# Forging cellular memories

Using the CRISPR-Cas system, we engineer cells to create permanent molecular records of dynamic intracellular processes.

Imagine trying to understand an unpredictable movie – such as the iconic *Psycho* – based only on a sequence of still images, or worse yet, a single image. Biologists currently face a very similar conundrum - we attempt to decipher dynamic processes occurring within cells using snapshots of biological data. This is because obtaining such data – for example, in the form of nucleic acid molecules – is a disruptive process; you have to destroy a cell in order to extract its nucleic acids, thus freezing the timepoint of collection. What if we could devise an intracellular system to record and remember such molecular information over time? Such a system could potentially be extracted and read out in a video-like fashion.

DNA (Deoxyribonucleic acid) has been the molecular information storage system of choice for evolution: all cellular organisms – from the nifty *Escherichia coli* to the clumsy *Homo sapiens* – use it for encoding and transmitting genetic information. A 'gene' describes a part of cellular DNA that is used to construct a functional product (such as a protein) through a process known as gene expression. Gene expression is pretty dynamic – for example, different genes are expressed (activated) over a cell's lifetime, or in response to different environmental stimuli. The first stage in gene expression is to make an RNA (Ribonucleic acid) copy of the gene's DNA sequence – this process is known as transcription. These RNA copies are transient – they quickly get degraded by cellular machinery. Consequently, if we extract and read RNA molecules present within cells at a particular time point, we get a snapshot of gene activity at that time point – this is currently done through a technology known as RNA-sequencing (RNA-seq). However, instead of extracting cellular RNA, we wanted to continuously record transcriptional activity onto an intracellular recording device.

One scientific tool that has become a buzzword is CRISPR, currently in the news for its genome-engineering applications. In nature, however, CRISPR–Cas systems play a role in adaptive immunity, enabling microbes (such as bacteria) to record molecular memories of previous infections in the form of short nucleic acid sequences derived from the invasive element (such as bacteriophage viruses), which then helps these microbes to detect and fight future invasions. The recording module of CRISPR-Cas systems has previously been reengineered for fascinating applications, such as encoding a digital movie into bacterial DNA.

We leveraged a novel CRISPR-Cas recording module, which has the capability to convert RNA sequences to DNA (through a process intuitively known as reverse transcription), to capture and sequentially store cellular RNA molecules onto a DNA-based recording device in *E. coli*. We thus record transcriptional events in the cell over time, along with a temporal sequence of events, and subsequently read them out by extracting our recorder at the end of the experiment. However, the efficiency of this

recording system is relatively low, which roughly meant that our output signal was rather weak - and to address this, we came up with a technique to selectively amplify and detect our output. We read this output through deep sequencing, and decode the transcriptional information thus obtained using a dedicated computational pipeline. We named this entire technology – the CRISPR-based recording system coupled with all the downstream steps – Record-seq.

We show that Record-seq samples transcriptional events in a dose-dependent manner, i.e. RNA molecules that are highly abundant in the cell are recorded more frequently. We further demonstrate that Record-seq is capable of capturing cell responses to environmental stimuli, and also highlight its ability to record transient stimuli which cannot be detected by conventional RNA-seq.

We believe that our technology can be broadly applied to describe the complex sequence of events that gives rise to normal biology and disease. Further, since we can detect environmental signals using Record-seq, cells carrying our recording device can potentially have applications as biological sensors and sentinels – for example acting as diagnostic tools in the gut microbiome or reporting levels of toxic heavy metals in industrial environments. While we may be far from recording the molecular and cellular world in high definition, this might be a turning point – moving forward from snapshots to motion pictures.