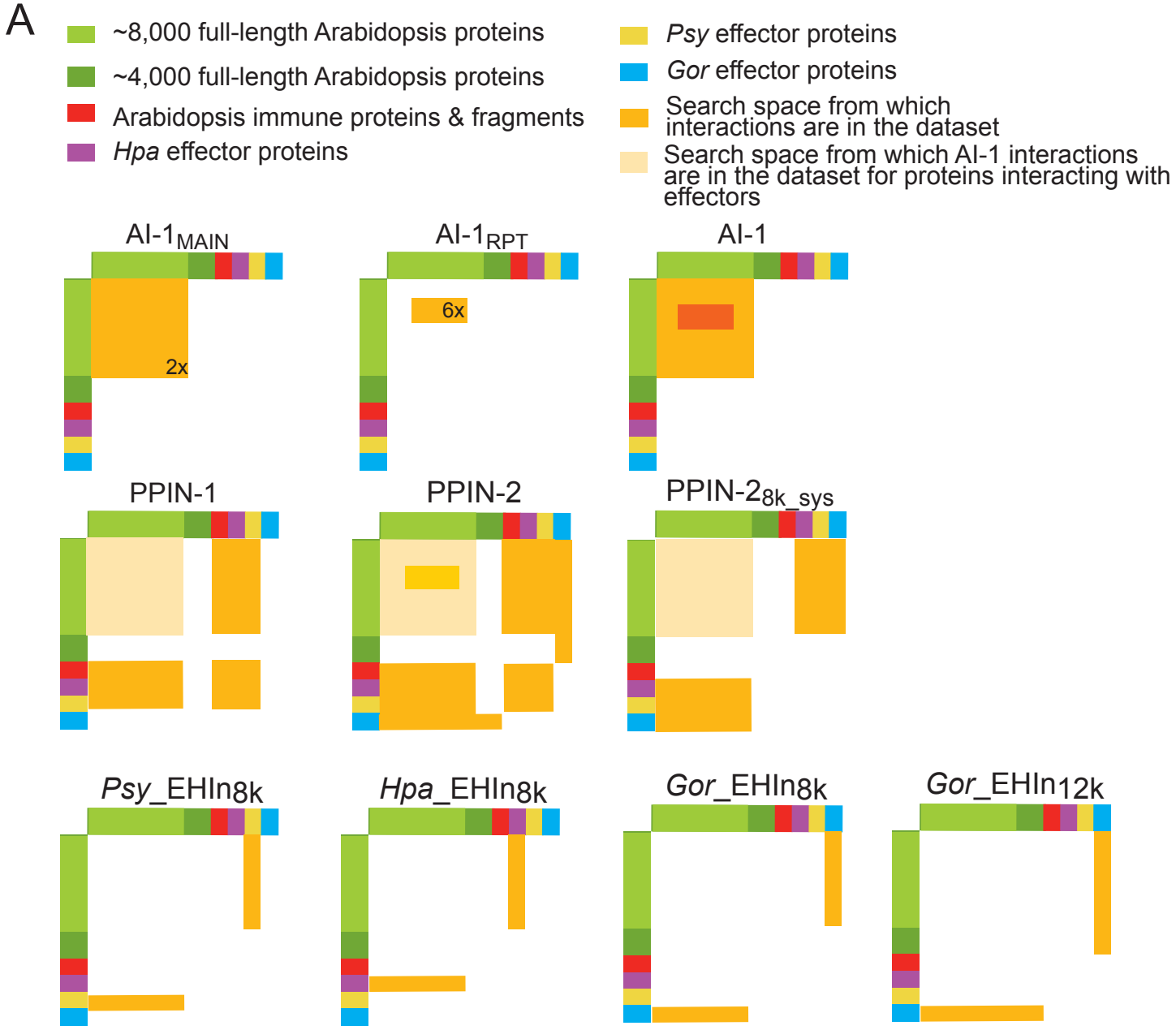
Figure S1, related to Figure 1

A



B

		Psy	*Hpa*	*Gor*	nr_Sum
8k Space	Effectors_8k	30	52	41	123
	Host-Interactors_8k	56	118	45	178
	Host-Interactors$_{AI\text{-}1MAIN}$	47	107	42	155
	Effector-host protein Interactions_8k	99	230	93	422
12k Space	Effectors_12k	-	-	46	46
	Interactors_12k	-	-	60	60
	Effector-Host-protein Interactions_12k	-	-	122	122
PPIN-1	Effectors	30	53	-	83
	Interactors_8K+Imm	61	122	-	165
	Effector-host-protein Interactions_8k+Imm	106	234	-	340

Figure 1C, related to Figure 1

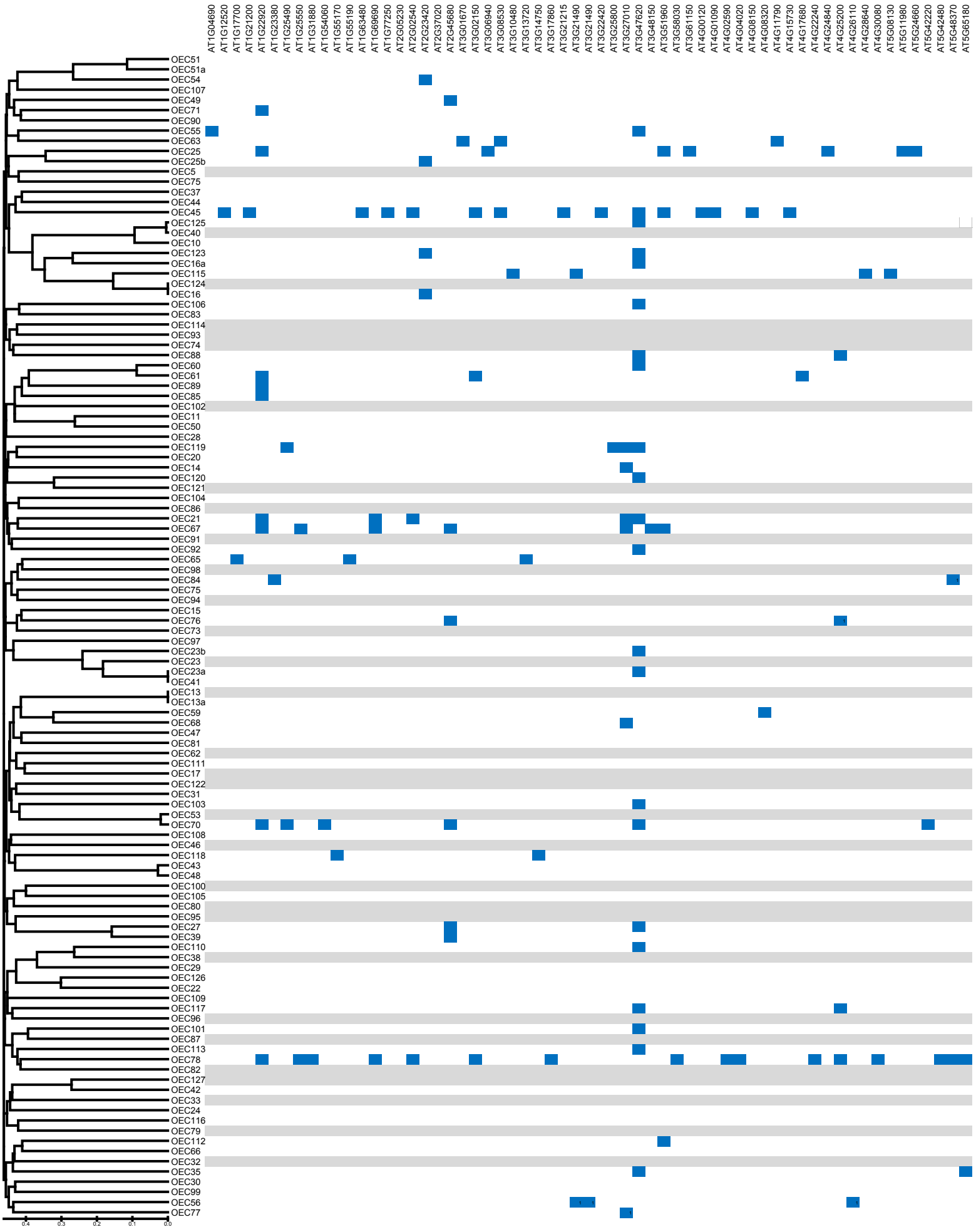
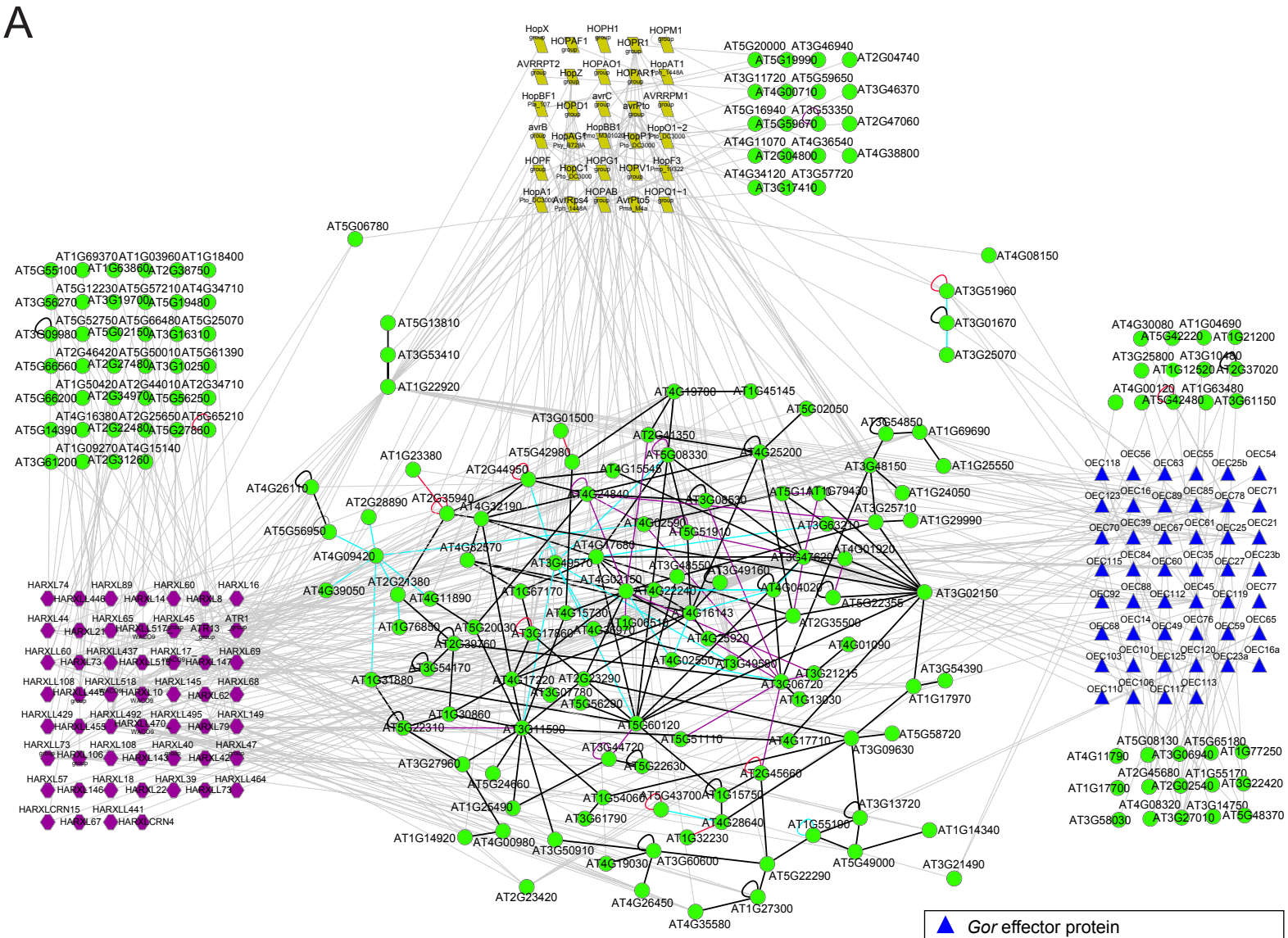


Figure S2, related to Figure 2



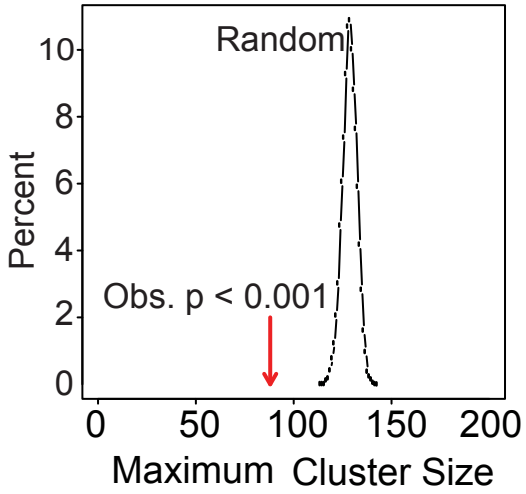
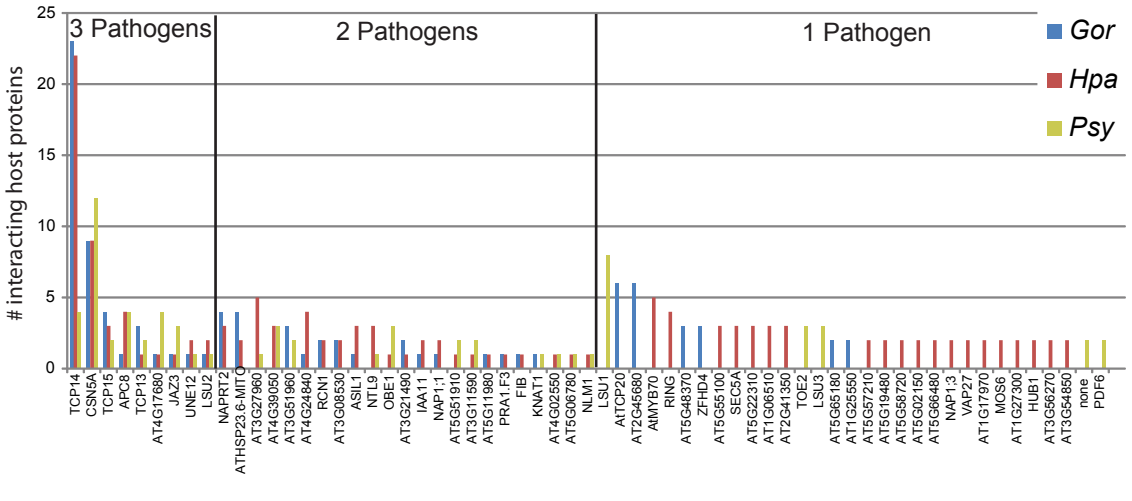
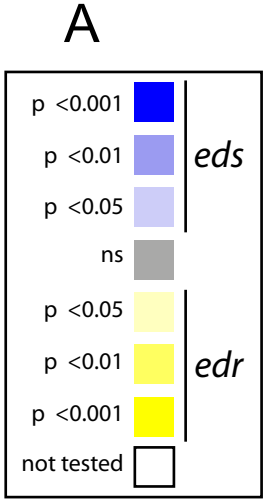


Figure S3, related to Figure 3





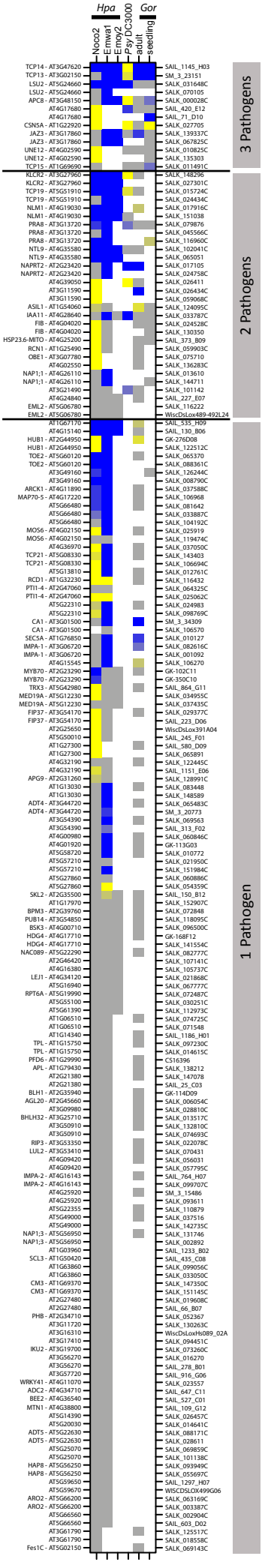


Figure S4, related to Figure 4

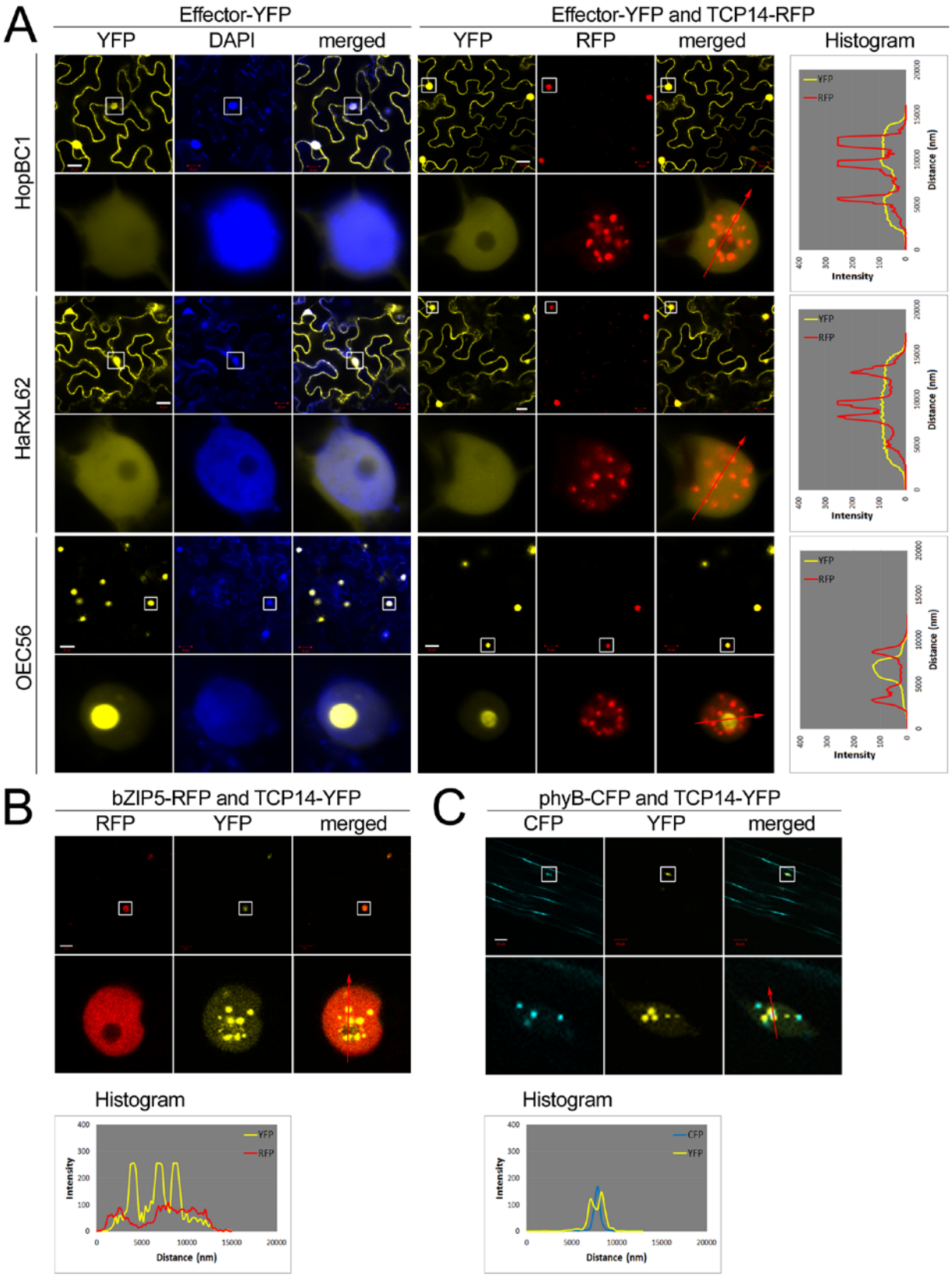


Figure S4, related to Figure 4

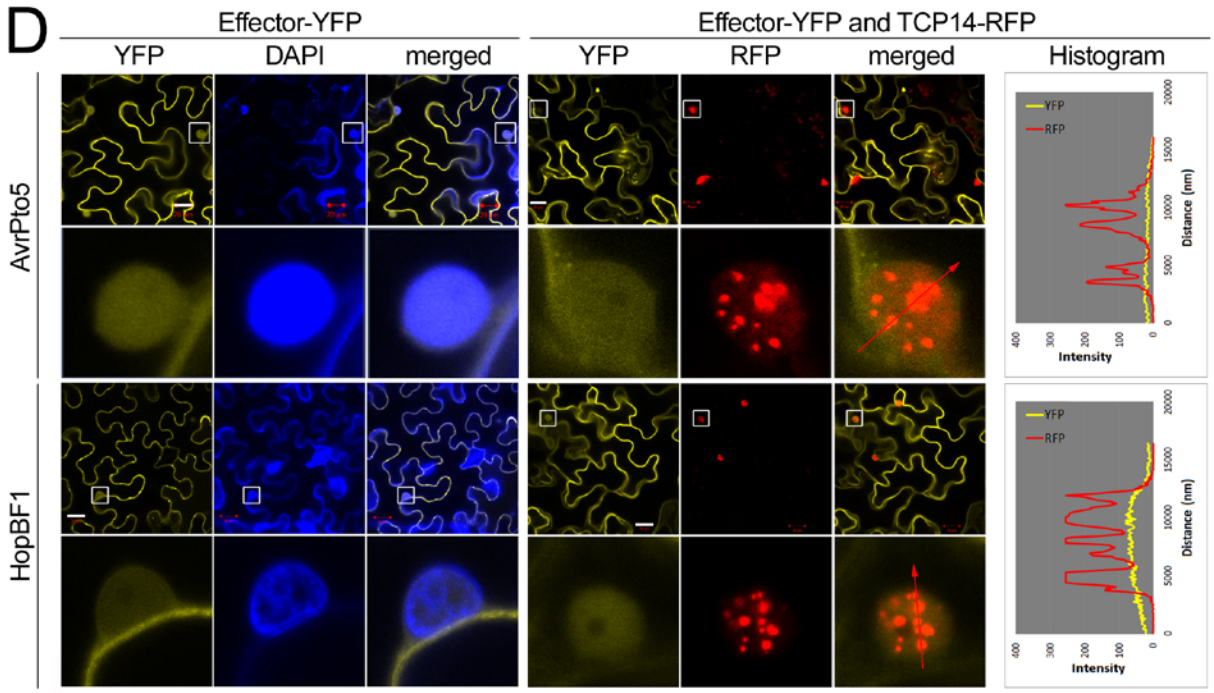


Figure S4, related to Figure 4

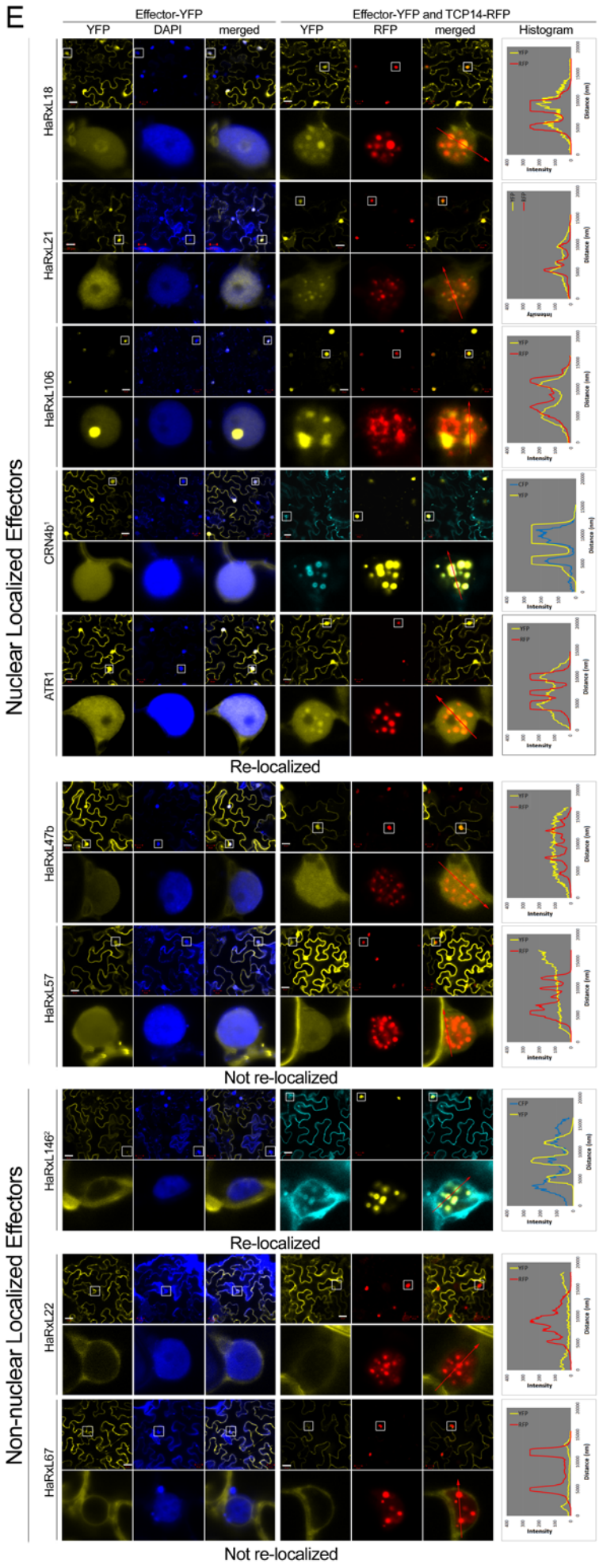


Figure S4, related to Figure 4

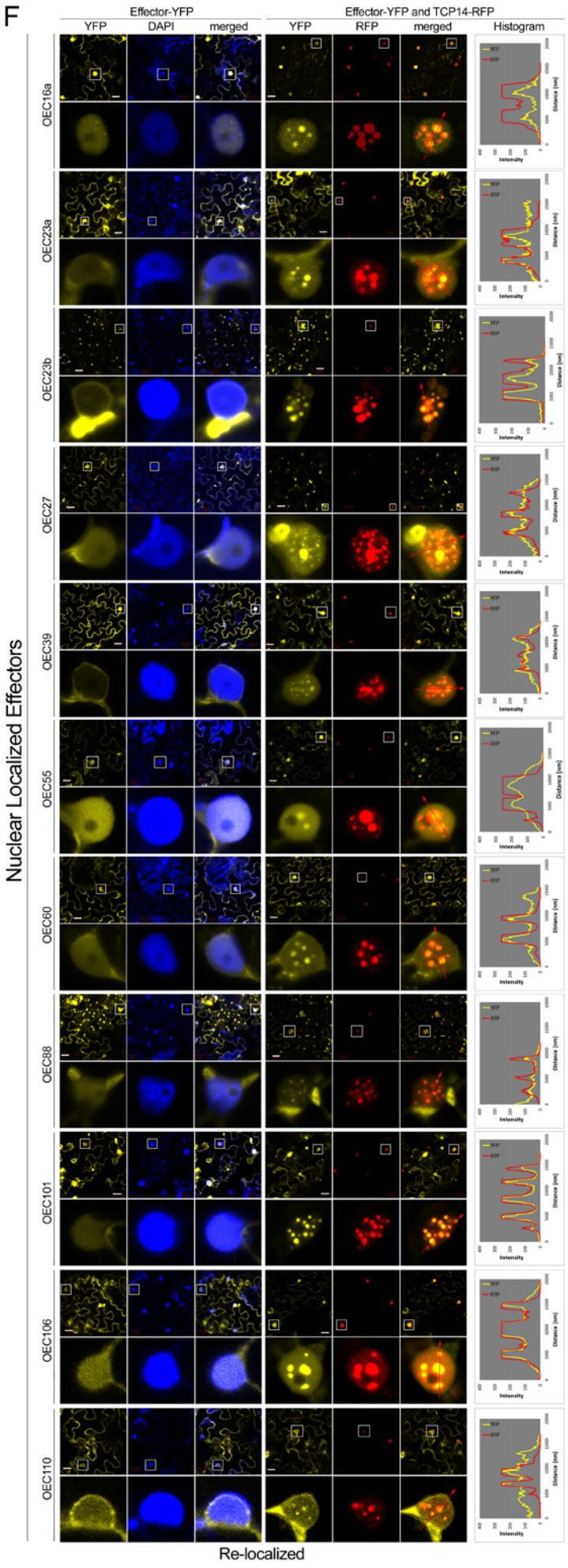


Figure S4, related to Figure 4

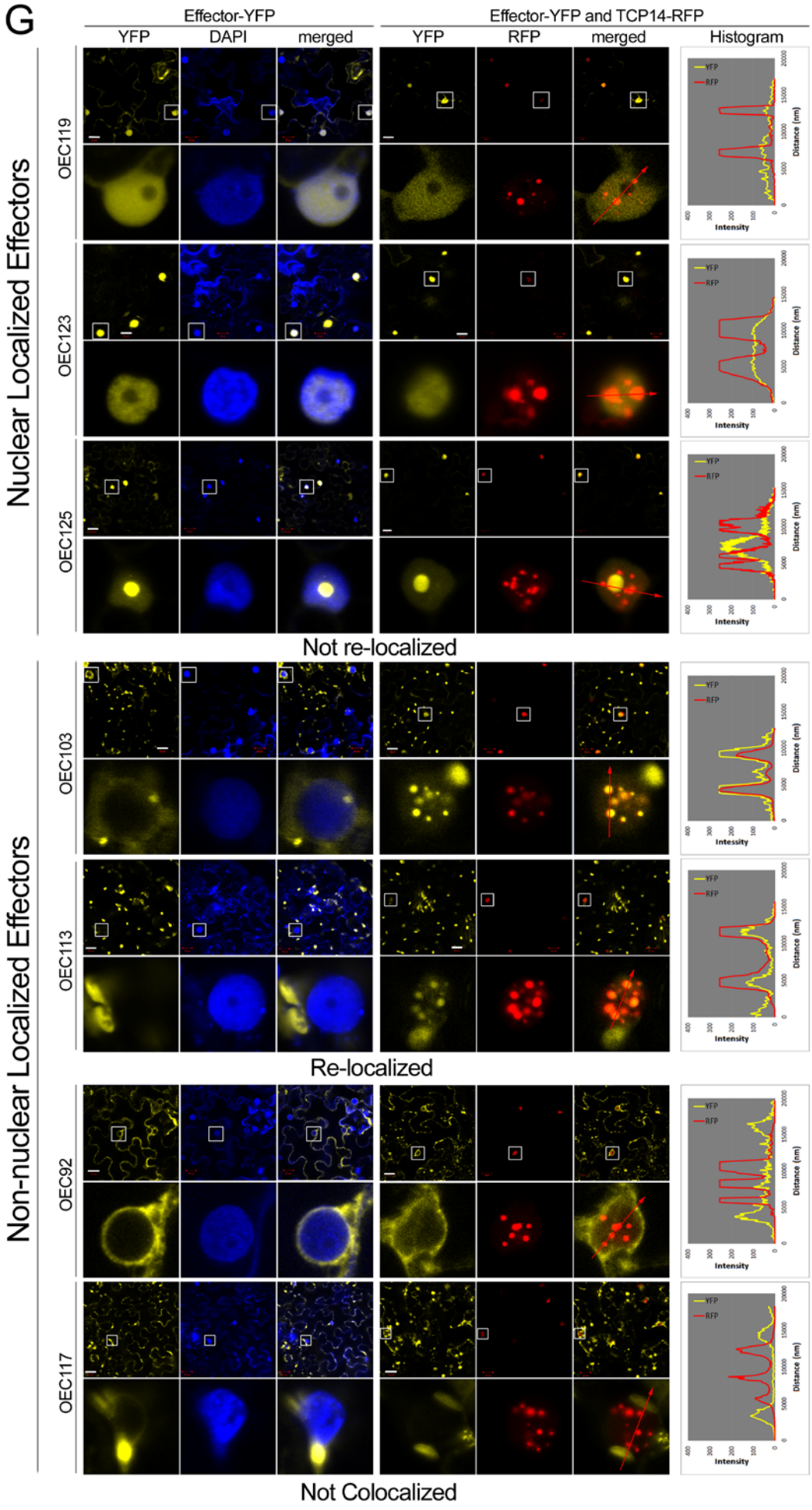


Figure S4, related to Figure 4

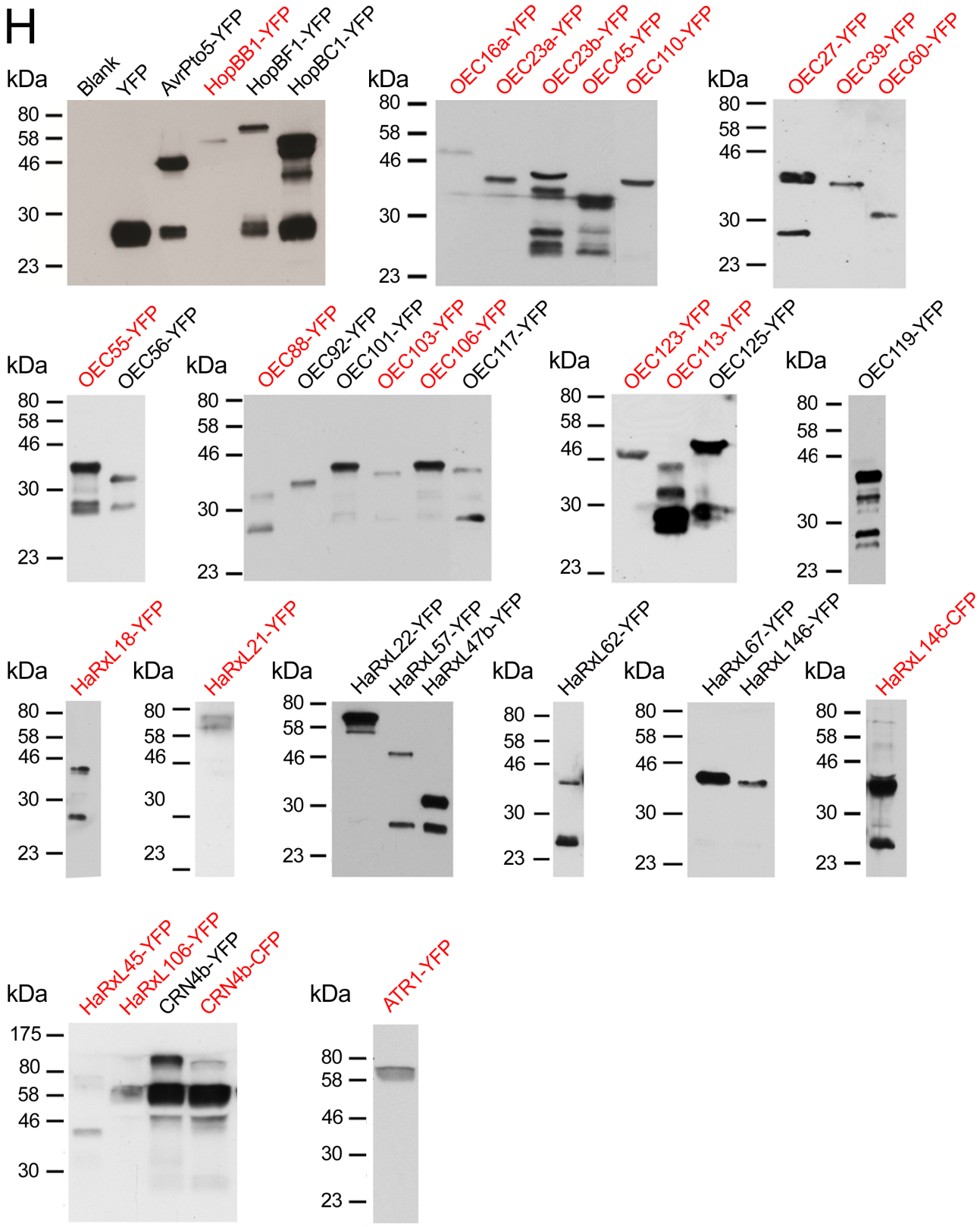
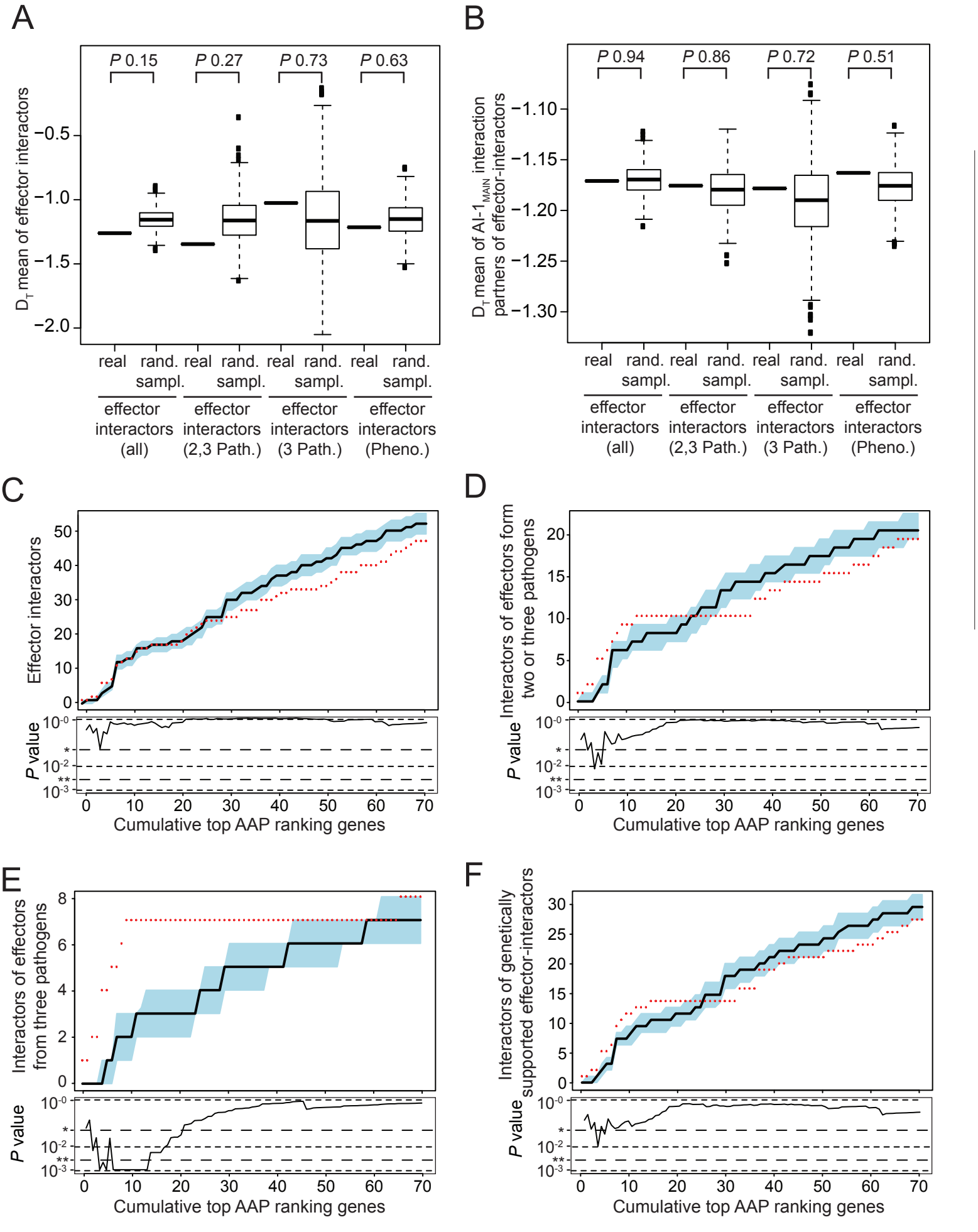


Figure S5, related to Figure 5



Legends to Supplementary Figures

Figure S1: Datasets and subsets used in the manuscript, related to Figure 1.
A. Graphic representation of the datasets. Names are provided on top of each dataset icon, explanation of the different ORFs used to map the respective datasets are in the legend on top and in the main text.
B. Tabulated summary of different interactions subsets and therein contained effectors, interactions, and interacting host proteins for each pathogen in the 8k, 12k and PPIN-1 search space. 12k and PPIN-1 include the 8k space, respectively. The 8k space interactors that are only in the search space versus those that are also part of the AI-1_{MAIN} network are listed separately.
C. Interactions of OECs, sorted according to phylogenetic relationship, with Arabidopsis host proteins are indicated by blue squares. Grey bars indicate OECs that were not cloned.

Figure S2: Complete plant pathogen interactome network 2 (PPIN-2), related to Figure 2
A. Complete PPIN-2 network contains interactions among host proteins found in PPIN-1, and AI-1_{RPT} and literature curated interactions (LCI).
B. Random network rewiring simulation shows that the effector interacting proteins are less connected than in degree-preserved randomly rewired networks. Shown is the size distribution of the largest connected component formed by the effector-interactors in 10,000 degree-preserved randomly rewired AI-1_{MAIN} networks compared to the observed value (red arrow).

Figure S3: Complete phenotyping data, related to Figure 3.
A. Complete heat-map displaying phenotyping results for all tested T-DNA insertion lines. Locus ID and gene symbol are indicated to the left, IDs of the insertion lines to the right.
B. Comparison of intra- and interspecies convergence. For each Arabidopsis protein interacting with at least two effectors the effector-degree is displayed in a color coded manner as a bar-graph. Arabidopsis proteins only interacting with a single effector are not shown.

Figure S4: TCP14 re-localizes effectors to sub-nuclear foci, related to Figure 4.
A-C. Controls demonstrating that TCP14 did *not* co-localize with controls with which it did not interact in Y2H. These included (**A**) three representative effectors; (**B**) the bZIP5 TF; and (**C**) the unrelated, sub-nuclear body localized PhyB protein. Note that the data in **C** represents images of hypocotyls of two-week-old Arabidopsis seedlings hemizygous for 35S:phyB-CFP and EST:TCP14-YFP. Seedlings were grown under white light. To induce the expression of TCP14-YFP, 20µM of estradiol was applied 24 hours before imaging.
D. TCP14 does not re-localize *Psy* effectors AvrPto5 and HopBF1.
E. TCP14 re-localizes five additional interacting *Hpa* effectors. HaRXL146 and CRN4b did not co-localize with TCP14-RFP as YFP-tagged fusions, but were re-localized as CFP-tagged fusions by TCP14-YFP.

F,G. TCP14 re-localizes 13 additional interacting *Gor* effectors. All confocal pictures were taken 40-48 hours after infiltration of Agrobacterium strains expressing the different xFP–tagged proteins in leaves of 5-6 week old *N. benthamiana* plants.
H. Western blots of effector fusions used in transient expression assay. Red denotes effectors re-localized to TCP14. Expected molecular masses are given in **Table S4**.

Figure S5: Evolutionary parameters of effectors interactions and high AAP proteins, related to Figure 5
A. For none of the four groups of effector targets is the observed mean D_T significantly different from random expectation. Shown are boxplots of the D_T mean of observed effector-interactors (real) compared to the distribution of means observed in 1,000 random samplings from AI-1_{MAIN} of the same size.
B. For the AI-1_{MAIN} interactors of none of the four groups of effector targets is the observed mean D_T significantly different from random expectation. The boxplots show the D_T mean of observed AI-1_{MAIN} interactors of effector-interactors (real) compared to the distribution of means observed in 1,000 random networks obtained by degree preserving random rewiring. Boxes bracket the 25^{th} and 75^{th} percentile; whiskers indicate the 1.5-fold interquartile distances; dots represent remaining outliers.
C. Real versus randomly observed interacting effector-interactors of top AAP-ranking proteins. Plotted on Y-axis are cumulative counts of effector-interactors interacting with proteins encoded by the top AAP-ranking x genes. Data from AI-1_{MAIN} are shown as red dots, the black line shows the median of 1,000 randomly rewired networks, grey shaded areas show the 25^{th} and 75^{th} percentiles of values found in the rewiring controls. The lower panel provides for each data point the experimentally determined *P* value (* 0.05; ** 0.005). Boxplots are laid out as in B.
D. Analysis as in B, but counting proteins interacting with effectors from two or three pathogens.
E. Analysis as in B, but counting proteins interacting with effectors from three pathogens.
F. Analysis as in B, but counting proteins interacting with effectors and whose genetic deletion caused an immune phenotype.

Glossary

AAP	Amino Acid Polymorphism – all nucleotide polymorphism that result in a given amino acid change at a specific position in the protein.
AD	GAL4 Activation Domain in the Y2H system
AI-1	Arabidopsis interactome 1, a large-scale interaction network map consisting of AI-1_{MAIN} and AI-1_{RPT}, previously published in Science 2011 (**Figure S1**).
AI-1_{MAIN}	Systematic dataset of the Arabidopsis interactome, obtained by screening 8000 ORFs of the 8k_space systematically against each other twice (**Figure S1**).
AI-1_{RPT}	A dataset obtained by screening a subset of 8k_space, consisting of 1000 x 2000 Proteins against each other 6 times (**Figure S1**).
degree	Number of interaction partners
Effector-degree	Number of virulence effectors interacting with a specific Arabidopsis protein.
EHIn (*Gor*_EHIn, *Hpa*_EHIn, *Psy*_EHIn	Effector Host Interactome – datasets describing interactions between effectors from the investigated (*Psy*, *Hpa*, *Gor*) pathogens with host proteins. Interactions within 8k_space and 12k_space (*Gor* only) and indicated by a respective subscript (**Figure S1**).
DB	GAL4 DNA Binding Domain in the Y2H system
D_T	Tajima's D
edge	Network term for connections between nodes, here: "interactions"
edr	Enhanced disease resistance
eds	Enhanced disease susceptibility
Gor	*Golovinomyces orontii*
Hpa	*Hyaloperonospora arabidopsidis*
node	Network analysis term for connected entities, here proteins
OEC	*G. orontii* effector candidates
ORF	Open Reading Frame
PPIN-1	Plant-Pathogen Interactome Network-1 obtained by screening of Hpa and Psy effectors twice against proteins in the 8k_space and against a selection of immune proteins as described in Mukthar *et al.*, Science, 2011 (**Figure S1**).
Psy	*Pseudomonas syringae*
θ_W	Watterson's estimator θ for the scaled mutation rate
8k_space	8,000 Arabidopsis proteins used to generate AI-1 and PPIN-1
12k_space	12,000 Arabidopsis ORFs, including all of the 8k_space

Supplemental Experimental Procedures

1. Yeast-2-hybrid interactome mapping of OECs

A detailed protocol of the Y2H pipeline used is presented in (Dreze et al., 2010). Briefly, the 84 cloned OECs were transferred into pDest-AD and pDest-DB vectors by Gateway recombination. Successful ORF transfer was confirmed by PCR analysis. Isolated destination clones were transferred into *S. cerevisiae* Y8930 (for DB clones; MATα) and *S. cerevisiae* Y8890 (for AD clones; MATa) by Lithium-Acetate based transformation. Transgenic clones were selected on selective medium and stored in 20% glycerol at -80°C before use. For autoactivator removal, DB- and AD-OEC clones were mated with yeast clones containing an empty bait or prey vector on YEPD medium. After o/n incubation, colonies were transferred to selective media for diploid yeast (Sc–Leu–Trp) and incubated o/n. Then, diploid colonies were transferred to interaction media (Sc–Leu–Trp–His + 3-amino-1,2,4-triazole (3-AT), incubated o/n and replica-cleaned (excess yeasts were removed by pushing plates onto a fine velvet on a replica plating block). Three days later, growth phenotypes were scored and autoactivators removed from the OEC libraries. The AD-OEC yeast clones were pooled by separate growth o/n and unification into one solution. Equal representation of clones in pools was confirmed by plating and colony PCR on 30 colonies. The Arabidopsis library used is described in (Mukhtar et al., 2011); Consortium, 2011). For the screen, single DB-OEC clones were mated with pools of 192 AD-At clones, while single DB-At clones were screened against the AD-OEC pool. The screen was repeated once. Five µl of freshly grown DB- and AD-yeast were spotted on top of each other on YEPD medium using a robotic fluid handling device. Plates were incubated o/n, colonies replated onto interaction medium as well as cycloheximide (CHX) autoactivator control plates (Sc–Leu–His + 3-AT + CHX (1 mg/l)), incubated o/n and replica-cleaned. After five days incubation, single colonies were isolated and rearrayed into 96-well plates. These primary positive interactors were reevaluated in a secondary screen. They were plated onto diploid-selection medium, incubated two days, and transferred to interaction medium plates (Sc–Leu–Trp–His + 3-AT). Three autoactivator plates (Sc-Leu-His + 1 mM 3-AT + CHX) were also included. Plates were replica-cleaned and incubated three days. Positive clones were restreaked to diploid selection medium, incubated two days and lysed. PCR was used to obtain sequence information on corresponding AD- and DB-clones per colony. The interactors were identified by BLAST searches, single clones of these interactors retrieved from the stock and rearrayed for the retest screen. Matings of single clones were performed as described above, but phenotypes were scored on both Sc–Leu–Trp–His + 3-AT and Sc–Leu–Trp–Ade plates. Interactions were scored as verified when they were positive in three of four repeated matings and autoactivation was never detected.

The experimental methods used to define *Gor*_EHIn were identical to those previously used for mapping bacterial and oomycete effector interactions, and for

producing the Arabidopsis Interactome AI-1 (Consortium, 2011; Mukhtar et al., 2011). Consequently, key parameters of the interactome screen such as sampling- and assay-sensitivity are identical between the experiments and integration of the data will not introduce biases due to experimental design (Venkatesan et al., 2009). Moreover, the 8k_space was systematically tested in all experiments and thus forms a common scaffold for integration. (**Figure S1A**).

2. T-DNA lines and pathogen assays

Homozygous insertion mutants were ordered from ABRC for 124 of 165 effector interactors. Homozygosity and correct insertion sites were verified by PCR using standard conditions. Plants were grown under short day conditions (9 h light, 21°C; 15 h dark, 18°C).

The phenotypic assays have different degrees of difficulty. We funneled the mutant collection through these assays from simplest (*Hpa*), for which we screened 179 mutants and second alleles extensively, to most difficult (*Pto*), where we focused on the mutants that had altered *Hpa* phenotypes and whose products interacted with the most effectors.

Hyaloperonospora arabidopsidis (Hpa) isolates, inoculations, and growth assays. Hpa isolates Emwa1, Emoy2, and Noco2 were propagated on the susceptible Arabidopsis ecotypes Ws-2, Oy-1 and Col-0, respectively (Dangl et al., 1992; Holub et al., 1994). Twelve day old seedlings were inoculated with sporangia suspended in water at a concentration of 30,000 spores/ml. Plants were kept covered with a lid to increase humidity and grown at 21°C with a 9 hrs light period. Sporangiophores were counted on cotyledons at 4 (*Hpa* Noco2) or 5 (*Hpa* Emwa1 and Emoy2) days post-infection (dpi) as described (Holt et al., 2005). The number of sporangiophores per cotyledon was determined on approximately 100 cotyledons / genotype.

Bacterial infection assays. P. syringae pv tomato DC3000 growth assays were performed as previously described (Holt et al., 2005) with modifications. Briefly, bacteria were resuspended in 10 mM $MgCl_2$ to ~$1x10^5$ cfu/ml and syringe infiltrated into leaves of ~5 week old wild type and mutant plants. Leaf discs were cut from the infiltrated area on the day of infiltration (0 dpi) and 3 dpi, and placed into Eppendorff tubes containing 3 glass beads and 400µl 10 mM $MgCl_2$. Tissue was ground using a Fastprep-24 Instrument (MP Biomedicals). Serial dilutions were plated on KB-agar plates and cfu/ml were determined. For day 0 samples, four leaf discs were transferred to the same microfuge tube and processed as described above. For day 3 samples, 8 x 4 leaf discs were processed. The experiment was repeated at least 3 times.

Fungal infection assays. Powdery mildew infections were carried out as described previously, except that spores were harvested at 7 dpi (Weßling et al., 2012). Inoculations were either performed on 18 day old seedlings or 4-5 week old plants. Briefly, inoculations were either performed on three pots per genotpye containing ~200 18 day old seedlings or four 4-5 week old plants were inoculated

in a settling tower with *G. orontii* spores harvested from four leaves 14-21 dpi. Each round of inoculation included nine pots of randomized genotypes, thus all genotypes were included in three separate inoculations. After one minute incubation the pots were returned to the growth chamber. At seven dpi, three times 200 mg (older plants) or 500 mg (seedlings) plant material was harvested across pots and *G. orontii* spores isolated by shaking in water. The number of spores/g fresh weight was determined by counting eight chambers in a hemocytometer.

3. Transient expression in *N. benthamiana*

Agrobacterium-mediated transient expression assay: N. benthamiana plants were grown in a growth chamber equipped with LGM550 Professional LED Grow Light (LED Grow Master Global LLC, USA) at 24°C(day)/20°C (night) under a 16-h light/8-h dark cycle. *A. tumefaciens* strain GV3101 containing protein expression constructs was grown at 28°C with appropriate antibiotics for 18-24 h. Agrobacterium cells were collected by centrifugation at 10,000 RPM for 1 min, and then resuspended in induction solution (10 mM MES (pH 5.6), 10 mM $MgCl_2$, and 150 µM acetosyringone). Cell suspensions were incubated at room temperature for 2 h before infiltration into *N. benthaminana*. For co-infiltration, Agrobacterium strains expressing different proteins were mixed together at the desired final OD_{600} values (effector OD_{600} =1.5; TCP14 OD_{600} =0.05; and p19 (silencing suppressor) expression plasmid OD_{600} = 0.1) and infiltrated into leaves of 5- to 6-week old *N. benthamiana* plants with a 1 ml needleless syringe.

DAPI staining: DAPI (1 µg/ml) was infiltrated into leaves 1 h before confocal imaging.

Estradiol treatment: 20 µM Estradiol in water with 0.004% Silwet L77 was applied to both abaxial and adaxial sides of leaves 6-8 hours before confocal imaging. The treatment was repeated 1 h later.

Confocal Microscopy Imaging: Leaf discs (5 mm diameter) were collected at 40-48 hours after infiltration. Each effector/TCP14 combination was assayed twice. The abaxial sides of three leaf discs from each co-infiltrated leaf were observed with a confocal microscope (LSM 7 DUO; Carl Zeiss). All samples were imaged with a 40x water objective. Between 5 and 15 nuclei were observed in each repetition. The confocal images were edited with Zen 2009 (Zeiss) and Adobe Photoshop CS2. Zen 2009 (Zeiss) and Excel (Microsoft) were used to create histograms. The excitation and detection wave lengths are listed in **Table S5**.

4. Western Blotting

Proteins were isolated in lysis buffer (20 mM HEPES pH 7.5; 13 % Sucrose; 1 mM EDTA; 1 mM Dithiothreitol (DTT); 0.01 % Triton, 1x complete protease inhibitor cocktail (Roche)) from two 0.9 cm leaf discs/experiment using metal beads and a mixer mill (Retsch, Haan, Germany). After addition of 1 volume loading buffer (125 mM Tris-HCl pH 6.8; 5 % sodium dodecyl sulfate (SDS); 25 % glycerol (v/v); 0.025 % bromphenol blue (w/v); 0.2 M DTT), sample were

denaturated for 5 min at 95°C and the supernatant used for gel electrophoresis and western blotting by standard methods. Fusion proteins were detected by anti-GFP (Roche), anti-HA (Roche) and anti-c-myc (Sigma-Aldrich) antibodies according to manufacturer's instructions.

Supplemental Bioinformatic Procedures

5. Convergence analyses

Intraspecies convergence: Significance of intraspecies convergence was determined experimentally based on the experimentally observed number of interacting *A. thaliana* host proteins within Space 8k_sys for effectors from *Hpa*, *Psy*, and *Gor* provided in **Figure S1B**. For each pathogen effector interactors were sampled randomly from a list of AI-1_{MAIN} proteins (not shown) and from a degree preserved list of AI-1_{MAIN} loci (**Figure 2C-E**) (Consortium, 2011) using the "*sample*" command in R. The second analysis is more stringent as it increases the probability of repeatedly picking the more connected proteins and therefore leads to a lower number of nodes expected by chance. The distribution obtained from 10,000 samplings were plotted and compared to the experimentally observed value. The experimental *P* value was calculated by dividing the number of samplings where the number of common targets is greater or equals the observed number of common targets by the number of samplings performed. If the observed number of targets is not seen in the simulation, the *P* value is set to < 0.001.

$$\text{observed p} - \text{value} = \frac{\begin{array}{c}\text{number of samplings where number of common targets} \\ \geq \text{observed number of common targets}\end{array}}{\text{number of samplings}}$$

Interspecies convergence statistics: Significance of the convergence of effectors from different pathogens interacting with common host proteins was determined experimentally. The convergence was determined for all possible pathogen combinations based on the numbers of common interaction partners provided in **Figure 2F**. For each pathogen the number of host interaction partners was sampled randomly from a unique list of proteins in AI-1_{MAIN} (Consortium, 2011) using the "sample" command in R with replacement. The observed number of common proteins for each pathogen combination in each of 10,000 samplings was plotted as a background expectation and compared to the experimentally observed value of common interaction partners provided in **Figure 2F**. The experimental *P* value was calculated by dividing the number of samplings where the number of common targets is greater or equals the observed number of common interactors by the number of samplings performed. If the observed value of common targets is not seen in the simulation, the *P* value is set to < 0.001.

$$\text{observed p} - \text{value} = \frac{\begin{array}{c}\text{number of samplings where number of common targets} \\ \geq \text{observed number of common targets}\end{array}}{\text{number of samplings}}$$

6. Gene Ontology (GO) enrichment

We used GO enrichment analysis to test, which functional processes are overrepresented i) among effector interacting proteins, and ii) among the top 55 ranking genes. To this end we performed a GO enrichment analysis using all loci in AI-1_{MAIN} as the background distribution. The analysis is based on GO annotations of TAIR10 (timestamp: 2013-09-03), which we downloaded from the TAIR ftp-server. We removed all annotations with the evidence code "inferred from electronic annotation" (IEA). 17 out of 2661 loci in AI-1_{MAIN} do not have any manually curated GO annotation. For enrichment analysis we used the GOstats package version 2.28.0 (Falcon and Gentleman, 2007). We used the function hyperGTest to perform a hypergeometric test on the GO terms. We used a *P* value cut-off of 0.005 and “conditional testing”, which means that parent terms are tested without genes, which already have been found to be significant in a children term.

7. Scoring of phenotypic assays

Scoring significant phenotypes: The three pathogens *H. arabidopsidis*, *P. syringae* and *G. orontii* used in our infection assays have different lifestyles and therefore the level of infection of wild-type and mutant plants is assessed using different statistical approaches. Depending on the data collected for each pathogen we used different statistical tests to determine if a T-DNA line shows a significant difference in pathogen infestation in the infection assay compared to the Col-0 control plants.

Values of *G. orontii* experiments of adult plants and seedlings have been derived from hemocytometer counts and represent spores/g fresh weight. A Gaussian generalized linear model was fitted on the data and used for ANOVA analysis (R package “car”)(Fox and Weisberg, 2011). As we have repeated the inoculation experiments up to four times with a T-DNA line and the control line we treat it as a block experiment, where every T-DNA line and the respective control plants of one batch are treated as one block. The block is treated as second factor in our ANOVA analysis beside the first factor of the knocked-out gene. Our data were analyzed as a two-way ANOVA experiment with the factors gene and batch. The Benjamini & Hochberg method (Benjamini and Hochberg, 1995) was applied for multiple testing correction.

P. syringae data are represented as cfu/ml and have the same characteristics like *G. orontii* data and have been analyzed the same way.

H. arabidopsidis isolates Emwa1, Emoy2 and Noco2 data sets were collected as counts of sporangiophores per cotyledon. These data do not satisfy the requirement of ANOVA for normality distribution. This requires the use of a non-parametric test. We used the Kruskal-Wallis test to calculate the *P* value and corrected the results with Bonferroni multiple testing correction method.

Determination of edr vs. eds phenotypes: To determine if a given insertion mutant shows an *eds* or *edr* phenotype compared to Col-0 accession, we calculated a $\log_2$ fold change.

For each pathogen / pathogen strain *p* the mean value *x* of raw spore / sporangiophores counts for each T-DNA and Col-0 control line were normalized by scaling values between 0 and 1.

$$\text{Normalized}\left(\mathrm{x_{i,p}}\right) = \frac{\mathrm{x_{i,p}} - \mathrm{X_{min,p}}}{\mathrm{X_{max,p}} - \mathrm{X_{min,p}}}$$

where $x_{i,p}$ is the raw mean value of the pathogen *p* and the value of the tested gene i. $X_{min,p}$ is the minimum value of pathogen *p* and $X_{max,p}$ is the maximum value of pathogen *p*.

The phenotype of the T-DNA line with respect to the Col-0 control plants was evaluated by calculating the fold change of the mean normalized values of all available batches for each T-DNA line. The average fold change of all batches for a given T-DNA line was converted to a $\log_2$ fold change. A $\log_2$ fold change of 0 means same pathogen infestation of T-DNA line and control line, a negative $\log_2$ fold change shows a lower infestation (enhanced disease resistance) and a positive $\log_2$ fold change indicates a higher infestation (enhanced disease susceptibility).

$$\mathrm{fc_{i,p,k}} = \frac{\text{Normalized}\ (\mathrm{x^{M}_{i,p,k}})}{\text{Normalized}\ (\mathrm{x^{C}_{i,p,k}})}$$

$fc_{i,p,k}$ is the fold change of T-DNA line i inoculated with pathogen p in batch k. Normalized($x^{M}_{i,p,k}$) is the normalized value of T-DNA line i inoculated with pathogen p in batch k and Normalized($x^{C}_{i,p,k}$) is the normalized value of control line of T-DNA line i inoculated with pathogen p in batch k.

$$\mathrm{Log_2 fc_{i,p}} = \log_2\left(\frac{\mathrm{fc_{i,p,k}}}{\mathrm{n_i}}\right)$$

$Log_2\ fc_{i,p}$ is the log2 fold change of T-DNA line i of pathogen p. $fc_{i,p,k}$ is divided by the number of batches n_i of the respective T-DNA line i.

Merging phenotypes for multiple T-DNA lines: For the summarized phenotypic analysis of mutants we combined the phenotypic outcome if more than one T-DNA line per gene was available. Therefore we compared the pathogen-specific

P values of the phenotypes on the different T-DNA lines representing the same gene. We selected the phenotypic outcome with the lowest (most significant) *P* value to obtain merged phenotypes for a gene. We found no contradictory phenotypes for any pathogen, i.e. we had no case where we observed an *edr* phenotype in one mutant line and an *eds* phenotype in the other mutant line. In 15 cases we observed a statistically significant pathogen-specific phenotype in one allele of a gene and no phenotype for the second allele of this gene. In these instances we selected the outcome of the line with the statistically significant result (i.e. the line showing the altered pathogen infection phenotype).

8. Arabidopsis Consensus Sequence Building and AAP evaluation

To evaluate natural variation in Arabidopsis accessions we used the complete genomes of 80 accessions sequenced in the context of the 1001Genomes project and mapped on the Col-0 reference genome. These were collected in eight regions distributed over Europe and Asia, where Arabidopsis naturally occurs and provide a large spatial and phylogenetic distribution adapted in different environments. This dataset was published by Cao et al., 2011 and can be downloaded from 1001genomes.org.

A challenge for the quantitative evaluation of coding variation is the fact that single nucleotide polymorphisms (SNPs) are reported relative to the Columbia (Col-0) reference genome (Cao et al., 2011). This results in numerous SNPs being called in all 80 accessions. In contrast, a conservative nomenclature would identify this as a SNP in Col-0 only. We used a majority voting scheme to define the consensus sequence of the Arabidopsis population consisting of the genomic sequences of Col-0 and the dataset MPICao2010 (Cao et al., 2011), which is available at 1001genomes.org (Altmann *et al.*, in preparation). In this scheme, the most common base at any position defines the consensus sequence and all other variants that occur in the population are counted as variant SNPs where they occur. The codons in coding regions of the consensus sequence of representative gene models as annotated in TAIR10 were compared against the respective codons in the 81 accessions in the genome matrix. In each accession we examined the codons for synonymous SNPs (sSNPs), non-synonymous SNPs (nsSNPs) and the resulting amino acid. A unique amino acid polymorphism (AAP) is defined qualitatively as a specific amino acid substitution at a given position, independent of the frequency of how often the specific substitution was found in the analyzed population. In other words a SNP leading to a hypothetical G->A substitution counts as a single unique AAP independent of how often this amino acid replacement occurred; a substitution resulting in a hypothetical G->T replacement is counted as a second unique AAP. For each protein we counted the unique number of AAPs observed for an individual position. We calculated the sum of unique AAPs per position of a protein, which results in the number of unique AAPs in a protein.

9. Calculation of Tajima's D (D_T) and Watterson's θ (θ_W)

In order to determine D_T and θ_W values for all genes in AI-1_{Main}, we extracted the aligned genomic sequences of all 80 accessions and Col-0 for the respective

representative gene models as fasta file from the whole genome alignment (TAIR10_genome_matrix_2012_03_13.txt.gz) from the dataset MPICao2010 (Cao et al., 2011) (1001genomes.org/projects/MPICao2010/). This information was used to calculate D_T and θ_W using the standard settings in the *compute* program of the *analysis* software package (version 0.8.4) developed by the Thornton lab (Thornton, 2003).

To rank genes for their genetic variation within the 81 accessions, we calculated a combined rank of D_T and θ_W (Dθ-ranking). We sorted all genes in descending order according to their D_T and θ_W values. Subsequently genes were ranked according to the ascending order of the mean rank of these two lists.

10. Statistics of D_T for effector targets and their AI-1_{MAIN} interactors

To determine whether the D_T values of effector-interactors, or that of their AI-1_{MAIN} interaction partners, deviate significantly from the background distribution, we performed sampling and random rewiring analyses for the four effector groups as indicated in the main text (all targets; targets of two or three pathogens; targets of three pathogens; and targets with phenotype).

To evaluate D_T for effector-interactors, we randomly drew without replacement 1,000 samples of the same size as the respective effector group from the unique set of genes encoding proteins in AI-1_{MAIN} as random control. To analyze the significance of D_T of the AI-1_{MAIN} interaction partners of effector-interactors, we generated 1,000 randomized degree-preserved networks and determined the mean D_T of all interaction partners of the respective effector-interactors in the random networks. For each of the four sets of effector-interactors and their AI-1_{MAIN} interaction partners, we calculated a two-sided *P* value for the observed mean compared to the distribution of the 1,000 mean values of the random controls. We counted the number of occurrences greater equal and lower equal than population mean +/- (population mean - sample mean), and divided it by the number of samples.

11. Statistical evaluation of top Dθ- and AAP-ranking genes with effector-interactors

Dθ-ranking: We investigated whether proteins encoded by top Dθ-ranking genes preferentially interact with effector-interactors. This analysis was performed for interactions with the four different sets of effector targets: all effector targets; proteins interacting with effectors from two or three pathogens; proteins interacting with effectors of three pathogens; and effector targets showing a phenotype in the phenotyping assay.

We performed 100,000 times a degree-preserving random rewiring of the AI-1_{MAIN} network by permuting two interaction partners of two randomly selected edges using the *rewire* function in the *igraph* R package v0.7.0 (Csardi and Nepusz, 2006). This was repeated to generate 1,000 rewired networks. In each random network we counted the number of effector-interactors interacting with

proteins encoded by the cumulative 1 – 70 top Dθ-ranking genes and the analysis was repeated for each class of effector-interactors. The data was used to calculate the experimental *P* value for the probability of finding the experimentally observed number of interactions between top Dθ-ranking genes with effector-interactors by chance. We calculated an observed *P* value by dividing the number of observations with a value greater equal than the real number of observation by the number of generated rewired networks.

Amino Acid Polymorphism Effector Interactors Evaluation: To determine, if loci having a high number of AAPs interact more often with effector targets than other loci, we evaluated the AAP in the same way than the combined Dθ ranking. We sorted the loci descending by their number of AAPs and determined the cumulative number of interacting effector-interactors. To calculate a *P* value we compared the real value against the distribution of number of effector targets from 1,000 rewired networks. *P* value calculation and network rewiring was conducted as for evaluation of combined ranking of D_T and θ_W.

12. Fisher Exact Contingency Tables

Proteins that are object of intraspecies convergence are also object of interspecies convergence (based on **Figure S2B**)

AI-1_{MAIN} Effector interactors	Intraspecies convergence	no intraspecies convergence	Total
Interspecies convergence	25	7	32
no intraspecies convergence	31	92	123
Total	56	99	**155**

Two-tailed Fisher's exact test: $P < 0.0001$

REFERENCES

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B-Methodological *57*, 289-300.

Cao, J., Schneeberger, K., Ossowski, S., Gunther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C.*, et al.* (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nat Genet *43*, 956-963.

Consortium, A.I.M. (2011). Evidence for network evolution in an Arabidopsis interactome map. Science *333*, 601-607.

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. InterJournal *Complex Systems*.

Dangl, J., Holub, E., Debener, T., Lehnackers, H., Ritter, C., and Crute, I. (1992). Genetic definition of loci involved in Arabidopsis-pathogen interactions. In Methods in Arabidopsis Research, C. Koncz, N.-H. Chua, and S. J, eds. (World Scientific), pp. 393-418.

Dreze, M., Monachello, D., Lurin, C., Cusick, M.E., Hill, D.E., Vidal, M., and Braun, P. (2010). High-quality binary interactome mapping. Methods Enzymol *470*, 281-315.

Falcon, S., and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. Bioinformatics *23*, 257-258.

Fox, J., and Weisberg, S. (2011). An R Companion to Applied Regression (SAGE Publications).

Holt, B.F., 3rd, Belkhadir, Y., and Dangl, J.L. (2005). Antagonistic control of disease resistance protein stability in the plant immune system. Science *309*, 929-932.

Holub, E.B., Beynon, L.J., and Crute, I.R. (1994). Phenotypic and genotypic characterization of interactions between isolates of *Peronospora-parasitica* and accessions of *Arabidopsis-thaliana*. Molecular Plant-Microbe Interactions *7*, 223-239.

Mukhtar, M.S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan, M., Galli, M., Hao, T., Nishimura, M.T.*, et al.* (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. Science *333*, 596-601.

Thornton, K. (2003). Libsequence: a C++ class library for evolutionary genetic analysis. Bioinformatics *19*, 2325-2327.

Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I.*, et al.* (2009). An empirical framework for binary interactome mapping. Nat Methods *6*, 83-90.