Automated Yeast Two-hybrid Screening for Nuclear Receptor-interacting Proteins*⒮

Michael Albers‡, Harald Kranz§, Ingo Kober¶, Carmen Kaiser‖, Martin Klink,
Jörg Suckow‡‡, Rainer Kern**, and Manfred Koegl‡§§**

High throughput analysis of protein-protein interactions is an important sector of hypothesis-generating research. Using an improved and automated version of the yeast two-hybrid system, we completed a large interaction screening project with a focus on nuclear receptors and their cofactors. A total of 425 independent yeast two-hybrid cDNA library screens resulted in 6425 potential interacting protein fragments involved in 1613 different interaction pairs. We show that simple statistical parameters can be used to narrow down the data set to a high confidence set of 377 interaction pairs where validated interactions are enriched to 61% of all pairs. Within the high confidence set, there are 64 novel proteins potentially binding to nuclear receptors or their cofactors. We discuss several examples of high interest, and we expect that communication of this huge data set will help to complement our knowledge of the protein interaction repertoire of this family of transcription factors and instigate the characterization of the various novel candidate interactors. *Molecular & Cellular Proteomics 4:205–213, 2005.*

Nuclear receptors are a family of transcription factors involved in the control of many physiological processes including development, sexual differentiation, inflammation, and metabolism (1, 2). They can bind to DNA directly or via interaction with other proteins. Nuclear receptor activity is regulated by the binding of small molecule ligands to the receptor and/or by posttranslational modifications. Activation of nuclear receptors involves a change in conformation that affects the interaction of the receptor with other proteins, which in turn brings about the effect of the receptor on gene expression (3, 4). Knowledge about the ligand-dependent binding of nuclear receptors to their cofactors is central to the understanding of their physiological function and their use as tar-

gets for drug discovery (5). However, the available knowledge is highly biased toward a few intensively studied receptors, and little is known for the potential interaction patterns of the rest of the family (6).

The elucidation of protein-protein interaction patterns provides an important basic data set in the functional analysis of the proteome. In the past, large data sets on protein interactions have been reported for model organisms such as yeast (7), fly (8), and worm (9). Here we present a protein interaction data set focusing on the family of human nuclear receptors generated by automated yeast two-hybrid (Y2H)[1] library screening. Application of statistical selection methods led to the generation of a high confidence subset of interaction pairs.

EXPERIMENTAL PROCEDURES

Yeast Two-hybrid Screening—"Bait" is the protein or protein fragment for which we tried to find interacting proteins in a cDNA library using the yeast two-hybrid method. The bait is always a fusion with the DNA binding domain of the GAL4 transcription factor. "Prey" is a protein or protein fragment isolated from a cDNA library in a yeast two-hybrid screen as potentially interacting with the bait. The prey is always a fusion with the activation domain of the GAL4 transcription factor. cDNAs encoding bait fragments were generated by PCR, cloned into pDONR201, and transferred into GATEWAY (Invitrogen)-

[1] The abbreviations used are: Y2H, yeast two-hybrid; LBD, ligand binding domain; PPAR, peroxisome proliferator-activated receptor; GCNF, germ cell nuclear factor; PNR, photoreceptor cell-specific nuclear receptor; PNRC, proline-rich nuclear receptor co-regulatory protein; SF1, steroidogenic factor 1; ER, estrogen receptor; ERR, estrogen-related receptor; ROR, retinoid-related orphan receptor; NR, nuclear receptor; TR2 and TR4, testicular orphan receptors 2 and 4; PR, progesterone receptor; EAR, eosinophil-associated ribonuclease; COUP-TF, chicken ovalbumin upstream promoter-transcription factor; GABARAP, γ-aminobutyric acid receptor type A receptor-associated protein; NCoA, nuclear receptor coactivator; SRC1, steroid receptor coactivator 1; TIF2, transcriptional intermediary factor 2; NCoR, nuclear receptor co-repressor; SMRT, silencing mediator for retinoid and thyroid hormone receptor; TRIP1, thyroid hormone receptor-interacting protein 1; PHLP, phosducin-like protein; NUCB, nucleobindin; LRH1, liver receptor homologue 1; pCAF, p300/CBP-associated factor (CBP, cAMP-response element-binding protein (CREB)-binding protein); SHP, short heterodimer partner; UTR, untranslated region; HNF4α, hepatocyte nuclear factor 4α; MR, mineralocorticoid receptor; SKIP, Ski-interacting protein; VDR, vitamin D receptor; TRAP, thyroid hormone receptor-associated protein; PXR, pregnane X receptor; RAR, retinoic acid receptor; SMIF, Smad-interacting factor; TRα, thyroid hormone receptor α; SHARP, SMRT/HDAC1-associated repressor protein (HDAC1, histone deacetylase 1); LXR, liver X receptor; eEF-1D, eukaryotic elongation factor-1D.

compatible versions of pGBT9 and pGAD424 by the LR reaction as specified by Invitrogen. Yeast strain CG1945 (Clontech) was transformed with the resulting vector. cDNA libraries (Clontech) were transformed into Y187 (Clontech). Both full-length receptor and fragments encompassing only the LBD of these receptors were screened against several cDNA libraries in the presence or absence of appropriate small molecule ligands. For screening, bait- and prey-expressing yeasts were mated in YPDA (yeast extract, peptone, dextrose, and adenine) in the presence of 10% polyethylene glycol 6000. Medium was changed to selective medium (synthetic dextrose) lacking Leu, Trp, and His with the following additives: 0.5% penicillin/streptomycin (50 μg/ml, Invitrogen), 50 μM 4-methylumbelliferyl-α-D-galactoside (Sigma), and varying concentrations of 3-aminotriazol (Sigma). In addition, low molecular weight ligands of nuclear receptors were added to some of the screens as indicated in Supplemental Data 1. The mating efficiency was determined by plating of cells on selective agar plates. Typically 5–40 million diploids were generated. The cell suspension was then aliquoted into microtiter plates (96 wells/plate, flat bottom, 200 μl/well) and incubated for 3–7 days. To identify wells containing positive clones, fluorescence was determined on a SpectraFluor fluorometer (Tecan) at 465 nm (excitation at 360 nm). Wells that displayed fluorescence above background were identified and automatically collected by a Tecan Genesis 200 robot. Selected cells were passaged to new wells twice and once to an agar plate before amplifying the library inserts by PCR. PCR products of sufficient quality for sequencing were collected, and the identity of the insert was determined by DNA sequencing in house or at GATC Biotech AG (Konstanz, Germany). We did not succeed in cloning NHR3B2/ERRβ. For the following nuclear receptors, we failed to find useful screening conditions: NR1F3/RORγ, NR2C1/TR2, NR2C2/TR4, NR2E1/TLX, NR3C3/PR, NR5A1/SF1, NR6A1/GCNF, NR4A1/Nurr77, and NR2E3/PNR. PPARα had to be removed from the data set after the project had been completed due to a mutation in the bait construct that had previously been overlooked.

Definition of a High Confidence Data Set—To determine the promiscuity of a prey, the number of different baits that led to the isolation of the prey was considered. For example, NM_003299/Hsp96 [2] has been found with six different baits; thus, the number of different baits is 6. Different fragments of the same bait were treated as the same bait. For example, NM_001584 has been found to bind to both full-length NR2F6/EAR2 and to fragments of the same protein encompassing only the ligand binding domain as well as to NR2F1/COUP-TFII. In this case, the number of different baits of NM_001584 is 2 because two of the baits are derived from the same gene and the fragments overlap (*i.e.* the ligand binding domain overlaps to its entirety with the full-length protein). To yield the promiscuity, the number of different baits per prey was divided by the total number of times this prey was isolated in our screens. This normalization step accounts for the abundance of the prey in the cDNA libraries screened. For example, NM_003299/Hsp96 has been found altogether eight times with six different baits so the promiscuity is 6/8 = 0.75, whereas for retinoid X receptor α, which has been found 329 times with 13 different baits, the promiscuity is 0.04.

To determine the promiscuity of the bait, the number of different prey proteins isolated with a given bait was divided by the number of different fragments analyzed with the bait. For example, for NR3A2/ERβ, we isolated 56 cDNA fragments that came from 43 different proteins. Thus, the promiscuity is 43/56 = 0.77. For the LBD of NR3A2/ERβ, we isolated 141 cDNA fragments that came from 13 different proteins such that the promiscuity is 0.09. In this case, the promiscuity of each different bait fragment was determined, and

overlapping fragments of the same protein were not pooled for the analysis (as for determining the promiscuity of the preys). To determine the promiscuity of protein pairs, the promiscuities of the respective bait and prey were multiplied.

To select pairs based on the interaction pattern, we defined groups of paralogues and highly related proteins. In the following, these groups are listed separated by semicolons: NM_003743/SRC1, NM_006540/TIF2, and NM_006534 NCoA3; NM_006311/NCoR1 and NM_006312/SMRT; the NR1 subfamily; the NR2 subfamily; the NR3 subfamily; the NR4 subfamily; NM_002805/TRIP1, NM_005388/PHLP, NM_006503/MIP221, and NM_002815/PSMD11; NM_007285/GABARAPL1 and NM_031412/GABARAPL2; NM_006813/PNRC1 and NM_017761/PNRC2; NM_005013/NUCB1 and NM_006184/NUCB2; NM_001455/FOXO3A and NM_002015/FOXO1A; the troponins; and the actinins. NR5A2/LRH1 was grouped with the NR3 subfamily since it displayed markedly similar interaction patterns. Pairs were selected based on their interaction pattern if (i) a protein A interacts with a protein B, and the identical protein A interacts with a protein B′ where B and B′ are related proteins as defined above (for example, pCAF is seen to interact with actinin α1 and also with the related protein actinin α2) or if (ii) an interaction of two proteins A and B is also observed for two proteins A′ and B′ where A is related to A′ and B is related to B′ (for example, NR0B2/SHP ("A") interacted with NR3B1/ERRα ("B") and NR0B1/DAX ("A′," related to NR0B2/SHP) interacted with NR3B3/ERRγ ("B′," related to NR3B1/ERRα). To select pairs based on the independent occurrence of the interaction, all pairs were considered where one of the following conditions was met: i) isolation of the same preys from different cDNA libraries (when screening multiple cDNA libraries with the same bait), ii) isolation of the same prey with different but overlapping fragments of the same bait, or iii) isolation of the bait as a prey when the prey is used as a bait protein in screens (reciprocal interaction).

For all interaction pairs, the part of the mRNA corresponding to the isolated fragments (the 5′-UTR, the coding region, or the 3′-UTR) was determined. All fragments that corresponded to the non-coding strand of cDNAs were automatically removed. For 7.7% of all pairs, only fragments corresponding to the 3′-UTR were found; these were removed in the final selection. All other clones were kept for further analysis. For 11.8% of interaction pairs only fragments with a fusion point mapping to the 5′-UTR were isolated, for 43.8% of pairs only fragments with a fusion point mapping to the coding region were isolated, and for 5.6% of pairs fragments with a fusion point mapping to several regions, including the coding region, were isolated. For 31% of the fragments, the coding region was not known at the time of the analysis. Translational frameshifting can lead to significant expressions of the correct reading frame even in cases when stop codons or frameshifts are generated for the hybrid proteins (8), and the reading frame is not a good predictor for the reliability of an interaction (9). For this reason, stops and frameshifts were not used as criteria to remove preys from the data set.

All steps of the process were directed by a dedicated laboratory information management system using a relational data base (Oracle). Assignment of experimentally detected sequences to known sequence entries was performed by a "blastn" search against a data base containing *Homo sapiens* RefSeq sequences from the National Center for Biotechnology Information (NCBI). In case there was no assignment at an e-value below 10^{-5}, a second blast against a *H. sapiens* subset of the UniGene data base (NCBI) was run.

<div align="center">RESULTS</div>

High Throughput Yeast Two-hybrid Screening

The goal of this project was an unbiased, systematic approach to the isolation of proteins potentially interacting with

---

[2] NM_003299 and all other accession numbers mentioned in this paper are from GenBank™.

nuclear receptors by use of the Y2H system. To make a project of this scale feasible, the throughput of the method was increased by performing all operations on microtiter plates and with pipetting robots. Among many other incremental improvements, the use of quantifiable reporters was an important step in reducing the background of spurious positive clones (see "Experimental Procedures"). Full-length versions or fragments from 38 of the 48 human nuclear receptors were successfully used as baits in screens against several cDNA libraries each in the presence of appropriate small molecule ligands. 23 additional proteins were included in our set of baits, most of which are known to bind nuclear receptors directly or indirectly. Over 12,000 different fragments were isolated initially. After automated removal of poor quality data and sequences mapping to the non-coding DNA strand, 6425 cDNA fragments were retained in our data base, forming 1613 different protein interaction pairs.

Generation of a High Confidence Data Set

Statistical parameters have been used successfully to evaluate the reliability of protein-protein interactions in large data sets (8–11). To evaluate the results of a selection of protein pairs based on statistical values, we defined positive and negative reference data sets. A positive reference data set was defined consisting of 257 previously published interaction pairs and of interactions of well known cofactors with nuclear receptors. A negative reference data set of 34 different pairs was defined consisting of proteins that are well known to reside outside the plasma membrane or within the mitochondrion. Four different parameters were found to be useful to enrich for the positive reference data set. First, the number of different interaction partners of the proteins forming the interaction pair (the promiscuity, see "Experimental Procedures" for details) can be used to deplete for interactions involving promiscuous proteins. Fig. 1*A* depicts the effects of using the promiscuity of a protein pair (*x axis*) as a selection criterion on the fractions of remaining pairs from the reference sets (*y axis*). As can be seen, the fraction of pairs from the positive reference set increases as the data set is depleted for protein pairs that involve highly promiscuous partners. Second, the number of times a given prey was isolated is a useful parameter to enrich for interaction pairs from the positive reference set. As seen in Fig. 1*B*, setting a minimum number of times that a prey has been isolated as a threshold criterion for selection of pairs leads to an increase in the fraction of pairs from the positive reference set as the threshold is increased.

Fig. 1*C* shows a plot of the different protein pairs according to the promiscuity of the pair against the number of times the respective prey has been isolated in our screens. In this plot, rare interactions of promiscuous proteins are localized to the *top left*, whereas frequently found preys that are involved in non-promiscuous interactions are localized to the *bottom*

right. To allow for seldom found preys involved in interactions with a low promiscuity, we divided the data set according to *diagonal lines* with varying slopes. A decreasing slope effectively enriches for the positive reference data (Fig. 1*D*). Selection of protein pairs according to the *separation lines* shown in Fig. 1, *C* and *D*, leads to an enrichment of the positive reference data set to 53% and to a loss of 76% of all interaction pairs but only 15% of the positive reference data.

As a third parameter, we looked for interactions that were picked up independently in several approaches, *e.g.* in screens against different cDNA libraries. This selection aims to deplete "technical" false positives where the interaction is not caused by an interaction of the bait and prey protein but by a spurious event specific to the Y2H method (12). As a fourth criterion, an evolutionary argument was used. If an interaction signal is based on a true affinity of two proteins and not on a technical artifact, similarity in protein sequence should be reflected in similar interaction properties. Based on this principle, pairs were selected if a protein interacted with two or more paralogues or if paralogous proteins interacted with proteins highly related to each other.

To derive a high confidence data set, we selected interaction pairs that meet at least three of the four selection criteria, yielding 377 different protein pairs (for details, see Supplemental Data). 61% (232) of these pairs reproduce previously published interactions or interactions of nuclear receptors with well known cofactors, 32% (120 pairs) involve novel proteins potentially associated with nuclear receptors, and 25 (7%) are interactions from the negative reference set (putative false positives). Fig. 2 summarizes the effects of applying the different selection criteria.

This data set contains numerous interaction pairs that had not previously been reported (Fig. 3, *A* and *B*; for details, see Supplemental Data). We believe that the fact that these novel interactors are found within a group of proteins where 61% of different interactions involve previously validated interactors allows the assumption that many of them of should be of biological significance. Several examples of interest are discussed below.

Examples of Potential Novel Nuclear Receptor-binding Proteins

NM_025082 and NR3A1/ERα—Fragments corresponding to the cDNA NM_025082 were isolated from five different libraries in screens using the LBD of NR3A1/ERα as a bait. In addition, use of NR5A2/LRH1-LBD as a bait led to the isolation of NM_025082 from two different cDNA sources. No other baits picked up this cDNA clone. When we tested the full-length coding sequence of this cDNA for interaction with NR3A1/ERα, NR3A2/ERβ, NR3B1/ERRα, and NR3B3/ERRγ, the clone gave strong interaction signals for all receptors
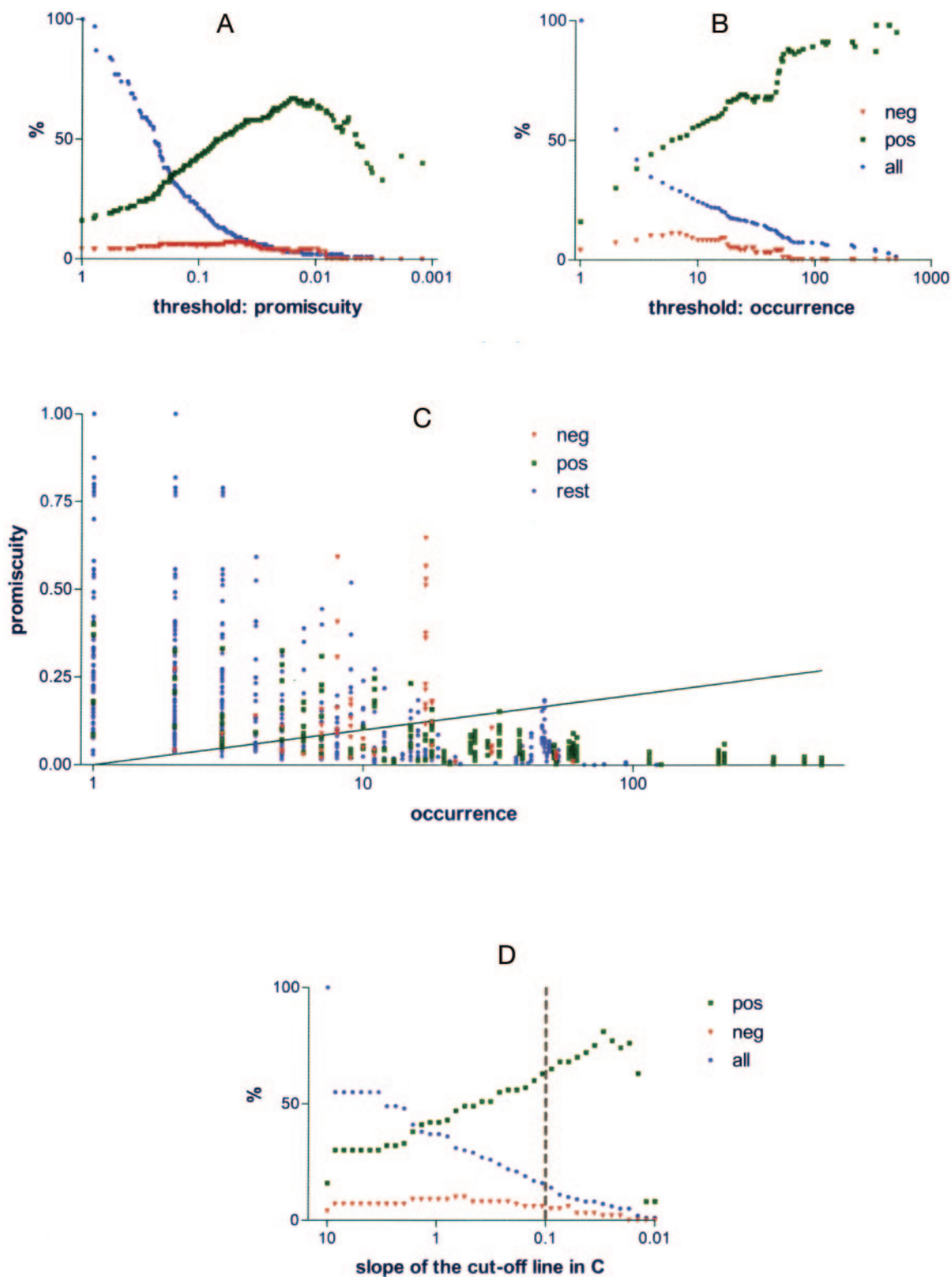
FIG. 1. *A*, *B*, and *D*, subsets of data were generated by excluding all interaction pairs that were beyond an increasing threshold level displayed on the *x axis*. The *y axis* of the graphs show the positive (*pos, green squares*) and negative (*neg, red triangles*) reference sets of different protein pairs as fractions of the remaining data as well as the overall number of pairs remaining (*all, blue circles*) as a function of the threshold level on the *x axis*. *A*, percentage of the three data sets as a function of decreasing promiscuities of the interaction pair as a threshold from *left* to *right*. *B*, percentage of the three data sets as a function of the occurrence of the prey (*i.e.* the total number of times a prey of a respective protein pair was isolated in our screens). *C*, plot of the promiscuity of a given pair against the occurrence of the respective prey. In *C*, *blue circles* depict

Total data set
(1613 different pairs)
16%

# of isolates + promiscuity
(425 different pairs)
53%

Multiple isolation
(395 different pairs)
43%

76%   85%

87%

Interaction pattern
(424 different pairs)
57%

2 of 3 criteria
(377 different pairs)
61%

All criteria
(187 different pairs)
78%

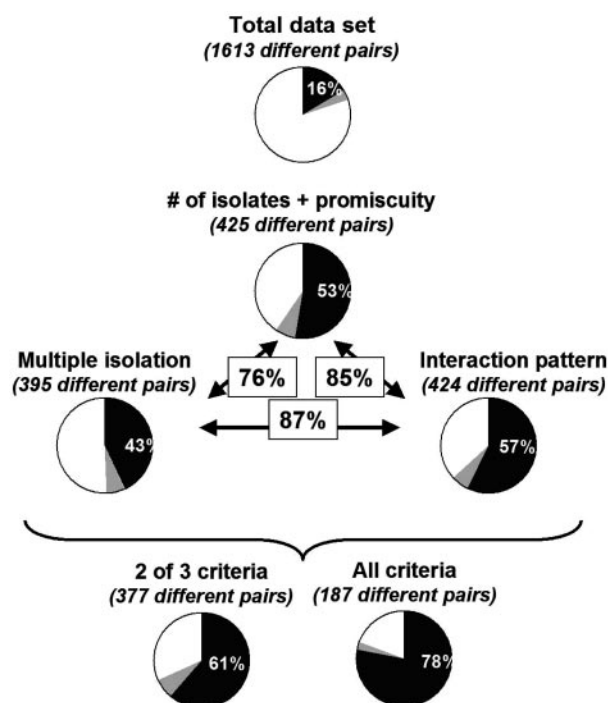Fig. 2. **Enrichment for the positive reference set by various criteria.** Subsets of the whole data set (*top pie chart*) were selected according to the criteria indicated. *Dark gray area*, positive reference set; *light gray area*, negative reference set; *white area*, unclassified interactions. The overlap of the subsets generated by the three criteria is indicated by *numbers over the arrows*. Selection of interactions that meet at least two of the three criteria (*bottom row*, *left pie chart*) and of interactions that meet all criteria (*bottom right*) after removal of fragments mapping to the 3′-UTR of the cDNA are shown.

except ERβ.[3] For NR3A1/ERα, the interaction was dependent on the presence of 17β-estradiol.

ERα is a well studied protein, and more than 90 different proteins have been published to bind this protein (6, 13). We were therefore surprised to see hitherto unpublished interactors of this receptor appearing in our screens. However, ongoing research has confirmed the interaction of NM_025082 and NR3A1/ERα by independent methods and shown that it depends on an L*XX*LL motif related to the one of NR0B2/SHP and appears to have a negative effect on ERα- but not ERβ-dependent transcription.[3] Interestingly use of this protein as a bait led to the isolation of 66 fragments of a single protein, Par4/NM_002583, which has been reported to play a role in apoptosis in prostate cancer cells and transcriptional regulation (14).

NM_002763/Prox1—Prox1 is a homeobox-containing transcription factor that is involved in eye development and lym-

phatic endothelial differentiation (15). Fragments of Prox1 were picked up as potential interactors of NR1C2/PPARβ, NR1F1/RORα, NR2A1/HNF4α, NR3B1/ERRα, NR3B3/ERRγ, and NR3C2/MR. Several independent fragments of this protein were isolated, all of them in a cDNA library prepared from testis. No other bait led to the isolation of Prox1. In quantitative PCR analysis of the tissue distribution of these receptors, we find that NR1C2/PPARβ, NR3B1/ERRα, NR3C2/MR and the splice variant 2 NR1F1/RORα are expressed to significant levels in testis.[4] An interaction of NR5A2/LRH1 and Prox1 has been reported (16), supporting the idea that Prox1 is a novel nuclear receptor-binding protein.

The NR2F Family—A strong case for specificity can be made when related baits selectively interact with related preys. Screens using NR2F1/COUP-TFI, NR2F6/EAR2, or the LBD of NR2F6/EAR2 as baits led to the isolation of multiple fragments of two proteins with homology to phosphoesterases NM_001584 and NM_001585 (Fig. 4). These two proteins are 79.3% identical in amino acid sequence. The fact that two related receptors independently lead to the isolation of two very similar proteins lends support to the assumption that these interactions should be of biological significance. Another case of similar interaction patterns of paralogues is the binding of COUP-TFI and COUP-TFII to the eukaryotic translation elongation factor eEF-1D.

NCoA62/SKIP—NCoA62/SKIP has been reported as a protein interacting with NR1I1/VDR as well as with several cofactors (17). Screens using NCoA62/SKIP (amino acids 209–336) as a bait led to the isolation of the nuclear receptors DAX/NR0B1, SHP1/NR0B2, and NR2F2/COUP-TFII, suggesting that the panel of receptors interacting with NCoA62/SKIP may be greater than hitherto assumed. The reciprocal isolation of NCoA62/SKIP in a screen using NR0B1/DAX as bait confirms this interaction. In addition, a cDNA corresponding to NM_138421 was isolated as a high confidence interactor of NCoA62/SKIP. NM_138421 has also been isolated with another cofactor, the TRAP220 protein (not shown). Potentially NM_138421 may be of importance in bridging NCoA62/SKIP and the TRAP-mediator complex for interaction with nuclear receptors.

PNRC2—PNRC2 is a cofactor of nuclear receptors isolated by Y2H screens using NR5A1/SF1 as bait (18). While our efforts to screen with SF1 as a bait failed, we isolated PNRC2 as a prey with 10 different nuclear receptor baits: NR3A1/ERα, NR3A2/ERβ, NR3B3/ERRγ, NR2A1/HNF4α, NR2A2/HNF4γ, NR5A2/LRH1, NR1I2/PXR, NR1B3/RARγ, NR1F1/RORα, and

---

---

interactions that were not part of the positive or negative reference set. For definition of the high confidence data set, only pairs mapping below the *diagonal cut-off line* were considered. *D*, percentage of the three data sets at decreasing slopes of the *diagonal cut-off line* displayed in *C*. A slope of 0.1 has been applied to our data set as the threshold to enrich for interaction pairs of the positive reference set as shown in *C* and *D*.
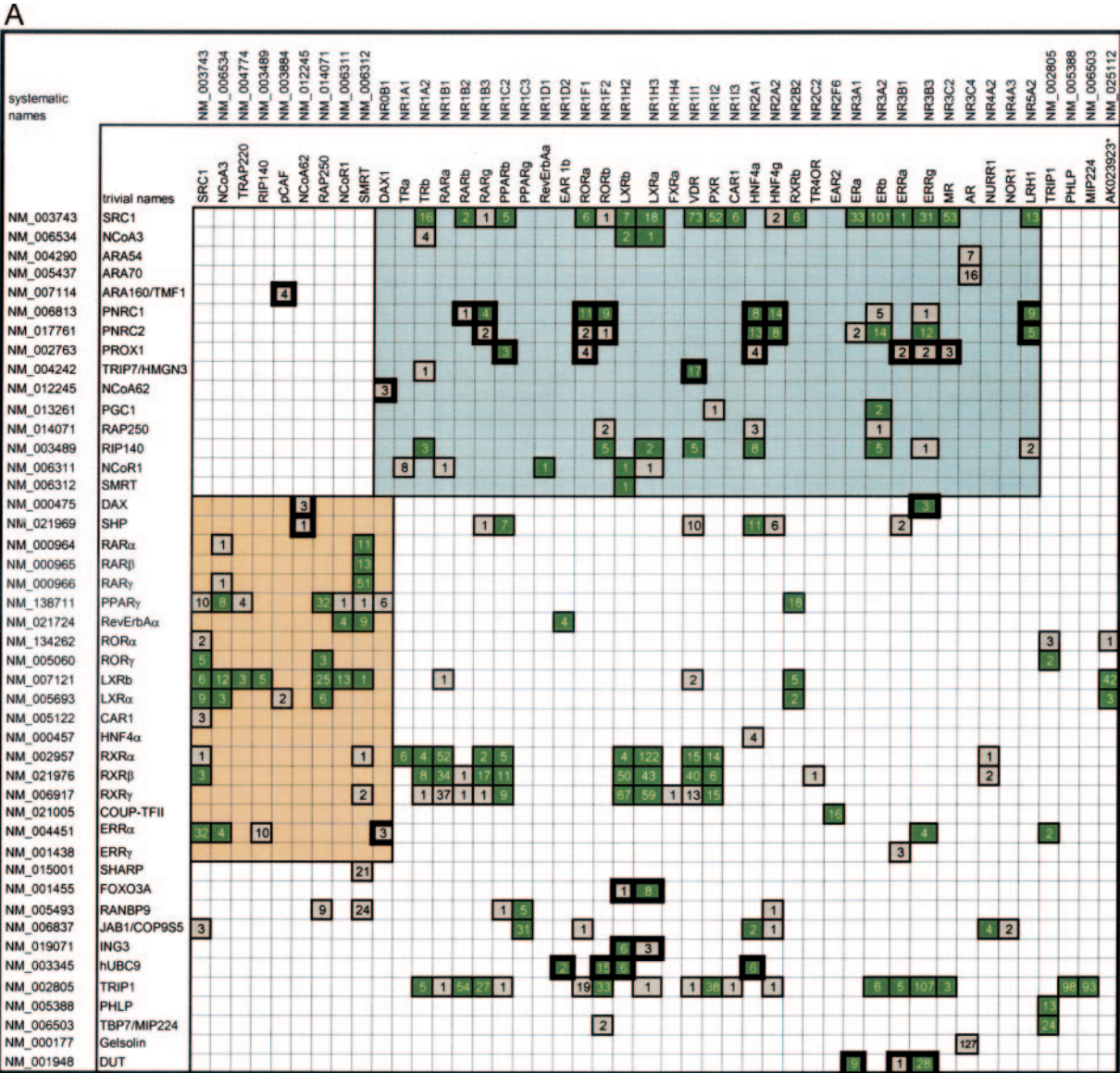
Fɪɢ. 3. **High confidence interaction pairs.** *Top*, baits; *left column*, preys. *Numbers in the boxes* represent the number of isolates recovered of a given prey with the respective bait. *Green* (in *A*) or *dark gray* (in *B*) boxes identify interactions that meet all selection criteria (see Fig. 2). *A*, interactions involving known nuclear receptor binding partners. *Orange box*, interactions of the cofactors with nuclear receptors (bait *versus* prey); *blue box*, interactions of the nuclear receptors with cofactors (bait *versus* prey). Some interactions of interest that have not been previously published are indicated by *bold borders* (see text). Note that the bait labeled *AK023923.1** is out of frame with the respective data base entry (see Supplemental Data 1). *B*, interactions involving preys not previously noted to bind to nuclear receptors. *Boxed* interactions are discussed in the text.

NR1F2/RORβ. Thus, PNRC2 appears to be a more general cofactor for nuclear receptor than previously appreciated. When full-length PNRC2 was used as bait in our screens, the only protein isolated was a transcription factor termed Smad-interacting factor (SMIF). When the screen was repeated with a fragment of PNRC2 (amino acids 74–139), again only SMIF was isolated as a prey. This protein binds to components of the transforming growth factor-β signal transduction pathway and cooperates with the transforming growth factor-β-stimulated transcriptional regulator SMAD4 (19). The potential in-

teraction of SMIF with PNRC2 is suggestive of a functional link between transforming growth factor-β signaling and the activity of the nuclear receptor cofactor PNRC2, an assumption that should be readily testable.

*Cullin 1 and NR1A1/TRα—*Cullin 1 (NM_003592) was found in screens with the LBD of NR1A1/TRα in the presence of triiodothyronine as several independent fragments in liver, testis, and kidney cDNA libraries. Cullins are a family of proteins that function in ubiquitin-dependent protein degradation (20). The importance of the ubiquitin system for nuclear re-
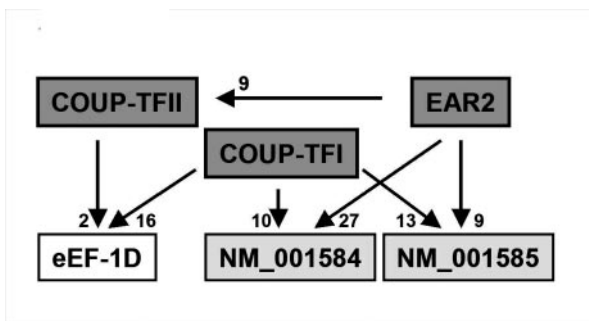
B

Column headers (systematic names): NM_003743, NM_017761, NM_004229, NM_004774, NM_003489, NM_003884, NM_006292, NM_012245, NM_014071, NM_014288, NM_025082, NM_006311, NM_006312, NR1A1, NR1A2, NR1B1, NR1B2, NR1B3, NR1C2, NR1C3, NR1D1, NR1D2, NR1F2, NR1H2, NR1H3, NR1H4, NR1I1, NR1I2, NR1I3, NR2A2, NR2B1, NR2B2, NR2B3, NR2F1, NR2F6, NR3A1, NR3B1, NR3B3, NR3C1, NR3C4, NR4A3, NR5A2, NM_002805, NM_005388, NM_006503, NM_001220, NM_025112, AK023923*, NM_006337

Column headers (trivial): SRC1, PNRC2, DRIP150, TRAP220, RIP140, pCAF, TSG101, NCoA62, RAP250, NRIF3, NM_025082, NCoR1, SMRT, TRa, TRb, RARa, RARb, RARg, PPARb, PPARg, RevErbAa, EAR 1b, RORb, LXRb, LXRa, FXRa, VDR, PXR, CAR1, HNF4g, RXRa, RXRb, RXRg, COUP-TF1, EAR2, ERa, ERRa, ERRg, GR, AR, NOR1, LRH1, TRIP1, PHLP, MIP224, CAMK2B, AK023923.1, MIF1

systematic name	trivial	values
NM_005013	NUCB2	PXR:11
NM_002136	HNRNPA1	MIP224:2
NM_007056	SWAP2	HNF4g:25
NM_002583	PAR-4	NRIF3:66
NM_014682	NM_014682	AR:4
NM_015640	PAI-RBP1	pCAF:4
NM_018403	SMIF	SRC1:21
NM_024309	ABIN-2	NCoR1:92
NM_003592	Cul1	NCoR1:11; EAR1b:1
NM_007285	GABARAPL2	VDR:14
NM_031412	GABARAPL1	VDR:10
NM_002109	HARS	CAMK2B:9
NM_001960	EEF1D	COUP-TF1:9
NM_004905	PRDX6	NCoA62:17; RARg:1; LXRb:1
NM_001220	CAMK2B	SMRT:3; COUP-TF1:4; AK023923:30
NM_004663	RAB11A	pCAF:37
NM_004726	POB1/REPS2	GR:15
NM_006118	HAX1	RXRa:11
NM_012278	Melusin	NCoA62:6; RAP250:2
NM_014958	GEF15	ERRa:10; ERRg:1
NM_003380	Vimentin	AR:39
NM_012090	MACF1	ERa:2
NM_144705	NM_144705	AR:4
NM_001102	Actinin alpha1	SRC1:1; PNRC2:1; RIP140:11; RXRa:1
NM_001103	Actinin alpha2	SRC1:118; PNRC2:108; RIP140:16; pCAF:29; TSG101:1; RAP250:18; NCoR1:35; RXRb:2
NM_000363	Troponin 1	ERRa:6
NM_003282	Troponin 1f	RARa:1; RORb:1; ERRa:46
NM_003283	Troponin 1s	ERRa:31
NM_006757	Troponin T3f	NCoA62:13
NM_033063	MAP6	NCoA62:10; SMRT:1; RevErbAa:2; COUP-TF1:1
NM_004082	Dynactin1	DRIP150:7
NM_002213	Integrin beta5	NM_025082:17
NM_007317	KIFF22	TRb:9; MIF1:3
NM_007186	CEP2	pCAF:8
NM_002065	GLUL	RARa:1
NM_014236	GNPAT	VDR:21
NM_002046	GAPDH	VDR:2; RXRb:3
NM_133491	SSAT2	pCAF:4
NM_006579	EBP	GR:4
NM_004068	AP2M1	TSG101:11
NM_018652	GLP	pCAF:41
NM_001584	NM_001584	COUP-TF1:10; EAR2:27
NM_001585	NM_001585	COUP-TF1:13; EAR2:9
NM_004683	RGN	TSG101:14; RORb:1
NM_006029	PNMA1	pCAF:9
NM_006384	calmyrin	RIP140:7
NM_013241	FHOD1	NCoA62:3; PPARb:10
NM_014342	NM_014342	NM_025082:2; ERRa:6
NM_017634	NM_017634	NR1I1/VDR:3; NR1I2/PXR:14
NM_017822	NM_017822	MIF1:72
NM_018003	uveal autoantigen	TRAP220:130
NM_019065	EFCBP2	TRAP220:23
NM_025082	NM_025082	ERa:5; AR:3
NM_025112	AK023923.1	RARg:7
NM_033296	PGR1	pCAF:20
NM_138421	NM_138421	NCoA62:35
NM_138423	NM_138423	FXRa:13
NM_138483	KIF1A	pCAF:13; NM_025082:6; RXRa:1; ERRa:3; AR:6
NM_152415	HCRP1	pCAF:78
AL134489	AL134489	LRH1:5
AL137505	AL137505	SMRT:11
AW292937	AW292937	TRb:3
BF668725	BF668725	PHLP:1; MIP224:1
BF981716	BF981716	SMRT:7

FIG. 3— *continued*

ceptor activity (21, 22) suggests that cullin-dependent ubiquitination may play a role in regulating the activity of the thyroid hormone receptors. Of note, the only other bait that led to the isolation of Cullin 1 was the LBD of NR1D1/EAR1b, a nuclear receptor of the same subfamily as NR1A1/TRα.

SMRT—When a fragment of the corepressor SMRT was

FIG. 4. **Protein interactions within the NR2F family.** *Numbers* indicate the number of prey fragments isolated in the screens.

used as bait, fragments from 21 potentially interacting different proteins were isolated. Among those, nine nuclear receptors were found that had previously been shown to interact with this corepressor as well as the SMRT-binding protein SHARP (23). The isolation of these established interactors of SMRT demonstrates that the screen has worked technically well and lends trust to other potentially novel interactors of SMRT. The most prominent protein fragment isolated with SMRT as a bait corresponds to a protein termed ABIN-2 (NM_024309), which has been isolated 92 times from six different cDNA libraries. ABIN-2 has been described as an inhibitor of the Iκ-B kinase complex (24, 25). Recently it has been reported that ABIN-2 has transcriptional activating functions and may be considered as a novel transcriptional cofactor (26). In our screens, ABIN-2 has also been isolated with the cofactor TSG101 as a bait as well as with NCoA62/SKIP, albeit only once. Its high prevalence in screens with SMRT and its isolation with two other nuclear receptor-binding proteins warrants further examination of the role of this protein in transcriptional regulation by nuclear receptors.

FOXO Transcription Factors—The FOXO family of transcription factors is important in the regulation of cell proliferation and survival. Inhibition of FOXO proteins by insulin-, androgen- or transforming growth factor-β-dependent signal transduction events is crucial in the prevention of apoptosis and the initiation of cell proliferation (27). Previously functional interactions between FOXO1a/FKHR and nuclear receptors (28–30) have been reported.

In our screens, we have isolated FOXO1a as a low confidence interactor of NR1C2/PPARβ, NR2A2/HNF4γ, and NR3C2/MR-LBD (not shown). Interestingly the related protein FOXO3a has been isolated as a high confidence interactor of NR1H3/LXRα and NR1H2/LXRβ but not with any other bait. The FOXO3a prey has been picked up in the screens only in presence of the synthetic LXR agonist T0901317 but not in the absence of ligand. A potential interaction of NR1H3/LXRα and NR1H2/LXRβ with FOXO3a would be expected to influence the activities of the transcription factors. Preliminary co-transfection experiments with full-length FOXO3a support this possibility.

DISCUSSION

The analysis of protein interaction networks in model organisms has shown that statistical parameters can be used for efficient enrichment of relevant interactions such as the detection of circular interaction patterns (*e.g.* Ref. 8). While the focus around a single family of proteins and the comparably low number of bait proteins in this project has prevented us from systematically applying all such methods, the higher depth of screening has allowed us to apply relatively simple but efficient criteria to select high confidence interactions. When all these criteria are applied to select a data set, pre-validated interactions are enriched to 78%. Based on the principle of "good neighborhood," it appears reasonable to assume that novel interactions found within a statistically selected data set consisting to 78% of prevalidated interactions are likely to be of biological relevance. Despite the dramatic enrichment for prevalidated interactions that can be achieved, it has to be mentioned that there is a caveat in the definition of the positive reference set: interactions that are published first tend to be the ones most easy to find such that clones that are found frequently in two-hybrid screens are more likely to be validated by publication than rare ones. For this reason, methods that enrich for a positive reference data set based on literature data are likely to enrich for frequently found interactions. At the same time, our enrichment methods will select against preys that are expressed in only a few tissues and/or at a low level. Thus, we expect that many interactions of biological significance can be found among the low confidence interactions. Enlargement of the data set and refinement of selection methods will allow some of the protein pairs missed in the present state of the analysis to be pinpointed.

None of the enrichment methods were efficient in depleting interactions from the negative data set. This is most likely a reflection of their being not technical artifacts of the Y2H method but biological artifacts that are caused by the co-expression of two proteins that indeed display affinity for each other but never meet unless artificially co-expressed. Within the limits of the Y2H method or any other method that measures the affinity of two proteins, these artifacts cannot be expected to be efficiently identified.

We publish this data set with the goal of instigating further research on potential novel NR-interacting proteins. Together with other functional genomic approaches, the supply of systematic data sets to the research community will both instigate hypothesis-driven research as well as provide the data basis for integrative approaches to cell biology.

REFERENCES

1. Giguere, V. (1999) Orphan nuclear receptors: from gene to function. *Endocr. Rev.* **20,** 689–725

2. Schwabe, J. W., and Teichmann, S. A. (2004) Nuclear receptors: the evolution of diversity. *Sci. STKE* http://stke.sciencemag.org/cgi/content/full/OC_sigtrans;stke.2172004pe4

3. Baek, S. H., and Rosenfeld, M. G. (2004) Nuclear receptor coregulators: their modification codes and regulatory mechanism by translocation. *Biochem. Biophys. Res. Commun.* **319,** 707–714

4. McKenna, N. J., and O'Malley, B. W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108,** 465–474

5. Smith, C. L., and O'Malley, B. W. (2004) Coregulator function: a key to understanding tissue specificity of selective receptor modulators. *Endocr. Rev.* **25,** 45–71

6. Albert, S., Gaudan, S., Knigge, H., Raetsch, A., Delgado, A., Huhse, B., Kirsch, H., Albers, M., Rebholz-Schuhmann, D., and Koegl, M. (2003) Computer-assisted generation of a protein-interaction database for nuclear receptors. *Mol. Endocrinol.* **17,** 1555–1567

7. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403,** 623–627

8. Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, and J., Rothberg, J. M. (2003) A protein interaction map of Drosophila melanogaster. *Science* **302,** 1727–1736

9. Fromont-Racine, M., Rain, J. C., and Legrain, P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat. Genet.* **16,** 277–282

10. Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M. (2004) A map of the interactome network of the metazoan C. elegans. *Science* **303,** 540–543

11. Deane, C. M., Salwinski, L., Xenarios, I., and Eisenberg, D. (2002) Protein interactions: two methods for assessment of the reliability of high-throughput observations. *Mol. Cell. Proteomics* **1,** 349–356

12. Vidalain, P. O., Boxem, M., Ge, H., Li, S., and Vidal, M. (2004) Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* **32,** 363–370

13. McDonnell, D. P., and Norris, J. D. (2002) Connections and regulation of the human estrogen receptor. *Science* **296,** 1642–1644

14. El-Guendy, N., and Rangnekar, V. M. (2003) Apoptosis by Par-4 in cancer and neurodegenerative diseases. *Exp. Cell Res.* **283,** 51–66

15. Hong, Y. K., and Detmar, M. (2003) Prox1, master regulator of the lymphatic vasculature phenotype. *Cell Tissue Res.* **314,** 85–92

16. Steffensen, K. R., Holter, E., Bavner, A., Nilsson, M., Pelto-Huikko, M., Tomarev, S., and Treuter, E. (2004) Functional conservation of interactions between a homeodomain cofactor and a mammalian FTZ-F1 homologue. *EMBO Rep.* **5,** 613–619

17. MacDonald, P. N., Dowd, D. R., Zhang, C., and Gu, C. (2004) Emerging insights into the coactivator role of NCoA62/SKIP in vitamin D-mediated transcription. *J. Steroid Biochem. Mol. Biol.* **89–90,** 179–186

18. Zhou, D., and Chen, S. (2001) PNRC2 is a 16 kDa coactivator that interacts with nuclear receptors through an SH3-binding motif. *Nucleic Acids Res.* **29,** 3939–3948

19. Bai, R. Y., Koester, C., Ouyang, T., Hahn, S. A., Hammerschmidt, M., Peschel, C., and Duyster, J. (2002) SMIF, a Smad4-interacting protein that functions as a co-activator in TGFβ signalling. *Nat. Cell Biol.* **4,** 181–190

20. Pan, Z. Q., Kentsis, A., Dias, D. C., Yamoah, K., and Wu, K. (2004) Nedd8 on cullin: building an expressway to protein destruction. *Oncogene* **23,** 1985–1997

21. Dace, A., Zhao, L., Park, K. S., Furuno, T., Takamura, N., Nakanishi, M., West, B. L., Hanover, J. A., and Cheng, S. (2000) Hormone binding induces rapid proteasome-mediated degradation of thyroid hormone receptors. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 8985–8990

22. Lonard, D. M., Nawaz, Z., Smith, C. L., and O'Malley, B. W. (2000) The 26S proteasome is required for estrogen receptor-α and coactivator turnover and for efficient estrogen receptor-α transactivation. *Mol. Cell* **5,** 939–948

23. Shi, Y., Downes, M., Xie, W., Kao, H. Y., Ordentlich, P., Tsai, C. C., Hon, M., and Evans, R. M. (2001) Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.* **15,** 1140–1151

24. Liu, W. K., Yen, P. F., Chien, C. Y., Fann, M. J., Su, J. Y., and Chou, C. K. (2004) The inhibitor ABIN-2 disrupts the interaction of receptor-interacting protein with the kinase subunit IKKγ to block activation of the transcription factor NF-κB and potentiate apoptosis. *Biochem. J.* **378,** 867–876

25. Van Huffel, S., Delaei, F., Heyninck, K., De Valck, D., and Beyaert, R. (2001) Identification of a novel A20-binding inhibitor of nuclear factor-κB activation termed ABIN-2. *J. Biol. Chem.* **276,** 30216–30223

26. Chien, C. Y., Liu, W. K., Chou, C. K., and Su, J. Y. (2003) The A20-binding protein ABIN-2 exerts unexpected function in mediating transcriptional coactivation. *FEBS Lett.* **543,** 55–60

27. Accili, D., and Arden, K. C. (2004) FoxOs at the crossroads of cellular metabolism, differentiation, and transformation. *Cell* **117,** 421–426

28. Li, P., Lee, H., Guo, S., Unterman, T. G., Jenster, G., and Bai, W. (2003) AKT-independent protection of prostate cancer cells from apoptosis mediated through complex formation between the androgen receptor and FKHR. *Mol. Cell. Biol.* **23,** 104–118

29. Schuur, E. R., Loktev, A. V., Sharma, M., Sun, Z., Roth, R. A., and Weigel, R. J. (2001) Ligand-dependent interaction of estrogen receptor-α with members of the forkhead transcription factor family. *J. Biol. Chem.* **276,** 33554–33560

30. Zhao, H. H., Herrera, R. E., Coronado-Heinsohn, E., Yang, M. C., Ludes-Meyers, J. H., Seybold-Tilson, K. J., Nawaz, Z., Yee, D., Barr, F. G., Diab, S. G., Brown, P. H., Fuqua, S. A., and Osborne, C. K. (2001) Forkhead homologue in rhabdomyosarcoma functions as a bifunctional nuclear receptor-interacting protein with both coactivator and corepressor functions. *J. Biol. Chem.* **276,** 27907–27912