

精要速览系列

Instant Notes

GENETICS (THIRD EDITION)

遗传学 (第三版)



· 导读版 ·

Hugh Fletcher, Ivor Hickey
& Paul Winter



科学出版社
www.sciencep.com



精要速览系列
Instant Notes in

Genetics

Third Edition

遗 传 学

(第三版, 导读版)

G.I. Hickey

Science Department,
St. Mary's University College, Belfast, UK

H.L. Fletcher

School of Biology & Biochemistry,
The Queen's University of Belfast,
Belfast, UK

and

P. Winter

Department of Haematology,
Belfast City Hospital, Belfast, UK

科 学 出 版 社

北 京

内 容 简 介

“精要速览系列(Instant Notes Series)”丛书是国外教材“Best Seller”榜的上榜教材。该系列结构新颖,视角独特;重点明确,脉络分明;图表简明清晰;英文自然易懂,被国内多所重点院校选用作为双语教材。

第三版在第二版基础上进行修订。对人类基因研究、RNA 调控、组蛋白修饰、表观遗传学、人类进化等新近研究进行重点补充和调整,其他各章节也进行了修订。

本书适合普通高等院校生命科学、医学、农学等相关专业使用,也可作为双语教学参考教材使用。

G. I. Hickey, H. L. Fletcher, P. Winter

Instant Notes in Genetics, 3rd edition

© 2007 by Taylor & Francis Group

ISBN0-4153-7619-x

All Right Reserved. Published by arrangement with Taylor & Francis Books Ltd, 2 & 4 Park Square, Milton Park, Abingdon, OX14 4RN, UK.

Licensed for sale in the Mainland of China only, booksellers found selling this title outside the Mainland of China will be liable to prosecution. Copies of this book sold without a Taylor & Francis sticker on the cover are unauthorized and illegal.

本授权版本图书仅可在中国大陆范围内销售,中国大陆范围以外销售者将受到法律起诉。本书封面贴有 Taylor & Francis 防伪标签,未贴防伪标签属未获授权的非法行为。

图书在版编目(CIP)数据

遗传学=Genetics; 导读本: 英文/(英)弗莱彻(Fletcher, H. L.)等主编. —3版. —北京: 科学出版社, 2009

(精要速览系列)

ISBN 978-7-03-025227-2

I. 遗… II. 弗… III. 遗传学-双语教学-高等学校-教材-英文 IV. Q3

中国版本图书馆 CIP 数据核字(2009)第 141412 号

责任编辑: 单冉东 / 责任校对: 张 琪

责任印制: 张克忠 / 封面设计: 耕者设计工作室

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

双 青 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

1999 年 3 月第 一 版 开本: 787×1092 1/16

2009 年 8 月第 三 版 印张: 27 3/4

2009 年 8 月第一次印刷 字数: 623 000

印数: 1—4 000

定价: 48.00 元

(如有印装质量问题, 我社负责调换)

导读编译者

谢 雍(香港科技大学生物学系)

张 博(北京大学生命科学学院)

高宇英(北京大学生命科学学院)

黄 鹏(北京大学生命科学学院)

肖 安(北京大学生命科学学院)

高 巍(北京大学生命科学学院)

徐琳杰(北京大学生命科学学院)

夏棣丹(北京大学生命科学学院)

梁 巍(北京大学生命科学学院)

王展翔(北京大学生命科学学院)

张 莹(北京大学生命科学学院)

前言

自本书第二版出版以来,人们对遗传学的研究取得了一些重要的进展。虽然基因组测序不再属于前沿领域,但是从中所获得的数据促进了诸多的研究工作。人类基因组仅有 20 000 个基因被鉴定出来,估计上限为 25 000 个基因;这远远低于先前的预测,将人类跟其他大多数的多细胞生物紧紧地联系在一起。人类的基因数目似乎比卷心菜的还要少。基因组比对结果还显示,有 67~83 个在黑猩猩体内依然活跃的基因,由于在进化过程中产生突变而在人类中成为没有活性的假基因。更令人惊异的是,人类基因组的某些部分与黑猩猩的差异比其余部分大得多。这表明当祖先群体向这两个不同的物种趋异进化时经历了很长一段时间,期间发生了一些遗传交换。

另一个新进展是发现 RNA 在控制染色质结构方面的作用,以及组蛋白 N 末端的共价修饰在调节染色体结构方面的重要性。某些组蛋白修饰(例如位点特异性甲基化)能够促使核小体紧密堆积,导致形成致密的染色质,造成转录装置难以接近,从而使其中的基因的表达受到限制。双链 RNA 能够引导特异的 DNA 序列发生这种修饰,使 CpG 序列(在植物中为 CpNpG)中的胞嘧啶被甲基化,并且这种修饰与组蛋白的去乙酰化相关。随着用以鉴定组蛋白特异修饰的单克隆抗体不断增多,人们对所谓的“组蛋白密码”的认识将会越来越清晰。

为了反映上述变化,我们减少了介绍基因组测序技术的篇幅,同时将表观遗传学移到了 A 单元中的“分子遗传学”部分,因为这样看起来更为恰当。人类的进化作为单独的一个单元出现。人类在地球上定居与迁移的地理学越来越清楚,而随着类人猿基因组的信息逐渐明朗,人们才刚刚开始认识人类与类人猿之间在分子水平上的差异。

有很多审稿人对本书的前一个版本提出了评论和修改建议,我们在此表示感谢。在力所能及的范围内,我们尽可能采纳了他们的建议。我们希望至少有一些审稿人对某些改动会感到满意。

PREFACE

There have been some significant changes in our understanding of genetics since the second edition. Genome sequencing is no longer a frontier project, but the data produced is driving most aspects of research. The human genome has only 20,000 genes identified, with an upper estimate of 25,000, putting us firmly in with most other multicellular organisms, far below earlier guesses. We appear to have fewer genes than a cabbage. Genome comparisons have also revealed between 67 and 83 genes that remain active in chimps, but are now inactive pseudogenes in humans due to mutations in the course of human evolution. A further surprise is that some parts of the human genome differ much more than others from chimpanzees. This implies a long period with some genetic exchange between the ancestral populations as they diverged into the two species.

Another new development is discovery of the role of RNA in the control of chromatin architecture, and the importance of covalent modifications to the N-terminal tails of histones in regulating chromatin structure. Some histone modifications (e.g. site specific methylation) promote tight nucleosome compaction and result in densely compacted chromatin which restricts access by the transcription machinery and restricts gene expression. Double-stranded RNA has a role in targeting specific DNA sequences for this modification, which is associated with histone deacetylation, while cytosines in CpG sequences (or CpNpG in plants) are methylated. The so-called "histone code" will become clearer as more monoclonal antibodies are produced to identify increasing numbers of specific histone modifications.

To reflect these changes, we have reduced the coverage of techniques used for genome sequencing, and moved the topic of epigenetics to Section A, Molecular Genetics, because that seemed appropriate. Human evolution has been given its own section. The geography of human colonization of the world is becoming clearer, and the molecular differences between humans and the other great apes are just beginning to appear as the apes' genomes become available.

We are grateful to the many reviewers who commented on the previous edition and the proposed changes. We have tried to respond to their suggestions where practicable and reasonable within the scope of a text of this type. We hope that some of the reviewers may be satisfied by some of the changes.

缩 略 词

3D	three dimensional	三维的
5BU	5-bromouracil	5-溴尿嘧啶
ADA	adenine deaminase	腺嘌呤脱氨酶
AIDS	acquired immune deficiency syndrome	艾滋病/获得性免疫缺陷综合征
ATP	adenosine triphosphate	腺苷三磷酸(三磷酸腺苷)
BAC	bacterial artificial chromosome	细菌人工染色体
bp	base pairs	碱基对
BrdU	bromodeoxyuridine	溴脱氧尿苷
bZIP ^①	basic leucine zipper	碱性亮氨酸拉链
CAP	catabolite activator protein	分解物激活蛋白
cAMP	cyclic adenosine monophosphate	环腺苷酸
cdk	cyclin dependent kinase	依赖细胞周期蛋白的激酶
cDNA	complementary or copy DNA	互补 DNA
CFTR	cystic fibrosis transmembrane conductance regulator	囊性纤维化跨膜传导调节蛋白
cM	centiMorgan	厘摩
DIG	digoxigenin	地高辛
DNA	deoxyribonucleic acid	脱氧核糖核酸
dATP	2'-deoxyadenosine 5'-triphosphate	2'-脱氧腺苷 5'-三磷酸
dCTP	2'-deoxycytosine 5'-triphosphate	2'-脱氧胞苷 5'-三磷酸
ddNTP	dideoxynucleotide triphosphate	双脱氧核苷三磷酸
dGTP	2'-deoxyguanosine 5'-triphosphate	2'-脱氧鸟苷 5'-三磷酸
dNTP	deoxynucleotide triphosphate	脱氧核苷三磷酸
dTTP	2'-deoxythymidine 5'-triphosphate	2'-脱氧胸苷 5'-三磷酸
ds	double stranded	双链
ES	embryonic stem	胚胎干(细胞)
EST	expressed sequence tags	表达序列标签
ETS	external transcribed spacers	外部转录间隔区
F	fertility	致育
FISH	fluorescent <i>in situ</i> hybridization	荧光原位杂交
GDP	guanosine diphosphate	鸟苷二磷酸
GTP	guanosine triphosphate	鸟苷三磷酸
Hfr ^②	high frequency recombination	高频重组
HIV	human immunodeficiency virus	人类免疫缺陷病毒
HLH	helix-loop-helix	螺旋-环-螺旋
HnRNA	heterogeneous nuclear RNA	核内不均一 RNA
HUGO	human genome organization	人类基因组组织
ICR	internal control region	(序列)内部控制区
ITS	internal transcribed spacers	内部转录间隔区
kb	kilo base	千碱基

kbp	kilo base pairs	千碱基对
LINE	long interspersed nuclear elements	长散在核元件
LTR	long terminal repeat	长末端重复
MCS	multiple cloning site	多克隆位点
mtDNA	mitochondrial DNA	线粒体 DNA
mRNA	messenger RNA	信使 RNA
NOR	nuclear organizer region	核仁组织区
ORF	open reading frame	可读框
PAC	P1 artificial chromosome	P1 人工染色体
PCR	polymerase chain reaction	聚合酶链式反应
PKU	phenylketonuria	苯丙酮尿症
pms	postmeiotic segregation	减数后分离
QTL	quantitative trait loci	数量性状基因座
R	resistance	抗性、耐受性
RF	replicative form	复制型
RFLP	restriction fragment length polymorphism	限制性片段长度多态性
RNA	ribonucleic acid	核糖核酸
RNAi	RNA interference	RNA 干扰
ROS	reactive oxygen species	活性氧组分
rRNA	ribosomal RNA	核糖体 RNA
RT-PCR ^①	reverse transcription polymerase chain reaction	反转录-聚合酶链式反应
SAR	scaffold attachment regions	支架附着区
SCE	sister chromatid exchanges	姐妹染色单体交换
SINE	short interspersed nuclear elements	短散在核元件
siRNA	small interfering RNA	小干扰 RNA
SNP	single nucleotide polymorphism	单核苷酸多态性
snRNP	small nuclear ribonucleoproteins	核内小核糖核蛋白
ss	single stranded	单链
SSB	single strand binding	单链结合
STS	sequence tagged sites	序列标签位点
TF	transcription factor	转录因子
tRNA	transfer RNA	转移 RNA
TIC	transcription initiation complex	转录起始复合体
VNTR	variable number tandem repeats	可变数目串联重复
YAC	yeast artificial chromosome	酵母人工染色体

注①:原文为 bZLP。

注②:原文为 HFr。

注③:原文为 RTPCR。

A 分子遗传学

A1 DNA 结构

要 点

核苷酸

DNA 是由核苷酸单体构成的链状聚合物,每个核苷酸由一个糖基、一个碱基和一个磷酸基团组成。DNA 的糖基是 2'-脱氧核糖,含有五个碳原子,分别称为 1'、2'……碳原子。DNA 的碱基分为四种:腺嘌呤和鸟嘌呤含有两个碳氮杂环,属于嘌呤碱;胸腺嘧啶和胞嘧啶只含有一个碳氮杂环,属于嘧啶碱。这些碱基与脱氧核糖的 1'碳原子相连,构成核苷。糖基的 5'碳原子可以与一个、两个或三个磷酸基团相连,构成一个核苷酸。核苷酸能够以单个分子或多聚体(如 DNA 或 RNA)的形式存在。

DNA 多聚核苷酸

分别对应四种碱基的核苷三磷酸相连接而形成了 DNA 多聚核苷酸链。核苷酸聚合时,其中的两个磷酸基团脱落,余下的 5'磷酸基与下一个核苷酸的 3'羟基形成磷酸二酯键。多聚核苷酸的一端为游离的 5'磷酸基(5'端),另一端为游离的 3'羟基(3'端)。遗传信息由碱基的顺序编码,其阅读方向为从 5'到 3'。多聚核苷酸链很长,可以产生 4^n (n 为核苷酸数)种不同的序列。

双螺旋

DNA 分子由两条互相缠绕成双螺旋结构的多聚核苷酸链组成。其中的糖基和磷酸基构成 DNA 分子的骨架,碱基则面向内部重叠排列。两条多聚核苷酸链走向相反。DNA 双螺旋结构是右旋的,每 10 对碱基构成一个螺旋周期。双螺旋结构的大沟可与蛋白质相互作用。目前已发现了各种不同的 DNA 结构,包括具有左旋结构的 Z-DNA。

碱基互补配对

两条 DNA 链上碱基之间形成的氢键维持了双螺旋结构的稳定性。由于受到双链内的空间限制,碱基配对只能发生在嘌呤和嘧啶之间,其中 A 只能与 T 配对,G 只能与 C 配对。这就是碱基互补配对原则。碱基配对的限制意味着两条链上的碱基顺序是相互关联的,一条链的顺序决定了另一条链的顺序。这样能够使遗传信息在 DNA 复制和基因表达的过程中得以维持。加热、化学物质或酶的作用能够导致 DNA 碱基间氢键的断裂,使双螺旋的两条链分开变成单链。

RNA 的结构

在 RNA 中,胸腺嘧啶被尿嘧啶取代,2'脱氧核糖被核糖取代。RNA 通常以单链多聚核苷酸的形式存在,但在互补的序列之间有可能产生短片段的碱基配对。

相关主题

基因转录(A4)

DNA 突变(B5)

DNA 复制(A9)

A2 基因

要 点

基因的结构

基因是遗传信息的基本单位,它对应于 DNA 上一个不连续的区段,编码一个多肽的氨基酸序列。^①人类细胞约含有 25000 个基因,分布于 23 条染色体上。基因是不连续的,被非编码的基因间 DNA 分开。遗传信息由模板链编码,用于指导 RNA 分子的合成。^②两条 DNA 链都能成为模板链。DNA 分子具有庞大的储存遗传信息的能力。

基因家族

有些基因成簇排列,称为操纵子或多基因家族。操纵子存在于细菌中,含有被共同调控、功能相关的基因。多基因家族存在于高等生物中,它含有一组相同或类似的基因,但是这些基因并不受协同调控。简单的多基因家族含有相同的基因,以满足细胞对该基因产物的大量需要。复杂的多基因家族则含有非常类似的基因,这些基因往往编码功能相关的蛋白质。

基因表达

编码于基因中的生物信息通过基因表达而得以表现。在这个过程中,基因首先指导合成 RNA,然后由 RNA 指导合成蛋白质。基因表达的中心法则指出:(遗传)信息总是由 DNA 传递给 RNA,再由 RNA 传递给蛋白质。细胞的功能通过许多蛋白的协同作用而实现。基因的表达保证了蛋白质在合适的位置和合适的时间被合成。

基因启动子

基因表达是受到高度调节的。细胞中并非所有的基因都处于活化状态,不同类型的细胞表达不同的基因。基因的表达受到位于编码区上游的一段 DNA 序列调控,称为启动子。它能够与 RNA 聚合酶以及有关的转录因子相结合,从而启动 RNA 分子的合成。

内含子和外显子

基因中的编码序列被分割成一系列的片段,称为外显子;它被称为内含子的非编码序列分开。内含子通常占据大部分的基因序列,其大小和数量因不同的基因而异。在蛋白质合成之前,内含子将被从 RNA 转录本中去除,这个过程称为剪接。细菌的基因中通常不存在内含子。

假基因

某些基因的拷贝含有在进化过程中获得的序列错误,使它们无法产生蛋白。这些基因被称为假基因,它们代表原始基因的进化遗迹。珠蛋白假基因即为一例。

相关主题

原核生物基因表达调控(A10)

原核生物基因组(B3)

真核生物基因表达调控(A11)

真核生物基因组(B4)

表观遗传学与染色质修饰(A12)

注①:关于基因的概念,更为普遍的看法是:基因是指一段特殊的 DNA,这一段 DNA 能够作为一个完整的功能单位指导合成一条多肽链或一个 RNA 分子。

注②:应为:遗传信息由有义链编码,与之互补的为模板链,该链用于指导 RNA 分子的合成。

A3 遗传密码

要 点

遗传表达

遗传信息由一系列基因的 DNA 碱基序列所编码。基因表达作为一个术语,是用来描述细胞如何将遗传信息解码,以合成细胞活动所必需的蛋白质。基因表达指合成互补的 RNA 链,这条 RNA 的序列决定了蛋白质的氨基酸序列。基因的 DNA 序列与多肽的氨基酸序列呈现线性相关。

遗传密码

20 种氨基酸由 64 种碱基三联体编码,称为密码子。大多数氨基酸具有一种以上的密码子。遗传密码的这种简并性能够减少突变的影响。编码同一个氨基酸的密码子称为同义密码子,它们通常在第三个碱基上有差异,这个碱基位置被称为“摆动”位置。AUG 编码甲硫氨酸,同时也是起始密码子。有三种终止密码子:UAG、UGA 和 UAA。

阅读框

如果可以选择不同位置的碱基开始读码,任何 DNA 序列都可以有三种不同的读码方式。每一种读码方式产生的一套密码子序列称为一个阅读框。起始密码子决定哪一种阅读框真正成为蛋白编码序列,其他阅读框往往包含终止密码子,无法用于指导蛋白质合成。阅读框是一段由起始密码子和终止密码子所限定的范围内的密码子序列。

遗传密码的普遍性

遗传密码普遍适用于所有生物,同样的密码子编码同一个氨基酸。但是,在线粒体基因组和一些单细胞生物中存在一些与标准密码子用法不同的例外。

相关主题

转移 RNA(A5)

翻译(A8)

A4 基因转录

要 点

转录

转录是基因表达的第一阶段,指的是以 DNA 为模板,在 RNA 聚合酶的催化下合成 RNA 的过程。RNA 根据 DNA 模板链合成,具有与非模板链(即有义/编码链)相同的碱基序列。在 RNA 合成过程中,核苷三磷酸的聚合方向为 $5' \rightarrow 3'$,而 DNA 模板链的方向则是 $3' \rightarrow 5'$ 。

原核生物的转录

原核生物的转录可分为三个阶段:起始、延伸和终止。大肠杆菌只有一种 RNA 聚合酶,称为全酶。全酶由五个亚基(两个 α 、一个 β 、一个 β' 和一个 σ 亚基)构成。 σ 亚基可以脱离全酶,而剩余的亚基则构成核心酶。

转录起始

转录从基因的启动子开始。大肠杆菌的 RNA 聚合酶可以识别 -10 区和 -35 区的两段序列, σ 可以结合到 -35 区上。最初形成一个闭合的启动子复合体,然后 DNA 双螺旋在 -10 区发生解离,形成一个开放的启动子复合体。 σ 亚基脱离 RNA 聚合酶后, RNA 合成开始。

转录延伸

在延伸过程中, RNA 聚合酶依照 DNA 模板的序列,把核苷酸依次添加到 RNA 分子的 $3'$ 端。RNA 聚合酶沿着 DNA 链向前移动并破坏碱基间的氢键(熔解),从而使 DNA 双链解旋。为避免 DNA 链上的张力太大,每次只有一小段 DNA 链(12~17 个碱基)会被解旋。新合成的 RNA 最初与 DNA 模板链配对结合,随后即被释放,以便重新形成 DNA 双链。

转录终止

RNA 中的回文序列可形成茎环结构,作为转录的终止信号。一般认为, RNA 聚合酶遇到一个茎环结构后会中止活动,随后茎环结构后面的一段较弱的 A-U 碱基对分开,从而释放出转录产物。另一种情况是因 ρ 蛋白破坏碱基配对而释放转录产物。转录终止以后, RNA 聚合酶从 DNA 模板上释放出来,并与 σ 亚基重新结合。

真核生物的转录

真核生物的转录起始很复杂,其终止并不需要茎环结构。三种不同的 RNA 聚合酶(I、II、III)转录不同的基因。

RNA 聚合酶 II

RNA 聚合酶 II 负责转录编码蛋白质的基因。RNA 聚合酶 II 的启动子序列通常包含一个 TATA 框,大约位于转录起始位点上游 25 碱基对处;TATA 框可以与 RNA 聚合酶 II 结合。相关的转录因子(TFII A、B 等)在 TATA 框附近按一定的顺序与 DNA 结合,为 RNA 聚合酶 II 的结合创造一个平台。缺乏 TATA 框的基因可能会含有一个起始元件,也可能两者都没有。其他启动子元件,例如 CAT 框,可以跟其他转录因子结合,从而影响转录起始的效率。增强子和沉默子等远端调控元件也能够影响转录效率。介导转录终止的信号尚不清楚。

RNA 聚合酶 I

RNA 聚合酶 I 转录 18S、28S 和 5.8S 核糖体 RNA(rRNA)。这类基因的启动子包含两个转录必需的元件。一个是与转录起始位点重叠的核心元件,另一个是在-100bp 附近的上游控制序列。在距离基因 3'端约 600bp 之后的位置有一个 18bp 的终止信号。

RNA 聚合酶 III

RNA 聚合酶 III 转录转移 RNA(tRNA)和 5S 核糖体 RNA(rRNA)等小基因。这类基因的启动子位于编码序列内,因而被称为内部控制区(ICR)。tRNA 基因有两个重要的序列元件,分别称为 A box 和 B box;5S rRNA 基因的转录需要的一段称为 C box 的序列,还有一段称为 A box 的序列也很重要。一段连续的 A 核苷酸残基是转录的终止信号。

相关主题

转移 RNA(A5)	真核生物基因表达调控(A11)
核糖体 RNA(A6)	表观遗传学与染色质修饰(A12)
信使 RNA(A7)	RNA 研究技术(E2)
原核生物基因表达调控(A10)	

A5 转移 RNA

要 点

tRNA 在翻译中的作用

转移 RNA(tRNA)是一类小分子,它以信使 RNA(mRNA)为模板,将氨基酸按顺序转运到一起,从而使蛋白质得以合成。细胞内 tRNA 的种类很多,每种 tRNA 都结合一种特定的氨基酸。每种 tRNA 可以与 mRNA 中特异的密码子结合,从而将其携带的氨基酸放在正确的位置。

tRNA 的结构

tRNA 通过分子内碱基配对形成三叶草形的结构。由碱基配对形成的茎环结构称为“臂”。这些“臂”包括:接纳臂,是氨基酸的结合位点;反密码子臂,识别 mRNA 中的密码子;二氢尿嘧啶(DHU)臂,含二氢尿嘧啶;额外臂和含有假尿嘧啶的胸苷-假尿苷-胞苷臂(TΨC 臂)。tRNA 的碱基配对区具有保守的核苷酸序列。根据三级结构预测的碱基配对情况与二维的三叶草结构类似,并显示接纳臂与反密码子臂分别位于 tRNA 分子的两端。

tRNA 的合成与加工

tRNA 基因以多拷贝的形式存在,每一个拷贝编码多个 tRNA。tRNA 首先由 RNA 聚合酶 III 转录成前 tRNA,之后被核糖核酸酶加工成为成熟的 tRNA。RNA 酶 P 含有核酶活性。转录结束后,CCA 序列被加到真核生物的 tRNA 上。

tRNA 中核苷酸的修饰

tRNA 分子被转录出来后还要进行修饰。修饰的类型包括:甲基化、碱基重排、双键饱和、脱氨、硫酸盐化和添加大基团。这些修饰的功能还不清楚。

相关主题

基因转录(A4)	翻译(A8)
核糖体 RNA(A6)	

A6 核糖体 RNA

要 点

核糖体

核糖体是由核糖体 RNA(rRNA)和蛋白质组成的大分子结构。它们在细胞质中大量存在,以信使 RNA 为模板合成蛋白质。核糖体由特定大小的大小亚基组成,它们的沉降系数(S)不同。原核生物核糖体沉降系数为 70S,由 50S 和 30S 亚基组成,包括三种 rRNA(23S、16S 和 5S),真核生物核糖体沉降系数为 80S,由 60S 和 40S 亚基组成,包括四种 rRNA(28S、18S、5.8S 和 5S)。

原核生物中
rRNA 的转录
和加工

大肠杆菌(*E. coli*)中的三种 rRNA 由一个基因转录而来。在基因组中该基因有 7 个拷贝。该基因首先转录出一个转录本,然后经过加工成为成熟的 rRNA。加工过程包括 RNA 的折叠、与核糖体蛋白质的结合、碱基的甲基化和核糖核酸酶的切割。

真核生物中
rRNA 的转录
和加工

真核生物 rRNA 从一种基因转录而来,该基因为多拷贝,且成簇存在。这些基因在细胞核仁中由 RNA 聚合酶 I 转录。首先合成一个单独的 rRNA 前体,然后经过加工,成为成熟的 28S、18S 和 5.8S rRNA。真核生物 rRNA 的加工包括 rRNA 的折叠、与核糖体蛋白质的结合、由小核糖核蛋白(snRNP)介导的核糖的甲基化,以及在核糖核酸酶作用下间隔序列的去除。5S rRNA 则是在 RNA 聚合酶 III 的作用下,由不连锁的基因转录而来。

相关主题

转移 RNA(A5)
信使 RNA(A7)

翻译(A8)

A7 信使 RNA

要 点

mRNA 的合成
与加工

信使 RNA(mRNA)是蛋白质合成的模板,它是在 RNA 聚合酶 II 的作用下,在细胞核中由蛋白质编码基因转录而来。mRNA 最初被转录成前体分子,称为前 mRNA;该前体包括非编码的内含子序列,在随后的剪接过程这些序列会被去掉。mRNA 的 5'端被“帽子”结构修饰,3'端被多聚腺苷酸修饰。这类 RNA 由 RNA 聚合酶 II 转录,以核内不均一 RNA 分子集合体的形式存在于细胞核中。

mRNA 剪接

mRNA 剪接是指从前 mRNA 中切除内含子的过程。内含子的末端分别为 GT 和 AG 序列,它们分别是 5'和 3'端剪接信号的一部分。另一个信号序列位于内含子内部,称为分支点序列。剪接过程包括内含子的 5'端被切开并被连接到分支点序列上,形成一个有尾环。随后,这个内含子在 3'端发生切割,从 mRNA 前体中被释放出来,相邻的两个外显子被连接到一起。剪接由核内小核糖核蛋白(snRNP)催化;U1 结合到 5'剪接位点上,U2 与分支点序列结合,之后 U5 和 U4/6 跟 U1 和 U2 形成一个复合物,称为剪接体。剪接体将 mRNA 置于正确的取向,并提供切除内含子、连接外显子的酶活性。

mRNA 加帽

真核 mRNA 的 5'端被一种特殊的核苷酸 7-甲基鸟嘌呤修饰。这一核苷酸通过一种不寻常的 5'→5'三磷酸键连接到 mRNA 的第一个核苷酸上。这种修饰被称为“加帽”,它能够保护 mRNA 不被 5'核酸外切酶降解。

mRNA 聚腺
苷酸化

大多数真核 mRNA 的 3'端具有多聚 A 尾巴(聚腺苷酸化)修饰。前 mRNA 在聚腺苷酸化信号(5'AAUAAA3')下游大约 20 个碱基处被切开,随后在聚腺苷酸聚合酶的作用下,加上一段腺苷酸残基。聚腺苷酸化被认为可保护 mRNA 的 3'端不被核酸外切酶降解。

mRNA 的稳定性

相对而言,mRNA 不如核糖体 RNA 和转运 RNA 稳定。这使得细胞能够通过改变基因转录的速率来调节蛋白质的水平。原核生物 mRNA 的半衰期较真核生物短得多。一些真核生物 mRNA 的稳定性受短干扰 RNA(siRNA)的调控,这些 siRNA 结合到 mRNA 上,与蛋白质相互作用,催化 mRNA 的降解。

mRNA 的可变加工

剪接方式的多样性造成一个前 mRNA 产生序列不同的 mRNA,从而产生不同的蛋白质。剪接方式可以包含或不包含一个或多个外显子。肿瘤细胞中常出现 mRNA 的异常剪接。聚腺苷酸化信号的变化也可导致生成不同的 mRNA。mRNA 的序列也可通过 RNA 编辑发生改变,这种编辑包括通过插入、缺失或单个碱基的置换而改变序列。

相关主题

- 基因转录(A4)
- 转移 RNA(A5)
- 核糖体 RNA(A6)
- 翻译(A8)
- 真核生物基因表达调控(A11)
- 基因与癌症(F3)

A8 翻译

要点

转移 RNA 的作用

转移 RNA(tRNA)以信使 RNA(mRNA)为模板,将氨基酸运送到核糖体上,从而合成蛋白质。每种氨基酸可与一种或多种可识别其密码子的 tRNA 结合,这些 tRNA 统称为同工 tRNA(isoacceptor)。在氨酰 tRNA 合成酶的作用下,氨基酸通过氨酰化与 tRNA 接纳臂的末端共价结合。

密码子识别

mRNA 中的密码子与 tRNA 中的反密码子之间的碱基互补配对保证了氨基酸在蛋白质合成过程中被放置在正确的位置。遗传密码具有简并性,大多数氨基酸被一个以上的密码子编码。由于反密码子中的 5'碱基能够结合不同密码子的 3'碱基,因而每个 tRNA 可以识别多个密码子。这就是密码子的摆动性。

翻译

真核生物和原核生物的翻译过程类似,都可分为三步(起始、延伸和终止)。每一步都需要许多辅助性的蛋白质。翻译所需的能量由腺苷三磷酸(ATP)和鸟苷三磷酸(GTP)的水解提供。

起始

在原核生物中,翻译起始于核糖体小亚基与 mRNA 的 SD 序列结合;在真核生物中,核糖体小亚基与 mRNA 5'帽子的结合标志着翻译的开始。小亚基沿 mRNA 迁移至 AUG 起始密码子的下游,起始甲硫氨酸 tRNA($tRNA_i^{met}$)结合上去,形成起始复合物。在原核生物中, $tRNA_i^{met}$ 的甲硫氨酸被甲酰化。IF1、IF2 和 IF3 是细菌的翻译起始因子。在翻译起始步骤完成前,IF1 和 IF3 阻止大亚基的结合;IF2 将 $tRNA_i^{met}$ 引导到起始复合物上。真核生物至少包含 9 种翻译起始因子,其中,eIF1 和 eIF2 的作用与 IF1 和 IF2 相似。一些因子能够打开 mRNA 的二级结构。

延伸

翻译起始步骤完成后,核糖体大亚基结合到起始复合物上,形成 A 位点和 P 位点。P 位点与 $tRNA_i^{met}$ 结合。第二个载有氨基酸的 tRNA 进入 A 位点,在肽基转移酶的作用下,两个氨基酸之间形成肽键。随后,tRNA 脱酰基酶切断甲硫氨酸与 tRNA 之间的化学键,使这个二肽结合在第二个 tRNA 上。在原核生物中,延长因子 EF-Tu 协助 tRNA 进入 A 位点,之后 GTP 水解,EF-Tu 被释放出来,与 GDP 结合。EF-Ts 可再生成新的 EF-Tu。在真核生物中,eEF1 的作用与 EF-Tu 相似。肽键形成之后,核糖体移动到下一个密码子上,与二肽结合的 tRNA 移至 P 位点,代替起始 tRNA,第三个携带氨基酸的 tRNA 进入 A 位点,如此进行延伸步骤的循环。在原核生物中,核糖体的移动需要 EF-G 的介导以及 GTP 的水解供能;真核生物中的 eEF2 具有类似的作用。

终止

终止密码子进入 A 位点标志着翻译过程的结束。释放因子进入 A 位点,导致多肽的释放。在大肠杆菌中,RF1、RF2 和 RF3 负责终止;在真核生物中,只需要 eRF 一个蛋白质。翻译终止后核糖体分离,释放出 mRNA。

翻译后修饰

翻译完成后,多肽可被修饰,例如在氨基酸侧链上、C 末端或 N 末端加上各种化学基团,或被蛋白水解酶切割。修饰有可能是某些蛋白质活化所必需的。

相关主题

遗传密码(A3)
转移 RNA(A5)

核糖体 RNA(A6)

A9 DNA 复制

要 点

DNA 复制

细胞在分裂前其 DNA 要进行自我复制。复制过程是以单链 DNA 为模板,在 DNA 聚合酶的作用下以 5'→3' 的方向进行。DNA 复制是半保留式的。在大肠杆菌中,DNA 聚合酶 I 和 III 具有 3'→5' 核酸外切酶的活性,这些酶因而能够对序列进行校正,以保证 DNA 复制过程中极低的错误率。

复制叉

DNA 合成发生在复制叉处。解旋酶分离双螺旋,单链结合蛋白(SSB)维持两条单链处于分离状态。DNA 的合成在先导链中是连续的,在后随链上则合成不连续的片段(冈崎片段)。在大肠杆菌中,由 DNA 聚合酶 III 参与的 DNA 合成起始于由引发酶 RNA 聚合酶合成短的 RNA 引物。随后,引物序列再由 DNA 聚合酶 I 置换为 DNA。在真核生物中,DNA 聚合酶 α 首先作为引物酶起始 DNA 合成,随后,先导链和后随链分别由 DNA 聚合酶 δ 和 α 合成。冈崎片段在 DNA 连接酶的作用下通过磷酸二酯键相互连接。

细菌 DNA 的复制

环状细菌 DNA 分子和线性真核生物染色体的复制方式不同。环状 DNA 分子从单一的起点复制,复制叉向两个方向前进,最后相遇并合并。环状 DNA 分子在双螺旋解旋过程中形成的超螺旋可以被拓扑异构酶 I 消除。复制过程中产生的子代分子会互相缠绕在一起,拓扑异构酶 II 可以将它们分开。

真核生物 DNA 的复制

真核生物的细胞分裂称为细胞周期,包含一系列的事件。真核生物染色体为多起点复制,先形成复制泡,最后相遇、合并。转录活跃的区域首先被复制。复制前 DNA 需要从核小体上解离。染色体的末端需要以特殊机制复制,端粒酶将非编码序列加到染色体末端以便使末端得到复制。

相关主题

DNA 结构(A1)

真核生物基因组(B4)

染色体(B1)

DNA 突变(B5)

原核生物基因组(B3)

诱变剂和 DNA 修复(B6)

A10 原核生物基因表达调控

要 点

基因表达调控

细菌可以调控其基因的活性,从而仅产生细胞活动所必需的基因产物。这使得细菌能够对环境变化作出反应。调控转录、mRNA 更新、mRNA 加工或翻译等的速率都能够改变基因产物的数量。其中改变基因转录的机制研究得最清楚。

细菌基因的结构	细菌的很多基因往往排列成协同调控的操纵子,这些基因往往编码功能上相关的蛋白质。可诱导的操纵子,例如乳糖操纵子,编码参与代谢途径的酶,并可被该途径的底物诱导。可抑制的操纵子,例如色氨酸操纵子,编码参与生物合成途径的酶;这些操纵子的基因表达受到该生物合成途径的最终产物或弱化作用的调节。	
乳糖操纵子	该操纵子包括三个基因(<i>lac Z,Y,A</i>),编码大肠杆菌乳糖代谢所需的酶。这些基因由同一个启动子转录,其表达可被乳糖诱导。有乳糖时,异乳糖与乳糖阻遏物结合,阻止后者跟乳糖操纵基因的结合,从而使操纵子得以转录。乳糖消耗完后,乳糖阻遏物恢复活性,又结合到乳糖操纵基因上,使转录受到阻遏。	
分解物阻遏	该调控机制使大肠杆菌可以在有葡萄糖存在的情况下抑制乳糖操纵子。分解物激活蛋白(CAP)与 cAMP 结合,并通过结合在乳糖操纵子启动子的上游激活乳糖操纵子的转录。葡萄糖可通过抑制腺苷酸环化酶来调节 cAMP 的含量。当葡萄糖存在时,cAMP 含量降低,CAP 无法与乳糖操纵子的启动子结合,操纵子仅有低水平的转录。葡萄糖含量低时,cAMP 含量上升,CAP 与乳糖操纵子的启动子结合,激活操纵子转录。分解物阻遏能够确保在乳糖和葡萄糖同时存在的情况下,细菌优先利用葡萄糖。	
色氨酸操纵子	该操纵子包括五个基因,编码色氨酸生物合成所需的酶,由同一个启动子转录。在色氨酸存在的情况下,色氨酸阻遏物结合在色氨酸操纵基因上,阻遏操纵子的转录。在没有色氨酸的情况下,色氨酸阻遏物无法与操纵子结合,色氨酸操纵子得以转录。	
弱化作用	该调控机制使色氨酸操纵子及其他操纵子的表达得到精确调节。位于色氨酸启动子和色氨酸操纵子的第一个基因之间的 DNA 序列,可以形成对转录无影响的大的茎环结构,也可以形成较小的终止环。上游一段短的编码区包含色氨酸密码子。当色氨酸足够时,RNA 聚合酶对该区转录之后,核糖体紧随其后,阻止较大的茎环结构形成,而只形成终止环,从而使转录停止。在色氨酸缺乏的情况下,核糖体停止移动,RNA 聚合酶则向前移动,从而形成大的茎环结构;由于终止环的形成受到抑制,操纵子得以继续转录。	
可变 σ 因子的调控	该机制能够使基因表达随环境的改变而发生重大变化。可变 σ 因子改变了细菌 RNA 聚合酶的特异性,使其能够识别不同基因的启动子。可变 σ 因子在大肠杆菌的热休克反应及在枯草芽孢杆菌的芽孢形成过程中激活基因转录。噬菌体也能够合成 σ 因子,以指导其基因的转录。	
相关主题	基因(A2) 信使 RNA(A7)	原核生物基因组(B3) 细菌噬菌体(B8)

A11 真核生物基因表达调控

要 点

基因表达

真核细胞中的基因具有复杂的调控模式。不同类型的细胞表达不同的基因,每个细胞表达的基因大约只占基因总量的 15%。基因表达的模式决定了细胞的特性及其在生物体中的作用。基因表达模式的改变导致细胞分化。基因表达模式的异常与肿瘤的发生有关。

转录调控

真核细胞主要通过改变基因的转录速率来调节基因表达。RNA 聚合酶 II 与基础转录因子之间相互作用,从而在 TATA 框处形成转录起始复合体(TIC)。其他转录因子通过与启动子序列结合和影响 TIC 的稳定性来改变转录起始的速率。被称为增强子和沉默子的远距离调控序列也可影响转录速率。

转录因子

基因的启动子含有多个可影响转录的转录因子结合位点。对转录的最终影响取决于所结合的转录因子的组合。转录因子具有模块化的结构特征,一般包括 DNA 结合域、二聚化结构域和反式激活域。每一个结构域都有各自的特征基序。DNA 结合域包括三种基序:螺旋-转角-螺旋,锌指以及与二聚化结构域组合而成的碱性结构域。二聚化结构域包含两种基序:亮氨酸拉链和螺旋-环-螺旋。二聚化可形成同源二聚体和异源二聚体,从而使转录因子具有多种功能。反式激活域没有可识别的基序,但是经常富含酸性氨基酸、谷氨酰胺或脯氨酸。反式激活域可能与 TIC 中多种不同的蛋白质产生相互作用,或在不同的转录阶段与各种不同的蛋白质产生相互作用。转录因子也可通过直接或间接的机制抑制转录。

激素和细胞因子对基因表达的调控

激素和细胞因子通过激活基因转录来影响靶细胞。类固醇激素进入细胞内部,与类固醇激素受体结合,将其从抑制蛋白中释放出来。受体二聚化并转移到细胞核内,与靶基因的启动子结合,从而激活转录。多肽类激素以及细胞因子与靶细胞表面的受体结合,一系列蛋白质通过磷酸化被依次激活,通过这样的信号转导触发基因的活化。

RNA 干扰(RNAi)介导的转录后调控

在 Drosha 和 Dicer 酶的作用下,双链 RNA 被加工成 22bp 的微 RNA(miRNA)或短干扰 RNA(siRNA),这标志着 RNA 干扰(RNAi)的开始。这些 RNA 片段跟包括 Argonaute 蛋白在内的蛋白质结合,形成 RNA 介导的沉默复合物(RISC)。此后,这些小 RNA 变成单链形式,将 RISC 引导到互补的 mRNA 上,进而结合 mRNA,并催化 mRNA 的切割。若两条 RNA 的碱基不完全匹配,则可能不发生切割,但是会阻止翻译。基因组的序列信息可以用来设计反义 RNA,这已成为一个非常有效的、通过实验手段降低基因表达以观察表型效应的方法。

分化与发育

在多细胞有机体中,细胞和组织的分化与发育是通过长期的表观遗传调控来完成的。

相关主题

基因(A2)

表观遗传学与染色质修饰(A12)

基因转录(A4)

人类基因组(B4)

A12 表观遗传学与染色质修饰

要点

概述

表观遗传效应(epigenetic effects,根据目前的用法)是指由于染色质结构修饰而造成的基因表达的长期变化,并不涉及 DNA 序列的改变。它们有可能在多细胞生物中维持细胞谱系(眼组织或骨细胞)的分化,或者将对基因表达的影响传递到下一代。许多表观遗传效应在配子形成或胚胎发育时被重新设置。

染色质修饰

染色体某个区域的凝缩(紧密化/异染色质化)能够抑制该区域基因的转录,究其原因可能是阻止了转录机器(对该区域)的接近(转录抑制)。这种凝缩伴随着 DNA 和组蛋白的修饰,并进而启动了一系列蛋白质的级联式结合,使染色质得到压缩。DNA 的修饰是通过将胞嘧啶甲基化为 5-甲基胞嘧啶而实现的,其中,在脊椎动物中甲基化主要发生在双核苷酸 5'-CpG-3'上,在植物中甲基化发生在三核苷酸 5'-CNG-3'上(在无脊椎动物中胞嘧啶没有甲基化现象)。DNA 复制之后,维持性的甲基化酶识别半甲基化的甲基-CpG·CpG 的回文序列(CpG 在两条链上的序列相同),将新链中的胞嘧啶甲基化,以与母链配对。从头甲基化酶则负责对未甲基化的 DNA 进行甲基化。组蛋白的修饰主要发生在 N 端;N 端从核小体中伸出,与其他蛋白质相结合。组蛋白的修饰包括赖氨酸的乙酰化或甲基化,以及丝氨酸的磷酸化;似乎存在一种复杂的修饰密码。转录失活(抑制)通过双链 RNA 指向特定的序列,其机制与 RNA 抑制类似(见 A11)。