PNAS

Supporting Information for

Bacterial tolerance to host-exuded specialized metabolites structures the maize root microbiome

Lisa Thoenen[1,2], Caitlin Giroud[2], Marco Kreuzer[3], Jan Waelchli[2], Valentin Gfeller[1], Gabriel Deslandes-Hérold[1], Pierre Mateo[1], Christelle A.M. Robert[1], Christian H. Ahrens[4], Ignacio Rubio-Somoza[5], Rémy Bruggmann[3], Matthias Erb[1], Klaus Schlaeppi[1, 2, *]

[1] Institute of Plant Sciences, University of Bern, 3013 Bern, Switzerland
[2] Department of Environmental Sciences, University of Basel, 4056 Basel, Switzerland
[3] Interfaculty Bioinformatics Unit, University of Bern, 3012 Bern, Switzerland
[4] Agroscope, Molecular Ecology and Swiss Institute of Bioinformatics, Zurich, 8046 Zürich, Switzerland
[5] Molecular Reprogramming and Evolution (MoRE) Lab, Centre for Research in Agricultural Genomics (CRAG), Carrer Vall Moronta Edifici CRAG, 08193, Barcelona, Spain

* Corresponding author: Klaus Schlaeppi, klaus.schlaeppi@unibas.ch

This PDF file includes:

- Supplementary Methods
- Supplementary Figures S1 to S14
- Supplementary Tables S1 to S5
- Supplementary Datasets S1 and S2 (Captions)
- References Supplementary Information

Other supporting materials for this manuscript include the following:

- Datasets S1 and S2

Supplementary Methods

Bacterial extracts

Various types of bacterial extracts were prepared for establishing the culture collection and quantifying root bacterial community size. First, bacterial extracts from 'dirty roots' (marked as 'RoRh' in Dataset S1) were prepared from 10 cm long root fragments (corresponding to the depth of -1 to -11 cm in soil) that were chopped into small pieces with a sterile scalpel after shaking off loose soil. These root fragments with firmly attached rhizosphere soil were then placed into 50 mL centrifuge tubes containing 10 mL sterile magnesium chloride buffer and Tween20 (10 mM $MgCl_2$ + 0.05% Tween; Sigma-Aldrich) for homogenization with a laboratory blender (Polytron, Kinematica, Luzern, Switzerland; 1 minute at 20'000 rotations per minute) followed by additional vortexing for 15 seconds. Extracts of washed roots (marked as 'root' in Dataset S1) were prepared analogously, except that the roots were washed twice in 50 mL centrifuge tubes with 25 mL of sterile deionized water and shaking the tubes 30 times vigorously to wash off the rhizosphere before cutting them in small pieces for homogenization. The rhizosphere fraction of the washing step was pelleted by sedimentation (supernatant was discarded) and resuspended in 5 mL $MgCl_2$-Tween (10 mM, 0.05%) to prepare the rhizosphere extracts for plating (marked as 'rhizo' in Dataset S1). Plating extracts from soil (marked as 'soil' in Dataset S1) were prepared by mixing 5 g of soil from the pot experiment with 5 mL $MgCl_2$-Tween (10 mM, 0.05%) and vortexing for 15 seconds.

Quantification of root bacterial community size with plating

We quantified the sizes of the root bacterial communities of B73 and *bx1*(B73) (1) plants in two greenhouse experiments (Table S2). In both experiments, one half of the roots was freshly used for plating the cultivable bacteria and the other half stored at -80 °C for culture-independent quantitative PCR (qPCR) analysis (see below). The first experiment consisted of 6-week-old plants and 7-week-old plants (same as 'Isolation 4') were analyzed in the second experiment. Extracts of washed roots were freshly prepared as described above, serially diluted for plating and 20 µL were plated on 10% TSA (tryptic soy medium amended with 15 g/L agar; both Sigma-Aldrich) plates containing cycloheximide (10 mg/L, Sigma-Aldrich). Plates were tilted to spread the 20 µL drops for counting, then incubated for six days at room temperature (23 °C). The forming colonies were counted, multiplied by the dilution factor and the volume plated, and then normalized with the sample fresh weight. The colony forming unit (CFU) data was transformed with log10 prior to statistical analysis (T-test) and visualization.

Quantification of root bacterial community size with qPCR

Complementary to CFU plating, we quantified bacterial community size on the second half of root samples using qPCR analysis. The frozen roots were lyophilized, and DNA was extracted using the Nucleo-Spin Soil DNA extraction kit (Macherey-Nagel, Düren, Germany) following the manufacturer's protocol. Additionally, we also utilized available DNA samples of our previous field

experiments in Changins (2), Reckenholz and Aurora (both (3)) for qPCR analysis. For all DNA samples, the concentration was measured using the AccuClear® Ultra High Sensitivity dsDNA Quantification Kit (Biotium, Fremont, United States) and adjusted to 1 ng/µL. qPCR reactions were set up in a total volume of 20 µL containing HOT FIREPol EvaGreen qPCR Mix Plus (Solis Biodyne, Tartu, Estonia), 250 nM of each primer, 0.3% bovine serine albumin, and 10 ng of root DNA. The size of the bacterial community was quantified on genomic DNA based on the bacterial 16S rRNA gene (primers 799F and 904R, Table S3) relative to the maize gene Actin (primers ZmActin1_F and ZmActin1_R1, Table S3). No-template-control reactions containing water were run in parallel as negative controls. qPCR reactions were set up (in triplicates for greenhouse experiments and in single reactions for samples from field experiment) using the Myra Liquid Handler (Bio Molecular Systems, Upper Coomera, Australia) and ran on a CFX96 Real Time System (Bio Rad, Hercules, California). The cycling program included an initial denaturation step at 95 °C for 15 min, followed by 80 cycles of 95 °C for 15 s, 63 °C for 40 s and 72 °C for 20 s, a hold phase at 72 °C for 10 min, followed by melting curve analysis (temperature incrementally increased by 0.5°C from 65 to 95 °C with steps held for 5 s). Raw data were exported directly from Bio-Rad CFX Manager 3.1 and imported into LinRegPCR version 2016.0 (4) to determine cycle threshold (Ct) and efficiency (E) using the default baseline limit option. The bacterial 16S rRNA gene signal was normalized to the plant signal using the following formula: 16S rRNA/plant gene = $E_{\text{plant gene}}$^$Ct_{\text{plant gene}}$/E_{16S}^Ct_{16S}, where Ct values of the individual reactions and mean E values over all reactions of a given primer pair and run were used for calculation (5). Data was transformed with log2 prior to statistical analysis (T-test) and visualization.

MRB isolate identification

The taxonomy of the purified isolates of the maize root bacteria (MRB) collection were identified by sequencing parts (base pairs 27 to 1492) of the 16S rRNA gene using Sanger technology. Liquid cultures were diluted 1:10 or 1:100 in sterile water and used as template for PCR. The PCR reactions were set up as follows: 15 µL sterile water, 15 µL 2x DreamTaq buffer (Thermo Fisher Scientific, Waltham, USA), 1.5 µL of each primer (stock concentration 10 µM, 27f and 1492r; sequences in Table S3) and 2 µL of the diluted liquid culture as DNA template. For some bacteria, the DNA was extracted using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich). PCR was performed in a Biometra T-advanced cycler according to the following program: 95°C for 3 min, 30 cycles with 95°C for 15 s, 55°C for 15 s and 72°C for 45 s followed by final elongation at 72°C for 5 min. PCR products were verified on an agarose gel (1%; Sigma-Aldrich) and sent for Sanger sequencing with the primers 1492r and/or 27f (Microsynth, Balgach, Switzerland). Sanger sequences were blasted against the NCBI database (National Center for Biotechnology Information, Rockville Pike, USA) for species identification. All metadata, sequences, and taxonomies of the MRB culture collection are listed in the Dataset S1.

Mapping MRB isolates to microbiome profiles

Rationales: We mapped the 16S rRNA gene (Sanger) sequences of the MRB strains to the 16S rRNA gene sequences of published maize community profiles. The first purpose was to investigate abundance of community members corresponding to MRB strains in profiles of maize roots from where the strains were isolated from. For this we mapped the MRB sequences to the data of the 'feedback experiment' reported in Hu et al. 2018 (2). This was a greenhouse experiment with pots filled with natural field soil from the Changins site. Second, to study presence and abundance of community members (corresponding to MRB strains) in root profiles of field grown maize, we mapped the MRB sequences to the datasets of Changins ('field experiment' reported in 2), Reckenholz and Aurora (3) and the pot experiment in Sheffield (6). The third purpose was to examine differential abundance of the MRB strain corresponding community members in profiles of BX-producing vs BX-deficient plants (the field data of Changins (2) also includes profiles of the *bx1* mutant maize line).

Bioinformatics: Because the published datasets (2, 3, 6) utilized different bioinformatic approaches, we re-processed the deposited raw sequence data to have uniformly analyzed microbiome data, to which we then mapped the MRB strains (see below). The raw sequence reads were quality checked with FastQC (7), demultiplexed by cutadapt (8) and then processed using the DADA2 pipeline with default options (9). The sequences were filtered by allowing maximal expected errors of two and with maximal zero Ns. Reads were truncated at the first instance of a quality score of less than three. The forward reads of the Changins, Reckenholz and Aurora data were trimmed to 250 bp and reverse reads to 170 bp. As the sequences of the Sheffield data were only 250 bp long, forward reads were not trimmed and reverse reads to 200 bp. Shorter reads were discarded. For each MiSeq run, a parametric error model was learned by the DADA algorithm and inferred to the previously dereplicated samples. Then the forward and reverse reads were merged if the overlap was identical and at least twelve bases long. A single amplicon sequence variant (ASV) table was created as all datasets used the same 16S rRNA gene primers. We removed chimeras and assigned taxonomy to the ASVs with the naive Bayesian classifier method (9) and the SILVA database (10). Scripts are available from https://github.com/PMI-Basel/Thoenen_et_al_BX_tolerance. The computations were performed at the Vital-IT (https://www.vital-it.ch) center for high-performance computing of the SIB Swiss Institute of Bioinformatics and at the sciCORE (http://scicore.unibas.ch/) scientific computing center of the University of Basel.

Mapping: We aligned the 16S rRNA sequences of the MRB obtained by Sanger sequences to overlap with the 16S rRNA gene region (primers 799F and 1193R; Table S3) of the microbiota profiles using the function *AlignSeqs* (R package DECIPHER, 11). Then, a distance matrix was calculated for all MRB sequences to the identified ASVs of the respective datasets using the function *DistanceMatrix* (DECIPHER). We did not consider mappings with <97% sequence

similarity, the typical threshold for defining operational taxonomic units (OTUs). Strains typically mapped to several ASVs (within 97%), hence their summed relative abundance was taken. The similarities of the MRB strain sequence to the microbiome members (i.e., the ASVs) are listed in Dataset S2. Scripts are available from https://github.com/PMI-Basel/Thoenen_et_al_BX_tolerance. Calculations were performed at sciCORE.

Isolation and purification of benzoxazinoids form maize plants

DIMBOA-Glc was isolated from maize plants as described below. Ca. 200 g of maize leaves (*Zea mays*, variety Akku) were frozen and ground in liquid nitrogen. The resulting powder was placed in 1.5 L MeOH and allowed to warm up to room temperature (23 °C). The resulting suspension was homogenized with an immersion disperser (PT-1035, Kinematica AG, Malters, Switzerland) and filtered through a P3 sintered glass filter equipped with two layers of filter paper, with suction. The filter cake was collected and suspended again in 0.6 L MeOH. After a second homogenization and a new filtration, the filtrates were combined and concentrated under reduced pressure with a rotary evaporator (RC900, KNF Neuberger AG, Balterswil, Switzerland). The aqueous residue obtained was lyophilized with a freeze-drier (LyoQuest-55, Telstar, Terrassa, Spain) to give 8.39 g of crude dry extract. Five runs of purification with ca. 1.7 g of raw material each were performed on an automated flash column chromatography apparatus (CombiFlash Rf+, Teledyne ISCO Inc., Lincoln NE, USA). Solid loading and 120 g silica cartridges were used. The elution gradient was as follow: 0-13% B over 7 min, 13-16% B over 9 min, 16-35% B over 9 min where A = $CHCl_3$ and B = MeOH. The fractions eluting between 19 and 25 min were collected, combined, concentrated under reduced pressure and submitted to new runs of purification. Batches of approx. 250 mg were purified separately with solid loading, 40 g silica gold cartridges, and eluting with 0-15% B over 1.2 min, 15-19% B over 7.2 min, 19-30% B over 1.2 min, 30% B over 3.6 min. The fractions eluting between 11 and 14 min were collected, combined, and concentrated under reduced pressure with a rotary evaporator (RC900, KNF Neuberger AG, Balterswil, Switzerland) to obtain 180 mg of a light-yellow foam (hygroscopic) containing ca. 70% DIMBOA-Glc, 15% DIM_2BOA-Glc, 15% HMBOA-Glc. The analytical data were in accordance with previous literature: UPLC m/z 194.04 $[M-Glc-MeOH]^+$, Mass window: 0.02 Da. Retention time: 1.62 min. HRMS calculated for $C_{15}H_{18}NO_{10}$ $[M-H]^-$: 372.0936, found: 372.0944; ^1H NMR (300 MHz,CD_3OD) δ 7.26 (d, J = 8.8 Hz, 1H), 6.75 (d, J = 2.6 Hz, 1H), 6.70 (dd, J = 8.8, 2.7 Hz, 1H), 5.91 (s, 1H), 4.67 (d, J = 7.8 Hz, 1H), 3.78 (s, 3H), 3.9-3.1 (m, 6H) (12).

High-throughput growth phenotyping of MRB strains

We have described our high-throughput chemical phenotyping system, which we have used to screen MRB strains for their tolerance against various BXs compounds, in detail (13). Here we document the specific settings used in this study.

Setting up an assay requires the preparation of liquid pre-cultures in a 96-well format from fresh bacterial cultures on solid media plates. Pre-cultures were prepared by transferring isolate colonies with inoculation needles (Greiner bio-one, Kremsmünster, Austria) to 1 mL of liquid 50% TSB (Table S5) in 2 ml 96-well deep-well plates (Semadeni, Ostermundigen, Switzerland). These pre-culture growth plates were covered with a Breathe-Easy membrane (Diversified Biotech, Dedham, USA) and grown until stationary phase for 4 days at 28°C and 180 rotations per minute.

Assays were set up by inoculating 4 µL of the pre-cultures to 200 µL fresh liquid 50% TSB (Table S5) in 96-well microtiter plates (Corning, Corning, USA) containing the compounds and concentrations to be tested: DIMBOA-Glc (500 and 2'500 µM), MBOA and BOA (250, 500, 625, 1'250, 2'500 and 5'000 µM), AMPO (10, 25 and 50 µM) or APO (10, 25 and 50 µM). These treatments were prepared by mixing their stock solutions into liquid 50% TSB. Stock solutions were prepared in the solvent DMSO (Sigma-Aldrich) depending on the solubility of the compounds (Table S5) and the DMSO concentration was kept constant in each treatment including the control.

All reactions and replicated plates were pipetted using a liquid handling system (Mettler Toledo, Liquidator 96™, Columbus, USA). All plates had lids and were piled up and inserted to a stacker (BioStack 4, Agilent Technologies, Santa Clara, United States), which was connected to a plate reader (Synergy H1, Agilent Technologies, Santa Clara, United States). Using this system, the optical density (OD600, absorbance at 600 nm) of every culture was recorded every 100 min over 68 hours. Prior to each measurement, the plates were shaken for 120 s. In each plate, wells with 50% TSB were included as 'no-bacteria-controls' and in each run one plate containing only media was included to monitor potential contaminations.

We set up separate runs for the different compounds. In one run, we always tested all concentrations of a compound against all 52 strains with 3 replicates per strain and an empty media control plate. For example, a typical run consisted of a total 23 plates that covered 11 treatments (e.g., 6 concentrations of MBOA + 3 concentrations of AMPO + 2 control treatments; 1 plate per treatment) * 162 cultures (e.g., 52 strains and 2 no bacteria controls, all with 3 replicates; distributed on 2 plates) plus 1 media plate without bacteria. Such a run yielded 1'782 single growth reactions. We have performed at least 2 full runs for every compound (except DIMBOA-Glc due to low availability of the compound). Data were exported from the software of the plate reader (Gen 5, Agilent Technologies, Santa Clara, United States) and imported into R for data analysis (see main methods).

Bacterial genomes

We generated the genomes of a subset of MRB strains in four sets (Dataset S1).

Set 1: The first set of MRB strains consisted of the following four bacteria: *Pseudomonas* LPB4.O, *Pseudomonas* LPD2, *Rhizobium* LRC7.O and *Rhizobium* LRH8 (Dataset S1). Genomic DNA was extracted from overnight cultures grown in liquid LB medium (Table S5) using the GeneElute Bacterial DNA kit (Sigma-Aldrich). 10 kb insert libraries were prepared from the genomic DNA (BluePippin size selection) and sequenced on a PacBio (Pacific Biosciences, Menlo Park, USA) RSII instrument (one RSII SMRT cell per strain; P6-C4 chemistry) at the Functional Genomics Centre Zurich (http://www.fgcz.ch).

Set 2: The second set of MRB strains consisted of 10 bacteria (Dataset S1). Genomic DNA was extracted as for the first set and used for library preparation using NEBNext® DNA Library Prep Kit (New England Biolabs, Ipswich, USA) following the manufacturer's recommendations. The libraries were sequenced on a NovaSeq 6000 instrument (paired-end 150 bp reads; Illumina, San Diego, USA) by Sequentia (www.sequentiabiotech.com) together with other samples of that company and the target to produce >1 Gb of data for each library.

Set 3: The majority of MRB strains (27 strains; Dataset S1) were sequenced in the third set. Total DNA was extracted using the DNeasy UltraClean Microbial Kit (Qiagen, Hilden, Germany) according to the protocol provided. Quantity, purity, and length of the total genomic DNA was assessed using a Qubit 4.0 fluorometer with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), a DS-11 FX spectrophotometer (DeNovix, Wilmington, USA) and a FEMTO Pulse System with a Genomic DNA 165 kb Kit (Agilent, Basel, Switzerland), respectively. Sequencing libraries were made using an Illumina DNA Prep Library Kit (Illumina, San Diego, USA) in combination with IDT for Illumina DNA/RNA UD Indexes Set C and Tagmentation according to the Illumina DNA Prep Reference Guide. The input DNA was set at 200 ng and 5 PCR cycles were employed to amplify the fragmented DNA. Pooled DNA libraries were sequenced paired end on a NovaSeq 6000 SP Reagent Kit v1 (300 cycles) on an Illumina NovaSeq 6000 instrument. The quality of the sequencing run was assessed using Illumina Sequencing Analysis Viewer (version 2.4.7) and all base call files were demultiplexed and converted into FASTQ files using Illumina bcl2fastq conversion software v2.20. All steps from gDNA extraction to sequencing data generation were performed at the Next Generation Sequencing Platform, University of Bern, Switzerland.

Set 4: Several Microbacteria strains (13 strains; Dataset S1) were subjected to PacBio sequencing. DNA was extracted following the GES method (14) from fresh agar plate cultures to ensure good quality DNA with low fragmentation. Briefly, 2-4 mL of each bacterial strain was grown overnight in liquid TSB (Table S3) at 28 °C, centrifuged for 10 min at 12'396 x g at room temperature (20-22 °C), the media was discarded, and the bacterial pellet was re-suspended in 200 µL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). For cell lysis 500 µL of GES solution (guanidium thiocyanate)

was added to each bacterial suspension and incubated for 10 min at RT, before the addition of 250 µL of 7.5 M ammonium acetate. The mixture was gently mixed and incubated on ice for 10 min. Thereafter, 500 µL phenol chloroform isoamyl alcohol mixture, 25:24:1 (Sigma-Aldrich) was added, vigorously mixed, and centrifuged for 15 min at 12'396 x g at 4 °C. The upper aqueous layer was transferred to a fresh tube and 500 µL of chloroform isoamyl alcohol mixture 24:1 (Sigma-Aldrich) was added, vigorously mixed, and centrifuged for 15 min at 12'396 x g at 4 °C. Once again, the upper layer of fluid was transferred to a new tube and mixed with 0.7 vol. 100% isopropanol, mixed well, and stored at -20 °C overnight. Precipitated DNA was recovered by centrifugation at 12'000 rpm for 15 min at 4 °C. The DNA pellet was washed once with 80% ethanol and twice with 70% ethanol. The pellet was dissolved slowly in 80 µL water with the aid of heating at 55 °C for 1 h. Prior to SMRTbell library preparation, bacterial genomic DNA was assessed for quantity, quality and purity using a Qubit 4.0 flurometer (Qubit dsDNA HS Assay kit; Thermo Fisher Scientific), an Advanced Analytical FEMTO Pulse instrument (Genomic DNA 165 kb Kit; Agilent) and a Denovix DS-11 UV-Vis spectrophotometer, respectively. Multiplexed SMRTbell libraries were prepared for sequencing on the Sequel exactly according to the PacBio guideline entitled: “Procedure & Checklist – Preparing Multiplexed Microbial Libraries Using SMRTbell® Express Template Prep Kit 2.0" - Part Number 101-696-100 Version 08 (November 2021). Concisely, 1 µg of gDNA in 100 µL was used to shear the gDNA using a Covaris g-TUBE (Covaris, Wolburn, US). Subsequently, the sheared gDNA was concentrated and cleaned using AMPure PB beads. The samples were then quantified and qualified to be in the range of 12-15 Kb using a Qubit 4.0 flurometer (Qubit dsDNA HS Assay kit, Thermo Fisher Scientific) and an Advanced Analytical FEMTO Pulse instrument (Genomic DNA 165 kb Kit, Agilent), respectively. The rest of the procedure as referenced above was followed including removal of single strand overhangs, DNA damage repair, end-repair & A-tailing, ligation of barcoded overhang adapters and then purification of the library using AMPure PB beads. The libraries were quality controlled using the steps described above and then were pooled using the PacBio microbial multiplexing calculator. Prior to and after size selection, the library pool was purified using AMPure PB beads. Size selection was performed a BluePippin instrument (Sage Science, Beverly, US) using BluePippin with dye free, 0.75% Agarose Cassettes and S1 Marker (Sage Science) wherein the selection cut-off was set at 6’000 bp. Library pool concentration and size was again assessed using a Thermo Fisher Scientific Qubit 4.0 flurometer and an Advanced Analytical FEMTO Pulse instrument (as described above), respectively. PacBio Sequencing primer v4 and Sequel DNA Polymerase 3.0 were annealed and bound, respectively, to the DNA template libraries. The polymerase binding time was 1 h and the complex was cleaned using 1.2 X AMPure PB beads. The libraries were loaded at an on-plate concentration of 150 pM using adaptive loading, along with the use of Spike-In internal control. SMRT sequencing was performed in CLR mode on the Sequel IIe with Sequel Sequencing kit 3.0, SMRT Cells 8M, a 2h pre-extension followed by a 15 h movie time and via PacBio SMRT Link v10.1. Thereafter, the CCS generation and barcode demultiplexing workflow was run in SMRT Link v10.1. All steps from gDNA extraction to sequencing

data generation were performed at the Next Generation Sequencing Platform, University of Bern, Switzerland.

The raw sequencing data of all genomes is available from the European Nucleotide Archive (http://www.ebi.ac.uk/ena) with the study accession PRJEB65362 (sample IDs ERS16291034 to ERS16291087; Dataset S1).

Genome assembly

We utilized similar pipelines to assemble the genomes of all MRB strains. For set 1 (PacBio *and* Illumina sequence data), the fasta sequences of the 'continuous long reads' (CLRs), as extracted from the BAM files using samtools v. 1.10 (15), were used for assembly conducted with Flye v. 2.9 (16). Since these strains were also sequenced on Illumina sequencers, the CLR assembly was corrected with Illumina reads. The reads were first mapped to the assembly using the Burrows-Wheeler Aligner BWA, v 0.7.8 (17). The resulting SAM file was then sorted and indexed using samtools v. 1.10 before using Pilon v. 1.24 (18) to correct the assemblies.

For sets 2 to 4 (generated on Illumina sequencers), the raw, paired end fastq sequences were trimmed using fastp v. 0.20.1 (19) with default options. Read quality was assessed with fastQC v. 0.11.7 (7). These genomes were assembled using the SPAdes assembler v. 3.14.0 (20) with the options `--isolate –k 21,33,55,77,99,127 --cov-cutoff 'auto'`. The quality of the assemblies was assessed with Quast v. 4.6.0 (21), BUSCO v. 5.1.3 (22) and checked for contamination with ConFindr v. 0.7.2 (23). The genomes were then annotated with the NCBI procaryotic genome assembly pipeline PGAP, v. 2022-04-14 (24). The annotated genomes were functionally annotated with EggNog v. 5.0.1 (25) and orthologue genes were determined using OrthoFinder v. 2.3.8 (26).

The annotated assemblies were then integrated into a local instance of OpenGenomeBrowser (27) hosted at the Interfacility Bioinformatics Unit (University of Bern). The genome assemblies and annotations have been deposited at NCBI (http://www.ncbi.nlm.nih.gov/) under the BioProject ID PRJNA1009252 (Dataset S1).

Supplementary Figures

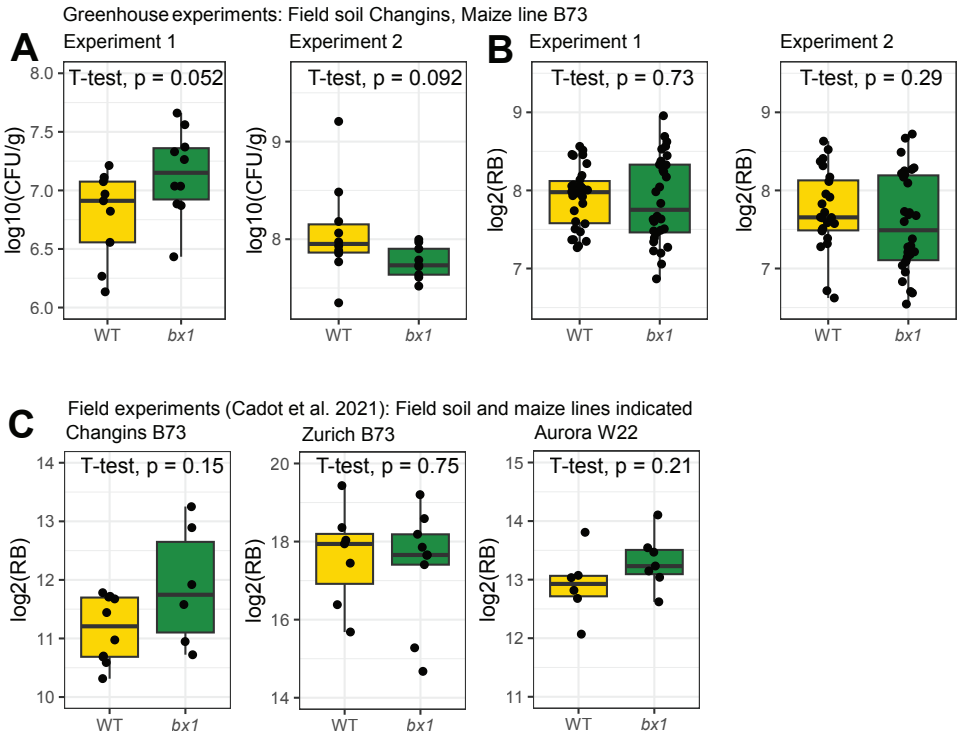
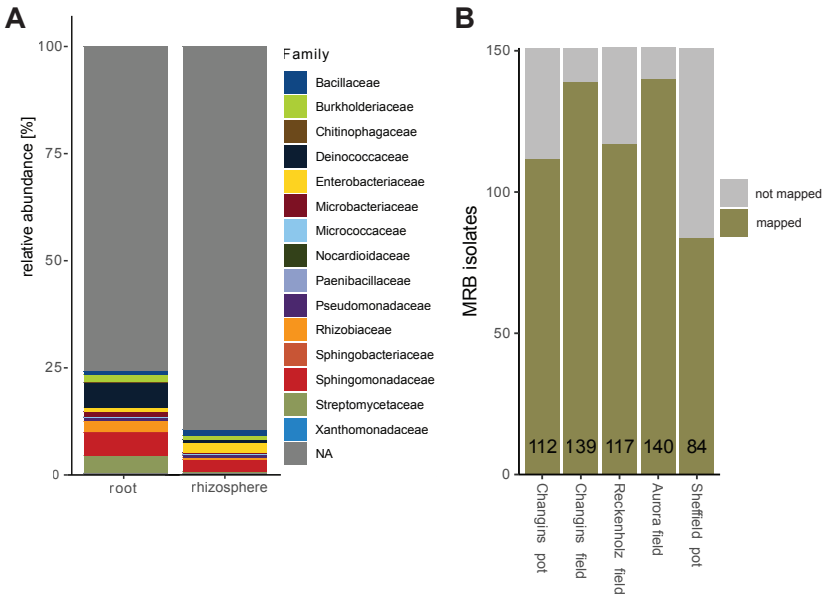


Figure S1: Bacterial community size on maize roots by microbiological and qPCR analyses

A) Bacterial root colonization was assessed by plating colony forming units (CFU) expressed as log10 CFU / g roots and tested statistically for differences between wild-type and *bx1* plants. Individual datapoints are shown (n = 9-10) **B)** DNA extracts from the same plants were used for qPCR analysis. The bacterial signal, derived from 16S rRNA primers 799F and 904R, was normalized relative to the plant signal of the plant actin gene (ZmActin1) expressed as log2(RB), RB = relative bacterial gene signal ($E_{\text{plant gene}}^{\text{Ct plant gene}} / E_{16S}^{\text{Ct16S}}$). Individual datapoints are shown (n = 27-30, 9-10 samples with three technical replicates each). **C)** DNA extracts from maize roots grown in three field experiments published in Cadot et. al. 2021. Individual datapoints are shown (WT Changins n = 10, *bx1* Changins n = 6, WT Zurich n = 7, *bx1* Zurich n = 9, WT Aurora n = 6, *bx1,* Aurora n = 7. Results from t-test between wild-type and *bx1* are shown in the panels.



318

319 **Figure S2: Mapping of MRB strains to maize root and rhizosphere microbiome datasets**

320 **A)** Cumulative abundance of MRB strains, reported at family level, in the root and rhizosphere
321 profiles of wild-type B73 maize plants, from which the MRB strains were isolated from. This was a
322 greenhouse experiment with pots filled with natural field soil from the Changins site. The
323 microbiome data corresponds to the feedback experiment reported in Hu et al. 2018. **B)** Number
324 of MRB isolates mapping to abundant community members (> 0.1% abundance) in root microbiome
325 datasets of maize grown in greenhouse and field experiments (Changins field data from Hu et al.
326 2018; Reckenholz and Aurora data from Cadot et al. 2021) or a greenhouse experiment with field
327 soil (Sheffield data from Cotton et al. 2019).

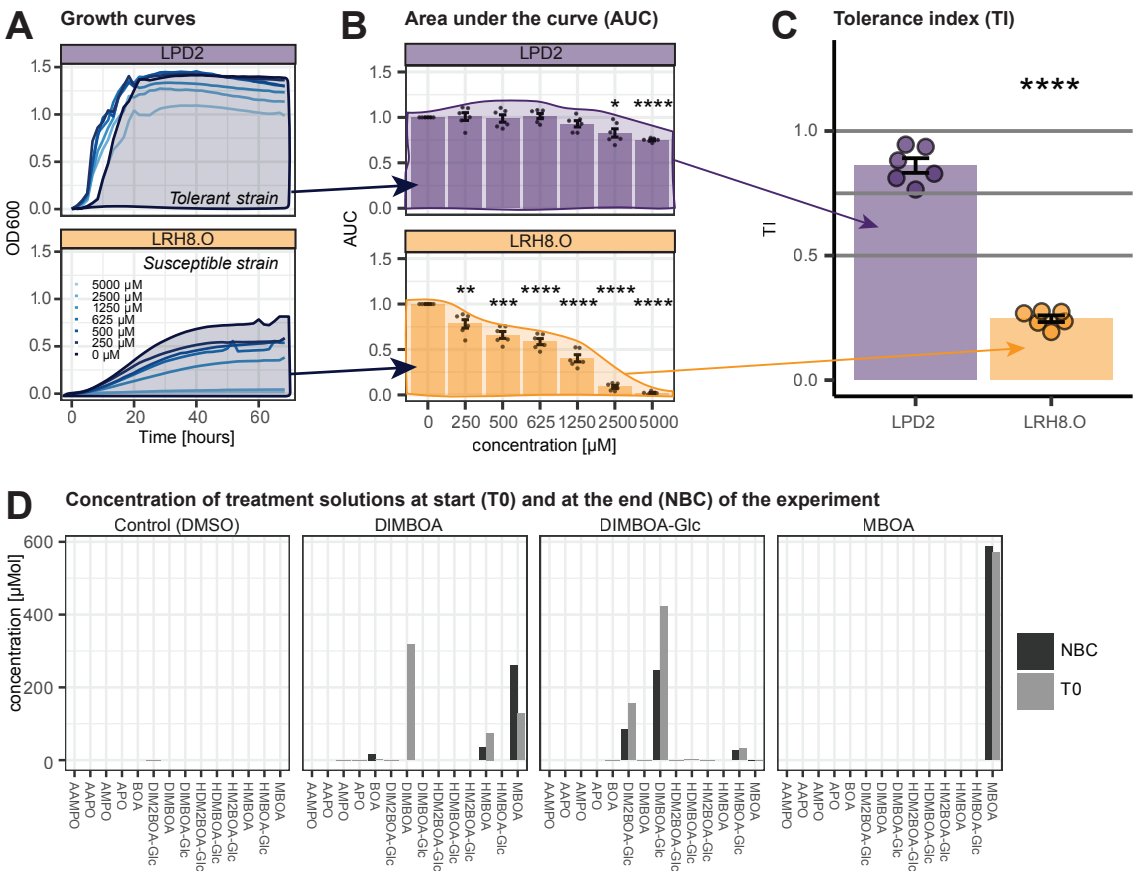


Figure S3: *In vitro* growth of maize root bacteria in MBOA and stability of benzoxazinoids in culture medium

A) Bacterial growth curves (OD_{600}) of a representative tolerant strain of Pseudomonadaceae (LPD2) and a representative susceptible strain of Rhizobiaceae (LRH8.O) at different concentrations of MBOA over a time course of 68 hours. **B)** Area under the curve (AUC), normalized to the BX-free control treatment **C)** Tolerance index (TI). Means ± SE bar graphs and individual datapoints are shown (n = 6). Results of pairwise t-test is shown inside the panels, p-value < 0.05 = *. **D)** We screened the treatment solutions of DIMBOA, DIMBOA-Glc and MBOA for all benzoxazinoid and aminophenoxazinone compounds of which we had standards. Graphs report the measured concentrations at the start (T0) and 68 h later at the end of the experiment (NBC).

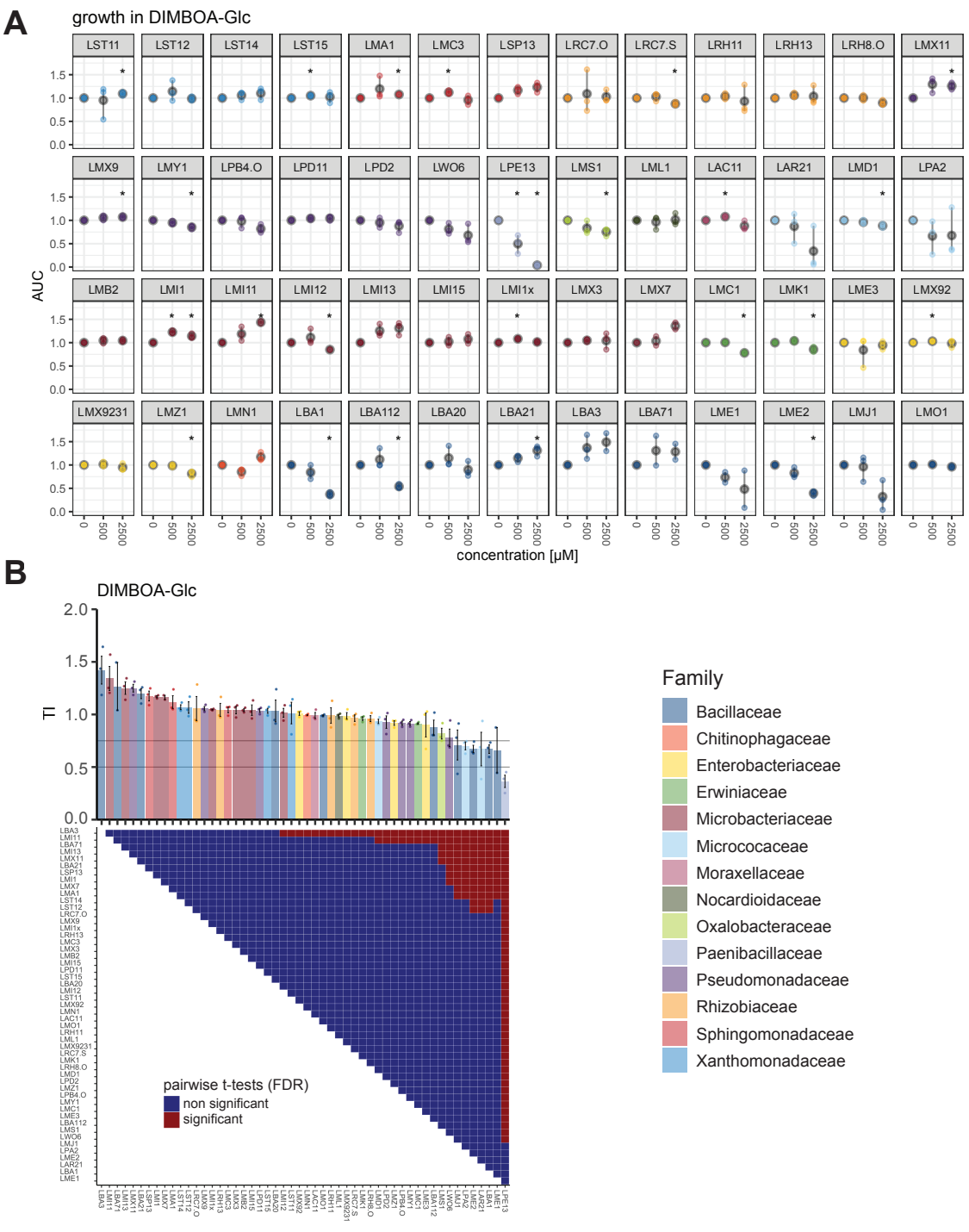


Figure S4: Screening maize root bacteria for tolerance to DIMBOA-Glc

A) Growth measured as area under the curve (AUC) of maize root bacteria in 0, 500 and 2'500 µM of DIMBOA-Glc. Data points are replicate measurements (n = 3). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: $p < 0.05$*). **B)** The upper panel repeats the tolerance TI data of the main figure and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant ($p > 0.05$) and significant ($p < 0.05$) differences are in blue and red, respectively. Colors by family taxonomy.

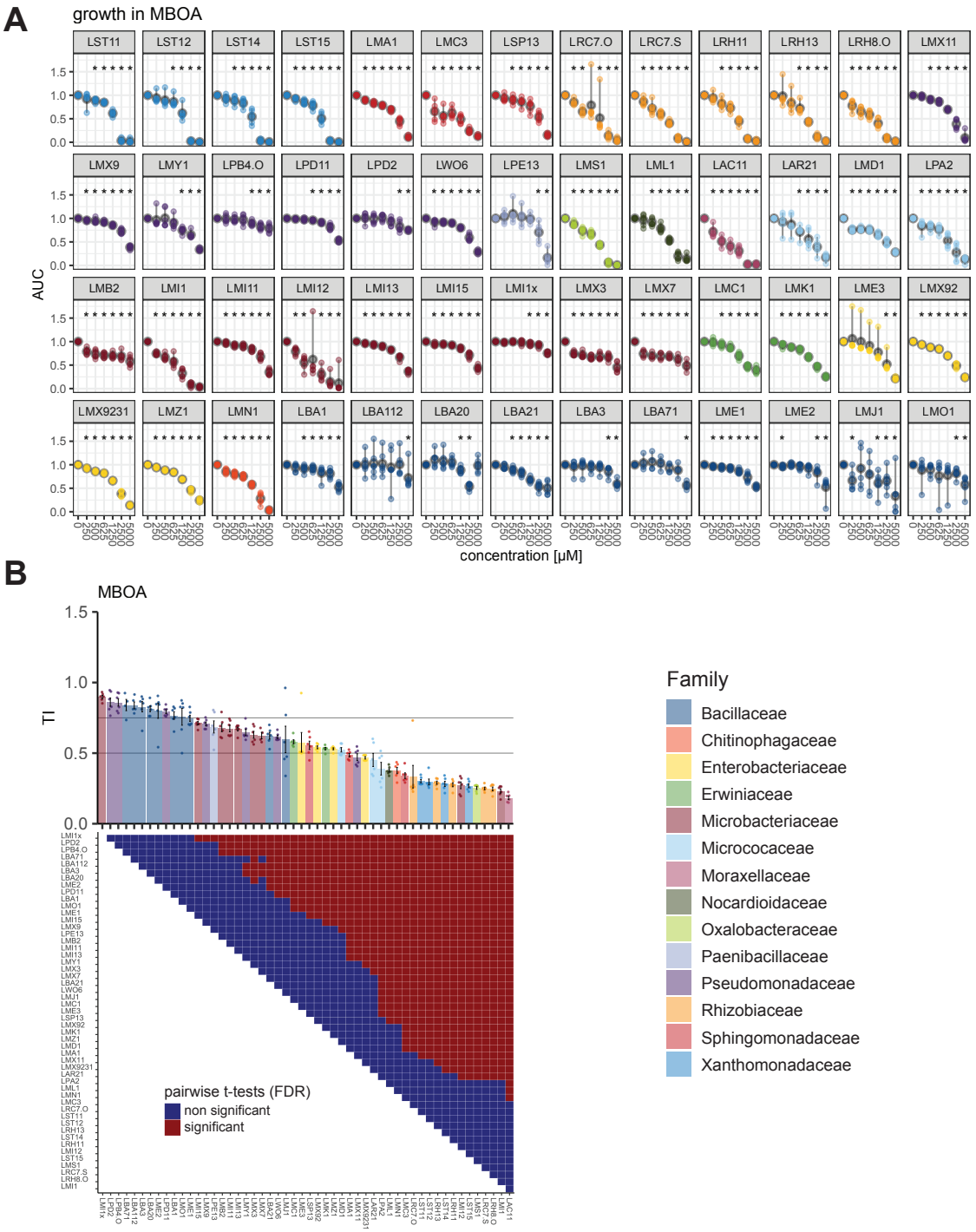
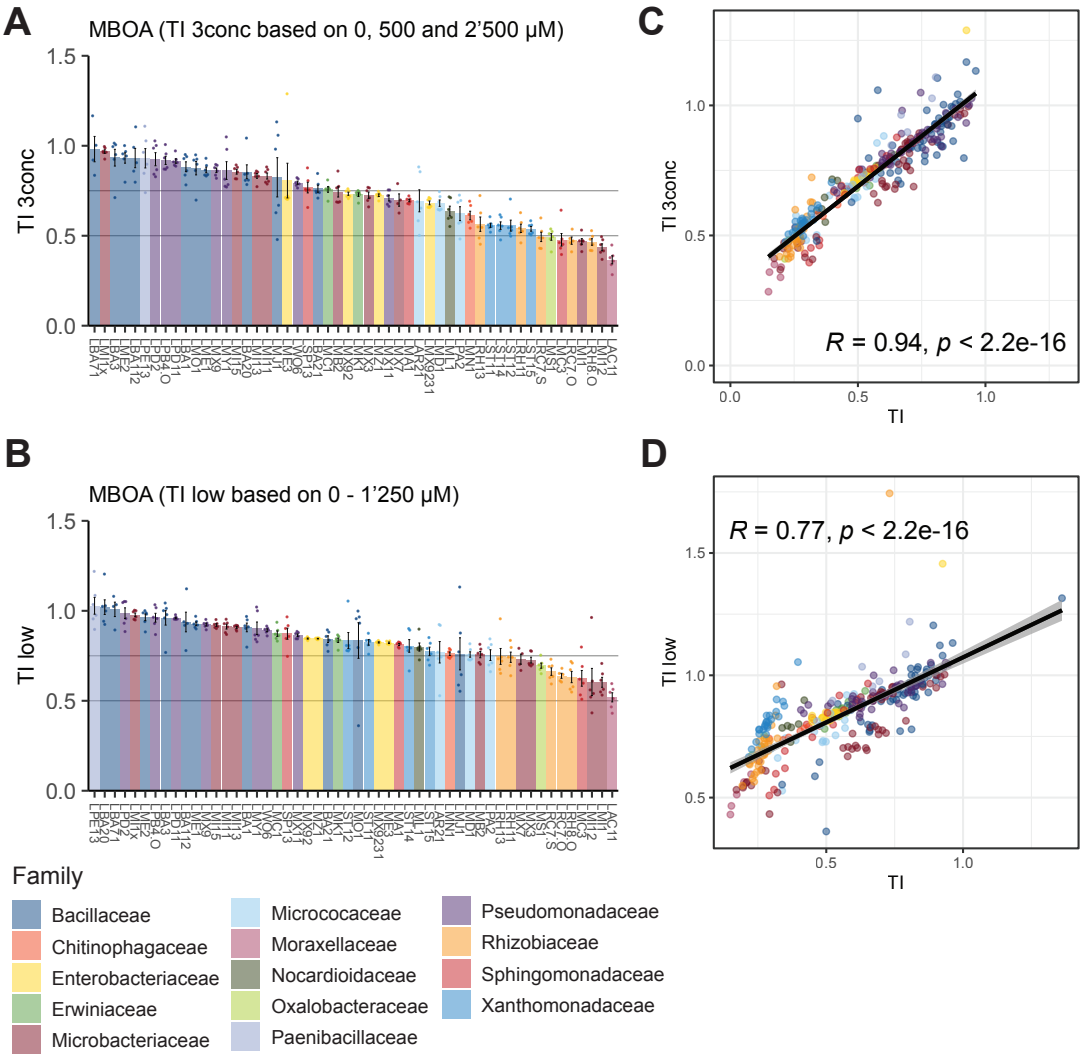


Figure S5: Screening of maize root bacteria for tolerance to MBOA

A) Growth measured as area under the curve (AUC) of maize root bacteria in 0 - 5'000 µM MBOA. Data points are individual measurements (n = 6). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: $p < 0.05$*). **B)** The upper panel repeats the tolerance TI data of the main figure and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant ($p > 0.05$) and significant ($p < 0.05$) differences are in blue and red, respectively. Colors by family taxonomy.



356

357 **Figure S6: Validation of tolerance index approach**

358 Robustness of TI-based findings were tested if they were affected by the number of concentrations
359 used for calculation and whether the TI was primarily driven by very high concentrations. **A)** Each
360 bacterium's tolerance to MBOA was re-calculated either only based on 3 concentrations (0, 500,
361 2'500 µM MBOA; defined as TI_{3conc}) or **B)** excluding data of the two highest concentrations (0 -
362 1'250 µM MBOA; defined as TI_{low}). Bargraphs report means ±SE (n = 6). Correlations between **C)**
363 TI_{3conc} and **D)** TI_{low} with the with the original TI reported in Fig. 2B with their correlation coefficient
364 R and p-value of the Pearson's product-moment test inside the panels. Colors by family taxonomy.

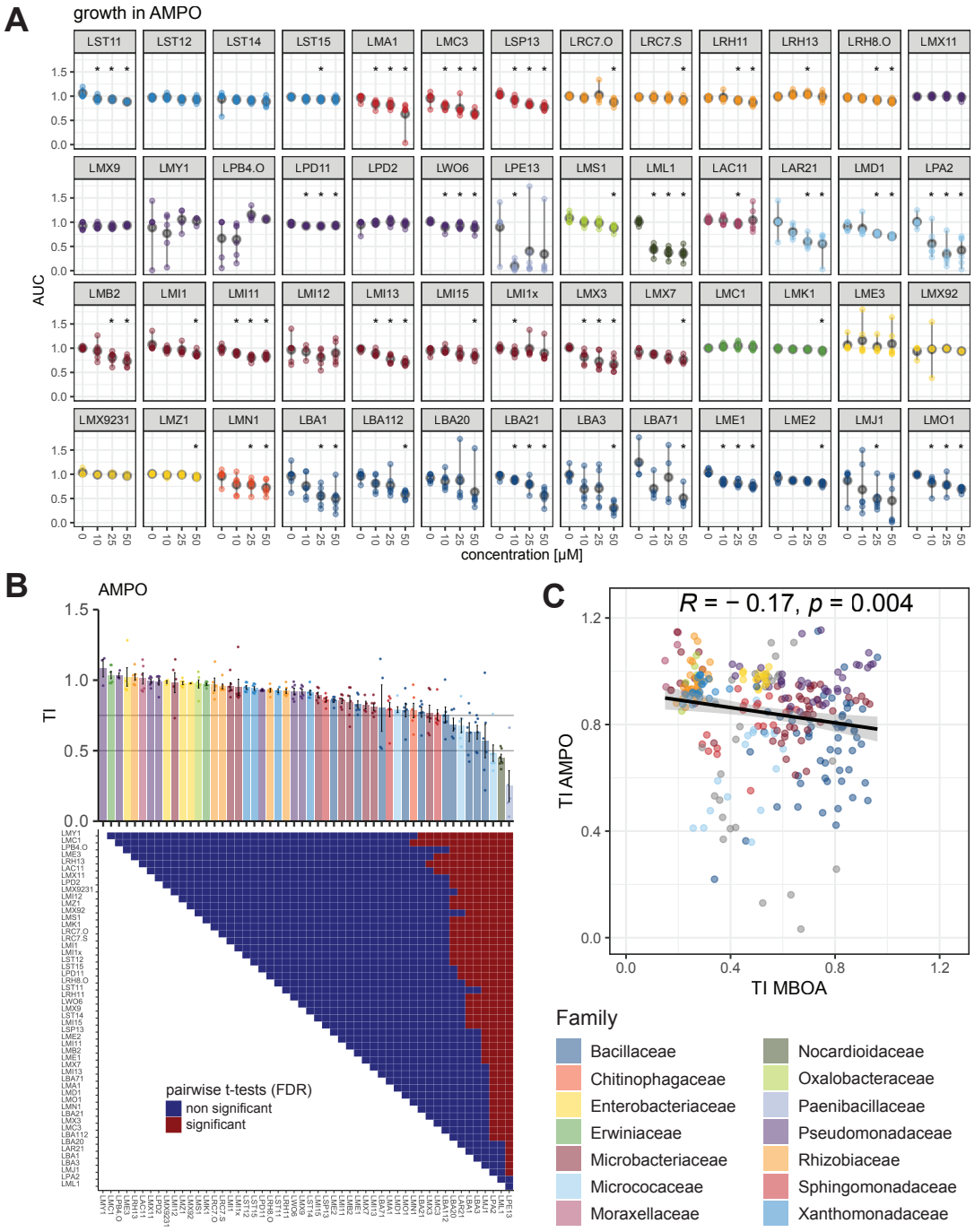


Figure S7: Screening of maize root bacteria for tolerance to AMPO

A) Growth measured as area under the curve (AUC) of maize root bacteria in 0, 10, 25 and 50 µM AMPO. Data points are individual measurements (n = 6). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: p < 0.05*). **B)** The upper panel repeats the tolerance TI data of the main figure and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant (p > 0.05) and significant (p < 0.05) differences are in blue and red, respectively. **C)** Correlation between TIs of AMPO and MBOA with their correlation coefficient R and p-value of the Pearson's product-moment test inside the panel. Colors by family taxonomy.

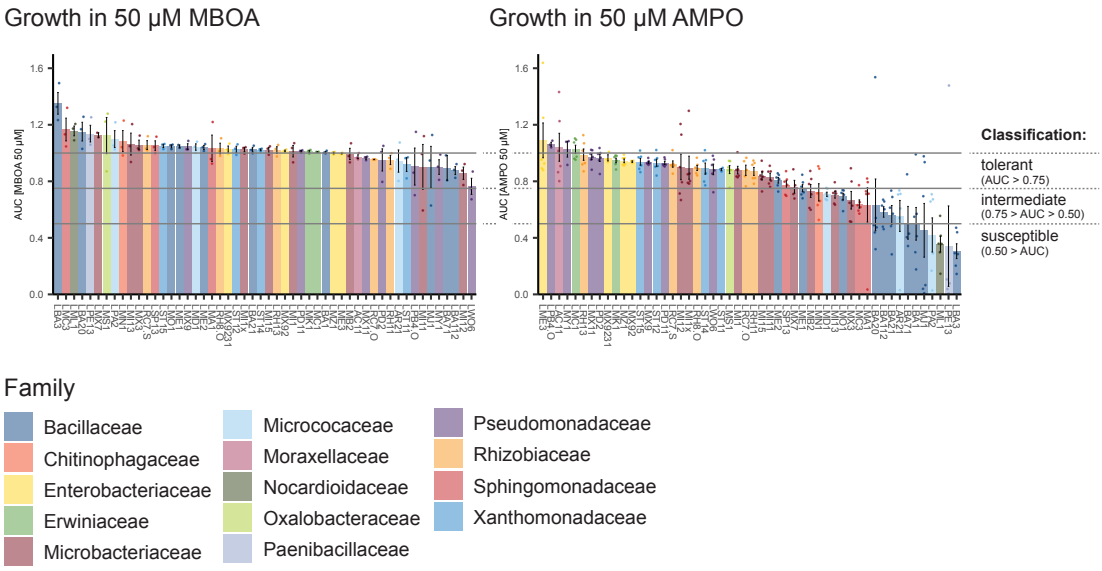


Figure S8: Growth of maize root bacteria in equimolar MBOA and AMPO

Direct comparison of growth (as area under the curve AUC) in 50 µM MBOA and 50 µM AMPO. Data points are individual measurements (n = 6). Strains were classified as tolerant (AUC > 0.75), intermediately tolerant (0.75 > AUC > 0.50) or susceptible (0.5 > AUC). Colors by family taxonomy.

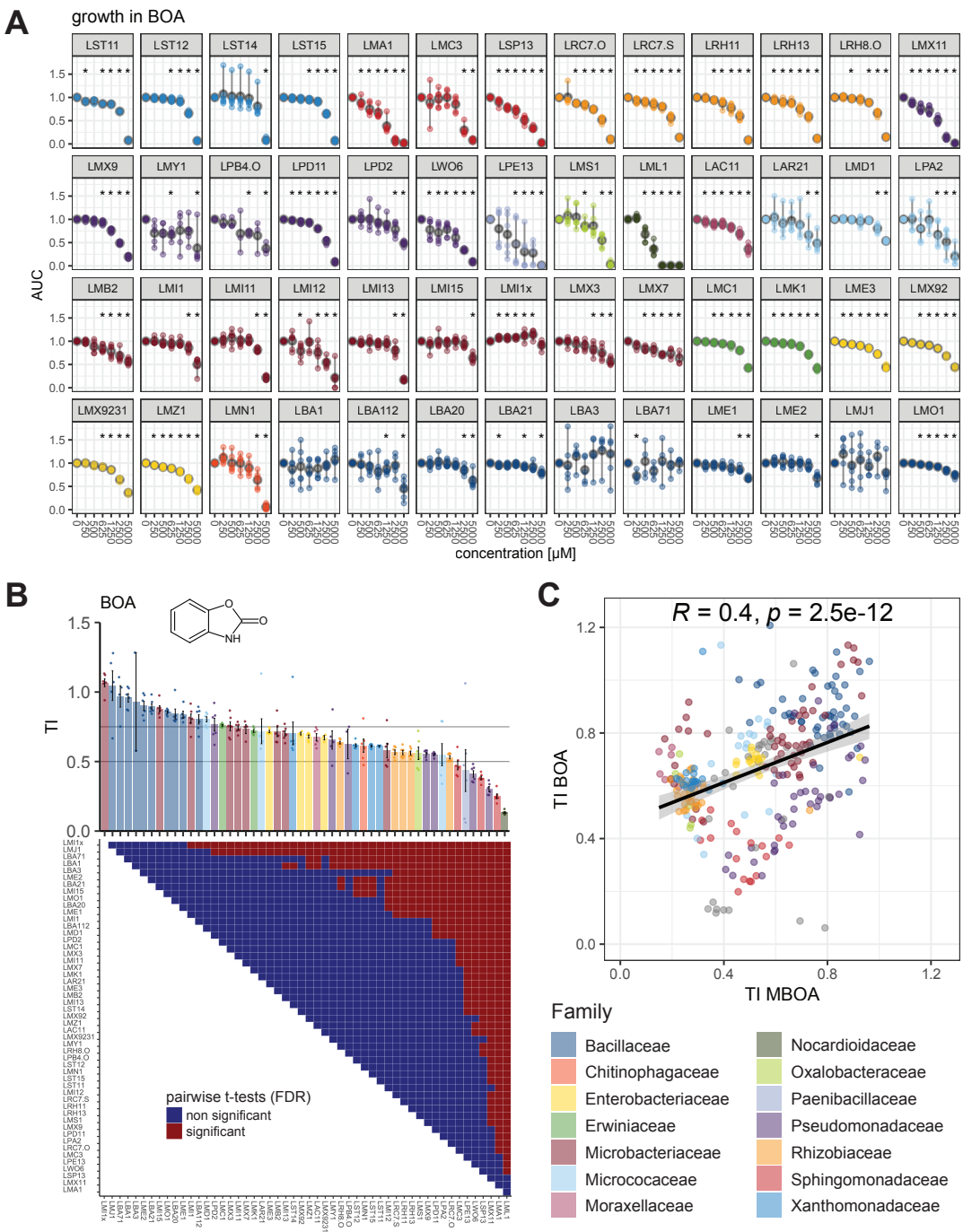


Figure S9: Screening of maize root bacteria for tolerance to BOA

BOA is the non-methoxylated relative compound of MBOA. **A)** Growth measured as area under the curve (AUC) of maize root bacteria in 0 - 5'000 µM BOA. Data points are individual measurements (n = 6). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: p < 0.05*). **B)** The upper panel reports the tolerance TI data (mean ±SE, n = 6) and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant (p > 0.05) and significant (p < 0.05) differences are in blue and red, respectively. **C)** Correlation between TIs of BOA and MBOA with their correlation coefficient R and p-value of the Pearson's product-moment test inside the panel. Colors by family taxonomy.

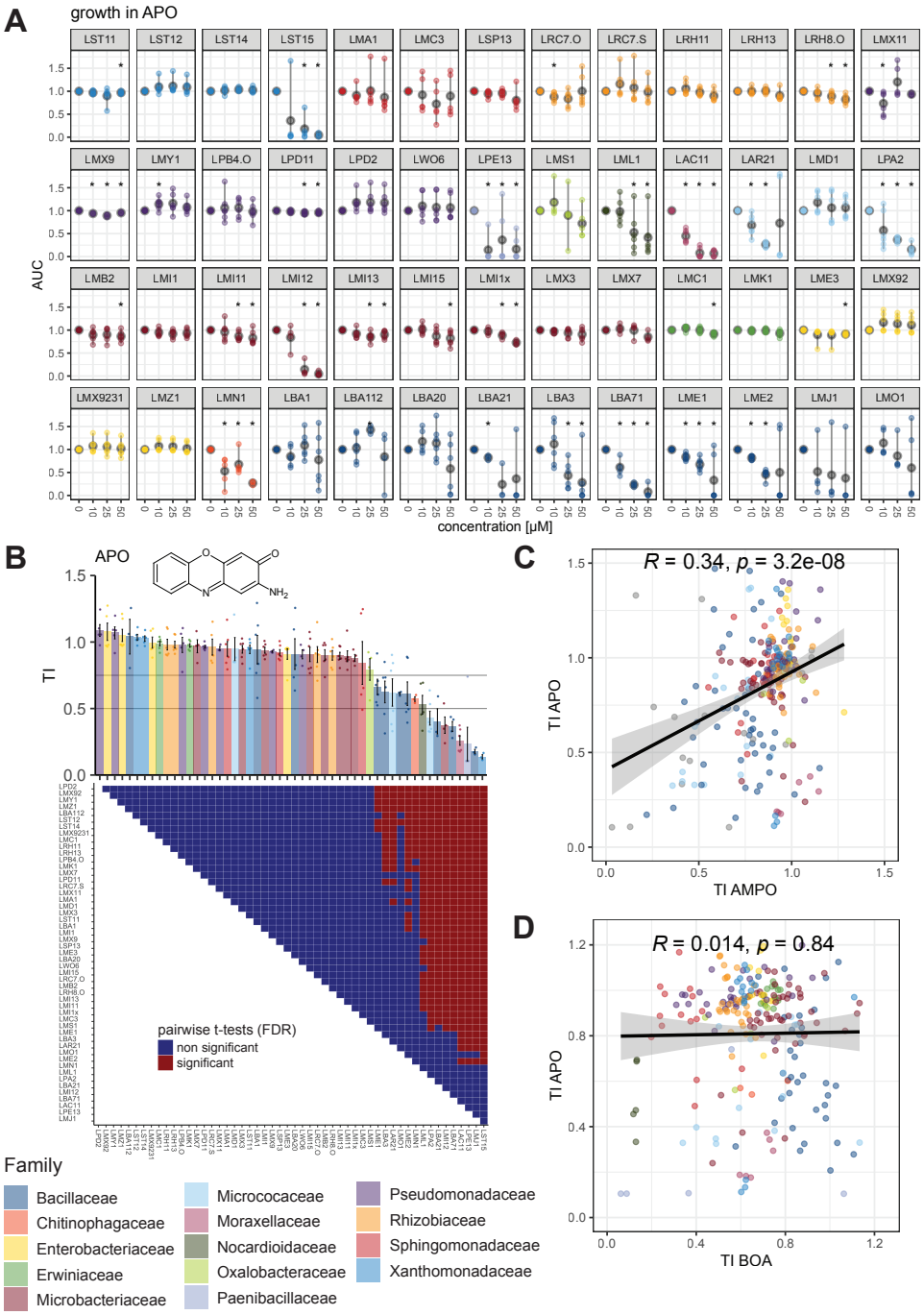


Figure S10: Screening of maize root bacteria for tolerance to APO

APO is the non-methoxylated relative compound of AMPO, both are aminophenoxazinones. **A)** Growth measured as area under the curve (AUC) of maize root bacteria in 0, 10, 25 and 50 µM APO. Data points are individual measurements (n = 6). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: $p < 0.05$*). **B)** The upper panel reports the tolerance TI data (mean ±SE, n = 6) and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant ($p > 0.05$) and significant ($p < 0.05$) differences are in blue and red, respectively. Correlations between TIs of **C)** APO and AMPO and **D)** APO and BOA with their correlation coefficient R and p-value of the Pearson's product-moment test inside the panels. Colors by family taxonomy.

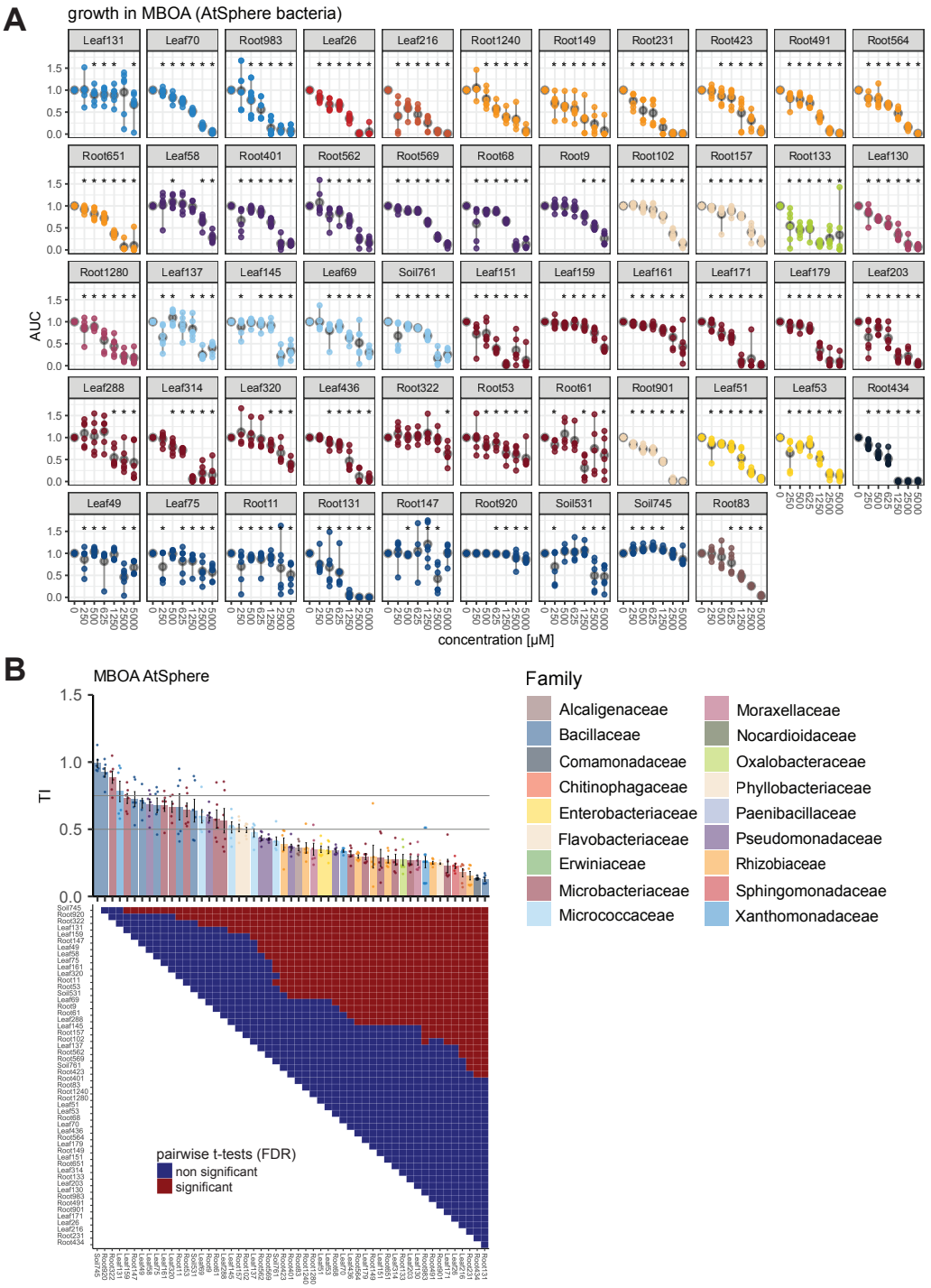
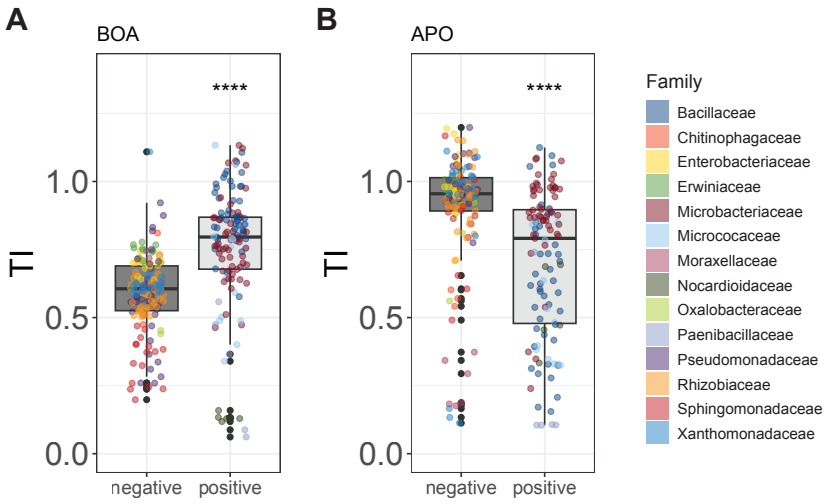


Figure S11: Screening of Arabidopsis bacteria for tolerance to MBOA

A) Growth measured as area under the curve (AUC) of AtSphere bacteria in 0 – 5'000 µM MBOA. Data points are individual measurements (n = 6). Statistical analysis relative to the 0 µM control (t-test, asterisks denote significance: $p < 0.05$*). **B)** The upper panel repeats the tolerance TI data of the main figure and the lower panel displays the corresponding statistical analysis with all strain-to-strain comparisons (pairwise t-tests, FDR adjusted p-values, non-significant ($p > 0.05$) and significant ($p < 0.05$) differences are in blue and red, respectively. Colors by family taxonomy.



410

411 **Figure S12: Tolerance of gram-negative vs. gram positive maize root bacteria**

412 The tolerance indices (TIs) to **A)** BOA and **B)** APO are summarized for gram-negative and gram-
413 positive maize root bacteria. Graphs report six measurements per strain and statistical analysis (t-
414 test, asterisks denote significance: $p < 0.0001$****). Colors by family taxonomy.

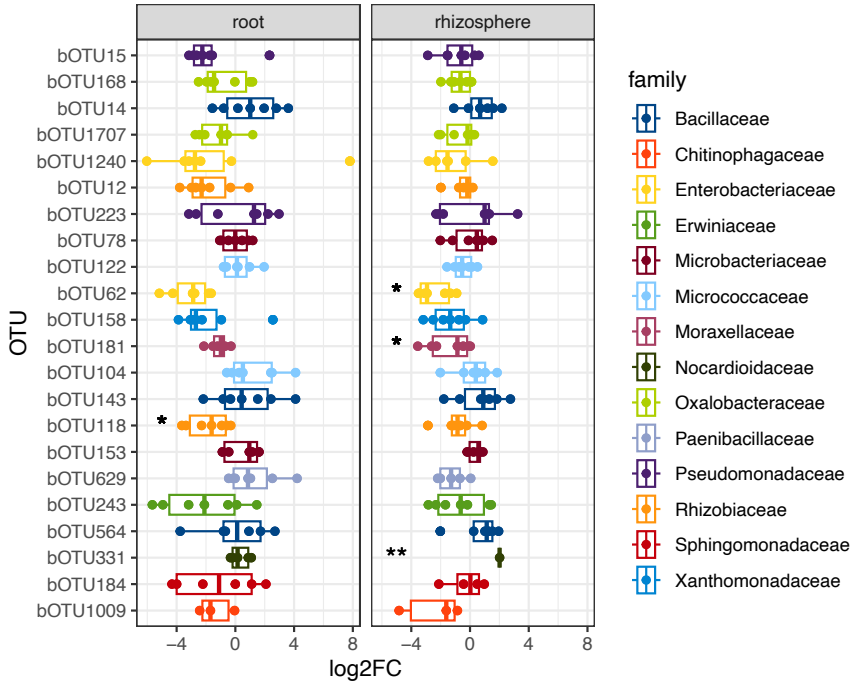


Figure S13: Differential abundance of OTUs corresponding to MRB strains on roots and in rhizosphere profiles of maize

MRB isolates were mapped to the bacterial operational taxonomic units (bOTUs) of the microbiome dataset of Hu et al. 2018. In this study, wild-type and *bx1* mutant maize lines were grown in a field experiment in Changins from where the soil was used for the isolation experiments. Differential abundance (i.e., the log2 fold change; log2FC) between wild-type and *bx1* plants was calculated for each bOTU on roots (n=7) and in the rhizosphere (n=7) microbiome profiles. Log2FC > 1 denote enrichment, while values < 1 refer to depletion on wild-type plants. Statistical analysis based on t-tests (asterisks denote significance: $p < 0.05$*). Colors by family taxonomy.

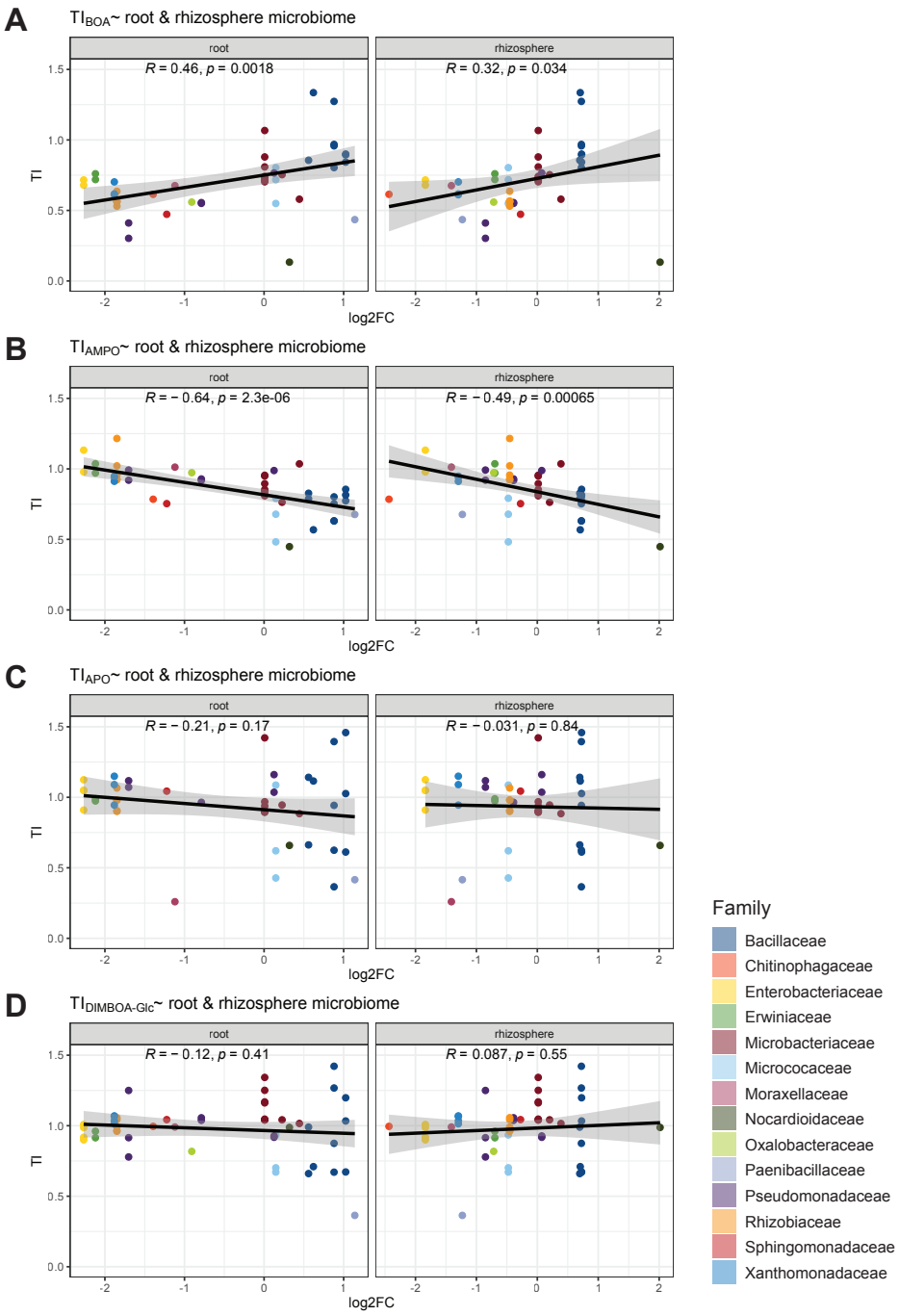


Figure S14: Correlations between in vitro tolerance and abundance in root and rhizosphere microbiomes of maize

Correlations between the tolerance indices (TIs) of the MRB strains with the abundance changes (log2FC, wild-type vs. *bx1*) of their corresponding OTUs in root and in rhizosphere microbiome profiles. This analysis was performed with the TIs of **A)** BOA, **B)** AMPO, **C)** APO and **D)** DIMBOA-Glc. The correlation coefficient *R* and the p-value of the Pearson's product-moment correlation test are reported inside the panels. Each data point reports the mean TI of a strain (means are based on n=6 replicates for BOA, AMPO and APO and n=3 for DIMBOA-Glc); colors by family taxonomy.

Supplementary Tables

Table S1: Abbreviation benzoxazinoid metabolites

Abbreviation	Full name	Class	Mass [g/mol]	Formula	Concentration Rhizosphere [µg/kg FW]*	Concentration Root exudates [µg/kg FW]*
AAMPO	2-acetylamino-7-methoxy-phenoxazin-3-one	Amino-phenoxazine	284.27	C15H12N2O4	ND	ND
AMPO	2-amino-7-methoxy-phenoxazin-3-one	Amino-phenoxazine	242.23	C13H10N2O3	0.39±0.10	ND
DIMBOA	2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one	Benzoxazinone	211.17	C9H9NO5	0.57±0.33	0.05±0.016
DIMBOA-Glc	4-hydroxy-7-methoxy-2-[3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxy-1,4-benzoxazin-3-one	Benzoxazinone glucoside	373.31	C15H19NO10	ND	**0.09±0.01**
HDM2BOA-Glc	2-(2-hydroxy-4,7,8-trimethoxy-1,4-benzoxazin-3-one)-β-d-glucopyranose	Benzoxazolinone glucoside	417.4	C17H23NO11	0.27±0.09	0.00±0.00
HDMBOA-Glc	4,7-dimethoxy-2-{[3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl]oxy}-3,4-dihydro-2H-1,4-benzoxazin-3-one	Benzoxazolinone glucoside	387.34	C16H21NO10	2.80±0.92	0.03±0.005
HMBOA-Glc	2-O-Glucosyl-7-methoxy-1,4(2H)-benzoxazin-3-one	Benzoxazinone glucoside	357.31	C15H19NO9	1.60±0.33	0.004±0.001
MBOA	6-methoxybenzoxazolin-2(3H)-one	Benzoxazolinone	165.15	C8H7NO3	**5.62±1.39**	0.002±0.0004

*Measurements from Supplementary Figures S1A & B of Hu et al. 2018, ND = not detected. Most abundant compounds are marked in bold.

Table S2: Experiments and plant growth conditions

Experiment	Changins soil*	Plant growth**	Isolation media***	Comment
Isolation 1	Summer 2014	10 weeks (no fertilization), setting 1	FlourA, PseudoA	re-streaked on LB
Isolation 2	Spring 2019	14 weeks (full period), setting 2	1/10 TSB MBOA	re-streaked on TSA
Isolation 3	Spring 2019	12 weeks (only weeks 1-5), setting 2	1/10 TSB MBOA	re-streaked on TSA
Isolation 4 and community size 2	Winter 2019	7 weeks (full period), setting 2	1/10 TSB & 1/10 TSB MBOA	re-streaked on TSA
Isolation 5	Spring 2019	12 weeks (full period), setting 2	1/10 TSB MBOA	re-streaked on TSA
Community size 1	Summer 2019	6 weeks (full period), setting 2	NA	NA

* Batches of Changins soil: the 'Summer 2014' batch corresponds to original study (2)

** Plant growth: age of harvested plants, period of fertilization (weeks 1 – 4: 100 ml; 0.2% Plantactive Typ K (Hauert HBG Duenger AG, Grossaffoltern, Switzerland), 0.0001% Sequestrene Rapid (Maag, Westland Schweiz GmbH, Dielsdorf, Switzerland); weeks 5 onwards: 200 ml; 0.2% Plantactive Typ K, 0.02% Sequestrene Rapid) and greenhouse settings (Settings 1: 14:10 h light/dark, 26 °C ± 2 °C, 55% relative humidity, 50'000 lm m−2; Setting 2: 16:8 light/dark, 26/23 °C, 50% relative humidity, ~550 µmol m-2s-1 light).

*** see Table S4

447 **Table S3: PCR primer sequences**

Name	Target	Direction	Sequence 5'-3'	Reference
27f	Bacterial 16S rRNA gene	forward	AGAGTTTGATCCTGGCTCAG	(28)
1492r	Bacterial 16S rRNA gene	reverse	GGTTACCTTGTTACGACTT	
ZmActin1_F	Maize Actin gene	forward	CCAGAGGCCACGTACAACT	(29)
ZmActin1_R1	Maize Actin gene	reverse	GGTAAAACCCCCACTGAGGA	
799F	Bacterial 16S rRNA gene	forward	AACMGGATTAGATACCCKG	(30)
904R	Bacterial 16S rRNA gene	reverse	CCCCGTCAATTCITTTGAGTTTYAR	

448

449 **Table S4: Media uses for isolation of maize root bacteria**

Medium*	Type	Ingredients (source)	Supplements** (source)
FlourA (31)	solid	6 g/L corn starch (Sigma-Aldrich) 0.3 g/L yeast extract (Duchefa Biochemie, Haarlem, Netherlands) 0.3 g/L sucrose (Fluka Chemie GmbH, Buchs, Switzerland) 0.3 g/L CaCO3 (Fluka Chemie GmbH) 18 g/L agar (Sigma-Aldrich)	10 mg/mL Cycloheximide (Sigma-Aldrich)
PseudoA	solid	45 g/L Pseudomonas Isolation Agar (Difco, Le pont de Claix, France)	40 µg/mL Ampicillin (Sigma-Aldrich) 13 µg/mL Chloramphenicol (Sigma-Aldrich)
LB	liquid	25 g/L Luria-Bertani broth (Carl Roth, Karlsruhe, D)	-
LBA	solid	40 g/L Luria-Bertani agar (Carl Roth)	-
TSB	liquid	30 g/L tryptic soy medium (Sigma-Aldrich)	-
TSA	solid	30 g/L tryptic soy medium (Sigma-Aldrich) 15 g/L agar (Sigma-Aldrich)	
50% TSB	liquid	15 g/L tryptic soy medium (Sigma-Aldrich)	-
50% TSA	solid	15 g/L tryptic soy medium (Sigma-Aldrich) 15 g/L agar (Sigma-Aldrich)	
10% TSA	solid	3 g/L tryptic soy medium (Sigma-Aldrich) 15 g/L agar (Sigma-Aldrich)	2 mL/L DMSO (Sigma-Aldrich) 10 mg/L Cycloheximide (Sigma-Aldrich)
10% TSA MBOA	solid	3 g/L tryptic soy medium (Sigma-Aldrich) 15 g/L agar (Sigma-Aldrich)	200 mg/L MBOA (Sigma-Aldrich) 10 mg/L Cycloheximide (Sigma-Aldrich)

450 *All media were sterilized by autoclaving.

451 **Supplements (Cycloheximide, MBOA, Ampicillin, Chloramphenicol) were filter sterilized.

452

453 **Table S5: Stock solutions of compounds used for in vitro growth assays**

Compound	mol. weight	stock conc. [mM]	stock mg/mL	solvent
DIMBOA-Glc	373.1	500	186.55	DMSO
MBOA	165	606	100	DMSO
BOA	135.1	500	67.55	DMSO
AMPO	242.23	15	3.6	DMSO
APO	212.21	15	3.18	DMSO
Ctrl	0	0	0	DMSO

454

Supplementary Datasets

Dataset S1: MRB strain collection sequences

This table contains detailed information about the taxonomic assignment of the MRB isolates. Further information on the isolation experiment, the plant, the extract, and the isolation media are included. The partial sequence of the 16S rRNA gene obtained by Sanger sequencing along with the primer used is listed. For each strain it is indicated if and with which method the genome was sequenced. Additionally basic information of the genomes is listed.

Dataset S2: MRB strain collection mapping

The mapping of MRB isolates to the microbiome profiles of the maize roots, where they were isolated from (pot experiment with Changins soil) indicating the identity to the taxonomic units.

References Supplementary Information

1. D. Maag, *et al.*, Highly localized and persistent induction of Bx1-dependent herbivore resistance factors in maize. *Plant Journal* **88**, 976–991 (2016).

2. L. Hu, *et al.*, Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature Communications* **9**, 2738–2738 (2018).

3. S. Cadot, *et al.*, Specific and conserved patterns of microbiota-structuring by maize benzoxazinoids in the field. *Microbiome* **9**, 103–103 (2021).

4. J. M. Ruijter, *et al.*, Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**, 0–0 (2009).

5. N. Bodenhausen, *et al.*, Relative qPCR to quantify colonization of plant roots by arbuscular mycorrhizal fungi. *Mycorrhiza* **31**, 137–148 (2021).

6. T. E. A. Cotton, *et al.*, Metabolic regulation of the maize rhizobiome by benzoxazinoids. *The ISME Journal* **13**, 1647–1658 (2019).

7. S. Andrews, FastQC: A quality control tool for high throughput sequence data (2010).

8. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

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

10. B. J. Callahan, Silva taxonomic training data formatted for DADA2 (Silva version 132) (2018) https:/doi.org/10.5281/zenodo.1172783.

11. E. S. Wright, Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *R Journal* **8**, 352–359 (2016).

12. K. Yamada, *et al.*, Isolation and Identification of Phototropism-regulating Sub- stances Benzoxazinoids from Maize Coleoptiles. *Heterocycles* **63**, 2707–2712 (2004).

13. L. Thoenen, *et al.*, “Customisable high-throughput chemical phenotyping of root bacteria” in (in press).

14. D. G. Pitcher, N. A. Saunders, R. J. Owen, Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Letters in Applied Microbiology* **8**, 151–156 (1989).

15. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

16. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* **37**, 540–546 (2019).

17. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

18. B. J. Walker, *et al.*, Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* **9**, 1–14 (2014).

19. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

20. A. Bankevich, *et al.*, SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19**, 455–477 (2012).

21. A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

22. M. Seppey, M. Manni, E. M. Zdobnov, BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods of Molecular Biology* **1962**, 227–245 (2019).

23. A. J. Low, A. G. Koziol, P. A. Manninger, B. Blais, C. D. Carrillo, ConFindr: rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. *PeerJ* **7**, e6995 (2019).

24. T. Tatusova, *et al.*, NCBI prokaryotic genome annotation pipeline. *Nucleic acids research* **44**, 6614–6624 (2016).

25. J. Huerta-Cepas, *et al.*, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47**, 309–314 (2019).

26. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 1–14 (2019).

27. T. Roder, S. Oberhänsli, N. Shani, R. Bruggmann, OpenGenomeBrowser: A versatile, dataset-independent and scalable web platform for genome data management and comparative genomics. *bioRxiv* (2022) https:/doi.org/10.1101/2022.07.19.500583.

28. S. Turner, K. M. Pryer, V. Miao, J. D. Palmer, Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis. *Journal of Eukaryotic Microbiology* **46**, 327–338 (1999).

29. M. Erb, *et al.*, Signal signature of aboveground-induced resistance upon belowground herbivory in maize. *Plant Journal* **59**, 292–302 (2009).

30. K. Schlaeppi, N. Dombrowski, R. G. Oter, E. V. L. van Themaat, P. Schulze-Lefert, Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 585–592 (2014).

31. J. T. Coombs, C. M. M. Franco, Isolation and Identification of Actinobacteria from Surface-Sterilized Wheat Roots. *Applied and Environmental Microbiology* **69**, 5603–5608 (2003).