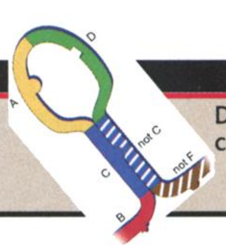
PAGE 1150
Sequencing endangered species



1152
DNA computing

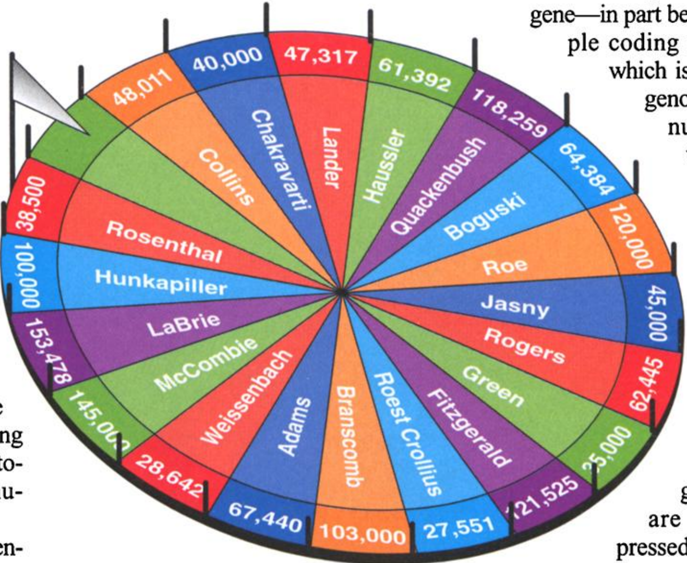
HUMAN GENOME PROJECT

And the Gene Number Is ...?

COLD SPRING HARBOR, NEW YORK— Even though a draft sequence of the human genome is nearing completion, biologists still don't know how many genes it contains. Indeed, the range of estimates seems to be growing rather than shrinking. The question lies at the core of our understanding of genetic complexity. If genomes are the books of life, then genes are the words that tell the story of each organism. Biologists have long assumed that microorganisms are short stories and complex organisms such as humans, great tomes.

But last week at a meeting* here, the generally accepted human "word" count of 80,000 to 100,000 took a battering when researchers from Germany, the United States, and France offered revised estimates of the number of genes—all of them well below 50,000. The talks, some of which will be published in the June issue of *Nature Genetics*, sparked heated debates, with at least one genomicist countering with a new 100,000-plus estimate. The liveliest discussions focused on the terms and estimates for a $1-a-bet gene-count sweepstakes. Early entries ranged from less than 30,000 to more than 150,000, and updates and rules are posted at www.ensembl.org/genesweep.html. The winner will be picked in 2003 and will take home not only the pot but a leather-bound copy of *The Double Helix*, signed by Cold Spring Harbor Laboratory's James Watson, co-discoverer of the structure of DNA. By then, the number of genes should be clear, or at least clearer, says Ewan Birney, the computational biologist from the European Bioinformatics Institute who started the competition.

Gerald Rubin, head of the Berkeley Drosophila Genome Project, set the stage for the current debate when he pointed out that the fruit fly has some 5000 fewer genes than the supposedly simpler nematode worm. "Complexity is not in any simple way related to gene number," he warned the audience.



Place your bet. Uncertainty over the number of human genes has sparked a debate—and a betting pool.

Then the next day, André Rosenthal of the Institute of Molecular Biotechnology in Jena, Germany, concluded his talk about the recent completion of chromosome 21—published in the 18 May issue of *Nature*—with a reality check on the current gene estimate. Rosenthal posited that the gene distribution of chromosomes 21 and 22, which was finished last December, should be roughly representative of the entire genome, as 22 seems to be gene rich and 21 gene poor. If the human genome is 3.3 billion bases long, then these two chromosomes and their genes (about 770) account for 2% of the genome. By this logic, he said, "we arrive at less than 40,000."

A few hours later, Hugues Roest Crollius, a molecular geneticist at Genoscope in Evry, France, came up with an even lower estimate, based on comparisons between the existing human sequence and the sequence of a freshwater puffer fish, *Tetraodon nigroviridis*. Roest Crollius, Jean Weissenbach, and their colleagues are using the puffer fish genome to identify evolutionarily conserved genes in the human genome. After sequencing about one-third of *Tetraodon*'s compact genome, he matched that DNA up against the finished human chromosomes and 800 million bases of rough draft human sequence deposited in GenBank by the end of 1999. Preliminary analysis with a new gene prediction program that identifies shared coding sequences indicated that there are 2.58 to 3.18 so-called evolutionarily conserved regions per human gene—in part because genes consist of multiple coding regions, more than one of which is likely to be present in both genomes. Extrapolating from the number of these regions, he thinks the human genome has between 27,700 and 34,300 genes.

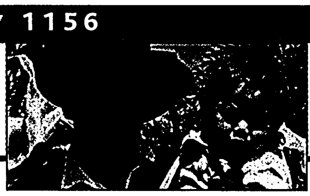
Then Phil Green, a genomicist at the University of Washington, Seattle, who has developed the widely used PHRED and PHRAP programs for assembling sequence data, said that he, too, had come up with a much lower gene count. Green's numbers are based on analyses of expressed sequence tags (ESTs), bits of genes identified during massive screenings of different tissues. Green, among others, has worked out computer programs that look for overlapping sequence between ESTs and connect those that belong to the same gene. Based on the EST representation of the genes in chromosome 22, he thinks the total number of genes is about 35,000. With Rosenthal's and Roest Crollius's data, "we now have three independent calculations that give these low numbers," says Green. "I'm pretty confident that it's in that range." Francis Collins, head of the National Human Genome Research Institute, came out on the low side too, placing his $1 bet on a mere 48,011 genes. Recent gene counts have been inflated, he said—particularly by genomics companies who boast that, "My list is bigger than yours."

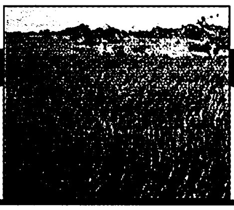
But some researchers are holding out for an *Anna Karenina*–sized word count, arguing that human complexity cannot be explained any other way. John Quackenbush of The Institute for Genomic Research (TIGR) in Rockville, Maryland, put his dollar on 118,259 genes, an estimate based on the genes in TIGR's human gene index. He also bases his estimate on ESTs but uses a different gene-building program from the one Green developed. What's more, he said, in the microbial genomes that TIGR has sequenced over the past several years, extra genes have always popped up. Says Quackenbush, "We still have a lot to learn about what's in the human genome." Nor was his

CREDIT: ILLUSTRATION BY C. FABER SMITH

---

* Genome Sequencing and Biology took place 10 to 14 May at the Cold Spring Harbor Laboratory in New York.





sweepstakes entry the highest. A 153,478-genes entry came in from Sam LaBrie of Incyte Genomics, a California-based company that in September 1999 announced that the human genome had at least 140,000 genes.

But Rubin and other gene minimalists were undeterred, arguing that complexity comes from how genes are regulated or expressed—not in the number of genes themselves. "We don't need large numbers of genes to be an intelligent species," Rosenthal explains. Weissenbach agrees, arguing that "once people go through the EST data, they will realize that many of the clusters [cover] the same gene." His bet: 28,700. And yours?

–ELIZABETH PENNISI

ASTRONOMY

Fly's Eye Spies Highs in Cosmic Rays' Demise

LONG BEACH, CALIFORNIA—On an October evening in 1991, an extraterrestrial intruder tore through the sky over North America. The interloper was a cosmic ray, a high-energy particle from deep space. Astronomers at the Fly's Eye observatory in Utah detected it as a flicker of light from the cascade of secondary particles it left in its wake. On analyzing the glow, they realized that the particle had packed a shocking 320 exa–electron volts (EeV) of energy, the same punch as an 88-kilometer-an-hour baseball pitch. Astrophysicists call it the "Oh-My-God" particle.

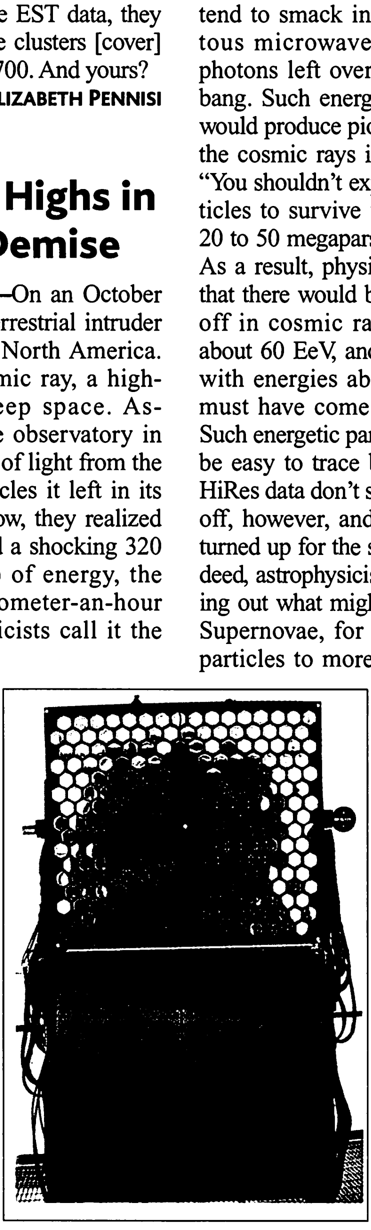
Now, observations from Fly's Eye's replacement have shown that the 320-EeV monster was not a freak of nature. This month at an American Physical Society meeting here, astronomers from the High-Resolution Fly's Eye detector, or HiRes, announced that they had found a handful of ultra-superhigh-energy cosmic rays that deepen the mystery about the origin of these strange particles.

Since it began operating in 1997, HiRes has nearly doubled the number of known ultrahigh-energy cosmic rays. It has recorded seven particles with energies greater than 100 EeV, including one that clocked in at an estimated 280 EeV, the second highest cosmic ray energy ever recorded. HiRes has also detected 13 other cosmic rays with energies over 60 EeV. "People are beginning to think these events are real," says Charles Jui, a physicist at the University of Utah in Salt Lake City.

They had reason to be skeptical. In the 1960s, scientists realized that particles with more than about 60 EeV of energy would tend to smack into the ubiquitous microwave background photons left over from the big bang. Such energetic collisions would produce pions, destroying the cosmic rays in the process. "You shouldn't expect these particles to survive for more than 20 to 50 megaparsecs," Jui says. As a result, physicists expected that there would be a sharp cutoff in cosmic ray energies at about 60 EeV, and any particles with energies above that level must have come from nearby. Such energetic particles, they thought, should be easy to trace back to their sources. The HiRes data don't show an abrupt 60 EeV cutoff, however, and no obvious sources have turned up for the superenergetic particles. Indeed, astrophysicists have trouble even figuring out what might give them such a punch. Supernovae, for instance, can't accelerate particles to more than about 1/1000 of an EeV of energy. "This is a burning astrophysical problem that needs to be solved," says University of Chicago astrophysicist Rene Ong.

Astrophysicists are hoping that HiRes, with its high sensitivity and ability to pinpoint a cosmic ray's direction and mass, will eventually provide the data that point to the answer. "Given that they've already doubled the sample, every few years they'll probably double it again," Ong says. "HiRes could make a very important statement."

–CHARLES SEIFE



Fly specs. Clusters of phototubes give HiRes an eye for cosmic rays.

CREDITS: (TOP) MICHAEL TURCO; (BOTTOM) D. BIRD

CONSERVATION BIOLOGY

Orangutans Face Extinction in the Wild

LISLE, ILLINOIS—Orangutans, our third-closest living relative, are in crisis, reported several leading primatologists at a meeting here last week on apes.* Indeed, the plight of the orangutans, whose range is now restricted to the shrinking forests of Borneo and Sumatra, dominated the 4-day meeting and sparked two late-night sessions at which researchers and zookeepers hatched a conservation plan. Without urgent action, warns Biruté Galdikas, a biological anthropologist and conservationist who teaches at Simon Fraser University in Burnaby, British Columbia, the apes could be extinct in the wild within 20 years. The culprit is wholesale logging and other habitat destruction, which will be difficult to halt.

Ignorance is bliss. A young orangutan plays in the water, heedless of the threat.

Exact numbers are hard to come by because orangutans, highly intelligent but mostly solitary apes, hang out high in the forest canopy where they are difficult to spot. But two recent surveys, one of which was presented at last week's meeting, show clearly that populations are crashing. Carel van Schaik of Duke University and his colleagues estimated the size of orangutan populations in a section of the 24,000-km^2 Leuser region on Sumatra in 1993 and 1999. Relying on data from satellite imaging and aerial photos of the animals' habitat—combined with their own knowledge from 24 years of field-

---

* "The Apes: Challenges for the 21st Century," 10 to 13 May, Lisle, Illinois, sponsored by the Brookfield Zoo.