

Look@NanoSIMS – a tool for the analysis of nanoSIMS data in environmental microbiology

Lubos Polerecky,* Birgit Adam, Jana Milucka,
Niculina Musat, Tomas Vagner and
Marcel M. M. Kuypers
Max-Planck Institute for Marine Microbiology,
Celsiusstrasse 1, 28359, Bremen, Germany.

Summary

We describe an open-source freeware programme for high throughput analysis of nanoSIMS (nanometre-scale secondary ion mass spectrometry) data. The programme implements basic data processing and analytical functions, including display and drift-corrected accumulation of scanned planes, interactive and semi-automated definition of regions of interest (ROIs), and export of the ROIs' elemental and isotopic composition in graphical and text-based formats. Additionally, the programme offers new functions that were custom-designed to address the needs of environmental microbiologists. Specifically, it allows manual and automated classification of ROIs based on the information that is derived either from the nanoSIMS dataset itself (e.g. from labelling achieved by halogen *in situ* hybridization) or is provided externally (e.g. as a fluorescence *in situ* hybridization image). Moreover, by implementing post-processing routines coupled to built-in statistical tools, the programme allows rapid synthesis and comparative analysis of results from many different datasets. After validation of the programme, we illustrate how these new processing and analytical functions increase flexibility, efficiency and depth of the nanoSIMS data analysis. Through its custom-made and open-source design, the programme provides an efficient, reliable and easily expandable tool that can help a growing community of environmental microbiologists and researchers from other disciplines process and analyse their nanoSIMS data.

Introduction

Nanometre-scale secondary ion mass spectrometry (nanoSIMS) allows quantitative analysis of elemental and isotopic composition of a sample with a submicrometre spatial resolution, and has been applied in diverse research fields ranging from geology and cosmochemistry to material and life sciences (e.g. Messenger *et al.*, 2003; Lechene *et al.*, 2006; McKeegan *et al.*, 2006; Herrmann *et al.*, 2007; Lozano-Perez *et al.*, 2008; Matzel *et al.*, 2011). In environmental microbiology, nanoSIMS is used to determine metabolic activities of individual microbial cells via incubations with substrates labelled with stable or radioactive isotopes (e.g. Musat *et al.*, 2008; Boxer *et al.*, 2009; Dekas *et al.*, 2009; Orphan and House, 2009; Ploug *et al.*, 2010; 2011). Furthermore, by combining this technique with fluorescence *in situ* hybridization of the 16S or 23S rRNA molecules (Orphan *et al.*, 2001), or by using isotopically labelled probes (e.g. I-labelled oligonucleotide probes; Li *et al.*, 2008) or halogenated tyramides (e.g. Br- and F-labelled tyramides; Behrens *et al.* 2008; Musat *et al.*, 2008; reviewed in Musat *et al.*, 2010), one can link the measured metabolic activity of single cells with their phylogenetic identity. This technique therefore opens many exciting possibilities for environmental microbiologists to study the interactions between microorganisms and their environment.

Presently, several software packages for nanoSIMS data analysis are available. Commercially available options include the WinImage software supplied by Cameca, the manufacturer of the nanoSIMS instrument (<http://www.cameca.com>), and the L'Image software developed by L. R. Nittler (http://www.dtm.ciw.edu/users/nittler/limage_manual.pdf). A freely available alternative is the Open_MIMS plug-in for ImageJ, developed at the National Resource for Imaging Mass Spectrometry at Harvard (<http://www.nrims.hms.harvard.edu>). These programmes offer many basic and advanced data processing and analytical functions, including display of scanned planes for the detected masses, drift-corrected plane accumulation, manual or automated definition of regions of interest (ROIs), ROI classification, quantification of the elemental and isotopic composition of the sample, and export of these as images, histograms, profiles or scatter plots, both in graphical and tabulated formats. Although these functions allow analysis of nanoSIMS data with

Received 3 April, 2011; revised 14 November, 2011; accepted 25 November, 2011. *For correspondence. E-mail lpolerec@mpi-bremen.de; Tel. (+49) (0) 421 2028 834; Fax (+49) (0) 421 2028 834.

© 2012 Society for Applied Microbiology and Blackwell Publishing Ltd

various levels of complexity, they do not provide a direct interface with a variety of other analyses that are routinely performed by environmental microbiologists (e.g. epifluorescence microscopy, TEM or SEM imaging, statistical analysis).

In microbial ecology, nanoSIMS is predominantly used to characterize the function and activity of specific microbial populations in a given habitat, and to understand how these are controlled by environmental conditions. In an increasing number of studies, single-cell measurements are extrapolated based on the cell abundance to estimate the contribution of a specific population to microbially mediated processes occurring in the environment. This is a delicate task considering that the activities of individual cells in a given microbial population and habitat vary greatly not only among different types of cells but also within one species (Lechene *et al.*, 2007; Popa *et al.*, 2007; Musat *et al.*, 2008; House *et al.*, 2009; Orphan and House, 2009; Ploug *et al.*, 2010; 2011). Therefore, to extrapolate and compare the results from single-cell analyses with those obtained by bulk analyses, a confident estimation of the activity of an 'average' cell in the population is required. This can be achieved by the analysis of a large number of cells. To aid such analysis, software tools should allow rapid and reproducible definition of many cells in a single scanned field of a sample, as well as offer the possibility to rapidly combine and statistically analyse results obtained for many fields and many samples.

Due to the fact that nanoSIMS data may not always contain sufficient information required for accurate and reliable definition and classification of objects of interest in the sample (e.g. different types of cells or sub-cellular structures), nanoSIMS measurements are often complemented by other high resolution imaging techniques. These include, e.g. scanning (SEM) or transmission (TEM) electron microscopy, atomic force microscopy (AFM), or, most routinely in environmental microbiology, fluorescence microscopy combined with fluorescence *in situ* hybridization (FISH; e.g. Orphan *et al.*, 2001). Therefore, to facilitate direct integration of these techniques with nanoSIMS, the software tools should allow import of images generated by these techniques and their accurate alignment with the ion images obtained by nanoSIMS.

Currently, the use of nanoSIMS in environmental microbiology is rapidly expanding due to the easier access to nanoSIMS facilities. The growing volumes of generated data require high throughput processing that can be

readily adjusted to the specific needs of the investigator. Thus, to promote widening of the scientific community that uses nanoSIMS and SIMS technology, as well as to allow rapid implementation of new types of analysis, the software tools should be easily accessible and expandable.

Here we describe a Matlab-based freeware and open-source programme for high throughput analysis of nano-SIMS data. The programme, called Look@NanoSIMS, can be downloaded from the website of the Max-Planck Institute in Bremen or from <http://www.microsen-wiki.net/doku.php?id=LANS>. It offers the same basic and advanced analytical functions as other programmes presently available, and can therefore be used as an alternative. However, our software also includes new functions that were custom-designed to address the needs of environmental microbiologists. Specifically, it allows definition and classification of ROIs based on the information provided externally (e.g. as a FISH, SEM, TEM or AFM micrograph), and implements post-processing routines for rapid synthesis and comprehensive statistical analysis of results obtained from multiple datasets. After validation of the programme, we demonstrate on two examples how these new analytical functions increase flexibility, efficiency and depth of the nanoSIMS data analysis.

Results

Programme validation

Validation was done by comparing the results obtained by Look@NanoSIMS with those obtained by two other programmes: Open_MIMS and WinImage. Comparison with L'Image was not possible because this software was not available to us. Datasets used for validation are described in detail in the next section.

Drift-corrected accumulation of secondary ion images by Look@NanoSIMS, Open_MIMS and WinImage produced results that were generally not identical but visually indistinguishable (Fig. 1). Detected drift corrections differed by a maximum of ± 1 pixel in not more than three planes per dataset. However, this depended on the type of ion image and area within the ion image used as a basis for drift correction, and not on the programme used. When drift correction was not applied, or was forced to be equal, the accumulated ion images calculated by all programmes were identical (data not shown).

Isotopic ratios calculated in ROIs were identical for all three programmes when the drift correction and ROI definition were identical (data not shown). However, when

Fig. 1. Comparison of images of accumulated secondary ion counts and their ratios, as calculated by Look@NanoSIMS, Open_MIMS and WinImage for an identical dataset. The ion counts in the top and right margins are lower due to the sample drift during data acquisition, which reached up to 17 pixels ($\sim 2 \mu\text{m}$). This effect is represented by an increased pixel noise level in the ratio images. The corresponding images are displayed in identical gray scales. Dwelling time in each pixel was 1 ms, the number of scanned planes was 35. Outlines and numbers mark ROIs for which the programmes were compared in more detail (see Table 1).

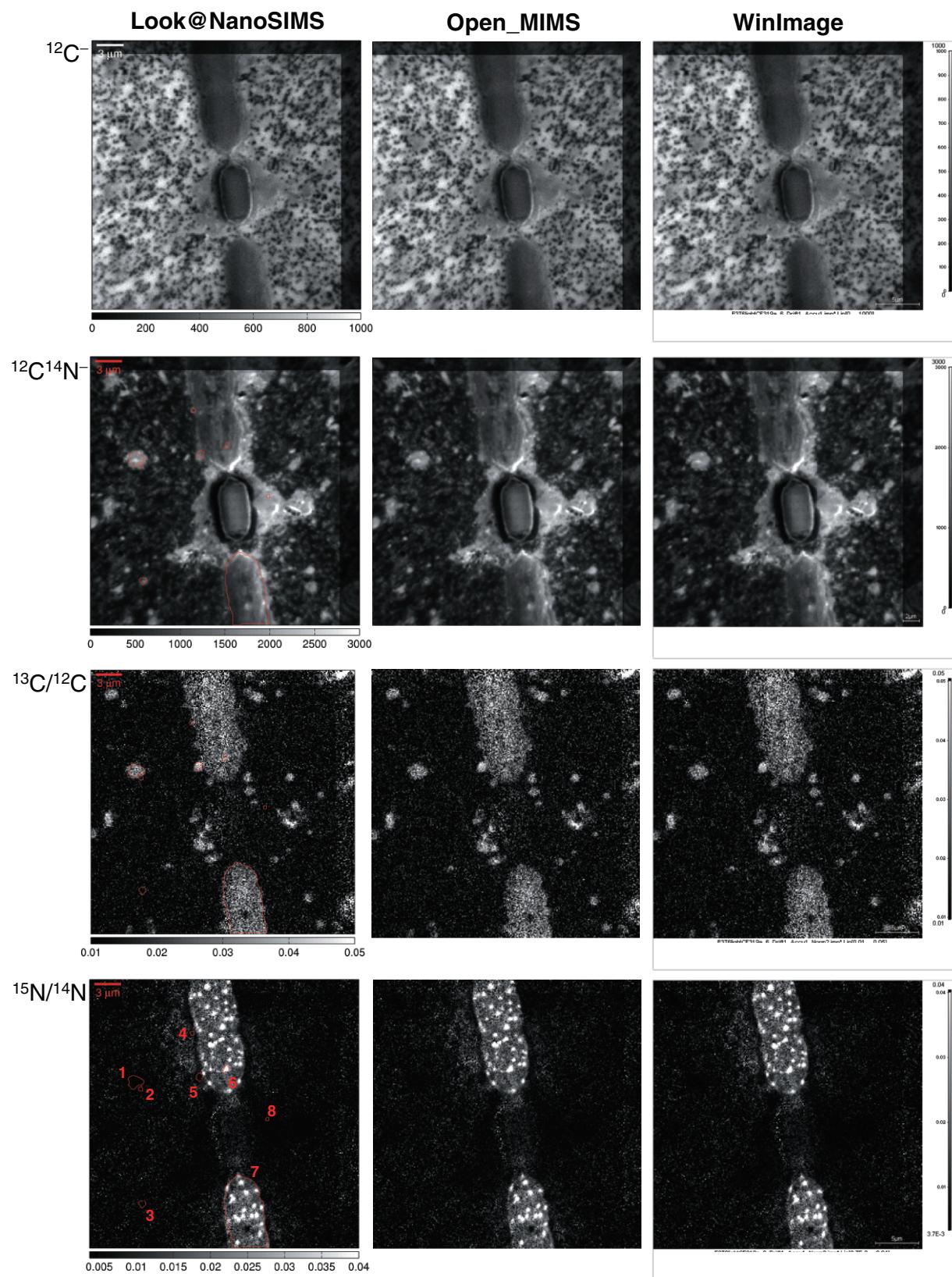


Table 1. Comparison of isotopic ratios (r) calculated by Look@NanoSIMS (LANS), Open_MIMS (OM) and WinImage (WI).

ROI	Pixels	$r = {}^{13}\text{C}/{}^{12}\text{C} \times 100$			$\Delta (\%)$			$\Delta_5 (\%)$	
		LANS	OM	WI	OM	WI	$\delta_{\text{LANS}} (\%)$	min	max
1	189	2.88	2.83	2.75	-1.8	-4.6*	1.9	-3.1*	6.3*
2	29	1.65	1.70	1.79	3.3	8.8*	8.3	-28.0*	11.1*
3	39	1.02	1.10	1.00	7.7	-2.2	8.8	-11.3*	14.2*
4	13	2.05	2.27	1.75	10.9	-14.4*	11.8	-14.7*	59.8*
5	65	3.64	3.65	3.65	0.3	0.3	3.0	-6.3*	4.2*
6	22	3.16	3.41	3.29	7.8*	4.1	7.3	-11.2*	6.8
7	2216	2.63	2.64	2.64	0.4	0.3	0.8	-0.2	0.2
8	14	1.08	1.07	1.03	-0.7	-4.7	12.7	-19.7*	20.1*

ROI	Pixels	$r = {}^{15}\text{N}/{}^{14}\text{N} \times 100$			$\Delta (\%)$			$\Delta_5 (\%)$	
		LANS	OM	WI	OM	WI	$\delta_{\text{LANS}} (\%)$	min	max
1	189	0.382	0.376	0.377	-1.5	-1.2	2.8	-0.6	3.2*
2	29	0.376	0.380	0.353	1.1	-6.2	8.3	-7.2	16.7*
3	39	0.399	0.407	0.394	2.0	-1.3	8.2	-5.6	8.1
4	13	0.749	0.701	0.712	-6.4	-5.0	10.7	-15.7*	23.0*
5	65	0.572	0.592	0.526	3.5	-8.0*	4.6	-2.2	31.8*
6	22	6.285	6.090	6.436	-3.1*	2.4	2.5	-13.4*	5.4*
7	2216	2.101	2.070	2.098	-1.5*	-0.2	0.44	-0.2	0.1
8	14	0.374	0.398	0.358	6.5	-4.1	9.7	-19.4*	19.2*

ROIs were defined independently in each programme by free-hand drawing.

Column Δ shows the percentage differences between the ratios calculated by Look@NanoSIMS and those calculated by OM and WI [$\Delta = (r - r_{\text{LANS}})/r_{\text{LANS}} \times 100\%$]. Column δ_{LANS} shows the Poisson percentage error of the ratio calculated by Look@NanoSIMS (Eq. 8). Column Δ_5 shows the minimum and maximum percentage differences between the ratios calculated in the original ROIs and in ROIs modified by the addition and removal of five pixels at the ROI perimeter.

Number of pixels in the original ROIs defined in Look@NanoSIMS are also shown.

Outlines of the original ROIs are shown in Fig. 1.

Asterisks mark significant differences between ratios calculated by LANS and OM or WI (when $|\Delta| > \delta_{\text{LANS}}$), or significant differences between ratios calculated by LANS for the original ROI and the modified ROI with 5 pixels removed or added ($|\Delta_5|_{\text{min}} > \delta_{\text{LANS}}$ or $|\Delta_5|_{\text{max}} > \delta_{\text{LANS}}$)

automatic drift correction was applied and the ROIs were drawn in each programme independently, the results differed by a few percent (Table 1, compare Δ and δ_{LANS}). Occasionally, the differences exceeded the theoretical precision of the ratio estimation given by the Poisson percentage error, and were therefore significant (Table 1, Δ values marked by an asterisk). This was, however, entirely due to the fact that the ROIs, defined in each programme independently, were not identical but differed by a few pixels.

To verify that such small differences in ROI definitions can lead to significant differences in the calculated ratios, random pixels were added and removed at the perimeter of the ROIs originally defined by the user, and the ratios were recalculated. This sensitivity analysis showed that the differences between ratios calculated for the original and modified ROIs could indeed become significant already for as few as five added or removed pixels (Table 1, Δ_5 values marked by an asterisk).

Analysis of C and N flow in a complex microbial community

Processing and statistical analysis of multiple nanoSIMS datasets in the Look@NanoSIMS programme is demon-

strated on data obtained for environmental samples from the Baltic Sea. The microbial community in the analysed samples was dominated by the cyanobacterium *Aphanizomenon* sp. and heterotrophic bacteria of the *Cytophaga* clade, as revealed by results obtained from clone libraries, FISH and autofluorescence measurements (B. Adam, R.A. Foster, N. Musat, H. Ploug, L. Polerecky, C.L. Moraru, G. Lavik and M.M.M. Kuypers, unpublished). Before nanoSIMS analysis, the cells were incubated in the light in seawater enriched with $\text{H}^{13}\text{CO}_3^-$ and ${}^{15}\text{N}_2$. Subsequently, they were hybridized with the oligonucleotide probe CF319a specific for *Cytophaga*, followed by a deposition of ${}^{19}\text{F}$ -labelled tyramides (see *Experimental procedures* for details).

Accumulated images of secondary ions (${}^{12}\text{C}^-$, ${}^{13}\text{C}^-$, ${}^{19}\text{F}^-$), molecular ions (${}^{12}\text{C}^{14}\text{N}^-$, ${}^{12}\text{C}^{15}\text{N}^-$) and secondary electrons (SE) demonstrated high complexity of the studied environmental sample (see Fig. S1). However, by combining the basic diversity information mentioned above with the visual information generated by the programme (Fig. 2), this complexity could be untangled. The large cells clearly visible in the SE, ${}^{12}\text{C}^-$ and ${}^{12}\text{C}^{14}\text{N}^-$ images corresponded to a heterocyst and vegetative cells of the cyanobacterium *Aphanizomenon* sp. (Figs S1 and

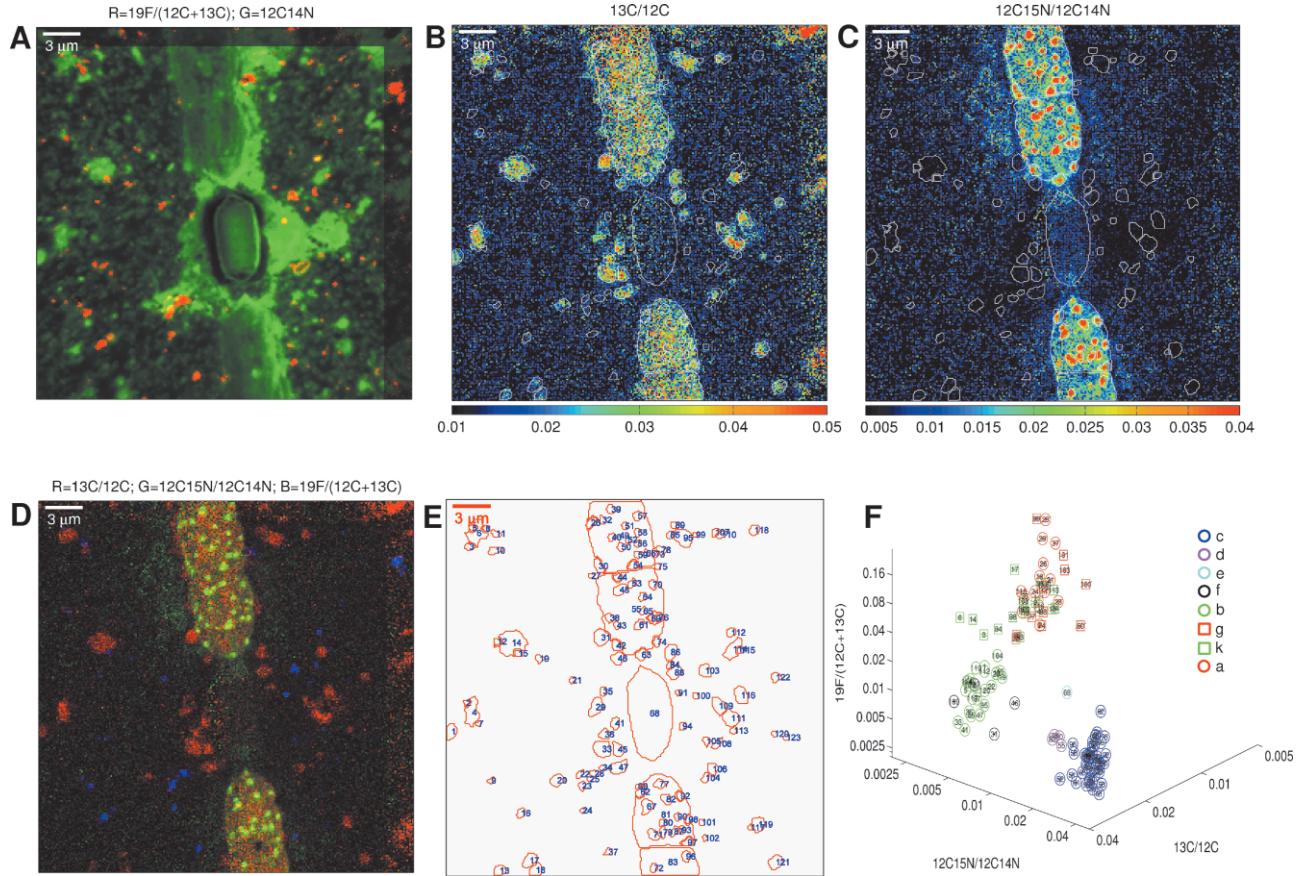


Fig. 2. Example images and graphs generated by Look@NanoSIMS for the environmental sample from the Baltic Sea.
A. Overlay between the $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratio (red) and total accumulated $^{12}\text{C}^{14}\text{N}$ ion counts (green).
B. Image of the $^{13}\text{C}/^{12}\text{C}$ ratio, with ROI outlines included (white).
C. Image of the $^{15}\text{N}/^{14}\text{N}$ ratio, calculated as $^{12}\text{C}^{15}\text{N}/^{12}\text{C}^{14}\text{N}$, with ROI outlines included (white).
D. Composite RGB image, constructed by combining the ratios $^{13}\text{C}/^{12}\text{C}$ (red), $^{15}\text{N}/^{14}\text{N}$ (green) and $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ (blue).
E. ROI outlines defined for the cells identified in the image. The ROI identification numbers are sorted such that the horizontal positions of their centres increase from left to right.
F. Scatter plot of the same ratios as in the RGB image in panel (D). Numbers inside the symbols correspond to ROI identification numbers (see panel E), different colours and symbols correspond to different ROI classes (see Table 2). The axes are log-scaled to enhance visualization of the clustering.

2A). The $^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$ images confirmed that this cyanobacterium fixed inorganic carbon and nitrogen during the incubation (Fig. 2B and C). The $^{15}\text{N}/^{14}\text{N}$ image further indicated the presence of highly ^{15}N -enriched spots inside the vegetative cells. These spots resembled nitrogen-storing cyanophycin granules observed previously for vegetative cells of *Anabaena oscillarioides* and *Trichodesmium* sp. (Popa *et al.*, 2007; Finzi-Hart *et al.*, 2009). Based on this similarity in shape, size and their exclusive presence in the vegetative cells, we refer to these ^{15}N -enriched spots as cyanophycin granules. Cells belonging to the *Cytophaga* group were identified based on colocalization of the $^{12}\text{C}^{14}\text{N}^-$ signal and highly elevated $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratios (Fig. 2A). The first signal indicated the presence of biomass, while the second resulted from the deposition of ^{19}F -labelled tyramides by the *Cytophaga*-specific probe CF319a (Manz *et al.*, 1996;

Musat *et al.*, 2008; 2010). Cells with elevated $^{13}\text{C}/^{12}\text{C}$ ratios (Fig. 2B) that were not hybridized by the CF319a probe remained unidentified. However, because of their ability to fix inorganic carbon, these cells are hereafter referred to as putative photosynthetic (PS) bacteria.

Regions of interest corresponding to the cells and cell organelles identified above were defined based on the $^{12}\text{C}^{14}\text{N}^-$ image as the primary template. The other elemental or isotopic ratio images, alone or overlaid with the $^{12}\text{C}^{14}\text{N}^-$ image (e.g. Fig. 2A), were used to fine-tune the ROI definition. This was possible because of the high contrast between the nitrogen content in the biological material (detected by nanoSIMS as a molecular ion $^{12}\text{C}^{14}\text{N}^-$) compared with that in the Au-Pd-coated polycarbonate filter, and because of the specific elemental or isotopic enrichment of the different types of cells. Due to this high contrast, most ROI outlines were defined

Table 2. Statistics characterizing the studied cell populations from the Baltic Sea environmental sample.

Class ID	Cell population	$^{13}\text{C}/^{12}\text{C} \times 100$	$^{15}\text{N}/^{14}\text{N} \times 100$	$^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C}) \times 100$	n			
c	Cyanophycin granules	2.66 ± 0.04	[1]*	3.04 ± 0.05	[1]*	0.99 ± 0.11	[3]	139
d	Vegetative cells	2.38 ± 0.07 [‡]	[2]*	1.89 ± 0.05 [‡]	[2]*	1.07 ± 0.14 [‡]	[3]	28
e	Heterocysts	1.19 ± 0.31	[3–4]*	0.62 ± 0.02	[3]*	0.67 ± 0.11	[3]	2
f	PS bacteria attached to d	2.62 ± 0.13	[1–2]*	0.75 ± 0.08	[3]*	1.14 ± 0.24	[3]	8
b	Free-living PS bacteria	2.43 ± 0.05	[2]*	0.36 ± 0.01	[4]	1.23 ± 0.10	[3]	86
g	CF319a attached to d	1.33 ± 0.07	[3]*	0.81 ± 0.06	[3]*	5.50 ± 0.86	[2]*	23
k	CF319a attached to b	1.52 ± 0.08	[3]*	0.38 ± 0.01	[4]	4.27 ± 0.33	[2]*	31
a	Free-living CF319a	1.03 ± 0.02	[4]	0.37 ± 0.01	[4]	10.71 ± 0.98	[1]*	73
n	All cells at time-point 0	1.01 ± 0.13	[4]	0.34 ± 0.05	[4]	—		71

[‡] Calculated excluding the cyanophycin storage granules.

Shown are mean values ± standard errors of the mean, ranking with respect to the mean (in brackets), and the numbers of analysed cells (n). To account for the fact that the mean values as well as variances of the $^{15}\text{N}/^{14}\text{N}$ and $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratios for the different cell populations differed by approximately an order of magnitude, class ranking for these variables was derived from the log-transformed values.

Asterisk indicates significant variation of the ratio among cells within the population (ANOVA, $P < 0.05$). The difference in ranking of the attached and free-living *Cytophaga* cells for the $^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$ ratios indicates a significant effect of attachment on their ^{13}C and ^{15}N enrichment, suggesting that these cells may be able to more rapidly gain fresh organic C and N by attaching to the bacterial cells that fix inorganic C (cyanobacteria and the putative PS bacteria) and N (cyanobacteria) respectively.

A similar benefit of attachment to the cyanobacteria is suggested also for the putative PS bacteria with respect to N transfer.

The ranking of cell populations with respect to the $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratio confirms consistency of the ROI classification.

semi-automatically using interactive thresholding, which allowed reproducible definition of tens of cells within minutes. Manual drawing of ROI outlines was employed only when cells were next to each other and thus difficult to detect semi-automatically. At the end, ROIs together with their identification numbers were exported graphically (Fig. 2E) and classified in accordance with the cell identities mentioned above. The *Cytophaga* and putative PS cells were further differentiated as free-living or attached (Table 2).

Elemental [$^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$] and isotopic ($^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$) ratios for each ROI were calculated using the total ion counts accumulated over ROI pixels and planes selected for analysis (Eqs 6–7). This calculation was justified by the fact that none of the ratios exhibited a significant trend with depth, as verified in the programme using the depth profile function (see Fig. S2). A scatter plot of these ratios revealed that the ROIs clustered according to the classes identified above (Fig. 2F). This suggested that cells from one population were functionally similar among each other but different when compared with cells from a different population.

To assess these similarities and differences quantitatively and with a higher level of confidence, five additional nanoSIMS datasets were analysed in the same way as described above. Four of these datasets corresponded to different fields of view of the same sample and one to a field of view of a control sample (taken at the beginning of the incubation). Subsequently, results generated by these separate analyses were combined into one multivariate dataset, which included depth profiles and mean values of the $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$ and $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratios for a total of 460 ROIs, differentiated into eight classes and two treatment categories (incubated versus control samples).

Finally, using the metafile processing function of the Look@NanoSIMS programme, these combined data were visualized in a scatter plot (Fig. 3) and statistically analysed (Table 2), which allowed in-depth assessment of differences between the individual ROIs, ROI classes and treatments. All of these steps were done in a few seconds,

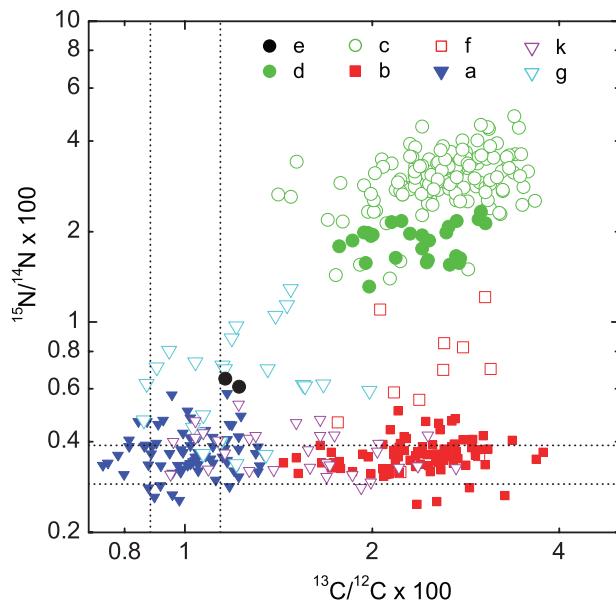


Fig. 3. Scatter plot of the $^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$ ratios for cells from the Baltic Sea environmental sample. Plot was generated by combining the results of the analysis of five fields of view in the incubated sample (390 cells) and one field of view in the control sample (70 cells), using the metafile processing function of the Look@NanoSIMS programme. Bands indicated by dashed lines represent mean ± standard deviation of the corresponding ratios for the control cells. See Table 2 for the definition of cell classes.

which demonstrated the utility of the programme for comprehensive and high throughput analysis of the nanoSIMS data.

Carbon uptake in a mixed bacterial community identified by FISH

A unique feature of the Look@NanoSIMS programme is the possibility to use an external image as a basis for ROI definition. This is demonstrated on a dataset obtained for a sample prepared artificially by mixing cells of *Escherichia coli* (*Gammaproteobacteria*) and *Azoarcus* sp. (*Betaproteobacteria*) grown in pure cultures. The *E. coli* cells were grown on ^{13}C -labelled glucose, whereas the *Azoarcus* sp. cells were cultured without a labelled substrate. After mixing and filtering, the cells were hybridized with the oligonucleotide probe specific for *Gammaproteobacteria* (Gam42a), followed by a deposition of fluorescently (Oregon Green) labelled and ^{19}F - containing tyramides and counterstaining with DAPI. Subsequently, the cells were marked with a laser micro-dissection microscope, imaged using a fluorescence microscope, and analysed by nanoSIMS (see *Experimental procedures* for details).

In the fluorescence image, the *Azoarcus* sp. cells are shown in blue due to DAPI fluorescence, whereas the *E. coli* cells are shown in cyan due to an overlay of green and blue fluorescence from Oregon Green and DAPI respectively (Fig. 4A). The same cells were identified in the secondary ion image detected by nanoSIMS (here $^{32}\text{S}^-$; Fig. 4B), which was possible because of the high abundance of sulfur in the biological material compared with that in the Au-Pd-coated polycarbonate filter.

To align the external (FISH) image with the nanoSIMS ion image, reference points that correspond to each other were manually defined in both images (numbers in Fig. 4A and B). Due to unequal magnification and distortion between the two images, 10 reference points were required to obtain satisfactory alignment. Subsequently, the alignment was automatically calculated and visually verified directly in the Look@NanoSIMS programme (Fig. 4C).

The rest of the data analysis proceeded as if the aligned FISH image was part of the nanoSIMS dataset. First, ROIs were defined by interactive thresholding or free-hand drawing using the FISH image template (Fig. 4D). Subsequently, the FISH image was overlaid with the $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ and $^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$ ratio images to visually inspect the relationship between the identities of the detected cells and their isotopic enrichment (Fig. 4E and G). Finally, the ^{13}C and ^{19}F abundance and the Oregon Green fluorescence intensity (OGFI) in the defined ROIs were exported and displayed in scatter plots. The $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ vs. OGFI plot (Fig. 4F) revealed consistency of the phylogenetic cell identification based on the HISH and FISH labelling procedures, and showed that while the FISH signal in the target *E. coli* cells was on average 40-fold higher than in the non-target *Azoarcus* sp. cells, the $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratio differed between the two populations on average only by a factor of 6. The $^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$ vs. OGFI plot (Fig. 4H) confirmed that all *E. coli* cells were highly ^{13}C -enriched, whereas no significant ^{13}C -enrichment was detectable for the *Azoarcus* sp. cells, as expected from the sample preparation procedure.

($^{12}\text{C} + ^{13}\text{C}$) vs. OGFI plot (Fig. 4F) revealed consistency of the phylogenetic cell identification based on the HISH and FISH labelling procedures, and showed that while the FISH signal in the target *E. coli* cells was on average 40-fold higher than in the non-target *Azoarcus* sp. cells, the $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratio differed between the two populations on average only by a factor of 6. The $^{13}\text{C}/(^{12}\text{C} + ^{13}\text{C})$ vs. OGFI plot (Fig. 4H) confirmed that all *E. coli* cells were highly ^{13}C -enriched, whereas no significant ^{13}C -enrichment was detectable for the *Azoarcus* sp. cells, as expected from the sample preparation procedure.

Discussion

Look@NanoSIMS is a new software for the analysis and processing of nanoSIMS data. It implements most of the functions provided by other commonly used programmes. However, it also offers additional features that make it particularly useful in the field of environmental microbiology (Table 3).

First, it allows external images to be used as template for the definition of ROIs. These images do not need to be pre-aligned with the ion images obtained by nanoSIMS because the alignment can be done directly within the programme. This feature makes it possible to directly couple data acquired by nanoSIMS with those obtained by routine imaging techniques used for identification of cells or other objects in the sample, such as FISH, SEM, TEM or AFM.

Second, Look@NanoSIMS implements post-processing functions that allow compilation, graphical presentation and statistical analysis of results obtained by the analysis of multiple datasets. All these steps are done automatically within the programme, which means that the obtained results do not need to be exported in a tabulated form on a one-by-one basis, manually reformatting and reimported into a third-party software for further processing. This makes the programme particularly suitable for comprehensive high throughput analysis of nanoSIMS data.

Third, Look@NanoSIMS offers the possibility to calculate an image defined by an arbitrary expression. This function is necessary when calculating elemental ratios in samples that contain comparable abundances of different isotopes of the same element [e.g. the ratio $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ in highly ^{13}C -enriched cells].

Fourth, Look@NanoSIMS implements the so-called interactive thresholding method for the definition of ROIs. Although this feature is not unique to Look@NanoSIMS (e.g. it is provided also by L'Image), our programme is the only freely available software that implements it. ROI definition is essential and often the most time-consuming part of nanoSIMS data analysis. It is also the most subjective step, largely determined by the user's experience to

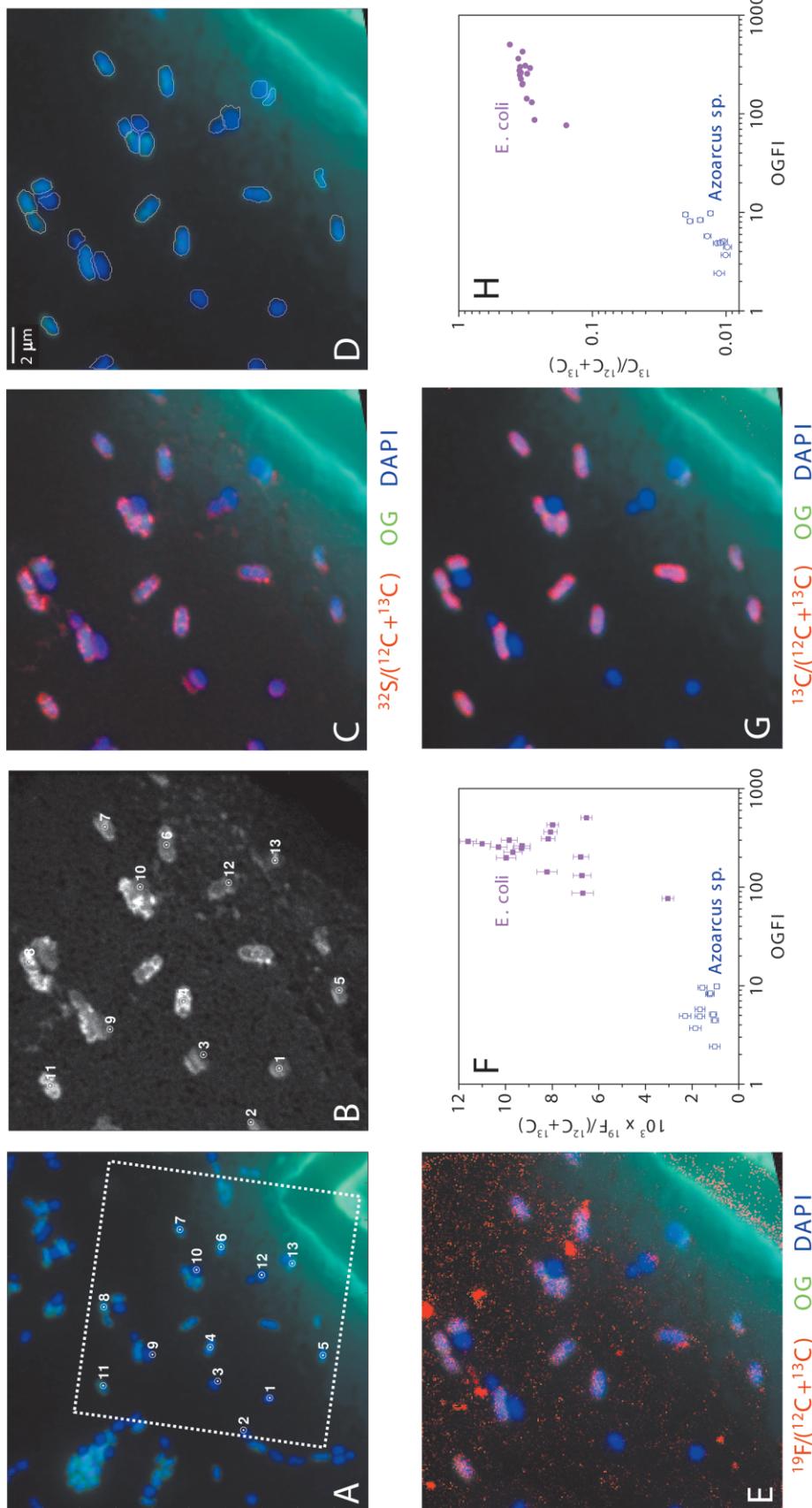


Fig. 4. The use of an external image as a template for ROI definition. A. The original FISH image, as detected by a fluorescence microscope. The blue signal due to DAPI fluorescence marks all cells, while the green signal due to Oregon Green (OG) fluorescence marks only the *E. coli* cells. Part of the mark produced by laser micro-dissection, which was used for localization of the FISH field of view in the nanoSIMS instrument, can be seen by the intense cyan colour at the bottom-right edge of the image. Dashed square shows approximate boundaries of the nanoSIMS field of view in the FISH image. B. Total $^{32}\text{S}^-$ ion counts detected by nanoSIMS. Numbers in panels (A) and (B) indicate corresponding reference points defined interactively by the user. C. Overlay of the aligned FISH and ^{32}S images. D. ROI outlines (white) defined based on the aligned FISH image. E. Overlay of the aligned FISH and $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ images, demonstrating consistency of cell identification based on the FISH and nanoSIMS data. F. $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ ratio versus the OGFI in ROIs shown in panel (D). G. Overlay of the aligned FISH and $^{19}\text{F}/(^{12}\text{C} + ^{13}\text{C})$ images, showing coupled information about cell identity and isotopic enrichment. H. ^{13}C -enrichment versus OGFI in ROIs shown in panel (D). In panels (C), (E) and (G), red channel shows the nanoSIMS data.

Table 3. Comparison of Look@NanoSIMS with other programmes used for nanoSIMS data analysis.

Feature/Function	Look@NanoSIMS	Open_MIMS	WinImage	L'Image
Display of scanned planes	+	+	+	+
Drift-corrected accumulation	+	+	+	+
ROI definition				
Manual	+	+	+	+
Semi-automated	+	+	-	+
Based on an external image	+	-	-	-
ROI classification	+	+	-	+
Quantification of elemental and isotopic compositions				
Images	+	+	+	+
Histograms	+	+	-	+
Depth profiles	+	+	+	+
Lateral profiles	+	+	+	+
Averages in ROIs	+	+	+	+
Scatter plots of averages in ROIs ^a	+/-	-/-	-/-	+/-
Dead-time and QSA correction	+	+	+	+
δ-notation	-	-	-	+
Arbitrary expressions	+	-	-	+
RGB composition	+	+	-	+
Image stitching	-	-	-	+
Statistical comparison of isotopic compositions				
In ROIs ^a	+/-	-/-	-/-	-/-
In ROI classes ^a	+/-	-/-	-/-	-/-
Open Source				
Platform	Matlab 2010b ^b (multiplatform)	ImageJ 1.43 ^b (multiplatform)	Aphelion (Windows XP ^b)	PV-WAVE (Windows XP ^b)
Availability	Free (MPI Bremen)	Free (NRIMS Harvard)	Commercial (Cameca)	Commercial (L.R. Nittler)

a. Single/multiple datasets.

b. Minimum version requirement.

Availability and absence of a feature is marked by '+' and '-' respectively. Although the L'Image programme was not tested in this study, its features are included based on the programme description available from http://www.dtm.ciw.edu/users/nittler/limage_manual.pdf.

recognize objects of interest and, subsequently, his/her ability to draw their outline reproducibly. Although the first step (object recognition) is a general challenge independent of the programme used, the second step (outline drawing) depends on the ROI definition methods offered by the programme. Typical methods for ROI definition include free-hand drawing and drawing of 'standard shapes' such as ellipses, rectangles or polygons. Clearly, when these methods are used, a few pixel differences can likely occur when ROIs for the same objects are defined by independent users or even by the same user but during independent trials. However, as shown by our sensitivity analysis, already such small differences may result in a significantly different isotopic or elemental composition of the object represented by the ROI. To avoid, or at least minimize the dependence of the data analysis on the user subjectivity, some programmes (e.g. Open_MIMS or L'Image) implement algorithms for automated object recognition. However, in our experience such algorithms are not able to satisfactorily recognize objects in images of complex samples, which leaves the ROI drawing 'by hand' as the only realistic option. On the other hand, for datasets where object recognition algorithms would be successful, it is likely that a quality check by an experienced user would also be required, lowering the overall speed of ROI definition. In our experience, interactive

thresholding, which combines reproducible drawing with an implicit quality check, is a good compromise for rapid and reliable definition of ROIs.

By design, Look@NanoSIMS is an open-source freeware programme. If required, anyone can modify it according to their specific needs. It is written in Matlab, allowing all powerful functions available in Matlab to be easily implemented in Look@NanoSIMS. Although Matlab is not free, it is widely used in universities and research institutions. Thus, many scientists and students are able to use Look@NanoSIMS at no additional cost.

Except for the dead-time and QSA corrections, Look@NanoSIMS presently does not allow correction for the instrument fractionation effects, introduced for example when samples are measured on different dates or with different instruments. Thus, when combining and comparing datasets by the metafile processing function, it is important that all data have been collected under the same instrumental conditions. The δ-notation, which is commonly used to describe isotopic enrichment of a sample relative to a standard, is also presently not implemented in Look@NanoSIMS. Instead, all elemental and isotopic ratios are calculated and exported as absolute values.

In conclusion, the newly developed programme Look@NanoSIMS facilitates rapid and comprehensive analysis

of a large number of nanoSIMS datasets. By implementing features ranging from robust plane accumulation to semi-automated interactive cell definition and classification to metafile processing and statistical analysis (Table 3), elemental and isotopic composition of hundreds of cells can be rapidly analysed and compared. The programme is available for free, enabling a growing community of researchers with access to a nanoSIMS facility to analyse their data at no additional cost. Furthermore, the programme is open-source, making it possible to customize the existing features and easily add new functions by others. Thus, in combination with the nanoSIMS measurements and other high-resolution analytical methods, the programme provides a valuable addition to the tools used for deciphering questions in environmental microbiology, microbial ecology and other research fields.

Experimental procedures

Sample preparation

Environmental samples from the Baltic Sea originated from surface waters of the Baltic Archipelago south of Stockholm ($N 58^{\circ}48' 28''$, $E 17^{\circ}37' 60''$). They were collected during a cyanobacterial bloom in August 2009 and incubated immediately after collection for 6 h in the light in seawater enriched with $H^{13}CO_3^-$ and $^{15}N_2$. Afterwards, the cells were chemically fixed with formaldehyde (1% final concentration) and filtered onto Au-Pd-coated polycarbonate filters (pore size 0.2 μm ; Millipore). Subsequently, the cells were hybridized with the horseradish peroxidase labelled oligonucleotide probe CF319a, which is specific for bacteria of the *Cytophaga-Flavobacteria-Bacteroidetes* clade (Manz *et al.*, 1996). The hybridization step was followed by the deposition of ^{19}F -containing tyramides, as previously described (HISH-SIMS; Musat *et al.*, 2008). Control samples were prepared identically following an incubation that was stopped by the addition of chemical fixatives in less than 5 min.

Artificial mixtures of *E. coli* (*Gammaproteobacteria*) and *Azoarcus* sp. (*Betaproteobacteria*) cells were prepared from cells grown in pure cultures. The *E. coli* culture medium contained ^{13}C -labelled glucose as the sole carbon source, whereas the *Azoarcus* sp. cells were grown on a minimal medium (previously described by Musat *et al.*, 2008) without labelled substrate. After culturing, cells were chemically fixed, mixed, and filtered onto Au-Pd-coated polycarbonate filters (0.2 μm pore size, Millipore). Subsequently, they were hybridized with the horseradish peroxidase labelled oligonucleotide probe specific for *Gammaproteobacteria* (Gam42a; Manz *et al.*, 1992). The hybridization step was followed by the deposition of fluorescently labelled (Oregon Green 488-X, Molecular Probes, Eugene, OR, USA) and ^{19}F -containing tyramides and DAPI staining, as previously described (Musat *et al.*, 2008).

Prior to nanoSIMS analysis, samples were visually inspected on a laser micro-dissection microscope (LMD, Leica, Germany). Fields of interest were marked with the laser, which allowed their fast and accurate localization

during the nanoSIMS analysis. After marking, the fluorescently labelled samples were imaged with an epifluorescence microscope (Axioskop 2 mot plus, Zeiss), using the standard filter sets for Oregon Green and DAPI imaging.

NanoSIMS measurements

Samples were analysed using the NanoSIMS 50 I instrument (Cameca). All analysed fields were first presputtered with Cs^+ ions to stabilize secondary ion intensity. Subsequently, the primary Cs^+ ion beam (current 1–2 pA, energy 16 keV) focused to a nominal spot size of 130 nm rastered the scanning area in 256 \times 256 pixels with dwelling time of 1 ms per pixel. Secondary ions were extracted from the sample surface and mass-separated in a magnetic sector of a mass detector. Secondary ions of desired mass-to-charge ratio (m/z) were counted simultaneously in separate electron multiplier detectors for each individual pixel. Mass resolving power above 6000 was used to separate interferences in the signal from molecular ions of close m/z ratio. Secondary ions and ion pairs detected for the environmental samples included $^{12}C^-$, $^{13}C^-$, $^{12}C^{14}N^-$, $^{12}C^{15}N^-$ and $^{19}F^-$, whereas $^{12}C^-$, $^{13}C^-$, $^{12}C^{14}N^-$, $^{19}F^-$ and $^{32}S^-$ were detected for the artificially prepared cell mixtures. Secondary electrons excited from the sample surface were also extracted and collected during the measurement, providing information about surface topography. Quality of data collected on different days was ensured by correct and constant tuning of the instrument. The need for QSA correction was avoided by keeping the secondary ion intensities sufficiently low.

Description of the Look@NanoSIMS programme

Look@NanoSIMS was developed in Matlab, and implements a graphical user interface (GUI) through which all functions are executed (Fig. S3). Here we briefly summarize these functions, and describe the structure and meaning of the results generated by the programme. A more detailed manual, as well as the programme itself, can be downloaded from the website of the Max-Planck Institute in Bremen or from <http://www.microsen-wiki.net/doku.php?id=LANS>.

Organization of the input and output data

Datasets belonging to a given experiment, including the raw ion image files (file extension **im**) and the information files (file extension **chk_im**), must be copied to a single directory. The processed information for a given dataset is stored in a directory of the same name as the input file name. This directory is further structured into sub-directories that contain exported information in the Matlab (**mat**), vector graphics (**eps** and **pdf**), bitmap (**tif**) and text (**dat**, **tex** and **tests**) formats (see Fig. S4).

Loading and drift-corrected accumulation of detected planes

Initial steps of the nanoSIMS data analysis are performed via the *Input* menu (Fig. S3B). First, the raw dataset is loaded

from disk and the detected planes are visually inspected. Subsequently, selected planes are drift-corrected and accumulated. This is controlled by settings specified in the *Accumulation options* box and in the *images* fields, which include the name of the base mass, the number of the base plane, the alignment region within the base mass, and the identification numbers of planes that should be included in the accumulation. For a given plane l , the optimum alignment relative to the base plane b is defined as such a translation T_{xy}^l by a whole number of pixels in the x and y directions that minimizes the sum

$$\sum_{i=1}^N [\mu_i^b - (T_{xy}^l \mu')_i]^2 \quad (1),$$

where N is the number of pixels in the alignment region, and μ' is the base mass image in plane l normalized such that its mean and variance over that region is 0 and 1 respectively. Subsequently, the translations T_{xy}^l are applied to all other masses and stored in the **xyalign.mat** file (Fig. S4B).

ROI definition

Definition of ROIs is done via the *ROIs* menu (Fig. S3C). First, the name of the image that will be used as a *ROI definition template* is chosen. This can be any of the accumulated ion images (e.g. 12C14N, 19F), a ratio image derived from the ion images [e.g. 13C/12C, 19F/(12C + 12C)], an RGB composite image constructed from the ion (*rgb1*) or ratio (*rgb2*) images, or an external image (*ext*). The external image can be any bitmap image stored in a common graphics format (e.g. TIF or PNG). Alignment between the external and ion images is done interactively in a dedicated GUI (Fig. S5).

Interactive ROI definition is done through the *Action* menu of a dedicated GUI (Fig. S6). ROI outlines can be defined as ellipses, rectangles, or freehand-drawn polygons. However, when the contrast between the background and the signal used for ROI recognition is high, interactive thresholding is the preferred method for ROI definition due to its speed and reproducibility. In this method, selection of a pixel on the image with a mouse-click leads to an automatic selection of a region that extends over pixels whose values do not fall below the value in the selected pixel multiplied by a threshold value. The default threshold value of 0.5 can be changed by pressing the 'up' or 'down' arrow key. Generally, the shape of the ROI outline defined by this method is a polygon. This can be, however, changed to an ellipse that circumscribes this polygon, whereby the axes of this ellipse can be changed by pressing the 'right' or 'left' arrow key. The current definition of a ROI can be cancelled by pressing *Esc* or confirmed by pressing *Enter*. Once confirmed, the ROI is defined as the pixels that are inside and on the edge of the drawn polygon (ellipse).

If the most recently defined ROI overlaps with one or more of the ROIs defined previously, the overlapping pixels will belong to the most recently defined ROI. Any previously defined ROI can be split by a line that is drawn through the ROI. Furthermore, previously defined ROIs can be removed, or combined into one. Any addition or removal of a ROI will update the ROI classification file (see section *ROI classification* below), if this file was defined and selected before ROI definition.

ROI definition is completed by saving the ROI file on the disk and exporting the ROI outlines in a graphical format. The latter step will automatically assign a unique identification number to each ROI. To facilitate rapid post-processing of different datasets in a later stage of the analysis (see section *Metafile processing* below), ROIs should be saved under the default file name **cells.mat** (Fig. S4B).

Text output

Export of ion counts and ion ratios in ROIs in a tabulated format is done through the *Output* menu (Fig. S3D). The results can be exported as depth profiles (files with extension **-z.dat**) or as values derived from total ion counts accumulated over all selected planes (files with extension **.dac**). In both cases the data are stored in the **dat** sub-directory (Fig. S4B).

When exporting depth profiles, ion counts in plane l and ROI k are calculated as

$$m_i^k = \sum_{i=1}^{P_k} M_{ii} \quad (2),$$

where M_{ii} is the ion count per dwelling time of mass M in pixel i and plane l , and P_k is the number of pixels in ROI k . Furthermore, the mean ratio of masses M and N in plane l and ROI k is calculated by Eq. 3, while Poisson percentage and standard errors of the mass ratios are calculated from Eqs 4 and 5 respectively. Poisson errors represent a theoretical precision of the mean, derived from the fact that the count of ions detected during dwelling time is a random variable with Poisson distribution (Cameca, WinImage User's Guide v 2.0.8, p. 82).

$$r_i^k = m_i^k / n_i^k \quad (3)$$

$$\delta r_i^k = \sqrt{1/m_i^k + 1/n_i^k} \times 100\% \quad (4)$$

$$PE(r_i^k) = r_i^k \delta r_i^k / 100\% \quad (5)$$

The depth profile output additionally contains the result of a test for a significant trend with depth. This is obtained by fitting the values of m_i^k and r_i^k plotted as a function of depth index i with a polynomial of order POL , and testing whether the variance of fit residuals is significantly lower than the variance in the original values. Maximal tested value of POL is 5. The significance testing is done using Levene's test (default significance level $\alpha = 0.05$). $POL > 0$ indicates significant trend with depth, whereas $POL = 0$ indicates no significant trend with depth.

When exporting values derived from total ion counts accumulated over all selected planes, the ion counts in ROI k are calculated as

$$m^k = \sum_{i=1}^L m_i^k \quad (6),$$

where m_i^k is calculated by Eq. 2 and L is the number of selected planes. Furthermore, the mean ratio of masses M and N in ROI k is calculated by Eq. 7, whereas the Poisson percentage and standard error is calculated from Eqs 8 and 9 respectively.

$$r^k = m^k / n^k \quad (7)$$

$$\delta r^k = \sqrt{1/m^k + 1/n^k} \times 100\% \quad (8)$$

$$PE(r^k) = r^k \delta r^k / 100\% \quad (9)$$

Additionally to these values, the exported **dac** file contains information about the ROI position, size and shape. The size is given both as the amount of pixels in the ROI and as the diameter of a circle (in μm) with the equivalent amount of pixels as the ROI. The shape is expressed as the ratio between the long and short axes of the ellipse circumscribing the ROI.

Graphical output as images, histograms, depth and lateral profiles, and scatter plots

Display and export of ion counts and ion ratios in a graphical format is done through the *Output* menu (Fig. S3D) and customized through the *Output options* box. Exported images and graphs are stored in the **eps** and **tif** sub-directories (Fig. S4B). Expressions for the ratio images, and the corresponding scales in which they should be displayed [given as (min max)], are defined in the provided text fields.

Images of total accumulated ion counts and their ratios are exported by selecting the *Display Images* option. They can be displayed in a linear or log-transformed scale, with ROI outlines included or not. Look@NanoSIMS additionally allows setting all values in pixels that do not belong to ROIs to zero. Although this may be useful for suppressing information that is biologically irrelevant or distracting the visual perception of the image (e.g. unspecific fluorine deposits that do not colocalize with biomass), this feature should not be misused to purposefully hide information that is not in line with the intended interpretation of the data.

Ion or ratio images can additionally be combined in a composite RGB image, which is done by selecting the *Combine images as RGB* option. Images that should be combined are specified by their identification numbers in the corresponding R, G and B fields (Fig. S3A). A specific colour is not used if the corresponding field is left empty. Images are combined either pixel by pixel or using average ratios in ROIs (Eq. 7). The contrast and saturation of each individual colour can be modified by changing the scale of the corresponding image.

Other types of graphical output include histograms, depth profiles, lateral profiles and scatter plots. Histograms are constructed from the values of ion counts or ion ratios in all pixels that belong to ROIs. If ROIs have been classified (see section *ROI classification* below), the histograms are constructed separately for each ROI class. Depth profiles are displayed using values calculated by Eqs 2 (ion counts) and 3 (ratios), and are plotted together with error bars (corresponding to the Poisson standard error; Eq. 5) and fitting polynomials (only when a significant trend with depth is detected, i.e. $\text{POL} > 0$; see section *Text output* above) in a dedicated GUI (Fig. S7). Lateral profiles are calculated using the values of accumulated ion counts and ratios in image pixels along a user-defined profile (Fig. S8). If the thickness

of the profile is set to more than 1 pixel, the value at each distance along the profile is displayed and exported as a mean \pm standard deviation of values in pixels perpendicular to the direction of the profile. Scatter plots of ratios are generated based on values calculated by Eq. 7, and are plotted together with error bars (corresponding to the Poisson standard error; Eq. 9). If ROIs have been classified, the different classes are displayed with different symbols and colours. The ratios that should be displayed in the scatter plot are specified by their identification numbers in the corresponding *x*, *y* and *z* fields (Fig. S3A).

ROI classification

ROI classification is done after ROI definition, by selecting *Classify ROIs* in the *ROIs* menu (Fig. S3C). ROI classes are defined by names that are one-character long (e.g. *a*, *b*, *c*, etc.), and stored in a text file with the recommended name **cells.dat** (Fig. S4B). When the number of ROIs is low, classification is best done *manually* (*ROI by ROI*), whereby the ROI identification numbers are added together with the corresponding class name in the provided GUI (Fig. S6B) and saved to disk. When the number of ROIs is large and the classification conditions can be formulated by a logical expression, ROI classification can be done *automatically*. In the provided GUI (Fig. S6C), classification variables (*x*, *y* or *z*) are first associated with the data, which is done by selecting the filename containing the data [e.g. a 19F-(12C + 13C).dac file containing the 19F/(12C + 13C) ratios] and specifying which data column should be considered as the corresponding classification variable (usually *mean ratio* or *ROI size*). Subsequently, classification conditions are specified in the provided fields using Matlab-like syntax (characters '&' and '|' represent logical AND and OR, respectively; Fig. S6C). The assignment of ROI identification numbers to classes is done in a sequential manner: all ROIs for which condition 1 is satisfied are assigned to the first class, all ROIs for which condition 2 is satisfied are assigned to the second class, irrespective whether some of them may have been already assigned to the first class, etc. ROIs that have not been assigned to a class by any of the specified conditions are by default assigned to class *otherwise*.

Statistical comparison of ROIs and ROI classes

Statistical comparison of ROIs for a currently analysed dataset is done by selecting the *Compare ROIs* option while calculating the ratios through the *Output* menu (Fig. S3), whereas comparison of ROIs from multiple datasets is done through the *Tools* menu during metafile processing (see section *Metafile processing* below). A dedicated GUI (see Fig. S9) offers statistical tests that can be performed for three levels: individual ROIs, ROI classes and treatments.

Comparison of ROI classes and treatments takes the mean ratios r^k (Eq. 7) as a basis for comparison. In contrast, comparison of individual ROIs is done based on the per-plane ratios r_i^k (Eq. 3). The latter is based on the assumption that r_i^k represent independent replicate measurements of the ratio in the object that the ROI represents. This assumption can be verified by plotting the ratio's depth profile (see

above). If r_i^k exhibit a significant trend with depth, it is likely that they originate from more than one type of object in the sample (e.g. a cell and the underlying filter or another cell with different isotopic composition) and are therefore not justified as a basis for ROI comparison.

For each comparison level, homogeneity of variance is tested using Levene's test (Fig. S9A), whereas significant differences between means are tested using either the Kruskal–Wallis or ANOVA test (Fig. S9B; default significance level $\alpha = 0.05$). All tests can be performed with both the original and log-transformed data to allow comparison of ROIs and ROI classes that exhibit small but also very large (e.g. an order of magnitude or more) differences in their isotopic or elemental composition. To identify which ROIs, ROI classes or treatments are significantly different from the others, the Kruskal–Wallis and ANOVA tests are coupled to a multiple comparison procedure based on Tukey tests. This analysis assigns to each tested ROI, ROI class or treatment a pair of ranks based on which it can be compared with the others. For example, assuming that ROI classes c_1 and c_2 were assigned pairs of ranks $[R_{\min}^{c_1} R_{\max}^{c_1}]$ and $[R_{\min}^{c_2} R_{\max}^{c_2}]$, respectively, their comparison is interpreted as follows: the inequality $R_{\max}^{c_1} < R_{\min}^{c_2}$ means that the mean ratio $\langle r^{c_1} \rangle$ for ROIs in class c_1 is significantly greater than the mean ratio $\langle r^{c_2} \rangle$ for ROIs in class c_2 . In contrast, the mean ratios $\langle r^{c_1} \rangle$ and $\langle r^{c_2} \rangle$ do not significantly differ within the instrumental precision if $R_{\max}^{c_1} \geq R_{\min}^{c_2}$.

The results of these statistical analyses are exported graphically or as text (Fig. S9). Both contain information about the ratio mean, 95% confidence interval of the mean, standard error of the mean and the pair of ranks $[R_{\min} R_{\max}]$.

Metafile processing

Metafile processing is done through the *Post-processing* menu (Fig. S3E). First, processing 'instructions' are defined in a dedicated GUI (Fig. S10A), including the names of the datasets that should be processed, the corresponding treatment identification numbers, ROI classes, and variables that should be processed. The variables can include ratios (e.g. 13C/12C or 12C15N/12C14N), or any other value stored in the text output files generated during the single dataset processing (e.g. ROI's size, x and y position). Subsequently, additional information is specified in the main GUI dedicated to metafile processing (Fig. S10B), including the base directory that contains the results generated for each single dataset, name of the metafile, and a generic name of the ROI (**cells.mat** by default) and ROI classification (**cells.dat** by default) file.

The actual metafile processing is done through the *Tools* menu (Fig. S10C), which allows display of variables specified in the metafile in 2D or 3D scatter plots, and comparison of ROIs, ROI classes and treatments. This plotting and comparison is done for all variables and all datasets specified in the metafile.

PDF output

A comprehensive summary of the results obtained during the analysis of a single nanoSIMS dataset is generated by select-

ing *Generate LaTeX + PDF output* from the *Output* menu of the main GUI (Fig. S3D), whereby the types of images and graphs included in the output depend on the specified *Output options* (Fig. S3A). Similarly, an overview of the results obtained by metafile processing is generated by selecting one of the *Export PDF* options in the *Tools* menu of the metafile processing GUI (Fig. S9C). Installation of LaTeX, a freely available text-processing programme (<http://www.latex-project.org>), is required to enable these features.

Acknowledgements

We are grateful to the Max-Planck Society for supporting the NanoSIMS facility, and to three anonymous reviewers for helpful comments on the manuscript.

References

- Behrens, S., Lösekann, T., Pett-Ridge, J., Weber, P.K., Ng, W.-O., Stevenson, B.S., et al. (2008) Linking microbial phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence *in situ* hybridization (EL-FISH) and Nano-SIMS. *Appl Environ Microbiol* **74**: 3143–3150.
- Boxer, S.G., Kraft, M.L., and Weber, P.K. (2009) Advances in imaging secondary ion mass spectrometry for biological samples. *Annu Rev Biophys* **38**: 53–74.
- Dekas, A.E., Poretsky, R.S., and Orphan, V.J. (2009) Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science* **16**: 422–426.
- Finzi-Hart, J.A., Pett-Ridge, J., Weber, K., Popa, R., Fallon, S.J., Gunderson, T., et al. (2009) Fixation and fate of C and N in the cyanobacterium Trichodesmium using nanometer-scale secondary ion mass spectrometry. *Proc Natl Acad Sci USA* **106**: 6345–6350.
- Herrmann, A.M., Ritz, K., Nunan, N., Clode, P.L., Pett-Ridge, J., Kilburn, M.R., et al. (2007) Nano-scale secondary ion mass spectrometry – a new analytical tool in biogeochemistry and soil ecology: a review article. *Soil Biol Biochem* **39**: 1835–1850.
- House, C.H., Orphan, V.J., Turk, K.A., Thomas, B., Pernthaler, A., Vrentas, J.M., and Joye, S.B. (2009) Extensive carbon isotopic heterogeneity among methane seep microbiota. *Environ Microbiol* **11**: 2207–2215.
- Lechene, C., Hillion, F., McMahon, G., Benson, D., Kleinfeld, A.M., Kampf, J.P., et al. (2006) High-resolution quantitative imaging of mammalian and bacterial cells using stable isotope mass spectrometry. *J Biol* **5**: 20.
- Lechene, C.P., Luyten, Y., McMahon, G., and Distel, D.L. (2007) Quantitative imaging of nitrogen fixation by individual bacteria within animal cells. *Science* **317**: 1563–1566.
- Li, T., Wu, T.-D., Mazeas, L., Toffin, L., Guerquin-Kern, J.-L., Leblon, G., and Bouchez, T. (2008) Simultaneous analysis of microbial identity and function using NanoSIMS. *Environ Microbiol* **10**: 580–588.
- Lozano-Perez, S., Kilburn, M.R., Yamada, T., Terachi, T., English, C.A., and Grovenor, C.R.M. (2008) High-resolution imaging of complex crack chemistry in reactor steels by NanoSIMS. *J Nucl Mater* **374**: 61–68.
- McKeegan, K.D., Aleon, J., Bradley, J., Brownlee, D., Busemann, H., Butterworth, A., et al. (2006) Isotopic

- compositions of cometary matter returned by Stardust. *Science* **314**: 1724–1728.
- Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K.-H. (1992) Phylogenetic oligonucleotide probes for the major subclasses of proteobacteria: problems and solutions. *Syst Appl Microbiol* **15**: 593–600.
- Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.H. (1996) Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* **142**: 1097–1106.
- Matzel, J.E.P., Ishii, H.A., Joswiak, D., Hutcheon, I.D., Bradley, J.P., Brownlee, D., et al. (2011) Constraints on the Formation Age of Cometary Material from the NASA Stardust Mission. *Science* **328**: 483–486.
- Messenger, S., Keller, L.P., Stadermann, F.J., Walker, R.M., and Zinner, E. (2003) Samples of stars beyond the solar system: silicate grains in interplanetary dust. *Science* **300**: 105–108.
- Musat, N., Halm, H., Winterholler, B., Hoppe, P., Peduzzi, S., Hillion, F., et al. (2008) A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. *Proc Natl Acad Sci USA* **105**: 17861–17866.
- Musat, N., Adam, B., and Kuypers, M.M.M. (2010) Nano-SIMS coupled with *in situ* hybridization for ecological research. In *Stable Isotope Probing in Microbial Molecular Ecology*. Murrell, J.C., and Whiteley, A.S. (eds). Washington, DC, USA: ASM press, pp. 295–303.
- Orphan, V.J., and House, C.H. (2009) Geobiological investigations using secondary ion mass spectrometry: microanalysis of extant and paleo-microbial processes. *Geobiology* **7**: 360–372.
- Orphan, V.J., House, C.H., Hinrichs, K.U., McKeegan, K.D., and DeLong, E.F. (2001) Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**: 484–487.
- Ploug, H., Musat, N., Adam, B., Moraru, C.L., Lavik, G., Vagner, T., et al. (2010) Carbon and nitrogen fluxes associated with the cyanobacterium *Aphanizomenon* sp. in the Baltic Sea. *ISME J* **4**: 1215–1223.
- Ploug, H., Adam, B., Musat, N., Kalvelage, T., Lavik, G., Wolf-Gladrow, D., and Kuypers, M.M.M. (2011) Carbon, nitrogen and O₂ fluxes associated with the cyanobacterium *Nodularia spumigena* in the Baltic Sea. *ISME J* **5**: 1549–1558.
- Popa, R., Weber, P.K., Pett-Ridge, J., Finzi, J.A., Fallon, S.J., Hutcheon, I.D., et al. (2007) Carbon and nitrogen fixation and metabolite exchange in and between individual cells of *Anabaena oscillarioides*. *ISME J* **1**: 354–360.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Images of secondary ions (¹²C⁻, ¹³C⁻, ¹⁹F⁻), ion pairs (¹²C¹⁴N⁻, ¹²C¹⁵N⁻) and secondary electrons (SE) detected for the environmental sample from the Baltic Sea. Gray-bar represents total ion counts accumulated over 35 planes (dwell-time per plane: 1 ms).

Fig. S2. A–C. Depth profiles of ratios ¹³C/¹²C, ¹⁵N/¹⁴N and ¹⁹F/(¹²C + ¹³C) in selected ROIs, derived from ion counts

accumulated over ROI pixels in each plane. Symbols represent mean ratio values, error bars correspond to Poisson errors (see Eqs 3 and 5 in *Experimental procedures*).

Fig. S3. A. Main graphical user interface (GUI) of the Look@NanoSIMS programme.

B. *Input* menu is used to define dead-time and QSA correction settings, load nanoSIMS datasets from disk, and to visually inspect, align and accumulate plane images. Alignment and plane accumulation are controlled by the parameters in the *Accumulation options* box and by *images* in the *Detected masses* box.

C. *ROIs* menu is used to interactively define ROIs, to classify ROIs, to export ROI outlines in a graphical output, and to interactively align an external image with nanoSIMS images if it was chosen as template for ROI definition. The ROI definition is controlled by the parameters in the *ROIs definition options* box.

D. *Output* menu is used to process and export data for the currently processed dataset. The type of data processing performed is controlled by the parameters in the *Output options* box, allowing display and export of detected masses and calculated mass ratios as values accumulated over ROIs, images, histograms, depth profiles, lateral profiles, composite RGB images and scatter plots. This menu also allows statistical comparison of ROIs and ROI classes.

E. *Post-processing* menu is used to process results obtained by the analysis of multiple datasets.

Fig. S4. Organization of the data generated by the Look@NanoSIMS programme. Datasets that belong to a specific experiment and are intended for later synthesis by metafile processing must be copied in the same base directory (panel A). For each nanoSIMS dataset (files with extension **im** and **chk_im**), a directory with the same base name as the dataset is automatically created during the analysis. This directory contains all information generated during the data analysis (panel B). File **prefs.mat** contains settings of the analysis, and must be saved manually by selecting *Store Preferences* in the *Output* menu after the analysis of each dataset.

Fig. S5. A. Graphical user interface for interactive alignment of the external and nanoSIMS ion images. Images are loaded through the *File* menu (B). Reference points that correspond to each other are defined interactively in parallel for both images through the *Action* menu (C). At least five reference points are usually needed to account for unequal magnification of, or correct for distortion between the two images to achieve satisfactory alignment. After visual verification of the alignment (D), the aligned external image as well as the reference points are saved on disk through the *Action* and *File* menu respectively.

Fig. S6. Graphical user interface for (A) interactive ROI definition (B) manual and (C) automatic ROI classification. ROI definition is controlled through the *Action* menu, allowing interactive addition and removal of ROIs. When ROIs are well separated, *interactive thresholding* is the preferred method for ROI definition due to its speed and reproducibility. When a composite RGB image is used as template for ROI definition, the active channel used in the interactive thresholding mode is selected through the *Select interactive channel* menu. The currently defined ROIs and their identification numbers (right panel in A) are displayed by selecting *Display ROIs* from the menu. Manual ROI classification (B) is done by manually

assigning a class name to each ROI identification number. The class name should consist of one alphabetical character. Automatic ROI classification (C) is based on logical conditions that can be specified for up to three variables. The variable values are taken from the files and specific columns defined in the *Data* fields, the conditions are specified in the *Condition* fields using Matlab syntax.

Fig. S7. Graphical user interface for interactive display of depth profiles of (A) masses and (B) ratios in defined ROIs. Symbols in (A) correspond to total ion counts accumulated over ROI pixels in a given plane per dwelling time, and are normalized to the amount of pixels in the ROI. Symbols in (B) correspond to ratios of total ion counts accumulated over ROI pixels in a given plane per dwelling time. Length of the error bars represents the corresponding Poisson errors (see Eqs 3 and 5 in *Experimental procedures*). The type of mass and mass ratio displayed is specified through the pull-down menu, ROIs for which depth profiles should be displayed are selected in the listbox. Due to limitations of the graphical user interface, maximum of 50 ROIs can be selected at once.

Fig. S8. A. Graphical user interface for interactive display of lateral profiles of masses and ratios. The profile is defined by clicking on the *Define lateral profile* button and specifying the profile's beginning and end-points. The profiles are displayed separately for each selected mass or ratio (B) or together in one graph for all selected masses or ratios (C). If the thickness of the profile is set to more than 1 pixel, the value at each distance along the profile is displayed and exported as a mean \pm standard deviation of values in pixels perpendicular to the direction of the profile.

Fig. S9. Graphical user interface for statistical comparison of (A) individual ROIs (B) ROI classes and (C) treatments. The level of comparison is selected through the *Compare* pull-down menu, whereas the type of statistical test is selected

through the *Test type* pull-down menu. When comparing individual ROIs or ROI classes, select one treatment and one or more ROI classes. In contrast, when comparing treatments, select one ROI class and multiple treatments. Data can optionally be log-transform before the comparison. Results of the statistical test are exported in a graphical (A–C) and text-based (D) format by clicking on the *Export results* button. In the graphical output, square symbols represent mean values, crosses represent boundaries of the standard error of the mean, and error bars represent 95% confidence intervals of the mean. Ranking of the mean values based on a multiple comparison procedure is graphically represented by differently coloured patches, which show boundaries of groups with significantly different means ($P < 0.05$). For example, with respect to the $^{13}\text{C}/^{12}\text{C}$ ratio (panel B), classes 'b' and 'd' do not significantly differ from each other because their symbols and error bars lie within one patch (rank 2), but do significantly differ from classes 'k' and 'g' (rank 3) as well as classes 'c' (rank 1) and 'a' (rank 4) because their patches do not overlap. Class 'f' is significantly not different from classes 'c', 'b' and 'd' because its error bar (95% confidence interval of the mean) overlaps with patches ranked as 1 and 2.

Fig. S10. Graphical user interface for (A) definition of metafile instructions and (B) metafile processing. The type of analysis is selected through the *Tools* menu (C), parameters of the graphical output are specified in the *Graph/Image options* box. Metafile processing allows rapid visualization and comparison of results from multiple nanoSIMS datasets, as demonstrated in Fig. 3 and Table 2 in the main text.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.