

ARTICLE

DOI: 10.1038/s41467-017-01753-4

OPEN

# Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions

Jordan C. Angle<sup>1</sup>, Timothy H. Morin<sup>2,3</sup>, Lindsey M. Soden<sup>1</sup>, Adrienne B. Narrowe<sup>4</sup>, Garrett J. Smith<sup>1</sup>, Mikayla A. Borton<sup>1,3</sup>, Camilo Rey-Sanchez<sup>2,3</sup>, Rebecca A. Daly<sup>1</sup>, Golnazalsat Mirfenderesgi<sup>2</sup>, David W. Hoyt<sup>1</sup>, William J. Riley<sup>6</sup>, Christopher S. Miller<sup>1</sup>, Gil Bohrer<sup>1</sup>, Kelly C. Wrighton<sup>1,3</sup>

The current paradigm, widely incorporated in soil biogeochemical models, is that microbial methanogenesis can only occur in anoxic habitats. In contrast, here we show clear geochemical and biological evidence for methane production in well-oxygenated soils of a freshwater wetland. A comparison of oxic to anoxic soils reveal up to ten times greater methane production and nine times more methanogenesis activity in oxygenated soils. Metagenomic and metatranscriptomic sequencing recover the first near-complete genomes for a novel methanogen species, and show acetoclastic production from this organism was the dominant methanogenesis pathway in oxygenated soils. This organism, *Candidatus Methanothrix paradoxum*, is prevalent across methane emitting ecosystems, suggesting a global significance. Moreover, in this wetland, we estimate that up to 80% of methane fluxes could be attributed to methanogenesis in oxygenated soils. Together, our findings challenge a widely held assumption about methanogenesis, with significant ramifications for global methane estimates and Earth system modeling.

<sup>1</sup> Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA. <sup>2</sup> Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, Columbus, OH 43210, USA. <sup>3</sup> Environmental Science Graduate Program, The Ohio State University, Columbus, OH 43210, USA.

<sup>4</sup> Department of Integrative Biology, University of Colorado Denver, Denver, CO 80217, USA. <sup>5</sup> Pacific Northwest National Laboratory, Richland, WA 99352, USA. <sup>6</sup> Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Jordan C. Angle and Timothy H. Morin contributed equally to this work.

Correspondence and requests for materials should be addressed to K.C.W. (email: [wrighton.1@osu.edu](mailto:wrighton.1@osu.edu))

**M**odeling and biological studies investigating methane flux from wetlands discount microbial methane production in surface, oxic soils<sup>1, 2</sup>. The basis of this assumption is that critical methanogen enzymes are inactivated by oxygen and methanogens are poor competitors with other microorganisms for shared substrates<sup>3, 4</sup>. Because of the assumed physiological constraint that oxygen has on methanogens, global terrestrial biogeochemical models limit soil methane production in the presence of dissolved oxygen (DO)<sup>5</sup>.

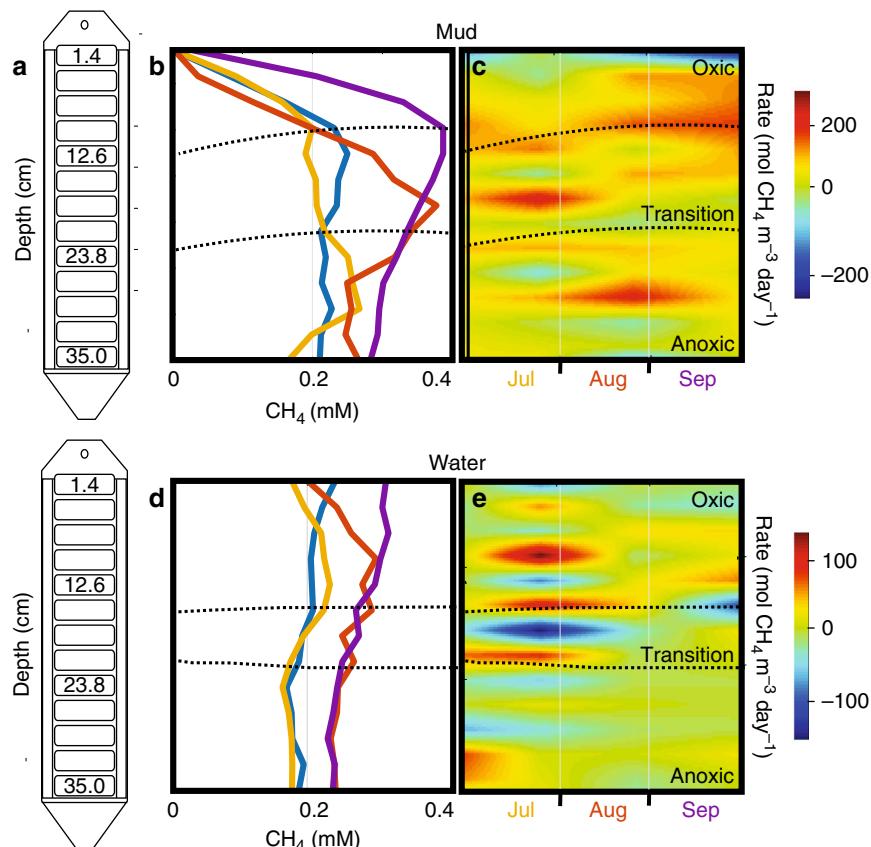
Recent reports present an alternative view that in some ecosystems methanogenesis also occurs in oxic environments, known as the methane paradox. In freshwater lakes, isotopic and molecular biology techniques provided evidence for the presence and activity of methanogens in well-oxygenated portions of the water column<sup>6–8</sup>. Similarly, isotopic signatures in oxygenated soils and activity measurements from soil laboratory enrichments have provided intriguing evidence for methanogenesis in soils with up to 19% oxygen<sup>9, 10</sup>. Despite this mounting, indirect evidence, comprehensive genomic investigations that link methanogens to methane production in any oxic habitat *in situ* are lacking.

Here we analyze observations from the Old Woman Creek (OWC) National Estuarine Research Reserve, a freshwater wetland at the shore of Lake Erie in Ohio. In this study, we experimentally assess biological methane production and emission in freshwater wetland soils across multiple spatial and temporal gradients. The results presented here provide the first ecosystem-scale demonstration of methane production in bulk-oxic soils, its microbial drivers, and the global significance of this currently underappreciated process.

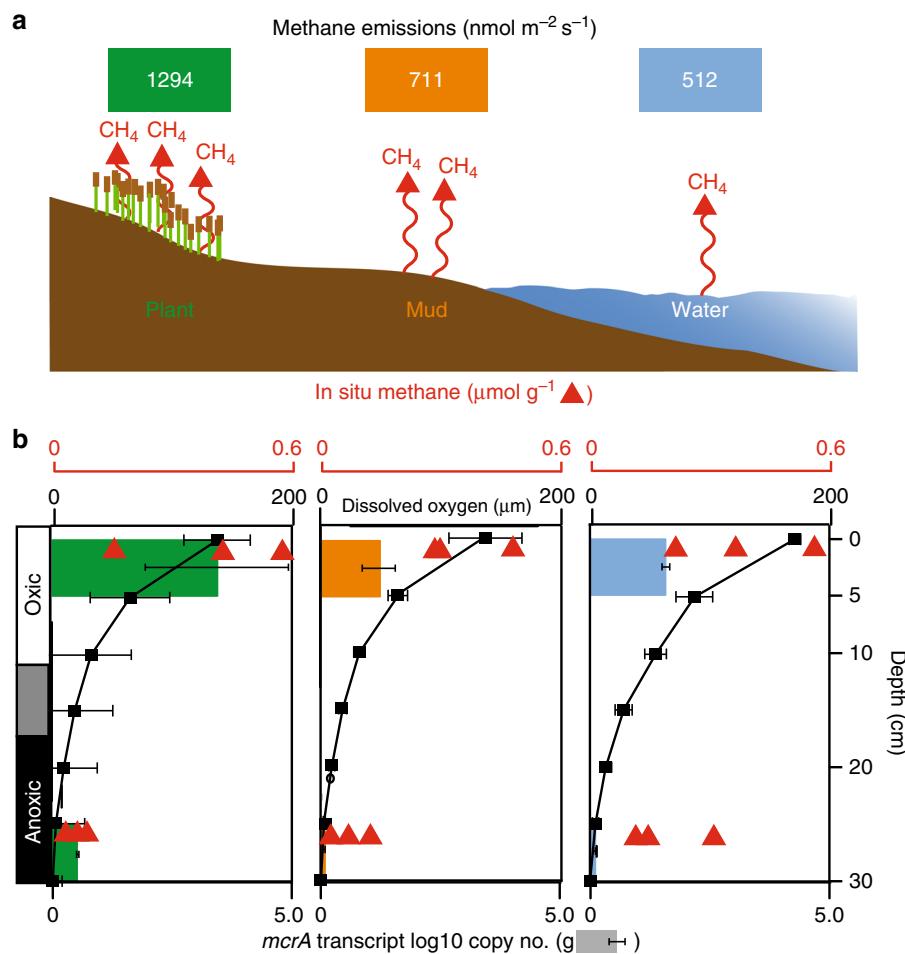
## Results

**Methanogens are most active in oxic surface soils.** To account for differences associated with distinct ecological sites in the wetland (ecosites), we sampled soils beneath three land coverage types: emergent vegetation (plant); periodically exposed mud flats (mud); and continuously submerged under open water (water) (Supplementary Fig. 1). Seasonal variability, especially the effects from photosynthesis and climate, was accounted for by sampling the three ecosites in summer (peak primary production) and late fall before freezing (senescence), while differences in vertical oxygen distributions were examined in 5 cm intervals up to 35 cm deep (Supplementary Data 1).

All ecosites were net methane emitting during both summer and fall sampling seasons (Supplementary Fig 2A). In summer, regardless of ecosite (plant, mud, water), the porewater DO profiles were similar; for instance, depths shallower than 10 cm were always oxic while soils deeper than 25 cm were always anoxic (Fig. 1, Supplementary Data 1). The *in situ* porewater dialysis samplers (peepers) measured the greatest methane concentrations in oxic, surface porewaters in the four summer months sampled (June–Sept). For mud and water ecosites, we paired these concentration measurements with direct surface flux measurements from static chambers, and used a dynamic diffusion model to calculate the net methane source (production and destruction) rate at each layer within the soil column (Supplementary Note 1). Compared to nonoxygenated soil layers, methane was frequently produced in larger amounts in oxygenated layers, in some instances up to an order of magnitude more, but the proportion varied with season and ecosite (Fig. 1).



**Fig. 1** Methane concentrations and production rates across soil depths. **a** Porewater dialysis peepers provide 2.8 cm resolved depth methane measurements. **b, d** Monthly in situ porewater dissolved methane concentrations in mud and water-covered soils with data collected from June (blue), July (yellow), August (red), and September (purple). Black dashed lines depict the 95% confidence interval for location of the oxic to anoxic transition. **c, e**, The calculated net methane volumetric fluxes in soils columns from mud and water ecosites show seasonal methane production (orange and red) in oxic soils (Supplementary Note 1)



**Fig. 2** The relationship between soil dissolved oxygen concentration and methanogenic activity with depth and ecosites from Summer. **a** Schematic of the three ecosites examined in this study with methane emissions shown in colored boxes and depicted by red lines (Supplementary Fig. 2A). **b** Dissolved oxygen concentrations (black boxes), transcripts for *mcrA* (colored bars), and porewater methane concentrations (red triangles) in soils. Error bars reflect s.e. (*mcrA*) and s.d. (oxygen);  $n = 3$

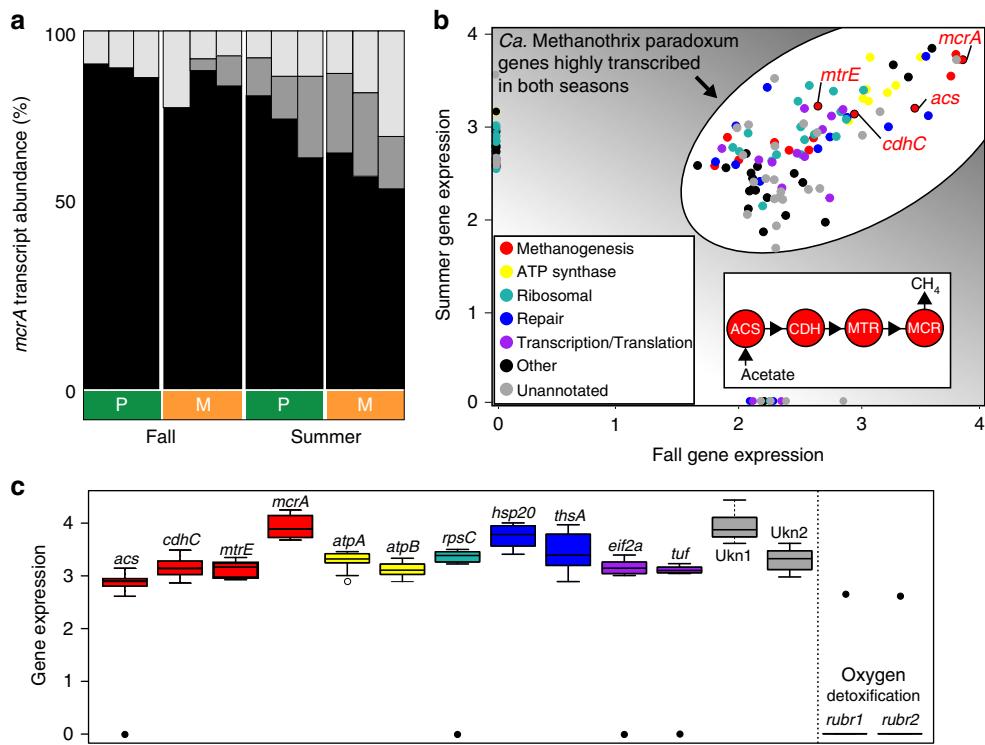
These findings demonstrate that the methane paradox occurs in wetland soils and provides the first evidence for the extent to which it operates over spatial and temporal gradients.

In order to measure methanogenesis activity from these surface and deep soils, we quantified *methyl-coenzyme reductase subunit A* (*mcrA*) gene transcripts, a key functional gene for inferring methanogenesis activity<sup>11</sup>. On average, across all ecosites and seasons, oxic soils contained nine times more *mcrA* transcripts and twice the methane concentration per gram of wet soil than anoxic soils (Fig. 2, Supplementary Data 1). Methanogen activity was positively correlated to porewater DOC and acetate concentrations, but not to other soluble methanogenic substrates like formate, methylamines, and methanol (Supplementary Fig. 2B and Supplementary Data 1). Taken together, these findings suggest that methanogens utilizing acetate may be responsible for sustaining the methane paradox in these soils.

**Candidatus Methanothrix paradoxum is active in oxic soils.** Paired metagenomic and metatranscriptomic sequencing provided the first holistic insight into the methanogens active in oxic environments. From metagenomics sequencing we reconstructed six (two estimated to be > 90% complete) genomes from oxic soils that represent a new species of methanogenic Archaea. Based on whole genome comparisons and phylogenetic analyses (e.g., 16S rRNA, concatenated ribosomal protein, and *mcrA*)

(Supplementary Fig. 3) these genomes clearly represent a new species within the genus *Methanothrix* (formerly *Methanosaeta*). Based on these analyses this new species was phylogenetically most closely related to *M. concili*, a methanogen species widely distributed in anoxic terrestrial methanogenic environments, such as flooded rice paddy soils and lake sediments<sup>12, 13</sup>. Comparative genomic analyses between these wetland genomes to four genomes from cultivated *Methanothrix* demonstrated the *Candidatus* *Methanothrix paradoxum* genomes expanded the *Methanothrix* pan-genome by 27%, with 467 genes uniquely encoded in our wetland genotypes. Of these unique genes, the majority (55%) lacked any functional annotation information (Supplementary Fig. 3C). Here we propose the name *Candidatus* *Methanothrix paradoxum*, after the implied role for this organism in the soil methane paradox (Supplementary Figs. 4–6 and Supplementary Note 2).

From our metatranscriptomic analyses, we conclude methanogenesis in oxic soils is conducted primarily via a canonical acetoclastic pathway (Supplementary Note 3 and Supplementary Data 2). Transcripts from these genomes were in the top 3% of all community-wide metatranscripts and accounted for on average 84% of the *mcrA* transcripts in surface soils (Fig. 3a, Supplementary Fig. 7, and Supplementary Data 3). In addition to the methanogenesis pathway, genes for protein synthesis and energy generation were consistently and highly expressed in both seasons and ecosites (Fig. 3b and Supplementary Fig. 8), signifying active



**Fig. 3** *Candidatus Methanothrix paradoxum* genes transcribed in oxic soils. **a** Taxonomic assignment and relative abundance of *mcrA* transcripts in surface soils assigned to *Candidatus Methanothrix paradoxum* (black), *Methanoregula* (dark gray), and other methanogens (light gray). **b** The relationship between the 100 most transcribed genes (by log<sub>10</sub> FPKM) in each season, with gene functional categories denoted in color and key steps of the methanogenesis pathway highlighted. **c** Gene expression levels for selected genes from (b), across all samples with color legend used from (b) (black line, boxes, and whiskers represent the median, quartiles, and minimum/maximum of the log<sub>10</sub>FPKM values). For comparison, oxygen detoxification genes are not consistently transcribed at detectable levels

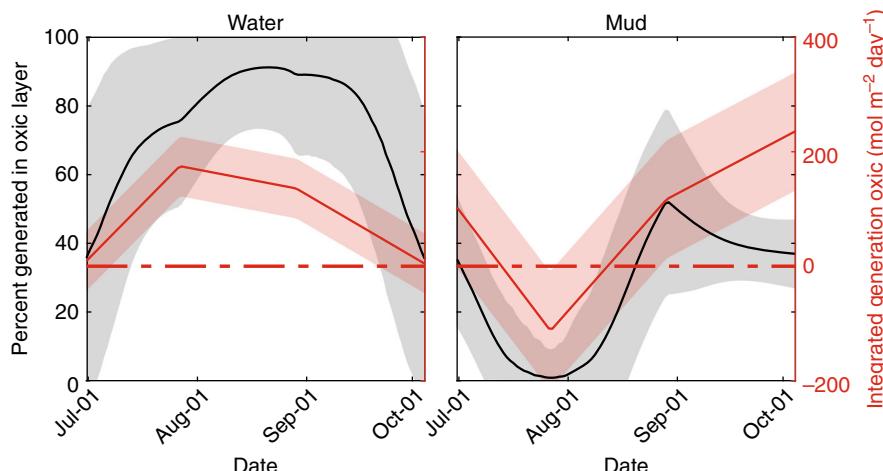
methanogenesis by this organism stably occurs in these oxic wetland soils.

Prior laboratory investigations have shown that methanogens in pure culture or from soil mesocosms upregulate antioxidant mechanisms to attenuate oxygen toxicity<sup>14–16</sup>. Consistent with those reports, *Candidatus Methanothrix paradoxum* genomes encode known oxygen detoxification genes including those for stabilizing free radicals, reducing toxic reactive oxygen species, and repairing oxidative disulfide damage (Supplementary Note 3 and Supplementary Data 2). However, these genes were not unique to our wetland genomes, present in similar abundances across all other *Methanothrix* spp. More notably, oxygen tolerance genes were not consistently transcribed in our oxic wetland soil samples by *Candidatus Methanothrix paradoxum* or any other methanogen. To illustrate the minimal transcript detection, we reported the two most abundant oxygen detoxification genes (rubrerythrin) alongside other more highly transcribed genes (Fig. 3c and Supplementary Fig. 8). Additionally, we cannot rule out that some of the highly and consistently transcribed genes lacking a known annotation in our surface methanogens (Fig. 3c) could play roles in oxygen detoxification; however, the use of remote homology detection via structural modeling<sup>17</sup> and hidden Markov model searches for PFAM domains failed to identify putative oxygen detoxifying genes in these highly expressed but unannotated genes. These metatranscriptomic findings demonstrate for the first time that oxygen detoxification is not a requirement for sustained anaerobic methanogen activity in oxic habitats.

Accounting for the black queen hypothesis<sup>18</sup>, we did consider that oxygen tolerance could be provided to *Ca. Methanothrix paradoxum* by other members in the soil community. In our metatranscriptomes, we recovered transcripts for a catalase gene

and several superoxide dismutase genes belonging to non-methanotroph Gammaproteobacteria; however, only one of these transcripts was detected in as many as 5 of the 12 samples, and at very low abundances. Importantly, none of these transcripts were highly abundant in our data set, nor correlated to methanogen activity, suggesting that the ability for methanogens to compensate for oxygen toxicity is not likely to originate from other community members. Together, these findings demonstrate that known oxygen detoxification mechanisms used by other methanogens in the laboratory are not a requirement to sustain methanogenesis in these oxic wetland soils.

Analysis of the methane paradox literature revealed several possible explanations for methane production in oxic habitats. In our data, we failed to find metabolite or molecular evidence supporting methane production from microbial decomposition of methylated compounds<sup>19</sup> or by protozoan endosymbionts<sup>20</sup> (Supplementary Discussion). Instead, the data convincingly demonstrate methane production by canonical methanogenic archaea as the driver of methane in these oxic soils. Together, the absence of transcripts for known oxygen tolerance genes and the activity from multiple methanogen genera (e.g., *Methanoregula* accounted for 16% of *mcrA* transcripts in summer (Fig. 3)), suggests a more general explanation for the methane paradox in these soils. We suggest that rather than having special adaptations, surface methanogen activity may be confined to anoxic subfractions (e.g., microsites, soil aggregates, or particles) with locally depleted soil oxygen concentrations relative to otherwise overall oxic surrounding soils. This hypothesis is not without warrant, as anoxic microsites were shown to facilitate anaerobic metabolism in bulk oxygenated soils (e.g., for denitrification, iron reduction) and particle-associated models explain methanogenesis in oxic lake waters<sup>6, 7</sup>.



**Fig. 4** Percent methane generated in ecosites over the season as determined from the diffusion and generation model. These data represent a synthesis of the 10% best performing realizations of the microbial activity terms ( $R(t,z)$ ) as determined by the Markov chain. Red lines show the integrated production/consumption of  $\text{CH}_4$  over the oxic zone, interpolated over time. The heavy red dashed line indicates the net neutral methane generation point. Black lines represent the fraction of methane production that can be attributed to the strictly oxic layer (i.e., production above the 97.5th percentile line of the oxic horizon). The shaded areas of both lines represent 1 s.d. of the 4000  $R(t,z)$  realizations. Oxic layers were almost always a net source, with the exception of August in the mud ecosite. The percent contribution depended on the total production within the soil column as well as the level of production in the oxic layers. As the footprint of the site is primarily open water (97% when accounting for only open water and mud, as we do here), the percent generation curve of open water dominated the site-level budget when calculating the total percent generation in the oxic layers.

#### *Candidatus Methanothrix paradoxum* is globally distributed.

To assess the contribution of *Candidatus Methanothrix paradoxum* to methanogenesis in this wetland and beyond, we mined our data and public databases for highly similar (> 99%) 16S rRNA gene sequences. In this wetland, this methanogen species is cosmopolitan, recovered from 97% of soil samples collected from various depths, ecosites, and time points over 3 years. Moreover, as we previously reported<sup>21</sup>, these methanogens are dominant members of the oxic soil community and unlike other methanogens show a strong enrichment in the top 5 cm of soil (Supplementary Fig. 9 and Supplementary Note 4). *Candidatus Methanothrix paradoxum* is also globally distributed, detected in 102 locations across 4 continents spanning a range of habitats including rice paddy, wetland, and peatland soils (Supplementary Data 4). From these analyses, we infer *Candidatus Methanothrix paradoxum* is well adapted to diverse hydric soils and sediments.

Our analysis of ribosomal 16S rRNA genes from previous methane paradox publications revealed that *Candidatus Methanothrix paradoxum* was detected and often acknowledged as a critical member in 10 studies where the methane paradox was previously reported (Supplementary Note 4 and Supplementary Fig. 10). For instance, many of the reported *Methanosaeta* sequences in oxic lake waters share > 99% 16S rRNA gene identity with *Candidatus Methanothrix paradoxum*<sup>7</sup> (Supplementary Fig. 10). We posit that perhaps the increased activity of *Candidatus Methanothrix paradoxum* over other acetoclastic methanogens in oxic soils may be due to its competitive substrate acquisition under low acetate concentrations<sup>22</sup> (< 1 mM) found in our (Supplementary Data 1) and others' soils<sup>23</sup>. Similarly, a recent report on the importance of acetoclastic methanogenesis to the methane paradox in lakes also alluded to low acetate concentrations in oxic surface waters as a potential contributor<sup>6</sup>. Our findings demonstrate that the *Candidatus Methanothrix paradoxum* genotypes reconstructed here are widespread and active, potentially contributing to methanogenesis in a wide variety of oxic, yet high-methane habitats globally.

#### Oxic soil methanogenesis contributes to the methane budget.

To understand the importance of methanogenesis in oxic soils,

we estimated the contribution of this process to the total methane budget in this wetland using simplifying assumptions (Supplementary Note 5). We first decomposed the eddy covariance flux signal into its ecosite-level contributors<sup>24</sup>. We then applied a diffusion model of porewater dissolved  $\text{CH}_4$  concentrations to determine the location of microbial  $\text{CH}_4$  activity within soil columns. We overlaid the microbial activity profile with the DO concentration profile to determine microbial activity in the oxic layers. Previously, Bogard et al.<sup>6</sup> used a scaling method to demonstrate methanogenesis in the oxic portion of the water column contributed to 20% of lake-wide emissions. Using a similar approach, when integrating over the course of this study, we estimated that between 40 and 90% of methane emitted originated in oxic soil layers (Fig. 4).

This study provides the first genome-resolved view of the methane paradox in any ecosystem and identifies the important contribution of a newly defined and globally distributed methanogen species, *Candidatus Methanothrix paradoxum*. We provide clear evidence for the extent of methanogenesis in oxic soils from multiple seasons and ecosites, and show this process is a significant contributor to overall wetland methane emissions. These findings have important ramifications for global biogeochemical models, as current simulations downregulate methanogenesis in surface soil layers due to oxygen concentrations, potentially greatly underestimating methane emissions. It is therefore critical to refine global biogeochemical models to account for methanogenesis in bulk-oxic surface soils, more accurately predicting net wetland methane emissions and their effects on climate.

#### Methods

**Field sampling.** The field location, OWC National Estuarine Research Reserve (41° 22' N 82° 30' W), is a 573-acre freshwater wetland and reserve located on the southern point of Lake Erie near Huron, Ohio. This site is co-operated by the National Oceanic and Atmospheric Administration (NOAA) and Ohio Department of Natural Resources. This is one of 28 coastal (only two are in the Great Lakes region) NOAA designated sentinel research sites. The site consists of a permanently flooded channel surrounded by marsh, mud flats, and forested upland habitat. We collected soil cores from three (~2 m<sup>2</sup>) ecologically differentiated sites (ecosites): plant, mud, and open. Four or more water-saturated soil cores were collected per ecosite to a depth of 35 cm (width 7 cm) using a modified Mooring

System soil corer. Cores were kept on ice in the field until processing in the laboratory (no more than 2 h), where soils were immediately hydraulically extruded<sup>25</sup>, sub-sectioned into surface (0–5 cm) and deep (23–35 cm), and then transferred into separate sterile Whirl-pak bags for RNA extraction (stored at -80 °C), DNA extraction (stored at -20 °C), and geochemical analysis (stored at 4 °C).

**Soil and porewater geochemical analyses.** Soil total carbon and porewater dissolved organic carbon (DOC) were analyzed via Shimadzu TOC-L with SSM-5000A solid sample combustion unit attachment using methods previously described<sup>26</sup>. Concentrations of soil and porewater acetate, nitrate, nitrite, and sulfate were determined via ion chromatography. For soils, 5 g of soil was mixed with 5 ml of MilliQ water (1:1 v/v), filtered with a 0.2 µm filter, and quantified using a Dionex ICS-2100 Ion Chromatography System with an AS18 column with standard curves performed for each anion. To more directly pair soil methane concentrations to microbiological soil data, in situ methane concentrations were calculated as described previously<sup>27</sup> following immediate transfer to 4 °C for transport and measurement on a Shimadzu GC-2014 gas chromatograph.

Soil porewater was extracted using methods and infrastructure previously described in detail from this wetland<sup>24, 25</sup>. Porewaters were then sent to the Pacific Northwest National Laboratory and metabolites were identified by proton nuclear magnetic resonance (<sup>1</sup>H NMR) as described previously<sup>28</sup>. Metabolomic responses were characterized using the EMSL 800 MHz and 600 MHz NMR spectrometers equipped with cryogenically cooled triple resonance probes for their high sensitivity and quantitation determined via <sup>1</sup>H NMR metabolite libraries (presently ~1000 metabolites). The two-dimensional (2D) NMR metabolomics methods including <sup>1</sup>H-<sup>13</sup>C correlation experiments (heteronuclear single quantum coherence), and connectivity experiments total correlation spectroscopy and correlation spectroscopy on a subset of samples (< 8) to enhance metabolite identification. Geochemical and metabolite data were analyzed in relationship to methanogenesis activity by linear correlations determined via Pearson's correlation ( $p < 0.05$ ).

**Collection of dissolved gases and greenhouse gas emission.** Surface emissions were measured by non-steady-state chambers, with floating chambers used for measurements over open water. Chambers were measured in duplicate in each ecosite and season<sup>24</sup> and sampling was coordinated to peeper measurement times. Additional greenhouse gas emissions were collected with an eddy covariance and meteorological station (3 m tall tower, site-wide footprint). We have previously shown that both chambers and eddy covariance measurements provide congruent measurements<sup>24</sup>.

Porewater dialysis samplers (peepers) were used to sample for dissolved CH<sub>4</sub>, CO<sub>2</sub>, and N<sub>2</sub>O below ground monthly, with a vertical resolution of 2.8 cm, throughout the upper 56 cm of soils with minimal disturbance to the soil<sup>29–31</sup>. Hydrogen was not measured in porewater from the dialysis peepers. The peepers feature 20 vertically stacked windows covered with a 0.1 µm dialysis membrane that allows the water inside the windows to equilibrate with dissolved gas concentrations outside. Gas concentrations in the peeper samples were quantified using a Shimadzu GC-2014 gas analyzer. The design and sample collection with both chambers and peepers followed protocols previously described<sup>24</sup>. Both chamber and peeper measurements were taken simultaneously, once a month during the 2015 growing season.

Temperature probes and co-located oxygen-level measurements (via a PreSens Fibox 4 handheld oxygen meter) provided vertical detail near each peeper measurement location. The oxic horizon was determined by fitting a reverse exponential curve to all DO data collected per patch per month. The horizon was taken as where the curve crossed 20 (µmol O<sub>2</sub> per kg H<sub>2</sub>O)<sup>32</sup>. To determine the upper and lower bounds of the horizon, we identified the soil depths at which the 2.5th and 97.5th% confidence intervals of the exponential fit crossed the same 20 (µmol O<sub>2</sub> per kg H<sub>2</sub>O) threshold. The oxic horizon and confidence bounds were interpolated linearly between measurement periods.

**Transport and production model.** A numerical model was used to combine chamber and peeper measurements to determine the rates of methane production/oxidation in different layers of the wetland soil. A diffusion model was separately created for mud and open ecosites (not plants), due to the complexity of including plant transport and roots. We discretized Fick's 2nd Law (Eq. 1) in one dimension using an implicit backwards Euler method to account for diffusive transport within the soil column. A production/oxidation term was included to account for the implied biological activity.

$$\frac{dC(t, z)}{dt} = \frac{d}{dz} \left( D(t, z) \frac{dC(t, z)}{dz} \right) + R(t, z) \quad (1)$$

Here, C is the soil porewater concentration of methane, z is the vertical position in the soil column, D is the temperature-dependent diffusion coefficient, t is the time in days, and R is a methane sink/source (generation/oxidation) term. The temperature profile was determined through measurements made with nearby soil temperature probes. A Neumann no-flux boundary condition was prescribed at the bottom of the soil column. We used a known flux top boundary condition (implemented by discretizing Fick's 1st Law) which was prescribed based on time

interpolated chamber measurements. Each month's measured concentration profile was used to model the next month's first using an ignorant guess of R (determined by solving the above with a month-long time step). We then refined the time step to 0.1 days and used a Markov Chain Monte Carlo Metropolis Hastings (MCMC-MH)<sup>33, 34</sup> approach with 40,000 repetitions to alter the value of R along the vertical column in order to minimize the error between the modeled future methane concentration profile and its measured value. We took the average of the 10% best performing MCMC runs as the microbial activity. Uncertainty was quantified as 1 s.d. of the 4000 selected runs. This simplistic model interprets observational concentration data as production/oxidation with no assumptions about the oxic conditions of the soil, providing a unique way of observing the data.

**Eddy covariance collection and data processing.** Eddy covariance data were collected from July to October 2015. The flux calculation approach was fully outlined previously<sup>35, 36</sup>. Briefly, a 3D rotation was applied to wind observations to force the vertical and cross wind components gathered from the sonic anemometer (CSAT3, Campbell Scientific, Logan, UT, USA) to average to 0 for each half-hour<sup>37</sup>. To correct for the separation of the sensors, the time series of concentration measurements were shifted in time using the maximal-covariance approach<sup>38</sup>. CO<sub>2</sub> (net ecosystem exchange), CH<sub>4</sub>, and water vapor flux (latent heat flux (LE)) were corrected as previously described<sup>39</sup> to account for the effects of changes in the densities of dry air and water vapor. Frequency response corrections for LE and CH<sub>4</sub> fluxes, which are based on concentration measurements from open-path gas analyzers (LI-COR Bioscience, LI-7500 for water vapor and CO<sub>2</sub>, and LI-7700 for CH<sub>4</sub>) were calculated and validated as previously described<sup>38, 40</sup>. The absorbance spectrum of CH<sub>4</sub> is temperature dependent. We therefore combined an absorbance spectrum correction with the Webb-Pearman-Leuning correction as detailed in the LI-7700 manual (LI-COR, 2010). Day-night transition was calculated using shortwave radiation observations from the nearby meteorological station. Night was defined as when shortwave radiation dropped to < 10 W m<sup>-2</sup>. The standard empirical approach of defining a seasonal threshold value of friction velocity (u\*) that indicates an insufficient level of turbulent mixing was used to reject invalid data<sup>41</sup>. The minimum value allowed for a u\* threshold was 0.2 m s<sup>-1</sup>. The u\* filter was used for both CO<sub>2</sub> and CH<sub>4</sub> fluxes on the assumption that if the turbulence is sufficient to provide adequate mixing conditions for carbon flux measurements it will be sufficient to do the same for CH<sub>4</sub>. Eddy covariance flux data were gap-filled to Morin et al.<sup>36</sup> using the automated neural network approach, an expanded version of the method commonly used in flux sites<sup>36, 41, 42</sup>, introduced by Papale and Valentini<sup>43</sup>.

**Site-level methane budget.** To determine the site-level methane budget and the oxic production contribution, we used an expanded version of the fixed frame eddy covariance scaling methodology developed for mud and open ecosites<sup>23</sup>. Briefly, this method combines eddy covariance data with a footprint method and ecosite flux measurements collected using the chamber method to decompose the eddy covariance flux signal into its contributing parts by ecosite. We used the Detto footprint method<sup>44</sup>, which is a 2D expanded version of the Hsieh model<sup>45</sup>. We used monthly varying displacement height and roughness lengths to represent the *Typha* spp. growing around the tower. There were four relevant ecosites with footprint contributions to the eddy covariance tower: open water, *Typha* spp., *Nelumbo* spp., and mud flat.

$$F_{\text{tower}} = F'_{\text{op}}e_{\text{op}} + F'_{\text{ty}}e_{\text{ty}} + F'_{\text{ne}}e_{\text{ne}} + F'_{\text{mu}}e_{\text{mu}} \quad (2)$$

where F' of each ecosite indicates the relative flux strength of that ecosite at the landscape level (different than that provided by the chamber measurements), and e is the footprint contribution of that patch per day.

Chamber flux data determined the relative flux strengths of each ecosite compared to open water fluxes. This provided a solvable system of equations for the relative flux strength contributing to the eddy covariance tower (F') with a daily temporal resolution.

$$\begin{aligned} m_1 &= \frac{F_{\text{ty}}}{F'_{\text{op}}} = \frac{F'_{\text{ty}}}{F'_{\text{op}}} \\ m_2 &= \frac{F_{\text{ne}}}{F'_{\text{op}}} = \frac{F'_{\text{ne}}}{F'_{\text{op}}} \\ m_3 &= \frac{F_{\text{mu}}}{F'_{\text{op}}} = \frac{F'_{\text{mu}}}{F'_{\text{op}}} \end{aligned} \quad (3)$$

We interpolated each ecosite's percent of production in the oxic layers to daily values over the course of the study. We scaled these values to the site level by integrating spatially over the site-level footprint and temporally over all times we were able to model methane production for (Fig. 4).

$$f_{\text{oxic}} = \frac{\sum_{t=t_{\text{peeper},0}}^{t_{\text{peeper},\text{final}}} (f_{\text{op}}F'_{\text{op}}P_{\text{op,ox}} + f_{\text{mu}}F'_{\text{mu}}P_{\text{mu,ox}})}{\sum_{t=t_{\text{peeper},0}}^{t_{\text{peeper},\text{final}}} (f_{\text{op}}F'_{\text{op}} + f_{\text{mu}}F'_{\text{mu}})} \quad (4)$$

where  $f$  is the site-level percent area contribution of each ecosites and  $p$  is the percent generation in the oxic zone (determined by the diffusion model) of each ecosite.

**Metagenomic analyses.** Genomic DNA was extracted from triplicate 0.5 g of soil using a MoBio PowerSoil DNA Isolation Kit following the manufacturer's protocol. DNA from three representative Fall and Summer surface soil samples (Plant, Mud, and Water) were sequenced at The Ohio State University and The Joint Genome Institute using an Illumina Library creation kit (KAPA Biosystems) with solid-phase reversible immobilization size selection. The quantified libraries were then prepared for sequencing on the Illumina HiSeq 2500 sequencing platform utilizing a TruSeq Rapid paired-end cluster kit, v4. We obtained 304 Gbp of metagenomic sequencing (Supplementary Table 1). Sequence assembly generated ~3.8 Gbp of contiguous sequences > 5 kbp from the six surface soil metagenome samples. Fastq files were generated using CASSAVA 1.8.2. and were trimmed from both the 5' and 3' ends using Sickle, then each sample was assembled individually using IDBA-UD with default parameters as previously described<sup>28, 46</sup>. Scaffold coverage was calculated by mapping reads back to the assemblies using Bowtie2<sup>47</sup>.

Genes on scaffolds ≥ 5 kb were annotated as described previously<sup>48, 49</sup> by predicting open reading frames using MetaProdigal<sup>50</sup>. Called genes were compared using USEARCH<sup>51</sup> to KEGG, UniRef90, InterproScan<sup>52</sup> with single and reverse best hit (RBH) matches > 60 bits reported. The collection of annotations for a protein were ranked: Reciprocal best BLAST hits (RBH) with a bit score > 350 given the highest (A) rank, followed by reciprocal best blast hit to Uniref with a bit score > 350 (B rank), blast hits to KEGG with a bit score > 60 (C rank), and UniRef90 with a bit score > 60 (D rank). The next rank represents proteins that only had InterproScan matches (E rank). The lowest (E) rank comprises the hypothetical proteins, with only a prediction from Prodigal but a bit score < 60.

Assembled scaffolds were binned into genomes based on GC, coverage, and taxonomic affiliation across samples using ESOM and Metabat<sup>53</sup>. For each bin, genome completion was estimated based on the presence of core gene sets (highly conserved genes that occur in single copy) for Bacteria (31 genes) and Archaea (104 genes) using Amphora<sup>54</sup> using a method previously reported<sup>28</sup>. Taxonomic placement of the genome bins was based on phylogenies of 16S rRNA genes recruited from the bin and/or ribosomal protein analyses. Overages (gene copies > 1 per bin) indicating potential misbins, and discrepant GC and phylogenetic signal were used to manually remove potential contaminating scaffolds from the bins.

Genome-wide average nucleotide identity values from our five > 50% complete reconstructed *Methanothrix* genomes (Supplementary Table 2) and comparisons to three existing *Methanothrix* genomes were calculated from the Kostas lab calculator (<http://envo-omics.ce.gatech.edu>). Metabolic analyses were performed largely using the M1 genome (Supplementary Data 5) as a model and are included in the Supplementary Note 1. To identify unique gene features in reconstructed genomes that differ from isolate *Methanothrix* genomes, we created an ITEP<sup>55</sup> database and compared all genes via all vs. all blast, orthoMCL clustering and sqlite database generation. A cluster is defined as having bidirectional best hits based on a percent identity cutoff.

**Transcriptomic analyses.** Metatranscripts were performed on three cores sampled in two ecosites (plant and mud) during two seasons (Fall and Summer) ( $n = 12$ ), as these soils demonstrated the highest surface methanogenesis activity. Total RNA was extracted from selected soils previously analyzed by metagenomics using ~2 g of soil via MoBio PowerSoil Total RNA Isolation Kit and DNA Elution Kit following the manufacturer's instructions. For metatranscriptomics, RNA was processed according to JGI established protocols. Briefly, ribosomal RNA (rRNA) was removed using the Ribo-Zero rRNA Removal Kit. Stranded complementary DNA (cDNA) libraries were generated using the Illumina Truseq Stranded RNA LT kit. The rRNA-depleted RNA was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second-strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 8 cycles of PCR.

We obtained 462 Gbp of metatranscriptomic sequencing (Supplementary Table 1). The resulting ~150 bp nucleotide sequences were trimmed as described above (see metagenomic section) and separately mapped via Bowtie2<sup>46</sup> to three databases (the metagenomic scaffolds > 5 kb, *mcrA* gene database from pure cultures and metagenome scaffolds > 1 kb, the *Candidatus Methanothrix paradoxum* genomes) allowing for up to three mismatches. The normalized transcript abundance for each gene was calculated via Cufflinks<sup>56</sup> using the rescue method for multi-reads and the mapped data were reported as fragments per kilobase per million mapped (FPKM). FPKM values for *Methanothrix* genes (from transcripts mapped to genes on *Methanothrix* scaffolds) were averaged across the 12 samples and ranked, with genes assigned to functional categories manually.

The top 100 most abundant *Candidatus Methanothrix paradoxum* transcripts in each season (Fall and Spring) were individually determined from the six samples in each season (triplicate samples in plant and mud ecosites). The resulting mean Fall and Summer log10 FPKM gene transcript relative abundance were plotted, and gene functional category was manually assigned and depicted by color. For the most expressed genes in each functional category across both seasons, we also plotted the log10 FPKM values from all 12 samples using boxplots, with mean

transcript relative abundance shown by black line. The responses of genes highly transcribed in both seasons and present in 50% of samples per season were also reported in a heatmap constructed using R pheatmap function.

Quantitative reverse transcription PCR using random-primed cDNA synthesis followed by cDNA quantification was performed to quantify *mcrA* transcripts. Briefly, RNA was synthesized to cDNA via manufacturer's protocol with SuperScript III Reverse Transcriptase kit, RNase OUT recombinant ribonuclease inhibitor, and random primers. Primers for *mcrA*<sup>57</sup> were previously reported. The quantification protocol and conditions followed from Franchini et al.<sup>58</sup>, with a slight deviation in reaction mix including 1  $\mu$ l of cDNA, 10  $\mu$ l 2X SsoAdvanced Universal SYBR Green Supermix reaction buffer, 0.5  $\mu$ M of each mls and *mcrA*-rev primer, balanced with 20  $\mu$ l total volume with PCR grade H<sub>2</sub>O. Serial diluted *mcrA* genes amplified from *M. acetivorans* DNA were used as standards. Statistical differences in *mcrA* transcript copy number between depths were evaluated via ANOVA (d.f. = 17,  $p < 0.05$ ).

**Phylogenetic analyses.** Single gene and concatenated genes analyses were performed as previously described<sup>45</sup>. Reference data sets for the 15 ribosomal proteins chosen as single-copy phylogenetic markers (RpL2, 3, 4, 5, 6, 14, 15, 18, 22, and 24 and RpS 3, 8, 10, 17, and 19), small subunit ribosomal protein 3 (*rps3*), and *mcrA* were created using sequences mined from the NCBI and Joint Genome Institute Integrated Microbial Genomes/Microbiome (JGI-IMG/M) databases (December 2016). Each individual protein data set was aligned using MUSCLE 3.8.31 and then manually curated to remove end gaps<sup>59</sup>. For amino acid phylogenetic analyses (S3 and concatenated ribosomal trees), we used ProtPipelin, a python script developed in-house for generation of phylogenetic trees (<https://github.com/lmsolden/protpipelin>). The pipeline runs as follows: alignments are curated with minimal editing by GBLOCKS<sup>60</sup>, and model selection conducted via ProtTest 3.4<sup>61</sup>. A maximum likelihood phylogeny for the concatenated alignment was conducted using RAxML version 8.3.1 under the LG model of evolution with 100 bootstrap replicates<sup>62</sup> and visualized in iTOL<sup>63</sup>. For the *mcrA* nucleotide tree, a similar analysis was used except under the GTRCAT model.

The V4 portion of 16S rRNA genes was sequenced at Argonne National Laboratory at the Next Generation Sequencing facility with a single lane of Illumina MiSeq using 2 × 251 bp paired-end reads and analyzed as described previously<sup>44</sup>. The full-length 16S rRNA sequence recovered from the *Ca. Methanothrix paradoxum* M1 genome was used to recover a single *Methanothrix* OTU from the V4 data set that was > 99% similar to the metagenomic recovered sequence. Additionally, the 16S rRNA fragment sequence from the M1 genome was searched against GenBank using BLASTN (e-value 1e-10, 100,000 alignments, 100,000 descriptions). Hits (> 500 bp) of at least 99% identity were retained. Genbank records for these sequences were parsed to find study title, and sequence source environment was assigned from the title and summarized (Supplementary Data 4). A maximum likelihood phylogenetic tree of *Candidatus Methanothrix paradoxum* 16S rRNA gene sequences was built using RAxML 8 in ARB<sup>64</sup> using all Silva SSURef NR99 (v128) *Methanothrix* sequences ≥ 900 bp, the M1 16S rRNA fragment sequence from this study, with *Methanoscincus* reference sequences as the outgroup. Additional reference sequences from the BLAST search above were aligned using SINA<sup>65</sup> and added to the tree using the parsimony add function in ARB.

**Data availability.** Amplicon sequencing and metagenomic sequencing data have been deposited in the sequence read archive under BioProject PRJNA338276. The draft *Candidatus Methanothrix paradoxum* genomes have been deposited in NCBI GenBank under accession numbers SAMN05908755, SAMN05908754, SAMN05908753, SAMN05908752, SAMN05908749, and SAMN05908746, with corresponding transcriptomic data also released. Flux data have been released to Ameriflux, site ID US-OWC (<https://ameriflux.lbl.gov/sites/siteinfo/US-OWC>) with additional meteorological and data available via National Estuarine Research Reserve System Centralized Data Management Office (<http://cdmo.baruch.sc.edu/get/landing.cfm>).

FASTA files used to construct all phylogenetic analyses in this manuscript are included as Supplementary Data 6–10. Necessary scripts and analyses to perform metagenome assembly, EMIRGE, annotation, and single-copy genes can be accessed from github ([https://github.com/TheWrightonLab/metagenome\\_analyses](https://github.com/TheWrightonLab/metagenome_analyses)). Code is available from individual repositories: [https://github.com/TheWrightonLab/metagenome\\_emirge/](https://github.com/TheWrightonLab/metagenome_emirge/); [https://github.com/TheWrightonLab/metagenome\\_assembly/](https://github.com/TheWrightonLab/metagenome_assembly/); <https://bitbucket.org/berkeleylab/metabat>.

Received: 7 April 2017 Accepted: 11 October 2017

Published online: 16 November 2017

## References

- Bodelier, P. L. & Dedysh, S. N. Microbiology of wetlands. *Front. Microbiol.* **4**, 1–4 (2013).

2. Bridgman, S. D., Cadillo-Quiroz, H., Keller, J. K. & Zhuang, Q. Methane emissions from wetlands: biogeochemical, microbial, and modeling perspectives from local to global scales. *Glob. Chang. Biol.* **19**, 1325–1346 (2013).
3. Cedervall, P. E. et al. Structural insight into methyl coenzyme M reductase chemistry using coenzymes B analogues. *Biochemistry* **49**, 7683–7693 (2010).
4. Bethke, C. M., Sanford, R. A., Kirk, M. F., Jin, Q. & Flynn, T. M. The thermodynamic ladder in geomicrobiology. *Am. J. Sci.* **311**, 183–210 (2011).
5. Riley, W. J. et al. Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. *Biogeosciences* **8**, 1925–1953 (2011).
6. Bogard, M. J. et al. Oxic water column methanogenesis as a major component of aquatic CH<sub>4</sub> fluxes. *Nat. Commun.* **5**, 5350 (2014).
7. Grossart, H. P., Frindte, K., Dziallas, C., Eckert, W. & Tang, K. W. Microbial methane production in oxygenated water column of an oligotrophic lake. *Proc. Natl. Acad. Sci. USA* **108**, 19657–19661 (2011).
8. Tang, K. W., McGinnis, D. F., Ionescu, D. & Grossart, H. P. Methane production in oxic lake waters potentially increases aquatic methane flux to air. *Environ. Sci. Technol. Lett.* **3**, 227–233 (2016).
9. Teh, Y. A., Silver, W. L. & Conrad, M. E. Oxygen effects on methane production and oxidation in humic tropical forest soils. *Glob. Chang. Biol.* **11**, 1283–1297 (2005).
10. Angel, R., Matthies, D. & Conrad, R. Activation of methanogenesis in arid biological soil crusts despite the presence of oxygen. *PLoS ONE* **6**, e20453 (2011).
11. Luton, P. E., Wayne, J. M., Sharp, R. J. & Riley, P. W. The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. *Microbiology* **148**, 3521–3530 (2002).
12. Großkopf, R., Janssen, P. H. & Liesack, W. Diversity and structure of the methanogenic community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S rRNA gene sequence retrieval. *Appl. Environ. Microbiol.* **64**, 960–969 (1998).
13. Schwarz, J. I., Eckert, W. & Conrad, R. Community structure of archaea and bacteria in a profundal lake sediment Lake Kinneret (Israel). *Syst. Appl. Microbiol.* **30**, 239–254 (2007).
14. Brioukhannov, A. L., Netrusov, A. I. & Eggen, R. I. L. The catalase and superoxide dismutase genes are transcriptionally up-regulated upon oxidative stress in the strictly anaerobic archaeon *Methanosarcina barkeri*. *Microbiology* **152**, 1671–1677 (2006).
15. Tholen, A., Pester, M. & Brune, A. Simultaneous methanogenesis and oxygen reduction by *Methanobrevibacter cuticularis* at low oxygen fluxes. *FEMS Microbiol. Ecol.* **62**, 303–312 (2007).
16. Liu, C.-T., Miyaki, T., Aono, T. & Oyaizu, H. Evaluation of methanogenic strains and their ability to endure aeration and water stress. *Curr. Microbiol.* **56**, 214–218 (2008).
17. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
18. Morris, J. J., Lenski, R. E. & Zinser, E. R. The black queen hypothesis: evolution of dependencies through adaptive gene loss. *mbio* **3**, e00036–12 (2012).
19. Repeta, D. J. et al. Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat. Geosci.* **9**, 884–887 (2016).
20. Fenchel, T. & Finlay, B. J. Production of methane and hydrogen by anaerobic ciliates containing symbiotic methanogens. *Arch. Microbiol.* **157**, 475–480 (1992).
21. Narrowe, A. B. et al. High-resolution sequencing reveals unexplored archaeal diversity in methane-emitting freshwater wetland soils. *Environ. Microbiol.* **19**, 2192–2209 (2017).
22. Jetten, M. S., Stams, A. J. & Zehnder, A. J. Acetate threshold values and acetate activating enzymes in methanogenic bacteria. *FEMS Microbiol. Lett.* **73**, 339–344 (1990).
23. Kemnitz, D., Chin, K.-J., Bodelier, P. & Conrad, R. Community analysis of methanogenic archaea within a riparian flooding gradient. *Environ. Microbiol.* **6**, 449–461 (2004).
24. Morin, T. H. et al. Combining eddy-covariance and chamber measurements to determine the methane budget from a small, heterogeneous urban floodplain wetland park. *Agric. For. Meteorol.* **237**, 160–170 (2017).
25. Chin, Y. P., Traina, S. J., Swank, C. R. & Backhus, D. Abundance and properties of dissolved organic matter in pore waters of a freshwater wetland. *Limnol. Oceanogr.* **43**, 1287–1296 (1998).
26. Stefanik, K. C. & Mitsch, W. J. Metabolism and methane flux of dominant macrophyte communities in created riverine wetlands using open system flow through chambers. *Ecol. Eng.* **72**, 67–73 (2014).
27. Chan, A. S. K. & Parkin, T. B. Methane oxidation and production activity in soils from natural and agricultural ecosystems. *J. Environ. Qual.* **30**, 1896–1903 (2001).
28. Daly, R. A. et al. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat. Microbiol.* **1**, 16146 (2016).
29. MacDonald, L. H., Paull, J. S. & Jaffé, P. R. Enhanced semipermanent dialysis samplers for long-term environmental monitoring in saturated sediments. *Environ. Monit. Assess.* **185**, 3613–3624 (2013).
30. Reid, M. C. et al. Tidal marsh methane dynamics: difference in seasonal lags in emissions driven by storage in vegetated versus unvegetated sediments. *J. Geophys. Res. Biogeosci.* **118**, 1802–1813 (2013).
31. Pal, D. S., Reid, M. C. & Jaffé, P. R. Impact of Hurricane Sandy on CH<sub>4</sub> released from vegetated and unvegetated wetland microsites. *Environ. Sci. Technol. Lett.* **1**, 372–375 (2014).
32. Wright, J. J., Kishori, M. K. & Hallam, S. J. Microbial ecology of expanding oxygen minimum zones. *Nat. Rev. Microbiol.* **10**, 381–394 (2012).
33. Haario, H., E. Saksman, E. & Tamminen, J. An adaptive metropolis algorithm. *Bernoulli* **7**, 223–242 (2001).
34. Haario, H., Laine, M., Mira, A. & Saksman, E. DRAM: efficient adaptive MCMC. *Stat. Comput.* **16**, 339–354 (2006).
35. Morin, T. H. et al. Environmental drivers of methane fluxes from an urban temperate wetland park. *J. Geophys. Res. Biogeosci.* **119**, 2188–2208 (2014).
36. Morin, T. H. et al. The seasonal and diurnal dynamics of methane flux at a created urban wetland. *Ecol. Eng.* **72**, 74–83 (2014).
37. Lee, X., Finnigan, J. & Paw U. K. T. in *Handbook of Micrometeorology, A Guide for Surface Flux Measurement and Analysis* (eds Lee, X. et al.) 33–66 (Kluwer Academic Publishers, Dordrecht, the Netherlands, 2004).
38. Detto, M., Verfaillie, J., Anderson, F., Xu, L. & Baldocchi, D. Comparing laser-based open- and closed-path gas analyzers to measure methane fluxes using the eddy covariance method. *Agric. For. Meteorol.* **151**, 1312–1324 (2011).
39. Webb, E. K., Pearman, G. I. & Leuning, R. Correction of flux measurements for density effect due to heat and water vapour transfer. *Q. J. Royal Meteorol. Soc.* **106**, 85–100 (1980).
40. Massman, W. J. A simple method for estimating frequency response corrections for eddy covariance systems. *Agric. For. Meteorol.* **104**, 185–198 (2000).
41. Reichstein, M. et al. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Glob. Chang. Biol.* **11**, 1424–1439 (2005).
42. Moffat, A. M. et al. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agric. For. Meteorol.* **147**, 209–232 (2007).
43. Papale, D. & Valentini, A. A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. *Glob. Chang. Biol.* **9**, 525–535 (2003).
44. Detto, M., Montaldo, N., Albertson, J. D., Mancini, M. & Katul, G. Soil moisture and vegetation controls on evapotranspiration in a heterogeneous Mediterranean ecosystem on Sardinia, Italy. *Water. Resour. Res.* **42**, W08419 (2006).
45. Hsieh, C. I., Katul, G. & Chi, T. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. *Adv. Water Resour.* **23**, 765–772 (2000).
46. Solden, L. M. et al. New roles in hemicellulosic sugar fermentation for the uncultivated Bacteroidetes family BS11. *ISME J.* **11**, 691–703 (2016).
47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie2. *Nat. Methods* **9**, 357–359 (2012).
48. Brown, C. T. et al. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208–211 (2015).
49. Wrighton, K. C. et al. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337**, 1661–1665 (2012).
50. Hyatt, D., LoCascio, P. F., Hauser, L. J. & Uberbacher, E. C. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* **28**, 2223–2230 (2012).
51. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
52. Quevillon, E. et al. InterProScan: protein domains identifier. *Nucleic. Acids Res.* **33**(suppl 2), W116–W120 (2005).
53. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
54. Wu, M. & Scott, A. J. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**, 1033–1034 (2012).
55. Benedict, M. N., Henricksen, J. R., Metcalf, W. W., Whitaker, R. J. & Price, N. D. ITEP: an integrated toolkit for exploration of microbial pan-genomes. *BMC Genomics* **15**, 8 (2014).
56. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
57. Steinberg, L. M. & Regan, J. M. mcrA-targeted real-time quantitative PCR method to examine methanogen communities. *Appl. Environ. Microbiol.* **75**, 4435–4442 (2009).
58. Franchini, A. G., Henneberger, R., Aepli, M. & Zeyer, J. Methane dynamics in an alpine fen: a field-based study on methanogenic and methanotrophic microbial communities. *FEMS Microbiol. Ecol.* **91**, fju032 (2015).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

60. Sasser, D. et al. Phylogenomic evidence for the presence of a flagellum and cbb (3) oxidase in the free-living mitochondrial ancestor. *Mol. Biol. Evol.* **28**, 3285–3296 (2011).
61. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
62. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
63. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
64. Ludwig, W. et al. ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
65. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).

## Acknowledgements

We are grateful to Kristin Arend, Frank Lopez, Dr Dave, and the rest of the OWC management team for site access and logistical support. We thank Kay Stefanik, Austin Rechner, Dominique Haddad, Sharon Acosta, and Chante Vines for their assistance in field measurements. This research was partially supported by Ohio Water Development Authority awards (6835 and 6560) to K.C.W and G.B. K.C.W. and G.B. thank Ohio Supercomputer center for allocation time for the data processing. W.J.R. was supported by the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under Contract No. DE-AC02-05CH11231 as part of their Regional and Global Climate Modeling program. This material is based partially upon work supported by Early Career Award from the U.S. Department of Energy to K.C.W., Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018022. Graduate student support was provided by the National Sciences foundation via Graduate Research Fellowship (to G.J.S.) and a Doctoral Dissertation Improvement Grant (to T.H.M.), and by the Department of Energy Office of Science Graduate Research Program, Solicitation 2 (to T.H.M.). Porewater metabolites (NMR) were performed using EMSL, a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research supported by DOE contract No. DE-AC05-76RL01830. DNA and RNA sequencing was conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility that is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. Flux observations were conducted with support from the US Department of Energy's Office of Science, Ameriflux Management project.

## Author contributions

K.C.W. and J.C.A. designed the experiment. K.C.W., J.C.A., T.H.M., G.B., and W.J.R. authored the manuscript, all other authors edited the manuscript. Wetland soil geochemical and biological sampling was conducted by J.C.A., T.M.H., G.J.S., and M.A.B. Methane field emission and porewater dialysis measurements were performed by J.C.A., T.H.M., G.J.S., M.A.B., and C.R.S. Modeling and scaling efforts were performed by T.H. M., G.B., and W.J.R. NMR was conducted by D.W.H. Metagenomics and metatranscriptomics analyses were conducted by K.C.W., J.C.A., L.M.S., G.J.S., R.A.D., and C.S.M. *Candidatus Methanothrix paradoxum* biogeographical analyses were conducted by A.B. N. and C.S.M.

## Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-01753-4.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**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) 2017