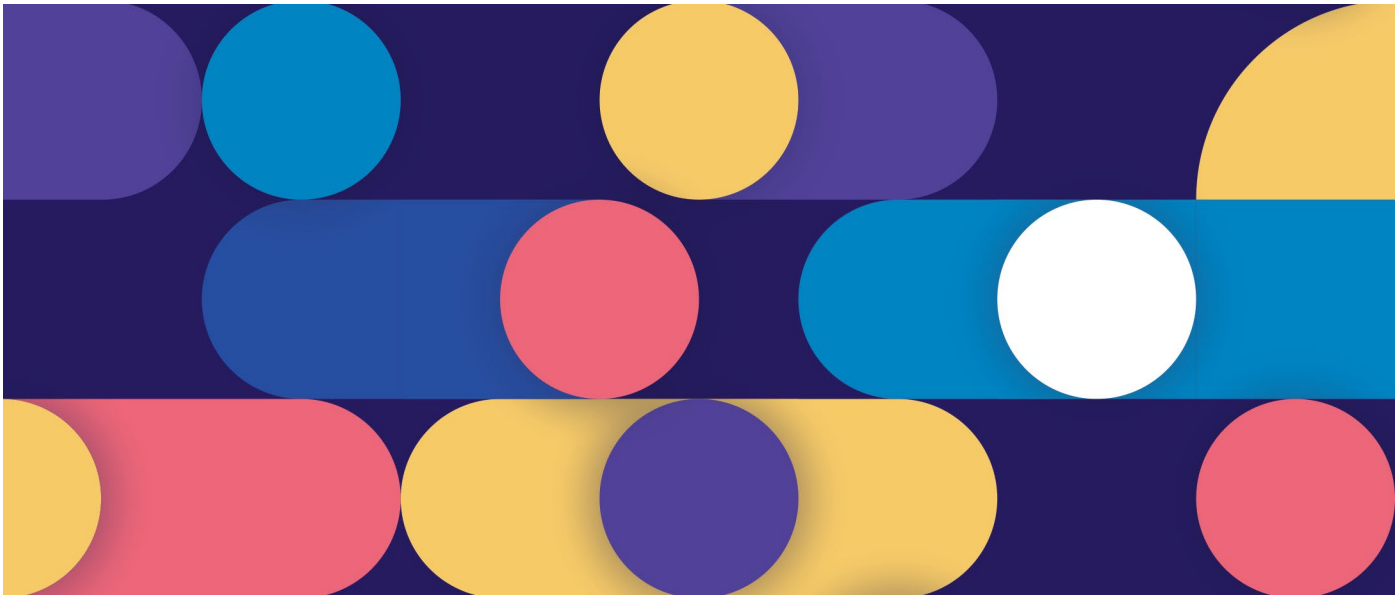
https://doi.org/10.1038/s41592-023-01985-x



A tough ask: high-efficiency, large-cargo prime editing

Researchers like the gene-editing method called prime editing for its precision and versatility. As they ask the method to do more, many factors shape what happens next.

By Vivien Marx

Which gene editing tool a lab chooses "really does depend on the context," says Prashant Mali from the University of California San Diego. "I love the versatility of prime editing, but will use base editing or direct genome editing for certain applications, such as, say, modifying stem cells, or for high-throughput screening efforts."

It won't call for extensively optimized efficiency when, for instance, a researcher wants to use prime editing[1] to generate a cell line, says David Liu, a Harvard University researcher and an investigator of the Howard Hughes Medical Institute who, along with his team, developed prime editing. For experiments with mammalian cells, they can consider small protocol tweaks[2], he says.

Optimization needs are growing as scientists are asking more of their gene-editing tools. MIT researcher Kristin Knouse studies the biological processes that underpin how the body's organs sense injury and react to it. Such insight could indicate how to repair damaged or diseased organs such as the liver. In her view, to realize the therapeutic potential of genome editing "we must improve our ability to deliver genetic tools to any cell type in the body at high efficiency." This efficiency exists with hepatocytes, but, she says, "efficient delivery to other cell types and tissues remains a challenge." Beyond overcoming the delivery barrier, she seeks editing efficiency, such that all cells receiving genome editing technology show DNA edited as intended.

In some conditions, a defect rests with production of a protein that is secreted systemically. In hemophilia, a mutation affects a clotting factor that the liver produces. To address this therapeutically, restoring protein production in a fraction of hepatocytes might, says Knouse, lead to enough circulating protein to have a curative effect. But when a disease mutation adversely affects every cell that harbors this change, such as in the case of protein storage diseases of the liver, she says, "reverting this mutation in only a fraction of cells may not be sufficient to restore liver function and offer therapeutic benefit."

Such questions in the realms of basic research and applied research, in mammalian cells or plants or in the clinical realm, drive scientists to address the tough ask of high-efficiency, large-cargo gene editing. High efficiency means the edit takes place in a large percentage of sites that are targeted for an edit and large cargoes involves ferrying in kilobase-sized pieces of DNA. Some scientists see prime editing as the method of choice to reach this goal but run into efficiency and cargo-loading challenges[3]. Here's how some, including the prime-editing method developers themselves, are tackling that.

One-stranded nick

As Columbia University researchers George Lampe and Samuel Sternberg note, CRISPR RNA-guided nucleases make double-stranded

breaks and set DNA repair processes in motion, which can lead to "heterogeneity in editing byproducts, and trigger aberrant chromosomal deletions and translocations"[4].

It's why some researchers turn to base and prime editing. In prime editing, after nicking one DNA strand, information is directly written onto a target site in a genome, and the method does not use double-stranded breaks. The method involves a programmable nickase, which is a partially catalytically impaired nuclease fused to a reverse transcriptase. The guide RNA is loaded with luggage for targeting a specific site in a genome and encodes the desired edit. "I am also one of many fans of gene editing without double-strand breaks," says Hyongbum Henry Kim at Yonsei University College of Medicine in South Korea.

Double-stranded breaks can suit for some experiments, he says, given that Cas9 is smaller than base- and prime-editing components. The guide RNA that points Cas9 to the genome location to be edited is shorter than prime-editing guide RNAs (pegRNAs). Once a break occurs with CRISPR–Cas9, DNA repair processes set in motion the possibility, for instance, of deleting a mutation. In prime editing, the mechanism is different: the prime editor proteins themselves install an edit. With CRISPR–Cas9, this process mainly works in cells that are dividing, and the major repair pathway in mammalian cells with CRISPR–Cas9 is non-homologous end joining. As Liu says, most therapeutically relevant types of cells do not support efficient homology directed repair (HDR), and the ratios of undesired indel to HDR are typically high. That's what makes prime editing, in his view, a robust method for installing or correcting many types of mutations and making precise small insertions, deletions and transversions.

"I am also one of many fans of gene editing without double-strand breaks," says Hyongbum Henry Kim.

Recently, Liu and his lab used phage-assisted continuous evolution and protein engineering (PACE) to generate sixth-generation prime editors that, he says, substantially enhance prime-editing efficiency. Double-stranded breaks can occur in prime editing, says Liu, but at around one hundred times below the rate at which they happen in the CRISPR–Cas9 system. It's a strength that prime editing does not require double-stranded breaks, as Liu phrases it, and can proceed, "if used carefully, with a minimum of indel byproducts and other undesirable consequences" that are associated with double-stranded breaks.

Prime editing "is virtually immune to bystander editing," he says, which is when nucleotides near the target sequence are edited unintentionally. The technique, he says, has perhaps the lowest average level of off-target editing among commonly used gene editing methods in mammalian cells. Such strengths matter especially in therapeutic gene editing. The new prime editor variants, he says "further improve editing efficiencies, especially for more challenging edits such as larger insertions."

Eye on large cargo

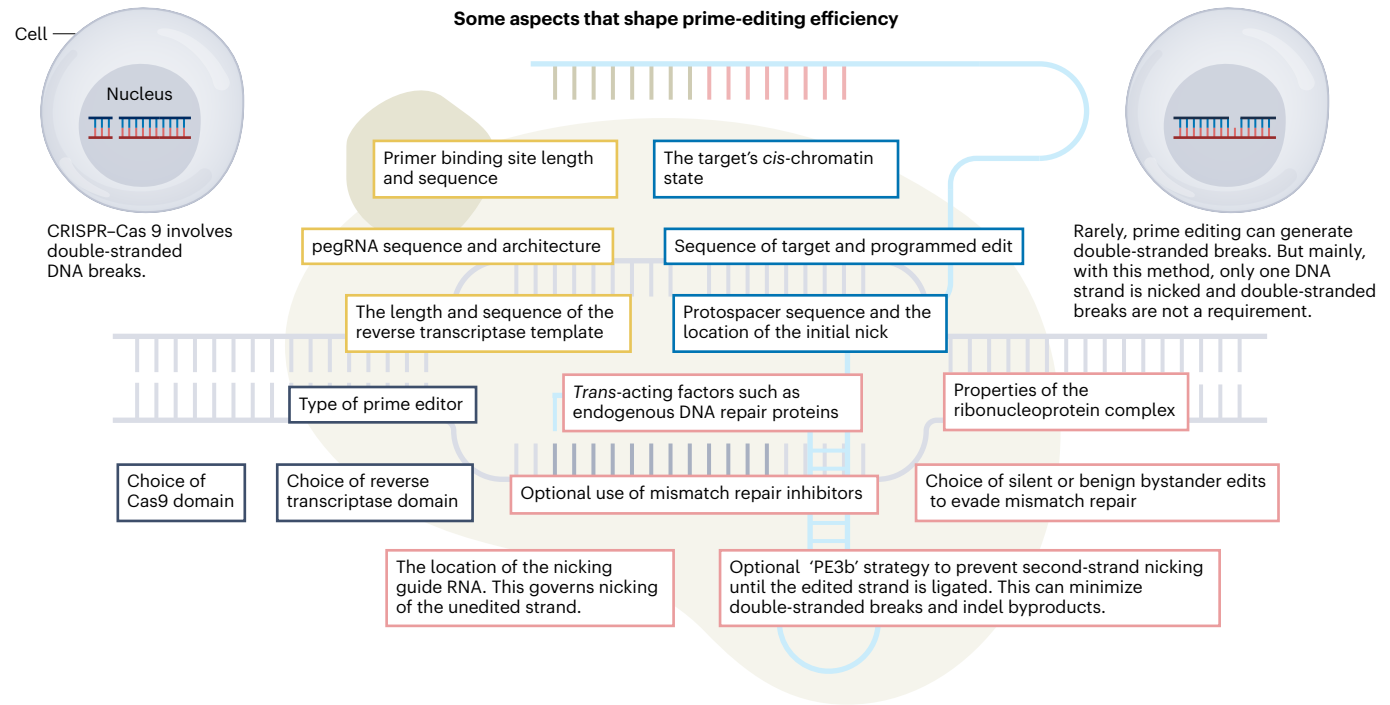
In publications[2,5], the Liu team points out that cargoes larger than 1 kilobase are hard to integrate into a genome through nuclease-directed double-stranded breaks. Stem cells and T cells have been successfully edited by adding donor DNA via HDR after double-stranded breaks, but HDR-mediated correction is cell-cycle dependent and "has proven inefficient in most therapeutically relevant cell types."

Standard prime editing can insert fragments several hundred bases in size into a genome, says Jeremy Duffield, chief scientific officer at Prime Medicine, a company David Liu co-founded. By adding integrases to prime editing, DNA fragments that are multi-kilobases in length can be ferried into a genome.

Other teams, too, use editing approaches combined with integrases, which are enzymes that can transfer phage genomes and various types of mobile generic elements between organisms[4–6]. With PASTE[6], which stands for programmable addition via site-specific targeting elements, the developer team at MIT, colleagues at other institutions and companies note they "married advances in programmable CRISPR-based gene editing, such as prime editing, with precise site-specific integrases." PASTE is in "a class of its own that is based off of prime editing," says MIT researcher Omar Abudayyeh, who runs his lab jointly with MIT researcher Jonathan Gootenberg. They co-led development of PASTE. In PASTE, pegRNAs have an attachment site encoded into their extensions. The Cas9-reverse transcriptase (RT) component inserts this attachment site into the genome, and that's where the integrase inserts its cargo, which can be as large as 36 kilobases. It can probably be larger, says Gootenberg, "we just didn't test beyond that." Phage genomes are on the order of 50 kb, he says, "so I would expect it to work to that size, likely." The team calls the guides attachment site-containing guide RNAs (atgRNAs). With the integrase, so-called pseudosites in the genome can be 'off-targets'. The team likes the integrase Bxb1, says Abudayyeh, "because there are no pseudosites as far as we are aware." PASTE involves engineering of CRISPR–Cas9 nickase combined with a reverse transcriptase and an integrase. The team reports editing efficiencies that match or are better than those of HDR and non-homologous end-joining-based methods and notes that PASTE works in dividing cells, non-dividing cells and in vivo. And there almost no indels, says Abudayyeh. The researchers have used PASTE with typical nicking guides and combined PASTE with the twinPE system that involves two atgRNAs, "which in many cases is more efficient," he says.

Says Liu, in 2016 he and his team had shown how one can combine Cas9 with recombinases, and later, with twinPE[7], they combined recombinases with paired pegRNAs. PASTE is similar to his team's approach, says Liu, "but in our hands PASTE is less efficient." Perhaps this is due to the way PASTE tethers prime editor and recombinase. Before the recombinase can act on the landing site and install its cargo in the genome, he says, the prime editor needs to move away from the target DNA site. Says Abudayyeh, in some versions of PASTE, the integrase is expressed separately via a linker. Overall, the team has found integrases to be active in every cell type they evaluated. Says Gootenberg, of the two steps, which are the Cas9-RT editing followed by integration via the integrase, between cell types there is more variability with the first step, the Cas9-RT editing. The integrase step, in their experience, performs more consistently across different cell types. It seems, says Abudayyeh, prime editing relies on some aspects of DNA flap repair or other DNA repair processes that might be cell dependent. "More work is needed here," he says.

Hao Yin from Wuhan University in Wuhan, China and his team developed GRAND editing[8] and used it to insert a fragment 1 kilobase in length into a genome. The efficiency of that edit was low, but for a 150-base-pair segment the efficiency was 63%. "We are working on increasing the efficiency of inserting 1 kb now," says Yin. GRAND generates two DNA flaps that are complementary to one another while prime editing involves one flap that complements the targeted sequence. Yin



Experimenters like prime editing for its precision. As with other gene-editing techniques, many factors influence this method's efficiency, which is a measure of how many targeted sites are edited as intended. Scientists are exploring ways to address these factors, which include general aspects about the process (pink) and parameters related to the DNA target (blue), the prime editors (black) and pegRNA design (gold).

and his team are addressing ways to curb the observed bystander mutations, which are changes a few bases from the site targeted for an edit. The scientists find they can insert longer sequences with GRAND editing than with traditional prime editing, and the process of GRAND editing appears to involve a molecular mechanism unlike the one in prime editing, says Yin.

Says Kim, "the size of the DNA fragments that PASTE can insert seems to be almost an order of magnitude higher than those by GRAND or twinPE." To his knowledge, GRAND, twinPE[7] and PASTE are still being improved, and "improving the efficiency of prime editing and large DNA insertion will be important." GRAND and twinPE could be used together with PASTE to improve editing efficiency, he says. The ability to insert big fragments in a targeted manner is, in his view, "the remaining important issue in the field of gene editing."

Vast 'influencers'

"PE is not a fast-food technology," says Liu. When optimizing prime editing to raise efficiency, a researcher has much to consider. The factors he sees include the protospacer sequence and the location of the initial nick; the length of the primer binding site and its sequence; the length and sequence of the reverse transcription template; the choice of reverse transcriptase domain; the choice of Cas9 domain; the choice of silent or benign bystander edits to include to evade mismatch repair; the optional use of mismatch repair inhibitors such as MLH1dn; the location of the nicking guide RNA, which nicks the unedited strand; and use of a 'PE3b' strategy to prevent second-strand nicking until the edited strand is ligated, which minimizes indel byproducts.

> **"PE is not a fast-food technology," says David Liu.**

In his view, the good news about these and other variables is that they give users many opportunities for optimization. That's in contrast to nuclease and base editing approaches, he says, "in which there are typically only a small handful of possible strategies to try before you've exhausted all possibilities." But the vastness of the prime editing parameter space also means that the search for an optimal prime-editing strategy "can take many, often hundreds, of different attempts," and most will require constructing and testing different pegRNAs.

In his interactions with researchers who reach out to his lab about what works well and less so, says Liu, he sees that people with the most efficient prime editing experiences have done the most extensive exploration of this parameter space. For instance, they might be using pegRNAs that had been previously optimized through an extensive search. But experimenters "who test only a handful of pegRNAs and PE conditions tend to have a lower likelihood of achieving high levels of prime editing efficiency," says Liu. It's fortunate that several machine learning models have begun to narrow the size of the parameter space to be explored in order to get the best-performing prime-editing systems for a given edit.

One set of methods pertains to machine learning models for design of pegRNAs, the prime-editing guide RNAs that have multiple components. It's challenging to design pegRNAs to maximize efficiency, says Kim. So he and his team developed the computational approaches DeepPrime and DeepPrime-FT[9] for predicting editing efficiencies in eight

SOURCES: KIM LAB, YONSEI COLL. OF MED.; LIU LAB, HARVARD UNIV.; SHENDURE LAB, UNIV. WASHINGTON; H. YIN, WUHAN UNIV. CREDIT: T. PHILLIPS, L. TANG, SPRINGER NATURE

prime editing systems in seven cell lines for edits up to three base pairs long. The tools let scientists predict efficiencies both for matched and mismatched targets for the different combinations of prime editors and pegRNAs.

DeepPrime, says Kim, can be used as a general deep learning model to design pegRNAs for various experimental conditions. DeepPrime-FT provides more optimized prediction of pegRNA efficiencies depending on the cell type used and prime editors and when considering the many different versions of prime editors, he says. He recommends choosing DeepPrime-FT results from experimental conditions similar to the user's intended ones. If they are unclear which one to select, he says, "users can simply use DeepPrime."

Junhong Choi, a postdoctoral fellow in the University of Washington lab of Jay Shendure, agrees that pegRNA design shapes efficiency, as does pegRNA architecture. To assess options, one can use the Kim lab's tools and others. Tweaking efficiency can involve additional RNA engineering, as with epegRNA from the Liu lab or petRNA by Xue and Sontheimer. Additional factors influencing prime editing include DNA mismatch repair. This corrects DNA mismatches that happen when DNA replicates, but in prime editing it can suppress prime editing efficiency and promote indels. Liu and team developed options for prime editing in vivo or in difficult-to-transfect cell types with a prime editor protein architecture experimenters can choose, called PEmax architecture[10], which, among other aspects, includes two Cas9 mutations that increase nuclease activity. Dual pegRNA prime editing approaches can be helpful, too, such as Bi-PE, GRAND, HOPE, twinPE, PEDAR and PRIME-Del.

Application: delivery

"I have worked on delivery of CRISPR for a decade." says Yin. Typically, labs use viral delivery, but with GRAND, he says, non-viral delivery such as methods using RNA and lipid nanoparticles is feasible. In gene-editing experiments with T cells and hematopoietic stem cells, says Yin, non-viral DNA delivery often kills the cells. His lab in collaboration with Wuhan University colleagues in the lab of Ying Zhang have worked out how to improve non-viral DNA delivery to these cell types and make gene-editing more precise and efficient[11]. "We identified the mechanism of cell death after non-viral DNA delivery and figured out ways to mitigate such death and to boost target insertion in these cells," he says.

Experimental electroporation of T cells appears to trigger an innate immune signaling pathway – cGAS-STING – that likely plays a role in disrupting generation of chimeric antigen receptor (CAR)-T cells. Given its broad roles in the cell, the cGAS-STING pathway itself cannot be knocked out. But cGAS activity is dampened by changing the osmolality of the electroporation buffer. Thus, buffer optimization improves non-virally mediated targeted insertion.

For some gene-editing applications, a viral vector such as adeno-associated virus delivers gene therapy components, says Duffield. In an approach that does not involve prime editing, Sarepta Therapeutics uses a viral vector in its recently FDA-approved treatment for Duchenne muscular dystrophy, in which people lack a muscle protein called dystrophin. With infusion therapy, the viral vector delivers a transgene with parts of the dystrophin gene to muscles. Delivery, says Prime Medicine's Duffield, "is one of the harder problems." To address the immune deficiency chronic granulomatous disease, the potential treatment the company is working on would be essentially an enzyme with a piece of RNA that must travel to the cell's nucleus. It's a much larger structure than typical small molecules, he says, and it involves RNA, a charged molecule.

Prime-editing components might be delivered ex vivo. In this approach, a person's cells are edited in the lab and then infused, as is the case with engineered T cells to treat cancer. A different approach, he says, builds on mRNA vaccine technology. "We think this RNA–lipid nanoparticle system is a great way to deliver prime editing," he says. To treat the lung, a nebulizer might be used; for other organs, infusion might be chosen. According to the company, the team has increased delivery efficiency for liver and brain.

Beyond overcoming the delivery barrier, says Kristin Knouse, she wishes for gene editing efficiency, such that all cells receiving genome editing technology show DNA edited as intended.

The company benefits from academic work, and some of Prime Medicine employees are former members of the Liu lab. One of the company's co-founders is Andrew Anzalone, who led the development of prime editing as a postdoctoral fellow.

Says Duffield, "we've really been able to industrialize prime editing." Companies are, in Duffield's view, more able than academic labs to industrialize screening and build functional and safety assays that are quality controlled and that work at scale. Beyond investing in high-throughput screening capabilities, the company has also developed machine learning algorithms that analyze screening data and "throw out poorly performing guides," he says. A screen might lead to a guide that has 5–10% efficiency. After optimization, the efficiencies are in the 70–90% range, he says. An efficiency of 70–90% means that, among all the DNA copies, 70–90% of them have the desired correction at the intended site after a single administration.

An important influence on prime editing outcomes is the repair pathway, he says. The scientists at the company are "very capable now of understanding that mismatch repair pathway and how to manipulate it to our advantage," he says. "And that's one of the ways that we are able to tune editing to very high efficiency." For a single target in the genome, Duffield and his team screen, in automated high throughput, hundreds to thousands of pegRNAs. This identifies the most active prime editors. Selection of guides is further optimized by using machine learning to filter according to factors that favor high efficiency.

Says Meredith Goldwasser, who directs company strategy and who was interviewed jointly with Duffield, she sees in publications from academic labs that scientists have often not optimized prime editing. To improve efficiency, "part of our special sauce has been this group of computational biologists employing machine learning," she says. They are also working to make prime editors smaller.

In work at the company using prime editing, they corrected a two-nucleotide mutation in the *NCF1* gene in CD34$^+$ hematopoietic stem cells. The gene encodes a protein subunit in NADPH oxidase. It is one of several contributors to chronic granulomatous disease, an inherited immunodeficiency that causes recurrent infections people can sometimes not fend off. Keith Gottesdiener, the company's CEO, says that in mice the team's prime editing reached 92% efficiency and enzyme activity was restored. They saw no off-target effects.

If and when the approach is approved for testing in people, a drug would be given that releases some bone marrow cells into a patient's bloodstream. Stem cells are purified from that sample, edited in the lab, and infused back into the patient. The cells home

to the bone marrow and make immune cells. Because stem cells are edited, the edits are passed on to daughter cells.

In the company's view, factors that make prime editing suited for gene therapy include the way it can be programmed to insert or delete DNA fragments in the kilobase size range, how it minimizes bystander editing and off-target edits, how it corrects mutations in both dividing and non-dividing cells, and that it can be used ex vivo and delivered with either viral or non-viral means. The company scientists apply prime-editing-assisted site-specific integrase gene editing (PASSIGE). It's similar to PASTE, says Duffield. PASSIGE involves placing a unique 'code' in the DNA as a recombinase target sequence, which a recombinase enzyme locates. The recombinase swaps in a piece of DNA there. One day, says Duffield, one might be able to insert an edited version of the large gene that causes cystic fibrosis.

Prime editors are "really precise," yet compared to the efficiency of other techniques, the efficiency of prime editing is "evolving," said Jay Shendure.

Application: developmental biology

At this year's annual meeting of the International Society for Stem Cell Research, Jay Shendure described how he and his team reconstruct mammalian development as they work toward a consensus ontogeny of mouse development. They profiled mouse embryos of different ages and built cell type lineage trees that span developmental phases from zygote to newborn pup. Among other methods, they used a single-cell combinatorial indexing and barcoding approach — sci-RNA-seq3 — to profile many cells.

They added to this single-cell RNA sequencing, an approach called DNA Typewriter[12] developed by Junhong Choi. The method leverages prime editing and uses pegRNAs to encode insertions. Each insertion writes a sequence into the DNA. This 'tape' can be followed as cells divide and the tape sequence changes. Shared sequences between tapes from different cells show how the cells relate and give a time-resolved recording of molecular events.

For lineage tracing, Choi has been thinking about programming synthetic mutations this way and using them to infer relationships among cells. He and the team are now optimizing the technology and are keen on achieving a balance with a good editing rate: "not too fast, as the array will be saturated with edits in a short time, or not too slow." Another factor that matters is how to enable good recovery of recorded information, especially in a readout from single-cell RNA sequencing.

In the lab's experience, Shendure said, prime editors are "really precise," yet compared to the efficiency of other techniques, the efficiency of prime editing is "evolving." Xiaoyi Li, a postdoctoral fellow in the Shendure lab, and her colleagues point out in their preprint[13] the factors they see shaping the efficiency of prime editing and note that the efficiency at an endogenous locus "is generally low and highly variable across target sites." Influences include the properties of ribonucleoproteins delivered to the cell, the sequence of pegRNAs and the type of prime editors, the primary sequence of the target and the programmed edit, and *trans*-acting factors such as endogenous DNA repair proteins. Then there's the *cis*-chromatin context of the target site. Of these factors, the way the *cis*-chromatin environment modulates the efficiency of prime editing "remains largely uncharacterized."

Chromatin context has an "enormous impact" on prime-editing efficiency, said Shendure. At one location, 2% of the alleles might be edited; in another location, 95% are edited. Chromatin compaction can physically shield DNA from damaging agents, and epigenetic factors can interact with DNA repair factors and modulate the kinetics of the DNA damage response. With CRISPR–Cas9 gene-editing, Cas9 activity is strongly influenced by chromatin, and the epigenetic environment overall influences gene-editing efficiency and shifts the balance of DNA repair pathways. The team studied whether different prime-editing loci respond differently to perturbations in DNA repair pathways, says Li.

Says Choi, different cells or organs can respond differently to prime editing owing to different chromatin states or a different dominant DNA repair pathway. Xiaoyi Li agrees on the role of the many factors that influence prime-editing efficiency. Li and her colleagues assessed the influence of chromatin state on prime editing[13]. They integrated chromatin 'sensors' into the genome that act as synthetic prime editing targets. Using methods such as a T7-assisted reporter mapping assay to track these locations, they observed "highly variable prime editing outcomes across genomic contexts." They identified epigenetic features predictive of high and low prime editing efficiencies.

When the target site is within a gene, prime editing efficiency correlates with the gene's transcriptional activity and negatively correlates with the distance between the target and the transcription start site. They see multiple ways to modulate prime editing outcomes with targeted epigenetic reprogramming steps at the gene that is to be targeted in prime editing.

Says Li, the readily available epigenetic editing tools are mostly CRISPR-based and can be used to both to tune up and tune down gene expression. "We think gene activation would be more useful in the setting of prime editing, as high expression generally correlates with high editing efficiency," she says. Methods such as CRISPRoff can tune down prime editing efficiency. Among the most direct technologies for modulating gene expression, she says, are CRISPRa, including CRISPR-based synergistic activation mediator, for gene activation and CRISPRi for transient repression and CRISPRoff to achieve prolonged repression by DNA methylation.

For some prime-editing applications, high efficiency may not be a main concern. But when it does carry weight, scientists will want to optimize as many influences as they can.

Vivien Marx ✉
Nature Methods.
✉e-mail: v.marx@us.nature.com

Published online: 25 August 2023

References

1. Anzalone, A. V. et al. *Nature* **576**, 149–157 (2019).
2. Doman, J. L., Sousa, A. A., Randolph, P. B., Chen, P. J. & Liu, D. R. *Nat. Protoc.* **17**, 2431–2468 (2022).
3. Ferreira da Silva, J. *Nat. Commun.* **13**, 760 (2022).
4. Lampe, G. D. & Sternberg, S. H. *Nat. Biotechnol.* **41**, 471–472 (2023).
5. Durrant, M. G. et al. *Nat. Biotechnol.* **41**, 488–499 (2023).
6. Yarnall, M. T. N. et al. *Nat. Biotechnol.* **41**, 500–512 (2023).
7. Anzalone, A. V. et al. *Nat. Biotechnol.* **40**, 731–740 (2022).
8. Wang, J. et al. *Nat. Methods* **19**, 331–340 (2022).
9. Yu, G. et al. *Cell* **10**, 2256–2272.e23 (2023).
10. Chen, P. J. et al. *Cell* **184**, 5635–5652.e29 (2021).
11. An, J. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-023-01073-7 (2023).
12. Choi, J. et al. *Nature* **608**, 98–107 (2022).
13. Li, X. et al. Preprint at *bioRxiv* https://doi.org/10.1101/2023.04.12.536587 (2023).