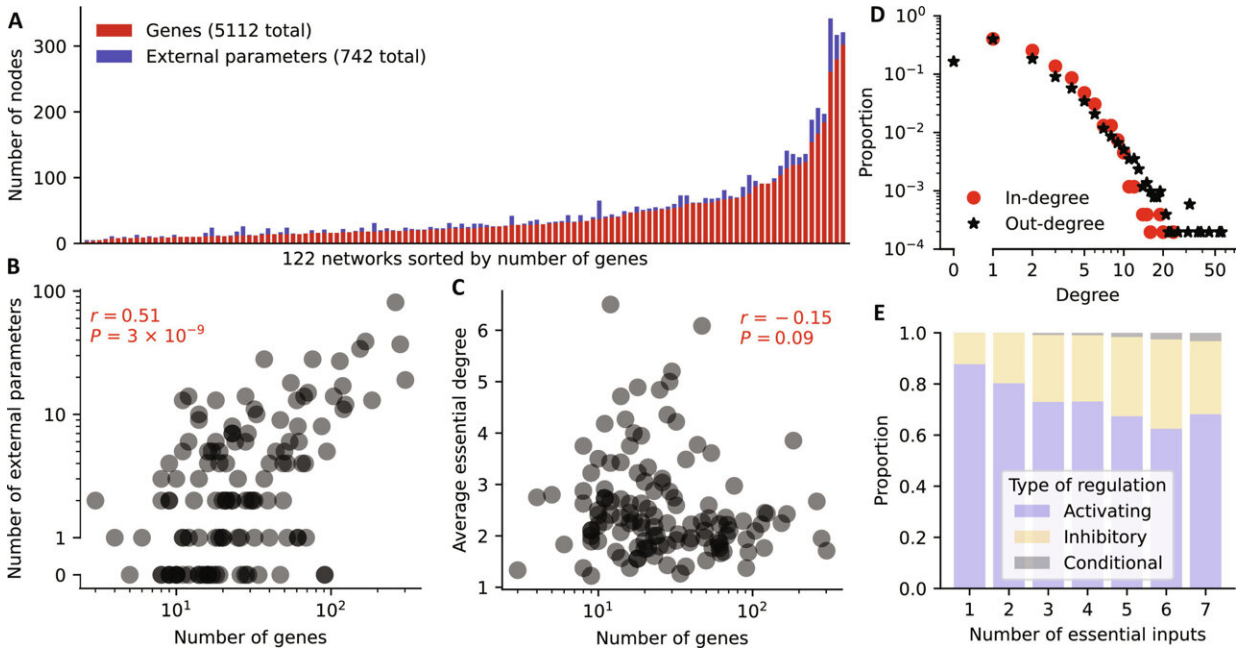
New research guides mathematical model-building for gene regulatory networks

January 23 2024



Summary statistics of the analyzed GRN models. (A) Plot of the number of genes and external parameters for each model sorted by number of genes. (B and C) For each model, the number of genes is plotted against (B) the number of external parameters and (C) the average essential in-degree of the genes. The Spearman correlation coefficient and associated P value are shown in red. (D) In-degree (red circles) and out-degree (black stars) distribution derived from all 5112 update rules. (E) Prevalence of each type of regulation (activation, blue; inhibition, orange; conditional, gray) stratified by the number of regulators (x axis). Nonessential regulations are excluded. Credit: *Science Advances* (2024). DOI: 10.1126/sciadv.adj0822

Over the last 20 years, researchers in biology and medicine have created Boolean network models to simulate complex systems and find solutions, including new treatments for colorectal cancer.

"Boolean network models operate under the assumption that each gene in a regulatory network can have one of two states: on or off," says Claus Kadelka, a systems biologist and associate professor of mathematics at Iowa State University.

Kadelka and undergraduate student researchers published a study in *Science Advances* that disentangles the common design principles in these mathematical models for gene regulatory networks.

He says showing what features have evolved over millions of years can "guide the process of accurate model building" for mathematicians, computer scientists and synthetic biologists.

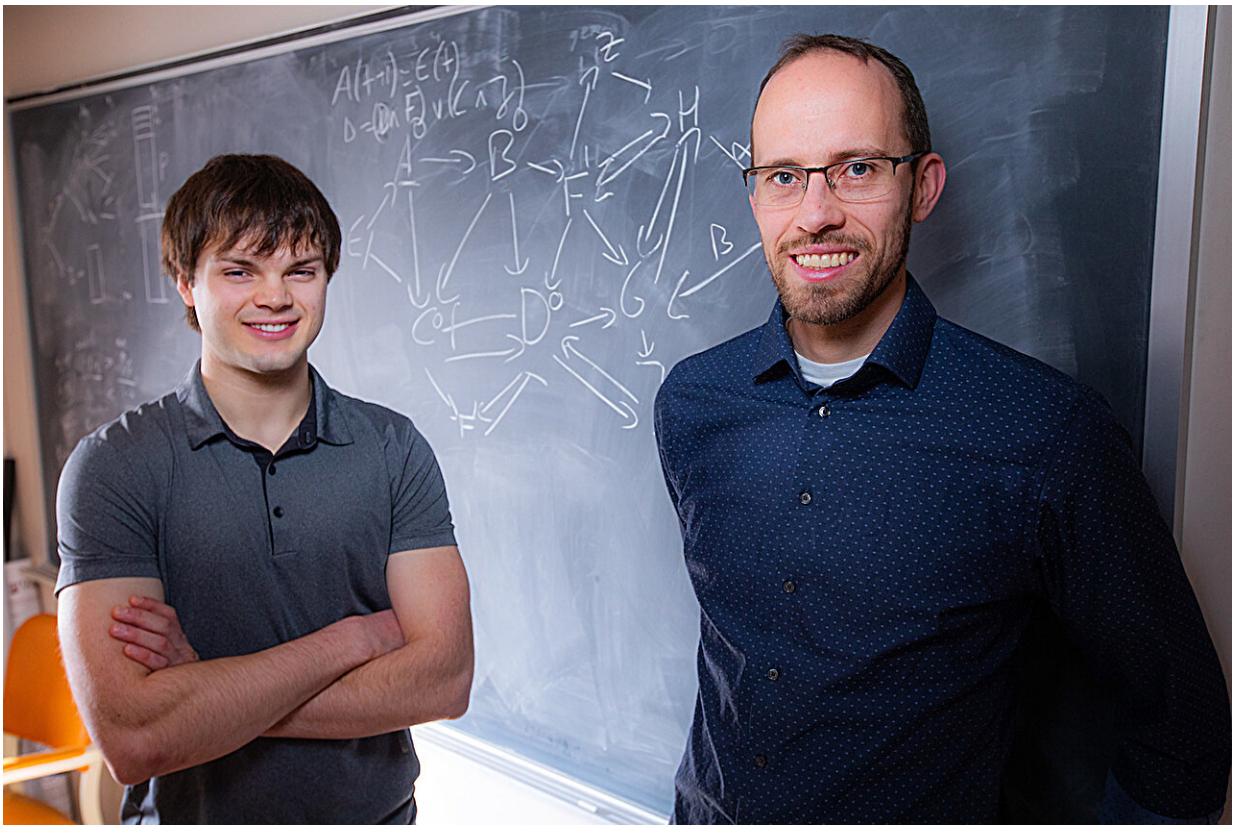
"Evolution has shaped the networks that control the decision-making of our cells in very specific, optimized ways. Synthetic biologists who try to engineer circuits that perform a particular function can learn from this evolution-inspired design," says Kadelka.

Gene regulatory networks determine what happens and where it happens in an organism. For example, they prompt cells in your stomach lining—but not in your eyes—to produce hydrochloric acid, even though all the cells in your body contain the same DNA.

On a piece of paper, Kadelka draws a simple, hypothetical gene regulatory network. Gene A produces a protein that turns on gene B, which turns on gene C, which turns off gene A. This negative feedback loop is the same concept as an air conditioner that shuts off once a room reaches a certain temperature.

But gene regulatory networks can be large and complex. One of the Boolean models in the researchers' dataset involves more than 300 genes. And along with negative feedback loops, gene regulatory networks may contain positive feedback loops and feed-forward loops, which reinforce or delay responses. Redundant genes that perform the same function are also common.

Among these and other design principles highlighted in the new paper, Kadelka says one of the most abundant is "canalization." It refers to a hierarchy or importance ordering among genes in a network.



Iowa State senior Addison Schmidt (left) and associate professor of mathematics Claus Kadelka (right) stand in front of a chalk board with depictions of a gene regulatory network in Carver Hall. Credit: Christopher Gannon/Iowa State University

Accessible data, bolstered with undergraduate research

Kadelka emphasizes that the project would have been difficult to complete without the "First-Year Mentor Program," which matches students in the Iowa State Honors Program with research opportunities across campus.

Undergraduate students helped Kadelka develop an algorithm to scan 30 million biomedical journal articles and filter those most likely to include Boolean biological network models. After reviewing 2,000 articles one by one, the researchers identified around 160 models with close to 7,000 regulated genes.

Addison Schmidt, now a senior in computer science, is one of the paper's co-authors. When he worked on the project as a freshman in 2021, he created an online database for the project.

"A major benefit of the research is that it collects and standardizes Boolean gene regulatory networks from many sources and presents them, along with a set of analysis tools, through a centralized web interface. This expands the accessibility of the data, and the web interface makes the analysis tools useable without a programming background," says Schmidt.

Kadelka says systems biologists have used the database for their research and expressed gratitude for the resource. He plans to maintain and update the website and investigate why evolution selects for certain design principles in gene regulatory networks.

As for Schmidt, he says working on the project as a freshman helped him expand his expertise with the Python programming language and

become more comfortable applying his skills to research.

"This project also motivated me to pursue other research at Iowa State where I developed other tools and, coincidentally, another website to present them," says Schmidt.

He adds that he appreciated Kadelka's mentorship and hopes the First-Year Mentor Program will continue to foster opportunities for undergraduate research at Iowa State.

More information: Claus Kadelka et al, A meta-analysis of Boolean network models reveals design principles of gene regulatory networks, *Science Advances* (2024). DOI: 10.1126/sciadv.adj0822

Provided by Iowa State University