

从碳到硅 ——

生物信息学

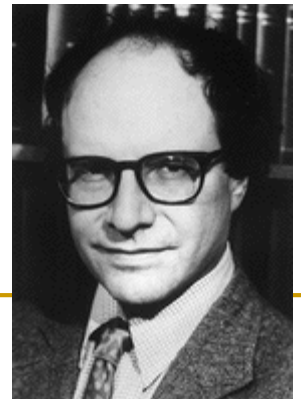
Bioinformatics

—— 生命科学的新纪元*

生物学研究模式的转变

基于全部基因都将知晓并以电子技术可操作的方式驻留在数据库中，新的生物学研究模式的出发点应该是理论的。科学家将从理论推测出发，再返回到实验中去，追踪或验证这些理论假设。.....生物学家不仅必须成为计算机学者，而且也要改变他们研究生命现象的途径。

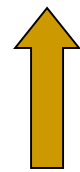
--- Gilbert (Nature, 1991)



Bioinformatics (v3) : Foreword (2005)

■ 生物学正在经历重大转变：

基因组信息的全面发掘，包括序列测序、大分子结构预测、功能注释以及调控网络的阐明，促使了“系统生物学”概念的出现。



生物信息学

其中数学、统计学、计算机科学具有重要地位（Central role*）。

第一章 绪论

- 生物信息学的定义和内容
- 生物信息学的产生和发展
- 生物信息学的应用和产业
- 生物信息学的教育和学习

第一章 绪论

- 生物信息学的定义和内容
- 生物信息学的产生和发展
- 生物信息学的应用和产业
- 生物信息学的教育和学习

什么是生物信息学？

生物信息学就是利用信息技术对生物信息进行获取、储存、分发（查询）和分析，以解释这些信息数据所蕴涵的生物学意义的学科。

姊妹学科——

计算生物学（Computational biology）

生命的信息内涵

什么是信息？

$$\text{Log}_2 2 = 1 \text{ bit}$$

■ 信息的产生：多种可能状态下进行**选择**。

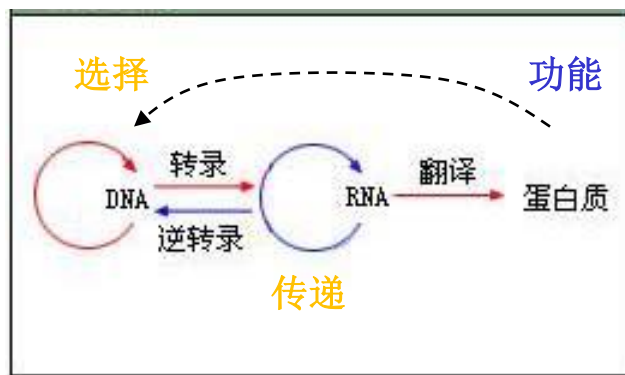
——生物信息的产生源于自然选择。

ATCGC ~ ? bits

■ 信息的“行为方式”——**传递**：一种选择引起另一种选择；**功能实现**。

特异性+功能性

中心法则：

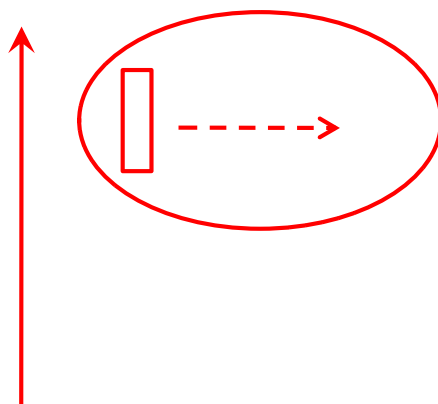


学科间的本质性联系

- For biology, **chemistry** is much more than just a method: biology is rooted in, and is firmly based on, the laws of **chemistry**.
- For biology, **informatics** is much more than just a method: biology is rooted in, and is firmly based on, the laws of **informatics**.

生物学中的信息内容

- 分子生物学：信息在生物体内的传递；
- 遗传学与进化生物学：
—— 信息在代间的传递和产生。



什么是生物信息学？

生物信息学就是利用 信息技术 对 **生物信息** 进行获取、储存、分发（查询）和分析，以解释这些信息数据所蕴涵的 **生物学意义** 的学科。

信息技术的应用

- 由于长期进化，生物信息如此丰富，其传递方式如此复杂，以至我们需要借助专门储存和分析它们的技术和工具——主要涉及数学和计算机科学。

什么是生物信息学？

生物信息学就是利用**信息技术**对**生物信息**进行获取、储存、分发（查询）和分析，以解释这些信息数据所蕴涵的**生物学意义**的学科。

——据1995年美国人类基因组计划第一个五年总结报告中的提法，略加概括而成。

参考定义： Bioinformatics is

the science that uses the instruments of informatics to analyze biological data in order to formulate hypotheses about life.

---- *Introduction to Bioinformatics. 2007.*

the application of computing and mathematics to the management, analysis, and understanding of data to solve biological questions.

---- *Instant Notes in Bioinformatics (2nd Edition). 2010.*

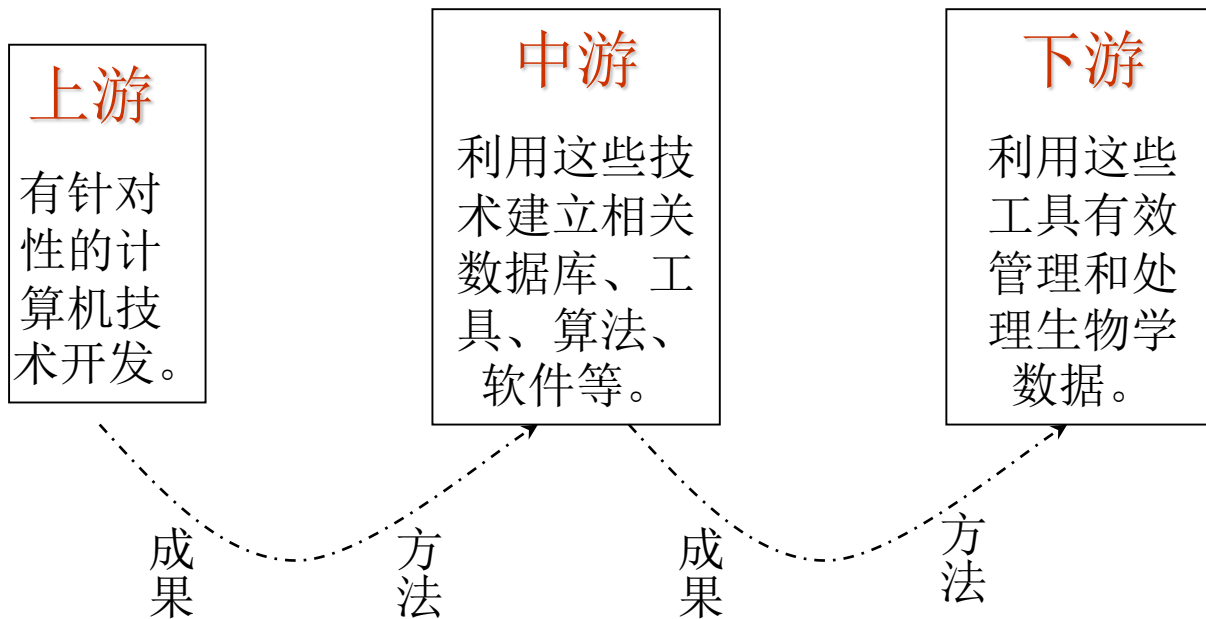
an interdisciplinary field that, by applying computer science, mathematics and statistics, develops methods and software tools for understanding biological data.

---- *Bioinformatics and Computational Biology. 2017.*

将计算机与信息科学技术应用到生命科学，尤其是分子生物学*中的交叉学科。

—— *生物信息学，陈铭，2022。*

研究内容---从方法（信息技术）角度：



研究内容 --- 从对象（生物信息）角度：

■ 目前主要包括：

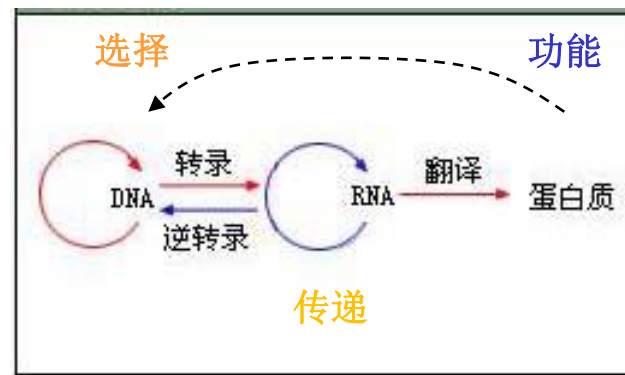
macromolecular sequences; （最初）

macromolecular structures;

expression profiles; (RNA-seq; 2D-PAGE)

biochemical network; (Interactions and reactions)

evolution history.



第一章 绪论

- 生物信息学的定义和内容
- 生物信息学的产生和发展
- 生物信息学的应用和产业
- 生物信息学的教育和学习

生物信息学的产生 之 孕育

- 1950s 末，计算机应用于生物科学研究中。

主要是利用数学模型、统计学方法和计算机处理宏观生物学数据。如 Numerical Taxonomy, Mathematical Ecology。
1960s末-70s 开始应用于分子水平：蛋白质数据库、序列比对算法.....70s末：Sanger测序法发明——

- 1980s，计算机应用于分子水平成为主要方向。

其中包括建立分子生物学数据库（如Genbank）以及蛋白质结构的计算机辅助分析与预测等。

——在上述研究中，人们已经逐步建立了理论基础和一批方法、模型、软件和数据库。

生物信息学的产生 之 促成（1）

■ 生物信息的全面发掘——巨大数据量的产生。

1990s后，以人类基因组计划（Human genome project, HGP）带动的数据爆炸：DNA sequencing, microarray, EST, 2D-PAGE, protein structure determination, protein interactions —— high-throughput technique.

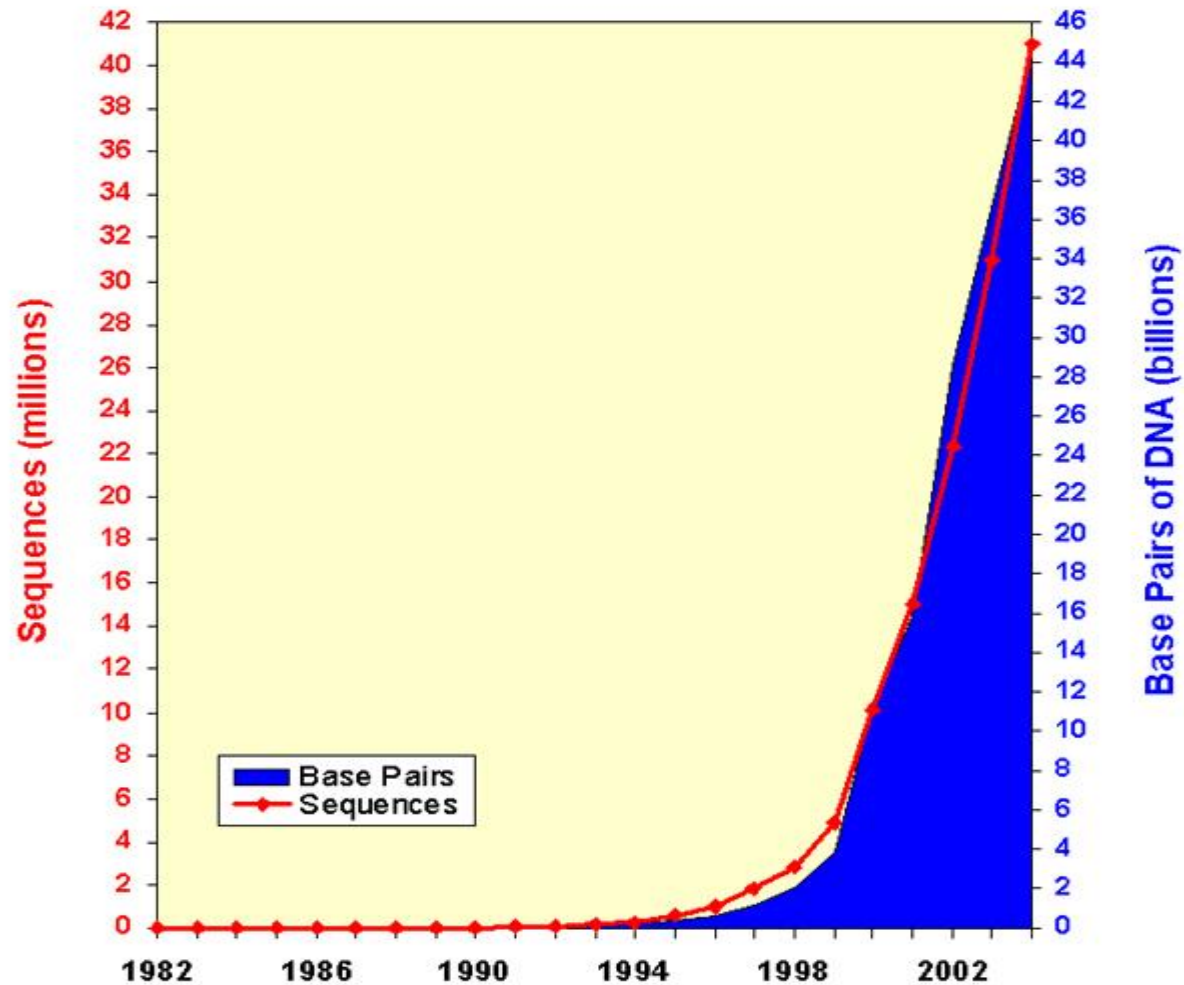
HGP：1990～2000年，10年时间实现了“工作草图”，2003年实现了“完成图”， 3×10^9 个碱基对（30亿美元），并对约30,000个基因进行了注释。越来越多的其他模式生物也完成了全基因组测序工作。

How many characters are in the “Heaven Book”?

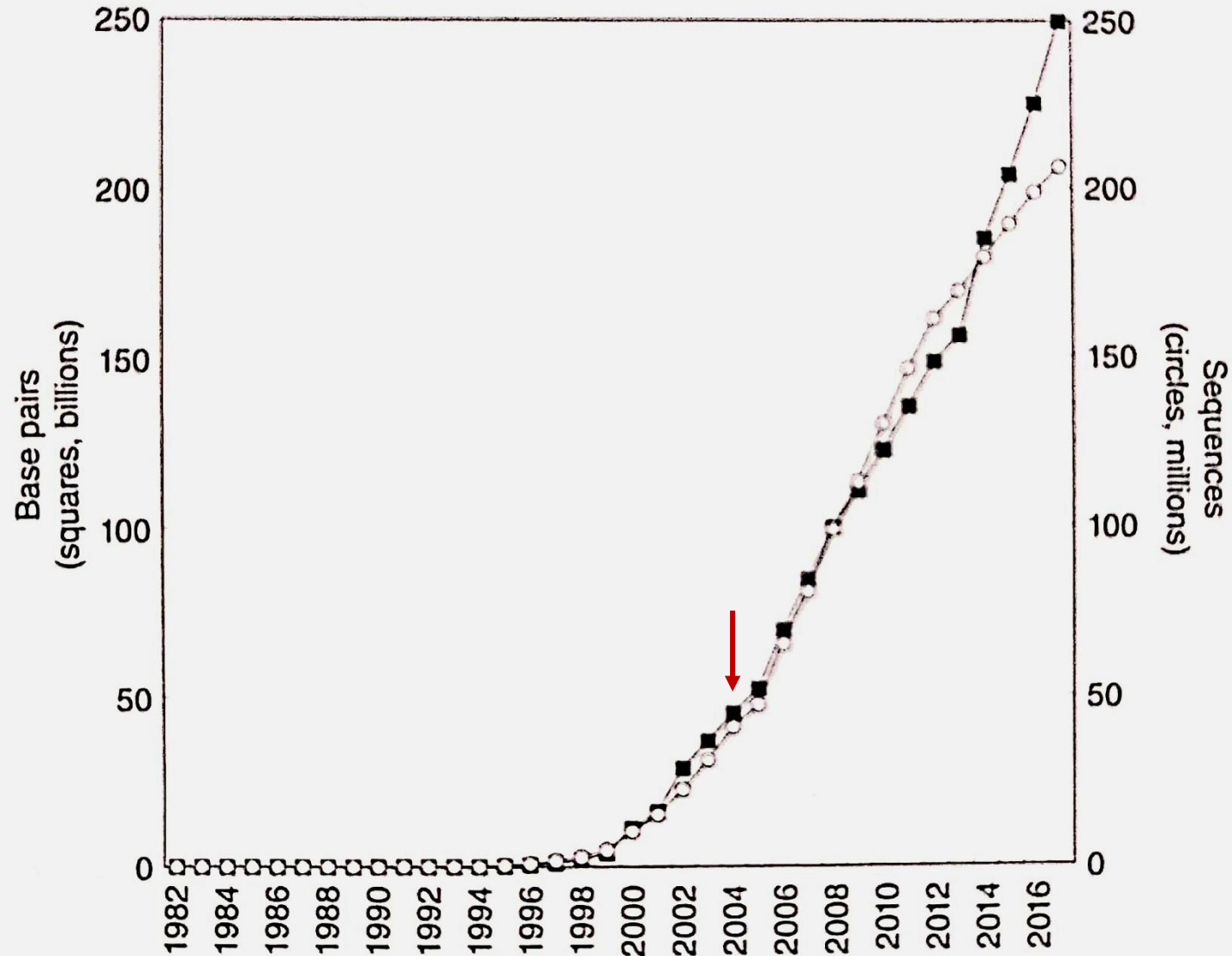
$3 \times 10^9 = 3,000 \text{ characters} \times 100 \text{ pages} \times 10,000 \text{ books}$

CCGGTCTCCCCGCCCGCGCGCAAGTAAAGGCCAGCGCAGCCCCGCGCTCCTGCCCTGGGGCCTCGTCTTTCTCCA
GGAAAACGTGGACCGCTCTCCGCCGACAGTCTCTTCCACAGACCCCTGTGCGCTTCGCCCCCGGTCTCTTCCGGTT
CTGTCTTTTCGCTGGCTCGATACGAACAAGGAAGTCGCCCCAGCGAGCCCCGGCTCCCCAGGCAGAGGGCGGCC
CGGGGGCGGAGTCAACGGCGGAGGCACGCCCTCTGTGAAAGGGCGGGGCATGCAAATTCGAAATGAAAGCCCGGG
AACGCCGAAGAAGCACGGGTGTAAGATTTCCCTTTTCAAAGGCGGGAGAATAAGAAATCAGCCCGAGAGTGTAAGGG
CGTCAATAGCGCTGTGGACGAGACAGAGGGAATGGGGCAAGGAGCGAGGCTGGGGCTCTCACCGCGACTTGAATGT
GGATGAGAGTGGGACGGTGACGGCGGGCGCGAAGGCGAGCGCATCGTTCTCGGCCCTTTGGCTAAGATCAAGTGT
AGTATCTGTTCTTATCAGTTTAAATATCTGATACGTCCTCTATCCGAGGACAATATATTTAAATGGATTGATCAATCCGCTT
CAGCCTCCCGAGTAGCTGGGACTACAGACGGTGCCATCACGCCAGCTCATTGTTGATTCCCGCCCCCTTGGTAGAG
ACGGGATTCCGCTATATTGCCTGGGCTGGTGTGCAACTCATAGAACAAGGATCCTCCCTCCTGGGCCTGGGCGTGG
GCTCGCAAAACGCTGGGATTCCCGGATTACAGGCGGGCGCACCACACCAGGAGCAAACACTTCCGGTTTTAAAAATT
CAGTTTGTGATTGGCTGTCATTAGTATTATGCTAATTAAGCATGCCCGGTTTTAAACCTCTTAAAACAACTTTTTAAAT
ACCTTTCCACCTAAAACGTTAAAAATTTGCAAGTGATAATATTCGACAAGCTGTTATTGCCAACTATTTTCTATTGTT
TCCTAATGGCATCGGAAGTAGCGAAAGTTTCTCGCCATCAGTTAAAAAGTTTTCGGCGAGATGTAGACCTAGCAGAGGTG
TGCGAGGAGGCCGTTAAGACTATACTTTCAGGGATCATTCTATAGTGTGTTACTAGAGAAGTTTCTCTGAACGTGTAG
AGCACCGAAAACCACGAGGAAGAGAGGTAGCGTTTTTCATCGGGTTACCTAAGTGCAGTGTCCCCCTGGCGCGCAAT
TGGGAACCCACACGCGGTGTAGAAATATATTTTAAGGGCGCGGGTAGCGTTTTTCATCGGGTTACCTAAGTGCAGTGT
CCCCCTGGCGCCCCGGTCTCCCCGCCCGCGCGCAAGTAAAGGCCAGCGCAGCCCCGCGCTCCTGCCCTGGGGCC
TCGTCTTTTCTCCAGAAAACGTGGACCGCTCTCCGCCGACAGTCTCTCCACAGACCCCTGTGCGCTTCGCCCCCG
GTCTCTCCGGTTCTGTCTTTTCGCTGGCTCGATACGAACAAGGAAGTCGCCCCAGCGAGCCCCGGCTCCCCCAGG
CAGAGGCGGCCCGGGGGCGGAGTCAACGGCGGAGGCACGCCCTCTGTGAAAGGGCGGGGCATGCAAATTCGAAA
TGAAAGCCCGGGAACGCCGAAGAAGCACGGGTGTAAGATTTCCCTTTTCAAAGGCGGGAGAATAAGAAATCAGCCCG
AGAGTGTAAGGGCGTCAATAGCGCTGTGGACGAGACAGAGGGAATGGGGCAAGGAGCGAGGCTGGGGCTCTCACC
GCGACTTGAATGTGGATGAGAGTGGGACGGTGACGGCGGGCGCGAAGGCGAGCGCATCGTTCTCGGCCCTTTTGGC
TAAGATCAAGTGTAGTATCTGTTCTTATCAGTTTAAATATCTGATACGTCCTCTATCCGAGGACAATATATTTAAATGGATT
GATCAATCCGCTTCAGCCCTCCCGAGTAGCTGGGACTACAGACGGTGCCATCACGCCAGCTCATTGTTGATTCCCGC
CCCTTGGTAGAGACGGGATTCCGCTATATTGCCTGGGCTGGTGTGCAACTCATAGAACAAGGATCCTCCCTCCTG
GGCCTGGGCGTGGGCTCGCAAAACGCTGGGATTCCCGGATTACAGGCGGGCGCACCACACCAGGAGCAAACACTTC
CGGTTTTAAAAATTCAGTTTGTGATTGGCTGTCATTAGTATTATGCTAATTAAGCATGCCCGGTTTTAAACCTCTTAA
ACAACTTTTAAAAATTACCTTTCCACCTAAAACGTTAAAAATTTGCAAGTGATAATATTCGACAAGCTGTTATTGCCAACT
ATTTTCTATTTGTTTCTAATGGCATCGGAAGTAAAGTTTCTCGCCATCAGTTAAAAAGTTTTCGGCGAGATGTAG
ACCTAGCAGAGTGTGCGAGGAGGCCGTTAAGACTATACTTTACGGGATCATTCTATAGTGTGTTACTAGAGAAGTTT
CTCTGAACGTGTAGAGCACCGAAACACGAGGAAGAGAGGTAGCGTTTTATCGGGTTACCTAAGTGCAGTGTCC
CCCTGGCGCGCAATTGGGAACCCACACGCGGTGTAGAAATATATTTTAAGGGCGCGTATTTGTTTCTAATGGCATC
GGAAGTAGCGAAAGTTTCTCGCCATCAGTTAAAAAGTTTTCGGCGAGATGTAGACCTAGCAGAGGTGTGCGAGGAGGCC
GTTAAGACTATACTTTACAGGATCATTCTATAGTGTGTTACTAGAGAAGTTTCTCTGAACGTGTAGAGCACCGAAAAC
CACGAGGAAGAGAGGTAGCGTTTTTCATCGGGTTACCTAAGTGCAGTGTCCCCCTGGCGCGCAATTGGGAACCCAC
ACGCGGTGTAGAAATATATTTTAAGGGCGCGTATTTGTTTCTAATGGCATCGGAAGTAGCGAAAGTTTCTCGCCATCA
GTTAAAAAGTTTTCGGCGAGATGTAGACCTAGCAGAGGTGTGCGAGGAGGCCGTTA

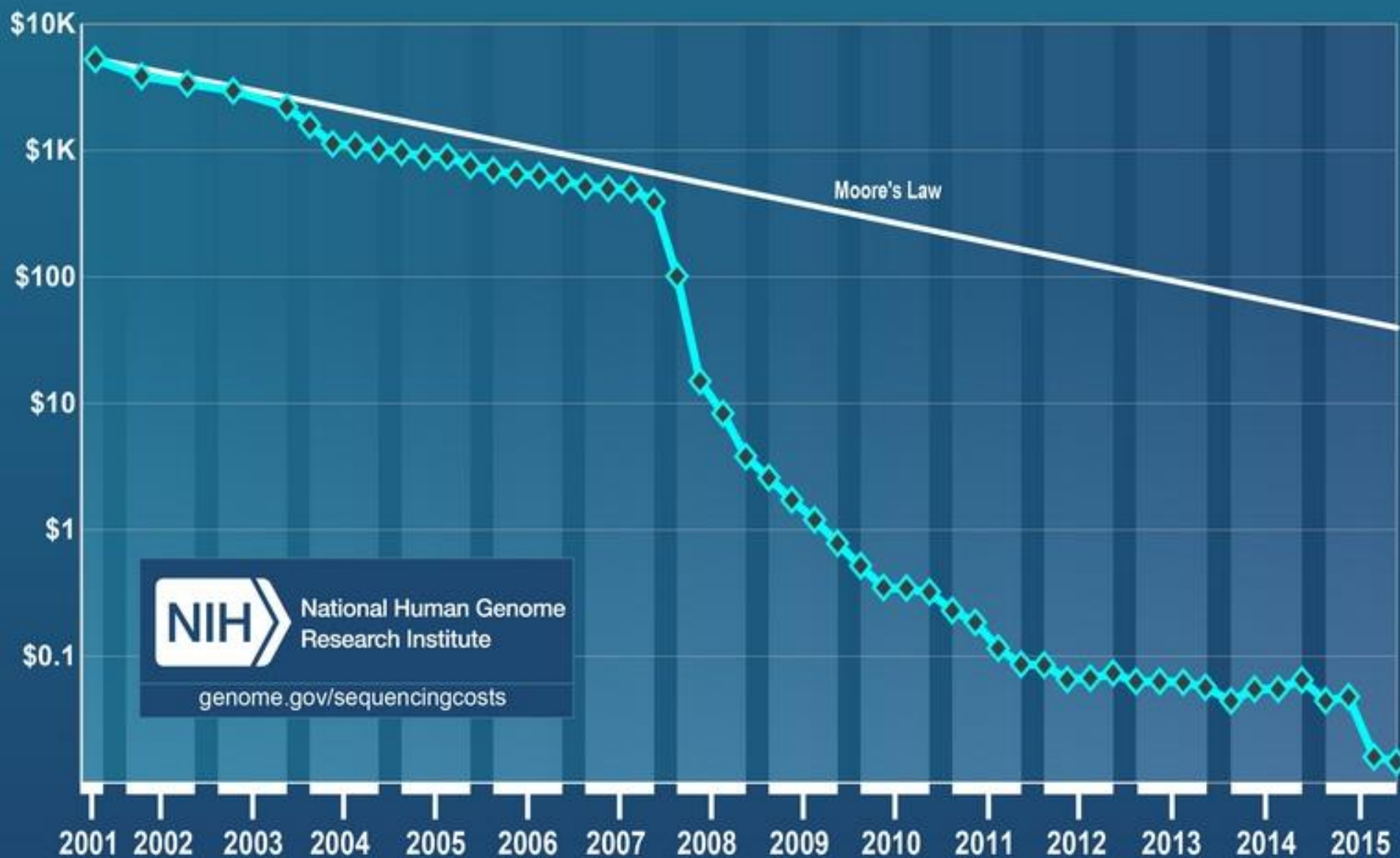
Growth of GenBank (1982 - 2004)



Growth of GenBank (1982 – 2016)



Cost per Raw Megabase of DNA Sequence

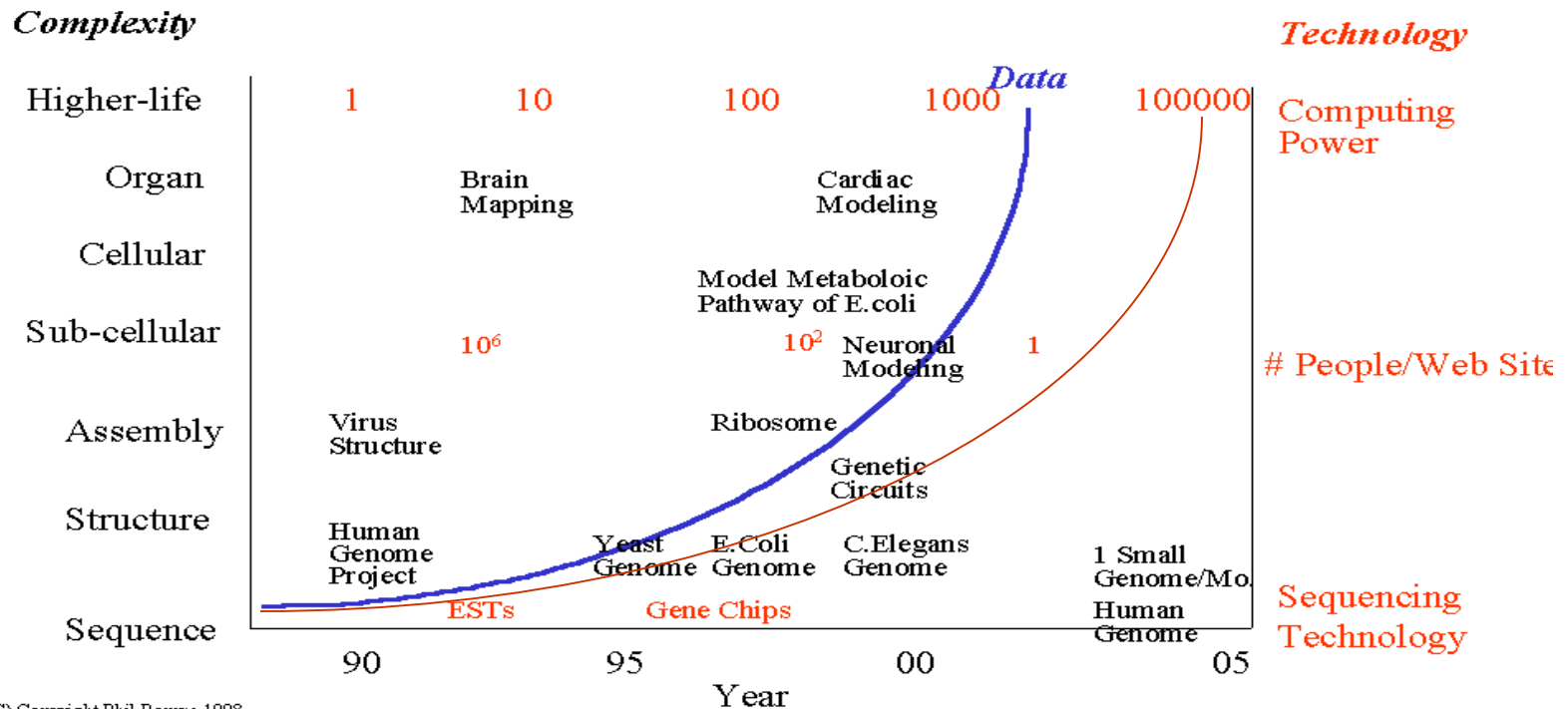
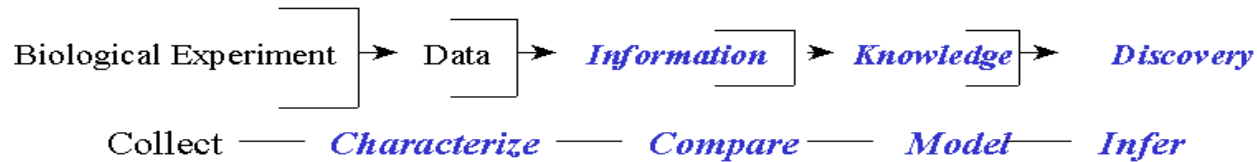


生物信息学的产生 之 促成（2）

- 生物信息的全面发掘——巨大数据量的产生。
- 计算机与信息科学的迅猛发展——为巨大数据量的处理提供了可能。

近三十年来，数据库技术、数据挖掘技术、网络技术高速发展，硬件性能迅速提高。

Bioinformatics - A Revolution

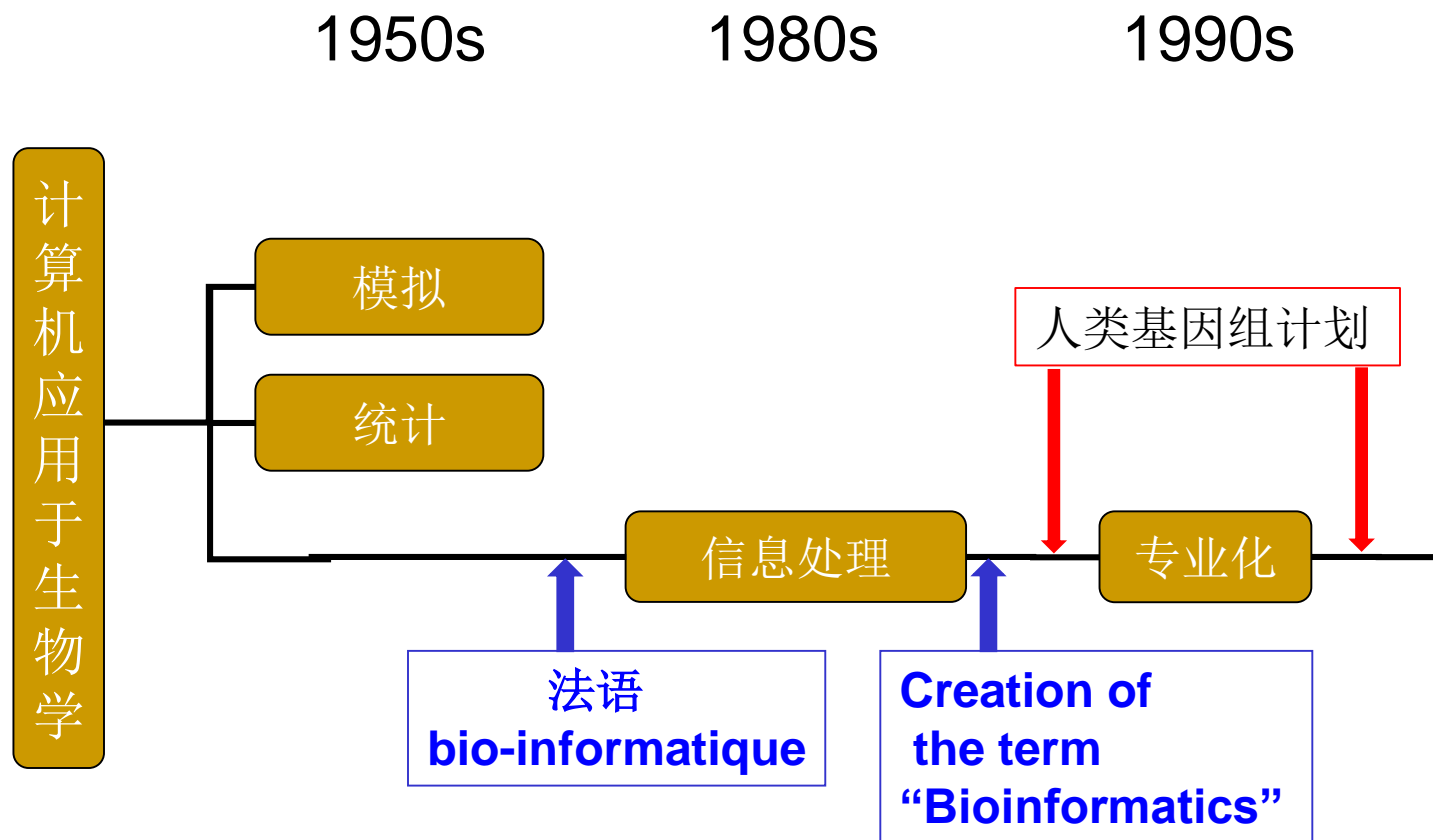


(C) Copyright Phil Bourne 1998

DNA序列数据: 14个月增长一倍;

计算机运算速度: 18个月增长一倍;

生物信息学的产生---示意图



生物信息学的发展阶段（30余年来）

- **前基因组时代：**生物数据库的建立、检索工具的开发以及DNA和蛋白质序列分析。如Genbank, Sequence alignment算法等。 (1980s)
- **基因组时代：**基因组测序，基因寻找和识别、网络数据库的建立和交互界面的开发等。如基因组信息组装，建立与发展EST（Expressed sequence tag）数据库等。 (1990s)
- **后基因组时代：**大规模基因组分析、蛋白质组分析及各种数据的比较和整合。例如，蛋白质组学（Proteomics）的产生以及功能基因组学的发展等。（系统生物学 Systems biology）。 (2000s~...)

生物信息学面临的挑战

- 建立生物信息系统相关的数据挖掘技术；
发展揭示大规模数据集合中不同组分之间关系的统计分析方法及优化算法.
- 开发各种类型的数据转换工具；
- 功能基因组学研究（系统生物学的深入探索）；
蛋白质相互作用网络；
基因组非编码区信息挖掘.
- 蛋白质结构的预测~~？
-

第一章 绪论

- 生物信息学的定义和内容
- 生物信息学的产生和发展
- 生物信息学的应用和产业
- 生物信息学的教育和学习

生物信息学的应用

- 生物工程： 酿造工艺、食品饮料.....
 - 农学： 遗传育种、作物抗病.....
 - 医学： 分子病理学、流行病学.....。
 - 药学： 药物开发（应用的范例）。
- 等等.....

应用于流行病传播途径推测

- **原理:** 变异较快的病原体，在传播的过程中会有不断的、明显的变异；所以通过它们的基因或基因组测序可以了解其传播途径。

(HIV的传播, 1992年, Science)

一个牙医和他的七个病人.....

生物信息学的应用

- 生物工程： 酿造工艺、食品饮料.....
 - 农学： 遗传育种、作物抗病.....
 - 医学： 分子病理学、流行病学.....。
 - 药学： 药物开发（应用的范例）。
- 等等.....

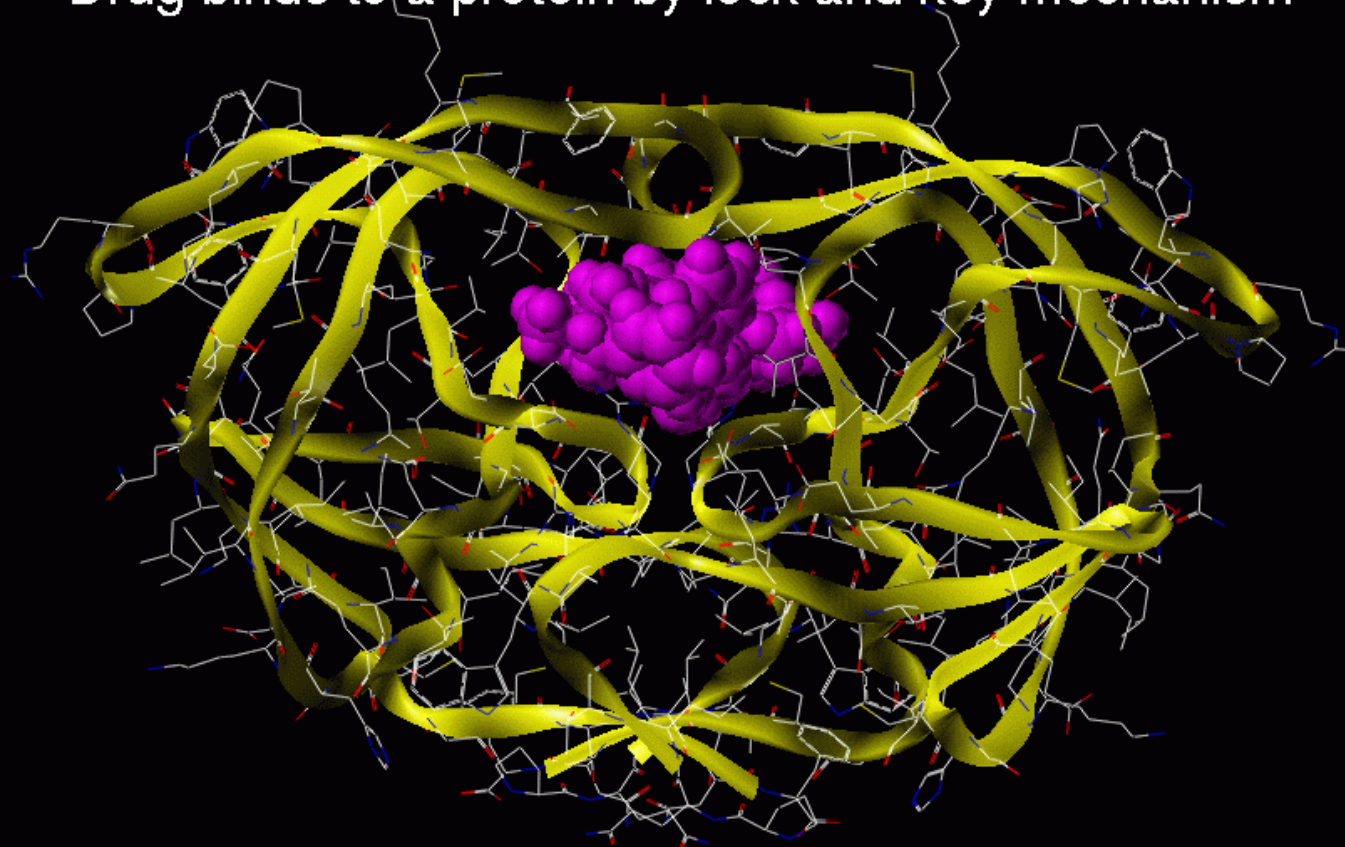
Computer-aided drug design (CADD)

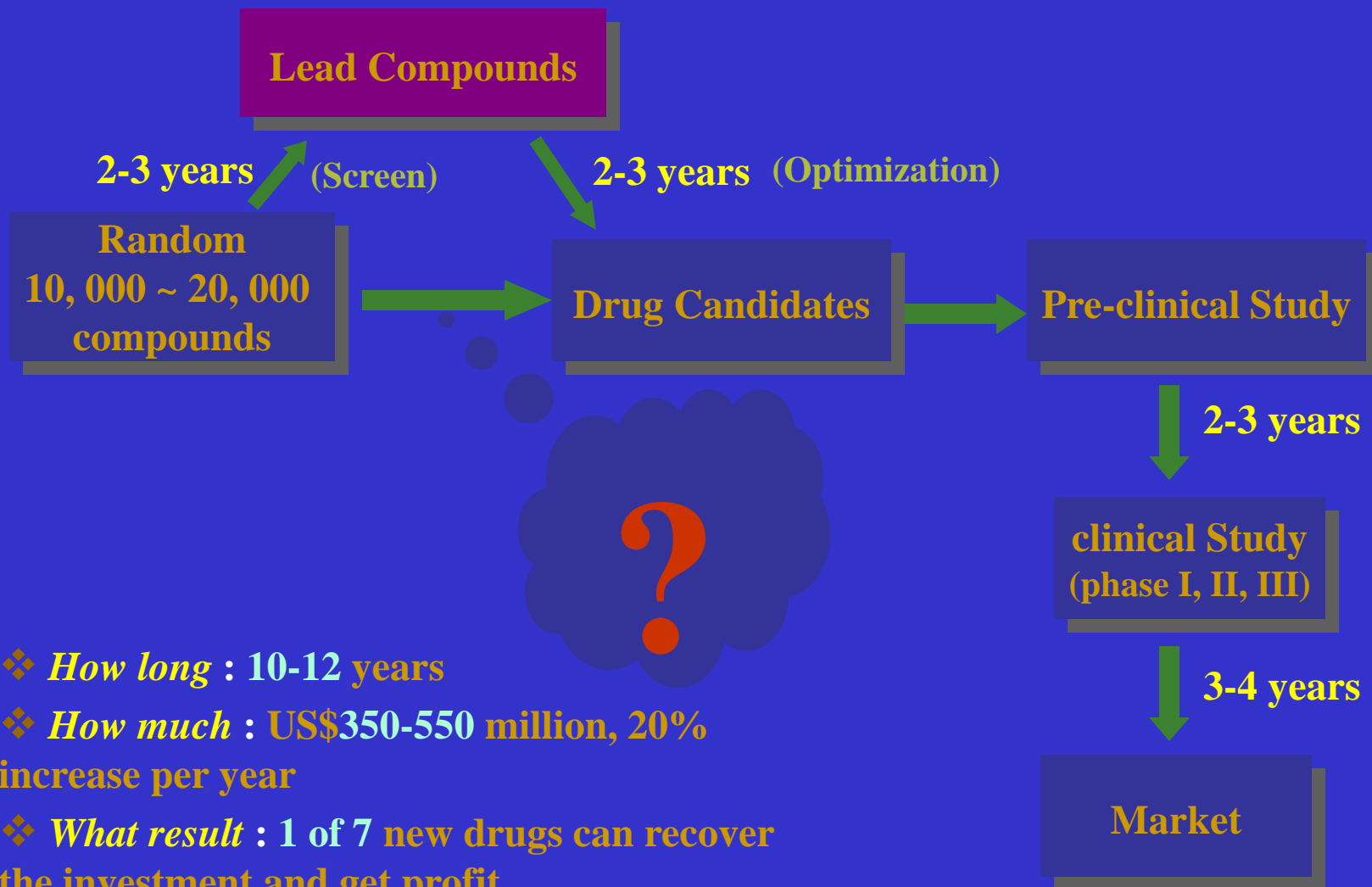
- 现在每项有一定规模的新药研究工作中，**CADD**的研究都是一项基本工作；世界上每一个大的制药公司都在致力于发展**CADD**技术。
- 主要方向——“面向靶蛋白”的药物设计。
原理：Drugs interact with targets, usually proteins (enzymes, acceptors, ion channels...), and thus cause physiological responses and curative effects.

糖尿病 —— 唐力、阿卡波糖；

Structure-Based Drug Design

Drug binds to a protein by lock and key mechanism





- ❖ *How long* : 10-12 years
- ❖ *How much* : US\$350-550 million, 20% increase per year
- ❖ *What result* : 1 of 7 new drugs can recover the investment and get profit

New drug discovery is one of the most exciting and stimulating challenges.

Accelrys Discovery Studio includes pharmacophore modeling software - Windows Internet Explorer

http://accelrys.com/products/discovery-studio/pharmacophore.html

Discovery studio (Insight II...)


accelrys®

Resource Center | Product Downloads | Customer Support | Contact

HOME | SOLUTIONS | PLATFORM & PRODUCTS | SERVICES | INNOVATION | COMMUNITIES | ABOUT ACCELRYRS

DISCOVERY STUDIO

Optimize your drug discovery process with a flexible application that delivers predictive science to its required depth.



Home > Products > Discovery Studio > Pharmacophore

Overview

ADME

Antibody Modeling

Biopolymer Building and Analysis

Pharmacophore Modeling and Analysis

Rapidly identify promising new molecular entities with or without target-structured data with the help of Discovery Studio Pharmacophore Modeling and Analysis tools, which enable:

Next Steps

CONTACT US

Register: Discovery

完成

Internet 100%

开始 生物信息学2011 Microsoft Power... 生物信息学-教... 7 Internet Exp... 11:12

第一章 绪论

- 生物信息学的定义和内容
- 生物信息学的产生和发展
- 生物信息学的应用和产业
- 生物信息学的教育和学习

生物信息学的教育与学习

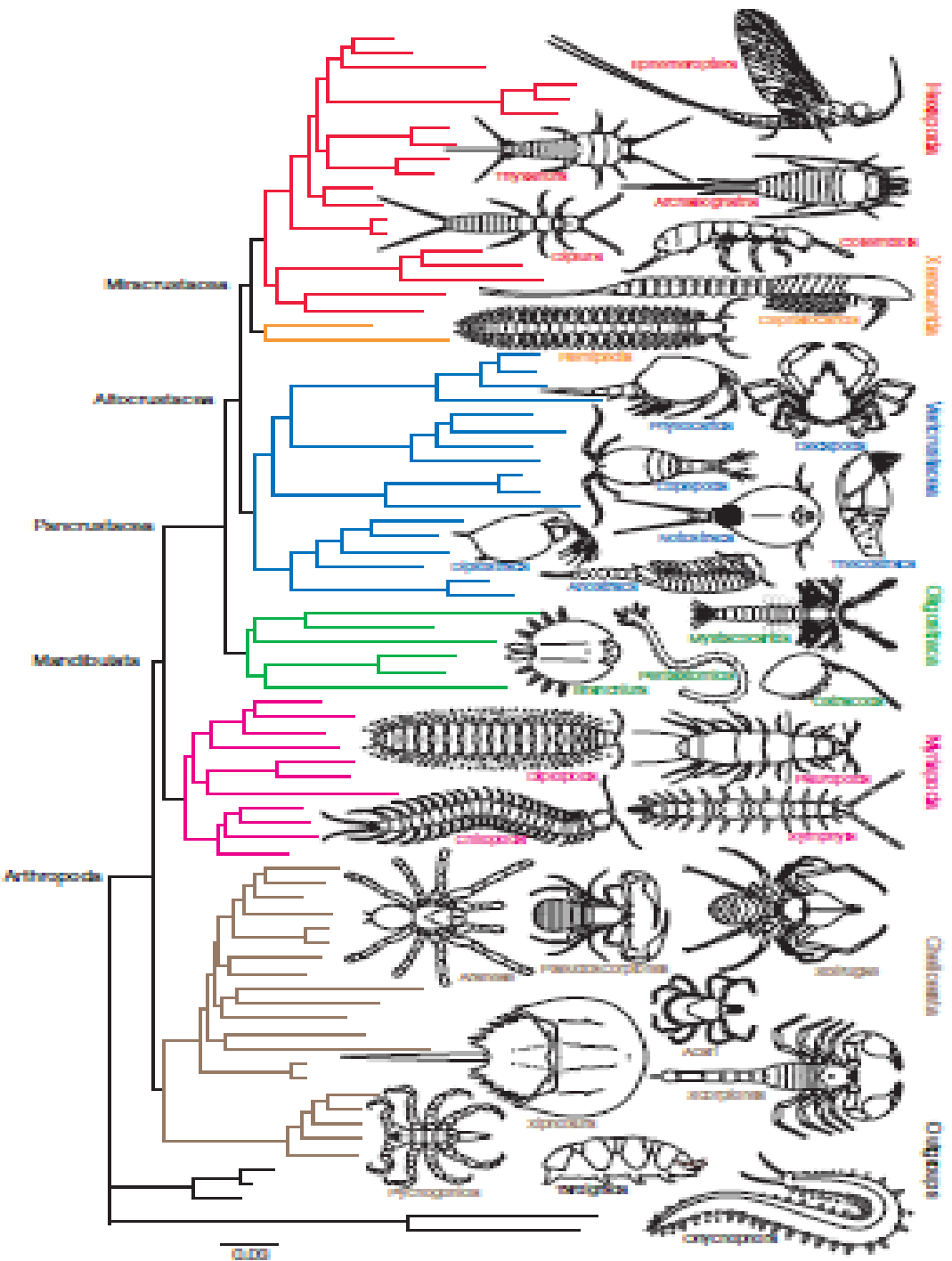
- 波士顿、宾州、加州、德州农工、印第安纳大学等于1990s中期即开始培养生物信息学博士和硕士。我国最早——北京大学生物信息学博士和硕士项目。
- 目前几乎每所美国和欧洲综合性大学均开设生物信息学本科课程以及各种类型的培训班和夏季课程等。我国情况也类似。
——众多网上课程。

生物信息学相关期刊

- Bioinformatics
- Briefings in Bioinformatics
- BMC Bioinformatics
- Applied Bioinformatics
- Genome Research
- Proteomics
- In Silico Biology
- J Computational Biology
- Plos Computational Biology
- Nucleic Acid Research
(一月数据库专辑;
七月软件资源专辑)

.....

Nature, 2010, 2



Nature, 20

The sequence and *de novo* assembly of the giant panda genome



Ruiqiang Li^{1,2*}, Wei Fan^{1*}, Geng Tian^{1,3*}, Hongmei Zhu^{1*}, Lin He^{4,5*}, Jing Cai^{3,6*}, Quanfei Huang¹, Qingle Cai^{1,7}, Bo Li¹, Yinqi Bai¹, Zhihe Zhang⁸, Yaping Zhang⁶, Wen Wang⁶, Jun Li¹, Fuwen Wei⁹, Heng Li¹⁰, Min Jian¹, Jianwen Li¹, Zhaolei Zhang¹¹, Rasmus Nielsen¹², Dawei Li¹, Wanjun Gu¹³, Zhentao Yang¹, Zhaoling Xuan¹, Oliver A. Ryder¹⁴, Frederick Chi-Ching Leung¹⁵, Yan Zhou¹, Jianjun Cao¹, Xiao Sun¹⁶, Yonggui Fu¹⁷, Xiaodong Fang¹, Xiaosen Guo¹, Bo Wang¹, Rong Hou⁸, Fujun Shen⁸, Bo Mu¹, Peixiang Ni¹, Runmao Lin¹, Wubin Qian¹, Guodong Wang^{3,6}, Chang Yu¹, Wenhui Nie⁶, Jinhuan Wang⁶, Zhigang Wu¹, Huiqing Liang¹, Jiumeng Min^{1,7}, Qi Wu⁹, Shifeng Cheng^{1,7}, Jue Ruan^{1,3}, Mingwei Wang¹, Zhongbin Shi¹, Ming Wen¹, Binghang Liu¹, Xiaoli Ren¹, Huisong Zheng¹, Dong Dong¹¹, Kathleen Cook¹¹, Gao Shan¹, Hao Zhang¹, Carolin Kosiol¹⁸, Xueying Xie^{1,3}, Zuhong Lu^{1,3}, Hancheng Zheng¹, Yingrui Li^{1,3}, Cynthia C. Steiner¹⁴, Tommy Tsan-Yuk Lam¹⁵, Siyuan Lin¹, Qinghui Zhang¹, Guoqing Li¹, Jing Tian¹, Timing Gong¹, Hongde Liu¹⁶, Dejin Zhang¹⁶, Lin Fang¹, Chen Ye¹, Juanbin Zhang¹, Wenbo Hu¹⁷, Anlong Xu¹⁷, Yuanyuan Ren¹, Guojie Zhang^{1,3,6}, Michael W. Bruford¹⁹, Qibin Li^{1,3}, Lijia Ma^{1,3}, Yiran Guo^{1,3}, Na An¹, Yujie Hu^{1,3}, Yang Zheng^{1,3}, Yongyong Shi⁵, Zhiqiang Li⁵, Qing Liu¹, Yanling Chen¹, Jing Zhao¹, Ning Qu^{1,7}, Shancen Zhao¹, Feng Tian¹, Xiaoling Wang¹, Haiyin Wang¹, Lizhi Xu¹, Xiao Liu¹, Tomas Vinar²⁰, Yajun Wang²¹, Tak-Wah Lam²², Siu-Ming Yiu²², Shiping Liu²³, Hemin Zhang²⁴, Desheng Li²⁴, Yan Huang²⁴, Xia Wang¹, Guohua Yang¹, Zhi Jiang¹, Junyi Wang¹, Nan Qin¹, Li Li¹, Jingxiang Li¹, Lars Bolund¹, Karsten Kristiansen^{1,2}, Gane Ka-Shu Wong^{1,2,5}, Maynard Olson²⁶, Xiuqing Zhang¹, Songgang Li¹, Huanming Yang¹, Jian Wang¹ & Jun Wang^{1,2}

Using next-generation sequencing technology alone, we have successfully generated and assembled a draft sequence of the giant panda genome. The assembled contigs (2.25 gigabases (Gb)) cover approximately 94% of the whole genome, and the remaining gaps (0.05 Gb) seem to contain carnivore-specific repeats and tandem repeats. Comparisons with the dog and human showed that the panda genome has a lower divergence rate. The assessment of panda genes potentially underlying some of its unique traits indicated that its bamboo diet might be more dependent on its gut microbiome than its own genetic composition. We also identified more than 2.7 million heterozygous single nucleotide polymorphisms in the diploid genome. Our data and analyses provide a foundation for promoting mammalian genetic research, and demonstrate the feasibility for using next-generation sequencing technologies for accurate, cost-effective and rapid *de novo* assembly of large eukaryotic genomes.

The giant panda, *Ailuropoda melanoleura*, is at high risk of extinction because of human population expansion and destruction of its habitat. The latest molecular census of its population size, using faecal samples and nine microsatellite loci, provided an estimate of only 2,500–3,000 individuals, which were confined to several small mountain habitats in Western China¹. The giant panda has several unusual biological and behavioural traits, including a famously restricted diet,

primarily made up of bamboo, and a very low fecundity rate. Moreover, the panda holds a unique place in evolution, and there has been continuing controversy about its phylogenetic position². At present, there is very little genetic information for the panda, which is an essential tool for detailed understanding of the biology of this organism.

A major limitation in obtaining extensive genetic data is the prohibitive costs associated with sequencing and assembling large eukaryotic

本课程教学计划：

- 生物信息学简介
- 生物信息学的计算机基础
- 生物信息学资源及其利用
- 序列分析（DNA/RNA/protein）
- 分子系统发育分析（进化分析）
- 基因组分析（转录组分析）
- 蛋白质组分析
- 结构分析（RNA/protein）

教学目标和考核方法

■ 教学目标：

结合学科前沿进展讲授生物信息学的基本理论知识，指导相关上机实践，培养学生主动利用生物信息资源和运用相关软件工具的能力，以及接受本学科新知识、新方法的能力。

■ 考核方法：

平时成绩：35%（包括三次上机作业）；

期末考试：65%。

本课程教学的关键点/特点:

- 形成“生物”与“信息”本质关联的理解——不仅仅从方法上理解“生物信息学”领域;
- 了解生物信息学中的算法思路——体验生物信息学的工作方式;
- 了解应用生物信息学数据和软件的正确方法——准确理解生物信息学工具在科研中的作用。

本课程教学的关键点/特点:

- 形成“生物”与“信息”本质关联的理解——不仅仅从方法上理解“生物信息学”领域;
- 了解生物信息学中的算法思路——体验生物信息学的工作方式;
- 了解应用生物信息学数据和软件的正确方法——准确理解生物信息学工具在科研中的作用。

教材：

《生物信息学》（第四版）陈铭主编，科学出版社，**2022**年。

参考书目：

“Bioinformatics” (4th Ed), Baxevanis AD, Bader GD and Wishart DS,
John Wiley & Sons, 2020.

“Bioinformatics and Functional Genomics” (3rd Ed), Pevsner J, Wiley-
Blackwell, 2015.

“Bioinformatics: Sequence and Genome Analysis” (2nd Ed), Mount DW,
CSHL Press, 2013.

“A Bioinformatics Guide for Molecular Biologists”, Aerni, SJ and Sirota
M, CSHL Press, 2014.

“Bioinformatics Data Skills”, Buffalo, V. O'Reilly, 2015.

NCBI ---- 最重要的生物信息数据中心

NCBI Resources How To MyNCBI Sign In

NCBI National Center for Biotechnology Information

Search All Databases Search Clear

NCBI Home

Site Map (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [Research](#) | [RSS Feeds](#)

Get Started

[Tools](#): Analyze data using NCBI software

[Downloads](#): Get NCBI data or software

[How-To's](#): Learn how to accomplish specific tasks at NCBI

[Submissions](#): Submit data to GenBank or other NCBI databases

Genomic Structural Variation

dbVar archives large scale genomic variation data and associates defined variants with phenotypic information.

Popular Resources

- BLAST
- Bookshelf
- Gene
- Genome
- Nucleotide
- OMIM
- Protein
- PubChem
- PubMed
- PubMed Central
- SNP

NCBI News

[Retirement of Peptidome, SRA & Trace Archive](#)

16 Feb 2011

Due to budget constraints, NCBI will be discontinuing the

http://www.ncbi.nlm.nih.gov/dbvar/

Internet 100%

开始 生物信息学2011 Microsoft PowerP... National Center ... 9:33

文件 (File) menu and other browser icons.

EBI ---- 最完备的生物信息学网站之一



Databases Tools EBI Groups Training Industry About Us Help Site Index

Explore the EBI:

FIND

Examples: [ROA1](#), [HUMAN](#), [tpi1](#), [Sulston](#)...

Help & Suggestions

Support & Feedback

Data Resources & Tools

- ENA
- UniProt
- ArrayExpress
- Ensembl
- InterPro
- PDBe
- Genomes
- Nucleotide Sequences
- Protein Sequences
- Macromolecular Structures
- Small Molecules
- Gene Expression
- Molecular Interactions
- Reactions & Pathways
- Protein Families
- Enzymes
- Literature
- Taxonomy
- Ontologies
- Patent Resources
- Sequence Similarity & Analysis
- Pattern & Motif Searches
- Structure Analysis
- Text Mining
- Downloads
- Web Services



European Bioinformatics Institute

Footer area with links like Latest News and Events.



Swiss Institute of
Bioinformatics



ExPASy Proteomics Server

Search for

[Databases](#) [Tools](#) [Services](#) [Mirrors](#) [About](#) [Contact](#)

You are here: [ExPASy CA](#)

The ExPASy (**Expert Protein Analysis System**) [proteomics](#) server of the [Swiss Institute of Bioinformatics](#) (SIB) is dedicated to the analysis of protein sequences and structures as well as 2-D PAGE ([Disclaimer](#) / [References](#) / [Linking to ExPASy](#)).

Databases

[UniProtKB](#), [PROSITE](#), [HAMAP](#), [SwissVar](#),
[ViralZone](#), [SWISS-MODEL Repository](#),
[neXtProt](#), [SWISS-2DPAGE](#), [World-2DPAGE](#)
[Repository](#), [MIAPEGelDB](#), [ENZYME](#),
[GlycoSuiteDB](#), [UniPathway](#)
[\[details\]](#) [\[full list\]](#)

Education & services

[Downloads](#), [Protein Spotlight](#),

Tools & Software

[Proteomics tools](#), [Blast](#), [ScanProsite](#),
[Melanie](#), [MSight](#), [Make2D-DB](#), [SWISS-](#)
[MODEL](#), [Swiss-PdbViewer](#), [SwissDock](#),
[SwissParam](#)
[\[full list\]](#)

Documentation

[What's New?](#), [E-mail alerts](#), [UniProtKB](#)

Latest News

ExPASy-beta - February 4, 2011

[ExPASy-beta](#), the first public version of the new ExPASy SIB Bioinformatics Resource Portal is now available online at:

<http://beta.expasy.org>. [ExPASy-beta](#) gives you access to scientific databases and software tools in different areas of life sciences including proteomics, genomics, phylogeny, systems biology, population genetics, transcriptomics etc. The wide range of resources is mainly provided by SIB groups and

http://www.nig.ac.jp/index-e.html

Google

NIG---日本著名的生物信息学中心



Research Organization of Information and Systems
National Institute of Genetics

ROIS

NIPR

NII

ISM

JAPANESE

ACCESS

SITEMAP

SEARCH

Go

HOME

About NIG

Research

Graduate Program

Database

Seminars

Open Seminars

Local Information

Information

>MORE

- DDBJ will continue Sequence Raw Data Archiving **NEW!**
- Undergraduate Summer Internship at NIG "NIGINTERN2011" **NEW!**
- NIG Postdoctoral Fellows 2011
- International Graduate Program at NIG **UPDATED**

News & Topics

>MORE



Dr. Tatsuo Fukagawa was appointed Associate Editor on the Editorial Board of Chromosome Research. **NEW!**



NIG International Symposium 2011 "1st Asia-Pacific Drosophila Research Conference"



NIG International Symposium "Future Perspectives of Biological Databases" October 11, 2010.



Dr. Hiroyuki Araki received the Kihara Prize of the Genetics Society of Japan.

Virtual Museum of Genetics [Japanese]

DNA Data Bank of Japan

National BioResource Project

China National GeneBank (CNGB)

CNGBdb

首页

数据资源

科学数据库

数据提交

计算

关于我们

搜索

国家基因库生命大数据平台(CNGBdb)

国家基因库生命大数据平台(China National GeneBank DataBase, CNGBdb)是一个为科研社区提供生物大数据共享和应用服务的统一平台(Science as a Service), 基于大数据和云计算技术, 提供数据归档、计算分析、知识搜索、管理授权和可视化等数据服务。



可信计算平台

安全可靠, 高效扩展



单细胞组学平台

32万+细胞的组学数据和可视化



GISAID EpiCoV™ 数据库

全球流感数据共享



数据资源

样本资源

数据归档

科学数据库

数据广场



CNCB

国家生物信息中心

全球生物信息的汇交、存储、管理及生物信息前沿交叉研究与转化应用



All databases

Find a bioproject, biosample, gene, protein, tc

Q Search



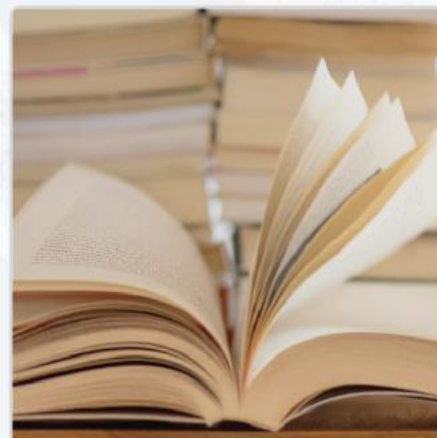
数据资源



计算分析



科学研究



文献情报



群名称:武大生信2023

群 号:650916900