Historically, boreal forests have burnt once every 70–120 years, she says, which gives the black-spruce forest that dominates the ecosystem time to regenerate and rebuild carbon in the soil. More-frequent fires can burn 'legacy' carbon that has accumulated over centuries[2] and can also kill off the black spruce (*Picea mariana*). That provides an opening for trees that do not promote the kind of carbon-rich soils that insulate permafrost.

Emissions win

Fire suppression could help to stave off some of these effects, buying humanity time to address the climate crisis. In a 2022 paper[3], researchers at Woodwell and the Union of Concerned Scientists, a non-profit organization in Cambridge, Massachusetts, found that fire-suppression efforts in Alaska tend to reduce the total area burnt. Their calculations suggest that investing in fire suppression could reduce carbon emissions at a lower cost than that of many technologies for reducing industrial emissions. With an investment of around US$700 million annually in suppression over the next decade, Alaska alone could reduce carbon emissions by up to 3.9 billion tonnes of carbon dioxide by mid-century. That is more than the annual greenhouse-gas emissions of the European Union.

The idea that governments should attempt fire suppression in remote boreal forests has encountered scepticism. Fires play an important part in the ecosystem, and research shows that suppression efforts allow fuel to build up in many forests, contributing to increasingly intense fires[4].

Still, because of the looming climate crisis, many researchers say that the suppression efforts at the Yukon Flats might be worthwhile. "Stopping fires across the boreal forests is an impossible feat, but targeted suppression in areas that are vulnerable seems like a great strategy," says Walker.

A refuge for permafrost

The pilot project at Yukon Flats began last year in 8 areas covering nearly 650,000 hectares of land. Those zones include 40% of the land underlain by a uniquely vulnerable type of permafrost called Yedoma, which contains deep ice wedges that often melt after fires. This causes the land to collapse, exposing ancient carbon to microorganisms whose activity releases greenhouse gases. The target areas contain some 1.1 billion tonnes of carbon, which, if released, would be equivalent to around 7 years of emissions from US coal burning.

Yukon Flats refuge manager Jimmy Fox says that he decided to move forward with the project after consulting with scientists and firefighters. He also involved Indigenous communities that own some 1.2 million hectares of land in the refuge in the discussion. These communities have voiced concerns about wildfire smoke and the negative effects of the fires on water quality and on the land used for hunting, berry picking and other subsistence activities.

Under the pilot programme, firefighters will be deployed on fires that start early in the season and have the potential to burn for several months. Firefighters will engage only if they determine that they can put out the fires within three days, so as to minimize costs and the use of firefighting resources.

Fox notes that Yedoma permafrost zones that burnt decades ago and then collapsed are slowly beginning to recover as vegetation thickens and produces carbon-rich soils that once again insulate the ice below. But with rising temperatures and increasing fire frequency, he says, the danger is that these areas will reach a tipping point at which point recovery would no longer be possible.

Ultimately, the goal is to hold fires to a more historical level. It could be a few decades before scientists can determine whether the effort pays off, Fox warns, "but we have to try".

1. Trehame, R. *et al*. Preprint at https://doi.org/10.21203/rs.3.rs-3909244/v1 (2024).
2. Walker, X. J. *et al. Nature* **572**, 520–523 (2019).
3. Phillips, C. A. *et al. Sci. Adv.* **8**, eabl7161 (2022).
4. Kreider, M. R. *et al. Nature Commun.* **15**, 2412 (2024).

ONLINE TOOL IDENTIFIES PAPERS DISCUSSED ON PUBPEER

Plug-in flags when studies — or their references — have been posted on a site for raising integrity concerns.

By Dalmeet Singh Chawla

A free online tool alerts researchers if a paper cites studies that are mentioned on the website PubPeer, a forum scientists often use to raise integrity concerns surrounding published papers.

Studies are usually flagged on PubPeer when readers have suspicions, for example about image manipulation, plagiarism, data fabrication or artificial intelligence (AI)-generated text. PubPeer already offers its own browser plug-in that alerts users if a study that they are reading has been posted on the site. The new tool, a plug-in released on 13 April by RedacTek, based in Oakland, California, goes further — it searches through reference lists for papers that have been flagged. The software pulls information from many sources, including PubPeer's database; data from the digital-infrastructure organization Crossref, which assigns digital object identifiers to articles; and OpenAlex, a free index of hundreds of millions of scientific documents.

"When you see a large number of problems in a bibliography, that just calls everything into question."

It's important to track mentions of referenced articles on PubPeer, says Jodi Schneider, an information scientist at the University of Illinois Urbana-Champaign, who has tried out the RedacTek plug-in. "Not every single reference that's in the bibliography matters, but some of them do," she adds. "When you see a large number of problems in somebody's bibliography, that just calls everything into question."

The aim of the tool is to flag potential problems with studies to researchers early on, to reduce the circulation of poor-quality science, says RedacTek founder Rick Meyler, who is based in Emeryville, California. Future versions might also use AI to automatically clarify whether the PubPeer comments on a paper are positive or negative, he adds.

Third-generation retractions

As well as flagging PubPeer discussions, the plug-in alerts users if a study, or a paper that it cites, has been retracted. There are existing tools that alert academics about retracted citations; some can do this during the writing process, so that researchers are aware of the publication status of studies when constructing bibliographies. But with the new tool, users can opt in to receive notifications about further 'generations' of retractions — alerts cover not only the study that they are reading, but also the papers it cites, articles cited by those references and even papers cited by the secondary references.

The software also calculates a 'retraction association value' for studies, a metric that measures the extent to which the paper is associated with science that has been withdrawn from the literature. As well as informing

individual researchers, the plug-in could help scholarly publishers to keep tabs on their own journals, Meyler says, because it allows users to filter by publication.

In its 'paper scorecard', the tool also flags any papers in the three generations of referenced studies in which more than 25% of papers in the bibliography are self-citations – references by authors to their previous works.

Meyler says that RedacTek is currently in talks with the scholarly-services firm Cabell's International in Beaumont, Texas, which maintains pay-to-view lists of suspected predatory journals. These publish articles without running proper quality checks for issues such as plagiarism but still collect fees from authors.

The plan is to use these lists to improve the tool so that it can also automatically flag any cited papers that are published in such journals.

'CHATGPT FOR CRISPR' CREATES NEW GENE-EDITING TOOLS

Some of the AI-designed gene editors could be more versatile than those found in nature.

By Ewen Callaway

In the never-ending quest to discover previously unknown CRISPR gene-editing systems, researchers have scoured microbes in everything from hot springs and peat bogs to poo and even yogurt. Now, thanks to advances in generative artificial intelligence (AI), they might be able to design these systems with the push of a button.

This week, researchers published details of how they used a generative AI tool called a protein language model – a neural network trained on millions of protein sequences – to design CRISPR gene-editing proteins. They also showed that some of these systems work as expected in the laboratory[1]. In February, another team announced that it had developed a model trained on microbial genomes, and used it to design fresh CRISPR systems, which are composed of a DNA or RNA-cutting enzyme and RNA molecules that direct the molecular scissors as to where to cut[2].

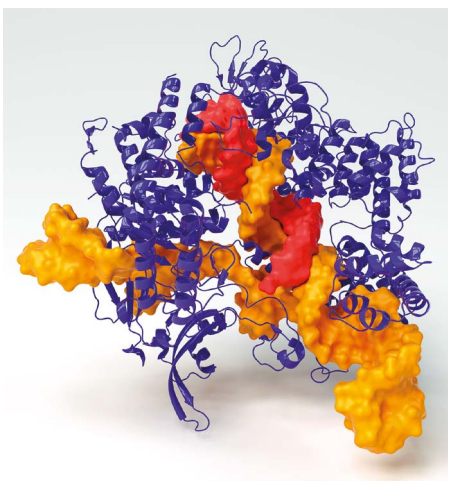
"It's really just scratching the surface. It's showing that it's possible to design these complex systems with machine-learning models," says Ali Madani, a machine-learning scientist and chief executive of the biotechnology firm Profluent, based in Berkeley, California. Madani's team reported what it says is "the first successful editing of the human genome by proteins designed entirely with machine learning" in a 22 April preprint[1] on bioRxiv (which hasn't been peer reviewed).

Alan Wong, a synthetic biologist at the University of Hong Kong, whose team has used machine learning to optimize CRISPR[3], says that naturally occurring gene-editing systems have limitations in terms of the sequences that they can target and the sort of changes that they can make. For some applications, therefore, it can be a challenge to find the right CRISPR. "Expanding the repertoire of editors, using AI, could help," he says.

Trained on genomes

Whereas chatbots such as ChatGPT train on text, the CRISPR-designing AIs were instead trained on vast troves of protein or genome sequences. The goal of this pre-training step was to imbue the models with insight into naturally occurring genetic sequences, such as which amino acids tend to go together. This information can then be applied to tasks such as the creation of totally new sequences.

Madani's team previously used a protein language model it developed, called ProGen, to come up with new antibacterial proteins[4]. To devise fresh CRISPRs, Madani's team retrained an updated version of ProGen with millions of diverse CRISPR systems, which bacteria and other single-celled microbes called archaea use to fend off viruses. Because CRISPR systems comprise not only proteins, but also RNA molecules that specify their target, Madani's team developed another AI model to design these 'guide RNAs'.



A bacterial CRISPR–Cas9 complex.

INDIGO MOLECULAR IMAGES/SPL

The team then used the neural network to design millions of CRISPR protein sequences that belong to dozens of different families of such proteins found in nature. To see whether AI-designed CRISPRs were bona fide gene editors, Madani's team synthesized DNA sequences corresponding to more than 200 protein designs belonging to the CRISPR–Cas9 system that is now widely used in the lab. When the researchers inserted these sequences into human cells, many of the gene editors were able to precisely cut their intended targets in the genome.

The most promising Cas9 protein – a molecule they've named OpenCRISPR-1 – was just as efficient at cutting targeted DNA sequences as a widely used bacterial CRISPR–Cas9 enzyme, and it made many fewer cuts in the wrong place. The researchers also used the OpenCRISPR-1 design to create a base editor – a precision gene-editing tool that changes individual DNA 'letters' – and found that it, too, was as efficient as other base-editing systems, as well as less prone to errors.

Another team used an AI model capable of generating both protein and RNA sequences. This model, called EVO, was trained on 80,000 microbial genomes, and has not been yet tested in the lab. But predicted structures of some of the CRISPR–Cas9 systems it designed resemble those of natural proteins. The work was described in a preprint[2] posted on bioRxiv, and has not been peer reviewed.

Precision medicine

"This is amazing," says Noelia Ferruz Capapey, a computational biologist at the Molecular Biology Institute of Barcelona in Spain. She's impressed by the fact that researchers can use the OpenCRISPR-1 molecule without restriction, unlike with some patented gene-editing tools. The ProGen2 model and atlas of CRISPR sequences used to fine-tune it are also freely available.

The hope is that these tools could be better suited to medical applications than are existing CRISPRs. Profluent is hoping to partner with companies that are developing gene-editing therapies. "It really necessitates precision and a bespoke design. And I think that just can't be done by copying and pasting" from naturally occurring CRISPR, says Madani.

1. Ruffolo, J. A. *et al*. Preprint at bioRxiv https://doi.org/10.1101/2024.04.22.590591 (2024).
2. Nguyen, E. *et al*. Preprint at bioRxiv https://doi.org/10.1101/2024.02.27.582234 (2024).
3. Thean, D. G. L. *et al. Nature Commun.* **13**, 2219 (2022).
4. Madani, A. *et al. Nature Biotechnol.* **41**, 1099–1106 (2023).