

doi:10.3969/j.issn.2095-1736.2012.04.092

基因是什么? 分子遗传学教学中的体会和理解

莫日根, 邢万金, 哈斯阿古拉
(内蒙古大学生命科学学院, 呼和浩特 010021)

摘要: 随着分子遗传学的飞速发展, 基因概念也在不断更新。笔者从事分子遗传学、分子生物学、遗传学教学以及基因与基因表达调控方面研究多年, 对基因的本质和概念有较深的理解和认识。回顾了经典基因概念的形成和发展过程, 并讨论了真核生物中的 RNA 遗传和朊病毒复制现象, 提出了新的基因概念。认为基因是携带遗传信息的、可遗传的核酸片段或者多肽分子, 它们可以编码具功能的 RNA 分子或多肽分子。

关键词: 基因; RNA 基因; 教学

中图分类号: Q341; G642

文献标识码: C

文章编号: 2095-1736(2012)04-0092-04

What is a gene——understanding of the gene concept in teaching of molecular Genetics

Morigen, XING Wan-jin, Hasi Agula

(School of Life Sciences, Inner Mongolia University, Hohhot 010021, China)

Abstract: The gene concept has been developed and challenged by new findings. Here the abstract concept of gene from Mendel, nature of gene and challenging discoveries for traditional gene concept and our own understanding of the gene concept was reviewed. Result proposed that a gene is a heritable nucleic acid or protein, which constitutes a functional molecule of RNA or protein.

Keywords: gene; RNA gene; teaching

分子遗传学是以在分子水平上阐明基因的本质、复制、变异、进化、表达调控、重组等生命现象的分子机理为研究目标的新兴学科。因而, 在分子遗传学教学中, 教师对基因本质的认识和基因概念的准确表述非常重要。早期的基因概念简单, 就是能够编码蛋白质的 DNA 片段。然而, 随着分子遗传学的快速发展, 当前用几句话很难把基因概念表述清楚。尤其小分子 RNA 及其功能的发现、某些特殊生命体遗传中 RNA 在遗传信息的传递过程中所起的作用、以及朊病毒的发现, 使基因概念更加模糊, 难以确切地定义。

Pearson^[1]和 Gerstein 等^[2]分别在 2006 和 2008 年回顾和讨论过基因概念, 但未能明确什么是基因。自 20 世纪 80 年代以来, 国内有 20 多篇综述讨论过基因概念的形成、发展和演变, 主要表述孟德尔基因概念、

Morgan 基因理论、断裂基因、跳跃基因、重复基因和假基因等概念^[3-4]。2010 年在《遗传》上发表的一篇综述回顾了基因概念的演绎, 重点讨论了 ENCODE(Encyclopedia of DNA Elements) 计划和 ncRNA(non-coding RNA)^[5]。

在此, 笔者将从孟德尔的抽象基因概念开始, 简单回顾和探讨基因的本质、概念的形成和发展以及挑战传统基因概念的新发现, 提出了新的基因概念。

1 基因概念的提出

经典豌豆杂交实验使孟德尔发现遗传因子以及“自由组合”、“独立分离”和“显性原则”等三大遗传学规律^[6], 奠定了经典遗传学的基础。孟德尔认为“杂合体中决定某一性状的成对遗传因子, 在减数分裂过程中, 彼此分离, 互不干扰, 使配子中只具有成对遗传因

收稿日期: 2011-11-15; 修回日期: 2011-12-05

基金项目: 国家自然科学基金项目(编号: 31060015)

作者简介: 邢万金(1965-) 男, 教授, 博士, 研究方向: 分子遗传学, E-mail: xwanjin@imu.edu.cn;

通讯作者: 莫日根(1961-) 男, 教授, 博士, 研究方向: DNA 复制与细胞周期, E-mail: morigenm@life.imu.edu.cn。

子中的一个,从而产生数目相等的两种类型配子,且独立地遗传给后代;具有两对(或更多对)相对性状的亲本进行杂交,在 F₁ 产生配子时,等位基因分离的同时,非等位基因表现为自由组合,这就是自由组合规律。也就是说,一对等位基因与另一对等位基因互不干扰,各自独立地分配到配子中”。

孟德尔所提出的“遗传因子”是人类对基因的最早认识。孟德尔认为“遗传因子(基因)是一个遗传的单位,可以决定从亲代遗传给子代的性状”。1909年, Johanssen 提出基因“gene”这个名词^[7],并认为孟德尔的“遗传因子”就是基因。然而,当时的基因只是个抽象概念,并不指特定的物质。

2 对基因本质的认识和基因概念的形成

1910年, Morgan 在果蝇杂交实验中发现果蝇眼睛颜色(白眼睛)与性别(雄性)相互连锁,这个结果使他认识到基因在染色体上,并证实基因是位于染色体上的实体,是物质。他认为基因是一个遗传、交换和突变的单位^[8]。1941年, Beadle 和 Tatum 用紫外线照射粗糙脉孢菌(*Neurospora crassa*),诱变获得了氨基酸和维生素的营养缺陷突变体。这些突变体不能产生具有某种功能的酶。例如,有一个突变体不能合成色氨酸是由于它不能产生色氨酸合成酶。而且这些突变性状可以稳定遗传。于是,他们提出了“一个基因一种酶”的假说^[9],认为性状是由基因决定的酶来实现的。这些结果提示基因在染色体上并决定蛋白质,后者控制性状。当时人们并不知道基因在染色体的 DNA 还是蛋白质分子上。当时很多科学家都认为蛋白质是遗传信息的携带者,因为组成蛋白质的 20 种氨基酸能够保证基因的多样性,确保物种多样性。然而, Avery 于 1944 年证明了 DNA 是遗传信息的携带者,是遗传物质。早在 1928 年 Griffith 发现光滑型(S 型)肺炎球菌的侵染会使小鼠死亡,而粗糙型(R 型)菌株的侵染不会。他还发现 R 型菌可以被 S 型菌的细胞提取液所转化而成为 S 型菌株。Avery 利用 Griffith 的这个体系进行进一步转化实验。他分别用蛋白酶(降解蛋白质)和 DNA 酶(降解 DNA)来处理 S 型菌株的细胞提取液,然后分别用来转化 R 型菌株,结果发现用蛋白酶处理的细胞提取液能够把 R 型菌转化为 S 型,而 DNA 酶处理的则不能。结果说明 S 型菌的 DNA 能够改变 R 型菌的遗传特性,而其蛋白质不能改变,由此间接证实了 DNA 是遗传物质,蛋白质则不是^[10]。1952 年 Hershey 和 Chase 分别标记噬菌体 DNA 和蛋白质并分析侵染细菌后的噬菌体特性来进一步证明遗传物质是 DNA 而不是蛋白质^[11]。这些实验不仅证明了 DNA 是遗传物

质,还揭示了遗传物质的化学本质。

1953 年, Watson 和 Crick 根据 Wilkins 和 Franklin 所进行的 X 射线衍射分析,提出了著名的 DNA 双螺旋结构模型,认为 DNA 是由反向平行排列的两条单链 DNA 组成,每条 DNA 链的骨架是由磷酸二脂键相连的戊糖组成;内侧碱基 A 和 T 之间、G 和 C 之间分别形成的 2 个和 3 个氢键保证双螺旋的相对稳定性^[12]。DNA 双螺旋结构全面揭示了 DNA 的化学特性,而且解决了当时困扰生物学家的基因复制和遗传难题。

综上所述,当时的实验证据认为基因就是染色体上的 DNA 分子的一个区段。每个基因由成千上万个脱氧核苷酸组成,一个 DNA 分子可以包含几个乃至几千个基因,基因表达产物是蛋白质。甚至 Benzer 认为基因是一个不能分割的功能单位^[13]。

3 基因理论的发展

在 Morgan 基因理论的影响下,20 世纪 50 年代,生物学家认为基因是在染色体 DNA 上,基因在染色体上线性排列其位置是固定的,而且基因不可分割,也不会随意移动。然而,同时代的 McClintock 发现玉米苗幼叶上有一对来自于姐妹细胞的同源区域的一半表现为色素减少,而另一半则表现为色素增多。从这一对相反关系中, McClintock 意识到,在有丝分裂期间,两个姐妹细胞中的一个得到了另一个细胞所丢失的因子,或者前者控制着后者的活性,结果使同源区域的色素表现出反向关系。这就是由 Ac(Activation)与 Ds 构成的一个控制系统,其中 Ac 的活动是自主的,能够自发转座;Ds 的活动是非自主的,因其中心部分产生缺失而不能自主转移,只有在基因组中有同一族的 Ac 存在时,才能够转座。于是, McClintock 认为基因可以从染色体的一个位点转移到另一个位点,甚至从一条染色体跳到另一条染色体上,即基因可以转座^[14]。当时这个观点并没有引起生物学界应有的反响,甚至有人反对,因为 Morgan 的基因理论深入人心,坚信基因在染色体上的位置是固定的。后来,20 世纪 60 年代在细菌研究中人们证实基因的确可以转座。基因转座现象挑战了 Morgan 的基因理论,同时也丰富了基因理论。

对传统基因理论的另一个挑战是内含子的发现。1977 年 Robert 和 Sharp 试图定位腺病毒(Adenovirus)的基因在其基因组上的位置。他们首先分离腺病毒的 mRNA,然后与其基因组 DNA 杂交,在电子显微镜下他们观察到,杂交后的 mRNA 和其 DNA 形成一些环状结构。于是,他们意识到 mRNA 分子明显比其相应 DNA 分子小,而且,某些 DNA 序列并没有转录成 mRNA。这个结果和其分析使他们得出结论“基因是可以断裂

的”。他们提出一个基因可以被一个或几个非编码的DNA序列分隔,非编码的序列被称为内含子(intron),编码序列为外显子(exon)^[15]。有趣的是,2005年Meng等发现细菌内含子可以表达具功能的蛋白质,这个发现使内含子的概念复杂化。人们还发现基因与基因之间有重叠,比如:大肠杆菌原噬菌体基因 $recE$ 的一部分同时也是相邻基因 $recT$ 的一部分。

由此,可以归纳出基因是由能够编码功能RNA分子的一个或几个DNA片段组成^[16],并且其位置也可以转座。

4 RNA基因

那么, RNA本身是基因吗?答案是肯定的,在RNA病毒中, RNA就是其基因组。近年来的不少实验证据说明,在真核生物中RNA也可以是基因,因为RNA直接携带遗传信息,并以RNA形式传递给下一代。2005年Lolle等发现一种RNA分子可以复制并遗传多代。拟南芥 $hth-8/hth-8$ 突变体的花序异常,当他们把两株 $hth-8/hth-8$ 纯合体进行杂交后发现,其后代中0.19%表型逆转为野生型。对其基因进行分析发现具有野生型表型的0.19%后代仍然是 $hth-8/hth-8$ 。他们认为这是因为野生型RNA分子所致,并提出RNA可以遗传^[17]。在小鼠研究中得到雷同结果。Rassoulzadegan等2006年发现灰色小鼠 $Kit^{tm1Aif/+}$ 杂合体的尾巴尖和脚趾为白色(杂色), $Kit^{+/+}$ 为全身灰色(单色), $Kit^{tm1Aif/tm1Aif}$ 不能存活。当他们把 $Kit^{tm1Aif/+}$ 杂合体进行杂交后,其后代基因型可能有 $Kit^{tm1Aif/+}$, $Kit^{+/+}$ 和 $Kit^{tm1Aif/tm1Aif}$,但后代能存活的小鼠表型应该是杂色和单色。然而,他们意外地发现杂交子一代的小鼠大部分为杂色。他们分析发现一部分杂色小鼠其基因型为 $Kit^{+/+}$,这一结果使他们意识到基因型为 $Kit^{+/+}$ 的小鼠表现为杂色是因为RNA遗传的结果^[18]。这两个实验说明在高等植物和动物中RNA可以遗传而且可以决定表型。

原生动物线虫*Oxytricha trifallax*是单细胞真核生物,有两个细胞核,一个是体积较大的体细胞核(macronucleus, MAC),为所有无性生长所需基因转录提供模板,另一个是较小的配子细胞核(micronucleus, MIC),为有性生殖过程中遗传物质交换所需要。相当于后生动物精子和卵子结合,线虫可以通过细胞质相连来交换两个细胞之间的遗传物质,在此过程中,体细胞核断裂、分解而大部分消失,只有配子细胞核相互融合。2007年Nowacki等发现在有性生殖中体细胞核不会直接传递到子细胞中,而是子细胞在后期发育中以RNA分子为模板重排基因组(DNA)建立起来的^[19]。

他们把改造的*Oxytricha trifallax* RNA分子注入生殖过程的细胞质,然后测定在其子代细胞的体细胞核(MAC)中是否有以他们注入的RNA分子为模板的DNA,结果是肯定的。这个实验很有力地证明携带有遗传信息的RNA分子通过细胞质可以传递给子细胞,使后者拥有建立体细胞核的遗传蓝本。

1998年Fire和Mello发现注射单链RNA分子或双链RNA分子能够沉默线虫*C. elegans*的相应基因表达,而且发现双链RNA分子可以遗传给下一代^[20]。他们这个研究掀起了小分子RNA(miRNA)研究热潮。2005年Lu等发现拟南芥基因组可以转录出1.5百万个miRNA,它们分别来自75000个不同DNA片段^[21];同年,Carninci等发现小鼠基因组转录出102281个RNA分子,包括32129个编码蛋白的和34030个非编码RNA(non-coding RNAs)^[22]。这些结果显示miRNA在生命过程中必定起着重要作用,特别在个体发育、基因表达调控、抑癌和诱癌等方面具重要作用。人类基因组中仅有1.5%的序列编码蛋白质,在果蝇、秀丽线虫及酵母菌中都发现类似的现象。那么,其余的98.5%序列的作用是什么呢?目前没有定论。有研究证明哺乳动物基因组的60%~70%可以转录为RNA分子^[22]。然而,这些RNA分子除了参与完成蛋白质合成(mRNA、rRNA和tRNA)、基因表达调控(miRNA)和RNA酶外,还执行哪些生命功能呢?其中的一部分RNA分子是否携带遗传信息,以非孟德尔方式遗传给下一代呢?

5 蛋白质是遗传物质吗?

朊病毒(Prion)的发现使一些科学家认为蛋白质也能遗传自身^[22]。朊病毒蛋白 PrP^{Sc} 本身是个体基因组的 PrP 基因编码的分子量为33~35 kDa的 PrP^C 蛋白的异构体。 PrP^C 蛋白对蛋白酶敏感,但其异常构型的异构体 PrP^{Sc} 对蛋白酶有抗性,能够引起疯牛病,即朊病毒。 PrP^C 和 PrP^{Sc} 蛋白的一级结构完全相同,但在蛋白质空间构象上不同, PrP^{Sc} 比 PrP^C 具有更多的 β 折叠结构,即 PrP^{Sc} 是 PrP^C 的一种异常的异构体(abnormal isoform)。朊病毒 PrP^{Sc} 侵染后能够把正常的 PrP^C 构型改变为具有侵染能力的 PrP^{Sc} 构型。从正常的 PrP^C 构型改变为 PrP^{Sc} 构型过程可以理解为 PrP^{Sc} 以其自身为模板,利用宿主的 PrP^C 复制出更多的 PrP^{Sc} ,虽然没有传统意义上的遗传信息控制和传递,但确实是在扩增繁殖。像朊蛋白这样的多肽既可以借助编码它的DNA基因遗传,也可以依靠自身直接“复制”扩增,这一现象拓宽了我们对遗传物质和遗传信息的实质及其传递方式的认识,甚至被一些学者^[23]称为“蛋白质基

因”。

6 小结

从孟德尔提出的抽象基因概念开始, Morgan 证实了基因是染色体上的实体, Beadle 和 Tatum 提出“一个基因一种酶”的理论, Avery 证实 DNA 是遗传物质, Sharp 和 Robert 发现基因被内含子分隔成几段以及 McClintock 发现的跳跃基因丰富和完善了基因理论。RNA 遗传现象和朊病毒增加了我们对基因的认识。也许我们可以说, 基因是携带有遗传信息的、可遗传的核酸片段或者多肽分子, 它们可以编码具有功能的 RNA 分子或多肽分子。

参考文献:

- [1] Pearson H. What is a gene? [J]. *Nature*, 2006, 441: 398–401.
- [2] Gerstein M B, Bruce C, Rozowsky J S, et al. What is a gene, post-ENCODE? history and updated definition [J]. *Genome Res*, 2007, 17(6): 669–681.
- [3] 杨洪日. 基因概念在现代遗传学发展中的地位及展望 [J]. *沈阳农业大学学报*, 1989, 20(3): 355–358.
- [4] 胡 楷. 基因概念的发展 [J]. *海南大学学报(自然科学版)*, 1985, 3(2): 54–67.
- [5] 谢兆辉. 基因概念的演绎 [J]. *遗传*, 2010, 32(5): 448–454.
- [6] Mendel J G. Versuche über Pflanzenhybriden [J]. *Verhandlungen des naturforschenden Vereines in Br nn*, 1865, *Abhandlungen*: 3–47.
- [7] Johannsen W. The genotype conception of heredity [J]. *The American Naturalist*, 1911, 45: 129.
- [8] Morgan T H. Sex-limited inheritance in *Drosophila* [J]. *Science*, 1910, 32: 120–122.
- [9] Bedle G W, Tatum E L. Genetic control of biochemical reactions in *Neurospora* [J]. *Proc Natl Acad Sci USA*, 1941, 27(11): 499–506.
- [10] Avery O T, MacLeod C M, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types [J]. *J*

Exp Med, 1944, 79: 137–158.

- [11] Hershey A D, Chase M. An upper limit to the protein content of the germinal substance of bacteriophage T2 [J]. *Virology*, 1955, 1: 108–127.
- [12] Watson J D, Crick F H C. A structure of deoxyribonucleic acid [J]. *Nature*, 1953, 171: 964–967.
- [13] Benzer S. Fine structure of a genetic region in bacteriophage [J]. *Proc Natl Acad Sci USA*, 1955, 41: 344–354.
- [14] McClintock B. Mutable loci in maize [M]. *Carnegie Inst Wash Year Book*, 1948, 47: 155–169.
- [15] Chow L T, Roberts J M, Lewis J B, et al. A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA: DNA hybrids [J]. *Cell*, 1977, 11(4): 819–836.
- [16] Alberts B, Johnson A, Lewis J, et al. *Molecular biology of the cell* [M]. 5th edition, GS Garland Science Taylor & Francis Group, 2008, 200.
- [17] Lolle S J, Victor J L, Young J M, et al. Genome-wide non-mendelian inheritance of extra-genomic information in *Arabidopsis* [J]. *Nature*, 2005, 434(7032): 505–509.
- [18] Rassoulzadegan M, Grandjean V, Gounon P, et al. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse [J]. *Nature*, 2006, 441: 469–474.
- [19] Nowacki M, Vijayan V, Zhou Y, et al. RNA-mediated epigenetic programming of a genome-rearrangement pathway [J]. *Nature*, 2008, 451(7175): 153–158.
- [20] Fire A, Xu S, Montgomery M K, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans* [J]. *Nature*, 1998, 391(6669): 806–811.
- [21] Lu C, Tej S S, Luo S, et al. Elucidation of the small RNA component of the transcriptome [J]. *Science*, 2005, 309(5740): 1567–1569.
- [22] Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome [J]. *Science*, 2005, 309(5740): 1559–1563.
- [23] 李振刚. 分子遗传学(第三版). 北京: 科学出版社, 2008, 203–208.

(上接 88 页)

参考文献:

- [1] 刘万勃, 宋 明, 刘富中, 等. RAPD 和 ISSR 标记对甜瓜种遗传多样性的研究 [J]. *农业生物技术学报*, 2002, 10(3): 231–236.
- [2] Bandraeco-Arnas S, Pitrat M. A genetic map of melon (*Cucumis melo* L.) with RFLP, RAPD, isozyme, disease resistance and morphological markers [J]. *Theor Appl Genet*, 1996, 93(1–2): 57–64.
- [3] Katzir N, Danin-Pole Y, Tzuri G, et al. Length polymorphism homologies of microsatellites in several *Cucurbitaceae* species [J]. *Theor Appl Genet*, 1996, 93(8): 1282–1290.
- [4] 杨柳花, 李先信, 曾柏全, 等. 新型分子标记 SRAP 的原理及其研究进展 [J]. *湖南农业科学*, 2009, 5: 15–17, 20.
- [5] 张永平, 朱为民, 崔辉梅. 上海地区番茄烟粉虱传双生病毒 PCR 检测 [J]. *上海农业学报*, 2007, 23(4): 100–101.
- [6] 马金骏. 不结球白菜种质资源遗传多样性的初步研究 [D].

2008, 27.

- [7] Ferriol M, Pico B, Pascual Fernandez C, et al. Molecular diversity of a germplasm collection of squash (*Cucurbita moschata*) determined by SRAP and AFLP marker [J]. *Crop Science*, 2004, 44: 653–664.
- [8] 王 刚, 潘俊松, 李效尊, 等. 黄瓜 SRAP 遗传连锁图的构建及侧枝基因定位 [J]. *生命科学(中国科学 C 辑)*, 2004, 34(6): 510–516.
- [9] 于拴仓, 柴 敏, 姜立纲. 主要番茄品种的分子鉴别研究 [J]. *中国农学通报*, 2005, 21(5): 84–89.
- [10] 艾呈祥, 陆 路, 马国斌, 等. SSR 标记在甜瓜杂交种纯度检测中的应用 [J]. *园艺学报*, 2005, 32(5): 902–904.
- [11] 柳剑丽, 周寿春, 宋晓华, 等. 利用 SSR 标记鉴定甜瓜杂交种的种子纯度 [J]. *分子植物育种*, 2006, 6(S): 23–26.
- [12] 李菊芬, 许 玲, 马国斌. 应用 SSR 分子标记鉴定甜瓜杂交种纯度 [J]. *农业生物技术学报*, 2008, 16(3): 494–500.