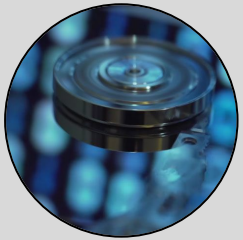
Storing any nucleic acid in the world.
Forever.

https://cache-dna.com

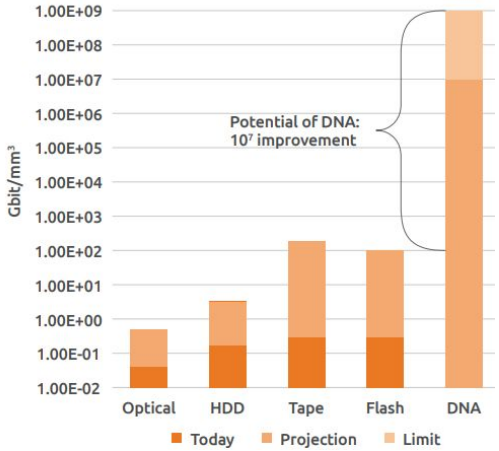
Data generation is exploding

Data generation is expected to reach 175 ZB by 2025.



- Feature sizes of conventional storage systems are close to their physical limit
- Optical & tape storage will not be able to address future demands

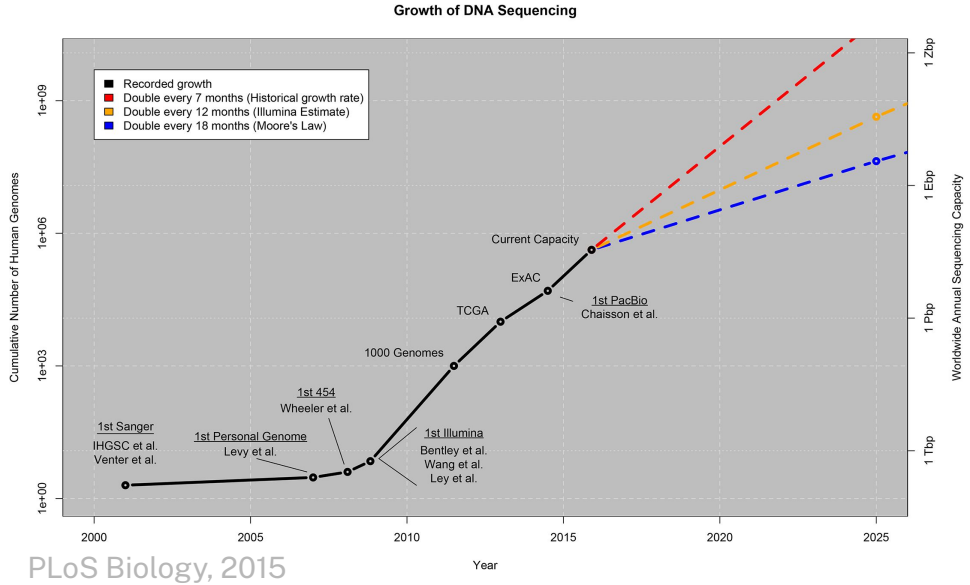
POTENTIAL FOR DNA DATA STORAGE



- Ultradense storage
- Easy copying
- Low energy for maintenance

Demand for storing nucleic acids is growing

Nucleic acids are the new information media of the 21st century.



PLoS Biology, 2015

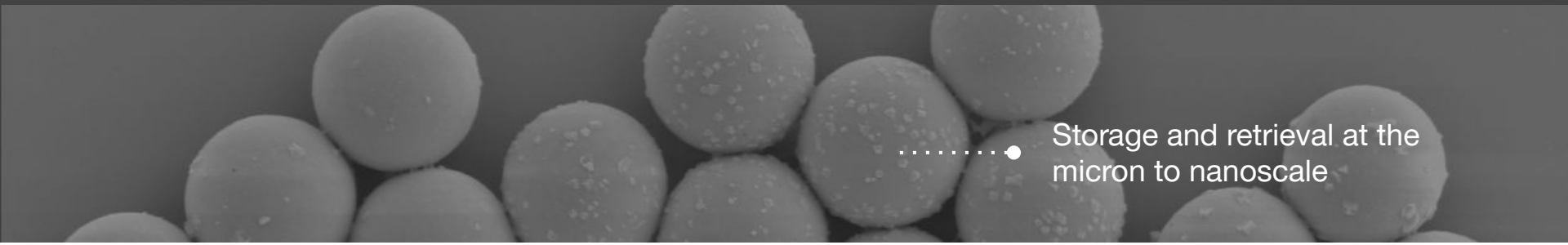
HEALTH

U.S. needs 193 million Covid-19 tests per month to reopen schools and keep up with pandemic, new report says

By KATE SHERIDAN @sheridan_kate / SEPTEMBER 9, 2020 Reprints

CACHE DNA

Revolutionizing Storage at the Smallest Scale



Storage and retrieval at the micron to nanoscale

18,000 t

CO2 reduction

$1.6 M

Annual cost savings

70%

Space reduction

From 2010 pilot study on the economics of room-temperature storage of nucleic acids by Stanford University's Sustainability and Energy Management

Cache DNA Has Broad Applications In the Information & Biological Storage Markets

BIOBANK STORAGE → BIOTHREAT SURVEILLANCE → ARCHIVAL DATA STORAGE → BIOLOGICAL COMPUTING

Go-to-Market

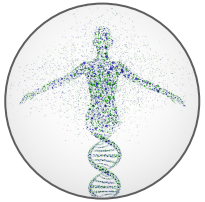
BIO STORAGE: BIOTHREAT SURVEILLANCE & BIOBANKING → ARCHIVAL DATA: DNA DATA STORAGE → NEW MARKETS: COMPUTING

NEAR-TERM STRATEGIC PRIORITIES

- Capture growing nucleic acid storage market

BIOBANKING

$0.5 B TAM



PERSONAL GENOMICS

$0.8 B TAM

- Offer new capabilities for biothreat surveillance ($1.0 B TAM)

LONG-TERM GOALS

- Near-term strategic priorities will provide:
  - Market demonstrations of scale for the Cache platform
  - Opportunities to build critical infrastructure & partnerships for DNA data storage ($8 B TAM)
  - Positive cash flow to augment product offering

- Leverage our experience in DNA data storage and massive biological storage to create new markets unique to the Cache platform, e.g. non-von Neumann architectures ($10 B TAM)

Market strategy and opportunity

3-5 year timeline

Pilot PoC
Reagent model

Partnerships
Reagent model

Partnerships
Device launch

Market-ready product

Beachhead market de-risks longer timeline market

5-10 year timeline

Partnerships
Operating system

Commercial PoC
Full stack data storage

Customer discovery

Validated value propositions

Cache DNA platform reduces cost to store samples more than half over 10 years and reduces footprint by 10^{10}-fold

Cache DNA platform requires almost no energy to store and access exabyte-scale data archives for a millenia

Select market research interviews

RUCDR INFINITE BIOLOGICS · MAYO CLINIC · 20th CENTURY STUDIOS · aws · Akamai · T

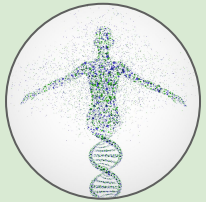
Cache DNA Platform Provides Immense Value to Multiple Growing Market Opportunities

NEAR-TERM MILESTONE Generate revenue

BIOBANKING

$0.5 B TAM

- Low-cost storage
- Scalable
- Small footprint
- Ambient



PERSONAL GENOMICS

$0.8 B TAM

BIOTHREAT SURVEILLANCE

$1.0 B TAM

- Long-term preservation
- Field-deployable
- Simple sample handling logistics

LONG-TERM MILESTONE Generate new markets

ARCHIVAL DATA STORAGE

$8.0 B TAM

- Ultradense storage
- Permanent recording
- Little to no energy required to maintain
- Easy to create copies

BIOCOMPUTING

$10.0 B TAM

- Massively parallelized computation
- Very low energy requirements for operation

Targeted Nucleic Acid Storage Value Chain

Sample collection

- *Collection*: Swabs, saliva collection tubes, serological sampling; *Processing*: QIAGEN & Beckman Coulter have huge market capture
- Highly commoditized

Sample storage

- Dominated by freezer & LIMS

ThermoFisher SCIENTIFIC · HAMILTON THE MEASURE OF EXCELLENCE · Brooks · QIAGEN

Sample retrieval

Cache DNA platform

- Compatible with molecular biology workflows
- Replace microtube with micron-to-nano size capsules (~10^{10} volume reduction)
- Ambient storage
- Scalable sample barcoding (~10^{24} samples can be labelled uniquely)

Sample analysis

- *DNA sequencing:* dominated by Illumina. Ion Torrent, PacBio, Oxford Nanopore are niche players.
- *Molecular Dx*: Roche, Abbott, and Thermo-Fisher are huge players. CRISPR-based analysis is emerging (e.g., Mammoth Biosciences, Sherlock Biosciences)

Report

- Clinical or application-driven

Use-Case Example: Persistent Biothreat Surveillance

Cache DNA can create longitudinal data sets to deliver actionable results for biothreat response & protect multiple stakeholders

Stakeholders

Schools
Hotels
Airlines
Nursing homes
Housing complexes
Insurance companies

CACHE DNA

LOW-COST SAMPLING

Human samples

Environmental surfaces

Wastewater samples

UNIQUE SAMPLE MANAGEMENT

Embedded rich metadata (e.g., spatiotemporal data) in sample storage

Detection/Diagnosis

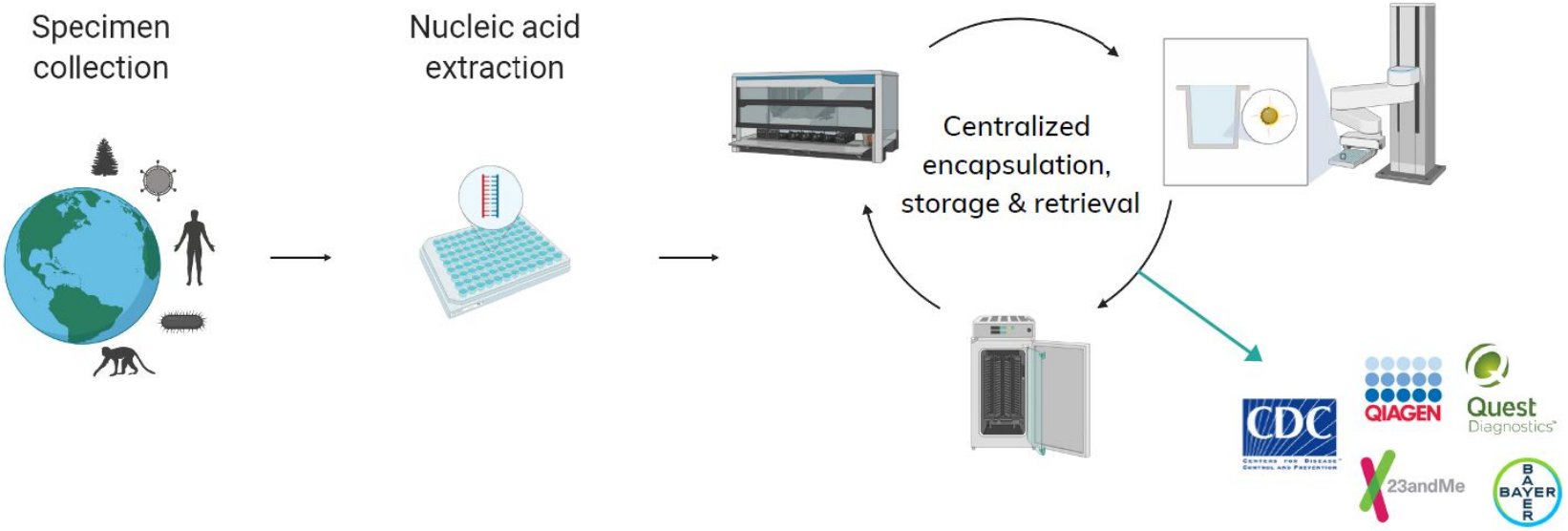
CDC
Reference labs
Sentinel labs

Biothreat Response

Cache DNA provides:

- Archival record of incidence
- Retrospective analyses of incidence

Massive Storage of Nucleic Acids From All Sources

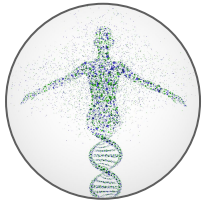


Cache DNA Aims to be the Eminent Data Storage & Computing Platform for the 21st Century

NEAR-TERM MILESTONE Generate revenue

BIOBANKING

$0.5 B TAM



PERSONAL GENOMICS

$0.8 B TAM

BIOTHREAT SURVEILLANCE

$1.0 B TAM

- Low-cost storage
- Scalable
- Small footprint
- Ambient

- Long-term preservation
- Field-deployable
- Simple sample handling logistics

LONG-TERM MILESTONE Generate new markets

ARCHIVAL DATA STORAGE

$8.0 B TAM

BIOCOMPUTING

$10.0 B TAM

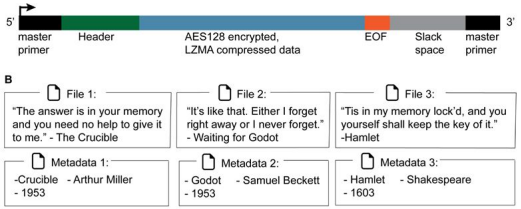
- Ultradense storage
- Permanent recording
- Little to no energy required to maintain
- Easy to create copies

- Massively parallelized computation
- Very low energy requirements for operation

Proof-of-Concept Full-Stack DNA Data Storage

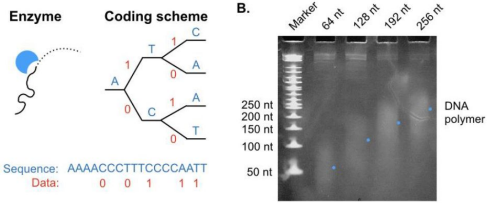
DNA WRITE

Biotech-scale production of plasmids



Shepherd, ..., Bathe. *Sci. Rep.* 2019

Parallelized enzymatic synthesis of DNA



MIT IP Bathe et al. US20180362969A1

RANDOM ACCESS & RETRIEVAL

Encapsulation-based random access

Banal, Shepherd, Berleant, ..., Bathe. *bioRxiv* 2020
MIT IP Bathe et al. WO2017189914A1

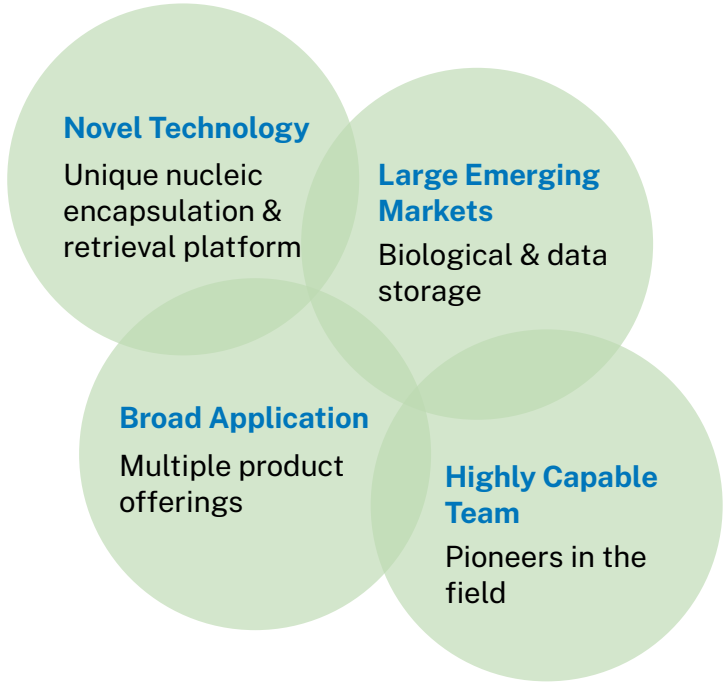
Cache DNA is developing revolutionary approaches to store, retrieve, and compute data in a scalable way and we're just getting started...

Team

	Mark Bathe, PhD Scientific Founder	Professor, Dept. of Biological Engineering Associate Member, Broad Institute
	James Banal, PhD Prospective Technical Founder	Postdoc, Dept. of Biological Engineering
	Paul Blainey, PhD Scientific Advisory Board Member	Professor, Dept. of Biological Engineering Core Member, Broad Institute
	George Church, PhD Scientific Advisory Board Member	Winthrop Professor of Genetics, Harvard Medical School Core Faculty, Wyss Institute
	Jeremiah Johnson, PhD Scientific Advisory Board Member	Professor, Dept. of Chemistry Member, Koch Institute at MIT Associate Member, Broad Institute

Why Cache DNA?

CACHE DNA



CACHE DNA

Appendix: Academic proof-of-concept on biobanking

Creating M  NSTER:

Biothreat surveillance for the genomic age

A hot take on cold storage

 $\sim 10^8 \times$  $365 = \sim 10^{10} =$  $= \sim 10^6$ 

Space:

Energy:

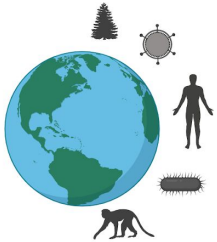
* No robotics included

MONSTER

Massive st**o**rage of **n**ucleic acid**s** for bio**t**hr**e**at su**r**veillance

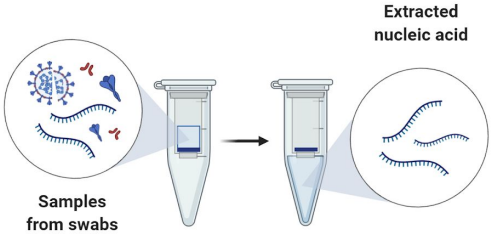
1 Sample collection

Sampling kits are used to collect samples from various sources around the world.



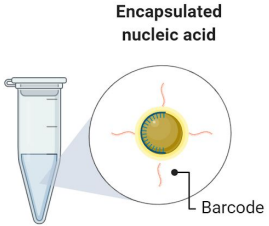
2 Nucleic acid extraction

Nucleic acid is extracted from sample using extraction kits.



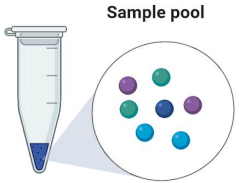
3 Nucleic acid encapsulation

Nucleic acid is encapsulated and barcoded using Cache DNA technology.



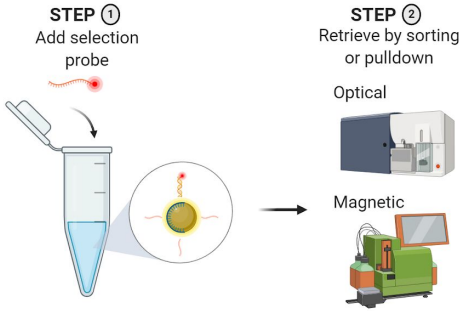
4 Nucleic acid storage

Encapsulated & barcoded samples can be pooled into one or several containers for transport or for archiving.



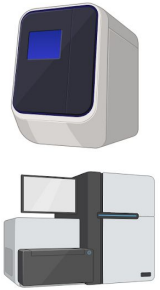
5 Nucleic acid retrieval

One or several encapsulated nucleic acids can be targeted for retrieval.



6 Nucleic acid analysis

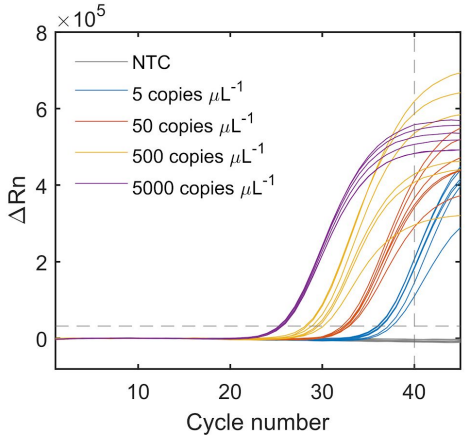
Nucleic acid is released from encapsulation and subjected to analysis via quantitive PCR or sequencing

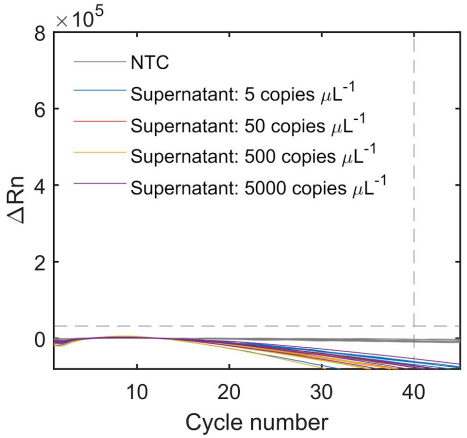


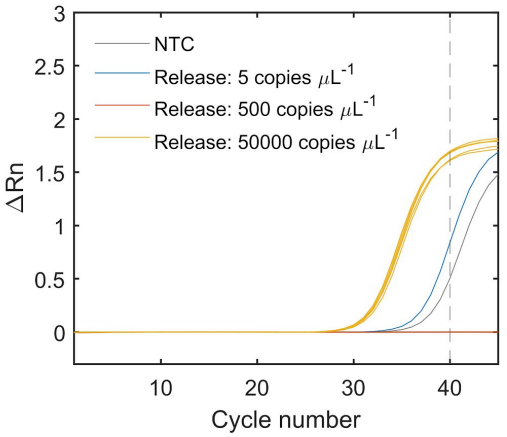
MONSTER

Massive st**o**rage of **n**ucleic acid**s** for bio**t**hr**e**at su**r**veillance

Quantitative encapsulation and reversible release of SARS-CoV-2 RNA:





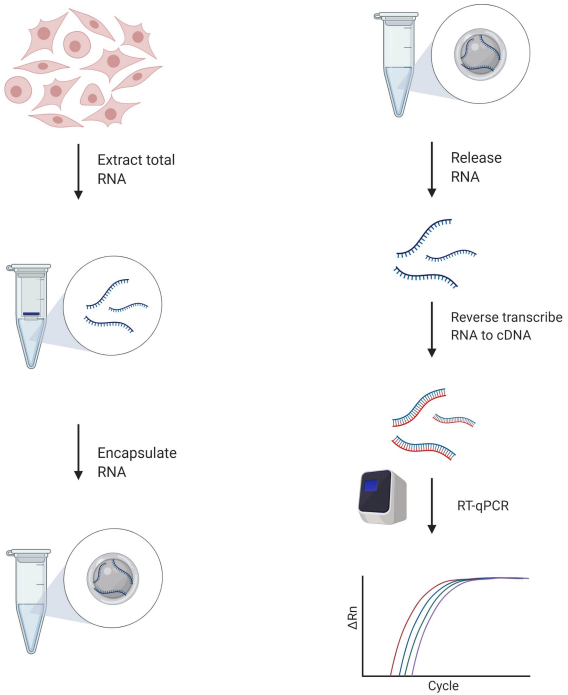


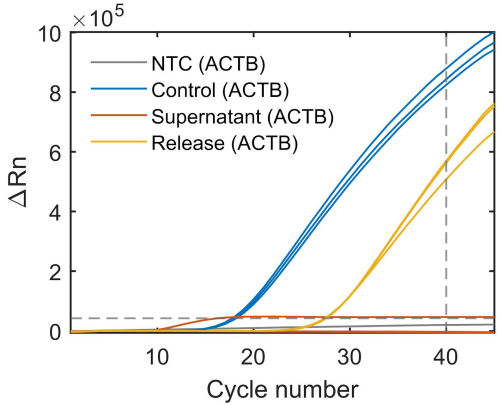
NTC = No template RNA control

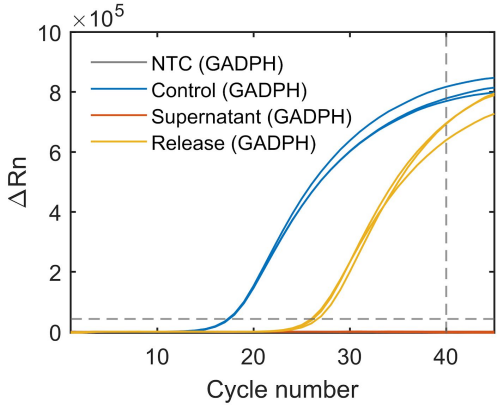
Biobanking proof-of-concept

RNA extraction & encapsulation

RNA release & RT-qPCR





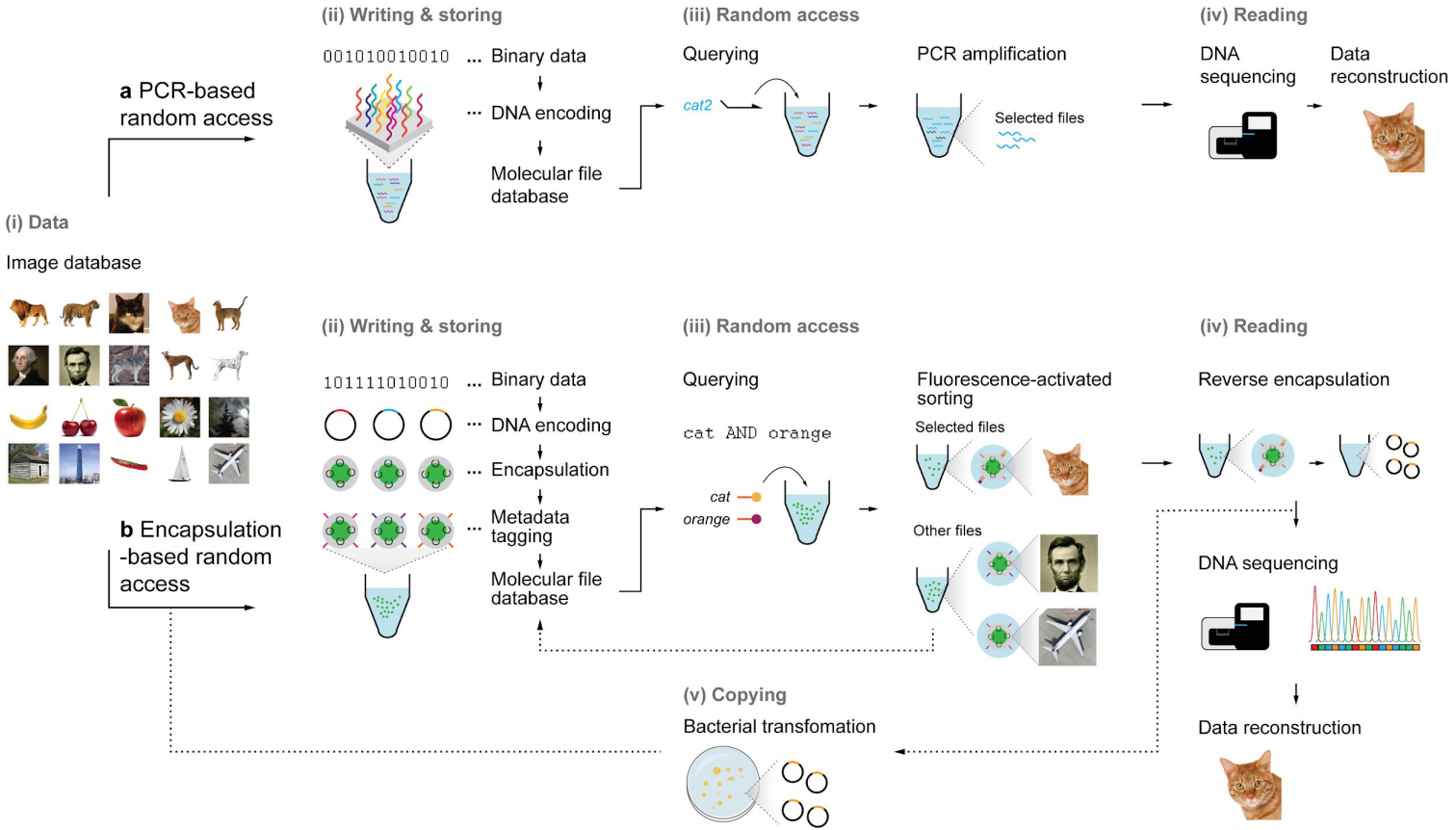


- No RNA detected on supernatant after encapsulation suggesting quantitative encapsulation of total RNA from HEK293 cells
- Cross validation of released RNA from encapsulation using two human housekeeping genes

NTC = No template RNA control

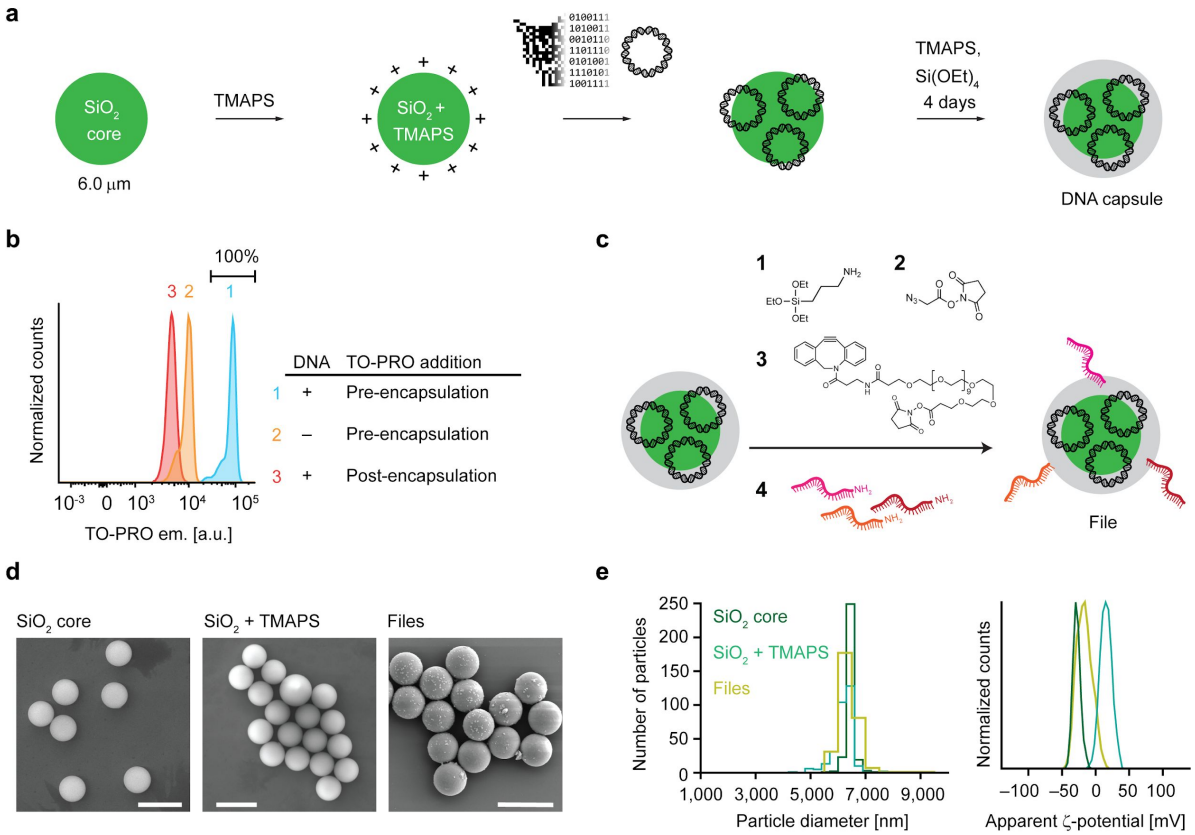
Appendix: Academic proof-of-concept on DNA data storage

DNA data storage
Reimagining computing using molecular packets

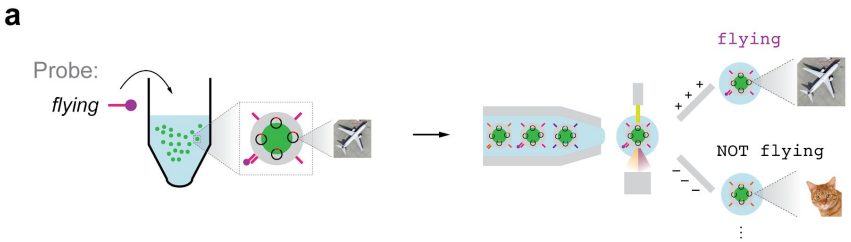


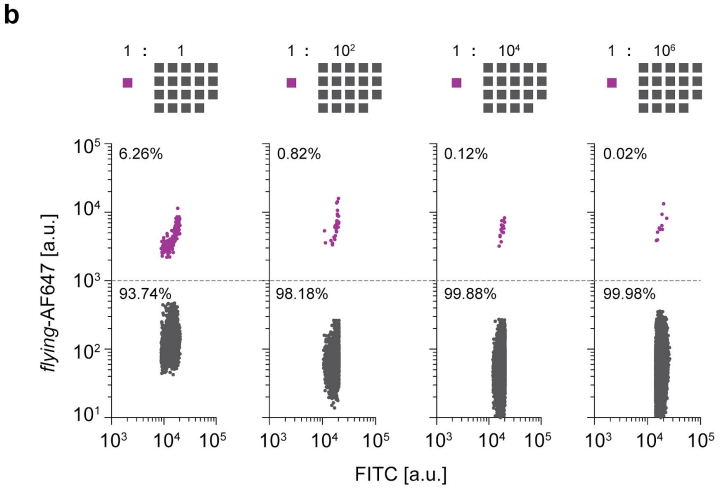
Unique implementation

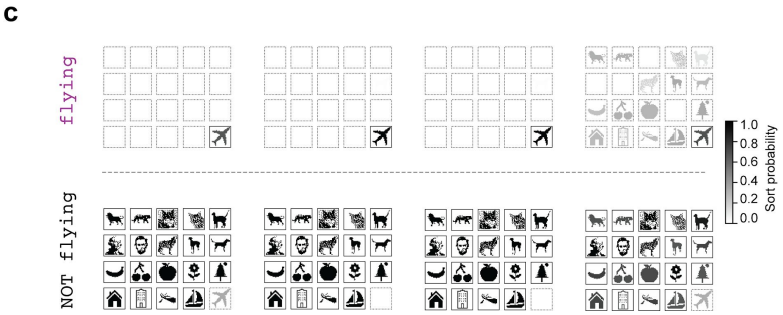
Encapsulation and barcoding of DNA in silica capsules



Searching for a needle in a huge haystack







Boolean search capable

A search engine for molecules using molecules

CACHE DNA Images

president AND NOT 18th century