

Anthropic says Claude agents found a new CRISPR-like enzyme system in phage DNA

Source copy kept by wheresthe.ai for the story "Anthropic says Claude agents found a new CRISPR-like enzyme system in phage DNA" (<https://wheresthe.ai/d/anthropic-claude-art-enzyme-discovery-2026-09/>).

Source: <https://www.anthropic.com/news/claude-discovers-novel-enzyme-system>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: ee12dba21b0409a64941388539f7f02de2f91f2d8f96c7aabe8981f60bfd86d9

Science Claude discovers a novel enzyme system with CRISPR-like repeats Sep 23, 2026

We're introducing a new life sciences research group and laboratory at Anthropic. Our focus is on fundamental biology research using Claude: exploring datasets of DNA to identify uncharacterized protein families, generating hypotheses at scale, and testing them through experiments in the lab. This post introduces the team behind this work and shares early results in which Claude discovered a novel enzyme system with properties reminiscent of CRISPR, with only high-level direction from our scientists.

Many discoveries that have revolutionized biology and medicine started with a scientist noticing something odd in the staggering diversity of molecular machines found in nature. Restriction enzymes, proteins that cut DNA at specific short sequences, were found in bacterial immune systems, where they destroy the DNA of invading viruses. Researchers realized they could use these enzymes to cut DNA at chosen places and splice genes from one organism into another, which launched the biotechnology industry. Taq polymerase, an enzyme that copies DNA at high temperatures, was identified in a bacterium in a Yellowstone hot spring. It became the basis for PCR, the DNA-copying method used in much of modern diagnostics. CRISPR was first noticed as an unusual repeat sequence in the DNA of certain bacteria, and is now the foundation of gene editing-based medicines.

In the spring of 2026, we formed a research group to see whether general AI models can systematize and accelerate such discoveries. We believe that this acceleration will come from establishing a new way of doing biology research, in which agents collaborate with humans in every step of the process. Developing this new way of working required that we build our own lab and a single team working on everything from training Claude in biology to running experiments in the lab. Today, we're sharing early results from one of our first research programs, in which Claude autonomously discovered a novel enzyme system that is associated with an array of DNA repeats, a pattern reminiscent of CRISPR. Although we don't yet know its function, the system that Claude discovered has a set of characteristics that have only ever been found together in a handful of other systems, all of which are programmable and perform operations like cutting, copying, and pasting DNA. Beyond CRISPR, which has already transformed science and medicine, several other such systems are now in development as promising tools. The system that Claude found is based on a reverse transcriptase (RT), enzymes that copy RNA into DNA. While this underlying RT, found in a jumbo phage, had been identified in previous studies, Claude appears to be the first to notice the system's defining features—an associated array of non-coding DNA sequences and an additional accessory protein of unknown function. After reviewing the pre-print, Feng Zhang, one of the pioneers of CRISPR genome editing and a professor at MIT and the Broad Institute, said: This is an exciting example of how AI agents can contribute to biological discovery. The identification of RNA-repeat arrays associated with reverse transcriptases is genuinely intriguing and merits further investigation. I hope this work encourages more scientists to explore how AI can support their research. We gave Claude a prompt to search through a massive database of DNA sequences for interesting new examples of RTs. Our involvement was limited to the initial prompt and the lab work, while Claude agents combed through the database, investigated the distinct RT families, and used their own judgment to identify interesting candidates. After 21 hours spent searching this data by roughly 950 agents using 210 million tokens, one of the agents spotted something remarkable: a repeating pattern of DNA sequences that occurs next to the gene for an odd-looking RT. After further analysis and testing in our lab, we recognized that this pattern marked a previously uncharacterized enzyme system found in bacteriophages (the viruses that infect bacteria) that we call array-associated reverse transcriptases (ART). Our work to understand the primary function of ARTs is ongoing. However, we think it is important to share such findings early, both to demonstrate Claude's capabilities and to give the broader community insight into what we're working on. We have released a pre-print ([here](#)) that discusses this in more detail. About our lab We are a team of scientists who have spent our careers exploring unusual proteins, and specialize in using computational approaches to systematically read DNA, interpret its evolution, and pick out biological systems for further

characterization. Our research prior to joining Anthropic has helped to better understand the evolution and regulation of CRISPR systems, discover new enzymes for next-generation cell and gene therapies, and build tools for accelerating the identification of anomalies in DNA, such as human pathogenic variants. We are part of Anthropic's life sciences organization, alongside teams whose work includes drug discovery, and training Claude in biology and chemistry. Our lab, located in the Bay Area, looks like a typical molecular biology lab. We do research that involves only the lower levels of the biosafety risk level (BSL-1 and BSL-2) and we do not handle pathogens that can infect humans. All of the lab work is performed by human scientists. Although we've experimented with using AI to accelerate lab work with initiatives like the Model Hardware Standard, this approach is less conducive to the sort of ad hoc workflows that are involved in our molecular biology research.

How we work Many of our workflows involve having Claude search through the vast collection of DNA sequences associated with proteins without a known function. One typical pattern begins with a survey of a given protein family. Claude reads the relevant literature and reproduces the established results from public data to check its methods. It then searches for family members or genomic neighbors that fit no described system, and writes a short, human-readable report for each candidate that proposes a function and describes the evidence supporting its claims. In follow-up analyses, Claude critically evaluates the evidence—typically most candidates are eliminated at this stage. A survey may end with a single candidate worth testing, or with none. When a candidate survives our review, we test it in the laboratory, expressing the protein in standard laboratory strains and characterizing it biochemically and structurally, with Claude helping to interpret the data. We do our work in Claude Science and Claude Code, the same tools available to any scientist, and sometimes with a harness of our own that coordinates many Claude sessions running in parallel. Because Claude produces hypotheses so prolifically, the hypotheses themselves have become an object of study for us. With hundreds to thousands of candidate reports from a single campaign, we have been asking what distinguishes the proposals we judge worth testing from those we set aside. What we learn goes back into the instructions we give Claude and teaches it to mimic our own scientific taste.

Claude finds ART In the past few years, researchers have discovered many more reverse transcriptases (RTs), most of them in bacteria, where they act as part of the immune system. Nearly all RT families were found by genomic analysis, or genome mining, which requires researchers to search sequence databases for genes that no one has characterized, notice the unusual ones, and work out what they do. Claude agents gathered over 200,000 RTs, picked out 3,500 new candidate systems, and narrowed those to the 20 most compelling candidates that they analyzed to produce human-readable reports. For an expert scientist, this type of analysis can take weeks to months of work. During the course of its research, Claude noticed an unusual RT family and decided to examine it in greater detail. While combing through the raw DNA sequence near the RT, the agent exclaimed: "[The DNA next to the RT] is spectacular: I can see by eye a tandem repeat array ... that's a CRISPR-like ... repeat array?"

The raw DNA Claude was reading when it detected a repeat pattern that no one had noticed

It then proceeded much as a human scientist would when faced with a potential discovery. It counted the repeats and measured their spacing, compared the layout with the known RT systems, and searched the literature for any previous report of the pattern. After a thorough analysis it was convinced that it had found a new biological system, and filed a report for human review. The system it found, ART, is found mainly in bacteriophages and consists of three parts: the RT, a partner gene beside it, and a long array of evenly spaced DNA repeat sequences. The repeat layout resembles a CRISPR array, which holds a bank of different RNA sequences that make CRISPR-Cas systems programmable biotechnological tools. Our first experiments show that the ART array is also expressed as a set of distinct short RNAs, suggesting that something analogous may be at play for this system. Further experiments are underway to determine how ART works, and we are sharing these early findings to show the community that Claude can autonomously detect anomalies and drive analyses to initiate biological discoveries. You can find more detail in our technical report ([here](#)).

Work with us We hope this work demonstrates the value of AI-driven hypothesis generation to the wider scientific community, and we would like to work with other scientists to extend this approach to a broad range of problems, in genomics and in other fields. If you have a proposal for a research question, we would like to hear from you.

Related content

An update on independent evaluators To evaluate the safety of our systems, Anthropic is giving employee-like access to a team from METR, a nonprofit that evaluates frontier AI models, and we're in early discussions with several other organizations, including Hugging Face. [Read more](#)

Preparing for the 2026 US midterms Earlier this year, we shared an update on our election safeguards. [Here](#), we outline how we've continued to strengthen these safeguards ahead of this year's US midterms. [Read more](#)

Anthropic opens Singapore office and signs MOU with

the Government of Singapore Anthropic has opened an office in Singapore and signed an MOU with the country's government to advance safe, responsible AI adoption. [Read more](#)

Subscribe to Anthropic Science Features on AI-assisted discoveries, practical workflows, and field notes across the sciences.