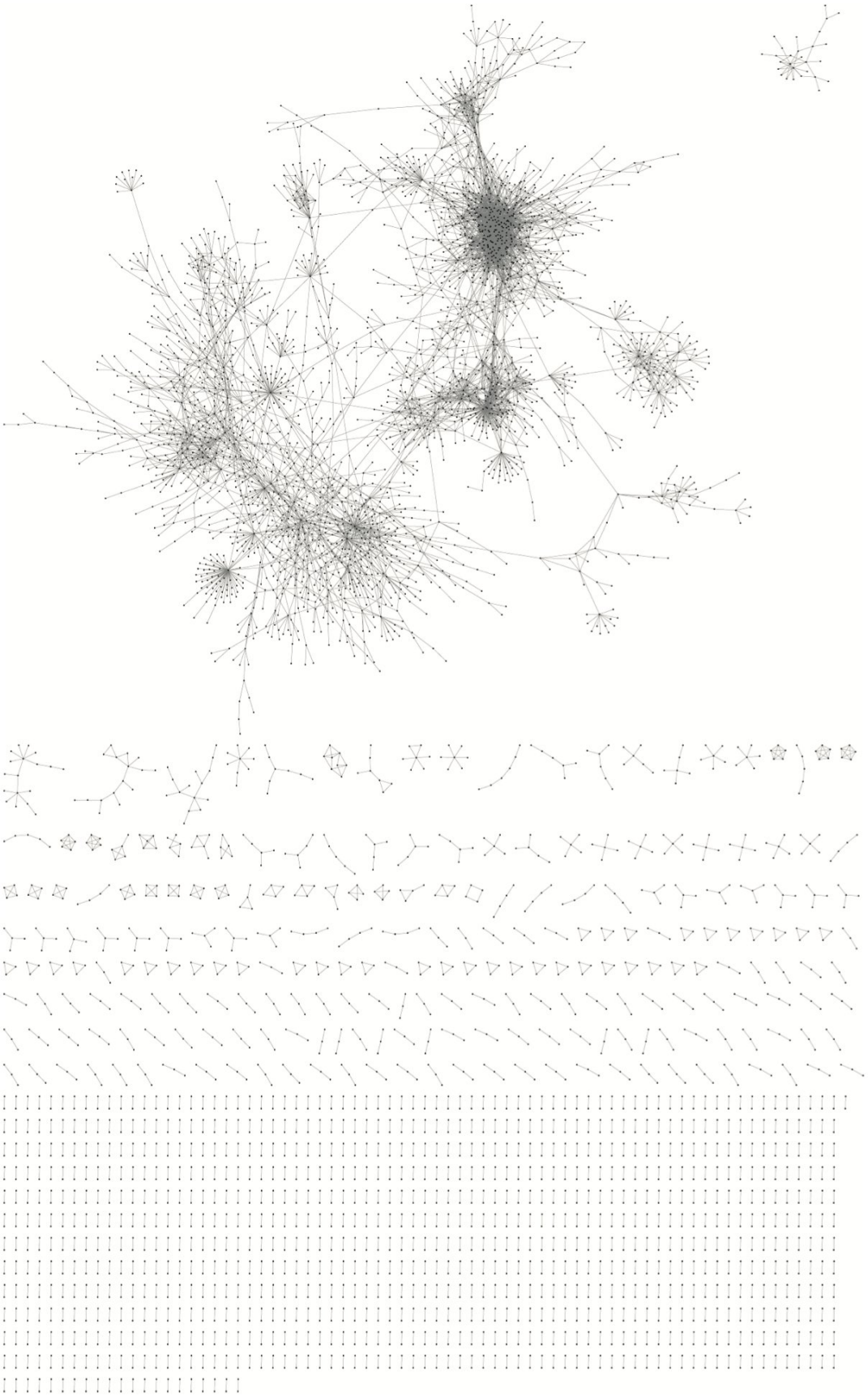
Undergraduate student uncovers genes associated with aggressive form of brain cancer

February 8 2018



In the NCBI brain network, each point represents a gene, and each line is indicative of gene co-expression. Credit: *Oncotarget*, doi:10.18632/oncotarget.24228

When Leland Dunwoodie, an undergraduate researcher in biochemistry, approached his PI about wanting to start research on "some human stuff" in the spring of 2016, he didn't imagine it would lead to the discovery of 22 genes that are implicated in glioblastoma, the most aggressive type of brain cancer.

"I definitely didn't come to Clemson thinking about brain cancer research," Dunwoodie said. "I was working on a project with grapes and other plants. I told Dr. (Alex) Feltus that I wanted to do some human stuff, and he said, 'That's cool - pick an organ.' "

After consulting with his family - should he study the brain or the heart? - Dunwoodie decided on the brain, and specifically on brain cancer. A prior summer internship at the Van Andel Institute had spurred his interest in cancer research.

Fast-forward two years later to a January 2018 publication in the journal *Oncotarget*, Dunwoodie's study is the first to describe glioblastoma -specific gene co-expression relationships between a group of 22 specific genes.

Heard in the news as the disease afflicting Senator John McCain and Beau Biden, the late son of U.S. Vice President Joe Biden, glioblastoma is highly malignant and is characterized by its lethality. Patients with glioblastoma have a median survival time of only 14.6 months after

diagnosis.

"Like many other tumors, diseases, and complex traits, glioblastoma is controlled by a variety of genetic and epigenetic factors," Dunwoodie said. "If there was one master-regulator of these cancers, we'd say, 'We're going to drug that, and we're going to save millions of lives every year,' but there are more things going on in glioblastoma than we can presently identify."

However, the complexity of glioblastoma is fitting for research in professor Feltus' Systems Genetics Lab in the department of genetics and biochemistry, where Dunwoodie is a student. Systems genetics, as the lab's name implies, uses computer- and mathematics-based approaches to analyze biological systems, such as genes and regulatory pathways.

To make the discovery, Dunwoodie first compiled data from two online public databases for genomic information: The Cancer Genome Atlas (TCGA) and the National Center for Biotechnology Information (NCBI).

From TCGA, more than 2,000 tumor expression datasets were downloaded, each one detailing how tumor cells differ from normal cells at the genetic level. Five different types of tumors, including those from bladder, ovarian, thyroid, lower-grade glioma and glioblastoma cancers, were included in the data to achieve a well-rounded case study.

The 2,000-plus datasets, each showing approximately 75,000 genes, were then organized into a gene expression matrix (GEM), a table that quantifies the expression level of each gene across every sample. For example, one of the genes pulled from TCGA, called LAPTM5, encodes a protein that is involved in the formation of blood cells. In the gene expression matrix, LAPTM5 was assessed across each tumor type to

gauge whether it is overly active (overexpressed) or underactive (underexpressed) in one tumor type versus another, indicated by a numerical ranking. The same scoring process was then conducted for the 74,999 remaining genes across the five tumor types in the TCGA data.

A separate GEM, encompassing 210,000 genes from 204 datasets from the NCBI database - including normal brain samples, glioblastoma brain samples and brain samples from patients with Parkinson's disease - was created independently for comparison. Will Poehlman, a graduate student in the Systems Genetics Lab, assisted Dunwoodie in preparing these GEMs.

Using novel computer software developed by Feltus and former graduate student Stephen Ficklin - who is now an assistant professor at Washington State University - Dunwoodie was then able to translate the GEMs into two different gene co-expression networks (GCNs), a visual representation of the data that provides insights into how the genes interact with one another.

The software package, known as Knowledge Independent Network Construction (KINC), is novel in that it finds expression relationships between genes without the researchers having to conduct any prior analyses. This knowledge-independent method reduces the amount of "noise" - from laboratory protocols or from natural variation between cells - that can prevent genetic interactions from being discovered.

"Through the two GCNs, we found a group of 22 genes that were co-expressed in a single module both in The Cancer Genome Atlas network and in the NCBI brain network," Dunwoodie said. "Only about 70 genes overlapped between the two networks, and 22 of them were in the same module - the same group of co-expressed genes. The overlap was really easy to spot."

While it's tempting to think that the genes - many of which function in the immune system - are feeding off one another to affect glioblastoma, Dunwoodie says this isn't exactly the case.

"It's hard to say that they're working together, because these are correlations. So, if person A runs eight miles on the same day that person B runs eight miles, it doesn't necessarily mean that they're running together," Dunwoodie said. "It's more likely that these genes are being regulated in the same way, and there are probably several things regulating them that we can't currently identify."

What's more - these 22 genes, when compared between glioblastoma and noncancerous samples, were found to have much stronger co-expression levels in glioblastoma, suggesting a disease-specific regulatory mechanism. The same finding was uncovered when comparing glioblastoma to lower-grade glioma, a less aggressive type of brain cancer, indicating a glioblastoma-specific activity to the 22 genes. The other notable finding of the study showed that the 22 genes are more associated with mesenchymal glioblastoma, a distinct subtype of the cancer, and that when the genes are highly expressed, they decrease survival time for patients in the mesenchymal group.

As is the case in research, where answering one question results in a plethora of new questions, the team's study is just one step toward understanding glioblastoma pathogenesis.

"It would be nice to find out what the 22 genes are specifically doing," Dunwoodie said. "Are they expressed in the surrounding immune cells? Are they a cause of cancer, or are they an effect of cancer? Does cancer propagate their expression? Why these genes are co-expressed there and what they're doing are questions that haven't been answered."

Dunwoodie - who plans to attend medical school to become a physician

informaticist - says the tools and methods he learned in the Systems Genetics Lab will stick with him long into his career.

"Cancer research is interesting because there are so many amazing people doing so many amazing things - but this is just one drop in the bucket," Dunwoodie said. "For me, the real purpose is patients being cured. Getting a paper published is great, but no one was immediately cured because of this, and that's the ultimate goal."

More information: Leland J. Dunwoodie et al, Discovery and validation of a glioblastoma co-expressed gene module, *Oncotarget* (2018). DOI: 10.18632/oncotarget.24228

Provided by Clemson University