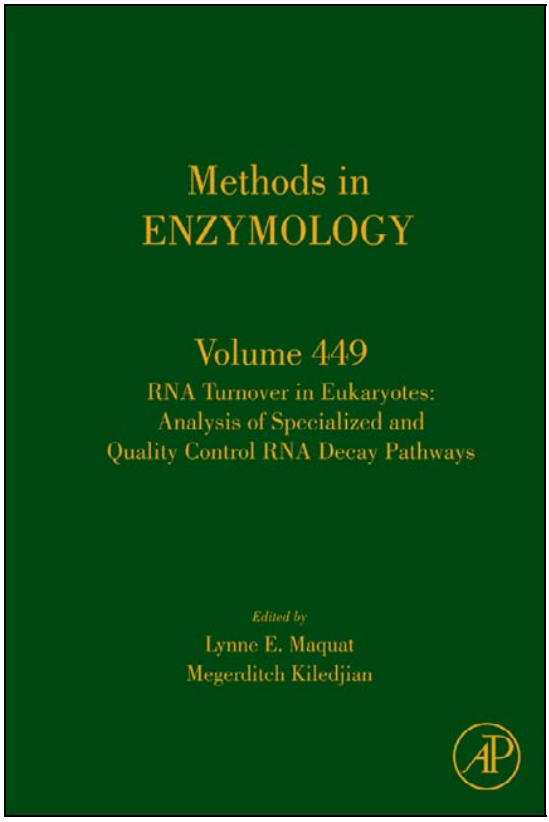

From: Craig R. Stumpf, Laura Opperman, and Marvin Wickens,
Analysis of RNA–Protein Interactions Using a Yeast Three-Hybrid System.
In Lynne E. Maquat and Megerditch Kiledjian, editors,
Methods in Enzymology, Vol 449,
Burlington: Academic Press, 2008, pp.295-315.
ISBN: 978-0-12-374584-2

CHAPTER FOURTEEN

Analysis of RNA–Protein Interactions Using a Yeast Three-Hybrid System

Craig R. Stumpf,[*,‡] Laura Opperman,[*,‡] *and* Marvin Wickens[†]

Contents

Abstract

RNA–protein interactions play an essential role in the maturation and regulation of RNAs within eukaryotic organisms. The three-hybrid system provides a simple,

* Program in Cellular and Molecular Biology, University of Wisconsin-Madison, Madison, Wisconsin, USA
† Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin, USA
‡ Authors contributed equally

Methods in Enzymology, Volume 449
ISSN 0076-6879, DOI: 10.1016/S0076-6879(08)02414-2

yet powerful means to study RNA–protein interactions within the eukaryote *Saccharomyces cerevisiae*. This chapter describes the basis of the system and applications in both examining specific RNA–protein interactions and screening libraries for novel interactions. We provide a detailed discussion on affinity versus reporter output, variations on library screening (e.g., randomization studies), some adaptations of the system, and updated reagents and protocols.

1. Introduction

The interactions between RNAs and proteins are critical for a wide variety of biological processes, ranging from development to memory to viral proliferation. For this reason, biochemical and genetic assays have been developed to analyze them. This chapter focuses on one genetic method, the yeast three-hybrid system (SenGupta *et al.*, 1996).

The yeast three-hybrid system monitors RNA–protein interactions by scoring reporter gene output. It allows analysis of an interaction independent of biological function. Like other genetic strategies, the three-hybrid system yields a clone of the RNA or protein of interest during the screen, establishing its identity early on. The system offers the possibility of connecting RNAs and proteins on a broad, and perhaps, genomic scale. The published applications of the method to date include discovery of proteins that bind to a known RNA sequence, identification of RNA sequences that bind known RNA-binding proteins, confirmation of suspected interactions between an RNA and protein, mutational analysis of interacting RNAs and proteins, and the discovery and analysis of multiprotein–RNA complexes.

This chapter describes the architecture of the system and summarizes its utility in identifying and analyzing RNA–protein interactions. We first present broad perspectives and background and then focus sequentially on specific applications. The utility of the system has benefited from adaptations developed by multiple laboratories over the past few years. This chapter focuses on recent developments and directs the reader to previous reviews that have provided other protocols in detail (Bernstein *et al.*, 2002; Kraemer *et al.*, 2000). Earlier reviews also include lists of specific published applications and citations (Bernstein *et al.*, 2002; Kraemer *et al.*, 2000).

In the spirit of the ''what's new'' blurbs that accompany software releases, we point out the following new or expanded aspects of the three-hybrid system since our last review (Bernstein *et al.*, 2002) in this series: a vector (p3HR2) useful in screening RNA libraries; a yeast strain (YBZ-1) that simplifies screening; the quantitative relationship between affinity and reporter output; the use of randomized libraries to study RNA–protein interactions; and updated protocols.

2. Principles of the Method

The general strategy of the three-hybrid system is shown in Fig. 14.1. DNA-binding sites are placed upstream of a reporter gene, which has been integrated into the yeast genome. The first hybrid protein consists of a DNA-binding domain linked to an RNA-binding domain. The RNA-binding domain interacts with its RNA-binding site in a bifunctional ("hybrid") RNA molecule. The other part of the RNA molecule interacts with a second hybrid protein consisting of another RNA-binding domain linked to a transcription activation domain. When this tripartite complex forms at the promoter, the reporter gene is turned on. Reporter expression can be detected by phenotype or simple biochemical assays. The specific molecules used most commonly for three-hybrid analysis are depicted in Fig. 14.1. The DNA-binding site consists of a 17-nucleotide recognition site for the *Escherichia coli* LexA protein and is present in multiple copies upstream of both the *HIS3* and the *lacZ* genes. The first hybrid protein consists of LexA fused to bacteriophage MS2 coat protein, a small polypeptide that binds to a short stem-loop sequence present in its RNA genome. The hybrid RNA (depicted in more detail in Fig. 14.2) consists of two MS2 coat protein-binding sites linked to the RNA sequence of interest, X. The second hybrid protein consists of the transcription activation domain of the yeast Gal4 transcription factor linked to an RNA-binding protein, Y.

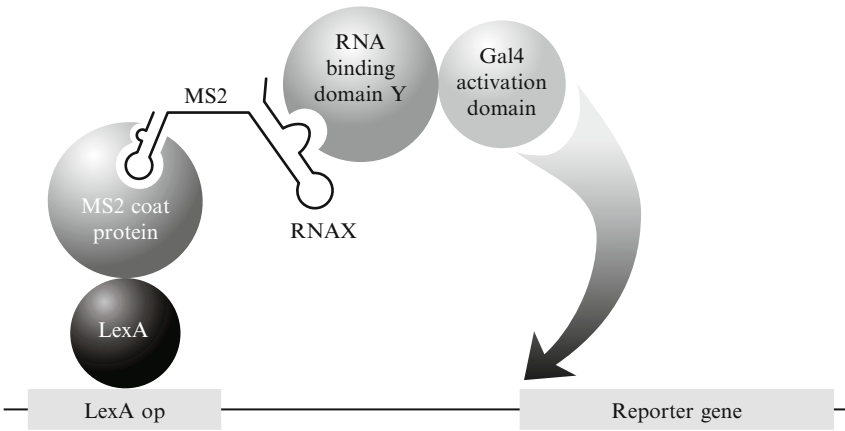

Figure 14.1 Three-hybrid system used to detect and analyze RNA–protein interactions. The diagram depicts the general strategy of the three-hybrid system. Specific protein and RNA components that are typically used are represented. For the sake of simplicity, the following points are not illustrated. In the strains L40coat and YBZ-1, both *lacZ* and *HIS3* reporter genes are present under the control of lexA operators (eight in the *lacZ* promoter and four in the *HIS3* promoter). The LexA protein binds as a dimer. The hybrid RNA contains two MS2 stem-loops and the MS2 coat protein binds one stem-loop as a dimer.

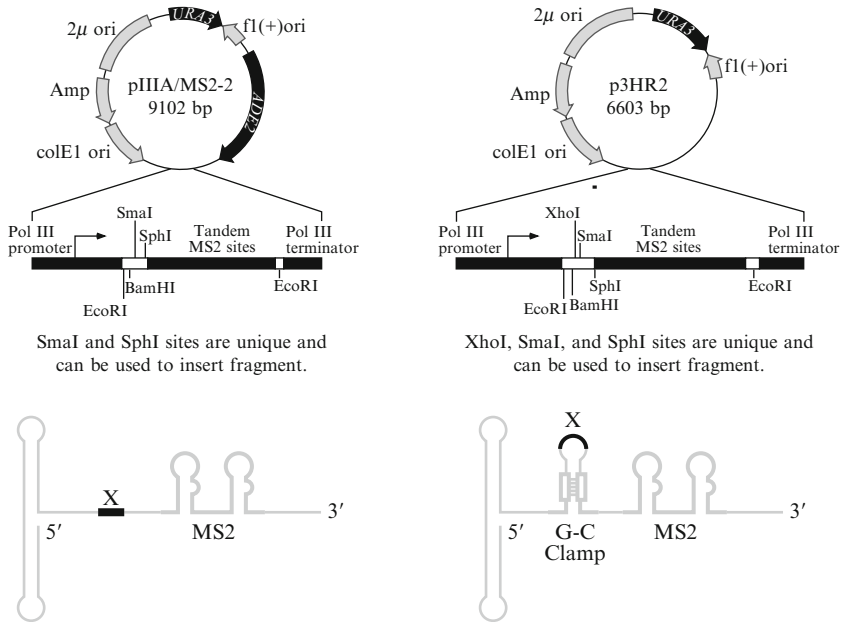

Figure 14.2 Plasmids for the expression of hybrid RNAs. p3HR2 displays the RNA sequence of interest in the loop of a G-C clamp, whereas pIIIA/MS2-2 does not (Zhang *et al.*, 1997). The sequence of interest can be inserted into the unique restriction sites. See text for more details.

By introducing libraries of RNA or protein, cognate partners can be identified. As in two-hybrid screens, the challenge then becomes to identify those molecules whose interaction is biologically relevant. For this purpose, mutations in the known RNA or protein component are very useful. The system also makes it possible to identify regions of an RNA or protein required for a known interaction and to test combinations of RNA and protein to determine whether they interact *in vivo*. In addition to the constraints of the two-hybrid system, the three-hybrid system has its own limitations, such as lower signal output and the efficient expression of RNA molecules, which are addressed later.

3. Key Components: RNAs, Vectors, and Strains

3.1. Hybrid RNAs

The RNA–protein interaction used to tether the hybrid RNA to the promoter must be specific and of high affinity. The RNA need not be structured. Several different interactions have been used; of these, by far, the

most common has been that between a short stem-loop RNA structure and bacteriophage MS2 coat protein. A few features of this interaction make it attractive: the interaction between the MS2 coat protein and its stem-loop binding site has an affinity in the subnanomolar range (Lowary and Uhlenbeck, 1987). This interaction has been the focus of intensive biochemical studies. The MS2 coat protein is a relatively small polypeptide that, along with the MS2 stem-loops, has been used in a variety of applications (Beach *et al.*, 1999; Bertrand *et al.*, 1998; Coller and Wickens, 2002).

RNA sequences to be tested can be inserted into the multiple cloning site of a high-copy RNA expression vector. The hybrid RNA molecule is transcribed from these plasmids by RNA polymerase III, using the RNase P promoter. The transcript is composed of the RNase P 5′ leader, the sequence of interest, two MS2 stem-loops, and the RNase P 3′ termination sequence. Use of the high copy plasmid, combined with the RNase P promoter, results in abundant expression of the hybrid RNA. Higher levels of RNA lead to increased expression of the reporter genes.

The relative orientation of the RNA sequence of interest and the MS2 stem-loops can affect signal output. In the few cases that have been tested systematically, both orientations yield activation and are specific. However, in some instances, placing the RNA of interest upstream of the MS2 sites results in a slightly higher level of transcription than the alternative arrangement. Although RNA folding programs can be used to predict whether one arrangement is more likely to succeed, the accuracy of their predictions *in vivo* is problematic. In most cases, we have placed the RNA sequence of interest upstream of the MS2 sites, as occurs in the pIIIA/MS2-2 and p3HR2 vectors depicted in Fig. 14.2. This orientation assures that any RNA containing the MS2 stem-loops also contains the RNA of interest. Published reports have used both orientations successfully.

RNA sequences to be analyzed are restricted in two respects at present, although neither restriction is absolute. First, runs of four or more uridines in succession can terminate transcription by RNA polymerase III. Second, RNA inserts of lengths less than ≈200 nucleotides typically yield higher signals; longer inserts commonly reduce the level of reporter activation.

Four or more uridines in succession can function as an RNA polymerase III terminator, and thus can prevent production of the desired hybrid RNA. The efficiency of termination at oligouridine tracts is context dependent. Northern blotting can be performed to determine if the hybrid RNAs are expressed at high levels. Termination in long runs of U's can be perturbed by appropriately placed single nucleotide substitutions. However, if little information exists about a protein–RNA interaction, the effect of single nucleotide substitutions on the RNA–protein interaction should be determined independently.

The size of the RNA insert appears to be an important determinant of three-hybrid activity. In reconstruction experiments using known

RNA–protein partners, RNA sequences that are less than 200 nucleotides in length typically yield substantial and specific reporter activation. Nevertheless, substantial signals can be detected from longer RNAs. For example, Rho and Martinis (2000) used a 1600-nucleotide RNA containing the yeast *cobI* intron to detect interactions with two different proteins.

3.2. Plasmids

3.2.1. Plasmids encoding the hybrid RNA

Several plasmids have been constructed to express hybrid RNA sequences (Bernstein *et al.*, 2002; SenGupta *et al.*, 1996); two are depicted in Fig. 14.2. Both plasmids are multicopy with origins of replication and selectable markers for propagation in yeast and bacteria. Additional information is available at Wickens laboratory Web page link: http://www.biochem.wisc.edu/faculty/wickens/lab/3h.aspx.

3.2.1.1. pIIIA/MS2-2 (Zhang et al., 1997) This is a yeast shuttle vector derived from pIIIEx426RPR (Good and Engelke, 1994). Sequences to be analyzed are inserted into the *Sma*I/*Xma*I and/or *Sph*I sites. This plasmid carries both *ADE2* and *URA3* markers for yeast selection and produces hybrid RNAs from the yeast RNase P RNA (*RPR1*) promoter, an RNA polymerase III promoter. The *RPR1* promoter is useful for two reasons. First, it is efficient, directing the synthesis of more than 1000 molecules per cell. Second, the transcripts produced from this promoter presumably do not enter the pre-mRNA processing pathway and may not leave the nucleus. The *ADE2* marker makes it possible to monitor the presence of the plasmid by colony color, which can be useful in screens of cDNA libraries.

Variants of the pIIIA/MS2-2 vector alter the orientation of the restriction sites and the MS2 binding sites, while others do not contain the *ADE2* marker (Bernstein *et al.*, 2002; SenGupta *et al.*, 1996).

3.2.1.2. p3HR2 (Stumpf et al., 2008) This vector was derived from pIIIA/MS2-2. The *ADE2* gene has been removed to reduce the size of the vector, thus making manipulations easier. A GC clamp has been inserted such that the restriction sites are placed within the loop region similar to what has been reported by Maher and colleagues (Fig. 14.2; Cassiday and Maher, 2001). A sequence of interest cloned into the *Xho*I, *Xma*I/*Sma*I, and/or *Sph*I sites will be presented within the loop region, thus helping prevent the formation of inhibitory structures with the MS2 sites and/or RNase P RNA 5′ leader. This arrangement is especially helpful in dealing with RNA libraries when one cannot predict if the arrangement is likely to form an inhibitory structure.

3.2.2. Plasmids encoding the activation domain fusion

Any activation domain (AD) fusion vector with a compatible marker (e.g., *LEU2, TRP1*) will work in the system. Most commonly, the plasmids used are multicopy with a strong promoter driving transcription of the fusion protein. Several AD fusion plasmids are in common use from a variety of sources.

The abundance of the AD-fusion protein influences the output of the reporter gene and is an important issue to consider. The relationship between the concentration of the activation domain-containing protein and the affinity of the interaction has been examined systematically. Not surprisingly, the lower the K_d, the less protein is required to detect the interaction. Beyond a certain protein concentration, the reporter gene output plateaus (Fig. 14.3; Hook *et al.*, 2005). The concentration at which the plateau occurs is a function of the K_d of the interaction (Hook *et al.*, 2005).

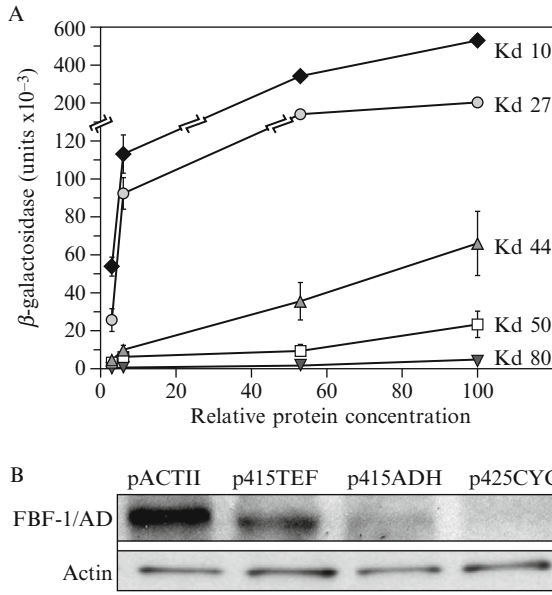

Figure 14.3 Reporter output as a function of protein abundance. (A) Graph indicates relationship between relative protein concentration (as determined by quantitative Western blotting of FBF-1/AD fusion protein) and β-galactosidase activity (relative light units; Hook *et al.*, 2005). This relationship was examined for RNAs having different affinities for *C. elegans* FBF-1 (Hook *et al.*, 2005). RNA6 (◆), RNA5 (○), RNA3 (▲), RNA2 (□), and RNA1 (▼) correspond to the RNAs described in Fig. 14.4. (B) Western blot depicting the expression of a Gal4 AD-FBF-1 fusion protein expressed from various promoters (Hook *et al.*, 2005). Actin was used as a loading control.

3.3. Yeast strains

The yeast reporter strain L40coat (SenGupta *et al.*, 1996) is derived from L40-ura3 (a gift of T. Triolo and R. Sternglanz, Stony Brook). The genotype of the strain is *MATa*, *ura3-52*, *leu2-3*, *−112*, *his3-200*, *trp1-1*, *ade2*, *LYS2*::(*lexA op*)-*HIS3*, *URA3*::(*lexA op*)-*lacZ,* and *LexA MS2 coat* (*TRP1*). The strain is auxotrophic for uracil, leucine, adenine, and histidine. Each of these markers can be exploited in the three-hybrid system. Both the *HIS3* and the *lacZ* genes have been placed under the control of the lexA operators, and hence are reporters in the three-hybrid system. A gene encoding the LexA–MS2 coat protein fusion has been integrated into the genome.

Strain R40coat is identical to L40coat, but of α mating type.

Strain YBZ-1 is a derivative of L40coat of genotype *MATa*, *ura3-52*, *leu2-3*, *−112*, *his3-200*, *trp1-1*, *ade2*, *LYS2*::(*LexAop*)-*HIS3*, *URA3*::(*lexA-op*)-*lacZ*, and *LexA-MS2 MS2 coat* (*N55K*; Hook *et al.*, 2005). It contains a MS2 coat protein mutant (N55K) that results in a 10-fold increase in the affinity of the RNA–protein interaction (Lim *et al.*, 1994). The coat protein mutants are placed in tandem, creating an intramolecular coat protein dimer that enhances binding to the MS2 stem-loops.

A direct comparison of the yeast strains L40coat and YBZ-1 has been performed in screens using a specific *Caenorhabditis elegans* RNA and a cDNA–AD library (Hook *et al.*, 2005). The YBZ-1 strain increased the number of genuine positives identified and, at the same time, decreased the number of false positives detected, as compared to L40coat (Hook *et al.*, 2005). This increased efficiency in screening cDNA libraries can be attributed to the tandem, head-to-tail dimer of the high-affinity variant of MS2 coat protein, N55K (Hook *et al.*, 2005). While the Hook *et al.* (2005) study was done with just one RNA–protein interaction, YBZ-1 generally enhances the efficiency of three-hybrid screening.

4. Methodology

Assaying the interaction between an RNA and a protein is straightforward whether performing a directed test or screening a library. The yeast strain with the integrated reporter gene and LexA-MS2 coat protein is transformed with two plasmids: an activation domain plasmid carrying the protein(s) of interest and a hybrid RNA plasmid carrying the desired RNA sequence(s). If the two molecules interact, the reporter gene is expressed. *HIS3* and *lacZ* are the most common reporters utilized. Control experiments are used to demonstrate that each component of the system is required for the assayed interaction.

4.1. Assaying interactions: β-Galactosidase activity and 3-aminotriazole resistance

The "strength" of an interaction is gauged by assaying the activity of a reporter gene. Typically, this is done either by assaying β-galactosidase activity or by determining the level of resistance to 3-aminotriazole (3-AT), which monitors His3 activity. These assays are described next.

4.2. Qualitative assays

Qualitative assays can be performed with both the *HIS3* and the *lacZ* reporter genes. β-Galactosidase activity can be assayed by measuring the conversion of a lactose analog to a chromogenic or luminescent product. This assay can be performed using either colonies permeabilized on a filter or a cell lysate. The filter assay yields qualitative results, while the liquid assay is more quantitative.

The filter assay is especially useful as a secondary selection of positive colonies from a three-hybrid screen using *HIS3* as the initial reporter output. In the filter assay, positives are replica plated onto a nitrocellulose filter and, after a day of growth, the β-galactosidase activity is measured. This use of the filter assay as a secondary selection removes positives that only activate one of the two reporters; these are likely to not represent real RNA–protein interactions. For a detailed protocol, see Bernstein *et al.* (2002).

The activity of the *HIS3* reporter gene can be measured in a qualitative growth assay by including the competitive inhibitor 3-AT in the medium. This is done during screens to help eliminate "false positives." In directed tests of particular protein–RNA interactions, yeast strains carrying the molecules of interest can be plated on medium containing various amounts of 3-AT. A common range includes 0.1–100 m*M* 3-AT, depending on the ability of the protein–RNA interaction to activate the *HIS3* reporter gene. For a detailed protocol, see Bernstein *et al.* (2002).

4.3. Quantitative assays

4.3.1. Relationship between reporter gene activity and affinity

Quantitative analysis of *lacZ* reporter gene expression can be measured by assaying β-galactosidase enzyme activity. This is particularly useful when analyzing the interaction between a specific protein and the RNA sequence.

The level of β-galactosidase activity has been examined as a function of *in vitro* affinity using a specific RNA–protein interaction (Fig. 14.4; Hook *et al.*, 2005). When the protein FBF-1 was tested with various RNAs,

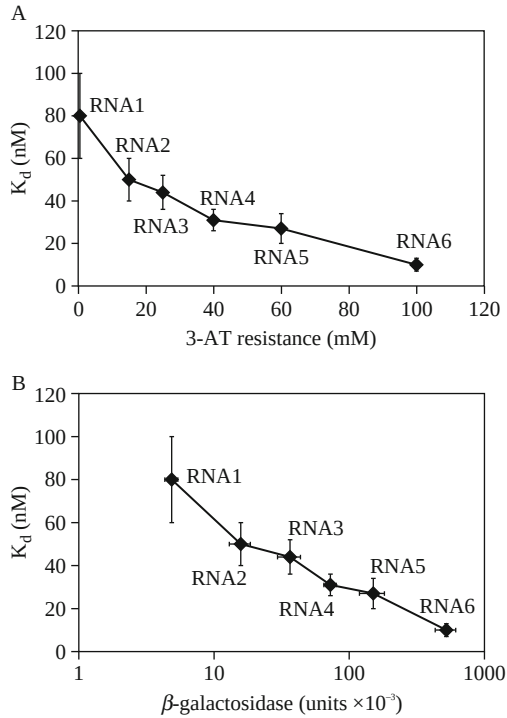

Figure 14.4 Direct relationship between affinity and reporter gene activation. RNA sequences (numbering correlates with Fig. 14.3) with increased affinity for *C. elegans* FBF-1 are capable of greater reporter output in the three-hybrid system (Hook *et al.*, 2005). (A) 3-AT resistance increases as K_d decreases. (B) β-Galactosidase activity (relative light units) increases as K_d decreases.

K_d was linearly related to the log of β-galactosidase activity over the range measured, from 1 to 100 n*M* (Fig. 14.4; Hook *et al.*, 2005).

4.3.2. Protocol for quantitative β-galactosidase assays

Numerous variations on the liquid β-galactosidase assay exist. The protocol described here uses the enzyme-coupled luminescent substrate Beta-glo (Promega Corporation; Hook *et al.*, 2005). This simple yet sensitive assay uses a luminometer to measure the output from the *lacZ* gene. The protocol described here uses a Turner 20/20n luminometer (Promega Corporation). Certain details, such as sample volumes, will vary depending on the instrument used.

1. Inoculate 5 ml cultures of selective medium in triplicate for each interaction to be tested. Grow to an OD_{660} of 0.1–0.2. Alternatively, grow cultures to saturation, dilute into fresh medium 1:40 (100 μl into 4 ml

selective media), and grow for about 2–2.5 h. Check to ensure an OD_{660} of 0.1–0.2.
2. Mix 50 μl of cells with 50 μl Beta-glo reagent. Vortex well for complete lysis.
3. Incubate for 1 h at room temperature in the dark.
4. Transfer 10 μl of cell:Beta-glo reagent mix to an appropriate tube.
5. Integrate the luminescent signal for 1 s.
6. Normalize signal to cell number to yield an activity/cell value.

5. Analyzing Known RNA–Protein Interactions

Assaying the interaction between a specific RNA and protein is straightforward. A yeast strain with the integrated reporter gene and LexA-MS2 coat protein is transformed with two plasmids: an activation domain plasmid containing the protein of interest and a hybrid RNA carrying the desired RNA sequence. If the protein and RNA interact, the reporter gene is expressed. *HIS3* and *lacZ* are commonly used as reporters; they can be assayed as described earlier. Control experiments can be used to demonstrate that the interaction is specific.

The three-hybrid system has been applied to known protein–RNA interactions for a variety of purposes (Bernstein *et al.*, 2002). For some applications (e.g., examining the effect of mutations on affinity), the correlation between reporter output and relative K_d is relevant (Fig. 14.4; Hook *et al.*, 2005). The relationship between these two varies as a function of the abundance of the protein (Hook *et al.*, 2005). Since abundances of AD fusion proteins vary, the range in which reporter gene output changes vs K_d also can differ.

Mutational analysis of the protein and RNA is relatively straightforward. For instance, protein mutants with altered or relaxed binding specificities can be identified (Martin *et al.*, 2000; Opperman *et al.*, 2005). Mutational analysis of the RNA can be performed to better understand RNA-binding specificity. Bernstein *et al.* (2005) undertook a detailed mutational analysis of two biological mRNA targets of the RNA-binding protein FBF. This allowed for the development of a consensus-binding element and the subsequent identification of putative biological targets.

This general strategy can be extended to select interaction-competent or -incompetent variants of the RNA or protein using randomized libraries. Selection experiments of this type can be used to identify specific amino acids or nucleotides that mediate interactions. Early applications include the identification of mutations in *Drosophila* Pumilio that disrupt its RNA-binding activity and analysis of the stem-loop target of SLBP (Edwards *et al.*, 2001; Martin *et al.*, 2000).

5.1. Assay limitations

Limitations of the three-hybrid system need to be taken into account whether performing directed tests or a screen. These limitations are the identification of "false positives" and "false negatives."

"False negatives" usually occur when either the RNA hybrid or the AD fusion is not expressed well or folded improperly. Alternatively, interference by endogenous yeast factors could affect the levels of either the RNA hybrid or the AD-fusion protein available for the assay.

"False positives" can be separated into two categories. The first refers to activation of a reporter in the absence of an interaction that requires all of the components of the three-hybrid system. This is often called "autoactivation." Either the RNA sequence or the protein being tested can cause autoactivation. The RNA sequence can do so by interacting with the transcription machinery while tethered to the promoter. The protein can do so by interacting directly with the promoter or the LexA-MS2 coat protein fusion. A second category of "false positives" consists of those interactions that occur in the three-hybrid system, but are not biologically relevant. There are two groups of false positives of this type.

1. *High-affinity, nonspecific interactions*. In principle, some proteins may interact with a restricted set of RNAs in the cell, but bind to many more RNAs in the three-hybrid system. Few such examples have been reported, but it seems likely they exist since the three-hybrid system measures RNA–protein interactions outside of their biological context. However, the success of three-hybrid screens of cDNA libraries with specific regulatory elements as "bait" tends to suggest that few proteins have nonspecific affinities (at least for these sequences) in the nanomolar range. Obviously, the mere fact that a protein binds to many RNA sequences with reasonable affinity does not preclude their being real targets *in vivo*.
2. *Bridges*. The protein that appears to interact with the RNA may in fact interact with a yeast cellular protein that binds to the RNA. This obviously is most germane when the activation domain hybrid is a yeast protein. For example, yeast She3 registers as binding to a portion of the 3′-UTR of the *ASH1* mRNA in the three-hybrid assay, but does so because it interacts with cellular She2, which is bound to the hybrid RNA (Long *et al.*, 2000).

6. Three-Hybrid Screens to Identify RNA–Protein Interactions

6.1. Types of screens

The three-hybrid system can be used to identify both proteins that bind a specific RNA sequence and RNA sequences that bind specific proteins. The methods for doing these screens are basically the same, with the

difference being which molecule is the "bait" and which is part of a library (the "prey"). Typically the plasmid containing the "bait" is transformed into yeast. Subsequently, a library containing the putative interacting molecules is transformed into that strain and positive interactions are identified by growth on selective media. These colonies are then selected for further scrutiny.

The first applications of the three-hybrid system involved the identification of proteins that bind a specific RNA sequence (Martin *et al.*, 1997; Wang *et al.*, 1996; Zhang *et al.*, 1997). Many such screens of this type have been reported [for a list as of 2002, see Bernstein *et al.* (2002)].

A second application seeks to identify the RNA partner(s) of a known protein. These screens complement biochemical techniques in identifying RNA–protein interactions (e.g., SELEX and coimmunoprecipitation experiments). The advantage of the yeast-three hybrid screen is, again, that the interactions occur within a living cell. Studies utilizing this technique have identified RNAs bound by a splicing factor and a PUF protein (Seay *et al.*, 2006; Sengupta *et al.*, 1999; Zhang *et al.*, 1997). Further work has utilized this technique to identify the consensus binding element of specific RNA-binding proteins (Bernstein *et al.*, 2005; Opperman *et al.*, 2005).

6.2. *A priori* considerations

Two general concerns that must be addressed when performing a three-hybrid screen are the stringency of the selection and the identity of the library. The level of selection is determined by the concentration of 3-AT in the medium. Prior to performing a screen, the level of selection should be determined based on prior knowledge of the molecules being tested or by determining the level of 3-AT required to eliminate the growth of any "background" colonies.

6.2.1. Stringency

The primary complication in interpreting results from a three-hybrid screen is identifying and discarding, "false positives." False positives usually arise from the "prey" molecule activating transcription of the reporter in the absence of its binding partner. This can occur when a protein fused to the activation domain can interact with the MS2 coat protein, the LexA protein, or the promoter directly. Alternatively, an RNA sequence tethered at the promoter can recruit RNA polymerase and initiate transcription of the reporter (Buskirk *et al.*, 2003; Saha *et al.*, 2003; Sengupta *et al.*, 1999). These colonies are discarded by testing for the activation of the reporter in the absence of the "bait" plasmid.

The addition of 3-AT to the media decreases the number of false positives by demanding a more stringent interaction between the RNA and the protein. However, 3-AT concentrations that are too high can result

in the loss of colonies containing legitimate RNA–protein interactions. Thus the concentration of 3-AT used in the screen is instrumental in mediating the balance between a high background (many false-positive colonies) and missing real interactions (false-negative results).

6.2.2. Libraries

The primary limitation in performing a three-hybrid screen is the population of unique plasmids represented in the library. Screening a library composed of more than 20 million plasmids is difficult using current techniques. The library selected to analyze in the screen will depend on the purpose for doing the screen. To identify proteins that bind to a particular RNA sequence, many activation domain fusion libraries exist, which have been used successfully in both two-hybrid and three-hybrid screens. Alternatively, a cDNA library can be cloned from a tissue of interest. To identify RNAs that interact with a protein of interest, one can clone fragments of a genome or expressed sequence tags from a tissue of interest.

As mentioned earlier, libraries containing randomized portions of a molecule of interest can be used to dissect a known interaction. To prepare a random RNA library, a randomized double-stranded oligonucleotide can be inserted into the RNA expression vector. This library can then be used in a screen to select for those RNA sequences that bind the protein of interest. Two simple methods can be used to prepare random mutations in the RNA-binding protein component. First, one can perform error-prone polymerase chain reaction (PCR) on a region of (or the entire) the gene. This fragment can then be inserted into the AD fusion vector. Alternatively, a random oligonucleotide can be inserted into a region of the protein. These mutant protein libraries can be screened against an RNA of interest to identify amino acids that are important for binding (Edwards *et al.*, 2001). After libraries are constructed, it is generally useful to amplify the library in *E. coli* or phage. Individual clones can then be sequenced to evaluate the diversity of the library. These libraries can then be used in a three-hybrid screen. Alternatively, libraries can be transformed into yeast directly following construction; however, this technique is inefficient.

In all screening applications, it is important to realize that the concentration of 3-AT used in the *HIS3* selection effectively sets a bar for K_d: interactions that are too weak register as negative, whereas interactions that are strong enough may not be discriminated in the first pass. The importance of this is that one may identify the tightest binding partner, yet other partners may bind with only slightly lower affinity.

Once these initial concerns have been addressed, the execution of a three-hybrid screen is relatively simple. A flowchart describing a typical screen is depicted in Fig. 14.5.

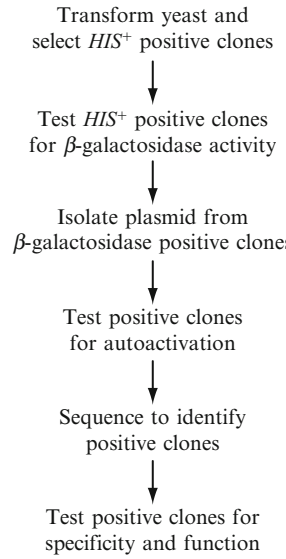

Figure 14.5 General scheme depicting a three-hybrid screen to detect RNA–protein interactions. See text for more details.

7. The Three-hybrid Screen: A General Protocol

Step 1. Transform yeast and select for growth

Yeast can be transformed with both plasmids simultaneously or the "bait" plasmid can be transformed first, followed by the library plasmid. Usually, transforming the plasmids sequentially yields more transformants. However, occasionally the "bait" plasmid is toxic to the yeast, requiring a cotransformation of the plasmids. The transformation is then plated on medium lacking histidine and containing the predetermined concentration of 3-AT. We have found that concentrations of 3-AT in the range of 1 to 5 m*M* are a reasonable starting point, but an initial titration of 3-AT in the absence of any "prey" is the ideal situation. We generally aim for a concentration of 3-AT that prevents growth of a yeast strain containing the "bait" plasmid, but lacking any "prey" plasmid.

Step 2. Assay β-galactosidase activity

As a secondary selection, the level of β-galactosidase in the cells is determined by performing a qualitative β-galactosidase assay. This protocol has been detailed previously (Bernstein *et al.*, 2002). Briefly:

1. Patch yeast onto nitrocellulose filters and grow overnight.
2. Lyse yeast by submersion in liquid nitrogen.
3. Expose yeast to X-gal at 30 °C for 30–60 min.
4. β-Galactosidase levels are indicated by the presence of a blue colony.

We have found that the majority of *HIS3+* colonies also express β-galactosidase. This assay is useful in weeding out any colonies that may have acquired an assay-independent mechanism for expressing *HIS3* during the course of the yeast transformation.

Step 3. Isolate plasmids

Once interacting colonies have been isolated, the ''prey'' plasmid from those colonies must be isolated for further analysis. Several methods have been developed for the isolation of plasmid DNA from yeast (Bernstein *et al.*, 2002; Robzyk and Kassir, 1992). A sample protocol using the Promega Wizard Plus SV Miniprep Kit is presented here.

1. Inoculate 5 ml of selective media and grow overnight.
2. Pellet 2 ml of yeast at 5000*g*.
3. Resuspend in 250 μl of cell resuspension buffer.
4. Spheroplast yeast with 5 μl Zymolase (Genotech) for 2 h at 37 °C.
5. Add 250 μl of lysis buffer and invert to mix.
6. Incubate for 5 min at room temperataure.
7. Incubate for 5 min at 65 °C. Cool to room temperature.
8. Add 10 μl of alkaline protease solution.
9. Incubate for 10 min.
10. Add 350 μl of neutralization solution. Invert to mix.
11. Centrifuge at >16,000*g* for 10 min.
12. Add lysate to binding column.
13. Centrifuge at >16,000*g* for 1 min.
14. Wash with 700 μl of wash buffer.
15. Wash with 500 μl of wash buffer.
16. Discard liquid and centrifuge at >16,000*g* for 5 min.
17. Elute in 100 μl of water.

Step 4. Test for bait dependence (autoactivation)

Once plasmid is purified from yeast, it can be introduced into *E. coli* using conventional methods. Since either ''bait'' or ''prey'' plasmids can be present in *E. coli*, colonies can be screened for the plasmid of interest by colony PCR. In some cases, different antibiotic resistance markers may be present on each plasmid used in the screen, which makes isolation of the ''prey'' plasmids more direct. The plasmids can then be amplified in *E. coli* for use in future applications.

"Prey" plasmids must be tested to determine if they activate the reporter genes in the absence of the "bait" plasmid. This is done by transforming each "prey" plasmid back into a yeast strain containing an empty "bait" vector. This strain can then be assayed for the expression of either the *HIS3* reporter or the β-galactosidase reporter as described previously.

Step 5. Identify positive clones

Once the plasmids are isolated and have been determined to be dependent on the "bait" plasmid to activate the reporters, they can be sequenced to determine their identity. Screens to identify protein partners for specific RNAs yield sequences that can be identified by comparison to common sequence databases. Screens designed to identify RNA sequences that interact with a particular protein yield sequences that can be analyzed by comparison to a sequence database or aligned and used to derive a consensus-binding element for the protein of interest.

Step 6. Determine binding specificity using mutant and control molecules

Clones that are dependent on the "bait" plasmid to activate the reporter genes can then be analyzed to determine if they interact specifically with the molecule of interest. Molecules used in specificity testing will depend on the nature of the individual screen. However, some examples are discussed here. The ideal scenario is to have small, directed (point) mutations in the "bait" molecule (perhaps disrupting an RNA-binding domain of a protein or mutating a nucleotide of known importance in an RNA), which disrupt the interaction between that molecule and its interacting partner. In this case, it may be helpful to know the identity of the candidates identified in the screen and use this information to guide mutagenesis studies. If no subtle mutations are available, even rudimentary analyses (such as using antisense RNA or an unrelated protein) can be informative.

Step 7. Functional tests or additional screens

Almost invariably, additional steps will be needed to identify those positives that are biologically meaningful. As stated earlier, each screen is unique. The interactions being analyzed and the organisms being studied will determine what additional steps need to be taken to determine the biological relevance of each interaction. It is not surprising, given that the assay is performed outside of most biological contexts, that some specific, high-affinity interactions may not be relevant to the biology of the system being studied.

8. Other Applications of the Three-Hybrid System

The three-hybrid system has spawned several unexpected applications. We mention a few of them to encourage others to exploit the versatility of the system.

8.1. RNA activators

In the first three-hybrid screen using a library of RNAs, Sengupta *et al.* (1999) discovered that certain RNAs tethered to a promoter activated transcription on their own, that is, without a cognate AD-protein fusion. These RNAs, while unwanted when studying RNA–protein interactions, have elicited considerable interest, as has the general selection scheme. In parallel to Sengupta *et al.* (1999), Saha *et al.* (2003) used the three-hybrid system to identify RNA molecules that can work as transcriptional activators when tethered to DNA. The three-hybrid system also has been used for *in vivo* evolution of RNA-based transcriptional activators and silencers (Buskirk *et al.*, 2003; Kehayova and Liu, 2007), some of which are ligand dependent (Buskirk *et al.*, 2004).

8.2. Examination of RNA aptamers

The system can be used to identify aptamers that are more effective in binding their cognate proteins *in vivo*. Cassiday and Maher (2003) optimized the ability of an RNA aptamer to block DNA binding by NF-κB. *In vitro* evolution was followed by use of the three-hybrid system to select an RNA aptamer with high affinity for NF-κB (Cassiday and Maher, 2003). The selected aptamers repress NF-κB *in vivo* (Cassiday and Maher, 2003). These studies demonstrate that the three-hybrid system can aid development of novel regulatory molecules.

8.3. Multiprotein complexes

The formation of biologically relevant multiprotein–RNA complexes has also been demonstrated in the three-hybrid system (Bieniasz *et al.*, 1998, 1999; Long *et al.*, 2000; Sonoda and Wharton, 1999, 2001). Screens have been performed to identify components of ribonucleoprotein particles (RNPs; Bouffard *et al.*, 2000). Further adaptations can readily be envisioned. One possible modification would allow for the identification of proteins and/or RNAs that either enhance or inhibit the formation of larger complexes. The effect of various components within an RNP on binding

affinity or specificity could also be analyzed. Along similar lines, adaptations of the three-hybrid system would allow for the study of the effect of biologically or clinically relevant compounds, which may influence the formation of RNA–protein complexes.

9. Concluding Remarks

The three-hybrid system has been used to analyze a wide variety of known or suspected interactions, to identify new RNA–protein interactions, and has been extended to study complexes containing multiple proteins and/or RNAs. Numerous studies have used the three-hybrid system to analyze the specificity of RNA–protein interactions. Screens utilizing the three-hybrid system have identified proteins that regulate mRNA translation and splicing. Conversely, specific RNA-binding sites have been identified for several RNA-binding proteins. Many of these studies have been summarized previously (Bernstein *et al.*, 2002).

We have mentioned a few adaptations to the basic architecture of the three-hybrid system. The ability to adapt the three-hybrid system to address different questions, in combination with current biochemical techniques, yields a multifaceted toolbox to help understand and dissect RNA–protein interactions.

ACKNOWLEDGMENTS

We thank members of the Wickens laboratory for their helpful comments on this manuscript. We are also grateful to the Media Laboratory of the Biochemistry Department of the University of Wisconsin for help with the figures. The Wickens laboratory is supported by research grants from the National Institutes of Health.

REFERENCES

Beach, D. L., Salmon, E. D., and Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. *Curr. Biol.* **9,** 569–578.

Bernstein, D., Hook, B., Hajarnavis, A., Opperman, L., and Wickens, M. (2005). Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF-1. *RNA* **11,** 447–458.

Bernstein, D. S., Buter, N., Stumpf, C., and Wickens, M. (2002). Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods* **26,** 123–141.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2,** 437–445.

Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., and Cullen, B. R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *EMBO J.* **17,** 7056–7065.

Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., and Cullen, B. R. (1999). Analysis of the effect of natural sequence variation in Tat and in cyclin T on the formation and RNA binding properties of Tat-cyclin T complexes. *J. Virol.* **73,** 5777–5786.

Bouffard, P., Barbar, E., Briere, F., and Boire, G. (2000). Interaction cloning and characterization of RoBPI, a novel protein binding to human Ro ribonucleoproteins. *RNA* **6,** 66–78.

Buskirk, A. R., Kehayova, P. D., Landrigan, A., and Liu, D. R. (2003). *In vivo* evolution of an RNA-based transcriptional activator. *Chem. Biol.* **10,** 533–540.

Buskirk, A. R., Landrigan, A., and Liu, D. R. (2004). Engineering a ligand-dependent RNA transcriptional activator. *Chem. Biol.* **11,** 1157–1163.

Cassiday, L. A., and Maher, L. J., 3rd. (2001). *In vivo* recognition of an RNA aptamer by its transcription factor target. *Biochemistry* **40,** 2433–2438.

Cassiday, L. A., and Maher, L. J., 3rd. (2003). Yeast genetic selections to optimize RNA decoys for transcription factor NF-kappa B. *Proc. Natl. Acad. Sci. USA* **100,** 3930–3935.

Coller, J., and Wickens, M. (2002). Tethered function assays using 3′ untranslated regions. *Methods* **26,** 142–150.

Edwards, T. A., Pyle, S. E., Wharton, R. P., and Aggarwal, A. K. (2001). Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* **105,** 281–289.

Good, P. D., and Engelke, D. R. (1994). Yeast expression vectors using RNA polymerase III promoters. *Gene* **151,** 209–214.

Hook, B., Bernstein, D., Zhang, B., and Wickens, M. (2005). RNA-protein interactions in the yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening. *RNA* **11,** 227–233.

Kehayova, P. D., and Liu, D. R. (2007). *In vivo* evolution of an RNA-based transcriptional silencing domain in *S. cerevisiae*. *Chem. Biol.* **14,** 65–74.

Kraemer, B., Zhang, B., SenGupta, D., Fields, S., and Wickens, M. (2000). Using the yeast three-hybrid system to detect and analyze RNA-protein interactions. *Methods Enzymol* **328,** 297–321.

Lim, F., Spingola, M., and Peabody, D. S. (1994). Altering the RNA binding specificity of a translational repressor. *J. Biol. Chem.* **269,** 9006–9010.

Long, R. M., Gu, W., Lorimer, E., Singer, R. H., and Chartrand, P. (2000). She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA. *EMBO J.* **19,** 6592–6601.

Lowary, P. T., and Uhlenbeck, O. C. (1987). An RNA mutation that increases the affinity of an RNA-protein interaction. *Nucleic Acids Res.* **15,** 10483–10493.

Martin, F., Michel, F., Zenklusen, D., Muller, B., and Schumperli, D. (2000). Positive and negative mutant selection in the human histone hairpin-binding protein using the yeast three-hybrid system. *Nucleic Acids Res.* **28,** 1594–1603.

Martin, F., Schaller, A., Eglite, S., Schumperli, D., and Muller, B. (1997). The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO J.* **16,** 769–778.

Opperman, L., Hook, B., DeFino, M., Bernstein, D. S., and Wickens, M. (2005). A single spacer nucleotide determines the specificities of two mRNA regulatory proteins. *Nat. Struct. Mol. Biol.* **12,** 945–951.

Rho, S. B., and Martinis, S. A. (2000). The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. *RNA* **6,** 1882–1894.

Robzyk, K., and Kassir, Y. (1992). A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Res.* **20,** 3790.

Saha, S., Ansari, A. Z., Jarrell, K. A., and Ptashne, M. (2003). RNA sequences that work as transcriptional activating regions. *Nucleic Acids Res.* **31,** 1565–1570.

Seay, D., Hook, B., Evans, K., and Wickens, M. (2006). A three-hybrid screen identifies mRNAs controlled by a regulatory protein. *RNA* **12,** 1594–1600.

Sengupta, D. J., Wickens, M., and Fields, S. (1999). Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *RNA* **5,** 596–601.

SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S., and Wickens, M. (1996). A three-hybrid system to detect RNA-protein interactions *in vivo*. *Proc. Natl. Acad. Sci. USA* **93,** 8496–8501.

Sonoda, J., and Wharton, R. P. (1999). Recruitment of Nanos to hunchback mRNA by Pumilio. *Genes Dev.* **13,** 2704–2712.

Sonoda, J., and Wharton, R. P. (2001). Drosophila brain tumor is a translational repressor. *Genes Dev.* **15,** 762–773.

Stumpf, C. R., Kimble, J., and Wickens, M. (2008). A Caenorhabditis elegans PUF protein family with distinct RNA binding specificity. *RNA* **14,** 1550–1557.

Wang, Z. F., Whitfield, M. L., Ingledue, T. C., 3rd, Dominski, Z., and Marzluff, W. F. (1996). The protein that binds the 3′ end of histone mRNA: A novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev.* **10,** 3028–3040.

Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997). A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* **390,** 477–484.