
研究揭示蛋白高阶特征的适应性趋同演化

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/35788.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

研究揭示蛋白高阶特征的适应性趋同演化

趋同演化是生物演化中的重要现象。长期以来，演化生物学研究致力于探索表型趋同背后的分子适应性演化机制。传统研究方法聚焦于蛋白序列中单个氨基酸位点的趋同变化。但是，越来越多的证据表明，即使没有明确的位点趋同，同源蛋白仍可能通过高阶结构或理化特征的趋同演化实现功能相似性。

近日，中国科学院动物研究所邹征廷团队提出ACEP计算分析框架，利用预训练蛋白语言模型，揭示了蛋白高阶特征在功能适应性趋同演化中的重要作用。

通过在海量蛋白序列数据上进行预训练，蛋白语言模型能够捕捉序列中复杂的上下文信息和高阶特征，将蛋白序列转化为包含这些演化特征的高维嵌入向量。研究发现，即使在已知不存在位点趋同的案例中，预训练蛋白语言模型生成的嵌入向量仍能反映蛋白高阶特征的相似性，表明高阶特征趋同是蛋白功能趋同的分子基础。

基于这一发现，研究团队设计了ACEP分析流程，其核心包括三个步骤，即计算目标类群同源蛋白嵌入向量的真实距离，通过模拟中性演化过程构建背景距离分布，基于分布对真实距离进行统计检验，判断是否存在显著的高阶特征趋同信号。同时，ACEP在多个已知案例中得到显著结果。

进一步，全基因组筛选显示，ACEP在蝙蝠与齿鲸中识别出数百个具有高阶特征趋同信号的候选基因，功能富集分析显著关联“感官感知”等条目，涵盖已知回声定位基因和多个新候选基因。部分候选基因还得到正选择检验的支持，增强了其发生适应性趋同的可信度。同时，为提升结果的可解释性，团队分析了嵌入向量趋同背后的具体高阶特征。

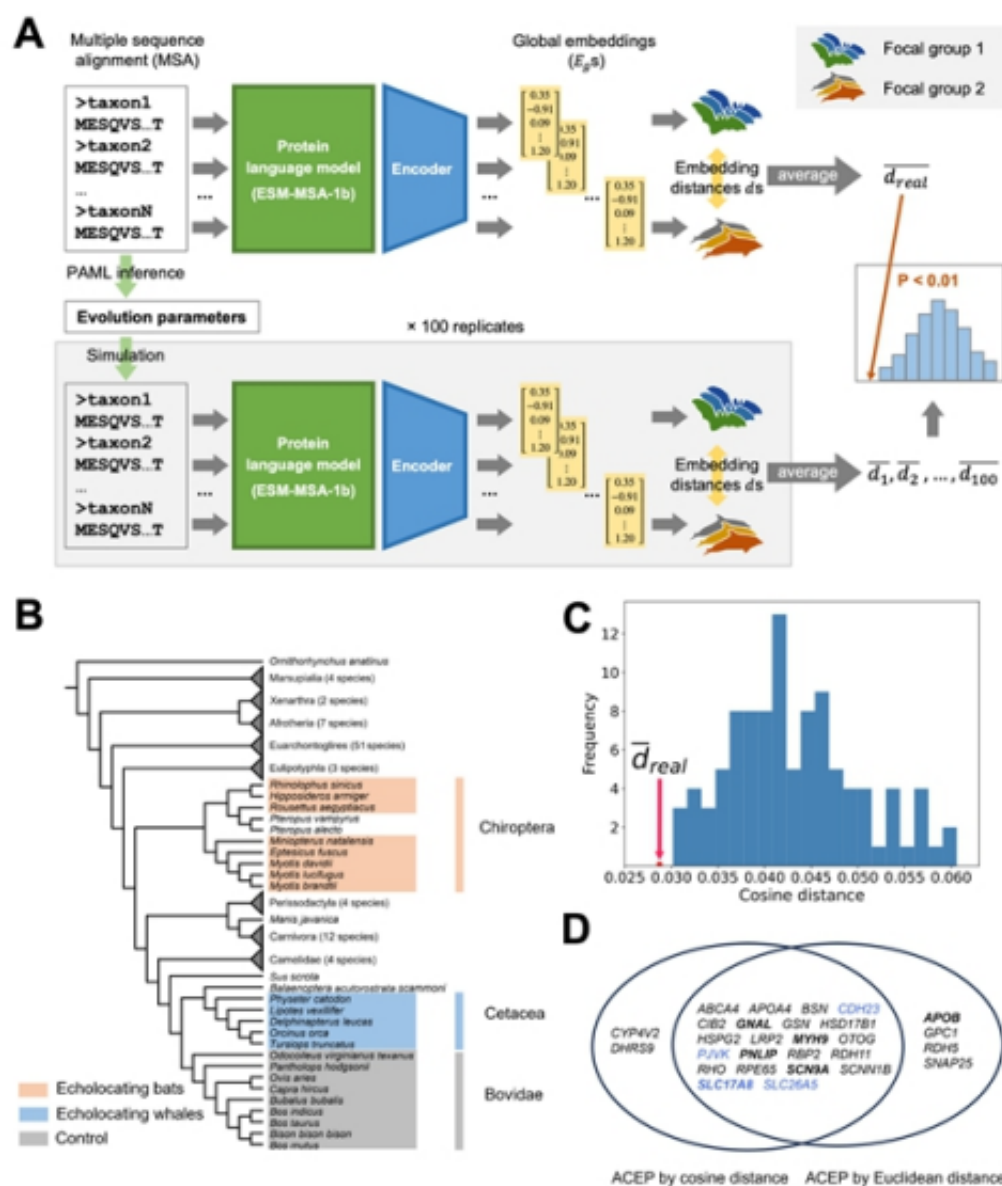
上述研究揭示了蛋白高阶特征趋同适应性演化的机制，展示了人工智能技术在针对复杂的基因型-表型映射进行演化生物学分析时的潜力。

ACEP框架为在全基因组水平系统挖掘基因的复杂适应性趋同模式提供了新工具，代码已通过HuggingFace平台开源。

相关研究成果发表在《美国国家科学院院刊》（PNAS）上。研究工作得到国家自然科学基金和中国科学院战略性先导科技专项等的支持。

[论文链接](#)

代码



ACEP流程示意图以及在回声定位哺乳类中检测出的与“感官感知”相关的候选适应性趋同基因

研究团队单位：动物研究所

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归作者所有，请勿用于商业用途，[爱科学iikx.com](http://iikx.com)转发