

# Drought re-routes soil microbial carbon metabolism towards emission of volatile metabolites in an artificial tropical rainforest

Received: 8 July 2022

Accepted: 19 June 2023

Published online: 31 July 2023

 Check for updates

Linnea K. Honeker<sup>1,2</sup>, Giovanni Pugliese<sup>3,4</sup>, Johannes Ingrisich<sup>3,5</sup>, Jane Fudyma<sup>6,12</sup>, Juliana Gil-Loaiza<sup>2</sup>, Elizabeth Carpenter<sup>7</sup>, Esther Singer<sup>7</sup>, Gina Hildebrand<sup>6</sup>, Lingling Shi<sup>8</sup>, David W. Hoyt<sup>9</sup>, Rosalie K. Chu<sup>9</sup>, Jason Toyoda<sup>9</sup>, Jordan E. Krechmer<sup>10,13</sup>, Megan S. Claflin<sup>10</sup>, Christian Ayala-Ortiz<sup>6</sup>, Viviana Freire-Zapata<sup>6</sup>, Eva Y. Pfannerstill<sup>4,14</sup>, L. Erik Daber<sup>3</sup>, Kathiravan Meeran<sup>5</sup>, Michaela A. Dippold<sup>8</sup>, Jürgen Kreuzwieser<sup>3</sup>, Jonathan Williams<sup>4</sup>, S. Nemiah Ladd<sup>3,11</sup>, Christiane Werner<sup>3</sup>, Malak M. Tfaily<sup>6</sup> & Laura K. Meredith<sup>1,2</sup>✉

Drought impacts on microbial activity can alter soil carbon fate and lead to the loss of stored carbon to the atmosphere as CO<sub>2</sub> and volatile organic compounds (VOCs). Here we examined drought impacts on carbon allocation by soil microbes in the Biosphere 2 artificial tropical rainforest by tracking <sup>13</sup>C from position-specific <sup>13</sup>C-pyruvate into CO<sub>2</sub> and VOCs in parallel with multi-omics. During drought, efflux of <sup>13</sup>C-enriched acetate, acetone and C<sub>4</sub>H<sub>6</sub>O<sub>2</sub> (diacetyl) increased. These changes represent increased production and buildup of intermediate metabolites driven by decreased carbon cycling efficiency. Simultaneously, <sup>13</sup>C-CO<sub>2</sub> efflux decreased, driven by a decrease in microbial activity. However, the microbial carbon allocation to energy gain relative to biosynthesis was unchanged, signifying maintained energy demand for biosynthesis of VOCs and other drought-stress-induced pathways. Overall, while carbon loss to the atmosphere via CO<sub>2</sub> decreased during drought, carbon loss via efflux of VOCs increased, indicating microbially induced shifts in soil carbon fate.

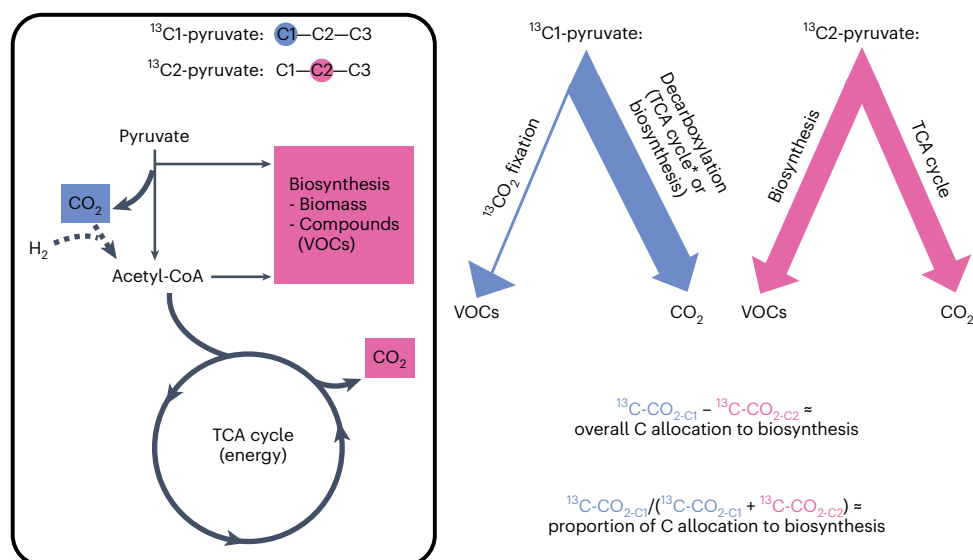
Microorganisms regulate terrestrial carbon (C) cycling in fundamental ways<sup>1</sup>, including transforming soil C into gaseous compounds that can escape to the atmosphere, primarily as CO<sub>2</sub> via microbial heterotrophic respiration. However, microbes also produce volatile organic

compounds (VOCs) as metabolic intermediates, signalling molecules and secondary metabolites<sup>2,3</sup>. In fact, volatile metabolites represent an often overlooked subset of the complete soil metabolome<sup>4,5</sup>, and although their emissions to the atmosphere only represent a small

<sup>1</sup>Biosphere 2, University of Arizona, Tucson, AZ, USA. <sup>2</sup>School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA.

<sup>3</sup>Ecosystem Physiology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany. <sup>4</sup>Max Planck Institute for Chemistry, Atmospheric Chemistry Department, Mainz, Germany. <sup>5</sup>Department of Ecology, Universität Innsbruck, Innsbruck, Austria. <sup>6</sup>Department of Environmental Sciences, University of Arizona, Tucson, AZ, USA. <sup>7</sup>Joint Genome Institute, Walnut Creek, CA, USA. <sup>8</sup>Geo-Biosphere Interactions, Department of Geosciences, University of Tuebingen, Tuebingen, Germany. <sup>9</sup>Environmental Molecular Science Laboratory (EMSL), Earth and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA. <sup>10</sup>Aerodyne Research, Inc., Billerica, MA, USA. <sup>11</sup>Department of Environmental Sciences, University of Basel, Basel, Switzerland. <sup>12</sup>Present address: Department of Plant Pathology, University of California, Davis, CA, USA.

<sup>13</sup>Present address: Bruker Daltonics Inc., Billerica, MA, USA. <sup>14</sup>Present address: Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA. ✉e-mail: [laurameredith@arizona.edu](mailto:laurameredith@arizona.edu)



**Fig. 1 | C Allocation from pyruvate into  $\text{CO}_2$  and VOCs.** Framework depicting how first (C1; blue) and second (C2; pink) carbon (C) positions of pyruvate from  $^{13}\text{C1}$ -pyruvate and  $^{13}\text{C2}$ -pyruvate, respectively, are expected to be allocated into  $\text{CO}_2$  and biosynthetic (including VOCs) pathways. As indicated by the arrow thickness, we expect most C1 to be released as  $\text{CO}_2$  during biosynthesis and TCA cycle, and C2 to be either released as  $\text{CO}_2$  in the TCA cycle or used for biosynthesis of biomass and products. The difference in

$^{13}\text{C-CO}_{2,\text{C1}} - ^{13}\text{C-CO}_{2,\text{C2}}$  approximates the total C allocation to biosynthesis, while the ratio  $^{13}\text{C-CO}_{2,\text{C1}} / (^{13}\text{C-CO}_{2,\text{C1}} + ^{13}\text{C-CO}_{2,\text{C2}})$  approximates the proportion of internal C allocation to biosynthesis. \*Some  $^{13}\text{C}$  from  $^{13}\text{C1}$ -pyruvate may end up in the TCA cycle due to anaplerotic  $\text{CO}_2$  assimilation, which would lead to a slight overestimate in the proportion of C allocation to biosynthesis and overall C allocation to biosynthesis. Figure adapted with permission from refs. 26,27, Elsevier.

soil C loss, they contribute substantially to atmospheric chemistry including ozone formation and cloud condensation nuclei<sup>6</sup>. Therefore, characterizing microbe-mediated C flow along the soil–atmosphere continuum is critical for understanding the fate of soil C and VOCs under projected environmental changes, including drought.

Drought stress induces well-characterized microbial physiological responses that impact C metabolism, such as biosynthesis of protective molecules (for example, osmolytes and extracellular polymeric substances) to preserve cell integrity<sup>7,8</sup> and concentrate resources<sup>9,10</sup>. Production of these biomolecules is energy intensive and may divert resources from biomass synthesis<sup>9</sup>, leading to decreased growth-fuelling heterotrophic respiration and  $\text{CO}_2$  emissions<sup>11,12</sup>. Drought also induces changes in soil C composition and availability<sup>13–15</sup>, further impacting microbial activity<sup>16,17</sup>, for example, by inducing shifts in substrate utilization<sup>18</sup>. Overall, it remains unclear how drought-induced shifts in microbial metabolism and soil C composition influence C allocation to volatile metabolites, which can mitigate drought stress in plants<sup>19</sup>. Moreover, soil water content has a strong impact on VOC emissions from soils<sup>20</sup>, including tropical soils<sup>21,22</sup>, perhaps driven by drought impacts on microbial VOC biosynthesis and/or consumption. Characterizing changes in microbial C cycling and allocation is particularly important in tropical rainforest soils where droughts will probably occur more frequently and last longer due to climate change<sup>23,24</sup>.

Detecting shifts in microbial C cycling and allocation within complex metabolic networks encompassing competing production and consumption pathways is challenging. This complexity can be navigated by tracking microbe-mediated C flow through soils using isotopically labelled central metabolites. Position-specific  $^{13}\text{C}$ -glucose and/or  $^{13}\text{C}$ -pyruvate labelling has been used to track microbial C allocation to  $\text{CO}_2$  and biomass in soil mesocosms<sup>25–27</sup>. However, studies on drought in the field and allocation to VOCs are missing<sup>28</sup>. The direct metabolic information derived from isotope labelling can be contextualized using powerful, information-rich constraints provided by ‘omics approaches that profile the gene content, gene expression and

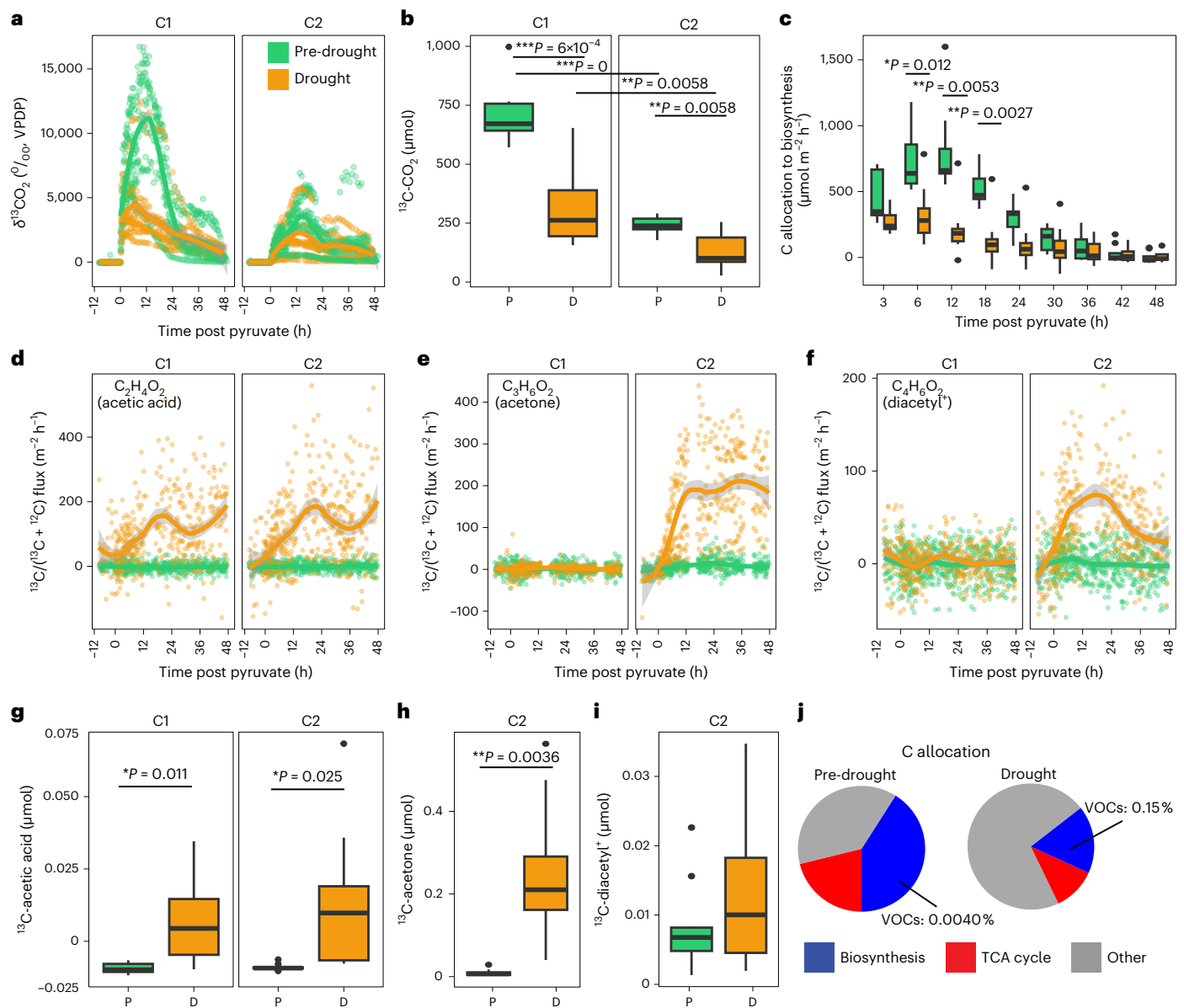
metabolomes of soil microbiomes<sup>26</sup>. Together, these approaches can uncover metabolic drivers of shifting microbial C cycling and allocation in soils under drought.

In this study, we performed a comprehensive assessment of how drought impacts soil microbial C allocation to  $\text{CO}_2$  and VOCs in an artificial enclosed tropical rainforest at Biosphere 2 using an integrated approach, combining position-specific  $^{13}\text{C}$ -pyruvate labelling with metatranscriptomics, metagenomics and metabolomics. Biosphere 2 (Oracle, Arizona) is a 12,700 m<sup>2</sup> steel and glass-enclosed building harbouring five distinct biomes, including a 1,900 m<sup>2</sup> tropical rainforest—an enclosed ecosystem with variable topography and microhabitats where rain and temperature can be controlled. In 2019, a 65-day drought was imposed on the tropical rainforest to study ecosystem-scale impacts of dry-down conditions<sup>29</sup>. With the assumption that the C1 position of pyruvate would enter  $\text{CO}_2$  via decarboxylation during biosynthesis or respiration (tricarboxylic acid (TCA) cycle), and the C2 position would enter VOCs (as biosynthesis) or  $\text{CO}_2$  during respiration<sup>26,27</sup>, we aimed to infer microbial C allocation to biosynthesis vs TCA cycle during both ambient and drought conditions. We hypothesized that during drought: (1) microbial C allocation to the TCA cycle would decrease as C allocation is diverted to increased biosynthesis of stress compounds, including VOCs and (2) changes in C allocation would be associated with shifts in metabolic composition and gene expression.

## Results

### Gaseous emissions from soil

We tracked  $^{13}\text{C}$  from the first and second positions of pyruvate ( $^{13}\text{C1}$ -pyruvate and  $^{13}\text{C2}$ -pyruvate, respectively) into  $\text{CO}_2$  and VOCs (Fig. 1) within soil chambers located among three sites across the tropical rainforest at Biosphere 2 (Extended Data Fig. 1). During drought, total soil  $\text{CO}_2$  efflux and its  $^{13}\text{C}$ -enrichment ( $\delta^{13}\text{C}_{\text{CO}_2}$ ) decreased after injection of  $^{13}\text{C1}$ -pyruvate and  $^{13}\text{C2}$ -pyruvate relative to pre-drought, with higher  $\delta^{13}\text{C}_{\text{CO}_2}$  from  $^{13}\text{C1}$ -pyruvate compared with  $^{13}\text{C2}$ -pyruvate chambers during both pre-drought and drought (Fig. 2a and Extended Data Fig. 2a). As expected, this pattern was reflected in the cumulative



**Fig. 2 | Drought induces a shift in C allocation, as measured using position-specific  $^{13}\text{C}$ -pyruvate labelling.** **a**,  $\delta^{13}\text{C}$  of  $\text{CO}_2$  efflux over time from 12 h pre to 48 h post pyruvate injection from both  $^{13}\text{C}1$ - and  $^{13}\text{C}2$ -pyruvate. **b**, Cumulative  $^{13}\text{C}$ - $\text{CO}_2$  soil efflux from 0 to 48 h post pyruvate injection. **c**, Overall C allocation to biosynthesis (calculated as  $^{13}\text{C}$ - $\text{CO}_{2\text{C}1}$  -  $^{13}\text{C}$ - $\text{CO}_{2\text{C}2}$ ). Continuous emission data were binned to 3 or 6 h intervals. **d–f**, Fluxes of  $^{13}\text{C}$ -enriched ( $^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$ ) acetic acid (**d**), acetone (**e**) and  $\text{C}_4\text{H}_6\text{O}_2$  (diacetyl $^+$ ) (**f**) from 0 to 48 h post pyruvate injection. **g–i**, Cumulative flux of  $^{13}\text{C}$ -acetic acid (**g**),  $^{13}\text{C}$ -acetone (**h**) and  $^{13}\text{C}$ -diacetyl $^+$  (**i**)  $^{13}\text{C}$ -diacetyl $^+$  from chambers that received  $^{13}\text{C}1$ - and/or  $^{13}\text{C}2$ -pyruvate, if  $^{13}\text{C}$ -enrichment was evident (see panels **d–f**). **j**, Pie charts of the percentage of  $^{13}\text{C}$  from pyruvate allocated to biosynthesis or the TCA cycle (or other which could represent unmetabolized pyruvate,  $\text{CO}_2$  fixation or uncharacterized pathways) for pre-drought (P)

and drought (D). For all time series (**a, d–f**), each point represents a single measurement per soil chamber ( $\sim 64$  measurements), of which there were 3 replicates each that were injected with either  $^{13}\text{C}1$ - or  $^{13}\text{C}2$ -pyruvate for each site (total  $n = 18$  each for pre-drought and drought), and lines show the smoothed data with the surrounding shaded area showing  $\pm$ s.e.m. Boxplots represent Q1–Q3, centre lines indicate the median and whiskers extend to the minimum and maximum values, exclusive of outliers (black points), of cumulative flux (measured per soil chamber) across all sites and timepoints (**b, g–i**) ( $n = 54$  each for pre-drought and drought) or overall C allocation to biosynthesis (measured per set of C1/C2 chambers) across sites and specific timepoints as indicated (**c**) ( $n = 9$  sets each for pre-drought and drought). *P* values in **b, c, g–i** are based on linear mixed effects models.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ .

$^{13}\text{C}$ - $\text{CO}_2$  effluxes (Fig. 2b), which decreased significantly during drought by 54.3% from chambers receiving  $^{13}\text{C}1$ -pyruvate ( $t$ -value =  $-5.85$ , d.f. = 7,  $P = 0$ , linear mixed effect (LME)) and 47.6% from chambers receiving  $^{13}\text{C}2$ -pyruvate ( $t$ -value =  $-3.92$ , d.f. = 7,  $P = 0.0058$ , LME).

Contrary to our hypothesis, microbes maintained their allocation to energy production during drought-induced stress. While overall C allocation to biosynthesis was higher during pre-drought conditions

between 3 and 18 h post pyruvate addition (Fig. 2c), internal partitioning of C to biosynthesis vs TCA cycle did not change (Extended Data Fig. 2b).

To further reveal C allocation strategies, we tracked  $^{13}\text{C}$  from pyruvate into  $^{13}\text{C}$ -enriched ( $^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$ ) volatile compounds and were only able to detect three, all of which are central metabolites: acetone, acetic acid and  $\text{C}_4\text{H}_6\text{O}_2$  (Fig. 2d–f). Cumulative effluxes of  $^{13}\text{C}$ -acetic

acid and  $^{13}\text{C}$ -acetone increased significantly by factors of 2.6 and 25.3, respectively, during drought from chambers receiving  $^{13}\text{C}$ -pyruvate ( $t$ -values = 2.83 and 4.30, respectively; d.f. = 7,  $P < 0.05$ , LME) (Fig. 2g,h), indicating synthesis along metabolic pathways after C1 decarboxylation to  $\text{CO}_2$ . Emissions of acetone also increased during low moisture in a previous study in the tropical rainforest<sup>21</sup>, and these results suggest a role of microbes in the production of acetone and consequent emissions during drought. Acetic acid also showed  $^{13}\text{C}$ -enriched continuous efflux (Fig. 2e,f) from chambers that received  $^{13}\text{C}$ -pyruvate where  $^{13}\text{C}$ -acetic acid cumulative efflux increased significantly by a factor of 1.9 during drought ( $t$ -value = 3.45; DF = 7;  $P < 0.05$ ; LME) (Fig. 2g).  $^{13}\text{C}$ -enrichment of acetic acid from chambers receiving  $^{13}\text{C}$ -pyruvate indicates active acetogenesis, where acetogens fix two  $\text{CO}_2$ , in this case  $^{13}\text{C}$ - $\text{CO}_2$ , and four  $\text{H}_2$  molecules to form acetyl-CoA, which is further reduced to acetate<sup>30</sup> (Fig. 1), an intermediate of central metabolism that can be protonated in soils and volatilized as acetic acid (hereafter referred to as acetate).  $\text{C}_4\text{H}_6\text{O}_2$  probably comprised diacetyl and/or one other unidentifiable compound (henceforth referred to as diacetyl<sup>+</sup>) based on structural identification of compounds with this formula from nearby locations (Extended Data Fig. 3).  $^{13}\text{C}$  may have been allocated to other non-volatile metabolites; however, no such  $^{13}\text{C}$ -enrichment was detected by  $^1\text{H}$ -nuclear magnetic resonance (NMR).

Drought induced a readjustment of overall C allocation. The proportion of  $^{13}\text{C}$  from pyruvate allocated to biosynthesis dropped from 41.0 to 17.3% and to the TCA cycle from 21.1 to 11.1% during drought (Fig. 2j). Despite this overall decrease in C allocation to biosynthesis,  $^{13}\text{C}$ -enriched emissions of acetate, acetone and diacetyl<sup>+</sup> represented a greater proportion of biosynthetic products during drought, increasing from 0.0040% to 0.15% (Fig. 2j). Meanwhile,  $^{13}\text{C}$  from pyruvate that did not track to emitted  $^{13}\text{C}$ - $\text{CO}_2$  ('other'), possibly representing unmetabolized pyruvate and/or autotrophic and heterotrophic  $\text{CO}_2$  fixation, increased from 37.9 to 71.6% during drought (Fig. 2j). Total  $\text{CO}_2$  fixation would be expected to decrease  $\text{CO}_2$  emissions by up to 5.6%, depending on soil type<sup>30,31</sup>, suggesting the potential for only a small impact on our estimates of microbial C allocation to biosynthesis vs TCA cycle. Overall, these data suggest an increase in biosynthesis of VOCs during drought, despite an overall decrease in microbial C allocation to biosynthesis.

### Acetate, acetone and diacetyl<sup>+</sup> cycling gene expression

To characterize active microbial metabolic pathways that could cycle acetate, acetone and diacetyl<sup>+</sup> in soils, we utilized a gene-centric approach using metatranscriptomics and metagenomics data to identify: (1) specific genes that may be driving the  $^{13}\text{C}$ -enriched VOC emissions (see Fig. 3a for possible metabolic pathways leading from pyruvate to acetate, acetone and diacetyl, and Extended Data Fig. 6 for associated gene expression (acetate and acetone only)) and (2) overall gene expression patterns that reflect ecosystem-scale microbial responses to drought. Drought and  $^{13}\text{C}$ -pyruvate injection (that is, comparison between 0 (before injection), 6 and 48 h post-injection timepoints) significantly affected overall gene expression profiles ( $P < 0.05$ , permutational multivariate analysis of variance (PERMANOVA)) (Extended Data Fig. 4a) and taxonomic composition of active microbial communities (Extended Data Fig. 4c–e) but had no impact on microbial function potential (Extended Data Fig. 4b) or taxonomic composition (Extended Data Fig. 4f–h). Only location (site) impacted microbial functional potential ( $P = 0.003$ , PERMANOVA) (Extended Data Fig. 4b). This demonstrates that fluctuating gene expression was driven by changes in microbial activity and not by drastic shifts in community composition. We point out the caveat that these results are based on known genes and pathways within the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, potentially leading to bias against some organisms that are underrepresented in the databases, such as protists<sup>32</sup>.

Both  $^{13}\text{C}$ -enriched acetate and acetone effluxes from chambers that received  $^{13}\text{C}$ -pyruvate (acetate-C2 and acetone-C2, respectively)

were driven by a combination of genes encoding for production and consumption, as well as soil moisture, which together explained 91% of acetate-C2 and 76% of acetone-C2 efflux (partial least square regression (PLSR)) (Table 1). Soil moisture was a major driver of both acetate-C2 and acetone-C2 effluxes (VIP = 1.02 and 1.00,  $w_1 = -0.42$  and  $-0.44$ ; PLSR) (Table 1), indicating that emission of these microbially produced VOCs was partly driven by drought-induced moisture changes. Previous research found that decreased soil moisture is associated with increased movement of volatile compounds through the soil<sup>33</sup>.

Microbial gene expression that drove acetate-C2 efflux primarily included acetate-producing enzymes, three of which were upregulated during drought (Fig. 3b). Expression of *ACH1* (K01067), encoding for an acetate-producing acetyl-CoA hydrolase which was upregulated during drought ( $P = 8.4 \times 10^{-14}$ , differential analysis using DESeq2) (Fig. 3b and Extended Data Fig. 5b), was the largest microbial driver of acetate-C2 efflux (VIP = 1.51;  $w_1 = 0.51$ ; PLSR) (Table 1). *ACH1*, only present in eukaryotes<sup>33–35</sup>, including fungi<sup>33,35</sup>, produces acetate and CoA from acetyl-CoA<sup>36</sup>, and can play a role in acetyl-CoA regulation. The two other largest microbial drivers of acetate-C2 efflux were expression of *poxB* (K00156; encoding pyruvate dehydrogenase (PDH)-quinone) and *poxL* (K00158; encoding pyruvate oxidase) (VIP = 0.90 and 0.92,  $w_1 = 0.17$  and 0.18, respectively; PLSR) (Table 1). The *poxB* gene, first identified in *Escherichia coli*<sup>37,38</sup>, encodes for the protein PDH-quinone, which catalyses the direct conversion of pyruvate to acetate and  $\text{CO}_2$  via oxidative decarboxylation, with quinone or menaquinone as the electron acceptor<sup>37,38</sup>. Quinone then shuttles the electrons to the electron transport chain where  $\text{O}_2$  is the final electron acceptor, thereby producing ATP<sup>39</sup>. The strong upregulation of *poxB* during drought suggests that this gene may improve microbial fitness during drought, perhaps to gain extra energy for biomolecule production, as discussed below.

Microbial gene expression that drove acetone-C2 efflux was limited to three genes: two encoding for acetone-consuming enzymes and one for acetone production. Expression of *adc* (K01574), encoding for the acetone-producing enzyme acetoacetate decarboxylase, was the largest microbial driver of acetone-C2 efflux (VIP = 1.05,  $w_1 = 0.61$ ; PLSR) (Table 1); however, *adc* expression did not change during drought (Fig. 3b). The gene *adc* is part of the acetone-butanol-ethanol (ABE) fermentation pathways in *Clostridium acetybuticum* and related species<sup>40,41</sup>; however, we did not see the expected concurrent  $^{13}\text{C}$ -enriched ethanol or butanol emissions, indicating either the presence of other routes of acetone production, or maintained consumption of ethanol and butanol under drought conditions. The other two genes that drove acetone-C2 efflux were *adh1* (K18382) and *acmA* (K18371), which encode for the acetone-consuming enzymes alcohol dehydrogenase (producing propanol, which was not detectable due to its extreme fragmentation during ionization with proton-transfer-reaction mass spectrometry (PTR-MS))<sup>42</sup> and acetone monooxygenase (producing methyl acetate), respectively (VIP = 0.84 and 0.64, respectively; PLSR) (Table 1). *AcmA* was also downregulated during drought, along with three other genes encoding for acetone-consuming enzymes (*acxA* (K10855), *acxB* (K10854) and *acxC* (K10586)) ( $P < 0.05$ ; DESeq2) (Fig. 3b). While  $\text{C}_3\text{H}_6\text{O}_2$  was detected, possibly indicating methyl acetate, it was not  $^{13}\text{C}$ -enriched. It is possible that methyl acetate was further broken down into methanol and  $^{13}\text{C}$ -acetate via methyl acetate hydrolase (*acm*; K18372); however, *acm* was not a major driver of acetone-C2 or acetate-C2 efflux. Therefore, the overall significant decrease in acetone-C2 efflux was driven by both increased production and decreased consumption during drought, signifying that changes in VOC fluxes from soils depend not only on production but also on internal consumption.

For diacetyl<sup>+</sup>-C2 efflux, gene expression and soil moisture accounted for only 23% of total variation (Table 1). Diacetyl is formed when *ilvB* (K01652), encoding for the enzyme acetolactate synthase, catalyses the conversion of pyruvate to  $\text{CO}_2$  and acetolactate, which





**Table 1 | PLSR model summaries of <sup>13</sup>C-enriched VOCs (acetate, acetone and diacetyl<sup>+</sup>) fluxes after labelling with <sup>13</sup>C<sub>2</sub>-pyruvate**

VOC model	Gene	KO	Produce (p) or consume (c)	Variable (enzyme/VOC efflux/physicochemical)	VIP	PLSR weight	
						w1	w2
Acetate	na	na	na	acetone-C2 efflux	1.51	0.61	0.28
	<i>ACH1</i>	K01067	p	acetyl-CoA hydrolase	1.21	0.48	0.15
	<i>aarC</i>	K18118	c	succinyl-CoA:acetate CoA-transferase	1.05	0.21	-0.23
	na	na	na	soil moisture (%)	1.02	-0.42	-
	<i>poxL</i>	K00158	p	pyruvate oxidase	0.92	0.17	-0.53
	<i>poxB</i>	K00156	p	PDH-quinone	0.90	0.18	-0.50
	<i>acyP</i>	K01512	p	acylphosphatase	0.86	-	0.28
	<i>acs</i>	K01895	c	acetyl-CoA synthase	0.82	-0.21	0.12
	<i>pct</i>	K01026	p	propionate CoA-transferase	0.82	-0.28	-
	<i>ackA</i>	K00925	c	acetate kinase	0.75	-	0.46
<b>R<sup>2</sup>- total 0.91 (6 comp.)</b>						<b>0.51</b>	<b>0.20</b>
Acetone	na	na	na	acetate-C2 efflux	1.36	0.61	-
	<i>adc</i>	K01574	p	acetoacetate decarboxylase	1.05	0.47	-
	na	na	na	soil moisture (%)	1.00	-0.44	-
	<i>adh1</i>	K18382	c	NAD <sup>+</sup> -dependent secondary alcohol dehydrogenase	0.80	0.36	-
	<i>acmA</i>	K18371	c	acetone monooxygenase	0.64	-0.29	-
<b>R<sup>2</sup>- total 0.76 (1 comp.)</b>						<b>0.76</b>	-
Diacetyl <sup>+</sup>	na	na	na	soil moisture (%)	1.12	-	0.99
	<i>butA</i>	K03366	c	meso-butanediol dehydrogenase	0.86	-0.99	-
<b>R<sup>2</sup>- total 0.23 (2 comp.)</b>						<b>0.15</b>	<b>0.08</b>

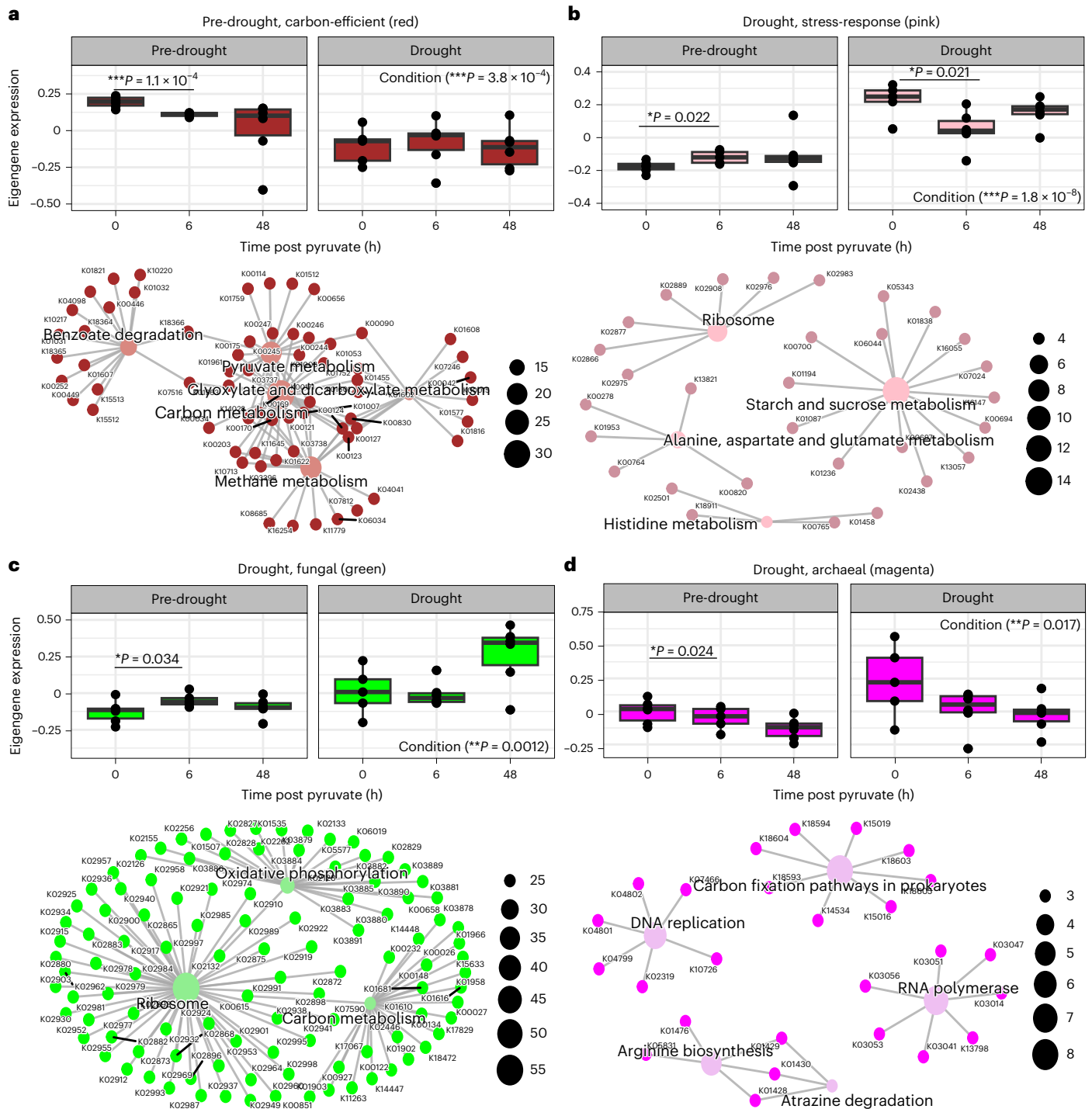
PLSR models were created with single response variables (fluxes for each of the <sup>13</sup>C-enriched VOCs) and predictor variables that included expression of genes known to produce or consume each <sup>13</sup>C-enriched VOC and soil moisture (%). The VIP values represent the degree of importance of each predictor variable on the response variable. The PLSR weights (w) depict the magnitude and direction of contribution of each predictor variable (+ or - relationship) on the response variable along the first and/or second component (w1 and w2), with a dash indicating a lack of contribution to that component. Genes or physicochemical properties that were determined to not contribute to each model are not shown (see Methods). <sup>13</sup>C-enriched VOC fluxes and gene expression data are from three collars that received <sup>13</sup>C<sub>2</sub>-pyruvate at 0, 6 and 48 h post pyruvate injection during pre-drought and drought conditions (n=18). The total percentage of variance explained (R<sup>2</sup>) from predictor variables for all components (comp., shown in parentheses) in the optimized model is shown in the bottom row for each VOC model in bold. The R<sup>2</sup> broken down by the first and second components are shown below w1 and w2, respectively, in bold. Acetone-C2 efflux was included in the model for acetate-C2, and acetate-C2 efflux was included in the model for acetone-C2 due to the high correlation between acetate-C2 and acetone-C2. na, not applicable; +, C<sub>4</sub>H<sub>6</sub>O<sub>2</sub> represents diacetyl and/or one other unidentifiable compound.

Next, we looked at overall shifts in microbial metabolic pathways of co-expressed gene modules using weighted gene co-expression network analysis (WGCNA). We identified a total of nine modules, four of which were associated with condition (pre-drought or drought) and acetate-C2 and acetone-C2 effluxes (Extended Data Fig. 8a,b).

The pre-drought-associated red module was enriched in C-cycling pathways. The association of the red module with pre-drought conditions was demonstrated by its negative correlation with ( $r = -0.6, P = 5.0 \times 10^{-4}$ , Pearson correlation coefficient (PCC)) (Extended Data Fig. 8b) and downregulation during ( $t$ -value =  $-4.24$ , d.f. = 30,  $P = 3.8 \times 10^{-4}$ , LME) (Fig. 4a) drought. The enriched central C metabolic pathways interwoven within the red module included C metabolism, butanoate metabolism (includes acetone and diacetyl cycling), and pyruvate and methane metabolism (both include acetate cycling) (Fig. 4a). These overlapping central C metabolic pathways suggest efficient use of intermediate C compounds produced, including volatile intermediates such as acetone and acetate, and are supported by negative correlations with acetate-C2 and acetone-C2 efflux ( $r = -0.32$  ( $P = 0.090$ , not significant (NS)) and  $-0.47$  ( $P = 0.010$ ), respectively; PCC) (Extended Data Fig. 8b). Therefore, efficient C cycling in the red module led to rapid consumption of metabolic intermediates, preventing losses to the atmosphere under moist conditions.

The drought-associated pink module played a role in osmotic-stress adaptation, quinone production and acetate/acetone biosynthesis. The pink module was positively correlated with ( $r = 0.8, P = 1.0 \times 10^{-7}$ , PCC) (Extended Data Fig. 8b) and upregulated during

( $t$ -value = 7.59, d.f. = 30,  $P < 1.8 \times 10^{-8}$ , LME) (Fig. 4b) drought, and it was also potentially involved in acetate and acetone production, as indicated by its positive correlations with acetate-C2 and acetone-C2 efflux ( $r = 0.29$  ( $P = 0.010$ , NS) and  $r = 0.43$  ( $P = 0.020$ ), PCC) (Extended Data Fig. 8b) and inclusion of the acetate-producing *poxB* and *poxL* genes. Furthermore, the pink module was enriched in starch and sucrose metabolism; alanine, aspartate and glutamine metabolism; and ribosome pathways (Fig. 4b). Biosynthesis of trehalose within the starch and sucrose metabolism pathway corresponds to the increase in trehalose concentrations observed during drought (Extended Data Table 1). At the sub-pathway level, biosynthesis of ubiquinone, a type of quinone, was also enriched, indicating a possible association with osmotic stress. We hypothesize two explanations that connect quinone biosynthesis, *poxB* expression and acetate biosynthesis as stress-response mechanisms during drought. First, a link between the electron transport chain and osmotic regulation is possible. As an immediate microbial response to osmotic stress, microbes actively pump in K<sup>+</sup>, which then promotes trehalose biosynthesis<sup>45</sup>. A supercomplex of H<sup>+</sup>-ATPase and K<sup>+</sup> pumps may form in the cellular membrane as part of the electron transport chain<sup>46</sup>. Therefore, increased production of quinone and transport of electrons to the electron transport chain could be linked with K<sup>+</sup> influx, leading to further osmotic adjustment. Second, due to the high demand for protective biomolecule synthesis (that is, trehalose), there is a high demand for energy production during drought, which may be limited by lower TCA cycle activity (Fig. 2j). To address this need, *poxB* may act as a bacterial PDH bypass route for production of acetyl-CoA



**Fig. 4 | Expression of pre-drought and drought-associated gene modules and their networks of enriched KEGG metabolic pathways. a–d.** Top: eigengene expression at 0, 6 and 48 h post pyruvate addition during pre-drought (0 and 48 h ( $n = 6$ ), 6 h ( $n = 5$ )) and drought (6 and 48 h ( $n = 6$ ), 0 h ( $n = 5$ )) conditions for the subset of modules: carbon-efficient, red (a); stress-response, pink (b); fungal, green (c) and archaeal, magenta (d). Expression values are arbitrary units. Each point represents one sample; boxes represent Q1–Q3 with centre line indicating

median, and bars extend to maximum and minimum values, excluding outliers. Bottom: metabolic network of enriched KEGG pathways within the indicated module. Central nodes represent the pathway and each branch represents the KO group within that pathway.  $P$  values in a–d are from linear mixed effects models between pre-drought and drought (as indicated after ‘Condition’) or 6 or 48 h vs 0 h (as indicated by lines between timepoints);  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ .

(in combination with *acs*)<sup>47</sup>, the precursor to many downstream secondary metabolites<sup>48</sup> (Fig. 3a). While experiments in *Corynebacterium* found that the PDH bypass route was not essential for growth<sup>38</sup>, *poxB* contributes to the aerobic growth efficiency in glucose-limited conditions in *E. coli*<sup>47</sup>. It is, therefore, possible that under times of stress

or within different bacterial species, this route improves fitness by increasing acetyl-CoA production to meet energy demands.

The drought-associated green and magenta modules were enriched in fungal and archaeal metabolic pathways, respectively. Both modules were significantly correlated with ( $r = 0.54$  and  $0.36$ ,

respectively;  $P \leq 0.05$ , PCC) (Extended Data Fig. 8b) and upregulated during ( $t$ -value = 3.94 and 2.26, respectively; d.f. = 30,  $P < 0.05$ , LME) (Fig. 4c,d) drought; however, while the green module was positively correlated with acetate-C2 and acetone-C2 ( $r = 0.62$  and  $0.77$ , respectively;  $P < 0.001$ , PCC) (Extended Data Fig. 8b) and contained the acetate-cycling genes *ACHI* and *adc* indicating its association with acetate and acetone efflux, the magenta module showed no correlation with  $^{13}\text{C}$ -enriched VOC emissions or genes of interest. The abundance of fungal genes within the green module's enriched KEGG pathways indicates a potential role for fungi in the drought response; specifically, the green module was enriched in genes encoding for eukaryotic ribosomal subunits within the ribosome pathway, as well as eukaryotic-specific enzymes (NADH:ubiquinone oxidoreductase and F-type ATPase) within the oxidative phosphorylation pathway (Fig. 4c) (KEGG pathway database<sup>49</sup>). Therefore, the association between the green module and acetate and acetone efflux, as well as fungal-specific genes, suggests a potential role of fungi in acetate and acetone fermentation during drought. While previous studies have found that fungal communities are generally more resistant to stresses, such as drought, compared with bacteria<sup>50–52</sup>, here we did not see a significant increase in total active fungi during drought based on taxonomic classification ( $0.16 \pm 0.7$  and  $1.1 \pm 0.3\%$  in relative abundance during pre-drought and drought conditions, respectively;  $t$ -value = 0.63, d.f. = 136,  $P = 0.52$ , LME); however, specific fungal taxa shifted in activity during drought, with *Ascomycota* increasing on average from 44.3 to 68.6% of the fungal community during drought ( $t$ -value = 3.02, d.f. = 31,  $P = 0.0050$ , LME) (Extended Data Fig. 4e), with *Fusarium* comprising a majority of this phylum particularly during drought at 48 h post pyruvate addition (52%). Specific studies examining fungal production of acetone are very limited; however, acetone production by *Penicillium brevicompactum* has been detected<sup>53</sup>, and *Fusarium* spp. have been found to produce several volatiles, including acetone and acetate<sup>54</sup>. In contrast, the magenta module was enriched in pathways specific to archaea. For example, the enriched DNA replication and RNA polymerase pathways (Fig. 4d) included polymerases specific to archaea. Despite the association between the drought-associated magenta module and drought, there was no concurrent increase in active archaeal abundance during drought ( $t$ -value = 0.87, d.f. = 276,  $P = 0.37$ , LME) (Extended Data Fig. 4d). Collectively, these results suggest that microbial acetate and acetone biosynthesis during drought may be associated more with fungal rather than archaeal metabolic activity, and we suggest further studies to experimentally test this hypothesis.

## Discussion

Here we show how microbially mediated soil C allocation strategies shift from typical (pre-drought) to drought conditions using a combination of  $^{13}\text{C}$ -pyruvate labelling and multi-omics in a unique controlled-drought experiment performed within an enclosed tropical rainforest. We found a distinct subset of emitted volatiles affected by drought that were formed from pyruvate, hence our findings are representative of pyruvate-dependent volatile emissions from forest floor areas that are not covered by understory vegetation. Under typical pre-drought conditions, there was a balance of C allocated to biosynthesis vs energy, and we were able to track  $^{13}\text{C}$  into several volatile compounds, all representing primary metabolic intermediates (acetate, acetone and diacetyl<sup>1</sup>). We hypothesize that most of the C was allocated to the biosynthesis of biomass as well, because the  $^{13}\text{C}$ -enriched volatile intermediates only made up a tiny fraction of the total C allocated to biosynthesis. Therefore, production of these volatiles was matched with immediate recapturing of these molecules during efficient C cycling and utilization. During the stressed conditions imposed by drought, there was a disruption to C allocation pathways. While the balance between C allocated to biosynthesis vs energy remained constant compared to pre-drought conditions, biosynthesis shifted (from presumably biomass) to stress biomolecules, including

trehalose and quinone. Furthermore, metabolic pathway shifts led to higher release of 'leaky' central metabolites (indicated by increased emissions of  $^{13}\text{C}$ -enriched acetate, acetone and diacetyl<sup>1</sup>), enhanced by C allocation to VOC production (that is, acetate) and reduced capacity to consume and recoup VOC-C (that is, acetone)<sup>55</sup>. Notably, this decreased C cycling efficiency during drought could be amplified by increased air-filled pore space, isolating microbial communities to microhabitats within the soil matrix that are cut off from an aqueous flow of non-volatile metabolites and substrates<sup>12</sup>. This signifies that during drought, microbes switch from investing in more-stable carbon pools such as biomass<sup>9</sup> to stress compounds that are prone to be quickly utilized by microbes upon rain rewet<sup>56</sup>. Overall, microbial survival responses to drought shift soil C cycling by interrupting C storage pathways, increasing C allocation to VOCs and loss in efficiency of recouping C from volatile intermediates, thus indicating potential shifts in soil carbon fate.

## Methods

### Biosphere 2 tropical rainforest drought experiment

Biosphere 2 in Oracle, Arizona is a 12,700 m<sup>2</sup> steel and glass-enclosed building harbouring five distinct biomes, including a 1,900 m<sup>2</sup> tropical rainforest—an enclosed ecosystem with variable topography and microhabitats with the ability to control rain and temperature inside, providing an optimal setting for studying drought effects<sup>57–59</sup>. This artificial tropical rainforest harbours approximately 70 species of trees and shrubs forming a canopy and understory, with soils that represent biogeochemical cycles present in natural rainforests<sup>60</sup>. In late 2019, a 65-day drought experiment was conducted as part of the Water, Atmosphere and Life Dynamics (WALD) campaign<sup>29</sup>. During ambient (pre-drought) conditions, rainfall events were simulated by spraying 15,000 l of irrigation water from the top of the Biosphere 2 tropical rainforest at a frequency of 3 d a week, with the last rainfall event before the drought occurring on 7 October 2019. During the drought, ambient temperatures were maintained between 20 and 26.7 °C in the lowland region. For more detailed characterization of the WALD drought experiment, please see ref. 29.

### Position-specific (C1 and C2) $^{13}\text{C}$ -pyruvate labelling

Bulk soil was labelled with position-specific (C1 or C2)  $^{13}\text{C}$ -pyruvate (henceforth referred to as  $^{13}\text{C1}$ -pyruvate and  $^{13}\text{C2}$ -pyruvate, respectively) to track C allocation into CO<sub>2</sub> and VOCs—a method adapted from one used in plants<sup>61</sup>. Similar  $^{13}\text{C}$ -pyruvate labelling was performed in other regions of the tropical rainforest during the WALD campaign, including plant roots<sup>62</sup> and leaves<sup>63</sup>. Three replicates each of  $^{13}\text{C1}$ -pyruvate or  $^{13}\text{C2}$ -pyruvate labelling were performed per site (Site 1, Site 2 and Site 3;  $n = 6$  per site), representing a vast expanse of the lowland region (Extended Data Fig. 1a) within soil chambers (Extended Data Fig. 1b) before (12–19 September) and during (7–19 November) drought. The night before labelling, two automatic chambers were placed onto the corresponding soil collars, which contained no plants but might have had small amounts of fine roots and were covered with a rain-out shelter. Each morning, these collars were labelled at around 10:00 by placing a 5 × 5 cm stencil with 1 × 1 cm openings into one side of each chamber (placed on different sides of the chamber during pre-drought and drought) and adding 100 μl of  $^{13}\text{C1}$ -pyruvate or  $^{13}\text{C2}$ -pyruvate solution (40 mg ml<sup>-1</sup>) (Cambridge Isotopes, CLM-8077-PK and CLM-8849-PK) to each opening to a depth of 1 cm (25 injections), for a total of 100 mg (Extended Data Fig. 1b). The stencil was then removed before soil chamber gas measurements.

Soil moisture and temperature were measured using a portable probe and LabQuest viewer (Vernier). For the pre-drought condition, soil moisture and temperature data were collected on 1 and 9 October and for drought on 11 and 18 November for a subset of collars (P11, P21, P23, P32 and P33 (Site # (P1, P2 or P3), replicate # (1, 2 or 3)), except for 18 November when all collars from the experiment were tested



(see Extended Data Fig. 1a). Soil moisture measured near the labelling sites decreased from  $26.0 \pm 6.9$  to  $13.8 \pm 2.6\%$  ( $P < 0.001$ ) between pre-drought and drought conditions, respectively. Soil temperature did not change significantly and ranged from  $23.0 \pm 0.6$  during pre-drought to  $23.3 \pm 1.2$  °C during drought.

### Continuous monitoring of VOCs and CO<sub>2</sub>

Before labelling, collars were measured at typical temporal resolution (~2 h) overnight. To capture any rapid changes in gas fluxes due to changes in microbial activity after pyruvate addition, measurement intervals were increased to high frequency (30 min) directly before labelling. After gas fluxes were expected to equilibrate, approximately 8 h post pyruvate labelling (-18:00), measurement intervals were decreased to low frequency (50 min) and remained at this frequency until measurements were stopped at 48 h post labelling, resulting in ~64 measurements for each soil collar. For each measurement, the automatic chamber closed over the collar for a total of 10 min (pre-purge, 2.5 min; measurements, 6.5 min; post-purge, 1 min). Fluxes were measured using an automated multiplexed Licor soil flux system (Licor 8100, Li-8150 16-port multiplexer and Lic 8100-104 long-term chambers with opaque lids, Licor). The system was coupled to a Picarro G2201-i analyser (Picarro) to measure CO<sub>2</sub> and isotopic composition and a proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF 8000, Ionicon) for VOCs (including <sup>13</sup>C-VOCs). The PTR-TOF-MS sampled the sub-flow from the soil flux system at 30 standard cubic centimeters per minute (scm). Perfluoroalkoxy tubing was used for the soil flux system, for the sub-sampling line and for the PTR inlet, with the aim to minimize the release and the retention of the VOCs from and on the tubing surface<sup>64</sup>. The PTR settings were as follows: inlet temperature was 60 °C, drift voltage was 600 V, drift temperature was 60 °C and drift pressure was 2.2 mbar, resulting in an  $E/N$  ( $E$  being the electric field strength and  $N$  the sample gas number density) ratio of 137 Td (Townsend). The time resolution was 10 ms and the mass range was up to 500 AMU. The PTR-TOF was operated in the H<sub>3</sub>O<sup>+</sup> mode, therefore only compounds having proton affinity higher than water ( $697 \text{ kJ mol}^{-1}$ ) underwent proton-transfer reactions and could be detected on their protonated mass to charge ratio ( $m/z$ ), which includes most VOCs. Measured ions were attributed to chemical formulae and specific chemical species based on the exact protonated  $m/z$ . PTR-TOF data were processed using the software PTRwid<sup>65</sup>. To account for possible variations in reagent ion signals, measured ion intensities were normalized to the H<sub>3</sub>O<sup>+</sup> counts in combination with the water-cluster ion counts<sup>66</sup>. At midnight, automatic calibrations were performed using standard gas cylinders containing different multi-VOC component calibration mixtures in Ultra-High Purity (UHP) nitrogen (Apel-Riemer Environmental). For a detailed description of the calibration setup, see ref. 29. The concentrations of compounds included in the standard were calculated with an uncertainty of  $\leq 23\%$ . Concentrations of compounds not included in the calibration standard cylinders were calculated by applying the kinetic theory of proton-transfer reaction<sup>67,68</sup> with an uncertainty of  $\leq 50\%$ .

Select additional soil experiments were conducted with a Vocus PTR-TOF (TOFWERK)<sup>69</sup> coupled to a custom-built gas chromatograph (GC) (Aerodyne Research)<sup>70</sup>. The GC contained an integrated two-stage adsorbent-based thermal desorption preconcentration system for in situ collection of VOCs before separation on the chromatographic column. For preconcentration, a multibed (Tenax TA/Graphitized Carbon/Carboxen 1000, Markes International) sorbent tube was used for the first stage of sample collection; this tube was then subjected to a post-collection water purge before the sample was thermally desorbed to a multibed focusing trap (Tenax, Carbopack X, Carboxen 1003, Markes International) before injection onto the GC column. For this study, the GC was equipped with a 30-m Rxi-624 column (Restek, 0.25 mm i.d., 1.4 µm film thickness) which resolves non- to mid-polarity VOCs in the C<sub>5</sub>-C<sub>12</sub> volatility range before ionization in the PTR detector. The GC-PTR sampled from in situ soil gas probes<sup>71</sup> on an alternating

timed schedule. The GC-PTR can speciate structural isomers and help identify some unknowns by matching to calibrated retention times.

### CO<sub>2</sub> data analysis

CO<sub>2</sub> fluxes and their isotopic composition were calculated on the basis of data from the Picarro isotope analyser. Fluxes were calculated with linear and exponential models fitted to each individual chamber measurement. A deadband of 30 s was used for each measurement to allow for mixing in the chamber. The linear models were calculated on the basis of the first 120 s and for the exponential model, the full closure period of 6.5 min was used. Fluxes were quality controlled visually. Fluxes based on the exponential fit were used preferentially but were replaced by the linear-fit flux where necessary. The isotopic composition was calculated on the basis of the individual efflux rates of the <sup>12</sup>C-CO<sub>2</sub> and <sup>13</sup>C-CO<sub>2</sub> isotopologues. These are reported separately by the gas analyser and linear fits based on the first 120 s were used to calculate efflux rates. Due to the high enrichment of the labelled soil CO<sub>2</sub> efflux, this method was found to be more reliable compared with conventional Keeling-plot approaches. The isotopic composition of the CO<sub>2</sub> efflux was then calculated from the ratios of <sup>12</sup>C-CO<sub>2</sub> to <sup>13</sup>C-CO<sub>2</sub> efflux rates and normalized to the Vienna Pee Dee Belemnite (VPDB) scale ( $\delta^{13}\text{C}_{\text{CO}_2} = ((^{13}\text{C}/^{12}\text{C})_{\text{CO}_2}) / ((^{13}\text{C}/^{12}\text{C})_{\text{VPDB}}) - 1$ ). C isotope fluxes were quality controlled visually for each individual chamber and outliers removed manually.

<sup>13</sup>C-CO<sub>2</sub> emitted from chambers that received <sup>13</sup>C1-pyruvate is formed as C is decarboxylated via pyruvate dehydrogenase (PDH) to form acetyl-CoA or via an alternate decarboxylation reaction leading to biosynthesis of compounds (for example, biomass, secondary metabolites, VOCs), while <sup>13</sup>C-CO<sub>2</sub> emitted from chambers that received <sup>13</sup>C2-pyruvate is primarily formed during decarboxylation in the TCA cycle during energy production (Fig. 1) (modified from refs. 26,27). Using this concept, we can approximate both total and relative (proportion of total) C allocated to biosynthesis by calculating  $^{13}\text{C-CO}_2_{\text{C1}} - ^{13}\text{C-CO}_2_{\text{C2}}$  and  $^{13}\text{C-CO}_2_{\text{C1}} / (^{13}\text{C-CO}_2_{\text{C1}} + ^{13}\text{C-CO}_2_{\text{C2}})$ , respectively, where  $^{13}\text{C-CO}_2_{\text{C1}}$  and  $^{13}\text{C-CO}_2_{\text{C2}}$  are efflux from chambers that received <sup>13</sup>C1-pyruvate or <sup>13</sup>C2-pyruvate, respectively. Within each site, there were three sets of 'C1' and 'C2' chambers located next to each other (Extended Data Fig. 1), and the calculations above were made for each set. To facilitate calculation of total and relative C allocation over time, continuous flux data were binned and averaged across 0–3 (3 h), 3–6 (6 h), 6–12 (12 h), 12–18 (18 h), 18–24 (24 h), 24–30 (30 h), 30–36 (36 h), 36–42 (42 h) and 42–48 (48 h). These calculations are approximations due to two reasons. First, heterotrophic and autotrophic CO<sub>2</sub> fixation, including anaplerotic reactions<sup>31,72</sup>, leads to <sup>13</sup>C-CO<sub>2</sub> being immediately re-consumed after production, causing an underestimate of actual <sup>13</sup>C-CO<sub>2</sub> emissions by up to 5.6% (based on soil CO<sub>2</sub> fixation rates calculated from a forest soil<sup>31</sup>). Second, due to anaplerotic CO<sub>2</sub> assimilation, some <sup>13</sup>C from <sup>13</sup>C1-pyruvate may have entered the TCA cycle and been released as <sup>13</sup>C-CO<sub>2</sub>, causing an overestimate of biosynthesis. However, assuming constant CO<sub>2</sub> fixation routes in pre-drought and drought conditions, this would not have impacted the conclusions we made from our calculations.

### VOC data analysis

The isotopic composition of VOC flux rates was calculated by applying the linear model to calculate the rate of change in the fractional abundance of <sup>13</sup>C ( $^{13}\text{C-VOC} / (^{13}\text{C-VOC} + ^{12}\text{C-VOC})$ ). For each 6.5 min chamber measurement, a deadband of 30 s was used to allow for mixing in the chamber and the linear model was applied to the successive 10 data points collected at 10 s intervals. To identify pyruvate metabolism pathways, we focused on those VOCs that showed <sup>13</sup>C isotopic enrichment in their emissions. To help clarify metabolic pathways and direction of reactions, we also considered the fluxes of VOCs independent of their <sup>13</sup>C-enrichment and metabolite concentrations immediately up- or downstream of the <sup>13</sup>C-enriched VOCs (Fig. 3a). For a complete

analysis of all VOCs that showed changes in flux patterns across the drought experiment, see ref. 73.

### Cumulative effluxes and distribution of $^{13}\text{C}$ from pyruvate

Cumulative  $^{13}\text{C}$ - $\text{CO}_2$  and  $^{13}\text{C}$ -VOC effluxes from 0 to 48 h post  $^{13}\text{C}$ -pyruvate injection were calculated using the area under the curve 'auc' function within the FLUX package (v.0.3)<sup>74,75</sup> in R on the basis of the total amount of  $^{13}\text{C}$ -pyruvate added, we could determine what percentage of total  $^{13}\text{C}$  ended up in  $\text{CO}_2$  or VOCs.

### Soil sample collection and processing

For metagenomics, metatranscriptomics and metabolomics, soil samples were collected just before  $^{13}\text{C}$ -pyruvate labelling (0 h) then at 6 and 48 h post labelling. For 0 h collection, soil (~8 g) was collected directly outside of the stencil using a 2.25 cm diameter sterile metal ring pushed into the soil to 2 cm depth and placed into a sterile 50 ml tube. For 6 and 48 h collections, soil samples were collected using the same method as for 0 h, but inside the metal frame where pyruvate labelling occurred. After each soil collection, samples were immediately brought to the lab and allocated for different downstream analyses: 1 g stored at  $-20^\circ\text{C}$  for metabolomics and 2 g in Lifeguard soil preservation solution (Qiagen, 12868–1000) stored at  $-80^\circ\text{C}$  for DNA/RNA extractions.

RNA and DNA were co-extracted from 1 g of soil using the RNeasy Powersoil Total RNA kit (Qiagen, 12866–25) coupled with the RNeasy Powersoil DNA Elution kit (Qiagen, 12867–25) following the manufacturer's protocol and eluted in 100  $\mu\text{l}$  of kit-provided elution buffer. RNA and DNA concentrations and quality were measured using a Qubit 4 fluorometer (Thermo Fisher) and NanoDrop spectrophotometer (Thermo Fisher). RNA was further treated with DNase (DNase Max, Qiagen, 15200–50) to remove any DNA contamination. Total RNA and DNA were sent to the Joint Genome Institute (JGI; Berkeley, California) for library prep and sequencing.

Water extractions were performed on samples for NMR, followed by solid phase extraction (SPE) for FTICR–MS. Water extraction procedures for bulk metabolite characterization were followed<sup>76–78</sup>. Briefly, 1 g of soil and 5 ml of double deionized water were vortexed in a 15 ml centrifuge tube, sonicated for 2 h at  $21^\circ\text{C}$ , then centrifuged for 20 min. One ml of water-extractable organic C (WEOC) was removed and stored at  $-80^\circ\text{C}$  for NMR analysis at the Pacific Northwest National Laboratory (PNNL), while 4 ml of WEOC was saved for downstream SPE.

SPE is an extraction technique used to clean and concentrate samples for the isolation and analysis of organic compounds<sup>79</sup>. Four ml of WEOC were acidified to pH 2 with 1 M HCl, then passed through methanol (MeOH)-preactivated Bond Elut PPL (Priority Pol/Lutant, Agilent, 12255002) barrel cartridges that contain a macroporous styrene-divinylbenzene crosslinked polymer to retain polar organic compounds, with a vacuum set no higher than  $-5$  psi. Next, the cartridges were washed three times with 9 ml 0.01 M HCl to remove impurities, air dried with a pressure air hose for 2–3 min and finally rinsed with 1.5 ml MeOH in a slow dropwise flow rate into 2 ml vials. The eluate was stored at  $-80^\circ\text{C}$  and sent to PNNL for FTICR–MS analysis.

### Metabolomics (NMR and FTICR) measurement and analysis

$^1\text{H}$ -NMR bulk metabolite characterization was performed on WEOC. Samples (180  $\mu\text{l}$ ) were combined with 2,2-dimethyl-2-silapentane-5-sulfonate- $d_6$  (DSS- $d_6$ ) in  $\text{D}_2\text{O}$  (20  $\mu\text{l}$ , 5 mM) and thoroughly mixed before transfer to 3 mm NMR tubes. Resonances corresponding to  $^{13}\text{C}$  labelling were identified by visual inspection, comparing labelled spectra to unlabelled spectra. Once differences were identified, the molecular location and quantification of  $^{13}\text{C}$  incorporation were determined by J-coupling pattern analysis and the 'linefitting' tool in MestReNova (v.14.2.014)<sup>80</sup>, respectively; however,  $^{13}\text{C}$ -labelled metabolites could not be identified in our samples. NMR spectra were acquired on a

Bruker Advance III spectrometer operating at a field strength of 17.6 T ( $^1\text{H}$   $\nu_0$  of 750.24 MHz) and equipped with a 5 mm Bruker TCI/CP HCN (inverse) cryoprobe with Z-gradient and at a regulated temperature of 298 K. The one-dimensional (1D)  $^1\text{H}$  spectra were acquired using a Nuclear Overhauser Effect Spectroscopy (NOESY) pulse sequence (noesypr1d). The  $90^\circ\text{H}$  pulse was calibrated before the measurement of each sample with a spectral width of 12 ppm and 1,024 transients. The NOESY mixing time was 100 ms and the acquisition time was 4 s, followed by a relaxation delay of 1.5 s during which presaturation of the water signal was applied. Time domain free induction decays (72,114 total points) were zero filled to 131,072 total points before Fourier transform, followed by exponential multiplication (0.3 Hz line-broadening) and semi-automatic multipoint smooth segments baseline correction. Chemical shifts were referenced to the  $^1\text{H}$  methyl or  $^{13}\text{C}$  signal in DSS- $d_6$  at 0 ppm. The 1D  $^1\text{H}$ -NMR spectra of all samples were processed, assigned and analysed using Chenomx NMR suite 9.2 (Chenomx) with quantification of spectral intensities of compounds in the Chenomx library relative to the internal standard. Candidate metabolites present in each of the complex mixtures were determined by matching chemical shift, J-coupling and intensity information of the experimental signals against signals of the standard metabolites in the Chenomx library. Signal to noise ratios (S/N) were measured using MestReNova with the limits of quantification and detection equal to an S/N of 10 and 3, respectively. Standard 2D experiments such as  $^1\text{H}/^{13}\text{C}$ -heteronuclear correlation (HSQC) or 2D  $^1\text{H}/^1\text{H}$  Total Correlation spectroscopy (TOCSY) further aided corroboration of several metabolite identifications where there was sufficient S/N.

A 12-Tesla Bruker FTICR mass spectrometer was used to collect high-resolution mass spectra of SPE-filtered WEOC by direct injection for secondary metabolite characterization. Approximately 100  $\mu\text{l}$  of water extract was mixed with MeOH (1:2) before injection to enhance ionization. Samples were directly injected into a standard Bruker electrospray ionization (ESI) source. The instrument settings were optimized by tuning on a Suwannee River fulvic acid standard (International Humic Substances Society), and the instrument was flushed between samples using a mixture of water and MeOH. Blanks (HPLC grade MeOH) were analysed at the beginning and end of the day to monitor potential carry over from one sample to another. The ion accumulation time varied to account for differences in C concentration between samples. For each sample, 144 individual scans were averaged and internally calibrated using an organic matter homologous series separated by 14 Da ( $\text{CH}_2$  groups). The mass measurement accuracy was  $<1$  ppm for singly charged ions across a broad  $m/z$  range (100–1,000  $m/z$ ). The mass resolution was  $\sim 240$  K at 341  $m/z$ . The transient was 0.8 s. Data Analysis software (Bruker Daltonik v.4.2) was used to convert raw spectra to a list of  $m/z$  values, applying the FTICR–MS peak picker module with an S/N threshold set to 7 and absolute intensity threshold set to the default value of 100. Putative chemical formulae were then assigned using Formularity software<sup>81</sup> on the basis of the following criteria: S/N  $> 7$ , mass measurement error  $< 1$  ppm and taking into consideration the presence of C, H, O, N, S and P and excluding other elements<sup>82</sup>. To ensure consistent formula assignment and eliminate mass shifts that could impact formula assignment, all sample peak lists were aligned to each other. The following rules were implemented to further ensure consistent formula assignment: (1) picking formulae with the lowest error between predicted and observed  $m/z$  and the lowest number of heteroatoms and (2) requiring the presence of at least four oxygen atoms for the assignment of one phosphorus atom<sup>82</sup>. The chemical character of thousands of peaks in each sample's ESI FTICR–MS spectrum was evaluated using van Krevelen diagrams, with biochemical compound classes reported as relative abundance values on the basis of counts of C, H and O as follows: lipids ( $0 < \text{O}:\text{C} \leq 0.3$  and  $1.5 \leq \text{H}:\text{C} \leq 2.5$ ), unsaturated hydrocarbons ( $0 \leq \text{O}:\text{C} \leq 0.125$  and  $0.8 \leq \text{H}:\text{C} < 2.5$ ), proteins ( $0.3 < \text{O}:\text{C} \leq 0.55$  and  $1.5 \leq \text{H}:\text{C} \leq 2.3$ ), amino sugars ( $0.55 < \text{O}:\text{C} \leq 0.7$  and  $1.5 \leq \text{H}:\text{C} \leq 2.2$ ), lignin ( $0.125 < \text{O}:\text{C} \leq 0.65$  and  $0.8 \leq \text{H}:\text{C} < 1.5$ ),

tannins ( $0.65 < \text{O:C} \leq 1.1$  and  $0.8 \leq \text{H:C} < 1.5$ ) and condensed hydrocarbons (aromatics;  $0 \leq \text{O:C} \leq 0.95$  and  $0.2 \leq \text{H:C} < 0.8$ )<sup>78</sup>. Analysis of FTICR data was performed using MetaboDirect (v.0.2.7)<sup>83</sup> to create profiles of biochemical compound classes and principal component analysis (PCA) plots.

### Metagenomics and metatranscriptomics sequencing and analysis

Metagenome and metatranscriptome libraries were created at the JGI. Plate-based DNA library preparation for Illumina sequencing was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using the Kapa-HyperPrep library preparation kit (Kapa Biosystems). Sample genomic DNA (200 ng) was sheared to 300–500 bp using a Covaris LE220 focused-ultrasonicator. The sheared DNA fragments were size selected by double-SPRI, then the selected fragments were end-repaired, A-tailed and ligated with Illumina-compatible sequencing adaptors (from IDT) containing a unique molecular index barcode for each sample library.

For metatranscriptome libraries, ribosomal (r)RNA was removed from 100 ng of total RNA using Qiagen FastSelect 5S/16S/23S for bacterial rRNA depletion (and additional FastSelect plant and/or yeast rRNA depletion) (Qiagen) with RNA blocking oligo technology. The fragmented and rRNA-depleted RNA was reverse transcribed to create first-strand complementary (c)DNA using Illumina TruSeq Stranded mRNA Library prep kit (Illumina), followed by the second-strand cDNA synthesis which incorporated dUTP to quench the second strand during amplification. The double-stranded cDNA fragments were then A-tailed and ligated to JGI dual indexed Y-adapters, followed by an enrichment of the library by 10 cycles of PCR. Metatranscriptomics of sample P35SSC1\_191108\_c were not completed due to library prep and sequencing issues.

The prepared libraries were quantified using KAPA Biosystems' next-generation sequencing library qPCR kit and run on a Roche Light-Cycler 480 real-time PCR instrument. Sequencing of the flowcell was performed on the Illumina NovaSeq sequencer using NovaSeq XP V1.5 reagent kits and S4 flowcell, following a  $2 \times 151$  indexed run recipe.

Sequence filtering, assembly, mapping and annotation were performed at JGI<sup>84</sup>. BBDuk (v.38.94), included in the BBtools package<sup>85</sup>, was used to trim reads that contained adapter sequences, homopolymers of G's of the size 5 or more at the ends of reads and reads where quality dropped to 0. BBDuk was also used to remove reads that contained 1 or more 'N' bases, had an average quality score across the read of <10, or had a minimum length of  $\leq 51$  bp or 33% of the full read length. Reads were mapped with BMap (v.38.44), included in the BBtools package, to masked human, cat, dog and mouse references at 93% identity, and common microbial contaminants were removed for downstream analysis. Filtered reads were error corrected before assembly using bbcms (v.38.90), included in the BBtools package and assembled using metaSPAdes (v.3.15.2)<sup>86</sup> with a minimum contig of 200 bp (for metatranscriptome libraries, there was an additional removal of rRNA during filtering and assembly was performed using MegaHit (v.1.2.9)<sup>87</sup>). Filtered reads were then mapped back to contigs using BMap to obtain coverage information. See Extended Data Table 2 for sequencing depth and mapping statistics. Feature prediction was next performed on assembled contigs by using tRNAscan-SE (v.2.0.6)<sup>88</sup> to predict transfer RNAs, INFERNAL (v.1.1.3)<sup>89</sup> to identify non-coding RNAs and rRNAs, CRT-CLI<sup>90</sup> to identify CRISPR regions, and Prodigal (v.2.6.3)<sup>91</sup> and GeneMarkS-2 (v.1.07)<sup>92</sup> to identify protein-coding genes (CDSs). Functional annotation was performed on CDSs with KEGG Orthology (KO) terms<sup>93</sup>. Taxonomy for each CDS was determined using the best LAST<sup>94</sup> hits of CDSs, then contigs were classified on the basis of the majority rule (the lowest taxonomic rank that at least 50% of CDSs on the contig matched to<sup>84</sup>). Gene copies per KO were calculated as the average coverage of the contigs each gene was predicted from, multiplied by the number of genes in the KO<sup>84</sup>.

Active metabolic pathways were determined by mapping KOs to KEGG pathways using the KEGG pathway mapper tool<sup>93</sup>. VOC cycling genes were chosen on the basis of their being immediately up- or downstream from the VOC in the metabolic pathway, as well as being downstream from pyruvate. Gene symbols used in the paper are based on the KEGG database. Whether the gene was involved in production or consumption of VOC was assessed by examining the reaction on the KEGG database and on the basis of the literature.

### Statistical analysis

All statistical analyses were performed in R (v.4.0.2)<sup>95</sup>. First, to compare <sup>13</sup>C-CO<sub>2</sub> and <sup>13</sup>C-VOC (acetate, acetone and diacetyl<sup>+</sup>) cumulative fluxes across pre-drought and drought conditions, LME modelling was performed using the LME function in the NLME package (v.3.1)<sup>96</sup> with 'Site' and 'ID' (location within each site) included as random variables to control for these confounding factors. Next, LME was used to compare differences in total and relative C allocation to biosynthesis (<sup>13</sup>C-CO<sub>2,C1</sub> - <sup>13</sup>C-CO<sub>2,C2</sub> and <sup>13</sup>C-CO<sub>2,C1</sub> / (<sup>13</sup>C-CO<sub>2,C1</sub> + <sup>13</sup>C-CO<sub>2,C2</sub>), respectively) between pre-drought and drought conditions. The following model equation was used for both of the above LME analyses (full scripts are included in GitHub at [github.com/linneakh/SoilPyruvate](https://github.com/linneakh/SoilPyruvate)):

```
lme ( response ~ Condition,
      random = list ( Site = -1,
                    ID = -1),
      data = data,
      weights = varIndent(form = -1|Condition)
```

WGCNA was performed on metatranscriptomic data to identify modules of co-expressed genes. First, gene copy data were normalized in DESeq2 (v.1.30.1)<sup>97</sup> using the variance stabilization transformation (VST) method and samples were clustered to detect outliers, which removed only one sample (P15SSC1\_191107\_c). Next, the normalized data were analysed using the WGCNA package (v.1.7)<sup>98,99</sup> in R employing the 'blockwiseModules' function with the following settings: soft-thresholding power of 6, minimum module size of 80, minimum KME of 0.35 and a 'signed' topology overlap matrix. Pearson correlation coefficients (*r*) were calculated between module eigengenes and the following: (1) condition (values for pre-drought and drought set to -1 and 1, respectively) and (2) <sup>13</sup>C-enriched VOC fluxes (acetate-C2 and acetone-C2) averaged at 0–2, 5–7 and 46–48 h post pyruvate addition to correspond with metatranscriptome soil collection times of 0, 6 and 48 h, respectively.

To identify expressed genes that were contributing to <sup>13</sup>C-enriched VOC efflux, PLSR analysis was performed using the pls package (v.2.8)<sup>100</sup>. This multivariate statistical test identifies predictive variables that contribute to a response variable while allowing for collinearity between variables. Three PLSR models were created using the <sup>13</sup>C-enriched flux (acetate-C2, acetone-C2 and diacetyl<sup>+</sup>-C2) as a single response variable for each model. To match flux measurements to gene expression data, we averaged VOC flux measurements for 0–2, 5–7 and 46–48 h post pyruvate addition to correspond with metatranscriptome soil collection times of 0, 6 and 48 h, respectively. As predictive variables, the VST-normalized gene copy data were used from collars that received <sup>13</sup>C2-pyruvate (Site2-P4, Site2-P6 and Site3-P4) from across three timepoints post pyruvate injection (0, 6 and 48 h) during pre-drought and drought conditions (*n* = 18). To account for the correlation between acetate-C2 and acetone-C2 at the 0, 6 and 48 h timepoints, which is probably due to the interconnection of the pathways for cycling of these compounds in metabolic pathways, acetate-C2 flux was added as a predictive variable to the acetone model, and acetone-C2 flux was added as a predictive variable to the acetate model. Furthermore, to improve the models, we included soil moisture (%) which likely impacts the movement of



VOCs through the soil and emissions to the atmosphere<sup>8</sup>. The optimal number of components for each model was determined by ‘leave one out’ cross-validation, calculating the root mean squared error in cross-validation (RMSECV) using 0–10 components and finding the number of components that had the lowest RMSECV value. The degree of importance of each predictive variable was assessed using the variable importance in prediction (VIP) values, which were calculated from the loadings, weights and scores of each variable<sup>101</sup> using the *plsVarSel* package (v.0.9.10)<sup>102</sup>. Additionally, to determine the significance of the model, cross-validated predictive residuals were compared to the residuals of the null model (that is, using the mean of the response variable) using an *F*-test. By comparing *F*-values for each model and their corresponding degrees of freedom, a *P* value could be calculated to determine whether the model was significantly different from the null. To determine which predictive variables significantly contributed to each model, the predictor variable with the lowest VIP was removed and a new model was created. This was repeated until the model with the highest % of variation explained ( $R^2 \times 100$ ) was obtained, indicating the predictive variables that were the largest drivers of the response variables. The variables removed from each model were as follows: *acmB* (K18372) for the acetate model; *acmB* (K18372), *acx*C (K10856), *acx*A (K10855) and *acx*B (K10854) for the acetone model; and *ilvB* (K01652) and *BDH* (K0004) for the diacetyl<sup>+</sup> model. The optimal numbers of components for each model were determined as follows: 6 components for acetate ( $R^2 = 0.51, 0.19, 0.16, 0.03, 0.03$  and  $0.002$  (total  $R^2 = 0.91$ )), 1 component for acetone ( $R^2 = 0.76$ ) and 2 components for diacetyl ( $R^2 = 0.15$  and  $0.08$  (total  $R^2 = 0.23$ )), with corresponding RMSECV values of 42.6, 62.5 and 31.6, respectively. This resulted in the following *P* values: 0.50 for acetate, 0.005 for acetone and 0.93 for diacetyl<sup>+</sup> models.

To find differences in metabolomic composition and metagenomic and metatranscriptomic gene copies between pre-drought and drought conditions, PCA was performed using the built-in R ‘*prcomp*’ function on log-transformed NMR data for all metabolites, FTICR data collapsed at the compound class level and VST-transformed normalized gene copy data using the *prcomp* function in base R and plotted with *ggplot2* (v.3.4.1)<sup>103</sup>. LME models were created as described above but with only ‘Site’ as a random variable to detect significant differences in all NMR metabolites and FTICR compounds classes between pre-drought and drought conditions. To determine which genes were up- or downregulated during drought, DESeq2 was used to find the  $\log_2$ FC and associated *P* values (false discovery rate (FDR) corrected) between pre-drought vs drought conditions while controlling for differences between ‘Sites’.

For WGCNA eigengene expression and taxonomic composition, LME models were created with only ‘Site’ as a random variable to determine significant differences between (1) pre-drought vs drought conditions across all timepoints and (2) 6 and 48 vs 0 h within each condition (pre-drought or drought, using ‘Time’ in place of ‘Condition’ in the model). *P* values reported for Pearson correlations between module eigengenes and <sup>13</sup>C-enriched VOC fluxes (acetate-C2 and acetone-C2) were corrected for FDR. Enriched KEGG metabolic pathways within each module were calculated using the ClusterProfiler package (v.3.18.1)<sup>104</sup>, with the KEGG reference database set to ‘ko’.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The metatranscriptomics and metagenomics sequence data are publicly available through Genbank SRA under the following BioProject IDs: PRJNA980752–PRJNA980834. FTICR, NMR, VOC and CO<sub>2</sub> data have been deposited at FigShare (<https://doi.org/10.6084/m9.figshare.20334537>). Individual raw.xml files for the FTICR data and

soil temperature/moisture data are included as Supplementary Data. Source data are provided with this paper.

### Code availability

All code used in data analysis and to create figures is available at [www.github.com/linneahkh/SoilPyruvate](http://www.github.com/linneahkh/SoilPyruvate).

### References

- Schimel, J. P. & Schaeffer, S. M. Microbial control over carbon cycling in soil. *Front. Microbiol.* **3**, 00348 (2012).
- Insam, H. & Seewald, M. S. A. Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils* **46**, 199–213 (2010).
- Penuelas, J. et al. Biogenic volatile emissions from the soil. *Plant Cell Environ.* **37**, 1866–1891 (2014).
- Honeker, L. K., Graves, K. R., Tfaily, M. M., Krechmer, J. E. & Meredith, L. K. The volatilome: a vital piece of the complete soil metabolome. *Front. Environ. Sci.* **9**, 649905 (2021).
- Meredith, L. K. & Tfaily, M. M. Capturing the microbial volatilome: an oft overlooked ‘ome. *Trends Microbiol.* **30**, 622–631 (2022).
- Koppmann, R. in *Hydrocarbons, Oils and Lipids: Diversity, Origin, Chemistry and Fate* (ed. Wilkes, H.) 267–277 (Springer, 2010).
- Warren, C. R. Response of osmolytes in soil to drying and rewetting. *Soil Biol. Biochem.* **70**, 22–32 (2014).
- Schimel, J., Balsler, T. C. & Wallenstein, M. Microbial stress-response physiology and its implications for ecosystem function. *Ecology* **88**, 1386–1394 (2007).
- Malik, A. A. & Bouskill, N. J. Drought impacts on microbial trait distribution and feedback to soil carbon cycling. *Funct. Ecol.* **36**, 1442–1456 (2022).
- Vardharajula, S. & Sk Z, A. Exopolysaccharide production by drought tolerant *Bacillus* spp. and effect on soil aggregation under drought stress. *J. Microbiol. Biotechnol. Food Sci.* **4**, 51–57 (2014).
- Sun, S., Lei, H. & Chang, S. X. Drought differentially affects autotrophic and heterotrophic soil respiration rates and their temperature sensitivity. *Biol. Fertil. Soils* **55**, 275–283 (2019).
- Schimel, J. P. Life in dry soils: effects of drought on soil microbial communities and processes. *Annu. Rev. Ecol. Evol. Syst.* **49**, 409–432 (2018).
- Brown, R. W., Chadwick, D. R., Zang, H. & Jones, D. L. Use of metabolomics to quantify changes in soil microbial function in response to fertiliser nitrogen supply and extreme drought. *Soil Biol. Biochem.* **160**, 108351 (2021).
- Bouskill, N. J. et al. Belowground response to drought in a tropical forest soil. I. Changes in microbial functional potential and metabolism. *Front. Microbiol.* **7**, 525 (2016).
- Bouskill, N. J. et al. Belowground response to drought in a tropical forest soil. II. Change in microbial function impacts carbon composition. *Front. Microbiol.* **7**, 323 (2016).
- Fang, H. et al. Changes in soil heterotrophic respiration, carbon availability, and microbial function in seven forests along a climate gradient. *Ecol. Res.* **29**, 1077–1086 (2014).
- Li, Y. et al. Effects of biochar application in forest ecosystems on soil properties and greenhouse gas emissions: a review. *J. Soils Sediment.* **18**, 546–563 (2018).
- Su, X. et al. Drought changed soil organic carbon composition and bacterial carbon metabolizing patterns in a subtropical evergreen forest. *Sci. Total Environ.* **736**, 139568 (2020).
- Peñuelas, J. & Llusà, J. BVOCs: plant defense against climate warming? *Trends Plant Sci.* **8**, 105–109 (2003).
- Asensio, D., Peñuelas, J., Llusà, J., Ogaya, R. & Filella, I. Interannual and interseasonal soil CO<sub>2</sub> efflux and VOC exchange rates in a Mediterranean holm oak forest in response to experimental drought. *Soil Biol. Biochem.* **39**, 2471–2484 (2007).



21. Boursoukidi, E. et al. Strong sesquiterpene emissions from Amazonian soils. *Nat. Commun.* **9**, 2226 (2018).
22. Jardine, K. et al. Dimethyl sulfide in the Amazon rain forest. *Glob. Biogeochem. Cycles* **29**, 19–32 (2015).
23. Reichstein, M. et al. Climate extremes and the carbon cycle. *Nature* **500**, 287–295 (2013).
24. Jentsch, A. & Beierkuhnlein, C. Research frontiers in climate change: effects of extreme meteorological events on ecosystems. *C. R. Geosci.* **340**, 621–628 (2008).
25. Bore, E. K., Apostel, C., Halicki, S., Kuzyakov, Y. & Dippold, M. A. Soil microorganisms can overcome respiration inhibition by coupling intra- and extracellular metabolism: <sup>13</sup>C metabolic tracing reveals the mechanisms. *ISME J.* **11**, 1423–1433 (2017).
26. Dijkstra, P. et al. Modeling soil metabolic processes using isotopologue pairs of position-specific <sup>13</sup>C-labeled glucose and pyruvate. *Soil Biol. Biochem.* **43**, 1848–1857 (2011).
27. Dijkstra, P. et al. Probing carbon flux patterns through soil microbial metabolic networks using parallel position-specific tracer labeling. *Soil Biol. Biochem.* **43**, 126–132 (2011).
28. Albright, M. B. N. et al. Differences in substrate use linked to divergent carbon flow during litter decomposition. *FEMS Microbiol. Ecol.* **96**, faa135 (2020).
29. Werner, C. et al. Ecosystem fluxes during drought and recovery in an experimental forest. *Science* **374**, 1514–1518 (2021).
30. Drake, H. L., Küsel, K. & Matthies, C. Ecological consequences of the phylogenetic and physiological diversities of acetogens. *Antonie Van Leeuwenhoek* **81**, 203–213 (2002).
31. Akinyede, R., Taubert, M., Schrupf, M., Trumbore, S. & Küsel, K. Rates of dark CO<sub>2</sub> fixation are driven by microbial biomass in a temperate forest soil. *Soil Biol. Biochem.* **150**, 107950 (2020).
32. Miao, W. et al. Protist 10,000 Genomes Project. *Innovation* **1**, 100058 (2020).
33. Chen, Y., Zhang, Y., Siewers, V. & Nielsen, J. Ach1 is involved in shuttling mitochondrial acetyl units for cytosolic C2 provision in *Saccharomyces cerevisiae* lacking pyruvate decarboxylase. *FEMS Yeast Res.* **15**, fov015 (2015).
34. Buu, L.-M., Chen, Y.-C. & Lee, F.-J. S. Functional characterization and localization of acetyl-coA hydrolase, Ach1p, in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 17203–17209 (2003).
35. Fleck, C. B. & Brock, M. Re-characterisation of *Saccharomyces cerevisiae* Ach1p: fungal coA-transferases are involved in acetic acid detoxification. *Fungal Genet. Biol.* **46**, 473–485 (2009).
36. Carman, A. J., Vylkova, S. & Lorenz, M. C. Role of acetyl coenzyme A synthesis and breakdown in alternative carbon source utilization in *Candida albicans*. *Eukaryot. Cell* **7**, 1733–1741 (2008).
37. Williams, F. R., Robert Williams, F. & Hager, L. P. Crystalline flavin pyruvate oxidase from *Escherichia coli*. *Arch. Biochem. Biophys.* **116**, 168–176 (1966).
38. Schreiner, M. E. & Eikmanns, B. J. Pyruvate:quinone oxidoreductase from *Corynebacterium glutamicum*: purification and biochemical characterization. *J. Bacteriol.* **187**, 862–871 (2005).
39. Carter, K. & Gennis, R. B. Reconstitution of the ubiquinone-dependent pyruvate oxidase system of *Escherichia coli* with the cytochrome o terminal oxidase complex. *J. Biol. Chem.* **260**, 10986–10990 (1985).
40. Awang, G. M., Jones, G. A. & Ingledew, W. M. The acetone-butanol-ethanol fermentation. *Crit. Rev. Microbiol.* **15**, S33–S67 (1988).
41. Maddox, I. S. The acetone-butanol-ethanol fermentation: recent progress in technology. *Biotechnol. Genet. Eng. Rev.* **7**, 189–220 (1989).
42. Karl, T., Striednig, M., Graus, M., Hammerle, A. & Wohlfahrt, G. Urban flux measurements reveal a large pool of oxygenated volatile organic compound emissions. *Proc. Natl Acad. Sci. USA* **115**, 1186–1191 (2018).
43. Branen, A. L. & Keenan, T. W. Biosynthesis of α-acetolactate and its conversion to diacetyl and acetoin in cell-free extracts of *Lactobacillus casei*. *Can. J. Microbiol.* **18**, 479–485 (1972).
44. Kakumanu, M. L., Ma, L. & Williams, M. A. Drought-induced soil microbial amino acid and polysaccharide change and their implications for C-N cycles in a climate change world. *Sci. Rep.* **9**, 10968 (2019).
45. Kempf, B. & Bremer, E. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* **170**, 319–330 (1998).
46. Trchounian, A. A. A direct interaction between the H<sup>+</sup>-F<sub>1</sub>F<sub>0</sub>-ATPase and the K<sup>+</sup> transport within the membrane of anaerobically grown bacteria. *Bioelectrochem. Bioenerg.* **33**, 1–10 (1994).
47. Abdel-Hamid, A. M., Attwood, M. M. & Guest, J. R. Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. *Microbiology* **147**, 1483–1498 (2001).
48. Wolfe, A. J. The acetate switch. *Microbiol. Mol. Biol. Rev.* **69**, 12–50 (2005).
49. KEGG Pathway Database (Kanehisa Laboratories.); [www.genome.jp/kegg/pathway.html](http://www.genome.jp/kegg/pathway.html) (accessed July 15, 2022).
50. Barnard, R. L., Osborne, C. A. & Firestone, M. K. Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J.* **7**, 2229–2241 (2013).
51. Waring, B. G. & Hawkes, C. V. Short-term precipitation exclusion alters microbial responses to soil moisture in a wet tropical forest. *Microb. Ecol.* **69**, 843–854 (2015).
52. Sun, Y. et al. Drought stress induced increase of fungi:bacteria ratio in a poplar plantation. *CATENA* **193**, 104607 (2020).
53. Börjesson, T., Stöllman, U. & Schnürer, J. Volatile metabolites produced by six fungal species compared with other indicators of fungal growth on cereal grains. *Appl. Environ. Microbiol.* **58**, 2599–2605 (1992).
54. Robinson, P. M. & Garrett, M. K. Identification of volatile sporostatic factors from cultures of *Fusarium oxysporum*. *Trans. Br. Mycol. Soc.* **52**, 293–299 (1969).
55. McBride, S. G., Osburn, E. D., Barrett, J. E. & Strickland, M. S. Volatile methanol and acetone additions increase labile soil carbon and inhibit nitrification. *Biogeochemistry* **145**, 127–140 (2019).
56. Slessarev, E. W. & Schimel, J. P. Partitioning sources of CO<sub>2</sub> emission after soil wetting using high-resolution observations and minimal models. *Soil Biol. Biochem.* **143**, 107753 (2020).
57. Pegoraro, E., Rey, A., Abrell, L., Haren, J. & Lin, G. Drought effect on isoprene production and consumption in Biosphere 2 tropical rainforest. *Glob. Change Biol.* **12**, 456–469 (2006).
58. van Haren, J. L. M. et al. Drought-induced nitrous oxide flux dynamics in an enclosed tropical forest. *Glob. Change Biol.* **11**, 1247–1257 (2005).
59. Pegoraro, E. et al. Effect of elevated CO<sub>2</sub> concentration and vapour pressure deficit on isoprene emission from leaves of *Populus deltoides* during drought. *Funct. Plant Biol.* **31**, 1137–1147 (2004).
60. Leigh, L. S., Burgess, T., Marino, B. D. V. & Wei, Y. D. Tropical rainforest biome of Biosphere 2: structure, composition and results of the first 2 years of operation. *Ecol. Eng.* **13**, 65–93 (1999).
61. Fasbender, L., Yáñez-Serrano, A. M., Kreuzwieser, J., Dubbert, D. & Werner, C. Real-time carbon allocation into biogenic volatile organic compounds (BVOCs) and respiratory carbon dioxide (CO<sub>2</sub>) traced by PTR-TOF-MS, <sup>13</sup>CO<sub>2</sub> laser spectroscopy and <sup>13</sup>C-pyruvate labelling. *PLoS ONE* **13**, e0204398 (2018).
62. Honeker, L. K. et al. Elucidating drought-tolerance mechanisms in plant roots through <sup>1</sup>H NMR metabolomics in parallel with MALDI-MS, and NanoSIMS imaging techniques. *Environ. Sci. Technol.* **56**, 2021–2032 (2022).

63. Ladd, S. N. et al. Leaf-level metabolic changes in response to drought affect daytime CO<sub>2</sub> emission and isoprenoid synthesis. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.04.29.490001> (2022).
64. Deming, B. L. et al. Measurements of delays of gas-phase compounds in a wide variety of tubing materials due to gas-wall interactions. *Atmos. Meas. Tech.* **12**, 3453–3461 (2019).
65. Holzinger, R. PTRwid: a new widget tool for processing PTR-TOF-MS data. *Atmos. Meas. Tech.* <https://doi.org/10.5194/amt-8-3903-2015> (2015).
66. de Gouw, J. et al. Sensitivity and specificity of atmospheric trace gas detection by proton-transfer-reaction mass spectrometry. *Int. J. Mass Spectrom.* **223–224**, 365–382 (2003).
67. Holzinger, R. et al. Validity and limitations of simple reaction kinetics to calculate concentrations of organic compounds from ion counts in PTR-MS. *Atmos. Meas. Tech.* <https://doi.org/10.5194/amt-12-6193-2019> (2019).
68. de Gouw, J. & Warneke, C. Measurements of volatile organic compounds in the earth's atmosphere using proton-transfer-reaction mass spectrometry. *Mass Spectrom. Rev.* **26**, 223–257 (2007).
69. Krechmer, J. et al. Evaluation of a new vocus reagent-ion source and focusing ion-molecule reactor for use in proton-transfer-reaction mass spectrometry. *Anal. Chem.* **90**, 12011–12018 (2018).
70. Clafflin, M. S. et al. An in situ gas chromatograph with automatic detector switching between PTR- and EI-TOF-MS: isomer-resolved measurements of indoor air. *Atmos. Meas. Tech.* **14**, 133–152 (2021).
71. Gil-Loaiza, J. et al. Versatile soil gas concentration and isotope monitoring: optimization and integration of novel soil gas probes with online trace gas detection. *Biogeosciences* **19**, 165–185 (2022).
72. Braun, A. et al. Reviews and syntheses: Heterotrophic fixation of inorganic carbon – significant but invisible flux in environmental carbon cycling. *Biogeosciences* **18**, 3689–3700 (2021).
73. Pugliese, G. et al. The effect of prolonged drought and recovery on soil VOC fluxes in an experimental rainforest. *Nat. Commun.* (in the press).
74. Koebsch, F., Glatzel, S. & Jurasinski, G. Vegetation controls methane emissions in a coastal brackish fen. *Wetl. Ecol. Manage.* **21**, 323–337 (2013).
75. Jurasinski, G., Koebsch, F. & Hagemann, U. Flux rate calculation from dynamic closed chamber measurements. R package v0.3-0.1 (2022).
76. Tfaily, M. M. et al. Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. *Anal. Chem.* **87**, 5206–5215 (2015).
77. Tfaily, M. M. et al. Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures. *J. Vis. Exp.* <https://doi.org/10.3791/59035> (2019).
78. Tfaily, M. M. et al. Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. *Anal. Chim. Acta* **972**, 54–61 (2017).
79. Dittmar, T., Koch, B., Hertkorn, N. & Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* **6**, 230–235 (2008).
80. Willcott, M. R. MestRe Nova. *J. Am. Chem. Soc.* **131**, 13180 (2009).
81. Tolić, N. et al. Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. *Anal. Chem.* **89**, 12659–12665 (2017).
82. Tfaily, M. M. et al. Vertical stratification of peat pore water dissolved organic matter composition in a peat bog in Northern Minnesota. *J. Geophys. Res. Biogeosci.* **123**, 479–494 (2018).
83. Ayala-Ortiz, C. O. et al. MetaboDirect: an analytical pipeline for the processing of FTICR-MS-based metabolomics data. *Microbiome* **11**, 38 (2023).
84. Clum, A. et al. DOE JGI metagenome workflow. *mSystems* **6**, e00804–e00820 (2021).
85. Bushnell, B., Rood, J. & Singer, E. BBMerge-accurate paired shotgun read merging via overlap. *PLoS One* **12**, e0185056 (2017).
86. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
87. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
88. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
89. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935 (2013).
90. Bland, C. et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* **8**, 209 (2007).
91. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
92. Lomsadze, A., Gemayel, K., Tang, S. & Borodovsky, M. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. *Genome Res.* **28**, 1079–1089 (2018).
93. Kanehisa, M. & Sato, Y. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci.* **29**, 28–35 (2020).
94. Kiełbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).
95. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing) (2020).
96. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team Linear and Nonlinear Mixed Effects Models. R package v.3 (2007).
97. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
98. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
99. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, i11 (2012).
100. Mevik, B.-H. & Wehrens, R. The pls package: principal component and partial least squares regression in R. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v018.i02> (2007).
101. Lu, B., Castillo, I., Chiang, L. & Edgar, T. F. Industrial PLS model variable selection using moving window variable importance in projection. *Chemometr. Intell. Lab. Syst.* **135**, 90–109 (2014).
102. Liland, K. H., Mehmood, T. & Sæbø, S. plsVarSel: Variable Selection in Partial Least Squares. R package v.0.9.6 (2020).
103. Wilkinson, L. ggplot2: elegant graphics for data analysis by WICKHAM, H. *Biometrics* **67**, 678–679 (2011).
104. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).

## Acknowledgements

A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) programme (proposal: <https://doi.org/10.46936/fics.proj.2019.50971/60000130>) awarded to M.M.T. and used resources at the DOE Joint Genome Institute (<https://ror.org/04xm1d337>) and the Environmental Molecular Sciences Laboratory (<https://ror.org/04rc0xn13>), which are DOE Office of Science User Facilities operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL). This material was based upon

work supported by the US Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0023189 awarded to L.K.M. This research was also supported in part by the European Research Council (ERC; Grant Number 647008) awarded to C.W., Department of Energy, Office of Science Biological and Environmental Research Grant (DE-SC0021349), and NSF CAREER Award (No. 2045332) to L.K.M. L.K.H. was supported by Biosphere 2 through the office of the Senior Vice President for Research Innovation and Impact at the University of Arizona, NSF CAREER Award (No. 2045332), and the BIO5 Postdoctoral Fellowship. G.P. was supported by the German Federal Ministry of Education and Research (BMBF contract 01LB1001A – ATTO+) and the Max Planck Society. We also acknowledge financial support from the Philecology Foundation.

### Author contributions

C.W., L.K.M. and S.N.L. conceived the project and experimental design. J.I., J.F., J.G.-L., J.E.K., M.A.D., L.E.D., S.N.L., C.W., M.M.T. and L.K.M. contributed to experimental design. J.I., J.F., J.G.-L., J.E.K., J.K., M.S.C., L.E.D., M.A.D., E.Y.P., L.S., S.N.L. and L.K.M. performed the experiments in the field. J.E.K., J.W., J.K., D.W.H., R.K.C., J.T. and M.M.T. contributed equipment and analysis tools. E.C., E.S., L.K.H., G.H. and D.W.H. contributed to sample preparation and processing. L.K.H., G.P., J.E.K., K.M., J.I., D.W.H., C.A.-O. and V.F.-Z. analysed the data. L.K.H., G.P. and J.E.K. contributed to making figures. L.K.H., L.K.M., S.N.L. and M.M.T. wrote the paper.

### Competing interests

The authors declare no competing interests.

### Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41564-023-01432-9>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41564-023-01432-9>.

**Correspondence and requests for materials** should be addressed to Laura K. Meredith.

**Peer review information** *Nature Microbiology* thanks Riikka Rinnan, Steve McBride, Steffen Kolb and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023, corrected publication 2023

**Extended Data Table 1 | Metabolic concentrations ( $\mu\text{M}$ ) detected using NMR**

Compound Name	Formula	Pre-drought					Drought					
		P15- Ohr	P24- Ohr	P26- Ohr	P34- Ohr	P35- Ohr	P12- Ctrl	P15- Ohr	P24- 6hr	P26- Ohr	P34- Ohr	P34- Ohr
2-Oxoisocaproate**P	$\text{C}_6\text{H}_{10}\text{O}_3$	6	5	D	5	D	ND	ND	ND	ND	ND	D
3-Hydroxybutyrate	$\text{C}_4\text{H}_8\text{O}_3$	ND	ND	ND	ND	D	ND	ND	ND	ND	ND	ND
Acetate	$\text{C}_2\text{H}_4\text{O}_2$	22	38	30	55	26	19	35	31	31	24	11
Acetone	$\text{C}_3\text{H}_6\text{O}$	ND	D	D	ND	ND	D	ND	ND	ND	ND	D
Alanine**P	$\text{C}_3\text{H}_7\text{NO}_2$	24	21	15	20	17	18	9	13	9	7	9
Benzoate	$\text{C}_7\text{H}_6\text{O}_2$	ND	ND	ND	D	ND	ND	ND	ND	ND	ND	ND
Betaine	$\text{C}_5\text{H}_{11}\text{NO}_2$	ND	ND	ND	ND	ND	D	D	ND	ND	ND	ND
Ethanol	$\text{C}_2\text{H}_6\text{O}$	12	17	21	19	18	28	24	12	22	22	12
Formate***P	$\text{CH}_2\text{O}_2$	17	38	40	23	31	ND	D	6	16	D	11
Glutamate	$\text{C}_5\text{H}_9\text{NO}_4$	12	12	8	6	10	11	7	10	10	10	6
Glycerol	$\text{C}_3\text{H}_8\text{O}_3$	6	8	8	6	7	5	6	7	5	8	6
Glycine*P	$\text{C}_2\text{H}_5\text{NO}_2$	12	9	6	8	7	5	D	7	6	D	D
Isoleucine	$\text{C}_6\text{H}_{13}\text{NO}_2$	7	7	D	D	D	D	D	D	D	D	D
Lactate	$\text{C}_3\text{H}_6\text{O}_3$	6	7	D	12	7	D	D	11	D	D	7
Leucine*P	$\text{C}_6\text{H}_{13}\text{NO}_2$	12	11	7	7	6	7	6	6	6	D	D
Methanol	$\text{CH}_4\text{O}$	150	375	367	222	326	508	329	203	370	372	322
Phenylalanine*P	$\text{C}_9\text{H}_{11}\text{NO}_2$	D	D	ND	D	ND	ND	ND	ND	ND	ND	ND
Pyroglutamate*P	$\text{C}_5\text{H}_7\text{NO}_3$	22	19	10	16	9	8	5	12	8	D	D
Pyruvate**P	$\text{C}_3\text{H}_4\text{O}_3$	11	13	10	11	10	D	5	9	5	D	7
Succinate	$\text{C}_4\text{H}_6\text{O}_4$	ND	ND	ND	ND	ND	ND	D	D	ND	ND	ND
Threonine	$\text{C}_4\text{H}_9\text{NO}_3$	6	7	D	D	D	D	D	D	D	D	D
Trehalose*D	$\text{C}_{12}\text{H}_{22}\text{O}_{11}$	ND	ND	ND	ND	ND	18	D	7	6	13	D
Uracil*P	$\text{C}_4\text{H}_4\text{N}_2\text{O}_2$	D	5	D	D	D	ND	ND	D	ND	ND	D
Valine*P	$\text{C}_5\text{H}_{11}\text{NO}_2$	15	13	8	9	8	8	6	8	6	5	5

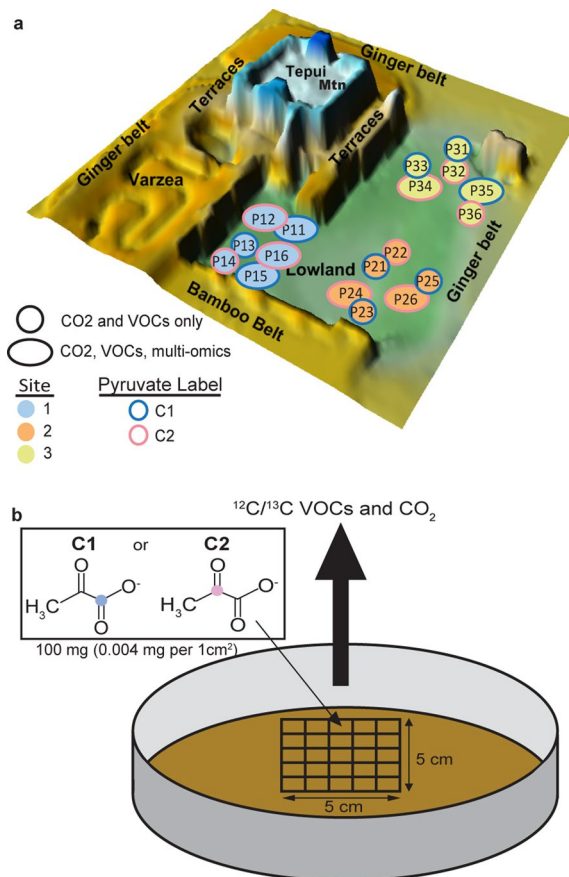
Samples displayed are a subset of total samples based on the quality of NMR data. Most samples are from 0h, except for P24 during drought which is from 6h post pyruvate injection. ND, not detected (below the level of detection where concentration  $< 1.9\mu\text{M}$ ); D, detected (below level of quantification, where concentration  $2 - 5\mu\text{M}$ )<sup>o</sup>, Compound higher during pre-drought (Linear Fixed Effect Model (d.f. = 7), where ND and D were given values of 1 and 4  $\mu\text{M}$  respectively, Oxoisocaproate  $P=0.0014$ , Alanine  $P=0.0074$ , Formate  $P=4.2\text{E-}4$ , Glycine  $P=0.017$ , Leucine  $P=0.016$ , Phenylalanine  $P=0.030$ , Pyroglutamate  $P=0.018$ , Pyruvate  $P=0.0011$ , Uracil  $P=0.022$ , Valine  $P=0.012$ )<sup>o</sup>, Compound higher during drought (Trehalose  $P=0.020$ )<sup>\*</sup>,  $P<0.05$ ; \*\*,  $P<0.01$ ; \*\*\*,  $P<0.001$



**Extended Data Table 2 | Sequence coverage for metagenomics and metatranscriptomics**

Sample	Metagenomics			Metatranscriptomics		
	Reads	Bases (Gb)	Mapped to assembly (%)	Reads	Bases (Gb)	Mapped to assembly (%)
P11SSC1_190916_c	142713986	21.4	49.1	124166928	18.0	77.9
P11SSC1_190916_f	224647504	33.6	60.0	134601358	19.5	81.2
P11SSC1_190916_i	155786656	23.3	53.4	101082054	14.5	77.6
P11SSC1_191110_c	227806600	34.1	64.9	82931902	12.0	81.1
P11SSC1_191119_f	132381506	19.8	55.9	107176982	15.4	82.3
P11SSC1_191119_i	257335184	38.5	64.8	76710842	11.0	81.1
P12SS_CTRL_191116_c	135568486	20.3	52.9	92715120	13.2	77.0
P12SS_CTRL_191116_f	103105310	15.4	48.0	146171918	20.8	73.4
P12SS_CTRL_191116_j	89432070	13.4	45.8	85117324	12.1	81.9
P15SSC1_190912_c	141457632	21.2	44.5	95565802	13.6	77.1
P15SSC1_190913_f	136093864	20.4	41.0	102381422	14.5	73.1
P15SSC1_190915_i	116566008	17.5	39.7	105544744	15.0	67.5
P15SSC1_191107_f	126857110	19.0	43.0	22699542	3.2	68.0
P15SSC1_191107_j	141678932	21.2	43.5	104936440	14.9	75.1
P15SSC1_191107_c	131734810	19.7	45.6	106270712	15.0	78.9
P24SSC2_190915_c	131640432	19.7	54.6	93137922	13.3	80.2
P24SSC2_190915_f	108754462	16.3	50.0	145088652	20.6	77.5
P24SSC2_190915_i	143103050	21.4	55.0	120662996	17.1	73.7
P24SSC2_191109_c	122604752	18.4	49.7	124732280	17.6	74.2
P24SSC2_191109_f	145461884	21.8	59.2	79879116	11.5	74.6
P24SSC2_191109_i	101194304	15.2	44.3	72649070	10.2	77.6
P26SSC2_190918_c	138246448	20.7	58.3	76395558	10.8	63.8
P26SSC2_190918_f	110179686	16.5	54.9	94868962	13.4	65.5
P26SSC2_190918_i	112067984	16.8	49.2	116842828	16.7	76.7
P26SSC2_191112_c	98449944	14.7	50.9	85137010	12.0	72.2
P26SSC2_191112_f	119958450	18.0	58.9	78330638	10.9	68.4
P26SSC2_191112_j	175881098	26.3	62.4	97683038	13.7	68.9
P34SSC2_190917_c	132462676	19.8	55.5	71554676	10.2	77.8
P34SSC2_190917_f	101846050	15.3	48.6	183218676	25.8	78.6
P34SSC2_190917_j	127521864	19.1	51.5	85630146	12.0	69.7
P34SSC2_191111_c	154773318	23.2	57.8	83204978	12.0	84.2
P34SSC2_191111_f	134298460	20.1	56.9	78791904	10.8	72.9
P34SSC2_191111_j	112046872	16.8	53.5	68499682	9.9	80.1
P35SSC1_190914_c	112270418	16.8	56.0	16633060	2.3	65.2
P35SSC1_190914_f	112035746	16.8	55.5	na	na	na
P35SSC1_190914_i	128627610	19.3	60.0	14394504	2.0	37.3
P35SSC1_191108_c	195950226	29.3	64.4	13623088	1.9	59.7
P35SSC1_191108_f	128984880	19.3	56.8	97460190	13.5	72.1
P35SSC1_191108_i	109718370	16.4	55.1	86907854	12.1	74.3

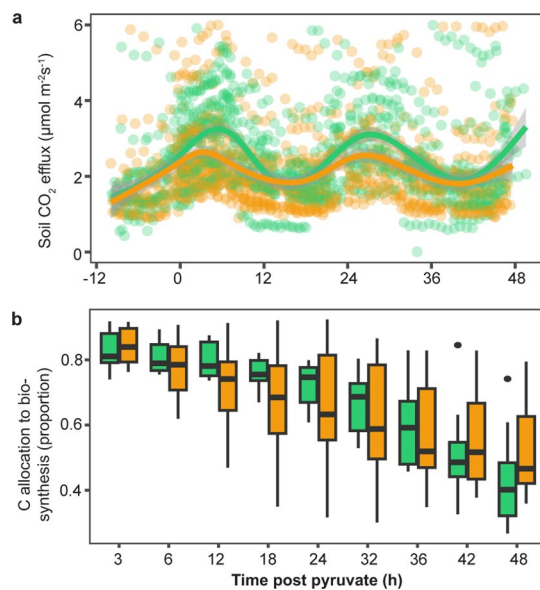
a, 0 h; f, 6 h; i, 48 h na, Metatranscriptomics of sample P35SSC1\_191108\_c were not completed due to sequencing issues Read counts are post-filtering (as described in the methods section).



### Extended Data Fig. 1 | Soil chamber locations and experimental setup.

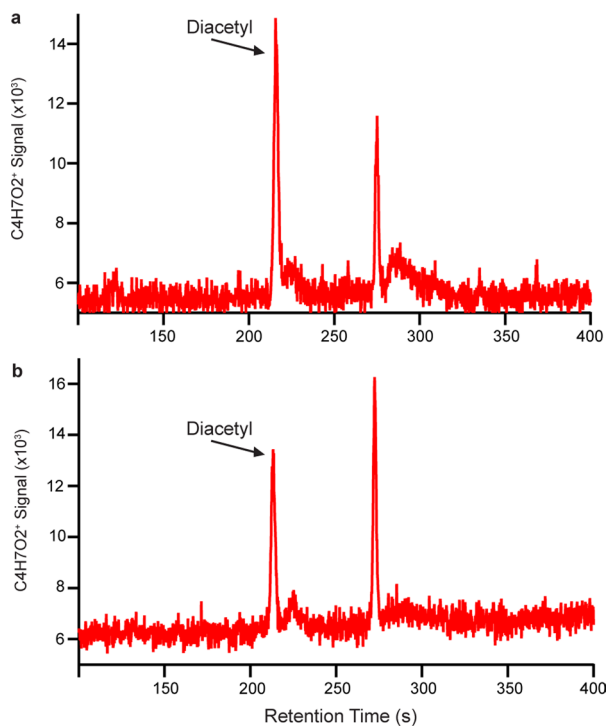
**a)** Topological map of Biosphere 2 Tropical Rainforest, with locations of Sites 1–3. All circles represent collar locations where CO<sub>2</sub> and VOC data were collected and large circles represent a sub selection of sites collars where additional soil samples were collected for metatranscriptomics, metagenomics, and

metabolomics. **b)** Diagram showing soil pyruvate experimental setup with stencil placed inside a soil collar where <sup>13</sup>C<sub>1</sub>- or <sup>13</sup>C<sub>2</sub>-pyruvate was added within each 1 cm×1 cm square. An automatic soil chamber was placed over the soil collar during the experiment and continuous measurements of <sup>12</sup>C and <sup>13</sup>C-CO<sub>2</sub> and VOCs were measured.



**Extended Data Fig. 2 | Continuous CO<sub>2</sub> emissions and proportion of C from pyruvate allocated to biosynthesis.** **a)** Total flux of CO<sub>2</sub> during pre-drought and drought conditions. Each point represents a single measurement per soil chamber (~64 measurements), of which there were 3 replicates each that were injected with either <sup>13</sup>C<sub>1</sub>- or <sup>13</sup>C<sub>2</sub>-pyruvate for each site (total n = 18 each for pre-drought and drought), and lines show the data smoothed with the surrounding shaded area showing ± SEM. **b)** Proportion of C from pyruvate

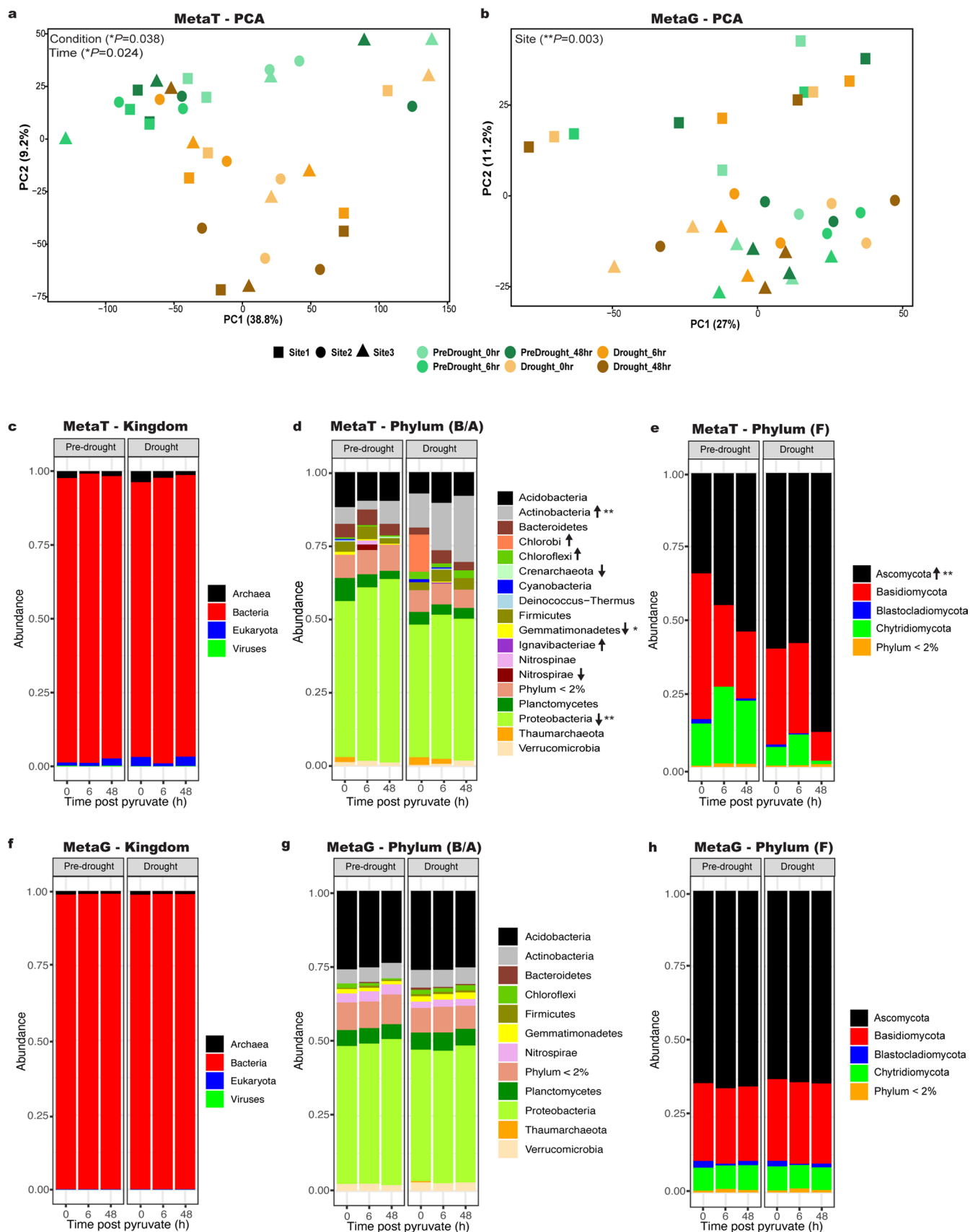
allocated to biosynthesis. This was calculated as  $\frac{^{13}\text{C-CO}_{2\text{-C1}}}{(^{13}\text{C-CO}_{2\text{-C1}} + ^{13}\text{C-CO}_{2\text{-C2}})}$  on each set of C<sub>1</sub>/C<sub>2</sub> chambers per site (n = 9 each for pre-drought and drought) with continuous emission data binned to 3 or 6 h intervals, from 0 to 48 h post pyruvate injection. (n = 9 each for pre-drought and drought). Boxes represent Q1 - Q3 with center line indicating median, and bars extending to maximum and minimum values, excluding outliers.



**Extended Data Fig. 3 | GC-PTR-TOF data to identify  $C_4H_6O_2$ .** Two nearby locations **a**) and **b**) had two peaks with retention times 217 s and 275 s for  $C_4H_6O_2$  corresponding to two different compounds. The retention time of 217 s matched the expected retention time for diacetyl, but the retention time

of 275 s did not match with any known compounds. Therefore,  $C_4H_6O_2$  may represent diacetyl and/or one additional unidentifiable compound, and is referred to as diacetyl\* in the manuscript. The diacetyl peak (retention time of 217 s) is indicated with an arrow.

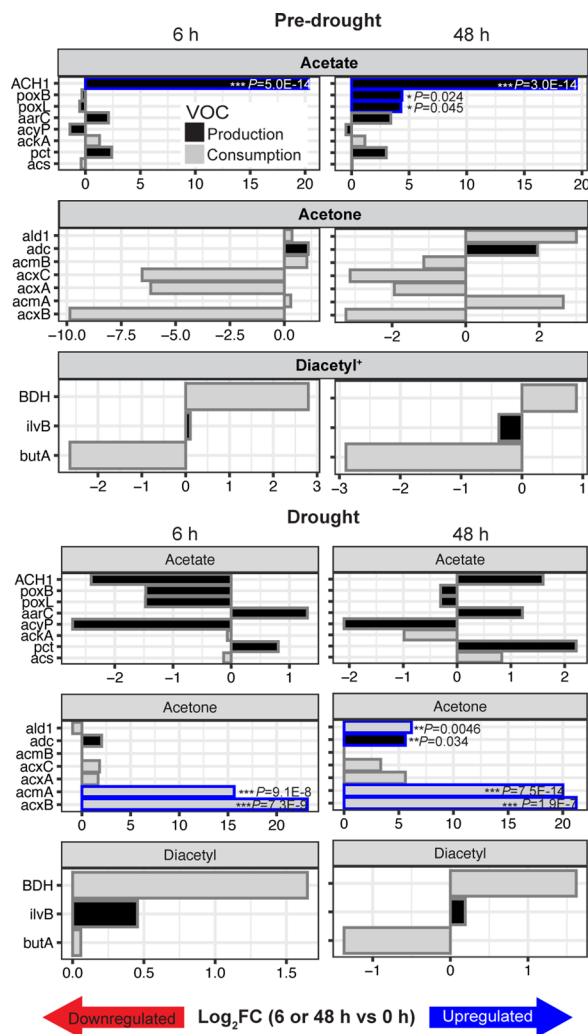




Extended Data Fig. 4 | See next page for caption.

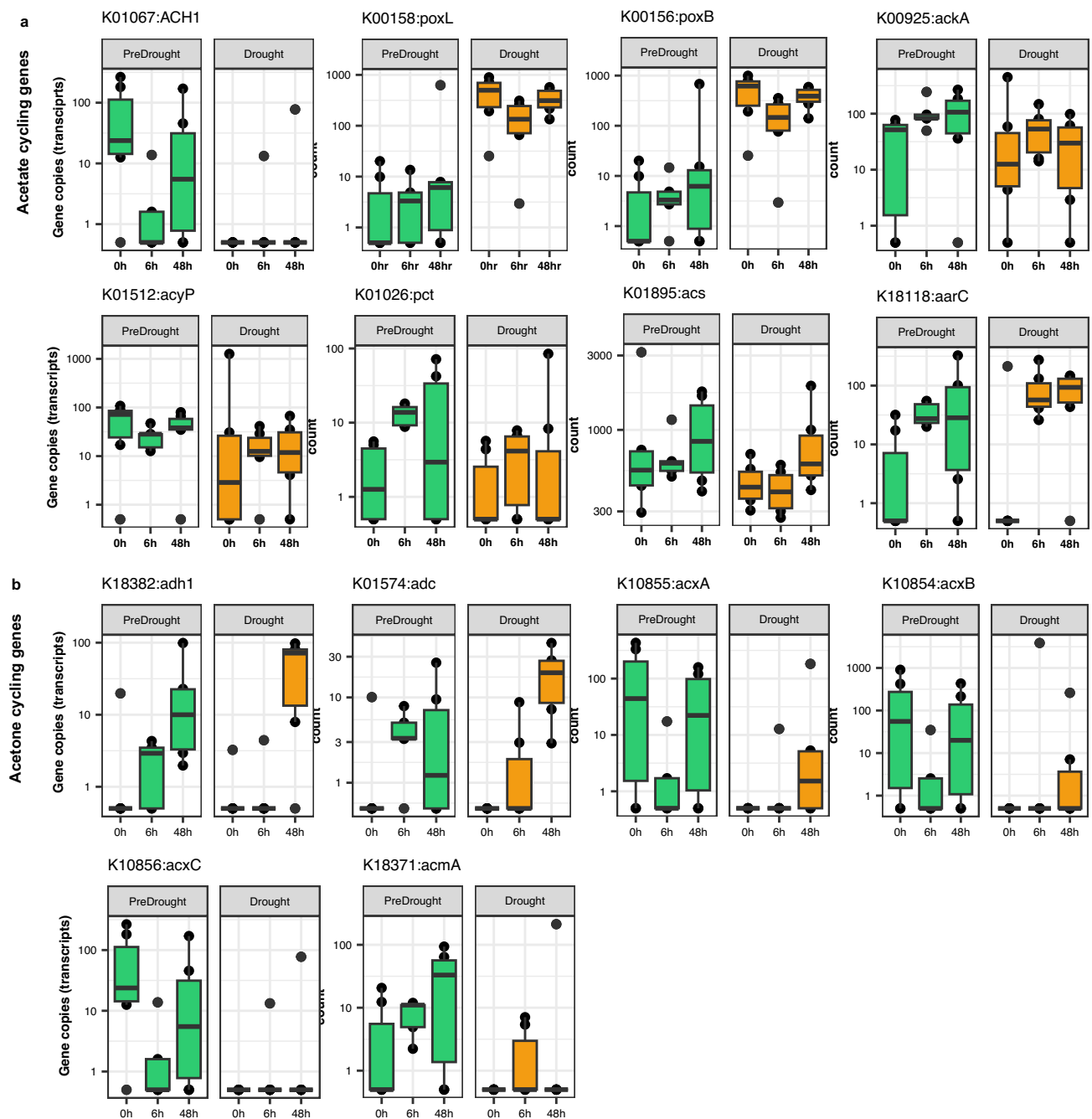
**Extended Data Fig. 4 | Metatranscriptomic and metagenomic functional and taxonomic diversity shows shifts in active microbial functional profiles and taxonomic diversity, but not in functional potential and total (active and inactive) taxonomic diversity.** PCA of (a) metatranscriptomic and (b) metagenomics data. Taxonomic profiles as inferred from metatranscriptomics data at the (c) kingdom-level for active archaea, bacteria, eukaryota, and viruses, (d) phylum-level for bacteria and archaea, and (e) phylum-level for fungi. Taxonomic profiles as inferred from metagenomics data at the (f) kingdom-level for total archaea, bacteria, eukaryota, and viruses, (g) phylum-level for bacteria and archaea, and (h) phylum-level for fungi. For taxonomic profiles, arrows

indicate direction of changes in relative abundance (up for increased, and down for decreased relative abundance in drought compared to pre-drought conditions). Arrows with no stars means the phylum was present in one condition (pre-drought or drought) and absent in the other condition. For panels a-b, P-values were determined using PERMANOVA on Bray-Curtis distance matrices. For panels c-h, P-values were determined using Linear Mixed Effect Models. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; (exact P-values:  $P = 0.010$  [Proteobacteria],  $P = 0.016$  [Gemmatimonadetes],  $P = 0.0012$  [Actinobacteria] (panel b),  $P = 0.0050$  [Ascomycota] (panel c)). metaT, metatranscriptomics; metaG, metagenomics; B, bacteria; A, archaea; F, fungi.



**Extended Data Fig. 5 | Up- or down-regulation of genes encoding for enzymes that cycle <sup>13</sup>C-enriched VOCs shift with pyruvate addition.** Log<sub>2</sub>-fold change (FC) of genes involved in cycling of acetate, acetone, and diacetyl\* at 6 and 48 h in relation to 0 h to show effect of pyruvate addition on microbial activity during **a)** pre-drought and **b)** drought conditions. Log<sub>2</sub>-FC values to indicate up- or

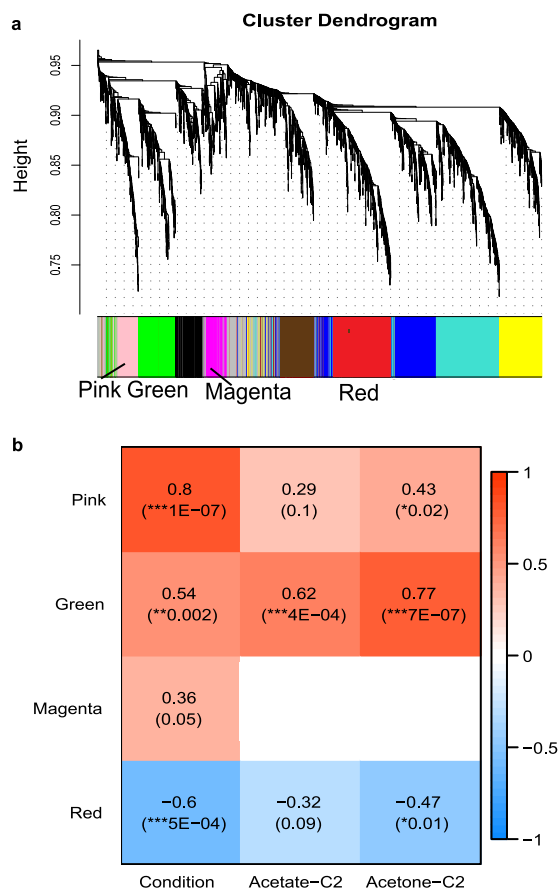
down-regulation calculated with DESeq2. Production vs. consumption indicated by black and grey bars, respectively. Significantly up- or down-regulated genes are outlined in color (blue = upregulated, red = downregulated), with exact *P*-values from DESeq2 analysis shown.); \*, *P* < 0.05; \*\*, *P* < 0.01; \*\*\*, *P* < 0.001.



**Extended Data Fig. 6 | Acetone and acetate cycling gene transcripts.** Gene copies (transcripts) of **a**) acetate- and **b**) acetone-cycling genes at 0, 6, and 48 h post pyruvate injection. Bars represent Q1 - Q3 with center line indicating median, and bars extending to maximum and minimum values, excluding outliers, across all chambers and sites ( $n = 9$  each for pre-drought and drought).







**Extended Data Fig. 8 | Modules of co-expressed genes that are associated with pre-drought or drought conditions and  $^{13}\text{C}$ -enrichment of VOCs.**  
**a)** WGCNA cluster dendrogram reveals 9 clusters of co-expressed genes, or modules. The four modules with significant correlations to drought or pre-drought (pink, green, magenta, and brown) are indicated. **b)** Correlations of the pink, green, magenta, and brown module eigengenes with environmental conditions (pre-drought = 0, drought = 1) and efflux of  $^{13}\text{C}$ -enriched (VOC

$^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$  acetate and acetone from chambers receiving  $^{13}\text{C}_2$ -pyruvate (acetate-C2 and acetone-C2, respectively) across three time points (0, 6, and 48 h post pyruvate injection; total  $n = 27$  [3 measurements each for each soil collar receiving  $^{13}\text{C}_2$ -pyruvate]). Pearson correlation coefficients are shown with FDR-corrected p-values below in parentheses. White boxes indicate no correlation ( $P < 0.1$ ). \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted <i>Give <math>P</math> values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input type="checkbox"/>	<input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	no software used
Data analysis	R v. 4.0.2, (packages: NLME v. 3.1, WGCNA v. 1.7, DESeq2 v. 1.30.1, ggplot2 v. 3.4.1, pls v. 2.8, plsVarSel v. 0.9.10, ClusterProfiler v. 3.18.1, FLUX v. 0.3), BBtools (BBDuk v. 38.94, BBDuk v. 38.44), metaSPAdes v. 3.15.12, INFERNAL v. 1.1.3, Prodigal v. 2.6.3, GeneMarkS-2 v. 1.07, tRNAscan-SE v. 2.0.6, BrukerDaltonik v. 4.2, PTRwid, MetaboDirect v. 0.2.7, Chenomx NMR suite 9.2, MestReNova, For)mularity Complete R code for all data analysis to produce figures is included in the following GitHub repository ( <a href="https://github.com/linneakh/SoilPyruvate">https://github.com/linneakh/SoilPyruvate</a> )

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The metatranscriptomics and metagenomics sequence data are publicly available through Genbank SRA under the following BioProject ID's: PRJNA980752-

PRJNA980834. FTICR, NMR, VOC, and CO2 data are deposited to FigShare (OOI: 10.6084/m9.figshare.20334537), and individual raw .xml files for the FTICR data are included as Supplementary Data. The KEGG database can be accessed at the following website: <https://www.genome.jp/kegg/pathway.html>

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

Population characteristics

Recruitment

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Our study examined the impact of drought on microbial carbon cycling using position-specific <sup>13</sup>C-pyruvate labeling, to determine how carbon is allocated to VOCs and CO<sub>2</sub>, combined with multi-omics. For VOC and CO<sub>2</sub> data, continuous measurements were made from 0 to 48 hrs post <sup>13</sup>C-pyruvate injection across 6 replicates each within 3 sites of the Biosphere 2 tropical rainforest (n = 18) during both pre-drought and drought conditions (total n = 36). For metabolomics, metatranscriptomics, and metagenomics, data are from discrete time points when soil was collected (0, 6, and 48 hr post pyruvate injection). Due to our limited number of samples that could be analyzed at EMSL, we sent a subset of the sites for analysis of 2 each from sites 1, 2 and 3 (for 3 time points, n= 18) for pre-drought and drought (total n=36).

Research sample

The sample unit originated from a region of soil enclosed by a soil gas chamber with a diameter of 20 cm. For gas measurements, the soil chamber was closed and continuous CO<sub>2</sub> and VOC measurements were made (every 30 min for first ~ 8 hr, then every 50 min until 48 hr). For soil samples, a 2.2 cm diameter metal ring was pushed into soil to a depth of 2 cm.

Sampling strategy

No calculations were made to determine a priori how many samples should be collected. A replicate size of 6 per site was chosen based on feasibility of instrumentation and connection to gas analyzers. A replicate size of 2 per site was chosen for multi-omics based on the limited number of samples we could have analyzed at JGI and EMSL with the FICUS grant. We selected 2 per site rather than 3 each for 2 sites because we wanted to capture the heterogeneity across the tropical rainforest biome and focus on the pre-drought vs drought conditions rather than comparisons between sites.

Data collection

CO<sub>2</sub> and VOC data, including isotopic composition, were collected by Johannes Ingrisch, Eva Pfanterstill, Jordan Krechmer, and Megan Claffin using multiplexed Licor soil flux system (Licor 8100, Li-8150, and Lic 8100-104) coupled to a Picarro G2201-i analyzer and proton-transfer-reaction time-of-flight spectrometer (PTR-TOF 8100). Metatranscriptomic and metagenomic data were collected by the joint genome institute (JGI) on an Illumina NovaSeq. Metabolomics data were collected at EMSL (PNNL) on an fourier transform ion-coupled resonance (FTICR) mass spectrometer for larger metabolites, and a nuclear magnetic resonance (NMR) spectrometer for small metabolites. Soil moisture data were collected by various researchers on a handheld probe coupled to a LabQuest viewer.

Timing and spatial scale

<sup>13</sup>C-pyruvate injection experiments and sample collections were performed between September 12 - 15 and November 17-19 for pre-drought and drought conditions, respectively. Per sampling location/chamber, gas measurements were made for 48 hours post <sup>13</sup>C-pyruvate injection, and for multi-omics, samples were collected and analyzed from 0, 6, and 48 hr time points post pyruvate injection.

Data exclusions

No data were excluded

Reproducibility

There were no attempts to reproduce the experiment due to the immense time, costs, and manpower (international research team) required to conduct the drought experiment. We tried to include sufficient replicates to ensure confidence in our experimental results.

Randomization

Samples were not randomized into groups, due to the nature of field collection from distinct locations and time points.



- Blinding Gas data acquisition could not be blinded due to the nature of the study. Genomic and metabolomic data acquisition was blinded because the individuals collected the data were not familiar with the project or sample IDs.
- Did the study involve field work?  Yes  No

## Field work, collection and transport

- Field conditions During ambient (pre-drought) conditions, rainfall events were simulated by spraying 15,000 L of irrigation water from the top of the B2 TRF at a frequency of 3 days a week. The last rainfall event before the drought was on October 7, 2019, and the first post-drought rain-fall was on December 12, 2019. During the drought, ambient temperatures were maintained in the lowland region between 20 to 26.7 °C.
- Location Tropical rainforest biome at Biosphere 2 in Oracle, AZ.
- Access & import/export Our study site was not part of a natural ecosystem, so we did not have to adhere to any specific guidelines for minimizing our disturbance to the site. The Biosphere 2 is a 1.27 ha steel and glass-enclosed building that is highly accessible for research. Samples were collected, stored in coolers with dry ice and then moved to refrigerators/freezers on site. From there, samples were transported to the University of Arizona, about 40 miles away.
- Disturbance Disturbance was minimal with soil samples collected only to a depth of 2 cm at 18 sites.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- | n/a                                 | Involvement  |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |

### Methods

- | n/a                                 | Involvement                                     |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |