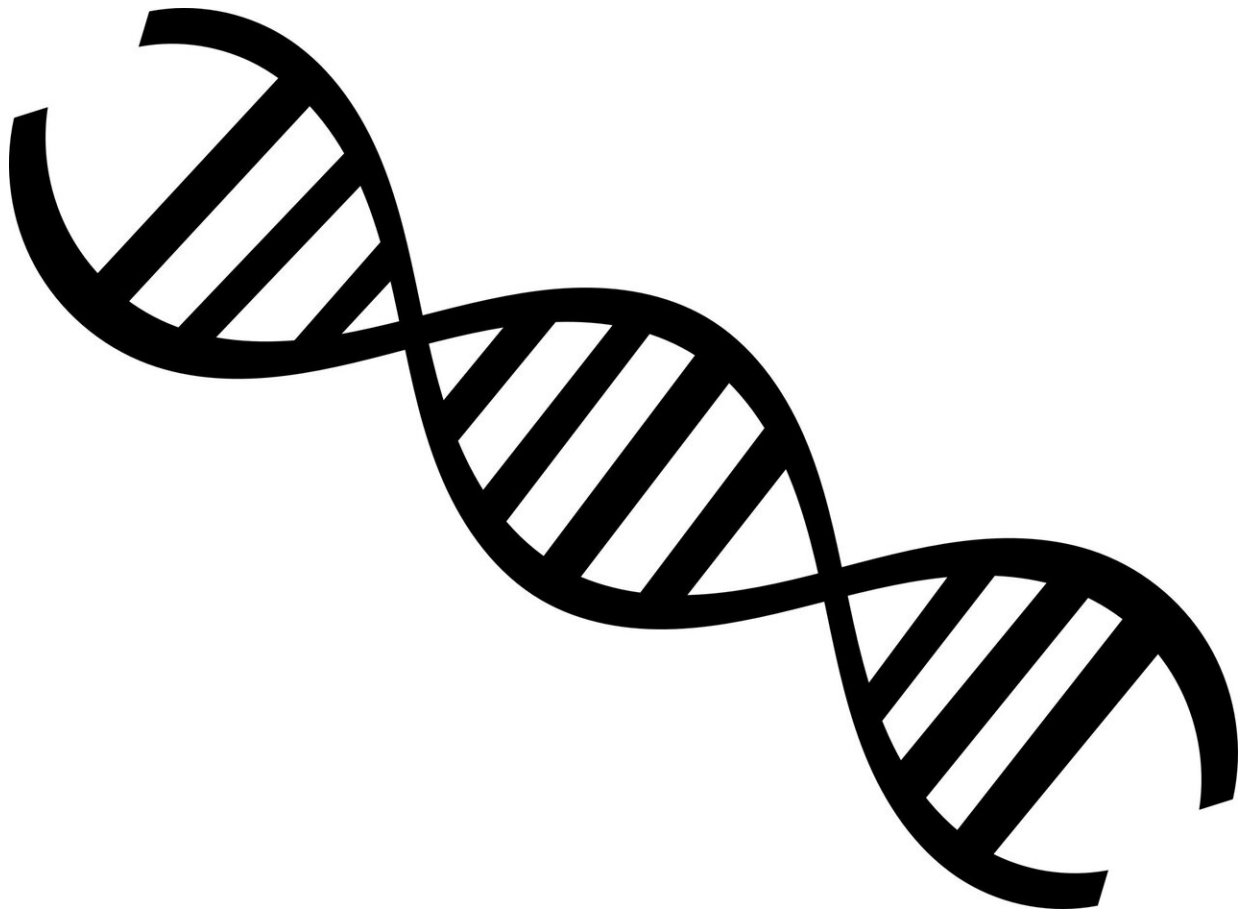


ENCODE 3 project details the inner workings of the human and mouse genome

July 29 2020



Credit: Pixabay/CC0 Public Domain

The [Encyclopedia of DNA Elements \(ENCODE\) Project](#) is a worldwide effort to understand how the human genome functions. With the

completion of its latest phase, the ENCODE Project has added millions of candidate DNA "switches" from the human and mouse genomes that appear to regulate when and where genes are turned on, and a new registry that assigns a portion of these DNA switches to useful biological categories. The project also offers new visualization tools to assist in the use of ENCODE's large datasets.

The project's latest results were published in *Nature*, accompanied by 13 additional in-depth studies published in other major journals. ENCODE is funded by the National Human Genome Research Institute, part of the National Institutes of Health.

"A major priority of ENCODE 3 was to develop means to share data from the thousands of ENCODE experiments with the broader research community to help expand our understanding of genome function," said NHGRI Director Eric Green, M.D., Ph.D. "ENCODE 3 search and visualization tools make these data accessible, thereby advancing efforts in open science."

To assess the potential functions of different DNA regions, ENCODE researchers studied biochemical processes that are typically associated with the switches that regulate [genes](#). This biochemical approach is an efficient way to explore the entire genome rapidly and comprehensively. This method helps to locate regions in the DNA that are "candidate functional elements"—DNA regions that are predicted to be functional elements based on these biochemical properties. Candidates can then be tested in further experiments to identify and characterize their functional roles in [gene regulation](#).

"A key challenge in ENCODE is that different genes and functional regions are active in different cell types," said Elise Feingold, Ph.D., scientific advisor for strategic implementation in the Division of Genome Sciences at NHGRI and a lead on ENCODE for the institute.

"This means that we need to test a large and diverse number of biological samples to work towards a catalog of candidate functional elements in the genome."

Significant progress has been made in characterizing [protein-coding genes](#), which comprise less than 2% of the human genome. Researchers know much less about the remaining 98% of the genome, including how much and which parts of it perform other functions. ENCODE is helping to fill in this significant knowledge gap.

The human body is composed of trillions of cells, with thousands of types of cells. While all these cells share a common set of DNA instructions, the diverse cell types (e.g., heart, lung and brain) carry out distinct functions by using the information encoded in DNA differently. The DNA regions that act as switches to turn genes on or off, or tune the exact levels of gene activity, help drive the formation of distinct cell types in the body and govern their functioning in health and disease.

During the recently completed third phase of ENCODE, researchers performed nearly 6,000 experiments—4,834 in humans and 1,158 in mice—to illuminate details of the genes and their potential regulators in their respective genomes.

ENCODE 3 researchers studied developing embryonic mouse tissues to understand the timeline of various genomic and biochemical changes that occur during mouse development. Mice, due to their genomic and biological similarity to humans, can help to inform our understanding of human biology and disease.

These experiments in humans and mice were carried out in several biological contexts. Researchers analyzed how chemical modifications of DNA, proteins that bind to DNA, and RNA (a sister molecule to DNA) interact to regulate genes. Results from ENCODE 3 also help explain

how variations in DNA sequences outside of protein-coding regions can influence the expression of genes, even genes located far away from a specific variant itself.

"The data generated in ENCODE 3 dramatically increase our understanding of the human [genome](#)," said Brenton Graveley, Ph.D., professor and chair of the Department of Genetics and Genome Sciences at UCONN Health. "The project has added tremendous resolution and clarity for previous data types, such as DNA-binding proteins and chromatin marks, and new data types, such as long-range DNA interactions and protein-RNA interactions."

As a new feature, ENCODE 3 researchers created a resource detailing different kinds of DNA regions and their corresponding candidate functions. A [web-based tool called SCREEN](#) allows users to visualize the data supporting these interpretations.

The ENCODE Project began in 2003 and is an extensive collaborative research effort involving groups across the U.S. and internationally, comprising over 500 scientists with diverse expertise. It has benefited from and built upon decades of research on gene regulation performed by independent researchers around the world. ENCODE researchers have created a [community resource](#), ensuring that the project's data is accessible to any researcher for their studies. These efforts in open science have resulted in over 2,000 publications from non-ENCODE researchers who used data generated by the ENCODE Project.

"This demonstrates that the encyclopedia is widely used, which is what we had always aimed for," said Dr. Feingold. "Many of these publications are related to human disease, attesting to the resource's value for relating basic biological knowledge to health research."

More information: Perspectives on ENCODE, *Nature* (2020). [DOI:](#)

[10.1038/s41586-020-2449-8](https://doi.org/10.1038/s41586-020-2449-8) ,
www.nature.com/articles/s41586-020-2449-8

Provided by NIH/National Human Genome Research Institute

Citation: ENCODE 3 project details the inner workings of the human and mouse genome (2020, July 29) retrieved 6 May 2024 from
<https://medicalxpress.com/news/2020-07-encode-human-mouse-genome.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.