



OPEN

Non-random genetic alterations in the cyanobacterium *Nostoc* sp. exposed to space conditions

Yuguang Liu^{1,2,13,15}, Patricio Jeraldo^{1,2,15}, William Herbert^{2,14}, Samantha McDonough³, Bruce Eckloff³, Jean-Pierre de Vera⁴, Charles Cockell⁵, Thomas Leya⁶, Mickael Baqué⁷, Jin Jen³, Dirk Schulze-Makuch^{8,9,10,11}✉ & Marina Walther-Antonio^{1,2,12}✉

Understanding the impact of long-term exposure of microorganisms to space is critical in understanding how these exposures impact the evolution and adaptation of microbial life under space conditions. In this work we subjected *Nostoc* sp. CCCryo 231-06, a cyanobacterium capable of living under many different ecological conditions, and also surviving in extreme ones, to a 23-month stay at the International Space Station (the Biology and Mars Experiment, BIOMEX, on the EXPOSE-R2 platform) and returned it to Earth for single-cell genome analysis. We used microfluidic technology and single cell sequencing to identify the changes that occurred in the whole genome of single *Nostoc* cells. The variant profile showed that biofilm and photosystem associated loci were the most altered, with an increased variant rate of synonymous base pair substitutions. The cause(s) of these non-random alterations and their implications to the evolutionary potential of single bacterial cells under long-term cosmic exposure warrants further investigation.

Microorganisms with a high level of resistance to desiccation, ultraviolet (UV) radiation, vacuum tolerance¹ and a capability of surviving in extreme environments such as cyanobacteria, lichen, fungi and green algae often serve as models for studies of the limits of terrestrial life². For example, lichens have survived a 14-day low-Earth orbit flight in the BIOPAN-5 facility³; cyanobacteria and lichens have survived a 10-day space exposure in the Lithopanspermia Space Experiment⁴; and cyanobacteria, lichens and Antarctic fungi survived for 15 months on the EXPOSE-E and EXPOSE-R platforms on the International Space Station (ISS)⁵.

Among the aforementioned microorganisms, *Nostoc* is a genus of cyanobacteria growing on soil, rocks and pool substrates as well as a symbiont in other plants from the tropics to temperate and mountain climates on our earth, but also in extreme habitats from the frigid Antarctic valleys to hot desert soils⁶. Its survival is largely due to its ability to recover its metabolic activities within days after rehydration even after years of desiccation. Recently, *Nostoc* sp. CCCryo 231-06, a strain isolated from permafrost soil and rock substrates in Antarctica, was used for the 15-month Biology and Mars Experiment (BIOMEX) on the ISS⁷. Understanding the genomic evolution of this ecologically widely adapted species, also with regard to extreme environmental conditions, can serve as a guide for further studies and will be vital for a better understanding of how life adapts to space environments.

Single cell whole genome sequencing (SC-WGS) is being increasingly applied for investigating the genetic diversity and heterogeneity of complex biological systems⁸. However, environmental samples often consist of a

¹Division of Surgical Research, Department of Surgery, Mayo Clinic, Rochester, MN 55905, USA. ²Microbiome Program, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA. ³Medical Genome Facility, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA. ⁴Space Operations and Astronaut Training, Microgravity User Support Center (MUSC), German Aerospace Center (DLR), Linder Höhe, 51147 Köln, Germany. ⁵School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, UK. ⁶Fraunhofer Institute for Cell Therapy and Immunology, Branch Bioanalytics and Bioprocesses IZI-BB, 14476 Potsdam, Germany. ⁷Astrobiological Laboratories, Planetary Laboratories Department, Institute of Planetary Research, German Aerospace Center (DLR), 12489 Berlin, Germany. ⁸Astrobiology Group, Center of Astronomy and Astrophysics, Technische Universität Berlin, 10623 Berlin, Germany. ⁹Section Geomicrobiology, GFZ German Research Center for Geosciences, Potsdam, Germany. ¹⁰Department of Experimental Limnology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Stechlin, Germany. ¹¹School of the Environment, Washington State University, Pullman, WA, USA. ¹²Department of Obstetrics and Gynecology, Mayo Clinic, Rochester, MN 55905, USA. ¹³Department of Immunology, Mayo Clinic, Rochester, MN 55905, USA. ¹⁴Mayo Clinic Graduate School of Biomedical Sciences, Mayo Clinic, Rochester, MN 55905, USA. ¹⁵These authors contributed equally: Yuguang Liu and Patricio Jeraldo. ✉email: schulze-makuch@tu-berlin.de; waltherantonio.marina@mayo.edu

biomass of complex cell communities and contaminants, making it difficult to identify genomic variants in low-abundance target cells⁹. With microfluidic technologies, it is feasible to isolate single cells from a population, and amplify femtograms of DNA for sequencing in a precisely controlled manner with minimal contamination¹⁰.

In this work, we performed microfluidic-based SC-WGS on single cells of the filamentous cyanobacterium *Nostoc* sp. CCCryo 231-06 exposed to space and Mars-simulated conditions for 23 months on the ISS. Experimental conditions included UV and cosmic radiation exposure, microgravity, and the use of lunar and martian analog soils on the ISS and in a Mars simulation chamber on Earth, to better understand the impact of these parameters. Our goal was to provide insights into genomic changes at the cellular level experienced by a species able to survive long-term exposure to space conditions.

Materials and methods

Sample preparation. The Antarctic strain CCCryo 231-06 (= UTEX EE21; CCME 391) of the cyanobacterium *Nostoc* sp. (hereafter, for readability addressed as "the *Nostoc*" only) was used as the target strain in this work. Briefly, the sample was prepared pre-flight, and the samples were on the ISS for 23 months, placed outside the ISS for 17 months, and exposed to UV for 15 months. These samples were returned to Earth and collected for post-flight processing before SC-WGA in a microfluidic device for sequencing and analysis. The overall workflow is illustrated in Fig. 1. Details of strain origin, pre-flight sample preparation, BIOMEX experimental parameters and post-flight sample preparation can be found in the supplemental information (SI) S1.

Single cell sequencing. *Microfluidic experimental setup.* SC-WGA was performed in an optofluidic platform at Mayo Clinic (Rochester, MN)^{11–13} with a customized microfluidic device for high-throughput SC-WGA. Briefly, this platform consists of a microscope (Nikon Eclipse), optical tweezers (Thorlabs) and a custom-built Polydimethylsiloxane (PDMS) microfluidic device with 4 identical reaction blocks that contain 14 parallel reaction lines in each block (Fig. 2a). Each reaction line has sets of valves that allow for the creation of isolated microchambers. The number of microchambers in each reaction line corresponds to the number of reagents that needs to be sequentially added to perform the SC-WGA reactions. The sample inlets of the microfluidic devices were designed to support the injection and isolation of cells under different experimental conditions in an individual manner, therefore minimizing cross-contamination between samples.

Each reaction block in the microfluidic device was designed to accommodate 13 single cells from the filaments of the *Nostoc* exposed to different conditions as listed in Table S1 and one negative control (PBS). Our original plan was to perform SC-WGA on 4 single cells from each of the 13 samples, however, it was increasingly difficult to identify *Nostoc* cells in sample No. 7 and sample No. 13, and thus, only three single cells and one single cell from these samples were recovered and sequenced (Table S1) respectively.

Microfluidic-based SC-WGA. The desiccated samples were rehydrated according to the procedure described in the supplemental information S1. Prior to introducing the cells into the microfluidic device, the sample channel in the microfluidic device was primed with chip diluent (0.04% Pluronic F127 in PBS) for 30 min to prevent the cells from sticking to the PDMS surface. The prepared cell suspension was then injected into the device. To reduce possible location-induced sequencing bias when performing SC-WGA in a microfluidic device, each sample was injected through channels at two different locations in the device, with two cells of each sample isolated into adjacent reaction lines at each location. The sample placement scheme is shown in Fig. 2b. The shaded areas represent samples from the ground simulation experiments.

Single cells were trapped and transported into microchambers by optical tweezers. Visually identifiable contaminating cells were trapped and moved out of the chambers to ensure only the target cell was in the chamber prior to the lysis step. The major advantages of using optical tweezers to isolate single cells from a population include high target single cell confidence, providing a way to visually ensure that only one cell is trapped into a microchamber and thus maintaining minimal possibility of sequencing contaminating cells unintendedly. A representative image of using the laser tweezers to trap a single *Nostoc* cell is shown in Fig. 2c. Lysis of the *Nostoc* cells and genome amplification procedure followed our optimized protocol for bacterial single cell lysis and whole genome amplification in microfluidic devices for downstream sequencing¹¹. The protocols for *Nostoc* lysis and library construction and sequencing can be found in SI.

Bioinformatic data analysis. *De novo genome assembly.* The sequenced reads were pre-processed prior to de novo genome assembly (see SI). One of the major challenges of bacterial SC-WGS is that the whole genome of a target single cell can only be partially recovered. However, with sufficient number of single cells (regardless of different experimental conditions), it is possible to co-assemble a consensus *Nostoc* genome to near completion. Therefore, in this work, we co-assembled using all 48 samples targeting the *Nostoc* isolate¹⁴. We emphasize the "consensus" aspect of this recovered reference sequence, and consequently this sequence will not completely match the genotypes of samples from each single exposure condition. This co-assembled consensus genome was to be used as a reference for the variant calling procedures and as a template for gene/function annotation to inform the significance of the detected variants. In this case, we determined that the consensus reference would be an acceptable substitute for the purpose of identifying variants of interest and offering insights into the functions present in *Nostoc* sp. CCCryo 231-06.

Specifically, we combined the reads from all 48 samples, and to offset the uneven coverage of sequenced reads introduced by the whole-genome amplification process, we digitally normalized the read coverage using the BBNorm tool from the BBTools suite version 38.26¹⁵ to a target coverage of 100X. We then took the combined, digitally-normalized reads and used the MEGAHIT de novo metagenomic assembler version 1.1.3¹⁶ using the "meta-sensitive" preset.

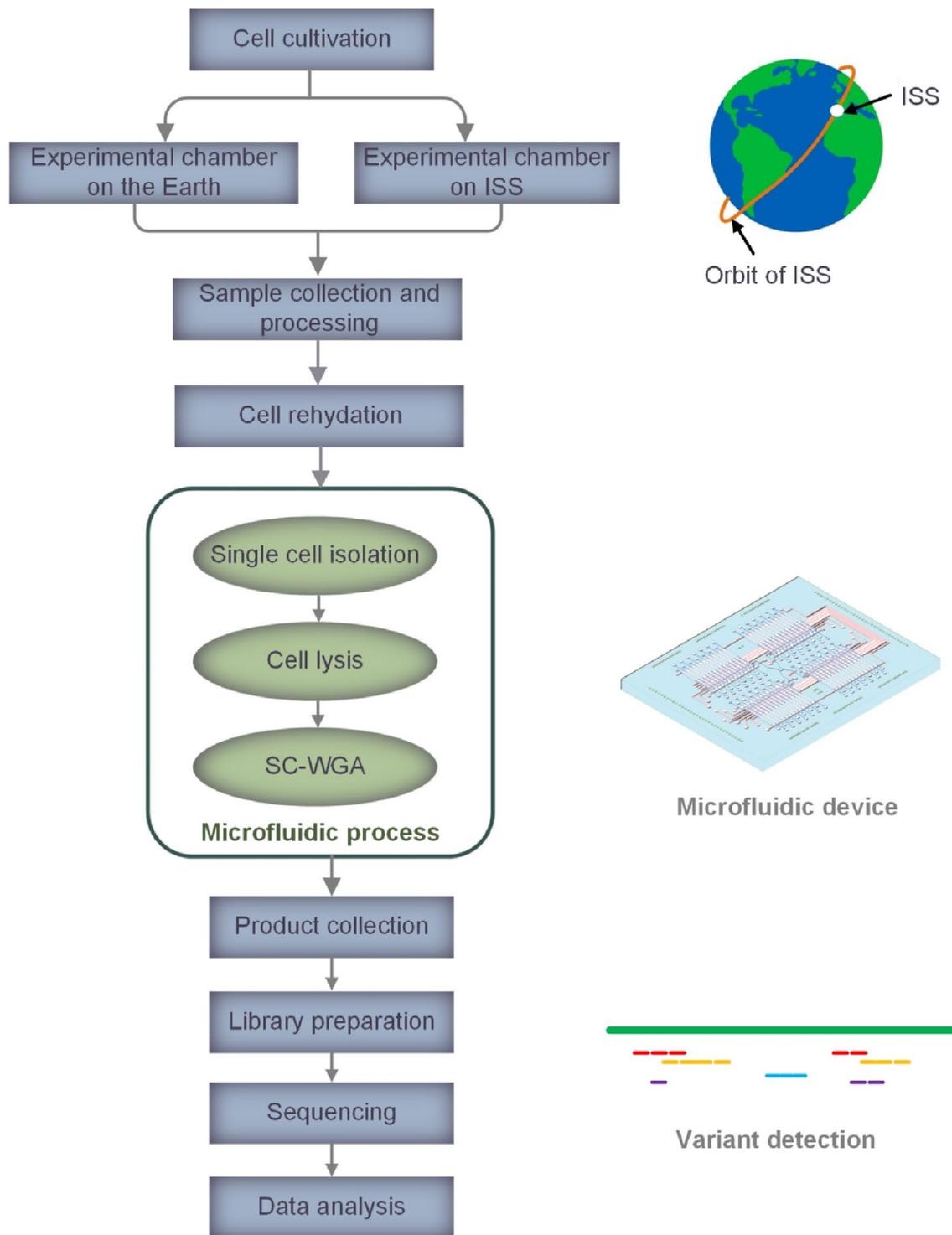


Figure 1. An overview of the workflow of SC-WGS for detecting variants in the *Nostoc* sp. strain CCCryo 231-06.

To reconstruct the consensus *Nostoc* sp. genome, we needed to separate its contigs from the contaminating contigs from other organisms (non-cyanobacterial bacteria, fungi). We used the BusyBee tool to identify and select the bin of our target organism¹⁷. We then assessed the initial quality of the recovered genome including the completeness and contamination based on a set of normally single copy gene markers using checkM version 1.0.13¹⁸, and then refined the bin using the refine tool following the procedure outlined in Parks, et al.¹⁹. Finally, we used the GTDBtk tool version 0.2.2²⁰ to putatively determine the taxonomic placement of the recovered genome. The genome was annotated using the RASTtk pipeline through the PATRIC service²¹. Details of metagenomics analysis and base quality score recalibration are described in SI.

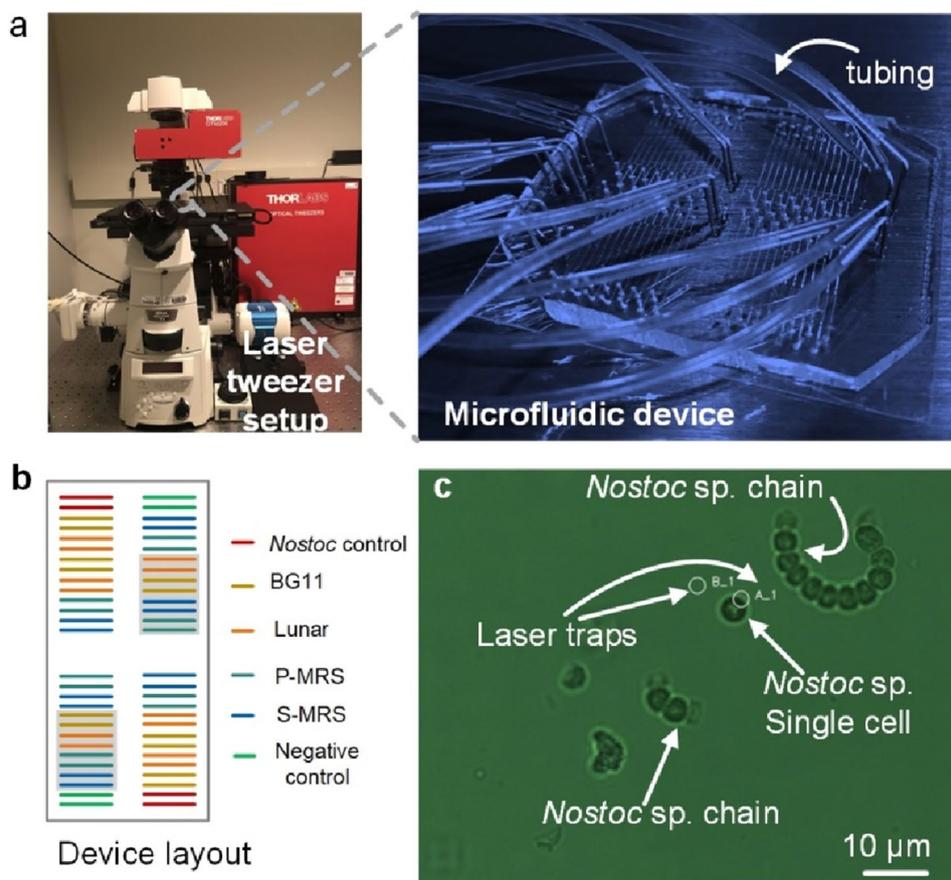


Figure 2. Optofluidic platform overview. **(a)** An optofluidic platform consists of a microscope, laser tweezers and a microfluidic device connected with tubings for high-throughput SC-WGA. **(b)** A schematic diagram of the microfluidic device with 4 identical reaction blocks with 14 reaction lines in each block. The location of samples and controls are shown as color coded. Experimental ground samples are indicated by the grey shades. **(c)** The use of laser tweezers to trap *Nostoc* CCCryo 231-06 single cells. Laser trap A_1 was turned on and trapped a single cell in the sample suspension. Laser trap B_1 was turned off. S-MRS sulfatic martian regolith, P-MRS phyllosilicatic martian regolith.

Variant calling. We used the recovered genome of the *Nostoc* sp. CCCryo 231-06 as a reference for variant calling. For each of the *Nostoc* samples and replicates, we mapped the reads with recalibrated base quality scores against this reference using BMap, then coordinate-sorted these reads using SAMtools. The variants were called using GATK4 version 4.0.1.0 HaplotypeCaller in haploid mode²². We used SnpEff version 4.3t to annotate variants and predict effect, and SnpSift version 4.3t to filter low quality variants²³. To remove the spurious variants resulting from the amplification procedure, we merged the replicates for each sample using BCFtools, requiring that each putative variant is present in at least two replicates. This way, most of the spurious variants will be removed. Finally, the variants were visualized using the web version of Integrative Genome Viewer.

Variant analysis. Whole genome and variant gene breadth of coverage were calculated using SAMtools. Variant analysis was performed on ISS/UV, ISS/Dark, and Ground/UV samples only as the gene coverage for samples grown on Mars and Lunar media was too poor (Fig. 3a). Synonymous to non-synonymous variant ratios (Ka/Ks ratio) were calculated by counting the respective variants in each replicate, gene-by-gene, and normalized according to the breadth of coverage of the area of interest²⁴. Genes were selected to be significant with Ka/Ks ratios greater than 1, indicating non-synonymous alterations or with Ka/Ks ratios less than 1, indicating synonymous alterations. Genes that met these criteria through at least two experimental conditions were analyzed for significance across replicates and within the genome with the Friedman test and Nemenyi post-hoc test. Probabilities of variant distributions on a gene-by-gene basis were calculated using Pascal's formula where N = length of gene and K = number of observed shared variants between conditions. Probabilities of genome-wide variant distribution were tested using an algorithm in which the whole genome was treated as a continuum. To assess whether the observed variants could be attributed to a random distribution, one million iterations of an equivalent set of simulated variants were created and randomly assigned to a location in the genome.

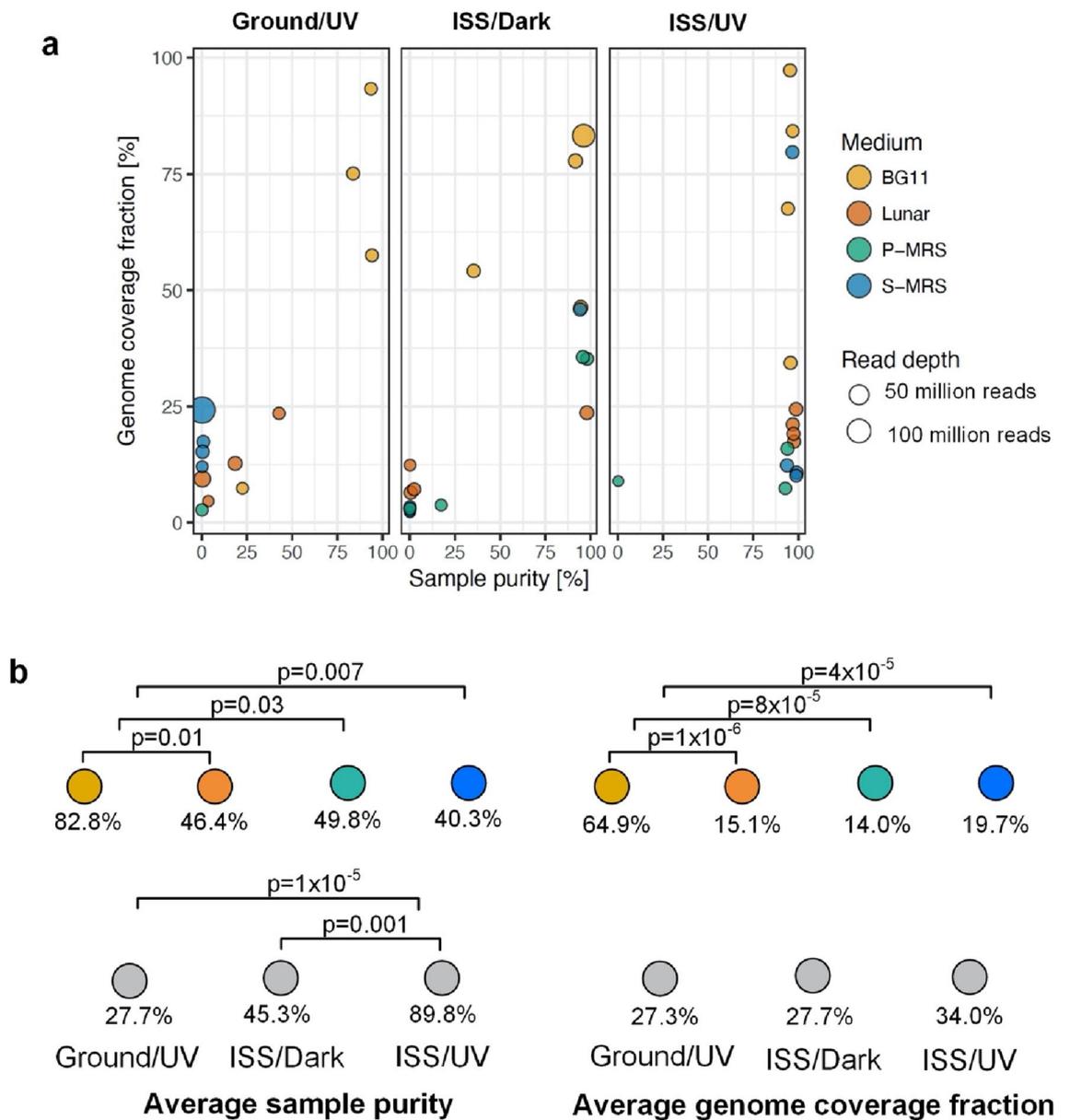


Figure 3. Sample purity and genome coverage of single cells of *Nostoc* sp. CCCryo 231-06 exposed to different media/substrates and exposure conditions during the BIOMEX experiment. Number of replicates per experimental conditions are described in Table S1. Different media/substrates are color-coded. (a) The horizontal axis represents the percentage of reads that were classified to the genus *Nostoc* after the metagenomic analysis of each sample and replicate. The vertical axis represents the fraction of the recovered *Nostoc* sp. CCCryo 231-06 that is mapped by the reads of each sample and replicate. The size of the circle represents the number of reads present in each sample and replicate. The components of the mineral substrates (lunar and Martian regoliths) are provided in Table S2. (b) Average sample purity and genome coverage fraction are sorted by medium and exposure condition. Samples on BG11 displayed significantly higher purity and genome coverage. Samples exposed to UV radiation on the ISS displayed significantly higher purity than those in other experimental conditions. Student's t test was used to calculate the p-values. BG11 blue-green medium, ISS international space station, P-MRS phyllosilicatic martian regolith, S-MRS sulfatic martian regolith, UV ultraviolet light, Ground/UV earth with UV exposure, ISS/Dark ISS without UV exposure, ISS/UV ISS with UV exposure.

Protein structure analysis. The consensus amino acid sequence of the *Nostoc*'s photosystem II D1 protein PsbA, obtained for the Ground/UV, ISS/UV, ISS/Dark experimental conditions, as well as for the reference genome, was submitted to the RaptorX Structure Prediction server to predict its tertiary structure²⁵. Each predicted structure was returned in PDB format, and they were structurally pairwise-aligned using PyMOL v2.3.0²⁶. PyMOL was then used to visualize the effects of variants across experimental conditions by showing structural differ-

ences and non-synonymous variants in the context of its tertiary structure. The nucleotide and amino acid sequences of proteins of interest were visualized using the method described in SI.

Contamination. Organisms in the samples that did not belong to the *Nostocaceae* were counted as contaminants. Sequenced single cells from space and ground samples on BG11 medium displayed an average of 82.8% sample purity, significantly higher than the purity of single cells on Lunar (46.4%), P-MRS (49.8%) and S-MRS (40.3%) regolith analogs with all p -values < 0.05 . As the regolith analogs are composed of multiple mineral components collected from various locations, there is a higher possibility of the mineral mixtures carrying contaminants. However, these minerals were autoclaved prior to plating the microorganism, thus it is more likely that the contaminants were introduced during sample transfer and handling. In either case, we assumed that contaminants did not influence the experimental results on the ISS and respective variants.

Results

Sample purity and genome coverage. To investigate possible changes in the genotype of *Nostoc* sp. CCCryo 231-06 under space conditions, we calculated sample purity and genome coverage of single cells exposed to different experimental conditions. The normalized sample purity and genome coverage of single cells of *Nostoc* sp. CCCryo 231-06 is shown in Fig. 3a. For samples on the ISS, the purity of those that were UV-exposed was $\sim 89.8\%$ ($SD = \pm 24.8\%$), approximately twice as pure as samples without UV exposure ($p \leq 0.03$, Student's t -test). This is in agreement with other studies demonstrating that UV radiation often leads to structural and genetic damage in various companion microorganisms in non-axenic cultures²⁷, leading to their lower survivability under extreme conditions. However, although the UV intensity in the simulation chamber on Earth was approximately twice that on the ISS, samples that remained on Earth were only $< 33\%$ as pure as those that were exposed to UV on the ISS (Fig. 3b). This raises the possibility that the combination of cosmic and UV radiations poses significant challenges to the survivability of most accompanying microorganisms. Further enhanced purity was observed under growth conditions on Blue-Green culture medium (BG11) relative to desiccated conditions on lunar or Mars-analog regolith ($p \leq 0.03$, Student's t -test). BG11 is one culture medium of several that is tailored to support growth of cyanobacteria like *Nostoc*, and therefore is likely to provide a competitive advantage over companion microbes. Samples of strain CCCryo 231-06 on the BG11 medium resulted in significantly higher genome coverage than those on lunar or Martian analogs. Specifically, the single cells on BG11 showed an average of 64.9% ($SD = \pm 26.3\%$) genome coverage, while in the cells on other substrates $< 20\%$ coverage was achieved ($p \leq 8 \times 10^{-5}$, Student's t -test) (Fig. 3b). While BG11 is widely used for cyanobacterial cultures and is well characterized and used in molecular assays, that is not the case for the lunar and Martian-analog substrates. It is possible that the latter contain compounds that have inhibitory effects on molecular reactions, interfering with the effectiveness of lysis and genomic amplification procedures. No statistically significant differences between the genome coverage of samples exposed on the ISS and on Earth were observed.

Genomic variant comparison. We used the recovered genome of the *Nostoc* sp. CCCryo 231-06 (NCBI BioProject accession number PRJNA721463; Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAHCSU010000000; genome announcement submitted elsewhere) as a reference for variant calling. To minimize artifacts or errors, we required that the same exact genomic variant be present in at least two single cell replicates exposed to the same condition to be called a variant. Genomic variants were only identified in samples on BG11, which may be a result of the low genomic coverage for the other substrates utilized (Fig. 3a). Therefore, variant analysis was performed on ISS/UV (ISS with UV exposure), ISS/Dark (ISS without UV exposure), and Ground/UV (Earth with UV exposure) samples on the BG11 medium only. We identified genomic variants in genes involved in biofilm production and/or photosynthesis. Biofilm associated variants appeared in hemagglutinin-related genes in samples from the ISS (in both UV and dark conditions) (Fig. 4a). While genomic variants related to photosynthetic genes were identified in all conditions (Photosystem II D1), some were identified only on Earth (high-light-inducible genes) (Fig. 4b,c). To examine the significance of these variants, genome-wide K_a/K_s ratios were calculated. This ratio represents the fraction of non-synonymous variants (K_a) to synonymous variants (K_s) (Table 1). Overall, all conditions exhibited a K_a/K_s ratio < 1 indicating that variants were preferentially synonymous.

When K_a/K_s ratios were calculated on a gene-by-gene basis with significant genes intersected across all experimental conditions, eight genes arose as regions of interest for examining the variant profiles in more depth. Four selected genes—a mobile element protein, *btuD*, *psbA*, and a transposase, displayed < 0.2 K_a/K_s ratios which indicates high density of synonymous variants (Table 2, Table S3). Three genes—two hypothetical proteins (proteins whose existence has been predicted but lack experimental evidence for *in vivo* expression) and a mobile element protein—displayed synonymous alterations only under certain conditions (Table 2, Table S3).

We also examined mutations within non-coding regions of the genome. There was a minimal number of intragenic noncoding mutations—3 for Ground/UV, 4 for ISS/Dark, and 1 for ISS/UV—but Ground/UV and ISS/Dark shared 3 mutations in the exact same position with the same base pair substitution in genes coding for a hypothetical protein, LSU rRNA, and SSU rRNA. All 3 samples shared the mutation within the same gene encoding for the small subunit ribosomal rRNA, but the mutation in ISS/UV occurred in a different base pair (Supplemental File S1). All intragenic mutations occurred in the same $\sim 35,000$ bp region. There was a larger amount of intergenic noncoding mutations—103 for Ground/UV, 159 for ISS/Dark, and 135 for ISS/UV. As with mutations seen in the coding regions of the genome, there appeared to be hotspots for intergenic noncoding mutations that were shared across the three conditions (Supplementary File S1). The distance to the closest gene was calculated for each intergenic mutation. The majority of intergenic mutations occurred within 10kbp of a gene—91/103 for Ground/UV, 113/159 for ISS/Dark, and 115/135 for ISS/UV. All mutations occurred within

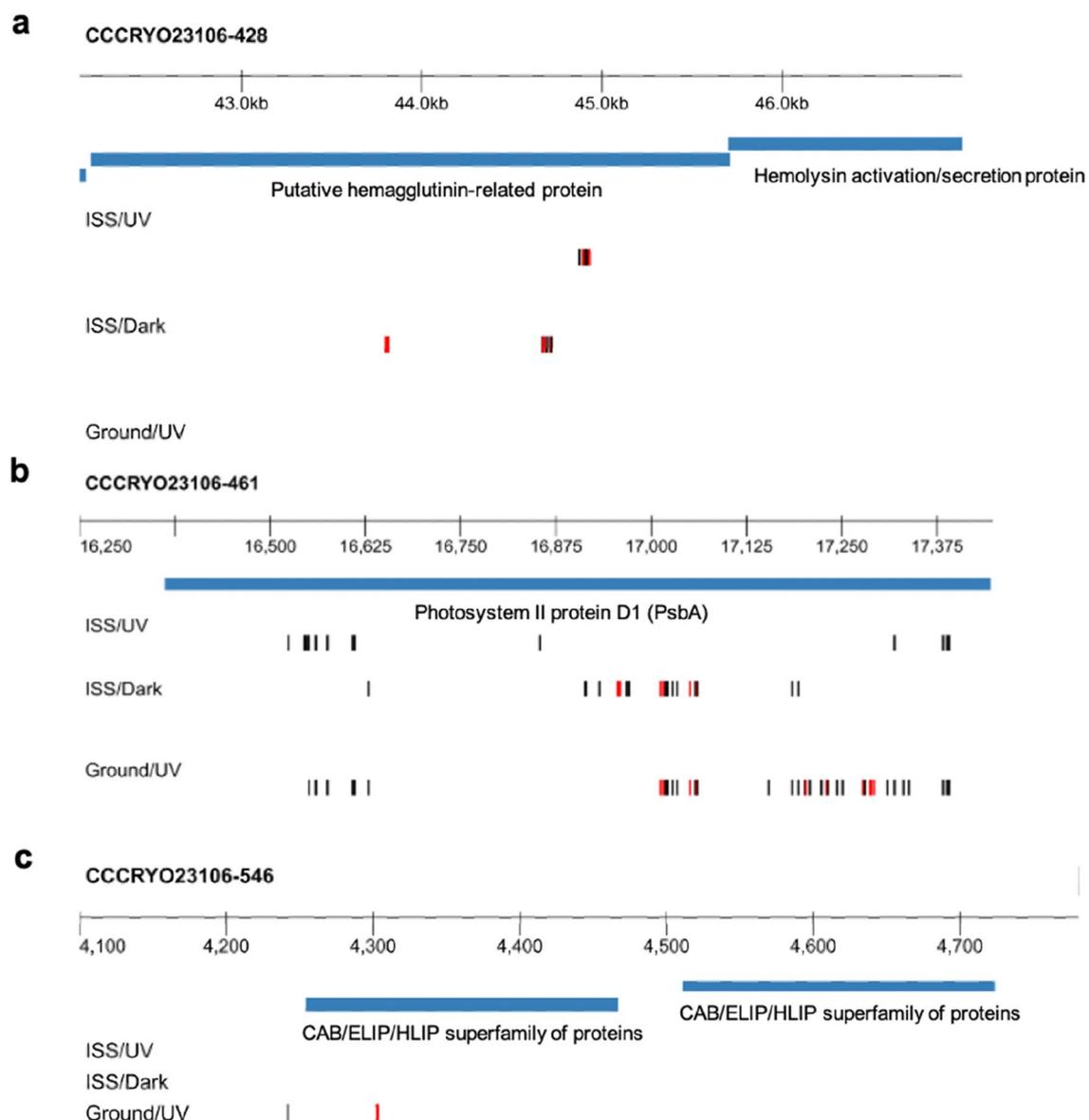


Figure 4. Plot of variants on BG11 medium detected in biofilm-associated hemagglutinin gene and photosynthesis-associated genes. **(a)** Biofilm-associated hemagglutinin genes. **(b)** Photosystem II D1 (*psbA*) gene. **(c)** High-light inducible protein genes. Each track represents a medium/exposure condition, and the blue markers represent variant positions with respect to the consensus reference genome for *Nostoc* sp. CCCryo 231-06. Black vertical lines represent synonymous variants (not impacting amino acid translation), red vertical lines represent non-synonymous variants (impacting amino acid translation). ISS international space station, UV ultraviolet light, ISS/UV ISS with UV exposure, ISS/Dark ISS without UV exposure, Ground/UV earth with UV exposure.

Sample	Non-synonymous variants	Synonymous variants	Ka/Ks Ratio
Ground/UV	64	117	0.54
ISS/UV	57	131	0.43
ISS/Dark	44	113	0.38

Table 1. Genome-wide Ka/Ks ratios. *Ka* non-synonymous variant, *Ks* synonymous variant, ISS international space station, UV ultraviolet light, Ground/UV earth with UV exposure, ISS/UV ISS with UV exposure, ISS/Dark ISS without UV exposure.

Gene name	Feature ID in annotation	Ground/UV	ISS/UV	ISS/Dark
Hypothetical protein	fig 1177.46.peg.2854	1	1	0.5
Hypothetical protein	fig 1177.46.peg.2237	1	0	1
Mobile element protein	fig 1177.46.peg.2271	1	1	0.17
Mobile element protein	fig 1177.46.peg.1675	0	0	0.05
Photosystem II D1 (PsbA) protein	fig 1177.46.peg.5245	0.17	0	0.15
Transposase	fig 1177.46.peg.4085	0.13	0.11	0.13
Vitamin B12 ABC transporter, ATP-binding protein BtuD	fig 1177.46.peg.2110	0.05	0	0

Table 2. Gene-by-gene K_a/K_s ratios. K_a non-synonymous variant, K_s synonymous variant, ISS international space station, UV ultraviolet light, $Ground/UV$ earth with UV exposure, ISS/UV ISS with UV exposure, $ISS/Dark$ ISS without UV exposure. Significant values are in bold.

80kbp across all conditions. As upstream regulatory elements are thought to be located up to 1 Mb from promoter regions of genes, it is possible that these mutations could have an impact on gene regulation, although it would require further work involving 3D conformation of the genome to confirm the full extent of this impact. When examining the specific base pair substitutions occurring in intergenic regions, a pattern begins to emerge (Table S6). A>G, G>A, C>T, and T>C are the most common reference to allele pairings across all experimental conditions for each respective reference nucleotide. A>G and T>C pairings would indicate increased GC content, and the reverse is true for C>T and G>A as it would decrease GC content. Furthermore, there are special cases of multiple base pair substitution in ISS/Dark and ISS/UV with 10 and 9 such instances respectively with Ground/UV having none, although no clear pattern emerges among these mutations.

These results suggest that these loci were preferentially altered. Variant profiles displayed alignment across experimental conditions, indicating the existence of variant hot spots (Figs. 4b,c, 5, Table S1). We calculated the probabilities of observing these variant patterns. A gene-by-gene analysis showed < 0.01% probability of the variants being random across conditions (Table S4). For genome wide variant space, the probability of the genes of interest randomly and independently emerging across all three conditions was < 10^{-8} .

Protein analysis. Exposing *Nostoc* sp. CCCryo 231-06 to the tested environmental conditions resulted in synonymous and non-synonymous variants. While synonymous variants do not change the encoded protein and thus are thought to be silent, non-synonymous variants have a translational impact, which can be quantified on how they impact the encoded protein. Most of the variants detected represented synonymous changes. To determine whether the non-synonymous changes were likely to impact the protein structure of these factors, we performed a tertiary structure protein prediction analysis (RaptorX Structure Prediction). We focused on the photosystem II D1 (PsbA) protein encoding gene because it was the gene with the highest overall number of variants across the 3 experimental conditions (Table S3). PsbA exhibits fast, light-dependent turnover that is typically related to the repair of photo-inactivated PSII complexes. Previous studies suggest the differences in the coding region of the respective *psbA* gene are assigned to different intrinsic sensitivities of PSII complexes containing D1:1 or D1:2 to photo damage²⁸. The predicted structures of the PsbA protein were compared between different experimental conditions and are presented in Fig. 6. For the samples on the ISS, non-synonymous variants of the protein occurred in samples kept in the dark, but not in those that experienced UV exposure (Figs. 4b, 6a). This suggests that the combinatorial nature of ionizing cosmic and UV radiations in ISS/UV may not have the same effect as each type of radiation alone. Compared with the samples exposed to UV radiation on the ISS, multiple variants emerged in the PsbA protein in the samples on the ground (Figs. 4b, 6b), which may indicate that the variants were due to unique aspects of Ground/UV, such as higher UV radiation levels (Table S1). For samples in all experimental conditions, no notable differences in protein structures were observed, except in the protein termini as marked in Fig. S1.

Sequence alignment of the PsbA protein is shown in Fig. S2. The alignment of the nucleotide sequences of the *psbA* gene (Fig. S3) across all conditions identified specific non-synonymous base pair substitutions shown in Fig. S1. Note that there were no insertions or deletions arising from different experimental conditions. Sequence alignment of a putative hemagglutinin-related protein is shown in Fig. S4. Sequence alignment of the nucleotide sequences of the putative hemagglutinin-related gene (Fig. S5) shows that only a few synonymous variants besides the non-synonymous variants highlighted in Fig. S4 emerged. Sequence alignment of a high light inducible protein (HLIP) family gene is shown in Fig. S6. In the sequence alignment of the nucleotide sequences of the HLIP family gene (Fig. S7), there was only a single non-synonymous change in the amino acid sequence. These results were found for all three experimental conditions.

Discussion

The high purity of the returned samples that travelled through space on the ISS, i.e. the apparently reduced abundance of contaminating (non-*Nostoc*) bacteria in the original culture, indicates that the UV radiation in the near-Earth environment, exposure to galactic cosmic rays, trapped radiation belt particles, solar energy particles and secondary particle radiation²⁹ likely exert complex effects that inflict lethal damage to a wide range of microorganisms. Compared to the simulated experimental conditions on Earth, radiation in space appeared to challenge the survival of microorganisms that are less resilient to these conditions. *Nostoc* species are capable of adapting to UV exposure to prevent radiation damage by producing UV-absorbing substances including

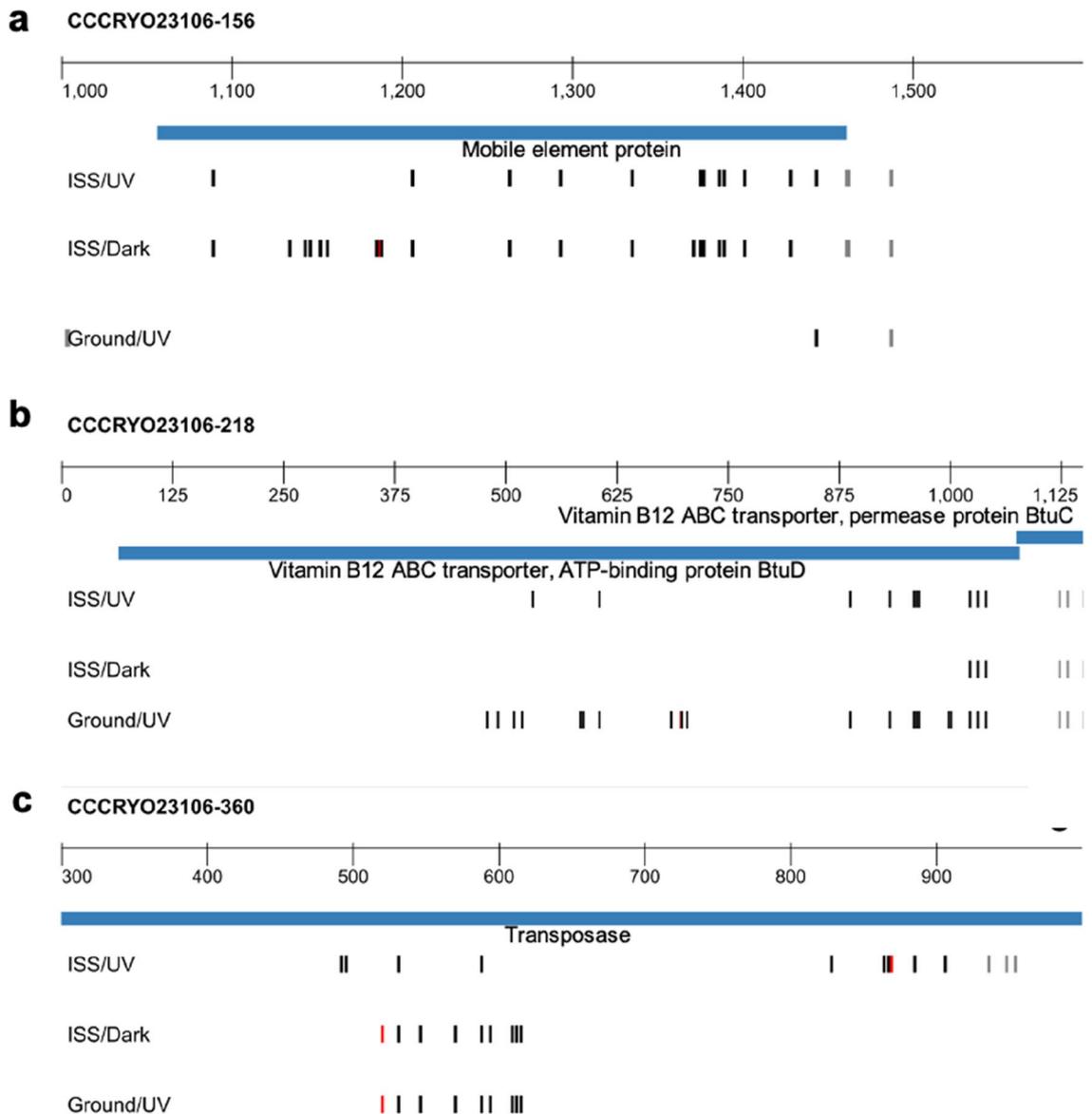


Figure 5. Selected synonymous (black) and non-synonymous (red) *Nostoc* sp. CCCryo 231-06 in BG11 medium variant profiles, where non-synonymous variants with respect to the consensus genome result in a change in amino acid. **(a)** Mobile element protein variant profile showing a high similarity between ISS/Dark and ISS/UV conditions. **(b)** BtuD variant profile showing similar variants for all three conditions, but more similarity between Ground/UV and ISS/UV throughout the whole sequence. **(c)** Transposase variant profile Ground/UV and ISS/Dark display the same sequence, while ISS/UV displays a different variant. ISS international space station, UV ultraviolet light, ISS/UV ISS with UV exposure, ISS/Dark ISS without UV exposure, Ground/UV earth with UV exposure.

shinorines, carotenoids or mycosporins³⁰ in their natural habitats under increased irradiation. In addition, *Nostoc* sp. produces extracellular polysaccharides in response to stress conditions, enhancing its tolerance to desiccation and freezing temperatures³¹ in such extreme locations as rock surfaces, cold mountain and alpine soils or permafrost. These factors could play a significant role in maintaining the cellular structure and DNA integrity of the *Nostoc* sp. on BG11 which led to higher genome coverage.

Despite high preservation of the genome of *Nostoc* sp., we detected non-random variant spots and patterning in genes involved in biofilm production and photosynthetic activity, hemagglutinin and *psbA* in particular. Hemagglutinin-related genes mediate contact between cells, leading to colony formation and biofilm maturation³² protecting *Nostoc* colonies from desiccation. It is not clear if the identified associated loci have any effect on the colony formation and cell contact. However, colony-forming and growth experiments performed after the experiment in space (Fig. S8) showed no obvious phenotypic effects caused by the observed variants. We observed that colonies from ISS/UV conditions initially appeared bleached along the outer parts of the colonies, but only on BG11 medium or lunar regolith, which we attribute to excessive radiation and bleaching of the photosynthetic pigments with likely damage to the photosystem. Interestingly, this was not observed in colonies embedded

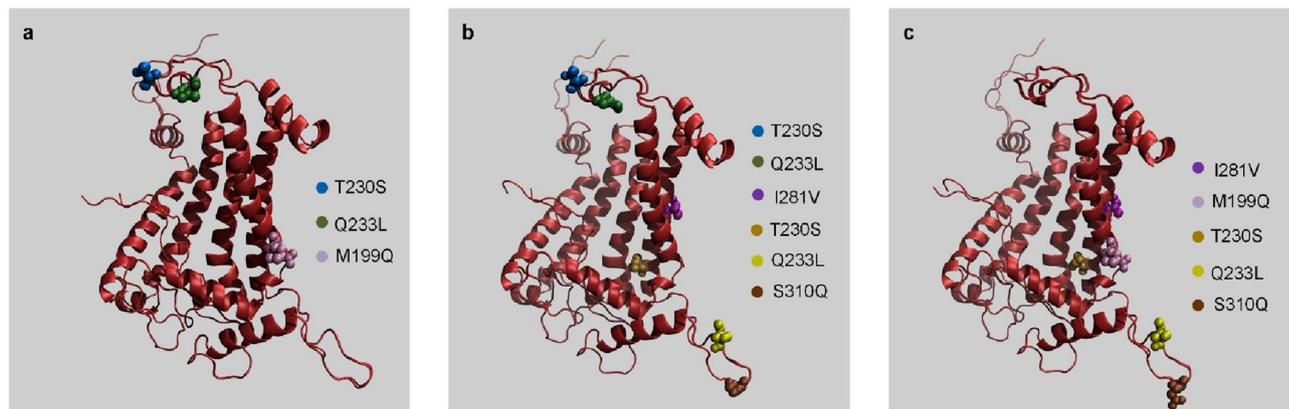


Figure 6. Non-synonymous sequence variants between the amino acid sequences from all three exposure conditions, overlaid on the respective predicted protein structures for the Photosystem II D1 (PsbA) protein. The marked locations represent the position of a non-synonymous variant between the sequences for the corresponding pair of conditions: (a) non-synonymous variants present in ISS/Dark but not ISS/UV conditions, (b) non-synonymous variants present in Ground/UV but not ISS/UV conditions, and (c) non-synonymous variants present in Ground/UV but not ISS/Dark conditions. ISS international space station, UV ultraviolet light, ISS/Dark ISS without UV exposure, ISS/UV ISS with UV exposure, Ground/UV earth with UV exposure.

within the martian regoliths (Mars-UV conditions). Colonies from those samples, no matter if from UV or dark conditions, looked intensely pigmented after rehydration and showed regrowth, young filaments and budding of new colonies, as well as heterocysts that are responsible for fixation of aerial nitrogen (Fig. S8). Differences in morphology, if at all, were noticeable between the substrates (BG11/Lunar versus P-MRS/S-MRS), but not between ground, dark or UV cultures from the ISS and thus, cannot be attributed to the variants detected. Differences in variants across substrates were unable to be examined accurately due to low genome coverage on Lunar and P-MRS/S-MRS analogs, which may have explained further these differences in morphology.

Photosystem-associated variants appeared in samples on BG11 both on the ISS and on Earth. The variants appeared in the *psbA* gene, which codes for the D1 polypeptide of the photosystem II (PS II) reaction center complex and is found in all photosynthetic organisms that perform oxygenic photosynthesis³³. The PS II reaction center in general, is essential for the light reaction, fundamental for the conversion of light energy into chemical energy, the photolysis of water and the start of the electron chain reaction through the membranes of photosynthesizing organisms. This leads to the formation of valuable ATP necessary for the subsequent reactions in the Calvin cycle and fixation of CO₂ derived carbon into carbohydrates (e.g. glucose, starch and other energy storage polysaccharides).

Based on our data, most of the genomic alterations observed were synonymous, and therefore no impact on survivability would be expected as a result. However, we cannot assess the survivability impact of the non-synonymous mutations, which primarily impacted protein termini structure. Follow up functional studies on these changes would have to be performed in order to make such an assessment. Space control samples that stayed on Earth were kept under Mars simulated conditions during the ground-based simulation, experiencing higher UV irradiation (up to 439 kJ/m² of UV_{200–400 nm}) than UV exposure at the ISS (up to 252 kJ/m² of UV_{200–400 nm})^{7,34}. During the ground-based simulation on Earth all three trays of the EXPOSE-R2 facility were exposed to an average UV fluence calculated for the wavelengths 200–400 nm by RedShift to represent the total irradiation fluences applied using a solar simulator (SOL2000, Dr. Hönle GmbH)³⁴. However, individual compartments on the ISS facility experienced widely different UV fluences⁷ due to complex shading and positions of the ISS in relation to the sun, which could not be reproduced individually for the ground experiment. Phycobilisomes, the light-harvesting antennas in cyanobacteria, play a significant role in photo- and radiation protection³⁵. Specifically, cyanobacteria are capable of uncoupling phycobilisomes from photosynthetic membranes which could be a first step in preventing radiation damage and overexcitation of the photosystems. Besides, stress conditions can also trigger photoprotection. These protective mechanisms prevent damage to the photosystems in cyanobacteria. However, we are not able to conclude if these variants would reduce or enhance the photosynthetic activity of the cell if fixed.

Comparing the final UV dose measured on the ISS for space and Mars samples with those on the ground (Table S1), the samples on Earth experienced both much stronger UV radiation intensity and final doses compared to the samples on the ISS. The samples on the ISS also received some shadowing from solar panels and other ISS modules while the ISS orbited the Earth. We note the extensive number of predicted changes in the *psbA* gene across all three exposures considered here, whereas neighboring genes were not as altered (with enough sequencing coverage for variant calling). This is possibly because the genes may have been preferentially damaged while in the desiccated state.

Due to the anhydrobiotic conditions during our experimental setting, we do not expect that significant biological activity, replication, or DNA repair activities took place. Therefore, our results are most likely solely a reflection of DNA damage occurred during the experimental conditions, as has been verified by other research groups under similar settings^{36–38}. Single cell amplification and sequencing technologies have been recently deployed to characterize the “damagenome” profile in human cells³⁹. This work represents a parallel technological

effort to characterize genomic alterations in bacterial single cells. It is important to note the different external stresses across all experimental conditions. While ISS/UV and Ground/UV had different UV radiation exposure, ISS/UV and ISS/Dark shared exposure to cosmic radiation and microgravity. This provides some insight into the drivers of variant hotspots. ISS/UV and ISS/Dark were similar in the mobile element variant profile, while the *btuD* variant profile showed a stark similarity between Ground/UV and ISS/UV conditions. This gene is found in the vitamin B12 complex which has known sensitivity to UV exposure⁴⁰, with reduced exposure under the ISS/Dark condition. Likewise, the *psbA* variant profile was exacerbated under Ground/UV conditions, which is also consistent with known sensitivity and degradation of the photosystem II, which includes inhibition of photosynthetic electron flow when under UV exposure²⁸ and likely exposure to radiation damage. Exposure to UV irradiation has also been observed to drive differential growth and enzymatic activity in cyanobacteria, including other species of *Nostoc*⁴¹.

Interestingly, the transposase gene showed high similarity between variant profiles in ISS/Dark and Ground/UV, which shared no particular drivers for damage, although transposases are known to be reactive to both ionizing⁴² and UV radiation⁴³. It is not clear why the variant profile between these 2 conditions shared such similarity, while ISS/UV conditions did not. It is likely that the combined effect of cosmic radiation and UV exposure in ISS/UV conditions generates a different outcome that is not simply combinatorial in nature. There are multiple explanations for observing consistent variants even under low probability. The flanking regions around the altered sites can increase the likelihood of point variations⁴⁴, which might be synonymous or non-synonymous. Neighboring base pairs mechanistically influence the frequency and type of variants in different genomic landscapes. There also is a preferential codon usage, particular to the *Nostoc* species being examined⁴⁵. The frequency of specific codon variant profiles as well as the DNA strand level mechanisms of the flanking motifs may contribute to the mechanisms for repeated variant events on a genomic level.

The combinatorial probabilities of genome-wide and gene-by-gene variants were used to evaluate the likelihood of observing certain variant profiles. The significance of these probabilities reduced the likelihood that the variant profiles observed across all three conditions were random. We therefore conclude that exposure of this *Nostoc* species to space conditions generates variants with preferential synonymous alterations of the biofilm and photosynthetic apparatus at the single cell level. Although all of us are used to contemplating a random mutation landscape upon which natural selection acts to select the fittest, our results suggest that the landscape may not always be random at the single cell level and under these unusual experimental conditions. Many variables may contribute to this outcome including a combination of DNA structure protected/vulnerable zones, bias in alterations caused by desiccation, UV and cosmic radiation for extended periods of time. Whatever those causes may be, the non-random nature of the alterations narrow the acting field for natural selection/genetic drift effects to subsequently occur when cells are revived. Our results suggest that under these particular experimental conditions, the non-random and largely synonymous alterations would create a scenario favoring stabilizing selection of the altered photosynthetic and biofilm genes upon revival.

Conclusions

We have found that *Nostoc* CCCryo 231-06 displayed single cell variant profiles consistent with non-random genomic alterations on the ISS and simulated UV exposure on ground (Earth). The genetic variant hotspots were clustered in filamentous hemagglutinin and Photosystem II D1 loci. The variant patterns were repeatedly observed, diminishing the explanatory power of stochastic effects. Based on our analyses, the combined effect of complex cosmic radiation and UV exposure may result in synergistic damage effects, with a higher number of synonymous variants with simultaneous exposure to cosmic and UV radiations. The cause(s) and evolutionary implications of the non-random synonymous genomic substitutions observed at the single cell level under long-term cosmic exposure warrants further investigation, and may indicate that we have to rethink how evolution takes place at the single cell, and also population level.

Data availability

Genome data for *Nostoc* sp. CCCryo 231-06 has been deposited at NCBI (BioProject accession number PRJNA721463). The Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAHCSU010000000. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials S1.

Received: 24 March 2022; Accepted: 15 July 2022

Published online: 22 July 2022

References

- Schulze-Makuch, D., & Irwin, L.N. Astrobiological potential of planetary bodies within the solar system. *Life in the Universe*: Springer; 2018. pp. 203–228.
- Cockell, C. S. *et al.* Habitability: A review. *Astrobiology* **16**(1), 89–117 (2016).
- Sancho, L. G. *et al.* Lichens survive in space: results from the 2005 LICHENS experiment. *Astrobiology* **7**(3), 443–454 (2007).
- de La Torre, R. *et al.* Survival of lichens and bacteria exposed to outer space conditions—results of the Lithopanspermia experiments. *Icarus* **208**(2), 735–748 (2010).
- Rabbow, E. *et al.* The astrobiological mission EXPOSE-R on board of the International Space Station. *Int. J. Astrobiol.* **14**(1), 3–16 (2015).
- Potts, M. *Nostoc* 465–504 (Springer, 2000).
- De Vera, J.-P. *et al.* Limits of life and the habitability of Mars: The ESA space experiment BIOMEX on the ISS. *Astrobiology* **19**(2), 145–157 (2019).
- Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: Current state of the science. *Nat. Rev. Genet.* **17**(3), 175–188 (2016).

9. Low, A. J., Koziol, A. G., Manninger, P. A., Blais, B. & Carrillo, C. D. ConFindr: Rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. *PeerJ* **7**, e6995 (2019).
10. Liu, Y. & Walther-Antonio, M. Microfluidics: A new tool for microbial single cell analyses in human microbiome studies. *Biomicrofluidics* **11**(6), 061501 (2017).
11. Liu, Y. *et al.* The development of an effective bacterial single-cell lysis method suitable for whole genome amplification in microfluidic platforms. *Micromachines* **9**(8), 367 (2018).
12. Liu, Y., Yao, J. & Walther-Antonio, M. Whole genome amplification of single epithelial cells dissociated from snap-frozen tissue samples in microfluidic platform. *Biomicrofluidics* **13**(3), 034109 (2019).
13. Liu, Y., Jeraldo, P., Jang, J., Eckloff, B., Jen, J., & Walther-Antonio, M. Bacterial single cell whole transcriptome amplification in microfluidic platform shows putative gene expression heterogeneity. *Anal. Chem.* (2019).
14. Liu, Y. *et al.* Whole genome sequencing of cyanobacterium *Nostoc* sp CCCryo 231–06 using microfluidic single cell technology. *Science*. **25**(5), 4291 (2022).
15. <https://sourceforge.net/projects/bbmap/>. BBMap (2019).
16. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**(10), 1674–1676 (2015).
17. Laczny, C. C. *et al.* BusyBee Web: Metagenomic data analysis by bootstrapped supervised binning and annotation. *Nucleic Acids Res.* **45**(W1), W171–W179 (2017).
18. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**(7), 1043–1055 (2015).
19. Parks, D. H. *et al.* Recovery of nearly 8000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**(11), 1533 (2017).
20. Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., & Chaumeil, P.-A., *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* (2018).
21. Davis, J. J. *et al.* The PATRIC bioinformatics resource center: Expanding data and analysis capabilities. *Nucleic Acids Res.* **48**(D1), D606–D612 (2020).
22. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., & DePristo, M.A. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* (2010).
23. Ruden, D. M. *et al.* Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program. *SnpSift. Front. Genet.* **3**, 35 (2012).
24. Yang, Z. & Bielawski, J. P. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**(12), 496–503 (2000).
25. Källberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**(8), 1511 (2012).
26. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8 (2015).
27. Pullerits, K. *et al.* Impact of UV irradiation at full scale on bacterial communities in drinking water. *NPJ Clean Water.* **3**(1), 1–10 (2020).
28. Golden, S. S. Light-responsive gene expression in cyanobacteria. *J. Bacteriol.* **177**(7), 1651 (1995).
29. Dachev, T. P. *et al.* Overview of the ISS radiation environment observed during the ESA EXPOSE-R2 mission in 2014–2016. *Space Weather* **15**(11), 1475–1489 (2017).
30. Sinha, R. P., Klisch, M., Helbling, E. W. & Häder, D.-P. Induction of mycosporine-like amino acids (MAAs) in cyanobacteria by solar ultraviolet-B radiation. *J. Photochem. Photobiol. B* **60**(2–3), 129–135 (2001).
31. Tamaru, Y., Takani, Y., Yoshida, T. & Sakamoto, T. Crucial role of extracellular polysaccharides in desiccation and freezing tolerance in the terrestrial cyanobacterium *Nostoc commune*. *Appl. Environ. Microbiol.* **71**(11), 7327–7333 (2005).
32. Guilhabert, M. R. & Kirkpatrick, B. C. Identification of *Xylella fastidiosa* antivirulence genes: hemagglutinin adhesins contribute to *X fastidiosa* biofilm maturation and colonization and attenuate virulence. *Mol. Plant Microb. Interact.* **18**(8), 856–868 (2005).
33. Nixon, R. A. & Sihag, R. K. Neurofilament phosphorylation: A new look at regulation and function. *Trends Neurosci.* **14**(11), 501–506 (1991).
34. Rabbow, E. *et al.* EXPOSE-R2: The astrobiological ESA mission on board of the International Space Station. *Front. Microbiol.* **8**, 1533 (2017).
35. Klementiev, K. E. *et al.* Radioprotective role of cyanobacterial phycobilisomes. *Biochim. Biophys. Acta* **2**, 121–128 (1860).
36. Mosca, C. *et al.* Revival of anhydrobiotic cyanobacterium biofilms exposed to space vacuum and prolonged dryness: implications for future missions beyond low Earth orbit. *Astrobiology* **21**(5), 541–550 (2021).
37. Kawaguchi, Y. *et al.* DNA damage and survival time course of deinococcal cell pellets during 3 years of exposure to outer space. *Front. Microbiol.* **11**, 2050 (2020).
38. Gladyshev, E. & Meselson, M. Extreme resistance of bdelloid rotifers to ionizing radiation. *Proc. Natl. Acad. Sci.* **105**(13), 5139–5144 (2008).
39. Zhu, Q., Niu, Y., Gundry, M. & Zong, C. Single-cell damagenome profiling unveils vulnerable genes and functional pathways in human genome toward DNA damage. *Sci. Adv.* **7**(27), 3329 (2021).
40. Heaven, G. H., Holiday, E. R., Johnson, E. A., Ellis, B. & Petrow, V. The chemistry of anti-pernicious anaemia factors: Part VI: The mode of combination of component alpha in vitamin B12. *J. Pharm. Pharmacol.* **2**(12), 944–955 (1950).
41. Sinha, R., Kumar, H. D., Kumar, A. & Häder, D. Effects of UV-B irradiation on growth, survival, pigmentation and nitrogen metabolism enzymes in cyanobacteria. *Acta Protozool.* **34**, 187 (1995).
42. Lynch, M. *et al.* Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* **17**(11), 704–714 (2016).
43. Aleshkin, G. I., Kadzhaev, K. V. & Markov, A. P. High and low UV-dose responses in SOS-induction of the precise excision of transposons tn1, Tn5 and Tn10 in *Escherichia coli*. *Mutat. Res.* **401**(1–2), 179–191 (1998).
44. Zhu, Y., Neeman, T., Yap, V. B. & Huttley, G. A. Statistical methods for identifying sequence motifs affecting point mutations. *Genetics* **205**(2), 843–856 (2017).
45. Knight, R. D., Freeland, S. J. & Landweber, L. F. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.* **2**(4), 10 (2001).

Acknowledgements

This research was funded by the following sources: Marina Walther-Antonio and Yuguang Liu acknowledge The Ivan Bowen Family Foundation. This project was supported by CTSA Grant Number KL2TR002379 from the National Center for Advancing Translational Science (NCATS). Contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. Dirk Schulze-Makuch and Jean-Pierre de Vera acknowledge support from ERC Advanced Grant “HOME” (# 339231), and in addition, Jean-Pierre de Vera acknowledges support from ESA for the BIOMEX project (ESA-ILSRA 2009-0834) and DLR for the DLR-FuW-Project BIOMEX (2474128). Mickael Baqué acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG—German Research Foundation) for the project “Raman Biosignatures for Astrobiology Research” (RaBi-oFAM; project number: 426601242). Charles S. Cockell was supported by the Science and Technology Facilities Council (STFC), Grant No. ST/M001261/1. In addition, we thank the Microbiome Program and the Center for

Individualized Medicine at Mayo Clinic for their support, and Dr. Alexander Revzin at Mayo Clinic for granting us the access to his microfabrication facilities.

Author contributions

Experimental methodology, Y.L. and M.W.-A.; Bioinformatic methodology, P.J.; Formal analysis, P.J., Y.L. and W.H.; Library preparation and sequencing, S.M.; Funding acquisition, M.W.-A.; Initial idea, D.S.-M.; Project administration, D.S.-M., J.-P.V., C.C., T.L., M.B. and M.W.-A.; Resources, D.S.-M., J.-P.V., C.C., T.L., M.B. and M.W.-A.; Supervision, D.S.-M., J.-P.V., C.C., T.L., M.B. and M.W.-A.; Sequencing supervision, J.J., B.E.; writing—original draft, Y.L., P.J., and W.H.; writing—review & editing, D.S.-M., J.-P.V., C.C., T.L., M.B. and M.W.-A.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16789-w>.

Correspondence and requests for materials should be addressed to D.S.-M. or M.W.-A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022