

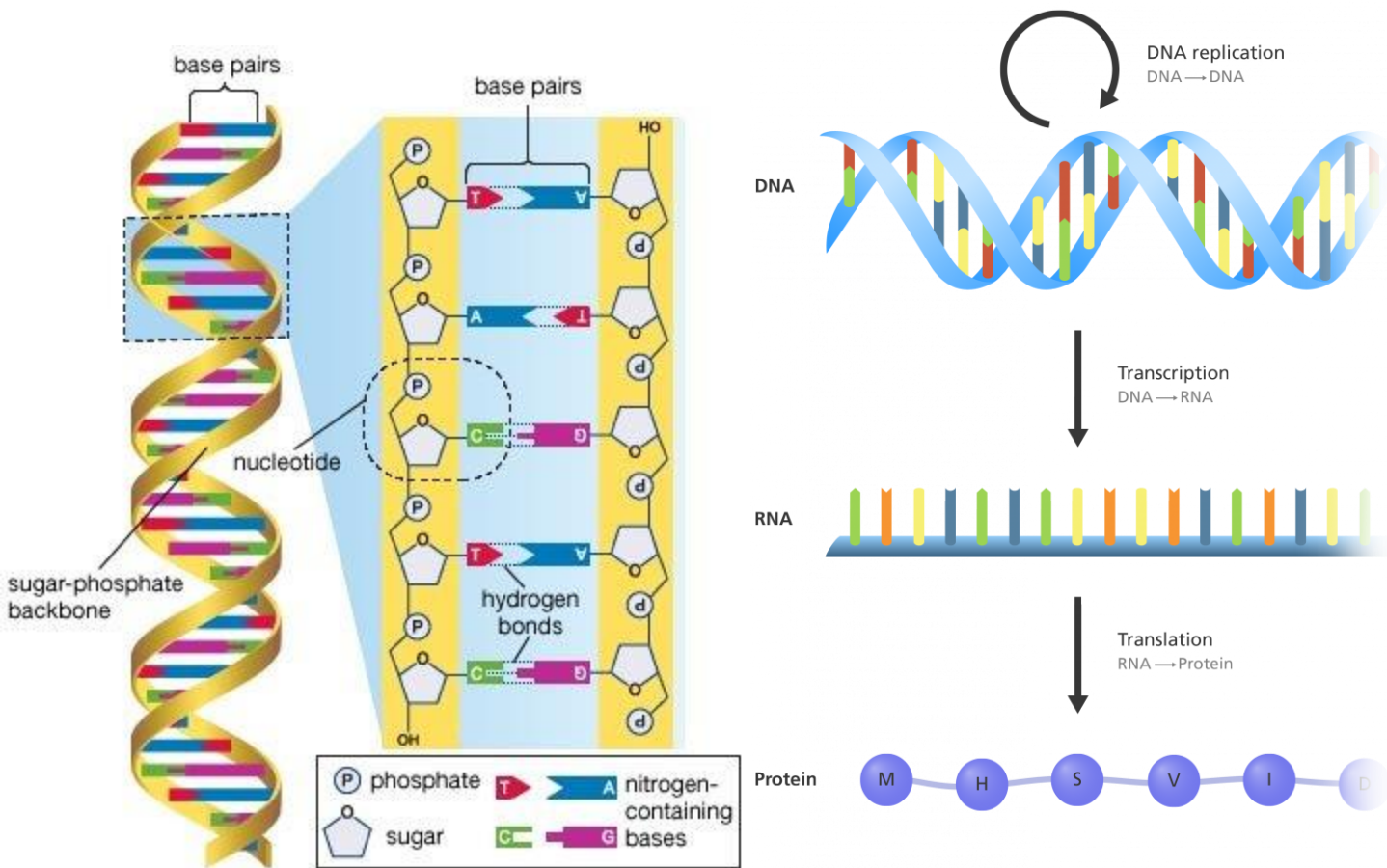


初代DNA测序

第五组：林晓倩、李灵讷、刘旭东、于华新、李文曦

2021.10.9

DNA是主要的遗传物质



- “中心法则”是将 DNA 中的碱基排布转化为功能性产物的过程。它于 1958 年由 DNA 结构的发现者 Francis Crick 首次提出。
- DNA 序列携带了生命的密码，从父母传递给后代，但在当时，人们根本不知道 DNA 序列是什么

- 70年代末, Walter Gilbert发明**化学修饰法**、Frederick Sanger发明**双脱氧终止法**测序
- Walter Gilbert和Frederick Sanger也因在测序技术中的贡献获得了1980年诺贝尔化学奖。

The Nobel Prize in Chemistry 1980

"for their contributions concerning the determination of base sequences in nucleic acids"



Walter Gilbert



Frederick Sanger

A new method for sequencing DNA

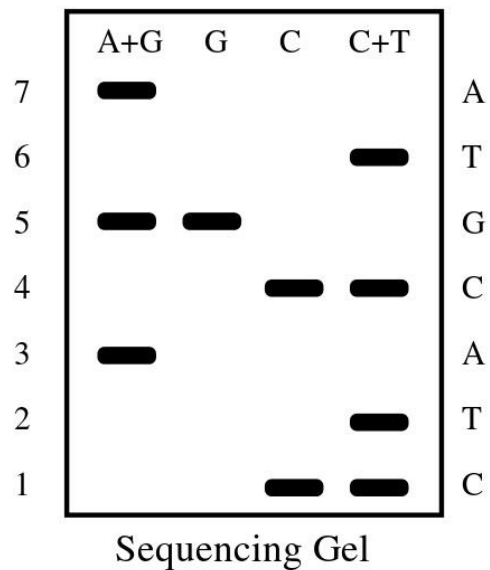
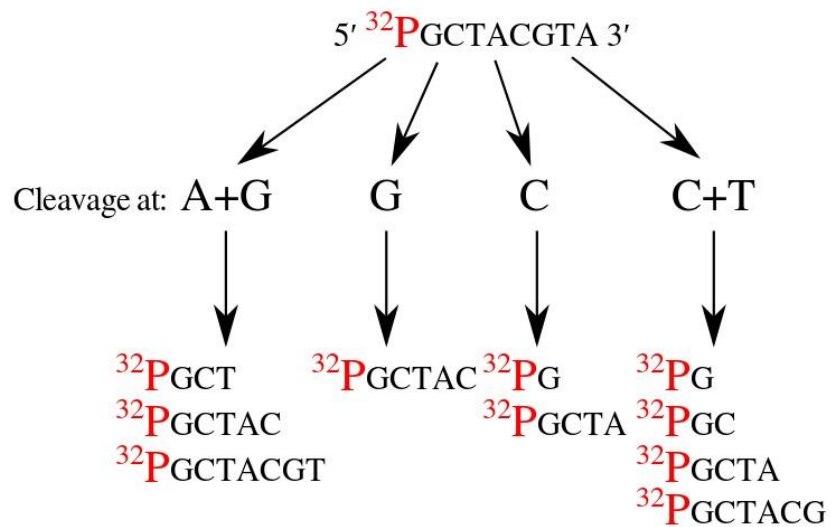
(DNA chemistry/dimethyl sulfate cleavage/hydrazine/piperidine)

ALLAN M. MAXAM AND WALTER GILBERT

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

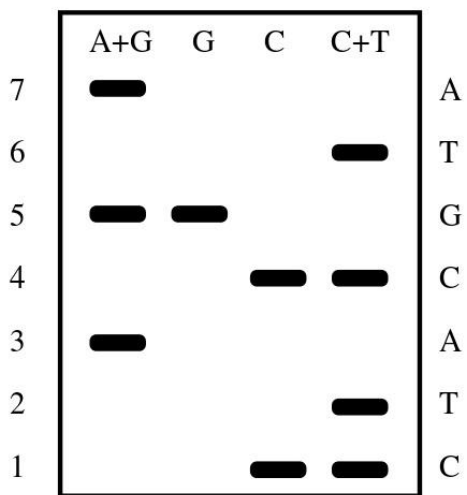
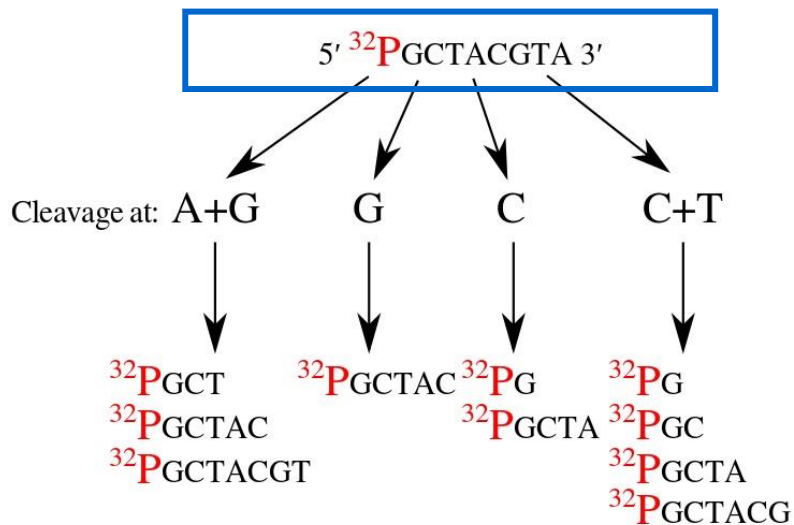
Contributed by Walter Gilbert, December 9, 1976

基本原理



1. 标记DNA末端
2. 分别用几种方法让DNA在指定碱基处断裂，形成各个长度大小的DNA片段
3. 通过聚丙烯酰胺凝胶电泳，可以看出每个条带的碱基，并且从最远的条带开始往上看，就能得到从标记的DNA末端开始的碱基序列

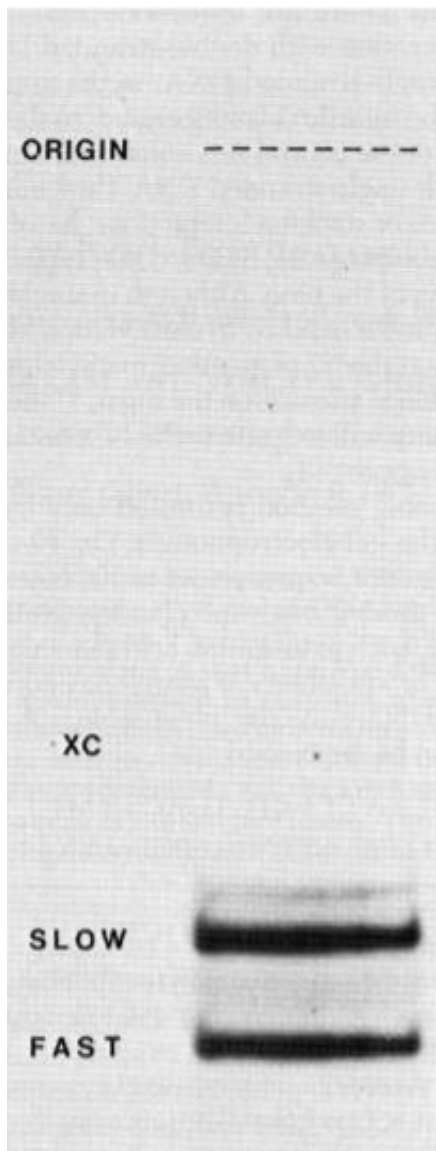
• 同位素标记DNA末端



Sequencing Gel

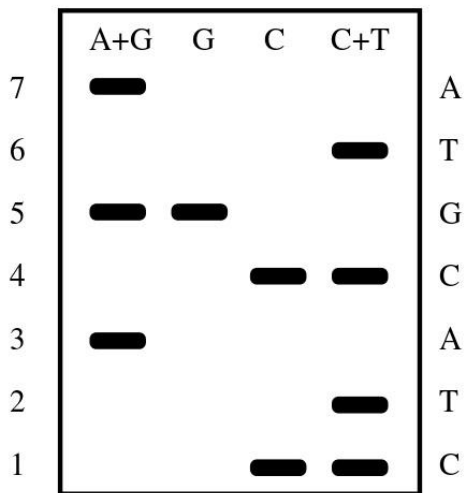
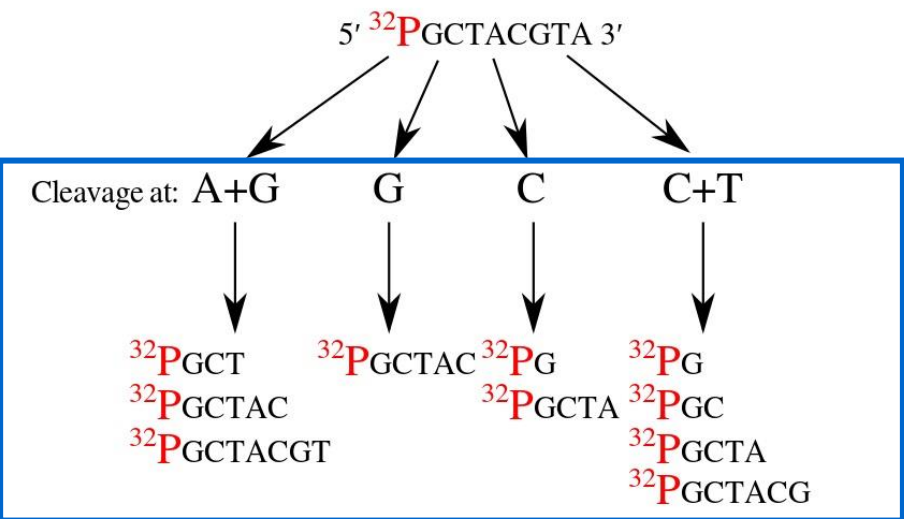
1. 首先攻击DNA，从五碳糖上移走一个碱基，此时暴露的五碳糖容易破裂，分别通过碱催化的β-消去反应和胺催化的β-消去反应可以分别断裂3' 和5' 端的磷酸基团
2. 用³²P标记DNA一条链的末端，可以是3' 端或5' 端（γ-³²P ATP的激酶反应）

- DNA双链的分离



- 以从 *Arthrobacter luteus* 的 *lac* 操纵子中截取的64-碱基对的DNA片段为例
- 标记后的片段经过水解，在电泳中分成了两条链。将两条链提取，进行后续四种断裂反应。

• 四种DNA断裂的化学反应

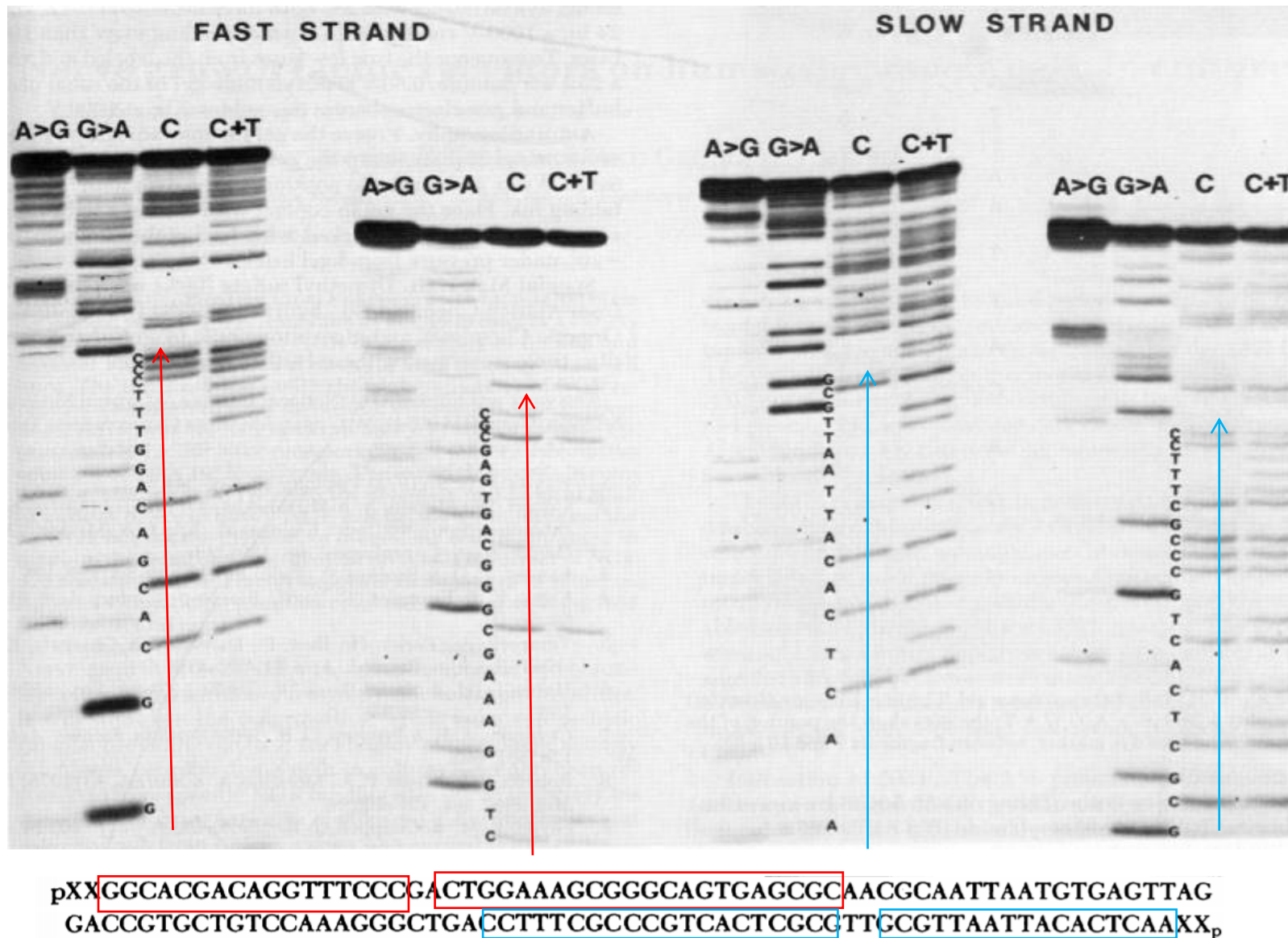


Sequencing Gel

碱基体系	化学修饰试剂	化学反应	处理方法	断裂部位
G>A	硫酸二甲酯	甲基化 (G的N7位置和A的N3位置)	中性pH下加热	G和A
A>G	硫酸二甲酯	甲基化 (G的N7位置和A的N3位置)	稀释的酸轻柔处理DNA	A和G
C+T	胍	打开嘧啶环	0.5 M 胍啶	C和T
C	胍+NaCl	打开胞嘧啶环	-	C

化学修饰法

- 碱基序列的读取



- 1、优点：化学处理容易控制；每一个碱基被击中的概率是相等的，因此每个条带都能显示出来，且不同碱基的反应之间都比较清晰。
- 2、单链和双链的DNA分子是等效的，这个测序方法仅受到凝胶电泳所能达到的分辨率的限制。

DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage ϕ X174)

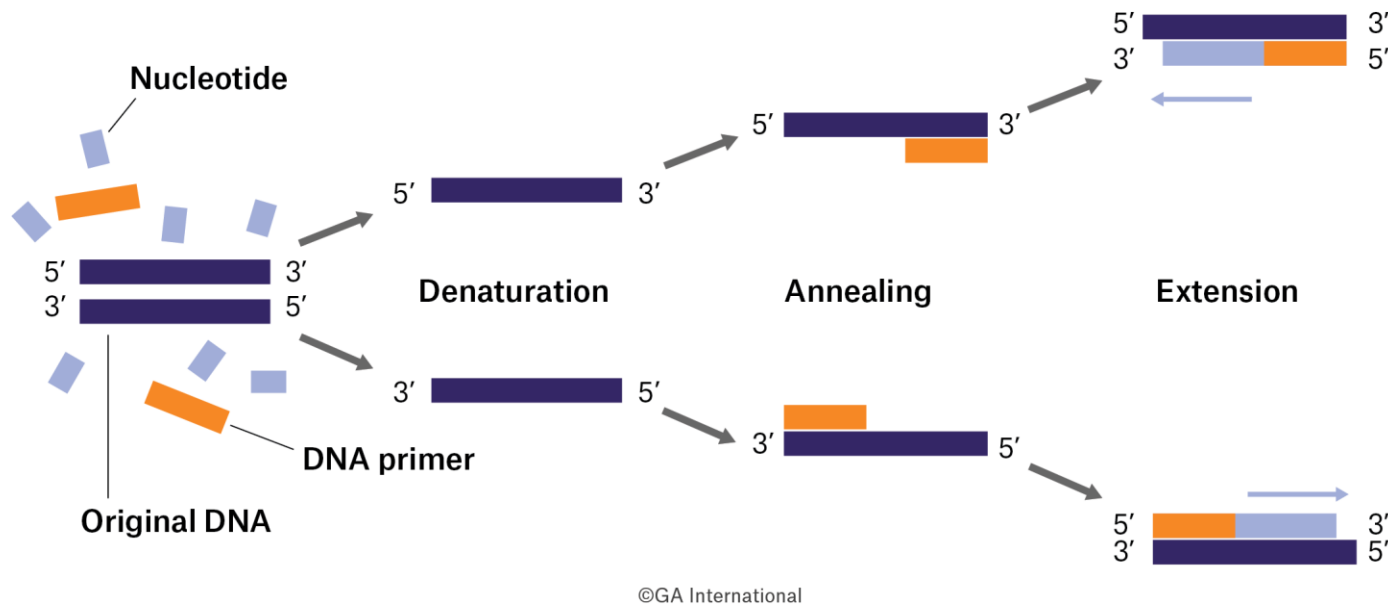
F. SANGER, S. NICKLEN, AND A. R. COULSON

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Contributed by F. Sanger, October 3, 1977

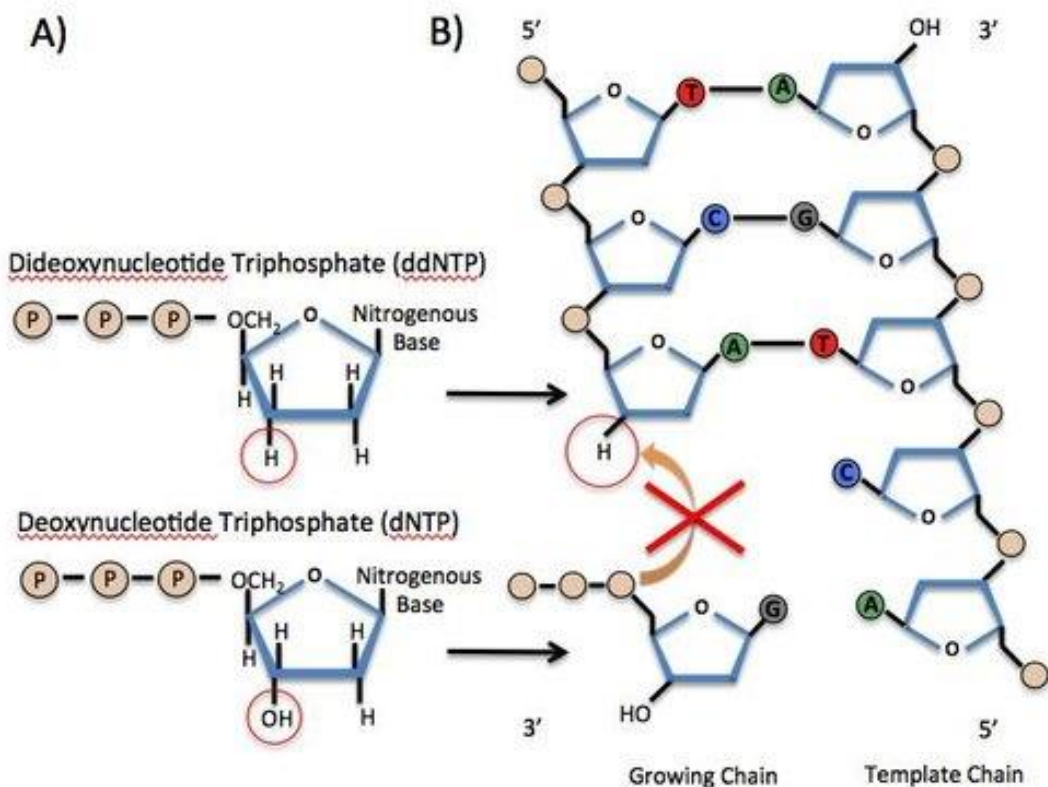
- 这个方法的创新点在于应用了2' , 3' -双脱氧核苷酸。
- 该技术已应用于噬菌体X174的测定——历史上第一个DNA序列。

- 基本原理-PCR



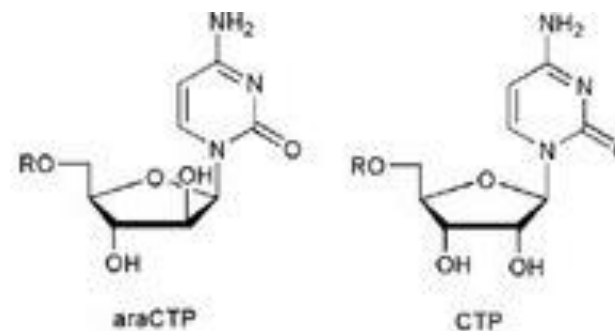
Sanger 测序法

基本原理-双脱氧末端终止



先前研究：

- 2', 3'-二脱氧三磷酸胸苷 (ddTTP) 对DNA聚合酶 I 的抑制活性取决于其与胸苷酸 (dT) 的结合
- 因为ddT不含3' -OH, 所以链不能进一步延伸, 从而在dT应结合的位置处发生终止
- 阿拉伯糖基核苷酸 (ara) 作为大肠杆菌DNA聚合酶 I 的链终止抑制剂, 其作用方式与ddT相当
- 为了获得读取合适的谱带, 必须具有合适的终止核苷酸与正常核苷酸的比率, 以便仅发生终止核苷酸的部分掺入。对于ddNTP, 该比率约为100, 而对于ara, 该比率约为5000。

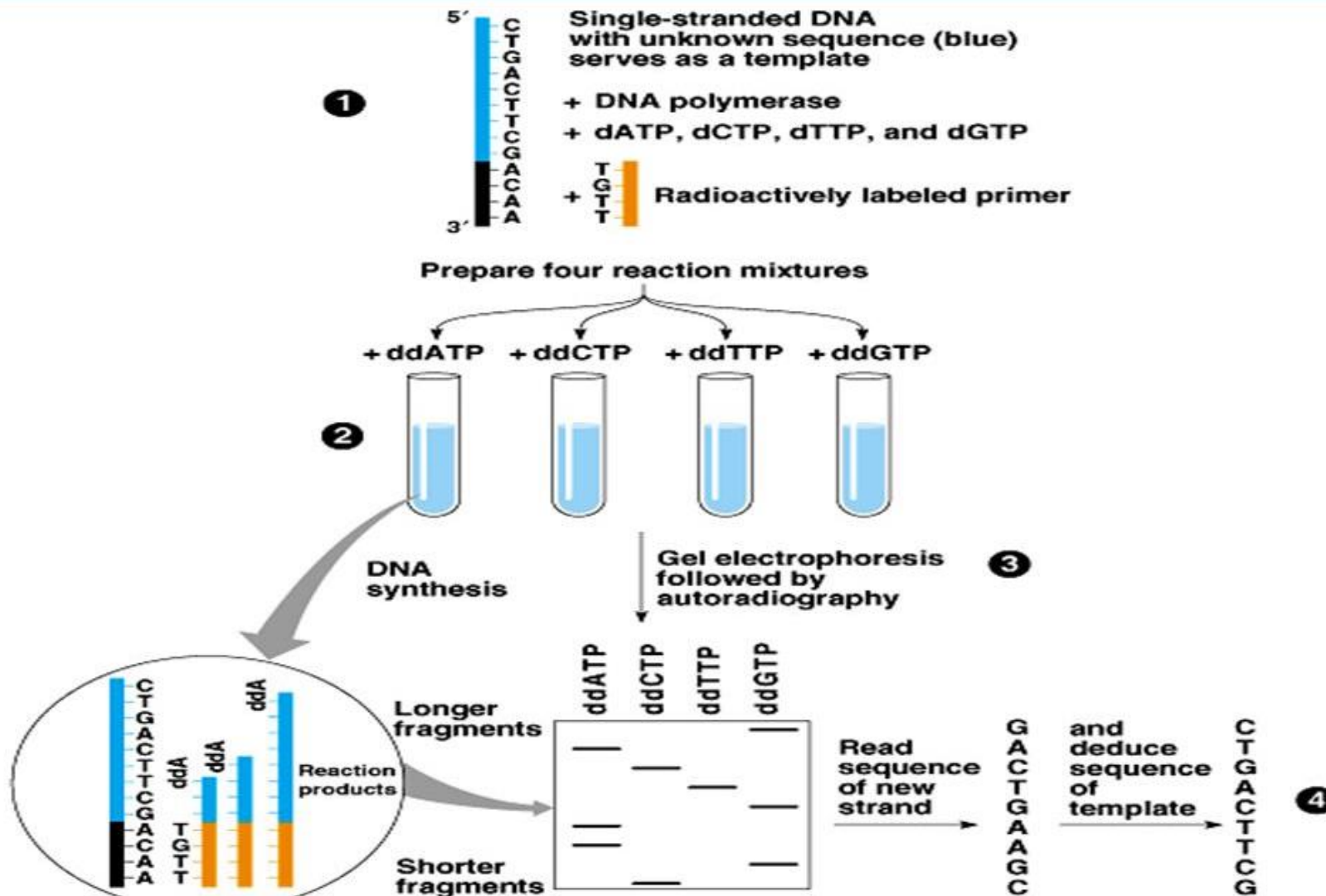
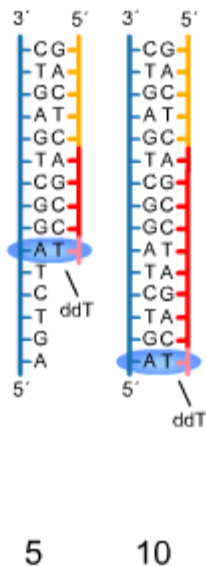


Sanger测序法

• 基本原理



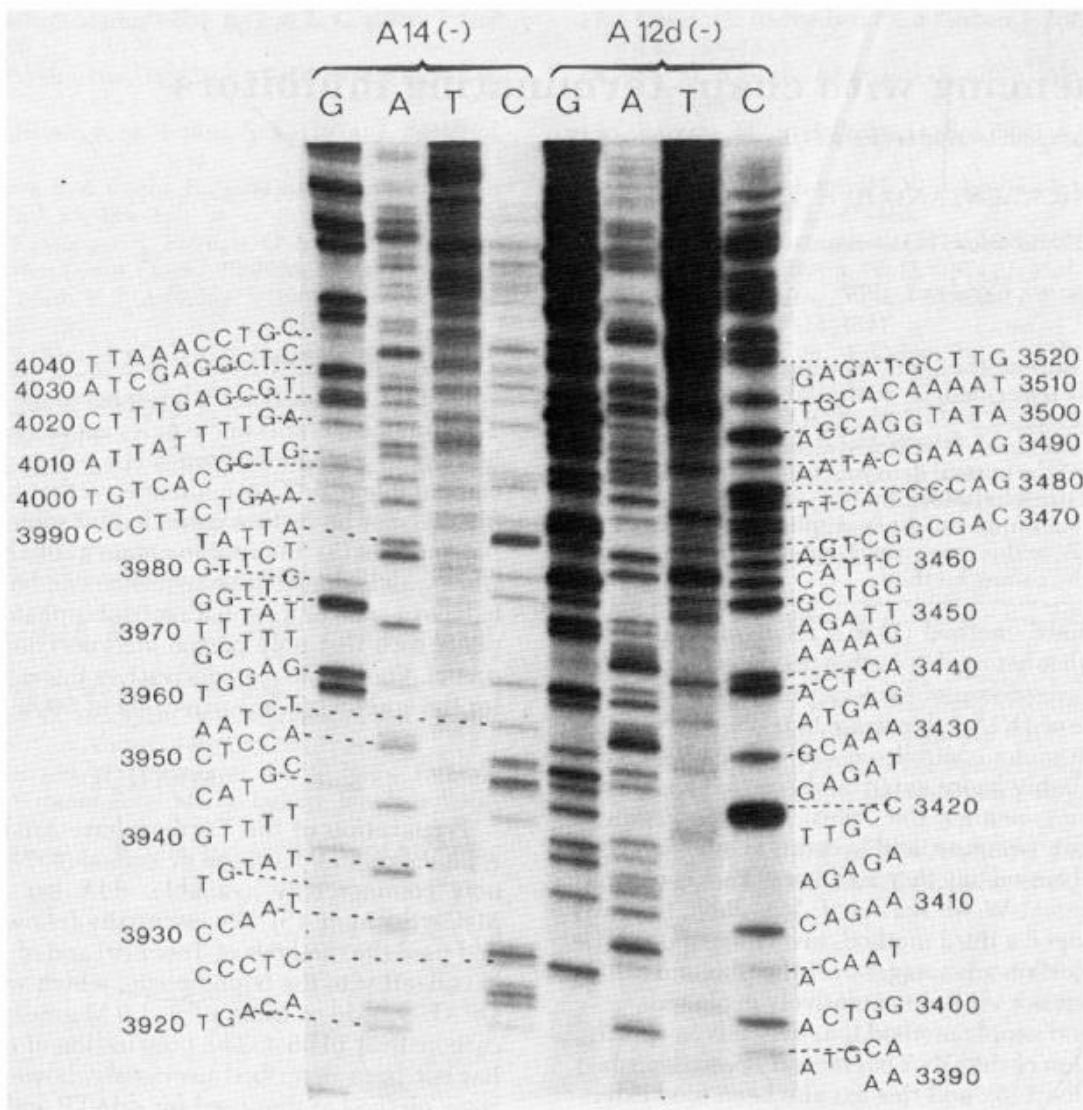
ddT reaction



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Sanger测序法

- 引物的选择-小限制性片段作为引物



- 用两个小的限制性片段A12和A14作为噬菌体X174互补链上的引物
- 使用的抑制剂为（左至右）
ddGTP, ddATP, ddTTP, araCTP
- 读取该序列约80个核苷酸

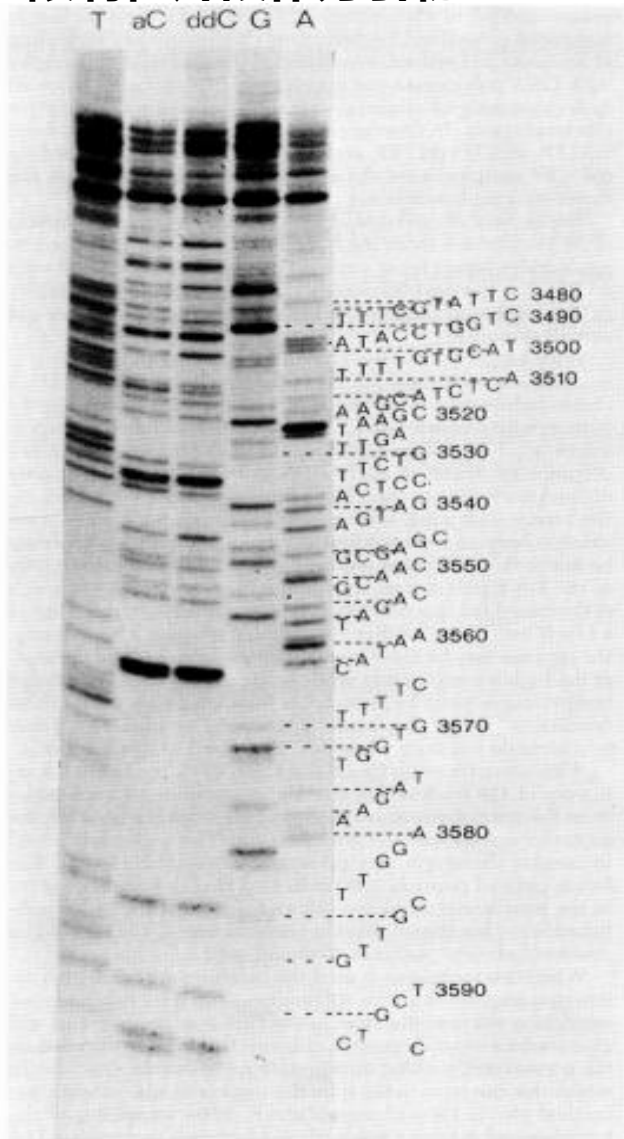
- 引物的选择-R4限制性片段作为引物



- 用较长的片段R4作为噬菌体X174互补链上的引物
- 这次的终止子只有ddCTP而不是araCTP
- 读取120个核苷酸的序列

Sanger测序法

- 引物的选择-A8限制性片段作为引物



- 以片段A8作为引物
- 使用的抑制剂为（从左到右）
ddTTP, araCTP, ddCTP, ddGTP, ddATP
- 从启动位点开始约110个核苷酸的序列可被读取。

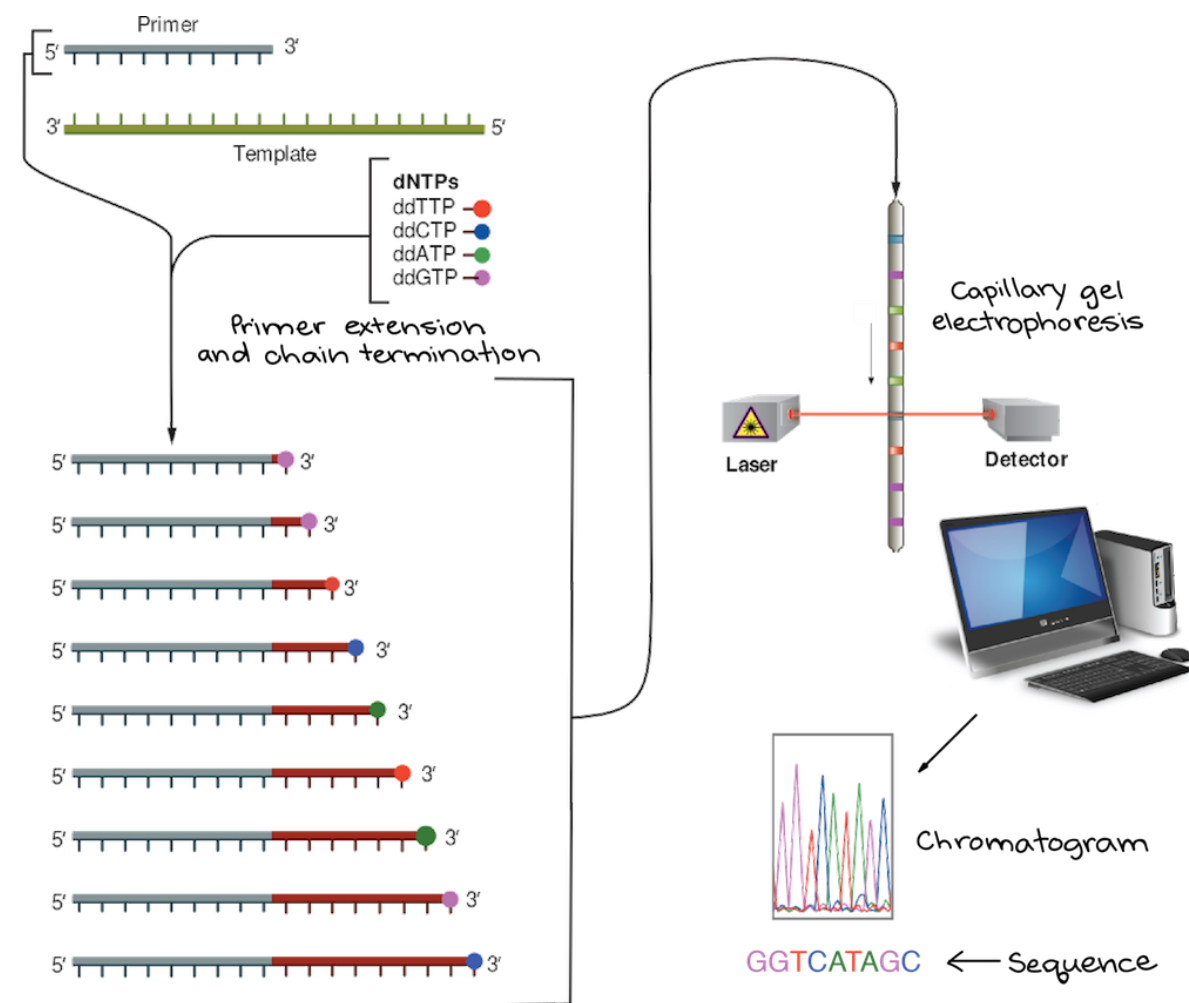
两种方法比较：

(随着新技术的出现，Maxam–Gilbert方法逐渐消失)

方法	相对优点	相对缺点
化学修饰法	对未经克隆的DNA片段可以直接测序	步骤多，各步骤耗时，难以排查错误； 测序片段长度有限，没有改进空间； 需要用到大量放射性物质，包括已知的神经毒素
Sanger测序法	步骤简单 试剂较为安全	需要扩增后产生单链DNA

为什么二代、三代测序出现了，而Sanger测序还在使用

- 随着技术的发展，生物素、荧光标记物等化学发光物质代替了对人体有危害的放射性同位素；凝胶电泳也被毛细管电泳法取代，实现了测序流程的**自动化**
- 测序读长长**，能达到800-1k bp
- 且**测序用时短**，只需要几十分钟即可完成一次测序
- 测序**准确度高**，准确性高达99.999%，目前仍是测序的金标准
- 缺点是通量低、成本高，影响了其真正大规模的应用



Sanger实际应用

- PCR产物测序:

对目的基因的PCR产物进行测序, 得到目的基因序列;

- 重测序:

突变、SNPs、插入或缺失克隆产物的验证;

- 分型分析:

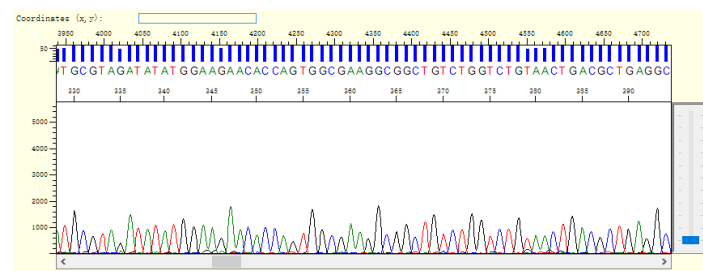
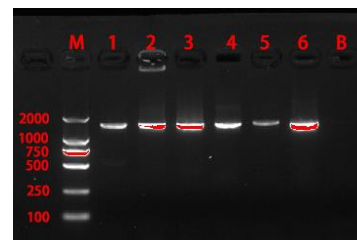
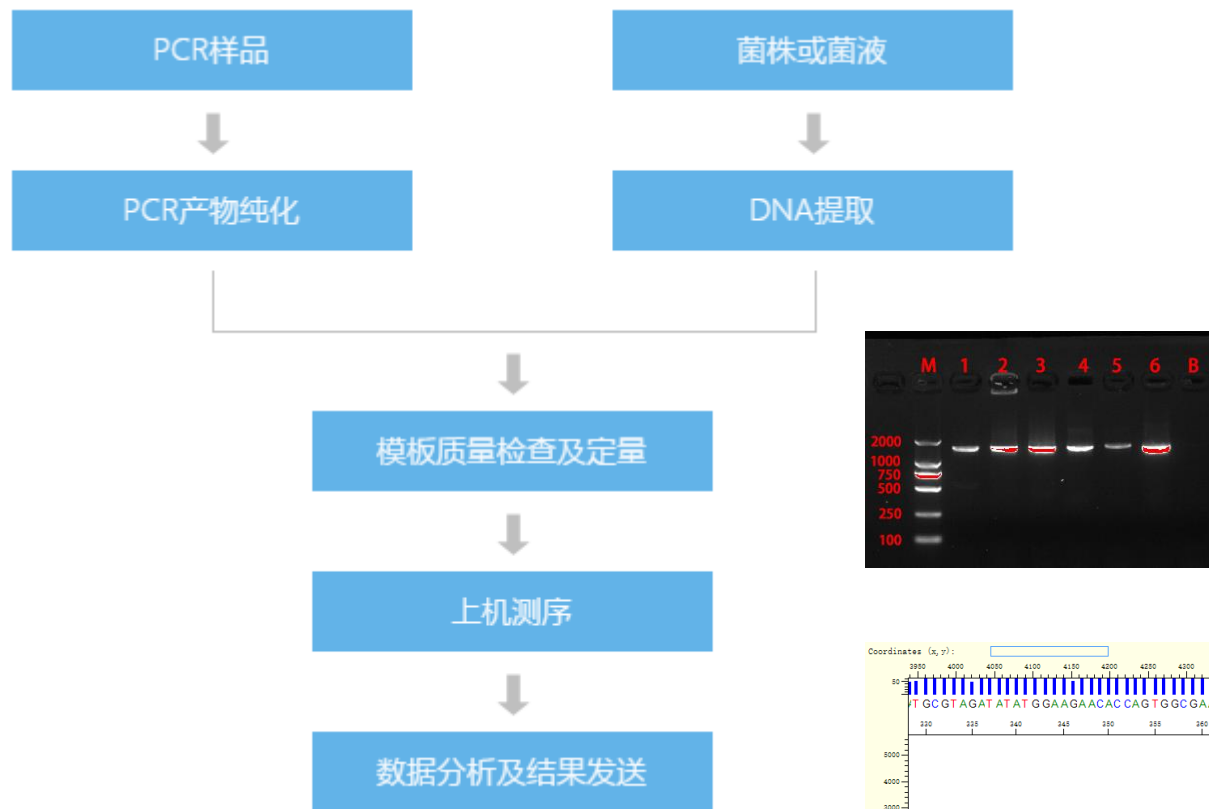
微生物和真菌分类学鉴定、HLA分型、病毒分型等;

- 临床应用:

肿瘤突变基因的检测和肿瘤个体化治疗, 致病基因位点明确并且数量有限的单基因遗传病检测;

- 对新一代测序技术的结果进行验证

Sanger测序拓展



峰图：

- 四色图谱
- 每种颜色对应一种碱基

Sanger测序拓展

序列质量:

- Base QV

$$QV = -10 \times \lg(pe)$$

Pe代表碱基读取的错误概率

QV10 : Pe=10%;

QV20 : Pe=1%;

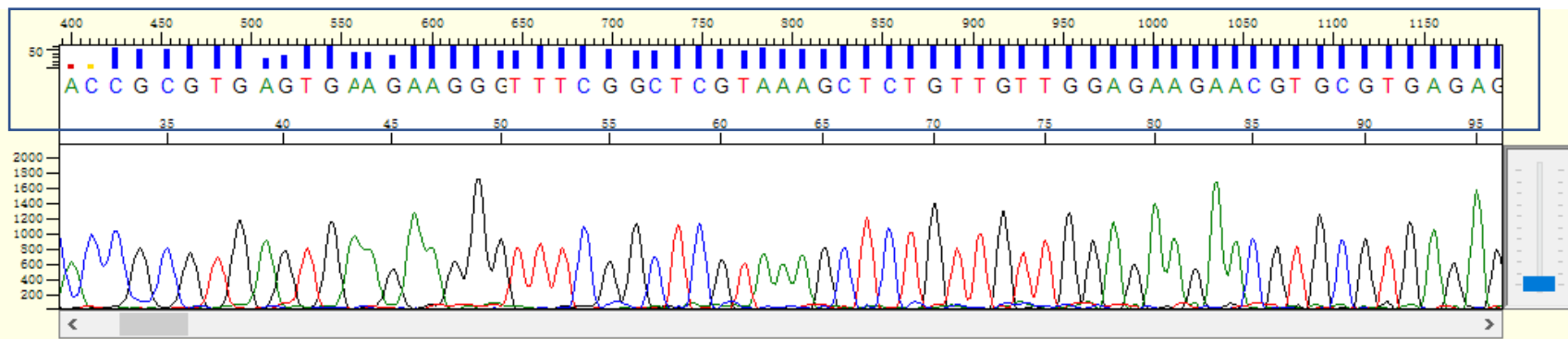
QV30 : Pe=0.1%

QV40 : Pe=0.01%

1	GGGGTCAATG	GGCGCAGCCT	GATGAGACAC	ACCGCGTGAG	TGAAGAAGGG	TTTCGGCTCG	TAAAGCTCTG	TTGTTGGAGA	AGAACGTGCG	TGAGAGTAAC	100
101	TGTTACGCA	GTGACGGTAT	CCAACCAGAA	AGTCACGGCT	AACTACGTGC	CAGCAGCCGC	GGTAATACGT	AGGTGGCAAG	CGTTATCCGG	ATTTATTGGG	200
201	CGTAAAGCGA	GCGCAGGCGG	TTGCTTAGGT	CTGATGTGAA	AGCCTTCGGC	TTAACCGAAG	AAGTGCATCG	GAAACCGGGC	GACTTGAGTG	CAGAAGAGGA	300
301	CAGTGGAACT	CCATGTGTAG	CGGTGGAATG	CGTAGATATA	TGGAAGAACA	CCAGTGGCGA	AGGCGGCTGT	CTGGTCTGCA	ACTGACGCTG	AGGCTCGAAA	400
401	GCATGGGTAG	CGAACAGGAT	TAGATACCTT	GGTAGTCCAT	GCCGTAAACG	ATGAGTGCTA	GGTGTGGAG	GGTTTCCGCC	CTTCAGTGCC	GGAGCTAACG	500
501	CATTAAGCAC	TCCGCCTGGG	GAGTACGACC	GCAAGGTTGA	AACTCAAAGG	AATTGACGG					559



0~50



影响测序质量的主要因素：

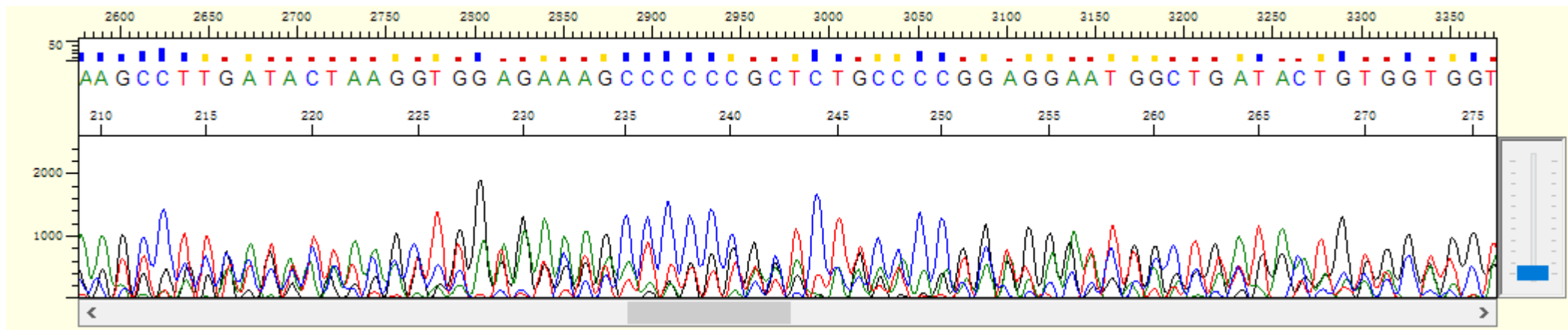
- DNA模板（最主要）

有多套模板、模板未纯化（片选）、模板有多个引物结合位点

- 引物

引物降解、引物与模板特异性结合不强

- 测序试剂、仪器
- 纯化方法
-



THANKS

OMICS FOR ALL

基因科技造福人类