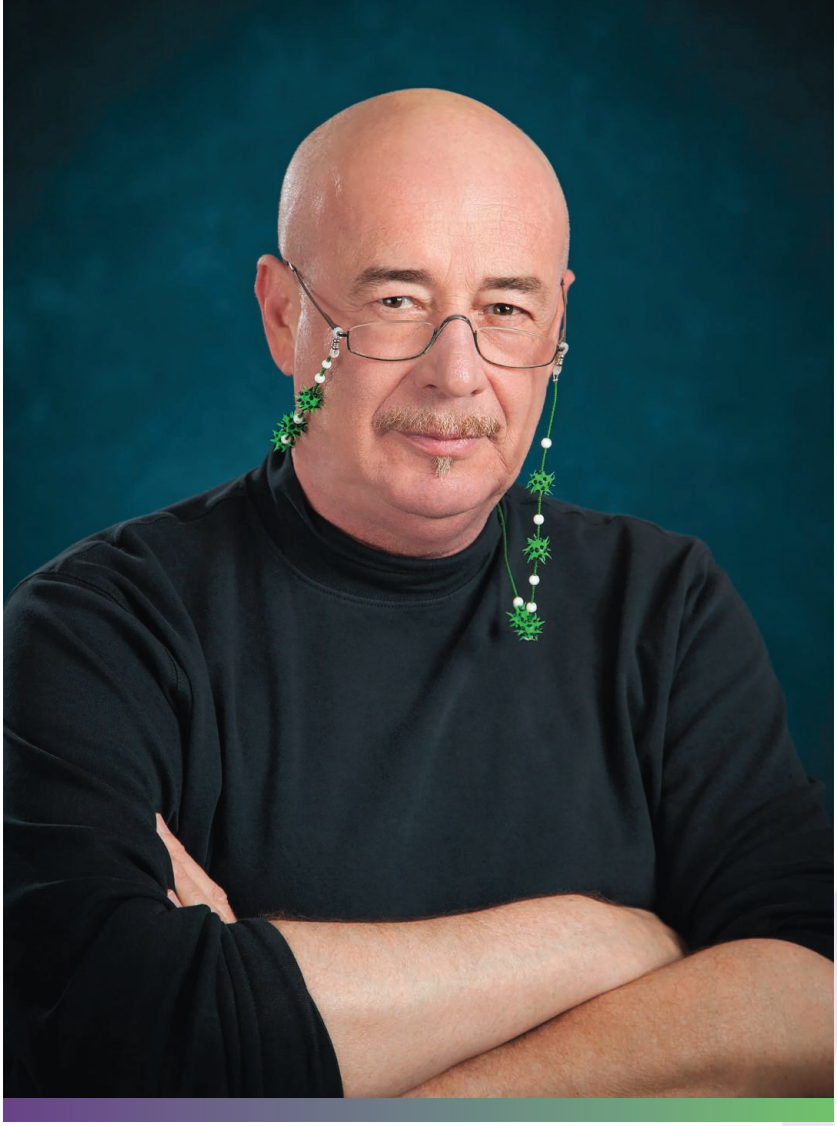

# The Vigilante

When the ENCODE Project declared that there is no such thing as junk DNA, Dan Graur counterattacked. But does he go too far?

Last year on 11 July, 2 weeks before his 60th birthday, Dan Graur was at the Society for Molecular Biology and Evolution's conference in Chicago, preparing to deliver a scathing criticism of ENCODE, the biggest genomics project funded by the U.S. National Institutes of Health (NIH) since the sequencing of the human genome. An imposing 6 feet 3 inches who likes to wear Hawaiian shirts that flow smoothly over his bulging midriff, Graur speaks with a strong Israeli accent and a deliberate enunciation that lends a scalpel-like sharpness to the sarcasm with which he dissects the world. Besides food and coffee, both of which he consumes immoderately, Graur relishes what he considers to be the unvarnished truth. When a student remarked to Graur—in response to his lament about turning 60—that age was all in the mind, Graur offered a trademark blunt retort. "No," he responded. "It's not in my mind. It's in my knees, my prostate, and my lower back. So go away."

Graur's talk that afternoon was an encore to a paper he had just published with two colleagues assailing the claims made by ENCODE, short for the Encyclopedia of DNA Elements. Launched in 2003 as a successor to the Human Genome Project, ENCODE's goal was to identify all the functional elements of the human genome, in addition to the 21,000 genes that make up a mere 1% of its 3 billion nucleotides. The co-author of a well-regarded textbook, *Fundamentals of Molecular Evolution*, Graur had been dimly aware of ENCODE's existence until the fall of 2012, when the consortium behind the project announced the first comprehensive results of the 6-year-long endeavor with the simultaneous publication of 33 papers in five journals, including *Nature* and *Science*. ENCODE's signal claim, highlighted by the team in the main *Nature* paper, was that its data "enabled us to assign biochemical functions for 80% of the genome, in particular outside of the well studied protein-coding regions."

ENCODE's leaders drove home that point in videos released by their institutions. "There is not a single place in the genome that doesn't have something that you might think could be controlling something else," said Ewan Birney, the lead analysis coordinator of ENCODE at the European Bioinformatics Institute near Cambridge, U.K., in one of the videos. In another video produced by the National Human Genome Research Institute in Bethesda, Maryland, Michael Pazin, the institute's program director for functional genomics, proclaimed: "Very little of our genomes are junk."

That finding challenged a widely held view, formed after decades of research in evolution and population genetics, that much of the human genome is nonfunctional junk. Other work had already found hints of function in some of the "junk." But Graur found ENCODE's blanket claim patently untrue. To a man of Graur's skeptical constitution, this made ENCODE an irresistible target, a plump duck calling out to a hound dog. Taking the podium in Chicago, he tore into the project.

The heart of his critique was that ENCODE researchers had made an unwarranted leap in the interpretation of their data. The project involved thousands of experiments. In some, researchers exposed cells to a multitude of transcription factors: molecules that bind to genomic DNA to initiate transcription into RNA, the first step in making a vast array of proteins required for metabolism. In other experiments, researchers identified and inventoried the different RNA molecules produced in various types of human cells. The results showed that more than 70% of DNA in the genome is transcribed into RNA; 8% latched on to transcription factors. Altogether more than 80% of the genome showed some kind of biochemical activity—the basis for ENCODE's claim that 80% of the genome is functional.

That inference, Graur inveighed, was utterly wrong because the mere transcription of a stretch of DNA or the binding to a transcription factor is not a function unto itself. He didn't say it simply; he said it with merciless mocking that, to some, undermined his message. "Graur wrote such a negative paper that it was hard to read," says Bradley Bernstein, an ENCODE researcher at Harvard University. Graur's criticism is so over-the-top that it's not worthy of a response, Bernstein adds. In his Chicago talk, Graur showed a photograph of chewing gum stuck to a shoe as an example of "a function that fits the ENCODE definition." "The fallacy of ENCODE's logic," he said, is this: "We know that some functional regions are transcribed. Ergo, all transcribed regions are functional." Toward the end of the presentation,

CREDIT: CHRIS WATTS/UNIVERSITY OF HOUSTON





Published by AAAS

he showed a photograph of dollar bills taped together in the shape of a toilet paper roll—his view of what ENCODE had achieved with the $288 million spent on the project so far.

Graur isn't the only one who has taken ENCODE to task. Others have made some of the same criticisms, including prominent biochemist W. Ford Doolittle of Dalhousie University in Halifax, Canada, who published a critique in the *Proceedings of the National Academy of Sciences* a month after Graur and his co-authors published theirs in *Genome Biology and Evolution* (*GBE*).

But if ENCODE has a bête noire, it is Graur. "The splashiest part of ENCODE was a conclusion that could not hold up, and Dan pointed it out in a way that was impossible to ignore," says Harmit Malik, an evolutionary geneticist at the Fred Hutchinson Cancer Research Center in Seattle, Washington. "No matter what anybody might think of his style, the points he has raised are very meaningful."

**Line of fire.** Claims about ENCODE's findings made by Ewan Birney (holding mic) and other project leaders at a September 2012 press briefing stoked Dan Graur's ire.

CREDIT: ANDREW MATTHEWS/PRESS ASSOCIATION VIA AP IMAGES

Doolittle, who calls Graur one of the "bad boys of molecular evolution," agrees. "As a reviewer of his manuscript, I did suggest he tone it down a little bit at certain places and he did," Doolittle says. "I think people like Dan are very useful. We simply do not do enough debunking in science these days. We have moved into a very positivist mode where everybody is expected to simply get with the program."

The few ENCODE scientists who've responded to Graur's criticisms say these are off the mark or blown out of proportion. And judging by the continuing flow of funds to the project—$30 million and counting since September 2012, for characterizing the behavior of genomic elements in more types of human cells—Graur's furious attacks have left ENCODE unscathed.

## Faultfinder

I met up with Graur on a rainy day last December at the University of Houston in Texas, where he has been a professor of molecular evolutionary bioinformatics since 2003. When he saw me watching squirrels, which routinely surprise visitors on campus by coming within stomping distance of people to beg for food, he noted dryly that the animals are simply "rats with good PR." Walking through the drizzle, he made a series of sardonic remarks about himself and the world, much as a standup comic might. He said he'd taken to wearing colorful shirts to work because he had been told that his earlier habit of wearing black intimidated students.

Born in Romania, Graur moved to Israel with his family in 1964, when he was 11. He describes himself as being a "goody two-shoes" growing up—a dubious claim in light of the fact that he was thrown out of school in 10th grade for writing off-color jokes in the school newspaper. (His only regret is that the jokes weren't funny.) He went to a technical school to be a lab technician, but was thrown out of there, too, after 2 years for making a political joke. (This one, he claims, was funnier.) He went on to serve in the Israeli army, where one of his field assignments was to lug generators for radio sets during the war with Egypt in 1973.

After his army stint, Graur studied chemistry for his undergraduate degree. Later, after getting a doctorate from the University of Texas, Houston, he taught at Tel Aviv University until 2003, when he turned 50 and grew restless. "At that age, people change their car or wife or computer system," he said. "I changed universities." The move brought him back to Houston, where he has spent the past decade producing papers on genomic evolution, with a focus on the comparative study of genomes. His other passion is collecting modern art, including a number of creations made from household junk. He wears his atheism on his sleeve: One of his pastimes is needling a devout Christian in his department with questions about the veracity of various biblical stories. Another is challenging antiabortion campaigns run by religious groups on campus.

Graur is given to intemperate griping over whatever he finds silly or stupid or wrong. By his own admission, he has a streak of vigilantism: On occasion he'll produce a serious paper that debunks someone else's finding. In 2001, he and a colleague at Tel Aviv University published a genetic analysis showing that a bacterium claimed to be 250 million years old was likely just a modern strain. Another team confirmed that Graur was right. When we met in December, he was getting ready to publish a study designed to poke statistical and analytical holes in a claim that the last common male ancestor of humans walked on Earth 338,000 years ago. On his personal blog, labeled Judge Starling (Judge is "Dan" in Hebrew; Graur is "starling" in Romanian), he regularly excoriates science in his field that he deems shoddy or hyped.

Graur's atheism inflamed his anger at ENCODE. He perceives an echo of intelligent design in the consortium's "80% claim," which he takes to imply that most of the genome exists because it serves a purpose. "What ENCODE researchers did not take into account," he contends, "is that everything is shaped by evolution." And evolution is slow to weed out useless features.

Genetic mutations—the drivers of evolution—occur at random, and those that are deleterious are weeded out, sometimes over many generations. Other mutations, salubrious and inconsequential alike, get passed down to progeny. As a result, species like humans and elephants that have a small effective population size are expected to accumulate a lot of junk in their genomes.



Published by AAAS

Various lines of evidence support the idea that vast genomic tracts in many species are littered with junk, he says. One is the surprising lack of correlation between an organism's complexity and the size of its genome. (The onion's genome is five times larger than ours.) Researchers have also discovered that more than 70% of the human genome is interspersed with repetitive stretches of DNA known as transposable elements, which are mostly inactive. Similarly, researchers have identified nearly as many defunct genes and pseudogenes in the human genome as genes.

The true benchmark of functionality, Graur and many others say, is whether a DNA sequence has been conserved over time. Because mutations in functional regions of the genome are likely to impair function, and thereby threaten survival, such mutations are expunged from the population. From this, researchers infer that functional regions evolve much more slowly than the rest of the genome and are conserved; that is, such regions can be expected to show up as identical or similar in genomes across and within species. By sequencing and comparing genomes of different species, researchers have estimated that only 5% to 15% of the human genome is functionally relevant.

"Is provocation necessary to get attention from a 'big science' consortium such as ENCODE?"

—*NATURE METHODS* EDITORIAL

To ENCODE researchers like Bernstein, conservation is too narrow a criterion for pronouncing a region of the genome to be functional. But Graur says that view is tantamount to saying that "evolutionary laws governing all known functions in the genome do not apply to the 'functions' defined by ENCODE."

He alleges that ENCODE leaders made such broad claims because they wanted to create a media splash that would justify the project's cost. "They needed to have something big to say," Graur says. "Why did they want to publish all the 30-some articles on the same day? Because they wanted a public relations impact."

"Science is not about abiding by a code of behavior put forward by Miss Manners."

—DAN GRAUR

Graur contends that ENCODE is an example of how big science can go wrong. "When the average grant size in the biomedical sciences has been halved compared to 10 years ago, this is a scandal," he says. "If you pour $288 million into one project, you do not fund 500 other projects. You kill the careers of young scientists. They are reduced to becoming technicians."

### No meeting of minds

Graur's strong words have struck a chord with some. On his webpage at the University of Houston's site, he has posted some 50 e-mails of endorsement he got from researchers soon after the publication of the March 2013 critique. "Thank you for publishing your paper about ENCODE in GBE," reads one. "[Y]ou proved what many of us thought, but didn't have the time or the courage to state." Since the Chicago conference, Graur says he has received several invitations to deliver his talk on ENCODE. "I seem to have tapped a very big anger," he says.

At the same time, Graur's combative approach has earned disapproval from some quarters. "Would a dispassionate and polite reply have been less visible?" *Nature Methods* asked in an editorial last fall that slammed Graur for engaging in what the journal saw as uncivil discourse. "Is provocation necessary to get attention from a 'big science' consortium such as ENCODE? We do not think so."

Birney and other ENCODE leaders have not engaged Graur directly. Birney did not respond to multiple requests from *Science* seeking comment on Graur's criticisms. On his blog, however, Birney appears to have backtracked from the use of the term "biological function" in summarizing ENCODE's results. He wrote that ENCODE had revealed 80% of the genome as having "specific biological activity," following up in a subsequent blog post that "we could have used different terminology to convey the concepts, consequence and massive extent of genomic events we observed." The consortium chose the 80% figure—he wrote—because it "brings home the impact of this work to a much wider audience."

Bernstein says ENCODE's value is evident in the hundreds of papers based on project data. As an example, he points to a paper in the *American Journal of Hematology* last November reporting the discovery of mutations associated with X-linked sideroblastic anemia. The mutations—identified through the genetic study of five families that suffer from the congenital disease—are located within a stretch of "junk" DNA that ENCODE had highlighted, which is now known to enhance expression of the *ALAS2* gene.

Graur dismisses that example. The mutations were not discovered because of ENCODE, he points out. After their discovery, the researchers found that the mutations' location was on ENCODE's long list of sequences with some biochemical function. "So what?" he asks. "ENCODE claimed that 80% of the genome is functional. Therefore 80% of all truly functional elements that have been discovered or will be discovered will be found in ENCODE by chance alone."

Graur and other critics place undue emphasis on the 80% figure, says John Stamatoyannopoulos, an ENCODE principal investigator at the University of Washington, Seattle. The real take-home lesson, he says, is that "there is a tremendous amount of activity encoded in the genome"—much more than researchers had suspected.

Given the current state of knowledge, Stamatoyannopoulos says, scientists need to remain "fairly agnostic" about the potential function of various genomic elements. In other words, while the likes of Graur are asking, How do you know it's functional? Stamatoyannopoulos and others are asking the opposite: How do you know it's not?

That logic infuriates Graur. "If you don't know a function, assume as a null hypothesis that it doesn't have function, and if you find a function, you'll refute the null hypothesis," he says.

I asked Graur if his detractors were right in calling him rude. He didn't think so; moreover, he felt rudeness was irrelevant to the discourse. "Science is not about abiding by a code of behavior put forward by Miss Manners," he told me. "In science, a strong voice is sometimes needed to fight self-promotion and self-delusion."

**–YUDHIJIT BHATTACHARJEE**



Published by AAAS