DOI:10.16262/j.cnki.1000-8217.2001.02.002

·学科进展·

基因组印迹基因及其生物学意义

訾晓渊 熊 俊 胡以平

(第二军医大学细胞生物学教研室,上海 200433)

[摘 要] 基因组印迹(genomic imprniting)是指在配子或合子发生期间,来自亲本的等位基因或染色体在发育过程中产生专一性的加工修饰,导致后代体细胞中两个亲本来源的等位基因有不同的表达活性,为一种后生论修饰。目前在人类和小鼠中鉴定的印迹基因已超过25个,它们具有一些共同的特点,如印迹基因分布的群集性、复制的不同步性、表达的时空特异性、遗传的保守性及编码RNAs等。基因组印迹的分子机理与印迹基因的甲基化尤其是CpG岛甲基化密切相关,特异性甲基化区域在印迹基因的表达中具有重要作用。基因组印迹基因与胎儿和胎盘的生长发育及细胞增殖有关,正常印迹的改变可引起包括肿瘤在内的多种遗传性疾病。

[关键词] 基因组印迹,印迹基因,甲基化,生物学意义

经典孟德尔遗传学认为所有父系及母系等位基因有同等表达,但随着对遗传学研究的深入,人们发现了一种被称为基因组印迹(genomic imprinting)的非孟德尔遗传的现象,它是指在配子或合子发生期间,来自亲本的等位基因或染色体在发育过程中产生专一性的加工修饰,导致后代体细胞中两个亲本来源的等位基因有不同的表达活性,又称遗传印迹或亲代印迹(parental imprinting)或配子印迹(gametic imprinting)。它是一种伴有基因组改变的非孟德尔遗传形式,可遗传给子代细胞,但并不包括DNA序列的改变。至今,在人类和小鼠中鉴定的印迹基因已超过25个,它们具有一些共同的特征,印迹基因的甲基化可能是基因组印迹的分子机制,特异性甲基化区域(DMRs)是控制印迹表达的重要因素。印迹基因中大多数在生长和分化中起重要作用,而且可引起一些人类遗传性疾病和肿瘤的发生。本文拟对基因组印迹基因的发现、印迹基因的特征及其生物学意义作一综述。

1 基因组印迹基因的发现

基因组印迹基因的发现最初开始于对全基因组的研究,接着是对个别染色体和染色体区域的研究,最终到鉴定特异的印迹基因。1989年,Swain和Lende等人发现了一个特异的印迹基因,但它并不是一个内源性基因。他们发现融合了免疫球蛋白增强子的C-myc转基因如来自父本,则在一些组织中表达,如果来自母本则表现为转录沉默,而且这些母源拷贝被甲基化。目前还发现了一些父源转基因甲基化的现象,但转录沉默现象比较少[1]。

DeChiara Efstradiatis等人通过基因剔除技术破坏小鼠胰岛素样生长因子Ⅱ(Igf2)基因发现了第一个内源性印迹基因,若被剔除的等位基因源于父本,则动物表现为侏儒,相反如为母源则无特殊表型,这些本身剔除了等位基因的雌鼠,其子代大小正常,原位杂交及RNase保护试验均证明剔除了父本等位基因的小鼠其组织中不表达Igf2,这些实验表明Igf2被印迹而且仅父源等位基因正常表达。这是一个里程性的研究,因为它们不仅表明基因组印迹可影响正常内源性基因,而且表明印迹具有组织特异性调节作用[2]。

在Igf2基因发现的同时,人们发现小鼠甘露糖6-磷酸/胰岛素样生长因子2型受体(M6p/Igf2r)基因也为印迹基因,其母源基因表达。有趣的是,Igf2和M6p/Igf2r在生化水平上相互作用,当小鼠M6p/

国家自然科学基金资助项目.
本文于2000年9月4日收到.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Igf2r 基因突变后,母源杂合子或纯合子胎鼠生长增加约 30%,这些表明,Igf2 和 M6p/Igf2r 印迹基因在胚胎发育和胎儿生长中都有重要作用。目前在人类和小鼠中已鉴定了 25 个印迹基因(见表 1)[4],而且新的印迹基因仍在不断的鉴定,如最近又发现了 IAC/plagl1 和 Tih1 等新基因[5,6]。

表 1　印迹基因的定位与表达

人　　类			小　　鼠		
基　因	定　位	等位基因表达	基　因	定　位	等位基因表达
NOEY2(ARHI)	1p31	父本			
p73	1p36	母本			
U2AFBPL	5q22-q31	双等位	U2afbp-rs	Proximal 11	父本
MAS1	6q25.3-q26	双等位/单等位(乳房)	Mas	Proximal 17	父本
M6P/IGF2R	6q26-q27	双等位/母本	M6p/Igf2r	Proximal 17	母本
			Igf2r-AS	Proximal 17	父本
GRB10	7q11.2-12	未报道	Meg1/Grb10	Proximal 11	母本
PEG1/MEST	7q32	父本	Peg1/Mest	Proximal 6	父本
WT1	11p13	双等位/母本	Wt1	2	未报道
ASCL2/HASH2	11p15.5	母本	Mash2	Distal 7	母本
H19	11p15.5	母本	H19	Distal 7	母本
IGF2	11p15.5	父本	Igf2	Distal 7	父本
			Igf2-AS	Distal 7	父本
IMPT1/BWR1A	11p15.5	母本	Impt1	Distal 7	母本
INS	11p15.5	双等位	Ins2	Distal 7	父本
IPL/TSSC3/BWR1C	11p15.5	母本	Ipl	Distal 7	母本
ITM	11p15.5	未报道	Itm	Distal 7	母本
KvLQT1	11p15.5	母本	Kvlqt1	Distal 7	母本
$p57^{KIP2}$/CDKN1C	11p15.5	母本	$p57^{KIP2}$	Distal 7	母本
TAPA1	11p15.5	双等位	Tapa1	Distal 7	母本
HTR2A	13q14	双等位/母本	Htr2	14,Band D3	母本
FNZ127	15q11-q13	父本			
GABRA5	15q11-q13	父本	Gabra5	Central 7	双等位
GABRB3	15q11-q13	父本	Gabrb3	Central 7	双等位
GABRG3	15q11-q13	父本	Gabrg3	Central 7	双等位
IPW	15q11-q13	父本	Ipw	Central 7	父本
NDN(necdin)	15q11-q13	父本	Ndn	Central 7	父本
PAR1	15q11-q13	父本			
PAR5	15q11-q13	父本			
PAR-SN	15q11-q13	父本			
SNRPN	15q11-q13	父本	Snrpn	Central 7	父本
UBE3A	15q11-q13	母本	Ube3a	Central 7	母本
ZNF127	15q11-q13	父本	Zfp127	Central 7	父本
PEG3	19q13.4	父本	Peg3/Apoc2	Proxima l7	父本
Neuronatin	20q11.2-q12	未报道	Peg5/Nnat	Distal 2	父本
GNAS1	20q13	父本	Gnas1	Distal 2	母本/父本
XIST	Xq13.2(XIC)	父本	Xist	Xic	父本
			$Grf1/Cdc25^{Mm}$	Distal 9	父本
			Impact	Proximal 18	父本
			Ins1	Distal 19	父本

2　印迹基因的特点

2.1　印迹基因分布的群集性

目前已知的一些印迹基因在基因组中的分布不是单一基因分散分布,而具有一定的群集性倾向,在这些群集区,父源和母源印迹基因均可找到。最大的印迹基因群集区之一位于小鼠 7 号染色体末端和人类 11p15.5 处,在约 1.5 Mb 的群集区内有 6 个印迹基因,在群集区的末端有 3 个母源表达基因:$p57^{KIP2}$,Kvlqt1 和 Mash2,中间为两个父源表达基因 Insulin-2(Ins-2)和 Igf2,在下游 90 kb 处为一母源表达基因 H19[7,8]。

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

另一个研究较多的印迹基因群集区位于人类 15 号染色体上，由于该区与 PWS 和 AS 两种人类遗传疾病有关，因此人们对这一群集区很感兴趣，目前已发现在此群集区至少包括 ZNF127、VBE3A、5NRPN、PAR-SN、PAR5、PAR1、NDN（necdin）、IPW、GABRG3、GABRB3、GABRA5 和 FNZ127 等 12 个印迹基因[4]。小鼠 2 号染色体远端 Gnas 位点处也有一印迹群，包括 4 个印迹基因：Nesp，Nespas，Gnasx1，Gnas，它们可形成一个印迹转录单位[9]。

最大的印迹基因群集区为 X 染色体本身，在真兽亚纲哺乳动物中雄性动物所有体细胞中两条染色体中的一条失活形成 X 连锁基因的剂量互补效应。虽然 X 染色体失活可在胚胎细胞中父源和母源 X 染色体间自由选择，但在胚胎发生时胚外细胞的胎盘滋养层和卵黄囊内胚层细胞中，这种选择由父系决定，也就是说父系 X 染色体优先失活，且父源基因沉默[10]。

以上几个群集区有以下几个共同特点：在一个很大的区域中群集分布着印迹基因，这些基因包含父源和母源表达基因，并且无编码 RNA。这些共同特点表明印迹可能不是一个或几个印迹基因[11]。

2.2 印迹基因 DNA 复制的不同步性

印迹基因的第 2 个特点是复制的不同步性。复制的不同步性首先做为 X 染色体失活的标记而发现，在细胞周期中失活 X 染色体的复制比有活性 X 染色体复制晚。Kitsberg 等人首先发现印迹基因复制的不同步性，他们用 FISH 技术通过对含有一个或两个杂交斑细胞所占的比例来分析复制的不同步性，结果发现 Igf2r、Igf2、H19 和 Snrpn 的两个亲源等位基因的复制不同步，它们的父源等位基因复制较早。细胞周期中的复制时间常与基因表达水平有关，早复制的基因有活性，但印迹基因不遵循这一原则，H19 和 Igf2r 基因为母源表达基因，但其父源基因复制早。还有少数基因不管哪个等位基因表达，其父源基因总是早复制[12]。Kawame 等人用 5-溴脱氧尿核苷直接脉冲标记细胞的方法来测定 DNA 复制的不同步性，结果发现，在人类淋巴细胞和原始淋巴细胞系中 Igf2 和 H19 的两个等位基因复制均晚，而这些细胞中不表达 Igf2 和 H19[13]。

2.3 基因组印迹基因表达的时空特异性

人们已发现了很多印迹基因表达的空间特异性的例子。如小鼠 Ins-2 基因仅在小鼠胚胎的胚外组织中印迹，但在胰岛细胞中双等位基因均表达；人类 KvLQT1 基因在大多数组织中为母源表达，但在心脏中无印迹，这可能是与该基因有关的长 QT 综合征无亲本遗传倾向的原因。

印迹基因表达的时间特异性的例子也很多：例如小鼠胚胎中的 Igf2 和 Mash2 开始为双等位基因表达，但在胚胎形成的后期则表现为母源等位基因表达。人类印迹基因中 Igf2 基因的表达也有时间改变，在个体发育期，Igf2 基因由 3 个独立的启动子启动父源染色体上的等位基因表达，出生后，肝脏中 Igf2 基因远端的启动子在两条染色体上均有活性，结果导致成人中双等位基因的表达。与上述例子相反，小鼠 H19 和 Snrpn 基因在发育的全过程中均印迹；人类 H19 基因仅在发育早期胎盘滋养层的少数细胞中无印迹，但在后期阶段完全印迹[11]。

2.4 真兽亚纲哺乳动物中印迹基因的保守性

人们对小鼠和人类基因组印迹基因研究发现人类和小鼠的印迹基因既有相似又有不同点，例如，小鼠印迹基因 H19、Igf2、$P57^{KIP2}$在人类也存在；相反，在小鼠中印迹的 Igf2r 基因在人类大部分为双等位基因表达，与此相同，拼接因子相关蛋白 V2afbp-rs 基因在小鼠中印迹而在人类中不印迹。至今还没有发现真兽亚纲以外的动物有基因组印迹现象的报道，但有袋类动物的 X 染色体上有印迹现象，它的失活不是随机的，在胚胎和胚外组织中均为父源 X 染色体优先失活[11]。

2.5 印迹基因编码 RNAs

人们发现很多印迹基因根本不编码蛋白质，但编码 RNAs。H19 基因是第一个被发现的该类基因，在比较人类和小鼠 H19 基因时发现，小鼠和人 H19 基因有保守的一级结构和部分二级结构，但无开放阅读框，因此认为 H19 尽编码 RNA 而不编码蛋白质。在胎儿发育期来自中胚层和内胚层的组织中聚积了大量的 H19 RNA，很难使这种不重要的转录假基因消失。第二个发现的编码 RNA 的印迹基因为小鼠的 Xist 和人的 XIST 基因，定位于 X 染色体的失活中心，该区域为 X 染色体失活所需的顺式作用区，表明这些 RNAs 在印迹形成中有一定作用。

上述两个编码 RNA 的基因具有一些共同特点：这些基因的第一个和最后一个外显子均大，中间含有多个小的外显子；在真兽亚纲动物中两个基因编码的 RNA 一级结构均很保守，虽然那些保守的核苷分布并不完全一致，但在每个基因中其出现具有周期性。根据潜在的基-环结构上互补碱基变化方式推测出 RNA 的二级结构其保守性更加显著。这些特点表明印迹基因编码的 RNA 在进化选择中有一

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

定的生物学功能[14]。

在人类染色体 PW 区发现了第 3 个编码 RNA 的 IPW 印迹基因,其父源等位基因表达,在小鼠中该基因与 Snrpn 基因连锁,也为父本等位基因表达。与 H19 和 Xist 所不同的是,IPW 的基因结构及一级顺序在小鼠和人类之间几乎无前述的保守性,人和小鼠之间仅有一 500 bp 的同源区,在人类为一段外显子区域,在小鼠为外显子与内含子的连接区[15]。

3　基因组印迹的分子机理

现知基因组印迹的分子机理与印迹基因 DNA 中胞嘧啶甲基化尤其是 CpG 岛的甲基化密切相关。在目前发现的印迹基因中几乎所有的印迹基因其某一亲代的等位基因都有一段甲基化序列,称特异性甲基化区域(DMRs)。剔除小鼠甲基转移酶的重要基因 Dnmt1 后,胞嘧啶不能甲基化,小鼠中印迹缺失,表明甲基化在维持印迹中有重要作用[16,17]。

胞嘧啶甲基化是 DNA 的一种共价修饰,是在胞嘧啶 DNA-5-甲基转移酶的作用下甲基从蛋氨酸腺嘌呤 S 位转移至胞嘧啶 C5 位置。在 CpG 二核苷酸处 DNA 甲基化现象是绝对地。DNA 甲基化模式是通过体细胞遗传的,并在甲基转移酶的作用下通过 DNA 复制来维持,甲基化转移酶对半甲基化 DNA(即亲源链甲基化,子代链非甲基化)的亲合力比对无甲基化的 DNA 的亲合力高 100 倍,但是配子期和胚胎生长期 DNA 甲基化富有戏剧性,既有甲基化缺失又可再甲基化,目前仍不知道促使 DNA 甲基化的机理[18]。

研究表明 DNA 甲基化在控制基因组印迹中有一定的作用。首先,小鼠中的印迹基因如 H19 表现为 CpG 岛亲源特异性的、组织非依赖性的甲基化。第二,DNA 甲基转移酶缺乏的基因剔除小鼠有广泛的基因组低甲基化,无 H19 特异等位基因 CpG 岛的甲基化,有 H19 双等位基因表达而无 Igf2 基因表达。母源 Igf2r 基因第一个内含子上的 CpG 岛也有相似的亲源特异性甲基化现象。甲基转移酶缺陷的基因剔除小鼠表现为 Igf2r 甲基化和基因后天沉默[19]。

通过同源重组试验证明小鼠中染色体的特异性甲基化区域(differentially methylated regions, DMRs)在印迹基因表达中具有重要作用。例如,小鼠 H19 基因的父源等位基因是甲基化的,其上游的 DMRs 是 H19 基因和 Igf2 基因印迹表达必需的,而且,将母源染色体上的该区缺失可致母源 Igf2 基因去抑制。根据上述实验事实以及转基因实验证据,目前人们认为母源染色体上非甲基化 H19 基因上游的 DMR 为一染色质界限,它可将 Igf2 与其增强子隔开。另一个已知的小鼠印迹基因为 17 号染色体上的 Igf2r 基因,该基因的母源等位基因表达,在母源染色体上有一个甲基化的内含子 DMR。其他小鼠印迹基因上的 DMRs 也已证明在生殖系中产生并在体细胞中维持,这些 DMRs 包括小鼠 11 号染色体上的 U2af1 - rs1 基因的 DMR,父源等位基因表达的 Snrpn 基因启动子和外显子 I 的 DMR,这两个基因均编码一个 RNA 缝接因子。将 Snrpn 基因 5'端的 DMR 缺失后,可致父源 Snrpn 基因表达降低,这可能是由于雄性生殖系中印迹形成受阻阶段[17,20]。

4　基因组印迹基因的生物学意义

印迹基因在生长发育中尤其是胎儿和胎盘的生长发育中有重要作用,它还与细胞增殖有关,正常基因组印迹模式改变会引起一系列人类遗传性疾病,包括神经和精神发育异常的遗传性疾病以及儿童和成人的一些肿瘤[21,22]。

在基因组印迹研究之初,人们就已证明印迹参与胎儿的生长发育,如有些染色体上的单亲二体(uniparental disomy, UPD)可调节生长发育:父本单亲二体可促进生长,而母本单亲二体则抑制生长。正常胚胎和孤雌胚胎融合而成嵌合体胚胎比正常同窝胚胎小的多,而孤雄嵌合体胚胎比正常胚胎大。研究嵌合体中 UPD 细胞的分布发现,雄性来源的孤雄细胞不对称分布于中胚层组织,如心脏和骨骼肌,但不影响脑组织。相反,孤雌细胞在中胚层组织中未发现,但在脑组织中发现。对第一个印迹基因 Igf2 基因的研究也证明印迹基因参与生长发育。Igf2 基因是胎儿生长的特异性因子,Igf2r 可抑制其活性。H19 基因可抑制 Igf2 和 Ins-2 基因的转录。这些说明基因组印迹对正常生理情况下生长发育有重要的作用[23]。细胞周期调节基因如 GNAS、GFR1 和 $p57^{KIP2}$ 基因为印迹基因,表明基因组印迹参与细胞增殖。

总之,基因组印迹基因在正常发育、细胞增殖及一些人类疾病中扮演着重要角色,是从分子水平弄清基因组印迹机理的突破点,对治疗一些遗传性疾病和肿瘤具有重大意义。

参　考　文　献

[1] Sapienza C, Paquete J, Tran T H et al. Epigenetic and genetic factors affect transgene methylation imprinting. Development, 1989, **107**: 165—168.

[2] De Chiara T M, Robertson E J, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor-2 gene. Cell, 1991, **64**: 849—

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

859.

[3] Barlow D P, Stger R, Herrmann B G et al. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme loucs. Nature, 1991, **349**:84—87.

[4] Falls J G, Pulfor D J, Wylie A A et al. Genomic imprinting: implications for human disease. Am. J. Patho., 1999, **154**:635—647.

[5] Kamiya M, Judson H, Okazaki Y et al. The cell cycle control gene IAC/PLAGL1 is imprinted — a strong candidate gene for transient neonatal diabetes. Hum. Mol. Genet., 2000, **9**:453—460.

[6] Frank D, Mendelsohn C L, Ciaone E et al. A novel pleckstrin homology-related gene family defined by Ipl/Tssc3, TDAG51 and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. Mamm Genome., 1999, **10**:1 150—1 159.

[7] Lee M P, Hu R, Johnson L A. Human KvLQT1 gene shows tissue – specific imprinting and encompasses Beckwrth – Wredemam syndrome chromosomal rearrangements. Nat. Genet., 1997, **15**:181—185.

[8] Maher E R and Reik W. Beckwith-Wiedemann syndrome: imprinting in clusters revisited. J. Clin. Invest., 2000, **105**:247—252.

[9] Wroe S F, Kelsey G, Skimer J et al. An imprinted transcript, antisense to Nesp, adds complexityy to the cluster of imprinted genes at the mouse Gnas locus. Proc. Natl. Acad. Sci. USA, 2000, **97**:3 342—3 346.

[10] Brown C J, Ballaboio A, Rupert J L et al. Agene from the region of the human X chromosome inactive centre is expressed exclusively from the inactive X chromosome. Nature, 1991, **349**:38—44.

[11] Bartolome M S, Tilghman S M. Genomic imprinting in mammals. Annu. Rev. Genet., 1997, **31**:493—525.

[12] Kitsberg D, Selig S, Brandeis M et al. Allele-specific replication timing of imprined gene regions. Nature, 1993, **364**:459—463.

[13] Kawame H, Gartler S M, Hansen R S. Allele-specific replication timing in imprinted domains: absence of asynchrony at several loci. Hum. Mol. Genet., 1995, **4**:2 287—2 293.

[14] Pteifer K, Tilghman S M. Allele-specific gene expression in mammals: the curious case of the imprinted RNAs. Genes. Dev., 1994, **8**:1 867—1 874.

[15] Wevrick R, Kerns J A, Francke U. Identification of a novel paternally expressed gene in the Parader-Willi syndrome region. Hum. Mol Genet., 1997, **6**:325—332.

[16] Razin A, Cedar H. DNA methylation and genomic imprinting. Cell, 1994, **77**:473—476.

[17] Feil R, Khosla S. Genomic imprinting in mammals: an interpaly between chromatin and DNA methylation? TIG, 1999, **15**:431—435

[18] Cedar H, Razin A. DNA methylation and development. Biochem. Biophys. Acta, 1990, **1 049**:1—8.

[19] Stöger R, Kubicka P, Liu C G et al. Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. Cell, 1993, **73**:61—71.

[20] Thorvaldsen J L. Deletion of the H19 differetially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes. Dev, 1998, **12**:3 693—3 702.

[21] Mannens M, Alders M. Genomic imprinting: concept and clinical consequences. Ann. Med., 1999, **31**:4—11.

[22] Tycko B, Ashkenas J. Epigenetics and its role in disease. J. Clin. Invest., 2000, **105**:245—246.

[23] Tilghman S M. The sins of the fathers ang mothers: genomic imprinting in mammalian development. Cell, 1999, **96**:185—193.

GENOMIC IMPRINTED GENE AND ITS BIOLOGICAL SIGNIFICANCE

Zi Xiaoyuan Xiong Jun Hu Yiping

(*Department of Cell Biology, Second Military Medical University, Shanghai* 200433)

Abstract Genomic imprinting is an epigenetic modification of a specific parental allele of a gene, or the chromosome on which is resides, in the gamete or zygote, leading to differential expression of the two alleles of the gene in somatic cells of the offspring. There are now more than 25 identified imprinted genes in human and mice. They have some common features such as clustering of imprinted genes, asynchrony of DNA replication of imprinted genes, temporal and spatial regulation of genomic imprinting, conservation of imprinting and imprinted genes code for RNAs. The molecular mechanisms of genomic imprinting is relative to DNA methylation of the imprinted genes especially CpG dinucleotides methylation closely. Differentially methylated regions (DMRs) are important for the expression of imprinted genes. Imprinted genes are involved in fetal and placental growth and cell proliferation, and the alteration of normal imprinting patterns gives rise to numerous human genetic disease including cancer.

Key words genomic imprinting, imprinted gene, methylation, biological significance

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net