基因组学研究中一些常用软件的概述 An Introduction of Several Programs Used in Genomic Analysis

吴清发 WU Qing-Fa

北京华大基因研究中心，北京 101300；　中国科学院遗传与发育生物学研究所人类基因组中心，北京 100101 Beijing Genomic Institute,Beijing,101300;Human Genome Center,Institute of Genetics and Developmental Biology,Chinese Academy of Sciences,Beijing 100101

摘要　基因组学是以一个物种的全部遗传信息为研究对象，在整体上研究遗传信息的分子组成、组织结构、表达调控和进化等内在机制的基础性学科。基因组学研究中海量数据的存储、管理和检索，以及对这些数据进行挖掘等过程，必须借助于生物信息学的方法。目前，大量成熟的软件广泛地应用在基因组学研究中，它们大都可通过互联网免费访问或索取。本文拟对人类基因组计划中常用的一些软件如序列比对、序列组装、重复序列鉴定和基因预测等软件的原理作一介绍，并结合典型软件加以说明。

Abstract:Genomics is a novel subject that has been developed accompanying with the progress of human genome project.Genomics deals with the chemistry component,structure organization and evolution of genome at global level.As genomics associated with huge data,bioinformatics plays an important role in these processes of data production,data management and data mining.At present,many reliable programs have been used in genomic research successfully,which are usually accessible and downloaded freely.We address here the principles of some programs used wildly in genomics such as sequence alignment,sequence assembly,repeat identification and gene prediction,which are exemplified with typical programs respectively.

Abstract

Key words

DOI:

通讯作者