# Nucleic Acid Extraction and Sequencing from Low-Biomass Synthetic Mars Analog Soils for In Situ Life Detection

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

| Citation               | Mojarro, Angel, et al. “Nucleic Acid Extraction and Sequencing from Low-Biomass Synthetic Mars Analog Soils for In Situ Life Detection.” Astrobiology 19, 9 (September 2019): 1139–52. © 2019 Angel Mojarro et al. |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| As Published           | http://dx.doi.org/10.1089/AST.2018.1929                                                                                                                                                           |
| Publisher              | Mary Ann Liebert Inc                                                                                                                                                                             |
| Version                | Final published version                                                                                                                                                                           |
| Citable link           | https://hdl.handle.net/1721.1/124639                                                                                                                                                              |
| Terms of Use           | Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.                                       |
Nucleic Acid Extraction and Sequencing from Low-Biomass Synthetic Mars Analog Soils for In Situ Life Detection

Angel Mojarro,1 Julie Hachey,2 Ryan Bailey,3 Mark Brown,3 Robert Doebler,3 Gary Ruvkun,4 Maria T. Zuber,1 and Christopher E. Carr1,4

Abstract

Recent studies regarding the origins of life and Mars-Earth meteorite transfer simulations suggest that biological informational polymers, such as nucleic acids (DNA and RNA), have the potential to provide unambiguous evidence of life on Mars. To this end, we are developing a metagenomics-based life-detection instrument which integrates nucleic acid extraction and nanopore sequencing: the Search for Extra-Terrestrial Genomes (SETG). Our goal is to isolate and sequence nucleic acids from extant or preserved life on Mars in order to determine if a particular genetic sequence (1) is distantly related to life on Earth, indicating a shared ancestry due to lithological exchange, or (2) is unrelated to life on Earth, suggesting convergent origins of life on Mars. In this study, we validate prior work on nucleic acid extraction from cells deposited in Mars analog soils down to microbial concentrations (i.e., \(10^4\) cells in 50 mg of soil) observed in the driest and coldest regions on Earth. In addition, we report low-input nanopore sequencing results from 2 pg of purified \(Bacillus subtilis\) spore DNA simulating ideal extraction yields equivalent to 1 ppb life-detection sensitivity. We achieve this by employing carrier sequencing, a method of sequencing sub-nanogram DNA in the background of a genomic carrier. After filtering of carrier, low-quality, and low-complexity reads we detected 5 \(B. subtilis\) reads, 18 contamination reads (including \(Homo sapiens\)), and 6 high-quality noise reads believed to be sequencing artifacts. Key Words: Life-detection instruments—Nucleic acids—Origins of life—Panspermia—Mars. Astrobiology 19, 1139–1152.

1. Introduction

Major strides in understanding the origins of life and meteorite transfer simulations support the notion that life on Mars, if it ever existed, may share a common genesis or perhaps share a common ancestry with life on Earth. Namely, analogous prebiotic environments (Johnson et al., 2008; Morris et al., 2010; Stoker et al., 2010; Grotzinger et al., 2014; Ranjan and Sasslov, 2017; Ranjan et al., 2017), molecular feedstocks (e.g., hydrogen cyanide) (Brack and Pillinger, 1998; Parker et al., 2011; Adcock et al., 2013), and plausible abiotic reactive pathways predicted on Earth and applicable on Mars may have resulted in parallel origin events in accordance with the RNA-world hypothesis (Powner et al., 2009, 2010; McKay, 2010; Ritson and Sutherland, 2012; Benner and Kim, 2015; Patel et al., 2015; Stairs et al., 2017). This hypothesis suggests that past or present martian life may have utilized known building blocks (e.g., nucleic acids, sugars, amino acids) and closely resembled life as we know it. Moreover, nonsterilizing lithological exchange between Mars and Earth from impact ejecta produced during the presumed Late Heavy Bombardment period (Gomes et al., 2005; Boehnke and Harrison, 2016) may have transported viable microbes between planets (Weiss, 2000; Shuster, 2005; Horneck et al., 2008; Abramov and Mojzsis, 2009), resulting in ancestrally related life (Isenbarger et al., 2008).

To test these hypotheses, we are developing the Search for Extra-Terrestrial Genomes (SETG) life-detection instrument for in situ extraction and nanopore sequencing of nucleic acids (Carr et al., 2016, 2017) from extant or preserved life on Mars. Assuming a convergent adoption of
nucleic acids as the unitary solution for genetic information storage and transmission (e.g., DNA or RNA), long-read nanopore sequencing could be capable of detecting nonstandard nucleic acids (Carr, 2016; Carr et al., 2017) possibly endemic to life on Mars (Ranjani et al., 2017). Furthermore, long-read sequencing is of particular significance for taxonomic identification at or below the species level (Greninger et al., 2015; Quick et al., 2015; Benitez-Paiz et al., 2016; Brown et al., 2017), which would permit the detection of microbial forward contamination. For instance, in the case of ancestrally related life, comparing sequence data detected on Mars to conserved genes (Makarova et al., 1999; Harris, 2003) on Earth (i.e., the ribosome) (Woese et al., 1975) could unambiguously discriminate forward contamination from a true life detection (Isenberger et al., 2008). Conversely, detecting a genetic sequence unlike anything found on Earth (including nonstandard bases) could signify a second genesis and perhaps indicate that nucleic acid–based life is common.

SETG operates by first extracting and isolating nucleic acids (DNA or RNA) from cells in solid or liquid samples using a modified Claremont BioSolutions solid-phase Purelyse bacterial genomic DNA extraction kit. Prior studies have utilized OmniLyse, which is Purelyse without extraction buffers, to lyse cells in an RNA extraction module aboard the International Space Station (Parra et al., 2017). Long-read sequencing is then conducted using the Oxford Nanopore Technologies MinION, which sequences nucleic acids via ion current monitoring (Lu et al., 2016) and has been validated to sequence DNA in microgravity (McIntyre et al., 2016; Castro-Wallace et al., 2017), lunar and martian gravity (Carr/Zuber, unpublished data), and under simulated martian temperature and pressure (Carr et al., 2019). Readers interested in the current status of SETG, including technology development, demonstration, and subsystem development, should refer to the works of Carr et al. (2017) and Bhattachar (2018).

Our goal for SETG is to be capable of analyzing a variety of environmental samples relevant to the search for life (related or otherwise to life on Earth) on Mars. However, complex soils, especially those containing iron oxides found on Mars (Bell et al., 2000), greatly inhibit nucleic acid extraction from cells due to competitive adsorption onto mineral surfaces (Jiang et al., 2012; Hurt et al., 2014) or destruction due to hydroxyl radicals (Inlay and Linn, 1988; Gates, 2009) during cell lysis (Mojarro et al., 2017b). Additional inhibition may occur in the presence of phyllosilicates (Greaves and Wilson, 1969; Melzak et al., 1996; Trevors, 1996) and salts (Henneberger et al., 2006) due to adsorption and/or DNA hydrolysis, denaturation, and depurination at alkaline to acidic conditions (Gates, 2009). These interactions are further exacerbated in low-biomass environments and by nanopore sequencing, which presently requires substantial (1 µg or 400 ng of high molecular weight) input DNA not likely to be acquired from recalcitrant soils (without amplification) due to limitations related to sequencing efficiency and nanopore longevity (Mojarro et al., 2018). Roughly five in a million nucleobases are sequenced (R9.4 flowcells and chemistry) while nanopores are expended before detection of sub-nanogram input DNA (Mojarro et al., 2017a, 2018) without library preparation modifications.

Extensive literature exists on methods used to mitigate soil-DNA interactions and yield quantifiable DNA from various soil species (e.g., Takada-Hoshino and Matsumoto, 2004; Barton et al., 2006; Henneberger et al., 2006; Herrera and Cockell, 2007; Direito et al., 2012; Lever et al., 2015). Competitive binders or blocking agents such as RNA, random hexamer primers, sodium pyrophosphate, and skim milk (e.g., Takada-Hoshino and Matsumoto, 2004) are applied to inhibit mineral adsorption sites while dialysis, desalting, or chelation (e.g., Barton et al., 2006) is employed to precipitate or flush soluble metals and salts from soils. Our previous work from Mojarro et al. (2017b) focused on adapting the standard Purelyse bacterial gDNA extraction kit to isolate DNA from Mars analog soils doped with tough-to-lyse spores of Bacillus subtilis. In that study, we concluded that with soil-specific mitigation strategies (e.g., desalting and competitive binders), we may be able to achieve adequate DNA extraction yields that are of sufficient purity for downstream sequencing.

In this study, we now focus on validating the modified Purelyse kit with low-biomass Mars analog soils containing cell densities observed in the driest regions of the Atacama Desert (Navarro-Gonzalez et al., 2003) and McMurdo Dry Valleys (Goordial et al., 2016). We perform baseline and modified DNA extractions from 50 mg of Mars analog soil containing 10⁴ spores of B. subtilis and vegetative Escherichia coli cells, respectively. These experiments stress the importance of the mitigation strategies used to reduce detrimental soil-DNA interactions (e.g., adsorption to mineral surfaces, DNA destruction) and conclude in the development of a “near-universal” extraction protocol for low-biomass Mars analog soils. As in our previous experiments (Carr et al., 2016, 2017), we require DNA extraction yields of at least 5% in order to achieve a minimum sensitivity target of 1 ppb for life detection (i.e., 2 pg of DNA from 10⁴ cells in 50 mg of soil).

Lastly, we investigate low-input nanopore sequencing by experimenting with 2 pg of purified B. subtilis spore DNA equivalent to a 5% DNA yield from 10⁴ spores. For many environmental samples, the total extractable DNA is far below the current input requirements of nanopore sequencing, preventing sample-to-sequence metagenomics from low-biomass or recalcitrant samples. On Mars, whole genome amplification could result in biasing microbial population results (Sabina and Leamon, 2015) while targeted amplicon sequencing such as 16S rRNA would assume a shared ancestry between Mars and Earth and could reduce taxonomic resolution (Poretsky et al., 2014). The absence of amplification is especially important in the case where nonstandard bases have been incorporated into the martian genome, as primers may not function or amplification could theoretically mask an alien signal due to promiscuous base-pairing resulting in an ATGC sequence (Pezo et al., 2014). Here we address these problems by employing carrier sequencing, a method to sequence low-input DNA by preparing the target DNA with a genomic carrier to achieve ideal library preparation and sequencing stoichiometry without amplification (Raley et al., 2014; Mojarro et al., 2018). We then use CarrierSeq (https://github.com/amojarro/carrierseq, Mojarro et al., 2018), a sequence analysis script used to identify low-input target (unknown) reads from the genomic carrier, and analyze the results in the context of life detection.

2. Materials and Methods

2.1. Mars analog soils

A total of six Mars analog soils and one lunar basalt analog (Orbitec, JSC-1A) (McKay et al., 1993) were utilized to
develop our extraction protocols. Five Mars analog soils were produced in accordance with in situ mineralogical and geochemical measurements collected by rover and lander missions as described by Schuerger et al. (2012, 2017). These soils represent various potentially habitable palaeoenvironments of astrobiological significance, for instance, ancient hydrothermal spring salt deposits in Gusev Crater (Ming et al., 2006; Ruff and Farmer, 2016). The five soils represent (1) the highly oxidized global aeolian dust (Bell et al., 2000; Ming et al., 2008a), (2) salt-rich (Burroughs subclass) soils of Gusev Crater (Ming et al., 2006; Morris et al., 2008), (3) jarosite-containing acidic soils (Paso Robles class) of Meridiani Planum (Klingelhofer, 2004; Morris et al., 2006; Ming et al., 2008a), (4) carbonate-rich alkaline Viking lander soils (Clark et al., 1982; Wänke et al., 2001), and (5) perchlorate-rich Phoenix lander soils (Ming et al., 2008b). In addition, we include a commercially available (6) aeolian spectral analog (Orbitech, JSC-Mars-1A) (Allen et al., 1998). We henceforth refer to these Mars analog soils as aeolian, salt, acid, alkaline, perchlorate, and JSC, respectively. The lunar analog represents our unaltered soil control and is referred to as basalt. Furthermore, the basalt sample closely resembles the reduced martian subsurface (Taylor, 2013; Grotzinger et al., 2014), which could theoretically harbor a deep modern biosphere (Jones et al., 2011). All soils were heat-sterilized at 130°C for 48 h prior to any experimentation in a Fisher Science Isotemp 282A vacuum oven, cooled overnight, and stored in air-tight borosilicate bottles (Corning, 1395-100) at room temperature (Schuerger et al., 2012). Details concerning the exact methods used to synthesize these soils can be found in the works of Schuerger et al. (2012) and Mojarro et al. (2017b).

2.2. Bacillus subtilis spores

Spore suspensions of Bacillus subtilis (ATCC 6633) similar to those used in clean-room sterilization effectiveness (Friedline et al., 2015) and bacteriostasis testing under Mars-like conditions (Kerney and Schuerger, 2011; Schuerger et al., 2017, 2003) were acquired from Crosstex (Part# SBS-08) to represent a worst-case DNA extraction scenario of a tough-to-lyse organism. Spore suspensions were DNase-treated (New England Biolabs, M0303L) in order to remove any extracellular DNA and counted using a viable spore assay on lysogeny broth agar plates (Mojarro et al., 2017b). We assume a single genome copy per DNase-treated spore of B. subtilis for calculating DNA yield (DNA_{out}/DNA_{in}).

2.3. Vegetative Escherichia coli cells

Vegetative Escherichia coli (OP50) cell cultures were grown overnight in flasks containing 200 mL of lysogeny broth medium inside an Innova 44 incubator at 37°C and 200 rpm. After 12 h, we measured optical density (OD600) on a DeNovix DS-11+ spectrophotometer, pelleted, and resuspended culture aliquots in phosphate-buffered saline solution (Thermo Fisher, 10010023). An estimated 1000 E. coli cells then underwent colony droplet digital polymerase chain reaction (ddPCR) with single copy metG (Wang and Wood, 2011) primers for absolute genome quantitation. In contrast to spores that contain a single genome copy, copy-number variation may exist in vegetative cells depending on growth stage (Skarstad et al., 1986; Pecoraro et al., 2011). Therefore, ddPCR results allow us to correct for genome copy-number variation (e.g., 2000 genome copies/1000 cells) and calculate DNA yield similar to B. subtilis (DNA_{out}/DNA_{in}).

2.4. DNA yield quantification

Absolute DNA yield quantification is possible through ddPCR paired with single copy primers. In contrast to traditional PCR, ddPCR fractionates one reaction into ~20,000 water-oil emulsion nanodroplets each capable of discrete amplification. When a sample is adequately dilute, Poisson statistics can precisely determine genome copy numbers independent of a standard curve (Brunetto et al., 2014; Yang et al., 2014). Paired with single copy primers and the presence of a DNA-specific fluorescent dye (EvaGreen), the results of a ddPCR reaction are either relatively weakly fluorescent droplets containing primer dimer or relatively highly fluorescent droplets containing an amplified product. We then equate one highly fluorescent droplet to one genome copy. DNA extracted from B. subtilis spores was amplified with spaC primers (forward: TGA GGA AGG ATG GGA CGA CA; reverse: AAC AGA TTG CTG CCA GTC CA) (Hu et al., 2004) while DNA extracted from vegetative E. coli cells was amplified with metG primers (forward: GGT GGA AGC TCT TAA AGA AGA AG; reverse: AGC AGT TTG TCA GAA CCT TCA AC) (Wang and Wood, 2011). Two ddPCR reactions per DNA extraction (one per elution) were performed and quantified on a BioRad QX200 ddPCR system. A total of 5 mL from each extraction elution was prepared in a 20 mL final reaction volume containing 3 mL of molecular-grade water (Thermo Fisher, 10977023), 10 mL of QX200 EvaGreen SuperMix (BioRad, 186-4033), 1 mL of 3.3 mM forward primer, and 1 mL of 3.3 mM reverse primer. All ddPCR reactions were prepared by an Andrew Alliance liquid handling robot within an AirClean 600 PCR Workstation while nanodroplets were generated using the automated QX200 AutoToDG system. Thermocycling conditions were as follows: (1) 95°C for 5 min, (2) 95°C for 30 s, (3) 60°C for 1 min, (4) repeat steps 2–3 40 times, (5) 4°C for 5 min, (6) 90°C for 5 min, (7) hold at 4°C until ready to measure.

2.5. Purelyse standard extraction protocol

The Purelyse bacterial gDNA extraction kit is a miniature (dime-sized) and battery-powered bead-beating lysis device capable of solid-phase nucleic acid extraction without the need for centrifugation. Purelyse kits work by shearing cells open at 30,000 rpm in the presence of a supplied low-pH (~3.5 pH) binding buffer that promotes the binding of negatively charged polymers to the surface of specialty oxide ceramic microbeads (~100 μm). The microbeads are then washed with a lower-concentration binding buffer, and DNA is eluted in a low-salt and high-pH (~8.0 pH) elution buffer. The standard extraction protocol for cell cultures dictates: (1) 2 min lysis with 1× binding buffer solution (100 μL of 8× binding buffer and 700 μL of molecular-grade water), (2) 45 s wash with 3 mL of 0.25× binding buffer wash solution (~100 μL of 8× binding buffer and 2.9 mL of molecular-grade water), (3) two 1 min elutions with 200 μL of 1× elution buffer all at 6V (equivalent to 4 AAA batteries).
2.6. Baseline extractions of spore and vegetative DNA from water

Two dilution series containing $10^5$, $10^6$, and $10^8$ spores of B. subtilis and equivalent vegetative E. coli cells in water were processed using the standard extraction protocol. The E. coli dilution series originated from a single culture prepared in phosphate-buffered saline solution, and DNA yield was corrected for genome copy-number variation using ddPCR.

2.7. Unmodified extractions of spore and vegetative DNA from Mars analog soils

An estimated $1.6 \times 10^4$ spores (about 70 pg of DNA) of B. subtilis were deposited on 50 mg of Mars analog soil and processed following the standard extraction protocol. For vegetative E. coli extractions, we utilized OD600 measurements to deposit approximately $1.6 \times 10^4$ cells on 50 mg of Mars analog soil or equivalent to spore concentrations. However, as E. coli genome copy number may vary (Skarstad et al., 1986; Pecoraro et al., 2011), each E. coli extraction from a Mars analog included accompanying colony ddPCR results used for correcting genome copy-number variation.

2.8. Modified extractions of spore and vegetative DNA from Mars analog soils

To parallel the unmodified extractions, an estimated $1.6 \times 10^4$ spores of B. subtilis and vegetative cells of E. coli were deposited on 50 mg of Mars analog soil, respectively. All samples were processed by using the “near-universal” modified extraction protocol: (1) Suspend sample in 800 μL of 8× binding buffer and vortex gently for 30 s, (2) desalt in a single 100K Amincon Ultra column (Z740183), (3) resuspend the sample in 400 μL of molecular-grade water and 400 μL of 8× binding buffer, (4) add 4−6 μg of random hexamer primers (Promega, C1181) and vortex gently for 30 s, (5) lyse cells and bind DNA with Purelyse at 6.5 V for 2 min, (6) wash with 2.5 mL of 1× binding buffer at 1.5 V for 1 min, (7) elute DNA with 200 μL of heated elution buffer to 70°C for 1 min, (8) repeat step 7 for second elution. Basalt and alkaline samples contained 4 μg of random hexamer primers as we observed a decrease in DNA yield with increasing amounts. We believe this is most likely due to hexamer-induced competitive binding since basalt and alkaline samples have a lower DNA binding affinity (Mojarro et al., 2017b). Meanwhile JSC, acid, salt, aeolian, and perchlorate samples required 6 μg.

2.9. Nanopore sequencing

The Oxford Nanopore Technologies MinION sequencer and flowcell perform single-strand DNA sequencing through monitoring changes in an ionic current produced by the translocation of k-mers through a nano-sized pore (Lu et al., 2016). Nanopore sequencing encompasses library preparation, sequencing, and basecalling. First, library preparation is the process by which genomic DNA is converted into a readable format for the sequencer. This includes adding tethers and a motor protein onto double-stranded DNA which guide and regulate the translocation rate (R9.4, 450 bp/s) through a nanopore. Second, the prepared library is loaded onto the flowcell, and sequencing is initiated. Up to 512 sequencing channels (pores) can be monitored, and raw current (translocation) events are recorded. Lastly, the raw events are then basecalled into corresponding nucleobase assignments or oligomer sequences.

2.10. Bacillus subtilis spore DNA extraction for low-input nanopore sequencing

To simulate an ideal sample-to-sequence scenario, an estimated $2.0 \times 10^8$ spores of B. subtilis in water were processed by using the standard extraction protocol. DNA yield was then quantified using a double-stranded, DNA-specific fluorometric assay (Invitrogen, Qubit dsDNA HS Assay Kit, Q32854) and a Qubit 2.0 fluorometer (Invitrogen, Q32866, limit of detection of 0.1 ng/mL dsDNA). Once the DNA concentration was known, an aliquot of the purified spore DNA was diluted in molecular-grade water to 2 pg/μL and subsequently verified with ddPCR prior to library preparation.

2.11. Low-input sequencing library preparation

The sequencing library was prepared following a modified one-direction (1D) Lambda control experiment protocol (Oxford Nanopore Technologies, SQK-LSK108). The ligation-based sequencing kit advises shearing 1 μg of Lambda genomic DNA to 8 kb fragments and spiking the sheared gDNA with a 3.6 kb positive control (Lambda genome 3′-end amplicon) in order to distinguish library preparation from sequencing failures. However, after validating our library preparation proficiency, we substituted 2 pg of B. subtilis spore DNA purified with Purelyse in lieu of the 3.6 kb positive control and replicated the Lambda control protocol without additional modifications. This substitution results in a low-input carrier library with ideal stoichiometry (Mojarro et al., 2018).

2.12. Sequencing and basecalling

The low-input carrier library was sequenced by a MinION Mk 1B sequencer and R9.4 spot-on flowcell for 48 h on an Apple iMac desktop computer running MinKNOW 1.5.18. The resulting raw nanopore reads were basecalled with Oxford Nanopore Technologies’ Albacore 1.10 offline basecaller.

2.13. Sequence analysis

All reads located in the Albacore workstation folder were compiled into a single fastq file and mapped directly to the B. subtilis reference genome using bwa (Li, 2013). In the context of an unknown sample, all reads were then processed with CarrierSeq (Mojarro et al., 2018), a sequence analysis script for carrier sequencing. CarrierSeq works by identifying all reads not belonging to the carrier (Lambda), applies quality-control filters, and implements a Poisson test to identify likely nanopore sequencing artifacts known as high-quality noise reads (HQNRs), which presumably originate from malfunctioning nanopores. The final subset or “target reads” should therefore only contain B. subtilis and likely contamination (Mojarro et al., 2018).

3. Results

3.1. Baseline extractions of spore and vegetative cell DNA from water

Water extractions of spore DNA remained consistent throughout the dilution series. At $10^5$, $10^6$, and $10^8$ spores, yields were 14.7%, 9.7%, and 13.3% (Fig. 1). In contrast, extractions
of vegetative cell DNA decreased with cell concentrations. At $10^8$, $10^6$, and $10^4$ vegetative cells, yields were 40%, 21.8%, and 13.2% (Fig. 1).

### 3.2. Unmodified extractions of spore and vegetative cell DNA from Mars analog soils

No detectable DNA was measured from any Mars analog soil extraction containing either spores or vegetative cells with the exception of the basalt and alkaline soils. Unmodified spore DNA extractions yielded 7.1% for basalt and 8% for alkaline while vegetative DNA extractions yielded 4.7% for basalt and 5.5% for alkaline (Fig. 2).

### 3.3. Modified extractions of spore and vegetative cell DNA from Mars analog soils

All modified Mars analog soil extractions of spore DNA achieved our 5% requirement by using the “near-universal” protocol. DNA yields were modestly increased for alkaline

---

**FIG. 1.** Baseline extractions of spore and vegetative cell DNA from water. Two dilution series containing $10^8$, $10^6$, and $10^4$ spores of *B. subtilis* and equivalent vegetative *E. coli* cells were processed using the standard extraction protocol. Water extractions of spore DNA remained consistent throughout the dilution series. In contrast, extractions of vegetative cell DNA decreased with cell concentrations (standard error shown, $n=3$).

**FIG. 2.** Unmodified extractions of spore and vegetative cell DNA from Mars analog soils. An estimated $1.6 \times 10^4$ spores (about 70 pg of DNA) of *B. subtilis* and equivalent vegetative cells of *E. coli* were deposited on 50 mg of Mars analog soil and processed following the standard extraction protocol. No detectable DNA was measured from any Mars analog soil extraction containing either spores or vegetative cells with the exception of the basalt and alkaline soils (standard error shown, $n=3$).
and basalt samples while aeolian, salt, acid, perchlorate, and JSC samples increased significantly from undetectable amounts. The spore results were basalt 9.2%, alkaline 10.7%, aeolian 6.1%, salt 6.7%, acid 5.1%, perchlorate 7.3%, JSC 7.8% (Fig. 3). In contrast, several vegetative cell DNA extractions, for example, JSC, aeolian, and perchlorate, failed to achieve 5% DNA yield. Similar to the spore extractions, yields for alkaline and basalt samples received a moderate increase while the remaining samples increased from undetectable amounts. The vegetative cell results were basalt 13.1%, alkaline 15.3%, aeolian 4.4%, salt 13.3%, acid 8.0%, perchlorate 4.5%, JSC 2.0% (Fig. 3).

3.4. Nanopore sequencing

From the resulting 48 h of sequencing of purified \textit{B. subtilis} spore DNA in the presence of Lambda, we detected a total of 8.7 gigabases or 1.3 million reads. Exactly 1,260,661 reads mapped to Lambda while 5 reads mapped to \textit{B. subtilis}. Given the sequencing library’s Lambda to \textit{B. subtilis} mass ratio, we expected approximately 2 \textit{B. subtilis} reads per 1,000,000 Lambda reads. However, we detected ~4 \textit{B. subtilis} reads per 1,000,000 Lambda reads. Assuming Lambda is the only known sequence, the CarrierSeq analysis ($Q=9$, $p=0.05$, default values) isolated 29 “target reads” (Table 1) that were identified using the NCBI blastn algorithm: 5 \textit{B. subtilis} reads, 13 \textit{Homo sapiens} reads, 1 \textit{Klebsiella pneumoniae} read, 1 human mastadenovirus read, 1 human DNA sequence (not able to be resolved to \textit{Homo sapiens}), 1 \textit{Streptococcus} sp. read, 1 \textit{Staphylococcus epidermis} read, and 6 unidentified HQNRs (Table 2). The basecalled and unanalyzed sequencing fastq files can be downloaded from FigShare at https://doi.org/10.6084/m9.figshare.5471824.v2, and the raw fast5 files can be downloaded from the NCBI SRA database at https://www.ncbi.nlm.nih.gov/biosample/9534236.

4. Discussion

4.1. Extraction of DNA from synthetic Mars analog soils

The extractions from spore DNA in water and soils at $10^4$ spores yielded similar results to our prior experiments at $10^8$ spores used to develop the initial modified protocol (Mojarro et al., 2017b). We previously hypothesized that low DNA yields from \textit{B. subtilis} were generally due to small acid-soluble proteins (SASPs) which bind to the DNA phosphate backbone and furnish protection from heat, salts, desiccation, and UV radiation in the spore state (Moeller et al., 2017b).

![FIG. 3. Modified extractions of spore and vegetative DNA from Mars analog soils. All DNA yields from $10^4$ vegetative cells and spores in 50 mg of Mars analog soils were increased with the “near-universal” extraction protocol. Our results indicate that a combination of desalting and compleitive binders is an adequate strategy for achieving our extraction goals (standard error shown, $n=3$).](image)

| Table 1. Low-Input Carrier Sequencing Metrics |
|-----------------------------------------------|
| **Number of reads** | **Min length** | **Max length** | **Median length** | **Total bases** |
| All reads (Lambda + \textit{B. subtilis} + Contamination + HQNRs) | 1,303,007 reads | 16 bases | 501,249 bases | 6,493 bases | 8,698,026,598 bases |
| Target reads (\textit{B. subtilis} + Contamination + HQNRs) | 29 reads | 267 bases | 1,559 bases | 595 bases | 19,981 bases |
| \textit{B. subtilis} reads | 5 reads | 848 bases | 1,559 bases | 967 bases | 5,270 bases |
| Contamination reads HQNRs | 6 reads | 267 bases | 865 bases | 453 bases | 2,933 bases |
Table 2. Target Reads

| Organism                          | Read length | Quality score | DUST score | Query cover | E value | Identity | Read ID                                                        |
|-----------------------------------|-------------|---------------|------------|-------------|---------|----------|---------------------------------------------------------------|
| Bacillus subtilis.1               | 848 bases   | 11.7          | 1.81       | 94 %        | 0       | 89 %     | @channel_69_7447cdae-d9cc-4ce8-a7f5-b0602b7f4f85_template     |
| Bacillus subtilis.2               | 967 bases   | 11.7          | 1.79       | 96 %        | 0       | 89 %     | @channel_200_5dbfb79-af02-47d8-9b80-4b1adaab20e88_template    |
| Bacillus subtilis.3               | 1559 bases  | 11.4          | 1.85       | 96 %        | 0       | 86 %     | @channel_217_4e9ba3a-4df3-4e85-9287-1067085687c5_template     |
| Bacillus subtilis.4               | 872 bases   | 10.0          | 1.70       | 94 %        | 0       | 83 %     | @channel_381_2ed7f3f3-ad20-4529-a63e-3a8639d1769_template     |
| Bacillus subtilis.5               | 1024 bases  | 11.2          | 1.77       | 96 %        | 0       | 86 %     | @channel_500_337c80ef-a4b9-403d-bb03-71e43339b85_template     |
| Streptococcus sp.                 | 1356 bases  | 11.1          | 2.08       | 94 %        | 0       | 84 %     | @channel_149_9b525218-4812-46a9-b8ca-ad8e06ccc3266c_template  |
| Staphylococcus epidermis          | 561 bases   | 9.39          | 2.09       | 92 %        | 8.0E-129| 83 %     | @channel_376_9f4bdc65-7083-44ca-8ee1-b202c2b2df2e2e_template   |
| Klebsiella pneumoniae             | 527 bases   | 11.6          | 1.59       | 91 %        | 0       | 91 %     | @channel_491_2e67e9b2-8a67-41da-b4ca-d3ba18428350_template    |
| Human mastadenovirus              | 595 bases   | 9.45          | 1.93       | 90 %        | 3.0E-148| 85 %     | @channel_434_cc744fa9-b62f-4274-86fd-7d957069ba5f_template    |
| Human DNA sequence                | 627 bases   | 9.47          | 1.95       | 57 %        | 7.0E-100| 85 %     | @channel_13_7738ebc8-3323-4ecf-4e98-9b1b81881c9424_template   |
| Homo sapiens.1                    | 681 bases   | 11.4          | 2.08       | 93 %        | 0       | 90 %     | @channel_182_5ed727df-dd19-48e1-9b70-f71f5c20de9_template     |
| Homo sapiens.2                    | 496 bases   | 10.2          | 3.12       | 91 %        | 1.0E-91 | 81 %     | @channel_227_f3bf082a-827e-4557-9db9-20662ddf3ce22e_template  |
| Homo sapiens.3                    | 416 bases   | 12.1          | 1.84       | 92 %        | 1.0E-127| 88 %     | @channel_86_e52fe73a-e4e7-45d5-9a72-71c473e18cb4e_template    |
| Homo sapiens.4                    | 989 bases   | 10.9          | 2.02       | 95 %        | 0       | 86 %     | @channel_98_ac4b4212-6f1d-49ed-9042-38fb46943ce_template      |
| Homo sapiens.5                    | 434 bases   | 11.9          | 1.89       | 89 %        | 2.0E-135| 90 %     | @channel_101_350b8de1-2288-4109-8b4a-a2210c0dd22e_template    |
| Homo sapiens.6                    | 638 bases   | 11.7          | 1.89       | 90 %        | 0       | 89 %     | @channel_139_23911960-d638-4db8-a2f2-f6f68ba0003c_template    |
| Homo sapiens.7                    | 569 bases   | 11.9          | 2.29       | 92 %        | 6.0E-180| 88 %     | @channel_368_2edf5ed6-1c62-42f0-ba32-0c5d29a9b2a2_template     |
| Homo sapiens.8                    | 933 bases   | 9.35          | 2.10       | 93 %        | 0       | 83 %     | @channel_357_97c139b2-fc1a-4a9f-9f68-4ce66247b2e2_template    |
| Homo sapiens.9                    | 554 bases   | 10.1          | 1.99       | 91 %        | 2.0E-150| 86 %     | @channel_447_472aee4-8a62-4209-b1da-0e06e35ee1e9f_template    |
| Homo sapiens.10                   | 661 bases   | 10.1          | 2.38       | 93 %        | 2.0E-164| 84 %     | @channel_290_d2806f7e-46b3-4af2-2404-02f18481630_template     |
| Homo sapiens.11                   | 654 bases   | 11.6          | 2.44       | 95 %        | 0       | 89 %     | @channel_300_a9532014-5520-48fd-9d3a-6f852e45e8f-template     |
| Homo sapiens.12                   | 445 bases   | 9.91          | 1.76       | 96 %        | 1.0E-132| 87 %     | @channel_318_b634ac5d-4a8d-401d-9ddf-13149858eaf_template     |
| Homo sapiens.13                   | 442 bases   | 9.65          | 2.00       | 91 %        | 9.0E-59 | 78 %     | @channel_334_1b36791e-2745-4e6b-adfd-7e5c19b494ce4f_template  |
| HQNR.1                            | 581 bases   | 12.7          | 2.44       | —           | —       | —        | @channel_343_1b36791e-2745-4e6b-adfd-7e5c19b494ce4f_template  |
| HQNR.2                            | 422 bases   | 9.18          | 5.99       | —           | —       | —        | @channel_143_a198d662-4306-42a5-ba53-caf57c7744e4f_template    |
| HQNR.3                            | 267 bases   | 9.22          | 2.21       | —           | —       | —        | @channel_54_103c06fa-3ac4-3f4c-a521-ccb7b461e5f4f_template     |
| HQNR.4                            | 865 bases   | 9.16          | 1.85       | —           | —       | —        | @channel_156_54e11b1-5323-a805-a7b8-530b644b040d40_de_template  |
| HQNR.5                            | 484 bases   | 11.7          | 1.94       | —           | —       | —        | @channel_166_f5014250-90d6-4977-9937-737e747a072d_template     |
| HQNR.6                            | 314 bases   | 9.42          | 1.98       | —           | —       | —        | @channel_241_19270801-598b-418b-a5fc-78a302197e3af-template   |
et al., 2009, 2012). In short, SASPs could inhibit interactions between the phosphate backbone and the Purelyse oxide ceramic microbeads, which would ideally attract negatively charged polymers (i.e., DNA). In the work of Mojarro et al. (2017b), we seemingly validated this hypothesis by increasing spore DNA yields in water from 15% to 43% using a proof-of-concept protein separation binding buffer/phenol cocktail. We therefore intuitively expected the DNA extraction yields from vegetative E. coli cells, which presumably do not contain SASPs, to be higher than B. subtilis spores. Extractions of spore DNA from our water control consistently yielded ~15% at all cell concentrations while extractions of vegetative DNA began with 40% at 10^8 cells and decreased to 13.3% at 10^6 (Fig. 1). These extractions share the exact parameters across cell lysis and DNA elution voltage, time, buffer volumes, and approximate cell concentrations with the exception of cells being either spores or vegetative. We are presently unable to offer a definitive explanation for the decrease in vegetative DNA with decreasing cell concentration and the apparent convergence of DNA yields at 10^4 cells (Fig. 1). Perhaps (1) cell concentration may affect DNA yield, (2) binding of nucleic acids to the Purelyse oxide ceramic microbeads is proportional to the concentration of lysed DNA, (3) vegetative DNA is more prone to destruction (e.g., hydrolysis, etc.) in solution without SASPs, or (4) vegetative DNA has a stronger affinity toward the Purelyse oxide ceramic microbeads and/or soil. Further work is ongoing in order to investigate this inconsistency between spore and vegetative DNA yields.

In regard to the Mars analog soil extractions, prior work from Mojarro et al. (2017b) had suggested that both alkaline and basalt soils would exert mild soil-DNA interactions due to adsorption effects by silicates (Melzak et al., 1996; Trevors, 1996; Zhou et al., 1996) while highly oxidized iron sulfate-containing Mars analog soils would undoubtedly destroy most DNA in an unmodified extraction (Fig. 2). Nevertheless, once disruptive metals/cations (e.g., Fe^{3+}, Ca^{2+}) were flushed and mineral adsorption sites were coated with competitive binders, DNA yields of at least 5% were achieved in all soils containing spores of B. subtilis (Fig. 3). Extractions from vegetative E. coli cells were only capable of achieving a 5% yield within the standard error with the exception of JSC, which yielded 2.0% (Fig. 3). These results indicate that perhaps DNA bound by SASPs has an increased resistance to destruction in the soil/binding buffer slurry immediately after cell lysis. If so, it would appear that DNA is more susceptible to damage from free radicals in Mars analog soils containing Fe^{3+}, Ca^{2+} species (JSC, aeolian, perchlorate) than from hydrolysis, depurination, and denaturation in acidic to alkaline soils (acid, salt, alkaline) (Fig. 3).

4.2. Ancient DNA on Mars

The relatively stable cryosphere on Mars since the Late Noachian period (Head and Marchant, 2014) makes it the ideal location for ancient DNA preservation. On Earth, the estimated half-life of DNA at -25°C is roughly on the order of 10^7 years (Millar and Lambert, 2013). However, local and global natural climate variability due to orbital cycles (Imbrie et al., 1992) and tectonics (Raymo and Ruddiman, 1992) limit the extent of cold environments favorable to-ward ancient DNA preservation on geologic timescales. Today, the average temperature at Gale Crater is ~48°C (Haberle et al., 2014) while similar conditions may have persisted since the Amazonian period (Head and Marchant, 2014), theoretically permitting the preservation of DNA beyond 10 million years. However, the likelihood of ancient DNA on the surface of Mars greatly decreases once we consider the effect of UV, cosmic ray exposure, and radioactive decay on nucleic acids (Kminek et al., 2003; Kminek and Bada, 2006; Hassler et al., 2014). Ancient DNA is prone to damage (e.g., hydrolysis, depurination) and fragmentation (Dabney et al., 2013; Millar and Lambert, 2013) once it is no longer actively repaired in a biologic system. In the unlikely case that nucleic acids are preserved and accessible

![FIG. 4. Cross-plots of target reads identified by Carrier-Seq displaying read length, quality score, and DUST score. There does not appear to be any separation of HQNRs and B. subtilis or contamination reads. We are presently only able to identify HQNRs through k-mer matching or via the NCBI blastn algorithm.](image-url)
at the martian surface, we then risk the possibility of destroying them once soils are hydrated during an extraction if they are not protected within cells (Gates, 2009; Mojarro et al., 2017b). The primary target for SETG is therefore extant or recently dead cells encapsulating nucleic acids that may survive the sample prep processes and yield long strands useful for taxonomic identification.

4.3. Low-input nanopore sequencing and contamination detection

A great advantage of nanopore sequencing is the ability to produce long reads capable of taxonomic identification (Greninger et al., 2015; Quick et al., 2015; Brown et al., 2017; Goordial et al., 2017). This has allowed us to directly map all basecalled reads to the reference genome and identify the five belonging to *B. subtilis* without further processing. However, in a true life-detection or unknown scenario, there must be a way to filter carrier (Lambda), low-quality, and low-complexity reads in order to identify those belonging to the unknown organisms. As mentioned earlier, we used CarrierSeq, which was developed in order to analyze carrier sequencing runs and identify the unknown or “target reads” in this context. This analysis identified 29 reads which did not map to Lambda, were not considered low quality (\(Q<9\), Albacore passing was \(Q>6\) at the time of analysis, currently \(Q>7\)) or low complexity (Morgulis et al., 2006), and did not originate from overly active “bad pores” that are known to produce spurious reads which we refer to as high-quality noise reads (HQNRs) in the work of Mojarro et al. (2018). These include the 5 *B. subtilis* reads detected by direct mapping, 6 HQNRs that originated from “good pores,” and most interestingly, 18 *Homo sapiens* and human microflora contamination reads (Table 2). Although we cannot recollect an exact point of contamination, this work was conducted in an open lab, with nonsterile nitrile gloves, and in an open bench without any nonstandard precautions (e.g., routine sporicidal/bleaching was conducted between sets of extractions).

The close abundance of contamination to *B. subtilis* reads possibly suggests comparable low levels of contamination (2 pg of *B. subtilis*), although there is no certain strategy to ascertain the true initial contamination quantities. We instead see this as an opportunity to highlight the effectiveness of nanopore sequencing to discriminate between true detection and contamination, particularly in the context of planetary protection where carrier sequencing could be employed as an effective method of detecting low levels of microbial contamination on the surface of spacecrafts (Rummel and Conley, 2018). The presence of HQNRs, however, complicates the certainty between a true unknown read and noise. Mojarro et al. (2018) proposed that HQNRs could possibly be the result of pore blockages from macromolecules (e.g., proteins) that may conceivably produce false signals. To date, we have not produced a fully satisfactory approach to identify and discard such reads. Identifying HQNRs through k-mer matching (Menzel et al., 2016) and via the NCBI blastn algorithm is presently the best approach, while plotting DUST (Morgulis et al., 2006), quality (Phred) score, and read length and statistical tests do not reveal any obvious relationships (Figs. 4 and 5). Recent work by Pontefract et al. (2018) on the failure modes of nanopore sequencing has suggested that improvements to basecalling algorithms, library preparation, and sequencing chemistries have suppressed the occurrence of these spurious reads. However, as Oxford Nanopore Technologies matures its sequencing technology, more robust testing under simulated low-input and unknown sample conditions is required in order to confidently identify HQNRs (if they are still an issue) where these artifactual reads may complicate the identification of a new organism or, specifically, extraterrestrial life.

![FIG. 5. Principal component analysis of the CarrierSeq target reads (n = 29). There does not appear to be any separability based on read length, quality score, and DUST score parameters.](image-url)
4.4. Detecting nonstandard bases

Raw nanopore sequencing data represent the translocation of unique k-mers, or nucleotide sequences of k bases, through a critical pore region which produces an associated ionic current signal. For simplicity, let us imagine a sliding window that is 3 bases wide (a 3-mer) moving one base at a time over a single-stranded DNA of sequence ATCG. As the window slides, we should observe AAT, ATC, and TCG each producing a unique signal. A moving 3-mer window over a DNA strand containing the four standard nucleobases (ATCG) could encounter up to 64 distinctive combinations. Therefore, a DNA strand containing a theoretical fifth nonstandard base would have 125 possible combinations, and so on. Raw data is then basecalled by applying algorithms such as Hidden Markov Models (Simpson et al., 2016) or Recurrent Neural Networks (Boža et al., 2017) that translate the collection of k-mers into consensus nucleic acid sequences. In previous work, we have demonstrated the capacity for these algorithms to detect nonstandard bases by sequencing poly(dexoxyinosinic-dexoxyctydelyd) acid, a synthetic DNA polymer composed of alternating deoxynosine (I) and deoxy-cytosine (C) bases (Carr et al., 2017). This proof of concept is of particular interest since the nonstandard inosine nucleotide contains hypoxanthine, a nucleobase that has been identified within meteorites (Martins et al., 2008). Meanwhile, work by others has also demonstrated the ability to detect base modification such as methylation (Simpson et al., 2016) using novel basecalling algorithms.

5. Conclusions

Life on Mars, if it exists or existed in the not-so-distant past, may potentially be detected via in situ nucleic acid extraction and nanopore sequencing. In this study, we have validated methodologies that enable the extraction of nucleic acids from low-biomass and recalcitrant analog soils ranging from a variety of surface environments (e.g., ancient hydrothermal spring deposits) to the subsurface of Mars. Our results indicate that a combination of desalting and competitive binding is a viable approach for achieving extraction yields equivalent to 1 ppb life detection (i.e., 2 pg of DNA from 10⁹ cells in 50 mg of soil). However, nucleic acid–based life detection on Mars may only be practical in the context of extant or recently dead cells as extracellular DNA preserved in soil may be destroyed during sample preparation. Furthermore, we have also demonstrated low-input nanopore sequencing from 2 pg of purified B. subtilis spore DNA simulating an ideal extractions yield from a Mars analog soil equivalent to 1 ppb life detection. This was accomplished by employing a genomic carrier (Lambda) to shuttle low-input amounts of B. subtilis DNA through library preparation and sequencing without the need for amplification. From the sequencing results, we were able to unambiguously discriminate between Lambda (carrier), B. subtilis (true detection), and contamination reads (human and human microflora). However, the detection of high-quality artifactual reads (HQNRs) not mapping to known life represents a false positive and complicates the identification of an unknown organism. Work is currently underway developing SETG through technology readiness level (TRL) 6, which integrates automated extraction, library preparation, and sequencing (Bhattaru, 2018). Validation of the automated system and pending challenges related to the concentration of volumes (e.g., for desalting, reducing elution volume, and library preparation) should soon enable a true sample-to-sequence demonstration. Nevertheless, here we have demonstrated the stepwise process and modifications required for the extraction and sequencing of nucleic acids from low-biomass Mars analog soils. We believe nucleic acids provide a sensitive and unambiguous indicator for life that facilitates distinguishing between forward contamination and putative nucleic acid–based martian life. Long-read nanopore sequencing could conceivably identify extraterrestrial sequences containing nonstandard bases and also discern conserved regions that may suggest a Mars-Earth shared ancestry (Isenbarger et al., 2008).

Acknowledgments

This work was supported by NASA MatISSE award NNX15AF85G and the MIT Shrock and OGE Diversity Graduate Fellowships.

Author Disclosure Statement

No competing financial interests exist.

References

Abramov, O. and Mojzsis, S.J. (2009) Microbial habitability of the Hadean Earth during the Late Heavy Bombardment. Nature 459:419–422.

Adcock, C.T., Haurust, E.M., and Forster, P.M. (2013) Readily available phosphate from minerals in early aqueous environments on Mars. Nat Geosci 6:824–827.

Allen, C.C., Morris, R.V., Jager, K.M., Golden, D.C., Lindstrom, D.J., Lindstrom, M.M., and Lockwood, J.P. (1998) Martian regolith simulant JSC Mars-1 [abstract 1690]. In 29th Annual Lunar and Planetary Science Conference, Lunar and Planetary Institute, Houston.

Barton, H.A., Taylor, N.M., Lubbers, B.R., and Pemberton, A.C. (2006) DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. J Microbiol Methods 66:21–31.

Bell, J.F., McSween, H.Y., Crisp, J.A., Morris, R.V., Murchie, S.L., Bridges, N.T., Johnson, J.R., Britt, D.T., Golombek, M.P., Moore, H.J., Ghosh, A., Bishop, J.L., Anderson, R.C., Brückner, J., Economou, T., Greenwood, J.P., Gunnlaugsson, H.P., Hargraves, R.M., Hviid, S., Knudsen, J.M., Madsen, M.B., Reid, R., Rieder, R., and Soderblom, L. (2000) Mineralogic and compositional properties of martian soil and dust: results from Mars Pathfinder. J Geophys Res: Planets 105:1721–1755.

Benítez-Páez, A., Fortune, K.J., and Sanz, Y. (2016) Species-level resolution of 16S rRNA gene amplicons sequenced through the MiniON™ portable nanopore sequencer. GigaScience 5, doi:10.1186/s13742-016-0111-z.

Benner, S.A. and Kim, H.-J. (2015) The case for a martian origin for Earth life. Proc SPIE 9606, doi:10.1117/12.2192890.

Bhattaru, S.A. (2018) Design, testing, and validation of the Search for Extra-Terrestrial Genomes instrument (Thesis). Massachusetts Institute of Technology, Cambridge, MA.

Boehnke, P. and Harrison, T.M. (2016) Illusory Late Heavy Bombardments. Proc Natl Acad Sci USA 113:10802–10806.
Boža, V., Brejová, B., and Vinař, T. (2017) DeepNano: deep recurrent neural networks for base calling in MinION nanopore reads. *PLoS One* 12, doi:10.1371/journal.pone.0178751.

Brack, A. and Pilling, C.T. (1998) Life on Mars: chemical arguments and clues from martian meteorites. *Extremophiles* 2:313–319.

Brown, B.L., Watson, M., Minot, S.S., Rivera, M.C., and Franklin, R.B. (2017) MinION™ nanopore sequencing of environmental metagenomes: a synthetic approach. *GigaScience* 6, doi:10.1093/gigascience/gix007.

Brunetto, G.S., Massoud, R., Leibovitch, E.C., Caruso, B., Johnson, K., Ohayon, J., Fenton, K., Cortese, I., and Jacobson, S. (2014) Digital droplet PCR (ddPCR) for the precise quantification of human T-lymphotropic virus 1 proviral loads in peripheral blood and cerebrospinal fluid of HAM/TSP patients and identification of viral mutations. *J Neurovirol* 20:341–351.

Carr, C.E. (2016) Detection of inosine via strand sequencing. Available online at https://doi.org/10.6084/m9.figshare.4276574.v1

Carr, C.E., Mojarro, A., Tani, J., Bhattachar, S.A., Zuber, M.T., Fuller, C.W., Finney, M., Church, G., and Ruvkun, G. (2016) Advancing the Search for Extra-Terrestrial Genomes. In 2016 IEEE Aerospace Conference, IEEE, Piscataway, NJ, doi: 10.1109/AERO.2016.7500859.

Carr, C.E., Mojarro, A., Hachey, J., Saboda, K., Tani, J., Bhattachar, S.A., Smith, A., Pontefract, A., Zuber, M.T., Doebler, R., Brown, M., Herrington, K., Talbot, R., Nguyen, V., Bailey, R., Ferguson, T., Finney, M., Church, G., and Ruvkun, G. (2017) Towards *in situ* sequencing for life detection. In 2017 IEEE Aerospace Conference, IEEE, Piscataway, NJ, doi:10.1109/AERO.2017.7943896.

Carr, C.E., Saboda, K., Mojarro, A., Hachey, J., Bhattachar, S.A., Ruvkun, G., and Zuber, M.T. (2019) Nucleic acid sequencing under Mars-like conditions [paper 2860]. In 2019 IEEE Aerospace Conference, IEEE, Piscataway, NJ, in press.

Castro-Wallace, S.L., Chiu, C.Y., John, K.K., Stahl, S.E., Rubins, K.H., McIntyre, A.B.R., Dworkin, J.P., Lupisella, M.L., Smith, D.J., Botkin, D.J., Stephenson, T.A., Juul, S., Turner, D.J., Izquierdo, F., Federman, S., Stryke, D., Somasekar, S., Alexander, N., Yu, G., Mason, C.E., and Burton, A.S. (2017) Nanopore DNA sequencing and genome assembly on the International Space Station. *Sci Rep* 7, doi:10.1038/s41598-017-18364-0.

Clark, B.C., Baird, A.K., Weldon, R.J., Tsuaki, D.M., Schnabel, L., and Candelaria, M.P. (1982) Chemical composition of martian fines. *J Geophys Res: Planets* 87:10059–10067.

Dabney, J., Meyer, M., and Paabo, S. (2013) Ancient DNA damage. *Cold Spring Harb Perspect Biol 3*:1000–1005.

Davila, A., Lacelle, D., Pollard, W., Marinova, M.M., Greer, C.W., DiRuggiero, J., McKay, C.P., and Whyte, L.G. (2016) Nearing the cold-arid limits of microbial life in permafrost of an upper dry valley, Antarctica. *ISME J* 10:1613–1624.

Goordial, J., Altshuler, I., Hindson, K., Chan-Yam, K., Marcolfes, E., and Whyte, L.G. (2017) *In situ* field sequencing and life detection in remote (79°26′N) Canadian High Arctic permafrost ice wedge microbial communities. *Front Microbiol* 8, doi:10.3389/fmicb.2017.02594.

Grinspoon, D., Bullock, M.A., Posner, A., Gomez-Elvira, J., Kohler, J., Martin, C., Reitz, G., Cinucotta, F.A., Kim, M.-H., Grinspoon, D., Bullock, M.A., Posner, A., Gomez-Elvira, J., Vasavada, A.R., and Zorzano-Mier, M.-F. (2014) Preliminary interpretation of the REMS pressure data from the first 100 sols of the MSL mission. *J Geophys Res: Planets* 119:440–453.

Harris, J.K. (2003) The genetic core of the universal ancestor. *Genome Res* 13:407–412.

Hassler, D.M., Zeitlin, C., Wimmer-Schweingruber, R.F., Ehresmann, B., Rafkin, S., Eigenbrode, J.L., Brinza, D.E., Weigle, G., Bottcher, S., Bohm, E., Burmeister, S., Guo, J., Kohler, J., Martin, C., Reitz, G., Cinucotta, F.A., Kim, M.-H., Henneberger, R.M., Walter, M.R., and Anitori, R.P. (2006) Extraction of DNA from acidic, hydrothermally modified volcanic soils. *Environmental Chemistry* 3:100–105.

Herrera, A. and Cockell, C.S. (2007) Exploring microbial diversity in volcanic environments: a review of methods in DNA extraction. *J Microbiol Methods* 70:1–12.
Horneck, G., Stöffler, D., Ott, S., Hornemann, U., Cockel, C.S., Moeller, R., Meyer, C., de Vera, J.-P., Fritz, J., Schade, S., and Artemieva, N.A. (2008) Microbial rock inhabitants survive hypervelocity impacts on Mars-like host planets: first phase of lithopanspermia experimentally tested. Astrobiology 8:17–44.

Hu, J., Wu, W.-C., and Sastry, S. (2004) Modeling subtilin production in Bacillus subtilis using stochastic hybrid systems. In Hybrid Systems: Computation and Control, Lecture Notes in Computer Science, presented at the International Workshop on Hybrid Systems: Computation and Control, Springer, Berlin, doi:10.1007/978-3-540-24743-2_28.

Hurt, R.A., Robeson, M.S., Shakhya, M., Moberly, J.G., Vish-nivetskaya, T.A., Gu, B., and Elias, D.A. (2014) Improved yield of high molecular weight DNA coincides with increased microbial diversity access from iron oxide cemented sub-surface clay environments. PLoS One 9, doi:10.1371/journal.pone.0102826.

Imbrie, J., Boyle, E.A., Clemens, S.C., Denny, A., Howard, W.R., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B., Morley, J.J., Peterson, L.C., Piasis, N.G., Prell, W.L., Raymo, M.E., Shackleton, N.J., and Toggweiler, J.R. (1992) On the structure and origin of major glaciation cycles 1. Linear responses to Milankovitch forcing. Paleoceanography 7, doi:10.1029/91PA02822.

Immler, J.A. and Linn, S. (1988) DNA damage and oxygen radical toxicity. Science 240:1302–1309.

Isernberger, T.A., Carr, C.E., Johnson, S.S., Finney, M., Church, G.M., Gilbert, W., Zubet, M.T., and Ruvkun, G. (2008) The most conserved genome segments for life detection on Earth and other planets. Orig Life Evol Biosph 38:517–533.

Jiang, S., Zhuang, J., Wang, C., Li, J., and Yang, W. (2012) Highly efficient adsorption of DNA on Fe 3+–iminodiacetic acid modified silica particles. Colloids Surf A Physicochem Eng Asp 409:143–148.

Johnson, A.P., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Lazcano, A., and Bada, J.L. (2008) The Miller volcanic spark discharge experiment. Science 322, doi:10.1126/science.1161527.

Jones, E.G., Lineweaver, C.H., and Clarke, J.D. (2011) An extensive phase space for the potential martian biosphere. Astrobiology 11:1017–1033.

Kerney, K.R. and Schuerger, A.C. (2011) Survival of Bacillus subtilis endospores on ultraviolet-irradiated rover wheels and Mars regolith under simulated martian conditions. Astrobiology 11:477–485.

Klingelhofer, G. (2004) Jarosite and hematite at Meridiani Planum from Opportunity’s Mossbauer spectrometer. Science 306:1740–1745.

Kminek, G. and Bada, J. (2006) The effect of ionizing radiation on the preservation of amino acids on Mars. Earth Planet Sci Lett 245, doi:10.1016/j.epsl.2006.03.008.

Kminek, G., Bada, J.L., Pogliano, K., and Ward, J.F. (2003) Radiation-dependent limit for the viability of bacterial spores in halite fluid inclusions and on Mars. Radiat Res 159:722–729.

Lever, M.A., Torti, A., Eichenbusch, P., Michaud, A.B., Šantl-Temkiv, T., and Jørgensen, B.B. (2015) A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. Front Microbiol 6:1–25.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997

Lu, H., Giordano, F., and Ning, Z. (2016) Oxford nanopore MinION sequencing and genome assembly. Genomics Proteomics Bioinformatics 14:265–279.

Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I., and Koonin, E.V. (1999) Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome Res 9:608–628.

Martins, Z., Botta, O., Fogel, M.L., Septon, M.A., Glavin, D.P., Watson, J.S., Dworkin, J.P., Schwartz, A.W., and Ehrenfreund, P. (2008) Extraterrestrial nucleobases in the Murchison meteorite. Earth Planet Sci Lett 270:130–136.

McIntyre, A.B.R., Rizzardi, L., Yu, A.M., Alexander, N., Rosen, G.L., Botkin, D.J., Stahl, S.E., John, K.K., Castro-Wallace, S.L., McGrath, K., Burton, A.S., Feinberg, A.P., and Mason, C.E. (2016) Nanopore sequencing in microgravity. npj Microgravity 2, doi:10.1038/npjmgavrl.2016.35.

McKay, C.P. (2010) An origin of life on Mars. Cold Spring Harb Perspect Biol 2, doi:10.1101/cshperspect.a003509.

McKay, D.S., Carter, J.L., Boles, W.W., Allen, C.C., and Allton, J.H. (1993) JSC-1: A new lunar regolith simulant. In 24th Lunar and Planetary Science Conference, Lunar and Planetary Institute, Houston, pp 963–964.

Melzak, K.A., Sherwood, C.S., Turner, R., and Haynes, C.A. (1996) Driving forces for DNA adsorption to silica in perchlorate solutions. J Colloid Interface Sci 181:635–644.

Menzel, P., Ng, K.L., and Krogh, A. (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7, doi:10.1038/ncomms11257.

Millar, C.D. and Lambert, D.M. (2013) Towards a million-year-old genome: ancient DNA. Nature 499:34–35.

Ming, D.W., Mittlefehldt, D.W., Morris, R.V., Golden, D.C., Gellert, R., Yen, A., Clark, B.C., Squires, S.W., Farrand, W.H., Ruff, S.W., Arvidson, R.E., Klingelhofer, G., McSween, H.Y., Rodionov, D.S., Schröder, C., de Souza, P.A., Jr., and Wang, A. (2006) Geochemical and mineralogical indicators for aqueous processes in the Columbia Hills of Gusev Crater, Mars. J Geophys Res 111, doi:10.1029/2005JE002560.

Ming, D.W., Gellert, R., Morris, R.V., Arvidson, R.E., Brückner, J., Clark, B.C., Cohen, B.A., d’Uston, C., Economou, T., Fleischer, I., Klingelhofer, G., McCoy, T.J., Mittlefehldt, D.W., Schmidt, M.E., Schröder, C., Squires, S.W., Tréguier, E., Yen, A.S., and Zipfel, J. (2008a) Geochemical properties of rocks and soils in Gusev Crater, Mars: results of the Alpha Particle X-Ray Spectrometer from Cumberland Ridge to Home Plate. J Geophys Res 113, doi:10.1029/2008JE003195.

Ming, D.W., Morris, R.V., Woida, R., Sutter, B., Lauer, H.V., Shinohara, C., Golden, D.C., Boynton, W.V., Arvidson, R.E., Stewart, R.L., Tamppari, L.K., Gross, M., and Smith, P. (2008b) Mars 2007 Phoenix Scout mission Organic Free Blank: method to distinguish Mars organics from terrestrial organics. J Geophys Res 113, doi:10.1029/2007JE003061.

Moeller, R., Selkow, P., Reitz, G., and Nicholson, W.L. (2009) Roles of small, acid-soluble spore proteins and core water content in survival of Bacillus subtilis spores exposed to environmental solar UV radiation. Appl Environ Microbiol 75:5202–5208.

Moeller, R., Schuerger, A.C., Reitz, G., and Nicholson, W.L. (2012) Protective role of spore structural components in determining Bacillus subtilis spore resistance to simulated Mars surface conditions. Appl Environ Microbiol 78:8849–8853.

Mojarro, A., Hache, J., Bailey, R., Brown, M., Doebler, R., Ruvkun, G., Zuber, M.T., and Carr, C.E. (2017a) Nucleic acid extraction and sequencing from low-biomass synthetic Mars analog soils [abstract 1585]. In 48th Lunar and Planetary Science Conference, Lunar and Planetary Institute, Houston.
Mojarrado, A., Ruvkun, G., Zuber, M.T., and Carr, C.E. (2017b) Nuclease acid extraction from synthetic Mars analog soils for in situ life detection. Astrobiology 17:747–760.

Mojarrado, A., Hackey, J., Ruvkun, G., Zuber, M.T., and Carr, C.E. (2018) CarrierSeq: a sequence analysis workflow for low-input nanopore sequencing. BMC Bioinformatics 19, doi: 10.1186/s12859-018-2124-3.

Morgulis, A., Gertz, E.M., Schäffer, A.A., and Agarwala, R. (2006) A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J Comput Biol 13:1028–1040.

Morris, R.V., Klingelhöfer, G., Schröder, C., Rodionov, D.S., Yen, A., Ming, D.W., de Souza, P.A., Wodwiak, T., Fleischer, I., Gelligt, R., Bernhardt, B., Bonnes, U., Cohen, B.A., Eivanov, E.N., Foh, J., Gültich, P., Kankeleit, E., McCoy, T., Mittlefeldt, D.W., Renz, F., Schmidt, M.E., Zubkov, B., Squyres, S.W., and Arvidson, R.E. (2006) Mössbauer mineralogy of rock, soil, and dust at Meridiani Planum, Mars: Opportunity’s journey across sulfate-rich outcrop, basaltic sand and dust, and hematite lag deposits. J Geophys Res: Planets 111, doi:10.1029/2006JE002791.

Morris, R.V., Klingelhöfer, G., Schröder, C., Fleischer, I., Ming, D.W., Yen, A.S., Gelligt, R., Arvidson, R.E., Rodionov, D.S., Crumpler, L.S., Clark, B.C., Cohen, B.A., McCoy, T.J., Mittlefeldt, D.W., Schmidt, M.E., de Souza, P.A., and Squyres, S.W. (2008) Iron mineralogy and aqueous alteration from Husband Hill through Home Plate at Gusev Crater, Mars: results from the Mössbauer instrument on the Spirit Mars Exploration Rover. J Geophys Res: Planets 113, doi:10.1029/2008JE003201.

Morris, R.V., Ruff, S.W., Gelligt, R., Ming, D.W., Arvidson, R.E., Clark, B.C., Golden, D.C., Siebach, K., Klingelhofer, G., Schröder, C., Fleischer, I., Yen, A.S., and Squyres, S.W. (2010) Identification of carbonate-rich outcrops on Mars by the Spirit rover. Science 329:421–424.

Navarro-Gonzalez, R., Rainey, F.A., Molina, P., Bagaley, D.R., Hollen, B.J., de la Rosa, J., Small, A.M., Quinn, R.C., Ranjan, S., Wordsworth, R., and Sasselov, D.D. (2017) Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. Nat Chem 9, doi:10.1038/nchem.2445.

Pezo, V., Schepers, G., Lambertucci, C., Marlière, P., and Herdewijn, P. (2014) Probing ambiguous base-pairs by genetic transformation with XNA templates. ChemBioChem 15:2255–2258.

Pontefract, A., Hackey, J., Zuber, M.T., Ruvkun, G., and Carr, C.E. (2018) Sequencing nothing: exploring failure modes of nanopore sensing and implications for life detection. Life Sci Res 18:80–86.

Poretzky, R., Rodriguez-R, L.M., Luo, C., Tsementzi, D., and Konstantinidis, K.T. (2014) Strengths and limitations of 16S rRNA gene amplification in revealing temporal microbial community dynamics. PLoS One 9, doi:10.1371/journal.pone.0093827.

Powner, M.W., Gerland, B., and Sutherland, J.D. (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242.

Powner, M.W., Sutherland, J.D., and Szostak, J.W. (2010) Chemoselective multicomponent one-pot assembly of purine precursors in water. J Am Chem Soc 132:16677–16688.

Quick, J., Ashton, P., Calus, S., Chatt, C., Gossain, S., Hawker, J., Nair, S., Neal, K., Nye, K., Peters, T., De Pinna, E., Robinson, E., Struthers, K., Webber, M., Catto, A., Dallman, T.J., Hawkey, P., and Loman, N.J. (2015) Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biol 16, doi:10.1186/s13059-015-0677-2.

Raley, C., Munroe, D., Jones, K., Tsai, Y.-C., Guo, Y., Tran, B., Gowda, S., Troyer, J.L., Soppet, D.R., Stewart, C., Stephens, R., Chen, J., Skelly, T.F., Heiner, C., Korlach, J., and Nisleys, D. (2014) Preparation of next-generation DNA sequencing libraries from ultra-low amounts of input DNA: application to single-molecule, real-time (SMRT) sequencing on the Pacific Biosciences RS II. bioRxiv 003566, doi:10.1101/003566.

Ranjan, S. and Sasselov, D.D. (2017) Constraints on the early terrestrial surface UV environment relevant to prebiotic chemistry. Astrobiology 17:169–204.

Ranjan, S., Wordsworth, R., and Sasselov, D.D. (2017) Atmospheric constraints on the surface UV environment of Mars at 3.9 Ga relevant to prebiotic chemistry. Astrobiology 17:687–708.

Raymo, M.E. and Ruddiman, W.F. (1992) Tectonic forcing of late Cenozoic climate. Nature 359:117–122.

Ritson, D. and Sutherland, J.D. (2012) Prebiotic synthesis of simple sugars by photoedox systems chemistry. Nat Chem 4: 895–899.

Ruff, S.W. and Farmer, J.D. (2016) Silica deposits on Mars with features resembling hot spring biosignatures at El Tatio in Chile. Nat Commun 7, doi:10.1038/ncomms13554.

Rummel, J.D. and Conley, C.A. (2018) Inadvertently finding Earth contamination on Mars should not be a priority for anyone. Astrobiology 18:108–115.

Sabina, J. and Leamon, J.H. (2015) Bias in whole genome amplification: causes and considerations. In Whole Genome Amplification, Methods in Molecular Biology, Humana Press, New York, pp 15–41.

Schuerger, A.C., Mancinelli, R.L., Kern, R.G., Rothschild, L.J., and McKay, C.P. (2003) Survival of endospores of Bacillus subtilis on spacecraft surfaces under simulated martian environments. Icarus 165:253–276.

Schuerger, A.C., Golden, D.C., and Ming, D.W. (2012) Biotoxicity of Mars soils 1. Dry deposition of analog soils on microbial colonies and survival under martian conditions. Planet Sci 72:91–101.
Schuerger, A.C., Ming, D.W., and Golden, D.C. (2017) Bio-
toxicity of Mars soils: 2. Survival of *Bacillus subtilis* and
*Enterococcus faecalis* in aqueous extracts derived from six
Mars analog soils. *Icarus* 290:215–223.
Shuster, D.L. (2005) Martian surface paleotemperatures from
thermochronology of meteorites. *Science* 309:594–600.
Simpson, J.T., Workman, R., Zuzarte, P.C., David, M., Dursi,
L.J., and Timp, W. (2016) Detecting DNA methylation using
the Oxford Nanopore Technologies MinION sequencer.
*bioRxiv* doi:10.1101/047142
Skarstad, K., Boye, E., and Steen, H.B. (1986) Timing of ini-
tiation of chromosome replication in individual *Escherichia
coli* cells. *EMBO J* 5:1711–1717.
Stairs, S., Nikmal, A., Bučar, D.-K., Zheng, S.-L., Szostak,
J.W., and Powner, M.W. (2017) Divergent prebiotic synthesis
of pyrimidine and 8-oxo-purine ribonucleotides. *Nat Commun*
8, doi:10.1038/ncomms15270.
Stoker, C.R., Zent, A., Catling, D.C., Douglas, S., Marshall,
J.R., Archer, D., Clark, B., Kounaves, S.P., Lemmon, M.T.,
Quinn, R., Renno, N., Smith, P.H., and Young, S.M.M.
(2010) Habitability of the Phoenix landing site. *J Geophys
Res* 115, doi:10.1029/2009JE003421.
Takada-Hoshino, Y. and Matsumoto, N. (2004) An improved
dNA extraction method using skim milk from soils that
strongly adsorb DNA. *Microbes Environ* 19:13–19.
Taylor, G.J. (2013) The bulk composition of Mars. *Chemie der
Erde Geochemistry* 73:401–420.
Trevors, J.T. (1996) DNA in soil: adsorption, genetic transfor-
mation, molecular evolution and genetic microchip. *Antonie
Van Leeuwenhoek* 70:1–10.
Wang, X. and Wood, T.K. (2011) ISS5 inserts upstream of the
master motility operon flhDC in a quasi-Lamarckian way.
*ISME J* 5, doi:10.1038/isemj.2011.27.
Wänke, H., Brückner, J., Dreibus, G., Rieder, R., and Ryab-
chikov, I. (2001) Chemical composition of rocks and soils at
the Pathfinder site. In *Chronology and Evolution of Mars,
edited by R. Kallenbach, J. Geiss, and W.K. Hartmann,
Springer, Dordrecht, the Netherlands, pp 317–330.
Weiss, B.P. (2000) A low temperature transfer of ALH84001
from Mars to Earth. *Science* 290:791–795.
Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L.,
Pechman, K., Lewis, B.J., and Stahl, D. (1975) Conservation
of primary structure in 16S ribosomal RNA. *Nature* 254:83–86.
Yang, R., Paparini, A., Monis, P., and Ryan, U. (2014) Com-
parison of next-generation droplet digital PCR (ddPCR) with
quantitative PCR (qPCR) for enumeration of Cryptospor-
idium oocysts in faecal samples. *Int J Parasitol* 44:1105–
1113.
Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery
from soils of diverse composition. *Appl Environ Microbiol*
62:316–322.

Address correspondence to:
Angel Mojarro
*Massachusetts Institute of Technology*
77 Massachusetts Ave, Room E25-647
*Cambridge, MA 02139*

E-mail: mojarro@mit.edu

Submitted 2 July 2018
Accepted 17 April 2019
Associate Editor: Christopher McKay

| Abbreviations Used |
|--------------------|
| ddPCR = droplet digital polymerase chain reaction |
| HQNRs = high-quality noise reads |
| SASPs = small acid-soluble proteins |
| SETG = Search for Extra-Terrestrial Genomes |