Supplementary Information (SI) for

Amazonian rainforest bioaerosol influenced by African dust intrusion

Jens Weber[1,2], Isabella Hrabe de Angelis[2,3], Sebastian Brill[2], Cybelli G. G. Barbosa[1,4], Stefanie Maier[1], Petya Yordanova[2], Rodrigo Paidano Alves[1], Cleo Quaresma[4], Alessandro Araújo[5], Paulo Artaxo[6], Ricardo H. M. Godoi[7], Ulrich Pöschl[2], Christopher Pöhlker[2], Bettina Weber[1,2]

[1] Institute of Biology, Institute of Plant Sciences, University of Graz, Holteigasse 6, 8010 Graz, Austria

[2] Max Planck Institute for Chemistry, Multiphase Chemistry Department, 55128 Mainz, Germany

[3] OceanX, 37 W 39th St New York, NY 10018

[4] Instituto Federal de Educação Ciência e Tecnologia do Pará, Belem, Brazil

[5] Brazilian Agricultural Research Corporation, Brasilia, Brazil

[6] Institute of Physics, University of São Paulo, 05508-900 São Paulo, Brazil

[7] Environmental Engineering Department, Federal University of Parana, Curitiba PR, Brazil

*e-mail: Jens Weber j.weber@mpic.de; Bettina Weber bettina.weber@uni-graz.at

Contents

Methods

Aerosol sampling. For decontamination prior to sampling, filters were baked in aluminum foil bags (Rotilabo aluminum foil, thickness 30 µm, Carl Roth GmbH + Co.KG, Karlsruhe, Germany) at 300 °C for 8 h. Prior to sampling, filter holders (Digitel Elektronik AG, Volketswil, Switzerland) were sterilized by washing with DNAexitus (AppliChem GmbH, Darmstadt, Germany), residues were rinsed off with Bacillol AF (Hartmann-Gruppe, Heidenheim an der Brenz, Germany) and the holders were airdried in a laminar flow cabinet while irradiated with UV light for 30 min. Tweezers were sterilized in 70 % ethanol and flamed in a Bunsen Burner before they were used to place the filter in the filter holder.

In the field, a handling and starting filter blanks were taken prior to the first sampling and subsequently every 7 days. For the handling blank, the filter within the filter holder was placed in the sampler for approx. 1 second and unloaded again. For the starting blank, the filter within the filter holder was placed in the sampler and the pump was switched on for 5 seconds before the filter was removed again. Additionally, bird feces were collected on the top platform of the tower to evaluate its relevance as a potential contamination source. For the transportation of the filters in the filter holders between the laboratory and the sampler, the filter holder was placed in a sterilized empty aluminum bag (baked for 8 h at 300 °C) using sterile nitrile gloves. After sampling, filters were handled only in the laboratory under a sterile laminar flow cabinet that was decontaminated with DNAexitus and Bacillol AF, followed by UV irradiation for 30 min prior to filter handling. Here, the filters were removed from the filter holder and placed inside the transportation bag. Filters were cut into 16 equally-sized pieces, and in total 2 filter pieces for the 42 m samples and

4 filter pieces for the 323 m samples from opposing sides of the filter were put on top of each other (loaded sides facing towards each other), rolled, and placed in 5 mL tubes, using sterilized tweezers. Then, 2 ml and 4 mL RNAlater (Invitrogen, Carlsbad, USA) were added to the 42 m and 323 m filter pieces, respectively, and the samples were incubated for 30 min at 4 °C. Samples were then stored at -20 °C until transportation to the permanent storage. The samples were transported in a threefold isolated Styrofoam box on dry ice (duration of transportation: 24 h). Afterwards, samples were stored at -80 °C until DNA extraction.

DNA extraction. DNA was extracted using the ChargeSwitch^{TM} Forensic DNA Purification Kit (Invitrogen, Carlsbad, USA). For extraction, two filter pieces were used for samples collected at 42 m and four for samples from 323 m height. The manufacturer instructions were followed with some modifications; prior to DNA extractions, the filter pieces were placed in a silica spinning column (PowerClean® Pro DNA Clean-Up Kit, QIAGEN Company, Hilden, Germany) and spun for 30 s at 4,000 rcf to remove RNAlater residues (ammonia sulfate residues), since ammonia sulfate would interfere with enzymatic lysis. Afterwards, the 400 µl 1x phosphate buffered saline was added and incubated for 1 min at RT. The column was spun again for 30 s at 4,000 rcf to wash off ammonia sulfate residues. Each filter piece was then transferred together with the silica spinning columns silica membrane to an InnuSPEED lysis tube W (Analytik Jena, Jena, Germany) and 1 mL lysis buffer of the ChargeSwitch™ Forensic DNA Purification Kit (Invitrogen Corporation, Waltham, USA) was added. Tubes were placed on a tissue lyser II (QIAGEN Company, Hilden, Germany) for 3 min at a frequency of 30,000 Hz to break down glass fiber, cells and spores. Afterwards, 26 µL of lysozyme (SIGMA-ALDRICH, Darmstadt, Germany) solution (10 mg mL-1) were added, shortly vortexed and incubated at 37 °C for 60 min (with a short vortex after 30 min). After lysozyme treatment, 15 µL of proteinase K solution

(ChargeSwitch™ Forensic DNA Purification Kit, Invitrogen Corporation, Waltham, USA) were added, vortexed, and incubated at 55 °C for 90 min (with a short vortex after 45 min). Afterwards, samples were vortexed for 5 s and inverted thrice. Tubes were centrifuged at 10,000 rcf for debris removal. The supernatant was transferred to a new 2 mL reaction tube. Filter debris was placed on the previously used silica spinning column and centrifuged for 1 min at 4,000 rcf. Supernatant and flowthrough were unified. In a washing step, 400 µL preheated (65 °C) PCR grade water (Biozym, Hessisch Oldendorf, Germany) was added to the spinning column. The column was first incubated for 5 min at 65 °C and afterwards centrifuged for 1 min at 10,000 rcf. Both flowthroughs were unified with the supernatant in the 2 mL reaction tube. Binding (5 min at 4 °C with 25 µL beads) and washing steps were performed according the manufacturer's instructions (ChargeSwitch™ Forensic DNA Purification Kit, Invitrogen Corporation, Waltham, USA). The washing step was performed twice. DNA was eluted in two steps. During the first step, 100 µL elution buffer were added to the beads and incubated at room temperature for 5 min. For the second elution, 100 µL elution buffer were added and the sample was incubated at 65 °C for 5 min. Both eluates were unified and incubated at 60 °C in a heat block shaker (VWR, Radnor, USA) until the volume was reduced to 30-40 µL.

DNA concentration was determined using the Denovix Ultra High Sensitivity Assay (DeNovix Inc., Wilmington, USA) on a Qubit 3 photometer (Invitrogen Corporation, Waltham, USA). For determination of the DNA concentration, 1 µL of extract was diluted in 9 µL of PCR grade water. The remaining sample was stored at -80 °C until sequencing. For DNA yield see SI Table S1 (eDNA concentration).

Bioinformatics. After demultiplexing, sequences were processed using DADA2 1.26[1]. For the prokaryotic reads, primer sequences were removed using cutadap[2] followed by trimming the

forward and reverse reads to 266 and 127 using the filterAndTrim function of dada2 to avoid the technical quality drop at the end of the reads. The maxEE input was 3 for forward and 6 for reverse reads. Reads were merged with a minimum overlap of 12 bases with the mergePairs function. Merged reads were controlled for chimeras with removeBimeraDenovo and classified against the SILVA training set v. 138[3].

For the fungal community, primer sequences were removed using cutadapt followed by quality control using the filterAndTrim function of dada2 ((maxN=0, truncQ=2, maxEE=(2,2), minLen=100, trimRight=(20,80)). Afterwards, reads were merged with a minimum overlap of 15 bases with the mergePairs function. Merged reads were controlled for chimeras by removeBimeradenovo and subsequently reads were classified using the UNITE 9.0 general FASTA release training set[4].

Before removal of potential contaminants, single- and doubletons were removed. Reads which were identified as potential contaminations by the decontam package[5] were removed. For the prokaryotic community, potential contaminants identified in the controls (AB, B21, A21) were removed from samples using the "either" method (reads identified either by frequency or prevalence were removed). For the fungal community, potential contaminants were removed using the frequency method. Afterwards, to extract the bacterial community from the prokaryotic sequences, mitochondrial, chloroplast and archaeal reads were excluded and sequencing depth was rarefied to an even depth of 7,627 reads per sample, using the rarefy_even_depth function of the PHYLOSEQ package v. 1.42.0[6]. For the fungal community the decontaminated reads were rarefied to an even depth of 20,022 reads.

Airborne endotoxin concentration. The filter pieces (one per filter) were rolled and placed into a 5 mL Eppendorf tube with the loaded side facing inwards and 5 mL of 0.05 % Tween 20 (Sigma Aldrich, St. Louis, USA) in endotoxin free water (Sigma Aldrich, St. Louis, USA) were added. The tubes were placed on a heat block shaker (VWR, Radnor, USA) at 15,000 Hz for 1 h and then frozen at -20 °C for 48 h to further break down cells. Subsequently, the tubes were thawed, vortexed to mix, and stored in 1.5 mL aliquots. For the endotoxin estimation, the kinetic chromogenic limulus amebocyte lysate method (Pierce LAL chromogenic endotoxin quantification kit, Thermo Fisher Scientific, Waltham, USA) was used and optical density (OD) was measured with a microplate reader.

Concentrations of endotoxins were analyzed in 1:10 diluted samples and blanks according to the manufacturer´s protocols, with an initial incubation of samples at 70 °C for 10 min to inactivate potential interferents. The optical density (OD) of 50 µL sample extract plus 200 µL stop solution (CB) was measured to make up for the naturally occurring coloration of the samples. Every measurement was performed in duplicates. All ODs were measured at a wavelength of 405 nm.

To calculate the endotoxin concentrations, the mean values of the OD of the duplicates of all samples and blanks were calculated. Then the CB was subtracted from ODs of the samples and the blank. Afterwards, the blank ODs were subtracted from the sample ODs and sample extract endotoxin concentrations were calculated using the standard dilution series made from the endotoxin standard supplied by the test kit. Endotoxin concentrations were expressed as endotoxin unit per m^3 (EU m^{-3}).

Backward trajectories calculation. The Hybrid Single-Particle Lagrangian Integrated Trajectory Model (HYSPLIT) of the National Oceanic and Atmospheric Administration (NOAA)[56] was used to calculate backward trajectories (BT). Therefore, the gridded meteorological

data of the U.S. National Center for Environmental Prediction by the Air Resources Laboratory (NCEP) Global Data Assimilation System (GDAS) model with 1° resolution was used. For each filter sample collected at 42 m and 323 m, a 240 h BT starting at 42 m and 323 m above ground level at the coordinates -59°00.2′W -02°08.9′S were calculated. For each filter sample, hourly backward trajectories were calculated for the duration of sampling.

Analysis and statistics. Samples were classified as dust influenced based on coarse mode mass concentration (M_{CM}), supported by elevated black carbon (BC) concentrations, BT analysis, and filter coloration. African dust plumes reaching ATTO are known to consist of a mixture of African dust and smoke aerosols since main biomass burning activity and dust season are happening simultaneously in Africa. For M_{CM} as a criterium for the presence of dust events we followed Moran-Zuloaga et al. 2018[7], with the following rules: (1) per filter averaged M_{CM} need to exceed 9 μg m^{-3}, (2) time series of M_{CM} need to show peak-like behavior with the whole event exceeding one day. Additionally, dust influenced filters showed a distinct coloration with grayish colors during forest background and brownish tones during dust influence. For meteorological conditions, M_{CM}, and BC see SI Fig. S1. M_{CM} was calculated as outlined in Moran-Zuloaga.[7] Backwards trajectories also support the source region of the dust event (SI Fig. S4 and S5). However, they need to be taken with caution because of the high standard deviation and uncertainty caused by the long run-time of 10 days. With this method we detected a five-day spanning dust event from 18. Feb to 22. Feb 2020. Every filter within this period is considered dust influenced, whereas filters outside of this period are classified as forest background. For detailed sample classification based on these criteria see SI table S1.

For height and day to day comparisons during background conditions, a dataset comprising 16 filters each taken at 42 m and 323 m was used. All of them (Sample ID: A04 - A09, A12, A19,

A22 - A28, A31, B04 - B09, B12, B19, B22 - B28, and B31) were used for the bacterial community while two (A23 and B31) were excluded from the fungal community analysis due to failed sequencing reaction. For comparison of dust and background conditions, five filters each collected at 42 m and 323 m on five dust event days (17 Feb 2020 to 21 Feb 2020; sample ID: A13, A14, A16 - A18, B13, B14, B16 - B18; for fungi, samples B14 and 16 could not be used) and on dust free days prior to the dust event (12 Feb 2020 to 16 Feb 2020; sample IDs: A06 - A09, A12, B06 - B09, and B12) were used.

Statistical analyses were performed in the software environment R v. 4.2.2[8]. Relative taxa abundance and alpha diversity metrics (e.g., Observed ASVs and Simpson evenness) were calculated using the PHYLOSEQ package v. 1.42.0[6]. To test for significant differences between sampling heights, as well as dust vs background condition, a t-test was used, using the STATS package v. 4.2.0[8]. Beta diversity of the community was calculated by means of Bray-Curtis dissimilarity and plotted as an unconstrained ordination plot containing environmental variables as vectors using the package vegan v.2.6-4[9]. For the unconstrained ordination plot, non-metric multidimensional scaling (NMDS) was selected.

SI Tables (S1-S2)

Table S1: Sampling details for individual total suspended particle samples. Dust Condition: Grouping of samples as defined in the analysis and statistics section, bg: background. Coarse mode is here defined as particles larger 1 μm. eDNA: environmental DNA. b.d.: Below detection limit.

Sample ID	Sampling start (UTC) [dd-mm-yyyy hh:mm]	Sampling end (UTC) [dd-mm-yyyy hh:mm]	Dust Condition	Sampled volume [m^3]	Samling height [m]	Coarse mode mass concentration 1-10 µm [$\mu g\ m^{-3}$]	Isolated eDNA amount [ng]	Sequenced
A04	10 Feb 2020 14:43	11 Feb 2020 14:43	42 m bg	720	42	6.09	59.9	yes
A05	11 Feb 2020 14:43	12 Feb 2020 14:43	42 m bg	720	42	6.03	76.1	Yes
A06	12 Feb 2020 14:43	13 Feb 2020 14:43	42 m bg	720	42	2.26	119.8	Yes
A07	13 Feb 2020 14:43	14 Feb 2020 14:43	42 m bg	720	42	1.16	93.4	Yes
A08	14 Feb 2020 14:43	15 Feb 2020 14:43	42 m bg	720	42	2.24	160.6	Yes
A09	15 Feb 2020 14:43	16 Feb 2020 14:43	42 m bg	720	42	2.92	353.5	Yes
A12	16 Feb 2020 15:33	17 Feb 2020 15:33	42 m bg	720	42	2.28	693.3	Yes
A13	17 Feb 2020 21:26	18 Feb 2020 15:46	42 m dust	550	42	11.13	156.5	Yes
A14	18 Feb 2020 15:46	19 Feb 2020 15:46	42 m dust	720	42	27.78	58	Yes
A16	19 Feb 2020 15:46	20 Feb 2020 15:46	42 m dust	720	42	32.64	109.6	Yes
A17	20 Feb 2020 15:46	21 Feb 2020 15:46	42 m dust	720	42	23.87	100.8	Yes
A18	21 Feb 2020 15:46	22 Feb 2020 15:46	42 m dust	720	42	15.26	138.3	Yes
A19	22 Feb 2020 15:50	23 Feb 2020 15:50	42 m bg	720	42	7.06	453.8	Yes
A22	23 Feb 2020 16:14	24 Feb 2020 16:11	42 m bg	718.5	42	2.17	485.1	Yes
A23	24 Feb 2020 16:18	25 Feb 2020 15:20	42 m bg	1382	42	1.29	82.0	Yes
A24	25 Feb 2020 15:33	26 Feb 2020 15:33	42 m bg	720	42	0.94	288.5	Yes
A25	26 Feb 2020 15:33	27 Feb 2020 15:33	42 m bg	720	42	1.51	440.9	Yes
A26	27 Feb 2020 15:33	28 Feb 2020 15:33	42 m bg	720	42	2.82	388.3	Yes
A27	28 Feb 2020 15:33	29 Feb 2020 15:33	42 m bg	720	42	2.46	497.7	Yes
A28	29 Feb 2020 15:33	01 Mar 2020 15:33	42 m bg	720	42	1.76	94.4	Yes
A31	01 Mar 2020 19:00	02 Mar 2020 19:00	42 m bg	720	42	1.00	166.9	Yes

B04	10 Feb 2020 13:47	11 Feb 202013:47	323 m bg	1,440	321	5.25	137.2	Yes
B05	11 Feb 202013:47	12 Feb 202013:47	323 m bg	1,440	321	2.44	111.5	Yes
B06	12 Feb 202013:47	13 Feb 202013:47	323 m bg	1,440	321	1.11	404.3	Yes
B07	13 Feb 2020 13:47	14 Feb 2020 13:47	323 m bg	1,440	321	1.30	419.6	Yes
B08	14 Feb 2020 13:47	15 Feb 2020 13:47	323 m bg	1,440	321	1.48	157.9	Yes
B09	15 Feb 2020 13:47	16 Feb 2020 13:47	323 m bg	1,440	321	3.11	314	Yes
B12	16 Feb 2020 14:19	17 Feb 2020 14:19	323 m bg	1,440	321	2.84	106.59	Yes
B13	17 Feb 2020 19:35	18 Feb 2020 15:15	323 m dust	1,180	321	9.16	44.7	Yes
B14	18 Feb 2020 15:18	19 Feb 2020 13:54	323 m dust	1,156	321	23.35	123.3	Yes
B16	19 Feb 2020 13:54	20 Feb 2020 13:54	323 m dust	1,440	321	25.37	124.5	Yes
B17	20 Feb 2020 13:54	21 Feb 2020 13:54	323 m dust	1,440	321	21.42	161.1	Yes
B18	21 Feb 2020 13:54	22 Feb 2020 13:54	323 m dust	1,440	321	13.52	167.0	Yes
B19	22 Feb 2020 14:30	23 Feb 2020 14:30	323 m bg	1,440	321	6.39	241.1	Yes
B22	23 Feb 2020 15:30	24 Feb 2020 15:30	323 m bg	1,440	321	1.30	209.3	Yes
B23	24 Feb 2020 15:30	25 Feb 2020 15:30	323 m bg	1,440	321	1.24	51.3	Yes
B24	25 Feb 2020 15:30	26 Feb 2020 15:30	323 m bg	1,440	321	0.96	131.5	Yes
B25	26 Feb 2020 15:30	27 Feb 2020 15:30	323 m bg	1,440	321	1.34	456.2	Yes
B26	27 Feb 2020 15:30	28 Feb 2020 15:30	323 m bg	1,440	321	2.79	112.3	Yes
B27	28 Feb 2020 15:30	29 Feb 2020 15:30	323 m bg	1,440	321	2.47	194.0	Yes
B28	29 Feb 2020 15:30	01 Mar 2020 15:30	323 m bg	1,440	321	1.75	208.7	Yes
B31	01 Mar 2020 19:16	02 Mar 2020 19:16	323 m bg	1,440	321	1.15	104.4	Yes
AB	23 Feb 2020 15:10	23 Feb 2020 15:10	Bird feces 323m	-	-	-	95.0	Yes

B21	23 Feb 2020 15:20	23 Feb 2020 15:20	Start-up blank 42 m	-	-	-	b.d.	Yes
A21	23 Feb 2020 16:12	23 Feb 2020 16:12	Start-up blank 323 m	-	-	-	b.d.	Yes
A01	09 Feb 2020 14:35	09 Feb 2020 14:35	handling blank 42 m	-	-	-	b.d.	No
A02	09 Feb 2020 14:40	09 Feb 2020 14:40	Start-up blank 42 m	-	-	-	b.d.	No
A10	16 Feb 2020 15:13	16 Feb 2020 15:13	handling blank 42 m	-	-	-	b.d.	No
A11	16 Feb 2020 15:14	16 Feb 2020 15:14	Start-up blank 42 m	-	-	-	b.d.	No
A20	23 Feb 2020 17:11	23 Feb 2020 17:11	handling blank 42 m	-	-	-	b.d.	No
A29	01 Mar 2020 17:40	01 Mar 2020 17:40	handling blank 42 m	-	-	-	b.d.	No
A30	01 Mar 2020 17:42	01 Mar 2020 17:42	Start-up blank 42 m	-	-	-	b.d.	No
B01	09 Feb 2020 13:40	09 Feb 2020 13:40	handling blank 323 m	-	-	-	b.d.	No
B02	09 Feb 2020 13:44	09 Feb 2020 13:44	Start-up blank 323 m	-	-	-	b.d.	No
B10	16 Feb 2020 13:51	16 Feb 2020 13:51	handling blank 323 m	-	-	-	b.d.	No
B11	16 Feb 2020 13:51	16 Feb 2020 13:51	Start-up blank 323 m	-	-	-	b.d.	No

B20	23 Feb 2020 15:20	23 Feb 2020 15:20	handling blank 323 m	-	-	-	b.d.	No
B29	01 Mar 2020 19:10	01 Mar 2020 19:10	handling blank 323 m	-	-	-	b.d.	No
B30	01 Mar 2020 19:12	01 Mar 2020 19:12	Start-up blank 323 m	-	-	-	b.d.	No

Table S2: Estimated richness (observed ASVs) and evenness (Simpson evenness) for background (bg) and dust condition (dust) at both heights for the bacterial and fungal community.

Sample ID	Dust Condition	Bacterial Richness (observed ASVs)	Bacterial Simpson evenness (%)	Fungal Richness (observed ASVs)	Fungal Simpson evenness (%)
A04	42 m bg	546	44.2	668	12.5
A05	42 m bg	555	41.7	566	17.1
A06	42 m bg	527	25.1	503	17.4
A07	42 m bg	725	39.0	539	15.9
A08	42 m bg	775	24.9	458	19.4
A09	42 m bg	540	22.7	708	14.2
A12	42 m bg	472	27.7	464	16.5
A13	42 m dust	607	25.1	582	18.9
A14	42 m dust	263	13.2	493	14.8
A16	42 m dust	494	12.5	605	15.5
A17	42 m dust	316	9.3	544	13.2
A18	42 m dust	645	20.2	546	13.7
A19	42 m bg	548	32.2	817	19.5

A22	42 m bg	728	35.3	856	17.3
A23	42 m bg	499	27.0	NA	NA
A24	42 m bg	331	18.5	752	16.5
A25	42 m bg	667	34.6	845	17.4
A26	42 m bg	493	13.5	708	11.6
A27	42 m bg	646	23.2	725	9.9
A28	42 m bg	346	25.6	581	11.2
A31	42 m bg	628	21.0	725	11.9
B04	323 m bg	485	36.9	788	7.9
B05	323 m bg	576	42.5	838	8.2
B06	323 m bg	368	20.7	638	10.5
B07	323 m bg	467	23.9	738	12.1
B08	323 m bg	295	8.4	673	7.9
B09	323 m bg	477	10.2	938	8.4
B12	323 m bg	490	13.1	821	7.2
B13	323 m dust	324	16.2	843	7.5
B14	323 m dust	265	5.9	NA	NA
B16	323 m dust	364	10.8	NA	NA
B17	323 m dust	917	5.4	978	8.6
B18	323 m dust	1021	13.5	567	7.9
B19	323 m bg	299	2.8	778	12.8
B22	323 m bg	427	19.1	814	13.1
B23	323 m bg	310	52.2	718	13.6
B24	323 m bg	311	15.5	780	10.5
B25	323 m bg	361	18.1	725	10.3
B26	323 m bg	574	15.3	914	9.8
B27	323 m bg	448	9.7	871	9.3

B28	323 m bg	648	12.5	644	9.4
B31	323 m bg	314	8.6	NA	NA

Table S3: Pairwise Bray-Curtis dissimilarities between consecutive samples collected at 42 m and 323 m.

Sample ID	Bacterial Bray-Curtis dissimilarities [%]	Fungal Bray-Curtis dissimilarities [%]
A04- A05	71.16	30.77
A05- A06	72.27	29.51
A06- A07	72.95	26.34
A07- A08	71.04	24.81
A08- A09	69.70	30.24
A09- A12	71.72	28.05
A19- A22	70.28	33.66
A22- A23	79.70	NA
A23- A24	76.78	NA
A24- A25	76.78	27.74
A25- A26	79.07	40.51
A26- A27	72.90	28.50
A27- A28	84.44	25.61
A28- A31	85.38	31.98
B04-B05	78.54	33.69
B05-B06	79.63	38.91
B06-B07	85.68	29.37
B07-B08	87.02	29.13
B08-B09	73.93	30.93

B09-B12	78.35	26.72
B19-B22	91.06	44.83
B22-B23	93.20	43.34
B23-B24	96.01	37.13
B24-B25	82.96	26.72
B25-B26	72.48	32.34
B26-B27	82.39	34.32
B27-B28	86.12	35.91
B28-B31	87.64	NA

Table S4: Pairwise Bray-Curtis dissimilarities between samples collected at the same time at 42 m and 323 m.

Sample ID	Dust Condition	Bacterial Bray-Curtis dissimilarities [%]	Fungal Bray-Curtis dissimilarities [%]
A04-B04	42 m bg	79.49	35.83
A05-B05	42 m bg	82.21	43.87
A06-B06	42 m bg	86.36	34.59
A07-B07	42 m bg	82.81	35.68
A08-B08	42 m bg	81.70	45.12
A09-B09	42 m bg	77.88	34.80
A12-B12	42 m bg	84.94	38.75
A13-B13	42 m dust	81.38	49.55
A14-B14	42 m dust	42.93	NA
A16-B16	42 m dust	59.74	NA
A17-B17	42 m dust	57.61	36.03
A18-B18	42 m dust	61.94	35.64

A19-B19	42 m bg	88.11	35.91
A22-B22	42 m bg	91.28	47.12
A23-B23	42 m bg	96.25	NA
A24-B24	42 m bg	86.01	39.21
A25-B25	42 m bg	86.86	44.32
A26-B26	42 m bg	77.30	33.79
A27-B27	42 m bg	87.87	24.73
A28-B28	42 m bg	97.10	35.32
A31-B31	42 m bg	87.58	NA

Table S5: Aerial endotoxin concentrations at 323 m per filter.

Sample ID	Aerial Endotoxin Concentration [EU m^{-3}]
B04	0,31
B05	0,53
B06	0,29
B07	0,56
B08	0,36
B09	0,35
B12	0,54
B13	0,55
B14	1,28
B16	2,10
B17	1,34
B18	1,08
B19	0,52
B22	0,70
B23	0,57
B24	0,26
B25	0,46

Figures (S1-S5)

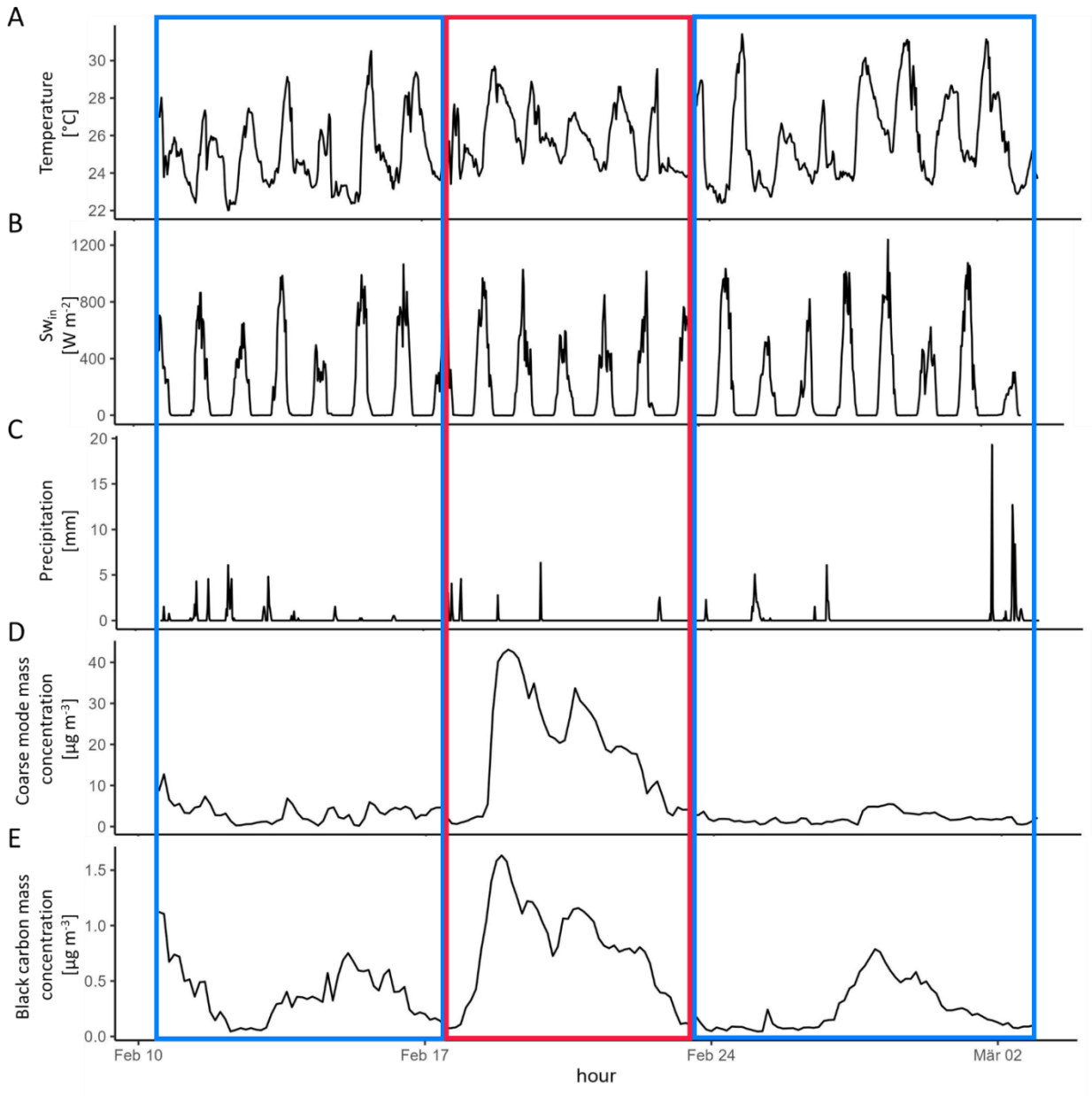

Figure S1. Local climatic conditions during filter sampling from 10 Feb to 02 Mar 2020. (A) Temperature measured at 81 m. (B) Shortwave radiation (SW_{in}) at 81 m. (C) Precipitation at 81 m. (D) Estimated coarse mode mass concentrations (size range: 1 to 10 µm). (E) Estimated black carbon mass concentrations. Time spans when filter samples were collected under clean conditions are framed in blue, whereas dust influenced time spans are framed in red.

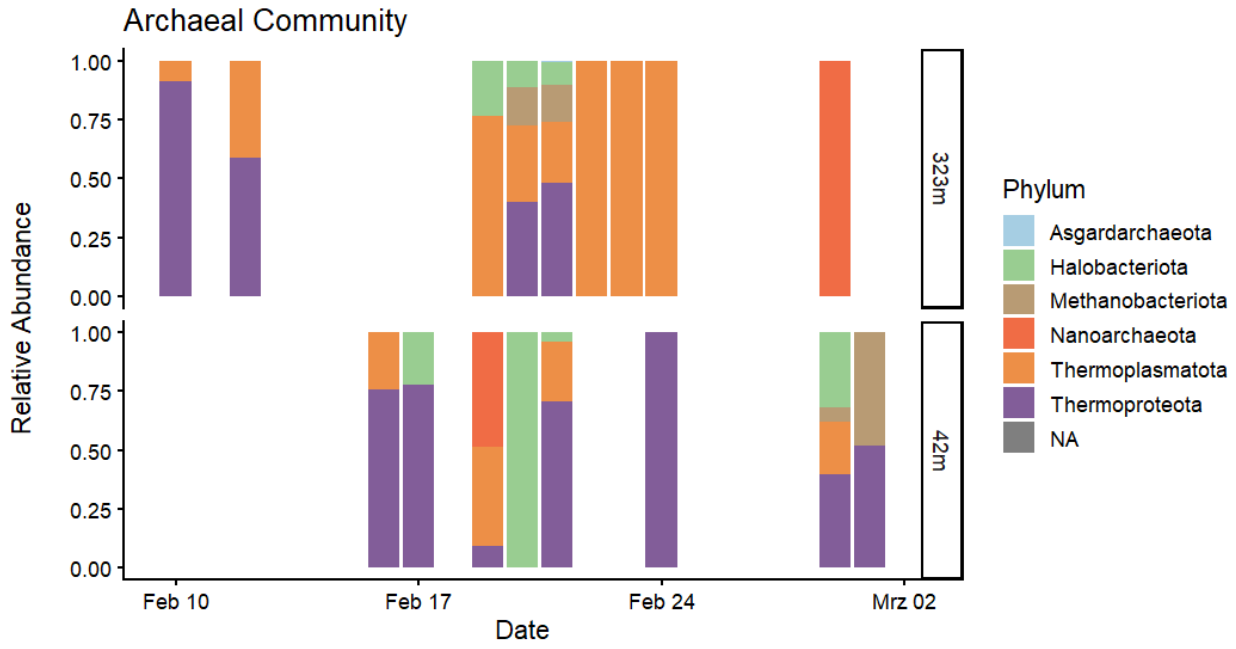

Figure S2. Stacked barcharts of the relative sequence abundance of the archaeal bioaerosol community on phylum level at 42 m (bottom) and 323 m height (top panel) at daily resolution from the 02 Feb 2020 to 02 Mar 2020. Samples with no archaeal reads detected appear as empty barcharts.

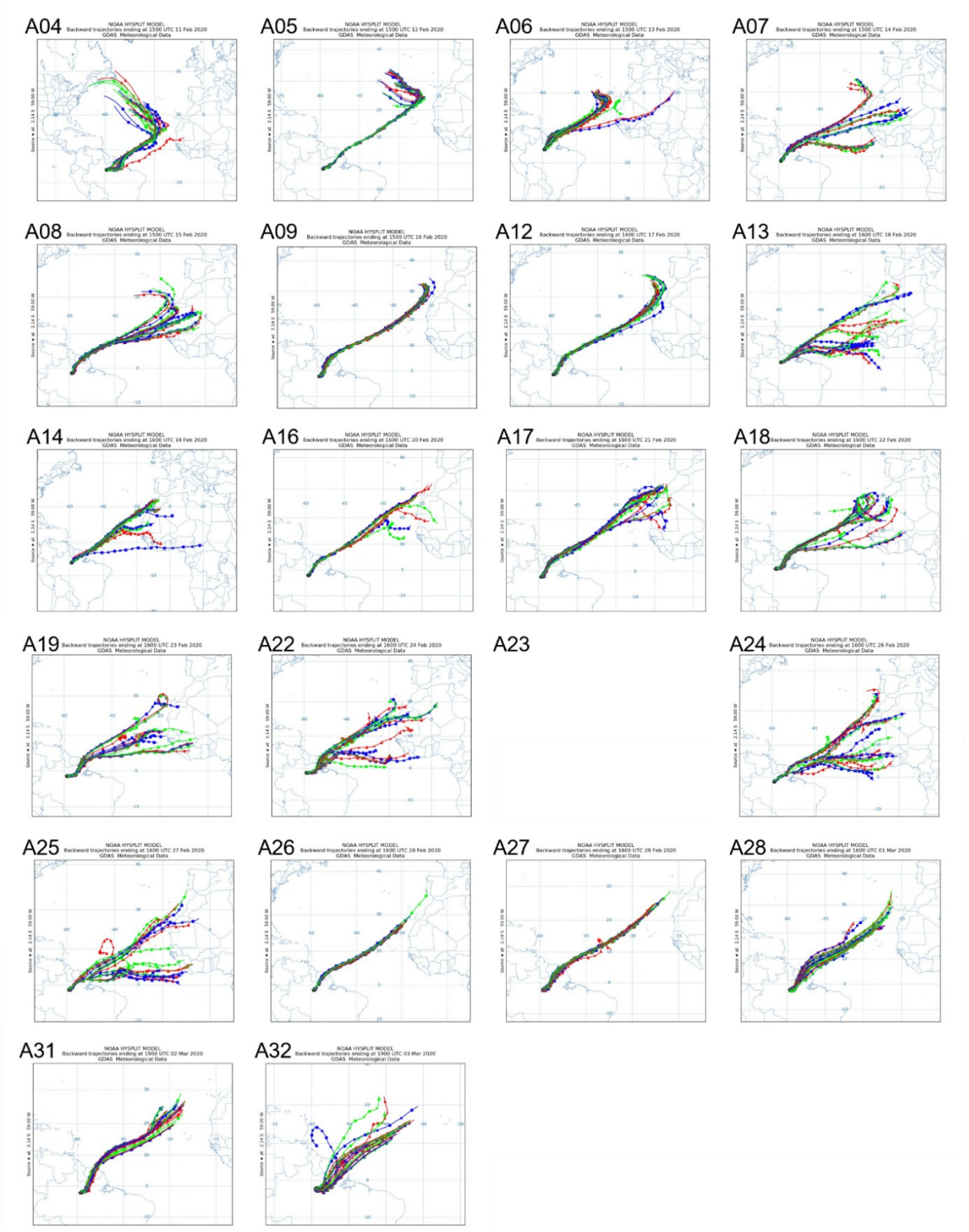

Figure S3. Backward trajectories of filter samples taken at 42 m height. Calculated for 240 h prior to sampling and starting every hour of the sampling duration. Start height 42 m. For calculation

the Hybrid Single-Particle Lagrangian Integrated Trajectory Model of the National Oceanic and Atmospheric Administration was used. As gridded meteorological data, the U.S. National Center for Environmental Prediction by the Air Resources Laboratory Global Data Assimilation System model with 1° resolution was used.

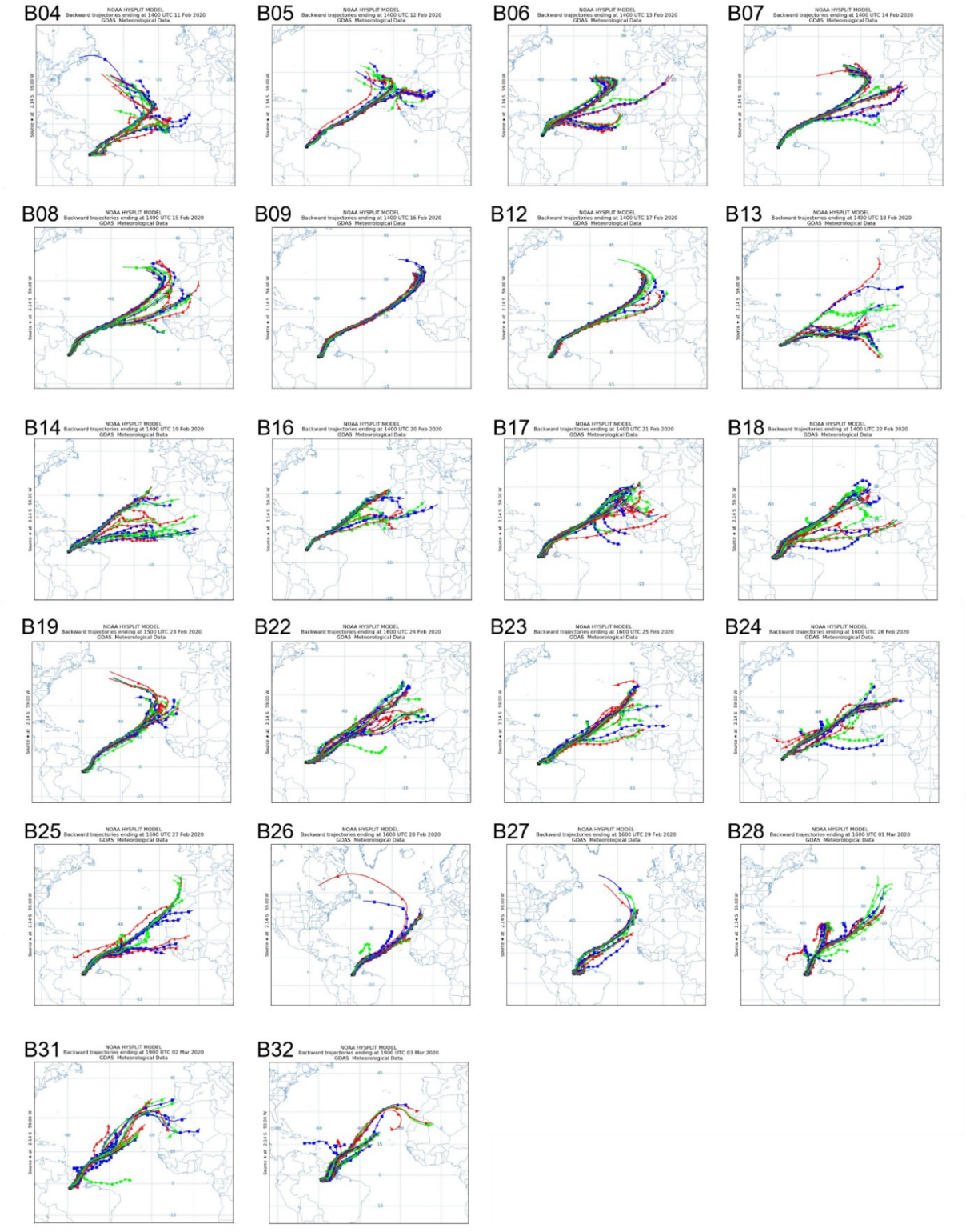

Figure S4. Backward trajectories of filter samples taken at 323 m height. Calculated for 240 h prior to sampling and starting every hour of the sampling duration. Start height 42 m. For

calculation, the Hybrid Single-Particle Lagrangian Integrated Trajectory Model of the National Oceanic and Atmospheric Administration was used. As gridded meteorological data the U.S. National Center for Environmental Prediction by the Air Resources Laboratory Global Data Assimilation System model with 1° resolution was used.

REFERENCES

1. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
2. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
3. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
4. Abarenkov, K. *et al.* The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. *Nucleic Acids Res.* **52**, D791–D797 (2024).
5. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Preprint at https://doi.org/10.1101/221499 (2017).
6. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
7. Moran-Zuloaga, D. *et al.* Long-term study on coarse mode aerosols in the Amazon rain forest with the frequent intrusion of Saharan dust plumes. *Atmospheric Chem. Phys.* **18**, 10055–10088 (2018).
8. R Core Team. R: Language and Environment for Statistical Computing. (2025).

9. Oksanen, J. *et al.* vegan: Community Ecology Package. 2.7-2 https://doi.org/10.32614/CRAN.package.vegan (2001).