



集成多聚体穿线和高通量实验的大肠杆菌 蛋白质 - 蛋白质相互作用网络及结构的预测

• 巩卫康

• 时间：2021 年 6 月 23 日 上午 10:30-11:30

• 线下：数学科学研究院 316 会议室

• 线上：腾讯会议

会议号：594 961 960



报告内容：

全基因组蛋白质 - 蛋白质相互作用 (PPI) 探测是结构生物学中一个重要但未解决的科学问题。由于高通量实验 (HTE) 在探测 PPI 时通常具有相对较高的假阳性率，同时 PPI 的四级结构预测比用传统结构生物学技术解析三级结构更难预测。我们提出一个计算方法 Threpp 用来解决这两个问题。Threpp 从一对单体序列开始，通过复合物结构库将两个序列联系起来，其中使用朴素贝叶斯分类器模型将比对分数与 HTE 数据相结合，用以预测两条链相互作用的可能性。紧接着，通过界面特定的结构对齐，将单体对齐于二聚体模板重新组装来预测复合物的结构。该方法应用于大肠杆菌 (*E. coli*) 基因组并预测了 35,125 对可信的 PPI，比单独的 HTE 高 4.5 倍。PPI 网络分析显示无标度属性，发现基因组进化的鲁棒性和对大肠杆菌生存至关重要的功能蛋白质。Threpp 构建了基于四级结构穿线对齐预测得到的所有 (*E. coli*) PPI 的复合物结构，其中 6771 个预测的复合物结构具有正确的折叠 (TM-score > 0.5)；尤其是 39 个预测的复合物结构和最近实验结晶出来的结构非常接近 (平均 TM-score = 0.73)。这些结果证明基于多聚体穿线的同源建模在全基因组 PPI 网络探测和复合物结构预测中有重要意义。

报告人简介：

巩卫康，北京工业大学环境与生命学部在读博士，导师为李春华教授，主要研究方向为生物物理、计算生物学及生物信息学。目前在新加坡南洋理工大学做交流访学。