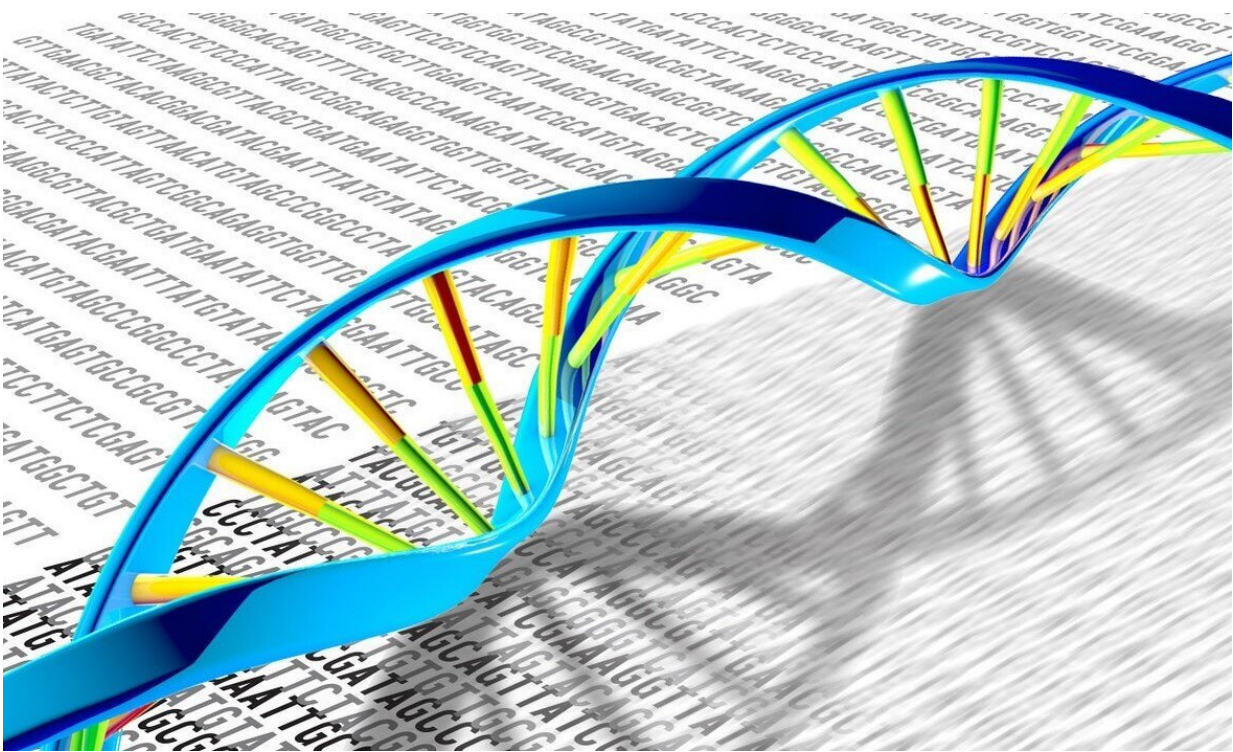


Synthetic DNA sheds light on mysterious difference between living cells at different points in evolution

March 6 2024



DNA, which has a double-helix structure, can have many genetic mutations and variations. Credit: NIH

"Random DNA" is naturally active in the one-celled fungi yeast, while such DNA is turned off as its natural state in mammalian cells, despite

their having a common ancestor a billion years ago and the same basic molecular machinery, a new study finds.

The new finding revolves around the process by which DNA genetic instructions are converted first into a related material called RNA and then into proteins that make up the body's structures and signals. In [yeast](#), mice, and humans, the first step in a gene's expression, [transcription](#), proceeds as DNA molecular "letters" (nucleobases) are read in one direction. While 80% of the human genome—the complete set of DNA in our cells—is actively decoded into RNA, less than 2% actually codes for [genes](#) that direct the building of proteins.

A longstanding mystery in genomics then is what is all this non-gene-related transcription accomplishing. Is it just noise, a side effect of evolution, or does it have functions?

A research team at NYU Langone Health sought to answer the question by creating a large, synthetic gene, with its DNA code in reverse order from its natural parent. Then they put synthetic gene into yeast and mouse stem cells and watched transcription levels in each.

Published in the journal *Nature*, the [new study](#) reveals that in yeast the genetic system is set so that nearly all genes are continually transcribed, while the same "default state" in the [mammalian cells](#) is that transcription is turned off.

Interestingly, say the study authors, the reverse order of the code meant that all of the mechanisms that evolved in yeast and mammalian cells to turn transcription on or off were absent because the reversed code was nonsense. Like a [mirror image](#), however, the reversed code reflected some basic patterns seen in the natural code in terms of how often DNA letters were present, what they fell near, and how often they were repeated.

With the reversed code being 100,000 molecular letters long, the team found that it randomly included many small stretches of previously unknown code that likely started transcription much more often yeast, and stopped it in mammalian cells.

"Understanding default transcription differences across species will help us to better understand what parts of the genetic code have functions, and which are accidents of evolution," said corresponding author Jef Boeke, Ph.D., the Sol and Judith Bergstein Director of the Institute for Systems Genetics at NYU Langone Health.

"This in turn promises to guide the engineering of yeast to make new medicines, or create new gene therapies, or even to help us find new genes buried in the vast code."

The work lends weight to the theory that yeast's very active transcriptional state is set so that foreign DNA, rarely injected into yeast for instance by a virus as it copies itself, is likely to get transcribed into RNA. If that RNA builds a protein with a helpful function, the code will be preserved by evolution as a new gene.

Unlike a single-celled organism in yeast, which can afford risky new genes that drive faster evolution, mammalian cells, as part of bodies with millions of cooperating cells, are less free to incorporate new DNA every time a cell encounters a virus. Many regulatory mechanisms protect the delicately balanced code as it is.

Big DNA

The new study had to account for the size of DNA chains, with 3 billion "letters" included in the [human genome](#), and some genes being 2 million letters long. While famous techniques enable changes to be made letter by letter, some engineering tasks are more efficient if researchers build

DNA from scratch, with far-flung changes made in large swaths of pre-assembled code swapped into a cell in place of its natural counterpart.

Because human genes are so complex, Boeke's lab first developed its "genome writing" approach in yeast, but then recently adapted it to the mammalian genetic code. The study authors use yeast cells to assemble long DNA sequences in a single step, and then deliver them into mouse embryonic stem cells.

For the current study, the research team addressed the question on how pervasive transcription is across evolution by introducing a synthetic 101 kilobase stretch of engineered DNA—the human gene hypoxanthine phosphoribosyl transferase 1 (HPRT1) in reverse coding order. They observed widespread activity of the gene in yeast despite the lack in the nonsense code of promoters, DNA snippets that evolved to signal for the start of transcription.

Further, the team identified small sequences in the reversed code, repeated stretches of adenosine and thymine building blocks, known to be recognized by transcription factors, proteins that bind to DNA to initiate transcription. Just five to 15 letters long, such sequences could easily occur randomly and may partly explain the very active yeast default state, the authors said.

To the contrary, the same reversed code, inserted into the genome of mouse embryonic stem cells, did not cause widespread transcription. In this scenario, transcription was repressed even though evolved CpG dinucleotides, known to actively shut down (silence) genes, were not functional in the reversed code.

The team surmises that other basic elements in the mammalian genome may restrict transcription much more so than in yeast, and perhaps by directly recruiting a protein group (the polycomb complex) known to

silence genes.

"The closer we get to introducing a 'genome's worth' of nonsense DNA into living cells, the better they can compare it to the actual, evolved genome," said first author Brendan Camellato, a graduate student in Boeke's lab.

"This could lead us to a new frontier of engineered cell therapies, as the capacity to put in ever longer synthetic DNAs enables better understanding of what insertions genomes will tolerate, and perhaps the inclusion of one or more larger, complete, engineered genes."

Along with Boeke and Camellato, NYU Langone study authors were Ran Brosh, Hannah Ashe, and Matthew Maurano.

More information: Jef Boeke, Synthetic reversed sequences reveal default genomic states, *Nature* (2024). [DOI: 10.1038/s41586-024-07128-2](https://doi.org/10.1038/s41586-024-07128-2).
www.nature.com/articles/s41586-024-07128-2

Provided by NYU Langone Health

Citation: Synthetic DNA sheds light on mysterious difference between living cells at different points in evolution (2024, March 6) retrieved 27 April 2024 from <https://phys.org/news/2024-03-synthetic-dna-mysterious-difference-cells.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.