IN-SITU SEQUENCING FOR LIFE DETECTION, HUMAN HEALTH, & PLANETARY PROTECTION.
C. E. Carr[1-3,*], G. Ruvkun[2-3] and M. T. Zuber[1]. [1]MIT Department of Earth, Atmospheric and Planetary Sciences, Cambridge MA. [2]Massachusetts General Hospital, Department of Molecular Biology, Boston, MA. [3]Harvard Medical School, Department of Genetics. *Correspondence to chrisc@mit.edu.

Introduction: Nucleic acid sequencing provides a powerful approach to characterize living systems that can address several fundamental elements of future robotic and human Mars exploration including: 1) Searching for life beyond Earth, 2) Assessing environmental and human health, and 3) Monitoring planetary protection. We first describe the best current technologies for in-situ sequencing, then describe how sequencing enables these activities.

Sequencing technology: The extreme mass, power, and volume constraints for in-situ sequencing make most modern sequencing technologies [1] inappropriate. While a microfluidic implementation of traditional Sanger sequencing [2] is potentially compatible, it is slow and not easily parallelizable. There are two current potential alternatives.

Semiconductor sequencing (SS) chips [3] can enable massively parallel sequencing, and are small (3 cm x 3 cm), fast (<2 hrs), and optics-free (Fig. 1A). These chips read out a nucleic acid sequence by synthesis, e.g. incorporation of nucleotides to generate a growing double-stranded DNA (dsDNA) from a single-stranded original. Nucleotide incorporation is sensed by detecting the pulse of hydrogen ions released during nucleotide incorporation. Each well within a multi-million well chip reads out a different sequence, enabling millions of reads in parallel. The current unoptimized reagent volume per run is around 0.5L. The most limiting aspect to these chips is the sample prepation, which requires assembling a sequencing library (DNA with known ends), amplifying this library on beads, and loading the beads into the sequencing chip. While this approach can achieve single-molecule sensitivity, it adds complexity and amplification bias.

Nanopore sequencing (NS) [4-7], not yet commercially available, may address these issues by eliminating the need for amplification and simplifying sample preparation, while further reducing mass, power, and volume (Fig. 1B). While reported (Oxford Nanopore) error rates are high (4%), read lengths up to tens of kb will facilitate contig and genome assembly.

The Search for Life Beyond Earth: If life exists beyond Earth and uses nucleic acids as informational polymers (Fig. 1C), we can use sequencing to characterize it with exquisite detail.

Is life on mars related to Life on Earth? Theoretical [8, 9] and experimental [10-13] studies suggest that viable microbes could have been transferred during meteoritic exchange between Earth and Mars; if so, any life there may be related to us and based on DNA or its likely precursor, RNA. Whole genome amplification approaches have enabled sequencing of single cells [14, 15], although in challenging samples realistic detection limits may be 10^2 cells [16]. NS will likely offer improved detection limits and avoid the biases of whole genome amplification. If the common ancestor lived in the RNA world, life on Mars might still be based on RNA, and could be reverse transcribed to DNA and sequenced. Direct RNA sequencing [17] in-situ may become feasible using NS approaches.

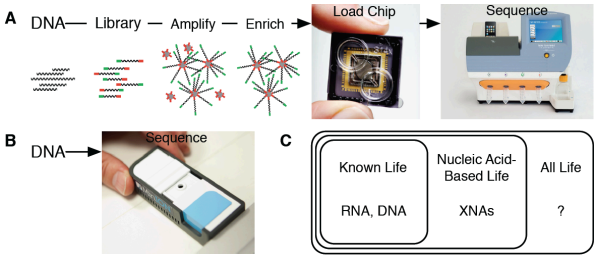

Fig. 1. A. Semiconductor sequencing workflow. **B.** Nanopore sequencing workflow. **C.** Hierarchy of life.

Sequence analysis can reveal signatures of common ancestry, such as conserved ribosomal sequence or structure, and identify potential contaminants. Protein-space comparisons such as BLASTX can be used to search for conserved protein modules. Both SS and NS may offer the ability to detect proteins through sequencing of DNA aptamers, molecules selected for their ability to bind specifically to their targets. The protein world may overlap with all three regions of the hierarchy of life (Fig. 1C); thus, protein detection may broaden the potential of sequencing for life detection.

A second genesis? If habitable environments beyond Earth evolved life independently, such life may also be based on nucleic acids. Habitable environments may share common sources of organic material, derived from organic synthesis in the stellar nebula [18, 19] and delivered through cometary and meteoritic impacts. These common sources may have biased the evolution of life towards common solutions.

Non RNA/DNA based life may be based on related polymers such as TNA [20]. Reading and writing between DNA and non-standard (xeno) nucleic acids (XNAs) is now possible [21]. Thus, XNAs could be converted to DNA and sequenced using SS chips, or, possibly, directly sequenced using NS.

These capabilities open the door to generic sequencing of a host of informational polymers that might be utilized by life beyond Earth, whether related to us or not. Thus, in-situ sequencing can be applied to the search for life in potential subsurface habitable environments on Europa or Enceladus. Enceladus is particularly attractive because sequencing could be performed in orbit or after flyby(s) by collecting and minimally processing water from its plume.

Environmental and Human Health: In-situ sequencing can inform toxicological characterization of the Mars environment during robotic missions, as well as enable monitoring of environmental and human health during long-duration spaceflight.

Characterizing environmental toxicity on Mars. Conventional toxicity testing relies upon chemical testing or testing in biological systems [22], e.g. animal models in which characterization of absorption, distribution, metabolism, and excretion (ADME) is performed. While some chemical analysis may be feasible in-situ, ADME toxicology screening is infeasible. However, in-situ sequencing enables a variety of toxicogenomics approaches including gene expression analysis (RNA-seq), epigenomics (NS identification of DNA methylation), and proteomics (using aptamers).

It was previously proposed that biological toxicity tests of Martian soil could be carried out in-situ [23], but in-situ sequencing may now enable a detailed readout of mechanism. For example, cells such as macrophages (used to study the lung inflammatory response to lunar and Martian simulants [24]), can be transported to Mars freeze-dried, then rehydrated with liquids exposed to relevant soil fractions. Differential gene expression, e.g. relative to a control, could identify the molecular response to exposure. Similarly, the response to space radiation could be monitored, with a suitable "no-radiation" control on Earth.

Monitoring human and environmental health. Sequencing for clinical diagnosis is being actively implemented. Sequencing could be used during future space missions to identify viral and microbial infections as well as to diagnose or rule out cancer. For example, the Ion Torrent SS platform has been used to de-novo sequence the new *E. coli* strain (O104:H4) responsible for the 2011 outbreak in Germany [25] and to detect a wide variety of mutations common in human cancers (Ion AmpliSeq™ Cancer Panel). Establishing the prognosis of an astronaut (whether a new probable melanoma is benign or malignant) could have significant value to the individual and the mission.

NASA has begun to explore applications of synthetic biology to space exploration such as biological in-situ resource utilization (ISRU): sequencing provides a valuable approach to monitor these systems.

Planetary Protection: NASA's broad planetary protection goal is to minimize biological cross-contamination during exploration. In-situ sequencing can contribute to this goal by verifying the level of forward contamination at the RNA, DNA, or (via aptamers) at the protein level. Detecting any forward contamination can also help to eliminate potential false positives during life detection activities. Furthermore, if life beyond Earth is based on nucleic acids and/or proteins, in-situ sequencing could be used to assess the risk of back contamination during two-way missions.

Conclusions: Rapid advances in high-throughput sequencing, specifically semiconductor sequencing and nanopore sequencing, are enabling in-situ sequencing. This capability will broadly support the search for life beyond Earth, assessments of environmental and human health, and planetary protection. The applications described here are predicated upon massively parallel sequencing and require substantial development to adapt them for in-situ use, particularly sample preparation. These technologies are already being applied to similar applications on Earth and have obvious benefits for studying the diversity of life, improving clinical outcomes, and minimizing forward contamination.

References: [1] Metzker M. L. (2010) Nat Rev Genet, 11(1), 31-46. [2] Blazej R. G. et al. (2006) PNAS, 103(19), 7240-5. [3] Rothberg J. M. et al. (2011) Nature, 475(7356), 348-52. [4] Venkatesan B. M., Bashir R. (2011) Nat Nano, 6(10), 615-4. [5] Postma H. (2010) Nano Letters, 10(2), 420-5. [6] He H. et al. (2012) J. Phys. Chem. C, 112(10), 3456-9. [7] Manrao E. A. et al. (2012) Nat Biotech, 30, 349-53. [8] Melosh H. J. (1988) Nature, 332(6166), 687-8. [9] Mileikowsky C. et al. (2000) Icarus, 145(2), 391-427. [10] Weiss B. P. et al. (2000) Science, 290(5492), 791-5. [11] Horneck G. et al. (2008) Astrobiology, 8(1), 17-44. [12] Burchell M. J. et al. (2004) MNRAS, 352(4), 1273-8. [13] Horneck G. et al. (2001) Orig Life Evol Biosph, 31(6), 527-47. [14] Woyke T. et al. (2010) PLoS ONE, 5(4), e10314. [15] Walker A. and Parkhill J., (2008) Nat Rev Microbiol, 6, 176-7. [16] Carr C. et al. (2012) Life Detection in Extraterrestrial Samples 2012, #6044. [17] Ozsolak F. et al. (2009) Nature, 461(7265), 814-8. [18] Ciesla F. J. and Sandford S. A. (2012) Science, 336(6080), 452-4. [19] Nuevo M. et al. (2012) Astrobiology, 12(4), 295-314. [20] Schöning K. et al. (2000) Science, 290(5495), 1347-51. [21] Pinheiro V. B. et al. (2012) Science, 336(6079), 341-4. [22] Mahadevan B. et al. (2011) Env Mol Mutagen, 52(5), 339-54. [23] Morgenthaler G. W. et al., (1998), ASCE Conf Proc doi:10.1061/40339(206)3. [24] Chatterjee A. et al. (2010) J. Tox Env Health A, 73(9), 623-6. [25] Rohde H. et al. (2011) N Engl J Med, 365(8), 718-724.