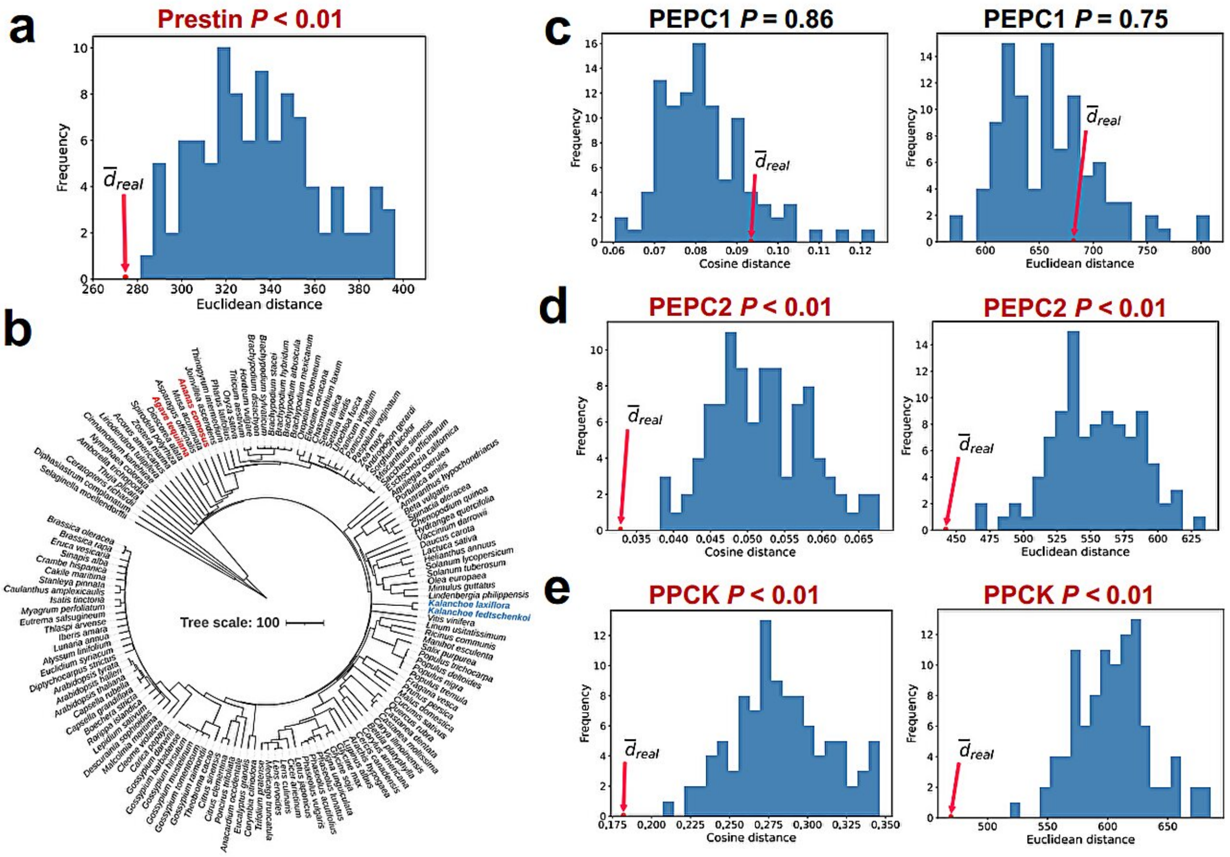
Deep learning tool developed to unravel molecular mechanisms of convergent evolution

September 30 2025, by Li Yali



Previously reported candidate genes underlying adaptive convergent evolution of echolocating mammals and CAM plants showed significance in ACEP tests based on Eg embeddings. Credit: *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2418254122

Convergent evolution—where distinct species independently evolve similar traits or functions, such as the wings of birds and bats—has long fascinated biologists. Now, a research team led by Prof. Zou Zhengting from the Institute of Zoology of the Chinese Academy of Sciences (CAS), has introduced a deep learning–based method to unravel the complex molecular mechanisms driving this phenomenon.

The study was published in the *Proceedings of the National Academy of Sciences* on Sept. 23.

Traditional approaches to studying convergent evolution focus on identifying changes at individual sites in specific proteins, as these offer critical insights into how organisms' genomes adapt to shared environmental pressures. Yet these methods have a key limitation: Protein function relies not just on individual amino acid residues, but on more complex high-order features—such as three-dimensional structures—that current tools fail to account for when assessing molecular-level adaptive convergence.

To address this gap, the team turned to the recently developed pretrained protein language models (PLMs). In several known biological cases, they demonstrated that the "embeddings" generated by PLMs can accurately reflect similarities in proteins' high-order features—even when the proteins show little similarity at the individual site level.

Building on this finding, the team developed an analytical framework called Adaptive Convergence by Embedding of Proteins (ACEP). This tool enables genome-wide detection of adaptive convergent evolution in protein high-order features by leveraging PLM embeddings.

The researchers then applied ACEP to the case of echolocating mammals. Comparing echolocating bats and toothed whales, ACEP identified both known and previously unreported candidate genes, each showing evidence of adaptive convergence in protein high-order features.

Further analysis revealed links between the significance of ACEP's detections and convergent physicochemical properties of proteins, such as net electric charge density—shedding light on potential molecular mechanisms behind the observed convergence.

This study highlights the critical role of protein high-order features as the molecular foundation of functional convergent evolution, which helps organisms adapt to their environments.

More information: Zhenqiu Cao et al, Language models reveal a complex sequence basis for adaptive convergent evolution of protein functions, *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2418254122

Provided by Chinese Academy of Sciences

Citation: Deep learning tool developed to unravel molecular mechanisms of convergent evolution (2025, September 30) retrieved 1 October 2025 from https://phys.org/news/2025-09-deep-tool-unravel-molecular-mechanisms.html