Backing visionary entrepreneurs

The European Innovation Council
EIC Pathfinder Challenges
Applicants day – 05/07/2022

DNA digital storage

European Innovation Council

Send us your questions by Sli.Do!

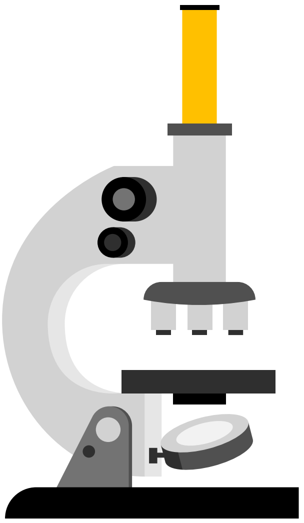
Join at
slido.com
#EICPathfinder

- General Q&A
- Carbon dioxide and nitrogen...
- Cardiogenomics
- DNA-based digital data stor...
- Energy storage
- Towards the Healthcare Con...
- Quantum Information

Join Sli.do #EICPathfinder and select the correct Breakout session

Outline

- Work Programme
  - Overall goal of the Challenge
  - Scope of the Challenge
  - Specific Objectives
  - Expected Outcomes
  - Expected Impacts
  - Eligibility conditions for applicants
- Challenge guide
  - Background
  - Goal of the portfolio
  - Categories
  - Portfolio considerations



Work Programme
Overall goal of the Challenge

European Innovation Council

This EIC Pathfinder Challenge is to explore scalable and reliable high-throughput approaches for using DNA as a general data-storage medium.

Work Programme
Scope of the Challenge

- Solutions would need to address the read/write/edit operations of digital data in synthetic DNA, capturing the expected advantages of high density and stability/longevity of this form of data storage.
- The use of DNA sequences as chassis for non-standard forms of information coding, or of other polymeric substrates and related coding/decoding techniques are also in scope, provided they entail at least similar benefits than state-of-the-art DNA approaches.
- Proposed techniques should deliver qualitative advances in key parameters such as throughput, DNA-length, reliability (coupling efficiency), speed and cost.
- Beyond the usual storage applications, there is also scope for radically different scenarios for such a technology, for instance for data-processing, in-vivo sensing or fingerprinting.

Work Programme
Specific Objectives

European Innovation Council

- New approaches for coding, decoding, modification or computational use of digital data in synthetic DNA or other sequence-controllable polymers with quantitative targets (theoretical and technological);
- Proof-of-Concept of technical feasibility with indications of at least state of the art benefits and major operational characteristics (e.g., extreme densities, longevity, stability) and going well beyond for some of them (e.g., speed, cost, accuracy);
- End-to-end scenarios of use, be it for data storage (archival, but also shorter-term storage) or other purposes (like sensing, cryptography or computation) that exploit the benefits of the technology.

Work Programme Expected Outcomes

- A range of new techniques with clear benefits and steps towards widening scope of applicability of DNA-based data storage;
- Broader range of scenarios and uses for DNA-based data technologies;

Work Programme Expected Impacts

European Innovation Council

- Emergence and anchoring of a European innovation eco-system on DNA-based data technologies and applications, including through involvement of relevant partners and end-users;
- Contribution to standardisation in the field and benchmarks to gauge progress.

Work Programme
Eligibility conditions for applicants

European Innovation Council

- Proposals for this Challenge can be submitted by single applicants or by consortia*, as dictated by the activities to be performed.

* In case of a consortium your proposal must be submitted by the coordinator on behalf of the consortium that includes at least two independent legal entities. The legal entities may for example be universities, research organisations, SMEs, start-ups, natural persons. In the case of single beneficiary projects, mid-caps and larger companies will not be permitted.

Challenge guide Background (1)

European Innovation Council

As the world becomes increasingly digitized, the number of data sources is rising dramatically. In addition to “standard” online activities, the Internet of Things, smart cities, personal devices, autonomous vehicles, and environmental sensors all contribute to a rapid rise in the volume of data generation.

A significant proportion of data that is generated is not yet stored beyond the short-term, but some longer-term projections forecast that overall demand for digital storage will exceed supply by up to three orders of magnitude by 2040.

Challenge guide
Background (2)

- Background challenges
  - Scalability - data scalability factors do not simply refer to storage capacity; they also refer to the number of different datasets to be handled or related to one another, the frequency with which datasets are updated and/or queried, the complexity of updates and/or queries, and the physical location, distribution, and replication of the stored datasets.
  - Sustainability - It is crucial to reduce the energetic requirements imposed by data storage. Existing storage methods require the mining of large amounts of gold, copper, and aluminium, as well as “rare earth” materials that are difficult to process and which generate significant toxic pollution.
  - Integrity and reliability - Existing media can offer data retention periods that last on the order of decades, but this comes at a cost in terms of proper environmental control, periodic data integrity checks, and other physical overheads.

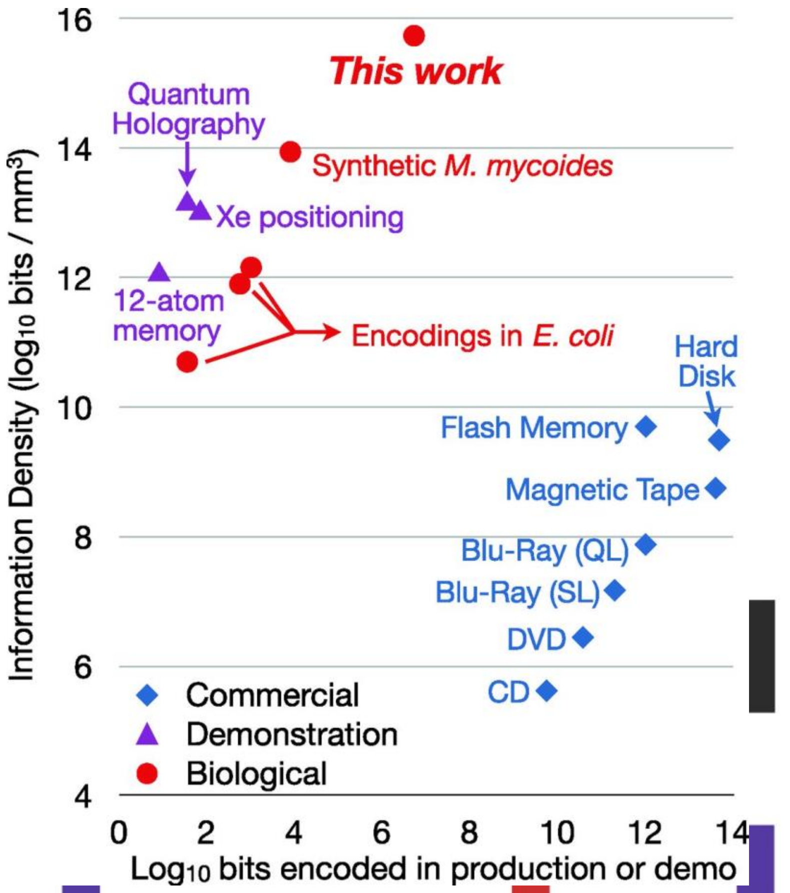
Challenge guide Background (3)

European Innovation Council

DNA as an alternative storage medium

- Storage density - The raw information density of DNA has been calculated as more than one exabyte (1000 petabytes) per cubic millimetre. Although traditional storage devices are moving beyond planar layouts to 3D structures, the inherent volumetric storage of DNA may also confer significant potential benefits.
- Robustness - DNA is remarkably stable, especially when stored in carefully controlled conditions (maintaining pH, temperature, humidity, etc.)

Church, G.M., Gao, Y. and Kosuri, S., 2012. Next-generation digital information storage in DNA. *Science*, 337(6102), pp. 1628-1628



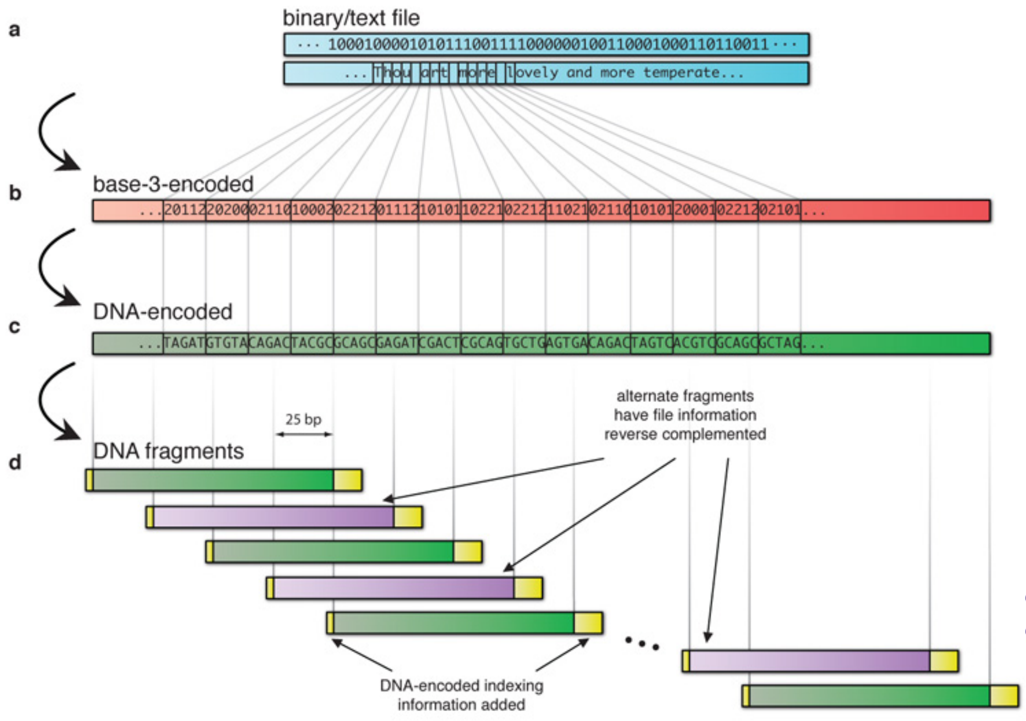
Challenge guide Background (4)

European Innovation Council

DNA as an alternative storage medium

- Universality - The genetic code (the mapping between nucleotide sequences and amino acids that make up proteins) is shared by all living organisms, meaning that it is one of the best-studied encoding schemes in existence.
- Resource requirements - Once data is properly encoded in DNA, it requires very little resource for its long-term storage.

Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E.M., Sipos, B. and Birney, E., 2013. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 494(7435), pp.77-80.



Challenge guide Background (5)

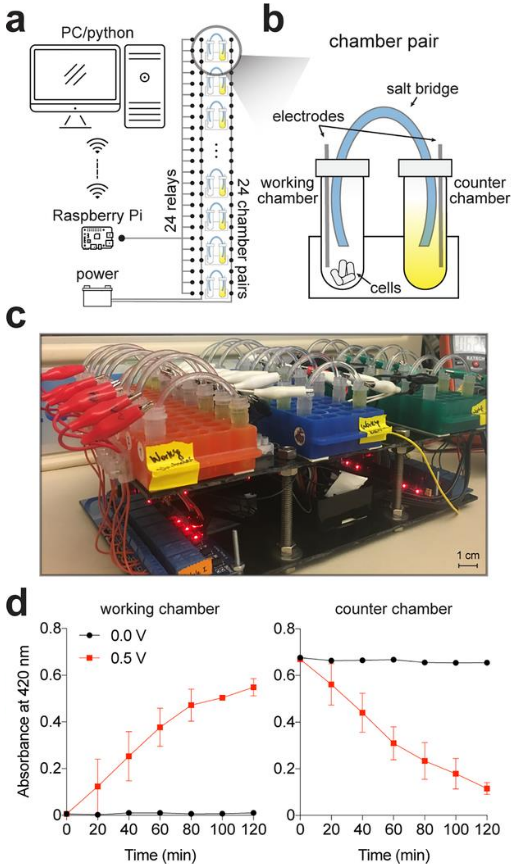
Dynamic DNA storage

- Although much recent attention has focussed on the use of DNA as a long-term storage medium, we should not overlook its potential as a dynamic medium.

Living cells as possible storage media

- The natural environment of DNA is the living cell, and evolution has produced a diverse toolbox of biological processes for copying, proof-reading and even propagating DNA sequences

Yim, S.S., McBee, R.M., Song, A.M., Huang, Y., Sheth, R.U. and Wang, H.H., 2021. Robust direct digital-to-biological data storage in living cells. *Nature Chemical Biology*, 17(3), pp.246-253.



Challenge guide
Goal of the portfolio

The goal is to achieve a diverse portfolio in terms of time-frames present, underlying substrate, and potential applications, while encouraging proposals that demonstrate end-to-end integration and interoperability where applicable.

Challenge guide
Categories

European Innovation Council

- <u>Storage duration:</u> The committee will seek to facilitate a set of projects that operate across the spectrum of time, from long-term, archival, “cold” storage, to medium-term, “working” storage and short-term “dynamic” storage.
- <u>Substrate:</u> Whilst acknowledging that the bulk of the portfolio may be comprised of projects that operate on synthetic DNA in vitro, the committee will encourage the investigation of alternative substrates such as non-natural polymers or living cells.

Challenge guide
Portfolio consideration

1) If present, at least one proposal will be added to the portfolio dealing with each of short-, mid, and long-term storage, since each of these time-frames presents its own set of challenges (in terms of the stability of the molecules used, and the technologies required to manipulate them) and opportunities (in terms of possible applications).
2) Portfolio construction will encourage a degree of diversity in the underlying substrate (DNA, other synthetic polymers, living cells, etc.), whilst acknowledging the current emphasis on synthetic DNA.
3) The committee will encourage proposals that demonstrate end-to-end integration and interoperability, if appropriate to the category.
4) Within each category, the committee will seek, where possible, diversity in terms of potential applications (so, for example, within the “long-term” window, they may look for applications that require different levels of fidelity of data storage).

Send us your questions by Sli.Do!

- General Q&A
- Carbon dioxide and nitrogen...
- Cardiogenomics
- DNA-based digital data stor...
- Energy storage
- Towards the Healthcare Con...
- Quantum Information

Join at
slido.com
#EICPathfinder

Join Sli.do #EICPathfinder and select the correct Breakout session

Thank you!

Roumen.BORISSOV@ec.europa.eu