MATERIALS AND METHODS

In-vitro plant growth conditions

Arabidopsis thaliana Columbia 0 (Col-0) seeds were surface-sterilized for 10 min with 70% bleach + 0.2% Tween-20, rinsed three times with sterile distilled water, and stratified at 4 °C for two days in the dark. Between 40 to 50 seeds were sowed on vertical square plates with half-strength MS supplemented with 0.5 g/L MES + 5 g/L sucrose + 10 g/L Bacto-agar (Difco, BD, Franklin Lakes, NJ, USA) with a final pH of 5.6-5.7. After 7 days, 10 seedlings were transferred to vertical square plates with half-strength MS supplemented with 0.5 g/L MES + 10 g/L Bacto-agar, with or without the synthetic community (SynCom) and the different glyphosate treatments with a final pH of 5.6-5.7. For the dose standardization experiment, $3.6x10^{-3}$ and $3.6x10^{-6}$ g a.e./L were assessed; for all the rest of the experiments, $3.6x10^{-6}$ g a.e./L was used as low dose of glyphosate (LDG). Each experiment included a no-glyphosate control treatment and an uninoculated control. Plates were set randomly in a growth chamber with a 16-h dark/8-h light regime at 21 °C day/18 °C night for 12 days.

Bacterial culture and plant-inoculation

The bacterial isolates used were previously obtained and sequenced in a previous study by Levy *et al.* [29]. Seven days before each experiment, glycerol stocks from each isolate were inoculated in 400 µL KB medium in a 96 deep well plate. The plates were incubated at 28°C at 250 rpm. After 5 days, 40 µL of the liquid culture was transferred to a new 96-well plate with fresh 400 µL KB medium and grow under the same conditions for 2 days. Ultimately, 7-day and 2-day liquid cultures were combined for each isolate. This procedure accounts for variable growth rates and aims to secure that non-stationary cells are present in the final inoculum, as previously describe by

Finkel *et al.* [12]. The number of cells for each isolate was normalized to an optical density at 600 nm (OD600) equal to 1 (Infinite M200 Pro plate reader, TECAN, Männedorf, Switzerland) in the final pool. The mixed culture was washed three times with 10 mM MgCl2 to remove the media and debris. Then, the washed mixed culture was diluted to a final OD600 of 0.2, and 100 µL of the inoculum was spread on 12 X 12 cm vertical square agar plates with the corresponding medium before transfer the seedlings.

DNA extraction

Roots were pooled from 6-8 plants for each plate and placed in 2.0 ml Eppendorf tubes with three sterile glass beads. Then, the samples were washed three times with sterile distilled water and frozen with liquid nitrogen. Roots were lyophilized for 48 hours (Labconco freeze dry system, Kansas City, MO, USA) and pulverized (tissue homogenizer MPBio, Munich, Germany). Agar from each plate was collected in 60 ml syringes with sterilized Miracloth (Millipore, Burlington, MA, USA) at the tip and store at -20 °C. After one week, syringes were thawed at room temperature and then gently pressed through the Miracloth into 50 ml tubes. The liquid samples were centrifuged at 4200rpm for 20 min and the supernatant was discarded. The remaining liquid with the pellet was transferred into a 2.0 ml Eppendorf tube, centrifuged, all supernatant was removed, and the pellet stored at -80 °C. DNA from root samples and pellets from agar were carried out using 96-well-format DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) following the manufacturer's instruction. Samples were randomized in the plates and maintained throughout library preparation and sequencing.

Library preparation and sequencing

Library preparation was done accordingly to [30] using a dual index approached. V3-V4 region of the bacterial 16S rRNA gene was amplified using the primers 338F (5′-ACTCCTACGGGAGGCAGCA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′). PCR reaction was performed as follows: 1.5 µL of 10 µM of each primer, 1 µL of 10 µM mitochondrial PNA, 1 µL of 10 µM plastid PNA, 6.5 µL PCR grade water, 12.5 µL Kappa Mastermix (Roche, Indianapolis, IN, USA) and 1 µL gDNA template; temperature cycling: 3 min. 95° C, 20 cycles of 15 sec. 95° C, 15 sec. 78° C, 15 sec. 50° C and 15 sec. 72° C. PCR reactions were made in triplicate and amplification was checked on 1.5% agarose gels at 100 V for 35 min. The triplicate reactions were pooled and purified using AMPure XP magnetic beads (Beckman Coulter, High Wycombe, UK) and quantified with Qubit BR DNA assay (Invitrogen, Carlsbad, CA, USA). Libraries were pooled in equal amounts and then diluted to 10 pM for sequencing on an Illumina MiSeq instrument (Illumina, San Diego, CA, USA) using a 600-cycle V3 chemistry kit.

Amplicon sequence data processing

Reads with 100% correct primer sequences were merged (MT-Toolbox) [31] and quality filtered (Sickle) [32] for Q-score > 20. The merge sequences were globally aligned to the 16S rDNA sequences of the 185 isolates in the SynCom (USEARCH v7.1090) [33] and were classified into 97 unique sequences (USeq). A USeq is a cluster of 100% identical sequences coming from a single or multiples isolates, as previously done by [12]. Match between USeq sequences and strain identities could be found in supplementary Table S1. On average, 80% of sequences assigned to an expected USeq. The unmapped sequences were clustered into Operational Taxonomic Units (OTUs) using UPARSE [34] at 97% identity. Representative OTU sequences were taxonomically

annotated with the RDP classifier [35] trained on the Greengenes database [36] (4 February 2011). Arabidopsis organellar and known bacterial contaminants were removed using the option 'usearch_global' at a 98% identity threshold (USEARCH v7.1090) [33]. USeq mapped sequences and OTU counts were used to produce a combined abundance table. The table was processed and analyzed with functions from the ohchibi package (https://github.com/isaisg/ohchibi).

The resulting count table was rarefied to 1,000 reads per sample. Beta diversity was analyzed with a Canonical Analysis of Principal Coordinates (CAP) based on Bray-Curtis dissimilarity calculated from the relative abundance matrices. The Fraction:Dose interaction analysis was performed constraining for the replica effect. Also, a Permutational Multivariate Analysis of Variance (PERMANOVA) was performed using the adonis function from vegan package v2.5-3 [37].

To establish the enrichment profiles in the comparison Low Dose of Glyphosate vs. No Glyphosate from each fraction, we employed the package DESeq2 v1.22.1 [38] to run the model: Abundance ~ Dose + Rep using the raw USeq/OTU combined count table. A USeq/OTU was considered statistically significant if it had a false discovery rate (FDR) adjusted p-value < 0.05.

Phylogenetic tree

The phylogenetic tree of the SynCom isolates was previously constructed by Finkel *et al.* [12]. We selected the same 47 markers and the same approach to create a super alignment and to infer the phylogeny utilizing the WAG model of evolution (FastTree v2.1) [39]. Then, we used the web-based tool (https://itol.embl.de/) to visualize the tree and to add the information of main root

elongation from each isolate in mono-association available in supplementary data S4 (https://www.biorxiv.org/content/10.1101/645655v1.supplementary-material) by [12].

Growth assessment

Shoot and root growth were measure after twelve days post-transferring to the specific media as described in the "*In vitro* plant growth conditions" section. For main root elongation, plates were imaged using a document scanner, and the primary root length from each plant was measured using ImageJ. Shoot was harvest for dry weight. Six to eight shoots, from one plate, were put in a preweighed 2.0 ml Eppendorf tubes and placed in an oven at 60° C for 72 h when the weight of the tubes was stable. To calculate the dry weight, the initial weight of the tube was subtracted from the weight of the tube with the shoot after 72 h and divided to the number of shoots placed in each tube. For the experiment presented in Figure 1, main root elongation and shoot dry weight were assessed, while for dose standardization (Supplementary Figure S1) and drop-out experiments (Figure 3 and Figure 4) only shoot dry weight was used.

Statistical analysis

Analyses of variance (ANOVA), controlling for the replicate effect, was used for Figure 1B and S1. Differences between treatments were shown using the confidence letter display (CLD) derived from the Tukey's post hoc test (package emmeans) [40].

Statistical analysis used in Figure 2 is explained in "Amplicon sequence data processing" section. For Figures 3 and 4, differences between No Glyphosate and Low Dose were analyzed using Student's t-test and adjusting the *p*-values for multiple testing using false discovery rate (FDR)

performed in the R package [41,42]. The number of replicates is given in the respective figure legends.

REFERENCES

12. Finkel, O. M., Salas-Gonzalez, I., Castrillo, G., Conway, J. M., Law, T. F., Teixeira, P. J. P. L., ... & Dangl, J. L. (2020). A single bacterial genus maintains root development in a complex microbiome. *bioRxiv*, 645655.
29. Levy, A., Gonzalez, I. S., Mittelviefhaus, M., Clingenpeel, S., Paredes, S. H., Miao, J., ... & Alvarez, B. R. (2018). Genomic features of bacterial adaptation to plants. *Nature genetics*, *50*(1), 138.
30. Gohl, D. M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., ... & Knights, D. (2016). Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nature biotechnology*, *34*(9), 942.
31. Yourstone, S. M., Lundberg, D. S., Dangl, J. L., & Jones, C. D. (2014). MT-Toolbox: improved amplicon sequencing using molecule tags. *BMC bioinformatics*, *15*(1), 284.
32. N. Joshi, J. Fass, Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle. (2011).
33. Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460-2461.
34. Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods*, *10*(10), 996.

35. Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*, *73*(16), 5261-5267.
36. DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.*, *72*(7), 5069-5072.
37. J. Oksanen *et al.*, "Package 'vegan'" (2015).
38. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 550.
39. Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*, *5*(3), e9490.
40. Package "emmeans" Type Package Title Estimated Marginal Means, aka Least-Squares Means (2019), doi:10.1080/00031305.1980.10483031
41. Team, R. C. (2013). R: A language and environment for statistical computing.
42. Allaire, J. (2012). RStudio: integrated development environment for R. Boston, MA, 770.