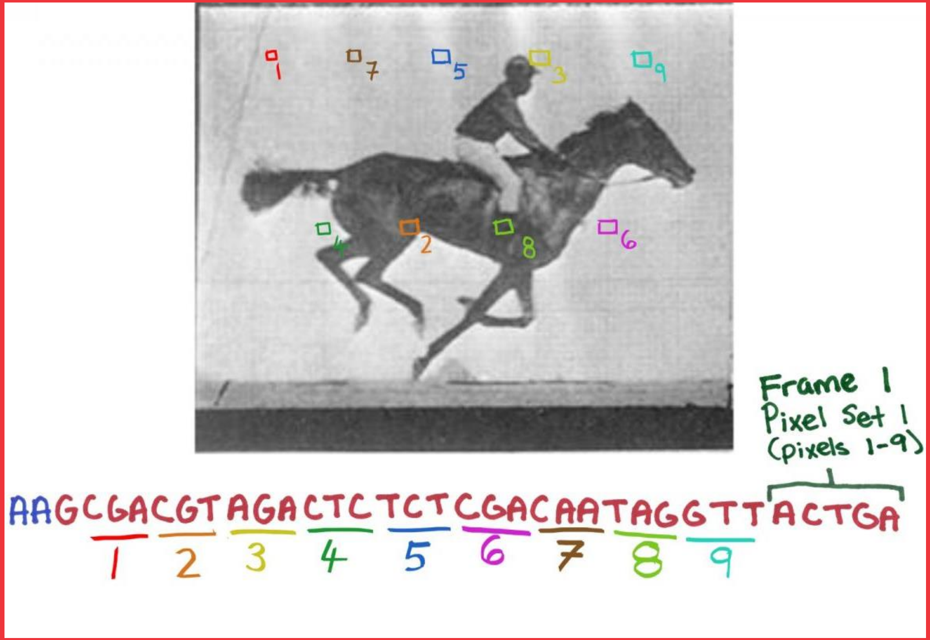
CRISPR-Cas – encoding image and a movie

Linh Tong, Filip Lewicki, Lotta Laihotie

April 26th, 2021

Aalto-yliopisto
Aalto-universitetet
Aalto University



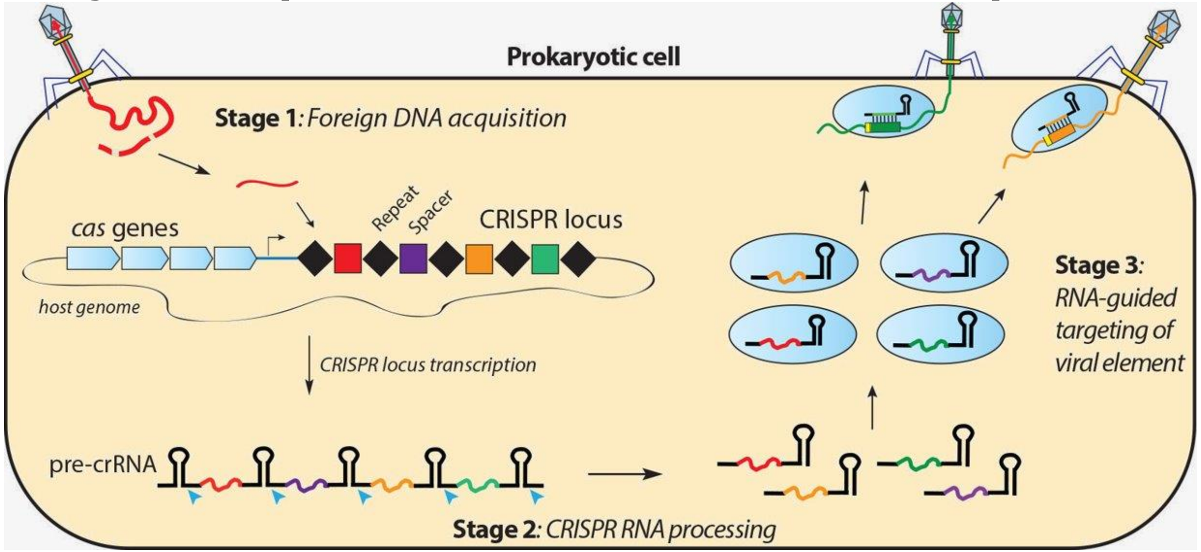
Credit: https://www.genengnews.com/

Contents

1. Basics of CRISPR

CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats

- **Adaptive immune system of prokaryotic organisms**
- **A tool for genome editing**
    - Nobel Prize in Chemistry 2020



https://doudnalab.org/research_areas/crispr-systems/

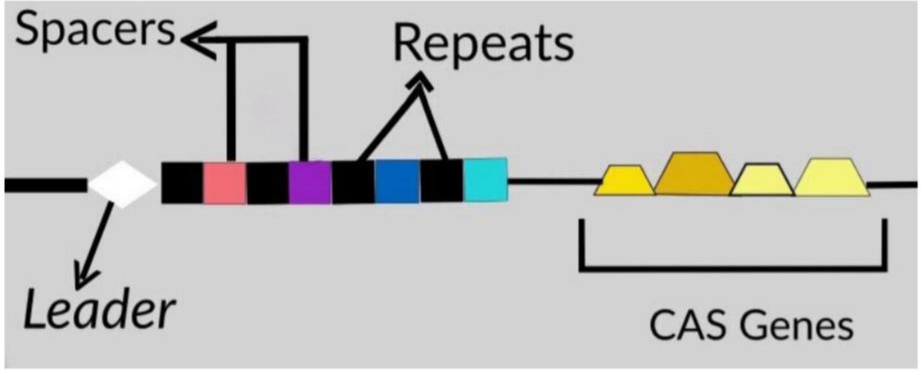
Marraffini, L. CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61 (2015). https://doi.org/10.1038/nature15386

1. Basics of CRISPR

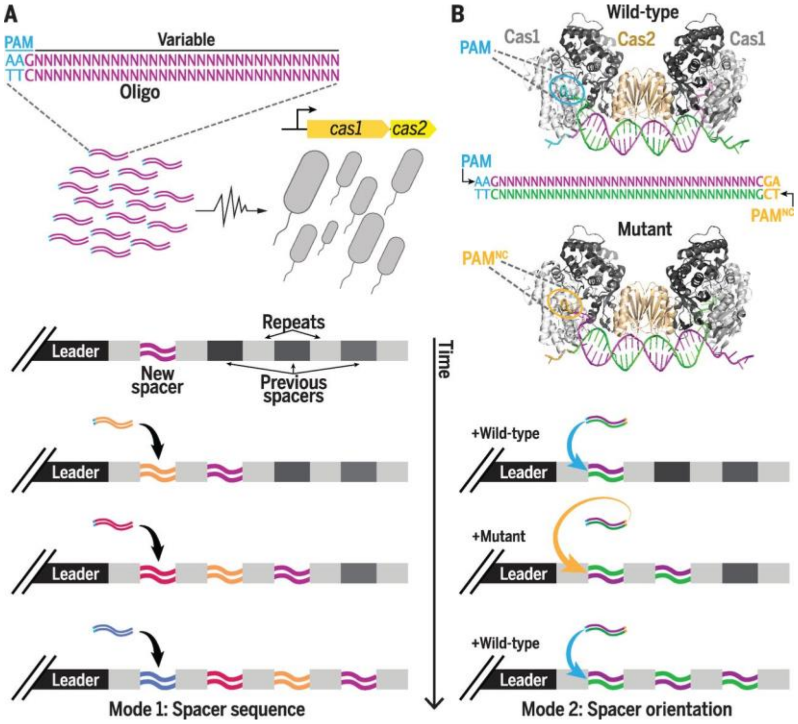
- **First identified in the *E. coli* genome in 1987**
    - The functional importance of the CRISPR loci remained elusive until 2005

- **CRISPR-associated (cas) genes encode cas proteins**
    - Helicases and nucleases
    - Cas 1 and 2 involved in spacer acquisition

CRISPR-Cas array



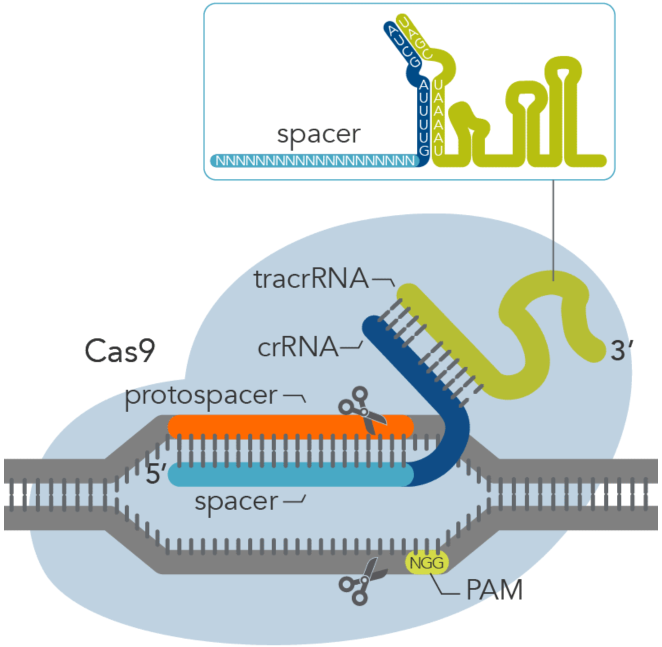
Structure of CRISPR array (Mohamadi et al,. 2020)



CRISPR-Cas array using Cas 1 and Cas 2 (Shipman, 2016)

CRISPR-Cas System

- **Spacers** = short DNA sequences between the short palindromic repeats, each segment is unique
    - **Protospacers** = sequences from which spacers are derived
- **Spacer acquisition** = adaption, integrating foreign genome into CRISPR array (McGinn & Marraffini, 2019)
- **PAM** = protospacer adjacent motif
- **AAM** = acquisition affecting motif



(Credits: idtdna.com)

2. CRISPR-Cas encoding a digital movie

LETTER

doi:10.1038/nature23017

CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria

Seth L. Shipman[1,2,3], Jeff Nivala[1,3], Jeffrey D. Macklis[2] & George M. Church[1,3]

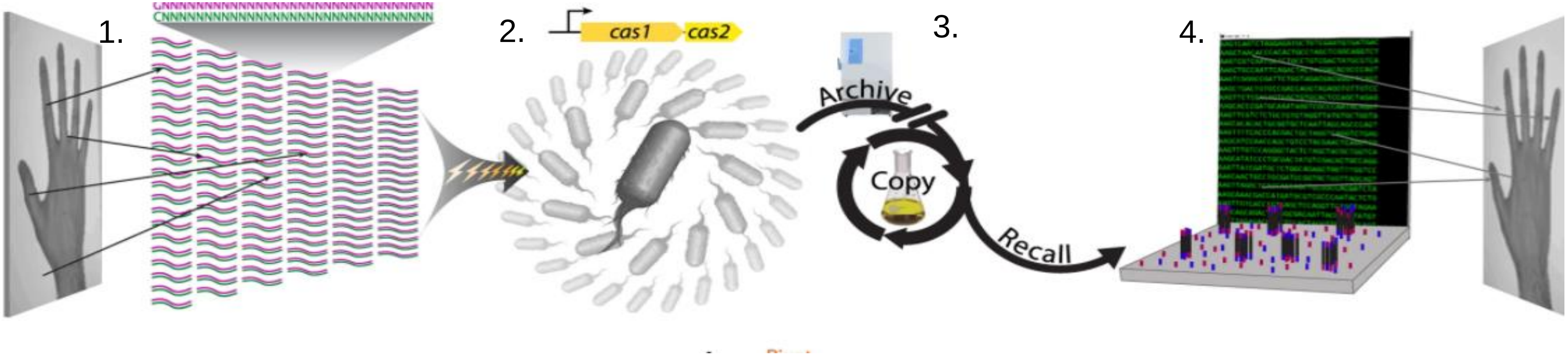
- **Encode pixel values into the genomes of a population of living bacteria**
- **Optimize delivery method, nucleotide content, and reconstruction method → DNA as a stable type of "molecular recorder"**
- **Discover principles of CRISPR-Cas adaptation system**

! This presentation mainly focuses on the methology

- Not focus: error quantification and optimization process

Materials and Methods

Encoding image into genome

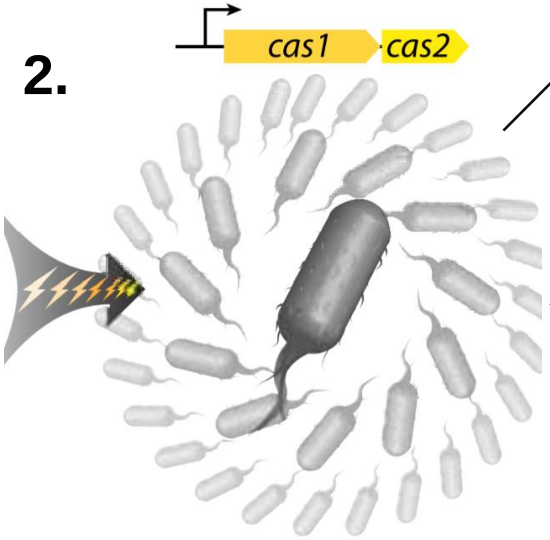


1. **Encoding of pixel values to protospacers**
2. **Electroporation of protospacers into bacteria (overexpress Cas1 and 2)**
3. **Bacteria can be archived and copied**
4. **In the final stage the bacteria are sequenced to recall the image**

S L Shipman *et al. Nature* 1–5 (2017) doi:10.1038/nature23017

Differing pixel-value-encoding strategies

2.



Hand^R
(A rigid strategy)

Hand^F
(A flexible strategy)

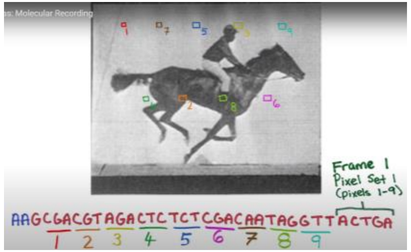
1st base	2nd base: T	C	A	G	3rd base
T	1	2	3	4	T
	5	6	7	8	C
	9	10	11	12	A
	13	14	15	16	G
C	17	18	19	20	T
	21	1	2	3	C
	4	5	6	7	A
	8	9	10	11	G
A	12	13	14	15	T
	16	17	18	19	C
	20	21	1	2	A
	3	4		5	G
G	6	7	8	9	T
	10	11	12	13	C
	14	15	16	17	A
	18	19	20	21	G

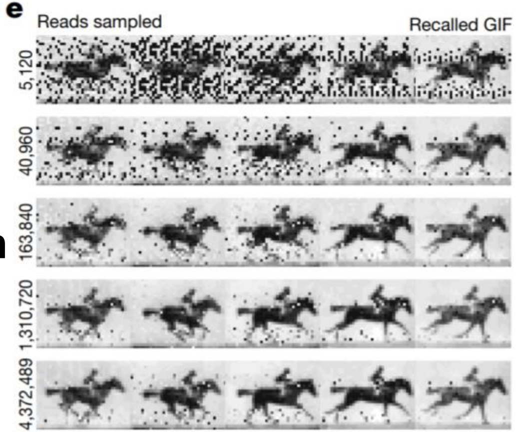
- Hand^R
  - 4 pixel colors defined by a differing base
- Hand^F
  - 21 pixel colors defined by a nucleotide table

S L Shipman *et al. Nature* 1–5 (2017) doi:10.1038/nature23017

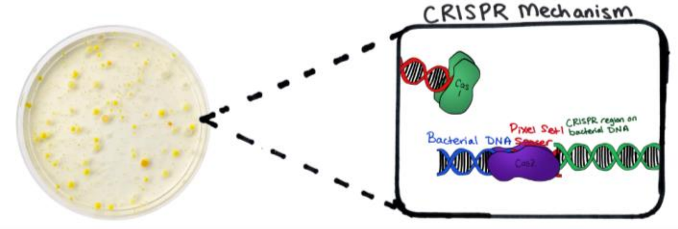
Encoding a GIF into genome

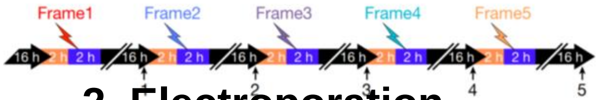
1. Encode frames using hand[F]



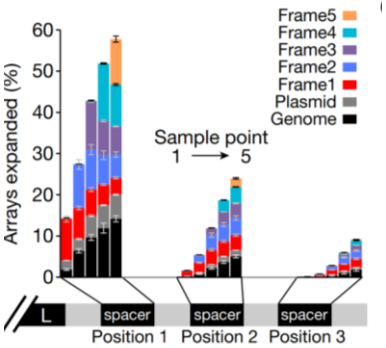


4. GIF recontruction





2. Electroporation



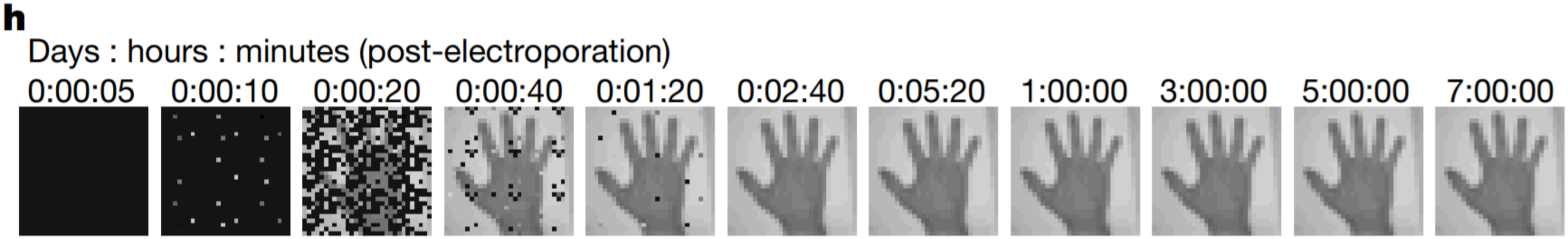
3. Spacer integration

Aalto-yliopisto
Aalto-universitetet
Aalto University

Shipman *et al. Nature* 1–5 (2017) doi:10.1038/nature23017

3. Results for encoding the image(s)

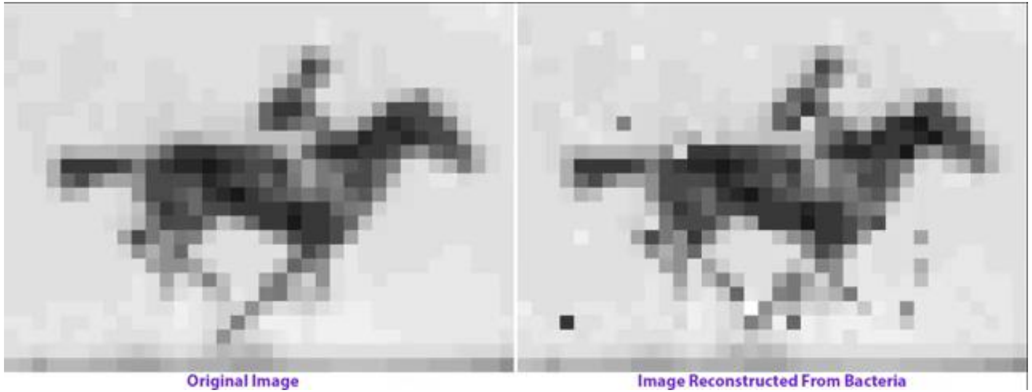
- **Reconstructed image from bacteria of hand[F]**
- **Out of 655,360 reads about 96% of pixet sequences were accurately recalled**
- **Reconstruction of GIF frames up to > 90% of overall accuracy**



Shipman *et al. Nature* 1–5 (2017) doi:10.1038/nature23017

3. Results for encoding the GIF

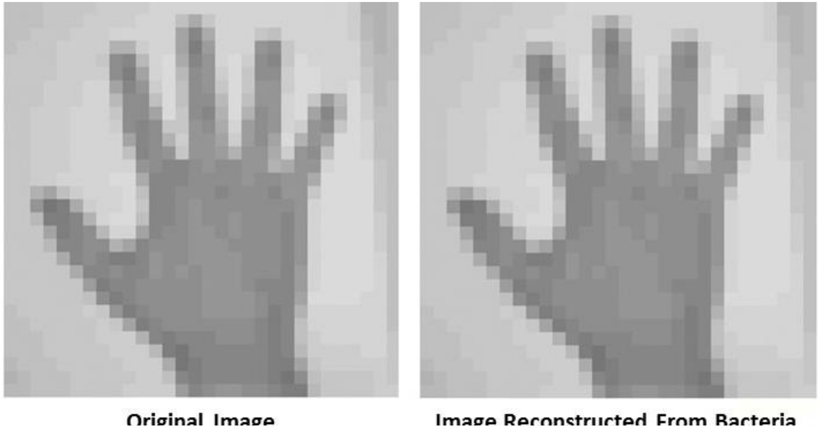
- **Not all protospacer sequences are equally effective → optimization**
- **Successfully tracked barcoded sequence elements over 5 times**
- **→ the system can record multidimensional biological information single-cell storage.**

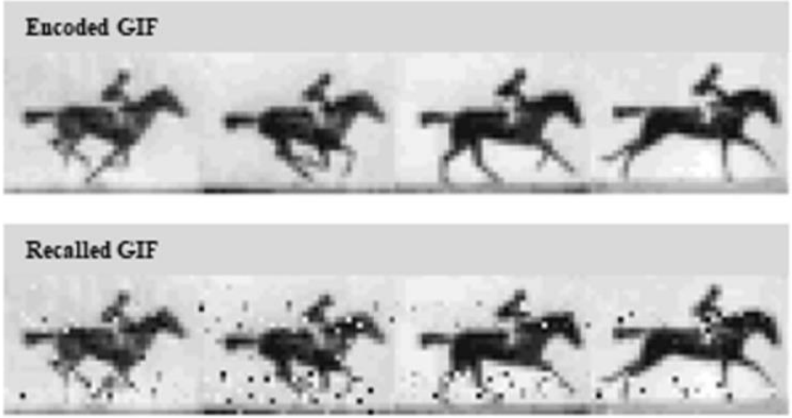


Shipman *et al. Nature* 1–5 (2017) doi:10.1038/nature23017

Summary

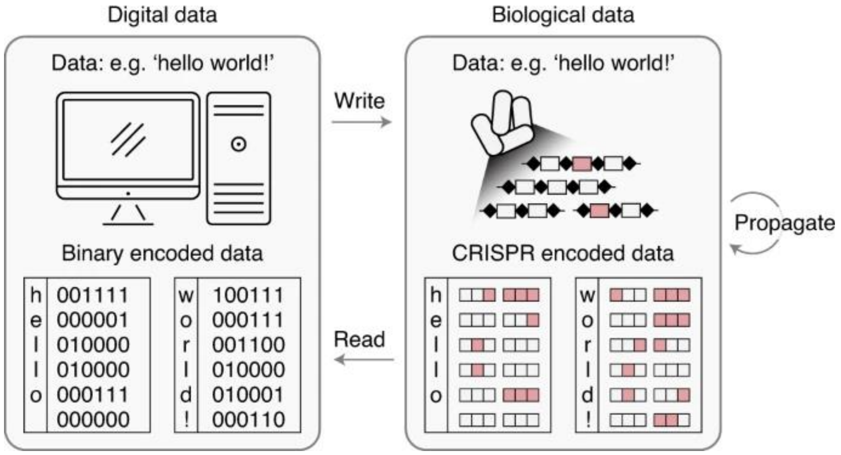
- **DNA can be utilized as a medium for storing data *in vitro***
- **Protospacer sequences differ at effectiveness of transferring data into genome**





Credit: Seth Shipman, Harvard Medical School, Boston

4. Cells (DNA) as information storage



(Yim, et al,. 2021)

Article | Open Access | Published: 22 October 2020

Low cost DNA data storage using photolithographic synthesis and advanced information reconstruction and error correction

Philipp L. Antkowiak, Jory Lietard, Mohammad Zalbagi Darestani, Mark M. Somoza, Wendelin J. Stark, Reinhard Heckel & Robert N. Grass

Nature Communications **11**, Article number: 5345 (2020) | Cite this article

3677 Accesses | **3** Citations | **19** Altmetric | Metrics

Limitation and future of DNA storage

Limitation

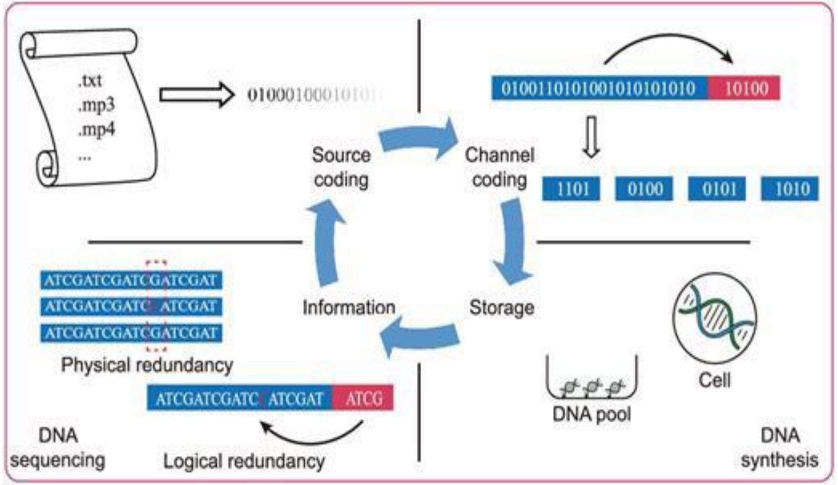
- **Time consuming to convert data**
- **Expensive to synthesize and read data (Service, 2021)**
- **Only short sequences are synthesized**
- **“Destructive process”**

Future (Lee, 2019)

- **DNA is stable**
- **Require less energy**
- **DNA has high storage capacity** *Escherichia coli* 10^{19} bits/cm^3

Conclusions

- **Ultilize the CRISPR system to integrate the encoding DNA into the genome of *E. coli*.** (Matsoukas, 2017)
    - an effective memory device.
    - scaled-up potential for recording in the genome memories/molecular experiences



General workflow of the DNA information storage process (Dong, et al., 2020)

References

- Lee, S.Y. 2019. *DNA Data Storage Is Closer Than You Think.* 7 1. Accessed 4 22, 2021. https://www.scientificamerican.com/article/dna-data-storage-is-closer-than-you-think/.
- Seth L. Shipman, Jeff Nivala, Jeffrey D. Macklis, George M. Church. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. Nature July 2017, Vol 547, 345-. doi:10.1038/nature23017
- Shipman, S. L., J. Nivala, J. D. Macklis, and G. M. Church. 2016. "Molecular Recordings by Directed CRISPR Spacer Acquisition." Science 353 (6298) (June 9): aaf1175–aaf1175. doi:10.1126/science.aaf1175.
- Service, R. 2021. *Scientists 'program' living bacteria to store data.* 1 11. Accessed 4 22, 2021. https://www.sciencemag.org/news/2021/01/scientists-program-living-bacteria-store-data.
- Marraffini, L. CRISPR-Cas immunity in prokaryotes. *Nature* 526**,** 55–61 (2015). https://doi.org/10.1038/nature15386
- Matsoukas, Ianis. (2017). Commentary: CRISPR–Cas Encoding of a Digital Movie into the Genomes of a Population of Living Bacteria. Frontiers in Bioengineering and Biotechnology. 5. 57. 10.3389/fbioe.2017.00057.
- McGinn, J., & Marraffini, L. A. (2019). Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nature Reviews.Microbiology, 17*(1), 7-12. doi:http://dx.doi.org/10.1038/s41579-018-0071-7
- Mohamadi, S., S. Z. Bostanabad and R. Mirnejad. "CRISPR Arrays: A Review on Its Mechanism." *Journal of Applied Biotechnology Reports* 7 (2020): 81-86.
- Yim, S.S., McBee, R.M., Song, A.M. *et al.* Robust direct digital-to-biological data storage in living cells. *Nat Chem Biol* **17,** 246–253 (2021). https://doi.org/10.1038/s41589-020-00711-4
- Yiming Dong, Fajia Sun, Zhi Ping, Qi Ouyang, Long Qian, DNA storage: research landscape and future prospects, *National Science Review*, Volume 7, Issue 6, June 2020, Pages 1092–1107, https://doi.org/10.1093/nsr/nwaa007
- Scientific Background on the Nobel Prize in Chemistry 2020. A Tool for Genome Editing. https://www.nobelprize.org/uploads/2020/10/advanced-chemistryprize2020.pdf

Thank you!

Aalto-yliopisto
Aalto-universitetet
Aalto University

Questions?

Aalto-yliopisto
Aalto-universitetet
Aalto University