**SUPPLEMENTARY MATERIAL 3. Refinement of reference genomes**

**Table 1.** Comparison of the BZR1159 genome with the genomes of the closest strains of the genus *Brevundimonas* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***Brevundimonas* sp.** | **LPMIX5** | **GCF\_003576505.1** | **98.8249** | **565/567** |
| ***Brevundimonas* sp.** | **S1H14** | **GCF\_047716865.1** | **97.8421** | **564/567** |
| *B. naejangsanensis* | TWI1127 | GCF\_039673185.1 | 97.7386 | 552/567 |
| ***B. naejangsanensis*** | **DSM 23858** | **GCF\_000421705.1** | **97.7121** | **564/567** |
| *B. naejangsanensis* | TWI573 | GCF\_039674225.1 | 94.7526 | 554/567 |

**Table 2.** Comparison of the BZR162 genome with the genomes of the closest strains of the species *Ochrobactrum quorumnocens* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **A44** | **GCF\_002278035.1** | **99.2473** | **1347/1618** |
| RPTAtOch1 | GCF\_008728025.1 | 97.4112 | 1310/1618 |

**Table 3.** Comparison of the BZR201 genome with the genomes of the closest strains of the species *Leucobacter aridicollis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **JUb111** | **GCF\_024806925.1** | **98.7855** | **1004/1038** |
| L-9 | GCF\_016758215.1 | 97.1685 | 936/1038 |
| JCM 13321 | GCF\_039528735.1 | 97.1401 | 927/1038 |
| DSM 17380 | GCF\_013409595.1 | 97.1321 | 935/1038 |
| CTOTU49848 | GCA\_032109725.1 | 96.0258 | 305/1038 |
| J1 | GCF\_024399335.1 | 86.1995 | 900/1038 |

**Table 4.** Comparison of the BZR201P genome with the genomes of the closest strains of the genus *Paenochrobactrum* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***P. pullorum*** | **LMG 28095** | **GCF\_041929845.1** | **99.2062** | **1055/1085** |
| *P. gallinarii* | CCUG 57736 | GCF\_041929905.1 | 86.906 | 858/1085 |
| *P. glaciei* | JCM 15115 | GCF\_039522165.1 | 86.8295 | 886/1085 |
| *P. gallinarii* | DSM 22336 | GCF\_014205685.1 | 86.8163 | 864/1085 |
| *P. glaciei* | JCM 15115 | GCF\_041929825.1 | 86.813 | 890/1085 |
| *Paenochrobactrum* sp. | Lapin-020\_bin4 | GCA\_947627865.1 | 77.5785 | 183/1085 |

**Table 5.** Comparison of the BZR206 genome with the genomes of the closest strains of the species *Leucobacter aridicollis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **JUb111** | **GCF\_024806925.1** | **98.7489** | **1053/1102** |
| DSM 17380 | GCF\_013409595.1 | 96.9946 | 980/1102 |
| JCM 13321 | GCF\_039528735.1 | 96.9888 | 977/1102 |
| L-9 | GCF\_016758215.1 | 96.9622 | 985/1102 |
| CTOTU49848 | GCA\_032109725.1 | 95.7548 | 320/1102 |
| J1 | GCF\_024399335.1 | 86.234 | 907/1102 |

**Table 6.** Comparison of the BZR278 genome with the genomes of the closest strains of the genus *Alcaligenes* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***Alcaligenes* sp.** | **YSL9** | **GCF\_030515105.1** | **98.4113** | **1313/1391** |
| ***A. phenolicus*** | **Arc10** | **GCF\_041435035.1** | **98.3586** | **1296/1391** |
| ***Alcaligenes* sp.** | **ZN1L-C2** | **GCF\_042653705.1** | **98.3259** | **1293/1391** |
| *A. phenolicus* | JC896 | GCF\_040430305.1 | 98.3128 | 1285/1391 |
| *A. faecalis* | DY-8 | GCF\_024584725.1 | 98.3058 | 1311/1391 |
| *A. nematophilus* | S7 | GCF\_027270935.2 | 98.3018 | 1243/1391 |
| *A. nematophilus* | S8 | GCF\_027270985.2 | 98.3007 | 1222/1391 |
| *A. faecalis* subsp. *faecalis* | NCTC10388 | GCF\_900445215.1 | 98.2773 | 1309/1391 |
| *A. faecalis* | AU14 | GCF\_005311025.1 | 98.2723 | 1300/1391 |
| *A. faecalis* | PSM4 | GCF\_023702785.1 | 98.2706 | 1162/1391 |
| *A. faecalis* | D334 | GCF\_023101245.1 | 98.2697 | 1309/1391 |
| *A. faecalis* | c16 | GCF\_016807785.1 | 98.2563 | 1317/1391 |
| *A. faecalis* | PSM8 | GCF\_023702805.1 | 98.2401 | 1178/1391 |
| *A. nematophilus* | MBGG1\_DBT-NECAB | GCF\_041920605.1 | 98.2273 | 1291/1391 |
| *A. phenolicus* | CTOTU50702 | GCF\_032082505.1 | 98.2208 | 1304/1391 |
| *A. phenolicus* | MB207 | GCF\_002082085.1 | 98.214 | 1313/1391 |
| *A. faecalis* | T17 | GCF\_022343965.1 | 98.1997 | 1306/1391 |
| ***A. nematophilus*** | **S1** | **GCF\_026799675.2** | **98.1924** | **1319/1391** |
| *A. nematophilus* | S9 | GCF\_027270955.2 | 98.1889 | 1279/1391 |
| *A. phenolicus* | CTOTU48189 | GCF\_032002825.1 | 98.1746 | 1283/1391 |
| *A. nematophilus* | A-TC2 | GCF\_026344155.1 | 98.1514 | 1285/1391 |
| *A. nematophilus* | EBU-23 | GCF\_034258645.1 | 98.1425 | 1269/1391 |
| *Alcaligenes* sp. | 1735tsa3 | GCF\_024134565.1 | 98.0936 | 1304/1391 |
| *A. faecalis* | MOR02 | GCF\_000770015.1 | 98.0881 | 1295/1391 |

**Table 7.** Comparison of the BZR466 genome with the genomes of the closest strains of the species *Leucobacter aridicollis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **JUb111** | **GCF\_024806925.1** | **98.6881** | **1048/1101** |
| DSM 17380 | GCF\_013409595.1 | 96.8974 | 985/1101 |
| JCM 13321 | GCF\_039528735.1 | 96.8626 | 978/1101 |
| L-9 | GCF\_016758215.1 | 96.8544 | 988/1101 |
| CTOTU49848 | GCA\_032109725.1 | 95.5807 | 321/1101 |
| J1 | GCF\_024399335.1 | 86.1829 | 894/1101 |

**Table 8.** Comparison of the BZR585 genome with the genomes of the closest strains of the genus *Alcaligenes* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***A. faecalis*** | **D334** | **GCF\_023101245.1** | **98.3905** | **1308/1392** |
| ***A. phenolicus*** | **JC896** | **GCF\_040430305.1** | **98.3841** | **1304/1392** |
| *A. nematophilus* | S7 | GCF\_027270935.2 | 98.3207 | 1232/1392 |
| *Alcaligenes* sp. | ZN1L-C2 | GCF\_042653705.1 | 98.3192 | 1284/1392 |
| *A. phenolicus* | Arc10 | GCF\_041435035.1 | 98.3068 | 1285/1392 |
| *A. nematophilus* | S8 | GCF\_027270985.2 | 98.2815 | 1216/1392 |
| ***A. phenolicus*** | **MB207** | **GCF\_002082085.1** | **98.261** | **1321/1392** |
| *A. nematophilus* | S9 | GCF\_027270955.2 | 98.2508 | 1269/1392 |
| *A. faecalis* | PSM4 | GCF\_023702785.1 | 98.2487 | 1163/1392 |
| *A. phenolicus* | CTOTU50702 | GCF\_032082505.1 | 98.2387 | 1294/1392 |
| *A. faecalis* | AU14 | GCF\_005311025.1 | 98.2254 | 1286/1392 |
| *Alcaligenes* sp. | YSL9 | GCF\_030515105.1 | 98.2229 | 1311/1392 |
| *A. nematophilus* | A-TC2 | GCF\_026344155.1 | 98.2134 | 1292/1392 |
| *A. nematophilus* | S1 | GCF\_026799675.2 | 98.2084 | 1313/1392 |

**Table 9.** Comparison of the BZR588 genome with the genomes of the closest strains of the genus *Paenochrobactrum* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***P. pullorum*** | **LMG 28095** | **GCF\_041929845.1** | **99.1523** | **1055/1094** |
| *P. gallinarii* | CCUG 57736 | GCF\_041929905.1 | 86.9528 | 854/1094 |
| *P. gallinarii* | DSM 22336 | GCF\_014205685.1 | 86.8589 | 867/1094 |
| *P. glaciei* | JCM 15115 | GCF\_039522165.1 | 86.8217 | 900/1094 |
| *P. glaciei* | JCM 15115 | GCF\_041929825.1 | 86.777 | 903/1094 |
| *Paenochrobactrum* sp. | Lapin-020\_bin4 | GCA\_947627865.1 | 77.5461 | 189/1094 |

**Table 10.** Comparison of the BZR635 genome with the genomes of the closest strains of the genus *Leucobacter* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***Leucobacter\_*sp.** | **G161** | **GCF\_001482305.1** | **89.5646** | **206/218** |
| ***L. komagatae*** | **JCM 9414** | **GCF\_039534185.1** | **85.8974** | **216/218** |
| ***L. komagatae*** | **DSM 8803** | **GCF\_006716085.1** | **85.8833** | **218/218** |
| *L. komagatae* | VKM ST2845 | GCF\_000834055.1 | 85.6026 | 207/218 |

**Table 11.** Comparison of the BZR736 genome with the genomes of the closest strains of the species *Bacillus cereus* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **CH** | **GCF\_016774535.1** | **99.5063** | **1728/1765** |
| **A01** | **GCF\_034078685.1** | **99.5021** | **1731/1765** |
| **CMCC P0011** | **GCF\_001635955.1** | **99.502** | **1732/1765** |
| **CMCC P0021** | **GCF\_001635915.1** | **99.4907** | **1732/1765** |
| 24-1 | GCF\_009935575.1 | 99.4891 | 1714/1765 |
| 1 | GCF\_009935655.1 | 99.4791 | 1715/1765 |
| 2-1 | GCF\_009935735.1 | 99.4777 | 1713/1765 |
| 4 | GCF\_009935625.1 | 99.4704 | 1711/1765 |
| 3 | GCF\_009935775.1 | 99.4664 | 1717/1765 |

**Table 12.** Comparison of the BZR762 genome with the genomes of the closest strains of the genus *Alcaligenes* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| ***A. phenolicus*** | **Arc10** | **GCF\_041435035.1** | **98.5442** | **1167/1184** |
| ***Alcaligenes* sp.** | **YSL9** | **GCF\_030515105.1** | **98.5418** | **1169/**1184 |
| ***A. faecalis*** | **DY-8** | **GCF\_024584725.1** | **98.5209** | **1171/**1184 |
| *A. faecalis* | AU14 | GCF\_005311025.1 | 98.5149 | 1165/1184 |
| *A. faecalis* subsp. *faecalis* | NCTC10388 | GCF\_900445215.1 | 98.5137 | 1167/1184 |
| *A. nematophilus* | S9 | GCF\_027270955.2 | 98.5088 | 1132/1184 |
| *A. faecalis* | c16 | GCF\_016807785.1 | 98.5053 | 1166/1184 |
| *A. nematophilus* | S1 | GCF\_026799675.2 | 98.4986 | 1166/1184 |
| *A. phenolicus* | MB207 | GCF\_002082085.1 | 98.4759 | 1166/1184 |
| *A. faecalis* | D334 | GCF\_023101245.1 | 98.4538 | 1165/1184 |
| *Alcaligenes* sp. | 1735tsa3 | GCF\_024134565.1 | 98.3709 | 1167/1184 |
| *A. faecalis* | BDB4 | GCF\_002205415.1 | 98.1334 | 1169/1184 |

**Table 13.** Comparison of the BZR920 genome with the genomes of the closest strains of the species *Bacillus velezensis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **UCMB5113** | **GCF\_000455585.1** | **99.1964** | **1174/1192** |
| **At1** | **GCF\_006489275.1** | **99.189** | **1177/1192** |
| **X-Bio\_1** | **GCF\_014230015.1** | **99.183** | **1175/1192** |
| B-14320 | GCF\_035811875.1 | 99.1529 | 1173/1192 |
| KS04AU | GCF\_022488425.1 | 99.147 | 1180/1192 |
| N85 | GCF\_018916765.1 | 99.1431 | 1180/1192 |
| IT-133MI5 | GCF\_949768215.1 | 99.1419 | 1179/1192 |
| 5RB | GCF\_003568755.1 | 99.1392 | 1178/1192 |
| JJ1043 | GCF\_040712135.1 | 99.1358 | 1184/1192 |

**Table 14.** Comparison of the BZR926 genome with the genomes of the closest strains of the species *Achromobacter marplatensis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **LMG 26219** | **GCF\_902859635.1** | **90.4544** | **1579/1622** |
| **CCM 7608** | **GCF\_042432955.1** | **90.4806** | **1565/1622** |
| CECT 7342 | GCF\_003315095.1 | 90.4533 | 1572/1622 |
| B2 | GCF\_002209535.1 | 90.4346 | 1574/1622 |
| ERR9969197\_bin.41 MetaWRAP\_v1.3\_MAG\_genomic | GCF\_963527865.1 | 90.4182 | 1553/1622 |
| CIP110817T | GCA\_965137325.1 | 90.3882 | 1522/1622 |
| GD03676 | GCF\_029844665.1 | 90.3776 | 1516/1622 |
| ANA | GCF\_000286415.1 | 90.3706 | 1565/1622 |

**Table 15.** Comparison of the BZR936 genome with the genomes of the closest strains of the species *Bacillus velezensis* based on the analysis of average nucleotide identity calculated using fastani

The most similar genomes are highlighted in bold.

|  |  |  |  |
| --- | --- | --- | --- |
| Strain | Identifier | ANI (%) | Orthologous fragments (found/total) |
| **UCMB5113** | **GCF\_000455585.1** | **99.1911** | **1178/1195** |
| **At1** | **GCF\_006489275.1** | **99.1883** | **1179/1195** |
| **X-Bio\_1** | **GCF\_014230015.1** | **99.182** | **1176/1195** |
| KS04AU | GCF\_022488425.1 | 99.1466 | 1182/1195 |
| JJ1043 | GCF\_040712135.1 | 99.1413 | 1184/1195 |
| B-14320 | GCF\_035811875.1 | 99.1403 | 1175/1195 |