

EVOLUTION

Jumping Genes Hop Into the Evolutionary Limelight

With genomes from ancient fish to modern humans in hand, researchers are gaining new respect for the role transposable elements play in evolution

Call it a molecular gold rush. Researchers sifting through the supposed junk DNA between genes—a whopping 98% of the human genome—have in the past few years hit a mother lode of functional sequence full of clues about how genomes operate and change through time. And, as junk DNA has gained respect, so have mobile bits of DNA called transposons that are often the source of this genomic clutter.

Most researchers have taken a dim view of transposons, considering them molecular parasites that clog chromosomes with seemingly useless sequence, sometimes disrupting genes. Now, comparative surveys, along with experimental studies of gene regulation, are showing that transposons can influence when, where, and how genes are expressed. These so-called parasites “might be better viewed as symbiotic,” says Eric Lander, director of the Broad Institute of Massachusetts Institute of Technology and Harvard University in Cambridge, Massachusetts. David Haussler of the University of California, Santa Cruz, adds that “people are underestimating the impact of transposons” in genome evolution.

Transposons, small packages of DNA that can splice into other sequences, seem to appear suddenly in a genome, copying, cutting, and pasting themselves throughout its chromosomes. Eventually, the genes for making them mobile are disabled by mutations, and the copied sequences themselves mutate until they become indistinguishable from the rest

of the genome’s junk DNA. But recent studies show that some transposons don’t decay. A few have lasted hundreds of millions of years relatively unchanged and are found in the same place in the genomes of many species. To be so highly conserved, they must play a role so important to survival that evolution keeps them intact, weeding out deleterious mutations. One family of transposons, for example, first made its appearance during the evolution of tetrapods; and its descendants are still recognizable, suggesting that they may have helped shape the evolution of that particular group of animals.

Beyond gene boundaries

A few researchers suggested early on that transposons may play important evolutionary roles. Nobel laureate Barbara McClintock called them control elements when she discovered them about 50 years ago, and 2 decades later, Roy Britten and Eric Davidson suggested that they provide fortuitous opportunities for evolutionary innovation. In 1969, they proposed that new branches on the tree of life and ever-more-complex organisms arose at least in part from changes in how genes were regulated. They also argued that these changes were often caused by repetitive elements, many later identified as transposons.

Sets of genes turn on at particular times during development in specific places to specify lungs instead of gills, brain instead of kidney, and so on. Certain regulatory DNA and proteins control these

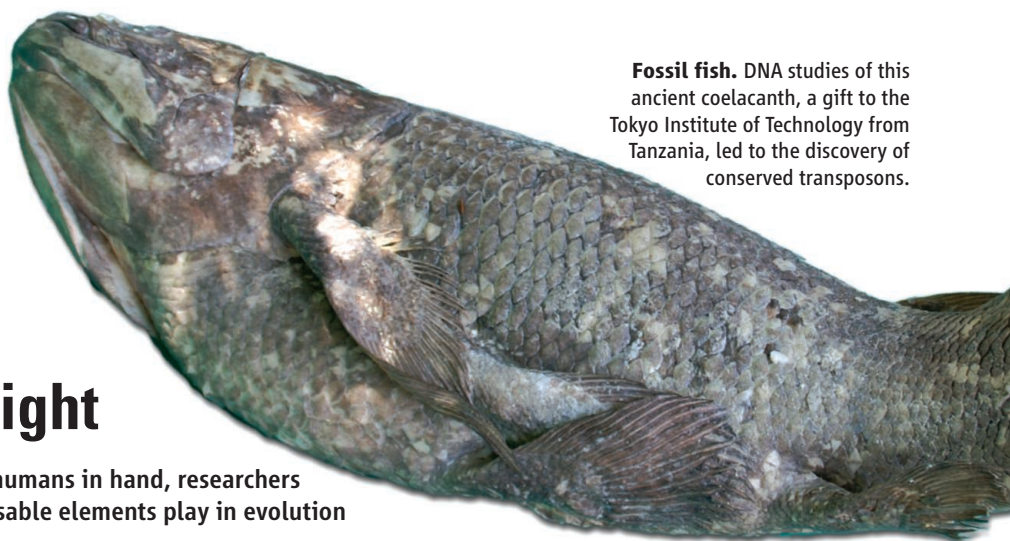
activities and, Britten and Davidson suggested, repetitive elements that copied themselves into different parts of the genome occasionally reconfigured these controls. In the new positions, this mobile genetic material could, for example, corral two independent gene networks and bring them under one regulatory roof, generating new cell types and consequently new structures. “You could have genes that are totally unrelated suddenly getting turned on in the same tissue at the same time,” explains Gill Bejerano, a molecular geneticist at Stanford University in Palo Alto, California.

Over the years, this idea lost momentum. “It was untestable,” recalls Davidson, a developmental biologist at the California Institute of Technology in Pasadena. But as researchers began comparing genomes rolling off the sequencing machines, it became clear that not all junk DNA was junk, and Davidson and Britten’s idea began to look more plausible.

In 2004, Bejerano, then at the University of California, Santa Cruz, and his colleagues described more than 400 stretches of at least 200 bases that were the same in human, rat, mouse, chicken, dog, and, to a lesser extent, fish. Three-quarters of these so-called ultraconserved regions resided outside genes (*Science*, 28 May 2004, p. 1321). Last year, Byrappa Venkatesh of the Institute of Molecular and Cell Biology in Singapore and colleagues compared DNA of elephant sharks and humans, which diverged 530 million years ago. Working with a very sketchy draft of the shark genome, they found 4800 conserved sequences. Like others, they found that these conserved sequences tended to cluster near genes for proteins that regulate transcription and DNA binding.

Greg Elgar of the University of London has found that many conserved elements burst onto the scene between the emergence of lampreys and sharks. The timing suggests

Fossil fish. DNA studies of this ancient coelacanth, a gift to the Tokyo Institute of Technology from Tanzania, led to the discovery of conserved transposons.



Transposons at work. A transposon-activated marker gene was turned on in the developing nervous system of a mouse embryo (*top*) in the same places as a key transcription-factor gene (*bottom*), indicating that the transposon helps control this gene’s expression.

that “the evolution of these sequences [was] key to the establishment of the vertebrate gene-regulatory network for development,” Elgar says.

Gradually, researchers began to realize that some of these conserved elements were transposons. The coelacanth, a “living fossil” species that dates back more than 400 million years, led Norihiro

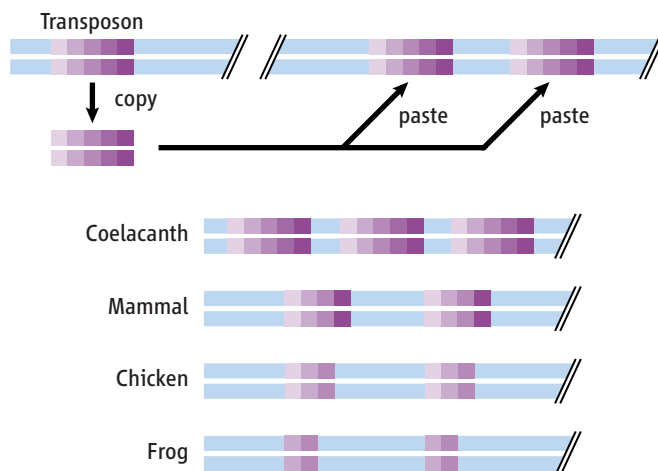
Okada of the Tokyo Institute of Technology to a whole conserved superfamily of transposons called short interspersed repetitive elements (SINEs). Okada and his colleagues first found two SINEs in the coelacanth whose sequences looked similar enough to each other to have a common origin; they then searched for the same sequences in genome databases. They found them in salmon, trout, hagfish, dogfish sharks, lancelets, catfish, zebrafish, and the sea urchin, they reported online in *Genome Research* on 22 May 2006. Yet another analysis unearthed 1000 copies of a subset of these SINEs, called AmnSINEs, in both humans and chickens. “It suggests that some of these [transposable elements] had acquired function very early on in evolution, and those functions have been retained,” says John Moran, a molecular geneticist at the University of Michigan, Ann Arbor.

When Broad Institute researchers Xiaohui Xie and Michael Kamal trekked through the human genome looking for stretches of DNA that occurred multiple times, they found one that looked quite a bit like the core of a zebrafish transposable element, also part of the SINE family. Eventually, they turned up 123 more copies, some more complete than others. They found this SINE’s 180-base core in the same places in other genomes and a few copies in the coelacanth. This work appeared in the 1 August 2006 issue of the *Proceedings of the National Academy of Sciences*. Their SINEs proved to be the same ones that Okada discovered.

When Bejerano took a close look at his ultraconserved sequences, he too discovered the remains of a family of transposable elements. He saw this first in coelacanth DNA, calling it LF-SINE, “LF” for “lobe-finned” fish. He estimates that 10,000 LF-SINEs

exist in this fossil fish and, from their sequences, knows they are still active. Through genome comparisons, he and his colleagues found LF-SINE variations in human, chicken, dog, and all the other tetrapod sequences in the public databases.

The conservation of the sequence in similar places in the genomes of all these species suggested they play a key role in genome function. Bejerano tested one that was located 500,000 bases from a gene coding for a transcription factor active in motor neuron development. That’s not close enough to be the primary controller of the gene’s activity, but Bejerano thought perhaps the transposon could exert a gene-activating effect from afar as a so-called enhancer. He linked the transposon’s sequence to a gene that would cause cells expressing that gene to turn blue when stained and put that DNA into fertilized mouse eggs. The marker gene turned devel-



Jumping genes frozen through time. Transposons copy and paste themselves, proliferating throughout a genome. Most slowly mutate beyond recognition, but sometimes they persist, albeit in truncated forms. Those conserved in many species likely serve important functions.

oping nervous system tissue blue precisely where Bejerano expected it, he and his colleagues reported 14 May 2006 in *Nature*.

Significant force

No one was sure, however, how frequently transposable elements were co-opted in this way. In theory, the odds are in the transposons’ favor. “A transposon has no problem making 50,000 copies and splattering them throughout the genome,” Haussler explains. Even if just one or two happen to land where they can be useful, they could still add up to be a powerful force in evolution. Yet “the impression was that there was a case here, a case there, that it was a really interesting fluke,” Bejerano says. Now work by Bejerano

and others has shown that functional transposable elements are more than a fluke.

Jerzy Jurka of the Genetic Information Research Institute in Mountain View, California, working with Kerstin Lindblad-Toh and Tarjei Mikkelsen, both from the Broad Institute, has found a surprisingly large role for transposons in the evolution of placental mammals. They compared the newly deciphered opossum genome to those of humans and other placental mammals to identify regions that were conserved in placental mammals but not found in this marsupial. These genetic innovations likely underlie the developmental and other differences between the two mammal groups. More than 95% of these innovations in the placental mammals were outside protein-coding genes, and 16% matched up to one of more than a dozen transposon families, the researchers reported in the 10 May issue of *Nature*. They conclude that transposable elements were likely instru-

mental for regulatory changes underlying features characteristic of placental mammals. “It’s such a significant fraction that it can’t be dismissed,” Lander says.

Overall, Bejerano and his colleagues have just found more than 10,000 conserved transposons in the human genome, many dating back well before the split between placental and marsupial mammals. He and his colleagues first identified these sequences by looking for conserved stretches across a range of vertebrate genomes, including human, and subtracting out any that represented genes. They pinpointed transposon-derived DNA by matching up the shared sequences with known mobile elements in a database. The matches represent

more than 5.5% of all the conserved noncoding sequence, he and his colleagues reported online 23 April in the *Proceedings of the National Academy of Sciences*.

At this point, the evidence for a role for most of these partially preserved transposons is circumstantial. Their conservation suggests they have a function; otherwise, they should slowly disappear. But new, more powerful tools for analyzing genomes in silico or for pinpointing where transcription factors bind DNA promise rapid progress toward understanding what these conserved transposons do (*Science*, 25 May, p. 1120). Says Bejerano: “We should have pretty spectacular answers pretty soon.”

—ELIZABETH PENNISI