Arce-Rodríguez *et al.*
Environmental Microbiome (2023) 18:6
https://doi.org/10.1186/s40793-023-00464-2

Environmental Microbiome

RESEARCH

Open Access

Out of the blue: the independent activity of sulfur-oxidizers and diatoms mediate the sudden color shift of a tropical river

Alejandro Arce-Rodríguez[1,2], Eduardo Libby[3], Erick Castellón[3,4], Roberto Avendaño[5], Juan Carlos Cambronero[6], Maribel Vargas[7], Dietmar H. Pieper[2], Stefan Bertilsson[8], Max Chavarría[3,5,6] and Fernando Puente-Sánchez[8*]

Abstract

Background Río Celeste ("Sky-Blue River") is a river located in the Tenorio National Park (Costa Rica) that has become an important hotspot for eco-tourism due to its striking sky-blue color. A previous study indicated that this color is not caused by dissolved chemical species, but by formation of light-scattering aluminosilicate particles at the mixing point of two colorless streams, the acidic Quebrada Agria and the neutral Río Buenavista.

Results We now present microbiological information on Río Celeste and its two tributaries, as well as a more detailed characterization of the particles that occur at the mixing point. Our results overturn the previous belief that the light scattering particles are formed by the aggregation of smaller particles coming from Río Buenavista, and rather point to chemical formation of hydroxyaluminosilicate colloids when Quebrada Agria is partially neutralized by Río Buenavista, which also contributes silica to the reaction. The process is mediated by the activities of different microorganisms in both streams. In Quebrada Agria, sulfur-oxidizing bacteria generate an acidic environment, which in turn cause dissolution and mobilization of aluminum and other metals. In Río Buenavista, the growth of diatoms transforms dissolved silicon into colloidal biogenic forms which may facilitate particle precipitation.

Conclusions We show how the sky-blue color of Río Celeste arises from the tight interaction between chemical and biological processes, in what constitutes a textbook example of emergent behavior in environmental microbiology.

Keywords Geobiology, Hydroxyaluminosilicates, Hydrothermal, Sulfur oxidizing bacteria, Diatoms, Río Celeste

*Correspondence:
Fernando Puente-Sánchez
fernando.puente.sanchez@slu.se
[1] Institute of Microbiology, Technical University of Braunschweig, 38106 Brunswick, Germany
[2] Microbial Interactions and Processes Research Group, Helmholtz Centre for Infection Research, 38124 Brunswick, Germany
[3] Escuela de Química, Universidad de Costa Rica, San José 11501-2060, Costa Rica
[4] Centro de Investigación en Ciencia e Ingeniería de Materiales (CICIMA), Universidad de Costa Rica, San José 11501-2060, Costa Rica
[5] Centro Nacional de Innovaciones Biotecnológicas (CENIBiot), CeNAT-CONARE, San José 1174-1200, Costa Rica
[6] Centro de Investigaciones en Productos Naturales (CIPRONA), Universidad de Costa Rica, San José 11501-2060, Costa Rica
[7] Centro de Investigaciones en Estructuras Microscópicas (CIEMic), Universidad de Costa Rica, San José 11501-2060, Costa Rica
[8] Deparment of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Lennart Hjelms Väg 9, 756 51 Uppsala, Sweden

Introduction

Costa Rica is located in the Pacific rim of fire and therefore has a number of active volcanoes, and multiple manifestations of hydrothermal and volcanic origin such as thermal springs, acidic rivers or mineral-rich streams [1, 3]. One of the most striking manifestations of that activity, that can be observed in the complex basaltic-andesitic volcanic massif of Tenorio (Guanacaste Volcanic Mountain Range), is Río Celeste ("Sky-Blue River"). Located within the Tenorio Volcano National Park, this river is recognized as one of the most beautiful rivers in the world for its characteristic sky-blue color that contrasts with the dark surrounding rainforest (Fig. 1). Because of this, the river and the landscapes of the National Park

have become an important hotspot for eco-tourism in Costa Rica [4] attracting more than 100,000 visitors per year [2, 5].

What makes Río Celeste unique is that it is possible to observe the specific site where the river turns blue. At this point known locally as 'Teñidero' (dyeing point), the two colorless streams Río Buenavista (Buenavista River) and Quebrada Agria (Sour Creek) merge, and the resulting waters quickly acquire the sky-blue color (Fig. 1). In a previous study, [6] showed that the sky-blue coloration of Río Celeste is caused by Mie light-scattering caused by particles suspended in its water stream. This phenomenon has been scarcely studied [7–9] and remains poorly understood. Because our previous measurements indicated that the particles present in Río Buenavista were smaller than the particles present in Río Celeste after the 'Teñidero' confluence [6], we initially proposed that the light-scattering particles originated via a pH-dependent aggregation mechanism when Río Buenavista (pH = 6.8) becomes acidified by Quebrada Agria (pH = 3.1), marking the beginning of Río Celeste (pH = 5.5). At that point we did not have measurements on the composition of the suspended particles, and we therefore alternatively examined the white material conspicuously deposited on the rocks which showed properties consistent with the presence of a poorly crystalline aluminosilicate mineral.

While our previous work demonstrated the role of light-dispersing aluminosilicates in giving Río Celeste its characteristic color, the formation mechanism of such particles was however not elucidated. Furthermore, our original study did not take into consideration the potential role of microorganisms in generating the chemical conditions conducive to this process. In the present study we provide a detailed analysis of the colloidal-sized particles that occur at the 'Teñidero' dyeing point, and examine the microbial communities that inhabit Río Buenavista, Quebrada Agria and the combined microbial community in Río Celeste. Our results challenge the previous perception in favor of a model where the particles causing Mie scattering mainly originate from dissolved aluminum and silicic acid coming from Quebrada Agria, with Río Buenavista acting as a secondary source of silicic acid and silica. Microbial community analyses show that Quebrada Agria hosts sulfur-oxidizing bacteria that are likely responsible for the acidic conditions in that source water, indirectly promoting aluminum and silicic acid mobilization. In contrast, Río Buenavista is a source of diatoms to Río Celeste whose presence influences silicon speciation and partitioning. Both the presence of sulfur-oxidizing bacteria in Quebrada Agria and the presence of diatoms in Río Buenavista are consistent with the chemical conditions of the respective stream. Our results shed new light on the combined biological and chemical processes that interactively cause formation of aluminosilicate particles that produce the beautiful sky-blue color that Río Celeste is well known for.

Materials and methods

Site description

The 'Teñidero' confluence (10°42′02.5″ N, 84°59′49.3″ W) is located in the Costa Rican Tenorio Volcano National Park, at the confluence of two streams: the neutral Río Buenavista and the acidic Quebrada Agria (Sour Creek). The whole area is embedded in the Volcanic Mountain Range of Guanacaste, and the region experiences frequent hydrothermal activity which gives rise to a

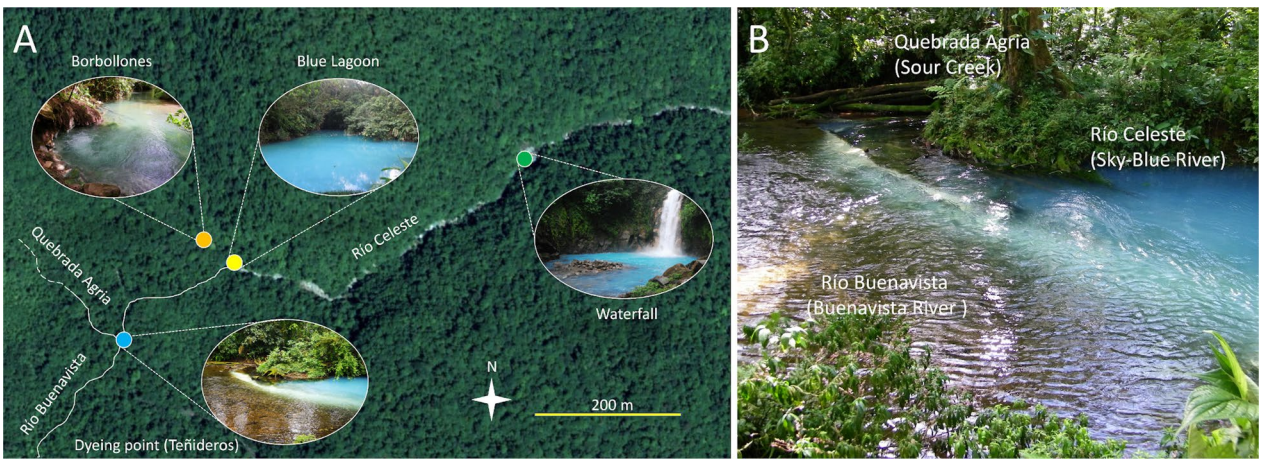

Fig. 1 Environmental setting. **a** Satellite view of the Río Celeste area. River paths are shown with continuous lines when they are not directly evident from the satellite imagery. Dashed lines in Quebrada Agria indicate that the upstream of this river is still uncharted. **b** Detail of the 'Teñidero' ('dyeing point') confluence, showing the confluence between Río Buenavista and Quebrada, and the origin of the sky-blue river Río Celeste

multitude of niches such as thermal waters, acidic rivers or mineral-rich springs [2, 3]. Even though other rivers in Costa Rica show some sky-blue color, the 'Teñidero' stands out as a unique point of research and popular interest since only here the confluence of two seemingly colorless streams instantly turn into the sky blue-colored Río Celeste (Fig. 1).

Sampling and field measurements

All necessary permits for sampling water and sediments were obtained from the National System of Conservation Areas (SINAC) of the Ministry of Environment and Energy (MINAE) of Costa Rica (Resolution No. 097-2014-ACAT). On 19 January 2014, samples of water and sediments were collected from Río Buenavista and Quebrada Agria about 5 m before their confluence (10°42′02.5″ N, 84°59′49.3″ W). Río Celeste's samples were taken 20 m after the confluence. At each sampling point, 3 surface water samples (~ 10 cm depth, 1L each) and 3 sediment samples (between 10 and 50 g) were taken across the river. All samples for DNA analysis were collected in clean and sterile glass bottles, chilled on ice, stored at 4 °C, pooled and processed within less than 24 h. For qualitative analysis of diatoms, three surface water samples (~ 10 cm depth, 50 mL) were collected in 50-mL Falcon tubes for each stream, fixed and preserved with 1% Lugol's iodine solution and 3% neutralized formaldehyde, and allowed to sediment for 24 h. Biofilm samples were also collected by brushing the surface of stones from the bottom of Río Celeste and preserved as described above. The chemical and physicochemical parameters for the three streams can be found at Additional file 1: Table S1 [6].

Characterization of suspended materials

Water samples from Quebrada Agria, Río Buenavista and Río Celeste (1 L each) were individually filtered through porous membranes with a nominal pore size of 0.4 µm. The collected filters were allowed to dry at room temperature (20–25 °C). No residues were observed on the filter from Quebrada Agria. The residues collected from the water samples of Río Celeste and Río Buenavista were adhered to electrically conductive carbon tape for their analysis on a scanning electron microscope (SEM, Jeol JSM-IT500). This apparatus was equipped with a detector for energy dispersive X-ray spectroscopy (EDS) to determine the bulk chemical composition of the suspended materials collected in the filters. Infrared spectra of the suspended and deposited particulate materials from Río Celeste were measured with a Fourier-transform infrared spectrometer (Perkin Elmer, Spectrum 1000). The samples (10 mg) were mixed with anhydrous KBr (990 mg), finely ground using a mortar and pestle, and pressed to form discs containing the particles. With these discs, the spectra were obtained in the wavelength interval 400–4000 cm^{-1}. Speciation calculations were carried out with PHREEQC [10]. For Río Buenavista, a measured alkalinity of 58.4 as bicarbonate was used.

For the precipitation experiments, water samples from Río Buenavista and Quebrada Agria (100 mL each) were respectively acidified with 0.005 M H_2SO_4 and neutralized with 0.01 M Na_2CO_3 until they reached the pH of Río Celeste (pH 5.5). Only the neutralization of the sample from Quebrada Agria produced suspended particles that were collected by centrifugation. The collected particles were resuspended in de-ionised water, centrifugated again and subsequently dried at room temperature. Such particles were pasted on carbon tape and analysed by SEM–EDS.

Total DNA isolation, construction of 16S rRNA gene libraries and illumina sequencing

Three water samples (1L each) taken at different points along the river width were filtered through a vacuum filtration system under sterile conditions using a membrane filter (pore size 0.22 µm; Millipore, GV CAT No GVWP04700). To prevent filter rupture, a support membrane (pore size 0.45 µm; Phenex, Nylon Part No AF0-0504) was placed below. The upper filter was collected and stored at − 80 °C until processing. The DNA was extracted from aseptically cut pieces of the filter using the PowerSoil® DNA Isolation Kit (MoBio, Carlsbad, CA, USA) as described by the manufacturer. Cells were disrupted by two steps of bead beating (FastPrep-24, MP Biomedicals, Santa Ana, CA, USA) for 30 s at 5.5 m s^{-1}. For the construction of microbial 16S rRNA amplicon libraries, the V5-V6 hypervariable regions were PCR-amplified with universal primers 807F and 1050R [11]. Barcoding of the amplicons and addition of Illumina adaptors were conducted by PCR as described previously [12, 13]. The PCR-generated amplicon libraries were subjected to paired-end 2 × 250 Illumina MiSeq sequencing (Illumina, San Diego, CA, USA).

Bioinformatic analysis of 16S rDNA amplicon data

Raw MiSeq sequences were quality-filtered and merged with *moira.py* v1.3.2 [14] with options–*paired–ambigs disallow–consensus_qscore posterior–qscore_cap 0–maxerrors 1–collapse True–alpha 0.005–output_format fastq*. A custom *python* script was used to expand the derreplicated files generated by *moira.py* so they could be used in subsequent analysis steps, and the contigs were then derreplicated and chimera screened with DADA2 v1.14.1 [15]. The resulting Amplicon Sequence Variants (ASVs) were aligned to the SILVA

nr 132 reference database [16] using *mothur v1.44.0* [17]. The alignment was curated with *mothur* using the *screen.seqs* (options: *start = 25,318, end = 33,595, maxhomop = 7, minlength = 200, maxlength = 275*) and *filter.seqs* (options: *vertical = T, trump = .*) commands, after which a UPGMA tree was constructed from the filtered alignment using *phangorn* v2.5.5 [18]. Variance-adjusted weighted UniFrac distances [19] were obtained from the tree with *GUniFrac* v1.1 [20]. Following McMurdie & Holmes [21], sequences were not rarefied prior to weighted UniFrac distance calculation. These distances were in turn used to compute a Non-metric Multidimensional Scaling with the *metaMDS* function from *vegan* v2.5.6 [22]. Further, we performed permutational multivariate analysis of variance [23] as implemented in the *PermanovaG* function from *GUniFrac* package to assess whether community composition was influenced by sample type (Water vs Sediment) and origin (Hydrothermally influenced vs Neutral). Finally, we used SINA [24] to taxonomically classify the ASV sequences, and SQMtools [25] to generate barplot figures. The following identity cutoffs to the closest match in the database were assigned to assign ASVs to the different taxonomic ranks [26]: phylum = 75%, class = 78.5%, order = 82%, family = 86.5%, genus = 94.5%. The taxonomy of each ASV is thus reported at the maximum resolution attainable without compromising a correct classification. Raw sequences were submitted to the sequence-read archive (SRA) under BioProject PRJNA747923. In order to be able to compare the composition of our samples to that of a known hydrothermally-influenced microbial community, we included the sequencing data from the hydrothermal spring Borbollones (located also inside the Tenorio Volcano National Park; Fig. 1) that was obtained in our previous study using the same sequencing methodology [3]. Reads assigned to Chloroplasts with SINA were further were annotated based on the naïve Bayesian classification with a pseudo-bootstrap threshold of 80% using RDP set18 [27]. Among them, those with more than 1000 total counts were manually annotated using the online Nucleotide BLAST tool from NCBI [28], and annotated against the Phyto REF database [29] using minimap2 [30].

Identification of diatoms by optical and electronic microscopy

Both the water samples and biofilms were sedimented at room temperature for 24 h. The sedimented material was examined under a light microscope (Model IX-51, Olympus) at 40X or 100X optical magnification. Selected samples were analyzed in a scanning electron microscope (Model S-3700 N, Hitachi) using an accelerating voltage of 15 kV.

Results

Formation of minerals at the 'Teñidero' confluence

In order to determine the nature and origin of the light-scattering particles, we isolated the suspended material from the streams of Río Buenavista and Río Celeste, and analyzed them in bulk with scanning electron microscopy (SEM, Fig. 2), energy dispersive X-ray spectroscopy (EDS, Additional file 2: Fig. S1) and Fourier transform infrared spectroscopy (FTIR, Fig. 2g). The white deposits covering the rocks downstream from 'Teñidero' were also examined by FTIR.

Suspended matter in Río Buenavista consisted of particles that were too small to effectively scatter light according to the Mie principle, mixed with a few larger particles (e.g. frustule fragments) (Fig. 2a, b). Particles in Río Celeste River water formed larger aggregates in the filtered sample (Fig. 2c, d). A composition map assessed on a broad region (11.0 μm × 8.3 μm) of the solid from Río Celeste indicated the presence of aluminum, silicon and oxygen together with small and variable amounts of sulfur, with an Al/Si ratio of 1.74 (Additional file 2: Fig. S1). The FTIR spectrum of the particles in 'Teñidero' was found to resemble that of the deposits on the rocks downstream (Fig. 2g). Quebrada Agria lacked relevant suspended particles.

Acidification of Río Buenavista water did not produce any particles while the elevation of pH in Quebrada Agria waters produced abundant particles rich in aluminum, silicon and oxygen, with an Al/Si ratio of 2.72 (Fig. 2e,f).

Microbial community composition at the Quebrada Agria (QA) stream

The microbial community in Quebrada Agria was dominated by Proteobacteria and Campylobacterota (Fig. 3a, QAW) assigned to known sulfur-oxidizing genera such as *Sulfuriferula* (Fig. 4, ASVs 1, 18), *Halothiobacillus* (ASV 3), *Sulfurimonas* (ASVs 4, 5), *Thiomonas* (ASV 44) or *Acidithiobacillus* (ASV 46). Archaea (Thaumarchaeota and Crenarchaeota) were also detected. Overall, the microbial composition of Quebrada Agria waters resembles to that of Borbollones (Fig. 3a, BBW, see hierarchical clustering), a nearby-located (at *ca* 300 m) hydrothermal spring whose microbial composition was recently described [3]. Overall, the sediment featured similar ASVs as the waters, but in different proportions (Fig. 3a, QAS). Campylobacterota were missing and Proteobacteria were present in lesser amounts and dominated by the iron-oxidizing family Gallionellaceae (Fig. 4, ASVs 14, 24), while the proportion of Archaea (Nitrosotaleaceae family, ASV 7, 26), Actinobacteria, Acidobacteria, Nitrospirota and Planctomycetota were higher. Some high-abundance ASVs were endemic to the QA sediments; this includes an Acidothermaceae

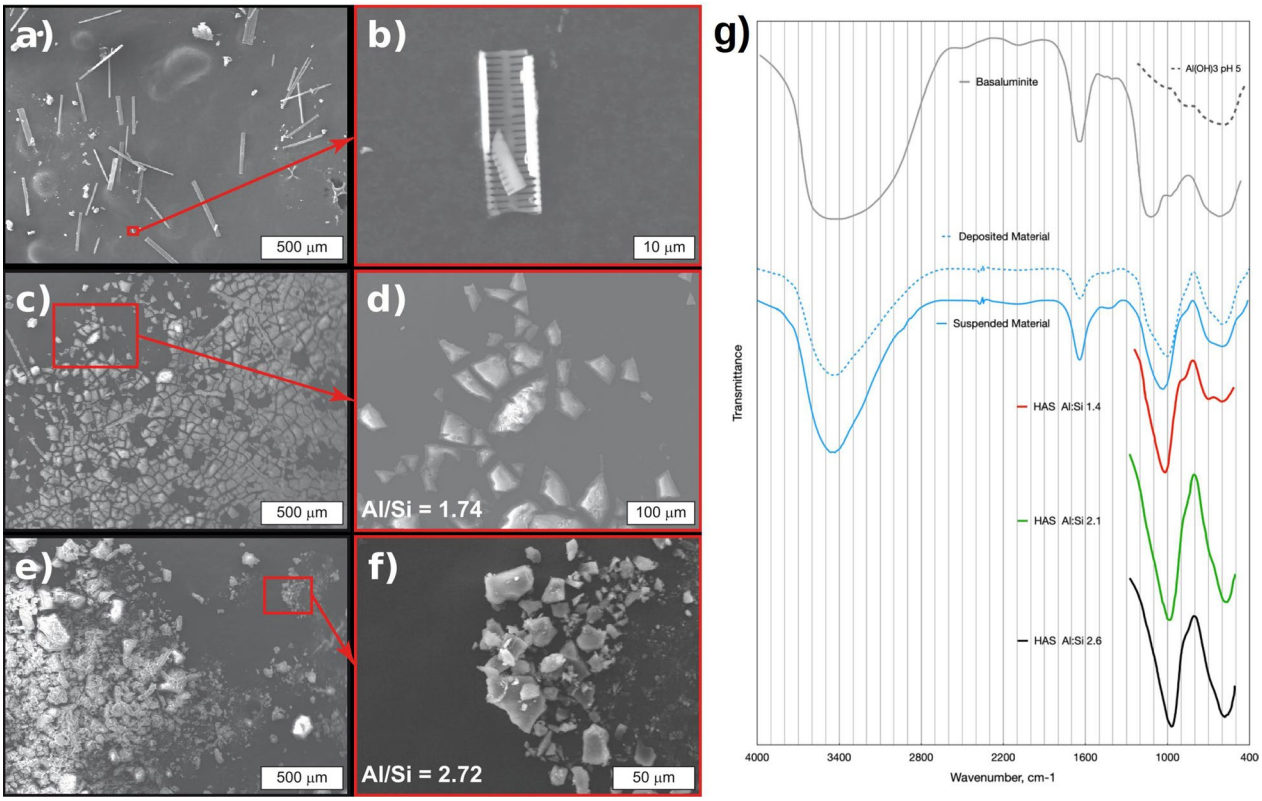

Fig. 2 Chemical characterization of filtered or precipitated materials from the water streams at the confluence that originate Río Celeste. **a-f** Scanning electron microscopy (SEM) images from different materials The atomic Al/Si ratios were obtained by EDS scans. **a** Filtered materials from Río Buenavista; the solids retained by the membranes are mainly frustule fragments. **b** Detail of a frustule fragment. **c** Filtered materials from Río Celeste; frustule fragments along with agglomerated particles are retained by the filter membranes. **d** Magnification showing a detail of the agglomerated filtered particles from Río Celeste. **e** Agglomerated particulate materials obtained by in vitro precipitation of a water sample from Quebrada Agria. **f** Magnification performed on the specimen in **e**, showing the agglomerated particles. **g** Fourier-transform infrared (FTIR) spectrum of the particles suspended in Río Celeste, the precipitates occurring downstream from 'Teñidero', and other minerals known to precipitate in similar settings, including hydroxyaluminosilicates (HAS)

actinobacterium (ASV 47) and a Nitrosomonadaceae gammaproteobacterium (ASV 81).

Microbial community composition at the Río Buenavista (RB) stream

The most abundant phyla in Buenavista were Proteobacteria, Campylobacteriota and Bacteroidota (Fig. 3a, RBW). Among them, the most abundant bacterial genus was *Sulfurovum* (Fig. 4, ASV 9) followed by *Aeromonas* (Fig. 4, ASV 23). Many of the abundant bacterial taxa were related to the recycling of sulfur and nitrogen, such as *Sulfurimonas* (ASV 32), *Thiovirga* (ASVs 30), *Sulfuricurvum* (ASVs 43), *Thiotrix* (Additional file 3: Table S2, ASV 53), Galliollenaceae (ASV 130) and Comamonadaceae (ASV 149). For eukaryotes, the second most abundant ASV in Buenavista's waters was a Bacillariophyta (diatom) chloroplast (Fig. 4, ASV 12), which in general represented the majority of the cyanobacteria-related reads obtained in this study (Additional file 4: Fig. S2). Since the results obtained with SINA and RDP do not distinguish between different Bacillariophyta, we manually annotated the sequence of ASV12 using the online Nucleotide BLAST tool from NCBI [28]. The first hit was a 100% match to the chloroplast of the Bacillariophyceae diatom *Halamphora coffeaeformis* (NCBI accession number NC_044465.1), but several other perfect matches to other diatom species were also found (Additional file 5: Table S3). We also annotated it against the Phyto REF database [29], which classified it into the Cymbellaceae family. However, other abundant Bacillariophyta ASVs still had 100% matches to Phyto REF reference sequences from different orders (Additional file 5: Table S3), implying that the V5-V6 16S rRNA region amplified in this study is not good at discriminating between chloroplasts from different diatom groups. In order to achieve a better classification, we instead resorted to optical and electronic microscopy (see "Diatom composition at the Río

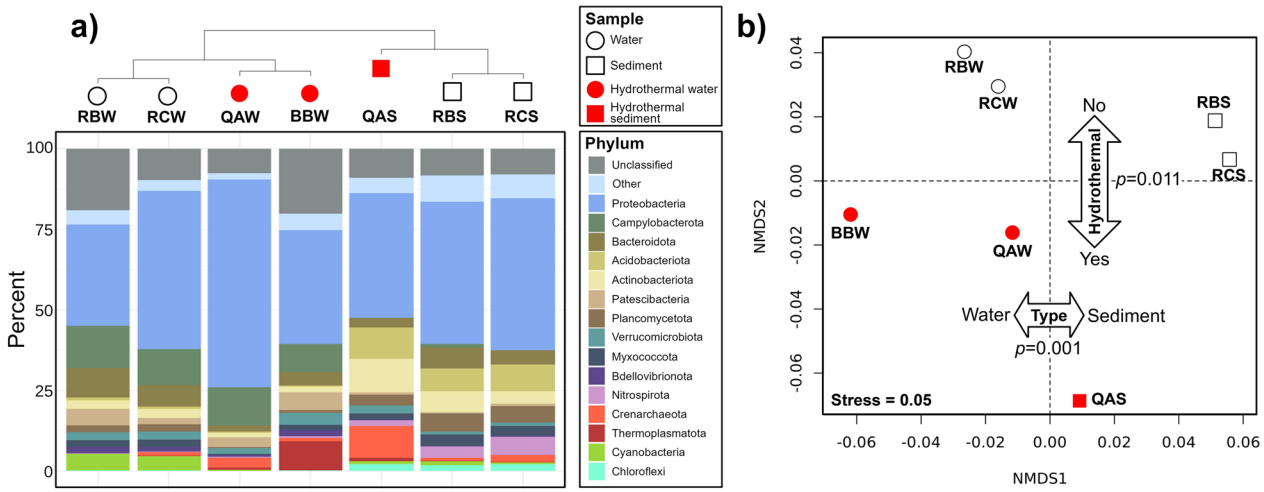

Fig. 3 Microbial community composition around the 'Teñidero' confluence. Samples were from the two streams converging into the dyeing point (sample codes starting by RB and QA) and the resulting mixed stream ('Río Celeste', sample codes starting by RC). For each location, we analyzed plankton from the water (W) and sediments (S). An extra water sample from a nearby hydrothermal spring (BBW) was also included for comparative purposes. **a** Hierarchical clustering of the samples based on their variance-adjusted weighted UniFrac distances (top) and barplot showing their phylum-level composition (bottom). **b** Non-metric multidimensional scaling (NMDS) of the samples included in this study. Plotting symbol and color denote sample type (circle: water, square: sediment) and origin (white: neutral samples; red: hydrothermally-influenced samples). Two-way arrows and the associated p-values indicates a significant influence (PERMANOVA test using UniFrac distances) of sample type and origin in microbial community composition

Celeste (RC) stream" section). As for Quebrada Agria, the microbial population in Buenavista's sediment was a mixture of endemic taxa and taxa which were also present in the water. Campylobacterota were at much lower relative abundances than in the water, and Actinobacteria, Proteobacteria, Acidobacteria, Nitrospirota and Planctomycetota increased (Fig. 3a, RBS). Among the taxa present in the sediments but not in the waters, the most abundant belonged to the Alphaproteobacteria and Gammaproteobacteria classes (Additional file 3: Table S2, ASVs 50, 58, 66, 86, 88) and the Bacteroidota and Nitrospirota phyla (ASV 185 and 120, respectively).

Microbial community composition at the Río Celeste (RC) stream

The composition of Río Celeste's planktonic microbiome was overall more similar to that of Río Buenavista than to that of Quebrada Agria (Fig. 3b), which is to be expected since Buenavista is the largest source of Celeste's waters (*ca.* a 3:1 proportion with Quebrada Agria according to mass balance of conservative elements). The contribution of Quebrada Agria was still apparent from the high relative abundance of Proteobacteria (*Sulfuriferula;* Fig. 4, ASVs 1, 18) and Archaea. The sediment of Río Celeste was very similar to that of Río Buenavista, except for the higher presence of a few ASVs (Fig. 4, ASV 7–Nitrosotaleaceae; Additional file 3: Table S2, ASVs 61–Nitrospira and 74—Uncultured Planctomycetales) attributed to the influence of Quebrada Agria.

Diatom composition at the Río Celeste (RC) stream

The analysis by optical- and electron-microscopy validated the results obtained by sequencing. A great variety of diatoms with different morphologies and sizes between 2 and 200 µm were observed (Additional file 4: Fig. S2). Morphologically, we identified organisms from the Bacillariophyta and Miozoa phyla. The identified species can be grouped into 3 classes (Bacillariophyceae, Mediophyceae and Dinophyceae) and 6 families (Diploneidaceae, Biddulphiaceae, Surirellaceae, Cymbellaceae, Catenulaceae and Peridiniaceae). Specifically, diatoms of genera *Diploneis*, *Biddulphia*, *Surirella*, *Cynbella*, *Amphora* were observed. The dinoflagellate *Peridinium* was also recorded being one of the most abundant taxa observed with light microscopy (Additional file 4: Fig. S2b).

Discussion

Hydroxyaluminosilicates are the main component of the minerals suspended in Río Celeste

Isolation of Río Celeste's light-scattering particles by filtration through a 0.45 µm filter afforded a residue that we examined by SEM/EDX (Fig. 2c). The aggregates were larger than the diameters previously predicted by dynamic light scattering [6]; this can be explained by flocculation during filtering, drying and centrifugation, as it has been described previously [31], as well as by the removal of the smaller-sized particles during filtering. We thus find it safe to assume that the aggregates

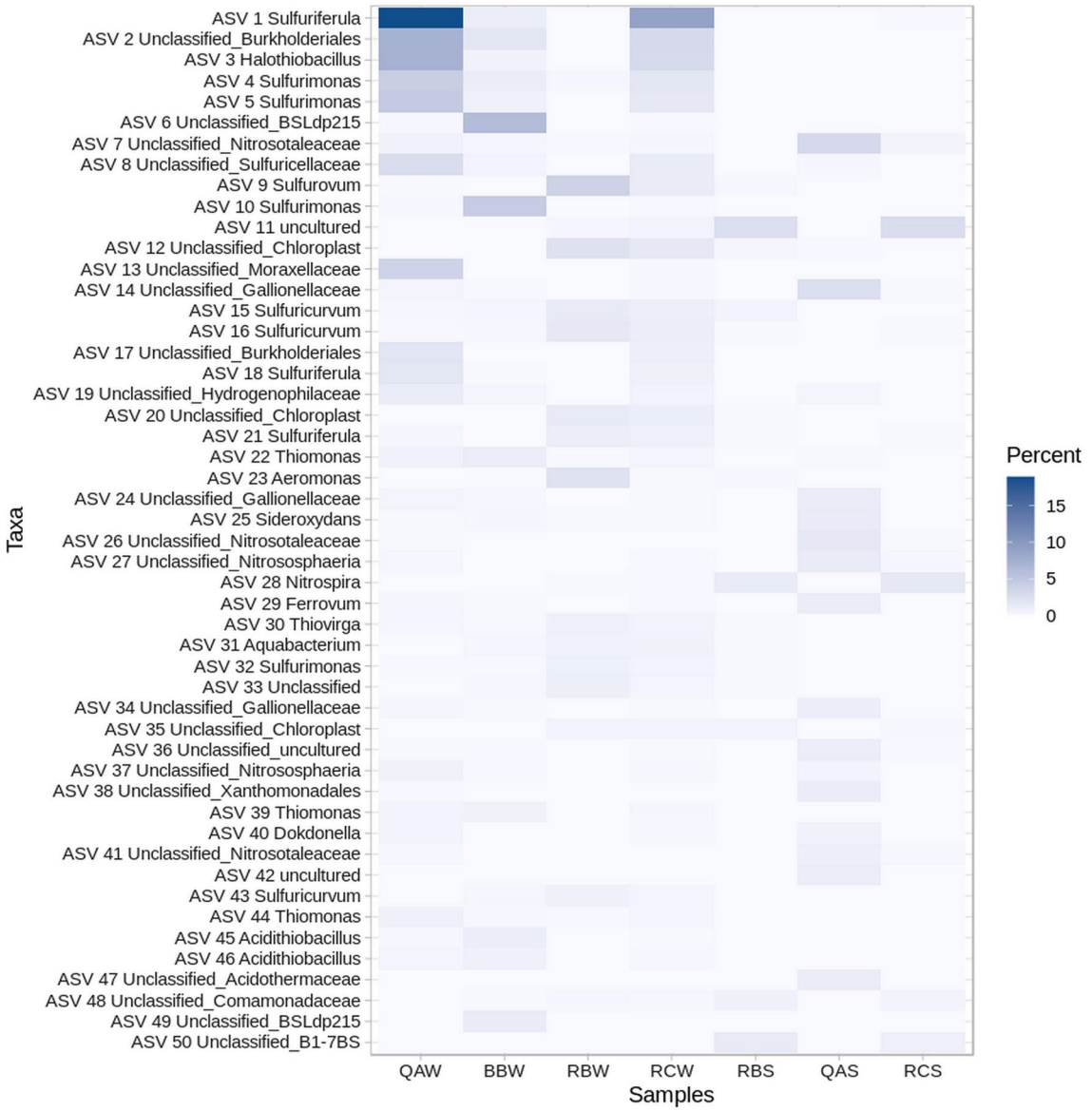

Fig. 4 Distribution and taxonomy of the 50 most abundant ASVs. Heatmap intensity represents percentage abundance. The first two letters in the sample code indicate the river (QA: Quebrada Agria, BB: Borbollones, RB: Río Buenavista, RC: Río Celeste), the third letter indicates the material (W: water, S: sediment)

were representative of the light-scattering suspended materials.

We initially thought the precipitated particles, and the white deposits observed downstream, would be similar to those observed in the mixing of other acidic and neutral rivers and would consist of $Al(OH)_3$ or the hydroxysulfates hydrobasaluminite or basaluminite [32, 33]. However, the composition of the particles suspended in Río Celeste was (in %) Al, 22.4; Si, 12.8; S 0.24; Fe 0.94, which not only contained a high amount of Si but also were too low in S compared to basaluminite's 7% S. (Additional file 2: Fig. S1).

The low-frequency IR-spectra range of both the suspended particles and the rock deposits show only two main bands at 1000 and 605 cm^{-1} with shoulders at *ca.* 1100 and 680 cm^{-1} respectively. They are very different from basaluminite, which has strong IR bands at 606 and 1130 cm^{-1} [34]. (Fig. 2g) We could not gather enough suspended particles to obtain an XRD pattern but in our previous study the material deposited in the rocks showed a poorly crystalline material with broad peaks at 3.3 and 2.25 Å (Cu-K alpha, 2θ at 27 and 40°; see Additional file 2: Fig. S1 in [6]). This contrasts markedly with the 9.68 and 4.55 Å peaks of basaluminite [35]. Initially we could only match our

XRD pattern to that of poorly crystalline aluminosilicates formed during sol–gel reactions and was the basis of our assignment. Other authors studying blue-colored lakes in Japan reached similar conclusions when studying the corresponding deposits of crystalline [9] or amorphous [36] aluminosilicate particles and suggested the amorphous particles have an "allophane-like" XRD pattern, with features at 3.4 and 2.3 Å similar to Río Celeste's deposits.

Precipitation of aluminosilicate solids from solutions containing Al^{3+} and H_4SiO_4 starts at pH 5–6 when Al ions hydrolyze and begin forming gibbsite-like sheets that react with monosilicic acid to form amorphous aluminosilicate precursors [37]. In andosols these precursors might eventually form the Short-Range Ordered Aluminosilicates (SROAS) allophane and imogolite but also might give rise to very insoluble hydroxyaluminosilicates (HAS) [38, 39] with end-member compositions similar to those of allophane, $(Al_2O_3)(SiO_2)_{1.3-2}{\cdot}2.5–3\ H_2O$ or imogolite or kaolinite, $Al_2SiO_3(OH)_4$. The first, named HAS_A forms if the concentration of $Si(OH)_4$ is less than or equal to the total aluminum and has an Al/Si ratio close to 1. Where the initial concentration of $Si(OH)_4$ is at least twice that of aluminum, further reaction with excess silicic acid yields HAS_B with a formula resembling imogolite and has Al/Si ratio close to 2. (Beardmore, 2016). For convenience, we will refer to the particles at the Teñidero as colloidal aluminosilicates which might consist of one or more of the above precursor phases, HAS or SROAS.

Due to their high insolubility the HAS are believed to be the main species controlling the solubility of aluminum in groundwater and in soils [40–42]. The IR spectra of all these SROAS show bands with frequency and relative intensity dependent on their Al/Si ratio and finally provided a good match and a way to characterize Río Celeste's particles. [43–45]. Our sampled particles' IR spectrum matches very well the asymmetric, high-frequency skewed bands of solids with Al/Si ratios around 1.4 as seen in Fig. 2g and compares well to Río Celeste's particles 1.7 Al/Si ratio. The powder XRD patterns of these substances even match the 3.3 and 2.25 Å features found in Río Celeste deposits (compare Additional file 2: Fig. S1 in [6] with Fig. 2 in [45]).

Chemical and environmental differences can explain precipitation of colloidal aluminosilicates in Río Celeste and precipitation of Basaluminite in other rivers.

Fast-mixing conditions at the Teñidero clearly produce a group of substances with structural and compositional properties similar to those of the HAS or SROAS. In contrast, in other studied bluish-water sites where acidic river water mixes with non-acidic tributaries, hydrobasaluminite and basaluminite are likely controlling aluminum solubility. Wanner et al. [34] characterized basaluminite deposits by IR spectroscopy at Ova Lavirun alpine creek. Water mixing calculations using PHREEQC were consistent with basaluminite controlling Al concentration as proposed for similar sites [32, 33]. At Colorado's Paradise Portal site [46] the authors were able to examine the particles by TEM but they decomposed under the electron beam and did not yield an electron diffraction pattern. As the nanoparticles contained Al, S and Si their composition was assigned to hydrobasaluminite and amorphous silica based on their finding that hydrobasaluminite was the only aluminum phase identified in the fresh riverbed precipitates. Finally, in two Spanish acidic pit-lakes, hydrobasaluminite and schwertmannite were clearly detected along with kaolinite. However, the authors were careful to notice that incorporation of silica into these minerals could not account for all the SiO_2 removal from water, and proposed that precipitation of proto-aluminosilicates would be consistent with the presence of kaolinite in the sediments [47].

In Additional file 6: Table S4 we show that Quebrada Agria has a much lower SO_4/Al molar concentration ratio of 4.4 compared to the ratios in Ova Lapirum (19.9) and Paradise Portal (12.9) which could help explain why we do not observe sulfate materials like basaluminite. Our Si/Al concentration ratio of 3.0 is also higher, also consistent with formation of a more Si-rich phase. In Additional file 7: Note S1, we carried out PHREEQC speciation calculations of Río Celeste water at pH 5 using our analysis data and found HAS_B is actually expected to precipitate under those conditions. However, if we instead input the aluminum, sulfate and silicic acid concentrations from the precipitation sites of the Ova Lapirum and Paradise Portal sites into the speciation calculation, the saturation index for HAS_B becomes negative showing the effect of the different Al, SO_4 and Si concentrations. Interestingly, the waters from the acidic pit lakes where hydrobasaluminite was found along with kaolinite also have low SO_4/Al concentration ratios (5.5 and 6.3) close to that of Río Celeste. The waters show however a Si/Al ratio between that from Río Celeste and the two hydrobasaluminite-only precipitating sites making this is an intermediate case where both aluminosilicates and aluminum oxo sulfates form. Precipitation of specific Si and Al- containing mineral phases seems to have a strong dependence on element concentrations and speciation in the waters as well as local kinetic factors.

Quebrada Agria (QA) is influenced by upstream hydrothermal activity

Quebrada Agria is a transparent and acidic stream (pH 3.1) largely devoid of suspended materials [6] while having a high content of dissolved sulfate (190 mg/L), chloride (71 mg/L), calcium (55 mg/L), and aluminum

(12 mg/L) (Additional file 1: Table S1; [6]). While this distinctive chemistry suggests the influence of hydrothermal activity, the source of Quebrada Agria's waters could not be determined during our sampling campaign and likely lies deep within the Tenorio Park rainforest. We therefore turned to the microbial community composition, and found that it was similar to that of the nearby (*ca* 300 m) Borbollones hydrothermal spring [3], sample code BBW). Samples from Quebrada Agria and Borbollones had a significantly different composition compared to the rest of the samples (Fig. 3b, PERMANOVA $p=0.011$), and both shared prominent and functionally defined microbial taxa such as the sulfur-oxidizers *Sulfurimonas* and *Sulfuriferula* and the iron-oxidizer *Gallionela* (Additional file 3: Table S2). These genera have also been found in other hydrothermally influenced settings [52–54].

Based on these chemical and microbiological similarities, we hypothesize that Quebrada Agria is influenced by one or more upstream hydrothermal springs akin to the one described in [3]. Under this scenario, sulfide originating from hydrothermal fluids would be oxidized by microbial activity [55, 56], causing a decrease in pH and high sulfate concentrations observed near the 'Teñidero' site. The acidity of the water would in turn promote weathering and result in extensive mobilization of dissolved silicon and aluminum [33].

Quebrada Agria is the main contributor to the generation of light-dispersing colloidal aluminosilicates

To further explore the origin of the particles, we performed an in vitro experiment in which water samples from the neutral Río Buenavista and the acidic Quebrada Agria were respectively acidified and neutralized to pH 5. Acidification of Río Buenavista water produced no noticeable changes upon reaching pH 5 indicating that a lowered Z-potential aggregation mechanism is not operating as we previously thought. However, a conspicuous proliferation of colloids was observed on the sample from Quebrada Agria upon taking it to pH 5 with sodium carbonate solution (Fig. 2c). This strongly suggests that, contrary to our original hypothesis, dissolved material from Quebrada Agria is the main source of the Río Celeste particles. The Al/Si atomic ratio in the synthetic particles was 2.7, higher than the 1.7 ratio observed in the naturally-occurring particles in Río Celeste. When bicarbonate alkalinity from Río Buenavista raises the pH of Quebrada Agria to the field-measured pH 5–6 values, precipitation starts and the high-silica river provides extra $Si(OH)_4$ to the reaction thus avoiding partial formation of amorphous $Al(OH)_3$ which could be thought to account for excess Al. It is believed that HAS form when silicic acid reacts with the growing $Al(OH)_3$ gibbsite sheet. As we are observing the initial stages in the precipitation, it is conceivable that some $Al(OH)_3$ could be present at some point. However, we only see evidence of its presence at small levels in the basification of Quebrada Agria experiment (Figure EDS) that produces a solid with a 2.7 Al/Si ratio. It is conceivable that the white rock-deposits downstream from the teñidero contain small amounts of other aluminum phases, accounting for excess aluminum in their composition (% Al, 33.8; Si, 10.7; S 2.40; Fe 4.99). Unfortunately, amorphous $Al(OH)_3$ precipitated at pH 5 lacks the strong 1000 cm^{-1} IR band and only shows a broad band at 570 cm^{-1} not detectable in our IR spectrum amongst other phases (Fig. 2g). Furthermore, when precipitated at pH 5, $Al(OH)_3$ has an essentially featureless XRD pattern and will not show in our XRD measurements [48] that only have features assignable to poorly crystalline aluminosilicates.

At the Teñidero, particles form within a few seconds and the resulting light-scattering stable colloidal phases are visible at least 14 km downstream. Smaller particles from Río Buenavista [6], probably become embedded with the precipitate at the Teñidero accounting for its smaller 1.7 Al/Si ratio and iron content. The chemistry of the Teñidero then bears some resemblance to that of flocculation water treatment processes that adsorb clay particles on positively charged Al/Si flocs prepared by reaction of aluminum potassium sulfate (alum) and silicic acid at pH 5 and above [49–51]. Our work is only concerned with the initial colloidal phase where the transformation of aluminum hydroxide into HAS or even SROAS is starting to take place. Enough dissolved species do remain in Río Celeste water and additional mineral phases are likely to form downstream but are not the focus of this report.

Río Buenavista (RB) is a circumneutral stream with standard chemical composition

The Río Buenavista has many of the features typically seen in freshwater ecosystems: pH is close to neutral (6.8) and the chemical composition of the water is within the established levels of metals and anions suitable for human consumption (Additional file 1: Table S1). The river is largely transparent and shows no noticeable suspended particles but do contain silica particles with an average diameter of 184 nm [6].

The most abundant ASV in Buenavista's waters were assigned to the *Sulfurovum* genus (Fig. 4, ASV 9). Other genera related to sulfur cycling such as *Sulfurimonas, Thiovirga* or *Thiotrix* were also present in moderate amounts. The dominance of *Sulfurovum* is somewhat

puzzling, as this genus is usually found in hydrothermal vents [52], cold seeps [57], groundwater aquifers [58], freshwater sediments [59] or as endosymbionts of larger organisms [60]. It is unclear if any of these conditions would apply in Río Buenavista waters: they could be originating from the sediments (where they are also present, albeit at lower relative abundance), be indicative of hydrothermal activity (although Buenavista's waters have an apparently standard chemical composition), or originate from some nearby groundwater seep.

The second most abundant bacteria (third most abundant ASV including non-bacterial sequences) was an ASV from the heterotrophic *Aeromonas* genus (Fig. 4, ASV 23). *Aeromonas* is commonly seen in freshwaters and drinking waters [61], and appears either as a free-living organism or associated to crustaceans or fish, for which they can act as pathogens [62]. The abundance of this genus in freshwaters has been considered an indicator of trophic status, analogous to Secchi disc depth, total phosphorus, or chlorophyll *a* [63]. Their high relative abundance suggest that Río Buenavista is somewhat eutrophic.

Finally, the second most abundant ASV in Río Buenavista's waters was classified as a Bacillariophyta (diatom) chloroplast (Fig. 4, ASV 12). Manual search against the Phyto REF database indicated that the chloroplast 16S rRNA gene sequence was affiliated to the Cymbellaceae family. (Additional file 5: Table S3). This was confirmed by light and electron microscopy, which revealed the presence of the diatom families Diploneidaceae, Biddulphiaceae, Surirellaceae, Cymbellaceae, Catenulaceae and Peridiniaceae (Additional file 4: Fig. S2). *Cymbellaceae* (found both by amplicon sequencing and microscopy) have been reported in other tropical rivers [64–66] including rivers from Costa Rica [67]. Other diatom taxa commonly reported in these rivers, such as *Nitzschia* or *Gomphonema* [64, 65, 67, 68] were however not detected in our study. We note that the abundance of diatoms in Río Buenavista might be much higher than suggested by the relative abundance of ASV 12, since our sample processing protocol was not optimized for retrieving DNA from hard-to-lyse eukaryotic cells. Our observations are nonetheless consistent with previous reports of high diatom activity in tropical rivers with low turbidity [69].

Diatoms as sources of colloidal silica in Río Buenavista

Diatoms incorporate dissolved silicate into their frustules, being the main source of biogenic silica in riverine settings [70, 71] and under some circumstances accounting for the majority of the total silicon budget [70, 72]. Most of the biogenic silica in rivers is not actually associated to living diatoms, but to detrital particles that originate when diatoms die or are preyed upon [73, 74]. Interestingly, diatom cultures have been shown to produce extracellular silica nanoparticles in the 150–400 nm diameter range [75], which is compatible with the 184 nm average diameter of the particles found in Río Buenavista [6]. Lithogenic silica (coming from the weathering of minerals such as clays, silts and sand) can also represent a large fraction of the total silicon in rivers [76]. The particle size from such processes is however in the micrometer scale [77, 78], much larger than the particles detected here [6]. Additionally, Río Buenavista was transparent and devoid of apparent particulate material. Overall, the data are consistent with the notion that diatoms, rather than lithogenic sources, are responsible for the colloidal silica present in the water from Río Buenavista. Diatom activity would therefore control the partitioning of silicon between the dissolved and colloidal forms in Río Buenavista, making them indirect contributors to the Río Celeste phenomenon.

Biological influence in the distinctive color of Río Celeste

Río Celeste arises at the 'Teñidero' confluence, when the acidic and aluminum-rich Quebrada Agria mixes with the neutral Río Buenavista. Our experiments suggest that the partial neutralization of Quebrada Agria initiates a precipitation reaction in which aluminum and silicic acid precipitate with dissolved silica and likely incorporate also smaller particles coming from Río Buenavista, resulting in colloidal hydroxyaluminosilicates that scatter light as described in [6]. The interaction of two streams with a characteristic physicochemical composition is thus necessary for the emergence of the unique Río Celeste phenomenon. Our data strongly suggest that, in both streams, this composition is biologically mediated (Fig. 5).

In Quebrada Agria, hydrothermal sulfide is oxidized by bacteria such as *Sulfuriferula and Sulfurimonas* resulting in an acidic fluid capable of mobilizing silicates and metals such as aluminum, as has been described for other ecosystems [33, 55, 56]. Meanwhile, the partitioning of silicon in Río Buenavista seems to be controlled by diatoms, which are in turn the most likely source for the colloidal silica particles suspended in the stream. Diatom particles are known for their ability to act as flocculants [79, 80], and incorporate aluminum both during biogenesis [81] and afterwards [82, 83], leading to deposition of aluminosilicates in riverine systems [84]. Diatoms might therefore be facilitating the chemical precipitation of minerals at 'Teñidero', as previously described for other

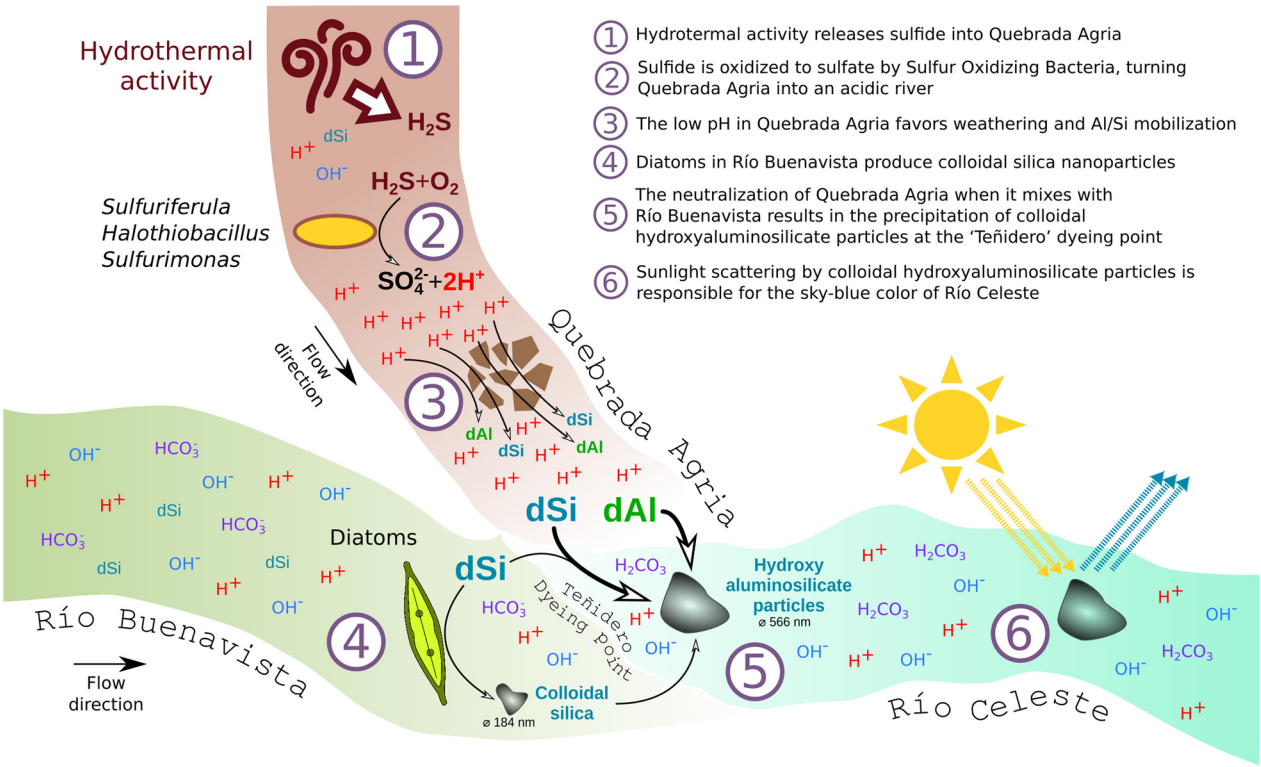

Fig. 5 Conceptual biogeochemical model of the origin of Río Celeste's sky-blue color

system [85], resulting in the precipitation of particles of the right size for a more intense blue coloration.

Conclusions

The sky-blue color of Río Celeste originates at the interface between chemistry and biology, and thus represents a textbook example of emergent behavior in environmental microbiology: by themselves, the biogeochemical processes operating in Quebrada Agria and Río Buenavista are fairly common in nature, but when put together they give rise to what could be described as a small natural wonder. Río Celeste is of relevance not only from the geobiologist's point of view, but also for its socioeconomic impact in the Costa Rican ecotourism industry. This unique phenomenon might however be fragile: the coloring of the river depends on just the right combination of factors, and this delicate equilibrium could very well be threatened by human activity and environmental change. Further research is therefore required in order to characterize the nature and temporal dynamics of its chemical and microbiological components, and assess its vulnerability to biotic and abiotic stressors.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40793-023-00464-2.

Additional file 1. Table S1 Chemical composition and physicochemical properties of the three streams studied in this work. Data taken from [6], DOI: 10.1371/journal.pone.0075165.

Additional file 2. Fig. S1 Energy dispersive X-ray spectroscopy (EDS) of the particles shown in Figure 2

Additional file 3. Table S2 Sequence, abundance distribution and taxonomy of the Amplicon Sequence Variants (ASV) detected in this work. Sequences are annotated with SINA, Chloroplast sequences were further annotated with RDP.

Additional file 4. Fig. S2 Presence of diatoms around the 'Teñidero' confluence. a) Relative abundance of cyanobacteria-like ASVs in the different samples. Most of the Bacillariophyta ASVs could be further assigned to different diatom groups (Additional file 5: Table S3). Sample codes are described in the legend for Figure 3b, b) Light microscopy image from a Río Celeste sample showing the presence of diatoms with different morphologies, c) Light microscopy image from a Río Celeste sample showing the presence of the dinoflagellate Peridinium, d-f) Electronic microscopy images showing details on diatom frustules.

Additional file 5. Table S3 Closest NCBI nr and Phyto REF matches of the most abundant (more than 1000 total counts) Chloroplastic ASVs. For each ASV, we include its best hit in the nr database to a sequence belonging to named species, as well as their best hit in the Phyto REF da-tabase. In case two or more hits were tied at the best score, all of them are included

Additional file 6. Table S4 Concentration of selected chemical species in waters where aluminum minerals precipitate.

Additional file 7. Note S1 Parameters used in PHREEQC speciation calculation.

Acknowledgements
We acknowledge the support during field work from the park rangers at Tenorio Volcano National Park and the SINAC administration.

Author contributions
AA-R: Investigation, Writing—Review & Editing; EL: Investigation, Formal analysis, Writing—Review & Editing; EC: Investigation, Formal analysis, Writing—Review & Editing; RA: Investigation; JCC: Investigation, Writing—Review & Editing; MV: Investigation; DP: Funding acquisition, Review & Editing; SB: Writing—Review & Editing; MC: Conceptualization, Funding acquisition, Supervision, Investigation, Writing—Original Draft; Fernando P-S: Conceptualization, Funding acquisition, Formal analysis, Visualization, Writing—Original Draft. All authors read and approved the final manuscript.

Funding
Open access funding provided by Swedish University of Agricultural Sciences. Research was funded by the Vice-rectory of Research of Universidad de Costa Rica (VI 809-B6-524), CENIBiot and by the European Research Council Advanced Grant ERC250350-IPBSL. E. Libby thanks support from University of Costa Rica's Vicerrectoría de Investigación through Grant Pry01-1747-2022, FPS was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 892961.

Availability of data and materials
Raw sequences were submitted to the sequence-read archive (SRA) under BioProject PRJNA747923.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 12 August 2022 Accepted: 10 January 2023
Published online: 19 January 2023

References

1. Arce-Rodríguez A, Puente-Sánchez F, Avendaño R, Libby E, Rojas L, Cambronero JC, Chavarría M. Pristine but metal-rich Río Sucio (Dirty River) is dominated by *Gallionella* and other iron-sulfur oxidizing microbes. Extremophiles. 2017;21(2):235–43.
2. Alvarado, G. E. Costa Rica y sus Volcanes; Editorial UCR, San José, 2021. ISBN 978-9968-46-776-6.
3. Arce-Rodríguez A, Puente-Sánchez F, Avendaño R, Martínez-Cruz M, de Moor JM, Pieper DH, Chavarría M. Thermoplasmatales and sulfur-oxidizing bacteria dominate the microbial community at the surface water of a CO2-rich hydrothermal spring located in Tenorio Volcano National Park. Costa Rica Extremophiles. 2019;23(2):177–87.
4. Jovanelly TJ, Rodríguez-Montero L, Sánchez-Gutiérrez R, Mena-Rivera L, Thomas D. Evaluating watershed health in Costa Rican national parks and protected areas. Sustain Water Resour Manag. 2020;6(5):1–14.
5. SINAC (Sistema Nacional de Áreas de Conservación). Informe Anual Estadísticas SEMEC 2018: SINAC en Números. Comp. B Pavlotzky. San José, CR, 2019. p 82.
6. Castellón E, Martínez M, Madrigal-Carballo S, Arias ML, Vargas WE, Chavarría M. Scattering of light by colloidal aluminosilicate particles produces the unusual sky-blue color of Río Celeste (Tenorio Volcano Complex, Costa Rica). PLoS ONE. 2013;8(9): e75165.
7. Ohsawa S, Kawamura T, Takamatsu N, Yusa Y. Rayleigh scattering by aqueous colloidal silica as a cause for the blue color of hydrothermal water. J Volcanol Geoth Res. 2002;113(1–2):49–60.
8. Oyama Y, Shibahara A. Simulation of water colors in a shallow acidified lake, Lake Onneto, Japan, using colorimetric analysis and bio-optical modeling. Limnology. 2009;10(1):47–56.
9. Takagai Y, Abe R, Endo A, Yokoyama A, Konno M. Unique aluminosilicate-based natural nanoparticles in the volcanogenic Goshiki-numa pond. Environ Chem Lett. 2016;14(4):565–9.
10. Parkhurst DL, Appelo CAJ. Description of input and examples for PHREEQC version 3—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. US Geol Survey Tech Methods. 2013;6(A43):497.
11. Bohorquez LC, Delgado-Serrano L, López G, Osorio-Forero C, Klepac-Ceraj V, Kolter R, Junca H, Baena S, Zambrano MM. In-depth characterization via complementing culture-independent approaches of the microbial community in an acidic hot spring of the Colombian Andes. Microb Ecol. 2012;63(1):103–15.
12. Camarinha-Silva A, Jáuregui R, Chaves-Moreno D, Oxley AP, Schaumburg F, Becker K, Wos-Oxley ML, Pieper DH. Comparing the anterior nare bacterial community of two discrete human populations using Illumina amplicon sequencing. Environ Microbiol. 2014;16(9):2939–52.
13. Burbach K, Seifert J, Pieper DH, Camarinha-Silva A. Evaluation of DNA extraction kits and phylogenetic diversity of the porcine gastrointestinal tract based on Illumina sequencing of two hypervariable regions. Microbiology open. 2016;5(1):70–82.
14. Puente-Sánchez F, Aguirre J, Parro V. A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. Nucleic Acids Res. 2016;44(4):e40–e40.
15. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3.
16. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic acids research. 2012;41(D1):D590–6.
17. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Weber CF. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537–41.
18. Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics. 2011;27(4):592–3.
19. Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol. 2007;73(5):1576–85.
20. Chen J, Chen MJ (2018) Package 'GUniFrac'. The Comprehensive R Archive Network (CRAN).
21. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol. 2014;10(4): e1003531.
22. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MH, Wagner H, Oksanen MJ. Package 'vegan.' Commun Ecol Package Vers. 2013;2(9):1–295.
23. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, Collman RG, Bushman FD, Li H. Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics. 2012;28(16):2106–13.
24. Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics. 2012;28(14):1823–9.
25. Puente-Sánchez F, García-García N, Tamames J. SQMtools: automated processing and visual analysis of'omics data with R and anvi'o. BMC Bioinformatics. 2020;21(1):1–11.
26. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014;12(9):635–45.
27. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014;42(D1):D633–42.

28. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y, Zaretskaya I. BLAST: a more efficient report with usability improvements. Nucleic Acids Res. 2013;41(W1):W29–33.
29. Decelle J, Romac S, Stern RF, Bendif EM, Zingone A, Audic S, Guiry MD, Guillou L, Tessier D, Le Gall F, Christen R. Phyto REF: a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. Mol Ecol Resour. 2015;15(6):1435–45.
30. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100.
31. Comba ME, Kaiser KL. Suspended particulate concentrations in the St. Lawrence River (1985–1987) determined by centrifugation and filtration. Sci Total Environ. 1990;97:191–206.
32. Nordstrom DK. The effect of sulfate on aluminum concentrations in natural waters: some stability relations in the system Al2O3-SO3-H2O at 298 K. Geochim Cosmochim Acta. 1982;46(4):681–92.
33. Bigham JM, Nordstrom DK. Iron and aluminum hydroxysulfates from acid sulfate waters. Rev Mineral Geochem. 2000;40(1):351–403.
34. Wanner C, Pöthig R, Carrero S, Fernandez-Martinez A, Jäger C, Furrer G. Natural occurrence of nanocrystalline Al-hydroxysulfates: Insights on formation, Al solubility control and As retention. Geochim Cosmochim Acta. 2018;238:252–69.
35. Sánchez-España J, Yusta I, Diez-Ercilla M. Schwertmannite and hydrobasaluminite: A re-evaluation of their solubility and control on the iron and aluminium concentration in acidic pit lakes. Appl Geochem. 2011;26(9–10):1752–74.
36. Takagai Y, Abe R. Concerning the chemical similarities and differences between the blue-coloured waters of Goshikinuma and blue pond in Biei Town, Hokkaido. Fukushima Univ Sci Eng Group Symbiotic Syst. 2014;14:80–7.
37. Du P, Yuan P, Thill A, Annabi-Bergaya F, Liu D, Wang S. Insights into the formation mechanism of imogolite from a full-range observation of its sol-gel growth. Appl Clay Sci. 2017;150:115–24.
38. Doucet FJ, Schneider C, Bones SJ, Kretchmer A, Moss I, Tekely P, Exley C. The formation of hydroxyaluminosilicates of geochemical and biological significance. Geochim Cosmochim Acta. 2001;65(15):2461–7.
39. Strekopytov S, Jarry E, Exley C. Further insight into the mechanism of formation of hydroxyaluminosilicates. Polyhedron. 2006;25(17):3399–404.
40. Beardmore J, Lopez X, Mujika JI, Exley C. What is the mechanism of formation of hydroxyaluminosilicates? Sci Rep. 2016;6(1):1–8.
41. Dobrzyński D. Silica solubility in groundwater from Permian volcanogenic rocks (the Sudetes Mts., SW Poland)–the role of reversible aluminosilicate solids. Geol Quarterly. 2006;50(4):407–17.
42. Dobrzyński D. Chemistry of neutral and alkaline waters with low Al3+ activity against hydroxyaluminosilicate HASB solubility. The evidence from ground and surface waters of the Sudetes Mts. (SW Poland). Aquatic Geochem. 2007;13(3):197–210.
43. Parfitt RL. Allophane in New Zealand-a review. Soil Res. 1990;28(3):343–60.
44. Montarges-Pelletier E, Bogenez S, Pelletier M, Razafitianamaharavo A, Ghanbaja J, Lartiges B, Michot L. Synthetic allophane-like particles: textural properties. Colloids Surf A. 2005;255(1–3):1–10.
45. Lenhardt KR, Breitzke H, Buntkowsky G, Reimhult E, Willinger M, Rennert T. Synthesis of short-range ordered aluminosilicates at ambient conditions. Sci Rep. 2021;11(1):1–13.
46. Caraballo MA, Wanty RB, Verplanck PL, Navarro-Valdivia L, Ayora C, Hochella MF Jr. Aluminum mobility in mildly acidic mine drainage: Interactions between hydrobasaluminite, silica and trace metals from the nano to the meso-scale. Chem Geol. 2019;519:1–10.
47. Sánchez-España J, Yusta I, Burgos WD. Geochemistry of dissolved aluminum at low pH: Hydrobasaluminite formation and interaction with trace metals, silica and microbial cells under anoxic conditions. Chem Geol. 2016;441:124–37.
48. Du X, Wang Y, Su X, Li J. Influences of pH value on the microstructure and phase transformation of aluminum hydroxide. Powder Technol. 2009;192(1):40–6.
49. Hem JD, Roberson CE, Lind CJ, Polxer WL (1973) Chemical interactions of aluminum with aqueous silica at 25 degrees Celsius (No. 1827-E). US Govt. Print. Off.
50. Letterman RD, editor. Water quality and treatment: a handbook of community water supplies. 5th ed. New York: McGraw-Hill; 1999.
51. Faust SD, Aly OM. Chemistry of water treatment. Boca Raton: CRC; 2018.
52. Meier DV, Pjevac P, Bach W, Hourdez S, Girguis PR, Vidoudez C, Amann R, Meyerdierks A. Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. ISME J. 2017;11(7):1545–58.
53. Lopez Bedogni G, Massello FL, Giaveno A, Donati ER, Urbieta MS. A deeper look into the biodiversity of the extremely acidic copahue volcano-río agrio system in Neuquén. Argentina Microorg. 2020;8(1):58.
54. Li J, Zhou H, Peng X, Wu Z, Chen S, Fang J. Microbial diversity and biomineralization in low-temperature hydrothermal iron–silica-rich precipitates of the Lau Basin hydrothermal field. FEMS Microbiol Ecol. 2012;81(1):205–16.
55. Amils R, González-Toril E, Fernández-Remolar D, Gómez F, Aguilera Á, Rodríguez N, Malki M, García-Moyano A, Fairén AG, de la Fuente V, Sanz JL. Extreme environments as Mars terrestrial analogs: The Rio Tinto case. Planetary Space Sci. 2007;55(3):370–81.
56. Arce-Rodríguez A, Puente-Sánchez F, Avendaño R, Libby E, Mora-Amador R, Rojas-Jimenez K, Martínez M, Pieper DH, Chavarría M. Microbial community structure along a horizontal oxygen gradient in a Costa Rican volcanic influenced acid rock drainage system. Microbial Ecol. 2020;80(4):793–808.
57. Niemann H, Linke P, Knittel K, MacPherson E, Boetius A, Brückmann W, Larvik G, Wallmann K, Schacht U, Omoregie E, Rehder G. Methane-carbon flow into the benthic food web at cold seeps–a case study from the Costa Rica subduction zone. PLoS One. 2013;8(10):e74894.
58. Hubalek V, Wu X, Eiler A, Buck M, Heim C, Dopson M, Bertilsson S, Ionescu D. Connectivity to the surface determines diversity patterns in subsurface aquifers of the Fennoscandian shield. ISME J. 2016;10(10):2447–58.
59. Duan JL, Sun JW, Ji MM, Ma Y, Cui ZT, Tian RK, Xu PC, Sun WL, Yuan XZ. Indicatory bacteria and chemical composition related to sulfur distribution in the river-lake systems. Microbiol Res. 2020;236:126453.
60. Salonen IS, Chronopoulou PM, Bird C, Reichart GJ, Koho KA. Enrichment of intracellular sulphur cycle–associated bacteria in intertidal benthic foraminifera revealed by 16S and aprA gene analysis. Sci Rep. 2019;9(1):1–12.
61. Holmes P, Niccolls LM, Sartory DP. The ecology of mesophilic Aeromonas in the aquatic environment. The Genus Aeromonas. 1996; 127–150.
62. Janda JM, Abbott SL. The genus Aeromonas: taxonomy, pathogenicity, and infection. Clin Microbiol Rev. 2010;23(1):35–73.
63. Rippey SR, Cabelli VJ. Use of the thermotolerant Aeromonas group for the trophic state classification of freshwaters. Water Res. 1989;23(9):1107–14.
64. Ali AD, Ezra AG, Abdul SD. Species composition and distribution of freshwater diatoms from Upper Dilimi River, Jos, Nigeria. 2015.
65. Molinero J, Barrado M, Guijarro M, Ortiz M, Carnicer O, Zuazagoitia D. The Teaone River: a snapshot of a tropical river from the coastal region of Ecuador. Limnetica. 2019;38(2):587–605.
66. Haque M, Jewel M, Sayed A, Akhi M, Atique U, Paul AK, et al. Seasonal dynamics of phytoplankton community and functional groups in a tropical river. Environ Monitor Assess. 2021;193(11):1–16.
67. Silva-Benavides AM. The epilithic diatom flora of a pristine and a polluted river in Costa Rica, Central America. Diatom Res. 1996;11(1):105–42.
68. Khan IN. Assessment of water pollution using diatom community structure and species distribution—a case study in a tropical river basin. Internationale Revue der gesamten Hydrobiologie und Hydrographie. 1990;75(3):317–38.
69. De Master DJ, Knapp GB, Nittrouer CA. Biological uptake and accumulation of silica on the Amazon continental shelf. Geochim Cosmochim Acta. 1983;47(10):1713–23.
70. Conley DJ. Riverine contribution of biogenic silica to the oceanic silica budget. Limnol Oceanogr. 1997;42(4):774–7.
71. Ran X, Liu S, Liu J, Zang J, Che H, Ma Y, Wang Y. Composition and variability in the export of biogenic silica in the Changjiang River and the effect of Three Gorges Reservoir. Sci Total Environ. 2016;571:1191–9.
72. Admiraal W, Breugem P, Jacobs DMLHA, Van Steveninck EDR. Fixation of dissolved silicate and sedimentation of biogenic silicate in the lower river Rhine during diatom blooms. Biogeochemistry. 1990;9(2):175–85.
73. Krause JW, Brzezinski MA, Landry MR, Baines SB, Nelson DM, Selph KE, Taylor AG, Twiningf BS. The effects of biogenic silica detritus, zooplankton grazing, and diatom size structure on silicon cycling in the euphotic zone of the eastern equatorial Pacific. Limnol Oceanogr. 2010;55(6):2608–22.

74. Carbonnel V, Vanderborght JP, Lionard M, Chou L. Diatoms, silicic acid and biogenic silica dynamics along the salinity gradient of the Scheldt estuary (Belgium/The Netherlands). Biogeochemistry. 2013;113(1):657–82.
75. Losic D, Mitchell JG, Voelcker NH. Diatom culture media contain extracellular silica nanoparticles which form opalescent films. In: Smart Materials V (Vol. 7267, p. 726712). International Society for Optics and Photonics. 2008.
76. Huang Y, Mi W, Hu Z, Bi Y. Effects of the three Gorges Dam on spatiotemporal distribution of silicon in the tributary: evidence from the Xiangxi River. Environ Sci Pollut Res. 2019;26(5):4645–53.
77. Le Meur M, Montargès-Pelletier E, Bauer A, Gley R, Migot S, Barres O, Delus C, Villiéras F. Characterization of suspended particulate matter in the Moselle River (Lorraine, France): evolution along the course of the river and in different hydrologic regimes. J Soils Sediments. 2016;16(5):1625–42.
78. Kravchishina MD, Lisitzyn AP. Grain-size composition of the suspended particulate matter in the marginal filter of the Severnaya Dvina River. Oceanology. 2011;51(1):89–104.
79. Passow U, Alldredge AL. Aggregation of a diatom bloom in a mesocosm: The role of transparent exopolymer particles (TEP). Deep Sea Res Part II. 1995;42(1):99–109.
80. Hamm CE. Interactive aggregation and sedimentation of diatoms and clay-sized lithogenic material. Limnol Oceanogr. 2002;47(6):1790–5.
81. Gehlen M, Beck L, Calas G, Flank AM, Van Bennekom AJ, Van Beusekom JEE. Unraveling the atomic structure of biogenic silica: evidence of the structural association of Al and Si in diatom frustules. Geochim Cosmochim Acta. 2002;66(9):1601–9.
82. Dixit S, Van Cappellen P. Surface chemistry and reactivity of biogenic silica. Geochim Cosmochim Acta. 2002;66(14):2559–68.
83. Wu B, Liu SM, Ren JL. Dissolution kinetics of biogenic silica and tentative silicon balance in the Yellow Sea. Limnol Oceanogr. 2017;62(4):1512–25.
84. Presti M, Michalopoulos P. Estimating the contribution of the authigenic mineral component to the long-term reactive silica accumulation on the western shelf of the Mississippi River Delta. Cont Shelf Res. 2008;28(6):823–38.
85. Stanton C, Cosmidis J, Kump L. A re-examination of the mechanism of whiting events: A new role for diatoms in Fayetteville Green Lake. New York, USA; 2021.

Publisher's Note