

# 哺乳动物 X 染色体失活机制

徐世永 刘红林\*

(南京农业大学动物科技学院, 南京 210095)

**摘要** 哺乳动物 X 染色体连锁基因的剂量平衡, 是通过雌性胚胎发育早期随机或印记失活一条 X 染色体来实现的, 这是一个复杂的过程, 包括: 启动、计数、选择、维持等一系列的步骤。X 染色体失活中心是 X 染色体失活的主控开关座位, 调节 X 失活的早期事件, 失活发生后, X 染色体的失活状态可稳定地存在并传递给后代, 这一过程涉及基因组印记的形成。此外, 在雄性动物, 精原细胞减数分裂早期也存在着短暂的 X 染色体失活现象。现对哺乳动物 X 染色体失活机制的最新进展进行综述。

**关键词** X 染色体失活; X 染色体失活中心; 基因组印记; 减数分裂性染色体失活

哺乳动物 X 和 Y 染色体是从一对常染色体进化而来, Y 染色体失去大部分的祖先基因, 而在 X 染色体这些基因得以保留, 从而在两性别间造成性连锁基因拷贝数的差异。已经证明胚胎或胚外组织伴性基因表达剂量的失衡是致死性的。哺乳动物采用剂量补偿机制来平衡两性间伴性基因的剂量。这种剂量补偿机制在不同发育阶段不同分化特征的细胞又有所区别, 首先是附植前囊胚期滋养外胚层(trophectoderm, TE)及随后的原始内胚层父源 X 染色体的印记失活, 然后是附植后原肠期胚胎细胞的 X 染色体随机失活。X 染色体失活(X-chromosome inactivation)的现象首先是 Barr(1949)在对雌猫神经核的观察中发现的。随后的研究表明 X 染色体失活是由 X 失活中心(X inactivation center, XIC)顺式控制, 包括一系列复杂的过程, 如: 启动(initiation)、扩展(spreading)、计数(counting)、选择(choice)、维持(maintenance)等。X 染色体失活一旦建立则保持稳定, 其所有子细胞均失活同一条 X 染色体。后来人们还发现雄性动物生殖细胞内也有 X 染色体失活现象, 并与体细胞 X 染色体失活有着不同的机制<sup>[1]</sup>。

## 1 X 染色体失活中心(XIC)

XIC 是 X 染色体失活的主控开关座位, 携带 XIC 的 X 染色体可以顺式启动失活, 相反, 缺失 XIC 的 X 染色体不被失活, 而包含 XIC 的常染色体易位失活可以蔓延到以 XIC 为中心的常染色体其余部分, 表明 XIC 与 X 染色体失活的发生有关。但

XIC 不是维持 X 失活所必需的, 缺失 XIC 后 X 染色体失活状态仍可以稳定地存在。人的 XIC 定位于 Xq13.3, 小鼠定位于 XD, 二者功能上有一定的保守性。XIC 是一个约 1 Mb 的编码序列, 包含一些 X 失活相关单元, 至少包括 4 个已知基因: (1)Xist (X inactive-specific transcript)。长约 17 kb, 至少含 8 个外显子, 产生一个非翻译 Xist RNA<sup>[2]</sup>, 由即将失活的 X 染色体转录, 在活性 X 染色体上其转录受到抑制。该基因上游存在两个启动子 P<sub>1</sub>、P<sub>2</sub>, 用以在即将失活 X 染色体上转录生成稳定聚集的 Xist RNA, 还有一个可能的启动子 P<sub>0</sub> 用以产生不稳定的 Xist RNA。Xist 转录产物是诱导 X 染色体失活的初始信号, 覆盖于即将失活 X 染色体的有限关键位点, 然后招募沉默复合物并向两侧扩展, 引起 X 染色体沉默的启动与传播, 但不涉及计数<sup>[3]</sup>。(2)Xce (X-chromosome controlling element)。位于 Xist 下游, 主要影响失活 X 染色体的选择。Xce 纯合子表现为正常的随机失活, 相反杂合子表现为非随机失活, 说明不同的 Xce 等位基因在决定哪一条 X 染色体失活上似乎有不同的强度。(3)Tsix。起始于 Xist 下游 15 kb 的一个顺式调控元件, 编码长度约 40 kb 的 RNA, 是 Xist 的反义链。X 染色体失活开始前, 两条 X 染色体都表达 Tsix, X 染色体失活一旦开始, Tsix 变为单链表达, 并结合在保持活性的 X 染色体上直到 Xist 关闭。X 染色体失活启动后, Tsix 不

能在失活 X 染色体上存在,表明 Tsix 仅在 X 失活启动时调节 Xist 的活性。定点删除 Tsix 5' 末端可产生 X 染色体非随机失活并影响原有的 X 染色体失活印记,表明这一位点及其转录产物涉及 X 染色体选择与印记<sup>[4]</sup>。(4)DXPas34。DXPas34 座位是一个 3 kb 的 CpG 富集区,包含一个位于 Xist 3' 端下游约 15 kb 的 34-mer 微卫星重复。DXPas34 在体细胞活性 X 染色体上超甲基化。Tsix 转录的重要起始位点位于 DXPas34 内。

XIC 上还有一个可能的调控单元候选区 2.1(2)P,位于 Xist P<sub>1</sub> 启动子上游 120 kb,包含一个 P<sub>0</sub> 启动子,在 X 染色体失活前后表现为不同的乙酰化状态,表明在 X 染色体失活启动之前或期间涉及 Xist 调控。只有细胞携带两个正常 Xist 功能拷贝时 2.1(2)P 区域才会表现超乙酰化,提示其可能与 X 染色体失活的计数有关<sup>[5]</sup>。

## 2 X 染色体失活的过程

### 2.1 计数与选择

X 染色体的失活首先涉及一种计数和选择机制,指导特定条件下哪(些)条 X 染色体失活,对成年细胞二倍体而言失活 X 染色体的数量遵循 N-1 规则(N, X 染色体数),以人细胞为例:46, XY 失活 0 条;46, XX 失活 1 条;47, XXX 失活 2 条;48, XXXX 失活 3 条,从而使二倍体细胞中仅有一条 X 染色体保持活性。X 染色体计数与选择受双向多因子调节并与基因组印记有关(见第 3 节),其研究还有待于进一步深入。长期以来人们知道 XIC 扮演着初始计数和选择的角色,但精细的功能区域定位一直比较困难,尤其是计数区域的定位,直到 1998 年才提供了计数单元的分子证据<sup>[6]</sup>,将其定位于 Xist 3' 端至第 6 外显子之间的 65 kb 区域,核型为 XO 的 ES 细胞缺失该区域后,虽然只有一条 X 染色体但仍能启动失活。2.1(2)P 乙酰化仅出现在 Xist 双拷贝型表明其与 X 染色体的计数也有一定的关联。而且 X 染色体计数似乎识别 Xist 印记。

影响选择的染色质区域则比较复杂,并与计数有明显区别。在小鼠,选择主要由 Xce 控制,其在 XIC 中有 3 个座位。基因突变试验显示 Xist 和 Tsix 也都与选择有关<sup>[7,8]</sup>。最近的研究确认了一种 X 染色体失活选择的候选反式作用因子 CTCF,其结合位点位于 Tsix 内,提示 CTCF 和 Tsix 可以共同建立 Xist 的外源开关<sup>[9]</sup>。

另外,常染色体上某些位点对 X 染色体的选择也有作用,Percec 等<sup>[10]</sup>在小鼠中用化学突变剂去确认涉及 X 失活的特异因子,报道了两个遗传学效应明显不同的常染色体优势突变(Xiaf1、Xiaf2),可影响胚胎形成早期 X 染色体的选择。

### 2.2 启动与扩展

Xist RNA 的聚集可以招募沉默复合物形成扩展失活状态的初始信号,并沿着 X 染色体建立一个可遗传的染色体结构。但不同的失活方式,似乎对 Xist 的依赖程度不同。随机 X 染色体失活启动之前,Xist 表达处于较低水平呈非稳定状态,分化信号刺激,Xist RNA 首先聚集并覆盖在即将失活的 X 染色体上,同时,活性 X 染色体上 Xist 基因失活(反义基因 Tsix 结合造成的)。印记失活的启动则有不同机制,RNA-FISH 分析证实 Xist 早在 8~16 细胞阶段即已稳定覆盖于父源 X 染色体上,此时两个 X 都有活性,给以分化信号刺激,父源染色体失活。由此可见,X 染色体失活与 Xist 在细胞内的聚集有密切联系,分化信号是分水岭,分化信号刺激后 Xist RNA 开始招募启动沉默相关因子,这些相关因子随后又被其他的外源沉默信号所替代和增强,如组蛋白的甲基化(约在分化信号后 1 天产生)、去乙酰化(4 天)、异常组蛋白变体(macroH2A)聚集(7~9 天)、CpG 岛甲基化(这是一个较迟的过程,约在分化信号后 21 天)等,从而形成稳定的失活状态,这一失活状态可以在有丝分裂中得以传递和维持。

然而并非所有的试验都支持以上的观点,Wutz 等<sup>[11]</sup>的研究表明在未分化条件下持续表达外源 Xist cDNA 同样能引起迅速而大范围的转录抑制(可逆);同时,在已分化的细胞,具有失活典型特征 X 染色体的沉默可不依赖于 Xist 的表达。这似乎预示着 Xist RNA 与 DNA 甲基化一样,扮演的应当是定位者的角色,还必须与其他沉默因子协同作用才能诱导 X 染色体沉默。

### 2.3 X 染色体失活状态的维持与复活

体细胞克隆胚胎常常仅能发育到囊胚期阶段,很少能附植发育,即使偶有获得,也常伴有严重的先天性缺陷。这种低成功率部分归咎于成年体细胞很难通过染色体重构,解除 X 染色体的失活状态。失活 X 染色体的结构与正常染色体不同,其分子机制包括:组蛋白 H4 低乙酰化、管家基因 CpG 岛甲基化、组蛋白 macroH2A1 的富集、S 期复制滞后及 Xist 的覆盖等<sup>[12]</sup>。大量证据显示 X 失活状态的维

持具有多种机制, 但各机制对沉默的贡献以及彼此间的相互关系, 还有待于进一步阐明。Xist 在其中扮演的角色颇有争议。一些试验证据显示分化后失活一旦建立, Xist 对于维持失活并非是必需的, 在缺失 Xist 状态下, X 染色体可以保持其失活状态, 然而 Xist RNA 在成熟雌体生命周期中的持续表达, 以及与失活 X 染色体的紧密连接, 预示其应当扮演一定的角色, 试验表明, 缺失正常功能的 Xist, 可扰乱 macroH2A1 在失活染色体中的定位, 从而导致沉默稳定性的降低, 使 X 连锁基因的复活率提高。

DNA 甲基化是维持失活的另一个重要因素, 缺失 CpG 岛甲基化, X 染色体失活是不稳定的。Csankovszki 等<sup>[12]</sup>发现 Xist/Dnmt1(DNA 甲基化酶)双突变, 可以使 30% 的基因恢复活性, 其效果远大于单突变。

组蛋白 H3、H4 氨基末端赖氨酸残基乙酰化的不足是失活 X 染色体的另一个重要特征。细胞经 TSA(组蛋白去乙酰化抑制因子)处理后, 可以发现沉默基因的转录激活。X 染色体的低乙酰化出现在分化及 X 染色体失活以后, 表明组蛋白去乙酰化是 X 染色体失活的维持机制而非建立机制。

DNA 甲基化与组蛋白乙酰化之间存在协同关系, DNA 甲基化作为一种初始机制, 似乎可以通过一种甲基化 DNA 结合蛋白的介导来实现沉默位点的组蛋白去乙酰化<sup>[13]</sup>。Csankovszki 等<sup>[12]</sup>建立一个 Xist 等位基因条件突变, 用 5-azadC 去甲基化和 TSA 乙酰化联合处理, 发现报告基因的复活表达提高 60 倍, 远大于单处理。

5-azadC、TSA 目前已被广泛应用于 X 染色体失活研究中, 借此探讨 X 染色体连锁基因解除失活状态的具体机制。事实上细胞中天然存在着 X 染色体失活逃逸现象, 如在人类 X 染色体中约 15% 的基因可以逃避失活<sup>[14]</sup>, 它们大多定位于 X 染色体的短臂(Xp), 且大部分在 Y 染色体上有同源序列。失活的逃避到底是由于缺乏失活的启动信号, 还是缺乏失活的维持信号, 目前还不清楚。同样在细胞当中也存在着整条失活 X 染色体的复活机制, 这是生殖系细胞传代所必需的<sup>[15]</sup>。已知原始生殖细胞(PGCs)中, X 染色体重新激活、DNA 去甲基化作用及特异性等位基因甲基化印记消失的起始过程是类似的, 似乎存在一种强的反式修饰作用。对这一作用的解释将有助于体细胞克隆工作的顺利进行。

### 3 基因组印记与 X 染色体失活

真兽类动物发育早期, Xist 为“印记”性表达, 致使最初分化的细胞中父源 X 染色体更倾向于失活, 这种倾向性最终导致滋养外胚层、原始内胚层选择性失活父源 X 染色体。X 染色体选择性失活据认为是由父母本基因组不同“印记”造成的, 而这种基因组印记发生在有袋目动物的所有组织当中。Keohane(1996)等据此认为 X 染色体的选择性失活是一种较古老的形式, 真兽类动物则进化形成随机失活机制。Xist 的外源印记标记是在配子发生过程中在 X 染色体上形成的<sup>[16]</sup>, 以一种非活性的表遗传学修饰状态进入合子, 在合子期任一 X 染色体的“印记”标记都可被移除和重新建立。Xist 启动子不同的甲基化状态是印记标记的明显特征, 现已证明卵母细胞中 Xist 因超甲基化而不表达, 母源 Xist 在含有两个雌原核的胚胎中保持甲基化直至囊胚期开始表达。印记重新建立后父源 X 染色体似乎逃避了 X 计数, 允许 Xist 程序性表达。另外有证据显示母源 X 染色体也携带一个稳定的基因印记, 用以抵制失活保持活性: 雌性胚胎中, 携带突变 Xist 的父源 X 染色体不能印记失活, 但删除母源 X 染色体的 Xist 后, 父源 X 染色体又恢复印记失活, 且不影响母源 X 染色体的正常功能。

与随机失活相比, 印记失活有如下不同的特征: (1)失活后 CpG 岛 DNA 甲基化似乎并未发生, 与此相一致, 结构性异染色质的甲基化水平也比正常胚胎细胞低。(2)复制发生于 S 期早期, 没有正常胚胎细胞的滞后复制现象。(3)在印记失活中变体组蛋白 macroH2A 与 X 染色体的结合在 X 染色体失活后即迅速发生<sup>[17]</sup>, 而在随机失活中这一过程则发生的较迟。(4) Xist RNA 覆盖在父源 X 染色体上时间不同, 印迹失活早在 8~16 细胞阶段即已经覆盖在父源 X 染色体上, 这一过程早于细胞分化信号, 同时也提前于印记失活信号约几个细胞分裂周期。而在随机失活中, Xist RNA 直到有细胞分化信号后才开始在即将失活的 X 染色体上聚集。(5)X 染色体印记失活的维持需要 PcG(Polycomb-group)基因的参与, PcG 蛋白复合物涉及失活 X 染色体组蛋白的正确修饰, 而随机失活的维持则不需要。

除以上启动 X 染色体失活的基因组印记外, X 染色体失活一旦发生其他的外源沉默标记也可形成能通过有丝分裂稳定传递的细胞记忆(基因组印记),

成为失活 X 染色体的特征性标记。对人和鼠科 X 连锁基因研究已经发现 5'CpG 岛甲基化与 X 失活状态间的密切相关: X 失活等位基因被甲基化, 相反活性等位基因未甲基化。Jegalian 等(1998)测定了 18 种真兽类动物 X 染色体失活中 X 连锁基因的甲基化状态, 表明 CpG 岛甲基化广泛伴随 X 染色体的失活。

然而真核生物调控的复杂性又必然使其不能完全依赖于 CpG 甲基化这种基于核苷酸结构的固定形式, 还必须寻求一种更为灵活的形式——组蛋白不同的转录后修饰(即“组蛋白密码”假说), 这个概念指出: 决定和维持染色体活性状态是通过调整关键的氨基酸位点, 最明显的是在组蛋白 H3、H4 的 N 末端, 这些修饰: 乙酰化、甲基化、磷酸化, 被认为是通过招募特异蛋白复合物或直接影响染色体结构来实现其功能。该模型的一个重要特点是这些修饰提供了可遗传的标记, 允许基因表达特征通过 DNA 复制和有丝分裂而稳定地传递。研究较多并与 X 染色体失活密切相关的主要是 H3 的甲基化, 关键的位点包括 K4(第 4 位赖氨酸)、K9、K27 等。H3-K9 甲基化是失活 X 染色体最重要的特征之一, 可在有丝分裂过程中保全, 并为失活状态的维持提供外源印记。H3-K9 的甲基化需要在 SUV39h 参与下完成, 这一过程在常染色体基因沉默中也有发生, 然而区别常染色体基因沉默机制的是该过程没有 HP1 蛋白的聚集, 是一种非依赖 SUV39h-HP1 调控通路<sup>[18,19]</sup>。试验表明 G9a 可以甲基化 H3-K9, 同时也可甲基化 H3-K27, 提示 H3-K27 在 X 染色体失活中的可能功能。H3-K27 与 PcG 复合物的协同作用是 X 染色体失活(随机与非随机)启动所必须的(但不充分)<sup>[20]</sup>。H3-K4 甲基化由 SET9 蛋白介导, 其与 H3-K9 甲基化是互相排斥的, H3-K4 去甲基化, H3-K9 的去乙酰化和甲基化发生在相同的时间范围内, H3-K4 甲基化总位于常染色体和 X 染色体活性基因上, 相反 H3-K9 总与失活基因的启动子相结合<sup>[19]</sup>。

#### 4 雄性生殖细胞 X 染色体失活

雄性动物生殖细胞中也存在 X 染色体失活现象, 称为减数分裂性染色体失活(meiotic sex chromosome inactivation, MSCI), 形态学上表现为 XY 染色体联结浓缩形成“性体”(sex body)。性体的产生可以禁止抑制精子发生基因的表达, 或掩蔽性染色体的非联合区, 避免缺陷基因接合导致生殖细胞死亡, 阻止性染色体非接合区重组。

雌性 X 染色体失活是为了雌雄两性间 X 连锁基因的剂量平衡, 而 MSCI 的功能角色, 人们尚无法知晓。MSCI 与体细胞染色体失活似乎有着相似的特征, 如复制滞后、组蛋白 macroH2A1.2 的富集。同时研究表明雄性动物中同样存在 Xist RNA、Tsix 的表达, 并且在胚胎发育早期即已出现, 但二者的表达都具有生殖细胞依赖性<sup>[1]</sup>。由此 Ayoub 等(1997)提出 MSCI 是由 Xist RNA 介导, 经由 XY 配对区域, 由 X 染色体扩散到 Y 染色体最终形成异固缩化 XY。然而 Xist 的定点突变不影响 XY 染色体遭受 MSCI, 也不影响性体的形成<sup>[1]</sup>, 说明 MSCI 是 Xist 非依赖性失活, MSCI 与体细胞 X 失活应当有着根本不同的发生机制。macroH2A1.2 在失活中的不同表现也说明了这一点, 如前所述 Xist 突变干扰 macroH2A1.2 在体细胞失活 X 染色体的正确定位, 但对于“性体”似乎并没有这种干扰。在 macroH2A1.2 出现的时间上, 体细胞中是在染色体异染色质后开始聚集<sup>[21]</sup>, 而在精原细胞中 macroH2A1.2 似乎迅速定位于 XY 染色体, 早于浓缩异染色质化<sup>[1]</sup>。问题是假如 MSCI 是完全不依赖于 Xist, 那么为什么 Xist 和 Tsix 还在睾丸组织中表达呢? 研究中发现睾丸中 Xist 的表达水平与雌性体细胞相比低的多, 而 Tsix 比 Xist 更低<sup>[22]</sup>。一个可能的解释是 Xist 与 Tsix 可能是由于雄性生殖细胞发生程序性去甲基化而产生的非必须转录表达物。

最近人们发现一些蛋白质与 MSCI 的发生有密切关系。如 macroH2A1.2 对于早期粗线期精原细胞“性体”形成有重要作用。Suv39h2(类 Suv39h1), 具有睾丸组织表达特异性, 编码一个 H3 组蛋白甲基化转移酶, 特异性聚集于第一次减数分裂阶段发生转录沉默的“性体”上, 表明 Suv39h2 在形成减数分裂异染色质化甚至标记外源印记上可能有一定的功能<sup>[22]</sup>。M31 是另外一个 MSCI 相关蛋白质, 属 HP1 着色区蛋白。雄性生殖系细胞 M31 定位研究表明, M31 与分化期间染色体改造有关, 涉及精子 DNA 结构的高级调控<sup>[23]</sup>。

正如文中所提到的, 虽然 X 失活研究自发现以来已经经历了半个多世纪的发展, 然而仍有许多相关的不确定问题, 如 X 失活有无其他的失活中心, 彼此间有何关系? 失活的精细过程如何? 以及怎样控制失活与解失活, 这些还都需要在今后的科学研究中不断的去补充完善。

#### 参考文献 (References)

- [1] Turner JMA *et al. J Cell Sci*, 2002, **115**: 4097  
[2] 刘红林. *遗传*, 2000, **22**: 269  
[3] Boumil RM *et al. Hum Mol Genet*, 2001, **10**: 2225  
[4] Lee JT *et al. Nat Genet*, 1999, **21**: 400  
[5] O'Neill LP *et al. EMBO J*, 1999, **18**: 2897  
[6] Clerc P *et al. Nat Genet*, 1998, **19**: 249  
[7] Avner P *et al. Nat Rev Genet*, 2001, **2**: 59  
[8] Lee JT *et al. Cell*, 1999, **99**: 47  
[9] Chao W *et al. Science*, 2002, **295**: 345  
[10] Percec I *et al. Science*, 2002, **296**: 1136  
[11] Wutz A *et al. Mol Cell*, 2000, **5**: 695  
[12] Csankovszki G *et al. J Cell Biol*, 2001, **153**: 773  
[13] Nan X *et al. Nature*, 1998, **393**: 386  
[14] Carrel L *et al. Proc Natl Acad Sci USA*, 1999, **96**: 14440  
[15] 韩 嵘等. *遗传*, 2002, **24**: 77  
[16] Eggen K *et al. Science*, 2000, **290**: 1578  
[17] Costanzi C *et al. Development*, 2000, **127**: 2283  
[18] Peters AHFM *et al. Nat Genet*, 2001, **30**: 77  
[19] Boggs BA *et al. Nat Genet*, 2002, **30**: 73  
[20] Plath K. *Science*, 2003, **300**: 131  
[21] Costanzi C *et al. J Biol Chem*, 2001, **276**: 21776  
[22] O'Carroll D *et al. Mol Cell Biol*, 2000, **20**: 9423  
[23] Hoyer-Fender S *et al. Exp Cell Res*, 2000, **254**: 72

## X Chromosome Inactivation in Mammalian

Shi-Yong Xu, Hong-Lin Liu\*

(College of Animal Science and Technology, Nanjing Agricultural University Nanjing 210095, China)

**Abstract** In mammals, equal dosage of X-linked genes between the sexes is achieved by X chromosome inactivation in a random or imprinted manner in early embryonic development of females. It is a complex process including initiation, counting, choice and maintenance. X inactivation center is the master control locus of the X chromosome inactivation and regulates the early events of X inactivation. Once the X inactivation happened the inactivation state can be stably delivered to generations. This involves genome imprinting. In addition, in males, the X chromosome is inactivated together with Y chromosome in spermatogenic cells shortly before or during early meiotic prophase. This article reviews the recent advances in X chromosome inactivation in mammalian.

**Key words** X chromosome inactivation; X inactivation center; genomic imprinting; meiotic sex chromosome inactivation

Received: December 30, 2003 Accepted: February 22, 2004

\*Corresponding author. Tel: 86-25-84395278, E-mail: ggpc@njau.edu.cn