

## Video Article

# Genome-wide Protein-protein Interaction Screening by Protein-fragment Complementation Assay (PCA) in Living Cells

Samuel Rochette\*<sup>1</sup>, Guillaume Diss\*<sup>1</sup>, Marie Filteau<sup>1</sup>, Jean-Baptiste Leducq<sup>1</sup>, Alexandre K. Dubé<sup>1</sup>, Christian R. Landry<sup>1</sup><sup>1</sup>Département de Biologie, Institut de biologie intégrative et des systèmes & PROTEO, Université Laval

\*These authors contributed equally

Correspondence to: Christian R. Landry at [Christian.Landry@bio.ulaval.ca](mailto:Christian.Landry@bio.ulaval.ca)URL: <https://www.jove.com/video/52255>DOI: [doi:10.3791/52255](https://doi.org/10.3791/52255)

Keywords: Cellular Biology, Issue 97, Protein-protein interaction (PPI); high-throughput screening; yeast; protein-fragment complementation assay (PCA); dihydrofolate reductase (DHFR); high-density arrays; systems biology; biological networks

Date Published: 3/3/2015

Citation: Rochette, S., Diss, G., Filteau, M., Leducq, J.B., Dubé, A.K., Landry, C.R. Genome-wide Protein-protein Interaction Screening by Protein-fragment Complementation Assay (PCA) in Living Cells. *J. Vis. Exp.* (97), e52255, doi:10.3791/52255 (2015).

## Abstract

Proteins are the building blocks, effectors and signal mediators of cellular processes. A protein's function, regulation and localization often depend on its interactions with other proteins. Here, we describe a protocol for the yeast protein-fragment complementation assay (PCA), a powerful method to detect direct and proximal associations between proteins in living cells. The interaction between two proteins, each fused to a dihydrofolate reductase (DHFR) protein fragment, translates into growth of yeast strains in presence of the drug methotrexate (MTX). Differential fitness, resulting from different amounts of reconstituted DHFR enzyme, can be quantified on high-density colony arrays, allowing to differentiate interacting from non-interacting bait-prey pairs. The high-throughput protocol presented here is performed using a robotic platform that parallelizes mating of bait and prey strains carrying complementary DHFR-fragment fusion proteins and the survival assay on MTX. This protocol allows to systematically test for thousands of protein-protein interactions (PPIs) involving bait proteins of interest and offers several advantages over other PPI detection assays, including the study of proteins expressed from their endogenous promoters without the need for modifying protein localization and for the assembly of complex reporter constructs.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/52255/>

## Introduction

Protein interaction networks (PINs) offer a low resolution map of how proteins are functionally organized in the cell<sup>1</sup>. Each physical connection between two proteins, or protein-protein interaction (PPI), may represent an association that is stable in time, such as those found within protein complexes and that contribute to the structural organization of the cell. These connections may also represent transient associations that regulate the activity, stability, localization and interactions of the two partners. Identifying the physical interaction partners of a given protein therefore provides rich information on the function and regulation of that protein<sup>2,3</sup>. For these reasons, large efforts have been put towards the mapping of PINs in model organisms, including *Escherichia coli*<sup>4-6</sup>, *Arabidopsis thaliana*<sup>7</sup>, *Saccharomyces cerevisiae*<sup>8-12</sup>, *Drosophila melanogaster*<sup>13</sup>, *Caenorhabditis elegans*<sup>14</sup> and *Homo sapiens*<sup>15</sup>. These studies have provided important insights into how proteins are organized in the cell and thus key information on proteins with previously unknown functions.

Several strategies have been developed over the years to study PINs. These technologies can be broadly grouped in three categories based on the kind of information they provide on PPIs (reviewed in<sup>16-18</sup>). The first one is based on yeast two-hybrid and its derivatives<sup>19</sup>. These technologies provide information on the direct association between pairs of proteins, which allows constructing binary networks. The second family is based on the affinity purification of bait proteins and the identification of their associated partners, such as affinity purification followed by mass spectrometry<sup>20</sup>. These approaches identify groups of proteins that are directly or indirectly associated, generally in a stable manner, and are extremely powerful to identify protein complexes. The third approach is based on protein-fragment complementation assays (PCAs)<sup>11,21</sup>. This approach provides an intermediate level of resolution between the two former approaches, as it allows detecting direct and proximal associations between proteins. Each technique has its own strengths and weaknesses, as recently reviewed<sup>18</sup>.

The best-described eukaryotic PIN is by far the one of the budding yeast *Saccharomyces cerevisiae*, in part because its proteome is relatively less complex than those of other model eukaryotes and because high-throughput assays to detect PPIs have first been assayed and are more efficiently implemented in this model organism<sup>9-12</sup>. A particularly powerful method for the yeast system is the dihydrofolate reductase protein-fragment complementation assay (DHFR-PCA), an assay that has been used in different contexts to study the yeast PIN in standard and perturbed conditions<sup>11,22-26</sup>. This method relies on a survival assay that enables the detection of direct and near-direct PPIs for a given bait protein at both endogenous expression levels and native subcellular localizations of the interaction partners<sup>11,21</sup> in a quantitative manner<sup>27</sup>. The signal obtained using this assay (*i.e.* colony size on high-density colony arrays) thus reflects the amount of protein complexes formed between the bait and prey in a cellular environment almost equivalent to the one of wild-type cells. The assay is based on the reconstitution of a reporter enzyme involved in folate metabolism, the dihydrofolate reductase (DHFR), whereby two complementary fragments of the DHFR that are fused

to the two proteins of interest are brought into proximity when the two proteins interact, which in turn leads to the reversible reconstitution of the enzyme activity<sup>11</sup> and growth of the strain on a medium containing methotrexate (MTX; **Figure 1**). This compound inhibits the endogenous DHFR enzyme, but not the mutated one used in the assay<sup>28</sup>. Two collections of PCA strains, one containing ~4,300 *MATa* strains with an ORF fused to the DHFR F[1,2] fragment and one containing ~4,800 *MATa* strains with an ORF fused to the DHFR [3] fragment, can be purchased to implement DHFR-PCA at small or large scale in any laboratory. Here, we describe a general but detailed protocol to screen for PPIs between one bait protein and ~4,800 prey proteins using this assay.

## Protocol

### 1. Construction/verification of Bait Strains

- If the bait strain of interest is available in the *MATa* DHFR F[1,2] collection, retrieve it from the collection as described in step 1.1.1, otherwise construct the strain as described in step 1.1.2.  
NOTE: The protocol described here uses a DHFR F[1,2] strain as a bait and the DHFR F[3] collection as preys, as this collection contains more strains than the DHFR F[1,2] collection. However, it is possible to perform the screen the other way around if the bait strain is only available in the DHFR F[3] collection or in both orientations if one requires higher coverage of the interactome.
  - Thaw the glycerol stock plate that contains the bait strain on ice for one hour. Sterilize the aluminum foil covering the plate using 95% ethanol. Pierce the foil with a sterile tip, pipette up and down to resuspend the cells and streak 2-3  $\mu$ l of glycerol stock on selective yeast extract peptone dextrose (YPD) + 100  $\mu$ g/ml Nourseothricin (Nat) in order to isolate single colonies. Incubate for two days at 30 °C.
  - Construction of a bait PCA strain (*MATa* DHFR F[1,2]).
    - Using high-fidelity polymerase and a standard PCR protocol, amplify the DHFR F[1,2] cassette from plasmid pAG25-linker-DHFR F[1,2]-ADHterm using oligonucleotides with overhanging ends homologous to the last 40 bp of the ORF's 3'-end excluding the stop codon (Forward primer) and to the first 40 bp of the gene's 3'-UTR (Reverse primer) (**Figure 1A**).
    - Transform the PCR product into competent yeast cells (usually in a BY4741 strain) using standard LiOAc/PEG yeast transformation protocol as in<sup>29</sup> (**Figure 1A**).
    - Plate on selective YPD + Nat medium to isolate positive transformants.
    - Perform a diagnostic colony PCR on isolated colonies to confirm the proper DHFR F[1,2] fusion. Use primers annealing 1) in the gene coding sequence (Forward oligo) about 100 bp upstream of the DHFR fusion and 2) in the ADH terminator of the cassette (Reverse oligo) (**Figure 1B**).
    - Sequence the PCR product by Sanger sequencing to confirm proper gene fusion.
    - Archive the confirmed bait strain in 25% glycerol at -80 °C.  
NOTE: The protocol can be paused at this step.

### 2. Pin-tool Sterilization and Printing Procedures

NOTE: The sterilization procedure described below was optimized for the pin-tools manipulated by the BM3-BC (S&P Robotics) robotic platform, but can be adapted to other platforms as well. This section describes the Pin-tool sterilization and printing procedures that are used to transfer cells from one medium to another for the rest of the protocol. In-house scripts used to perform these routines can be obtained upon request. Note that all steps can be performed without the need of a robotic platform using a manual pin-tool<sup>30</sup>.

- Mount the appropriate pin-tool on the robotic platform.
- Prepare cleaning and wet stations as follows:
  - Add 500 ml of sterile water in the water bath station.
  - Add 320 ml of sterile water in the brush station.
  - Add 380 ml of 70% ethanol in the sonicator when replicating from agar plate (86 x 128 mm omnitray, containing 35 ml of solidified medium) to agar plate, or 400 ml when replicating from a microtiter plate containing liquid cultures to an agar plate.
  - Add 35 ml of sterile water in the wet station (consisting of a sterile empty omnitray).
- At the beginning of each day that requires the robotic platform, sterilize the pin-tools five times for one min in the sonicator bath. Meanwhile, turn on the UV lamp for five min to sterilize the robot enclosure.  
NOTE: This is not required if the robot is housed under a sterile hood.
- Between each pin-tool replication round, sterilize the pin-tool as follows:
  - Soak the pin-tool five times for 10 sec in the water bath station to remove cell clumps.
  - Soak the pin-tool twice back and forth in the brush station.  
NOTE: The rotating brush will remove residual cells.
  - Soak the pin-tool twice for 20 sec in the sonicator station.  
NOTE: Remaining cells on the pins will be removed by sonication or killed by ethanol.
  - Make sure that the dipping depth of the pins increases at each successive bath to ensure proper sterilization.
- Dry the pin-tool in the air-dryer station for 25 sec.
- Before taking cells on the source plates, wet the pins in the wet station, which contains 35 ml of sterile water in an omnitray.
- Dip the pins twice in the colonies from the source plate.
- Print on a freshly poured destination plate by touching the agar surface twice (hereafter referred to as the action of "printing" an array).

### 3. Condensation of the DHFR F[3] Collection in 1,536 Arrays Using an Automated Robotic Platform

1. Thaw on ice the DHFR F[3] collection (60 96-well plates) and an additional 96-well plate filled with the L-DHFR F[3] control strain (**Figure 1C**), which contains the DHFR F[3] fragment and the upstream linker expressed alone as a negative control.  
NOTE: In principle, this fragment should not interact with any DHFR-fragment fusion protein (see discussion for details).
2. Centrifuge plates (quick spin) before removing the aluminum foil to avoid risks of cross-contamination among wells.
3. Condense the collection onto 16 arrays of 384 strains (**Figure 2A**). To do this, for each 384 array, print four glycerol plates on the four quadrants of a selective YPD + 250 µg/ml Hygromycin B (HygB) omnitrays using the 96 pin-tool (here, a 384 array can be subdivided in four equally interspaced quadrants, each consisting of 96 positions in a 2 x 2 matrix layout). Insert four 96-well plates containing the L-DHFR F[3] negative control in between the 60 other plates in order to have a final set of 64 plates that exactly fill four 1,536 arrays. Sterilize the pin-tool between each replication cycle as described in step 2.  
NOTE: Insert the L-DHFR F[3] plates in order to have one such plate on each of the four final 1,536 arrays.
4. Incubate plates for two days at 30 °C. NOTE: At this stage, the DHFR collection can be stored in a 384 format at 4 °C for up to one month on agar plates.
5. Condense the collection into four arrays of 1,536 strains (**Figure 2A**). To do this, for each 1,536 array, print four arrays of 384 strains on the four quadrants of a plate of the same medium as in step 3.2 using the 384 pin-tool (here, a 1,536 array can be subdivided in four equally interspaced quadrants each consisting of 384 positions in a 2 x 2 matrix layout).
6. Incubate plates for two days at 30 °C. NOTE: At this stage, the DHFR collection can be stored in a 1,536 format at 4 °C for up to one month on agar plates.
7. Replicate the four arrays on the same medium to standardize colony size using a 1,536 pin-tool.
8. Incubate plates for two days at 30 °C.

### 4. High-throughput DHFR-PCA Procedure

1. Inoculate a culture of the bait strain (DHFR F[1,2] fusion) obtained from step 1.1.1 or 1.1.2 in 20 ml of liquid YPD + Nat in a 50 ml tube (**Figure 2B**).
2. Incubate for two days at 30 °C with shaking at 250 rpm to allow the culture to reach saturation.
3. After two days of incubation, plate 5 ml of the culture on a YPD + Nat omnitrays. Let cells adsorb on the surface for 5-10 min and remove the excess liquid (**Figure 2B**). Repeat twice to make three replicates.
4. Incubate for two days at 30 °C.
5. Print the bait strain from steps 4.3 and 4.4 on 12 YPD plates (enough to mating four plates of the DHFR F[3] collection × three replicates) with a 1,536 pin-tool using each cell lawn no more than four times.
6. Print the appropriate array of the DHFR F[3] collection on top of the bait cells using the 1,536 pin-tool (**Figure 2C**).
7. Let the strains mate by incubating for two days at 30 °C.
8. Select diploid cells by printing colonies on omnitrays containing YPD + HygB + Nat (**Figure 2C**).
9. Incubate for two days at 30 °C.
10. Repeat diploid selection as described in steps 4.8 and 4.9 (**Figure 2B**).
11. Prepare plates with media containing MTX (MTX medium) a day before use following these steps<sup>21</sup> (CAUTION: be careful when working with MTX as it is a toxic compound. Always wear gloves, protection glasses and a lab coat when handling it) (**Figure 2C**):
  1. Prepare the 10x -lys/met/ade amino acid drop-out in deionized water. Filter-sterilize the drop-out in a sterile bottle using a 0.2 µm sterile syringe filter or, if needed in large quantities, a 0.2 µm bottle top filtration funnel (**Table 1**).
  2. Prepare a 10 mg/ml MTX stock solution in dimethyl sulfoxide (DMSO). Use immediately after preparing the solution and freeze the remainder at -20 °C. Protect it from light as it is photosensitive. Do not freeze MTX again after thawing it.
  3. Prepare the media as follows (medium ingredients and quantities are the same as those used in Tarassov *et al.*<sup>11</sup>):
    1. For one liter of medium, mix in two separate flasks: 1) 6.69 g of yeast nitrogen base without amino acids and without ammonium sulfate and 330 ml of deionized water; 2) 25 g of noble agar and 500 ml of deionized water.
    2. Autoclave flasks at 121 °C for 20 min.
    3. Equilibrate temperature in a water bath at 55 °C for at least one hour.
    4. Mix the two flasks together and add 50 ml of sterile 40% glucose, 100 ml of 10x sterile drop-out, and 20 ml of MTX 10 mg/ml.
    5. Pour 35 ml (see Note below) of medium in omnitrays. Let solidify for at least an hour and a half. Shield plates from light.  
NOTE: Here, pouring 35 ml of medium per omnitrays is critical to ensure equal plate thickness, which is important for all downstream steps.
12. Image plates from the second round of diploid selection with the robotic platform or with a regular digital camera with uniform plate illumination. Use these pictures to identify empty positions on the arrays when performing the downstream analysis. Make sure that the parameters of the camera are always the same and that the robot light is turned on.
13. Print diploid cells on MTX media using the 1,536 pin-tool.
14. Incubate for four days at 30 °C in plastic bags to prevent drying.
15. Prepare a second batch of omnitrays containing MTX medium as described in step 4.11.
16. After four days of incubation, image plates using the robotic platform or regular digital camera. Make sure that the parameters of the camera are always the same and that the robot light is turned on.
17. Perform a second round of MTX selection by replicating the cells on the second batch of MTX media.  
NOTE: This will decrease the background growth of PCA strains and increase the quantitative resolution.
18. Incubate for four days at 30 °C in plastic bags to prevent drying.
19. Image plates as described in step 4.16.

## 5. Image Analysis

- Analyze images of colony arrays with custom ImageJ<sup>31</sup> scripts or using published softwares such as Colonizer, Ht colony grid analyzer, Cell profiler, Colony Imager, ScreenMill, YeastXtract and gitter (compiled in<sup>32</sup>). Image analysis should output one or several spreadsheets containing colony sizes for each position of each array, use these colony sizes for all downstream analyses.  
NOTE: In this study, we used a custom ImageJ script described in Leducq *et al.*<sup>33</sup> (see discussion section for more details).

## 6. Data Analysis

NOTE: Results from image analysis can be processed in a tabulator such as excel or using a scripting language such as R<sup>34</sup>. The following steps describe the procedure using a custom ImageJ<sup>31</sup> script.

- Using a custom script, concatenate output files from image analysis and annotate each row with the plate and strain information as in Supplementary **Table 1**.
- Log<sub>2</sub> transform the colony size values (Integrated density or colony area; here, column "IntDenBackSub" from Supplementary **Table 1** was used).  
NOTE: Value distribution will look as in **Figure 3A**.
- Normalize these values by subtracting the median value of each plate.  
NOTE: This step controls for plate bias that may result from unequal media quantity or variation in automatic image acquisition, and reduces the inter-replicate variance (**Figure 3B**).
- Verify that replicates correlate with each other (**Figure 3C**) to assess the reproducibility of the experiments.
- To differentiate interacting from non-interacting bait-prey pairs, set a high-confidence threshold corresponding to the 95<sup>th</sup> percentile of the distribution of the L-DHFR F[3] controls.  
NOTE: In this experiment, this corresponds to 3.39 (**Figure 3D**). Alternatively, a threshold based on the overlap with known physical interactors such as those reported in the BioGRID<sup>35</sup> can be used. See discussion for more details.
- For any bait, filter preys identified as involved in false positive interactions in DHFR-PCA screens (see discussion for more details) and listed in Supplementary **Table 2** (identified as "1" in column "filtered").
- Average the Log<sub>2</sub> normalized colony sizes of the three replicates of each interaction (column "Average score" in Supplementary **Table 2**).

## 7. Validation of Physical Interactors Using Small-scale Experiments

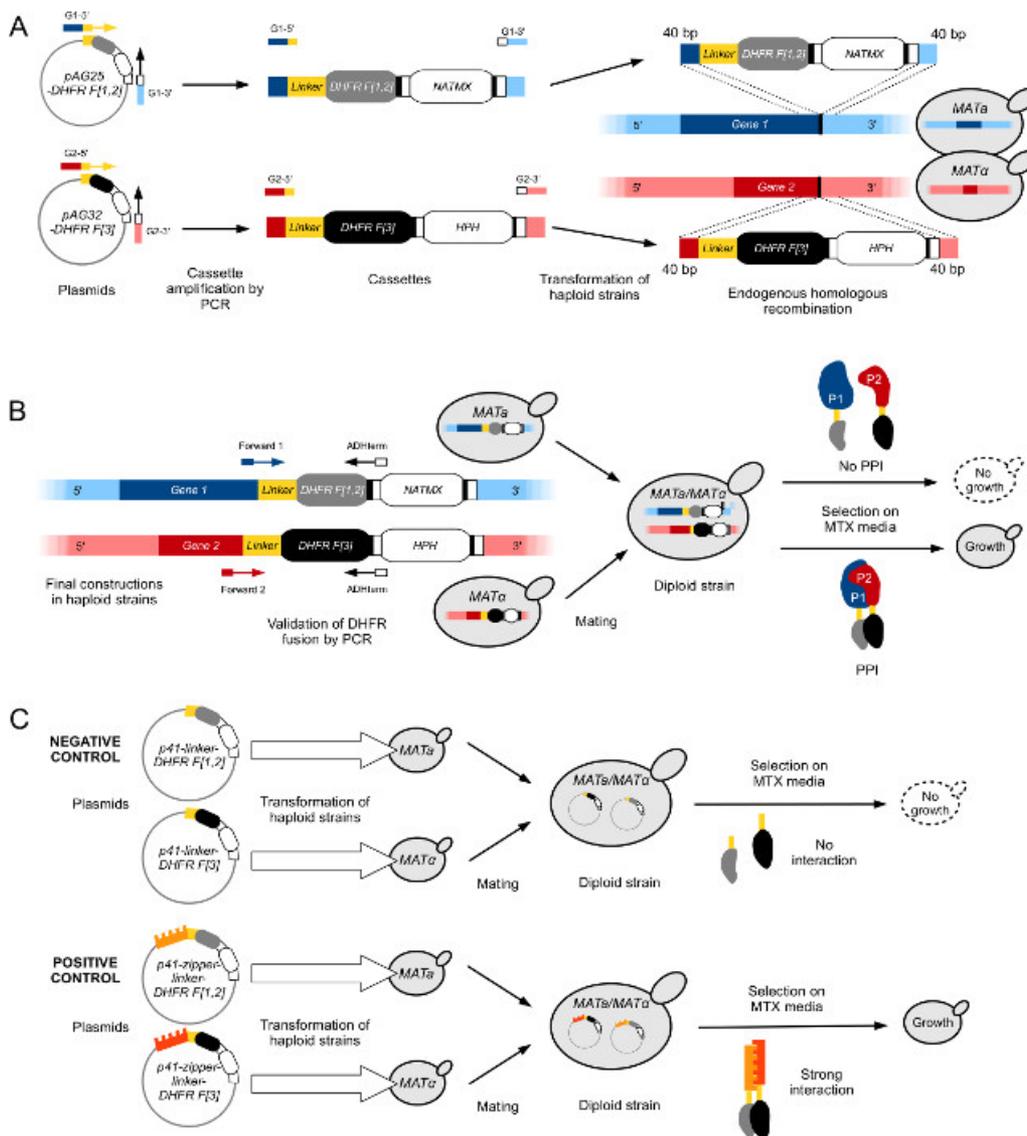
NOTE: Any PPI of particular interest having a score above or close to the applied threshold can be validated using the DHFR-PCA assay in a small-scale experimental design using a growth assay on solid or liquid MTX medium. The steps below show the procedure to manually construct diploid PCA strains and perform spot assays on MTX medium. The experimenter should perform these steps for all necessary controls (Bait-DHFR F[1,2] x L-DHFR F[3], Zipper-linker-DHFR diploid strains and linker-DHFR diploid strain).

- Plate 2-3  $\mu$ l of the glycerol stock of the bait strain generated in 1.1.2.6), the L-DHFR F[3] control strain, the Zipper-linker-DHFR and linker-DHFR diploid strains on YPD+Nat, YPD+HygB and twice YPD+Nat+HygB media, respectively.
- Retrieve the prey of interest in the DHFR F[3] collection and follow instructions in step 1.1.1, but streak the strain on YPD+HygB medium instead of YPD+Nat medium.
- Perform a diagnostic PCR as in 1.1.2.4 to confirm the DHFR F[3] fusion at the prey locus and sequence the product.
- Inoculate 1 ml of liquid YPD medium with the haploid strains to mate (Bait x Prey, Bait x L-DHFR F[3] control) and grow at least two days at 30 °C to allow the diploids to form.
- Select diploids by streaking 4-5  $\mu$ l of the culture in 7.4 on solid YPD+HygB+Nat medium. Grow two days at 30 °C.
- Select one isolated colony and grow overnight in liquid culture (1 ml) to perform the growth assay.
- Prepare MTX and DMSO (same ingredients as MTX medium, but without MTX) plates a day before use. See step 4.11 for more details.
- Perform growth assay by spotting serial dilutions of the different cultures on the control (DMSO) and selection plates (MTX).
  - Dilute precultures to O.D.= 1.
  - Perform five-fold dilutions (up to a dilution factor of 625) in a sterile 96-well plate.
  - Spot 4  $\mu$ l of each dilution on the PCA media (DMSO and MTX).
  - Incubate at 30 °C in plastic bags to prevent drying.
  - Image plates from days 1 to 7 of incubation using the robotic platform or a regular digital camera.

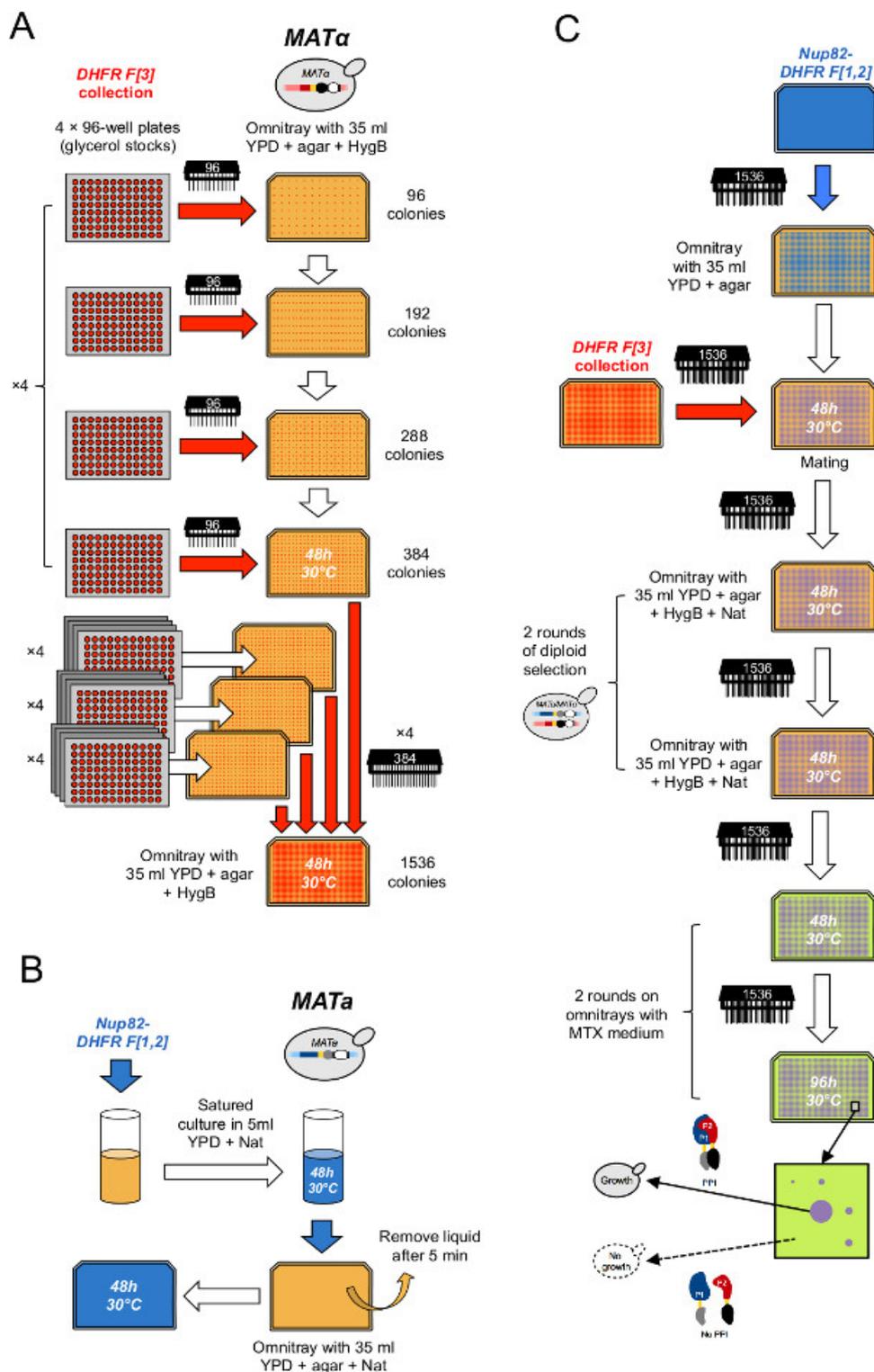
## Representative Results

Supplementary **Table 2** is an example of representative results obtained using the yeast protein Nup82 fused to the DHFR F[1,2] fragment as a bait. The threshold defined with the L-DHFR F[3] controls can be used as an empirical threshold to determine high confidence hits (**Figures 3D & 3E**). Alternatively, the score ranking can be used to perform Gene Ontology enrichments or other functional analyses<sup>36</sup> based on gold standards<sup>37</sup>. The known physical interactors of the bait can be retrieved from databases like BioGRID<sup>35</sup> and overlaid on the data (**Figures 3E & 3F**). In this example, five out of eight high confidence hits have been previously reported as Nup82 interactors and two are part of the Nup82 subcomplex, Nup116 and Nup159 (**Figure 3F & 3G**). The other member of the complex, Nsp1, does not show any interaction in our experiment. Two preys, Ade17 and Tef2 (not shown in **Figure 3F**), had scores above the hard threshold applied, but these are likely to be false positives as they interact with almost any bait protein in PCA screens we have performed (unpublished results). On the other hand, Pex30 may represent a novel physical interactor of Nup82 and we were able to confirm this interaction using DHFR-PCA at low-throughput (**Figure 3G**). Pex30 is a peroxisomal membrane protein and a few direct interactions have been reported between the nuclear pore complex (NPC) and this organelle. A two-hybrid screen identified two other NPC protein, Nup53 and Asm4 (Nup59), as physical interactors of Pex30<sup>38</sup>, and a genetic interaction

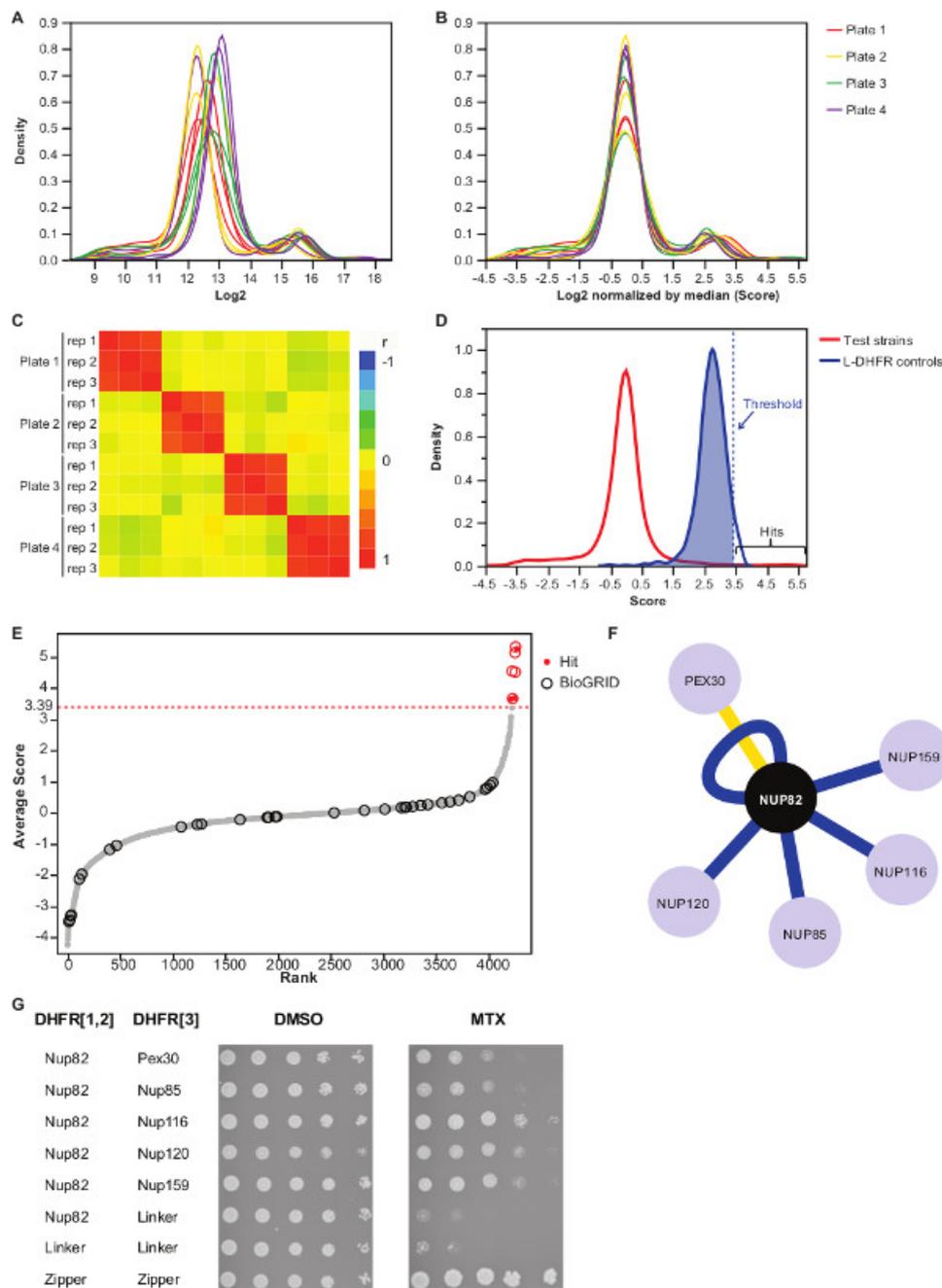
between Pex30 and Nup170 has been reported<sup>39</sup>. Two other interaction partners detected, Nup120 and Nup85 (Figure 3F & 3G), are not part of the Nup82 sub-complex, illustrating the ability of the DHFR-PCA to detect interactions within and between subcomplexes in larger complexes<sup>11</sup>.



**Figure 1: Engineering yeast strains for high-throughput DHFR-PCA (figure adapted from Leducq et al. 2012<sup>33</sup>).** (A) Construction of haploid *MATa* and *MATα* strains to fuse *Gene1* (G1) and *Gene2* (G2) with the *DHFR F[1,2]-NatMX* and the *DHFR F[3]-HPH* cassettes, respectively. Cassettes are amplified from plasmids *pAG25-DHFR F[1,2]* and *pAG32-DHFR F[3]* with forward primers G1-5' and G2-5', and reverse primers G1-3' and G2-3', and then inserted in the genome at the 3' end of the target gene by homologous recombination. The resulting proteins, P1 and P2, are respectively fused to the DHFR F[1,2] fragment (*MATa*) and the DHFR F[3] fragment (*MATα*) via a flexible linker. (B) Verification of the construction in (A) is performed by sequencing the junctions between *Gene1*'s and *Gene2*'s ORFs and the DHFR cassettes. The constructed PCA strains of opposing mating types are then mated to form a diploid. Diploid strains grow on MTX medium if the two complementary DHFR fragments are brought into proximity by an interaction between P1 and P2, which reconstitutes the activity of the DHFR enzyme. (C) Construction of control diploid PCA strains for DHFR-PCA screens. The negative controls (L-DHFR) are constructed by transforming separately haploid *MATa* and *MATα* strains with plasmids *p41-linker-DHFR F[1,2]* and *p41-linker-DHFR F[3]*<sup>11</sup>, respectively. The two strains are mated resulting in a negative control diploid strain in which the DHFR fragments are unable to complement each other (top). Positive controls are constructed using the same approach as for the negative controls, but the plasmids transformed in haploid strains (*p41-zipper-linker-DHFR F[1,2]* (*p41-ZL-DHFR F[1,2]*) and *p41-zipper-linker-DHFR F[3]*) (*p41-ZL-DHFR F[3]*) contain two GCN4 parallel leucine zipper fragments fused to the complementary DHFR fragments, which leads to a strong and constitutive interaction that reconstitutes DHFR activity (bottom). Please click here to view a larger version of this figure.



**Figure 2: High-throughput DHFR-PCA procedure (figure adapted from Leducq *et al.*<sup>33</sup>).** (A) The *MAT $\alpha$*  DHFR F[3] collection is condensed in a 1536 format through two successive rounds of condensation. First, glycerol stock plates are combined by groups of four on selective YPD + HygB medium in a 384 format using the 96 pin-tool. Second, 384 arrays are combined by groups of four on selective YPD + HygB medium in a 1536 format using the 384 pin-tool. (B) Cell lawns of the *MAT $\alpha$*  PCA bait strain are prepared by growing a saturated culture of the bait strain in selective YPD + Nat medium and plating the culture on a YPD + Nat omnitray. (C) These lawns are used to mate the bait strain with the DHFR F[3] collection on YPD medium. Cells are successively transferred twice on YPD + HygB + Nat medium to select for diploids and twice on MTX medium to perform PCA. Growth will only be observed on MTX medium if the DHFR fragments complement each other following an interaction between the bait and the prey proteins. [Please click here to view a larger version of this figure.](#)



**Figure 3: Data analyses through steps of normalization, determination of a significance threshold and identification of positive interactions.** (A) Density distribution of colony size on each plate ( $\log_2$ ). (B) Normalization by the median of each image corrects for biases associated with plate-to-plate effects. (C) Heatmap showing the Spearman correlation coefficient among the plates, confirming the reproducibility of the procedure. (D) Distribution of scores for the tested PPIs and the L-DHFR F[3] controls. A hard threshold can be set to the 95<sup>th</sup> percentile of the L-DHFR F[3] distribution to identify high-confidence PPIs (represented by a dotted vertical line). (E) Rank order distribution of the average score of each prey. Previously reported physical interaction partners of Nup82p in the BioGRID<sup>35</sup> are identified by circles and those reported in this study are identified by red dots. The threshold defined in (D) is shown as a dotted line. (F) Network showing the high-confidence PPIs identified in this study. Blue edges show previously reported physical interactors (BioGRID<sup>35</sup>) and yellow edges show a previously unreported interaction with Pex30. (G) Spot-dilution assay of PCA diploid strains involving Nup82-DHFR F[1,2] and prey-DHFR F[3] pairs identified as physical interactors of Nup82p in the present study. Growth assay was performed in DMSO medium (MTX solvent, left panel) and MTX medium (right panel). Negative controls consisting of Nup82-DHFR F[1,2] - Linker-DHFR F[3] and Linker-DHFR F[1,2] - Linker-DHFR F[3] and a positive control consisting of the strong interaction between two leucine zipper moieties (zipper-DHFR F[1,2] - zipper-DHFR F[3]) were included. Cell growth superior to the negative controls in MTX medium should be interpreted as a physical interaction. [Please click here to view a larger version of this figure.](#)

Amino acid	Quantity (g)
Adenine sulfate*	0.2

L-Tryptophane	0.4
L-Tyrosine	0.3
L-Phenylalanine	0.5
L-Glutamic acid (monosodium salt)	1.0
L-Asparagine	1.0
L-Valine	1.5
L-Threonine	2.0
L-Serine	3.75
Uracile	0.2
L-Histidine HCl	0.2
L-Arginine HCl	0.2
L-Methionine*	0.2
L-Lysine*	0.2
L-Leucine	0.6

\* Withdrawn when performing standard PCA (as in this protocol) but can be added for other purposes.

**Table 1: Composition of the 10X -lys/met/ade drop-out in the MTX medium.** Quantities are for one liter of 10x drop-out. Asterisks denote the compounds that should be withdrawn for the standard MTX medium, but that can be added for other purposes.

**Supplementary Table 1: Combined data from the 12 test plates.** Table 1 contains the concatenated data from the ImageJ analysis performed using a custom ImageJ script with each row corresponding to a single position on each 1,536 array. In addition, each row has been annotated with the information about the image file, DHFR collection plate, replicate, ORF and protein name.

**Supplementary Table 2: Average normalized growth per strain.** Table 2 contains the average  $\text{Log}_2$  normalized growth for each strain of the collection along with the standard deviation. Filtered preys, hits and known physical interactors are identified in separate columns.

## Discussion

We describe a protocol based on the DHFR-PCA assay enabling the systematic identification of physical interactors for any given bait protein at high-throughput. This protocol can be adapted by screening for more baits, and this at any desired level of replication. We demonstrate the reliability of this protocol by the identification of physical interaction partners for a bait protein involved in the Nuclear Pore Complex: Nup82. Our analysis enabled to find five previously reported interactors and one previously unreported interactor (**Figures 3F & 3G**), highlighting the ability of the method to study the yeast protein interactome.

The protocol described here includes several critical steps to which the experimenter should pay attention. We recommend to 1) Make sure that the bait DHFR F[1,2] fusion is correct (**Figure 1B**); this can be achieved by sequencing the construction and measuring proper protein expression using an anti-DHFR F[1,2] or anti-DHFR F[3] antibody; 2) Prior to beginning the screen, it is recommended to verify if any bait of interest exhibits promiscuous interactions in PCA screens. This can be done by performing control screens with baits crossed with the appropriate L-DHFR control or by manually mating the bait with the appropriate L-DHFR control and performing a growth assay in MTX medium. 3) Plates should be poured the day before they are used so that moisture is optimal for cell adherence on the agar surface during the printing process; 4) Source plates should not be used more than four times to transfer enough cells on the destination plate. Increasing the number of copies of the destination plate can be done by successive steps of expansion (e.g. 4 copies -> 16 copies -> 64 copies). Alternatively, cells can be picked at different positions on the lawns or in the colony between different rounds of replication; 5) If several positions are missing after the diploid selection(s), make sure that the source plates were not used too many times in the mating step (steps 4.5 to 4.7); 6) Ensure that the MTX medium contains all essential ingredients at the right concentrations. Indeed, if no growth at all is observed on the MTX medium, it can be either because no interaction is detectable by PCA for the proteins of interest or because the MTX medium was not prepared properly. To ensure that the medium allows growth of strains showing DHFR fragments complementation, a constitutive interaction can be added at empty positions of the collection and used as a positive control such as DHFR-fragments fused to leucine zipper moieties<sup>33</sup> (**Figure 1C**). Parallel tests using the linker-DHFR fragments or the zipper-linker-DHFR fragments will allow to discriminate between conditions that allow all cells to grow (low MTX concentration or bait protein that tend to make false positive interactions, as described below) and conditions that prevent growth of all strains (MTX concentration too high or essential ingredient missing in the medium); 7) Given that PCA is performed through successive rounds of replications from one medium to another, cross-contamination among strains between different plates may occur if, for instance, the pin-tool is not sterilized properly between replication rounds and/or the last water bath (i.e. wet station) in the sterilization procedure is contaminated by colonies of previous replication rounds. Several positions on the arrays are empty and can thus be used as control positions where no growth should be observed to detect cross-contaminations.

Image analysis can be performed using several published softwares (see section 5 of the protocol) or any custom script. In this study, the custom script executes the following steps: 1) The script subtracts pixel values of an empty plate to pixel values of each plate in order to correct for illumination biases. 2) The script converts each background-corrected picture to binary using a pixel value threshold of 10. 3) For each 1,536 positions of each plate, determined by overlaying a rectangle on the edge colonies, the script runs ImageJ "Analyze particles ..." function in a circular selection. The circular selection is set with a radius equal to the interval between two positions minus 10 pixels. 4) The script selects the closest particle from the center of the selection and confirms it as a colony if its location is not more than half the interval between two colonies

away from the center of the selection. 5) The script measures pixel values of the selected particle on the background-corrected picture. 6) To further correct any remaining background illumination biases, the script subtracts the mean value of all the pixels from the circular selection that were not part of a particle to pixel values of the colony. The sum of these corrected pixel values, stored in the column "IntDenBackSub" of Supplementary Table 1, are used as a measure of colony size.

A critical step within the analysis part is the choice of the significance threshold. Here, we chose a threshold based on the distribution of the negative L-DHFR F[3] controls, but depending on the objective of the screen, such threshold may be too stringent. Indeed, L-DHFR F[3] controls are overexpressed (strong TEF promoter) such that the complementary fragments may spontaneously complement each other and these are thus not representative of the expression of most proteins. This is highlighted by the fact that the distribution of the L-DHFR F[3] controls is higher than the average of the background growth (Figure 3D). Thus, some interactions having scores below this stringent threshold but that are clearly outside of the background growth distribution can be considered as putative hits that may represent, for instance, transient or weak interactions. These can be further studied and cross-validated if, for instance, the two proteins are not expressed at levels that can allow the spontaneous complementation of the DHFR fragments like the L-DHFR controls. As an alternative, one could set a significance threshold based on the proportion of overlap with reported physical interactors in databases like BioGRID<sup>35</sup> in order to maximize the proportion of true positives over false positives. However, unlike the use of the L-DHFR distribution, this alternative may not always be feasible if, for instance, the number of known physical interactors is not sufficiently high. Moreover, the choice of the significance threshold has an impact on the proportion of false positives and false negatives in the final data set. Indeed, like any other PPI detection assay, false positives can result from unspecific interaction of a protein with the DHFR-fusion protein if, for example, the protein is highly abundant as mentioned earlier. This is exemplified by the fact that some preys systematically interact with all bait proteins in PCA screens and, thus, need to be removed from the analysis<sup>11</sup> (e.g. Tef2 and Ade17 and **Supplementary Table 2**). To circumvent this problem, a control PCA screen of the two collections against the appropriate L-DHFR control (F[1,2] or F[3]) to identify baits and preys exhibiting spontaneous DHFR fragments complementation can be performed in the specific conditions of each screen. Moreover, performing a Gene Ontology enrichment analysis can increase the confidence in the data if the function of a given bait is known. On the other hand, DHFR-PCA can give rise to false negatives for several reasons: 1) not all proteins can be fused to the DHFR fragments as these may destabilize the proteins or modify their localization if, for instance, the DHFR fusion to the C-terminus interferes with a localization signal; 2) DHFR reconstitution in some cellular compartments may not produce folate if, for instance, an essential precursor for folate synthesis is not available; 3) C-termini need to be within a distance of 8 nm for DHFR complementation to occur<sup>11</sup>. Thus, a well-known interaction may not be detected if their C-termini are not close enough in space. This is exemplified here by the fact that a large fraction of Nup82 physical interactions reported in databases, most of which are indirect, were not detected in our assay. Similarly, interactions between membrane proteins for which the C-termini are in *trans* relative to the membrane will not lead to DHFR fragments complementation and will not be detected<sup>11</sup>. Limitations 1) and 3) can be circumvented relatively simply by fusing the DHFR fragment to the N-termini of the protein. Doing so may prevent to interfere with a localization signal near the C-termini and may allow to detect an interaction between membrane proteins whose N and C-terminus are in *cis* relative to the membrane.

Several challenges remain in the study of PINs (reviewed in<sup>2,3</sup>). The maps of PINs produced so far have largely been described in a single experimental condition for each species and thus offer a single snapshot of how protein networks might be organized. There is therefore a need for the exploration of other experimental conditions to see how PINs may be reorganized in response to environmental changes, specific stimuli, across development or following mutations. These challenges will be overcome by the development of new technologies for interrogating PPIs in real-time, in living cells and by adapting current techniques so that they can be used by a larger community of laboratories. As a quantitative technique that can detect changes in the amount of DHFR complementation complexes<sup>27</sup>, DHFR-PCA can be adapted to overcome these challenges and has been used to study how PPIs are affected by a DNA damaging agent<sup>22</sup>, chemical agents<sup>25</sup>, gene deletions<sup>23,26</sup> or in other yeast species and their hybrids<sup>33</sup>. Exploring these new dimensions will become more and more important to reveal the dynamic of the PIN.

## Disclosures

Part of the open access publication fees for this article were paid by S&P Robotics.

## Acknowledgements

This work was supported by Canadian Institute of Health Research (CIHR) Grants 191597, 299432 and 324265, a Natural Sciences and Engineering Research Council of Canada Discovery grant and a Human Frontier Science Program grant to CRL. CRL is a CIHR New Investigator. Guillaume Diss is supported by a PROTEO fellowship. Samuel Rochette is supported by NSERC and FRQNT fellowships.

## References

1. Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*. **92**, 291-294 (1998).
2. Diss, G., *et al.* Integrative avenues for exploring the dynamics and evolution of protein interaction networks. *Curr Opin Biotechnol*. **24**, 775-783 (2013).
3. Vidal, M., Cusick, M. E., Barabasi, A. L. Interactome networks and human disease. *Cell*. **144**, 986-998 (2011).
4. Hu, P., *et al.* Global functional atlas of Escherichia coli encompassing previously uncharacterized proteins. *PLoS biology*. **7**, e96 (2009).
5. Arifuzzaman, M., *et al.* Large-scale identification of protein-protein interaction of Escherichia coli K-12. *Genome Res*. **16**, 686-691 (2006).
6. Rajagopala, S. V., *et al.* The binary protein-protein interaction landscape of Escherichia coli. *Nature biotechnology*. **32**, 285-290 (2014).
7. Arabidopsis-Interactome-Mapping-Consortium. Evidence for network evolution in an Arabidopsis interactome map. *Science*. **333**, 601-607 (2011).
8. Babu, M., *et al.* Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. *Nature*. **489**, 585-589 (2012).
9. Krogan, N. J., *et al.* Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*. **440**, 637-643 (2006).
10. Gavin, A. C., *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. **415**, 141-147 (2002).
11. Tarassov, K., *et al.* An in vivo map of the yeast protein interactome. *Science*. **320**, 1465-1470 (2008).

12. Uetz, P., *et al.* A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*. **403**, 623-627 (2000).
13. Guruharsha, K. G., *et al.* A protein complex network of *Drosophila melanogaster*. *Cell*. **147**, 690-703 (2011).
14. Li, S., *et al.* A map of the interactome network of the metazoan *C. elegans*. *Science*. **303**, 540-543 (2004).
15. Rual, J. F., *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature*. **437**, 1173-1178 (2005).
16. Landry, C. R., Levy, E. D., Abd Rabbo, D., Tarasov, K., Michnick, S. W. Extracting insight from noisy cellular networks. *Cell*. **155**, 983-989 (2013).
17. Berggard, T., Linse, S., James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics*. **7**, 2833-2842 (2007).
18. Wodak, S. J., Vlasblom, J., Turinsky, A. L., Pu, S. Protein-protein interaction networks: the puzzling riches. *Current opinion in structural biology*. **23**, 941-953 (2013).
19. Fields, S., Song, O. A novel genetic system to detect protein-protein interactions. *Nature*. **340**, 245-246 (1989).
20. Dunham, W. H., Mullin, M., Gingras, A. C. Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics*. **12**, 1576-1590 (2012).
21. Michnick, S. W., Ear, P. H., Landry, C., Malleshaiah, M. K., Messier, V. A toolkit of protein-fragment complementation assays for studying and dissecting large-scale and dynamic protein-protein interactions in living cells. *Methods Enzymol*. **470**, 335-368 (2010).
22. Rochette, S., Gagnon-Arsenault, I., Diss, G., Landry, C. R. Modulation of the yeast protein interactome in response to DNA damage. *Journal of proteomics*. **100**, 25-36 (2014).
23. Diss, G., Dube, A. K., Boutin, J., Gagnon-Arsenault, I., Landry, C. R. A systematic approach for the genetic dissection of protein complexes in living cells. *Cell Rep*. **3**, 2155-2167 (2013).
24. Gagnon-Arsenault, I., *et al.* Transcriptional divergence plays a role in the rewiring of protein interaction networks after gene duplication. *Journal of proteomics*. **81**, 112-125 (2013).
25. Schlecht, U., Miranda, M., Suresh, S., Davis, R. W., St Onge, R. P. Multiplex assay for condition-dependent changes in protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*. **109**, 9213-9218 (2012).
26. Lev, I., *et al.* Reverse PCA, a systematic approach for identifying genes important for the physical interaction between protein pairs. *PLoS Genet*. **9**, e1003838 (2013).
27. Freschi, L., Torres-Quiroz, F., Dube, A. K., Landry, C. R. qPCA: a scalable assay to measure the perturbation of protein-protein interactions in living cells. *Mol Biosyst*. **9**, 36-43 (2013).
28. Pelletier, J. N., Campbell-Valois, F. X., Michnick, S. W. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proceedings of the National Academy of Sciences of the United States of America*. **95**, 12141-12146 (1998).
29. Gietz, R. D., Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol*. **350**, 87-96 (2002).
30. Schuldiner, M., Collins, S. R., Weissman, J. S., Krogan, N. J. Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. *Methods*. **40**, 344-352 (2006).
31. Schneider, C. A., Rasband, W. S., Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature methods*. **9**, 671-675 (2012).
32. Wagih, O., Parts, L. gitter: A Robust and Accurate Method for Quantification of Colony Sizes From Plate Images. *G3 (Bethesda)*. **4**, (3), 547-552 (2014).
33. Leducq, J. B., *et al.* Evidence for the robustness of protein complexes to inter-species hybridization. *PLoS Genet*. **8**, e1003161 (2012).
34. *Development Core Team: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria (2008).
35. Stark, C., *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res*. **34**, D535-D539 (2006).
36. Vinayagam, A., *et al.* Protein complex-based analysis framework for high-throughput data sets. *Science signaling*. **6**, rs5 (2013).
37. Jansen, R., Gerstein, M. Analyzing protein function on a genomic scale: the importance of gold-standard positives and negatives for network prediction. *Current opinion in microbiology*. **7**, 535-545 (2004).
38. Ito, T., *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences of the United States of America*. **98**, 4569-4574 (2001).
39. Costanzo, M., *et al.* The genetic landscape of a cell. *Science*. **327**, 425-431 (2010).