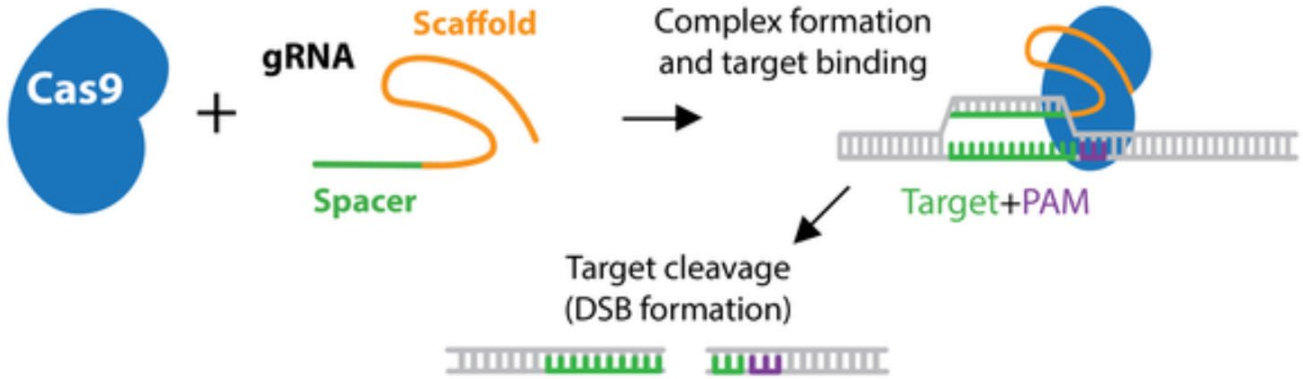
Prime editing and CRISPR screens for studying of DNA repair

Technical Journal Club

Lukas Frick

31 May 2022

CRISPR and DNA repair

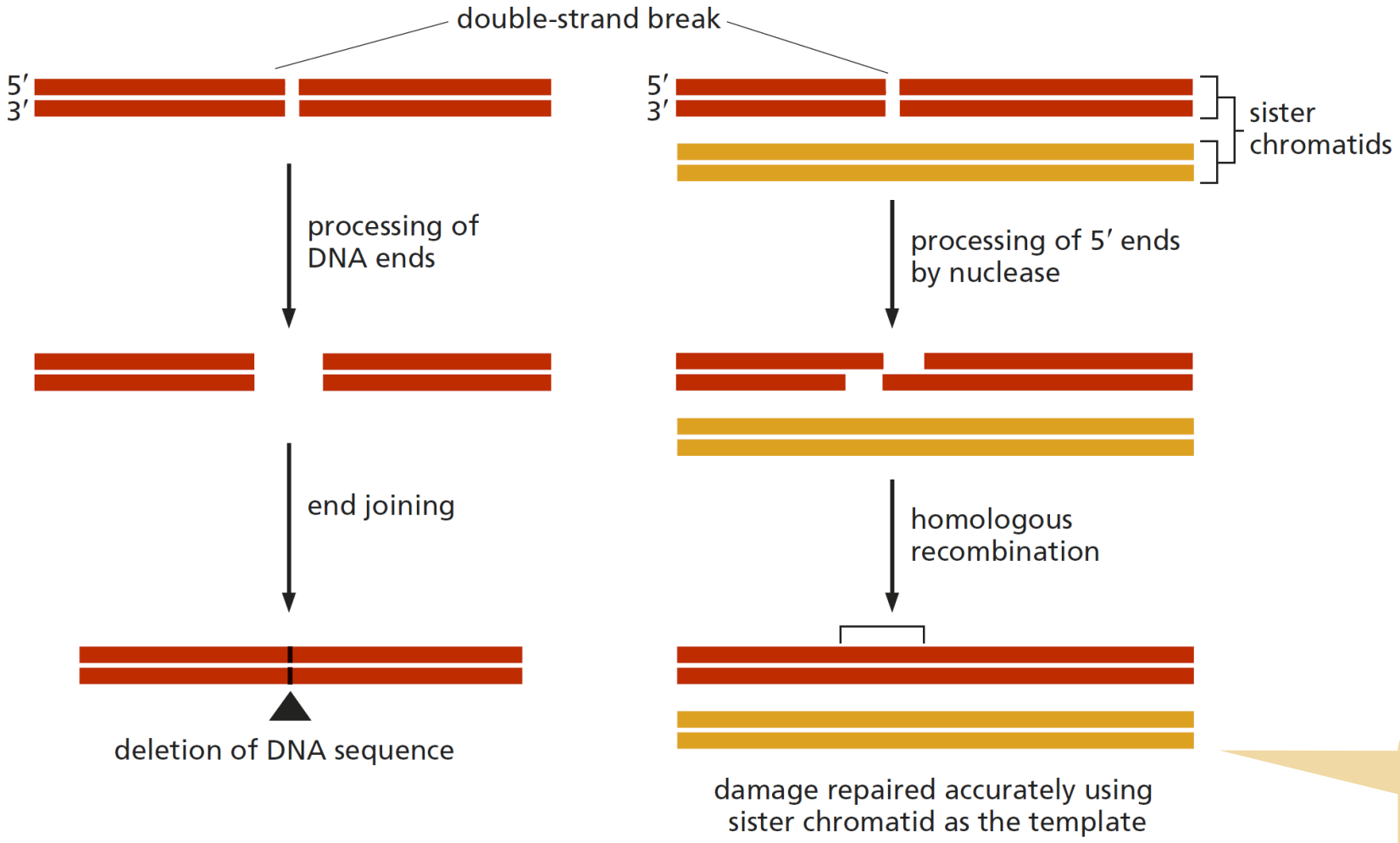


- CRISPR systems that aim to modify the genome (including the original – CRISPR-cut) **rely on the cell's built-in DNA repair mechanisms**.
  - Typically, a range of repair outcomes are seen, even at a single locus.
  - The determinants of these outcomes are poorly understood.

- I will present two papers that use **CRISPR inhibition screens** to study DNA repair.
  - Goal #1: **Improve CRISPR technology** by skewing DNA repair in the desired direction.
  - Goal #2: Advance the **general understanding** of DNA repair with an unbiased dataset.

Non-homologous end-joining (NHEJ)

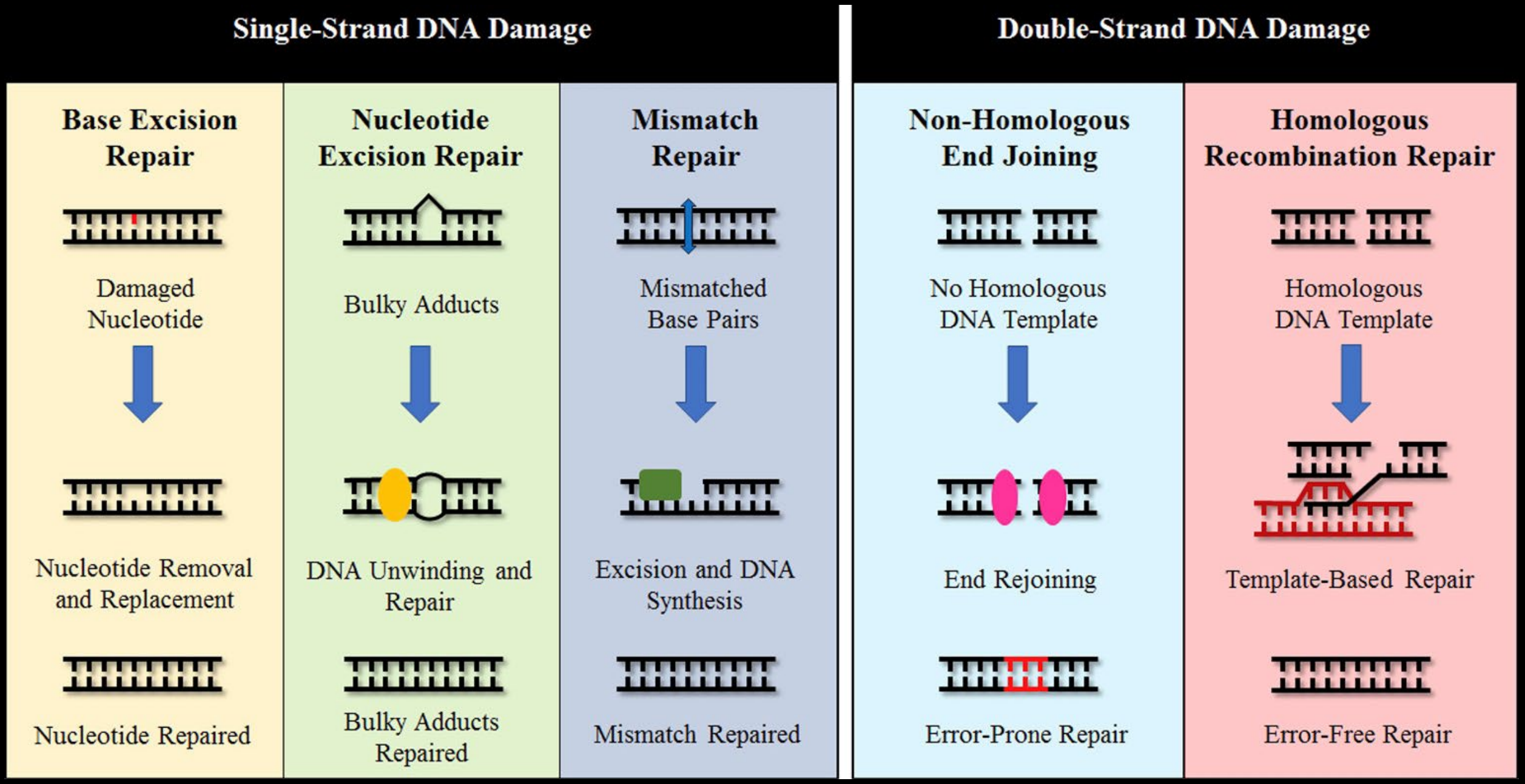
Homologous recombination (HR)



- Double-strand breaks induced by **CRISPR-cut** are usually resolved by **NHEJ**.
- Homologous recombination can provide **error-free** repair of mistakes that occur during DNA replication. However, it requires a template.

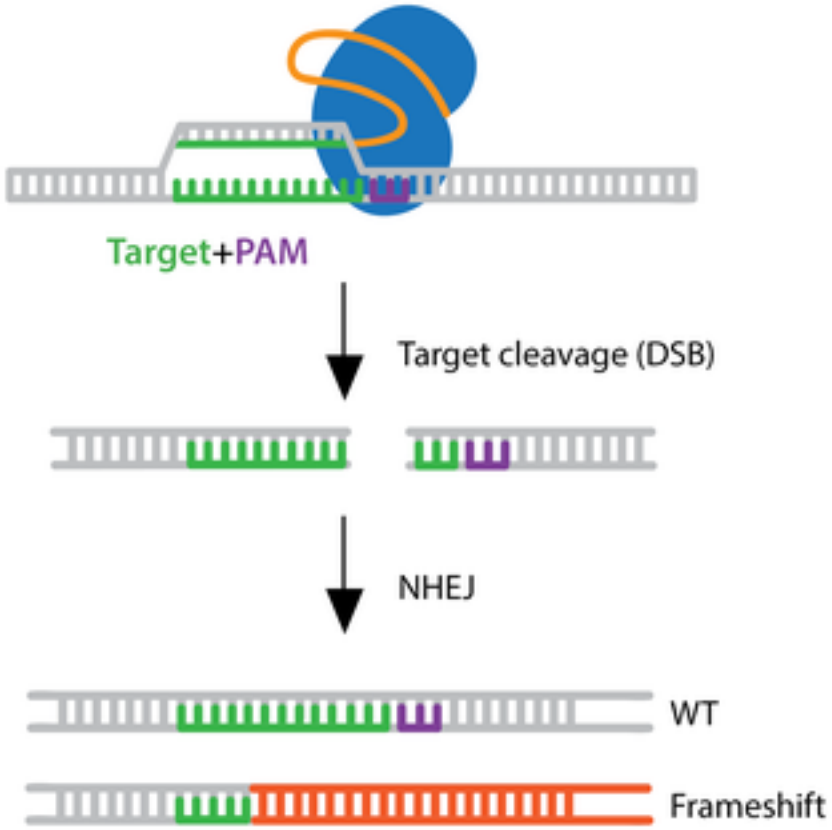
Only in S or G2 phases of the cell cycle!

Albert's Molecular Biology of the Cell, 6e



Minten and Yu, 2019

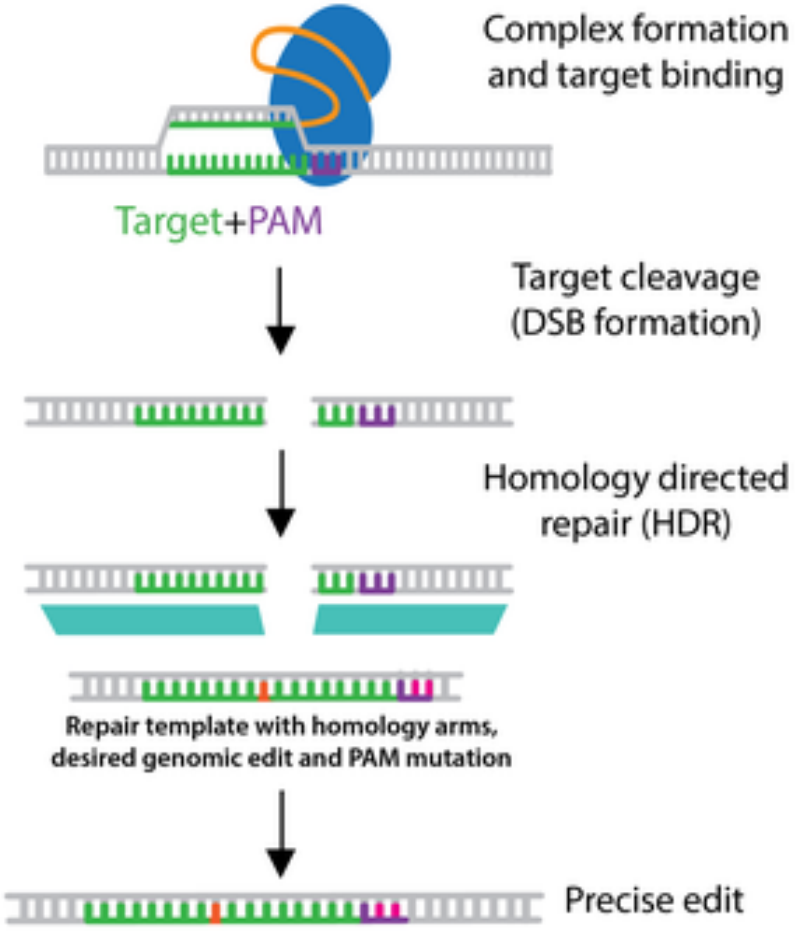
Cas-nuclease and NHEJ lead to indels (short insertions or deletions) that often cause a **frameshift** → gene knockout



- **NHEJ can also result produce perfect repairs.** In fact, only ~5% of NHEJ events after CRISPR cutting (at blunt DNA ends without damaged nucleotides) result in indels.
- However, since the target sequence and the PAM remain intact, it will be cut again!
- (NHEJ takes ~1 hour to complete, so we can expect gene knockout within 24 hours.)

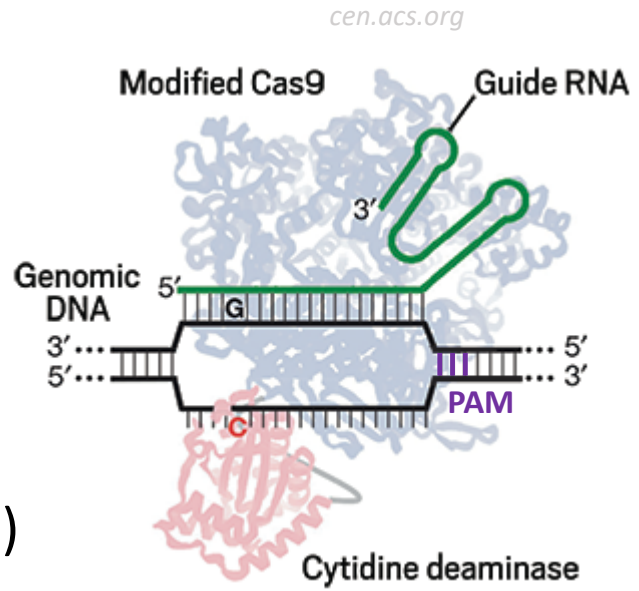
blog.addgene.org/crispr-101-non-homologous-end-joining

The other pathway for DSB repair – homology directed repair – can also be exploited with CRISPR to introduce specific sequences at a chosen site



- A DNA template with the desired sequence and homologous arms is provided, along with the CRISPR-cut system.
- Drawbacks:
  - Only works in dividing cells (the necessary factors for homologous recombination are not expressed in the G1 phase).
  - Very inefficient!

Two new and efficient CRISPR systems permit precise genome modifications without double-strand breaks

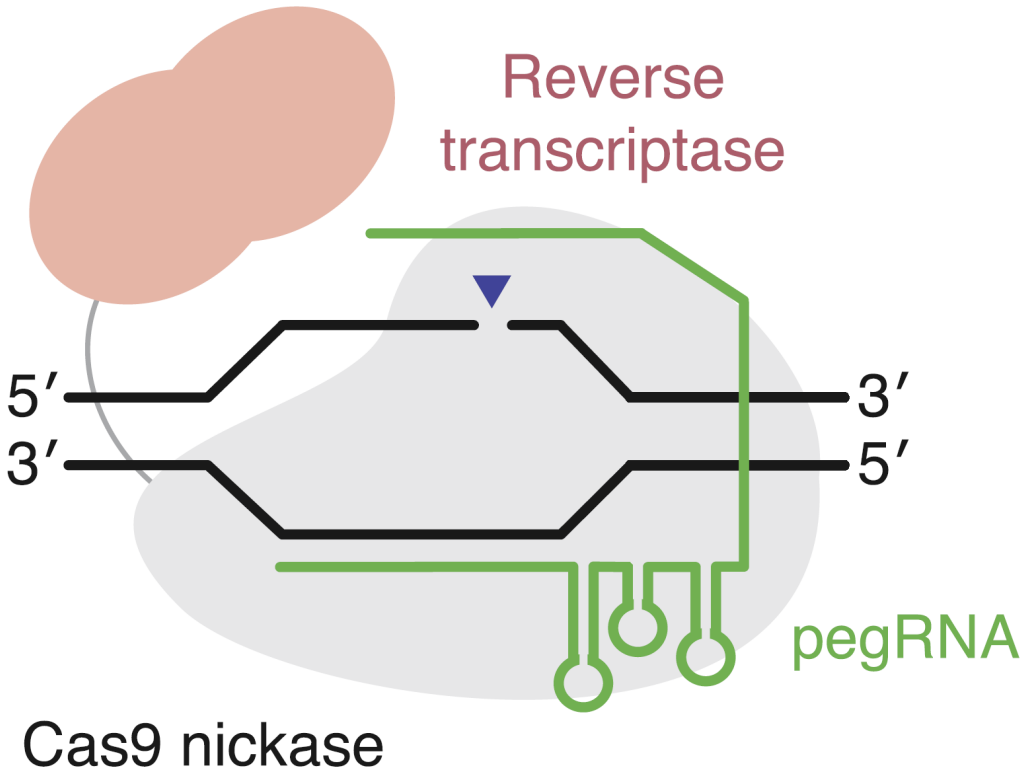


1. **Base editors**
   - Cas9 coupled to a cytosine or adenosine deaminase
   - Efficient, but with limitations:
     - Only supports single-nucleotide **transition** mutations

       → 4 out of 12 possibilities (C→T, but not C→A or C→G)
     - Requires a nearby PAM (protospacer-adjacent motif, **NGG**)
     - Can be unspecific if there are multiple target nucleotides near the PAM

2. **Prime editors**

Prime editors

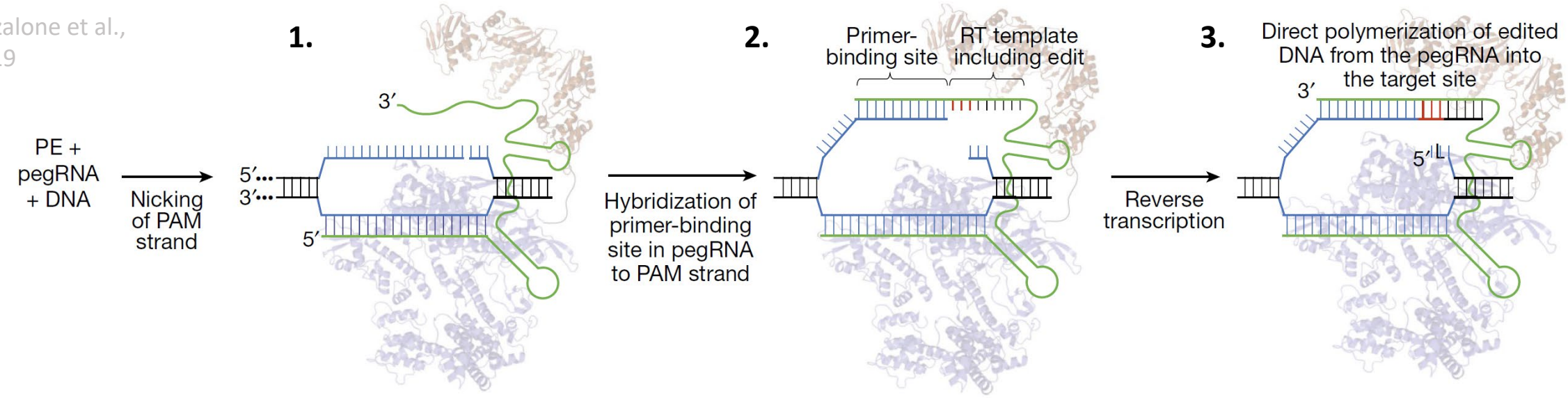


- In prime editing, a **guide RNA** and an **RNA template** are combined in a single RNA molecule – the prime editing guide RNA (**pegRNA**).
- Prime editing is **versatile**:
  - → can produce **any** point mutation
  - → supports small insertions or deletions (~ 50 bp)

Cas9 nickase (Cas9n) is a Cas9 variant that produces single-strand nicks instead of double-strand breaks

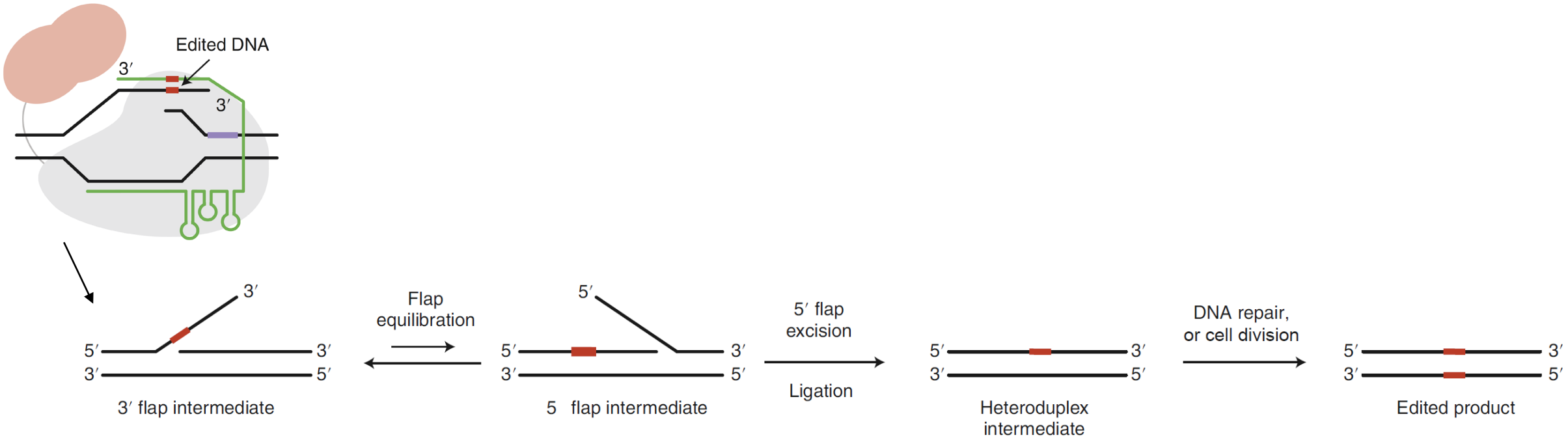
Anzalone, Koblan and Liu, 2020

Anzalone et al., 2019



The prime editor consists of a Cas9 nickase fused to a **reverse transcriptase**.

1. The guide RNA part of the pegRNA directs the Cas9 nickase to a specific target sequence. One strand is nicked 3 nucleotides upstream of the PAM.
2. The 3’ end of the pegRNA consists of a sequence complementary to the nicked strand – the **primer-binding sequence (PBS).** The PBS hybridizes to the 3’ DNA flap.
3. The hybridized 3’ flap acts as a **primer for the reverse transcriptase**, which starts to polymerize onto the DNA. It **incorporates the RT template** containing the desired edit.



- The next steps are driven by cellular processes and are incompletely understood.
- In a percentage of cases, the 3’ flap with the newly incorporated templated sequence displaces the original 5’ flap of DNA, which is excised.
- A DNA heteroduplex forms. In some cases, this heteroduplex is replaced by the edited sequence also on the opposite strand (by DNA repair or replication), producing the desired edit.

The efficiency of prime editing is locus- and cell-line dependent.

In “good” cell lines, it can reach 20-50%.

Article

Search-and-replace genome editing without double-strand breaks or donor DNA

Andrew V. Anzalone[1,2,3], Peyton B. Randolph[1,2,3], Jessie R. Davis[1,2,3], Alexander A. Sousa[1,2,3], Luke W. Koblan[1,2,3], Jonathan M. Levy[1,2,3], Peter J. Chen[1,2,3], Christopher Wilson[1,2,3], Gregory A. Newby[1,2,3], Aditya Raguram[1,2,3] & David R. Liu[1,2,3]*

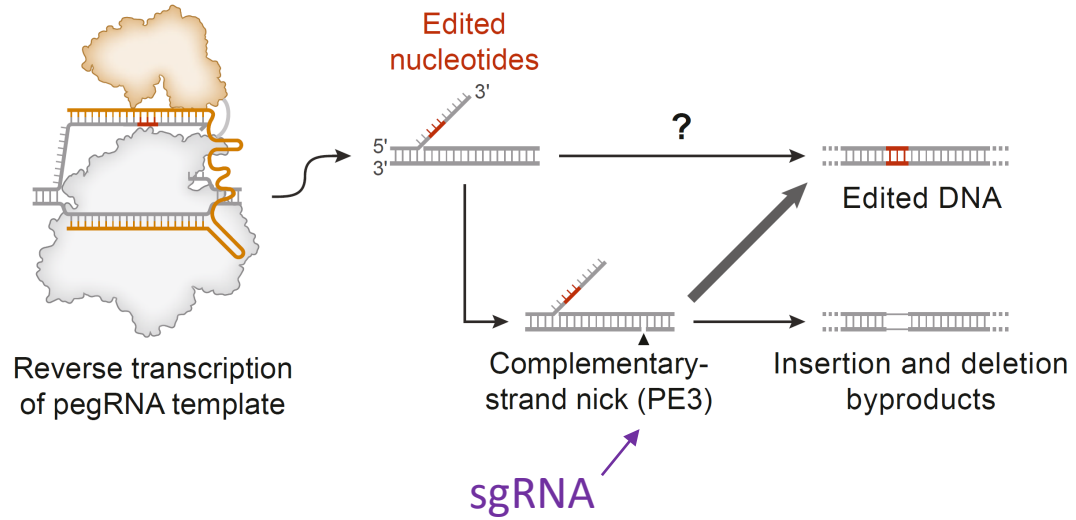
Nature, October 2019

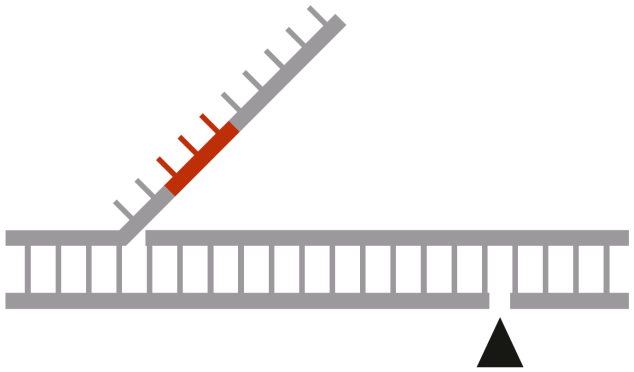
Prime editing was first described by the group of David Liu in 2019.

Already in the first paper, they describe 3 modifications to improve efficacy.

Search-and-replace genome editing without double-strand breaks or donor DNA

- **PE1**: First-generation the prime editor. Consists of the M-MLV RT (Moloney Murine Leukemia Virus Reverse Transcriptase) fused to Cas9(H840A) nickase.
- **PE2**: Over 30 variants of reverse transcripase were tested. An M-MLV RT variant with 5 single-nucleotide substitutions was found to have the highest efficiency.
- **PE3:** An additional sgRNA is introduced that targets the opposite strand. The resulting nick promotes edits by encouraging the removal of the unedited strand.



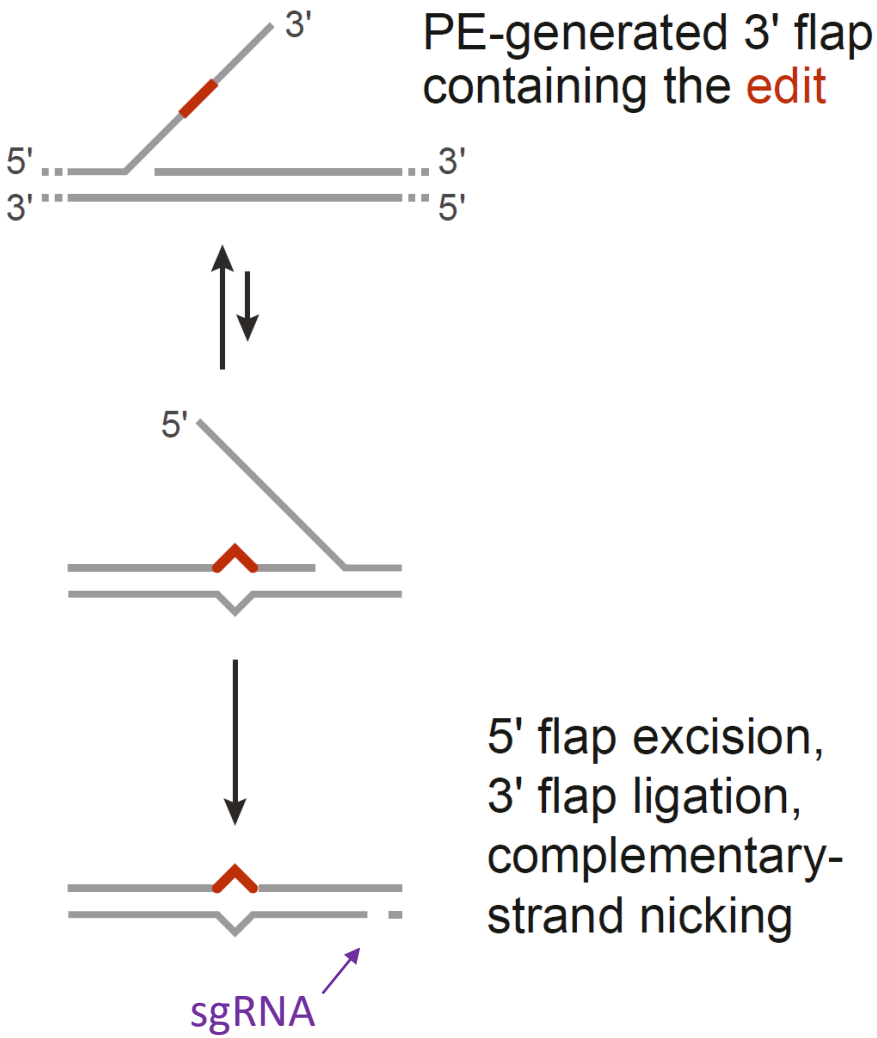


Complementary-strand nick (PE3)

The same strategy (using a Cas9 nickase to edit the unmodified strand) was previously found to improve the efficacy of base editors.

DNA repair is biased against nicked strands, and preferentially excises them.
(In freshly replicated DNA, both new strands contain nicks – the lagging has more than the leading strand.)

Search-and-replace genome editing without double-strand breaks or donor DNA



Unfortunately, while the PE3 system improved editing efficacy, it led to increased side effects (indels) compared to PE2.

In a final step, the authors developed **PE3b:**

- In PE3b, the nicking sgRNA is **specific to the edited sequence**, so it only nicks once the edited 3' flap has been incorporated.
- This greatly reduces the risk of doubly nicked DNA, double-strand breaks and indels.
- However, there needs to be a suitable PAM!

Cell

Article

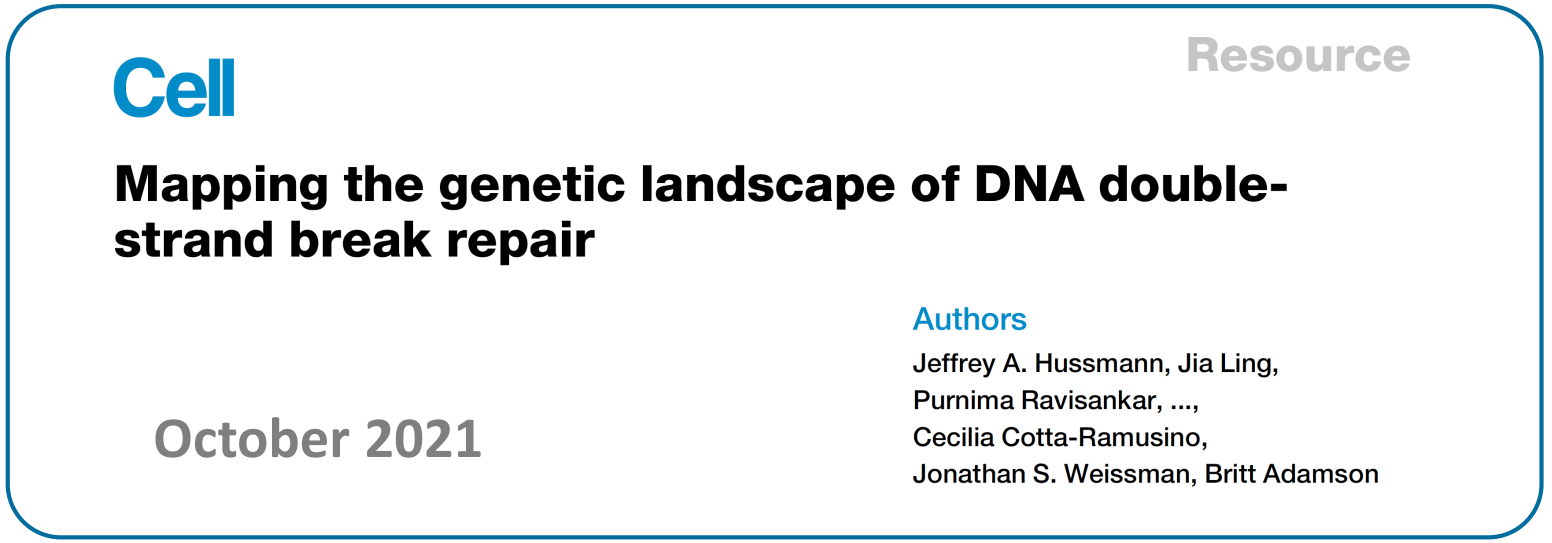
Enhanced prime editing systems by manipulating cellular determinants of editing outcomes

October 2021

Authors

Peter J. Chen, Jeffrey A. Hussmann, Jun Yan, ..., Jonathan S. Weissman, Britt Adamson, David R. Liu

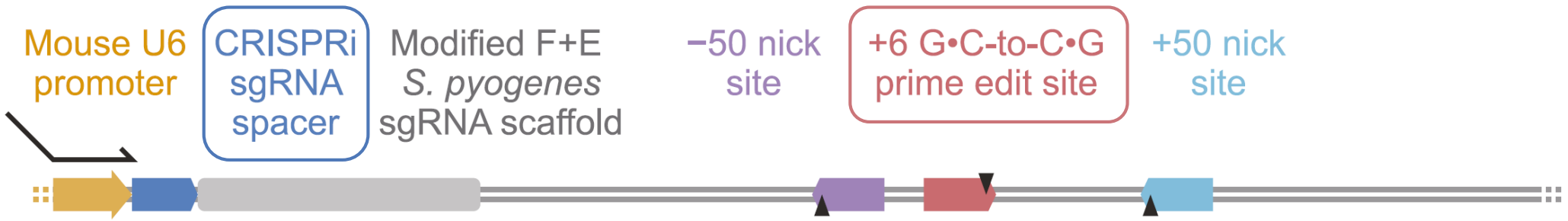
To identify determinants of prime editing outcomes, the authors teamed up with the inventors of **Repair-seq**



The developers of Repair-seq published their paper in the same issue of *Cell.*

Repair-seq features a linked CRISPRi sgRNA and a target site (where the repair occurs) on the same vector

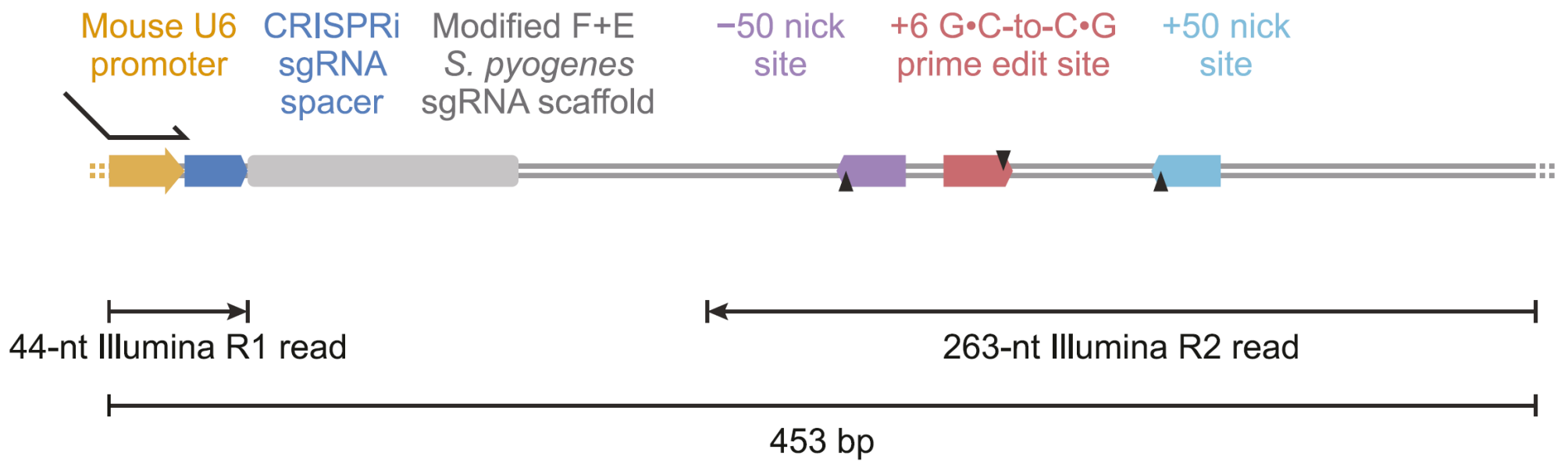
- A pooled screen is done with CRISPR inhibition (CRISPRi) with a dCas9-KRAB repressor. (In a screen for DNA repair, CRISPR-cut would cause confounding effects.)
- The CRISPRi sgRNA varies for each plasmid in the library, but the target region is constant.
- A few days after transduction of the library, the **target region is altered in all cells** (e.g. by transfection with plasmids coding for a prime editor and pegRNA against the target site).
- The **read-out** of the screen is the effect of the sgRNA on the **repair outcomes** at the target.



The authors screened for genes whose inhibition altered prime editing outcomes

The Repair-seq vector allows for sequencing of both the sgRNA sequence and the repair outcomes at the target sequence with **paired-read sequencing**.

much cheaper than long-read sequencing!

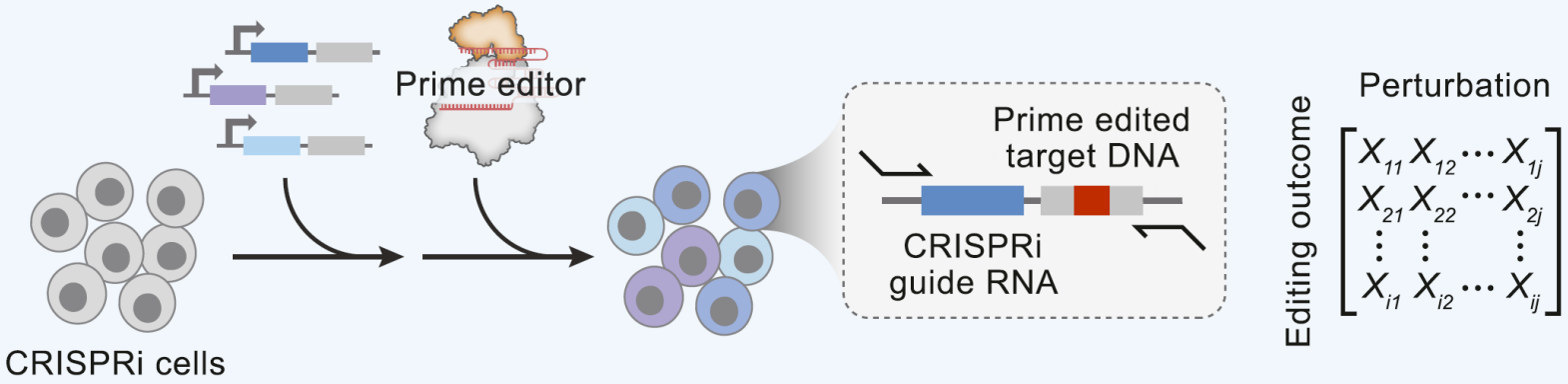


A **focussed screen** was done with a library of 476 genes, curated mostly for their roles in DNA repair.

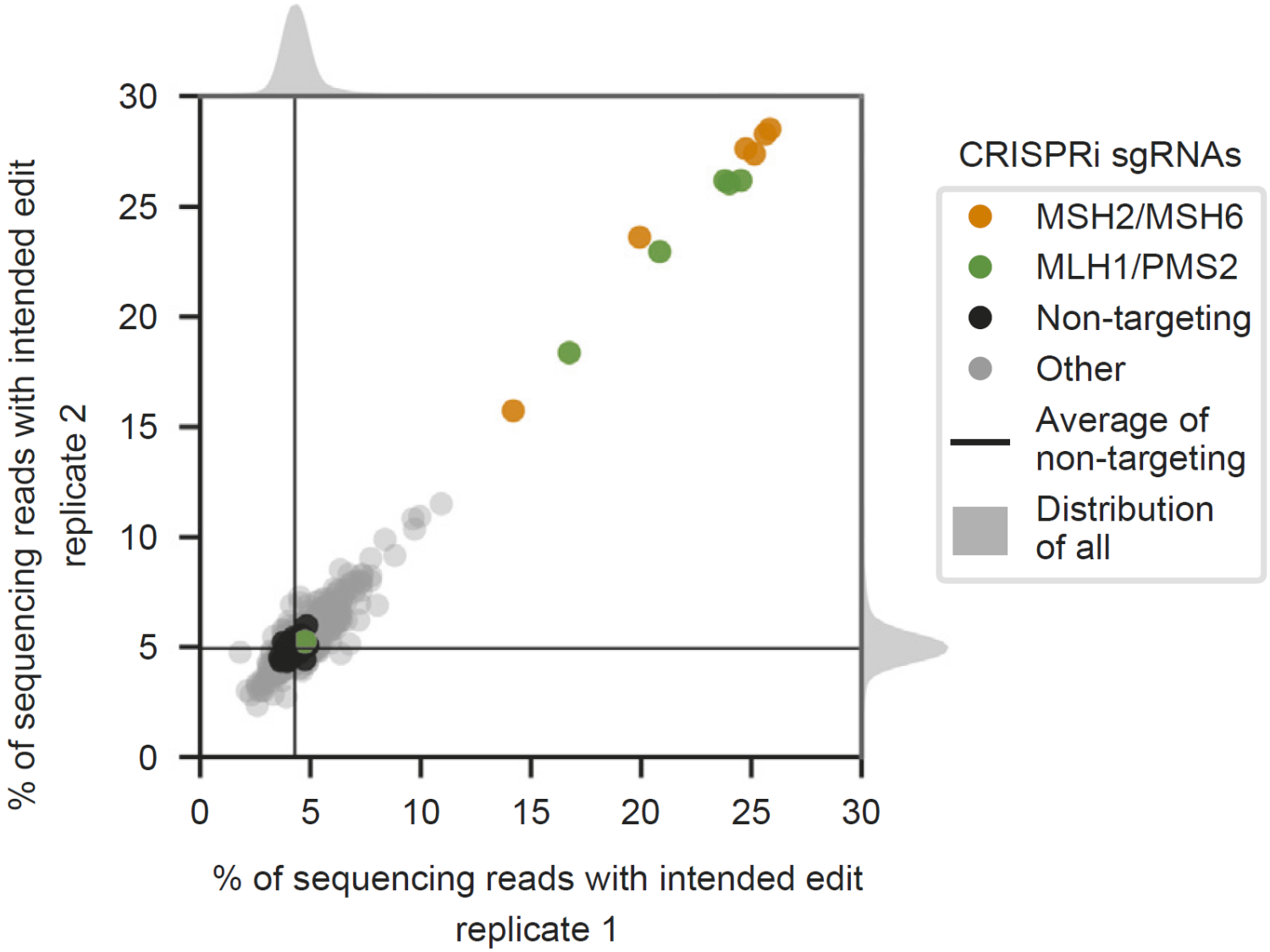
- Because of the complex read-out (repair outcome, rather than survival), more statistical power was needed.
- 3 sgRNAs per gene,
  60 non-targeting controls,
  1513 sgRNAs in total

Functional annotation class	Number of genes in CRISPRi sgRNA library
Recombinational repair	41
Double-strand break repair	36
Nucleotide-excision repair	30
Base-excision repair	23
Postreplication repair	16
Mismatch repair	16
DNA repair (other)	50
Cellular response to stimulus	4
Nuclear division	12
Cell cycle checkpoint	12
mRNA processing	21
DNA replication	23
Protein modification	24
Other	39
Regulation of transcription	46
Chromatin processes	83
Total	476

Steps of the screen



1. A CRISPRi cell line is used (e.g. HeLa cells stably expressing dCas9-KRAB).
2. The pooled CRISPRi Repair-seq library is transduced into the cells, followed by puromycin selection.
3. After **5 days**, plasmids coding for the prime editor and the pegRNA were both transfected into cells, followed by blasticidin selection for the PE protein. Prime editing occurs (or fails to occur) at the target site.
4. **3 days** later, cells were harvested for DNA extraction.



The authors looked for genes whose repression improves the efficiency of prime editing → higher percentage with the desired edit.

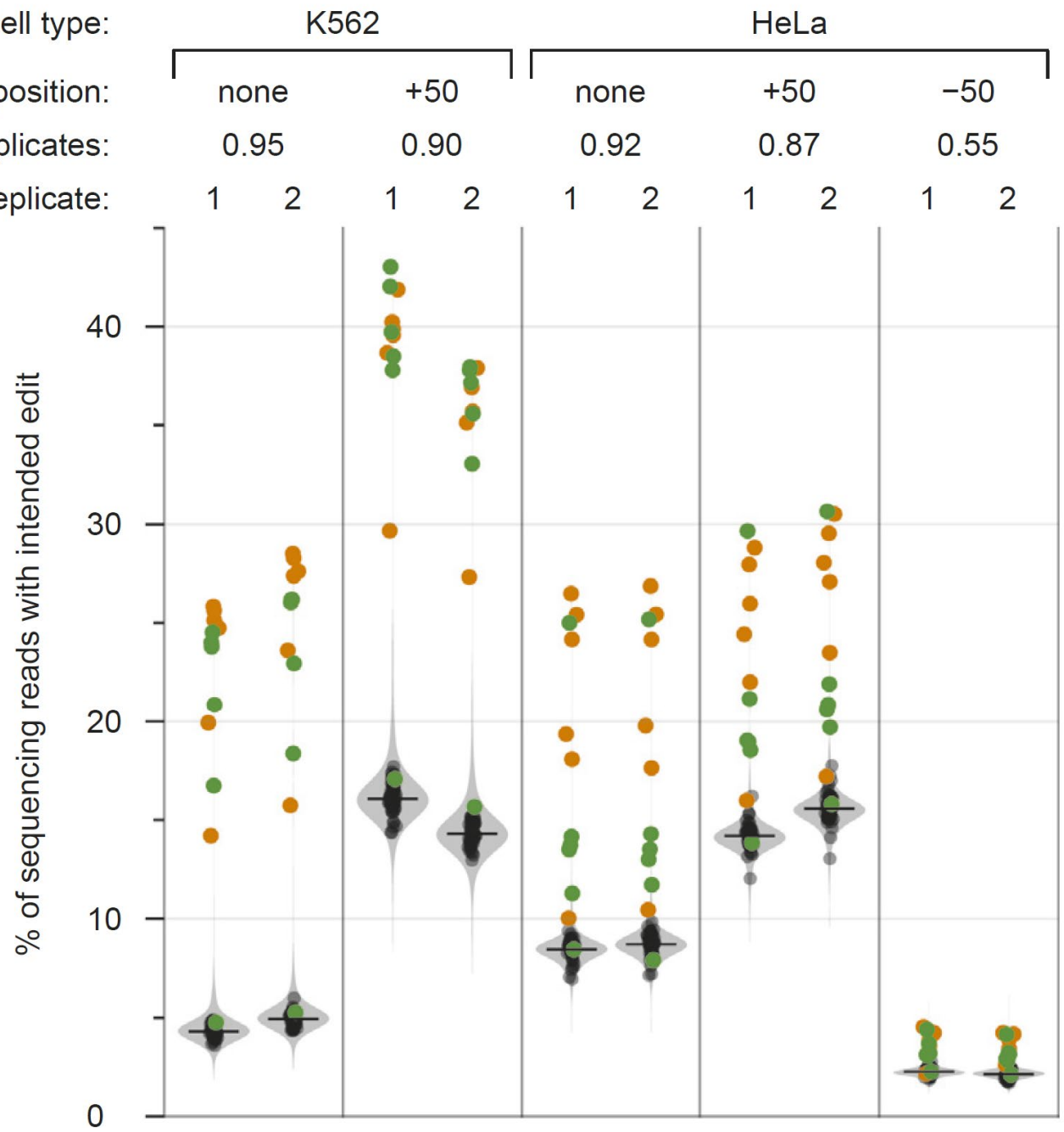
The **top 4 genes** were all components of the **mismatch repair** (MMR) pathway!

Consistent results were seen in two cell lines, and with the PE2 (plain) or PE3 (opposite-strand nick) prime editing systems.

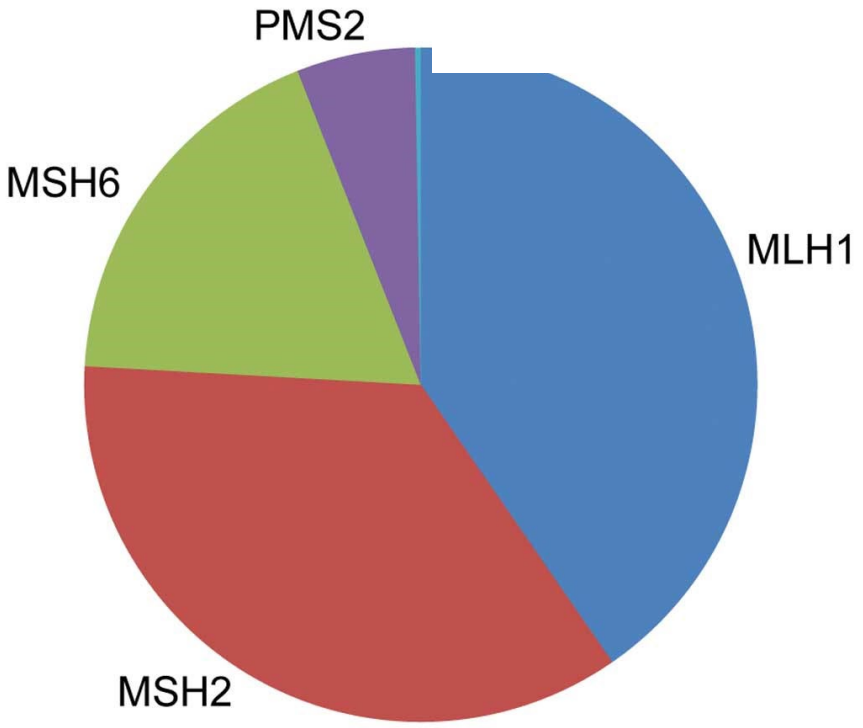
MSH2 or MSH6

MLH1 or PMS2

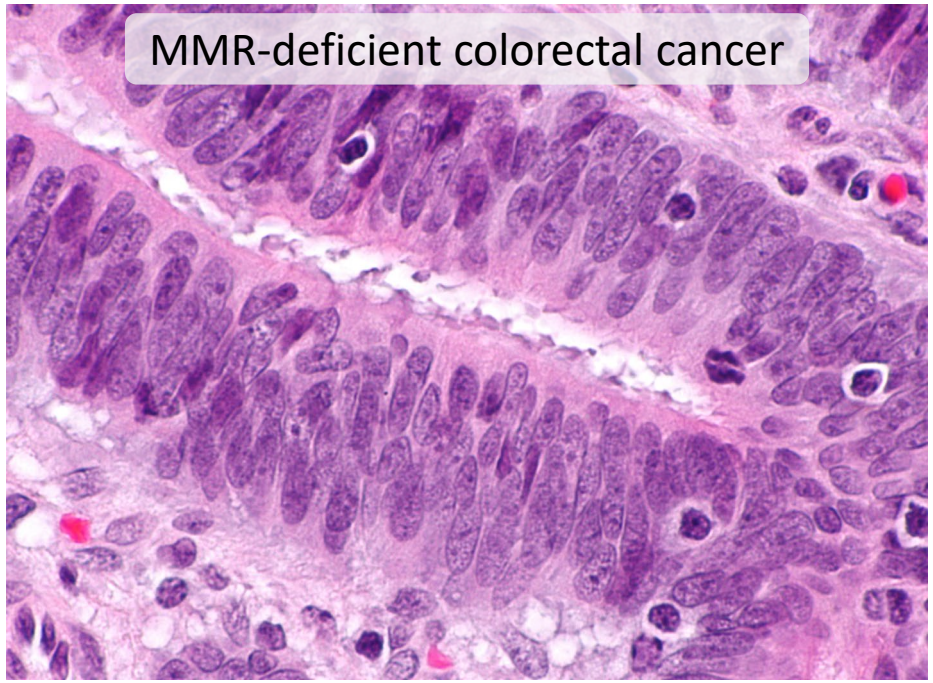
all 3 sgRNAs are shown



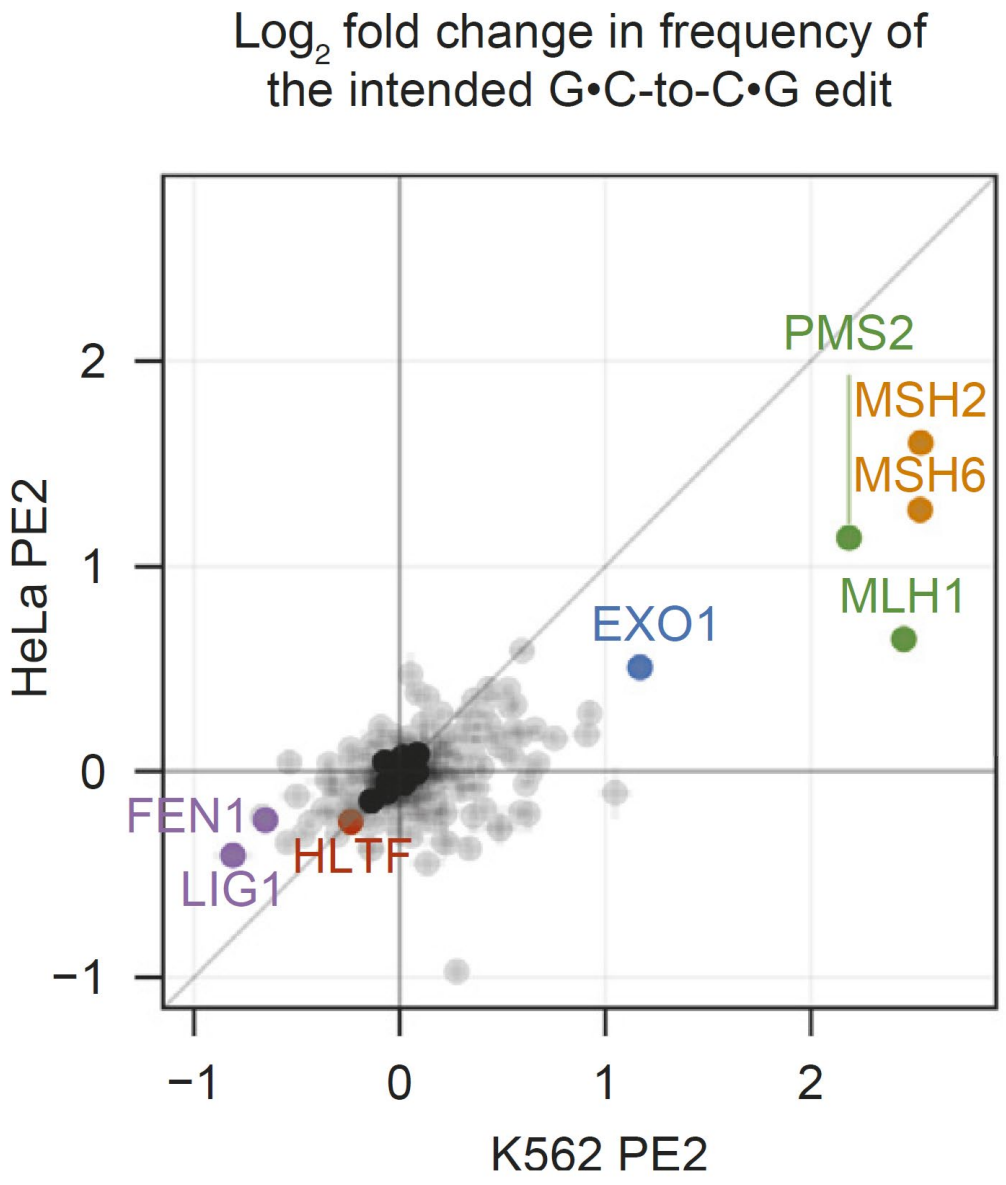
The exact same four genes are most commonly mutated in Lynch syndrome (= hereditary non-polyposis colon cancer, HNPCC)



Lagerstedt-Robinson et al., Swedish population



In Lynch syndrome, heterozygous loss-of-function mutations in MMR genes predispose to colorectal cancer (lifetime risk 52-58%) and endometrial cancer (25-60%), among others.



EXO1 (exonuclease 1), which is also involved in mismatch repair, had a weaker effect.

Conversely, two genes were also found whose inhibition **reduced** PE efficiency:

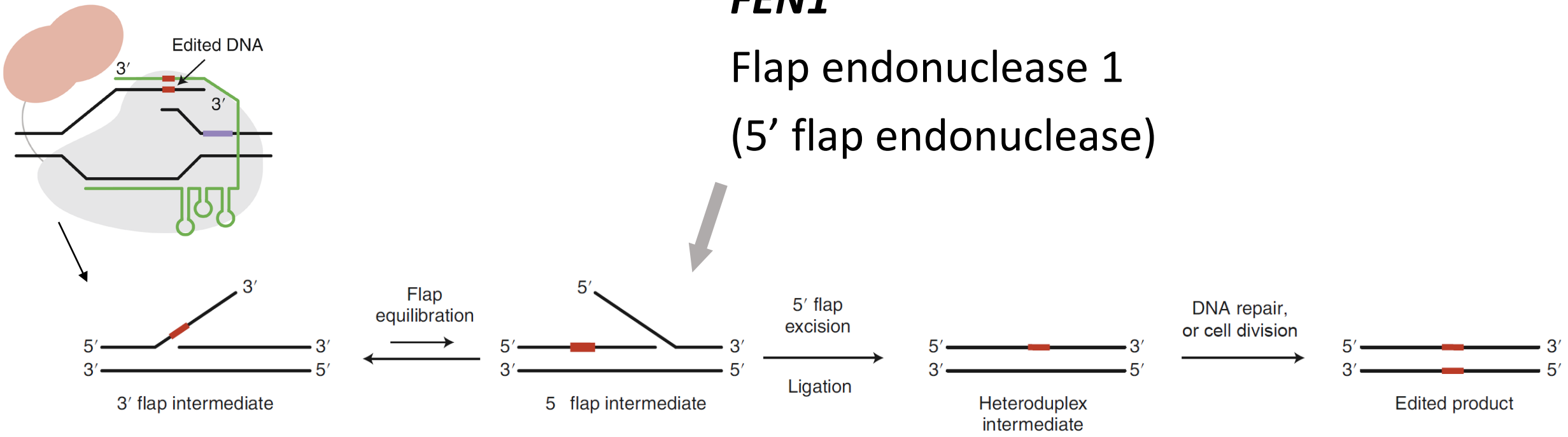
FEN1 Flap endonuclease 1 (5’ flap endonuclease)

LIG1 DNA ligase 1 (nick ligase)

FEN1

Flap endonuclease 1

(5’ flap endonuclease)

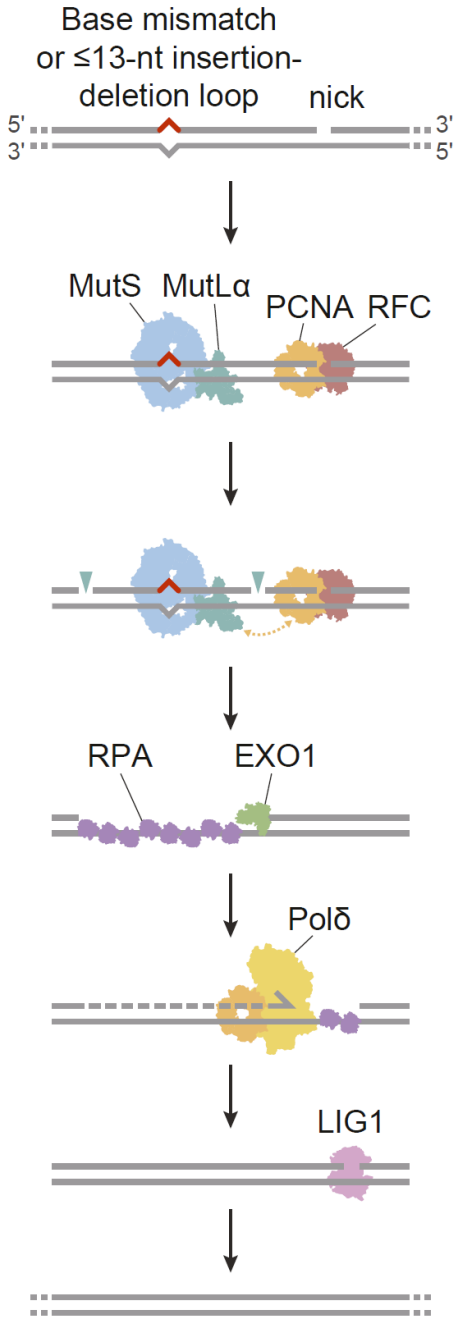


LIG1

DNA ligase 1

(nick ligase)

Mismatch repair (MMR)



MutSα (MSH2–MSH6) or MutSβ (MSH2–MSH3) recognizes the mismatch or insertion-deletion loop and recruits MutLα (PMS2–MLH1). RFC assymetrically loads PCNA at the nicked DNA.

PCNA signals MutLα to incise the nick-containing DNA strand around the heteroduplex.

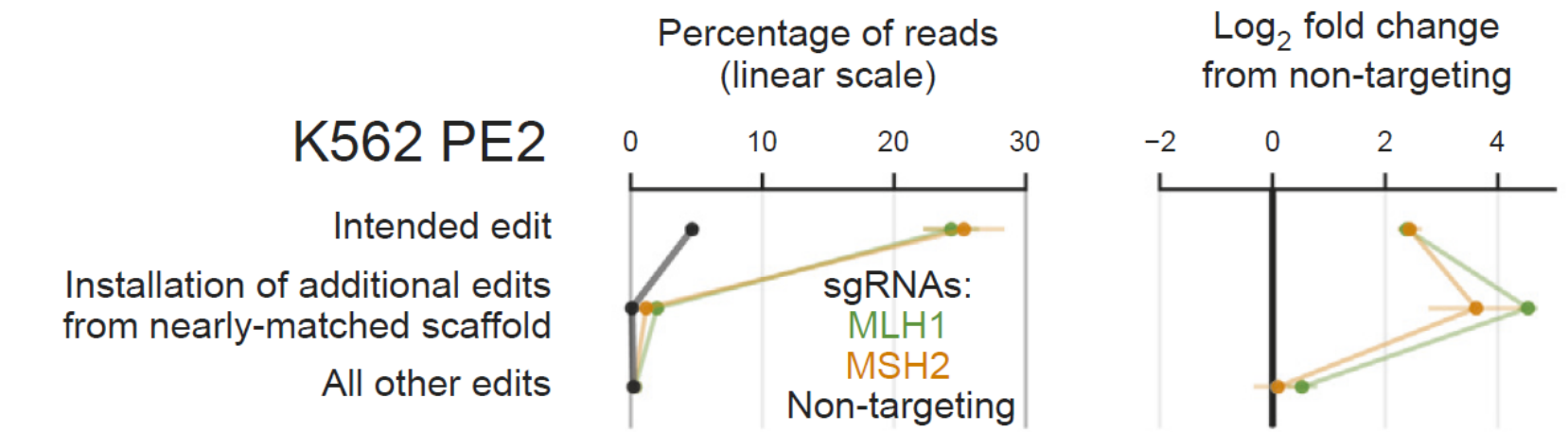
MS2-MS6 recognizes a mismatch, and recruits **PMS2-MLH1**, which nicks the DNA on both sides of the heteroduplex.

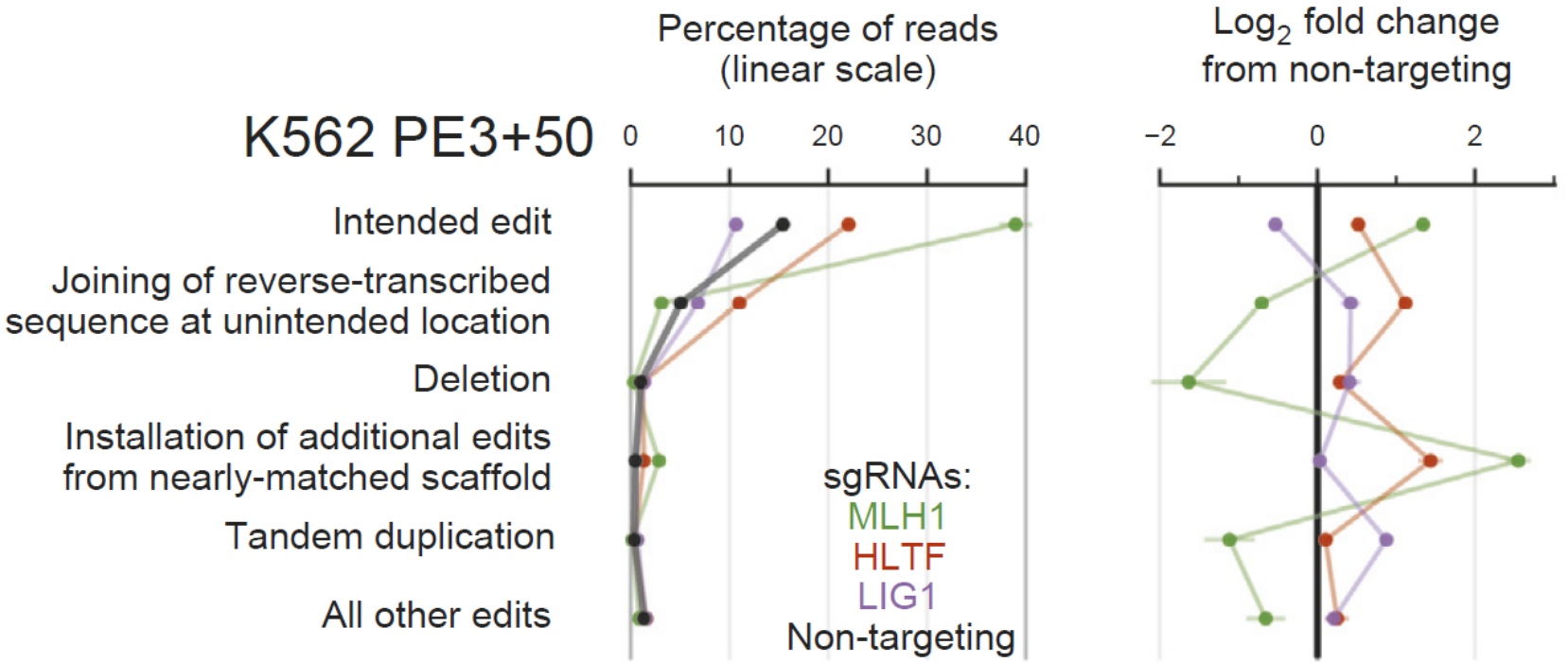
EXO1 excises the nicked strand from these incisions and the resulting ssDNA is bound by RPA.

EXO1 removes the nicked strand.

Polδ resynthesizes the excised strand from the exposed 3' DNA end, using the non-nicked strand as a template.

LIG1 seals the remaining nick.

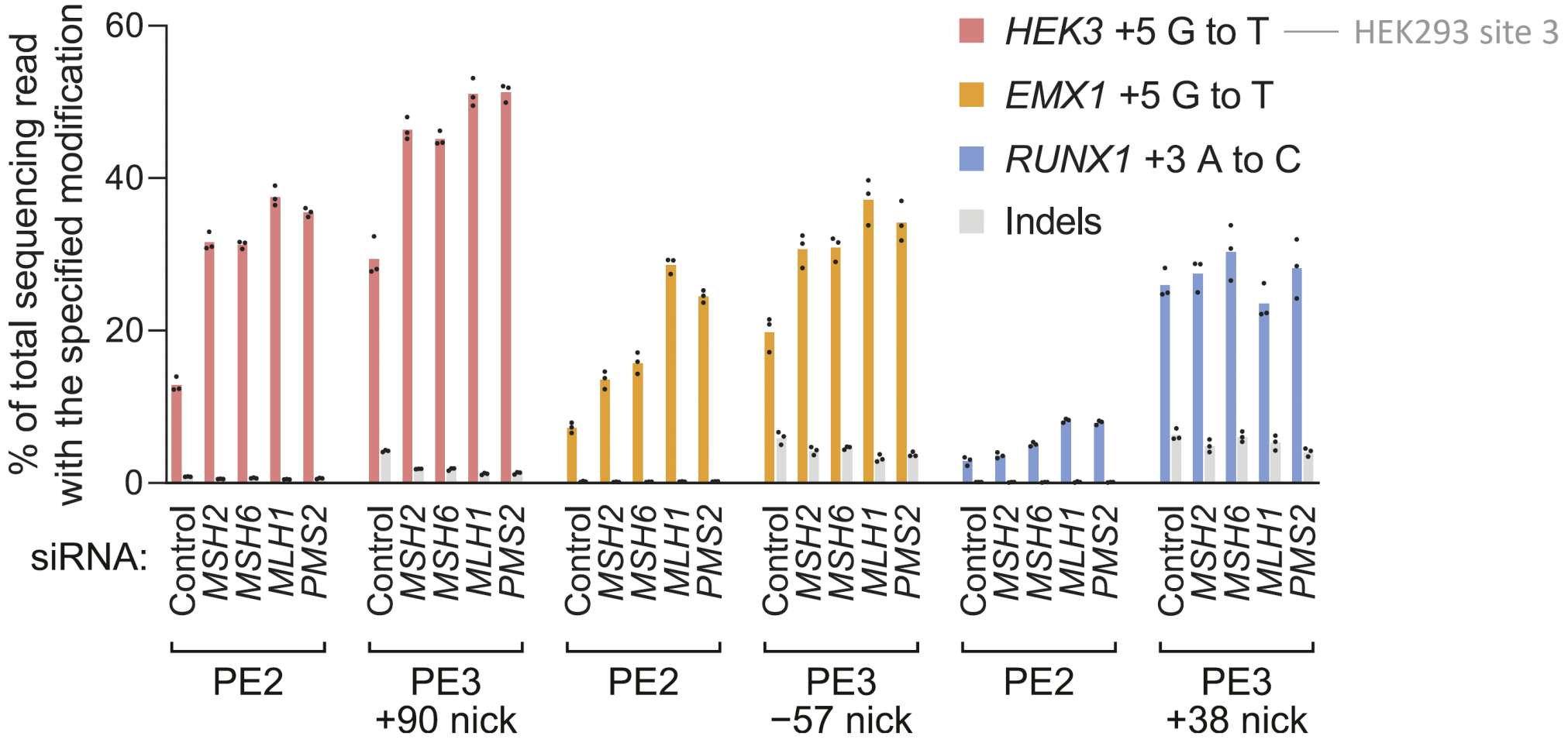


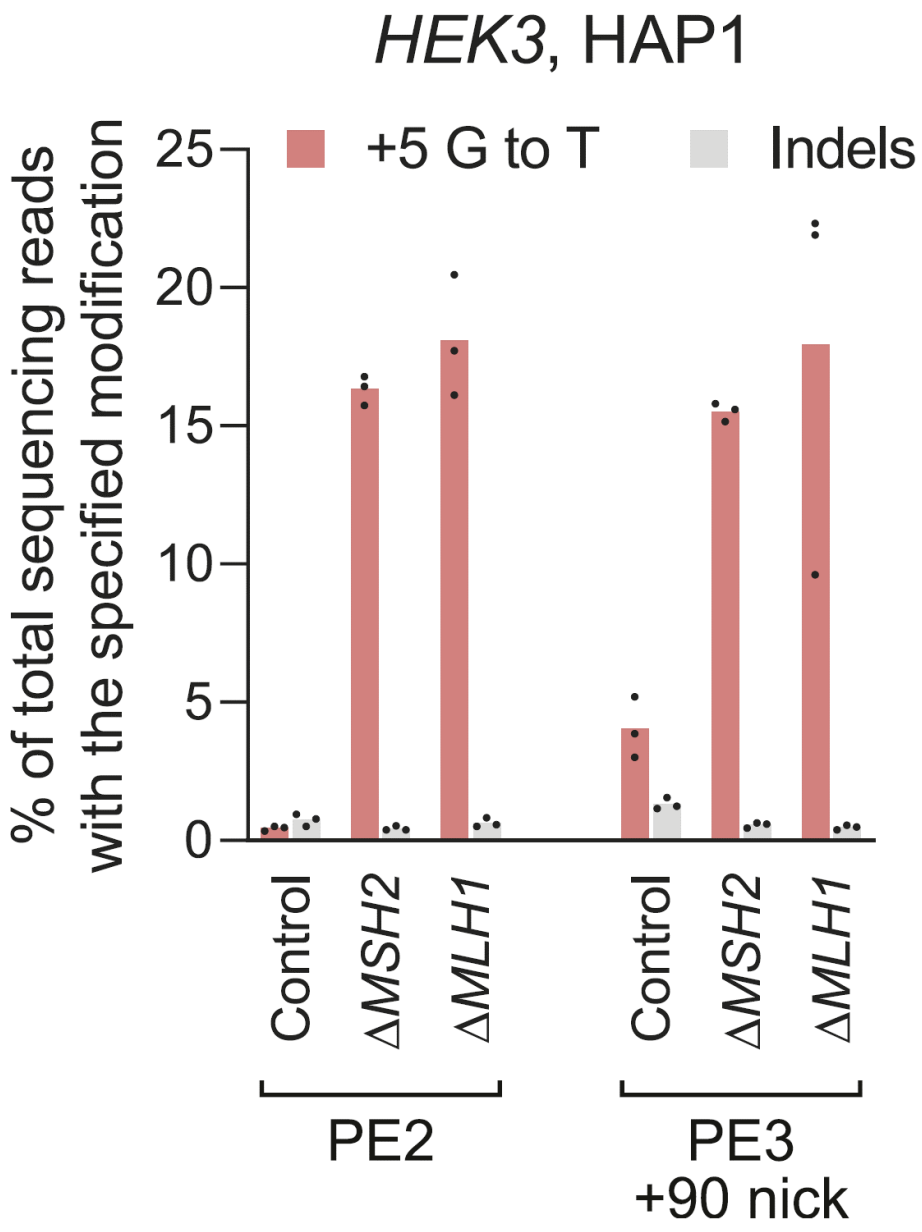


In the PE2 system (no opposite-strand nick), there was a >4-fold increase in efficiency with CRISPRi for MLH1 or MSH2.

In the PE3 system, there was still a >2.5-fold increase in efficiency. Importantly, there was a **reduction in unintended outcomes** too (deletions, tandem duplications)!

The improvement in **prime editing efficiency** and **outcome purity** was validated with siRNA knockdown at 3 new loci in HEK293 cells

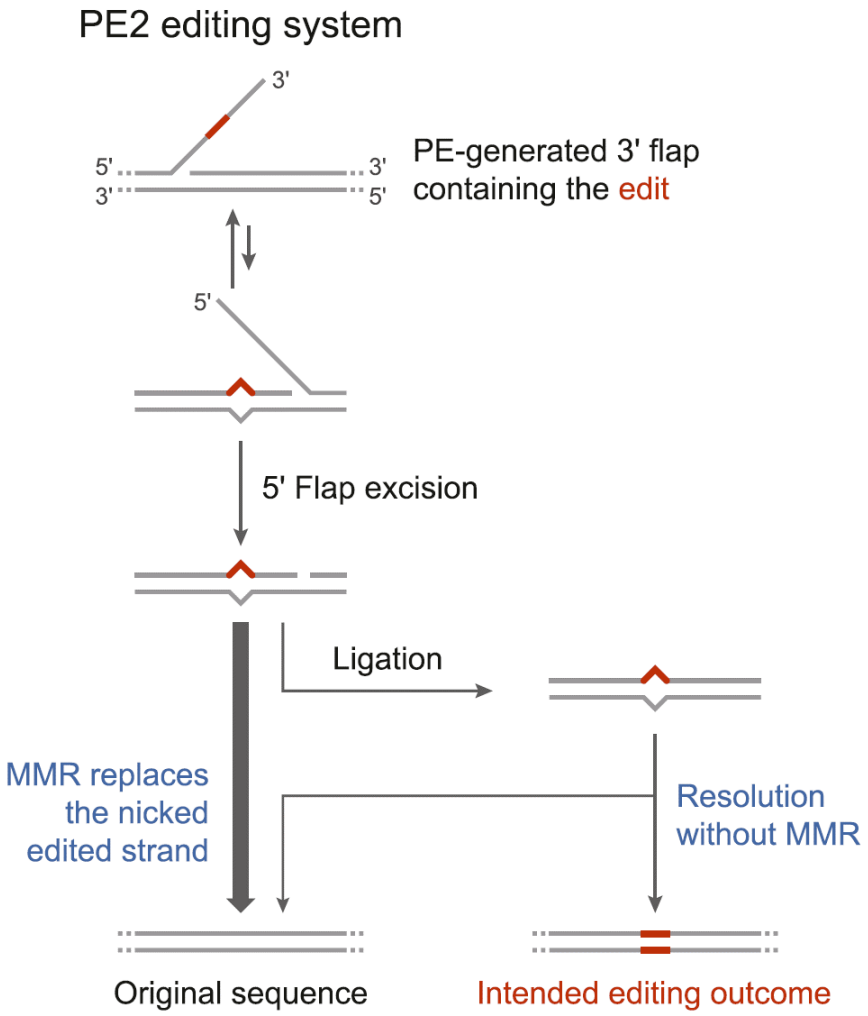




Prime editing efficiency and purity was also improved in **stable knockout** cell lines for MSH2 or MLH2.

HAP1 = human haploid cell line

The authors propose a model for how MMR impacts prime editing



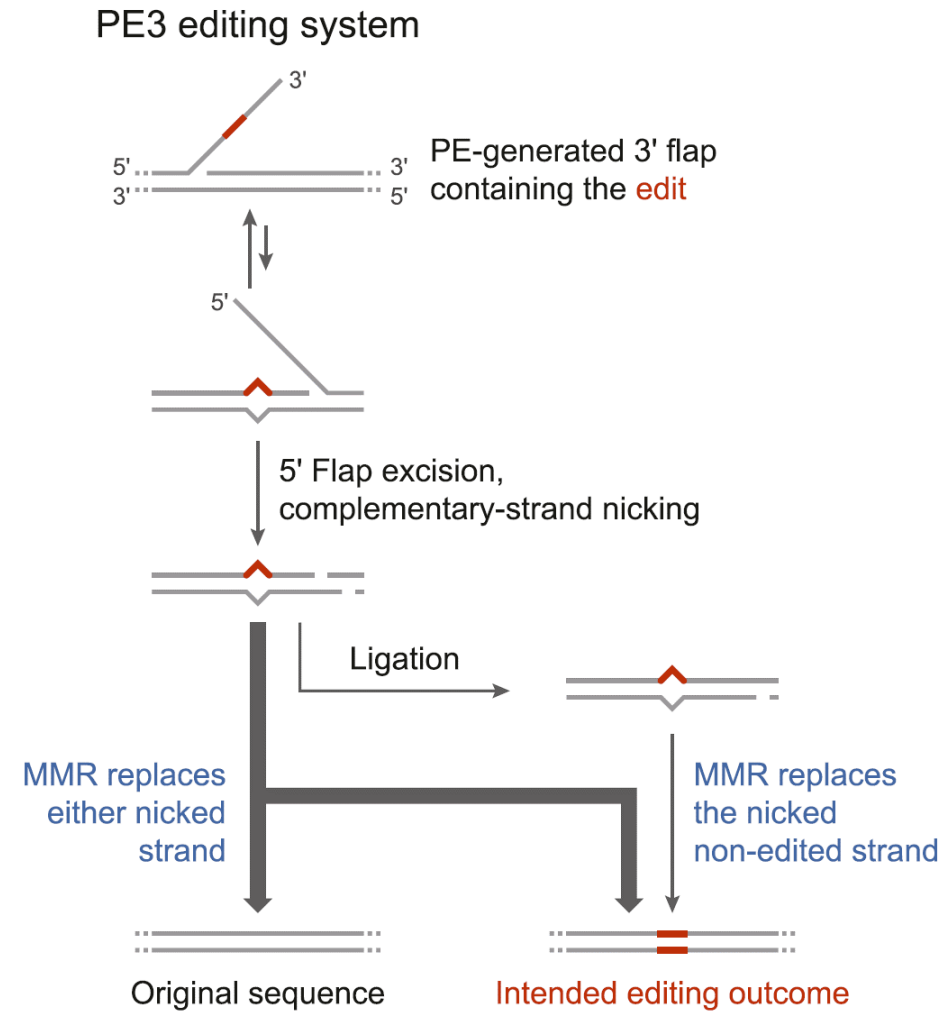
With PE2, in the absence of a nick on the opposite strand, MMR is strongly biased towards excising the edited strand, which has a nick.

→ Prime editing is inefficient.

The authors propose a model for how MMR impacts prime editing

With PE3, the additional nick on the opposite strand makes it a target for removal by MMR.

This explains the increased efficacy of PE3 compared to PE2.

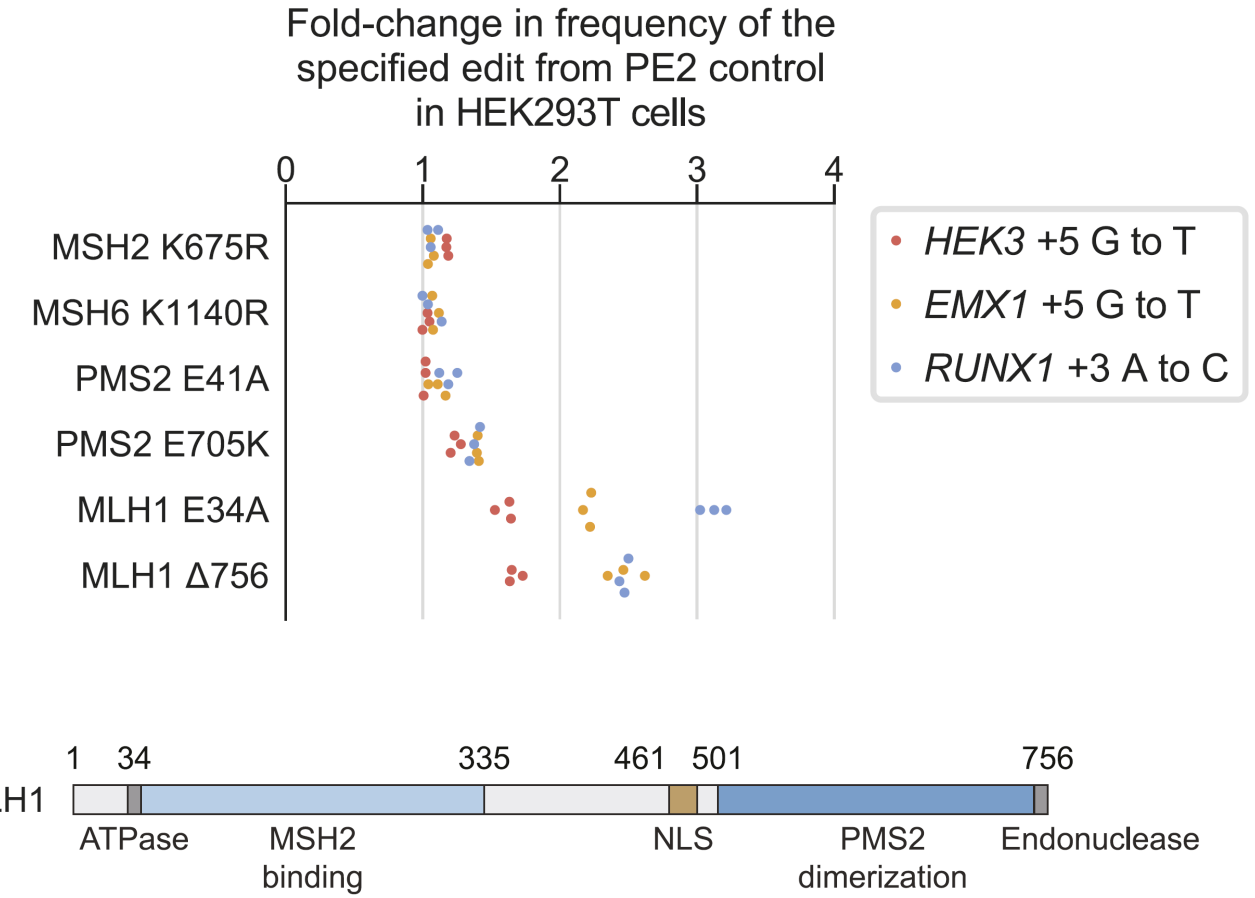


MMR in prime editing

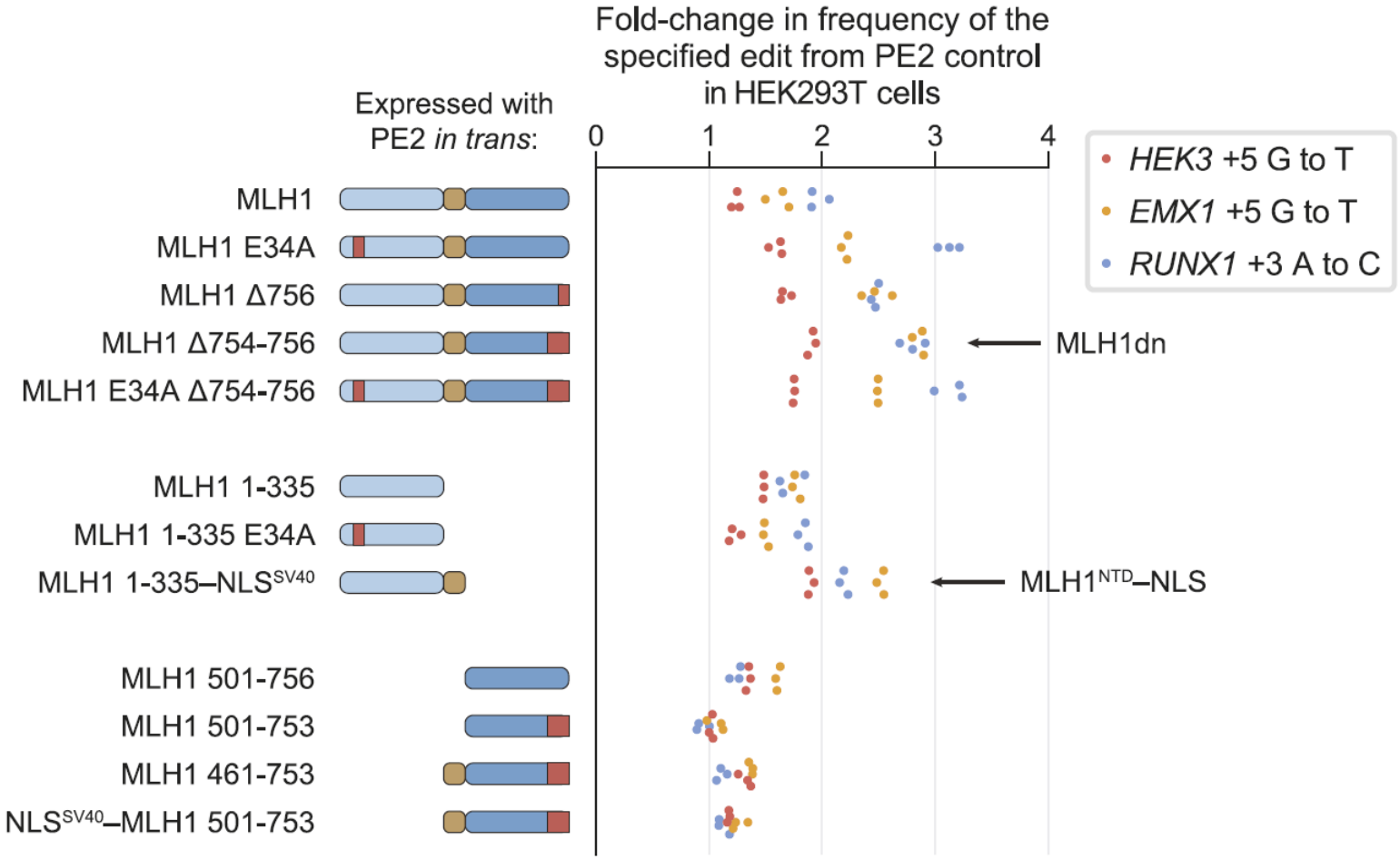
- It is clear why the absence of MMR would increase the efficacy of the PE2 system.
- In the PE3 system, it is less clear. It seems that unbiased MMR is better than biased MMR, but **no MMR** is best for efficacy!
  - The incidence of unintended outcomes (indels) is also improved without MMR. **MMR may cause double-strand breaks** in a minority of cases.
- It remains unclear how the heteroduplex is resolved in the absence of MMR…

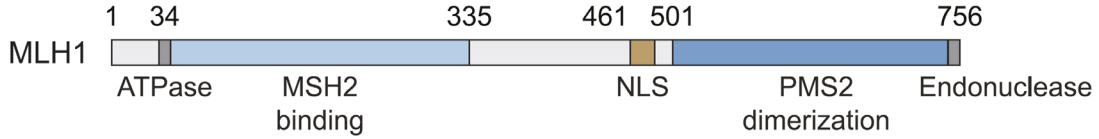
The authors use this knowledge to engineer PE3 and PE5 systems

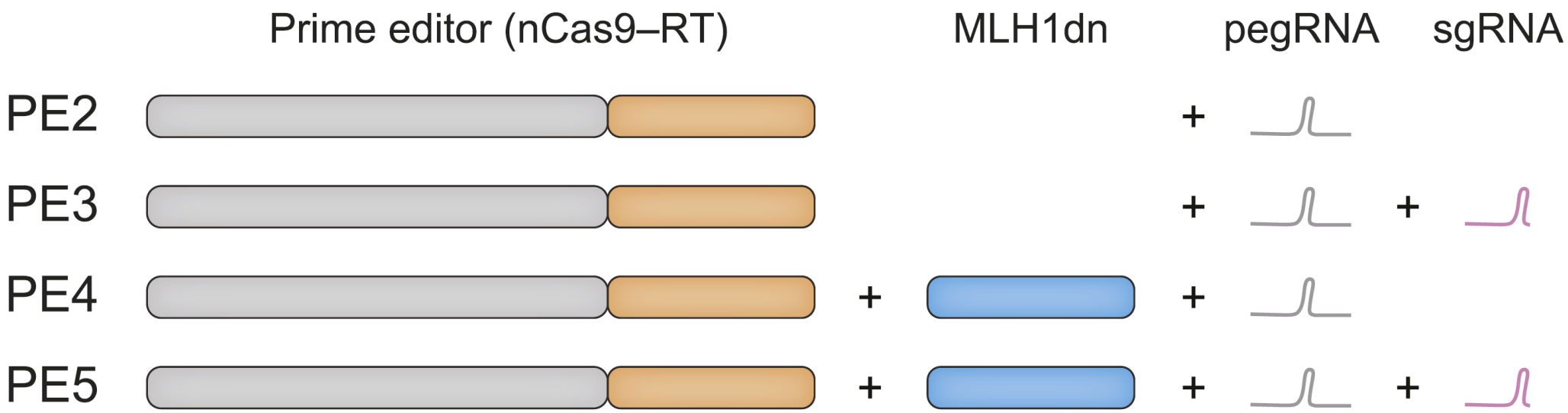
- Knockdown of MMR genes 3 days prior to prime editing improved outcomes.
- However, co-transfection of the siRNAs simultaneous to PE had no effect!
- The authors aim to create **dominant-negative** mutants of MMR proteins, to be co-delivered with the prime editor.



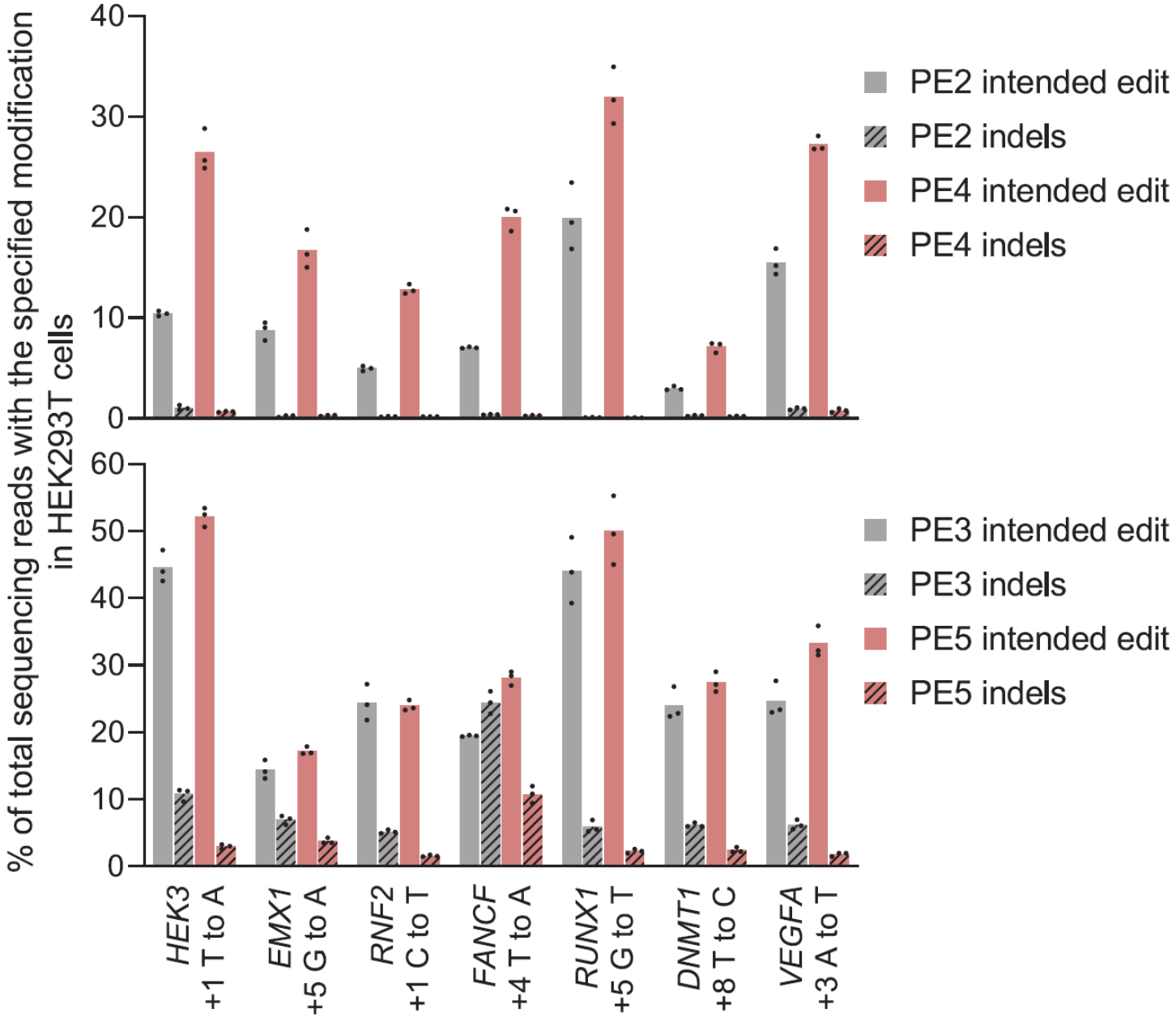
MLH1 with a defective endonuclease domain (Δ754-756) performed the best → **MLH1dn**



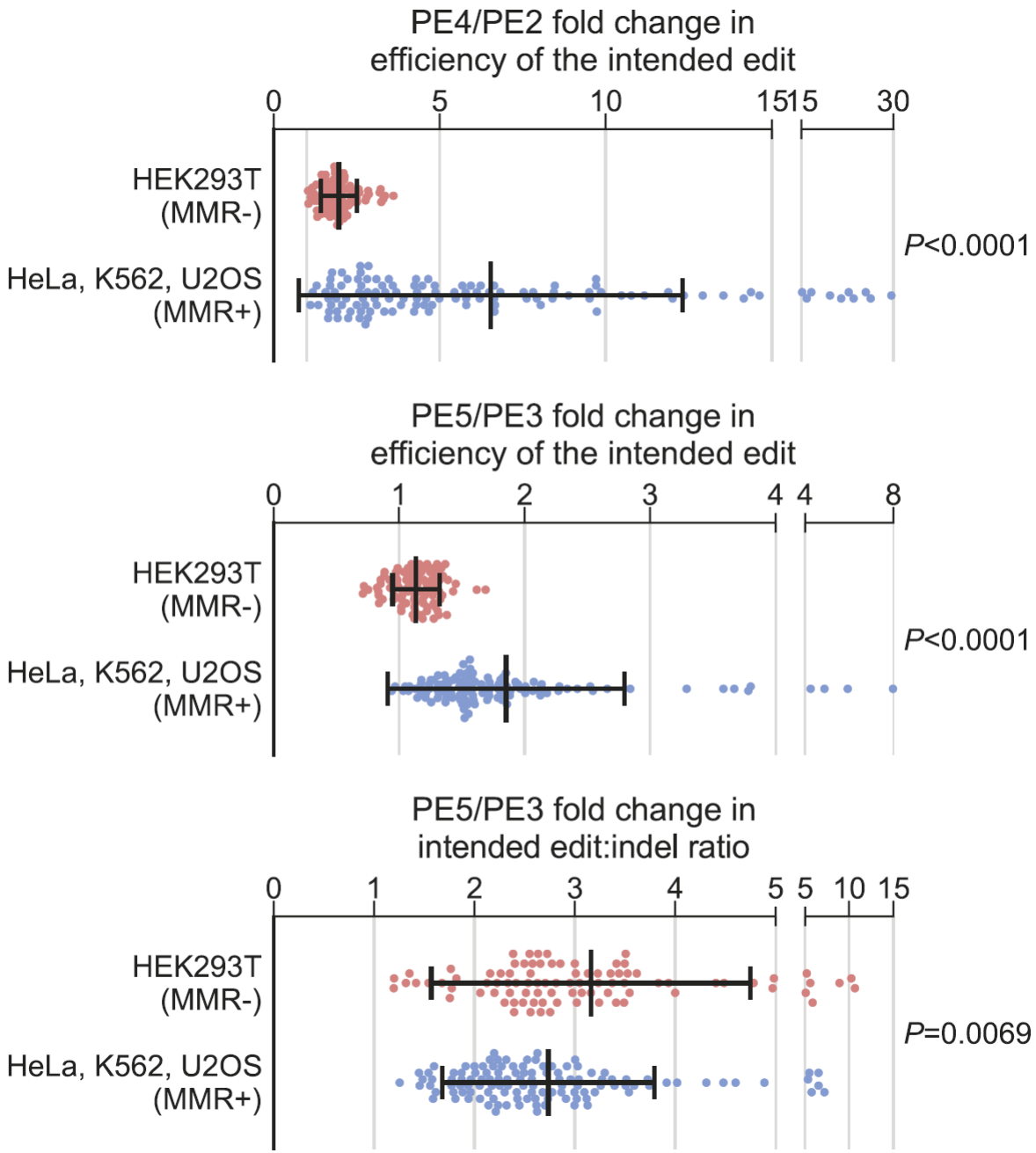




Expression plasmids for these elements are co-transfected.



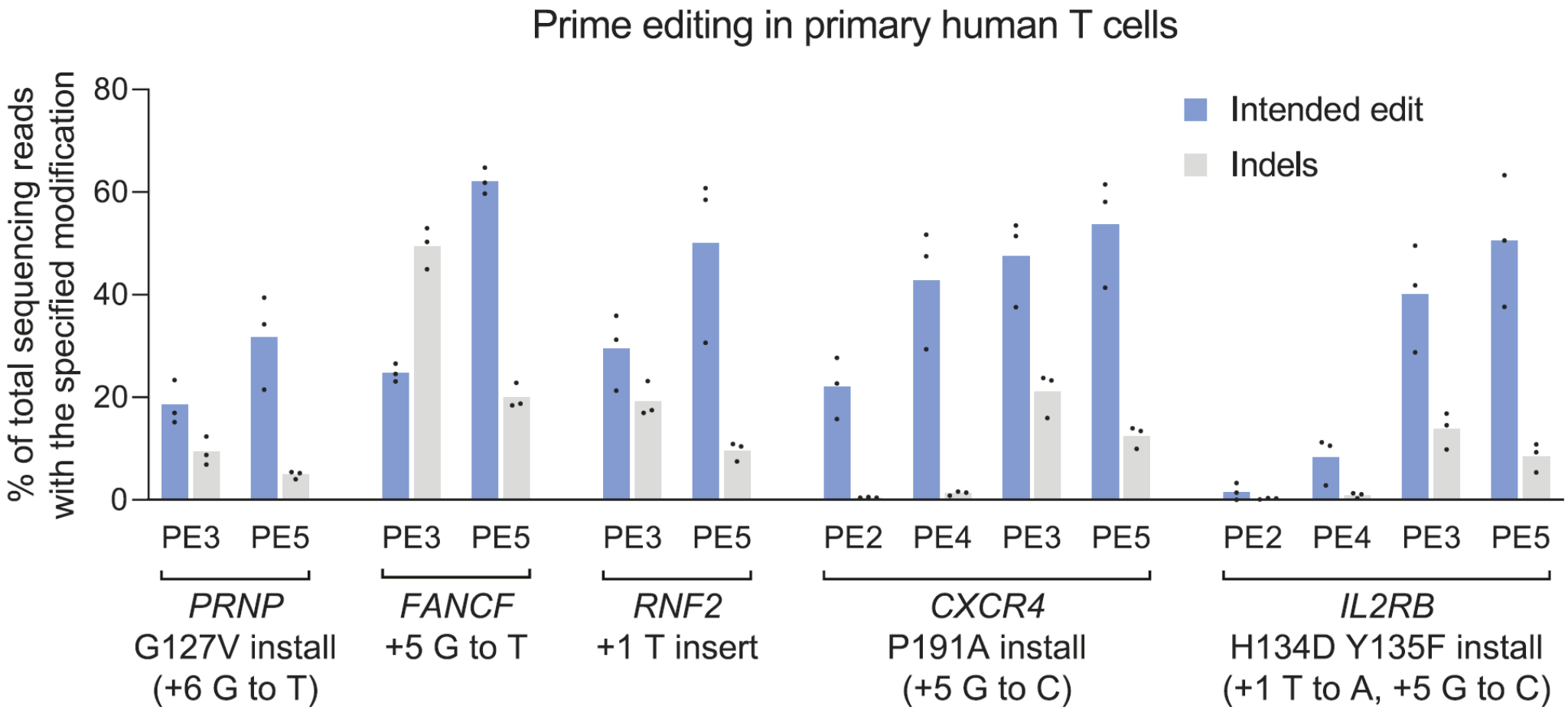
In PE3 and PE5, MLH1dn increases efficacy while improving purity.

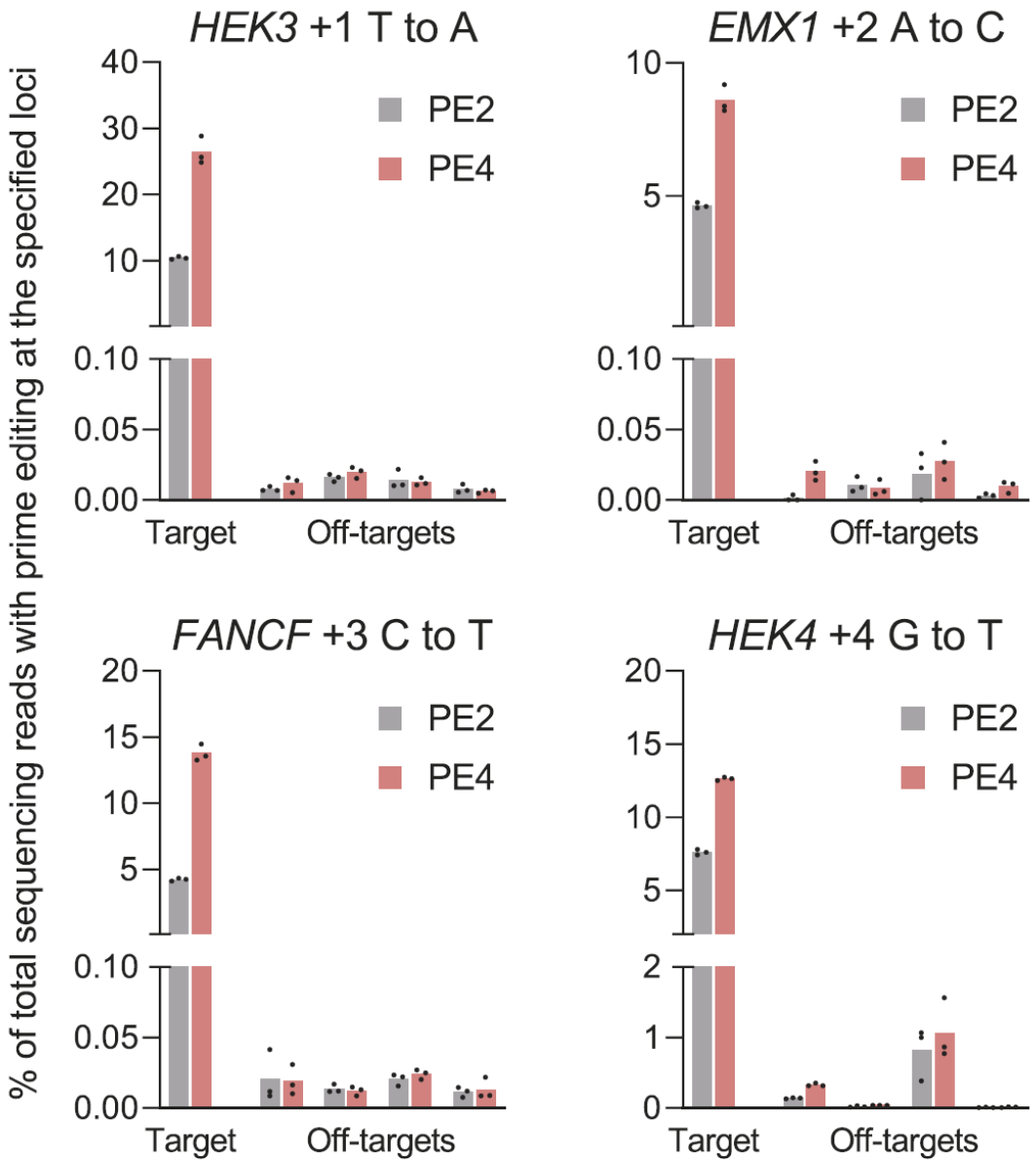


HEK293 cells are partially MMR-deficient due to hypermethylation of the *MLH1* promoter.

Accordingly, MLH1dn provides an **even greater boost** for prime editing efficacy in MMR-competent cell lines!

MLH1dn improved prime editing outcomes in primary T cells (and iPSCs)

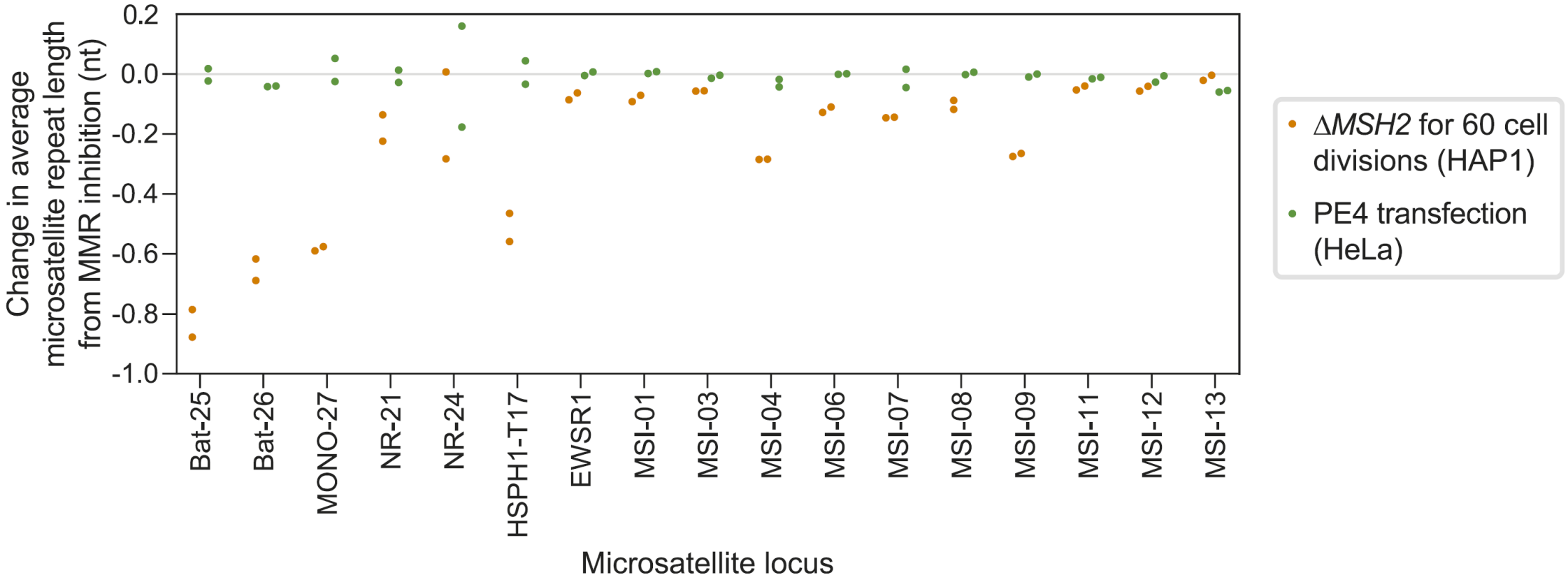




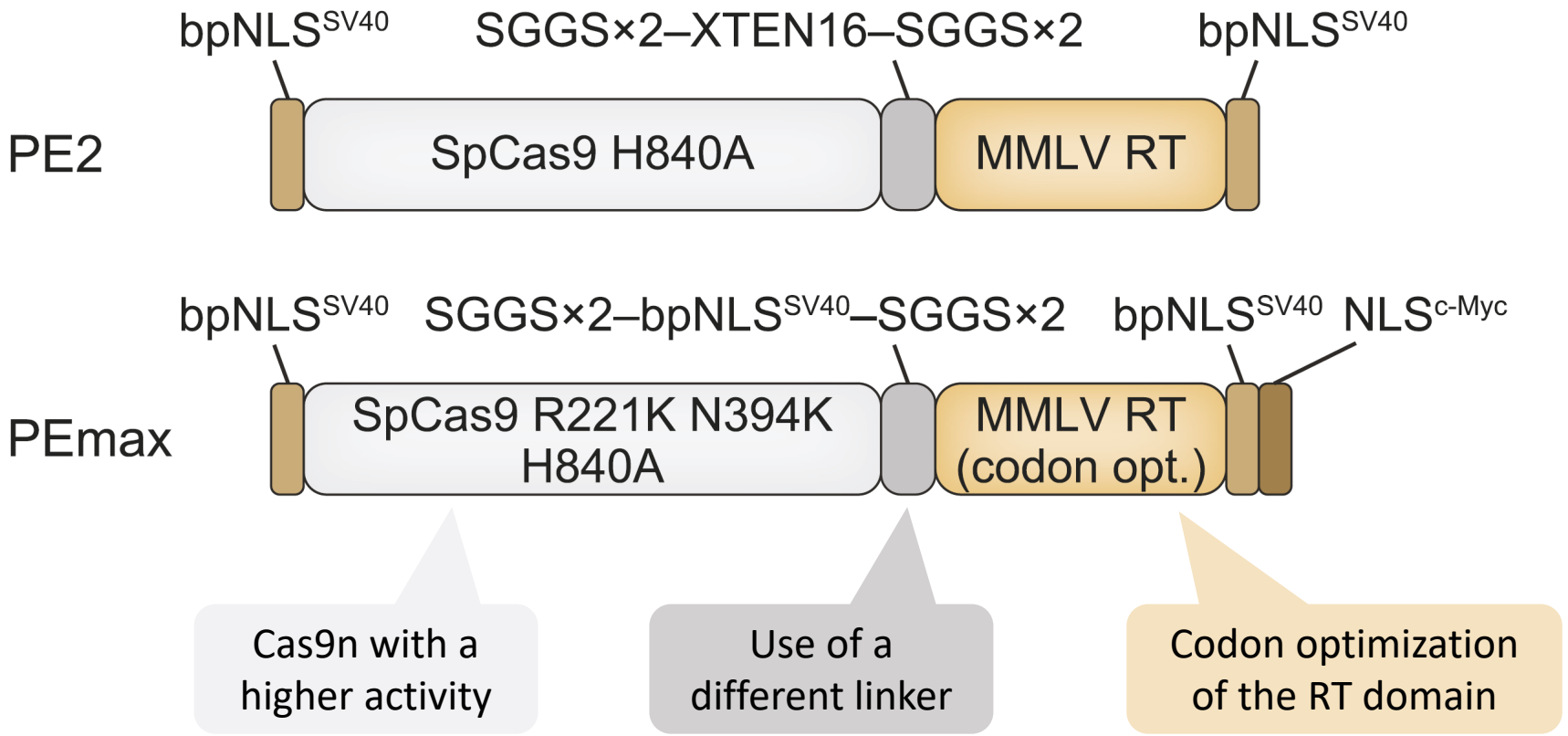
Off-target effects are known to be less of an issue with prime editing than Cas nucleases.

MLH1dn did not cause an increase in off-target effects.

Safety check: MLH1dn did not cause microsatellite instability



In a parallel effort, the authors tested many variants to establish an optimized prime editor protein (PEmax)



The same group also developed engineered pegRNAs.

Might these synergize with MLH1dn?

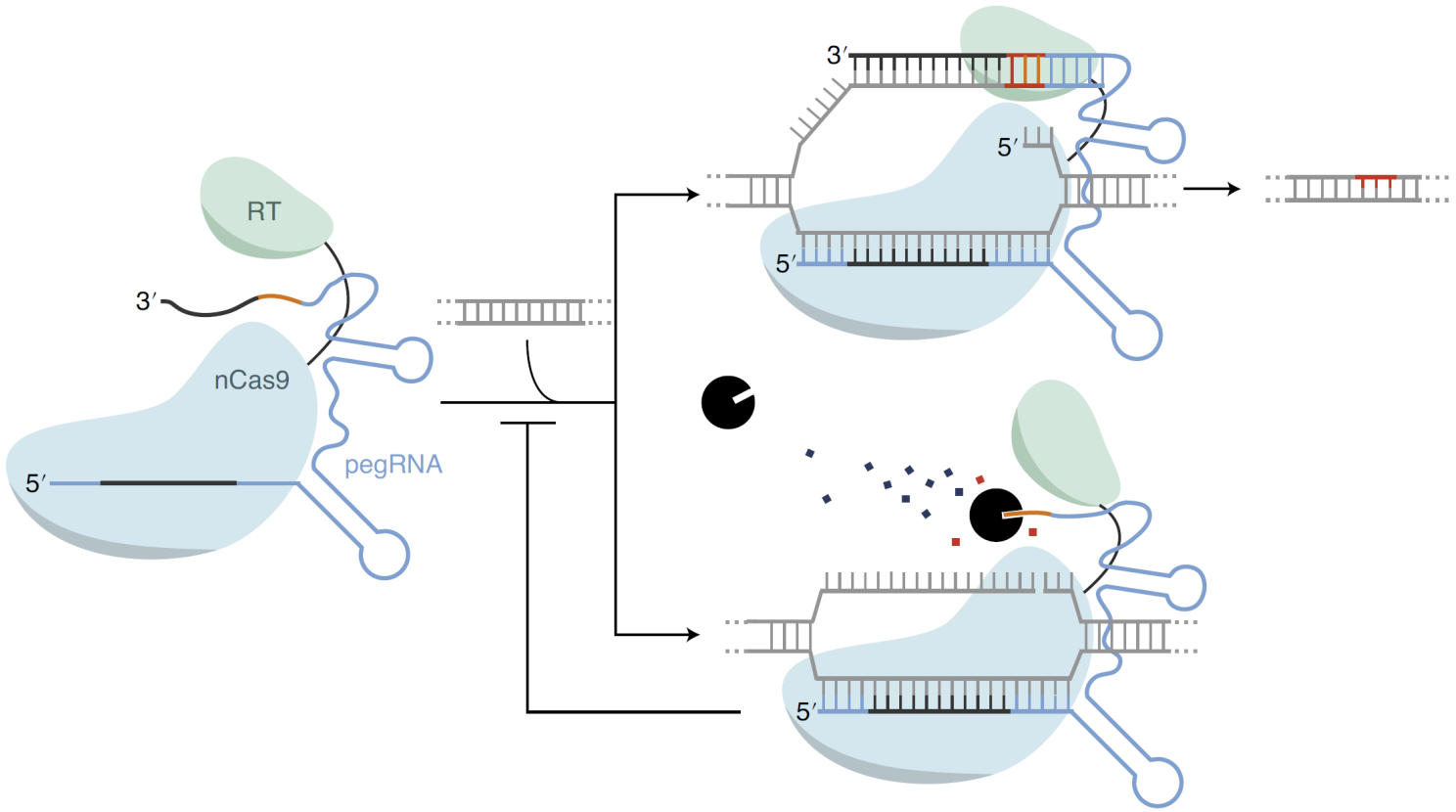
nature biotechnology

Engineered pegRNAs improve prime editing efficiency

James W. Nelson[1,2,3,4], Peyton B. Randolph[1,2,3,4], Simon P. Shen[1,2,3], Kelcee A. Everette[1,2,3], Peter J. Chen[1,2,3], Andrew V. Anzalone[1,2,3], Meirui An[1,2,3], Gregory A. Newby[1,2,3], Jonathan C. Chen[1,2,3], Alvin Hsu[1,2,3] and David R. Liu[1,2,3] ✉

October 2021

Engineered pegRNAs

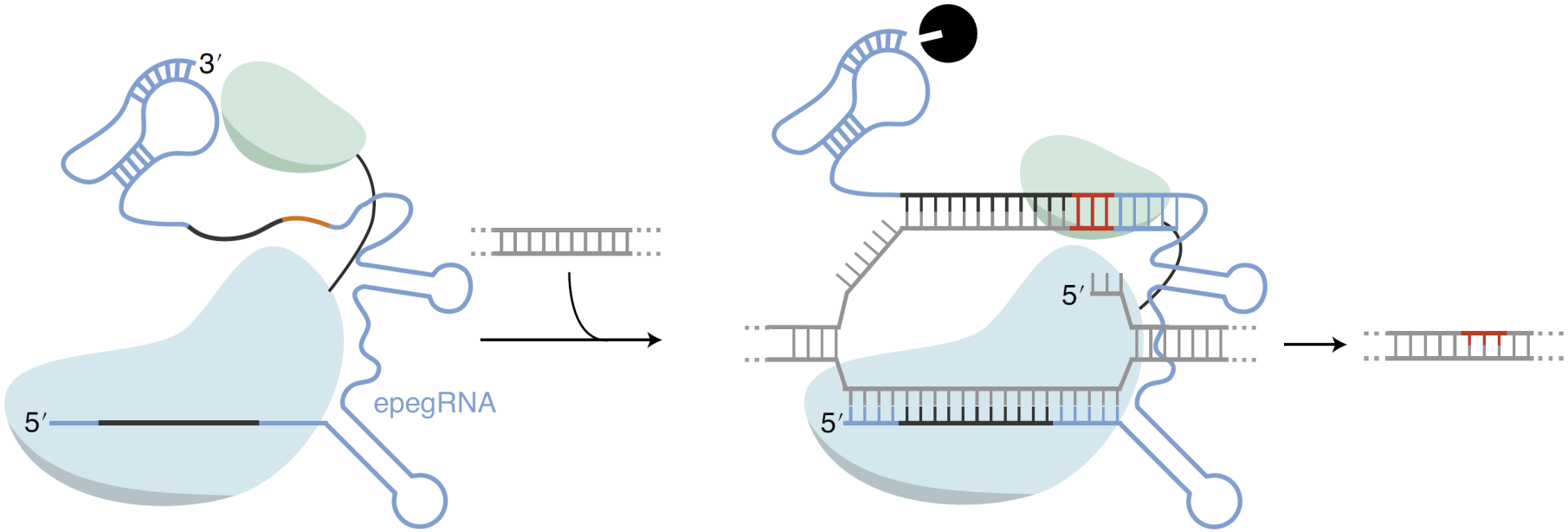


It was hypothesized that endogenous exonucleases can digest the unprotected 3’ end of pegRNAs.

Such pegRNAs are ineffective and compete for the prime editor.

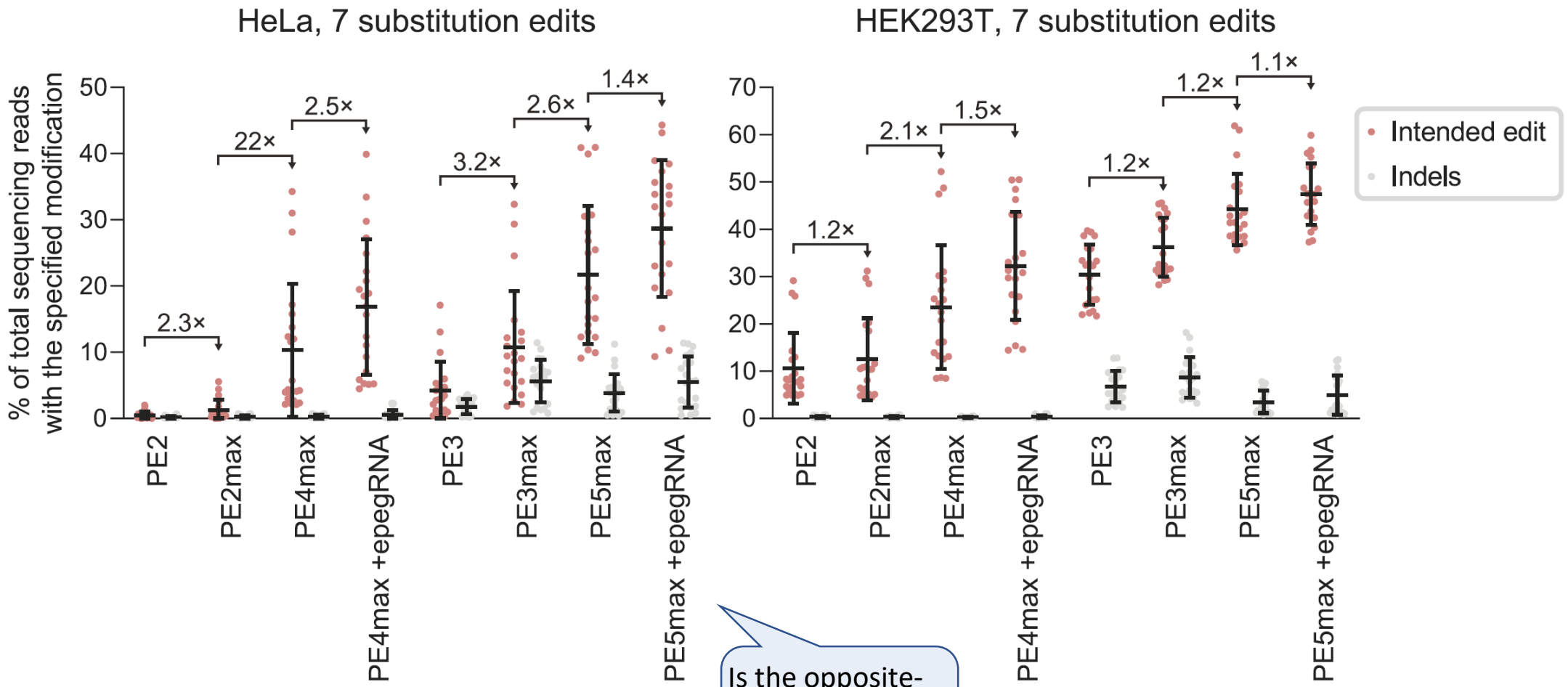
Engineered pegRNAs provide the solution.

Engineered pegRNAs



In epegRNAs, an RNA motif that produces a **small hairpin structure** is added to the 3’ end. If a suitable linker is chosen (that does not interfere with the primer-binding sequence), efficacy is improved.

In combination, there was a striking improvement in efficacy!



Conclusions

- In a focussed CRISPR screen using Repair-seq, Mismatch repair (MMR) components were discovered to limit both the efficacy and accuracy of prime editing.
- Co-transfection of a dominant-negative mutant of the MMR gene *MLH1,* in combination with other improvements, led to a greatly improved PE system.

  → Increases the advantage over homology-directed repair for efficacy and purity!

Cell

Resource

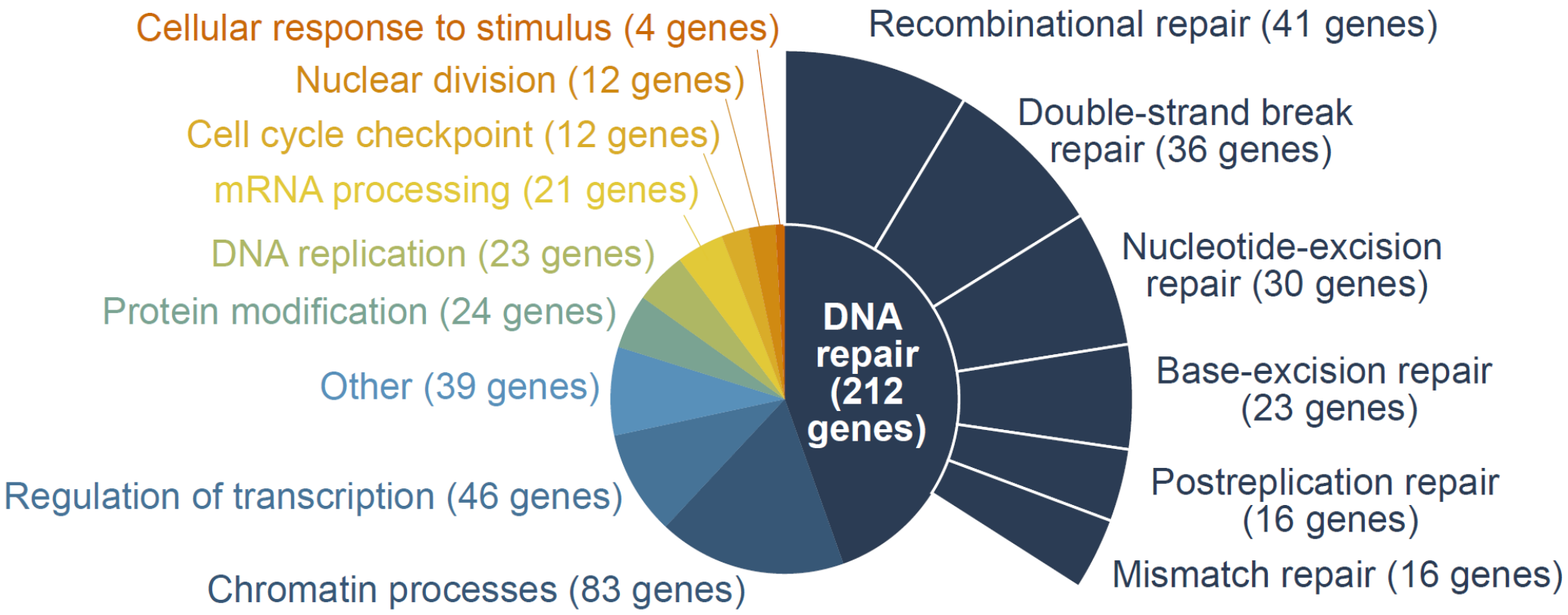
Mapping the genetic landscape of DNA double-strand break repair

October 2021

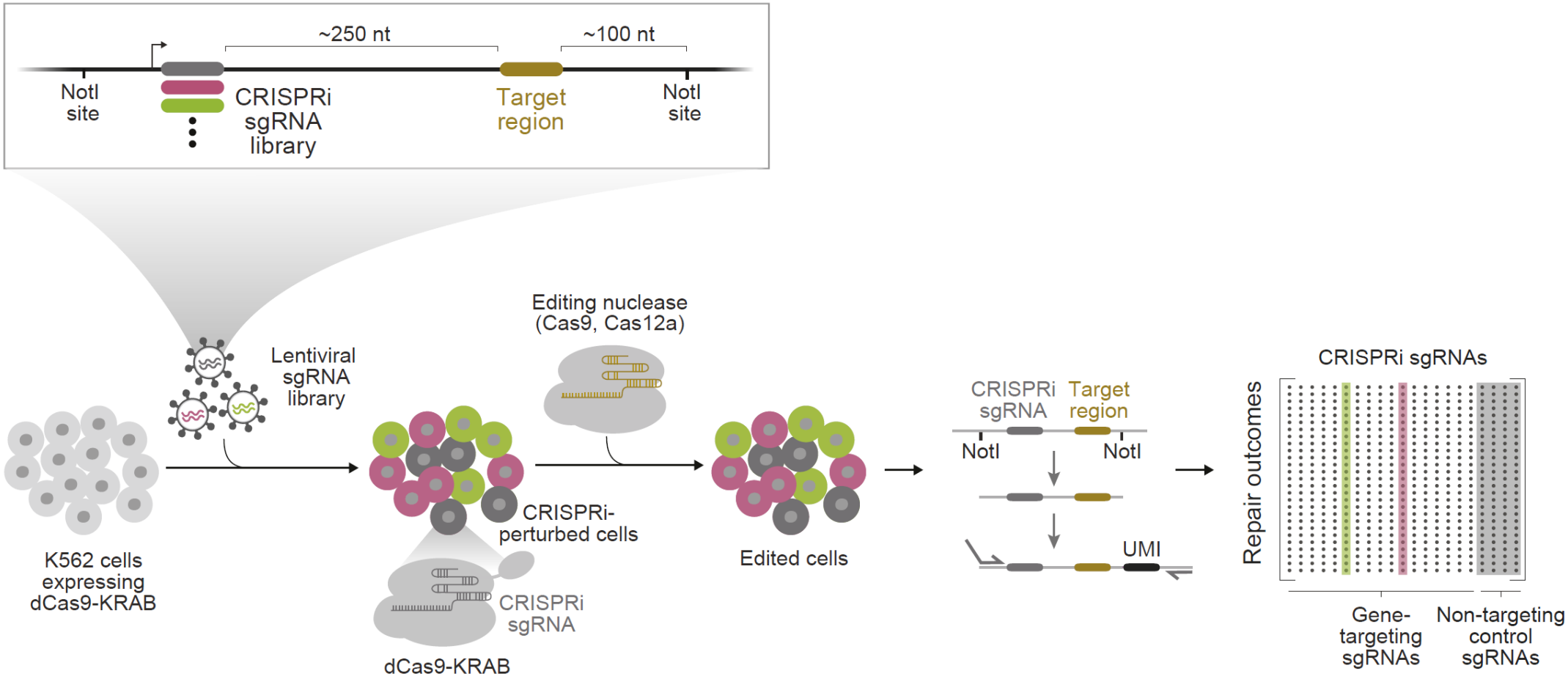
Authors

Jeffrey A. Hussmann, Jia Ling,
Purnima Ravisankar, ...,
Cecilia Cotta-Ramusino,
Jonathan S. Weissman, Britt Adamson

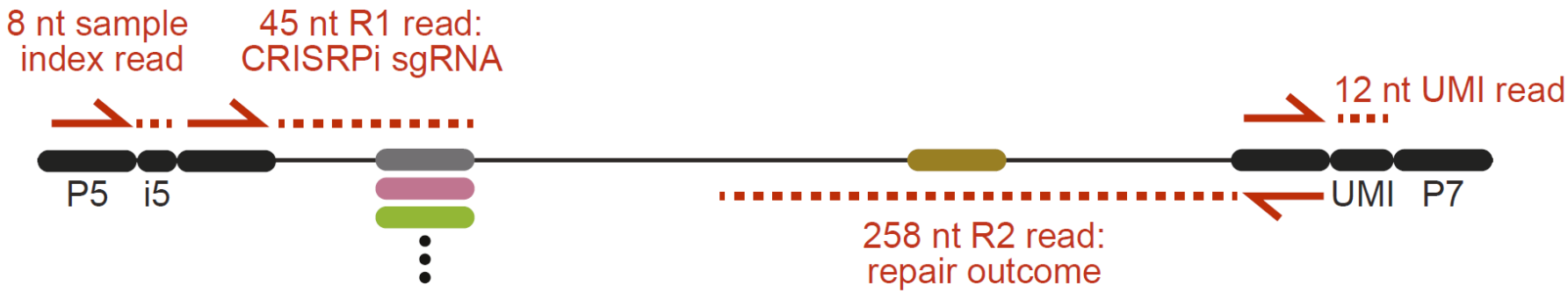
The same CRISPRi library of 476 repair-related genes was used

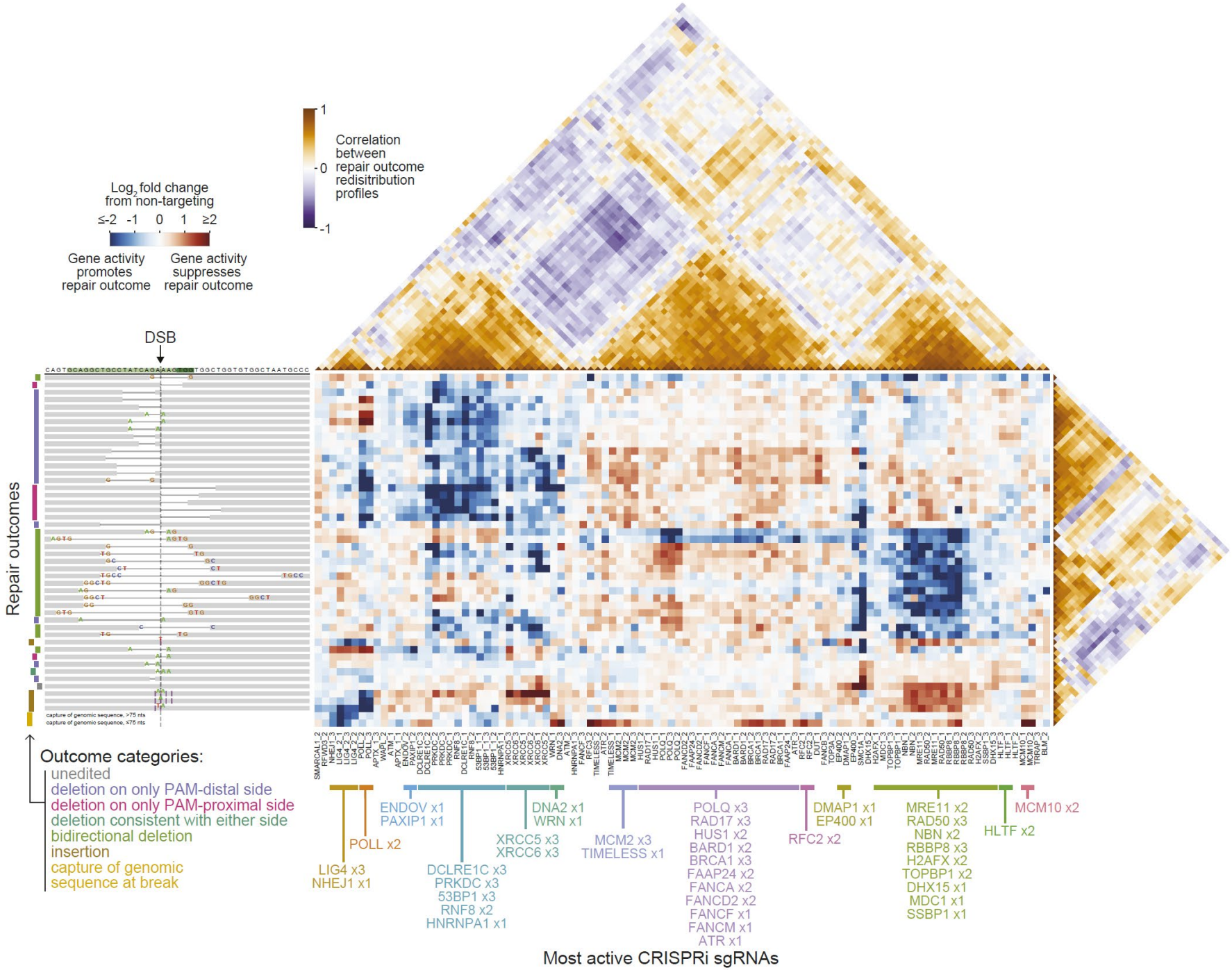


The authors performed Repair-seq to study the repair of double-strand breaks. These were induced in the target region by Cas nucleases (Cas9 or Cas12a).



- In this study, the authors were much more interested in the **diversity of repair outcomes**.
- In every pooled CRISPR screen, PCR is done to amplify the region containing the sgRNAs. In Repair-seq, paired-read sequencing allows the capture of the target region as well.
- The authors were concern that **PCR bias** may lead to the preferential amplification of certain genomic variants over others.
  - To address this, **unique molecular identifiers** (UMIs) were ligated to the DNA fragments after DNA extraction, digestion, size selection and purification.
  - Only then is PCR performed.
  - The number of UMIs, not the number of reads, is used to determine sgRNA counts.

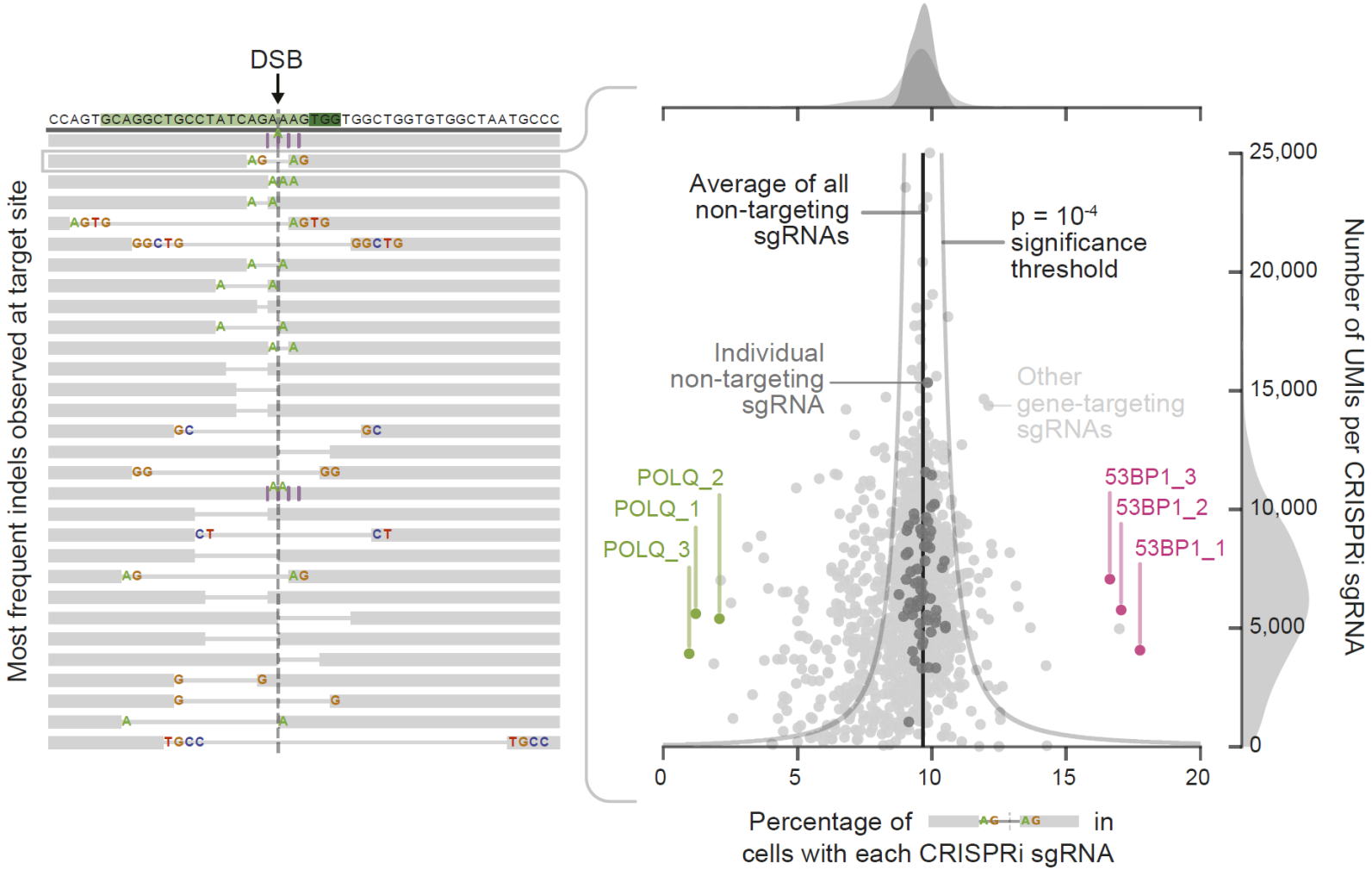




Repair-Seq provides a **global** and unbiased view on the **relationships** of genes involved in DNA repair.

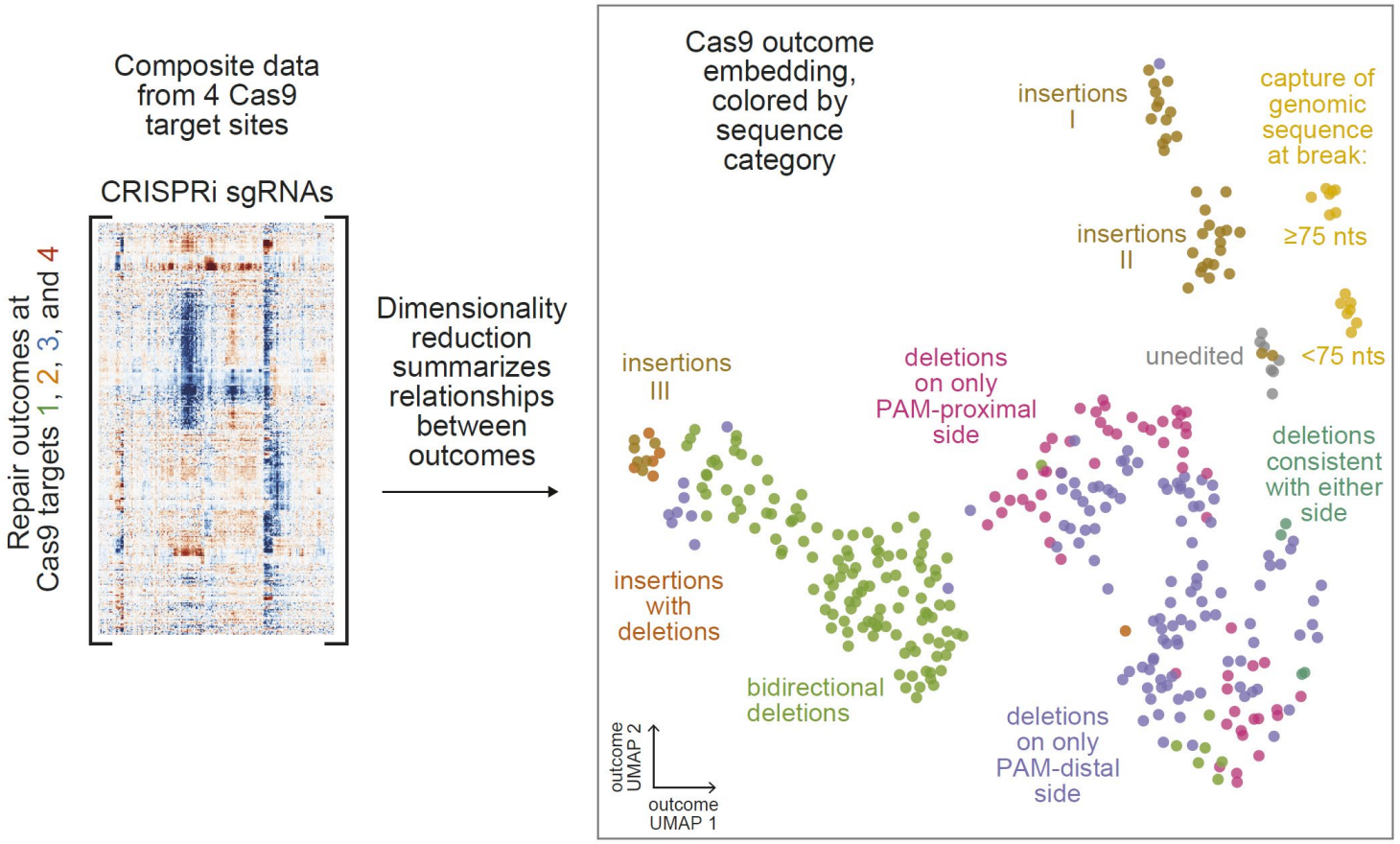
(Because of the bias in the library, we can only discover new relationships, rather than new genes involved in DNA repair.)

The screen identifies previously known functions of repair genes



- CRISPRi of *POLQ* (DNA polymerase theta, involved in theta-mediated end-joining) decreases the incidence of microhomology-flanked deletions.
- Inhibition of *53BP1* increases the rate of these alterations.

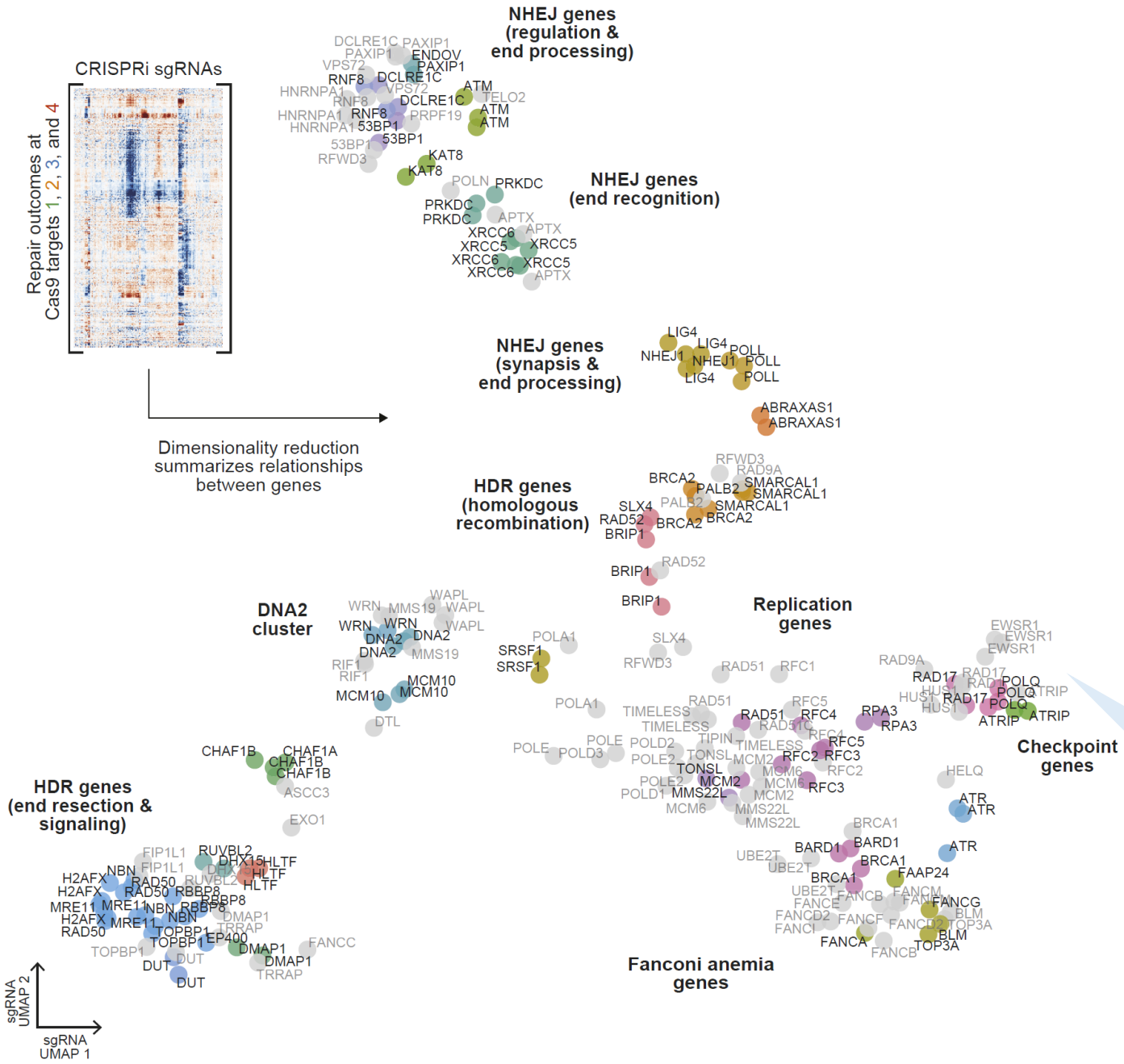
Genomic alterations were clustered based on **which CRISPRi sgRNAs affect the frequency of their occurrence**



Sequence features were **not** used for clustering!

Nevertheless, alterations of a given type often cluster together.

However, superficially similar alterations often falls into distinct sub-clusters, which hints to their genesis via distinct molecular pathways.



Genes were clustered by their impact on the **frequency of each alteration**.

This is an unbiased method for sorting repair genes into pathways.

The clusters reflect known pathways, but also some unexpected relationships.

The strong relationship between the replication checkpoint gene *RAD17* and the *POLQ*, the mediator of alternative end-joining, was consistent across replicates, and was previously unknown.

Conclusions

- The authors provide a purely data-driven approach for defining DNA repair pathways.
- The size of the library (1513 sgRNAs, 476 genes) was chosen based on experimental practicality. They achieved a good signal-to-noise ratio, however, there are likely genes with unexpected roles in DNA repair that remain to be discovered.
- The detection of a sequence requires the presence of two flanking sequences. If one of them is lost, sequencing is impossible. Thus, it is impossible to discover the regulators of large deletions, chromosomal rearrangements or chromothripsis with Repair-seq.

Thank you for your attention!