

表观遗传学研究进展

康静婷^{1,2}, 梁前进^{1,3}, 梁辰¹, 王鹏程¹

1. 北京师范大学生命科学学院; 抗性基因资源与分子发育北京市重点实验室, 北京 100875
2. 北京师范大学生命科学学院; 基因工程药物及生物技术北京市重点实验室, 北京 100875
3. 北京师范大学生命科学学院; 细胞增殖及调控生物学教育部重点实验室, 北京 100875

摘要 概述了表观遗传调节模式、表观遗传调节的效应、植物表观遗传学的研究进展等。在每种细胞中, 都会发生一部分特异基因激活、另一部分基因抑制的现象, 形成多种基因表达模式。表观遗传指 DNA 序列不发生变化, 而基因表达发生可遗传改变的现象。表观遗传学改变包括 DNA 甲基化、组蛋白修饰、非编码 RNA 作用等, 产生基因组印记、母性影响、基因沉默、核仁显性、休眠转座子激活等效应。表观遗传变异是环境因素和细胞内遗传物质间交互作用的结果, 其效应通过调节基因表达, 控制生物学表型来实现。正是因为表观修饰对于维持生物体内环境和各器官系统功能的重要性, 表观遗传的异常会引发疾病, 这也成为药物和治疗方案设计的着眼点。

关键词 表观遗传; DNA 甲基化; 组蛋白修饰; 非编码 RNA; 基因表达调控

中国分类号 Q341

文献标志码 A

doi 10.3981/j.issn.1000-7857.2013.19.011

Overview on Epigenetics and Its Progress

KANG Jingting^{1,2}, LIANG Qianjin^{1,3}, LIANG Chen¹, WANG Pengcheng¹

1. Beijing Key Laboratory of Gene Resource and Molecular Development; College of Life Sciences, Beijing Normal University, Beijing 100875, China
2. Beijing Key Laboratory of Gene Engineering Drugs & Biology Technology; College of Life Sciences, Beijing Normal University, Beijing 100875, China
3. Key Laboratory of Cell Proliferation and Regulation Biology of Ministry of Education; College of Life Sciences, Beijing Normal University, Beijing 100875, China

Abstract Epigenetic regulation modes, epigenetic regulatory effects, and the research progresses in plant epigenetic are summarized. According to classical genetics, nucleic acids are the foundation of heredity, and the hereditary information of life is stored in the base sequence of nucleic acids. The change of base sequence might lead to phenotypic variations of organisms, and these variations might pass on to the next generation. Within one organism, the different types of cells have their own unique gene expression patterns, although they share the same genetic background. It is caused by activating some special sets of genes, meanwhile suppressing the other. This phenomenon is known as epigenetic modification. Epigenetics is a phenomenon for altering gene expression pattern; however changes in the underlying DNA sequence of the gene are not involved. Its process includes DNA methylation, histone modification, and non-coding RNA regulation, etc. Its manifestation includes genomic imprinting, maternal effect, gene silencing, nucleoli dominance, and dormant transposon activation, etc. Epigenetic modification is the result of interactions between environmental factors and intracellular genetic materials. Through gene expression regulation, epigenetic modification is able to control the phenotype of an organism. Due to its importance for maintaining stable environment inside an organism and normal functions of organs, abnormal epigenetic modification could lead to disease. It is why today epigenetics becomes a hot research spot on the drug design and therapy development.

Keywords epigenetics; DNA methylation; histone modification; non-coding RNA; gene expression regulation/control

收稿日期: 2013-03-15; 修回日期: 2013-05-11

基金项目: 北京市自然科学基金项目(5122017); 细胞增殖及调控生物学教育部重点实验室开放基金项目(201101); 教育部2011年度“大学生创新创业训练计划”项目(北京师大·2011-60); 抗性基因与分子发育北京市重点实验室开放基金项目(201204)

作者简介: 康静婷, 博士研究生, 研究方向为肿瘤细胞生物学, 电子信箱:kjt0225@163.com; 梁前进(通信作者), 副教授, 研究方向为分子细胞遗传及功能基因, 电子信箱:lqj@bnu.edu.cn

0 引言

经典遗传学认为,核酸是遗传的分子基础,生命的遗传信息储存在核酸的碱基序列。人类基因组(genome)有20000多个基因,但在成人体内的200种左右细胞中每种细胞内都只有一部分特定基因会表达^[1]。通俗地说,每个个体内虽然所有细胞都含相同的遗传信息,但由于基因表达模式不同,这些本来由同一个受精卵分裂而成的细胞经过分化后成了具有不同功能和形态的细胞(如肝细胞、上皮细胞和血细胞等),从而组成了不同的组织和器官。DNA序列不发生改变的情况下,基因表达发生可遗传改变的现象,就被定义为表观遗传(epigenetic)现象。

在经典遗传学创立数十年后,Waddington于1942年提出表观遗传学的概念,并将其定义为研究生物发育机制的学科^[2]。最初,人们认为表观遗传只是一种表型,但随着生命科学的研究发展,基因组学不能解释的问题越来越多,表观遗传学在这样的情况下不断发展。20世纪70年代中期,人们对表观遗传的理解开始变化 Holliday对其进行了系统表述,即现在广为接受的表观遗传学概念——研究非DNA序列变化所致的可遗传的基因表达变化^[3]。在分子生物学空前发展的形势下,表观遗传学也在分子水平上得到了更为系统的研究,人们不仅发现了多种表观遗传修饰方式,而且探究了其错综复杂的生物学作用。表观遗传学现已成为生命科学领域的研究热点之一,形成了独立的分支学科。

近年研究发现,表观遗传与人类的很多疾病相关,环境压力诱导的一些分子也会引起表观遗传变异(epigenetic variation)^[4]。从根本上看,表观遗传现象就是由环境因素引起的生物细胞内遗传物质变化的结果。在相当长一段时间内,表观遗传学的研究集中在甲基化(methylation)、小RNA(small RNA)和染色质重塑(chromatin remodeling)等方面。前期发表的论文中曾概述了副突变(paramutation)、亲代印记(parental imprinting)、性别相关的基因剂量补偿效应(gene dosage compensation effect)和转基因沉默(transgene silencing)等典型的表观遗传现象及其研究进展^[5,6],本文论述主要的调节机制。

表观遗传学的调节机制主要包括DNA甲基化(DNA methylation)、组蛋白修饰(histone modification)、非编码RNA(noncoding RNA,ncRNA)作用等。这些调节模式易受环境影响,因此表观遗传学更加关注环境诱导的表观遗传变异。最为直观的例子是环境诱导的动物免疫调节反应和胚胎发育后期母体和胚胎的改变。受环境影响,任何一种调节机制发生异常都可能导致细胞状态或细胞增殖发生紊乱,进而引起各种疾病。与DNA序列改变引起的疾病不同,由于许多表观遗传变异是可逆的,表观遗传异常引发的疾病相对容易治疗,这是表观遗传学成为生物医学研究领域新热点的原因之一。

1 表观遗传调节模式

1.1 DNA甲基化及其生理生化效应

DNA甲基化(DNA methylation)是目前研究得最多、最清

楚的表观遗传修饰方式。DNA高度甲基化首先会影响DNA结构,进而阻遏基因转录,引起基因沉默。真核细胞内甲基化状态有3种:持续的低甲基化状态(如管家基因的甲基化)、诱导的去甲基化状态(如一些发育阶段特异性基因的修饰)和高度甲基化状态(如人类女性细胞内缢缩-失活的X染色体的甲基化修饰)。DNA甲基化的具体反应过程是DNA甲基转移酶(DNA methyltransferase, DNMT)将S-腺苷甲硫氨酸上的甲基转移到DNA双链中胞嘧啶的第5位碳原子上,形成5-甲基胞嘧啶(5-mC)。催化该反应的DNA甲基转移酶主要有4种:DNMT1、DNMT3A、DNMT3B和DNMT3L。在DNA复制完成后,DNMT1是催化甲基转移至新合成的DNA链上的甲基化位点反应中最主要的酶,这一现象称为维持甲基化(maintenance methylation);而DNMT3A和DNMT3B则负责催化核酸链上新的甲基化位点发生反应,称为形成甲基化(de novo methylation)。DNMT3L在DNA甲基转移酶家族中属于不具有甲基转移酶活性的调节酶,其主要作用是调节其他甲基转移酶的活性^[7]。

哺乳动物细胞内的DNA甲基化主要发生在胞嘧啶和鸟苷酸(CpG)二核酸中的胞嘧啶(C)上基因组中的CpG约有60%~90%会发生甲基化^[8]。在DNA双链中5'-CpG-3'及其互补链中的3'-GpC-5'中的C都会被甲基化,这些CpG是基因组中的维持甲基化位点。在结构基因的启动子或转录起始位点有大量未甲基化的CpG(在大约200bp碱基对中CpG含量超过60%^[9]),这些CpG簇被称为CpG岛(CpG island)。如果CpG岛发生高甲基化,基因表达就会被完全抑制。DNA甲基化对基因表达影响的机制已经研究得较为透彻,可简单地阐述为:甲基化会使DNA双链的大沟在三维结构上发生变化,阻滞甲基化敏感的转录因子(TFs,包括E2F、CREB、AP2、cMyc/Myn、NF-κB、cMyb和ETS等)的DNA结合活性。与此同时,甲基化不敏感的methyl-CpG结合蛋白(如Sp1、CTF和YY1等)会结合在DNA上,这些蛋白是转录抑制因子,它们都含有保守的甲基化DNA结合结构域(methylated DNA-binding domain, MBD)。DNA甲基化影响基因表达的方式如图1^[9]所示,选择性地结合于甲基化DNA的特异转录抑制子MeCP2(methyl-CpG binding protein 2),即甲基化CpG结合蛋白,与组蛋白去乙酰化酶(histone deacetylase, HDAC)共存于一个复合物中。DNA甲基化调节基因表达与组蛋白去乙酰化之间有密切的关系,MeCP2可抑制依赖或不依赖HDAC的基因转录^[10]。在图1中,上面的基因[含不同的外显子(exons)]的启动子(promoter, Pro)和转录起始点未被甲基化,尽管基因的下游有碱基被甲基化(有MBD结合),但基因依旧能被转录,而增强子(enancer)能发挥作用是因为沉默子(silencer)和绝缘子(insulator)被甲基化。相应地,下面基因的启动子和转录起始位点被甲基化,因此不能在起始位点正常转录同时沉默子和绝缘子与其结合蛋白(图1中GCF2是肝癌相关抗原,CTCF是CCCTC结合因子——一种多功能转录因子)也发挥作用阻止增强子的活性,而LINE(长散在重复序列)有可能在

未被甲基化的情况下进行转录。

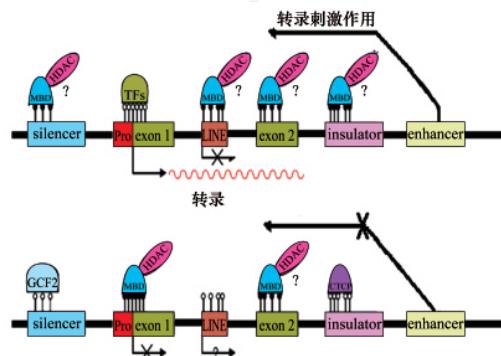


图 1 哺乳动物细胞内 DNA 甲基化对基因表达的影响方式
Fig. 1 Effect manner of intracellular DNA methylation on gene expression in mammals

DNA 甲基化在生物体内有多方面的重要生理意义。正常的甲基化对于维持细胞的生长及代谢等是必需的,具体的体现如维持染色质结构、基因印记、X 染色体失活、细胞分化和胚胎发育等。而异常的 DNA 甲基化则会引发疾病(如肿瘤),因为异常的甲基化一方面可能使抑癌基因无法转录(如在很多肿瘤组织中 *p53* 基因发生了高度甲基化),另一方面也会导致基因组不稳定(如 5-mC 在脱胺反应中转化为胸腺嘧啶,如没有及时修补,就会发生基因变异,引发可遗传疾病)。显然,研究 DNA 甲基化对于了解生物生长发育及疾病治疗是非常有帮助的。在将来的研究中,如果可以去除基因治疗过程中细胞对外来基因的甲基化,将会增加提高疾病治愈的机会。

1.2 组蛋白修饰及其生理生化效应

由于被修饰,一些蛋白失去活性,一些蛋白获得活性,一些蛋白改变功能,因此蛋白修饰是功能蛋白质库的“扩增”因素。组蛋白修饰(histone modifications)是表观遗传修饰的一种重要方式,具有特殊的生理生化功能。在细胞的生长状态下,DNA 以染色质形式存在于细胞核当中。染色质的基本单位是核小体(nucleosome),核小体由 145~147 对 DNA 碱基缠绕在组蛋白 H2A、H2B、H3 和 H4 各 2 个单位组成的八聚体核心周围而形成,每个核小体间由长度约为 60bp 的 DNA 连接,组蛋白 H1 就结合在这些接头 DNA(linker DNA)上。也就是说,染色质由 DNA 结合组蛋白形成的核小体串组成,组蛋白是染色质的基本结构蛋白。组蛋白折叠基序(folding domain)常位于 C-端,参与组蛋白分子间互作并与 DNA 缠绕有关。组蛋白的另一个重要结构域称为组蛋白尾(histone tail),约占组全长的 25%,常位于 N-端(但在组蛋白 H2A 则处于 C-端)^[10],可与 DNA、调节蛋白、酶和其他染色质蛋白相互作用,大部分的组蛋白翻译后修饰都发生在这个结构域的第 15~38 个氨基酸残基上^[11,12]。另外,组蛋白尾在染色质组装和凝聚成高度有序结构的过程中发挥重要作用^[13],而染色质的凝聚程度会直接影响 DNA 复制、重组和转录^[13]。

组蛋白表观遗传修饰方式有甲基化(methylation)、乙酰化(acetylation)、磷酸化(phosphorylation)、泛素化(ubiquitination)、SUMO 化(sumoylation)、腺苷酸化(adenylation)、ADP-核糖基化(ADP-ribosylation)、生物素化(biotinylation)和脯氨酸异构化(proline isomerization)等,这些修饰方式灵活地影响着染色质的结构与功能,既可以阻遏也可以促进基因的转录。参与组蛋白修饰的酶主要有组蛋白甲基转移酶(histone methyltransferase,HMT)、组蛋白乙酰转移酶(histone acetyltransferase,HAT)、组蛋白激酶(histone kinase)和组蛋白泛素化酶(histone ubiquitylase)等,这些酶是催化相应的基团结合到组蛋白氨基残基上所必需的酶。例如 HAT 的底物特异性、催化机制、调节途径和转录调控或染色质修饰功能的研究,证明组蛋白乙酰化广泛影响着生物的正常发育和病理发生过程^[14]。同时相应地,也有组蛋白去甲基化酶(histone demethylase,HDM)、组蛋白脱乙酰基酶(histone deacetylase,HDAC)、组蛋白磷酸酶(histone phosphatase)和组蛋白去泛素化酶(histone deubiquitylase),这些酶可去除结合在组蛋白端氨基残基上的分子基团。大量的研究证明组蛋白修饰是一个动态可逆的过程,基团的添加和去除就是由以上一系列的酶催化反应形成的。一般的组蛋白修饰需要一个或多个不同的共价修饰发生协同或拮抗作用,这些多样性修饰及它们时间和空间上的组合形成大量的特异信号,这些信号类似于密码并可被相应的调节蛋白识别,影响一系列蛋白质的活动,从而调控真核生物的基因表达,这就是“组蛋白密码假说”(histone code hypothesis)^[12,15]。组蛋白密码的组合变化繁多,因此组蛋白共价修饰是精细、有序的基因表达和生理调控方式,研究清楚这样的过程不论在理论上还是在医疗等实践中都可能取得巨大的成果。

组蛋白翻译后修饰类型多,且相互之间息息相关,彼此影响,形成一个错综复杂但井然有序的网络来影响基因的表达。一般认为,组蛋白翻译后修饰影响基因表达的途径有 3 种^[15]:① 改变其周围的环境(如电荷量和 pH 值等),增强或减弱转录因子或转录辅因子与 DNA 间的作用;② 直接改变染色质结构和凝集状态,进而影响蛋白间和蛋白与 DNA 间的相互作用;③ 作为信号影响下游蛋白,进而调控基因表达。例如组蛋白乙酰化(图 2^[16]),当乙酰基结合在组蛋白上时,就会中和后者的正电荷,这样组蛋白末端就能以比较弱的作用力结合在带负电荷的 DNA 链上,这种宽松的染色质结构便可以使特异转录因子等相关蛋白与 DNA 结合。相同原理,当 ADP-核糖基和泛素基团结合在组蛋白 N-末端时,也会使核小体结构变得较为松散。有研究认为,组蛋白 H4 的 SUMO(一种结构上与泛素相似的分子)化可能通过引起 HDAC 和异染色质相关蛋白 HP1 的异常招募来发挥转录抑制作用^[17]。除了组蛋白翻译后修饰的不同方式会在相互影响中发挥作用外,外源物质也会通过影响组蛋白修饰来改变基因的表达。例如,丙戊酸作为一种抗癫痫药物和神经递质 γ -氨基丁酸受体激活剂,也是 HDAC 的抑制剂,可重复激活肿瘤抑制

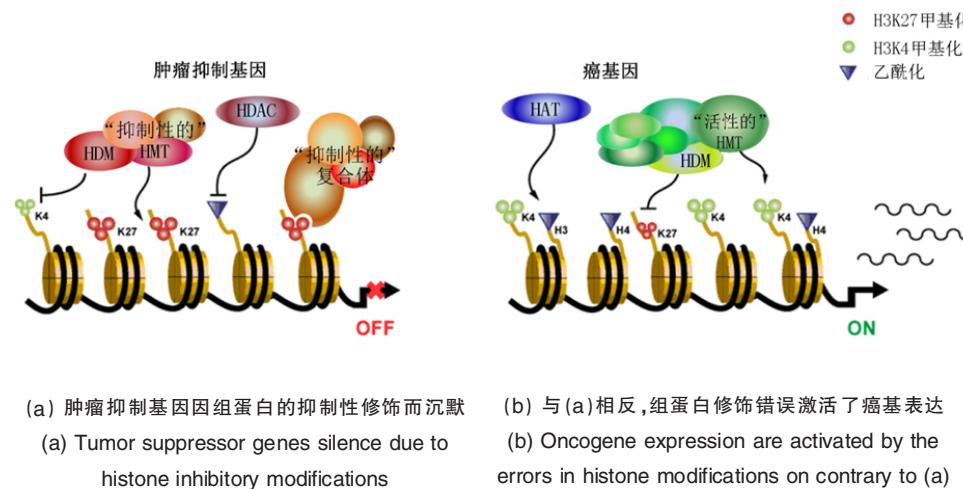


图 2 肿瘤抑制基因和癌基因表达调控错误

Fig. 2 Expression control errors of tumor suppressing genes and oncogene

基因,现已作为抗癌药物投入临床试验^[18]。

目前组蛋白翻译后修饰已成为生物学及基础医学的研究热点,也取得了很多成果,下面重点介绍组蛋白甲基化方面的研究进展。

组蛋白甲基化通常发生在组蛋白 H3 和 H4 N-末端的精氨酸(Arg,R)和赖氨酸(Lys,K)残基上,另外组蛋白的球状结构域有时也会被甲基化,如组蛋白 H3 的 79 位赖氨酸(表示为 H3K79me)会被 DOT1L 催化发生反应^[1]。根据每一位点甲基化的程度不同,组蛋白甲基化形式可分为单甲基化、二甲基化和三甲基化。基于甲基化位点和受作用基因的不同,组蛋白甲基化对基因表达的影响也不同,如图 2 中 H3K27 甲基化对特异基因表达的抑制和 H3K4 甲基化对特异基因表达的激活。催化组蛋白甲基化的 HMT 主要分为两大类:一类是含有保守的 SET (suppressor of variegation,enhancer of zeste, trithorax)催化结构域的蛋白家族(如 SUV39 蛋白家族),广泛存在于真核生物细胞内;另一类不含该结构域(如催化赖氨酸甲基化的甲基转移酶 DOT1L)。与 HMT 作用相反的 HDM 最初由 Shi 等发现^[19]。研究发现 HMT 和 HDM 可与同一个蛋白复合体结合,组蛋白是否被甲基化很大程度上直接取决于这个蛋白复合即时结合的对象。因此,HMT 和 HDM 的表达、活性和招募都会改变组蛋白的状态,从而改变转录平衡。在很多基因敲除小鼠实验中已经发现 HMT 基因的缺失会严重影响小鼠的发育,HMT 的错误调节会导致疾病的产生。目前已经发现了人类基因组中编码了大约 50 种精氨酸和赖氨酸 HMT,而其中至少有 22 种与癌症或其他疾病相关^[1]。

1.3 非编码 RNA 及其生理生化作用

在复杂的生物体内,基因表达受到很多因素影响。在近期高通量基因组水平的分析中,越来越多的证据表明非编码 RNA 在调控基因表达过程中发挥了很大作用。事实上,在所

有输出的转录本中,编码蛋白的 RNA 数量不足 1.5%,剩下的是非编码 RNA(noncoding RNA,ncRNA)^[20]。根据长度分类,介导表观遗传修饰的 ncRNA 可分为 long ncRNA (lncRNA) 和 small ncRNA(sncRNA)。

lncRNA 指长度超过 200nt 的非编码 RNA,其长度范围大概从 50kb 到几百 kb,其序列不具保守性,且不与任何目的基因同源,通过顺式作用调节基因表达,使之沉默。目前认为 lncRNA 的来源途径主要有:蛋白编码基因受多种因素作用而断裂,形成 lncRNA;染色质重排中两分开区域紧密靠拢,形成 lncRNA;非编码基因转录形成 lncRNA;小非编码 RNA 中某段序列多次复制形成 lncRNA;转录因子中插入一段序列形成 lncRNA 等^[21,22]。这些 lncRNAs 虽不编码蛋白,但可调节表观遗传过程,多种肿瘤的发生与其表达异常有关。lncRNA 在基因表达的表观遗传调节中发挥重要作用的实例已屡有报道。如 Xist(17kb)RNA 可引起 X 染色体失活,而起始于 Xist 基因下游的 Tsix(40kb)RNA 产物是其调节子^[23];Air(108kb)Kcnq1(>60kb)和 H19(2.3kb)分别作用于 Igf2r 和 Kcnq1 基因簇,通过与染色质相互作用使基因沉默^[19]。

sncRNA 长度通常小于 30nt,包括 micro-RNA(miRNA)、small interfering RNA (siRNA) 和 piwi -interacting RNA (piRNA)。sncRNA 一般是在 2 个水平上对基因表达进行调控:① 转录水平,被称为转录基因沉默(transcriptional gene silencing,TGS);② 转录后水平,即转录后基因沉默(post-TGS,PTGS)。TGS 抑制转录的发生是通过染色质修饰和异染色质化(heterochromatinization),而 PTGS 则通过降解 mRNA 或阻止 mRNA 翻译来影响 RNA 的翻译。总之,TGS 和 PTGS 最终都是使基因沉默。sncRNA 调控基因表达的机理相对于 lncRNA 来说简单而且单一,因为不论是 TGS 和 PTGS,其作用机制都是 sncRNA 的序列与目的基因相匹配,两者配对、结

合使基因不能发生转录,或 mRNA 不能发生翻译。

2 表观遗传调节的效应

2.1 基因组印记

在一个基因或基因组域上发生双亲来源信息的生化标记的生物学现象,叫做基因组印记(genomic imprinting),或遗传印记、基因印记。所发生的“印记”可以是共价标记(如 DNA 甲基化),也可以是非共价标记(如 DNA-蛋白质互作、DNA-RNA 互作和核基因组定位等),印记方式与整个细胞周期中维持双亲表观记号的特定核内酶的作用有关。基因组印记使得基因依亲代的不同而有不同的表达,还可能导致细胞中两

个等位基因的一个表达而(不同亲源的)另一个不表达。基因组印记病主要体现为生长过度或迟缓、智力障碍和行为异常等。在肿瘤研究中发现,印记缺失是引起肿瘤最常见因素之一。对鸡的 DNA 甲基化所致基因组印记效应的研究发现,在 L-精氨酸:甘氨酸胺基转移酶(GATM)基因的第 8 内含子内有差异甲基化区域(DMR),经亚硫酸氢钠测序法(bisulfite sequencing)分析表明,6d 胚胎的雄性原始生殖细胞(PGC)比雌性的甲基化水平高出 1.6 倍。由于肌酸生物合成(图 3)中的问题会导致严重的神经学缺陷等病症,这样的表观遗传修饰必定需要精细调控。这种 GATM 雌雄差异甲基化状态下的双等位基因表达对揭示基因组印记的机理很有意义^[23]。

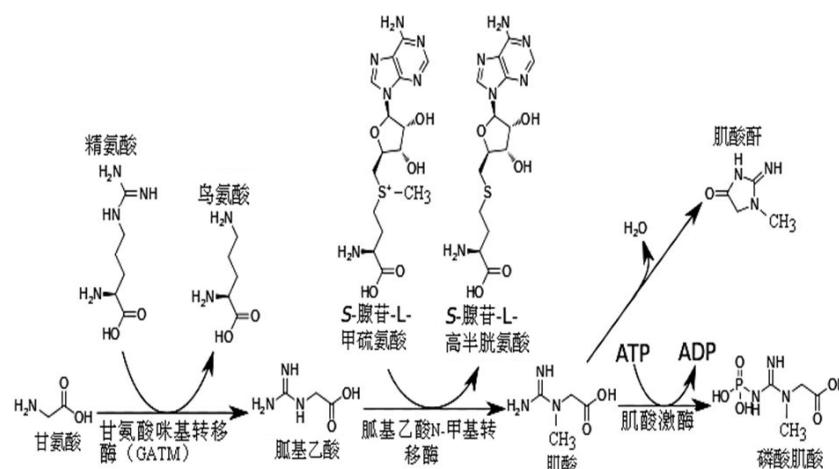


图 3 肌酸生物合成途径

Fig. 3 Biosynthesis path of creatine

2.2 母性效应

母性效应(maternal effect),或称母性影响,是指子代某一性状的表型由母体核基因型决定,而不受本身基因型支配。有名的例子是椎实螺(pond snail)的螺壳,有左旋和右旋之分,旋转方向的遗传符合母性效应(图 4)。最新研究表明,椎实螺交配前行为的偏好性与其螺旋方向(手征性,chirality)相关;不仅如此,脑结构的镜像性分化和整个身体结构的不对称布局也是发育早期形成的,而且都与螺壳手征性的发生有关,很明显是由母性效应基因座位决定的^[24]。母性效应常与印记效应相关,研究表明相关基因的差异性甲基化、磷酸化以及选择性的蛋白互作与母性效应的形成和维持有很大关系。



图 4 椎实螺螺旋方向的母性效应

Fig. 4 Maternal effect of the shell coil chirality of pond snails

2.3 基因沉默

基因沉默(gene silencing)也称基因沉寂,是真核生物细胞基因表达调节的重要手段之一。基因沉默常缘于异染色质(heterochromatin)形成,被沉默的基因区段高度浓缩。受控于 ncRNA 的 RNAi(RNA 干扰)与转录后基因沉默在分子层次上实际是同一现象。基因沉默的反应过程包括组蛋白 N-端赖氨酸残基的去乙酰基化、甲基化修饰、以及甲基化组蛋白与其结合蛋白(MBP)诱发异染色质形成等。目前尚未真正完全弄清楚组蛋白密码与基因沉默的关系。研究者将能够与甲基化组蛋白结合的一组蛋白质 Sir1/2/3/4 称为沉默信息抑制因子,其中 Sir2 就是去乙酰化酶,Sir1/3/4 则负责与甲基化组蛋白结合,发挥染色质“沉默”作用(使之形成异染色质)。相关研究表明,端粒重复 RNA 区域(TERRA)是控制异染色质和端粒酶的因素,在端粒中,DNA 结合蛋白 Rap1 的 C 端结构域可招募 Sir2/3/4 和 Rif1/2(Rap1-interacting factor 1/2)复合物,进而造成基因沉默,促进 Rat1-核酸酶依赖性 TERRA 降解^[25]。基因沉默一方面是遗传修饰生物实用化、商品化的障碍,另一方面也是生物抗逆性(如植物抗病毒)的重要反应,为植物工程育种等提供了策略(例如 RNA 介导病毒抗性技

术的发展)。

2.4 核仁显性

核仁显性(nucleolar dominance)指在动植物杂合体中,核糖体位点基因受到抑制,使染色体遗传自父母中的一方,而表现出的显性效应。其机理是,来自父或母方的 RNA 聚合酶 I 在核糖体 RNA(rRNA)基因转录过程中呈现出可逆的沉默。由于 rRNA 基因跨越数百万 bp, 成簇存在于核仁组织区, 不难理解核仁显性是染色体沉默中的一种主要机制, 本质上是 rRNA 染色质因化学修饰而沉默的结果, 但具体机制并没有定论。近期有人分析了黑腹果蝇(*Drosophila melanogaster*)性染色体中的核仁显型现象, 证明果蝇中的核仁显型只部分地依赖于已知的异染色质基因沉默, 而另外还有作用于 rRNA 基因的明显不同的异染色质调节方式,rRNA 基因阵列的大规模组织形式在其中发挥了重要作用, 这也是有关果蝇等位基因失活研究揭示的首例成果^[24]。由于核仁显型, 杂交动植物中整组亲代 rRNA 基因可能被关闭; 核仁显性机制与癌症等疾病的某些失控机制有相似性, 因而在医学方面也有重要的研究意义。

2.5 表观遗传修饰的其他效应

文献[5]、[6]阐述了染色质重塑(chromosome remodeling), 核小体在真核细胞 DNA 上重新定位的过程, 引起染色质变化, 与组蛋白修饰和核小体结构改变有关)、副突变(paramutation, 一个等位基因可以使其同源基因的转录产生稳定可遗传变化的途径, 涉及配子与合子之间的 RNA 转移)和 RNA 编辑(RNA editing, 基因转录产生的 mRNA 分子中, 因核苷酸缺失、插入或置换, 而造成转录物序列不与基因编码序列互补, 产生不同于基因编码信息的蛋白质氨基酸组成的现象), 这里不再赘叙。除此之外, 休眠转座子激活(dormant transposon activation) 也是逐渐展开研究的表观遗传修饰效应。研究表明, 原癌基因 *CSF1R* 的转录可激活 *MaLR* 基因家族的内源长末端重复(LTR)信息, 即“垃圾 DNA”产生了特殊活性。这对于霍金奇淋巴瘤之类肿瘤细胞的生长起到了推波助澜的作用^[27]。

生物的生理和发育过程均具有明显的全息(holographic)特性, 一种过程往往联系着另一种过程。李艳等^[28]研究了比 H3K9me2 更特异的 X 染色体失活修饰 H3K9me3 的细胞学现象后, 近期又证明这样的表观遗传标记在肿瘤发生中失去特异性^[29]。

3 植物表观遗传学研究

在一定范围内, 植物表观遗传学的相关进展相当突出。表观遗传学的一些重要现象, 如转基因沉默、副突变等开始就是在植物体系中发现的。植物表观遗传学也有一些独特之处, 如 DNA 甲基化的种类和调控 DNA 甲基化的酶类等。植物表观遗传学特异性现象和机制的揭示是理解表观遗传学全貌的重要内容。

3.1 植物转基因沉默

经过努力, 数 10 种转基因植物已经问世, 有的已经用于

农田生产, 但所转入的目的基因沉默而不表达, 或不稳定、不完全表达却是令人尴尬的局面, 这已经成为植物遗传改良的突出障碍; 沉默基因的消除是一个重要的课题^[30]。转基因沉默就是转基因失活, 指的是当把外源基因导入生物体内时, 相应序列内源基因被抑制而不能表达的基因调控现象。不难看出, 转基因沉默属于同源性决定的或重复序列诱导的基因失活。此时, 其他基因并不受影响; 沉默基因恢复表达活性的时机是所转入外源基因与内源同源基因重组分离或减数分裂分离之后^[31]。植物转基因沉默发现于 1990 年, 研究人员在将查尔酮合成酶(chalcone synthase)基因向紫色牵牛花(*Petunia*)中转移时, 发现外源基因与内源同源查尔酮合成酶基因一起发生了抑制(共抑制, co-suppression)^[32]。后来, 人们发现植物受病毒侵染也会造成基因沉默, 与上述转基因沉默一样属于转录后基因沉默(post-transcription gene silencing, PTGS)^[33]。

20 世纪末到 21 世纪初的大量研究成果随转基因沉默的发现而产生。人们在反义基因诱导果蝇(*Drosophila pal-1*)基因沉默的研究中发现了双链 RNA(dsRNA)可致基因沉默, 随即又在线虫(*Caenorhabditis elegans*)中发现 dsRNA 比单链的反义 RNA(antisense RNA)能更有效诱导基因失活, 于是产生了 RNA 干扰(RNA interference, RNAi; 即 RNA 介导的基因沉默)的概念: 这是一种 PTGS 式的基因调节模式, 其中非编码 dsRNA 分子(小干扰 RNA, small interfering RNA, siRNA)介导靶 mRNA 的序列特异性降解^[34-36]。siRNAs 是 21~23 个核苷酸的 dsRNA, 含有对称的 2~3 个核苷酸的 3' 突出末端, 并有 5' 和 3' 羟基端(图 5)。



图 5 siRNA 的结构

Fig. 5 Structure of siRNA

发生 RNAi 时, 长的 dsRNA 分子被 Dicer 酶裂解, 产生 siRNA。随即, siRNA 分子掺入一个多蛋白因子的 RNA 诱导沉默复合体(multiprotein RNA-inducing silencing complex, RISC)中; dsRNA 解链, 其中的反义链(anti-sense strand)引导 RISC 到互补的 mRNA 上, 执行后续的核酸内切裂解反应(endonuclease cleavage)(图 6)。

人们研究真菌中导入类胡萝卜素基因发生的基因抑制, 研究拟南芥、斑马鱼、西红柿、涡虫、烟草和小鼠等各种生物的 RNA 介导性基因沉默, 逐渐探清了 RNAi 的机理。由于 RNAi, 即使能证实转基因已经整合到宿主基因组内并拷贝完整, 也常得不到稳定表达产物, 甚至完全得不到表达产物。

总体看, 植物的转基因沉默可以发生在 DNA 水平(位置效应, position effect)、转录水平(转录失活, transcription inactivation)和转录后水平(转录后基因沉默, PTGS)。

具体讲, 所导入的外源基因向宿主基因组的插入具有随机性, 若插入到转录不活跃区, 就会发生位置效应所致基因

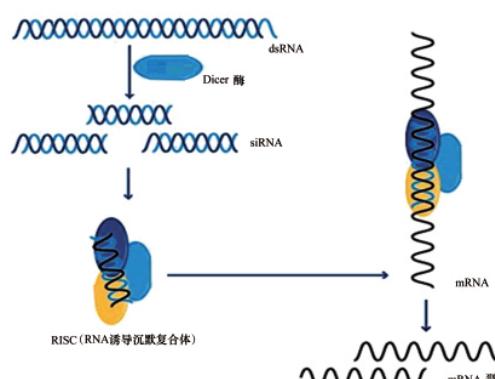


图 6 RNAi 的作用途径
Fig. 6 Acting path of RNAi

低表达或不表达。在染色质中,两个核基质结合区(nuclear matrix attachment region, MAR)间的基因片段被界定成一个独立的染色质环,作为隔离子(insulator)阻止附近顺式调控元件对环内基因表达的干扰^[37]。当转基因在受体基因组内整合后,有可能在 MAR 作用下形成环形结构单元。这可以解释 DNA 水平转基因沉默的成因及提高转基因表达水平的机理。

转基因所形成的 DNA 异位配对可造成异染色质化或从头甲基化,进而抑制转录,这一效应也可由 DNA-RNA 协同作用造成。这就诱发了转录水平的基因沉默。甲基化是活细胞中最常见的 DNA 共价修饰形式,通常发生于 GC 和 GNC 序列的 C 上,在高等植物和哺乳动物中尤其常见。几乎所有植物转基因沉默都与启动子甲基化有关;GC 和 GNC 序列的 C 上的甲基化虽不是转录水平的转基因沉默的前提,但却是维持这种沉默的必要方式。另外,多拷贝重复基因在宿主基因组内的整合,形成异位配对,引起基因组防御系统识别而被甲基化或造成异染色质化;转基因受染色体包装影响,转录因子的接触机会发生改变等,均是造成转录水平转基因失活的可能机理。

相比之下,转录后的转基因沉默更是研究热点。PTGS 的特点是外源基因可形成 mRNA 而并不发生累积,因为被立即降解或被反义 RNA、特异蛋白因子“封杀”了。这里除涉及 RNAi 以外,过量 RNA 也可导致同源基因重新甲基化而失活。有研究认为,细胞能容纳的外源基因量有一定的阈值,超过阈值就会启动监控机制而排除超量 RNA——内源 RNA 依赖的 RNA 聚合酶将多余的转基因 mRNA 反转录为 RNA 拷贝,去配对结合转基因和内源基因的 mRNA,形成的 dsRNA 被胞内 RNase 识别而遭到降解。另外,内外源异源配对的 RNA 会降低正常的 RNA 加工、转译效率,也是解释 PTGS 的模型。更多地研究表明,转基因在个体发育的某阶段会受到细胞内因子的后成修饰(epigenetic modification,即表观修饰),这与受体植物的核型有关;光控因子、种子特异发育因子和热休克调控元件等会在环境因素诱导下综合作用于转基因和内源基因,造成 PTGS。

转基因沉默说明了基因免疫(gene immunology)的存在,是基因组抵抗外源“入侵”基因的本领^[38]。

3.2 副突变

20世纪50年代,有研究者在玉米中发现了副突变(paramutation)现象。所谓副突变,是指具有同一位点的2个等位基因间的互作,导致其中一个等位基因发生可遗传的变化。随后,其他生物体中的副突变现象也被发现。副突变不符合孟德尔遗传模式,分子生物学技术研究表明其发生的基础机制由 RNA 引发,涉及一系列表观遗传修饰和染色体结构动态。副突变的研究有助于农作物改良和遗传疾病治疗。

需指出的是,副突变遗传现象也已在动物中被发现。研究小鼠的 *Kit* 基因(这种基因发生变异时,小鼠的尾巴上长出白色斑点)时发现,与无效突变体杂交后野生表型并不充分表达;虽然 *Kit*⁺/*Kit*⁺基因型按预期频率形成,但副突变使大多数个体仍有白点突变。这里涉及配子与合子间的 RNA 转移——起调控作用的微 RNA(microRNA)介导了基因活性调控,而且还传到了下一代(虽拥有 2 个正常 *Kit* 基因)身上。研究发现变异的 *Kit* 基因产生了大量 microRNA,并且也积聚在基因变异小鼠的精子中^[39]。

表观遗传的研究还在不断取得进展,尚有很多不明之处。对本课题组发现的一种新纺锤体蛋白 INMAP 的研究表明,这种蛋白基因表达异常(基因沉默)时细胞呈现着丝粒(centromere)结构松弛,荧光显示时出现月晕(halo)状表型,染色体不均等分离。利用间接免疫荧光实验检测这种纺锤体蛋白基因发生沉默后细胞的基因组变化,发现在这样的基因表达异常情况下,细胞中对于维持着丝粒中染色质 DNA 周期回折结构起重要作用的着丝粒蛋白 CenpB 发生裂解,致使着丝粒结构因有序的 DNA-蛋白复合结构不能维持而松解,造成细胞染色体分离异常和基因组不稳定。由此可以推测,重要功能基因表达异常可能通过修饰着丝粒蛋白而表观性改变遗传模式。

4 展望

表观遗传学使人们更深刻地认识到基因与表型间的关系,同时也很好地补充了“中心法则”中没有涉及到的 2 个问题:一是哪些因素影响了基因的正常转录与翻译;二是核酸是不是储存遗传信息的唯一载体^[40]。目前,研究发现的表观遗传修饰主要包括 DNA 甲基化、组蛋白修饰和非编码 RNA 等。这 3 种主要因素对于基因表达的影响机制的研究已经取得了一定的成果,即在分子水平上人们对于表观遗传如何发挥作用已经有了一定的认识,但这三者之间的相互关系及它们是如何共同调节基因表达的,还需要更深入的研究。

在生物体内,分子水平的变化会引起相应的表型变化。在研究表观遗传分子机制的同时,人们还发现了表观遗传修饰在生物生长、发育和疾病发生中发挥了很重要的作用,因此,表观遗传对于生物表型整体的影响是不可忽视的。相较于 DNA,表观遗传范畴的可遗传信息更易受环境的影响,更

容易发生动态、可逆的变化,因此也可以说表观遗传是灵活多变的遗传信息。正常的表观遗传过程对于胚胎发育、细胞分化等生命活动是非常重要的。简单地说,表观遗传在发育上的效力体现为早期环境因素对生物体晚期表型及后代表型的影响^[41]。越来越多的研究发现,很多疾病的发生与异常的表观遗传修饰相关,如哮喘^[42]、肝母细胞瘤^[43]和HMT异常引起的多种肿瘤^[43]、与涉及lncRNA作用的细胞特异性基因表达调节相联系的复杂疾病等。比如,一种lncRNA调节了干扰素IFN- γ 编码基因染色质的表观印记,影响该基因表达,与病毒、细菌病原体感染敏感性有明显关系^[44]。在发现表观遗传与疾病相关的同时,相应的药物研发与治疗方法也在逐渐摸索当中^[45]。可见,表观遗传学不仅在生物学领域内是研究热点,而且在医学方面也占有很重要的位置。系统和深入地研究表观遗传学有利于揭示生物的生长、发育和人类疾病等许多生命现象的本质。

参考文献(References)

- [1] Alber M, Helin K. Histone methyltransferases in cancer [J]. Seminars in Cell & Developmental Biology, 2010, 21(2): 209–220.
- [2] Choudhuri S. From Waddington's epigenetic landscape to small noncoding RNA: Some important milestones in the history of epigenetics research[J]. Toxicology Mechanisms and Methods, 2011, 21(4): 252–274.
- [3] 董玉玮, 侯进, 朱必才, 等. 表观遗传学的相关概念和研究进展 [J]. 生物学杂志, 2005, 22(1): 1–3.
Dong Yuwei, Hou Jinhui, Zhu Bicai, et al. Journal of Biology, 2005, 22 (1): 1–3.
- [4] Lovinsky-Desir S, Miller R L. Epigenetics, Asthma, and allergic diseases: A review of the latest advancements[J]. Current Allergy and Asthma Reports, 2012, 12(3): 211–220.
- [5] 梁前进. 表观遗传学——理论·方法·研究进展 (1)[J]. 生物学通报, 2007, 42(10): 4–7.
Liang Qianjin. Bulletin of Biology, 2007, 42(10): 4–7.
- [6] 梁前进. 表观遗传学——理论·方法·研究进展 (2)[J]. 生物学通报, 2007, 42(11): 11–12.
Liang Qianjin. Bulletin of Biology, 2007, 42(11): 11–12.
- [7] Nelissen E C, van Montfoort A P, Dumoulin J C, et al. Epigenetics and the placenta[J]. Human Reproduction Update, 2011, 17(3): 397–417.
- [8] Nakao M. Epigenetics: Interaction of DNA methylation and chromatin[J]. Gene, 2001, 278(1/2): 25–31.
- [9] Jones P A, Takai D. The role of DNA methylation in mammalian epigenetics[J]. Science, 2001, 293(5532): 1068–1070.
- [10] Theisen J W, Gucwa J S, Yusufzai T, et al. Biochemical analysis of histone deacetylase-independent transcriptional repression by MeCP2[J]. The Journal of Biological Chemistry, 2013, 288(10): 7096–7104.
- [11] Furumatsu T, Ozaki T. Epigenetic regulation in chondrogenesis[J]. Acta Medica Okayama, 2010, 64(3): 155–161.
- [12] 蒋智文, 刘新光, 周中军. 组蛋白修饰调节机制的研究进展 [J]. 生物化学与生物物理进展, 2009, 36(10): 1252–1259.
Jiang Zhiwen, Liu Xinguang, Zhou Zhongjun. Progress in Biochemistry and Biophysics, 2009, 36(10): 1252–1259.
- [13] Wolffe A P, Hayes J J. Chromatin disruption and modification[J]. Nucleic Acids Research, 1999, 27(3): 711–720.
- [14] Roth S Y, Denu J M, Allis C D. Histone acetyltransferases[J]. Annual Review of Biochemistry, 2001, 70: 81–120.
- [15] 王维, 孟智启, 石放雄. 组蛋白修饰及其生物学效应 [J]. 遗传, 2012, 34(7): 810–818.
Wang Wei, Meng Zhiqi, Shi Fangxiong. Hereditas, 2012, 34 (7): 810–818.
- [16] Albert M, Helin K. Histone methyltransferases in cancer[J]. Seminars in Cell & Developmental Biology, 2010, 21(2): 209–220.
- [17] 翟斐, 高鹏, 郑杰. 组蛋白SUMO化与肿瘤 [J]. 生命的化学, 2010, 30 (3): 333–337.
Zhai Fei, Gao Peng, Zheng Jie. Chemistry of Life, 2010, 30 (3): 333–337.
- [18] LeBaron M J, Rasoulpour R J, Klapacz J, et al. Epigenetics and chemical safety assessment[J]. Mutation Research, 2010, 705(2): 83–95.
- [19] Shi Y J, Matson C, Lan F, et al. Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell, 2005, 19(6): 857–864.
- [20] Mohammad F, Mondal T, Kanduri C. Epigenetics of imprinted long noncoding RNAs[J]. Epigenetics, 2009, 4(5): 277–286.
- [21] 孙甜甜, 房静远. 长非编码RNA与其他表观遗传机制在肿瘤中的相互调控[J]. 上海交通大学学报: 医学版, 2011, 31(7): 1022–1026.
Sun Tiantian, Fang Jingyuan. Journal of Shanghai Jiaotong University: Medical Science Edition, 2011, 31(7): 1022–1026.
- [22] 徐宏, 梁尚栋. 长非编码RNA与其他表观遗传的相互调控与神经系统疾病[J]. 中国药理学通报, 2012, 28(10): 1348–1351.
Xu Hong, Liang Shangdong. Chinese Pharmacological Bulletin, 2012, 28 (10): 1348–1351.
- [23] Jang H J, Lee M O, Kim S, et al. Biallelic expression of the L-arginine: glycine amidinotransferase gene with different methylation status between male and female primordial germ cells in chickens [J]. Poultry Science, 2013, 92(3): 760–769.
- [24] Davison A, Frend H T, Moray C, et al. Mating behaviour in lymnaea stagnalis pond snails is a maternally inherited, lateralized trait[J]. Biology Letters, 2009, 5(1): 20–22.
- [25] Iglesias N, Redon S, Pfeiffer V, et al. Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast[J]. EMBO Reports, 2011, 12(6): 587–593.
- [26] Greil F, Ahmad K. Nucleolar dominance of the Y chromosome in Drosophila melanogaster[J]. Genetics, 2012, 191(4): 1119–1128.
- [27] Lamprecht B, Walter K, Kreher S, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma[J]. Nature Medicine, 2010, 16(5): 571–579.
- [28] 李艳, 梁前进, 陶伟. 伴随X染色体失活的甲基化修饰[J]. 高技术通讯, 2007, 17(12): 1307–1311.
Li Yan, Liang Qianjin, Tao Wei. Chinese High Technology Letters, 2007, 17(12): 1307–1311.
- [29] Li Y, Tan T, Zong L, et al. Study of methylation of histone H3 lysine 9 and H3 lysine 27 during X chromosome inactivation in three types of cells[J]. Chromosome Research, 2012, 20(6): 769–778.
- [30] 邹永海, 施季森, 诸葛强, 等. 遗传转化植物中沉默基因的消除[J]. 分子植物育种, 2006, 4(1): 95–102.
Zou Yonghai, Shi Jisen, Zhuge Qiang, et al. Molecular Plant Breeding, 2006, 4(1): 95–102.
- [31] 刘悦萍, 赵晓萌, 宫飞. 转基因植物中外源基因沉默机制的研究进展 [J]. 中国农学通报, 2005, 21(4): 80–83.
Liu Yueping, Zhao Xiaomeng, Gong Fei. Chinese Agriculture Science Bulletin, 2005, 21(4): 80–83.

- [32] Napoli C, Lemieux C, Jorgensen R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible cosuppression of homologous genes in trans[J]. Plant Cell, 1990, 2(4): 279–289.
- [33] Romero I, Tikhunov Y, Bovy A. Virus induced gene silencing in detached tomatoes and biochemical effects of phytoene desaturase gene silencing[J]. Journal of Plant Physiology, 2011, 168(10): 1129–1135.
- [34] Dykxhoorn D M, Novina C D, Sharp P A. Killing the messenger: Short RNAs that silence gene expression[J]. Nature Review Molecular Cell Biology 2003, 4(6): 457–467.
- [35] Hammond S M. Dicing and slicing the core machinery of the RNA interference pathway[J]. FEBS Letters, 2005, 579(26): 5822–5829.
- [36] Lingel A, Sattler M. Novel modes of protein–RNA recognition in the RNAi pathway[J]. Current Opinion in Structural Biology, 2005, 15(1): 107–115.
- [37] Kalos M, Fournier R E. Potition-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain[J]. Molecular and Cellular Biology, 1995, 15(1): 198–207.
- [38] 王红梅, 张献龙, 蔡忠民, 等. 植物转基因沉默机制及消除对策[J]. 棉花学报, 2003, 15(4): 248–251.
Wang Hongmei, Zhang Xianlong, Cai Zhongmin, et al. Cotton Science, 2003, 15(4): 248–251.
- [39] Morelle R. Spotty mice flout genetics laws [EB/OL]. 2006–05–24. <http://news.bbc.co.uk/2/hi/science/nature/5011826.stm>.
- [40] 张永彪, 褚嘉祜. 表观遗传学与人类疾病的研究进展[J]. 遗传, 2005, 27(3): 466–472.
Zhang Yongbiao, Chu Jiayou. Hereditas, 2005, 27(3): 466–472.
- [41] Calle A, Fernandez-Gonzalez R, Ramos-Ibeas P, et al. Long-term and transgenerational effects of in vitro culture on mouse embryos[J]. Theriogenology, 2012, 77(4): 785–793.
- [42] Tomlinson G E, Kappler R. Genetics and epigenetics of hepatoblastoma [J]. Pediatr Blood Cancer, 2012, 59(5): 785–792.
- [43] Bartkova J, Moudry P, Hodny Z, et al. Heterochromatin marks HP1 γ , HP1 α and H3K9me3, and DNA damage response activation in human testis development and germ cell tumours[J]. International Journal of Andrology, 2011, 34(4pt2): e103–e113.
- [44] Gomez J A, Wapinski O L, Yang Y W, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- γ locus[J]. Cell, 2013, 152(4): 743–754.
- [45] Egger G, Liang G, Aparicio A, et al. Epigenetics in human disease and prospects for epigenetic therapy[J]. Nature, 2004, 429(6990): 457–463.

(责任编辑 刘志远, 陈广仁)

· 学术动态 ·



2013 年中国流动科技馆全国巡展启动

2013年6月21日,以“体验科学”为主题的中国流动科技馆全国巡展在青海省海北州刚察县启动。中国科协党组书记、书记处第一书记申维辰宣布2013年中国流动科技馆全国巡展启动。全国23个省、自治区、直辖市,以流动科技馆形式为尚未建设科技馆地区的公众提供科普服务的巡展工作全面开启。

中国流动科技馆以“体验科学”为主题,通过声光体验、电磁探秘、运动旋律、数学魅力、健康生活、安全生活、数字生活7个主题展区的50件互动展品,与科学表演、科学实验、科普影视相结合,为公众提供参与科学实践的场所,观众可以在这里感受“体验科学”的快乐。

2011年7月,中国科协启动在四川、贵州、青海等9省、自治区开展流动科技馆巡展试点工作。中国流动科技馆迄今已在84个县(市)巡展,受益人数300余万,其中中小学生参观比例占90%以上,取得了良好的社会效益。

详见中国科协网 <http://www.cast.org.cn/n35081/n35096/n10225918/14813302.html>。