

# Junk DNA in birds may hold key to safe, efficient gene therapy

February 20 2024

---



Retrotransposons found in the genomes of the white-throated sparrow and the zebra finch are shown to safely shepherd transgenes into the human genome, providing a gene therapy approach complementary to CRISPR-Cas9 gene editing. Credit: Briana Van Treeck, UC Berkeley

The recent approval of a CRISPR-Cas9 therapy for sickle cell disease demonstrates that gene editing tools can do a superb job of knocking out genes to cure hereditary disease. But it's still not possible to insert whole

genes into the human genome to substitute for defective or deleterious genes.

A new technique that employs a retrotransposon from birds to insert genes into the genome holds more promise for [gene therapy](#), since it inserts genes into a "safe harbor" in the human genome where the insertion won't disrupt essential genes or lead to cancer.

Retrotransposons, or retroelements, are pieces of DNA that, when transcribed to RNA, code for enzymes that copy RNA back into DNA in the genome—a self-serving cycle that clutters the genome with retrotransposon DNA. About 40% of the human genome is made up of this "selfish" new DNA, though most of the genes are disabled, so-called junk DNA.

The new technique, called Precise RNA-mediated INsertion of Transgenes, or PRINT, leverages the ability of some retrotransposons to efficiently insert entire genes into the genome without affecting other genome functions. PRINT would complement the recognized ability of CRISPR-Cas technology to disable genes, make point mutations and insert short segments of DNA.

A description of PRINT, which was developed in the laboratory of Kathleen Collins, a professor of molecular and cell biology at the University of California, Berkeley, is published on Feb. 20 in the journal *Nature Biotechnology*.

PRINT involves the insertion of new DNA into a cell using delivery methods similar to those used to ferry CRISPR-Cas9 into cells for genome editing. For PRINT, one piece of delivered RNA encodes a common retroelement protein called R2 protein, which has multiple active parts, including a nickase—an enzyme that binds and nicks double-stranded DNA—and reverse transcriptase, the enzyme that generates the

DNA copy of RNA. The other RNA is the template for the transgene DNA to be inserted, plus gene expression control elements—an entire autonomous transgene cassette that R2 protein inserts into the genome, Collins said.

A key advantage of using R2 protein is that it inserts the transgene into an area of the genome that contains hundreds of identical copies of the same gene—each coding for ribosomal RNA, the RNA machine that translates messenger RNA (mRNA) into protein. With so many redundant copies, when the insertion disrupts one or a few ribosomal RNA genes, the loss of the genes won't be missed.

Putting the transgene into a safe harbor avoids a major problem encountered when inserting transgenes via a human virus vector, which is the common method today: The gene is often inserted randomly into the genome, disabling working genes or messing with the regulation or function of genes, potentially leading to cancer.

"A CRISPR-Cas9-based approach can fix a mutant nucleotide or insert a little patch of DNA—sequence fixing. Or you can just knock out a gene function by site-specific mutagenesis," said Collins, who holds the Walter and Ruth Schubert Family Chair.

"We're not knocking out a gene function. We're not fixing an endogenous gene mutation. We're taking a complementary approach, which is to put into the genome an autonomously expressed gene that makes an active protein—to add back a functional gene as a deficit bypass. It's transgene supplementation instead of mutation reversal. To fix loss-of-function diseases that arise from a panoply of individual mutations of the same gene, this is great."

**'The real winners were from birds'**

Many hereditary diseases, such as cystic fibrosis and hemophilia, are caused by a number of different mutations in the same gene, all of which disable the gene's function. Any CRISPR-Cas9-based gene editing therapy would have to be tailored to a person's specific mutation. Gene supplementation using PRINT could instead deliver the correct gene to every person with the disease, allowing each patient's body to make the normal protein, no matter what the original mutation.

Many academic labs and startups are investigating the use of transposons and retrotransposons to insert genes for gene therapy. One popular retrotransposon under study by biotech companies is LINE-1 (Long INterspersed Element-1), which in humans has duplicated itself and some hitchhiker genes to cover about 30% of the genome, though fewer than 100 of our genome's LINE-1 retrotransposon copies are functional today, a miniscule fraction of the genome.

Collins, along with UC Berkeley postdoctoral colleague Akanksha Thawani and Eva Nogales, UC Berkeley Distinguished Professor in the Department of Molecular and Cell Biology and a Howard Hughes Medical Institute investigator, [published](#) a cryoelectron microscopy structure of the enzyme protein encoded by the LINE-1 retroelement on Dec. 14 in the journal *Nature*.

That study made it clear, Collins said, that the LINE-1 retrotransposon protein would be hard to engineer to safely and efficiently insert a transgene into the human genome. But previous research demonstrating that genes inserted into the repetitive, ribosomal RNA encoding region of the genome (the rDNA) get expressed normally suggested to Collins that a different retroelement, called R2, might work better for safe transgene insertion.

Because R2 is not found in humans, Collins and senior researcher Xiaozhu Zhang and postdoctoral fellow Briana Van Treeck, both from

UC Berkeley, screened R2 from more than a score of animal genomes, from insects to the horseshoe crab and other multicellular eukaryotes, to find a version that was highly targeted to rDNA regions in the human genome and efficient at inserting long lengths of DNA into the region.

"After chasing dozens of them, the real winners were from birds," Collins said, including the zebra finch and the white-throated sparrow.

While mammals do not have R2 in their genomes, they do have the binding sites needed for R2 to effectively insert as a retroelement—likely a sign, she said, that the predecessors to mammals had an R2-like retroelement that somehow got kicked out of the mammalian genome.

In experiments, Zhang and Van Treeck synthesized mRNA-encoding R2 protein and a template RNA that would generate a transgene with a fluorescent protein expressed by an RNA polymerase promoter. These were cotransfected into cultured human cells. About half the cells lit up green or red due to fluorescent protein expression under [laser light](#), demonstrating that the R2 system had successfully inserted a working fluorescent protein into the genome.

Further studies showed that the transgene did indeed insert into the rDNA regions of the genome and that about 10 copies of the RNA template could insert without disrupting the protein-manufacturing activity of the rDNA genes.

## **A giant ribosome biogenesis center**

Inserting transgenes into rDNA regions of the genome is advantageous for reasons other than it gives them a safe harbor. The rDNA regions are found on the stubby arms of five separate chromosomes. All of these stubby arms huddle together to form a structure called the nucleolus, in

which DNA is transcribed into ribosomal RNA, which then folds into the ribosomal machinery that makes proteins.

Within the nucleolus, rDNA transcription is highly regulated, and the genes undergo quick repairs, since any rDNA breaks, if left to propagate, could shut down protein production. As a result, any transgene inserted into the rDNA region of the genome would be treated with kid gloves inside the nucleolus.

"The nucleolus is a giant ribosome biogenesis center," Collins said. "But it's also a really privileged DNA repair environment with low oncogenic risk from gene insertion. It's brilliant that these successful retroelements—I'm anthropomorphizing them—have gone into the ribosomal DNA. It's multicopy, it's conserved, and it's a safe harbor in the sense that you can disrupt one of these copies and the cell doesn't care."

This makes the region an ideal place to insert a gene for human gene therapy.

Collins admitted that a lot is still unknown about how R2 works and that questions remain about the biology of rDNA transcription: How many rDNA genes can be disrupted before the cell cares? Because some cells turn off many of the 400+ rDNA genes in the [human genome](#), are these cells more susceptible to side effects of PRINT?

She and her team are investigating these questions, but also tweaking the various proteins and RNAs involved in retroelement insertion to make PRINT work better in cultured cells and primary cells from human tissue.

The bottom line, though, is that "it works," she said. "It's just that we have to understand a little bit more about the biology of our rDNA in

order to really take advantage of it."

**More information:** Harnessing eukaryotic retroelement proteins for transgene insertion into human safe-harbor loci, *Nature Biotechnology* (2024). DOI: [10.1038/s41587-024-02137-y](https://doi.org/10.1038/s41587-024-02137-y)

Akanksha Thawani et al, Template and target-site recognition by human LINE-1 in retrotransposition, *Nature* (2023). DOI: [10.1038/s41586-023-06933-5](https://doi.org/10.1038/s41586-023-06933-5).

Provided by University of California - Berkeley

Citation: Junk DNA in birds may hold key to safe, efficient gene therapy (2024, February 20) retrieved 27 April 2024 from <https://phys.org/news/2024-02-junk-dna-birds-key-safe.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.