Supplementary Information

First observation of unicellular organisms concentrating arsenic in ACC intracellular inclusions in lake waters

Agathe Martignier[1], Montserrat Filella[2*], Jean-Michel Jaquet[1], Mathieu Coster[3] and Daniel Ariztegui[1]

[1]*Department of Earth Sciences, University of Geneva, Geneva, Switzerland*

[2]*Department F.-A. Forel, University of Geneva, Geneva, Switzerland*

[3]*Service de l'écologie de l'Eau, Geneva, Switzerland*

Supplementary Information index

SI_excel_1.xlsx. Databases and statistical results for As-bearing micropearls from Lakes Geneva and Titicaca.

SI_excel_2.xlsx. Example of computation of summary statistics with a regression equation on probability scale (ROS technique).

Table S1

Sampling campaigns at location GE3 (46.2994 °N / 6.2197 °E) in Lake Geneva (Switzerland/France). Sampling depths generally correspond to the depth of Chl *a* maximum.

Year	Month	Sampling date	Depth / m
2013	June	21.06.2013	15
	August	06.08.2013	27
	September	02.09.2013	23
2014	January	28.01.2014	2
	November	04.11.2014	20
2015	March	24.03.2015	34
	May	19.05.2015	18
2016	March	30.03.2016	25
	April	21.04.2016	10
	June	14.06.2016	10
	July	12.07.2016	18
2017	February	21.02.2017	15
	March	21.03.2017	6
	April	24.04.2017	10
	May	16.05.2017	20
	June	13.06.2017	10
	July	11.07.2017	16
	August	10.08.2017	10
	October	31.10.2017	8
	December	05.12.2017	2.5

Table S2

Proportion of non-detects in descriptors and morpho-chemical groups for: (a) *lab* and (b) *left-censored* datasets. In yellow and green, descriptors retained for analysis.

a

MC	Number	Lab	Na	Mg	Al	Si	P	S	Cl	K	Ca	As	Sr	Ba
1	16	**Proportion of zero values**	0.69	0.25	0.38	0.81	1.00	1.00	0.69	0.81	0.00	0.00	0.13	0.00
2	50		0.50	0.00	0.66	0.72	1.00	1.00	0.54	0.82	0.00	0.00	0.30	0.00
3	70	Conservation criterion: proportion < 0.6	0.67	0.43	0.77	0.93	0.67	0.86	0.60	0.23	0.00	0.00	0.03	0.99
4	34		0.09	0.32	0.97	0.97	0.94	1.00	0.06	0.24	0.00	0.00	0.00	1.00
5	97		0.68	0.29	0.75	0.47	0.99	1.00	0.74	0.93	0.00	0.00	0.82	0.99
Total	267		0.57	0.27	0.75	0.72	0.90	0.96	0.58	0.63	0.00	0.00	0.37	0.75

b

MC	Number	Left-censored	Na	Mg	Al	Si	P	S	Cl	K	Ca	As	Sr	Ba
1	16	**Proportion of values**	0.94	0.81	0.69	0.88	1.00	1.00	0.88	0.94	0.00	0.00	0.25	0.00
2	50	**below the 1.0 mol% detection level**	0.72	0.00	0.70	0.72	1.00	1.00	0.54	0.98	0.00	0.00	0.42	0.00
3	70		0.80	0.83	0.86	0.93	0.71	0.87	0.79	0.31	0.00	0.00	0.07	1.00
4	34	Conservation criterion: proportion < 0.7	0.09	0.65	0.97	0.97	0.94	1.00	0.06	0.41	0.00	0.00	0.00	1.00
5	97		0.72	0.29	0.76	0.59	0.99	1.00	0.77	1.00	0.00	0.00	0.98	1.00
Total	267													

Table S3

Summary statistics differences between *lab* and *left-censored* data (Values of Table 1 minus values of Table 2). Pink: descriptor present in *lab* and eliminated in *left-censored* data. Unreliable values in MC2/Cl due to poor ROS fit (see SI_excel_2.xlsx).

Morpho Gr	Ns	Nv	Stat	Na	Mg	Al	Si	Cl	K	Ca	As	Sr	Ba
MC1	16	5	Mean		0.6	0.1				0.9	0.0	-0.1	0.7
Big Ba			SD										
			Median		0.3	0.3				1.2	0.0	0.1	1.0
MC2	50	6	Mean	0.7	0.2			-1.1		0.3	0.1	-0.2	1.2
Small Ba			SD										
			Median	0.1	0.0			-3.0		1.0	0.2	0.0	0.9
MC3	70	4	Mean		0.5				-0.1	3.6	0.1	0.1	
Sr-Tetraselmis			SD										
			Median		0.3				0.0	2.5	0.1	0.2	
MC4	34	7	Mean	0.0	-0.4			0.0	-0.1	0.3	0.1	0.1	
Sr-Titicaca			SD										
			Median	0.0	-0.8			0.0	0.0	0.0	0.0	0.0	
MC5	97	4	Mean		-0.6		-0.2			2.3	0.2		
MgAs			SD				0.0						
			Median		0.2		-0.5			2.4	0.1		

Table S4

Discrepancies between tables of difference/similarity (Table 3A and 3B) for *lab* and ROS datasets. L pink: the groups differ with respect to given descriptor in *lab* dataset, but not in ROS. R pink: the groups differ with respect to given descriptor in ROS dataset, but not in *lab*. L green: the groups do not differ with respect to given descriptor in *lab* dataset. R green: the groups do not differ with respect to given descriptor in ROS dataset.

Na

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4		D			
MC5					

Mg

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2	D				
MC3	S	D			
MC4	S		S		
MC5	D		D		

Al

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

Si

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

Cl

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4		D			
MC5					

K

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4			S		
MC5					

Ca

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

As

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

Sr

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

Ba

	MC1	MC2	MC3	MC4	MC5
MC1					
MC2					
MC3					
MC4					
MC5					

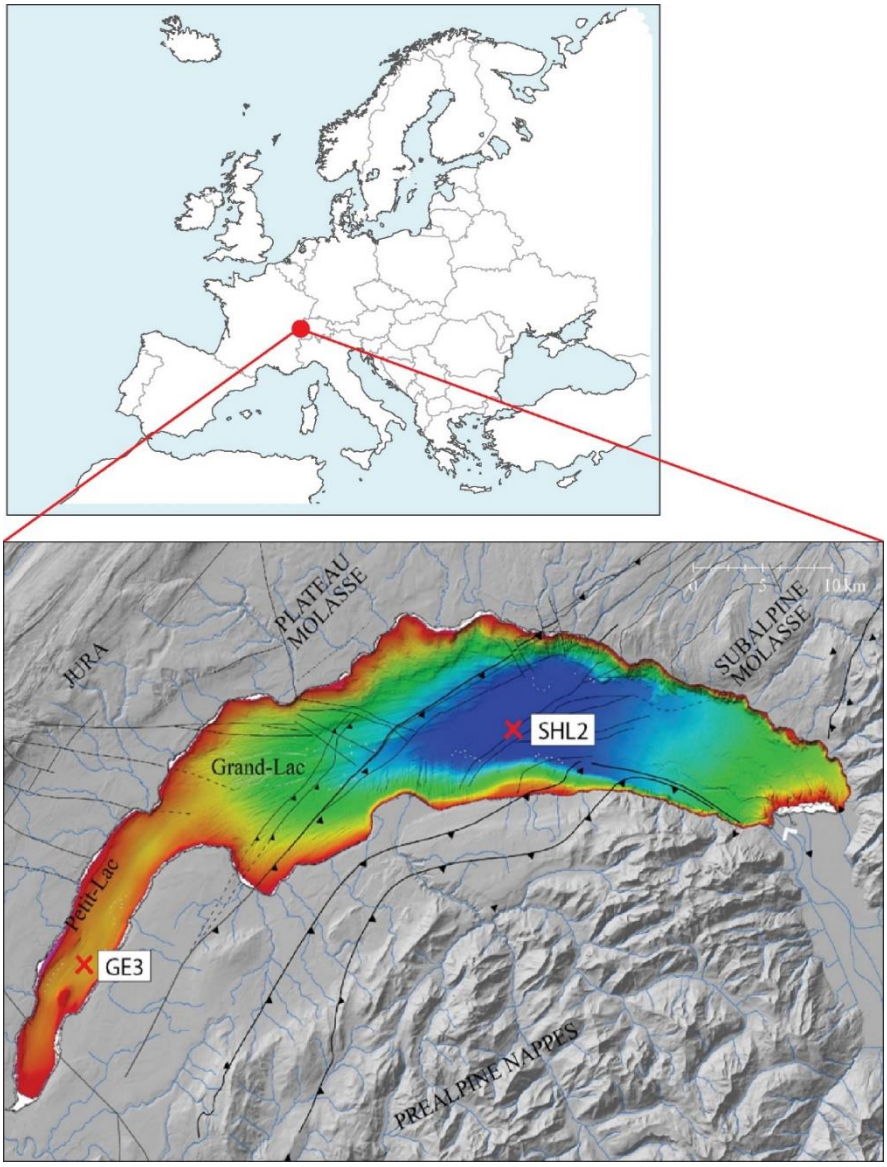
Table S5

Tests on descriptors distribution type per morpho-chemical group (*lab* data). Normality tested by Shapiro-Wilks W criterion ($p \geq 0.05$). Blue: normal; green: lognormal. Bimodality (gold) evaluated by visual inspection.

Morpho group	Number of descriptors	Distribution	Na	Mg	Al	Si	Cl	K	Ca	As	Sr	Ba
MC1	3	Normal										
		Lognormal										
		Bimodal										
MC2	6	Normal										
		Lognormal										
		Bimodal										
MC3	5	Normal										
		Lognormal										
		Bimodal									?	
MC4	7	Normal										
		Lognormal										
		Bimodal					?					
MC5	4	Normal										
		Lognormal										
		Bimodal										

Majority:	Normal							X			X
	Lognormal	X	X	X	X	X	X	X	X	X	
	Bimodal	X					X			X	

FIGURE S1. Location of the two sampling points in Lake Geneva (Switzerland/France): GE3 (46.2994 °N / 6.2197 °E) and SHL2 (46.452 ° N / 6.589 °E). Bathymetry drawn by University of Geneva and University of Bern (data from Canton de Vaud). Topographic layout from Kremer et al., 2014[a] (data from Swisstopo).



[a]Kremer, K., Marillier, F., Hilbe, M., Simpson, G., Dupuy, D., Yrro, B.J.F., Rachoud-Schneider, A.-M., Corboud, P., Bellwald, B., Wildi, W., and Girardclos, S. 2014. Lake dwellers occupation gap in Lake Geneva (France–Switzerland) possibly explained by an earthquake–mass movement–tsunami event during Early Bronze Age. *Earth and Planetary Science Letters* 385:28-39.

FIGURE S2. Location of the sampling point in Lake Titicaca (Bolivia/Peru). Coordinates: 16.19991 °S / 68.78534 °W. Source: Stephane Guédron.

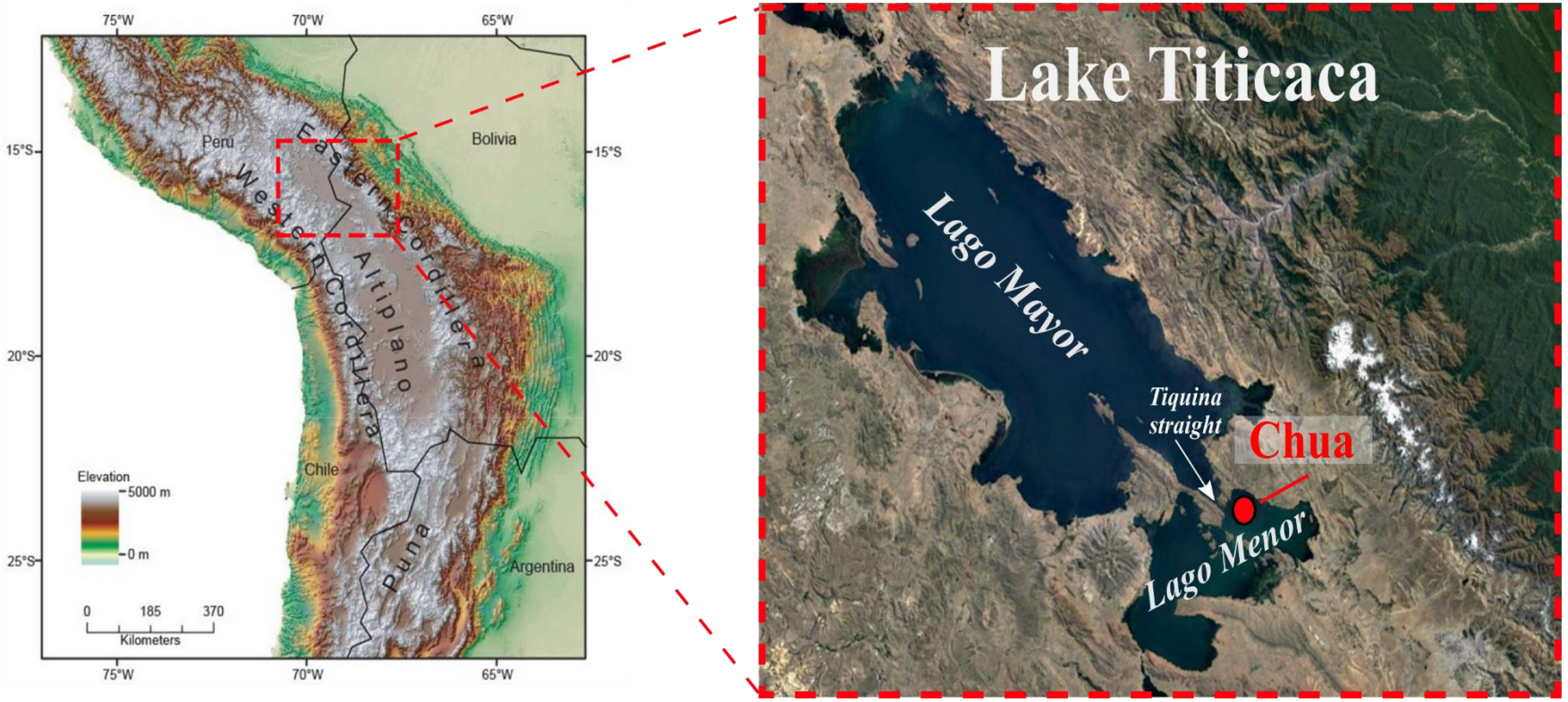


FIGURE S3. Image of three different types of micropearls present in Lake Geneva. Backscattered SEM image of a dried filter. Water sample taken in Lake Geneva (Switzerland) at SHL2 location, on 11 May 2015, at 10 m depth. The three different micropearl types are indicated by red arrows. The micropearls highly enriched in barium (small Ba micropearls, MC2) appear in white due to their higher atomic mass. The other micropearl types (Sr (MC3) and MgAs (MC5) micropearls) appear light grey. Note the globular cell of the *Tetraselmis cordiformis* algae (Sr-*Tetraselmis* micropearls). The small black dots in the background are the pores of the filter (0.2 µm). Most of the other items present on the image are frustules of diatoms.

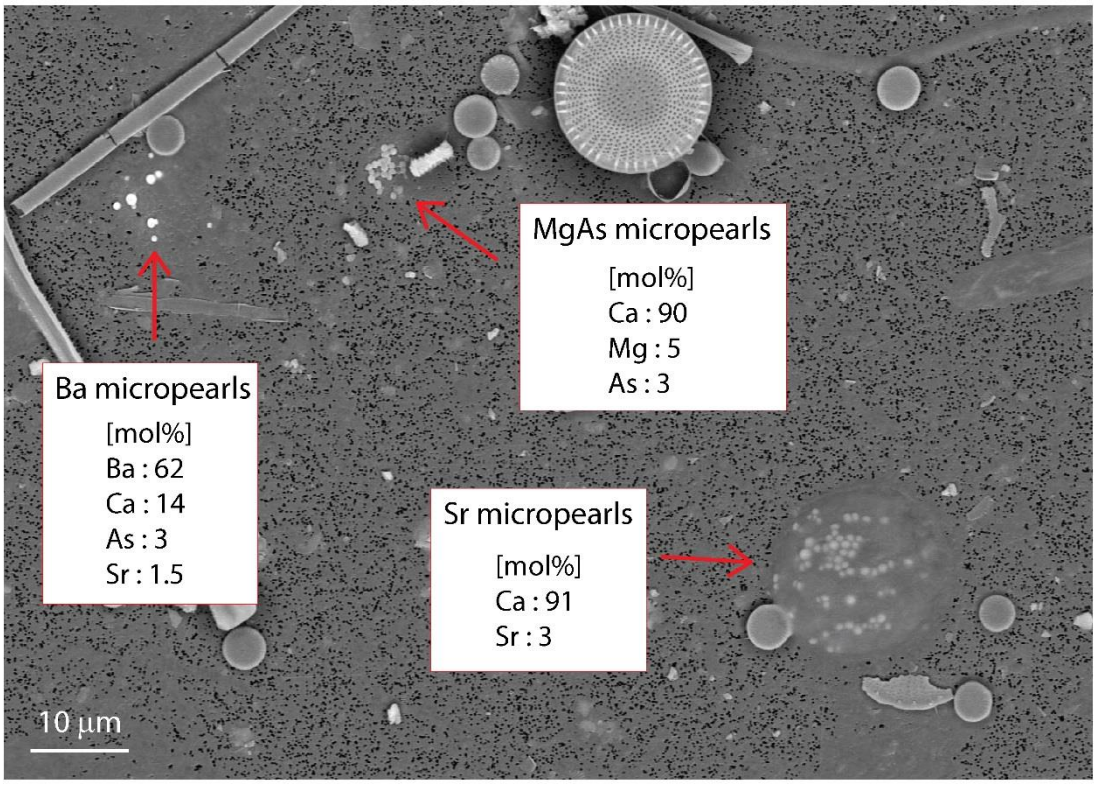
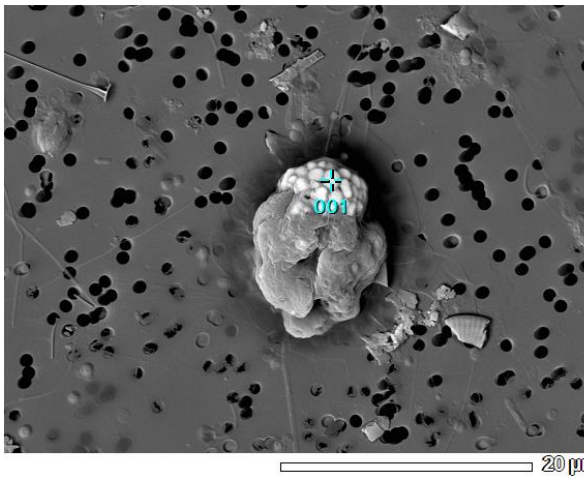
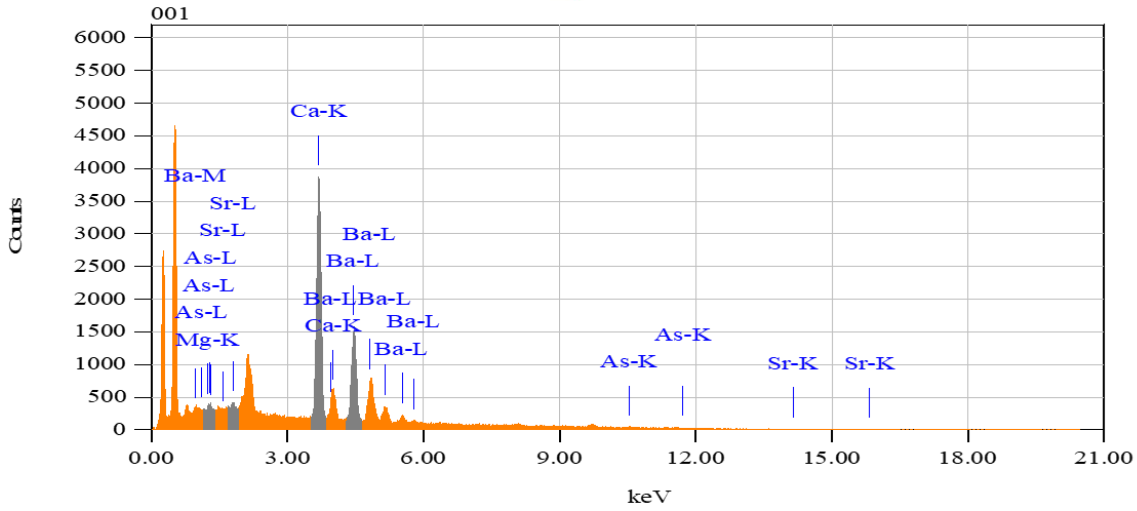


FIGURE S4. Raw EDXS data for samples in Figure 2 showing the whole energy range measured.

Sample a1



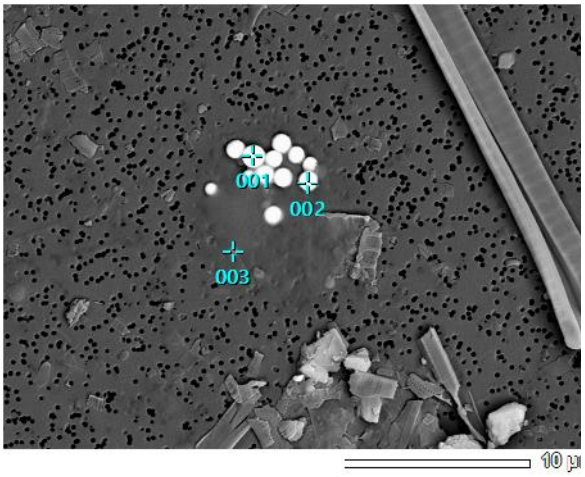

```
Title          : IMG1
---------------------------
-
Instrument     : 7001F
Volt           : 15.00 kV
Mag.           : x 2,500
Date           : 2018/07/13
Pixel          : 1024 x 768
```



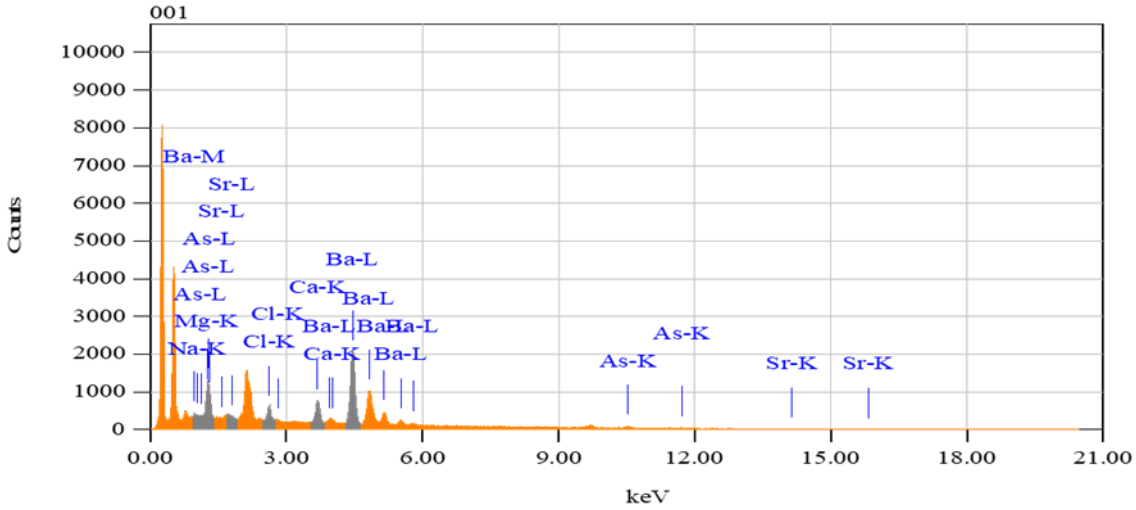
ZAF Method Standardless Quantitative Analysis
Fitting Coefficient : 0.4704

Element	(keV)	Mass%	Counts	Sigma	Atom%	Compound	Mass%	Cation	K
Mg K*	1.253	0.18	227.84	0.05	0.54				0.0958
Ca K	3.690	35.73	41395.16	0.22	64.53				44.1562
As L*	1.282	1.30	713.75	0.11	1.26				0.9487
Sr L*	1.806	1.89	1196.73	0.14	1.56				1.4829
Ba L	4.464	60.90	15288.63	0.46	32.10				53.3164
Total		100.00			100.00				

Sample a2



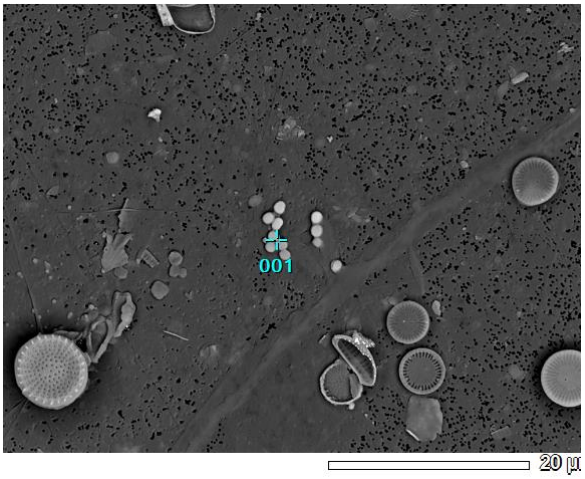

```
Title          : IMG1
--------------------------
-
Instrument     : 7001F
Volt           : 15,00 kV
Mag.           : x 3,700
Date           : 2016/09/16
Pixel          : 1024 x 768
```



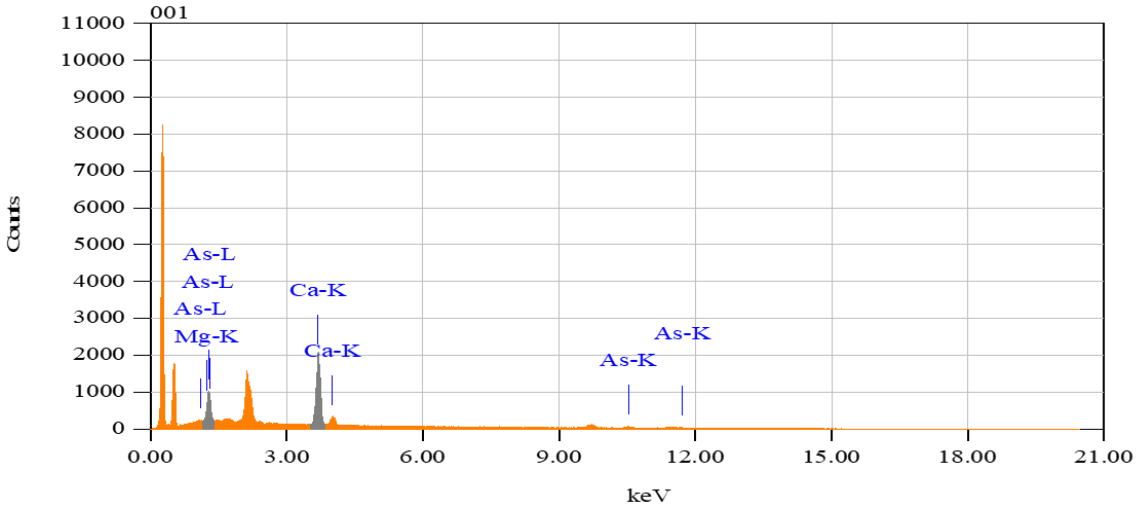
ZAF Method Standardless Quantitative Analysis
Fitting Coefficient : 0.6804

Element	(keV)	Mass%	Sigma	Atom%	Compound	Mass%	Cation	K
Na K*								
Mg K	1.253	2.32	0.07	9.25				1.2963
Cl K	2.621	2.85	0.04	7.80				3.1145
Ca K	3.690	5.83	0.07	14.11				7.8492
As L*	1.282	10.03	0.16	12.98				7.7927
Sr L*	1.806	0.29	0.14	0.32				0.2252
Ba L	4.464	78.68	0.27	55.55				79.7220
Total		100.00		100.00				

Sample c




```
Title          : IMG1
--------------------------
-
Instrument     : 7001F
Volt           : 15,00 kV
Mag.           : x 2,000
Date           : 2016/03/16
Pixel          : 1024 x 768
```



ZAF Method Standardless Quantitative Analysis
Fitting Coefficient : 0.8279

Element	(keV)	Mass%	Sigma	Atom%	Compound	Mass%	Cation	K
Mg K	1.253	3.39	115.28	6.13				2.2626
Ca K	3.690	72.83	679.07	79.91				75.9640
As L	1.282	23.78	396.38	13.96				21.7734
Total		100.00		100.00				

FIGURE S5. Comparison of the composition of *Sr-Tetraselmis* (MC3) in Lake Geneva and *Sr-Titicaca* (MC4). Asterisk: statistically significant difference (p = 0.95).

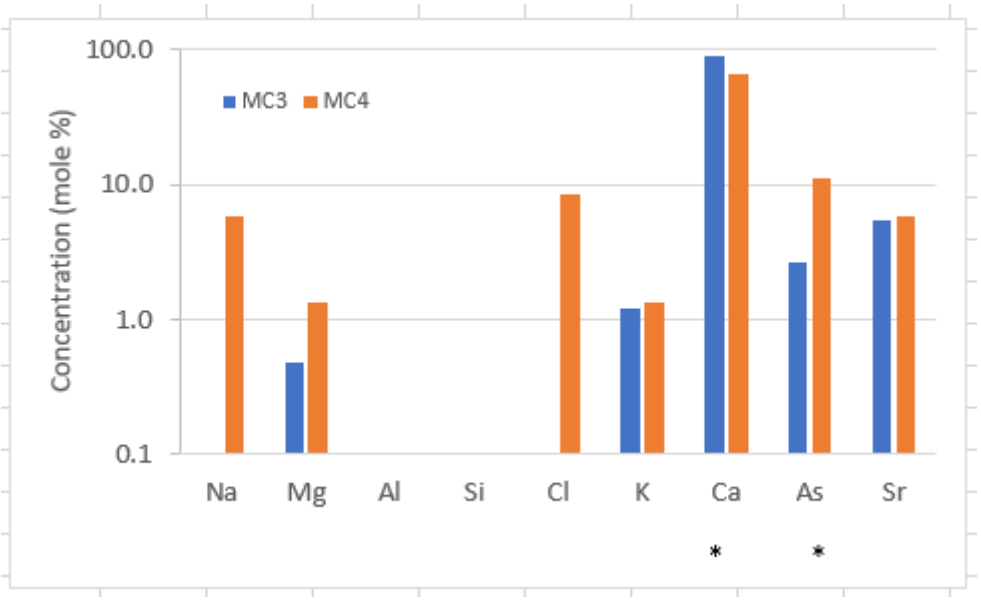


FIGURE S6. Relationship between the size of micropearls and their arsenic concentration. SEM secondary images of small areas of dried filters where MgAs micropearls expelled from their cell during filtration can be observed. Water sample taken from location GE3 in Lake Geneva, on 19 May 2015, at 18 m depth. (a) and (b): zoomed-in images of micropearls. Some of them appear partially dissolved. (c): general view of the MgAs micropearl cluster. Areas where the pores are not seen are patches of organic matter, remnants of the cell which produced the micropearls. SEM-EDXS measurements show that larger micropearls show higher As concentrations than small ones. The small black dots scattered through the image are the pores of the filter (0.2 µm). Thin scale bar (a and b): 500 nm. Thick scale bar (c): 1 µm.