

**Table S1: Plant samples investigated for leaf endophytes.** Closest relatives were determined using the EzBioCloud identification service (<https://www.ezbiocloud.net/identify>; accessed March 2022). Species marked with \* are likely misidentified. <sup>a</sup> Estimation for *F. homblei* made by shotgun sequencing of whole leaf tissue <sup>b</sup> Presence of *Burkholderia* s.l. symbiont unclear (see results). <sup>c</sup> based on only 808bp of 16S rRNA. Ratio bacteria/plant and *Burkholderiaceae*/total bacteria determined by blastn identification of a subset of 1 million reads against the NCBI nucleotide database. Endophyte isolation: +: endophyte isolated successfully; -: endophyte isolation failed; NA: endophyte isolation not attempted

| Species                     | Voucher       | Collection Location  | Investigated material                | Specialised endophyte structure | Ratio bacteria/plant | Ratio <i>Burkholderiaceae</i> /total bacteria | Endophyte detected | Endophyte isolation | Closest relative (16S rRNA; %ID)                            |
|-----------------------------|---------------|--|--------------------------------------|---------------------------------|----------------------|---|--------------------|---------------------|---|
| <i>Ardisia cornudentata</i> | GU 19862434   | Ghent University Botanical Garden, Belgium                         | Fresh leaves                         | Leaf galls                      | 14.30                | 95.88   | +                  | NA                  | <i>Caballeronia udeis</i> LMG 27134 (98.48%)                |
| <i>Ardisia mamillata</i>    | GU 19730151   | Ghent University Botanical Garden, Belgium                         | Fresh leaves                         | Leaf galls                      | 1.23                 | 86.01   | +                  | NA                  | <i>Caballeronia choica</i> LMG 22940 (98.83%)               |
| <i>Ardisia virens</i>       | ED 200420025  | Royal Botanic Garden Edinburgh, United Kingdom                     | Silica-dried leaves                  | Leaf galls                      | 0.62                 | 94.49   | +                  | NA                  | <i>Caballeronia udeis</i> LMG 27134 (98.62%)                |
| <i>Fadogia homblei</i>      | PRU 128010    | Roodeplaat, South Africa   | Bacterial isolate (from leaf tissue) | -                               | 2.84 <sup>a</sup>    | 99.15 <sup>a</sup>                            | +                  | +                   | <i>Paraburkholderia strydomiana</i> Wk1.1f (100%)           |
| <i>Pavetta capensis</i> *   | ED 19697408   | Royal Botanic Garden Edinburgh, United Kingdom                     | Silica-dried leaves                  | -                               | 0.09                 | 0.36  | -                  | NA                  | NA  |
| <i>Pavetta hochstetteri</i> | ED 19671929   | Royal Botanic Garden Edinburgh, United Kingdom                     | Silica-dried leaves                  | Leaf galls                      | 18.60                | 97.70   | +                  | NA                  | <i>Caballeronia pterochthonis</i> LMG 29326 (97.52%)        |
| <i>Pavetta indica</i> *     | BR 2014024336 | Meise Botanic Garden, Belgium                                      | Silica-dried leaves                  | -                               | 0.39                 | 0.75  | -                  | NA                  | NA  |
| <i>Pavetta revoluta</i>     | NA            | The Manie van der Schijff Botanical Garden, Pretoria, South Africa | Fresh leaves                         | Leaf galls                      | 0.04                 | 69.28   | (+) <sup>b</sup>   | -                   | <i>Caballeronia calidae</i> LMG 29321 (98.89%) <sup>c</sup> |
| <i>Pavetta ternifolia</i>   | BR 2002160529 | Meise Botanic Garden, Belgium                                      | Silica-dried leaves                  | Leaf galls                      | 0.07                 | 0.70  | -                  | NA                  | NA  |
| <i>Psychotria capensis</i>  | NA            | The Manie van der Schijff Botanical Garden,                        | Fresh leaves                         | Leaf galls                      | 0.06                 | 1.82  | -                  | -                   | NA  |

|                                   |               |  |                     |   |      |       |   |   |  |
|-----------------------------------|---------------|--|---------------------|---|------|-------|---|---|--|
|                                   |               | Pretoria, South Africa   |                     |   |      |       |   |   |  |
| <i>Psychotria zombamontana</i>    | NA            | The Manie van der Schijff<br>Botanical Garden,<br>Pretoria, South Africa | Fresh leaves        | - | 0.00 | 3.92  | - | - | NA   |
| <i>Vangueria dryadum</i>          | PRU 128005    | Lowveld National Botanic<br>Gardens, South Africa                        | Fresh leaves        | - | 1.07 | 97.13 | + | - | <i>Paraburkholderia insulsa</i> PNG-April (97.65%)     |
| <i>Vangueria esculenta</i>        | BR 2012116870 | Meise Botanic Garden,<br>Belgium   | Silica-dried leaves | - | 4.60 | 98.58 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.59%) |
| <i>Vangueria infausta</i>         | PRU 126086    | Voortrekker Monument,<br>Pretoria, South Africa                          | Fresh leaves        | - | 0.06 | 85.15 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.66%) |
| <i>Vangueria infausta</i>         | PRU 128003    | Lowveld National Botanic<br>Gardens, South Africa                        | Fresh leaves        | - | 9.85 | 98.76 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.66%) |
| <i>Vangueria macrocalyx</i>       | PRU 128006    | Lowveld National Botanic<br>Gardens, South Africa                        | Fresh leaves        | - | 1.67 | 97.78 | + | - | <i>Paraburkholderia insulsa</i> PNG-April (97.65%)     |
| <i>Vangueria madagascariensis</i> | PRU 128004    | Lowveld National Botanic<br>Gardens, South Africa                        | Fresh leaves        | - | 2.39 | 99.30 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.64%) |
| <i>Vangueria madagascariensis</i> | ED 19715346   | Royal Botanic Garden<br>Edinburgh, United Kingdom                        | Silica-dried leaves | - | 2.73 | 98.40 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.66%) |
| <i>Vangueria madagascariensis</i> | BR 19584419   | Meise Botanic Garden,<br>Belgium   | Silica-dried leaves | - | 3.29 | 98.11 | + | - | <i>Paraburkholderia phenoliruptrix</i> AC1100 (99.66%) |
| <i>Vangueria pygmaea</i>          | PRU 126008    | Cullinan, South Africa   | Fresh leaves        | - | 0.57 | 99.39 | + | - | <i>Paraburkholderia strydomiana</i> Wk1.1f (100%)      |
| <i>Vangueria pygmaea</i>          | PRU 126088    | Cullinan, South Africa   | Fresh leaves        | - | 1.63 | 96.97 | + | - | <i>Paraburkholderia</i>                                |

|  |            |   |              |   |      |       |   |   |   |
|--|------------|---|--------------|---|------|-------|---|---|---|
|  |            |   |              |   |      |       |   |   | <i>strydomiana</i><br>Wk1.1f (100%)                                 |
| <b><i>Vangueria randii</i></b>                       | PRU 128008 | Lowveld National Botanic<br>Gardens, South Africa | Fresh leaves | - | 1.74 | 98.24 | + | - | <i>Paraburkholderia</i><br><i>phenoliruptrix</i> AC1100<br>(99.52%) |
| <b><i>Vangueria</i><br/><i>soutpansbergensis</i></b> | PRU 128002 | Lowveld National Botanic<br>Gardens, South Africa | Fresh leaves | - | 2.73 | 99.35 | + | - | <i>Paraburkholderia</i><br><i>phenoliruptrix</i> AC1100<br>(99.45%) |

**Table S2: Sample details and accessions numbers for re-assembled and reference genomes.** Samples marked with \* are references that were not re-assembled.

| Species                           | Voucher       | Collection Location  | Investigated material                               | Specialised endophyte structure | Reference                                      | NCBI Accession           |
|-----------------------------------|---------------|--|---|---------------------------------|--|--------------------------|
| <i>Ardisia crenata</i>            | BR 19073685   | Meise Botanic Garden, Belgium  | Dissected leaf galls from fresh leaves              | Leaf galls                      | Carlier et al., 2015                           | PRJNA253365              |
| <i>Fadogia homblei</i>            | NA            | Meise Botanic Garden, Belgium<br>(Grown from seeds originally collected in South Africa) | Bacterial isolate R-49542<br>(isolated from leaves) | -                               | Verstraete et al., 2014                        | Not previously assembled |
| <i>Pavetta schumanniana</i>       | BR 2000194257 | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253363              |
| <i>Psychotria brachyanthoides</i> | BR 2009844596 | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253362              |
| <i>Psychotria humilis</i>         | BR 2009135940 | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253360              |
| <i>Psychotria pumila</i>          | BR 2004143571 | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253357              |
| <i>Psychotria umbellata</i>       | BR 2007130262 | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253361              |
| <i>Psychotria verschuerenii</i>   | BR 19750204   | Meise Botanic Garden, Belgium  | Dissected leaf galls from silica-dried leaves       | Leaf galls                      | Pinto-Carbó et al., 2016                       | PRJNA253359              |
| <i>Psychotria kirkii</i> *        | NA            | University of Zurich Botanic Garden, Switzerland   | Dissected leaf galls from fresh leaves              | Leaf galls                      | Carlier & Eberl, 2012;<br>Carlier et al., 2015 | GCF_000234195.1          |
| <i>Psychotria punctata</i> *      | NA            | University of Zurich Botanic Garden, Switzerland   | Dissected leaf galls from fresh leaves              | Leaf galls                      | Pinto-Carbó et al., 2016                       | GCF_001189345.1          |

**Table S3: *Burkholderia*, *Caballeronia*, and *Paraburkholderia* reference genomes used for comparative analysis.**

| Species                               | NCBI Accession | Species                                 | NCBI Accession |
|---------------------------------------|----------------|---|----------------|
| <i>Burkholderia thailandensis</i>     | GCF_000012365  | <i>Caballeronia concitans</i>           | GCF_0015444615 |
| <i>Burkholderia lata</i>              | GCF_000012945  | <i>Caballeronia arvi</i>                | GCF_0015444695 |
| <i>Burkholderia pseudomallei</i>      | GCF_000756125  | <i>Caballeronia catudaia</i>            | GCF_0015444755 |
| <i>Burkholderia plantarii</i>         | GCF_000835205  | <i>Caballeronia temeraria</i>           | GCF_0015444795 |
| <i>Burkholderia oklahomensis</i>      | GCF_000959365  | <i>Caballeronia fortuita</i>            | GCF_0015444835 |
| <i>Burkholderia dolosa</i>            | GCF_000959505  | <i>Caballeronia hypogeia</i>            | GCF_0015444875 |
| <i>Burkholderia multivorans</i>       | GCF_000959525  | <i>Caballeronia pedi</i>                | GCF_0015444915 |
| <i>Burkholderia pyrrocinia</i>        | GCF_001028665  | <i>Caballeronia glebae</i>              | GCF_001545035  |
| <i>Burkholderia humptydooensis</i>    | GCF_001462435  | <i>Caballeronia ptereochthonis</i>      | GCF_001545075  |
| <i>Burkholderia singularis</i>        | GCF_001523725  | <i>Caballeronia calidae</i>             | GCF_900044055  |
| <i>Burkholderia vietnamiensis</i>     | GCF_001523785  | <i>Caballeronia arationis</i>           | GCF_900230245  |
| <i>Burkholderia territorii</i>        | GCF_001527205  | <i>Caballeronia novacaledonica</i>      | GCF_900258035  |
| <i>Burkholderia seminalis</i>         | GCF_001718535  | <i>Caballeronia glathei</i>             | GCF_902833485  |
| <i>Burkholderia metallica</i>         | GCF_001718555  | <i>Caballeronia zhejiangensis</i>       | GCF_902833575  |
| <i>Burkholderia ubonensis</i>         | GCF_001718655  | <i>Paraburkholderia xenovorans</i>      | GCF_000013645  |
| <i>Burkholderia cenocepacia</i>       | GCF_001718895  | <i>Paraburkholderia phymatum</i>        | GCF_000020045  |
| <i>Burkholderia stagnalis</i>         | GCF_001718955  | <i>Paraburkholderia phytofirmans</i>    | GCF_000020125  |
| <i>Burkholderia mallei</i>            | GCF_002346025  | <i>Paraburkholderia kururiensis</i>     | GCF_000341045  |
| <i>Burkholderia reimsis</i>           | GCF_003294055  | <i>Paraburkholderia dilworthii</i>      | GCF_000472525  |
| <i>Burkholderia contaminans</i>       | GCF_004723625  | <i>Paraburkholderia mimosarum</i>       | GCF_000472825  |
| <i>Burkholderia cepacia</i>           | GCF_009586235  | <i>Paraburkholderia nodosa</i>          | GCF_000519185  |
| <i>Burkholderia glumae</i>            | GCF_009931375  | <i>Paraburkholderia acidipaludis</i>    | GCF_000684975  |
| <i>Burkholderia guangdongensis</i>    | GCF_013403875  | <i>Paraburkholderia bannensis</i>       | GCF_000685015  |
| <i>Burkholderia stabilis</i>          | GCF_900240005  | <i>Paraburkholderia ferrariae</i>       | GCF_000685035  |
| <i>Burkholderia gladioli</i>          | GCF_900608535  | <i>Paraburkholderia oxyphila</i>        | GCF_000685075  |
| <i>Burkholderia anthina</i>           | GCF_902498995  | <i>Paraburkholderia heleia</i>          | GCF_000739775  |
| <i>Burkholderia paludis</i>           | GCF_902499105  | <i>Paraburkholderia sacchari</i>        | GCF_000785435  |
| <i>Burkholderia arboris</i>           | GCF_902499125  | <i>Paraburkholderia monticola</i>       | GCF_001580545  |
| <i>Burkholderia aenigmatica</i>       | GCF_902499295  | <i>Paraburkholderia ginsengiterrae</i>  | GCF_001645135  |
| <i>Burkholderia ambifaria</i>         | GCF_902829835  | <i>Paraburkholderia sprentiae</i>       | GCF_001865575  |
| <i>Burkholderia diffusa</i>           | GCF_902830815  | <i>Paraburkholderia acidophila</i>      | GCF_002097715  |
| <i>Burkholderia latens</i>            | GCF_902832795  | <i>Paraburkholderia aromaticivorans</i> | GCF_002278075  |
| <i>Burkholderia pseudomultivorans</i> | GCF_902832925  | <i>Paraburkholderia fungorum</i>        | GCF_002891075  |
| <i>Burkholderia puraquae</i>          | GCF_902859845  | <i>Paraburkholderia terrae</i>          | GCF_002902925  |
| <i>Caballeronia insecticola</i>       | GCF_000402035  | <i>Paraburkholderia insulsa</i>         | GCF_003002115  |
| <i>Caballeronia grimmiae</i>          | GCF_000698555  | <i>Paraburkholderia eburnea</i>         | GCF_003003375  |
| <i>Caballeronia mineralivorans</i>    | GCF_001028175  | <i>Paraburkholderia unamae</i>          | GCF_003096875  |
| <i>Caballeronia sordidicola</i>       | GCF_001544455  | <i>Paraburkholderia silvatlantica</i>   | GCF_003217075  |
| <i>Caballeronia humi</i>              | GCF_001544475  | <i>Paraburkholderia bryophila</i>       | GCF_003269035  |
| <i>Caballeronia telluris</i>          | GCF_001544495  | <i>Paraburkholderia dokdonella</i>      | GCF_003286395  |
| <i>Caballeronia terrestris</i>        | GCF_001544515  | <i>Paraburkholderia graminis</i>        | GCF_003330785  |
| <i>Caballeronia choica</i>            | GCF_001544535  | <i>Paraburkholderia terricola</i>       | GCF_003330825  |
| <i>Caballeronia udeis</i>             | GCF_001544555  | <i>Paraburkholderia lacunae</i>         | GCF_003353175  |
| <i>Caballeronia cordobensis</i>       | GCF_001544575  | <i>Paraburkholderia caffeinilytica</i>  | GCF_003368325  |
|                                       |                | <i>Paraburkholderia phosphatilytica</i> | GCF_003443895  |
|                                       |                | <i>Paraburkholderia dinghuensis</i>     | GCF_003837865  |
|                                       |                | <i>Paraburkholderia quartelaensis</i>   | GCF_004353905  |
|                                       |                | <i>Paraburkholderia rhizosphaerae</i>   | GCF_004366595  |

|   |               |
|---|---------------|
| <i>Paraburkholderia dipogonis</i>         | GCF_004402975 |
| <i>Paraburkholderia azotifigens</i>       | GCF_007995085 |
| <i>Paraburkholderia panacisoli</i>        | GCF_008369935 |
| <i>Paraburkholderia franconis</i>         | GCF_009362735 |
| <i>Paraburkholderia bonniea</i>           | GCF_009455625 |
| <i>Paraburkholderia agricolaris</i>       | GCF_009455635 |
| <i>Paraburkholderia hayleyella</i>        | GCF_009455685 |
| <i>Paraburkholderia madseniana</i>        | GCF_009690905 |
| <i>Paraburkholderia acidiphila</i>        | GCF_009789655 |
| <i>Paraburkholderia acidisoli</i>         | GCF_009789675 |
| <i>Paraburkholderia youngii</i>           | GCF_013366925 |
| <i>Paraburkholderia caribensis</i>        | GCF_013378095 |
| <i>Paraburkholderia tropica</i>           | GCF_014171495 |
| <i>Paraburkholderia atlantica</i>         | GCF_014200895 |
| <i>Paraburkholderia ginsengisoli</i>      | GCF_016128195 |
| <i>Paraburkholderia caledonica</i>        | GCF_902833635 |
| <i>Paraburkholderia piptadeniae</i>       | GCF_900007165 |
| <i>Paraburkholderia ribeironis</i>        | GCF_900019265 |
| <i>Paraburkholderia lycopersici</i>       | GCF_900096975 |
| <i>Paraburkholderia phenazinium</i>       | GCF_900100735 |
| <i>Paraburkholderia tuberum</i>           | GCF_900101795 |
| <i>Paraburkholderia caballeronis</i>      | GCF_900104845 |
| <i>Paraburkholderia sartisoli</i>         | GCF_900107685 |
| <i>Paraburkholderia diazotrophica</i>     | GCF_900108945 |
| <i>Paraburkholderia megapolitana</i>      | GCF_900113825 |
| <i>Paraburkholderia aspalathi</i>         | GCF_900116445 |
| <i>Paraburkholderia hospita</i>           | GCF_900167965 |
| <i>Paraburkholderia susongensis</i>       | GCF_900177725 |
| <i>Paraburkholderia rhynchosiae</i>       | GCF_902859775 |
| <i>Paraburkholderia sediminicola</i>      | GCF_902859805 |
| <i>Paraburkholderia phenoliruptrix</i>    | GCF_902859825 |
| <i>Paraburkholderia humisilvae</i>        | GCF_902859855 |
| <i>Paraburkholderia solisilvae</i>        | GCF_902859875 |
| <i>Paraburkholderia ultramafica</i>       | GCF_902859915 |
| <i>Paraburkholderia fynbosensis</i>       | GCF_902859935 |
| <i>Paraburkholderia coffeinotolerans</i>  | GCF_902859945 |
| <i>Paraburkholderia kirstenboschensis</i> | GCF_904848585 |
| <i>Paraburkholderia metrosideri</i>       | GCF_904848625 |
| <i>Paraburkholderia sabiae</i>            | GCF_904848645 |
| <i>Paraburkholderia hiikae</i>            | GCF_904848665 |
| <i>Paraburkholderia strydomiana</i>       | GCF_004334935 |

**Table S4: Average nucleotide identities between leaf endophyte genomes.** ANI values are based on blastn alignments. Abbreviations: Acor: *Ca. C. ardisicola* Acor; Acre: *Ca. B. crenata* UZHbot9; Amam: *Ca. C. ardisicola* Amam; AvirE: *Ca. B. crenata* Avir; FhomR: *P. caledonica* R-49542; FhomSA: *P. caledonica* R-82532; Pbra: *Ca. B. brachyanthoides* UZHbot7; Phoc: *Ca. C. hochstetteri* PhocE; Phum: *Ca. B. humilis* UZHbot5; Pkir: *Ca. B. kirkii* UZHbot1; Ppum: *Ca. B. pumila* UZHbot3; Ppun: *Ca. B. kirkii* UZHbot2; Psch: *Ca. B. schumannianae* UZHbot8; Pumb: *Ca. B. calva* UZHbot6; Pver: *Ca. B. verschuerenii* UZHbot4; Vdry: *Ca. P. dryadicola* Vdry; Vesc: *P. phenoliruptrix* Vesc; Vinf: *P. phenoliruptrix* Vinf; Vmac: *Ca. P. dryadicola* Vmac; Vmad: *P. phenoliruptrix* VmadSA; VmadE: *P. phenoliruptrix* VmadEBG; VmadM: *P. phenoliruptrix* VmadMBG; Vpyg08: *P. caledonica* Vpyg08; Vpyg88: *P. caledonica* Vpyg88; Vran: *P. phenoliruptrix* Vran; Vsou: *Ca. P. soutpansbergensis* Vsou.

|        | Acor | Acre | Amam | Avir | FhomR | FhomSA | Pbra | Phoc | Phum | Pkir | Ppum | Ppun | Psch | Pumb | Pver | Vdry | Vesc | Vinf | Vmac | Vmad | VmadE | VmadM | Vpyg08 | Vpyg88 | Vran | Vsou |
|--------|------|------|------|------|-------|--------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|--------|--------|------|------|
| Acor   | 1.00 | 0.94 | 0.96 | 0.94 | 0.75  | 0.75   | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.78 | 0.77 | 0.76 | 0.77 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76  | 0.76  | 0.75   | 0.75   | 0.76 | 0.76 |
| Acre   | 0.94 | 1.00 | 0.94 | 0.99 | 0.76  | 0.76   | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.77 | 0.78 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77  | 0.77  | 0.76   | 0.76   | 0.77 | 0.77 |
| Amam   | 0.96 | 0.94 | 1.00 | 0.94 | 0.75  | 0.75   | 0.77 | 0.77 | 0.76 | 0.77 | 0.77 | 0.77 | 0.77 | 0.76 | 0.77 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76  | 0.76  | 0.75   | 0.75   | 0.76 | 0.76 |
| AvirE  | 0.94 | 0.99 | 0.94 | 1.00 | 0.76  | 0.76   | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.77 | 0.78 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77  | 0.77  | 0.76   | 0.76   | 0.77 | 0.77 |
| FhomR  | 0.76 | 0.77 | 0.75 | 0.77 | 1.00  | 0.98   | 0.77 | 0.78 | 0.76 | 0.77 | 0.77 | 0.77 | 0.78 | 0.77 | 0.76 | 0.82 | 0.86 | 0.86 | 0.82 | 0.86 | 0.86  | 0.86  | 0.98   | 0.98   | 0.86 | 0.86 |
| FhomSA | 0.76 | 0.77 | 0.75 | 0.77 | 0.98  | 1.00   | 0.77 | 0.78 | 0.76 | 0.77 | 0.77 | 0.77 | 0.78 | 0.77 | 0.76 | 0.82 | 0.86 | 0.86 | 0.82 | 0.86 | 0.86  | 0.86  | 0.98   | 0.98   | 0.86 | 0.85 |
| Pbra   | 0.77 | 0.78 | 0.77 | 0.78 | 0.76  | 0.77   | 1.00 | 0.83 | 0.82 | 0.83 | 0.82 | 0.83 | 0.83 | 0.91 | 0.86 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78  | 0.78  | 0.77   | 0.77   | 0.78 | 0.78 |
| Phoc   | 0.77 | 0.78 | 0.77 | 0.78 | 0.77  | 0.77   | 0.82 | 1.00 | 0.82 | 0.94 | 0.82 | 0.87 | 0.95 | 0.82 | 0.83 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78  | 0.78  | 0.77   | 0.77   | 0.78 | 0.78 |
| Phum   | 0.77 | 0.78 | 0.77 | 0.78 | 0.76  | 0.76   | 0.83 | 0.82 | 1.00 | 0.82 | 0.82 | 0.83 | 0.83 | 0.82 | 0.83 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77  | 0.77  | 0.76   | 0.76   | 0.77 | 0.77 |
| Pkir   | 0.77 | 0.78 | 0.77 | 0.78 | 0.76  | 0.76   | 0.82 | 0.94 | 0.82 | 1.00 | 0.82 | 0.87 | 0.96 | 0.82 | 0.82 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78  | 0.78  | 0.77   | 0.77   | 0.78 | 0.78 |
| Ppum   | 0.77 | 0.78 | 0.77 | 0.78 | 0.76  | 0.76   | 0.82 | 0.82 | 0.82 | 0.82 | 1.00 | 0.82 | 0.82 | 0.82 | 0.82 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77  | 0.77  | 0.76   | 0.76   | 0.77 | 0.77 |
| Ppun   | 0.77 | 0.78 | 0.77 | 0.78 | 0.77  | 0.77   | 0.83 | 0.87 | 0.82 | 0.87 | 0.82 | 1.00 | 0.88 | 0.82 | 0.83 | 0.78 | 0.79 | 0.79 | 0.78 | 0.79 | 0.79  | 0.79  | 0.77   | 0.77   | 0.79 | 0.79 |
| Psch   | 0.77 | 0.78 | 0.77 | 0.78 | 0.77  | 0.77   | 0.83 | 0.95 | 0.82 | 0.96 | 0.82 | 0.88 | 1.00 | 0.82 | 0.83 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78  | 0.78  | 0.77   | 0.77   | 0.78 | 0.78 |
| Pumb   | 0.76 | 0.78 | 0.76 | 0.78 | 0.76  | 0.76   | 0.91 | 0.82 | 0.82 | 0.82 | 0.82 | 0.83 | 0.82 | 1.00 | 0.85 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77  | 0.77  | 0.76   | 0.76   | 0.77 | 0.77 |
| Pver   | 0.77 | 0.78 | 0.77 | 0.78 | 0.76  | 0.76   | 0.86 | 0.83 | 0.83 | 0.83 | 0.82 | 0.83 | 0.83 | 0.86 | 1.00 | 0.77 | 0.78 | 0.78 | 0.77 | 0.78 | 0.78  | 0.78  | 0.76   | 0.76   | 0.78 | 0.78 |
| Vdry   | 0.76 | 0.77 | 0.76 | 0.77 | 0.82  | 0.82   | 0.78 | 0.78 | 0.77 | 0.78 | 0.77 | 0.78 | 0.78 | 0.77 | 0.77 | 1.00 | 0.83 | 0.83 | 1.00 | 0.83 | 0.83  | 0.83  | 0.82   | 0.82   | 0.83 | 0.83 |
| Vesc   | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.79 | 0.77 | 0.78 | 0.77 | 0.78 | 0.79 | 0.77 | 0.78 | 0.83 | 1.00 | 1.00 | 0.83 | 1.00 | 0.99  | 0.99  | 0.85   | 0.85   | 1.00 | 0.95 |
| Vinf   | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.79 | 0.77 | 0.78 | 0.77 | 0.78 | 0.79 | 0.77 | 0.77 | 0.83 | 1.00 | 1.00 | 0.83 | 1.00 | 0.99  | 0.99  | 0.85   | 0.85   | 1.00 | 0.95 |
| Vmac   | 0.76 | 0.77 | 0.76 | 0.77 | 0.82  | 0.82   | 0.78 | 0.78 | 0.77 | 0.78 | 0.77 | 0.78 | 0.78 | 0.77 | 0.77 | 1.00 | 0.83 | 0.83 | 1.00 | 0.83 | 0.83  | 0.83  | 0.82   | 0.82   | 0.83 | 0.83 |
| Vmad   | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.78 | 0.77 | 0.78 | 0.77 | 0.78 | 0.79 | 0.77 | 0.77 | 0.83 | 1.00 | 1.00 | 0.83 | 1.00 | 0.99  | 0.99  | 0.85   | 0.85   | 1.00 | 0.95 |
| VmadE  | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.79 | 0.77 | 0.78 | 0.77 | 0.79 | 0.79 | 0.78 | 0.78 | 0.83 | 0.99 | 0.99 | 0.83 | 0.99 | 1.00  | 1.00  | 0.85   | 0.85   | 0.99 | 0.95 |
| VmadM  | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.79 | 0.77 | 0.78 | 0.77 | 0.79 | 0.79 | 0.78 | 0.78 | 0.83 | 0.99 | 0.99 | 0.83 | 0.99 | 1.00  | 1.00  | 0.85   | 0.85   | 0.99 | 0.95 |
| Vpyg08 | 0.76 | 0.77 | 0.75 | 0.77 | 0.97  | 0.97   | 0.77 | 0.78 | 0.76 | 0.78 | 0.77 | 0.78 | 0.78 | 0.77 | 0.77 | 0.82 | 0.86 | 0.86 | 0.82 | 0.86 | 0.86  | 0.86  | 1.00   | 1.00   | 0.86 | 0.86 |
| Vpyg88 | 0.76 | 0.77 | 0.75 | 0.77 | 0.97  | 0.97   | 0.77 | 0.78 | 0.76 | 0.78 | 0.77 | 0.78 | 0.78 | 0.77 | 0.77 | 0.82 | 0.86 | 0.86 | 0.82 | 0.86 | 0.86  | 0.86  | 1.00   | 1.00   | 0.86 | 0.86 |
| Vran   | 0.76 | 0.77 | 0.76 | 0.78 | 0.85  | 0.85   | 0.78 | 0.79 | 0.77 | 0.78 | 0.78 | 0.78 | 0.79 | 0.77 | 0.77 | 0.83 | 1.00 | 1.00 | 0.83 | 1.00 | 0.99  | 0.99  | 0.85   | 0.85   | 1.00 | 0.95 |
| Vsou   | 0.76 | 0.77 | 0.76 | 0.77 | 0.85  | 0.85   | 0.78 | 0.78 | 0.77 | 0.78 | 0.78 | 0.79 | 0.78 | 0.78 | 0.77 | 0.83 | 0.95 | 0.95 | 0.83 | 0.95 | 0.95  | 0.95  | 0.85   | 0.85   | 0.95 | 1.00 |

**Table S5: Non-essential core genes of leaf endophytes.** Gene identifiers are from the genome of *Ca. Burkholderia kirkii* UZHbot1 (NCBI accession GCF\_000234195). Abbreviations: COG – Cluster of Orthologues Genes. COG Category meanings: D – Cell cycle control, cell division, chromosome partitioning; E – Amino acid transport and metabolism; F – Nucleotide metabolism and transport; G – Carbohydrate transport and metabolism; H – Coenzyme transport and metabolism; L – Replication, recombination and repair; M – Cell wall/membrane/envelope biogenesis; O – Post-translational modification, protein turnover, and chaperones; P – Inorganic ion transport and metabolism; Q – Secondary metabolite biosynthesis, transport and catabolism; S – Function unknown; T – Signal transduction mechanisms.

| Gene        | COG category | Functional prediction (EGGNOG)   | Best blastp hit in NCBI nr protein database  | Best blastp hit in UniPept/SwissProt  |
|-------------|--------------|--|--|---|
| CCD_35550.1 | H            | Removes the pyruvyl group from chorismate, with concomitant aromatization of the ring, to provide 4-hydroxybenzoate (4HB) for the ubiquinone pathway | Chorismate lyase ( <i>Caballeronia pterochtonis</i> )  | Probable chorismate pyruvate-lyase ( <i>Paraburkholderia xenovorans</i> )                       |
| CCD_35680.1 | M            | Membrane protein   | Colicin transporter ( <i>Caballeronia pterochtonis</i> )   | No hit  |
| CCD_36330.1 | M            | Membrane protein   | Phage holin family protein ( <i>Caballeronia novacaledonica</i> )                                  | Uncharacterized membrane protein YvID ( <i>Bacillus subtilis</i> )                              |
| CCD_36810.1 | Q            | Carboxymethylenebutenolidase   | Dienelactone hydrolase family protein ( <i>Caballeronia calidae</i> )                              | Putative carboxymethylenebutenolidase/Dienelactone hydrolase ( <i>Azospirillum brasilense</i> ) |
| CCD_37159.1 | M            | (Lipo)protein  | Outer membrane protein assembly factor BamC ( <i>Caballeronia catudaia</i> )                       | Outer membrane protein assembly factor BamC ( <i>Thiobacillus denitrificans</i> )               |
| CCD_37257.1 | D            | Cell division protein ZapD   | Cell division protein ZapD ( <i>Caballeronia</i> )   | Cell division protein ZapD ( <i>Burkholderia lata</i> )   |
| CCD_37658.1 | S            | Trm112 family protein  | Acyl-ACP desaturase ( <i>Caballeronia pedi</i> )   | No hit  |
| CCD_37721.1 | S            | Protein of unknown function (DUF2909)  | Twin transmembrane helix small protein ( <i>Burkholderiaceae</i> )                                 | No hit  |
| CCD_37723.1 | S            | Signal sequence binding sco1   | Cytochrome C oxidase subunit I ( <i>Caballeronia glebae</i> )                                      | No hit  |
| CCD_37888.1 | S            | Bacterial protein of unknown function (DUF883)   | DUF883 family protein ( <i>Caballeronia pterochtonis</i> )   | Uncharacterized protein YgjD ( <i>Escherichia coli</i> )  |
| CCD_37889.1 | S            | Membrane protein   | Phage holin family protein ( <i>Caballeronia pterochtonis</i> )                                    | No hit  |
| CCD_37890.1 | S            | Protein of unknown function (DUF3318)  | DUF3318 domain-containing protein ( <i>Caballeronia calidae</i> )                                  | No hit  |
| CCD_37930.1 | F            | Phosphoribosylaminoimidazolesuccinocarboxamide synthase purC   | Phosphoribosylaminoimidazolesuccinocarboxamide synthase ( <i>Caballeronia megalochromosomata</i> ) | Phosphoribosylaminoimidazole-succinocarboxamide synthase ( <i>Cupriavidus metalluridans</i> )   |
| CCD_37995.2 | O            | Peptide-methioine (S)-S-oxide reductase MsrA   | Peptide-methioine (S)-S-oxide reductase MsrA ( <i>Caballeronia calidae</i> )                       | Peptide methionine sulfoxide reductase MsrA ( <i>Ralstonia solanacearum</i> )                   |
| CCD_38181.1 | P            | Catalase activity  | Ferritin-like domain-containing protein ( <i>Caballeronia temeraria</i> )                          | No hit  |
| CCD_38266.1 | GM           | Nad-dependent epimerase dehydratase  | SDR family oxidoreductase ( <i>Caballeronia pterochtonis</i> )                                     | UDP-glucose 4-epimerase ( <i>Vibrio cholerae</i> )  |
| CCD_38823.1 | S            | STAS domain-containing protein   | anti-anti-sigma regulatory factor ( <i>Caballeronia jiangsuensis</i> )                             | No hit  |
| CCD_39033.1 | S            | Protein of unknown function (DUF2863)  | DUF2863 family protein ( <i>Caballeronia pterochtonis</i> )  | No hit  |
| CCD_39432.1 | L            | Involved in DNA repair and RecF pathway recombination  | DNA repair protein RecO ( <i>Caballeronia jiangsuensis</i> )                                       | DNA repair protein RecO ( <i>Paraburkholderia xenovorans</i> )                                  |



|             |   |  |  |  |
|-------------|---|--|--|--|
| CCD_39439.1 | T | Regulatory protein   | MucB/RseB C-terminal domain-containing protein<br>( <i>Caballeronia fortuita</i> ) | Sigma-E factor regulatory protein RseB<br>( <i>Haemophilus influenzae</i> )          |
| CCD_39635.1 | S | Sterol-binding domain protein  | Sterol-binding protein ( <i>Caballeronia peredens</i> )                            | No hit   |
| CCD_39933.1 | S | Hypothetical protein   | Regulator ( <i>Caballeronia pterochtonis</i> )                                     | No hit   |
| CCD_39967.1 | M | (Lipo)protein  | Outer membrane protein assembly factor BamC<br>( <i>Caballeronia catudaia</i> )    | Outer membrane protein assembly factor BamC<br>( <i>Thiobacillus denitrificans</i> ) |
| CCD_40265.1 | H | ATP-dependent carboxylate-amine ligase which exhibits weak glutamate--cysteine ligase activity | Carboxylate-amine ligase ( <i>Caballeronia turbans</i> )                           | Putative glutamate-cysteine ligase 2<br>( <i>Paraburkholderia phytofirmans</i> )     |
| CCD_40266.1 | P | Sodium:hydrogen antiporter   | Sodium/hydrogen exchanger ( <i>Caballeronia peredens</i> )                         | No hit   |
| CCD_40282.1 | E | Arginine/lysine/ornithine decarboxylase  | Arginine/lysine/ornithine decarboxylase<br>( <i>Caballeronia pterochtonis</i> )    | Lysine decarboxylase LdcA ( <i>Pseudomonas aeruginosa</i> )                          |
| CCD_40366.1 | O | Arginyl-tRNA-protein transferase   | Arginyl-tRNA-protein transferase ( <i>Caballeronia hypogeia</i> )                  | Aspartate/glutamate leucyltransferase<br>( <i>Paraburkholderia phymatum</i> )        |
| CCD_40735.1 | S | Fe-S cluster assembly protein IscX   | Fe-S cluster assembly protein IscX ( <i>Burkholderiaceae</i> )                     | Protein IscX ( <i>Haemophilus influenzae</i> )                                       |

**Table S6: Leaf endophyte core genes not conserved in *Burkholderia*, *Caballeronia*, and *Paraburkholderia* genomes.** Gene identifiers are from the genome of *Ca. Burkholderia kirkii* UZHbot1 (NCBI accession GCF\_000234195). Abbreviations: COG – Cluster of Orthologues Genes. COG Category meanings: C – Energy production and conversion; E – Amino acid transport and metabolism; F – Nucleotide metabolism and transport; G – Carbohydrate metabolism and transport; H – Coenzyme metabolism and transport; J – Translation, ribosomal structure and biogenesis; K – Transcription; L – Replication, recombination and repair; M – Cell wall/membrane/envelope biogenesis; O – Post-translational modification, protein turnover, and chaperones; P – Inorganic ion transport and metabolism; Q – Secondary metabolite biosynthesis, transport and catabolism; S – Function unknown; T – Signal transduction mechanisms; V – Defence mechanisms.

| Gene        | COG category | Functional prediction (EGGNOG)   | Best blastp hit in NCBI nr protein database  | Best blastp hit in UniPept/SwissProt  |
|-------------|--------------|--|--|---|
| CCD_35375.1 | T            | histidine kinase A domain protein  | Molecular chaperone DnaK ( <i>Caballeronia pterochtonis</i> )  | Chaperone protein DnaK ( <i>Paraburkholderia phymatum</i> )                                     |
| CCD_36748.1 | O            | Binds to Cpn60 in the presence of Mg-ATP and suppresses the ATPase activity of the latter  | Co-chaperone GroES ( <i>Burkholderiaceae</i> )   | 10kDa chaperonin/GroES/Cpn10 ( <i>Paraburkholderia phytofirmans</i> )                           |
| CCD_36810.1 | Q            | Carboxymethylenebutenolidase   | Dienelactone hydrolase family protein ( <i>Caballeronia calidae</i> )  | Putative carboxymethylenebutenolidase/Dienelactone hydrolase ( <i>Azospirillum brasilense</i> ) |
| CCD_36831.1 | H            | dihydroneopterin aldolase  | Dihydroneopterin aldolase ( <i>Caballeronia insecticola</i> )  | No hit  |
| CCD_37258.2 | S            | Protein conserved in bacteria  | Dephospho-CoA kinase ( <i>Caballeronia pterochtonis</i> )  | Dephospho-CoA kinase ( <i>Burkholderia lata</i> )   |
| CCD_37382.1 | V            | PFAM ABC transporter related   | ABC transporter ( <i>Caballeronia cordobensis</i> )  | Uncharacterized ABC transporter ATP-binding protein YadG ( <i>Escherichia coli</i> )            |
| CCD_37823.1 | C            | Belongs to the citrate synthase family   | Citrate (Si)-synthase ( <i>Caballeronia</i> )  | Citrate synthase ( <i>Bradyrhizobium diazoefficiens</i> )                                       |
| CCD_38121.1 | L            | TIGRFAM hydrolase, TatD family   | TatD family hydrolase ( <i>Caballeronia pterochtonis</i> )   | Uncharacterized metal-dependent hydrolase ( <i>Haemophilus influenzae</i> )                     |
| CCD_38703.1 | K            | Belongs to the ParB family   | ParB/RepB/SpoJ family partition protein ( <i>Caballeronia pedi</i> )   | Probable chromosome-partitioning protein ParB ( <i>Pseudomonas putida</i> )                     |
| CCD_38731.1 | M            | Belongs to the D-alanine–D-alanine ligase family   | D-alanine-D-alanine ligase ( <i>Caballeronia novaealedonica</i> )  | D-alanine-D-alanine ligase ( <i>Paraburkholderia xenovorans</i> )                               |
| CCD_39839.1 | L            | Histone-like DNA-binding protein which is capable of wrapping DNA to stabilize it, and thus to prevent its denaturation under extreme environmental conditions | HU family DNA-binding protein ( <i>Caballeronia zhejiangensis</i> )  | DNA-binding protein HU-beta ( <i>Pseudomonas aeruginosa</i> )                                   |
| CCD_40423.1 | H            | Methyltransferase type 11  | Bifunctional 2-polyprenyl-6-hydroxyphenol methylase/3-demethylubiquinol 3-O-methyltransferase UbiG ( <i>Caballeronia zhejiangensis</i> ) | Ubiquinone biosynthesis O-methyltransferase ( <i>Paraburkholderia phymatum</i> )                |
| CCD_40926.1 | EGP          | PFAM major facilitator superfamily MFS_1   | Lysophospholipid transporter LpIT ( <i>Caballeronia zhejiangensis</i> )  | Lysophospholipid transporter LpIT ( <i>Yersinia enterocolitica</i> )                            |
| CCD_41480.1 | FJ           | Catalyses the deamination of adenosine to inosine at the wobble position 34 of tRNA(Arg2)  | tRNA adenosine(34) deaminase TadA ( <i>Caballeronia insecticola</i> )  | tRNA-specific adenosine deaminase ( <i>Haemophilus influenzae</i> )                             |

**Table S7: Top 25 orthologues groups of genes enriched in endophytes compared to other members of *Burkholderia*, *Caballeronia*, and *Paraburkholderia*.** The representative gene of *Ca. Burkholderia kirkii* UZHbot1 was chosen as representative if possible, otherwise it was decided randomly. EEVS-cluster denotes if the gene is present in one of the two EEVS (2-*epi*-5-*epi*-valiolone synthase) gene clusters identified in *Ca. Burkholderia kirkii*. Closest relative represent the genus of the closest relative of that orthogroup by blastp searches against the RefSeq protein database (accessed June 2021). Endophyte/BCP genomes and proportion denotes the number and proportion of endophyte/BCP genomes present in the orthogroup. The difference is calculated as Endophyte proportion – BCP proportion. Abbreviations: EEVS - 2-*epi*-5-*epi*-valiolone synthase; BCP – *Burkholderia/Caballeronia/Paraburkholderia*; COG – Cluster of Orthologues Genes; COG category meaning: C – Energy production and conversion; E – Amino acid metabolism and transport; G – Carbohydrate metabolism and transport; H – Coenzyme metabolism and transport; J – Translation, ribosomal structure and biogenesis; K – Transcription; M – Cell wall/membrane/envelope biogenesis; O – Post-translational modifications, protein turnover, and chaperones; Q – Secondary metabolite metabolism, transport, and catabolism; S – Function unknown.

| Representative gene | EEVS-cluster | Closest relative (RefSeq) | Endophyte genomes | BCP genomes | Endophyte proportion | BCP proportion | Difference | COG | Functional annotation   |
|---------------------|--------------|---------------------------|-------------------|-------------|----------------------|----------------|------------|-----|---|
| CCD_36718.2         | X            | <i>Pseudomonas</i>        | 24                | 1           | 0.92                 | 0.01           | 0.92       | E   | 2- <i>epi</i> -5- <i>epi</i> -valiolone synthase  |
| CCD_39382.1         |              | <i>Xenorhabdus</i>        | 21                | 1           | 0.81                 | 0.01           | 0.80       | S   | Unknown   |
| CCD_39384.2         |              | <i>Xenorhabdus</i>        | 20                | 0           | 0.77                 | 0.00           | 0.77       | J   | Radical SAM superfamily   |
| CCD_36712.1         | X            | <i>Pseudomonas</i>        | 17                | 0           | 0.65                 | 0.00           | 0.65       | G   | Mannose-6-phosphate isomerase, cupin superfamily  |
| CCD_36715.1         | X            | <i>Pseudomonas</i>        | 17                | 0           | 0.65                 | 0.00           | 0.65       | G   | Glycoside Hydrolases Family 4; Likely 6-phospho-beta-glucosidase  |
| CCD_39395.1         | X            | <i>Noviherbaspirillum</i> | 15                | 0           | 0.58                 | 0.00           | 0.58       | G   | Trehalose-6-phosphate synthase  |
| CCD_39415.1         |              | <i>Burkholderia</i>       | 19                | 27          | 0.73                 | 0.20           | 0.53       | S   | Unknown   |
| CCD_39396.1         | X            | <i>Noviherbaspirillum</i> | 13                | 0           | 0.50                 | 0.00           | 0.50       | E   | Medium chain reductase/dehydrogenase (MDR)/zinc-dependent alcohol dehydrogenase-like family                         |
| CCD_39398.1         | X            | <i>Noviherbaspirillum</i> | 13                | 0           | 0.50                 | 0.00           | 0.50       | G   | Major Facilitator Superfamily Transporter   |
| PCALR49542_2702     |              | <i>Paraburkholderia</i>   | 13                | 4           | 0.50                 | 0.03           | 0.47       | K   | NA-binding transcriptional regulator, IclR family   |
| CCD_39397.1         | X            | <i>Noviherbaspirillum</i> | 12                | 0           | 0.46                 | 0.00           | 0.46       | S   | HAD-family hydrolase  |
| PPHERAN_6119        |              | <i>Pseudomonas</i>        | 12                | 2           | 0.46                 | 0.02           | 0.45       | C   | Pyruvate-formate lyase-activating enzyme  |
| PPHERAN_2366        |              | <i>Paraburkholderia</i>   | 13                | 8           | 0.50                 | 0.06           | 0.44       | Q   | Homospermidine synthase   |
| CCD_39196.1         |              | <i>Caballeronia</i>       | 22                | 56          | 0.85                 | 0.42           | 0.42       | H   | NAD-synthase  |
| CCD_39408.1         |              | <i>Nitrospira</i>         | 10                | 0           | 0.38                 | 0.00           | 0.38       | H   | N-Acyltransferase superfamily; possibly -acyl-L-homoserine lactone synthetase                                       |
| PCALR49542_2579     |              | No Hit                    | 10                | 0           | 0.38                 | 0.00           | 0.38       | S   | Unknown   |
| PPHERAN_5893        |              | <i>Breoghania</i>         | 10                | 0           | 0.38                 | 0.00           | 0.38       | E   | Amidinotransferase; possibly N-Dimethylarginine dimethylaminohydrolase or N-Dimethylarginine dimethylaminohydrolase |

|                             |                         |    |    |      |      |      |   |  |
|-----------------------------|-------------------------|----|----|------|------|------|---|--|
| <b>PPHERAN_58<br/>95</b>    | <i>Breoghania</i>       | 10 | 0  | 0.38 | 0.00 | 0.38 | J | Aspartyl-tRNA synthetase                       |
| <b>PPHERAN_61<br/>20</b>    | <i>Pseudomonas</i>      | 10 | 0  | 0.38 | 0.00 | 0.38 | S | Unknown  |
| <b>PCALR49542<br/>_6990</b> | <i>Paraburkholderia</i> | 11 | 8  | 0.42 | 0.06 | 0.36 | O | Glycosyltransferase Family 4 protein           |
| <b>PCALR49542<br/>_6995</b> | <i>Paraburkholderia</i> | 11 | 8  | 0.42 | 0.06 | 0.36 | M | SAM-dependent methyltransferase                |
| <b>CCD_35310.1</b>          | <i>Caballeronia</i>     | 23 | 70 | 0.88 | 0.53 | 0.35 | S | Bacterial protein of unknown function (DUF883) |
| <b>PPHERAN_54<br/>81</b>    | <i>Paraburkholderia</i> | 10 | 4  | 0.38 | 0.03 | 0.35 | M | RfaE bifunctional ADP-heptose synthase         |
| <b>PCALR49542<br/>_6993</b> | <i>Paraburkholderia</i> | 11 | 11 | 0.42 | 0.08 | 0.34 | S | GNAT family N-acetyltransferase                |
| <b>PPHERAN_14<br/>63</b>    | <i>Paraburkholderia</i> | 10 | 6  | 0.38 | 0.05 | 0.34 | S | Unknown  |

**Table S8: EEVS cluster organisation of other EEVS-clusters in endophyte genomes.** Genomes of the same host with the same cluster layout are merged. Genera in brackets represents the genus of the closest protein relative. \*A large region has three predicted genes in different frames that show homology with EEVS genes, possibly due to one or multiple frameshift mutations. Abbreviations: EEVS – 2-*epi*-5-*epi*-valiolone synthase; IS – Insertion element;  $\Psi$  – gene predicted to be a pseudogene.

|  |       |  |  |   |   |  |  |  |   |               |
|--|-------|--|--|---|---|--|--|--|---|---------------|
| <b><i>Ca. Burkholderia verschuerenii</i></b><br><b>UZHbot4</b> | IS5   | <b>EEVS</b>  | Group II intron  |   |   |  |  |  |   |               |
|  |       | (Pver_5505)<br>( <i>Pseudomonas</i> )              | reverse transcriptase/<br>maturase<br>( <i>Burkholderia</i> )                                | Gfo/Idh/MocA family<br>oxidoreductase<br>( <i>Pedobacter</i> )  | ATP-grasp<br>domain protein<br>( <i>Pseudomonas</i> )   | Inosamine-phosphate<br>amidinotransferase 1<br>( <i>Streptomyces</i> )   | Branched-chain<br>amino acid<br>aminotransferase<br>(Mixed origin)                     | 2OG-Fe(II)<br>oxygenase<br>( <i>Pseudomonas</i> )  | Argininosuccinate<br>synthase<br>( <i>Burkholderia</i> /<br><i>Salmonella</i> ) | Contig<br>end |
| <b><i>Ca. Caballeronia ardisicola</i></b><br><b>Acor</b>       | IS630 | <b>EEVS (<math>\Psi</math>)*</b>                   |  |   |   |  |  |  |   |               |
|  |       | (CBARDCOR_42<br>00)<br><i>Pseudomonas</i> )        | ROK family<br>( <i>Pseudomonas</i> )   | DegT/DnrJ/EryC1/StrS<br>family<br>aminotrasferase<br>( <i>Pseudomonas</i> )                           | 6-phospho-<br>beta-<br>glucosidase<br>( <i>Pseudomonas</i> )  | Gfo/Idh/MocA family<br>oxidoreductase<br>( <i>Pseudomonas</i> )  | 3-phosphoshikimate<br>1-<br>carboxylvinyltransfe<br>rase ( <i>Pseudomonas</i> )        | HAD family<br>hydrolase<br>( <i>Pseudomonas</i> )  | Class-I Dependent<br>methyltransferase<br>( $\Psi$ ; <i>Pseudomonas</i> )       | Contig<br>end |
| <b><i>Ca. Paraburkholderia dryadicola</i></b><br><b>Vdry</b>   | IS3   | <b>EEVS</b>  |  |   |   |  |  |  |   |               |
|  |       | Hypothetical<br>protein<br>( <i>Burkholderia</i> ) | (CPDRYDRY_65<br>70)<br>( <i>Burkholderia</i><br><i>singularis</i> /<br><i>Streptomyces</i> ) | SDR family<br>oxidoreductase<br>( <i>Burkholderia</i><br><i>singularis</i> /<br><i>Streptomyces</i> ) | GMC family<br>oxidoreductase<br>( <i>Burkholderia</i><br><i>singularis</i> /<br><i>Streptomyces</i> ) | DegT/DnrJ/EryC1/StrS<br>family aminotransferase<br>( <i>Burkholderia</i><br><i>singularis</i> /<br><i>Streptomyces</i> ) | GNAT-family N-<br>acetyltransferase<br>( <i>Burkholderia</i> /<br><i>Pseudomonas</i> ) | Carbamoyl<br>transferase<br>( <i>Pseudomonas</i> ) | Contig end  |               |

**Table S9. Mobile genetic element context of EEVS clusters in endophyte genomes.** The families of the IS elements as predicted by ISEscan using default settings are listed, together with their total abundance in the genome assembly in parenthesis. Left flanks are defined as the 3 kb upstream the first gene of the putative cluster, and right flanks as the 3 kb downstream of the last gene of the cluster. \*: Orphan EEVS outside of a corresponding cluster; nd: no IS element detected; contig end: the contig harbouring the EEVS clusters ends within 3 kb of the cluster boundaries.

| Strain                                | Cluster | Left flank    | Right flank   | Most abundant in genome assembly |
|---------------------------------------|---------|---------------|---------------|----------------------------------|
| <i>Ca. C. ardisicola</i> Acor         | S       | contig end    | IS21 (0.26%)  | IS21 (0.26%)                     |
| <i>Ca. C. ardisicola</i> Amam         | S       | IS6 (0.04%)   | IS6 (0.04%)   | IS21 (0.29%)                     |
| <i>Ca. B. brachyanthoides</i> UZHbot7 | K       | IS5 (1.67%)   | contig end    | IS5 (1.67%)                      |
| <i>Ca. B. calva</i> UZHbot6           | K       | IS5 (0.76%)   | IS256 (0.13%) | IS5 (0.76%)                      |
| <i>Ca. C. hochstetteri</i> PhocE      | K       | nd            | IS256 (0.04%) | IS21 (0.24%)                     |
| <i>Ca. C. hochstetteri</i> PhocE      | S       | nd            | IS66 (0.23%)  | IS21 (0.24%)                     |
| <i>Ca. B. humilis</i> UZHbot5         | S       | IS5 (0.31%)   | nd            | IS5 (0.31%)                      |
| <i>Ca. B. humilis</i> UZHbot5         | K       | contig end    | IS110 (0.21%) | IS5 (0.31%)                      |
| <i>Ca. B. kirkii</i> UZHbot1          | S       | IS630 (0.21%) | IS3 (2.41%)   | IS5 (4.84%)                      |
| <i>Ca. B. kirkii</i> UZHbot1          | K       | IS5 (4.84%)   | IS5 (4.84%)   | IS5 (4.84%)                      |
| <i>Ca. B. kirkii</i> UZHbot2          | K       | IS630 (2.92%) | IS630 (2.92%) | IS630 (2.92%)                    |
| <i>Ca. B. kirkii</i> UZHbot2          | S       | IS3 (2.33%)   | IS630 (2.92%) | IS630 (2.92%)                    |
| <i>Ca. B. pumila</i> UZHbot3          | K       | IS30 (0.47%)  | contig end    | IS630 (1.58%)                    |
| <i>Ca. B. schumanniana</i> UZHbot8    | K       | IS110 (0.34%) | contig end    | IS110 (0.34%)                    |
| <i>Ca. B. schumanniana</i> UZHbot8    | S       | contig end    | contig end    | IS110 (0.34%)                    |
| <i>Ca. B. verschuerenii</i> UZHbot4   | K       | IS5 (0.41%)   | contig end    | IS5 (0.41%)                      |
| <i>Ca. P. dryadicola</i> Vdry         | Other   | contig end    | nd            | IS66 (0.23%)                     |
| <i>Ca. P. soutpansbergensis</i> Vsou  | K*      | IS110 (0.14%) | IS21 (0.17%)  | IS5 (0.20%)                      |
| <i>P. caledonica</i> R-49542          | K*      | IS3 (0.21%)   | IS110 (0.24%) | IS110 (0.24%)                    |
| <i>P. caledonica</i> R-82532          | K*      | IS5 (0.12%)   | contig end    | IS110 (0.33%)                    |
| <i>P. caledonica</i> Vpyg08           | S       | IS3 (0.03%)   | IS21 (0.23%)  | IS21 (0.23%)                     |
| <i>P. caledonica</i> Vpyg88           | S       | IS3 (0.07%)   | IS21 (0.16%)  | IS5 (0.21%)                      |
| <i>P. dryadicola</i> Vmac             | Other   | nd            | contig end    | IS66 (0.25%)                     |
| <i>P. phenoliruptrix</i> Vesc         | S       | IS5 (0.31%)   | IS630 (0.06%) | IS5 (0.31%)                      |
| <i>P. phenoliruptrix</i> Vinf         | S       | IS5 (0.35%)   | IS630 (0.08%) | IS5 (0.35%)                      |
| <i>P. phenoliruptrix</i> VmadBGE      | S       | IS5 (0.30%)   | IS630 (0.05%) | IS5 (0.30%)                      |
| <i>P. phenoliruptrix</i> VmadMBG      | S       | IS5 (0.38%)   | IS5 (0.38%)   | IS5 (0.38%)                      |
| <i>P. phenoliruptrix</i> VmadSA       | S       | IS5 (0.40%)   | IS630 (0.07%) | IS5 (0.40%)                      |
| <i>P. phenoliruptrix</i> Vran         | S       | IS5 (0.47%)   | IS630 (0.06%) | IS5 (0.47%)                      |

