# Taking life to bits

**How many genes does an organism need to live? *William Wells* talks to researchers who are on a mission to build the smallest bug in the world**

AT HOME in the human body, *Mycoplasma genitalium* enjoys an enormously varied diet of delicious chemicals. Outside the body, it demands no less. Clyde Hutchison describes its daily meal as "horribly complex", a bacterial banquet including several ingredients that most microbiologists call "undefined", and Hutchison calls "cow bits".

But while *M. genitalium*'s diet is complex, its inner workings are simple. A model of efficiency, *M. genitalium* gets by with only 468 genes, making it the simplest free-living organism known. Viruses are far simpler, but they need help from the proteins of a host cell. *M. genitalium* makes all the proteins that it needs.

Hutchison, who works at the Institute for Genome Research (TIGR) in Rockville, Maryland, believes he can simplify things further. He is dismantling the bug, one gene at a time, in an attempt to work out the minimum genome, the smallest group of genes capable of sustaining independent life. Others are taking a bottom-up approach to the same end. Researchers like Evgeni Selkov at the Russian Academy of Science in Pushchino are using computers to model the minimum number of biochemical pathways, and, by working backwards, the minimum number of genes, needed to create a virtual bug.

With such information, Hutchison, Selkov and their colleagues may finally be able to understand how everything in a cell—that's every single one of the hundreds of thousands of molecules—interacts with everything else to keep life ticking over. The path to this ultimate understanding could help to create super-efficient, living chemical factories, and give hints about how life emerged from the primordial slime.

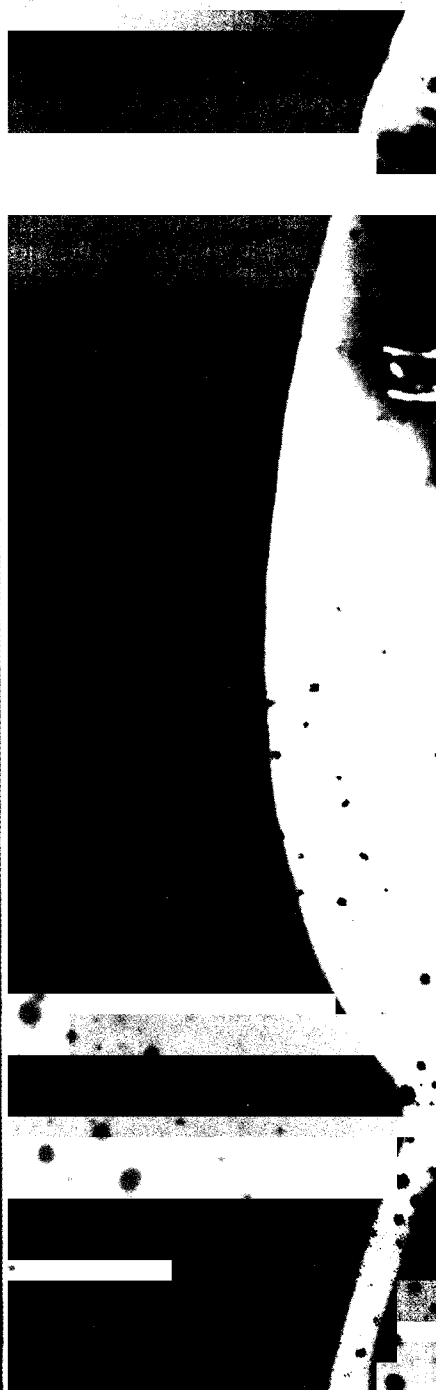
One of the participants calls it "a basic research project of breathtaking beauty". But not everyone is so enthusiastic. One sceptic calls the whole concept of a minimum organism "fallacious".

**'It's just a list of genes. Nobody knows for sure whether, if it were a real bug, it could survive'**

## Genetic jumbling

Minimum organism research is only possible because of genome projects—efforts to decode the entire DNA encyclopedias of various bacteria, yeast and (by 2005) humans. TIGR scientists were the first to blast their way through an entire genome, initially for a usually harmless resident of our lungs called *Haemophilus influenzae*, and then for the diminutive *M. genitalium*, which infects the lung and the genital tract. As these genomes became available, Eugene Koonin, a computer scientist at the National Center for Biotechnology Information near Washington DC, lined them up against each other. Any genes that the two bugs had in common are genes that have weathered the genetic jumbling of the 1·5 billion years of evolution since their last common ancestor. Such genes are likely to be vital for survival since natural selection has preserved them unchanged for so long. Koonin, aping nature, kept the shared 240 genes and threw out the rest.

Next, where more than one gene could do the same job, Koonin disposed of the surplus. He also threw out the genes that allow the two bacteria to cling to human cells, figuring that they are only essential for infecting a human body, not life.



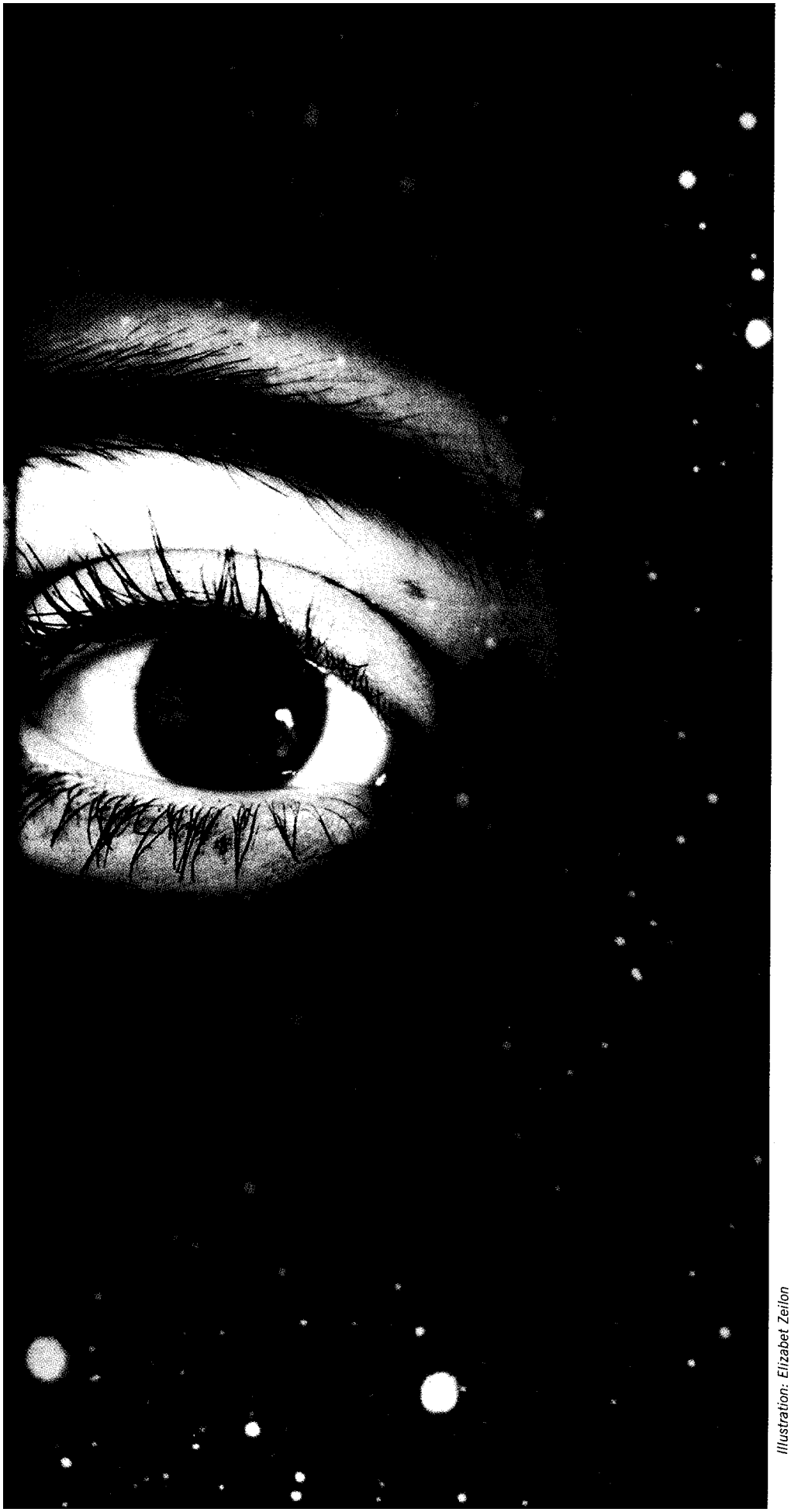


Illustration: Elizabet Zeilon

Finally, it turned out that 22 genes that look completely different in the two bug species actually perform the same essential tasks. Although these genes had been eliminated in the initial analysis, Koonin added back the *M. genitalium* versions of the genes. That gave him 256 genes, the final minimum set that Koonin says is "necessary and sufficient" for the life of a modern cell. Gone were all the genes for consuming food of all but the simplest sort, gone were the genes for manufacturing anything that could be scavenged from the environment. The theoretical bug could barely repair any damage that its DNA sustained, and it had lost the ability to fine-tune the activity of each of its remaining genes. "The take-home message is how little metabolism there is [in a minimum organism]," says Koonin.

## Out of action

Koonin's theoretical bug is a starting point, but it is still just a list of genes. Nobody knows for sure whether, if it were a real bug, it could survive. The next step is to chop surplus genes out of a real organism. Hutchison is putting random *M. genitalium* genes out of action by inserting bits of DNA called transposons into them. Already he has destroyed 70 genes that were not essential. He calculates that the job is only half done, so another 70 should follow. "We're thinking the minimum gene set may be a little bigger than [Koonin] predicted," says Hutchison. "But so far things are in remarkable agreement."

Hutchison is creating lots of strains of the bacterium, each with one gene missing. His next job is trickier. He has to make one strain with all of the genes missing.

One way to target and destroy specific genes is a technique called homologous recombination. Homologous recombination is temperamental, with a ballpark success rate of one in a million. It works best within the single-celled yeast, *Saccharomyces cerevisiae*. So Hutchison will remove the DNA from *M. genitalium* and insert it into yeast, and then employ a reusable marker called URA3 to label the genes that have been successfully destroyed.

Others have used homologous recombination to destroy genes (albeit not so many at once), and put largish pieces of DNA into yeast—so that part of the project has a good chance of working. The hard part will be transferring the pockmarked DNA back into an *M. genitalium*

cell. But the concept "is certainly not outlandish", says Patrick Brown, a genome researcher at Stanford University in Palo Alto, California. If the technique works, Hutchison will have created the simplest living cell in the world.

But Hutchison acknowledges that those who scorn the concept of creating the simplest cell in the world have a point. That's partly because of genetic redundancy—more than one gene can do a given job. This doesn't means that a very simple organism cannot be constructed, rather it means there is no one, ultimate microbug. For instance, if you remove gene 1, *M. genitalium* may be kept alive by gene 2; but if you started by removing gene 2, the bug may survive using gene 1. "As long as you realise that [the concept] has certain flaws," says Carl Woese, an evolutionary biologist at the University of Illinois in Urbana-Champaign, "it will be useful."

## The ultimate Tamagotchi

That doesn't mean that minimum bugs are going to be easy to understand. A cell is a bag of seemingly ever-changing chemical reactions and, even for a minimum cell, the only thing capable of keeping track of them all will be a computer.

With that end in mind, Hutchison has teamed up with computer scientist Masaru Tomita of Keio University in Fujisawa, Japan, who by 2000 plans to have created a "virtual" bacterium—an electronic entity that Tomita compares to a Tamagotchi, the Japanese virtual pet.

But Tomita's toy will be a little more complex. For a start it will require a supercomputer to run it. At the International Conference on Intelligent Systems for Molecular Biology, held in Greece in June, Tomita described a prototype *M. genitalium* called E-CELL. At the moment, E-CELL is pretty good at metabolism—for example, working out how to convert a sugar molecule into energy, cell building blocks, and waste products. It also handles the conversion of genes into proteins quite well. But it is less reliable when it comes to transporting chemicals across membranes, replicating its DNA, or dividing to form two daughter cells.

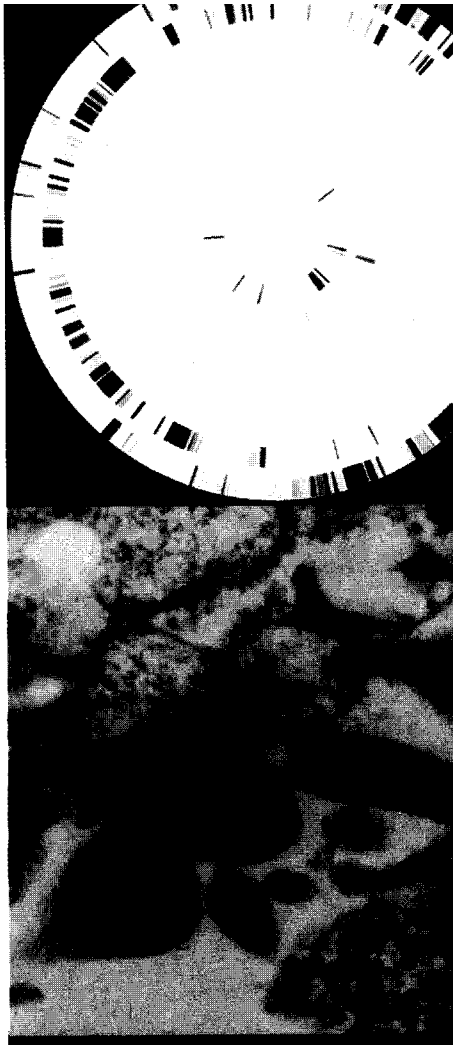
Tomita wants his final virtual *M. genitalium* to be an approximation of all the biochemical pathways that make up the living cell. He predicts it will comprise perhaps 2000 interacting virtual genes, proteins, chemicals, and so on, governed by several thousand rules that control all of the myriad possible interactions between these components.

Meanwhile, Selkov at the Russian Academy and Ross Overbeek at Argonne National Laboratories in Illinois have devised a computer program that constructs rough models of networks of biochemical pathways directly from the genome sequence of an organism. If the program comes across a gene whose function is unknown, it matches it up with a gene with a similar sequence whose function is already known and slots its protein product in a pathway accordingly. So far, Selkov and Overbeek have simulations of the main energy-producing circuits of several organisms, including *M. genitalium*. But their virtual circuits can act up, perhaps failing to respond to an environmental change. This is a sign of a missing link in the virtual chain of biochemical reactions. As more connections are added, the model should gradually come to resemble a real cell.

**'Those who scorn the concept of creating the simplest bug in the world have a point: a very simple organism can be constructed. But there is no ultimate microbug'**

Both the Selkov and Tomita teams plan to use their virtual bacteria to work out what happens when different genes are removed, generating other versions of the minimum organism.

But few are convinced that such ambitious models can be made to work. Harley McAdams, an independent researcher in Palo Alto, California, is one such sceptic. He has created a computer model of the rise and fall of viral protein production in a bacterium when a virus called lambda infects it. This process involves few genes, and yet, each simulated infection took considerable supercomputer muscle. "One could say pretty quickly that one couldn't model a [whole] bacterium with that approach," says McAdams.

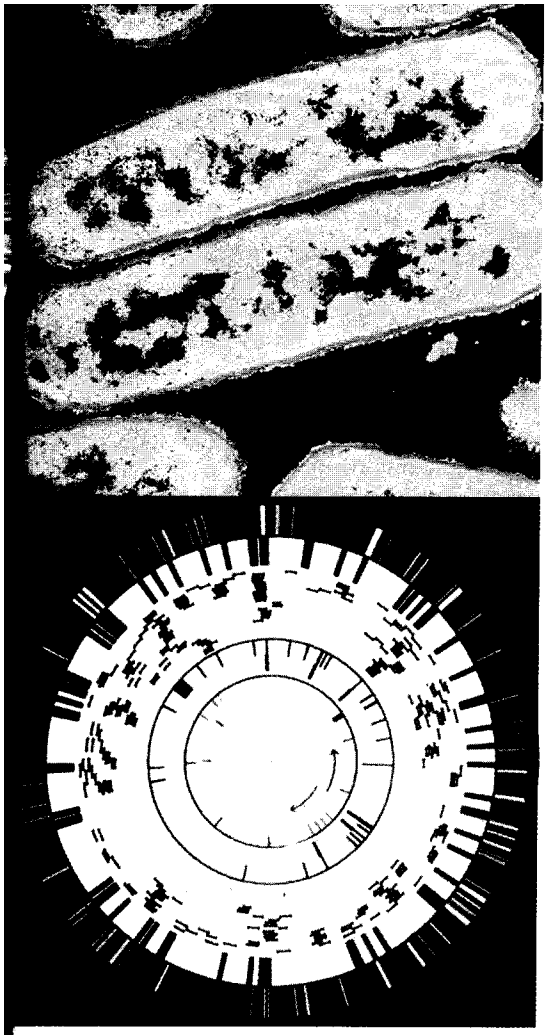


Bare necessities: genes shared by *M. genitalium* (above) and *H. influenzae* (right) are probably essential for life. The circles represent the bacterial chromosomes. Coloured stripes represent genes, and other genetic landmarks

A bacterium is more than a group of linear chemical reactions. In the cell, chemical pathways interconnect, layer upon layer, forming criss-crossing complex "circuits". "The cell has layers and layers of feedback that keep things stable," says McAdams. "If you take account of all these effects, you end up with such a complicated mess that it's hard to model."

## Thick skins

But the way the cell creates that stability could come to the modellers' rescue. Princeton biophysicists Stanislaws Leibler and Naama Barkai reported in the 26 June issue of *Nature* that they had constructed a computer model of the biochemical pathway that allows bacteria to sense food. When Leibler changed the characteristics of the virtual proteins to change



CNRI/Science Photo Library

the speed of the chemical reactions in the food-sensing pathway, the bacteria often remained surprisingly good at seeking out food. Leibler concluded that, rather than fine-tune every protein in every pathway, evolution has designed the pathways so that they are insensitive to certain changes—the changes that might occur naturally as genes mutate and code for slightly different proteins.

Tomita takes this as a sign that creating a virtual bacterium is a manageable project—modellers can get away with making a few errors. "One can easily make any problem unsolvable," adds Selkov, by demanding an unreasonable level of perfection.

A cell's ability to control its internal environment also interests "metabolic engineers", who see minimum organisms as a new type of chemical factory. Regular-sized bacteria like *Escherichia coli* have been used for decades to make edible molecules, such as monosodium glutamate, and they are now being genetically engineered to make industrial chemicals like 1,3-propanediol, used in polyesters.

The trick is to get the bacterium to make the right chemical, and stop it from wasting energy and contaminating the end-product by making too many others. One solution is to add enzymes to drive the reaction you want. But the cell, with its vigilant environmental surveillance, usually notices and retaliates by shutting down the overactive pathway. It can also siphon off chemicals for other uses, or modify chemicals in unexpected ways. A minimum bacterium could be the basis for a souped-up bacterial factory in which production lines are ultra-efficient, and wasteful and modifying pathways absent.

Cells are designed to cope with an ever-changing natural environment. Industrial fermenters are, by contrast, eminently predictable. Consequently, the genes that deal with changing temperatures and irregular food supplies could be stripped out, and genes for enzymes that make the chemical of interest added. "The rest of the cell [would be] there for the purpose of supporting those enzymes," says Gregory Stephanopoulos, a chemical engineer from the Massachusetts Institute of Technology in Cambridge.

**'A minimum bacterium could be the basis for a souped-up bacterial factory in which production lines are ultra-efficient, and wasteful pathways absent'**

Metabolic cost-cutting can go too far, however. If having fewer genes were always better, all chemical companies might opt to use tiny bacteria like *M. genitalium*. But *M. genitalium* needs everything but the kitchen sink in order to grow, whereas bugs used in industry live off simple mixtures of salts and a sugar. This is another argument against a single "minimum organism". What constitutes the smallest gene set also depends on what you're prepared to do for the bug. Witness *M. genitalium*—a bacterium with a less complicated diet is going to need a slew of extra genes to be able to manufacture the missing nutrients.

## Downsizing

Karl Sanford, a microbiologist at Genencor International in Palo Alto, California, is kicking off his genome-reduction campaign with *E. coli* and another industrial microbe, *Bacillus subtilus*. Sanford is not giving much away, except that he has started by deleting the genes for the cellular pumps and channels that let in nutrients that are not present in your average industrial reactor.

Meanwhile, Woese is identifying the genes that are common to the growin number of genomes from all thre kingdoms—the Bacteria, other single celled organisms called the Archaea and the Eukaryota—to create yet anothe minimum genome, in this case an ancien one. The project is allowing him to mak educated guesses about how the ancestor common to all three kingdoms lived. Con troversially, he has concluded that RNA molecules were not responsible for th emergence of primordial life ("Let ther be life", *New Scientist*, 6 July 1996, p 22).

Now he has shown that when i comes to DNA duplication, only th Archaea and the Eukarya show similari ties—bacteria go about it in their own idiosyncratic way. This suggests, says Woese, that the different strategies arose after the kingdoms split apart, perhaps because the ancestor organism did not have the genome structure that is now common to all three kingdoms. Rather than long stringy chromosomes, the first sort of genetic material might have been small fragments of DNA.

The final, minimum gene set of Woese's ancestor organism will be tentative, and he believes the same is true for any minimum organism that could be created. "It's not something that you go to the top of the mountain and proclaim as the word of God," he says. But the minimum genome researchers are confident that any small genome will be better than none. "You can't avoid it," says Overbeek. "We're finally going to get a sense of how the cell works." □

***William Wells*** is a freelance science writer based in San Francisco

***Further reading:*** "A minimum gene set for cellular life derived by comparison of complete bacterial genomes" by Arcady R. Mushegian and Eugene V. Koonin, *Proceedings of the National Academy of Science*, USA, vol 93, p 10268 (1996)