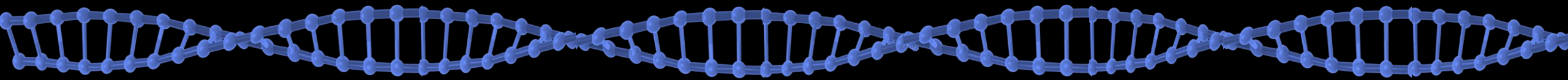# **Biomolecule Sequencer**:
Next-Generation DNA Sequencing Technology for In-flight Environmental Monitoring, Research, and Beyond

David J. Smith, Ph.D.
NASA Ames Research Center
ASGSR, Health Care Technologies Sessions
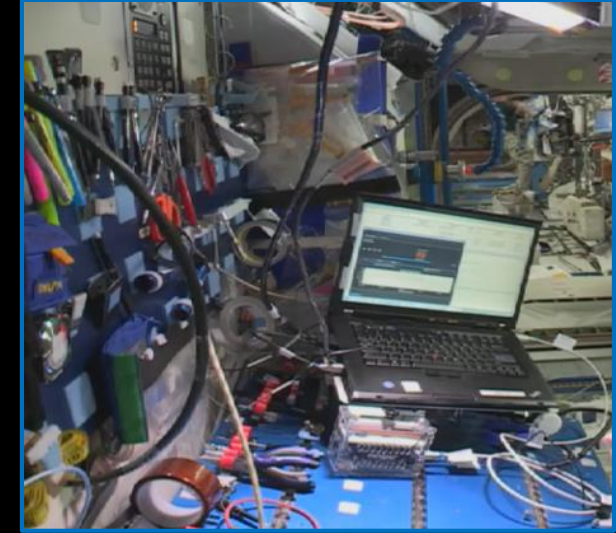October 28, 2016

2016: The Molecular Space Age

April 19: The first molecular biology assay in space is completed, as DNA is amplified using the miniPCR™ thermal cycler

April 29: RNA isolation, reverse transcription, and DNA amplification data obtained on the ISS using the Wetlab-2 qPCR platform

2016

2016

2016: The Molecular Space Age Continues

MinION™ by Oxford Nanopore Technologies

Biomolecule Sequencer Payload

- **First attempt at DNA sequencing in the microgravity environment of space**

- Enabled by the MinION™, developed by Oxford Nanopore Technologies
- COTS miniature DNA Sequencer
- 3 ¾ x 1 ¼ x 5/8 inches
- Less than 120 grams (with USB cable)
- Powered via USB connection
- Capable of DNA, RNA, and protein sequencing

Biomolecule Sequencer: The Payload

Johnson Space Center Class 1E Payload

- ~1 year to certify a payload for flight
- Class 1E is a streamlined certification process
- Reduce the time it takes to get scientific payloads to the ISS and increase utilization as a National Laboratory

- Authority to proceed February, 2015
- Hardware delivered on December 18, 2015
- Launched July 18, 2016 (SpaceX CRS-9)
- Technology Demonstration operations occurred on August 26, Sept. 3 and Sept. 7, 2016
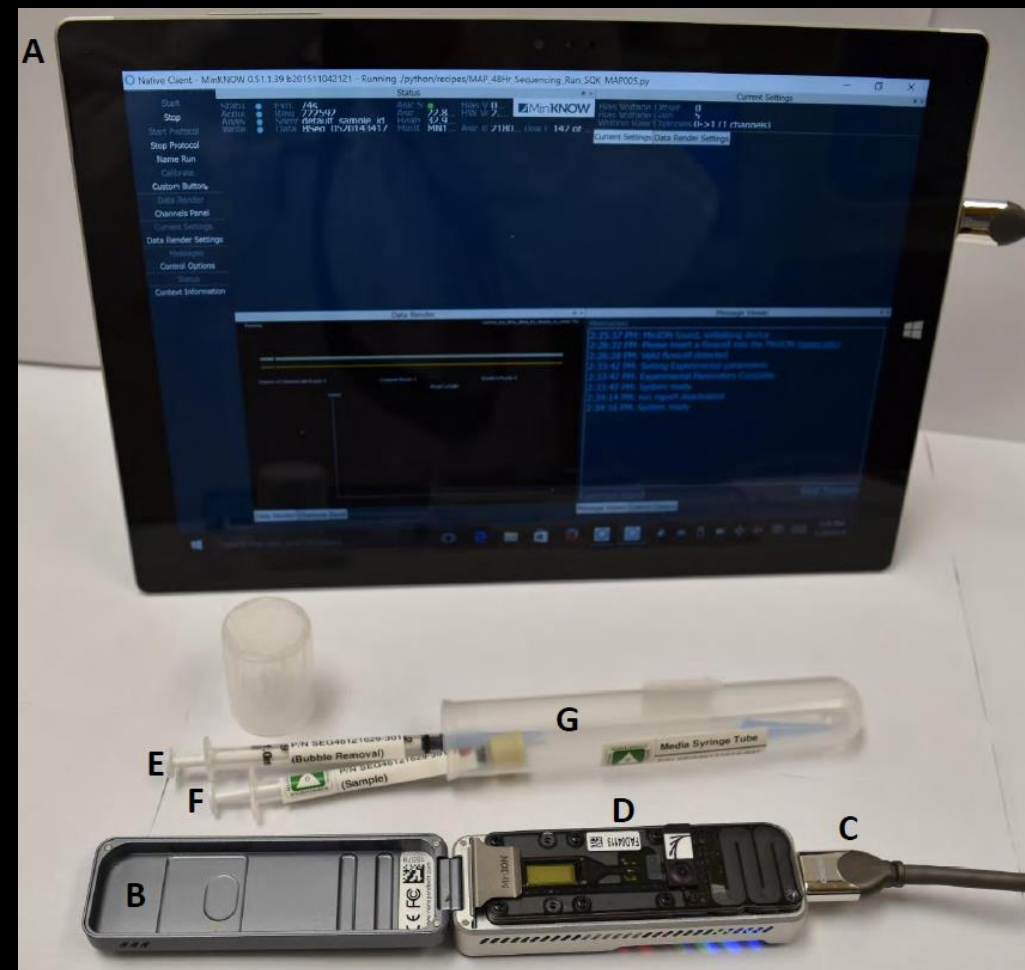
Biomolecule Sequencer: The Team

Payload Development

- Aaron Burton, Ph.D. (PI)
  NASA JSC
- Sarah Castro-Wallace, Ph.
  NASA JSC
- Kristen John, Ph.D. (Depu
  and PE)
  NASA JSC
- Sarah Stahl, M.S., (PS)
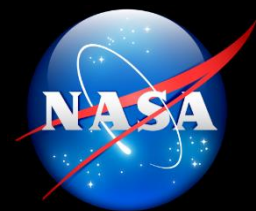 NASA JSC

Astronaut

- Kate Rubins, Ph.D.
  NASA JSC

External Science Team

- Charles Chiu, Ph.D. (UCSF)
  - Scot Federman
  - Sneha Somasekar
  - Doug Stryke
  - Guixia Yu

- Chris Mason, Ph.D. (WCMC)
 - Noah Alexander
  - Alexa McIntyre

Biomolecule Sequencer: The Need for In-flight Sequencing

Why do we need a DNA sequencer to support the human exploration of space?

- Operational environmental monitoring
  - Identification of contaminating microbes
  - Infectious disease diagnosis
- Research
  - Human
  - Animal
  - Microbes/Cell lines
  - Plant
- Med Ops
  - Response to countermeasures
  - Radiation
- Functional testing for integration into robotics for Mars exploration missions

Clogged SRV-K line

FGB panel
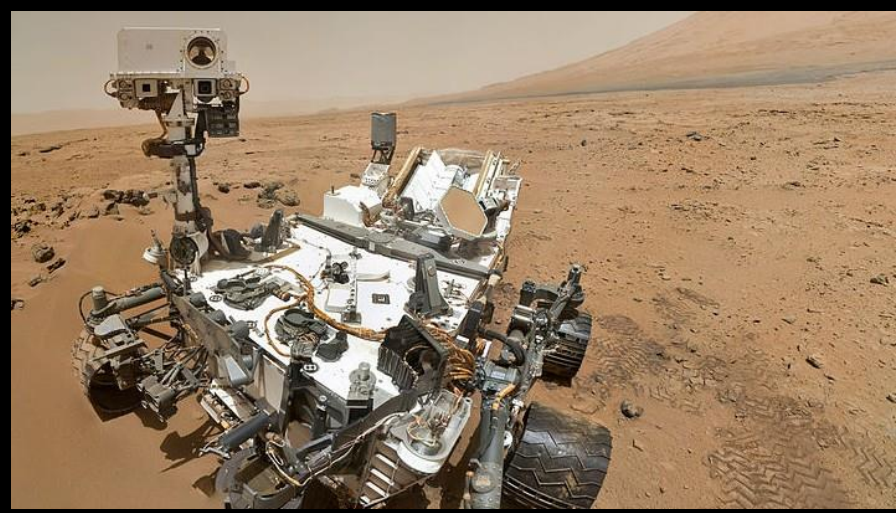
Biomolecule Sequencer: The Benefits

- Benefits to In-flight Sequencing
  - Sequencing on the ISS can inform real-time decisions (remediation strategies, research, med ops, etc.)
  - Unlike other technologies, sequencing is not limited to the detection of specific targets, but rather will provide data on the entirety of a sample
  - Reduce down mass (sample return for environmental monitoring, crew health, etc.)
  - Real-time analysis can influence medical intervention
  - Support astrobiology science investigations
    - Technology superiorly suited to *in situ* nucleic acid-based life detection
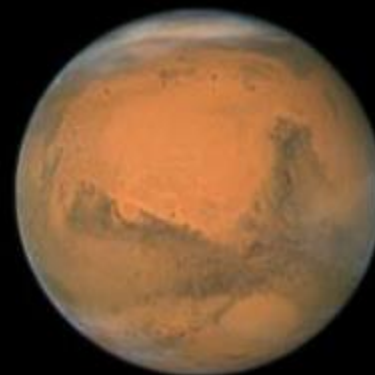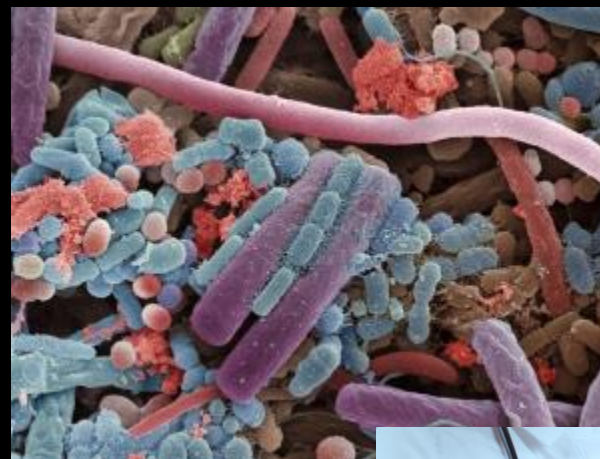
Biomolecule Sequencer: The Hardware
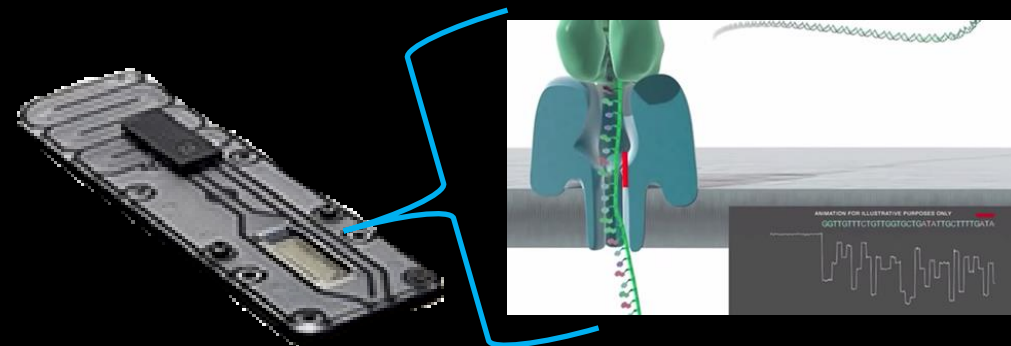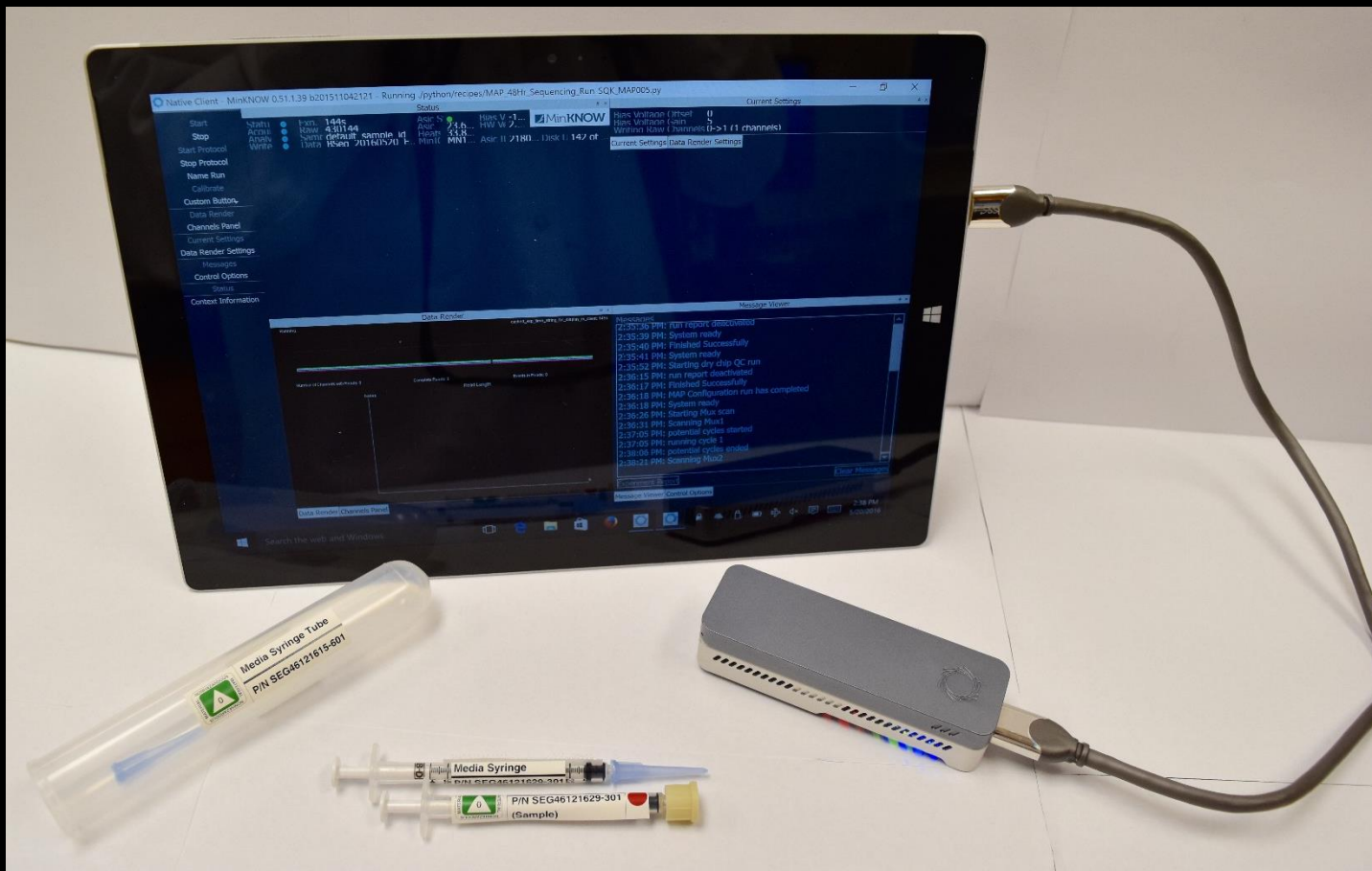
Flow Cell: Contains the nanopore sensing technology that is required to perform the sequencing reaction

Biomolecule Sequencer: Nanopore Sequencing

Nanopore-based sequencers measure changes in current caused by DNA strands migrating through the pore. The changes in current are characteristic of the sequence of migrating DNA.
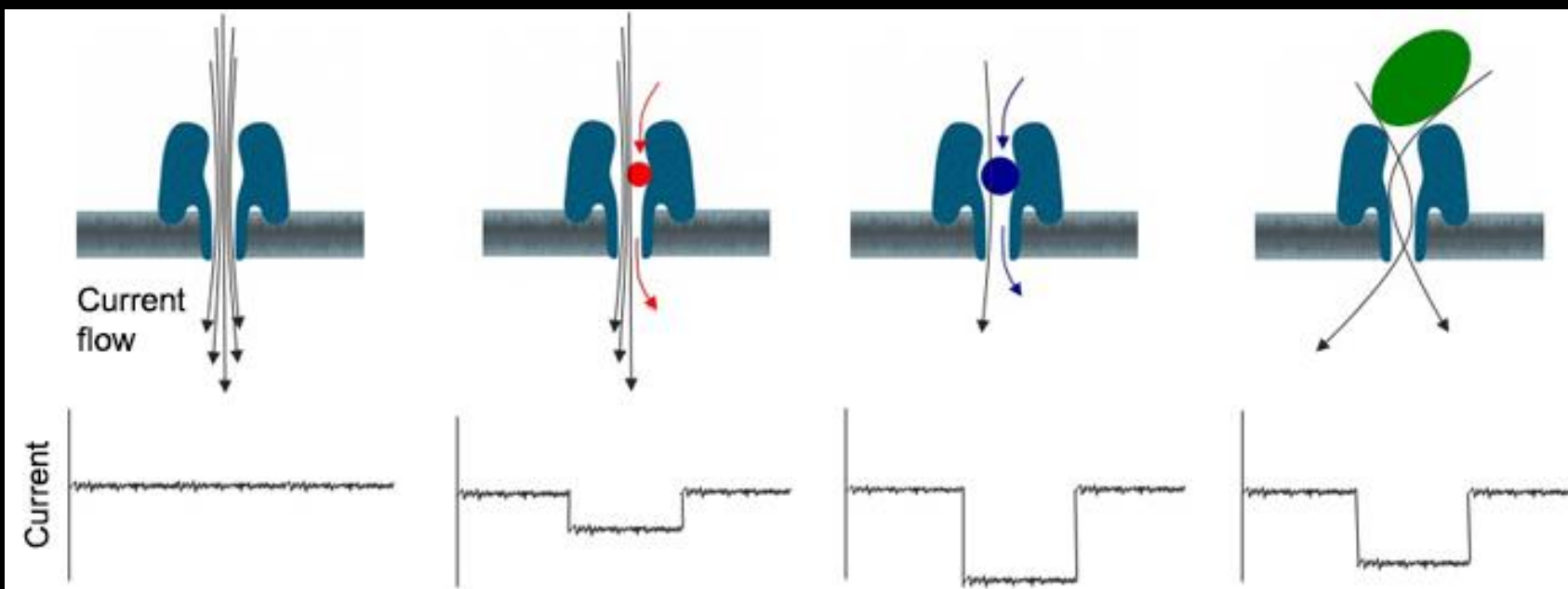
Biomolecule Sequencer: Comparison to other Sequencers

Manufacturer	Model	Sequencing methodology	Mass-based cost to transport hardware to the ISS	Mass of sequencer hardware (lb)	Volume (cu in)	Power requirements	Special requirements
Illumina	MiSeq	Fluorescence: each nucleotide has a different fluorophore	$1,200,000	120	6,940	100 - 240V AC, 10A, 400 W	
Illumina	MiniSeq	Fluorescence: each nucleotide has a different fluorophore	$990,000	99	12,350	100 - 240V AC, 15A	
PacBio	Sequel	Fluorescence: each nucleotide has a different fluorophore	$7,800,000	780	81,900	208 - 240V AC, 30A	Pressurized N_2 (50 PSI)
IonTorrent	Ion PGM	Electrochemical: measures voltage changes caused by nucleotide addition	$650,000	65	10,080	100 - 240V AC, 9A, 200 - 300 W	Pressurized N_2 (35 - 45 PSI)
Oxford Nanopore Technologies	MinION	Electrochemical: measures DNA-mediated changes in current passing through nanopores	$2,000	0.2	5	USB3 (5V, < 1A) 1 W	

Biomolecule Sequencer: DNA Samples

Project Goals:

1. Test the basic functionality by comparing ISS sequencing results of pre-determined samples to ground results
2. Evaluate crew operability and potential for degrees of autonomy

Experimental Design:

Sequence a ground-prepared sample containing a mixture of genomic DNA from:

- Bacteriophage lambda
- *Escherichia coli*
- Mouse – BALB/C (female)

National Aeronautics and Space Administration
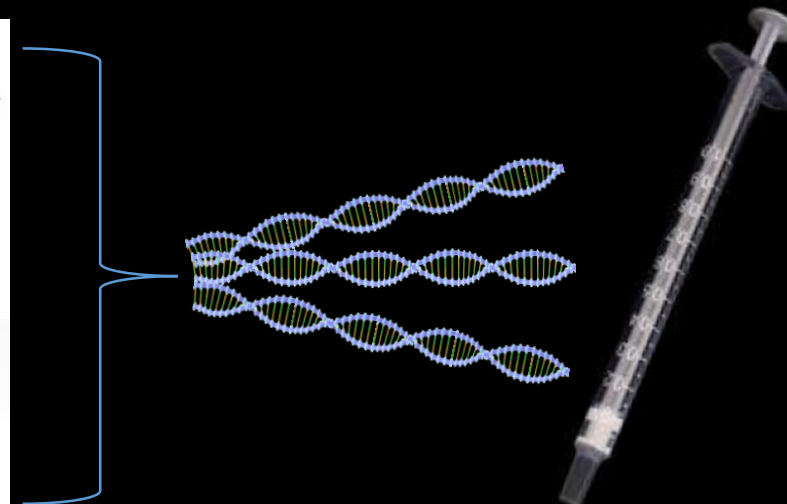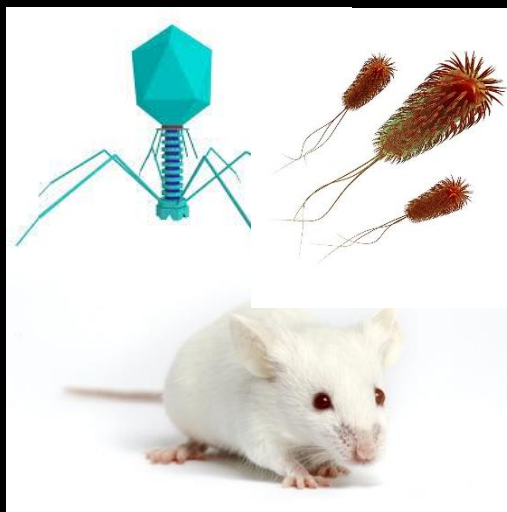
4. Destow and connect MinION to Surface Pro3

5. Sample injection

6. Dispose of sample syringe

7. Initiate the sequencing experiment
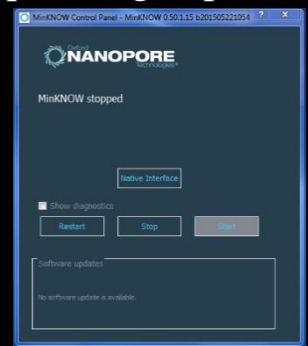
8. Data collection
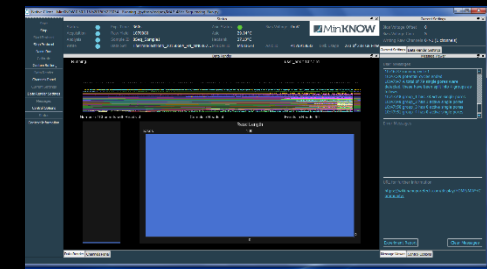
Sequencing stops after 6 hrs

3. Remove flow cell and sample syringe from cold stowage and allow to equilibrate to ambient ISS temperature

9. Stow used flow cell for return

2. Stowage (Ambient, -90°C, + 4°C)
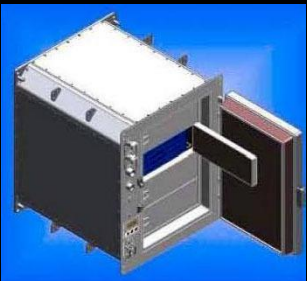
10. Data downlink

The First DNA Sequencer in Space!
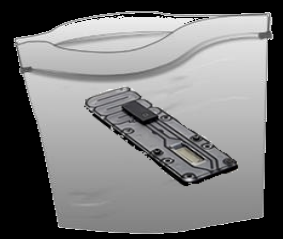
11. Stow MinION, Surface Pro3, power & USB cords

1. Launch packaged items

12. Return of payload
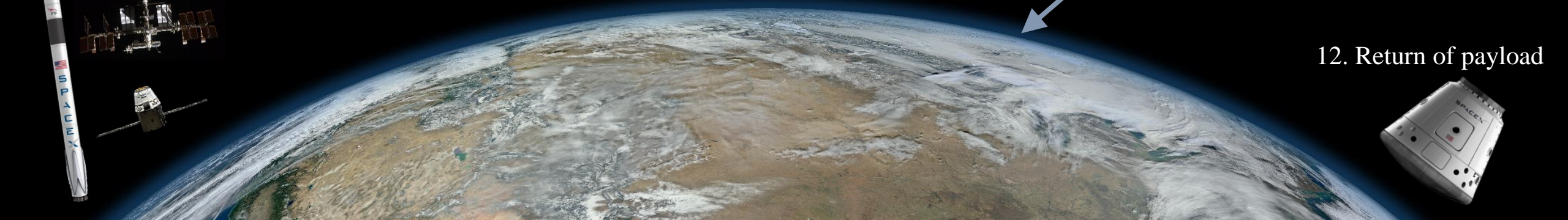
Biomolecule Sequencer: ISS Test Demo

- 4 batches of libraries were prepared containing the genomes to be sequenced
- From the 4 batches, 18 samples were produced: 9 for ISS and 9 identical ground controls
- Aliquots of all libraries were sequenced for quality control
- Synchronous ground controls were performed
- 3 sequencing experiments have been conducted to date (Aug. 26, Sept. 3, Sept. 7, 2016); additional runs are planned

National Aeronautics and Space Administration

Did it work?

Yes!

The emergence of nanopore-based sequencers greatly expands the reach of sequencing into low-resource field environments, enabling in situ molecular analysis. In this work, we evaluated the performance of the MinION DNA sequencer (Oxford Nanopore Technologies) in-flight on the International Space Station (ISS), and benchmarked its performance off-Earth against the MinION, Illumina MiSeq, and PacBio RS II sequencing platforms in terrestrial laboratories. The samples contained mixtures of genomic DNA extracted from lambda bacteriophage, Escherichia coli (strain K12) and Mus musculus (BALB/c). The in-flight sequencing experiments generated more than 80,000 total reads on with mean 2D accuracies of 85 to 90%, mean 1D accuracies of 75 to 80%, and median read lengths of approximately 6,000 bases. We were able to make directed assemblies of the ~4.7 Mb E. coli genome, ~48.5 kb lambda genome, and a representative M. musculus sequence (the ~16.3 kb mitochondrial genome), at 100%, 100%, and 96.7% pairwise identity, and de novo assemblies of the lambda and E. coli genomes solely with yielded 100% and 99.8% genome coverage, respectively, at 100% and 98.5% pairwise identity. Across all surveyed metrics (base quality, throughput, stays/base, skips/base), **no observable decrease in MinION performance was observed while sequencing DNA in space**. Simulated runs of in-flight nanopore data using an automated bioinformatic pipeline demonstrated the feasibility of real-time sequencing analysis and metagenomic identification of microbes in space. Additionally, cloud and laptop based-assembly illustrated the plausibility of automated, de novo genomic assembly from nanopore data on the ISS. Applications of sequencing for space exploration include infectious disease diagnosis, environmental monitoring, evaluating biological responses to spaceflight, and even potentially the detection of extraterrestrial life on other planetary bodies.

(A) A mixture of equimolar DNA from mouse, *E. coli* and lambda phage genomes was sequenced in parallel on Earth ("Ground") and in-flight on the ISS (after being delivered by a SpaceX Dragon capsule). Synchronous nanopore sequencing runs were performed from August 26 to September 13, 2016.

(B) Plot of mean current intensity in picoAmperes (pA; Y-axis) against k-mers (x-axis) in order of increasing mean current based on a model distribution from Oxford Nanopore Technologies (black). Current distributions are tightly clustered with the exception of lower-quality ground #2.

(C) Density plots showing pairwise identity of nanopore reads collected on the ISS (left panel) and on the ground (right panel) relative to lambda (left box), E. coli (middle box) and mouse (right box) genomes. Abbreviations: 2D, high-quality two-dimensional reads; template, 1D template read; complement, complement read.

(D) Pie charts of the read distributions corresponding to each ISS run and pooled ISS runs 1 – 4, in comparison to that obtained from a ground Illumina MiSeq run of the same sample mixture.
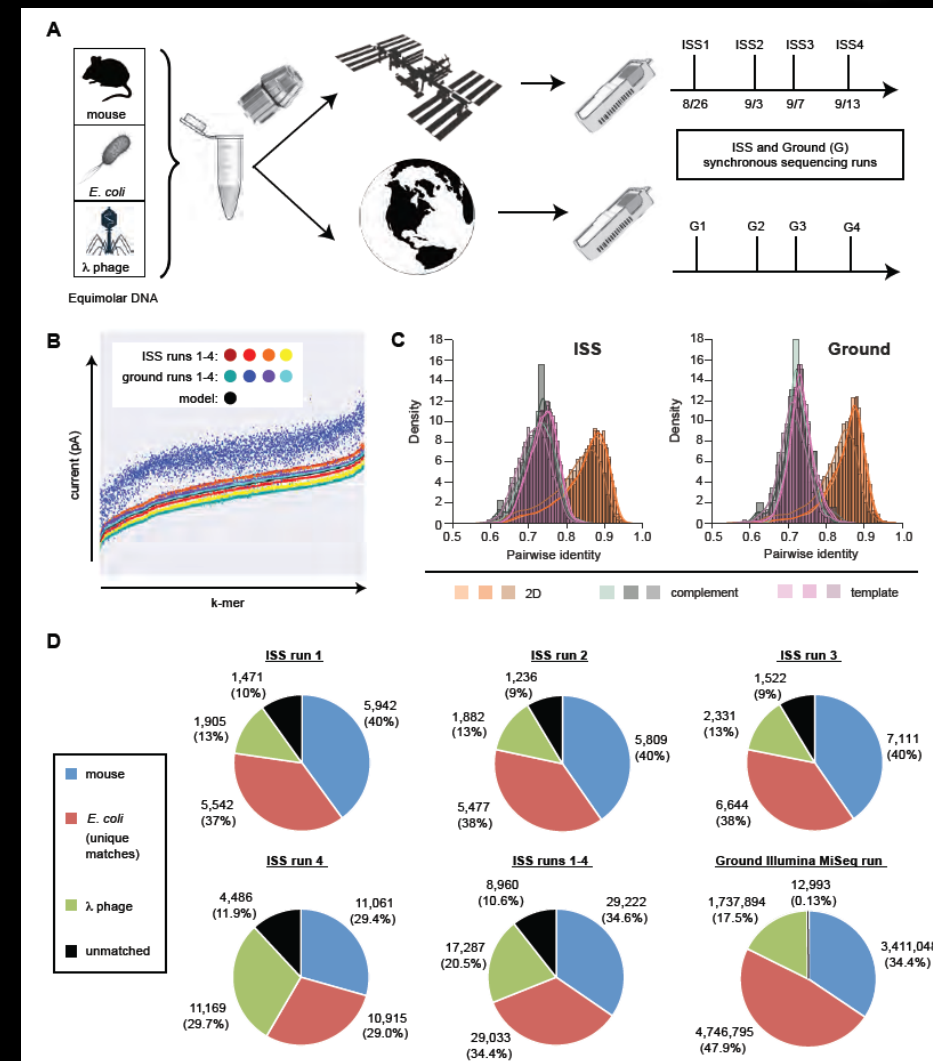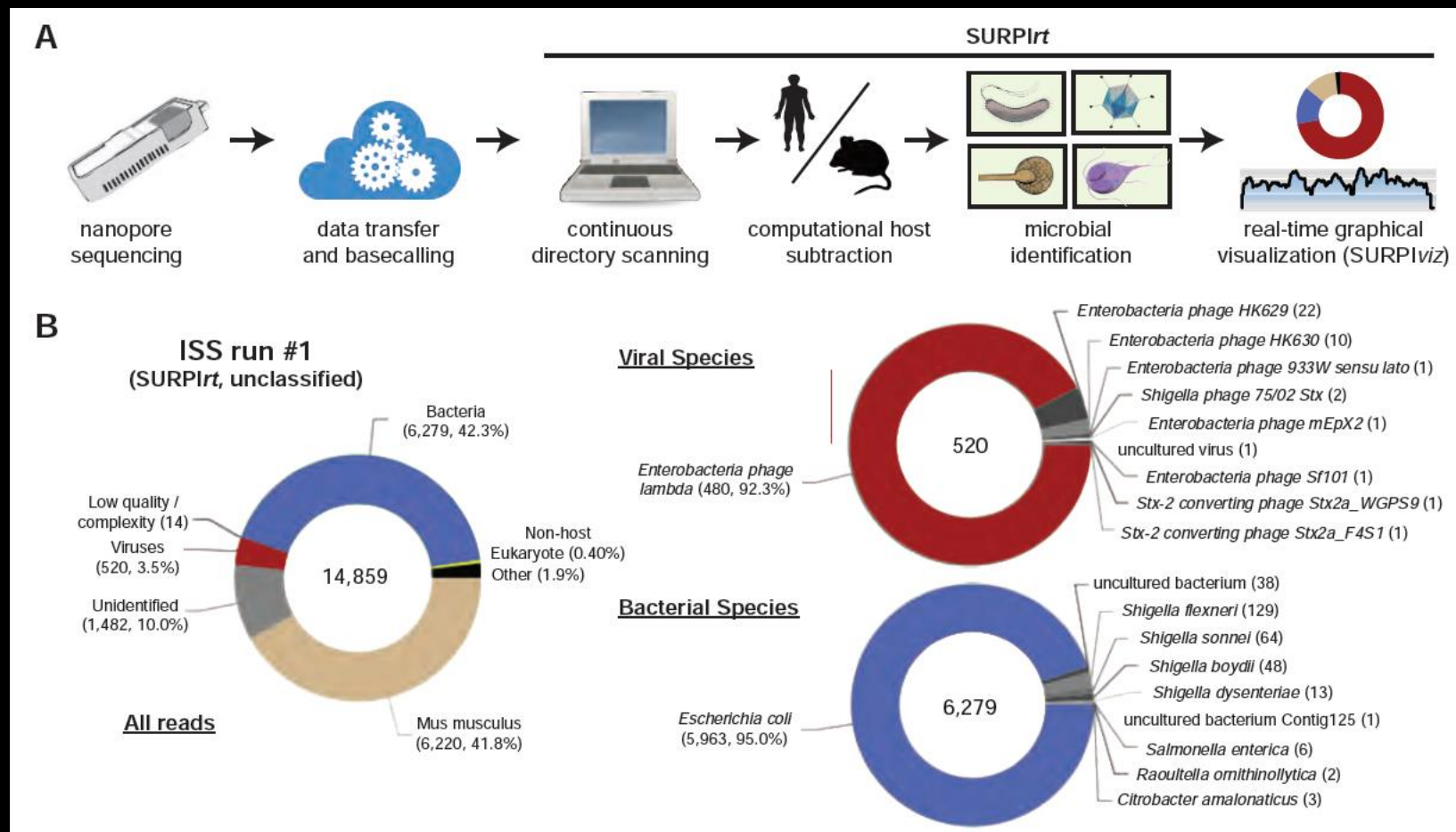
Biomolecule Sequencer: Automated Analysis Pipeline

(A) Flow chart of the SURPI*rti* bioinformatics pipeline for real-time microbial detection from nanopore data.

(B) Donut charts of read distributions corresponding to all reads (left), viruses, (upper right), and bacteria (upper right) from ISS run 1. These charts were generated dynamically as part of a real-time sequencing analysis simulation using SURPI*rt*.
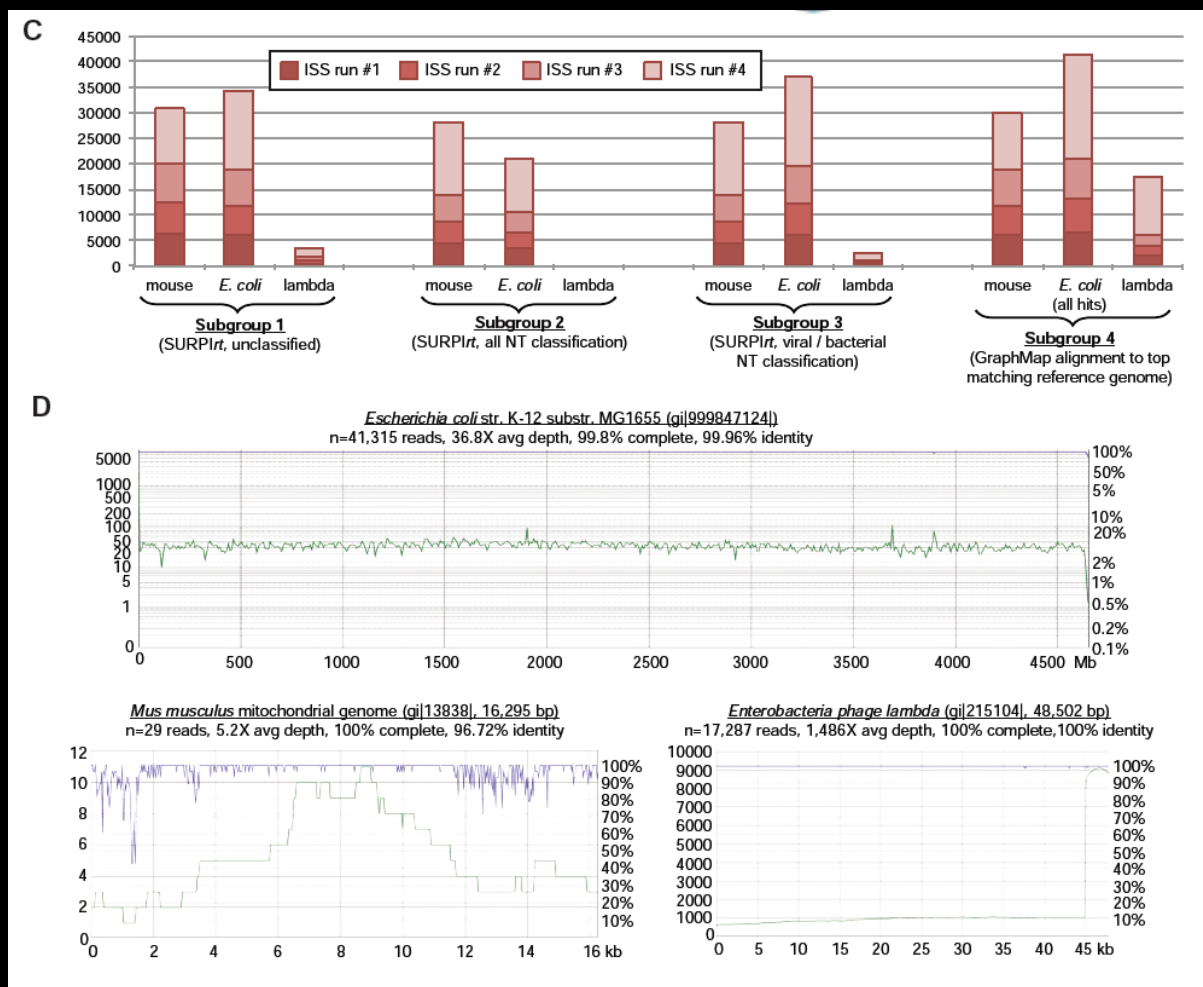
Biomolecule Sequencer:
Automated Analysis Pipeline

(C) Stacked distributions of reads from ISS runs 1 through 4 aligning to mouse, *E. coli*, and lambda. Subgroup 1 shows raw SURPI*rt* output in the absence of taxonomic classification, while subgroups 2 and 3 show the effects of classification using the GenBank NT database and separate viral or bacterial databases, respectively. The relative proportions of read counts from SURPI*rt* differ from those obtained by GraphMap alignment to the most closely matched reference genome in NCBI NT (subgroup 4).

(D) Coverage (green) and pairwise identity plots (purple) of raw nanopore reads mapped the *E. coli* (upper panel), the mouse mitochondrial (lower left panel), and lambda genomes (lower left panel). Reads are mapped to the most closely matched reference genome identified by SURPI*rt*.
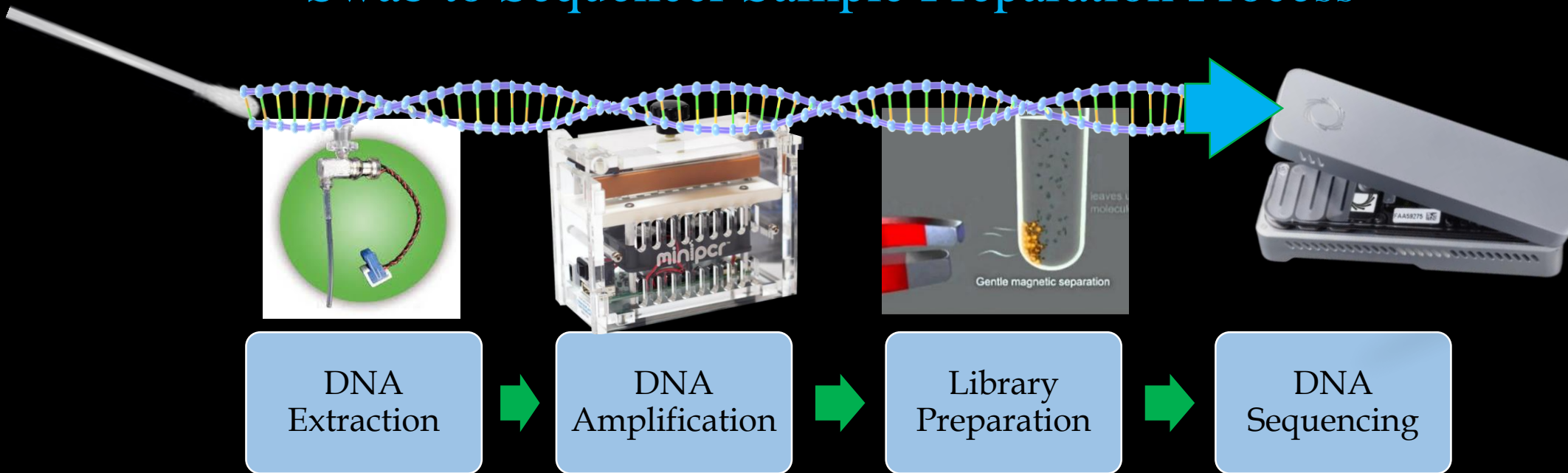
Looking Ahead: Need for Sample Preparation

Swab to Sequencer Sample Preparation Process

Gentle magnetic separation

| DNA Extraction | → | DNA Amplification | → | Library Preparation | → | DNA Sequencing |

Biomolecule Sequencer: The Future

- Maintain the Biomolecule Sequencer on the ISS as a permanent operational and research facility.
- Continue initiatives to develop the capabilities to perform the sample collection and preparation on orbit, allowing an endless number of potential experiments.
- Nanopore sequencing can go far beyond DNA and can enable methylation, epigenetics, RNA, modified bases, and protein studies.

Biomolecule Sequencer: The Team

The Payload Development Team

- Aaron Burton, Ph.D. (PI)
  NASA JSC
- Sarah Castro-Wallace, Ph.D. (PM)
  NASA JSC
- Kristen John, Ph.D. (Deputy PM and PE)
  NASA JSC
- Sarah Stahl, M.S., (PS)
  NASA JSC

Astronaut

- Kate Rubins, Ph.D.
  NASA JSC

The Science Team

- Jason Dworkin, Ph.D.
  NASA GSFC
- Mark Lupisella, Ph.D.
  NASA GSFC
- David Smith, Ph.D.
  NASA ARC
- Tim Stephenson, Ph.D.
  NASA GSFC
- Doug Botkin, Ph.D.
  Independent consultant

The Science Team

- Charles Chiu, Ph.D.
  - Scot Federman
  - Sneha Somasekar
  - Doug Stryke
  - Guixia Yu
  UCSF
- Chris Mason, Ph.D.
  - Noah Alexander
  - Alexa McIntyre
  WCMC

Oxford Nanopore Technologies

- James Brayer
- Sissel Juul, Ph.D.
- Daniel Turner, Ph.D.
- Michael Micorescu, Ph.D.

Special thanks to Marc Reagan for testing during NEEMO 21 engineering week and Drs. Lindsay Rizzardi and Andy Feinberg for parabolic flight testing!

Biomolecule Sequencer: Questions?

Thank you!

David.J.Smith-3@nasa.gov