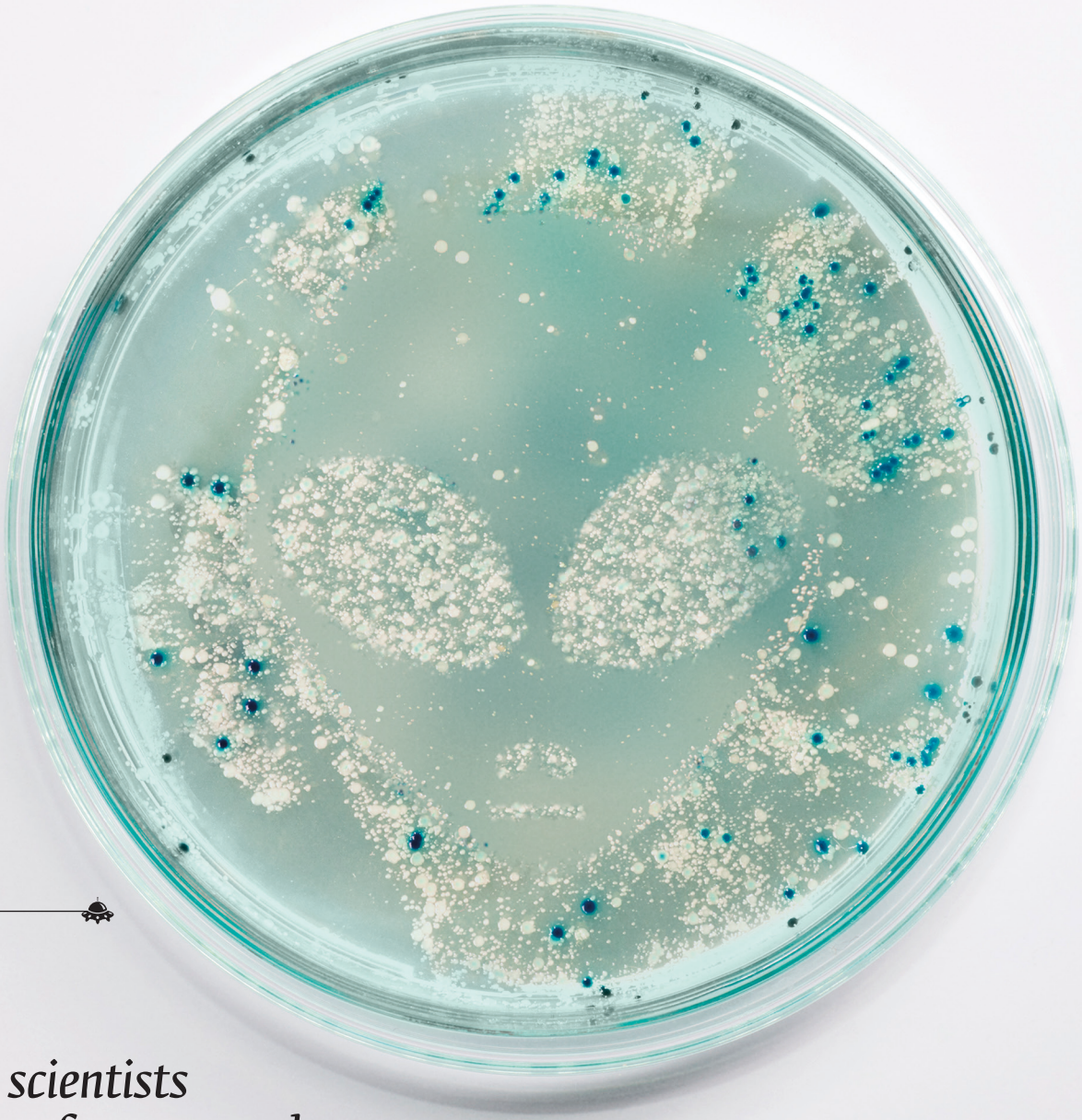
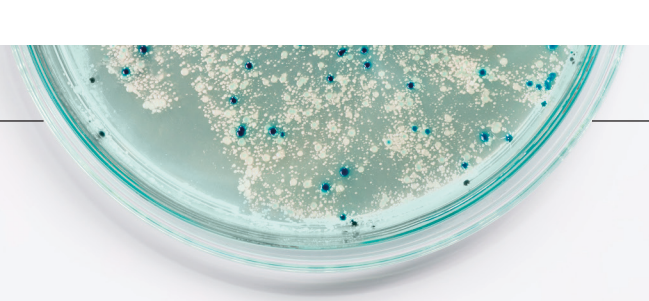


Some scientists went as far as to explore whether biological material could be a means of communication for extraterrestrial civilisations



SYNTHETIC DNA

FROM EXTRATERRESTRIAL INTELLIGENCE TO FUTURE BIOMOLECULES

Following the creation of artificial *E. coli*, we look at the history of those pushing the boundaries in the field of synthetic biology.

Researchers in the UK have reached another milestone in the field of synthetic biology. A team from the Medical Research Council laboratory of molecular biology has fashioned a synthetic version of the *E. coli* bacteria, rewriting its genetic code to create an artificial living organism.

The synthetic genome of Syn61, as this organism is referred to, is four times larger than the ones previously designed by scientists. In 2010, researchers had already reached a turning point with the design of the first complete synthetic genome. One million base pairs long, it was a remarkable feat, but it has now been surpassed by this new four million base pairs long *E. coli* genome.

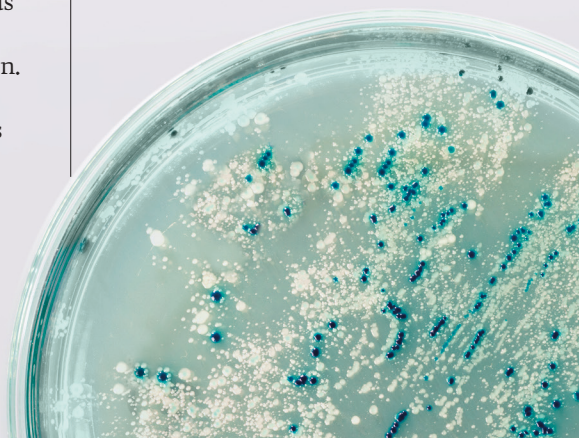
Syn61 does grow more slowly than natural *E. coli*, but its cells are functional and the production of proteins happens normally. "The exciting thing is that the genetic code of this organism has been rewritten and the cells are still viable," says Julius Fredens, postdoctoral researcher in the Medical Research Council laboratory of molecular biology and one the scientists on the team.

New amino acids

Yet, the point of this research, now published in the journal *Nature*, was not just to create a large synthetic organism. Rather, the team wanted to understand the redundancies in the genetic code which have long puzzled scientists. Their objective in the long run is to create new amino acids that do not exist in nature.

The basic idea behind this project revolves around the fact that the genetic code of all living things consists of 64 codons. These sequences of three nucleotides encode the 20 amino acids – the building blocks of the proteins needed by every organisms to function. This means that some codons are redundant in some way; synonymous codons encode the same amino acids. By embarking on this project, one of the questions that the scientists wanted to answer was whether all synonymous codons are necessary. They wanted to investigate whether it is possible to reduce the number of codons used to encode amino acids, by entirely replacing three specific codons in the genome with their synonymous codons.

To do this, it was easier to re-design a synthetic organism from scratch and to rewrite its genetic code than to take a natural *E. coli* bacteria and replace the millions of target codons. This resulted in the creation of Syn61, a live organism with a genetic code uniquely made up of 61 codons. For the team, the ultimate objective is to use these genetically unique organisms as a basis to create new amino acids that do not exist in nature. "All life forms as we know them are using 20



amino acids to build all the proteins and enzymes we know of. Our lab has the goal of including new unnatural amino acids into biological systems. Then we can include new chemical groups, like a chemical handle on proteins and enzymes, that will give us a new way of controlling biological processes and developing new ones," Julius Fredens explains.

The rise of synthetic biology

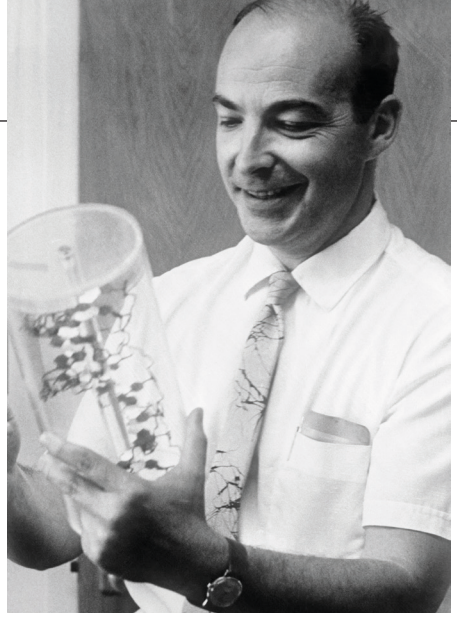
Designing artificial organisms made of synthetic DNA and recoding their genomes is now a reality, which can be done far more cheaply than just a few years ago. However, the path to get there has not been straightforward, and it was not until the "genomics revolution", at the close of the 20th century, that the field of synthetic biology started to grow.

The origins of synthetic DNA can be traced back to the 1950s, and the pioneering work of American biochemist Arthur Kornberg. With his team, he identified the mechanisms of the biological synthesis of DNA and RNA – a discovery which earned him the Nobel Prize in medicine in 1959.

Kornberg had been interested in studying how nucleotides are strung together to form DNA molecules. In 1956, he discovered that an enzyme-catalyzed polymerisation reaction was at the root of this process.

He and his colleagues were then able to isolate and purify the enzyme that could produce precise replicas of short DNA molecules in test tubes, which we now refer to as the DNA polymerase enzyme. He then went on to work with the phi X 174 bacteriophage, a single-stranded DNA virus that infects *E. coli*, using it as a model to prove that DNA synthesised *in vitro* using purified polymerase enzymes could produce all the features of a natural virus.

This was at a time when the field of synthetic biology as whole, which combines engineering and biology to redesign biological systems or create new ones (including enzymes, genetic circuits, and cells) was emerging. In 1961,



The origins of synthetic DNA can be traced back to the 1950s, and to the pioneering work of American biochemist Arthur Kornberg

researchers Francois Jacob and Jacques Monod had published a study about *E. coli*, in which they hypothesised the existence of regulatory circuits that underpin the response of a cell to its environment. Their work allowed scientists to start exploring the possibility of assembling new regulatory systems from molecular components. In the following decades, the development of molecular cloning techniques led to genetic manipulations becoming more common in labs around the world. Scientists started to gain new understandings of the genetic code.

It is in this productive environment that they achieved the first synthesis of a complete gene – a yeast tRNA – in 1972. This was followed by the synthesis of the first peptide- and protein-coding genes. "Scientists first started by synthetically making short oligonucleotides to crack the genetic code. Then, the first synthetic gene was for a very small transfer RNA, the same kind of molecule that's being deleted in the current recoding studies," says George Church, Professor of Genetics at Harvard Medical School.

In parallel, other majors advances linked to DNA sequencing were taking place, leading the scientific community to get a better understanding of the "code of life" as well as valuable insights to improve DNA synthesis. In 1977, a research team sequenced the genome of the phi X 174 bacteriophage. It was the first DNA-based genome to be sequenced. This kind of work generated a lot of excitement, leading to the publications of hundreds of studies, ranging from the most serious academic work on recombinant gene expression to the most bizarre hypotheses. In a 1979 paper in the scientific journal *Icarus* (current impact factor – 2.981), which is still published today by Elsevier, some scientists went as far as to explore whether biological material could be a means of communication for extraterrestrial civilisations. They studied the DNA of phi X174, to determine whether it might "carry a message from an advanced society".

The abstract states: "It is speculated that advanced civilizations could manipulate viral or bacterial DNA so that its base sequence would carry a coded message in addition to specifying compounds necessary for survival and send a microorganism containing the message to a planet with conditions similar to those of the sending planet, where the microorganism would replicate."

At this point, however, the scientific community was still a long way away from using this newly acquired knowledge, and the techniques of DNA synthesis, to create complex artificial organisms.

The revolution came in the mid-1990s, with the development of new tools, such as automated DNA sequencing, which meant that complete microbial genomes could be sequenced, starting with the genomes of *Helicobacter pylori* in 1994 and *Haemophilus influenzae*, in 1995.

As researchers gained more insight into molecular components that make up biological systems, new milestones were reached. In 2000, the first reports of genetic circuits being engineered to carry out specific functions started to emerge.

In the years that followed, more improvements in parts and genetic circuit design in *E. coli* continued. But it was not until 2010 that a real step forward was made. A team at the J. Craig Venter Institute in the US reported the design, synthesis and assembly of the first synthetic genome, the 1.08–mega–base pair *Mycoplasma mycoides* bacteria. This was followed by the first recoded genome in 2013, which paved the way for the creation of Syn61 a few years later.

Looking to the future

This most recent achievement was also made possible by the dramatic reduction in the cost of producing synthetic DNA. "This is one of those areas where the price is decreasing exponentially. If you looked at the cost 10 years ago, it would have been impossible," Finn Stirling, researcher at Harvard University's molecular and cellular biology department, points out. The next logical steps will be for the creation of much bigger synthetic organisms. More recoding studies, with a focus on rewriting their genetic codes will also take place.

All this research holds great potential for fundamental genetic research. Recoding the genome of synthetic organisms would indeed help researchers to learn about cell response to massive codon replacements, and could allow them to discover new properties associated with global transcription mechanisms.

At the Medical Research Council laboratory of molecular biology, the team will continue working to create new amino acids that do not exist in nature, with potential medical implications.

"The unnatural amino acid aspect will allow us to make new enzymes with new functions. One of the ideas that has been discussed for many years now is to use these amino acids to link an antibody with a chemotherapy drug. You can make an antibody that is very specific and very controlled, and you can link a very specific drug to a specific cell. That's probably the best medical application," Julius Fredens explains.

Other areas of medicine and environmental sciences also stand to benefit from synthetic genome design and recoding. For example, synthetic biologists have long worked to design synthetic bacteria with the purpose of decontaminating pesticide-contaminated fields, or to non-invasively diagnose the presence of chemicals in the gut. However, they are concerned that such organisms could exchange DNA with other un-engineered bacteria, which could have unpredictable and potentially dangerous consequences.

In the long term, working these synthetic bacteria and recoding their genomes could be a solution, as it is a way of genetically isolating them from everything that exists in nature. "What it effectively does is make a genetic firewall between that species and another species. Bacteria evolve by picking up DNA in their environment and from other bacteria. If you have recoded an organism, it cannot express those genes which have been picked up. In medical sciences, it enables you to engineer a bacteria that you can be much more confident it can be released safely into the environment because it is isolated from another species," Finn Stirling says.

Another example of a potentially very useful application is the design of bacteria strains that can degrade plastic and help deal with the global plastic waste epidemics in the safest possible way.

"Obviously you don't want other bacterial species to gain this capacity because plastic has a lot of applications that rely upon its durability. If you could isolate it from other strains and make sure it stays alive just within the confines of a landfill, that would be a helpful technology," Stirling adds.

As the technical capacities continue to expand, scientists will continue to find numerous applications for synthetic biology. "In contrast to most other engineering fields, synthetic biology allows us to make trillions of precisely designed prototypes and then select the best ones. Now that we have powerful tools, we can focus on functional goals – like enhanced metabolic and ecosystem engineering, resistance to all viruses, gene and cell therapy," George Church concludes. BMS