FEMS Microbiology Letters, 367, 2020, fnaa051

https://doi.org/10.1093/femsle/fnaa051
Advance Access Publication Date: 19 March 2020
Research Letter

RESEARCH LETTER – Environmental Microbiology

Metagenomic investigation of African dust events in the Caribbean

Samantha Marie Waters[1,*], S. K. Purdue[2], R. Armstrong[3] and Y. Detrés[3]

[1]Universities Space Research Association, Space Biosciences Division, Ames Research Center, Moffett Field, CA 94035, USA, [2]Atmospheric Science, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL 33149, USA and [3]NOAA Center for Atmospheric Sciences (NCAS), Department of Marine Sciences, University of Puerto Rico-Mayaguez, Puerto Rico 00682, USA

*Corresponding author: Universities Space Research Association, Space Biosciences Division, NASA Ames Research Center, N-239, B037, Moffett Field, CA 94035, USA. Tel: +1-650-604-0077; E-mail: samantha.m.waters@nasa.gov

One sentence summary: Metagenomic investigations of Saharan dust events show biological changes from the source environment, the Sahara, after intercontinental transit and these changes may result in dust events being biologically similar year to year.

Editor: Sergio Casella

ABSTRACT

African dust from the Sahara and Sahel regions of Northern Africa is blown intercontinental distances and is the highest portion of atmospheric dust generated each year. During the Northern Hemisphere summer months (boreal summer), these dust events travel into the Caribbean and southern United States. While viability assays, microscopy and bacterial amplicon analyses have shown that dust-associated microbes may be diverse, the specific microbial taxa that are transported intercontinental distances with these dust events remain poorly characterized. To provide new insights into these issues, five metagenomes of Saharan dust events occurring in the Caribbean, collected in the summer months of 2002 and 2008, were analyzed. The data revealed that similar microbial composition existed between three out of the five of the distinct dust events and that fungi were a prominent feature of the metagenomes compared to other environmental samples. These results have implications for better understanding of microbial transport through the atmosphere and may implicate that the dust-associated microbial load transiting the Atlantic with Saharan dust is similar from year to year.

Keywords: Saharan dust; fungi; metagenomes; aerobiology; bioaerosols; Caribbean

INTRODUCTION

It has been known for more than a century that microorganisms are present in the air (Morris *et al.* 2011; Després *et al.* 2012). Mechanisms that act to aerosolize microbes include, but are not limited to, bubble bursting at liquid–air interfaces, release from plant surfaces when heated (heat advection) and wind (e.g. storms). Once aerosolized, microorganisms are subjected to physical processes limiting distance traveled (wet and dry deposition and size selection) and acting on viability (high desiccation, ultraviolet [UV] exposure, etc.) (Lighthart and Shaffer 1994; Jones and Harrison 2004; Wilkinson *et al.* 2012; Choudoir *et al.* 2018; Caro *et al.* 2019). Despite the variety of parameters acting against microbial transport through the atmosphere, it is known that viable microorganisms do travel intercontinental distances (Griffin *et al.* 2003; Prospero *et al.* 2005; Smith *et al.* 2011; Stres *et al.* 2013).

Viable microorganisms are associated with transiting dust plumes (Prospero *et al.* 2005; Griffin 2007; Perfumo and Marchant 2010; Creamean *et al.* 2013; Favet *et al.* 2013; Giongo *et al.* 2013; Vila-Costa *et al.* 2013; Barberán *et al.* 2015; Itani and Smith 2016). The dust plumes that originate in the Sahara and Sahel regions of Northern Africa are the largest contributors to atmospheric dust on the planet. Each year, these dust plumes travel intercontinental distances into Europe, Asia and the Americas. During the Northern Hemisphere's summer (boreal summer), Saharan

Received: 3 October 2019; **Accepted:** 17 March 2020

dust events (SDEs) travel into the Caribbean and southeastern United States (Prospero *et al.* 1970; Moulin *et al.* 1997; Prospero 1999; Prospero and Mayol-Bracero 2013). Research has shown that viable fungi and bacteria are associated with SDEs after travel across the Atlantic Ocean (Griffin *et al.* 2001, 2003; Prospero *et al.* 2005).

Despite decades of research on microorganisms sampled from the atmosphere, only a few culture-independent, shotgun metagenomic studies exist on airborne microbial populations (Be *et al.* 2014; Cao *et al.* 2014) and both of these works were conducted on urban bioaerosol samples. Previous works on the taxonomic analysis of microbes associated with African dust events occurring in the American, European and Asian continents have used clone and ribonsomal ribonucleic acid (rRNA) amplicon libraries to look at bacterial, and sometimes fungal, groups (Jeon *et al.* 2011; Katra *et al.* 2014; Meola, Lazzaro and Zeyer 2014; Itani and Smith 2016; Mazar *et al.* 2016; Yamaguchi *et al.* 2016; Cáliz *et al.* 2018). To the best of our knowledge, there are currently no shotgun metagenome studies of how long-distance transport into the Caribbean may influence the microbial composition of SDEs. This study investigated the question of how dust-associated microbial populations may have changed from their source environment (Schuerger *et al.* 2018) by investigating five SDEs that were sampled in the Caribbean during the summers of 2002 and 2008 (Table 1).

MATERIALS AND METHODS

Dust sample collection and air mass characterization analysis

Samples analyzed were collected at various sites in collaboration with the NOAA Center for Atmospheric Sciences at the University of Puerto Rico at Mayaguez. Since aerosols of African origin are transported by the trade winds across the Atlantic Ocean and impact the Eastern Caribbean during the summer months, the samples analyzed were collected from May to July of 2002 and 2008 using air samplers. Samples collected and the associated metadata are summarized in Table 1 (Saharan Dust-Puerto Rico, SDPR) . Briefly, samples from 2002 were collected in the east coast of Dominica and at Cabezas de San Juan light house in Fajardo (northeastern coast), Puerto Rico. Other samples were collected at Isla Magueyes Field Station located at La Parguera (Southwestern Puerto Rico) in 2008. These events were sampled across multiple days using a Partisol Plus 2025 (Rupprecht and Patashnick Co., Inc., New York) sequential multifilter air sampler and a RAAS 2.5-200 Audit Sampler (Andersen Instruments, Smyrna, GA, USA) at a standard flow of 16.7 liters per minute. Particulate matter was collected on Whatman Polytetrafluoroethylene (PTFE) membrane filters (GE Healthcare UK, Buckinghamshire, United Kingdom) with diameter of 46.2 mm and pore size of 2 μm. Samples were covered with stainless steel caps and placed in filter cassette magazines to ensure sample integrity and protect them during transport. The filtration devices were wiped down with sterile, dry wipes before filter installation to control contamination. The filters were kept in freezer until shipped to Georgia Institute of Technology, where they were immediately put into a −80°C freezer until processed.

The transport of airborne particles from Africa was supported with satellite remote sensing products, trajectory models and ground-based data. Initial determination of an SDE used the Navy Aerosol Analysis and Prediction System global aerosol model (http://www.nrlmry.navy.mil/aerosol/).

Table 1. Filters and corresponding sample metadata used for metagenomes.

	Filters included in metagenomic paper								
Metagenome ID	Sample ID	Collection year	Month	Date	$\mu g/m^3$	Description	Location	Filter	Instrument	Flow rate
SDPR-002	4049690	2008	May	9,10,11	n/a	Surface Air/PM 2.5	Lajas, PR	PM 2.5 PTFE	RAAS 2.5-200 Audit Sampler	16.7 L/m
SDPR-003	4049695	2008	July	7,8,9	n/a	Surface Air/PM10	Lajas, PR	PM10 PTFE	RAAS 2.5-200 Audit Sampler	16.7 L/m
SDPR-004	1006449	2002	May	20,21,22,23	16.81	Surface Air/PM 2.5	Dominica	PM2.5 PTFE	Partisol Plus Model 2025 (Rupprecht and Patashnick Co., Inc.)	16.7 L/m
SDPR-005	1036388	2002	June	16,17,18,19	7.15	Surface Air/PM 2.5	Dominica	PM2.5 PTFE	Partisol Plus Model 2025 (Rupprecht and Patashnick Co., Inc.)	16.7 L/m
SDPR-008	1036367	2002	May–June	29,30,31,1	27.69	Surface Air/PM 10	Fajardo, PR	PM10 PTFE	Partisol Plus Model 2025 (Rupprecht and Patashnick Co., Inc.)	16.7 L/m

This system operates in near real-time to predict the distribution of tropospheric aerosols. MODIS imagery (NASA GSFC) provided daily and weekly mean mapped τ_{865} (aerosol optical thickness from the 865-nm channel) level 3 images at 4-km spatial resolution, while NOAA's HYSPLIT (Hybrid Single Particle Lagrangian Integrated Trajectory) backward trajectory model was utilized for tracking air mass sources (Figure S1, Supporting Information; Fig. 1). The AErosol Robotic NETwork (AERONET, NASA GSFC) stations located in the island of Guadeloupe, Isla Magueyes Field Station and Cabezas de San Juan sites provided aerosol optical thickness data (degree to which aerosols prevent the transmission of light by absorption or scattering of light), which was indicative of dust.

HYSPLIT back trajectories were obtained using reanalysis data on the READY website, http://www.ready.noaa.gov (Stein *et al.* 2015). Ground-level height at the different sampling sites were used as input for starting points in the model, and trajectories were calculated back for 5–10 days prior to sample times. The HYSPLIT data for each event was then downloaded and plotted with respect to longitude against both latitude and height (Draxler 1988, 1992) (Fig. 1). Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data were then obtained from the NASA Langley Research Center Atmospheric Science Data Center and used to confirm dust was present at the starting and end points of the HYSPLIT trajectories (Figure S2, Supporting Information), as well as at the correct height (Winker 2009). Further confirmation of dust events was completed with MODIS images (obtained from http://modis-atmos.gsfc.nasa.gov/) of the sampling sites and starting points of the trajectories.

DNA extraction and shotgun sequencing

Total DNA was extracted off filters as described previously (Yuan *et al.* 2012). Briefly, lysis buffer (40 mM Tris–HCl, pH 8.5; 50 mM EDTA; 0.73 M sucrose) was applied to filters (500 μL for cellulose; 1 mL for quartz) with addition of lysis buffer containing an enzyme cocktail (lysozyme 5 mg/mL; mutanolysin 500 U/mL; stapholysin 200 U/mL): 100 μL of enzymes for cellulose filters, 200 μL for quartz. Enzymes were incubated with filters for 1 h at 37°C with slow rotation. Hot phenol:chloroform:isoamyl alcohol mixture (24:1:1; 60°C) of equal volume was applied to filters and vortexed. In addition, 600 mg of 0.1 mm beads were added to each filter and vigorously shaken for 1 min to further aid in cell lysis. Mechanical lysis (e.g. bead beating) has previously been used to aid in bacterial and eukaryotic cell lysis (Fredricks, Smith and Meier 2005; Yuan *et al.* 2012). After centrifugation, the aqueous phase was cleaned using a QIAamp DNA micro kit (QIAGEN, Hilden, Germany); elution of DNA off column was done with EB buffer. DNA was quantified using a QuBit fluorometer and the High Sensitivity dsDNA Quantification kit (QuBit Invitrogen, Carlsbad, CA, USA).

Extracted DNA samples were sequenced at the DOE Joint Genome Institute. Libraries were prepared using the Mondrian SP System (NuGEN Technologies, Inc., CA, USA); sequencing was performed on an Illumina HiSeq (2 × 150 bp; Table S1, Supporting Information; Illumina, San Diego, CA, USA)). Reads were filtered and processed for contaminating artifacts by JGI's in-house methods: artifacts; reads with 'N' calls; low quality (phred < 10); DNA spike-in removal; PhiX reads. The subsequent 'raw' reads were processed for quality score and length by SolexaQA (quality cutoff of 20; both sister reads length greater than 50 nt) (Cox, Peterson and Biggs 2010). Trimmomatic V0.32 was used to remove common Illumina adaptor sequences; only reads with length greater than 50 nt were kept. Processed reads were analyzed and confirmed for good quality by FastQC V0.11.2 (Andrews 2010).

Analysis of microbial composition

Cleaned reads were assembled by IDBA-UD (option: pre_correction) (Peng *et al.* 2011) and the N50 of all resulting contigs was determined (Table 2). Due to poor assembly of reads, we conducted taxonomic analyses on merged read fragments, using PEAR (Zhang *et al.* 2014) to merged the paired-end reads. 16S and 18S rRNA gene-encoding reads were extracted from each sample using Parallel-META 2.0 (Su *et al.* 2014) using the GreenGenes and Silva databases, respectively (DeSantis *et al.* 2006; Quast *et al.* 2012). Combined 18S and 16S SSU rRNA reads were put through standard QIIME workflows (Caporaso *et al.* 2010; Edgar 2010; Price, Dehal and Arkin 2010; Quast *et al.* 2012); sequences were clustered into operational taxonomic units (OTUs) using UCLUST at 97% similarity. An OTU phylogenetic table was created using the Silva database (release 104) as a reference, and taxonomy was summarized by absolute counts. Summarized tables were normalized in R using the DESeq and DESeq2 packages (Anders and Huber 2010; Love, Huber and Anders 2014). Normalized OTU counts were analyzed and visualized in R (Dixon and Palmer 2003; Wickham 2009; McMurdie and Holmes 2013). Non-multidimensial scaling (nMDS) analysis using the vegan package in R (distance = 'bray', k = 2, try.max = 50) was performed on family-level OTU summarized tables (Fig. 2). Hierarchical clustering was performed using vegan package (method = bray) (Oksanen *et al.* 2015) (Figure S3, Supporting Information).

Nucleotide sequence accession numbers

Raw paired-end reads from the generated metagenomes in this study were submitted to the NCBI Sequence Read Archive (SRA) under project: PRJNA345472. Accession number of individual metagenomic datasets are: 5864048, 5864049, 5864050, 5864051 and 5864052.

RESULTS AND DISCUSSION

To confirm air masses had associated dust originating from the Sahara, a combination of back trajectories (HYSPLIT model) and satellite data was utilized (Fig. 1). Confirmation of the 2008 dust events were shown with CALIOP Lidar 3D data from the CALIPSO, which confirmed the presence of dust particles at the longitude and latitude of collection (Figure S2, Supporting Information). HYSPLIT showed the exact path of the air masses collected for each event (Fig. 1). CALIPSO data for dust events occurring before 2007 were not available. Therefore, the 2008 events (SDPR002 and SDPR003; Fig. 1A and B) were used here to support the HYSPLIT models, which showed origin of air masses from North African coast (Fig. 1C–E).

Raw reads were processed through standard pipelines for read quality and length, adaptor trimming, and assembly using IDBA-UD. The resulting contigs were short (N50 < 600 nt) with a small number of reads assembled (Table S1, Supporting Information). Two other shotgun metagenomic reports of environmental air samples have been published at the time of writing (Be *et al.* 2014; Cao *et al.* 2014). Neither of these two studies reported on assembled reads; the Cao metagenomes were publicly available. The raw reads from the Cao study were put through the same quality and trimming controls as the SDPR samples and

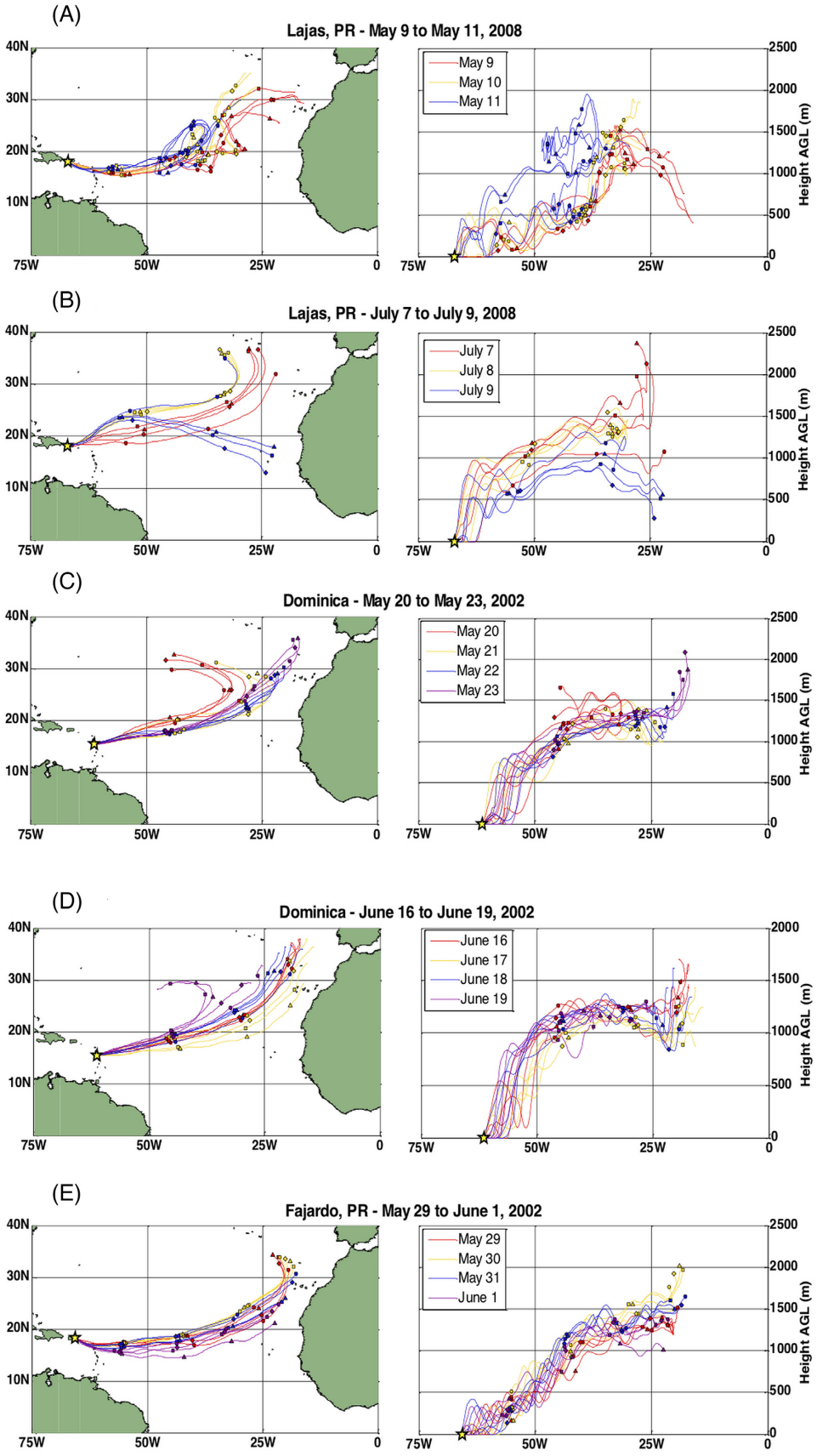

Figure 1. HYSPLIT back trajectories for SDEs, each day includes time points at 12 am, 6 am, 12 pm, 6 pm (GMT) with markers every 48 h going back 13 days. Stars indicate collection location. **(A)** Data for Lajas on May 9–11, 2008 (SDPR002). **(B)** Data for Lajas on July 7–9, 2008 (SDPR003). **(C)** Data for Dominica on May 20–23, 2002 (SDPR004). **(D)** Data for Dominica on June 16–19, 2002 (SDPR005). **(E)** Data for Fajardo on May 29–31 and June 1, 2002 (SDPR008).

Table 2. Summary of sequencing statistics.

Sample	Total DNA extracted (ng)	Total raw reads	Number of reads after QC	N50	# All contigs	Percentage of reads assembled	Percent reads merged (PEAR)	N50 of merged reads
SDPR-002	4	50388018	18574605	467	21001	25	37	206
SDPR-003	9	78640676	26040745	399	43740	18	33	215
SDPR-004	1.8	68619724	25208920	365	9387	7	37	208
SDPR-005	–	73466514	22102918	418	64509	20	30	217
SDPR-008	–	119586440	38329487	598	54546	8	32	216

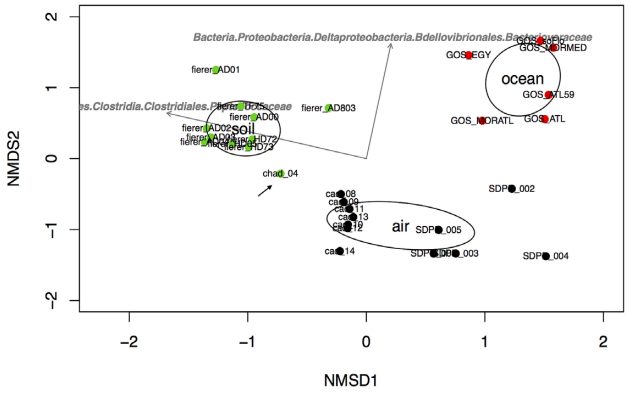

Figure 2. Normalized SSU rRNA analyses, at the family level, of SDPR and comparison of environment metagenomes. nMDS of metagenomes colored by environment type (green is soil environments, black is air environments and red is ocean environments). Ordination ellipses were generated for each environmental type; vectors indicate statistically significant influences of clustering. Black arrow indicates sample from Chad (i.e. source environment).

then assembled as described above for consistency purposes. The Cao reads also assembled poorly more so than the SDPR samples (Table S2, Supporting Information). It is hypothesized here that air samples are diverse and low in cell numbers of any particular organism and thus, do not assemble well. For this reason, paired reads were merged using PEAR and resulted in read length of ~200 nt; ~30% of total reads were merged. Subsequent analyses are focused on the merged reads with no assembly.

After dust from the Sahara becomes airborne, intercontinental transport into the Caribbean may take several days or up to a week (Petit *et al.* 2005; Ben-Ami *et al.* 2012; Prospero *et al.* 2014). This passage takes place over the Atlantic Ocean and it was expected that such long-term transport would cause the original microbial composition to undergo changes due to: (i) deposition of larger particles (either large microbial cells and aggregates or microorganisms associated with dust grains) due to gravitational settling, (ii) mixing of aerosolized surface ocean water organisms, and (iii) the theory that the atmosphere acts as a homogenizer of microbial signals. For the purpose of investigating the shifts in airborne microbial composition during transit, small subunit (SSU) ribosomal DNA (rRNA) gene fragments encoded on metagenomic reads were analyzed using the QIIME workflow and compared to previously reported metagenomes (Table S3, Supporting Information) of the source environment (Sahara) (Giongo *et al.* 2013), ocean surface water (global Ocean Sampling Day [GOS]; Kopf *et al.* 2015), other desert (hot and cold) environments (Fierer *et al.* 2012), forest soil (Fierer *et al.* 2012) and surface air (Cao *et al.* 2014) (geography of metagenomes; Figure S5, Supporting Information).

The SSU rRNA fragments were analyzed at the family taxonomic level after normalization for the size of each dataset using DESeq2 (Love *et al.* 2014) and compared in similarity ordination plots (nMDS) and with hierarchical clustering (Figure S4, Supporting Information). The top 20 families from each metagenome were plotted with other families summed into the 'Others' category (Figure S4, Supporting Information). As may be expected after long-distance transport across the Atlantic Ocean, SDPR metagenomes did not cluster closely to the metagenome collected in Chad (Fig. 2; Figure S3, Supporting Information). It was not expected that all of the SDPR metagenomes would be closely clustered with one another, as it is known that each dust event varies in terms of intensity (Prospero and Lamb 2003; Yu *et al.* 2015), as well as, the likely variable influence of the long-term transit across the Atlantic Ocean for each individual microbial population. In addition, the SDPR samples were temporally disparate from one another (months to years). Despite these possible influences, three of the SDPR metagenomes clustered together (SDPR003, -005 & -008), which indicated that SDEs may not differ greatly in terms of (i) microbial composition from one dust event to the next and (ii) processes that may transform the samples, especially with respect to particle size. The dominant OTUs influencing the clustering of SDPR003, -005 and -008 (% reads matching: 62, 42 and 51%, respectively) with other air metagenomes are from known soil fungi: Saccharomyceta and Agaricomycotina (Figure S4, Supporting Information). The former has been recently described as being negatively correlated with increasing moisture content (Pereira de Castro *et al.* 2016), i.e. these organisms may be found in well drained or arid soils. Agaricomycotina are common soil organisms, dominated by the mushrooms and their spores (Hibbett 2006).

It was also interesting that the SDPR (-003, -005 and -008) and Beijing metagenomes clustered together with 40–68% of SSU rRNA fragment assignments to fungi (Fig. 2; Figure S3, Supporting Information). It is important to note there may be an overestimation of abundance at the rRNA level due to the multiple copy number occurrence of rRNA genes in certain fungal species. This copy number bias of fungal rRNA genes also confounds the ability to compare bacterial versus fungal microbial fractions. Despite this, it is apparent that when talking of the taxonomic influences of the shotgun metagenomic reads, the SDPR003, -005 and -008 were influenced by a fungal fraction that could also be seen to occur in the urban air shotgun metagenomes. The Cao samples were collected in Beijing during a relatively intense urban smog event, with an increase in $PM_{2.5}$- and PM_{10}-sized particles (Cao *et al.* 2014). From these data and other reports, it is clear that fungi are an influencing signal in sequencing data from the atmosphere, during dust events or not (Griffin *et al.* 2001; Kellogg *et al.* 2004; Fröhlich-Nowoisky *et al.* 2012; Oh *et al.* 2014).

The SDPR metagenomes have eukaryotic matches previously described to be present in the Sahara (Favet *et al.* 2013; Giongo *et al.* 2013): Dikarya and Streptophyta. Though within the

Dikarya, the SDPR samples showed a stronger influence from Basidiomycota than the Chad and Beijing (Cao) metagenomes, which was in line with previous results indicating a higher prevalence of Basidiomycota in air samples in the Caribbean compared to Asian air samples (Quintero, Rivera-Mariani and Bolaños-Rosero 2010; Fröhlich-Nowoisky *et al*. 2012). This latter observation may indicate that during dust events in the Caribbean, the local background (i.e. marine source) of microbial signals was still present and detectable, as Basidiomycota has been found to be associated with coastal air collection previously (Urbano *et al*. 2011).

It was expected that the oceanic influences were significant on the SDPR samples and this may explain changes in microbial composition from the source environment (e.g. the Sahara). Indeed, the HYSPLIT trajectories for SDPR002 and -004 (Fig. 1A and C) did have back trajectories that indicated some of the air masses sampled had more of 'mixed' origins compared to SDPR003, -005 and -008 (Fig. 1). Further confirmation of their variance to the other SDPR metagenomes was shown in the relatively low assignment of SSU rRNA fragments to fungi of SDPR002 and -004, 3 and 13%, respectively. The two samples were not only dissimilar to the other SDPR metagenomes but also to each other, further suggesting a more 'mixed' environmental source. For instance, SDPR002 had more SSU rRNA fragments assigned to common soil- and plant-associated bacteria (e.g. Bacillaceae and Xanthomonadaceae); 93% of all SSU rRNA fragments in SDPR002 had bacterial assignments compared to 20–33% range of the other metagenomes. Consistent with these interpretations, the SDPR002 and -004 metagenomes contained matches to known oceanic orders of cyanobacteria and Alphaproteobacteria: SAR11, Rhodobacterales and Synechococcales (e.g. Prochlorococcus), Rickettsiales (e.g. Pelagibacteraceae).

There was no way, however, from this dataset, to tell how much the background environment may have played on the sampled microbial population, especially as sampling was performed on an island. The Basidiomycota signal, SDPR002 and -004 having oceanic signals, and the bias that may occur with atmospheric sampling at ground locations together all highlight the need for controls and non-dust day sampling (i.e. background environment sampling) for future sampling (Schuerger *et al*. 2018).

The similarities in SDPR003, -005 and -008 were in line with previous reports of Saharan dust collected in Europe where dust days were very similar among themselves (16S rRNA and 18S rRNA gene amplicon data) when compared to non-dust days even though these dust days occurred in separate years (Mazar *et al*. 2016; Cáliz *et al*. 2018). These results also highlight the importance of monitoring air masses when sampling and the need for sampling and handling controls. While it is not possible to say for certain that SDPR002 and -004 would have clustered better with the other SDPR metagenomes had monitoring of air masses occurred, it is more than likely that there was a 'diluting out' of dust sampled during these two events due to collection of air masses that did not have a clear African origin.

It should be also be noted that the Chad metagenome might not be completely representative of the entire Saharan soil microbial communities as it represents a single sample. However, Figure S3 (Supporting Information) shows that the Chad metagenome clustered more closely with other hot desert metagenomes (Fierer *et al*. 2012). Figure S4 (Supporting Information) shows that the arid soils (cold and hot deserts, including Chad metagenome) were heavily dominated by a number of actinobacterial families, most notably: Solirubrobacterales, AK1W543 and Acidimicrobiales. It was interesting that the air metagenomes from Beijing showed higher similarities to Chad and arid soils (Figure S3, Supporting Information) compared to the SDPR samples. This may have been due to higher particulate matter concentrations in Beijing urban air and/or dust input from desert regions in China (Wang *et al*. 2004; Ginoux *et al*. 2012).

This study indicated that African dust events occurring in the Caribbean were from year to year more similar to one another and that fungi may be the dominant microbial group associated with these dust events. In addition, the data indicated that in terms of taxonomy (SSU rRNA), air transit (and its associated processes, e.g. mixing and gravitational settling of larger particles) of a source material resulted in more similarity to other air-processed samples from different environments. However, this may also be a result of which sources have the strongest emission of microbial cells into the atmosphere (i.e. soil and vegetative environments). In conclusion, while this study furthers the field in terms of describing microbial populations transported by dust from Africa (Schuerger *et al*. 2018), it also raises more questions to be answered. Future work should focus on high volume and more frequent sampling, coupled with downstream viability and microscopy analysis, in order to better link the dynamics in abundance of the airborne organisms with weather conditions and the influence of local background and oceanic air mass transport.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSLE* online.

FUNDING

This work was supported by the Camille and Henry Dreyfus Foundation Postdoctoral Program in Environmental Chemistry. The work was conducted by the United States Department of Energy Joint Genome Institute, a Department of Energy Office of Science User Facility, under Contract No. DE-AC02–05CH11231. Field sampling in Puerto Rico and Dominica was supported by the National Oceanic and Atmospheric Administration, Educational Partnership Program, United States Department of Commerce, under Agreement No. NA11SEC4810003.

ACKNOWLEDGMENTS

The authors would like to thank Dr Nenes for their valuable feedback on this work and Drs DeLeon-Rodriguez and Konstantinidis for facilitating the DNA sequencing of the samples and their feedback on the bioinformatic analyses.

Conflicts of interests. None declared.

REFERENCES

Anders S, Huber W. Differential expression analysis for sequence count data, *Genome Biol* 2010;**11**:R106.

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Version 0.11.2 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Barberán A, Ladau J, Leff JW *et al*. Continental-scale distributions of dust-associated bacteria and fungi. *Proc Natl Acad Sci USA* 2015;**112**:5756–61.

Ben-Ami Y, Koren I, Altaratz O *et al*. Discernible rhythm in the spatio/temporal distributions of transatlantic dust. *Atmos Chem Phys* 2012;**12**:2253–62.

Be NA, Thissen JB, Fofanov VY *et al.* Metagenomic analysis of the airborne environment in Urban Spaces. *Microb Ecol* 2014;**69**:346–55.

Cao C, Jiang W, Wang B *et al.* Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event. *Environ Sci Technol* 2014;**48**:1499–507.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Caro TA, Wendeln M, Freeland M *et al.* Ultraviolet light measurements (280–400 nm) acquired from stratospheric balloon flight to assess influence on bioaerosols. *Aerobiologia* 2019;**35**:771–6.

Choudoir MJ, Barberán A, Menninger HL *et al.* Variation in range size and dispersal capabilities of microbial taxa. *Ecology* 2018;**99**:322–34.

Cox MP, Peterson DA, Biggs PJ. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 2010;**11**:1–6.

Creamean JM, Suski KJ, Rosenfeld D *et al.* Dust and biological aerosols from the Sahara and Asia influence precipitation in the Western U.S. *Science* 2013;**339**:1572–8.

Cáliz J, Triadó-Margarit X, Camarero L *et al.* A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proc Natl Acad Sci USA* 2018;**115**:12229.

DeSantis TZ, Hugenholtz P, Larsen N *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006;**72**:5069–72.

Després V, Huffman JA, Burrows SM *et al.* Primary biological aerosol particles in the atmosphere: a review. *Tellus B* 2012;**64**:15598.

Dixon P, Palmer MW. VEGAN, a package of R functions for community ecology. *J Veg Sci* 2003;**14**:927–30.

Draxler RR. *Hybrid Single-Particle Lagrangian Integrated Trajectories (HY-SPLIT), Version 3.0: User's Guide and Model Description*. US Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Air Resources Laboratory, 1992.

Draxler RR. Hybrid single-particle Lagrangian integrated trajectories (RT-SPLIT) model description. 1988.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Favet J, Lapanje A, Giongo A *et al.* Microbial hitchhikers on intercontinental dust: catching a lift in Chad. *ISME J* 2013;**7**:850–67.

Fierer N, Leff JW, Adams BJ *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci USA* 2012;**109**:21390–5.

Fredricks DN, Smith C, Meier A. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *J Clin Microbiol* 2005;**43**:5122–8.

Fröhlich-Nowoisky J, Burrows SM, Xie Z *et al.* Biogeography in the air: fungal diversity over land and oceans. *Biogeosciences* 2012;**9**:1125–36.

Ginoux P, Prospero JM, Gill TE *et al.* Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. *Rev Geophys* 2012;**50**:RG3005.

Giongo A, Favet J, Lapanje A *et al.* Microbial hitchhikers on intercontinental dust: high-throughput sequencing to catalogue microbes in small sand samples. *Aerobiologia* 2013;**29**:71–84.

Griffin D, Garrison V, Herman J *et al.* African desert dust in the Caribbean atmosphere: microbiology and public health. *Aerobiologia* 2001;**17**:203–13.

Griffin DW, Kellogg CA, Garrison VH *et al.* Atmospheric microbiology in the northern Caribbean during African dust events. *Aerobiologia* 2003;**19**:143–57.

Griffin DW. Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin Microbiol Rev* 2007;**20**:459–77.

Hibbett DS. A phylogenetic overview of the Agaricomycotina. *Mycologia* 2006;**98**:917–25.

Itani GN, Smith CA. Dust rains deliver diverse assemblages of microorganisms to the Eastern Mediterranean. *Sci Rep* 2016;**6**:22657.

Jeon EM, Kim HJ, Jung K *et al.* Impact of Asian dust events on airborne bacterial community assessed by molecular analyses. *Atmos Environ* 2011;**45**:4313–21.

Jones AM, Harrison RM. The effects of meteorological factors on atmospheric bioaerosol concentrations – a review. *Sci Total Environ* 2004;**326**:151–80.

Katra I, Arotsker L, Krasnov H *et al.* Richness and diversity in dust stormborne biomes at the Southeast Mediterranean. *Sci Rep* 2014;**4**:5265.

Kellogg CA, Griffin DW, Garrison VH *et al.* Characterization of aerosolized bacteria and fungi from desert dust events in Mali, West Africa. *Aerobiologia* 2004;**20**:99–110.

Kopf A, Bicak M, Kottmann R *et al.* The ocean sampling day consortium. *GigaScience* 2015;**4**:1–5.

Lighthart B, Shaffer B. Bacterial flux from chaparral into the atmosphere in mid-summer at a high desert location. *Atmos Environ* 1994;**28**:1267–74.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:1.

Mazar Y, Cytryn E, Erel Y *et al.* Effect of dust storms on the atmospheric microbiome in the Eastern Mediterranean. *Environ Sci Technol* 2016;**50**:4194–202.

McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;**8**:e61217.

Meola M, Lazzaro A, Zeyer J. Bacterial composition and survival on Sahara dust particles transported to the European Alps. *Front Microbiol* 2014;**6**:1454.

Morris CE, Sands DC, Bardin M *et al.* Microbiology and atmospheric processes: research challenges concerning the impact of airborne micro-organisms on the atmosphere and climate. *Biogeosciences* 2011;**8**:17–25.

Moulin C, Lambert CE, Dulac F *et al.* Control of atmospheric export of dust from North Africa by the North Atlantic Oscillation. *Nature* 1997;**387**:691–4.

Oh SY, Fong JJ, Park MS *et al.* Identifying airborne fungi in Seoul, Korea using metagenomics. *J Microbiol* 2014;**52**:465–72.

Oksanen J, Blanchet FG, Kindt R *et al. vegan: Community ecology package*. 2015.

Peng Y, Leung HC, Yiu SM *et al.* Meta-IDBA: a *de novo* assembler for metagenomic data. *Bioinformatics* 2011;**27**:i94–i101.

Pereira de Castro A, Sartori da Silva MRS, Quirino BF *et al.* Microbial diversity in Cerrado Biome (Neotropical Savanna) soils. *PLoS One* 2016;**11**:e0148785.

Perfumo A, Marchant R. Global transport of thermophilic bacteria in atmospheric dust. *Env Microbiol Rep* 2010;**2**:333–9.

Petit R, Legrand M, Jankowiak I *et al.* Transport of Saharan dust over the Caribbean Islands: study of an event. *J Geophys Res-Atmos* 2005;**110**:004748.

Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490.

Prospero JM, Blades E, Mathison G *et al*. Interhemispheric transport of viable fungi and bacteria from Africa to the Caribbean with soil dust. *Aerobiologia* 2005;**21**:1–19.

Prospero JM, Bonatti E, Schubert C *et al*. Dust in the Caribbean atmosphere traced to an African dust storm. *Earth Planet Sci Lett* 1970;**9**:287–93.

Prospero JM, Collard F, Molinié J *et al*. Characterizing the annual cycle of African dust transport to the Caribbean Basin and South America and its impact on the environment and air quality. *Global Biogeochem Cy* 2014;**28**:757–73.

Prospero JM, Lamb PJ. African droughts and dust transport to the Caribbean: climate change implications. *Science* 2003;**302**:1024–7.

Prospero JM, Mayol-Bracero OL. Understanding the transport and impact of african dust on the Caribbean Basin. *Bull Am Meteorol Soc* 2013;**94**:1329–37.

Prospero JM. Long-term measurements of the transport of African mineral dust to the southeastern United States: implications for regional air quality. *J Geophys Res* 1999;**104**:15917–27.

Quast C, Pruesse E, Yilmaz P *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;**41**:D590–6.

Quintero E, Rivera-Mariani F, Bolaños-Rosero B. Analysis of environmental factors and their effects on fungal spores in the atmosphere of a tropical urban area (San Juan, Puerto Rico). *Aerobiologia* 2010;**26**:113–24.

Schuerger AC, Smith DJ, Griffin DW *et al*. Science questions and knowledge gaps to study microbial transport and survival in Asian and African dust plumes reaching North America. *Aerobiologia* 2018;**34**:425–35.

Smith DJ, Griffin DW, McPeters RD *et al*. Microbial survival in the stratosphere and implications for global dispersal. *Aerobiologia* 2011;**27**:319–32.

Stein AF, Draxler RR, Rolph GD *et al*. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bull Am Meteorol Soc* 2015;**96**:2059–77.

Stres B, Sul WJ, Murovec B *et al*. Recently deglaciated high-altitude soils of the Himalaya: diverse environments, heterogenous bacterial communities and long-range dust inputs from the upper troposphere. *PLoS One* 2013;**8**:e76440.

Su X, Pan W, Song B *et al*. Parallel-META 2.0: enhanced metagenomic data analysis with functional annotation, high performance computing and advanced visualization. *PLoS One* 2014;**9**:e89323.

Urbano R, Palenik B, Gaston C *et al*. Detection and phylogenetic analysis of coastal bioaerosols using culture dependent and independent techniques. *Biogeosciences* 2011;**8**:301–9.

Vila-Costa M, Barberan A, Auguet J-C *et al*. Bacterial and archaeal community structure in the surface microlayer of high mountain lakes examined under two atmospheric aerosol loading scenarios. *FEMS Microbiol Ecol* 2013;**84**:387–97.

Wang X, Dong Z, Zhang J *et al*. Modern dust storms in China: an overview. *J Arid Environ* 2004;**58**:559–74.

Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2009.

Wilkinson DM, Koumoutsaris S, Mitchell EA *et al*. Modelling the effect of size on the aerial dispersal of microorganisms. *J Biogeogr* 2012;**39**:89–97.

Winker D. *CALIPSO LID L1 ValStage1 HDF File - Version 3.01*. 2009.

Yamaguchi N, Baba T, Ichijo T *et al*. Abundance and community structure of bacteria on Asian dust particles collected in Beijing, China, during the Asian dust season. *Biol Pharm Bull* 2016;**39**:68–77.

Yuan S, Cohen DB, Ravel J *et al*. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* 2012;**7**:e33865.

Yu H, Chin M, Yuan T *et al*. The fertilizing role of African dust in the Amazon rainforest: a first multiyear assessment based on data from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations. *Geophys Res Lett* 2015;**42**:1984–91.

Zhang J, Kobert K, Flouri T *et al*. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics* 2014;**30**:614–20.