PNAS
www.pnas.org

Supplementary Information for

Specific modulation of the root immune system by a community of commensal bacteria

Paulo J. P. L. Teixeira, Nicholas R. Colaianni, Theresa F. Law, Jonathan M. Conway, Sarah Gilbert, Haofan Li, Isai Salas-González, Darshana Panda, Nicole M. Del Risco, Omri M. Finkel, Gabriel Castrillo, Piotr Mieczkowski, Corbin D. Jones, Jeffery L. Dangl*

* To whom correspondence should be addressed.
Email: dangl@email.unc.edu

This PDF file includes:

Supplementary text
Figures S1 to S10
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S7

Supplementary Information Text

Materials and Methods

Preparation of bacterial synthetic communities (SynComs). We employed SynComs to evaluate the effect of the root microbiota on plant immune responses. A 35-member SynCom (SynCom35) was assembled using a set of genome-sequenced strains that represent the typical taxonomic diversity of plant-associated bacteria (1). This SynCom included 32 strains that were isolated from *Brassicaceae* roots (mainly Arabidopsis), two strains that were isolated from unplanted soil (Mason Farm - North Carolina, USA; +35° 53′ 30.40″, −79° 1′ 5.37″) and *Escherichia coli* DH5α as a control. Smaller SynComs were assembled with subsets of strains from this 35-member SynCom (i.e., SynCom5-S: five suppressors; and SynCom5-NS: five non-suppressors). A detailed description of the strains that comprise the SynComs used in this study is provided in SI Appendix, Dataset 1.

Assembly of SynComs followed the procedures described by Castrillo et al. (1). Briefly, a single colony of each strain was inoculated into 4 mL 2xYT medium (16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl) in a test tube and kept at 28°C under agitation at 250 rpm. After one to two days, bacterial cultures were washed twice with 10 mM $MgCl_2$ (by centrifugation and resuspension) and OD_{600nm} was determined by spectrophotometry. Cultured bacterial strains were then added to 500 mL of medium to a final concentration of 10^5 cfu/mL each (assuming that one OD_{600nm} unit equals to 10^9 cfu/mL). In some experimental conditions, the bacterial elicitor flagellin 22 (flg22) was added along with bacteria at a final concentration of 1 µM (described below).

Preparation of plants for bacteria and flg22 treatments. *Arabidopsis thaliana* seeds (ecotype Columbia; Col-0) were surface-sterilized with a cleaning solution (70% bleach and 0.2% Tween-20) for 8 min and rinsed three times with sterile distilled water to eliminate any seed-borne microbes on the seed surface. After stratification at 4°C for 2 days in the dark, seeds were transferred to square plates containing Johnson medium (5 g/L sucrose; 85 mg/L KH_2PO_4; 0.6 g/L KNO_3; 0.9 g/L $Ca(NO_3)_2{\cdot}4H_2O$; 0.2 g/L $MgSO_4{\cdot}7H_2O$; 3.8 mg/L KCl; 1.5 mg/L H_3BO_3; 0.8 mg/L $MnSO_4{\cdot}H_2O$; 0.6 mg/L $ZnSO_4{\cdot}7H_2O$; 0.1 mg/L $CuSO_4{\cdot}5H_2O$; 16.1 µg/L H_2MoO_4; 1.1 mg/L $FeSO_4{\cdot}7H_2O$; 0.1 g/L Myo-Inositol; 0.5 g/L MES; pH 5.6–5.7) solidified with 1% bacto-agar. Plates were kept in the vertical position for seven days in a growth chamber under a 16-hour light/8-hour dark regime at 21°C day/18°C night. The resulting seedlings were then ready for experiments.

Treatment of plants with flg22 in the presence of SynComs. We evaluated the effect of SynCom35 on the plant response to flg22 (Fig. 1) by cultivating plants in Johnson medium with or

without 1 µM of flg22 (synthesized by PhytoTech Labs) in the presence or not of the SynCom (inoculum: 10^5 cfu/mL of each strain; prepared as described above). A control condition in which plants were exposed to the heat-killed SynCom was included in the experiment by heating the bacteria at 100°C for 2h in an oven immediately before inoculating them into the medium at the same concentration as the SynCom alive. The complete design of this experiment as well as the parameters measured are shown in SI Appendix, Fig. S1.

Seven-day-old seedlings were obtained by germinating the seeds in Johnson medium as described above and then transferred to the six different experimental conditions (SI Appendix, Fig. S1). Plates were kept in a growth chamber under a 9-hour light/15-hour dark regime at 21°C day/18°C night. The entire root system of the seedlings was harvested for RNA extraction at two time-points: 1 and 12 days after the transfer to the treatment plates. Tubes were stored at −80°C until processing. The experiment included 9 biological replicates per condition, which were sub-divided into 3 batches (each containing 3 replicates per condition) that were performed independently (i.e., different days, medium, seedlings and bacterial inoculum). Each replicate corresponds to the root systems of ten seedlings grown in a single plate.

Treatment of plants with flg22 in the presence of individual strains. We evaluated the effect of each member of SynCom35 on the plant response to 1 µM flg22 in the context of mono-associations (Fig. 2 and SI Appendix, Fig. S5). For this, each strain was cultured and added individually to 500 mL of medium as described above at a concentration of 10^5 cfu/mL. Treatment of seedlings with flg22 in mono-association was performed as described for SynCom35 (see above). Seven-day-old seedlings were transferred to the different conditions of the experiment and harvested after 12 days for RNA extraction. Tubes were stored at −80°C until processing. The experiment included 9 biological replicates per condition, which were sub-divided into 3 batches (each containing 3 replicates per condition) that were performed independently (i.e., different days, medium, seedlings and bacterial inoculum). Each replicate corresponds to ten seedlings grown in a single plate. SI Appendix, Fig. S5 summarizes the design of this experiment.

Root length measurements. Before harvesting samples for RNA extraction at day 12 (end of the experiments), plates were scanned, and the root length of each seedling (10 per plate) was determined using ImageJ (2). Results from the SynCom35 experiment (Fig. 1 and SI Appendix, Fig. S4) were visualized with a boxplot made in R using the ggplot2 package (3). Due to the large number of treatments (i.e., strains) in the mono-association experiment (Fig. 2), results were visualized with a heatmap. For this, the raw root length data was fit with a linear model using the lm function in R. We identified a significant batch effect using the model:

$$Root\_length \sim Bacteria * flg22 + Batch$$

“Root_length” is the root length from each replicate, “Bacteria” is the strain inoculated into the plant growth medium, “flg22” is the presence of 1 µM flg22 or not, and “Batch” is the experimental effect (i.e., each of three independent batches of the experiment). To correct the root lengths for the batch effect, we used the beta effects of bacteria, flg22, and bacteria:flg22 from the linear model to calculate fitted root length measurements for each bacterial strain, one with flg22 and one without. These fitted measurements were then clustered using a Euclidean distance matrix and the ward.d2 clustering method implemented in the hclust function in R. The data was plotted with the R package ComplexHeatmap (4). The heatmap was annotated to indicate the Class of each strain as well as the presence or not of a type-3 secretion system (T3SS), which was predicted using macsyfinder and hidden Markov Models described by Abby et al. (5). The code used for this analysis can be found at GitHub (https://github.com/ncolaian/MAE).

RNA extraction and RNA-seq library preparation. Total RNA was extracted following the procedures described by Logemann et al. (6). Frozen plant tissue was pulverized inside a 2 mL tube containing three 4 mm glass beads using a Qiagen TissueLyser II instrument. Samples were homogenized in 400 µL of Z6-buffer (8 M guanidinium- HCl, 20 mM MES, 20 mM EDTA, pH 7.0). Following the addition of 400 µL phenol:chloroform:isoamyl alcohol (25:24:1), samples were vortexed and centrifuged (20,000 *g*, 10 min) for phase separation. The aqueous phase was transferred to a new 1.5 mL tube and 0.05 volumes of 1 N acetic acid and 0.7 volumes 96% ethanol were added. The RNA was precipitated at −20 °C overnight. Following centrifugation (20,000 *g*, 10 min, 4 °C) the pellet was washed with 200 µL sodium acetate (pH 5.2) and then 70% ethanol. The RNA was dried and dissolved in 30 µL of ultrapure water and stored at −80 °C until use.

Illumina mRNA-seq libraries were prepared from 1,000 ng RNA using the protocol described by Castrillo et al. (1). Briefly, mRNA was purified from total RNA using Sera-mag oligo(dT) magnetic beads (GE Healthcare Life Sciences) and then fragmented in the presence of divalent cations (Mg^{2+}) at 94°C for 6 min. The resulting fragmented mRNA was used for first-strand cDNA synthesis using random hexamers (ThermoFisher Scientific) and the EnzScript reverse transcriptase (Enzymatics), followed by second strand cDNA synthesis using DNA polymerase I (Enzymatics) and RNaseH (Enzymatics). Double-stranded cDNA was end-repaired using T4 DNA polymerase (Enzymatics), T4 poly-nucleotide kinase (Enzymatics) and Klenow polymerase (Enzymatics). The DNA fragments were then adenylated using Klenow exo-polymerase (Enzymatics) to allow the ligation of Illumina Truseq HT adapters (D501–D508 and D701–D712). Following library preparation, quality control and quantification were performed using a 2100 Bioanalyzer instrument (Agilent) and the Quant-iT PicoGreen dsDNA Reagent (ThermoFisher Scientific), respectively. Libraries were sequenced on the Illumina HiSeq4000 instrument to generate 50-bp single-end reads. A total of 100 and 666 libraries were produced for the SynCom35 (Fig. 1) and the mono-

association (Fig. 2) experiments, respectively. SI Appendix, Datasets 3 and 4 present details about each sample included in these experiments.

Processing of RNA-seq reads. Initial quality assessment of raw sequences was done with FastQC version 0.11.7 (Babraham Bioinformatics, Cambridge, UK). Trimmomatic version 0.36 (7) was used to remove adaptor-containing and low-quality sequences with parameters set at ILLUMINACLIP:TruSeq3-SE.fa:2:30:10, SLIDINGWINDOW:4:5, LEADING:5, TRAILING:5, MINLEN:50. The resulting high-quality reads were then aligned against the TAIR10 Arabidopsis reference genome using HiSat2 version 2.1.0 (8) using default parameters. The featureCounts function from the Sub-read package version 1.6.3 (9) was used to count reads that mapped to each one of the 27,206 nuclear protein-coding genes. Evaluation of the results of each step of the analysis was done with MultiQC version 1.7 (10). Raw RNA-seq reads and the count matrices generated in this study are available at the NCBI Gene Expression Omnibus under the accession number GSE156426.

Evaluation of the transcriptional response to flg22 in plants growing in the presence of SynComs. Identification of differentially expressed genes was performed using the generalized linear model (glm) approach (11) implemented in the R package edgeR (12). In the SynCom35 experiment (Fig. 1), weakly expressed genes were filtered out by removing those genes that did not achieve a minimum expression level of 1 count per million in at least five libraries. Normalization was performed using the trimmed mean of M-values method (TMM; function calcNormFactors in edgeR) (13). The glmFit function was used to fit the counts to a negative binomial generalized linear model with a log link function. In this experiment, our goal was to evaluate the root transcriptional response to flg22 in plants grown in the presence of the SynCom35 (alive or dead) or in the absence of bacteria (Fig. 1*B*). For this, we employed a one-way-layout model to compare each treatment ('no bacteria.flg22', 'SynCom35.flg22' or 'heat-killed SynCom35.flg22') to its control condition ('no bacteria.mock, 'SynCom35.mock' or 'heat-killed SynCom35.mock') at each time-point (T1h and T12h). The model included the covariate 'Experiment' to control for batch effects associated with the three independent repetitions of the experiment. The Benjamini–Hochberg method (false discovery rate; FDR) was used for the correction of multiple comparisons (14). Genes with an FDR corrected p-value below or equal to 0.01 and a fold-change of at least 1.5x were considered differentially expressed. We compared the sets of differentially expressed genes in each condition ('no bacteria', 'SynCom' and 'Heat-killed SynCom') using a Venn diagram (15) (Fig. 1*C*). Pearson correlations of flg22 responses (Fig. 1*D*) were estimated by comparing the fold-changes of all 715 genes that responded to flg22 in any of our conditions using the R package corrplot (16). A Constrained Analysis of Principal Coordinates (CAP) was performed to allow the visualization of the relative effects of our experimental conditions on the plant transcriptome (Fig. 1*E*). For this, we

selected the 500 genes with the highest standard deviation across all samples from the matrix of normalized gene counts generated by edgeR and used the capscale function from the vegan package (17) to compute the CAP. We constrained the ordination on the presence of the SynCom and the addition of flg22, while conditioning for the batch effect. Hierarchical clustering (Fig. 1*F* and SI Appendix, Fig. S3*C*) was performed with the 'heatmap.2′ function from the R package gplots (18) using those genes that responded to flg22 in at least one condition of the experiment (i.e, 'no bacteria', 'SynCom35' or 'Heat-killed SynCom35') within each time-point. Genes were clustered on the basis of the Euclidean distance with the complete-linkage method. The cutree function was used to define eight clusters based on dendrogram distances. To generate a representative view of each cluster (Fig. 1*G*), RPKM expression values were normalized by z-score transformation and presented in a boxplot. Gene Ontology (GO) enrichment analyses (Fig. 1*H* and SI Appendix, Figs S1 and S3) were performed with the compareCluster function of the R package clusterProfiler (19) and with the PlantGSEA platform (20).

Evaluation of the transcriptional response to flg22 in plants growing in the presence of individual strains. In the mono-associations experiment (Fig. 2), weakly expressed genes that did not have greater than 1 count per million in at least five libraries were removed from subsequent analyses. We then normalized the gene count matrix using the weighted trimmed mean of M-values (TMM) implemented in the calcNormFactors function from edgeR. We were interested in identifying (1) the plant genes that respond to bacteria in mono-association (bacteria_mock vs axenic_mock); and (2) plant genes that respond to flg22 in the presence of each strain in mono-association (bacteria_flg22 vs bacteria_mock) or in the 'no bacteria' control (axenic_flg22 vs axenic_mock). To perform these comparisons, we first subset the raw gene count table to only include the conditions being tested and created a DGEList object with this data using edgeR. We normalized the counts based on normalization factors calculated from the full gene count table. We then estimated the gene-wise dispersion parameters within each of the smaller datasets using the estimateGLMCommonDisp, estimateGLMTrendedDisp and estimateGLMTagwiseDisp functions from edgeR. Our design matrices were as follows for each of the contrasts listed above:

(1) expression ~ bact*experiment

(2) expression ~ treatment+experiment

"bact" is the bacterial strain, "experiment" is the batch effect (i.e., each of the three independent repetitions of the experiment), and "treatment" is the flg22 status. We then fit a gene-wise negative binomial generalized linear model with our calculated dispersion parameters using the glmFit function. After fitting the models, we conducted gene-wise log ratio tests contrasting our factor of

interest, (1) bacterial strain and (2) flg22 added, to the controls: (1) 'no bacteria' and (2) 'no flg22 added' using the glmLRT function. We used the decideTestsDGE function in edgeR to calculate FDR corrected p-values and report differentially expressed genes that had a log fold-change of at least 1.5 and an FDR corrected p-value of less than 0.01.

Principal Component Analysis (PCA) was performed with the set of 428 flg22-responsive genes defined in the control condition of the experiment (i.e., plants growing in the absence of bacteria) using the median expression values from the normalized counts matrix (log-transformed counts per million; with a prior count of 1) of the nine biological replicates from each condition. The function PCA from the R package AMOR (21) was employed with default parameters. The ellipses were defined using a multivariate t-distribution with a 70% confidence level. Code used for all analysis is found at https://github.com/ncolaian/mae.

DNA extraction for 16S sequencing. Plants were grown using MS medium (0.5X Murashige and Skoog basal medium with vitamins at pH 5.7 solidified with 1% bacto-agar) under the conditions and treatments described previously. Five roots were pooled and placed in 2.0 mL tubes with three sterile 4 mm glass beads. Samples were vortexed rigorously three times in sterile distilled water to remove agar particles and weakly associated microbes. Tubes containing the samples were stored at -80˚C until processing. Agar from each plate was collected in 30 mL syringes with a square of sterilized Miracloth (Millipore) at the bottom and stored at -20 °C until processing. Root samples were lyophilized for 48 hours using a Labconco freeze dry system and pulverized for 45 s in a FastPrep-24 Classic Instrument (MP Biomedicals). Agar containing syringes were thawed at room temperature and samples were squeezed gently through the Miracloth into 50 mL falcon tubes. Samples were centrifuged at max speed for 20 min and most of the supernatant was discarded. The remaining 1-2 mL of supernatant, containing the pellet, was transferred into clean 2 mL tubes. Samples were centrifuged again, supernatant was removed, and pellets were stored at -80 °C until DNA extraction. DNA extractions were carried out on ground root and agar pellets using 96-well format MoBio PowerSoil Kit (MOBIO Laboratories; Qiagen) following the manufacturer's instruction.

16S rDNA Sequencing Preparation. We amplified the V3-V4 regions of the bacterial 16S rRNA gene using the primers 357F (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**CCTACGGGAGGCAGCAG**-3′) and 806R (5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GGACTACHVGGGTWTCTAAT**-3′).
Each PCR reaction was performed in triplicate. PCR conditions were as follows with the Kapa HiFi Hotstart readymix: 1.2 µL Kapa Buffer, 0.18 µL dNTPs, 0.3 µL DMSO, 0.12 µL Kapa HiFi Polymerase, 0.3 µL 10 µM 357F, 0.3 µL 10 µM 806R, 0.9 µL mixed plant rRNA gene-blocking

peptide nucleic acids (PNAs; 1:1 mix of 10 µM plastid PNA and 10 µM mitochondrial PNA), 1.7 µL dH_2O, 1 µL DNA; temperature cycling: 95°C for 5 min; 28 cycles of 98°C for 20 s; 78°C (PNA) for 15 s; 55°C for 15 s; 72°C for 1 min; 4°C until use. Triplicate PCR products were pooled and then diluted twice; 2 µL of PCR products into 18 µL of dH20 and 5 µL of Dilution 1 into 45 µL of dH_2O. Following dilution, the PCR product was indexed using 16 indexed primers 357F (5'-AATGATACGGCGACCACCGAGATCTACACXXXXXXXXTCGTCGGCAGCGTC-3') and 24 indexed primers 806R (5'-CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTCTCGTGGGCTCGG-3'). PCR conditions were as follows with the Kapa HiFi Hotstart readymix: 2 µL Kapa Buffer, 0.3 µL dNTPs, 0.5 µL DMSO, 0.2 µL Kapa HiFi Polymerase, 0.5 µL 10 µM 357F index, 0.5 µL 10 µM 806R index, 0.15 µL PNAs; 1:1 mix of 100 µM plastid PNA and 100 µM mitochondrial PNA), 0.85 µL dH2O, 5 µL DNA; temperature cycling: 95°C for 5 min; 10 cycles of 98°C for 20 s; 78°C (PNA) for 15 s; 55°C for 15 s; 72°C for 1 min; 4°C until use. PCR products were purified using Sera-mag SpeedBeads (Fisher # 09-981-123) and quantified with a Qubit 2.0 fluorometer (Invitrogen) and then diluted to 6 pM for sequencing. Sequencing was performed on an Illumina MiSeq instrument using a 600- cycle V3 chemistry kit with 10% PhiX.

Analysis of the 16S sequencing data. Raw reads were trimmed with trimmomatic version 0.36 (7) using the parameters "HEADCROP:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15". Due to poor-quality base outputs in the beginnings and ends of reads we used headcrop and the leading/trailing trimming procedures. The sliding window command trims reads after the mean of 4 continuous base pairs reaches below a mean of 15. Once filtered, the resulting paired reads were analyzed using the DADA2 package version 1.14.1 (22) in R using the custom script code_for_16S_analysis.Rmd (https://github.com/ncolaian/MAE). Briefly, reads were filtered again by removing sequences containing uncalled bases, truncating reads after bases with a quality score of 2 and reads with no more than 3 expected errors using the FlilterAndTrim function in DADA2. We then learned the error rates for the forward and reverse reads separately with the learnErrors function. These error rates were then used to infer amplicon sequence variants (ASVs) with the function dada on the reverse and forward reads separately. After, we merged the forward and reverse strand results with the mergePairs function. The merged ASVs were used to construct an ASV sequencing table with only ASVs between 420 and 450 bp long, which reflect a majority of the reads from the V3 and V4 region, using the makeSequenceTable function. We then removed chimeras from the sequence table with the removeBimeraDenova function. The resulting ASVs were then mapped to the known V3 and V4 regions of the isolates in our 35-member SynCom at 98% identity using vsearch's -usearch_global function (2.14.2). We only considered ASVs that mapped to SynCom members in all further 16S analysis.

Analysis of flg22 effect on the assembly of SynCom35. We performed our constrained analysis of principal coordinates (CAPs) similar to Finkel et al. (23). Briefly, the number of reads from ASVs that mapped to each isolate were summed together to reflect a single abundance value for each member of SynCom35. Relative abundances were calculated by dividing each count by the total number of reads mapping to the SynCom35 members in that sample. We performed a constrained analysis of principal coordinates (CAPs) on the root sample's relative abundance data using the capscale function from the vegan version 2.5-6 package in R. We used this function to create a square-rooted Bray-Curtis distance matrix with the vegdist function, and ordinate the matrix based on the model formula: relative_abundance ~ flg22 + experiment_batch. We then performed a PERMANOVA to determine the significance of the model and each of the model terms with 5000 permutations using the anova.cca function from vegan. We used the sum of variance explained calculated from the anova.cca function to determine the constrained and unconstrained variance explained.

Differential abundance of microbes with the addition of flg22 in root samples. We took the raw reads for each bacterial strain and used the DESeq2 version 1.26.0 (24) package in R to call differential abundant microbes between flg22- and mock-treated root samples. We created a DEseq object using the DESeqDataSetfromMatrix function and used the model – reads ~ Biological.Replicate + flg22. To perform the differential abundance analysis, we first estimated the normalization factors using the estimateSizeFactors function implemented in DESeq2 using the poscounts method, which handles bacteria with 0 counts in some of the samples by calculating a modified geometric mean. We then estimated the dispersion estimates for the negative binomial distribution that will be used to call differential abundances with the estimateDispersions function in DESeq2. We then fit coefficients to each microbe indicating the change in abundance between treatments and tested these coefficients for significance using a Wald test implemented in the nbinomWaldTest function in DESeq2. We used an alpha value of 0.05 to assign significance.

Enrichment of members in root or agar samples. Each 16S experiment had a paired root and agar sample. We utilized this design to compare the relative abundance values for each suppressor (SynCom5-S1) and non-suppressor (SynCom-NS1) strain between the root and agar samples. We counted the number of times each strain was found at higher relative abundance values in either the agar or root sample. We did not compare relative abundance values if a strain was absent from both the root and agar samples in an experiment. The maximum number of experiments counted for each strain was 18, the minimum was 0 (for CL18 which was never found in any of our samples), and the median was 18. The counts (we called enrichments) for each strain were combined into the respective SynCom groupings for a total enrichment count of 68 for SynCom5-NS1 and 80 for

SynCom5-S1. The agar and root enrichments were then tabulated (SynCom X enrichment) and a chi-square test was performed on the table with the chisq.test function in R.

Evaluation of the root acute response to flg22. We evaluated the ability of each member of SynCom35 to suppress the roots response to flg22 in the context of an acute exposure to the elicitor (i.e., 5 hours) using the procedures described by Millet et al. (25) (SI Appendix, Fig. S6*D*). For this, we used an Arabidopsis line carrying the *pCYP71A12::GUS* reporter construct, which is activated by flg22 specifically in the root elongation zone (25). Approximately 10 to 15 bleach-sterilized seeds (procedure described above) were inoculated in each well of a 12-well microtiter plate containing 1 mL MS medium (Murashige and Skoog basal medium with vitamins containing 0.5 g/L MES hydrate and 0.5% sucrose at pH 5.7). The plate was kept in a growth chamber under 16-hour light/8-hour dark regime at 21°C day/18°C night for seven days. At the end of the seventh day, the medium was removed with a pipette and replaced by 1 mL MS medium only ('no bacteria' control) or 1 mL MS containing a single bacterial strain at OD_{600nm} of 0.002. The plate was returned to the growth chamber and incubated overnight (~14 hours) to allow bacterial colonization of the roots. The elicitor, flg22, was then added to the medium at a final concentration of 100 nM. Seedlings treated with *Pseudomonas simiae* WCS417 (overnight incubation) or with 10 µM MeJA (added along with flg22) were used as positive controls for the suppression of the flg22 response (25). Plates were kept in the growth chamber for 5 hours after flg22 addition and then submitted to the GUS histochemical assay.

GUS histochemical assays. After the treatment with bacteria and/or flg22, seedlings were washed twice with 50 mM sodium phosphate buffer, pH 7. Then, 1 mL of freshly prepared GUS substrate solution (50 mM sodium phosphate, pH 7; 10 mM EDTA; 0.5 mM $K_4[Fe(CN)_6]$; 0.5 mM $K_3[Fe(CN)_6]$; 0.5 mM X-Gluc; and 0.01% Silwet L-77) was added to each well of the plate, which was subsequently incubated in the dark for 4 hours at 37°C. GUS substrate solution was then replaced by 1 mL of a 3:1 ethanol:acetic acid solution and left at 4°C overnight (~14 hours). Before imaging, this solution was replaced with 1 mL 95% ethanol. Roots were visualized using a Leica M205FA stereoscope coupled to a Leica DFC310FX camera.

Root immune suppression by bacterial culture. We evaluated if culture filtrates of suppressor strains are sufficient to prevent the root response to flg22 following the procedures described by Yu et al. (26) (SI Appendix, Fig. S8*A*). Briefly, Arabidopsis seedlings were surface sterilized and grown for 10 days in a 12-well microtiter plate containing 1 mL MS medium as described above. After this, the plant growth medium was filtered using a 0.22 µM Millipore syringe filter and stored at -20°C until use. Bacteria were grown in 2xYT medium, washed in 10 mM $MgCl_2$ and then inoculated in a 12-well plate containing the plant exudate filtrate at a final OD_{600nm} of 0.002. After

incubating the plate for 22 hours in a growth chamber, cultures were passed through a 0.22 µM Millipore syringe filter, resulting in a bacteria filtrate. This filtrate was then supplied to 10-day-old *pCYP71A12::GUS* Arabidopsis seedlings for 1.5 hours, and 100 nM of flg22 (or a mock control) was then added. After five hours of flg22 treatment, seedlings were submitted to the GUS histochemical assay as described above. Plants growing in the absence of any bacteria or in the culture filtrate of the non-suppressor strain *Ochrobactrum sp*. MF370 were used as controls. The experiment also included *Pseudomonas simiae* strain WCS417, whose acidic culture filtrate suppresses the plant response to flg22 (26). To contrast our suppressor strains to *P. simiae* WCS417, we measured the pH of all bacterial filtrates using an Accumet ab15 pH meter (Fisher Scientific).

Phylogenetic tree construction and annotation. To construct a phylogenetic tree of the 35-member SynCom strains, we trimmed a previously constructed tree (23) that contained a total of 185 strains, including all members of the 35-member SynCom, using the drop.tip function from the ape package version 5.3 in R (27). This phylogenetic tree was based on 47 single copy genes that are present in every isolate of the larger 185-member collection. Single gene alignments were performed using MAFFT (28) and low-quality columns were filtered with trimAl (29). Gene alignments were then concatenated into a single super-alignment and used in the phylogenetic inference using FastTree version 2.1 with the WAG evolution model (30). The resulting tree was annotated with the results from five of our experiments that evaluated the ability of each strain to suppress/activate the plant immune responses: (I) Root-growth inhibition (RGI) triggered by flg22 (Fig. 2*A*); (II) Expression of the flg22 regulon (Fig. 2*B*); (III) Expression of the flg22 core gene set (SI Appendix, Fig. S6*C*); (IV) Interference with the plant response to acute exposure to flg22 (SI Appendix, Fig. S6*D*); (V) Expression of specific sectors of the flg22 regulon (Clusters M3, M4 and M5; Fig. 2*C*).

Quantification of bacterial colonization of plant roots. We evaluated the bacterial levels in plants colonized by communities of suppressor or non-suppressor strains in two independent assays. The first assay (Fig. 3*C*) was based on a community of five suppressors (SynCom5-S1) and five taxonomically matching non-suppressors (SynCom5-NS1). In this experiment, Arabidopsis Col-0 were grown axenically in Johnson medium for seven days (as described above for the SynCom35 experiments) and then transferred to new plates containing either SynCom5-S1 or SynCom5-NS1 (10^5 cells/mL each strain; prepared as described above for SynCom35). Plants transferred to medium without bacteria were used as controls. Plates were kept in a growth chamber under a 9-hour light/15-hour dark regime at 21°C day/18°C night. The root systems of approximately 5 plants per plate were harvested after 12 days for bacteria counting. The experiment was repeated 5 independent times, each with 3 biological replicates per condition. The

second assay (SI Appendix, Fig. S9) utilized a slightly modified community of five suppressors (SynCom5-S2) and five different, not taxonomically matching, non-suppressors (SynCom5-NS2). In this experiment, Arabidopsis Col-0 plants were germinated in 0.5xMS medium in the presence of SynComs ($5x10^5$ cells/mL each strain) or no bacteria and kept in a growth chamber under a 9-hour light/15-hour dark regime at 21°C day/18°C night. The root system of 3-6 seedlings were harvested after 16 days. Strains that comprise all SynComs are listed in SI Appendix, Dataset 1.

To isolate and quantify plant-colonizing bacteria, roots were harvested, rinsed and vortexed vigorously three times with sterile 10 mM $MgCl_2$ to remove agar particles and weakly-associated microbes. Plant material was weighed and then homogenized for 45s in a FastPrep-24 Classic Instrument (MP Biomedicals) in 2 mL tubes containing three 4 mm glass beads and 400 µL 10 mM $MgCl_2$. Homogenized samples were brought to 1 mL by adding 600 µL 10 mM $MgCl_2$, then submitted to a serial dilution. Four microliters of each dilution were inoculated onto LB agar plates (Luria-Bertani; 10g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl and 15 g/L agar). Colony forming units were counted after 1-3 days incubation at 28°C and used to determine the original bacterial abundance per gram of root tissue based on the serial dilution used for counting.

Evaluation of the Arabidopsis transcriptional response to SynComs of suppressors and non-suppressors. We evaluated the transcriptomes of plants grown in the presence of five suppressors (SynCom5-S1) or five taxonomically matching non-suppressors (SynCom5-NS1) (SI Appendix, Fig. S9*B*-*E*). This analysis is part of the experiment described above for the quantification of bacterial growth (shown in Fig. 3*C*). Briefly, seven-day-old Col-0 seedlings were transferred to plates containing no bacteria, SynCom5-S1 or SynCom5-NS1 (10^5 cells/mL each strain; SI Appendix, Dataset 1). No exogenous flg22 was added. Whole seedlings were harvested after twelve days and used for RNA extraction and sequencing as described above. The experiment included six biological replicates per conditions and was repeated two independent times (each with 3 biological replicates). The processing of RNA-seq reads, identification of differentially expressed genes and gene ontology enrichment analyses were performed as described for the other RNA-seq experiments described above. We compared the transcriptomes of plants grown with each SynCom against the 'no bacteria' control using the model: expression ~ SynCom + experiment, where "SynCom" refers to the presence of bacteria (no bacteria, SynCom5-S1 or SynCom5-NS1) and "experiment" refers to each of the two experimental batches. SI Appendix, Dataset 6 presents the metadata information of the experiment, the lists of differentially expressed genes and the complete results of the enrichment analyses.

Evaluation of the growth of commensals in the presence of SynComs. We measured if the root colonization capacity of commensal bacteria (*Ochrobactrum sp.* MF370 or *Pseudomonas*

viridiflava OTU5 p5.e6) is influenced by communities of suppressors (SynCom5-S2) or non-suppressors (SynCom5-NS2) (Fig. 4 and SI Appendix, Dataset 1). For this, approximately 500 Col-0 seeds were sterilized by shaking in 4.25% sodium hypochlorite for 4 min. The sodium hypochlorite was removed, and the seeds were washed four times in sterile water and then resuspended in 0.1% agar. Seeds were vernalized at 4° C for 2-3 days. Seeds were germinated on 0.5X MS medium containing either the suppressor SynCom (SynCom5-S2), the non-suppressor SynCom (SynCom5-NS2) or no bacteria. Plates were kept in a growth chamber under a 9-hour light/15-hour dark regime at 21^{o}C day/18^{o}C night. All bacteria strains were cultured from a single colony and grown in 3 mL of LB medium for 24 h at 28° C and 250 rpm. Cells were centrifuged for 6 min at 4,342 x g and washed three times with 10 mM $MgCl_2$. Media containing SynComs were prepared by embedding each strain at an OD_{600nm} of 0.0005. After two weeks, plants were flooded for 5 min with 7 mL of *Ochrobactrum sp*. MF370 or *Pseudomonas viridiflava* OTU5 strain p5.e6 containing pBBR1MCS-5 at an OD_{600nm} of 0.0001. Seedlings were then transferred to new 0.5X MS agar plates without bacteria. After 30 min, the flooding and transfer step was repeated with the strain in 10 mM $MgCl_2$ containing 0.005% silwet. After 48 h, roots from 3-6 seedlings per replicate were harvested, weighed, and ground for 45 s in a FastPrep-24 Classic Instrument (MP Biomedicals) in a 2 mL microcentrifuge tube containing 400 µL of 10 mM $MgCl_2$ and three 4 mm glass beads (Fisher, Cat. No. 11-312B). Homogenized samples were brought up to 1 mL by adding 600 µL 10 mM $MgCl_2$, then submitted to a serial dilution. Four microliters of each dilution were inoculated onto LB agar plates (Luria-Bertani; 10g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl and 15 g/L agar) containing 25 µg/mL gentamycin or 50 µg/mL kanamycin for antibiotic selection of OTU5 and MF370, respectively. Members of SynCom5-S2 and SynCom5-NS2 were susceptible to these antibiotics. Experiments were performed with three biological replicates and repeated three to four times. All experiments were analyzed together using an ANOVA that controlled for a significant batch effect using the model: log.CFU ~ treatment + batch. We identified truly significant groups with a Tukey test and an alpha of 0.05 using the HSD.test function from the agricolae package in R (31).

Type-3 secretion system deletion in MF79. Two unmarked Type-3 Secretion System (T3SS) deletion mutants of *Dyella japonicum* UNC79MFTsu3.2 (henceforth MF79) were constructed by two-step allelic exchange to test the involvement of the T3SS in suppression of the flg22 response observed by wild-type MF79. The first strain, MF79 Δ*sctUTSR*, has the genes for the inner membrane export apparatus of the T3SS deleted (*sctUTSR,* IMG database gene ID 2558297553-2558297556). The second strain, MF79 ΔT3SS, has a larger deletion comprising the full T3SS (*sctUTSRVCDJLN,* IMG database gene ID 2558297553-2558297569). Knockout vectors pJMC151 and pJMC152 for construction of strains MF79 Δ*sctUTSR* and MF79 ΔT3SS, respectively, were designed based on a genetic system developed for *Burkholderia* spp. and the suicide vector

pMo130 (32). All PCR steps were performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs) including the optional GC enhancer. The pMo130 vector backbone was amplified using primers JMC203-JMC204, cleaned-up, and treated with *DpnI* (New England Biolabs). 800-1000bp regions for homologous recombination flanking genes to be deleted were amplified (pJMC151/152 5': primers JMC403-JMC404, pJMC151 3': primers JMC405-JMC406, pJMC152 3': primers JMC407-JMC408). Vectors pJMC151 and pJMC152 were assembled using HiFi DNA Assembly Mastermix (New England Biolabs), transformed into NEB 5-alpha chemically competent *E. coli* (New England Biolabs) and selected on LB agar with 50 µg/mL kanamycin at 37°C. Plasmid DNA was isolated from clones using the ZR Plasmid Miniprep Classic Kit (Zymo Research) and sequenced by Sanger sequencing. Sequencing-confirmed pJMC151 and pJMC152 were transformed into biparental mating *E. coli* strain WM3064 and selected on LB agar plates containing 50 µg/mL kanamycin and 0.3 mM diaminopimelic acid (DAP) at 37°C.

Biparental mating was performed by growing the bacteria overnight: *Dyella japonicum* MF79 in 2xYT medium at 28°C and *E. coli* WM3064 containing either pJMC151 or pJMC152 in LB with 50 µg/mL kanamycin and 0.3 mM DAP at 37°C. Bacteria were washed three times with 2xYT medium and resuspended in 1/10 of the original volume and mixed in equal proportion donor to recipient and plated on LB agar plates containing 0.3 mM DAP and grown overnight at 28°C. The following day, exconjugants of MF79 were selected by streaking on LB agar plates containing 50 µg/mL kanamycin and lacking DAP and grown at 28°C. Several putative first crossover strains were grown from individual colonies in 2xYT medium containing 50 µg/mL kanamycin and screened using primers outside the regions of homologous recombination (JMC409, JMC410, JMC411) and primers on the pMo130 suicide vector backbone (JMC321, JMC322) to confirm insertion of knockout vectors in the MF79 chromosome at the appropriate location. Positive first crossovers were resolved by passaging the strains once on 2xYT lacking antibiotics, then passaging on media containing 10 g/L tryptone, 5 g/L yeast extract, 100 g/L sucrose, and 1mM IPTG, and finally plating on the same medium containing 1.5% agar. The resulting strains were screened for gene deletion using primers JMC409-JMC410 for strain MF79 Δ*sctUTSR* and primers JMC409-JMC411 for strain MF79 ΔT3SS. Putative positive strains were plate-purified by streaking two additional times on LB agar plates. The gene deletions in the final strains were confirmed again using PCR as described above. Purity of the final deletion strains were confirmed by performing PCR using forward primer JMC409 and reverse primer JMC426 inside the deleted gene *sctU*, which resulted in a 1.2 kb product for wild-type MF79 and no product in the deletion strains. Primers are listed in SI Appendix, Dataset 7 and the confirmatory DNA gel is shown in SI Appendix, Fig. S10*A*. Growth of the knockout mutants was compared to wild-type MF79 by growing the strains in Minimal Medium A (33) with 10 mM sucrose and 0.1% tryptone in an Infinite M200 Pro plate reader (Tecan) at 28°C while measuring OD_{600nm}.

MF79 Transposon insertion library screening. Arabidopsis *pCYP71A12::GUS* seeds were surface sterilized and stratified as described above for 24-48 hours in sterile distilled water, then arrayed in 96-well plates, with 1-3 seeds per well. 80 µL of MS medium was added and seedlings were grown at 16-hour light/8-hour dark regime at 21°C day/18°C night for 7 days.

The RB-TnSeq library for MF79 (34) was plated on LB agar plates containing 50 µg/mL kanamycin and grown at 28°C. Approximately 4500 individual colonies were placed into single wells of 96-deep well plates containing 2xYT medium with 50 µg/mL kanamycin and grown at 28°C. The bacteria were washed 3 times with 10 mM $MgCl_2$, resuspended in MS medium, and then 1 µL was added to each well containing Arabidopsis *pCYP71A12::GUS* seedlings in 80 µL MS. To store the arrayed bacteria, a final concentration of 20% glycerol was added to each well of washed bacteria and the library was frozen at -80ºC. In total, 49 plates were screened comprising approximately 4500 individual transposon insertion strains.

Approximately 16 hours after adding bacteria to the seedlings, 1 µM flg22 peptide was added to the wells. After 5 hours of flg22 treatment, seedlings were submitted to the GUS histochemical assay as described above. Plates were manually inspected for any wells containing plants with blue root tips indicating a failure to suppress the flg22 induction of *pCYP71A12*-driven GUS expression. All putatively positive strains were regrown by streaking from the 96-well plate glycerol stock of the arrayed library to LB plates containing 50 µg/mL kanamycin. Colonies were picked and grown 2xYT medium containing 50 µg/mL kanamycin, washed three times with 10mM $MgCl_2$, and assayed again using the GUS histochemical assay in 48-well plates with 100 nM flg22 peptide to confirm their ability to suppress the flg22 induced GUS expression. In this assay, bacteria were inoculated at an OD_{600nm} of 0.002. Six strains from the mutant library screened in this confirmatory assay were unable to suppress the flg22 induced GUS expression and were subjected to transposon insertion mapping. Growth of the selected Tn mutants were compared to wild-type MF79 by growing the strains in Minimal Medium A (33) with 10 mM sucrose and 0.1% tryptone in an Infinite M200 Pro plate reader (Tecan) at 28°C while measuring $OD600_{nm}$.

Transposon insertion mapping. An arbitrary PCR approach was used to map the transposon insertions in the six strains that lost the ability to suppress the plant response to flg22. PCR steps were performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs) using the optional GC enhancer according to the manufacturer instructions. A first PCR was performed using a transposon kanamycin marker specific primer (JMC525) and arbitrary primer ARB6 (35) with an annealing temperature of 35°C and extension time of 90s for 30 cycles. Following this, a second round of PCR was performed on the product using a primer nested inside the kanamycin resistance gene of the expected product (JMC567) and primer ARB2 (35) with annealing temperature of 71°C

and extension time of 90s for 30 cycles. The resulting PCR product was cleaned up and Sanger sequenced using primer JMC567. BLASTn was used to identify the MF79 genomic sequence adjacent to the transposon insertion and determine the Tn insertion location and, thus, the gene disrupted by the insertion. Primers are listed in SI Appendix, Dataset 7.

Suppression of the flg22 response by *D. japonica* MF79 cell-free supernatant. *D. japonica* MF79 wild-type and the *gspD::Tn* and *gspE::Tn* mutant strains were grown to stationary phase in 2xYT medium. Each strain was washed three times with 10 mM $MgCl_2$ and inoculated at OD_{600nm} of 0.02 into 1xMS medium containing 5 g/L sucrose and 0.1% (w/v) casamino acids (Bacto). Cultures were grown at 21°C with 250 rpm shaking for 18 hours. To remove the bacteria, cultures were centrifuged at 4000 x g for 15 min and the supernatant was filtered through as 0.2 µm filter to prepare cell-free supernatant. To evaluate the potential size of the active suppressive molecule(s), supernatant from the MF79 wild type was passed through a 10,000 MWCO Amicon Ultra-15 spin concentrator (Millipore). The 10 kDa flow-through is the sample that passed through the concentrator. The 10 kDa retentate sample was retained in the concentrator and brought to its original volume with 1xMS medium. Both the 10 kDa retentate and flow-through were sterile filtered prior to use.

pCYP71A12::GUS seedlings were grown in 24-well plates with 500 µL 1xMS medium with sucrose. This growth medium was removed from 7-day-old seedlings and replaced with the sterile cell-free supernatant, 10 kDa retentate, or 10 kDa flow-through samples from MF79 wild-type, *gspD::Tn*, and *gspE::Tn* or a blank medium control. Then, 1 µM flg22 was added to the wells for 5 hours, following which GUS histochemical staining was performed as described above.

Non-suppressor MF370 growth with MF79 10 kDa suppressive fraction. We evaluated if suppression of the flg22 response by *D. japonica* MF79 supernatant enhanced the growth of a non-suppressor commensal strain (Fig. 5*E*). Arabidopsis *pCYP71A12::GUS* seeds were surface sterilized as described above, arrayed in 24-well plates with 5-6 seeds per well, and grown in 500 µL 1xMS medium containing sucrose at 16-hour light/8-hour dark regime at 21°C day/18°C night for 7 days. Then, the wells and seedlings were washed 3 times with 750 µL 0.5x MS lacking sucrose.

Cell-free supernatant from *D. japonica* MF79 wild-type, *gspD::Tn*, and *gspE::Tn* were prepared as described above. To remove sucrose and casamino acids, the cell free supernatant for each was concentrated in a 10,000 MWCO concentrator and exchanged into 0.5xMS lacking sucrose by adding this medium and concentrating three times. The final exchanged 10 kDa retentate was brought to the original volume in 0.5xMS lacking sucrose. 500 µL of 10 kDa retentate exchanged

into 0.5xMS or medium alone as a control was added to the wells of the washed 7-day old seedlings and they were returned to the growth chamber. Approximately 18 hours later, 1 μM flg22 or mock treatment was added to wells for 5 hours. The non-suppressor strain *Ochrobactrum sp.* MF370 was grown to stationary phase in 2xYT medium containing 100 μg/mL ampicillin and washed three times in 10 mM $MgCl_2$. *Ochrobactrum sp.* MF370 was diluted in 0.5xMS and added to the wells at a final OD_{600nm} of 0.0002. Plates were returned to the growth chamber for 2 days. To assess the colonization by *Ochrobactrum sp.* MF370, plants were harvested as described above for the plant colonization competition assay and colony forming units of *Ochrobactrum sp.* MF370 per g plant fresh weight were determined on LB plates containing 100 μg/mL ampicillin.

Statistical information. All statistical tests were performed using R software. All boxplots display the median, 25th and 75th quantiles (box), and the smallest value, largest value, or 1.5*IQR below or above the 25th and 75th quantiles (whiskers). All statistical tests used an alpha value of 0.05 to assign significance unless otherwise stated. When possible, all data points are presented in the figures from all replications performed unless otherwise stated. ANOVAs were performed as a two-way ANOVA to control for experimental biases when analyzing multiple experiments at once. Any time multiple comparisons were made, we corrected our comparisons using a Tukey test. When many statistical tests were performed on the same data and many p-values were obtained we performed Benjamini and Hochberg p-value correction method (FDR) to control for false discovery rate. All statistical analysis methods are described in the figure legends. Code for statistical tests can be found at GitHub (https://github.com/ncolaian/MAE).

Data and code availability. All data generated in this project is publicly available. Raw sequences of 16S amplicon sequencing are available at the NCBI Short Read Archive (SRA) under the accession number PRJNA657936. The metadata file, read count table, relative abundance matrix and DESeq results of this experiment are provided in SI Appendix, Dataset 2. Raw sequences and read count matrices from RNA-seq experiments are available at the NCBI Gene Expression Omnibus (GEO) under the accession number GSE156426. The metadata file, edgeR results, list of heatmap cluster members and raw results from Gene Ontology (GO) enrichment analyses are provided in SI Appendix, Datasets 3 and 4. Scripts used for plotting and data analysis are available at GitHub (https://github.com/ncolaian/MAE).

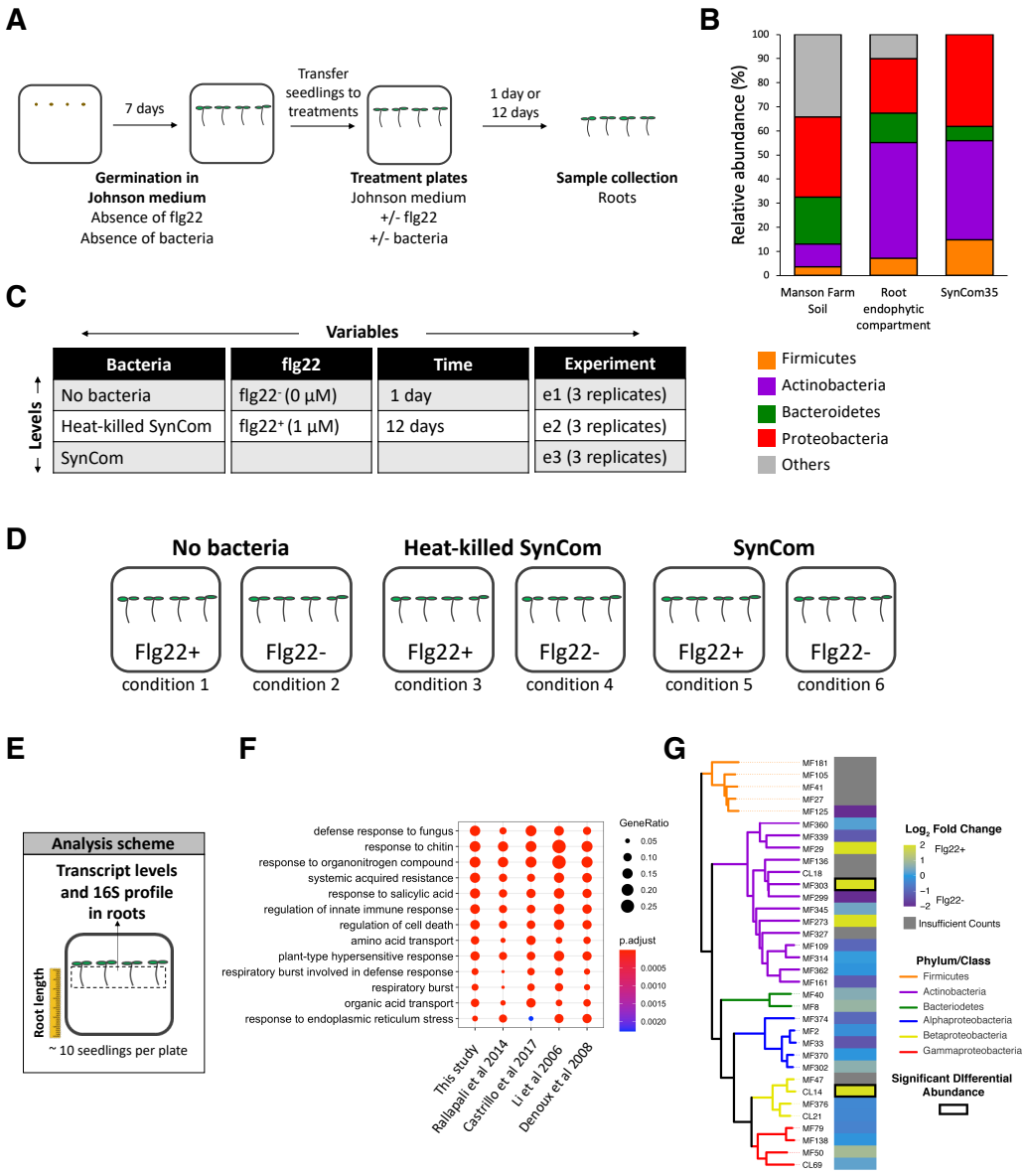


Fig. S1. Evaluating the Arabidopsis root transcriptome and bacterial community in response to the MAMP flg22. (A) Representation of the experimental system used to evaluate the effect of SynCom35 on the root response to flg22. Seeds of wild-type Arabidopsis (Col-0 ecotype) were germinated in Johnson medium under gnotobiotic conditions in the absence of flg22. After seven days, the resulting seedlings were transferred to plates (10 per plate) containing Johnson medium embedded with SynCom35 (10^5 cfu/mL) and/or 1 µM flg22. Roots were harvested for RNA extraction one or twelve days later. (B) Comparison of the bacterial taxonomic composition of Manson Farm soil, the root endophytic compartment of Arabidopsis plants grown in the same soil and of SynCom35. (C) Variables accounted for in the experiment: presence of bacteria, presence of flg22, length of the treatment and experimental batch. (D) A total of six different conditions were evaluated, corresponding to plants treated or not with 1 µM flg22 in the absence of bacteria, with heat-killed SynCom35 or with SynCom35. (E) Cartoon representation of the outputs measured: root length, root transcriptional program (RNA-seq) and root bacterial community profile (16S amplicon sequencing). (F) Seedlings treated with flg22 for 12 days activate defense-related genes in this experimental system. The figure shows biological processes (Gene Ontology) enriched among genes that were up-regulated by flg22 in control plants (in the absence of SynCom35). Four other independent studies are shown for comparison. Note that these other studies are based on plants of different ages, different tissues, different treatment times and flg22 concentration. See SI Appendix, Dataset S5 for details about each experiment. (G) Variation in relative abundance of each member that comprises SynCom35. Only two strains (MF303 and CL14) displayed significant differences in abundance in roots treated with flg22 relative to untreated roots after 12 days.

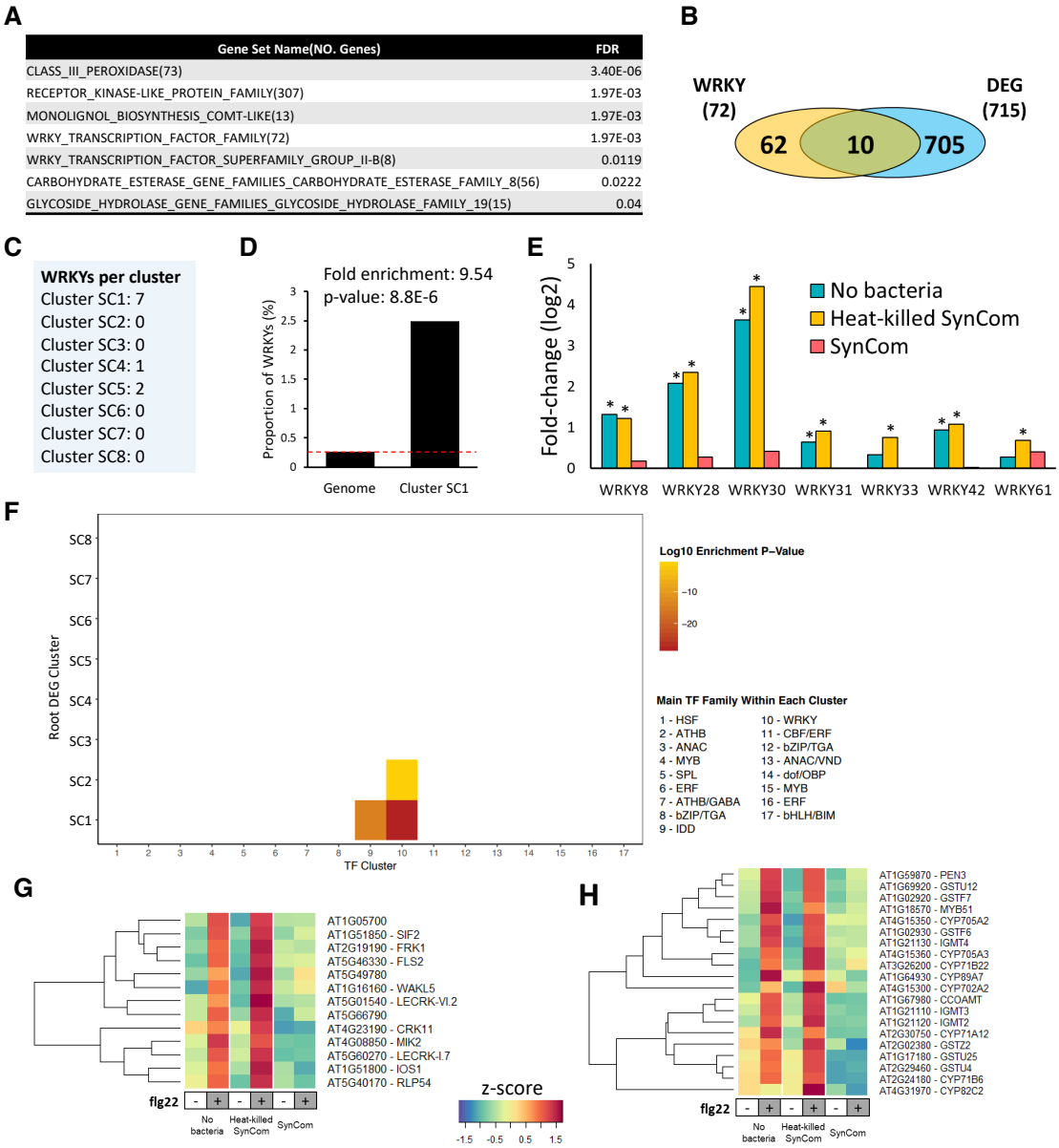


Fig. S2. SynCom35 suppresses key components of the plant immune system. (A) Gene families enriched among the group of flg22-responsive genes that were suppressed by SynCom35 (cluster SC1; Fig. 1*F*). Note the prevalence of defense-related families. (B) A total of 72 genes encoding WRKY transcription factors were annotated in the Arabidopsis Col-0 ecotype at The Arabidopsis Information Resource (TAIR). Among these, 10 were differentially expressed in the roots of plants treated with flg22 in our study. (C) Distribution of *WRKYs* per cluster (Fig. 1*F*). Eight of the ten flg22-responsive *WRKY* genes were up-regulated by flg22 (clusters SC1 and cluster SC4), and two were down-regulated (cluster SC5). Note that seven of the eight *WRKYs* that are activated by flg22 are included in cluster SC1 (i.e., they are suppressed by SynCom35). (D) Enrichment analysis showing that cluster SC1 contains approximately 10 times more *WRKY* genes than expected by chance (hypergeometric test; p-value = 8.8E-6). (E) Fold-changes (flg22 vs mock) of the seven *WRKY* genes included in cluster SC1. Note that these genes are not activated by flg22 when living SynCom35 is also present. Asterisks indicate genes that are differentially expressed based on the edgeR analysis. (F) Genes containing WRKY-binding motifs in their promoters (i.e., putative WRKY targets) are highly enriched in cluster SC1 and, to a lesser extent, in cluster SC2. (G) Expression profile of receptor kinase-like genes found in cluster SC1 (Fig. 1*F*). FRK1 is a widely used as a marker of MTI and FLS2 is the flg22 receptor. (H) Expression profile of genes involved in the synthesis of secondary metabolites found in cluster SC1 (Fig. 1*F*).

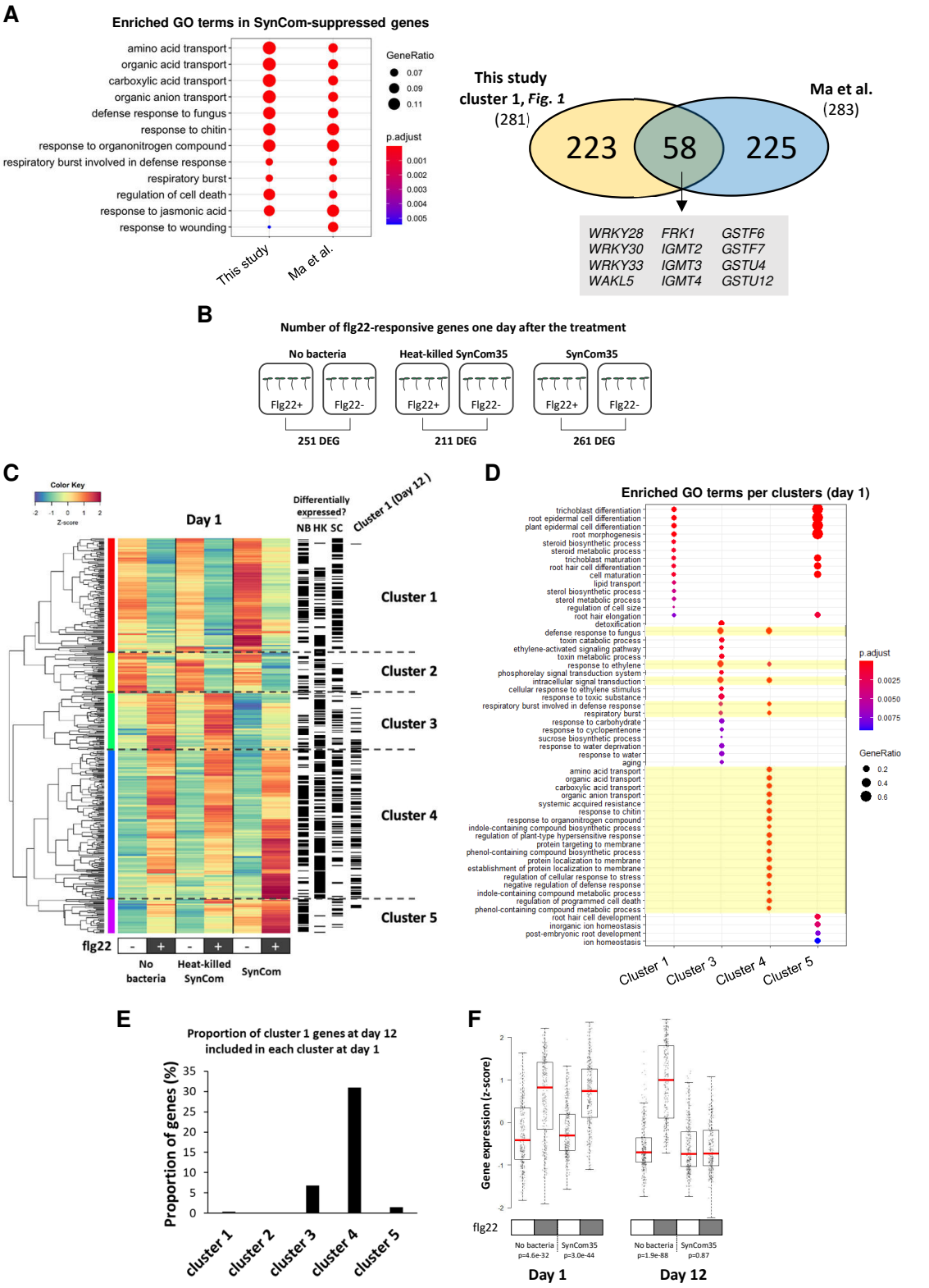


Fig. S3. SynCom35 actively suppresses an ongoing immune response that is also observed in other plant-microbiota interactions. (A) Suppression of defense-related genes is also observed with an independently derived SynCom (Ma et al., 2020). The left side shows that the genes suppressed by the SynCom reported by Ma et al. are enriched in the same biological processes identified in the genes suppressed by SynCom35 in this study (cluster SC1, Fig. 1*F*). p-values were adjusted with the FDR method.

The right side shows the overlap between the two sets of genes, highlighting 58 genes that are commonly suppressed by the SynComs in these independent studies (p=2.53E-58; hypergeometric test). Representative shared genes are shown in the grey box. The complete list of shared genes is shown in SI Appendix, Dataset S3. (B) Number of differentially expressed genes one day after the treatment of roots with flg22 (or mock) in the absence of bacteria, with heat-killed SynCom35 or with SynCom35 alive. (C) Hierarchical clustering of the 449 genes that responded to flg22 in at least one of the experimental conditions one day after treatment. In contrast to the results observed 12 days after treatment (Fig. 1*F*), SynCom35 does not cause a major interference in the root response to flg22 at day 1, indicating that SynCom35 interferes with an ongoing immune response. Indeed, flg22-responsive genes that are suppressed by SynCom35 at day 12 (Cluster 1, Fig. 1*F*) are activated at day 1 (black marks in the "Cluster 1" lane on the right of the heatmap). Genes that respond to flg22 significantly in each of the treatments at day 1 according to edgeR are indicated with black marks on the right in the "Differentially expressed" lanes (NB: no bacteria; HK: heat-killed SynCom35; SC: SynCom35). (D) Gene ontology enrichment analyses showing enriched biological processes in each of the five clusters. Cluster 4, which contains defense-related genes, is highlighted in yellow. No biological processes were enriched in Cluster 2. p-values were adjusted with the FDR method. (E) Proportion of SynCom35-suppressed genes at day 12 (Cluster 1; Fig. 1*F*) that is present in each cluster at day 1. Note that over 30% of the genes suppressed by SynCom35 at day 12 are included in Cluster 4 at day 1. (F) Expression profile of the 281 flg22-responsive genes that are suppressed by SynCom35 at day 12 (Cluster SC1; Fig. 1*F*). Note that the expression of these genes is increased by flg22 in the absence of bacteria at both time points (1 and 12 days after treatment). However, their expression is reduced in the presence of SynCom35 only 12 days after treatment. The p-values on the bottom refer to the comparison between groups (flg22 vs mock) for each condition using t-tests.

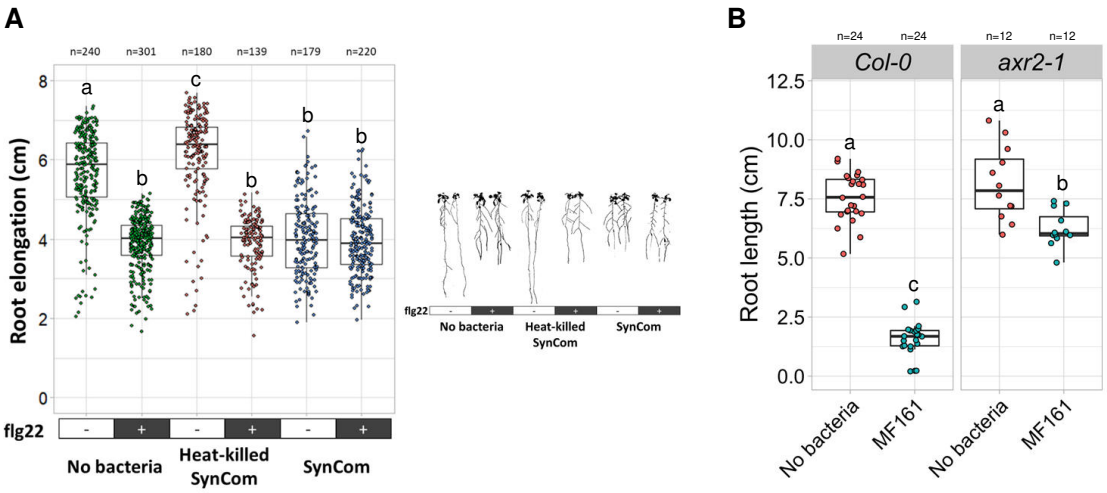


Fig. S4. SynCom35 effects on the Arabidopsis root morphology. (A) Root elongation measured in Arabidopsis seedlings grown in the absence of bacteria, with heat-killed SynCom35 and with SynCom35. For each of these three conditions, plants were exposed or not to 1 µM flg22, resulting in six treatments. Root growth inhibition (RGI) was induced by flg22 in plants grown in the absence of bacteria or with heat-killed SynCom35. Seedlings grown with SynCom35 displayed shorter roots even in the absence of flg22. Multiple comparisons were performed with ANOVA followed by a Tukey test (α = 0.05). Representative images of seedlings from each treatment are shown on the right. (B) The auxin-producing strain *Arthrobacter sp*. MF161 included in SynCom35 induces the RGI phenotype when in mono-association with Arabidopsis. This phenotype is significantly reduced in the auxin-resistant mutant line *axr2-1* suggesting that it is mediated by an auxin response. Multiple comparisons were performed with ANOVA followed by a Tukey test (α = 0.05).

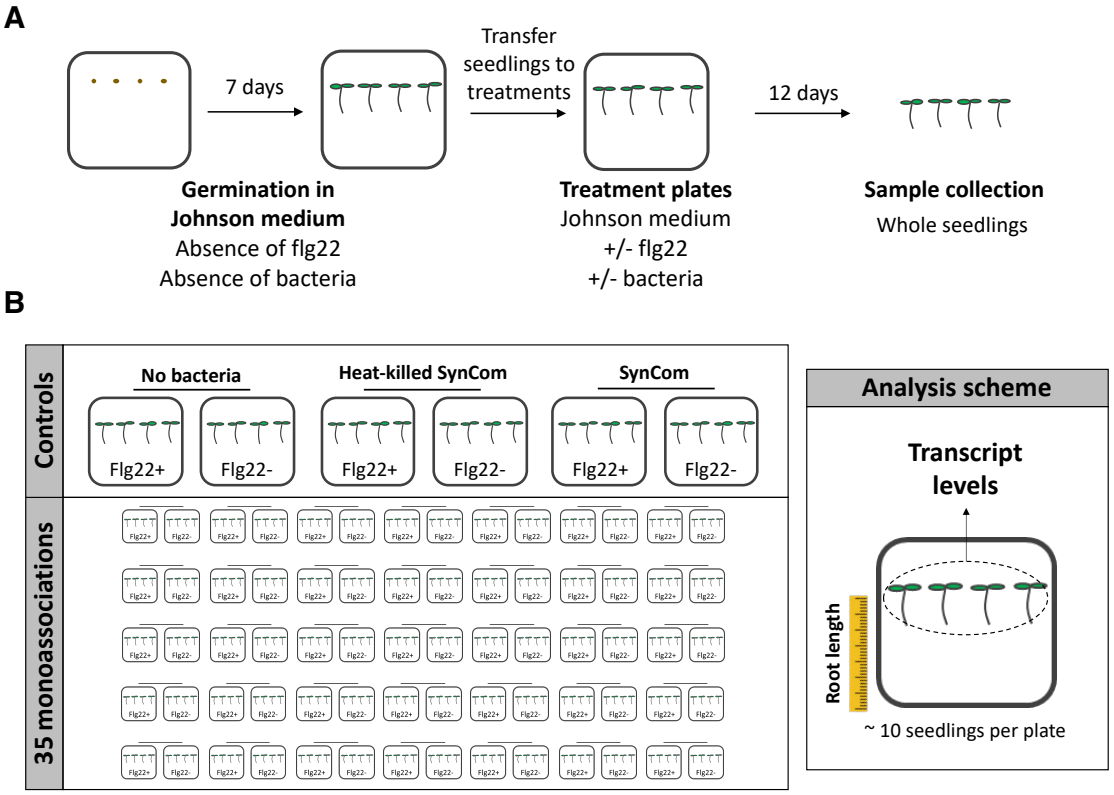


Fig. S5. Design of the mono-association experiment. (A) Representation of the experimental system used to evaluate the effect of each of the 35 members of SynCom35 on Arabidopsis seedlings. Seeds of wild-type Arabidopsis (Col-0 ecotype) were germinated in Johnson medium under gnotobiotic conditions in the absence of flg22. After seven days, the resulting seedlings were transferred to plates (10 per plate) containing Johnson medium embedded with a single bacterial strain (10^5 cfu/mL) in the presence or not of 1 µM flg22. Whole seedlings were harvested for RNA extraction twelve days later. Plants grown in the absence of bacteria and with heat-killed SynCom35 and SynCom35 were used as controls. (B) Representation of the conditions included in the experiment. A total of 76 different conditions were evaluated (35 strains and 3 controls x 2 treatments) in three independent experiments, each containing 3 biological replicates (total of 9 biological replicates per condition). The root length and the plant transcriptional program (RNA-seq) was assessed in each sample.

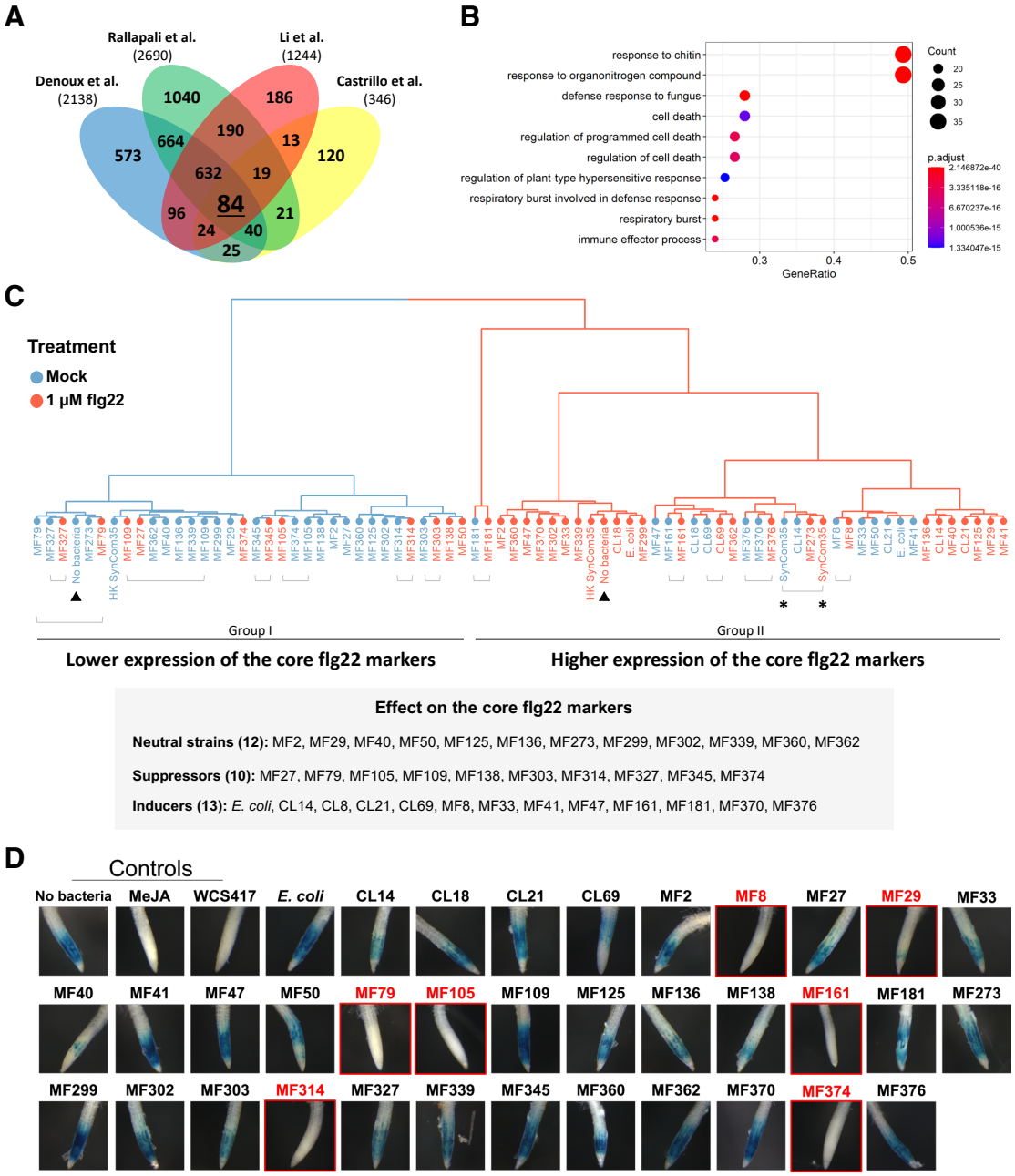


Fig. S6. Members of SynCom35 interfere with the expression of a robust set of 84 core flg22-responsive genes in a chronic treatment and also suppress the plant acute response to the elicitor. (A) Venn diagram with the overlap among genes found up-regulated by flg22 in four different gene expression experiments (1, 38–40). The intersection (84 genes; SI Appendix, Dataset 5) was used as a robust marker set of the flg22 response. (B) A Gene Ontology enrichment analysis showed that this core set of genes is highly enriched in immunity-related biological processes. p-values were adjusted with the FDR method. (C) Plants grown in mono-association with SynCom35 members were clustered based on the expression of the 84 markers genes of the flg22 response (core flg22 markers). Plants grown with SynCom35 (asterisks), heat-killed SynCom35 and in the absence of any bacteria ("no bacteria"; triangles) were also included in the experiment. For each condition, plants were exposed to 1 µM flg22 (red labels) or a mock treatment (blue labels). Note that samples are organized in two main clusters: the cluster on the left includes most of the mock-treated plants (blue) and displays lower expression of the core flg22 marker set. The cluster on the right includes most of the flg22-treated plants (red) and displays higher expression of the same marker set. In the presence of some strains, plants treated with flg22 (red) displayed low expression of the flg22 response marker set and, thus, grouped with mock-treated samples on the left. These were considered suppressor strains. In

other cases, bacteria triggered higher expression of the flg22 marker set even in the absence of exogenous MAMP (blue), grouping with flg22-treated plants on the right. These were considered inducer strains. Lines under the labels connect plants treated with a specific strain for which addition of flg22 had no major effect on the expression of the 84 marker genes (i.e., mock- and flg22-treatments grouped together). The chart below lists strains with neutral, suppressive or inducive effect on the flg22 markers. (D) At least seven members of SynCom35 (labeled in red) suppress the acute flg22 response in Arabidopsis roots. The reporter line carrying the *pCY71A12::GUS* construct (25) was grown for seven days in gnotobiotic conditions and then inoculated with each single strain individually at a final OD of 0.002. After 14h of co-cultivation, plants were treated with 100 nM flg22 for 5h. A blue signal in the elongation zone indicates that the *pCYP71A12* was activated by flg22. MeJA and *Pseudomonas simiae* WCS417 were previously shown to suppress MTI in the roots and were used as controls in the experiment (25). The experiment was performed at least two independent times for each strain. Each time included 2 repetitions (wells) consisting of approximately 10 seedlings each.

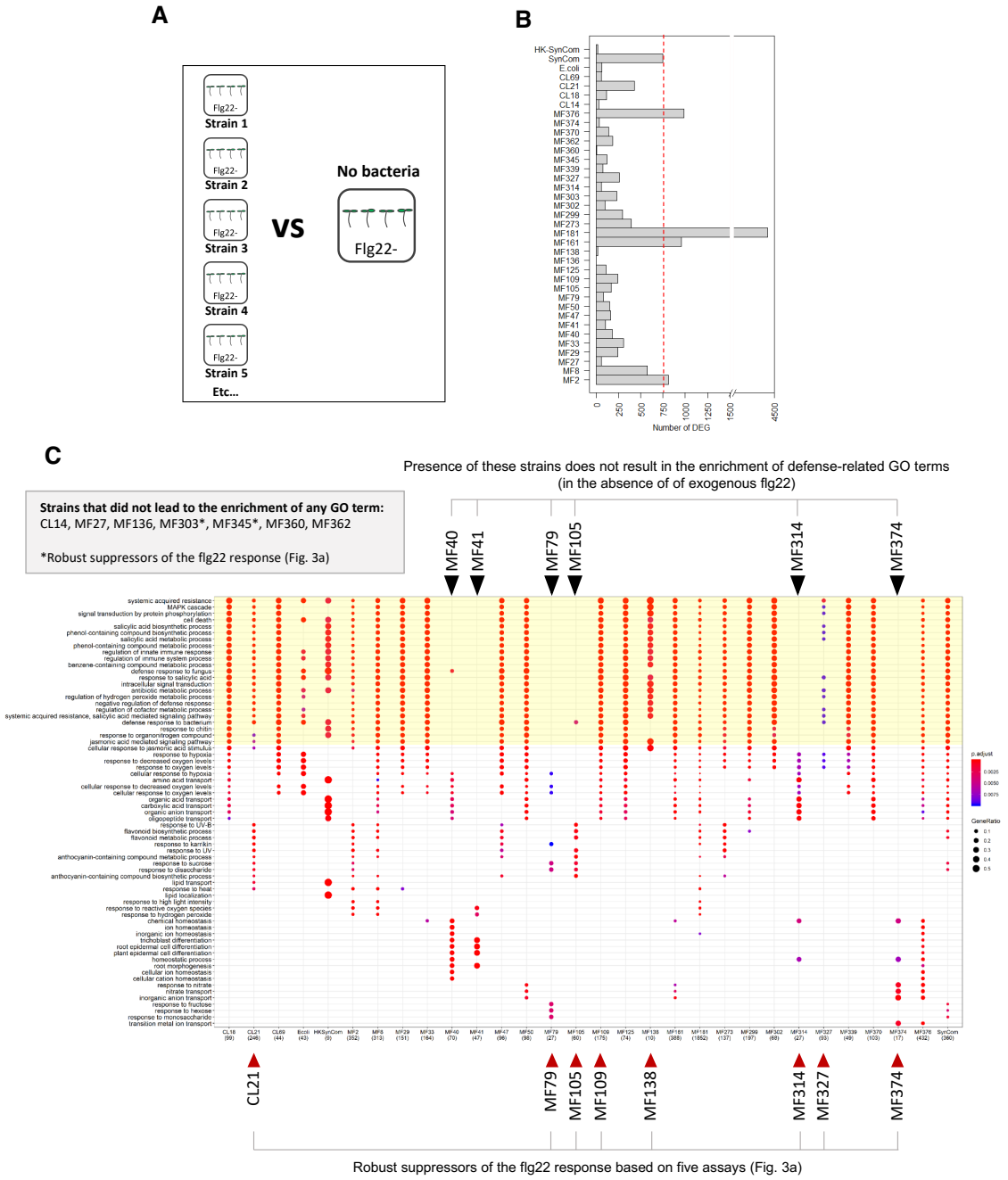


Fig. S7. Most, but not all, members of SynCom35 trigger the activation of plant defense genes to some extent when in mono-association. (A) Representation of the comparisons performed to evaluate the effect that each member of SynCom35 has on the plant. The transcriptome of plants grown with individual strains in the absence of flg22 was compared to the control condition (no addition of bacteria or flg22). (B) Number of differentially expressed genes identified in plants grown with each SynCom35 member. The red line depicts the number of differentially expressed genes identified in plants grown in the presence of SynCom35. (C) Gene Ontology (GO) enrichment analysis showing biological processes enriched in the sets of genes up-regulated by each strain. The area shaded in yellow corresponds to GO terms related to immune responses. Note that most strains trigger a transcriptional program that is enriched in defense-related genes. The six strains indicated with arrows on the top (MF40, MF41, MF79, MF105, MF314 and MF374) do not lead to the activation of the plant immune response but trigger the activation of other biological processes. Seven other strains (MF27, MF136, MF303, MF345, MF360, MF362 and CL14) did not result in the enrichment of GO processes and are not shown in the figure. Strains defined as robust suppressors (Fig. 3*A*) are labeled on the bottom. p-values were adjusted with the FDR method.

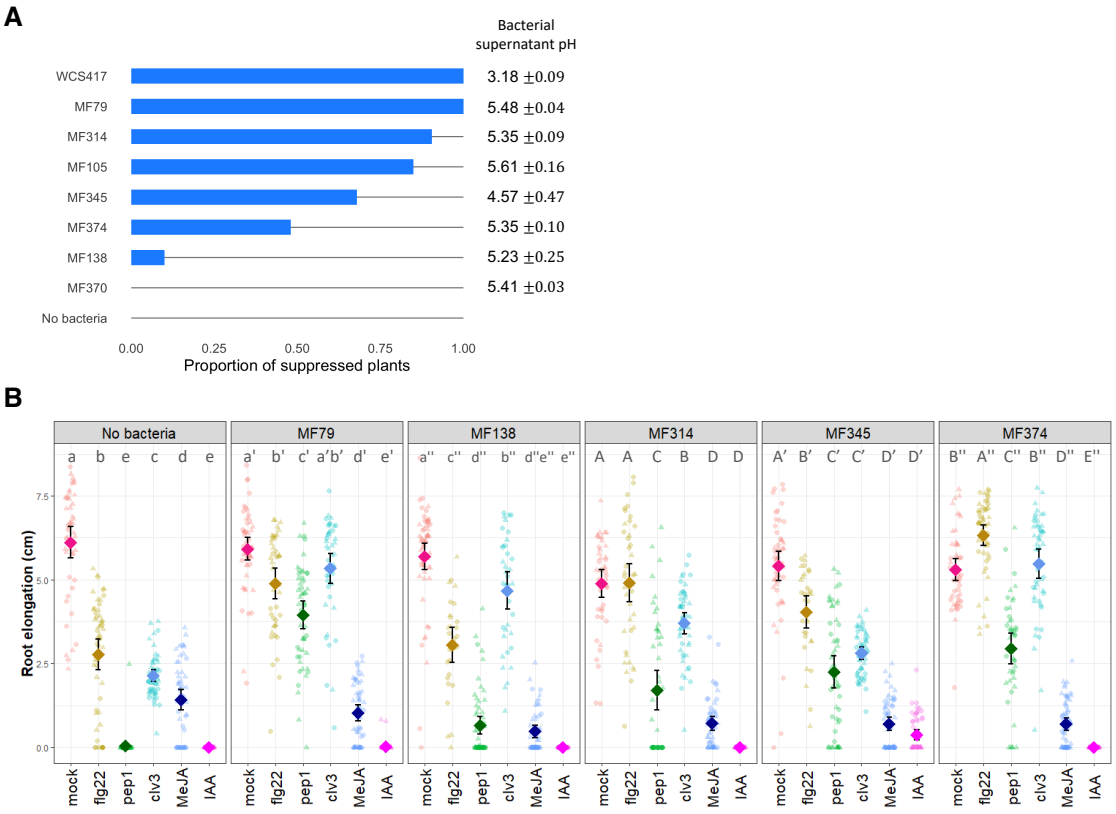


Fig. S8. Suppression of the plant immune system by commensal bacteria likely occurs through multiple different mechanisms. (A) Six robust suppressor strains were evaluated for their ability to prevent the root response to flg22 through acidification of the extracellular medium. The assay was performed as described by Yu et al. (2019) using cell-free bacterial supernatants and the Arabidopsis *pCYP71A12::GUS* reporter line. *Pseudomonas simiae* WCS417, which suppresses the flg22 response through medium acidification, was used as a positive control. The non-suppressor strain *Ochrobactrum sp.* MF370 was used as a negative control. The bars show the proportion of plants that did not respond to flg22 in each treatment (i.e., proportion of suppressed plants). A total of 20-25 seedlings were evaluated in the treatment of each suppressor strain. Note that the cell-free supernatant of most strains was sufficient to suppress the root response to flg22. Consistent with previous observations (Yu et al., 2019), *P. simiae* WCS417 lowered the extracellular pH to ~3. However, the supernatant of our suppressor strains remained at pH of 5-6, which is sufficient for a normal response to flg22 (Yu et al., 2019). This indicates that these strains do not employ the acidification of the extracellular medium as a strategy to suppress MTI. (B) Evaluation of the ability of suppressor strains to interfere with the root growth inhibition (RGI) phenotype induced by various molecules. The peptides flg22 (1 µM), pep1 (1 µM) and clv3 (1 µM) as well as the hormones methyl jasmonate (MeJA; 10 µM) and indole-3-acetic acid (IAA; 1 µM) were used. Seven-day-old Col-0 seedlings were transferred to Johnson medium embedded with bacteria (OD_{600nm}=0.0001) and the molecules of interest and grown at 9h light/15h dark (21°C light/18°C dark). Root elongation was Note that all molecules induce RGI in plants grown in the absence of bacteria (control treatment). measured after 12 days. Most strains suppress the RGI phenotype caused by peptides but not by the two hormones. Diamonds represent means with two times standard error. Multiple comparisons within each 'bacteria' condition were performed with ANOVA followed by a Tukey test (letters on the top).

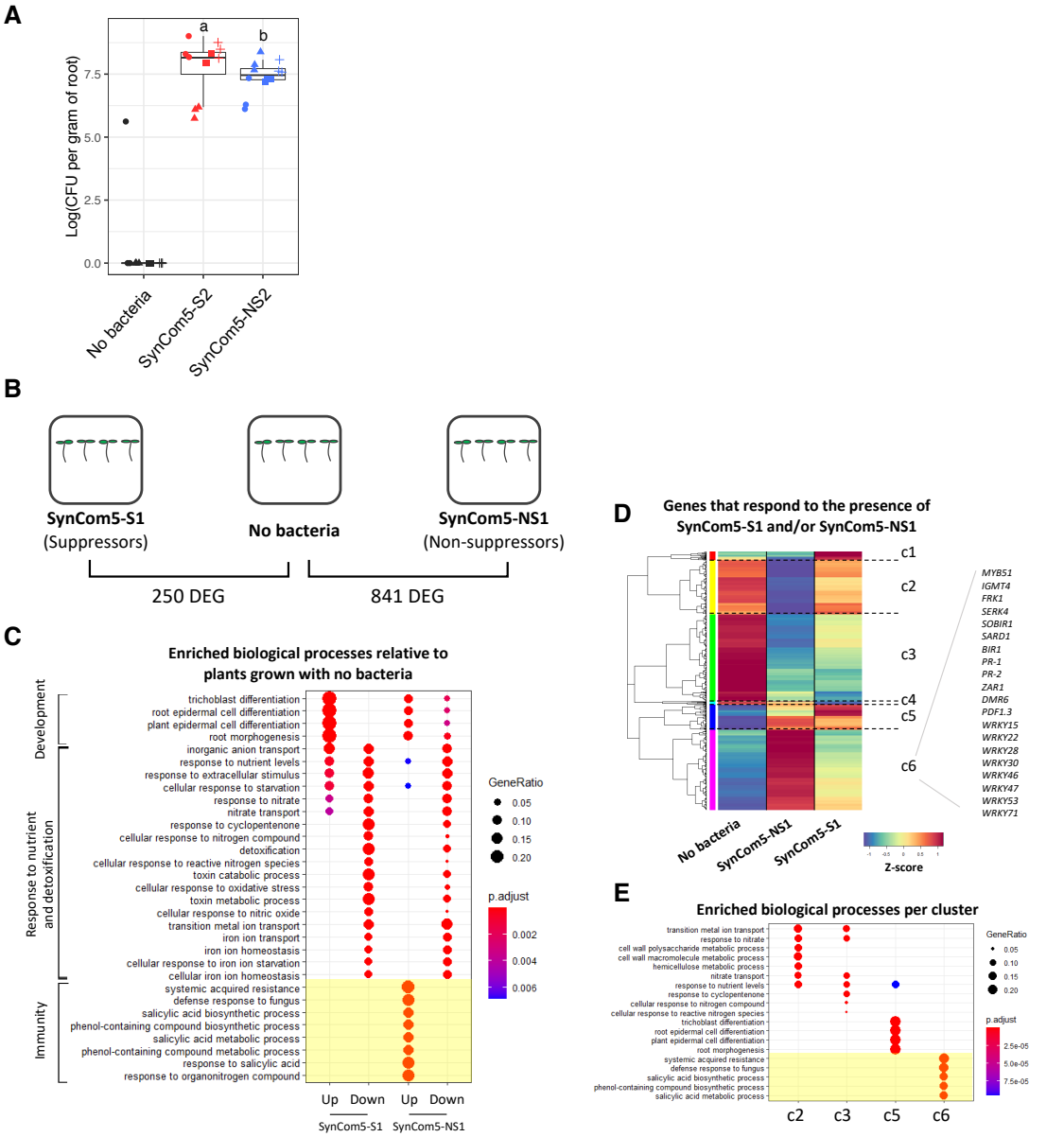
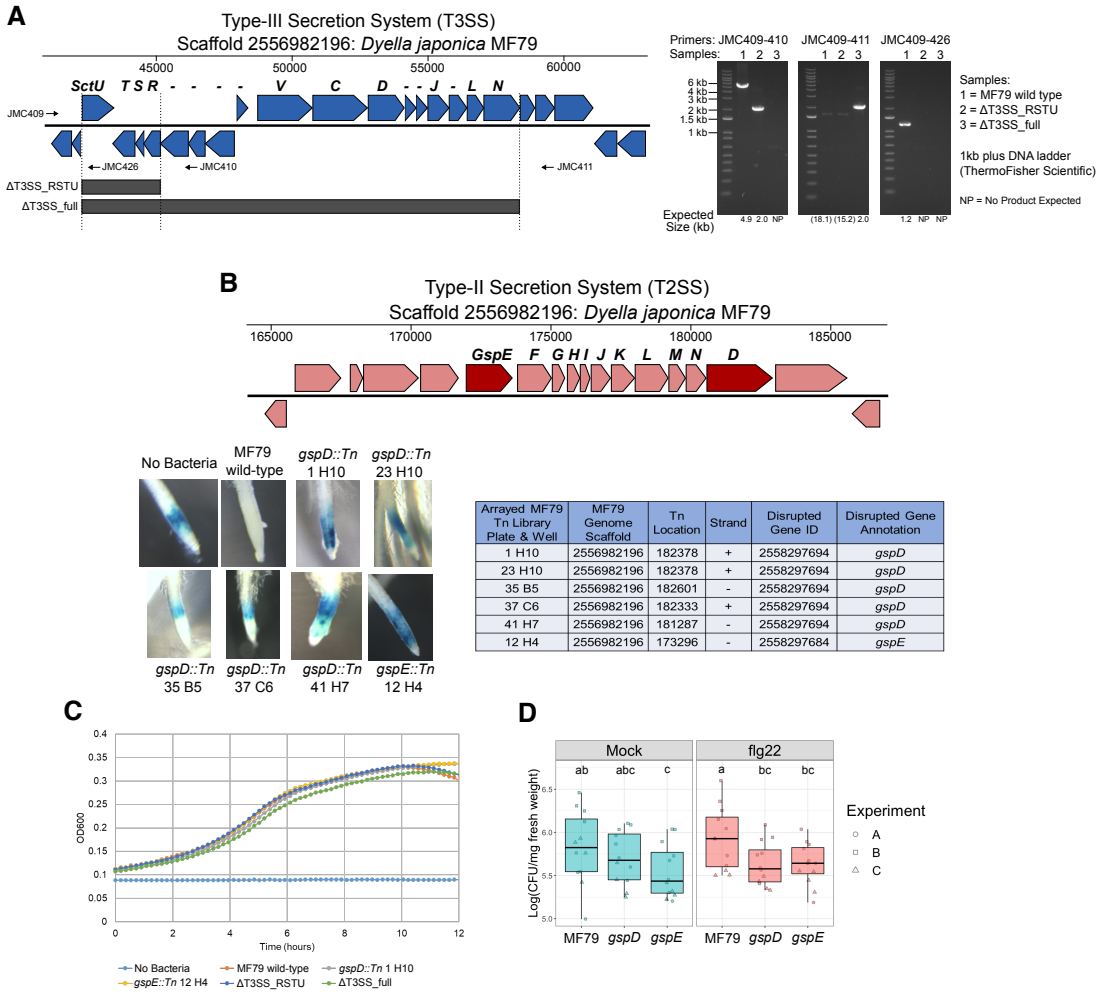


Fig. S9. A community of non-suppressor bacteria triggers defense responses in Arabidopsis. (A) Evaluation of bacteria colonization of Arabidopsis roots. A SynCom made of five suppressors (SynCom5-S2) grows to higher levels in the roots than a SynCom made of five non-suppressors (SynCom5-NS2). The experiment was performed four independent times, each with three biological replicates per condition (n=12). Different symbols represent different experimental repetitions. This experiment complements the results presented in Fig. 3*C* and was performed in a different experimental system (see methods) using different SynComs. (B) Transcriptional analysis of plants grown in the absence of bacteria or with communities of suppressors (SynCom5-S1) and non-suppressor (SynCom5-NS1) strains. The number of differentially expressed genes (DEG) is shown under each comparison. No exogenous flg22 was added. (C) Gene ontology enrichment analysis showing biological processes enriched in the sets of genes up- or down-regulated in plants grown with SynCom5-S1 or SynCom5-NS1 relative to axenic conditions. Note the absence of immunity-related terms in plants colonized with SynCom5-S1 (shaded in yellow). (D) Hierarchical clustering of 925 genes that responded to SynCom5-S1, SynCom5-NS1 or both. Six clusters were defined. Example of defense-related genes are shown on the right. (E) Gene ontology enrichment analysis of each cluster defined in panel (D). Note that immunity-related terms are enriched in cluster c6, which includes genes more expressed in the presence of SynCom5-NS1. Clusters c1 and c4 do not have enriched terms. SI Appendix, Dataset 6 presents metadata information of the experiment, the lists of differentially expressed genes and the complete results of the enrichment analyses.



Arrayed MF79 Tn Library Plate & Well	MF79 Genome Scaffold	Tn Location	Strand	Disrupted Gene ID	Disrupted Gene Annotation
1 H10	2556982196	182378	+	2558297694	*gspD*
23 H10	2556982196	182378	+	2558297694	*gspD*
35 B5	2556982196	182601	-	2558297694	*gspD*
37 C6	2556982196	182333	+	2558297694	*gspD*
41 H7	2556982196	181287	-	2558297694	*gspD*
12 H4	2556982196	173296	-	2558297684	*gspE*

Fig. S10. The T2SS and not the T3SS of *Dyella japonica* MF79 is required for suppression of the flg22 response. (A) Gene map for the *Dyella japonica* MF79 type-III secretion system (T3SS, *Sct* genes). DNA coordinates on Scaffold 2556982196 are identified above the map. Genes deleted in the MF79 ΔT3SS_RSTU and MF79 ΔT3SS_full mutants are marked below the map. DNA gels confirming gene knockouts and strain purity in these mutants are shown to the right. Expected product sizes are shown below the gels. The location of primers used to screen these mutants are marked in the T3SS map. Primers sequences are available in SI Appendix, Dataset 7. (B) Gene map for the *Dyella japonica* MF79 type-II secretion system (T2SS, *Gsp* genes). Six strains with loss of flg22-response suppression were identified from the arrayed MF79 transposon insertion library. Transposon insertion locations were identified by arbitrary PCR. At left, GUS expression assay demonstrating loss of flg22-response suppression by the 6 strains (5 insertions in *GspD*, 1 insertion in *GspE*). Genes disrupted in the MF79 *gspE::Tn* and *gspD::Tn* strains are highlighted in dark red in the map of the T2SS. (C) Growth of the MF79 wild-type, T3SS deletion, and T2SS transposon insertion mutant strains in Minimal Medium A with 10 mM sucrose and 0.1% tryptone. (D) The ability to suppress the plant immune system correlates with enhanced colonization of Arabidopsis roots. Although *gspD* and *gspE* mutants did not display impaired growth *in vitro*, they are poorer colonizers of Arabidopsis roots than the wild-type strain. Plants were cultivated with each strain in mono-colonization for 5 days in the presence (red) or not (blue) of 1 μM flg22.

Dataset S1 (separate file). Bacterial strains used in this study.
Spreadsheet with detailed information about the bacteria used in the experiments and the composition of each SynCom constructed. The 16S sequence of each strain is also provided.

Dataset S2 (separate file). 16S amplicon sequencing.
Spreadsheet with information about the sequencing of SynCom 16S amplicons. The file contains the metadata of the experiment, raw counts, relative abundance and the results of the differential abundance analysis.

Dataset S3 (separate file). SynCom35 RNA-seq experiment.
Spreadsheet with information about the RNA-seq experiment of plants exposed to SynCom35 and flg22. The file contains the metadata of the experiment, the results of differential expression analyses, a list of members in each heatmap cluster and Gene Ontology enrichment analyses.

Dataset S4 (separate file). Mono-associations RNA-seq experiment.
Spreadsheet with information about the RNA-seq experiment of plants exposed to individual bacteria and flg22. The file contains the metadata of the experiment, the results of differential expression analyses, a list of members in each heatmap cluster and Gene Ontology enrichment analyses.

Dataset S5 (separate file). Marker genes of the flg22 response.
Spreadsheet with a set of literature-curated genes that were up-regulated by flg22 in four independent studies.

Dataset S6 (separate file). SynCom5-S1 and SynCom5-NS1 RNA-seq experiment.
Spreadsheet with information about the RNA-seq experiment of plants exposed to SynCom5-S1 (suppressors) and SynCom5-NS1 (non-suppressors). The file contains the metadata of the experiment, the results of differential expression analyses, a list of members in each heatmap cluster and Gene Ontology enrichment analyses.

Dataset S7 (separate file). Genetic analyses of *Dyella japonica* MF79
Spreadsheet with primers used for the generation of T3SS knockout strains and for transposon mapping. Detailed information about the non-suppressing mutants identified in the RB-TnSeq library is also provided.

SI References

1. G. Castrillo, *et al.*, Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513–518 (2017).

2. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

3. H. Wickham, *ggplot2: elegant graphics for data analysis* (springer, 2016).

4. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

5. S. S. Abby, B. Néron, H. Ménager, M. Touchon, E. P. C. Rocha, MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS ONE* **9**, e110726 (2014).

6. J. Logemann, J. Schell, L. Willmitzer, Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* **163**, 16–20 (1987).

7. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

8. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

9. Y. Liao, G. K. Smyth, W. Shi, The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).

10. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

11. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

12. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

13. M. D. Robinson, A. Oshlack, A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

14. Y. Benjamini, Y. Hochberg, 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 (1995).

15. H. Heberle, G. V. Meirelles, F. R. da Silva, G. P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169 (2015).

16. T. Wei, V. Simko, *R package "corrplot": Visualization of a Correlation Matrix* (2017).

17. J. Oksanen, *et al.*, *vegan: Community Ecology Package* (2019).

18. G. R. Warnes, *et al.*, *gplots: Various R Programming Tools for Plotting Data* (2020).

19. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

20. X. Yi, Z. Du, Z. Su, PlantGSEA: a gene set enrichment analysis toolkit for plant community. *Nucleic Acids Res.* **41**, W98-103 (2013).

21. S. Herrera-Paredes, *AMOR: Abundance Matrix Operations in R* (2018).

22. B. J. Callahan, *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

23. O. M. Finkel, *et al.*, A single bacterial genus maintains root growth in a complex microbiome. *Nature* (2020) https:/doi.org/10.1038/s41586-020-2778-7.

24. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

25. Y. A. Millet, *et al.*, Innate immune responses activated in Arabidopsis roots by microbe-associated molecular patterns. *Plant Cell* **22**, 973–990 (2010).

26. K. Yu, *et al.*, Rhizosphere-Associated *Pseudomonas* Suppress Local Root Immune Responses by Gluconic Acid-Mediated Lowering of Environmental pH. *Current Biology* **29**, 3913-3920.e4 (2019).

27. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

28. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

29. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

30. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).

31. F. de Mendiburu, *agricolae: Statistical Procedures for Agricultural Research* (2020).

32. M. A. Hamad, S. L. Zajdowicz, R. K. Holmes, M. I. Voskuil, An allelic exchange system for compliant genetic manipulation of the select agents *Burkholderia pseudomallei* and *Burkholderia mallei*. *Gene* **430**, 123–131 (2009).

33. F. Ausubel, *et al.*, *Current Protocols in Molecular Biology* (Wiley Interscience, New York, 1996).

34. M. N. Price, *et al.*, Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* **557**, 503–509 (2018).

35. G. A. O'Toole, *et al.*, Genetic approaches to study of biofilms. *Meth. Enzymol.* **310**, 91–109 (1999).

36. D. S. Lundberg, *et al.*, Defining the core Arabidopsis thaliana root microbiome. *Nature* **488**, 86–90 (2012).

37. K.-W. Ma, *et al.*, Coordination of microbe-host homeostasis via a crosstalk with plant innate immunity. *Research Square (pre-print)* (2020).

38. C. Denoux, *et al.*, Activation of defense response pathways by OGs and Flg22 elicitors in Arabidopsis seedlings. *Mol Plant* **1**, 423–445 (2008).

39. G. Rallapalli, *et al.*, EXPRSS: an Illumina based high-throughput expression-profiling method to reveal transcriptional dynamics. *BMC Genomics* **15**, 341 (2014).

40. B. Li, *et al.*, Phosphorylation of trihelix transcriptional repressor ASR3 by MAP KINASE4 negatively regulates Arabidopsis immunity. *Plant Cell* **27**, 839–856 (2015).