

# Genetics research sheds light on 'dark' portion of genome

March 11 2020

---



Clemson Professors Trudy Mackay and Robert Anholt address human genetics questions by studying the common fruit fly, *Drosophila melanogaster*, because many genes are conserved between humans and fruit flies, meaning research results can be extrapolated to human health and disease. Credit: College of Science

Just as there is a mysterious dark matter that accounts for 85 percent of our universe, there is a "dark" portion of the human genome that has perplexed scientists for decades. A study published March 9, 2020, in *Genome Research* identifies new portions of the fruit fly genome that, until now, have been hidden in these dark, silent areas.

The collaborative paper titled "Gene Expression Networks in the *Drosophila* Genetic Reference Panel" is the culmination of years of research by Clemson University geneticists Trudy Mackay and Robert Anholt. Their groundbreaking findings could significantly advance science's understanding of a number of genetic disorders.

The "dark" portion refers to the approximate 98 percent of the genome that doesn't appear to have any obvious function. Only 2 percent of the human genome codes for proteins, the building blocks of our bodies and the catalysts of the chemical reactions that allow us to thrive. Scientists have been puzzled by this notion since the 1970s when gene sequencing technologies were first developed, revealing the proportion of coding to noncoding regions of the genome.

Genes are traditionally thought to be transcribed into RNAs, which are subsequently translated into proteins, as dictated by the central dogma of molecular biology. However, the entire assemblage of RNA transcripts in the genome, called the transcriptome, contains RNA species that appear to have some other function, apart from coding for proteins. Some have proposed that noncoding regions might contain regulatory regions that [control gene expression](#) and the structure of chromosomes, yet these hypotheses were difficult to study in past years as diagnostic technology was developing.

"Only in recent years, with the sequencing of the entire transcriptome complete, have we realized how many RNA species are actually present. So, that raises the whole new question: if they aren't making the

proteins—the work horses of the cell—then what are they doing?" said Mackay, director of Clemson University's Center for Human Genetics (CHG), which is part of the College of Science.

For Mackay and Anholt, also of the CHG, these [human genetics](#) questions can be probed by studying the common fruit fly, *Drosophila melanogaster*. Because many genes are conserved between humans and fruit flies, findings revealed by analyzing the *Drosophila* genome can be extrapolated to human health and disease.

Mackay and Anholt's former postdoctoral researchers, Logan Everett and Wen Huang, led the charge on this latest research, which identified more than 4,500 new transcripts in *Drosophila* that have never been uncovered before. Referred to by the researchers as "novel transcribed regions," these 4,500 transcripts consist primarily of noncoding RNAs that appear to be involved in regulating networks of genes and that could contribute to genetic disorders.

"Most disease-causing mutations are known to occur in the protein-coding portion of the genome, known as the exome, but when you're only sequencing the exome, you miss other disease-related factors in other parts of the genome, such as these long noncoding RNAs," said Anholt, Provost's Distinguished Professor of Genetics and Biochemistry at Clemson University. "Now that the cost of whole genome sequencing has gone down considerably, and we have the capability of sequencing whole genomes rapidly, we can look at elements of the genome that have traditionally been considered unimportant, and we can identify among them potential disease-causing elements that have never been seen before."

By probing several hundred inbred *Drosophila* fly lines, each containing individuals that are virtually genetically identical, the researchers discovered that many of the novel long noncoding RNAs regulate genes

in heterochromatin, a tightly packed form of DNA in the genome that is usually considered "silent." Because heterochromatin is so condensed, it was thought to be inaccessible to the molecular machinery that transcribes DNA into RNA. Thus, any genes contained within heterochromatin are kept off, silent and unexpressed—or are they?

"What we think is that the repression of gene expression in heterochromatin is somewhat leaky, and that there is variation in how those genes are repressed," Mackay said. "The network of RNAs we've discovered may have to do with actually regulating chromatin state."

"These noncoding RNAs may play an important role in opening up such regions of the genome for expression of genes in a way that varies among different individuals depending on their genetic background," Anholt added.

Another outcome of the study is the expression of "jumping genes," known as transposons, that are pieces of DNA able to move around the genome. As transposons cut and paste into other genes, they may cause genome instability that leads to cancer, neurodegenerative disorders and other diseases.

These transposons were also located in heterochromatin, but the identification of transcripts of these transposons shows that they are actually being expressed, despite residing in a usually silent portion of the genome. Identifying regulators of transposable elements, as the researchers found among these 4,500 "novel transcribed regions," could prove useful in treating disorders that stem from transposon interference.

Overall, the study lends toward a greater understanding of gene regulatory networks that contribute to human health and disease.

"These observations open up an entirely new area of biology that hasn't

been explored and has unlimited potential for future follow-up," Anholt said.

The team's own follow-up studies are using CRISPR gene editing technology to uncover what happens when genes revealed by this study are altered or deleted from the *Drosophila* [genome](#). If the expression of other [genes](#) is altered by knocking one out, important conclusions can be drawn about the role that deleted gene plays in development or progression of disease.

**More information:** Logan J Everett et al, Gene expression networks in the *Drosophila* Genetic Reference Panel, *Genome Research* (2020). [DOI: 10.1101/gr.257592.119](https://doi.org/10.1101/gr.257592.119)

Provided by Clemson University

Citation: Genetics research sheds light on 'dark' portion of genome (2020, March 11) retrieved 6 May 2024 from <https://phys.org/news/2020-03-genetics-dark-portion-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.