

# Chemoautotrophic carbon fixation in thermokarst lakes on the Tibetan Plateau

Received: 9 January 2024

Accepted: 2 December 2025

Published online: 16 December 2025

 Check for updates

Futing Liu<sup>1,8</sup>, Luyao Kang<sup>2,3,4,8</sup>, Ziliang Li<sup>2,3</sup>, Josep Peñuelas<sup>5,6</sup>, Benjamin W. Abbott<sup>7</sup>, Weijie Xu<sup>2,3</sup>, Wei Zhou<sup>2,3,4</sup>, Xuning Liu<sup>2,3</sup>, Leiyi Chen<sup>2,3</sup>, Shuqi Qin<sup>2,3</sup>, Dianye Zhang<sup>2,3</sup>, Yunfeng Peng<sup>2,3</sup> & Yuanhe Yang<sup>2,3,4</sup> ✉

Dissolved organic carbon (DOC) derived from thermokarst lakes is usually considered to be prone to microbial degradation and releases substantial carbon dioxide to the atmosphere, potentially enhancing the positive permafrost carbon (C)-climate feedback. In contrast to this long-term standing view, here we show that dark C fixation exceeds DOC degradation in ~1/3 of the investigated thermokarst lakes on the Tibetan Plateau, based on the combination of large-scale water and sediment sampling across seasons and years, biodegradable DOC experiments and <sup>14</sup>C-labeling bicarbonate (NaH<sup>14</sup>CO<sub>3</sub>) assimilation incubation experiment. By employing qPCR, amplicon sequencing and metagenomic analyses, we find that microbial C fixation is mainly driven by nitrifying microorganisms via the Calvin-Benson-Bassham cycle carried out by the *cbbL* gene (encoding form I ribulose-1,5-bisphosphate carboxylase/oxygenase). These findings demonstrate that chemoautotrophic C fixation predominates in part of thermokarst lakes, which could partly offset C emissions upon permafrost thaw and thus weaken the positive permafrost C-climate feedback.

Abrupt permafrost thaw, also called thermokarst formation, could profoundly alter the hydrology, ecology, and biogeochemistry of the permafrost zone, and subsequently impact permafrost carbon (C) cycling<sup>1–4</sup>. Thermokarst lake, as a typical landform of abrupt permafrost thaw, covers an area of 1.3 million km<sup>2</sup>, occupying about 7% of the Arctic permafrost area<sup>2</sup>. The occurrence and widespread expansion of thermokarst lakes not only trigger substantial methane release<sup>5–7</sup>, but also transfer large quantities of dissolved organic C (DOC) from nearby soils to water bodies<sup>8,9</sup>. This DOC pool could be mineralized and emitted to the atmosphere as carbon dioxide (CO<sub>2</sub>), eventually aggravating climate warming<sup>10–12</sup>. A comprehensive understanding of DOC dynamics in thermokarst lakes is thus essential for projecting the direction and strength of permafrost C-climate feedback.

Theoretically, both decomposers (mineralizing C) and primary producers (fixing C) co-exist in lakes<sup>13</sup>. Previous studies, however, concentrated on the decomposition process, and reported that DOC derived from thermokarst lakes was labile to biological and photochemical degradation<sup>8,11,14–17</sup>, acting as a potential C source which could reinforce climate warming. By contrast, the assimilation of C by chemoautotrophs has mostly been ignored among the global change research community, given that the C fixation in aquatic ecosystems is traditionally considered to be driven by photosynthetic microorganisms or phytoplankton<sup>18</sup>. If chemoautotrophs exist and have a high C fixation rate, they could work together with photosynthetic organisms to partially offset CO<sub>2</sub> emissions from thermokarst lakes, thereby weakening the permafrost C-climate feedback. More importantly, this

<sup>1</sup>Key Laboratory of Forest Ecology and Environment of National Forestry and Grassland Administration, Ecology and Nature Conservation Institute, Chinese Academy of Forestry, Beijing, China. <sup>2</sup>State Key Laboratory of Forage Breeding-by-Design and Utilization; Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China. <sup>3</sup>China National Botanical Garden, Beijing, China. <sup>4</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>5</sup>CSIC, Global Ecology Unit CREAF-CSIC-UAB, Barcelona, Spain. <sup>6</sup>CREAF, Cerdanyola del Vallès, Barcelona, Spain. <sup>7</sup>Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT, USA. <sup>8</sup>These authors contributed equally: Futing Liu, Luyao Kang.

✉ e-mail: [yhyang@ibcas.ac.cn](mailto:yhyang@ibcas.ac.cn)

dark C fixation could be carried out continuously even at night, which takes two to three times longer to sequester C than photosynthesis. Nevertheless, up to now, dark C fixation by chemoautotrophs has not yet been reported in thermokarst lakes around the world.

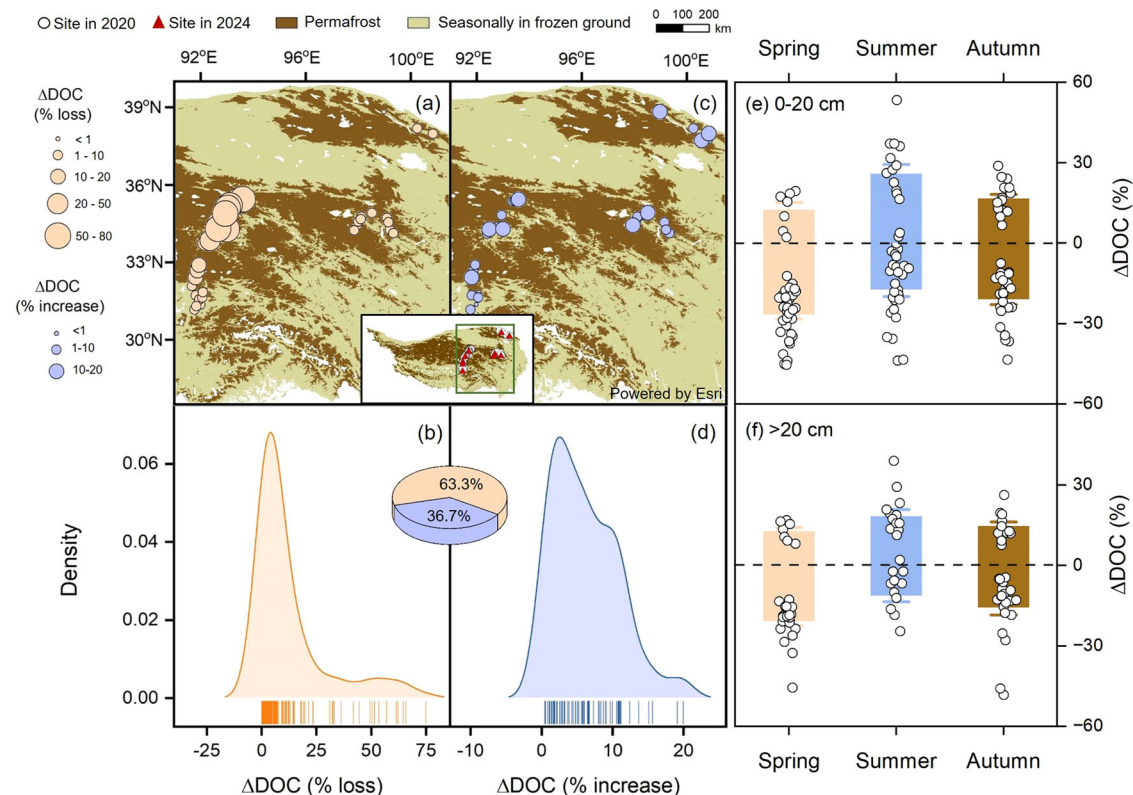
As the Earth's largest alpine permafrost region, the Tibetan Plateau has undergone rapid climatic warming and widespread development of thermokarst landscapes<sup>19,20</sup>. This area provides an ideal platform for exploring the DOC dynamics in thermokarst lakes. In this study, we identified chemoautotrophic C fixation in these aquatic ecosystems by sampling surface water (0–20 cm) in summer of 2020, and re-sampling the whole water column of thermokarst lakes in spring, summer and autumn of 2024, along a ~1100-km permafrost transect across the plateau (Fig. 1 and Supplementary Fig. 1). This study was designed to test the following two hypotheses. First, we hypothesized that, in partial thermokarst lakes, chemoautotrophs would have an advantage over heterotrophs<sup>13,21,22</sup> because of the oligotrophic environments in these lakes<sup>16</sup>. This could lead to a situation where chemoautotrophic C fixation exceeds DOC decomposition, leading to a net accumulation of DOC during the incubation period. Second, we hypothesized that the chemoautotrophic C fixation would be mainly carried out through the Calvin–Benson–Bassham (CBB) cycle rather than the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB)

pathway due to their different sensitivities to sunlight<sup>13</sup>. Particularly, although both are primary C fixation pathways for chemoautotrophs in oxic water, the shallow water depth and high elevation of thermokarst lakes across the Tibetan Plateau create high dissolved oxygen<sup>10</sup> and heavy light radiation (caused by high altitude)<sup>16,23</sup>, potentially favoring the CBB cycle. In consistent with these hypotheses, our results show that dark C fixation exceeds DOC degradation during the incubation across nearly one-third of the studied thermokarst lakes, and this dark C fixation is performed by chemoautotrophs mainly through the CBB cycle fueled by nitrification.

## Results and discussion

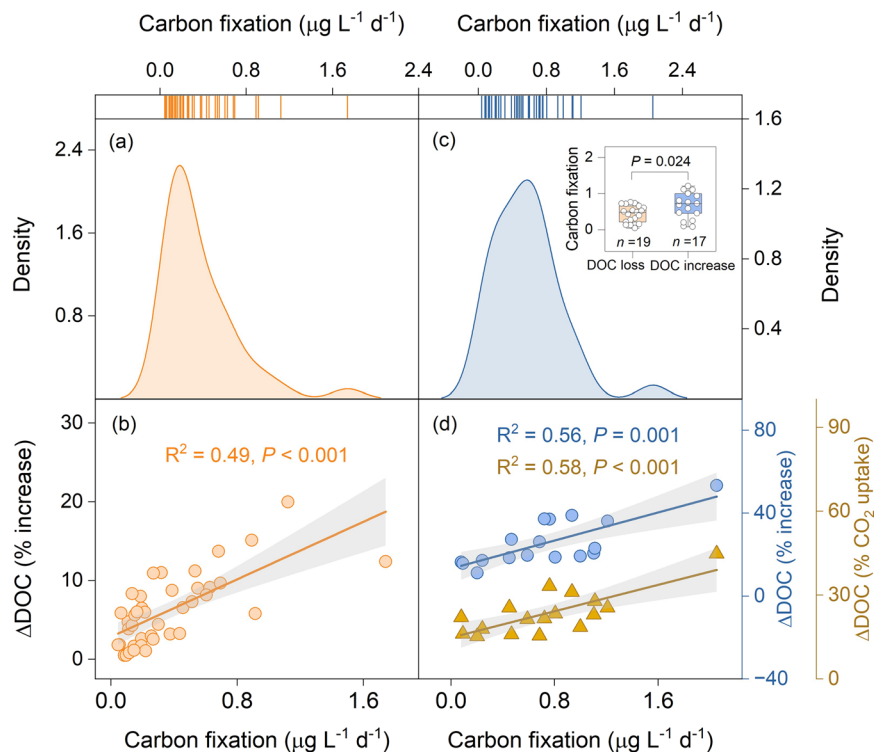
### Dark C fixation in thermokarst lakes

By conducting a 28-day laboratory incubation, we observed a net decline of DOC concentration in 119 of the 188 studied thermokarst lakes sampled in the summer of 2020. DOC loss ranged from 0.03% to 74.9%, with a mean of  $12.3 \pm 1.6\%$  (Fig. 1a, b; hereafter, values are shown as mean  $\pm$  standard error unless otherwise noted), suggesting that the DOC in these lakes was biodegraded<sup>8,11</sup>. Surprisingly, we found that the DOC concentration of the remaining 69 lakes showed an increasing pattern during the incubation period. DOC increase ranged from 0.5% to 20.0%, with an average of  $6.2 \pm 0.6\%$ , reflecting that about one-third



**Fig. 1 | Changes in dissolved organic carbon (DOC) of thermokarst lakes during the incubation across different seasons.** DOC biodegradation [ $\Delta$ DOC (% loss); **a**] and DOC accumulation [ $\Delta$ DOC (% increase); **c**] of surface water (0–20 cm) during a 28-day incubation across the alpine thermokarst lakes sampled in the summer of 2020. Circle size indicates the magnitude of DOC loss or accumulation. Insets in (**a**, **c**) indicate the distribution of 188 thermokarst lakes collected from 46 sites in the summer of 2020, in which 40 lakes re-collected from 10 sites in various seasons of 2024 were marked with a red triangle. The density distribution of  $\Delta$ DOC (% loss; **b**) and  $\Delta$ DOC (% increase; **d**) in surface water samples collected in the summer of 2020. The lines at the bottom of panels denote the values of  $\Delta$ DOC during incubation in each lake, in which the data are normally distributed. The pie chart shows the proportion of surface water from the 188 thermokarst lakes that experienced DOC loss or accumulation. The changed percentage of DOC in the surface

(0–20 cm; **e**) and subsurface water (>20 cm; **f**) of thermokarst lakes throughout spring, summer, and autumn in 2024 during a 28-day incubation. The dashed lines in (**e**, **f**) distinguish water samples in thermokarst lakes occurring with DOC accumulation and loss. Data are shown as mean  $\pm$  standard error. Sample sizes of water in thermokarst lakes collected in spring, summer, and autumn were 70, 64, and 72, respectively. Notably, considering that the depth of ~85% of thermokarst lakes was below 50 cm (Supplementary Table 3), we divided the investigated lakes into surface layer (0–20 cm) and sublayer (>20 cm) for data analysis. The permafrost map was drawn using ArcGIS Pro (Environmental Systems Research Institute, Inc., Redlands, CA, USA) powered by Esri (<https://www.esri.com>). The source data used in the permafrost map were derived from Zou et al.<sup>83</sup> (<https://tc.copernicus.org/articles/11/2527/2017/>)/CC BY (<https://creativecommons.org/licenses/by/3.0/>).



**Fig. 2 | Dark C fixation rates from the investigated thermokarst lakes and their associations with  $\Delta$ DOC (% increase) and  $\Delta$ DOC (%  $\text{CO}_2$  uptake).** Dark C fixation rates in surface water (0–20 cm) of thermokarst lakes collected in the summer of 2020 (a) and the relationship between dark C fixation and  $\Delta$ DOC (% increase) (b). Rates of dark C fixation in the whole water column of thermokarst lakes re-collected in the summer of 2024 (c) and their associations with  $\Delta$ DOC (% increase) and  $\Delta$ DOC (%  $\text{CO}_2$  uptake) (d). Inset in (c) shows the comparisons of dark C fixation rates among the thermokarst lakes occurring with DOC loss ( $n = 19$ ) and those with DOC accumulation ( $n = 17$ ). The whiskers in each box indicate the 5th and 95th quartiles, and the box ends denote the 25th and 75th quartiles (interquartile range). The

central line within the box represents the median value. Dark C fixation rates of thermokarst lakes were determined on the basis of  $^{14}\text{C}$ -labeling incubation<sup>25</sup>.  $\Delta$ DOC (% increase) and  $\Delta$ DOC (%  $\text{CO}_2$  uptake) denote the increased percentage of DOC during the incubation determined based on DOC<sup>11</sup> and  $\text{CO}_2$  measurements<sup>24</sup>, respectively. A one-way ANOVA with LSD multiple comparisons (two-sided) was used to test the statistical significance among the respective groups. The solid lines indicate the fitted ordinary least-squares model, with gray areas denoting the 95% confidence intervals. Notably, statistically significant relationships between  $\Delta$ DOC (% increase),  $\Delta$ DOC (%  $\text{CO}_2$  uptake), and dark C fixation rates remained stable even after removing the rightmost point in (d).

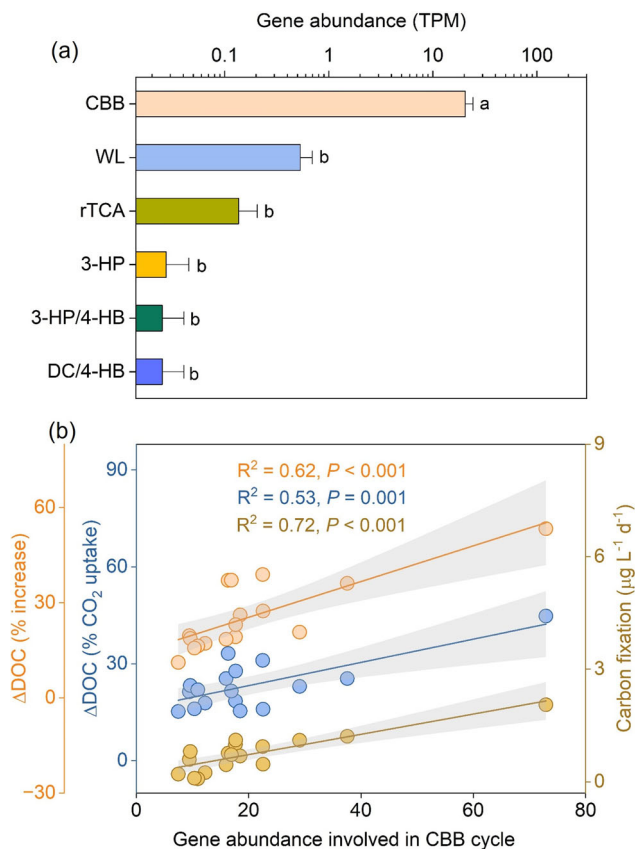
of thermokarst lakes accumulated rather than lost DOC during the incubation period (Fig. 1c, d). Our additional analyses based on water samples collected in the whole water column from 40 selected lakes across various seasons of 2024 revealed that this phenomenon occurred throughout the entire water column and the seasons. Over 28-day incubation, organic C in surface layer (0–20 cm) and sublayer (>20 cm) increased across 30.8% and 36.0% of water samples in thermokarst lakes, respectively (Fig. 1e, f). Likewise, organic C accumulated across 20.0%, 42.2% and 37.5% of lake samples in spring, summer and autumn, respectively (Fig. 1e, f). The bioincubation experiment based on  $\text{CO}_2$  measurements<sup>24</sup> confirmed that chemoautotrophic C fixation exceeded DOC decomposition in part of the studied thermokarst lakes (Supplementary Fig. 2), supporting our first hypothesis.

To further demonstrate the existence of microbial C fixation, we conducted a  $\text{NaH}^{14}\text{CO}_3$  assimilation experiment<sup>25</sup> for the 40 lakes (sampled in 2020) where DOC accumulation was detected over 28-day incubation and for the 20 lakes (sampled in 2024) where both DOC loss and accumulation were observed during the laboratory incubation. Our results revealed that dark C fixation occurred in the entire water column of all investigated lakes, including those suffering either DOC loss or accumulation, with rates ranging from 0.04 to 2.05  $\mu\text{g C L}^{-1} \text{d}^{-1}$ , while the rates in lakes experiencing DOC loss were significantly lower than those occurring with DOC accumulation ( $P = 0.024$ ;  $0.44 \pm 0.05$  vs.  $0.74 \pm 0.12 \mu\text{g C L}^{-1} \text{d}^{-1}$ ; Fig. 2a, c). The dark C fixation rates in thermokarst lakes were similar to or even higher than those reported in freshwater and groundwater, and also within the lower limit range in

marine ecosystems (Supplementary Table 1). Moreover, the dark C assimilation rates were positively associated with the increased percentage of DOC during biodegradable incubation regardless of the lakes sampled in 2020 and 2024 (all  $P < 0.01$ ; Fig. 2b, d), reflecting that DOC accumulation was indeed caused by microbial C fixation. In addition, chemoautotrophic C fixation existed in these lake sediments, with the rates ( $3.42$ – $25.70 \text{ mg C m}^{-2} \text{d}^{-1}$ ) being positively associated with those detected in water samples (Supplementary Fig. 3). These results illustrated that chemoautotrophic C fixation occurred in both the whole water column and lake sediments of the studied thermokarst lakes.

### Dark C fixation via the Calvin–Benson–Bassham pathway

To explore the chemoautotrophic communities and their pathway of C assimilation in thermokarst lakes, we analyzed the related functional genes involved in C fixation. Given that the pathways of chemoautotrophic C fixation under aerobic conditions mainly include the 3-HP/4-HB and CBB cycles<sup>26–28</sup>, we examined the key genes involved in these two pathways. Specifically, we used the related primers to amplify the *accA* and *hcd* functional genes in the 3-HP/4-HB pathway and the *cbbL* and *cbbM* genes involved in the CBB cycle<sup>13,29,30</sup>. Our results showed that the C fixation genes, including the *cbbL* and *cbbM*, were discovered (Supplementary Fig. 4), but neither the *accA* nor the *hcd* genes were detected in the studied thermokarst lakes, even though we employed multiple primers and changed PCR conditions (Supplementary Fig. 5 and Supplementary Table 2). Likewise, the



**Fig. 3 | Normalized abundances of functional genes involved in chemoautotrophic C fixation pathways and their associations with DOC accumulation and dark C fixation in thermokarst lakes during the incubation. a** Normalized abundances of genes involved in C fixation pathways of CBB, WL, rTCA, 3-HP, 3-HP/4-HB, and DC/4-HB. Data are shown as mean  $\pm$  standard error ( $n = 17$ ).

**b** Relationships of gene abundance in the CBB cycle with  $\Delta$ DOC (% increase),  $\Delta$ DOC (% CO<sub>2</sub> uptake), and dark C fixation rates in thermokarst lakes.  $\Delta$ DOC (% increase) and  $\Delta$ DOC (% CO<sub>2</sub> uptake) denote the increased percentage of DOC during the incubation, determined based on DOC<sup>11</sup> and CO<sub>2</sub> measurements<sup>24</sup> for water samples collected in the summer of 2024, respectively. Dark C fixation rates of thermokarst lakes were determined on the basis of <sup>14</sup>C-labeling incubation<sup>25</sup> for water samples collected in the summer of 2024. CBB Calvin–Benson–Bassham cycle, WL Wood–Ljungdahl pathway, rTCA reverse tricarboxylic acid cycle, 3-HP 3-hydroxypropionate bicycle, 3-HP/4-HB 3-hydroxypropionate/4-hydroxybutyrate cycle, DC/4-HB dicarboxylate/4-hydroxybutyrate cycle, TPM transcripts per million. Different lowercase letters denote significant differences ( $P < 0.05$ ) among the respective groups by one-way ANOVAs with LSD multiple comparisons (two-sided). The solid lines indicate the fitted ordinary least-squares model, with gray areas denoting the 95% confidence intervals. Notably, statistically significant relationships of functional genes in the CBB cycle with  $\Delta$ DOC (% increase),  $\Delta$ DOC (% CO<sub>2</sub> uptake), and dark C fixation in thermokarst lakes during the incubation remained stable even after removing the rightmost point in (b).

metagenomic sequencing analysis revealed that the functional genes related to the 3-HP/4-HB cycle were not observed in 92% of the examined water samples, and the abundance of functional genes in the 3-HP/4-HB pathway was much lower than those in the CBB cycle (Fig. 3a). Moreover, the abundance of genes involved in CBB cycle was positively associated with both the dark C fixation rates and increased proportion of DOC over 28-day incubation (Fig. 3b), while the functional genes related to the 3-HP/4-HB cycle had no significant association with either the dark C fixation rates or the increased percentage of DOC determined through bioincubation (Supplementary Fig. 6). Additionally, the functional genes involved in other anaerobic chemoautotrophic C fixation pathways, such as WL, rTCA, 3-HP, and DC/4-HB, were not only low, but also not related to the dark C fixation rates

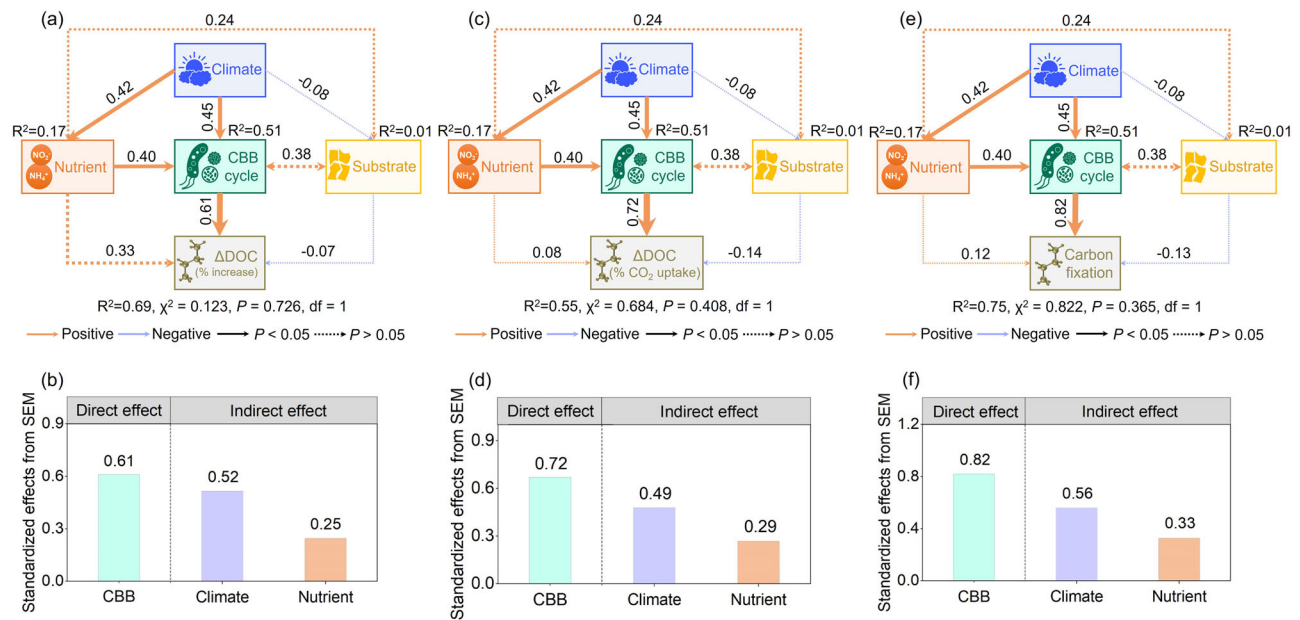
(Fig. 3a and Supplementary Fig. 6). These results demonstrated that the CBB cycle was the primary C fixation pathway of the studied thermokarst lakes, supporting our second hypothesis. Such a phenomenon could be ascribed to that chemoautotrophs engaged in the CBB cycle are less susceptible to photo-inhibition (thermokarst lakes involved in this study receive intense ultraviolet radiation caused by high altitude on the Tibetan Plateau) and thus widely distributed in aerobic water of lakes<sup>13</sup>.

Our results also indicated that the abundance of *cbbL* gene was more than ten times higher than that of the *cbbM* gene ( $P < 0.001$ ; Supplementary Fig. 4a). Similarly, metagenomic sequencing recovered 17 metagenome-assembled genomes (MAGs) associated with the CBB cycle, of which 16 encoded form I Rubisco (*cbbL*), and one encoded form II Rubisco (*cbbM*) (Supplementary Fig. 7). More importantly, the *cbbL* gene copies were positively related to both C fixation rates and  $\Delta$ DOC (% increase) (all  $P < 0.01$ ), while the *cbbM* gene abundance had no association with either of them (all  $P > 0.05$ ; Supplementary Fig. 4b, c). These results demonstrated that, as the primary CO<sub>2</sub> fixation pathway of the studied thermokarst lakes, the CBB cycle was mainly carried out by the *cbbL* rather than the *cbbM* genes. This situation could be explained by the different affinities for CO<sub>2</sub> between the *cbbL* and *cbbM* genes under aerobic conditions<sup>31</sup>. Generally, chemoautotrophs retain genes encoding forms I and II of Rubisco that allow efficient CO<sub>2</sub> fixation under different O<sub>2</sub> levels, because of their distinct affinities for CO<sub>2</sub> with O<sub>2</sub> present<sup>13,31</sup>. Specifically, the *cbbL* gene (encoding form I Rubisco) could function normally at high levels of O<sub>2</sub><sup>31</sup>. By comparison, the *cbbM* gene (encoding form II Rubisco) is adapted to functioning in low-O<sub>2</sub> conditions<sup>31</sup>, usually existing in oxic/anoxic boundaries of deep lakes<sup>13</sup>. In our case, water samples of the studied thermokarst lakes had a relatively higher dissolved oxygen (ranged from 2.0 to 17.0 mg L<sup>-1</sup>). Hence, the *cbbL* rather than the *cbbM* gene was more ready for CO<sub>2</sub> fixation through the CBB cycle under this condition<sup>13,31</sup>.

Our results further revealed that, besides microbial properties, the abiotic factors also significantly affected the DOC variations and C fixation of thermokarst lakes (all  $P < 0.05$ ; Supplementary Fig. 6). The structural equation model (SEM) analysis showed that either the increased proportion of DOC over 28-day incubation or the chemoautotrophic C fixation in thermokarst lakes were directly correlated with functional genes involved in CBB cycle (i.e., *cbbL*), with the standardized path coefficients of 0.61, 0.72 and 0.82, respectively (Fig. 4). Moreover, climate and nutrients were indirectly linked to DOC variations during the incubation and dark C fixation via affecting the functional genes of CBB cycle. Nevertheless, compared to climate, nutrients, and substrates, the functional genes involved in the CBB cycle (i.e., *cbbL*) were the dominant factor shaping variations in DOC accumulation and chemoautotrophic C fixation across thermokarst lakes on the Tibetan Plateau (Fig. 4).

### Chemoautotrophs dominated by nitrifiers

To identify the community composition of chemoautotrophs from thermokarst lakes, we analyzed the taxa encoding the *cbbL* gene using amplicon sequencing. Based on phylogenetic tree analysis for the core operational taxonomic units (OTUs) with abundance  $>0.1\%$ , our results illustrated that the number of species affiliating with *Nitrosospira* and *Nitrosomonas* genera and their relative abundance were the highest in the studied thermokarst lakes (Fig. 5a). At the order level, the *Nitrosomonadales* order was dominant, accounting for 51.1% of the C-fixing microbial community composition in thermokarst lakes (Fig. 5b). *Nitrosospira* and *Nitrosomonas* together made up 98.3% of the genus composition within the *Nitrosomonadales* order. Specifically, the *Nitrosospira* genus was mainly composed of *Nitrosospira lacus*, while the *Nitrosomonas* genus was primarily composed of *Nitrosomonas* sp. AL212, both of which are obligate chemoautotrophs fueled by nitrification<sup>13,32</sup>. Nitrifying bacteria could thus be the dominant



**Fig. 4 | Structural equation models (SEMs) assessing the effects of climate, nutrients, substrates, and chemoautotrophic genes involved in the CBB cycle on the microbial C fixation of thermokarst lakes collected in the summer of 2024 across the Tibetan Plateau.** SEMs for  $\Delta$ DOC (% increase; **a**) and standardized effects (**b**). SEMs for  $\Delta$ DOC (%  $\text{CO}_2$  uptake; **c**) and standardized direct and indirect effects (**d**); SEMs for the C fixation rates (**e**) and corresponding effects (**f**). The solid orange and blue arrows represent positive and negative relationships, respectively, while the dotted arrows denote an insignificant relationship. Arrow width is proportional to correlation strength, with adjacent numbers indicating standardized path coefficients. Climate, nutrients, and substrates denote the first component

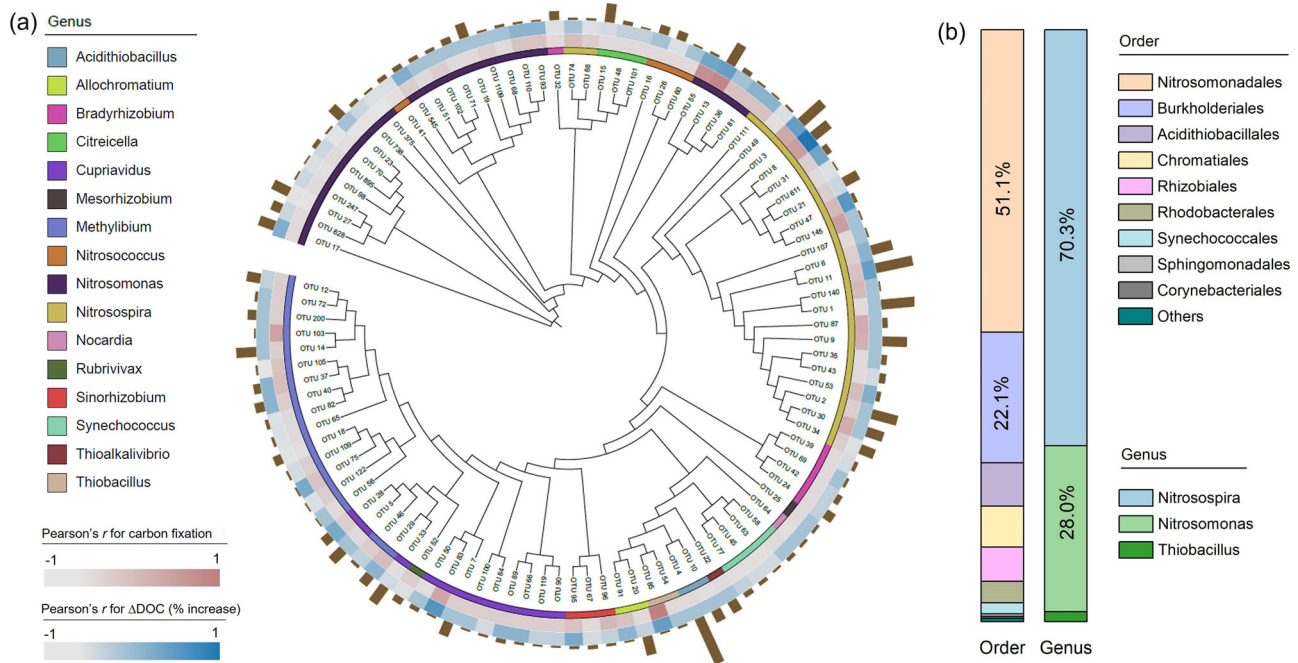
derived from PCA for the corresponding variables (Supplementary Table 4). The first components of climate, nutrients, and substrates could well represent mean annual temperature (MAT) and mean annual precipitation (MAP), ammonium ( $\text{NH}_4^+\text{-N}$ ) and nitrite ( $\text{NO}_2^-\text{-N}$ ), biological index (BIX) and humification index (HIX), respectively. Chemoautotrophic gene abundance involved in the CBB cycle was determined based on metagenomic sequencing.  $\Delta$ DOC (% increase) and  $\Delta$ DOC (%  $\text{CO}_2$  uptake) denote the increased percentage of DOC during the incubation determined based on DOC<sup>11</sup> and  $\text{CO}_2$  measurements<sup>24</sup>, respectively. Dark C fixation rates were determined on the basis of  $^{14}\text{C}$ -labeling incubation<sup>25</sup>. All the graphics used in SEMs were drawn by F.T. Liu.

microbial group in thermokarst lakes across the Tibetan Plateau. This deduction was supported by the following three lines of evidence: (1) the positive relationships of C fixation rates and  $\Delta$ DOC (% increase) with the relative abundance of OTUs affiliating *Nitrosospira* and *Nitrosomonas* genera (Fig. 5a). (2) Both concentrations of ammonium ( $\text{NH}_4^+\text{-N}$ ) and nitrite ( $\text{NO}_2^-\text{-N}$ ) were closely related to C fixation rates of thermokarst lakes (all  $P < 0.05$ ; Fig. 6a, b), illustrating that the first (from  $\text{NH}_4^+$  to  $\text{NO}_2^-$ ) and second (from  $\text{NO}_2^-$  to  $\text{NO}_3^-$ ) reaction steps of the nitrification provided the energy source for chemoautotrophic C fixation as an electron donor<sup>33</sup>. (3) Both the gene abundance of nitrification derived from metagenomic sequencing and the rates of nitrification based on  $^{15}\text{N}$ -isotope incubation experiments<sup>33</sup> were positively correlated with dark C fixation rates and the increased proportion of DOC during the incubation (Fig. 6c, d), confirming that nitrifying bacteria were the dominant chemoautotrophs driving dark C fixation in the studied thermokarst lakes.

The nitrifying bacteria-induced chemoautotrophic C fixation may be related to the high availability of  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_2^-\text{-N}$  in thermokarst lakes on the Tibetan Plateau. It has been reported that chemoautotrophs can obtain electrons through nitrogen ( $\text{NH}_4^+\text{-N}$  and  $\text{NO}_2^-\text{-N}$ ), sulfur (sulfide,  $\text{S}_2\text{O}_3^{2-}$  and  $\text{S}^0$ ), metal elements ( $\text{Fe}^{2+}$ ), and specific gases ( $\text{CO}$ ,  $\text{CH}_4$ , and  $\text{H}_2$ ) to carry out chemoautotrophic C fixation<sup>31,34</sup>. In our case, the concentrations of  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_2^-\text{-N}$  were much higher than those of reduced sulfur and metal elements within the studied thermokarst lakes (Supplementary Fig. 8), which may be due to the fact that sulfur and metal ions mainly exist in anaerobic and acidic water bodies<sup>35,36</sup>. Particularly, the alpine thermokarst lakes involved in this study are aerobic ecosystems with neutral or alkaline water (dissolved oxygen: 2.0–17.0  $\text{mg l}^{-1}$ ; pH: 7.4–10.8), thus resulting in less reduced sulfur and metallic elements. Moreover, given that  $\text{CO}$ ,  $\text{CH}_4$ , and  $\text{H}_2$  (mainly exist in anaerobic environments) are also lower in the water

with aerobic environments, both  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_2^-\text{-N}$  became the dominant electron donors for chemoautotrophs than the other reduced substrates mentioned above. Consequently, chemoautotrophs fixed C mainly through the electron donors provided by nitrification in the studied thermokarst lakes. Notably, we also detected the *Thiobacillus* (*Thiobacillus denitrificans*) and *Thioalkalivibrio* (*Thioalkalivibrio* sp.AKL11) genera (Fig. 5a), indicating that autotrophic organisms utilizing reduced sulfur as substrates for chemoautotrophic C fixation are present<sup>13</sup>. Nevertheless, given that the relative abundance of the above two types of chemoautotrophs was much lower than that of microorganisms engaged in nitrification (2.3% vs. 50.2%; Fig. 5a), nitrifying bacteria were likely the dominant chemoautotrophs in thermokarst lakes across the Tibetan Plateau.

Although this research advances our understanding about patterns and drivers of DOC dynamics in thermokarst lakes across the Tibetan Plateau, there remain some limitations that need to be tackled in future studies. First, owing to the logistical constraints of large-scale sampling and the challenges of preserving samples under remote and harsh field conditions, this study was limited to DNA-based analyses, including high-throughput sequencing and metagenomics, to characterize chemoautotrophs in the studied thermokarst lakes. While DNA-level data shed insights into the potential functional capacity of microbial communities, they do not directly capture in situ activity because the gene abundances are not always equal to the real activities of functional microbes. Second, despite the abundance of functional genes (i.e., *cbbl*) in the CBB cycle, which was closely associated with chemoautotrophic C fixation, statistical correlations do not necessarily reflect causal mechanisms. Therefore, to elucidate the underlying mechanisms of chemoautotrophs more rigorously, future studies should incorporate RNA-based approaches (metatranscriptomics) and DNA-stable isotope probing (DNA-SIP) that systematically explore the



**Fig. 5 | Phylogeny, relative abundances of chemoautotrophs encoding the *cbbL* gene, and their associations with microbial C fixation.** **a** The phylogenetic relationships of chemoautotrophic bacteria containing the *cbbL* gene. Only OTUs with relative abundances >0.1% were included in the tree. The color in the first ring represents the different genera of chemoautotrophs. The second and third rings show the relationships of C fixation rates and  $\Delta$ DOC (% increase) with the corresponding abundances of OTUs, where the variations of color indicate the strength of the correlations. The outside bar of the third ring corresponds to the relative

abundances of OTUs. **b** The first pillar shows the relative abundances of chemoautotrophs at the order level, and the second pillar shows the relative abundances of *Nitrosospora*, *Nitrosomonas*, and *Thiobacillus* genera within the *Nitrosomonadales* order. Dark C fixation was determined based on the  $\text{NaH}^{14}\text{CO}_3$  assimilation experiment<sup>25</sup> for water samples collected in the summer of 2020.  $\Delta$ DOC (% increase) denotes the increased percentage of DOC during the incubation, determined based on DOC concentration measurement<sup>11</sup> for water samples collected in the summer of 2020.

active microbial processes underpinning dark C fixation in these environments, and further illustrate whether *cbbL* takes advantage in the CBB cycle within the studied thermokarst lakes.

In summary, based on two kinds of biodegradable DOC experiments and  $\text{NaH}^{14}\text{CO}_3$  assimilation incubation, this study proved that DOC accumulation driven by chemoautotrophic C fixation predominated in nearly one-third of the investigated thermokarst lakes across the Tibetan Plateau. This C fixation was performed mainly through the CBB cycle rather than the 3-HP/4-HB pathway. Moreover, we found that the dark C fixation was primarily fueled by nitrification, and nitrifying bacteria were the dominant group of chemoautotrophs in these lakes. These findings provide two important implications for understanding the permafrost C cycle and its feedback to climate warming. First, in the traditional view, DOC in thermokarst lakes is highly biodegradable and vulnerable to being photo-degraded or decomposed by microorganisms, which is regarded as a hot spot with substantial  $\text{CO}_2$  release<sup>5,8,14,15</sup>. This study, however, challenges this traditional view, revealing that chemoautotrophic C fixation rather than DOC mineralization predominates in a portion of thermokarst lakes, possibly weakening the feedback between the permafrost C cycle and climate warming. This finding highlights the prominent role of microbially mediated C fixation in affecting DOC dynamics in partial thermokarst lakes, and provides insights for understanding the basic characteristics of the permafrost C cycle. Second, chemoautotrophs and their related functional genes were identified as the dominant factors affecting dark C fixation compared to other variables such as climate, nutrients, and substrates. This finding suggests that the specific functional genes and associated C fixation pathway exert a crucial role in mediating DOC dynamics in thermokarst lakes, which could then impact the permafrost C cycle. It is thus fundamental to incorporate the key C fixation genes and microbially mediated chemoautotrophic C fixation process into Earth system models to enhance the

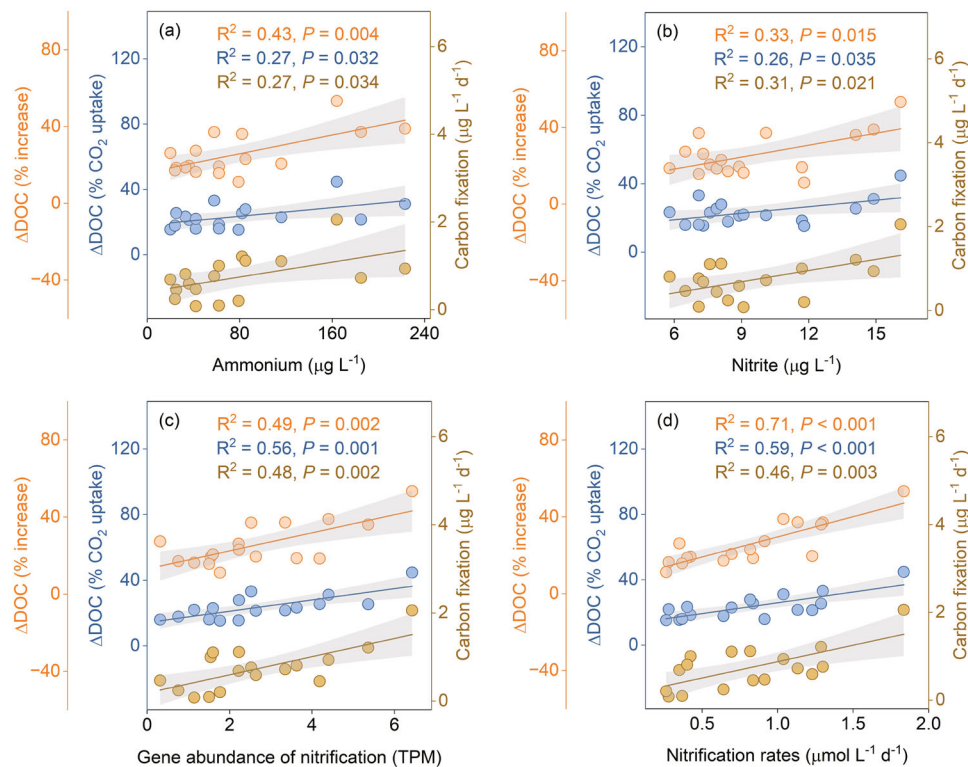
accuracy of projecting the direction and intensity of permafrost C-climate feedback.

## Methods

### Study sites and field sampling

The Tibetan Plateau, an important region of permafrost in the Northern Hemisphere, covers a permafrost area of 1.06 million  $\text{km}^2$ , and stores approximately 14.1–46.2 Pg (1 Pg =  $10^{15}$  g) of organic C within soils in the top 3 m<sup>37–40</sup>. The average annual temperature varies from  $-2.9$  to  $7.0$  °C, and the mean annual precipitation ranges from 129 to 590 mm<sup>7</sup>. The main vegetation types consist of alpine steppe, alpine meadow, and swamp meadow<sup>41</sup>, and the corresponding dominant species in these types of grassland are *Stipa purpurea* and *Carex moorcroftii*, *Kobresia pygmaea* and *K. humilis*, and *K. tibetica*, respectively<sup>42</sup>. The soil is classified as Cambisol according to the World Reference Base for Soil Resources. The mean active layer thickness is about 1.9 m on average across this permafrost area<sup>43</sup>. Climate warming has extensively thawed the permafrost in this region<sup>44</sup>, accompanied by an increase in the thickness of the active layer ( $-8.6$  cm year<sup>-1</sup> in the past decade)<sup>45</sup> and the widespread expansion of thermokarst landscapes<sup>19</sup>. Thermokarst lakes, a typical form of thermokarst landscape, are distributed across an area of approximately 2800  $\text{km}^2$ , where more than 160,000 lakes exist across the entire plateau<sup>19</sup>. These lakes are characterized by a small area and shallow depth, accounting for 4/5 of the total thermokarst lakes on the plateau<sup>19</sup>. The area and number of thermokarst lakes caused by climate warming across the Tibetan Plateau have increased by 83.1% and 58.8% respectively, during the past three decades<sup>46</sup>.

During July and August in 2019 and 2020, we collected water samples from 188 thermokarst lakes along a  $\sim$ 1100-km permafrost transect across the Tibetan Plateau<sup>47</sup>. To simplify the narrative, samples collected in both years are jointly termed the 2020 samples, given



**Fig. 6 | Associations of DOC accumulation and chemoautotrophic C fixation during the incubation with the abiotic and biotic variables involving the nitrification process in thermokarst lakes.** Relationships of  $\Delta$ DOC (% increase),  $\Delta$ DOC (% CO<sub>2</sub> uptake), and dark C fixation rates with ammonium (a), nitrite (b), gene abundance of nitrification (c), and microbial nitrification rates (d) in water samples of thermokarst lakes collected in the summer of 2024, respectively. Gene abundance of nitrification was examined through metagenomic sequencing.

Nitrification rates were determined based on <sup>15</sup>N-isotope incubation experiments<sup>33</sup>.  $\Delta$ DOC (% increase) and  $\Delta$ DOC (% CO<sub>2</sub> uptake) denote the increased percentage of DOC during the incubation determined based on DOC<sup>11</sup> and CO<sub>2</sub> measurements<sup>24</sup>, respectively. Dark C fixation rates were determined on the basis of <sup>14</sup>C-labeling incubation<sup>25</sup>. The solid lines indicate the fitted ordinary least-squares model, with gray areas denoting the 95% confidence intervals.

that the majority were obtained in that year. In each lake, we collected three surface (0–20 cm) water samples with the same volume from the shore to the center of the lake, and then mixed them into one sample. These mixed samples were then filtered via pre-combusted GF/F filters (0.7  $\mu$ m), stored in brown high-density polyethylene (HDPE) bottles at  $-20^{\circ}\text{C}$ , and transferred to the laboratory for subsequent analysis. To ensure the universality of our findings across various seasons and water depths, we further selected 40 of the 188 thermokarst lakes collected in summer of 2020, and re-collected water samples in the whole water column from these selected 40 lakes in spring (sampling from 25th May to 15th June), summer (sampling from 20th July to 26th August), and autumn (sampling from 5th October to 20th October) in 2024. Specifically, water samples were collected from the 0–20, 20–50, 50–80, and 80–110 cm water layers according to different lake depths (Supplementary Table 3). In spring, 70 samples were collected from these 40 lakes, 64 samples were collected in summer, and 72 samples were collected in autumn depending on the variations in lake depth, with a total of 206 water samples being collected. The procedures of collection, filtration, and storage for water samples obtained in 2024 were the same as those collected in the summer of 2020. While collecting water samples, we also sampled sediments of thermokarst lakes in the summer of 2024. In each lake, surface sediments (0–1 cm) from the shore to the center of the lakes were collected and then mixed as one sample for subsequent <sup>14</sup>C-labeling incubation (surface sediments are the area with the most active metabolism of chemoautotrophs and could represent the in situ dark C fixation<sup>48,49</sup>).

Apart from collecting water and sediment samples in the field, we took photographs of the lakes with an unmanned aerial vehicle (Phantom 4, DJI, Shenzhen, China) at each sampling time (including

2020 and 2024), and calculated the area of each lake using ImageJ (National Institutes of Health, Bethesda, USA). The altitude and geographic location of each site were recorded using a portable global positioning system (Jisibao G120, Beijing Hezhongshizhuang Technology, Beijing, China). Lake depth was determined using a measuring stick. Moreover, we used a Professional Plus (Pro Plus) multiparameter instrument (ProSolo Digital Water Quality Meter, Yellow Springs Instrument, Brannum Lane, USA) to examine in situ pH, dissolved oxygen, conductivity, and salinity within lakes. Across all the 40 selected thermokarst lakes, the lake area ranged from 16 to 2110 m<sup>2</sup>, and the average depth was about 35 cm (Supplementary Table 3)<sup>47</sup>. The pH varied from 7.4 to 10.8, and the concentration of dissolved oxygen ranged from 2.0 to 17.0 mg L<sup>-1</sup> among the studied thermokarst lakes<sup>47</sup>, indicating that these lakes are alkaline aerobic lakes.

### Biodegradable DOC experiment

A bioincubation experiment was conducted to examine the DOC dynamics in thermokarst lakes collected in the summer of 2020 (including 188 lakes from 46 sites) and throughout spring, summer, and autumn of 2024 (selecting 40 of the 188 lakes). The specific incubation procedure was conducted with the following two steps. First, 100 mL of filtered (0.7  $\mu$ m) water sample in each lake was placed into an HDPE bottle and then incubated at  $20^{\circ}\text{C}$  for 28 days (both parameters are recommended to conduct DOC incubation in permafrost-impacted aquatic systems, as these are the most common, relatively easy to maintain, and allow comparison among studies)<sup>11</sup>. Second, the incubated sample was swirled each day to ensure a sufficient supply of oxygen. To analyze the shifts in DOC concentration and its potential drivers, we re-filtered (0.7  $\mu$ m) water samples at day 0

( $t = 0$ ) and day 28 ( $t = 28$ ) and divided each sample into two parts: one of them was stored in a refrigerator at  $-20\text{ }^{\circ}\text{C}$  for the subsequent determination of DOC, ammonium ( $\text{NH}_4^+\text{-N}$ ), nitrite ( $\text{NO}_2^-\text{-N}$ ), nitrate ( $\text{NO}_3^-\text{-N}$ ), total dissolved nitrogen (TDN), total dissolved phosphorus (TDP), and metallic ions; the other part was refrigerated at  $4\text{ }^{\circ}\text{C}$  in the dark for subsequent spectral analysis. We then applied the modified method reported by ref. 11 to calculate the variations of DOC using Eq. (1):

$$\Delta\text{DOC}(\%) = (\text{DOC}_{t=28} - \text{DOC}_{t=0}) / \text{DOC}_{t=0} \times 100\% \quad (1)$$

where  $\text{DOC}_{t=0}$  and  $\text{DOC}_{t=28}$  represent the DOC concentration ( $\text{mg L}^{-1}$ ) before and after the 28-day incubation, respectively. It should be noted that, if  $\Delta\text{DOC}$  was positive, we defined it as  $\Delta\text{DOC}$  (% increase), reflecting the increased proportion of DOC during the incubation. By contrast, if  $\Delta\text{DOC}$  was negative, we used the absolute value of the data and defined it as  $\Delta\text{DOC}$  (% loss), which was adopted to denote the biodegradability of DOC within lakes. The concentration of DOC was examined using a multi C/N analyzer (Multi-N/C 3100, Analytik Jena AG, Jena, Germany).

To demonstrate that chemoautotrophs could fix  $\text{CO}_2$  from the atmosphere, we conducted another biodegradable DOC experiment using the modified method based on  $\text{CO}_2$  measurements<sup>24</sup>. Specifically, each lake sample collected in spring, summer, and autumn in 2024 was divided into six subsamples (9 mL each), and then placed into amber jars with airtight lids. Three of them were used as the control group collected at day 0. The remaining three subsamples were selected as the treatment group, which were incubated at  $20\text{ }^{\circ}\text{C}$  for 28 days in a dark environment and swirled every day. The headspace  $\text{CO}_2$  concentration in the jar before and after 28-day incubation was examined using a gas chromatograph (Agilent 7890A, Agilent Technologies, Santa Clara, USA) after adding 1 mL of  $\text{H}_3\text{PO}_4$  (40%) to translate dissolved inorganic C in water to headspace  $\text{CO}_2$ . Microbial  $\text{CO}_2$  productions were determined based on the variations of  $\text{CO}_2$  concentration in the headspace of jars examined on day 0 and day 28. The shifts in  $\text{CO}_2$  production before and after 28-day incubation could be regarded as the C respired or fixed by microbes. We then used the modified method reported by ref. 24 to calculate the variations of DOC using Eq. (2):

$$\Delta\text{DOC}_{\text{CO}_2}(\%) = \left[ (\text{C} - \text{CO}_2)_{t=0} - (\text{C} - \text{CO}_2)_{t=28} \right] / \text{DOC}_{t=0} \times 100\% \quad (2)$$

where  $(\text{C} - \text{CO}_2)_{t=0}$  and  $(\text{C} - \text{CO}_2)_{t=28}$  represent the amount of C converted by  $\text{CO}_2$  productions on day 0 and day 28 in the jars, respectively.  $\text{DOC}_{t=0}$  represents the DOC concentration ( $\text{mg L}^{-1}$ ) before the 28-day incubation. If  $\Delta\text{DOC}_{\text{CO}_2}$  was positive, we defined it as  $\Delta\text{DOC}$  (%  $\text{CO}_2$  uptake), which was then used to denote the increased proportion of organic C via fixing  $\text{CO}_2$  during the incubation. If  $\Delta\text{DOC}_{\text{CO}_2}$  was negative, we defined it as  $\Delta\text{DOC}$  (%  $\text{CO}_2$  release), reflecting the proportion of DOC loss via microbial degradation during the incubation.

### Radiolabeled bicarbonate ( $\text{NaH}^{14}\text{CO}_3$ ) incubation

To verify the existence of microbial C fixation in thermokarst lakes, we employed  $\text{NaH}^{14}\text{CO}_3$  (Hartmann Analytic GmbH, Braunschweig, Germany) to trace microbial  $\text{CO}_2$  assimilation, and then calculated the C fixation rate for each lake<sup>25</sup>. Considering the high experimental cost, we randomly selected 40 of the 69 thermokarst lakes, with the occurrence of DOC accumulation determined in the summer of 2020. To elucidate whether the dark C fixation occurred in the whole water column and in the lakes experiencing DOC loss, we further randomly selected 20 of the 40 thermokarst lakes collected in the summer of 2024 for  $^{14}\text{C}$ -labeled incubation. In these 20 selected lakes, we obtained a total of 36 water samples (17 of them experienced DOC accumulation, and the others suffered DOC loss during the bioincubation) in the

entire water column of thermokarst lakes and traced chemoautotrophic  $\text{CO}_2$  assimilation as follows. First, each lake sample was divided into five subsamples (20 mL each) and then placed into 50-mL tubes. Three of them were selected as the treatment group, and the remaining two subsamples were used as the control group by adding formaldehyde (2%, final concentration) to prevent microbial C fixation. Second,  $\text{NaH}^{14}\text{CO}_3$  was added to each tube ( $0.25\text{ }\mu\text{Ci mL}^{-1}$ , final concentration)<sup>27</sup>. Then, 20  $\mu\text{L}$  water sample taken from each tube was put into a scintillation vial to measure the total radioactivity. Third, all samples labeled with  $\text{NaH}^{14}\text{CO}_3$  (including the treatment and control groups) were incubated at  $20\text{ }^{\circ}\text{C}$  for 48 h under dark and aerobic conditions. The incubated samples were filtered (0.22  $\mu\text{m}$  Sterivex GP, Millipore), and then fumigated with hydrochloric acid for 2 h<sup>25</sup>. Disintegrations per minute (dpm) was measured for each filter using a liquid scintillation counter (Hidex 300 SL, Hidex Oy, Turku, Finland). The C fixation rate was then calculated according to Eq. (3) reported by the ref. 25:

$$\text{Carbon fixation} = \frac{\text{dpm}_{tr} - \text{dpm}_{cont} \times \text{DIC}}{\text{dpm}_i \times t} \quad (3)$$

where  $\text{dpm}_{tr}$  and  $\text{dpm}_{cont}$  denote the dpm of filters obtained from the treatment and control water samples, respectively.  $\text{dpm}_i$  indicates the dpm counts in the initial water samples. DIC represents the ambient concentration ( $\mu\text{g L}^{-1}$ ) of dissolved inorganic C in each sample, and  $t$  indicates the incubation time (days).

We also conducted a  $\text{NaH}^{14}\text{CO}_3$  assimilation experiment<sup>48</sup> for the surface lake sediments (0–1 cm). To constrain experimental cost, we selected sediment samples from 20 of the 40 thermokarst lakes (same as those used for  $^{14}\text{C}$ -labeling incubation of water samples collected in 2024). For each selected lake, 3 g of sediment was mixed with 3 mL of overlying water filtered through 0.22  $\mu\text{m}$  filters<sup>48</sup>. This obtained slurry was evenly separated into five parts (1 mL each) and poured into 2-mL tubes. Two of them were added with formaldehyde (2%, final concentration) as the control group, and another three were regarded as the treatment group. Both groups were added with  $\text{NaH}^{14}\text{CO}_3$  solution to a final concentration of  $\sim 2\text{ }\mu\text{Ci mL}^{-1}$ , and a partially labeled slurry (10  $\mu\text{L}$ ) was taken out for the determination of total radioactivity<sup>48</sup>. Under dark and aerobic conditions, all of the samples were incubated at  $20\text{ }^{\circ}\text{C}$  for 48 h. Afterwards, the incubated slurry was added with 0.5 mL of 3 M HCl, and purged with compressed nitrogen to remove remaining  $\text{D}^{14}\text{C}$  completely. Finally, this dried slurry was mixed with 9 mL of pure water and scintillation cocktail (Insta-Gel plus, Perkin Elmer, Waltham, USA), which was then used to examine dpm with a liquid scintillation counter (Hidex 300 SL, Hidex Oy, Turku, Finland). The C fixation rates in sediments were then calculated according to Eq. (3) mentioned above.

### Chemoautotrophic community and function analyses

**DNA extraction.** To determine the characteristics of chemoautotrophs, we extracted DNA from the water samples collected from the 40 thermokarst lakes in the summer of 2020. Specifically, 700 mL of water was put into an HDPE bottle and incubated at  $20\text{ }^{\circ}\text{C}$  for 28 days under aerobic and dark conditions (consistent with the biodegradable DOC experiment). The water was then filtered using a 0.22  $\mu\text{m}$  polyethersulfone membrane (Sterivex GP, Millipore, Boston, USA). We collected these filters and then extracted DNA from each filter. DNA concentration and quality were examined using a NanoDrop-2000 (Thermo Fisher Scientific, Waltham, USA).

**qPCR and high-throughput sequencing.** After the DNA extraction, we employed real-time fluorescence quantitative PCR to quantify the abundances of *cbbL* and *cbbM*, which are key genes encoding Rubisco enzymes for C fixation based on the CBB cycle and widely used for exploring chemoautotrophs in previous studies<sup>13,26,27</sup>. Notably, the key

genes, such as *accA* and *hcd*, involved in 3-HP/4-HB pathway (the other of the two chemoautotrophic C fixation pathways under aerobic conditions), were also examined, but neither of them was detected despite applying various primers and changing PCR conditions (Supplementary Fig. 5 and Supplementary Table 2). The *cbbl* gene was amplified using the primer pair (*cbbl*-F: 5'-ACCAYCAAGCCSAAGCTSGG-3'; *cbbl*-R: 5'-GCCTTCSAGCTTGCCSACCRC-3') and the thermal-cycling conditions reported by ref. 50. The primer set (*cbbm*-F: 5'-GGCACCATCATCAAGCCCAAG-3'; *cbbm*-R: 5'-TCTTGCCGTAGCCCATGGTGC-3') was applied to amplify the *cbbm* gene. PCR amplification conditions were carried out following the methods of ref. 51. Each DNA template was conducted with triplicate measurements. Standard curves ( $R^2 > 0.99$ ) were made by using 10-fold serial dilutions of purified plasmids containing the target gene. The copy numbers of the gene per unit volume of the water sample were calculated based on the standard curves.

To analyze the community composition and structure of chemoautotrophs, we conducted amplicon sequencing with the aforementioned primers of the *cbbl* gene rather than the *cbbm* gene because the CBB cycle was mainly carried out by the former rather than the latter revealed by this study. A total of 39 DNA samples were used for amplicon sequencing, and one sample was not examined due to insufficient DNA content. PCR conditions were: initial denaturation at 95 °C for 3 min, 40 cycles of 95 °C for 30 s, 62 °C for 30 s and 72 °C for 30 s, followed by 72 °C for 10 min. The quantity and quality of PCR products were determined using a 1.2% agarose gel. The obtained PCR products were then purified with a DNA gel extraction kit (Axygen, Silicon Valley, USA) and quantified using an FTC-3000 real-time PCR system (Funglyn, Shanghai, China). The final PCR pools were generated by mixing equimolar ratios of amplicons from different samples, and then sequenced on the MiSeq platform (2 × 300 bp) (Illumina, San Diego, USA). Raw sequence data were trimmed using Trimmomatic<sup>52</sup> with a sliding window size of 4 and an average phred score in the window of 20. Then Trimmed reads were merged using FLASH<sup>53</sup> with a minimum 10-bp overlap and an allowable error ratio <0.2. Subsequently, clean reads were clustered into OTUs using UPARSE<sup>54</sup> with 97% nucleotide identity. Representative sequences of each OTU were taxonomically classified using the FunGene database<sup>55</sup> with a confidence threshold  $\geq 0.8$ . Finally, we aligned representative sequences by MUSCLE<sup>56</sup> and constructed maximum-likelihood trees using FastTree<sup>57</sup>.

**Metagenomic sequencing.** To further elucidate the chemoautotrophic C fixation pathways, we performed metagenomic sequencing for water samples of thermokarst lakes (same as the water samples used for <sup>14</sup>C-labeled incubation in the summer of 2024). Specifically, 5–8 L of water sample collected in the field was filtered using a 0.22 μm polyethersulfone membrane (Sterivex GP, Millipore, Boston, USA). Filters were immediately frozen in a liquid nitrogen container and then stored at –80 °C before extraction. DNA was extracted using the DNeasy PowerWater Kit, following the instructions. The resulting DNA was then purified using the DNeasy PowerClean CleanUp Kit. DNA quality and concentration were detected with a NanoDrop-2000 (Thermo Fisher Scientific, Waltham, USA). Then, the ALFA-SEQ DNA Library Prep Kit was used to construct the metagenome libraries, in accordance with the manufacturer's protocols. Finally, the Illumina Nova PE150 platform was used to sequence the libraries, and 150 bp paired-end reads were generated subsequently.

Metagenomic assembly and binning processes were performed following the methods of ref. 58. Specifically, adapters trimming and raw sequences quality control were processed by fastp v0.21.0<sup>59</sup>, and then we obtained 1.25 Tb clean reads, with a mean value of 34.8 Gb data per sample (ranging from 28.3 to 47.4 Gb). Afterwards, we used megahit v1.2.9<sup>60</sup> to assemble the clean reads into contigs with default k-mers. Protein-coding ORFs finding was performed by Prodigal

v2.6.3<sup>61</sup> using contigs with a length longer than 1000 bp. A non-redundant gene catalog was generated by clustering ORFs globally with a 95% nucleotide identity cutoff using CDHIT v4.8.181<sup>62</sup>. Gene abundance was estimated with Salmon v1.5.1<sup>63</sup> and normalized as transcripts per million (TPM), according to gene length and sequencing depth. Finally, eggNOG-mapper v2.1.3 with the DIAMOND mode was used to annotate the gene functional profiles using the eggNOG 5.0 database<sup>64</sup>.

The assembled contigs underwent binning analysis using metaWRAP v1.3.2<sup>65</sup> to construct MAGs, leveraging the built-in MaxBin2, metaBAT2, and CONCOCT modules. Resulting MAGs were then refined using metaWRAP with the bin\_refinement module and deduplicated by dRep v3.4.3<sup>66</sup>. Meanwhile, the quality of MAGs was assessed with checkM v1.2.2<sup>67</sup>, and 973 good-quality MAGs ( $\geq 70\%$  completeness and  $\leq 10\%$  contamination) were kept for further analysis. We then used the coverM v0.7.0<sup>68</sup> to calculate the MAGs coverage. The taxonomic information was annotated by GTDB-Tk v2.1.1 using the GTDB r207 database<sup>69</sup>. The phylogenetic tree of MAGs was generated based on single-copy concatenated proteins using GTDB-Tk and visualized using iTOL<sup>70</sup>. METABOLIC V4.0 was used to annotate the proteins by integrating multiple databases, including KEGG, TIGRFam, Pfam, custom hidden Markov model profiles, dbCAN2, and MEROPS databases<sup>71</sup>. Metabolic pathways (or a module here) with completeness exceeding 75% were assigned as present within MAGs during functional annotation using METABOLIC V4.0<sup>71</sup>.

#### <sup>15</sup>N-isotope incubation for examining nitrification rates

To verify whether the dark C fixation was fueled by nitrification, we measured the rates of nitrification in the whole water column of thermokarst lakes collected in the summer of 2024 via a <sup>15</sup>N-isotope incubation experiment. Specifically, we used the modified method reported by ref. 33 to determine the nitrification rates of water samples (same as the samples used for <sup>14</sup>C-labeled incubation in 2024). First, 150 mL of filtered (0.7 μm) water samples was added to a 500 mL HDPE bottle. Then sterile filtered solution of <sup>15</sup>N-labeled ammonium sulfate [(<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 98.5% <sup>15</sup>N] was put into each bottle at a final concentration of 50 μM. Second, all the bottles labeled with (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> were incubated at 20 °C under dark and aerobic conditions for 5 days. During this incubation period, 20 mL of water sample was taken from each bottle after 12, 24, 48, 72, 96, and 120 h, which was then re-filtered with 0.2 μm filters and divided into two subsamples. One subsample (10 mL) was used for the measurements of NH<sub>4</sub><sup>+</sup>-N, NO<sub>2</sub><sup>-</sup>-N, and NO<sub>3</sub><sup>-</sup>-N based on a flow injection analyzer (Autoanalyzer 3 SEAL, Bran and Luebbe, Norderstedt, Germany) and Griess colorimetry. The other part (10 mL) was applied for determining the production of <sup>15</sup>NO<sub>3</sub><sup>-</sup>-N + <sup>15</sup>NO<sub>2</sub><sup>-</sup>-N.

For examining the production of <sup>15</sup>NO<sub>3</sub><sup>-</sup>-N + <sup>15</sup>NO<sub>2</sub><sup>-</sup>-N, 5 mL of water sample was placed into 12-mL glass vials (Exetainer; Labco Ltd., Lampeter, Ceredigion, UK), followed by adding 0.5 g spongy cadmium<sup>72</sup>. All the glass vials were capped with Teflon-lined septa and then shaken for 12 h on a horizontal shaker with 120 cycles/min, which could convert NO<sub>3</sub><sup>-</sup>-N to NO<sub>2</sub><sup>-</sup>-N. Afterwards, these water samples were transferred to fresh 12-mL glass vials and flushed with helium for 10 min to remove any N<sub>2</sub> in the vials. Finally, the NO<sub>2</sub><sup>-</sup>-N in water samples was further reduced to N<sub>2</sub> by adding sulfamic acid<sup>33</sup>. The produced N<sub>2</sub> (<sup>15</sup>N<sup>15</sup>N and <sup>14</sup>N<sup>15</sup>N) was determined using a gas chromatography isotope ratio mass spectrometer (253 Plus, Thermo Fisher Scientific, Waltham, USA). Nitrification rates were calculated based on the slope of the linear regression of the produced <sup>15</sup>N with incubation time.

#### Physical and chemical analyses of water samples

To explore the effects of water physiochemical properties on the DOC dynamics in thermokarst lakes, we not only analyzed DOC content, but also examined other parameters such as nitrogen, phosphorus, and

dissolved metallic ions in the lake water sampled in 2024. Specifically, we determined the concentration of total dissolved C (TDC) and TDN of lake samples using a multi N/C 3100 analyzer (Analytick Jena AG, Jena, Germany). The dissolved inorganic C (DIC) content was calculated by subtracting DOC from TDC. TDP, potassium, and manganese ions were determined using a spectrophotometer (ICAP6300, Thermo Fisher Scientific, Waltham, USA). Both the total irons (added with hydroxylamine hydrochloride) and ferrous ions (without hydroxylamine hydrochloride) in lakes were measured based on the ferrozine method<sup>73</sup>. The content of ferric iron was obtained based on the difference between total iron and ferrous ions. The concentration of reduced sulfur in water samples was determined by the methylene blue colorimetric method and iodometric titration method<sup>74</sup>.

To characterize the DOC fractions, we measured the optical properties (including ultraviolet/visible absorbance and fluorescence spectroscopy) of DOC in thermokarst lakes. Specifically, the absorbance of water samples ranging from 200 to 600 nm was examined with an ultraviolet-visible spectrophotometer (Lambda35, Perkin Elmer Inc., Waltham, USA). We then analyzed optical indices such as the specific UV absorbance at 254 nm ( $S_{254}$ ) and the absorption spectral slope at 275–295 nm ( $S_{275-295}$ )<sup>75</sup>.  $S_{254}$  was determined by dividing the absorption at 254 nm by the DOC content, which was positively associated with DOM aromaticity<sup>76</sup>.  $S_{275-295}$  was determined using Eq. (4) reported by the ref. 77:

$$\alpha_{\lambda} = \alpha_{\lambda_{ref}} e^{-S(\lambda - \lambda_{ref})} \quad (4)$$

where  $\alpha$  ( $m^{-1}$ ) is the Napierian absorption coefficient at wavelength  $\lambda$  (nm), and  $\lambda_{ref}$  is the reference wavelength (nm).  $S_{275-295}$  is negatively related to the molecular weight of DOC<sup>77</sup>.

Three-dimensional fluorescence excitation-emission matrices (EEMs) were examined using a fluorometer (F-4500, Hitachi Ltd., Tokyo, Japan). Before fluorescence scanning, water samples should be diluted to an absorption of <0.3 at 254 nm when necessary to diminish inner filter effects<sup>78</sup>. The EEMs were then measured across an excitation range from 250 to 500 nm (5 nm increments), and an emission range from 250 to 600 nm (2 nm increments) at room temperature<sup>10</sup>. The obtained EEMs were blank subtracted, inter-filter corrected, Raman and Rayleigh scattering eliminated, and Raman normalized using MATLAB version R2018a (MathWorks, MA, USA)<sup>79</sup>. Two parameters, including the biological index (BIX) and humification index (HIX), were determined using the corrected EEM data. The BIX could be used to reflect the proportion of freshly synthesized or autochthonous origin DOC, which is obtained as the ratio of the fluorescence emission intensity of 380 to 430 nm at the excitation wavelength of 310 nm<sup>80</sup>. The HIX is generally applied to represent the proportion of complex molecules and the degradation degree of DOC<sup>78</sup>. This parameter is calculated by dividing the intensity of fluorescence emission at 435–480 nm by the sum of the intensities of fluorescence emission (including 300–345 nm and 435–480 nm) at an excitation wavelength of 254 nm<sup>78</sup>.

### Statistical analyses

All data were checked and log-transformed if necessary to satisfy the normality of variance. We then analyzed the data with the following three steps. First, one-way ANOVAs coupled with the least significant difference (LSD) test were performed to examine the differences between the dark C fixation rates in thermokarst lakes experiencing DOC loss and accumulation during the incubation. We also applied the same statistical method to examine the differences in abundances of functional genes among the six chemoautotrophic C fixation pathways, such as CBB, WL, rTCA, 3-HP, 3-HP/4-HB, and DC/4-HB, and compare the differences between the concentrations of reduced inorganic compounds (inorganic nitrogen, ferrous ion, and reduced sulfur) for chemoautotrophs in these lakes. A paired *t*-test was used to

assess the differences between the *cbbL* and *cbbM* gene abundances in thermokarst lakes.

Second, regression analyses were conducted to examine the relationships of dark C fixation rates with  $\Delta$ DOC (% increase) and  $\Delta$ DOC (% CO<sub>2</sub> uptake) in thermokarst lakes, and the relationships between dark C fixation rates in water and sediments. We further performed the above analyses to explore the associations of increased proportion of DOC and dark C fixation in thermokarst lakes during the incubation with the concentrations of NH<sub>4</sub><sup>+</sup>-N and NO<sub>2</sub><sup>-</sup>-N, gene abundances of nitrification, and microbial nitrification rates. Regression analyses were also carried out to explore the linkages of  $\Delta$ DOC (% increase),  $\Delta$ DOC (% CO<sub>2</sub> uptake) and C fixation rates with climate (MAT and MAP), environments (Fe<sup>2+</sup>, Fe<sup>3+</sup>, manganese ion, reduced sulfur, potassium, dissolved oxygen, pH, salinity, conductivity, total dissolved solids, water temperature, and oxidation-reduction potentiality), substrates ( $S_{254}$ ,  $S_{275-295}$ , BIX, and HIX), nutrients (NH<sub>4</sub><sup>+</sup>-N, NO<sub>2</sub><sup>-</sup>-N, NO<sub>3</sub><sup>-</sup>-N, and TDP), chemoautotrophic taxa encoding *cbbL* gene and functional genes involved in chemoautotrophic C fixation pathways.

Third, based on the SEM, we identified the dominant factors shaping the C fixation rates,  $\Delta$ DOC (% increase), and  $\Delta$ DOC (% CO<sub>2</sub> uptake) across thermokarst lakes. Before SEM analysis, we established a priori model for microbial C fixation according to our empirical knowledge and significant correlations among these examined factors (Supplementary Fig. 9). In this model, microbial C fixation (including  $\Delta$ DOC (% increase),  $\Delta$ DOC (% CO<sub>2</sub> uptake) and chemoautotrophic C fixation) was assumed to be directly affected by the functional genes in CBB cycle, nutrients and substrates, and indirectly influenced by climate and nutrients. We conducted a principal component analysis (PCA) for each group of variables (such as climate, nutrients, and substrates) before performing the SEM analysis to reduce collinearity. The first component (PC1), explaining 74.5–87.0% of the total variance for the three groups of parameters (Supplementary Table 4), was introduced into the model as a new variable for the SEM analysis. We then parameterized the base model based on our measurements. The fitness of the final model was assayed using a Chi-squared test, the root mean-squared error, and Akaike's information criterion. All of these statistical analyses were performed using R 4.2.2<sup>81</sup> at a significance level of  $P < 0.05$ , except that the SEM was conducted with AMOS 21.0 (Amos Development Corporation, Chicago, USA).

### Reporting summary

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

### Data availability

All data supporting the findings in this study are available in the figshare database (<https://doi.org/10.6084/m9.figshare.30081886.v3>)<sup>82</sup> and Supplementary Information. The sequence data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under accession number PRJNA1363675. Source data are provided with this paper.

### References

- Abbott, B. W., Jones, J. B., Godsey, S. E., Larouche, J. R. & Bowden, W. B. Patterns and persistence of hydrologic carbon and nutrient export from collapsing upland permafrost. *Biogeosciences* **12**, 3725–3740 (2015).
- Olefeldt, D. et al. Circumpolar distribution and carbon storage of thermokarst landscapes. *Nat. Commun.* **7**, 13043 (2016).
- Wang, X. et al. Contrasting characteristics, changes, and linkages of permafrost between the Arctic and the Third Pole. *Earth Sci. Rev.* **230**, 104042 (2022).
- Turetsky, M. R. et al. Permafrost collapse is accelerating carbon release. *Nature* **569**, 32–34 (2019).

5. Walter Anthony, K. et al. 21st-century modeled permafrost carbon emissions accelerated by abrupt thaw beneath lakes. *Nat. Commun.* **9**, 3262 (2018).
6. Wik, M., Varner, R. K., Anthony, K. W., Macintyre, S. & Bastviken, D. Climate-sensitive northern lakes and ponds are critical components of methane release. *Nat. Geosci.* **9**, 99–105 (2016).
7. Yang, G. et al. Characteristics of methane emissions from alpine thermokarst lakes on the Tibetan Plateau. *Nat. Commun.* **14**, 3121 (2023).
8. Payandi-Rolland, D. et al. Dissolved organic matter biodegradation along a hydrological continuum in permafrost peatlands. *Sci. Total Environ.* **749**, 141463 (2020).
9. Schuur, E. A. G. & Mack, M. C. Ecological response to permafrost thaw and consequences for local and global ecosystem services. *Annu. Rev. Ecol. Evol. Syst.* **49**, 279–301 (2018).
10. Li, Z. et al. Accelerated organic matter decomposition in thermokarst lakes upon carbon and phosphorus inputs. *Glob. Change Biol.* **29**, 6367–6382 (2023).
11. Vonk, J. E. et al. Biodegradability of dissolved organic carbon in permafrost soils and aquatic systems: a meta-analysis. *Biogeosciences* **12**, 6915–6930 (2015).
12. Abbott, B. W., Larouche, J. R., Jones, J. B., Bowden, W. B. & Balsler, A. W. Elevated dissolved organic carbon biodegradability from thawing and collapsing permafrost. *J. Geophys. Res. Biogeosci.* **119**, 2049–2063 (2014).
13. Alfreider, A. et al. CO<sub>2</sub> assimilation strategies in stratified lakes: diversity and distribution patterns of chemolithoautotrophs. *Environ. Microbiol.* **19**, 2754–2768 (2017).
14. Cory, R. M., Crump, B. C., Dobkowski, J. A. & Kling, G. W. Surface exposure to sunlight stimulates CO<sub>2</sub> release from permafrost soil carbon in the Arctic. *Proc. Natl. Acad. Sci. USA* **110**, 3429–3434 (2013).
15. Cory, R. M., Ward, C. P., Crump, B. C. & Kling, G. W. Sunlight controls water column processing of carbon in arctic fresh waters. *Science* **345**, 925–928 (2014).
16. Hu, J. et al. Photo-produced aromatic compounds stimulate microbial degradation of dissolved organic carbon in thermokarst lakes. *Nat. Commun.* **14**, 3681 (2023).
17. Nalven, S. G. et al. Experimental metatranscriptomics reveals the costs and benefits of dissolved organic matter photo-alteration for freshwater microbes. *Environ. Microbiol.* **22**, 3505–3521 (2020).
18. Raven, J. A. Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquat. Microb. Ecol.* **56**, 177–192 (2009).
19. Wei, Z. et al. Sentinel-based inventory of thermokarst lakes and ponds across permafrost landscapes on the Qinghai-Tibet Plateau. *Earth Space Sci.* **8**, e2021EA001950 (2021).
20. Gao, T. et al. Accelerating permafrost collapse on the eastern Tibetan Plateau. *Environ. Res. Lett.* **16**, 054023 (2021).
21. Vick-Majors, T. J. & Priscu, J. C. Inorganic carbon fixation in ice-covered lakes of the McMurdo Dry Valleys. *Antarctic Sci.* **31**, 123–132 (2019).
22. Baltar, F., Aristegui, J., Gasol, J. M., Sintes, E. & Herndl, G. J. Evidence of prokaryotic metabolism on suspended particulate organic matter in the dark waters of the subtropical North Atlantic. *Limnol. Oceanogr.* **54**, 182–193 (2009).
23. Dahlback, A., Gelsor, N., Stamnes, J. J. & Gjessing, Y. UV measurements in the 3000–5000 m altitude region in Tibet. *J. Geophys. Res.* **112**, D09308 (2007).
24. Wickland, K. P. et al. Biodegradability of dissolved organic carbon in the Yukon River and its tributaries: seasonality and importance of inorganic nitrogen. *Glob. Biogeochem. Cycles* **26**, GBOE03 (2012).
25. Berg, C., Listmann, L., Vandieken, V., Vogts, A. & Jurgens, K. Chemolithoautotrophic growth of ammonia-oxidizing Thaumarchaeota enriched from a pelagic redox gradient in the Baltic Sea. *Front. Microbiol.* **5**, 786 (2015).
26. Berg, I. A. et al. Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* **8**, 447–460 (2010).
27. Yakimov, M. et al. Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary production in Tyrrhenian deep waters (Central Mediterranean Sea). *ISME J.* **5**, 945–961 (2011).
28. Berg, I. A., Kockelkorn, D., Bucket, W. & Fuchs, G. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* **318**, 1782–1786 (2007).
29. Yakimov, M. M., Cono, V. L. & Denaro, R. A first insight into the occurrence and expression of functional *amoA* and *accA* genes of autotrophic and ammonia-oxidizing bathypelagic Crenarchaeota of Tyrrhenian Sea. *Deep Sea Res. II Top. Stud. Oceanogr.* **56**, 748–754 (2009).
30. Offre, P., Nicol, G. W. & Prosser, J. I. Community profiling and quantification of putative autotrophic thaumarchaeal communities in environmental samples. *Environ. Microbiol. Rep.* **3**, 245–253 (2011).
31. Badger, M. R. & Bek, E. J. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO<sub>2</sub> acquisition by the CBB cycle. *J. Exp. Bot.* **59**, 1525–1541 (2008).
32. Urakawa, H. et al. Ammonia availability shapes the seasonal distribution and activity of archaeal and bacterial ammonia oxidizers in the Puget Sound Estuary. *Limnol. Oceanogr.* **59**, 1321–1335 (2015).
33. Overholt, W. A. et al. Carbon fixation rates in groundwater similar to those in oligotrophic marine systems. *Nat. Geosci.* **15**, 561–567 (2022).
34. Bochet, O. et al. Iron-oxidizer hotspots formed by intermittent oxic-anoxic fluid mixing in fractured rocks. *Nat. Geosci.* **13**, 149–155 (2020).
35. Mardanov, A. V., Beletsky, A. V., Kadnikov, V. T., Slobodkin, A. I. & Ravin, N. V. Genome Analysis of thermosulfurimonas dismutans, the first thermophilic sulfur-disproportionating bacterium of the phylum thermodesulfobacteria. *Front. Microbiol.* **7**, 950 (2016).
36. Wasmund, K., Mumann, M. & Loy, A. The life sulfuric: microbial ecology of sulfur cycling in marine sediments. *Environ. Microbiol. Rep.* **9**, 323–344 (2017).
37. Ding, J. et al. The permafrost carbon inventory on the Tibetan Plateau: a new evaluation using deep sediment cores. *Glob. Change Biol.* **22**, 2688–2701 (2016).
38. Wang, T. et al. Permafrost thawing puts the frozen carbon at risk over the Tibetan Plateau. *Sci. Adv.* **6**, eaaz3513 (2020).
39. Wang, D. et al. A 1 km resolution soil organic carbon dataset for frozen ground in the Third Pole. *Earth Sys. Sci. Data* **13**, 3453–3465 (2021).
40. Chen, L. et al. Permafrost carbon cycle and its dynamics on the Tibetan Plateau. *Sci. China Life Sci.* **67**, 1833–1848 (2024).
41. Editorial Committee for Vegetation Map of China. *Vegetation Atlas of China* (Science Press, 2001).
42. Zhang, J., Wang, J., Chen, W., Li, B. & Zhao, K. *Vegetation of Xizang (Tibet)* (Science Press, 1988).
43. Zhao, L. & Sheng, Y. *Permafrost and its changes on the Qinghai-Tibetan Plateau* (Science Press, 2019).
44. Bai, Y. et al. Heating up the roof of the world: tracing the impacts of in-situ warming on carbon cycle in alpine grasslands on the Tibetan Plateau. *Nat. Sci. Rev.* **12**, nwae371 (2025).
45. Fu, Z. et al. Non-temperature environmental drivers modulate warming-induced 21st-century permafrost degradation on the Tibetan Plateau. *Nat. Commun.* **16**, 7556 (2025).
46. Mu, M. et al. Thermokarst lake changes along the Qinghai-Tibet Highway during 1991–2020. *Geomorphology* **441**, 108895 (2023).
47. Kang, L. et al. Patterns and drivers of prokaryotic communities in thermokarst lake water across Northern Hemisphere. *Glob. Ecol. Biogeogr.* **32**, 2244–2256 (2023).

48. Santoro, A. L., Bastviken, D., Tranvik, L. & Enrich-Prast, A. Simultaneous measurements of dark carbon fixation and bacterial production in lake sediment. *Limnol. Oceanogr. Methods* **11**, 298–303 (2013).
49. Dykma, S., Bischof, K., Fuchs, B. M., Hoffmann, K. & Mumann, M. Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *ISME J.* **10**, 1939–1953 (2016).
50. Tolli, J. & King, G. M. Diversity and structure of bacterial chemolithotrophic communities in pine forest and agroecosystem soils. *Appl. Environ. Microbiol.* **71**, 8411–8418 (2005).
51. Alfreider, A., Vogt, C., Hoffmann, D. & Babel, W. Diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from groundwater and aquifer microorganisms. *Microb. Ecol.* **45**, 317–328 (2003).
52. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
53. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
54. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
55. Fish, J. et al. FunGene: the functional gene pipeline and repository. *Front. Microbiol.* **4**, 1–14 (2013).
56. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
57. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
58. Kang, L. et al. Metagenomic insights into microbial community structure and metabolism in alpine permafrost on the Tibetan Plateau. *Nat. Commun.* **15**, 5920 (2024).
59. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
60. Dinghua, L., Chi-Man, L., Ruibang, L., Kunihiko, S. & Tak-Wah, L. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
61. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **11**, 119 (2010).
62. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
63. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
64. Cantalapiedra, C. P., Ana, H.-P., Ivica, L., Peer, B. & Jaime, H. C. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
65. Uritskiy, G. V., Jocelyne, D. R. & James, T. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).
66. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
67. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
68. Aroney, S. T. N. et al. CoverM: read alignment statistics for metagenomics. *Bioinformatics* **41**, p1 (2025).
69. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2019).
70. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, 256–259 (2019).
71. Zhou, Z. et al. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome* **10**, 33 (2022).
72. McIlvin, M. R. & Altabet, M. A. Chemical conversion of nitrate and nitrite to nitrous oxide for nitrogen and oxygen isotopic analysis in freshwater and seawater. *Anal. Chem.* **77**, 5589–5595 (2005).
73. Anastácio, A. S., Harris, B., Yoo, H. I., Fabris, J. D. & Stucki, J. W. Limitations of the ferrozine method for quantitative assay of mineral systems for ferrous and total iron. *Geochim. Cosmochim. Acta* **72**, 5001–5008 (2008).
74. Zopfi, J., Ferdelman, T. G. & Fossing, H. Distribution and fate of sulfur intermediates-sulfite, tetrathionate, thiosulfate, and elemental sulfur-in marine sediments. *Geol. Soc. Am.* **379**, 97–116 (2004).
75. Drake, T. W., Wickland, K. P., Spencer, R. G., Mcknight, D. M. & Striegl, R. G. Ancient low-molecular-weight organic acids in permafrost fuel rapid carbon dioxide production upon thaw. *Proc. Natl. Acad. Sci. USA* **112**, 13946–13951 (2015).
76. Weishaar, J. L. et al. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* **37**, 4702–4708 (2003).
77. Helms, J. R. et al. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* **55**, 955–969 (2008).
78. Ohno, T. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environ. Sci. Technol.* **36**, 742–746 (2002).
79. Murphy, K. R., Stedmon, C. A., Graeber, D. & Bro, R. Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal. Methods* **5**, 6557–6566 (2013).
80. Fellman, J. B., Hood, E. & Spencer, R. G. M. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. *Limnol. Oceanogr.* **55**, 2452–2462 (2010).
81. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. (2022).
82. Liu, F. et al. Chemoautotrophic carbon fixation in thermokarst lakes on the Tibetan Plateau. Figshare <https://doi.org/10.6084/m9.figshare.30081886.v3> (2025).
83. Zou, D., Zhao, L., Sheng, Y., Chen, J. & Cheng, G. A new map of permafrost distribution on the Tibetan Plateau. *Cryosphere* **11**, 2527–2542 (2017).

## Acknowledgements

The authors would like to thank the Isotope Laboratory Platform from the Institute of Biophysics, Chinese Academy of Sciences, and the Analytical and Testing Center of Beijing Normal University for supporting  $\text{NaH}^{14}\text{CO}_3$  assimilation experiment. The authors would also be grateful to Drs. Hongjie Zhang (from Chinese Academy of Sciences), Guohua Jiang and Jianhua Huang (both are from Beijing Normal University) for their technical assistance with the isotopic analyses. This work was supported by the National Natural Science Foundation of China (32588202, 32425004, and 32571870), the National Key Research and Development Program of China (2022YFF0801903), the New Cornerstone Science Foundation through the XPLORER PRIZE, and the Fundamental Research Funds of the Chinese Academy of Forestry (CAFYBB2023QA004 and CAFYBB2024ZA034).

## Author contributions

Y.Y. conceived the idea. Y.Y., F.L., and L.K. designed the study. L.K., Z.L., F.L., W.Z., W.X., and X.L. conducted the field water and sediment sampling. F.L., L.K., Z.L., and W.X. performed the experiments. F.L. and L.K. analyzed the data. F.L., Y.Y., and L.K. wrote the manuscript with input from Z.L., J.P., B.W.A., L.C., S.Q., D.Z., and Y.P.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67478-x>.

**Correspondence** and requests for materials should be addressed to Yuanhe Yang.

**Peer review information** *Nature Communications* thanks Karl Romanowicz and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025