Full title(s):

Y2H-SCORES: A statistical framework to infer protein-protein interactions from next-generation yeast-two-hybrid sequence data

Short title:

Statistical inference of protein-protein interactions with Y2H-SCORES

Valeria Velásquez-Zapata[a,b], J. Mitch Elmore[b,c], Sagnik Banerjee[a,d], Karin S. Dorman[a,d,e], Roger P. Wise[a,b,c]

[a]Program in Bioinformatics & Computational Biology, Iowa State University, Ames, IA 50011

[b]Department of Plant Pathology & Microbiology, Iowa State University, Ames, IA 50011

[c]Corn Insects and Crop Genetics Research, USDA-Agricultural Research Service, Ames, IA 50011

[d]Department of Statistics, Iowa State University, Ames, IA 50011

[e]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011

Correspondence to: roger.wise@ars.usda.gov

Abstract

Interactomes embody one of the most efficient representations of cellular behavior by revealing function through protein associations. Building such models requires reproducible and efficient high-throughput implementation of the prevailing molecular techniques to ascertain interacting partners. Among them, yeast two-hybrid (Y2H) and its high-throughput version, termed next-generation interaction screening (NGIS), comprise promising approaches to generate extensive protein-protein interaction networks. However, challenges remain to mining reliable information from these screens and thus, limit its broader implementation. Here, we describe a statistical framework, designated Y2H-SCORES, for analyzing high-throughput Y2H screens that considers key aspects of experimental design, normalization, and controls. Three quantitative ranking scores were implemented to identify interacting partners, comprising: **1)** significant enrichment under selection for positive interactions, **2)** degree of interaction specificity among multi-bait comparisons, and **3)** selection of *in-frame* interactors. Using simulation and an empirical dataset, we provide a quantitative assessment to predict interacting partners under a wide range of experimental scenarios, facilitating independent confirmation by one-to-one bait-prey tests. Simulation of Y2H-NGIS identified conditions that maximize interactor mining efficiency, which can be achieved with protocols such as prey library normalization, maintenance of larger culture volumes and replication of experimental treatments. Y2H-SCORES can be implemented in different yeast-based interaction screenings, accelerating data analytics of those technologies. Proof-of-concept was demonstrated by discovery and validation of a novel interaction between the barley powdery mildew effector, AVR_{A13}, with the vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1.

Keywords: Interactome inference; high throughput yeast two hybrid (Y2H); ranking interactions; interaction prediction; barley; *Blumeria graminis*; powdery mildew

Author Summary

Organisms respond to their environment through networks of interacting proteins and other biomolecules. In order to investigate these interacting proteins, many *in vitro* and *in vivo* techniques have been used. Among these, yeast two-hybrid (Y2H) has been adapted for use in combination with next generation sequencing (NGS) to approach protein-protein interactions on a genome-wide scale. The fusion of these two methods has been termed next-generation-interaction screening, abbreviated as Y2H-NGIS. However, the massive and diverse data sets resulting from this technology have presented unique challenges to analysis. To address these challenges, we optimized the computational and statistical evaluation of Y2H-NGIS to provide metrics to identify high-confidence interacting proteins under a variety of dataset scenarios. Our proposed framework can be extended to different yeast-based interaction settings, utilizing the general principles of enrichment, specificity, and *in-frame* prey selection to accurately assemble protein-protein interaction networks. Lastly, we showed how the pipeline works experimentally, by identifying and validating a novel interaction between the barley powdery mildew effector AVR_{A13} and the barley vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1. Y2H-SCORES software is available at GitHub repository https://github.com/Wiselab2/Y2H-SCORES.

Introduction

Investigations into the molecular interactions among hosts and pathogens has benefited from the plethora of omics datasets that can be used for the prediction of gene and protein networks. For plants, this knowledge can be used to guide modern plant breeding efforts in response to pathogens that present diverse challenges to the host through the identification of resistance proteins and their function (Dangl et al., 2013; Struk et al., 2019). In animal models, accurate networks can lead to the determination of drug targets and vaccines by providing a detailed view

of the host-pathogen interactions and the triggered immune response (Adams et al., 2011; Arkin et al., 2014; Koff et al., 2013; Oberg et al., 2011).

Well-characterized signaling networks are one of the most efficient models to understand molecular events during the infection process (Braun et al., 2013). These can be represented using protein interactome networks, where nodes represent proteins and edges represent physical interactions (De Las Rivas and Fontanillo, 2010). The yeast two-hybrid (Y2H) screen is one of most powerful tools for uncovering new protein-protein interactions (PPI), discerning connections between bait and prey proteins while correcting for biases in their cell concentrations and affinity (Vidal and Fields, 2014). However, traditional Y2H screens involve a labor-intensive step where distinct yeast colonies growing on selective media are picked, grown, and sequenced to identify prey cDNA fragments. Thus, to optimize genome-scale screens and obtain interactome data in a time-efficient manner, more recent approaches, collectively termed next-generation interaction screening (NGIS), use deep sequencing to score the result of Y2H screens. These innovations facilitate quantitative measures of bait-prey interactions, and importantly, do not require an ordered open-reading-frame library from the host of interest (Lewis et al., 2012; Pashkova et al., 2016; Suter et al., 2015; Trigg et al., 2017; Weimann et al., 2013; Yachie et al., 2016).

Despite the methodological advantages of Y2H-NGIS, there remain overlooked informatics and statistical challenges. **1)** Current pipelines map and quantify total reads while ignoring prey-fusion reads (reads containing both Y2H plasmid and prey cDNA) that provide frame information of the cDNA fusion protein. **2)** There is no consensus regarding what control screen is more appropriate to signify the background interactivity of the preys and to help to distinguish true interactions, having a non-selected control or a selected control with a random bait as options (Erffelinck et al.,

2018; Kessens et al., 2018; Pashkova et al., 2016; Trigg et al., 2017; Weimann et al., 2013; Yachie et al., 2016; Zong et al., 2020). **3)** Despite its importance, most existing methods do not assess data normalization, or implement inappropriate normalization methods for Y2H-NGIS data. High-throughput sequencing datasets, where read counts quantify signal strength, e.g., RNA-Seq, require normalization, as there are external factors, aside from experimental treatments, that influence read counts (Evans et al., 2018). Normalization methods, such as those used for RNA-Seq, assume that most genes in the sample are not differentially enriched (DE), but in Y2H-NGIS experiments, the enrichment of each prey is determined by completely different factors under the two conditions. In non-selected conditions, the prey's relative abundance in the library determines the enrichment, while in selection, it is the prey's ability to activate the reporter, via interaction with the bait, or by autoactivation.

Most if not all prey will therefore be DE in the selected condition. However, most Y2H-NGIS platforms do not use any normalization method nor methods that satisfy this assumption (Erffelinck et al., 2018; Pashkova et al., 2016; Trigg et al., 2017; Weimann et al., 2013; Yachie et al., 2016; Zong et al., 2020). Finally, **4)** with no consensus on the appropriate data analysis, nor even how to report the results, whether ratios of counts, log fold change from DE analysis, or a custom score function, it is nearly impossible to compare Y2H-NGIS studies, and there are reports that Y2H-NGIS-based interaction predictions suffer high false positive rates (Erffelinck et al., 2018; Kessens et al., 2018; Lewis et al., 2012; Pashkova et al., 2016; Trigg et al., 2017; Weimann et al., 2013; Yachie et al., 2016; Zong et al., 2020). Hence, there is a need for robust and consistent statistical models that make use of all the available information in Y2H-NGIS data.

We optimized the protocol proposed by Pashkova and colleagues (2016) to sub-culture diploid yeast populations that carry bait and prey plasmids under two batch conditions: **1)** diploid growth

obtained in what we call the non-selected condition (SC-Leu-Trp), used as background level of library prey abundance and **2)** interaction test, or the selected condition (SC-Leu-Trp-His), which theoretically only allow the growth of diploid populations with positive bait-prey interactions through activation of the reporter gene. Diploids that grow under selection comprise a group of preys that can be split into true and false interactors. True interactors are those preys that can be verified as bait-specific in a protein-protein binary interaction test. In contrast, false interactors represent the group of selected preys which do not produce a positive binary test, nor are bait-specific (autoactivating the reporter). Previously, we developed a robust informatics pipeline, designated NGPINT, to signify candidate interacting partners obtained from Y2H-NGIS, including mapping reads to the reference prey genome, reconstruction of prey fragments and fusions with the prey plasmid (Banerjee et al., 2020). Here, we propose statistical methods to rank the resulting preys, distinguish false vs. true interactors, providing to the user a high confidence list of candidates. This ranking system, designated Y2H-SCORES, is an aggregate of three experimental outcomes: **1)** the non-selected population as a baseline to detect which preys are significantly enriched under selection, **2)** selected samples as a control baseline to measure the specificity of a prey, and **3)** fusion read information to identify *in-frame* enrichment of the prey fragments under selection.

Using simulation and experimental validation we assessed the Y2H-SCORES success. Simulation of Y2H-NGIS allowed us to demonstrate its robustness under different dataset scenarios. Additionally, as a proof of concept, we used Y2H-SCORES to identify and confirm the interaction of the barley powdery mildew effector, AVR_{A13}, with the barley vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1. This interaction, accompanied with previous evidence (Rajaraman et al., 2018) and expression quantitative trait loci (eQTL) associations (Surana et al., 2017), support the involvement of HvTHF1 with resistance to powdery mildew.

Results

The effect of normalization on Y2H-NGIS data

Principal component analysis (PCA) of the logged Y2H-NGIS raw read counts identified selection as the major source of variability (Fig 1A). The effect of selection and baits (colors in Fig 1A) under selection are expected sources of variation, but we also expect all non-selected samples to resemble the cDNA used to build the prey library and the three replicates to cluster. Indeed, if there is more variation among replicates than baits, it could prove difficult to reproducibly identify bait interactors.

We considered four normalization methods to reduce the experimental variation, particularly to reduce variation across replicates. All preys are expected to be differentially enriched (DE) in selected compared to non-selected samples; interactors since they grow exponentially and non-interactors since they cannot reproduce under selection. Our goal is to identify prey whose relative abundance in the selected samples *increases* over the relative abundance in the non-selected samples. Normalization methods appropriate for this goal include library size (Dillies et al., 2013), transcripts (or in this case, preys) per million (TPM) (Wagner et al., 2012), and remove unwanted variation (RUVs) (Risso et al., 2014). Many normalization methods are designed to detect enrichment relative to unchanging reference genes, which simply do not exist in Y2H-NGIS data. Specifically, we used median-of-ratios (Anders and Huber, 2010), which assumes the majority of genes are not DE, as a control method that should fail to normalize Y2H-NGIS data. Fig 1B-E show the PCA plots of the total counts after implementation of the different normalization methods. TPM, RUVs, and library size reduced the variability in the non-selected samples to

varying degrees but retained most of the other variation. The median-of-ratios method, in contrast, removed over half of the selected vs. non-selected variation.

Ideally, all non-selected samples should resemble the prey library. Indeed, Pearson correlations of the count data between all pairs of non-selected samples exceeded 0.98 for all but RUVs normalization (Fig 1F), indicating that non-selected prey counts are largely bait-independent, as expected. In contrast, prey counts were much less correlated among selected samples, presumably reflecting the effect of the baits. Library size and TPM normalizations increased the non-selected sample correlation (Wilcoxon signed-rank test p-value of 9.4×10^{-39} and 5.5×10^{-18}, respectively) over the raw counts (S1 Table), but RUVs normalization substantially decreased it (Wilcoxon signed-rank test p-value of 2.7×10^{-81}). RUVs, which seeks factors explaining variation across replicates, may have retained greater variation within non-selected samples because we used just one factor to explain the technical variation.

Appropriate normalization should reduce inter-replicate variability. We measured the variability across replicates using the coefficient of variation (CV), computed for each prey. As illustrated in Fig 1F and S1 Fig, we found that RUVs, library size, and TPM normalizations reduced the CV compared to raw counts (Wilcoxon signed-rank test p-values <0.05, S2 Table). However, there is no one single method that works the best in all cases (S2 Table). For non-selected conditions, we found that all normalizations performed very well with CV peaks within 0 - 0.5, which indicate a low variation between replicates. TPM and median-of-ratios had the best performance with a narrower density. The median-of-ratios method can perform well in non-selected conditions since as compared with the selected conditions, they have only a small portion of DE preys that come from the transcriptome used to make the prey library (Hunt et al., 2019). Contrast in performance was highest when normalizing selected samples. In this case, the ranking of methods was bait

specific (S1 Fig, S2 Table), though the median-of-ratios consistently failed to reduce the variation. CV distributions for library size, TPM and RUVs peaked within 0 - 1, but the CV distribution for the median-of-ratios method peaked in the range 1 - 1.5. After evaluating the results from the Pearson correlation and CV analyses we decided to use library size normalization as the main method for the Y2H-NGIS dataset.

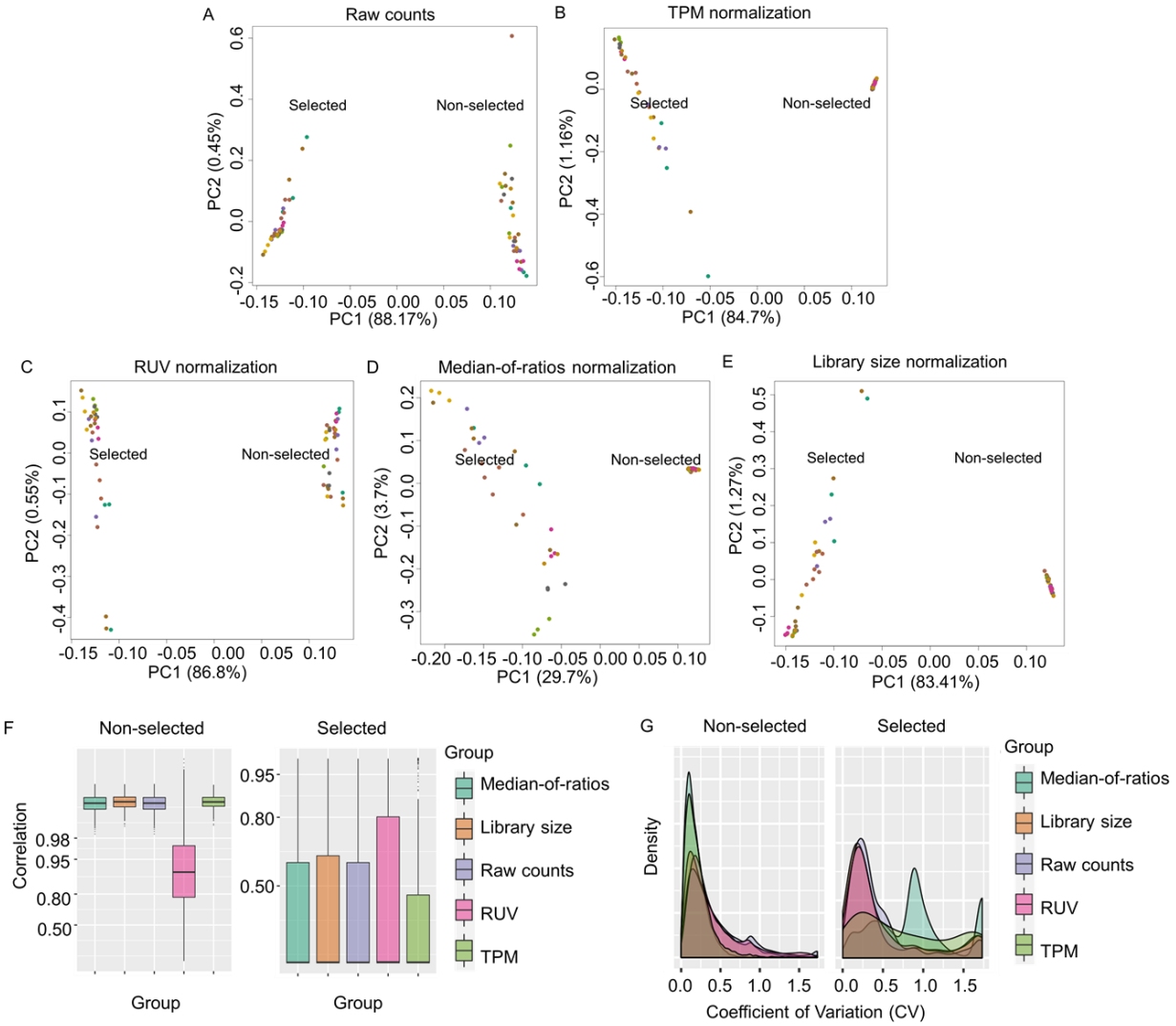

Fig 1. Effect of count normalization in Y2H-NGIS. A) PCA analysis of raw read counts and B) TPM, C) RUVs, D) Median-of-ratios and E) Library size normalized reads for selected (left) and non-selected (right) samples for 10 bait screenings (color coded). F) Boxplot of the pairwise correlation coefficients for raw and normalized read counts for all samples in non-selected and,

separately, selected conditions. G) Coefficient of variation (CV) for each prey using different normalization methods in the bait AVR_{A13}. Higher CV values may indicate poor performance because of a high variation between replicates.

Y2H-SCORES identifies true interacting partners

After optimizing normalization, we proposed a set of ranking scores based on statistical assessments of the count data to predict interacting partners. We considered the biological principles that define PPIs and key aspects of Y2H-NGIS, such as the experimental design, normalization, and controls (see methods). Summarizing, we modeled the total prey counts using a Negative Binomial (NB) regression and the *in-frame* fusion counts using the Binomial distribution. We designed a set of three quantitative ranking scores, called Y2H-SCORES, to identify interacting partners: **1)** Enrichment score: a measure of significant enrichment under selection for positive interactions, using as control the non-selected samples; **2)** Specificity score: a measure of the specificity of a bait-prey interaction, using other selected baits as controls; and **3)** *In-frame* score: a measure of the enrichment for *in-frame* proteins in selected samples. To test Y2H-SCORES, we designed a Y2H-NGIS simulator, motivated by real data. The simulator includes non-specific interactors, which are genes that survive selection because the product protein interacts with multiple baits, and true interactors, genes that are strongly selected in the presence only of their cognate bait.

We began by simulating idealized conditions of 10 bait screens with three replicates, a cDNA prey library of 20,000 genes, 1 - 20 true interactors per bait, a stickiness (percentage of sticky preys in the library) of 0.1%, and a strength of interaction for true interactors above the 99.9th percentile. The strength of true interactors was quantified with a fitness coefficient e_{ik}, which we estimated

for all prey in our real data. In this simulation, we reserved the top 0.1% of all estimated fitnesses for the true interactors. Fig 2A-C demonstrates that all scores performed well, via the Receiver Operating Characteristic (ROC) curves, to separate true and non-specific interactors. In this scenario all the scores achieved good performance: the enrichment, specificity and *in-frame* scores had an ROC Area Under the Curve (ROC AUC) of 0.98 for enrichment, 1 for specificity and 0.99 for *in-frame*. The Precision Recall (PR) AUC were 0.47, 0.62, and 0.54, respectively. We plotted the PCA of the Y2H-SCORES under this scenario (Fig 2D) and we found the enrichment and specificity scores are more related with each other than with the *in-frame* score, nonetheless the three scores seem to provide different information based on their position in the plot.

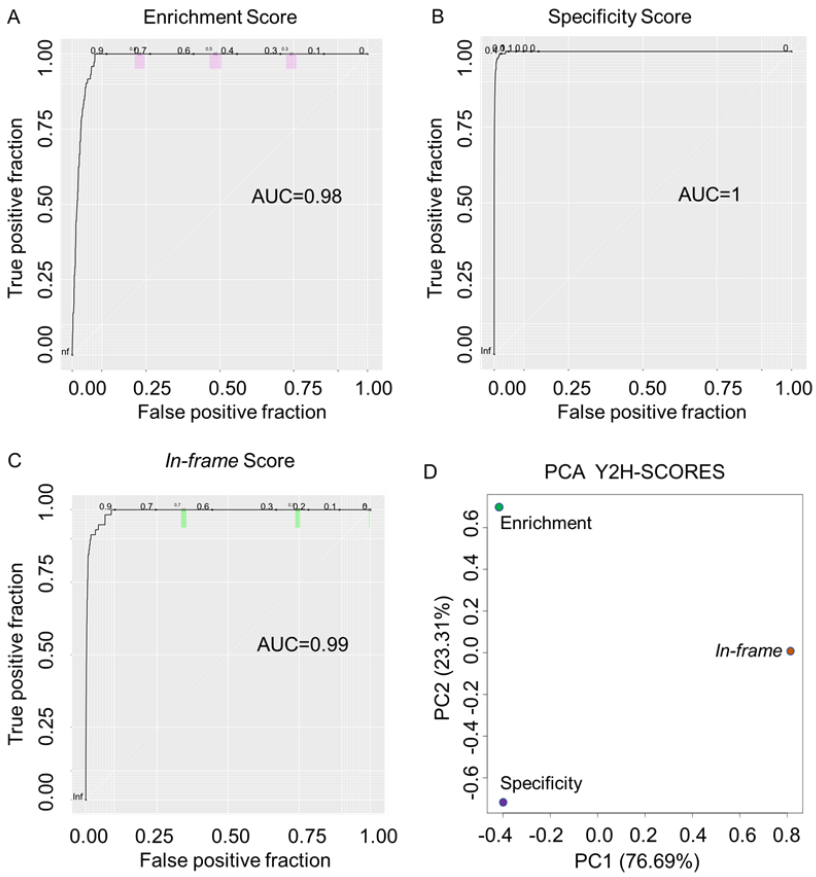

Fig 2. Performance of the ranking scores in an ideal scenario. A) to C) ROC curves of the enrichment, specificity, and *in-frame* scores. Colored sections represent 95% confidence intervals

for the score values 0.7, 0.5, and 0.3. D) PCA of the Y2H-SCORES calculated under the ideal scenario.

Y2H-SCORES overcomes challenging Y2H-NGIS scenarios

The ideal condition simulated for Fig 2 was discovered from an extensive simulation study where we explored the effect of several parameters that vary in experimental datasets, as defined in S3 Table. The simulator uses a Galton-Watson branching process followed by a NB model for generating total counts, and a binomial model for fusion counts (see S2 Text). Hence, we generated scenarios varying the size of the prey library to assess scalability, the stickiness of the samples as well as strength of true interactors to vary the signal-to-noise ratio, overdispersion of true interactors to consider increasing levels of biological and experimental variation, initial proportion of true interactors in selected samples to assess the role of genetic drift, and the number of baits and replicates to increase the power. Except for the true interaction strength and the stickiness, most of these parameters are directly associated with the cDNA prey library and experimental conditions, which can be improved by the researcher. A graphical summary of the results is shown in Fig 3. Briefly, the scores were able to correctly identify true interactors even in extreme conditions, but the variation in performance helps us identify ideal experimental conditions for detecting these interactors.

The scalability of the Y2H-SCORES was evaluated by testing three prey library sizes (8000, 20000 and 40000 preys). We found that increasing the library size maintained the performance of the scores (S3 Table). The PR AUC values of the enrichment score oscillated between 0.47 and 0.68, the specificity from 0.62 and 0.80 and the *in-frame* score from 0.54 to 0.76, while the

ROC AUC remained constant. This result shows that even with large library sizes the Y2H-SCORES perform well and therefore, they can still be used to identify protein-protein interactions.

We then tested the effect of the stickiness of the samples and the strength of true interactors on the Y2H-SCORES performance. The results from our simulations, shown in Fig 3A-B, suggest that Y2H-SCORES performance is less influenced by changes in the stickiness than by the strength of true interactors. Keeping the strength of true interactors above the 99.9 percentile and variations of the stickiness between 0.1% to 10%, did not cause major changes in the ROC and PR AUC values. This result indicates that Y2H-SCORES is able to identify sticky interactors, even in when they comprise 10% of the preys in the sample. In contrast, the strength of true interactors had a greater effect on the performance of Y2H-SCORES. As we decreased the strength of true interactors from the 99.9th to the 95th percentile, we found that the PR AUC values dropped to near zero. ROC curves were more stable, showing a gradual decrease.

To evaluate the effect of experimental variation we tested changes in the overdispersion. We simulated two scenarios, either a high or random overdispersion in both the selected and non-selected condition. After estimating the overdispersion parameters observed in real data, we jointly sampled the proportion of preys and the overdispersion φ_{kN} in the non-selected samples, and the fitness and overdispersion values φ_{ikS} in the selected samples, from the joint empirical distributions. In the overdispersed scenario, we resampled φ_{kN} and φ_{ikS} values higher than the 90th percentile of their densities. The scores' performance was maintained in scenarios with high overdispersion as measured by both PR and ROC AUC values (Fig 3C).

The initial proportion of each prey before culture expansion depends on the composition of the cDNA prey library, which can be controlled through experimental library normalization. We assumed the post-expansion prey proportions in the non-selected samples were identical to the unobserved prey proportions at the beginning of selection. Thus, we sampled these initial proportions from the observed non-selected proportions. We expect more inter-replicate variability and lower power to detect true interactors when the initial true interactor proportion is low because of initial sampling variation and greater genetic drift during culture growth. To simulate the effect of a low concentration of preys in the library we used the minimum proportion q_{ik} that we observed in our experimental dataset, ~1×10^{-8}, as reference value and assigned it to the true interactors in the "Low" condition (Fig 3D). Results from this analysis showed that the PR AUC decreased for all three scores. The enrichment and specificity scores decreased from 0.47 to 0.30, and from 0.62 to 0.17, respectively. The largest decrease was observed for the *in-frame* score, going from 0.54 to 0.005. Scenarios with low initial proportions of true interactors in non-selected samples caused a low number of total prey reads in the non-selected condition, and a reduction or even absence of fusion reads (which normally represent a small fraction of the total number of reads). This trend was observed in experimental datasets too, where we observed a large number of preys with no fusion reads available in the non-selected samples.

Detection of DE prey with statistical confidence requires replication, but more replicates increase the time and cost of the sequencing project. We evaluated the effect of having two, three and five replicates. Increasing the number of replicates increased the performance of the enrichment and the specificity scores, while the *in-frame* score was not affected (Fig 3E, S3 Table). The *in-frame* score maintained a good performance even in cases with two replicates, with PR AUC around 0.53, but the enrichment and specificity scores had reduced performance. The enrichment score had the greatest reduction in the PR AUC values going from 0.61 (five replicates) to 0.40 (two

replicates), and the specificity PR AUC values went from 0.65 to 0.55. Finally, we tested the effect of the number of baits using values from two to ten (Fig 3F). The enrichment and *in-frame* scores showed a decrease in their PR AUC values only in the case with two replicates with values of 0.35 and 0.44, respectively. In contrast, the performance of the specificity score improved with more baits in the simulation, with PR AUC values increasing from 0.47 to 0.71. The specificity information provided by the additional bait screenings increases the resolving power of the specificity score.

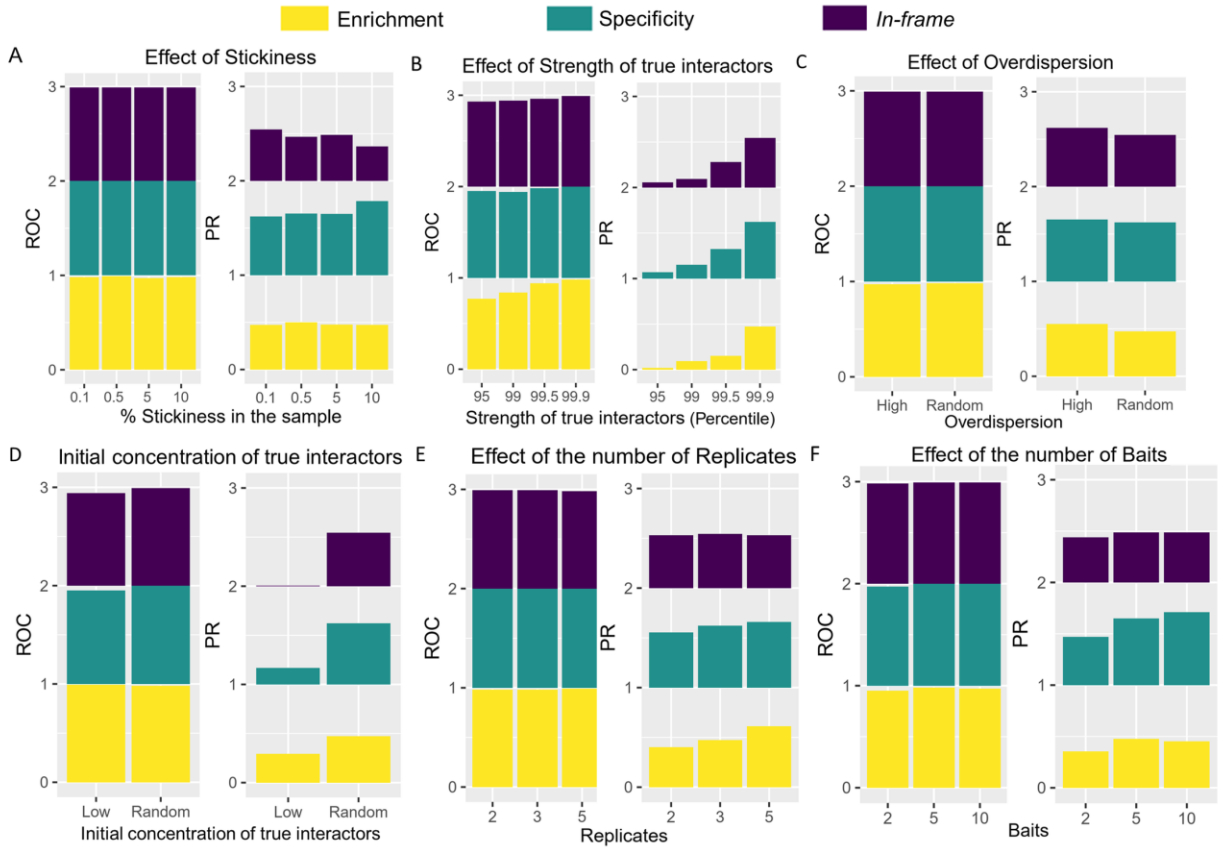

Fig 3. Effect of changes in the parameters that define Y2H-NGIS simulation. Examples of challenging scenarios were simulated to determine the Y2H-SCORES classification power. A) Stickiness, B) Strength of true interactors, C) Overdispersion, D) Concentration of true interactors before selection, E) Number of replicates, and F) Number of baits. ROC and PR AUC values were reported for the enrichment, specificity, and *in-frame* scores.

Y2H-SCORES discards sticky interactions and identifies a new interaction between AVR_{A13} and HvTHF1

As proof of concept, we used Y2H-SCORES to identify interactors of the AVR_{A13} effector from the barley powdery mildew pathogen, *Blumeria graminis* f. sp. *hordei (Bgh)* (Lu et al., 2016). Our framework allowed us to discard sticky interactors and to confirm the interaction between AVR_{A13} and the barley vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1. We screened AVR_{A13} using an experimental setting of three replicates in the non-selected and selected condition. We ran our bioinformatic pipeline to map and quantify reads in the samples and obtained the prey regions for cloning. After this, we calculated the three Y2H-SCORES and created a Borda ensemble (Lin, 2010) to obtain a list of candidate interactors. Interestingly, after running Y2H-SCORES with different normalizations we found an increase in the number of highly ranked candidates with the median-of-ratios method. When we dissected this trend, we observed an increase in the number of candidates with high enrichment and *in-frame* scores and low specificity score (Wilcoxon ranked-sum test, S4 Table; S2 Fig), theoretically indicating sticky preys. The top-scoring preys unique to this list had low specificity scores across all normalization methods (S5 Table). We performed binary Y2H with two of these preys, corresponding to the gene IDs HORVU2Hr1G060120 (TCP family transcription factor 4) and HORVU2Hr1G024160 (Chaperone protein DnaJ-related), and confirmed that they were sticky (S3 Fig).

Using the Y2H-SCORES calculated from library size normalization, we focused on preys with high Borda ensemble scores, and therefore, high Y2H-SCORES values. HvTHF1, corresponding to the barley gene ID HORVU2Hr1G041260, was identified as the top candidate interactor of AVR_{A13} with the following Y2H-SCORES: enrichment of 0.9743, specificity of 0.8986, *in-frame* of 1 and Borda ensemble of 1203.73. To test the binary interaction, the AVR_{a13} bait sequence was fused

with the GAL4 transcription factor activation domain (GAL4-AD), while the *HvThf1* prey sequence was fused with the GAL4 DNA binding domain (GAL4-BD). After mating, diploid selection in SC-LW media and selection of the interaction with SC-LWH and SC-LWH + 3AT, we obtained a strong and specific positive interaction. Fig 4B shows the results from this binary test, validating the interaction between this protein and the effector.

The prey consisted of a fragment of 860 nucleotides and an open reading frame of 170 amino acids (S1 Text contains fasta files with the prey sequence and protein translation). We analyzed the prey fragment peptide to identify the interacting domains and we found it contains a thylakoid formation protein domain. To position HvTHF1 in a signaling context we predicted protein-protein interactors using interlogs (Gu et al., 2011; Jafari et al., 2015; Nakajima et al., 2018; Stark, 2005). We found 18 predicted interactions with four main protein families: Homeobox protein BEL1, DOF zinc finger protein, protein kinase, receptor kinase and leucin-rich receptor-like kinase. In addition, previous expression quantitative trait locus (eQTL) analysis indicated that *Thf1* associates significantly with the *Mla1* (*mildew resistance locus a1*) *trans*-eQTL at 32 hours after inoculation (HAI), the timeframe coinciding with development of *Bgh* haustorial feeding structures (Surana et al., 2017). As illustrated in Fig 4, many of the predicted interactors for the THF1 protein are also associated with significant *trans*-eQTL at different powdery mildew infection stages, *MlLa (Laevigatum* resistance locus) at penetration and *Mla1* during haustorial development (Surana et al., 2017).

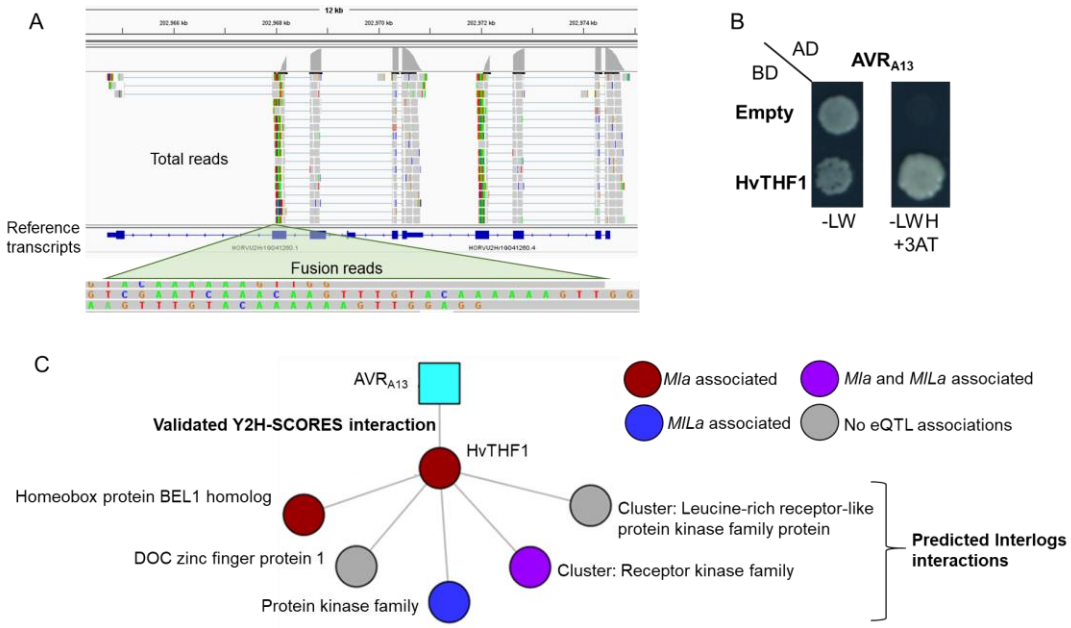

Fig 4. Experimental validation of the interaction between the barley powdery mildew effector, AVR$_{A13}$ and the vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1. A) IGV mapping of reads to the prey region and zoom to fusion reads, used for primer design and cloning. B) Binary Y2H test between AVR$_{A13}$ and HvTHF1 showing the diploids control (SC-LW) and the stringent interaction (SC-LWH+3AT). C) Prediction of interlogs for the AVR$_{A13}$ protein target, HvTHF1 (validated using Y2H-SCORES). *Trans*-eQTL associations (Surana et al., 2017) with the *Mla1* (powdery mildew resistance) and *MlLa (Laevigatum* resistance locus) are color coded.

Discussion

Analysis of Y2H-NGIS data

Analysis of Y2H-NGIS data is challenging due to the complexity of the raw dataset (composed of total and fusion prey reads under both selective and non-selective conditions) and the substantial variability across replicates. Banerjee and colleagues (2020) proposed a bioinformatic pipeline to

map and quantify raw reads from Y2H-NGIS, providing both total and fusion prey counts under selective and non-selective conditions. In this report, we outline Y2H-SCORES, a framework to rank candidate prey/bait interactions based on these count data. Using a Negative Binomial (NB) regression we modeled total counts, and the *in-frame* fusion counts were analyzed using the Binomial distribution. From these models we designed Y2H-SCORES to identify interacting partners based on their properties: **1)** Enrichment score: measures the enrichment under selection for positive interactions, as compared with non-selected conditions; **2)** Specificity score: assigns higher values to unique bait-prey interactions, as compared to prey selected in multiple bait screens; and **3)** *In-frame* score: measures the enrichment for *in-frame* proteins in selected samples. We validated the method and used simulation to evaluate the impact of several experimental factors on the power to detect true interactions and the accuracy of the rankings. We found that normalization methods and controls have a profound impact on the amount of information that can be used to identify interactors.

Normalization significantly modified the variation among replicates. Utilizing methods whose assumptions are satisfied by the Y2H-NGIS dataset leads to a more successful interpretation of the observed variation to infer protein interactors. Library size, TPM and RUVs are appropriate normalization methods for Y2H-NGIS data, but their ability to reduce variance within replicates varied, therefore, we recommend that users evaluate them individually and decide which one works better for their experiment. Median-of-ratios normalization is not appropriate for Y2H-NGIS data, and applying this method increased the number of candidate interactors with high enrichment and *in-frame* scores and low specificity score, compared to the other three normalization methods. Median-of-ratios normalization appears to promote non-interactors and sticky preys (which activate the HIS3 promoter without an interaction with the bait), since the top-scoring preys unique to this list had low specificity scores across all normalization methods (S5

Table) and two of them were confirmed as sticky (S3 Fig). Performing appropriate normalizations removed these sticky preys from the top ranked list. Overall, median-or-ratios normalization produced lists with higher enrichment and *in-frame* scores, but lower specificity scores than the other methods (Wilcoxon rank-sum test, S4 Table).

Additionally, different controls that have been proposed for Y2H-NGIS can be used to measure different properties of an interactor. First, we used non-selected controls as an enrichment baseline of preys, allowing the implementation of the enrichment and *in-frame* scores. Second, we demonstrated the advantage of using screenings under selection for multiple baits as a second type of control that provides information for the specificity score. The baits used for this purpose may contain an empty bait, a random bait, or a set of baits of interest. Our PCA analysis of Y2H-SCORES (Fig 2D) suggest independent information coming from each of the three scores, hence we recommend using both kinds of controls and all three scores to obtain a high-confidence list of interactors.

Development and testing of Y2H-SCORES has yielded some suggestions for the design of Y2H-NGIS experiments. Calculating the enrichment and the *in-frame* scores requires selected and non-selected samples. As we demonstrated (Fig 1F), composition of the non-selected sample is almost identical regardless of the bait (Pearson correlation > 0.98), information that can be utilized to reduce the number of sequenced samples, e.g., using a few random baits as non-selected controls for multiple baits in the same mating design. In contrast, the specificity score requires at least two different bait screenings, with better results as the number is increased. This strategy exploits the count information of multiple selected baits to discard sticky preys, giving priority to specific candidate interactors.

Experimental setting and optimization of Y2H-SCORES

Different experiment scenarios allowed us to test the robustness of Y2H-SCORES and identify the most challenging dataset types (Fig 3). The three scores were affected differently depending on the simulation scenario. Scenarios with low strength and low initial concentration of true interactors imposed the most challenging conditions for these scores, reducing the ROC and PR AUC values. The enrichment score was more affected by the strength of true interactors while the *in-frame* score was more affected by the initial concentration of true interactors, due to the inherent lower number of fusion reads. This analysis led us to explore aspects of experimental design that could be adjusted to increase the accuracy and sensitivity of interactor detection via Y2H-NGIS. These include the experimental prey library normalization, the number of replicates and baits in the experiment, sequencing depth, and scaling of the experiment setting.

First, library normalization determines the proportion of each prey in the library, in the non-selected samples and at the start of selection. Normalizing the prey library to avoid low-abundance preys in the library reduces the stochasticity and noise in prey counts. After normalizing the prey library, experimentalists should also ensure the number of yeast recipient cells are sufficient to represent such library in the screens, for which they can use procedures as described by Krishnamani et al., 2019. Our simulations found that low abundance interactor preys in the library (200 times lower than the expected prey abundance) can be detected in Y2H-NGIS, but the lower the initial proportion, the stronger the interaction with the bait needed to be relative to the interaction strength of the sticky prey. Thus, normalizing the prey library can reveal weaker true interactors.

A second parameter, the number of replicates in Y2H-NGIS, represents a cost-power trade off as in most “omics” experiments. For small numbers of replicates, as tested in the simulations, we found at least three replicates of Y2H-NGIS were needed to maintain the performance of the three scores. As expected, we observed better results as the number of replicates increased, especially for the enrichment score. We did not test replicate numbers greater than 5 since this does not represent typical experimental practices, and if one had to choose, increasing the number of baits would yield more biological information, since it would increase the performance of the specificity score. We anticipate that increasing the number of replicates would decrease the false discovery rate as it is reported for techniques such as RNA-Seq (Schurch et al., 2016). Controlling for false discovery rate also informed our selection of DESeq2 as the tool for calculating differential enrichment due to the documented outperformance in low replicate numbers (Evans et al., 2018; Schurch et al., 2016).

It has been reported that reducing overdispersion in counts can improve sensitivity and accuracy of Y2H-NGIS (Pashkova et al., 2016). We did not observe a decrease in the performance of the Y2H-SCORES as we increased the overdispersion, which may be explained by the strategies to control it. As a result, the overdispersions estimated from our data may already be lower than what we could have observed with a different experimental setting. The main recommended strategies to control overdispersion among samples include maintaining a large-scale mating (in our experiment, $1x10^8$ bait and $1.17x10^7$ prey cells, obtaining ~$2x10^9$ diploid cells afterwards) and subsequent high culture volumes of non-selected and selected samples (typically 800 ml per sample in 2-liter fluted Erlenmeyer flasks for ca. 40,000 preys in the library). Increasing the volume reduces the stochasticity of the prey population before mating and during culture expansion. Stochasticity is most notable in selected growth, where genetic drift dominates as the viable prey population shrinks. Population bottlenecks must be avoided throughout the experiment, which

implies increasing the aliquot size in every culture step, including the final sampling for sequencing, as reported by Pashkova and associates (2016). We also recommend adjusting the sequencing depth to match the prey library size and specifically increasing depth for the more complex, non-selected samples. Having a high depth in non-selected samples also increases the number of fusion reads, the major pitfall for the successful implementation of the *in-frame* score in our simulations.

Extrapolation of Y2H-SCORES to other yeast-based interaction methods

The analytical framework we propose here can be extrapolated to other yeast-based interaction settings, e.g., testing for DNA-protein interactions through yeast one-hybrid (Reece-Hoyes and Marian Walhout, 2012) or multiplexed yeast two-hybrid (Trigg et al., 2017). The similarities and limitations among these techniques make them suitable candidates for the implementation of the Y2H-SCORES framework. All share the same true interactor properties: they should be enriched in selected samples, be specific to a bait screen and be selected *in-frame*. In addition, these techniques require control for false positives with appropriate statistics. Currently, these techniques propose a solid media selection of interactors, nonetheless the batch culture experimental setting proposed by Pashkova and colleagues (2016) can also be applied to these contexts, increasing the reproducibility and facilitating the experimental workflow. Once output counts are obtained, it will be possible to calculate the Y2H-SCORES and use them to identify true interactors.

Experimental validation of Y2H-SCORES identifies a new interaction between AVR_{A13} and HvTHF1

As an example of the effectiveness of our ranking scores, we identified and validated a new interaction between the barley powdery mildew effector, AVR_{A13}, and the vesicle-mediated thylakoid membrane biogenesis protein, HvTHF1. HvTHF1 has been found to interact in yeast and *in planta* with proteins encoded by the U-box/armadillo-repeat E3 ligase *HvPUB15* and a partial gene duplicate, *HvARM1* (for *H. vulgare* Armadillo 1) (Rajaraman et al., 2018). Neo-functionalization of *HvARM1* increases resistance to powdery mildew and provides a link between plastid function and colonization by biotrophic pathogens. In the broader context of plant-pathogen interactions, THF1 and its homologs interact with different protein mediators of plant resistance and susceptibility. The wheat homolog of *HvThf1*, *TaToxABP1,* encodes a target of the necrotizing Toxin A effector from the tan-spot fungal pathogen, *Pyrenophora tritici-repentis,* which has been associated with ROS (reactive oxygen species) burst (Manning et al., 2007; Pandelova et al., 2012). In addition, THF1 has been found to destabilize several NB-LRR-type resistance proteins by binding their I2-like coiled-coil (CC) domains (Hamel et al., 2016), interactions that were also predicted by our interlogs search.

Previous eQTL analysis performed by our group, combined with an interlogs search, revealed significant associations of *Thf1* and its primary interactors with the *Mla1* and *MlLa trans*-eQTL (Fig 4; Surana et al., 2017). These *trans*-eQTL associations were observed at different powdery mildew infection stages, *MlLa* at penetration, and *Mla1* during haustorial development**.** This initial analysis provides an additional entry point to support the involvement of HvTHF1 with resistance to powdery mildew disease.

Methods

Normalization

We implemented library size (Dillies et al., 2013), transcripts per million (TPM) (Wagner et al., 2012), removing unwanted variation (RUVs) with replicate control samples (Risso et al., 2014) and the median-of-ratios (Anders and Huber, 2010) normalization methods. RUVs was applied to selected and non-selected samples separately, using k=1 factor and grouping the three replicates for each bait in the selected condition, and all replicates for all baits in the non-selected condition. Median-of-ratios was applied for all baits and conditions, grouping the replicates for each bait-condition combination. The coefficient of variation (CV) for each prey was calculated analyzing each bait-condition combination separately and grouping the three replicates for each prey. Pairwise differences between the distributions of the CV were detected with a Wilcoxon signed-rank test on the prey CVs computed after application of each normalization method for each bait. Pearson correlation was calculated for each method separately and within replicates for each bait-condition combination. Differences between correlations for each normalization method in each condition were addressed using a Wilcoxon signed-rank test.

Differential enrichment analysis of prey counts

We modeled the optionally normalized prey count data with Negative Binomial (NB) regression (Love et al., 2014). This distribution allows for the effects of selection on the mean counts, while accounting for overdispersion across replicates, a consequence of biological and experimental variability. Specifically, we calculated the significance and magnitude of the enrichment for each prey gene interacting with each bait using the DESeq2 model (Love et al., 2014). Two different baseline controls were used to identify interactors: the non-selected condition for enrichment and the other selected baits for specificity.

Modeling fusion reads

Let Y_{ikc} be the number of *in-frame* reads out of a total F_{ikc} fusion reads for prey k mating with bait i in condition $c = N, S$ (*N*=non-selected, *S*=selected). We modeled $Y_{ikc} \sim Bin(F_{ikc}, \pi_{ikc})$, where π_{ikc} is the proportion of *in-frame* reads. To test *in-frame* enrichment under selection, we pose the hypothesis: $H_o: \pi_{ikN} = \pi_{ikS}$ vs. $H_a: \pi_{ikN} < \pi_{ikS}$, testing for an increase in the *in-frame* read proportion under selection. We evaluated this hypothesis using the *Z*-score statistic ρ_{ik}:

$$\rho_{ik} = \frac{\hat{\pi}_{ikS} - \hat{\pi}_{ikN}}{\sqrt{\hat{\pi}_{ik}\left(1 - \hat{\pi}_{ik}\right)\left(\frac{1}{f_{ikS}} + \frac{1}{f_{ikN}}\right)}} \sim N(0,1) \quad with \;\; \hat{\pi}_{ik} = \frac{f_{ikS}\;\hat{\pi}_{ikS} + f_{ikN}\;\hat{\pi}_{ikN}}{f_{ikS} + f_{ikN}},$$

where $\hat{\pi}_{ikc}$ is the observed *in-frame* read proportion and f_{ikc} is the observed number of fusion reads for prey *k* mated with bait *i* in condition *c*.

Y2H-SCORES

We implemented and validated a ranking score system, designated Y2H-SCORES, for identifying interacting partners from Y2H-NGIS. It is comprised of three elements, each in the range [0,1], with values close to 1 indicating high support for a true interaction. Throughout, n_p is the number of prey and n_b is the number of baits.

Enrichment score.

This score quantifies the level of enrichment of a prey under selection relative to non-selection. Let p_{ik} be the *p*-value and f_{ik} the log2 fold change in the normalized counts of prey k interacting with bait *i* in selected over non-selected condition as given by DESeq2. The score consists of a

system of ranks within ranks of these values to prioritize interactors. We first define the rank for p_{ik}: Consider $G_\alpha = \{(i,k): 1 \leq i \leq n_b,\ 1 \leq k \leq n_p,\ p_{ik} \leq \alpha\}$, the set of putatively interacting prey/bait combinations with *p*-values $p_{ik} \leq \alpha$. All *p*-values larger than α are assigned a value of zero. The *p*-value enrichment score is defined only for *i, k* bait/prey combinations with $p_{ik} \leq \alpha$ as

$$E(p_{ik}) = \frac{R_\alpha(\alpha) - R_{pik}(p_{ik})}{R_\alpha(\alpha)},$$

where $R_\alpha(p)$ is the rank of *p*-value p among the p_{ik} with indices in the set G_α. To further resolve prey/bait interactions, we score the effect size by partitioning G_α into $b = \frac{\alpha}{w}$ subsets $\{G_{\alpha 1}, G_{\alpha 2}, \dots, G_{\alpha b}\}$, where $G_{\alpha l} = \{(i,k): 1 \leq i \leq n_b,\ 1 \leq k \leq n_p, (l-1)w \leq p_{ik} \leq lw\}$ for $1 \leq l \leq b$, contains a subset of the $n_p * n_b$ prey/bait combinations with similar *p*-values. The rank of the log2 fold changes is calculated within the $G_{\alpha l}$ subsets to obtain the fold change enrichment score as

$$E(f_{ik}) = \frac{max\big(R_l(f)\big) - R_l(f_{ik})}{max\big(R_l(f)\big)},$$

and $max\big(R_l(f)\big)$ is the maximum rank of fold-changes with indices in $G_{\alpha l}$. Finally, we combined these two scores, ranking first by *p*-value and second by log2 fold-change, to obtain the enrichment score for (i,k) bait/prey combination as

$$E(ik) = E(p_{ik}) + \frac{max\Big(E\big(p_{G_{\alpha l}}\big)\Big) - min\Big(E\big(p_{G_{\alpha l}}\big)\Big)}{N(G_{\alpha l})} E(f_{ik}),$$

where the (i,k) combination is contained in $G_{\alpha l}$, $E\big(p_{G_{\alpha l}}\big)$ is the set of *p*-value enrichment scores from $G_{\alpha l}$ and $N(G_{\alpha l})$ is the size of $G_{\alpha l}$. Finally, to rescale $E(ik)$ between [0,1] we divided the score by the maximum value as: $E_{ik} = \frac{E(ik)}{max_{i,k}\,(E(ik))}$.

Specificity score.

True interactors should interact with few specific partners. To develop the specificity score we penalized preys that were enriched under selection with multiple bait screenings. Define p_{ijk} as

the *p*-value and f_{ijk} as the log2 fold change obtained from DESeq2, of prey k mated with bait i over prey k mated with bait $j \neq i$, both in selected conditions. We define the *p*-value specificity score $S(p_{ijk})$, just as $E(p_{ik})$ was defined before. Let $G_{s\alpha} = \{(i,j,k): 1 \leq i \leq n_b,\ \ j < i,\ 1 \leq k \leq n_p,\ p_{ijk} \leq \alpha\}$. Then:

$$S(p_{ijk}) = \frac{R_{s\alpha}(\alpha) - R_{s\alpha}(p_{ijk})}{R_{s\alpha}(\alpha)}$$

where $R_{s\alpha}(p)$ returns the rank of p among the *p*-values in set $G_{s\alpha}$. If $f_{ijk} < 0 \vee p_{ijk} > \alpha \rightarrow S(p_{ijk}) = 0$. We average the *p*-value specificity scores by $n_b - 1$ number of bait comparisons to obtain the final specificity score of bait/prey combination (i,k),

$$\mathrm{S}(p_{ik}) = \frac{1}{n_b - 1}\sum_{\mathrm{j}\neq\mathrm{i}} \mathrm{S}(p_{ijk})$$

We defined $S(f_{ijk})$, just as we did before for $E(f_{ik})$, partitioning $G_{s\alpha}$ into $b = \frac{\alpha}{w}$ subsets $\{G_{s\alpha 1}, G_{s\alpha 2}, \dots, G_{s\alpha b}\}$, where $G_{s\alpha l} = \{(i,k): 1 \leq i \leq n_b,\ j < i,\ 1 \leq k \leq n_p,\ (l-1)w \leq p_{ijk} \leq lw\}$ and $S(p_{ijk}) = 0 \rightarrow S(f_{ijk}) = 0$. We average over the $n_b - 1$ number of scores $S(f_{ijk})$ for (i,k) to obtain $S(f_{ik})$:

$$S(f_{ijk}) = \frac{max(R_{sl}(f)) - R_{sl}(f)}{max\ (R_{sl}(f))} \qquad S(f_{ik}) = \frac{\sum_{j\neq i} Sf_{ijk}}{n_b - 1}$$

where p_{ijk} is in $G_{s\alpha l}$ and $R_{sl}(f)$ the rank of the fold change f among those indexed in the lth subset $G_{s\alpha l}$. The combined specificity score is given by:

$$S(ik) = S(p_{ik}) + \frac{max\left(S(p_{G_{s\alpha l}})\right) - min\left(S(p_{G_{s\alpha l}})\right)}{(N(G_{s\alpha l}))} S(f_{ik})$$

With a rescaling to [0,1], as: $S_{ik} = \frac{S(ik)}{max_{i,k}\ (S(ik))}$

In-frame score.

We expect that true interactors will tend to appear *in-frame* under selective conditions. We convert the *in-frame* test of proportions statistic ρ_{ik} into the *in-frame* score. Let $G = \{(i,k): 1 \leq i \leq n_b,\ 1 \leq k \leq n_p\}$ be the complete set of prey/bait combinations. Then, the *in-frame* score is:

$$IF_{ik} = \frac{R_G(\rho_{ik})}{max(R_G(\rho_{ik}))}$$

where $R_G(\rho_{ik})$ represents the rank of the ρ_{ik} statistic formulated to test the increase of π_{ik} for prey *k* interacting with bait *i* among all prey/bait interactions. For prey with no fusion reads in either non-selected or selected conditions, IF_{ik} was set to zero.

Simulation of the Y2H-NGIS dataset

To test the performance of Y2H-SCORES under different conditions we developed a framework for Y2H-NGIS simulation, using empirical data to motivate the simulation model and parameter values. Fig 5 shows the experimental workflow we wish to simulate. We simulated both total and fusion read counts under selected and non-selected conditions.

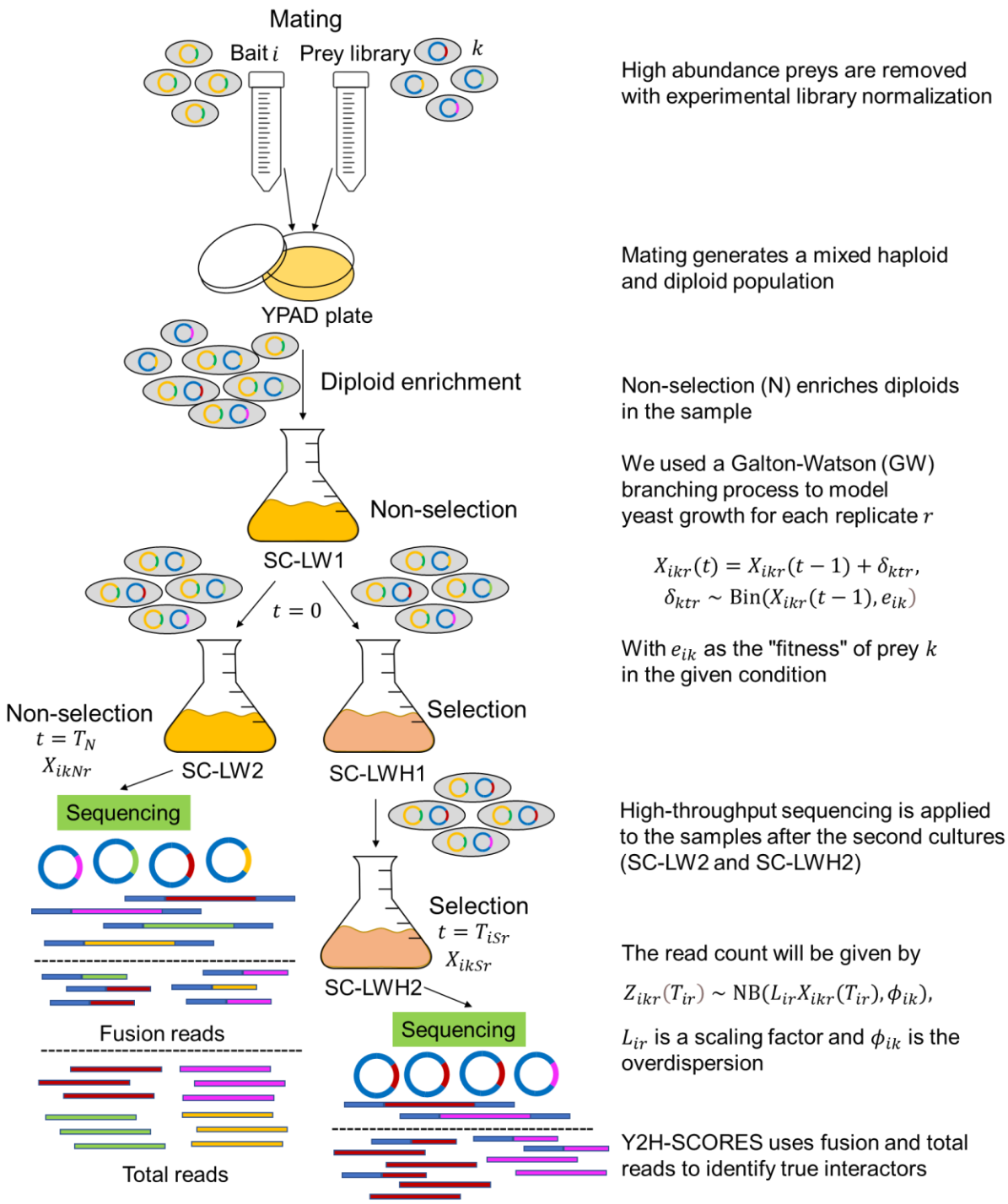

Fig 5. Experimental workflow for Y2H-NGIS. After the mating between bait and prey, diploids go through a non-selective culture to reach exponential phase. Once there ($t = 0$), the culture is split into two flasks, one for non-selection and another for selection. The objective of the second subculture is to grow yeast exponentially until it reaches saturation, a process that is repeated

twice under selective conditions. After T_N generations in the non-selected condition and T_{iSr} generations in the selected condition, culture aliquots are taken to be sequenced.

Model

We used a Galton-Watson (GW) branching process to model yeast growth in each condition $c \in \{S, N\}$. In this presentation of the model, we drop the index c from the notation for simplicity. The rth replicate culture in the presence of bait i starts with $M_{ir}(0) = M_0 = 3.84 \times 10^9$ total yeast, and is grown for a potentially random number of T_{ir} generations until the exponential growth phase ends. While the population size $M_{ir}(T_{ir})$ at the end of the experiment will be about 7.5×10^{10}, there is enough variation in this number that we do not consider it necessary to condition on its value.

Let $X_{ikr}(t)$ be the number of yeast containing prey k at generation t. We assume $X_{ikr}(t)$ follows a simple Galton-Watson branching process,

$$X_{ikr}(t) = X_{ikr}(t-1) + \delta_{ktr},$$

where $\delta_{ktr} \sim \mathrm{Bin}(X_{ikr}(t-1), e_{ik})$ and e_{ik} is the "fitness" of prey k in the given condition with bait i. We will generally assume each prey is experiencing differential growth rates e_{ik} because of selection, but the model also applies to non-selection conditions, where we assume all yeast grow at the same rate $e_{ik} = e_N$. The initial number of prey k is $X_{ikr}(0) = M_{ikr}$, with $M_{ikr} \sim Bin(M_0, q_{ik})$, and given the true proportion q_{ik} of prey k in the prey library.

At the end of the experiment (selection or non-selection), at generation T_{ir}, we do not observe $X_{ikr}(T_{ir})$ directly. Instead, we observe read counts $Z_{ikr}(T_{ir}) \sim NB(L_{ir}X_{ikr}(T_{ir}), \phi_{ik})$, from a Negative Binomial distribution with mean and variance:

$$\begin{aligned}\mathbb{E}[Z_{ikr}(T_{ir}) \mid X_{ikr}(T_{ir})] &= L_{ir}X_{ikr}(T_{ir}) \\ \mathrm{Var}[Z_{ikr}(T_{ir}) \mid X_{ikr}(T_{ir})] &= L_{ir}X_{ikr}(T_{ir}) + \phi_{ik}L_{ir}^2X_{ikr}^2(T_{ir}),\end{aligned}$$

where $L_{ir} \approx \frac{V_{ir}}{M_{ir}(T_{ir})}$ is a scaling factor (also called “size factor") accounting for sequencing depth V_{ir} and the population size $M_{ir}(T_{ir})$ at generation T_{ir}. Parameter $\phi_{ik} \geq 0$ is an overdispersion parameter that accounts for extra variation not already explained by the randomness in the initial prey count M_{ikr} and the branching process. We treat diploid enrichment and the second round of selection as deterministic in the model (Fig 5), but either may cause overdispersion relative to our stochastic growth model. Possible overdispersion is accommodated in with this NB observation model.

For full details about the modelling and parameter estimation please refer to the S2 text file.

Design of simulation scenarios and performance of scores

We designed different simulation scenarios to test the performance of the scores under different conditions. We took samples with 8000, 20000, and 40000 preys and 2 to 10 bait screenings. We used the following parameters to evaluate the performance scores as reported in S3 Table: 1) the stickiness of the samples, which we define as the percentage of sticky preys in the library; 2) the strength of the true interactors (defined as the minimum percentile of the observed fitness parameter $\hat{e}_{ik}$ calculated from real data and used to simulate the true interactor group); 3) overdispersion of true interactors under selection, given by the parameters φ_{kN}, φ_{ikS}, and sampled randomly or from a subset with high values (over the 90th percentile observed in real data); 4) the proportion of true interactors at the beginning of selection, given by q_{ik}; 5) the number of replicates; and 6) the number of baits. S4 Fig shows sampling distributions of some of these parameters as estimated from real data. Receiver Operating Characteristic (ROC) and Precision

Recall (PR) curves were constructed, and their respective Area Under the Curve (AUC) values calculated for comparison between simulations.

Experimental procedures

We generated experimental data to estimate parameters for the simulation, and to test the efficiency of Y2H-SCORES. Using an established Gateway-compatible CEN/ARS GAL4 system (Dreze et al., 2010; Yu et al., 2015), we created a normalized, three-frame cDNA expression library of 1.1 x 10^7 primary clones from pooled RNA isolated from a time-course experiment of barley, *Hordeum vulgare* L. (*Hv*) infected with the powdery mildew fungus, *Blumeria graminis* f. sp. *hordei* (*Bgh*) (Hunt et al., 2019; Surana, 2017). Baits were mated with a prey strain expressing the cDNA library and grown on selective media to identify protein-protein interactions. To initiate screening, mating of bait and prey cDNA library was performed on solid YPAD media. Diploids were enriched in SC-Leu-Trp (SC-LW) liquid media and sub-cultured under two conditions: 1) non-selected diploid growth (SC-LW) and 2) selected for reporter activation in SC-Leu-Trp-His (SC-LWH). Diploids expressing a positive PPI activate the HIS3 reporter construct and multiply in SC-LWH media whereas diploids expressing two non-interacting proteins are unable to grow under this selection. After sub-culturing the samples and reaching saturation (OD = 2.5 - 3), cells were collected, plasmids were isolated, and prey cDNA was amplified and sequenced using the Illumina HiSeq 2500 platform. We performed three independent biological replicates, collecting 5-10 million reads per sample.

Data from the *Bgh* effector protein AVR_{A13} (CSEP0372) and luciferase acting as baits were analyzed using the NGPINT pipeline. Outputs were taken to compute the Y2H-SCORES. We applied different normalization methods and calculated the Y2H-SCORES and their ensemble using Borda counts to obtain a ranked list of interactors. We compared the score values of the

top 5% of the ranked interactors with each normalization method, using a Wilcoxon ranked-sum test (S4 Table). Quantiles of the specificity score for the top 100 interactions ranked using median-of-ratios normalization, and unique or non-unique across other normalizations, were calculated to show the lower values of the non-unique list (S5 Table). Using the Y2H-SCORES calculated from library size normalization and the Borda ensemble, we predicted a top list of candidate interactors to be validated. The validation consisted in identifying candidate true interactors based on the list, determining the interacting prey fragments using the IGV alignments obtained from the NGPINT pipeline (Banerjee et al., 2020) and the *in-frame* prey transcripts with the highest *in-frame* score. After the determination of the exact fragment, primers were designed for Gateway cloning, and subsequent insertion into the prey vector. After cloning the candidate prey into yeast, we concluded the validation with a binary Y2H test in a series of media and controls: 1) Diploid selection (SC-LW), interaction selection (SC-LWH) and stringent selection (SC-LWH+3AT) as shown in Fig 4.

Determination of interlogs

Predicted protein-protein interactors of HvTHF1 were inferred using interlogs (Jafari et al., 2015; Nakajima et al., 2018). Orthologs of *HvThf1* with *Arabidopsis thaliana, Zea mays* and *Oryza sativa* were obtained using the Plant Compara tables from Ensembl Plants (Howe et al., 2020). Experimentally validated interactions for these plants were mined from BioGRID 3.5.171 version (Stark, 2005), the Protein-Protein Interaction database for Maize (PPIM) (Zhu et al., 2016) and the Predicted Rice Interactome Network (PRIN) database (Gu et al., 2011). Barley interlogs were inferred by assigning the mined interactions from the corresponding orthologs, with *Oryza sativa* as the species with experimental interactions reported. Visualization of the network was done using Cytoscape (Shannon et al., 2003).

Abbreviations

AUC: Area Under the Curve

Bgh: *Blumeria graminis* f. sp. *hordei*

CV: Coefficient of Variation

DE: Differentially Enriched

eQTL: expression Quantitative Trait Loci

Hv: *Hordeum vulgare*

Mla1: powdery mildew resistance locus a1

MlLa: Laevigatum resistance locus

NGPINT: **N**ext-**g**eneration **p**rotein-protein **int**eraction software

NB: Negative Binomial

PCA: Principal Component Analysis

PPI: Protein-Protein Interaction

PR: Precision Recall

ROC: Receiver Operating Characteristic

ROS: Reactive oxygen species

RUV: Remove Unwanted Variation

TPM: Transcripts Per Million

Y2H: Yeast Two-Hybrid

Y2H-NGIS: Yeast Two-Hybrid Next-Generation Interaction Screening

Availability of code, data, and materials

R code and ReadMe file for the Y2H-SCORES software are provided at GitHub (https://github.com/Wiselab2/Y2H-SCORES). Additionally, we implemented a python script to link the score functions with the NGPINT pipeline (Banergee et al., 2020), which integrates both software packages by allowing Y2H-SCORES to run on the NGPINT outputs. Users can find the instructions in the same repository. Y2H-NGIS data supporting the conclusions of this article will be available in NCBI's Gene Expression Omnibus (GEO) upon publication.

Supplemental files

S1 Fig. Coefficient of variation (CV) for each prey using different normalization methods.

S2 Fig. Top candidate interactors inferred with Y2H-SCORES and different normalization methods.

S3 Fig. Binary Y2H for the candidate preys HORVU2Hr1G060120 and HORVU2Hr1G024160.

S4 Fig. Distributions of the parameters used for Y2H-NGIS simulation.

S1 Table. Wilcoxon signed-rank test for the Pearson correlation of samples grouped by condition, for different normalization methods.

S2 Table. Wilcoxon signed-rank test for CV densities of different normalization methods.

S3 Table. Y2H-NGIS simulation scenarios and performance of Y2H-SCORES.

S4 Table. Wilcoxon ranked-sum test for the top 5% ranked interactions using the Borda ensemble of the Y2H-SCORES under different normalization methods.

S5 Table. Quantiles of the specificity score for the top 100 interactions ranked using median-of-ratios normalization and unique or non-unique across other normalizations.

S1 Text. *HvThf1* prey sequences.

S2 Text. Y2H-NGIS simulation model.

Acknowledgments

The authors thank Greg Fuerst for conducting the time-course infection experiment and for expert isolation of RNA for the 3-frame Y2H library. Research supported in part by Fulbright - Minciencias 2015 & Schlumberger Faculty for the Future fellowships to VVZ, USDA-NIFA-ELI Postdoctoral Fellowship 2017-67012-26086 to JME, Oak Ridge Institute for Science and Education (ORISE) under U.S. Department of Energy (DOE) contract number DE-SC0014664 to SB, USDA-National Institute of Food and Agriculture (NIFA) Hatch project IOW03617 to KSD, and National Science Foundation - Plant Genome Research Program grant 13-39348 and USDA-Agricultural Research Service project 3625-21000-067-00D to RPW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA, NIFA, ARS, DOE, ORAU/ORISE or the National Science Foundation. USDA is an equal opportunity provider and employer.

References

Adams, L.G., Khare, S., Lawhon, S.D., Rossetti, C.A., Lewin, H.A., Lipton, M.S., Turse, J.E., Wylie, D.C., Bai, Y., Drake, K.L., 2011. Enhancing the role of veterinary vaccines reducing zoonotic diseases of humans: Linking systems biology with vaccine development. Vaccine 29, 7197–7206. https://doi.org/10.1016/j.vaccine.2011.05.080

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11. https://doi.org/10.1074/jbc.272.7.4310

Arkin, M.R., Tang, Y., Wells, J.A., 2014. Small-molecule inhibitors of protein-protein interactions: Progressing toward the reality. Chem. Biol. 21, 1102–1114. https://doi.org/10.1016/j.chembiol.2014.09.001

Braun, P., Aubourg, S., Van Leene, J., De Jaeger, G., Lurin, C., 2013. Plant protein interactomes. Annu. Rev. Plant Biol. 64, 161–187. https://doi.org/10.1146/annurev-arplant-050312-120140

Dangl, J.L., Horvath, D.M., Staskawicz, B.J., 2013. Pivoting the plant immune system from dissection to deployment. Science (80-.). 341, 746–751. https://doi.org/10.1126/science.1236011

De Las Rivas, J., Fontanillo, C., 2010. Protein–protein interactions issentials: Key concepts to building and analyzing interactome networks. PLOS Comput. Biol. 6, e1000807.

Dillies, M.A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, N.S., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloë, D., Le Gall, C., Schaëffer, B., Le Crom, S., Guedj, M., Jaffrézic, F., 2013. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Brief. Bioinform. 14, 671–683. https://doi.org/10.1093/bib/bbs046

Dreze, M., Monachello, D., Lurin, C., Cusick, M.E., Hill, D.E., Vidal, M., Braun, P., 2010. High-quality binary interactome mapping. Methods Enzymol. https://doi.org/10.1016/S0076-6879(10)70012-4

Erffelinck, M.L., Ribeiro, B., Perassolo, M., Pauwels, L., Pollier, J., Storme, V., Goossens, A., 2018. A user-friendly platform for yeast two-hybrid library screening using next generation sequencing. PLoS One 13, 1–21. https://doi.org/10.1371/journal.pone.0201270

Evans, C., Hardin, J., Stoebel, D.M., 2018. Selecting between-sample RNA-Seq normalization

methods from the perspective of their assumptions. Brief. Bioinform. 19, 776–792. https://doi.org/10.1093/bib/bbx008

Gu, H., Zhu, P., Jiao, Y., Meng, Y., Chen, M., 2011. PRIN: A predicted rice interactome network. BMC Bioinformatics 12. https://doi.org/10.1186/1471-2105-12-161

Hamel, L.P., Sekine, K.T., Wallon, T., Sugiwaka, Y., Kobayashi, K., Moffett, P., 2016. The chloroplastic protein THF1 interacts with the Coiled-Coil domain of the disease resistance protein N´ and regulates light-dependent cell death. Plant Physiol. 171, 658–674. https://doi.org/10.1104/pp.16.00234

Howe, K.L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D.M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., Gil, L., Gupta, P., Hammond-Kosack, K.E., Haskell, E., Hunt, S.E., Jaiswal, P., Janacek, S.H., Kersey, P.J., Langridge, N., Maheswari, U., Maurel, T., McDowall, M.D., Moore, B., Muffato, M., Naamati, G., Naithani, S., Olson, A., Papatheodorou, I., Patricio, M., Paulini, M., Pedro, H., Perry, E., Preece, J., Rosello, M., Russell, M., Sitnik, V., Staines, D.M., Stein, J., Tello-Ruiz, M.K., Trevanion, S.J., Urban, M., Wei, S., Ware, D., Williams, G., Yates, A.D., Flicek, P., 2020. Ensembl Genomes 2020-enabling non-vertebrate genomic research. Nucleic Acids Res. 48, D689–D695. https://doi.org/10.1093/nar/gkz890

Hunt, M., Banerjee, S., Surana, P., Liu, M., Fuerst, G., Mathioni, S., Meyers, B.C., Nettleton, D., Wise, R.P., 2019. Small RNA discovery in the interaction between barley and the powdery mildew pathogen. BMC Genomics 20, 1–21. https://doi.org/10.1186/s12864-019-5947-z

Jafari, M., Mirzaie, M., Sadeghi, M., 2015. Interlog protein network: An evolutionary benchmark of protein interaction networks for the evaluation of clustering algorithms. BMC Bioinformatics 16. https://doi.org/10.1186/s12859-015-0755-1

Kessens, R., Sorensen, N., Kabbage, M., 2018. An inhibitor of apoptosis (SfIAP) interacts with SQUAMOSA promoter-binding protein (SBP) transcription factors that exhibit pro-cell death characteristics. Plant Direct 2, 1–17. https://doi.org/10.1002/pld3.81

Koff, W.C., Burton, D.R., Johnson, P.R., Walker, B.D., King, C.R., Nabel, G.J., Ahmed, R., Bhan, M.K., Plotkin, S.A., 2013. Accelerating next-generation vaccine development for global disease prevention. Science (80-.). 340. https://doi.org/10.1126/science.1232910

Krishnamani, V., Stamnes, M.A., Piper, R.C., 2019. MALTA: A calculator for estimating the coverage with shRNA, CRISPR, and cDNA libraries. SoftwareX 9, 154–160. https://doi.org/10.1016/j.softx.2019.01.006

Lewis, J.D., Wan, J., Ford, R., Gong, Y., Fung, P., Nahal, H., Wang, P.W., Desveaux, D., 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. https://doi.org/10.1186/1471-2164-13-8

Lin, S., 2010. Rank aggregation methods. Wiley Interdiscip. Rev. Comput. Stat. 2, 555–570. https://doi.org/10.1002/wics.111

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, X., Kracher, B., Saur, I.M.L., Bauer, S., Ellwood, S.R., Wise, R., Yaeno, T., Maekawa, T., Schulze-Lefert, P., 2016. Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. Proc. Natl. Acad. Sci. U. S. A. 113, E6486–E6495. https://doi.org/10.1073/pnas.1612947113

Manning, V.A., Hardison, L.K., Ciuffetti, L.M., 2007. Ptr ToxA interacts with a chloroplast-

localized protein. Mol. Plant-Microbe Interact. 20, 168–177. https://doi.org/10.1094/MPMI-20-2-0168

Nakajima, N., Hayashida, M., Jansson, J., Maruyama, O., Akutsu, T., 2018. Determining the minimum number of protein-protein interactions required to support known protein complexes. PLoS One 13, 1–17. https://doi.org/10.1371/journal.pone.0195545

Oberg, A.L., Kennedy, R.B., Li, P., Inna G. Ovsyannikova, G.A.P., 2011. Systems Biology Approaches to New Vaccine Development. Curr. Opin. Immunol. 23, 436–443. https://doi.org/10.1038/jid.2014.371

Pandelova, I., Figueroa, M., Wilhelm, L.J., Manning, V.A., Mankaney, A.N., Mockler, T.C., Ciuffetti, L.M., 2012. Host-selective toxins of pyrenophora tritici-repentis induce common responses associated with host susceptibility. PLoS One 7, 13–20. https://doi.org/10.1371/journal.pone.0040240

Pashkova, N., Peterson, T.A., Krishnamani, V., Breheny, P., Stamnes, M., Piper, R.C., 2016. DEEPN as an approach for batch processing of yeast 2-hybrid interactions. Cell Rep. 17, 303–315. https://doi.org/10.1016/j.celrep.2016.08.095

Rajaraman, J., Douchkov, D., Lück, S., Hensel, G., Nowara, D., Pogoda, M., Rutten, T., Meitzel, T., Brassac, J., Höfle, C., Hückelhoven, R., Klinkenberg, J., Trujillo, M., Bauer, E., Schmutzer, T., Himmelbach, A., Mascher, M., Lazzari, B., Stein, N., Kumlehn, J., Schweizer, P., 2018. Evolutionarily conserved partial gene duplication in the Triticeae tribe of grasses confers pathogen resistance. Genome Biol. 19, 1–18. https://doi.org/10.1186/s13059-018-1472-7

Reece-Hoyes, J.S., Marian Walhout, A.J., 2012. Yeast one-hybrid assays: A historical and technical perspective. Methods 57, 441–447. https://doi.org/10.1016/j.ymeth.2012.07.027

Risso, D., Ngai, J., Speed, T.P., Dudoit, S., 2014. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat. Biotechnol. 32, 896–902. https://doi.org/10.1038/nbt.2931

Schurch, N.J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G.G., Owen-Hughes, T., Blaxter, M., Barton, G.J., 2016. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? Rna 22, 839–851. https://doi.org/10.1261/rna.053959.115

Shannon, P., Markiel, A., Owen Ozier, 2, Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite

Stark, C., 2005. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 34, D535–D539. https://doi.org/10.1093/nar/gkj109

Struk, S., Jacobs, A., Sánchez Martín-Fontecha, E., Gevaert, K., Cubas, P., Goormachtig, S., 2019. Exploring the protein–protein interaction landscape in plants. Plant Cell Environ. 42, 387–409. https://doi.org/10.1111/pce.13433

Surana, P., 2017. Membrane trafficking in resistance gene-mediated defense against the barley powdery mildew fungus. Iowa State University.

Surana, P., Xu, R., Fuerst, G., Chapman, A.V.E., Nettleton, D., Wise, R.P., 2017. Interchromosomal transfer of immune regulation during infection of barley with the powdery mildew pathogen. G3 Genes, Genomes, Genet. 7, 3317–3329. https://doi.org/10.1534/g3.117.300125

Suter, B., Zhang, X., Gustavo Pesce, C., Mendelsohn, A.R., Dinesh-Kumar, S.P., Mao, J.H.,

2015. Next-generation sequencing for binary protein-protein interactions. Front. Genet. 6, 1–6. https://doi.org/10.3389/fgene.2015.00346

Trigg, S.A., Garza, R.M., MacWilliams, A., Nery, J.R., Bartlett, A., Castanon, R., Goubil, A., Feeney, J., O'Malley, R., Huang, S.S.C., Zhang, Z.Z., Galli, M., Ecker, J.R., 2017. CrY2H-seq: A massively multiplexed assay for deep-coverage interactome mapping. Nat. Methods 14, 819–825. https://doi.org/10.1038/nmeth.4343

Vidal, M., Fields, S., 2014. The yeast two-hybrid assay: Still finding connections after 25 years. Nat. Methods 11, 1203–1206. https://doi.org/10.1038/nmeth.3182

Wagner, G.., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285.

Weimann, M., Grossmann, A., Woodsmith, J., Özkan, Z., Birth, P., Meierhofer, D., Benlasfer, N., Valovka, T., Timmermann, B., Wanker, E.E., Sauer, S., Stelzl, U., 2013. A Y2H-seq approach defines the human protein methyltransferase interactome. Nat. Methods 10, 339–342. https://doi.org/10.1038/nmeth.2397

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., Yi, S., Tyagi, T., Sheykhkarimli, D., Roth, J.F., Wong, C., Musa, L., Snider, J., Liu, Y., Yu, H., Braun, P., Stagljar, I., Hao, T., Calderwood, M.A., Pelletier, L., Aloy, P., Hill, D.E., Vidal, M., Roth, F.P., 2016. Pooled-matrix protein interaction screens using Barcode Fusion Genetics. Mol. Syst. Biol. 12, 863. https://doi.org/10.15252/msb.20156660

Yu, C.P., Chen, S.C.C., Chang, Y.M., Liu, W.Y., Lin, H.H., Lin, J.J., Chen, H.J., Lu, Y.J., Wu, Y.H., Lu, M.Y.J., Lu, C.H., Shih, A.C.C., Ku, M.S. Ben, Shiu, S.H., Wu, S.H., Li, W.H., 2015. Transcriptome dynamics of developing maize leaves and genomewide prediction of cis elements and their cognate transcription factors. Proc. Natl. Acad. Sci. U. S. A. 112,

E2477–E2486. https://doi.org/10.1073/pnas.1500605112

Zhu, G., Wu, A., Xu, X.-J., Xiao, P.-P., Lu, L., Liu, J., Cao, Y., Chen, L., Wu, J., Zhao, X.-M., 2016. PPIM: A Protein-Protein Interaction database for Maize. Plant Physiol. 170, 618–626. https://doi.org/10.1104/pp.15.01821

Zong, T., Yin, J., Jin, T., Wang, L., Luo, M., Li, K., Zhi, H., 2020. A DnaJ protein that interacts with soybean mosaic virus coat protein serves as a key susceptibility factor for viral infection. Virus Res. 281, 197870. https://doi.org/10.1016/j.virusres.2020.197870