can change to any of the other three bases, generating about 10^{12} different potential versions of VR. The biochemical basis of the specificity for mutation at A is unknown, but presumably the bacteriophage RT could be 'sloppy by design' whenever an A is copied.

Insightful genetic analyses by Doulatov *et al.* also demonstrate that various genetic-selection protocols can produce bacteriophages that have undergone the diversity-generation process. The selection criteria for the altered bacteriophages include natural ones, such as being able to infect a host cell with an altered surface, and artificial ones, such as surviving DNA cleavage by a restriction enzyme. Depending on the nature of the selection, different patterns of survivor mutations were observed.

Restriction enzymes cleave DNA in strictly defined, short sequences. So it is easier to identify and interpret the specific mutations that enable bacteriophages to survive this threat (that is, mutations that prevent DNA cleavage) than to identify those that allow infection of bacteria with altered surface properties. In addition to mutations that occurred directly in the specific restriction-enzyme-recognition site, each variant typically contained a 'patch' of additional mutations flanking the site. However, the patterns of inherited mutation differed dramatically in the survivors of two different restriction-site selections, with patches centred precisely on the enzyme-recognition sites. Thus, it is the type of selection imposed that apparently determines the pattern of mutations in survivors. These results are consistent with a mechanism in which segments of the variable cDNA are randomly directed to replace the existing VR segment (but not the TR segment).

Reverse transcriptases are central to many RNA viruses, or retroviruses, where they are key to the viral replication process, through an RNA→DNA→RNA mechanism. Being error-prone is an intrinsic property of RTs, which lack the ability to 'proof-read' mistakes made in the copying reaction[3] — so the RTs also generate genetic diversity in the viral sequences. Such diversity enables the virus to evolve resistance to host defence mechanisms and drugs. However, there is a crucial and noteworthy difference between the *Bordetella* bacteriophage RT and that of all true retroviruses. The *Bordetella* bacteriophage has a large DNA genome, and it is predicted to replicate by a DNA→DNA mechanism. The bacteriophage RT is therefore completely dispensable for replication — mutants lacking it are merely unable to generate variants. Thus, it appears that the sole 'purpose' of the bacteriophage RT gene is to create sequence diversity. Perhaps similar logic can explain the long-sought 'function' of another class of RTs, the retrons[4], long known to generate peculiar but very specific nucleic-acid species through reverse transcription[5].

The success of this diversity-generating system is underlined by Doulatov and colleagues' discovery that similar cassettes occur in various bacteria and blue-green algae[1]. These cassettes all consist of an RT gene and suitably oriented TR and VR sequences, one of which is mutated at the A residues relative to the other. This widespread structural conservation suggests that RT sequences lacking any apparent replicative capability can provide a strong selective advantage, based simply on their ability to generate high levels of DNA diversity within specific genes. ■

Jef D. Boeke is at the High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.
e-mail: jboeke@jhmi.edu

1. Doulatov, S. *et al. Nature* **431,** 476–481 (2004).
2. Liu, M. *et al. Science* **295,** 2091–2094 (2002).
3. Gabriel, A., Willems, M., Mules, E. H. & Boeke, J. D. *Proc. Natl Acad. Sci. USA* **93,** 7767–7771 (1996).
4. Rest, J. S. & Mindell, D. P. *Mol. Biol. Evol.* **20,** 1134–1142 (2003).
5. Lampson, B., Inouye, M. & Inouye, S. *Prog. Nucleic Acid Res. Mol. Biol.* **67,** 65–91 (2001).

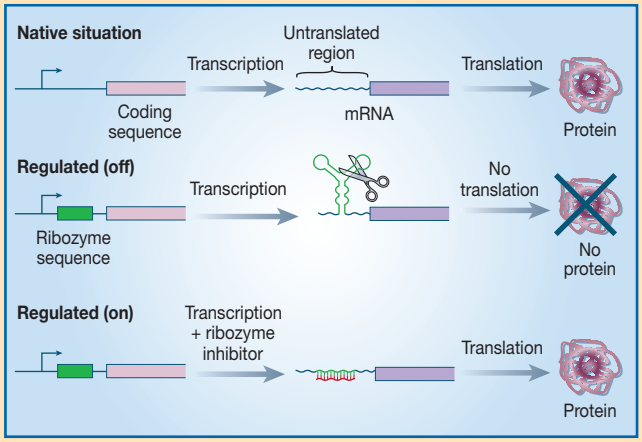
Biological techniques

Tailor-made riboswitches

Studies of certain model systems — bacteria, yeast, flies and worms — that can be manipulated genetically have been extremely useful in outlining fundamental molecular signalling pathways that are common to most organisms. But the genomes of more complex creatures, including mammals, can contain an order of magnitude more genes, many of which do not resemble those known from the model systems. Without a means of directly regulating the expression of these genes, it has been difficult to dissect their roles, particularly if they are involved in development- or tissue-specific processes. That looks set to change, however, thanks to an exciting new technique described by Laising Yen and colleagues elsewhere in this issue (*Nature* **431,** 471–476; 2004).

The discovery of ribozymes — RNA sequences that can mediate their own cleavage, in the absence of any protein cofactor — showed that RNAs had the potential to act as regulatory molecules. This potential was realized with the identification of riboswitches, messenger RNAs containing specialized ribozyme sequences whose self-cleavage is prevented by the binding of a small molecule. Natural riboswitches that recognize ligands such as metabolic intermediates or nucleotides have been described in bacterial systems, where they function to regulate gene expression.

Yen and colleagues exploited this information to develop a way of regulating gene expression in mammalian cells (see illustration). The system consists of two parts: a ribozyme sequence that is both highly efficient and constitutively active *in vivo*, and a means of modulating the ribozyme's self-cleavage activity (using either a ligand or an oligonucleotide complementary to the ribozyme sequence). If a ribozyme-encoding DNA sequence is placed upstream of the coding region of a gene of interest, the gene's expression will be repressed unless the small inducer molecule or antisense oligonucleotide is provided.



Needless to say, a certain amount of fiddling was required to find a workable ribozyme–inducer pair. Once done, however, the system could regulate gene expression in many cultured cell types and in mice; expression was almost undetectable unless the inducer molecule was present. Proof of this principle was obtained with a transgene that the authors inserted into the cells.

In conjunction with other emerging technologies, such as aptamers, it may be possible to tailor gene-regulation systems to respond to virtually any small molecule or metabolite. If ribozyme sequences are incorporated into an endogenous gene, it should also be possible to monitor that gene's function during development or in specific tissues by supplying specific concentrations of intracellular molecules. And having different ribozyme sequences will allow several genes to be examined either singly or together. **Angela K. Eggleston**