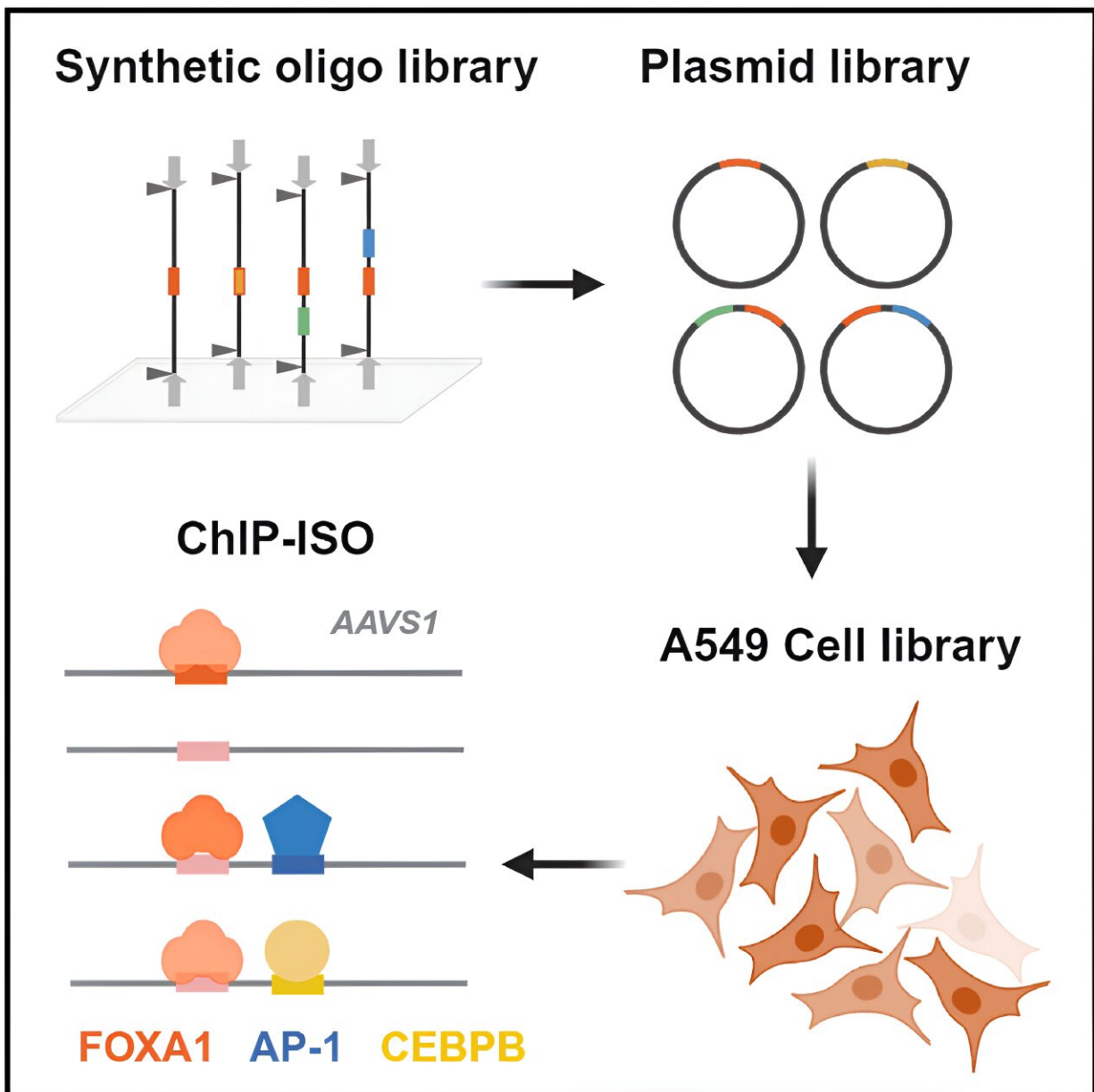


# How proteins bind to closed regions of the genome to facilitate cell differentiation and development

September 11 2024, by Sam Sholtis



Credit: *Molecular Cell* (2024). DOI: 10.1016/j.molcel.2024.06.022

A new experimental method allows researchers to dissect how certain proteins, called pioneer factors, can bind to selective regions of the genome that are inaccessible to other DNA binding proteins.

The Penn State researchers who developed the method [published](#) their approach in the journal *Molecular Cell*. The work provides what they called a "powerful" way to gain insight into how genes are regulated.

The genome contains the entirety of the genetic makeup of an organism, but only a subset of this information is used in individual cells. That subset helps determine how the cells develop and differentiate to perform their specialized functions.

Proteins called [transcription factors](#) interact with the DNA in the genome to control the specific set of genes expressed in a cell type by binding to short patterns of DNA sequence called binding motifs, but often those motifs are inaccessible to the transcription factor proteins because of how the DNA is packaged in the cell.

Led by Lu Bai, professor of biochemistry and [molecular biology](#) and of physics in the Eberly College of Science at Penn State, researchers have now developed a new technique that can test thousands of sequence variants of binding motifs in a single experiment.

With this technique, the researchers can begin to identify features of the motifs that allow the specialized transcription factors, called pioneer factors, to access these typically inaccessible genomic regions and open

them up for additional access.

The technique also allows the researchers to parse out how the pioneer factors work together with other co-factor proteins, which could inform which motifs are bound in which specific cells.

The new method is named ChIP-ISO or Chromatin Immunoprecipitation with Integrated Synthetic Oligonucleotides.

"All of the cells of an organism contain the same genome, but not all cells are the same," Bai said. "Different cell types are different because of the set of genes that they express. Gene expression is regulated by transcription factors that bind to the DNA at specific short sequence motifs that are found across the genome. But only a very small portion of these sequence motifs are actually used at any one time in a cell, and we are interested in how this specificity is determined."

The chromosomes in the nucleus of a cell are composed of long strands of DNA packaged with various proteins into a structure called chromatin. Such packaging limits the accessibility of many DNA-binding factors, including transcription factors, allowing them to only bind to a subset of their motifs.

Pioneer factors, however, are special because they can bind to DNA even in tightly packed, or closed, chromatin. Despite this ability, even pioneer factors only bind, or associate, with a small fraction of their motifs, the researchers said.

"It's perplexing why pioneer factors only bind to some of their motifs, so we designed ChIP-ISO as a method to pick apart characteristics of the DNA motifs that allow them to associate or not associate with pioneer factors," Bai said. "A benefit of the technique is that is extremely high-throughput, allowing us to test thousands of variations of the binding

motif's DNA sequences in a single experiment."

To test their new method, the research team focused on the well-known pioneer factor, FOXA1, along with several co-factor proteins that potentially influence when and in what cell types FOXA1 binds to DNA.

In the ChIP-ISO experiment, the researchers designed short synthetic DNA sequences that contain variants of the binding sites for the pioneer factor, FOXA1, as well as binding sites of the co-factors proteins. The researchers then integrate thousands of these short sequence variants into the genome of millions of cells grown in the lab. They can then determine which variants are actually bound by FOXA-1 and its co-factors in the cells.

"If FOXA1 binds the synthetic sequences we introduced into a cell, we can capture that bit of DNA and determine its sequence to learn how the variants impact binding," Bai said. "As expected, the FOXA1 motif itself is important for FOXA1 binding. However, what surprised us is that some other transcription factors near the FOXA1 motif can be almost as essential for FOXA1 binding in some sequence contexts."

The team found that mutating the binding sites for the co-factors, AP-1 and CEBPB, led to a significant drop in FOXA1 binding. Mutating the AP-1 binding site had a particularly strong effect, suggesting that it plays a crucial role in directing FOXA1 binding in these cells. They also found that local sequence variations had a larger impact on FOXA1 binding than chromatin context, or how densely DNA is packed into the chromatin structure.

"In our cells, AP-1 is the most important co-factor for FOXA1, but in different cell types, other co-factors may play that role," Bai said. "There are many other pioneer factors, and potentially many other co-factors, and we are interested in dissecting this complex network. The ChIP-ISO

method gives us a platform to test the combination of different factors and start to understand their role in determining cell identity."

The researchers also used neural networks to analyze variations in FOXA1 binding across different cell types using publicly available data from previous studies using different methods to study FOXA1.

"Our neural network analyses confirmed that the motifs of other transcription factors can explain why FOXA1 binds to different locations in different cell types," said Shaun Mahony, associate professor of biochemistry and molecular biology at Penn State and an author of the paper.

"This combination of ChIP-ISO's ability to test thousands of customized DNA sequences with machine learning-based analyses is a very powerful way to gain insight into gene regulatory systems."

**More information:** Cheng Xu et al, Systematic dissection of sequence features affecting binding specificity of a pioneer factor reveals binding synergy between FOXA1 and AP-1, *Molecular Cell* (2024). [DOI: 10.1016/j.molcel.2024.06.022](https://doi.org/10.1016/j.molcel.2024.06.022)

Provided by Pennsylvania State University

Citation: How proteins bind to closed regions of the genome to facilitate cell differentiation and development (2024, September 11) retrieved 11 September 2024 from <https://phys.org/news/2024-09-proteins-regions-genome-cell-differentiation.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.