High-resolution protein–protein interaction mapping using all-*versus*-all sequencing (AVA-Seq)

Received for publication, April 7, 2019, and in revised form, June 9, 2019 Published, Papers in Press, June 10, 2019, DOI 10.1074/jbc.RA119.008792

Simeon S. Andrews[‡1], **Stephanie Schaefer-Ramadan**[‡1], **Nayra M. Al-Thani**[‡], **Ikhlak Ahmed**[‡], **Yasmin A. Mohamoud**[§], **and Joel A. Malek**[‡§2]

From the [‡]*Department of Genetic Medicine and* [§]*Genomics Laboratory, Weill Cornell Medicine in Qatar, Doha 24144, Qatar*

Edited by Karin Musier-Forsyth

Two-hybrid systems can be used for investigating protein–protein interactions and may provide important information about gene products with unknown function. Despite their success in mapping protein interactions, two-hybrid systems have remained mostly untouched by improvements in next-generation DNA sequencing. The two-hybrid systems rely on one-*versus*-all methods in which each bait is sequentially screened against an entire library. Here, we developed a screening method that joins both bait and prey as a convergent fusion into one bacterial plasmid vector that can then be amplified and paired-end sequencing by next-generation sequencing (NGS). Our method enables all-*versus*-all sequencing (AVA-Seq) and utilizes NGS to remove multiple bottlenecks of the two-hybrid system. AVA-Seq allows for high-resolution protein–protein interaction mapping of a small set of proteins and has the potential for lower-resolution mapping of entire proteomes. Features of the system include ORF selection to improve efficiency, high bacterial transformation efficiency, a convergent fusion vector to allow paired-end sequencing of interactors, and the use of protein fragments rather than full-length proteins to better resolve specific protein contact points. We demonstrate the system's strengths and limitations on a set of proteins known to interact in humans and provide a framework for future large-scale projects.

Methods of studying protein–protein interactions can broadly be categorized as one-*versus*-one or one-*versus*-all studies. The goal of this study was to develop and apply a novel methodology that allows screening in an all-*versus*-all fashion to compare complex libraries. Here, we have applied our method to define the interacting regions of a set of human proteins with high resolution.

Yeast and bacterial two-hybrid screens and their derivatives have long been a staple of large-scale protein interaction mapping (1, 2). Although there are multiple forms of two-hybrid systems, the most common approaches fuse a protein of interest ("bait") to a DNA-binding domain (DBD)[3] and test it against a library of proteins ("prey") that are fused to a transcription activation domain (AD). If the bait and prey proteins interact, they drive the overproduction of a survival gene by bringing the AD and DBD into proximity (Fig. 1). The interacting proteins are then identified by sequencing bait and prey DNA from proliferating colonies. In the case of the bacterial two-hybrid (B2H) system used here, the bait protein is encoded as a C-terminal fusion of the λcI protein (DBD), whereas the prey is encoded as a C-terminal fusion of the RNA polymerase α subunit (AD). If the bait and prey interact, the DBD and AD can jointly drive production of HIS-3, a gene that is essential for growth in histidine-free medium and that can be inhibited by the small molecule 3-amino-1,2,4-triazole (3-AT).

Although two-hybrid screens have been used for many years, only recently has next-generation sequencing (NGS) been applied to facilitate it (3–5). However, all these methods still require the use of one-*versus*-all methodologies: each bait must be sequentially screened against the entire library. Based on the B2H system by Dove and Hochschild (2), we have developed a new method that joins both bait and prey as a convergent fusion into one vector called pAVA. This all-*versus*-all sequencing (AVA-Seq) system is amenable to the power of NGS technology and allows for high-resolution mapping of protein interaction sites.

Results

AVA-Seq overview

The pAVA system is unique in that a two-hybrid screening of a small subset of proteins or even an entire genome can be completed in a relatively short amount of time and can take advantage of the power of NGS. Fig. 2 illustrates the AVA-Seq system. First, the protein-encoding DNA is sheared into random fragments and is size-selected (Fig. 2*a*). These fragments are then ligated into two different pBORF vectors (pBORF-DBD and pBORF-AD) to allow for open reading frame (ORF) filtering using carbenicillin (a more stable alternative to ampi-

This work was supported by a Qatar Foundation Biomedical Research Program (BMRP) grant and a Weill Cornell Medicine in Qatar pilot grant. The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1 and S2 and Tables S1–S3.

The data have been deposited with links to BioProject accession number PRJNA542588 in the DDBJ BioProject database.

[1] Both authors share first authorship.

[2] To whom correspondence should be addressed: Dept. of Genomic Medicine, Weill Cornell Medicine in Qatar, Doha 24144, Qatar. Tel.: 974-4492-8420; E-mail: jom2042@qatar-med.cornell.edu.

[3] The abbreviations used are: DBD, DNA-binding domain; AD, activation domain; B2H, bacterial two-hybrid; 3-AT, 3-amino-1,2,4-triazole; NGS, next-generation sequencing; AVA-Seq, all-*versus*-all sequencing; fs, frameshift(s); JUN, c-Jun; FOS, c-Fos; PKAR/PKA, cAMP-dependent protein kinase regulatory subunit 2; AKAP5, protein kinase A–anchoring protein 5; MDM2, murine double minute 2; KAN, kanamycin; CB, carbenicillin; FDR, false-discovery rate; KanR, kanamycin-resistance gene; AmpR, ampicillin-resistance gene; CAM, chloramphenicol; AA, amino acid.

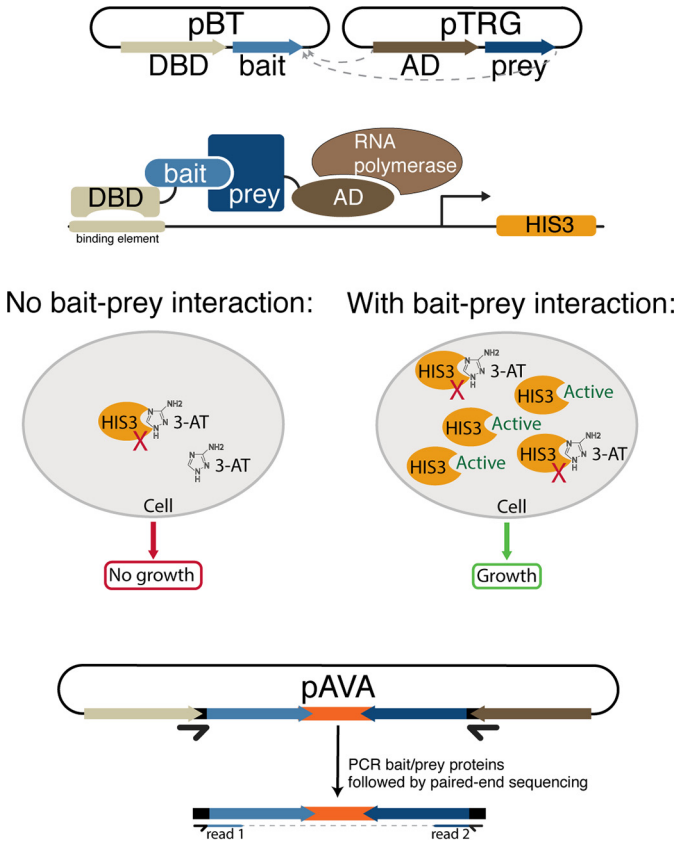

Figure 1. The AVA-Seq modification to two-hybrid systems. Normal two-hybrid screening uses separate bait and prey vectors. If the bait and prey proteins interact when expressed in a bacterial cell, they drive production of a HIS-3 survival gene. The presence of the HIS-3 inhibitor 3-AT silences background expression of HIS-3, providing selective growth advantage to clones with interacting proteins to drive HIS-3 overproduction. pAVA can be constructed by moving the activating domain portion of the prey vector into the bait vector. Once bait and prey proteins are cloned in with a stop-codon linker, the identity of the proteins can be determined by paired-end sequencing.

cillin) (Fig. 2*b*). The ORF-filtered fragments are amplified with primers allowing tail-to-tail orientation with a linker region with stop codons in all six frames (Fig. 2*c*) and fused using overlap extension PCR (Fig. 2*e*). Amplification of the fused fragments allows for cloning into the convergent-fusion vector pAVA that contains the DBD and AD central to the two-hybrid system (Fig. 2*f*). This is followed by liquid growth in nonselective or selective conditions (Fig. 2*g*) where interacting fragments have a growth advantage and become a larger fraction of the population. These fragments undergo paired-end sequencing in batches of several million to identify changes in the population of paired fragment that signify interactors (Fig. 2*h*). This allows for high-resolution snapshots of the interacting domains.

Convergent-fusion method validation

Initial validation of the pAVA convergent-fusion system was conducted with two positive and three negative controls. Growth (OD_{600}) was measured following a 9-h incubation in minimal media in the presence of 0, 2, or 5 mM 3-AT, a competitive inhibitor of the HIS-3 gene (2). The pAVA vector with no insert (Fig. 3, *a* and *f*) has strong growth in the unselected (absence of the DMSO vehicle and 3-AT) and 0 mM 3-AT conditions and minimal growth in the presence of 2 and 5 mM 3-AT. To create a positive control of known interactors, Gal11p and LGF2 yeast protein fragments were amplified from the original B2H (2) and inserted into the pAVA vector in convergent orientation. The schematics of the convergent-fusion positive control in both orientations are shown in Fig. 3, *b* and *c*. Both constructs show robust growth in 2 and 5 mM 3-AT (Fig. 3, *g* and *h*) irrespective of fragment orientation (adjacent to DBD or AD), which indicates a strong interaction between the Gal11p and LGF2, as expected.

To generate negative controls for system validation, nucleotides were inserted at the beginning of the LGF2 sequence (Fig. 3*d*) or at the beginning of both Gal11p and LFG2 sequences (Fig. 3*e*), which results in frameshifts (fs). Both negative controls show diminished growth as the concentration of 3-AT increased, indicating that the frameshifted proteins are not able to overcome the HIS-3 inhibition (Fig. 3, *i* and *j*). These results indicate that the pAVA system can detect the same protein–protein interaction in both fusion orientations.

Application to six known interacting proteins

After successful validation using yeast Gal11p and LGF2 protein fragments, the pAVA system was tested for high-resolution interaction mapping with three pairs of human proteins known to interact *in vivo*. The first pair was transcription factors of the AP-1 complex c-Jun (JUN) and c-Fos (FOS) (6, 7). The second pair of interacting proteins was cAMP-dependent protein kinase regulatory subunit 2 (PKAR; referred to as PKA in the text) and protein kinase A–anchoring protein 5 (AKAP5) (8, 9). The final pair was the human tumor suppressor protein p53 (TP53; referred to as p53 in the text) and the negative regulator of p53 known as murine double minute 2 (MDM2) (10, 11).

Open reading frame selection

These six human genes were amplified via PCR, pooled together in equimolar amounts, and sheared into ~450-bp fragments (Fig. 2*a*). A portion of the ligation was subjected to ORF filtering, whereas the remaining portion was not (non-ORF). For ORF selection, the fragments were ligated into pBORF vectors (AD and DBD) that are modified pBluescript II SK(+) vectors conferring kanamycin (KAN) resistance and containing a split AmpR (β-lactamase) gene. The fragments of interest are cloned between the localization signal and enzyme-coding portion of AmpR. Fragments that are in-frame and do not encode stop codons will permit expression of AmpR and survival in the presence of both KAN and carbenicillin (CB; β-lactamase expression). Most fragments of 450 bp with no stop codon are physiological open reading frames. pBORF-AD and -DBD differ only in the N-terminal sequence, which permits use of nonidentical primers in the later step of tail-to-tail fusion PCR. The libraries of fragments in pBORF were challenged with KAN and CB using a method similar to Di Niro *et al.* (12). Fragments cloned in the pBORF-DBD and pBORF-AD resulted in 1.3 and 4.8% survival on KAN/CB, respectively, relative to growth on KAN alone. The over 95% drop in clones between the two antibiotics is indication of ORF selection.

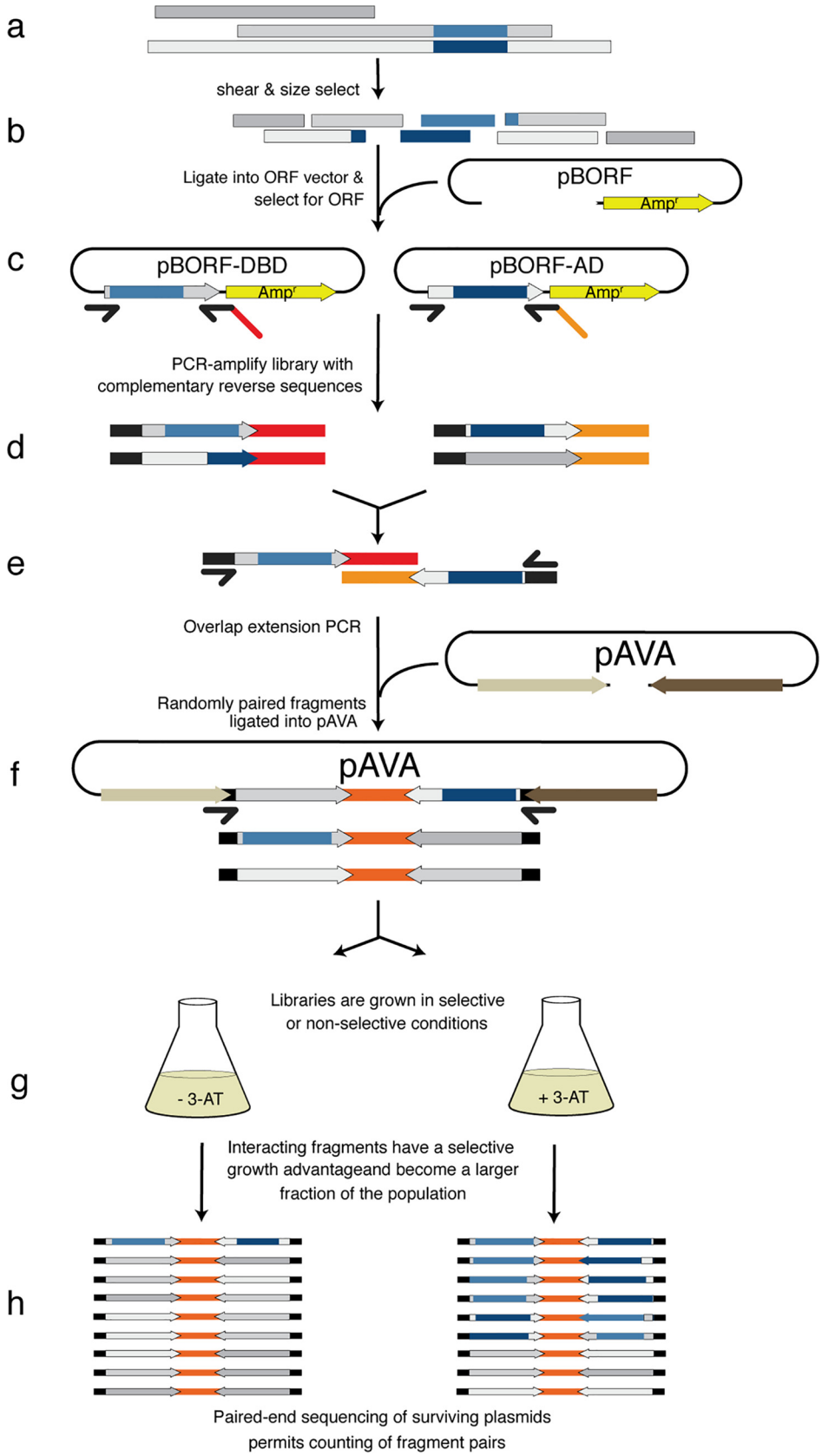

Figure 2. The AVA-Seq workflow. To screen a library of protein-encoding DNA (interacting regions colored *darker*), the protein-encoding DNAs are sheared and size-selected (*a*) and cloned upstream of a AmpR in the pBORF vectors (*b*). Growth on carbenicillin-containing medium selects for ORF-containing fragments. These can be amplified with complementary reverse primers (*d*), which allows for overlap extension PCR to create a fragment pair (*e*) for cloning into pAVA (*f*). *g*, liquid culture growth in the presence of 3-AT creates a growth advantage for bacteria carrying interacting pairs. *h*, paired-end sequencing of millions of inserts allows counting to determine which pairs have increased as a fraction of the population.

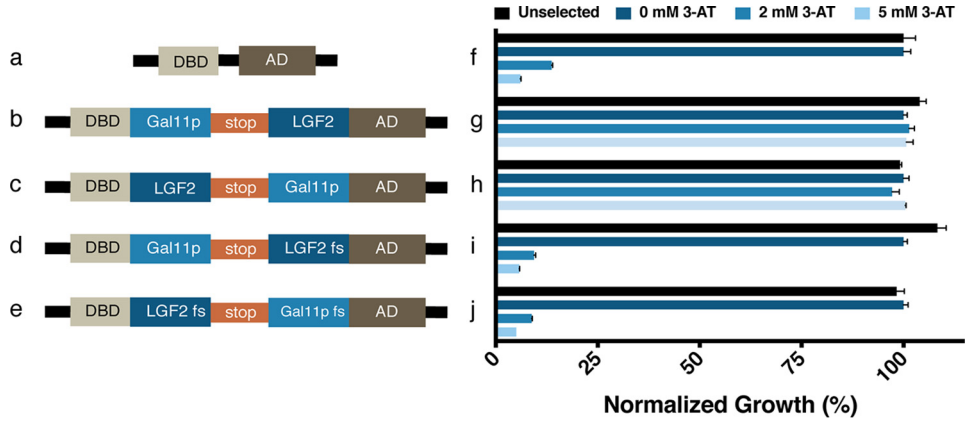

Figure 3. Schematic and growth of controls used in this study. *a*, pAVA with no DNA insert (negative control). *b*, pAVA with Gal11p and LGF2 (positive control). *c*, pAVA with LGF2 and Gal11p (positive control). *d*, pAVA with Gal11p and LGF2 that has been frameshifted by the insertion of one nucleotide (negative control). *e*, pAVA with one-nucleotide insertion for both LGF2 and Gal11p that results in frameshifts (negative control). *f–j*, competitive inhibition of HIS-3 gene using 3-AT. Unselected samples are grown in the absence of 3-AT and DMSO. 0, 2, and 5 mM 3-AT samples have increasing amounts of 3-AT and decreasing amounts of DMSO. Growth charts *f–j* correspond to schematics *a–e*, respectively. Data represent an average of three replicates with *error bars* representing S.E. The optical density (OD_{600}) was normalized to 0 mM growth after 9 h of expression.

Sequencing and analysis of the ORF-selected fragments revealed that on average 48% of vectors from pBORF-AD contained in-frame fragments of one of the six known proteins used in this study. This compared with ~16% in the non-ORF–selected libraries in the same vector. ORF selection was less efficient for fragments in pBORF-DBD with ~24% being in-frame *versus* 13% for the non-ORF–selected libraries.

Convergent fragment generation and interaction screening

The ORF-selected fragments in the pBORF-AD and pBORF-DBD were amplified, "stitched" together using overlap extension PCR, and inserted into pAVA (Fig. 2, *d* and *e*). When convergent-fusion fragments were considered, ~19% of paired products analyzed contained both DBD- and AD-fused fragments expressed in-frame with one of the six human proteins compared with only 1.5% in the non-ORF–selected libraries. ORF selection provides at least a 12-fold improvement in screening efficiency over non-ORF–selected libraries, especially given that both fusion fragments are required to be in-frame for interaction. The convergent-fusion fragments in the pAVA vector were transformed into BacterioMatch II electrocompetent reporter cells (2) using a maximum of 2 ng of DNA to ensure that only one vector was present per cell (13). To screen for interactions, the fragments were then challenged in triplicate with 0, 2, or 5 mM 3-AT (Fig. 2*g*) in histidine-free media. The same process was repeated for the non-ORF fragments.

Library construction and sequencing

Libraries were constructed using the NEBNext Ultra II DNA Library Prep kit for Illumina according to the manufacturer's protocol. A total of nine libraries, three replicates for each 0, 2, and 5 mM 3-AT selection, were combined for sequencing using MiSeq. A total of 6.1 million paired-end sequences were generated for which both the DBD and AD sequences were of high enough quality to allow translation of the fused fragments. This resulted in ~680,000 paired reads for each replicate. As a control for the selection and sequencing analysis process, the preconstructed positive control with Gal11p and LFG2 domains (Fig. 3*b*) was spiked in at low concentration prior to the 3-AT selection.

Sequence analysis

Paired-end sequencing reads were translated in-frame with the DBD or AD fragments they were fused to. Sequencing primers sit upstream of the of DBD- or AD-specific sequence allowing enough sequence (~150 bp) downstream to identify the fused fragment and whether the fusion is to DBD or AD. The translated sequences were then aligned to a database of the six human proteins with BLASTP. The gene ID and starting amino acid position to which a fragment aligned were documented and considered as a unique identifier. Paired sequences that revealed that both fused fragments were in-frame with a known protein were then carried forward. This process was repeated for all replicates in the analysis, and the results were combined in a table of counts for each unique fragment pair in each replicate. We observed a total of 146,531 unique fragment pairs (distinct protein/amino acid start point combinations) detected in any of the replicates of the ORF-selected libraries and 10,564 in replicates of the non-ORF–selected libraries. The ORF-selected libraries had ~120,000 paired-end, in-frame read counts per replicate distributed across the 146,531 unique fragment pairs. The non-ORF–selected libraries had ~6,500 paired-end, in-frame read counts per replicate distributed across the 10,564 unique fragment pairs.

Differential growth in the higher concentrations of 3-AT *versus* 0 mM 3-AT is an indication of a potential protein interaction. Fragment pairs were then tested for a statistically significant increase in the number of read counts, after normalization, in the 2 and 5 mM 3-AT replicates using DESeq2 (14). A fold-change cutoff (based on read counts) of at least 3 and a false-discovery rate of less than 5% was applied. As expected, the positive control (Gal11p-LGF2) showed a highly significant increase in read counts in selective media (Tables 1 and 2).

Analysis of read counts for fragment pairs with at least one observed read in the libraries with selection *versus* 0 mM 3-AT libraries showed a Pearson product-moment correlation (r) of 0.67, 0.97, and 0.97 for 0, 2, and 5 mM replicates, respectively

Table 1

Significant interaction pairs of ORF-filtered fragments: 2 versus 0 mM 3-AT

Significantly interacting ORF-filtered fragment pairs are listed with the gene name of protein 1:starting amino acid of the fragment:gene name of protein 2:starting amino acid of the fragment. The first protein in a fragment pair was fused to DBD, whereas the second was fused to AD. Only fragment pairs that show a positive $\log_2$ -fold change with a *p*-adjusted (FDR) <0.05 in the presence of 2 mM 3-AT when compared with 0 mM 3-AT were deemed as significantly interacting.

Fragment pair	$\log_2$ -fold change	*p* value	*p*-adjusted
JUN:275:FOS:172	7.15	2.61E−68	2.23E−64
JUN:282: FOS:172	8.24	1.16E−52	4.95E−49
JUN:275: FOS:96	6.93	6.37E−45	1.82E−41
JUN:282: FOS:183	8.13	8.67E−36	1.48E−32
JUN:282: FOS:96	7.22	7.92E−36	1.48E−32
JUN:282: FOS:163	8.01	6.58E−29	9.39E−26
JUN:281:FOS:95	6.48	2.37E−24	2.89E−21
JUN:282:FOS:181	5.99	6.19E−23	6.62E−20
JUN:282:FOS:58	4.83	1.11E−22	1.06E−19
AKAP5:366:JUN:23	5.36	3.87E−20	3.32E−17
JUN:281:FOS:181	5.96	5.23E−20	4.07E−17
JUN:282:FOS:164	5.47	6.28E−15	4.48E−12
JUN:282:FOS:94	5.44	1.09E−13	7.18E−11
PKA:174:FOS:73	3.46	1.47E−13	9.00E−11
JUN:282:FOS:174	5.50	1.93E−11	1.10E−08
JUN:281:FOS:59	5.25	2.85E−10	1.53E−07
JUN:282:FOS:173	5.18	6.42E−10	3.23E−07
JUN:282:FOS:83	5.05	2.05E−09	9.75E−07
JUN:275:FOS:173	4.64	2.28E−09	1.03E−06
Gal11p:59:LGF2:271	4.87	1.06E−08	4.55E−06
AKAP5:351:PKA:14	4.84	1.38E−08	5.62E−06
p53:256:p53:160	3.21	3.08E−08	1.20E−05
PKA:6:AKAP5:347	4.73	4.03E−08	1.50E−05
JUN:282:FOS:161	4.81	7.12E−08	2.54E−05
AKAP5:180:JUN:45	2.75	1.02E−07	3.49E−05
FOS:51:MDM2:259	3.42	4.56E−07	0.000150037
JUN:282:FOS:195	4.03	5.80E−07	0.000183903
JUN:282:FOS:95	4.31	1.12E−06	0.000342329
p53:157:AKAP5:150	1.30	3.45E−06	0.001018361
FOS:181:JUN:287	4.14	3.81E−06	0.001087944
FOS:171:p53:297	2.53	4.50E−06	0.001242992
AKAP5:235:JUN:23	3.75	7.10E−06	0.001899643
AKAP5:368:FOS:204	2.01	9.03E−06	0.002342664
AKAP5:348:PKA:3	3.93	1.60E−05	0.004016045
AKAP5:1:FOS:56	3.57	2.51E−05	0.006129581
JUN:282:FOS:165	3.78	3.91E−05	0.009292876
AKAP5:291:PKA:4	3.76	4.24E−05	0.009812216
JUN:29:p53:322	3.73	5.09E−05	0.011459083
FOS:167:p53:317	2.38	6.78E−05	0.014879163
FOS:31:AKAP5:114	3.69	7.08E−05	0.01515947
JUN:78:p53:120	1.64	7.37E−05	0.015386613
JUN:275:FOS:163	3.51	0.00017438	0.035540297

Table 2

Significant interaction pairs of ORF-filtered fragments: 5 versus 0 mM 3-AT

Significantly interacting ORF-filtered fragment pairs are listed with the gene name of protein 1:starting amino acid of the fragment:gene name of protein 2:starting amino acid of the fragment. The first protein in a fragment pair was fused to DBD, whereas the second was fused to AD. Only fragment pairs that show a positive $\log_2$ -fold change with a *p*-adjusted (FDR) <0.05 in the presence of 5 mM 3-AT when compared with 0 mM 3-AT were deemed as significantly interacting.

Fragment pair	$\log_2$ -fold change	*p* value	*p*-adjusted
JUN:275:FOS:172	6.71	8.35E−71	1.45E−66
JUN:282:FOS:172	8.47	3.11E−56	2.70E−52
JUN:275:FOS:96	6.57	1.45E−42	8.39E−39
JUN:282:FOS:183	8.23	9.11E−38	3.95E−34
JUN:282:FOS:164	6.38	4.49E−37	1.56E−33
JUN:282:FOS:163	8.42	1.01E−33	2.93E−30
JUN:282:FOS:96	6.20	8.87E−26	2.20E−22
JUN:282:FOS:181	6.10	8.59E−25	1.86E−21
AKAP5:366:JUN:23	4.97	2.64E−17	4.58E−14
JUN:281:FOS:95	5.58	2.45E−17	4.58E−14
AKAP5:351:PKA:14	6.16	1.90E−15	3.00E−12
JUN:282:FOS:58	4.06	4.21E−15	6.09E−12
Gal11p:59:LGF2:271	5.88	7.76E−14	1.04E−10
JUN:281:FOS:181	4.99	1.99E−13	2.46E−10
JUN:282:FOS:174	5.62	1.73E−12	2.00E−09
JUN:282:FOS:94	5.15	1.91E−12	2.07E−09
JUN:282:FOS:161	5.55	5.53E−12	5.64E−09
JUN:282:FOS:173	5.25	1.19E−10	1.15E−07
JUN:282:FOS:83	4.84	7.56E−09	6.90E−06
JUN:275:FOS:173	4.37	1.57E−08	1.36E−05
JUN:281:FOS:162	4.42	2.84E−07	0.00023425
JUN:281:FOS:59	4.41	3.22E−07	0.000253657
AKAP5:348:PKA:3	4.24	1.22E−06	0.000920409
JUN:275:FOS:163	4.03	5.66E−06	0.004091383
JUN:282:FOS:165	3.82	2.07E−05	0.014374318
AKAP5:188:p53:299	2.36	3.82E−05	0.025509153
p53:132:FOS:41	2.28	7.50E−05	0.048172968
PKA:50:PKA:154	2.83	7.92E−05	0.049071126

(Fig. S1). The lower correlation among the 0 mM replicates is likely due to the fragment pairs with significant numbers of reads in 2 or 5 mM 3-AT having very few reads in 0 mM 3-AT, making correlation less likely. We observed a correlation of −0.09 and −0.07 when 0 mM 3-AT libraries were compared with 2 and 5 mM 3-AT, respectively. Comparison of a library with no carrier (no DMSO, no 3-AT) *versus* the 0 mM 3-AT library showed correlation of 0.64, agreeing well with the correlation among 0 mM 3-AT replicates. This suggests that the 0 mM 3-AT–selected libraries do not appear to contain bias from the DMSO carrier. Whether the growth in minimal media creates an initial selective pressure that could result in bias of fragment content is not clear, but all comparisons in this study were controlled with 0 mM growth levels.

Protein interaction analysis

Using this method, we have tested at high-resolution 96.14% (5,686,520/5,914,624 pairwise amino acid combinations) of the possible interaction space when all libraries are considered (Fig. 4*a*). This was possible by generating protein fragments with multiple starting points. The tested interaction space was reduced to 5.6 (331,580/5,914,624) and 1.9% (111,597/5,914,624) of significantly interacting amino acids in the 2 and 5 mM concentrations of 3-AT, respectively. Fig. 4, *b* and *c*, show the refined interaction space of the 36 protein pairs (6 proteins × 6 proteins) used in this study. By looking at the change in interaction regions moving from 2 mM 3-AT (Fig. 4*b*) to an increase in selective pressure using 5 mM 3-AT (Fig. 4*c*), there are obvious changes in the interaction landscape. These data indicate the depth of information generated using mild and strong selective pressure (correlating, in theory, to mild and strong interactions).

Fig. 5 illustrates several of the most significant (*p*-adjusted < 0.05) fragment pairs of the expected interactions (Tables 1 and 2). The JUN–FOS protein fragments demonstrated the most abundant interaction of the 36 possible interaction combinations (not including the different starting amino acids) comprising 23 of 42 and 21 of 28 significant interactions in 2 and 5 mM, respectively (Tables 1 and 2). Additionally, in the non-ORF–filtered data, this interaction was 35 of 38 and 28 of 31 significant interactions in 2 and 5 mM, respectively (Tables S2 and S3). Importantly, the fragments align to the interaction regions for the proteins (Fig. 5, *light gray*) and show fragments in both orientations, demonstrating the level of detail this method affords (Tables 1 and 2). The JUN–FOS interaction also showed similar fragment pairs for the ORF-filtered and non-ORF–filtered libraries, which also attests to the strength of the interaction. The multiple interactions of JUN and FOS frag-

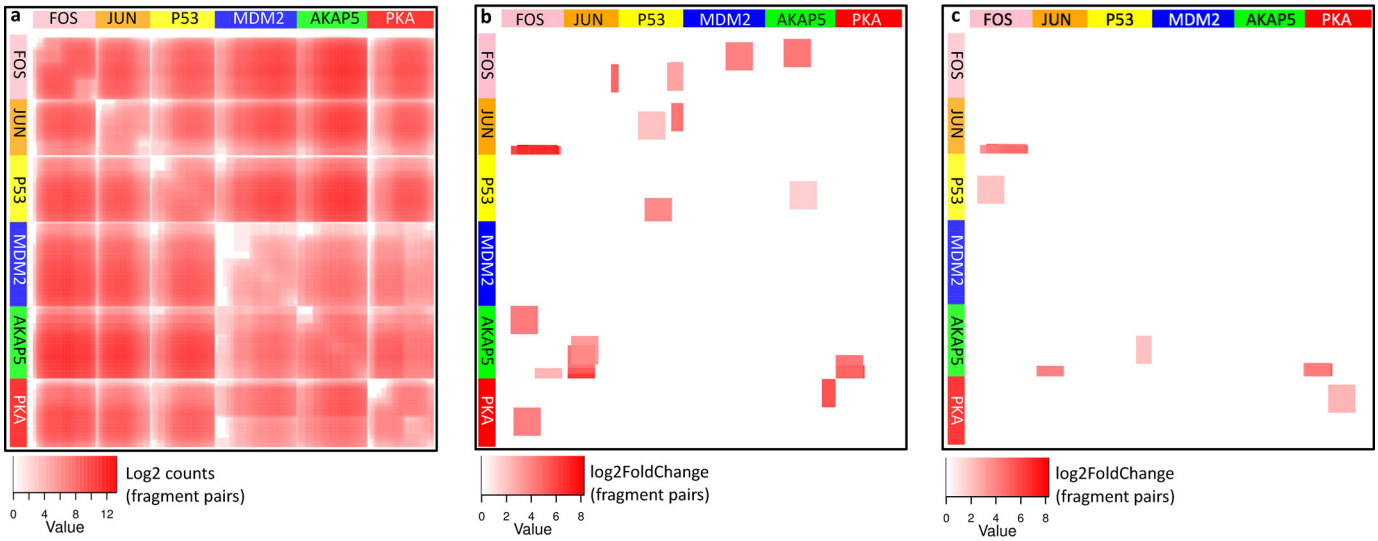

Figure 4. Matrix heat maps of all amino acid (AA) pairs tested or interacting in this study. *Colored panels* on the *top* and *left* show location of proteins in the matrix. *a*, detection of the AA pair (*red*) or no detection of the AA pair (*white*) in any of the libraries sequenced. *b* and *c* show $\log_2$-fold change in 2 and 5 mM libraries with respect to the 0 mM library for fragment pairs that show a statistically significant enrichment. Enriched pairs in *b* and *c* imply interaction between the corresponding regions of the proteins. Orientation of the tested pair is included to illustrate that not all interactions are symmetrical in the matrix; *i.e.* fusions to DBD are listed *vertically*, whereas fusions to AD are *horizontal*.

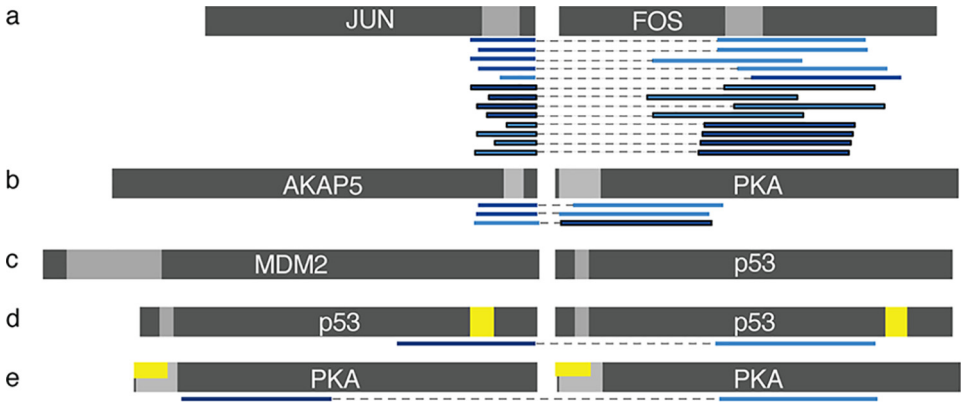

Figure 5. Significant interactions detected ($p < 0.05$) for expected protein pairs. *a*, interacting pairs of Jun and Fos. *b*, interacting pairs between AKAP5 and PKA. *c*, expected interaction regions of MDM2 and p53. *d*, p53 homodimer. *e*, PKA homodimer. The *light gray* region of the protein schematic indicates interaction regions reported in the literature (6, 9, 11, 15, 16). ORF-filtered fragments are shown with *no outline*, and the non-ORF–filtered fragments have a *black outline*. *Dark blue* fragments are associated with the AD, and the *light blue* fragments are associated with the DBD. *Yellow bars* indicate residues involved in homodimer association.

ments are consistent with the interaction sites reported in the literature (6, 15).

In addition, Fig. 5*b* shows that AKAP5 and PKA also had interacting fragment pairs in both orientations. For the ORF-filtered fragments, the significant interactions for AKAP5 and PKA comprised 4/23 and 2/28 significant interactions in 2 and 5 mM, respectively (Tables 1 and 2). An advantage to this fragment-based method, rather than more traditional full-length protein, is the ability to identify the interaction region(s) between the proteins. Looking at Fig. 5*b* more closely shows that alignment of the significant interaction pairs corresponds to the interaction regions described in the literature for AKAP5 and PKA as well (9).

Although the literature shows that p53 and MDM2 interact, the AVA-Seq method was not able to detect statistically significant interactions between the two proteins (Fig. 5*c*). This is likely due to two major reasons: first, ensuring sufficient coverage of the N- and C-terminal regions of gene fragments, and second, the complex nature of the interaction. First, it is possible we did not see the MDM2–p53 interaction because the interacting region was not well-covered by fragments (Fig. 4*a*). Fig. 4*a* shows an underrepresented portion of approximately the first one-third of the MDM2 sequence, which is exactly where the interaction region with p53 is located. The small portion of the N- and C-terminal sequences is underrepresented due to the difficulty in representing those regions during ORF filtering. We tried to minimize this by using primers that bind up- and downstream of the gene on the vector. Second, it is possible we did not see the MDM2–p53 interaction because the interaction site of MDM2 is large, and the interaction complex of MDM2–p53 is dependent primarily on van der Waals interactions, which is different from most identified proteins. Furthermore, interaction occurs in a buried surface that consists mostly of hydrophobic interaction between pseudosymmetry domains (11, 16). Fragments from these proteins were, however, detected as significant interactions with other pro-

teins (Fig. S2). Interestingly, there was a significant self-interaction for p53–p53 and PKA–PKA (Fig. 5, *d* and *e*, respectively). The *yellow bars* in Fig. 5, *d* and *e*, represent residues involved in the p53–p53 dimer interface (17) or PKA–PKA dimer interface (18). Two of the four fragments involved in the self-interaction align to these regions involved in the dimer interaction.

This method not only identified known self-dimerization interactions (Fig. 5, *d* and *e*), but it also shines light on possible novel interactions that have not been explored in the literature to the best of our knowledge (Fig. S2, Tables 1 and 2, and Tables S2 and S3). Several of the interactions indicate fragments in both orientations, increasing the probability of a real, strong interaction and worth investigating further. The AVA-Seq method is able to confirm at high resolution the interactions of two of three of the human proteins pairs as well as two homodimer interactions and other, possibly novel interactions among the six proteins.

Discussion

Protein–protein interaction data can provide important information for understanding how an individual protein functions and its system-wide role in the context of other proteins. Despite the importance, methods such as the two-hybrid system have not seen significant reductions in labor and cost since its first use. Although some improvements have been made using next-generation sequencing, deep screening or large-scale two-hybrid studies remain labor-intensive.

At the center of our approach to NGS based protein-interaction mapping is the convergent-fusion vector, pAVA. The novelty of pAVA is that it joins the traditionally individual bait and prey DNA sequences on a single DNA molecule, allowing them to be amplified and paired-end sequenced. This combined with the high transformation efficiency has allowed us to test almost the entire interaction space of the six proteins at high resolution (146,000 paired fragments tested). As with methods such as RNA-Seq, this system is limited simply by the diversity of the library and depth of sequencing. Higher resolution of interacting domains is achievable with deeper sequencing of diverse libraries. An important feature of the system is the ability to "dial-in" the level of selection by changing the concentration of 3-AT. Here, we used 2 and 5 mM concentrations, but this could be changed depending on the strength of the interactions being targeted. Our results show that many of the interactions detected in the more stringent 5 mM selection are also found in the 2 mM 3-AT selection, whereas some interactions are only detected in one of the conditions.

Although we expect weaker interactions to be detected in the 2 mM 3-AT conditions that subsequently drop out in the more stringent 5 mM conditions, we did indeed observe two interactions in 5 mM 3-AT that were not considered significant in the less stringent 2 mM 3-AT. One of those interactions (AKAP5:188:p53:299) was just above the significance cutoff in 5 mM, whereas it was just below in 2 mM 3-AT. The other interaction (p53:132:FOS:41) was not observed altogether in the 2 mM 3-AT sequencing. The fact that both interactions were just above the significance cutoff in 5 mM 3-AT may be an indication of why this occurs. It is important to note that increasing 3-AT will select for stronger interactions but that strong interactions do not necessarily distinguish themselves by increased growth at lower stringency such as in 0 or 2 mM 3-AT. That is, in 2 mM 3-AT many more interactions can be detected because weaker interactions can continue to grow in those conditions, competing with the stronger interactions. So, to detect these interactions more thoroughly would likely require even deeper sequencing of the 2 mM 3-AT to ensure that enough reads are distributed across the increased number of growing fragment pairs to allow more interacting fragments to achieve significance in the statistical tests. In contrast, in the more stringent 5 mM 3-AT, the weaker interactions fail to grow, and stronger interactions more rapidly take over the population of clones such that more sequencing reads are distributed across fewer interaction pairs. This allows the statistical tests to show significance, especially in the borderline interactions. This is a feature of conducting the screening in liquid culture *versus* selection of colonies, which is a binary decision process.

These results will be important for future "tuning" of the system to better understand what levels of cutoff and what depth of sequencing should be used and how a user of the system will make decisions on what strength of interaction to consider. It may be possible, in the future, to rank the strength of interactions in the system relative to the positive control (Gal11p-LGF2) as it is included in every experiment. Here, for example, the JUN–FOS fragment interactions in general were stronger relative to the positive control in the 2 mM 3-AT selection, whereas the PKA–AKAP5 fragments were weaker. By providing data from both 2 and 5 mM 3-AT, we hope that future users can be more informed on the benefits of multiple conditions for helping better understand the strength of the interactions.

We observed two of three known interactions among the six proteins tested. Those are the JUN–FOS and PKA–AKAP5 interactions. Homodimerization of p53 and PKA were also observed, albeit by one fragment pair each (Fig. 5, *d* and *e*). The p53–MDM2 interaction was not detected at a statistically significant level by the system, showing its limitation in detecting interactions between large domains (Fig. 5*c*). Frequent criticism of two-hybrid systems centers around the potential for false positives, or interactions that should not be detected, and false negatives, that is interactions that were missed. The AVA-Seq system employs various features to mitigate both types of errors. Although this does not guarantee that the interactions detected occur *in vivo*, it increases the likelihood that interactions are not spurious within the system or were simply missed. These features include interactions being observed in both orientations, that is with the bait and prey fragments fused to the DBD and AD and vice versa. Additionally, applying a requirement that multiple, overlapping fragments from the same genes be observed to interact decreases the possibility that the interaction is invalid. Autoactivators, those fragments that always activate expression of the selective marker, can be removed by filtering out those fragments that have extremely high numbers of interacting partners. Although current limitations of NGS fragment lengths require that in general we test fragments rather than full-length genes, we observed benefits from testing multiple fusions at various amino acids positions

for each protein. We noted that not all fragment pairs expected to interact were observed and that likely not every fusion point creates a functionally active protein to allow interaction (12). Thus, using multiple fusion points within a gene, rather than a full-length gene, may help overcome both false positives and false negatives. However, the limitation of fragment length was clear in the missed p53–MDM2 interaction, likely due to larger fragments being required to span the interaction region (11). Lastly, the ability to screen with multiple levels of selective pressure in a cost-effective manner should increase the chances of detecting a range of interactions.

The reduction of the potential interaction space by ~17- (2 mM) and 50-fold (5 mM) to a small portion of interactors is significant. It is even more significant a reduction when considering that three pairs of the proteins were expected to interact. The additional interactions detected in this study (Fig. S2) provide a starting point for future validation studies. Some of these interactions such as that between AKAP5 and p53 show interaction between multiple overlapping fragment pairs with at least one in the reverse-fusion orientation.

With the improvements we have made to the two-hybrid system, we envision two uses for AVA-Seq. The first, as demonstrated here, is the high-resolution protein–protein interaction mapping of a small set of proteins either all-*versus*-all or few-*versus*-all. This might include proteins of unknown function being tested against a whole cDNA library or those from a single pathway. The number of transformants and sequencing depth are readily achievable for high-resolution domain mapping. The second application will expand to large all-*versus*-all interaction screening for entire bacterial genomes and beyond. By utilizing the ORF selection process, the number of screening events have been drastically reduced while maintaining the benefits of fragment-based mapping described above. Indeed, using the high transformation efficiency of a single vector and deep sequencing, we believe whole-genome protein-interaction mapping could be achieved by small laboratory groups in a relatively short amount of time.

Experimental procedures

Design of pBORF vectors

The pBORF vector was designed by replacing the existing AmpR in the pBluescript II SK(+) vector (Stratagene, catalog number 212205) with a kanamycin-resistance gene (KanR) adjacent to the original AmpR promoter. The result is a pBluescript II SK(+) vector with kanamycin resistance. Next, the β-lactamase localization sequence and AmpR-encoding DNA were inserted in the multiple-cloning-site region of the modified pBluescript II SK(+) vector between XhoI and KpnI/Acc65I sites. Next, ~25-bp inserts were designed and inserted between the β-lactamase localization sequence and the AmpR gene to allow for ORF filtering and to differentiate pBORF vectors. pBORF-DBD is associated with DBD (λcI), and pBORF-AD is associated with AD (RNA polymerase alpha). Primer A and primer B were used to create an insert for pBORF-AD, and primer C and primer D were used to create an insert for pBORF-DBD. Primer pairs were subjected to an annealing program (95 °C for 2 min followed by slow cooling to 25 °C) and then diluted 1:1,000 from 50 μM to 50 nM. To prepare for ligation, the modified pBluescript vector was linearized with primers E and F. Ligation of the linearized vector was set up using 3:1 insert to vector. The resulting ligations created pBORF-DBD or pBORF-AD.

When needed, pBORF-DBD and pBORF-AD were linearly amplified the same day as ligation. The PCR used primers G and H for the pBORF-AD vector and primers G and I for the pBORF-DBD vector. The column-purified linearized pBORF-DBD and -AD vectors were then subjected to DpnI (New England Biolabs, R0176S) digestion followed by PCR column cleanup.

Design of pAVA

The pAVA vector was constructed from the BacterioMatch II two-hybrid system (Agilent) vectors pBT and pTRG. First, the AD domain in pTRG was amplified using primers that included XhoI and NotI restriction sites. Restriction digestion was performed for both pBT vector and the amplified AD PCR product using XhoI and NotI enzyme sites. Next, ligation was performed and resulted in the AD in convergent orientation with DBD in pBT vector. This new construct is referred to as pAVA (vector all-*versus*-all). The pAVA vector was linearized using primers N and O. Column cleanup was performed after amplification using a GenElute PCR Cleanup kit and then subjected to DpnI treatment using the standard protocol. To introduce BstXI restriction enzyme sites, primers P and Q were used. Ligation was performed with 3:1 insert to vector ratio using linearized pAVA vector digested with BstXI and BstXI-digested insert obtained from P and Q primer amplification.

pAVA controls

Gal11p and LGF2 sequences (sp P04386 GAL4_YEAST and sp P19659 MED15_YEAST, respectively) were amplified from the pBT-LGF2 and pTRG-Gal11p control vectors from the BacterioMatch II two-hybrid system and ligated into both pBORF-AD and pBORF-DBD. The final converging positive control constructs (pAVA-Gal11p-LGF2 and pAVA-LGF2-Gal11p) were made as described (Fig. 3, *b* and *c*). The first negative control is the empty vector (pAVA) containing AD and DBD without DNA insert (Fig. 3*a*). The second and third negative controls are the positive control constructs with the addition of one nucleotide to introduce a frameshift to LGF2 only (pAVA-Gal11p-LGF2(fs)) or both LGF2 and Gal11p (pAVA-Gal11p(fs)-LGF2(fs)) (Fig. 3, *d* and *e*). Each control vector was tested separately in the presence of 0, 2, and 5 mM 3-AT in DMSO as well as unselected sample that includes no DMSO or 3-AT as an additional control. All vectors (pBORF-AD (127462), pBORF-DBD (127463), pAVA (127464), pAVA-Gal11p-LGF2 (127480), pAVA-Gal11p-LGF2(fs) (127482), pAVA-LGF2-Gal11p (127481), and pAVA-LGF2(fs)-Gal11p(fs) (127483)) have been submitted to Addgene.

Amplification of six human genes

The following six human genes were purchased from Origene as 10-μg stocks in pCMV6-Entry vector with SgfI and MluI cloning sites: PKAR2 (RC220376; 1212 bp), AKAP5 (RC221314; 1281 bp), MDM2 (RC219518; 1491 bp), p53 (RC200003; 1179 bp), FOS (RC202597; 1140 bp), and JUN (RC209804; 993 bp). All sequences were confirmed by in-house

Sanger sequencing. Each gene in pCMV6-Entry was amplified separately using T7 and M13 reverse primers.

Shearing of DNA and ligation of fragmented genes into pBORF vectors

30 nM PKA, AKAP5, MDM2, p53, FOS, and JUN were pooled in a total volume of 50 μl and transferred to a Covaris microTUBE (Covaris, part number 520045, lot number 002563). The sample was sheared using the settings appropriate for 400-bp median size. Next, end repair (NEBNext DNA Library Prep Master Mix Set for Illumina, New England Biolabs, catalog number E6040S) was performed as directed. Blunt ligation of the end-repaired fragments into linearized pBORF-DBD and pBORF-AD vectors was performed using a molar ratio of 6:1 (insert to vector).

ORF filtering

Transformations were performed for each pBORF-AD and pBORF-DBD ligation product using CopyCutter EPI400™ electrocompetent *Escherichia coli* (Lucigen, catalog number C400EL10; 1.8 mV). The transformations for pBORF-DBD or pBORF-AD were pooled. To have an idea of how many unique fragments are represented in the library, dilutions were made and plated on LB-agar supplemented with KAN (30 μg/ml), allowing for the total number of colonies to be counted. The colonies on the KAN antibiotic alone represent the non-ORF fragments. Transformed cells plated directly on LB-agar plates with KAN (30 μg/ml) and CB (15 μg/ml) represent the ORF-filtered fragments. We typically see a 95–99% decrease in survival when comparing the number of non-ORF:ORF fragments.

Extracting DNA

Colonies grown on KAN were gently scraped and resuspended with sterile LB. The same was done for the colonies grown on KAN and CB. Approximately 100 μl of the resuspended colonies were pelleted, and DNA was extracted with a GenElute Plasmid MiniPrep (Sigma, catalog number PLN350). Samples were then subjected to a 0.7% agarose gel (110 V; 1 h) and gel-purified.

Extracting ORF-filtered DNA product from the pBORF vectors: overlap extension PCR

Extracting the ORFs from the vector was carried out by PCR. PCR amplification was performed using ORF-filtered DNA in pBORF-DBD or pBORF-AD vectors and primers J and K for pBORF-DBD or primers L and M for pBORF-AD. Next, overlap extension PCR was performed using the resulting products above plus primers J and L. The PCR cycle was as follows: 98 °C for 2 min, 12 × (98 °C 15 s, 57 °C 20 s, 72 °C 45 s), 72 °C 3 min, 4 °C hold. The resulting PCR product combined ORF-filtered pBORF-DBD and pBORF-AD fragments into one larger product. This whole process was repeated for the non-ORF using the same protocol.

Ligation of PCR product into pAVA

The gel-purified pAVA was linearized with BstXI using a standard protocol. Next, the samples were run on a 0.7% agarose gel, and the band was excised and extracted using a GenElute Gel Extraction kit. The PCR product from the overlap extension PCR was digested with BstXI in the same manner as above and run on a 1.5% agarose gel, and the bands between 700 and 1200 bp were excised and extracted.

The gel-purified, BstXI-digested products were ligated with a 1:6 vector to insert ratio with T4 DNA ligase (New England Biolabs) overnight followed by GenElute PCR cleanup. The purified ligation was then transformed into New England Biolabs Turbo electrocompetent cells and plated on LB-agar plates supplemented with 10 μg/ml chloramphenicol (CAM). The colonies were then scraped with LB as described above, and DNA was then extracted. A maximum of 2 ng of DNA/50 μl of cells ($\sim 2.8 \times 10^9$ cells) were transformed into the BacterioMatch II electrocompetent reporter cells (Stratagene, catalog number 200195) to ensure one vector per cell.

3-AT selection

The resulting colonies in the BacterioMatch II reporter strain were scraped with LB broth. 500 μl of the slurry were diluted into 5 ml of minimal media and centrifuged (5 min; 3,000 × *g*; room temperature). The supernatant was decanted, and the pellet was washed with another 5 ml of minimal media. This wash of the pellet was repeated for a total of four washes followed by resuspension in 5 ml of fresh minimal media. The OD_{600} of the cells was measured and diluted to $OD_{600} = 0.05$ in 75 ml of minimal media.

A 5-ml culture of the Gal11p-LGF2 (positive control in pAVA and VR) was inoculated with fresh colonies and grown in LB broth in the presence of 25 μg/ml CAM for 3 h. The sample was then centrifuged (5 min; 3,000 × *g*; room temperature) and washed identically to the slurry above. The final resuspension for the positive control was $OD_{600} = 0.05$ in 1 ml of minimal media. From this, a 1:1,000 dilution was made in 1 ml of minimal media. 7.5 μl were removed from the 75-ml cell mixture of the washed sample and replaced with 7.5 μl of the 1:1,000 dilution of the positive control for a final $OD_{600} = 5 \times 10^{-9}$ spike-in.

One 15-ml culture tube was set up for an unselected sample that contained 5 ml of cells ($OD_{600} = 0.05$ with the positive control spike-in) in minimal media. Three additional tubes with 5 ml of cells for each of the following at 0 mM (25 μl of DMSO), 2 mM (15 μl of DMSO and 10 μl of 1 M 3-AT), and 5 mM 3-AT (25 μl of 1 M 3-AT) for a total of 10 culture tubes. Cells were allowed to grow for 9 h at 250 rpm at 37 °C. After 9 h of growth, OD_{600} was measured. Samples were centrifuged at 3,000 × *g* for 5 min. DNA was extracted using GenElute MiniPrep.

Library construction

To begin library construction, interacting fragments from the 3-AT selection were amplified from the pAVA vector using primers R and S (Table S1). The sequencing libraries were prepared using NEBNext Ultra II DNA Library Prep kit (catalog number E7645S) according to the manufacturer's protocol. The library concentrations were determined using KAPA Library Universal Quantification kit (KK4827). Agilent High Sensitivity DNA kit (5067-4626) for Bioanalyzer electrophoresis was used to determine average fragment length. Final libraries were prepared as directed in the MiSeq System Denature and Dilute Libraries Guide (Illumina) and run using MiSeq V2 Reagent kit (300 cycles) (MS-102-2002).

Preparation of media and reagents for 3-AT selection

For each selection experiment, 100 ml of fresh minimal medium were made using the following recipe adapted from the BacterioMatch II system manual. In the following order, 2 ml of 20% glucose, 1 ml of 20 mM adenine HCl, and 10 ml of 10× His-dropout amino acid mixture (Clontech, 630415; autoclaved according to manufacturer's directions and stored at 4 °C until use) were combined. Then 100 μl of each of the following were added: 1 M $MgSO_4$, 1 M thiamine HCl, 10 mM $ZnSO_4$, 100 mM $CaCl_2$, and 50 mM isopropyl 1-thio-β-D-galactopyranoside. After mixing well, 76 ml of autoclaved Millipure water, 10 ml of 10× M9 salts, and 100 μl of 25 mg/ml CAM were added. To make 10× M9 salts, 14 g of disodium phosphate ($Na_2HPO_4{\cdot}7H_2O$), 6 g of potassium dihydrogen phosphate (KH_2PO_4), 1 g of sodium chloride (NaCl), and 2 g of ammonium chloride (NH_4Cl) were dissolved in 200 ml of Millipure water, filtered, autoclaved, and stored at room temperature. 1 M 3-AT stocks were made in 100% DMSO and stored at −20 °C for no more than 30 days. Adenine, isopropyl 1-thio-β-D-galactopyranoside, thiamine HCl, and 3-AT aliquots were used once, and any remaining was discarded.

Analysis of significant interactions

Higher growth in the presence of 3-AT at 2 mM and 5 mM concentrations *versus* 0 mM is indication of a potential protein–protein interaction. Statistical significance of differential growth for each comparison was evaluated from three replicates in each growth condition at a positive predictive value of 95% (FDR < 0.05) using DESeq2 (14). For each comparison, only those protein fragment pairs that were observed in at least four of the samples of six replicates were taken for differential growth analysis to estimate the significant differences. DESeq2 performs an internal normalization step where the geometric mean is calculated for each row across all samples, and counts in each sample are then divided by the mean. The median of the ratios in a sample is used as the size factor for that sample to correct for the library size and composition bias. Rows containing count outliers are automatically removed using Cook's distance. In addition, an optimization procedure further removes the fragment pairs with low counts by filtering the rows where mean of normalized counts is below a determined threshold. Finally, a negative-binomial generalized linear model is fitted to determine differential growth using the Wald test for significance testing, which computes *p* value and the adjusted *p* values (FDR) for each protein fragment pair. Only fragment pairs that show a positive $\log_2$ -fold change with an FDR < 0.05 in the presence of 3-AT when compared with 0 mM 3-AT were deemed as significantly interacting.

Author contributions—S. S. A. and J. A. M. conceptualization; S. S. A., S. S.-R., I. A., Y. A. M., and J. A. M. data curation; S. S. A., S. S.-R., Y. A. M., and J. A. M. supervision; S. S. A., S. S.-R., I. A., and J. A. M. validation; S. S. A., S. S.-R., N. M. A.-T., and J. A. M. investigation; S. S. A., S. S.-R., and J. A. M. methodology; S. S.-R., N. M. A.-T., and J. A. M. writing-original draft; S. S.-R., Y. A. M., and J. A. M. project administration; S. S.-R. and J. A. M. writing-review and editing; I. A. and J. A. M. formal analysis; Y. A. M. and J. A. M. resources; J. A. M. funding acquisition; J. A. M. visualization.

Acknowledgment—We thank the members of the Genomics Laboratory at Weill Cornell Medicine in Qatar for assistance.

References

1. Fields, S., and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340,** 245–246 CrossRef Medline
2. Dove, S. L., and Hochschild, A. (2004) A bacterial two-hybrid system based on transcription activation. *Methods Mol. Biol.* **261,** 231–246 CrossRef Medline
3. Suter, B., Zhang, X., Pesce, C. G., Mendelsohn, A. R., Dinesh-Kumar, S. P., and Mao, J. H. (2015) Next-generation sequencing for binary protein-protein interactions. *Front. Genet.* **6,** 346 CrossRef Medline
4. Lewis, J. D., Wan, J., Ford, R., Gong, Y., Fung, P., Nahal, H., Wang, P. W., Desveaux, D., and Guttman, D. S. (2012) Quantitative interactor screening with next-generation sequencing (QIS-Seq) identifies *Arabidopsis thaliana* MLO2 as a target of the *Pseudomonas syringae* type III effector HopZ2. *BMC Genomics* **13,** 8 CrossRef Medline
5. Yachie, N., Petsalaki, E., Mellor, J. C., Weile, J., Jacob, Y., Verby, M., Ozturk, S. B., Li, S., Cote, A. G., Mosca, R., Knapp, J. J., Ko, M., Yu, A., Gebbia, M., Sahni, N., *et al.* (2016) Pooled-matrix protein interaction screens using barcode fusion genetics. *Mol. Syst. Biol.* **12,** 863 CrossRef Medline
6. Martin, M. L., Lieberman, P. M., and Curran, T. (1996) Fos-Jun dimerization promotes interaction of the basic region with TFIIE-34 and TFIIF. *Mol. Cell. Biol.* **16,** 2110–2118 CrossRef Medline
7. Rauscher, F. J., 3rd, Sambucetti, L. C., Curran, T., Distel, R. J., and Spiegelman, B. M. (1988) Common DNA binding site for Fos protein complexes and transcription factor AP-1. *Cell* **52,** 471–480 CrossRef Medline
8. Wong, W., and Scott, J. D. (2004) AKAP signalling complexes: focal points in space and time. *Nat. Rev. Mol. Cell Biol.* **5,** 959–970 CrossRef Medline
9. Newlon, M. G., Roy, M., Morikis, D., Carr, D. W., Westphal, R., Scott, J. D., and Jennings, P. A. (2001) A novel mechanism of PKA anchoring revealed by solution structures of anchoring complexes. *EMBO J.* **20,** 1651–1662 CrossRef Medline
10. Meng, X., Franklin, D. A., Dong, J., and Zhang, Y. (2014) MDM2-p53 pathway in hepatocellular carcinoma. *Cancer Res.* **74,** 7161–7167 CrossRef Medline
11. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274,** 948–953 CrossRef Medline
12. Di Niro, R., Sulic, A. M., Mignone, F., D'Angelo, S., Bordoni, R., Iacono, M., Marzari, R., Gaiotto, T., Lavric, M., Bradbury, A. R., Biancone, L., Zevin-Sonkin, D., De Bellis, G., Santoro, C., and Sblattero, D. (2010) Rapid interactome profiling by massive sequencing. *Nucleic Acids Res.* **38,** e110 CrossRef Medline
13. Goldsmith, M., Kiss, C., Bradbury, A. R., and Tawfik, D. S. (2007) Avoiding and controlling double transformation artifacts. *Protein Eng. Des. Sel.* **20,** 315–318 CrossRef Medline
14. Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 CrossRef Medline
15. Glover, J. N., and Harrison, S. C. (1995) Crystal structure of the heterodimeric bZIP transcription factor c-Fos–c-Jun bound to DNA. *Nature* **373,** 257–261 CrossRef Medline
16. Böttger, A., Böttger, V., Garcia-Echeverria, C., Chène, P., Hochkeppel, H. K., Sampson, W., Ang, K., Howard, S. F., Picksley, S. M., and Lane, D. P. (1997) Molecular characterization of the hdm2-p53 interaction. *J. Mol. Biol.* **269,** 744–756 CrossRef Medline
17. Jeffrey, P. D., Gorina, S., and Pavletich, N. P. (1995) Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 Å. *Science* **267,** 1498–1502 CrossRef Medline
18. Scott, J. D., Stofko, R. E., McDonald, J. R., Comer, J. D., Vitalis, E. A., and Mangili, J. A. (1990) Type II regulatory subunit dimerization determines the subcellular localization of the cAMP-dependent protein kinase. *J. Biol. Chem.* **265,** 21561–21566 Medline