# Supplementary Information for

## Ultrahigh-throughput functional profiling of microbiota communities

Stanislav S. Terekhov[1], Ivan V. Smirnov[1,2], Maja V. Malakhova[3], Andrei E. Samoilov[3], Alexander I. Manolov[3], Anton S. Nazarov[1], Dmitry V. Danilov[1], Svetlana A. Dubiley[4,5], Ilya A. Osterman[5,6], Maria P. Rubtsova[5,6], Elena S. Kostryukova[3], Rustam H. Ziganshin[1], Maria A. Kornienko[3], Anna A. Vanyushkina[3], Olga N. Bukato[3], Elena N. Ilina[3], Valentin V. Vlasov[7], Konstantin V. Severinov[5,8], Alexander G. Gabibov[1] & Sidney Altman[9,10]

[1]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia.
[2]Department of Life Sciences, Higher School of Economics, Moscow, Russia.
[3]Federal Research and Clinical Centre of Physical-Chemical Medicine, Federal Medical and Biological Agency, Moscow, Russia.
[4]Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia.
[5]Skolkovo Institute of Science and Technology, Skolkovo, Russia.
[6]Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia.
[7]Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia.
[8]Waksman Institute for Microbiology Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA.
[9]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, USA.
[10]School of Life Sciences, Arizona State University, Tempe, USA.
Correspondence should be addressed to K. S., A. G. & S. A. (severik@waksman.rutgers.edu, gabibov@ibch.ru, gabibov@gmail.com & sidney.altman@yale.edu).

**This PDF file includes:**

Supplementary text
Figs. S1 to S18
Table S1
References for SI reference citations



www.pnas.org/cgi/doi/10.1073/pnas.1811250115

## Supplementary Information Text

## SI Materials and Methods.

**Microfluidics.** Microfluidic chips with a flow-focusing geometry were produced using the standard soft lithographic techniques, i.e., SU-8/Si masters were used to make PDMS slides with 20×22 μm (width/height) orifices. Inlets and outlets were made using a 1.2 mm biopsy punch (Harris Uni-Core). PDMS slides were bound with glass slides using Atto plasma cleaner (Diener). Immediately after bonding, chips were treated with 1% (w/v) poly(vinyl alcohol) Mowiol 23-88 (Kuraray Specialties Europe) for hydrophilic chips or Aquapel (PPG Industries) for hydrophobic chips. MDE generation was made using piezoelectric pressure controller OB1 MkII (Elveflow) in flow control mode. The inner water streams of cells (IW1) and substrate/antibiotic (IW2) were mixed before the hydrophobic chip with MicroTee (P-890, IDEX). The joint stream was sequentially emulsified in hydrophobic and hydrophilic chips using fluorocarbon oil (O) Novec 7500 with 2% (w/v) Pico-Surf 2 (Dolomite) and outer water (OW) phase (2% (w/v) Pluronic F127 (Sigma-Aldrich), 0.1% Mowiol 23-88, and 50 mM potassium phosphate pH 7.2). The characteristic flow rates were 3:3:3:20-50 μl/min for IW1:IW2:O:OW phases.

**Microbiota collection and storage.** The study was approved by the Local Ethics Committee of the Federal Research and Clinical Centre of Physical-Chemical Medicine (FRCC PCM), conclusion No. 2017/02 from 13.04.2017. All donors provided written informed consent. The healthy donor (D) was examined before collecting the fecal sample. The examination includes: general blood test, biochemical blood test, ELISA blood screening for Lamblias, Toksokary, Opisthorchy, Ascaris and Trichinella, hepatitis B and C test, HIV test, syphilis test, urinalysis, and stool screening for Clostridium difficile, Campylobacter, Salmonella, EIEC, Shigella, Rotavirus, Norovirus, Adenovirus, Cryptosporidium, Cyclospora, Giardia, and Isospora. The patient (P) was diagnosed with ulcerative colitis confirmed by a colonoscopy before the stool sample collection. Stool samples were suspended in 10 volumes of PBS, filtered through 40-μm cell strainers (Greiner Bio-One), and washed with sterile medium for microbiota cryopreservation composed from BHI (BD), 20% glycerol and 30% heat inactivated FBS (Gibco). Oral microbiota samples (B) were collected from a healthy 3-4 year old East Siberian brown bear (Ursus arctos collaris) using noninvasive probe immediately after capture in Siberian taiga (Novosibirsk Region, Maslyaninsky District). Noninvasive probe containing the collected microbiota samples was thoroughly washed with sterile medium for microbiota cryopreservation. The cell suspension was immediately frozen in liquid nitrogen, transported in dry ice and stored in liquid nitrogen.

**Selection of bacteria displaying anti-*S. aureus* activity.** The target *S. aureus* cells producing GFP reporter were vitally stained with sulfo-Cyanine5 NHS (Lumiprobe), washed, filtered using 20-μm solvent filters (A-313, IDEX), and coencapsulated with microbiota suspension in MDE droplets using 20-μm microfluidic chips. Microbiota samples were unfrozen directly prior encapsulation, resuspended in BHI broth (BD), and filtered through 40-μm cell strainers (Greiner Bio-One). After overnight incubation at 35°C, Calcein Violet AM (Thermo Fisher Scientific) was added to the droplet emulsion to the final concentration of 10 μM. Subsequently, the droplets with simultaneous $sCy5^{high}$, $GFP^{low}$, and Calcein $Violet^{high}$ fluorescence were sorted using FACSAria III cell sorter (BD). Bacterial colonies were regenerated after plating on BHI-agar and Columbia agar (Oxoid Ltd., UK) supplied with a 5% of sheep blood and tested for anti-*S. aureus* activity using agar overlay assay. Bacterial clones demonstrating inhibition of *S. aureus* growth were identified by mass spectrometry.

**Identification of bacteria using mass spectrometry.** Bacterial cells were spotted on a sample spot of a MALDI target plate (MSP 96 target, ground steel; Bruker Daltonics) and were overlaid with 2 μL of matrix solution HCCA (saturated solution of α-4-cyano-hydroxycinnamic acid; Bruker Daltonics) in 50% acetonitrile (Sigma-Aldrich) and 2.5% trifluoroacetic acid solution (Sigma-Aldrich). Mass spectra profiles were acquired using a Microflex spectrometer (Bruker

Daltonics). The molecular ions were measured automatically in linear positive ion mode with the instrument parameters optimized for a range of 2,000–20,000 m/z. The software packages flexControl 3.0 (Bruker Daltonics) and flexAnalysis 3.0 (Bruker Daltonics) were used for mass spectra recording and processing. Spectra identification and analysis were carried out using the MALDI Biotyper 3.0 (Bruker Daltonics). The identification was performed by comparing the obtained spectra with the MALDI Biotyper 3.0 library (version 3.2.1.1).

**Antimicrobial activity.** Bacterial collection of clinical isolates was kindly provided by Lytech Co. Ltd. The MICs for bacteria were determined by a serial two-fold dilution (0.1-256 μg/ml) in Mueller-Hinton Broth (BD) or Anaerobe Basal Broth (Oxoid) in anaerobic conditions for *Bacteroides*. MIC was defined as the lowest concentration of amicoumacins that prevented the growth of the test organism in 96-well plate after 10-hour cultivation. Prolonged cultivation is undesirable due to spontaneous Ami self-hydrolysis reaction (Figure 2D) and leads to overestimated MICs. Bacterial growth time course was monitored using absorbance at 600 nm measured using Varioskan Flash multimode reader (Thermo Fisher Scientific).

**NGS sequencing.** The selected MDE droplets were freeze dried and total DNA was isolated using the QIAamp DNA Investigator Kit (Qiagen). Whole-genome amplification was performed using the REPLI-g Single Cell Kit (Qiagen). Fragment libraries were prepared using the NEBNext® DNA Library Prep Reagent Set for Illumina and the NEBNext® Multiplex Oligos for Illumina® (96 Index Primers) (Illumina) according to the manufacturer's instructions. Sequencing of libraries was performed using the genetic analyzer HiSeq2500, the HiSeq® PE Cluster Kit v4 – cBot™ and the HiSeq® SBS Kit v4 (250 cycles) (Illumina) according to the manufacturer's instructions. For individual strains, genomic DNA (100 ng for each sample) was disrupted into 400–550 bp fragments by Covaris S220 System (Covaris, Woburn, Massachusetts, USA). Fragment libraries were prepared using the NEBNext® DNA Library Prep Reagent Set for Illumina and the NEBNext® Multiplex Oligos for Illumina® (96 Index Primers) (Illumina) according to the manufacturer's instructions. Sequencing of libraries was performed using the genetic analyzer HiSeq2500, the HiSeq® PE Cluster Kit v4 – cBot™, and the HiSeq® SBS Kit v4 (250 cycles) (Illumina) according to the manufacturer's instructions.

**Bioinformatics.** Genome assemblies were performed using SPAdes 3.9.0 (1). Genomes were annotated with prokka (2). Bacteria abundance in metagenome samples was assessed with Metaphlan2 (3). Comparative analysis of Ami gene clusters (Fig. 2A) were obtained using EMBOSS Stretcher and resulting protein identity was color-coded. Package genoplotR was used to draw the plot. To detect the presence of the clusters similar to Ami cluster in various genomes, genome annotations of all genomes available from NCBI Refseq database were downloaded in a genebank format and amino acid sequences were blasted to proteins of Ami cluster of *B. pumilus* 124 as reference. Blastp search with 1e-5 e-value threshold was performed. Then we searched for clusters of proteins with blast hits. We demanded that proteins with hits in these clusters were separated from each other no longer than 10 kbp. A chain of such genes with the maximum number of genes in it was further considered. To calculate synteny length, we summed length of all fragments of reference sequence covered by blast HSPs (High Scoring Pairs) to a detected cluster in query genome. To calculate the total blast score we summed scores of best hits for each query sequence. R script to perform these operations is available at github (https://github.com/paraslonic/ami). To build phylogenetic tree, we inferred 16S gene with barrnap (https://github.com/tseemann/barrnap), aligned sequences with MUSCLE v3.8.31 (4) and a Maximum Likelihood tree was built using the PhyML 3.1 tool (5). We performed all reciprocal blastp alignments of protein sequences to obtain maximum likelihood tree for comparison of different biosynthetic clusters plotted in. Blast scores of all hits were summed and normalized by dividing it on the scores of self-hits: $Score(i,j) = Score\,(i,j)/\sqrt{Score(i,i) * Score(j,j)}$, where $i$ and $j$ are different biosynthetic clusters. Then the maximum likelihood tree was built using $1 - Score$ as an input to bionj function from R ape library. Tree then was rooted by the midpoint.

Comparative analysis of gene clusters homologous to Ami cluster (Fig. 3) were obtained using MultiGeneBlast (6). To obtain phylogenetic trees of different *B. pumilis* strains, we inferred one-copy orthologs with Orthofinder 2.2.1 (7), and aligned their nucleotide sequences with MUSCLE v3.8.31. Aligned sequences were concatenated by strain, and a Maximum Likelihood tree was built using the PhyML 3.1 tool. Ami genes in different strains were considered as present if they belonged to the same homology group (inferred with Orthofinder) as genes in *B. pumilus* 124 strain. Taxonomy tree of the bacteria encoding gens homologous to *amiN* was built using blastp hits of *B. pumilus* 124 *amiN* on nt database ordered by e-value using 1e-65 as a cutoff. Multiple cases of organisms with unidentified taxonomy were merged into one record for every e-value range.

**Cultivation of *B. pumilus* 124 and Ami purification.** *B. pumilus* 124 was cultivated using LB agar (BD) plates and 2YT broth (BD) supplemented with 100 mM $NH_3$ and 50 mM potassium phosphate buffer pH 7.5 at 30°C. For metabolomic studies, *B. pumilus* 124 was cultivated in different conditions with «activated» (F) and «inactivated» (N) Ami production. The induction of Ami production (F) was performed by cultivation for 24 hours using 25 $cm^2$ culture flasks without agitation in 5 ml volume. The cultivation conditions with reduced Ami production (N) were achieved by cultivation in 10 ml volume using 50 ml centrifuge tubes with 250 rpm shaking. *B. pumilus* 124 was cultivated for 24 hours using 750 ml flasks in 50 ml volume with 220 rpm shaking for a high-scale Ami production. The cells were centrifuged at 10,000 g for 10 min. Ami was extracted from supernatant with ethylacetate and dried by rotary evaporation. The extract was dissolved in DMSO and fractionated on RP-HPLC Zorbax C8 (Agilent) column using buffer A (20 mM $NH_4OAc$ pH 5.0, 5% ACN), buffer B (20 mM $NH_4OAc$ pH 5.0, 80% ACN); flow rate 5 ml/min; gradient rate 0–10 min (0% B), 10–12 min (0–40% B), 12–16 min (40–52% B), 16–17.5 min (52–67% B), 17.5–19 min (67–100% B). Finally, Ami was purified on Symmetry C18 (Waters) RP-HPLC column using buffer A and B; flow rate 1 ml/min; gradient rate 0–3 min (0% B), 3–4 min (0–30% B), 4–11 min (30–37% B), 11–12 min (37–100% B). Amicoumacin and its derivatives were monitored by absorbance, $\varepsilon_{315\ nm}^{MeOH} = 4380\ M^{-1}cm^{-1}$ for Ami (8).

**Metabolomic analysis of Ami and its derivatives.** Culture medium was collected after 24 hours of cultivation and frozen in liquid nitrogen. Cells count in F and N samples were normalized using absorbance at 600 nm. Cells were washed with 50 mM Tris-HCl pH 7.2 and lysed by vortexing with acid-washed glass beads (Sigma-Aldrich) on ice. Cell lysates were filtered using 3 kDa Amicon Ultra Centrifugal Filters (Millipore) and frozen in liquid nitrogen. Samples were analyzed with an LC-MS/MS system consisting of a nanopump (G2226A) with a four-channel microvacuum degasser (G1379B), a microfluidic chip cube (G4240-64000) interfaced to a Q-TOF mass spectrometer (6530), a capillary pump (G1376A) with degasser (G1379B), and an autosampler with thermostat (G1377A) (all Agilent technologies). All modules were controlled by Mass Hunter software (version B.06.00, Agilent technologies). A microfluidic reversed-phase HPLC chip (Zorbax 300SB-C18, 5-µm particle size, 75-µm i.d., and 150-mm length) was used to separate the samples. A mixture of 96.9% water, 3% acetonitrile, and 0.1% formic acid (v/v) was used as the sample loading buffer and buffer A. Buffer B was 90% ACN, 9.9% water, and 0.1% formic acid (v/v). The samples were loaded on a trap-column at a flow rate of 3 µl/min for 5 min and eluted with a linear 20-min gradient of 0–80% buffer B at a flow rate of 300 nl/min. After each gradient, the column was washed with 100% buffer B for 5 min and reequilibrated with buffer A for 5 min. Raw MS/MS data were analyzed by Agilent MassHunter workstation software (B.06.00) with Qualitative Analysis (6.0.633.0) software. Acceptance criteria included a match to retention time (within 0.015 minutes), isotope spacing and abundance, accurate mass (within 5 ppm), and fragmentation spectra.

**Ami phosphorylation by AmiN and dephosphorylation by AmiO.** *AmiN* and *amiO* genes were amplified from *B. pumilus* DNA by PCR and cloned into DHFR Control Template plasmid (NEB) using *Nde*I/*Kpn*I restriction sites. Recombinant AmiN, AmiO and the respective 6xHis-tagged proteins were produced in *E.coli* BL21(DE3) using autoiduction protocol (9).

Recombinant proteins were purified using metal affinity chromatography on cOmplete™ His-Tag column (Roche), anion exchange chromatography on Mono Q (GE), and size exclusion chromatography on Superdex 75 GL (GE). *In vitro* phosphorylation was performed in AmiN reaction buffer (50 mM Tris-HCl pH 8.0, 10 mM $Mg(OAc)_2$, 0.1 mM EDTA, 1 mM ATP) using 1.4 nM AmiN and 60 μM Ami. *In vitro* dephosphorylation was conducted in AmiO reaction buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 0.2 mM $CoCl_2$, 0.2 mM $ZnCl_2$) using 5 nM AmiO and 40 μM AmiA-P. Reaction mix was analyzed using Symmetry C18 (Waters) RP-HPLC column (1 mL/min, linear gradient 5–80% ACN with 20 mM $NH_4OAc$ pH 5.0, in 20 min). Reaction products were subsequently analyzed by LC-MS/MS spectrometry.

**Reconstitution of AmiN and AmiO activity in *B. subtilis* 168.** *B. subtilis* 168 *yerI* gene encodes putative phosphotransferase, which is highly homologous to *B. pumilus* AmiN kinase. Construction of *B. subtilis* 168 *ΔyerI* strain was performed as described in (10). Briefly, a fragment of the *yerI* gene generated by PCR using yerI_F (5'-TTATATGGATCCTATATGATGGGAGAGATGGAATG-3') and yerI_R (5'-TTATATCTCGAGCGTAGTTGTCCTGGCTTTTC-3') primers was inserted into temperature-sensitive pSC vector between *BamH*I and *Xho*I sites. Electrocompetent *B. subtilis* 168 cells were transformed with pSC-yerI vector and cells carrying insertion at *yerI* gene were selected at 37°C on LB agar supplemented with 5 μg/ml erythromycin. Curing of the integrated vector was carried out using pCRE-PAS plasmid. *B. pumilis amiN* and *amiO* genes were amplified with amiN_F (5'-TTATATGGATCCCAGCATGCATAAAGATGTAAAAGC-3') and amiN_R (5'-TTATATGACGTCTTAAGATTGACTTAGCTTTGTAAAGTCA-3'), and amiO_F (5'-TTATATGGATCCCAGCATGGGCTATATGGATCGTAAAAAAC-3') and amiO_R (5'-TTATATGACGTCTTATGGGCGATAGGTGGACG-3') primer pairs, respectively. Obtained PCR products were inserted into the pHT-01 vector (MoBiTech GmbH) between BamHI and AatII. For the phenotypic assay, the *B. subtilis* 168 Δ*yerI* cells were transformed with either the pHT-amiN, pHT-amiO or control pHT01 plasmid, and expression of the genes was induced with 1mM IPTG.

**MS sample preparation and measurements.** Dry pellets prior to mass spectrometry analysis were resuspended in 500 μl of 20% acetonitrile in water (v/v). After brief rigorous vortexing the samples were shaken for 10 min at 4°C, sonicated in an ice-cooled sonication bath for 10 min, and centrifuged 10 min at 14.000 g at 4°C. Then 900 μl of mixture of 20% acetonitrile in water with 0.1% formic acid (v/v) solution was added to 100 μl of supernatant. After brief rigorous vortexing, 100 μl of diluted sample were transferred to Hamilton glass syringe (100 μl, Hamilton Company, Romania) and placed into the KD Scientific Syringe Pumps (USA). The flow rate was set to 500 nl/min. The mass spectra (Supplementary Fig. 13-18) were acquired in direct injection and positive ionization mode using TripleTOF 5600+ mass spectrometer with a NanoSpray III ion source (AB Sciex, Framingham, MA). We applied the following source parameters: the nanospray needle voltage was set to 3,500 V, curtain gas was set to 50, ion source gas 1 was set to 20, spray temperature was 150°C. MS1 and MS2 spectra were obtained with the following parameters: mass range 50–2000 m/z, number of cycles 100, period cycle time 1000 ms, and accumulation time 1000 ms. Collision-activated dissociation was performed with nitrogen gas with the collision energy ramped from 10 to 70 V with step 1 V per 1 sec.

**Sample preparation for proteomics.** Reduction, alkylation, and digestion of the proteins were performed as described previously (11). Briefly, sodium deoxycholate (SDC) reduction and alkylation buffer pH 8.5 were added to a sample contained 100 μg protein so that the final concentrations of protein, TRIS, SDC, TCEP, and 2-chloroacetamide were 1 mg/ml, 100 mM, 1% (w/v), 10 mM and 40 mM, respectively. The solution was boiled for 10 min and the equal volume of trypsin solution in 100 mM TRIS pH 8.5 was added in a 1:100 (w/w) ratio. After overnight digestion at 37°C, peptides were acidified by 1% trifluoroacetic acid (TFA) for SDB-RPS binding, and 20 μg was loaded on two 14-gauge StageTip plugs. Ethylacetate/1% TFA (125 μl) was added, and the StageTips were centrifuged at 300 g. After washing the StageTips, peptides

were eluted by 60 µl 80% acetonitrile/5% ammonia mixture. The collected material was stored at -80°C. The peptides were suspended in 2% acetonitrile/0.1% TFA buffer and sonicated for 2 min before the analyses.

**Liquid chromatography and mass spectrometry of proteome samples.** Approximately 1 µg of peptides were loaded for 2 h gradient. Peptides were separated on a 25-cm 75-µm i.d. column packed in-house with Aeris Peptide XB-C18 2.6 µm resin (Phenomenex). Reverse-phase chromatography was performed with an Ultimate 3000 Nano LC System (Thermo Fisher Scientific), which was coupled to the Q Exactive HF benchtop Orbitrap mass spectrometer (Thermo Fisher Scientific) via a nanoelectrospray source (Thermo Fisher Scientific). Peptides were loaded in buffer A (0.2% (v/v) formic acid) and eluted with a linear 120-min gradient of 4–45% buffer B (0.1% (v/v) formic acid, 80% (v/v) acetonitrile) at a flow rate of 350 nl/min. The column was washed with 95% buffer B for 5 min and reequilibrated with buffer A for 5 min after each gradient. Column temperature was maintained at 40°C. Peptides were analyzed on a mass spectrometer, with one full scan (300–1,400 *m*/*z*, *R* = 60,000 at 200 *m*/*z*) at a target of 3e6 ions and max ion fill time 60 ms, followed by up to 15 data-dependent MS/MS scans with higher-energy collisional dissociation (HCD) (target 1e5 ions, max ion fill time 30 ms, isolation window 1.2 m/z, normalized collision energy (NCE) 28%, underfill ratio 2%), detected in the Orbitrap (*R* = 15,000 at fixed first mass 100 *m*/*z*). Other settings: charge exclusion - unassigned, 1, >6; peptide match – preferred; exclude isotopes – on; dynamic exclusion - 30 s was enabled.

**Proteomic data analysis.** Label-free protein was quantified by MaxQuant software version 1.5.6.5 and a common contaminants database by the Andromeda search engine (12), with cysteine carbamidomethylation as a fixed modification and N-terminal acetylation and methionine oxidations as variable modifications. Peak lists were searched against the database of CDS sequences which was obtained after genome annotation with prokka (2). The FDR was set to 0.01 for both proteins and peptides with a minimum length of seven amino acids. Peptide identification was performed with an allowed initial precursor mass deviation up to 20 ppm and an allowed fragment mass deviation of 20 ppm. The downstream bioinformatics analysis was carried out with Perseus (13) (versions 1.5.5.1). Protein groups only identified by site, only from peptides identified also in the reverse database, or belonging to the common contaminants database, were excluded from the analyses. For Student's t-test, missing values were imputed with a width of 0.3 and a downshift of 1.8 over the total matrix. Two sample tests were performed in Perseus with s0 set to 0. Label free quantification was performed with a minimum ratio count of 1 (14). To quantify proteins in each sample, we used the iBAQ algorithm, implemented into MaxQuant software (15). The normalization of each protein's iBAQ value to the sum of all iBAQ values generates a relative iBAQ (riBAQ) value corresponding to the mole fraction of each protein (16).

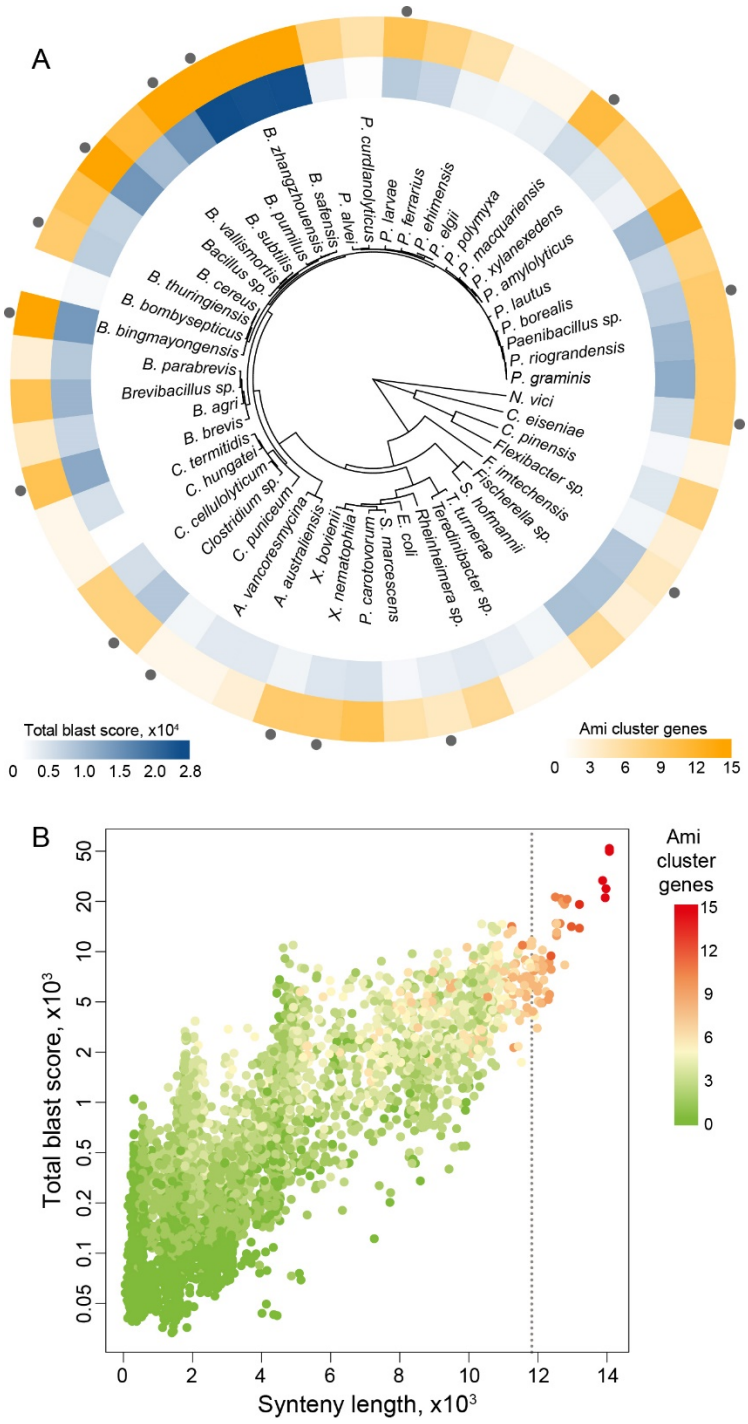


**Fig. S1.** *(A)* Phylogenetic tree with a total blast score to *B. pumilus* 124 Ami cluster and homologous genes count coded with color. The best match from all strains of species is chosen. The species with the top 50 synteny length are illustrated. Representative species with clusters discussed in paper (Fig. 3) are designated with gray dots. Phylogeny is reconstructed based on 16S gene. *(B)* Total blast score to *B. pumilus* 124 Ami cluster and synteny length of clusters in different species. Count of different Ami genes encoded with a color gradient. The dotted line shows the top 50 hits ordered by a synteny length.

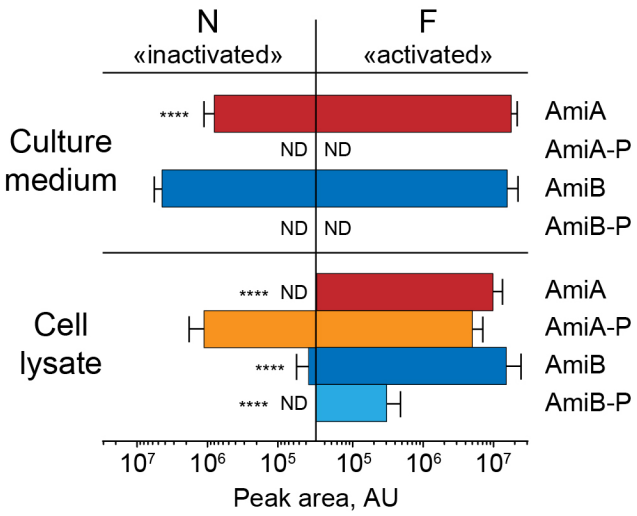


**Fig. S2.** Comparative metabolomic analysis of Ami and its derivatives after *B. pumilus* 124 cultivation in «activated» (F) and «inactivated» (N) conditions. **** represents how F differs from N, $p<0.0001$. The data represent mean ± SD. ND indicates that the peak area is less than $10^4$ AU. AmiA (Ami) denotes amicoumacin A, AmiB – amicoumacin B, AmiA-P and AmiB-P – phosphorylated AmiA and AmiB, respectively.

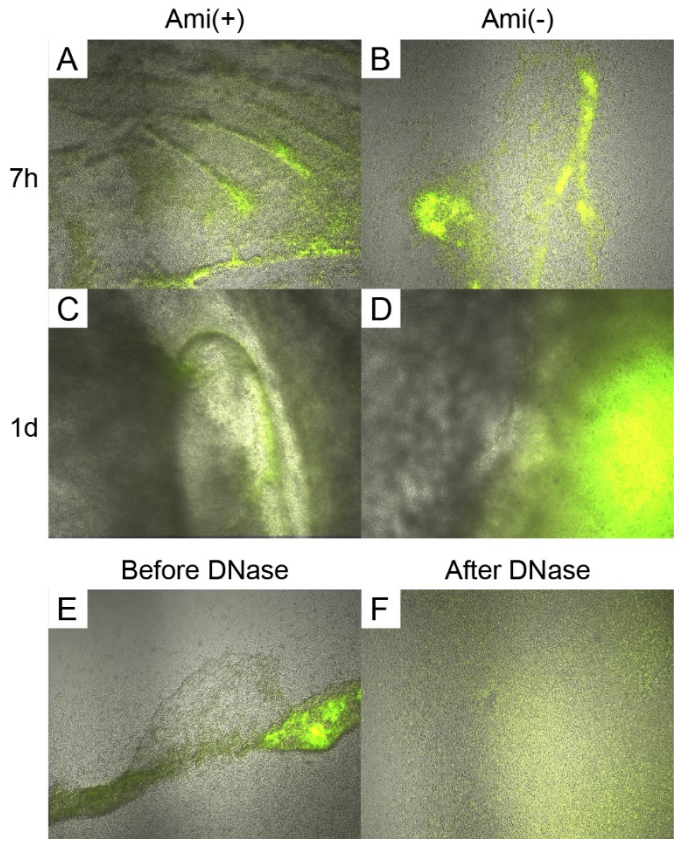


**Fig. S3.** Slimy biofilm produced by *Bacillus pumilus* and its role in suppressing the cell growth of *Staphylococcus aureus*. Coculture of *S. aureus* producing GFP reporter and Ami producing *B. pumilus* 124 *(A, C)* or reference *B. pumilus* 123 without Ami cluster *(B, D)* after 7 hours and 24 hours of cocultivation respectively. Coculture of *S. aureus* and *B. pumilus* 124 following 9 hours of coincubation prior *(E)* and after *(F)* treatment with DNase for 10 min.

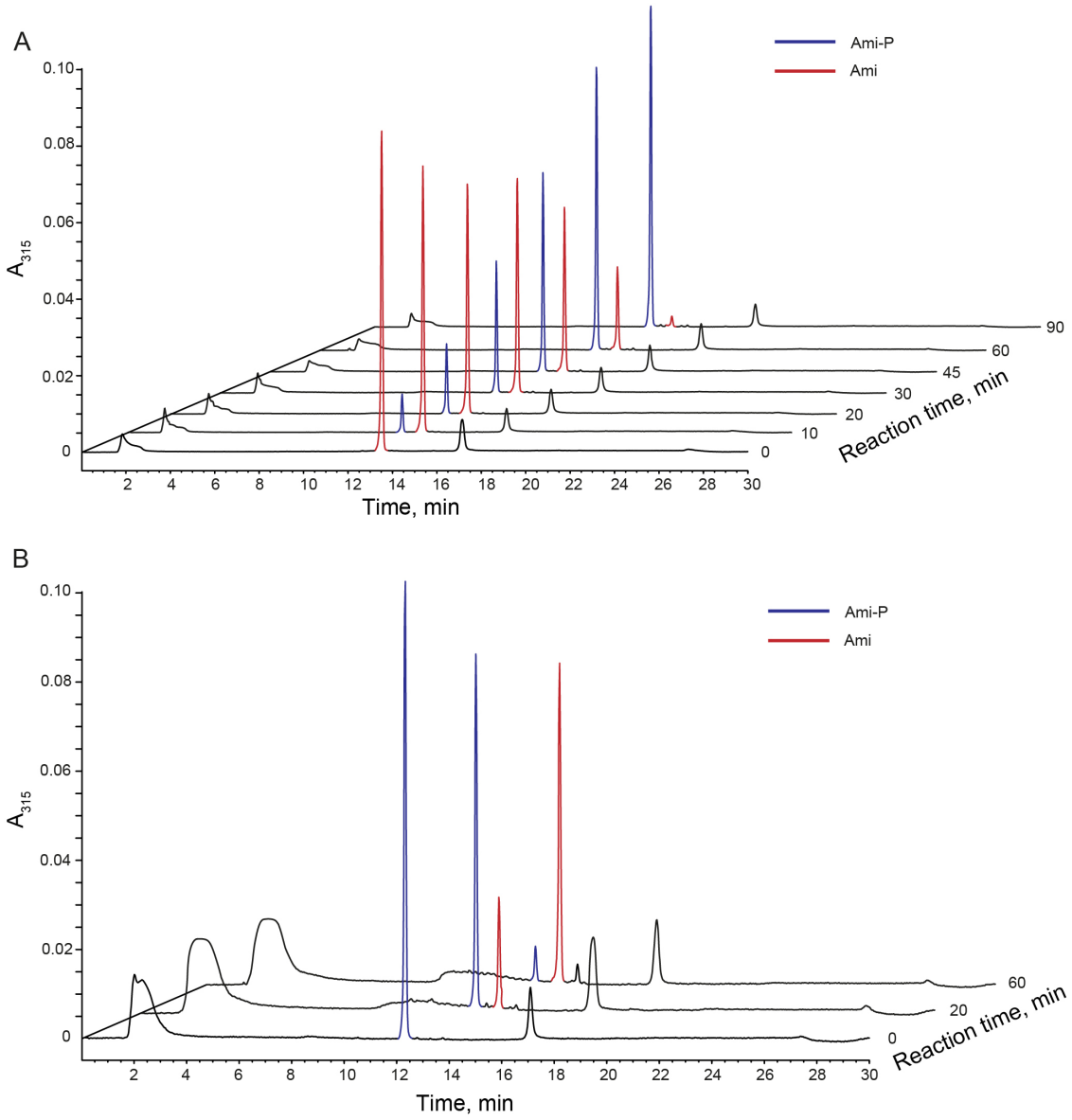


**Fig. S4.** *In vitro* activity of enzymes AmiN and AmiO. *(A)* Phosphorylation of Ami by kinase AmiN. *(B)* Dephosphorylation of the phosphorylated Ami by phosphatase AmiO. Ami is marked with blue; the phosphorylated Ami - with red. Ami-P stands for the phosphorylated Ami.

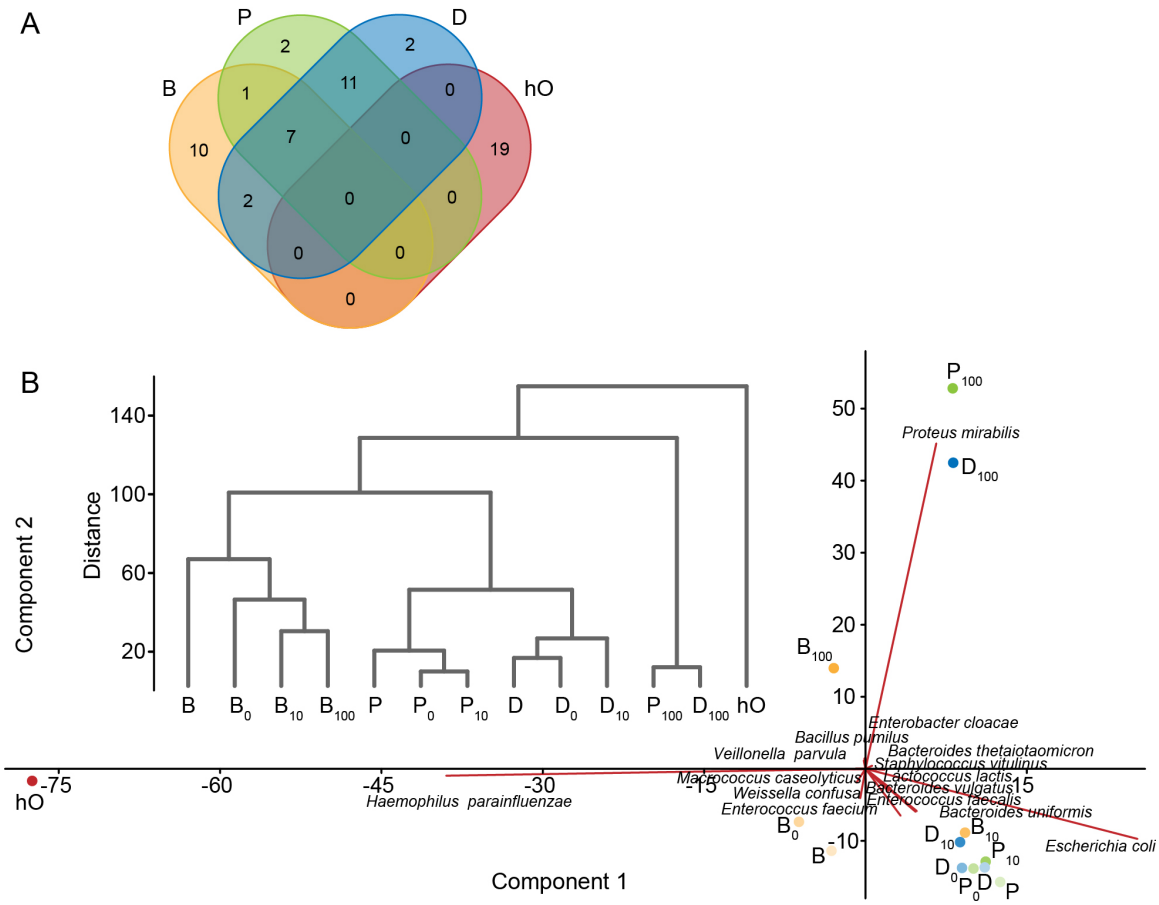


**Fig. S5.** Composition of microbiota samples before and after cultivation in MDE microcompartments. *(A)* Venn diagram shows composite resemblance between isolated microbiota samples on the level of bacterial species. *(B)* Principal component analysis of microbiota samples before and after cultivation in droplets. Bear oral microbiota (B), human fecal microbiota from a patient with colitis (P), a healthy human donor (D), and human oral microbiota (hO). Subscripts indicate the samples selected after cultivation in droplets with various Ami concentrations (0, 10 and 100 µg/ml respectively).


**Fig. S6.** Cultivation of the microbiota samples in droplets results in the fast-growing bacteria enrichment as well as depletion of particular slow-growing or unculturable population. The heat map indicates bacterial species that are culturable in MDE microcompartments (Culturable), partially depleted after cultivation in MDE (Slow-growing), or deplete below the level of less than 0.1% in overall population (Rare/unculturable).

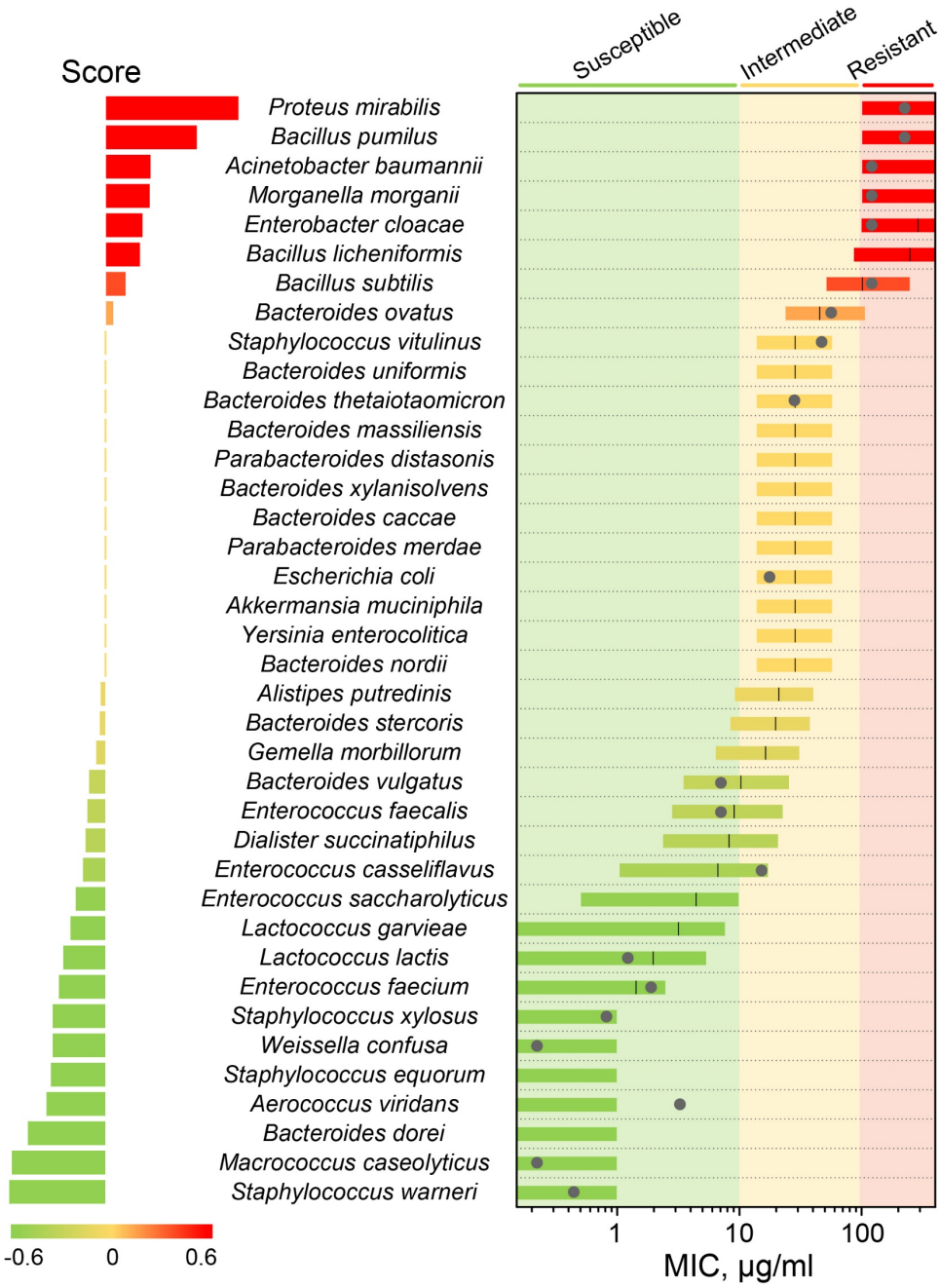


**Fig. S7.** Values of scores obtained from the shift of microbiota composition after single-cell cultivation in MDE compartments in the presence of Ami (left). Bars of the plot (right) represent confidential intervals (α = 0.05) of predicted MIC obtained according to the scores (left). Grey points indicate mean MIC obtained *in vitro* for particular bacterial strains, the vertical lanes show mean values of predicted MIC.

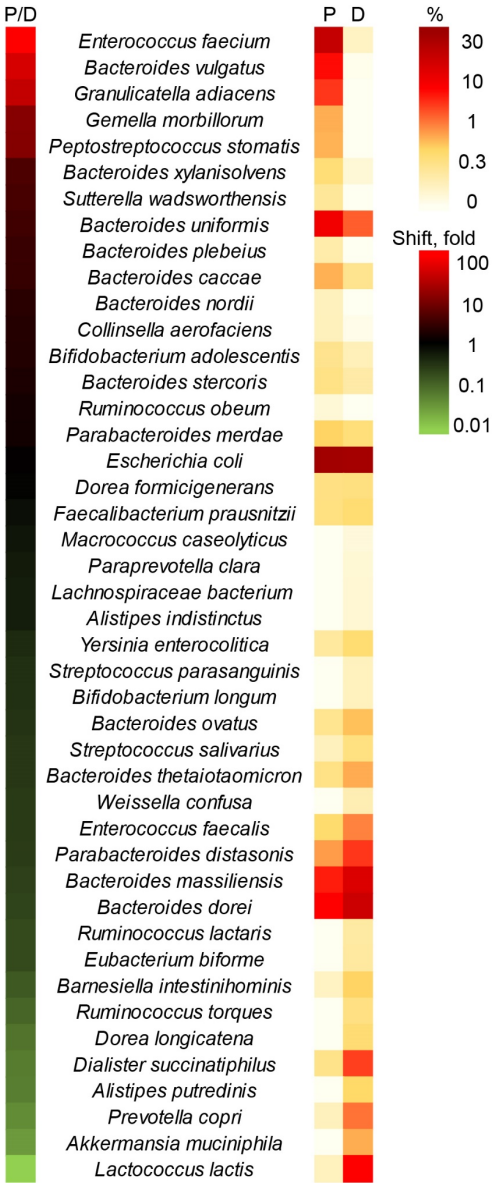


**Fig. S8.** Comparison between microbiota of the patient with colitis (P) and a healthy donor (D). The heat maps indicate the portion of bacteria in each microbiota sample (left) and bacterial species that are overrepresented (red) or underrepresented (green) in P vs D (P/D).

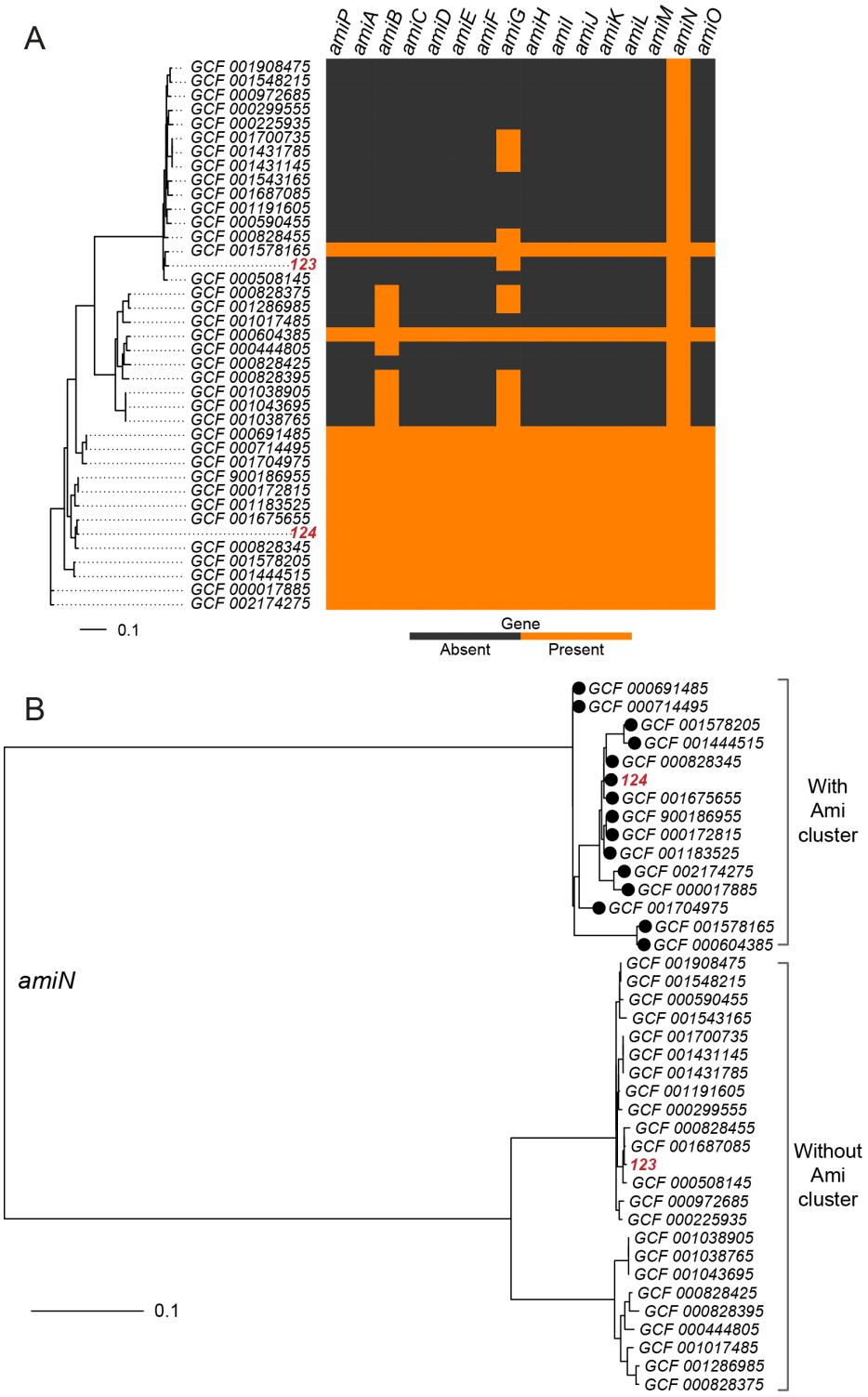


**Fig. S9.** *(A)* Presence of Ami cluster genes in different strains of *B. pumilus* is designated with orange color. Unrooted phylogenetic tree is shown on the left. *(B)* Maximum likelihood tree based on alignment of *amiN* gene. The circles designate the strains with Ami cluster present. A huge difference is observed between *amiN* genes in strains with and without cluster.

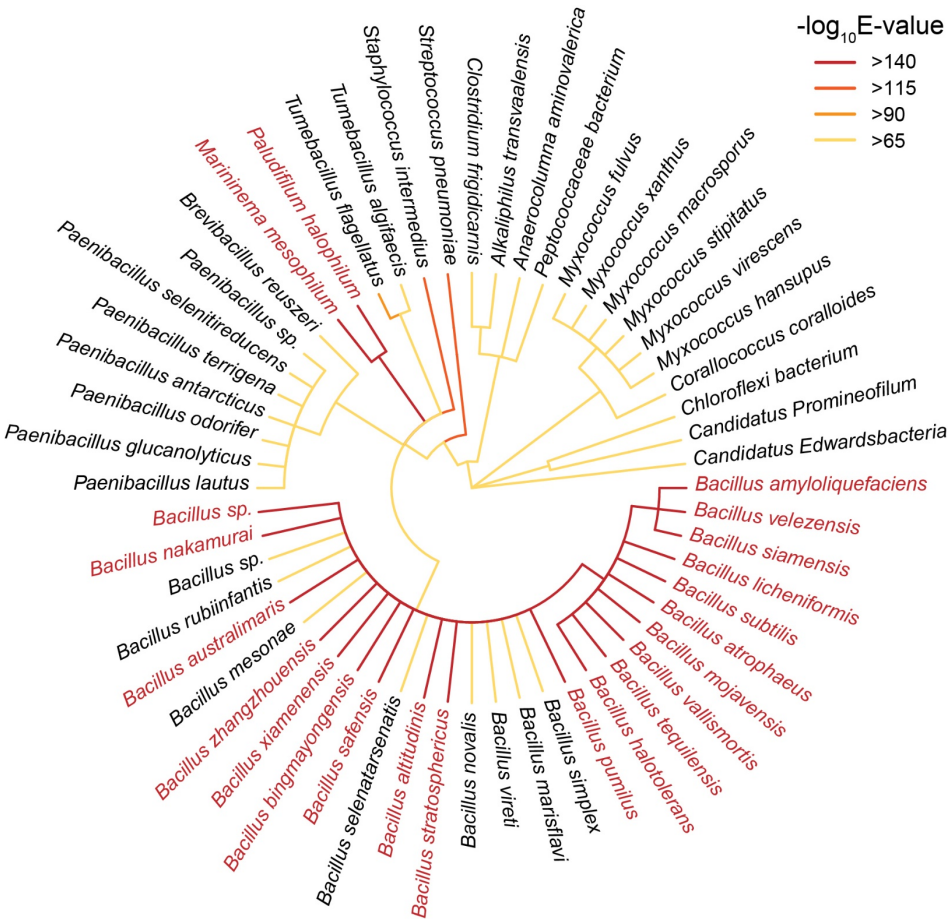


**Fig. S10.** Taxonomy tree of bacteria encoding gens homologous to *amiN* of *B. pumilus* 124, color represents e-value of blastp hit.

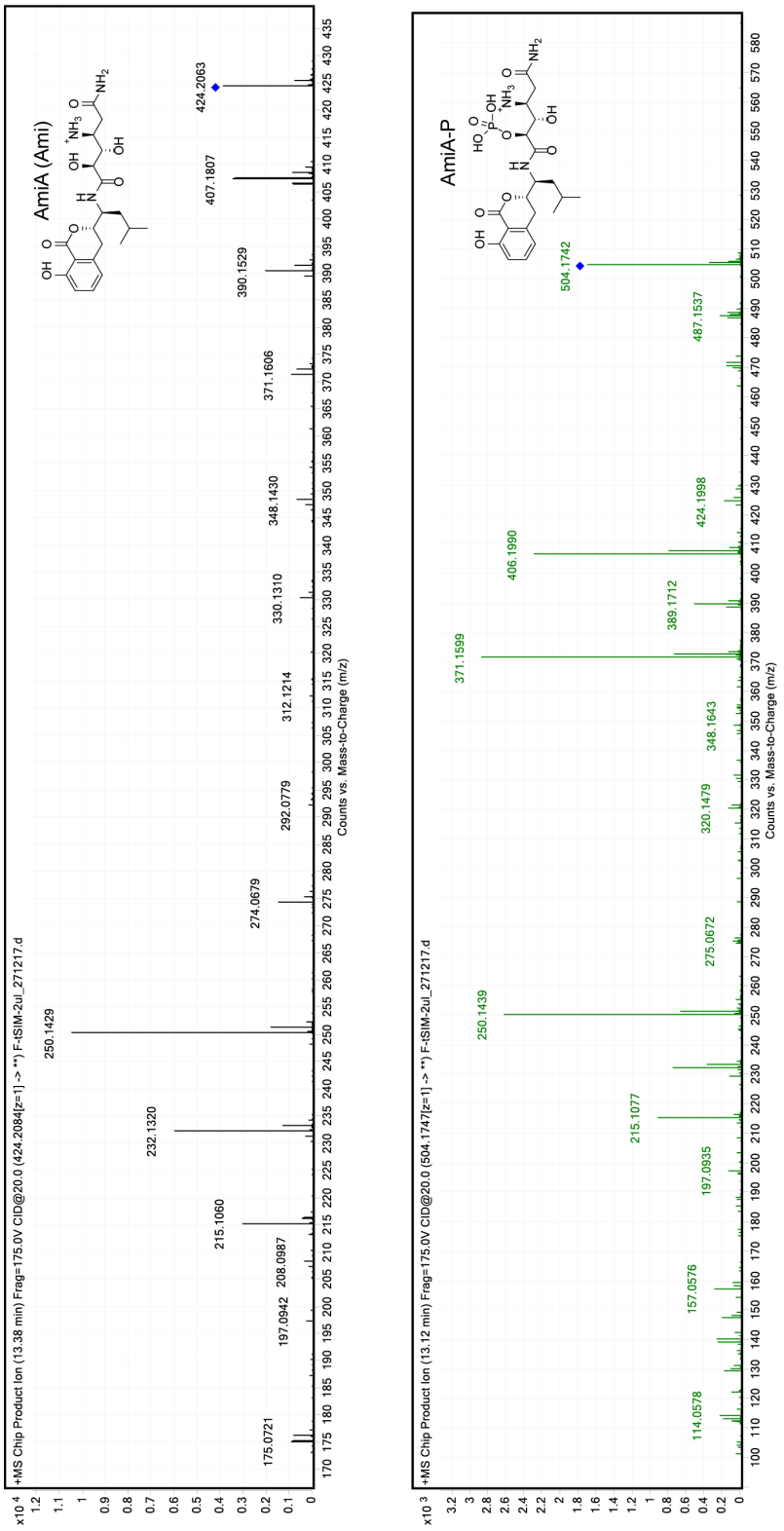


**Fig. S11.** MS/MS spectrum of amicoumacin A (AmiA) and its phosphorylated analog (AmiA-P).

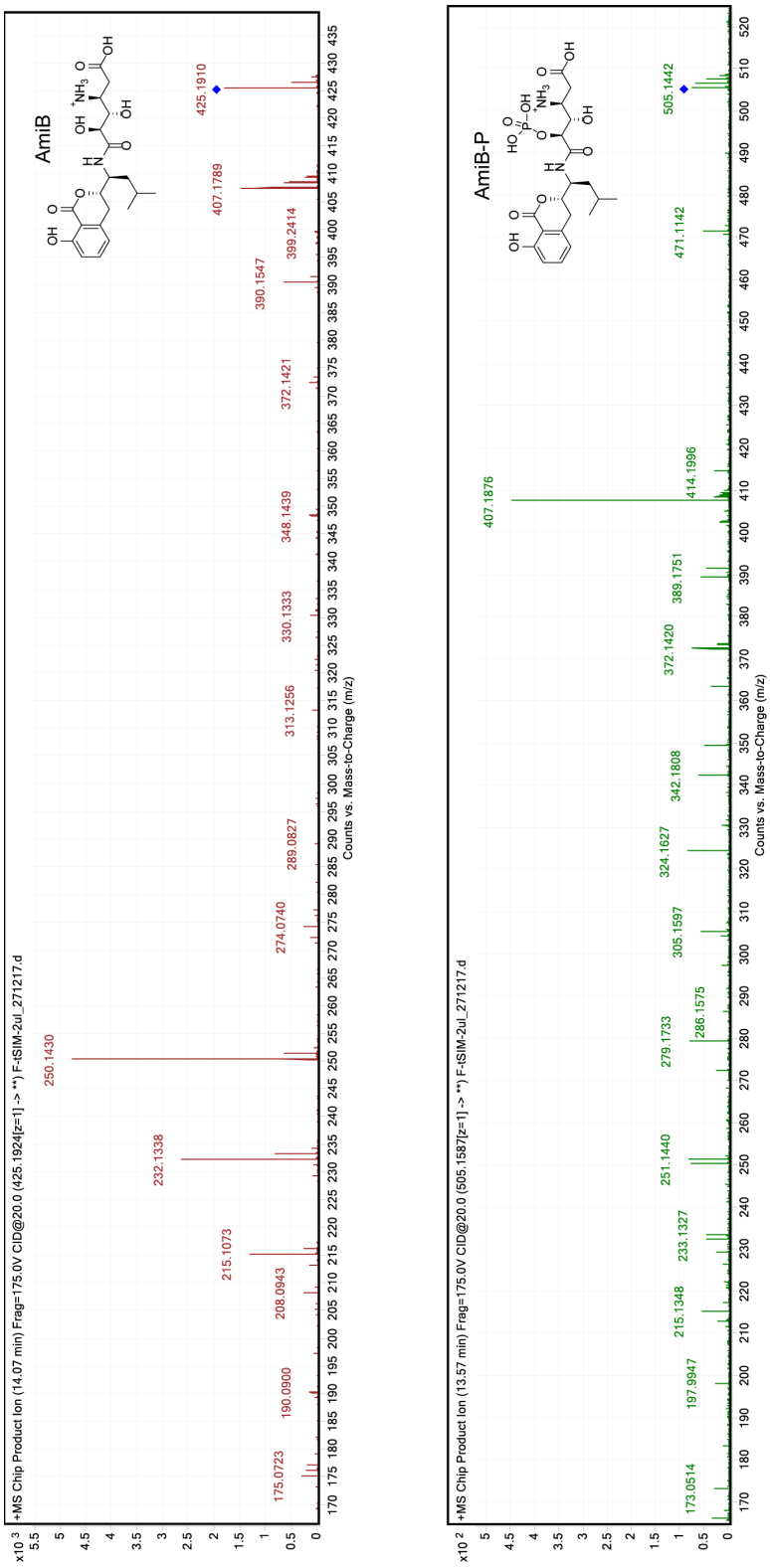


**Fig. S12.** MS/MS spectrum of amicoumacin B (AmiB) and its phosphorylated analog (AmiB-P).

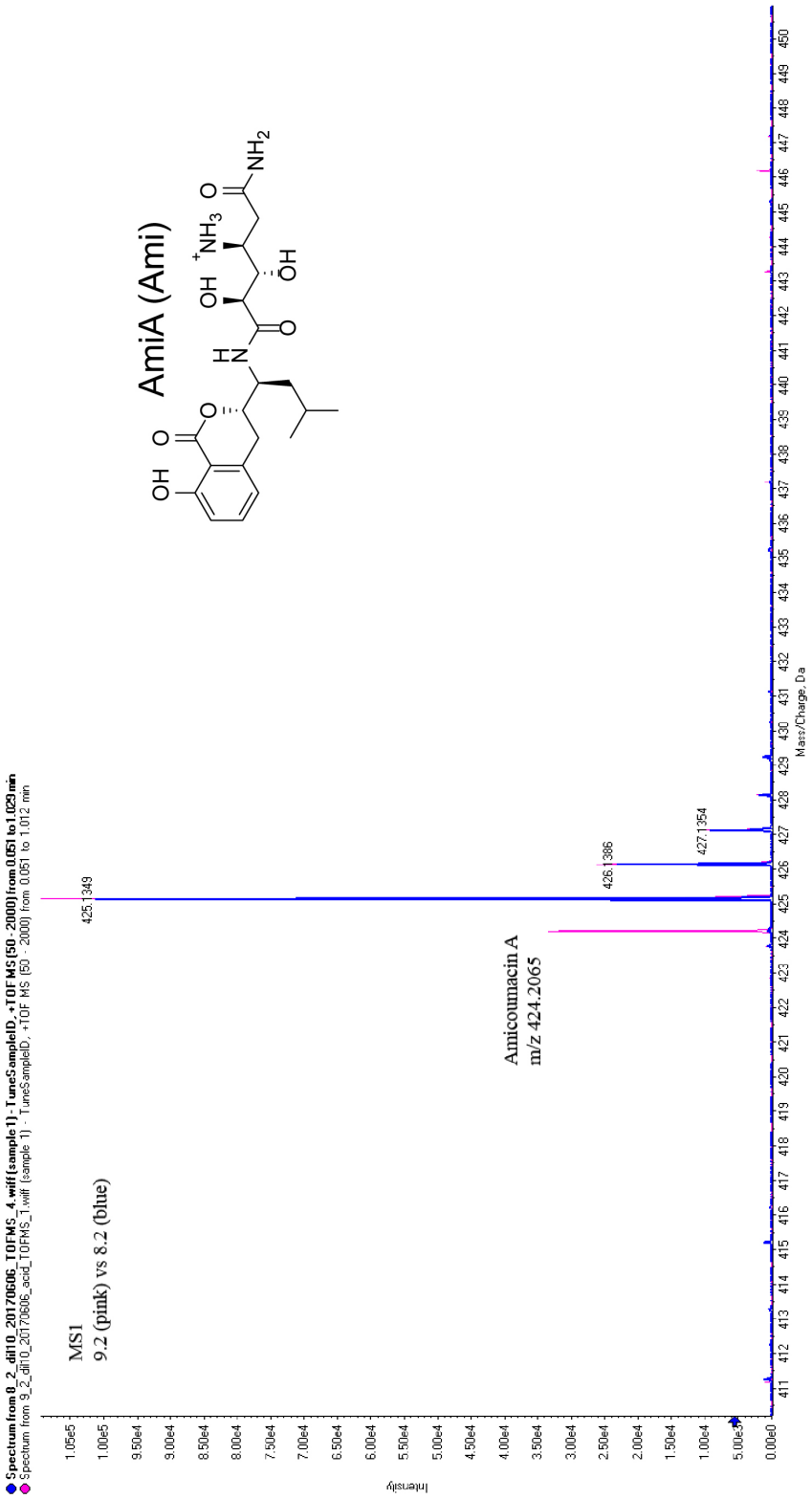


**Fig. S13.** MS spectra of amicoumacin A in direct injection, positive ionization mode in the scan range 410-450 m/z. Amicoumacin A [M+H]+ m/z=424.2065, Δ=5 ppm.

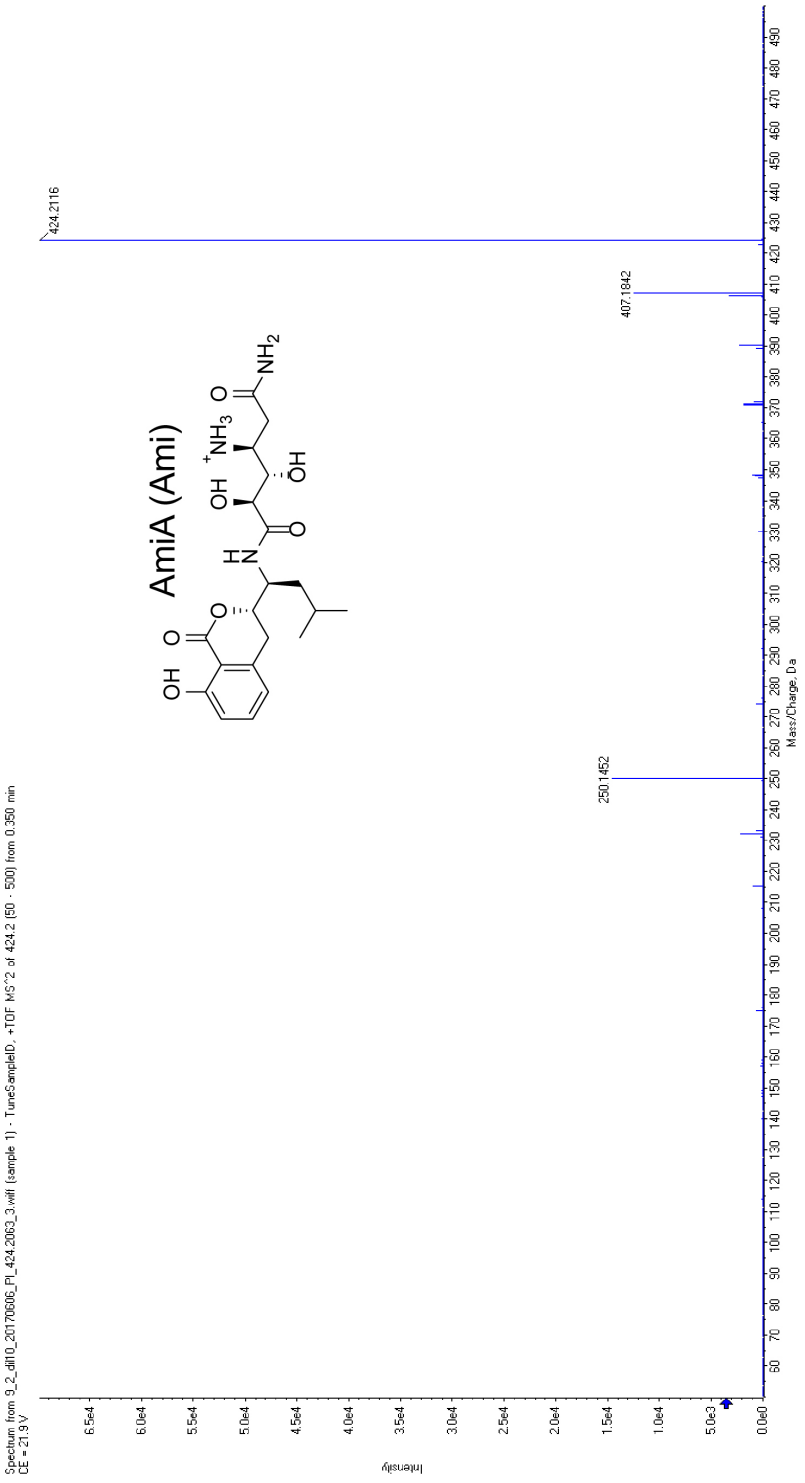


**Fig. S14.** MS/MS spectrum of amicoumacin A ion. Fragmentation spectrum was obtained through direct injection method in positive ionization mode with collision energy 21.9 eV (ramp mode used); amicoumacin A parent ion [M+H]+ m/z=424.2116, Δ=8 ppm. Detected fragments m/z 250.1452, 407.1842.

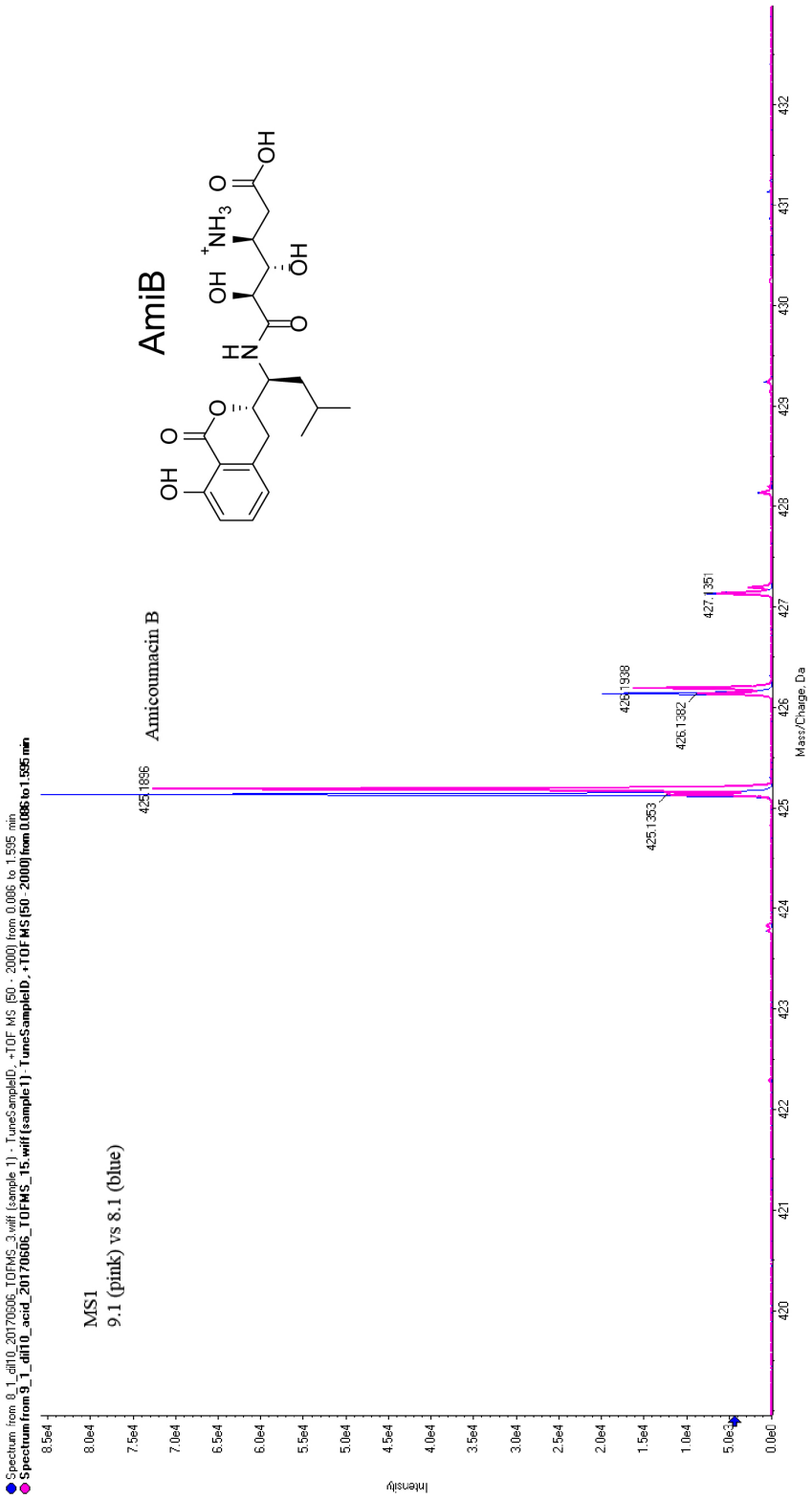


**Fig. S15.** Full-scan MS spectra of amicoumacin B in direct injection, positive ionization mode in scan range 419-432 m/z. Amicoumacin B [M+H]+ m/z=425.1896, Δ=7 ppm.

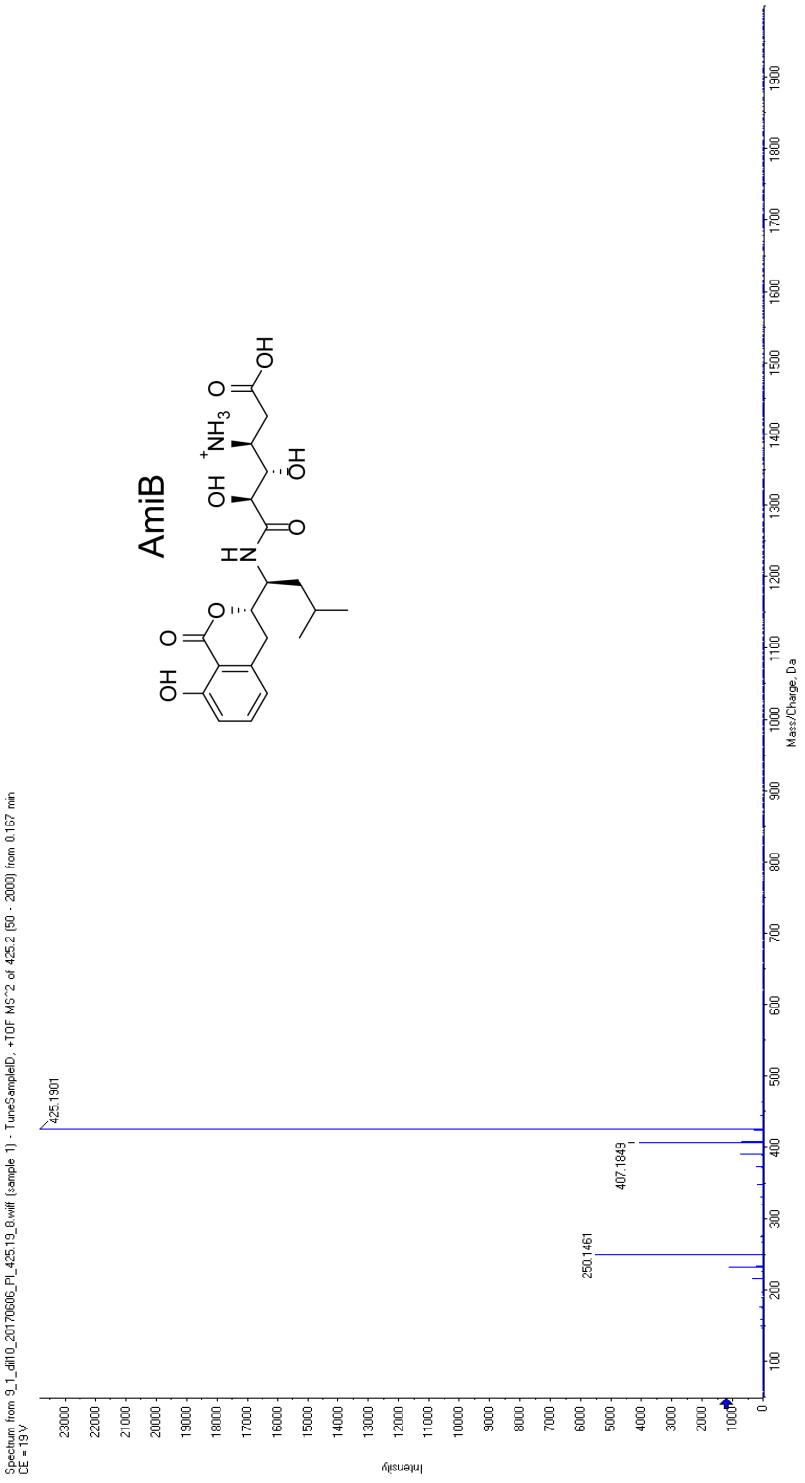


**Fig. S16.** MS/MS spectrum of amicoumacin B ion. Fragmentation spectrum was obtained using direct injection method in positive ionization mode with collision energy 19 eV (ramp mode used); amicoumacin B parent ion [M+H]+ m/z=425.1901, Δ=5 ppm. Detected fragments m/z 250.1461, 407.1849.

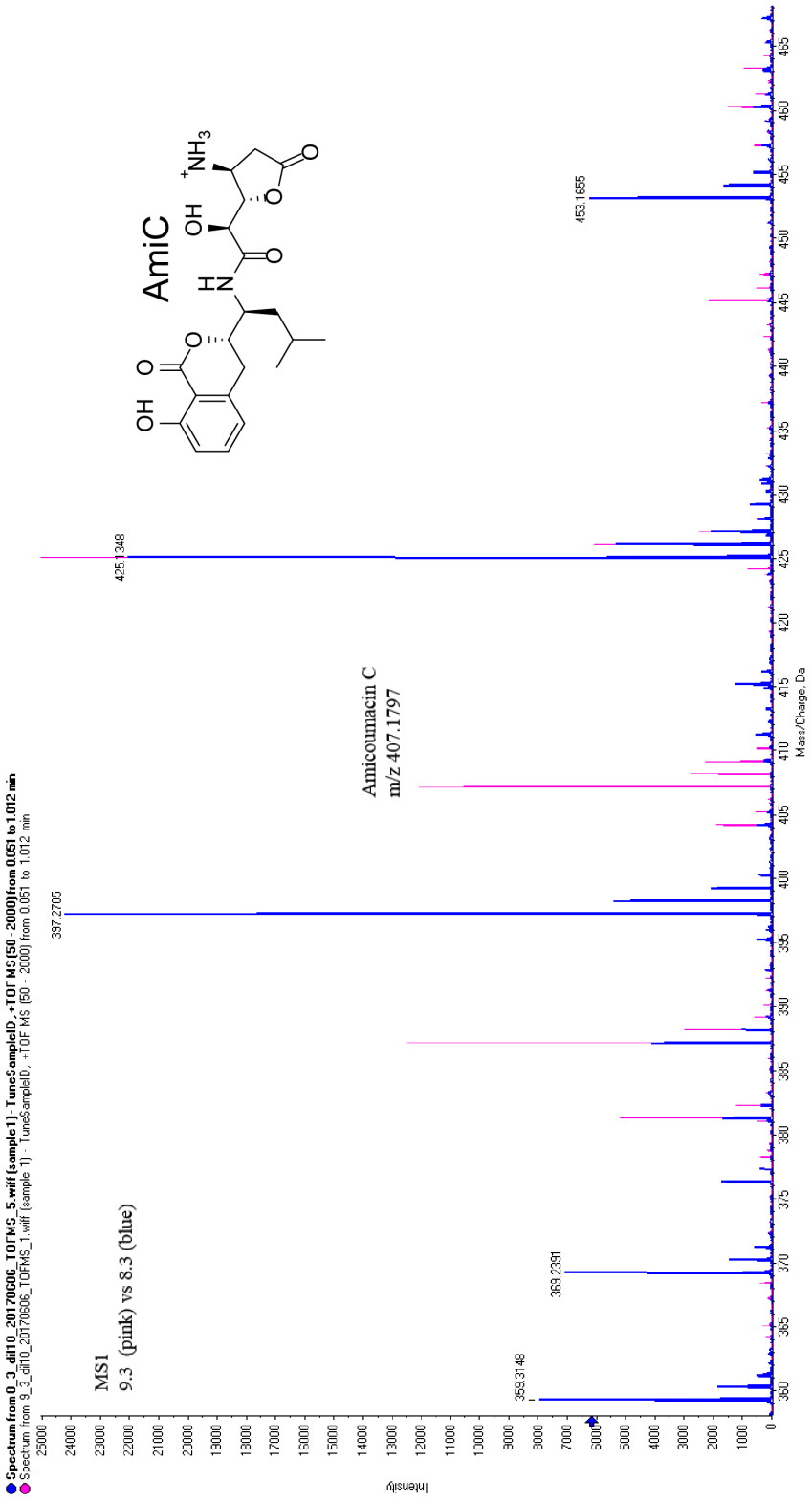


**Fig. S17.** Full-scan MS spectra of amicoumacin C in direct injection, positive ionization mode in the scan range 360 - 465 m/z. Amicoumacin C [M+H]+ m/z=407.1797, Δ=4 ppm.

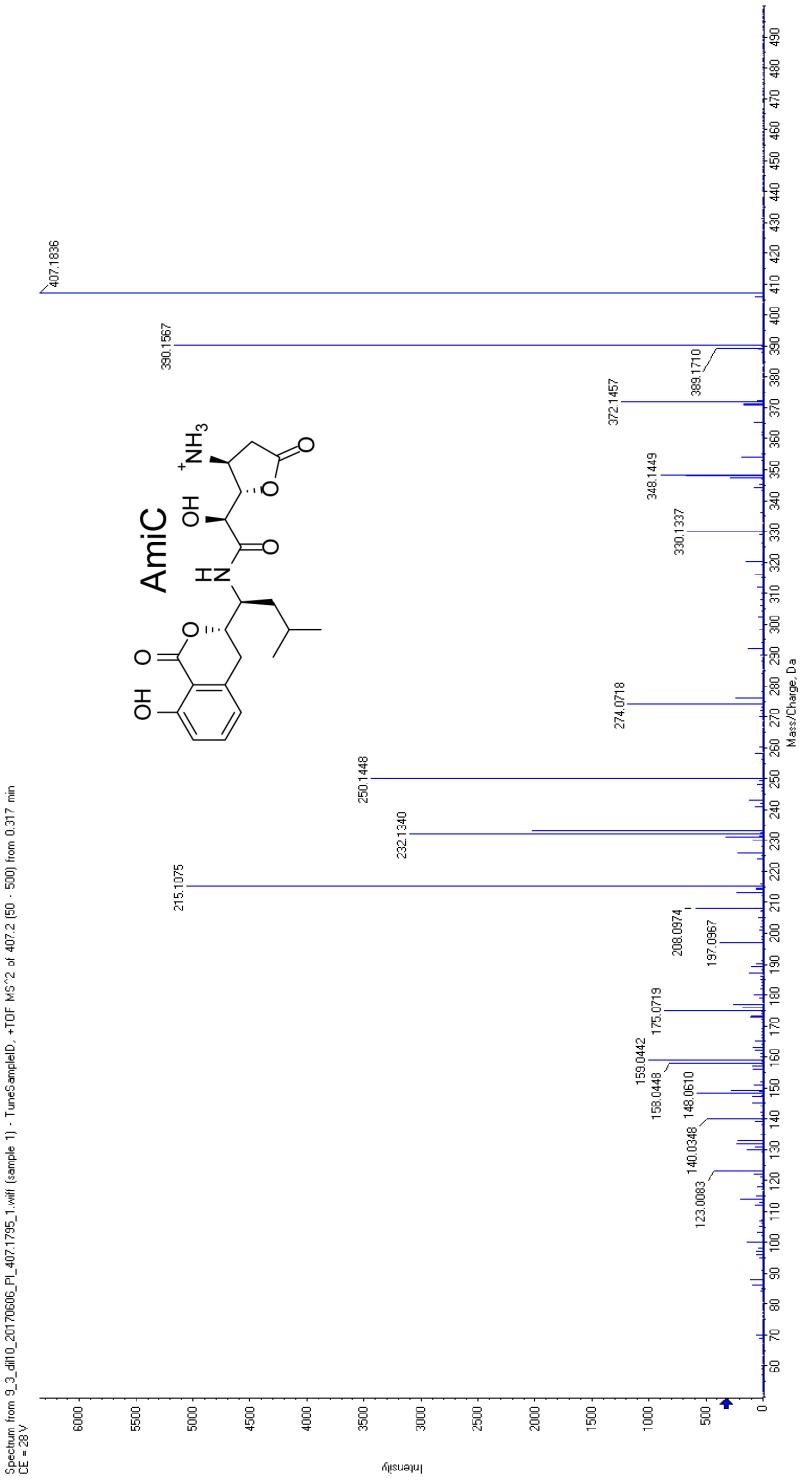


**Fig. S18.** MS/MS spectrum of amicoumacin C ion. Fragmentation spectrum was obtained using direct injection method in positive ionization mode with collision energy 28 eV (ramp mode used); amicoumacin C parent ion [M+H]+ m/z=407.1836, Δ=5 ppm. Detected fragment m/z 250.1448, 215.1075, 232.1340, 390.1567.

**Table S1. The influence of heterologous expression of *amiN* on MIC of model bacteria *E. coli* BL21(DE3), *B. subtilis* 168 and *B. subtilis* 168 *ΔyerI* gene knock out.**

| Strain | MIC, μg/ml | |
|---|---|---|
| | WT | +*amiN* |
| *E. coli* BL21(DE3) | 16 | >100 |
| *B. subtilis* 168 | >100 | – |
| *B. subtilis* 168 *ΔyerI* | 0.8 | >100 |

## References


1. Bankevich A, *et al.* (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19(5):455-477.
2. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068-2069.
3. Truong DT*, et al.* (2015) MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods* 12:902.
4. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5(1):113.
5. Guindon S*, et al.* (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* 59(3):307-321.
6. Medema MH, Takano E, & Breitling R (2013) Detecting Sequence Homology at the Gene Cluster Level with MultiGeneBlast. *Molecular Biology and Evolution* 30(5):1218-1223.
7. Emms DM & Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* 16(1):157.
8. Itoh J, Omoto S, Nishizawa N, Kodama Y, & Inouye S (1982) Chemical Structures of Amicoumacins Produced by *Bacillus pumilus*. *Agricultural and Biological Chemistry* 46(11):2659-2665.
9. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41(1):207-234.
10. Pomerantsev AP, Camp A, & Leppla SH (2009) A New Minimal Replicon of Bacillus anthracis Plasmid pXO1. *Journal of Bacteriology* 191(16):5134-5146.
11. Kulak NA, Pichler G, Paron I, Nagaraj N, & Mann M (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods* 11:319.
12. Cox J & Mann M (2011) Quantitative, High-Resolution Proteomics for Data-Driven Systems Biology. *Annual Review of Biochemistry* 80(1):273-299.
13. Tyanova S*, et al.* (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* 13:731.
14. Cox J*, et al.* (2014) Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics* 13(9):2513-2526.
15. Schwanhäusser B*, et al.* (2011) Global quantification of mammalian gene expression control. *Nature* 473:337.
16. Shin J-B*, et al.* (2013) Molecular architecture of the chick vestibular hair bundle. *Nature Neuroscience* 16:365.