Single-cell trait diversity explains niche and fitness differences in aquatic microbial communities

Simone Fontana[a,b,c,#], Désirée A. Schmitz[d,e,f], Michael Daniels[d,e], Francesco Danza[g], Thomas Röösli[d,e], Hannah Bruderer[d,e], Jean-Claude Walser[h], Julien Dekaezemacker[i], Stephane Escrig[j], Anders Meibom[j,k], Francesco Pomati[a], Frank Schreiber[d,e,l,#]

[a]Eawag, Swiss Federal Institute for Aquatic Science and Technology, Dept. of Aquatic Ecology, Dübendorf, Switzerland.

[b]University of Freiburg, Nature Conservation and Landscape Ecology, Freiburg, Germany.

[c]Kanton St. Gallen, Amt für Natur, Jagd und Fischerei, Abteilung Natur und Landschaft, St. Gallen, Switzerland.

[d]ETH Zurich, Dept. of Environmental Systems Science, Institute for Biogeochemistry and Pollutant Dynamics, Zurich, Switzerland.

[e]Eawag, Swiss Federal Institute for Aquatic Science and Technology, Dept. of Environmental Microbiology, Dübendorf, Switzerland.

[f]University of Zurich, Department of Quantitative Biomedicine, Zurich, Switzerland.

[g]University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Department of Environment, Constructions and Design, Institute of Microbiology, Via Flora Ruchat-Roncati 15, 6850 Mendrisio, Switzerland.

[h]ETH Zurich, Dept. of Environmental Systems Science, Genetic Diversity Centre (GDC), Zurich, Switzerland.

[i]Max Planck Institute for Marine Microbiology, Dept. of Biogeochemistry, Bremen, Germany.

[j]Laboratory for Biological Geochemistry, School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

[k]Center for Advanced Surface Analysis, Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland.

[l]Federal Institute for Materials Research and Testing (BAM), Department of Materials and Environment, Division of Biodeterioration and Reference Organisms, Berlin, Germany.

#Corresponding authors: Simone Fontana (simonefontana.j@gmail.com) and Frank Schreiber (frank.schreiber@bam.de)

Abstract

Two fundamental questions in ecology are how biodiversity is maintained and how it affects ecosystem functioning. Until now, it has been difficult to study the above mechanisms in natural microbial communities, yet they are important drivers of biogeochemical ecosystem functions. Here, we use a new approach to define and measure biodiversity in complex lake microbial communities (Lake Cadagno, Switzerland) based on the cell-to-cell variation in multiple functionally relevant phenotypic traits. We use stable isotope probing coupled to correlative imaging using confocal laser scanning microscopy (CLSM) and nanometer-scale secondary ion mass spectrometry (NanoSIMS) to obtain morphological (size), physiological (pigments) and metabolic (carbon and nitrogen isotope uptake and sulfur content) traits for a large number of individual cells along the environmental gradient found across lake depth. We show that cell-to-cell trait variation is significantly correlated with cell densities as a proxy for ecosystem functioning, whereas genetic diversity measured at the level of 16S and 18S is not. Our single-cell analysis provides evidence for a simultaneous increase in niche partitioning (measured as increased evenness in pigment composition) and decrease in fitness differences (measured as decreased variability in sulfur content) due to light limitation and competition for nutrients in deep layers of the lake. This leads to a negative relationship between niche and fitness differences. Our results suggest that niche and fitness differences in natural microbial communities can be understood at the level of single-cell traits, providing a mechanistic understanding of the relationship between microbial diversity and ecosystem functioning.

Introduction

Modern and past biodiversity levels on Earth have raised two central questions in ecological research: how did this diversity arise and how is it maintained in natural ecosystems? The principle of competitive exclusion[1-3] and the closely related concept of ecological niche[4-6] initially led to the recognition that species have different requirements and use distinct resources, thus explaining their coexistence. However, the number of different types of organisms – especially in microbial communities – greatly exceeds the number of distinct limiting resources (e.g., 'the paradox of the plankton'[7]), and other mechanisms must also play a role. These mechanisms have broadly been divided into two categories: increasing niche partitioning and decreasing fitness differences[8-10], which are not mutually exclusive. Indeed, both niche and fitness effects can act simultaneously[11,12] and be considered interdependent[13]. In addition, evolutionary (genetic) mechanisms also act as sources of diversity in microbial communities[14,15]. For example, rapid evolution has been shown to maintain taxonomic diversity in aquatic communities, even in the absence of niche differentiation[16,17].

In addition to the drivers of diversity, ecological research has also addressed its consequences, i.e., biodiversity and ecosystem functioning relationships (e.g.[18,19]). The mechanisms that generate diversity are arguably related to those determining ecosystem functioning[20-22]. As microbial communities are fundamental drivers of biogeochemical functions, such as the remineralization of polymers and the nitrogen cycle, it is essential to elucidate these mechanisms to better understand the contribution of microbial communities to ecosystem functioning.

Interestingly, both niche partitioning and fitness differences can be the outcome of variation in functional traits that, by definition, are features measurable at the level of individual organisms[23]. In other words, the outcome(s) of coexistence observed at the community level are in fact the result of interactions occurring among individuals[24]. However, studies of niche and fitness

differences at the level of the effective competitive units, i.e., individual organisms are still lacking. Addressing this gap is particularly challenging for microorganisms in the natural environment, as it is difficult to measure their relevant traits at the single-cell level. Most research to date in this field has used low specificity measures of trait diversity, typically including multiple traits that provide limited mechanistic understanding, such as broad classes of functionally unrelated genes from metagenomic datasets[25]. However, the advantage of this approach is that it often provides good predictive power of ecosystem functioning[24]. We capitalize on recent advances in the understanding of microbial function through single-cell trait-based approaches to investigate niche and fitness differences in microbial communities based on single-cell phenotypes.

Modern high-throughput single-cell technologies allow for a more mechanistic and hypothesis-driven investigation of microbial adaptive strategies based on individual-level traits[26]. The integration of such technologies has the potential to reveal niche and fitness differences at the level of the individual organisms and assess whether these traits drive stabilizing and equalizing mechanisms at the population and community level. First, the definition of microbial species (assuming they exist) is challenging and not trivial[27-29]. Second, the extreme phenotypic variability observed even within clonal microbial populations[30] or closely related sub-taxonomic groups[31] might jeopardize the promise of modern taxonomic approaches linked to rRNA gene sequencing to elucidate relationships between biodiversity and ecosystem function[32-34]. Taken together, a single-cell perspective will improve our understanding of how niche and fitness differences relate to ecosystem functioning in nature, with possible implications for the monitoring of biodiversity and related functions.

In this study, we aim to investigate niche and fitness differences within a natural microbial community and their relationship with ecosystem functioning based on the characteristics of individual microbial cells. In this context, it is crucial to identify which traits might cause niche partitioning, and which may be used to quantify fitness differences. We use correlative

104 fluorescence microscopy and stable isotope probing coupled to nanometer-scale secondary ion
105 mass spectrometry (NanoSIMS)[35] to determine photosynthetic pigment composition, size, and
106 metabolic activity for carbon, nitrogen and sulfur simultaneously in thousands of individual
107 microbial cells from different samples. We leverage the idea that environmental gradients explain
108 the diversification of ecological strategies or niches[36], as well as fitness differences between
109 phenotypes[37]. On the one hand, pigment-related traits (i.e., photosynthesis/physiology) enable
110 niche partitioning, as cells can decrease competition for light by targeting slightly different
111 wavelengths through fine-tuning of the relative expression of different pigments[38]. On the other
112 hand, traits defining the assimilation or cell content of elements involved in fundamental metabolic
113 pathways (C, N, and S) - which often show great intraspecific variation[39-41] - can be considered
114 as proxies for fitness, as they influence the maintenance, growth and reproduction of
115 organisms[42,43]. Here, we sampled single cells living in structured microbial communities at
116 different depths in the meromictic Lake Cadagno (Switzerland), and considered decreasing levels
117 of trait integration to predict ecosystem functioning with multidimensional trait diversity indices
118 and detect niche and fitness differences determined by the variation in physiological and
119 metabolic traits[24]. Accordingly, we formulate the following expectations:
120 **1)** Ecosystem functioning (i.e., cell density as a proxy for primary production) is better
121 explained by trait diversity metrics that integrate all available individual-based traits
122 (morphological, physiological, and metabolic) rather than by taxonomic diversity metrics based
123 on rRNA gene sequencing[33,44]. **2)** In deep layers of the lake, competition for light drives increasing
124 niche partitioning in pigment-related traits (i.e., increasing trait evenness[38]), while limitation for
125 essential nutrients mediates decreasing fitness differences in metabolic traits (i.e., trait
126 convergence[45]), as compared to the upper layers of the lake. **3)** These simultaneous shifts in trait
127 distribution lead to a negative relationship between single-cell niche and fitness differences across
128 water depth.

129 ## Results

130 ### Environmental variables structure Lake Cadagno into distinct layers

131 First, we set out to characterize the environmental variables that shape the microbial
132 communities in Lake Cadagno and that potentially drive trait variability, mediating niche and
133 fitness differences. The vertical structure of physical, chemical and biological parameters of the
134 meromictic Lake Cadagno during one day of our sampling campaign is shown in Fig. 1A (see
135 Extended Data Fig. 1 for the profiles of the other two days). Lake Cadagno is perfectly suited to
136 study niche and fitness differences in microbial communities because its steep gradients of light
137 intensity, oxygen and sulfide concentration (Fig. 1A) create ideal conditions for very distinct
138 microbial communities at different depths[46]. Differences in salt concentration stably stratify the
139 lake into three main layers. The upper layer, called mixolimnion, is dominated by oxygenic
140 photosynthetic organisms. Oxygen is consumed around 11 m water depth, below which light can
141 still penetrate. A chemocline then forms in this largely anoxic water, which is dominated by high
142 cell densities of anoxygenic phototrophic purple and green sulfur bacteria. As electron donors,
143 these bacteria use sulfide and other reduced sulfur compounds that diffuse upward from the lake's
144 bottom waters (Fig. 1A; see also[47]). Below the chemocline, in the bottom waters, light penetration
145 is no longer measurable, and sulfide is produced by sulfate-reducing bacteria (Fig. 1A).

146

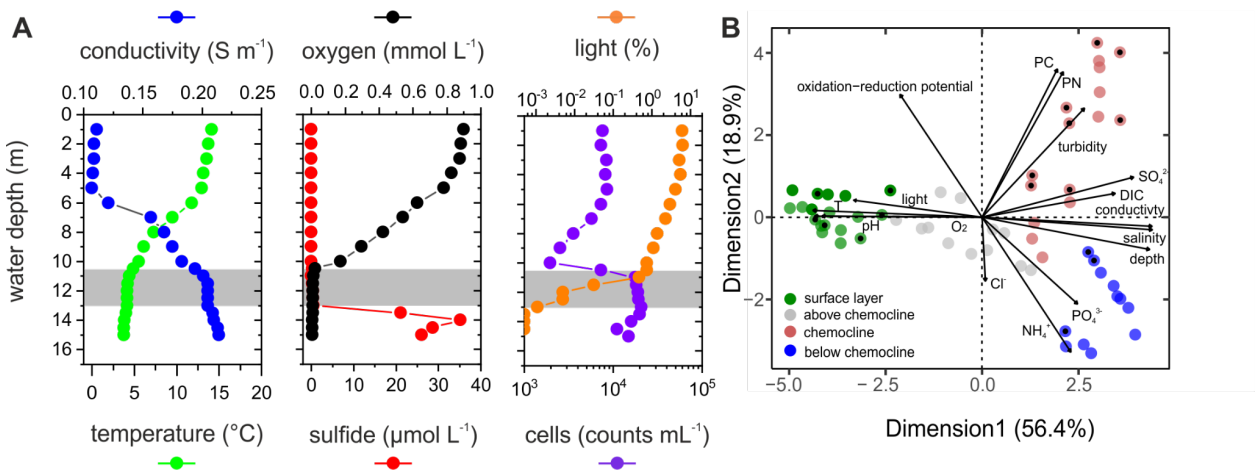

Fig. 1. Vertical gradients of physical, chemical and biological parameters in Lake Cadagno separate the ecosystem into three distinct layers. **(A)** Vertical profiles of conductivity (blue), temperature (green), oxygen (O_2, black), total sulfide (red), light (orange), and cell density (purple) measured on day C (19/08/2014) of the sampling campaign. Profiles from the other sampling days are shown in Extended Data Fig. 1. The gray area represents the chemocline, where sulfide and O_2 are depleted, the cell density increases, and light gets severely limited. Light values are presented as % photosynthetic active radiation of the light above the water. Light values shown as half circles below the chemocline were below the detection limit of the sensor. **(B)** Principal component analysis (PCA) of individual samples based on environmental factors (abbreviations: T=temperature; Cl^-=chloride; NH_4^+=ammonium; PO_4^{3-}=ortho-phosphate; DIC=dissolved inorganic carbon; SO_4^{2-}; sulfate; PN=particulate nitrogen; PC=particulate carbon). The PCA plot shows individual samples (circles) taken at different days distributed across environmental factors. Samples belonging to the same sampling layer (color) are grouped more closely in two-dimensional space. Samples for which individual-based trait diversity (*'Trait Onion Peeling'* TOP, *'Trait Even Distribution'* TED and *'Functional Dispersion'* FDis) were calculated for further analysis are marked by a black point. The surface layer is defined hereafter as 'upper layer' (as the gray dots are not considered for further analyses), and the layer below chemocline as 'lower layer'.

Our analysis of the lake stratification based on environmental factors (Fig. 1B) shows that Lake Cadagno is separated into distinct layers and that different environmental factors shape the structure and function of the microbial communities in each layer. Specifically, the upper layer is strongly separated from the chemocline and the lower layer by one axis of the principal component analysis (PCA; dimension 1), which explains 56.4% of the variance. The variables that most

168 strongly contribute to variation in this dimension are depth and light intensity. This indicates that
169 the microbial communities found in these samples are, as expected, stratified based on the
170 availability of light as a resource for photosynthesis. Samples in the chemocline are most strongly
171 separated from the lower layer by other environmental parameters (dimension 2 of the PCA,
172 explaining 18.9% of the variance): high turbidity, oxidation-reduction potential, particulate carbon,
173 and ammonium concentration. Samples in the lower layer correlate most strongly with the
174 concentration of ammonium and ortho-phosphate. These environmental factors thus separate
175 Lake Cadagno into different layers and provide the physio-chemical basis for the structuring of
176 the lake's microbial communities.

177

178 **Single-cell trait diversity explains ecosystem functioning**

179 Next, we tested whether ecosystem functioning can be better explained by integrative trait
180 diversity metrics (including physiological and morphological traits from CLSM and metabolic traits
181 from NanoSIMS; Fig. 2) than taxonomic diversity metrics based on rRNA gene sequencing. To
182 do that, we analyzed trait and taxonomic diversity data across the different depths of the lake and
183 time (days) with linear models to assess their explanatory power for ecosystem functioning,
184 represented here by the density of cells as a proxy for primary production. The analysis of our
185 large dataset across time and depths shows that trait diversity is a significant driver of ecosystem
186 functioning. Specifically, the relative regularity in distances between individual cells (TED, *'Trait*
187 *Even Distribution'*) and their spread around the average phenotype (FDis, *'Functional Dispersion'*)
188 in high-dimensional trait space was significantly related to ecosystem functioning, whereas the
189 total trait space covered (TOP, *'Trait Onion Peeling'*) did not. Thus, the two best models retained
190 for model averaging explained 56.8% (explanatory variables: TED, FDis) and 71.9% (explanatory
191 variables: TED, FDis, Day of sampling) of the variance in cell density relative to the maximum

192 achieved in a given layer (Extended Data Table S1). Model averaging showed a negative effect
193 of TED on cell density (standardized regression coefficient=-0.195; 95% CI=[-0.339; -0.050]),
194 whereas FDis had a positive effect of similar strength (standardized regression coefficient=0.153;
195 95% CI=[0.028; 0.279]) (Fig. 3). Contrary to this, taxonomic diversity metrics based on rRNA gene
196 sequencing did not show any significant correlation with cell density. Indeed, the best model
197 according to AICc (Akaike information criterion corrected for small sample sizes) was the null
198 model (including the intercept only). The only other model retained for model averaging also
199 included the Simpson index based on 18S sequencing and only explained 5.7% of the variance
200 in cell density relative to the maximum achieved in a given layer (Extended Data Table S1). As a
201 consequence, the effect size of the Simpson index based on 18S sequencing is not significantly
202 different from 0 (standardized regression coefficient=0.026; 95% CI=[-0.079; 0.130]) (Fig. 3).
203 Taken together, these data show that trait diversity explains ecosystem functioning in Lake
204 Cadagno while taxonomic diversity does not.

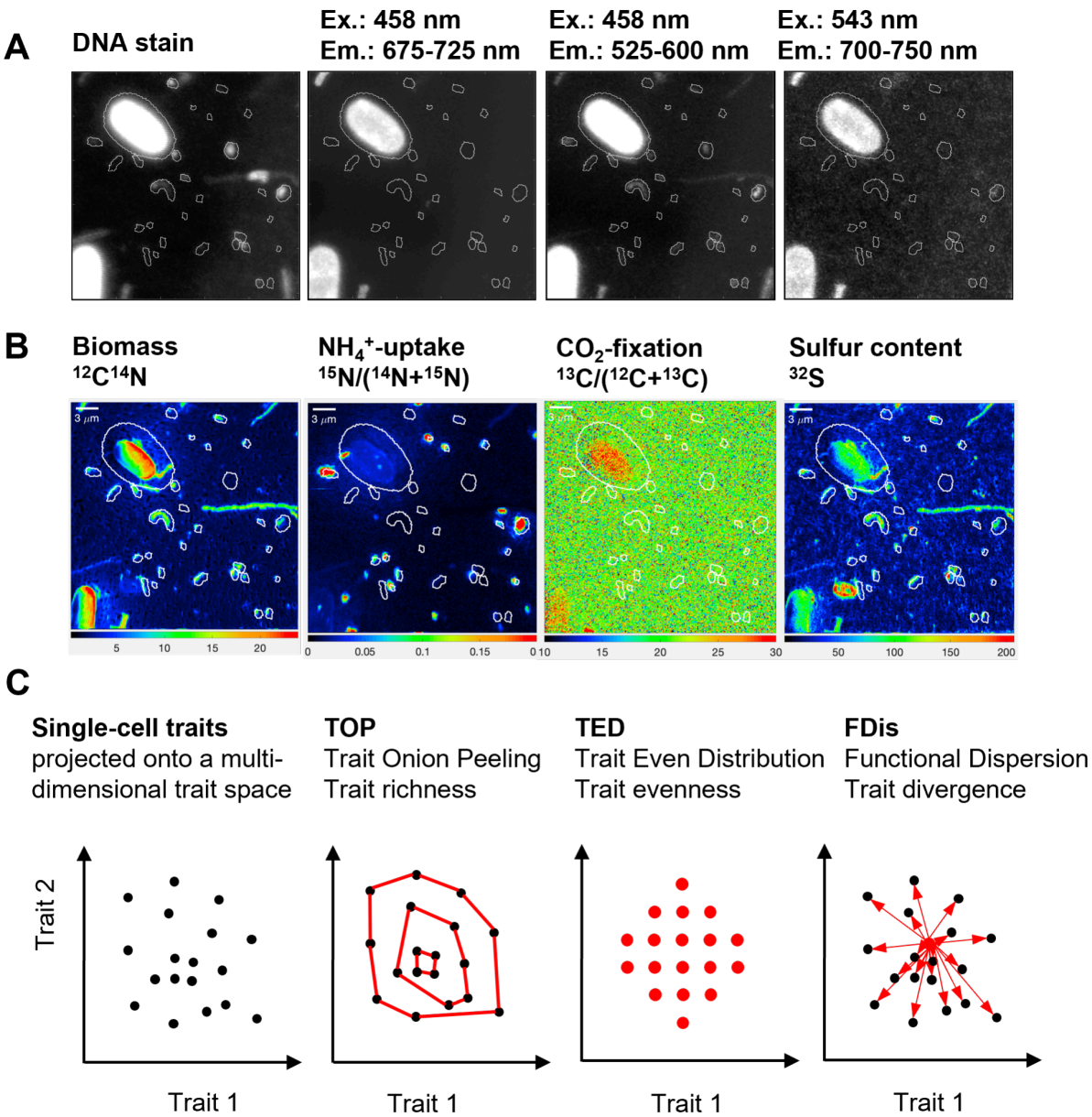

Fig. 2. Correlative microscopy to determine single-cell trait-based diversity indices. Panels **(A)** and **(B)** show example images for the same region of interest (ROI) measured with confocal laser scanning microscopy and NanoSIMS, respectively. The lines around cells show the cell segmentation that was based on DNA stain and the $^{12}C^{14}N$ NanoSIMS signal. Cells that are at the edge of each ROI were not included in the analysis. The excitation (Ex.) and emission (Em.) wavelengths are shown above each microscopy image. The intensity scale for isotope counts ($^{12}C^{14}N$ and ^{32}S) and isotope fractions ($^{15}N/[^{14}N+^{15}N]$ and $^{13}C/[^{12}C+^{13}C]$) are shown below each NanoSIMS image. Scale bars are only shown in the NanoSIMS images, but they are also valid for the microscopy images since they show the same ROI. **(C)** Scheme summarizing how the measured single-cell traits are used to calculate the three components of trait diversity: richness

(TOP index), evenness (TED index) and divergence (FDis index). Note that the scheme shows the indices on a 2-dimensional trait axis to better understand the principle. In this study, the trait axes form a multi-dimensional trait space with seven dimensions represented by fluorescence data (three channels) and cell size from CLSM, and NanoSIMS data (C and N assimilation, as well as S cell content).

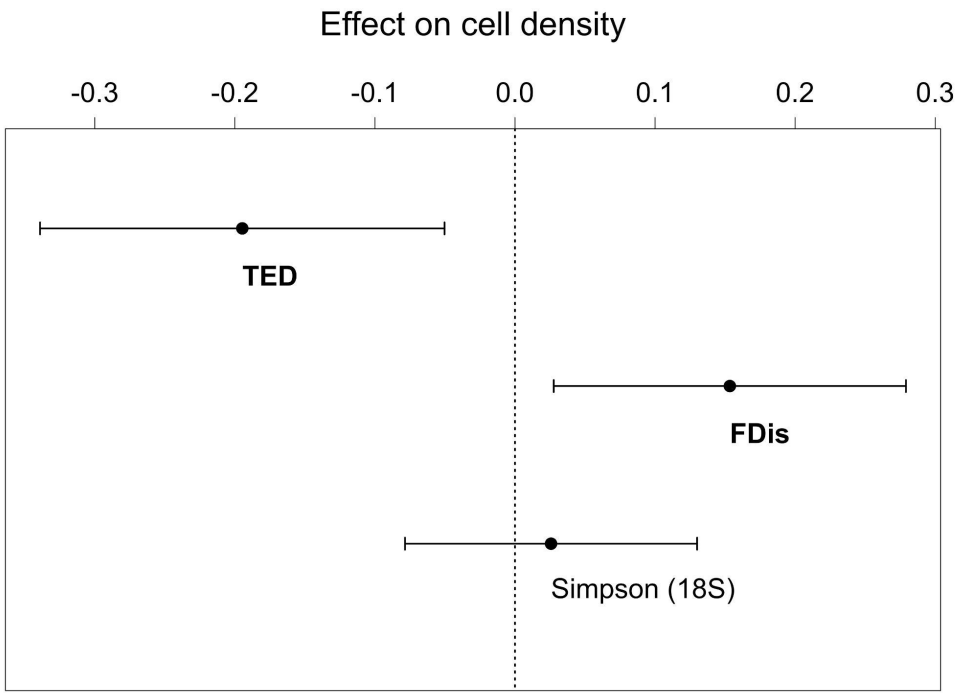

Fig. 3: Single-cell-based trait diversity explains ecosystem functioning in Lake Cadagno, whereas taxonomic diversity based on rRNA gene sequencing does not. Effect size of trait diversity metrics (TOP index of trait richness, TED index of trait evenness, and FDis index of trait divergence, calculated using three fluorescence channels, cell size from CLSM, as well as C assimilation, N assimilation, and S cell content) and genetic diversity metrics (Chao1 and Simpson index using 16S and 18S sequencing data) on cell density. Only metrics included in the list of models used for model averaging are shown here (predictors with a significant effect are in bold). Trait diversity (N=17) and genetic diversity (N=24) were modeled separately (see Methods). Values are standardized model-averaged regression coefficients shown as black circles with their 95% confidence intervals depicted as error bars (details about single models can be found in Extended Data Table S1). Cell density (quantified by scanning flow cytometry) is relative to the maximum achieved in a given layer.

231 **Single-cell traits drive niche and fitness differences in distinct lake layers**

232 We analyzed our extensive single-cell dataset to investigate whether the highly competitive
233 environment in deep layers of the lake drives increasing niche partitioning in pigment-related traits
234 and decreasing fitness differences in metabolic traits. The data showed that the evenness of
235 pigment-related traits was significantly higher in the chemocline (under low light intensity
236 combined with high cell density) as compared to the upper layer (under highest light availability)
237 (Fig. 4, Table 1A). The lower layer showed intermediate evenness of pigment-related traits, but
238 differences with the upper layer and chemocline were not significant (Table 1A). In contrast, the
239 evenness of metabolic traits did not show any significant difference between the three layers of
240 the lake (Fig. 4 and Table 1B). This suggests that the evenness in traits related to light acquisition
241 is positively related to increased niche partitioning under light limitation.

242 The analysis further revealed that metabolic traits exhibit a significantly lower multidimensional
243 metabolic trait divergence in the chemocline and lower layer than in the upper layer (Fig. 4 and
244 Table 1D). In contrast, no significant difference in the divergence of pigment-related traits was
245 observed between the three layers of the lake (Table 1C). This suggests that decreased
246 divergence in metabolic traits related to resource acquisition is linked to decreased fitness
247 differences. The latter occurred when competition for limiting resources (i.e., nutrients) increased.
248 Taken together, our data suggest that niche partitioning and fitness effects act simultaneously
249 between single-cell phenotypes in microbial communities along the depth gradient of Lake
250 Cadagno and are linked to different sets of traits.

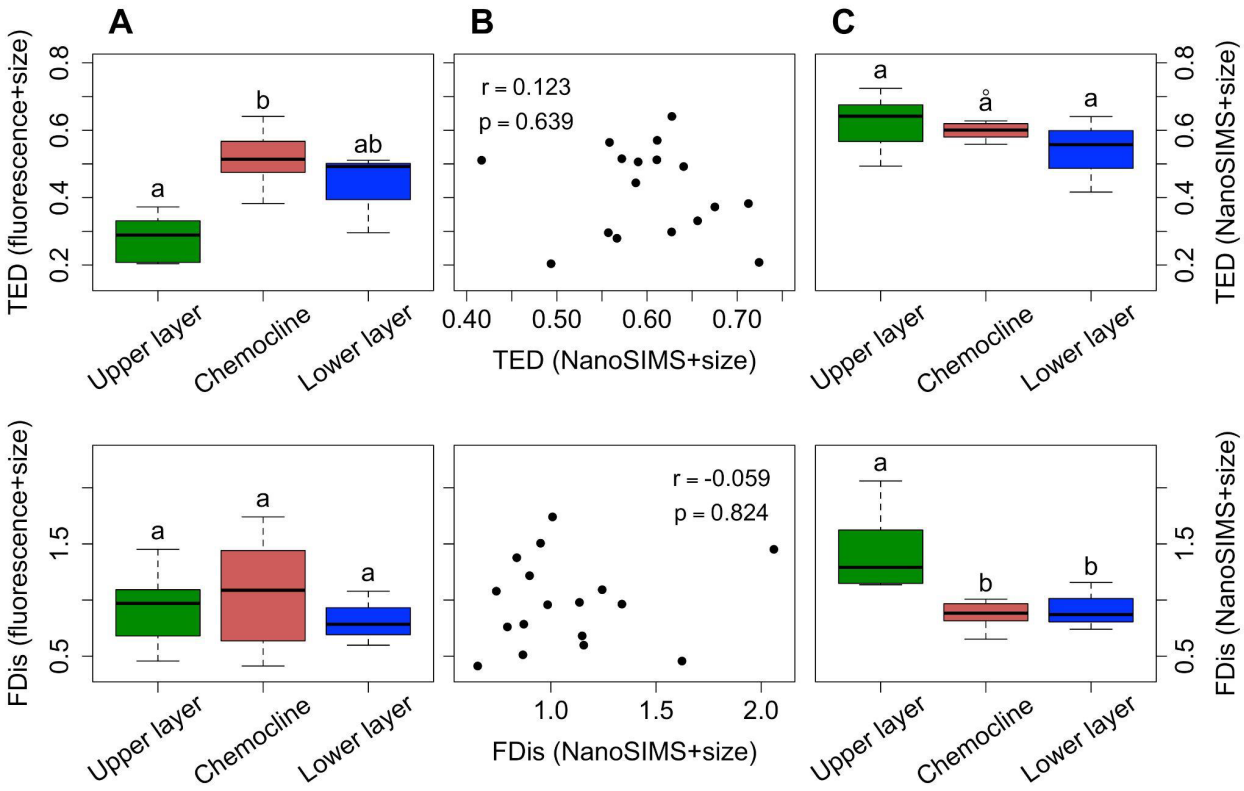

Fig. 4: Niche and fitness differences are linked to different sets of traits. One-way ANOVA analyses and post hoc Tukey tests (see Table 1; significance level at $p<0.05$, N=17) show that the evenness of fluorescence traits (TED) is higher in the chemocline than in the upper layer, whereas trait divergence (FDis) does not show any significant difference (column **A**). Conversely, the evenness of metabolic traits measured with NanoSIMS does not change significantly between layers, but trait divergence is highest in the upper layer (column **C**). Fluorescence and metabolic trait distributions show weak and non-significant Spearman correlations (column **B**; function *cor.test*, N=17), further supporting the idea that fluorescence and metabolic traits reflect distinct mechanisms. Each boxplot reports minimum and maximum (whiskers), first and third quartile (lower and upper side, respectively), as well as median (horizontal black line). We note that, for the calculation of TED and FDis, cell size (from CLSM) was included along with three fluorescence channels from CLSM on one side (see Fig. 2A), as well as C assimilation, N assimilation and S cell content on the other side (see Fig. 2B).

265 **Table 1:** The results of the one-way ANOVA analyses and post hoc Tukey tests for the data shown in Fig. 4. All p-
266 values are adjusted for multiple comparisons.
267

	A) TED (fluorescence+size)		B) TED (NanoSIMS+size)	
	difference	**p-value**	**difference**	**p-value**
chemocline *vs.* upper layer	0.235 [0.118; 0.351]	<0.001	-0.015 [-0.119; 0.089]	0.924
lower layer *vs.* upper layer	0.151 [-0.001; 0.303]	0.052	-0.086 [-0.222; 0.050]	0.258
lower layer *vs.* chemocline	-0.084 [-0.230; 0.062]	0.319	-0.071 [-0.201; 0.059]	0.356

	C) FDis (fluorescence+size)		D) FDis (NanoSIMS+size)	
	difference	**p-value**	**difference**	**p-value**
chemocline *vs.* upper layer	0.123 [-0.451; 0.697]	0.842	-0.551 [-0.894; -0.207]	0.002
lower layer *vs.* upper layer	-0.117 [-0.868; 0.635]	0.913	-0.502 [-0.952; -0.052]	0.028
lower layer *vs.* chemocline	-0.240 [-0.960; 0.480]	0.666	0.049 [-0.382; 0.479]	0.953

268
269

Niche and fitness differences are negatively related because of resource limitation

After we confirmed that individual-level niche and fitness differences are evident in our dataset, we analyzed how they are linked to each other across the main environmental gradient of the lake. To this end, we formulated the hypothesis that single-cell niche and fitness differences are negatively related across depth. More specifically, our analysis showed that the cell-to-cell coefficient of variation in fitness proxies was negatively related to the evenness of traits related to pigments. This is arguably not unexpected because, with increasing depth, both increased niche partitioning of the light spectrum and decreased fitness differences due to more similar metabolic traits should occur. The results of the linear model (estimate= -2.378, SE=0.752, t=-3.164, df=15, R^2=0.400; p=0.006) suggested that the individual-level niche partitioning of the light spectrum in deeper layers was strongly associated with more similar fitness as estimated by S content. Accordingly, cells in the upper layer were characterized by a high coefficient of variation in S content and uneven distribution of pigment-related traits (Fig. 5A, green dots), whereas cells in the chemocline and lower layer showed the expected combination of high evenness of pigment-related traits and low coefficient of variation in S content (Fig. 5A, red and blue dots). We also investigated the relationship between the evenness of pigment-related traits and the coefficient of variation in the single-cell assimilation rates of two other elements (C as CO_2 and N as NH_4^+; Fig. 5B and C, respectively). We did not find any significant relationship between trait evenness of pigments and C assimilation rate (estimate=0.221, SE=1.537, t=0.144, df=15, R^2=0.001; p=0.888), which suggests that CO_2 assimilation is not a good proxy for fitness in this system. In contrast, the coefficient of variation of NH_4^+ assimilation rates showed a significant and positive relationship with the evenness of pigment-related traits (estimate=1.332, SE=0.584, t=2.280, df=15, R^2=0.257; p=0.038). Cells in the upper layer (Fig. 5C, green dots) were generally more similar to each other in terms of NH_4^+ assimilation rates than cells in the chemocline and lower layer (Fig. 5C, red and blue dots).

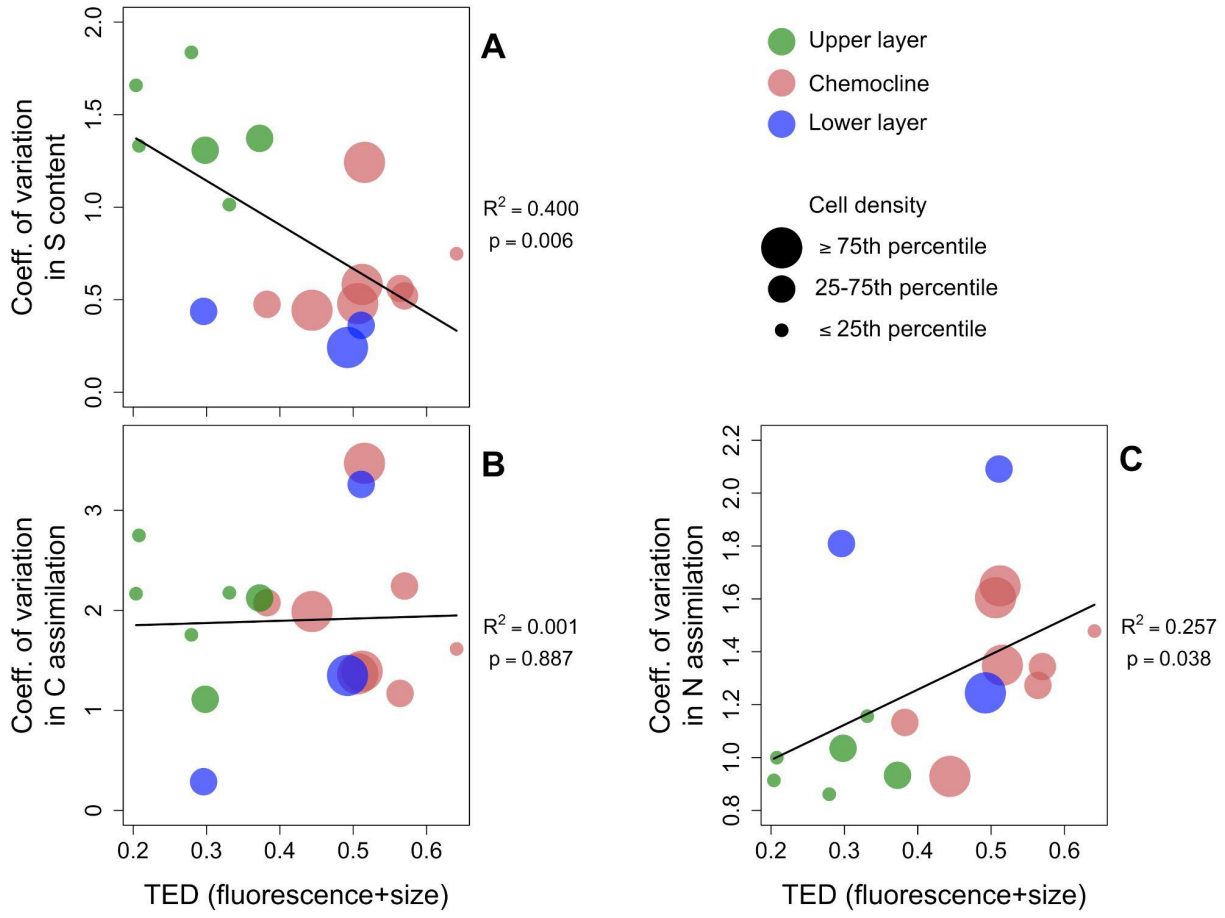

Fig. 5: Relationship between niche and fitness differences across different layers of Lake Cadagno. Niche partitioning of the light spectrum was quantified using the TED index of trait evenness calculated with fluorescence data (three channels) plus cell size from CLSM data (see Fig. 2A & C). Fitness differences were quantified from NanoSIMS data (see Fig. 2B) using the coefficient of variation in S content (panel **A**), C assimilation (panel **B**) and N assimilation rates (panel **C**) among individual cells, whereby fitness differences decrease with decreasing coefficient of variation. Each circle represents a sample taken at a specific water depth and day of sampling. The color of the circles represents the association of the sample to a specific lake layer as defined in Fig. 1B. The size of the circle represents the association of each sample to a category of the magnitude of relative cell density in this layer as a proxy for ecosystem functioning. Black regression lines, R^2, and p-values were obtained by fitting simple linear regressions (N=17).

308 ## Discussion

309 In this study, trait diversity metrics integrating individual-level morphological, physiological and
310 metabolic traits correlate significantly with ecosystem functioning, measured as cell densities, in
311 contrast to taxonomic diversity metrics based on rRNA gene sequencing. This is consistent with
312 previous findings: for example, multidimensional indices of functional diversity based on
313 individual-based traits have been shown to outperform taxonomic diversity metrics in explaining
314 phytoplankton biomass in natural communities[34]. Furthermore, the appropriate definition of
315 microbial strategies (trait combinations) can help explain the contribution of microbial processes
316 to carbon fluxes in soil ecosystems[48]. The lack of correlation between ecosystem functioning and
317 taxonomic diversity, however, contradicts some previous findings[21]. One potential explanation for
318 this is the importance of resource limitation in Lake Cadagno as compared to previously
319 investigated systems. Previous studies have shown a positive relationship between resource
320 limitation and intraspecific phenotypic heterogeneity in microbial populations[38,41,49]. Phenotypic
321 heterogeneity can result from physiological acclimation and stochastic gene expression, and
322 these physiological responses are not captured by taxonomic analysis. It is important to note that,
323 in this case, physiology might have a key role in predicting ecosystem functioning, more than
324 community composition or richness from taxonomic analysis. An alternative explanation could be
325 that rRNA-based diversity has much greater interspecies redundancy than single-cell functional
326 diversity for the traits we measured[50] or that the measured traits show only a shallow phylogenetic
327 conservation[51]. Indeed, to detect differences in trait diversity indices, it was necessary to first
328 determine the minimum number of single cells per sample required for a robust analysis
329 (Extended Data Fig. 2). It is noteworthy that 250 cells are sufficient for this purpose despite the
330 presence of a high taxonomic diversity. This suggests that many different species and strains rely
331 on similar functions and metabolic pathways[52]. Another explanation could be the extremely high
332 accuracy with which we can measure differences in traits; for example, the consideration of the

333 relative absorbance of multiple light wavelengths on a continuous scale allows us to distinguish
334 individual cells based on slight differences in pigment profiles.

335 It has been recognized that trait variation among individuals must be considered to understand
336 species coexistence[53,54]. These differences among individuals, attributable for example to
337 phenotypic plasticity or phenotypic heterogeneity, affect both niche and fitness differences and
338 therefore coexistence[30,55]. This may explain, for example, why the average cell or colony size of
339 freshwater cyanobacterial species, while affecting both their niche and fitness differences, does
340 not predict their coexistence even in simplified experimental communities[56]. This is consistent
341 with our hypothesis that increasing niche partitioning (evenness in pigment-related traits) and
342 decreasing fitness differences (convergence in metabolic traits) are better defined at a single-cell
343 level when seeking a mechanistic understanding of microbial diversity. Niche and fitness
344 differences are not necessarily independent but may be directly correlated, especially if they are
345 both involved in metabolic activity and more specifically in photosynthesis. This interdependence
346 should become apparent when using more specific and less integrative proxies for fitness. In this
347 study, we find the cell content of S to be a good proxy for fitness, as sulfur globules ensure the
348 availability of electron donors despite the extremely variable environmental conditions in the
349 chemocline[57]. Indeed, sulfur bacteria can accumulate sulfur globules through sulfide uptake, and
350 further use them as electron donors under conditions of sulfide limitation[58,59]. An even partitioning
351 of the light spectrum among cells could also be directly responsible for the more similar values of
352 cellular S (Fig. 5A). Indeed, this form of niche partitioning allows for a more uniform access to
353 photons by decreasing competition for light, and the light-dependent accumulation of sulfur
354 globules[58,59]. In contrast, N and C assimilation rates might not be good proxies for fitness, as a
355 species dominating the uptake of both elements (i.e. *Chromatium okenii*) is present at very low
356 densities in the chemocline of Lake Cadagno, whereas the most abundant species (i.e.,
357 *Chlorobium clathratiforme*) contributes much less to N and C cycles[57]. This may explain why we

358 did not find any relationship between the evenness of pigment-related traits, our measure of niche
359 partitioning, and the variability of C assimilation rates (Fig. 5B). On the contrary, N assimilation
360 rates became increasingly variable among cells with increasing evenness of pigment-related traits
361 (Fig. 5C). This suggests that sulfide (electron donor) limitation in the chemocline and lower layer
362 drives heterogeneity in N assimilation, as previously shown experimentally for the fixation of N_2
363 in a strain of the green sulfur bacterium *Chlorobium phaeobacteroides* isolated from Lake
364 Cadagno[60].

365 Our results support the idea that niche and fitness differences, even at the individual level, can
366 be associated with different sets of traits, which can be measured simultaneously on the same
367 single cells with the methods used in this study. In the chemocline and lower layer, cells show a
368 higher evenness of pigment-related traits (i.e., niche partitioning of the light spectrum under light
369 limitation, as observed in experimental phytoplankton populations[38]) and at the same time a lower
370 variation in cell S content (i.e., more similar fitness). In other words, the environment of deep
371 water layers, which we assume to be highly competitive because of high cell densities along with
372 light and nutrient limitation, requires both high niche partitioning and low fitness differences for
373 organisms to coexist. It should be noted that the strong correlation between these two
374 mechanisms of coexistence (Fig. 5) does not imply a causal relationship. The fact that the cells
375 increasingly partition the light spectrum might not necessarily cause the very similar cell S content
376 but could otherwise indicate that S is a metabolic intermediate of an essential metabolic pathway
377 in the chemocline and lower layer as opposed to the upper layer. Low S content variation in deep
378 layers of the lake might reflect the role of S as an electron donor in anoxygenic photosynthesis.
379 In order to survive there, cells need both enough photons to sustain photosynthesis (leading to
380 high evenness in pigment-related traits) and enough S to channel their metabolism through this
381 element (leading to low variation in cell S content).

382 Our findings have important implications for microbial ecology, biogeochemistry and
383 community ecology. By measuring focal functional traits at the single-cell level, we connect the
384 three above-mentioned fields in a novel conceptual framework that uses niche and fitness
385 differences among individual cells to explain emerging properties at the community and
386 ecosystem level, such as cell densities and standing biomass that sustains natural food webs.
387 Fluorescence and metabolic traits represent fundamentally different axes of variation (see lack of
388 correlation in Fig. 4), which only in combination can explain the occurrence of the corresponding
389 phenotypes in a highly competitive environment. Our results further imply that, despite the high
390 taxonomic complexity of natural microbial communities, single-cell trait-based approaches can
391 explain community-level functions such as the nutrient cycling by reducing this complexity to a
392 few relevant traits[24], i.e., morphological, physiological and metabolic. Such strong links between
393 traits and ecosystem properties across different environments have also been shown to exist at
394 the level of microbiomes, aiding the identification of functional groups or guilds in metagenomic
395 datasets[61,62], as well as ecological clusters based on representative genomes[63]. Our approach
396 might benefit the understanding of ecological communities in general, including those composed
397 of larger organisms, and their links to broad-scale ecosystem functions[64]. Our findings support
398 the need for manipulative experiments and emerging theoretical concepts that explain species
399 coexistence using the fundamental unit at which interactions and selection occur in ecosystems,
400 namely the individual organisms. Such concepts should eventually encompass all levels at which
401 eco-evolutionary processes linked to environmental change occur: i.e. intraspecific trait variation
402 resulting from phenotypic heterogeneity[30] and phenotypic plasticity[55,65], as well as rapid
403 evolutionary adaptations[16].
404
405
406

407 # Methods

408 ## Study site, design, and sampling

409 The study was conducted in August 2014 in the meromictic lake 'Lake Cadagno' located in the
410 Swiss Alps (1923 m above sea level) in Ticino, Switzerland. The maximum depth of the lake is 21
411 m. The water of the lake is infiltrated through gypsum-rich dolomite rock. The infiltrated water
412 therefore transports salts, including sulfate, to the bottom water of the lake. This process
413 establishes an anaerobic, euxinic monimolimnium with high salinity, and an aerobic, low-salinity
414 mixolimnion separated by a permanent chemocline at ca. 10-14 m depth. The chemocline is
415 characterized by sharp gradients and a strong turbidity maximum (ca. 1 m vertical thickness)
416 dominated by populations of purple and green sulfur bacteria, which grow by anoxygenic
417 photosynthesis with sulfide[57]. The exact depth of the turbidity maximum can fluctuate (between
418 days or even within hours) as internal waves build up depending on the wind conditions.

419 Fieldwork was conducted for 3 consecutive days in August 2014. The days had varying
420 sunlight exposure, introducing differences in photosynthetic activity. Each day, a multiparametric
421 profile (temperature, conductivity, pH, salinity, dissolved oxygen, redox potential, and turbidity)
422 was measured in the water column using a YSI 6000 profiler (Yellow Spring, Inc., Yellow Spring,
423 OH). Water was retrieved through gas-tight tubing (Viton, d=6 mm) mounted to the
424 multiparametric probe with a vacuum-driven pump (0.3 bar, 200 mL min^{-1}) during the profile
425 measurements. Samples were retrieved throughout the water column in 1 m steps from 1 to 10
426 m depth and in 0.5 m steps from 10.5 to 15 m depth. An overview on all samples and the
427 associated analysis conducted with them is shown in Supplementary Dataset 1. Samples for
428 stable isotope incubations were filled anoxically into N_2-rinsed, septum-sealed, 250 mL serum
429 bottles without head-space. Samples for chemical and flow cytometric analyses were sampled
430 into 250 mL serum bottles that were cleaned by consecutive rinsing with 0.1% HCl and double-

431 deionized water. Samples for DNA extraction and amplicon sequencing were sampled into 1 L
432 bottles that were cleaned by consecutive rinsing with 2.5% NaClO, 0.1% HCl and 0.2 µm-filtered
433 double-deionized water. Moreover, we measured a profile for photosynthetically active radiation
434 as the percentage of transmission (PAR; Two LI-193SA spherical quantum sensors from LI-COR
435 Ltd, Lincoln, NE) in the afternoon of each day after samples had been retrieved and stable isotope
436 incubations had been placed in the lake.

437 **Chemical analysis**

438 Samples for chemical analyses were filtered onsite through 0.47 µm nitrocellulose filters
439 (Sartorius), stored cold and dark in autoclaved, crimp-sealed serum bottles without headspace,
440 and analyzed within 10 days after sampling. Samples for ammonium were treated differently;
441 these samples were filtered through 0.2 µm polycarbonate filters (GTTP, Millipore) and stored in
442 50 mL polypropylene tubes (consecutively rinsed with 0.1% HCl and 0.2 µm-filtered double-
443 deionized water). Sulfate, chloride and nitrate were analyzed by ion chromatography (Metrohm
444 930 Compact IC Flex with chem. suppression; Metrohm Metrosep A Supp 5 100/4 mm column;
445 conductivity detector). Ammonium was measured with Berthelot's reaction (DIN 38406-5:1983-
446 10). Ortho-phosphate was measured after the formation of a phosphorous-molybdenum blue-
447 complex (DIN EN ISO 6878, 2004). Dissolved inorganic carbon was measured with TOC-L CSH
448 (Shimadzu, Japan). Sulfide was measured one day from 12 mL samples that were immediately
449 transferred to screw-capped tubes containing 0.8 mL 4% zinc acetate solution. These samples
450 were stored in the dark and analyzed colorimetrically using a commercial kit (Spectroquant Sulfide
451 kit, Merck, Schaffhausen, Switzerland). Total particulate nitrogen and carbon were determined by
452 filtration of 100 to 220 mL lake water on pre-combusted (400°C, 3 h) 0.7-µm-pore glass microfibre
453 filters (GF/F, Whatman) and subsequent measurement with isotope ratio mass spectrometry
454 (IRMS). Briefly, the glass microfibre filters for IRMS were immediately stored at -20°C after
455 filtration. After the fieldwork, samples were dried for at least 3 h at 65°C and decalcified in an

456 atmosphere generated by 37% hydrochloric acid. The filters were kept dry in a desiccator at room
457 temperature until further processing. The filters were measured by N_2 and CO_2 released by flash
458 combustion in excess oxygen in an automated Thermo Flash EA 1112 elemental analyzer
459 coupled to an isotopic ratio mass spectrometer (Thermo Delta Plus XP, Thermo Fisher Scientific).
460 Standards were prepared from caffeine and total particulate carbon and nitrogen were calculated
461 from the summed peak areas of masses 44, 45, and 46 for carbon and masses 28, 29, and 30
462 for nitrogen in each sample.

463 **Definition of lake layers**

464 In order to determine the contribution of environmental factors to the distribution of samples in
465 our dataset, we performed a principal component analysis (PCA; Fig. 1B). Individual samples
466 were grouped into three major layers based on their position relative to the chemocline: upper
467 layer, chemocline and lower layer. Samples from above the chemocline were further divided into
468 a surface layer (initial 6 m of Lake Cadagno) and a layer just above the chemocline. This further
469 subdivision was based on the fact that many environmental factors such as conductivity,
470 temperature and oxygen concentrations were stable in the surface layer and gradually changed
471 thereafter (Fig. 1A). In order to maximize differences between layers, we excluded from further
472 analyses the layer between surface and chemocline, which shows intermediate properties and is
473 therefore less clearly defined (Fig. 1B). For the PCA the following environmental factors were
474 considered: depth [m], light [% PAR above the water], temperature [°C], conductivity [S/m], salinity
475 [practical salinity units, psu], oxygen [mg L^{-1}], pH, oxidation-reduction potential [mV], turbidity
476 [formazine turbidity unit, FTU], ammonium [µmol N/L], dissolved inorganic carbon (DIC) [mg/L],
477 ortho-phosphate [µg/L], chloride [mg/L], sulfate [mg/L], particulate nitrogen [µmolN/L], particulate
478 carbon [µmolC/L]. For a small subset of points, individual values were not collected due to device
479 failure. For these points, we calculated the mean of the respective variables using the samples
480 directly above and below the missing value.

481 **Flow cytometry**

482 Using the scanning flow cytometer Cytobuoy (www.cytobuoy.com, Woerden,The
483 Netherlands), we counted and characterized phytoplankton cells from water sampled at different
484 depths during each sampling day. Samples were fixed after stable isotope incubation with a
485 fixative solution containing 0.01% paraformaldehyde and 0.1% glutaraldehyde at pH=7 and were
486 stored at 4°C for 14 weeks until flow cytometric measurements. Cytobuoy contains two laser
487 beams (coherent solid-state sapphire, wavelengths 488 and 635 nm) and three different detectors
488 to measure the fluorescence emitted by photosynthetic pigments (ranges: 668–734 nm, 601–668
489 nm, 536–601 nm). Measurements were triggered by the sideward scattering signal. More details
490 on the instrument can be found elsewhere[34].

491 Cytometry data were successively cleaned by separating living phytoplankton cells from
492 organic debris, suspended solids, and non-photosynthetic (heterotrophic) bacteria. First, a
493 clustering algorithm was applied to a random subset of particles using the library *flowPeaks*[66],
494 and junk clusters were identified visually (characterized by low fluorescence in all pigments).
495 Second, a random forest classifying algorithm[67,68] was trained using the library *randomForest*[66] to
496 identify junk particles in the whole datasets and remove them. In addition, only particles in the
497 size range between 1 µm and 1 mm are reliably characterized by Cytobuoy and were therefore
498 retained in the dataset. The number of living phytoplankton cells identified was divided by the
499 volume processed from each sample to calculate the cell density of photosynthetic cells as a
500 proxy for primary production (ecosystem functioning).

501 **Stable isotope incubations**

502 Stable isotope incubations were performed with anoxically-sampled, freshly-retrieved lake
503 water as described under 'Study site, design, and sampling'. The samples in the serum bottles
504 were spiked with 100 µL stock solutions of ^{15}N-NH_4Cl (25 mM, Sigma) and ^{13}C-$NaHCO_3$ (0.5 M,

505 Sigma) to a final concentration of 10 μM and 200 μM, respectively. The final labeling percentage
506 (^{13}C = 10 to 30%; ^{15}N = 30 to 100%) was calculated for each sample retrospectively based on
507 total NH_4 and DIC measured through chemical analysis in separate water samples. After the
508 addition of the isotopically-labeled substances, serum bottles were attached to a rope that was
509 intruding into the water column and that was mounted on a stably-anchored research platform.
510 The bottles were attached at different positions on the rope in a way that they incubated at the
511 depth from which they had been retrieved. The bottles incubated for an entire day/night/day cycle
512 (approximately 22 h) starting from about noon each day. The serum bottles were retrieved from
513 the lake at the end of the incubations and kept in the dark and on ice until further processing (2
514 to 3 hours). Samples from the incubations were fixed with a fixative solution for later analysis with
515 correlative imaging and flow cytometry (30 to 50 mL).

516 **Correlative imaging for autofluorescence and isotopic enrichment**

517 Single-cell correlative imaging was performed with confocal laser scanning microscopy
518 (CLSM, Leica SP5) for autofluorescence and nanometer-scale secondary ion mass spectrometry
519 (NanoSIMS 50L, CAMECA) for isotopic enrichment within the Laboratory for Biological
520 Geochemistry at EPFL and University of Lausanne. The cells were placed on polycarbonate filters
521 (0.2 μm GTTP, Millipore) by vacuum filtration. The filters had a diameter of 5 mm. The diameter
522 of the filtration area was 3 mm. Etched marks were distributed across the filtration area with laser
523 micro-dissection (PALM micro-dissection, Zeiss 200M equipped with a 355 nm pulsed UV laser)
524 before filtration. The filtered volumes of fixed cells were adjusted between 0.2 to 2 mL to have an
525 evenly distributed, dense single-cell layer. Filtration of cells was immediately followed by filtration
526 of 200 μL double-deionized water to wash the filters. Air-dried filters were coated with a thin film
527 of low melting agarose by dipping them into a 0.1% low melting agarose solution at 35°C,
528 removing excess agarose with a tissue by carefully touching the filter edge with the tissue, and

529 air-drying the filters. The filters were stored dry and in the dark at room temperature until CLSM
530 and NanoSIMS measurements.

531 First, the filters were imaged with CLSM to detect single-cell autofluorescence signals. To this
532 end, cells were initially stained with a general DNA stain to enable cell detection. The filters were
533 embedded in a solution containing the general DNA stain Hoechst (10 µg mL^{-1}) and the antifading
534 reagents Citifluor AF1 (Citifluor Ltd., PA; containing phosphate-buffered saline and glycerol) at 3
535 parts and Vectashield (Vecta Laboratories, CA) at 1 part. The cells on the embedded filters were
536 imaged with CLSM by measuring z-stacks that were adjusted to extend from the filter surface until
537 the top of the biggest cell including large algae and diatoms. For each sample, we obtained
538 mosaics with 4 tiles from about 3 marks. The acquired images had the following dimensions: z-
539 stack step-size of 2 µm (open pinhole) and xy-dimensions between 140 to 145 µm with a
540 resolution of 1024×1024 pixels, resulting in a pixel size of about 0.14 µm (similar to the pixel size
541 of NanoSims images). Images were recorded with bidirectional scanning at 400 Hz with a signal
542 depth of 16 bit and averaging between 3 frames. Fluorescence was recorded in 17 channels
543 varying in excitation (ex.; λ in nm) and emission (em.; λ in nm) during 9 sequential scans.
544 Channels were selected based on (i) lambda scans of emitted light from samples obtained from
545 different depths, (ii) spectra (in nm) of known photopigments (chlorophyll-a, chl-a; phycocyanin,
546 PC; phycoerythrin, PE; fucoxanthin, Fuxo; accessory and decaying pigments, ADP) and (iii) equal
547 fluorescence emission as recorded by the flow cytometer Cytobuoy. Scan 1: ex.: 633 / em.: 660-
548 688 (chl-a/PC) & 725-750; scan 2: ex.: 633 / em.: 668-734 (Cytobuoy 1; chl-a/PE); scan 3: ex.:
549 543 / em.: 660-700 (PE/Fuxo); scan 4: ex.: 543 / em.: 625-650 & 700-750; scan 5: ex.: 488 / em.:
550 601-668 (Cytobuoy 2; PC) & 536-601 (Cytobuoy 3; ADP); scan 6: ex.: 488 / em.: 670-734 (chl-
551 a/Fuxo) & 735-775; scan 7: ex.: 488 / em.: 518-548 & 565-605; scan 8: ex.: 458 / em.: 525-600 &
552 675-725 (chl-a) & reflected light; scan 9: ex.: 405 / em.: 423-503 (Hoechst stain) & reflected light.
553 Imaging was done from higher to lower excitation wavelengths to reduce photobleaching. Gains

554 were kept constant between all images. After imaging, the filters were washed in 100% ethanol
555 and double-deionized water.

556 Next, the filters were prepared for NanoSIMS analysis by sputter-coating with a conductive
557 gold layer (ca. 20 nm). Individual images or image mosaics were measured in the same regions
558 that were imaged with CLSM previously. The respective areas were located using the etched
559 marks and the CCD camera built into the NanoSIMS 50L. The measured areas were pre-
560 sputtered with a Cs^+ primary ion beam of about 4 pA to remove surface contamination, implant
561 Cs^+ ions, and achieve an approximately stable secondary ion emission rate. A primary Cs^+ ion
562 beam with a beam current between 1 and 1.2 pA and a beam diameter of around 150 nm was
563 rastered across the cells for analysis with a dwell time of 5 ms per pixel. Secondary ion images
564 for $^{12}C^{12}C^-$, $^{13}C^{12}C^-$, $^{12}C^{14}N^-$, $^{12}C^{15}N^-$, and $^{32}S^-$ were simultaneously recorded from analysis areas
565 of 30 µm × 30 µm with a resolution of 256×256 pixels, resulting in a pixel size of 0.117 µm. Ten
566 planes from each individual area were measured. Mass resolving power was around 10,000
567 (Cameca definition), enough to resolve all measured isotopes from potential interferences in the
568 mass spectrum.

569 Images were analyzed with a custom-written Matlab code that integrated Schnitzcells[69] with
570 Look@NanoSIMS[70]. First, a maximum intensity projection (MIP) image was created from the z-
571 stack of the individual Hoechst images (scan 9). These MIP images, in which all cells were
572 stained, were aligned with the corresponding NanoSIMS image in Look@NanoSIMS. The
573 alignment was applied to all fluorescence images. Next, the aligned MIP image of scan 9 was
574 used for automated segmentation of cells using Schnitzcells. This segmentation was manually
575 corrected with an overlay image of Hoechst fluorescence and the $^{12}C^{14}N^-$ (biomass) NanoSIMS
576 image for 268 to 662 cells per analyzed image. The segmentation was applied to all fluorescence
577 images and the fluorescence intensities and cell dimensions for all channels were compiled.

578 Finally, the segmentation was imported into Look@NanoSIMS to compile the corresponding
579 isotope counts, isotopic fractions, and cell dimensions for each segmented cell. C and N
580 assimilation for each segmented cell were calculated from their isotopic enrichment $^{15}N/(^{14}N+^{15}N)$
581 and $^{13}C/(^{12}C+^{13}C)$. Overall, images from 17 different samples coming from different days and
582 depths were analyzed corresponding to a total of 6440 cells.

583 **Calculation of trait diversity**

584 To obtain single-cell trait data, we combined stable isotope incubations - by adding $^{15}N\text{-}NH_4+$
585 and $^{13}C\text{-}CO_2$ into samples from different water depths and days - with correlative fluorescence
586 microscopy and NanoSIMS (see previous two sections, as well as Fig. 2A and B). For the
587 calculation of trait diversity, we first selected seven traits in total to avoid collinearity (Spearman
588 correlation <0.7) and thus assure trait complementarity. Four traits were extracted from CLSM
589 images: the mean fluorescence intensity of 3 channels: (i) excitation 543 nm; emission 700-750
590 nm, (ii), excitation 458 nm; emission 525-600 nm, (iii) excitation 458 nm; emission 675-725 nm,
591 and (iv) cell size. Three traits were derived from NanoSIMS data: C and N assimilation, as well
592 as S cell content. Those traits were then transformed to the same unit (mean=0, SD=1) and used
593 to calculate three complementary indices of trait diversity, representing the distribution of
594 individual cells of a microbial community in a high-dimensional trait space (Fig. 2C): TOP (*'Trait*
595 *Onion Peeling'*)[71], TED (*'Trait Even Distribution'*)[71] and FDis (*'Functional Dispersion'*)[72]. TOP, as
596 a measure of trait richness, is sensitive to expansions and contractions of the trait space covered
597 by the individuals of a community. TED, as a measure of trait evenness, is designed to respond
598 to competitive interactions by quantifying niche partitioning as relative regularity in distances
599 between individuals in the trait space. FDis, as a measure of trait divergence, reflects the degree
600 to which individuals are spread around the average phenotype of the community. Please note
601 that the list of traits included in the calculation of trait diversity varied according to the aim (and
602 underlying hypotheses) of each single analysis, always being a selection of the above-mentioned

603 seven traits (see 'Statistical analysis' below for more details). To robustly perform such analyses
604 for a large array of samples and make trait diversity completely independent of the number of
605 cells, we first had to determine the number of single cells in each sample that were required to
606 detect differences in trait diversity indices between upper layer and chemocline (sensitivity
607 analysis). To do this, we selected the two samples (A3 and A13) with the highest number of
608 measured cells (662 and 540, respectively), and bootstrapped 100 times 50 to 450 cells (in
609 intervals of 50) to calculate trait diversity indices (including all above mentioned seven traits). We
610 then fitted linear regressions to test differences in trait diversity between the two samples. The
611 data show that differences for each trait diversity index can be resolved by measuring
612 approximately 250 cells (Extended Data Fig. 2). More specifically, goodness of fit and p-values
613 indicate that at least 200 cells are necessary to distinguish upper layer and chemocline
614 communities with a reasonable degree of confidence, especially with respect to the TED index of
615 trait evenness (Extended Data Fig. 2). As the analyzed 17 samples had at least 269 cells, we
616 decided to bootstrap 250 cells from each of them to further increase the representativeness of
617 our subsamples. For further analyses, we considered the average trait diversity values of the 999
618 bootstrapped communities.

619 **Community analysis and calculation of genetic diversity**

620 Samples for DNA extraction and amplicon sequencing were immediately filtered on 0.2 µm
621 polycarbonate filters with a diameter of 25 mm (Millipore). The filtration volume was adjusted to
622 filter approximately 10^7 to 10^8 cells on each filter, corresponding to 0.2 to 0.4 L per filter. Filters
623 were stored at -20°C until further processing. An empty polycarbonate filter was processed as a
624 negative control (DNA extraction and sequencing).

625 DNA was extracted from the filters using hot phenol/chloroform extraction. Cell lysis was
626 performed with glass beads (106 µm) in 450 µL SDS-lysis buffer (35 mM sodium acetate, 7 mM

627 EDTA, 2.9% SDS, pH 8) mixed with 1 mL phenol:chloroform:isoamyl alcohol (25:24:1 at pH 8) at
628 70°C for 12 min and bead beating for 2 min (Tissue Lyser II, Qiagen). The mixture was centrifuged
629 at 4°C for 3 min at maximum speed and the aqueous phase was transferred to a new reaction
630 tube. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to this tube.
631 This extraction and centrifugation process was repeated for two more rounds. The final round of
632 extraction was performed by adding chloroform:isoamyl alcohol (24:1) and again collecting the
633 aqueous phase (~200 µL). DNA was then precipitated from the recovered aqueous phase by
634 adding 0.1 volume of 3 M sodium acetate (pH 8), gently mixing, and adding 2 volumes of 100 %
635 ethanol and 1 µL of glycogen (20 mg/mL). Samples were incubated at 20°C for at least 3 hours,
636 centrifuged at 4°C for 60 min at maximum speed and the supernatant was removed. The pellet
637 was then washed twice with 200 µL of 80% ethanol, centrifuged at 4°C for 5 min and the
638 supernatant removed. The pellet was dried under vacuum for 7 min and dissolved in 25 µL
639 nuclease-free water. The DNA concentration was determined fluorometrically (Qbit, Invitrogen)
640 and the concentration was normalized between samples to a maximum of 20 ng/µL by the addition
641 of appropriate volumes of nuclease-free water.

642 Sequencing library preparation and DNA amplicon sequencing were performed by MR DNA
643 (Molecular Research LP, Shallowater, TX) using a modified bTEFAP® library preparation
644 protocol[73] and sequencing on an Illumina MiSeq platform with a minimum of 20,000 reads per
645 sample and read lengths of 2x300 bp. The 16S sequencing library was prepared with Bact341F:
646 5'-CCTACGGGNGGCWGCAG-3' and Bact785R:5'-GACTACHVGGGTATCTAATCC-3' [74], and
647 the 18S sequencing library was prepared with euk516f:GGAGGGCAAGTCTGGT and
648 euk1055r:CGGCCATGCACCACC[75]. A mock bacterial community was sequenced as a positive
649 control (BEI Resources, Genomic DNA from Microbial Mock Community B, staggered, high
650 Concentration, v5.2H, for Whole Genome Shotgun Sequencing, HM-277D).

651 We had to use two slightly different methods to process the sequence data of the two genes.
652 The detailed protocol files describing the application, version, and parameters of the data
653 processing for both genes can be found in the Appendix. In the first step, the sequence data were
654 filtered to remove PhiX residues and sequences of low complexity. For the 16S amplicon data,
655 read pairs were merged and primer site sequences were removed. For the 18S amplicon data,
656 only the primer site trimmed forward reads were used as the reads could not be merged. We used
657 UNOISE3[76] to obtain zOTUs (zero-radius OTUs), which are considered equivalent to amplicon
658 sequence variants (ASVs). We used SINTAX[76] in combination with the reference SILVA SSU
659 (v138 - https://www.arb-silva.de/) for 16S and pr2 SSU (v4.14.0 - https://pr2-database.org/) for
660 18S to predict the taxonomic affiliation of the zOTUs. The sequence data is deposited at
661 NCBI/ENA under the BioProject number PRJNA1014378.

662 Finally, we calculated genetic diversity indices (Chao1 and Simpson) using the library *phyloseq*[66]
663 from unclustered zOTUs data.

664 **Statistical analysis**

665 To test our first hypothesis, that ecosystem functioning is better explained by trait diversity than
666 taxonomic diversity, we used a model averaging approach[77]. To this end, we formulated a list of
667 linear models including as predictors all possible combinations (=16) of three trait diversity metrics
668 (TOP, TED and FDis, calculated using three pigment-related traits, cell size, C assimilation, N
669 assimilation, and S cell content) and day of sampling. In addition, in a second list of linear models
670 we included as predictors all possible combinations (=32) of four genetic diversity metrics (Chao1
671 and Simpson index using 16S and 18S sequencing data) and the day of sampling. For both lists
672 of models, the response variable was cell density relative to the maximum achieved in a given
673 layer. Models with a difference in the Akaike information criterion corrected for small sample sizes
674 (delta AICc<2) were defined as the most parsimonious set of models[78] and were consequently

675 used for model averaging (separately for trait and taxonomic diversity). Model selection and
676 averaging were performed using the *MuMIn* R-package[66]. To address our second hypothesis that
677 the highly competitive environment in deep layers of the lake drives increasing niche partitioning
678 in pigment-related traits, and decreasing fitness differences in metabolic traits, we performed one-
679 way ANOVA analyses and post hoc Tukey tests using the *TukeyHSD* function of the R-package
680 *stats*. This procedure takes multiple comparisons into account when calculating p-values[66]. This
681 allowed us to assess differences in the evenness of pigment-related traits and the divergence of
682 metabolic traits between the three layers. To test our third hypothesis that single-cell niche and
683 fitness differences are negatively related across depth, we fitted three simple linear models (*lm*
684 function) with the evenness of pigment-related traits as a predictor and the coefficient of variation
685 in S content and C and N assimilation rates as response variables.

Acknowledgements

We thank the personnel of Piora Centro Biologia Alpina, especially Mauro Tonolla, for making it possible to use the laboratory facilities and housing during field work at lake Cadagno. We thank Martin Ackermann, Jukka Jokela and Dave Johnson (all Eawag) for discussions and support during the initial stages of the project. We thank Ilona Szivak (Eawag) for acquiring CLSM images. We thank the AUA chemical analysis teaching laboratory at Eawag for chemical analysis of sulfate, chloride, nitrate, ortho-phosphate, and dissolved inorganic carbon. We thank Nicola Storelli for reviewing the manuscript. This study was supported by Eawag, by the Swiss National Science Foundation Project 31003A_144053, an Eawag academic transition grant to F.S., a Synthesis Grant of the ETH Zurich Center for Adaptation to a Changing Environment (ACE) to F.S, and a Marie-Curie-Intra-European fellowship for career development (FP7-MC-IEF;271929;Phenofix) provided by the European Union to F.S. The NanoSIMS instrument in the Laboratory for Biological Geochemistry was funded in part by the European Union through a European Research Council Advanced Grant 246749 (BIOCARB) to A.M.

Credit Authorship

Simone Fontana: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration

Désirée A. Schmitz: Formal analysis (correlative imaging data), Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization

Michael Daniels: Formal analysis (PCA analysis), Writing - Review & Editing, Visualization

Francesco Danza: Investigation (field work), Resources, Writing - Review & Editing

Thomas Röösli: Software (for correlative image analysis), Writing - Review & Editing

Hannah Bruderer: Methodology, Investigation, Writing - Review & Editing

Jean-Claude Walser: Formal analysis (sequencing data), Data Curation, Writing - Review & Editing

Julien Dekaezemacker: Investigation (IRMS measurements), Writing - Review & Editing

Stephane Escrig: Investigation (NanoSIMS measurements), Writing - Review & Editing

Anders Meibom: Resources (NanoSIMS), Writing - Review & Editing, Supervision

Francesco Pomati: Conceptualization, Writing - Review & Editing, Supervision

Frank Schreiber: Conceptualization, Methodology, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition