

# 基因组学

汇报人：秦雅婷

2010.11.19



- Definition
- Background
- Major publications
- Classic papers
- Important research institutions and domain experts
- Major achievements
- Important new frontiers



# What is -omics

## Genomics

- Mapping, sequencing and analysis of genome

## Structure genomics

- high-throughput determination of 3D crystal structure

## Functional genomics

- Simultaneous study of function of all gene in pathway

## Proteomics

- Simultaneous study of all proteins in a pathway or system

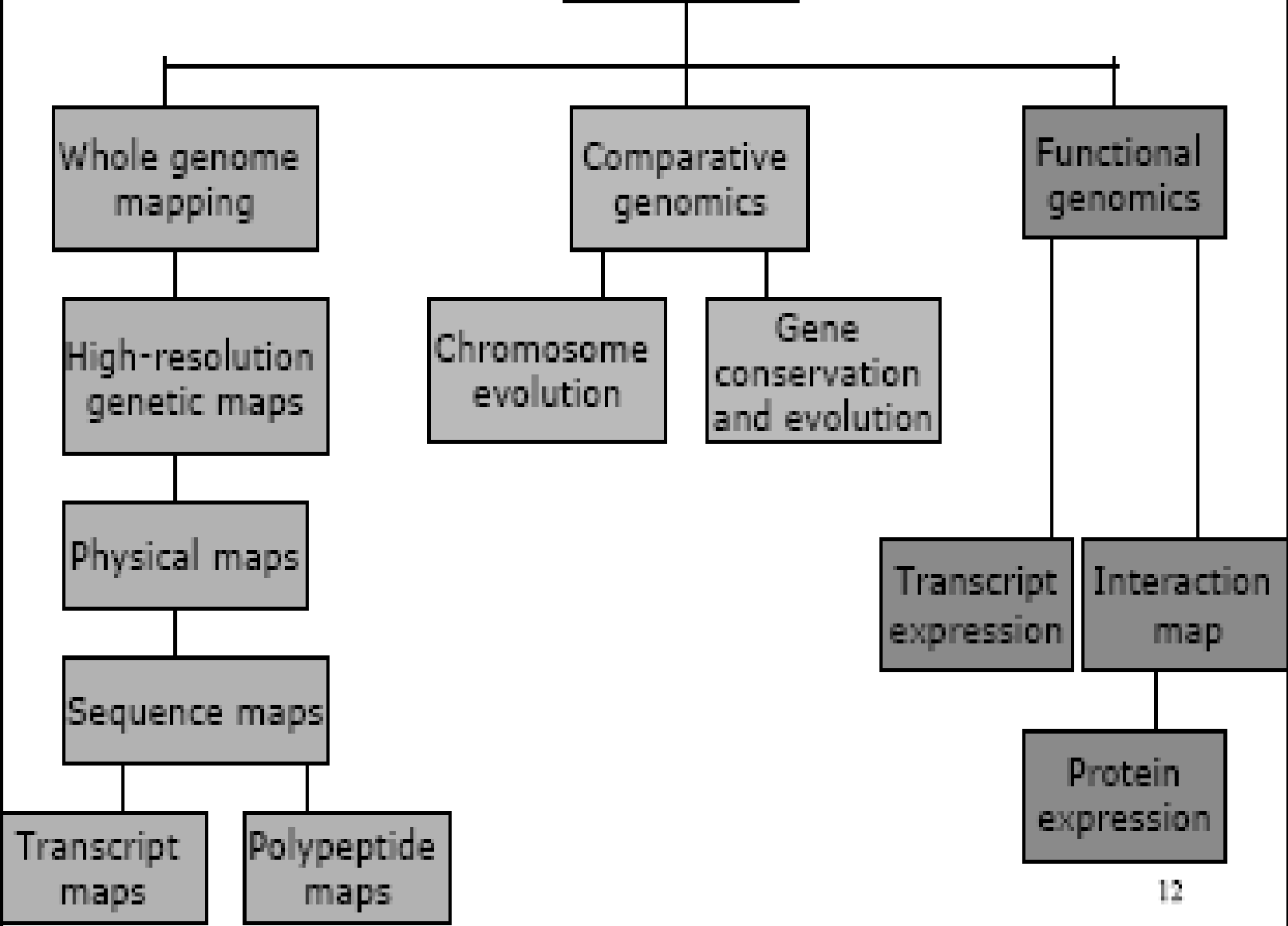




■ 基因组学（**Genomics**):研究基因组及其基因的科学。

■ 最初是**Thomas Roderick**于**1986**年提出，其主要内容是指基因组作图（**Mapping**）和测序（**Sequencing**）。

# Genomics



## 科学发展历程

### 1953 New Discovery

**(April) James Watson and Francis Crick discover the double helical structure of DNA (*Nature*).**



# 1977 Methods for Sequencing DNA



1972

(October) Paul Berg and co-workers create the first recombinant DNA molecule (*PNAS*).



1977

Allan Maxam and Walter Gilbert (pictured) at Harvard University and Frederick Sanger at the U.K. Medical Research Council (MRC) independently develop methods for sequencing DNA (*PNAS*, February; *PNAS*, December)



# 1985 HGP first be discussed



**(May) Robert Sinsheimer hosts a meeting at the University of California (UC), Santa Cruz, to discuss the feasibility of sequencing the human genome.**



**(December) Kary Mullis and colleagues at Cetus Corp. develop PCR, a technique to replicate vast amounts of DNA (*Science*).**



# 人类科学史上的三大工程



阿波罗登月计划

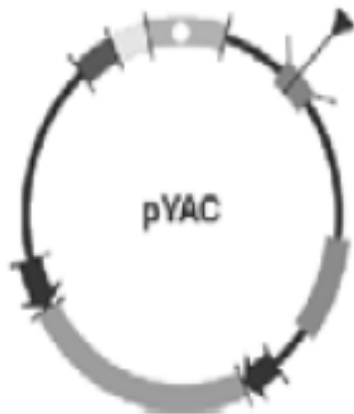


人类基因组计划



曼哈顿原子计划

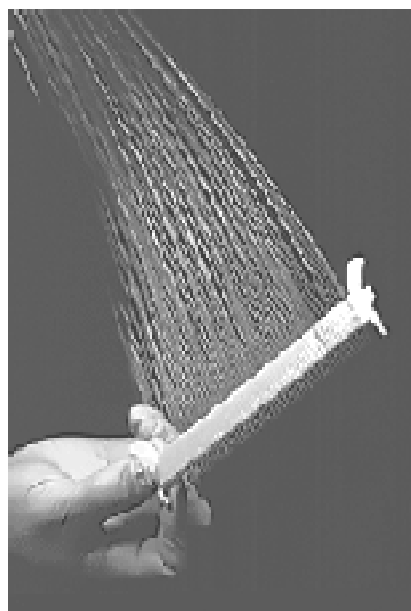
# 1987 Develop YACs for Cloning



(May) David Burke, Maynard Olson, and George Carle of Washington University in St. Louis develop YACs (left) for cloning, increasing insert size 10-fold (*Science*).

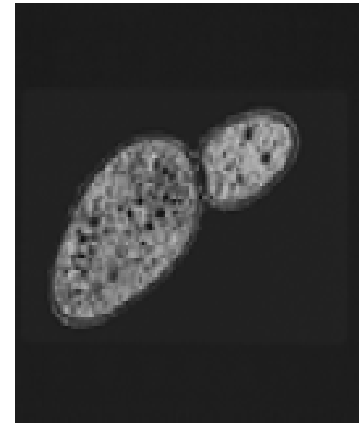
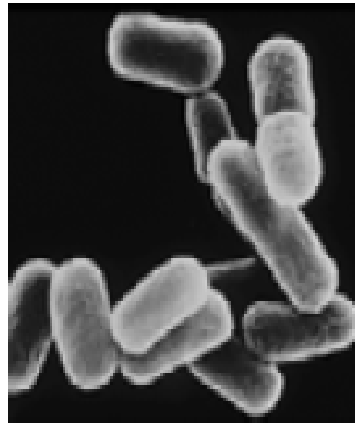
Applied Biosystems Inc. puts the first automated sequencing machine, based on Hood's technology, on the market.

# 1990 Capillary Electrophoresis



Three groups develop capillary electrophoresis, one team led by Lloyd Smith (*Nucleic Acids Research*, August), the second by Barry Karger (*Analytical Chemistry*, January), and the third by Norman Dovichi (*Journal of Chromatography*, September).

# 1990 large-scale sequencing on model organisms



(August) NIH (National Institutes of Health) begins large-scale sequencing trials on four model organisms: *Mycoplasma capricolum* (支原体), *Escherichia coli* (大肠杆菌) (up, pink), *Caenorhabditis elegans* (线虫) (up, rainbow), and *Saccharomyces cerevisiae* (啤酒酵母) (up, ovals). Each research group agrees to sequence 3 Mb at 75 cents a base within 3 years.

# 1991 rice genome sequencing begins

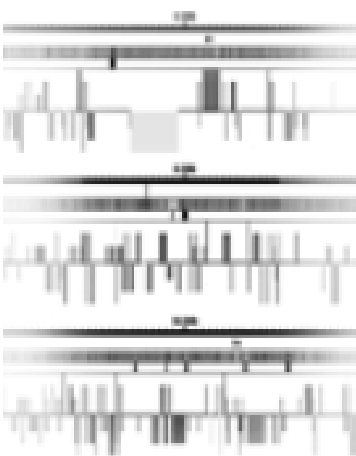
(June) NIH biologist **J. Craig Venter** announces a strategy to find expressed genes, using ESTs (*Science*). A fight erupts at a congressional hearing 1 month later, when **Venter** reveals that NIH is filing patent applications on thousands of these partial genes.



(October) The Japanese rice **genome** sequencing effort begins.



# 2000 New Development



(March) Celera and academic collaborators sequence the 180-Mb genome of the fruit fly *Drosophila melanogaster* (果蝇) (left), the largest genome yet sequenced and a validation of Venter's controversial whole-genome shotgun method (*Science*).

# 2000 New Development



(October) DOE (Department of Energy) and MRC (Medical Research Council) launch a collaborative project to sequence the genome of the puffer fish, *Fugu rubripes* (河豚) (left), by March 2001.

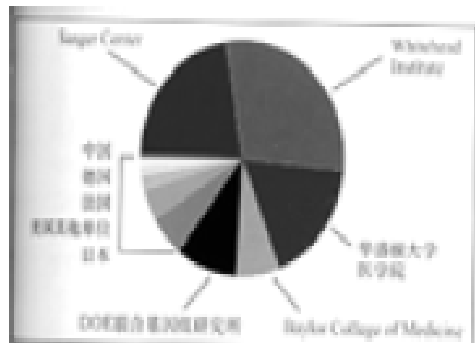


(December) An international consortium completes the sequencing of the first plant, *Arabidopsis thaliana* (拟南芥) (left), 125 Mb.



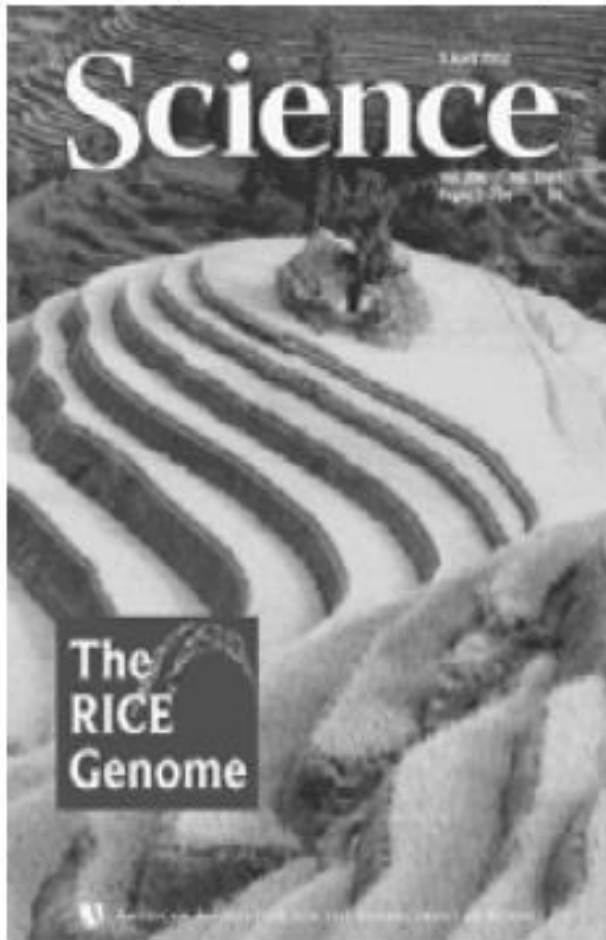
# 2001 Draft Sequence of Human Genome Published

The HGP consortium publishes its working draft in *Nature* (15 February), and Celera publishes its draft in *Science* (16 February).



For more details, we can refer the web:  
<http://www.sciencemag.org/feature/plus/sfg/human/timeline.shtml>

# And 2002 April 5, A Draft Sequence of Rice Genome published by Chinese scientist



Refer Science Online:

<http://intl.sciencemag.org/cgi/content/abstract/296/5565/79>

A Draft Sequence Assembly of the Rice Genome  
*Science* 2002 April 5; 296(5565): p. 79-92

# Sequenced Organisms

Organism	Genome size	Completion date	Estimated no. of genes
<i>H. influenzae</i>	1.8 MB	1995	1,740
<i>S. cerevisiae</i>	12.1 Mb	1996	6,034
<i>E. coli</i>	4.7 Mb	1997	4288
<i>C. elegans</i>	97 Mb	1998	19,099
<i>A. thaliana</i>	125 Mb	2000	25,000
<i>D. melanogaster</i>	180 Mb	2000	13,061
<i>M. musculus</i>	3000 Mb	2002	~30,000?
<i>O. sativa</i>	389 Mb	2000	~3,200
<i>H. sapiens</i>	3000 Mb	2001	~30,000?
<i>G. gallus</i>	1000 Mb	2004	20000~30000
<i>B. Mori</i>	428.7 Mb	2004	18510



## A Draft Sequence for the Genome of the Domesticated Silkworm (*Bombyx mori*)

Biology analysis group: Qingxin Xia,<sup>1\*</sup> Ziyang Zhou,<sup>1\*</sup> Cheng Lu,<sup>1\*</sup> Dejun Cheng,<sup>1</sup> Fenglin Dai,<sup>1</sup> Bin Li,<sup>1</sup> Ping Zhao,<sup>1</sup> Jingfu Zhu,<sup>1</sup> Tingzai Cheng,<sup>1</sup> Chundi Chai,<sup>1</sup> Guojing Fan,<sup>1</sup> Jiechen Xu,<sup>1</sup> Chun Liu,<sup>1</sup> Ying Lin,<sup>1</sup> Jifeng Qian,<sup>1</sup> Yang Hui,<sup>1</sup> Zhangli Wu,<sup>1</sup> Guorong Li,<sup>1</sup> Huihui Fan,<sup>1</sup> Chunfang Li,<sup>1</sup> Yitang Shen,<sup>1</sup> Xijian Lan,<sup>1</sup> Ganwei Yuan,<sup>1</sup> Tian Li,<sup>1</sup> Hanfu Xu,<sup>1</sup> Guangwei Yang,<sup>1</sup> Yongli Wan,<sup>1</sup> Yang Zhu,<sup>1</sup> Haode Yu,<sup>1</sup> Weide Shen,<sup>1</sup> Dayang Wu,<sup>1</sup> Zhonghui Wang<sup>1</sup>  
Genome analysis group: Jun Ye,<sup>2,3\*</sup> Jun Wang,<sup>2,3\*</sup> Ruiqiang Li,<sup>2,3\*</sup>

draft sequence has just been announced by the International Silkworm Genome Project (ISGP). Our project is independent of, but complementary to, that of the consortium. Our sequence has been submitted to the DNA Data Bank of Japan/Toronto Library of Science/Linköping/Genome Research Center accession number AAZ000000. A ZENODO/000000 and is also available on Web site <http://www.genoscope.cns.fr/1551/seq/seq.html>. The draft sequence can be found at GenBank accession numbers CK488034 to CK501194. ISGP's first genome sequence from an animal domesticated in temperate silk gland, *Bombyx mori*, of 1201 mbps. A whole-genome/finchbird was used, and a 2.5x resolution was achieved.

2004.12.10 Science



2004.12.9 Nature

我国科学家以加入“人类基因组计划”为起点，独立启动和阶段性地完成了“中国杂交水稻基因组计划”、“鸡基因组计划”和“家蚕基因组计划”等重要经济物种的基因组学研究，建立了规模化基因组学研究平台，组建了具有国际竞争力的研究团队。

# 主要刊物

- 生物技术通报
- 生物化学与生物物理
- 植物学通报
- Genomics
- Annu Rev Genet
- J Struct Funct Genomics
- Biotechnol Lett
- THEORETICAL AND APPLIED GENETICS
- Mol Biochem Parasitol
- Int Psychogeriatr
- SCIENCE
- PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS



# 经典论文

- Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types, Journal of Experimental Medicine, 1994, 79:173-158.
- Collins FS, Guyer MS, Chakravarti A. Variations on a theme: cataloging human DNA sequence variation. Science, 1997, 278:1580-1581.



# Important research institutions

- UNIV CALIF SAN DIEGO
- SCRIPPS RES INST
- NOVARTIS RES FDN
- BURNHAM INST MED RES
- DUKE UNIV
- SLAC NATL ACCELERATOR LAB



# Domain Experts

 ELSLIGER, MA

 WILSON, IA

 GODIIK, A

 WOOLEY, J

 DEACON, AM

 KRISHNA, SS

 LESLEY, SA

 WEEKES, D





# 主要成果



## On Feb 16<sup>th</sup> 2001... What happened?

2001年2月中旬,《Nature》与《Science》分别发表了人类基因组工作框架图(the draft genome),这是人类基因组计划(human genome project, HGP)实施以来所取得的最重大进展,也是生命科学领域中的一个里程碑。自HGP提出以来,到现在一共16年。这16年来科学家们所走过的是一条充满艰辛的路,同时也是一条谱写辉煌的路。回顾这段将在人类历史上永载史册的光辉历程,相信对每一位科学殿堂里的探索者都会大有启迪。我们应该记住他们,记住他们所作出的成就。

# 近期研究成果

- PLoS Biology: 中国人单碱基精确度的全基因组DNA甲基化图谱绘制完成
- Nature: 千人基因组计划发布最详尽的人类基因多态性图谱
- Nat.Genetics: 首份日本人基因组图谱绘制成功
- Nature: 第三个人类基因组单体型图完成
- Nature Genetics: 人类全基因组同步分析方法诞生

# 重要前沿课题

## ■ Nature Genetics: 复杂疾病或由低频遗传变异决定健康的功能

北京时间2010年10月4日，美国加州大学伯克利分校、丹麦哥本哈根大学等单位合作的研究成果“对200个人类外显子的测序揭示大量低频率非同义突变的存在”在国际著名学术杂志*Nature Genetics*上。

该项研究对200个丹麦个体蛋白质编码基因的外显子组进行了深度测序，发现了大量以往未知的单核苷酸多态性位点（SNP），其中大部分在人群中都以较低频率出现。该研究完成了目前在人类外显子区域规模最大、分辨率最精细的遗传图谱，并以翔实的数据证明，人群当中的低频率多态性位点富集了大量能引起蛋白质氨基酸序列改变的变异，而这类变异在人群中受到自然选择作用，可能具有影响人类。



## Genome Res.: 大脑发育和衰老研究 新进展

大脑在生命的早期阶段经历了快速的生长和发育，接着，当人们进入老龄之后，大脑开始退化；然而，人们对区别大脑发育和衰老的生物学过程却所知甚少。在近期发表的网络版 *Genome Research* 的一篇报道中，研究人员发现了在年轻和衰老的大脑的变化之间发生的一种基因调控联系。

## PLoS One: 2型糖尿病基因在何处?

复旦大学上海医学院内分泌系，瑞典Lund大学的研究人员在2型糖尿病的研究方面取得新的进展相关成果文章公布在*PLoS One*上。

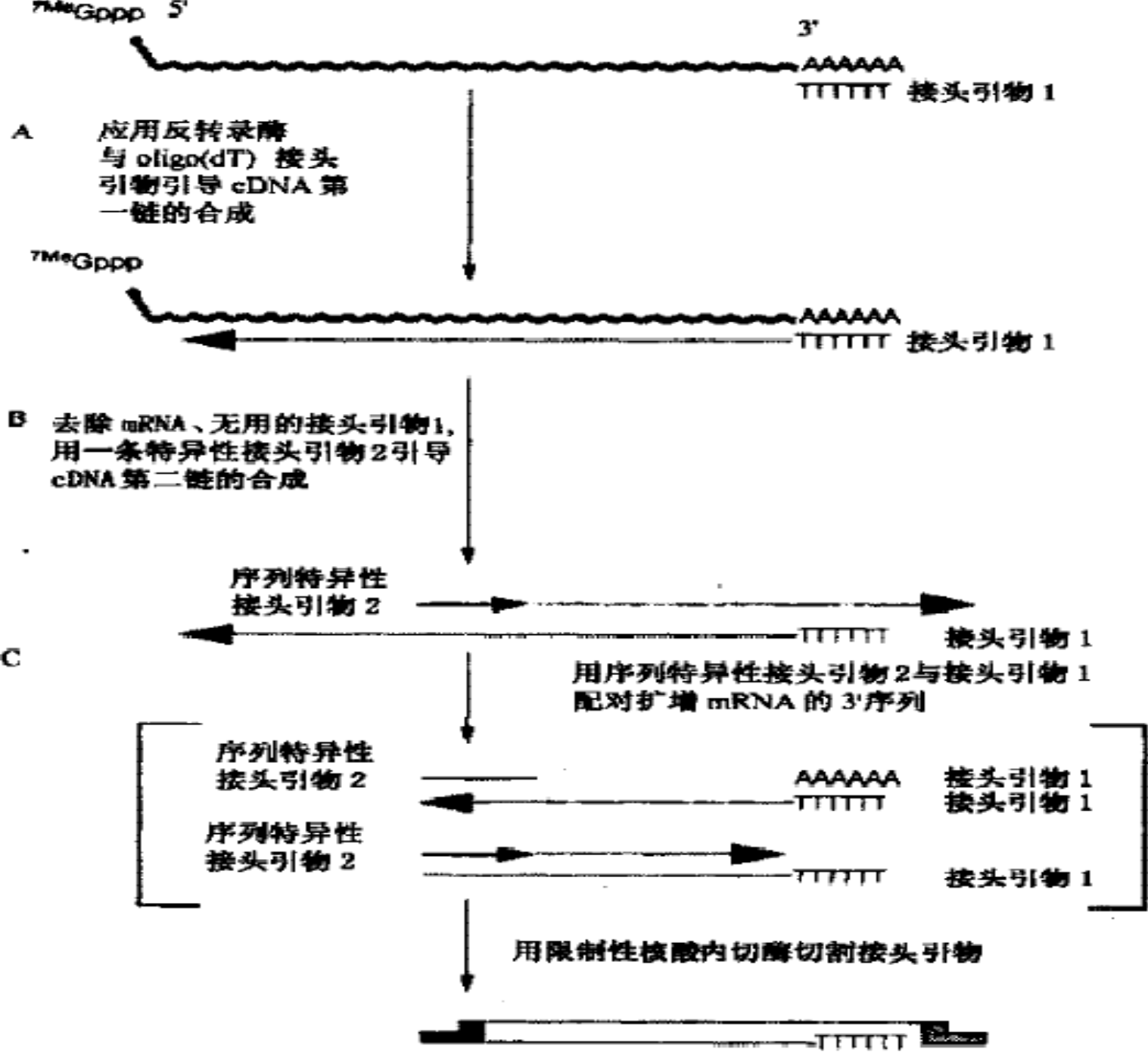
近期，复旦大学上海医学院的胡仁明教授与瑞典合作者使用GWAS技术在中国人群中鉴定2型糖尿病的遗传标记，据估计，20年后，中国糖尿病患者将翻倍。

胡仁明研究组对1165个患有2型糖尿病的患者和1136个正常人进行GWAS分析，结果发现，患有2型糖尿病的患者普遍地在CDKN2A/B，CDKAL1和TCF7L2基因上发生变异。



# CDNA3'末端的快速扩增 (3'RACE)

- 在用oligo(dT)引物构建的cDNA文库中，偶尔发生部分cDNA克隆确实了与靶mRNA3'末端互补的序列。这部分序列到底是在cDNA合成过程中还是在分子克隆中，怎样缺失和在哪个环节中缺失尚不清楚。
- 3'RACE反应通常用于分离3'末端未知的序列及对mRNA的3'末端序列的定序。





Thank you!

