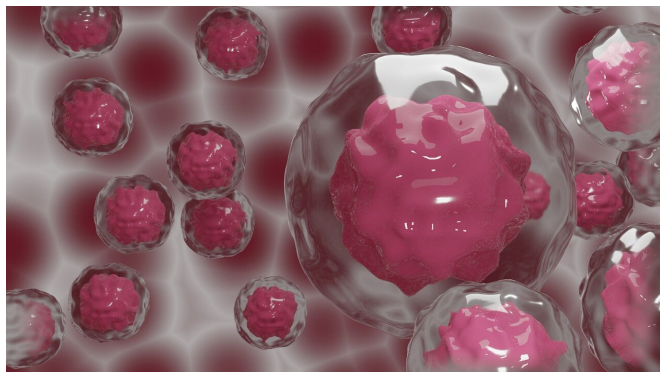


Researchers make glycomics data AI-ready

8 September 2021



Credit: CC0 Public Domain

Researchers at the University of California San Diego have created a tool that allows glycomics datasets to be analyzed using explainable Artificial Intelligence (AI) systems and other machine learning approaches. In a recent paper published in *Nature Communications*, the team demonstrated that glycomics data require extra care to be properly used for statistical analysis or machine learning. They also offer a new preprocessing solution to prepare glycomics data to substantially boost the power of its use with machine learning and AI. They named the approach GlyCompare. It takes a systems level perspective that accounts for shared biosynthetic pathways of glycans within and across samples.

To introduce GlyCompare, the team demonstrated their ability to enhance comparisons of glycomics datasets by shining light on the hidden relationships between glycans in several contexts, including [gastric cancer](#) tissues. Cancer is a useful example given the importance of glycan changes to cancer and its utility for early-stage diagnosis.

"We applied GlyCompare to cancer tissues and showed that while one couldn't find cancer specific glycans using standard statistical methods, novel biomarkers emerge when processed using our method," said UC San Diego professor of

Bioengineering and Pediatrics Nathan Lewis, who is the corresponding author on the paper. Lewis co-directs the CHO Systems Biology Center, and glycoengineered CHO cell lines were used to produce diverse proteins used in the study.

In another analysis, the team showed the method substantially boosts statistical power, such that one needs half as many samples to get equivalent power to detect biomarkers. In the paper, the researchers outline how the methods behind GlyCompare will be transformative for bringing glycomics to the clinic. In fact, Lewis is part of the founding team of a new start-up that is licensing related intellectual property to commercialize this technology for high value applications, including cancer diagnostics.

One of the keys to the GlyCompare approach is that it looks at the biological steps needed to synthesize the subunits that make up glycans, rather than just looking at only the whole glycans themselves, greatly improving the accuracy of statistical analyses of glycomics data. The researchers believe this approach will enable the detection of more subtle changes in glycosylation in many applications, including early stage cancer. Moreover, GlyCompare could lead to new insights on the mechanisms behind the observed changes in [glycans](#) that are present.

Bokan Bao and Benjamin P. Kellman, the co-first-authors on the paper, are both in the Bioinformatics and Systems Biology Graduate Program, and members of the Department of Bioengineering at the UC San Diego Jacobs School of Engineering.

More information: Bokan Bao et al, Correcting for sparsity and interdependence in glycomics by accounting for glycan biosynthesis, *Nature Communications* (2021). [DOI: 10.1038/s41467-021-25183-5](https://doi.org/10.1038/s41467-021-25183-5)

Provided by University of California - San Diego

APA citation: Researchers make glycomics data AI-ready (2021, September 8) retrieved 1 December 2021 from <https://medicalxpress.com/news/2021-09-glycomics-ai-ready.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.