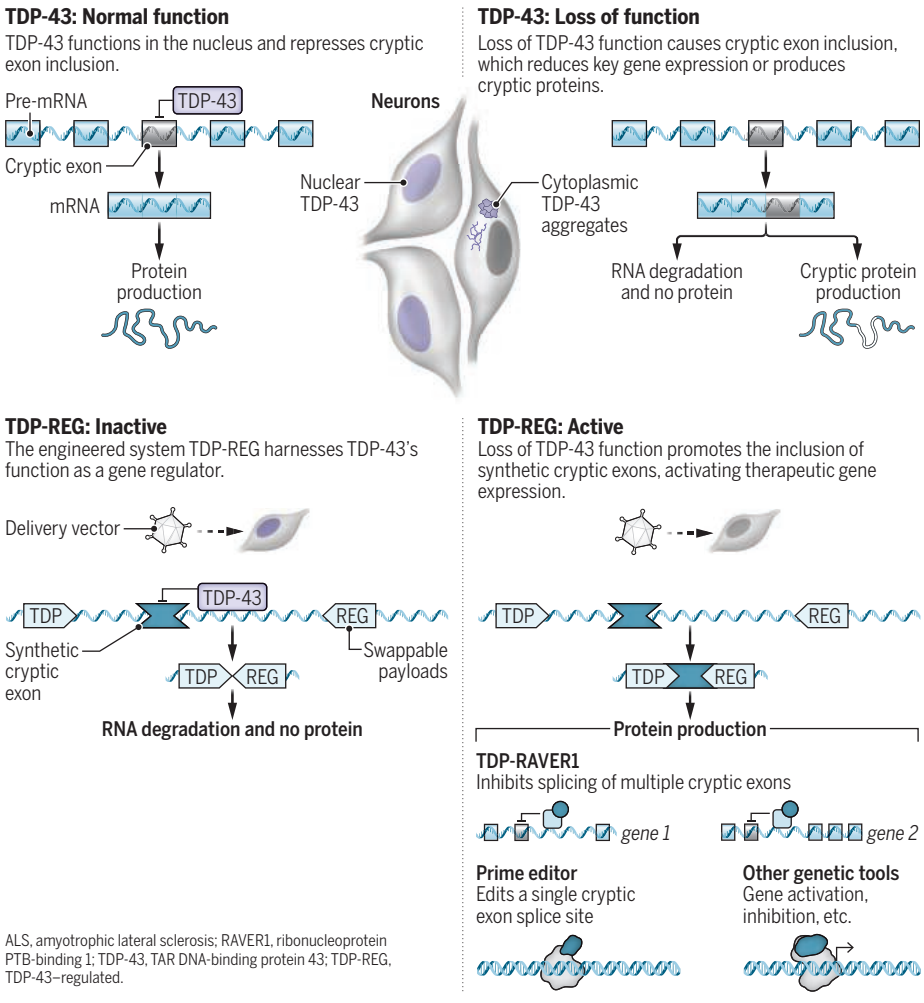
A precision medicine strategy for ALS

Insights from cryptic exon processing may pave the way for targeted treatments for neurodegenerative disorders.



ALS, amyotrophic lateral sclerosis; RAVER1, ribonucleoprotein PTB-binding 1; TDP-43, TAR DNA-binding protein 43; TDP-REG, TDP-43–regulated.

GRAPHIC: A. MASTIN/*SCIENCE*

simultaneously. They did so by using a fusion protein between TDP-43 and the splicing repressor domain of the ribonucleoprotein PTB-binding 1 (RAVER1) protein. This work builds on previous studies that used similar approaches (*6*, *14*) but extends this strategy to only be active in those neurons that need it (i.e., ones with loss of TDP-43 function), avoiding potential deleterious effects of too much TDP-43 in healthy neurons.

Wilkins *et al.* present an elegant strategy to deliver therapies to correct ALS pathologies in those neurons that need them and not the ones that do not. ALS is a devastatingly aggressive and fast progressing disease, so a concern is that once diagnosis is made, it might be too late for therapies to be effective. However, because of the safety features built into TDP-REG, Wilkins *et al.* envision patients at risk for disease (perhaps because of a genetic susceptibility factor) receiving one of these therapies at a presymptomatic stage, where it lays dormant until activated by incipient TDP-43 dysfunction. TDP-REG's modular design strategy is also likely to support future modifications to target new disease-relevant pathologies, cell types, and cell states as they are discovered. This work brings us one step closer to precision medicine therapies for neurodegenerative diseases such as ALS. ■

REFERENCES AND NOTES

1. C. H. June, R. S. O'Connor, O. U. Kawalekar, S. Ghassemi, M. C. Milone, *Science* **359**, 1361 (2018).
2. O. G. Wilkins *et al.*, *Science* **386**, 61 (2024).
3. J. P. Taylor, R. H. Brown Jr., D. W. Cleveland, *Nature* **539**, 197 (2016).
4. M. Neumann *et al.*, *Science* **314**, 130 (2006).
5. S. C. Ling, M. Polymenidou, D. W. Cleveland, *Neuron* **79**, 416 (2013).
6. J. P. Ling, O. Pletnikova, J. C. Troncoso, P. C. Wong, *Science* **349**, 650 (2015).
7. Z. Melamed *et al.*, *Nat. Neurosci.* **22**, 180 (2019).
8. J. R. Klim *et al.*, *Nat. Neurosci.* **22**, 167 (2019).
9. X. R. Ma *et al.*, *Nature* **603**, 124 (2022).
10. A. L. Brown *et al.*, *Nature* **603**, 131 (2022).
11. K. E. Irwin *et al.*, *Nat. Med.* **30**, 382 (2024).
12. S. Seddighi *et al.*, *Sci. Transl. Med.* **16**, eadg7162 (2024).
13. M. W. Baughn *et al.*, *Science* **379**, 1140 (2023).
14. J. P. Ling *et al.*, *Nat. Commun.* **13**, 5773 (2022).

10.1126/science.ads5951

MOLECULAR BIOLOGY

Tricking phages with a reverse move

An antiviral gene is absent in DNA but expressed by rolling circle reverse transcription

By **Ilya Osterman** *and* **Rotem Sorek**

The textbook definition of a protein-coding gene requires it to be linearly encoded along the DNA axis so that the beginning of the gene is always upstream of the gene's end. On pages 40 and 41 of this issue, Tang *et al.* (*1*) and Wilkinson *et al.* (*2*), respectively, challenge this paradigm and reveal peculiar antiphage proteins in bacteria that are not directly encoded in the bacterial DNA. To produce these proteins, DNA is first transcribed into a noncoding RNA (ncRNA), which is converted back to DNA by a rolling circle reverse transcription reaction. The resulting DNA is subsequently transcribed into an mRNA that encodes the protein. The protein assists bacteria to overcome infection by viruses (phage).

Reverse transcriptases (RTs), enzymes that synthesize DNA using an RNA template, are widely distributed across all domains of life. These enzymes have roles in multiple processes, including in the life cycle of retroviruses and mobile genetic elements as well as in telomere biology. In bacteria, RTs are especially important for antiphage defense and are used by diverse genetic systems whose role is to protect bacteria against phage infections. For example, some CRISPR-Cas systems use reverse transcription to acquire new immune cassettes of nucleic acids ("spacers") against RNA phages (*3*). RTs are also used in antiphage genetic systems called retrons, which consist of three genes encoding an RT, an ncRNA, and an "effector" toxin. Through the process of reverse transcription, retrons produce a chimeric nucleic acid chain in which DNA and RNA are covalently linked. The role of this chimeric DNA-RNA molecule is yet unclear, but it was shown to

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. Email: ilya.osterman@weizmann.ac.il; rotem.sorek@weizmann.ac.il

exert antiphage activity by controlling effector protein toxicity (*4*). More than a dozen other bacterial defense systems encode RTs, but the role of reverse transcription in most of these systems is unknown (*5*).

Tang *et al.* and Wilkinson *et al.* independently decided to focus on an RT-containing defense system in the bacterium *Klebsiella pneumoniae* called defense-associated RT 2 (DRT2). Both studies found that this system provides robust defense against phages and that during infection, bacteria that encode DRT2 cease to grow and do not allow phage replication. The DRT2 system encodes merely an RT and an ncRNA, seemingly without any other protein domains or factors involved, and both studies examined how such a minimal system can both recognize infection and cause growth arrest. What they found was that DRT2 actually encodes another protein, but its production requires a complex and unexpected chain of molecular events.

Tang *et al.* and Wilkinson *et al.* observed that the RT uses a specific piece of about 120 bases in the associated ncRNA as a template to generate complementary DNA (cDNA). The RT does not stop at one pass of reverse transcription but jumps back to the starting point of the 120-base sequence and continues in a so-called "rolling circle" reaction to generate a long cDNA containing multiple repeats of the reverse-transcribed template (see the figure). Most commonly, the generated cDNA contained five concatenated repeats of the template, but some cDNAs contained more than a hundred such repeats.

What could be the role of the cDNA? Both studies noted that the cDNA contains sequence motifs that are almost identical to the consensus sequence of bacterial sigma-70 promoters. These promoters comprise two sequence motifs, the −35 and the −10 motifs, which are thus called because they are found 35 and ~10 bases upstream of the transcription start site, respectively. The template repeat in the DRT2 ncRNA sequence contains both motifs but in a permuted manner such that the single repeat does not form a valid promoter; but when concatenated on the cDNA, the −35 motif at the end of one repeat is placed in the correct orientation and distance from the −10 motif at the beginning of the next repeat, forming a strong promoter that activates a high rate of transcription.

Both studies further determined that this promoter drives the transcription of an mRNA, which, in turn, encodes a protein. The mRNA transcribed from the concatenated cDNA does not contain any stop codon, and because the protein could potentially be very long (depending on the number of repeats in the cDNA), Tang *et al.* and Wilkinson *et al.* both named it Neo [nearly endless ORF (open reading frame)].

How does this system block bacterial infection by phages? The research teams found that DRT2 produces a single-stranded concatenated cDNA but that the synthesis of a second strand is only initiated when the cell is infected by a phage. Once a double-stranded cDNA is formed, the promoter drives transcription, followed by translation of the mRNA into Neo. This toxic protein rapidly arrests bacterial growth, thus preventing the phage from replicating and exploiting the cell's resources.

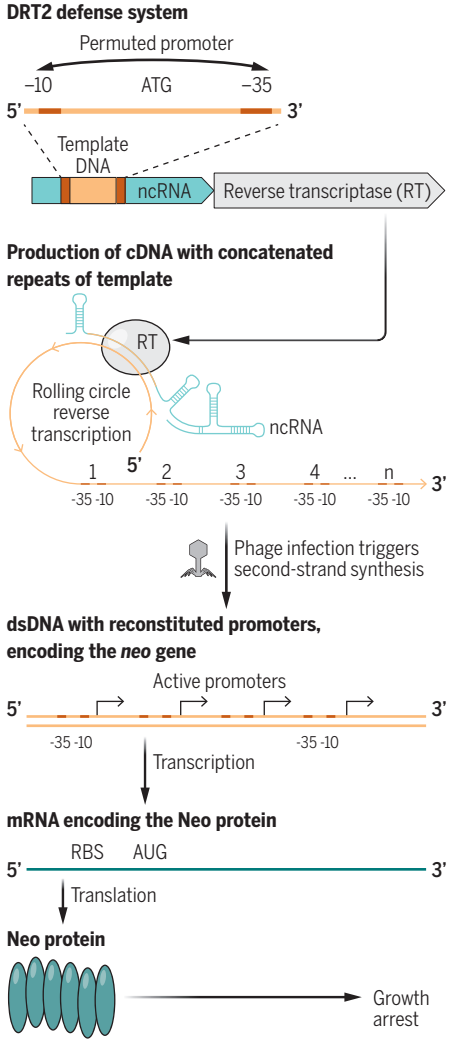
The information flow necessary to encode Neo is remarkably unusual. What is the benefit of such an unconventional protein-encoding mechanism? Tang *et al.* and Wilkinson *et al.* offer two possible explanations. The expression of toxic genes is difficult to control, and the transcriptional leakage of such genes incurs a fitness cost. This is thought to be the reason why defense systems that encode toxins tend to be frequently lost in short evolutionary timescales (*6*). The problem of autoimmunity by transcriptional leakage is solved in the DRT2 system, because the single repeat does not contain a promoter and is itself not toxic. Another explanation is that lytic phages tend to degrade the genome of their host early upon infection, in part to shut down the cell's ability to respond to infection through transcription and translation of immune proteins. The DRT2 system ingeniously overcomes this challenge by relying on an ncRNA that has already been transcribed prior to infection, allowing for the conversion of this ncRNA back into DNA once the DNA degradation activity of the phage has ceased.

It is unclear how DRT2 senses the presence of the phage, and what molecular mechanism activates the synthesis of the second cDNA strand. Although Wilkinson *et al.* showed that the second-strand synthesis initiates from a short DNA primer covalently linked to the end of the ncRNA, it is unknown how this primer is generated. Perhaps the most intriguing question is the mechanism of growth arrest by Neo, as it does not resemble any known protein, and the mechanism by which it ceases bacterial growth remains a puzzle.

RTs from bacterial defense systems, and specifically RTs from retrons, have been adopted for genome editing applications because of their ability to generate cDNA of choice (*7*). The ability of DRT2 to generate double-stranded DNA by rolling circle reverse transcription could contribute to biotechnological applications, exploiting its ability to amplify and concatenate a template sequence. In this context, it is noteworthy that there are still dozens of RTs encoded by bacterial defense systems whose mechanisms are unclear. ■

Hidden gene

A bacterial defense-associated reverse transcriptase 2 (DRT2) system to fight phage infection uses rolling circle reverse transcription and a noncanonical path to express a toxic protein.



ATG/AUG, start codons; cDNA, complementary DNA; dsDNA, double-stranded DNA; ncRNA, noncoding RNA; Neo, nearly-endless ORF (open reading frame); RBS, ribosome binding site.

REFERENCES AND NOTES

1. S. Tang *et al.*, *Science* **386**, eadq0876 (2024).
2. M. E. Wilkinson *et al.*, *Science* **386**, eadq3977 (2024).
3. S. Silas *et al.*, *Science* **351**, aad4234 (2016).
4. A. Millman *et al.*, *Cell* **183**, 1551 (2020).
5. M. R. Mestre *et al.*, *Nucleic Acids Res.* **50**, 6084 (2022).
6. A. Bernheim, R. Sorek, *Nat. Rev. Microbiol.* **18**, 113 (2020).
7. S. C. Lopez, K. D. Crawford, S. K. Lear, S. Bhattarai-Kline, S. L. Shipman, *Nat. Chem. Biol.* **18**, 199 (2022).

Published online 29 August 2024
10.1126/science.ads3638