Earth's species have more in common than previously believed
17 June 2020

The Earth hosts an abundance of life forms—from well-known animals and plants to small, more hardy life forms such as archaea, viruses and bacteria. These life forms are fundamentally different all the way down to the cell level—or so scientists thought.

Now, an international team of researchers has analyzed the proteins found in 100 species—from bacteria and archaea to plants and humans. It is the largest protein mapping ever to be conducted across different species.

They have learned that these life forms have a number of common characteristics. The study is a collaboration between researchers in Professor Matthias Mann's group in the Novo Nordisk Foundation Center for Protein Research and the Max Planck Institute of Biochemistry. It has been published in the top scientific journal Nature.

"We have mapped the proteins, together called the proteome, of 100 different species. And it is obvious that they are extremely different. At the same time, though, they have more in common than we thought. In all these life forms, a large share of the proteins focus on metabolism and on maintaining a protein balance," says Professor Matthias Mann.

Doubling of Experimentally Confirmed Proteins

Previously, researchers were mainly interested in the DNA of various organisms, for example, how much genetic material humans share with different animals. However, with advancements in the technology used for studying organisms at molecular level, researchers have turned to proteins, the workhorses of the cell.

"A common characteristic of all these life forms is the fact that a high percentage of their proteomes focus on maintaining a sort of balance, what is called as homeostasis. Another common characteristic is the fact that a large share of the proteins help to generate energy. Even though the ways in which this is done differ—from photosynthesis to carbohydrate burning," says Alberto Santos Delgado who during the studies was employed at the Novo Nordisk Foundation Center for Protein Research.

The researchers have used an advanced technology called mass spectrometry to study all 100 species. The technology enabled them to double the number of proteins confirmed experimentally.

Previous research has predicted how many and which proteins exist based only on the genetic code and bioinformatic calculations. However, the new protein mapping has provided actual data on the existence of a very large number of new proteins.

Machine Learning Can Reveal New Correlations

"Our work connecting quantitative mass spectrometry-based proteomics with database resources has resulted in a data set of eight million data points with 53 million interconnections. We made all the data publicly available, enabling other researchers to use it to identify new correlations."
New technologies enabled by machine learning are on the rise and we expect those to benefit from the large and uniform dataset we provide publicly," says Ph.D. Student Johannes Mueller from the Max Planck Institute of Biochemistry.

The researchers at the University of Copenhagen focussed on data processing and bioinformatics analysis, while the researchers at the Max Planck Institute of Biochemistry in Munich focussed on mass spectrometry.

On the website Proteomes of Life, the researchers will make all data from the project publicly available.


---

Provided by University of Copenhagen