Radiation Resistance of Sequencing Chips for in situ Life Detection

Citation: Carr, Christopher E. et al. "Radiation Resistance of Sequencing Chips for in Situ Life Detection." Astrobiology 13.6 (2013): 560–569. © 2013 Mary Ann Liebert, Inc.

As Published: http://dx.doi.org/10.1089/ast.2012.0923

Publisher: Mary Ann Liebert

Persistent URL: http://hdl.handle.net/1721.1/79754

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

ASTROBIOLOGY
Volume 13, Number 6, 2013

DOI: 10.1089/ast.2012.0923

Radiation Resistance of Sequencing Chips for *in situ* Life Detection

Christopher E. Carr,[1,2,3] Holli Rowedder,[2,3] Clarissa S. Lui,[1,2,3] Ilya Zlatkovsky,[4] Chris W. Papalias,[4] Jarie Bolander,[4] Jason W. Myers,[4] James Bustillo,[4] Jonathan M. Rothberg,[4] Maria T. Zuber,[1] and Gary Ruvkun[2,3]

Abstract

Life beyond Earth may be based on RNA or DNA if such life is related to life on Earth through shared ancestry due to meteoritic exchange, such as may be the case for Mars, or if delivery of similar building blocks to habitable environments has biased the evolution of life toward utilizing nucleic acids. In this case, *in situ* sequencing is a powerful approach to identify and characterize such life without the limitations or expense of returning samples to Earth, and can monitor forward contamination. A new semiconductor sequencing technology based on sensing hydrogen ions released during nucleotide incorporation can enable massively parallel sequencing in a small, robust, optics-free CMOS chip format. We demonstrate that these sequencing chips survive several analogues of space radiation at doses consistent with a 2-year Mars mission, including protons with solar particle event–distributed energy levels and 1 GeV oxygen and iron ions. We find no measurable impact of irradiation at 1 and 5 Gy doses on sequencing quality nor on low-level hardware characteristics. Further testing is required to study the impacts of soft errors as well as to characterize performance under neutron and gamma irradiation and at higher doses, which would be expected during operation in environments with significant trapped energetic particles such as during a mission to Europa. Our results support future efforts to use *in situ* sequencing to test theories of panspermia and/or whether life has a common chemical basis. Key Words: Laboratory simulation experiments—Life-detection instruments—Nucleic acids—Mars—Panspermia. Astrobiology 13, 560–569.

1. Introduction

Life on Mars, if it exists, may be related to life on Earth due to significant meteoritic transfer between the planets (Gladman and Burns, 1996; Gladman *et al.,* 1996) at low temperature (Weiss *et al.,* 2000; Fritz *et al.,* 2005; Shuster and Weiss, 2005) and survivable shock pressures (Horneck *et al.,* 2008). Under a scenario of shared ancestry, modern molecular biology tools can be applied to finding and characterizing any extant or recently dead martian life, in particular by isolating, detecting, and sequencing nucleic acids such as DNA or RNA (Ruvkun *et al.,* 2002; Isenbarger *et al.,* 2008a, 2008b). One strategy would be to analyze samples returned from Mars. However, under current mission scenarios, samples collected on Mars would experience conditions that could degrade nucleic acids, such as irradiation during extended storage on the surface of Mars or temperature fluctuations during the return cruise. *In situ* sequencing is a promising alternative. Traditional Sanger sequencing has been implemented in a microfluidic format (Blazej *et al.,* 2006), but this approach has limited throughput: it is slow and not easily parallelizable. Thus, the number of sequences generated on Mars would be very limited. Second-generation DNA sequencing technologies enable massively parallel sequencing (Metzker, 2010) but have been too large and complex to consider *in situ* sequencing. However, massively parallel sequencing *in situ* is now feasible through semiconductor sequencing: a small standard CMOS chip that enables concurrent sequencing in millions of wells, requires no imaging or optics, and is extremely small, fast, robust, and relies upon simple chemistry (Rothberg *et al.,* 2011).

In the present study, we characterized the ability of semiconductor sequencing chips to withstand simulated space radiation conditions associated with a 2-year mission

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts.
[2]Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts.
[3]Department of Genetics, Harvard Medical School, Boston, Massachusetts.
[4]Life Technologies, South San Francisco, California.

to Mars. Space instruments must survive hazards of the space environment such as extremes of temperature and pressure; space radiation is generally the main concern for semiconductor devices. These sequencing chips rely on ion-sensitive field-effect transistors (ISFETs) to detect the release of hydrogen ions during addition of nucleotides to a growing DNA strand. Space radiation, such as protons, heavy ions, and neutrons (produced through secondary reactions) may impact semiconductor sequencing in two fundamental ways: First, it may cause lattice displacement, leading to recombination centers and depletion of minority carriers (holes for p-type silicon, electrons for n-type silicon). Second, it may cause ionization effects resulting in random transients. We focus on the former, which can lead to charge buildup within the silicon dioxide gate, leading to easier activation of n-type ISFETs such as those used in the sequencing chips. When the accumulated charge is high enough, it can keep transistors permanently open (n-type) or closed (p-type). The potential impact of irradiation on the sequencing chips could be an increased risk of false-positive nucleotide incorporation signals, reduced performance such as lower read length, or complete loss of sensors.

2. Materials and Methods

2.1. Approach

Semiconductor sequencing works as follows (Fig. 1A): First a sequencing library, consisting of fragments of DNA with known ends, is generated. This can be done by using polymerase chain reaction (PCR) to target specific genomic loci, or through metagenomic approaches such as fragmentation of DNA followed by ligation of known sequences to the fragments. The library is amplified via clonal PCR on beads in an emulsion. The amplified library is then enriched to eliminate beads without clonal products. The second strand is removed to produce single-stranded DNA, and a sequencing primer and polymerase are added. Then a sequencing chip containing millions of microwells is loaded with beads, which are sized to achieve one bead per well. Sequencing takes 1–3 h depending upon the desired read length and works as follows: Beneath each well is an ISFET. Consider one well on a chip, occupied by a bead covered by a clonal DNA molecule: When a matching deoxynucleotide triphosphate (dNTP) flows by, a polymerase enzyme sitting on the DNA incorporates the dNTP into the 3′ end of a growing double-stranded DNA molecule, releasing a hydrogen ion. When this happens concurrently on the $\sim 10^6$ identical molecules in the well, the resulting transient change in pH is detected as a change in the source voltage. By flowing the four standard dNTPs one by one, and looking for transients, the target sequence can be determined in each of the occupied wells. By fitting these transients to a model of nucleotide incorporation, each base can be called and scored for quality. Reads can be mapped to a known genome or some collection of relevant sequence data for comparison. Finally, control beads, templated with known sequences, are added at the 1–2% level to provide a built-in control with every run.

We used Ion 314 chips with 1.3 million wells (or equivalently, pixels) per chip. To characterize sequencing chips, we electrically tested them and confirmed their ability to perform DNA sequencing after irradiation. Electrical testing included quantifying gain, reference voltage, and sensitivity to pH for each ISFET sensor (pixel), and requires chip wetting for the ISFETs to function properly. The chips were designed for single use; thus, to ensure that we could also test the ability of the chips to sequence DNA after irradiation, we split 40 total chips into two groups of 20 chips (Fig. 1B). One group was electrically tested, dried, irradiated, and tested again, while the second group was irradiated and

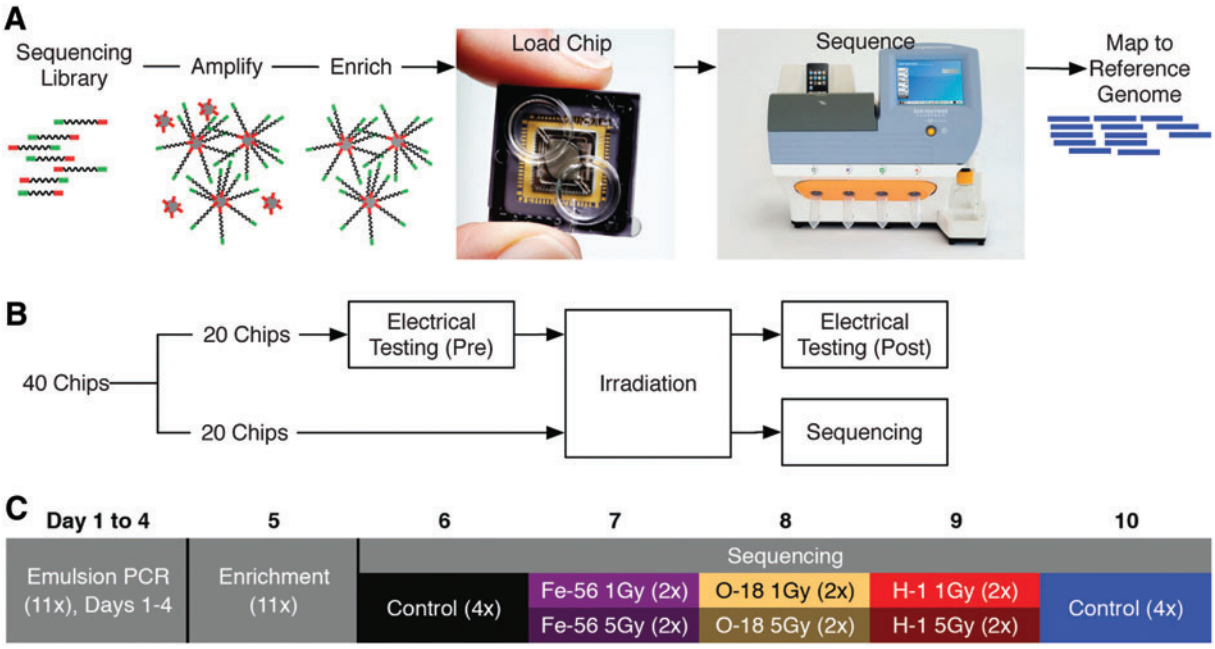


FIG. 1. Experiment overview. (**A**) Semiconductor sequencing. (**B**) Partitioning of sequencing chips into electrical testing and sequencing cohorts. (**C**) Sequencing was carried out over 10 days. See text for details. Color images available online at www.liebertonline.com/ast

used for DNA sequencing. The sequencing group of 20 included two chips for each of six irradiation conditions (3 ions × 2 doses, see below), four control chips, and four additional control chips from a separate manufacturing lot. Chips were similarly allocated for electrical testing; in this group, all chips came from the identical manufacturing lot and wafer (Supplementary Table S1; Supplementary Data are available online at www.liebertonline.com/ast).

2.2. Irradiations

We carried out these sequencing chip exposures (Table 1) concurrently with exposures of reagents required for DNA amplification via PCR (Carr *et al.,* 2013), utilizing the same ions and water-equivalent doses of 1 and 5 Gy. We exposed the chips to proton (H-1), oxygen (O-18), and iron (Fe-56) ions at the NASA Space Radiation Lab, using the 200 MeV Linear Accelerator (H-1) and Electron Beam Ion Source (O-18 and Fe-56). The H-1 exposure simulated the August 1972 solar particle event (SPE) energy distribution (Supplementary Table S2), whereas the O-18 and Fe-56 exposures utilized 1 GeV/nucleon particles. Sequencing chips were exposed in electrostatic packaging in groups of four, with typical beam uniformity of $<2\%$ variation over a 13×13 cm area (Supplementary Fig. S1). We used water-equivalent calibrated dose to estimate fluence as previously described (Carr *et al.,* 2013), then used this fluence to estimate silicon-equivalent dose. Linear energy transfer (LET) values used to calculate silicon-equivalent dose (Supplementary Table S1) were determined with LET124 (http://tvdg10.phy.bnl.gov/let.html). For detailed beam characterization data including Bragg peaks see the NSRL website at http://www.bnl.gov/medical/nasa.

At the higher dose, the H-1 fluence corresponds to the expected total SPE fluence of protons with energies above 30 MeV for a 2-year period under a situation of zero shielding (Carr *et al.,* 2013). This proton exposure is higher than expected under martian surface conditions, because of shielding due to the martian surface and some shielding provided by the atmosphere. Each heavy ion dose corresponds to a fluence above the total expected fluence (2.5×10^6 cm^{-2}) of all heavy ion particles, estimated as the galactic cosmic radiation heavy ion contribution of 0.04 cm^{-2} s^{-1} × 2 years. Neutrons are expected to contribute significantly to total dose on the surface of Mars (Le Postollec *et al.,* 2009), with a worst-case fluence over 2 years of $\sim 10^{10}$ cm^{-2}. However, in the present study we focused on other particles because CMOS devices tend to be relatively insensitive to neutron exposure up to above 10^{14} cm^{-2} (Walker *et al.,* 1990). Gamma rays, also associated with secondary reactions, contribute significantly to dose on the surface of Mars (Le Postollec *et al.,* 2009) and will be explored in future work. While gamma irradiation at sterilization doses (>25,000 Gy) has been shown to impair pH ISFETs (Shimada *et al.,* 1980), these doses are far beyond what would be expected for a 2-year Mars mission.

2.3. Electrical testing

Electrical testing of the semiconductor (CMOS) chip was done with the Ion Personal Genome Machine (PGM) System. Measurements of gain, reference voltage, and pixels in range were measured during the standard chip calibration step. This calibration step verifies that the chip is operating correctly and reports if the sensor is good enough to be used to sequence. A special PGM System test mode allows independent measurement of these parameters without requiring a sequencing run. In addition to the standard electrical tests, a separate manual pH sensitivity test was also run on a subsample of parts to ensure there had been no shifts in the pH sensitivity of the device after radiation exposure. Control parts were tested and then held in storage for the same amount of time as the irradiated parts. This was to determine any aging effect due to the surface being wet and to assess any associated drift.

The gain test measures the voltage gain from the input of the ion sensing transistor (ISFET) to the output of the chip, which is then converted to a digital code with an on-PGM analog-to-digital converter. The nominal gain of the chip is ~ 0.6. Any gain near 0.6 meets sequencing requirements. Any damage to the ISFET or other internal circuitry would cause a shift in gain due to offsets, charging, or destruction of the transistors from the applied radiation.

The reference voltage measures the voltage applied to the reference electrode in the W2 solution (80 mL of 100 m*M* KMT in 1.92 L milli-Q H_2O), which biases the ISFET to the proper operating voltage. A DAC code is recorded during calibration, which ideally would be set to 1.9 V but can vary depending on the calibration routines optimizing for the maximum number of pixels in range. Any damage to the ISFET such as charging from the radiation would cause a shift in this reference voltage.

The "pixels in range" test monitors how many of these pixels are active to provide sequencing data. Bad pixels are defined as pinned "high" or pinned "low," corresponding to

Table 1. Summary of Exposures and Achieved Fluence and Dose

Source	*Particle*	*Energy (MeV/n)*	*Flux ($cm^{-2}s^{-1}$)*	*Exposure time (min)*	*Fluence (cm^{-2})*	*Water-equivalent dose (Gy)*	*Silicon-equivalent dose (Gy)*
Accelerator	H-1	SPE	N.A.	4.39	5.4×10^8	1.0	4.4
Accelerator	H-1	SPE	N.A.	12.03	2.7×10^9	5.1	21.8
Accelerator	O-16	1000	3.4×10^5	2.16	4.4×10^7	1.0	4.5
Accelerator	O-16	1000	2.0×10^5	18.36	2.2×10^8	5.0	22.3
Accelerator	Fe-56	1000 [967]	4.4×10^4	1.55	4.1×10^6	1.0	4.5
Accelerator	Fe-56	1000 [967]	4.9×10^4	6.95	2.1×10^7	5.0	22.3

Energy is target and [achieved] in MeV/nucleon; SPE indicates solar particle event energy distribution (Supplementary Table S2). N.A. = not applicable. We controlled the water-equivalent dose, and the fluence and flux are calculated from LET values (Supplementary Table S3).

their bias point being either above or below the optimal electrical operating conditions. The software reports pixels in range rounded to the nearest thousand pixels.

Statistically, there could be a small percentage of transistors that are defective in a pixel array. Damage to the ISFETs from radiation would result in a significant shift in pinned pixels pre- and post-exposure. Also, parts that have been wet, baked (to remove moisture), and sitting for some time may produce a slight increase of pinned pixels. Note that sequencing chips are intended for single use only, so exposure of used chips to external environmental conditions is not a source of concern. A spatial plot of pixels in range is also displayed on the PGM instrument screen, which provides a visual aid to discern whether there is any pattern or localized area of pinned pixels.

Differences in gain, reference voltage, and pixels in range were assessed by a one-way analysis of variance (ANOVA) with the MATLAB (Mathworks, Natick, MA) *anova1* function. When there were only two categories to compare, we used either a two-tailed (equivalent to ANOVA) or one-tailed *t* test (*ttest2* function).

The pH sensitivity test measures the output voltage in response to changes in pH at the ISFET (mV/pH). This test requires a special setup on the PGM instrument: Tris buffer is placed in the reagent tubes instead of the nucleotides. Each tube has a buffer with a different pH (8, 7.6, 7.2, 7.4). The output voltage is then measured in response to each of these buffers. A linear extrapolation of the pH sensitivity is then obtained from post-processing of the data. Histograms and heat maps of the pH step are also generated. Tris buffer is sensitive to temperature, and the pH sensitivity test can vary from measurement to measurement. This test is performed to assure that the pH sensitivity of the sensor has not changed after being irradiated.

2.4. Sequencing and sequence analysis

For sequencing (Fig. 1C), we used an *Escherichia coli* DH10B control library (Ion Sphere Quality Control Kit 4468656, Life Technologies, Carlsbad, CA), amplified it at 1×concentration with the Ion OneTouch System, loaded beads onto Ion 314 chips (with 1.26 million wells), and sequenced on the Ion Personal Genome Machine (PGM) Sequencer. To control for run-to-run variability in bead preparation, we generated 11 sets of amplified beads (days 1–4), enriched them (day 5), and pooled them. Controls were sequenced on the first day of sequencing (day 6) and on the last day (day 10) to control for any aging phenotypes (such as loss of loadable amplified beads through aggregation). During each sequencing day, we carried out four sequencing runs following this pattern: initialization, sequencing (2×), water clean (2×), initialization, sequencing (2×), water clean (2×). In addition, chlorite cleans were carried out before the first day of sequencing (Day 5) and preceding the final water cleans on days 7, 9, and 10 to limit any growth within the sequencer fluidics. Manufacturing lot controls were sequenced in the third and fourth runs of each control sequencing day.

The automated chip data analysis first determined which wells contain beads; next, alive beads (those producing measurable signals) were identified. Recognition of an initial key sequence separated control beads from library beads. Flows were then analyzed to produce sequences for each well, eliminating those with poor signal or signatures of polyclonality and trimming 3′ ends to remove adaptor sequences or low-quality bases.

Reads were aligned to the *E. coli* DH10B genome with the Torrent Mapping Alignment Program (TMAP; https://github.com/iontorrent/TMAP). Briefly, TMAP determines candidate mapping locations with the BWA-short (Li and Durbin, 2009), BWA-long (Li and Durbin, 2010), and SSAHA (Ning *et al.*, 2001) algorithms; scores these candidates with the Smith-Waterman (Smith and Waterman, 1981) algorithm; and finds an aggregate best alignment. Each sequencing run results in a single BAM file containing the alignments. These files were uploaded to Galaxy (Giardine *et al.*, 2005; Blankenberg *et al.*, 2010; Goecks *et al.*, 2010) and analyzed with the SAM/BAM Alignment Summary Metrics feature to identify alignment statistics including mismatch rate and insertion/deletion (indel) rate. These metrics were combined with ion and dose information and analyzed with MATLAB via regression (using *regress* and *regstats* functions) and ANOVA (using *anovan* function).

3. Results

3.1. Electrical testing

All chips, including controls, passed all pre- and post-irradiation electrical tests (Supplementary Tables S4–S6).

All gain values remained within the normal range of variability. We first examined gain differences within the control (C) and irradiation (I) subgroups in the pre- and post-irradiation conditions (Fig. 2A) and found a nearly significant difference in the means by ANOVA ($p=0.0545$). However, there was no association between irradiation and changes in the gain (Fig. 2B; ANOVA or two-sided *t* test, $p=0.91$), whereas there was a significant difference between electrical tests (Fig. 2C; one-sided *t* test, $p=0.0035$), consistent with an effect of the wetting and drying process associated with electrical testing.

Similarly, all variations in the reference electrode voltage across all chips stayed within the normal range, never exceeding 1.3% (Fig. 2D), even while there were differences in group means by ANOVA ($p=0.0035$). Once again, there was no association between irradiation and change in the reference voltage (Fig. 2E; ANOVA or two-sided *t* test, $p=0.78$), whereas there was a significant difference between electrical tests (Fig. 2F; one-sided *t* test, $p=0.00032$).

Pixels in range analysis revealed significant but inconsequential differences by ANOVA (Fig. 2G; $p=7.5\times10^{-7}$). The control parts had a maximum shift of pixels in the range of 0.5%, where the irradiated parts had a maximum shift of pixels in the range of 0.2%. These small shifts can be caused by bubbles in the wash solution but are well within normal statistical variation. Comparing pre- and post-irradiation conditions, we found that the irradiated parts had more pixels in range than the controls (Fig. 2H; one-sided *t* test, $p=0.00016$). This difference was caused primarily by a reduction in pixels in range in the controls, although there was also a significant reduction in pixels in range in the irradiated group (post vs. pre, one-sided *t* test, $p=0.0018$). Overall, effects of the wetting and drying during electrical testing and/or bubbles in the wash buffer contributed to significant but inconsequentially lower numbers of pixels in range in the

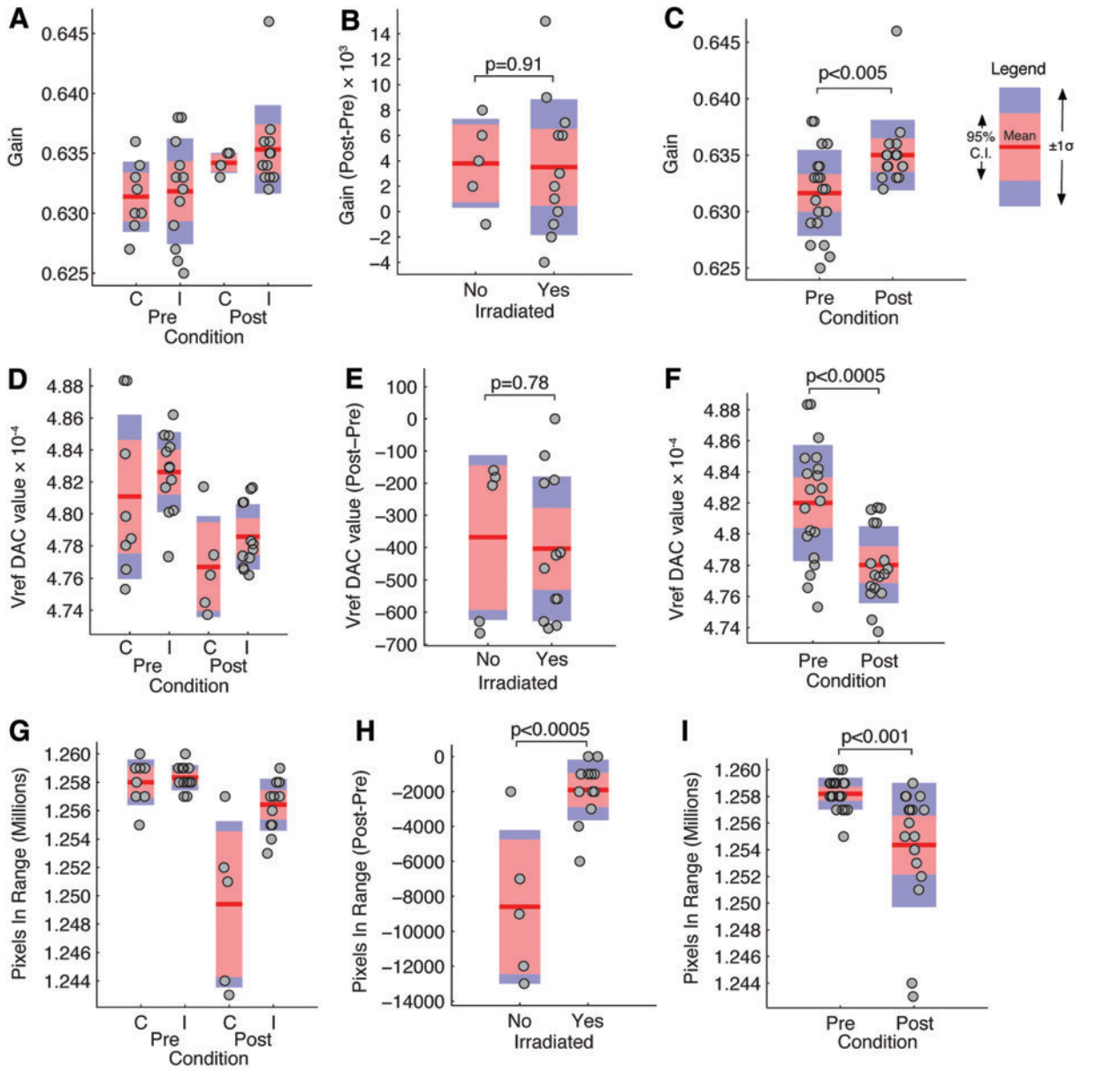


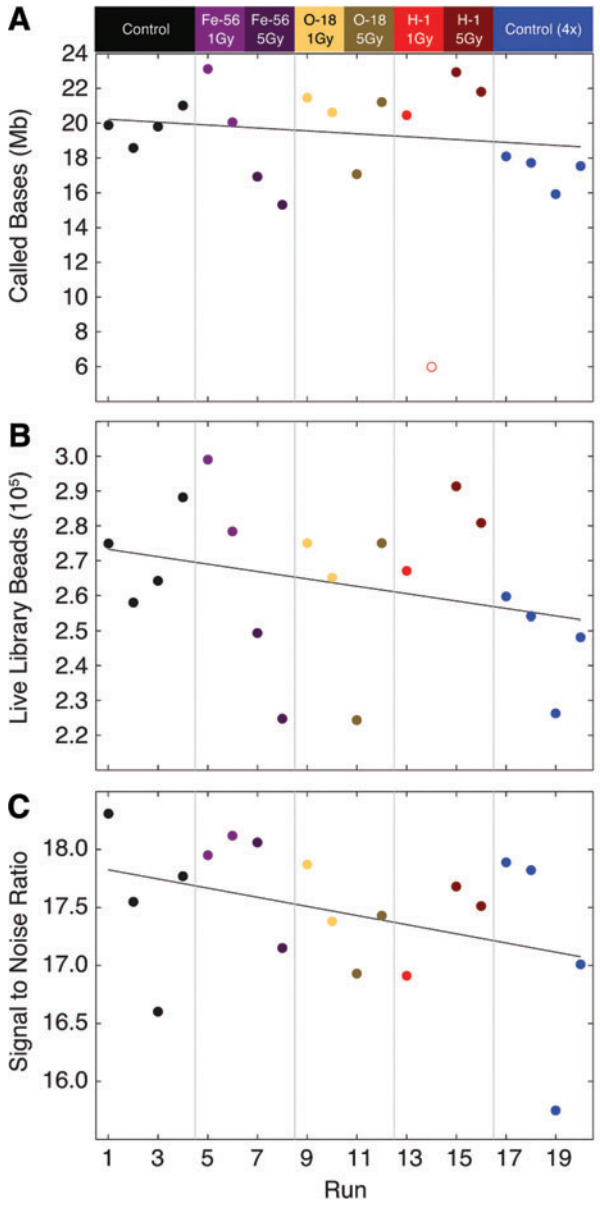
FIG. 2. Electrical testing results. (**A**) Gain values grouped by pre- or post-irradiation condition and chip assignment to control (C) or irradiation (I) subgroups. (**B**) Gain differences pre- and post-irradiation. (**C**) Gain with control and irradiated subgroups pooled. (**D–F**) Reference voltage DAC values as in (A–C). (**G–I**) Pixels in range values as in (A–C). All panels: Gain and reference voltage (Vref) are mean values across each chip as measured during electrical testing. For details see text and Supplementary Tables S4–S6. Color images available online at www.liebertonline.com/ast

second electrical test (Fig. 2I; post vs. pre, one-sided t test, $p=0.00054$).

The pH sensitivity test was performed on a sample of the various radiation conditions and control parts to assure that the sensor still functioned as a pH sensor. A measured value >50 mV/pH indicates the sensor is functioning properly. Due to sensitivity of the Tris buffer to temperature, variations of pH were observed from measurement to measurement in both control and irradiated parts. However, all tested parts were properly functioning as a pH sensor, above 50 mV/pH, and were within expected statistical variations (Supplementary Tables S4 and S5).

3.2. Sequencing

All chips successfully generated sequence data (Fig. 3A). Run 14 generated much lower sequence data because of a known preparation error and was excluded from further analysis. By ANOVA, we found no effect of ion ($p=0.61$) or dose ($p=0.71$) on signal-to-noise. To assess bead aging, we regressed run number against a number of performance metrics. We found a trend toward lower total called bases (-0.33 Mb/day, $p=0.37$), lower numbers of live library beads (Fig. 3B; -4.3×10^3/day, $p=0.22$), and lower signal-to-noise (Fig. 3C; -0.16/day, $p=0.11$) as beads aged, but no



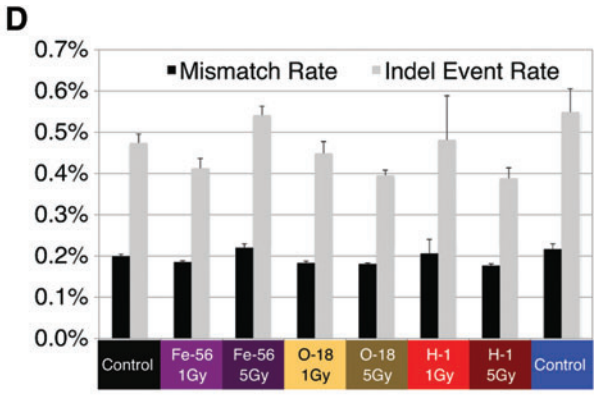


FIG. 3. Sequencing results. (**A**) Called bases (in megabases) across runs were consistently around 20 Mb. (**B**) Live library beads ranged from 0.2 to 0.3 million. (**C**) Signal-to-noise ratio declined across runs, although this effect was not significant. In (A–C), black lines are linear regressions as described in the text. (**D**) Error characterization after mapping of sequencing reads reveals consistent error rates with no consistent dose effect. Color images available online at www.liebertonline.com/ast

significant effect of run number on these or other metrics, including the fraction of beads that were live, the total called bases, the average number of reads per library bead, the mapped bases per live library bead, the mismatch rate, and the indel event rate. As expected from prior analysis, indels are the largest source of alignment error (Fig. 3D). However, the error rates did not vary significantly across runs by ANOVA when ion and dose were used as factors (mismatch rate $p = 0.09$, 0.22; indel event rate $p = 0.25$, 0.5, respectively). The ANOVA coefficients for mismatch rate were positive for control and Fe-56, negative for O-18 and H-1, and corresponded to a different manufacturing lot; a separate ANOVA for which only manufacturing lot was used as a factor was highly significant ($p = 0.0042$).

Although we found no functional impact of irradiation at the level of sequence accuracy, we also assessed whether irradiation might subtly modify nucleotide incorporation signals as measured by the ISFETs. We compared measured nucleotide incorporation dynamics during the first eight flows (TACGTACG), which are the flows required to read out the key sequence (TCAG) associated with library beads (Fig. 4A). Note that the peaks read out the key sequence, and while the T, C, and A peaks are expected to be of similar magnitude (because all live library beads have the key sequence), the final G peak is expected to be higher due to some beads having homopolymer incorporations during this flow. Subtle differences in the flows were observed between control chips (black, blue) and O-18 (yellow) or H-1 (red) chips but not Fe-56 (purple) chips (Fig. 4B). In all cases, 1 Gy and 5 Gy conditions were virtually identical. When controlling for manufacturing lot (Fig. 4C), we found that all the differences are associated with a specific lot, whereas the two other lots were virtually indistinguishable (Fig. 3D). The most parsimonious explanation is that these differences result from manufacturing variability, given the lot association and lack of dose effect.

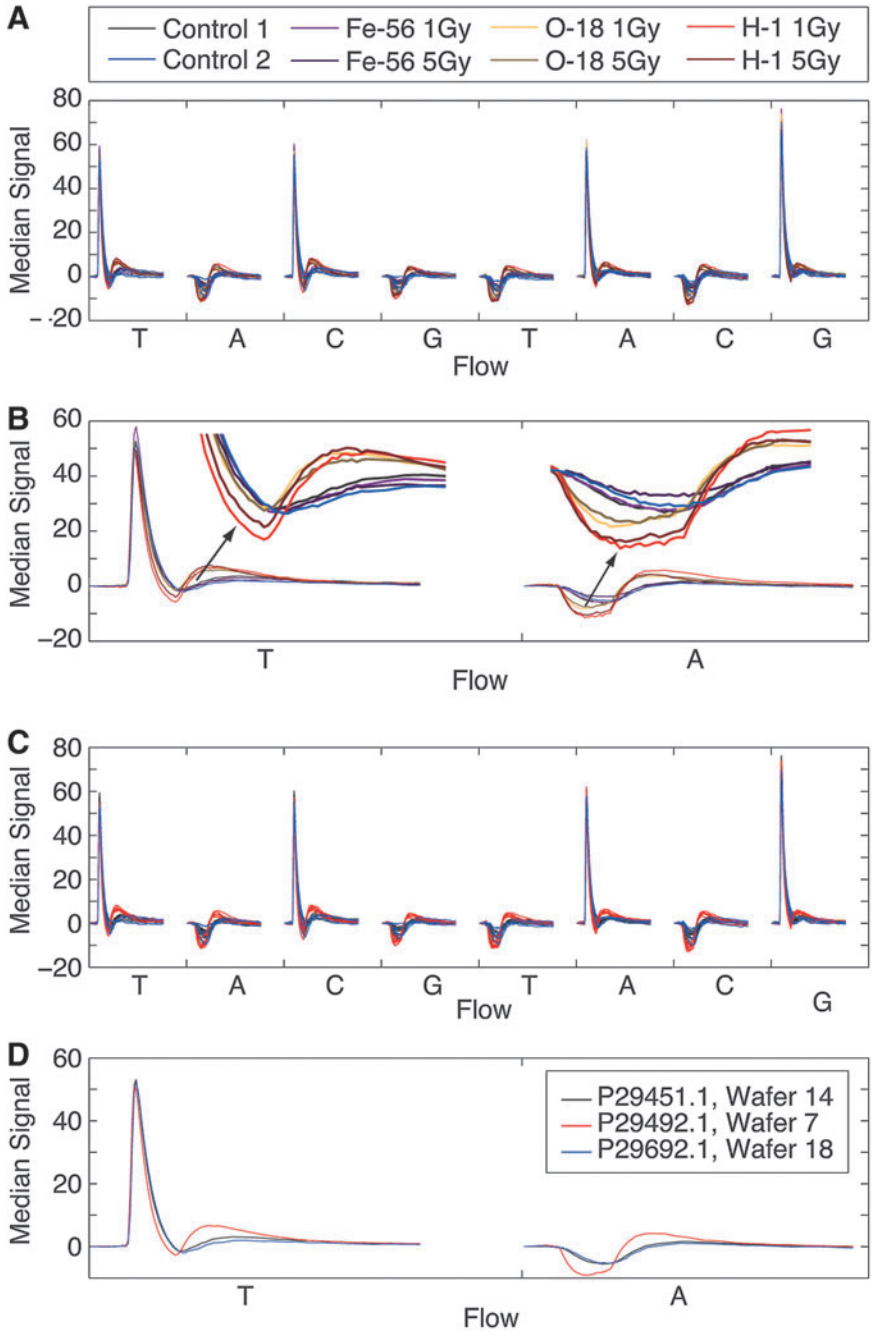


FIG. 4. No impact of irradiation on nucleotide incorporation dynamics. (**A**) Median voltage signals across all live beads during the first eight flows show minor differences apparently associated with O-18 and H-1 irradiation. (**B**) Details of the first two flows [mean across chips of median signal in (A)] show a lack of dose effect for O-18 and H-1 and Fe-56 dynamics similar to controls. (**C**) Same data in (A) colored according to manufacturing lot reveals that the differences are associated with manufacturing lot, not irradiation. (**D**) Detail view of first two flows [mean across chips of median signal in (C)] confirms this finding. Panels (A–B) share the top legend, while panels (C–D) share the legend in (D).

4. Discussion

An *in situ* life-detection instrument, as conceived here, would need to operate under pressure and controlled temperature to allow proper functioning of DNA polymerase during the sequencing process. Long-term (1+ years) storage of lyophilized reagents is now common and would provide adequate shelf life of reagents required for *in situ* sequencing. This leaves radiation resistance as the main requirement unique to the space environment. The lack of measurable effects of irradiation under the conditions studied suggests that *in situ* sequencing on Mars may be feasible with the semiconductor sequencing technology, at least from a total dose perspective.

In particular, the results of electrical testing for the controls, pre-irradiation, and post-irradiation parts are all within

normal statistical variation. Furthermore, there was no evidence that irradiation contributed to any changes in electrical parameters. Although there was a slight shift in pixels in range in irradiated chips, this shift was smaller than the shift in un-irradiated chips, suggesting that this shift was not associated with irradiation.

These findings indicate that irradiation did not cause charging of the ISFET sensor plate, damage to the transistors, or other shifts in the operating conditions of the sensor circuitry. Any shifts would have been observed as either nonfunctioning devices or significant shifts in pixels in range, gain, pH sensitivity, or reference electrode voltages. We can therefore conclude that under the conditions studied, irradiation did not affect device electrical performance.

The sequencing runs were very consistent, although total output was low compared to typical runs, which may generate 50–80 Mb when using equivalent reagents and chips. This may be associated with a poor amplification or loss of beads during the pooling process. In support of the former hypothesis, our pre-enrichment bead assay suggested that around 4% of beads were templated (ideal range 5–30%), corresponding to low polyclonality and many beads without clonal products. The post-enrichment value of 35% indicated a significantly reduced number of beads without clonal products but is still low, which is suggestive of poor emulsion PCR and/or bead performance. However, in the present experiment consistency was far more important than total output, and in this regard we achieved extremely stable output from run to run.

Much higher throughput is feasible during future *in situ* sequencing due to scaling of well numbers and improvements in read length. For example, commercially available sequencing chips for the PGM have 1.26, 6.34, and 11.3 million wells (314, 316, and 318 chips respectively), and commercially available sequencing kits achieve read lengths up to 400 bp. These chips have 3 mm wells, which have been scaled to 0.5 mm wells in chips compatible with the newer Ion Proton System. The commercially released Ion Proton I chip has 165 million wells, and future chips may have as many as 1.2 billion wells. This has implications for radiation survival: Under all conditions except the lowest fluence (Fe-56 1 Gy), the mean number of particles per well is more than one for the chips tested (314) but is approximately 100× lower for the Proton I chip (Fig. 5A). Thus, at 1 and 5 Gy doses, most wells of a Proton I chip would not be hit by an iron or oxygen ion, whereas most would be hit by a proton (Fig. 5B). Treating irradiation as a Poisson process, the variance in particles per well is equal to the average number of particles per well, so that at ~100× higher fluences we can expect that essentially all wells would be hit for all ions even in the Proton I chip.

Assume for a moment that the smaller features of the Proton I chip make it more sensitive to a given number of ions per well, resulting in failure and degradation in 80% and 15% of the wells, respectively. If using only the remaining 5% (8.25 million) of the wells and assuming 70% loading density, 95% live bead fraction, 60% final library reads at mean read length of 250 bp, an *in situ* run might generate 822 Mb, which is enough to sequence the *E. coli* DH10B genome at 176-fold coverage. This coverage is adequate for genome assembly; while it may not be possible to generate a complete, closed whole genome *de novo* from these reads, they are more than adequate to re-sequence the entire genome of a moderately sized prokaryote or generate nearly complete contigs via *de novo* assembly. Thus, if an *in situ* sampled environment contained a dominant organism, it may be possible to sequence its genome *in situ* even with 95% loss of sequencing chip wells. Our recent results illustrate that these numbers are reasonable; in one recent run using a 316 chip (6.34 million wells) we generated over 824 and 488 Mb of Q20 (1% predicted error rate) data, enough to sequence the *E. coli* DH10B genome at an average of 175-fold coverage, or 104-fold coverage at 1% error.

Inherent in the process of semiconductor sequencing is identifying which wells are performing. Current algorithms identify a number of different well states, including but not limited to whether a well contains a bead, whether that bead is a library bead or a control bead, and whether a bead contains a clonal product or is polyclonal. If impacts of radiation are identified at higher doses, or with other ions or energies, algorithms can be developed that adapt to induced

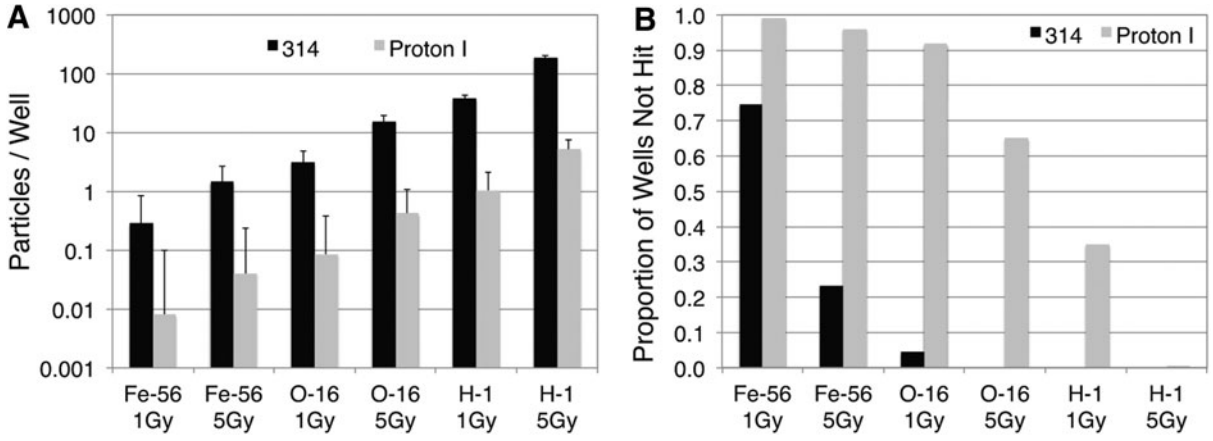


FIG. 5. Impact of sequencing chip well geometry and number. (**A**) Estimated particles/well for current (Ion 314) and future (Ion Proton I) sequencing chips using a Poisson model of irradiation. (**B**) Estimated proportion of wells not hit during irradiation, neglecting secondary particles. See text for chip geometry and well counts.

changes in measured nucleotide incorporation signals and, when interpretation may lead to sequencing errors, ignore data from damaged wells.

We have already demonstrated that the key reagents required for *in situ* sequencing, essentially those required for PCR, function after the same irradiation conditions described herein (Carr *et al.,* 2013). In that study, at Mars-relevant doses we found no apparent damage to DNA oligonucleotides nor decrease in our ability to PCR 1–1.5 kb products from irradiated genomic DNA. This suggests that DNA oligonucleotides, carried from Earth and used as part of the sequencing process, will not be adversely affected by space radiation. Similarly, we would not expect any DNA extracted *in situ* to be adversely affected by space radiation during processing, although space radiation may contribute to DNA degradation before sample collection in samples with less than on the order of 1 m of shielding.

In situ sequencing is a complex endeavor that will require extensive technology development in other areas. Nevertheless, we have now demonstrated that *in situ* semiconductor sequencing is a viable candidate for future missions such as those on the martian surface, where a lack of extensive trapped charged particles yields only moderate doses during a 2-year mission.

Beyond searching for life on Mars that is potentially related to life on Earth, *in situ* sequencing may be a viable approach to detect non-ancestrally related life, too, if it is based on RNA or DNA. Because complex organics, including nucleic acids and their precursors, have been identified in meteorites (Engel and Macko, 1997; Martins *et al.,* 2008; Schmitt-Kopplin *et al.,* 2010; Callahan *et al.,* 2011; Cooper *et al.,* 2011) and in interstellar space (Hollis *et al.,* 2000), delivery of this material to habitable environments could have steered the development of life toward these biomolecules. Thus, it makes sense to search for RNA- or DNA-based life within potential habitable zones even outside the context of meteoritic exchange, such as the probable liquid water oceans beneath Europa (Carr *et al.,* 1998; Kivelson *et al.,* 2000), Enceladus (Porco *et al.,* 2006; McKay *et al.,* 2008; Postberg *et al.,* 2011), and possibly Titan (Iess *et al.,* 2010, 2012; Baland *et al.,* 2011). Testing at higher doses and alternate conditions including gamma radiation is required to determine what kind of radiation hardening is required to enable *in situ* semiconductor sequencing in environments like Europa, where Jupiter's magnetosphere traps high-energy particles and leads to doses orders of magnitude higher than evaluated here. In contrast, Enceladus is of particular interest due to its potential habitability, easy access to subsurface material via sampling its liquid plume, and much less intense radiation environment.

Acknowledgments

We thank Peter Guida, Adam Rusek, and the rest of the NSRL-associated staff for assistance with the irradiations, and Alan Natapoff at MIT for statistical advice. This work was funded by NASA's ASTID program (NNX08AX15G).

Author Disclosure Statement

C.E.C., H.R., C.S.L., M.T.Z., and G.R. have no competing financial interests. I.Z., C.W.P., J.B., J.W.M., J.B., and J.M.R. are employees of Ion Torrent, and I.Z., C.W.P., J.B., J.B., and J.M.R. own shares of Life Technologies stock.

Abbreviations

ANOVA, analysis of variance; dNTP, deoxynucleotide triphosphate; indel, insertion/deletion; ISFETs, ion-sensitive field-effect transistors; LET, linear energy transfer; PCR, polymerase chain reaction; PGM, Personal Genome Machine; SPE, solar particle event.

References

Baland, R.M., Van Hoolst, T., Yseboodt, M., and Karatekin, Ö. (2011) Titan's obliquity as evidence of a subsurface ocean? *Astron Astrophys* 530:A141.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010) Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* Chapter 19, Unit 19.10.1–19.10.21.

Blazej, R.G., Kumaresan, P., and Mathies, R.A. (2006) Microfabricated bioprocessor for integrated nanoliter-scale Sanger DNA sequencing. *Proc Natl Acad Sci USA* 103:7240–7245.

Callahan, M.P., Smith, K.E., Cleaves, H.J., Ruzicka, J., Stern, J.C., Glavin, D.P., House, C.H., and Dworkin, J.P. (2011) Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proc Natl Acad Sci USA* 108:13995–13998.

Carr, C.E., Rowedder, H., Vafadari, C., Lui, C., Cascio, E., Zuber, M., and Ruvkun, G. (2013) Radiation resistance of biological reagents for *in situ* life detection on Mars. *Astrobiology* 13: 68–78.

Carr, M.H., Belton, M.J., Chapman, C.R., Davies, M.E., Geissler, P., Greenberg, R., McEwen, A.S., Tufts, B.R., Greeley, R., Sullivan, R., Head, J.W., Pappalardo, R.T., Klaasen, K.P., Johnson, T.V., Kaufman, J., Senske, D., Moore, J., Neukum, G., Schubert, G., Burns, J.A., Thomas, P., and Veverka, J. (1998) Evidence for a subsurface ocean on Europa. *Nature* 391:363–365.

Cooper, G., Reed, C., Nguyen, D., Carter, M., and Wang, Y. (2011) Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. *Proc Natl Acad Sci USA* 108:14015–14020.

Engel, M.H. and Macko, S.A. (1997) Isotopic evidence for extraterrestrial non-racemic amino acids in the Murchison meteorite. *Nature* 389:265–268.

Fritz, J., Artemieva, N., and Greshake, A. (2005) Ejection of martian meteorites. *Meteorit Planet Sci* 40:1393–1411.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W.J., and Nekrutenko, A. (2005) Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* 15:1451–1455.

Gladman, B.J. and Burns, J.A. (1996) Mars meteorite transfer: simulation. *Science* 274:161–165.

Gladman, B.J., Burns, J.A., Duncan, M., Lee, P., and Levison, H.F. (1996) The exchange of impact ejecta between terrestrial planets. *Science* 271:1387–1392.

Goecks, J., Nekrutenko, A., Taylor, J., and Team, G. (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11:R86.

Hollis, J., Lovas, F., and Jewell, P. (2000) Interstellar glycolaldehyde: the first sugar. *Astrophys J* 540:L107–L110.

Horneck, G., Stöffler, D., Ott, S., Hornemann, U., Cockell, 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.

Iess, L., Rappaport, N.J., Jacobson, R.A., Racioppa, P., Stevenson, D.J., Tortora, P., Armstrong, J.W., and Asmar, S.W. (2010) Gravity field, shape, and moment of inertia of Titan. *Science* 327:1367–1369.

Iess, L., Jacobson, R.A., Ducci, M., Stevenson, D.J., Lunine, J.I., Armstrong, J.W., Asmar, S.W., Racioppa, P., Rappaport, N.J., and Tortora, P. (2012) The tides of Titan. *Science* 337: 457–459.

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

Isenbarger, T.A., Finney, M., Rios-Velazquez, C., Handelsman, J., and Ruvkun, G. (2008b) Miniprimer PCR, a new lens for viewing the microbial world. *Appl Environ Microbiol* 74:840–849.

Kivelson, M.G., Khurana, K.K., Russell, C.T., Volwerk, M., Walker, R.J., and Zimmer, C. (2000) Galileo magnetometer measurements: a stronger case for a subsurface ocean at Europa. *Science* 289:1340–1343.

Le Postollec, A., Incerti, S., Dobrijevic, M., Desorgher, L., Santin, G., Moretto, P., Vandenabeele-Trambouze, O., Coussot, G., Dartnell, L., and Nieminen, P. (2009) Monte Carlo simulation of the radiation environment encountered by a biochip during a space mission to Mars. *Astrobiology* 9:311–323.

Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595.

Martins, Z., Botta, O., Fogel, M.L., Sephton, 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.

McKay, C.P., Porco, C.C., Altheide, T., Davis, W.L., and Kral, T.A. (2008) The possible origin and persistence of life on Enceladus and detection of biomarkers in the plume. *Astrobiology* 8:909–919.

Metzker, M.L. (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46.

Ning, Z., Cox, A.J., and Mullikin, J.C. (2001) SSAHA: a fast search method for large DNA databases. *Genome Res* 11:1725–1729.

Porco, C.C., Helfenstein, P., Thomas, P.C., Ingersoll, A.P., Wisdom, J., West, R., Neukum, G., Denk, T., Wagner, R., Roatsch, T., Kieffer, S., Turtle, E., McEwen, A., Johnson, T.V., Rathbun, J., Veverka, J., Wilson, D., Perry, J., Spitale, J., Brahic, A., Burns, J.A., Delgenio, A.D., Dones, L., Murray, C.D., and Squyres, S. (2006) Cassini observes the active south pole of Enceladus. *Science* 311:1393–1401.

Postberg, F., Schmidt, J., Hillier, J., Kempf, S., and Srama, R. (2011) A salt-water reservoir as the source of a compositionally stratified plume on Enceladus. *Nature* 474:620–622.

Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., Hoon, J., Simons, J.F., Marran, D., Myers, J.W., Davidson, J.F., Branting, A., Nobile, J.R., Puc, B.P., Light, D., Clark, T.A., Huber, M., Branciforte, J.T., Stoner, I.B., Cawley, S.E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J.A., Namsaraev, E., McKernan, K.J., Williams, A., Roth, G.T., and Bustillo, J. (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475:348–352.

Ruvkun, G., Finney, M., Church, G., Zuber, M., and Gilbert, W. (2002) A Robotic-PCR detector for DNA-based life on other planets. In *Signs of Life: A Report Based on the April 2000 Workshop on Life Detection Techniques,* Space Sciences Board, The National Academies Press, Washington, DC, pp 137–140.

Schmitt-Kopplin, P., Gabelica, Z., Gougeon, R., Fekete, A., Kanawati, B., Harir, M., Gebefuegi, I., Eckel, G., and Hertkorn, N. (2010) High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc Natl Acad Sci USA* 107:2763–2768.

Shimada, K., Yano, M., Shibatani, K., Komoto, Y., Esashi, M., and Matsuo, T. (1980) Application of catheter-tip i.s.f.e.t. for continuous *in vivo* measurement. *Med Biol Eng Comput* 18:741–745.

Shuster, D.L. and Weiss, B.P. (2005) Martian surface paleotemperatures from thermochronology of meteorites. *Science* 309:594–600.

Smith, T. and Waterman, M. (1981) Identification of common molecular subsequences. *J Mol Biol* 147:195–197.

Walker, C., Boyle, D., Jennings, S., Ellis, J., and Ellis, R. (1990) A comparison of the radiation tolerance of the STC bipolar, CMOS and merged bipolar/CMOS processes. *Journal of Electronic Materials* 19:617–627.

Weiss, B.P., Kirschvink, J.L., Baudenbacher, F.J., Vali, H., Peters, N.T., Macdonald, F.A., and Wikswo, J.P. (2000) A low temperature transfer of ALH84001 from Mars to Earth. *Science* 290:791–795.

Address correspondence to:
Gary Ruvkun
MGH Department of Molecular Biology
Simches Research Building
185 Cambridge Street
CPZN-7250
Boston, MA 02114

E-mail: ruvkun@molbio.mgh.harvard.edu

Submitted 9 September 2012
Accepted 19 March 2013