Poplar phyllosphere harbors disparate isoprene-degrading bacteria

Andrew T. Crombie\textsuperscript{a,1}, Nasmille L. Larke-Meja\textsuperscript{b}, Helen Emery\textsuperscript{b}, Robin Dawson\textsuperscript{b}, Jennifer Pratscher\textsuperscript{c}, Gordon P. Murphy\textsuperscript{d}, Terry J. McGinity\textsuperscript{a}, and J. Colin Murrell\textsuperscript{b}

\textsuperscript{a}School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom; \textsuperscript{b}School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom; \textsuperscript{c}LYLE Centre, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom; and \textsuperscript{d}School of Biological Sciences, University of Essex, Colchester CO4 3SQ, United Kingdom

Edited by Caroline S. Harwood, University of Washington, Seattle, WA, and approved November 1, 2018 (received for review July 24, 2018)

The climate-active gas isoprene (2-methyl-1,3-butadiene) is released to the atmosphere in huge quantities, almost equaling that of methane, yet we know little about the biological cycling of isoprene in the environment. Although bacteria capable of growth on isoprene as the sole source of carbon and energy have previously been isolated from soils and sediments, no microbiological studies have targeted the major source of isoprene and examined the phyllosphere of isoprene-emitting trees for the presence of degraders of this abundant carbon source. Here, we identified isoprene-degrading bacteria in poplar tree-derived microcosms by DNA stable isotope probing. The genomes of isoprene-degrading taxa were reconstructed, putative isoprene metabolic genes were identified, and isoprene-related gene transcription was analyzed by shotgun metagenomics and metatranscriptomics. Gram-positive bacteria of the genus Rhodococcus proved to be the dominant isoprene degraders, as previously found in soil. However, a wider diversity of isoprene utilizers was also revealed, notably Variovorax, a genus not previously associated with this trait. This finding was confirmed by expression of the isoprene monooxygenase from Variovorax in a heterologous host. A Variovorax strain that could grow on isoprene as the sole carbon and energy source was isolated. Analysis of its genome confirmed that it contained isoprene metabolic genes with an identical layout and high similarity to those identified by DNA-stable isotope probing and metagenomics. This study provides evidence of a wide diversity of isoprene-degrading bacteria in the isoprene-degrading tree phyllosphere and greatly enhances our understanding of the biodegradation of this important metabolite and climate-active gas.

Isoprene (2-methyl-1,3-butadiene) is emitted to the atmosphere at a rate of \SI{\sim 500}{\text{Tg\,y}^{-1}}, on a par with methane and approximately one-third of the total volatile organic compound (VOC) emissions (1, 2). The vast majority originates from terrestrial plants (400–600 Tg\,y\textsuperscript{-1}), with a small but uncertain flux from marine algae (0.1–12 Tg\,y\textsuperscript{-1}) (2–4), bacteria, fungi, and animals (5, 6). A reactive diene, isoprene is rapidly photochemically oxidized (1), with a significant and complex effect on global climate (7). Hydroxyl (OH) and nitrate (NO\textsubscript{3}) radicals and ozone (O\textsubscript{3}) in the atmosphere react with isoprene depending on prevailing conditions (1). In pristine environments, isoprene reacts directly with ozone and hydroxyl radicals, resulting in ozone depletion. However, the high NO\textsubscript{3} levels typical of urban environments result in the formation of tropospheric ozone, with important negative effects on human health and on yields of ozone-sensitive crops (8). Globally, these reactions result in a net radiative forcing of 0.09 W m\textsuperscript{-2}, with an additional indirect effect since depletion of OH radicals increases the atmospheric lifetime of methane (9). The isoprene oxidation products form secondary organic aerosols and cloud condensation nuclei with implications for planetary albedo, air quality, and climate (10, 11).

Plants produce isoprene in the chloroplast and release it to the atmosphere from the abaxial surface of leaves via the stomata (12). Although isoprene production is not a universal trait among plants, it protects against heat and oxidative stress (13, 14) and has roles in plant/insect signaling and plant energy dynamics (5, 15). High-isoprene–emitting trees worldwide include oil palm (Elaeis guineensis), eucalyptus (Eucalyptus spp.) and, in temperate regions, English oak (Quercus robur), poplar (Populus spp.), and willow (Salix spp.), with reported emissions of 77 (English oak) and 175 (oil palm) μg g\textsuperscript{-1} (dry leaves) h\textsuperscript{-1} (16). Due to its short atmospheric lifetime, isoprene concentrations vary from a few parts per billion by volume (ppbv) (17), to tens of ppbv in high-isoprene–emitting forests (18). Soils acted as a biological sink, both in closed chambers (19, 20) and in continuous systems using isoprene concentrations of 2–200 ppbv (21), suggesting that soil microbes act as a significant sink for isoprene, consuming up to 4% of global emissions.

Several isoprene-consuming bacterial strains have been isolated from soils and marine sediments (6, 22–25), and recently several strains have been characterized biochemically and genetically (26–30). Before this work, all strains characterized by sequence data were members of the Actinobacteria, and all

Significance

Isoprene is a climate-active gas, produced in huge amounts by trees, yet we know little about its biogeochemical cycle. Bacteria able to grow on isoprene have been isolated from soils and sediments, but the phyllosphere, the principal isoprene source, has remained unexplored. Using targeted cultivation-independent techniques, we show that the phyllosphere of an isoprene-emitting tree contains a diverse and active isoprene-degrading population. We reconstruct the genome of an isoprene-degrading Variovorax strain and show that it contains a functional isoprene monooxygenase. This detailed study targets isoprene degraders from the phyllosphere, applies metatranscriptomics to isoprene degradation, and isolates and sequences an isoprene-degrading member of the Proteobacteria.

Author contributions: A.T.C., N.L.L.-M., G.P.M., T.J.M., and J.C.M. designed research; A.T.C., N.L.L.-M., H.E., R.D., and G.P.M. performed research; A.T.C., N.L.L.-M., H.E., J.P., and G.P.M. analyzed data; and A.T.C., T.J.M., and J.C.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The genome sequences of Variovorax sp. W511 and megaplasmid-cured Rhodococcus AD45-1D have been deposited at DNA Data Bank of Japan/European Nucleotide Archive/GenBank under accession nos. PXZZ01000000 and PYHL0000000. Versions described here are PXZZ01000000 and PYHL00000000, respectively. Sequence reads have been deposited at GenBank Sequence Read Archive (accession no. SRP101805).

1To whom correspondence should be addressed. Email: a.crombie@uea.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1812668115/-/DCSupplemental.

Published online November 29, 2018.
terrestrial examples were rhodococci. They all contain six genes (isoABCDEFG) encoding a multicomponent isoprene monoxygenase (IsoMO) (Fig. 1) with four additional genes, isoGHJL, just upstream. van Hylckama Vlieg et al. (26, 27) showed that iso emissions atribute to isoprene s-transferase that conjugates glutathione with the epoxide product of IsoMO, and that isoH encodes a dehydrogenase that acts on the product of IsoL. IsoG and IsoL are involved in the subsequent metabolic pathway but have yet to be characterized. In Rhodococcus AD45, the entire cluster isoGHJLABCDEF is cotranscribed as an operon (28). Significantly, all previously characterized isoprene gene clusters include the glutathione biosynthesis genes ghk/hb, suggesting a specific role for glutathione in isoprene metabolism. This small thiol is not usually found in Gram-positive bacteria (31), although a role in isoprene metabolism is also suggested for clostridia (32).

The plant microbiome plays an essential role in plant health and development (33), and the phyllosphere, although an unstable environment, constitutes an extensive habitat for microorganisms, mainly bacteria, the diverse community of which typically comprises \(10^{10} - 10^{12}\) cells per cm\(^2\) (34). In the intercellular spaces of leaves, near the point of emission from stomata, isoprene concentrations reach 30 ppmv, approximately three orders of magnitude higher than atmospheric concentrations (12, 35–37). Apart from isoprene, plants produce a range of other VOCs. Although leaf-dwelling methylotrophs can reduce plant methanol to formaldehyde (38), the plant microbiome moderates release of other VOCs is not known (39, 40). Human intervention is likely to alter global emissions of isoprene, partly because emissions increase with temperature but are inversely related to atmospheric carbon dioxide concentrations, but also due to the development of high-isoprene-emitting agroforestry (41, 42). Therefore, our aim was to better understand the diversity of microbial taxa involved in isoprene consumption and their mechanisms of action. Specifically, this study addresses the following questions: Does the microbiome of a leaf contain isoprene degraders able to take advantage of this abundant carbon source? And are novel genes expressed in response to isoprene?

**Results and Discussion**

**Isoprene Degraders Isolated from Leaves and Soil.** Leaf washings from high-isoprene—emitting tree species (oak, poplar, willow) consumed isoprene in microcosms. After 15 d, 65% of the microcosms showed >30% depletion of added isoprene (SI Appendix, Table S1), confirming the presence of isoprene-degrading microbes. We obtained three Rhodococcus isolates that grew on isoprene as the sole source of carbon and energy—two strains (ACPA1 and ACPA4) from poplar leaves and one (ACS1) from soil beneath trees. All contained isoprene metabolic genes similar to those identified previously (28, 43).

**Stable Isotope Probing Identifies Active Degraders.** To extend the revealed diversity beyond those amenable to cultivation, we used DNA-stable isotope probing (DNA-SIP) in two independent experiments. Since sun-exposed leaves may produce more isoprene than shaded leaves (44), leaves of white poplar (Populus alba) taken from locations on the tree exposed to full sun were used. Initially, microbial cells were dislodged from the leaves and incubated in minimal medium with \(^{13}\)C-labeled or unlabeled isoprene (headspace concentration ~500 ppmv). Added isoprene was consumed rapidly, and cells were harvested after 7 or 8 d (SI Appendix, Fig. S1). Extraction, fractionation (based on incorporation of 13C label), and quantification of DNA showed that labeled DNA from the 13C-isoprene incubations migrated to “heavy” and “light” regions of the ultracentrifuge tubes depending on incorporation of label, whereas DNA from 13C-isoprene incubations was restricted (>99%) to light fractions (SI Appendix, Fig. S2A). Analysis of 16S rRNA gene amplicons showed that the unenriched bacterial community, highly consistent across six replicates, was composed mainly of Proteobacteria [Sphingomonas (mean relative abundance (RA) 36%), Pseudomonas (17%), members of the Oxalobacteraceae (8%), Methylobacterium (1%), and members of the Comamonadaceae (25%)], Bacteroidetes [Hymenobacter (17%) and Sphingobacteriaceae (1%)], and Actinobacteria (4%) (Fig. 2 and SI Appendix, Fig. S3), in agreement with previous studies of plant-associated bacteria (45, 46). Following enrichment, the active isoprene-degrading community was dominated by Rhodococcus (78 ± 8% mean RA ± SD), although several proteobacterial taxa were also labeled, principally members of the Xanthomonadaceae (8 ± 7%) and Comamonadaceae (3 ± 1.6%) (Fig. 2 and SI Appendix, Fig. S3). So far, all well-described terrestrial isoprene degraders in cultivation are members of the genera Rhodococcus, Rhodocyclus, and Xanthomonas. We hypothesized the presence of sequences related to Proteobacteria in DNA-SIP incubations with labeled isoprene (29), despite the fact that no publicly available proteobacterial genomes contain recognizable isoprene-degradation–related gene sequences. We investigated these nonactinobacterial isoprene degraders in more detail, again using DNA-SIP.

The second DNA-SIP experiment, with isoprene at 150 ppmv, included incubations without isoprene (no substrate) as an additional control. RNA was also extracted from the incubations, although not subjected to ultracentrifugation or separation into heavy and light fractions. RNA and DNA from the enriched bacterial community was dominated by Rhodococcus (96% RA by amplicon and shotgun sequencing, respectively), confirmed the dominance of this operational taxonomic unit (OTU) among the unenriched bacterial community, highly consistent across six replicates, was composed mainly of Proteobacteria (73–99%) to light fractions (SI Appendix, Fig. S2B). In contrast to the first experiment, nucleic acids were analyzed using a shotgun approach. Community composition was assessed by clad-specific marker genes using MetaPhlAn2 (47) (Fig. 2, SI Appendix, Fig. S3), metaSPPH (48) (Fig. 3), and amplicon sequencing (49). This approach, albeit based on an extremely restricted community, was shown to accurately identify dominant families and genera (50). The second DNA-SIP experiment, with isoprene at 150 ppmv, included incubations without isoprene (no substrate) as an additional control. RNA was also extracted from the incubations, although not subjected to ultracentrifugation or separation into heavy and light fractions. RNA and DNA from the enriched bacterial community was dominated by Rhodococcus (96% RA by amplicon and shotgun sequencing, respectively), confirmed the dominance of this operational taxonomic unit (OTU) among the unenriched bacterial community, highly consistent across six replicates. Analysis of the same DNA, from the heavy fraction of the 13C-isoprene incubations of the first experiment, by both 16S rRNA ampicron and metagenomic sequencing, enabled direct comparison of the microbial community revealed by these two methods. Both confirmed the dominance of Rhodococcus sequences (78 and 94% RA by amplicon and shotgun sequencing, respectively), together with a contribution from Comamonadaceae (3 and 1.5%, respectively). However, the Xanthomonadaceae, identified as 8% among 16S rRNA amplicons, were not identified in significant numbers by the shotgun approach (<0.1% RA), possibly indicative of bias in the PCR-based method.

In the second SIP experiment, the time-point-zero community was dominated by the Bacteroidetes Hymenobacter and Pedobacter (50% RA). Sphingomonas, the most abundant genus of the first experiment, was a minor component (5%) in the second. Following isoprene enrichment, the labeled community was again dominated by Actinobacteria (mainly Rhodococcus, average 74% RA) together with, as before, Proteobacteria, notably Comamonadaceae, principally Variovorax, which averaged 16% of the labeled community, albeit with considerable intersample variability (SI Appendix, Fig. S3). In contrast, these taxa comprised a small fraction (0.5–1%) of the DNA from the pooled light fractions. The community profile of control incubations without added substrate was similar to that of the light fractions, further confirming that the heavy fractions contained the labeled DNA from isoprene consumption.

The transcriptionally active taxa were characterized by profiling the mRNA transcriptome reads using MetaPhlAn2 (SI Appendix, Fig. S4). This approach, albeit based on an extremely restricted subset of taxonomically informative marker genes, identified the Bacteroidetes genus Pedobacter and the Gammaproteobacterium
Pseudomonas as transcriptionally active, not only in isoprene enrichments, but also in unenriched time-point-zero samples and in incubations without added substrate, suggesting that these taxa were able to scavenge nutrients from endogenous organic matter and/or dead microbial cells. Transcripts of Rhodococcus and Variovorax, as well as the Bacteroidetes Chrysobacterium and Riemerella were specifically enriched in isoprene incubations, indicating that these taxa were directly or indirectly stimulated by isoprene (SI Appendix, Fig. S4).

Assembly Identifies isoA Sequences. To identify isoprene-degrading sequences, reads from pooled unenriched time-point-zero samples from each 13C-isoprene–enriched heavy fraction, from pooled light-fraction DNA, and from no-substrate incubations were coassembled, resulting in a total of 1.84 Gbp of sequence (SI Appendix, Table S3). Using, as query, the amino acid sequence of isoA (encoding the alpha subunit of IsoMO), we identified 11 sequences with 40% amino acid identity to isoA from Rhodococcus AD45 (Fig. 3). These scaffolds were examined in detail, and five contained sufficient sequence data to also identify one or more of the genes, specific for isoprene metabolism in Rhodococcus AD45 (28), immediately flanking the monoxygenase (Fig. 4) (36–100% amino acid identity with the Rhodococcus AD45 gene products), indicating a likely role in isoprene degradation. Similarly, reads originating from rRNA-depleted RNA were assembled, resulting in 220,637–434,247 transcripts containing 83.6–223.5 Mbp of sequence for each sample (SI Appendix, Table S4). Eighteen transcripts containing distinct isoA-related sequences were identified (Fig. 3), 8 of which contained other isoA-related genes in addition to the IsoMO genes isoABCDEF (SI Appendix, Table S5). Many of these genes and transcripts were closely related to those of the Rhodococcus isolates obtained in this study, as well as to other strains identified previously (28, 29) (Figs. 3 and 4). Within-sample quantification of global transcript abundance showed that these Rhodococcus-like isoprene gene transcripts were among the most highly expressed, many at the top 0.2 or 1% of community-wide transcripts (Fig. 3 and SI Appendix, Table S5). In addition to sequences grouping with those of characterized isolates, three highly expressed transcripts, centered around metagenome scaffold MG_3829, formed a distinct cluster (97% identity to isoA from Rhodococcus opacus PD630), suggesting that there is yet more diversity to discover among these isoprene-degrading Actinobacteria. Apart from these Rhodococcus-like isoA genes, more divergent sequences were identified, as described below. Significantly, low-level transcription of isoA was also observed at time-point zero (unenriched) and in incubations without isoprene (no substrate), suggesting that there is a degree of constitutive expression of these isoprene-degradation genes.

Genome Reconstruction Through Binning. To reconstruct individual genomes from the metagenome, scaffolds were assigned to 266 bins based on abundance and nucleotide composition. Bin quality was assessed and refined, and taxonomy was assigned, resulting in 27 genome bins with predicted completeness >90% and contamination <5%. From the 266 bins, 18 were identified as of interest based on the presence of an isoA sequence or predicted RA of >2% in the labeled community, and these were examined in greater detail (SI Appendix, Tables S6 and S7). In total, nine isoA-containing scaffolds were assigned to genome bins, allowing taxonomy to be inferred for most of these sequences (Fig. 3). Several sequences, similar to those of characterized isolates, were assigned to actinobacterial bins (Fig. 3). The most abundant (average 31% RA across three 13C-isoprene enrichments), predicted to be 99% complete and with 1.1% contamination (SI Appendix, Table S6), had 98.7% average nucleotide identity to isolate Rhodococcus sp. ACPA4. Several other (less complete) DNA-RNA bins were at least 93% identified. However, since DNA-SIP revealed a considerable diversity of labeled organisms, we also looked at more dissimilar sequences. Genome bin 197, assigned to Pseudonocardia, contained two isoA homologs. The first of these, on scaffold MG_3829, contained a typical isoprene metabolic gene cluster with 75–93% amino acid identity to six isoprene metabolic gene products, including isoA, from Rhodococcus AD45 (Fig. 4). The same genome also contained scaffold MG._720, with a less similar IsoA homolog (48% amino acid identity with isoA from Rhodococcus AD45). This region of scaffold MG._720 contained genes with high sequence identity and an identical layout to Rhodococcus jostii DSM44719 and Pseudonocardia dioxanivorans CB1190 (48) annotated as toluene-4 monoxygenases (SI Appendix, Fig. S5). These monoxygenase genes are of low similarity and arranged
in a different order to isoABCDEF from isoprene degraders and are not flanked by any other recognizable isoprene-related genes, suggesting that this region was not responsible for isoprene metabolism by known or predicted pathways. Interestingly, although they could not be identified as directly involved in isoprene degradation, these divergent genes on scaffold MG_720 of genome bin 197 were transcribed at moderately high levels (Fig. 3 and SI Appendix, Table S5). Genomic bin 095 contained a sequence with homology to a putative toluene monoxygenase from Myxococcales bacterium 68–20 (OY25088.1), with a similar gene layout. Again, although this genome contained no identifiable isoprene-related genes apart from the monoxygenase, the isoA homolog was represented in the transcriptome (Fig. 3 and SI Appendix, Table S5). These two examples suggest that isoprene is capable of inducing genes not central to its metabolism, although those monoclonal genes are constitutively expressed at considerable levels, or that they form part of a novel and so-far-undescribed isoprene metabolic pathway.

Genome bin 232, assigned to the betaproteobacterial genus Variorovax, contained an isoA homolog (scaffold MG_478), which was represented by a moderately expressed transcript and aligned most closely with xamoA of the propylene-degrader Xanthobacter autotrophicus Py2 (Fig. 3 and SI Appendix, Table S5). In contrast to characterized isoprene degraders, however, X. autotrophicus Py2 does not use glutathione in its alkene metabolic pathway (49), does not contain any homologs of the isoprene metabolic genes surrounding isoA-F at the xamoA locus (50), and does not grow on isoprene (30). Scaffold MG_478 comprised 76,719 bp of contiguous DNA, allowing examination of the genomic context of isoA (Fig. 4). In addition to the six genes (isoABCDEF) encoding IsoMO, isoGHIJ and aldhl1 were present in an identical layout to those of many isoprene-degrading isolates (Fig. 4), although sequence identity of the gene products with those of Rhodococcus AD45 ranged from 42–71%, much lower than those of characterized isolates. Glutathione biosynthesis genes gshA and gshB were not present in this isoprene cluster but were found on scaffold MG_3916 of this genome bin (85 and 76% amino acid identity to ACS17089.1 and ACS17095.1 from Myxococcus xanthus). Genomic context of isoABCDEF was represented by a moderately expressed transcript and highly expressed transcripts were not central to metabolism, that these monooxygenase genes are not flanked by any other recognizable isoprene-related genes, suggesting that this region was not responsible for isoprene metabolism by known or predicted pathways. Interestingly, although they could not be identified as directly involved in isoprene degradation, these divergent genes on scaffold MG_720 of genome bin 197 were transcribed at moderately high levels (Fig. 3 and SI Appendix, Table S5). Genomic bin 095 contained a sequence with homology to a putative toluene monoxygenase from Myxococcales bacterium 68–20 (OY25088.1), with a similar gene layout. Again, although this genome contained no identifiable isoprene-related genes apart from the monoxygenase, the isoA homolog was represented in the transcriptome (Fig. 3 and SI Appendix, Table S5). These two examples suggest that isoprene is capable of inducing genes not central to its metabolism, although those monoclonal genes are constitutively expressed at considerable levels, or that they form part of a novel and so-far-undescribed isoprene metabolic pathway.

Genome bin 232, assigned to the betaproteobacterial genus Variorovax, contained an isoA homolog (scaffold MG_478), which was represented by a moderately expressed transcript and aligned most closely with xamoA of the propylene-degrader Xanthobacter autotrophicus Py2 (Fig. 3 and SI Appendix, Table S5). In contrast to characterized isoprene degraders, however, X. autotrophicus Py2 does not use glutathione in its alkene metabolic pathway (49), does not contain any homologs of the isoprene metabolic genes surrounding isoA-F at the xamoA locus (50), and does not grow on isoprene (30). Scaffold MG_478 comprised 76,719 bp of contiguous DNA, allowing examination of the genomic context of isoA (Fig. 4). In addition to the six genes (isoABCDEF) encoding IsoMO, isoGHIJ and aldhl1 were present in an identical layout to those of many isoprene-degrading isolates (Fig. 4), although sequence identity of the gene products with those of Rhodococcus AD45 ranged from 42–71%, much lower than those of characterized isolates. Glutathione biosynthesis genes gshA and gshB were not present in this isoprene cluster but were found on scaffold MG_3916 of this genome bin (85 and 76% amino acid identity to ACS17089.1 and ACS17095.1 from Myxococcus xanthus). Genomic context of isoABCDEF was represented by a moderately expressed transcript and highly expressed transcripts were not central to metabolism, that these monooxygenase genes are not flanked by any other recognizable isoprene-related genes, suggesting that this region was not responsible for isoprene metabolism by known or predicted pathways. Interestingly, although they could not be identified as directly involved in isoprene degradation, these divergent genes on scaffold MG_720 of genome bin 197 were transcribed at moderately high levels (Fig. 3 and SI Appendix, Table S5). Genomic bin 095 contained a sequence with homology to a putative toluene monoxygenase from Myxococales bacterium 68–20 (OY25088.1), with a similar gene layout. Again, although this genome contained no identifiable isoprene-related genes apart from the monoxygenase, the isoA homolog was represented in the transcriptome (Fig. 3 and SI Appendix, Table S5). These two examples suggest that isoprene is capable of inducing genes not central to its metabolism, although those monoclonal genes are constitutively expressed at considerable levels, or that they form part of a novel and so-far-undescribed isoprene metabolic pathway.

Genome bin 232, assigned to the betaproteobacterial genus Variorovax, contained an isoA homolog (scaffold MG_478), which was represented by a moderately expressed transcript and aligned most closely with xamoA of the propylene-degrader Xanthobacter autotrophicus Py2 (Fig. 3 and SI Appendix, Table S5). In contrast to characterized isoprene degraders, however, X. autotrophicus Py2 does not use glutathione in its alkene metabolic pathway (49), does not contain any homologs of the isoprene metabolic genes surrounding isoA-F at the xamoA locus (50), and does not grow on isoprene (30). Scaffold MG_478 comprised 76,719 bp of contiguous DNA, allowing examination of the genomic context of isoA (Fig. 4). In addition to the six genes (isoABCDEF) encoding IsoMO, isoGHIJ and aldhl1 were present in an identical layout to those of many isoprene-degrading isolates (Fig. 4), although sequence identity of the gene products with those of Rhodococcus AD45 ranged from 42–71%, much lower than those of characterized isolates. Glutathione biosynthesis genes gshA and gshB were not present in this isoprene cluster but were found on scaffold MG_3916 of this genome bin (85 and 76% amino acid identity to ACS17089.1 and ACS17095.1 from Myxococcus xanthus). Genomic context of isoABCDEF was represented by a moderately expressed transcript and highly expressed transcripts were not central to metabolism, that these monooxygenase genes are not flanked by any other recognizable isoprene-related genes, suggesting that this region was not responsible for isoprene metabolism by known or predicted pathways. Interestingly, although they could not be identified as directly involved in isoprene degradation, these divergent genes on scaffold MG_720 of genome bin 197 were transcribed at moderately high levels (Fig. 3 and SI Appendix, Table S5). Genomic bin 095 contained a sequence with homology to a putative toluene monoxygenase from Myxococales bacterium 68–20 (OY25088.1), with a similar gene layout. Again, although this genome contained no identifiable isoprene-related genes apart from the monoxygenase, the isoA homolog was represented in the transcriptome (Fig. 3 and SI Appendix, Table S5). These two examples suggest that isoprene is capable of inducing genes not central to its metabolism, although those monoclonal genes are constitutively expressed at considerable levels, or that they form part of a novel and so-far-undescribed isoprene metabolic pathway.

Genome bin 232, assigned to the betaproteobacterial genus Variorovax, contained an isoA homolog (scaffold MG_478), which was represented by a moderately expressed transcript and aligned most closely with xamoA of the propylene-degrader Xanthobacter autotrophicus Py2 (Fig. 3 and SI Appendix, Table S5). In contrast to characterized isoprene degraders, however, X. autotrophicus Py2 does not use glutathione in its alkene metabolic pathway (49), does not contain any homologs of the isoprene metabolic genes surrounding isoA-F at the xamoA locus (50), and does not grow on isoprene (30). Scaffold MG_478 comprised 76,719 bp of contiguous DNA, allowing examination of the genomic context of isoA (Fig. 4). In addition to the six genes (isoABCDEF) encoding IsoMO, isoGHIJ and aldhl1 were present in an identical layout to those of many isoprene-degrading isolates (Fig. 4), although sequence identity of the gene products with those of Rhodococcus AD45 ranged from 42–71%, much lower than those of characterized isolates. Glutathione biosynthesis genes gshA and gshB were not present in this isoprene cluster but were found on scaffold MG_3916 of this genome bin (85 and 76% amino acid identity to ACS17089.1 and ACS17095.1 from Myxococcus xanthus). Genomic context of isoABCDEF was represented by a moderately expressed transcript and highly expressed transcripts were not central to metabolism, that these monooxygenase genes are not flanked by any other recognizable isoprene-related genes, suggesting that this region was not responsible for isoprene metabolism by known or predicted pathways. Interestingly, although they could not be identified as directly involved in isoprene degradation, these divergent genes on scaffold MG_720 of genome bin 197 were transcribed at moderately high levels (Fig. 3 and SI Appendix, Table S5). Genomic bin 095 contained a sequence with homology to a putative toluene monoxygenase from Myxococales bacterium 68–20 (OY25088.1), with a similar gene layout. Again, although this genome contained no identifiable isoprene-related genes apart from the monoxygenase, the isoA homolog was represented in the transcriptome (Fig. 3 and SI Appendix, Table S5). These two examples suggest that isoprene is capable of inducing genes not central to its metabolism, although those monoclonal genes are constitutively expressed at considerable levels, or that they form part of a novel and so-far-undescribed isoprene metabolic pathway.

Genome bin 232, assigned to the betaproteobacterial genus Variorovax, contained an isoA homolog (scaffold MG_478), which was represented by a moderately expressed transcript and aligned most closely with xamoA of the propylene-degrader Xanthobacter autotrophicus Py2 (Fig. 3 and SI Appendix, Table S5). In contrast to characterized isoprene degraders, however, X. autotrophicus Py2 does not use glutathione in its alkene metabolic pathway (49), does not contain any homologs of the isoprene metabolic genes surrounding isoA-F at the xamoA locus (50), and does not grow on isoprene (30). Scaffold MG_478 comprised 76,719 bp of contiguous DNA, allowing examination of the
average amino acid identity (51) was compared between these genomes and 52 Variovorax genome assemblies available in GenBank. The metagenome-derived genome was most similar to Variovorax sp. CF079 (GCA_900101545.1) (82.3%) whereas strain WS11 was most similar to Variovorax sp. B2 (GCA_002891695.1) (88%), and similarity between bin 232 and V. sp. WS11 was 79.5%, suggesting that both are novel species. All Variovorax genomes in GenBank were searched for isoprene-related gene sequences like those present in the Variovorax strains described here, but with negative results, suggesting that much diversity still exists outside the reference sequences.

**Conclusions.** Although isoprene degraders are present in all isoprene-exposed environments tested (6), no studies have identified isoprene degraders residing on the leaves of isoprene-emitting trees at the source of emission. Here, the isoprene degraders retrieved from the poplar phyllosphere were dominated by Rhodococcus, but included other Actinobacteria (Pseudonocardia) and Proteobacteria (Variovorax). Taxonomy alone is insufficient to identify isoprene-degrading bacteria, with extremely closely related strains differing in terms of isoprene-degrading ability, and phylogeny based on 16S rRNA genes is not congruent with isoA-based analyses (30). These data suggest lateral transfer of the isoprene metabolic genes and imply that surveys that rely on 16S rRNA gene analysis are not able to identify isoprene degraders.

We used metagenomics to reconstruct the genome of a Variovorax strain and conventional methods to obtain related isolate Variovorax WS11. Both genomes contained the entire isoprene metabolic gene cluster. Despite being more similar in sequence to alkene monooxygenase from Xanthobacter autotrophicus Py2 than IsoMO from known isoprene degraders (Fig. 3), the presence of other genes unique to isoprene metabolism and the expression of the monooxygenase in a heterologous host proved that these are genuine isoprene metabolic gene clusters. Variovorax are metabolically versatile bacteria capable of degradation of natural products and xenobiotics, frequently plant-associated and with plant-growth-promoting effects (52, 53), and have been identified as part of the core bacterial microbiome of both Arabidopsis and poplar (54, 55).

Interestingly, we also detected a significant level of monoxygenase transcripts, similar or identical to 13C-labeled DNA scaffolds, which have not, so far, been implicated in isoprene degradation and are possibly indicative of novel isoprene metabolic pathways. This study shows that the leaves of an isoprene-emitting tree provide a habitat for taxonomically disparate isoprene degraders and forms a basis for continued development of molecular tools to detect isoprene degraders. This is a prerequisite for quantification of isoprene-related genes and transcripts and comparison of the activity of microbes associated with isoprene-emitting and nonemitting environments (including comparisons between high- and low-isoprene-emitting tree species), and hence to establishing the extent to which the tree microbiome is able to take advantage of, and mitigate the release of, this abundant carbon source and climate-active gas.

**Materials and Methods**

For full details, see SI Appendix.

**Enrichment, Isolation, and Stable Isotope Probing.** Isoprene degraders were enriched from soil or from cells dislodged from leaves by ultrasonication and purified by standard methods. For SIP enrichments, cells were washed from 5 g leaves, resuspended in minimal medium, and supplied with 13C-labeled or unlabeled isoprene. Isoprene consumption was followed by gas chromatography (GC), and cells were harvested when they had consumed ~11 or 6 μmol ml−1 (isoprene C), for the first or second experiment, respectively. Each treatment was carried out in triplicate.

**Nucleic Acid Extraction and Purification.** DNA and RNA were extracted using standard methods. Total RNA was depleted of rRNA using Ribo-Zero (Illunima). Labeled and unlabeled DNA were separated by density gradient ultracentrifugation and fractionation as described previously (29). Fractions containing labeled (“heavy”) and unlabeled (“light”) DNA were identified based on the data presented in SI Appendix, Fig. S2.

**Sequencing of Nucleic Acids.** Communities profiled by 16S rRNA gene amplicons were sequenced using Illumina MiSeq and analyzed using Qiime (56). Metatomic libraries were sequenced using Illumina HiSeq (SI Appendix, Tables S3 and S4). Quality-filtered reads were taxonomically profiled using Metaphlan2 v2.5.0 (47), coassembled using IDBA-UD v1.1.1 (57) (SI Appendix, Table S3), binned using MaxBin v2.2 (58), and quality-checked and refined using CheckM v1.0.5 (59) and RefineM v0.0.23 (60). Filtered transcript reads were de novo assembled using Trinity v2.3.2 (61) (SI Appendix, Table S4). Local Blast databases were searched using tblastn v2.2.28 (62). Reads were mapped to assembled transcripts and quantified using kallisto v0.43.1 (63). Normalized expression levels, as transcripts per million (TPM), of each transcript were ranked for each sample as a percentile.

**Expression of IsoMO.** The IsoMO genes isoABCD of metagenome bin 232 were PCR-amplified from the pooled heavy fractions of DNA-SIP enrichments and expressed from a vector in a strain of Rhodococcus (R. AD45-ID) cured of the megaplasmid that contains the isoprene metabolic genes. Isoprene uptake of IsoMO-expressing cell suspensions was quantified by GC.

**Accession Numbers.** The genomes of Variovorax sp. WS11 and Rhodococcus AD45-ID have been deposited at DNA Database of Japan/European Nucleotide Archive/GenBank under accessions PXZZ00000000 and PYHL00000000. Versions described here are PXZZ01000000 and PYHL01000000, respectively.

---

Crombie et al.
Sequence reads have been deposited at GenBank Sequence Read Archive (accession no. SRP101805) (SI Appendix, Tables S2–S4).

ACKNOWLEDGMENTS. Plasmid pTipQC1 was a gift from Tomohiro Tamura. This work was funded by the Earth and Life Systems Alliance at the University of East Anglia; Natural Environment Research Council (NERC) Grants NE/J009725/1 (to J.C.M.) and NE/J009555/1 (to T.J.M.); European Research Council Advanced Grant 694578—IoMet (to J.C.M.); and NERC Fellowship NE/L010771/1 (to J.P.). We acknowledge receipt of a Colciencias Colombian Government Scholarship (to N.L.L.-M.).

1. Atkinson R, Arey J (2003) Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. Atmos Environ 37(Suppl 2):197–219.
2. Guenther AB, et al. (2012) The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. Geosci Model Dev 5:1471–1492.
3. Shaw SL, Gantt B, Mleshkhide N (2010) Production and emissions of marine iso-mono and monoterpenes: A review. Adv Meteorol 2010:1–24.
4. Dani SKG, et al. (2017) Relationship between isoprene emission and photosynthesis in diatoms and its implications for global marine isoprene estimates. Mar Chem 198:17–24.
5. Sanadze GA (2017) Biogenic isoprene emission as expression of dissipativity, a fundamental cell property. Russ J Plant Physiol 63:134–140.
6. Guenther AB, et al. (2012) The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. Geosci Model Dev 5:1471–1492.
7. Sanadze GA (2017) Biogenic isoprene emission as expression of dissipativity, a fundamental cell property.
8. Ashworth K, Wild O, Hewitt CN (2013) Impacts of biofuel cultivation on mortality and morbidity of birds. Environ Health Perspect 121:275–279.
9. Sanadze GA (2017) Biogenic isoprene emission as expression of dissipativity, a fundamental cell property.
10. McGenity TJ, Crombie AT, Murrell JC (2018) Microbial cycling of isoprene, the most abundant volatile organic compound on Earth. ISME J 12:931–941.
11. Carlton AG, Wiedinmyer C, Kroll JH (2009) A review of Secondary Organic Aerosol (SOA) formation from isoprene. Atmos Chem Phys 9:4867–5005.
12. Parniske M, et al. (2017) Tethered balloon measurements of biogenic VOCs in the atmospheric free-tropopause layer. Atmos Chem Phys 17:897–912.
13. Sharkey TD, Wiberley AE, Donohue AR (2008) Isoprene emission from plants: Why and how? Ann Rev Plant Biol 59:492–496.
14. Zeinali N, Altarawneh M, Li D, Al-Dujaili J, Dlugogorski BZ (2016) New mechanistic insights: Why do plants produce isoprene? ACS Omega 1:220–225.
15. Kroll JH, Carlton AG, Wiedinmyer C (2010) A review of the formation and implications of isoprene-emitting bacteria. Front Microbiol 1:220–225.
16. Kesselmeier J, Staudt M (1999) Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. J Atmos Chem 33:23–88.
17. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
18. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
19. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
20. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
21. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
22. Srivastva N, Singh A, Bhardwaj Y, Dubey SK (2018) Biotechnological potential for isoprene degradation. Biotechnol Adv 36:982–999.
23. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
24. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
25. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
26. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
27. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
28. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
29. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
30. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
31. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
32. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
33. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
34. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
35. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
36. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
37. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
38. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
39. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
40. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.
41. Cleveland CC, Yavitt JB (1998) Microbial consumption of atmospheric isoprene in a forest. J Geophys Res Atmos 103:D19:9159–9165.
42. Cleveland CC, Yavitt JB (1997) Consumption of atmospheric isoprene in soil. J Geophys Res Atmos 102:D19:9159–9165.