
科学家构建RNA结合蛋白的剪接调控作用预测模型

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/2828.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！

科学家构建RNA结合蛋白的剪接调控作用预测模型。基因组研究结果显示，人体内超过90%的基因存在选择性剪接(alternative splicing)。该过程在不同组织以及不同生理阶段受到严格的调控，其失调会导致多种疾病的发生。选择性剪接的体内调控主要由前体mRNA中的顺式元件(cis-elements) 招募反式剪接作用因子(trans-acting splicing factors)来实现。通常情况下，反式剪接作用因子存在一个模块化的构成，其包括一个或多个RNA结合域以及不同的功能模块。但目前科学家对这些功能域的研究还停留在少数几种典型的剪接因子上，如SR蛋白家族和hnRNP蛋白家族，对广大的其他RNA结合蛋白中的功能模块却知之甚少。而深入理解这些功能模块会为科学家进一步研究以至从头合成新型RNA剪接因子提供依据。

11月7日，中国科学院上海营养与健康研究院计算生物学研究所、分子细胞科学卓越创新中心、中科院计算生物学重点实验室RNA系统生物学课题组王泽峰研究团队在国际学术期刊Cell Systems在线发表了题为Modeling and predicting the activities of trans-acting splicing factors with machine learning 的研究长文。在该项工作中，研究人员首次将机器学习方法用于构建RNA结合蛋白的剪接调控作用预测模型，揭示了RNA结合蛋白的序列组成偏好性对其调控作用的影响，为研究RNA结合蛋白的剪接活性提供了重要指导意义，也为人工合成剪接因子提供了可行性。

在之前的研究中，该课题组发现在RNA结合蛋白中存在大量的序列低复杂区域。该研究在此基础上对这些序列低复杂区域在RNA选择性剪接中所扮演的功能进行了系统性研究。研究人员通过构建人工剪接因子的方法检测了多达12种代表性序列低复杂区域在不同RNA位置的剪接活性，发现了这些低复杂区域在RNA选择性剪接中具有位置依赖性(context dependent)且相似的序列组成具有相似的剪接活性。

研究人员进而根据这些功能模块的序列偏好性与其剪接活性，构建了一个以机器学习方法为核心的多肽剪接活性预测模型。利用该机器学习模型，他们还发现了一些此前从未报道过的具备剪接活性的序列特征。并且以此序列特征为基础，他们在世界上首次从头合成带有特定活性的人工剪接因子，取得了极高的成功率(10/11)。该研究的此项发现也为以后发展以人工剪接因子为基础的基因治疗方法扫除了障碍。

该研究工作由中科院-马普学会计算生物学伙伴研究所、华东理工大学以及美国国立环境健康科学研究所(NIEHS)等机构合作完成。该工作在研究员王泽峰的指导下，由华东理工大学联合培养博士生茅缪伟、计算生物所胡玥和杨赞等完成，并得到华东理工大学教授杨弋与NIEHS高级研究员Xiaoling Li的大力支持。

该研究得到国家基金委、上海市科委和国家留学基金委等的支持。

文章链接

RNA结合蛋白剪接调控作用预测模型流程图

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

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