



面向世界科技前沿, 面向国家重大需求, 面向国民经济主战场, 率先实现科学技术跨越发展,
率先建成国家创新人才高地, 率先建成国家高水平科技智库, 率先建设国际一流科研机构。

——中国科学院办院方针



官方微博



官方微信

首页 组织机构 科学研究 人才教育 学部与院士 资源条件 科学普及 党建与创新文化 信息公开 专题

搜索

首页 > 科研进展

北京基因组所开发长非编码RNA数据库LncBook

文章来源: 北京基因组研究所 发布时间: 2018-10-26 【字号: 小 中 大】

我要分享

近日, 由中国科学院北京基因组研究所开发的人类长非编码RNA (long non-coding RNA, lncRNA) 数据库LncBook正式上线。该项研究成果以*LncBook: a curated knowledgebase of human long non-coding RNAs* 为题在国际学术期刊《核酸研究》(*Nucleic Acids Research*) 在线发表。

近年来, 有关lncRNA的研究一直是国际热点, 研究表明, lncRNA在多种生物过程中发挥了重要作用并与疾病的发生密切相关, 但lncRNA的注释信息及质量仍远远落后于蛋白编码基因。LncBook数据库不仅提供了丰富的高质量人类lncRNA数据集, 还进行了大规模的多组学数据分析, 以及系统的功能与疾病注释, 为功能实验研究及生物信息学分析提供了大量可用信息与数据。

基于严格的审编标准, LncBook整合已有lncRNA数据并鉴定新的lncRNA, 共获得270, 044个lncRNA转录本。在此基础上, LncBook在lncRNA表达、甲基化、变异、miRNA-lncRNA相互作用这些多组学层面进行大规模深层次的数据分析。在表达层面, 绘制lncRNA在32或者53种正常人类组织中的表达图谱, 并鉴定出49, 115个高度组织特异 (tissue-specific) 和819个持家 (housekeeping) lncRNA; 在甲基化层面, 构建lncRNA在9种癌症中正常与癌症状态下的promoter及body区的甲基化图谱; 在变异层面, 基于dbSNP数据库SNP位点注释lncRNA区域92, 725, 757个SNP最小等位频率 (基于千人基因组数据)、ClinVar与COSMIC疾病关联信息; 预测了128, 392, 451个lncRNA-miRNA的相互作用条目。上述结果以图或表的形式展示在LncBook数据库中, 相关信息均可免费下载。基于以上数据, LncBook还预测了97, 998个潜在的疾病关联lncRNA。此外, 在LncRNAWiki的基础上LncBook对1, 867个文献报道lncRNA进行了系统的功能与疾病信息注释。

LncBook作为重要的lncRNA资源库, 提供了目前为止数据量最为丰富的人类lncRNA数据。作为LncRNAWiki数据库的补充, LncBook具备友好的查询、浏览与可视化功能。用户可通过ID / symbol、功能、疾病名称等检索lncRNA信息, 浏览指定lncRNA的多组学信息, 并通过ftp下载所有相关注释信息与分析结果。此外, LncBook还提供了可用于lncRNA序列比对、分类、编码能力预测等研究的工具, 方便在线分析。

该研究与沙特阿卜杜拉国王科技大学 (King Abdullah University of Science & Technology, KAUST) 教授Vladimir Bajic合作开展。研究得到中科院战略性先导科技专项、中科院国际伙伴计划、中科院“十三五”信息化专项等的资助。

[论文链接](#)

[数据库链接](#)

(责任编辑: 叶瑞优)



© 1996 - 2018 中国科学院 版权所有 京ICP备05002857号 京公网安备110402500047号 联系我们
地址: 北京市三里河路52号 邮编: 100864

热点新闻

中科院党组传达学习贯彻中央经...

中科院党组2018年冬季扩大会议召开
中科院与大连市举行科技合作座谈
中科院老科协工作交流会暨30周年总结表...
白春礼: 中国科学院改革开放四十年
《改革开放先锋 创新发展引擎——中国科...

视频推荐



【新闻联播】“率先行动”计划 领跑科技体制改革



【新闻联播】三北防护林工程区生态环境明显改善

专题推荐

