Draft Genome Sequences of Three Terrestrial Isoprene-Degrading Rhodococcus Strains

Andrew T. Crombie, Helen Emery, Terry J. McGenity, J. Colin Murrell

School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom; School of Biological Sciences, University of Essex, Wivenhoe Park, Colchester, Essex, United Kingdom

ABSTRACT Isoprene is produced in abundance by plants and constitutes a carbon source for microbes. The genomes of three isoprene degraders isolated from tree leaves or soil from the campus of the University of East Anglia were sequenced. These high-GC-content isolates are actinobacteria belonging to the genus Rhodococcus.

The emissions of isoprene to the atmosphere from terrestrial plants, principally trees, are similar in magnitude to those of methane (approximately 550 Tg per year). Some bacteria are capable of using isoprene as a sole source of carbon and energy, but their diversity and contribution to cycling of this climate-active trace gas have not been intensively studied until recently (1, 2). So far, genome sequences for a relatively small number of isoprene-degrading strains have been published (3–5).

Rhodococcus sp. strains ACPA1 and ACPA4 were isolated from the leaves of a white poplar tree (Populus alba) and Rhodococcus sp. strain ACS1 was isolated from soil in the vicinity of willow trees (Salix fragilis) located on the campus of the University of East Anglia, Norwich, United Kingdom. Isolates were grown in liquid culture supplied with isoprene, as described previously (3). Genomic DNA was extracted using a conventional phenol-chloroform method (3). For each strain, genomic DNA was sequenced by Edinburgh Genomics (Edinburgh, UK), following the construction of three libraries with inserts of 330, 550, and 4,500 bp, on an Illumina MiSeq instrument generating 300-nucleotide (nt) paired-end reads. Reads were trimmed using Cutadapt version 1.8.3 (6) using parameters -q 30 and -m 50, assembled using SPAdes version 3.7.0 (7) (removing contigs shorter than 200 bp), and annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The basic genome statistics are shown in Table 1.

The large genome sizes (7 to 11 Mbp) are typical of metabolically versatile rhodococci (8), although the genome of Rhodococcus sp. ACPA4 is significantly smaller and of lower GC content than those of the other two strains. Based on analysis of the 16S rRNA genes, Rhodococcus sp. strains ACPA1 and ACPA4 are most closely related to the isoprene degraders Rhodococcus opacus PD630 (9) and Rhodococcus sp. strain AD45 (3), respectively, and Rhodococcus sp. strain ACS1 is related most closely to a non-isoprene-degrading Rhodococcus koreensis strain (10). All three genomes contain high-similarity homologues (>80% amino acid identity) of the isoprene metabolic genes described in Rhodococcus sp. AD45 (3, 11), including those encoding the soluble diiron center isoprene monoxygenase (isoABCDEF), glutathione-S-transferase (isoI), dehydrogenase (isoH), and genes for enzymes predicted to perform subsequent metabolic steps (isoG and isoJ). As in other isoprene degraders, isoGHJ are duplicated nearby, while Rhodococcus sp. ACPA4 also contains a third copy of isoH and isoJ. The glutathione biosynthesis genes gshA and gshB are also present in all three strains, consistent with the observation that conjugation of isoprene epoxide with glutathione appears to be
TABLE 1  Genome statistics and accession numbers

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Genome size (Mbp)</th>
<th>Coverage (x)</th>
<th>No. of contigs</th>
<th>N\textsubscript{50} (Mbp)</th>
<th>G+C content (%)</th>
<th>No. of rRNA operons</th>
<th>No. of tRNAs</th>
<th>No. of CDS\textsuperscript{a}</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACPA1</td>
<td>10.06</td>
<td>238</td>
<td>47</td>
<td>1.38</td>
<td>66.9</td>
<td>1</td>
<td>68</td>
<td>9,193</td>
<td>NSDX00000000</td>
</tr>
<tr>
<td>ACPA4</td>
<td>7.07</td>
<td>296</td>
<td>9</td>
<td>5.07</td>
<td>61.6</td>
<td>3</td>
<td>55</td>
<td>6,473</td>
<td>NSDY00000000</td>
</tr>
<tr>
<td>ACS1</td>
<td>10.89</td>
<td>172</td>
<td>40</td>
<td>1.74</td>
<td>67.1</td>
<td>1</td>
<td>70</td>
<td>10,062</td>
<td>NSDZ00000000</td>
</tr>
</tbody>
</table>

\textsuperscript{a}CDSs, coding sequences.

universal among isoprene degraders, despite the uncommon usage of this small thiol in Gram-positive bacteria (12, 13).

Interestingly, the genomes of \textit{Rhodococcus} sp. strains ACPA1 and ACS1 contain an additional soluble diiron center monoxygenase in a different region of the genome, with high similarity (>90% amino acid identity) to propane monooxygenase from \textit{Gordonia} TY-5 (14), indicative of the ability of many isoprene-degrading strains to grow on short-chain alkanes in addition to isoprene (5, 15).

These genome sequences extend the diversity of known \textit{iso} genes and will enable the development of improved gene probes and molecular ecology methods for the detection of isoprene degraders in the environment.

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers shown in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work was funded by the Earth and Life Systems Alliance (ELSA) at the University of East Anglia, Natural Environment Research Council (NERC) grants to J.C.M. (NE/J009725/1) and T.J.M. (NE/J009555/1), and an ERC advanced grant to J.C.M. (694578 — IsoMet).

REFERENCES