狮城新闻

15 October 2026

Translated from Chinese

NTU develops cell imaging analysis tool that can identify thousands of cell changes at once

Proteins within cells can also clump together like oil droplets in water, separating from their surroundings. This process, known as "phase separation," helps cells respond quickly to changes, but it can also lead to abnormal aggregation when it gets out of control, which is associated with diseases such as Alzheimer's.

Now, a research team at Nanyang Technological University (NTU Singapore) has used artificial intelligence to provide researchers with a more systematic "protein phase separation map".

The team developed the AI-driven benchmark database PLATO and the analysis platform PhaseHub to help researchers more accurately predict which proteins will undergo phase separation, which molecules they may interact with, and how this behavior varies in different species.

Cross-species analysis scale

1,106 species

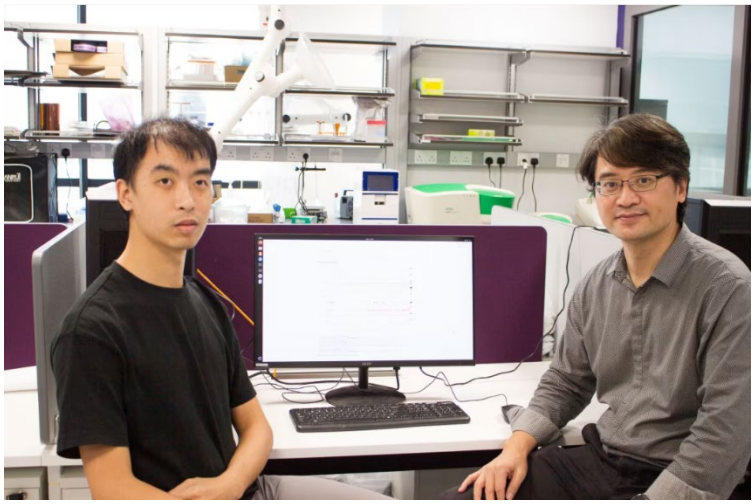
Covering life forms such as animals, plants, fungi, and bacteria

Why do proteins "clump together"?

Phase-separating proteins detach from their surroundings and aggregate into droplet-like structures. Cells can use this to aggregate specific molecules, accelerating or inhibiting certain reactions, thus responding to stress more quickly.

However, if this process malfunctions, proteins may form harmful aggregates or condensates. Such changes are associated with neurodegenerative diseases such as cancer and Alzheimer's disease; in plants, abnormal phase separation can also impair stress response and the ability to resist microbial pathogens.

Therefore, identifying which proteins undergo phase separation and what factors regulate this behavior is an important question in disease, aging, and crop research.



From left: Dr. Liang Qiyu and Professor Miao Yansong, researchers at the NTU Institute for Digital Molecular Analysis and Science. Image credit: NTU Singapore

PLATO: Giving AI a more reliable "standard answer"

Past protein phase separation datasets have often focused on whole proteins, rarely including experimentally confirmed results of "no phase separation," and lacking data on truncated proteins, mutants, and engineered variants. This data imbalance can easily lead predictive models to overestimate the tendency of proteins to aggregate.

To fill the gap, the NTU team used a large language AI model system to search scientific papers indexed in PubMed, extract information on proteins and their phase behavior that had been tested in experiments, and then had researchers manually verify the information to ensure scientific accuracy.

PLATO database includes:

• 5,118 unique protein sequences

• More than 3,600 proteins that undergo phase separation

• Approximately 1,300 proteins that do not undergo phase separation

• Also covers truncated proteins, mutants, and engineered variants.

This more balanced and diverse data can provide a benchmark for training and evaluating next-generation predictive tools.

PhaseHub: More than just prediction, it explores "who interacts with whom".

The research team also established the PhaseHub platform. They first used the previously developed machine learning algorithm MolPhase to predict proteins that may undergo phase separation in 1,106 species, and then combined experimental data

on protein abundance and interactions to analyze multi-component protein condensates.

Researchers can search for proteins or submit sequences in PhaseHub to view their phase separation predictions, visualize potential interactions, and explore features that influence molecular interactions, such as consecutively repeating amino acid fragments.



What did the team discover, spanning across species of life?

Analysis shows that eukaryotes, whose cells contain a nucleus, typically have a higher proportion of potentially phase-separating proteins. For organisms with smaller genomes, the proportion of potentially phase-separating proteins increases with genome expansion, then gradually plateaus.

This suggests a possible balance between genomic DNA content and phase-separating ability: a larger genome can encode more phase-separating proteins, supporting more complex cellular activities; but an excessive tendency to aggregate may also increase the risk of forming harmful aggregates.

Who completed this study?

The research was led by Professor Miao Yansong of the College of Life Sciences and the Institute for Digital Molecular Analysis and Science (IDMxS) at NTU , in collaboration with Professor Gao Weibo of the College of Electrical and Electronic Engineering at NTU .

IDMxS researcher Dr. Liang Qiyu is the first author of both research papers. The related AI tools and research findings have been published in the peer-reviewed journal *Molecular Plant*.

What does this research mean?

It is not a treatment technology that has entered clinical trials, but rather a more accurate and user-friendly tool for basic research. In the future, it may help researchers

screen candidate proteins more quickly, understand disease and aging mechanisms, and study how to enhance plant stress resistance and immunity.

From single protein sequences to life maps covering over a thousand species, AI is helping scientists connect evidence scattered across papers and experiments. What's truly promising is not just faster predictions, but also the provision of a shared foundation for disease research and sustainable agriculture that can be continuously accumulated, compared, and validated.

https://www.shicheng.news/v/mgz5p