Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

DNA event recorders send past information of cells to the time of observation

Soh Ishiguro[1,2,3], Hideto Mori[1,2,3] and Nozomu Yachie[1,2,3,4,5]

While current omics and single cell technologies have enabled measurements of high-resolution molecular snapshots of cells at a large scale, these technologies all require destruction of samples and prevent us from analyzing dynamic changes in molecular profiles, phenotypes, and behaviors of individual cells in a complex system. One possible direction to overcome this issue is the development of a cell-embedded 'event recorder' system, whereby molecular and phenotypic information of a cell (s) can be obtained at the time of observation with their past event information stored in 'heritable polymers' of the same cell. This concept has been demonstrated by many synthetic cellular circuits that monitor and transmit a certain set of environmental and intracellular signals into DNA, and have now been further accelerated by recent CRISPR-related technologies. Notably, the discovery of the RT-Cas1–Cas2 system, which acquires sequences of cellular transcripts into a specific host genomic region, has enabled recording of a broader range of molecular profile histories in the DNA tapes of cells, to understand the dynamics of complex biological processes that cannot be addressed by current technologies.

Addresses
[1] Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo 153-8904, Japan
[2] Institute for Advanced Biosciences, Keio University, Tsuruoka 997-0035, Japan
[3] Graduate School of Media and Governance, Keio University, Fujisawa 252-0882, Japan
[4] Department of Biological Sciences, School of Science, The University of Tokyo, Tokyo 113-0033, Japan
[5] PRESTO, Japan Science and Technology Agency (JST), Tokyo 153-8904, Japan

Corresponding author:
Yachie, Nozomu (yachie@synbiol.rcast.u-tokyo.ac.jp)

Current Opinion in Chemical Biology 2019, **52**:54–62

This review comes from a themed issue on **Synthetic biology**

Edited by **Hirohide Saito** and **Yohei Yokobayashi**

For a complete overview see the Issue and the Editorial

Available online 11th June 2019

https://doi.org/10.1016/j.cbpa.2019.05.009

Introduction

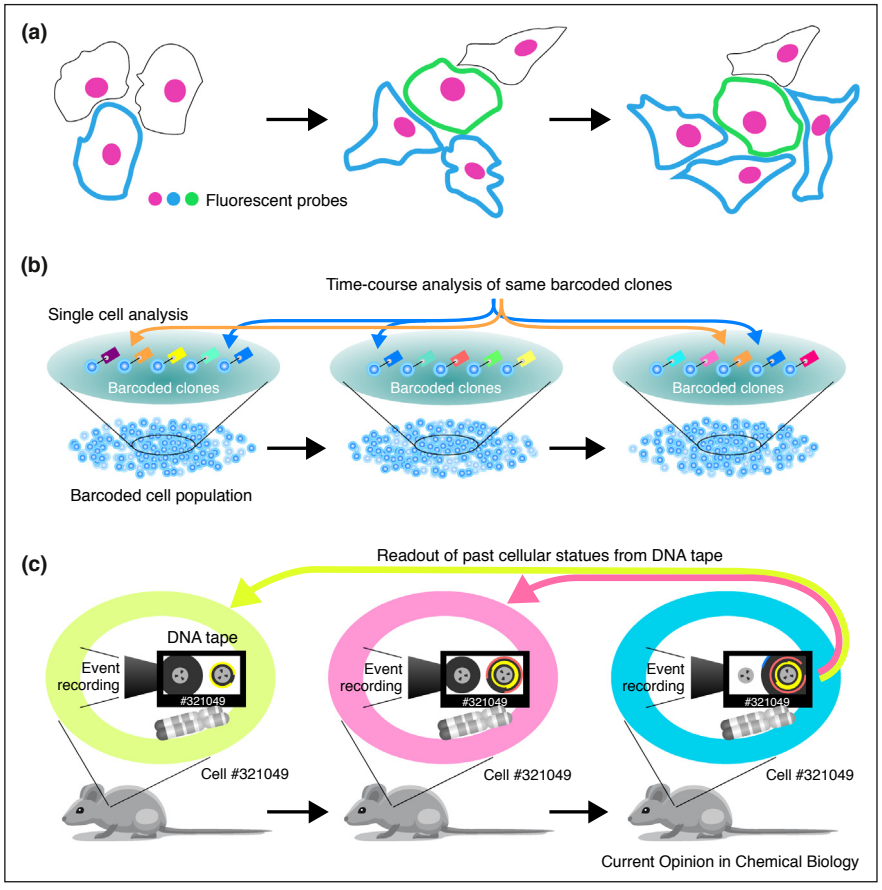
Massively parallel sequencing and mass spectrometry technologies obtain various types of cellular information, such as chromatin structures and transcriptomic, proteomic and metabolomic profiles, as well as post-transcriptional and post-translational modifications at a large scale. While these technologies have uncovered how different biological systems dynamically change their 'average' molecular landscapes, recent single cell technologies have identified many unknown but distinct types of cells in multicellular organisms. Cells with different molecular phenotypes dynamically emerge and orchestrate in the progression and homeostasis of tissues, organs, and individuals. Therefore, it is extremely important to monitor how different single cells (or clones) in such systems change their molecular profiles and interact with others. However, high content molecular analyses of cells using current omics technologies require destruction of samples and preclude performing time course analysis of the same biological objects. For example, it is extremely difficult to analyze how a given blood cell clone of a mouse changes its molecular profiles throughout the hematopoietic differentiation process.

Live cell imaging is capable of analyzing spatiotemporal dynamics of molecules and cells with fluorescent proteins or probes, but only for a limited number of objects, which can be observed by microscopy (Figure 1a). Although single cell technologies coupled with molecular barcoding of cells may be a potential solution to obtain transcriptomic and (epi-) genomic information of the same cell clones (or 'close relative' cells) across time course samples, the scalabilities of current single cell technologies are up to the order of 10^5 cells per sample, limiting the detection sensitivity to capture the same barcoded cells across different samples (Figure 1b). Because of its high information capacity, physical density, durability, and ease of duplication, DNA has been proposed to store artificial information [1], and several high capacity DNA data storage methods have been accomplished [2–5]. In synthetic biology, there is the concept of recording environmental and intracellular events in DNA that replicates with cell division such that the previous event histories of cells can be obtained from the DNA media with other high content omics information at the time of observation (Figure 1c). Although none of the current synthetic circuits have yet succeeded to record high content, time course event information in cellular DNA to address practical biological issues, the methodologies harnessing DNA recombination and CRISPR genome editing have been rapidly heading toward this goal.

Synthetic event recording using DNA recombination

To transmit extracellular event signals to permanent alterations of a DNA sequence, site-specific DNA

Figure 1



Different types of methods to analyze dynamic changes in molecular profiles of the same cells or clones. **(a)** Live cell imaging enables direct observation of progressing cell populations by labeling target molecules with fluorescent markers, but only for a limited number of objects that can be observed under a microscope. **(b)** Single cell transcriptome sequencing (scRNA-seq or single cell RNA sequencing) of barcoded cells enables analysis of dynamic changes in transcriptome profiles. Cells in an initial population are uniquely tagged with DNA barcodes that can be identified by an scRNA-seq pipeline. Clonal dynamics of transcriptomes can be observed for clones which barcodes can be observed across different time course scRNA-seq data. **(c)** The DNA event recorder concept. Each cell in a complex sample (e.g. a whole mouse body) involves a DNA event recorder system that sequentially records cellular and environmental event information in a DNA tape. The past event history information of cells can be derived from their DNA tapes with their high content omics profiles at the time of observation.

recombinases, such as Cre, Flp, PhiC31, and Bxb1, have been widely used with signal-inducible gene expression promoters. In the simplest case, upon exposure to a specific signal, a site-specific DNA recombinase is induced in cells to delete or flip a given DNA sequence sandwiched by two recombination sites, which can be then readout, for example, by sequencing or as a restoration of fluorescent gene expression. Extensive efforts have been made to screen orthogonal sets of recombinase-target site pairs, and to build more complex signal-dependent circuits by concatenating multiple different recombination sites and various promoters and terminators [6–10,11••,12]. Examples include circuits that can count the number of extracellular stimuli [6], circuits that exhibit specific responses depending on the order of different stimuli [7], and various logic gate repertoires, such as AND and OR gates as well as NAND, NOR, XNOR, and others [8,9]. Another study has also reported establishment of stringent band-path filters for precision computation using multiple signal inputs [10]. While most of these circuits have been demonstrated in *Escherichia coli* cells, BLADE (Boolean logic and arithmetic through DNA excision) [11••] enables flexible construction of various circuits in mammalian cells, which has succeeded to construct 109 two-input, two-output circuits in human embryonic kidney and Jurkat T cells; three-input, two-output circuits; a six-input, one-output circuit; and circuits that incorporate CRISPR–Cas9 to regulate endogenous genes.

In contrast to the fast-reacting DNA recombinase-based methods with limited scalability owing to the number of available recombinase enzymes as 'DNA writer' modules, another scalable and slow-reacting method, SCRIBE

(synthetic cellular recorders integrating biological event), has been developed using bacterial retron and recombineering technologies, which might be suitable to record long-term histories of cells [13•]. In this system, an engineered retron system is employed, whereby a chemical stimulus causes an arbitrary sequence harbored by common moiety sequences that form a specific RNA secondary structure to be transcribed and reverse transcribed by an RT enzyme. Thus produced single-stranded DNA (ssDNA) is then recombined with a target DNA by homologous recombination enhanced by Beta recombinase (from bacteriophage λ). SCRIBE has been demonstrated to successfully trace long-term exposure patterns of multiple compounds in *E. coli*. The scalability of SCRIBE is high because orthogonal retron-target DNA pairs can be freely designed by changing their donor ssDNA sequences and target sites. Although the retron system is originated from a bacterial species, it is known to be also active in yeast [14] and mammalian cells [15]. Moreover, by harnessing CRISPR–Cas9 and the yeast retron technology, CRISPEY (Cas9 retron precise parallel editing via homology) has been established as a highly efficient single base pair genome editing method in yeast [16], showing the potential of SCRIBE-type event recording in other eukaryotic cells.

CRISPR–Cas9 as a DNA writer module

Recently, CRISPR-related systems have started to be employed for the DNA event recording concept. The CRISPR–Cas system is a prokaryotic adaptive immune system in which DNA sequences of foreign plasmids and phages are immunized into a genomic CRISPR (clustered regularly interspaced short palindromic repeats) locus. Upon second exposure to the same invaders, crRNAs (CRISPR RNAs) transcribed from the immunized CRISPR locus recruit a Cas (CRISPR-associated) protein(s) to the target DNA molecules via sequence complementarity and digest them by the DNA cleavage activity of Cas protein(s). Among the different types of CRISPR–Cas immunity response systems, Cas9 from the type II-A CRISPR system has been widely used for genome editing [17,18]. While a trans-activating crRNA (tracrRNA) is also required to form a secondary structure with crRNA to activate Cas9 at the target locus, a synthetic crRNA-tracrRNA fusion, hereafter referred to as guide RNA (gRNA) [19], is commonly used in practical genome editing. In summary, CRISPR–Cas9 genome editing is a simple two component system in which Cas9 with a double-stranded DNA cleavage activity is recruited by gRNA to a target DNA region containing a protospacer adjacent motif (PAM) in its 3′-end. PAM is necessary for target recognition by Cas9, conferring the distinction of the foreign target DNA and host crRNA-encoding region both having the same sequence. Because DNA double-strand breaks induce various DNA repair pathways, such as non-homologous end joining (NHEJ) and homologous recombination (HR), CRISPR–Cas9 has been used to enhance gene knockout and knock-in efficiencies in model and non-model organisms.

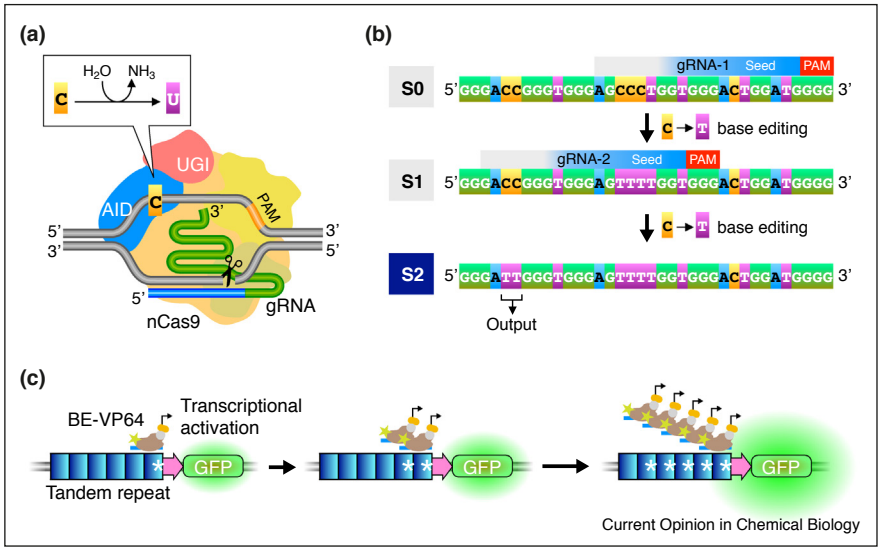
Nuclease-deficient Cas9 (dCas9) and nickase Cas9 (nCas9) mutants can be tethered to various effector protein modules (e.g. transcription activation complexes and epigenetic modifiers) to enable synthetic control of target gene expression, epigenetic modifications, and chromatin structures [20,21]. Unlike previous synthetic biology methods, this strategy allows flexible construction of various synthetic circuits with minimal genetic engineering effort in various organisms. Furthermore, base editing technologies have been developed by fusing Cas9 mutants to nucleoside-modifying enzymes, which enables direct substitution of target nucleotides [22–24]. Among such technologies, fusing activation-induced cytidine deaminase (AID) and uracil glycosylase inhibitor (UGI) to nCas9 (or dCas9), highly efficient cytidine base editors for targeted C:G-to-T:A editing has been developed and improved by many groups [22,23,25–27] (Figure 2a). In such targeted single base editing, gRNA recruits AID and UGI to a target locus via nCas9 (or dCas9). AID converts cytidines in the target region to uridines by deamination; and UGI blocks the glycosylation-initiated base excision repair pathway and enhances the conversion of the uracil bases to thymine bases through DNA replication [28]. While various applications of base editors have been proposed, such as *in vivo* correction and creation of disease mutations [29–31], several studies have also used them to construct DNA event recorder circuits in cells.

Molecular event recording using CRISPR–Cas9

Among the DNA event recording circuits based on CRISPR–Cas9, mSCRIBE (mammalian synthetic cellular recorder integrating biological events) employed the concept of self-targeting gRNA (stgRNA; see another article in this issue [32]) [33] and successfully recorded combinatorial chemical conditions of human cell culture, in which two stgRNAs were regulated under the control of doxycycline-inducible and IPTG-inducible promoters. The mSCRIBE concept was also used to establish an inflammation-recording cell line using an NF-kB-inducible Cas9 expression cassette and record lipopolysaccharide-induced inflammation levels of individual mice in the synthetic DNA of the cells.

CAMERA (CRISPR-mediated analog multi event recording apparatus) has been also proposed to construct synthetic circuits that can transmit order information of different environmental signals into DNA [34••]. In CRIPSR–Cas9, the gRNA/target DNA heteroduplex tends not to require perfect base pairing in the PAM distal region, whereas the PAM-proximal 'seed' sequence is necessary for the efficacy of Cas9 [35]. In CAMERA, by harnessing this feature and using a C:G-to-T:A base

Figure 2



Synthetic DNA circuits using CRISPR base editors. **(a)** Conceptual diagram of the cytidine base editor. **(b)** An example circuit demonstrated in the CAMERA study. The S1-to-S2 state transition by base editing using gRNA-2 can be induced only after the S0-to-S1 transition by gRNA-1 generates the seed sequence for gRNA-2. **(c)** An example circuit demonstrated in the DOMINO study. The iterative domino-like editing of a repetitive sequence by a C:G-to-T:A base editor tethered to a transcription factor increases the fluorescent protein expression level over time.

editor, a circuit was established in *E. coli* cells, in which data recording by a secondary gRNA could not be achieved until the first gRNA made a specific C:G-to-T:A edit to create the seed sequence for the secondary gRNA (Figure 2b). CAMERA was also used to record the intracellular status of mammalian cells. For example, using the TCF/LEF promoter linked to the base editor, lithium chloride-dependent recording of Wnt signaling was achieved.

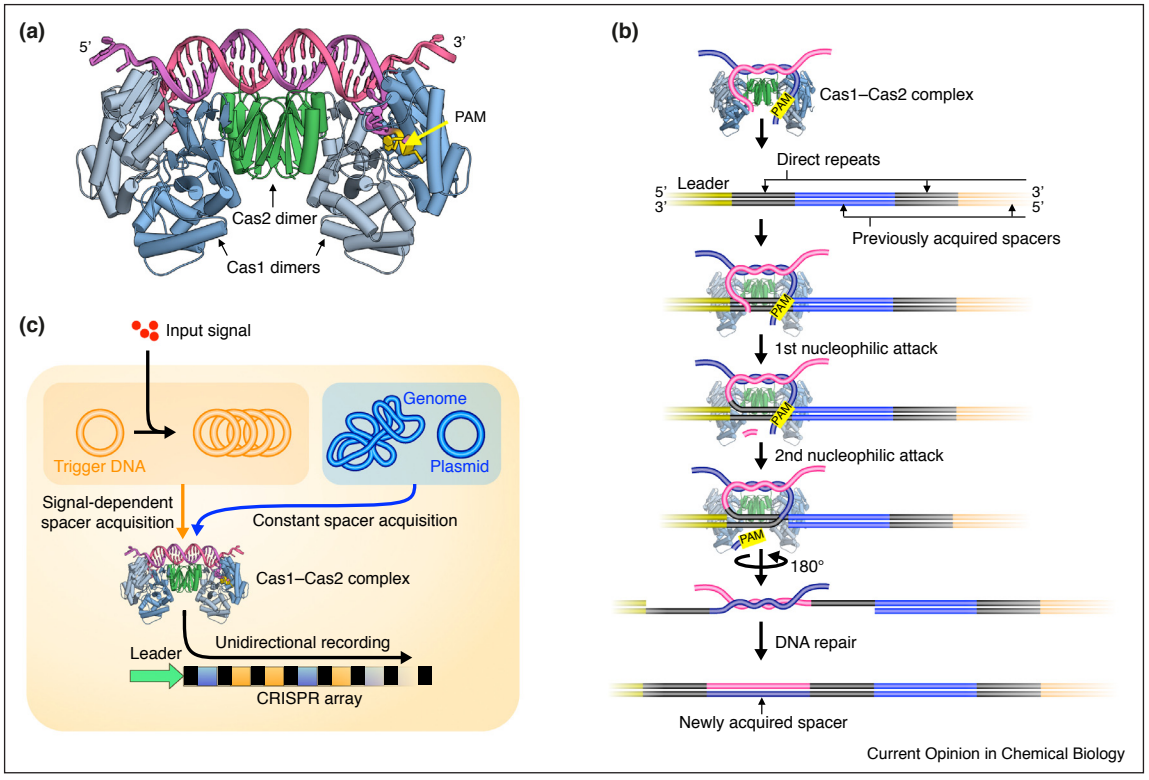
DOMINO (DNA-based ordered memory and iteration network operator) employs the same concept used in CAMERA, but achieves more complex circuits, in which, for example, an IPTG-inducible gRNA sequentially restores the seed sequences of its three tandem overlapping target sites so that the fourth editing by an arabinose-inducible gRNA can be made in accordance with the concentration of IPTG and duration of IPTG exposure [36••]. In this circuit, the sequential rounds of base editing increase the copy number of targetable sequences for the same gRNA. By fusing the transcription factor VP64 to the base editor, a similar DOMINO circuit was also used to demonstrate an online state reporting circuit in mammalian cells, where the GFP expression level was increased along with time progression (Figure 2c).

CRISPR–Cas1–Cas2 as another class of DNA event recorder systems

While CRISPR immunity systems are evolutionary diverged, one of which is Cas9 from the type II-A CRISPR system, the upstream CRISPR immunization process is relatively conserved across species, in which the Cas1–Cas2 protein complex is commonly involved in the acquisition of invasive DNA into the CRISPR locus (Figure 3a) [37–39]. The genomic CRISPR sequence consists of a leader sequence followed by an array of periodic repeats interspaced by previously immunized spacer sequences (as its acronym stands for) and the immunization of a new 'protospacer' DNA always occurs on the leader-proximal repeat, creating a new interspaced repeat unit (Figure 3b). Therefore, the order of spacers in the CRISPR array from the tail to head direction can be seen as an old-to-new historical immunization or invasion record of foreign DNA fragments in the host cell.

Active recording of high content artificial information into an *E. coli* cell population was recently demonstrated using the type I-E CRISPR acquisition system [40]. In this study, three sets of artificial protospacers were prepared separately for five rounds of information recording. For each round every 20 hours, three protospacers each from a different set were pooled and introduced into the cell population. Although a single bacterial cell cannot efficiently record all protospacer queries in its CRISPR locus, newly queried protospacers are always acquired in the upstream region of the older spacers. Using this property, a computational framework enabled prefect reconstruction of the orders of protospacer queries for all three sets from the massively parallel sequencing data of the population's CRISPR arrays. This approach was also

Figure 3



Spacer acquisition by CRISPR–Cas1–Cas2. **(a)** Structure of the Cas1–Cas2 protein complex bound to a protospacer DNA (PDB number: 5DQZ). **(b)** Schematic representation of spacer acquisition to a CRISPR array by the Cas1–Cas2 complex. **(c)** Conceptual diagram of TRACE. The copy number of the trigger plasmid is induced by a signal input, by which the spacer acquisition from the trigger plasmid is also enhanced, while that from the reference DNA remains at the same level.

demonstrated in another study, in which the world's first movie '*The Horse in Motion*' by Edward Muybridge in 1878 was converted into a five-frame 36 × 26 pixel animation and recorded in a cell population using a pool of 104 protospacers for each frame [41]. The five sets of the complex protospacer pools were sequentially introduced into a cell population from which more than 90% of the movie data were recovered.

Temporal recording of extracellular signals using Cas1–Cas2

The CRISPR–Cas1–Cas2 system can be also used to record environmental event histories of cells in their CRISPR arrays. As seen in high content artificial data recording examples, the directionality of spacer acquisition in the CRISPR locus allows simple implementation of a synthetic circuit to record time course event signals. TRACE (temporal recording in arrays by CRISPR expansion) is the first example that enables recording of temporal environmental stimuli of an *E. coli* cell population into their CRISPR arrays using CRISPR–Cas1–Cas2 [42••]. This method employs the concept of a trigger plasmid (pTrig), in which the phage P1 lytic replication protein RepL and its replication origin are placed downstream of a chemically inducible promoter, so that the copy number of the plasmid in cells can be elevated by the corresponding environmental compound. As opposed to the pTrig plasmid, the other cellular DNA, such as genomic DNA and the other plasmid encoding the anhydrotetracycline (aTc)-inducible Cas1–Cas2 system, is relatively stable and serves as reference DNA. While the acquisition of spacers from pTrig and the reference DNA occur at a constant rate when aTc induces Cas1–Cas2, the spacer acquisition frequency from pTrig is elevated when cells are also exposed to its corresponding chemical signal (Figure 3c). In TRACE, four variants of pTrig plasmids have been developed for IPTG, cooper, trehalose, and fructose signals, allowing time course event recording of combinatorial signal patterns into CRISPR arrays of a cell population at a certain degree.

Transcriptome-scale molecular recording into cell populations using RT-Cas1–Cas2

A common drawback in all of the abovementioned systems is that the recording targets need to be predefined in the circuit and are restricted to a limited number and kinds of

chemical compounds or cellular events that can be linked to specific promoters available in a target system. The ideal DNA event recorder would autonomously and efficiently record a wider range of temporal molecular and environmental information of cells to trace a variety of cell statuses.

Because type III CRISPR systems have activities to digest RNA [43–46] and some of their Cas1 genes are naturally fused to reverse transcriptase domains (RT-Cas1) in various species [47–49], it has been suggested that there is a conserved mechanism of spacer acquisition from RNA such as for the defense against RNA viruses. The first example of RNA-derived spacer acquisition has been recently discovered in *Marinomonas mediterranea* (MMB-1), in which intracellular RNAs are reverse transcribed and acquired in the genomic CRISPR array by the RT-Cas1–Cas2 complex [50•]. While functional reconstitution of MMB-1 RT-Cas1–Cas2 to yet to be performed in *E. coli*, another recent study has demonstrated that *Fusicatenibacter saccharivorans* RT-Cas1–Cas2 (*Fs*RT-Cas1–Cas2) can work heterologously in *E. coli* to acquire sequences of both endogenously expressed RNAs and viral RNAs into a synthetic CRISPR array. By developing a method to efficiently capture newly expanded CRISPR spacers, namely SENECA (selective amplification of expanded CRISPR arrays), for massively parallel sequencing, this study also established the Record-seq method that enables transcriptome-scale molecular recording in cell populations (Figure 4) [51••].

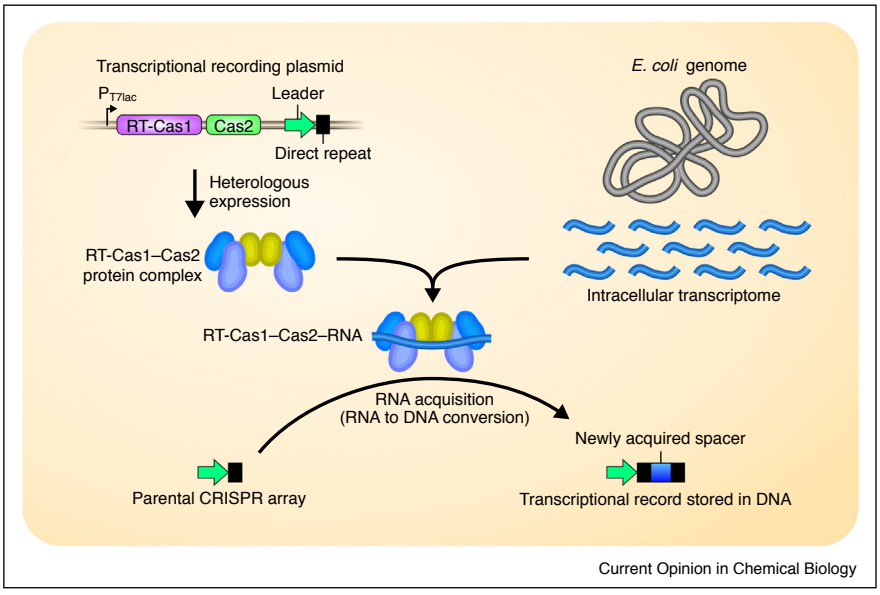
In this study, Record-seq successfully captured RNA-derived spacer acquisition from a chemically inducible transgene by which the acquisition frequency in the populational CRISPR array sequences was also elevated along with the increased input level of the chemical compound. Furthermore, while transcriptome and Record-seq profiles of *E. coli* cells in each of the different growth phases correlated at a certain level, a higher correlation was observed between a transcriptome profile of the log phase and a Record-seq profile of the stationary phase, showing that Record-seq captured past molecular profile information of cells observed at the time of the experiment. Record-seq also distinguished whether cell populations were under oxidative (or acid) stress or not, and whether cell populations were transiently experiencing herbicide exposure, which could not be differentiated by their transcriptome profiles. This is the first demonstration of a synthetic DNA recorder system that enabled retrospective reconstruction of past cell population statuses by recording global molecular profiles without defining specific target events.

Because SENECA identifies newly acquired spacer units separately, it currently limits the capacity of the Record-seq method that can readout only an additive (or cumulative) transcriptome profile for some ambiguous duration in the past of the cell population without their time course information. However, similar to the other Cas1–Cas2-based recording methods that successfully reconstruct the order of different event signals from CRISPR arrays, further developments would enable serial recording of dynamic changes in transcriptome profiles into cellular CRISPR arrays.

Outlook and discussion

The permanent recoding of environmental stimuli into cellular DNA has been achieved by synthetic circuits

Figure 4



Recording of a bacterial transcriptome in a synthetic CRISPR array by RT-Cas1–Cas2.

using DNA recombinases and more recently by CRISPR-related approaches. However, all of the current methods remain at the proof-of-concept stage, which have demonstrated molecular recording of event histories only at the cell population level. None of them have yet accomplished analysis of heterogeneous cells or clones with diverse phenotypic trajectories such as differentiation. This is because the numbers of target signals that can be efficiently recoded in synthetic DNA tapes have been currently limited to distinguish diverse cell statuses, and/or the signal-to-memory transmission efficiencies of the DNA writer modules have been too low to record sufficient amount of information in DNA. While mSCRIBE, CAMERA, and DOMINO have been demonstrated to achieve highly efficient recording of cellular events in synthetic DNA introduced into mammalian cells, the signal events that can be recoded simultaneously in these systems have been limited to those with available corresponding promoters. Although RT-Cas1–Cas2 has the potential to record transcriptome profiles into DNA, its spacer acquisition efficiency is extremely low [51••]. Furthermore, neither Cas1–Cas2 nor RT-Cas1–Cas2 has been successfully reconstituted heterologously in eukaryotic cells.

The DNA event recorder concept would be useful to study heterogeneous progression of mammalian systems only when it renders recording of a broader range of past information in DNA tapes at the single cell resolution or at least at the level of cell type groups, each of which corresponds to a distinct cell progression trajectory. While the establishment of RT-Cas1–Cas2-based molecular recording in mammalian cells could be one of the goals, it would require discovery of a new RT-Cas1–Cas2 variant that can heterologously work in mammalian cells or cryptic bacterial factors that support the function of the previously characterized RT-Cas1–Cas2 in mammalian systems, as well as enhancement of their efficiencies. Therefore, another possible direction could be increasing the scalability of Cas9-based approaches. In the inflammation recording by mSCRIBE and the Wnt signaling recording by CAMERA, the gene expression of the DNA writer modules, Cas9 or the C:G-to-T:A base editors, was under the control of specific promoters. However, a better approach would be regulating expression of multiple gRNAs in accordance with their respective promoters linked to specific gene expression. While gRNAs cannot be naturally expressed from variety of RNA polymerase type II (pol II) promoters, recent studies have shown that gRNAs concatenated with self-cleaving ribozymes [52] or tRNAs [52,53] can be expressed by pol II promoters. Furthermore, riboswitch systems have been demonstrated to regulate folding of the mature gRNA secondary structure in response to a 'trigger' RNA [54], showing the potential to construct large RNA circuits that transmit a broader range of RNA expression patterns to synthetic DNA tapes. For example, embryonic stem cells harboring a circuit that enables simultaneous recording of Oct4, Sox2, Klf4, cMyc, and other maker gene expression histories would be of great use in studying cell differentiation and reprogramming processes.

Taken together, the current progress in this field suggests some potential approaches to establish high capacity cell-embedded DNA event recorder systems that overcome the limitations of the current technologies to trace changes in molecular profiles and environmental conditions of different cells in complex biological systems including mammalian tissues and individuals. Moreover, as discussed in another article of this issue [32], the development of 'evolving DNA barcodes' has also been rapidly accelerated for high resolution cell lineage tracing of multicellular organisms with the concept of DNA barcodes and CRISPR–Cas9 to establish a new framework to discuss heterogeneous cellular phenotypes in the context of cell lineages. Further developments in both of these fields would enable us to understand how mammalian and other heterogeneous systems develop and function with their molecular and cellular contexts.

Conflict of interest statement

Nothing declared.

Acknowledgements

We thank members of the Yachie laboratory for discussions about different aspects of the current and past cell lineage tracing approaches. The research projects of the Yachie laboratory in this field are supported by grants from the Japan Science and Technology Agency (JST) PRESTO and CREST programs, the New Energy and Industrial Technology Development Organization (NEDO) Smart Cell Project, the Japan Agency for Medical Research and Development (AMED) PRIME program, the Shimadzu Science Foundation, the Takeda Science Foundation, the SECOM Science and Technology Foundation, the Naito Foundation, the Nakajima Foundation, and the Asahi Glass Foundation.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Clelland CT, Risca V, Bancroft C: **Hiding messages in DNA microdots**. *Nature* 1999, **399**:533-534.

2. Church GM, Gao Y, Kosuri S: **Next-generation digital information storage in DNA**. *Science* 2012, **337**:1628.

3. Goldman N, Bertone P, Chen S, Dessimoz C, LeProust EM, Sipos B, Birney E: **Towards practical, high-capacity, low-maintenance information storage in synthesized DNA**. *Nature* 2013, **494**:77-80.

4. Erlich Y, Zielinski D: **DNA Fountain enables a robust and efficient storage architecture**. *Science* 2017, **355**:950-954.

5. Organick L, Ang SD, Chen YJ, Lopez R, Yekhanin S, Makarychev K, Racz MZ, Kamath G, Gopalan P, Nguyen B *et al.*: **Random access in large-scale DNA data storage**. *Nat Biotechnol* 2018, **36**:242-248.

6. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ: **Synthetic gene networks that count**. *Science* 2009, **324**:1199-1202.

7. Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK: **Synthetic recombinase-based state machines in living cells**. *Science* 2016, **353** eaad8559.

8. Bonnet J, Yin P, Ortiz ME, Subsoontorn P, Endy D: **Amplifying genetic logic gates**. *Science* 2013, **340**:599-603.

9. Siuti P, Yazbek J, Lu TK: **Synthetic circuits integrating logic and memory in living cells**. *Nat Biotechnol* 2013, **31**:448-452.

10. Rubens JR, Selvaggio G, Lu TK: **Synthetic mixed-signal computation in living cells**. *Nat Commun* 2016, **7**:11658.

11. •• Weinberg BH, Pham NTH, Caraballo LD, Lozanoski T, Engel A, Bhatia S, Wong WW: **Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells**. *Nat Biotechnol* 2017, **35**:453-462.
This paper reported BLADE, which enabled computing of various logic circuits in mammalian cells using many different orthogonal site-specific DNA recombinases.

12. Yang L, Nielsen AA, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, Voigt CA: **Permanent genetic memory with >1-byte capacity**. *Nat Methods* 2014, **11**:1261-1266.

13. • Farzadfard F, Lu TK: **Synthetic biology. Genomically encoded analog memory with precise *in vivo* DNA writing in living cell populations**. *Science* 2014, **346** 1256272.
This paper reported SCRIBE, which enabled multiple signal recording in DNA of a living bacterial population by coupling the bacterial retron and recombineering technologies.

14. Miyata S, Ohshima A, Inouye S, Inouye M: ***In vivo* production of a stable single-stranded cDNA in *Saccharomyces cerevisiae* by means of a bacterial retron**. *Proc Natl Acad Sci U S A* 1992, **89**:5735-5739.

15. Mirochnitchenko O, Inouye S, Inouye M: **Production of single-stranded DNA in mammalian cells by means of a bacterial retron**. *J Biol Chem* 1994, **269**:2380-2383.

16. Sharon E, Chen SA, Khosla NM, Smith JD, Pritchard JK, Fraser HB: **Functional genetic variants revealed by massively parallel precise genome editing**. *Cell* 2018, **175**:544-557.e16.

17. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.*: **Multiplex genome engineering using CRISPR/Cas systems**. *Science* 2013, **339**:819-823.

18. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM: **RNA-guided human genome engineering via Cas9**. *Science* 2013, **339**:823-826.

19. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity**. *Science* 2012, **337**:816-821.

20. Dominguez AA, Lim WA, Qi LS: **Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation**. *Nat Rev Mol Cell Biol* 2016, **17**:5-15.

21. Adli M: **The CRISPR tool kit for genome editing and beyond**. *Nat Commun* 2018, **9**:1911.

22. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR: **Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage**. *Nature* 2016, **533**:420-424.

23. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY *et al.*: **Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems**. *Science* 2016, **353** eaaf8729.

24. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR: **Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage**. *Nature* 2017, **551**:464-471.

25. Komor AC, Zhao KT, Packer MS, Gaudelli NM, Waterbury AL, Koblan LW, Kim YB, Badran AH, Liu DR: **Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity**. *Sci Adv* 2017, **3** eaao4774.

26. Rees HA, Komor AC, Yeh WH, Caetano-Lopes J, Warman M, Edge ASB, Liu DR: **Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery**. *Nat Commun* 2017, **8**:15790.

27. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR: **Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions**. *Nat Biotechnol* 2017, **35**:371-376.

28. Hess GT, Tycko J, Yao D, Bassik MC: **Methods and applications of CRISPR-mediated base editing in eukaryotic genomes**. *Mol Cell* 2017, **68**:26-43.

29. Ryu SM, Koo T, Kim K, Lim K, Baek G, Kim ST, Kim HS, Kim DE, Lee H, Chung E *et al.*: **Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy**. *Nat Biotechnol* 2018, **36**:536-539.

30. Kim K, Ryu SM, Kim ST, Baek G, Kim D, Lim K, Chung E, Kim S, Kim JS: **Highly efficient RNA-guided base editing in mouse embryos**. *Nat Biotechnol* 2017, **35**:435-437.

31. Villiger L, Grisch-Chan HM, Lindsay H, Ringnalda F, Pogliano CB, Allegri G, Fingerhut R, Haberle J, Matos J, Robinson MD *et al.*: **Treatment of a metabolic liver disease by *in vivo* genome base editing in adult mice**. *Nat Med* 2018, **24**:1519-1525.

32. Masuyama N, Mori H, Yachie N: **DNA barcodes evolve for high-resolution cell lineage tracing**. *Curr Opin Chem Biol* 2019, **52**:63-71.

33. Perli SD, Cui CH, Lu TK: **Continuous genetic recording with self-targeting CRISPR-Cas in human cells**. *Science* 2016, **353** eaag0511.

34. •• Tang W, Liu DR: **Rewritable multi-event analog recording in bacterial and mammalian cells**. *Science* 2018, **360** eaap8992.
This paper reported CAMERA, which was successfully used to record order information of environmental stimuli in a synthetic DNA tape using a cytidine base editor. Using CAMERA, temporal Wnt activity of human cell culture was also successfully recorded in a cell-embedded DNA tape. See also Ref. [36••].

35. Hsu PD, Lander ES, Zhang F: **Development and applications of CRISPR-Cas9 for genome engineering**. *Cell* 2014, **157**:1262-1278.

36. •• Farzadfard F, Gharaei N, Higashikuni Y, Jung G, Cao J, Lu TK: **Single-nucleotide-resolution computing and memory in living cells**. *bioRxiv* 2018. 263657.
This paper reported DOMINO, which was successfully used to record order information of environmental stimuli in a synthetic DNA tape using a cytidine base editor. See also Ref. [34••].

37. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH *et al.*: **An updated evolutionary classification of CRISPR-Cas systems**. *Nat Rev Microbiol* 2015, **13**:722-736.

38. Jackson SA, McKenzie RE, Fagerlund RD, Kieper SN, Fineran PC, Brouns SJ: **CRISPR-Cas: adapting to change**. *Science* 2017, **356** eaal5056.

39. McGinn J, Marraffini LA: **Molecular mechanisms of CRISPR-Cas spacer acquisition**. *Nat Rev Microbiol* 2019, **17**:7-12.

40. Shipman SL, Nivala J, Macklis JD, Church GM: **Molecular recordings by directed CRISPR spacer acquisition**. *Science* 2016, **353** eaaf1175.

41. Shipman SL, Nivala J, Macklis JD, Church GM: **CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria**. *Nature* 2017, **547**:345-349.

42. •• Sheth RU, Yim SS, Wu FL, Wang HH: **Multiplex recording of cellular events over time on CRISPR biological tape**. *Science* 2017, **358**:1457-1461.
This paper reported TRACE, which was the first demonstration for simultaneous recording of multiple temporal stimuli into synthetic CRISPR arrays of an *E. coli* cell population using CRISPR–Cas1–Cas2.

43. Peng W, Feng M, Feng X, Liang YX, She Q: **An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference**. *Nucleic Acids Res* 2015, **43**:406-417.

44. Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A, Marraffini LA: **Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity**. *Cell* 2015, **161**:1164-1174.

45. Tamulaitis G, Kazlauskiene M, Manakova E, Venclovas C, Nwokeoji AO, Dickman MJ, Horvath P, Siksnys V: **Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus***. *Mol Cell* 2014, **56**:506-517.

46. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP: **RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex**. *Cell* 2009, **139**:945-956.

47. Toro N, Nisa-Martinez R: **Comprehensive phylogenetic analysis of bacterial reverse transcriptases**. *PLoS One* 2014, **9**:e114083.

48. Simon DM, Zimmerly S: **A diversity of uncharacterized reverse transcriptases in bacteria**. *Nucleic Acids Res* 2008, **36**:7219-7229.

49. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV: **A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action**. *Biol Direct* 2006, **1**:7.

50. • Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, Lambowitz AM, Fire AZ: **Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein**. *Science* 2016, **351** eaad4234.

This paper reported the first characterization of the RT-Cas1–Cas2 complex from a marine bacterial species *Marinomonas mediterranea*, which incorporates intracellular RNA sequences into the genomic CRISPR array.

51. •• Schmidt F, Cherepkova MY, Platt RJ: **Transcriptional recording by CRISPR spacer acquisition from RNA**. *Nature* 2018, **562**:380-385.

This paper reported the first heterologous reconstitution of RT-Cas1–Cas2 for RNA derived spacer acquisition in *E. coli*. By establishing the transcriptome-scale molecular recording using RT-Cas1–Cas2, the Record-seq method enabled retrospective reconstruction of various temporal cell population statuses from populational CRISPR arrays.

52. Nissim L, Perli SD, Fridkin A, Perez-Pinera P, Lu TK: **Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells**. *Mol Cell* 2014, **54**:698-710.

53. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK: **Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing**. *Nat Biotechnol* 2014, **32**:569-576.

54. Siu KH, Chen W: **Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function**. *Nat Chem Biol* 2019, **15**:217-220.