

<<生物信息学算法导论>>

图书基本信息

书名：<<生物信息学算法导论>>

13位ISBN编号：9787122001696

10位ISBN编号：7122001695

出版时间：2007-7

出版时间：第1版 (2007年7月1日)

作者：N.C.琼斯

页数：345

字数：411000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<生物信息学算法导论>>

### 内容概要

这是一本关于生物信息学算法和计算思想的导论性教科书，原著由国际上的权威学者撰写，经国内知名专家精心翻译为中文，系统介绍推动生物信息学不断进步的算法原理。

全书强调的是算法中思想的运用，而不是对表面上并不相关的各类问题进行简单的堆砌。

体现了以下特色：阐述生物学中的相关问题，涉及对问题的模型化处理并提供一种或多种解决方案；简要介绍生物信息学领域领军人物；饶有趣味的小插图使得概念更加具体和形象，方法更容易被领会，激励学生学习的兴趣并鼓励他们加入到生物信息学研究工作中来。

书中的大量论述表明：较少的几种设计思想就能解决大量的生物学难题。

目前，可供本科学生使用的生物信息学著作为数不多，本书恰恰是其中的一本。

国内生物信息学，计算生物学、计算数学等领域的本科生、研究生和其他研究人员，会从书中汲取基本的算法原理、解决实际问题的方法和技巧，进而更好地从事相关研究工作。

## &lt;&lt;生物信息学算法导论&gt;&gt;

## 书籍目录

## 1 绪论2 算法与复杂性 2.1 算法是什么？

2.2 生物学算法与计算机算法 2.3 找钱问题 2.4 正确的与错误的算法 2.5 递归算法 2.6 迭代算法与递归算法的比较 2.7 快速算法与慢速算法的比较 2.8 大O记号 2.9 算法设计技术 2.10 易处理与不易处理问题的比较 2.11 附注 人物天地：Richard Karp 2.12 问题3 分子生物学简介 3.1 生命是由什么组成的？

3.2 什么是遗传物质？

3.3 基因是干什么的？

3.4 哪些分子编码基因？

3.5 DNA的结构是怎样的？

3.6 在DNA和蛋白质间传递信息的物质是什么？

3.7 蛋白质是由什么组成的？

3.8 我们该如何去分析DNA？

3.9 一个物种的个体差异是怎样产生的？

3.10 不同物种间有怎样的差异？

3.11 为什么要搞生物信息学？

人物天地：Russell F.Doolittle4 穷举搜索 4.1 限制酶切作图 4.2 不实用的限制酶切作图算法 4.3 一个实用的限制酶切作图算法 4.4 DNA序列上的调控基序 4.5 序列剖面 4.6 基序发现问题 4.7 检索树 4.8 发现基序 4.9 发现一个中间字符串 4.10 附注 人物天地：Gary Stormo 4.11 问题5 贪婪算法

5.1 基因组重排 5.2 反序排序法 5.3 近似算法 5.4 断点：贪婪的另一面 5.5 贪婪方法与基序发现

5.6 附注 人物天地：David Sankoff 5.7 问题6 动态规划算法 6.1 DNA序列比较的力量 6.2 找钱问题重述 6.3 曼哈顿游客问题 6.4 编辑距离与联配 6.5 最长共同子序列 6.6 全局序列联配 6.7 得分联配 6.8 局部序列联配 6.9 缺口罚分联配 6.10 多重联配 6.11 基因预测 6.12 基因预测的统计方法

6.13 基于相似性的基因预测方法 6.14 剪接联配 6.15 附注 人物天地：Michael Waterman 6.16 问题7 分而治之算法 7.1 排序问题的分治法 7.2 空间效率高的序列联配 7.3 模序联配和四个俄罗斯人的加速法 7.4 在亚二次时间内构建联配 7.5 附注 人物天地：Webb Miller 7.6 问题8 图算法 8.1 图

8.2 图与遗传学 8.3 DNA测序 8.4 最短超字符串问题 8.5 作为可选择测序技术的DNA阵列 8.6 杂交测序 8.7 SBH与Hamilton路问题 8.8 SBH与欧拉路问题 8.9 DNA测序中的片段装配 8.10 蛋白质测序和鉴定 8.11 肽测序问题 8.12 谱图 8.13 基于数据库搜索的蛋白质鉴定 8.14 谱的卷积 8.15 谱联配 8.16 附注 8.17 问题9 组合模式匹配 9.1 重复序列发现 9.2 哈希表 9.3 精确模式匹配 9.4 关键词树 9.5 后缀树 9.6 启发式相似性搜索算法 9.7 近似模式匹配 9.8 BLAST：依靠数据库的序列比较 9.9 附注 人物天地：Gene Myers 9.10 问题10 聚类和树 10.1 基因表达分析 10.2 系统聚类 10.3 k-均值聚类 10.4 聚类和有瑕团 10.5 进化树 10.6 基于距离的树重构 10.7 由可加矩阵重构树 10.8 进化树与系统聚类 10.9 基于字符的树重构 10.10 小简约问题 10.11 大简约问题 10.12 附注 人物天地：Ron Shamir 10.13 问题11 隐马氏模型 11.1 CG岛和“公平赌场” 11.2 公平赌场和隐马氏模型

11.3 解码算法 11.4 隐马氏模型参数估计 11.5 剖面隐马氏模型联配 11.6 附注 人物天地：David Haussler 11.7 问题12 随机化算法 12.1 排序问题回顾 12.2 吉布斯抽样 12.3 随机投影 12.4 附注

12.5 问题参考文献索引

## &lt;&lt;生物信息学算法导论&gt;&gt;

## 媒体关注与评论

前言 20世纪90年代初，当我们中的一位第一次讲生物信息学课时，他还不能确定会有多少学生来听。

虽然那时Smith?Waterman算法和BLAST算法已经被开发出来了，但在生物学家中还不像现在这样广为流传，甚至“bioinformatics（生物信息学）”一词还没有被创造出来。

那时，大多数人认为DNA阵列只不过是尚未确定实际应用的智力玩具，只有少数狂热分子认为这是一项拥有巨大潜力的技术。

为数不多的生物信息学专家开始为尚不存在的数据集发展新的算法思想：DavidSankoff设立了一项基金，专门鼓励在缺乏基因次序资料时所作的基因组重排研究；MichaelWaterman和GaryStormo在几乎没有启动子样本可用时发展了基序（或称为模体，motif）查找算法；在还没有装配好细菌基因组时，GeneMyers就研发了一个精致的片段装配工具；当长为172282个核苷酸的Epstein?Barr病毒还是GenBank中最长的序列时，WebbMiller已在梦想比对十亿核苷酸长度的DNA序列了。

GenBank也恰好刚刚完成了自身的转变，从一系列（纸张！

）装订的册子转变为记录在磁带上的电子数据库，从而可以发送给世界各地的科学家们了。

我们必须回到20世纪80年代中期和90年代早期去感谢在这十年中生物学所发生的变革。

然而，生物信息学比生物学的影响更大——它对计算科学也产生了深远的影响。

生物学已快速成为新算法和统计问题的一大数据源，而且应用在生物学上的算法之多也超越了其他任何一种基础科学。

将计算机科学和生物学联系起来具有重要的教育意义，它不仅改变了我们教生物学家计算思想的方法，也教给了计算机科学家们懂得如何应用算法。

过去，只有计算机科学家们才学习计算机科学这门课程，而且仅局限于来自其他学科的少数学生。

在20世纪90年代早期，几乎不可能（虽然欢迎）看到生物学专业学生去上算法课。

但是现在情况改变了，许多生物专业的学生都了解了一些浅显的算法知识。

同时，好奇的计算机科学专业的学生也会学习一些遗传学和生物信息学的基础知识。

虽然这样做的学生相对来说还是少数，但要知道在20世纪90年代初期几乎还没有生物信息学课程，而现在看来数量已经不少了。

我们预见，大学本科生的生物信息学课程将成为所有一流大学的固定课程。

这是必然的趋势，而不是空想。

这是一本关于生物信息学算法及其计算思想的导论性教科书，近20年来，正是这些计算思想推动着生物信息学算法的发展。

还有许多重要的概率和统计技术以及当前生物信息学专家们正在研究的重要课题尚未涉及到。

我们故意不包括计算生物学的所有领域，例如，像蛋白质折叠这样重要的课题甚至也没讨论。

最早的生物信息学教科书是由Waterman在1995年写的 [ 108 ]，此书极好地介绍了DNA统计；

而Gusfield在1997年写的书 [ 44 ] 可以说是字符串算法的大全。

Durbin等在1998年写的书 [ 31 ] 和Baldi与Brunak在1997年写的书 [ 7 ] 则着重于隐马氏模型和机器学习技术；

Baxevanis和Ouellette在1998年写的书 [ 10 ] 是一本优秀的生物信息学实用指南；Mount在2001年写的书 [ 76 ] 很好地描述了生物学问题与生物信息学技术之间的联系；

Bourne和Weissig在2002年写的书 [ 15 ] 则将重点放在蛋白质生物信息学方面。

网络上也有许多生物信息学课程笔记，我们从互联网（WWW）上由SerafimBatzoglou、DickKarp

、RonShamir、MartinTompa等人提供的材料学到了许多有关生物信息学的教学法。

<<生物信息学算法导论>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>