北京大学 PEKING UNIVERSITY

生命科学学院 School of Life Sciences

A time-resolved, multi-symbol molecular recorder via sequential genome editing

UHPB Journal Club

Jiaxi Liu (刘家希)

School of Life Sciences

2026.5.30

Paper

DNA Typewriter

Article | Open access | Published: 06 July 2022

A time-resolved, multi-symbol molecular recorder via sequential genome editing

Junhong Choi ✉, Wei Chen, Anna Minkina, Florence M. Chardon, Chase C. Suiter, Samuel G. Regalado, Silvia Domcke, Nobuhiko Hamazaki, Choli Lee, Beth Martin, Riza M. Daza & Jay Shendure ✉

Nature **608**, 98–107 (2022) | Cite this article

Author

Jay Shendure

- Lead Scientific Director, Seattle Hub for Synthetic Biology (Allen-CZI-UW)
- Scientific Director, Allen Discovery Center for Cell Lineage Tracing
- Scientific Director, Brotman Baty Institute for Precision Medicine
- Full Professor (with tenure), Dept. of Genome Sciences, University of Washington
- Investigator, Howard Hughes Medical Institute

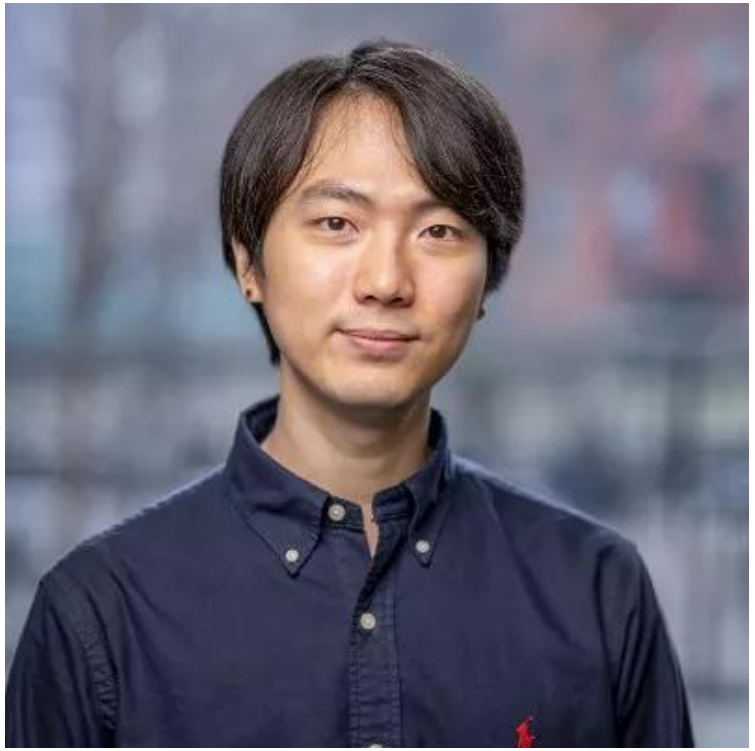
Research Interest:

DNA sequencing technologies, Multiplex molecular methods, Human genetics & medical diagnostics, Mammalian gene regulation, Developmental biology

- AB, Princeton University
- PhD, Harvard University
- MD, Harvard Medical School

Author

Junhong Choi



- Assistant Professor, Sloan Kettering Institute
- PostDoc, Genome Sciences, University of Washington (advisor: **Jay Shendure**)

Research Interest

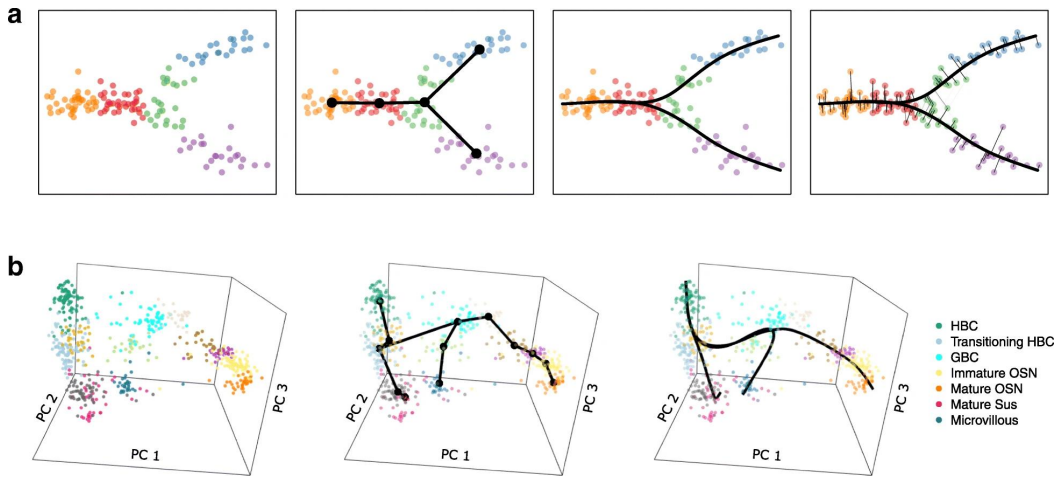
develops new synthetic biology tools to study cell-fate decisions in development. focusing on developing such new tools through:

- Cellular engineering
- Molecular circuitry
- DNA as an information storage medium, and
- Synthetic models for cellular development

- B.S., University at Buffalo
- PhD, Stanford University (advisor:Joseph Puglisi & Zev Bryant)

How do we learn the order of molecular events in living systems?

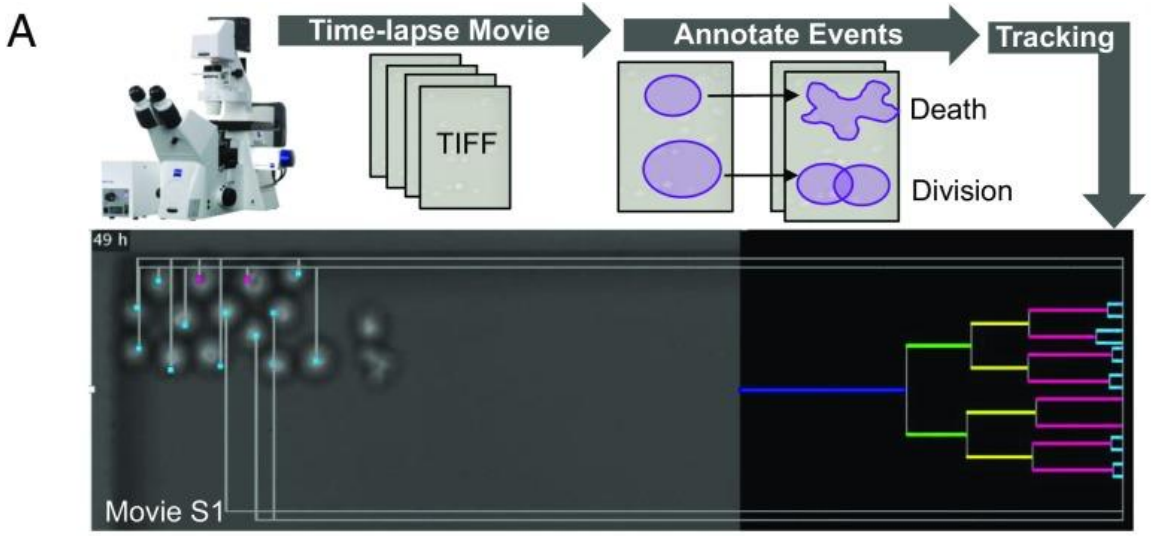
Traditional method



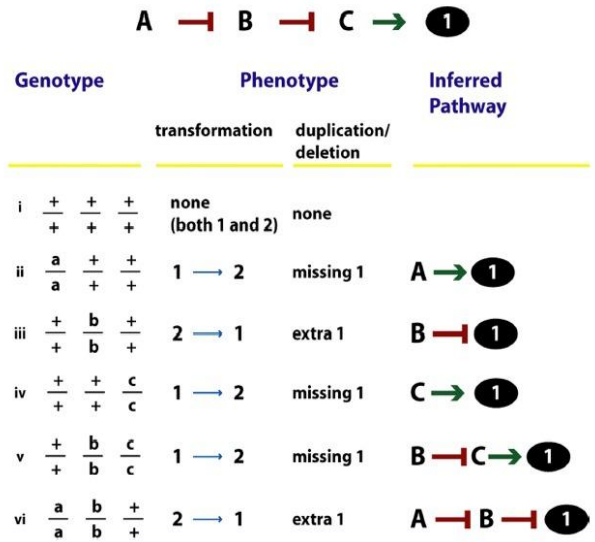
Time-series scRNA–seq

Street K, et al. *BMC Genomics*. 2018

Live-cell imaging



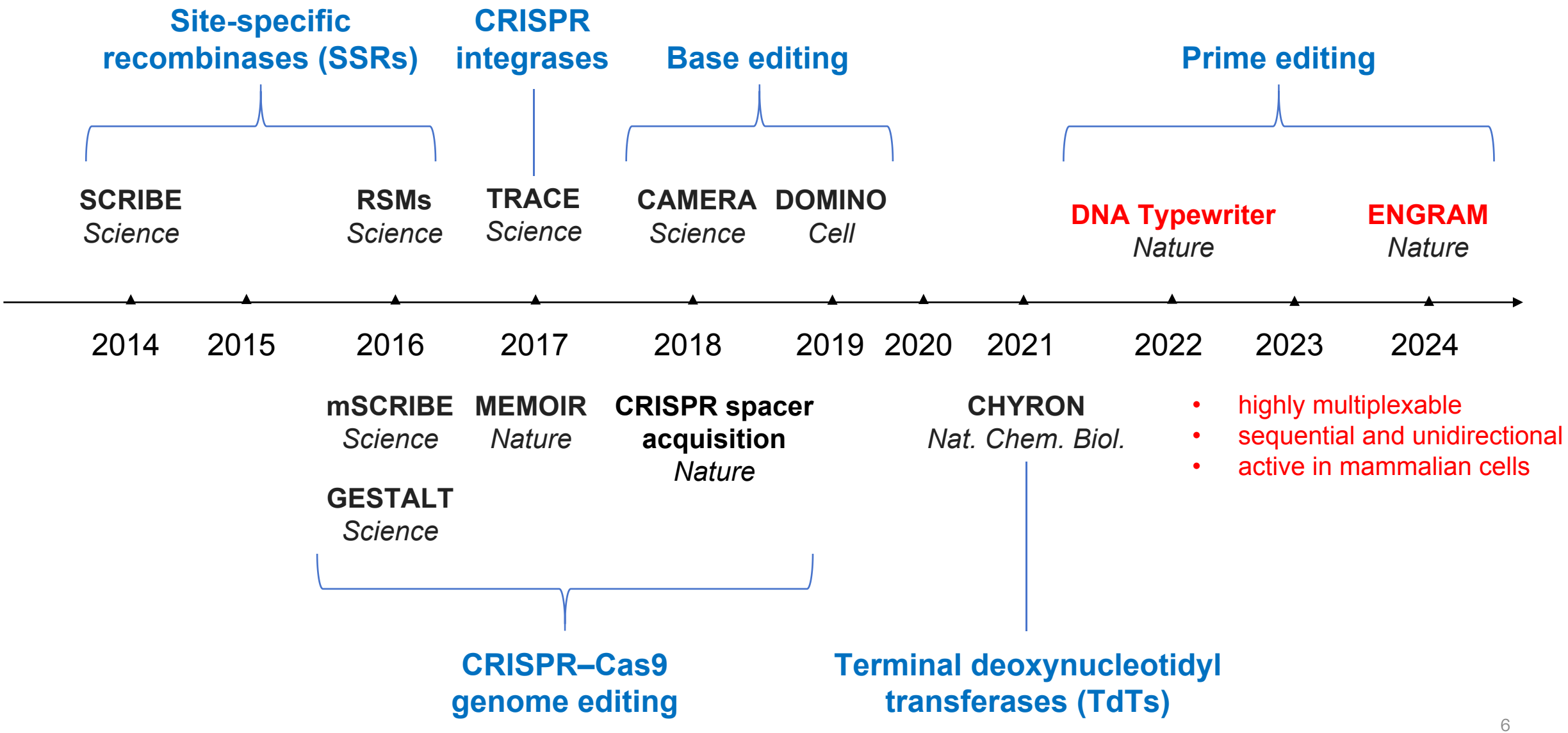
Mitchell S, et al. *PNAS*. 2018



Epistasis analysis

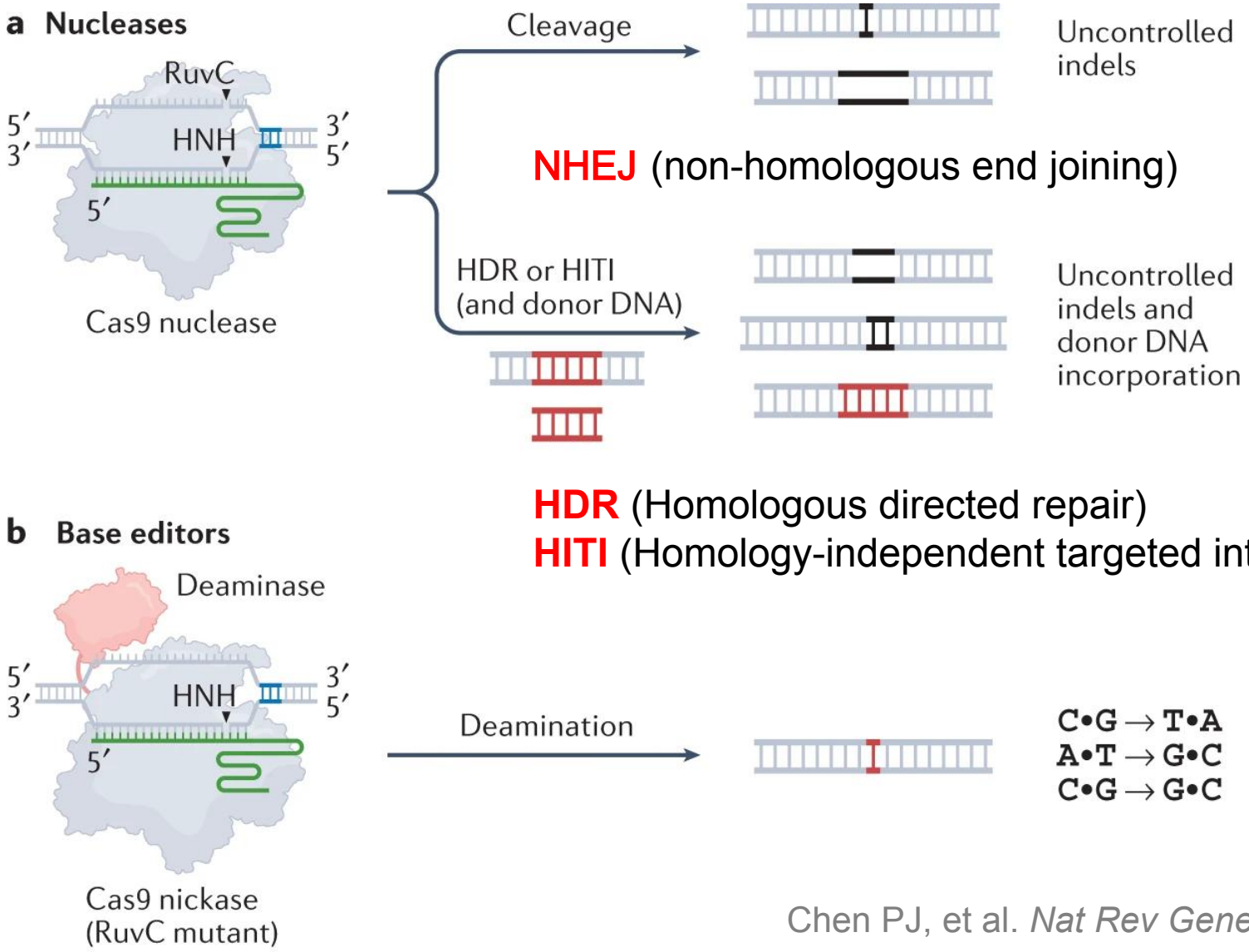
Huang LS, et al. *WormBook*. 2006

History of DNA memory device



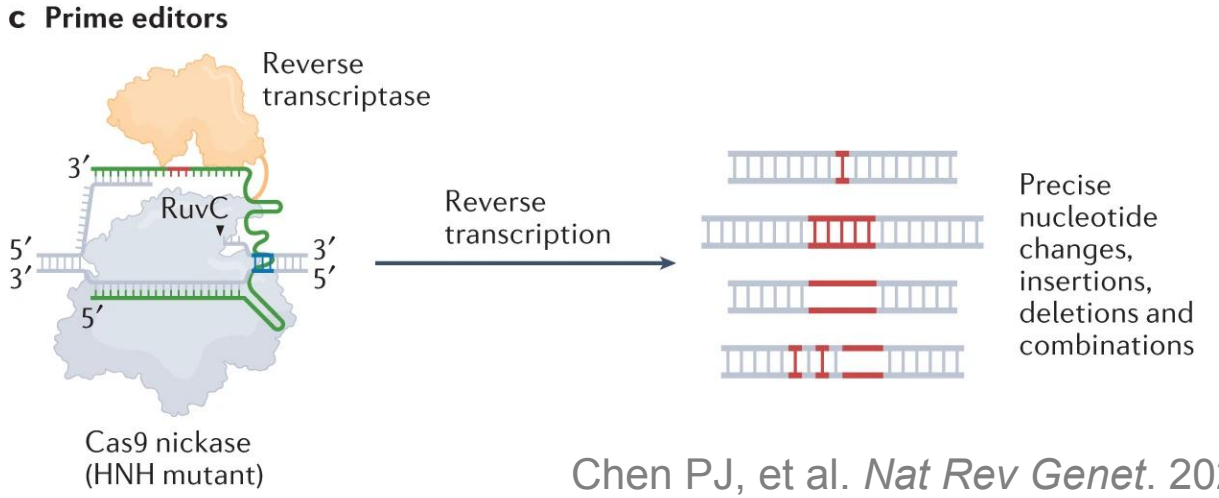
Cas Nucleases and Base editors

DSB
(DNA double-strand break)

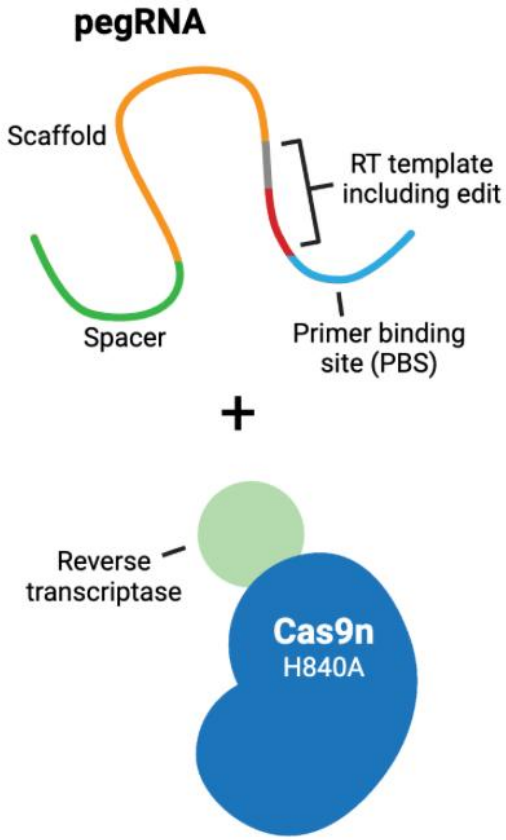


Chen PJ, et al. *Nat Rev Genet*. 2023

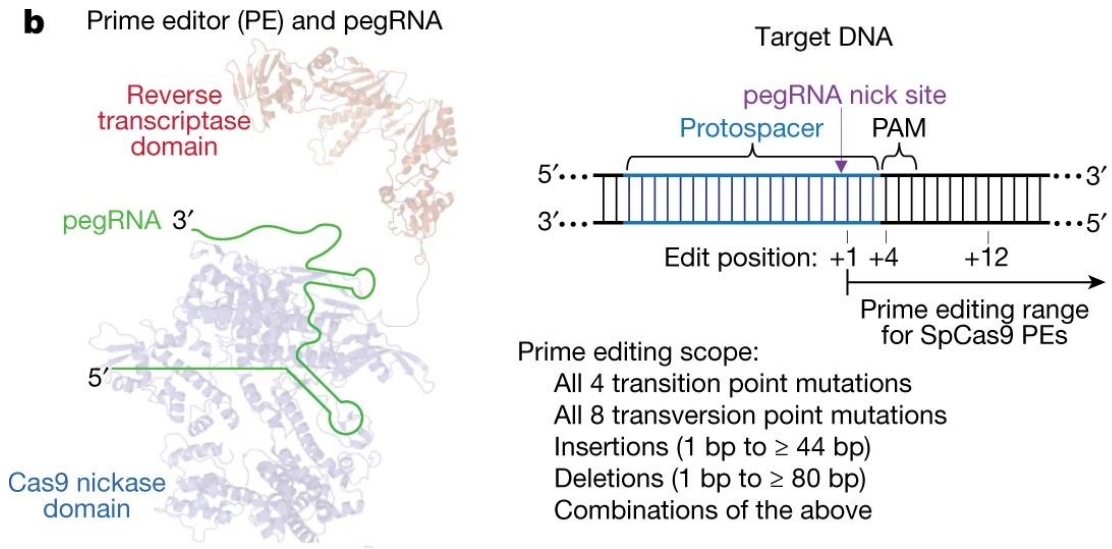
Prime Editing



Chen PJ, et al. *Nat Rev Genet*. 2023

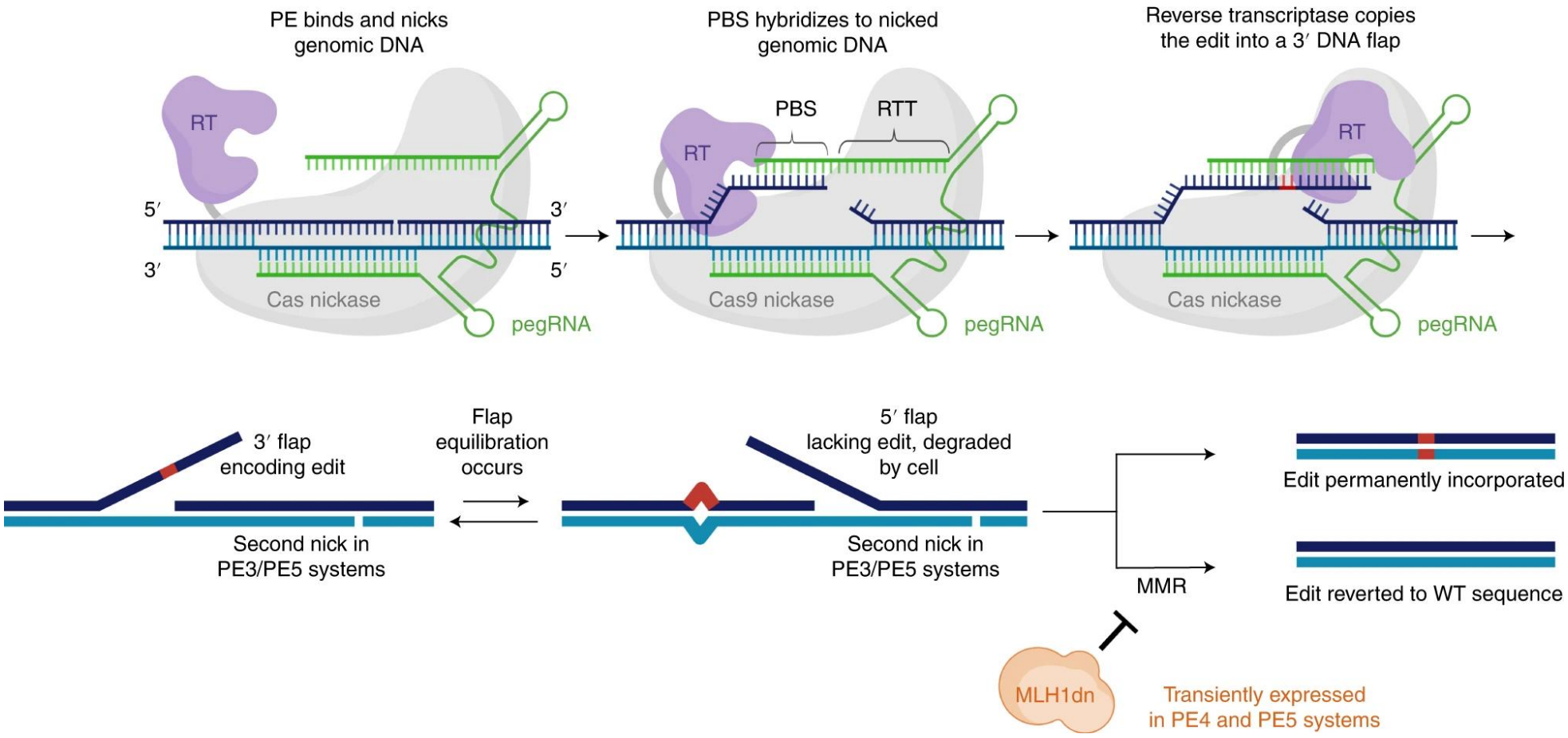


https://www.addgene.org/crispr/prime-edit/



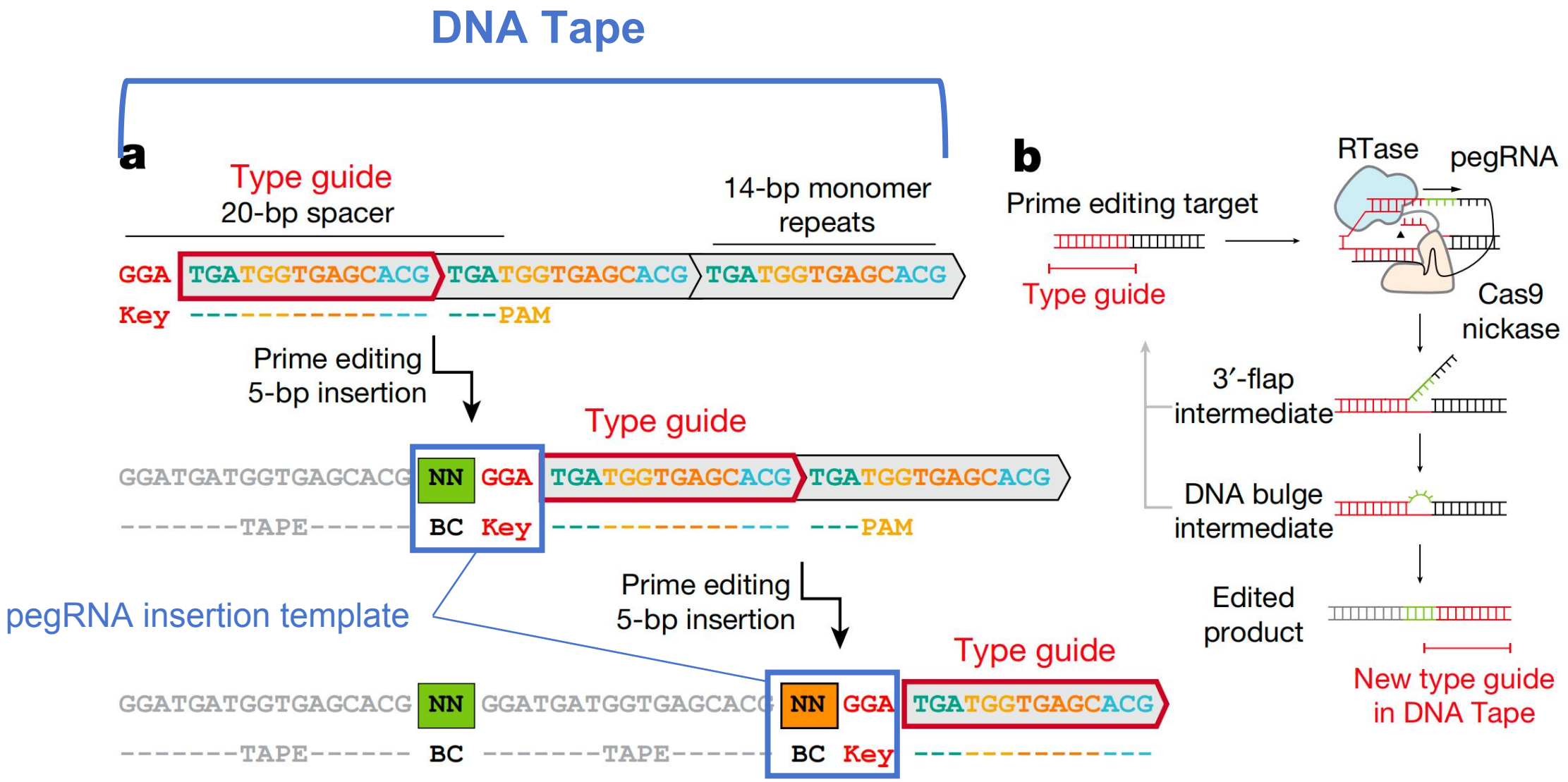
Anzalone AV, et al. *Nature*. 2019

Mechanism of Prime Editing

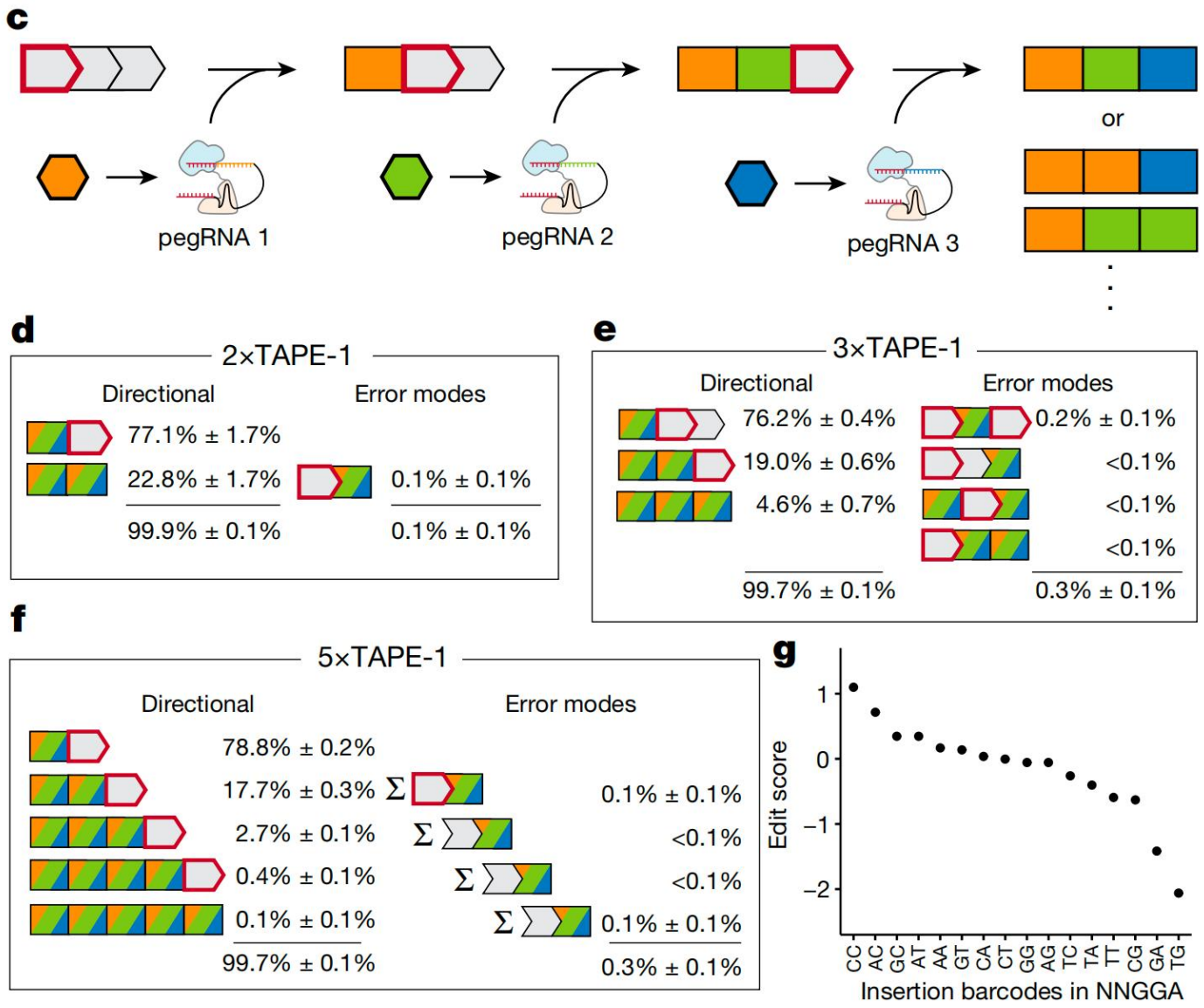


Doman JL, et al. *Nat Protoc*. 2022

DNA Typewriter



Ordered recording



cells: HEK293T piggyBac（2×TAPE-1、3×TAPE-1、5×TAPE-1）

Transient transfection: pCMV-PE2-P2A-GFP + 16x pegRNA

readout:PCR amplify DNA Tape + sequencing

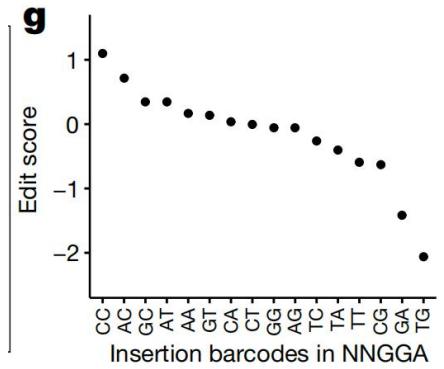
2×TAPE-1: 4.7% ± 0.5%
3×TAPE-1: 5.2% ± 0.6%
5×TAPE-1: 5.9% ± 0.8%

exhibited any editing

~6%

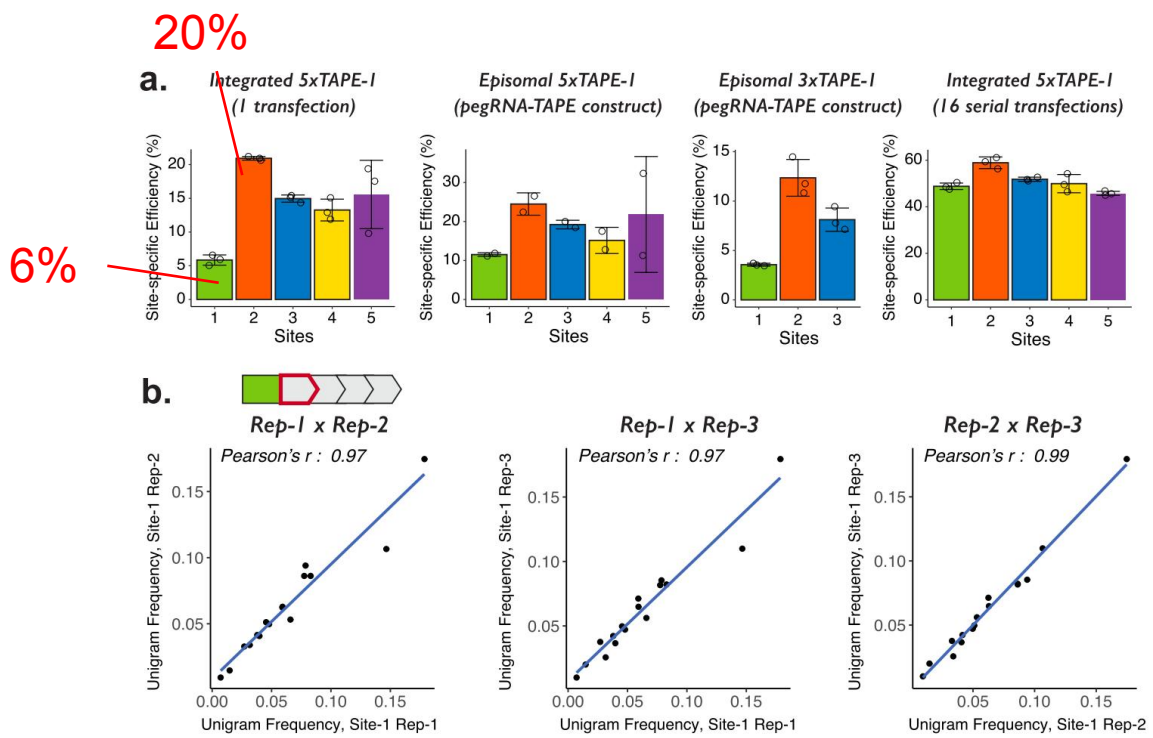
- **editing efficiency**
- **sequential fidelity**
- **symbol bias**

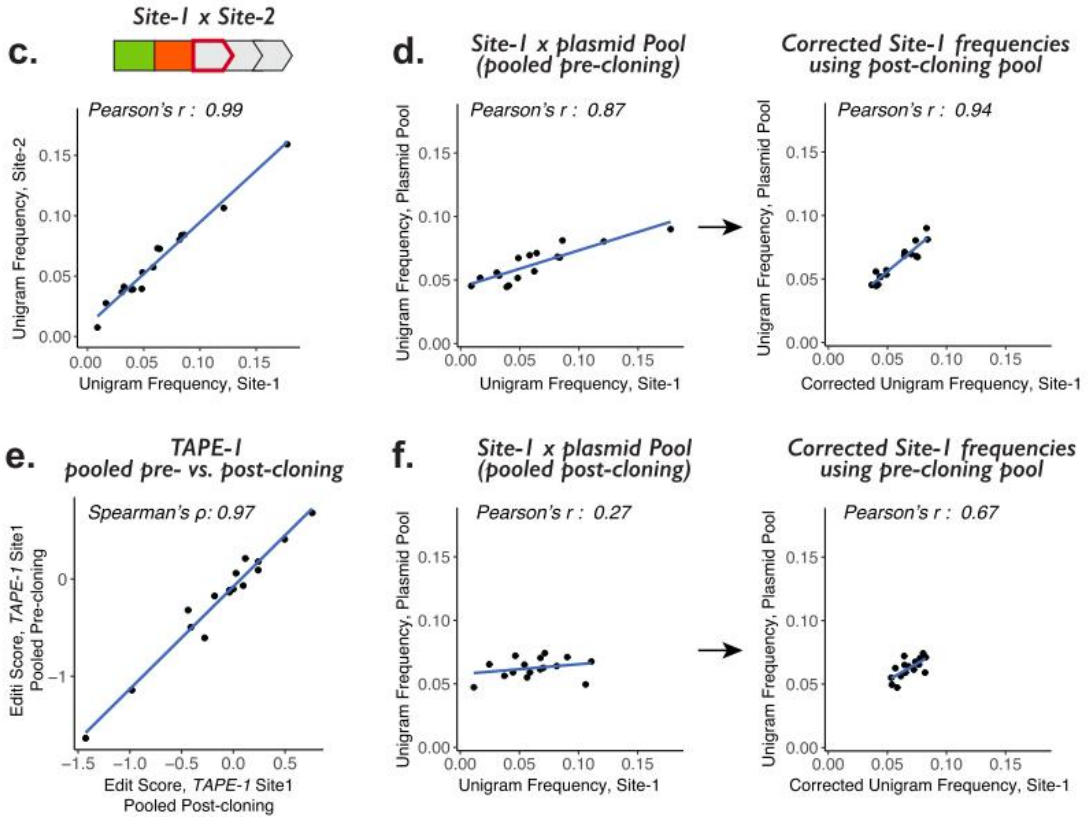
barcode-specific editing bias and pseudo-processivity



$$\mathrm{EditScore}_i = \log_2\left[\frac{N_i^{\mathrm{ins}} / \sum_j N_j^{\mathrm{ins}}}{N_i^{\mathrm{pegRNA}} / \sum_j N_j^{\mathrm{pegRNA}}}\right]$$

- editing efficiency
- sequential fidelity
- symbol bias



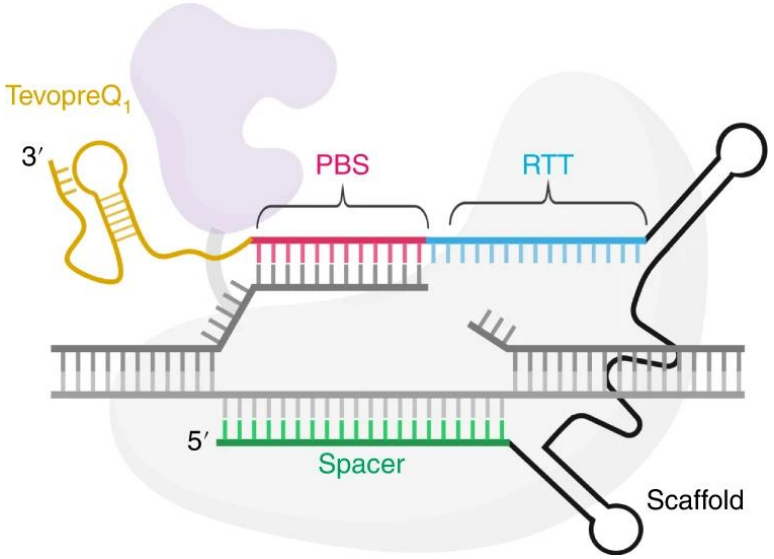


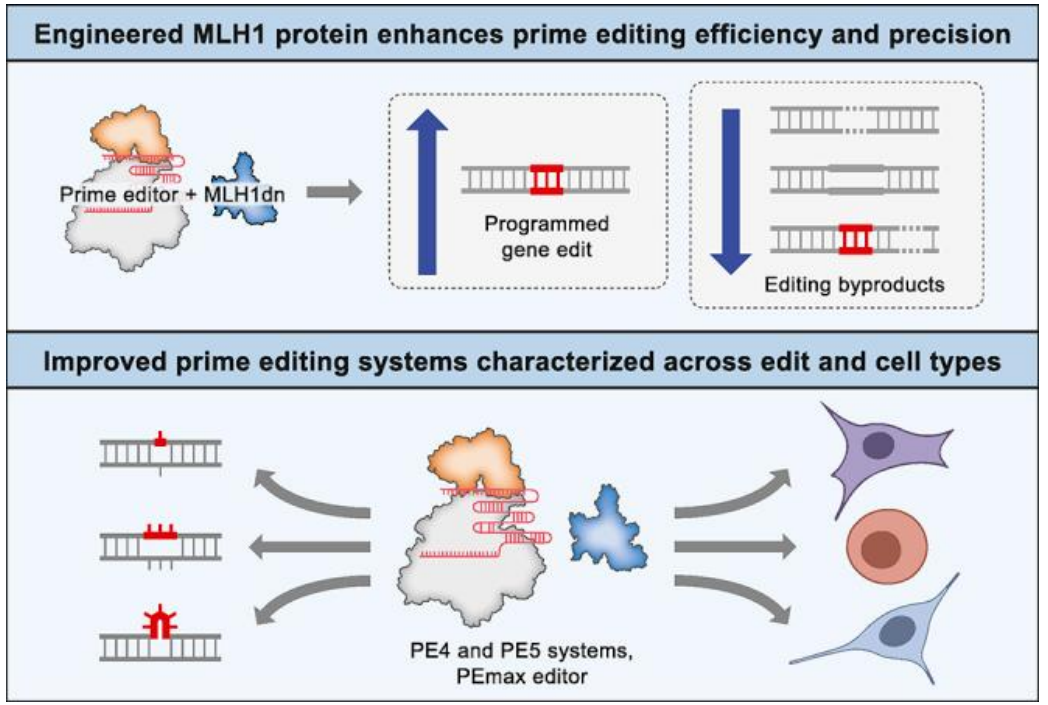
Enhanced prime editing of DNA Tape

epegRNA: addition of degradation-resistant secondary structure to the 3′ end of the pegRNA

hMLH1dn: introduction of human MLH1 dominant-negative peptide (hMLH1dn) to favour the intended edit

PEmax: modifications to the primary sequence of the prime editing enzyme





Chen PJ, et al. *Cell*. 2021

Cells: HEK293T piggyBac（5×TAPE-1）

Transient transfection: pCMV-**PEmax**-P2A-**hMLH1dn** +16x **epegRNA (U6-driven)**

Readout:PCR amplify DNA Tape + sequencing

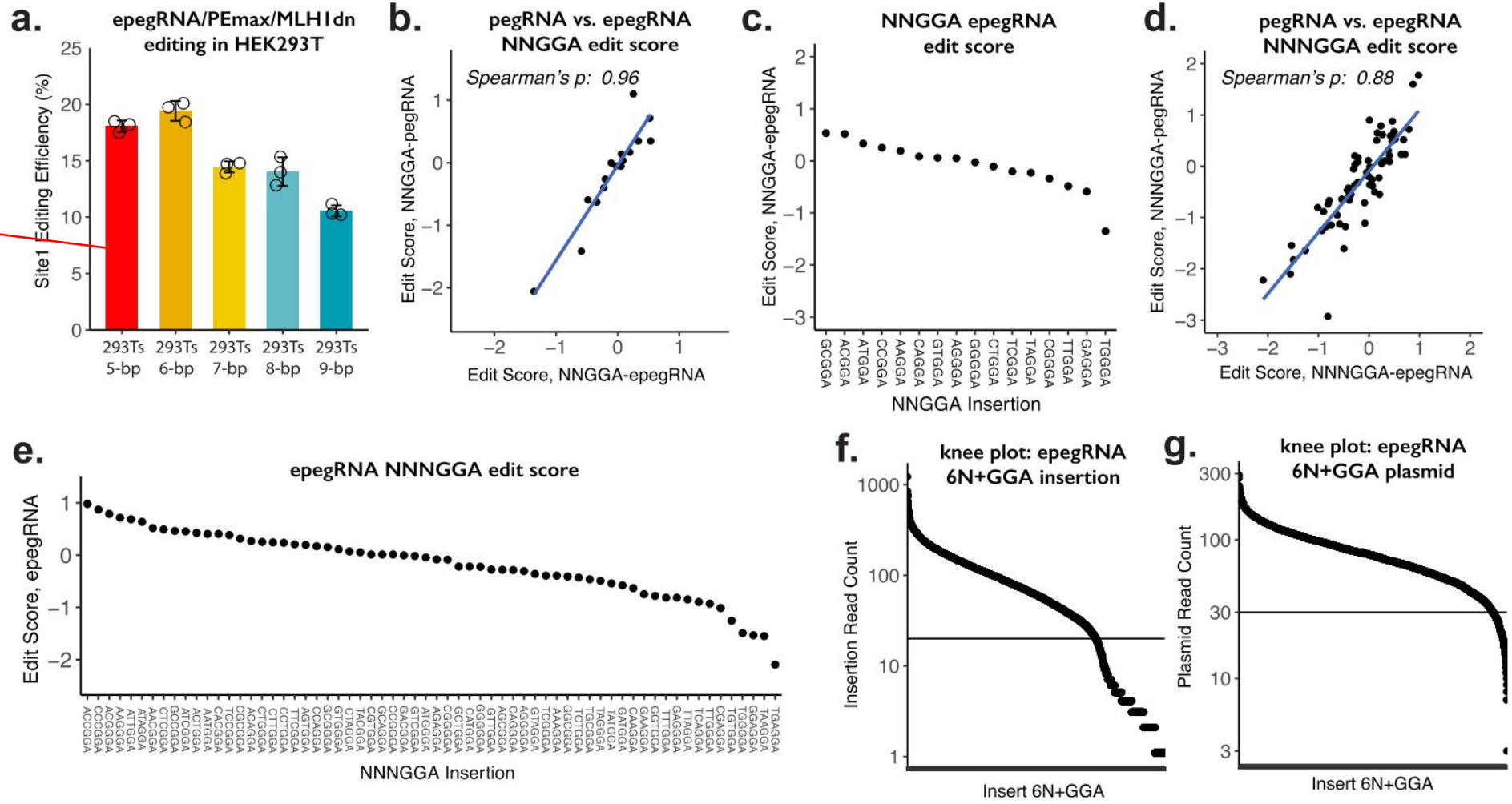
Enhanced prime editing increase editing efficiency

pCMV-**PEmax**-P2A-**hMLH1dn** +16x **epegRNA**

NNGGA -6N+GGA

~6%

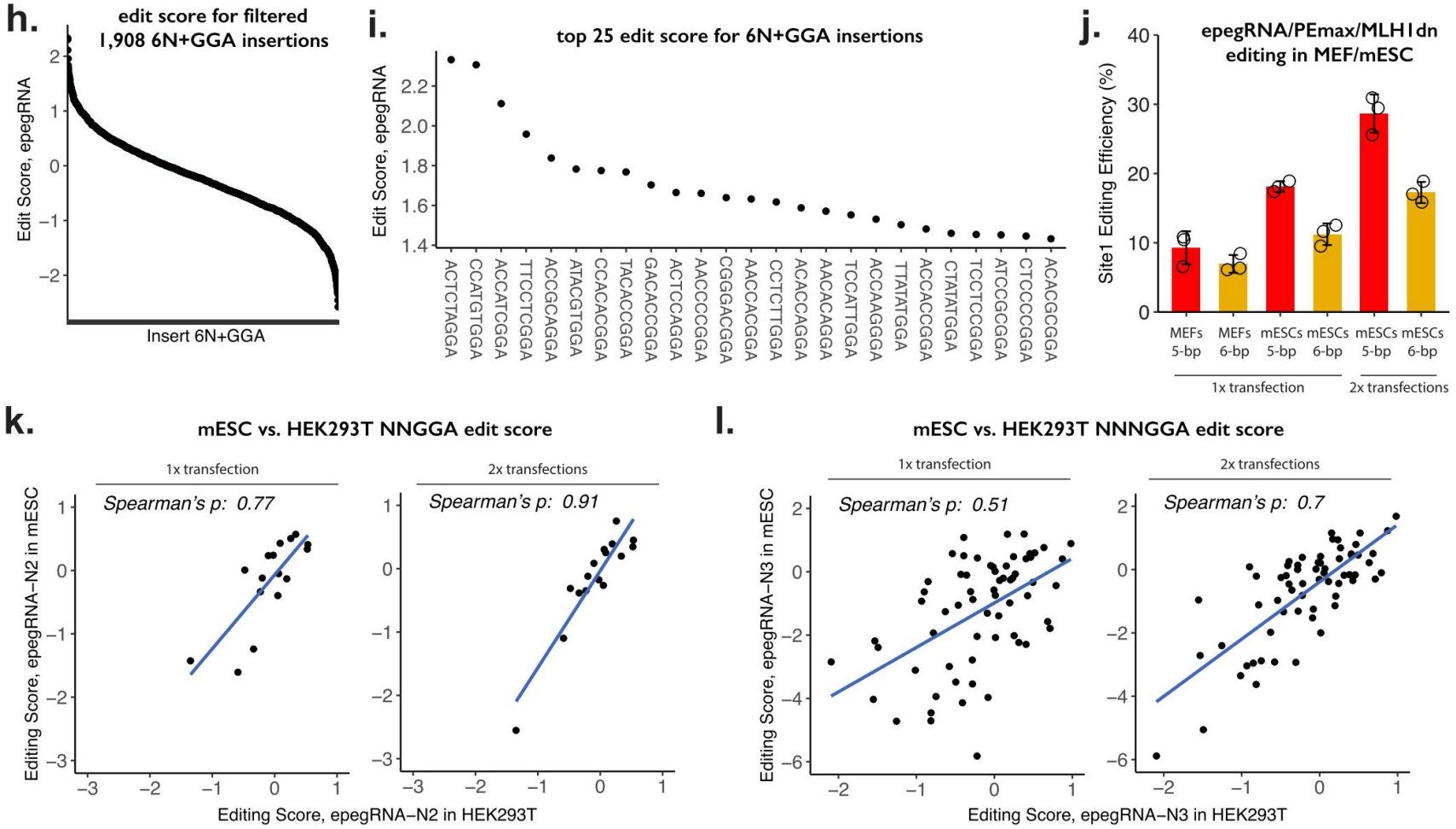
18.1%



DNA recording

HEK293T → mouse embryonic fibroblasts (MEF)
mouse embryonic stem cells (mESC)

pCMV-**PEmax**-P2A-**hMLH1dn** +16x **epegRNA**



Screening additional DNA Tape sequences

TAPE-1 ➡ TAPE**48**

8 basal spacers x **6** design rules

cells: HEK293T piggyBac（3×TAPE-1 ～48）

Transient transfection: pCMV-PE2-P2A-GFP + 48x pegRNA-by 3xTAPE pool

readout:PCR amplify DNA Tape + sequencing

pegRNA-expressing cassette
4–6bp insert (2bp NN +2–4bp key)

a.

HEK3	GGCCCAGACTGAGCACG TGATGGCAGAGGAA	(Base-1)
TAPE-1	GGATGATGGTGAGCACG TGATGGTGAGCACG	(Design-1)
TAPE-2	GGATGATGGCAGAGGAG TGATGGCAGAGGAG	(Design-2)
TAPE-3	GGACCAGGGTGAGCACG CCAGGGTGAGCACG	(Design-3)
TAPE-4	GACTGATGGTGAGCACG TGATGGTGAGCACG	(Design-4)
TAPE-5	GATGATGGGTGAGCACC TGATGGGTGAGCACC	(Design-5)
TAPE-6	GGACTGATGGTGAGACG TGATGGTGAGACG	(Design-6)

K14	GTTGTTGTTCACCTCCT CGATGGCATTCACT	(Base-2)
TAPE-7	GGACGATGGCACCTCCT CGATGGCACCTCCT	(Design-1)
TAPE-8	GGACGATGGCATTCACG CGATGGCATTCACG	(Design-2)
TAPE-9	GGAGTTGGGCACCTCCT GTTGGGCACCTCCT	(Design-3)
TAPE-10	GACCGATGGCACCTCCT CGATGGCACCTCCT	(Design-4)
TAPE-11	GACGATGGTCACCTCCT CGATGGTCACCTCCT	(Design-5)
TAPE-12	GGACCGATGGCACTCCT CGATGGCACTCCT	(Design-6)

K21	GGTCCCTCCAGCATCTG CTCTGGCTCCATGG	(Base-3)
TAPE-13	GGACTCTGGAGCATCTG CTCTGGAGCATCTG	(Design-1)
TAPE-14	GGACTCTGGCTCCATGG CTCTGGCTCCATGG	(Design-2)
TAPE-15	GGACCCTGGAGCATCTG CCCTGGAGCATCTG	(Design-3)
TAPE-16	GACCTCTGGAGCATCTG CTCTGGAGCATCTG	(Design-4)
TAPE-17	GTCTCTGGGAGCATCTG CTCTGGGAGCATCTG	(Design-5)
TAPE-18	GGATCTCTGGAGCACCT CTCTGGAGCACCT	(Design-6)

K22	GCGAGTTCAAGTGCTAC CCGAGGTGCGAGGC	(Base-4)
TAPE-19	GGACCGAGGAGTGCTAC CCGAGGAGTGCTAC	(Design-1)
TAPE-20	GGACCGAGGTGCGAGGC CCGAGGTGCGAGGC	(Design-2)
TAPE-21	GGAAGTTGGAGTGCTAC AGTTGGAGTGCTAC	(Design-3)
TAPE-22	GATCCGAGGAGTGCTAC CCGAGGAGTGCTAC	(Design-4)
TAPE-23	GACCGAGGGAGTGCTAC CCGAGGGAGTGCTAC	(Design-5)
TAPE-24	GGATCCGAGGAGTGCAC CCGAGGAGTGCAC	(Design-6)

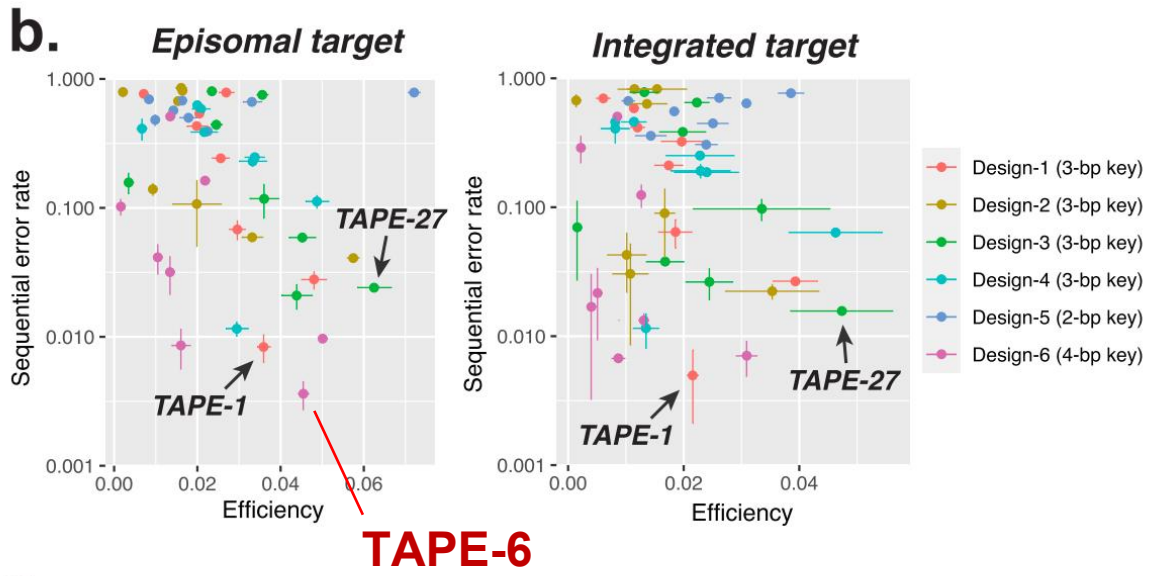
FANCF	GGAATCCCTTCTGCAGC ACCTGGATCGCTTC	(Base-5)
TAPE-25	GAGACCTGGTCTGCAGC ACCTGGTCTGCAGC	(Design-1)
TAPE-26	GAGACCTGGATCGCTTC ACCTGGATCGCTTC	(Design-2)
TAPE-27	GAGATCCGGTCTGCAGC ATCCGGTCTGCAGC	(Design-3)
TAPE-28	GTAACCTGGTCTGCAGC ACCTGGTCTGCAGC	(Design-4)
TAPE-29	GAACCTGGGTCTGCAGC ACCTGGGTCTGCAGC	(Design-5)
TAPE-30	GCAGACCTGGTCTGCAC ACCTGGTCTGCAC	(Design-6)

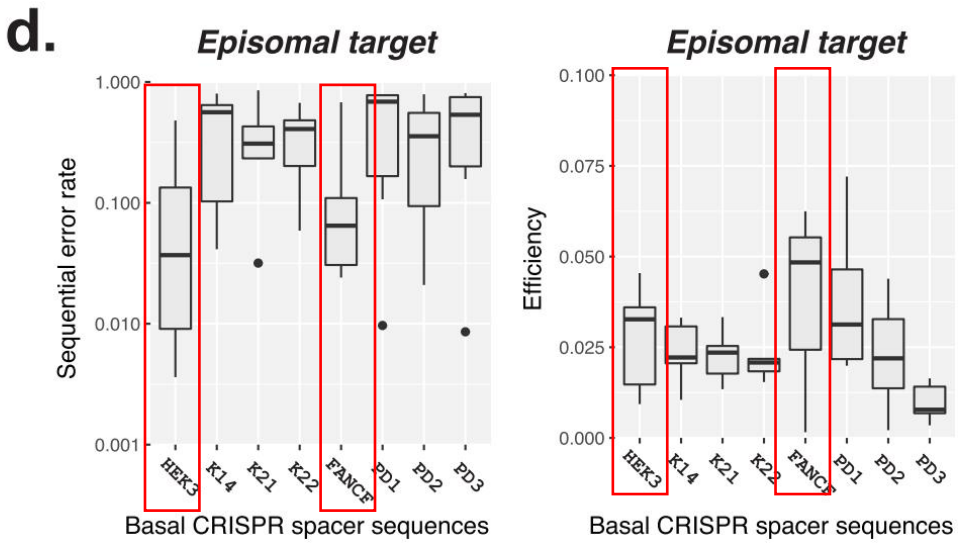
PD1	AACCTCTCGGCTTTCCC GCGCGGCGCCGCCC	(Base-6)
TAPE-31	GGAGCGCGGGCTTTCCC GCGCGGGCTTTCCC	(Design-1)
TAPE-32	GGAGCGCGGCGCCGCCC GCGCGGCGCCGCCC	(Design-2)
TAPE-33	GGACTCTGGGCTTTCCC CTCTGGGCTTTCCC	(Design-3)
TAPE-34	GATGCGCGGGCTTTCCC GCGCGGGCTTTCCC	(Design-4)
TAPE-35	GAGCGCGGAGCTTTCCC GCGCGGAGCTTTCCC	(Design-5)
TAPE-36	GGATGCGCGGGCTTTCC GCGCGGGCTTTCC	(Design-6)

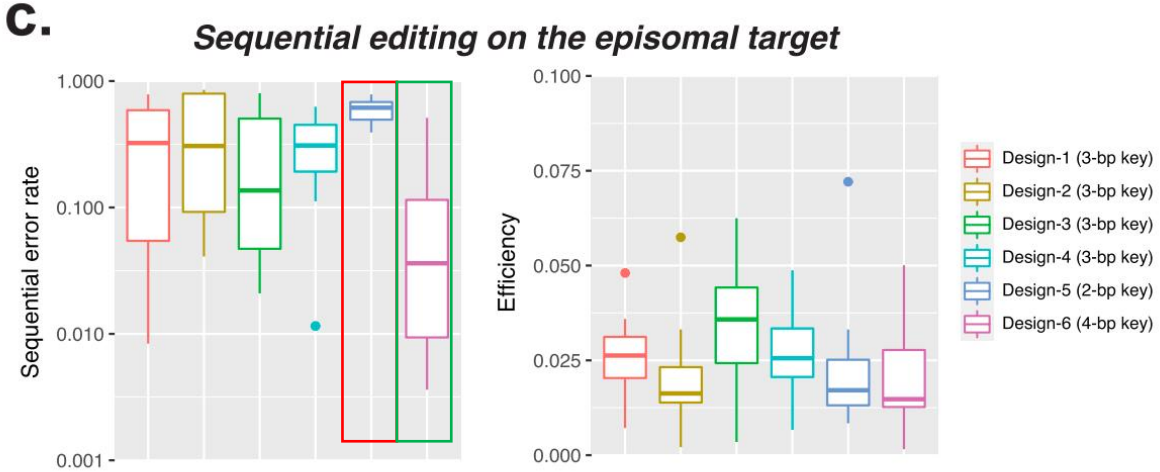
PD2	GCCTGCAAACTGGTAGG CGCCGGCGTAGGCG	(Base-7)
TAPE-37	GCACGCCGGCTGGTAGG CGCCGGCTGGTAGG	(Design-1)
TAPE-38	GCACGCCGGCGTAGGCG CGCCGGCGTAGGCG	(Design-2)
TAPE-39	GCATGCAGGCTGGTAGG TGCAGGCTGGTAGG	(Design-3)
TAPE-40	GATCGCCGGCTGGTAGG CGCCGGCTGGTAGG	(Design-4)
TAPE-41	GACGCCGGACTGGTAGG CGCCGGACTGGTAGG	(Design-5)
TAPE-42	GCAGCGCCGGCTGTAGG CGCCGGCTGTAGG	(Design-6)

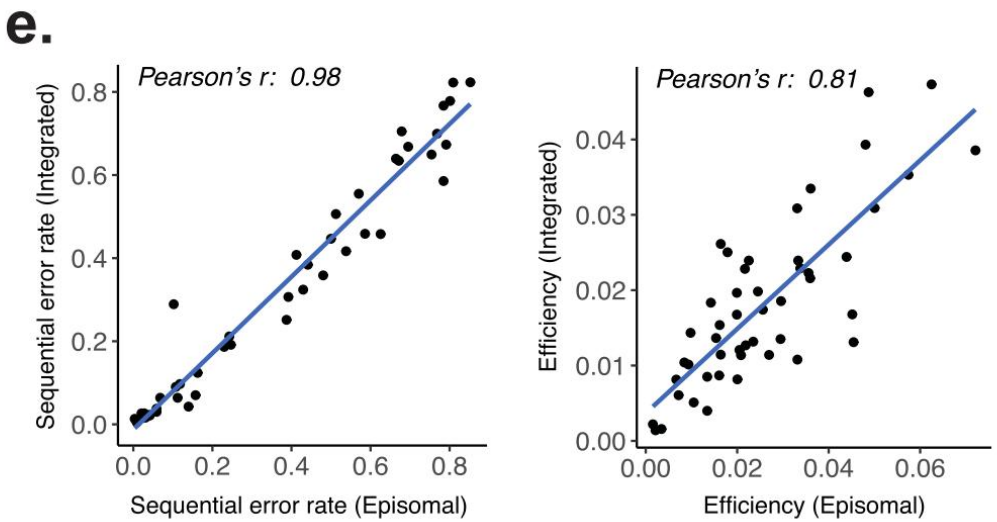
PD3	AGCTGCTCACCACGACG CCAGGGCTGCGGGG	(Base-8)
TAPE-43	GGACCAGGGCCACGACG CCAGGGCCACGACG	(Design-1)
TAPE-44	GGACCAGGGCTGCGGGG CCAGGGCTGCGGGG	(Design-2)
TAPE-45	GGATGCTGGCCACGACG TGCTGGCCACGACG	(Design-3)
TAPE-46	GTCCCAGGGCCACGACG CCAGGGCCACGACG	(Design-4)
TAPE-47	GACCAGGGACCACGACG CCAGGGACCACGACG	(Design-5)
TAPE-48	GGACCCAGGGCACGACG CCAGGGCACGACG	(Design-6)

Screening additional DNA Tape sequences









Transfection programmes for 16 sequential epochs

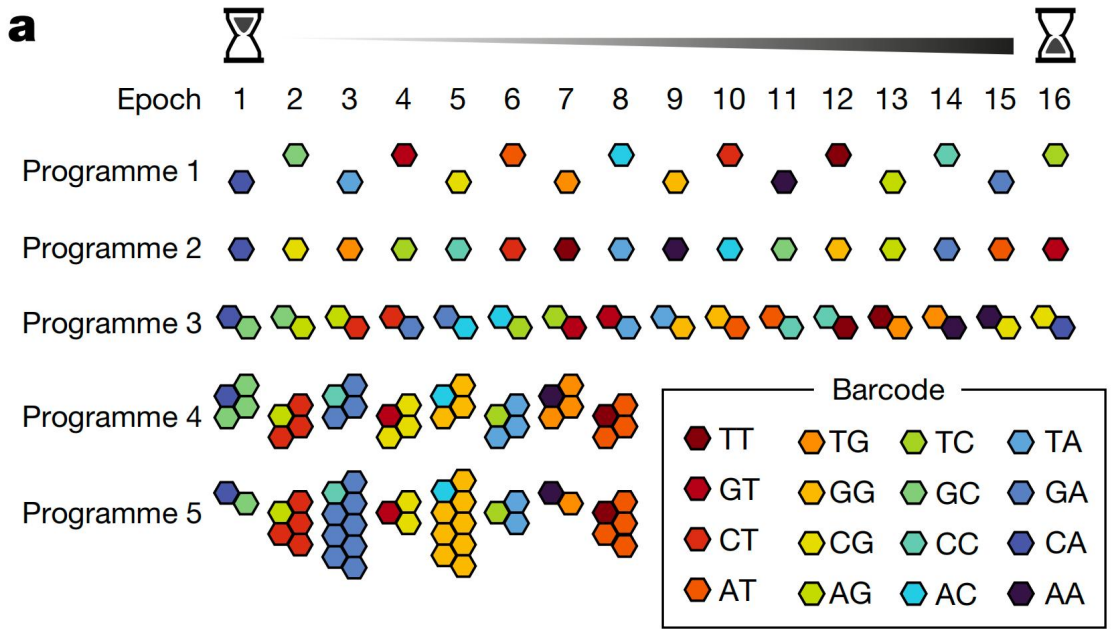
Programme	epochs	barcode/epoch	ratio	Others
Programme 1	16	1	–	MAX edit distances
Programme 2	16	1	–	MIN edit distances
Programme 3	16	2	1:1	1BC shared between adjacent
Programme 4	8	2	1:3	constant ratio
Programme 5	8	2	1:1, 1:2, 1:4, 1:8	varying ratios

cells: HEK293T 5×TAPE-1

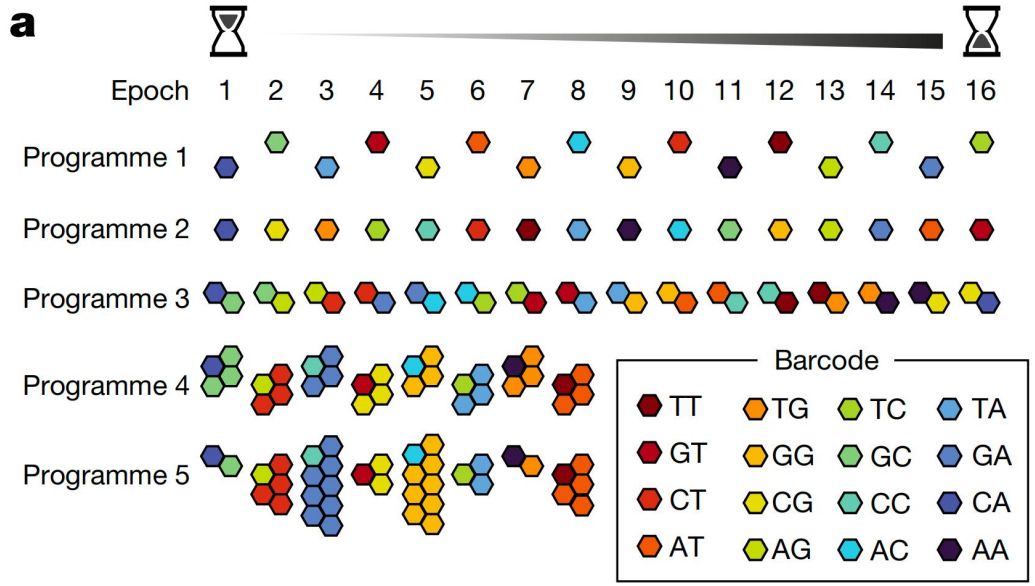
Transfection: pCMV-PE2-P2A-GFP + 16x pegRNA

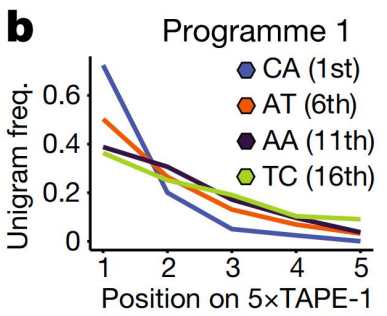
Time: 3d/epoch

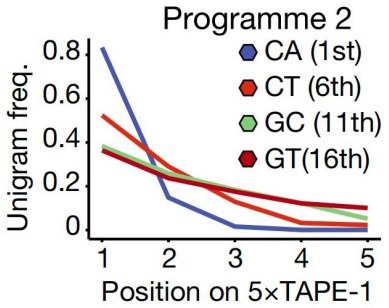
readout:PCR amplify DNA Tape + sequencing

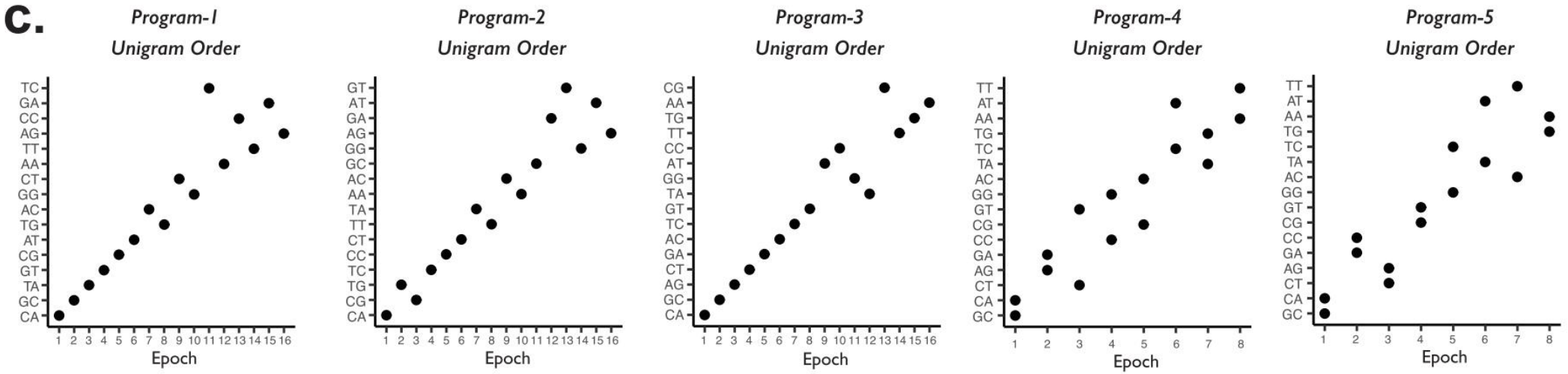


site 1 unigram frequency is not sufficient to decode the order

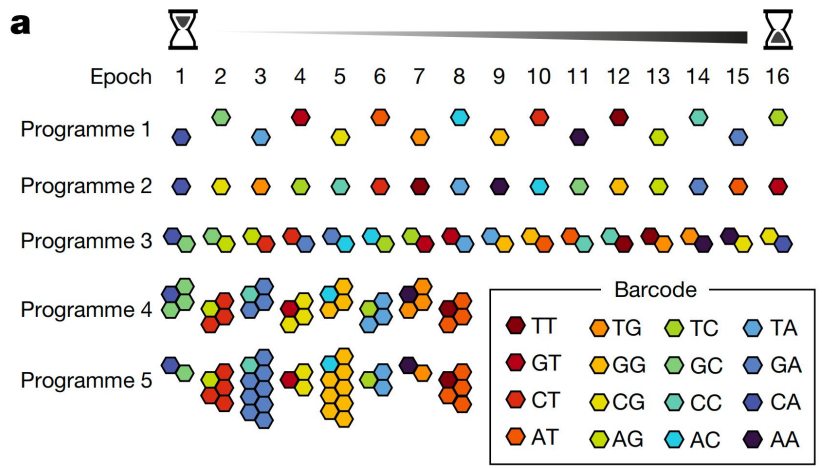


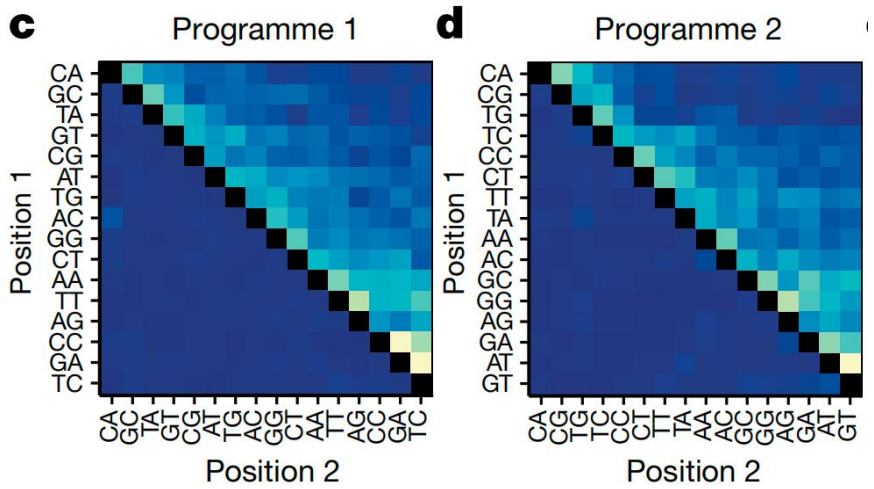




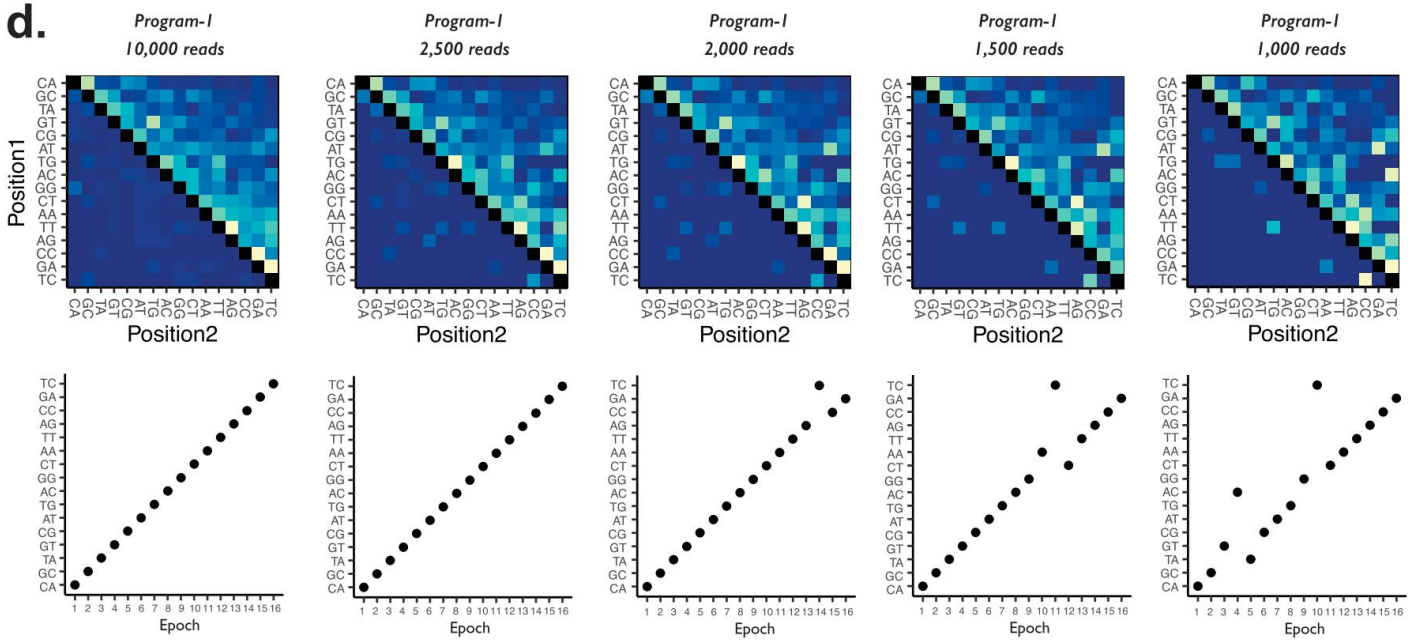


bigram matrix restores the sequence of events

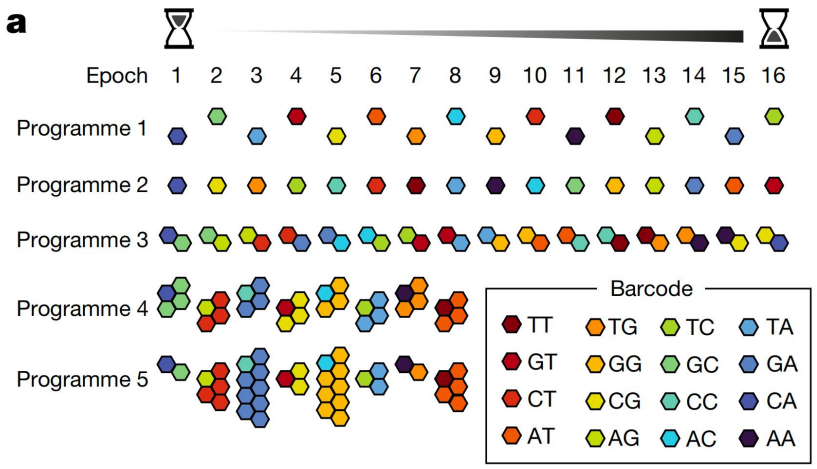


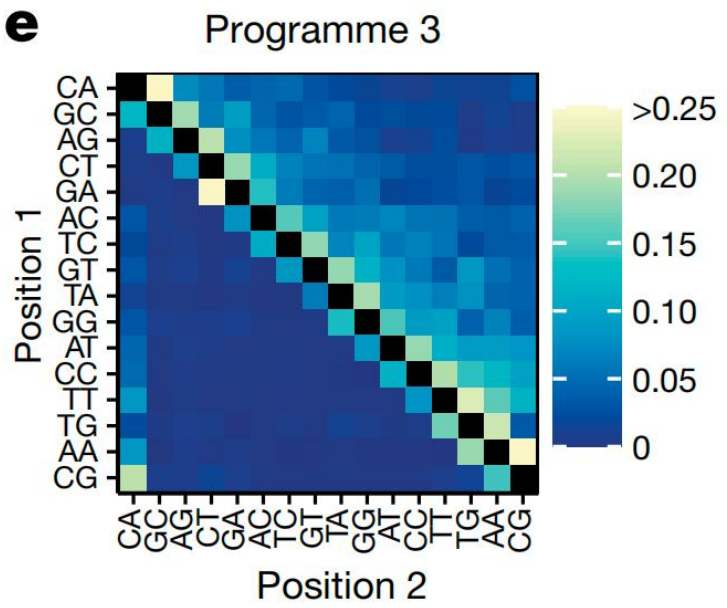


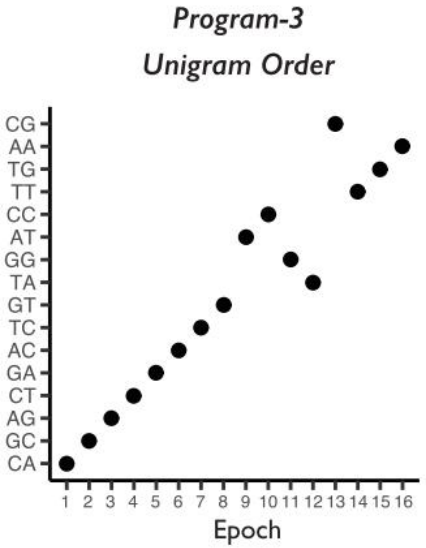
The decoding algorithm is a **bubble-sort-like** adjacent-swap heuristic based on pairwise bigram preferences.

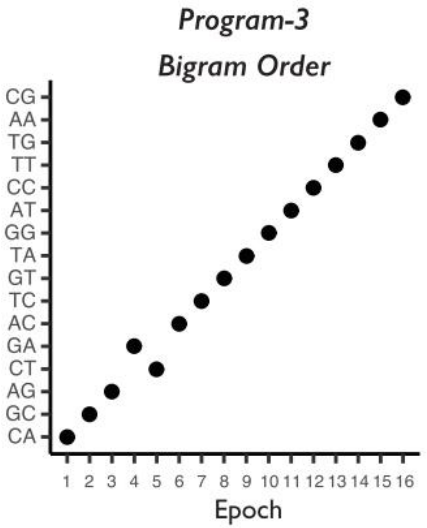


overlapping signals

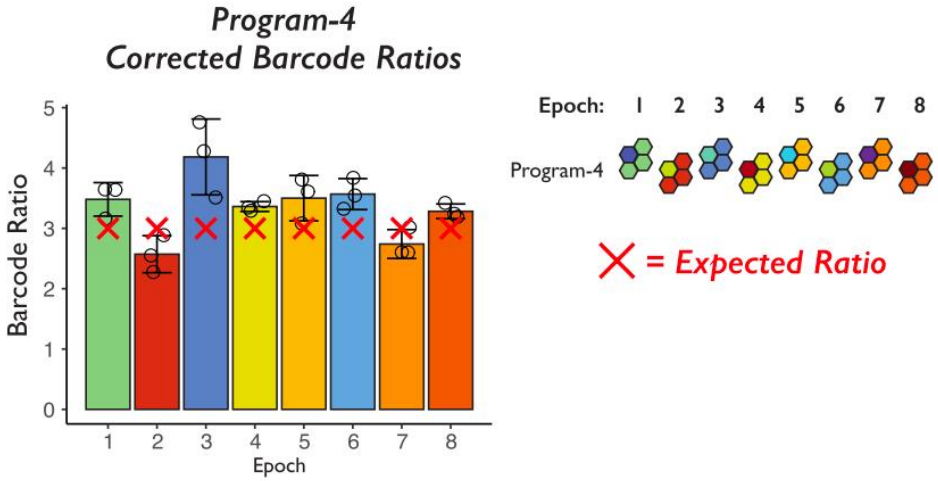


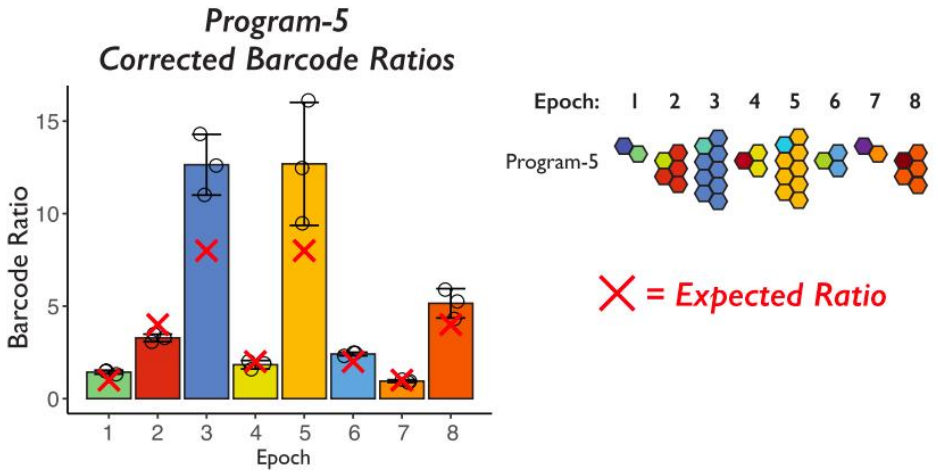


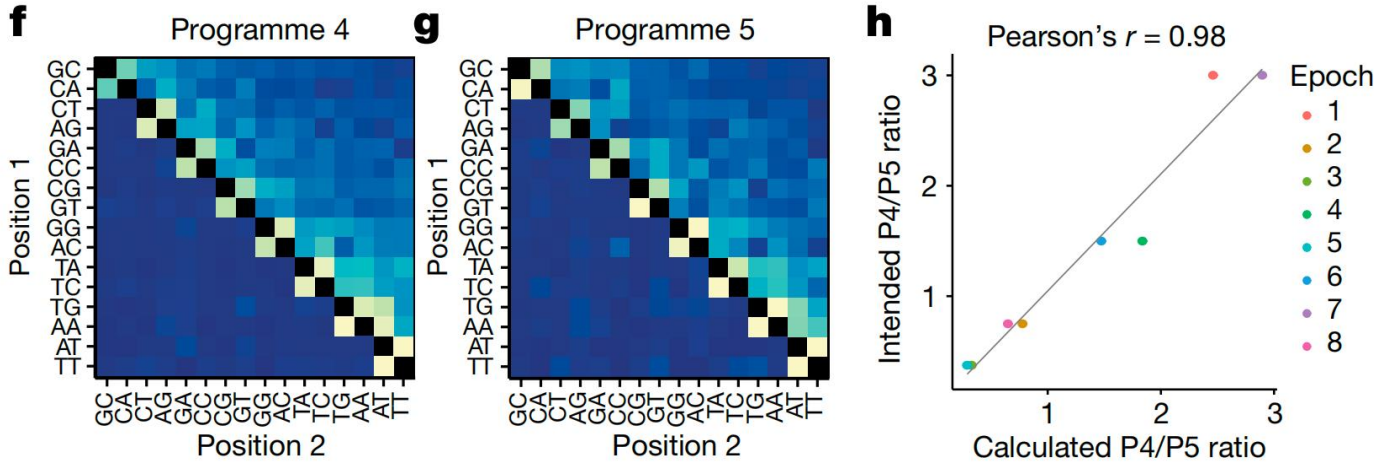


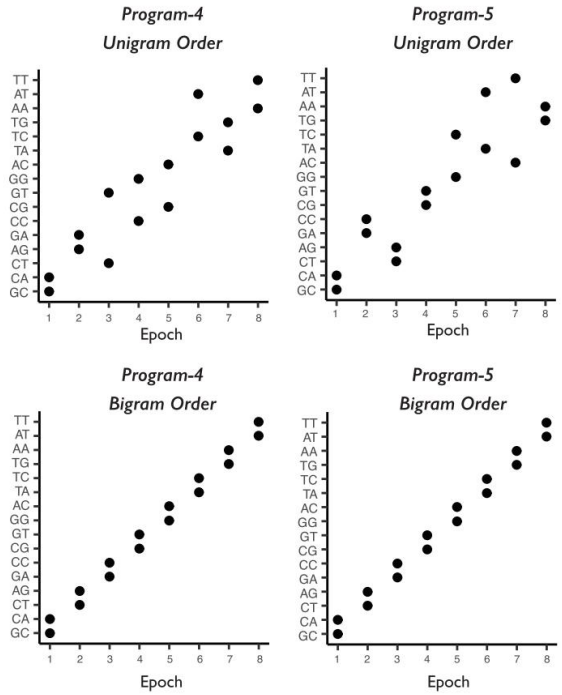


relative ratio



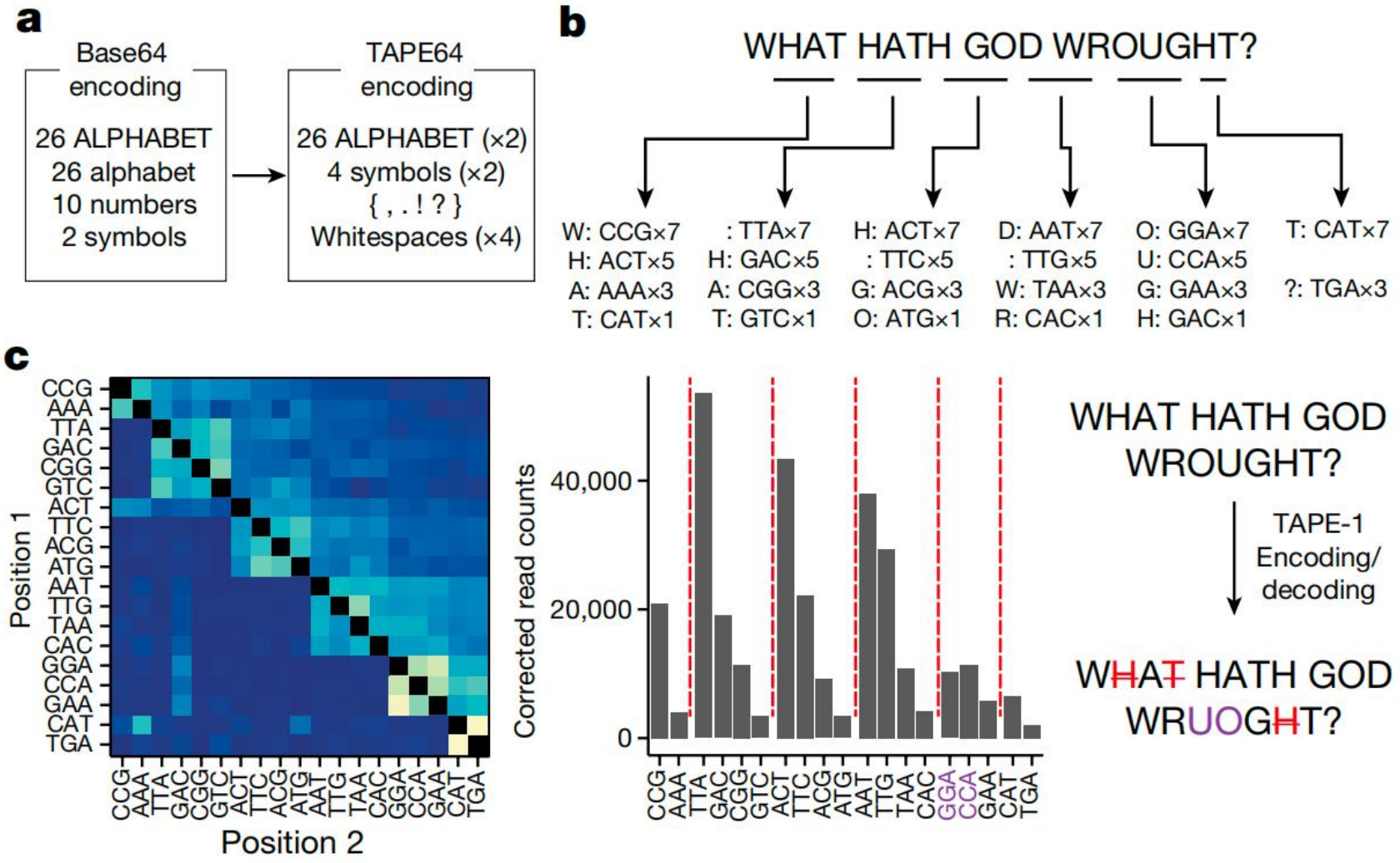




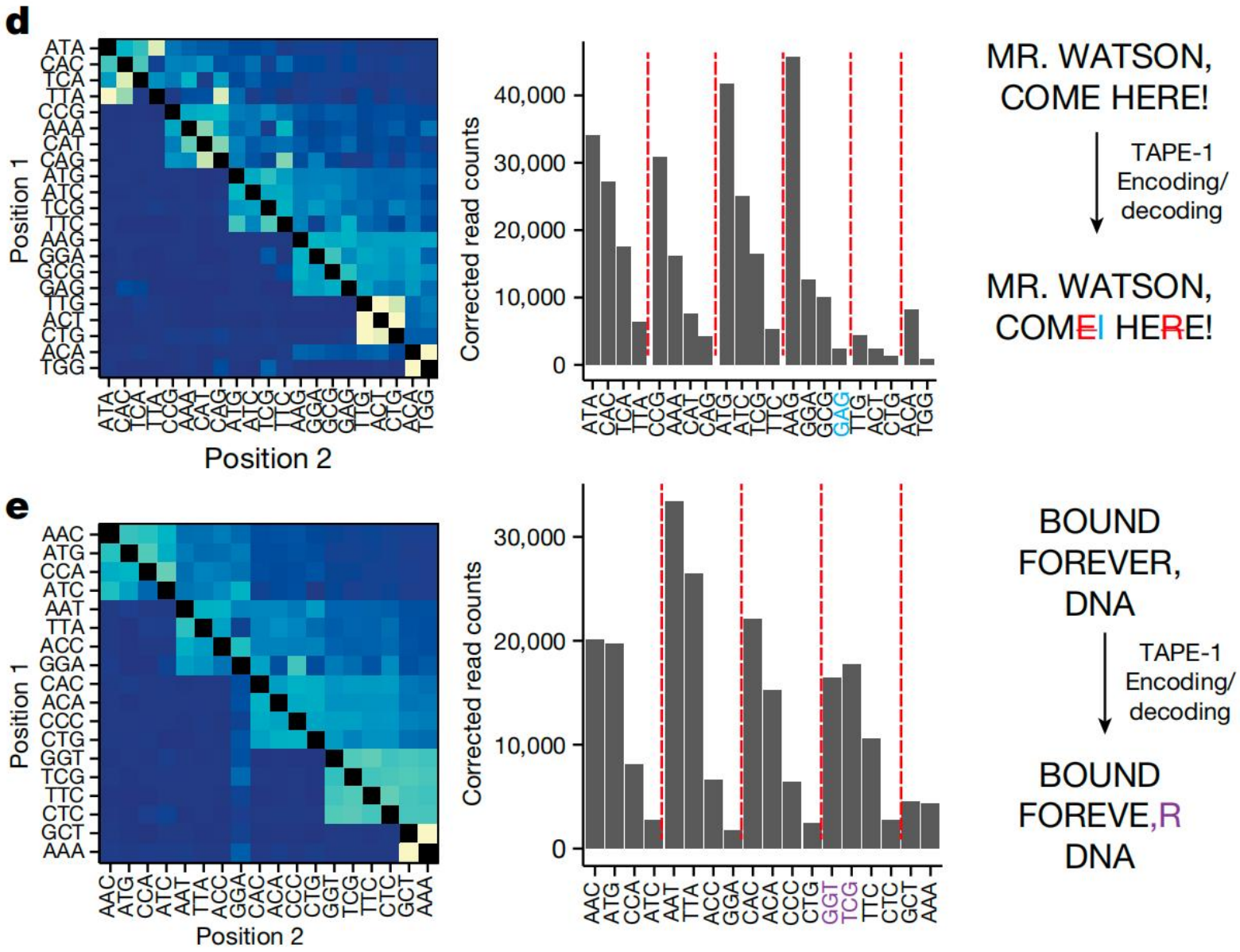


DNA Typewriter can record, recover and decode complex event histories including the order, overlap and relative strength of signals.

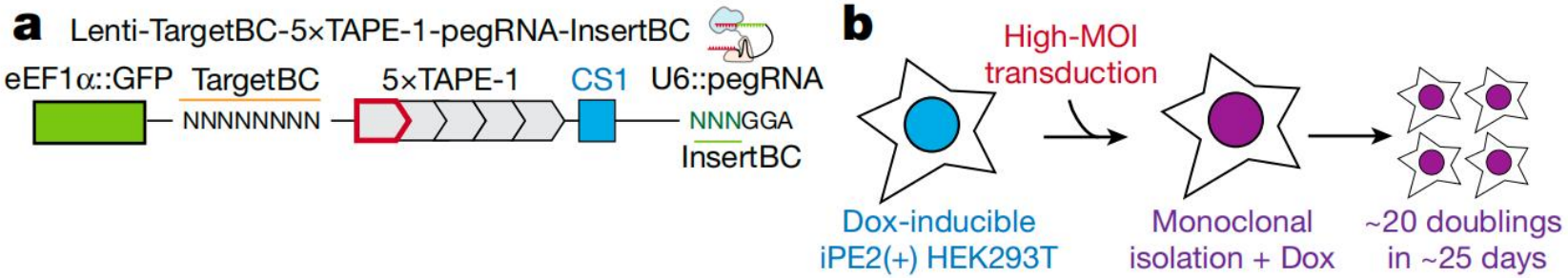
Recording and recovering short texts

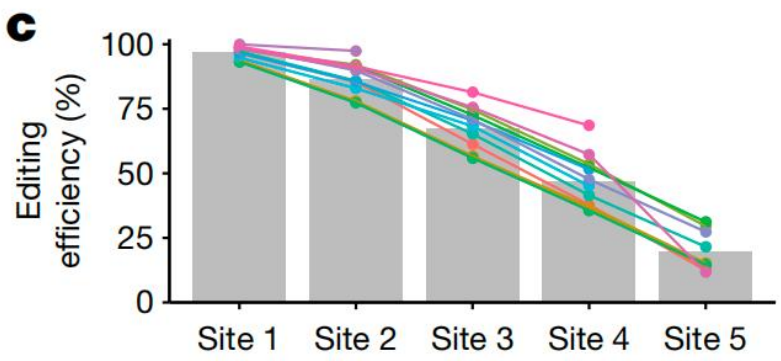


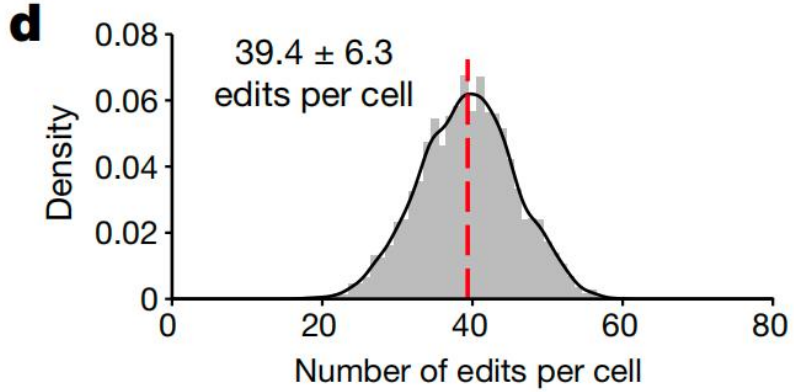
Recording and recovering short texts

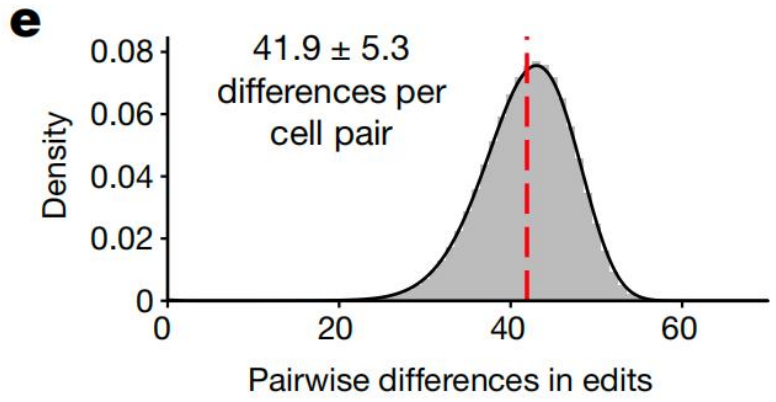


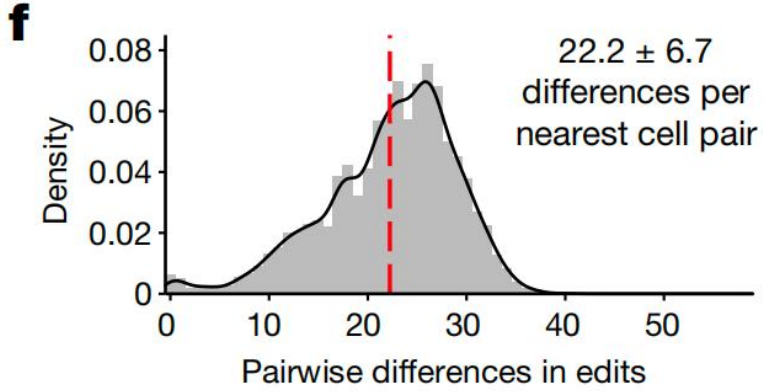
Ordered recording of cell lineage











Cell：Dox-inducible PE2 HEK293T

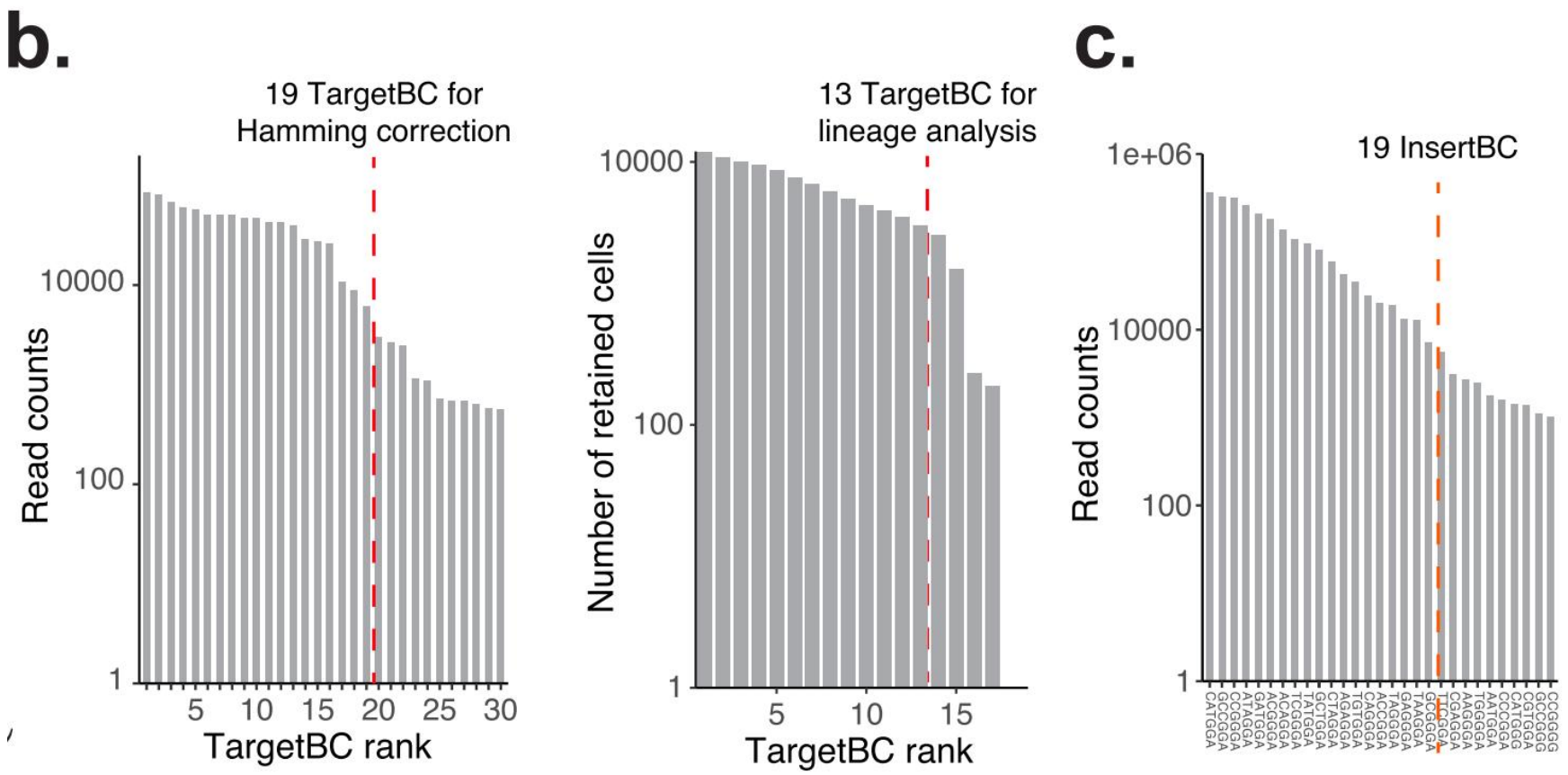
Vector：lentiviral construct

Core element：

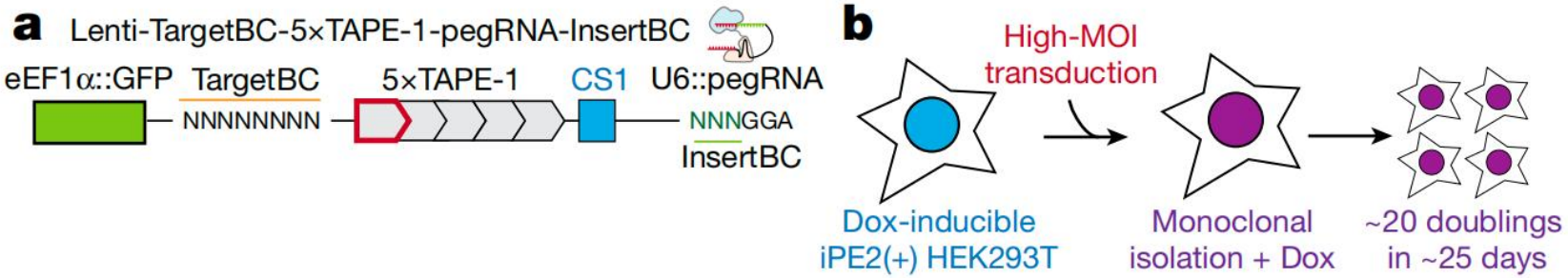
- 5×TAPE-1
- 5′ random 8-bp **TargetBC**
- TargetBC-5×TAPE-1 transcript cassette，with scRNA-seq capture sequence
- constitutive pegRNA expression cassette
- pegRNA with NNNGGA, NNN =**InsertBC**，GGA = key

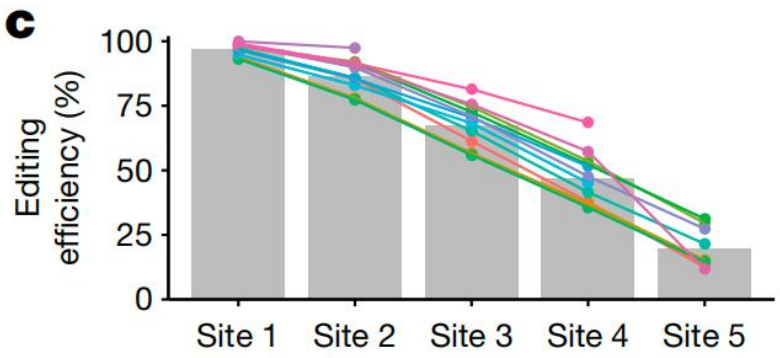
scRNA-seq: **CellBC**

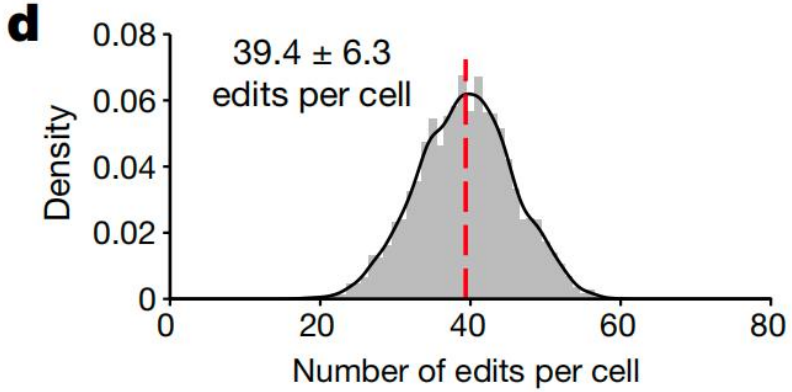
Ordered recording of cell lineage

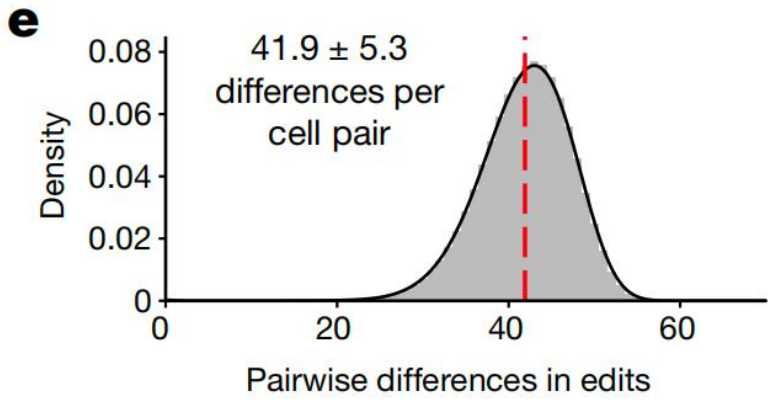


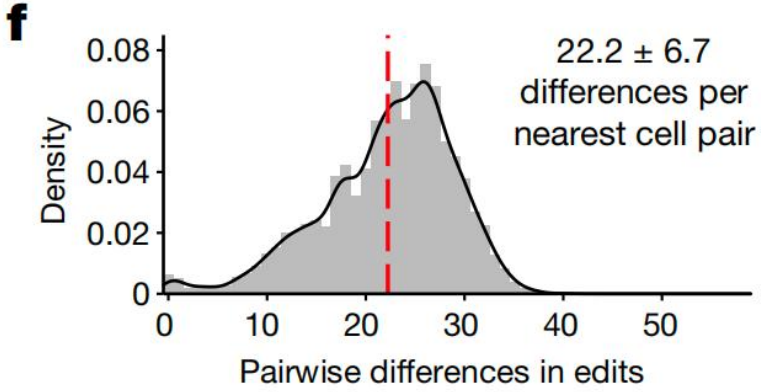
Ordered recording of cell lineage











Cell：Dox-inducible PE2 HEK293T

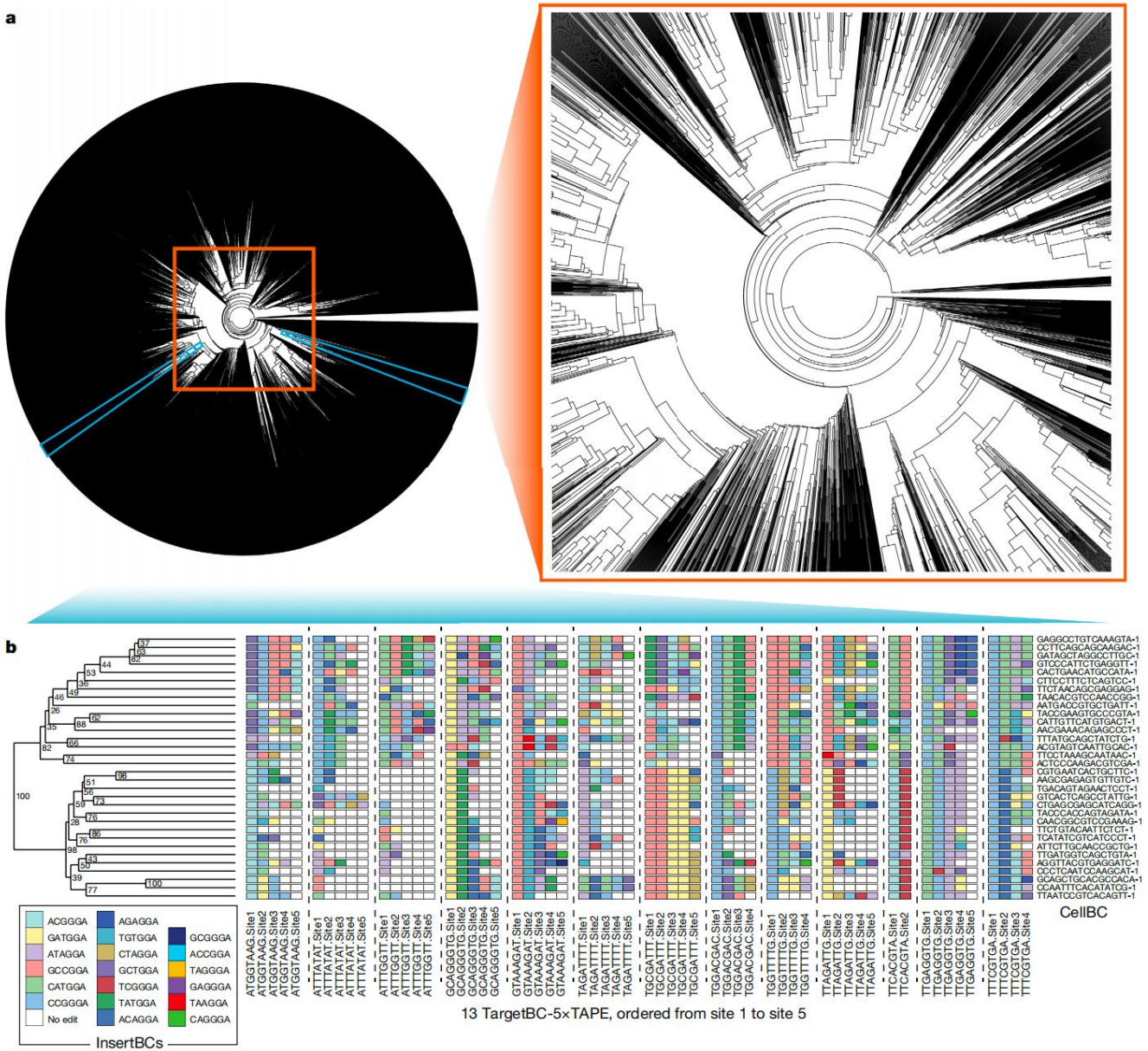
Vector：lentiviral construct

Core element：

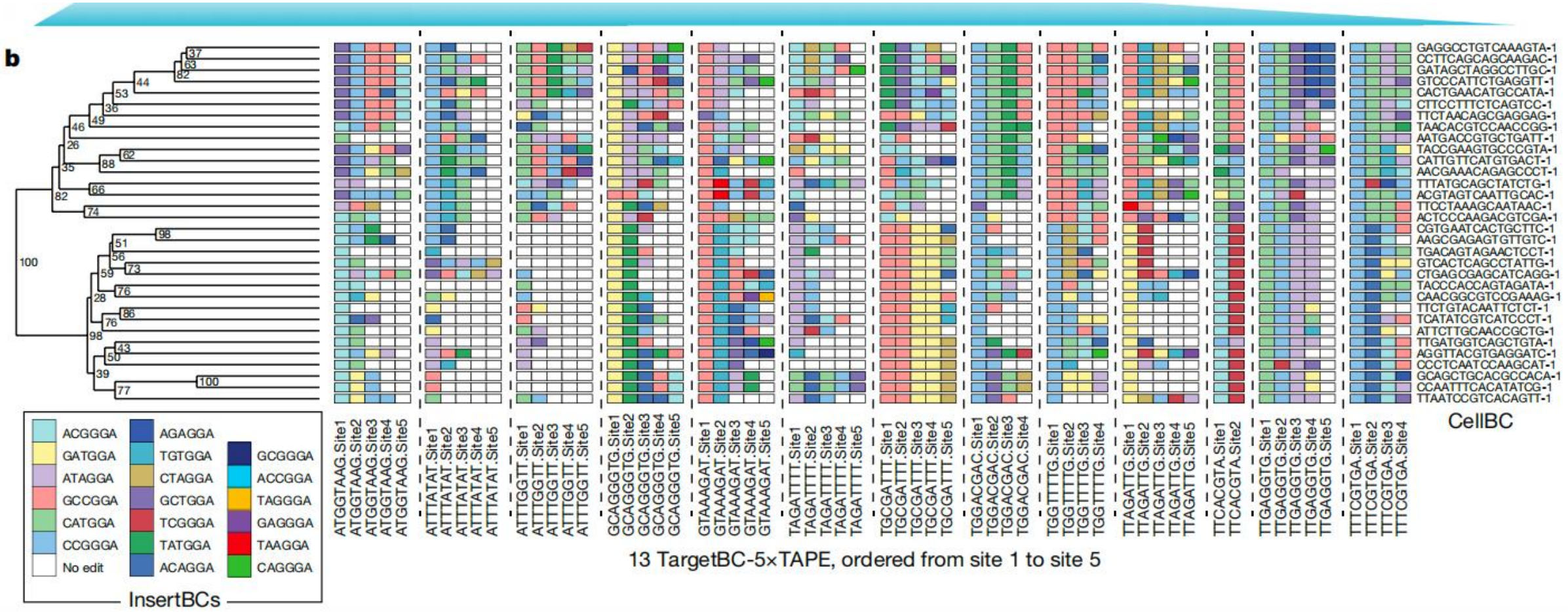
- 5×TAPE-1
- 5′ random 8-bp **TargetBC**
- TargetBC-5×TAPE-1 transcript cassette，with scRNA-seq capture sequence
- constitutive pegRNA expression cassette
- pegRNA with NNNGGA, NNN =**InsertBC**，GGA = key

scRNA-seq: **CellBC**

Reconstruction of a monophyletic cell lineage tree



Reconstruction of a monophyletic cell lineage tree



Limitations

- most of the current proof-of-concept symbols are artificial transfection or random expression of pegRNA, rather than endogenous biological events themselves. True biological recording requires coupling of NF-κB, Wnt, small molecule response elements, or other signals to pegRNA production.
- the prime editing rate is on the order of days, so the DNA Typewriter is more suitable for the days/weeks scale process and is not suitable for the fast signal of minutes/hours.
- issues of delivery, cell-type differences, tape recovery, spatial compatibility, and editor perturbation still need to be addressed for in vivo complex tissue applications. In the future, a molecular flight recorder locus can be established to record lineage and biological signals, but this is still an engineering goal

Connecting biological signals to DNA recording

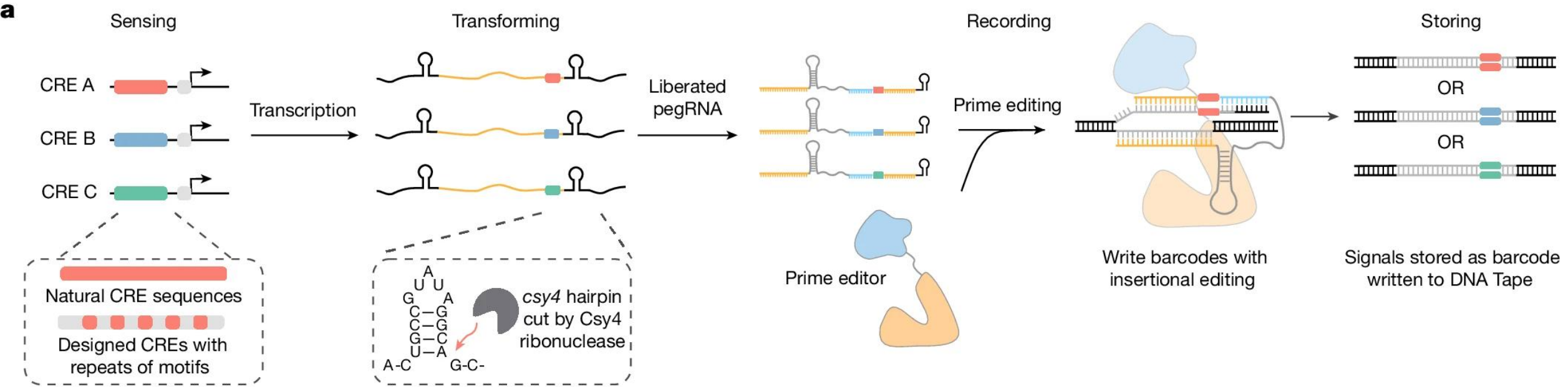
ENGRAM

Article | Open access | Published: 17 July 2024

Symbolic recording of signalling and *cis*-regulatory element activity to DNA

Wei Chen ✉, Junhong Choi, Xiaoyi Li, Jenny F. Nathans, Beth Martin, Wei Yang, Nobuhiko Hamazaki, Chengxiang Qiu, Jean-Benoît Lalanne, Samuel Regalado, Haedong Kim, Vikram Agarwal, Eva Nichols, Anh Leith, Choli Lee & Jay Shendure ✉

Nature **632**, 1073–1081 (2024) | Cite this article



Reference

1. Chen W, Choi J, Li X, et al. Symbolic recording of signalling and cis-regulatory element activity to DNA. Nature. 2024;632(8027):1073-1081. doi:10.1038/s41586-024-07706-4
2. Chen PJ, Hussmann JA, Yan J, et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell. 2021;184(22):5635-5652.e29. doi:10.1016/j.cell.2021.09.018
3. Doman JL, Sousa AA, Randolph PB, Chen PJ, Liu DR. Designing and executing prime editing experiments in mammalian cells. Nat Protoc. 2022;17(11):2431-2468. doi:10.1038/s41596-022-00724-4
4. Chen PJ, Liu DR. Prime editing for precise and highly versatile genome manipulation. Nat Rev Genet. 2023;24(3):161-177. doi:10.1038/s41576-022-00541-1
5. Street K, Risso D, Fletcher RB, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018;19(1):477. doi:10.1186/s12864-018-4772-0
6. Mitchell S, Roy K, Zangle TA, Hoffmann A. Nongenetic origins of cell-to-cell variability in B lymphocyte proliferation. Proc Natl Acad Sci U S A. 2018;115(12):E2888-E2897. doi:10.1073/pnas.1715639115
7. Sheth RU, Yim SS, Wu FL, Wang HH. Multiplex recording of cellular events over time on CRISPR biological tape. Science. 2017;358(6369):1457-1461. doi:10.1126/science.aao0958
8. Frieda KL, Linton JM, Hormoz S, et al. Synthetic recording and in situ readout of lineage information in single cells. Nature. 2017;541(7635):107-111. doi:10.1038/nature20777
9. McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science. 2016;353(6298):aaf7907. doi:10.1126/science.aaf7907
10. Farzadfard F, Lu TK. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. Science. 2014;346(6211):1256272. doi:10.1126/science.1256272
11. Synthetic recombinase-based state machines in living cells. Science. Accessed May 29, 2026. https://www.science.org/doi/10.1126/science.aad8559
12. Choi J, Chen W, Minkina A, et al. A time-resolved, multi-symbol molecular recorder via sequential genome editing. Nature. 2022;608(7921):98-107. doi:10.1038/s41586-022-04922-8

北京大学
PEKING UNIVERSITY

生命科学学院
School of Life Sciences

Thanks for your Listening!
Any Questions?

UHPB Journal Club

Jiaxi Liu (刘家希)

School of Life Sciences

2026.5.30