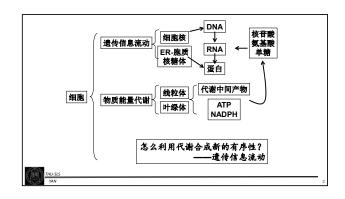
现代生物学导论

V 基因的秘密A

(书上第五章)

闫永彬 Yong-Bin YAN, Ph.D. 清华大学 生命科学学院



〈基因的秘密〉两次课的主要内容

- 1 基因的发现(逻辑重要)
- 2 基因和基因组 (重要)
- 3 DNA半保留复制(中学知识自行复习)
- 4 RNA的组成和作用 (中学知识自行复习)
- 5 遗传密码的破译 (了解, 自学)
- 转录及蛋白质的合成 (重要)
- 7 基因表达调控 (概念重要)
- 8 DNA相关生物技术 (了解)

基因的秘密 系列科学问题

- 谁通过什么实验发现了遗传基本单位? 谁给了基因这 个名字?
- 怎么去鉴定基因的物理化学本质?
- 知道了基因是什么以后,怎么去探究基因怎么工作的?
- 知道了基因怎么工作的。怎么去应用、设计和改造?
- 为啥说基因的发现很伟大?

经典遗传学

- Mendel
 - ■遗传单位(等位基因)、纯合子、杂合子 ■分离、自由组合、连锁互换、伴性遗传
- ▶ Mendel遗传学定律准确地反映了一些有性生殖过程中遗传 性状的传递规律,但并不能代表所有遗传因子表现的性状 及遗传规律。
- ▶ 遗传规律具有多样性,Mendel遗传学定律的延伸和变化

5.1 基因的发现

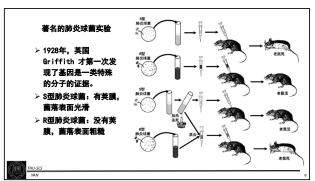
- ▶ 1828,1831,1838——,1859,1865——1909年,丹麦学者约翰逊将孟德尔的 "遗传因子"正式改称为Gene,是 "开始"、"生育" 的意思。中文的 "基因"得益于谈家桢院士的生动音译。
- 十九世纪六十年代——孟德尔同时代人瑞士化学家米歇尔(Miescher)发现细 胞核含有一种以前不知道的含磷丰富的物质——核酸。
- ▶ 1924——福尔根 (F. Feulgen) 证明了DNA是染色质的一个主要组分。
- ▶ 1920s—Morgan (1933 Nobel Prize) 摩尔根发现了染色质就是基因的载体!
- ▶ 1930s之前——生物化学以及生理学的发展使人们逐渐认识到生物体所有的性 状和生物体内代谢过程中的生化反应密切相关,而生物体内各种各样的酶在在 这些生理生化过程当中起着关键作用。So…

■ 基因的发现

- > 在20世纪的前40年,困扰科学家的两个最基本的问题依 然没有解决:
- (1) 基因是由什么物质组成的?
- (2) 基因是如何工作的?
- > 在Mendel和Morgan时代,使用的实验材料主要是豌豆和 果蝇等,它们都是一些非常复杂的多细胞生物。
- > 后来, 在对细菌和病毒这些极其简单的生命形式的研究 中, 科学家才发现了遗传物质的踪影。

■ 沃森《双螺旋》-

- 著名理论物理学家薛定谔写的《生命是什么?》一书。 这本书非常清楚地提出一个信念。即基因是活细胞 的关键组成部分, 以及懂得什么是生命, 必须知道 基因是如何发挥作用的。
- 薛定谔写这本书时(1944年),人们普遍认为基因是特 殊类型的蛋白质分子。但是, 几乎与此同时, 细菌 学家艾弗里 (0. T. Avery) 正在纽约洛克菲勒研究所 进行实验,实验表明,遗传的特性能够被纯化的DNA 分子由一种细菌传递给另一种细菌。



結果说明: (转化, transformation) 加热杀死的s型肺炎球菌 中一定有某种特殊的生物分子或遗传物质,可以使无害的R型 肺炎球菌转化为有害的s型肺炎球菌(为什么?)

> 这种生物分子或遗传物质是什么呢?

- > 纽约洛克非勒研究所 Avery
- 从加热杀死的S型肺炎球菌将蛋白质、核酸、 多糖、脂类分离出来,分别加入到无害的R型肺炎球菌中;
- > 结果发现,惟独只有核酸可以使无害的R型 肺炎球菌转化为有害的S型肺炎球菌。
- > 1944年Avery正式得出结论: DNA是生命的遗 传物质



Fig. 8.2. O. T. Avery (I

更有说服力的噬菌体实验

- > 1952年, Hershey 和 Chase 病毒(噬菌体)
- ➢ 放射性同位素³⁵S标记病毒的蛋白 质外壳,³²P标记病毒的DNA, 然 后分别感染细菌。
- 新复制的病毒,检测到了^{32P}标记的DNA,没有检测到³⁵S标记的蛋白质。
- ▷ DNA在病毒和生物体复制或繁殖 中的关键作用。
- > 8年的时间



, 0 1 85.51

>结论: 生命的遗传物质是DNA。

■ DNA双螺旋结构是结构分析和 遗传学概念相结合的光辉典范, 它开辟了分子生物学的新纪元。

DNA分子是由两条脱氧核糖核酸长 链互以碱基配对相连而成的螺 旋状双链分子。

由于碱基互补配对的高度精确性, 在细胞分裂前DNA复制的时候, 可以使贮藏在DNA分子中以4种 核酸碱基编码的遗传信息得以 稳定的向下一代传递。

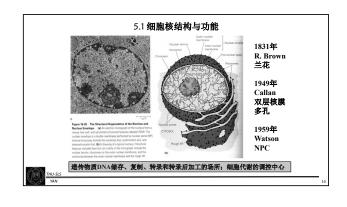


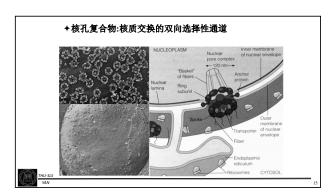
THU-SLS

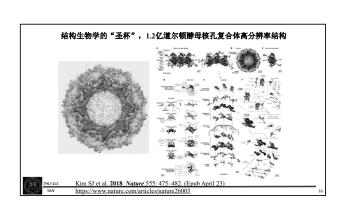
■ 1958年沃森&克里克提出了从DNA到蛋白质的"中心法则"假说;

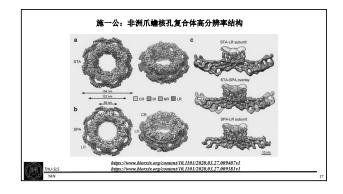
- 1958年美国科学家梅塞森通过DNA的同位素标记实验,证实了DNA的半保留
- 1960年,布伦纳通过一系列实验,证明了mRNA的存在和mRNA传递遗传信息的机制;
- 1961年,法国科学家莫诺提出了操纵子模型,开创了基因调控研究的先河;
- 1965年,霍利首次分离了tRNA,并阐明了tRNA的序列与结构;
- 1968年,科拉纳等人破解了遗传密码并阐明了其在蛋白质合成中的作用。
- 至此,从基因的化学本质到基因控制生物体性状的机制,一套完整的理论 体系被逐步建立了起来,同时也宣告了分子生物学的诞生。

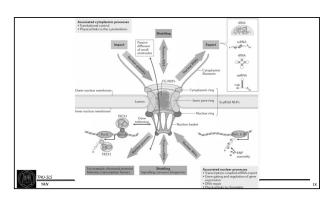
THU-S

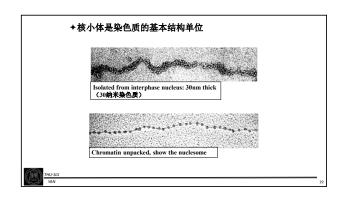


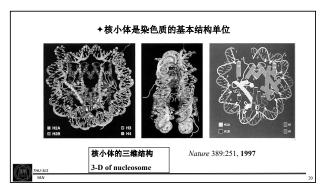


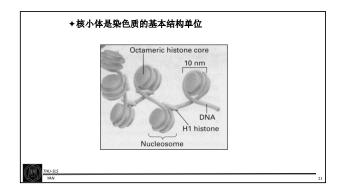


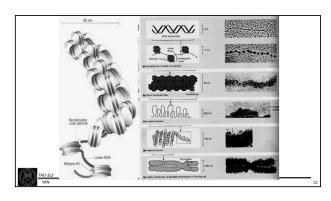


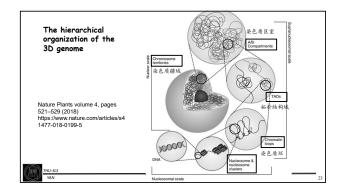












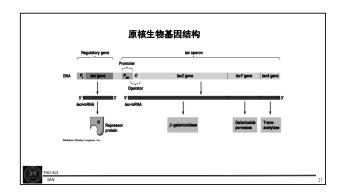
5.3 基因和基因组
■ DNA和基因的概念 (复习)
> DNA分子是由两条脱氧核糖核苷酸(脱氧核苷酸)长链互以碱基配对相连而成的螺旋状双链分子,是遗传的物质基础,遗传信息的携带者。在真核生物中DNA主要存在于细胞核内(线粒体和叶绿体中也有很少量DNA)。
> 基因是贮存遗传信息的特殊DNA片段,它编码蛋白质的氨基酸序列,从而决定蛋白质的功能。通过蛋白质的作用,DNA实际上控制着细胞和生物体的生命过程。

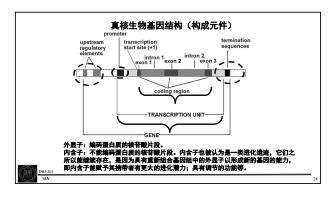
_

- DNA/基因的空间结构 (复习) 是由许多顺序排列的核苷酸单体组成的大分子,具有三
- DNA的一级结构是指4种核苷酸的连接和排列顺序,代表 了DNA分子的化学组成,也影响了DNA的高级结构;
- DNA的二级结构是指两条多核苷酸长链以反向平行盘绕而 成的双螺旋状结构,它又包括三种构象,即B-DNA、 A-DNA及 Z-DNA:
- DNA三级结构是DNA的高级结构,是指DNA双螺旋进一步扭 曲盘绕所形成的特定空间结构,又分为正超螺旋和负超 螺旋,它们之间可以互换,但以负超螺旋为主。

■ 基因的构成(功能性的)

- > 原核生物基因没有内含子,而且其序列中功能相关的RNA 和蛋白质基因往往成串排列在基因组的一个或几个特定 的部位,形成一功能单位或转录单元,它们可以在一个 启动子的控制下被一起转录为多个mRNA(多顺反子)。
- > 真核生物基因由上游的调节元件(upstream regulator elements)、启动子(promoter)、 起始位点 (transcription start site) 、<u>外显子</u>(exon) 、内 含子(intron)以及终止子序列(termination sequence)





The Nobel Prize in Physiology or Medicine





Split genes were independently discovered by Richard J. Roberts and Phillip A. Sharp in 1977, for which they shared the 1993 Nobel Prize in Physiology or Medicine.

- 基因组是生物体内遗传信息的集合,是某一个特定物种 细胞内部全部DNA分子的总和。基因组学从总体的角度 解析生物体整个基因组的全部遗传信息。
- 从整体上看,人类个体的基因是相同的,人类只有一个 基因组。由于不同的人可能拥有不同的等位基因,因此 决定了人与人之间个体上的差异。这种遗传性差异被称 为"单核苷酸多态性"(SNP)。
- 人类基因组主要是指核基因组,因此人类基因组包括了 24条染色体上的全部基因。

Comparison of the number and types of proteins



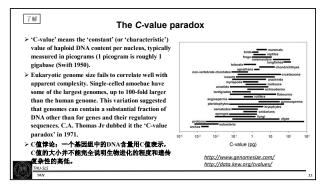
人类基因组的结构

人类绝大多数细胞为二倍体, 即细胞核内有23对、总数为 46条染色体。其中22对为常 染色体,每对都有相同的2条 染色体:另一对为性染色体, 性染色体有X染色体和Y染色 体两种。人类基因组主要是 指核基因组,因此人类基因 组包括了24条染色体上的全 部基因。

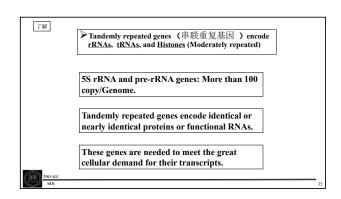
encoded in the genomes of different eukaryotes

is not directly related to its biological complexity

一般来说,基因组的大小与生物的复杂度正相关,但并不一定是定量关系

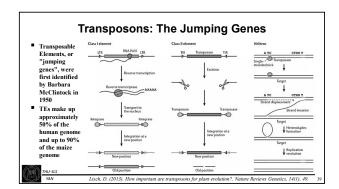




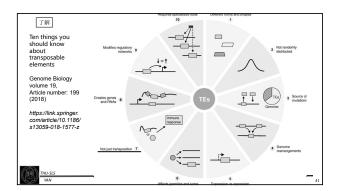


了解 Tandemly repeated DNA串联重复序列: Simplesequence, Highly repeated, Noncoding Satellite DNA: Centromeres, Telomeres 卫星DNA由大量串连重复的短序列(长度约5至几百bp)形成含有几百万DNA碳基对 Microsatellite DNA: 序列量短的(长1-5bp),聚集成50-100bp的小熊、蒙卫星均匀地散在于基因组中,人类 基因组中至少有30000个不同的蒙卫星基因座; 特定小卫星基因座的长度在群体中是高度可 变的,即使在同一家族的成员中也是如此, Minisatellite DNA: 小卫星序列长约12-100bp,成簇分布,多达3000次重复.用于DNA指纹技 特定小卫星某些基因座的改变可能会引起癌症和糖尿病

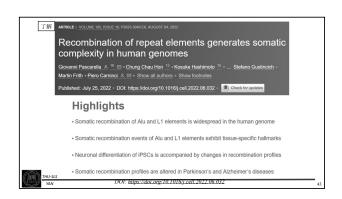
了解 DNA fingerprinting depends on differences in length of minisatellite or microsatellite DNA The restriction fragments from the two individuals a separated on a gel, and then detected with a piece of radioactive RNA complementary to the repeat sequen Because these two individuals have different numbers of repeal between these otherwise identical genes, their DNA produces different patterns on the gel. If enough repeats are separated and tested in this way, the pattern is unique. 了解 DNA精准匹配应该被推广吗? 你的DNA数据还属于你吗? • 也许在不久的将来,就能从公司同事残留在水杯上的 口水中, 探查出他们的身世秘密, 虽然听起来很荒谬 , 但已经有公司正在收集公寓大楼的狗粪, 以查明是 哪些主人没有清理他们家狗狗的排泄物。 • 除了法律之外,个人几乎没有办法保护他们的基因数 据。即使从未进行过DNA测试,但是在这种基因检测技 术发展之下,只需要一滴唾液就能识别出包括本人在 内的三代表兄弟的DNA数据。 • 美国已经开始采集在美一些华人的DNA数据。 · 也许下一代身份证里会储存你的DNA数据。

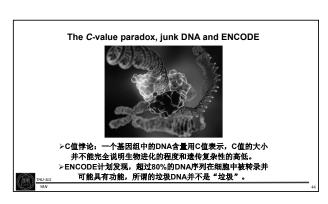


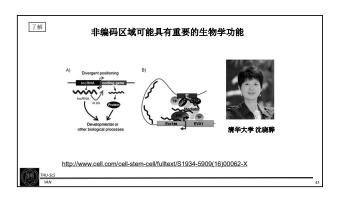




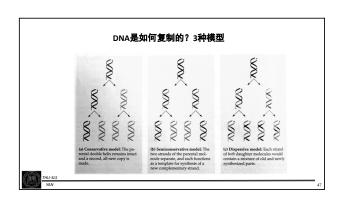


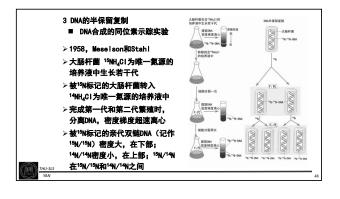


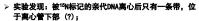












- > 繁殖后第一代大肠杆菌的DNA离心后也只有一条带 (?), 分布于离心管中部;
- > 繁殖后的第二代大肠杆菌DNA离心后出现两条带(?,?)。 -条分布于离心管中部,另一条分布于离心管上部, 证明新合成的DNA分子的两条多核苷酸链中有一条来自 亲代DNA, 一条则是新合成的。

DNA的复制是以亲代的一条DNA为模板,按照碱基互补的原则,合成另一条具有互补碱基的新链,因此,细 胞中DNA的复制被称为半保留复制。



② 单链附着蛋白稳定 解旋的母链DNA

冈崎片断 (Okazaki fragment): DNA半不 连续复制形成的DNA片

■ DNA的复制总是由5'向

■ DNA的半保留复制保证了

所有的体细胞都携带相 同的遗传信息,并可以

将遗传信息稳定地传递

3'方向进行

给下一代。

先导锋 滞后链

■ DNA的半保留复制过程

▷ DNA的复制发生在细胞周期的S期,在解旋酶的作用下, 首先双螺旋的DNA可以同时在许多DNA复制的起始位点 局部解螺旋并拆开为两条单链,如此在一条双链上可 形成许多"复制泡",解链的叉口处称为复制叉。







4 RNA的组成和作用

- RNA结构特点:
- (1) RNA大多是单链分子;
- (2) 含核糖而不是脱氧核糖;
- (3) 4种核苷酸中,不含胸腺嘧啶(T),而是由尿 嘧啶(U)代替了胸腺嘧啶(T)。

细胞中主要有3种RNA, 即信使RNA (messenger RNA, mRNA), 核糖体RNA (ribosome RNA, rRNA) 和转运RNA (transfer RNA, tRNA) .

此外,细胞内还有大量的非编码RNA。

■ mRNA, 即信使RNA

- mRNA是遗传信息的携带者。线形单链结构,在细胞核中 转录DNA上的遗传信息,再进入细胞质,作为指导合成蛋
- 真核细胞mRNA的5′端有甲基化结构 (帽子),抗水解, 并作为蛋白质合成的起始因子识别位置,也保证前体 mRNA的正确剪切;
- 有前导序列,用于定位的功能;
- · 3'含有polyA尾结构,有利于进入细胞质以及提高稳定性。
- 除了编码区(开放阅读框, ORF),还含有非编码区调控 基因的表达



■ tRNA. 即转移RNA

- tRNA局部为双链,在 3′、5′端相反一端的 环上具有由3个核苷酸 组成的反密码子, 在蛋 白质合成时与mRNA上互 补的密码子相结合。 tRNA起识别密码子和携 带相应氨基酸的作用。
- Science, 2013: 神经 退行性病变与tRNA有关

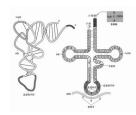
5 遗传密码的破译 ■ 遗传密码的概念

> mRNA上每三个核苷酸翻译成蛋白质多肽上的一个氨基 酸。这三个核苷酸就称为密码子。也叫三联子密码。

> 当蛋白质合成时,遗传密码子的翻译是从起始密码子

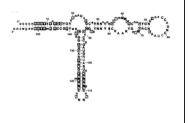
开始(AUG), 沿着 mRAN 5' \rightarrow 3'方向连续阅读, 直

到终止密码子(UGA)出现,生成一条多肽链-蛋白质。



■ rRNA, 核糖体RNA

- > rRNA和蛋白质共同组成 核糖体,即蛋白质合成的场所。在原核生物和 真核生物细胞中rRNA的 种类以及核糖体的组成 不同。其功能是作为 NA的支架,使m 子在其上展开,实现蛋 白质的合成。
- ≻ rRNA占RNA总量的82%左





- ▶ 遗传信息是如何储藏在4种核苷酸中的?
- > 如何破逢遗传察码?
- > 数学家、物理学家-辑运算或推导,不能够得 到结果
- > 分子生物学家──实验验 证得到



> 1955年 1955年 纽约大学 Grunberg-Manago 将核苷酸连接起来的酶 形成RNA聚合体 分寫得到 A连接成多 聚A (polyA, A-A-A-A-A-A) polyC polyG polyU polyAU

问题: 什么样的核苷酸组合可以被翻译成多肽片段?

> 1960年 Matthei 31岁 德国人 美国国家健康研究所 老板 33岁的Nirenberg 烟草花叶病毒 试管中合 成多肽 将ATP和游离的氨基酸加入到从细胞中提取的核 糖体、核酸和酶的混合物中

问题:哪一种RNA可促进多肽的合成?

> 对RNA高度敏感及时检测多肽合成的试管实验系统 在试管中加入了ATP、游离的氨基酸、酶和核糖体及核糖体 RNA----没有蛋白质的合成

问题:需要其他带有遗传信息的RNA?

▶ Matthei, Nirenberg列出200多种RNA,烟草花叶病毒RNA 神秘的蛋白质合成成功了!

蛋白质体外翻译(in vitro protein translation)

≻ Marianne Grunberg-Manago方法人工合成RNA 加入不同的酶、核糖体、ATP、氨基酸 加入poly U、poly A、poly AU poly U产生了许多蛋白质

问题: Matthei, poly U主要利用了哪些氨基酸呢?

▶ 不同的氨基酸分别加入 到poly U试管系统中 5天通宵达旦 星期六早 **晨,熬红了眼的Matthei** 得到了答案: poly U合 成的肽链全部是苯丙氨 酸 (Phe)。

> Matthei成为世界上破译 第一个遗传密码的人。

问题:几个U决定一个苯丙 氨酸的合成?

> Matthei 是世界上首次破译遗传密码的人

≻ 1966年, Nirenberg和Khorana

61个负责20种氨基酚翻译。

诺贝尔奖

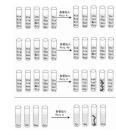
Nobel Prize

in Physiology or Medicine 1968

全部遗传密码字典 64个密码子

≻ Nirenberg 和 Khorana 1968年

3个无义密码子



| Watch | Watc

> Nirenberg 莫斯科 第五届国际生物化学大会 不善于推销自己 小组会上 Meselson认为非同小可 全体大会上重新做学术报告 Francis Crick

问题: 几个U决定一个苯丙氨酸的合成?

- ➢ Nirenberg全力组织其他遗传密码的破译 Matthei回德 国 Nirenberg跟他人合作,发现并定义了3个核苷酸为一个密码子 决定一个氨基酸的翻译
- 按需要连接任意核苷酸 > Khorana ACACACACACACACAC thr-his-thr-his锌

ACA----苏氨酸的密码子 CAC-组氨酸的密码子

胜链 Thr His Thr His

■ 遗传密码的性质

- 1. 密码的简并性:蛋白质存在20种氨基酸,但却有61 种密码子,所以许多氨基酸有多个密码子,只有色 氨酸有一个密码子。因此,称由一种以上密码子编 码同一个氨基酸的现象为简并性(degeneracy),对 应于同一氨基酸的密码子称为同义密码子 (synoymous codon);
- 2. 密码的普遍性与特殊性: 遗传密码子普遍适用于病 毒、细菌、动物以及植物,包括体内和体外。说明 其十分保守。但目前也发现了两个例外,在支原体 种,终止密码子UGA被用来编码色氨酸;在嗜热四膜 虫中,另一个终止密码子UAA被用来编码谷氨酰氨;

4. 密码子的方向性: 是按5'→3'方向编码、不重 复、无标点的三联体密码子。若插入或删除一个 或者两个核苷酸,就会使读码框发生移位,称为

3.	具有起始密码子和终止密码子: 原核生物大部分
	使用AUG为起始密码子,少数使用GUG;真核生物
	则全部使用AUG为起始密码子,而以UAA
	(TAAvsTTC), UAG (TTGvsCAA), UGA (TGAvsTCA)
	三种均作为终止密码子,有时连用两个密码子,
	以确保有效的终止肽链的合成;

移码突变。

Griffith和Avery通过著名的肺炎球离实验提出了DNA 是遗传物质,Hershey和Chase通过噬菌体实验证实了这一 结论。 DNA分子是由两条脱氧核糖核酸长链互以碱基配对相 连而成的螺旋状双链分子,是遗传的物质基础。可以变性

连而成的螺旋状双链分子,是遗传的物质基础。可以变性和复性。
DNA的复制以亲代的一条DNA为模板,在DNA聚合酶的作用下,按照碱基互补的原则,由5′向3′方向合成另一条具有互补碱基的新链,复制的DNA子链与亲代双链完全相同(半保留复制)。
细胞中主要有3种RNA。mRNA是遗传信息的携带者,在细胞质作为蛋白质合成的模板。tRNA起着识别密码子和携带相应氨基酸的作用。rRNA与蛋白质共同组成的复合体就是核糖体,即蛋白质合成的场所。

homework!

见网络学堂

下节课主要内容:基因的秘密B (书5.3-5.4, 第十二章部分内容)