MOLECULAR BIOLOGY

A treasure trove of molecular scissors

Ancestral RNA-guided nucleases greatly expand the toolbox for gene editing

By **François Rousset** *and* **Rotem Sorek**

Over the past decade, the CRISPR revolution has shaken the fields of molecular biology and biomedicine (*1*). Discovered as defense mechanisms that protect bacteria from viral infection (*2*), the CRISPR-Cas systems cut DNA at a specific sequence, in a highly precise and programmable manner (*3*, *4*). This property has enabled researchers to harness CRISPR-Cas for genome editing applications, a promising advance for treating genetic disorders that were previously considered incurable (*5*). Because CRISPR-Cas genome editing tools are not without flaws, tremendous efforts have been devoted to finding other editing enzymes for biotechnological and clinical needs (*6*–*9*). On page 57 of this issue, —Altae-Tran *et al*. (*10*) report a large untapped resource for programmable DNA-cutting enzymes that expand our genome editing toolbox.

The most widely used CRISPR-Cas enzyme for genome editing applications is Cas9. It is shepherded by a specific "guide RNA" to cut DNA that is complementary to the RNA sequence (*3*, *4*). As such, Cas9 can be repurposed as a programmable molecular scissor simply by providing it with a custom guide RNA to cleave a DNA sequence of choice (*4*). But Cas9 comes with some imperfections as a molecular tool. Its large size makes delivery into target cells challenging. In addition, Cas9 requires a short sequence called a protospacer adjacent motif (PAM) to be present at a specific position in the target DNA molecule (*11*), rendering certain genes inaccessible for editing. As a result, many studies have attempted to identify alternatives for Cas9 that are smaller and have better properties (*6*–*9*).

"What is the evolutionary origin of Cas9?"

Altae-Tran *et al*. report an enormous untapped resource of RNA-guided nucleases encoded in bacterial genomes. This discovery was primed by a "basic research" question that was initially unrelated to any biotechnological application: What is the evolutionary origin of Cas9? In bacterial and archaeal cells, Cas9 functions as an antiviral enzyme programmed by its guide RNAs to destroy the DNA of invading viruses (*3*, *4*). But a recent computational analysis suggested that Cas9 evolved from a family of transposon proteins called IscB (*12*). Transposons are selfish genetic elements that can copy themselves and insert their copies into other places in the genome, and IscB-containing transposons are frequent in bacterial genomes. Altae-Tran *et al*. observed that IscB proteins can occasionally be found in close proximity to CRISPR arrays, genomic loci that encode the short guide RNAs that operate in CRISPR-Cas systems. This association prompted the authors to test whether, like Cas9, IscB could exploit these short guide RNAs to mediate cleavage of a complementary DNA sequence—a hypothesis they experimentally verified. These findings show that some of the Cas9 ancestors were already RNA-guided nucleases.

Still, the vast majority of IscB homologs are not located near CRISPR arrays. If all IscB proteins were to be RNA-guided nucleases, what would guide IscB in the absence of the CRISPR array? By examining a large number of IscB loci, Altae-Tran *et al*. discovered that they encode "hidden" guide RNAs next to the *iscB* gene, which they named omega RNAs (ωRNAs). The authors showed that ωRNAs guide numerous IscB proteins to cleave complementary DNA and that cleavage can be reprogrammed by replacing a certain portion of the ωRNAs with a sequence of choice. They also demonstrated similar activity for the ancestors of IscB, a protein family called IsrB, thereby solving the evolutionary history of the Cas9 protein family (see the figure).

Whereas the antiviral role of CRISPR-Cas9 is well established, the biological function of IscB remains an unsolved puzzle. Contrary to the RNAs derived from CRISPR arrays, ωRNAs rarely target viral DNA, excluding a role in antiviral immunity. Current hypotheses postulate that these transposon-encoded RNA-guided nucleases could be involved in transposi-

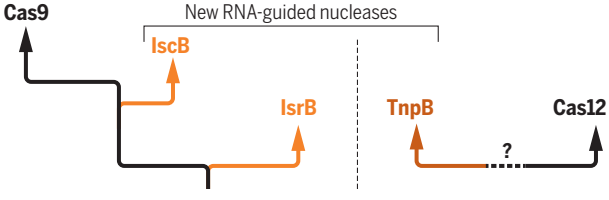
Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. Email: rotem.sorek@weizmann.ac.il

Transposon-encoded nucleases

A wealth of transposon-encoded nucleases found in bacteria may be programmed with specific guide RNAs for targeted DNA editing. ωRNAs, omega RNAs.

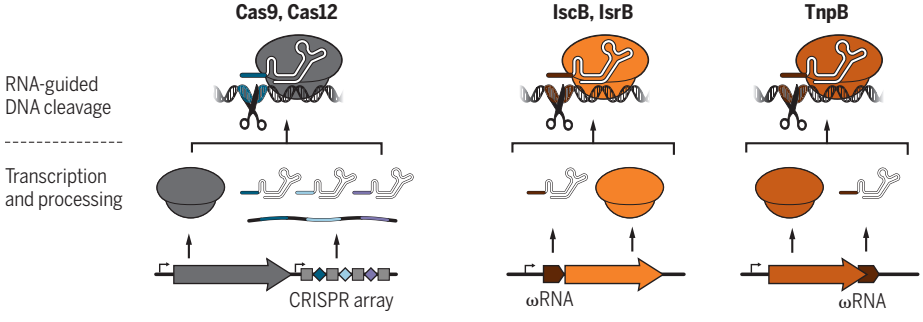
Evolutionary relationships

The origins of Cas9 are transposon-encoded enzymes of the IscB and IsrB families that are also guided by RNA. Likewise, the TnpB transposons may be the ancestors of Cas12.



Potentially programmable

RNA-guided endonucleases encoded by bacterial transposons function similarly to CRISPR nucleases Cas9 and Cas12. If tethered to guide RNAs of choice, they could be leveraged as new genome editing tools.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

tion or play a role in the maintenance of the transposon.

The discovery that transposon-residing genes can encode RNA-guided nucleases proved to be an entry point to explore transposon-encoded nucleases in general. Altae-Tran *et al.* discovered that TnpB, another family of transposon-encoded nucleases, is also RNA-guided. This finding is of exceptional importance for two main reasons. TnpB is the probable evolutionary ancestor of Cas12, another CRISPR nuclease that has been used for genome editing and diagnostic purposes (*13*, *14*). And with over a million homologs identified in microbial genomes to date, TnpB is one of the most abundant prokaryotic gene families, suggesting that RNA-guided nuclease activities are much more prevalent than previously envisioned.

Altogether, the IscB and TnpB protein families represent a giant collection of programmable enzymes that can be harnessed for gene editing. Their diversity in sequence, protein size, and PAM-like motif requirements should make them an invaluable addition to the molecular toolbox. As a proof of concept, Altae-Tran *et al.* demonstrated that DNA in human cells could be edited by using IscB proteins.

Long considered as genomic junk, transposons are now emerging as key components in the evolution of adaptive immune systems. It is well established that transposons called *Transib* gave rise to the V(D)J recombination system that generates the diversity of immunoglobulins in vertebrates (*15*), and it is now clear that transposons gave rise to the CRISPR-Cas adaptive immune system of prokaryotes. As transposons are highly diverse and have probably been associated with bacterial genomes for billions of years, it is not unlikely that they contributed to the evolution of additional immune systems, opening interesting avenues for future studies. ■

REFERENCES AND NOTES

1. E. Pennisi, *Science* **341**, 833 (2013).
2. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
3. S. J. J. Brouns *et al.*, *Science* **321**, 960 (2008).
4. M. Jinek *et al.*, *Science* **337**, 816 (2012).
5. S. S. Wu, Q.-C. Li, C.-Q. Yin, W. Xue, C.-Q. Song, *Theranostics* **10**, 4374 (2020).
6. D. Burstein *et al.*, *Nature* **542**, 237 (2017).
7. P. Pausch *et al.*, *Science* **369**, 333 (2020).
8. J. H. Hu *et al.*, *Nature* **556**, 57 (2018).
9. R. T. Walton, K. A. Christie, M. N. Whittaker, B. P. Kleinstiver, *Science* **368**, 290 (2020).
10. H. Altae-Tran *et al.*, *Science* **374**, 57 (2021).
11. F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* **155**, 733 (2009).
12. V. V. Kapitonov, K. S. Makarova, E. V. Koonin, *J. Bacteriol.* **198**, 797 (2015).
13. B. Zetsche *et al.*, *Cell* **163**, 759 (2015).
14. J. S. Chen *et al.*, *Science* **360**, 436 (2018).
15. V. V. Kapitonov, J. Jurka, *PLOS Biol.* **3**, e181 (2005).

10.1126/science.abm2239

CANCER

Identifying cancer drivers

Analysis of protein interaction networks can identify previously unknown oncogenic drivers

By **Ran Cheng**[1] *and* **Peter K. Jackson**[2]

A comprehensive protein-protein interaction (PPI) network is a critical tool for understanding how pathways are organized in normal cells and altered in cancer. For many cancers, there is an extensive catalog of genetic mutations, but a consolidated map that organizes these mutations into pathways that drive tumor growth is missing. On pages 49 and 50 of this issue, Swaney *et al.* (*1*) and Kim *et al.* (*2*), respectively, use affinity purification–mass spectrometry (AP-MS) to examine PPI networks in head and neck squamous cell carcinoma (HNSCC) and breast cancer (BC) and find many previously unknown interactions in cancer cells. On page 51 of this issue, Zheng *et al.* (*3*) combine the new PPI data with existing public data to generate a structured map of protein pathways to help validate these PPIs, suggesting a framework for future analyses that could drive our understanding of oncogenic transformation and identify therapeutic targets.

Although genome sequencing has cataloged alterations in thousands of patient tumors, the overwhelming complexity of mutations points to a single driver oncogene in only a few tumor types. Activating mutations in *KRAS* are the classical example in non–small cell lung, pancreatic, and colorectal tumors. Further, individual cancer-associated genes show a complex variety of specific coding mutations. A clearer picture would emerge if mechanisms critical for tumor growth were better consolidated into specific pathways. Identifying and consolidating these pathways and identifying how combinations of pathways drive cancer will simplify our search for effective cancer therapies. PPIs are critical because they extend far beyond gene lists to define the protein biochemistry of tumor pathways and druggable targets.

PPI studies now include increasing amounts of publicly available AP-MS and yeast two-hybrid data (*4*) that span a substantial fraction of the human proteome, but most PPIs have been identified in workhorse cell lines, such as human embryonic kidney 293 (HEK293) cells (*5*, *6*). Swaney *et al.* and Kim *et al.* show that PPIs found in cancer cells are substantially different in HNSCC and BC cell lines compared with normal cell lines from the same lineage or compared with public data from workhorse cell lines.

Swaney *et al.* focus on two human HNSCC cell lines versus a normal esophageal line and add epitope tags to 31 proteins that represent highly mutated cancer genes found in 99% of HNSCC tumors. They tag both wild-type and frequent mutant proteins, with a focus on the phosphatidylinositol-3 kinase (PI3K) pathway, critical in HNSCC (*7*). PI3K responds to nutrients, growth factors, and environmental cues to activate cell growth, metabolism, cell motility, and angiogenesis. The gene encoding the catalytic subunit of PI3K, *PIK3CA*, is mutated in tumors to create a constitutively active pathway, which makes PI3K an attractive target; indeed, there are many US Food and Drug Administration (FDA)–approved PI3K inhibitors available. PI3K is also a critical effector of the RAS oncoproteins. However, PI3K inhibitors cause toxicity in patients, and the use of combination therapies to improve efficacy is a major focus of research. Further defining PI3K interactors could yield insight into more effective and less toxic combination therapies.

Using a standardized AP-MS and analysis workflow (see the figure), Swaney *et al.* find more than 770 interactions involving ~650 proteins. The majority were not previously observed. Moreover, the cancer lines showed distinct interactions compared with both normal lines and each other. To identify the cancer-specific interactions, the authors developed a differential interaction score for which PPIs were enriched in tumor cells but not normal cells. They found many previously unknown interacting partners, which themselves had coding mutations in HNSCC compared with noncoding alterations, suggesting that specific cancer pathways are activated by coding mutations in multiple pathway members.

These new candidate cancer drivers were exemplified by a link between PI3K and the HER3 receptor—a protein that lacks clear kinase activity but potentially regulates

[1]Department of Biology, Stanford University School of Medicine, Stanford, CA, USA. [2]Baxter Laboratory, Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, USA. Email: pjackson@stanford.edu

A treasure trove of molecular scissors

François RoussetRotem Sorek

Science, 374 (6563), • DOI: 10.1126/science.abm2239