Digital soup: DNA as a computational device

By MARK FISCHETTI

A SALESMAN PLANS to visit 50 cities. What's the most efficient route for him to follow? No conventional computer can calculate the answer without years of computing time. But in November 1994, Len Adleman at the University of Southern California unveiled a model for an unconventional computer that just might lead one day to a solution of the traveling salesman problem: a test-tube containing one-fiftieth of a teaspoon of water teeming with DNA.

Biologists and computer scientists had only daydreamed about DNA computing. But since Adleman's experiment, the handful of scientists trained in both computer science and biochemistry have begun to combine these fields of expertise to make test-tube trials possible. Any practical device is decades away, but the potential is awesome. A soup of 1018 DNA strands would possess 10,000 times the power of today's best supercomputers, opening the door to applications like scheduling mass transit, designing telephone networks, optimizing the layout of large integrated circuits, predicting global weather patterns, and simulating bomb explosions. Furthermore, a DNA computer could store information in one-trillionth the space of current media-if researchers can get 1018 strands of DNA to do what they're told.

Virtually no one was working on DNA computing when Adleman, known for his work on computer security and for coining the term "computer virus," published his results in Science in 1994. The problem he solved involved only seven cities, easy for most computers; the computation took about one second, after a week of lab preparation. However, it provided a blueprint for a new type of computer architecture.

Whether the approach can be scaled up to numerous cities or applied to different problems is unknown. Princeton's Richard Lipton and two graduate students, Dan

Boneh and Chris Dunworth, have designed (on paper) a DNA computing process that could solve one of the toughest problems known: cracking the Data Encryption Standard, invented by the National Security Agency and used by banks and governments for secure data transmission. Breaking this code means trying a phenomenal number of alphanumeric combinations-a centuries-long process even for a computer performing 100,000 operations per second. Advanced parallel processors have a better shot but still need massive amounts of time.

Last spring, Lipton's team unveiled a design for a DNA parallel processor. Though a single computation would take hours, 1018 DNA strands could test billions of combinations simultaneously. Lipton and colleagues described how to create this soup with gene-manipulation techniques perfected for the Human Genome Project, plus a way to encode the DNA strands to represent DES data. When a reaction is initiated, the sheer number of strands ensures that one or more will reconfigure to represent the keys to the code. "The DNA computer can break the code in about four months," says Boneh.

NO ONE HAS actually cooked up this soup. Manipulating so much DNA is painstaking work, often causing uncontrollable random errors; DNA molecules interact on an electron level and tend to self-organize. Error rates need assessment before work can proceed, Dunworth says. "For a practical device, I'd say we'd have to do about a thousand-fold better in reducing errors." More molecular algorithms will also be needed.

Finding a way to manufacture huge quantities of pure DNA is an obstacle. The polymerase chain reaction can crank out identical DNA strands-in volumes of a drop or two. Solving a traveling salesman problem with 25 variables would require kilograms of DNA and bathtub quantities of enzymes. The number of molecules required "rises exponentially" to "unrealistic amounts," according to three Oxford scientists who wrote to critique Adleman's results in the April 28, 1995, issue of Science.(1) Undaunted, Adleman is assembling a multidisciplinary group and seeking grants to investigate biological solutions that can compute.

DNA computing expands on what some call the fourth mode of computing machines (after mechanical tabulators, World War II-era vacuum-tube computers, and the transistor-based electronic computers of the 1960s): molecular computers using enzymes and proteins. The reactions of these biological materials to stimuli such as light, heat, or other molecules constitute the computer's output. DNA computers, however, are programmed like digital electronics. Base pairs represent bits of data. Nothing happens until the DNA is given an instruction to react. "DNA computing is unbiological, really, though a stimulating development," says Michael Conrad at Wayne State University, a molecular computing pioneer with credentials in biology, physics, mathematics, and computing. "It's formal computation." To foster interdisciplinary learning in this new field, Conrad and others formed the International Society for Molecular Electronics and Biocomputing in 1991.

"Basic properties of DNA are universal to every living thing-to a germ, an elephant, a redwood tree," says Isidore Edelman, professor emeritus of biochemistry and molecular biophysics at Columbia and director of Columbia's new Genome Center. "These properties are being used in genome research and in computing." Edelman describes how the Human Genome Project is inducing biologists and computer scientists to converge on a common object (DNA) and a common goal (mapping a complete consensus genome).

As for a common object for future work, Alfred Aho, chairman of Columbia's computer science department, suggests everyone look in the mirror. "What biological processes enable us to store knowledge and process it? What algorithms does our brain use?" Aho, a leader in pattern matching (algorithms used in everything from database searches to translations between computer languages), maintains that DNA computing offers a "brilliant opportunity at the intersection of biology and computer science to understand how the brain encodes itself. Knowing this, we might be able to understand mental problems."

This knowledge may prompt disturbing Turingesque questions. If DNA can be programmed, Aho says, "it is conceivable that we could clone the functionality of a human being in a computer. If you communicated with this computer and the real person and got the same responses to questions about love, religion, fear, then what are we as human beings: just higher-order data processors?"
We may never confront such questions. The brain contains billions more neuronal connections than there are base pairs in a human genome; its information-encoding processes may be incalculable. Respondents to Adleman's article have also detailed complex difficulties in scaling up his procedures, possibly limiting DNA computing to simple problems, for which conventional computers are already sufficient (or superior, since DNA computers are too slow at serial processing to compete with PCs or mainframes).

Proponents of DNA computing acknowledge the challenge-and, having learned a lesson from cold fusion about grandiose claims, allow for levity. "Our colleagues keep asking us if we have a computing vending machine yet," says Dunworth. "You know, a computer in a coffee cup: If it solves your problem, then you can drink it."

In his response to the April 28 Science letters, Adleman acknowledges the tenuous nature of claims for molecular computing. "Will it ever compute?" he writes. "It is too early for either great optimism or pessimism. Today's electronic computers are the product of a half century of extraordinary development. Molecular computers are less than a year old. Perhaps they will mature well-perhaps not."

A practical device, he explains, is not the point. Investigation of DNA computing creates a benchmark by which to measure progress-for biologists and chemists in understanding cellular mechanisms, for computer scientists and mathematicians in finding molecular algorithms, and for physicists and engineers in designing large-scale computers. The real excitement is in bringing these fields together, and perhaps developing new ones.

"We should not lose sight of the fact that the primary reason for research in this area is to elucidate fundamental aspects of computation and biology," Adleman writes. "In this regard there is reason for optimism."

1. Lo, Y.M.D., Jui, K.F.C., Wong, S.L. Letter. Science 268.28 (April 1995):481.

MARK FISCHETTI is a free-lance science and business writer whose work has appeared in Smithsonian, Scientific American, Omni, The New York Times, and the Harvard Business Review. He is a contributing editor to Issues in Science and Technology. His e-mail address is markfis@aol.com.

ART CREDIT: Debra Solomon
