complex can be isolated. Biochemical purifications are technically difficult, and isolating the genes that encode these proteins is often cumbersome, requiring multiple steps of purification, peptide sequencing, and cDNA isolation. However, it is important to note that

direct complex purification has become more attractive, because each of these steps has been significantly simplified by recent technical advances such as epitope tagging proteins, mass spectrometry, and genome/expressed sequence tag (EST) sequencing projects.

The two-hybrid approach revolutionized the identification of protein–protein interactions by directly identifying the gene of an interacting protein. A second reason that two-hybrid systems are so widely used is that the techniques are well established and can be performed in most laboratory settings. IVEC also has both of these advantages. The two techniques are actually complementary because some of the limitations of two-hybrid systems are the strengths of IVEC. First, some important interactions may not occur in yeast nuclei because of cellular compartmentalization. IVEC interactions occur in cytoplasmic lysates that are cleared of compartmentalizing membranes. Two-hybrid screens cannot identify interactions that require more than two metazoan proteins. As discussed earlier, the experimental malleability of an in vitro reaction allows the easy addition of protein or small-molecule components. Third, two-hybrid interactions are prone to false positives, which must be excluded using the kind of interaction assay that is the basis of IVEC screening. Fourth, metazoan proteins may not be properly modified posttranslationally in budding yeast nuclei. Reticulocyte lysates most often generate well-folded proteins that are modified posttranslationally in a manner analogous to metazoan interphase cytoplasm. Thus, for metazoan systems the binding interactions using IVEC may actually occur at more physiological concentrations and conditions than in two-hybrid reactions.

14-3-3 binds phosphorylated proteins. Thus, the identification of 14-3-3 interacting proteins is an example of a screen that may have been difficult using typical two-hybrid techniques, because the "bait" proteins must be properly modified posttranslationally. In the screen presented here, we depended on the kinases in the reticulocyte lysate to phosphorylate the translated protein pools properly.

REFERENCES

Andresson T. and Ruderman J.V. 1998. The kinase Eg2 is a component of the *Xenopus* oocyte progesterone-activated signaling pathway. *EMBO J.* **17:** 5627–5637.

Cryns V.L., Byun Y., Rana A., Mellor H., Lustig K.D., Ghanem L., Parker P.J., Kirschner M.W., and Yuan J. 1997. Specific proteolysis of the kinase protein kinase C-related kinase 2 by caspase-3 during apoptosis. Identification by a novel, small pool expression cloning strategy. *J. Biol. Chem.* **272:** 29449–29453.

Kanai F., Marignani P.A., Sarbassova D., Yagi R., Hall R.A., Donowitz M., Hisaminato A., Fujiwara T., Ito Y., Cantley L.C., and Yaffe M.B. 2000. TAZ: A novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *EMBO J.* **19:** 6778–6791.

Kothakota S., Azuma T., Reinhard C., Klippel A., Tang J., Chu K., McGarry T.J., Kirschner M.W., Koths K., Kwiatkowski D.J., and Williams L.T. 1997. Caspase-3-generated fragment of gelsolin: Effector of morphological change in apoptosis. *Science* **278:** 294–298.

Li H., Zhu H., Xu C.-J., and Yuan Y. 1998. Cleavage of BID by caspase-8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell* **94:** 491–501.

Lustig K.D., Stukenberg P.T., McGarry T.J., King R.W., Cryns V.L., Mead P.E., Zon L.I., Yuan J., and Kirschner M.W. 1997. Small pool expression screening: Identification of genes involved in cell cycle control, apoptosis, and early development. *Methods Enzymol.* **283:** 83–99.

McGarry T.J. and Kirschner M.W. 1998. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93:** 1043–1053.

Mead P.E., Zhou Y., Lustig K.D., Huber T.L., Kirschner M.W., and Zon L.I. 1998. Cloning of mix-related homeodomain proteins using fast retrieval of gel shift activities (FROGS), a technique for the isolation of DNA-binding proteins. *Proc. Natl. Acad. Sci.* **95:** 11251–11256.

Stukenberg P.T., Lustig K.D., McGarry T.J., King R.W., Kuang J., and Kirschner M.W. 1997. Systemic identification of mitotic phosphoproteins. *Curr. Biol.* **7:** 338–348.