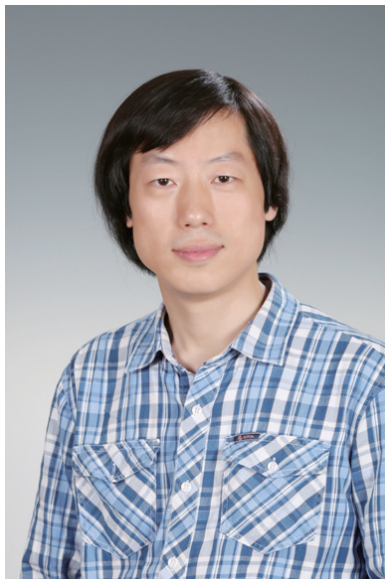


2015年度“中国生命科学领域十大进展”:发育过程中人类原始生殖细胞基因表达网络的表现遗传调控



← 牵头科学家汤富酬, 北京大学生命科学学院研究员, 北大-清华生命科学联合中心研究员。

中国科协生命科学学会联合体: 生殖细胞(精子和卵细胞)是人类生命繁衍的种子和纽带, 在胚胎发育过程中来自原始生殖细胞。人类原始生殖细胞基因表达特征及其表观遗传学调控影响生殖细胞的发育。北京大学汤富酬研究团队与北京大学第三附属医院乔杰研究团队紧密合作, 采用单细胞转录组高通量测序等一系列关键技术, 解析了人类原始生殖细胞多个发育阶段的转录组和DNA甲基化组的动态变化, 揭示了人类原始生殖细胞基因表达调控的一系列关键独特特征。该项研究有助于更好地理解人类生殖细胞和早期胚胎发育的根本规律, 并对未来辅助生殖技术安全性评估、以及临床上生殖、发育异常相关疾病诊断具有重要意义。该研究成果2015年6月在《Cell》上发表。

人类原始生殖细胞的转录组和DNA甲基化组概观

闫丽盈, 郭帆, 汤富酬, 乔杰

北京大学生命科学学院, 北京 100871

1 生命的起始与早期胚胎发育

细胞是人体内结构和功能的基本单位, 在成年人体内大约有40万亿到60万亿个细胞, 包含了200多种不同的细胞类型。然而如此庞大的细胞群体在发育的最早期阶段是由一个单细胞——受精卵发育而来。精子与卵子在输卵管里奇迹般地会合后, 形成一个受精卵, 生命开始了。卵子受精后, 分裂为2个细胞, 大约每隔12 h分裂1次。这团细胞从输卵管进入子宫时, 分泌出液体, 于是膨胀成一个空心球, 叫做胚泡。这个空心球在几天内会变成2层, 球内含有微量液体, 内细胞团堆在球的一侧。球壁以后会变成胎盘和羊膜, 里

面的细胞则会变成胎儿。受精卵靠输卵管的收缩来到子宫, 输卵管内壁的许多纤毛, 不断推动管内的液体, 对输送受精卵也起辅助作用。一般在排卵后4 d左右受精卵到达子宫腔, 约在卵子脱离卵细胞的第9天, 胚胎钻进子宫内膜。到受精后约第10天, 整个胚泡位于子宫内膜中, 第11天在子宫内膜处有小血块和细胞碎片构成的闭锁栓, 第12天胚泡几乎全部被增生的上皮所覆盖, 并形成一个小隆起, 着床即完成。在第2周胚泡植入时, 内细胞团的细胞也增殖分化, 逐渐形成一个圆盘状的胚盘, 此时的胚盘有内、外2个胚层。外胚层为邻近滋养层的一层柱状细胞, 内

胚层是位居胚泡腔侧的一层立方细胞, 2层紧贴在一起。至第3周初, 胚盘外层细胞增殖, 在胚盘外胚层尾侧正中线上形成1条增厚区, 称原条。从第4周初至第8周末的发育过程, 胚胎不仅初具人形, 而且胚盘的3胚层分化发育, 建成各器官系统的雏形; 胎膜和胎盘也于此时期发育形成。胚体形成的同时, 3个胚层也逐渐分化形成各器官的原基。脊索形成后, 诱导其背侧中线的外胚层增厚呈板状, 称神经板。中胚层在脊索两旁从内侧向外侧依次分化为轴旁中胚层、间介中胚层和侧中胚层。在胚体形成的同时, 内胚层卷折形成原始消化管。原始消化管将分化为消化管、

引用格式: 闫丽盈, 郭帆, 汤富酬, 等. 人类原始生殖细胞的转录组和DNA甲基化组概观[J]. 科技导报, 2016, 34(13): 70-76; doi: 10.3981/j.issn.1000-7857.2016.13.011

消化腺、呼吸道和肺的上皮组织,以及中耳、甲状腺、甲状旁腺、胸腺、膀胱和阴道等的上皮组织。

所以从受精卵发育为一个新个体历经复杂的演变过程,包括细胞增殖、死亡、分化、识别、迁移和功能表达,以及组织和器官的形成等。这些变化具有严密规律,具有精细的时间顺序和空间关系。来自同一受精卵的细胞,它们的基因结构是相同的,胚胎发育过程中,细胞基因的表达起决定作用,并同时受内、外环境因素的影响。因而胚胎发育机理成为现代发育生物学中重大的研究课题。千百年来,无数人想破解其中的玄机,但即便到了科技如此发达的今天,这其中仍然有大量发育生物学问题未被完全解决。

2 生殖细胞

生殖细胞又称配子,是人体内起繁殖后代功能的细胞类群,包括精子和卵子,均为单倍体细胞,即仅有23条染色体,其中1条是性染色体。生殖细胞是人类及其他众多多细胞物种得以繁衍和传承的关键种类细胞,对于维持物种的稳定性及延续性起着最为关键的作用。对生殖细胞的形成、发育和维持过程进行深入研究,不但是发育和生殖生物学的核心问题,也是整个生物学的核心问题之一,更是理解人类作为一个物种的延续性、稳定性以及各种遗传疾病的发生机制所必不可少的理论基础。

在胚胎发育的早期阶段,受精卵会在数次分裂之后形成囊胚,其中的一些细胞会形成内细胞团。在囊胚形成的最初几周之内,内细胞团中的干细胞会开始逐渐分化,其中一部分形成原始生殖细胞。原始生殖细胞是两性生殖细胞的“始祖”,它具有分化为卵原细胞或精原细胞的能力,并可进一步形成精子和卵子。精卵结合后继而会发育成新的个体,并将遗传物质传递给下一代以维持种族的延续。生殖细胞(精子和卵子)是人类维持生命延续、代代相传的种子和纽带。

早在妈妈的肚子里,胎儿除了要完成自身发育,形成内脏及四肢等组织器官,还要为自己的后代做好准备,有一类特殊的胚胎细胞将会形成原始生殖细胞(primordial germ cell, PGC)并进行性腺发育,以保证性成熟后形成正常的精子和卵。这类特殊的原始生殖细胞与其他细胞有何不同?基因表达调控的特征是什么?除了遗传序列本身,祖父母或外祖父母及父、母亲还把哪些表观遗传记忆留在了PGC细胞中?哪些表观遗传记忆信息必须要清除?生殖细胞是如何建立自己的表观遗传印记?人类对其还缺乏深刻的认识。因此,对人类早期胚胎以及原始生殖细胞的发育过程进行深入的研究,对于理解人类胚胎发育特征以及对于反复流产、胚胎停育、不孕不育等疾病发病的机制具有重要意义。

3 DNA甲基化

一般认为,同一个生物个体内几乎所有不同类型的细胞都拥有几乎完全相同序列的基因组DNA。然而各种细胞的形态和功能千差万别,这说明同样序列的DNA分子在不同种类的细胞中发挥着不同的作用,指导细胞实现各种不同的功能。人类的基因组含有大约30亿个碱基对,相同序列的DNA分子可以有大量不同的修饰状态,正是这些修饰状态的不同,使得每种类型的细胞,甚至是每个细胞都有自己特定的基因表达特点。

广义来说,表观遗传学修饰不只是包括DNA碱基上的修饰和组蛋白的修饰,还包括DNA折叠组装成核小体的位置和动态,染色体上转录因子以及其他染色体结合蛋白的分布和动态,以及染色体的折叠,常染色质、异染色质的分布,甚至还包括RNA的各种共价修饰等等。这些不同的修饰以及修饰的组合,使得不同种类的细胞,或者同种细胞类型不同细胞间产生异质性。这种细胞类型特异的基因组修饰状态,是表观遗传学领域的重要研究方向。活细胞的神奇之处还在于它们在细胞分

裂过程中能准确“记住”各种修饰信息,然后在新生的子细胞中精确地加上这些修饰。如果其中某些修饰不能重新从母细胞中被准确完整地复制下来,新生子细胞的基因组有可能因此演变为肿瘤细胞或者诱发细胞凋亡。目前国内大量研究组在这些领域取得新的突破。

DNA甲基化是在DNA甲基转移酶的作用下将甲基选择性地添加到胞嘧啶上形成5-胞嘧啶的过程,刚被发现时被定义为第5种碱基,实际上它是一种重要的表观遗传学标记,在调控基因表达、维持染色质结构、基因印记、X染色体失活以及胚胎发育等生物学过程中发挥着重大的作用。它就像DNA的一项最神奇的“帽子”,带给世人层出不穷的惊喜。

在真核生物基因组中,基因仅占一小部分,例如在人类基因组中基因的编码序列还不到2%,那么在大量非编码DNA存在的情况下,实现精确控制基因的表达,降低周围的转录噪音对生物体至关重要。DNA甲基化作为一种可遗传的修饰方式,非编码DNA(内含子、重复元件以及潜在的具有活性的转座子)的长期沉默提供了一种有效的抑制机制。DNA复制后胞嘧啶的甲基化会改变DNA的构象,使DNA的大沟无法与DNA结合蛋白正常结合,从而使这些非编码区长期保持无表达活性的状态。而有转录活性的基因可利用非甲基化的启动子进行转录表达,即使在相邻的非转录区是高度甲基化的,其启动子仍然可以起始转录并被调控。

过去人们一直以为遗传和环境两大因素共同决定生物体的性状,然而人们无法合理解释马和驴的正反交后代、同卵双胞胎差别以及X染色体失活等现象。1942年,Waddington首次提出了表观遗传学(epigenetics)的概念,它针对研究基因型与表型的关系,使经典的孟德尔核内遗传规律无法解释的现象得到了合理而完美的解释。基因组表观遗传修饰主要包括DNA甲基化修饰与核小体中组蛋白的修饰等,使得被修

饰 DNA 的空间结构发生改变或使染色体结构发生改变,导致基因的沉默或过度表达。这 2 种修饰都是在 不改变 DNA 碱基种类与数量的前提下使生物体表型呈现出多样化。DNA 甲基化对基因表达模式以及基因组稳定性均起着至关重要的作用,并在印迹基因与 X 染色体失活等典型的表观遗传现象中起重要作用(图 1)。

DNA 甲基化作为一种可遗传的表观遗传修饰,在体细胞增殖过程中通过依赖于 DNA 复制的 DNA 甲基转移酶 Dnmt1 稳定地传递给子细胞。但在胚胎发育的不同时期,基因组范围内的 DNA 甲基化水平会发生剧烈的改变,改变最剧烈的阶段为配子形成期与早期胚胎发育阶段,甲基化模式在配子形成时已经建立。DNA 甲基化对胚胎正常发育和等位基因的选择表达至关重要。错误甲基化模式的建立将引起人类的疾病,如 Prader-Willi 综合征、Angelman 综合征和脆性 X 染色体综合征等。

4 DNA 去甲基化

细胞的 DNA 甲基化状态虽然比其他的表观遗传修饰更稳定,但并不是一成不变的。DNA 甲基化修饰,在某些特定的生理条件下能被细胞“聪明”地

擦去,DNA 去甲基化的机制是近些年来表观遗传学的一大热点,现在普遍认为主要存在两种去甲基化途径:主动去甲基化和被动去甲基化,而这两种途径中只有主动去甲基化需要羟甲基胞嘧啶的参与。羟甲基胞嘧啶是胞嘧啶的另外一种修饰状态,是指胞嘧啶的嘧啶环第 5 号位置上发生了羟甲基化修饰,现在普遍接受的观点是羟甲基化的修饰在很大程度上和甲基化修饰的性质不同,是去甲基化修饰的最主要途径,也即去甲基化的中间产物。而且羟甲基化发生的区域比较稀少,因此其程度远远低于甲基化修饰,这也就是为什么直到最近几年随着检测灵敏度的提高才被发现、被重视。

另外一条 DNA 去甲基化的途径就是被动去甲基化,它和上面提到的主动去甲基化的原理完全不同。被动去甲基化主要是依赖于细胞分裂,DNA 复制而发生的。随着 DNA 的复制,新生的子链上不会按原来母链上甲基化修饰的胞嘧啶(CpG)对应的位置上重新加上甲基,从而被被动地稀释掉。

哺乳动物一生中 DNA 甲基化水平经历 2 次显著擦除,第一次发生在受精卵最初几次卵裂过程中,去甲基化酶清除了 DNA 分子上几乎所有从亲代遗传来的甲基化标志;而在胚胎植入子宫

时,新加的甲基化遍布整个基因组,甲基化酶使 DNA 重新建立一个全新的甲基化模式。细胞内新的甲基化模式一旦建成,即可通过甲基化以“甲基化维持”的形式将新的 DNA 甲基化传递给所有子细胞 DNA 分子。第二次发生在原始生殖细胞迁移的过程中,DNA 甲基化被大规模地擦除,在配子成熟和分化的过程中 DNA 甲基化谱式又会被重新建立。

相对于基因组范围内的大规模主动去甲基化,在体细胞中会发生局部的、高度位点特异性的主动去甲基化。DNA 的去甲基化与 DNA 甲基化这两个过程相互平衡,维持了 DNA 甲基化谱式的稳定。任何一方的失调都会导致 DNA 甲基化谱式的紊乱,进而引起多种神经退行性疾病、免疫系统疾病以及癌症。特别是两次基因组范围的主动去甲基化事件对于生命的起始以及生命的传承具有非常重要的意义。

5 哺乳动物胚胎发育 DNA 甲基化水平动态变化

在小鼠为代表的哺乳动物的一生当中,DNA 甲基化的状态长时间内都基本保持不变,这得益于活细胞体内有一套非常严密的 DNA 甲基化的调控机制。但在受精前后,着床前后,原始生殖细胞(PGC)的发生、迁移,配子的生成等重要发育过程中,DNA 甲基化会有剧烈的重编程发生。现在的研究在小鼠体内已经比较清楚(图 2^[11])。

成熟的雌、雄配子在受精前 DNA 甲基化状态是不同的,其中精子的甲基化程度远高于成熟的卵母细胞。单个精子与卵母细胞结合完成受精后,随着雌雄原核出现并且逐渐靠近到融合之前,父方的基因组和母方的基因组主要以不同的方式进行去甲基化。当雌雄原核靠近并且融合之后,此时的受精卵(早已经开始了 DNA 复制)的 DNA 甲基化程度大大降低,远低于受精之前的卵母细胞和精子的 DNA 甲基化水平。尤其是在重复序列上,DNA 去甲基化更加明显。等到融合的合子再分裂成二

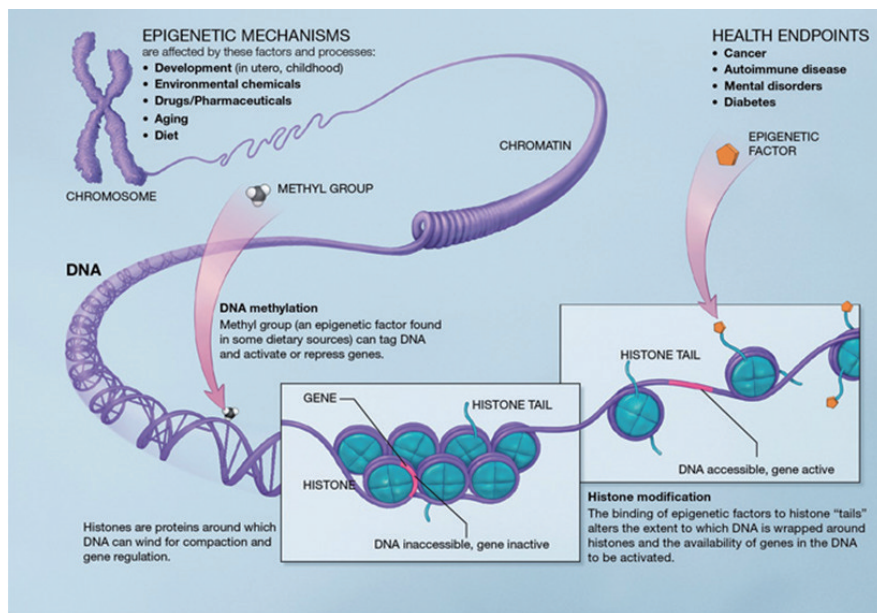


图 1 表观遗传学及其主要修饰(图片来源:Wikipedia)

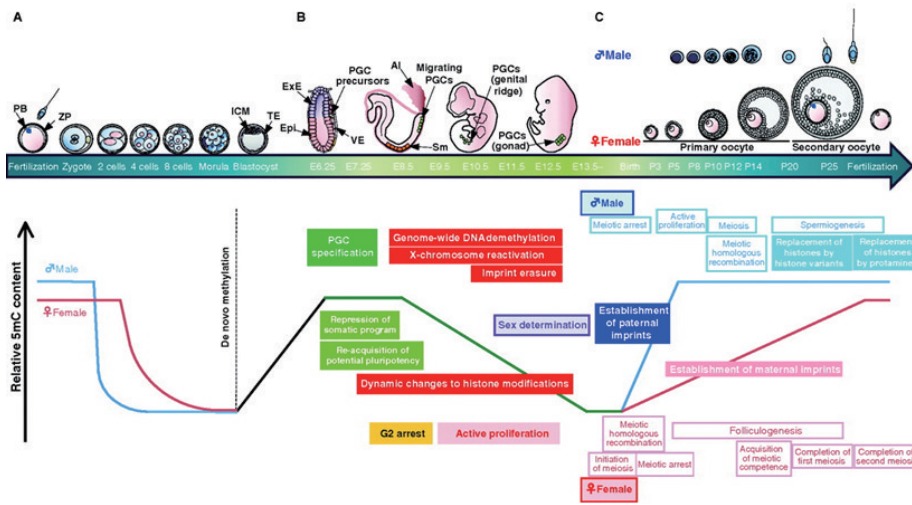


图2 DNA甲基化在小鼠胚胎发育过程中的动态变化

细胞的时候,其DNA甲基化水平进一步降低。受精到分裂成二细胞时间很短,细胞需要在这个较短的时间内迅速但是精确地完成上述重编程过程,才能顺利发育成多细胞的胚胎。

虽然说受精卵能发育到二细胞已经顺利度过了一个难关,二细胞期胚胎的DNA甲基化水平降到了很低水平,但并不意味着DNA甲基化在之后的发育过程中就不重要了。在胚胎早期卵裂期间,DNA的甲基化状态也至关重要。虽然甲基化的水平在卵裂期间很低,而且一直都维持在很低的水平,但是细胞分裂伴随着DNA复制一直都在进行,从二细胞到八细胞阶段,甚至到桑葚胚阶段,DNA甲基化的水平基本都维持在很低的水平,到了囊胚阶段,尤其是晚期囊胚的内细胞团(ICM)阶段,DNA甲基化水平下降到最低值,而这个阶段的胚胎已经经过很多次细胞分裂,已经从1个细胞变为上百个细胞。需要指出的是,囊胚细胞的DNA甲基化水平很低,但是这两种细胞的DNA甲基化状态略微不同,滋养外胚层细胞的DNA甲基化水平略低于内部的内细胞团,推测可能是由于滋养外胚层细胞中DNMT家族的酶的丰度低于内细胞团的。

在着床之后,体细胞的基因组甲基化程度迅速增加,伴随着3个胚层的分

化,在胚胎内部特定的部位出现原始生殖细胞。首先形成的是生殖细胞的前体细胞即原始生殖细胞,并且伴随着胚胎发育,原始生殖细胞增殖并迁移,最终进入到生殖嵴,与生殖嵴内的中胚层细胞一起形成原始的性腺(睾丸或者卵巢),与此同时也伴随着性别决定。此时的原始生殖细胞呈碱性磷酸酶阳性,多能性基因高表达(如*Sox2*,*Oct4*等),体细胞基因低表达(如*Hox*家族的基因)。得益于特异性很高的转基因标记的小鼠,现在能比较容易地分离纯化到不同发育阶段的原始生殖细胞,方便我们做各种研究。因此,目前关于小鼠原始生殖细胞各个发育阶段的基因表达或者是DNA甲基化变化的研究有较多报道。小鼠中,在受精之后的第6.5天(E6.5)就开始出现了原始生殖细胞,这个阶段的PGC的DNA甲基化水平和它周围的体细胞的水平比较一致,当原始生殖细胞大量增殖并且开始迁移,其DNA甲基化水平开始降低,到E11.5天时降到最低水平。

6 人类原始生殖细胞基因组甲基化概貌

人类的发育过程是由胚胎自身携带的遗传信息和外界环境协同调控各种基因的表达来实现的。对基因表达的调控主要有遗传学调控和表观遗传

学调控两种方式。表观遗传学调控是整合遗传信息与环境因素的主要方式。而表观遗传学调控主要是通过染色体中基因组DNA的甲基化修饰以及组蛋白的共价修饰来实现的。其中基因组DNA的甲基化是最早发现的,是最重要的表观遗传修饰方式之一,也是调控细胞分化过程中基因表达的主要机制之一。它虽然不改变基因序列但可以遗传给后代,而且容易受外界环境的影响而发生改变,在胚胎发育、干细胞分化、癌症的发生等方面发挥着重要作用。在胚胎的大脑、心脏、肝脏等体细胞的发育过程中DNA甲基化可以是单向改变的,因为这些体细胞来源的器官只在1个世代的个体中存在,并不需要传递到下一代。但是在胚胎的种系细胞、也就是生殖系细胞中,DNA甲基化的改变必须是可逆的,就是在上一代的精子和卵细胞中先形成生殖细胞特异的DNA甲基化状态之后,在下一代的胚胎发育早期必须擦掉这些生殖细胞特异的DNA甲基化“记忆”,之后在胚胎的大脑、心脏、肝脏等体细胞中从头建立体细胞特异的DNA甲基化状态,同时在胚胎的生殖系细胞中重建生殖细胞特异的DNA甲基化状态。简单来说,就是在上一代个体的生殖细胞到下一代个体的生殖细胞的发育过程中,DNA甲基化变化要形成“环路”,回到“原点”,必须依次经历DNA甲基化的“编程”和“重编程”过程。实际的哺乳动物胚胎发育中DNA甲基化状态变化要经过两轮“编程”—“重编程”的过程才能完成,以保证“正确”的DNA甲基化记忆从上一代“传递”到下一代。

在精子和卵细胞结合形成受精卵后,精子和卵细胞中生殖细胞特异的DNA甲基化会被擦除(只保留印记基因的甲基化状态),然后在胚胎着床后整个胚胎中从头建立体细胞(大脑、心脏、肝脏等)特异的DNA甲基化状态,完成第一轮“编程”—“重编程”过程。之后在胚胎中的生殖前体细胞(原始生殖细胞)中会把体细胞特异的DNA甲基化擦除,并重建生殖细胞特异的

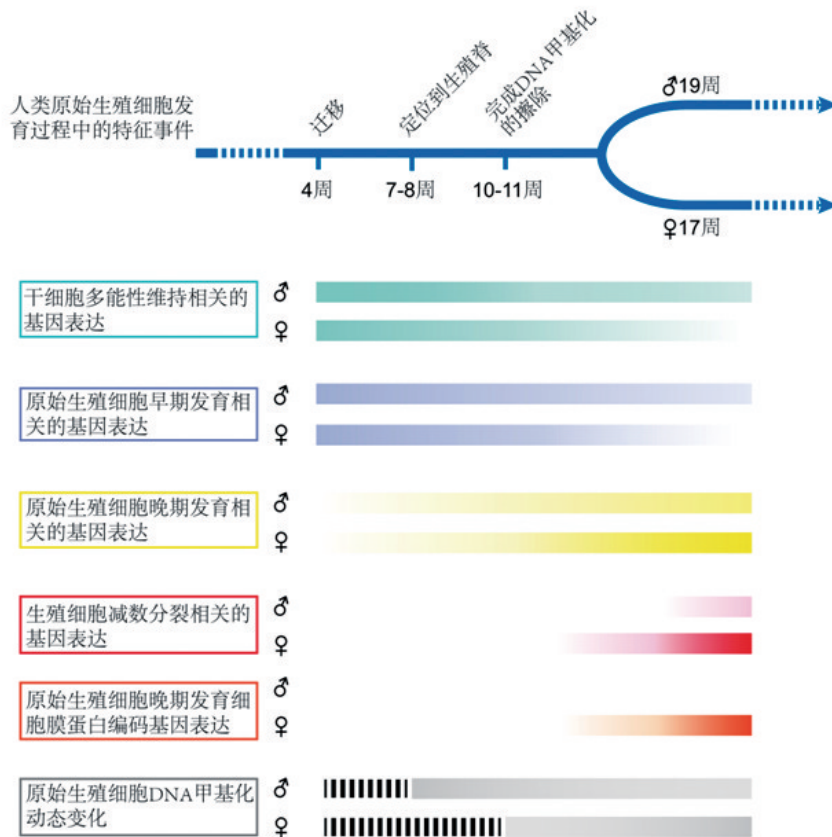


图3 人类原始生殖细胞代表性基因的表达水平以及DNA甲基化随发育时间变化示意

DNA甲基化状态,完成第二轮“编程”——“重编程”过程。

北京大学研究团队首先对人类着床前胚胎的DNA甲基化图谱进行了单碱基分辨率的、全基因组测序分析,发现人类精卵结合后,早期胚胎的基因组DNA发生大规模的去甲基化,甲基化程度从精子中的86%(中位数)降低到囊胚期胚胎中的43%,而在这一过程中印记基因控制区域的DNA甲基化得以精确的维持。在着床后的胚胎中,基因组DNA被大规模重新甲基化,DNA甲基化水平增加到92%(图3)。该项研究还发现了DNA甲基化组重编程对于基因表达网络的关键调控特征。相关研究成果于2014年7月发表在《Nature》上^[2]。该项研究为人们提供了一个深入分析人类早期胚胎发育过程中基因表达网络表观遗传调控的坐标系。

在此基础上,该研究团队对人类胚胎中第二轮DNA甲基化组重编程过程及其对基因表达网络的调控关系进行了深入、全面的分析。正常情况下,一个着床后胚胎内大部分组织器官的基因组DNA加上甲基化标记后就基本维持稳定、不再擦去,而含有要传递给后代遗传信息的原始生殖细胞则要经历大规模的DNA甲基化擦除和重建的过程。该团队利用精细的流式细胞分选技术对不同发育阶段以及不同性别的人类原始生殖细胞进行了分选、收集,再结合单细胞转录组高通量测序、微量细胞DNA甲基化组高通量测序、微量细胞DNA羟甲基化组高通量测序等技术以及细胞免疫荧光技术,对人类原始生殖细胞的转录组、DNA甲基化组以及组蛋白共价修饰等关键特征进行了深入的分析^[3]。

该项研究发现,处于发育早期阶段

的人类原始生殖细胞也会表达一系列与干细胞多能性相关的基因(例如*OCT4*、*NANOG*和*REX1*等),印证了用小鼠作为模式动物研究原始生殖细胞发育过程的重要性。另一方面,与小鼠不同的是,人类原始生殖细胞并不表达关键的转录因子基因*SOX2*,取而代之的是表达另外2个*SOX*家族基因——*SOX15*和*SOX17*。同时,人类原始生殖细胞还表达一系列与生殖细胞发育相关的基因(例如*KIT*、*TNAP*、*AP2γ*和*NANOS3*等)。与早期处于有丝分裂阶段的细胞相比,进入减数分裂阶段的人类原始生殖细胞其转录组发生明显改变,同时细胞之间转录组的差异也大大增强,提示原始生殖细胞在进入减数分裂停滞期时同一个胚胎的不同生殖细胞处于显著不同的发育状态。其次,在人类着床后雌性胚胎中,每个细胞的两条X染色体中的1条会随机失活,以保持两性的X染色体剂量相同(两性都保持每个细胞中只有1条活跃转录的X染色体)。而在人类雌性胚胎的生殖细胞中失活的那条X染色体会被重新激活,该团队发现这一过程在发育第4周的雌性原始生殖细胞中就已经完成了,明显早于小鼠原始生殖细胞中X染色体的重新激活。

前述植入前胚胎的研究结果显示,在受精之后的第一轮去甲基化过程中,基因印记区域的甲基化能得以精确维持。在生殖细胞如此彻底的去甲基化过程中,基因印记区域的甲基化该如何变化?我们通过文献报导,将已知的印记基因分为了母源印记基因(maternal imprinting DMRs)、父源印记基因(paternal imprinting DMRs)、体细胞特异的印记基因(somatic DMRs)以及胎盘特异的印记基因(placental imprinting DMRs)。我们发现母源和父源印记基因在着床后胚胎的大脑和心脏以及对应的体细胞中,是基本能维持50%左右的甲基化水平的;在雌性或者雄性PGC细胞中,基本都擦除到了很低水平,甚至到了0%。而体细胞特异的印记区域以及胎盘特异的印记区域,

在着床后胚胎的大脑和心脏组织以及性腺的体细胞中,有些基因区域能维持50%左右甲基化水平,一部分的区域是超过90%以上的高甲基化,而另一部分区域是基本未甲基化,所有这些体细胞特异的印记区域以及胎盘特异的印记区域在雌性和雄性PGC细胞中,都是很低水平的甲基化。综上,所有的印记基因区域,在生殖细胞中都完全被擦除,来自父母本的印记信息不能精确维持并传递到后代中。同样的,我们观察了一些生殖细胞特异性表达的基因,这些基因只在PGC细胞中有高表达,而在体细胞和大脑、心脏等组织中基本不表达,而这些基因的启动子在大脑心脏和性腺体细胞中几乎为100%甲基化水平,而在PGC细胞中几乎不被甲基化。同时,我们也进一步观察了在不同的基因组元件上,DNA甲基化的各种变化情况。

除了印记基因区域和生殖细胞特异性基因的启动子区域在生殖细胞中甲基化程度很低外,我们想进一步观察是否某些特定区域的DNA甲基化有残留并且会遗传给后代,因此对不同基因组区域进行细分和比较,发现虽然整体DNA甲基化水平在PGC细胞中都降到很低,但在某些重复元件上,例如SINE(short interspersed nuclear element),LINE(long interspersed nuclear element)和LTR(long terminal repeat)等大类的重复序列上,就残留有一定程度的甲基化;而在将大类的细分为各种亚类后,我们发现在ERV(endogenous retroviruse)和ALR(alpha satellite repeat)上残留最多甲基化,ERV中尤其以ERVK(endogenous retroviruse K)最多。

在生殖细胞中DNA甲基化与基因表达的关系如何呢?我们进一步将甲基化分为启动子甲基化与基因区域甲基化,启动子甲基化与基因表达呈现明显负相关,即便是整体DNA甲基化下降到了10%以下后,这种负相关关系始终存在,这种负相关关系在之前的着床前胚胎中也被发现。另一方面,我们将基因根据表达分为高表达基因、中等表

达量基因以及低表达基因3类,发现高表达量的基因其基因区域的甲基化程度反而高于中等表达量和低表达量的基因,而且在不同阶段的细胞中基本是一致的,这也就说明基因区域的甲基化与基因表达是一个较为微弱的正相关关系,而这种正相关关系其实也在之前着床前的胚胎中得到过证明。

提到DNA甲基化与基因表达的关系,我们自然也会联想到之前提到过的DNA甲基化与重复序列表达的关系。在着床前的胚胎中,我们观察到了进化上不同时间的重复序列的DNA甲基化与其表达的关系,同样,我们根据进化时间的远近,分别计算了进化上年轻的L1和Alu,以及进化上比较古老的L2和MIR重复序列在着床后胚胎的心脏和大脑组织以及生殖细胞中的DNA甲基化程度以及对应的表达情况,发现L1和L2同属于LINE家族,Alu和MIR同属于SINE家族,L1比L2进化上更为年轻,对应的拥有更高的转录水平,而在第二轮大规模去甲基化过程中,L1在生殖细胞中倾向于残留有更高程度的甲基化,以防止其过度活跃而损害DNA的稳定性;同理Alu比MIR更为年轻,也有更高的表达水平,同时相比于MIR,Alu在PGC细胞中残留更高程度的DNA甲基化。

我们在全基因组范围内检测到了两轮大规模的DNA去甲基化,这两次大规模的甲基化相比较,我们发现:1)在第二轮去甲基化过程中,也就是在生

殖细胞中,甲基化擦除地更为彻底和充分。第一轮去甲基化到囊胚阶段还残留有接近40%左右的全基因组甲基化水平,但是在10~11周的PGC细胞中,全基因组甲基化水平只有7%左右,绝大部分基因,包括印记基因区域的甲基化也都被擦除,只在某些重复序列上残留有一定程度的甲基化水平。胚胎非常机智地在第二轮去甲基化过程中擦除原来的甲基化记忆,之后在晚期生殖细胞变成精子和卵母细胞的过程中会重建DNA甲基化,这样确保成熟的配子不会携带父母本的遗传记忆到下一代中,遗传给下一代的只是自己重新建立的甲基化信息。胚胎在对待自己的生殖细胞时是非常“严格”和“负责”的,以确保父母本的遗传信息(可能有些是环境诱导形成的有害突变)不会传递到子代中,这样为物种的稳定延续提供了一定的遗传学基础。PGC中在10~11周时甲基化擦除很低,但是基因表达仍然能正常维持,推测可能有其他表观遗传学因素存在,例如组蛋白修饰和piRNA来调控基因表达。2)甲基化对于重复元件的抑制作用始终存在,不管是在着床前胚胎和着床后的PGC细胞中,虽然某些阶段整体甲基化程度降到很低,但是这种抑制作用是始终存在的。而且我们在着床前胚胎和着床后PGC细胞中,都发现了年轻的重复序列在去甲基化过程中倾向于残留有更高程度的甲基化水平,以防止其过度活跃的转座作用而影响胚胎的发育。但是

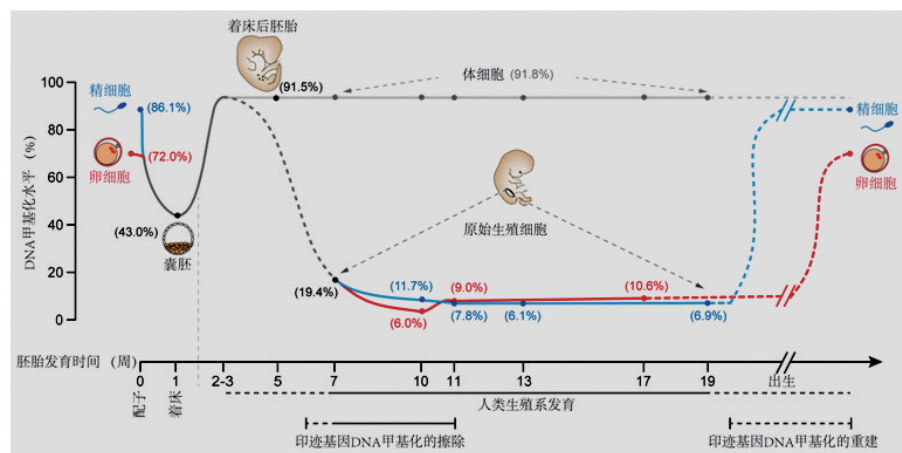


图4 人类早期胚胎与原始生殖细胞甲基化擦除与建立特征

细胞并没有完全将重复序列用 DNA 甲基化锁起来。实际上我们在细胞中均能检测到一定量重复序列 RNA 的表达。这一方面可能产生有害的基因转座和序列改变,另一方面也可能产生一些有益的基因转座,使得物种在稳定的延续过程中也能出现一些基因组序列的变化,有些有益的突变使得个体更适合生存而被保留下来,而那些有害的突变在强大的生存压力下被逐步淘汰,这也为个体之间遗传变异提供了一定的理论基础。

综上,我们通过对人类着床前早期胚胎,以及着床后生殖系细胞全基因组范围甲基化的检测,发现了人类胚胎发育过程中 2 次大规模的 DNA 甲基化重编程过程的关键特征(图 4),特别是很多不同于小鼠胚胎发育的独特特点,系统地阐释了 DNA 甲基化对人类早期胚胎以及生殖系细胞发育过程的调控机理。

该项工作首次分别在单细胞以及单碱基分辨率对人类原始生殖细胞的转录调控网络和 DNA 甲基化重编程过程进行了深入、系统的分析,加深了对人类原始生殖细胞的发育以及表观遗传重编程过程的认识。为人类生殖细胞的表观遗传重编程、早期胚胎全能性的建立、DNA 甲基化的隔代遗传以及胚胎干细胞向精卵定向分化等科学问题的探究提供了理论基础。尤其对辅助生殖技术的安全性评估、疾病对后代影响的评价以及临床上生殖细胞发育异常相关疾病的研究具有非常重要的社会意义。

《Cell》杂志同期刊发美国与台湾的合作研究团队、英国剑桥大学团队关于人类原始生殖细胞发育过程的基因表达及 DNA 甲基化表观遗传特征的研究结果^[4,5]。在同一期同时发表 3 个来自不同国家的团队独立开展的研究,充分体现了该研究在科学领域的重要

性。这 3 个研究,内容与结果略有不同,但又相互验证,共同为人类认识自身配子发生的特征做出了突出贡献。在这 3 项研究中,北京大学团队的研究结果是发育阶段覆盖最全面、分辨率最高、数据最精确、分析最深入的。比如北京大学团队不但对原始生殖细胞本身的转录组进行了高精度的分析,而且对这些生殖细胞的微环境细胞也进行了一对一匹配的全方位分析,为胚胎干细胞体外定向分化为生殖细胞的体外微环境方面的探索提供了重要的线索。国际表观遗传学权威学者 Reik 和 Meyenn 在同期《Cell》发表述评,称这 3 项研究为人类提供了详尽的人类原始生殖细胞发育过程中基因表达网络及其表观遗传调控景观图,这是理解基因组潜能重置、表观遗传记忆擦除、人类生殖细胞建立的基础。这方面的知识有助于更好地理解人类生殖细胞和早期胚胎发育特征与规律。

参考文献

- [1] Saitou M, Kagiwada S, Kurimoto K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells[J]. *Development*, 2012, 139(1): 15-31.
- [2] Guo H, Zhu P, Yan L, et al. The DNA methylation landscapes of human early embryos[J]. *Nature*, 2014, 511(7511): 606-610.
- [3] Guo F, Yan L, Guo H, et al. The transcriptome and DNA methylome landscapes of human primordial germ cells[J]. *Cell*, 2015, 161(6): 1437-1452.
- [4] Tang W W, Dietmann S, Irie N, et al. A unique gene regulatory network resets the human germline epigenome for development[J]. *Cell*, 2015, 161(6): 1453-1467.
- [5] Gkoutela S, Zhang K X, Shafiq T A, et al. DNA demethylation dynamics in the human prenatal germline[J]. *Cell*, 2015, 161(6): 1425-1436.

(责任编辑 王媛媛)