

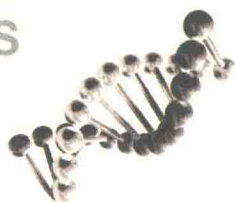


北京市科学技术委员会  
科普专项经费资助



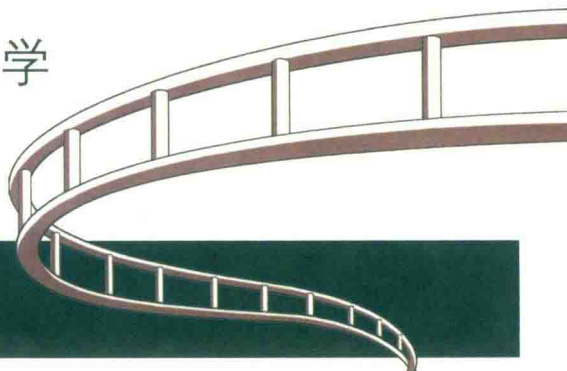
科学文化工程  
公民科学素养系列

# INHERITANCE IN LIFE THE SECRET OF EPIGENETICS



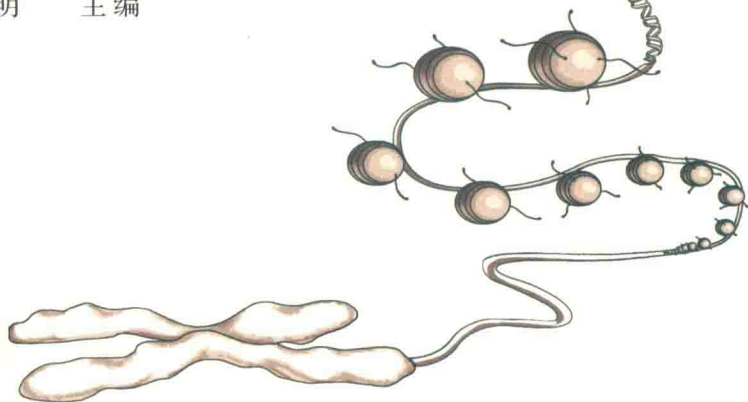
## 承 续 的 魅 力

令人着迷的表观遗传学



科技前沿“故事汇”

曹晓风 许瑞明 主编



科学出版社



北京市科学技术委员会  
科普专项经费资助



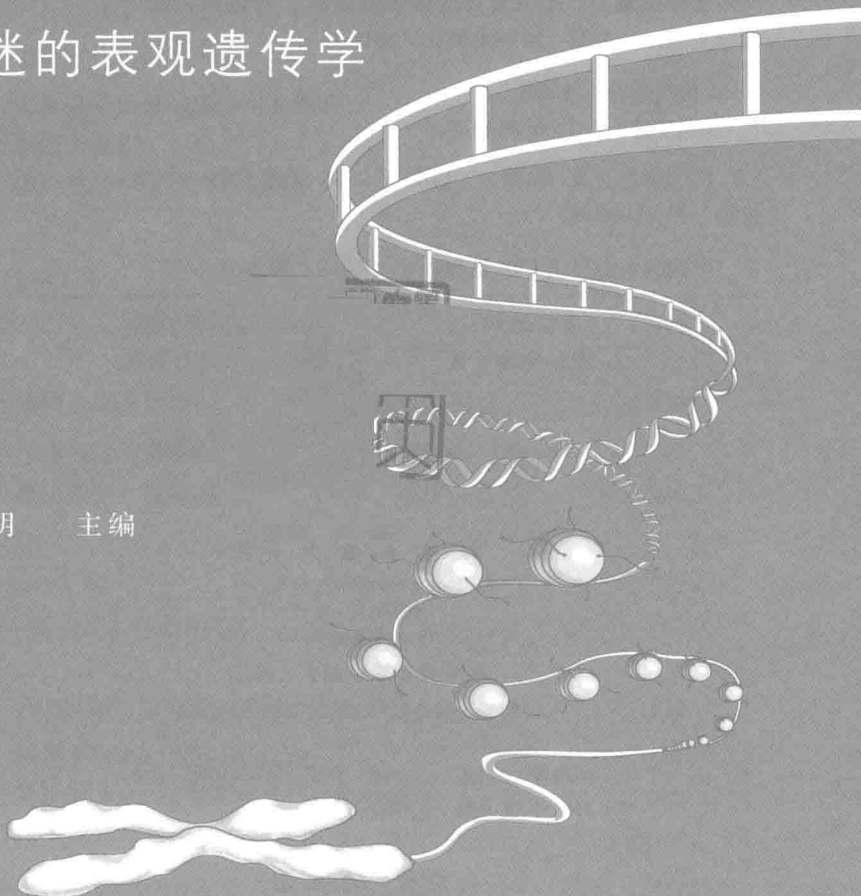
科学文化工程  
公民科学素养系列

科技前沿“故事汇”

# 承 续 的 魅 力

## 令人着迷的表现遗传学

曹晓风 许瑞明 主编



科学出版社

## 内 容 简 介

近二十年来,表观遗传的研究正如火如荼地开展,研究内容涵盖了从动物、植物到真菌,从干细胞、胚胎发育到衰老、癌症发生等生命科学研究方方面面。那么什么是表观遗传,为什么它的研究占据了生命科学研究的重要地位呢?本书邀请了我国表观遗传领域的十余位优秀科学家对各自的研究方向进行了介绍。

### 图书在版编目(CIP)数据

承续的魅力:令人着迷的表观遗传学 / 曹晓风, 许瑞明主编. —北京: 科学出版社, 2018.11

(科技前沿“故事汇”)

ISBN 978-7-03-059640-6

I. ①承… II. ①曹… ②许… III. ①遗传学—普及读物 IV. ①Q3-49

中国版本图书馆 CIP 数据核字 (2018) 第 262949 号

责任编辑: 王亚萍 / 责任校对: 杨 然  
责任印制: 师艳茹 / 内文设计: 楠竹文化

编辑部电话: 010-64003228

E-mail: wangyaping@mail.sciencep.com

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

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销



2018 年 11 月第 一 版 开本: 720×1000 1/16

2018 年 11 月第一次印刷 印张: 14 3/4

字数: 240 000

定价: 60.00 元

(如有印装质量问题, 我社负责调换)



# 目 录

## Chapter 1

### 表观遗传的奥秘

- |                    |     |
|--------------------|-----|
| 01 表观遗传的前世今生       | 003 |
| 02 表观遗传信息的遗传与重编程   | 008 |
| 03 组蛋白修饰           | 023 |
| 04 表观遗传中重要的非编码 RNA | 051 |

## Chapter 2

### 表观遗传与植物发育

- |                      |     |
|----------------------|-----|
| 01 植物表观遗传调控自然界的“转基因” | 089 |
| 02 植物再生的表观遗传调控机制     | 112 |
| 03 春化作用：植物的低温记忆与表观遗传 | 121 |

## Chapter 3

### 表观遗传与人类健康

- |                       |     |
|-----------------------|-----|
| 01 DNA 甲基化与胃肠癌的诊断和防治  | 139 |
| 02 组蛋白修饰与胃肠癌的防治       | 146 |
| 03 非编码 RNA 与胃肠癌的诊断和预警 | 152 |
| 04 衰老调控过程中的表观遗传       | 187 |
| 05 表观遗传调控和男性生殖健康      | 202 |
| 06 表观遗传干预的药物研发        | 220 |



# Chapter 1

## 表观遗传的奥秘

- / 01 表观遗传的前世今生
- / 02 表观遗传信息的遗传与重编程
- / 03 组蛋白修饰
- / 04 表观遗传中重要的非编码 RNA



## 01

## 表观遗传的前世今生



## 遗传物质的继承是生命繁衍的根本

生命体都是由细胞构成的。遗传物质就是细胞内承载生命繁衍的基本物质，可以从亲代传递到子代。奥地利修道士孟德尔（Gregor Johann Mendel）通过豌豆杂交实验，研究了不同的性状在亲本和子代间传递的基本规律，提出了性状遗传的单位是遗传因子，即我们现在所熟知的基因。美国遗传学家摩尔根（Thomas Hunt Morgan）进一步证实了基因是细胞核中清晰可见的物质，排列在“可以被染色的小体”，即染色体上。孟德尔和摩尔根因在遗传规律和遗传物质的发现中作出重要贡献，分别被誉为“现代遗传学之父”“现代实验生物学奠基人”。



## 分子遗传学开启了遗传学的新篇章

1869年，瑞士生物学家弗雷德里希·米歇尔（Friedrich Miescher）从分离的白细胞细胞核中鉴定出一种含有磷元素和氮元素的物质，命名



为“核素”（Nuclein），揭示了遗传物质的分子本质，将遗传学研究带入了分子遗传学时代。分子遗传学大大推进了人类对遗传现象的理解和对遗传规律的探索。我们的染色体一半来自父亲、一半来自母亲，遗传物质决定了我们的性状。现在我们都知道，基因是一种带有遗传信息的脱氧核糖核酸（deoxyribonucleic acid, DNA）片段。DNA 分子由核苷酸组成，核苷酸的含氮碱基为腺嘌呤（adenine, A）、鸟嘌呤（guanine, G）、胞嘧啶（cytosine, C）及胸腺嘧啶（thymine, T）。真核生物的 DNA 并不是裸露存在的，而是缠绕在蛋白质上形成核小体，核小体串又进一步折叠压缩形成染色体。组成核小体的蛋白质是一类富含碱性氨基酸残基、具有特征折叠类型的组蛋白，主要包括 H3、H4、H2A 和 H2B 四种核心组蛋白，以及核小体连接处的组蛋白 H1。



## 分子遗传学的拓展和补充——表观遗传学

表观遗传学是分子遗传学发展的必然产物，其兴起主要是通过通过对一些“异常”现象的观察而来的。人们逐渐发现一些现象无法用经典的遗传学规律解释，如突变果蝇的“花眼睛”（花斑位置效应）、后腿“多余的”性器官（多梳基因沉默），雌性动物为何有一条失活的 X 染色体（剂量补偿效应）等。还有一些更为直观的“异常”现象，例如，同卵双胞胎虽然拥有完全相同的基因，但随着逐渐成长，不同的性格、外表会越来越凸显；再者，随着体细胞克隆技术的发展，人们希望通过克隆基因获得性状一模一样的生物，然而现实并不那么“理想”，特别是第一只克隆动物 CC 猫（Copy Cat）的诞生，让她的主人和大众大跌眼镜，因为 CC 长得和她的母亲实在是相差太多了；另一个有意思的现象则是蜂王和工蜂拥有相同的基因，为何只有一个成为“女王”，其他的则沦为“仆人”呢？这与食用蜂王浆有什么关系呢？还有一个值得思考的问题，那就是人体存在 200 多种细胞，它们都是由一个受精卵分化发育而

来的，具有相同的基因，但为何会选择性地发育成不同类型的细胞，在细胞分化过程中“走”上了不同的“道路”？上述这些问题似乎提示着“基因不是我们的全部宿命”，存在一些“超越”基因的秘密可以影响遗传性状。

表观遗传的概念翻译自英文 Epigenetic，其中“Epi”是“外在”或“超越”的意思，“genetic”是遗传，合在一起意味着“超越遗传信息本身的其他因素”，即 DNA 序列未发生变异而性状发生可遗传变化的现象。简单来讲，就是虽然基因相同，但基因表达的模式不同，得到的功能行使者——蛋白质是不同的，因此会产生不同的性状。

这些影响基因表达的因素是表观遗传学研究的主要内容，包括 DNA 的甲基化修饰、组蛋白上的多种修饰（合称表观修饰），还有染色质重塑、组蛋白变体、非编码 RNA 等。这些因素共同影响染色体的高级结构，调节染色体局部结构的动态变化，简而言之，就是如果染色体缠绕得很紧密，则不利于基因表达；如果染色体结构松散，则利于该部位的基因表达。

表观修饰是表观遗传分子机制研究的最早、最为核心的内容，其研究通常分为三个方面，包括书写者（Writer，指加上修饰的酶）、擦写者（Eraser，指去掉修饰的酶）及读取者（Reader，指识别修饰的因子），三者协同作用调节表观遗传信息的建立、维持和解读。举一个简单的例子，有助于大家理解这些化学修饰是如何改变染色质结构的。我们知道 DNA 是带负电荷的，而组蛋白带有很多正电荷，所以 DNA 缠绕在组蛋白上主要是由正负电荷相互作用而形成的。组蛋白中带正电荷最多的就是赖氨酸残基，如果赖氨酸的侧链发生了乙酰化修饰会中和其所带的正电性，不仅会减弱 DNA 与组蛋白之间的电荷作用，使得 DNA 更容易从组蛋白上脱落，而且这种化学修饰有利于“招募”转录活化复合物，导致该染色质区域的开放和基因激活；DNA 的甲基化则多与组蛋白的去乙酰化密切相关，因此沉默了基因的表达；而组蛋白的甲基化修饰则要复杂精细得多，对基因表达的调控也是位点依赖和修饰程度依赖决定的。例如，

组蛋白 H3 的第 9 位赖氨酸的三甲基化修饰（简称为 H3K9me<sub>3</sub>，后同）和 H3K27me<sub>3</sub> 及 H4K20me<sub>2</sub> 都与基因沉默密切相关，而 H3K4me<sub>3</sub> 和 H3K36me<sub>3</sub> 则与开放的基因表达上调相关。与乙酰化对电荷作用影响不同的是，这些甲基化的信号主要依赖一类称为“效应因子”的蛋白“识别和读取”表观遗传修饰，将信息“解码”并传递给下游效应分子。



## 表观遗传是如何遗传的？

表观遗传的一个重要特征就是可以被遗传。比如，之前提到的细胞分化问题，我们人体内 200 多种细胞都是由一个受精卵分化发育来的，一旦分化完成后，在自然条件下是不可逆转的。若是心肌细胞，那再分裂后依然是心肌细胞；若是神经细胞，那其“后代”也是代代“神经”。这种遗传非常重要，如果组成器官的细胞可以随意变为其他种类的细胞，那是一件很危险的事！因此，这种伴随体细胞增殖（有丝分裂）而产生的表观遗传，可以将表观信息传递给下一代细胞。特别强调的是，有些学者将表观遗传与“拉马克主义”<sup>①</sup>联系在一起，但现有的证据尚不足以支持后天获得性状的世代间遗传，在此要特别声明两点，表观遗传不挑战遗传（基因）的重要性，不代表拉马克主义的复兴。

本书旨在通过对不同研究领域的介绍，向广大读者传播表观遗传中一些有趣的现象和这些现象背后的分子机理，特别是表观遗传对人类健康生活影响的方方面面。第一篇章首先介绍了表观遗传的概念，继而通过三个例子介绍了表观遗传的继承性和细胞重编程，即遗传因素和非遗传因素，然后详细介绍了三个主要研究内容，包括 DNA 甲基化、组蛋白修饰及非编码 RNA。因为篇幅所限，组蛋白变体、染色质重塑等部分

<sup>①</sup> 19 世纪初期，法国生物学家拉马克继承和发展了前人的生物进化思想，进一步提出生物是由低级向高级发展进化，提出“用进废退”和“获得性遗传”原则，即认为经常使用的器官就发达，不用则会退化；后天获得的新性状也可能遗传下去。

的内容未单独列出，但在书中后续的部分也会有所涉及。第二篇章介绍了植物中的表观遗传现象，生动有趣地描述了植物的“天然转基因”“自体再生”和“春化开花”过程，以及这些过程如何受表观因素的影响和调控，其研究都与农业的增产和农产品的改良密切相关，特别是那些天然的转基因优良品种，让人看着不禁想立即品尝一番。第三篇章则介绍了表观遗传与人类健康的关系，如肿瘤、衰老、肥胖及男性生殖健康等严重影响我们生活质量和生命价值的重要疾病。与疾病对应的自然就是药物，当前国际上以表观因子为靶点的药物研究方兴未艾，书中最后一节也概述了表观遗传干预的药物研究的进展及未来发展方向。综上所述，希望读者通过对本书的阅读，了解表观遗传这门新兴学科的概念、发展历程及其研究内容，通过揭示“表观”的内涵，共同体会“承续”的魅力，见证生命的伟大之处。



许瑞明

### 作者简介：

中国科学院生物物理研究所研究员、中国科学院大学教授、首批“千人计划”引进人才、国家杰出青年科学基金获得者、全国优秀科技工作者、美国科学促进会会士、973首席科学家、科技部重点领域创新团队负责人、国家基金委创新群体学术带头人。主要研究方向为基因表达的表观遗传调控。

## 02

# 表观遗传信息的遗传与重编程



表观遗传研究领域的一个重要而有趣的问题是：表观遗传信息能否像 DNA 序列一样，从父母传递到下一代呢？为了回答这个问题，让我们先从两性生殖说起。

我们知道，人类的出生需要父母的遗传物质结合。我们享受着双亲无微不至的照顾，在他们的悉心呵护下茁壮成长。但是，一个很有意思的问题是，为何包括人类在内的很多高等生物，必须通过父母有性生殖才能产生后代？

### 两性生殖的奥秘：DNA 甲基化介导的表观遗传信息的遗传及重编程

为了探索上述问题，早在 20 世纪 80 年代，科学家们就进行了一项有趣的实验。在自然条件下，两只雄性小鼠或两只雌性小鼠是无法通过正常受精产生后代的。科学家通过显微操作，将受精卵中雄（雌）原核<sup>①</sup>

---

<sup>①</sup> 原核指的是父母双方的细胞核还没有融合的时期。

置换成来自另外一个细胞的雌（雄）原核，这样人工制造形成具有两套母本基因组或两套父本基因组的“受精卵”，科学家们称之为孤雌或孤雄胚胎（图 1-1）。检测发现，这种“单亲”胚胎是不能正常发育成小鼠的，而只有那些同时具有双亲细胞核的胚胎才可以正常出生。当胚胎发育一段时间后，科学家们发现，正常胚胎中一部分细胞会分化为胚胎外部的营养和支持组织（胚外组织，如一些细胞后续会发育成提供营养的胎盘），另一部分细胞则将分化为胚胎个体本身（胚内组织）。相反，孤雌生殖的胚胎中细胞更多地发育为胚外组织；而孤雌胚胎中的细胞则恰恰相反，会更多地向胚内组织发育。因此仅仅具有单亲基因组的胚胎无法正常发育成为完整的个体<sup>[1-4]</sup>。

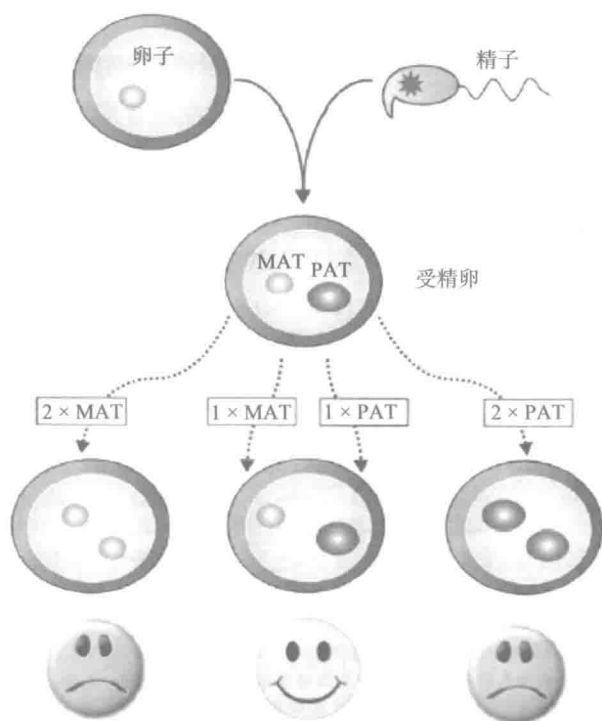


图 1-1 孤雌及孤雄生殖示意图<sup>[1-4]</sup>

注：MAT 为母本基因组；PAT 为父本基因组。

从这些尝试中科学家们发现，无论是母本还是父本基因组，似乎都带有某些该亲本所特有的信息，而这些信息是子代胚胎发育所必需的。进一步的实验表明，这些信息似乎不是来源于 DNA 遗传序列的差别，因为科学家可以用父本和母本 DNA 序列几乎完全相同的小鼠进行以上实验，但是拥有两套母本或父本基因组的小鼠仍然不能正常发育和出生。这表明，父母双方的差别不仅在于 DNA 遗传序列方面，也包括 DNA 序列之外的一些信息。这些和 DNA 序列本身无关，但是可以遗传给子代的信息被称为表观遗传信息。缺失父母任何一方的表观遗传信息，胚胎都无法正常发育。

随着研究越来越深入，科学家们发现这些异常发育的胚胎中有一些基因，其来源于父本和母本的表达有很大差别。比如，*H19* 基因可以控制胚胎大小及细胞生长。它的异常表达不仅导致胚胎发育异常，还可能引发某些癌症的发生<sup>[5,6]</sup>。这个基因的特点是在正常发育的胚胎中，来自于母亲的基因拷贝表达，而来自于父亲的基因拷贝不表达。如果有两个母本的拷贝，*H19* 就会过量表达；而如果有两个父本的拷贝，*H19* 表达量就会太低（图 1-2）。这种亲本间基因表达差别是由于调控该基因表达

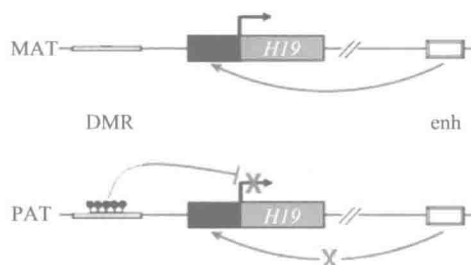


图 1-2 印记基因 *H19* 调控表达机制<sup>[13]</sup>

DMR 为 Differentially Methylated Region，差异甲基化区域；enh 为 enhancer，即增强子，可调控基因表达，是提高基因转录水平的基因组 DNA 元件；MAT 为母本基因组；PAT 为父本基因组。其中差异甲基化区域（DMR）在母本基因组中保持低甲基化，因而能够“招募”一类绝缘子蛋白——CTCF 蛋白，结合到该区域，从而促使该增强子与 *H19* 基因发生相互作用，增强 *H19* 在母本基因组的表达；而在父本基因组上，该区域甲基化水平较高，导致 CTCF 蛋白无法结合到该区域，使得增强子无法与 *H19* 基因联系，抑制了 *H19* 基因的表达。

的 DNA 序列上甲基化程度存在差异。这种来源于不同亲本的 DNA 序列有差异甲基化的现象被称为基因组印记 (genomic imprinting)。

由此可见, DNA 甲基化对基因组印记的形成非常重要。那么何为 DNA 甲基化呢?

DNA 甲基化是指在 DNA 甲基转移酶的作用下, 将供体上甲基 ( $\text{CH}_3$ ) 基团转移至胞嘧啶 (C) 的 5' 碳原子上, 形成甲基化胞嘧啶的过程。在哺乳动物体内, 该反应主要发生在基因组序列 CpG 位点上<sup>[7]</sup>。一般来说, 启动子区域甲基化水平与基因表达水平呈现负相关。启动子区高甲基化与基因沉默相关; 相反, 基因表达通常伴随着启动子区甲基化的降低或缺失<sup>[8]</sup>。DNA 甲基化的建立主要依赖于 DNA 甲基转移酶 DNMT 家族。该家族蛋白能通过 S-腺苷甲硫氨酸 (S-adenosyl-methionine, SAM) 提供的甲基来修饰胞嘧啶。DNMT 蛋白家族主要包括 DNMT1、DNMT2、DNMT3A、DNMT3B 和 DNMT3L 五个成员。最新研究发现了一种新的 DNA 甲基转移酶, 被命名为 DNMT3C<sup>[9]</sup>。而 DNA 去甲基化的机制比较复杂, 其中被广泛研究的主要是氧化去甲基化过程中的 TET (ten-eleven translocation) 蛋白家族, 包括 TET1, TET2 及 TET3。TET 蛋白是一种依赖于  $\alpha$ -酮戊二酸及二价铁离子的双加氧酶, 通过氧化催化 5-甲基胞嘧啶 (5mC) 转化为 5-羟甲基胞嘧啶 (5hmC), 并最终被未被修饰的胞嘧啶替代, 从而实现甲基化主动去除<sup>[10]</sup>。

DNA 甲基化是生物体内最重要的表观遗传标记之一, 既参与调控基因转录、维持染色体结构, 也在基因印记、X 染色体失活等方面发挥重要作用<sup>[7,11]</sup>。DNA 甲基化异常也是导致癌症发生的原因之一<sup>[12]</sup>, 因此, 维持体内的甲基化水平正常对生物体生长发育十分重要。

基因印记主要指在配子形成过程中, 针对某些特别的基因, 父本和母本携带不同的表观遗传修饰, 导致受精后在胚胎发育过程中不同亲本来源的两个等位基因在子代细胞中差异化表达。当印记基因在亲代和子代的传递过程中发生障碍时, 这些基因就会发生表达失调, 最终导致胚胎发育异常<sup>[13]</sup>。



在了解到印记基因的原理及其功能后，科学家们大胆设想：如果能够人为干预亲代遗传物质上的印记基因，使两套同样亲本（父本或母本）基因组带有正常双亲所具有的印记信息，那么是否可以通过同性生殖产生正常后代呢？经过不懈的努力，他们成功地实现了这个听上去有些匪夷所思的设想。日本科学家通过对两种印记基因 *H19* 及 *Dkl1* 进行基因编辑，使其表达水平与正常胚胎相近，从而得到了仅有母本基因组的小鼠胚胎<sup>[13,14]</sup>。而该实验也证实了，基因印记是决定胚胎发育成功与否最主要的关卡之一。

印记基因不仅在胚胎发育过程中发挥重要作用。在临床研究中，印记基因的异常表达也是导致很多遗传疾病的原因之一。

Prader-Willi 综合征（Prader-Willi syndrome, PWS）和 Angelman 综合征（Angelman syndrome, AS）是两种临床表征差异很大的神经遗传性疾病。Prader-Willi 综合征患者表现为身材肥胖、矮小，有智力障碍，手足异常，促性腺激素分泌不足。Angelman 综合征患者表现为智力低下，出现癫痫的情况，语言、运动障碍，面容特征为下颌巨大、张口吐舌<sup>[16]</sup>（图 1-3）。虽然两种疾病的表征差异很大，但是它们都是由同一染色体区段（15q11-13）缺陷所引起的。不同的是，PWS 是父本染色体 15 号染色体存在缺陷，而 AS 则为母本染色体同一区段存在缺陷。PWS 和 AS 的致病机理之一就是该区段的 DNA 缺失及 DNA 甲基化异常。例如，在 PWS 患者的父本染色体上发现了母本特有的甲基化印迹，而 AS 患者则相反<sup>[17]</sup>，从而导致相关印记基因表达异常。事实上，目前所发现的印记基因大多与胎儿的生长发育或其他疾病相关，并且它们的调控很多都和 DNA 甲基化紧密相连。

上述例子证明，父母的基因印记确实对生殖发育至关重要。然而，在形成下一代卵子和精子的过程中，细胞要从原始生殖细胞（primordial germ cell, PGC，指精子和卵子的前体细胞）经过减数分裂形成单倍体的配子，这时候父母的染色体会随机分配到配子中。比如，新形成的卵子中的基因印记就会有两种可能，一种是来源于母亲的，而另一种是来源