MESSE : Molecular Encoded Storage for Space Exploration

Chayapatr Archiwaranguprok
chayapatr@spaceth.co
Spaceth.co
Bangkok, Thailand

Prompt Suathim
prompt.sua@mail.kmutt.ac.th
FREAK lab
Bangkok, Thailand

Saran Seehanam
saran.hch@gmail.com
FREAK lab
Bangkok, Thailand

Phoomparin Mano
poom@hey.com
BRIKL BV
Bangkok, Thailand

Chottiwatt Jittprasong
chottiwatt@spaceth.co
Spaceth.co
Bangkok, Thailand

Chanud Sitthipredanant
chanud@spaceth.co
Spaceth.co
Bangkok, Thailand

Watcharin Unwet
watcharin.unw@gmail.com
Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University
Bangkok, Thailand

Puey Ounjai
puey.oun@mahidol.edu
Department of Biology, Faculty of Science, Mahidol University
Bangkok, Thailand

Pakpoom Subsoontorn
pakpoomton@gmail.com
Department of Biochemistry, Faculty of Medical Science, Naresuan University
Phitsanulok, Thailand

Nattanon Dungsunenarn
nattanon@spaceth.co
Spaceth.co
Bangkok, Thailand

Pat Pataranutaporn
patpat@media.mit.edu
MIT Media Lab
Cambridge, Massachusetts, USA

ABSTRACT

Molecular Encoded Storage for Space Exploration (MESSE) is an anti-disciplinary space mission that demonstrates the future of bio-digital information system for interplanetary communication. The team developed a proof-of-concept data encoding mechanism, assembly method, and deployment strategies for DNA storage as a digital message carrier for the sub-orbital space flight. The team encoded musical composition with poetic meanings in the DNA storage to illuminate the dialogues and tensions between art and science at the crossroad of existential philosophy, synthetic-biology, and space exploration. The payload was integrated and flew to space on Blue Origin NS-13 sub-orbital flight. The experiment can be divided into 2 sets. (I) Study the fundamental process of DNA storage synthesis and self-assembly. We synthesized three fragments of DNA, which will later be assembled into a full storage device using Gibson assembly technique. Sending complete assembled and fragmented DNA into space helps us understand spaceflight's impact on a synthesized DNA by investigating the DNA damage and decoding the data inside. (II) To hide our DNA in unprecedented objects secretly within cloth, mylar film, golden leaf, and a piece of 3D printing filament. If the DNA's message survives and can remain readable when back down to earth, it may be tolerable to make a DNA heist on an existing space mission.

KEYWORDS

DNA Storage, Biosecurity, Biodigital, Synthetic Biology, Gibson Assembly, Space, Optimization, Algorithm

Figure 1: The design of the payload with artwork on the top.

1 INTRODUCTION

> "the biggest innovations of the 21st century will be at the intersection of biology and technology"
> – Steve Jobs

With the advancement in synthetic biology, humans have developed engineering methods for writing and reading information at the molecular scale. Scientists and innovators have studies the possibility of using DNA as the future generation of data storage [2]. The concept of post-digital data storage is to use the molecular structure of genetic material to represent states of data similar to the use of binary in digital data storage. In the present, many research groups have successfully encoded various kinds of data e.g.,

Figure 2: Image captured inside the payload aboard the Blue Origin's rocket on a sub-orbital flight. The shadow of space dinosaur can visibly seen in the background floating in zero-gravity.

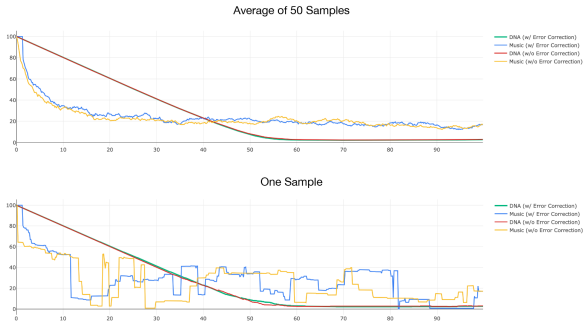

Figure 3: The similarity between original data and decoded data (x-axis: 0%-100%) (compared using Levenshtein Distance) after random mutation (y-axis: 0%-100%) (compared using Smith-Waterman Sequence Aligner) in one sample and the average of 50 samples. The data from the strand with Reed Solomon in the first 1.25% of mutation is completely restorable as expected as we set a small fragment of RS parity to correct eight words.

text, number or bitmap (image) into DNA form. [6] [4], [5] These studies involved developing the algorithm and techniques to prevent data from genetic mutation and molecular level damage.

Space is an unarguable future for human exploration as it has become much more accessible than before. Many modern scientists, e.g., Carl Sagan or Stephen Hawking have foresight that for the future of humanity, we have to prevent the loss of our civilization by preserving our knowledge in a recoverable format. We considered the genetic material to be suitable for this application, as life always fight to survive and reproduce.

2 METHODOLOGY

We stored the song (The dream and the universe) by a famous Thai rock band Bodyslam. Using procedures described above, successfully encoded the musical notation of the song within 640 Base Pairs of a DNA strand.

The final algorithmic output was cut into three smaller DNA strands with the addition of an overlapping parts to consequently conduct Gibson Assembly.

2.1 Encoding Scheme

Similar to character encoding schemes in computers, six bits are used to define one musical notation element. It is enough to store four-octave notes and essential musical components, which are enough for a proof of concept experiment and can easily be referred to as and converted into three nitrogenous bases. The more extensive n-bits size notation can be discussed in the future. The source code is available on Github. [1]

2.1.1 Text-based Musical Notation Scheme. The notation is a concatenation of musical notes defined using two fundamental music note conditions, i.e., pitch and duration, in the form of string. The string is then compressed using the tendency of repeat patterns in a musical score which gives the capability to define a symbol used in the compressed result instead of long repeats chunks. Subsequently, The musical notation string is mapped to a binary string.

2.1.2 Error Correction Algorithm. By conditions and limitations of DNA Storage methods, there is a probability of mutations in the strand. Therefore, the usage of the error correction scheme is advised by some recent papers. As the Reed-Solomon algorithm was utilized in recent protocol[5]. We also employ this correction scheme using cho45's Javascript implementation of the ZXing Reed-Solomon library.

2.1.3 Randomization. To strengthen base pairing, we averaged the content of four types of nitrogenous bases by XOR the binaries content with another pre-created binary string. The Pseudorandom Number Generator (PRNG) is used to create a binary string; we select the Xorshift to do the procedure.

As a tribute to Carl Sagan, in which many of his projects and writings influence our work, we've used 9111934 and 20121996, the dates of his birth and death, as a seed for a randomization algorithm.

The distribution of 01 content of the seed generated is tested before the final product usage, and the result is at an acceptable rate.

2.1.4 Conversion to Nitrogenous Bases. Two bits of string are converted to one base. The binaries are mapped into nitrogenous bases.

2.2 Gibson Assembly of Space Exposed DNA

We studied the fundamental process of DNA formation: self-assembly. The method, Gibson Assembly, allows for assemble of DNA fragments or genomes[3]. We synthesized three fragments of DNA, which can be Gibson Assembled. By sending complete assembled and fragmented DNA into space, we can understand spaceflight's impact on synthesized DNA. The researchers will inspect the DNA and look for any damage and decode the data inside.

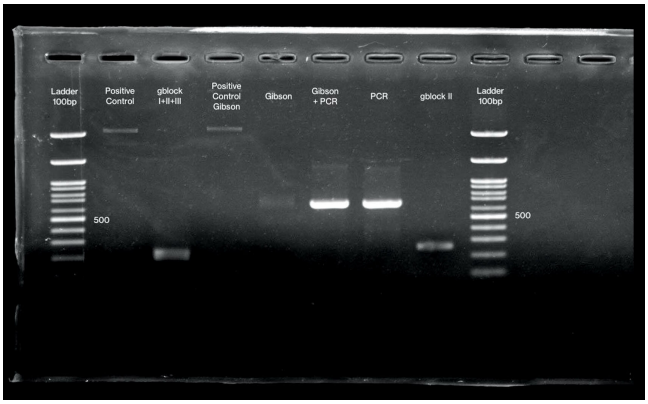

Figure 4: Result of the first Gel Electrophoresis. The total of an around 640 base-pairs match our designated DNA sequence resulted from the algorithm. The assembled products were PCR amplified, and compared with kit(NEB E2621) positive control and pre-assembled fragments

2.3 Counter Bioterriorism Experiment

The second experiment of the mission is to hide our DNA in unprecedented objects secretly within cloth, mylar film, golden leaf, and a piece of 3D printing filament. It was shown[7] in that DNA could be transported onto various parts of a rocket, and with some still retaining its function and integrity. If the DNA's message survives and can remain readable when back down to earth, it may be tolerable to make a DNA heist on an existing space mission.

2.4 Payload Design

The experiment fitted in a 10x10x10 (1U) payload container. We 3D-printed the entire uni-body frame and each modular segment. 42 PCR tubes containing our DNA are situated in a double anti-leaked configuration. The design went trough Power Spectrum Density (PSD) random vibration test.

A small 3D-printed dinosaur was included to indicate zero-gravity during the first stage separation and Apogee. We managed to record the in-flight footage by a small ultra-wide-angle camera inside.

2.5 A cultural perspective

In nature, there is no separation between art and science; analytics and aesthetics; or expression and exploration. They co-exist in the symbiotic relationship that foster human creativity and the search for meaning in the universe. What DNA and Digital Storage has in common is their ability to embody the abstract form of information in a physical medium. They are very similar in functionality, but different in scale and computability. One is human made and the other makes humans. It is the role of innovators and thinkers to create hope, show humanity new possibilities, make the impossible possible, and bring light to the darkness. Not only do our experiment involve the popular Thai song; our payload design will symbolically represent the enthusiasm, dream, and hope for our nation's position in the era of space exploration. Leonardo Da Vinci's Vitruvian man inspired the art cover of the payload. The artists and

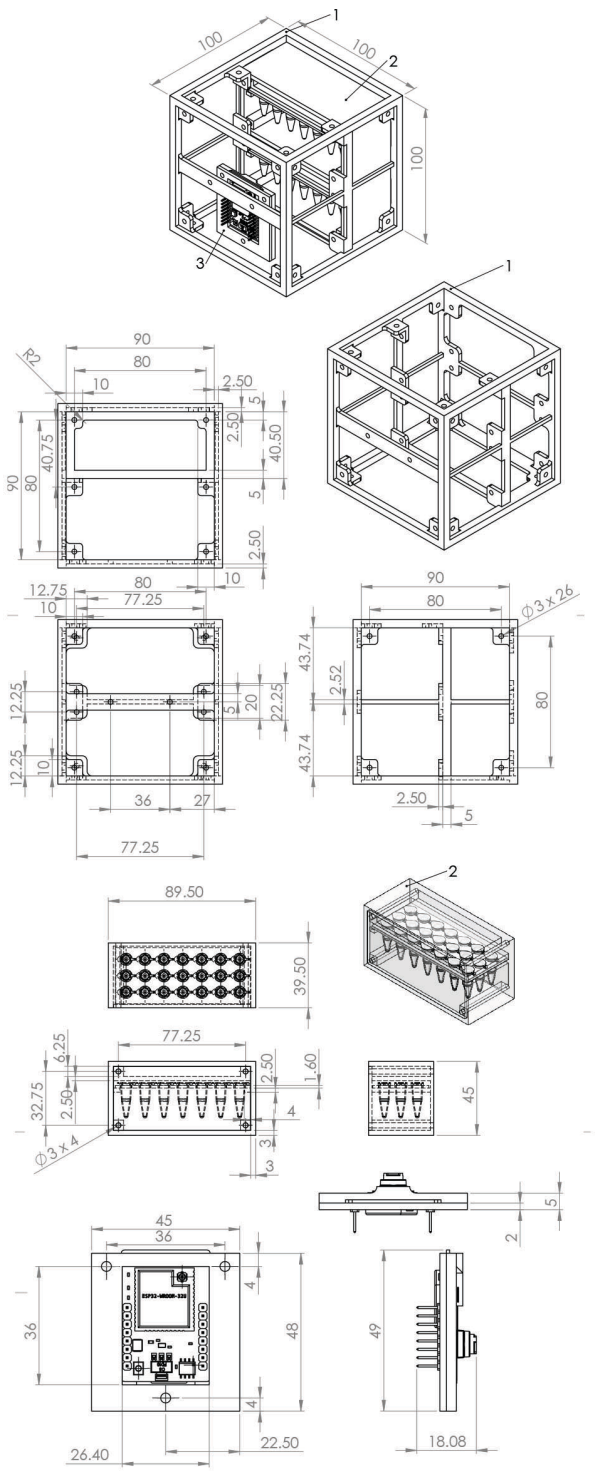

Figure 5: Main structure and sub-compartment of the payload. The structure was designed in modular compartment make it become reusable.

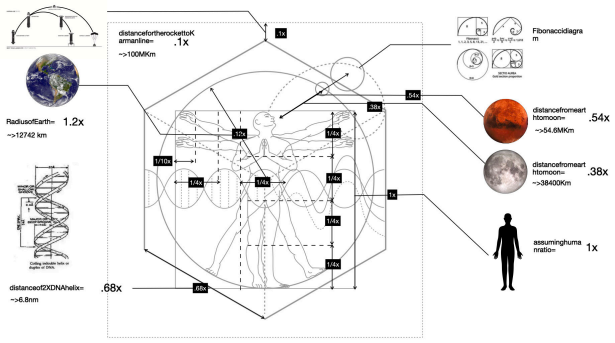

Figure 6: Overall mission explanation graphic.

researchers together came up with an idea of the symbol representing mankind's analytical ability, humility, and at the same time, pushing the boundary of being human. It is intended to illustrate the verse of the song encoded in our DNA.

3 CONCLUSION

After the experiment landed back to earth and has been shipped back to Thailand in November 2020, we are studying the result of the experiment by benchmark the result from controlled sets of DNA as well as recovering DNA from Counter Bioterriorism Experiment.

The result will address and discuss the recovery rate of our algorithm, base pair damage caused by various factors, and chances of DNA storage surviving in harsh scenarios (e.g., moon, Mars, or Interstellar travel) by comparing to the data heritage of Voyager 1 and 2 interstellar mission.

This mission aims to enhance engagement from all disciplines in Thai society on a unique mission. The collaboration involves artists, scientists, space companies, science communicators, and public engagement. This mission will mark a distinguished chapter for Thai space exploration.

In terms of global impact, this mission demonstrates an alternative way of storing music in DNA by implementing the foundation of using musical notation and turning them into base pairs of DNA.

4 ACKNOWLEDGMENTS

Mu Space Corp essentially supported this research. The payload has flown aboard its designated flight to help researchers and students get access to space. We appreciate their engineering team on a significant contribution to payload integration and mission design, including Supanat March Limjitti, Grin Rhothjirathanin, Borinot Hongdilokkul, Nunticha Sirinak, and Phurit Pookayaporn. The model of the dinosaur used in the mission was from the generosity of Suravech Suteethorn. The song was written and permitted for the research purpose by Artiwara Kongmalai, Poonsak Jaturaboon from Genie Record. The lab procedure was supported by Dr.Subsoontorn's colleagues at the Naresuan University, Orapan Meesungnoen, Panjaporn Wongpayak, and Dr.Ounjai's Mahidol University colleagues, Tanatchakorn Asawasriworanan, Ponlawoot Raksat. We are also immensely appreciative to Natthaporn Kobkulkalyakorn, faculty of ICT, Mahidol University, who convert and rearrange the music before the algorithmic process. Patcharapol Sankaew helps us to select the suitable logic board for the payload. Henry Tan, Jitsupa Chin, Nottapon Boonprakob, Yossawat Sittiwong helps us connect to the public via their owned media channels. The tremendous support from Space Zab company via Wares Chanchoren and Potiwat Ngamkajornwiwat. The concept of art and science intersection was inspired by Werasak Surareungchai, Freak Lab, King Mongkut's University of Technology Thonburi, and Pataranutaporn's colleagues at the MIT Media Lab. And lastly, the ultimate inspiration from Carl Sagan.

REFERENCES

[1] Chayapatr Archiwaranguprok. 2020. Golden Record. https://github.com/spaceth/goldenrecord.

[2] George M. Church, Yuan Gao, and Sriram Kosuri. 2012. Next-Generation Digital Information Storage in DNA. *Science* 337, 6102 (2012), 1628–1628. https://doi.org/10.1126/science.1226355 arXiv:https://science.sciencemag.org/content/337/6102/1628.full.pdf

[3] Lei Young Ray-Yuan Chuang J. Craig Venter Clyde A. Hutchison Gibson, Daniel G. and Hamilton O. Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature* 6, 5 (12 Apr 2009), 343–345. https://doi.org/10.1038/nmeth.1318

[4] Nick Goldman, Paul Bertone, Siyuan Chen, Christophe Dessimoz, Emily M. LeProust, Botond Sipos, and Ewan Birney. 2013. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* 494, 7435 (01 Feb 2013), 77–80. https://doi.org/10.1038/nature11875

[5] Linda C. Meiser, Philipp L. Antkowiak, Julian Koch, Weida D. Chen, A. Xavier Kohll, Wendelin J. Stark, Reinhard Heckel, and Robert N. Grass. 2020. Reading and writing digital data in DNA. *Nature Protocols* 15, 1 (01 Jan 2020), 86–101. https://doi.org/10.1038/s41596-019-0244-5

[6] Seth L. Shipman, Jeff Nivala, Jeffrey D. Macklis, and George M. Church. 2017. CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature* 547, 7663 (01 Jul 2017), 345–349. https://doi.org/10.1038/nature23017

[7] Svantje Tauber Andreas Schütte Burkhard Schmitz Harald Nuesse Ralf Moeller Thiel, Cora S. and Oliver Ullrich. 2014. Functional activity of plasmid DNA after entry into the atmosphere of earth investigated by a new biomarker stability assay for ballistic spaceflight experiments. *PLoS* 9, 11 (26 Nov 2014), 1–24. https://doi.org/10.1371/journal.pone.0112979