Scientists cut "Gordian knot" in the human genome

Allison Huseman
(713) 798-4710
Houston, TX - Jul 18, 2016

Females have two X chromosomes in each of their cells. Fully unfolded, each copy is two inches long. One of these two X chromosomes is inactive – its genes are turned off. This copy folds into a structure called the Barr body, a mysterious configuration that was discovered in 1949. Recently, scientists have shown that the Barr body contains massive superloops bringing DNA sequences at opposite ends of the chromosome together inside the nucleus of a cell.

Now, a team of scientists at Baylor College of Medicine, Florida State University and the Broad Institute of MIT and Harvard has determined which part of the DNA code is responsible for these superloops and has shown that it is possible to use this information to change the structure of the Barr

An artwork inspired by the problem of genome folding, by Broad Institute artist-in-residence Guhapriya Ranganathan. This originally appeared in her exhibition, Unfolding.

body as a whole. The report, which sheds light on female development in mammals, appears today in *Proceedings of the National Academy of Sciences*.

"X inactivation is fundamentally important for female development," said Dr. Miriam Huntley, co-first author on the study. "Without it, females would generate too much of every gene product of the X chromosome."

Huntley recently received her Ph.D. at Harvard University, where she worked with co-senior author Dr. Erez Lieberman Aiden, assistant professor of molecular and human genetics, a McNair scholar and director of the Center for Genome Architecture at Baylor. In earlier work, Huntley and her colleagues in the Aiden lab created the first genome-wide map of loops – positions where the genome folds back on itself in the nucleus of a cell and points that lie far apart along the contour of a chromosome come together in 3-D.

In the process, they demonstrated that the Barr body – the inactive X chromosome that is present in females – contains superloops, structures that have no analog in men. "The typical loop in a male genome spans about 200,000 letters of the DNA code. If fully stretched out, it would be about three thousandths of an inch long. But the loops in the Barr body can span as many as 77 million DNA letters – an inch of DNA," said Huntley. "We call these giant structures superloops."

Independently, the laboratory of Dr. Brian Chadwick at Florida State University had been studying X inactivation with a different set of techniques. "We had found that the human inactive X was organized into at least two different types of silent DNA that alternated along the chromosome. At one of the intersections was a strange DNA element called DXZ4, where a single sequence repeated, over and over again, for hundreds of thousands of letters," said Chadwick, co-senior author of the new study. "I've always been fascinated with what the purpose of this repeat was. Other large repeats in our genome perform structural roles, such as the centromeres and telomeres. I was convinced that the role of DXZ4 was just waiting to be discovered."

In the new study, Huntley, co-first author Emily Darrow and their co-authors show that superloops are not unique to humans; they are found in rhesus macaque monkeys as well as in mice. In all three species, DXZ4 was at one end of superloops. "The presence of DXZ4 at one end of human superloops could have been a coincidence," said Aiden, who also is a faculty member at Rice University in computer science and at the Center for Theoretical Biological Physics. "But the fact that DXZ4 lies at the site of superloops in all three species was a critical clue. It's a bit like seeing the same suspect at the scene of multiple crimes. The chances are higher that you've found the culprit."

The team also showed that the superloops create a hub in which many distant pieces of the X chromosome come together. "We introduced an approach called concatemer ligation assay, or COLA, which allows us to track collisions between three or more positions in the genome," said Huntley. "It's similar to looking through someone's

Media Queries?

During business hours call: (713) 798-4710
After business hours call: (713) 775-6912

photo album. Eventually, you might notice patterns, such as the same group appearing together in many of the pictures. COLA made it clear that the superloops were bringing many pieces of the inactive X chromosome together, all at once."

"The inactive X hub is a unique structure in the mammalian genome," said Aiden. "The loops that form it are macroscopic objects – but, like a Gordian knot, they pack that extraordinary length into a tiny space, a microscopic corner of the cell nucleus."

With so much evidence pointing to DXZ4, the team took the final step; they checked what happens when you delete it. All of a sudden, the hub was disrupted, and the inch-long loops anchored near DXZ4 were gone. "This work assigns the first known function to this type of tandem repeat and confirms that, like centromeres and telomeres, DXZ4 performs a structural role that in this case is unique to the inactive X chromosome," said Chadwick.

The results also have consequences for the emerging field of 3-D genome engineering. In a recent study, scientists at the Center for Genome Architecture showed that it was possible to engineer the smaller loops that are present in both males and females, adding and removing them by design. "Disrupting the inactive X hub was much more ambitious," said Darrow, a graduate student in the Chadwick lab. "After all, it is vastly larger."

"The ability to modify how a genome folds in 3-D is an emerging frontier," said Dr. Eric Lander, director of the Broad Institute and co-corresponding author. "As we understand more and more of the DNA elements that play a role in genome folding, our toolbox will continue to expand."

"These results suggest a promising future for the field of 3-D genome engineering," said Huntley. "The deletion of DXZ4, several hundred thousand DNA letters, led to predictable changes in folding at the scale of an entire chromosome, spanning over 150 million letters. This is another step toward the dream of precisely sculpting the 3-D organization of the genome."

Other contributors to this work include Olga Dudchenko, Elena K. Stamenova, Neva C. Durand, Zhuo Sun, Su-Chen Huang, Adrian L. Sanborn, Ido Machol, Muhammad S. Shamim, and Andrew P. Seberg at Florida State University, Baylor College of Medicine, Harvard University, the Broad Institute of MIT and Harvard, the Massachusetts Institute of Technology, Harvard Medical School, Rice University, and Stanford University.

This research was supported by NIGMS grant R01GM073120; a National Science Foundation Graduate Research Fellowship under Grants DGE0946799 and DGE1144152; an NIH New Innovator Award (1DP2OD008540-01), an NIH 4D Nucleome Grant (U01HL130010), an NSF Physics Frontier Center (PHY-1427654, Center for Theoretical Biological Physics), the NHGRI Center for Excellence for Genomic Sciences (HG006193), an NVIDIA Research Center Award, an IBM University Challenge Award, a Google Research Award, a Cancer Prevention Research

Center Award, an IBM University Challenge Award, a Google Research Award, a Cancer Prevention Research Institute of Texas Scholar Award (R1304), a McNair Medical Institute Scholar Award, the President's Early Career Award in Science and Engineering, and an NHGRI grant (HG003067).

For a complete list of conflict of interests see the publication.