February 13, 2001, Tuesday

READING THE BOOK OF LIFE; Genome Shows Evolution Has an Eye for Hyperbole

By NATALIE ANGIER

In keeping with the current fashionableness of all things Extreme -- Extreme Football, Extreme Wrestling and Extremely Large Tax Cuts -- scientists now present the Extreme Genome.

The publications this week by two teams of scientists of the rough draft of all three billion chemical bases that make up human DNA run for scores of pages and list hundreds of contributors, and their very presentations bellow the monumentality of the event.

"One doesn't want to get carried away, but I have to say I'm pretty carried away," said Dr. Francis S. Collins, director of the National Human Genome Research Institute, which coordinated the global team of scientists called the International Human Genome Sequencing Consortium. "I'd put this on the short list of big moments in biology," right up there with Darwin's work on evolution by natural selection, Mendel's discovery of the laws of genetics and James Watson and Francis Crick's discovery of the double helical structure of DNA.

Dr. J. Craig Venter of Celera Genomics, who led the private sector's counterpart sequencing project, was no less extravagant in his assessment, declaring, "It will change science and our view of ourselves as we go forward."

For some time now, scientists have been aware of most of the overall features of the human genome, and gradually have been gathering clues to the sticky, stringy, springy, dynamic, garrulous, gorgeous and preposterous molecule of life that resides in nearly every cell of every human being on earth. Now, with the entire genetic landscape laid out before them, everything looks even more exaggerated than before. The heights are higher, the brights are brassier, the exceptions are the rule, and all rules are out the window.

For example, researchers knew that only a small fraction of the genome ranked as genuine genes -- the chemical instructions for building the body's panoply of hard-working proteins. They knew that the bulk of the genome consisted of so-called noncoding sequences, rambling stretches of A's, T's, G's and C's that some dismissed as "junk DNA." By the standard estimate, about 5 to 8 percent of the genome ranked as coding sequences, with the remaining 2.8 billion or so bases the apparent product of a typing pool of drunken baboons.

Now it seems that the coding region is even tinier than imagined, comprising not about 5 percent, but a trifling 1 to 1.5 percent of the genome. It's the genome by Gertrude Stein or Dr. Seuss: is there any there anywhere?

Which brings us to another extremity revealed. Scientists knew that human DNA was lumpy -- that the genes were not distributed evenly across the cell's 23 pairs of chromosomes, but instead was arranged in patches, with some regions of the chromosomes being gene-rich, others gene-deprived.

But there's lumpy, and then there's the oatmeal in your high school cafeteria. Genes, it turns out, really like to stick together. As a result, the map of the entire genome resembles a population map of the United States, with urban areas of dense habitation, and vast rural tracts occupied by three people and their sport utility vehicles. Not only are genes distributed unevenly across the chromosomes, but so are the types of noncoding repeat sequences that make up the bulk of the genome. Repetitive sequences with an excess of the nucleotides C and G (for cytosine and guanine) tend to be found in the neighborhood of genes, while repeat sequences heavy on A's and T's (adenines and thymines) generally dominate throughout the non-gene "deserts."

The sequencing of the genome offers a definitive explanation why we see the distinctive dark and light banding patterns on chromosomes familiar to anybody who has had her, or her fetus's, chromosomes tested, or karyotyped. As it turns out, the light bands represent regions rich in G's and C's, and so are areas of comparatively high gene concentration; while the dark bands signal neighborhoods thick on A's and T's and thin on genes.

For reasons that remain unknown, some chromosomes hit the gene jackpot. Chromosome 19, for example, is among the smallest of the 23 chromosomes, yet it is the most densely packed with genes, as well as with noncoding sequences that lean toward C's and G's. As a result, Chromosome 19 in a karyotype displays very few dark bands.

Dr. Eric S. Lander of the Whitehead Institute for Biomedical Research, another consortium center, points out that the ruggedness of our genome distinguishes it from that of other species whose genomes have been sequenced, including the roundworm, the fruit fly, yeast and several species of bacteria. "Other genomes are like the plains of Kansas and tend to be flat," he said. "Ours is much more like the Rocky Mountains."

Dr. Collins suggests that the architecture of the human genome is no accident, but has been shaped over millions of years by evolution. "Otherwise there would be a deterioration toward the mean," he said. "But the genome has maintained a broad spectrum of mountainous landscapes, which means that it must be good for us."

The exaggerated topography of the genome reveals nothing so much as evolution's taste for hyperbole, theatrics and opportunism, though in this case played out on a microscopic scale. It is a genuine jungle in there, a wild ecosystem of competing "species" of DNA freeloaders, DNA symbionts, endangered DNA sequences and rotting DNA fossils.

Our genome is a La Brea tar pit, a spectacular record of the genetic history of life on earth. A number of our sequences date back 700 million years or more, when we lived a tidy, unicellular life, while other chemical phrases bespeak our membership in the fraternity of apes, and may explain how we came to be human. Some stretches of base pairs are ancient gifts left behind by infecting viruses and bacteria. The human immune system, which is capable of cutting and pasting together different genetic segments to spin off a staggering variety of warrior immune cells and antibodies, may well have

adopted the editing trick from a virus that was cutting and pasting DNA into its host many millions of years ago. The gene that encodes monoamine oxidase, an important degradative enzyme for the central nervous system, was bequeathed to an ancestor's cell by bacteria, as were about 230 other genes in our genomic toolbox.

"The fact that there's been lateral transfer of genetic information from bacteria to vertebrates suggests that the architecture of our genome is not sacrosanct or static, but extraordinarily fluid and dynamic," said Dr. Leroy Hood of the Institute for Systems Biology in Seattle, a member of the international consortium.

For better or worse, the genome passes around reams of information intramurally as well. The great bulk of the noncoding sequences in human DNA are not foreign-born, but represent the offspring of bits of genetic material that long ago broke away from a chromosome or part of the cell's RNA and protein synthesis machinery and decided to go freelance. Called jumping genes or transposons, the little entrepreneurs figured out how to reproduce themselves and reinsert their copies -- their progeny -- back into the mother genome.

Researchers examining the landscape of the human genome have identified four classes of transposons. One class, called DNA transposons, appears to be dead, mere fossils in the DNA that have lost the signals they need for effective replication, are incapable of ever jumping anywhere again and are gradually decaying out of existence. Another class, called the LTR transposons, is, said Dr. Lander, "on the critically endangered list" and may soon go extinct.

Only two transposon families in the genome are considered active, replicating themselves tidily in the course of transmission from parent to child. One is a genuine parasite called a LINE, for Long Interspersed Element, which encodes instructions for everything it needs, from copying its DNA into the intermediary form, RNA, and copying the RNA back again into DNA, and then, voilà, hopping back onto its little chromosomal niche.

"It's the ultimate selfish element that evolved at the beginning of eukaryotes," said Dr. Lander, referring to any animal, from yeast on up, whose cells have nucleuses in them. "It's been wildly successful. It's the perfect parasite."

Significantly, most of the LINE's in the human genome are located in AT-rich regions of the chromosomes, far from genes, and therefore far from the DNA quality control machinery that attend to the genes' well-being and might try to sniff them out and eliminate them.

Ah, but as a flea hath smaller fleas that on him prey, so the LINE parasite itself has a parasite -- the infamous Alu element. With more than a million copies scattered across the chromosomes, Alu elements, which run about 300 bases in length, are the most abundant sequences in the human genome. Molecular biologists have long despised them for getting in the way of their efforts to clone genuine genes.

Alus cannot replicate on their own, but instead "borrow" the machinery of the larger LINE elements to reproduce. At first glance, they seem to be even more perfect in their parasitism, and like a real parasite, they can harm their host. On occasion, in the course of the creation of an egg or sperm cell, a replicating Alu sequence will be inserted in the midst of a critical gene, resulting in a child with a genetic disease.

Yet the sequencing of the genome lends strong support to a theory proposed a couple of years ago by Dr. Carl Schmid of the University of California at Davis and his colleagues -- that the Alus may have begun life long ago as rank parasites, but they have since been coopted by the human genome to do useful work. Dr. Schmid proposed that the genome used the Alu elements to help modulate the body's response to stress, from, say, excessive exposure to heat, or a few too many martinis. Through experiments in which mice were dunked in hot tubs or made to drink lots of alcohol, Dr. Schmid and his co-workers observed that the rodent equivalent of Alu sequences were activated by such stressful events.

A scan of the human genome suggests that genes do benefit from having Alus in the neighborhood. Whereas most species of transposons are located in the AT deserts of the chromosomes, Alus are preferentially clustered in the GC regions, shoulder to shoulder with genes. "It looks like they're being selectively held onto, that the people in the living room -- the genes -- like having these guys as companions because they're useful to human biology," Dr. Collins said. "That's quite a shock, and the natural conclusion is, this part of our junk DNA isn't junk."

It's just possible that Alu sequences helped make us human. Dr. Wanda F. Reynolds, a molecular biologist at the Sidney Kimmel Cancer Center in San Diego who studies Alu elements, points out that Alu elements are found only in higher primates, and that the core of the Alu sequence is responsive to a large family of receptor proteins, the so-called nuclear receptor superfamily.

These receptors are the cell's way of recognizing potent hormones like estrogen, retinoic acid and thyroid hormone. Hence the presence of an Alu sequence in or around a gene may result in the gene being pitched a little higher or lower, turned up or down, should the appropriate hormone come calling.

"This gives the genome a great deal of plasticity, and a lot of choices," she said. "The presence of so many Alu sequences throughout the genome allows evolution to ask, What would happen if I raise the expression of Gene X threefold during development -- does that make a better primate? And what would happen if I slightly modify the expression of a thousand genes -- does that make a human being?"

After all, Dr. Reynolds says, "our genes are almost identical to chimpanzees. Something had to happen that made us different from chimpanzees."

In her view, a change in the expression of genes, in the volume and timing of their activation, could have a more revolutionary effect on development than would a mutation in the genetic sequence proper.

In evolution as in life, junk is in the eyes of the beholder, and one person's flea market giveaway is another person's windfall on eBay.

Organizations mentioned in this article:

Related Terms:
Genetics and Heredity; Deoxyribonucleic Acid; Proteins
