Supplementary Note

Supplementary Figures 1-13

Mapping phyllosphere microbiota interactions *in planta* to establish genotype - phenotype relationships

Martin Schäfer, Christine M. Vogel, Miriam Bortfeld-Miller, Maximilian Mittelviefhaus, Julia A. Vorholt*

Institute of Microbiology, ETH Zurich, Zurich, Switzerland

*corresponding author: jvorholt@ethz.ch

Supplementary note

Focal community single strain drop-out and replacement experiments

To test how robust the focal community was in terms of community assembly, we conducted an experiment with single strain drop-outs from the inocula (Supplementary Fig. 9a) and assessed the effect on the composition of the remaining strains (Supplementary Fig. 10a). Only one interaction was identified. Removal of *Rhizobium* Leaf371 - the second most abundant strain - led to an increased abundance of *Aeromicrobium* Leaf272, suggesting a negative interaction between the two strains (log2 fold-change (log2FC): 2.2, p_{adj}: 0.0005). This interaction was reproducible when quantified by CFU counts instead of 16S rRNA gene sequencing and could be recreated in binary strain inoculations without community context (Supplementary Fig. 10b). The occurrence of only one interaction was not surprising, given that the focal strains captured the diversity of the *At*-LSPHERE strain collection at the phylogenetic but also the functional level[1] and hence only partially overlapping niches would be expected.

To assess the representativeness of the focal strains, we also replaced each strain in the focal community one at a time with a close relative within the culture collection (Supplementary Fig. 9b). These strain exchanges did not lead to emerging interactions with the other 14 strains (Supplementary Fig. 11a) demonstrating functional similarity of these strains. For the genera *Methylobacterium*, *Rhizobium* and *Sphingomonas* that are represented by many isolates in the *At*-LSPHERE collection, we additionally tested a more distant relative within the genus (Supplementary Fig. 11b, for phylogenetic tree see Supplementary Fig 1). For two of these, we found interactions indicating diverging strain properties.

Identification of active molecules in the culture supernatant of Leaf245

To identify the active molecule, we fractionated cell suspension supernatants of Leaf245 by size-exclusion chromatography and analyzed the active fractions by SDS-PAGE (Fig. 3c,d). These fractions from two independent experiments showed two prominent bands. Excision of these from the gel and subsequent protein identification led to the identification of an annotated serine protease (ASF05_09325) and a putative M23 family endopeptidase (ASF05_00205), both of which could potentially explain the observed bacteriolytic activity. These two proteins, as well as three other candidates identified in pooled active fractions from the size exclusion chromatography (Extended Data Fig. 5a, Supplementary Table 6), were selected for further analysis. We tested codon-optimized versions of all five proteins for heterologous expression in *E. coli*, and more specifically whether cell lysates caused Leaf374 inhibition and cell lysis (Extended Data Fig. 5b). All constructs showed inhibition of Leaf374, but strong inhibition was observed for only three of them (ASF05_00205, ASF05_05375 and ASF05_12180). All three proteins inhibited Leaf 374 also when purified (Fig. 3e) and both ASF05_00205 as well as ASF05_12180 resulted in lysis of the target cells (Extended Data Fig. 5c). These results suggest that a cocktail of inhibitory and lytic proteins is produced by Leaf245.

Description of regulatory EMS mutants and analysis of their culture supernatants

Because E_60 was the only mutant in an effector candidate, we wondered whether there were other predicted regulatory proteins besides ASF05_00210 mutated in the EMS screen (Supplementary Table 7). Two clones stood out because they showed complete loss of inhibition and had very few mutations. One clone (E_12) was affected in the *gsiA* gene, encoding a glutathione import ATP-binding protein[2], the other (E_48) showed a mutation in a gene with predicted histidine kinase domains that may be part of a two-component signal transduction system (ASF05_16175, SMART domains[3] HisKA and HATPase_c). To assess if potential regulatory mutants were impaired in effector production, we fractionated cell free supernatants of wild type and regulatory mutant clones and examined them by SDS-PAGE (Fig. 4c, Supplementary Fig. 12). While only the band for ASF05_00205 was absent in the ASF05_00210 mutants (clones E_13 and E_34), the bands of both ASF05_00205 and ASF05_09325 were absent in the *gsiA* (E_12) and histidine kinase (E_48) mutants (Fig. 4c). Congruent results were obtained when the proteins identified by shotgun proteomics were compared (Supplementary Fig. 13), demonstrating that ASF05_00205 endopeptidase secretion was linked to the presence of the ASF05_00210 transcriptional activator.

Attempts to transform Leaf245 by plasmid based methods

We attempted to generate a transposon library of Leaf245 with transposons himar1 (pAK415[4]) and Tn5 (pAG408[5]). Plasmid was transferred by bi-parental mating with *E. coli*-S17 λpir (24 h) on minimal medium without carbon source, supplemented with 500 µM vanillate for pAK415 plasmid. Cells were suspended in 0.9% (w/v) NaCl solution and dilutions were plated on R-2A+M supplemented with colistin (10 µg mL^{-1}) and kanamycin (5 µg mL^{-1}). No transformants were obtained. In addition, we tried marker exchange mutagenesis by homologous recombination to introduce an apramycin cassette, which was not successful. Also tri-parental mating with a dam-/dcm- *E.coli* donor strain did not yield any apramycin resistant clones.

References

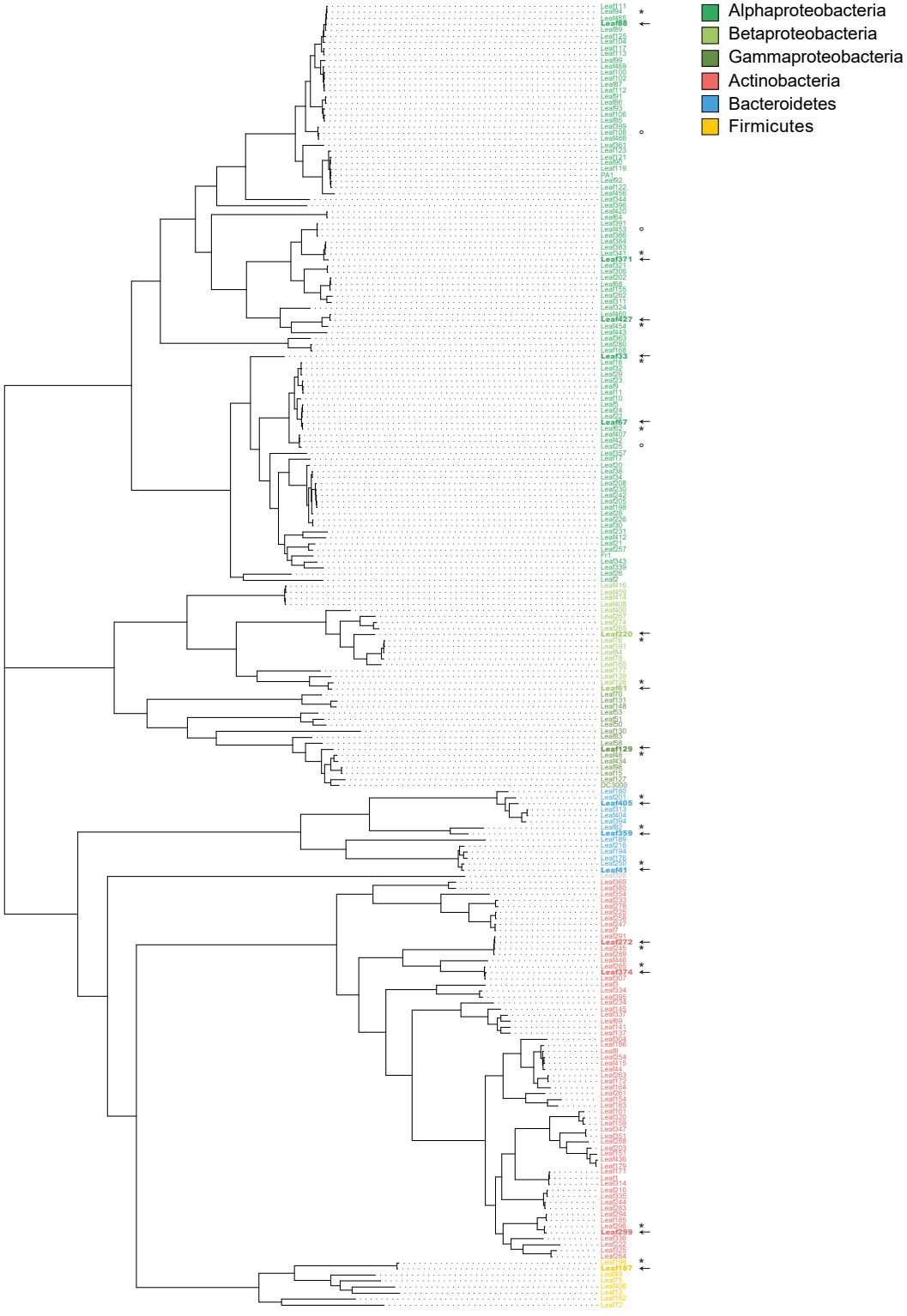
1 Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E. *et al.* Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528:** 364-369 (2015).

2 Suzuki, H., Koyanagi, T., Izuka, S., Onishi, A. & Kumagai, H. The yliA, -B, -C, and -D genes of *Escherichia coli* K-12 encode a novel glutathione importer with an ATP-binding cassette. *J Bacteriol* **187:** 5861-5867 (2005).

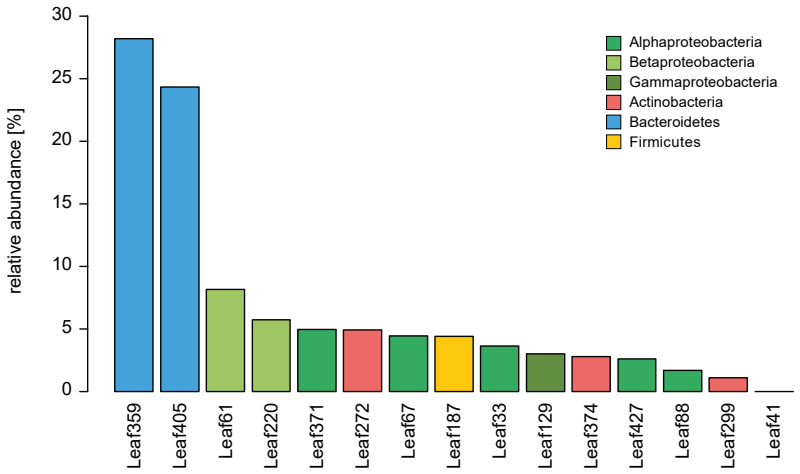
3 Letunic, I., Khedkar, S. & Bork, P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res* **49:** 458-460 (2020).

4 Kaczmarczyk, A., Hochstrasser, R., Vorholt, J. A. & Francez-Charlot, A. Complex two-component signaling regulates the general stress response in Alphaproteobacteria. *P Natl Acad Sci USA* **111:** 5196-5204 (2014).

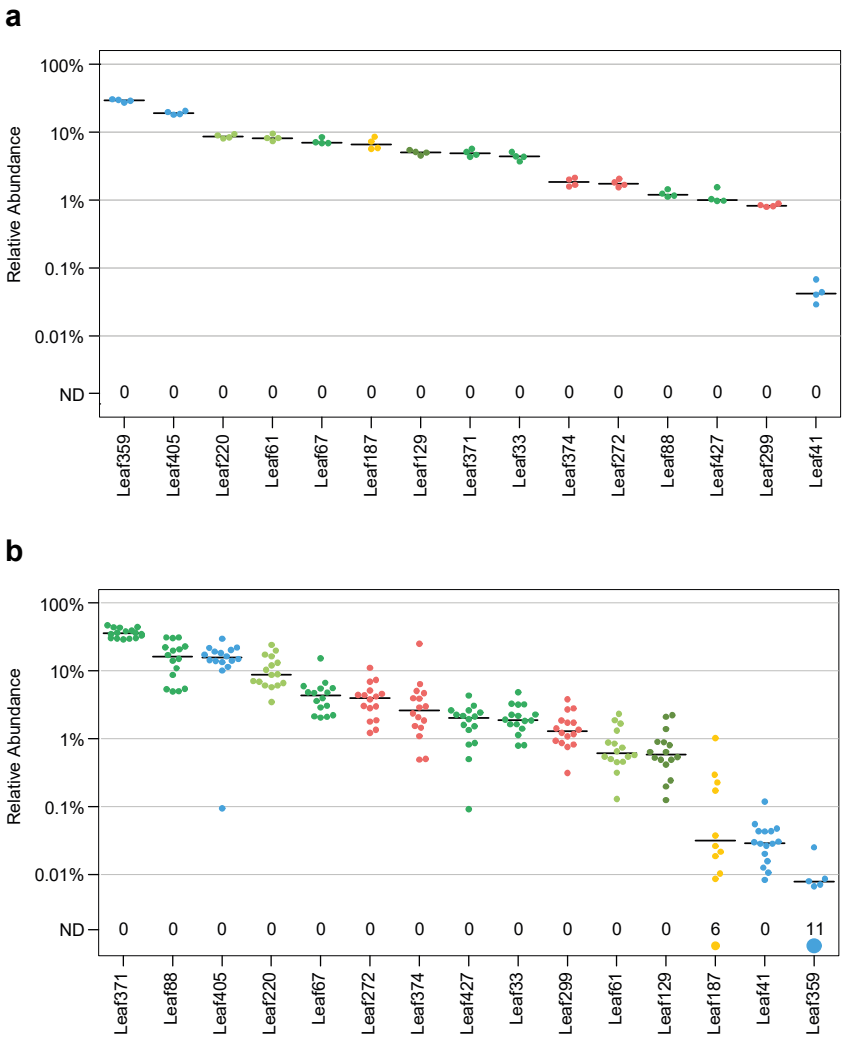
5 Suarez, A., Guttler, A., Stratz, M., Staendner, L. H., Timmis, K. N. *et al.* Green fluorescent protein-based reporter systems for genetic analysis of bacteria including monocopy applications. *Gene* **196:** 69-74 (1997).



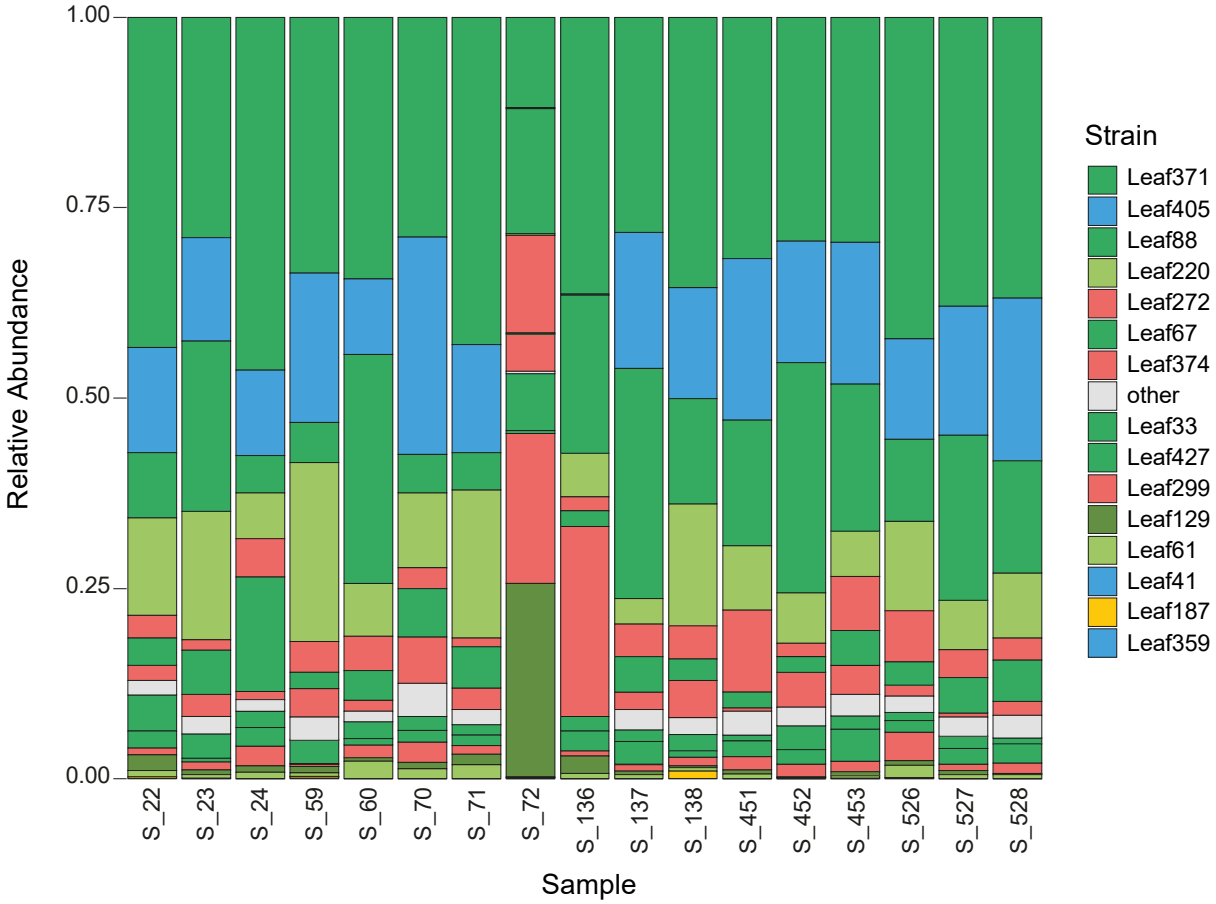
Supplementary Fig. 1: Strains included in this study. Phylogenetic tree based on full-length 16S rRNA gene sequences of *At*-LSPHERE strains included in this study as well as additional model organisms. The color indicates the phylum and in case of Proteobacteria the class. Focal community strains are highlighted with bold letters and indicated with an arrow. Additional cultivated isolates used in replacement experiments are indicated with an * for a close relative, and a ◦ for a distant relative, of a replacement focal strain (see Supplementary Note and Supplementary Figure 11).



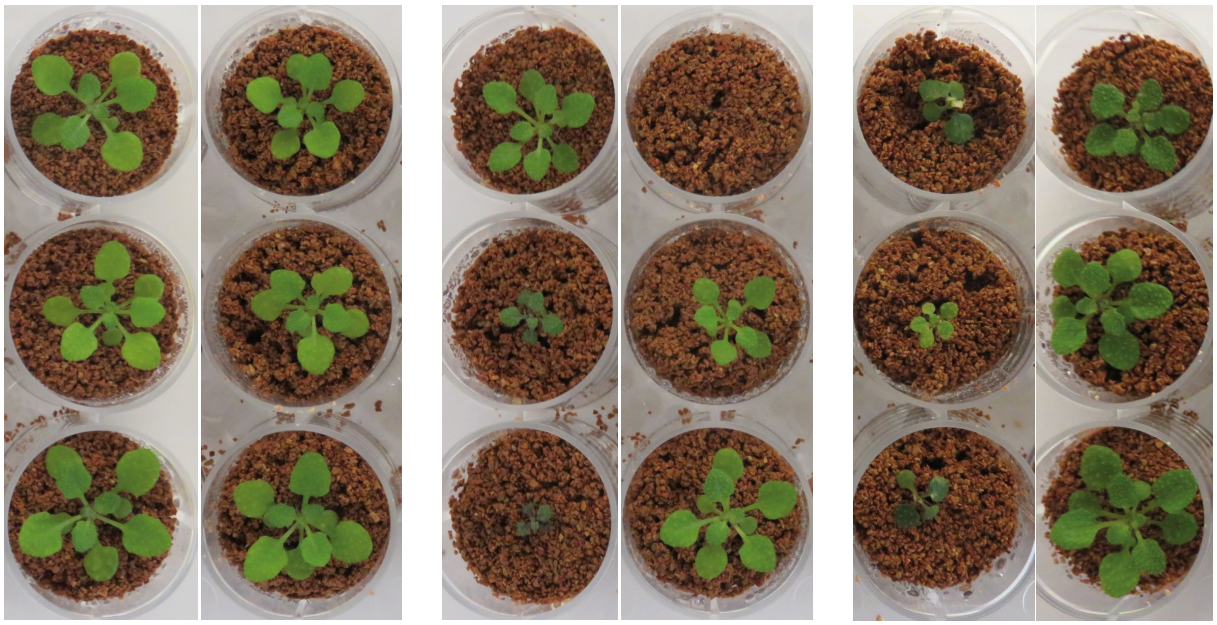
Supplementary Fig. 2: Inoculum composition of the focal community. Relative abundance of each strain in the 15 strain focal community mix used for plant inoculation. Strains are colored by phylum or class.



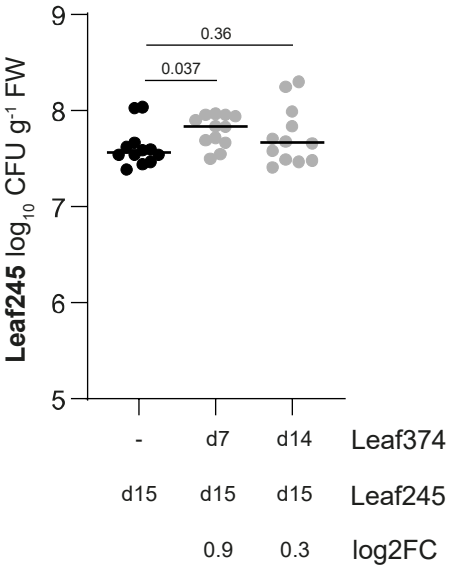
Supplementary Fig. 3: Composition of the control focal community observed in the drop-in screen. **a)** Composition of the focal community inoculum (n=4). **b)** Relative abundance of each focal strain in the community (n=16). The median and individual data points are shown. Replicates where a given strain was not detected (ND) are indicated by circle size and count above the strain name. Dots are colored according to phylum or Proteobacteria class.



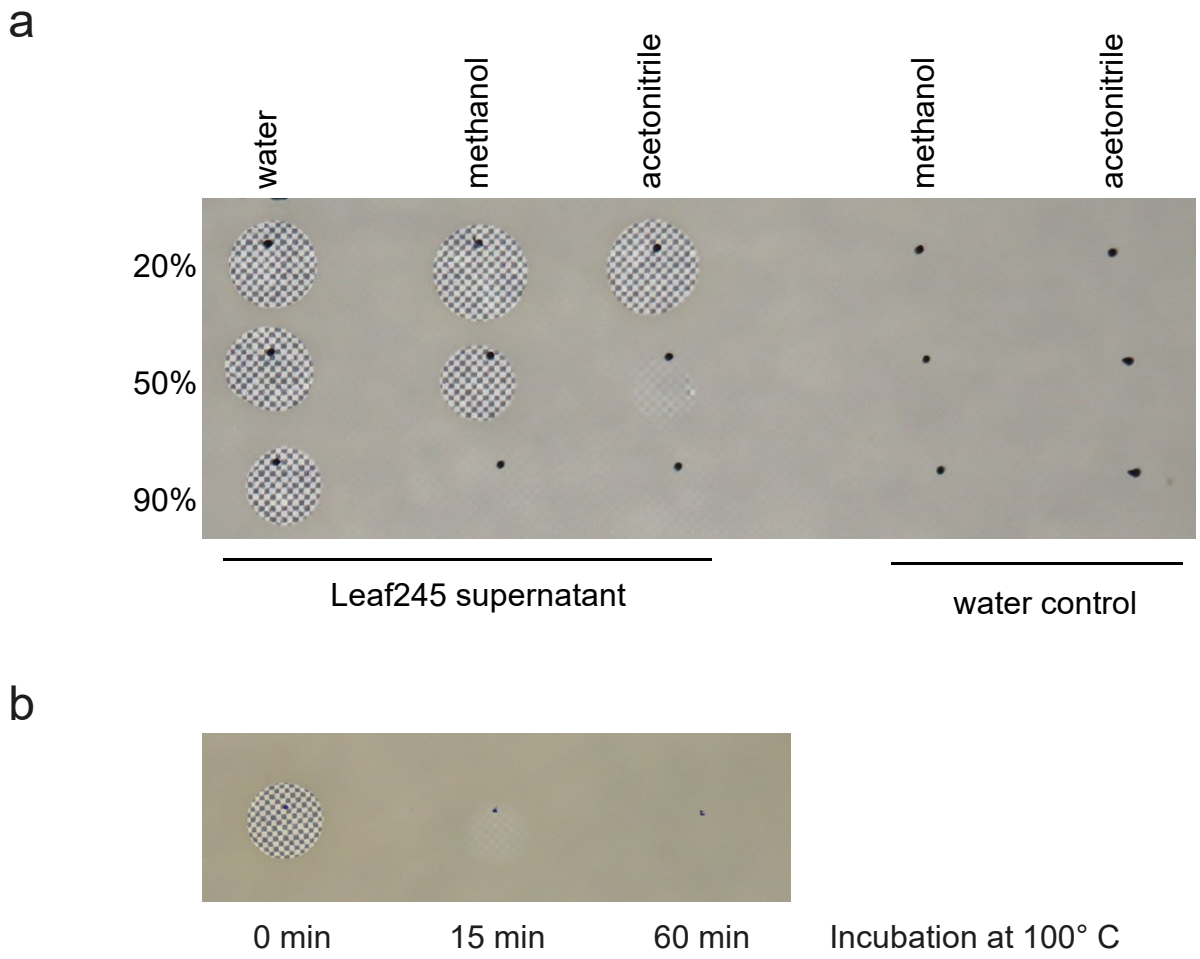
Supplementary Fig. 4: Relative composition of focal community samples used as the basis for comparison in the drop-in experiments. Stacked barplot of the focal community composition for the sample indicated on the x-axis. Strains are colored by phylum or class. Mapped reads not corresponding to a focal strain were summed up and are shown as "other". Sample S_72 was omitted from the comparisons as an outlier.



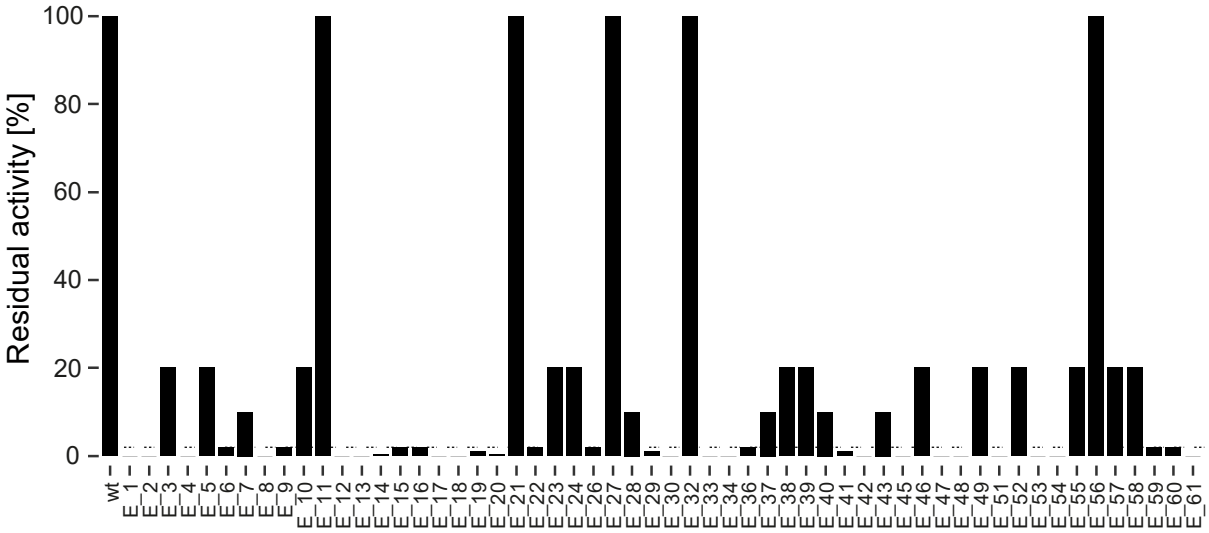
Supplementary Fig. 5: Plant phenotypes observed upon strain addition to the focal community. Images of 28 day old Arabidopsis Col-0 plants inoculated with the 15-strain focal community (left) or focal community with drop-in of *Serratia* Leaf50 (middle) or *Pseudomonas syringae* pv. tomato DC3000.



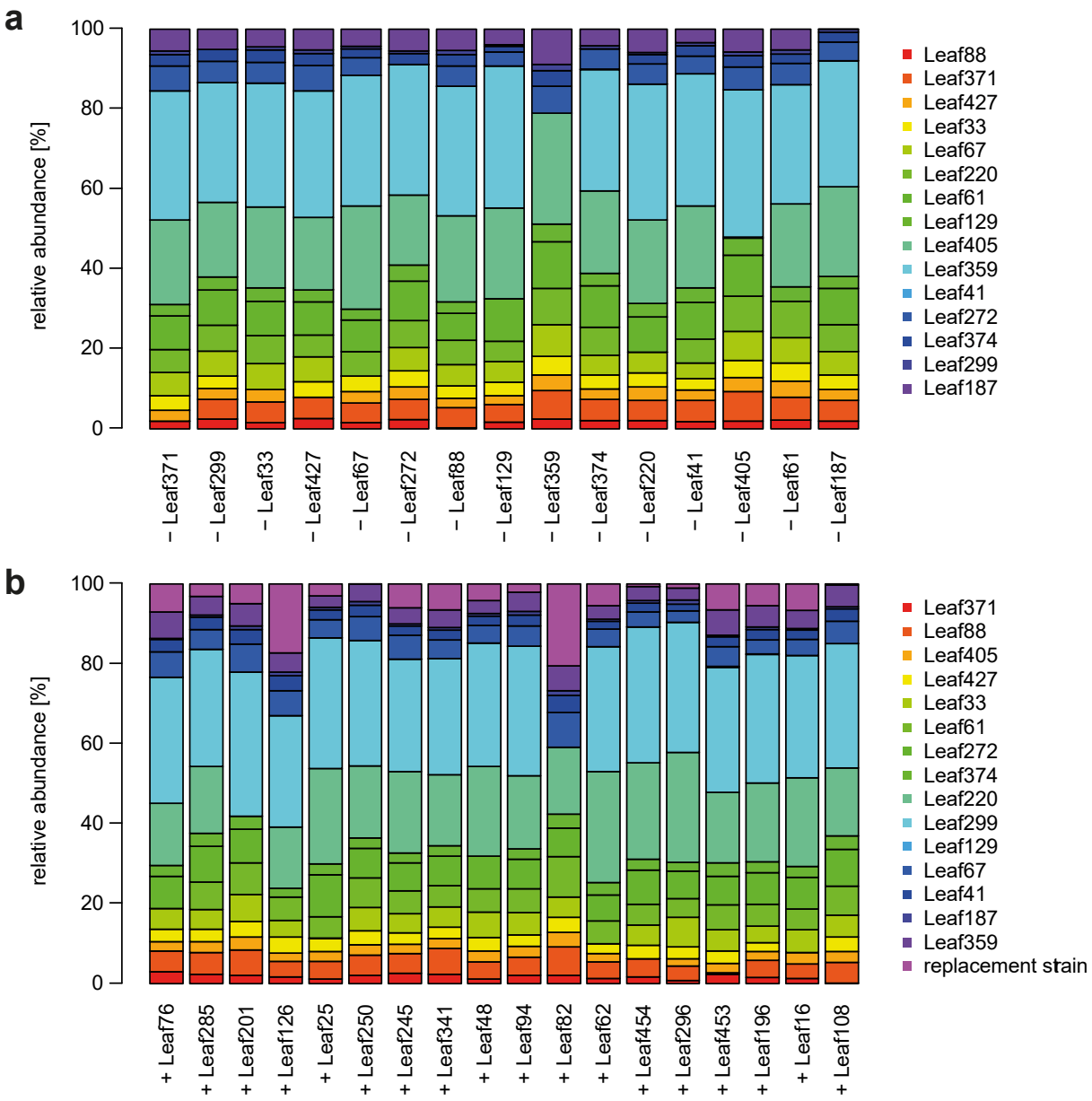
Supplementary Fig. 6: ***Aeromicrobium*** **Leaf245 colonization level.** Log10-transformed CFU per g plant fresh weight recovered after phyllosphere colonization in mono-association or in combination with Leaf374. The timepoint of inoculation with Leaf245 or Leaf374 is indicated below the figure as days (d) after planting. Shown are the median and individual data points (n=12). Exact p-values (two-sided Wilcoxon rank sum test) and log2 fold changes (log2FC) compared to the mono-association control are indicated above or below the graph, respectively. For the corresponding Leaf374 colonization see Fig. 3a.



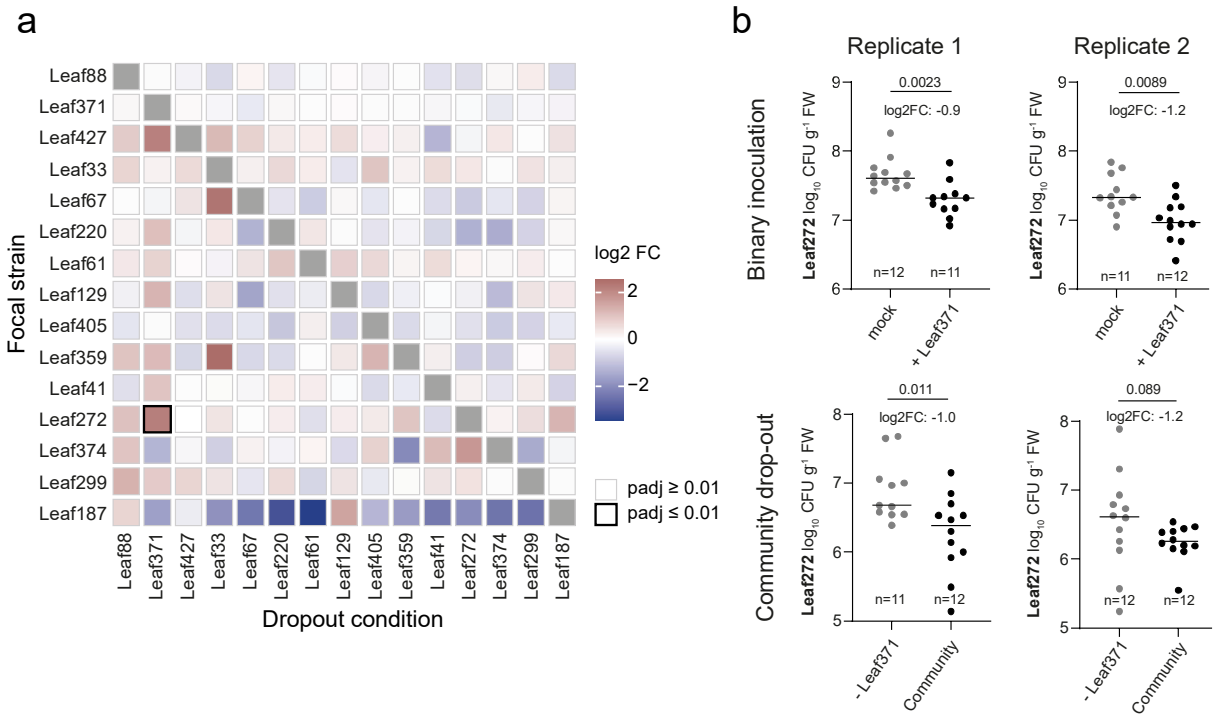
Supplementary Fig. 7: Effect of solvent and heat treatment on activity of Leaf245 supernatant. Samples were applied on top of Leaf374 overlay agar plate and inhibition was scored after 48 h. **a)** Solvent stability of Leaf245 bioactive supernatant was tested by diluting supernatant into methanol or acetonitrile. Dilution in water was included as control to account for dilution. Percentage of solvent in the final sample is indicated on the left. To exclude interference of the solvent with Leaf374 growth, a control where water instead of supernatant was used for dilution was included. **b)** Activity of unboiled (0 min) and boiled (15 or 60 min) Leaf245 supernatant.



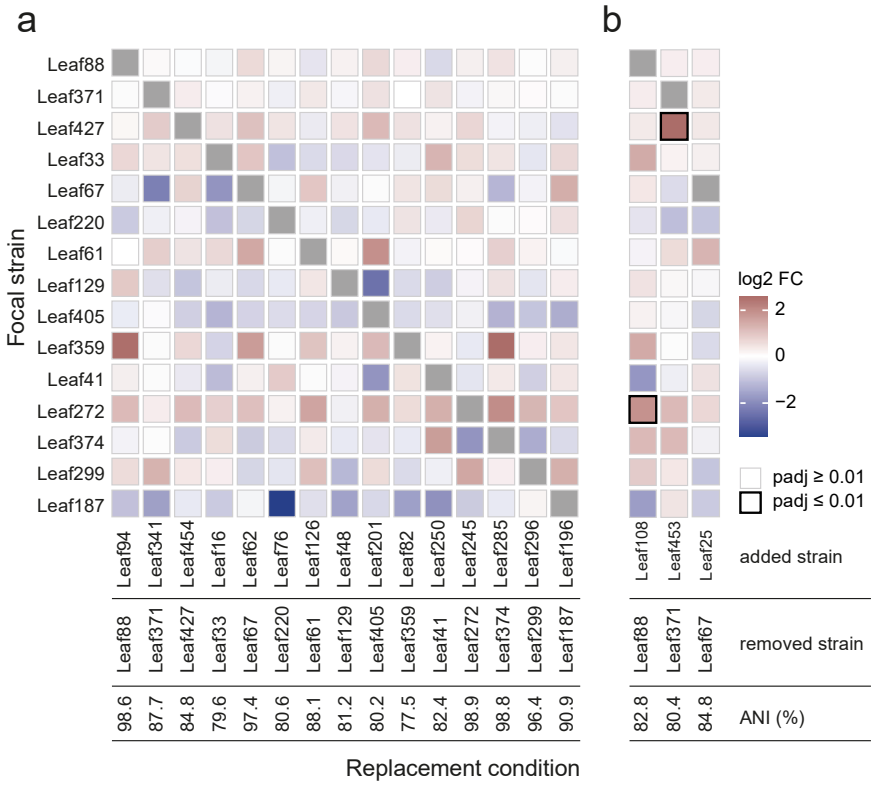
Supplementary Fig. 8: Validation of identified loss of function EMS mutants by supernatant activity. Leaf245 EMS mutants and wild type were suspended in 100 mM ammonium bicarbonate buffer pH 7.5 (OD_{600} = 25-50). Supernatant was diluted up to 500-fold and activity was assayed on Leaf374 overlay plates. Shown are residual activities normalized to wild type activity. The activity cut-off chosen for clones to be genome re-sequenced is shown as a dotted line.



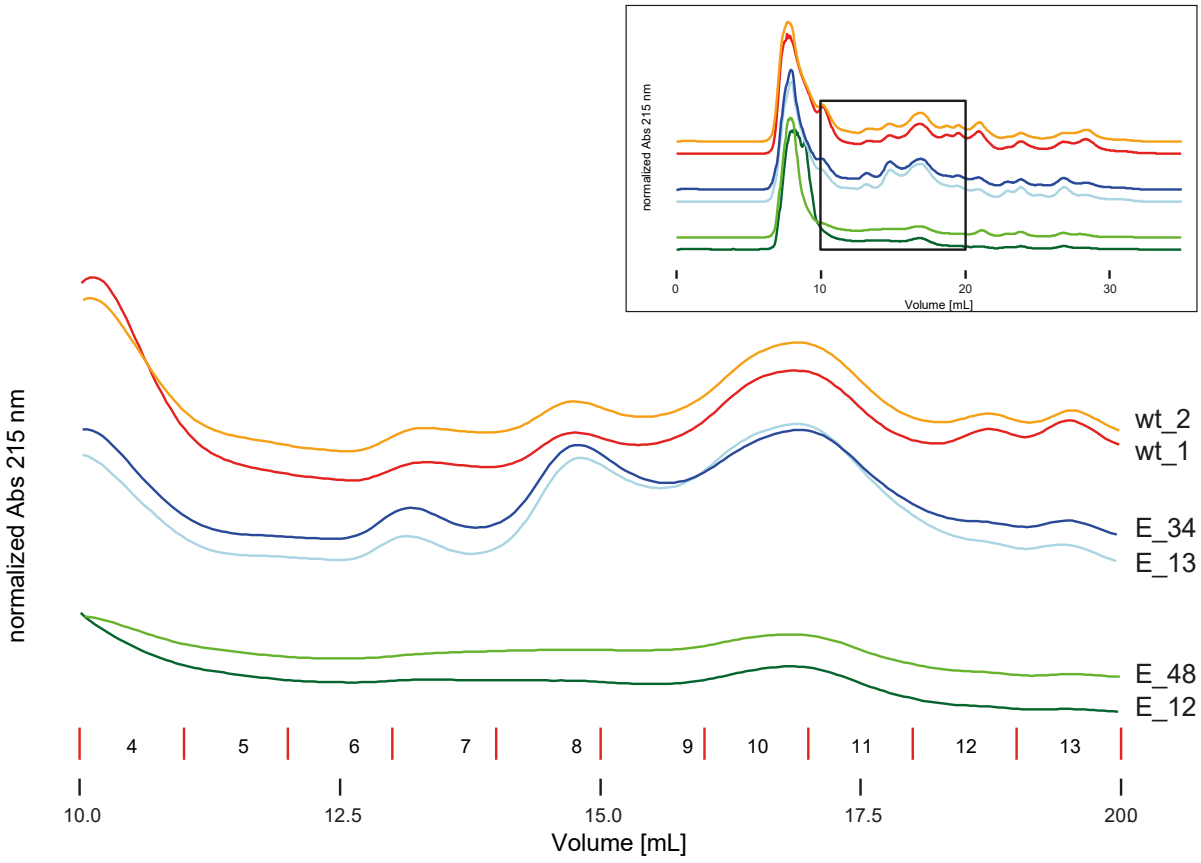
Supplementary Fig. 9: Inoculum composition of the drop-out and replacement conditions. Stacked barplots of the inoculum composition for each mix for which a focal strain was **a)** dropped out or **b)** was replaced with a close relative. The drop out or replacement strain is indicated below each bar.



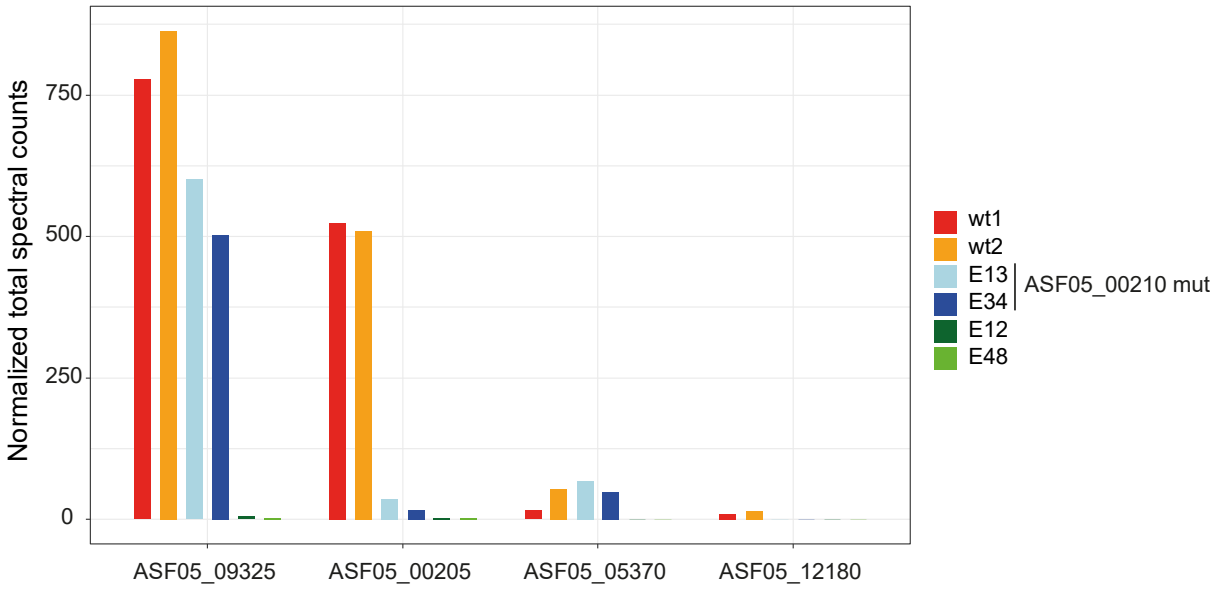
Supplementary Fig. 10: Effect of single strain drop-outs on the composition of the remaining focal community in the phyllosphere. a) Heatmap showing the log2-transformed fold changes (treatment vs control) of each focal strain (left) upon drop-out of one focal community strain (bottom) compared to the unperturbed 15-strain community. Significant changes are marked with a black frame (DESeq normalized counts, Wald-test, Benjamini-Hochberg adjusted $p \leq 0.01$, n=12). Grey boxes indicate that a strain was not present in a given condition. Exact p-values are in Supplementary Data 7. **b)** Validation of interaction between *Aeromicrobium* Leaf272 and *Rhizobium* Leaf371 by CFU enumeration. Top: Comparison of Leaf272 colonization in mono-association and in combination with Leaf371. Bottom: Comparison of Leaf272 colonization in the 15 strain focal community and in the Leaf371 drop-out condition. Shown are the median and individual data points of log10-transformed CFU g^{-1} plant fresh weight recovered on R-2A+M Rif plates for two biological replicates. Exact p-values (two-sided Wilcoxon rank sum test) and log2 fold-changes are indicated.



Supplementary Fig. 11: Strain abundance changes upon replacement of focal community strains. Heatmap showing log2-fold changes (treatment vs control) of focal community strains (left) upon exchange of individual focal strains with **a)** a close relative or **b)** more distant relative. The removed strain and the corresponding replacement strain (added strain) are shown for each treatment (bottom) and the average nucleotide identity (ANI) of the two strains is shown below. Significant changes are marked with a black frame (DESeq normalized counts, Wald-test, Benjamini-Hochberg adjusted $p \leq 0.01$, n=12). Grey boxes indicate that a strain was not present in a given condition. Exact p-values are provided in Supplementary Data 8.



Supplementary Fig. 12: Trace of absorption at 215 nm observed for Leaf245 wild type and EMS mutant supernatants during size exclusion chromatography. Absorption was normalized for each sample by the maximum absorption and stacked for comparison. Samples were grouped by sample type. Two replicates of the wild type (red and orange), two ASF05_00210 mutants (E_13: light blue; E_34: blue) as well as GsiA (E_12: dark green) and histidine kinase (E_48: light green) mutants are shown. Red vertical lines and numbers in between indicate the collected fractions. The top right box shows the trace of the complete fractionation run and indicates the enlarged region shown in the main panel. Fractions 7-9 were pooled for SDS-PAGE and mass spectrometry analysis (Fig. 4c and Supplementary Fig. 13).



Supplementary Fig. 13: Approximate quantification of protein abundances based on proteomics for Leaf245 wild type and EMS mutant supernatant fractions. Pooled active fractions of wild type and corresponding (non active) mutant supernatant were sent for protein identification by shotgun proteomics. Normalized total spectra counts were calculated for each protein indicated on the x-axis based on the total number of spectra recorded for each sample.