Published version: *Nature* **615**, 592-594 (2023)

doi: https://doi.org/10.1038/d41586-023-00702-0

An antiviral molecular language barrier

*'*Synthetic bacterial genome upgraded for viral defence and biocontainment'

Benjamin A. Blount

Biodiscovery Institute, School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK.

e-mail: benjamin.blount@nottingham.ac.uk

Bacteria with a synthetic genome were engineered to alter the way that the DNA code instructs cells to make proteins. This 'language barrier' serves to isolate the cells genetically, and makes them immune to viral infection.

As scientists gain an increasing ability to build synthetic genomes to specific design criteria, this is enabling the production of cells that have beneficial properties not found in the natural world. Writing in Nature, Nyerges *et al*.[1] report that they have engineered bacteria to be immune to viral infections and to acquire properties that could be useful for microbial biocontainment. The work has implications for the development of safe and efficient applications in biotechnology, and is a compelling demonstration of the possibilities that are opened up by the use of synthetic genomes.

Unlike genome-engineering techniques in which an existing genome in living cells is modified, synthetic genomes can be designed and built from scratch. This means that the scale of changes to a genome, and consequently to a cell's behaviour, is no longer limited by the ability to edit existing DNA.

A functioning synthetic genome has previously been built[2] for the model bacterium *Escherichia coli*. A feature of this genome is a change to how the bacterium translates the information encoded by its DNA into proteins.

DNA determines the order and content of amino acids in a protein, and sequences of three DNA bases, termed codons, encode a given amino acid. This code is almost universally evolutionarily conserved across biology. To produce a protein, DNA-sequence information is copied into a corresponding messenger RNA. A transfer RNA (tRNA) recognizes each codon and then 'translates' it into a specific amino acid to be incorporated into a growing protein chain. There is redundancy in this code, because most amino acids can be encoded by several codons. This makes it possible to swap equivalent codons, changing the DNA sequence while maintaining the amino-acid content.

By swapping codons on a genomic scale, the synthetic *E. coli* genome was designed to exclude all instances of three particular codons. Two of these normally encode the amino acid serine and the other instructs protein assembly to stop. The tRNAs responsible for translating these excluded codons can be removed from the cells or repurposed for a new function. This process of recoding a

genome to change the way a cell interprets the DNA code has many implications, not least for how cells interact with viruses.

Viruses typically lack the resources for producing the proteins they need to make copies of themselves and to attack the host cell. Instead, they hijack the host's resources, including its pool of tRNAs. In this case, it would follow that, in synthetic cells in which some tRNAs have been removed from the genome, the viral genetic information wouldn't be translated properly, rendering the virus incompatible with the host. Previous work has indeed shown that cells that have synthetic genomes and have some tRNAs removed can evade infection when challenged with certain viruses[3].

Nyerges and colleagues identified viruses from environmental and wastewater samples that carry their own copies of tRNAs, and showed that these viruses did not rely on host tRNAs and could infect the recoded cells. To protect the engineered cells from this class of virus, Nyerges *et al*. developed tRNAs that changed the link between the DNA code and the protein content. These tRNAs recognize two of the codons removed from the synthetic genome. However, rather than translating them into the hydrophilic amino acid serine during protein assembly, as would normally occur for the universal genetic code, these tRNAs directed the incorporation of the hydrophobic amino acid leucine (Fig. 1a).

This change has consequences if any DNA sequences containing the codons corresponding to those recognized by these tRNAs are introduced into these cells. During protein synthesis, the resulting insertion of leucine, which has different chemical properties from those of serine, would probably alter a protein's structure and properties, inactivating it. These bacteria therefore speak a different 'language' from the rest of nature, including viruses.

Parallel work[4] by the team behind the *E. coli* synthetic genome has shown that this approach is sufficient to prevent infection by selected viruses that have their own tRNAs. Surprisingly, however, Nyerges *et al.* found that the viruses they had isolated could nevertheless infect and kill cells with the modified tRNAs. Further analysis of cells undergoing infection revealed that virally encoded tRNAs were produced at high levels, rapidly outnumbering the host tRNAs and leading to most of the viral proteins being properly assembled and functional.

When faced with this superiority of the viral tRNAs, the team decided to turn that strength back on the invading viruses by co-opting viral tRNAs and generating new versions of them to force a coding switch from serine to leucine. With these virus-derived alternative tRNAs, the engineered bacterial cells could drive the incorporation of the 'incorrect' amino acids into viral proteins. The problematic viruses, along with all other viruses tested, could not overcome this molecular language barrier and couldn't infect the engineered cells.

Crucially, the authors considered that bacteria with an engineered synthetic genome that grants such an unprecedented capacity to evade viral infection could have a competitive advantage over natural bacterial populations. The microorganisms might thus pose a serious challenge if accidentally released outside a controlled environment.

The authors therefore set about enhancing the biosecurity of the engineered strains by further exploiting the codon changes made to the synthetic genome. Building on some of their previous work[5], Nyerges and colleagues made the strains 'addicted' to a synthetic amino acid called L-4,4'-biphenylalanine (BipA), which doesn't occur naturally. Cells were modified to recognize a repurposed codon as an instruction to incorporate BipA into a protein that the bacterial cells need to survive, thus limiting the microbes' growth to environments in which this non-natural molecule is supplied to them (Fig. 1b).

Finally, the authors demonstrated that the recoding strategies can be used to prevent the spread of synthetic DNA on sequences called mobile genetic elements. Microbes naturally exchange genetic information between cells in various ways, including by the transfer of circular DNA molecules called plasmids — elements that researchers use to transfer DNA into cells. The authors created a set of plasmids that use the modified codon language of the engineered cells to encode components required for plasmid replication in bacteria. These plasmids can function only in cells with a synthetic genome and engineered tRNAs, a scenario that strikingly reduces the risk of engineered DNA being unintentionally transferred to wild bacterial populations (Fig. 1c).

It currently takes a huge effort to establish a working synthetic genome, with only a handful completed so far. Our capabilities on this front are slowly scaling up, with a full synthetic genome for a eukaryotic cell (one that contains a nucleus) expected to be finalized within the next few years[6], and work towards a human synthetic genome project also under way[7]. As the number, size and ambition of synthetic-genome projects increases, so, too, will our ability to study and manipulate biology. The impressive feats achieved through codon repurposing in this work will be immensely valuable to bacterial biotechnology, in which viral contamination is a persistent and expensive problem.

The biggest impact of this work will probably be in providing a foundation for similar strategies in synthetic genomes for other organisms. Increasingly, key medical products, such as vaccines and protein therapeutics, depend on the use of mammalian or human cell-culture systems that are vulnerable to viral infection, with substantial implications for cost and product safety[8]. Controlled, reliable manufacturing processes that are protected from problems of viral infection will be crucial for maximizing these industries' positive impact on health and well-being, while ensuring that the processes are safe, contained and retain public confidence.

1. Nyerges, A. *et al*. Nature https://doi.org/10.1038/s41586-023-05824-z (2023).

2. Fredens, J. *et al*. Nature 569, 514–518 (2019).

3. Robertson, W. E. *et al*. Science 372, 1057–1062 (2021).

4. Zürcher, J. F. *et al*. Science 378, 516–523 (2022).

5. Mandell, D. J. *et al*. Nature 518, 55–60 (2015).

6. Richardson, S. M. *et al*. Science 355, 1040–1044 (2017).

7. Boeke, J. D. *et al*. Science 353, 126–127 (2016).

8. Barone. P. W. *et al*. Nature Biotechnol. 38, 563–572 (2020).

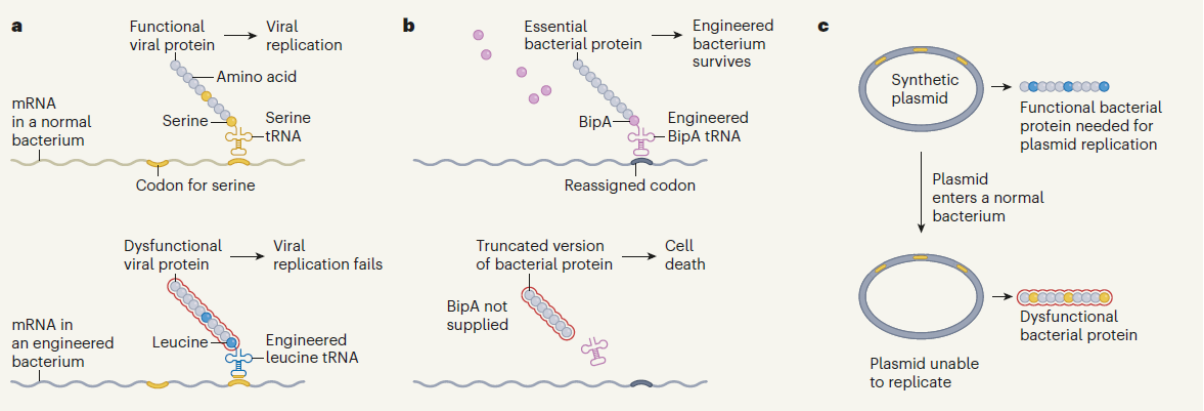

Figure Legend: Engineering bacteria to evade viral infection and to need human intervention for survival. (a) Nyerges *et al.*[1] engineered a strain of the bacterium *Escherichia coli* containing a synthetic genome to change some aspects of how the microorganism's genetic code is recognized. Nucleotide sequences in messenger RNA (mRNA), called codons, are recognized by transfer RNA (tRNA) molecules that carry a particular amino acid, such as serine. Thus, the amino acid encoded by the mRNA is added during protein synthesis. When a virus invades a bacterium, it can hijack bacterial machinery to produce the viral proteins needed for replication. The authors used bacteria that lack certain codons for serine, and engineered the bacteria to have tRNAs that recognize the serine codons but carry the amino acid leucine. Viruses can't replicate in these cells because the abnormal leucine content disrupts viral proteins. (b) The authors reassigned a codon in the engineered bacteria to be recognized by a tRNA that carries the artificial amino acid BipA. The mRNA for an essential bacterial protein contains this reassigned codon, and these bacteria need BipA to survive, offering a potential strategy for biocontainment. (c) Nyerges *et al.* generated circular DNA molecules called plasmids that rely on the modified system for codon recognition used in the engineered bacteria. These plasmids could not replicate in *E. coli* that have a standard system for codon recognition, reducing the risk of synthetic DNA spreading in wild populations.