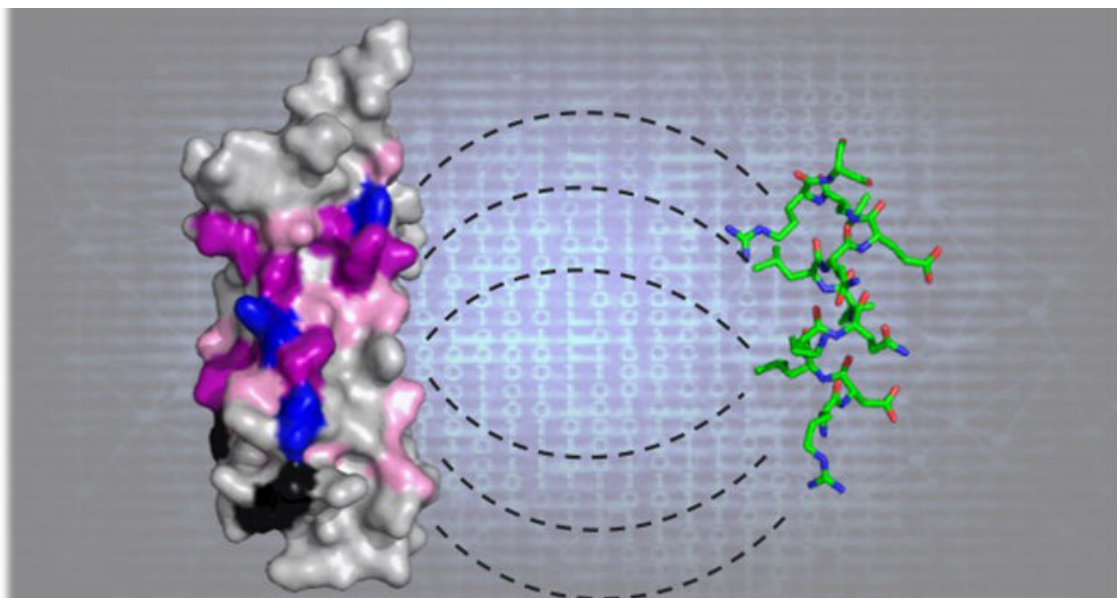


LD Motif Finder locates ancient hidden protein patterns

January 6 2020



The researchers used AI methods to predict the interaction (shown by the dashed lines) of short, linear protein segments (green stick model) with their target proteins (molecular representation on the left). Predictions were experimentally confirmed as shown by the imprint of the ligand on the surface of the target protein (blue, magenta and pink colored surfaces). Credit: King Abdullah University of Science and Technology

An iterative machine learning approach has identified elusive 800 million-year-old amino acid patterns that are responsible for facilitating protein interactions.

Leucine-aspartic acid (LD) [motifs](#) are short amino acid sequences embedded within some proteins to link them to cellular molecules that control cell adhesion, motility and survival. They are known to also play a role in cancer cell spreading and in cardiovascular and infectious diseases. LD motifs were first revealed in 1996 in a family of proteins called paxillin. Only three other LD motif-containing proteins have been discovered since then, and scientists do not know the importance of LD motifs or how many other types of proteins contain them.

KAUST structural biologist Stefan Arold and computational bioscientists Xin Gao and Vladimir Bajic combined the efforts of their teams to develop a [machine learning tool](#) that they called LD Motif Finder (LDMF) to scan through the human proteome and identify LD motif patterns. This was no small task given the tiny number of known LD-motif-containing proteins that could be used to train the tool.

The team "taught" their computational tool using biophysical and structural data from known LD motifs and their proteins. To improve the accuracy of their algorithm, they included a round of experimental testing of its initial predictions and trained the tool to learn from these results.

A final step, performed in collaboration with KAUST colleagues Mariusz and Lukasz Jaremko, involved three-dimensional structural analyses of the association between newly identified LD motifs and known LD motif-binding proteins.

Using this integrative approach, the researchers were able to identify 12 new human proteins that carry functional LD motifs. "This gives us a good idea of how many of these motifs exist within the human proteome," says Arold. "It seems there are far fewer than researchers initially suggested. Of course, this does not mean that they are biologically irrelevant."

The researchers found that these proteins containing LD motifs had functions related to cell adhesion and morphogenesis, suggesting that LD motifs significantly define the proteins' cellular roles. Indeed, the researchers observed alterations in cell adhesion or spreading when fluorescently labeled LD motifs were injected into cultured human cells.

Given that the machine learning tool made it easy to scan whole proteomes, the team also investigated the genomes of mammals, birds, fish, worms, insects and microbes for LD motifs. This large-scale analysis allowed them to conclude that LD motif signaling evolved more than 800 million years ago in unicellular organisms, possibly by co-opting ancestral interaction sequences that label proteins for export out of the nucleus.

"The model, which is freely available online, is highly accurate and sensitive, but there is still room for improvement," says Ph.D. student Meshari Alazmi, first author of the study.

The team hopes to continue developing their model to study the evolution and prevalence of other short protein-[protein](#) interaction motifs across species.

More information: Tanvir Alam et al, Proteome-level assessment of origin, prevalence and function of leucine-aspartic acid (LD) motifs, *Bioinformatics* (2019). [DOI: 10.1093/bioinformatics/btz703](https://doi.org/10.1093/bioinformatics/btz703)

Provided by King Abdullah University of Science and Technology

Citation: LD Motif Finder locates ancient hidden protein patterns (2020, January 6) retrieved 28 April 2024 from <https://phys.org/news/2020-01-ld-motif-finder-ancient-hidden.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.