The plant NADPH oxidase RBOHD is required for microbiota homeostasis in leaves

Sebastian Pfeilmeier, **Gabriella C. Petti**, **Miriam Bortfeld-Miller**, **Benjamin Daniel**, **Christopher M. Field**, **Shinichi Sunagawa**, **Julia A. Vorholt**[*]
Institute of Microbiology, ETH Zurich, 8093 Zurich, Switzerland

Abstract

The plant microbiota consists of a multitude of microorganisms that can impact plant health and fitness. However, it is currently unclear how the plant shapes its leaf microbiota and what role the plant immune system plays in this process. Here, we evaluated *Arabidopsis thaliana* mutants with defects in different parts of the immune system for an altered bacterial community assembly using a gnotobiotic system. While higher order mutants in receptors that recognize microbial features and in defense hormone signaling showed pronounced microbial community alterations, the absence of the plant NADPH oxidase RBOHD caused the most substantial change in the composition of the leaf microbiota. The *rbohD* knockout resulted in an enrichment of specific bacteria. Among these, we identified *Xanthomonas* strains as opportunistic pathogens that colonized wild-type plants asymptomatically but caused disease in *rbohD* knockout plants. Strain dropout experiments revealed that the lack of RBOHD unlocks the pathogenicity of individual microbiota members driving dysbiosis in *rbohD* knockout plants. For full protection, healthy plants require both a functional immune system and a microbial community. Our results show that the NADPH oxidase RBOHD is essential for microbiota homeostasis and emphasizes the importance of the plant immune system in controlling the leaf microbiota.

Keywords

Microbiota; plant immunity; dysbiosis; NADPH oxidase; phyllosphere; synthetic community; RBOHD; opportunistic pathogen; Xanthomonas

In nature, plants are colonized by a diverse microbiota. The phyllosphere constitutes a vast microbial habitat in which microorganisms populate plant surfaces as epiphytes and the intercellular space inside leaves as endophytes[1,2]. The functional repertoire of the

microbiome expands the capacity of the plant to adapt to its environment, as beneficial microbes can promote plant growth and increase tolerance to abiotic and biotic stress[3–5]. Plants must be prepared to recognize pathogens and mount defense responses against them[6,7] while being colonized by a complex microbiota that can potentially also trigger immunity and provide indirect plant protection[8,9]. Therefore, plant immunity must be balanced to allow accommodation of a microbiota[4,10]. Emerging evidence suggests that the plant immune system has a significant impact on the bacterial microbiota in the rhizosphere[9,11,12] and phyllosphere[13–15]. However, our current understanding of specific host factors involved in the perception and structuring of the microbiota and the molecular mechanisms underlying them is limited.

In recent years, experiments with synthetic communities (SynCom) have been introduced to identify host genetic factors that drive microbiota assembly under controlled environmental conditions[16]. The availability of representative collections of strains isolated from *Arabidopsis thaliana* and the reproducible reconstitution of the microbiota to similar phylogenetic structures as found in nature[11,17,18] provide opportunities to quantify community changes and probe available genetic resources to establish gene-phenotype causal relationships.

Here, we used a reductionist approach to link host immune components to the assembly of the bacterial microbiota. To test the impact of the plant immune system on the microbiota, we used single and higher order mutants of well-described key immunity factors, and probed changes in the microbiota composition compared to the one that establishes on wild type plants. To identify specific plant-microbiota interactions we adopted a SynCom approach to achieve strain-level resolution. We found a number of *A. thaliana* immunity mutants with altered bacterial community composition and demonstrate a critical function of an immunity-related NADPH oxidase in shaping the leaf microbiota. Ultimately, we explore the interplay between the host and individual community members and their roles in plant health and community composition.

Results

Bacterial community in the *A. thaliana* phyllo- and endosphere

For microbiota reconstitution experiments, we used an established gnotobiotic growth system consisting of *A. thaliana* and a bacterial synthetic community based on the *At*-LSPHERE collection of genome-sequenced strains[17,19]. The microbiota inoculum communities consisted of either 222 or 223 strains (termed SynCom-222 or SynCom-223), which represent the maximal genomic diversity of the *At*-LSPHERE and included strains with identical amplicon sequence variant (ASV) but different genomes. In follow-up experiments, we chose a subset of 137 strains (SynCom-137) in which each strain represent an ASV with a unique 16S rRNA gene (rDNA) amplicon sequence (Supplementary Table; Methods). Germ-free 10-day-old *A. thaliana* seedlings were inoculated with the SynCom and the aboveground part of the plants were harvested 5.5 weeks after germination to examine the bacterial communities by 16S rDNA amplicon sequencing (Fig. 1a).

The relative abundance of phyllosphere strains (Fig. 1b) was similar to the natural microbiota of *A. thaliana* at the phylum level and was consistent with SynCom experiments from previous studies[17,19]. The community structure in the *A. thaliana* Col-0 phyllosphere comprised Alphaproteobacteria (36%), Betaproteobacteria (30%), Gammaproteobacteria (2%), Actinobacteria (19%), Bacteroidetes (12%) and Firmicutes (0.0005%) (Fig. 1b). In addition to these phyllosphere samples, we analyzed the endosphere-enriched bacterial fraction (Fig. 1a; Methods). Comparing the community composition, we observed a relative enrichment of Gammaproteobacteria in the endosphere compared to the entire phyllosphere (Fig. 1b and Extended Data Fig. 1). The quantification of bacterial abundance by qPCR using 16S rDNA copies per plant gene revealed reduced colonization of the endosphere. Similarly, colony forming units (CFU) per gram of fresh weight recovered from endosphere samples were approximately 1000-fold lower than those recovered from phyllosphere samples (Fig. 1c), as expected[14,20].

The analysis of the relative abundance of strains revealed that a few bacteria were significantly enriched or depleted in the endosphere across three SynCom experiments (Fig. 1d). *Xanthomonas* Leaf131 in particular but also *Pseudomonas* Leaf59 were enriched in the endosphere samples, thus explaining the higher relative abundance of Gammaproteobacteria (Fig. 1b). In contrast, *Methylophilus* Leaf416, *Aeromicrobium* Leaf350 and *Pseudorhodoferax* Leaf265 were depleted, indicating that these strains or ASVs have a greater relative abundance on the leaf surface compared to the community in the apoplast (Fig. 1d). The experiment was repeated with SynCom-223 and SynCom-137 and revealed similar results (Source Data Fig. 1d), which shows that the strains in SynCom-137 are representative for the ASVs.

Because the differences in relative abundance suggest a distinct adaptation to an endophytic or epiphytic lifestyle of some strains, we conducted leaf infiltration experiments to test bacterial growth in the apoplast. Indeed, the endophytic-enriched strains *Xanthomonas* Leaf131 and *Pseudomonas* Leaf59 were able to grow to a greater extent (37-fold and 3.5-fold increase, respectively) after low-dose infiltration into leaves of soil-grown Col-0 plants, in contrast to the depleted strain *Pseudorhodoferax* Leaf265 (2.5-fold increase) (Extended Data Fig. 2), confirming the potential of the former, especially *Xanthomonas* Leaf131, to thrive as endophytes. Importantly, no disease symptoms were observed after leaf infiltration, and bacterial titers were lower than those of the model pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (Pst) (Extended Data Fig. 2).

Screening *A. thaliana* mutants for an altered bacterial community

To identify host genetic factors that impact the bacterial community composition in the phyllosphere, we tested *A. thaliana* mutants with defects in distinct parts of the immune system. These included mutants for immune signaling or regulators that are involved in pattern-triggered immunity (PTI)[21], effector-triggered immunity (ETI)[22] and homeostasis of defense-associated plant hormones[6] (Supplementary Table).

A. thaliana Col-0 wild-type and different mutants were inoculated with SynCom-222, and the community composition was determined by 16S rDNA amplicon sequencing. We examined the impact of the host genotype on the community composition in a principal

component analysis (PCA), which quantifies how much variation in the strain (or ASV) abundances can be explained by individual genotypes, and determined the effect size and statistical significance by PERMANOVA. Bacteria inhabiting the apoplast are in closer contact with host cells and, as such, may be subject to stronger host control than bacteria on the leaf surface. Therefore, we primarily focused our analysis on the bacterial community in the endosphere compartment.

The largest impact on the endophytic community was observed in the *A. thaliana* mutant *rbohD*[23], which lacks the plasma membrane-localized NADPH oxidase Respiratory Burst Oxidase Homologue (RBOH) D, with a 25% effect size (Fig. 2a-c). Among the genotypes with defects in PTI signaling, the effect of *rbohD* was larger than that of other genotypes with defects in pattern recognition receptors (e.g., *fls2efrcerk1* (*fec*)[24], *lym3lym1*[25], *pepr1pepr2*[26]) or coreceptors (e.g., *bak1-5bkk1cerk1* (*bbc*)[27]). The genotypes compromised in plant hormone biosynthesis or perception had an effect size between 17% by *jar1* (JA signaling)[28] and 7% by *ein2* (ethylene signaling)[29]. Notably, the higher-order mutants with defects in multiple hormone systems, *jar1ein2npr1* (*jen*)[30] and *dde2ein2pad4sid2* (*deps*)[31], did not have a stronger impact on the community composition than the single mutants tested (Fig. 2a). The salicylic acid-dependent immunoregulatory mutant *npr1*[32] and the stomatal opening regulator *ost1*[33] had a minor impact on the endophytic community composition. The quadruple mutants *min7fec* (*mfec*) and *min7bbc* (*mbbc*)[27] also affected community assembly, corroborating a recent study reporting that these mutants are involved in bacterial microbiota homeostasis[14]. The hyperimmune genotypes *cpr5*[34] and *cpr6*[35], which have constitutively high salicylic acid (SA) levels, also had an impact on the community.

In fact, all genotypes with large effects on the endosphere community also affected the entire phyllosphere community (Extended Data Fig. 3a). The malectin-like receptor-like kinase Feronia knockout *fer*[36] had the highest effect size (28%) on the phyllosphere community but was not tested for the endosphere community due to poor germination and impaired development[37]. In addition to the genotype effect on data variance in a PCA, we also analyzed the variability of community composition upon colonization of immunity mutants (Fig. 2e). The microbiota variability in *rbohD* plants was higher than that in wild-type plants, while it was equally variable in *jar1* and wild-type plants in all SynCom experiments.

Analysis of the community composition at higher taxonomic levels, such as phylum (or class for Proteobacteria), between individual genotypes revealed that several taxa were significantly altered in their relative abundance in *rbohD*, *mbbc*, *bbc*, *mfec*, *cpr5*, *jar1* and *fer* (*fer* tested only for phyllosphere) compared to wild-type plants (Fig. 2d and Extended Data Fig. 3b).

Hierarchical clustering of the differential strain abundances between individual genotypes and Col-0 of endosphere-enriched strains grouped *rbohD* separately from the other genotypes, but no other significant clusters were observed (Extended Data Fig. 4a). Cluster analysis of the phyllosphere community profiles revealed clustering of all genotypes but *fer* and *cpr5* separately (Extended Data Fig. 4b).

Overall, we found striking genotype effects on the leaf bacterial community in both endosphere-enriched and total phyllosphere samples.

rbohD assembles a dysbiotic microbiota

Because *rbohD* showed the largest impact on the endosphere community with an effect size of 25% (Fig. 2a), compared to 17% and lower for other immunity mutants, we examined the impact of RBOHD on the microbiota more closely. RBOHD is part of a family of NADPH oxidases that generate extracellular reactive oxygen species (ROS)[38] and function in diverse physiological processes, including biotic and abiotic stress signaling[39,40]. In particular, RBOHD is regulated through activated immune receptors and is responsible for inducible apoplastic ROS production during PTI and ETI[23,41,42], playing critical roles in microbe-associated molecular pattern (MAMP)-induced stomatal closure[43], cell wall damage-induced lignification[44,45] and resistance to pathogens[23,46,47].

In addition to assembling an altered endophytic community, as described above, *rbohD* plants developed disease symptoms after inoculation with a SynCom (Fig. 3a), while axenic *rbohD* plants were indistinguishable from the wild-type. SynCom-induced disease was also reflected in the reduced average plant weight (Fig. 3b) and points towards a dysbiosis phenotype[48]. Although symptoms could be observed in most *rbohD* plants, the disease severity varied (Fig. 3a). We classified SynCom-induced disease into five categories (1, healthy, to 5, dead), and increasing disease severity matched with decreasing plant fresh weight (Extended Data Fig. 5b). Next, we examined the dysbiotic microbiota of *rbohD* at the strain level by using the SynCom-137. A defined set of bacteria was affected in their relative abundance in the endosphere and phyllosphere compared to Col-0 (Fig. 3c). Mostly Gammaproteobacteria (*Xanthomonas* Leaf131, *Serratia* spp. Leaf50 and Leaf51, *Pseudomonas* spp. Leaf83, Leaf58, Leaf127, Leaf15 and Leaf434), an Alphaproteobacterium (*Sphingobium* Leaf26) and two Actinobacteria (*Sanguibacter* Leaf3 and *Microbacterium* Leaf161) showed a significant increase in relative abundance in *rbohD* compared to wild-type plants (Fig. 3c) and were among the top colonizers in the endosphere of *rbohD* (Fig. 3d). Notably, the same strains were consistently enriched in *rbohD* in three SynCom experiments, which underlines the robustness of the observed effect. SynCom-137 recapitulated the results of the full community, thus showing that single strains were representative of the ASV with respect to their behavior in *rbohD*. Furthermore, most of the affected strains were increased in both the endosphere and phyllosphere of *rbohD* (Fig. 3c).

To investigate whether disease in *rbohD* is accompanied by an increased bacterial load, we assessed the bacterial cell numbers in the phyllosphere and endosphere. Overall bacterial abundance measured by CFU and qPCR normalized to plant weight or plant genomes, respectively, was not significantly increased in *rbohD* compared to Col-0 across multiple experiments (Extended Data Fig. 5c). This finding is in line with previous plant pathology studies that did not observe a correlation between disease severity and pathogen titer[49–51], although a correlation had been described in other studies[20,52–54]. In fact, the outcome might depend on the specific plant-microbe interaction, the microbiota background, and environmental conditions. To test whether certain bacteria are linked to disease in *rbohD*,

we correlated plant weight and the relative abundance of bacterial families. Indeed, we observed a strong negative correlation for *Xanthomonadaceae* (Spearman coefficient ρ=-0.82 and ρ=-0.83 for experiments with SynCom-222 and SynCom-137, respectively) in the phyllosphere samples of *rbohD* (Extended Data Fig. 5d).

As mentioned above, in our experiments, *rbohD* had a larger effect on the endosphere community profile and on individual strains than other PTI mutants, such as *fec*, *bbc*, *mfec*, and *mbbc* (Fig. 2a, Extended Data Fig. 6a). For example, beside *Xanthomonas* Leaf131 that was more abundant also in other PTI mutants, additional *Pseudomonas* strains, Leaf83, Leaf58, Leaf127, Leaf15, Leaf434, showed increased abundance only in *rbohD*. In addition to RBOHD, RBOHF also plays a role in ROS production and disease resistance, although to a lesser extent[23,47,55]. We tested the double knockout mutant *rbohDrbohF*[23] and found that it had an even stronger impact on SynCom-137 than *rbohD* alone. For example multiple strains were specifically enriched in the endosphere of *rbohDrbohF* but not in *rbohD* (Extended Data Fig. 6b), such as *Rhizobium* Leaf311 and *Brevundimonas* Leaf168 (Alphaproteobacteria); *Pseudorhodoferax* Leaf265 and *Acidovorax* Leaf160 (Betaproteobacteria); *Serratia* Leaf50, *Erwinia* Leaf53 and *Pseudomonas* Leaf59 (Gammaproteobacteria); *Arthrobacter* strains Leaf145 and Leaf141, and *Plantibacter* Leaf314 (Actinobacteria).

The substantial impact of *rbohD* and *rbohDrbohF* compared to other PTI-associated genotypes in our experiments (Fig. 2a and Extended Data Fig. 6) could be explained by the central roles of NADPH oxidases in converging signaling pathways in immunity and beyond[40,56]. Overall, our results suggest that dysbiosis of *rbohD* is caused by a shift in community structure with an increased abundance of certain taxa dominating the community.

The bacterial leaf community contains opportunistic pathogens

Many plant-associated bacteria carry immunogenic MAMPs and can potentially trigger immunity[8]. We selected phylogenetically diverse species, including *rbohD*-enriched strains, and tested whether they can elicit an apoplastic ROS burst. Extracts of most *rbohD*-enriched strains and other strains triggered plant ROS, suggesting that ROS-eliciting potential is common, especially among Beta- and Gammaproteobacteria (Extended Data Fig. 5e).

To identify the causal agent(s) of SynCom-induced disease in *rbohD*, we tested the pathogenicity of the individual SynCom-137 strains. We inoculated single strains onto germ-free *rbohD* seedlings and monitored the plant phenotype. Based on this screening we identified 28 phenotype-inducing strains that caused plant symptoms, such as necrotic lesions or curled leaves, indicating a strong host-microbe interaction (Extended Data Fig. 5a), with *Xanthomonas* Leaf131 triggering the most severe disease symptoms when inoculated alone (Fig. 4a and 4b). The *At*-LSPHERE strain collection contains another *Xanthomonas* strain, Leaf148, which shares the same 16S rDNA amplicon sequence but has only an 87% average nucleotide identity (ANI) with Leaf131. We tested both *Xanthomonas* strains for their phenotype on Col-0, *rbohD* and *rbohD/RBOHD* complementation lines. While *rbohD* plants became diseased or died after inoculation with Leaf131 or Leaf148, Col-0 and the complemented *rbohD/RBOHD* looked healthier and had higher plant weights

(Fig. 4a and 4b). In addition, bacterial extracts of both *Xanthomonas* strains triggered a plant ROS burst, indicating that the plant is able to perceive these bacteria (Fig. 4c). Moreover, infiltration of *Xanthomonas* Leaf131 into leaves of soil-grown Col-0 wild-type and *rbohD* plants resulted in mild disease symptoms only in *rbohD* (Fig. 4d). We concluded that Leaf131 and Leaf148 are opportunistic pathogens that develop their deleterious potential on immunocompromised *rbohD* plants.

RBOHD activation in *A. thaliana* occurs through N-terminal phosphorylation by receptor-like cytoplasmic kinases, such as Botrytis-Induced Kinase 1 (BIK1)[21,42]. Multiple complementary and overlapping RBOHD phosphorylation sites have been reported to be critical for PTI signaling[56–60]. We probed whether RBOHD phosphorylation sites can be linked to microbiota-induced disease. After validation of impaired MAMP-induced ROS production in *RBOHD* mutants (Extended Data Fig. 7a)[57,58], we tested whether RBOHD with mutated phosphorylation sites, i.e., *rbohD/RBOHD-S39A-S339A-S343A* and *rbohDrbohF/RBOHD-S343A-S347A*, could complement the *rbohD* disease phenotype after inoculation with SynCom-137 or single strains of *Xanthomonas* Leaf131 and Leaf148. The presence of SynCom-137 caused a significant weight reduction of *rbohD* and *rbohDrbohF* and, to a lesser extent, of the phosphorylation site mutants compared to their respective controls (Fig. 4b). The intermediate disease phenotype of the RBOHD phosphorylation site mutants compared to the knockout lines *rbohD* and *rbohDrbohF* could be due to regulatory redundancy of multiple posttranslational modifications resulting in tunable activation of RBOHD and ROS production[56–60]. Inoculation with *Xanthomonas* Leaf131 and Leaf148 caused disease symptoms and reduced the average plant weight of *rbohD* and the two phosphorylation site mutants compared to their control lines Col-0, *rbohD/RBOHD*, and *rbohDrbohF/RBOHD*, respectively (Fig. 4b and Extended Data Fig. 7b). In contrast, inoculation with Pst reduced the plant weight to a similar level on all genotypes irrespective of RBOHD presence, while colonization by the nonvirulent Pst *hrcC-* did not affect the plant weight of the different genotypes (Extended Data Fig. 7b).

In summary, our data highlight the importance of RBOHD and its regulation via immunity-associated phosphorylation sites for plant health during colonization with opportunistic *Xanthomonas* pathogens and for microbiota homeostasis.

Dysbiosis of *rbohD* is driven by opportunistic pathogens

Due to the enrichment of *Xanthomonas* Leaf 131 in *rbohD* (Fig. 3c and 3d) and the phenotype it caused in mono-association (Fig. 4a), we next wondered whether the opportunistic pathogen was sufficient to explain the phenotype in a community context and speculated on the relative contribution of plant immunity and the microbiota to plant health. We designed several synthetic communities by removing either the opportunistic pathogen *Xanthomonas* Leaf131 or all *rbohD*-enriched and plant phenotype-inducing (REPI) strains (Supplementary Table). We inoculated germ-free Col-0, *rbohD* and *rbohD/RBOHD* with the selected bacterial mixes. The full community control, i.e., SynCom-137, reduced plant health and weight compared to axenic control plants only in *rbohD*, not in Col-0 and *rbohD/RBOHD* (Fig. 5), confirming our previous observation that SynCom causes disease in *rbohD* (Fig. 3a). *Xanthomonas* Leaf131 single inoculation was detrimental to

rbohD and reduced plant weight in Col-0 and *rbohD/RBOHD*, corroborating its pathogenic potential (Fig. 4). We then mimicked the dysbiotic community observed in *rbohD* plants by assembling SynCom-REPI, consisting of the 32 strains for which we observed phenotypes or enrichment in *rbohD* (Fig. 3c and Extended Data Fig. 5a). SynCom-REPI severely reduced the plant weight of *rbohD* and had a mild effect on Col-0. Remarkably, when Leaf131 or all SynCom-REPI members were removed from SynCom-137, the plant weight difference between Col-0 and *rbohD* disappeared and both genotypes had phenotypes that were not distinguishable from axenic control plants (Fig. 5). Because *Xanthomonas* Leaf131 is also part of the REPI strains, we assumed that it is responsible for the negative effect on *rbohD*. Thus, we repeated the dropout SynCom experiment with an additional condition consisting of Leaf131 dropout from SynCom-REPI. Again, we observed that whenever the SynCom did not contain Leaf131, plant health and weight were increased in Col-0, *rbohD* and *rbohD/RBOHD* (Extended Data Fig. 8). This shows that *Xanthomonas* Leaf131 is responsible and sufficient for SynCom-induced disease in *rbohD*. Importantly, a diverse bacterial community can alleviate the detrimental effects of the opportunistic pathogen *Xanthomonas* Leaf131 on immunocompromised *rbohD*; however, a healthy plant requires functional RBOHD (Fig. 5).

Discussion

The assembly and maintenance of a healthy microbiome is crucial for host fitness[10,48]. Prior knowledge about the function of key plant immunity genes makes it now possible to test how these genes influence not only pathogens, but the entire plant microbiota and interactions within these. By using a reverse genetic screen, we found multiple plant immunity mutants that impact bacterial leaf community assembly. In particular, we uncovered the pivotal role of the NADPH oxidase RBOHD to prevent microbiota dysbiosis in the phyllosphere. NADPH oxidases are evolutionarily conserved in multicellular eukaryotes with functions in defense, development and redox-dependent signaling[61,62]. Strikingly, NADPH oxidases play a major role in gut microbe homeostasis in mammals[63,64], zebrafish[65] and insects[66,67], highlighting the importance of understanding the interplay between ROS production, host immunity and microbiota.

Likewise to dysbiosis in the human intestine, which has been characterized by an increase of relative abundance of Gammaproteobacteria[68] and dysbiotic communities being more diverse than communities in healthy patients[69], we also found an enrichment of multiple Gammaproteobacteria (Fig. 3) and a higher variability among the microbiota between individual *rbohD* plants compared to wild-type (Fig. 2e). However, in contrast to observations made in the gut, we identified *Xanthomonas* spp. as the most conspicuous, acting as opportunistic pathogens when inoculated on plants alone (Fig. 4).

Xanthomonas spp. are common plant pathogens and are ubiquitously found in the phyllosphere microbiota on a variety of plants[70–74]. The type-3 secretion system (T3SS) represents a major virulence determinant in pathogenic *Xanthomonas*[74]. A preliminary genome analysis of *Xanthomonas* Leaf131 and Leaf148 revealed the absence of a T3SS17 which is in accordance with their opportunistic lifestyle. Note also that the opportunistic *Xanthomonas* pathogens identified here have been isolated

from phenotypically asymptomatic *A. thaliana* plants in their natural environment[17]. Correspondingly, numerous ecological studies have detected pathogenic species, including *Xanthomonas*, as members of the plant microbiota that are asymptomatic for the host in a community context but cause disease during mono-association[53,75–77]. Our data support the concept of conditional pathogens[78], which defines species with pathogenic potential as an integral part of the microbiota that are kept in check by other community members and the host, and underlines the importance of both the genotype and the microbiota to plant protection (Fig. 5).

The mechanisms by which the microbiota reduces the prevalence of opportunistic pathogens might be manifold and remain to be elucidated. Screening of bipartite interactions between *At*-LSPHERE strains *in vitro* did not reveal growth inhibition of *Xanthomonas* Leaf131 and Leaf148 by other community members[79]; nonetheless, the plant could be protected through a concerted action of multiple commensals or indirectly via stimulation of plant immunity[80–84]. Furthermore, *Xanthomonas* WCS2014-23, a close relative to the opportunistic pathogen *Xanthomonas* Leaf148 (97.5% ANI), has been isolated from *A. thaliana* roots and interacts with other bacteria to promote plant growth and to induce disease resistance[85].

Apoplastic ROS production by RBOHD is a typical response during plant-pathogen interactions. As the community shift in *rbohD* could be observed in endosphere and phyllosphere samples, the internal and external bacterial populations might form a continuum[1]. The absence of RBOHD could facilitate endophytic growth of certain taxa, which then relocate to the leaf surface. RBOHD-produced ROS can act directly as antimicrobials as well as in local and systemic signaling[23,38,40]. In our experiments, many bacteria were able to induce a ROS burst, indicating that microbiota members have the capacity to trigger PTI. Future experiments will help to disentangle whether RBOHD selectively impacts individual bacteria, for example through antimicrobial ROS, cell wall modification via cross linking or downstream immune responses[23,40,58,44], and whether pathogenic *Xanthomonas* act as keystone species by affecting other microbes either directly or indirectly via the plant.

Recently, Chen et al. proposed a genetic network to control the microbiota and prevent dysbiosis[14]. Our data indicate that RBOHD and RBOHF could be central parts of such a network and integrate signals from converging pathways, as *rbohD* and *rbohDrbohF* knockout lines had a more pronounced impact on the endophytic community than other immunity mutants under our experimental conditions. While Gammaproteobacteria generally benefited from gene knockouts in *bbc mbbc*, *mfec*, *rbohD* and *rbohDrbohF*, we could see overlap but also differences at the strain level between the genotypes (Fig. 2d and Extended Data Fig. 6). Notably, RBOHD has been identified as a central link between the immune signaling pathways PTI and ETI[86,87], which could explain its extraordinary effect on the microbiota. In addition, RBOHD could be a convergence point of biotic and abiotic stress signaling[39,40] that jointly regulate microbiota homeostasis. It remains to be determined how the different host components interact to control the microbiota. The impact of RBOHD on selected members of the bacterial microbiota potentially has strong

implications for plant defense against pathogens as these bacteria may prime plant immunity and convey health-protective properties.

In summary, we identified a pivotal function of the NADPH oxidases RBOHD and RBOHF in maintaining microbiota homeostasis in the phyllosphere to prevent dysbiosis. Our findings emphasize the role of the commensal microbiota together with a functional immune system in the control of opportunistic pathogens.

Methods

Plant growth conditions

Gnotobiotic plants were grown in calcined clay (Diamond Pro Calcined Clay Drying Agent, Arlington, Texas, USA) supplemented with 0.5x Murashige and Skoog (MS) medium including vitamins, pH 5.8 (M0222.0050; Duchefa, Haarlem, Netherlands), in round gamma-irradiated microboxes (No. O118/80+OD118 with green filter lid; Saco2, Deinze, Belgium) as described previously[19].

A. thaliana seeds were surface sterilized using 70% EtOH for two minutes and 7% NaOCl containing 0.2% Triton X-100 for eight minutes[88]. Seeds were washed with water six times and stratified for 4 days in the dark at 4°C. Several seeds were individually placed in five positions per microbox to account for unsuccessful seed germination. Additional seedlings were removed one day before inoculation to keep five plants per microbox. Microboxes only contained plants of one genotype and treatment for the initial screening experiment with SynCom-222. Each plant was watered with 200 μl of 0.5x MS, pH 5.8, at 9, 24 and 38 days after germination. A list of plant genotypes and their references can be found in Supplementary Table.

Growth chambers (CU-41L4; Percival, Perry, Iowa, USA) had air flow from the bottom of each shelf and were equipped with full-spectrum lights (Philips Master TL-D 18 W/950 Graphica; Philips, Amsterdam, Netherlands) and UVA-/UVB-emitting light (Sylvania Reptistar F18 W/6500 K; Sylvania, Etobicoke, Ontario, Canada) to simulate natural conditions. The combined light intensity was set to 220 μmol m^{-2} s^{-1} for wavelengths of 400-700 nm (measured with PAR quantum sensor SKP215; Skye Instruments, Powys, UK) and 5.4 μmol m^{-2} s^{-1} for wavelengths of 300-400 nm (UV light measured with MU-250; Apogee, Utah, USA). The growth chambers were set to 22°C and 54% relative humidity and run under an 11-hour light cycle. Due to light activity during the day, the temperature inside microboxes increased to 24°C, and water condensates appeared on sidewalls.

Synthetic community selection

SynCom composition was selected based on the *At*-LSPHERE strain collection[17]. SynCom-222 contained 222 strains (Supplementary Table) and covered maximal phylogenetic diversity to investigate changes in the community assembly on *A. thaliana* genotypes. SynCom-223 was used in a follow-up experiment and contained 223 strains (SynCom-223 additionally included *Serratia* Leaf50, which was left out in SynCom-222 due to concerns of being a strong pathogen, which turned out not to be the case). The V5-V7 region of the 16S rDNA gene amplified with the primers 799F[89] and 1193R[90] allows us

to distinguish 137 of these strains with 100% sequence identity representing ASVs. To be able to attribute changes in strain abundance to individual strains, we selected SynCom-137 containing one strain as a representative for each ASV. To distinguish whether the data shown in the Figures belongs to individual strains or an ASV containing multiple strains (as indicated in Supplementary Table), the ASVs with more than one strain are marked by a superscript circle after the strain name (i.e., *Xanthomonas* Leaf131°) in the axis labels. Experiments with SynCom-137 were always done with a single strain for each ASV.

Synthetic community mixing and plant inoculation

Bacteria were individually streaked on R-2A-agar (Sigma-Aldrich) plates containing 0.5% v/v methanol (R2A-MeOH), grown for 4 days at 22°C, restreaked on fresh R2A-MeOH plates and grown for 3 days. To prepare the SynCom inoculum, a "loopfull" of biomass from each strain was scraped off with a sterile 1-μl plastic loop and resuspended individually in 1 ml of 10 mM $MgCl_2$. Tubes containing the resuspended cells were vortexed for 5 min to disperse cell aggregates. Strains that were not fully resuspended after vortexing were treated with TissueLyzer II (Qiagen) at 30 Hz for 2x 45 sec. The SynCom inoculum was obtained by mixing each strain in an equal volume ratio. The SynCom inoculum was adjusted to an OD_{600} of 0.02, and each plant was inoculated with 200 μl of bacterial solution or 10 mM $MgCl_2$.

Aliquots of 1 ml from the SynCom mixtures were taken for DNA extraction to determine the community composition of the inoculum. Almost all strains could be detected in the SynCom inoculum mixes, and undetected strains were most likely due to insufficient sequencing depths (Extended Data Fig. 1 and Supplementary Table). In addition, a dilution series was made from all resuspended strains, and 4 μl was spotted on R2A-MeOH plates to count colony forming units (CFU) and verify the viability of each strain in the inoculum.

Dropout SynCom inocula was prepared as described previously[19]. The resuspended strains were mixed in an equal ratio for SynCom-137 without (w/o) REPI containing the least number of strains. For inoculum preparation, 10 mM $MgCl_2$ was added to account for "missing strains" compared to SynCom-137 containing most strains. The remaining strains were added at an equal ratio to an aliquot of SynCom-137 w/o REPI to obtain SynCom-137 w/o Leaf131, and $MgCl_2$ was added to account for missing Leaf131 in the inoculum. Similarly, SynCom-137 was assembled by proportionally adding Leaf131 to an aliquot of SynCom-137 w/o Leaf131. SynCom-137 was adjusted to an OD_{600} of 0.02, and all dropout SynComs as well as single strain inoculum *Xanthomonas* Leaf131 were diluted accordingly to ensure an equal number of cells from each strain in every inoculum mix. Different plant genotypes (n=20) were grown in the same microbox to minimize box variations, and treatments were blinded after inoculum preparation to avoid unconscious biases.

Plant sampling and surface sterilization

Plants were inoculated at 10 days after germination and harvested 5-5.5 weeks after germination. Four plants per microbox were harvested for DNA extraction (n=12-16 plants per condition). The aboveground part of the plant was cut with a scalpel under sterile conditions, the cotyledons were removed with sterilized forceps, and the plant was placed

into a 2-ml screw cap sampling tube (FastDNA SPIN Kit for Soil, MP Biomedicals). Plant fresh weight was determined, and phyllosphere samples were immediately frozen in liquid nitrogen. To obtain endosphere samples, plants were placed in regular 2-ml tubes, weighed, and endophytic bacteria were enriched by surface sterilization (in bulk for each genotype) using 70% ethanol for 30 sec and 3% NaOCl for 2 min to lyse epiphytic bacterial cells and degenerate their DNA[75,91]. Plants were washed three times in sterile water, placed in 2-ml screw cap sampling tubes, frozen in liquid nitrogen and stored at -80°C until further processing. Samples were randomized after harvest to avoid batch effects during 16S rDNA gene amplification and sequencing library preparation.

To determine bacterial abundance, plants were harvested as described above, weighed and placed in 2-ml tubes filled with 200 μl of phosphate buffer, pH 7, and a metal bead (5 mm diameter). Plants were homogenized in a TissueLyzer II (Qiagen) at 30 Hz for 2x 45 sec, and a dilution series was plated on R2A-MeOH plates to determine CFU. Mock treated plants using 10 mM $MgCl_2$ were harvested to verify axenic growth conditions.

DNA extraction and 16S rDNA amplicon sequencing

Samples were freeze-dried at -40°C and 0.12 mbar for 16 h (Alpha 2-4 LD Plus, Christ, Osterode am Harz, Germany) and pulverized at 30 Hz for 2x 45 sec (TissueLyzer II, Qiagen). DNA was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals) following the manufacturer's instructions. The samples were transferred to DNA low-binding 96-well plates (Frame Star 96, semiskirted), and the DNA concentration was quantified using dsDNA QuantiFluor (Promega) and normalized to 1 ng/μl. The 16S rDNA amplicon library was generated as previously described[17]. PCR amplification was performed with primers 799F and 1193R using Taq polymerase (Bioron DSF Taq) in triplicate and the following program: 94°C for 2 min, followed by 25 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 60 sec, and a final extension at 72°C for 10 min. Technical triplicates were combined, and 5 μl of PCR run on a 2% agarose gel to visualize the amplification product. The PCR product was cleaned up using Antarctic phosphatase (New England Biolabs, NEB) and exonuclease I (NEB) to digest unused primers. The reaction mix was incubated at 37°C for 30 min and at 85°C for 15 min to inactivate enzymes. Sample 96-well plates were spun down at 4000 g for 10 min to pellet denatured enzymes. Barcoding of 16S rDNA amplicons was performed using barcoded forward and reverse primers as previously described[17]. Barcoding PCR was done in triplicate with the following program: 94°C for 2 min, followed by 10 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 60 sec, and a final extension at 72°C for 10 min. Triplicate reaction mixes were pooled, and the PCR product was visualized on a 2% agarose gel. To assemble the final library, individual samples were quantified according to the band intensity of the bacterial 16S rDNA amplicon on the agarose gel picture using ImageJ to calculate equal amounts of barcoded DNA for each sample. To account for differences between agarose gels, amplicon samples from each agarose gel were first added into separate subpools. DNA from each subpool was cleaned using commercial AMPure or homemade SPRI magnetic beads according to the manufacturer's protocol. Subpools were run on a 2% agarose gel, the bacterial 16S rDNA amplicon band was excised from the gel, and DNA was extracted using a QIAquick Gel Extraction Kit (Qiagen, Venlo, Netherlands). The DNA concentration of each subpool was measured and combined in

volumes according to respective sample numbers. The pooled 16S rDNA amplicon library was cleaned twice with AMPure magnetic beads using a bead-to-DNA ratio of 0.7 to remove small DNA fragments. The amplicon length distribution of the library was assessed on a 2200 TapeStation using HS D1000 (Agilent, Santa Clara, California, USA), resulting in an effective library size of 554-643 bp. Sequencing was performed on a MiSeq desktop sequencer (Illumina) at the Genetic Diversity Centre Zurich using the MiSeq reagent kit v.3 (paired end, 2x 300 bp). Denaturation, dilution and addition of 15% PhiX to the DNA library were performed according to the manufacturer's instructions. Custom sequencing primers were used as described previously[17].

16S rDNA amplicon data processing

16S rDNA reference sequences of SynCom strains were extracted as described previously[19] and each strain is represented by one ASV reference. Paired-end sequencing reads were merged using the USEARCH v.11.0.667-i86 linux64 command fastq_mergepairs[92], with a minimum overlap of 16 bp and a minimum identity of 90%. Merged reads were quality filtered using -fastq_filter with a maximum expected error of 1 and a minimum length of 200 bp. The reads were classified and counted using -otutab with 100% identity to 16S rDNA reference sequences to generate an initial ASV table with a count for each reference per sample. Unclassified sequences of each sample (potentially originating from additional 16S rRNA gene copies of SynCom strains, or from sample or plant contamination, or from DNA artifacts of PCR amplification or sequencing) were dereplicated (-fastx_uniques) and clustered using the UPARSE algorithm (-cluster_otus[93], with a minimum cluster size of 1 and a fixed identity threshold of 97%). *De novo* OTU clusters were annotated using the SILVA SSU Ref NR database (release 132) and added to the initial ASV table. The unclassified reads that could not be assigned to the reference sequences were summed and included in the final ASV table as additional variables (Supplementary Table). 16S rDNA sequencing samples of axenic plants and water controls were used to determine the origin of OTUs from unclassified reads.

16S rDNA amplicon data analysis and visualization

The final ASV table of each experiment was processed in R v.3.6.3 as described previously[19]. Briefly, the ASV table was arranged in compatible datasets, each consisting of samples from a test condition and control. To account for varying sequencing depths between samples, the ASV table was log-normalized and variance-stabilized by DESeq2 v.1.14.194. The data from each SynCom experiment were analyzed separately.

To examine the effect on individual strains between the test and control conditions, the output of DESeq2 provided log2 fold-change values and *p*-values (Wald test, Benjamini-Hochberg adjusted). The differential strain abundances between the test and control conditions were visualized as heatmaps or volcano plots.

To assess the overall effect on communities, PCA was performed with the transformed ASV table using the *prcomp* command. The effect size represents the variance explained by the compared factor and was calculated on Euclidean distances followed by a PERMANOVA to test for statistical significance using the *adonis* command of the package vegan v.2v.5-4.

To summarize the relative abundance of each strain or taxon in a sample, the relative abundance values were calculated by proportional normalization of each sample by its sequencing depth.

The following R packages were used during analysis and visualization: agricolae v.1.3-3[95], ape v.5.4[96], ggplot2 v.3.3.1[97], vegan v.2.5-4[98], DESeq2 v.1.14.1[94].

Phylogenetic analysis

The strain phylogeny was constructed based on full-length 16S rRNA gene sequences (see Supplementary Table for accession numbers). Sequences were aligned using SINA v.1.3.399 and SILVA SSU Ref NR database release 132. To build a maximum likelihood phylogeny from the alignment, PhyML v.3.3.2018214 was used with default parameters[100].

qPCR for bacterial abundance quantification

qPCR assays were run on a QuantStudio 7 Flex using FastStart Universal SYBR green Master mix (Roche Applied Science) in a 20 μl volume containing 300 nM of each primer and approximately 5 ng of DNA per reaction. Primer sequences have been described previously[13]. Bacterial 16S rDNA was amplified using primers 799F (5'-AACMGGATTAGATACCCKG) and 904R (5'-CCCCGTCAATTCITTTGAGTTTYAR) in a 3-step qPCR run with 95°C for 30 sec, 55°C for 60 sec, and 72°C for 60 sec for 40 cycles followed by a melting curve. The reference plant gene *AT4G33380* was used to normalize the 16S rDNA copy number to the number of plant genomes. *AT4G33380* was amplified using primers ExpF1 (5'-ATACAGAAACAACCACCCAAAAG) and ExpF2 (5'-ACGGTACTCCAATTTTCAAGACTC) in a 2-step qPCR run with 95°C for 30 sec and 72°C for 60 sec for 40 cycles followed by a melting curve. We tested the primers for background amplification by titrating increasing amounts of bacterial DNA to plant DNA and did not observe amplification of plant organelle *16S rDNA* within the amplification range Ct12 to Ct32. In addition, linear DNA amplification was tested for both primer pairs with a dilution series of template DNA. Phyllosphere and endosphere samples from 16S rDNA amplicon sequencing were pooled (four technical replicate DNA samples per reaction) for the qPCR run. All samples were run in technical duplicates. The cycle number threshold (Ct) was set by evaluating the levels of background amplification in the water controls. The amplification efficiency (E) was determined by the following equation: $E = -1+10^{(-1/slope)}$. The *16S rDNA* gene copy number was normalized to the plant gene and calculated as described[13]: *16S rRNA*/plant gene $= E_{plant}^{\text{Ct plant}}/E_{16S}^{\text{Ct 16S}}$.

ROS burst measurements

ROS production after treatment with bacterial extracts was measured using a luminol-based assay as described previously[101]. Eight leaf discs (4 mm diameter) per treatment or *A. thaliana* genotype were collected, placed in white 96-well plates (Nr: 655075; Greiner Bio-One) and incubated in sterile water overnight in the dark at 22°C. The water was replaced with a solution containing 17 mg/mL luminol (Sigma-Aldrich, St. Louis, Missouri, USA), 200μg/ml horseradish peroxidase (HRP, type VI-A; Sigma-Aldrich) and either bacterial extracts (1:10 volume ratio of final reaction mix, i.e., 20 μg/ml) or 10 nM flg22 peptide (EZBiolab, Westfield, IN, USA). Luminescence was recorded as photon counts with a 0.5

millisecond exposure time in 90-sec intervals over a total period of 60 min in a Victor3 plate reader (Perkin Elmer, Waltham, MA, USA).

To obtain crude bacterial extracts, bacteria were grown on R2A-MeOH agar plates for three days at 22°C. Cells were resuspended in 10 mM $MgCl_2$ with an OD_{600} between 0.2 and 0.8, and 100 µl of bacterial suspension was plated on R2A-MeOH agar plates. Plates were incubated at 22°C for 3-5 days until bacteria covered the plate in a continuous lawn. Bacteria were scraped off the plate and resuspended in sterile, deionized water, and the OD_{600} was adjusted to 10 in a 1-ml volume. Cells were lysed by boiling at 100°C for 10 min with intermediate vortexing and immediately placed on ice. The bacterial crude extract was obtained by centrifugation of the cell lysate for 10 min at 16,000 g at 4°C, and the supernatant was collected and stored in aliquots at -20°C for further analysis. Total protein concentration of bacterial extracts was determined in a BCA assay and extracts adjusted to 200 µg/ml.

Bacterial infections by leaf infiltration

A. thaliana seeds were sterilized using 70% ethanol for 1 min and stratified for 4 days at 4°C in the dark. Plants were grown in potting soil (Substrat 1, Klasmann-Deilmann, Germany) for 4-5 weeks in growth chambers (CU-41L4, Percival) fitted with full spectrum lights (Philips Master TL-D, 18 W/840) at 22°C and 60% relative humidity under an 11-hour light cycle. Bacteria were grown on R2A-MeOH agar plates for four days, restreaked and incubated for another three days at 22°C. Cells were resuspended in 10 mM $MgCl_2$, and the OD_{600} was adjusted to 0.002 (~10^6 CFU/ml). Three to four leaves of four plants were infiltrated per treatment. Plants were covered with a plastic dome to increase humidity, and infection was monitored for 5 days.

Samples were taken from infiltrated leaves at the indicated time points. Leaves were surface sterilized by submerging them in 70% ethanol for 30 sec and subsequently washed in water three times. Leaf discs were taken with a cork borer (diameter 7.5 mm), and two leaf discs from the same plant were combined in tubes filled with 200 µl of 10 mM $MgCl_2$ and two glass beads (3 mm diameter, Merck, Darmstadt, Germany). Leaf discs were homogenized with TissueLyzer (QIAGEN, Hilden, Germany) at 30 Hz for 2x 45 sec. The bacterial solution was plated as a dilution series on R2A-MeOH agar plates to count CFUs.

Gnotobiotic plant inoculation with single strains

Plants were grown in microboxes as described for SynCom experiments. Bacteria were grown on R2A-MeOH plates and bacterial suspension for inoculum was prepared as described above for the SynCom experiments. Germ-free plants were inoculated with 200 µl bacterial solution of OD_{600} of 0.02 in 10 mM $MgCl_2$. For screening of bacteria that induce a plant phenotype during mono-association on *rbohD*, each inoculum was prepared from a 1:10 dilution of the resuspension of a "loopful" bacterial biomass in 1 ml of 10 mM $MgCl_2$ (final inocula had an OD_{600} between 0.02 and 0.08). The plant phenotype was examined at 25 days after inoculation.

Extended Data
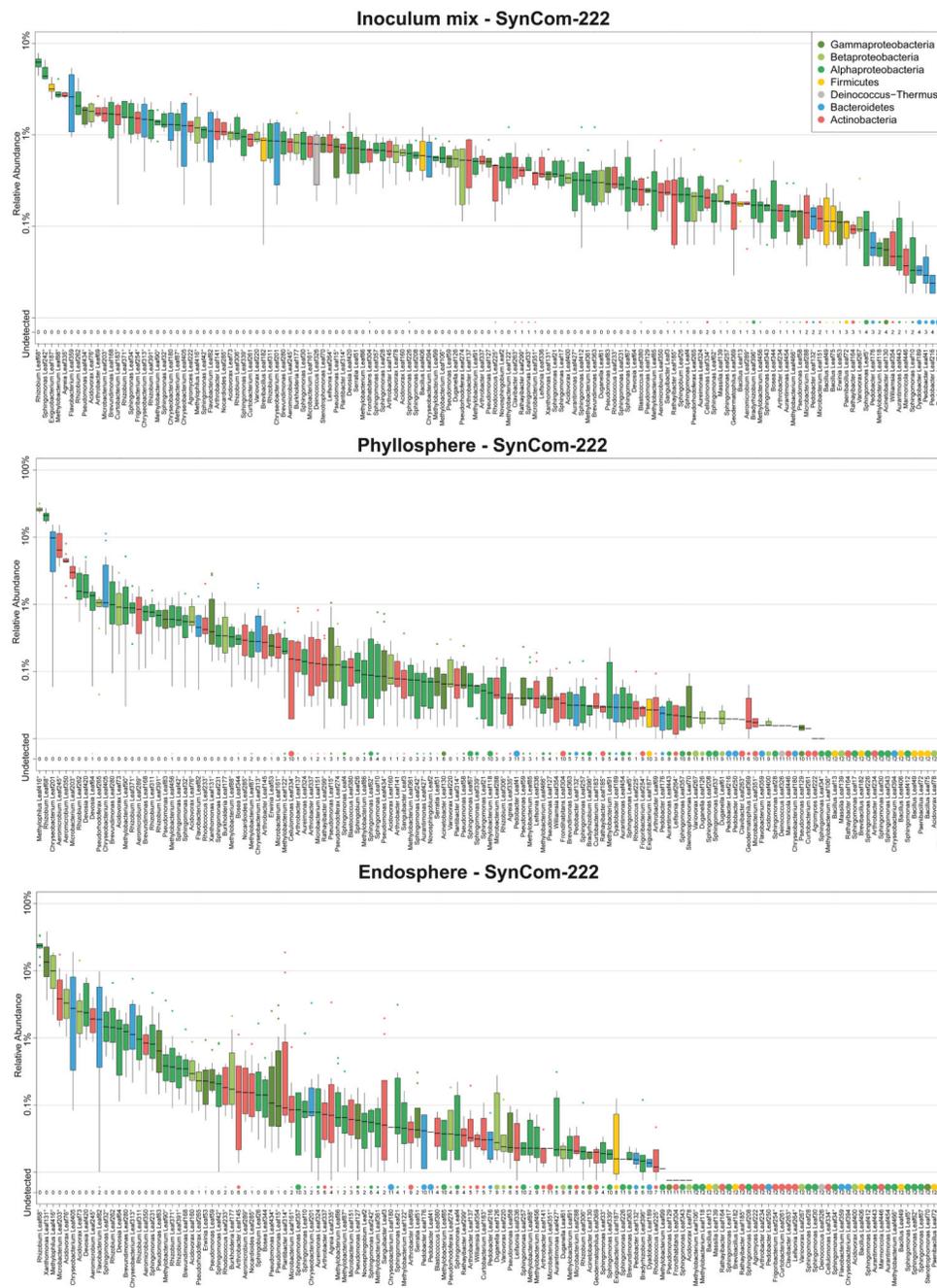
Extended Data Fig. 1. SynCom-222 composition of inoculum, phyllosphere and endosphere.
Relative abundance of strains (or ASVs indicated by superscript circle) determined by 16S
rDNA amplicon sequencing of samples from SynCom-222 inoculum mix, phyllosphere and
endosphere samples of *A. thaliana* Col-0. Box plots show the median with upper and lower
quartiles and whiskers present 1.5x interquartile range. "Undetected" and dot size indicate

the number of plant replicates where a given strain was not detected (n=12). Colors represent strain phylogeny.

a

b
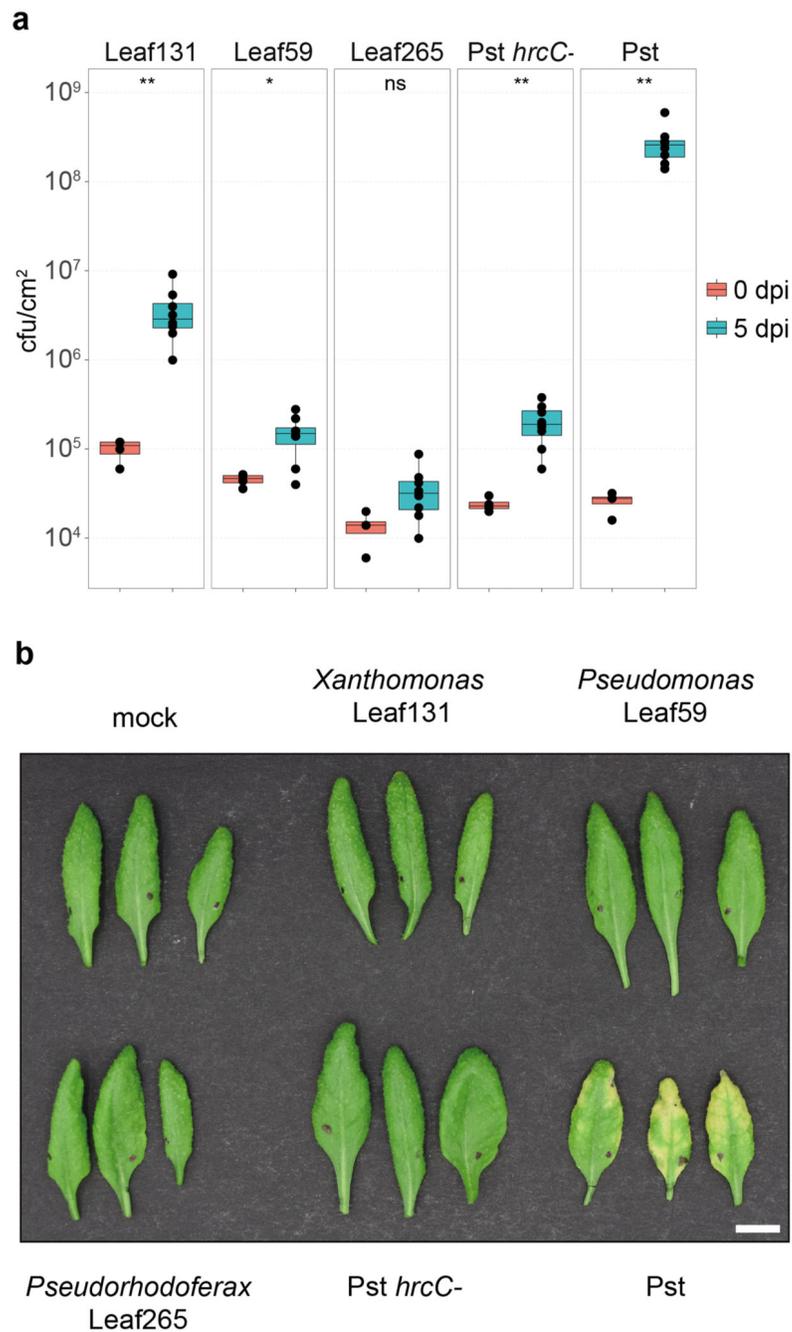
Extended Data Fig. 2. *A. thaliana* **Col-0 leaves after infiltration with bacterial strains.**
a) Bacterial load measured as colony forming units (CFU)/cm^2 at 0 and 5 days post infiltration (dpi). *Xanthomonas* Leaf131, *Pseudomonas* Leaf59, *Pseudorhodoferax* Leaf265, *Pseudomonas syringae* pv. *tomato* DC3000 (Pst) *hrcC-* and wild-type Pst were infiltrated at OD=0.002 (~10^6 CFU/ml) into leaves of soil-grown, four-week-old Col-0 plants. Infiltrated

leaves were harvested, surface sterilized in 70% ethanol and homogenized before serial dilution plating. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. Statistical differences were calculated with two-tailed Mann-Whitney U-test (0 dpi, n=4; 5 dpi, n=8; ns, not significant; * p<0.05, ** p<0.01). b) Photographs of leaves from Col-0 five days after treatments as described above. White bar indicates one cm.
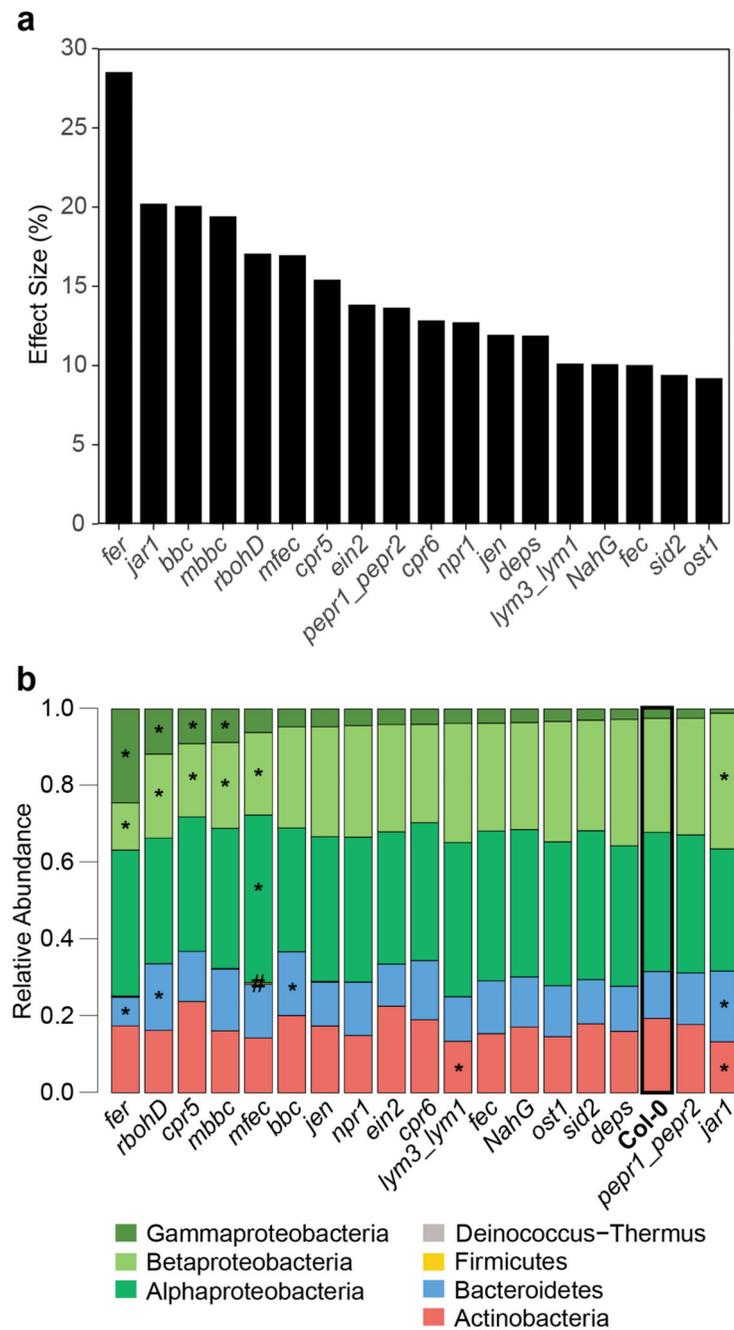
Extended Data Fig. 3. Genotype effect on bacterial community in phyllosphere.

a) Effect of plant genotype on phyllosphere community. The bacterial community of each genotype was compared to Col-0 wild-type in principal component analysis (PCA, n=12) followed by PERMANOVA (permutations=10,000), and the effect size of the genotype was plotted in decreasing order. Effect size represents variance explained by genotype (p-value<0.05, Benjamini-Hochberg adjusted; n=12). b) Relative abundance of phyllosphere bacteria on the indicated plant genotypes. Asterisks (or hash tags for Firmicutes) denote significant differences in taxa on genotypes compared to Col-0 in a two-sided t-test (p<0.05, Benjamini-Hochberg adjusted, n=12).
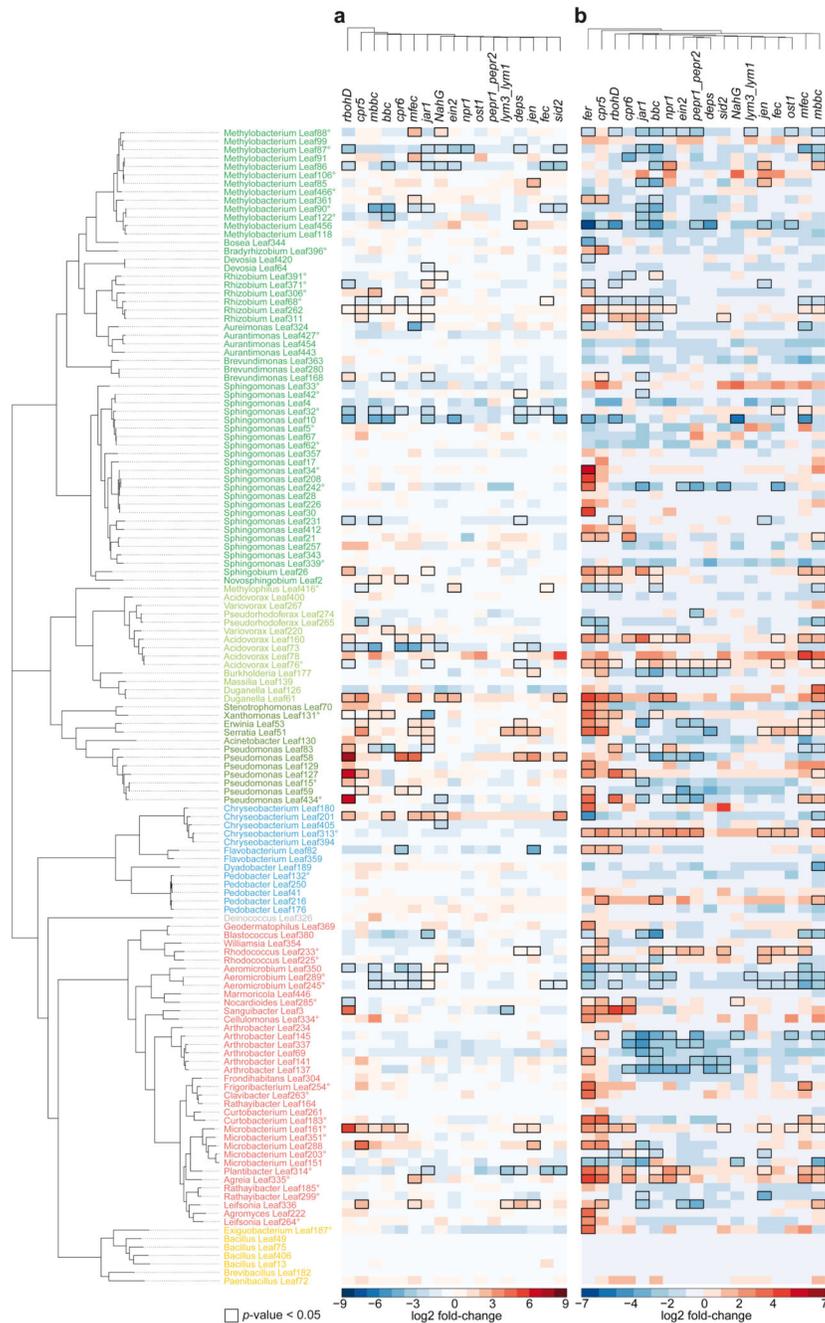
Extended Data Fig. 4. Overview and clustering of community profiles on genotypes versus Col-0.
a) Strain changes in endosphere communities are displayed in a heatmap as log2 fold-change of strains in the endosphere of the individual genotypes versus Col-0 (columns). Strains or ASVs (indicated by superscript circle) of SynCom-222 are ordered and colored by phylogeny. Hierarchical clustering (R command *hclust*, method "single") of genotypes was performed on an Euclidean distance matrix of log2 fold-changes between test conditions and controls. b) Strain changes in phyllosphere communities shown in the heatmap with genotype clustering as described above. Differential strain abundance was calculated using

DESeq2, and statistical significance was expressed with *p*-values (two-sided Wald test, Benjamini-Hochberg adjusted): the black cell rectangle highlights significant changes *p*<0.05.
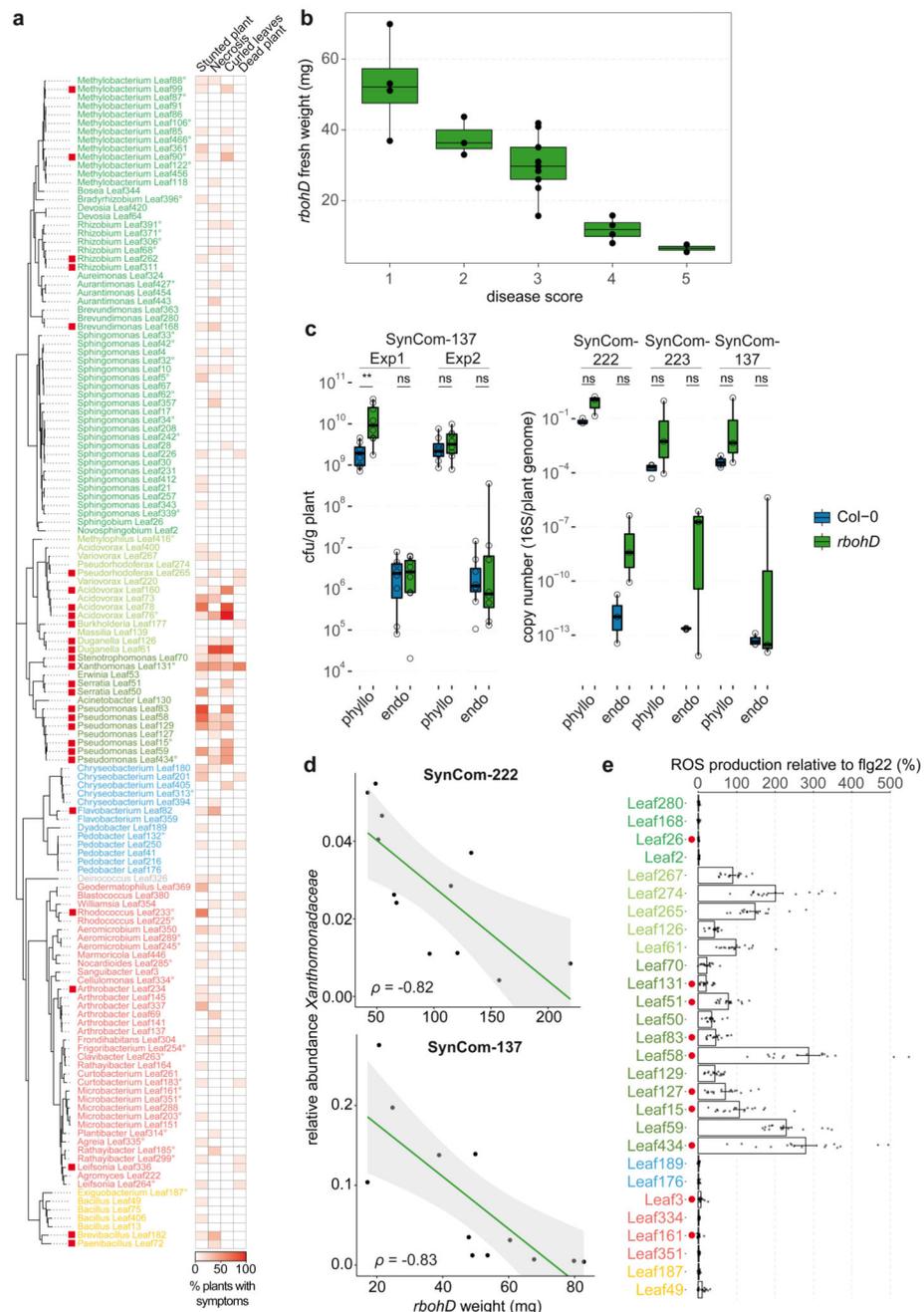
Extended Data Figure 5. Microbiota-induced disease in *rbohD* is linked to the enrichment of *Xanthomonadaceae*.

a) Screening of plant phenotypes and disease symptoms in *rbohD* after colonization with individual bacterial strains. Germ-free, 10-day-old *rbohD* seedlings were inoculated with a bacterial suspension (OD$_{600}$ ranging from 0.02 to 0.08), and the plant phenotype was

examined after 3.5 weeks (n=8). Inoculated plants showed a variety of phenotypes, e.g., stunted plants, necrotic lesions, curled leaves or dead plants. Phenotype-inducing strains were selected based on symptoms and are highlighted by red rectangles in the phylogeny. b) Disease index was assessed in *rbohD* plants: 1, healthy; 2, mild symptoms on individual leaves; 3, stronger symptoms on multiple leaves; 4, strong symptoms on whole plant; 5, severe symptoms or dead plant. The graph displays the plant fresh weight (mg) of SynCom-137-inoculated *rbohD* plants (n=20) per disease category. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. c) Bacterial load in SynCom-137-inoculated Col-0 and *rbohD* plants measured as colony forming units (CFU) per gram of plant fresh weight isolated from the endosphere and phyllosphere (n=16; two-tailed Mann-Whitney *U*-test; ns, not significant; **, $p < 0.01$). Bacterial load in Col-0 and *rbohD* plants inoculated with the indicated SynCom represented by qPCR of the bacterial 16S rDNA gene relative to plant gene. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. Statistical significance was calculated by two-tailed Mann-Whitney *U*-test (n=3, each pool of four DNA samples; ns, not significant). qPCR data for SynCom-222 in Col-0 are the same as in Figure 1c. d) Correlation analysis between the relative abundance of *Xanthomonadaceae* in phyllosphere samples of *rbohD* inoculated with the indicated SynCom and plant fresh weight. The Spearman coefficient ρ (*p*-value<0.01) was calculated using *ggscatter* command (ggpubr, R), grey area shows 95% confidence interval of regression line in green. Correlation data on endosphere samples was not possible due to bulk surface sterilization of plants. e) ROS accumulation was measured in leaf discs with a luminol-based assay after treatment with extracts from heat-killed bacteria. ROS production was recorded for 45 min, and luminescence counts were integrated over time. ROS triggered by individual treatments were normalized to ROS production by 10 nM flg22. Normalized ROS accumulation is shown for each bacterial strain. Barplots show mean and error bars show standard deviation (n=16; combined data from two independent experiments). Red dots indicate *rbohD*-enriched strains.
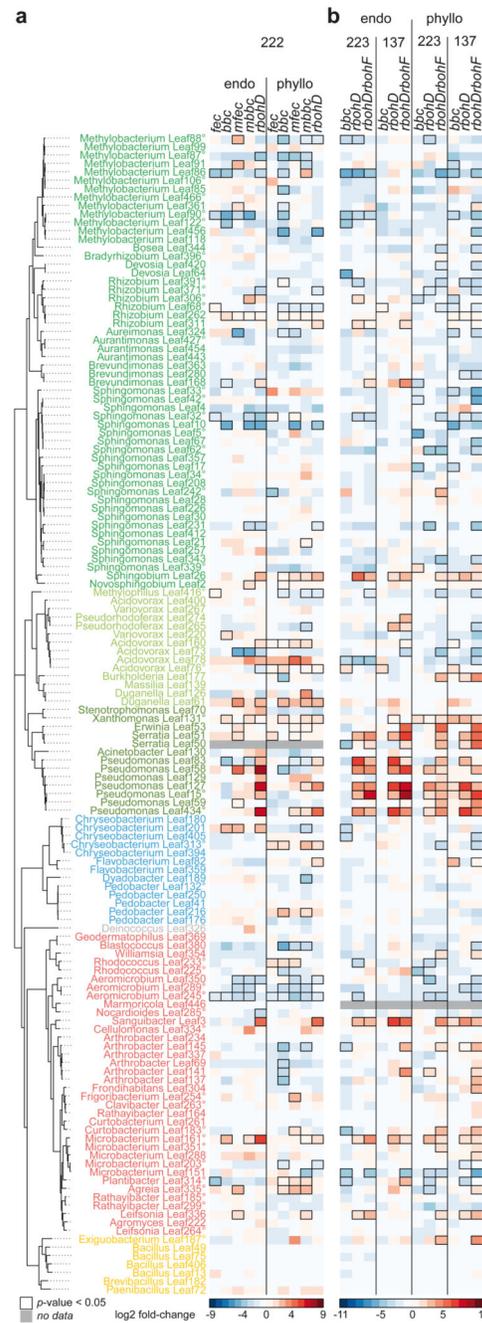
Extended Data Fig. 6. The community assembly of *rbohD* and *rbohDrbohF* substantially differs from that of other PTI genotypes.

The heatmap shows log2 fold-changes in strain abundance on different genotypes compared to Col-0 wild-type. Columns show endosphere and phyllosphere samples from plants inoculated with either a) SynCom-222 or b) SynCom-223 or SynCom-137. Strains (or ASVs indicated by superscript circles) are ordered and colored by phylogeny. Statistical significance of differential strain abundances was calculated with the two-sided Wald test (DESeq2 package, R) and highlighted with a black cell rectangle for *p*<0.05 (n=12; Benjamini-Hochberg adjusted).
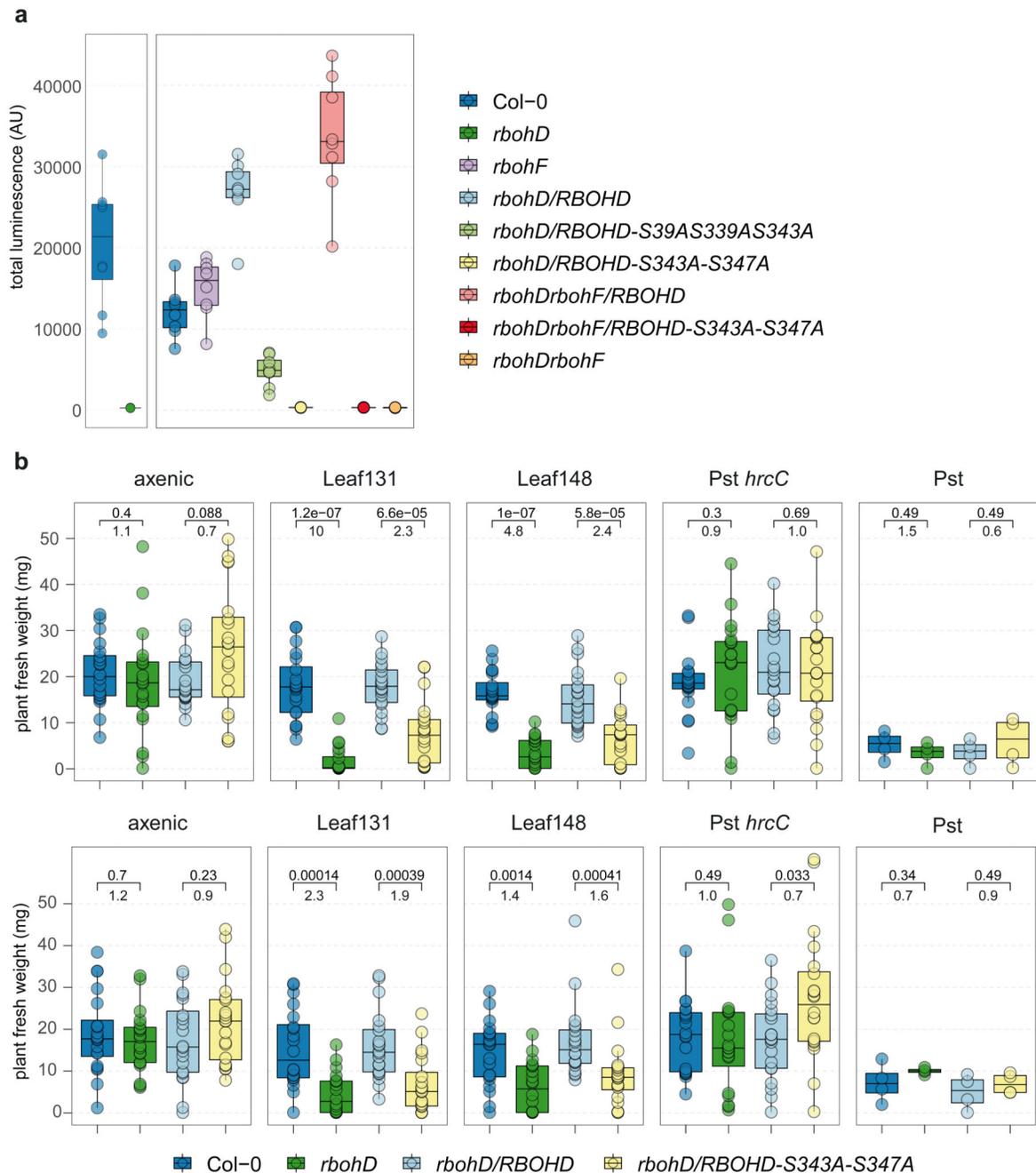
Extended Data Fig. 7. PTI-associated phosphorylation sites of RBOHD are involved in resistance to opportunistic pathogens.

a) ROS production of *A. thaliana* Col-0, *rbohD*, *rbohF*, *rbohD/RBOHD*, *rbohD/RBOHD-S39AS339AS343A*, *rbohD/RBOHD-S343A-S347A*, *rbohD/RBOHD-S343A-S347A*, *rbohDrbohF/RBOHD*, *rbohDrbohF/RBOHD-S343A-S347A* and *rbohDrbohF*, after treatment with 100 nM flg22. ROS production was measured in leaf discs from soil-grown plants with a luminol-based assay and expressed as integrated luminescence over 45 min (AU, arbitrary units). Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. b) Fresh weight of Col-0, *rbohD*, *rbohD/RBOHD*,

and *rbohD/RBOHD-S343A-S347A* inoculated with individual strains of *Xanthomonas* spp. Leaf131, Leaf148, Pst *hrcC-*, and Pst wild-type or mock-inoculated with 10 mM MgCl$_2$. Germ-free 10-day-old seedlings were inoculated with OD=0.02 of the respective strains. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. Significant differences were calculated between the mutants and their respective controls (n=20; two-tailed Mann-Whitney *U*-test). Brackets above bar plots indicate comparison groups with *p*-values displayed above and fold-change below. Data from two independent experiments are shown in separate graphs.
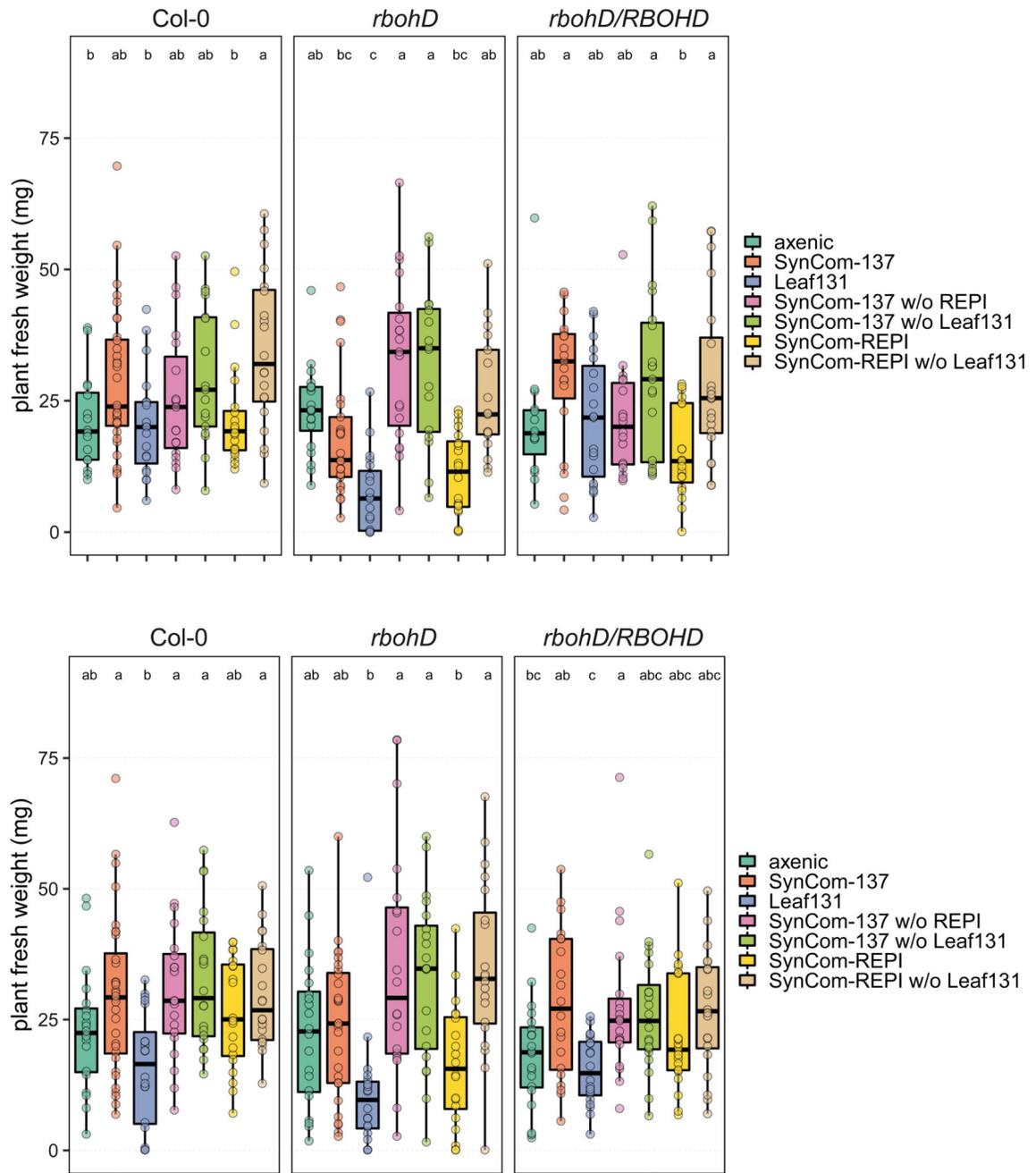
Extended Data Fig. 8. Leaf131 is required and sufficient for disease in *rbohD*.
Two independent replicate experiments with dropout synthetic communities as presented
in Figure 5. In addition, SynCom-REPI without (w/o) Leaf131 was tested and compared
to SynCom-REPI. Significant differences were calculated with ANOVA and Tukey's HSD
post-hoc test (n=20, letters indicate significance groups, α=0.05).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data Availability

Raw reads of demultiplexed 16S rRNA amplicon samples can be found in the European Nucleotide Archive (ENA) under accession number PRJEB44158.

Code availability

The code used to analyse all data and generate figures can be found at https://github.com/MicrobiologyETHZ/phylloR/releases/tag/v1.1. No unpublished algorithms or methods were used.

References

1. Beattie GA, Lindow SE. Bacterial colonization of leaves: a spectrum of strategies. Phytopathol. 1999; 89: 353–359.

2. Vorholt JA. Microbial life in the phyllosphere. Nat Rev Microbiol. 2012; 10: 828–840. [PubMed: 23154261]

3. Pieterse CM, et al. Induced systemic resistance by beneficial microbes. Annu Rev Phytopathol. 2014; 52: 347–375. [PubMed: 24906124]

4. Müller DB, Vogel C, Bai Y, Vorholt JA. The plant microbiota: systems-level insights and perspectives. Annu Rev Genet. 2016; 50: 211–234. [PubMed: 27648643]

5. Lugtenberg B, Kamilova F. Plant-Growth-Promoting Rhizobacteria. Annu Rev Microbiol. 2009; 63: 541–556. [PubMed: 19575558]

6. Berens ML, Berry HM, Mine A, Argueso CT, Tsuda K. Evolution of hormone signaling networks in plant defense. Annu Rev Phytopathol. 2017; 55: 401–425. [PubMed: 28645231]

7. Cook DE, Mesarich CH, Thomma BPHJ. Understanding plant immunity as a surveillance system to detect invasion. Annu Rev Phytopathol. 2015; 53: 541–563. [PubMed: 26047564]

8. Teixeira PJPL, Colaianni NR, Fitzpatrick CR, Dangl JL. Beyond pathogens: microbiota interactions with the plant immune system. Curr Opin Microbiol. 2019; 49: 7–17. [PubMed: 31563068]

9. Yu K, Pieterse CMJ, Bakker PAHM, Berendsen RL. Beneficial microbes going underground of root immunity. Plant, Cell Environ. 2019; 42: 2860–2870. [PubMed: 31353481]

10. Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay between innate immunity and the plant microbiota. Annu Rev Phytopathol. 2017; 55: 565–589. [PubMed: 28645232]

11. Lebeis SL, et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science. 2015; 349: 860–864. [PubMed: 26184915]

12. Castrillo G, et al. Root microbiota drive direct integration of phosphate stress and immunity. Nature. 2017; 543: 513–518. [PubMed: 28297714]

13. Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. PLoS Genet. 2014; 10 e1004283 [PubMed: 24743269]

14. Chen T, et al. A plant genetic network for preventing dysbiosis in the phyllosphere. Nature. 2020; 580: 653–657. [PubMed: 32350464]

15. Horton MW, et al. Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. Nat Commun. 2014; 5: 5320. [PubMed: 25382143]

16. Vorholt JA, Vogel C, Carlstrom CI, Müller DB. Establishing causality: opportunities of synthetic communities for plant microbiome research. Cell Host & Microbe. 2017; 22: 142–155. [PubMed: 28799900]

17. Bai Y, et al. Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature. 2015; 528: 364–369. [PubMed: 26633631]

18. Durán P, et al. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. Cell. 2018; 175: 973–983. e914 [PubMed: 30388454]

19. Carlström CI, et al. Synthetic microbiota reveal priority effects and keystone strains in the *Arabidopsis* phyllosphere. Nat Ecol Evol. 2019; 3: 1445–1454. [PubMed: 31558832]

20. Karasov TL, et al. The relationship between microbial population size and disease in the *Arabidopsis thaliana* phyllosphere. bioRxiv. 2020. 828814

21. Couto D, Zipfel C. Regulation of pattern recognition receptor signalling in plants. Nat Rev Immunol. 2016; 16: 537–552. [PubMed: 27477127]

22. Peng Y, van Wersch R, Zhang Y. Convergent and divergent signaling in PAMP-triggered immunity and effector-triggered immunity. Mol Plant-Microbe Interact. 2018; 31: 403–409. [PubMed: 29135338]

23. Torres MA, Dangl JL, Jones JD. Arabidopsis gp91^{phox} homologues *AtrbohD* and *AtrbohF* are required for accumulation of reactive oxygen intermediates in the plant defense response. Proc Natl Acad Sci USA. 2002; 99: 517–522. [PubMed: 11756663]

24. Gimenez-Ibanez S, Ntoukakis V, Rathjen JP. The LysM receptor kinase CERK1 mediates bacterial perception in *Arabidopsis*. Plant Signal Behav. 2009; 4: 539–541. [PubMed: 19816132]

25. Willmann R, et al. *Arabidopsis* lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. Proc Natl Acad Sci USA. 2011; 108: 19824–19829. [PubMed: 22106285]

26. Yamaguchi Y, Huffaker A, Bryan AC, Tax FE, Ryan CA. PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in *Arabidopsis*. Plant Cell. 2010; 22: 508–522. [PubMed: 20179141]

27. Xin XF, et al. Bacteria establish an aqueous living space in plants crucial for virulence. Nature. 2016; 539: 524–529. [PubMed: 27882964]

28. Staswick PE, Tiryaki I, Rowe ML. Jasmonate response locus *JAR1* and several related *Arabidopsis* genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. Plant Cell. 2002; 14: 1405–1415. [PubMed: 12084835]

29. Alonso JM, Hirayama T, Roman G, Nourizadeh S, Ecker JR. EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. Science. 1999; 284: 2148–2152. [PubMed: 10381874]

30. Clarke JD, Volko SM, Ledford H, Ausubel FM, Dong X. Roles of salicylic acid, jasmonic acid, and ethylene in *cpr*-induced resistance in *Arabidopsis*. Plant Cell. 2000; 12: 2175–2190. [PubMed: 11090217]

31. Tsuda K, Sato M, Stoddard T, Glazebrook J, Katagiri F. Network properties of robust immunity in plants. PLoS Genet. 2009; 5 e1000772 [PubMed: 20011122]

32. Cao H, Glazebrook J, Clarke JD, Volko S, Dong X. The *Arabidopsis* NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell. 1997; 88: 57–63. [PubMed: 9019406]

33. Mustilli A-C, Merlot S, Vavasseur A, Fenzi F, Giraudat J. Arabidopsis OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. Plant Cell. 2002; 14: 3089–3099. [PubMed: 12468729]

34. Bowling SA, Clarke JD, Liu Y, Klessig DF, Dong X. The *cpr5* mutant of *Arabidopsis* expresses both NPR1-dependent and NPR1-independent resistance. Plant Cell. 1997; 9: 1573–1584. [PubMed: 9338960]

35. Clarke JD, Liu Y, Klessig DF, Dong X. Uncoupling *PR* gene expression from NPR1 and bacterial resistance: characterization of the dominant *Arabidopsis cpr6-1* mutant. Plant Cell. 1998; 10: 557–569. [PubMed: 9548982]

36. Stegmann M, et al. The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. Science. 2017; 355: 287–289. [PubMed: 28104890]

37. Deslauriers SD, Larsen PB. FERONIA is a key modulator of brassinosteroid and ethylene responsiveness in Arabidopsis hypocotyls. Mol Plant. 2010; 3: 626–640. [PubMed: 20400488]

38. Suzuki N, et al. Respiratory burst oxidases: the engines of ROS signaling. Curr Opin Plant Biol. 2011; 14: 691–699. [PubMed: 21862390]

39. Torres MA, Dangl JL. Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. Curr Opin Plant Biol. 2005; 8: 397–403. [PubMed: 15939662]

40. Miller G, et al. The plant NADPH oxidase RBOHD mediates rapid systemic signaling in response to diverse stimuli. Sci Signal. 2009; 2: ra45. [PubMed: 19690331]

41. Zhang J, et al. A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. Cell Host & Microbe. 2007; 1: 175–185. [PubMed: 18005697]

42. Qi J, Wang J, Gong Z, Zhou J-M. Apoplastic ROS signaling in plant immunity. Curr Opin Plant Biol. 2017; 38: 92–100. [PubMed: 28511115]

43. Mersmann S, Bourdais G, Rietz S, Robatzek S. Ethylene signaling regulates accumulation of the FLS2 receptor and is required for the oxidative burst contributing to plant immunity. Plant Physiol. 2010; 154: 391–400. [PubMed: 20592040]

44. Denness L, et al. Cell wall damage-induced lignin biosynthesis is regulated by a reactive oxygen species- and jasmonic acid-dependent process in *Arabidopsis*. Plant Physiol. 2011; 156: 1364–1374. [PubMed: 21546454]

45. Hamann T, Bennett M, Mansfield J, Somerville C. Identification of cell-wall stress as a hexose-dependent and osmosensitive regulator of plant responses. The Plant J. 2009; 57: 1015–1026. [PubMed: 19036034]

46. Pogány M, et al. Dual roles of reactive oxygen species and NADPH oxidase RBOHD in an *Arabidopsis-Alternaria* pathosystem. Plant Physiol. 2009; 151: 1459–1475. [PubMed: 19726575]

47. Fagard M, et al. *Arabidopsis thaliana* expresses multiple lines of defense to counterattack *Erwinia chrysanthemi*. Mol Plant-Microbe Interact. 2007; 20: 794–805. [PubMed: 17601167]

48. Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. Dysbiosis and the immune system. Nat Rev Immunol. 2017; 17: 219–232. [PubMed: 28260787]

49. Fryda SJ, Otta JD. Epiphytic movement and survival of *Pseudomonas syringae* on spring wheat. Phytopathology. 1978; 68: 1064–1067.

50. Tsuji J, Somerville SC, Hammerschmidt R. Identification of a gene in *Arabidopsis thaliana* that controls resistance to *Xanthomonas campestris* pv *campestris*. Physiol Mol Plant Pathol. 1991; 38: 57–65.

51. Obrien RD, Lindow SE. Effect of plant-species and environmental-conditions on epiphytic population sizes of *Pseudomonas syringae* and other bacteria. Phytopathol. 1989; 79: 619–627.

52. Mew TW, Kennedy BW. Seasonal variation in populations of pathogenic pseudomonads on soybean leaves. Phytopathol. 1982; 72: 103–105.

53. Jakob K, et al. *Pseudomonas viridiflava* and *P. syringae* - Natural pathogens of *Arabidopsis thaliana*. Mol Plant-Microbe Interact. 2002; 15: 1195–1203. [PubMed: 12481991]

54. Ercolani GL, Crosse JE. Growth of *Pseudomonas phaseolicola* and related plant pathogens in vivo. J Gen Microbiol. 1966; 45: 429.

55. Morales J, Kadota Y, Zipfel C, Molina A, Torres M-A. The Arabidopsis NADPH oxidases RbohD and RbohF display differential expression patterns and contributions during plant immunity. J Exp Bot. 2016; 67: 1663–1676. [PubMed: 26798024]

56. Kadota Y, et al. Quantitative phosphoproteomic analysis reveals common regulatory mechanisms between effector- and PAMP-triggered immunity in plants. New Phytol. 2019; 221: 2160–2175. [PubMed: 30300945]

57. Li L, et al. The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. Cell Host & Microbe. 2014; 15: 329–338. [PubMed: 24629339]

58. Kadota Y, et al. Direct regulation of the NADPH oxidase RBOHD by the PRR-associated kinase BIK1 during plant immunity. Mol Cell. 2014; 54: 43–55. [PubMed: 24630626]

59. Lee D, et al. Regulation of reactive oxygen species during plant immunity through phosphorylation and ubiquitination of RBOHD. Nat Commun. 2020; 11: 1838. [PubMed: 32296066]

60. Kimura S, et al. CRK2 and C-terminal phosphorylation of NADPH oxidase RBOHD regulate reactive oxygen species production in *Arabidopsis*. Plant Cell. 2020; 32: 1063–1080. [PubMed: 32034035]

61. Sumimoto H. Structure, regulation and evolution of Nox-family NADPH oxidases that produce reactive oxygen species. FEBS J. 2008; 275: 3249–3277. [PubMed: 18513324]

62. Bedard K, Krause K-H. The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. Physiol Rev. 2007; 87: 245–313. [PubMed: 17237347]

63. Grasberger H, et al. Increased expression of DUOX2 is an epithelial response to mucosal dysbiosis required for immune homeostasis in mouse intestine. Gastroenterology. 2015; 149: 1849–1859. [PubMed: 26261005]

64. Sommer F, Bäckhed F. The gut microbiota engages different signaling pathways to induce Duox2 expression in the ileum and colon epithelium. Mucosal Immunol. 2015; 8: 372–379. [PubMed: 25160818]

65. Flores MV, et al. Dual oxidase in the intestinal epithelium of zebrafish larvae has anti-bacterial properties. Biochem Biophys Res Commun. 2010; 400: 164–168. [PubMed: 20709024]

66. Ha E-M, Oh C-T, Bae YS, Lee W-J. A direct role for dual oxidase in *Drosophila* gut immunity. Science. 2005; 310: 847–850. [PubMed: 16272120]

67. Xiao X, et al. A Mesh–Duox pathway regulates homeostasis in the insect gut. Nat Microbiol. 2017; 2 17020 [PubMed: 28248301]

68. Winter SE, Bäumler AJ. Dysbiosis in the inflamed intestine. Gut Microbes. 2014; 5: 71–73. [PubMed: 24637596]

69. Halfvarson J, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol. 2017; 2 17004 [PubMed: 28191884]

70. Kniskern JM, Traw MB, Bergelson J. Salicylic acid and jasmonic acid signaling defense pathways reduce natural bacterial diversity on Arabidopsis thaliana. Mol Plant-Microbe Interact. 2007; 20: 1512–1522. [PubMed: 17990959]

71. An S-Q, et al. Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. FEMS Microbiol Rev. 2019; 44: 1–32.

72. Bartoli C, et al. *In situ* relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. ISME J. 2018; 12: 2024–2038. [PubMed: 29849170]

73. Buell CR. Interactions between *Xanthomonas* species and *Arabidopsis thaliana*. The Arabidopsis book. 2002; 1: e0031. [PubMed: 22303203]

74. Jacques M-A, et al. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. Annu Rev Phytopathol. 2016; 54: 163–187. [PubMed: 27296145]

75. Karasov TL, et al. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. Cell Host & Microbe. 2018; 24: 168–179. e164 [PubMed: 30001519]

76. Agler MT, et al. Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLoS Biol. 2016; 14 e1002352 [PubMed: 26788878]

77. Essakhi S, et al. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical type III secretion system. Appl Environ Microbiol. 2015; 81: 5395–5410. [PubMed: 26048944]

78. Jochum L, Stecher B. Label or concept – what is a pathobiont? Trends Microbiol. 2020; 28: 789–792. [PubMed: 32376073]

79. Helfrich EJN, et al. Bipartite interactions, antibiotic production and biosynthetic potential of the *Arabidopsis* leaf microbiome. Nat Microbiol. 2018; 3: 909–919. [PubMed: 30038309]

80. Innerebner G, Knief C, Vorholt JA. Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. Appl Environ Microbiol. 2011; 77: 3202–3210. [PubMed: 21421777]

81. Romero FM, Rossi FR, Gárriz A, Carrasco P, Ruíz OA. A bacterial endophyte from apoplast fluids protects canola plants from different phytopathogens via antibiosis and induction of host resistance. Phytopathol. 2019; 109: 375–383.

82. Ritpitakphong U, et al. The microbiome of the leaf surface of Arabidopsis protects against a fungal pathogen. New Phytol. 2016; 210: 1033–1043. [PubMed: 26725246]

83. Stromberg KD, Kinkel LL, Leonard KJ. Interactions between *Xanthomonas translucens* pv. *translucens*, the causal agent of bacterial leaf streak of wheat, and bacterial epiphytes in the wheat phyllosphere. Biol Control. 2000; 17: 61–72.

84. Vogel C, Bodenhausen N, Gruissem W, Vorholt JA. The *Arabidopsis* leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health. New Phytol. 2016; 212: 192–207. [PubMed: 27306148]

85. Berendsen RL, et al. Disease-induced assemblage of a plant-beneficial bacterial consortium. ISME J. 2018; 12: 1496–1507. [PubMed: 29520025]

86. Yuan M, et al. Pattern-recognition receptors are required for NLR-mediated plant immunity. Nature. 2021.

87. Ngou BPM, Ahn H-K, Ding P, Jones JDG. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. Nature. 2021.

88. Schlesier B, Bréton F, Mock H-P. A hydroponic culture system for growing *Arabidopsis thaliana* plantlets under sterile conditions. Plant Mol Biol Rep. 2003; 21: 449–456.

89. Chelius MK, Triplett EW. The diversity of Archaea and Bacteria in association with the roots of *Zea mays* L. Microb Ecol. 2001; 41: 252–263. [PubMed: 11391463]

90. Bulgarelli D, et al. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. Nature. 2012; 488: 91–95. [PubMed: 22859207]

91. Reinhold-Hurek B, Bünger W, Burbano CS, Sabale M, Hurek T. Roots shaping their microbiome: global hotspots for microbial activity. Annu Rev Phytopathol. 2015; 53: 403–424. [PubMed: 26243728]

92. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinform. 2010; 26: 2460–2461.

93. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013; 10: 996–998. [PubMed: 23955772]

94. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Gen Biol. 2014; 15: 550.

95. de Mendiburu F, Yaseen M. agricolae: Statistical procedures for agricultural research. R package version 1.4.0. 2020.

96. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinform. 2004; 20: 289–290.

97. Wickham, H. ggplot2: Elegant graphics for data analysis. Springer-Verlag; 2009.

98. Jari Oksanen F, et al. vegan: community ecology package. R package version 2.5-6. 2019.

99. Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinform. 2012; 28: 1823–1829.

100. Guindon S, et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010; 59: 307–321. [PubMed: 20525638]

101. Pfeilmeier S, Saur IM-L, Rathjen JP, Zipfel C, Malone JG. High levels of cyclic-di-GMP in plant-associated *Pseudomonas* correlate with evasion of plant immunity. Mol Plant Pathol. 2016; 17: 521–531. [PubMed: 26202381]
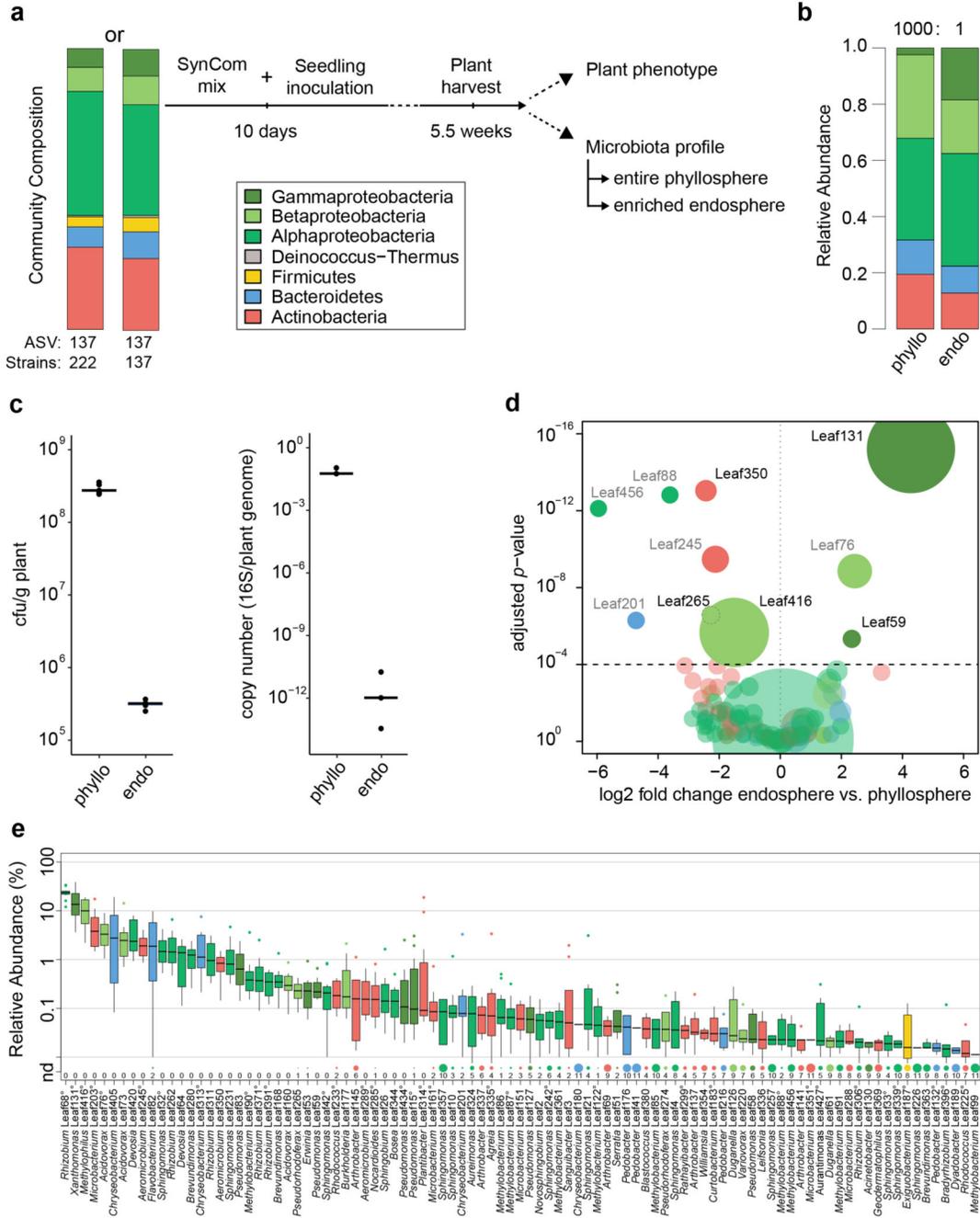
Figure 1. Characterization of the *At*-LSPHERE model community in the phyllosphere and endosphere of *Arabidopsis thaliana*.

a) Schematic workflow of SynCom experiments. Two types of SynCom (SynCom-222 and SynCom-137) with similar taxonomic compositions but different numbers of strains as indicated below the stacked bar plot were used (see Supplementary Table). Colors represent bacterial phyla and classes for Proteobacteria. In the experiments, germ-free *Arabidopsis thaliana* seedlings were inoculated after 10 days, and aboveground plant tissue was harvested 5.5 weeks after germination. Plant phenotype was characterized by scoring disease symptoms and fresh weight measurements. Bacterial community profiles were determined

for the entire phyllosphere community and the endosphere-enriched community by 16S rDNA amplicon sequencing. b) Community composition in the *A. thaliana* phyllosphere ("phyllo") and endosphere ("endo") after inoculation with SynCom-222 on phylum and class level for Proteobacteria. Numbers above bar stacks indicate a 1000-fold difference between bacterial loads in both compartments. c) Overall bacterial load on SynCom-inoculated plants measured by colony forming units per gram plant fresh weight (n=5) and by qPCR of bacterial 16S rDNA relative to plant gene abundance in the phyllosphere and endosphere (n=3). d) Volcano plot shows log2-fold changes in strain abundances of SynCom-222 between phyllosphere and endosphere samples (n=12). Dot size represents relative abundance in endosphere. Statistical significance of differential strain abundance is expressed with *p*-values determined by two-sided Wald test and Benjamini-Hochberg adjusted. Black labels highlight strains that were affected in all three SynCom experiments (see Source Data Figure 1). e) Relative abundance of strains (or ASVs indicated by superscript circles) determined by 16S rDNA amplicon sequencing of the SynCom-222 endosphere. Detected strains with relative abundance >0.01% are displayed (see entire graph in Extended Data Figure 1). Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. "nd" and dot size indicate the number of plant replicates where a given strain was not detected (n=12). Colors represent strain phylogeny.
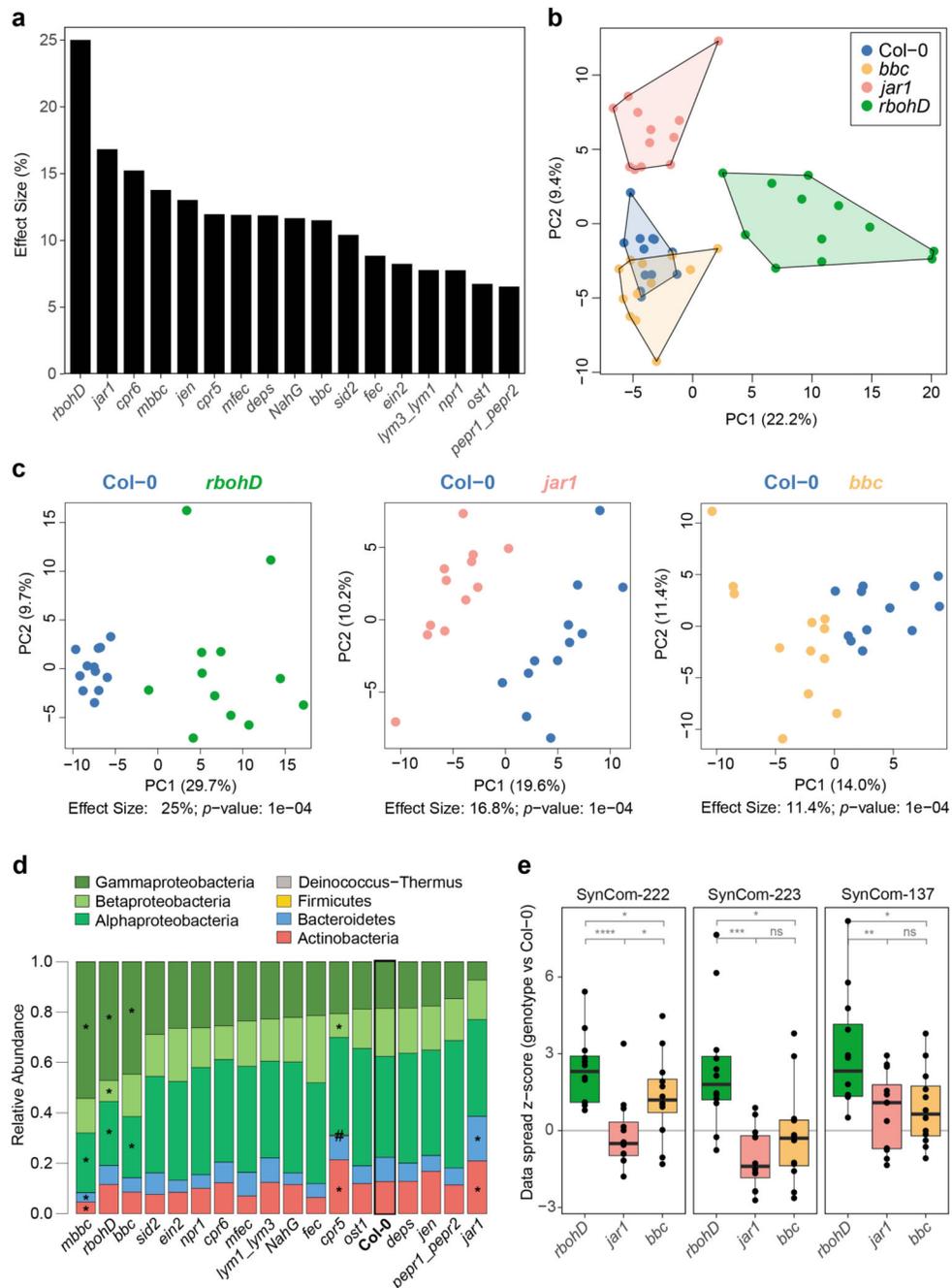
Figure 2. Effect of plant genotype on the leaf endosphere community.
a) The endosphere bacterial community (SynCom-222) of each plant genotype was
compared to Col-0 wild-type plants in principal component analysis (PCA, n=12) followed
by PERMANOVA (permutations=10,000), and the effect size of the genotype was plotted in
decreasing order. b) PCA of bacterial endosphere communities in Col-0 (blue), *bbc* (orange),
jar1 (pink) and *rbohD* (green). PC1 and PC2 are principal components PC1 and PC2 with
their explained variance (%). c) Exemplary PCA of endosphere communities in *rbohD*, *jar1*
and *bbc*. Effect size represents variance explained by genotype, and statistical significance

is expressed with *p*-values determined by PERMANOVA (Benjamini-Hochberg adjusted, n=12). PCA results for all genotypes can be found in Source Data of Fig. 2a. d) Relative abundance of phyla (and classes for Proteobacteria) of endosphere bacteria on the indicated plant genotypes. Genotypes are ordered by decreasing abundance of Gammaproteobacteria. Asterisks (or hash tags for Firmicutes) denote significant differences in taxa on a genotype compared to Col-0 in a two-sided t-test ($p<0.05$, Benjamini-Hochberg adjusted, n=12). e) Community spread of genotypes *rbohD*, *jar1* and *bbc* relative to Col-0 in PCA. The community spread was calculated as the Euclidean distance of data points to the centroid in PCA. Distances of genotypes were normalized to the median distance of Col-0 as the z-score. Community spread was calculated for SynCom-222, SynCom-223 and SynCom-137. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. Statistical differences were determined by two-sided t-test (n=12; ns, not significant; *p<0.05; **p<0.01; ***p<0.001).
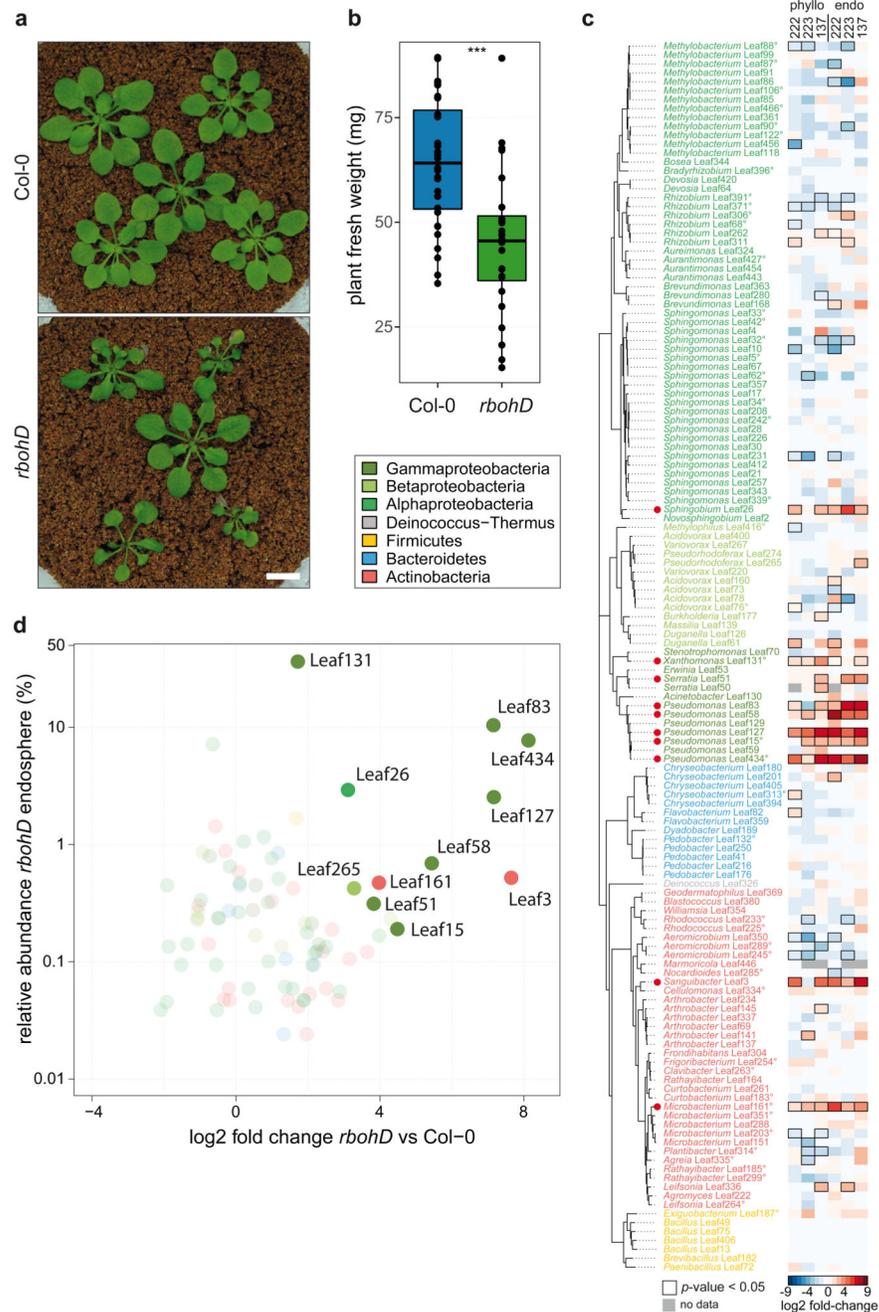
Figure 3. The plant mutant *rbohD* shows a dysbiosis phenotype and assembles a microbiota enriched in Gammaproteobacteria.

a) Representative pictures of five-week-old Col-0 and *rbohD* inoculated with SynCom-137. White bar indicates one cm. b) Fresh weight of aboveground plant tissue of Col-0 and *rbohD* inoculated with SynCom-137. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range (n= 12; two-tailed Mann-Whitney *U*-test; ***, p<0.001). c) Heatmap shows log2 fold-changes of strains in phylogenetic order in *rbohD* compared to Col-0 wild-type plants. Columns show phyllosphere and endosphere samples from plants inoculated with either SynCom-222, SynCom-223 or SynCom-137.

Black rectangles show significant changes, p<0.05 (n=12; Wald test, Benjamini-Hochberg adjusted). Strains (or ASVs indicated by superscript circles) are colored according to phylogeny. Red dots highlight enriched strains in *rbohD* across multiple experiments. d) Volcano plot shows the relative abundance of *rbohD*-enriched strains of SynCom-137 in *rbohD* endosphere (log2-fold-changes in *rbohD* compared to Col-0 with adjusted *p*-value <0.05).
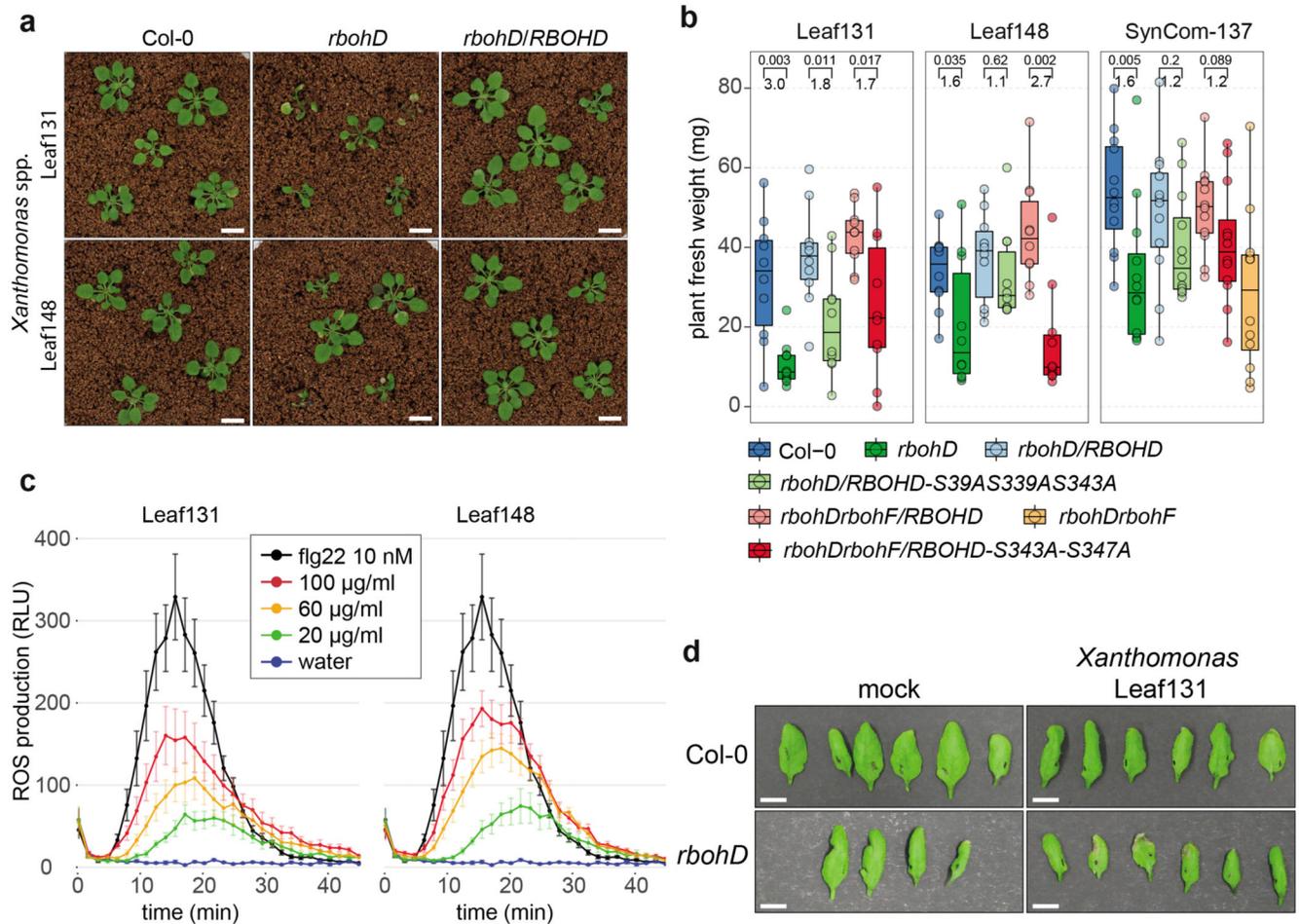
Figure 4. *Xanthomonas* Leaf131 and Leaf148 are opportunistic pathogens in immunocompromised *rbohD*.

a) Pictures of Col-0, *rbohD* and *rbohD/RBOHD* plants inoculated with *Xanthomonas* Leaf131 or Leaf148. Germ-free 10-day-old seedlings were inoculated with bacterial solution at OD=0.02 (~10^7 CFU/ml). White bar indicates one cm. b) PTI-associated phosphorylation sites of RBOHD are involved in resistance to opportunistic pathogens and in microbiota homeostasis. Fresh weight of phyllosphere plant tissue from Col-0, *rbohD*, *rbohD/RBOHD*, *rbohD/RBOHD-S39AS339AS343A*, *rbohDrbohF/RBOHD*, *rbohDrbohF/RBOHD-S343A-S347A* and *rbohDrbohF* after inoculation with either single strain *Xanthomonas* Leaf131 or Leaf148 or SynCom-137. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range. Brackets denote comparison groups, and the number above indicates the *p*-value of two-tailed Mann-Whitney *U*-test and the number below the fold-change in plant weight (n=10). c) Apoplastic ROS production was measured over time in leaf discs with a luminol-based assay after treatment with water, 10 nM flg22 or extracts from heat-killed *Xanthomonas* Leaf131 and Leaf148 adjusted to the indicated total protein concentrations. ROS production is displayed as relative light units (RLU) and curves show the mean (n=8) and error bars the standard error. d) Photographs of leaves from Col-0 and *rbohD* five days after infiltration with 10 mM $MgCl_2$ or *Xanthomonas* Leaf131 at OD=0.002 (~10^6 CFU/ml). White bar indicates one cm.
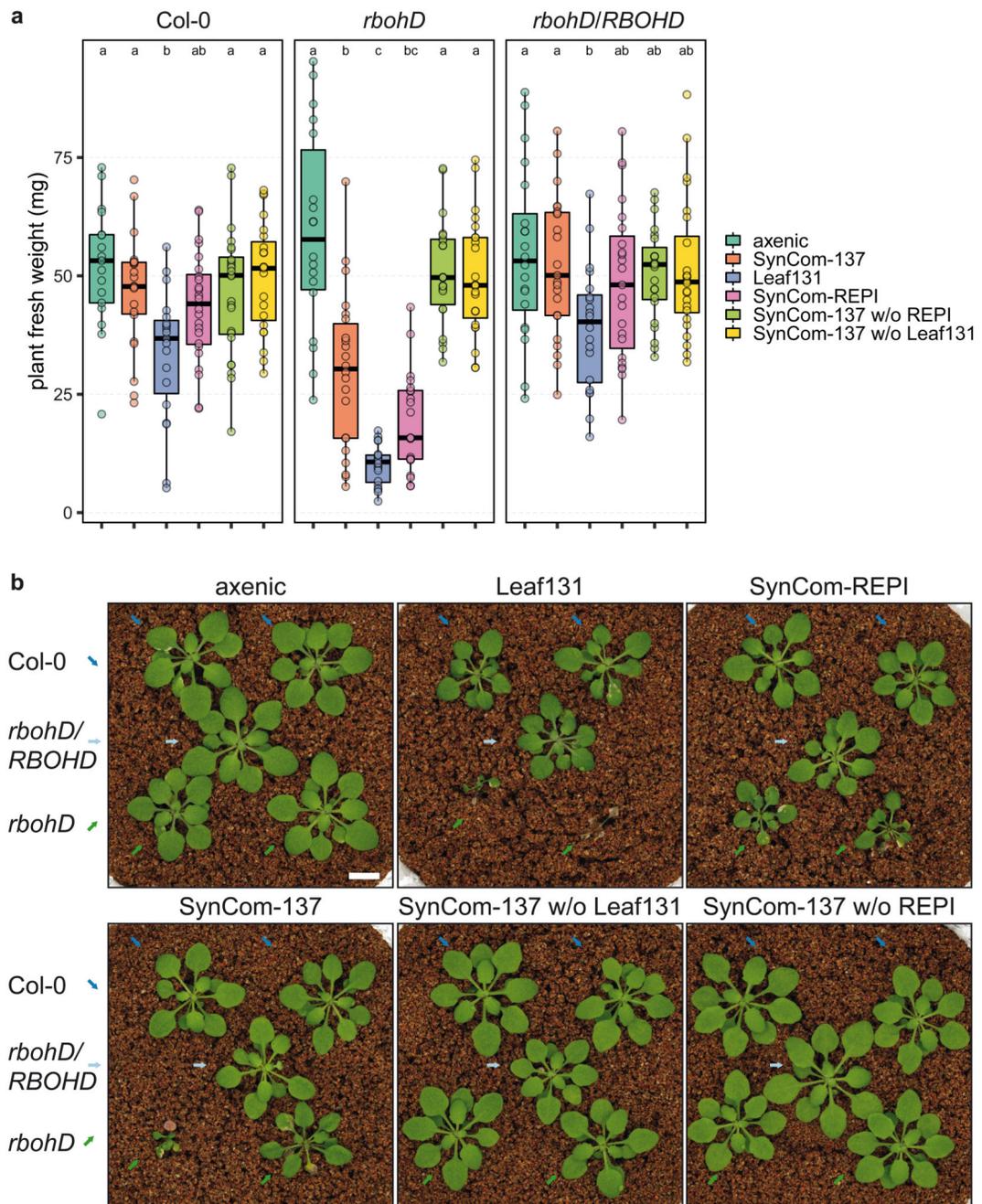
Figure 5. *Xanthomonas* Leaf131 causes dysbiosis in *rbohD*.

a) Fresh weight of Col-0, *rbohD* and *rbohD*/*RBOHD* plants inoculated with either 10 mM MgCl$_2$ (axenic), SynCom-137, single strain Leaf131, SynCom-REPI (containing 32 *rbohD*-enriched and phenotype-inducing strains), SynCom-137 without (w/o) Leaf131, SynCom-137 w/o REPI. Germ-free 10-day-old seedlings were inoculated with OD=0.02 of SynCom-137, and other inocula with lower strain numbers were diluted with 10 mM MgCl$_2$ to obtain equal amounts of cells from each strain in the inoculum. Box plots show the median with upper and lower quartiles and whiskers present 1.5x interquartile range.

Significant differences were calculated with ANOVA and Tukey's HSD post-hoc test (n=20, letters indicate significance groups, α=0.05). b) Pictures of plants treated as indicated above. Genotypes Col-0, *rbohD* and *rbohD/RBOHD* were grown in the same microbox in rows, and different genotypes are highlighted by colored arrows (Col-0, blue; *rbohD/RBOHD*, light blue; *rbohD*, green). White bar indicates one centimeter. The experiment was repeated twice with additional treatments (see Extended Data Figure 8).