

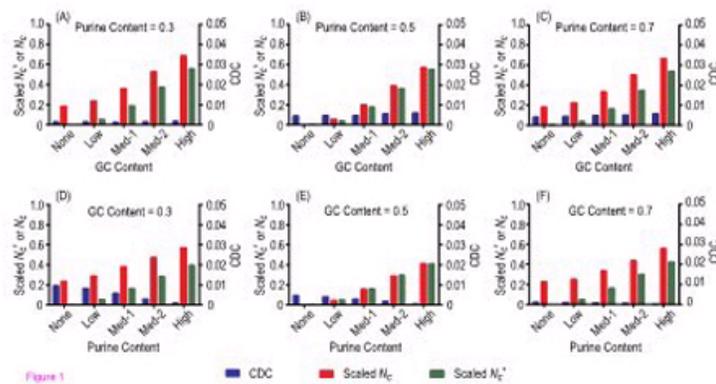


基因组所设计开发出检测密码子使用偏好的新算法

文章来源：北京基因组研究所

发布时间：2012-04-05

【字号：小 中 大】



密码子使用偏好的不同算法比较

近日，中国科学院北京基因组研究所基因组科学与信息重点实验室“百人计划”章张研究员带领其团队，成功设计开发出检测密码子使用偏好（Codon Usage Bias，简称CUB）的新算法：密码子偏差系数模型（Codon Deviation Coefficient，简称CDC）。该研究成果发表在*BMC Bioinformatics*杂志上。

此项工作原创性地将概率论中的交、并、补操作应用到组分分析，用GC含量（S）和嘌呤含量（R）来表示四个核苷酸组分，并在此基础上推导出密码子和氨基酸的组分，从而设计出基于S和R的组分模型，应用该模型考察基因的CUB，进而提出了CDC算法。不同于现有的CAI、ENC等相关算法，CDC通过GC含量和嘌呤含量考虑了不同序列的背景组分特异性，独创性地运用自展重抽样法（Bootstrap Resampling）检测CUB的显著性，且不需要高表达基因作为先验信息。

经验证，CDC在模拟数据中优于现有的多个相关算法，在真实数据中CDC与基因表达含量的关联系数（Correlation Coefficient）高于其它算法，并且在大肠杆菌中发现CUB的显著性与基因功能有着紧密联系。

该项成果的发布，使科研工作者能更准确快速地分析CUB，进而更深入地研究在自然选择压力下的基因突变、基因表达，蛋白质功能等的进化。

[论文链接](#)

打印本页

关闭本页