Isolation of isoprene degrading bacteria from soils, development of *isoA* gene probes and identification of the active isoprene degrading soil community using DNA-stable isotope probing

Myriam El Khawand,[1] Andrew T. Crombie,[1] Antonia Johnston,[1] Dmitrii V. Vavlline,[2] Joseph C. McAuliffe,[2] Jacob A. Latone,[2] Yuliya A. Primak,[2] Sang-Kyu Lee,[2] Gregg M. Whited,[2] Terry J. McGenity[3] and J. Colin Murrell[1]

[1]University of East Anglia, Norwich Research Park, Norwich, UK; [2]DuPont Industrial Biosciences, 925 Page Mill Road, Palo Alto, CA 94304, USA; [3]University of Essex, Wivenhoe Park, Colchester, UK.

Classification:

Biological Sciences, Microbiology

Corresponding author:

Andrew T. Crombie

School of Environmental Science, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. Email: a.crombie@uea.ac.uk Telephone +44 (0)1603 592239

Keywords:

Isoprene, 2-methyl-1,3-butadiene, biodegradation, culture independent methods, molecular ecology, *Rhodococcus*, biogeochemistry, stable isotope probing, DNA-SIP, [U-^{13}C]-isoprene biosynthesis.

Originality-Significance Statement

Microorganisms with the ability to degrade the important climate-active gas isoprene are abundant in the environment yet we know very little about them. Here we applied cultivation-dependent and -independent methods (including the first example of the use of DNA-stable isotope probing with this substrate) to greatly expand our knowledge of isoprene degraders, which is a prerequisite to developing an understanding of isoprene biodegradation. We provide genome sequence data on two new isolates, solid evidence for isoprene degradation in genera not previously implicated and tools to investigate biodegradation of this important atmospheric trace gas.

Abstract

Emissions of biogenic volatile organic compounds (bVOCs), are an important element in the global carbon cycle, accounting for a significant proportion of fixed carbon. They contribute directly and indirectly to global warming and climate change and have a major effect on atmospheric chemistry. Plants emit isoprene to the atmosphere in similar quantities to emissions of methane from all sources and each accounts for approximately one third of total VOCs. Although methanotrophs, capable of growth on methane, have been intensively studied, we know little of isoprene biodegradation. Here we report the isolation of two isoprene-degrading strains from the terrestrial environment and describe the design and testing of PCR primers targeting *isoA*, the gene encoding the active-site component of the conserved isoprene monooxygenase, which are capable of retrieving *isoA* sequences from isoprene-enriched environmental samples. Stable isotope probing experiments, using biosynthesized ^{13}C-labelled isoprene, identified the active isoprene-degrading bacteria in soil. This study identifies novel isoprene-degrading strains using both culture-dependent and, for the first time, culture-independent methods and provides the tools and foundations for continued investigation of the biogeography and molecular ecology of isoprene-degrading bacteria.

Introduction

Isoprene accounts for approximately one third of the total flux of volatile organic compounds to the atmosphere, an amount that is similar to the methane flux (Guenther et al., 2012; Kirschke et al., 2013). In the atmosphere, isoprene is rapidly photochemically oxidized resulting in a short lifetime (of the order of hours) and consequent low concentrations. Attack by hydroxyl or nitrate radicals or ozone leads to a variety of products depending on temperature and pollutant (nitrogen oxides, NO_x) levels (Atkinson and Arey, 2003). Overall, isoprene has a significant effect on atmospheric chemistry and hence climate change, due both to the production of greenhouse gases (principally ozone) and by reducing the hydroxyl radical-mediated oxidizing capacity of the atmosphere, which increases the lifetime of methane (Pacifico et al., 2009). In addition, isoprene oxidation-products form secondary organic aerosols and cloud condensation nuclei, with implications for air quality and climate (Fiore et al., 2012).

About 600 Tg y^{-1} isoprene is emitted to the atmosphere by terrestrial plants, although not all plant species produce isoprene (Sharkey, 2013; Loreto and Fineschi, 2015). Isoprene synthesis occurs in the chloroplast, via the enzyme isoprene synthase, which converts dimethylallyl pyrophosphate to isoprene (Logan et al., 2000). Isoprene protects plants against heat stress by reducing heat-induced cell-membrane damage, enhances tolerance of reactive oxygen species and may affect plant-insect interactions (Loivamäki et al., 2008; Sharkey et al., 2008; Vickers et al., 2009; Sharkey, 2013). In the marine environment isoprene is released by phytoplankton and macroalgae (Broadgate et al., 2004; Exton et al., 2015). Some bacteria, including soil-dwelling species such as *Bacillus subtilis*, release isoprene, as do some fungi (Kuzma et al., 1995; Julsing et al., 2007; Bäck et al., 2010),

although we lack a clear understanding of why these organisms produce isoprene. A non-enzymatic reaction resulting in isoprene was reported in humans and other animals, associated with the mevalonate pathway of cholesterol synthesis (Gelmont et al., 1981).

Although atmospheric isoprene concentrations are low (<1 – 4 ppbv in one wide-ranging study (Greenberg et al., 1999)), in the vicinity of isoprene sources (around or below tree canopy level) concentrations are significantly higher. For example, Wiedinmyer and colleagues measured ground level isoprene concentrations of 11 and 36 ppbv at sites in Texas and Missouri, respectively (Wiedinmyer et al., 2001; Wiedinmyer et al., 2005). Soils can act as a biological sink for isoprene, at or below these concentrations. In field chambers set up in temperate forest soils, isoprene was rapidly depleted to below the 5 ppbv limit of detection (Cleveland and Yavitt, 1997; Cleveland and Yavitt, 1998). In continuous flow experiments conducted by Gray et al. (2015), soils supplied with isoprene at concentrations of 2 – 200 ppbv consumed isoprene at all concentrations, with a rate of 62 pmol g^{-1} h^{-1} at 20 ppbv. These data demonstrate both the potential of soils to consume isoprene released locally in soils and also to take up atmospheric isoprene, conclusions also reached in mesocosm experiments by Pegoraro et al. (2005). Several bacterial strains, tentatively assigned to the genera *Nocardia, Rhodococcus* (Actinobacteria) and *Alcaligenes* (Betaproteobacteria), were isolated from isoprene enrichment cultures and shown to grow on isoprene as sole growth substrate, (van Ginkel et al., 1987b; van Ginkel et al., 1987a; Ewers et al., 1990; Cleveland and Yavitt, 1997) and, more recently, *Pseudomonas, Alcaligenes* and *Klebsiella* isoprene-degrading strains were isolated from rubber-contaminated soil (Srivastva et al., 2015). Strains were also obtained from the marine environment, including representatives of Actinobacteria,

Bacteroidetes and Alpha- and Gammaproteobacteria (Acuña Alvarez et al., 2009). None of these terrestrial isolates was extensively characterized and the best documented isoprene degrader to date is *Rhodococcus* sp. AD45, a Gram-positive actinobacterium isolated nearly 20 years ago from freshwater sediment by the group of Dick Janssen (van Hylckama Vlieg *et al.,* 1998).

In *Rhodococcus* sp. AD45, isoprene is oxidized to epoxyisoprene (1,2-epoxy-2-methyl-3-butene) by a four-component soluble diiron centre monooxygenase (SDIMO) with homology to enzymes including the soluble methane monooxygenase (sMMO) and alkene/aromatic monooxygenases (van Hylckama Vlieg et al., 2000; Leahy et al., 2003). The epoxide is then conjugated with glutathione, catalyzed by glutathione-*S*-transferase (GST) (IsoI) and oxidized in two steps by a dehydrogenase (IsoH) resulting in 2-glutathionyl-2-methyl-butenoic acid (Fig. 1) (van Hylckama Vlieg et al., 1998; van Hylckama Vlieg et al., 1999). Interestingly, conjugation with glutathione in *Rhodococcus* sp. AD45 contrasts with other alkene utilizers, which often overcome the toxicity of epoxides by forming coenzyme M conjugates or by hydrolysis (Ensign, 2001; Kottegoda et al., 2015). The genes encoding the monooxygenase (*isoABCDEF*) and two subsequent enzymes, together with two additional genes of unknown function, were cloned and sequenced (van Hylckama Vlieg et al., 2000). Recently, we sequenced the genome of *Rhodococcus* sp. AD45 and showed, by mutagenesis, that isoprene monooxygenase (IsoMO) was essential for isoprene metabolism. Using RNAseq, a cluster of 22 genes was identified, all of which were induced by isoprene or the immediate product of isoprene oxidation, epoxyisoprene (Crombie *et al.,* 2015).

DNA stable isotope probing (DNA-SIP) is a cultivation-independent technique with the ability to identify active substrate-consuming organisms in environmental

126 samples (Radajewski et al., 2000; Dumont and Murrell, 2005). The method relies on
127 incubation of samples with stable-isotope-labelled growth substrate. The
128 incorporation of isotope (typically ^{13}C or ^{15}N) into biomass (including DNA), enables
129 the identification of active microorganisms following separation of labelled and
130 unlabelled DNA by isopycnic (density gradient) centrifugation.

131 Despite its abundance and climatic importance, our knowledge of isoprene in the
132 environment is heavily skewed towards production in plants and atmospheric
133 oxidation, with only a few studies investigating isoprene bio-degradation. Isoprene is
134 an abundant plant secondary metabolite, also produced in soils, and would provide a
135 good source of carbon and energy for bacteria. Strains capable of growth on
136 isoprene have frequently been isolated from diverse environments, albeit generally
137 not characterized at the molecular level. Our hypothesis was that isoprene degraders
138 are widely distributed and may play an important role in the biogeochemistry of this
139 environmentally important trace gas. Our aim was to isolate and sequence isoprene
140 degraders, and identify putative isoprene metabolic genes. We aimed to develop
141 gene probes to target key diagnostic markers of isoprene degradation and to identify
142 the active isoprene-assimilating organisms in soil enrichments, using DNA-SIP.

143

144

145 **Results and Discussion**

146 *Enrichment and isolation of two novel terrestrial isoprene-utilizing bacteria*

147 Our initial aim was to isolate isoprene degraders from contrasting environments in
148 order to provide sequence data for later cultivation-independent approaches. Since

isoprene consumption has previously been observed in soils and since isoprene is emitted from tree leaves, we used these as source material. Separate enrichments were set up, using either garden soil or Horse Chestnut (*Aesculus hippocastanum*) leaves, in minimal medium and incubated with isoprene. Both enrichments consumed isoprene and two isolates (designated SC4 and LB1), capable of growing on isoprene as sole source of carbon and energy, were obtained from soil and leaf samples, respectively. The nearly complete (1521 nucleotides) 16S rRNA gene sequences of both strains SC4 and LB1 were identical to strains of both *Rhodococcus opacus and R. wratislaviensis* (Fig. S1). Strains SC4 and LB1 also grew on acetate, succinate, glucose, fructose, propane and butane as sole source of carbon and energy (Table S1), in contrast to *Rhodococcus* sp. AD45, which does not grow on propane or butane (Crombie et al., 2015).

Genome sequencing

To identify isoprene-related gene sequences, we sequenced the genomes of strains SC4 and LB1. Their genomes, 10.6 and 10.7 Mbp, are considerably larger than that of *Rhodococcus* sp. AD45 (6.9 Mbp) and closer to that of *R. jostii* RHA1 (9.7 Mbp) (McLeod et al., 2006; Crombie et al., 2015), whereas the GC contents (66.7 and 66.6% respectively) are typical of the genus (Table S2).

Identification of isoprene related genes in isolates

Our previous work (Crombie *et al.,* 2015) detected a cluster of 22 genes induced by isoprene, which appeared to be the complete inventory of isoprene-responsive

genes. In addition to *isoABCDEF,* encoding the monooxygenase, and *isoGHIJ* that encode a protein of unknown function, a dehydrogenase and two glutathione-*S*-transferases previously described (van Hylckama Vlieg et al., 2000), the cluster includes glutathione biosynthesis genes, predicted aldehyde dehydrogenases and a coenzyme-A disulfide reductase (Crombie et al., 2015). We therefore searched for homologous sequences in the genomes of strains SC4 and LB1, which are highly similar to each other in this region (over 99% nucleotide identity). Using the isoprene-responsive gene products previously identified in *Rhodococcus* sp. AD45 as query sequences in tBLASTn searches, we identified homologues of all of these 22 genes, with amino acid sequence identity ranging from 50–96% (Table S3). The most highly conserved were *isoABCDEF* encoding the multi-component IsoMO (81–96%), whereas a predicted protein of unknown function (SZ00_06083), highly induced by isoprene in *Rhodococcus* sp. AD45, shared 50% amino acid identity with sequences from these strains. We observed the same duplication of *isoGHIJ* (Fig. 2, 77 – 88% amino acid identity between copies) as is present in *Rhodococcus* sp. AD45. Between *isoA* and *isoJ*, both of the new strains contain a second copy (81% amino acid identity between copies), not present in *Rhodococcus* sp. AD45, of an aldehyde dehydrogenase (*aldh1*) which is located approximately 10,000 nucleotides (nt) upstream of the monooxygenase in *Rhodococcus* sp. AD45. A gene encoding a predicted coenzyme-A disulfide reductase is also present in two copies (69% amino acid identity between copies), although in strain LB1 (but not in strain SC4) one copy has a nucleotide insertion approximately 255 nt from the end, resulting in a frameshift mutation, suggesting this may not encode a functional protein. In comparison with *Rhodococcus* sp. AD45, five additional genes are present, in both strains, in the middle of the cluster. These are predicted to encode two hypothetical

proteins and an alpha/beta hydrolase domain-containing protein of unknown function, an acetyl-CoA acetyltransferase, and a 3-hydroxyacyl-CoA dehydrogenase. Interestingly, in strain LB1, another insertion in the acetyl-CoA acetyltransferase has resulted in a frameshift mutation. None of these five genes are present in this region of the *Rhodococcus* sp. AD45 genome, nor were remotely-located homologous sequences induced by isoprene (Crombie et al., 2015), implying that they are not essential for isoprene metabolism. In comparison with *Rhodococcus* sp. AD45, strains LB1 and SC4 are more similar in this region of the genome, in terms of both sequence identity and gene layout, with *R. opacus* PD630 (also included in Fig. 2). Previously, using sequence data, we predicted, and confirmed, that *R. opacus* PD630 could grow on isoprene (Crombie et al., 2015), although, beyond this, we have no direct experimental data regarding isoprene-related gene function in this strain.

Development and validation of functional gene markers targeting isoA

Using the sequence data from these isoprene-degrading strains, we designed probes to detect isoprene-related genes in environmental samples. *isoA* codes for the alpha-subunit of the hydroxylase of IsoMO and contains the diiron centre active site and was shown to be essential for isoprene degradation in *Rhodococcus* sp. AD45 (Crombie *et al.,* 2015). We therefore selected *isoA* as target for a functional gene probe. To expand the diversity of *isoA* sequences, we also retrieved *isoA* genes from draft sequences of a *Gordonia* strain and a *Mycobacterium* strain previously isolated from surface sediment from the Colne estuary (Essex, UK), identified as *Gordonia* i37 and *Mycobacterium* AT1 (Acuña Alvarez et al., 2009;

Johnston, 2014). The *isoA* sequences of strains SC4, LB1, i37 and AT1, together with those of *Rhodococcus* sp. AD45 (SZ00_06091) and *R. opacus* PD630 (Pd630_LPD03572), were aligned at the amino acid level, and conserved regions were used to design primers targeting *isoA*. To exclude other non-isoprene-degrading members of the SDIMO family, sequences of *mmoX* and *xamoA*, encoding the alpha subunits of sMMO from *Methylosinus trichosporium* OB3b (Cardy et al., 1991), and alkene monooxygenase from *Xanthobacter autotrophicus* Py2 (Zhou et al., 1996), respectively, were also included in the alignment.

The *isoA* primers were validated by PCR amplification of template DNA from contrasting sources (Table S4): (i) genomic DNA from 15 isoprene-degrading isolates, (ii) DNA extracted from three isoprene-enriched soils and four isoprene enrichments of marine and estuarine water and sediment, and (iii) control DNA extracted from eight non-isoprene-degrading isolates able to grow on alkanes, alkenes or aromatic compounds. We obtained PCR products of the expected size using DNA extracted from all the isoprene-degrading isolates and enrichments, but not from any of the non-isoprene degraders. PCR products from enrichments were cloned and analysed by restriction fragment length polymorphism (RFLP) (Table S5). Representatives of each operational taxonomic unit (OTU) were sequenced, (all of which appeared to be *isoA* sequences) and aligned at the amino acid level with IsoA sequences obtained from the sequenced genomes. A phylogenetic tree of the *isoA* nucleotide sequences (1011 nt) was constructed from the alignment (Fig. 3). All the sequences, although from diverse phylogenetic groups including both Gram-positive and Gram-negative strains, were relatively similar (> 86% amino acid identity between sequences), but could be broadly separated into two groups in which the terrestrial sequences and those from the low-salinity environment of Hythe, on the

Colne estuary, were distinct from marine and other estuarine sequences similar to IsoA of *Gordonia* i37.

Active isoprene-assimilating bacteria identified by DNA-stable isotope probing (SIP)

Since all four of our sequenced terrestrial isolates were rhodococci and all *isoA* sequences retrieved were relatively similar, we used cultivation-independent methods to test whether a greater diversity of isoprene degraders existed in soils in a DNA-stable isotope experiment (DNA-SIP) using ^{13}C-labelled isoprene, biosynthesized as described in Experimental Procedures. Soil microcosms (in triplicate for labelled substrate and duplicate for unlabelled-substrate controls) were incubated with 0.5% (v/v) isoprene, without any other amendments. Isoprene was consumed without an appreciable lag phase, and incubations were terminated after consumption of 10 µmol substrate g^{-1} soil (15 days). DNA was extracted from ^{13}C- and control ^{12}C-isoprene incubations and separated into heavy and light fractions and used for 454 pyrosequencing of 16S rRNA genes. The unenriched soil at timepoint zero displayed a typically diverse community, comprising 50% Proteobacteria, with Planctomycetes, Actinobacteria, Bacteroidetes, Chloroflexi and Firmicutes contributing an additional 40% (Fig. 4). The major effect of the incubations was to greatly increase the relative abundance of Actinobacteria, while having a relatively minor effect on the remaining phyla (Fig. 4). Multivariate analysis (Fig. S2) shows that the unlabelled bacterial community of the ^{13}C-isoprene incubations (i.e. light DNA fraction) was extremely similar to that of the timepoint zero community, suggesting that the change in community profile during the incubations was due to an enrichment of isoprene degraders. As expected, the community represented by the light fraction of the ^{12}C-isoprene incubations, derived

271 from both isoprene-consumers and non-consumers, was extremely similar to the
272 total (unfractionated) DNA from ^{13}C-incubations. The heavy fraction of ^{12}C-isoprene
273 incubations, which contained only a very small fraction of the total DNA of ^{12}C-
274 incubations, was less similar. This DNA results mainly from incomplete separation,
275 with minor effects due to factors such as GC content of DNA. In contrast to all of
276 these, the active isoprene degraders represented by the heavy fractions of the ^{13}C
277 incubations formed a distinct community (Fig. S2) and displayed a dramatically
278 altered taxonomic profile (Fig. 5), becoming dominated by *Rhodococcus* sequences,
279 which comprised 88 ± 5% (mean ± s.d.) of the heavy fraction of ^{13}C incubations.
280 These sequences were predominantly those of *Rhodococcus wratislaviensis, R.*
281 *koreensis* and *R. globerulus* (Fig. 5) and were 147- to 161-fold more abundant in the
282 heavy fractions compared to light fractions of ^{13}C incubations, but without any
283 corresponding enrichment in the heavy fraction of ^{12}C-incubations (Fig. S3). Also
284 enriched in the heavy fractions of ^{13}C-incubations were members of the
285 Betaproteobacteria, *Comamonas* spp. and *Variovorax* spp., which together
286 comprised 6.5 ± 1.2% of the ^{13}C heavy fractions (but were not detected in the ^{13}C
287 light fractions), and were also enriched during the incubations, increasing from 0.3%
288 to 0.7% of the total community. These data clearly demonstrate that *Rhodococcus*
289 and, to a lesser extent, *Comamonas* and *Variovorax* spp. had assimilated carbon,
290 directly or indirectly, from labelled isoprene. Searches of the publicly available
291 databases did not identify high-similarity putative isoprene-related genes in the
292 family Comamonadaceae (which encompasses both *Comamonas* and *Variovorax*)
293 (see Experimental Procedures). Comparison with the PCR-based *isoA* analysis,
294 which generated similar sequences from diverse phylogenetic groups, suggests that
295 the isoprene-degrading members of the Comamonadaceae implicated in the SIP

296 experiment are not represented in the NCBI databases, or that they use genes or
297 possibly pathways dissimilar to those of the characterized strains.

298 Previous studies of terrestrial environments have isolated *Alcaligenes, Klebsiella* and
299 *Pseudomonas* isoprene-degrading strains in addition to the Actinobacteria *Nocardia*
300 and *Rhodococcus*, although some of these identifications were not based on
301 molecular data and most strains were not extensively characterized (van Ginkel et
302 al., 1987b; van Ginkel et al., 1987a; Ewers et al., 1990; Cleveland and Yavitt, 1997;
303 van Hylckama Vlieg et al., 1998; Srivastva et al., 2015). In our SIP incubations, 16S
304 rRNA gene sequences of *Alcaligenaceae* spp. were slightly enriched during the
305 incubations, increasing from 0.3% to 0.6% of the total community, but were
306 concentrated (14:1) in the non-labelled light DNA fractions. *Pseudomonas* spp.,
307 present at 0.9% of the initial community, decreased to 0.1% following isoprene
308 incubations and were also not labelled, whereas *Klebsiella* were not detected at any
309 point. These data indicate that these taxa had not assimilated carbon from isoprene
310 under our experimental conditions.

311

312 *Conclusions*

313 The isoprene concentrations used here are considerably in excess of those normally
314 encountered in the environment and the aim was not to replicate environmental
315 conditions, but rather to expand the diversity of known isoprene degraders. Draft
316 genome sequences showed that isoprene monooxygenase was present in all our
317 isoprene-degrading isolates, which enabled the development of *isoA* primers to
318 effectively target this enzyme with high specificity. While we cannot exclude the
319 possibility that some sequences may be missed, or that other isoprene-degrading

enzymes or pathways exist, these primers constitute an effective tool to identify isoprene-related gene sequences in environmental samples. DNA-SIP, to our knowledge the first time this technique has been used to identify isoprene assimilators, showed that the major isoprene utilisers in the microcosms were *Rhodococcus* strains and that members of the *Comamonadaceae* were also active in isoprene degradation. The fact that we did not obtain isolates from this family suggests that they may be resistant to cultivation under our laboratory conditions, emphasizing the importance of cultivation-independent techniques. This study confirms that soils readily consume isoprene and harbor a diverse community of isoprene degraders. Investigation of their diversity, abundance and mechanisms of isoprene degradation is essential to assess the environmental relevance of the global biological isoprene sink and the extent to which biodegradation moderates the effect on the atmosphere of this abundant and climate-active trace gas. This study provides the tools and foundations to further investigate these topics. Future experiments should search for novel genes and pathways involved in isoprene degradation, perhaps using SIP coupled with metagenomics. The isoprene-degrading community in the phyllosphere is also worthy of investigation, since our isolation of a strain from this environment, to our knowledge the first published example, suggests that isoprene degraders may be present or abundant on leaves, close to the major source of isoprene to the atmosphere.

Experimental procedures

Cultivation of bacterial strains

The terrestrial isoprene-degrading strains were routinely grown in minimal medium (CBS medium) which contained (per 1 litre): 0.1 g $MgSO_4.7H_2O$, 0.8 g NH_4Cl, 1.5 g

344 KH_2PO_4, 6.3 g Na_2HPO_4, and 10 ml of trace element solution (Tuovinen and Kelly,
345 1973) (pH 7.0). Marine isoprene degrading-bacteria were grown in mineral salts
346 medium prepared according to Schaefer et al., (2002), except containing 20 g l^{-1}
347 NaCl and supplemented with Na_3VO_4 and Na_2SeO_3 (5 ng L^{-1} each). Cultures were
348 set up in serum vials (120 ml) sealed with grey butyl rubber seals or in Quickfit flasks
349 (250 ml or 2 L) fitted with SubaSeal stoppers (Sigma-Aldrich) and isoprene was
350 added (as gas) to a final concentration of 0.6% - 1% (v/v) by injection through the
351 septum. The cultures were incubated at 30 °C, shaking at 150 rpm.
352 For growth tests on other gaseous substrates, 25 ml of CBS medium in serum vials
353 (120 ml) was inoculated with isoprene-grown culture (5% inoculum) and incubated
354 with 10% (v/v) substrate. Cultures with succinate, glucose, fructose or acetate (10
355 mM) as growth substrates were prepared in universal bottles (20 ml) containing 5 ml
356 CBS medium inoculated with 5% of isoprene-grown culture. Isoprene (catalogue no.
357 I19551) was obtained from Sigma Aldrich.
358

359 *Isolation of isoprene-degrading strains*

360 Isoprene enrichment cultures were set up using garden soil from Leamington Spa
361 (UK) or leaves of a Horse Chestnut tree from the campus of the University of
362 Warwick (Coventry, UK). Isoprene (0.6 % v/v) was added to 50 ml CBS minimal
363 medium in flasks (250 ml) and inoculated with either 0.3 g soil or one leaf, cut into
364 small pieces. The optical density of enrichment cultures was followed
365 spectrophotometrically at 540 nm and isoprene uptake was monitored with a gas
366 chromatograph fitted with a flame ionization detector (GC-FID) as described
367 previously (Crombie et al., 2015). Enrichment cultures were streaked on CBS agar

plates and incubated at 30 °C in a desiccator with isoprene vapour (approximately 5% v/v). Colonies were subcultured until pure, confirmed by phase contrast microscopy (Zeiss Axioscop, UK). Marine strains for primer design and validation were isolated as described previously (Acuña Alvarez et al., 2009; Johnston, 2014).

DNA extraction, amplification and sequencing of 16S rRNA genes

DNA was extracted from isolates and enrichment cultures using the FastDNA Spin Kit for Soil (MP Biomedicals), following the manufacturer's instructions. For identification of strains, 16S rRNA genes were amplified using 27f / 1492R primers (Lane, 1991). Amplicons were purified, cloned into pGEMT-easy vector (Promega), and sequenced with M13 primers (Invitrogen).

Genome sequencing, annotation and mining

High molecular-mass genomic DNA was extracted from 500 ml mid-late exponential cultures of isoprene-degrading strains following the Marmur extraction method (Marmur, 1961) except that the sodium dodecyl sulfate (SDS) concentration was increased to 2% (w/v) and the incubation period at 55 °C in sucrose/ethylenediaminetetraacetic acid/Tris (SET) buffer and SDS extended to 5 hours to achieve better cell lysis.

The genome of *Gordonia* i37 was sequenced at Oregon State University (USA) using a Roche 454 pyrosequencing platform. The genomes of *Mycobacterium* AT1, *Rhodococcus* SC4 and *Rhodococcus* LB1 were sequenced using Illumina GAIIx at the University of Warwick Genomics Facility (Coventry, UK). Reads were assembled into contigs using CLC Genomics Workbench for de novo assembly (CLC bio,

Aarhus, Denmark). The genome sequences were uploaded to RAST (Rapid Annotation using Subsystem Technology) for annotation. Local nucleotide databases were constructed using NCBI BLAST in BioEdit. The database was mined using tBLASTn with the amino acid sequences of genes shown to be important in isoprene metabolism in *Rhodococcus* sp. AD45 as query sequences. These Whole Genome Shotgun projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers LTCZ00000000 (SC4) and LSBM00000000 (LB1). Versions described in this paper are LTCZ01000000 and LSBM01000000. Cloned *isoA* sequences and *isoA* gene sequences of *Mycobacterium* AT1 and *Gordonia* i37 have been deposited under accession numbers KU870702 – KU870744 and KU870745 and KU870746, respectively.

isoA *Primer design and PCR*

Deduced amino acid sequences of *isoA* and other related enzymes were aligned using ClustalW in Mega6 (Tamura et al., 2013). The alignment was visually inspected for conserved regions specific to isoprene degraders, which were not conserved among homologous sequences from related SDIMO enzymes. Primers for *isoA* amplification, containing a maximum of three degenerate bases, were designed based on these regions (5´-TGCATGGTCGARCAYATG-3´ and 5´-GRTCYTGYTCGAAGCACCACTT-3´), yielding a predicted amplicon of 1015 bp (*Rhodococcus* sp. AD45 template). These primers were used with a touchdown PCR protocol: an initial step at 94 °C for 3 min was followed by 19 cycles of denaturation at 94 °C for 30 seconds, annealing at variable temperatures for 45 seconds, and extension at 72 °C for 60 seconds. The annealing temperature, initially 72 °C, was

decreased by 1 °C per cycle until 54 °C and maintained at this temperature for a further 25 cycles, followed by a final extension at 72 °C for 5 min.

Clone libraries

Clone libraries were constructed from *isoA* amplicons using DNA extracted from isoprene enrichments of soil, sediment or water. The purified *isoA* amplicons were cloned into pGEMT Easy vector (Promega) prior to transformation into *E. coli* TOP10 cells (Invitrogen) following the manufacturers' instructions. Clones were screened by PCR using M13 primers and restriction fragment length polymorphism (RFLP) using *Eco*RI and *Msp*I restriction enzymes (Fermentas). Clones yielding identical restriction fragment profiles on agarose gels were designated operational taxonomic units (OTUs) and representatives of each OTU were sequenced using M13 primers.

Search for iso genes in the *Comamonadaceae*

Tblastn (Altschul et al., 1990) was used to search the NCBI nr and genome databases and 153 whole genome shotgun projects deposited in NCBI, for isoprene-degrading gene sequences among the Comamonadaceae. Using IsoA from *Rhodococcus* sp. AD45 as query, genes with 46 – 48% amino acid identity were identified in the genomes of *Hydrogenophaga* sp. T4, *Comamonas badia* DSM 17552 and *Variovorax paradoxus* ZNC0006. Genes encoding the other subunits of the monooxygenase were identified in the *Hydrogenophaga* and *Comamonas* genomes, but no identifiable additional isoprene metabolic genes were found nearby, and no hits to IsoI were found in any of the genomes in any location (evalue 0.0001).

The *isoA* homologues were more similar to characterized toluene monooxygenase genes, such as *touA* from *Pseudomonas stutzeri* OX1 (67 – 73% amino acid identity) (Bertoni et al., 1998), suggesting that isoprene may not be their natural substrate.

Biosynthesis of [1,2,3,4,5-^{13}C]-2-methyl-1,3-butadiene ([U-^{13}C]-isoprene) from D-[U-^{13}C]-glucose

Uniformly labelled ^{13}C-isoprene was prepared utilizing an isoprene-producing engineered variant of *E. coli* BL21, expressing a functional mevalonate pathway and an isoprene synthase, as biocatalyst. Specifically, the genes encoding thiolase, mevalonate synthase, and 3-hydroxy-3-methyl-glutaryl- (HMG-) CoA reductase (*mvaE, mvaS*) were derived from *Enterococcus faecalis*, the gene encoding mevalonate kinase was derived from *Methanosarcina mazei*, and the genes encoding phosphomevalonate kinase, mevalonate pyrophosphate decarboxylase and isopentyl-PP isomerase were from *Saccharomyces cerevisiae*. This recombinant pathway serves to channel carbon from glucose to the isoprene precursor dimethylallyl pyrophosphate (DMAPP). The final intracellular enzymatic step to convert DMAPP to isoprene was catalyzed by recombinant isoprene synthase from *Populus alba*. This strain was designated *E. coli* CMP1082 and has been described in detail elsewhere (McAuliffe et al., 2015). Details of fermentation and isoprene recovery and its ^{1}H NMR spectrum are included in Supporting Information.

DNA stable isotope probing (SIP) experiments

For SIP incubations, 5 g soil (pH 7.4), collected from the upper 5 cm (after removal of vegetation and leaf litter), in the vicinity of Willow (*Salix fragilis*) trees on the

University of East Anglia campus, was incubated in serum vials (120 ml volume) with 0.5% (v/v) isoprene in triplicate (labelled substrate) or duplicate (unlabelled substrate). Vials were incubated at room temperature (22 °C) in the dark. Headspace isoprene concentration was monitored by GC-FID. When substrate was depleted, additional isoprene was injected through the septum to the original concentration. Incubations were terminated and DNA extracted from the soil when 10 μmol g^{-1} had been consumed. Control incubations with autoclaved soil were also conducted. DNA extracted from samples was separated into heavy (^{13}C-labelled) and light (unlabelled) fractions as previously described (Neufeld et al., 2007). Briefly, 1-2 μg DNA was added to caesium chloride solution (final density of 1.725 g ml^{-1}) and subjected to density gradient ultracentrifugation (177,000 x *g*, 40 h, 20 °C, Beckman Vti 65.2 rotor). The contents of each tube was separated into 12 – 15 fractions and the density of each fraction measured by refractometry (AR200 digital refractometer, Reichert Inc., Buffalo, USA). DNA from each fraction was precipitated, resuspended in nuclease-free water and characterized by 16S rRNA gene analysis using denaturing gradient gel electrophoresis (DGGE). DGGE showed unique profiles in heavy, compared to light, fractions of ^{13}C-incubations while the ^{12}C incubations exhibited similar profiles between fractions, indicating that heavy fractions from ^{13}C incubations contained the DNA of isoprene degraders (data not shown). Based on fraction density and DGGE profiles, fractions 7 and 11 were identified as containing labelled and unlabelled DNA, respectively, designated “heavy” and “light” and used for downstream analysis. DNA was obtained from six different treatments: extracted from timepoint zero (one sample); extracted following ^{13}C-isoprene incubation but prior to fractionation (unfractionated); and after separation into heavy and light fractions for both ^{13}C- and ^{12}C-isoprene incubations. For ^{13}C heavy fractions triplicate

samples were analysed individually, but for the other treatments the DNA was pooled prior to analysis.

Denaturing gradient gel electrophoresis (DGGE)

Bacterial 16S rRNA genes were amplified using primers 341F-GC (Muyzer et al., 1993) and 907R (Muyzer et al., 1998). The PCR products were run on an 8% (w/v) polyacrylamide gel with a 30% - 70% linear denaturant gradient. Electrophoresis was carried out for 16 h, at 80 V, 60 °C using the DCode™ Universal Mutation Detection System (Bio-Rad), stained with SYBR® Gold Nucleic Acid Gel Stain (Invitrogen) and imaged using a BioRad GelDoc system.

Bacterial 16S rRNA gene analysis by 454 pyrosequencing

Labelled (heavy) and unlabelled (light) DNA from SIP incubations was characterized by sequencing of 16S rRNA gene amplicons generated by PCR using the primers 27Fmod (5′-AGRGTTTGATCMTGGCTCAG-3′) and 519Rmodbio (5′-GTNTTACNGCGGCKGCTG-3′) using a Roche 454 FLX titanium instrument at MR DNA (Molecular Research LP), Shallowater, USA. Sequence data were processed at MR DNA using a published pipeline (Dowd et al., 2008; Capone et al., 2011). Briefly, the Q25 reads were stripped of barcodes and primers. Short sequences (<200 bp), sequences with ambiguous base calls and those with >6 bp homopolymer runs were removed. Remaining sequences were denoised using a custom pipeline, OTUs clustered at 97% sequence identity, chimeric sequences were removed using Uchime (Edgar et al., 2011) and taxonomy was assigned using BLASTn against the

RDPII/NCBI database (v 11.1) (Cole et al., 2014). An average of 3,320 sequences per sample were used for analysis.

Statistical Analysis

The Bray-Curtis distance measure was applied to log-transformed relative abundance and used to generate non-metric multidimensional scaling (NMDS) plots using Primer 6 (Primer-E, Plymouth, UK).

Acknowledgements

We thank Dan Arp, Luis Sayavedra-Soto and colleagues at Oregon State University for *Gordonia* sequence information. This research was funded by a Natural Environment Research Council (NERC) grant to JCM (NE/J009725/1) and TJM (NE/J009555/1) entitled *Microbial degradation of isoprene in the terrestrial environment* and by the Earth and Life Systems Alliance (ELSA) at the University of East Anglia. We acknowledge the support of a University of Warwick Postgraduate Research Scholarship awarded to MK and a NERC studentship awarded to AJ. We thank Hendrik Schäfer for help with genome assembly.

Acuña Alvarez, L., Exton, D.A., Timmis, K.N., Suggett, D.J., and McGenity, T.J. (2009) Characterization of marine isoprene-degrading communities. *Environ Microbiol* **11**: 3280-3291.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403 - 410.

Atkinson, R., and Arey, J. (2003) Atmospheric degradation of volatile organic compounds. *Chem Rev* **103**: 4605-4638.

Bäck, J., Aaltonen, H., Hellén, H., Kajos, M.K., Patokoski, J., Taipale, R. et al. (2010) Variable emissions of microbial volatile organic compounds (MVOCs) from root-associated fungi isolated from Scots pine. *Atmos Environ* **44**: 3651-3659.

Bertoni, G., Martino, M., Galli, E., and Barbieri, P. (1998) Analysis of the gene cluster encoding toluene/o-xylene monooxygenase from *Pseudomonas stutzeri* OX1. *Appl Environ Microbiol* **64**: 3626-3632.

Broadgate, W.J., Malin, G., Küpper, F.C., Thompson, A., and Liss, P.S. (2004) Isoprene and other non-methane hydrocarbons from seaweeds: a source of reactive hydrocarbons to the atmosphere. *Mar Chem* **88**: 61-73.

Capone, K.A., Dowd, S.E., Stamatas, G.N., and Nikolovski, J. (2011) Diversity of the human skin microbiome early in life. *J Invest Dermatol* **131**: 2026-2032.

Cardy, D.L., Laidler, V., Salmond, G.P., and Murrell, J.C. (1991) Molecular analysis of the methane monooxygenase (MMO) gene cluster of *Methylosinus trichosporium* OB3b. *Mol Microbiol* **5**: 335-342.

Cleveland, C.C., and Yavitt, J.B. (1997) Consumption of atmospheric isoprene in soil. *Geophys Res Lett* **24**: 2379-2382.

Cleveland, C.C., and Yavitt, J.B. (1998) Microbial consumption of atmospheric isoprene in a temperate forest soil. *Appl Environ Microbiol* **64**: 172-177.

555 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y. et al. (2014)
556 Ribosomal Database Project: data and tools for high throughput rRNA analysis.
557 *Nucleic Acids Res* **42**: D633-642.
558 Crombie, A.T., Khawand, M.E., Rhodius, V.A., Fengler, K.A., Miller, M.C., Whited,
559 G.M. et al. (2015) Regulation of plasmid-encoded isoprene metabolism in
560 *Rhodococcus*, a representative of an important link in the global isoprene cycle.
561 *Environ Microbiol* **17**: 3314-3329.
562 Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G.,
563 and Edrington, T.S. (2008) Evaluation of the bacterial diversity in the feces of cattle
564 using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP).
565 *BMC Microbiol* **8**: 125.
566 Dumont, M.G., and Murrell, J.C. (2005) Stable isotope probing - linking microbial
567 identity to function. *Nat Rev Micro* **3**: 499-504.
568 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME
569 improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.
570 Ensign, S.A. (2001) Microbial metabolism of aliphatic alkenes. *Biochemistry* **40**:
571 5845-5853.
572 Ewers, J., Freier-Schroder, D., and Knackmuss, H.J. (1990) Selection of
573 trichloroethene (TCE) degrading bacteria that resist inactivation by TCE. *Arch*
574 *Microbiol* **154**: 410-413.
575 Exton, D.A., McGenity, T.J., Steinke, M., Smith, D.J., and Suggett, D.J. (2015)
576 Uncovering the volatile nature of tropical coastal marine ecosystems in a changing
577 world. *Glob Change Biol* **21**: 1383-1394.
578 Fiore, A.M., Naik, V., Spracklen, D.V., Steiner, A., Unger, N., Prather, M. et al.
579 (2012) Global air quality and climate. *Chem Soc Rev* **41**: 6663-6683.

580 Gelmont, D., Stein, R.A., and Mead, J.F. (1981) Isoprene — The main hydrocarbon
581 in human breath. *Biochem Bioph Res Co* **99**: 1456-1460.
582 Gray, C.M., Helmig, D., and Fierer, N. (2015) Bacteria and fungi associated with
583 isoprene consumption in soil. *Elem Sci Anth* **3**: 000053.
584 Greenberg, J.P., Guenther, A., Zimmerman, P., Baugh, W., Geron, C., Davis, K. et
585 al. (1999) Tethered balloon measurements of biogenic VOCs in the atmospheric
586 boundary layer. *Atmos Environ* **33**: 855-867.
587 Guenther, A.B., Jiang, X., Heald, C.L., Sakulyanontvittaya, T., Duhl, T., Emmons,
588 L.K., and Wang, X. (2012) The Model of Emissions of Gases and Aerosols from
589 Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling
590 biogenic emissions. *Geosci Model Dev* **5**: 1471-1492.
591 Johnston, A. (2014) Molecular ecology of marine isoprene degradation. PhD thesis,
592 School of Environmental Sciences, University of East Anglia, Norwich, UK
593 Julsing, M., Rijpkema, M., Woerdenbag, H., Quax, W., and Kayser, O. (2007)
594 Functional analysis of genes involved in the biosynthesis of isoprene in *Bacillus*
595 *subtilis*. *Appl Microbiol Biotechnol* **75**: 1377-1384.
596 Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J.
597 et al. (2013) Three decades of global methane sources and sinks. *Nature Geosci* **6**:
598 813-823.
599 Kottegoda, S., Waligora, E., and Hyman, M. (2015) Metabolism of 2-methylpropene
600 (isobutylene) by the aerobic bacterium *Mycobacterium* sp. strain ELW1. *Appl Environ*
601 *Microbiol* **81**: 1966-1976.
602 Kuzma, J., Nemecek-Marshall, M., Pollock, W., and Fall, R. (1995) Bacteria produce
603 the volatile hydrocarbon isoprene. *Curr Microbiol* **30**: 97-103.

Lane, D.J. (1991) 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*. Stackebrandt, E., and Goodfellow, M. (eds). New York: Wiley, pp. 115-175.

Leahy, J.G., Batchelor, P.J., and Morcomb, S.M. (2003) Evolution of the soluble diiron monooxygenases. *FEMS Microbiol Rev* **27**: 449-479.

Logan, B.A., Monson, R.K., and Potosnak, M.J. (2000) Biochemistry and physiology of foliar isoprene production. *Trends Plant Sci* **5**: 477-481.

Loivamäki, M., Mumm, R., Dicke, M., and Schnitzler, J.-P. (2008) Isoprene interferes with the attraction of bodyguards by herbaceous plants. *Proc Natl Acad Sci* **105**: 17430-17435.

Loreto, F., and Fineschi, S. (2015) Reconciling functions and evolution of isoprene emission in higher plants. *New Phytol* **206**: 578-582.

Marmur, J. (1961) A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J Mol Biol* **3**: 208-IN201.

McAuliffe, J.C., Muir, R.E., Nielsen, A.T., Peres, C.M., Vaviline, D.V., and Wells, D.H. Enhanced production of isoprene using host cells having decreased *ispA* activity. US patent 8,975,051 B2, March 10, 2015.

McLeod, M.P., Warren, R.L., Hsiao, W.W.L., Araki, N., Myhre, M., Fernandes, C. et al. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci* **103**: 15582-15587.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.

628 Muyzer, G., Brinkhoff, T., Nuebel, U., Santegoeds, C., Schaefer, H., and Waver, C.
629 (eds) (1998) *Denaturing gradient gel electrophoresis (DGGE) in microbial ecology*.
630 Dordrecht, The Netherlands: Kluwer Academic Publishers.
631 Neufeld, J.D., Vohra, J., Dumont, M.G., Lueders, T., Manefield, M., Friedrich, M.W.,
632 and Murrell, J.C. (2007) DNA stable-isotope probing. *Nat Protocols* **2**: 860-866.
633 Pacifico, F., Harrison, S.P., Jones, C.D., and Sitch, S. (2009) Isoprene emissions
634 and climate. *Atmos Environ* **43**: 6121-6135.
635 Pegoraro, E., Abrell, L., Van Haren, J., Barron-Gafford, G., Grieve, K.A., Malhi, Y. et
636 al. (2005) The effect of elevated atmospheric CO2 and drought on sources and sinks
637 of isoprene in a temperate and tropical rainforest mesocosm. *Glob Change Biol* **11**:
638 1234-1246.
639 Radajewski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. (2000) Stable-isotope
640 probing as a tool in microbial ecology. *Nature* **403**: 646-649.
641 Schaefer, J.K., Goodwin, K.D., McDonald, I.R., Murrell, J.C., and Oremland, R.S.
642 (2002) *Leisingera methylohalidivorans* gen. nov., sp. nov., a marine methylotroph
643 that grows on methyl bromide. *Int J Syst Evol Microbiol* **52**: 851-859.
644 Sharkey, T.D. (2013) Is it useful to ask why plants emit isoprene? *Plant Cell Environ*
645 **36**: 517-520.
646 Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from
647 plants: why and how. *Ann Bot-London* **101**: 5-18.
648 Srivastva, N., Shukla, A.K., Singh, R.S., Upadhyay, S.N., and Dubey, S.K. (2015)
649 Characterization of bacterial isolates from rubber dump site and their use in
650 biodegradation of isoprene in batch and continuous bioreactors. *Bioresource*
651 *Technol* **188**: 84-91.

652 Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6:
653 Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**: 2725-2729.
654 Tuovinen, O.H., and Kelly, D.P. (1973) Studies on the growth of *Thiobacillus*
655 *ferrooxidans*. *Archiv Mikrobiol* **88**: 285-298.
656 van Ginkel, C.G., Welten, H.G.J., and de Bont, J.A.M. (1987a) Oxidation of gaseous
657 and volatile hydrocarbons by selected alkene-utilizing bacteria. *Appl Environ*
658 *Microbiol* **53**: 2903-2907.
659 van Ginkel, C.G., de Jong, E., Tilanus, J.W.R., and de Bont, J.A.M. (1987b)
660 Microbial oxidation of isoprene, a biogenic foliage volatile and of 1,3-butadiene, an
661 anthropogenic gas. *FEMS Microbiol Lett* **45**: 275-279.
662 van Hylckama Vlieg, J.E.T., Kingma, J., van den Wijngaard, A.J., and Janssen, D.B.
663 (1998) A glutathione *S*-transferase with activity towards cis-1,2-dichloroepoxyethane
664 is involved in isoprene utilization by *Rhodococcus* sp. strain AD45. *Appl Environ*
665 *Microbiol* **64**: 2800-2805.
666 van Hylckama Vlieg, J.E.T., Kingma, J., Kruizinga, W., and Janssen, D.B. (1999)
667 Purification of a glutathione *S*-transferase and a glutathione conjugate-specific
668 dehydrogenase involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. *J*
669 *Bacteriol* **181**: 2094-2101.
670 van Hylckama Vlieg, J.E.T., Leemhuis, H., Spelberg, J.H.L., and Janssen, D.B.
671 (2000) Characterization of the gene cluster involved in isoprene metabolism in
672 *Rhodococcus* sp. strain AD45. *J Bacteriol* **182**: 1956-1963.
673 Vickers, C.E., Gershenzon, J., Lerdau, M.T., and Loreto, F. (2009) A unified
674 mechanism of action for volatile isoprenoids in plant abiotic stress. *Nat Chem Biol* **5**:
675 283-291.

Wiedinmyer, C., Friedfeld, S., Baugh, W., Greenberg, J., Guenther, A., Fraser, M., and Allen, D. (2001) Measurement and analysis of atmospheric concentrations of isoprene and its reaction products in central Texas. *Atmos Environ* **35**: 1001-1013.

Wiedinmyer, C., Greenberg, J., Guenther, A., Hopkins, B., Baker, K., Geron, C. et al. (2005) Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "isoprene volcano". *J Geophys Res-Atmos* **110**: D18307.

Zhou, N.Y., Chan Kwo Chion, C.K., and Leak, D.J. (1996) Cloning and expression of the genes encoding the propene monooxygenase from *Xanthobacter* Py2. *Appl Microbiol Biotechnol* **44**: 582-588.

Figure Legends

Figure 1. Isoprene metabolism in *Rhodococcus* sp. AD45. Enzymes: IsoABCDEF, isoprene monooxygenase; IsoI, glutathione-S-transferase; IsoH, dehydrogenase. HGMB, 1-hydroxy-2-glutathionyl-2-methyl-3-butene; GMB, 2-glutathionyl-2-methyl-3-butenal; GMBA, 2-glutathionyl-2-methyl-3-butenoic acid; SG, glutathione; GSH, reduced glutathione.

Figure 2. The isoprene metabolic gene cluster from *Rhodococcus* sp. AD45 (Crombie et al., 2015), together with similar regions from *R. opacus* PD630 and strains SC4 and LB1. The contigs containing the genes are identified by horizontal lines and numbers below. The monooxygenase genes are shown in red, and other colours indicate genes of the corresponding predicted function between strains. Locus tags and gene names are indicated with angled text. Locus tag prefixes: *R*.

699 sp. AD45, SZ00_; *R*. opacus PD630, Pd630_LPD; *R*. sp. SC4, AXA44_; *R*. sp. LB1,
700 AZG88_.

701

702

703 **Figure 3**. Phylogenetic tree of *isoA* genes from isolates and environmental clones,
704 constructed using the Maximum Likelihood method in MEGA6 (Tamura et al., 2013).
705 All positions containing gaps and missing data were eliminated and there were 1011
706 nucleotide positions in the final dataset. The scale bar shows nucleotide substitutions
707 per site. Bootstrap values (500 replications) greater than 50% are shown at the
708 nodes. Isoprene-degrading isolates are shown in bold. Cloned *isoA* sequences
709 (Tables S4 and S5) are labelled with the sample site followed by clone identification
710 number. WCO_L4, Western Channel Observatory station L4; FW, freshwater.

711

712

713 **Figure 4**. Soil bacterial community (based on 16S rRNA gene amplicons) of labelled
714 isoprene DNA-SIP microcosms. The pie charts show (left) the unenriched soil
715 community from timepoint zero and (right) the ^{13}C- isoprene-enriched total
716 community (prior to isopycnic centrifugation and DNA fractionation). For the right
717 hand chart, DNA from triplicate enrichments was pooled prior to analysis.

718

719

720 **Figure 5**. Bacterial communities derived from labelled isoprene DNA-SIP
721 enrichments and fractionation. The bar chart shows 16S rRNA gene-based
722 composition of the unenriched soil community (timepoint zero), the isoprene-
723 enriched total community (unfractionated) and labelled (heavy) and unlabelled (light)

components separated by isopycnic centrifugation and fractionation. The symbols on the x-axis correspond with those shown in Fig. S2. The isoprene-assimilating community is represented by the heavy fractions of the three replicate ^{13}C incubations (solid red diamonds). T-0, timepoint zero; U-F, unfractionated; H, heavy fraction; L, light fraction. The inset (b) shows the species composition of the rhodococci in the heavy fractions (mean of three replicates)