**SUPPORTING INFORMATION**

# Endophytes are Hidden Producers of Maytansine in *Putterlickia* Roots

Souvik Kusari,*[,†] Marc Lamshöft,[†,‖,#] Parijat Kusari,[‡,#] Sebastian Gottfried,[†] Sebastian Zühlke,[†] Kathrin Louven,[†] Ute Hentschel,[§] Oliver Kayser,[‡] Michael Spiteller*[,†]

[†]Institute of Environmental Research (INFU), Department of Chemistry and Chemical Biology, Chair of Environmental Chemistry and Analytical Chemistry, TU Dortmund, Otto-Hahn-Str. 6, D-44221 Dortmund, Germany

[‡]Department of Biochemical and Chemical Engineering, Chair of Technical Biochemistry, TU Dortmund, Emil-Figge-Str. 66, D-44227 Dortmund, Germany

[§]Department of Botany II, Julius-von-Sachs Institute for Biological Sciences, University of Würzburg, Julius-von-Sachs-Platz 3, 97082 Würzburg, Germany

[‖]Present address: Bayer CropScience, Alfred-Nobel-Str. 50, 40789 Monheim, Germany

---

***Corresponding Authors**

Tel.: +49-231-755-4086. Fax: +49-231-755-7484. E-Mail: souvik.kusari@infu.tu-dortmund.de

Tel.: +49-231-755-4080. Fax: +49-231-755-7485. E-Mail: m.spiteller@infu.tu-dortmund.de

---

[#]These authors contributed equally to this work

**TABLE OF CONTENTS**

## I. GENERAL EXPERIMENTAL PROCEDURES

***In vitro* biosynthesis of maytansine by root-specific endophytic bacterial community.** For isolation of endophytes, the same *Putterlickia verrucosa* and *Putterlickia retrospinosa* plants were independently sampled twice (in 2013 and 2014). The primary as well as secondary roots of both plants containing their endophytic microbial community were thoroughly washed in running tap water. Small fragments of approximately 20 mm (length) were cut with the aid of a razor blade. The excised explants were washed thoroughly under running tap water followed by deionized water (DI) to remove any dirt attached to them. The cut tissues were air-dried at RT for 15 min and their weight was noted. The surface sterilization of these root tissues was done following previously established procedures,[1] suitably modified. Briefly, the small tissue fragments were surface sterilized by sequential immersion in 70% ethanol for 1 min, 1.3 M sodium hypochlorite (3-5% available chlorine) for 3 min, and 70% ethanol for 30 s. Finally, these surface-sterilized tissue pieces were rinsed thoroughly in sterile, double-distilled water for a couple of minutes, to remove excess surface sterilants. The excess moisture was blotted on sterile filter paper. The surface-sterilized tissue fragments were placed in sterile mortar-pestle and crushed with the addition of sterile double-distilled water. The macerated tissues, thus obtained, were immersed in 250 mL Erlenmeyer flasks, each containing 50 mL of optimized *Streptomyces* broth medium (SM; Sigma-Aldrich, cat. no. 85883). As negative control, fully sterilized stem and root segments were suspended in SM in parallel after autoclaving the plant tissues, wrapped in aluminum foil, at 121 °C at 15 lbs pressure for 15 min. Unseeded SM served as a blank control. Each setup was prepared in triplicates. The seeded and unseeded flasks were incubated at 28 ± 2 °C with shaking (150 rev $min^{-1}$) on a rotary shaker (INFORS HT Multitron 2, Einsbach, Germany). Samples (1 mL) were aseptically drawn from each flask at zero hour (2 h after initial seeding), and at an interval of 12 h till 60

h, for extraction and determination of maytansine. The optical density at 570 nm of the suspension was taken as a measure of microbial growth. The endophytic bacterial community harbored in the stems and leaves of both plants were isolated and processed in the same manner.

**Plate-based microbiological assay (endophytic community versus *Hamigera avellanea*).** The test organism, a type strain of *Hamigera avellanea* (syn. *Penicillium avellaneum*; DSM 2208; obtained from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, DSMZ, Braunschweig, Germany; the activation of the bacterial strain was performed according to DSMZ guidelines) was cultured in 100 mL potato dextrose broth (PDB; Sigma-Aldrich, cat. no. P6685). After incubation at 28 ± 2 °C with proper shaking (150 rev $min^{-1}$) for 72 h, the medium was decanted. The *H. avellanea* biomass, thus obtained, was transformed into a suspension with a sterile mortar and pestle through the addition of 4 mL sterile double-distilled water. For the endophytic microbial communities, 24-hour liquid cultures (*Streptomyces* broth medium; SM; Sigma-Aldrich, cat. no. 85883) were prepared and used for the assay, since the production of maytansine was detected at this time-point using HPLC-ESI-$HRMS^n$. For the plate-based assay, 15 µL of the endophytic community cultures were transferred onto sterile antibiotic test discs with a diameter of 6 mm (Macherey-Nagel, cat. no. MN 827 ATD) and dried under the biological safety cabinet, also shown schematically in Figure 2A in the manuscript. In the meantime, 200 µL of *H. avellanea*-suspension was spread on plates containing SM agar medium (1.7% agar). The discs were then transferred onto the plates using sterile forceps. Unseeded sterile discs served as blank. The plates were sealed using Parafilm, then incubated at 28 ± 2 °C, and observed every 24 h for 5 days.

**Genomic DNA extraction, PCR amplification and sequencing.** The total genomic DNA (gDNA) of the bacterial endophytic communities, positive control *Actinosynnema pretiosum* subsp. *auranticum* (DSM 44131; obtained from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, DSMZ, Braunschweig, Germany; the activation of the bacterial strain was performed according to DSMZ guidelines) and negative control *Escherichia coli* (DSM 682; also obtained from DSMZ; the activation of the bacterial strain was performed according to DSMZ guidelines) were extracted from the *in vitro* cultures using peqGOLD bacterial DNA kit (Peqlab Biotechnologie GmbH, Germany). Briefly, a set of conical flasks with 500 mL capacity each with 100 mL *Streptomyces* broth medium (SM; Sigma Aldrich, cat. no. 85883) were used with proper autoclaving. The bacterial communities and positive control strain were inoculated in respective flasks and incubated at 28 ± 2$^{o}$C with proper shaking (150 rev min$^{-1}$) on a rotary shaker (INFORS HT Multitron 2, Einsbach, Germany). Further, total genomic DNA of *P. verrucosa* leaf (second negative control) was extracted using Macherey Nagel Nucleo Spin Plant II kit. Each genomic DNA was extracted strictly following the manufacturer's guidelines.

For 16S rRNA analysis of bacterial endophytic communities, the genomic DNA was then subjected to PCR amplification using the primers 27f and 1492r.[2] The PCR amplification was performed in a 50 µL reaction mixture containing 45 µL Red Taq DNA Polymerase Master Mix (1.1x), 0.5 µL forward primer (100 µM), 0.5 µL reverse primer (100 µM), 3 µL template DNA and 1 µL of sterile double-distilled water. The PCR cycling protocol consisted of an initial denaturation at 95$^{o}$C for 2 min, 30 cycles of denaturation, annealing and elongation at 95$^{o}$C for 30 s, 60$^{o}$C for 40 s and 72$^{o}$C for 30 s. This was followed by a final elongation step at 72$^{o}$C for 5 min. As a negative control, the template DNA was replaced by sterile double-distilled water. The PCR amplified products spanning around 1500 bp were checked by agarose gel electrophoresis. The PCR products were further purified using GFX$^{TM}$ PCR

DNA and Gel Band Purification kit (GE Healthcare Life Sciences, Germany) following manufacturer's instructions. The amplified products were then sequenced from both directions using primers 27f and 1492r at GATC Biotech (Cologne, Germany).

For the amplification of AHBA synthase gene, a pair of degenerate primers were used as described by Huitu *et al*. (2009).[3] This primer pair amplifies the two conserved amino acid regions of five AHBA synthase genes namely *asnF*, *asm24*, *mitA*, *napF* and *rifK*.[4-7] The PCR amplification was performed in a 50 µL reaction mixture containing 45 µL Red Taq DNA Polymerase Master Mix (1.1x), 0.5 µL (100 µM) forward primer (F1: 5'-AGAGGATCCTTCGAGCRSGAGTTCGC-3'), 0.5 µL (100 µM) reverse primer (R1: 5'-GCAGGATCCGGAMCATSGCCATGTAG-3'), where R= A/G, S= G/C, M= A/C, 3 µL template DNA and 1 µL of sterile double-distilled water. The PCR cycling protocol consisted of an initial denaturation at 95 $^{o}$C for 2 min, 35 cycles of denaturation, annealing and elongation at 95 $^{o}$C for 30 s, 59 $^{o}$C for 40 s and 72 $^{o}$C for 30 s. This was followed by a final elongation step at 72 $^{o}$C for 5 min. As a positive control, the template DNA of *A. pretiosum* subsp. *auranticum* (DSM 44131) was used. The template DNA of *E. coli* (DSM 682) and *P. verrucosa* leaf were used as negative controls. The PCR amplified products spanning around 755 bp were checked and confirmed by agarose gel electrophoresis. The PCR products were further extracted from the agarose gel and purified using GFX$^{TM}$ PCR DNA and Gel Band Purification kit (GE Healthcare Life Sciences, Germany) strictly following manufacturer's instructions. The amplified products were then sequenced from both directions using primers F1 and R1 at GATC Biotech (Cologne, Germany).

**Phylogenetic evaluation and detection of AHBA synthase gene in bacterial endophytic communities of *Putterlickia* plants.** For the phylogenetic evaluation, the sequences of each bacterial endophytic

community were aligned using the EMBOSS-Needle Pairwise Sequence Alignment. This created an optimal global end to end alignment of the sequences based on Needleman-Wunsch algorithm.[8,9] The pairwise sequences of all the communities were further subjected to multiple sequence alignment for deducing the phylogenetic relationships using the T-Coffee Multiple Alignment software.[10,11] The software processed the sequences following the ClustalW alignment format with an alignment length of 80. For the construction of the phylogenetic tree, Drawgram in PHYLIP 3.66 was used.[12,13] The alignments (Newick format) were used to construct a phenogram-like rooted tree diagram using centered ancestral nodes and taking branch lengths into consideration.

The sequences of 755 bp (approx.) agarose gels excised and purified PCR products of positive control (DSM 44131) and positive endophytic communities (PRR1, PRR2, PRL, PRS1, PRS2, PVR1 and PVR2) were aligned using the EMBOSS-Needle Pairwise Sequence Alignment based on Needleman-Wunsch algorithm.[8,9] The aligned sequences, thus obtained, were matched against the public nucleotide database using the Basic Local Alignment Search Tool (BLASTn) of the US National Centre for Biotechnology Information (NCBI) for the identification of the genes. The similarity of the community sequences with phylogenetically-related reference sequences were identified using the EMBL-European Nucleotide Archive (ENA). The coding sequences were further translated into protein sequences using BLASTx of UniProt Knowledgebase (UniProtKB). The maximum homology to the respective protein in each case was identified using the UniProtKB identifier. The sequences of all the products have been deposited at EMBL-Bank.

**References**


(1) Kusari, S.; Zuehlke, S.; Spiteller, M. *J. Nat. Prod.* **2009**, *72*, 2-7.

(2) Lane, D. J. (1991) In *Nucleic acid techniques in bacterial systematics*; Stackebrandt, E.; Goodfellow, M. Eds.; Wiley, New York, 1991; pp 115–175.

(3) Huitu, Z.; Linzhuan, W.; Aiming, L.; Guizhi, S.; Feng, H.; Qiuping, L.; Yuzhen, W.; Huanzhang, X.; Qunjie, G.; Yiguang, W. *J. Appl. Microbiol.* **2009**, *106*, 755-763.

(4) August, P. R.; Tang, L.; Yoon, Y. J.; Ning, S.; Mueller, R.; Yu, T. W.; Taylor, M.; Hoffmann, D.; Kim, C. G.; Hutchinson, C. R.; Floss, H. G. *Chem. Biol.* **1998**, *5*, 69-79.

(5) Mao, Y.; Varoglu, M.; Sherman, D. H. *Chem. Biol.* **1999**, *6*, 251-263.

(6) Chen, S.; von Bamberg, D.; Hale, V.; Breuer, M.; Hardt, B.; Mueller, R.; Floss, H. G.; Reynolds, K. A.; Leistner, E. *Eur. J. Biochem.* **1999**, *261*, 98-107.

(7) Yu, T.-W.; Bai, L.; Clade, D.; Hoffmann, D.; Toelzer, S.; Trinh, K. Q.; Xu, J.; Moss, S. J.; Leistner, E.; Floss, H. G. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 7968-7973.

(8) Needleman, S. B.; Wunsch, C. D. J. Mol. Biol. **1970**, *48*, 443–453.

(9) Rice, P.; Longden, I.; Bleasby, A. *Trends Genet.* **2000**, *16*, 276-277.

(10) Notredame, C.; Higgins, D. G.; Heringa, J. *J. Mol. Biol.* **2000**, *302*, 205-217.

(11) Tommaso, P. D.; Moretti, S.; Xenarios, I.; Orobitg, M.; Montanyola, A.; Chang, J.-M.; Taly, J.-F.; Notredame, C. Nucleic Acids Res. **2011**, *39*, W13-W17.

(12) Felsenstein, J. *Amer. Nat.* **1985**, *125*, 1-15.

(13) Felsenstein, J. PHYLIP (PHYLogeny Inference Package) version 3.66, Distributed by the author, Department of Genetics, University of Washington, Seattle, WA, 2006.

## II. FIGURES

```
T-COFFEE, Version_11.00.8cbe486 (2014-08-12 22:05:29 - Revision 8cbe486 - Build 477)
Cedric Notredame
CPU TIME:0 sec.
SCORE=890
*
 BAD  AVG  GOOD
*
PRR1    :  90
PRR2    :  39
PRL     :  88
PRS1    :  88
PRS2    :  88
PVR1    :  88
PVR2    :  88
PVL     :  88
PVS1    :  88
PVS2    :  92
cons    :  89

PRR1    GCTACCATGCAGTCGACGGTAGCACAGAGAGCTTGCT----CTCGGGTGACGAGTGGCGGACGGGTGAGTAATGT
PRR2    AAGT-GGTAGCGCCCTCCCGAGG-TT-AAAGCTACCTACTTCTTTTGNNCCCA---------------CTCCCAT
PRL     CCTA-NNTGCAGTCGAGCGATGGATTAAGAGCTTGCT----CTTATGAAGTTAGCGGCGGACGGGTGAGTAACAC
PRS1    TG-------CAGTCGAGCGATGGATTAAGAGCTTGCT----CTTATGAAGTTAGCGGCGGACGGGTGAGTAACAC
PRS2    CTAT-ANTGCAGTCGAGCGATGGATTGAGAGCTTGCT----CTCAAGAAGTTAGCGGCGGACGGGTGAGTAACAC
PVR1    CC--------AG-----------ATT---------------CNTA----------------CGGGAGG-CAGCA-
PVR2    CC--------AG-----------ATT---------------CNTA----------------CGGGAGG-CAGCA-
PVL     CTAT-ANTGCAGTCGAGCGATGGATTGAGAGCTTGCT----CTCAAGAAGTTAGCGGCGGACGGGTGAGTAACAC
PVS1    GC--------AGTCGAGCGATGGATTGAGAGCTTGCT----CTCAAGAAGTTAGCGGCGGACGGGTGAGTAACAC
PVS2    AC--------GGN-----------------------------------AG------------------GCAGCA-

cons    

PRR1    CTGGGAAA-CTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTC-------GCAA
PRR2    G-GTGTGACGGGC-G----GTGT-GTACAAGGCCCGGGAAC-GT---ATTCACCG--TAGCATTCTGATCTAC--
PRL     GTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAACCGCAT
PRS1    GTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAANNNCAT
PRS2    GTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAATATTTTGAACTGCAT
PVR1    GT--------------------------------------------------------------------------
PVR2    GT--------------------------------------------------------------------------
PVL     GTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAATATTTTGAACTGCAT
PVS1    GTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAACCGCAT
PVS2    GT--------------------------------------------------------------------------

cons    

PRR1    GA--CCAAA--GAGGGGGACCTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGTGGGGT
PRR2    GATT--------ACTAGCGATTCCGA--CT------T---CATGGAGTCGAGTTGCAGACT--------------
PRL     GGTTCGAAATTGAAAGGCGGCTTCGG--CTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGT
PRS1    GGTTCGAAATTGAAAGGCGGCTTCGG--CTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGT
PRS2    GGTTCGAAATTGAAAGGCGGCTTCGG--CTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGT
PVR1    ----------------------------------------------------------------------------
PVR2    ----------------------------------------------------------------------------
PVL     GGTTCGAAATTGAAAGGCGGCTTCGG--CTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGT
PVS1    GGTTCGAAATTGAAAGGCGGCTTCGG--CTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGT
PVS2    ----------------------------------------------------------------------------

cons    

PRR1    AATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAACTGAGACACGGTCCA
PRR2    -----CCAATCCGGACTACGACGCNC------TTTATGAGG---TCCGCTTGCTCTCG---CGAGT--TCGCT--
PRL     AACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCA
PRS1    AACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCA
PRS2    AACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCA
```

```
PVR1    ------------------------------------------------------------------------
PVR2    ------------------------------------------------------------------------
PVL     AACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCA
PVS1    AACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCA
PVS2    ------------------------------------------------------------------------

cons

PRR1    GACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGA
PRR2    -TCTCTT--TGTAT-----------------GCGCCATTGTAGCACGTGTGTA-GCCCTAC-TCGTAAGGGCCA
PRL     GACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGA
PRS1    GACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGA
PRS2    GACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGA
PVR1    ----------------------GGGGAATATTGCNNAATGGGNNCAAGCCTGATGCAGCCATGCCGGNNNTATGA
PVR2    ----------------------GGGGAATATTGCNNAATGGGNNCAAGCCTGATGCAGCCATGCCGGNNNTATGA
PVL     GACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGA
PVS1    GACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGA
PVS2    ----------------------AGGGANNTTNNCGCAATGGACGAANNTNTGACGGAGCAACGCCGCNTGAGTGA

cons

PRR1    AGAAGGCCTTNNGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGTGTTGTGGT-TAATA--ACCGCAGCAATTGA
PRR2    TGATGAC-TTG----ACGTCATCCCCACC--------------TTCCTCCAGTTTATCACTGGCAGTCTCCTTTG
PRL     TGAAGGC-TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA
PRS1    TGAAGGC-TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA
PRS2    TGAAGGC-TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA
PVR1    AGAAGGCCTTNNGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGNNNNAAGGT-TAATA--ACCTCNNNNATTGA
PVR2    AGAAGGCCTTNNGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGNNNNAAGGT-TAATA--ACCTCNNNNATTGA
PVL     TGAAGGC-TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA
PVS1    TGAAGGC-TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA
PVS2    TGAAGGC-TTTCGGGTCGTAAAANTNTNTTGTTAGGGAAGAACAAGTGCTAGTTGAATA--AGCTGGCACCTTGA

cons

PRR1    CGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCG
PRR2    -AGTTCCCGGCCGG-------------ACCG-CTGGCAACAA---AGG--ATAA--GGGTTGCGCTCGTTG-CGG
PRL     CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG
PRS1    CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG
PRS2    CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG
PVR1    CGTTACTCGCAGAAGAAGCACCGGCTAANTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCG
PVR2    CGTTACTCGCAGAAGAAGCACCGGCTAANTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCG
PVL     CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG
PVS1    CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG
PVS2    CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCG

cons

PRR1    GAATTACTGGGCGTA-----------------------------------------------------AAGCGCACGC
PRR2    GACT---------TA-----------------------------------------------------A-CC-----C
PRL     GAATTATTGGGCGTA-----------------------------------------------------AAGCGCGCGC
PRS1    GAATTATTGGGCGTAAAGCGCGCGCAGGTGGTTTCTTAAGTCGTTATCCGGAATTATTGGGCGTAAAGCGCGCGC
PRS2    GAATTATTGGGCGTA-----------------------------------------------------AAGCGCGCGC
PVR1    GAATTACTGGGCGTA-----------------------------------------------------AAGCGCACGC
PVR2    GAATTACTGGGCGTA-----------------------------------------------------AAGCGCACGC
PVL     GAATTATTGGGCGTA-----------------------------------------------------AAGCGCGCGC
PVS1    GAATTATTGGGCGTA-----------------------------------------------------AAGCGCGCGC
PVS2    GAATTATTGGGCGTA-----------------------------------------------------AAGCGCGCGC

cons

PRR1    AGGCGGTCTGTCAAGTCGGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAGGCTAGAGT
PRR2    AA--CATTTCACAACAC---------GAGCTGACGA--CAGCCATGCAGCACCT---G--TCTCAGAG--TT----
PRL     AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT
PRS1    AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT
PRS2    AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT
```

```
PVR1    AGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTTGAAACTGGCAAGCTAGAGT
PVR2    AGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTTGAAACTGGCAAGCTAGAGT
PVL     AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT
PVS1    AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT
PVS2    AGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGT

cons    *    * *   **  *        * *    *   ** **  * *   *    *   **   *   *

PRR1    CTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGC
PRR2    CCCGAAG------GCACCAATCCATCTCTGG----AAAGTT-------CTCTGGATG-TCAAGAGTAGGTAAGGT
PRL     GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC
PRS1    GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC
PRS2    GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC
PVR1    CTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGC
PVR2    CTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGC
PVL     GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC
PVS1    GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC
PVS2    GCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGC

cons      * **      *    * ****  * * *     ** *         * **** *   *   ** *  ****

PRR1    GGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCANACAGGATTAGATACCCTGGTAGTCCA
PRR2    T-C-TTCGCG-TTGCATC-GA-ATTAAA---------CCACATGCTCC--ACCGC-TT----G---TGCGGGCCCC
PRL     GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCANACAGGATTAGATACCCTGGTAGTCCA
PRS1    GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
PRS2    GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCANACAGGATTAGATACCCTGGTAGTCCA
PVR1    GGCCCCCTGGACNAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
PVR2    GGCCCCCTGGACNAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
PVL     GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCANACAGGATTAGATACCCTGGTAGTCCA
PVS1    GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCANACAGGATTAGATACCCTGGTAGTCCA
PVS2    GACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA

cons      *   *  *       * **   * *          *    *   *  ** *  **        **   * **

PRR1    CGCCGTANACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCG-TGGCTTCCGGAGCTAACGCGTTAAGTCGACCG
PRR2    CGTCAATTCATTTGAGTTTTAACC--TTGCGGCCGTACTCCC-CAGGCGGTCGACTTAACGCGTTA-G--CTCCG
PRL     CGCCGTANACGATGAGTGCTANGTGTTAGAGGGTTTCCGCCCTTNAGTGCTGNAGTTAACGCANTAAGCACTCCG
PRS1    CGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCG
PRS2    CGCCGTANACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCG
PVR1    CGCCGTAAACGATGTCGACTTGGAGGTTGTNCCCTTGAGGNG-TGGCTTCCGGAGCTAACGCGTTAAGTCGACCG
PVR2    CGCCGTAAACGATGTCGACTTGGAGGTTGTNCCCTTGAGGNG-TGGCTTCCGGAGCTAACGCGTTAAGTCGACCG
PVL     CGCCGTANACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCG
PVS1    CGCCGTANACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCG
PVS2    CGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCG

cons    ** *        **    *      * *    *              *  ******  ** *    ***

PRR1    CCTGGGGAGTACGGCCGCAAG-GTTAAAACTCANATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PRR2    -----GAAGCCACGCCTCAAGGGCACAACCTCCAAG---TCGACAT-----CGTTTACGGCGTGGACTACCAGGG
PRL     CCGGGGGAGTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PRS1    CCTGGGGAGTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PRS2    CCTGGGGAGTACGGCCGCAAG-GCTGANACTCANAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PVR1    CCTGGGGAGTACGGCCGCAAG-GTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PVR2    CCTGGGGAGTACGGCCGCAAG-GTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PVL     CCTGGGGAGTACGGCCGCAAG-GCTGAAACTCANAGGAATTGACGGGGGCCCGCACAAGCGGTGGANCATGTGGT
PVS1    CCTGGGGAGTACGGCCGCAAG-GCTGAAACTCANAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT
PVS2    CCTGGGGAGTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT

cons        * **    *** ****  *   *  ***  *    * ***       **   * *  *****  *   **

PRR1    TTAATTCGATGCAACGCGAANAACCTTACCTACTCTTGACATCCA-GAGAACTTTCCANAGATGNNTTGGTGCCT
PRR2    T-A--TCTAATCC---TG--TTTGCTCCCCACGCTTTCGCACCTGAGCGTCAGTCTTTG-TCCAGGGGGCCGCCT
PRL     TTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGAAAACCCTAGAGATAGGGCTTCTCCT
PRS1    TTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGAAAACCCTAGAGATAGGGCTTCTCCT
PRS2    TTAATTCGAAACAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGAAAACCCTAGAGATAGGNCTTCTCCT
```

```
PVR1    TTAATTCGATGCAACGCGAAGAACCTTACCTACTCTTGACATCCA-GAGAACTTTCCAGAGATGGATTGGTGCCT
PVR2    TTAATTCGATGCAACGCGAAGAACCTTACCTACTCTTGACATCCA-GAGAACTTTCCAGAGATGGATTGGTGCCT
PVL     TTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGAAAACCCTAGAGATAGGGCTTCTCCT
PVS1    TTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGANAACCCTAGAGATAGGGCTTCTCCT
PVS2    TTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCT-CTGAAAACCCTAGAGATAGGGCTTCTCCT

cons    * *  ** *  *     *      **  **     **  ** *  *                      ***

PRR1    TC-GGGAACTCTGAG---ACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGC
PRR2    TC-GCCACCGGTATTCCTCCAGAT-CTCTACGCATTTCACC-GCT--ACACCTGGAATT-----CTACCCCCCTC
PRL     TC-GGGAGCAGAGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGC
PRS1    TC-GGGAGCAGAGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGC
PRS2    TCNGGGAGCANNGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGC
PVR1    TC-GGGAACTCTGAG---ACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGC
PVR2    TC-GGGAACTCTGAG---ACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGC
PVL     TC-GGGAGCANAGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCNN
PVS1    TC-GGGANCAGAGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTNGTGNNNTGTTGGGTTAAGTCCCGC
PVS2    TC-GGGAGCAGAGTG---ACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGC

cons    **  *  * *       *** *  *  * *  * **  * ***       **   *       **   ***

PRR1    AACGAG-CGCAACCCT-TATCCTTT-GTTGCCAGCGGTCCGGCCGGGAACTCANAGGAGACTGCCA--GTGATAA
PRR2    TACAAGACTCTAGCCTGCCAGTTTCGAATGCAGTTC-CCAGGTTGAGCCCGGGGATTTCACATCCGACTTGACAG
PRL     AACGAG-CGCAACCCT-TGATCTTA-GTTGCCATCA-TTAAGTTGGGCACTCTAAGGTGACTGCCG--GTGACAA
PRS1    AACGAG-CGCAACCCT-TGATCTTA-GTTGCCATCA-TTAAGTTGGGCACTCTAAGGTGACTGCCG--GTGACAA
PRS2    AACGAG-CGCAACCCT-TGATCTTA-GTTGCCATCA-TTAAGTTGGGCACTCTAAGGTGACTGCCG--GTGACAA
PVR1    AACGAG-CGCAACCCT-TATCCTTT-GTTGCCAGCGGTTCGGCCGGGAACTCAAAGGAGACTGCCA--GTGATAA
PVR2    AACGAG-CGCAACCCT-TATCCTTT-GTTGCCAGCGGTTCGGCCGGGAACTCAAAGGAGACTGCCA--GTGATAA
PVL     AACGAG-CGCAACCCT-TGATCTTA-GTTNNCATCA-TTAAGTTGGGCACTCTAAGGTGACTGCCG--GTGACAA
PVS1    AACGAG-CGCAACCNT-TGATCTTA-GTTGCCATCA-TTAANTTGGNNNCTCTAAGGTGACTGCCG--GTGACAA
PVS2    AACGAG-CGCAACCCT-TGATCTTA-GTTGCCATCA-TTAAGTTGGGCACTCTAAGGTGACTGCCG--GTGACAA

cons     ** **  * * * * *      **   *      *    *    *    **  **   *** *

PRR1    ACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGAGTAGGGCTACACACGTGCTACAATGGCGC
PRR2    ACCGCCTG--CGTGCGCTTTACGCCCAGTAAT---TCCGATTAACGCTTGCAC--CCTCCGTATTACCGCGGCTG
PRL     ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG
PRS1    ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG
PRS2    ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG
PVR1    ACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGAGTAGGGCTACACACGTGCTACAATGGCGC
PVR2    ACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGAGTAGGGCTACACACGTGCTACAATGGCGC
PVL     ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG
PVS1    ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG
PVS2    ACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG

cons    ** *   *   *** *  * *** * * * **     *  ***      *  *  *   ***  ***   **

PRR1    ATA-CAAAGAGAAGCGAACTCGCGAGAGCAAGCGGACCTCATAAAGNGCGTCGTAGTCCGGATTGGAGTCTGCAA
PRR2    CTGGCACGGAGTT---AG----CC-GGTGCTTCT--TCTGCGGGTAACGTCAATTGCTGCGGTTATTAACCACAA
PRL     GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA
PRS1    GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA
PRS2    GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA
PVR1    ATA-CAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGNNNGTCGTAGTCCGGATTGGAGTCTGCAA
PVR2    ATA-CAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGNNNGTCGTAGTCCGGATTGGAGTCTGCAA
PVL     GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA
PVS1    GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA
PVS2    GTA-CAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAA

cons     *  ** ***     *     *  *     *    **          *    * **     *  ***

PRR1    CTCGACTCCATGAAGTCGGAATCGCTAGTAATCGTAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTA
PRR2    CAC-CTTCCTCCCCGCTGAAAGTAC------------TTTACAAC-C----CNNAAGGCCTTC------TTCATA
PRL     CTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTA
PRS1    CTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTA
PRS2    CTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTA
```


**Figure S1.** The multiple sequence alignment of all the bacterial endophytic communities (PRR1, PRR2, PRL, PRS1, PRS2, PVR1, PVR2, PVL, PVS1, and PVS2) using T-Coffee, combined into one final alignment, based on their 16S rRNA regions.

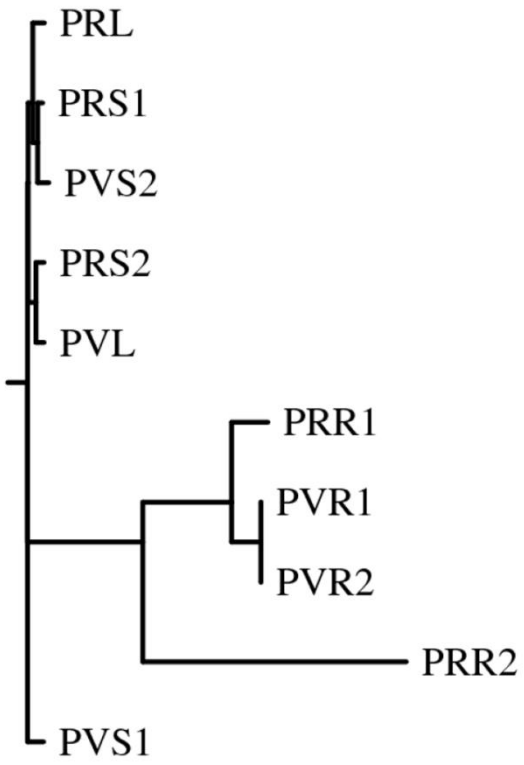


**Figure S2.** A phenogram-like rooted phylogenetic tree representing the correlation between the tested/evaluated endophytic communities on horizontal levels.

**III. TABLES**

**Table S1.** Overview of the AHBA synthase genes found in different endophytic communities isolated from different tissues of *Putterlickia* plants.

| **Bacterial endophytic community (Community code)** | **AHBA synthase genes* (bp)** | **Most closely related gene (EMBL accession number)** | **Maximum homology with EMBL (% identity)** | **Most closely related translated protein (UniProt identifier)** | **Maximum homology with UniProt identifier (% identity)** |
|---|---|---|---|---|---|
| ***Putterlickia retrospinosa*** **Primary root (PRR1)** | 755 | AAC13997.1 | 100 | Q44131 | 100 |
| ***Putterlickia retrospinosa*** **Secondary root (PRR2)** | 749 | AAC13997.1 | 99 | Q44131 | 99 |
| ***Putterlickia retrospinosa*** **Leaf (PRL)** | 754 | AAC13997.1 | 99 | Q44131 | 99 |
| ***Putterlickia retrospinosa*** **Thick stem (PRS1)** | 777 | AAC13997.1 | 98 | Q44131 | 98 |
| ***Putterlickia retrospinosa*** **Twigs (PRS2)** | 778 | AAC13997.1 | 99 | Q44131 | 99 |

| | | | | | |
|---|---|---|---|---|---|
| ***Putterlickia verrucosa*** **Primary root** **(PVR1)** | 772 | AAC13997.1 | 98 | Q44131 | 98 |
| ***Putterlickia verrucosa*** **Secondary root** **(PVR2)** | 756 | AAC13997.1 | 99 | Q44131 | 99 |

* All the sequences and products verified/discovered in the present study have been deposited at the EMBL-Bank.