

The Literature

GENETICS

CRISPR Predictions

THE PAPERS

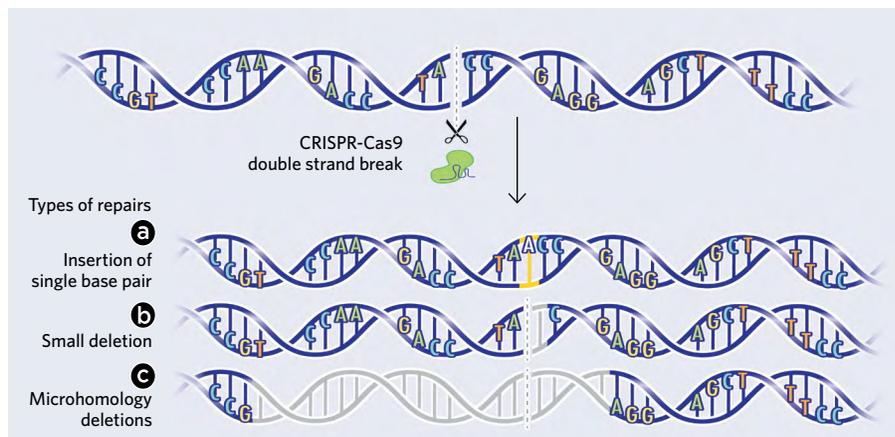
M.W. Shen et al., “Predictable and precise template-free CRISPR editing of pathogenic variants,” *Nature*, 563:646–51, 2018.

F. Allen et al., “Predicting the mutations generated by repair of Cas9-induced double-strand breaks,” *Nat Biotechnol*, 37:64–72, 2019.

During gene editing with CRISPR technology, the Cas9 scissors that cut DNA home in on the right spot to snip with the help of guide RNA. The way the genetic material is stitched back together afterward isn't terribly precise, though; in fact, scientists have long thought that without a template, the process is random. However, “there's been anecdotal evidence that cells don't repair DNA randomly,” geneticist Richard Sherwood of Brigham and Women's Hospital tells *The Scientist*. A 2016 paper also suggested patterns in the repairs. Sherwood wondered if artificial intelligence could predict these outcomes.

In a study published last year in *Nature*, Sherwood and colleagues describe how they trained a machine learning algorithm called inDelphi to predict repairs made to DNA snipped with Cas9, using experimental data from 1,872 target sequences cut and then restitched in mouse and human cell lines. The algorithm showed that 5–11 percent of the guide RNAs used induced a single, predictable repair genotype in the human genome in more than 50 percent of editing products. In other words, the edits aren't random, the team reports.

Separately, Felicity Allen and Leopold Parts of the Wellcome Sanger Institute in the UK and colleagues created an algorithm called FORECasT (favored outcomes of repair events at Cas9 targets) to do the same thing. Based on a library of 41,630 guide RNAs and the sequences of the targeted loci before and after repair—a dataset that totaled more



CRYSTAL BALL: CRISPR guide RNAs target specific spots in the genome for the Cas9 enzyme to cut, forming a double-strand break. A machine learning algorithm predicts which types of repairs will be made at a site targeted by a specific guide RNA. Possibilities include an insertion of a single base pair **a**, a small deletion **b**, or a larger change known as a microhomology deletion **c**.

than 1 billion repairs in various cell types—the model showed that the majority of repairs are either single base insertions, small deletions, or longer deletions called microhomology-mediated deletions, and are based on specific sequences that exist at the Cas9-cut site. The algorithm was then able to use the sequences that determine each repair to predict Cas9 editing outcomes, the researchers reported in *Nature Biotechnology*. The predicted repairs are similar to Sherwood's, but based on much more data, Allen and Parts say.

“It's the right place and the right time for these predictions to occur,” says Rich Stoner, the chief science officer at Synthego, a genome engineering company interested in developing repair-prediction programs, similar to inDelphi, FORECasT, and a third one called SPROUT, for commercialization. However, Stoner notes, a still-unpublished analysis of the three algorithms' results reveals that at times they all made vastly different predictions for the same cuts in the same types of cells,

suggesting that the algorithms' accuracy needs improvement.

Accurate predictions of sequence repair could allow researchers to computationally predict the precise guide RNAs that will reproduce exact human patient mutations, leading to the development of better research models to study genetic disease. Sherwood and his colleagues also showed that their algorithm could predict which guide RNAs would be needed to—without a repair template—correct disease-causing mutations found in human patients, a clinical application of CRISPR that is still years, if not decades, from becoming a reality. The predicted repairs worked on cell lines from patients with a rare genetic disease that causes a blood clotting deficiency and albinism, and another that includes growth failure and nervous system deterioration.

Next, Sherwood says, “we would want to test whether we can fix disease-causing mutations in animal models, with an eventual goal of doing so for human patients.”

—Ashley Yeager