



(12) 发明专利申请

(10) 申请公布号 CN 112272516 A

(43) 申请公布日 2021.01.26

(21) 申请号 201980037877.9

(22) 申请日 2019.04.05

(30) 优先权数据

62/654,199 2018.04.06 US

(85) PCT国际申请进入国家阶段日

2020.12.04

(86) PCT国际申请的申请数据

PCT/US2019/026074 2019.04.05

(87) PCT国际申请的公布数据

W02019/195738 EN 2019.10.10

(71) 申请人 儿童医疗中心有限公司

地址 美国马萨诸塞州

(72) 发明人 Y·张 S·马托巴

(74) 专利代理机构 北京戈程知识产权代理有限公司 11314

代理人 程伟

(51) Int.Cl.

A01K 67/027 (2006.01)

C12N 9/02 (2006.01)

C12N 15/113 (2006.01)

C12N 15/85 (2006.01)

C12N 15/873 (2006.01)

C12N 15/877 (2006.01)

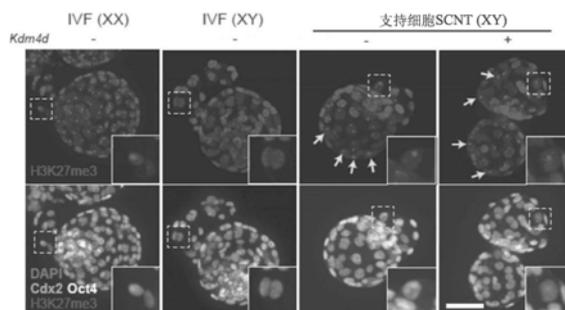
权利要求书1页 说明书67页 附图25页

(54) 发明名称

用于体细胞重新编程和调整印记的组合物和方法

(57) 摘要

本发明提供用于改善克隆效率和调整印记控制区域的方法。在具体实施方案中,本发明提供用于激活印记控制区域中被阻遏的等位基因,从而治疗印记相关病变的方法。在其它实施方案中,本发明提供用于改善体细胞核转移效率的方法,其牵涉Xist敲除的供体细胞中的Kdm4d过度表达。



1. 一种获得克隆囊胚的方法,所述方法包括将从缺乏Xist活性的体细胞获得的供体细胞核转移至去核卵母细胞内,并且在所述卵母细胞内表达Kdm4d,从而获得克隆囊胚。
2. 根据权利要求1所述的方法,其中向所述卵母细胞注射Kdm4d mRNA。
3. 根据权利要求1所述的方法,其中所述供体细胞核是从包含Xist删除或包含无活性形式的Xist的胚胎成纤维细胞获得的。
4. 根据权利要求1所述的方法,其中所述供体细胞核是从人类、猫、牛、狗、猪或马获得的。
5. 根据权利要求1所述的方法,还包括将所述囊胚转移至宿主子宫内进行妊娠。
6. 根据权利要求5所述的方法,其中,相对于传统的体细胞核转移,所述方法将活产率增加了至少约10%至20%。
7. 一种获得以供移植入受试者体内的细胞或组织的方法,所述方法包括:
 - (a) 在从受试者获得的培养细胞中令Xist失活或降低Xist活性或表达;
 - (b) 将来自所述培养细胞的细胞核转移至去核卵母细胞内,从而激活所述卵母细胞;以及
 - (c) 向在步骤(b)中获得的所述激活卵母细胞注射Kdm4d mRNA并且培养所得细胞,从而获得适用于移植入所述受试者体内的细胞或组织。
8. 根据权利要求7所述的方法,其中通过基因组编辑令Xist失活。
9. 根据权利要求7所述的方法,其中使用CRISPR系统将删除或失活突变引入基因组Xist多核苷酸中。
10. 根据权利要求7所述的方法,其中使用siRNA或shRNA来降低Xist多核苷酸表达或活性。
11. 一种根据权利要求1所述的方法生产的囊胚。
12. 一种细胞,其包含Xist删除或具有降低水平的Xist表达并且包含编码Kdm4d的异源多核苷酸。
13. 一种根据权利要求7所述的方法生产的细胞或组织。
14. 一种通过将根据权利要求1所述的囊胚植入宿主子宫内生产的克隆有机体。
15. 一种卵母细胞,其包含从缺乏Xist活性的体细胞获得的供体细胞核并且包含含量相对于传统卵母细胞增加的Kdm4d。

用于体细胞重新编程和调整印记的组合物和方法

[0001] 相关申请的交叉引用

[0002] 本申请主张2018年4月6日提交的下述美国临时申请62/654,199的权益,该临时申请的整体内容通过引用并入本文。

[0003] 联邦政府资助的研究成果的发明权声明

[0004] 本发明在政府支持下完成,由美国国立卫生研究院资助(项目编号:HD092465)。政府享有本发明的某些权利。

背景技术

[0005] 哺乳动物卵母细胞能够通过体细胞核转移(SCNT)将体细胞重新编程为全能状态。SCNT用于治疗性克隆中,其牵涉从基因上等同于或类似于所预期宿主的供体有机体生成组织。SCNT也令动物克隆成为可能。这一技术在农业生物技术以及濒危物种保护中具有很大潜力。然而,极低的克隆成功率使得这一技术难以实际使用。以小鼠为例,仅约30%的SCNT胚胎发育成囊胚,并且仅1%至2%的转移至代孕母亲的胚胎能达到足月。此外,在存活的胚胎中,几乎在所有克隆的哺乳动物物种的胚胎外组织诸如胎盘和脐带中经常观察到异常。这些观察结果暗示,SCNT重新编程有些缺陷,妨碍了胚胎发育过程。已经表明,DNA甲基化、组蛋白修饰和基因组印记中的各种表观遗传异常是SCNT的低成功率的肇因。对于改善克隆效率存在显著需求。

发明内容

[0006] 本发明提供用于改善克隆效率的方法。本发明提供用于改善克隆效率的方法。在具体实施方案中,本发明提供用于改善体细胞核转移效率的方法,其牵涉Xist敲除的供体细胞中的Kdm4d过度表达。

[0007] 一方面,本发明提供一种获得克隆囊胚的方法,所述方法包括将从缺乏Xist活性的体细胞获得的供体细胞核转移至去核卵母细胞内,并且在所述卵母细胞内表达Kdm4d,从而获得克隆囊胚。在该方法的一些实施方案中,向卵母细胞注射Kdm4d mRNA。在一些实施方案中,供体细胞核是从包含Xist删除或包含无活性形式的Xist的胚胎成纤维细胞获得的。在一些实施方案中,供体细胞核是从人类、猫、牛、狗、猪或马获得的。在一些实施方案中,该方法也包括将囊胚转移至宿主子宫内进行妊娠。在一些实施方案中,相对于传统的体细胞核转移,所述方法将活产率增加了至少约10%至20%。本发明的一些方面包括通过上述方法生产的囊胚。本发明的一些方面包括通过植入通过上述方法生产的囊胚而生产克隆有机体。

[0008] 另一方面,本发明提供获得以供移植入受试者体内的细胞或组织的方法,该方法包括:在从受试者获得的培养细胞中令Xist失活或降低Xist活性或表达;将来自所述培养细胞的细胞核转移至去核卵母细胞内,从而激活所述卵母细胞;以及向激活卵母细胞注射Kdm4d mRNA并且培养所得细胞,从而获得适用于移植入所述受试者体内的细胞或组织。在本发明的一些方面,提供通过这一方法生产的细胞或组织。在一些实施方案中,通过基因组

编辑令Xist失活。例如,在一些实施方案中,使用CRISPR系统将删除或失活突变引入基因组Xist多核苷酸中。在该方法的其它实施方案中,使用siRNA或shRNA来降低Xist多核苷酸表达或活性。

[0009] 在本发明的其它方面,提供一种细胞,其具有Xist删除或降低水平的Xist表达并且具有编码Kdm4d的异源多核苷酸。

[0010] 另外的方面包括一种卵母细胞,其包含从缺乏Xist活性的体细胞获得的供体细胞核并且包含含量相对于传统卵母细胞增加的Kdm4d。

[0011] 定义

[0012] 除非另做定义,否则本文中使用的所有科技术语均具有本发明所属领域技术人员所一般理解的意义。下述参考文献对技术人员提供本发明中使用的多个术语的一般性定义:Singleton等人所著《微生物学和分子生物学词典(第二版)》(Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed.1994));《剑桥科技词典》(The Cambridge Dictionary of Science and Technology (Walker ed.,1988));Rieger等人编撰的《遗传性术语表(第五版)》(The Glossary of Genetics,5th Ed.,R.Rieger et al. (eds.), Springer Verlag (1991));以及Hale和Marham所著《哈珀·柯林斯生物学词典》(Hale&Marham, The Harper Collins Dictionary of Biology (1991)).如本文中所示,除非明确排除,否则下述术语具有归于其下方的意义。

[0013] “KDM4D多肽”意为一种多肽或其片段,其与NCBI参考序列Q6B0I6具有至少约85%的氨基酸序列一致性并且具有去甲基酶活性。例示性的KDM4D氨基酸序列提供于下:

[0014] >sp|Q6B0I6|KDM4D_HUMAN赖氨酸特异性去甲基酶4D OS=智人(Homo sapiens)

[0015] OX=9606 GN=KDM4D PE=1 SV=3

[0016] METMKSKANCAQNPNCNIMIFHPTKEEFNDFDKYIAYMESQGAHRAGLAKIIPPKEWKARETYDNISEI
LIATPLQQVASGRAGVFTQYHKKKKAMTVGEYRHLANSKKYQTPPHQNFEDLERKYWKNRIYNSPIYGADISGSLFD
ENTKQWNLGHLGTIQDLLEKECGVVIEGVNTPYLYFGMWKTTFAWHTEDMDLYSINYLHLGEPKTWYVVPPEHGQRL
ERLARELFPGSSRGCAGFLRHKVALISPTVLKENGIPFNRIQTQEAGEFMVTFPYGYHAGFNHGFNCAEAINFATPRW
IDYGKMASQCSCGEARVTFSMDFVRILQPERYDLWKRQDRAVVDHMEPRVPASQELSTQKEVQLPRAALGLRQL
PSHWARHSPWPMAARSGTRCHTLVCSLPRRSVAVSGTATQPRAAAVHSSKKPSSTPSSTPGPSAQIIHPSNGRRGRG
RPPQKLRAQELTLQTPAKRPLLAGTTCTASGPEPEPLPEDGALMDKPVPLSPGLQHPVKASGCSWAPVP

[0017] “KDM4D多核苷酸”意为编码KDM4D多肽的核酸分子。例示性的KDM4D核酸提供于下:

1 aaggggcggg gccgaagcgg ccagggggc gggcgtttga aatcagtgcc ttagagtaga
61 ccctaaacct cttttatac cttcaagaac caattactta atgtctcttc cgtcttttcc
121 gtccccgacc cctcccaga ctcttcatt cggtaactgc gtggacggaa agccccgggt
181 agccgacacc acgtccccgg ctacggggag agagcgtgga aaaggattac accaaactgt
241 ttaaatccaa cgactcctgc ttccatcctt tctcctgagc tagaaccaac aaacctagag
301 agttgggctt cggaaaaact agtgttttca ttttaattgga tatgaagaaa gaacaaat
361 gtacggggca accacgatct ttacaaagaa cataagttcc aggaaagcag gaacctgtc
421 tctcttggtc actgggtgta tctctgcat atagaacagt gcctggcaca taatagggtc
[0018] 481 tgaattttgt tctaaacct gaggacatc tctgctacat ttgggtcgta cccccaggtc
541 tgagtaattc aatagactta agaagacaga gccagcagc aaccgaaaca taacagagtt
601 gcaggatcag ctaacgtcaa tgctgggca aagctgctgc ccagagtgga atctcactag
661 tgaataaaca agcccaagaa agattatcat ctcatctgca aaaaaaaaaag tacgctggta
721 gatcctgcta cctcatagat aacaccagtc aaatTTTTTT ttaaagtagc attttctac
781 attgtcaact atctagaaca tacctaaaaa ctaagagttt actgcttatt aatggaaac
841 tatgaagtct aaggccaact gtgccagaa tccaaattgt aacataatga tatttcatcc
901 aaccaagaa gagtttaatg attttgataa atatattgct tacatggaat cccaagggtc
961 acacagagct ggcttggtc agataattcc acccaaagaa tggaaagcca gagagaccta

1021 tgataatatac agtgaaatct taatagccac tcccctccag caggtggcct ctgggagggc
 1081 aggggtgttt actcaatacc ataaaaaaaa gaaagccatg actgtggggg agtatcgcca
 1141 tttggcaaac agtaaaaaat atcagactcc accacaccag aatttcgaag atttgagcgc
 1201 aaaatactgg aagaaccgca totataatc accgatttat ggtgctgaca tcagtggctc
 1261 cttgtttgat gaaaacacta aacaatggaa tcttgggcac ctgggaacaa ttcaggacct
 1321 gctggaaaag gaatgtgggg ttgtcataga aggcgtcaat acaccctact tgtactttgg
 1381 catgtgaaa accacgtttg cttggcatac agaggacatg gacctttaca gcatcaacta
 1441 cctgcacctt ggggagccca aaacttggtg tgtggtgccc ccagaacatg gccagcgctc
 1501 ggaacgcctg gccagggagc tcttcccagg cagttcccgg ggttgtgggg ccttccctgcg
 1561 gcacaagggtg gccctcatct cgcctacagt tctcaaggaa aatgggatte cttcaatcg
 1621 cataactcag gaggctggag agttcatggt gacctttccc tatggctacc atgctggctt
 1681 caacatggt tccaactgcg cagaggccat caattttgcc actccgcgat ggattgatta
 1741 tggcaaatg gctcccagt gtagctgtgg ggaggcaagg gtgacctttt ccatggatgc
 1801 cttcgtgcg atcctgcaac ctgaacgta tgacctgtgg aaacgtgggg aagaccgggc
 1861 agttgtggac cacatggagc ccagggtacc agccagccaa gagctgagca ccagaagga
 1921 agtccagta cccaggagag cagcgtggg cctgagacaa ctcccctccc actgggcccg
 [0019] 1981 gcattcccct tggcctatgg ctgcccgcag tgggacacgg tgccacaccc ttgtgtgctc
 2041 ttcactcca cgccgatctg cagttagtgg cactgctacg cagccccggg ctgctgctgt
 2101 ccacagctct aagaagccca gctcaactcc atcatccacc cctggccat ctgcacagat
 2161 tatccaccg tcaaatggca gacgtggctg tggctgccc cctcagaaac tgagagctca
 2221 ggagctgacc ctccagactc cagccaagag gccctcttg gcgggcacaa catgcacagc
 2281 ttccggccca gaacctgagc cctacctga ggatggggct ttgatggaca agcctgtacc
 2341 actgagccca gggctccagc atcctgtcaa ggcttctggg tgcagctggg cccctgtgcc
 2401 ctaagtccac gggctgtctt tatatcccac tgccctgctg tgtgacagtt tgatgaaact
 2461 ggttacattt acatcccaaa actttggttg agtttgcagg actctaggca tgcagaaag
 2521 agccccctg gtgatgcctt tggatgctgc caagtccatg gtagttttca atttgccat
 2581 acttttgctt ttcctaccgg accctggaat gtctttggat attgctaaaa tctatttctg
 2641 cagctgaggt tttatccact ggacacattt gtgtgtgaga actaggtctt gttgaggtta
 2701 gcgtaacctg gtatatgcaa ctaccatcct ctgggccaac tgtggaagct gctgcacttg
 2761 tgaagaatcc tgagcttga ttcctcttca gtctacgat ttctctcttc cctccctca
 2821 cccccctttt cttataaaac taggttcttt atacagataa ggtcagtaga gttccagaat
 2881 aaaagatatg acttttctga gttatttatg tacttaaaat atgttgcac agtatttgtt
 2941 cccaaatata ttaaaggtaa ccaaatggtt aaaaaaaaa aaaaaaaa

[0020] “EZH1多肽” (组蛋白-赖氨酸N-甲基转移酶EZH1) 意为一种蛋白质或其片段, 其与提供在NCBI参考序列: NP_001982的序列的氨基酸一致性为至少约85%并且具有甲基转移酶活性。例示性的H3K27甲基转移酶氨基酸序列提供于下:

1 meipnpptsk citywkrkvk seymrlrqlk rlqanmgaka lyvanfakvq ektqilneew
 61 kklrvqpvqs mkpvsghpfl kkctiesifp gfasqhmlmr slntvalvpi myswsplqqn
 [0021] 121 fmvedetvlc nipymgdevk eedetfiel innydgkvhg eeemipgsvl isdavfilelv
 181 dalnqysdee eeghndtsdg kqddskedlp vtrkrkrhai egnkksskkq fpndmif sai
 241 asmfpengvp ddmkeryrel temsdpnalp pqctpnidgp naksvqreqs lhsfhtl fcr

301 rcfkydcflh pfhatpnvyk rknkeikiepepcgtdcfl1 legakeyam1 hnprskcsgr
 361 rrrrhivsa scsnasasav aetkegdsdr dtgndwasss seansrcqtp tkqkaspapp
 421 qlcvveapse pvewtgaees lfrvfhgtyf nncfsiar11 gktkckqvfg favkeslik
 [0022] 481 lptdelmnp qkkkrkhr1w aahrkiqlk kdnsstqvyn yqpcdhprp cdstpcimt
 541 qnfcekfcqc npdcqnrifpg crcktqcntk qcpcylavre cdpdlcltgc asehwdckvv
 601 skncsiqrg lkkhlll1aps dvagwgtfik esvqknefis eycgelisqd eadrrgkvyd
 661 kymssflfn1 nndfvvdatr kgnkirfanh svnpncyakv vmvngdhrig ifakraiqag
 721 eelffdyr1s qadalkyv1gi eretdvl

[0023] “EZHI多核苷酸”意为编码EZHI多肽的核酸分子。示例性EZHI多核苷酸序列提供于NM_001991.4,其转载于下:

1 aggaggcgcg gggcggggca cggcgcaggg gtggggccgc ggcgcgcatg cgtcctagca
 61 gcgggacccg cggctcggga tggaggtcgg acacctgttc tgetgttggt tccctgccatt
 121 ctctgaaga acagaggcac actgtaaac ccaacacttc ccctgcatt ctataagatt
 181 acagcaagat ggaaatacca aatcccccta cctccaaatg tatcacttac tggaaaagaa
 241 aagtgaatc tgaatacatg cgacttcgac aacttaaacg gcttcaggca aatatgggtg
 301 caaaggcttt gtatgtggca aat111t1gca aggttcaaga aaaaaccag atcctcaatg
 361 aagaatggaa gaagcttctg gtccaacctg ttcagtcaat gaagcctgtg agtggacacc
 421 cttttctcaa aaagtgtacc atagagagca ttttccggg atttgcaagc caacatatgt
 481 taatgaggtc actgaacaca gttgcattgg tcccatcat gtattcctgg tcccctctcc
 541 aacagaactt tatggtagaa gatgagacgg tttt1gt1gca tattccctac atgggagatg
 601 aagtgaaga agaagatgag acttttattg aggagctgat caataactat gatgggaaag
 661 tccatggtga agaagagatg atccctggat ccgttctgat tagtgatgct gttttctgg
 721 agttggtcga tgccctgaat cagtactcag atgaggagga ggaaggcacc aatgacacct
 781 cagatggaaa gcaggatgac agcaaagaag atctgccagt aacaagaaag agaaagcgc
 841 atgctattga aggcaacaaa aagagttcca agaaacagtt cccaaatgac atgatcttca
 901 gtgcaattgc ctcaatgttc cctgagaatg gtgtcccaga tgacatgaag gagaggatc
 [0024] 961 gagaactaac agagatgtca gacccaatg cacttcccc tca111t1gca cccaacatg
 1021 atggcccaaa tgccaagtct gtgcagcggg agcaatctct gactccttc cacacacttt
 1081 tttgcccggc ctgctttaa tacgactgct tccctcacc ttttcatgcc acccctaag
 1141 tatataaac caagaataaa gaaatcaaga ttgaaccaga accatgtggc acagactgct
 1201 tcc1111t1gct ggaaggagca aaggagtatg ccatgctcca caacccccgc tccaagtgct
 1261 ctggctcgtcg ccggagaagg caccacatag tca111t1gct ctgctccaat gcctcagcct
 1321 ctgctgtggc tgagactaaa gaaggagaca gtgacagggc cacaggcaat gactgggct
 1381 ccagttcttc agaggctaac tctcgtctgc agactccac aaacagaag gctagtccag
 1441 cccacctca actctgcgta gtggaagcacc cctcggagcc tgtggaatgg actggggctg
 1501 aagaatctct ttttcgagtc tccatggca cctacttcaa caactctgt tcaatagcca
 1561 ggcttctggg gaccaagacg tgcaagcag tctt1cagtt tgca111t1gca gaatcactta
 1621 tcc111t1gagct gccacagat gagctcatga acccctcaca gaagaagaaa agaaagcaca
 1681 gattgtgggc tgcacactgc aggaagattc agctgaagaa agataactct tccacacaag
 1741 tgtacaacta ccaacccctgc gaccaccag accgcccctg tgacagcacc tgcccctgca
 1801 tcatgactca gaat111t1gct gagaagttct gccagtgca cccagactgt cagaatcgtt
 1861 tccctggctg tcgctgtaag acccagtgca ataccaagca atgtccttgc tatctggcag
 1921 tgcgagaatg tgaccctgac ctgtgtctca cctgtggggc ctcagagcacc tgggactgca

1981 aggtggtttc ctgtaaaaac tgcagcatcc agcgtggact taagaagcac ctgctgctgg
 2041 cccctctga tgtggccgga tggggcacct tcataaagga gtctgtgcag aagaacgaat
 2101 tcaattctga atactgtggt gagctcatct ctcaggatga ggctgatcga cgcgaaaagg
 2161 tctatgacaa atacatgtcc agcttctct tcaacctcaa taatgatttt gtagtggatg
 2221 ctactcgaa aggaacaaa attcgatttg caaatcattc agtgaatccc aactgttatg
 2281 ccaaagtggc catgggtgaat ggagaccatc ggattgggat ctttgccaag agggcaattc
 2341 aagctggcga agagctcttc tttgattaca ggtacagcca agctgatgct ctcaagtacg
 2401 tggggatcga gagggagacc gacgtccttt agccctccc ggccccacgg cagcacttat
 2461 ggtagcggca ctgtcttggc tttcgtgctc acaccactgc tgctcgagtc tctgactg
 2521 tgtctcccac actgagaaaac cccccaccc actccctctg tagtgaggcc tctgccatgt
 2581 ccagagggca caaaactgtc tcaatgagag gggagacaga ggcagctagg gcttggctc
 2641 ccaggacaga gagttacaga aatgggagac tgtttctctg gcctcagaag aagcagcac
 2701 aggtgggggt ggatgactta tgcgtgattt cgtgtcggct cccaggctg tggcctcagg
 2761 aatcaactta ggcagttccc aacaagcgtc agcctgtaat tgtagcttcc cacatcaaga
 2821 gtccctatgt tattgggatg caggcaaacc tctgtggtcc taagacctgg agaggacagg
 2881 ctaagtgaag tgtggctcct ggagcctaca agtggctctg gttagaggcg agcctggcag
 2941 gcagcacaga ctgaactcag aggtagacag gtcaccttac tacctcctcc ctctggcag
 3001 ggctcaaact gaaagagtgt gggttctaag tacaggcatt caaggctgg ggaaggaaag
 3061 ctacgccatc ctcccttagc cagagagga gaaccagcca gatgatagta gttaaactgc
 3121 taagcttggg ccaggaggc tttgagaaag ccttctctgt gtactctgga gatagatgga
 3181 gaagtgtttt cagattcctg ggaacagaca ccagtgtcc agctcctcca aagtctggc
 3241 ttagcagctg caggcaagca ttatgctgct attgaagaag cattaggggt atgcctggca
 [0025] 3301 ggtgtgagca tctggctcgt ctggatttgt ggggtgtttc aggcctcca tccccatag
 3361 aggcaaggcc caatggccag tgttgcttat cgcttcaggg taggtgggca caggcttggc
 3421 ctagagagga gaaagattgg tgtaactctc tttcctgtct gtagtgctg ctgtttgaa
 3481 aggtgtgatt agaatagtt ccaaggttgg tgaggggcta aattgcacgc gtttaggctg
 3541 gcacccctgt tgcagggcac actggcagag ggtatctgaa gtgggagaag aagcaggtag
 3601 accacctgtc ccaggctgtg gtgccacct ctctggcatt catgcagagc aaagcacttt
 3661 aaccatttct tttaaaagg ctatagattg gggtagagtt tggcctaagg tctctagggt
 3721 cctgctctaa atccactcc tgagggagg ggaagaagag aggggtgggag attctcctcc
 3781 agtctgtct catctcctgg gagaggcaga cgagtgtgatt tcacacagaa gaatttcatg
 3841 tgaatggggc cagcaagagc tgccctgtgt ccatgggtgg tgtgccgggc tggctgggaa
 3901 caaggagcag tatgttgagt agaaaggggt tgggcgggta tagattggcc tgggagtgtt
 3961 acagtggga gcaggcttct ccttcttctc tgggactcag agccccgctt cttccactc
 4021 cacttgttgt cccatgaagg aagaagtggg gttcctcctg acccagctgc ctcttacggt
 4081 ttggtatggg acatgcacac acactcacat gctctcactc accacactgg agggcacaca
 4141 cgtacccgc acccagcaac tctgacaga aagctcctcc cacccaaatg ggccaggccc
 4201 cagcatgate ctgaaatctg catccgccgt ggtttgtatt cattgtgcat atcagggata
 4261 cctcaagct ggactgtggg ttccaaatta ctcatagagg agaaaaccag agaaagatga
 4321 agaggaggag ttaggtctat ttgaaatgcc aggggctcgc tgtgaggaat aggtgaaaaa
 4381 aaacttttca ccagcctttg agagactaga ctgacccac ccttcttca gtgagcagaa
 4441 tcaactgtgt cagtctcctg tcccagcttc agttcatgaa tactctgtt cctccagttt
 4501 cccatccttt gtccctgctg tccccactt ttaaagatgg gtctcaacc cttcccacca
 4561 cgtcatgatg gatggggcaa ggtgggtggg actaggggag cctgggatac atgocgcttc
 4621 attgccaata aatttcatgc actttaaagt cctgtggctt gtgacctctt aataaagtgt
 [0026] 4681 tagaatccaa aaaaaaa
 [0027] “EZH2多肽” (组蛋白-赖氨酸N-甲基转移酶EZH2) 意为一种蛋白质或其片段, 其与

提供在UniProtKB/Swiss-Prot:Q15910.2的序列的氨基酸一致性为至少约85%并且具有甲基转移酶活性。例示性的H3K27甲基转移酶氨基酸序列提供于下：

```

1 mgqtgksek gpvcwrkrvk seymrlrqlk rfrradevks mfssnrqkil erteilnqew
61 kqrriqpvhi ltsvsslrgt recsvtsdld fptqviplkt lnavasvpim yswsplqqnf
121 mvedetvlhn ipymgdevld qdgtfieeli knydgkvhgd recgfindei fvelvналgq
181 ynddddddg ddpeereekq kdledhrddk esrpprkfps dkifeaissm fpdkgtaeel
241 kekykelteq qlpgalpec tpnidgpnak svqreqslhs fhtlfcrrcf kydcflhpfh
301 atpntykrkn tetaldnkpc gpqcyqhleg akefaaalta eriktpkkrp grrrrgrlpn
[0028] 361 nssrpstpti nvlesktdts dreagtetgg enndkeeeek kdetssssea nsrcqtpikm
421 kpnieppenv ewsgaeasmf rvligtyydn fcaiarligt ktcrqvyefr vkessiiapa
481 paedvdtppr kkkrkhlwa ahcrkiqlkk dgssnhvyny qpcdhprqpc dsscpcviaq
541 nfcekfcqs secqnrfgc rckaqcntkq cpcylavrec dpdlcltoga adhwsknvs
601 cknsciqrgs kkhlllpsd vagwgifikd pvqknefise ycgeiisqde adrrgkvydk
661 ymcsflfnln ndfvvdatrk gnkirfanhs vnpncyakvm mvngdhrigi fakraiqtge
721 elffdyrysq adalkyvgie remeip

```

[0029] “EZH2多核苷酸”意为编码EZH2多肽的核酸分子。示例性EZH2多核苷酸序列提供于NM_001203248.1,其转载于下：

```

1 ggcggcgctt gattgggctg ggggggcca ataaaagcga tggcgattgg gctgccgctg
61 ttggcgctcg gtcgggtcgc gtccgacacc cggtgaggact cagaaggcag tggagccccg
121 gcggcgcgcg cggcggcgcg cgggggagac gcgcgggaac aacgcgagtc ggcgcgcggg
181 acgaagaata atcatgggac agactgggaa gaaatctgag aagggaccag tttgttggcg
241 gaagcgtgta aatcagagt acatgcgact gagacagctc aagaggttca gacgagctga
301 tgaagtaaag agtatgttta gttccaatcg tcagaaaatt ttggaaagaa cggaaatctt
361 aaaccaagaa tggaaacagc gaaggataca gctgtgcac atcctgactt cttgttccgt
421 gaccagtgc ttggattttc caacacaagt catccatta aagactctga atgcagttgc
481 ttcagtacc ataattgtatt cttggtctcc cctacagcag aattttatgg tggagatga
541 aactgtttta cataacattc cttatatggg agatgaagtt ttagatcagg atggtacttt
[0030] 601 cattgaagaa ctaataaaaa attatgatgg gaaagtacac ggggatagag aatgtgggtt
661 tataaatgat gaaatTTTTG tggagttggt gaatgccctt ggtcaatata atgatgatga
721 cgatgatgat gatggagacg atcctgaaga aagagaagaa aagcagaaag atctggagga
781 tcaccgagat gataaagaaa gccgccacc tcggaaattt cttctgata aaatTTTTga
841 agccatttcc tcaatgtttc cagataaggg cacagcagaa gaactaaagg aaaaatataa
901 agaactcacc gaacagcagc tcccaggcgc acttctctct gaatgtacc ccaacataga
961 tggaccaa at gctaaatctg ttcagagaga gcaaagctta cactcctt atacgctttt
1021 ctgtaggcga tgttttaaat atgactgctt cctacatcct tttcatgcaa cacccaacac
1081 ttataagcgg aagaacacag aaacagctct agacaacaaa cttgtggac cacagtgtta
1141 ccagcatttg gagggagcaa aggagtttgc tgctgctctc accgctgagc ggataaagac

```

1201 cccaccaaaa cgtccaggag gccgcagaag aggacggctt cccaataaca gtagcaggcc
 1261 cagcaccccc accattaatg tgctggaatc aaaggataca gacagtgata gggaagcagg
 1321 gactgaaacg gggggagaga acaatgataa agaagaagaa gagaagaaag atgaaacttc
 1381 gagctcctct gaagcaaatt ctcggtgtca aacaccaata aagatgaagc caaatattga
 1441 acctcctgag aatgtggagt ggagtgggtgc tgaagcctca atgttttagag tcttcattgg
 1501 cacttactat gacaatttct gtgccattgc taggttaatt gggaccaaaa catgtagaca
 1561 ggtgtatgag tttagagtca aagaatctag catcatagct ccagctcccg ctgaggatgt
 1621 ggatactcct ccaaggaaaa agaagaggaa acaccggttg tgggctgcac actgcagaaa
 1681 gatacagctg aaaaaggacg gtcctctaa ccatgtttac aactatcaac cctgtgatca
 1741 tccacggcag cttgtgaca gttcgtgccc ttgtgtgata gcacaaaatt tttgtgaaaa
 1801 gttttgtcaa tgtagttcag agtgtcaaaa ccgctttccg ggatgccgct gcaaagcaca
 1861 gtgcaacacc aagcagtgcc cgtgctacct ggctgtccga gagtvtgacc ctgacctctg
 [0031] 1921 tcttacttgt ggagccgctg accattggga cagtaaaaat gtgtcctgca agaactgcag
 1981 tattcagcgg ggctccaaaa agcatctatt gctggcacca tctgacgtgg caggctgggg
 2041 gatttttatc aaagatcctg tgcagaaaaa tgaattcatc tcagaatact gtggagagat
 2101 tattttctca gatgaagctg acagaagagg gaaagtgtat gataaatata tgtgcagctt
 2161 tctgttcaac ttgaacaatg attttgtggt ggatgcaacc cgcaagggtg acaaaaattcg
 2221 ttttgcaaat cattcggtaa atccaaactg ctatgcaaaa gttatgatgg ttaacggtga
 2281 tcacaggata ggtatTTTTT ccaagagagc catccagact ggcgaagagc tgtTTTTTga
 2341 ttacagatac agccaggctg atgccctgaa gtatgtcggc atcgaaagag aaatggaaat
 2401 cccttgacat ctgctacctc ctccccctc ctctgaaaca gctgccttag cttcaggaac
 2461 ctcgagtact gtgggcaatt tagaaaaaga acatgcagtt tgaatttctg aatttgcaaa
 2521 gtactgtaag aataatttat agtaatgagt ttaaaaatca actttttatt gccttctcac
 2581 cagctgcaaa gtgttttga ccagtgaatt tttgcaataa tgcagtatgg tacatttttc
 2641 aactttgaat aaagaatact tgaacttgtc cttgttgaat c

[0032] “KDM6A多肽”（赖氨酸特异性去甲基酶6A，也称为组蛋白去甲基酶UTX）意为一种蛋白质或其片段，其与提供在NCBI参考序列：015550.2的序列的氨基酸一致性为至少约85%并且具有去甲基酶活性。例示性的KDM6A氨基酸序列提供于下：

1 mkscgvslat aaaaaafgd eekkmaagka sgeseasps ltaeereal gldsrifgfv
 61 rfhedgartk allgkavrcy eslilkaegk vesdffcqig hfnlilledyp kalsayqryy
 121 slqsdwyna aflyglglvy fhynafqwai kafqevlyvd psfcrakeih lrlglmfkvn
 181 tdyesslkhf qlalvdcnpc tlnsaeiqfh iahlyetqrk yhsakeayeq llqtenlsaq
 241 vkatvlqqlg wmhhtvdllg dkatkesyai qylqkslead pnsqgsyfl grcyssigkv
 [0033] 301 qdafisyrrs idkseasadt wcsigvlyqq qnqpmalqa yicavqlhdg haaawmdlgt
 361 lyescnqpqd aikcylnatr skscsntsal aarikylqag lcnlpqgslq nktklpsie
 421 eawslpipae ltsrqgamnt aqntsdnws gghavshppv qqqahswclt pqklqhleql
 481 ranrnnlnpa qklmleqles qfvlmqqhqm rptgvaqvrz tgipngptad sslptnsvsg
 541 qppqlaltrv psvsqpgvrp acpgqplang pfsaghvpcs tsrtlgstdt ilignnhitg
 601 sgsngnvpvl qrnaltlphn rntltssaee pwknqlsnst qglhkgqssh sagpngerpl
 661 sstgpsqhlq aagsgiqnqn ghptlpsnsv tqgaalnhls shtatggggq gitltkeskp

721 sgniltvpet srhtgetpns tasveglpnh vhmtdavc spshgdskspp gllssdnppq1
 781 sallmgkann nvgtgtcdkv nnihpavhtk tdnsvassps saistatpsp ksteqtttns
 841 vtslnsphsg lhtingegme esqspmktdl llvnhkpspq iipsmsvsiy pssaevlkac
 901 rnlgknglsn ssilldkcpp prppsspypp lpkdklnppt psiylenkrd affpplhqfc
 961 tnpnnpvtvi rglagalkld lglfstktlv eannehmvev rtqllqpade nwdptgtkki
 [0034] 1021 whcesnrsh tkiakyaqyqa ssfqeslree nekrshhkdh sdsestssdn sgrrrkqpfk
 1081 tikfgtnidl sddkkwklql heltklpafv rvvsagnlls hvghtilgmn tvqlymkvpg
 1141 srtpghqenn nfcsvninig pgdcewfvvp egywgvln df ceknnlnflm gswwpnledl
 1201 yeavpvyrf iqrpgdlvwi nagtvhwvqa igwcnniawn vgpltacqyk laveryewnk
 1261 lqsvksivpm vhlswnmarn ikvsdpklfe mikycllrtl kqcqtlreal iaagkeiiwh
 1321 grtkeepahy csicevevfd llfvtnesns rktyivhcqd carktsgnle nfvvleqykm
 1381 edlmqvdydqf tlapplpsas s

[0035] “KDM6A多核苷酸”意为编码KDM6A多肽的核酸分子。例示性的KDM6A多核苷酸序列提供于NM_001291415.1。

[0036] “KDM6B多肽”（赖氨酸特异性去甲基酶6，也称为含JmjC结构域的蛋白质3）意为一种蛋白质或其片段，其与提供在NCBI参考序列：015054.4的序列的氨基酸一致性为至少约85%并且具有去甲基酶活性。例示性的KDM6B氨基酸序列提供于下：

1 mhravdppga raareafalg glscagawss cpphppprsa wlpggrcsas igqpplpap1
 61 ppshgsssgh pskpypapga ptprrlhgkl eslhgcvcqal lrepaqpglw eqlgqlyese
 121 hdseeatrcy hsalryggsf aelgprigr l qqaqlwnfht gscqhrakvl ppleqvwnll
 181 hlehkrnyga krggppvkra aeppvvqvpv paalsgpsge eglsppgkrr rgcnseqtgl
 241 ppglplpppp lpppppppppp pppplpplat sppfqltkpg lwstlhgdaw gperkgsapp
 301 erqeqrshlp hpyypapay tahppghrlv paappgpgpr ppgaeshgcl patrppgsdl
 361 resrvqrsrm dssvspaatt acvpyapsrp pglpgtttss sssssntgl rgvepnpqip
 421 gadhyqtpal evshhgrlgp sahssrkpfl gapaatphls lppgpssppp ppcprllrpp
 481 pppawlkpca craaredei leelffgteg pprpappplp hregflgppa srfsvgtqds
 [0037] 541 htpptpopt tsssnnsngs hssspagpvs fppppylars idplprppsp aqnpqdplv
 601 pltlalppap psschqntsg sfrrpesrpr rvsfpktpev gpgpppgpls kapqpvppgv
 661 gelpargprl fdfpoptled qfeepaefki lpdglanimk mldesirkee eqqqheagva
 721 pqpplkepfa slqspfptdt aptttapava vttttttttt ttatqeeek pppalppppp
 781 lakfpppsqp qpppppppsp asllkslasv legqkykyrg tgaavstrpg plpttqyspg
 841 ppsgatalpp tsaapsaags pppsassssq fstsggpwar errageepvp gpmtptqppp
 901 plslpparse sevleeisra cetlvervgr satdpadpvd taepadsgte rllppaqake
 961 eaggvaavsg sckrrqkehq kehrrhrrac kdsvglrrpre grakakakvp keksrrvlgn
 1021 ldlqseeiqg reksrpdlgg askakpopt appsapapsa qptppsasvp gkkareeapg
 1081 ppgvsradml klrslsegpp kelkirlikv esgdketfia severeirlm adltishcaa
 1141 dvvrasrnak vkgkfresyl spaqsvkpi nteeklprek lnpptpsiyl eskrdafspv

1201 llqfctdprn pitvirglag slrlnlglfs tktlveasge htvevrtqvq qpsdenwdlt
 1261 gtrqiwpces srshttiaky ayyqassfqe slqeekesed eeseepdstt gtpssapdp
 1321 knhhiikfgt nidlsdakrw kpqlqellkl pafmrvtstg nmlshvghti lgmntvqlym
 [0038] 1381 kvpgsrtphg qennnfcsvn inigpgdcew favhehywet isafcdrhgv dyltgswwpi
 1441 lddlyasnip vyrfvqrpgd lwwinagtvh wvqatgwcn iawnvgplta yyyqlalery
 1501 ewnevknvks ivpmihvswv vartvkisdp dlkfmikfcl lqsmkhcqvq reslvragkk
 1561 iayqgrvkde payycnedv evfnilfvts engsrntylv hcegcarrs aglqgvvvle
 1621 qyrteelaqa ydaftlapas tsr

[0039] “KDM6B多核苷酸”意为编码KDM6B多肽的核酸分子。示例性KDM6B多核苷酸序列提
 供于NM_001080424.2,其转载于下:

1 ggcaacatgc cagccccgta gcaactgcca ccccacccac tgtggtctgt tgtacccccac
 61 tgctgggggtg gtggttcaa tgagacaggg cacaccaaac tccatctggc tgttactgag
 121 ggggagacac ggggtgatgat tggctttctg gggagagagg aagtcctgtg attggccaga
 181 tctctggagc ttgccgacgc ggtgtgagga cgctcccacg gaggccgga ttggtctgta
 241 aaggactgag gcagccatct ggggtagcg ggcactctta tcagagcggc tggagccgga
 301 ccatcgtccc agagagctgg ggcagggggc cgtgccaat ctccagggct cctggggcca
 361 ctgctgacct ggctggatgc atcgggcagt ggaccctca ggggcccgcg ctgcacggga
 421 agcctttgcc cttgggggccc tgagctgtgc tggggcctgg agctcctgcc cgcctcatcc
 481 ccctcctcgt agcgcacgac tgccctggagg cagatgctca gccagcattg ggcagcccc
 541 gcttctcgtc ccctacccc cttcacatgg cagtagttct gggcaccaca gcaaaccata
 601 ttatgctcca ggggagccca ctccaagacc cctccatggg aagctggaat cctgcatgg
 661 ctgtgtgcag gcattgctcc gggagccagc ccagccaggg ctttgggaac agcttgggca
 721 actgtacgag tcagagcacg atagttagga ggccacacgc tgcaccaca gcgcccttcg
 781 atacggagga agcttcgctg agctggggcc ccgcatggc cgactgcagc aggccagct
 841 ctggaacttt catactggct cctgccagca ccgagccaag gtcttgcacc cactggagca
 [0040] 901 agtgtggaac ttgctacacc ttgagcaca acggaactat ggagccaagc ggggaggtcc
 961 cccggtgaag cgagctgctg aacccccagt ggtgcagcct gtgcctcctg cagcactctc
 1021 agggccctca ggggaggagg gcctcagccc tggaggcaag cgaaggagag gctgcaactc
 1081 tgaacagact ggccttcccc cagggtgccc actgcctcca ccaccattac caccaccacc
 1141 accaccacca ccaccaccac caccaccctt gcctggcctg gctaccagcc cccatttca
 1201 gctaaccaag ccagggtgtg ggagtacct gcatggagat gcctggggcc cagagcga
 1261 ggggttcagca ccccagagc gccaggagca gcggcaactc ctgcctcacc catatccata
 1321 ccagctcca gcgtacaccg cgcaccccc tggccaccgg ctggtcccgg ctgtccccc
 1381 agggccagc cccgcccc caggagcaga gagccatggc tgctgctc ccaccgctcc
 1441 ccccggaagt gaccttagag agagcagagt tcagaggtcg cggatggact ccagcgttcc
 1501 accagcagca accaccgct gcgtgcctta cgcctctcc cggccccctg gcctccccgg
 1561 caccaccacc agcagcagca gtagcagcag cagcaacact ggtctccggg gcgtggagcc
 1621 gaaccagge attcccggcg ctgaccatta ccaaactccc gcgctggagg tctctacca
 1681 tggcgcctg gggccctcg cacacagcag tcggaaaccg ttcttggggg ctcccgtgc
 1741 cactccccac ctatccctgc cacctggacc ttctcacc cctccacccc cctgtccccg
 1801 cctcttacgc ccccaccac ccctgctg gttgaaggt ccggcctgcc gggcagcccg
 1861 agaggatgga gagatcttag aagagctct ctttgggact gagggacccc cccgcccctg

[0041]

1921 cccaccaccc ctccccatc gcgagggcct cttggggcct cggcctccc gcttttctgt
 1981 gggcaactcag gattctcaca cccctcccac tcccccaacc ccaaccacca gcagtagcaa
 2041 cagcaacagt ggcagccaca gcagcagccc tgctgggcct gtgtcctttc ccccaccacc
 2101 ctatctggcc agaagtatag accccttcc cggcctccc agcccagcac agaaccccc
 2161 ggaccacact cttgtacccc tgaactctgc cctgcctcca gccctcctt cctcctgcca
 2221 ccaaaatacc tcaggaagct tcaggcgccc ggagagcccc cggcccaggg tctcctccc
 2281 aaagaccccc gaggtggggc cggggccacc cccaggcccc ctgagtaaag cccccagcc
 2341 tgtgcgcgcc ggggttgggg agctgcctgc cggaggccct cgactctttg attttcccc
 2401 cactcgcctg gaggaccagt ttgaggagcc agccgaattc aagatcctac ctgatgggct
 2461 ggccaacatc atgaagatgc tggacgaatc cattcgcaag gaagaggaac agcaacaaca
 2521 cgaagcaggc gtggcccccc aacccccgct gaaggagccc tttgatctc tgcagtctcc
 2581 tttccccacc gacacagccc ccaccactac tgctcctgct gtgcgcgtca ccaccaccac
 2641 caccaccacc accaccacca cggccacca ggaagaggag aagaagccac caccagcct
 2701 accaccacca ccgctctag ccaagttccc tccacctct cagccacagc caccaccacc
 2761 cccaccccc agcccggcca gcctgctcaa atccttggcc tccgtgctgg agggacaaaa
 2821 gtactgttat cgggggactg ggcagctgt tccaccgg cctgggcct tgcccaccac
 2881 tcagtattcc cctggcccc catcaggtgc taccgccctg ccgccacct cagcggcccc
 2941 tagcgcagc ggctccccac agcctctgc ttctcgtca tctcagttct ctacctcagg
 3001 cgggccttg gcccgggagc gcagggcggg cgaagagcca gtcccgggcc ccatgacccc
 3061 caccacaccg ccccccccc tatctctgcc cctgctcgc tctgagtctg aggtgctaga
 3121 agagatcagc cgggcttgcc agaccctgt ggagcgggtg ggccggagtg ccaactgacc
 3181 agccgaccca gtggacacag cagagccagc ggacagtggg actgagcagc tgctgcccc
 3241 cgcacaggcc aaggaggagg ctggcggggg ggcggcagtg tcaggcagct gtaagcggcg
 3301 acagaaggag catcagaagg agcatcggcg gcacaggcgg gcctgtaagg acagtgtggg
 3361 tcgtcggccc cgtgagggca gggcaaaggc caaggccaag gtcccaaaag aaaagagccg
 3421 ccgggtgctg gggaacctg acctgcagag cgaggagatc cagggtcgtg agaagtcccg
 3481 gccgatctt ggcggggcct ccaaggccaa gccaccaca gctccagccc ctccatcagc
 3541 tctcgcacct tctgccagc ccacaccccc gtcagcctct gtccctggaa agaaggtcgc
 3601 ggaggaagcc ccagggccac cgggtgtcag ccgggcccag atgctgaagc tgcctcact
 3661 tagtgagggg ccccccaagg agctgaagat ccggctcctc aaggtagaga gtggtgacaa
 3721 ggagacctt atcgcctctg aggtggaaga gcggcggctg cgcattggcag acctcaccat
 3781 cagccactgt gctgctgacg tcgtgcgcgc cagcaggaat gccaaagtga aagggaagtt
 3841 tcgagagtcc tacctttccc ctgcccagtc tgtgaaaccg aagatcaaca ctgaggagaa
 3901 gctgccccgg gaaaaactca acccccctac acccagcatc tatctggaga gcaaacggga
 3961 tgctttctca cctgtcctgc tgcagttctg tacagacct cgaaatcca tcacagtgat
 4021 ccggggcctg ggggctccc tgcggctcaa cttgggcctc ttctccacca agacctggg
 4081 ggaagcaggt ggcgaacaca ccgtggaagt tcgcaccag gtgcagcagc cctcagatga
 4141 gaactgggat ctgacaggca ctgcgcagat ctggccttgt gagagctccc gttcccacac
 4201 caccattgcc aagtacgac agtaccagge ctcatcctc caggagtctc tgcaggagga
 4261 gaaggagagt gaggatgagg agtcagagga gccagacagc accactggaa cccctcctag
 4321 cagcgcacca gaccgaaga accatcacat catcaagttt ggcaccaaca tcgacttgc
 4381 tgatgctaag cgggtggaag cccagctgca ggagctgctg aagctgccc ccttcagcg
 4441 ggtaacatcc acgggcaaca tgctgagcca cgtgggccac accatcctgg gcatgaacac
 4501 ggtgcagctg tacatgaagg tgcccggcag ccgaacgcca ggccaccagc agaataacaa
 4561 cttctgctcc gtcaacatca acattggccc aggcgactgc gagtggttcg cgggtcacga

4621 gcactactgg gagaccatca gcgctttctg tgatcggcac ggcgtggact acttgacggg
4681 ttctctggtgg ccaatcctgg atgatctcta tgcattccaat attcctgtgt accgcttctg
4741 gcagcgcacc ggagacctcg tgtggattaa tgcggggact gtgcactggg tgcaggccac
4801 cggctggtgc aacaacattg cctggaacgt ggggccctc accgcctatc agtaccagct
4861 ggcctcggaa cgatacaggt ggaatgaggt gaagaacgtc aaatccatcg tgcccatgat
4921 tcaagtgtca tggaacgtgg ctgcacaggt caaaatcagc gaccccgact tgttcaagat
4981 gatcaagttc tgctgtctgc agtccatgaa gcaactgccag gtgcaacgcg agagcctggt
5041 gcgggcaggg aagaaaatcg cttaccaggg ccgtgtcaag gacgagccag cctactactg
5101 caacgagtgc gatgtggagg tgtttaacat cctgttctgtg acaagtgaga atggcagccg
5161 caacacgtac ctggtacact gcgagggctg tgcccggcgc cgcagcgcag gcctgcaggg
5221 cgtggtggtg ctggagcagt accgcactga ggagctggct caggcctacg acgccttoac
5281 gctggtgagg gcccgggggc cgcgcgggca gcggaggagg gcaactgggc aggctgcagg
5341 gacgggcttc gggagcccgg ccgcgccttt cctgagccc ccgcccgtt tctccccca
5401 ggcctcagcc agcaactcgc gatgaggcgc gacgcccgc ccgctgctt gcccgcgaa
5461 ggcgcgcgga gccaccagc acatgcctgg gctggacctt ggtcccgcct gtggccgaga
5521 agggggtcgg gccagccct tccaccatc tggcagctcc cctcaactaa tttattaaga
5581 aaaacttttt ttttttttt agcaaatatg aggaaaaaag gaaaaaaaat gggagacggg
[0042] 5641 ggagggggct ggcagcccct cgcaccagc cgcctcccct caccgacttt ggcctttta
5701 gcaacagaca caaggaccag gctccggcgg cggcgggggt cacatacggg ttccctcacc
5761 ctgccagccg cccgcccgc ccggcgcagat gcacgcggct cgtgtatgta catagacggt
5821 acggcagccg aggtttttta tgagattctt tctatgggct ttaccctcc cccggaacct
5881 ctttttttac ttccaatgct agctgtgacc cctgtacatg tctctttatt cacttggtta
5941 tgatttgtat tttttgtct tttctgttt tttgttttt aatttataac agtcccactc
6001 acctctattt attcattttt gggaaaacc gacctcccac acccccagc catcctgcc
6061 gccctccag ggaaccgccc tcgcccggct ctcccgcgc ccagtggtg gtccgggccc
6121 ggcccagacc tctccaccg tccgcccgc gctccagcc ggttctcatg gtgctcaaac
6181 ccgctcccct ccctacgct ctgcacttct tcggaccagt cccccactc ccgaccgac
6241 cccagcccc cctgaggggt agcaactcct gtactgtagg ggaagaagt ggaactgaaa
6301 tggtattttg taaaaaaaaa aaataaaata aaaaaattaa aggttttaaa gaaagaacta
6361 tgaggaaaag gaaccccgct cttcccagc ccggccaact taaaaaaaaa cagaccttca
6421 cccccacccc cttttctttt taagtgtgaa acaaccagc gccagggcct cactggggca
6481 gggacacccc ggggtgagtt tctctgggct tttatcttcg tttgttggg tgtttttct
6541 ccacgctggg gctgcggagg ggtggggggg ttacagtccc gcacctcgc actgcactgt
6601 ctctctgccc caggggcaga ggggtcttcc caaccctacc cctatcttcg gtgatttttg
6661 tgtgagaata ttaatattaa aaataaacgg agaaaaaaaa aaaaaaaaaa aaaaaaaaaa
6721 aaaaaaaaaa a

[0043] “KDM6C多肽” (组蛋白去甲基酶UTY, 也称为在Y染色体上广泛转录的TPR蛋白) 意为一种蛋白质或其片段, 其与提供在NCBI参考序列:014607.2的序列的氨基酸一致性为至少约85%并且具有去甲基酶活性。例示性的KDM6C氨基酸序列提供于下:

[0044] 1 mkscavsltt aavafgdeak kmaegkasre seesvsltv eerealggmd srlfgfvrh
61 edgartktll gkavrcyesl ilkaegkves dffcqlghfn llledyskal sayqryyslq

121 adywknaafl yglglvyfyy nafhwaikaf qdvlyvdpsf crakeihlrl glmfkvntdy
 181 ksslkhfqla lidcnpctls naeiqfhiah lyetqrkyhs akeayeqlq tenlpaqvka
 241 tvlqqlgwmh hnmldvgdka tkesyaiqyl qksleadpns gqswyflgrc yssigkvqda
 301 fisyrqsidk seasadtwcs igvlyqqnq pmdalqayic avqldhghaa awmdlgtlye
 361 scnqppdaik cylnaarskr csntstlaar ikflqngsdn wnggqslshh pvqqvyslcl
 421 tpqklqhleq lranrdlnp aqkhqleqle sqfvlmqmr hkevaqvrvt gihngaitds
 481 slptnsvsnr qphgaltrvs svsqppvrpa cveklssga fsagcipcgt skilgsttdi
 541 llgsnciags esngnvpvlq qnthtlphnh tdlnssteep wrkqlnsaq glhksqsscl
 601 sgpneeplf stgsaqyhqa tstgikkane hltlpsnsvp qgdadshlsc htatsggqqg
 661 imftkeskps knrslvpets rhtgdtngc advkglshv hqliadavss pnhgdspll
 [0045] 721 iadnpqlsal ligkangnvg tgtcdkvnni hpavhtktdh svasspsai statpspkst
 781 eqrsinsvts lnsphsglht vngeglgksq sstkvdplla shrstsqilp smsvsicpss
 841 tevlkacrnpgknslnsnci lldkcppprp ptspyplpk dklnpptpsi ylenkrdaff
 901 pplhqfctnp knpvtvirgl agalkldlgl fstktlvean nehmvevrtq llqpadenwd
 961 ptgtkkiwrc esnrshittia kyaqyqassf qeslreenek rtqkhdhdsn estssensgr
 1021 rrkqpfktik fgtnidlsdn kkwklqlhel tklpafarvv sagnllthvg htilgmntvq
 1081 lymkvpgrt pghqennnfc svninigpgd cewfvvpedy wgvlndfcek nlnflmssw
 1141 wpledlyea nvpvyrfiqr pgdlvwinag tvhwvqavgw cniawnvvp ltacqyklav
 1201 eryewnlks vkspvpmvhl swmarnikv sdpklfemik ycllkilkqy qtlrealvaa
 1261 gkeviwhgrt ndepahycsi cevefnllf vtneantqkt yivhchdcar ktskslenfv
 1321 vleqykmedl iqvydqftla lslssss

[0046] “KDM6C多核苷酸”意为编码KDM6C多肽的核酸分子。示例性KDM6A多核苷酸序列提供于NM_001258249.1,其序列转载于下:

1 gctcatcggt tgttggttag ataatatcat gaactgataa atgcagttgc cacgttgatt
 61 ccctagggcc tggcttaccg actgaggtca taagatatta tgccttctct ttagacttgg
 121 tcagtggaga ggaaatgggc aaagaaccag cctatggagg tgacaaggcc ttagggccaa
 181 aagtcttgag ggtgaaggtt tagggcctgc gcagcttccc tgccatgccc cgcaaggtct
 241 cgcattcgca aggcttgta cagtgggagc ctactacgg actctcctaa agtccatggt
 301 gtcctctttt cgcatttgcg ccccgtaggt gatgcccgat gccgccttc ccatcgctct
 361 cttccccttc aagcgtatcg caactgcaa aacaccagc acagacactc cttttctat
 421 cttaatgcat ttaactagca caacctacag gttgttccat cccagagact accctttct
 481 ccatagacgt gaccatcaac caaccagcg tcagaatcag tcagcctctg tcatgttct
 [0047] 541 aggtccttgg cgaactggct gggcggggct ccagcagcct aggagtagc tggagcaatg
 601 cctgacgtaa gtcaacaaag atcacgtgag acgaatcagt cgcctagatt ggctacaact
 661 aagtgggttg gagcggggag gtcgcgccg ctgctgggg ttcgcccgtg acacaattac
 721 aactttgtgc tgggtgctgc aaagtttgtg attttaagaa attctgctgt gctctccagc
 781 actgcgagct tetgcttcc ctgtagtttc ccagatgta tccaggtagc cgagttccgc
 841 tgcccgtgct tcggtagctt aagtctttgc ctgagctttt ttccttgag ccgctgagga
 901 ggcgataaaa ttggcgtcac agtctcaagc agcgattgaa ggcgtctttt caactactcg
 961 attaaggttg ggtatcgtag tgggacttgg aaatttgttg tttccatgaa atcctgagca
 1021 gtgtcgctca ctaccgccgc tgttgcttc ggtgatgagg caaagaaaat ggcggaagga

[0048]

1081 aaagcgagcc gcgagagtga agaggagtct gttagcctga cagtcgagga aagggaggcg
1141 cttggtggca tggacagccg tctcttcggg ttcgtgagge ttcataaga tggcgccaga
1201 acgaagacce tactaggcaa ggctgttcgc tgctacgaat ctttaatctt aaaagctgaa
1261 ggaaaagtgg agtctgactt cttttgcaa ttaggtcact tcaacctctt gttggaagat
1321 tattcaaaag cattatctgc atatcagaga tattacagtt tacaggetga ctactggaag
1381 aatgctgcgt ttttatatgg ccttggtttg gtctacttct actacaatgc atttcattgg
1441 gcaattaaag catttcaaga tgtcctttat gttgacccca gcttttgtcg agccaaggaa
1501 attcatttac gacttgggct catgttcaaa gtgaacacag actacaagtc tagtttaaag
1561 cattttcagt tagccttgat tgactgtaat ccatgtactt tgtccaatgc tgaattcaa
1621 tttcatattg cccatttgta tgaaaccag aggaagtatc attctgcaa ggaggcatat
1681 gaacaacttt tgcagacaga aaaccttct gcacaagtaa aagcaactgt attgcaacag
1741 ttaggttggg tgcatacaaa tatggatcta gtaggagaca aagccacaaa ggaaagctat
1801 gctattcagt atctcaaaa gtctttggag gcagatccta attctggcca atcgtggtat
1861 tttcttggaa ggtgttattc aagtattggg aaagttcagg atgcctttat atcttacagg
1921 caatctattg ataaatcaga agcaagtga gatacatggt gttcaatagg tgtgtgtat
1981 cagcagcaaa atcagcctat ggatgcttta caggcatata tttgtgctgt acaattggac
2041 catgggcatg ccgcagcctg gatggaccta ggtactctct atgaatcctg caatcaacct
2101 caagatgcca ttaaagtcta cctaaatgca gctagaagca aacgtttag taataacct
2161 acgcttctg caagaattaa atttctacag gctcagttgt gtaacctcc acaaagtagt
2221 ctacagaata aaactaaatt acttctagt attgaggagg catggagcct accaatcccc
2281 gcagagctta cctccaggca gggtgccatg aacacagcac agcaggctta tagagctcat
2341 gatocaaata ctgaacatgt attaaaccac agtcaaacac caattttaca gcaatccttg
2401 tcactacaca tgattacttc tagccaagta gaaggcctgt ccagtcctgc caagaagaaa
2461 agaacatcta gtccaacaaa gaatggttct gataactgga atgggtggca gagtcttca
2521 catcatccag tacagcaagt ttattcgttg tgtttgacac cacagaaat acagcacttg
2581 gaacaactgc gagcaaatag agataattta aatccagcac agaagcatca gctggaacag
2641 ttagaaagtc agtttgtctt aatgcagcaa atgagacaca aagaagttgc tcaggtacga
2701 actactggaa ttcataacgg ggccataact gattcatcac tgctacaaa ctctgtctct
2761 aatogacaac cacatggtgc tctgaccaga gtatctagcg tctctcagcc tggagtctgc
2821 cctgcttctg ttgaaaaact tttgtccagt ggagcttttt ctgcaggctg tattccttgt
2881 ggcacatcaa aaattctagg aagtacagac actatcttgc taggcagtaa ttgtatagca
2941 ggaagtgaag gtaatggaaa tgtgccttac ctgcagcaaa atacacacac tctacctcat
3001 aatcatacag acctgaacag cagcacagaa gagccatgga gaaaacagct atctaactcc
3061 gctcaggggc ttcataaaag tcagagttca tgtttgtcag gacctaatga agaacaacct
3121 ctgttttcca ctgggtcagc ccagtatcac caggcaacta gcactggtat taagaaggcg
3181 aatgaacatc tcaactctgc tagtaattca gtaccacagc gggatgctga cagtcacctc
3241 tctgtcata ctgctacctc aggtggacaa caaggcatta tgtttacaa agagagcaag
3301 cctcaaaaa atagatcctt ggtgcctgaa acaagcagc atactggaga cacatctaat
3361 ggctgtgctg atgtcaaggc acttttctaat catgttcac agttgatagc agatgctgtt
3421 tccagtccta accatggaga ttcacaaat ttattaattg cagacaatcc tcagctctct
3481 gctttgttga ttggaaaagc caatggcaat gtgggtactg gaacctgtga caaagtgaat
3541 aatattcacc cagctgttca tacaagact gatcattctg ttgcctcttc acctcttca
3601 gccatttcca cagcaacacc ttctcctaaa tccactgagc agagaagcat aaacagtgtt
3661 accagcctta acagtcctca cagtggatta cacacagtca atggagaggg gctggggaag
3721 tcacagagct ctacaaaagt agacctgct ttagctagcc acagatctac ttctcagatc

3781 ttaccatcaa tgtcagtgtc tatatgcccc agttcaacag aagttctgaa agcatgcagg
3841 aatccaggta aaaatggctt gtctaatagc tgcattttgt tagataaatg tccacctcca
3901 agaccaccaa cttcaccata cccacccttg ccaaaggaca agttgaatcc acccacacct
3961 agtatttact tggaaaataa acgtgatgct ttctttctct cattacatca atttgtaca
4021 aatccaaaaa accctgttac agtaatacgt ggccttgctg gagctcttaa attagatctt
4081 ggacttttct ctacccaaac tttggtagaa gctaacaatg aacatattgt agaagtgagg
4141 acacagttgc tgcaaccagc agatgaaaac tgggatccca ctggaacaaa gaaaatctgg
4201 cgttgtgaaa gcaatagatc tcatactaca attgccaat acgcacaata ccaggcttcc
4261 tccttcagg aatcattgag agaagaaaat gagaaaagaa cacaacacaa agatcattca
4321 gataacgaat ccacatcttc agagaattct ggaaggagaa ggaaaggacc ttttaaaacc
4381 ataaaatttg ggaccaacat tgacctctct gataacaaaa agtggaaagt gcagttacat
4441 gaactgacta aacttctgc ttttgctgct gtgggtgctg caggaaatct tctaaccat
4501 gttgggcata ccattctggg catgaatata gtacaactgt atatgaaagt tccaggagg
4561 cggacaccag gtcaccaaga aaataacaac ttctgctctg ttaacataaa tattggtcca
4621 ggagattgtg aatggtttgt tgtacctgaa gattattggg gtgttctgaa tgacttctgt
4681 gaaaaaata atttgaattt tttaatgagt tcttgggtggc ccaacctga agatctttat
4741 gaagcaaatg tccctgtgta tagatttatt cagcgacctg gagatttggg ctggataaat
4801 gcaggcactg tgcattgggt tcaagctggt ggctgggtgca ataacattgc ctggaatggt
4861 ggtccactta cagcctgcca gtataaattg gcagtggaac ggtatgaatg gaacaaattg
4921 aaaagtgtga agtcaccagt acccatgggt catctttctt ggaatatggc acgaaatct
4981 aaagtctcag atccaaagct ttttgaaatg attaagtatt gtcttttgaa aattctgaag
5041 caatatcaga cattgagaga agctcttgtt gcagcaggaa aagagggtat atggcatggg
[0049] 5101 cggacaaatg atgaaccagc tcattactgt agcatttctg aggtggagggt ttttaactg
5161 ctttttgtca ctaatgaaag caatactcaa aaaacctaca tagtacattg ccatgattgt
5221 gcacgaaaaa caagcaaaag tttggaaaat tttgtggtgc tcgaacagta caaatggag
5281 gacctaatcc aagtttatga tcaatttaca ctagctcttt cattatcact ctcatctga
5341 tatagttcca tgaatattaa atgagattat ttctgctctt caggaaattt ctgcaccact
5401 ggtttttagt ctgtttcata aaactgttga ctaaaagcta tgtctatgca acctccaag
5461 aatagtatgt caagcaactg gacacagtgc tgcctctgct tcaggactta acatgctgat
5521 ccagctgtac ttcagaaaaa taatattaat catatgtttt gtgtacgtat gacaaactgt
5581 caaagtgaca cagaatactg atttgaagat agcctttttt atgtttctct atttctgggc
5641 tgatgaatta atattcattt gtattttaa cctgcagaat tttccttagt taaaaact
5701 ttctagctg gtcatttctt cataagatag caaatttaa tctctctctg atcagctttt
5761 aaaaaatgtg tactattatc tgaggaagtt ttttactgct ttatgtttt gtgtgtttg
5821 aggcatgat gattacattt gtggttccaa aataattttt taaatatta atagccata
5881 tacaagata atggattgca catagacaaa gaaataaact tcagatttgt gatttttgtt
5941 tctaaacttg atacagattt acactattta taaatagta tttattgctt gaaaatattt
6001 gtgaatggaa tgttgtttt ttcagacgt aactgccatt aaatactaa gaggttctgta
6061 gttttaaaca ctactctat tacattttat atgtgtagat aaaactgctt agtattatac
6121 agaaatttt attaaaattg ttaaatgtt aaagggttc ccaatgtttg agtttaaaaa
6181 agactttctg aaaaaatcca cttttgttc attttcaaac ctaatgatta tatgtattt
6241 atatgtgtg gtatgtgtac acacatgtat aatatataca gaaacctga tatataattg
6301 tatagattt aaaaatttta tttttacat ctatggtagt ttttgagggt cctattataa
6361 agtattacgg aagtttgctg tttttaaagt aaatgtctt tagtgtgatt tattaagttg
6421 tagtcacat agtgatagcc cataaataat tgctggaaaa ttgtatttta taacagtaga

6481 aaacatatag tcagtgaagt aaatatttta aaggaaacat tatatagatt tgataaatgt
 6541 tgtttataat taagagtttc ttatggaaaa gagattcaga atgataacct cttttagaga
 [0050] 6601 acaataaagt gacttatttt tttaaagcta gatgactttg aaatgctata ctgtcctgct
 6661 tgtacaacat ggtttggggg gaaggggagg aaagtattaa aaaatctata tcgctagtaa
 6721 attgtaataa gttctattaa aacttgatt tcatatgaaa aatttgctaa tttaatatta
 6781 actcatttga taataaact tgtcttttct acctctc

[0051] “Gab1多肽”(GRB2相关结合蛋白1)意为一种蛋白质或其片段,其与提供在NCBI参考序列:NP_997006.1的序列的氨基酸一致性为至少约85%。例示性的Gab1氨基酸序列提供于下:

1 msggevvcsq wlrksppek lkryawkrrw fvlrsgrltg dpdvleyykn dhakkpirii
 61 dlnlcqqvda gltfnkkefe nsyifdinti drifylvads eeemnkwwrc icdicgfntp
 121 eedpvkppgs slqapadlpl aintappstq adsssatlpp pyqlinvpph letlgiqedp
 181 qdyllllincq skkpeptrth adsakstsse tdcndnvpsh knpassqskh gmngffqqqm
 241 iydsppsrp sasvdsslyn lprsyshdvl pkvpspsstea dgelyvfnfp sgtssvetqm
 301 rhvsiydip ptpgntyqip rtfpegtlqg tskltdtipdi ppprppkphp ahdrspvetc
 [0052] 361 siprtasdtd ssysiptagm spsrntist vdlnlkrlkda ssqdcydipr afpsdrsssl
 421 egfhnfhkvv nvlvtgsvss eeldenyvpm npnspprqhs ssftepiqea nyvpmtpgtf
 481 dfssfgmqvp ppahmgfrss pktpprrvp vadcepppvd rnlkpdrkgq spkilrlkph
 541 glertdsqti gdfatrrkvk papleikplp eweelqapvr spitrfsard ssrfpmsprp
 601 dsvhsttsss dshdseenyv pmnplnsed pnlfgsnsld ggsspmikpk gdkqveyldl
 661 dldsgkstpp rkqkssgsgs svadervdyv vvdqqktlal kstreawtdg rqstesetpa
 721 ksvk

[0053] “Gab1多核苷酸”意为编码Gab1多肽的核酸分子。示例性Gab1多核苷酸序列提供于NM_002039.3,其转载于下:

1 agggggcgga ggcgcaaagga cagaagctcc ggcaccgagt cggggcagag tcccgtgag
 61 tccgagcgct gctgaggcag ctggcgagac ggcacgtctg gaggcgaggc gggcgactg
 121 aaaggaggcc ggcgcgcccg cgccccggc tcgcttctg ttcaggttcg tgggcctgca
 181 gaggagagac tcgaactcgt ggaaccgcg cacctggag tctgtccgcc cagtccgtcc
 241 ggggtgcgcg accaggagag ctaggttctc gccactgcgc gctcggcagg cgtcggctgt
 301 gtcgggagcg cgccgcgc cctcagctg cccggcccgg agcccagac gcgcgacca
 361 tgagcgggtg tgaagtggc tgctccgat ggctccgcaa gtccccccg gagaaaaagt
 [0054] 421 tgaagcgtta tgcattggaag aggagatggt tctgtttacg cagtggcct ttaactggag
 481 atccagatgt tttggaatat tacaataatg atcatgcaa gaagcctatt cgtattattg
 541 atttaaattt atgtcaacaa gtagatgctg gattgacatt taacaaaaa gagtttgaa
 601 acagctacat ttttgatata aactattg accggatatt ctacttggtg gcagacagcg
 661 aggaggagat gaataagtgg gttcgttgta tttgtgacat ctgtgggtt aatccaacag
 721 aagaagatcc tgtgaagcca cctggcagct ctttacaagc accagctgat ttaccttag
 781 ctataaatac agcaccacca tccaccagg cagattcaco ctctgctact ctacctctc
 841 catatcagct aatcaatggt ccaccacacc tggaaactct tggcattcag gaggatcctc

[0055]

901 aagactacct gttgctcacc aactgtcaaa gcaagaagcc cgaaccacc agaacgcatg
961 ctgattctgc aaaatccacc tttctgaaa cagactgcaa tgataacgtc cttctcata
1021 aaaatcctgc ttctcccag agcaaactg gaatgaatgg ctttttcag cagcaaatga
1081 tatacgactc tocacctca cgtgccccat ctgcttcagt tgactccagc ctttataacc
1141 tgcccaggag ttattccat gatgttttac caaagggtgc tccatcaagt actgaagcag
1201 atggagaact ctatgttttt aataccccat ctgggacatc gagtgtagag actcaaatga
1261 ggcattgtatc tattagttat gacattctc caacacctgg taatacttat cagattccac
1321 gaacatttcc agaaggaacc ttgggacaga catcaaagct agacactatt ccagatattc
1381 ctccacctcg gccaccgaaa ccacatccag ctcatgaccg atctctgtg gaaacgtgta
1441 gtatcccacg caccgcctca gacactgaca gtagtactg tatccctaca gcagggatgt
1501 cgccttcacg tagtaatacc atttccactg tggatttaa caaattgca aaagatgta
1561 gttctcaaga ctgctatgat attccacgag catttccaag tgatagatct agttcaactg
1621 aaggttcca taaccacttt aaagtcaaaa atgtgttgac agtgggaagt gttcaagtg
1681 aagaactgga tgaaaattac gtcccaatga atcccaatc accaccacga caacattcca
1741 gcagttttac agaaccaatt caggaagcaa attatgtgc aatgactcca ggaacatttg
1801 attttctctc atttggaatg caagttctc ctctgctca tatgggttc aggtccagcc
1861 caaaaacccc tcccagaagg ccagttctg ttgcagactg tgaaccacc cccgtggata
1921 ggaacctcaa gccagacaga aaagtcaagc cagcgccttt agaaataaaa ctttggcag
1981 aatgggaaga attacaagcc ccagttgat ctccatcac taggagttt gctcgagact
2041 cttccaggtt tccatgtcc ccccgaccag attcagtga tagcacaact tcaagcagtg
2101 actcacacga cagtgaagag aattatgttc ccatgaacce aaacctgtcc agtgaagacc
2161 caaatctctt tggcagtaac agtcttgatg gaggaagcag ccctatgatc aagcccaaag
2221 gagacaaaca ggtggaatac ttagatctcg acttagatc tgggaaatcc acaccaccac
2281 gtaagcaaaa gagcagtggc tcaggcagca gtgtagcaga tgagagagtg gattatgttg
2341 ttgttgacca acagaagacc ttggctctaa agagtaccg ggaagcctgg acagatggga
2401 gacagtccac agaatcagaa acgccagcga agagtgtgaa atgaaaatat tgccttgcca
2461 tttctgaaca aaagaaaact gaattgtaa gataaatccc tttgaagaa tgacttgaca
2521 cttccactct aggtagatcc tcaaatgagt agagttgaag tcaaaggacc tttctgacat
2581 aatcaagcaa tttagactta agtgggtgctt tgtggatct gaacaattca taacatgtaa
2641 ataatgtggg aaaatagtat tgtttagctc ccagagaaac atttgttcca cagttaacac
2701 actcgtagta ttactgtatt tatgcacttt tcatctaaa acattgttct gggtttccc
2761 aatgtacctt accataatc ctttgggagt tcttgtttt tgtcacacta ctttatataa
2821 caatactaag tcaactaagc tactttttaga tttgaaatt gctgtttaca gtctaacaac
2881 attaaaatga gaggtagatt cacaagttag ctttctacct gaagcttcag gtgataacca
2941 ttagcttata cttggactca tcatttgttg ctttccaaa tgctgaggat aatgtatgta
3001 ctgggtgctg gacctagtct tctggttaat gtacatttag ttttaaatgg tggaaacttg
3061 ttatattttg ttaattacag tgttttgggt tcattgagtg aagattctgc cgggtgggat
3121 cttgcacctt tgaaagactg aataattaca ctaccaagta agcctgcaaa tcattgatgg
3181 catgcagtga tgatgtgctc ttacacttgt taacatgat taagtgttat ttgcaaaagg
3241 tagattatgt aaccaatcag gtacgtacca ggcagtgatg tgctaataca ctgatcaggt
3301 ttagacaatg agctttgggt gtgttcttgt tagtcctaat attggtttct agtttggat
3361 taataaagca gttgacattc actgttagtt acagcaacat actgtgattt ttaattagat
3421 agtaattcag atttattact ctatgaaatt ctgtcttttg acaccatagt gccctttcta
3481 tgattttttt tacttaatat tcttcttggc cttatattta attccctatg caattaatat
3541 tttatatctg cattttttta aaaaaatag atgttatata agtgattctc gtatgtagca

[0056]

3601 cctggtgctt ttocactgaa agaattacgg attttgtact gtgatttata ttocactgcc
3661 caattcaaga aatattggag ccttgctaca atgtgaaatg ttatagtcac ggactccttc
3721 caaccagatt tctgaaaaca ccagagggat ggtataatc tgtctcacct ataacatggt
3781 cctgtgacat agatattaag accacaagtt gtagtgaggc tacaattata ttctgtctgc
3841 ttggctttgc aacataattht agaaagcacg tatagttgtht ttttaaccaa gttacataca
3901 atctcatgta ctgatttgag acttataaca atttttggag ggggcataga gaaaggagtg
3961 cccacagttg aggcacagacc cctccattc agacctctaa ctggttgctg agtacacaga
4021 tgtgccctga tttctggccc attggccata gtagctgccc taatcaatgt aataggtht
4081 ttttcccaat cctcaacta aaaatgtht taacaagatg aattgtagac tagtaacatt
4141 ttagcttht aatatttgct ttttttht acaaaaacta aaaccagaa gtgaatttht
4201 aggtggattt ttaataaaaa aagattgatt gagtttggtg tgcaagctgt tttataatga
4261 aacaacaaaa tgaaatctaa aatcctgaaa tgtgcctaaa ctatcaaac acacgataca
4321 gctaatgtgt aaagatgcta aattctgtht ctggaggat gaatatattt aagatttht
4381 acacaataat aaatacatga ttaattcaaa aataaaaaatc tttacagctg cctatcaagg
4441 gtctaaagca cttaatgaat gtttttagtc taacttatca ttaacttht acaagtcacc
4501 atatttgaag atctgtagca ctctgatttht cagaaaattht ttcattctga ataatttht
4561 aatggtgatg tattagaaag gcagtttgcct ttagaaaact aatcacatt gaacattgta
4621 ttagagaatt aaattaaaag tttcttacag agcagtttht tccaaacatt ttagcacta
4681 gaatcttht agatgaaatt ttatgtataa cccaataca taaagcctga aaactcaatt
4741 ttatcaatat aatgtattt tgggttcaca tttatgctta ttcatttht ctactacta
4801 agcataataa gattctgagc tttctgaa taacacaaat gtggagttat acatagttga
4861 tgaaaccagc agccaattta tagctatgcc ctgthtatt tgtatactat caagaaatt
4921 ttgattcaca caaatgtaag caaaaataat aggtthttht catacatctc aggaaattct
4981 ttaattagag atagctaaag ttattcaagg tctatacaaa aataagttat cctggtagtg
5041 gaagttaata cataagcagc ctccagtggt gtaaagtagg gtatgtaaca catcagaatg
5101 tgcgthttht ttaggthttht aatattgcac gtataaaaac taaattgaa tcaaacctt
5161 ttaactcacc tcaagaagc tagacttht ccaggaatgg gctaaaaacc actggttaac
5221 gatgtgacag ttatgatctt ggagattgga aatcttht cccattaga gttctthtacc
5281 ttaattcctt attctgaaaa attgtaagat tttatgaagg tttgaact gaagcacagc
5341 tctgctthtca aaaattaaaa ttcaaacttg aaaaagctgt ttaaccatg gaagatatca
5401 tttagtaaga tgtaaaagat tttthttht tacacttcag tttatacatc tttatcatta
5461 tcaatactat ataagttact gtgagcattt tagagaatc cataaagga ctatgagtg
5521 gtctgtatgt gtgtgtatat atagcattgt atttaatcat agactaaatt taatttgata
5581 tagaaact actttacttg tacattaagg tcataattc tgctggactc tttatattt
5641 aattaatggg gattatagtc ttcctcata aatgcattta aacctgaaat tgaacaccag
5701 tgtthtctt tttctactta tgggaagttg tctgcttccc ctttagaga aaacagtatt
5761 tttatatttht gthtthttht taactacttht atgcctacac actatgctgt agatactgat
5821 cataattctt ggggttccac aaacactcct agtgctctt tttggccc ttgaaagtgt
5881 tggattact actttcacta cagagccttht ggccctctaa taatgctgag gtgggctgat
5941 ccttcccatt tctgtcttgc ggtcattctg gtaggtcttc tctccactg tcaagtaagc
6001 aatcaggtcc gtgacagga ttggacatat gaacaaatta agtgatgata cacagtgaga
6061 aagatacatg cattctatgg taacaactac tgtcaataac atctgatgtht acatgcacat
6121 ttatatatat ataatttht aaactgaact atgagaagcc atggataaa tgaatattgt
6181 ggacatcatg gacttgatag gatagaaatc aattgtcagc ttgagaaagt tgtthttht
6241 ctgtctaaat agttcatgca ttactacagc taaaatagc ttcatttht tctatagac

6301 ttaatTTTTat tccggttcag tataatctct gttaacagag tttcagcaaa ctgattggtc
 6361 aaggTattaa catagcttct acttccttta cttaaaaaga tgtggTTTTa tgtaagttct
 6421 tgattactga tgatcatccc aaatTTTgac aacaaaatca tatgtataaa tttatttctc
 6481 cctcttTgtt catcatcttt tgtaaaggtc ccattgtaga tcttttctgc taccaaataa
 6541 aactTTTcaa acaattTgtt ttcaagacct taaatagaca agttggatac taagattTgtg
 6601 aactgataag gacataataa tttatatttc cagcccttcc ttagagtctt tatctgcatc
 6661 aaaaacccaa ttctgccatt aactgtgctt cccagtccca cctctatatg tcaactatTT
 6721 tctgcaacaa agatctcact aaatcatgTT gaaacacaag tcatgatcct ctctaagtaa
 6781 atagaaaaag ctccctggaa aaactctgTT gccacatgca cgtgccctgt tactctcca
 6841 gccagccagt gctgccagca ttttattTgtg taaaagtcca aataaataag ggctgcatg
 6901 caacctTTat cttcagaaac taggTTTTat atgtaaaatg tgactTggga aatgattctg
 6961 tttattaact ggctgggatt tttcatttct atgaaagTTT caaacatctc cagtactTTa
 [0057] 7021 taaaatccca acaattgctg taagtcagca cttTggTcca ctCagcccaC ccagcccaCt
 7081 tgcaactctg actcttcaCt gaatcatatt TgggaagTTT gggtagggTg aggctatctt
 7141 cttcaagatt atTTtctcat atgtctgtct gtcacctTgt aaacatgag actcctgggt
 7201 atTTgcatgt aacttctTtg aggaagttac caccatctct gatataGaca cactTTTtga
 7261 gttgcagTTt ctgttagaat tttTgggaga ctaactTgC aattctgtga atgttattga
 7321 atattTaaaa agctgggtct gtaatgggag gcattTTtatt agctgtTgtg attgggtaac
 7381 atgtcccctt agatttctct atTTaaaatt atacaaaatt actattTTTg ataaaaataa
 7441 ggaacaccta cagaaaatta agtttctaag atgttctat acttcattag aaaagattTT
 7501 attactatta cttatggTta ttggTgatta acacttaatg cgtctcctct gattttgtgt
 7561 tccatgaggt gcttggaaCa tttggagtgc tctgtgcgag ggacatacag tgatataGga
 7621 aattTaaaaa tTaaaataat acccaaaacc cactttatca gatatggTat tgtgatggTt
 7681 aatattatgt gtcaactTgg tgaggctatg gcgcccattg gttTggTcaa acactagcct
 7741 agatgtTgct gtgaatatat tttgtagatg tgattaacat ttacaatcag ttgattTtaa
 7801 gTaaagcaga ttctcatcca aaaaaaaaaa aaaaaa

[0058] “Sfmbt2多肽” (具有四个MBT结构域的scm样蛋白2) 意为一种蛋白质或其片段, 其与提供在NCBI参考序列: NP_001018049.1的序列的氨基酸一致性为至少约85%。例示性的Sfmbt2氨基酸序列提供于下:

1 mestlsasnm qdpssspIek clgsangngd ldseegssle etgfnwgeyl eetgasaaph
 61 tsfkHveisi qsnfqpgmkl evanknnpdt ywvatiitte gqllllrycg ygedrradfw
 121 cdvviadlhp vgwctqnnkv lmpPdaikek ytdwteflir dltgsrtapa nlllegplrgk
 181 gpidlItvgs lielqdsqnp fqywivsvie nvgrlrIry vglEdtesyd qwlfyldyrI
 241 rpvGwcqenk yrmdppseiy plkmasewkc tleKslidaa kfplpmevfk dhadlrshff
 [0059] 301 tvgmKletvn mcepfyispa svtkvfnnhf fqvtiddlrp epsklsmIch adslgilpvq
 361 wclknGvsIlt ppgySgqdf dwadyhkqhg aqeappfcfr ntsfsrgftk nmkleavnpr
 421 npgelcvasv vsvkgrlmwl hleglqtPvp evivdvesmd ifpvGwcean syplTaphkt
 481 vsqkkrkiav vqpekqlpPt vpvkkiphdl clfphldttg tvngkyccpQ lfinhrCfsg
 541 pylnkgriae lpqsvpgpkc vlvlkevIsm iinaaykpgr vlrelqlved phwnfQeetI
 601 kakyrGktyr avkivrTsd qvanfCrrvc akleccpnlF spvlisencp encsihtkK
 661 ytyyygkrkk iskppigesn pdsghpkpar rrrkrksifv qkrrrSsavd ftagSgeese
 721 eedadamddd taseetgsel rddqtdtssa evpsarprra vtlrsgsepv rrpPpertrr
 [0060] 781 grgapaassa eegekcpptk pegtedtkqe eerlvlesn plewtvtdvV rfikltdcap
 841 lakifqeqdi dgqalllltI ptvqecmelk lgpaiKlchq iervkvafya qyan

[0061] “Sfmbt2多核苷酸”意为编码Sfmbt2多肽的多核苷酸。示例性Sfmbt2多核苷酸序列提供于NM_001018039.1,其转载于下:

```
1 cgccttgtgt gtgctggatc ctgctggggt agatccccga gtaatttttt ctgcaggatg
   61 aattaagaga agagacactt gctcatcagg catggagagc actttgtcag cttccaatat
  121 gcaagacctt tcattctcac ccttggaaaa gtgtctcggc tcagctaata gaaatggaga
  181 ccttgattct gaagaaggct caagcttggg ggaactggc ttaactggg gagaatattt
  241 ggaagagaca ggagcaagtg ctgctcccc caccatcatc aaacacgttg aaatcagcat
  301 tcagagcaac ttccagccag gaatgaaatt ggaagtggct aataagaaca acccggacac
  361 gtactgggtg gccacgatca ttaccacgtg cgggcagctg ctgcttctgc gctactgcgg
  421 ttacggggag gaccgcaggg ccgacttctg gtgtgacgta gtcatcgcgg atttgaccc
  481 cgtgggggtg tgcacacaga acaacaagg gttgatgccg ccggacgcaa tcaaagagaa
  541 gtacacagac tggacagaat ttctcatac gtaacttgact ggttcgagga cagcaccgc
  601 caacctctg gaagtcctc tgcgagggaa aggcctata gacctatta cagttggttc
  661 cttaatagaa cttcaggatt ccagaacct ttttcagta cggatagtta gtgtgattga
  721 aaatggttga ggaagattac gccttcgta tgtgggattg gaggacactg aatcctatga
  781 ccagtgggtt ttttacttgg attacagact tcgaccagtt ggttgggtgc aagagaataa
  841 atacagaatg gaccacctt cagaaatct tctttgaag atggcctctg aatggaaatg
  901 tactctggaa aaatccctt ttgatgctg caaatttct cttccaatgg aagtgtttaa
  961 ggatcacgca gatttgcgaa gccatttctt cacagttggg atgaagcttg agacagtgaa
1021 tatgtgcgag cctttttaca tctctcctgc gtcgggtgact aaggttttta acaatcactt
[0062] 1081 ttttcaagtg actattgatg acctaagacc tgaaccaagt aaactgtcaa tgctgtgcc
1141 tgcagattct ttggggattt tgccagtaca gtggtgcctt aaaaatggag tcagcctcac
1201 tcctcccaaa ggttactctg gccaggactt cgactgggca gattatcaca agcagcatgg
1261 ggcgcaggaa gccctcctt tctgcttccg aaatacatca tcagtcgag gtttcacaaa
1321 gaacatgaaa cttgaagctg tgaaccccag gaatccagga gaactgtgtg tggcctccgt
1381 tgtgagtgtg aaggggcggc taatgtggct tcacctggaa gggctgcaga ctctgttcc
1441 agaggtcatt gttgatgtg aatccatgga catcttccca gtgggctggt gtgaagccaa
1501 ttcttatcct ttgactgcac cacacaaaac agtctcaca aagaagagaa agattgcagt
1561 cgtgcaacca gagaaacaat tgccgccac agtgctggtt aagaaaatac ctcatgacct
1621 ttgtttattc cctcacctg acaccacagg aaccgtcaac gggaaatac gctgtcctca
1681 gctcttcac aaccacagg gtttctcagg ccttacctg aacaaaggaa ggattgcaga
1741 gctacctcag tcggtgggac cgggcaaat cgtgctggtt cttaaagagg ttcttagcat
1801 gataatcaac gcagcctaca agcctggaag ggtattaaga gaattacagc tggtagaaga
1861 tccccactgg aatttcag gaaagacgct gaaggccaaa tacagaggca aacatacag
1921 ggctgtggtc aaaatcgtac ggacatctga ccaagtcgca aatttctgcc gccgagtctg
1981 tgccaagcta gagtgcctg caaatttgtt tagtctgtg ctgatctctg aaaactgcc
2041 agagaactgc tccattcata ccaaaaccaa atacacctat tactatggaa agagaaagaa
2101 gatctccaag ccccccctc gggaaagcaa cccgcacagc ggacaccca aaccgccag
2161 gcggaggaag cgacggaaat ccatttctgt gcagaagaaa cggaggctt ctgccgtgga
```

2221 cttcaccgcg ggctcggggg aggaaagtga agaggaggac gctgacgcca tggacgatga
 2281 caccgccagt gaggagaccg gctccgagct ccgggatgac cagacggaca cctcgtcggc
 2341 ggaggtgccc tggccccgc cccggagggc cgtcaccctg cggagcggct cagagccctg
 2401 ggcocggcca cccccagaga ggacacgaag gggccgcggg ggcocggctg cctcctcagc
 2461 agaggaaggg gagaagtgcc gcgccacca gcccgagggg acagaggaca cgaaacagga
 2521 ggaggaggag agactgggtt tggagagcaa cccgttgag tggacggta cgcagctggt
 2581 gaggttcatt aagctgacag actgtgcccc cttggccaag atatttcagg agcaggatat
 2641 tgacggccaa gcaactcctg ttctgacctc tccgacggtg caggagtga tggagctgaa
 2701 gctggggcct gccatcaagt tatgccacca gatcgagaga gtcaaagtgg ctttctacgc
 2761 ccagtacgcc aactgagtct gccctcggga ggtggcccat tattgctggg atgcggtggt
 2821 ggtaaagggt tccaggactg aaactttgat tttccgggat atgttaaag gtacagccac
 2881 taagtatcac cagaaaacca gaagcccagg atcttctgcc tccgccagcc tgtgagctgt
 2941 ttocatgttt tcaaagcaca gcagcagtcg cttctgggga gtgccagta aagtcatgca
 3001 tcagaccctg ccagacgtgg gctgtctctc tggctcacc acgttttgcc tttctcctgc
 3061 cccaaatcag gcagctccct tggagcaggg tttcctcaga tgaggactgc attctttgaa
 3121 acaaagaat gtgcacaagg aagaaacctc acgccatgct gtagtgttc ctgtaatcac
 3181 acgagcacat ttatatatgc agtttcccat ggataggcgt gtgaccctgg ttgagtggca
 3241 cttgcgggtt catcttggtg gcaactcctt tgcaatgcag ctggcagcga catccttata
 3301 aaaacatgtg ctaaagctct gtcctctggt agaggtgcct tttaggaata cggggagtga
 3361 aggaaggccg gcaggcatct ccattgcaact agatggtttg tttgtttgt tgtttgtttg
 3421 ttgttcattt tgtttgtttt tttgagacag ggtcttgctc tgcgcgccag gttgtaatgc
 3481 agtggcgcaa tctcagctca ctgcaacctc tctctcccgg gttcaagtga tctcctgcc
 3541 tcagcctccc aagtagctgg gattacagge acccaccacc atgctggct aatttttghta
 3601 tttttggtag agacaggggt tcacatggtt ggtcaggcta gtcttgaact cccaacctca
 3661 agtgatctgc ccgctcggc ctcccaactg gctgggatta caggtgtgag ccactacgcc
 3721 ccggcccaac tggatgggtt ttgattgaag cctagaacat ctgtagagac aaactctacc
 3781 cagtcttttc tagacctca actatctcca gtgttggtgt ttaatcgtag ccggtacagg
 3841 gagtgagtct tttaggcaaa tgttggtgatta tatatcaaag gaaaagctta gtttcagaga
 3901 ggaggaaggg aaagagatgt gagggagca tttcatcaac cagctacgtc ccccttagaa
 3961 ggatcactgc agcaggtcac cgagcaggag tccctctgag cgtccctct gtctcgttct
 4021 gccctagctg gcagcatatg aaccaggcat gatgcagcag gagcagtga tctggagtca
 4081 gccacttggc accctgggtt cgctgagaac aaactctgag atcttgggtg acttctcact
 4141 actctggacc tccattcctg tgaagtgaca ggtgtggacc ctgaggggtg ggtggtgagc
 4201 aactgtctc ctgctggcat tcacccact catgctggaa aggaagatcc agatcgtaca
 4261 aaaattagaa aaagaaagaa taagaaggggt ctgggtcccag ttctgactcg gccattctta
 4321 cagctctttc tggttttgag tttgcttggt gaatttctc ggcagttgtg ttaaaccgc
 4381 caggtcacgt gcagacaaag ctgtggctgc gagagttggt tggcctctg gaccagaagc
 4441 catctccata tctcatgag cgattccata tctccactca gacctgtg actacagtgt
 4501 tccgctgtgg tggctgcaa gatgcctct taaacttatg caaggaaacc aaacctccc
 4561 acagttccca agcagacact ggaagcagag gcttctcacc cttcctgctt tttaccaca
 4621 atcacttga gctcgtccct tggactagag tctccacagt tccagtaaaa tctcgcggtg
 4681 ggctgatgag ctgottgcat ttctgtgaca tttccagata tgattctcag tgggattttg
 4741 gaaactttga ttgtcaagc tcaccttct taacattctg taatggttac agatgagaat
 4801 ggaaaacaca tttttatgg atgagggctt ttggctccc ctgcagtcga tttctagaat
 4861 caagttttag agttcggctg atgcatctgc ctggggacct cagatgggag gagtgtgta

[0063]

4921 gttgtacccc gacagaaatg tctctgggat ctgtggctgg cttgccccgg gcatctctcc
 4981 ttttaagctca agttttgaac tctctgcggt tttccacccc tgccttctca gccacatgct
 5041 tttggcctta aacgctcagt cttgtggagt tcaactctgt caaacgattg gaaagggcat
 5101 ccatttccag atctttggca ttttccccgc gctgactctt tgatgatcct tcaactgtgg
 5161 cttttcaagc tcagctgttc ctgttgatt tgagacgagg gtgaggggaat gtggtggcca
 5221 caaaagaaca gggacttgca gcacaaatgt cacttctgtc tcccttttca gtggtagcac
 5281 ggaggaggag gtgctgcggt ggagggaggg gatcctccag gagctctctg gagcccatct
 5341 aggaagctag agtgtgtggc ccgccaggag ctcaaggaagg atacagccac tgtcgcaggg
 5401 gaaagtgttt gcttcccggt gagccaagcg cccaagactc tccgtatcct tcacctgac
 5461 agtttaactt cagcgtttct ctgtgcagtt gcggtcacca tgggtgagca ctgtctgtgc
 5521 acgtgccagg gaggagatgg ctgggaccac tgcacaggag ggcgcagcct ggcgtgcca
 5581 tgaaagtgtt ctctgtgcca tctctccggt ccttgaggag agcccagaaa gattttagga
 5641 ccaggaggt gcttttctc cagctgttgc cagtgtcctt ctgagcctgg attctccggg
 5701 gatttccgtc gtggtggatg gacttcacat cagcagcagt tctggtacag aattgtaatg
 5761 tgttttcatt tctctgtagg attcacctct caccagcgtc tgtcttaaag gtagggccaa
 5821 tttcatggag ctttttctg tgtgtgtcct tgttgctttt gccagaaaaa gtggatttga
 5881 catgcgtgcc ccgatgccac catagccctt aggccaacaa tgtcatggtc taaacaccaa
 5941 aaagtgatgc ccgcattcc ttcctggat ggtaccgttt cttctccgtc tctctttgat
 6001 gattctttgg gaccaaagtc ctctccttag tgcgcctact tctgtgggc atcatgccac
 6061 ttggaactta ttggaactgg ccggggagac tctgcagtct gcgccgtttg aaaaccctga
 6121 gaaagagatg ccacctcaac ttgaatcatg acagcccatc gctcagctc acctaaact
 6181 catggagctt gtttcagctc ctcaactctt gactgtattt gtactatgtt gaaaaaatat
 6241 cctgtccaca aagacataag cctaacaacc tagaaaaaca acagggtaact actggcatta
 6301 cagaacttct ttgcctttca aaacaaaagc aaaacacagt gaacttcacc acggagctgc
 6361 acagcgtggg gaactcatcc atcaacttca aaattagagt catttgatcc aagttggagt
 6421 cagacacagt atttgagctg cacggcttct gggttctccc accttatttg atcatattcg
 6481 aaagattatt tctgtgttt gctttgattt gttcctcagt acattaaaat gatccacacc
 6541 ttgaaactg cctctctag aaggttgatt ttgatcagcc ttttgaagat ggggtgcggt
 6601 tccctaactt atctcacaga attttgagtg ttgtatttgg caagttctga gatttgcctt
 6661 ctgtcttatg ccaaacaccc ctttctaaga gctgtccccg cttagtttta gaagtactag
 6721 gggttttcat acttatttta tagaacaccc atttatattt atttctgtat atagaactaa
 6781 aaaaaacagt agtgtaaaa atctttgttg tggtttgagc atctttgctg cttttggatt
 6841 gagatggcga atcaaggctt cacttctctc ctcttctgtc tttagaaagc tgtgatcgtg
 6901 cgtgcaatta tttgaaaggc aacatagtca attaagaaac ctgtagtgtt taaggaagaa
 6961 attgttggca agatatccat actgcccata tctcgttggg gcaataatta aatagcaaa
 7021 gaaatctgta ttggcaacta ttataattca ataattcttt tgtttactgc cttttctgt
 7081 tcaagaattt tctggaaatt actcccttcc acatggttga actcttaagt tgaccagttc
 7141 tcatagctct atcaactagaa tggtttgcag ataccccaaa catactatga taaaatcaaa
 7201 ttgtgctact tttgacctt gtaatttacc taaaagtgtt aattgctgac agagtactgc
 7261 cttgaatttt ggtttaaaac ctctctagtt tcaatgacaa gtaacaactc aaataattcc
 7321 atattgtttg aggaagaggc cataatcctt ctgaattgtt ggcactaagt aatgggattt
 7381 ggcccagtaa gtatgacggg cgtgtgcgct aaccaacgca gagcagtgct ttttgtgtgg
 7441 ctgaagcgat gtgctgacga aaaaaggaaa attctaggac aatcgttggc taaaaatcac
 7501 cttaggatga aaaatttgag gcaaattttt ttaaatgaca gaaaaagata atcatctcac
 7561 ttgcttgaag caggagccag catgatctct ggaagcatca actatcctc gtcgtgattg

[0064]

[0065] 7621 ttgaaagctc tttcactggt ttgcattcta gtttgaatag tttgtattga aattggattc
 7681 ctatcttgtg tatgtttttg gtgcgtaaaa gggaaaaatt ggtgtcatta cttttgaaat
 7741 ttgcaggacg aagggcatgc ttttggtttg ctgtaagatt gtattctgta tatatgtttt
 7801 catgtaaata aatgaaaatc tatatcagag ttatatttta atttttattc taaatgaaa
 7861 aaaccctttt tacttcaaaa aaattgtaag ccacattggt aataaagtaa aaataaatc
 7921 ta

[0066] “Smoc1多肽”(SPARC相关模块化钙结合蛋白1)意为一种蛋白质或其片段,其与提供在NCBI参考序列:NP_001030024的序列的氨基酸一致性为至少约85%。例示性的Smoc1氨基酸序列提供于下:

[0067] 1 mlparcarll tphllllvlvq lsparghrtt gprflisdrd pqcnlhcsrt qpkpicasdq
 61 rsyesmceyq rakordptlg vvhrgrckda gqskcrlera qaleqakkpq eavfvpeceg
 121 dgsftqvqch tytgycwcvt pdgkpisgss vqnktpvcsg svtdkplsqq nsgrkddgsk
 181 ptptmetqpv fdgdeitapt lwikhlvikd sklnntnirn sekvyscdqe rqsaleeaqq
 241 npregivipe capgglykpv qchqstgycw cvlvdtdgrpl pgtstryvmp scesarakt
 301 teaddpfkdr elpgcpegkk mefitslllda lttmvmqain saaptgggrf sepdpshle
 361 ervvhwyfsq ldsnssndin kremkpfkry vkkkakpkkc arrftdycdl nkdkvislpe
 421 lkgclgvske vgrlv

[0068] “Smoc1多核苷酸”意为编码Smoc1多肽的核酸分子。示例性Smoc1多核苷酸序列提供于XM_005267995.1,其转载于下:

[0069] 1 ataacgggaa ttcccatggc cggggetcag gcgtecaacc tgctgcccgc tgggccccgc
 61 cgagcgggagc tagcgcgcgc cgcagagcac acgctcgcgc tccagctccc ctctgcgcgc
 121 gttcatgact gtgtcccctg accgcagcct ctgcgagccc ccgccgcagg accacggccc
 181 gctccccgcc gccgcgaggg ccccgagcga aggaaggaag ggaggcgcgc tgtgcgcccc
 241 gcggagcccc cgaacccccgc tcgctgcccg ctgcccagcc tggctggcac catgctgccc
 301 gcgcgctgcg ccgcctgct cacgccccac ttgctgctgg tgttggtgca gctgtcccct
 361 gctcgcggcc accgcaccac aggccccagg tttctaataa gtgaccgtga cccacagtgc
 421 aacctccact gctccaggac tcaacccaaa cccatctgtg cctctgatgg caggctctac
 481 gagtccatgt gtgagtacca gcgagccaag tgccgagacc cgaccctggg cgtggtgcat
 541 cgaggtagat gcaaagatgc tggccagagc aagtgtogcc tggagcgggc tcaagccctg
 601 gagcaagcca agaagcctca ggaagctgtg tttgtcccag agtgtggcga ggatggctcc
 661 tttaccocagg tgcagtgcca tacttacact gggtaactgt ggtgtgtcac cccggatggg
 721 aagcccatca gtggctcttc tgtgcagaat aaaactcctg tatgttcagg ttcagtcacc
 781 gacaagccct tgagccaggg taactcagga aggaaagtct cctttcgatt ctttttaacc
 841 ctcaattcag atgacggggtc taagccgaca cccacgatgg agaccagcc ggtgttcgat
 901 ggagatgaaa tcacagcccc aactctatgg attaaacact tggatgacaa ggactccaaa
 961 ctgaacaaca ccaacataag aaattcagag aaagtctatt cgtgtgacca ggagaggcag
 1021 agtgccctgg aagaggccca gcagaatccc cgtgagggta ttgtcatccc tgaatgtgcc
 1081 cctgggggac tctataagcc agtgcaatgc caccagtcca ctggctactg ctggtgtgtg
 1141 ctggtggaca cagggcgccc gctgcctggg acctccacac gctacgtgat gccagttgt
 1201 gagagcgacg ccagggccaa gactacagag gcggatgacc ccttcaagga cagggagcta

1261 ccaggctgtc cagaagggaa gaaaatggag tttatcacca gcctactgga tgctctcacc
 1321 actgacatgg ttcaggccat taactcagca gcgcccactg gaggtgggag gttctcagag
 1381 ccagacccca gccacaccct ggaggagcgg gtagtgcact ggtatttcag ccagctggac
 1441 agcaatagca gcaacgacat taacaagcgg gagatgaagc ccttcaagcg ctacgtgaag
 1501 aagaaagcca agcccaagaa atgtgcccgg cgtttcaccg actactgtga cctgaacaaa
 1561 gacaaggtea tttcactgcc tgagctgaag ggctgcctgg gtggttagcaa agaagtagga
 1621 cgctctgtct aaggagcaga aaaccaaggg gcaggtggag agtccagggga ggcaggatgg
 1681 atcaccagac acctaacctt cagcgttgcc catggccctg ccacatcccg tgtaacataa
 1741 gtggtgcca ccatgtttgc acttttaata actcttactt gcgtgttttg tttttggttt
 1801 ctttttaaaa caccaatata taataccaca gtgggaaaag gaaaggggaag aaagacttta
 1861 ttctctctct tattgtaagt ttttgatct gctactgaca acttttagag ggttttgggg
 1921 ggggtgggga ggggtgtgtt ggggctgaga agaaagagat ttatatgctg tatataaata
 1981 tatatgtaa ttgtatagtt cttttgtaca ggcattggca ttgctgtttg tttatttctc
 2041 tccctctgcc tgctgtgggt ggtgggact ctggacacat agtccagctt tctaaaatcc
 2101 aggactctat cctgggccta ctaaacttct gtttggagac tgacccttgt gtataaagac
 2161 gggagtctct caattgtact gcggactcca cgagttcttt tctgggtggga ggactatatt
 2221 gccccatgcc attagttgtc aaaattgata agtcacttgg ctctcggcct tgtccagggga
 2281 ggttgggcta aggagagatg gaaactgcc tgggagagga agggagtcca gatcccatga
 2341 atagcccaca caggtaccgg ctctcagagg gtccgtgcat tctgtctctc cggaccccca
 2401 aagggccag cattggtggg tgcaccagta tcttagtgac cctcggagca aattatccac
 [0070] 2461 aaaggatttg cattacgtca ctcgaaacgt tttcatccat gcttagcctc tactctgtat
 2521 aacgcatgag aggggaggca aagaagaaaa agacacacag aagggccttt aaaaaagtag
 2581 atatttaata tctaagcagg ggaggggaca ggacagaaag cctgcactga ggggtgcggt
 2641 gccaacaggg aaactcttca cctccctgca aacctaccag tgaggtccc agagacgcag
 2701 ctgtctcagt gccaggggca gattgggtgt gacctctcca ctctccatc tctgtctgtt
 2761 gtcttagtgg ctatcacagg cctgggtggg tgggttgggg gaggtgtcag tcacctgtt
 2821 ggtaacacta aagttgtttt gttggttttt taaaaacca atactgaggt tcttctgtt
 2881 cctcaagtt ttcttatggg cttccaggct ttaagctaat tccagaagta aaactgatct
 2941 tgggtttcct attctgcctc cctagaaggg gcaggggtga taaccagct acagggaaat
 3001 cccggcccag ctttccacag gcatcacagg catcttccgc ggattctagg gtgggctgcc
 3061 cagccttctg gtctgaggcg cagctccctc tgcccagggtg ctgtgcctat tcaagtggcc
 3121 ttcaggcaga gcagcaagtg gcccttagcg ccccttcca taagcagctg tgggtggcagt
 3181 gagggagggt gggtagccct ggactggtcc cctctcaga tcacccttgc aaactctggcc
 3241 tcatcttgta ttccaaccog acatccctaa aagtaacctc accggttccg ggtctggaag
 3301 gcgttggcac cacaagcact gtccctgtgg gaggagcaca accttctcgg gacaggatct
 3361 gatggggtct tgggctaaag gaggtccctg ctgtcctgga gaaagtccta gaggttatct
 3421 caggaatgac tgggtggcct gccccaacgt ggaaaggtgg gaaggaagcc ttctcccatt
 3481 agcccctatg agagaactca acgtgccgga gctgagtggg ccttgacga gacactggcc
 3541 ccactttcag gcctggagga agcatgcaca catggagacg gcgctgcct gtagatgttt
 3601 ggatcttca gatctccca ggcattctgt cctccacagg atcgtgtgtg taggtgtgtg
 3661 tgtgtggttt tctttgtga aggagagagg gaaactattt gtagcttgtt ttataaaaaa
 3721 taaaaaatgg gtaaactctg

[0071] “在赖氨酸27处三甲基化的组蛋白H3 (H3K27me3)”意为在组蛋白H3蛋白亚基的赖氨酸27处的三甲基化。H3K27me3修饰通常与基因阻遏有关。

[0072] “剂”意为肽、核酸分子或小化合物。

[0073] “等位基因”意为在染色体上相同位置处发现的两个或更多个可供选择形式的基因中的一个。

[0074] “改变”意为基因或多肽的表达水平或活性的变化(增加或减少),如通过该领域已知的方法如本文所述的那些方法所检测。如本文中所示,改变包括表达水平变化至少10%,优选变化25%,更优选变化40%,且最优选表达水平变化50%或更高。

[0075] “缓解”意为减少、阻抑、衰减、缩减、迟滞或稳定化疾病的发展或进展。

[0076] 本公开中,“包含”、“含有”和“具有”等可具有美国专利法中规定的属于它们的意义,且可意为“包括”等;“主要由...组成”等具有美国专利法中规定的属于它们的意义,且该术语是开放性的,允许超过其所引述者的存在,只要其所引述的基本或新颖特征没有被超过其所引述者的存在改变即可,但不包括先前技术实施方案。

[0077] “检测”指的是证实待检测的被分析物的存在、不存在或量。

[0078] “疾病”意为损害或干扰细胞、组织或器官的正常功能的任何病症或病变。病变的示例包括与通过H3K27me3依赖性印记进行的不希望的阻遏相关的那些。小眼畸形例示性病变与H3K27me3依赖性相关,而后者与印记病变相关。

[0079] “DNA”意为脱氧核糖核酸。在多种实施方案中,术语DNA是指基因组DNA、重组DNA或cDNA。在具体实施方案中,DNA包含“靶标区域”。本文设想的DNA文库包括基因组DNA文库以及从RNA构建的cDNA文库,例如,RNA表达文库。在各种实施方案中,DNA文库包含一个或多个附加DNA序列和/或标签。

[0080] “有效量”意为相对于未治疗患者缓解疾病症状所需的化合物的量。用来实践本发明的用于治疗性处理疾病的活性化合物的有效量可变,取决于给药模式以及受试者的年龄、体重和一般健康情况。基本上,主治医师或兽医将决定适宜的量 and 用药方案。这一量被称为“有效”量。

[0081] “片段”意为多肽或核酸分子的一部分。这一部分优选地含有参考核酸分子或多肽总长度的至少10%、20%、30%、40%、50%、60%、70%、80%或90%。片段可包含10、20、30、40、50、60、70、80、90、或100、200、300、400、500、600、700、800、900、或1000个核苷酸或氨基酸。

[0082] 术语“分离的”、“纯化的”或“生物学纯的”指的是物质没有不同程度的在其天然状态下发现的正常伴生组分。“分离”表示与原始来源或周围物质分隔的程度。“纯化”表示高于“分离”的分隔程度。“纯化的”或“生物学纯的”蛋白质足够不含其它物质,使得任何杂质均不在物质上影响该蛋白质的生物学性质或不造成其它负面后果。换言之,如果本发明的核酸或肽基本上在通过重组DNA技术生产时不含细胞物质、病毒物质或培养基,或者当化学合成时不含化学前体或其它化学品,则该核酸或肽是纯化的。典型使用分析化学技术例如聚丙烯酰胺凝胶电泳或高效液相色谱来测定纯度和均质性。术语“纯化的”可表示核酸或蛋白质在电泳凝胶中给出实质上的一条谱带。对于可进行修饰如磷酸化或糖基化的蛋白质,不同的修饰可给出不同的分离的蛋白,它们可独立纯化。

[0083] “分离的多核苷酸”意为一种核酸(例如,DNA),其不含在本发明的核酸分子所源自的有机体的天然出现的基因组中位于该基因侧翼的基因。该术语因此包括,举例而言,重组DNA,其被并入载体内、并入自主复制质粒或病毒内、或并入原核生物或真核生物的基因组DNA内,或作为独立于其它序列的单独分子而存在(举例而言,通过PCR或限制性内切酶消化

制备的cDNA或基因组或cDNA片段)。此外,该术语包括从DNA分子转录的RNA分子,以及作为编码附加多肽序列的杂交基因的一部分的重组DNA。

[0084] “分离的多肽”意为已经与其天然伴随组分分离的本发明的多肽。典型地,以重量计,当多肽的至少60%不含其天然关联的蛋白质和天然出现的有机分子时,则该多肽是分离的。优选该制剂含有至少75重量%、更优选至少90重量%、最优选99重量%的本发明的多肽。可通过例如从天然来源提取、编码多肽的重组核酸的表达、或化学合成蛋白质来获得本发明的分离的多肽。可通过任何适宜的方法例如柱层析、聚丙烯酰胺凝胶电泳或HPLC分析来测量纯度。

[0085] 本文中提供的范围理解为该范围内所有值的略写。例如,1至50的范围理解为包括来自1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49或50所组成的组的任意数字、数字组合或子范围。

[0086] “降低”意为至少10%、25%、50%、75%或100%的负向改变。

[0087] “参考”意为标准或对照条件。

[0088] “参考序列”是用作序列比对基础的定义序列。参考序列可以是特定序列的一部分或整体;例如,全长度cDNA或基因序列的链段、或完整的cDNA或基因序列。对于多肽,参考多肽序列的长度将通常为至少约16个氨基酸,优选至少约20个氨基酸,更优选至少约25个氨基酸,且甚至更优选约35个氨基酸、约50个氨基酸或约100个氨基酸。对于核酸,参考核酸序列的长度将通常为至少约50个核苷酸,优选至少约60个核苷酸,更优选至少约75个核苷酸,且甚至更优选约100个核苷酸或约300个核苷酸或与之接近或之间的任何整数。

[0089] 可用于本发明方法的核酸分子包括编码本发明的多肽或其片段的任何核酸分子。此类核酸分子无须与内源性核酸序列100%一致,但典型将会展现基本一致性。具有与内源性序列“基本一致性”的多核苷酸典型能与双链核酸分子的至少一条链杂交。可用于本发明方法的核酸分子包括编码本发明的多肽或其片段的任何核酸分子。此类核酸分子无须与内源性核酸序列100%一致,但典型将会展现基本一致性。具有与内源性序列“基本一致性”的多核苷酸典型能与双链核酸分子的至少一条链杂交。“杂交”意为在互补多核苷酸序列(例如,本文所述的基因)或其部分在多种严格条件下配对以形成双链分子。(见,例如,Wahl, G.M. and S.L.Berger (1987) *Methods Enzymol.* 152:399; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507)。

[0090] 举例而言,严格的盐浓度一般为少于约750mM NaCl和75mM柠檬酸三钠,优选少于约500mM NaCl和50mM柠檬酸三钠,且更优选少于约250mM NaCl和25mM柠檬酸三钠。低严格度杂交可在不存在有机溶剂如甲酰胺的情况下获得,而高严格度杂交可在存在至少约35%甲酰胺且更优选至少约50%甲酰胺的情况下获得。严格的温度条件一般将包括至少约30°C,更优选至少约37°C,且最优选至少约42°C的温度。可变的附加因素,如杂交时间、洗涤剂如十二烷基硫酸钠(SDS)的浓度、以及包含或不饱和载体DNA,是该领域技术人员所周知的。通过按需要组合这些多种条件来实施各种水平的严格度。在优选的实施方式中,杂交将出现在30°C的750mM NaCl、75mM柠檬酸三钠和1%SDS中。在更优选的实施方式中,杂交将出现在37°C的500mM NaCl、50mM柠檬酸三钠、1%SDS、35%甲酰胺和100μg/ml变性鲑鱼精DNA(ssDNA)中。在最优选的实施方式中,杂交将出现在42°C的250mM NaCl、25mM柠檬酸三钠、

1% SDS、50% 甲酰胺和200 μ g/ml ssDNA中。对这些条件的有用改变对于该领域技术人员是显而易见的。

[0091] 对于大多数应用,杂交后的洗涤步骤的严格度也将改变。可通过盐浓度和温度来界定洗涤严格度条件。如上,可通过降低盐浓度或增加温度来增加洗涤严格度。举例而言,洗涤步骤的严格的盐浓度优选为少于约30mM NaCl和3mM柠檬酸三钠,且最优选少于约15mM NaCl和1.5mM柠檬酸三钠。用于洗涤步骤的严格的温度条件一般将包括至少约25 $^{\circ}$ C,更优选至少约42 $^{\circ}$ C,且甚至更优选至少约68 $^{\circ}$ C的温度。在优选的实施方式中,洗涤步骤将出现在25 $^{\circ}$ C的30mM NaCl、3mM柠檬酸三钠和0.1% SDS中。在更优选的实施方式中,洗涤步骤将出现在42 $^{\circ}$ C的15mM NaCl、1.5mM柠檬酸三钠和0.1% SDS中。在更优选的实施方式中,洗涤步骤将出现在68 $^{\circ}$ C的15mM NaCl、1.5mM柠檬酸三钠和0.1% SDS中。对这些条件的额外改变对于该领域技术人员是显而易见的。杂交技术是该领域技术人员所周知的,且揭示在例如Benton and Davis (Science 196:180,1977); Grunstein and Hogness (Proc.Natl.Acad.Sci.,USA 72:3961,1975);《分子生物学现代技术》(Ausubel et al.(Current Protocols in Molecular Biology,Wiley Interscience,New York,2001));《分子克隆技术指南》(Berger and Kimmel(Guide to Molecular Cloning Techniques,1987,Academic Press,New York));以及《分子克隆:实验室手册》(Sambrook et al.,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory Press,New York)中。

[0092] “基本相同”意为多肽或核酸分子表现出与参考氨基酸序列(例如,任何一种本文所述的氨基酸序列)或核酸序列(例如,任何一种本文所述的核酸序列)的至少50%一致性。优选地,此序列在氨基酸水平或核酸水平上与用于比较的序列的一致性为至少60%,更优选80%或85%,且更优选90%、95%或甚至99%。

[0093] 通常使用序列分析软件(例如,威斯康星大学生物技术中心的遗传学计算机公司(Genetics Computer Group,University of Wisconsin Biotechnology Center,1710University Avenue,Madison,Wis.53705)的序列分析软件包BLAST、BESTFIT、GAP、或PILEUP/PRETTYBOX程序)测量序列一致性。该软件通过设定多种替换、删除、及/或其它修饰的同源性程度来匹配一致或相似的序列。保守替换典型包括下述各组的组内替换:甘氨酸、丙氨酸;缬氨酸、异亮氨酸、亮氨酸;天冬氨酸、谷氨酸、天冬酰胺、谷氨酰胺;丝氨酸、苏氨酸;赖氨酸、精氨酸;以及苯丙氨酸、酪氨酸。在例示性的测定一致性程度的途径中,可使用BLAST程序,其中介于 e^{-3} 与 e^{-100} 之间的可能性得分指示密切相关的序列。

[0094] “体细胞核转移”或“SCNT”意为供体细胞核从体细胞转移到去核卵母细胞内。该过程可用于生殖性或治疗性克隆中并且可通过体细胞与去核卵母细胞的融合、将细胞核注入去核卵母细胞内完成或通过任何其他方法完成。

[0095] 体细胞的细胞核提供遗传信息,而卵母细胞提供胚胎发育所必需的营养物质和其它产生能量的物质。一旦融合发生,细胞就是多能的,并最终发育为囊胚,此时内部细胞团被隔离。

[0096] 如本文中所用,术语“核转移”是指一种基因操纵技术,其允许通过将去核卵母细胞与核遗传物质或体细胞核人工组合而获取相同的特征和质量。在一些实施方案中,核转移过程是,其中来自供体的体细胞核或核遗传物质被转移到去核卵子或卵母细胞(其核/原核已经被移除的卵子或卵母细胞)内。供体细胞核可来自体细胞。

[0097] 术语“核遗传物质”是指在细胞核内发现的结构和/或分子,其包含编码关于个体的信息的多核苷酸(例如,DNA)。核遗传物质包括染色体和染色质。该术语也称为通过细胞分裂诸如亲代细胞分类为子代细胞产生的核遗传物质(例如,染色体)。核遗传物质不包括线粒体DNA。

[0098] 术语“SCNT胚胎”是指已经与体细胞的核或核遗传物质融合的去核卵母细胞的细胞或其全能后代。SCNT胚胎可发育为囊胚并且在着床后发育为成活的后代。SCNT胚胎可以为1细胞胚胎、2细胞胚胎、4细胞胚胎或在变成囊胚之前任意阶段的胚胎。

[0099] 术语“供体人类细胞”或“供体人类体细胞”是指被转移到作为核受体或接纳者的卵母细胞内的人类细胞的体细胞或核。

[0100] 术语“体细胞”是指一种植物或动物细胞,其并非生殖性细胞或生殖性细胞前体。在一些实施方案中,已分化的细胞并非生殖细胞。体细胞与多能细胞或全能细胞无关。在一些实施方案中,体细胞为“非胚胎性体细胞”,其意为胚胎中不存在或不是从胚胎获得并且不是此类细胞在体外增殖所致。在一些实施方案中,体细胞为“成体体细胞”,其意为存在于除胚胎或胎儿外的有机体中或从该有机体获得或者是此类细胞在体外增殖所致。

[0101] 如本文中所示,术语“卵母细胞”是指已经达到减数分裂中期II的成熟卵母细胞。卵母细胞也用来描述牵涉到生殖中的雌配子或生殖细胞,并且也常常成为卵子。成熟的卵母细胞具有一套母系染色体(人类和灵长动物为23个,X)并且止于中期II。

[0102] “杂交卵母细胞”是指具有来自第一个人类卵母细胞的细胞质(称为“接纳者”)但不具有接纳者卵母细胞的核遗传物质的去核卵母细胞;其具有来自另一个人类细胞(称为“供体”)的核遗传物质。在一些实施方案中,杂交卵母细胞也可包含线粒体DNA(mtDNA),该线粒体DNA并非来自接纳者卵母细胞而是来自供体细胞(其可为作为核遗传物质的相同供体细胞,或来自不同供体例如来自供体卵母细胞)。

[0103] 如本文中所示,术语“去核卵母细胞”是指其核已经被移除的人类卵母细胞。

[0104] 如本文中所示,术语“去核”是指借以移除核遗传物质而仅留下细胞质的过程。当应用于卵子时,去核是指移除不为核膜环绕的母系染色体。术语“去核卵母细胞”是指其中核物质或核被移除的卵母细胞。

[0105] 如本文中所示,“接纳者人类卵母细胞”是指在移除其原始细胞核之后接受来自人类核供体细胞的细胞核的人类卵母细胞。

[0106] 如本文中所示,术语“融合”是指核供体细胞与接纳者卵母细胞的脂质膜的合并。例如,脂质膜可以为细胞的浆膜或核膜。融合可在当将核供体细胞和接纳者卵母细胞放置为彼此相邻时或当将核供体细胞放置在接纳者卵母细胞的卵周隙时,将电刺激施加到它们之间后发生。

[0107] 如本文中所示,术语“成活的后代”意为可在子宫外存活的动物。优选地,它是可存活一秒、一分钟、一天、一周、一个月、六个月或超过一年的动物。该动物可能不需要子宫内的存活环境。

[0108] 术语“产前”是指在出生之前存在或发生。类似地,术语“产后”是在出生之后存在或发生。

[0109] 如本文中所示,术语“囊胚”是指胎盘哺乳动物的大约30至150个细胞的植入前胚胎(对于小鼠为受精后大约3天,对于人类为受精后大约5天)。囊胚阶段在桑椹胚阶段之后,

并且可通过其独特的形貌予以区分。囊胚由以一层细胞形成的球体(滋养外胚层)、流体填充的腔体(卵裂腔或囊胚腔体)和内部上的一簇细胞(内细胞团,或ICM)组成。ICM由未分化细胞组成,如果将囊胚植入子宫中,则ICM产生将会变为胎儿的细胞。这些相同的ICM细胞如果以簇的形式生长,则可产生胚胎干细胞系。在植入的时间点,小鼠囊胚由大约70个滋养层细胞和30个ICM细胞形成。

[0110] 如本文中所示,术语“囊胚”是指胚胎发育的早期阶段,由包封被称为卵裂腔的流体填充腔体的中空细胞球组成。术语囊胚泡有时与囊胚可互换地使用。

[0111] 通篇所用的术语“卵裂球”指代从植入前胚胎获得的至少一个卵裂球(例如,1个、2个、3个、4个等)。术语“两个或更多个卵裂球的簇”与“源自卵裂球的外长物”可互换地使用,是指在卵裂球的体外培养期间生成的细胞。例如,在从SCNT胚胎获得卵裂球并初步培养之后,它通常分裂至少一次以产生两个或更多个卵裂球的簇(也称为源自卵裂球的外长物)。该簇可与胚胎或胎儿细胞进一步培养。最终,源自卵裂球的外长物将继续分裂。在该培养方法进程中,将会从这些结构将发育出ES细胞、全能干(TS)细胞和部分已分化的细胞类型。

[0112] 如本文中所示,术语“克隆的(或克隆)”是指一种基因操纵技术,其用于制备新的个体单元以具有与另一个体单元相同的基因集。在本发明中,如本文所用,术语“克隆的”是指细胞、胚胎细胞、胎儿细胞和/或动物细胞具有基本上与另一细胞、胚胎细胞、胎儿细胞、已分化细胞和/或动物细胞的核DNA序列相似或相同的核DNA序列。术语“基本上相似”和“相同”在本文中描述。克隆的SCNT胚胎可以产生自一个核转移,或者,克隆的SCNT胚胎可以产生自包括至少一个再次克隆步骤的克隆过程。

[0113] 如本文中所示,术语“转基因有机体”是指一种有机体,已经将来自另一个有机体的遗传物质实验性地转移到其中,使得宿主在其染色体组成中获得被转移基因的遗传特性。

[0114] 如本文中所示,参考本文中所公开的SCNT胚胎的术语“植入”是指使用本文所述的SCNT胚胎使代孕雌性动物受孕。这一技术是本领域技术人员周知的。参见,例如,《奶牛胚胎移植》(Seidel and Elsdon,1997,Embryo Transfer in Dairy Cattle,W.D.Hoard&Sons,Co.,Hoards Dairyman)。可使胚胎在子宫内发育,或者,可以在分娩之前将胎儿从子宫环境中移除。

[0115] “受试者”意为哺乳动物,包括但不限于,人类或非人类哺乳动物,诸如农业意义的哺乳动物(例如,牛、马、羊、猪)、宠物(例如,狗、猫)或稀有或濒危哺乳动物(例如,熊猫)。

[0116] 如本文中所示,“治疗”(动词、动名词、名称)等指的是减轻或缓解与其相关的病变和/或症状。应知晓,尽管未排除,但治疗病症或症状并不需要完全消除与该病症或症状相关的病症、症状或症候。

[0117] 除非明确指出或从语境中明显可见,否则如本文中所示,术语“或”理解为包含的。除非明确指出或从语境中明显可见,否则如本文中所示,术语“一”和“该”理解为单数或复数(即,至少一个)。举例而言,“元件”意为一个元件或超过一个元件。

[0118] 除非明确指出或从语境中明显可见,否则本文中使用的术语“约”理解为处于该领域正常公差范围内,例如,处于均值的2标准偏差内。“约”可理解为处于所指出值的10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%或0.01%内。除非从语境中明确排除,否则本文中提供的所有数值均以术语“约”修饰。

[0119] 本文中,对任何变量定义中一系列化学基团的描述包括该变量作为任何单一基团

或作为所列基团的组合的定义。对本文中变量或方面的实施方案的描述包括该实施方案作为任何单一输送方案或作为与任何其它实施方案或其部分的组合。

[0120] 本文中提供的任何组合物或方法可与本文中提供的一种或多种任何其它组合物和方法组合。

附图说明

[0121] 图1A至1F显示Xist KO供体细胞和Kdm4d mRNA注射的合用并不完全恢复SCNT胚胎的发育潜能。

[0122] 图1A包括使用抗H3K27me3抗体、抗Cdx2抗体、抗Oct4抗体和DAPI染色的IVF囊胚和SCNT囊胚的代表性图像。箭头指示点状的H3K27me3信号,表示异位失活的X染色体。注意,无论是否注射Kdm4d mRNA,都观察到了异位XCI。比例尺为50 μ m。

[0123] 图1B提供柱状图,显示IVF囊胚和SCNT囊胚中具有或不具有点状的H3K27me3信号(表示失活的X染色体)的细胞的比率。每一栏表示单个囊胚。

[0124] 图1C提供柱状图,显示通过在E19.5进行剖腹产检查的SCNT胚胎的幼崽生成率。注意,使用Xist KO供体细胞与Kdm4d mRNA注射的组合额外地改善了使用卵丘细胞、支持细胞(Sertoli cell)和MEF细胞作为供体的SCNT胚胎的足月率。

[0125] 图1D显示了通过使用与Kdm4d mRNA注射联合的Xist KO支持细胞的SCNT衍生的成年雄性小鼠的图像,及其通过与野生型雌性自然交配生成的幼崽的图像。

[0126] 图1E提供箱型图,显示通过在E19.5进行剖腹产检查的胎盘重量。须表示最大值和最小值。***p<0.001.ns,不显著。

[0127] 图1F提供使用高碘酸希夫氏(PAS;右)染色的足月胎胚的组织学切片的代表性图像。注意,无论供体细胞中的Xist等位基因为何种基因型,在SCNT胎盘中,PAS阳性的海绵状滋养层都已经侵入到迷路层中。比例尺为1mm。

[0128] 图2A至2C显示SCNT胚胎的着床后发育停止。

[0129] 图2A提供柱状图,显示在所指示的时间点,使用与Kdm4d mRNA注射联合的Xist KO MEF细胞生成的SCNT胚胎的发育率。

[0130] 图2B为在E4.5采集的SCNT胚胎的图像。

[0131] 图2C为在E104.5采集的SCNT胚胎的图像。注意,SCNT胚胎在每个阶段的胚胎/身体尺寸上表现出较大的差异。比例尺为100 μ m(图2B)和1mm(图2C)。

[0132] 图3A至3D显示SCNT囊胚中DNA甲基化的广泛重新编程。

[0133] 图3A为实验方法的示意图。使用通过IVF或SCNT(Xist KO供体与Kdm4d注射的组合)生成的囊胚进行全基因组亚硫酸氢盐测序(WGBS)和RNA-seq。

[0134] 图3B包括箱型图,其比较SCNT囊胚和IVF囊胚的基因组中全部被覆盖的CpG的DNA甲基化水平,以及MEF、合子、精子和卵母细胞中的DNA甲基化水平。箱中的出现指示中值,而叉号表示均值。须表示第2.5个百分位和第97.5个百分位。Sp+Oo表示精子和卵母细胞的平均值。MEF、精子和卵母细胞的WGBS数据集从GSE56151和GSE56697获得。

[0135] 图3C为比较每个样品之间的DNA甲基化水平的图。注意,重度甲基化的供体MEF细胞基因组通过SCNT全局重新编程,导致与IVF囊胚类似的DNA甲基化谱。

[0136] 图3D为比较IVF囊胚与SCNT囊胚的基因表达谱的散点图。SCNT胚胎中的上调基因

(n=37 (深色簇); 倍变(FC) >3.0) 和下调基因 (n=55, 浅色簇; FC>3.0)。

[0137] 图4A和4B显示, SCNT囊胚和IVF囊胚具有类似的DNA甲基化组和转录组。

[0138] 图4A提供柱状图, 比较在包括IVF囊胚和SCNT囊胚中的重复序列在内的各种基因组特征处的平均甲基化水平。

[0139] 图4B包括比较IVF囊胚和SCNT囊胚的生物学复本的转录组的散点图。

[0140] 图5A至5H显示SCNT囊胚中差异性甲基化区域 (DMR) 的鉴定和表征。

[0141] 图5A显示箱型图, 显示SCNT囊胚和IVF囊胚在高DMR和低DMR的DNA甲基化水平。箱中的出现指示中值, 而叉号表示均值。也指示了DMR的数目。

[0142] 图5B包括比较高DMR与低DMR长度的箱型图。

[0143] 图5C为显示基因组中高DMR和低DMR分布的饼形图。

[0144] 图5D为所指示的样品在低DMR的平均DNA甲基化水平与它们的侧翼区比较的图。

[0145] 图5E为显示IVF囊胚和SCNT囊胚在低DMR的父系 (Pat) 和母系 (Mat) 等位基因特异性DNA甲基化水平与它们的侧翼区比较的图。

[0146] 图5F为显示所指示发育阶段的IVF囊胚和SCNT囊胚在低DMR的父系 (Pat) 和母系 (Mat) 等位基因特异性DNA甲基化水平与它们的侧翼区比较的图。

[0147] 图5G为所指示的样品在高DMR的平均DNA甲基化水平与它们的侧翼区比较的图。

[0148] 图5H为所指示的样品在高DMR的平均DNA甲基化水平与它们的侧翼区比较的图。所使用的数据集来自GSE11034。

[0149] 图6A至6D提供SCNT囊胚中的低DMR和高DMR特征。

[0150] 图6A为高DMR和低DMR的代表性基因组浏览器视图。

[0151] 图6B为代表性基因组浏览器视图, 显示与IVF囊胚中的那些重叠的卵母细胞中的甲基化峰。

[0152] 图6C为高DMR相关基因的基因本体论分析。

[0153] 图6D包含峰图, 显示PGC发育过程中在高DMR的均值甲基化水平 (5mC) 和羟甲基化水平 (5hmC)。

[0154] 图7A至图7D显示SCNT囊胚中的H3K27me3依赖性印记的损失。

[0155] 图7A提供柱状图, 显示SCNT囊胚中H3K27me3印记基因的相对基因表达水平。显示了以可靠检测水平在IVF囊胚中表达的26个基因 (每百万个映射片段中外显子的每千碱基片段数 (FPKM) >1)。IVF囊胚的表达水平设定为1。通过比较将SCNT囊胚与IVF囊胚中的表达变化, 将基因归类为上调、下调和未改变 (FC>1.5)。

[0156] 图7B提供柱状图, 显示IVF囊胚和SCNT囊胚中H3K27me3印记基因的等位基因表达比率 (父系/母系)。显示了在26个被表达的基因 (FPKM>1) 中, 每个样品中均有17个基因具有>10SNP读数。星号表示100%偏向父系等位基因。注意, 在SCNT囊胚中, 全部17个基因均失去了它们的父系等位基因偏向。

[0157] 图7C显示两个代表性H3K27me3印记基因的H3K27me3 ChIP-seq信号的代表性基因组浏览器视图。

[0158] 图7D显示各种细胞类型 (卵母细胞、精子、MEF、ESC) 和组织在76个H3K27me3印记基因处的平均H3K27me3 ChIP-seq强度与3Mb侧翼区比较。

[0159] 图8A至图8F例示已知的126个印记基因的印记状态和它们的已知ICR。

[0160] 图8A提供柱状图,显示SCNT囊胚中23个已知印记控制区(ICR)的相对DNA甲基化水平。IVF囊胚的甲基化水平设定为1。虚线指示IVF囊胚甲基化水平的50%。注意,在SCNT囊胚中,23个ICR中的21个保持了IVF甲基化水平的至少50%,但Slc38a4 ICR和Snrpn ICR(标记为红色)显示了低于IVF水平的50%。

[0161] 图8B提供柱状图,显示在20个具有足够等位基因特异性甲基化信息的ICR处的DNA甲基化的等位基因偏向(在IVF囊胚和SCNT囊胚两者中的两个等位基因中所检测的CpG均>5)。注意,在SCNT囊胚中,全部20个ICR均保持了等位基因偏向的DNA甲基化。

[0162] 图8C提供柱状图,显示SCNT囊胚中已知印记基因的相对基因表达水平。显示了在IVF囊胚中可靠检测的45个印记基因(FPKM>1)。IVF囊胚的表达水平设定为1。基于它们在SCNT囊胚中的表达水平与在IVF胚胎中的表达水平的比较,将基因分类为上调、下调和未改变(FC>1.5)。

[0163] 图8D提供柱状图,显示IVF囊胚和SCNT囊胚中已知印记基因的等位基因表达比率(母系/父系)。显示了6个母系表达的基因(MEG;母系/父系>2.0),其以可靠检测水平在IVF囊胚中表达,具有足够的SNP跟踪读取(在每个样品中,FPKM>1,均值SNP读数>10)。星号表示100%偏向母系等位基因。注意,在SCNT囊胚中,全部6个MEG均保持了它们的母系等位基因偏向。

[0164] 图8E提供柱状图,显示IVF囊胚和SCNT囊胚中已知印记基因的等位基因表达比率(父系/母系)。显示了13个父系表达的基因(PEG;父系/母系>2.0),其以可靠检测水平在IVF囊胚中表达,具有足够的SNP跟踪读取(在每个样品中,FPKM>1,均值SNP读数>10)。星号表示100%偏向父系等位基因。箭头指示在SCNT囊胚中丢失父系偏向表达的基因。Slc38a4、Sfmbt2、Phf17和Gab1是H3K27me3依赖性的印记基因。

[0165] 图8F呈递了非典型印记基因的H3K27me3 ChIP-seq信号的代表性基因组浏览器视图。

具体实施方式

[0166] 本发明提供用于改善克隆效率的方法。在具体实施方案中,本发明提供用于改善体细胞核转移效率的方法,其牵涉Xist敲除的供体细胞中的Kdm4d过度表达。

[0167] 本发明至少部分地基于下述发现,Xist敲除的供体细胞与Kdm4d mRNA注射联合可改善体细胞核转移效率。这一联合方法在使用已分化的体细胞供体细胞进行的小鼠克隆中导致的有史以来报告的最高效率。然而,很多SCNT胚胎仍表现出着床后发育停止并且存活的胚胎具有异常的大胎盘,暗示一些重新编程缺陷仍持续存在。比较性甲基组和转录组分析表明,在SCNT囊胚胚胎中的异常DNA甲基化和H3K27me3依赖性印记的丢失,这可能是所观察到的发育缺陷的肇因。

[0168] H3K27me3是DNA甲基化依赖性的印记机制

[0169] 哺乳动物精子和卵母细胞具有不同的表观遗传特征并且以不通过的方式组织化。在受精之后,除了包括印记控制区(ICR)在内的某些基因座以外,最初截然不同的亲本表观基因组变成在很大程度上相同。亲本染色质如何变成相同以及ICR如何从这一重新编程中逃脱在很大程度上是未知的。这里,在小鼠合子和桑椹胚中表征了亲本等位基因特异性DNase I高敏位点(DHS),并且探究了等位基因DHS的底层表观遗传机制。DNA甲基组和H3K27me3 ChIP-seq数据集的综合分析表明,76个具有父系等位基因特异性DHS的基因(表

1) 没有进行DNA甲基化,但携带母系等位基因特异性H3K27me3。

[0170] 表1:H3K27me3依赖性印记基因

基因_名称	基因_chr	基因_始点	基因_终点
Rbp2	chr9	98390956	98410190
Runx1	chr16	92601711	92826311
Sfmbt2	chr2	10292078	10516880
Slc38a2	chr15	96517823	96530129
Slc38a4	chr15	96825254	96886387
Gramd1b	chr9	40105492	40263349
Bbx	chr16	50191957	50432502
Sox21	chr14	118632456	118636252
Mbnl2	chr14	120674891	120830920
Prdm11	chr2	92815063	92886301
1700067G17Rik	chr1	90912688	90918785
1700095B10Rik	chr5	113222312	113230721
Mir692-2b	chr4	125181992	125182101
Sh3gl3	chr7	89319728	89455927
Etv6	chr6	133985725	134220165
Tle3	chr9	61220173	61266304
Hunk	chr16	90386642	90499798
Gab1	chr8	83288333	83404378
Matn1	chr4	130500300	130511391
Chst1	chr2	92439864	92455409
Clic6	chr16	92498392	92541486
1700110K17Rik	chr9	40141057	40150922
Foxl1	chr8	123651585	123654544
Mir6241	chr14	118657855	118657958
Otog	chr7	53496357	53566804
1700017J07Rik	chr2	168803769	168804406
4930404H11Rik	chr12	72641594	72657120
Gm5086	chr13	98329955	98353949
Tshz2	chr2	169459146	169714004
Bmp7	chr2	172695189	172765794
G730013B05Rik	chr16	50526358	50559572
Rftn1	chr17	50132632	50329822
C430002E04Rik	chr3	41291603	41297121
Myoz2	chr3	122709124	122737905
Six3os1	chr17	86001272	86017736
Slc38a1	chr15	96401849	96473344
Rbms1	chr2	60590010	60801261

[0171]

	Flt1	chr5	148373772	148537564
	Sall3	chr18	81163113	81183317
	Otx2os1	chr14	49288963	49413023
	1700006F04Rik	chr14	120148449	120150786
	2300005B03Rik	chr15	74573269	74577117
	4931430N09Rik	chr6	118830176	118835561
	Gas7	chr11	67346500	67502494
	Phf17	chr3	41359656	41420786
	Igsf21	chr4	139582767	139802726
	Otx2	chr14	49277859	49282547
	Klhdc7a	chr4	139518088	139523941
	1700125H03Rik	chr8	70892358	70899609
	Lpar3	chr3	145883925	145949178
	Mir6239	chr14	118352964	118353069
	Epas1	chr17	87153204	87232750
	Slc6a1	chr6	114232629	114267519
	Cdh26	chr2	178165312	178222071
	1700025C18Rik	chr2	164904193	164916250
	Prox1	chr1	191945658	191994559
[0172]	1700121N20Rik	chr12	107680862	107685876
	Adamts2	chr11	50415587	50617551
	Gad1l	chr9	115818573	115985294
	Dnase2b	chr3	146244337	146278562
	Inhbb	chr1	121312042	121318825
	E2f3	chr13	29998444	30077932
	Ajap1	chr4	152747330	152856939
	BC049762	chr11	51067153	51076453
	Edn3	chr2	174586274	174609543
	Enc1	chr13	98011060	98022995
	4930465M20Rik	chr12	108961953	108973698
	9630028H03Rik	chr2	135406266	135408956
	Cd44	chr2	102651300	102741822
	Epgn	chr5	91456543	91464238
	Syt13	chr2	92755258	92796208
	Myb	chr10	20844736	20880790
	Lrig3	chr10	125403275	125452415
	Fam198b	chr3	79689852	79750200
	Smoc1	chr12	82127795	82287401
	1700084F23Rik	chr13	70142928	70167226

[0173] 有趣的是,这些基因在植入前胚胎中进行父系表达,而H3K27me3的异位移除诱导了母系等位基因表达。H3K27me3依赖性印记在胚胎细胞谱系中很大程度上丢失,但至少5个基因在胚胎外细胞谱系中保持了它们的印记。这5个基因包括所有先前鉴定的不依赖于DNA

甲基化的印记常染色体基因。母系H3K27me3是DNA甲基化依赖性的印记机制。在一个实施方案中,本发明的方法牵涉H3K27me3选择性甲基化酶的用途。

[0174] H3K27me3对于X染色体失活是重要的

[0175] 在包括啮齿动物在内的某些雌性正兽亚纲哺乳动物中,两条X染色体中的一条是失活的以实现基因量补偿。这一被称为X染色体失活(XCI)的现象提供了用于理解表观遗传沉默机制的优异模型。在发育过程中,XCI可以印记模式或随机模式发生。对于印记XCI,父系X染色体(Xp)在植入前发育期间被选择性失活。尽管印记XCI在胚胎外细胞谱系中得以保持,但它在晚期囊胚的内细胞团(ICM)中丢失。在植入前阶段,外胚层细胞经历随机XCI,导致Xp或母系X染色体(Xm)的沉默。先前的研究已经证明了Xist(一种X连锁长非编码RNA)在印记XCI和随机XCI中的关键角色。通过以顺式包覆X染色体并令其失活,Xist RNA参与到XCI中。

[0176] 基因组印记允许进行亲源特异性基因调控。为了在印记XCI期间选择性地沉默Xp,Xist基因以一种长期寻找但尚未鉴定的机制在Xm中被印记进行沉默。先前使用核转移途径进行的研究已经暗示,像常染色体印记基因一样,在卵子发生期间建立了Xist的基因组印记。在小鼠植入前胚胎和胚胎外细胞中,仅父系X染色体(Xp)失活。印记父系X染色体失活(XCI)的中心为长非编码RNA,即Xist,它从Xp表达并且以顺式作用来包覆并沉默整个Xp。为了实现Xp特异性失活,必须沉默母系Xist基因,但沉默基质仍尚不明朗。如本文所报告,在小鼠卵母细胞中,Xist基因座包覆有宽H3K27me3结构域,其通过植入前发育过程而持续存在。H3K27me3的异位移除诱导了母系Xist表达和母系XCI。因此,母系H3K27me3用作Xist的印记标记。

[0177] 在一些实施方案中,本发明的方法牵涉给药包含选择性H3K27me3去甲基酶抑制剂的组合物。

[0178] H3K9me3和SCNT

[0179] 供体体细胞中的组蛋白H3赖氨酸9三甲基化(H3K9me3)是SCNT重新编程的表观遗传屏障。在小鼠和人类中,供体细胞中的H3K9me3均防止合子基因组激活时相关区的转录激活并且导致SCNT胚胎在植入前阶段的发育停止。重要的是,通过过表达H3K9me3特异性去甲基酶Kdm4d来移除H3K9me3屏障,使得SCNT胚胎以类似于IVF的速率发育到囊胚阶段。因此,足月率的总体克隆效率增加了8至9倍。尽管在SCNT中使用Kdm4d导致了与IVF相当的着床率,但少于15%的着床SCNT胚胎发育到足月。此外,在注射Kdm4d的SCNT胚胎中仍然观察到了异常的大胎盘。这些结果暗示,H3K9me3重新编程屏障主要阻碍了植入前发育,而其它屏障影响着床后发育。

[0180] Xist对于小鼠SCNT胚胎的着床后发育是重要的。Xist从母系X染色体的异常表达在植入前胚胎中造成异位X染色体失活(XCI)和全局转录改变,导致SCNT胚胎的着床后发育失败。重要的是,由异位Xist表达造成的这一发育失败可通过使用Xist敲除(KO)体细胞作为供体细胞或通过将对Xist的小干扰RNA注射入1细胞雄性SCNT胚胎中而克服,造成足月率增加8至10倍。

[0181] 抑制性核酸

[0182] 抑制性核酸分子是那些抑制多肽或多核苷酸(例如,Xist多核苷酸)的表达或活性的寡核苷酸。此类寡核苷酸包括结合编码Xist多核苷酸的核酸分子(例如,反义分子、

siRNA、shRNA)的单链核酸分子和双链核酸分子(例如,DNA、RNA及其类似物)。

[0183] siRNA

[0184] 21个至25个核苷酸的短双链RNA在下调基因表达中有效(Zamore et al.,Cell 101:25-33;Elbashir et al.,Nature 411:494-498,2001,通过引用并入)。McCaffrey等人在体内证明了siRNA途径在哺乳动物中的有效性(Nature 418:38-39,2002)。

[0185] 根据靶标基因的序列,可将siRNA设计成令该基因失活。例如,此类siRNA可直接给药到受影响的组织或系统性给药。基因的核酸序列可用来设计小干扰RNA(siRNA)。例如,可使用21个至25个核苷酸的siRNA来减少Xist表达。

[0186] 本发明的抑制性核酸分子可用作双链RNA干扰(RNAi)介导的表达抑制。在一个实施方案中,体细胞中的Xist基因表达减少。RNAi是降低所关注的特异性蛋白的细胞表达的方法(在Tuschl,Chembiochem 2:239-245,2001;Sharp,Genes&Devel.15:485-490,2000;Hutvagner and Zamore,Curr.Opin.Genet.Devel.12:225-232,2002和Hannon,Nature 418:244-251,2002中回顾)。通过dsRNA的转染或通过siRNA的表达使用基于质粒的表达系统将siRNA引入细胞内越来越多地被用来在哺乳动物细胞中创建功能丢失表型。

[0187] 在本发明的一个实施方案中,制备了双链RNA(dsRNA)分子,其包括本发明核酸碱基低聚物的8至10个之间的接续核酸碱基。dsRNA可以是具有双链体结构的RNA的两条不同的链,或具有自身双链体的单个RNA链(小发夹(sh)RNA)。通常,dsRNA为大约21个或22个碱基对,但若需要,可以更短或更长(至多大约29个碱基对)。dsRNA可使用标准技术(例如,化学合成或体外转录)制备。试剂盒可从例如Ambion(Austin, Tex.)和Epicentre(Madison, Wis.)获得。在哺乳动物细胞中表达dsRNA的方法下各自通过引用并入本文的下列文献中有所描述:Brummelkamp et al.Science 296:550-553,2002;Paddison et al.Genes&Devel.16:948-958,2002;Paul et al.Nature Biotechnol.20:505-508,2002;Sui et al.Proc.Natl.Acad.Sci.USA 99:5515-5520,2002;Yu et al.Proc.Natl.Acad.Sci.USA 99:6047-6052,2002;Miyagishi et al.Nature Biotechnol.20:497-500,2002;和Lee et al.Nature Biotechnol.20:500-505 2002。

[0188] 小发夹RNA(shRNA)包含具有茎环结构的RNA序列。“茎环结构”是指具有二级结构的核酸,其包括已知或预计形成双链或双链体的核苷酸区(茎部分),该茎部分在一侧链接有主要为单链核苷酸的区(环部分)。本文中,术语“发夹”也用来指代茎环结构。此类结构是该领域中熟知的,并且该术语以其在该领域中已知者一致的意义使用。如该领域中已知的,二级结构不具有精确的碱基配对。因此,该术语可包括一处或多处碱基误配或凸起。另选地,碱基配对可以是精确的,即,不包括任何误配。多个茎环结构可通过链接基诸如例如核酸链接基、miRNA侧翼序列、其它分子或它们的一些组合彼此链接。

[0189] 如本文中所用,术语“小发夹RNA”包括传统茎环shRNA,其形成前体miRNA(pre-miRNA)。尽管在范围上可能存在一定变动,但传统茎环shRNA可包括在19至29个bp范围内的茎和在4至30个bp范围内的环。“shRNA”也包括嵌有微小RNA的shRNA(基于miRNA的shRNA),其中miRNA双链体的引导链和随从链并入现存(或天然)miRNA中或并入经修饰或合成(设计)的miRNA中。在一些实例中,前体miRNA分子可包括超过一个茎环结构。微小RNA是大约22个核苷酸长的内源性编码的RNA分子,其通常以高度组织特异性或发育阶段特异性的方式被表达,并且转录后地调控靶标基因。在植物和动物中已经鉴定了超过200种截然不同的

miRNA。据信,这些小调控性RNA通过两种占优的作用模式发挥重要的生物功能:(1)通过压制靶标mRNA的转译,以及(2)通过RNA干扰(RNAi),换言之,mRNA的裂解和降解。在后一种情况下,miRNA类似于小干扰RNA(siRNA)地发挥作用。因此,可基于现存miRNA基因的特征来设计并表达人工miRNA。

[0190] shRNA可从DNA载体表达以向几乎任何细胞类型提供持续的沉默和高良率递送。在一些实施方案中,载体是病毒载体。示例性的病毒载体包括逆转录病毒,包括慢病毒、腺病毒、杆状病毒和禽病毒载体,并且包括此类允许进行稳定的单拷贝基因组整合的病毒。可自其递送逆转录质粒载体的逆转录病毒包括但不限于,莫洛尼鼠白血病病毒、脾坏死病毒、劳斯氏肉瘤病毒、哈维氏肉瘤病毒、禽白血病病毒、长臂猿白血病病毒、人类免疫缺陷病毒、骨髓增生性肉瘤病毒和乳腺肿瘤病毒。逆转录病毒质粒载体可用来转导包装细胞系以形成生产细胞系。可被转染的包装细胞的示例包括但不限于,PE501、PA317、R-2、R-AM、PA12、T19-14x、VT-19-17-H2、RCRE、RCRIP、GP+E-86、GP+envAm12和DAN细胞系,如通过引用而以其整体并入本文的Miller, *Human Gene Therapy* 1:5-14 (1990) 中所述。载体可通过该领域中已知的任何手段转导包装细胞。生产细胞系生成感染性逆转录病毒载体颗粒,该载体颗粒包括编码DNA复制蛋白的多核苷酸。然后,此类逆转录病毒载体颗粒可用来在体外或体内转导真核细胞。被转导的真核细胞将会表达DNA复制蛋白。

[0191] 包括本发明的反义序列的催化性RNA分子或核糖酶可用来抑制体内核酸分子(例如,Xist)的表达。核糖酶序列包含在反义RNA内可赋予它们以RNA裂解活性,从而增加构造体的活性。靶标RNA特异性核糖酶的设计和用途在各自通过引用并入的Haseloff et al., *Nature* 334:585-591.1988和美国专利公布2003/0003469A1中有所描述。

[0192] 据此,本发明也提出催化性RNA分子,其包括位于结合臂中的反义RNA,该反义RNA具有8个至19个之间的接续核酸碱基。在本发明的优选实施方案中,催化性核酸分子在锤头或发夹基序中形成。此类锤头基序的示例在Rossi et al., *Aids Research and Human Retroviruses*, 8:183,1992中有所描述。发夹基序的示例在Hampel等人于1989年9月20日提交的“RNA Catalyst for Cleaving Specific RNA Sequences”(它是1988年9月20日提交的美国专利申请07/247,100的部份接续申请案)、Hampel and Tritz, *Biochemistry*, 28:4929,1989和Hampel et al., *Nucleic Acids Research*, 18:299,1990中有所描述。这些特异性基序不限制在本发明中,该领域技术人员应认识到,在本发明的酶核酸分子中,最重要的是它具有与靶标基因RNA区的一者或多者互补的特异性底物结合位点,以及它具有位于赋予该分子以RNA裂解活性的底物结合位点内或周围的核苷酸序列。

[0193] 基本上可采用任何用于将核酸构造体引入细胞内的方法。引入核酸的物理方法包括注射含有该构造体的溶液,使用覆盖有该构造体的颗粒轰击,将细胞、组织或有机体浸泡该核酸的溶液中,或在构造体的存在下对细胞膜进行电穿孔。包封在病毒颗粒中的病毒构造体可用来实施将表达构造体有效引入细胞内和被编码shRNA的转录。可使用该领域中已知的其它用于将核酸引入细胞内的方法,诸如脂质介导的载剂转运、化学介导的转运诸如磷酸钙等。因此,可将编码shRNA的核酸构造体与实施以下一种或多种活性的组分一起引入:增强细胞对RNA的摄取、促使双链体链的解链、稳定化解链的链、或以其它方式增加对靶标基因的抑制。

[0194] 对于在细胞内表达,可采用DNA载体,例如包含RNA聚合酶II或RNA聚合酶III启动

子的质粒载体。内源性miRNA的表达受RNA聚合酶II (Pol II) 启动子的控制,并且在一些情况下,与RNA聚合酶III启动子相比,shRNA更有效地由Pol II启动子驱动(Dickins et al., 2005, Nat. Genet. 39:914-921)。在一些实施方案中,shRNA的表达可能受诱导型启动子或条件性表达系统的控制,包括而不限于,II型RNA聚合酶启动子。在本发明的情境中可用的启动子的示例为四环素诱导型启动子(包括TRE-tight)、IPTG诱导型启动子、四环素反式激活因子系统、以及反四环素反式激活因子(rtTA)系统。也可使用构成性启动子,也可使用细胞或组织特异性启动子。很多启动子将无处不在,使得它们在所有细胞和组织类型中被表达。某些实施方案使用四环素响应启动子,它是体外和体内研究中最有效的条件性基因表达系统之一。见国际专利申请案PCT/US2003/030901(公布号WO 2004-029219A2)和Fewell et al., 2006, Drug Discovery Today 11:975-982关于诱导型shRNA的描述。

[0195] 多核苷酸的递送

[0196] 裸多核苷酸或其类似物能够进入哺乳动物细胞并且抑制所关注基因的表达。尽管如此,可能希望使用辅助将寡核苷酸或其它核酸碱基低聚物递送至细胞的制剂(见,例如,美国专利5,656,611、5,753,613、5,785,992、6,120,798、6,221,959、6,346,613和6,353,055,其各自通过引用并入本文)。

[0197] 寡核苷酸和其它核酸碱基低聚物

[0198] 至少两种类型的寡核苷酸通过RNase H诱导RNA的裂解,RNase H是具有磷酸二酯(PO)或硫代磷酸酯(PS)链结的多核苷酸。尽管2'-OMe-RNA序列表现出对于RNA靶标的高亲和性,但这些序列并非RNase H的底物。所希望的寡核苷酸是基于2'-修饰寡核苷酸的,其包含寡脱氧核糖核苷酸间隙,其中一些或全部核苷酸间链结被修饰成硫代磷酸酯以具有耐核酸酶性。甲基磷酸酯修饰的存在增加了寡核苷酸对于其靶标RNA的亲合性,并因此减少了IC₅₀。这一修饰也增加了被修饰寡核苷酸的耐核酸酶性。应理解,本发明的方法和试剂可与可能被研发的任何技术合用,包括共价闭合的多反义(CMAS)寡核苷酸(Moon et al., Biochem J. 346:295-303, 2000; PCT公布WO 00/61595)、带式反义(RiAS)寡核苷酸(Moon et al., J. Biol. Chem. 275:4647-4653, 2000; PCT公布WO 00/61595)和大环状反义寡核苷酸(美国专利申请案公布US 2002/0168631 A1)。

[0199] 如该领域中已知的,核苷是核酸碱基-糖组合。核苷的碱基部分正常情况下为杂环碱基。两类最常见的此类杂环碱基是嘌呤和嘧啶。核苷酸是还包括共价链接至核苷的糖部分的磷酸酯基团的核苷。对于那些包括呋喃戊糖基糖的核苷,该磷酸酯基团可链接至糖的2'、3'或5'羟基部分。在形成寡核苷酸时,磷酸酯基团将相邻核苷彼此共价链接以形成线性聚合的化合物。这一线性聚合结构的各端可转而接合以形成环状结构;开放的线性结构通常是优选的。在寡核苷酸结构中,磷酸酯基团一般称为形成该寡核苷酸的主链。RNA和DNA的正常链结或主链为3'至5'磷酸二酯链结。

[0200] 可用于本发明的优选核酸碱基低聚物的具体示例包括,包含经修饰的主链或非天然核苷间链结的寡核苷酸。如本说明书中限定的,具有经修饰的主链的核酸碱基低聚物包括那些磷原子保留在主链中的和那些主链中不具有磷原子的。出于本说明书的目的,在其核苷间主链中不具有磷原子的经修饰的寡核苷酸也被视为核酸碱基低聚物。

[0201] 就有经修饰的寡核苷酸主链的核酸碱基低聚物包括,例如,硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基-磷酸三酯、甲基和其它烷基的磷酸酯(包括3'-

亚烷基磷酸酯和手性磷酸酯)、亚膦酸酯和、磷酰胺化物(包括3'-氨基磷酰胺化物和氨基烷基磷酰胺化物)、硫羰基磷酰胺化物、硫羰基烷基磷酸酯、硫羰基烷基磷酸三酯、具有正常3'-5'链结的硼磷酸酯、这些的2'-5'链接类似物和那些具有反极性的,其中相邻的核苷单元对被链接为3'-5'与5'-3'或者2'-5'与5'-2'。也包括各种盐、混合盐和游离酸形式。教导上述含磷链结的制备的代表性美国专利包括但不限于,各自通过引用并入本文的美国专利3,687,808、4,469,863、4,476,301、5,023,243、5,177,196、5,188,897、5,264,423、5,276,019、5,278,302、5,286,717、5,321,131、5,399,676、5,405,939、5,453,496、5,455,233、5,466,677、5,476,925、5,519,126、5,536,821、5,541,306、5,550,111、5,563,253、5,571,799、5,587,361和5,625,050。

[0202] 具有主链中不包括磷原子的经修饰寡核苷酸主链的核酸碱基低聚物具有通过以下形成的主链:短链烷基或环烷基核苷间链结、混合杂原子和烷基或环烷基核苷间链结、或一种或多种端链杂原子或杂环状核苷间链结。这些包括具有以下的那些:吗啉基链结(部分地从核苷的糖部分形成);硅氧烷主链;二硫醚、亚砷和砷主链;甲乙酰基(formacetyl)和三甲乙酰基主链;亚甲基甲乙酰基和三甲乙酰基主链;含有亚烷基的主链;氨基硫酸酯主链;亚甲基亚胺基和亚甲基胍基主链;磺酸酯和磺酰胺主链;酰胺主链;和其它具有混合的N、O、S和CH₂组分部分的主链。教导上述寡核苷酸的制备的代表性美国专利包括但不限于,各自通过引用并入本文的美国专利5,034,506、5,166,315、5,185,444、5,214,134、5,216,141、5,235,033、5,264,562、5,264,564、5,405,938、5,434,257、5,466,677、5,470,967、5,489,677、5,541,307、5,561,225、5,596,086、5,602,240、5,610,289、5,602,240、5,608,046、5,610,289、5,618,704、5,623,070、5,663,312、5,633,360、5,677,437和5,677,439。

[0203] 在其它核酸碱基低聚物中,糖和核苷间链结(即主链)两者都被替换为新颖基团。维持核酸碱基单元,以便与所列的Xist基因杂交。一种此类核酸碱基低聚物称为肽核酸(PNA)。在PNA化合物中,寡核苷酸的糖-主链被替换为含有酰胺的主链,尤其是氨基乙基甘氨酸主链。核酸碱基得以保留,并且直接或间接地结合至主链的酰胺部分中的氮杂氮原子。制备和使用这些核酸碱基低聚物的方法在例如P.E.Nielsen编撰的《肽核酸:实验指南和应用》(Peptide Nucleic Acids:Protocols and Applications,Horizon Press,Norfolk,United Kingdom,1999)中有所描述。教导PNA的制备的代表性美国专利包括但不限于,各自通过引用并入本文的美国专利5,539,082、5,714,331和5,719,262。关于PNA化合物的进一步教导可在Nielsen et al.,Science,1991,254,1497-1500中找到。

[0204] 在本发明的具体实施方案中,核酸碱基低聚物具有硫代磷酸酯主链和具有杂原子主链的核苷,尤其是-CH₂-NH-O-CH₂-、-CH₂-N(CH₃)-O-CH₂- (称为亚甲基(甲基亚胺基)或MMI主链)、-CH₂-O-N(CH₃)-CH₂-、-CH₂-N(CH₃)-N(CH₃)-CH₂-和-O-N(CH₃)-CH₂-CH₂-。在其它实施方案中,寡核苷酸具有美国专利5,034,506中描述的吗啉基主链结构。

[0205] 核酸碱基低聚物也可包含一个或多个取代的糖部分。核酸碱基低聚物包含位于2'位置的下列一者:OH、F、O-烷基、S-烷基、N-烷基、O-烯基、S-烯基、N-烯基、O-炔基、S-炔基、N-炔基或O-烷基-O-烷基,其中该烷基、烯基和炔基可以是取代或未取代的C₁至C₁₀烷基或C₂至C₁₀烯基和炔基。尤其优选O[(CH₂)_n]_nCH₃、O(CH₂)_nOCH₃、O(CH₂)_nNH₂、O(CH₂)_nCH₃、O(CH₂)_nONH₂和O(CH₂)_nON[(CH₂)_nCH₃]₂,其中n和m为1至约10。其它优选的核酸碱基低聚物包括位于2'位置的下列一者:C₁至C₁₀低级烷基、取代的低级烷基、烷芳基、芳烷基、O-烷芳基或O-芳烷

基;SH;SCH₃;OCN;Cl;Br;CN;CF₃;OCF₃;SOCH₃;SO₂CH₃;ONO₂;NO₂;NH₂;杂环烷基;杂环烷基芳基;氨基烷基胺基;聚烷基胺基;取代的硅基;RNA裂解基团;报告基团;嵌入剂;用于改善核酸碱基低聚物的药代动力学性质的基团;或用于改善核酸碱基低聚物的药效动力学性质的基团;以及其它具有类似性质的取代基。优选的修饰为2'-O-甲基和2'-甲氧基乙氧基(2'-O-CH₂CH₂OCH₃,也称为2'-O-(2-甲氧基乙基)或'-MOE)。另一个希望的修饰为2'-二甲基胺基氧乙氧基(即, O(CH₂)₂ON(CH₃)₂),也称为2'-DMAOE。其它修饰包括2'-氨基丙氧基(2'-OCH₂CH₂CH₂NH₂)和2'-氟(2'-F)。类似的修饰也可在寡核苷酸或其它核酸碱基低聚物的其它位置做出,尤其是在3'末端核苷酸上或2'-5'链接寡核苷酸中的糖的3'位置和5'末端核苷酸的5'位置。核酸碱基低聚物也可具有替代呋喃戊糖基糖的糖模拟物诸如环丁基部分。教导此类经修饰的糖结构的制备的代表性美国专利包括但不限于,各自通过引用而以其整体并入本文的美国专利4,981,957、5,118,800、5,319,080、5,359,044、5,393,878、5,446,137、5,466,786、5,514,785、5,519,134、5,567,811、5,576,427、5,591,722、5,597,909、5,610,300、5,627,053、5,639,873、5,646,265、5,658,873、5,670,633和5,700,920。

[0206] 核酸碱基帝爵也可包括核酸碱基修饰或替换。如本文中所示,“未修饰的”或“天然的”核酸碱基包括嘌呤碱基腺嘌呤(A)和鸟嘌呤(G),以及嘧啶碱基胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U)。修饰的核酸碱基包括其它合成的和天然的核酸碱基,诸如5-甲基胞嘧啶(5-me-C);5-羟甲基胞嘧啶;黄嘌呤;次黄嘌呤;2-氨基腺嘌呤;腺嘌呤和鸟嘌呤的6-甲基和其它烷基衍生物;腺嘌呤和鸟嘌呤的2-丙基和其它烷基衍生物;2-硫代尿嘧啶、2-硫代胸腺嘧啶和2-硫代胞嘧啶;5-卤尿嘧啶和5-卤胞嘧啶;5-丙炔基尿嘧啶和5-丙炔基胞嘧啶;6-偶氮尿嘧啶、6-偶氮胞嘧啶和6-偶氮胸腺嘧啶;5-尿嘧啶(假尿嘧啶);4-硫代尿嘧啶;8-卤、8-氨基、8-巯基、8-硫代烷基、8-羟基和其它8-取代的腺嘌呤和鸟嘌呤;5-卤(例如,5-溴)、5-三氟甲基和其它5-取代的尿嘧啶和胞嘧啶;7-甲基鸟嘌呤和7-甲基腺嘌呤;8-氮杂鸟嘌呤和8-氮杂腺嘌呤;7-去氮鸟嘌呤和7-去氮腺嘌呤;以及3-去氮鸟嘌呤和3-去氮腺嘌呤。其它核酸碱基包括美国专利3,687,808中公开的那些、Kroschwitz, J. I. 编撰的《高分子科学与工程简明百科全书》(The Concise Encyclopedia Of Polymer Science And Engineering, John Wiley & Sons, 1990) 第858-859页中公开的那些、Englisch et al., *Angewandte Chemie, International Edition*, 1991, 30, 613中公开的那些、以及Sanghvi, Y. S. 在《反义研究与应用》(Antisense Research and Applications, Crooke, S. T. and Lebleu, B., ed., CRC Press, 1993) 第15章第289-302页中公开的那些。这些核酸碱基中的某些尤其可用于增加本发明的反义寡核苷酸的结合亲和性。这些包括5-取代的嘧啶;6-氮杂嘧啶;以及N-2、N-6和O-6取代的嘌呤,包括2-氨基丙基腺嘌呤、5-丙炔基尿嘧啶和5-丙炔基胞嘧啶。已经显示,5-甲基胞嘧啶替换将核酸双链体稳定性增加了0.6至1.2°C (Sanghvi, Y. S., Crooke, S. T. and Lebleu, B., eds., *Antisense Research and Applications*, CRC Press, Boca Raton, 1993, pp. 276-278) 并且是所希望的碱基替换,当与2'-O-甲氧基乙基或2'-O-甲基糖修饰组合时尤甚。教导上述标注的经修饰的核酸碱基中某些的制备的代表性美国专利包括但不限于,各自通过引用并入本文的美国专利4,845,205、5,130,302、5,134,066、5,175,273、5,367,066、5,432,272、5,457,187、5,459,255、5,484,908、5,502,177、5,525,711、5,552,540、5,587,469、5,594,121、5,596,091、5,614,617、5,681,941和5,750,692。

[0207] 本发明的核酸碱基低聚物的另一修饰牵涉将核酸碱基低聚物化学链接至一个或

多个提升寡核苷酸的活性、细胞内分布或细胞摄取的部分或缀合物。此类部分包括但不限于,脂质部分诸如胆固醇部分(Letsinger et al., Proc. Natl. Acad. Sci. USA, 86:6553-6556, 1989)、胆酸(Manoharan et al., Bioorg. Med. Chem. Lett., 4:1053-1060, 1994)、硫醚例如己基-S-三苯甲基硫醇(Manoharan et al., Ann. N.Y. Acad. Sci., 660:306-309, 1992; Manoharan et al., Bioorg. Med. Chem. Lett., 3:2765-2770, 1993)、巯基胆固醇(Oberhauser et al., Nucl. Acids Res., 20:533-538:1992)、脂肪族链例如十二碳二醇或十一烷基残基(Saison-Behmoaras et al., EMBO J., 10:1111-1118, 1991; Kabanov et al., FEBS Lett., 259:327-330, 1990; Svinarchuk et al., Biochimie, 75:49-54, 1993)、磷脂质例如二-十六烷基-外消旋甘油或1,2-二-0-十六烷基-外消旋甘油-3-H-磷酸三乙铵(Manoharan et al., Tetrahedron Lett., 36:3651-3654, 1995; Shea et al., Nucl. Acids Res., 18:3777-3783, 1990)、聚胺或聚乙二醇链(Manoharan et al., Nucleosides & Nucleotides, 14:969-973, 1995)、金刚烷乙酸(Manoharan et al., Tetrahedron Lett., 36:3651-3654, 1995)、棕榈基部分(Mishra et al., Biochim. Biophys. Acta, 1264:229-237, 1995)或十八烷基胺或己基胺基-羰基-氧胆固醇部分(Crooke et al., J. Pharmacol. Exp. Ther., 277:923-937, 1996)。教导此类核酸碱基低聚物缀合物的制备的代表性美国专利包括:各自通过引用并入本文的美国专利4,587,044、4,605,735、4,667,025、4,762,779、4,789,737、4,824,941、4,828,979、4,835,263、4,876,335、4,904,582、4,948,882、4,958,013、5,082,830、5,109,124、5,112,963、5,118,802、5,138,045、5,214,136、5,218,105、5,245,022、5,254,469、5,258,506、5,262,536、5,272,250、5,292,873、5,317,098、5,371,241、5,391,723、5,414,077、5,416,203、5,451,463、5,486,603、5,510,475、5,512,439、5,512,667、5,514,785、5,525,465、5,541,313、5,545,730、5,552,538、5,565,552、5,567,810、5,574,142、5,578,717、5,578,718、5,580,731、5,585,481、5,587,371、5,591,584、5,595,726、5,597,696、5,599,923、5,599,928、5,608,046和5,688,941。

[0208] 本发明也包括作为嵌合化合物的核酸碱基低聚物。“嵌合的”核酸碱基低聚物是核酸碱基低聚物,尤其是寡核苷酸,其含有两个或更多个化学式截然不同的区,每个区由至少一种单体单元构成,即,在寡核苷酸的情况下由核苷酸构成。这些核酸碱基低聚物典型含有至少一个区,在该区处,核酸碱基低聚物被修饰以赋予该核酸碱基低聚物以增加的耐核酸酶降解性、增加的细胞摄取和/或增加的对靶标核酸的结合亲和性。核酸碱基低聚物的附加区可用作能够裂解RNA:DNA杂交体或RNA:RNA杂交体的酶的底物。举例而言,RNase H是细胞核酸内切酶,其裂解RNA:DNA双链体的RNA链。因此,RNase H的激活导致RNA靶标的裂解,从而极大地提升了核酸碱基低聚物抑制基因表达的效率。因此,当使用嵌合核酸碱基低聚物时,使用较短的核酸碱基低聚物往往可获得与使用杂交至相同靶标区的硫代磷酸酯脱氧核糖核苷酸相当的结果。

[0209] 本发明的嵌合核酸碱基低聚物可形成为两种或更多种上述核酸碱基低聚物的复合结构。当此类核酸碱基低聚物为寡核苷酸时,在该领域中其也被称为杂交体或间隙体(gapmers)。教导此类杂交结构的制备的代表性美国专利包括但不限于,各自通过引用而以其整体并入本文的美国专利5,013,830、5,149,797、5,220,007、5,256,775、5,366,878、5,403,711、5,491,133、5,565,350、5,623,065、5,652,355、5,652,356和5,700,922。

[0210] 根据本发明使用的核酸碱基低聚物可通过周知的固相合成技术便利且常规地制

备。用于此类合成的设备由若干商家包括例如Applied Biosystems (Foster City, Calif.) 销售。可额外地或另选地采用该领域已知的用于此类合成的任何其它手段。使用类似的技术来制备寡核苷酸诸如硫代磷酸酯和烷基化衍生物是众所周知的。

[0211] 本发明的核酸碱基低聚物也可与其它分子、分子结构或化合物的混合物混合、缀合或以其它方式缔合或使用其它分子、分子结构或化合物的混合物包封,该其它分子、分子结构或化合物的混合物为,例如,用于辅助摄取、分配和/或吸收的脂质体、受体靶向的分子、口服制剂、直肠制剂、外用制剂或其它制剂。教导此类摄取、分配和/或吸收辅助制剂的制备的代表性美国专利包括但不限于,各自通过引用并入本文的美国专利5,108,921、5,354,844、5,416,016、5,459,127、5,521,291、5,543,158、5,547,932、5,583,020、5,591,721、4,426,330、4,534,899、5,013,556、5,108,921、5,213,804、5,227,170、5,264,221、5,356,633、5,395,619、5,416,016、5,417,978、5,462,854、5,469,854、5,512,295、5,527,528、5,534,259、5,543,152、5,556,948、5,580,575和5,595,756。

[0212] 基因组编辑以敲除Xist

[0213] 基因编辑是生物医学研究的一个主要焦点,其属于基础科学和临床科学的交叉学科。新颖“基因编辑”工具的研发提供了一下能力:操纵位于特定染色体基因座处的细胞DNA序列而不在该基因组其它位点引入突变。这一技术有效地令研究者能够在体外或体内操纵受试者细胞的基因组。在SCNT的情境中,使用CRISPR生成包含Xist中KO的细胞。

[0214] 在一个实施方案中,基因编辑牵涉将核酸内切酶(造成DNA在DNA分子内进行内部断裂的酶)靶向至基因组的特定位点,从而触发染色体双链断裂(DSB)在所选择位点处的形成。如果伴随着染色体断裂的引入而引入供体DNA分子(例如,通过质粒或寡核苷酸引入),则断裂的染色体与引入的DNA之间可能发生相互作用,如果两个序列共享同源性则尤其如此。在这一实例中,可发生被称为“基因靶向”的过程,在该过程中,染色体的DNA末端通过同源性重组(HR)侵入供体DNA的同源序列中。使用供体质粒序列作为HR的模板,可实施所关注基因的无缝敲除。重要的是,如果供体DNA分子包括位于靶标基因(例如,Xist)内的删除,则HR介导的DSB修复将会把供体序列引入染色体中,导致该删除被引入染色体基因座内。通过将核酸酶靶向至含有靶标基因的基因组位点,该概念为使用DSB形成来刺激HR并因而将功能性靶标基因替换为该基因的删除形式。HR路径的优点是,它具有在原先的野生型等位基因位置处无缝生成基因敲除的潜力。

[0215] 目前,基因组编辑工具使用双链断裂(DSB)诱导来增强细胞基因操纵。此类方法包括锌指核酸酶(ZFN;例如在美国专利6,534,261、6,607,882、6,746,838、6,794,136、6,824,978、6,866,997、6,933,113、6,979,539、7,013,219、7,030,215、7,220,719、7,241,573、7,241,574、7,585,849、7,595,376、6,903,185和6,479,626中以及美国专利公布20030232410和US2009020314中有所描述,其通过引用并入本文)、转录激活子样效应子核酸酶(TALEN;例如在美国专利8,440,431、8,440,432、8,450,471、8,586,363和8,697,853中以及美国专利公布20110145940、20120178131、20120178169、20120214228、20130122581、20140335592和20140335618中有所描述,其通过引用并入本文)、和CRISPR(成簇规律间隔短回文重复序列)/Cas9系统(例如在美国专利8,697,359、8,771,945、8,795,965、8,871,445、8,889,356、8,906,616、8,932,814、8,945,839、8,993,233和8,999,641中以及美国专利公布20140170753、20140227787、20140179006、20140189896、20140273231、20140242664、

20140273232、20150184139、20150203872、20150031134、20150079681、20150232882和20150247150中有所描述,其通过引用并入本文)。例如,ZFN DNA序列识别能力和特异性可能是不可预测的。类似地,TALEN和CRISPR/Cas9不仅在所希望的位点处裂解,而且经常也在其它“脱靶”位点处裂解。这些方法具有与脱靶双链断裂诱导和潜在的有害突变(包括与这些脱靶效应有关的indel、基因组重排和染色体重排)相关的显著问题。ZFN和TALEN一定要使用模块化序列特异性DNA结合蛋白来生成对于基因组中~18bp序列的特异性。

[0216] RNA引导的核酸酶介导的基因组编辑,基于2型CRISPR(成簇规则间隔短回文重复序列)/Cas(CRISPR结合的)系统,为改变基因组提供了一个有价值的途径。简而言之,Cas9(由单向导RNA(sgRNA)引导的核酸酶)结合至紧随前间区序列邻近基序(PAM)之后的靶向基因组基因座并生成双链断裂(DSB)。然后,通过导致插入/删除(indel)突变的非同源性末端接合(NHEJ),或通过需要外源性模板并且可在靶标基因座生成精确修饰的同源导向修复(HDR)来修复DSB(Mali et al., Science. 2013Feb 15; 339(6121): 823-6)。不同于将功能性或部分功能性的基因拷贝添加到患者细胞但保留该基因的最初功能失调拷贝的其它基因疗法,这一系统可移除缺陷。使用工程化核酸酶进行的基因矫正已经在组织培养细胞和罕见病啮齿动物模型中得以证明。

[0217] CRISPR已经被广泛用于有机体中,包括面包酵母(*S.cerevisiae*)、斑马鱼、线虫(*C.elegans*)、植物、小鼠和数种其它有机体。此外,已经对CRISPR做出修饰以制备可编程的转录因子,该转录因子允许科学家靶向并激活或沉默特定基因。现在可获得数万种向导RNA的文库。

[0218] 自2012年起,CRISPR/Cas系统已经被用于基因编辑(沉默、增强或改变特定基因),它甚至在真核生物如小鼠和灵长动物中也发挥作用。通过插入包含cas基因和特异性设计的CRISPR的质粒,可在任何所希望的位置切割有机体的基因组。

[0219] CRISPR重复序列的尺寸在24至48个碱基对范围内。它们往往显示某种二元对称性,暗指二级结构诸如发夹的形成,但不是真正的回文。重复序列通过类似长度的间隔序列分隔开来。一些CRISPR间隔序列准确地匹配来自质粒和噬菌体的序列,但一些间隔序列匹配原核生物基因组(自我靶向间隔序列)。可快速地添加新闻隔序列以应对噬菌体感染。

[0220] CRISPR相关(cas)基因往往与CRISPR重复序列-间隔序列阵列相关。截至2013年,已有超过四十个不同的Cas蛋白质家族被描述。这些蛋白质家族中,Cas1在不同的CRISPR/Cas系统中似乎无处不在。已经使用cas基因与重复序列结果的具体组合来定义8个CRISPR亚型(*Ecoli*、*Ypest*、*Nmeni*、*Dvulg*、*Tneap*、*Hmari*、*Apern*和*Mtube*),其中的一些与编码重复序列相关神秘蛋白(RAMP)的额外基因模块有关。单个基因组中可能出现超过一种CRISPR亚型。CRISPR/Cas亚型的零星分布暗示,该系统在微生物进化期间受到水平基因转移的影响。

[0221] 外源性DNA显然被由Cas基因编码的蛋白质加工为小元件(长度为约30个碱基对),然后,这些小元件以某种方式插入到邻近前导序列的CRISPR基因座中。来自CRISPR基因座的RNA被构成性的表达并且被Cas蛋白加工为小RNA,这些小RNA由单个的具有侧翼重复序列的外源性衍生序列元件构成。这些RNA引导其它Cas蛋白以RNA或DNA水平沉默外源性基因元件。证据表明了CRISPR亚型间的功能多样性。Cse(Cas亚型*Ecoli*)蛋白(称为大肠杆菌(*E.coli*)中的CasA-E)形成功能性复合物Cascade,其将CRISPR RNA转录物加工成Cascade保留的间隔序列-重复序列单元。在其它原核生物中,Cas6加工CRISPR转录物。有趣的是,在

大肠杆菌中进行基于CRISPR的噬菌体失活需要Cascade和Cas3,但不需要Cas1和Cas2。在强烈火球菌(*Pyrococcus furiosus*)和其它原核生物中发现的Cmr(Cas RAMP模块)与识小CRISPR RNA形成功能性复合物,该复合物识别并裂解互补性靶标RNA。RNA引导的CRISPR酶被分类为V型限制性内切酶。

[0222] 也见美国专利公布2014/0068797,其通过引用而以其整体并入本文。

[0223] Cas9

[0224] Cas9是一种核酸酶,一种专门切割DNA的酶,具有两个活性切割位点,双螺旋中每条链一个。研究团队证明,它们可以使一个或两个位点失效,同时保存Cas9的起始定位其靶标DNA的能力。Jinek et al. (2012)将tracrRNA和间隔序列RNA组合成“单向导RNA”分子,当与Cas9混合时,该分子可发现并切割正确的DNA靶标。已经有人提议,此类合成的向导RNA可能用于基因编辑(Jinek et al., Science. 2012 Aug 17; 337 (6096): 816-21)。

[0225] Cas9蛋白在致病菌和共生菌中高度富集。CRISPR/Cas介导的基因调控可能对内源性细菌基因的调控有所贡献,在细菌与真核宿主相互作用期间尤其如此。例如,新凶手弗朗西丝菌(*Francisella novicida*)的Cas蛋白Cas9使用一种独特的、小的、CRISPR/Cas相关的RNA(scaRNA)来阻遏编码细菌脂蛋白的内源性转录物,而细菌脂蛋白对于新凶手弗朗西丝菌减弱宿主响应并提升毒力至关重要。将Cas9 mRNA和sgRNA共同注射到生殖系细胞(合子)内生成了具有突变的小鼠。也设想了Cas9 DNA序列的递送。

[0226] gRNA

[0227] 作为RNA引导的蛋白质,Cas9需要短RNA来引导对DNA靶标的识别。尽管Cas9优先探究含有PAM序列NGG的DNA序列,但它可结合在这里而无需原间隔序列靶标。然而,Cas9-gRNA复合物需要与gRNA密切匹配以创建双链断裂。细菌中的CRISPR序列被表达在多个RNA中,然后经处理以创建RNA的引导链。因为真核细胞系统缺少处理CRISPR RNA所需的一些蛋白质,创建了合成性构造体gRNA以将Cas9靶向的RNA的基本碎片组合成以RNA聚合酶2I型启动子U6表达的单个RNA。合成gRNA的最小长度略微超过100bp,并且含有一个以在PAM序列NGG前与之紧邻的20个原间隔序列核苷酸为靶标的部分;gRNA不含PAM序列。

[0228] 在一种方法中,改变受试者的一个或多个细胞以使用CRISPR-Cas系统将Xist删除或灭活。Cas9可用来靶向Xist基因。一旦识别靶标,Cas9就将双链断裂引入Xist靶标基因中。在双链断裂位点处的同源导向修复(HDR)可允许插入灭活或删除形式的Xist序列。

[0229] 以下美国专利和专利公布通过引用并入本文:8,697,359、20140170753、20140179006、20140179770、20140186843、20140186958、20140189896、20140227787、20140242664、20140248702、20140256046、20140273230、20140273233、20140273234、20140295556、20140295557、20140310830、20140356956、20140356959、20140357530、20150020223、20150031132、20150031133、20150031134、20150044191、20150044192、20150045546、20150050699、20150056705、20150071898、20150071899、20150071903、20150079681、20150159172、20150165054、20150166980、和20150184139。

[0230] 治疗方法

[0231] 调节存在于印记控制区内的H3K27me3印记的药剂可用于使用SCNT生成克隆的完全足月有机体。加入H3K27me3印记的药剂可与注射Kdm4d多核苷酸的Xist KO细胞合用。

[0232] 在一种方法中,将抑制H3K27me3去甲基酶的药剂与SCNT合用。所给药的药剂的剂

量取决于大量因素,包括个体患者的体积和健康情况。对于任何特定受试者,应根据个体需要和给药该组合物或监督该组合物给药的人的专业判断随时间推移而调节具体用药方案。

[0233] 提出以下实施例,以向该领域技术人员提供对任何制作和使用本发明的试验、筛查和治疗方法的完全公开和说明,而不试图限制发明人所认为的他们的发明的范畴。

[0234] 体细胞核转移

[0235] 体细胞核转移(SCNT)是可用于例如家畜(例如,牛、马、绵羊、山羊、猪)的生殖性克隆或用于治疗性克隆的技术,在该技术中,生产所希望的组织用于细胞替代疗法。不幸的是,所克隆的动物苦于因不适宜的印记而生成的某些缺陷,诸如组蛋白H3蛋白亚基上赖氨酸27的三甲基化中的缺陷。可以通过提供编码在SCNT过程中执行该三甲基化事件的酶的mRNA来弥补这一缺陷。在一个实施方案中,在SCNT过程之前或期间,将编码能够执行该三甲基化事件的酶(例如,EZH1、EZH2、PRC2)的mRNA注射到接纳者细胞或核供体细胞内。

[0236] 体细胞核转移包括获得核供体细胞,然后将这一核供体细胞融合到去核的接纳者(最优选去核卵母细胞)细胞内以形成核转移胚胎,激活这一胚胎,以及最终培养该配体或将这一胚胎转移到母系宿主体内。在核转移期间,将来自一个细胞的核DNA的完整补体引入去核细胞中。细胞核转移方法是本领域技术人员周知的。参见,1991年2月19日授权给Prather等人的题为“Multiplying Bovine Embryos”的美国专利4,994,384;1991年10月15日授权给Massey的题为“Bovine Nuclear Transplantation”的美国专利5,057,420;1999年11月30日授权给Stice等人的题为“Production of Chimeric Bovine or Porcine Animals Using Cultured Inner Cell Mass Cells”的美国专利5,994,619;2000年1月19日分别授权给Campbell等人和Wilmut等人的题为“Quiescent Cell Populations For Nuclear Transfer”的英国专利GB 2,318,578和GB2,331,751;2000年1月4日授权给Strelchenko等人的题为“Method of Cloning Bovines Using Reprogrammed Non-Embryonic Bovine Cells”的美国专利6,011,197;以及题为“Method of Cloning Porcine Animals”的美国专利申请09/753,323(代理人案号:030653.0026.CIP1,2000年12月28日提交),其各自通过引用以其包括所有附图、表格和图示的整体并入本文。可使用不为透明到所环绕的卵母细胞实现核转移。

[0237] 在核转移过程中,将核供体细胞或其细胞核引入接纳者细胞中。接纳者细胞优选卵母细胞并且优选是去核的。然而,本发明部分地涉及核转移,其中卵母细胞的细胞核并未从卵母细胞物理性地抽取。可建立核转移胚胎,其中来自供体细胞的核DNA在细胞分裂过程中复制。参见,例如,Wagoner et al.,1996,Functional enucleation of bovine oocytes:effects of centrifugation and ultraviolet light,Theriogenology 46:279-284。此外,也可通过将核供体与超过一个去核卵母细胞组合来实现核转移。而且,可通过将一个核供体、一个或多个去核卵母细胞、以及一个或多个去核卵母细胞的细胞质组合来实现核转移。所得的核供体细胞与接纳者细胞的组合可称为“杂交细胞”。

[0238] 如本文所用,术语“核供体”是指具有可被转运到卵母细胞内的核DNA的任何细胞或其细胞核。核供体可以是已经从细胞中分离的细胞核。该领域技术人员可使用多种技术从细胞中分离细胞核,然后使用该细胞核作为核供体。见,例如,各自通过引用而以其整体并入本文的美国专利4,664,097、6,011,197和6,107,543。任何类型的细胞均可用作核供体。核供体细胞的示例包括但不限于,培养和非培养的从由两个配子在体外或体内结合而

生成的胚胎中分离的细胞;由培养的胚胎细胞(例如,前囊胚细胞和内细胞团细胞)生成的胚胎干细胞(ES细胞);培养和非培养的由从胚胎中分离的内细胞团细胞生成的细胞;培养和非培养的前囊胚细胞;培养和非培养的胎儿细胞;培养和非培养的成年细胞;培养和非培养的原始生殖细胞;培养和非培养的生殖细胞(例如,胚胎生殖细胞);培养和非培养的从动物分离的体细胞;培养和非培养的卵丘细胞;培养和非培养的羊膜细胞;培养和非培养的胎儿成纤维细胞;培养和非培养的生殖嵴细胞;培养和非培养的已分化细胞;同步群体中的培养和非培养细胞;异步群体中的培养和非培养细胞;培养和非培养的血清饥饿细胞;培养和非培养的永久细胞;以及培养和非培养的全能细胞。参见,例如,Piedrahita et al.,1998, Biol.Reprod.58:1321-1329;Shim et al.,1997,Biol.Reprod.57:1089-1095;Tsong et al.,1995,Shih Yen Sheng Wu Hsueh Pao 28:173-189;和Wheeler,1994, Reprod.Fertil.Dev.6:563-568,其各自通过引用以其包括所有附图、图示和表的整体并入本文。此外,核供体可以是先前冷冻或低温保存的细胞。

[0239] 通过核转移过程制备的杂交细胞可用于例如生殖性克隆或再生性克隆中。

[0240] SCNT实验显示,来自成年已分化体细胞的细胞核可重新编程为全能状态。据此,使用本文所公开方法生成的SCNT胚胎可在体外于何时的培养基中培养,以生成全能干细胞或胚胎干细胞或干细胞样细胞和细胞集落。适用于胚胎的培养和突变的培养自是该领域中周知的。可用于牛胚胎培养和维持的已知培养基的示例包括Ham's F-10+10%胎牛血清(FCS)、组织培养基-199(TCM-199)+10%胎牛血清、Tyrodes白蛋白-乳酸盐-丙酮酸盐(TALP)、Dulbecco磷酸盐缓冲盐水(PBS)、Eagle培养基和Whitten培养基。最常用于卵母细胞的收集和突变的培养基之一是TCM-199以及1%至20%血清补充,该血清补充包括胎牛血清、新生儿血清、发情牛血清、高养血清或阉牛血清。优选的维持培养基包括TCM-199和Earl盐、10%胎牛血清、0.2M丙酮酸盐和50ug/ml硫酸庆大霉素。上述任一者也可牵涉与各种细胞类型共培养,诸如粒膜细胞、输卵管细胞、BRL细胞和子宫细胞以及STO细胞。

[0241] 具体地,子宫内膜上皮细胞在着床前和着床期过程中分泌白血病抑制因子(LIF)。因此,在一些实施方案中,涵盖将LIF加入培养基中以提升源自SCNT的胚胎的体外发育。LIF用于胚胎和干样细胞培养用途已经在美国专利5,712,156中有所描述,其通过引用并入本文。

[0242] 另一种维持培养基在Rosenkrans等人的美国专利5,096,822中有所描述,其通过引用并入本文。这一被称为CR1的胚胎培养基含有支持胚胎所必需的营养物质。CR1含有L-乳酸半钙,其含量在1.0mM至10mM,优选1.0mM至5.0mM的范围内。L-乳酸半钙是其上并入有半钙盐的L-乳酸盐。而且,用于维持培养基中的人类胚胎干细胞的合适培养基在Thomson et al.,Science,282:1145-1147(1998)和Proc.Natl.Acad.Sci.,USA,92:7844-7848(1995)中有所讨论。

[0243] 在一些实施方案中,饲养细胞量包含小鼠胚胎成纤维细胞。用于制备合适的成纤维细胞饲养层的手段在下述实施例中描述,并且处于普通专业人士的技能范围内。

[0244] 从囊胚阶段的SCNT胚胎(或其等效物)衍生ES细胞(例如,NT-ESC或hNT-ES)的方法是该领域中周知的。此类技术可用来从SCNT胚胎衍生ES细胞(例如,hNT-ESC),其中,与未使用KDM4甲基化酶家族成员和/或组蛋白甲基转移酶SUV39h1/SUV39h2抑制剂处理的SCNT相比,用来生成hNT-ESC的SCNT胚胎在其由体细胞供体捐赠的核遗传物质中的H3K9me3水平有

所下降。此外或另选地，hNT-ESC可衍生自处于发育的更早阶段的克隆SCNT胚胎。

[0245] 在某些实施方案中，使用本文所公开的方法、组合物和试剂盒从SCNT胚胎生成的卵裂球可使用玻璃移液管解离以获得全能细胞。在一些实施方案中，解离可在存在0.25%胰蛋白酶的情况下发生(Collas and Robl, 43BIOL.REPROD.877-84,1992;Stice and Robl, 39BIOL.REPROD.657-664,1988;Kanka et al., 43MOL.REPROD.DEV.135-44,1996)。

[0246] 在某些实施方案中，从SCNT胚胎获得的囊胚或囊培养簇可用来获得胚胎干细胞系，例如，核转移ESC (ntESC) 细胞系。此类细胞系可例如根据通过引用而以其整体并入本文的Thomson et al., Science, 282:1145-1147 (1998) 和Thomson et al., Proc.Natl.Acad.Sci., USA, 92:7544-7848 (1995) 中所报导的培养方法获得。

[0247] 也可由从SCNT胚胎中移除的单个卵裂球生成多能胚胎干细胞，而不干扰胚胎的正常发育到出生。参见，以下美国专利申请：2004年11月4日提交的60/624,827、2005年3月14日提交的60/662,489、2005年6月3日提交的60/687,158、2005年10月3日提交的60/723,066、2005年10月14日提交的60/726,775、2005年11月4日提交的11/267,555；2005年11月约4日提交的PCT申请PCT/US05/39776，其通过引用而以其整体并入本文；也参见，Chung et al., Nature, Oct.16, 2005 (早于出版的电子公布) 和Chung et al., Nature V.439, pp.216-219 (2006)，其各自的完整公开通过引用而以其整体并入本文。在这种情况下，不摧毁SCNT胚胎就生成了多能干细胞。

[0248] 在本发明的一个方面，该方法包括在研究和治疗中使用衍生自SCNT胚胎的细胞。此类多能干细胞 (PSC) 或全能干细胞 (TSC) 可分化为身体中的任何细胞，包括而不限于，皮肤细胞、软骨细胞、骨骼细胞、骨骼肌细胞、心肌细胞、肾细胞、肝细胞、血液细胞和成血细胞、脉管前体细胞和脉管上皮细胞、胰腺β细胞、神经元细胞、神经胶质细胞、视网膜细胞、内耳滤泡细胞、肠细胞或肺细胞。

[0249] 在本发明的另一实施方案中，可将SCNT胚胎、或囊胚、或从SCNT胚胎获得的多能或全能干细胞 (例如，NT-ESC) 暴露于一种或多种分化诱导剂中，以获得其它可用于治疗的细胞，诸如视网膜色素上皮细胞、造血前体细胞和成血液血管祖细胞、以及外胚层、中胚层和内胚层的很多其它可用细胞类型。此类诱导剂包括但不限于：细胞因子诸如白介素αA、干扰素αA/D、干扰素β、干扰素γ、干扰素γ诱导的蛋白质-10、白介素-1至白介素-17、角化细胞生长因子、瘦素、白血病抑制因子、巨噬细胞集落刺激因子、和巨噬细胞炎性蛋白-1α、1β、2、3α、3β、和单核细胞趋化蛋白1至3、6kine、激活素A、双调蛋白、血管生成素、B-内皮细胞生长因子、β-细胞素、脑源性神经营养因子、C10、心脏营养因子-1、睫状神经营养因子、细胞因子诱导的中性粒细胞趋化因子-1、嗜酸细胞激活趋化因子、上皮生长因子、上皮中性粒细胞激活肽-78、促红细胞生成素、雌激素受体α、雌激素受体β、成纤维细胞生长因子 (酸性和碱性)、肝素、FLT-3/FLK-2配体、胶质细胞源性神经营养因子、Gly-His-Lys、粒细胞集落刺激因子、粒细胞巨噬细胞集落刺激因子、GRO-α/MGSA、GRO-β、GRO-γ、HCC-1、肝素结合上皮生长因子、肝细胞生长因子、调节蛋白-α、胰岛素、胰岛素生长因子结合蛋白-1、胰岛素样生长因子结合蛋白-1、胰岛素样生长因子、胰岛素样生长因子II、神经生长因子、神经营养因子-3和4、制瘤素M、胎盘生长因子、多效生长因子、趋化因子 (rantes)、干细胞因子、基质细胞源性因子1B、血小板生成素、转化生长因子-(α、β1、2、3、4、5)、肿瘤坏死因子(α和β)、血管内皮生长因子、和骨形态形成性蛋白质；改变以下激素和激素拮抗剂表达的酶：诸如17β-雌二醇、促

肾上腺皮质激素、肾上腺髓质素、 α -促黑素细胞激素、绒毛膜促性腺激素、皮质类固醇结合球蛋白、皮质酮、地塞米松、雌三醇、促卵泡激素、胃泌素1、胰高血糖素、促性腺激素、L-3,3',5'-三碘甲状腺氨酸、黄体生成素、L-甲状腺素、褪黑素、MZ-4、催产素、甲状旁腺素、PEC-60、垂体生长激素、孕酮、促乳素、分泌素、性激素结合球蛋白、促甲状腺激素、促甲状腺素释放因子、甲状腺素结合球蛋白、和血管加压素；细胞外基质成分，诸如纤连蛋白、纤连蛋白的蛋白酶水解片段、层粘连蛋白、腱生蛋白、血小板反应蛋白；以及蛋白聚糖，诸如聚集蛋白聚糖、硫酸乙酰肝素蛋白聚糖、硫酸软骨素蛋白聚糖、和多配体蛋白聚糖。其它诱导剂包括源自来自限定组装的细胞的细胞或组分，用来向衍生自本发明的衍生自重新编程细胞的分化中细胞提供诱导信号。此类诱导细胞可源自人类、非人类哺乳动物或鸟类，诸如无特异性热原 (SPF) 的胚胎或成年细胞。

[0250] 卵裂球培养在一个实施方案中，SCNT胚胎可用来生成卵裂球，并利用涉及目前使用的植入前遗传诊断 (PGD) 的体外技术从通过本文所述方法生成的SCNT胚胎分离单个卵裂球，而不摧毁SCNT胚胎或者以其它方式显著改变其生存能力。如本文所证明的，可由从如本文所公开SCNT胚胎中移除的单个卵裂球生成多能胚胎干细胞 (hES) 和细胞系，而不干扰胚胎的正常发育到出生。

[0251] Wilmut等人在克隆山羊“多莉”中的发现 (Wilmut, et al, Nature 385, 810 (1997)，连同Thomson等人在衍生hESC过程中的发现 (Thomson et al., Science 282, 1145 (1998))，已经产生了对于再生细胞移植的极大热情，该再生细胞移植基于源自SCNT胚胎或的患者特异性hESC或由患者自身细胞核生成的SCNT工程化细胞团的建立。这一以通过自体移植避免免疫排斥为目标的策略，可能是SCNT的最强有力的临床理论基础。出于同样的原因，复杂疾病特异性SCNT-hESC的衍生可能加速对疾病机制的发现。对于细胞移植，用衍生自个体小鼠自身SCNT的mESC对小鼠SCID和PD模型进行创新治疗是令人类鼓舞的 (Rideout et al, Cell 109, 17 (2002) ; Barberi, Nat. Biotechnol. 21, 1200 (2003))。最后，创建具有广泛组织相容性的衍生自SCNT的干细胞银行的能力将减少对于持续供应新卵母细胞的需求。

[0252] 在本发明的某些实施方案中，从SCNT胚胎获得的多能或全能细胞 (例如，hNT-ESC) 可任选地经分化，并且可将其引入组织内，它们正常地驻留在在该组织内，以便表现出治疗可用性。例如，可将来自SCNT胚胎获得的多能或全能细胞引入组织内。在某些其它实施方案中，可将来自SCNT胚胎的多能或全能细胞系统地引入或引入到与希望进行治疗之处有一定距离的部位。在此类实施方案中，从SCNT胚胎获得的多能或全能细胞可远程发挥作用或者可磨砺到所希望的部位。

[0253] 在本发明的某些实施方案中，从SCNT胚胎获得的克隆细胞、多能细胞或全能细胞可用于诱导其它多能干细胞的分化。从单个细胞生成能够在维持胚胎的基因表达模式的同时在体外繁殖的细胞的群体，对于诱导其它多能干细胞的分化是有用的。细胞-细胞诱导是在早期胚胎中引导分化的常见手段。很多具有潜在医用性的细胞类型受到正常胚胎发育过程中诱导信号的影响，包括脊椎神经元、心肌细胞、胰 β 细胞和确定性造血细胞。从单个细胞生成的能够在维持胚胎的基因表达模式的同时在体外繁殖的细胞的群体，可在各种体外、卵内或体内条件下培养，以诱导其它多能干细胞分化而变成所希望的细胞或组织类型。

[0254] 从SCNT胚胎获得的多能或全能细胞 (例如，ntESC) 可用来获得任何所希望的已分化细胞类型。此类已分化细胞的治疗性应用是无与伦比的。如本文中讨论的，可根据本文所

公开的方法,使用KDM4组蛋白二甲基化酶激活因子和/或H3K9甲基转移酶抑制剂处理供体细胞、接纳者卵母细胞、杂交卵母细胞或SCNT胚胎。

[0255] 另选地,供体细胞可以是来自具有病变的受试者的成年体细胞,并且所生成的SCNT胚胎可用来生产疾病的动物模型或疾病特异性多能或全能细胞,该多能或全能细胞可在分化条件下培养以产生疾病的细胞模型。本发明的最大优点是,通过提高SCNT的效率,提供了基本上无限制的等基因或同基因ES细胞的供应,尤其是非诱导多能干细胞(例如,不是iPSC)的多能干细胞的供应。与iPSC相比,此类NT-ESC具有一定优势,并且适用于移植,因为它们并非部分多能的,并且不具有病毒转基因或重新编程因子的强制表达以指导它们的重新编程。

[0256] 在一些实施方案中,从SCNT生成的NT-ESC是从SCNT胚胎获得的患者特异性多能细胞,其中供体细胞从待使用多能干细胞或其已分化的子代治疗的受试者获得。因此,它将避免与目前的移植方法有关的重要问题,即,可能因为宿主抗移植物或移植物抗宿主排异而发生的对移植组织的排异。传统上,通过给药抗排异药物诸如环孢菌素来预防或减轻排异。然而,此类药物具有明显的副作用,例如,免疫抑制、致癌性,并且非常昂贵。本发明应消除或至少大大减少对抗排异药物诸如环孢菌素、imulan、FK-506、糖皮质激素和雷帕霉素以及它们的衍生物的需求。

[0257] 除非明确指定,否则本发明的实践采用传统的分子生物学(包括重组技术)、微生物学、细胞生物学、生物化学和免疫学技术,这些技术处于该领域技术人员知识范围内。此类技术在文献中完整地诠释,如,《分子克隆:实验室手册(第二版)》(Molecular Cloning:A Laboratory Manual,second edition(Sambrook,1989);《寡核苷酸合成》(Oligonucleotide Synthesis,(Gait,1984);《动物细胞培养》(Animal Cell Culture,(Freshney,1987);《酶学方法》(Methods in Enzymology)《实验免疫学手册》(Handbook of Experimental Immunology,(Weir,1996);《用于哺乳动物细胞的基因转移载体》(Gene Transfer Vectors for Mammalian Cells,(Miller and Calos,1987);《分子生物学现代方法》(Current Protocols in Molecular Biology,(Ausubel,1987);《PCR:聚合酶链反应》(PCR:The Polymerase Chain Reaction,(Mullis,1994);《免疫学现代方法》(Current Protocols in Immunology,(Coligan,1991))。这些技术可用来生产本发明的多核苷酸和多肽,且因此可在作成并实践本发明中考虑这些技术。尤其可用于特定实施方式的技术将在下文中讨论。

[0258] 提出下述实施例以向该领域技术人员提供对于如何制作并使用本发明的测定、筛选和治疗方法的完全公开和说明,但并不试图限制发明人所认定的发明范畴。

[0259] 实施例

[0260] 实施例1:Kdm4d注射不减轻SCNT相关的异常Xist激活

[0261] 失活的X染色体通过其使用抗H3K27me3抗体进行的点状染色来标记。一直以来,在体外受精的(IVF)中,此类点状染色仅在雌性(XX)细胞中可检测,而在雄性(XY)细胞中检测不到。相比之下,当使用雄性支持细胞作为供体细胞时,在SCNT配体中观察到了此类点状染色(图1A),暗示了SCNT胚胎中的异常Xist激活。重要的是,与不注射的对照SCNT配体相比,Kdm4d注射不改变点状染色图案或其频率(图1A、图1B)。这些结果证明,供体细胞中的H3K9me3以及SCNT配体中Xist的异位激活是SCNT重新编程中的两个独立的屏障。

[0262] 实施例2: Xist突变供体细胞和Kdm4d mRNA注射的合用极大改善了克隆效率

[0263] 两个重新编程屏障彼此独立的事实引起了以下问题: 使用Xist KO供体细胞并且注射Kdm4d的联合方法是否将具有协同效应或加性效应以实现克隆效率的增加。尝试了使用卵丘细胞作为供体的SCNT。在具有B6D2F1背景的野生型对照中, 转移到代孕母亲的胚胎中仅1.2%发育到足月(图1C; 表2)。

[0264] 表2. SCNT胚胎的着床后发育, 关于图1

接纳者的数目	所转移的2细胞胚胎的数目	所移植的数目(%每ET)	幼崽的数目(%每ET)	出生时的体重(g ± SD)	出生时的胎盘重量(g ± SD)
6	171	52 (30.4)	2 (1.2)	1.56 ± 0.04	0.36 ± 0.06
8	179	110 (61.5)	15 (8.4)	1.59 ± 0.14	0.34 ± 0.06
[0265] 4	75	46 (61.3)	14 (18.7)	1.50 ± 0.12	0.30 ± 0.04
3	55	25 (45.5)	1 (1.8)	1.50	0.28
4	77	47 (61.0)	7 (9.1)	1.48 ± 0.11	0.26 ± 0.10
4	85	57 (67.1)	20 (23.5)	1.46 ± 0.16	0.27 ± 0.08
5	40	15 (37.5)	0 (0.0)	N/A	N/A
5	53	36 (67.9)	2 (3.8)	1.70 ± 0.18	0.38 ± 0.04
5	29	23 (79.3)	2 (6.9)	1.89 ± 0.25	0.40 ± 0.13

[0266] 所注射的Kdm4d mRNA为1500ng/μl。N/A, 不可用。ET: 胚胎转移。

[0267] Kdm4d mRNA注射将幼崽生成率增加至8.4%, 与先前的观察结果一致 (Matoba et al., Proc. Natl. Acad. Sci. U.S.A. 108, 20621-20626, 2014)。当使用源自Xist杂合小鼠的卵丘细胞作为供体细胞并且与Kdm4d mRNA注射合用时, 幼崽生成率增加到18.7% (图1C; 表S1)。同样, 源自支持细胞的SCNT胚胎的幼崽生成率通过Kdm4d mRNA注射从1.8%改善到9.1%, 并且通过将Xist KO与Kdm4d mRNA注射合用进一步增加到23.5% (图1C; 表2), 这是有报导以来最高的小鼠克隆率 (Ogura et al., Phil. Trans. R. Soc. B 368, 20110329, 2013)。重要的是, 在杂交 (129S1/Svj x CAST/EiJ) 遗传背景下, 使用不同Xist突变系的MEF细胞也观察到了这一加性效应 (图1C; 表2)。使用给联合方法生成的幼崽生长到成年期并表现出正常的生育能力 (图1D)。

[0268] 上述结果指示, 本发明通过使用联合方法提供了最高的小鼠克隆效率。然而, 即便是最高的克隆效率23.5%仍低于IVF幼崽生成率的一半, 使用IVF, 超过50%的被转移胚胎发育成幼崽。实际上, 在联合支持细胞SCNT组中, 65% (57个中的37个) 的植入胚胎在着床后发育过程中停止发育 (表2)。在不同胚胎阶段进行的仔细形态学检查显示, 胚胎发育停止在着床之后就开始了并且在发育进行过程中逐渐增加 (图2A)。此外, 形态学分析和组织学分析显示了大胎盘表型 (图1E), 这与PAS阳性的海绵状成纤维细胞侵入迷路层有关, 不能通过Kdm4d mRNA注射或联合方法减少 (图1E和图1F; 表2)。因此, 尽管Xist KO和Kdm4d mRNA注射在改善SCNT胚胎发育方面有联合的积极效应, 但其它重新编程屏障可能对这些SCNT胚胎的发育失败有所贡献。

[0269] 实施例3: 在组合重新编程的SCNT囊胚中实现了广泛的DNA甲基化重新编程

[0270] 由于组合重新编程的SCNT胚胎在植入之后就表现出发育缺陷 (图2A、图2B、图

2C),有人假设,重新编程相关的表观遗传缺陷可能已经存在于SCNT囊胚中,但它们在形态上表现正常。为了鉴定此类表观遗传缺陷,生成了两个源自Xist KO MEF细胞(129S1/Svj x CAST/EiJ注射)与Kdm4d mRNA注射联合的SCNT囊胚的全基因组亚硫酸氢盐测序(WGBS)数据,并与基因上相匹配的IVF囊胚的数据比较(图3A)。

[0271] 分别为来自IVF囊胚和SCNT囊胚的2060万和2090万个CpG位点的DNA甲基化信息(表3)。

[0272] 表3.WGBS文库的总结,关于图2

样品	总测序读数	映射的读数	映射的读数的百分比(%)	1x 覆盖的 CpG	5x 覆盖的 CpG	亚硫酸氢盐转化率 (%)
IVF 囊胚	416,465,544	297,678,300	71.5	20,611,901	12,717,260	99.2
SCNT 囊胚	412,706,010	298,966,969	72.4	20,676,648	12,506,862	99.2

[0274] 为了进行比较,也从公共数据库获得了MEF(Yu et al., Proc.Natl.Acad.Sci.U.S.A.111,5890-5895,2014)、精子和卵母细胞(Wang et al.,Cell 157,979-991,2014)的WGBS数据集。首先,计算全部所覆盖CpG的平均DNA甲基化水平。发现精子和卵母细胞中的平均甲基化水平分别为82.2%和58.8%。用作IVF合子的起始甲基化水平的精子和卵母细胞的平均甲基化水平为70.5%(图3B)。然而,高度甲基化的配子通过囊胚阶段被全局地重新编程到低甲基化水平(19.1%),可能是由于在植入前发育期间发生的主动和被动去甲基化过程。

[0275] 之后,计算常用的所覆盖CpG的MEF细胞和SCNT囊胚的DNA甲基化水平,发现供体MEF被高度甲基化(78.0%),但SCNT囊胚被以类似于IVF囊胚甲基化水平(19.1%)的甲基化水平(15.6%)被低甲基化,指示DNA甲基化状态的成功的全球编程(图3B)。实际上,DNA甲基化的成对比较显示,与配子或MEF细胞相比,IVF囊胚和SCNT囊胚两者均具备极低的DNA甲基化(图3C)。不仅全局甲基化水平是相似的,而且已甲基化CpG的分布也是相似的(图5A)。与相似的全局DNA甲基化模式一致,RNA-seq显示了IVF囊胚和SCNT囊胚之间的高度相似的转录组($R=0.988$)(图3D、图4A、图4B和图5B)。实际上,在所检测的8,921个基因(在至少一个样品中,FPKM>3)中,仅92个基因在SCNT囊胚中差异表达($FC>3$)(图5D,表4)。

[0276] 表4:SCNT囊胚中的经差异性调节的基因

	符号	状态	Chr
	Msc	下调	Chr1
	Crygd	下调	Chr1
	Ren1	下调	Chr1
	Mael	下调	Chr1
	Car4	下调	Chr11
	Cbr2	下调	Chr11
	Cd7	下调	Chr11
	Scarna3b	下调	Chr12
	Acot1	下调	Chr12
	Cox8c	下调	Chr12
	Pi1	下调	Chr13
	Bhmt2	下调	Chr13
[0277]	Sult4a1	下调	Chr15
	Tnp2	下调	Chr16
	Fetub	下调	Chr16
	Kng1	下调	Chr16
	Impact	下调	Chr18
	Pde6a	下调	Chr18
	1700123I	下调	Chr19
	01Rik		
	Fabp9	下调	Chr3
	Fabp4	下调	Chr3
	Slc25a31	下调	Chr3
	S100a3	下调	Chr3
	Tdpoz4	下调	Chr3
	Gstm6	下调	Chr3

Scarna2	下调	Chr3
Fam154a	下调	Chr4
Cox7b2	下调	Chr5
Asz1	下调	Chr6
Npy	下调	Chr6
Tuba3b	下调	Chr6
Tex101	下调	Chr7
Klk7	下调	Chr7
Rbmx12	下调	Chr7
Adrb3	下调	Chr8
Tat	下调	Chr8
Tex12	下调	Chr9
2410076I	下调	Chr9
21Rik		
Rbp2	下调	Chr9
AU0227	下调	ChrX
51		
1700013	下调	ChrX
H16Rik		
Xlr	下调	ChrX
Gm773	下调	ChrX
4930550	下调	ChrX
L24Rik		
Xlr3a	下调	ChrX
[0278] Xlr5a	下调	ChrX
Xlr5b	下调	ChrX
Xlr3b	下调	ChrX
Xlr4b	下调	ChrX
Xlr4c	下调	ChrX
Xlr3c	下调	ChrX
Magea8	下调	ChrX
Magea5	下调	ChrX
Gm6432	下调	Chr9
LOC100	下调	Chr12
861571		
Scarna3a	上调	Chr1
Derl3	上调	Chr10
Ddit3	上调	Chr10
Krt19	上调	Chr11
Jdp2	上调	Chr12
Serpina3	上调	Chr12
m		
BC05166	上调	Chr13
5		
Slc39a2	上调	Chr14
Slc7a8	上调	Chr14
Entpd1	上调	Chr19
Aif11	上调	Chr2

	Elf5	上调	Chr2
	Chac1	上调	Chr2
	Trib3	上调	Chr2
	Defb25	上调	Chr2
	My19	上调	Chr2
	Ptk6	上调	Chr2
	Glipr2	上调	Chr4
	Tspan1	上调	Chr4
	Plac8	上调	Chr5
	Apoc2	上调	Chr7
	LOC100302567	上调	Chr7
	Parva	上调	Chr7
	Nupr1	上调	Chr7
	H19	上调	Chr7
[0279]	Gm514	上调	Chr9
	Ostb	上调	Chr9
	Praf2	上调	ChrX
	Rhox6	上调	ChrX
	Plac1	上调	ChrX
	0610009	上调	Chr11
	L18Rik		
	1700019	上调	Chr2
	E08Rik		
	4930461	上调	Chr9
	G14Rik		
	4930591	上调	Chr2
	A17Rik		
	AF357426	上调	Chr12
	Gm5480	上调	Chr16
	LOC100861570	上调	Chr12

[0280] 这些结果指示,DNA甲基化和转录组通过SCNT在囊胚阶段很大程度地重新编程。

[0281] 实施例4:鉴定和表征SCNT囊胚中的差异甲基化区 (DMR)

[0282] 尽管DNA甲基化组被成功地全局重新编程,但SCNT囊胚的平均甲基化水平(15.6%)仍略微但明显地低于(p值 $<2.2e-16$) IVF囊胚(19.1%) (图5B)。执行全基因组扫描分析(10CpG作为最小窗口尺寸)以鉴定SCNT囊胚和IVF囊胚的差异甲基化区(DMR),发现了56,240个绝对甲基化差异大于10%的DMR(图5A)。这些DMR中的大多数(48,315;85.9%)显示在SCNT中的DNA甲基化低于在IVF的甲基化,并且称其为hypoDMR。鉴定了7,925个在SCNT中的DNA甲基化高于在IVF中的甲基化的区,并且称其为hyperDMR。有趣的是,hyperDMR的平均长度(741bp)比hypoDMR的平均长度(5,743bp)短得多(图5B)。实际上,代表性的hyperDMR作为尖峰存在于启动子/增强子区中,而代表性的hypoDMR覆盖整个基因编码区(图6A、图6B)。一致地,hyperDMR和hypoDMR表现出截然不同的基因组分布,hyperDMR富集在基因间区中,而hypoDMR富集在基因本体中(图5C)。

[0283] 为了理解这些DMR是如何形成的以及它们是否可能对SCNT胚胎的着床后发育失败有所贡献,以hypoDMR为研究对象进行分析。在精子中,hypoDMR处的甲基化水平明显高于侧

翼区 (~90%对~80%) (图5D)。在卵母细胞中, hypoDMR和侧翼区之间的甲基化差异甚至更大 (~75%对~60%) (图3D)。相比之下, 在MEF中, hypoDMR和侧翼区之间的甲基化差异小得多 (图5D)。因此, SCNT囊胚的hypoDMR与其配子中相对较高的DNA甲基化水平密切相关, 如果SCNT胚胎和IVF胚胎都经历相同数量的复制依赖性稀释, 则它应在IVF囊胚保持较高的水平。与这一观点一致, 对基因组浏览器视图中的代表性hypoDMR的可视化检查显示, 卵母细胞中的甲基化峰清楚地与IVF囊胚中的甲基化峰重叠 (图6B)。等位基因DNA甲基化分析也支持这一观点, 因为IVF囊胚中的甲基化模式明显偏向母系等位基因 (图5E)。实际上, 对已公布的植入前胚胎的WGBS数据集的分析 (Wang et al., Cell 157, 979-991, 2014) 显示, 母系等位基因在4细胞阶段之前一直维持其在hypoDMR处的高DNA甲基化, 而父系等位基因快速失去了其在这些区处的甲基化 (图5F)。

[0284] 之后, 分析了hyperDMR。在卵母细胞中, hyperDMR处和侧翼区的甲基化水平是相似的 (~50%), 而在精子中, hyperDMR处的甲基化水平显著低于侧翼区 (~55%对~80%) (图5G)。尽管它们在甲基化水平上有所差异, 但在IVF囊胚中, hyperDMR和侧翼区两者均去甲基化至一个非常低的水平 (~20%) (图5G)。相比之下, 在MEF中, hyperDMR被重度甲基化 (~80%), 其甲基化水平甚至高于侧翼区 (图5G)。hyperDMR在MEF中被重度甲基化而在配子中并非如此的事实表明, 在这些区的低甲基化可能是生殖细胞系独有的特征。实际上, 对不同细胞类型的公共DNA甲基化组数据集的分析显示, hyperDMR实际上在全部所分析的体细胞类型中均被重度甲基化, 但在精母细胞、精细胞和卵母细胞中仅较低程度地甲基化 (图5H)。一致地, 对与hyperDMR有关的基因的GO分析表明了生殖细胞系相关功能诸如精子发生和配子发生的显著丰富 (图6C)。在原始生殖细胞 (PGC) 发育期间, HyperDMR似乎随着羟甲基胞嘧啶在 (5hmC) 被显著富集在PGC中的hyperDMR内而通过Tet1被甲基化 (Yamaguchi et al., Nature 504, 460-464, 2013) (图6D)。这暗示hyperDMR主要与生殖系细胞发育有关而与胚胎发育无关。

[0285] 实施例5: SCNT囊胚中H3K27me3依赖性印记的丢失

[0286] 缺陷性的胎盘发育是SCNT胚胎的重要特征。先前的研究已经确定, 基因组印记在胎盘发育中扮演关键角色。因此, 除了已鉴定的可能潜在地对SCNT胚胎的着床后缺陷有所贡献的DRM之外, 重要的是评估联合方法对基因组印记的影响。有鉴于此, 分析了23个已知印记控制区 (ICR) 的DNA甲基化水平, 因为印记基因的等位基因表达极大受控于等位基因ICR甲基化。比较性DNA甲基化分析显示, 尽管大多数ICR处的DNA甲基化水平略有下降, 但23个ICR中的21个维持在IVF囊胚水平的至少一半 (图8A), 指示DNA甲基化介导的基因组印记得到很大程度的维持。实际上, 全部20个具有足够等位基因特异性甲基化信息 (在IVF囊胚和SCNT囊胚两者中的两种等位基因中都检测到了>5个CpG) 的ICR, 示出了在IVF囊胚和SCNT囊胚之间一致的等位基因特异性DNA甲基化 (图8B)。

[0287] 为了进一步评估DNA甲基化加到基因组印记在SCNT缺陷中的潜在角色, 分析了以126个已知印记基因为研究对象的RNAseq数据集。在45个可在IVF囊胚中可靠检测 (FPKM>1) 的印记基因中, 仅6个在SCNT囊胚中比在IVF囊胚中显著上调 (FC>1.5) (图8C)。等位基因表达分析 (在每个样品中, FPKM>1, 平均SNP读数>10) 显示, 在36个具有足够数量的SNP读数的印记基因中, 6个在IVF囊胚中显示偏向母系等位基因的表达 (母系/父系>2.0) 而13个在IVF囊胚中显示偏向父系等位基因的表达 (父系/母系>2.0) (图8D和图8E, 较浅色的柱)。全部6

个母系表达基因 (MEG) 在SCNT囊胚中均维持其偏向母系的表达 (图8D)。在13个父系表达基因 (PEG) 中,7个在SCNT囊胚中失去等位基因偏向并且变成双等位基因表达 (图8E中箭头所指)。有趣的是,7个在SCNT囊胚中失去印记表达的PEG包括Slc38a4、Sfmbt2、Phf17和Gab1 (图8E中深色的柱),已知其印记表达不依赖于DNA甲基化但依赖于母系沉积的H3K27me3 (Inoue et al., Nature 547, 419-424, 2017)。

[0288] 该分析以最近鉴定但不包括在上述分析中的H3K27me3依赖性印记基因为研究对象。在76个在桑椹胚中表现出H3K27me3依赖性印记表达的基因 (Inoue et al., Nature 547, 419-424, 2017) 中,26个在IVF囊胚中以能可靠检测的水平表达 (FPKM>1)。有趣的是,它们中的多数 (15/26) 在SCNT囊胚中得以显著上调 (FC>1.5) (图7A)。等位基因表达分析显示,在23个具有足够SNP读数 (在美国样品中,FPKM>1,平均SNP读数>10) 的基因中,在所分析的基因背景 (129S1/Svj x CAST/EiJ) 下,有17个在IVF囊胚中示出了偏向父系的表达 (父系/母系>2.0)。很明显,全部17个PEG均失去了偏向父系等位基因的表达并且示出了双等位基因表达 (图7B)。这些结果清楚地证明,H3K27me3依赖性印记基因在SCNT囊胚中完全失去了它们的印记。

[0289] 为什么这些H3K27me3依赖性印记基因在SCNT囊胚中失去了印记?由于这些基因的印记状态是由卵子发生过程中沉积的母系等位基因特异性H3K27me3结构域调控的,供体MEF中的H3K27me3模式可能不同于卵母细胞中的模式。对可获得的长成的卵母细胞和MEF细胞的H3K27me3 ChIP-seq数据集的分析显示,卵母细胞中这些印记基因处的H3K27me3结构域在MEF细胞中完全不存在 (图7C和图8F)。将分析拓展至包括其它体细胞类型。这一分析发现,这些印记基因中的H3K27me3结构域通常不存在于所分析的体细胞类型中,因此是卵母细胞基因组所独有的 (图8D)。这些结果指示,H3K27me3甲基化在供体体细胞中这些印记基因的母系等位基因处的缺失可能是SCNT之后印记丢失 (LOI) 的肇因。

[0290] 尽管通过SCNT = 成功克隆了超过20个哺乳动物物种,但克隆效率普遍低下,并且在几乎全部所克隆的哺乳动物物种中均观察到了包括胎盘过度生长在内的发育异常。有人推测表观遗传异常是造成所克隆动物发育失败的原因。在这一研究中,将两种方法联合以克隆先前鉴定的两种妨碍小鼠SCNT胚胎发育的重新编程屏障,Xist KO供体体细胞和Kdm4d mRNA注射的合用实际上将克隆效率 (足月率) 增加了20倍,以达到有使用体细胞供体进行小鼠生殖性克隆的报导以来最高的幼崽生成率 (例如,23%,使用支持细胞)。这一效率是引人类瞩目的,因为它接近于包括类似的细胞核注射的细胞质内精子注射或圆形精细胞注射 (ICSI/ROSI) 的效率 (Ogonuki et al., PLoS One 5, 2010; Ogura et al., International Review of Cytology, pp.189-229 2005)。这一成就清楚地证明,供体细胞中的H3K9me3和Xist的异常激活表示阻碍成功克隆的两个主要屏障,因此构成了理解SCNT介导的重新编程的分子机制的基础。

[0291] 尽管存在引人类瞩目的改善,但很多使用该联合方法生成的SCNT胚胎未能在植入之后发育。此外,无论供体细胞类型任何,仍观察到了胎盘过度生长,指示存在自爱高效率动物克隆的其它屏障。为了鉴定这些其它重新编程屏障,鉴定了发育失败开始的时间点,并且从恰好处于发育表型出现之前的SCNT囊胚生成了第一个WGBS收据及。比较性DNA甲基化组分析显示,通过SCNT成功进行了全局DNA甲基化组重新编程,指示卵质和裂解胚胎中的DNA去甲基化机构在SCNT胚胎中发挥的功能与在IVF胚胎中发挥的功能相似。即便如此,对

DNA甲基化组的详细比较性分析显示了很多横跨IVF囊胚和SCNT囊胚基因组的DMR。有趣的是,hyperDMR富集在生殖细胞系中去甲基的基因组区中,这与以下事实一致:在生殖细胞发育期间,尤其是在原始生殖细胞(PGC)阶段,生殖细胞特异性基因通过Tet1被去甲基化。而SCNT绕过了这一去甲基化过程。hyperDMR相关基因的列举不包括以下被报导在1细胞SCNT胚胎阶段快速去甲基化的生殖细胞系,指示一些生殖细胞系基因但并非多数在SCNT之后经历了去甲基化。另一方面,在卵母细胞中,hypoDMR主要与去甲基化区重叠。在IVF胚胎中,这些区内的母系DNA甲基化似乎逃脱了该去甲基化过程,尤其是在8细胞阶段之前。在8细胞阶段之前的DNA甲基化母系等位基因特异性维持的底层机制是进一步研究的热点。综上所述,看来SCNT个异性DMR(或高或低)是由于配子发生的独特机制形成的,而配子发生是通过正常受精遗传给囊胚的。一经发现,通过受精从卵母细胞传递到胚胎的母系DNA甲基化在早期滋养层发育阶段中扮演重要角色(Branco et al., Dev. Cell 36, 152-163, 2016)。因此,卵母细胞样DNA甲基化模式在SCNT囊胚中的丢失可能对于SCNT囊胚的发育表型有所贡献。

[0292] 如本文所报告,DNA甲基化和对DNA甲基化印记基因的转录组分析显示,大多数ICR在很大程度上维持其正常印记状态,并且大多数经典印记基因实际上在SCNT囊胚中维持等位基因表达模式。相比之下,最近发现的H3K27me3介导的非经典印记基因(Inoue et al., 2017)是整体失调的并且在SCNT囊胚中表现出双等位基因表达。SCNT囊胚中的失调非经典因基因的列举包括Slc38a4、Sfmbt2和Gab1,与先前报导的这三个基因在E13.5 SCNT胚胎的胎盘中的印记丢失(LOI)一致(Okabe et al., Hum. Mol. Genet. 23, 992-1001, 2014)。鉴于已经发现这三个基因全部在胎盘生长中扮演重要角色,这些基因的LOI似乎对于SCNT胚胎的胎盘过度生长表型有所贡献。此外,已经发现Runx1、Otx2和Etv6在小鼠早期胚胎发育中扮演关键角色,因此这些基因在囊胚阶段的LOI可能对于着床后CNT胚胎的胚胎致死表型有所贡献。非经典印记基因在SCNT中的LOI的肇因很可能是由于供体体细胞中的这些基因座处不存在H3K27me3。对H3K27me3印记基因的调节机制的更详细研究将为改善SCNT胚胎发育提供线索。

[0293] 总而言之,除了通过将Kdm4d mRNA注射与使用Xist KO供体细胞联合来构建最有效的小鼠克隆方法之外,还发现H3K27me3印记基因是阻碍有效动物克隆的潜在屏障。不欲受缚于理论,基于小鼠SCNT囊胚中H3K27me3依赖性印记基因处的LOI与它们在胚胎发育中关键功能的清晰关联,在H3K27me3印记基因处的LOI最有可能是SCNT胚胎的着床后表型的原因,但不能排除在这一研究中鉴定的异常DNA甲基化具有潜在贡献的可能性。见于在哺乳动物物种中普遍观察到了SCNT胚胎的陷性着床后发育和异常胎盘表型,可能需要进一步探究其它物种的克隆胚胎中H3K27me3依赖性印记状态。

[0294] 使用下述方法和材料获得所描述的结果。

[0295] 从PN5阶段合子分离母系和父系原核

[0296] 全部动物研究均根据位于哈佛医学院的机构动物管理及使用委员会的指导方针实施。从通过注射7.5I.U. Of的PMSG(密理博公司(Millipore))和hCG(密理博公司)促排的8周龄B6D2F1/J(BDF1)雌性小鼠采集MII期卵母细胞。对于体外受精(IVF),在以10mg/ml牛血清白蛋白补充的HTF培养基(BSA,西格玛-奥德里奇公司(Sigma-Aldrich))中,用从成年BDF1雄性小鼠附睾尾部获得的激活精子对MII卵母细胞授精。通过在HTF培养基中培养1小时达到精子获能。将合子在37.8°C于具有5%CO₂/95%空气的人类源化气氛中培养。在受精

后10小时(10hpf),将合子转移到含有10 μ g/ml细胞松弛素B的M2培养基(西格玛-奥德里奇公司)中。通过电压式显微操作仪(日本茨城的普瑞姆技术公司(Prime Tech Ltd., Ibaraki, Japan))七个透明带,并从合子分离原核。在12hpf(PN5阶段),分离的原核用0.2%BSA/PBS洗涤,转移到Eppendorf LoBind 1.5ml管中,并且在进行DNase I处理之前一直置于冰上。对于每个实验,采集并制备150至200个原核进行1iDNase-seq。通过以下项区分亲本原核:(1)与第二极性体的距离和(2)原核的尺寸。

[0297] 雄性单性生殖(AG)胚胎和雌性单性生殖(GG)胚胎的制备

[0298] 从8周龄促排BDF1雌性小鼠采集MII卵母细胞,并使用BDF1精子授精。在7hpf,将合子转移到含有5 μ g/ml细胞松弛素B的M2培养基中,使用压电式显微操作仪交换亲本原核。使用仙台病毒(HVJ, Cosmo-bio)将核体与先前所述的细胞质融合。重构之后,在KSOM中培养胚胎。

[0299] 当采集胚胎进行RNA-seq或/和1iDNase-seq时,通过短时间暴露于酸性台氏液(Acid tyrode's solution)(西格玛-奥德里奇公司)移除透明带(ZP),然后,先后用M2培养基和0.2%BSA/PBS洗涤胚胎。对于1iDNase-seq,将10个桑椹胚转移到Eppendorf LoBind 1.5ml管中,并且在进行DNase I处理之前一直置于冰上。对于RNA-seq,将七至十个胚胎转移到不含RNase的薄壁PCR管(Ambion)中。分别在30hpf和78hpf采集2细胞胚胎和桑椹胚。当制备 α -鹅膏蕈碱处理的2细胞胚胎时,将5hpf合子转移到含有25 μ g/ml α -鹅膏蕈碱(西格玛-奥德里奇公司)的KSOM中,并且在采集(30hpf)之前一直在存在 α -鹅膏蕈碱的情况下培养。将ICM与TE分离。简而言之,用酸性台氏液处理120hpi的AG胚胎和GG胚胎以移除ZP。用M2培养基洗涤之后,将胚胎在37 $^{\circ}$ C于含有兔抗鼠淋巴细胞血清的KSOM(赛达兰公司(Cedarlane),1:8稀释)中培养45分钟。用M2培养基洗涤之后,将它们转移到含有豚鼠补体的KSOM(MP Biomedicals,1:3.3稀释)。在37 $^{\circ}$ C培养30分钟之后,通过用玻璃毛细管移液移除溶解的TE细胞。将剩余的ICM丛在37 $^{\circ}$ C于0.25%胰蛋白酶/EDTA(赛默飞世尔公司,25200)培养10分钟,然后解离成单细胞以避免溶解TE细胞的污染。采集100至200个细胞进行RNA-seq。

[0300] 从长成的卵母细胞分离GV细胞核

[0301] 在注射5I.U.PMSG之后44至48小时,从3周龄BDF1小鼠获得长成的GV阶段卵母细胞。将卵巢转移到M2培养基中。用30号针穿刺卵泡,使用窄口径玻璃移液管从卵丘-卵母细胞复合物轻柔地移除卵丘细胞。然后,将卵母细胞转移到以5%胎牛血清(FBS)(西格玛-奥德里奇公司,F0926)、10ng/ml上皮生长因子(西格玛-奥德里奇公司,E4127)和0.2mM 3-异丁基-1-甲基黄嘌呤(IBMx;西格玛-奥德里奇公司)补充的 α -MEM(生命技术公司(Life technologies),12571-063)中。采集之后一小时,捕获表现出可见的卵周隙的GV卵母细胞,其具有绕核仁(SN)型染色质。然后,将它们在含有10 μ g/ml细胞松弛素B、0.1 μ g/ml秋水仙酰胺(西格玛-奥德里奇公司)和0.2mM IBMx的M2培养基中培养15分钟。然后,使用压电式显微操作仪分离GV细胞核。用0.2%BSA/PBS洗涤之后,将GV细胞核转移到Eppendorf LoBind 1.5ml管中。对于每个实验,采集115至150个GV细胞核进行1iDNase-seq。

[0302] E6.5胚胎的解剖和GFP阳性E9.5胎盘细胞的FACS分选

[0303] 使用天然交配方案获得C57BL6(B6)/PWK杂交胚胎。为了获得PWK/B6杂交胚胎,使用B6精子进行PWK卵母细胞的体外受精,并且将2细胞胚胎转移到代孕的ICR系母体内。将

E6.5胚胎解剖为EPI、EXE和VE。为了采集E9.5胎盘细胞,从杰克逊实验室(Jackson laboratory)购买B6GFP小鼠[C57BL/6-Tg(CAG-EGFP)1310sb/LeySopJ,物料编号006567]。从促排的8周龄B6GFP或PWK小鼠采集MII卵母细胞和精子。进行体外受精之后,将2细胞胚胎转移到代孕的ICR系母体内。在E9.5收获态样,切割为~0.5mm的片,转移到50ml试管内,在以200rpm震荡下,用2ml的0.25%胰蛋白酶-EDTA(赛默飞世尔科技公司(Thermo Fisher Scientific),25200)在30℃处理15分钟以解离胎盘细胞。通过加入含有10%FBS的2ml DMEM终止胰蛋白酶处理。移液之后,将试管离心,用0.2%BSA/PBS将沉降的细胞洗涤三次。加入DAPI,其在最终的细胞悬浮液中的最终浓度为1 μ M。用BD FACStaria机器(BD Biosciences)分选GFP阳性细胞,将DAPI阳性细胞作为死亡细胞排除。从每个胎盘采集到大约10,000至20,000个GFP阳性细胞,对应于总胎盘细胞的40%至60%。

[0304] 质粒构建和mRNA制备

[0305] 为了生成Kdm6b^{WT}构造体,扩增了编码含有催化剂结构域的羧基端部分的cDNA(氨基酸1025至末端)。克隆了位于pcDNA3.1-Flag-poly(A)83质粒的Flag标签和poly(A)之间的PCR扩增子。使用PrimeSTAR突变发生(TAKARA)生成了H1390A Kdm6b^{MUT}构造体。用于突变发生的引物是5'-CCAGGCgctCAAGAGAATAACAATTTCTGCTCAGTCAACATCAAC-3'和5'-CTCTTGagcGCCTGGCGTTCGGCTGCCAGGGACCTTCATG-3'。所有构造体均通过DNA测序验证。先前描述了用于野生型和H189A突变Kdm4d的质粒。

[0306] 通过限制性内切酶进行线性化之后,构造体用苯酚-氯仿萃取进行纯化,使用mMESSAGE mMACHINE T7 Ultra试剂盒(生命技术公司)根据制造商的使用说明书进行体外转录而合成mRNA。所合成的mRNA通过氯化锂沉淀进行纯化,使用无核酸酶水稀释。mRNA等量小样在使用之前一直在-80℃储存。

[0307] mRNA注射

[0308] 从促排的8周龄BDF1雌性小鼠采集MII卵母细胞,并使用BDF1精子授精。在2.5hpf,将受精的卵母细胞转移到M2培养基中,使用压电式显微操作仪注射mRNA。在4hpf完成mRNA注射。Kdm6b^{WT}和Kdm6b^{MUT}的mRNA浓度为1.8 μ g/ μ l,而Kdm4d^{WT}和Kdm4d^{MUT}的mRNA浓度为1.5 μ g/ μ l。当制备注射Kdm6b的PG胚胎时,以含有5 μ g/ml细胞松弛素B的SrCl₂在无Ca²⁺KSOM中的3mM溶液处理MII卵母细胞以将其化学激活。在激活后4小时(4hpa),用KSOM洗涤胚胎。在5hpa,对它们注射mRNA。

[0309] 全封固免疫染色

[0310] 将合子在多聚甲醛(PFA)在含有0.2%Triton的PBS的3.7%溶液中固定20分钟。用含有10mg/ml BSA的PBS(PBS/BSA)洗涤4次后,将合子以一级抗体在4℃过夜处理。这一研究中使用的一级抗体是鼠抗H3K27me3(1/500,Active Motif公司,61017)、兔抗H3K9me3(1/500,密理博公司,07-442)和兔抗FLAG(1/2000,西格玛-奥德里奇公司,F7524)。用PBS/BSA洗涤3次之后,用荧光素异硫氰酸酯偶联的抗鼠IgG(杰克逊免疫研究实验室(Jackson Immuno-Research))或Alexa Flour568驴抗兔IgG(生命技术公司)的1:250稀释液培养1小时。然后,在具有4',6-二脒基-2-苯基吲哚(DAPI)(加利福尼亚州伯林盖姆的维克多实验室(Vector Laboratories,Burlingame,CA))的Vectashield抗漂白溶液中,将合子封固在载玻片上。在具有旋转盘(CSU-10,横河公司(Yokogawa))和EM-CCD相机(ImagEM,滨松公司(Hamamatsu))或蔡司(Zeiss)LSM800的激光扫描共焦显微镜下,检测荧光。

[0311] 所有图像均使用Axiovision软件(卡尔蔡司公司(Carl Zeiss))获取并分析。用Axiovision软件将荧光信号强度定量。简而言之,测定了母系原核中的信号强度,并且将细胞质信号作为背景减去。然后,将未注射的对照合子的平均信号强度设定为1.0。

[0312] 低输入DNase-seq

[0313] 如先前所述并做最小改变,制备低输入DNase-seq文库。将收集在1.5ml管中的配体和原核再次悬浮于36 μ l细胞溶解缓冲液(10mM Tris-HCl,pH 7.5、10mM NaCl、3mM MgCl₂、0.1% Triton X-100)中并在冰上培养5分钟。加入DNase I(10U/ μ l,罗氏公司(Roche)),最终浓度为80U/ml(对于GV细胞核样品)或40U/ml(对于全部其它样品),并且在37 $^{\circ}$ C培养,培养时间为精确的5分钟。通过加入含有2 μ l蛋白酶K(20mg/ml,生命技术公司)的80 μ l终止缓冲液(10mM Tris-HCl,pH 7.5、10mM NaCl、0.15% SDS、10mM EDTA)终止反应。然后,加入20ng的环状载体DNA(一种不含任何哺乳动物基因的纯质粒DNA,以0.5x Beckman SPRIselect珠(贝克曼库尔特公司(Beckman Coulter))纯化以移除小DNA片段)。将混合物在50 $^{\circ}$ C培养1小时,然后,通过用苯酚-氯仿萃取并通过乙醇在存在线性丙烯酰胺(生命技术公司)的情况下在-20 $^{\circ}$ C过夜沉淀而纯化DNA。将沉淀的DNA再次悬浮在50 μ l TE(2.5mM Tris,pH 7.6、0.05mM EDTA)中,使用整个体积进行测序文库的构建。

[0314] 使用伊露米娜(Illumina)(新英格兰生物实验室(New England Biolabs))的NEBNext Ultra II DNA文库制备试剂盒,根据制造商的使用说明书制备测序文库,但进行一下更改:使用0.03 μ M接头(adaptor)在20 $^{\circ}$ C进行30分钟的连接反应实施接头连接,并且使用Kapa高保真热启动预混液(Kapa Hifi hotstart readymix)(KAPA生物系统公司(Kapa Biosystems))进行8个循环而实施PCR扩增。PCR产物用x1.3体积的SPRIselect珠(贝克曼库尔特公司)纯化,然后,先后使用x0.65体积和x0.7体积的SPRIselect珠选择尺寸。样品在24 μ l TE中洗脱。使用1 μ l的1:1,000稀释样品通过qPCR确定二次PCR扩增所需的循环数。然后,使用Kapa高保真热启动预混液扩增剩余的23 μ l样品(对于这一研究中的所有样品,我们均使用了7个循环)。PCR产物用x1.3体积的SPRIselect珠纯化,然后,先后使用x0.65体积和x0.7体积的SPRIselect珠选择尺寸。DNA在30 μ l的TE中洗脱,通过Qubit dsDNA HS检验试剂盒(赛默飞世尔科技公司,Q32854)和安捷伦高灵敏度检验试剂盒(安捷伦技术公司(Agilent Technologies))定量。将这些文库在具有单端100bp读数HiSeq2500(伊露米娜公司(Illumina))的上测序。

[0315] RNA测序

[0316] 如先前所述制备RNA-seq文库。简而言之,使用SMARTer超低输入RNA cDNA制备试剂盒(Clontech公司,634890),使用全胚胎溶解液实施逆转录和cDNA扩增。当处理2细胞的AG胚胎样品、GG胚胎样品和 α -鹅膏蕈碱处理的IVF胚胎样品时,在细胞溶解步骤将1 μ l的1:40,000稀释的ERCC(外部RNA对照物联合体)标准RNA(生命技术公司)加入每管内。然后,使用具有microTUBE-50(Covaris公司)的Covaris M220超声发生器(Covaris公司)将cDNA片段化为150至160bp的片段。根据制造商的使用说明书,使用伊露米娜公司的NEBNext Ultra DNA文库制备试剂盒(新英格兰生物实验室)对片段化的cDNA进行末端修复、接头连接和扩增。在HiSeq2500测序仪(Illumina)上实施单端100bp测序。

[0317] liDNase-seq数据分析

[0318] 首先,使用trim_galore从liDNase-seq数据的读数修剪掉低质量读数和接头读

数,然后使用Bowtie v0.12.9映射到小鼠基因组(mm9)。“-m1”参数以保持唯一的映射命中。使用SAMtools移除映射质量(MAPQ) ≤ 10 的读数或映射到具有相同取向的相同定位的冗余读数。通过Hotspot程序鉴定liDNase-seq数据中的DHS峰,FDR ≤ 0.01 。使用bedtools中的“bedtools merge”将来自全部33个文库中的DHS峰合并。使用deepTools中的“multiBamSummary”计算每个文库的每个DHS中的读数的数量,并将其归一化至映射读数的总数量和DHS的长度(标签定位在每百万个映射读数中每1kb一个位置处的可能性)。移除性染色体的读数,因为性染色体数量在亲本原核之间是不同的,并且在雄性单性生殖胚胎和雌性单性生殖胚胎之间也是不同的。计算全基因组DHS中tag密度的皮尔森相关系数(r)以测量复本之间的相关性。对于合子和桑椹胚中亲本等位基因特异性DHS的鉴定,使用严格的截止点(在偏向等位基因的全部复本中,RPKM均值 >2 ,RPKM >1 ,并且两种等位基因之间的均值倍变大于4)。通过对Ps-DHS应用额外的准则“显微注射合子的亲本PN的全部复本中,RPKM >1 ”,鉴定了431个最可靠的Ps-DHS。如图2D和图2E中的基因组特征分布分析,使用了来自UCSC基因组浏览器数据库的RefSeq基因组装(mm9)和先前定义的CGI。

[0319] RNA-seq数据分析

[0320] 构建了将小鼠基因组(mm9)与ERCC对照物组合的定制参考序列。使用TopHat v2.0.6或STAR(github.com/alexdobin/STAR)将RNA-seq的读数映射到参考基因组。除非另做指定,否则全部程序均使用缺省参数运行。然后,使用subread-v1.5.1中的featureCounts,将唯一映射的读数组装成由参考注释(UCSC基因模型)指导的转录本。对于所有2细胞的RNA-seq文库,仅采用ERCC读数计数,使用R包DESeq的“estimateSizeFactors”函数估算文库尺寸因子。将文库尺寸归一化之后,用归一化的FPKM(每百万个映射片段中每千碱基外显子的片段数)定量每个基因的表达水平。计算基因表达水平的皮尔森相关系数(r)以指示两个复本之间的相关性。对于在2细胞阶段新合成的转录本的鉴定,滤除AG或GG胚胎与 α -鹅膏蕈碱处理的2细胞胚胎见的统计学不显著基因。有鉴于此,采用负二项模型,用R包DESeq的“nbinomTest”函数计算经调节的P值,并仅选择FDR <0.05 的基因。然后,应用额外的截止点(平均FPKM(AG或GG) >2 ,并且倍变(FC)(AG/Ama或GG/Ama) >2)。结果,分别将4,381个和3,916个基因鉴定为在AG和GG 2细胞胚胎中新合成的基因。对于鉴定2细胞胚胎中的AG特异性DEG和GG特异性DEG,从AG胚胎和GG胚胎减去每种基因在 α -鹅膏蕈碱2细胞胚胎中的表达水平(FPKM)。将显示FC(AG/GG或GG/AG) >10 的基因鉴定为DEG。

[0321] WGBS和H3K27me3 ChIP-seq数据分析

[0322] 使用methpipe v3.4.2计算DHS处的DNA甲基化水平。当计算在每个DHS处的DNA甲基化水平以获得对WGBS读数的足够覆盖时,将每个DHS拓展到上游和下游各2kb处以包括更多的附近CpG位点。通过在卵母细胞中的 $>80\%$ 甲基化和在精子中的 $<20\%$ 甲基化,定义卵母细胞甲基化的gDMR。对于图5A,使用“bedtools makewindows”为Ps-DHS的 ± 100 kb侧翼区生成一组不重叠的1kb类(bin)。对于H3K27me3 ChIP-seq分析,从Zheng et al.,2016下载Bed文件,并使用来自UCSC基因组浏览器数据库的“bedClip”和“bedGraphToBigWig”转化为bigWig格式。使用deepTools中的“multiBigwigSummary”计算DHS和周围区域的H3K27me3信号。

[0323] 统计分析和数据可视化

[0324] 使用R(www.r-project.org/)执行统计分析。使用“cor”函数以缺省参数计算皮尔

森r系数。图6B和图10D是使用R函数“heatmap.2”生成的。图7D、图10C和图12A至图12D是使用R函数“pheatmap”生成的。图1B和图7B是使用deepTools中的“computeMatrix”和“plotHeatmap”函数生成的。通过使用macs2 v2.1.0归一化到文库中的总唯一映射读数并可可视化为IGV基因组浏览器中的定制轨迹,测定了测序读数的基因组位置覆盖率。

[0325] 已知的印记基因信息

[0326] 从www.geneimprint.com/site/genes-by-species.Mus+musculus下载已知的印记信息。

[0327] 代码可用性

[0328] 基于SNP信息,使用定制管线将杂交RNA-seq数据分割到它们的亲本起源。代码可在github.com/lanjiangboston/UniversalSNPsplit找到。

[0329] 数据可用性声明

[0330] 这一研究中生成的所有liDNase-seq数据集和RNA-seq数据集都汇集在登录号为GSE92605的GEO数据库中。精子liDNase-seq数据集来自先前发布(GSE76642)。从www.nodai-genome.org/mouse.html?lang=en下载精子和GV卵母细胞的WGBS数据集。从先前发布(GSE76687)下载精子、MII卵母细胞以及1细胞胚胎的SNP示踪母系和父系等位基因的H3K27me3 ChIP-seq数据集。

[0331] 小鼠植入前胚胎的采集

[0332] 全部动物研究均根据机构动物管理及使用委员会的指导方针在哈佛医学院实施。从通过注射7.5I.U.0f的PMSG(密理博公司(Millipore))和hCG(密理博公司)促排的8周龄B6D2F1/J(BDF1)雌性小鼠采集MII期卵母细胞。对于体外受精(IVF),在以10mg/ml牛血清白蛋白补充的HTF培养基(BSA,西格玛-奥德里奇公司)中,用从成年BDF1或PWK雄性小鼠(杰克逊实验室,003715)附睾尾部获得的激活精子对MII卵母细胞授精。通过在HTF培养基中培养1小时达到精子获能。将合子转移到KSOM,并在37.8°C于具有5%CO₂/95%空气的人类源化气氛中培养。

[0333] mRNA注射

[0334] 在受精后4小时(4hpf),将合子转移到M2培养基中,使用压电式显微操作仪(日本茨城的普瑞姆技术公司)注射mRNA。mRNA的构建和制备如上所述。Kdm6b^{WT}和Kdm6b^{MUT}的注射mRNA浓度为1.8μg/μl,并且Kdm4d^{WT}和Kdm4d^{MUT}的注射mRNA浓度为1.5μg/μl。

[0335] 用于荧光原位杂交的探针

[0336] 根据制造商的使用说明书,使用Nick转译试剂盒(雅培分子公司(Abbott Molecular),07J00-001)和Cy3-dCTP(GE保健公司(GE healthcare),PA53021A)制备用于Xist RNA的探针。用于探针制备的模板DNA是编码全长度小鼠Xist基因的质粒,是Rudolf Jaenisch赠送的礼物(pCMV-Xist-PA,26760)(Wutz and Jaenisch,2000)。使用同一试剂盒和Green-dUTP(雅培分子公司,02N32-050)制备用于DNA FISH的探针。模板DNA是含有Rnf12基因座(RP23-36C20)的BAC克隆物。将该荧光探针与5μg Cot-1 DNA(生命技术公司)、5μg鲱鱼精子DNA(赛默飞世尔科技公司)和2.5μg酵母tRNA(赛默飞世尔科技公司,AM7119)一起进行乙醇沉淀,然后用20μl甲酰胺(赛默飞世尔科技,17899)溶解。将探针在4°C储存。使用前,将探针(每次0.75μl)与已经进行乙醇沉淀并溶解在甲酰胺中的0.75μl Cot-1 DNA、以及2.25μl的4xSSC/20%葡萄糖(密理博公司,S4030)混合。将探针混合物在80°C加热30分钟,

然后转移到37°C培养箱中(“预退火的探针”)。

[0337] 全封固RNA/DNA荧光原位杂交

[0338] 在室温下,在78hpf将桑椹胚在2%多聚甲醛(PFA)在含有0.5%Triton X-100的PBS中的溶液中固定20分钟。用含有1mg/ml BSA(PBS/BSA)的PBS洗涤3次之后,在4°C将胚胎用含有0.02%Triton X-100的0.1N HCl处理15分钟。用含有0.1%BSA的2xSSC洗涤3次之后,将胚胎在15 μ l的10%甲酰胺/2xSSC中在玻璃皿(电子显微镜科学公司(Electron Microscopy Science),705430-30)内培养。通过轻柔移液令全部胚胎下沉并粘附到玻璃皿的底部。5分钟之后,加入15 μ l的30%甲酰胺/2xSSC。5分钟止呕,加入90 μ l的60%甲酰胺/2xSSC,使得最终的甲酰胺浓度为50%,在室温下将胚胎再培养30分钟。用矿物油覆盖含有胚胎的甲酰胺溶液。将样品在80°C加热30分钟,然后转移到37°C培养箱中培养至少30分钟。将胚胎拾取到玻璃移液管中,转移到另一玻璃皿上覆盖有矿物油的4.5 μ l“预退火的探针”中,并且在37°C培养至少24小时。将胚胎用含有0.1%BSA的预热(42°C)2xSSC洗涤,并且在最后一滴洗涤液中停留30分钟。用1%BSA/PBS洗涤3次之后,将它们在具有4',6-二脒基-2-苯基吲哚(DAPI)(加利福尼亚州伯林盖姆的维克多实验室)的Vectashield抗漂白溶液中封固在载玻片上。在蔡司LSM800型激光扫描共焦显微镜下检测荧光。

[0339] 全封固免疫染色

[0340] 免疫染色和定量的过程如上所述。

[0341] 偏向母系等位基因的H3K27me3的计算鉴定

[0342] 从GEO下载编号为GSE76687的用于内部细胞团(ICM)中H3K27me3 ChIP-seq的包括RPKM值在内的100bp类的bed文件。将标记为母系或父系的含有两个亲本你等位基因和等位基因读数的RPKM值的Bed文件归一化至读数的总数。使用“bedtools makewindows”生成用于mm9基因组的1000bp类,然后通过“bedtools map”计算每个类的RPKM值。使用信号截止点1和倍变截止点4将所有的类分为三大类:无信号、双等位基因、偏向母系。使用滑动窗法鉴定含有偏向母系的H3K27me3类的窗口,鉴定标准是:窗口尺寸为20kb,最小类数目为3,以及偏向母系的H3K27me3类的百分比大于50%。重叠窗口用“bedtools merge”合并。在基因组中鉴定了总计5986个窗口。

[0343] RNA测序

[0344] 如上文所述并做最小改变,制备RNA-seq文库。简而言之,使用SMARTer超低输入RNA cDNA制备试剂盒(Clontech公司,634890),使用全胚胎溶解液实施逆转录和cDNA扩增。然后,使用Nextera XT DNA文库制备试剂(Illumina)进行片段化而将cDNA片段化。根据制造商的使用说明书,使用Nextera PCR预混液扩增片段化的cDNA。在HiSeq2500测序仪(Illumina)上实施单端100bp测序。

[0345] RNA-seq数据分析

[0346] 使用STAR(github.com/alexdobin/STAR)将RNA-seq的读数映射到参考基因组。除非另做指定,否则全部程序均使用缺省参数运行。然后,使用subread-v1.5.1中的featureCounts,将唯一映射的读数组装成由参考注释(UCSC基因模型)指导的转录本。将文库尺寸归一化之后,用归一化的FPKM(每百万个映射片段中每千碱基外显子的片段数)定量每个基因的表达水平。计算基因表达水平的皮尔森相关系数(r)以指示两个复本之间的相关性。

[0347] 使用R (www.r-project.org/) 执行统计分析。使用“cor”函数以缺省参数计算皮尔森r系数。

[0348] 代码可用性

[0349] 基于SNP信息,使用定制管线将杂交RNA-seq数据分割到它们的亲本起源。代码可在github.com/lanjiangboston/UniversalSNPsplit找到。

[0350] 数据可用性

[0351] 这一研究中生成的所有RNA-seq数据集都汇集在登录号为GSEXXXXXX.的GEO数据库中。从www.nodai-genome.org/mouse.html?lang=en下载GV卵母细胞的WGBS数据集。将来自同一100bp类的WGBS读数聚集在一起以计算平均甲基化水平,并且需要覆盖最少10个读数。从先前研究(GSE76687)下载精子、MII卵母细胞以及1细胞、2细胞胚胎和囊胚内细胞团的SNP示踪母系和父系等位基因的H3K27me3 ChIP-seq数据集。卵母细胞DNaseI-seq数据集来自于上(GSE92605)。

[0352] 小鼠

[0353] 使用B6D2F1/J (BDF1) 小鼠来及用于SCNT的接纳者卵母细胞。对于小鼠胚胎成纤维细胞(MEF)的细胞制备,令维持在129S1/SvImj背景下的Xist KO雌性小鼠(Marahrens et al., 1997)与CAST/EiJ雄性小鼠交配以生成129/CAST F1背景下的Xist杂合KO胚胎。对于卵丘和支持细胞制备,令C57BL/6N背景下的Xist KO雌性小鼠(Sado et al., 2005)与DBA/2N雄性小鼠交配以生成BDF1背景下的Xist杂合KO胚胎。全部动物实验均获得哈佛医学院机构动物管理和使用委员会和日本理化学研究院筑波研究所(RIKEN Tsukuba Institute)机构动物管理和使用委员会的批准。

[0354] 供体细胞制备

[0355] 原代MEF细胞来自13.5dpc的Xist KO雄性小鼠胚胎。移除头部和全部器官之后,将来自剩余躯体的组织糜在37°C于具有1mM EDTA(赛默飞世尔科技公司,#25200056)的500 μ l的0.25%胰蛋白酶中解离10分钟。用含有10%FBS和青霉素/链霉素(赛默飞世尔科技公司,#15140-022)的等量DMEM(赛默飞世尔科技公司,#

[0356] 11995-073) 稀释细胞悬浮液,并且上下移液20次。用新鲜培养基稀释细胞悬浮液,放置在100mm培养皿中,并在37°C培养。两天后,收获MEF细胞并冷冻。将MEF细胞的冷冻物料融化,并在传代一代之后用于实验中。

[0357] 从通过注射7.5IU的孕马血清促性腺激素(PMSG;密理博公司,#367222)和7.5IU的人类绒毛膜性腺激素(hCG;密理博公司,#230734)促排的野生型(WT)和Xist杂合KO成年雌性小鼠(日本理化学研究院生物资源中心(RIKEN BioResource Center),RBRC01260)采集卵丘细胞。注射hCG后15至17小时,从输卵管采集卵丘-卵母细胞复合体(COC),并使用含有300U/ml牛睾丸透明质酸酶(卡尔生化公司(Calbiochem),#385931)的Hepes缓冲钾单纯形优化培养基(KSOM)简短处理以获得解离的卵丘细胞。

[0358] 如所述者,从3至5日龄的WT或Xist KO雄性小鼠的睾丸采集支持细胞(Matoba et al., 2011)。在37°C,将睾丸团在含有0.1mg/ml胶原酶(赛默飞世尔科技公司,#17104-019)的PBS中培养30分钟,之后在室温用具有1mM EDTA的0.25%胰蛋白酶处理5分钟。用含有3mg/ml牛血清白蛋白的PBS洗涤四次之后,将解离的细胞悬浮在Hepes-KSOM培养基中。

[0359] Kdm4d mRNA合成

[0360] 如先前所述,通过体外转录 (IVT) 合成Kdm4d (Matoba et al., 2014)。简而言之,通过XbaI将含有全长度小鼠Kdm4d以及之后的 (A) 83 (Addgene, #61553) 的pcDNA3.1质粒线性化。纯化之后,使用线性化的质粒DNA作为模板,以使用mMESSAGE mMACHINE T7 Ultra试剂盒 (赛默飞世尔科技公司, #AM1345) 进行IVT。将合成的mRNA溶解于无核酸酶水中,并通过NanoDrop ND-1000分光光度计 (纳米滴技术公司 (NanoDrop Technologies)) 定量。将mRNA稀释至1500ng/ μ l之后,将等量小样在-80 $^{\circ}$ C储存。

[0361] SCNT

[0362] 如先前所述,进行小鼠体细胞核转移 (Matoba et al., 2014; Ogura et al., 2000)。简而言之,从通过注射7.5IU的PMSG和7.5IU的hCG促排的成年BDF1雌性小鼠采集接纳者MII卵母细胞。注射hCG后15至17小时,从输卵管采集卵丘-卵母细胞复合体 (COC), 并使用含有300U/ml牛睾丸透明质酸酶的Hepes-KSOM简短处理以MII卵母细胞。使用压电式显微操作仪 (普瑞姆, #PMM-150FU), 在含有7.5 μ g/ml的细胞松弛B (卡尔生化公司, #250233) 的Hepes缓冲KSOM培养基中将分离的MII卵母细胞去核。将卵丘或支持细胞的细胞核注射到去核的卵母细胞内。通过灭活的仙台病毒包膜 (GenomOne CF; 石原产业株式会社 (Ishihara Sangyo), #CF001) 将MEF细胞与去核的卵母细胞融合。在KSOM中培养1小时后,通过在含有3mM氯化锶 (SrCl₂) 和5 μ g/ml细胞松弛素B的无钙KSOM中培养1小时而激活重构的SCNT卵母细胞,并且进一步在具有5 μ g/ml细胞松弛素B的KSOM中培养4小时。在启动SrCl₂处理后5小时 (激活后5小时, 5hpa), 洗涤激活的SCNT胚胎, 并且在37.8 $^{\circ}$ C于具有5%CO₂的湿润气氛下在KSOM中培养。在5至6hpa, 使用压电式显微操作仪向一些SCNT胚胎注射 \sim 10p1的1500ng/ μ l小鼠Kdm4d mRNA。

[0363] 胚胎转移

[0364] 将2细胞阶段的SCNT胚胎转移到假孕 (E0.5) ICR雌性小鼠的输卵管中。在分娩当天 (E19.5) 行剖腹产回收幼崽, 并且由哺乳期ICR雌性小鼠照料。在E4.5和E10.5牺牲一些雌性小鼠, 以检查胚胎发育。

[0365] 胎盘组织学分析

[0366] 将E19.5的胎盘在4%多聚甲醛 (PFA) 于4 $^{\circ}$ C固定过夜, 并常规地包埋在石蜡中。对系列切片 (厚度4 μ m) 进行周期性酸性席夫试剂 (PAS) 染色。

[0367] 囊胚H3K27me3的免疫染色

[0368] 在室温将囊胚使用4%多聚甲醛 (PFA) 固定20分钟。用含有10mg/ml BSA (PBS/BSA) 的PBS洗涤之后, 将固定化的胚胎用0.5% Triton-X 100培养15分钟以将其通透化。在室温下于PBS/BSA中封闭1小时之后, 将它们在包括兔抗H3K27me3抗体 (1/500, 密理博公司, 07-449)、山羊抗Oct4抗体 (1/500, 圣克鲁兹生物技术公司 (SantaCruz), sc-8628) 和鼠抗Cdx2抗体 (1/100, 拜尔基因公司 (BioGenex), AM392-5M) 的一级抗体混合物中于4 $^{\circ}$ C培养过夜。使用PBS/BSA洗涤三次之后, 将胚胎用包括荧光素异硫氰酸酯缀合的抗鼠IgG (1/400, 杰克逊免疫研究实验室)、Alexa Flour 546驴抗兔IgG (1/400, 赛默飞世尔科技公司) 和Alexa Flour 647驴抗山羊IgG (1/400, 赛默飞世尔科技公司) 的二级抗体混合物在室温培养1小时。最后, 将它们用具有4', 6-二脒基-2-苯基吲哚 (DAPI) (维克多实验室, #H-1200) 的Vectashield封固。使用激光扫描共焦显微镜 (蔡司LSM510型) 和EM-CCD相机 (滨松ImagEM型) 观察荧光信号。

[0369] WGBS

[0370] 使用EZ DNA甲基化直接试剂盒(酶研究公司(Zymo Research), D5020)对早期囊胚阶段(受精或激活后96小时)的IVF胚胎和SCNT胚胎直接进行亚硫酸氢盐转化。分别使用39个胚胎和36个胚胎制备IVF样品和SCNT样品。在进行亚硫酸盐转化之前,将少量(0.01ng)的非甲基化λDNA(普洛麦格公司(Promega), D152A)加入每个样品中,以用作评估亚硫酸氢盐转化效率的跟踪标定对照物。使用EpiGnome Methyl-Seq试剂盒(Epicenter, EGMK81312),遵循制造商使用说明书,制备测序文库。仅将文库扩增12个循环,然后使用Agencourt AMPure XP珠(贝克曼库尔特公司, A63880)纯化。在具有PhiX跟踪标定对照物的HiSeq 2500测序仪(伊鲁米娜)上对最终文库进行单读数(100bp)测序。

[0371] RNA-seq

[0372] 将早期囊胚阶段(受精或激活后96小时)的6个IVF或SCNT胚胎直接溶解,并用于使用SMARTer超低输入RNA cDNA制备试剂盒(克隆技术公司(Clontech), 634936)进行的cDNA合成。扩增之后,使用M220超声发生器(Covaris公司)将cDNA样品片段化。根据制造商的使用说明书,使用伊鲁米娜的NEBNext Ultra DNA文库制备试剂盒(新英格兰生物实验室, E7370),使用片段化的cDNA作成测序文库。在HiSeq 2500测序仪(Illumina)上实施单读数50bp测序。

[0373] 量化和统计分析**[0374] WGBS和RRBS数据分析**

[0375] 首先,使用trim_galore修剪WGBS和简并代表性亚硫酸氢盐测序(RRBS)读数,以移除低质量序列和接头序列。使用Bismark(版本0.15.0)将读数与亚硫酸氢盐转化的参考基因组(mm9)对齐。使用bismark_methylation_extractor从对齐的读数中提取出每个胞嘧啶的覆盖深度和甲基化水平。当计算CpG位点的甲基化水平时,将来自两条链的信息合并,并且需要覆盖至少五个读数。使用methpipe(版本3.4.3)鉴定DMR,并进一步过滤,需要至少10个CpG位点和至少10%的甲基化差异。使用R中的clusterProfiler(版本2.4.3)实施DMR相关基因(即,具有定位在TSS±3kb区内的DMR的基因)的功能注释。对于已知ICR的等位基因特异性甲基化分析,将一个ICR内的所有已检测CpG聚集在一起,需要覆盖两个等位基因中至少5个已检测CpG以进行进一步甲基化比较。

[0376] RNA-seq数据分析

[0377] 使用TopHat(版本2.0.14)以参数“无覆盖检索-无新连接点-文库类型=fr-unstranded”将RNA-seq数据映射到小鼠基因组(mm9)。然后,使用Cufflinks(版本2.2.1)中的featureCounts,将唯一映射的读数组装成由参考注释(UCSC基因模型)指导的转录本。用归一化的FPKM(每百万个映射片段中每千碱基外显子的片段数)定量每个基因的表达水平。计算基因表达水平的皮尔森相关系数以指示两个复本之间的相关性。

[0378] 等位基因特异分析

[0379] 将杂交WGBS和RNA-seq数据(129S1/Svj x CAST/EiJ)映射到小鼠基因(mm9)之后,基于从小鼠基因组项目(ftp://ftp-mouse.sanger.ac.uk/REL-1211-SNPs_Indels/)下载的SNP信息,使用定制的Perl脚本将所映射的读数拆分成它们的亲本来源。然后,个体地处理等位基因读数。

[0380] ChIP-seq和DIP-seq数据分析

[0381] 使用Bowtie (版本2.1.0) 以参数“-D 20-R 3-N 1-L 20-i S,1,0.50”将下载的ChIP-seq和DIP-seq读数映射到小鼠基因组(mm9), 以仅获得那些具有最多3个误配的唯一映射的读数。为了将信号以基因组浏览器形式可视化, 通过将唯一映射的读数(在同一基因组位置保留最多两个读数) 向3' 扩展200bp并将读数计数分类为50bp间隔, 我们使用MACS2 (版本2.1.1) 生成了每个数据集的wig跟踪文件。将每个类中的Tag计数进一步归一化至唯一映射读数的总数(每百万个读数中的读数,RPM)。使用deepTools的“computeMatrix”程序计算DMR上或ICR上和它们的侧翼区上的ChIP-seq和DIP-seq信号。

[0382] 统计分析和数据可视化

[0383] 使用R (版本3.4.1, <http://www.r-project.org>) 执行统计分析和绘图。使用“cor”函数以缺省参数计算皮尔森相关系数。使用“t.test”函数以缺省参数实施学生t检验(双尾, 等方差)。将ChIP-seq信号和DNA甲基化水平可视化为整合基因组观察器(Integrative Genomics Viewer) 基因组浏览器形式的定制跟踪。

[0384] 数据和软件可用性

[0385] 这一研究中生产的WGBS和RNA-seq数据集已经保存在基因表达综合数据库(Gene Expression Omnibus (GEO)) 的登录号GSE109214下。

[0386] 本研究中使用的公开数据库

[0387] 从GSE56697 (Wang et al., 2014) 获得了植入前胚胎(2细胞、4细胞和ICM) 的母系和父系DNA甲基化。从GSE11034和GSE43719 (Soumillon et al., 2013) 获得了不同细胞和体组织的RRBS数据。从GSE49847 (Yue et al., 2014) 和GSE76687 (Zheng et al., 2016) 获得了H3K27me3 ChIP-seq数据。从SRP016940 (Hackett et al., 2013) 下载了PGC发育过程中5mC和5hmC的DIP-seq数据。

[0388] 其它实施方案

[0389] 从前述说明书可知, 可对本文所述的发明作出变更和修改以令其适应各种用途和条件。此类实施方案也处于所附权利要求书的范畴内。

[0390] 本文中, 对任何变量定义中一系列元件的描述包括该变量作为任何单一元件或作为所列元件的组合(或亚组合) 的定义。对本文中实施方案的描述包括该实施方案作为任何单一实施方案或作为与任何其它实施方案或其部分的组合。

[0391] 本说明书中提及的全部专利和出版物通过引用并入本文, 其并入程度与各独立的专利和出版物具体且独立地指明以待通过引用并入相同。

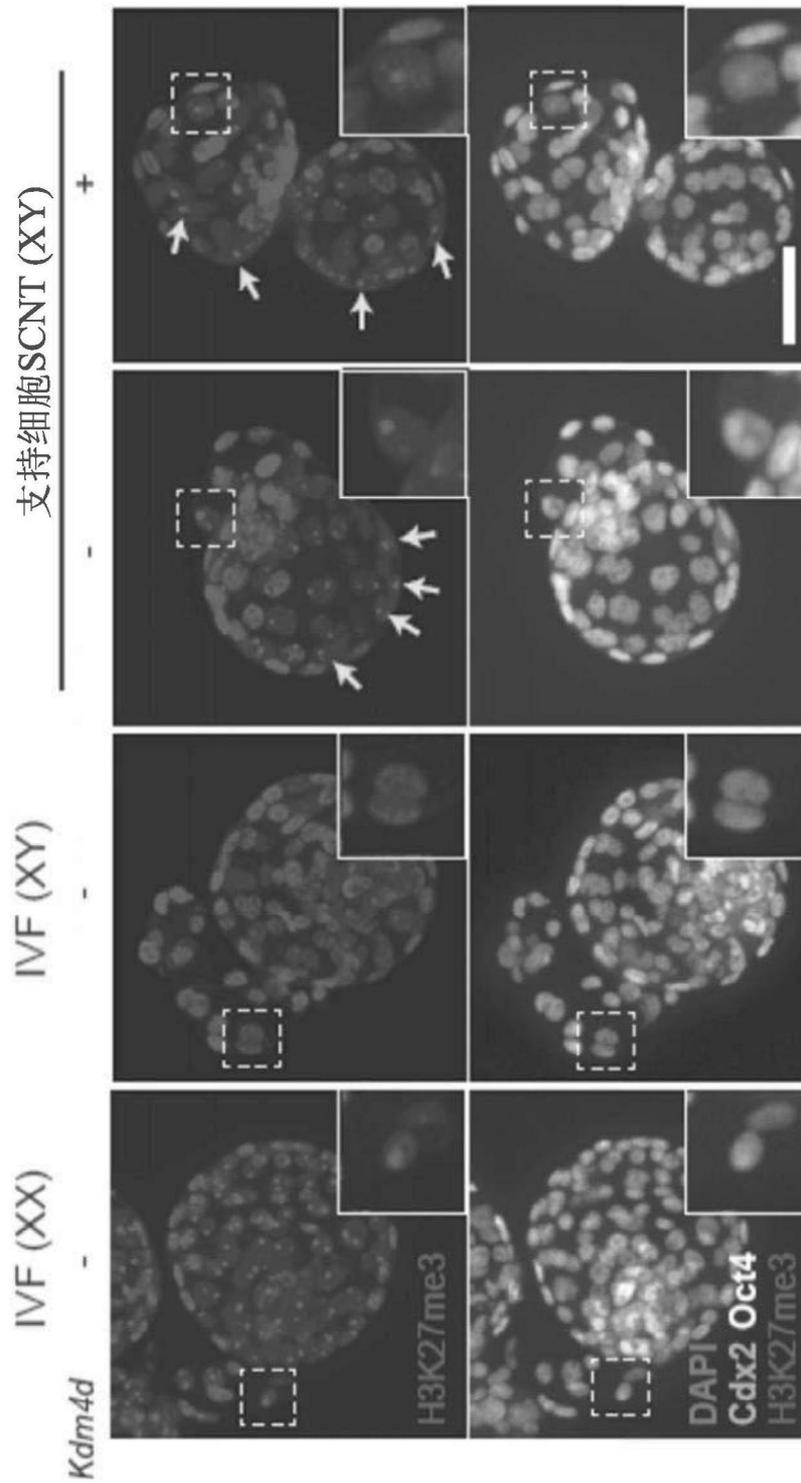


图1A

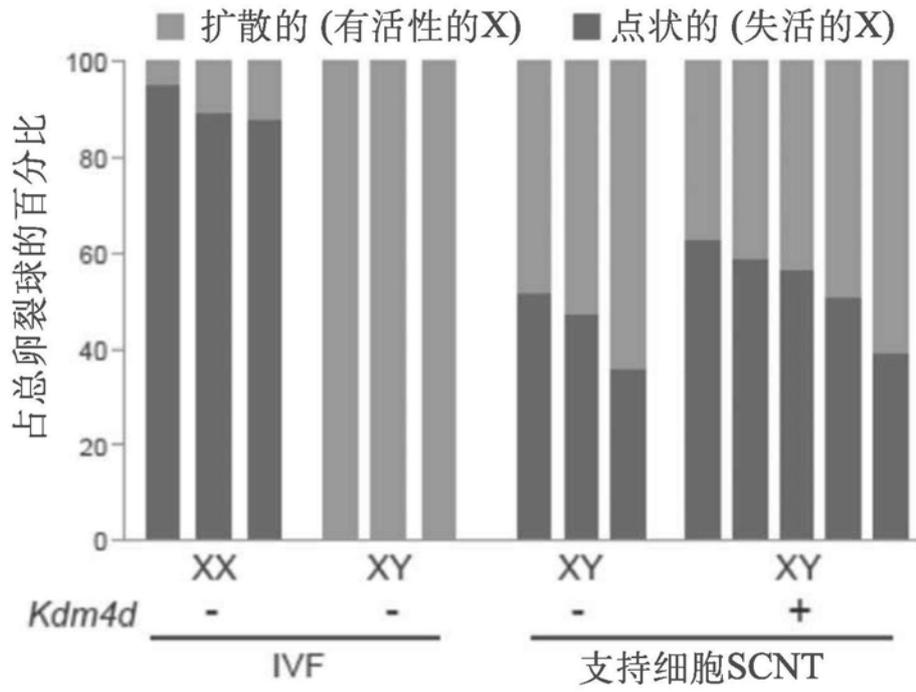


图1B

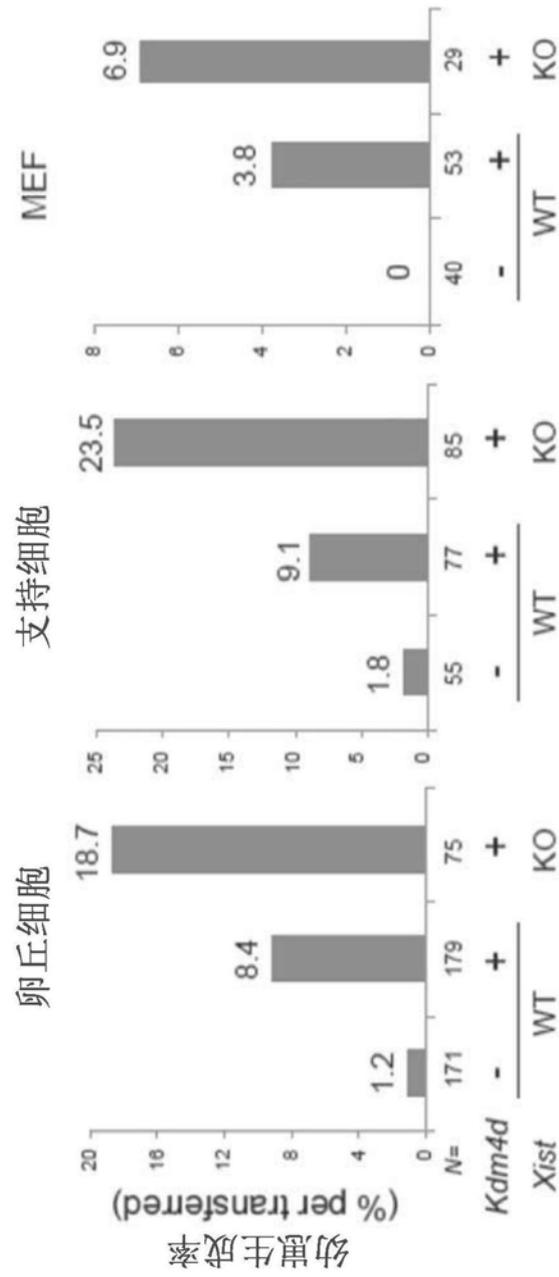


图1C



图1D

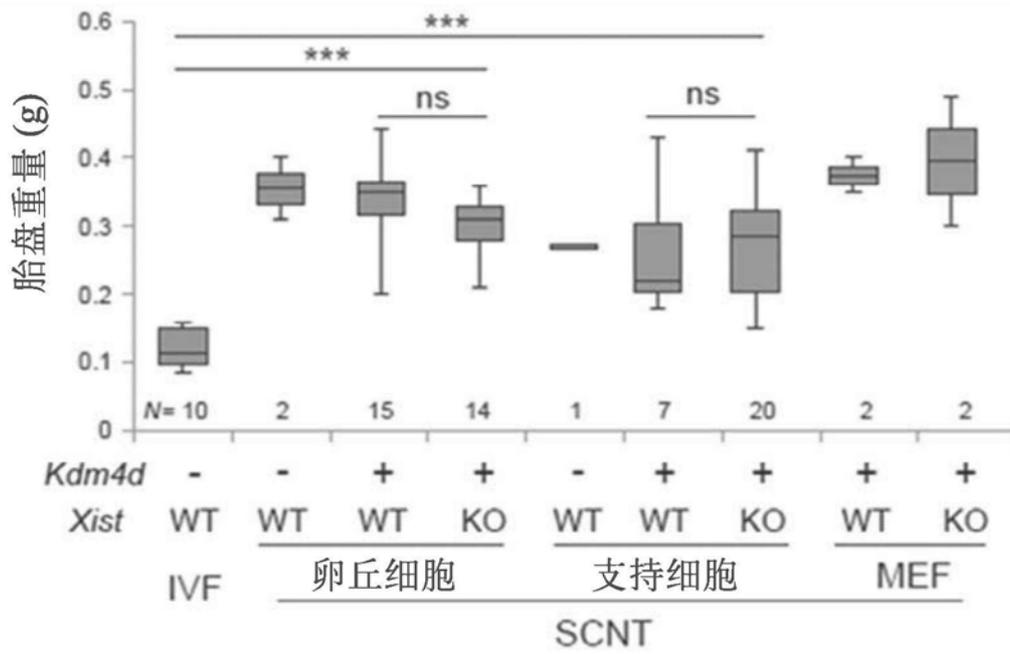


图1E

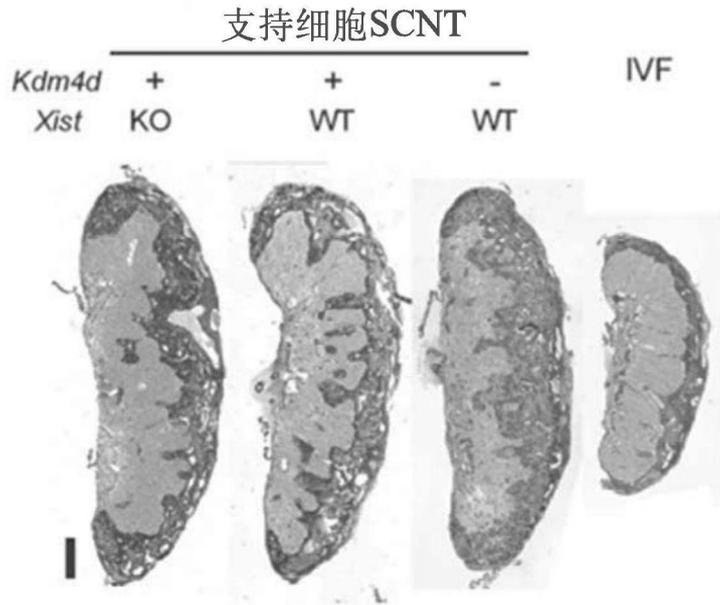


图1F

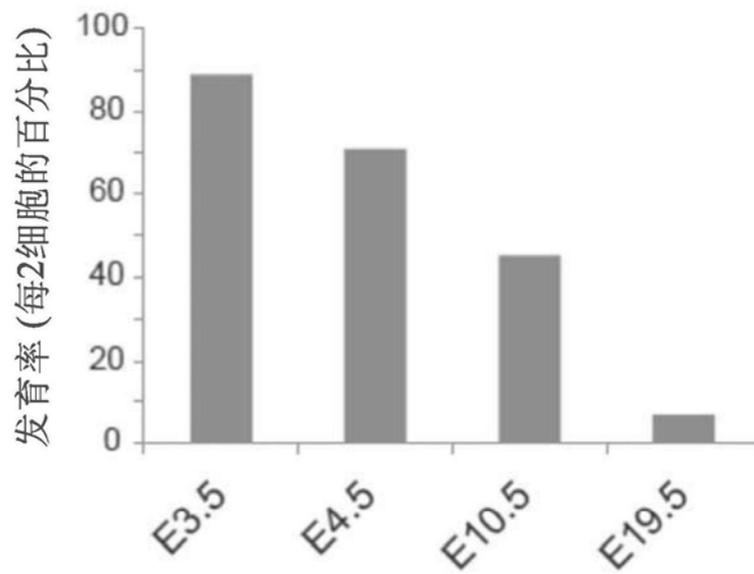


图2A

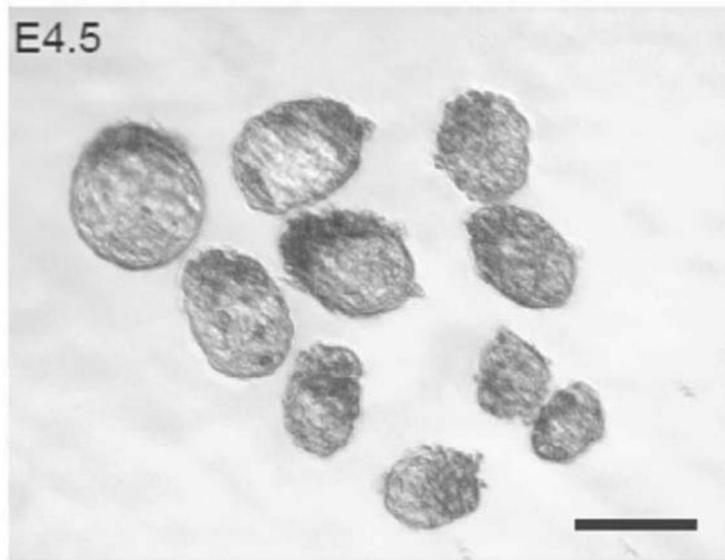


图2B

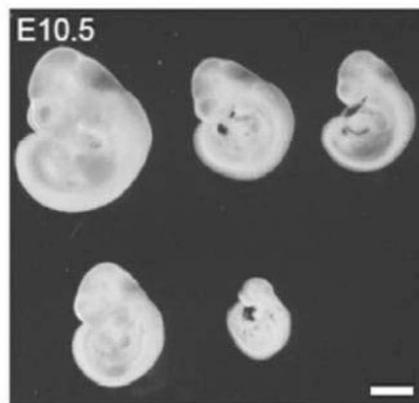


图2C

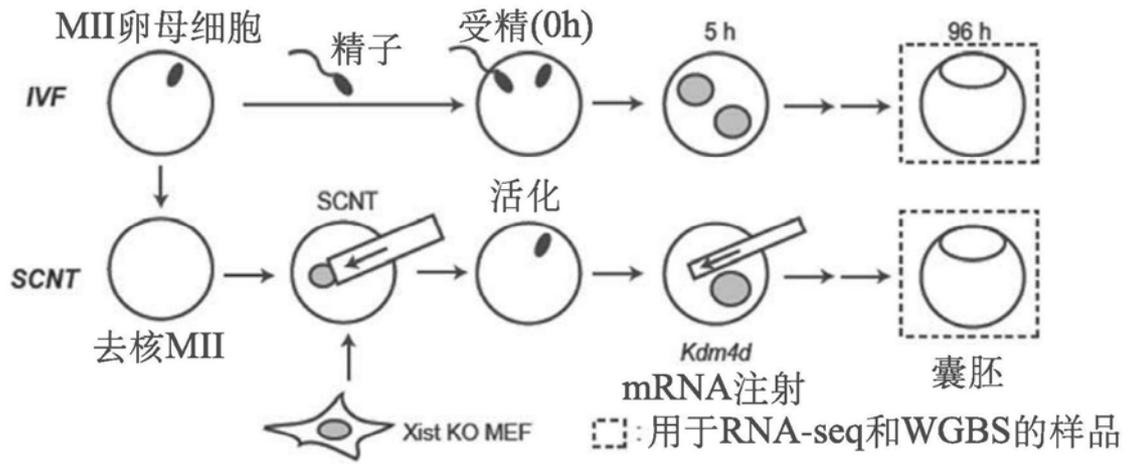


图3A

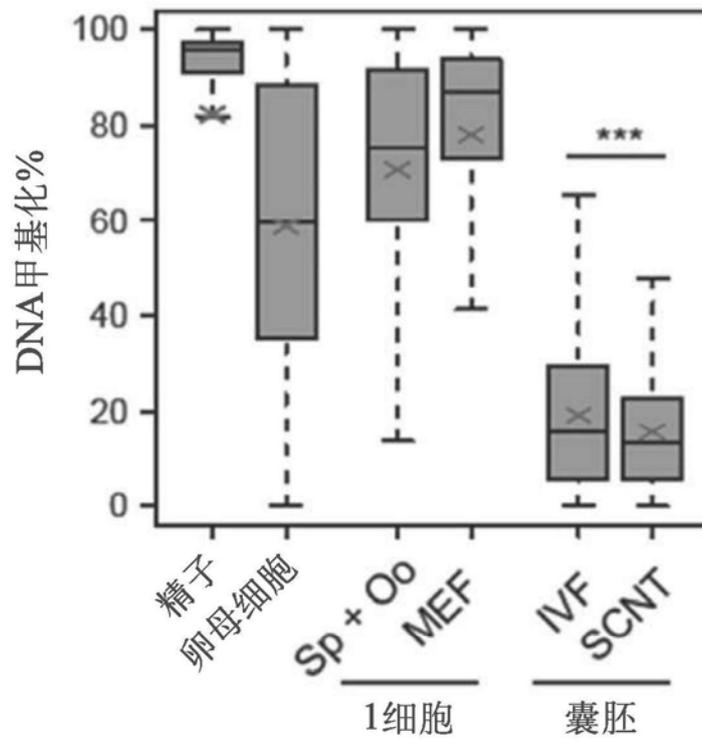


图3B

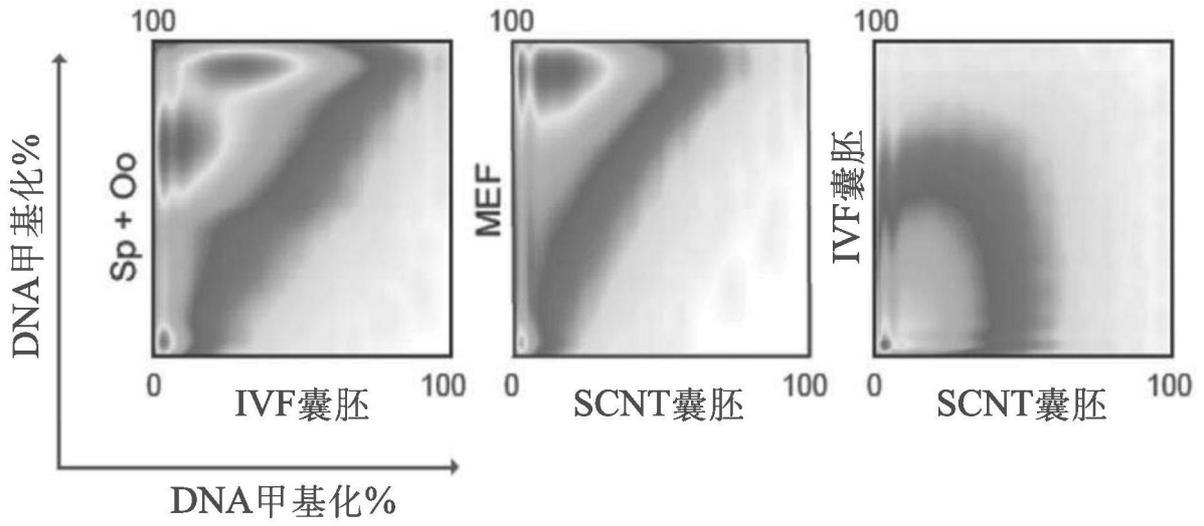


图3C

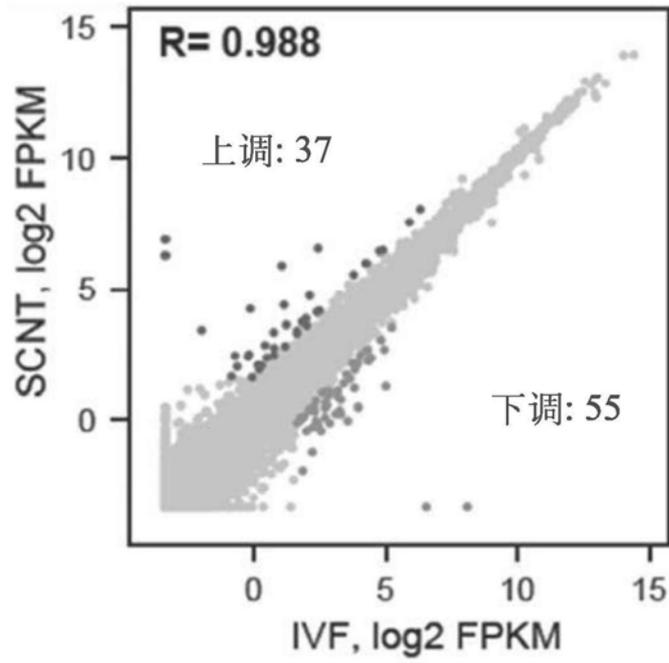


图3D

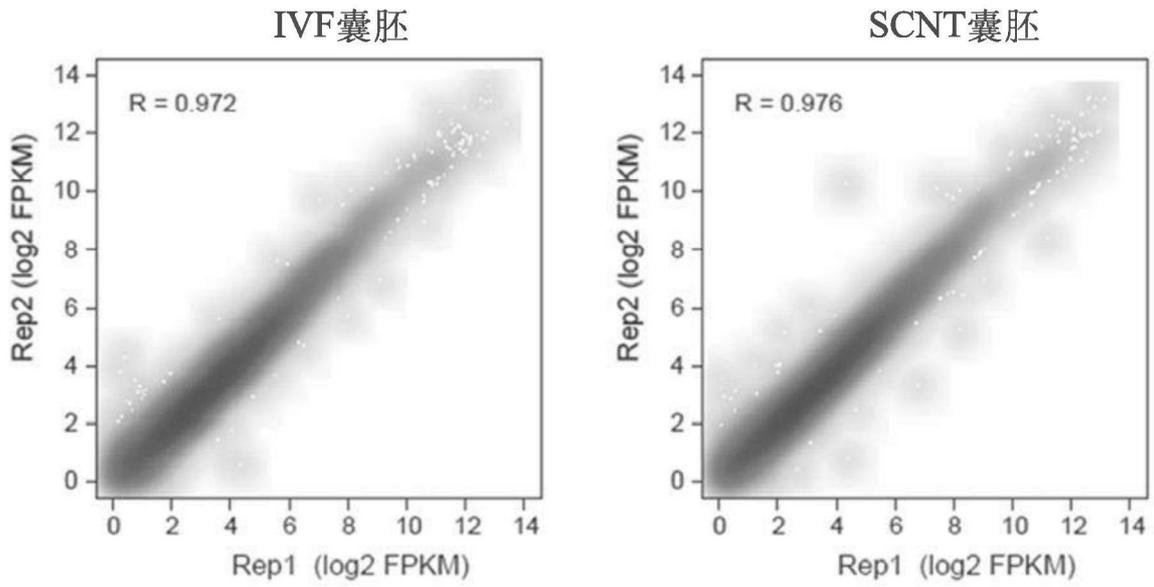


图4B

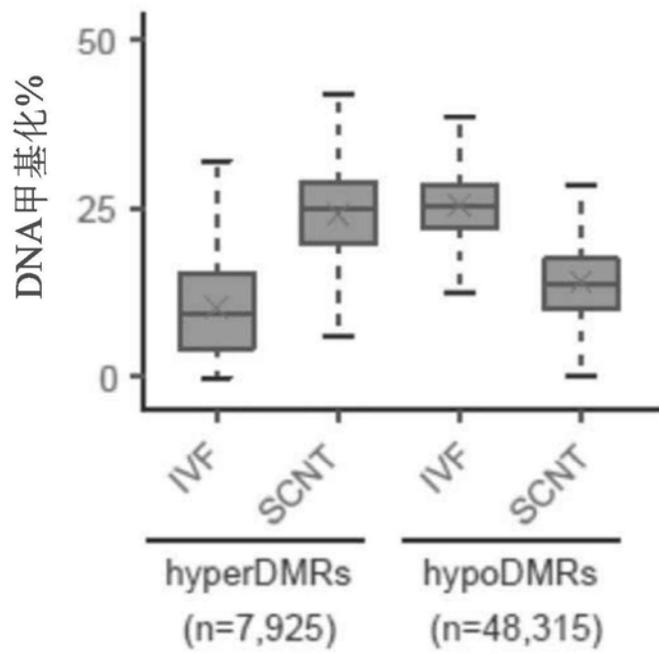


图5A

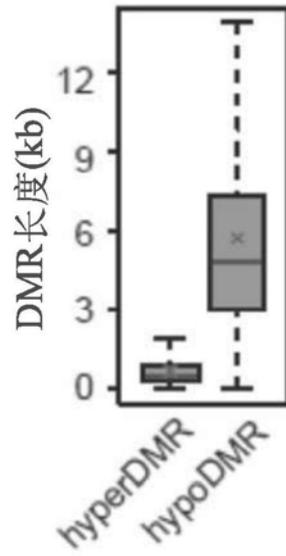


图5B

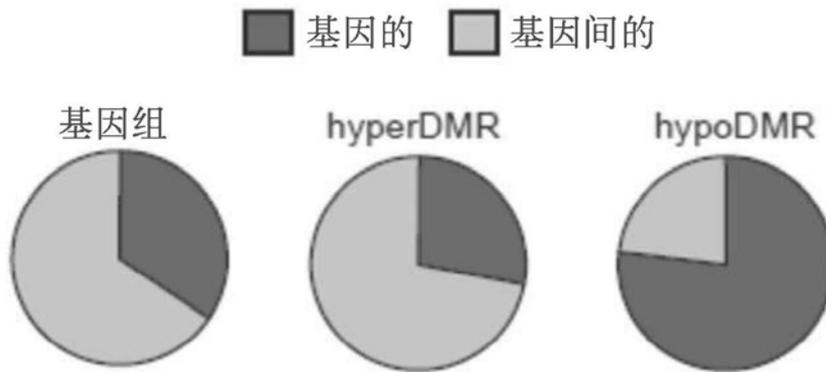


图5C

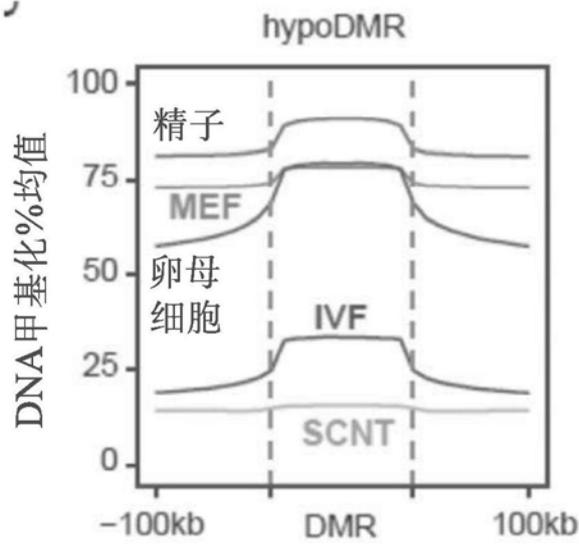


图5D

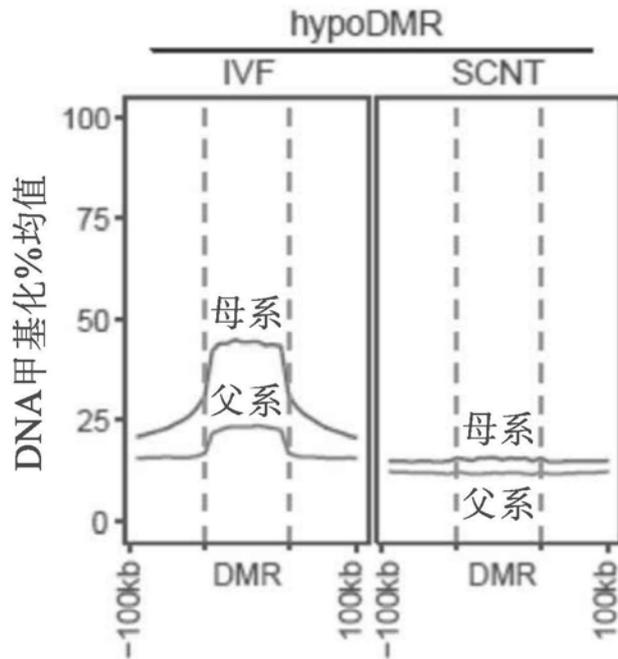


图5E

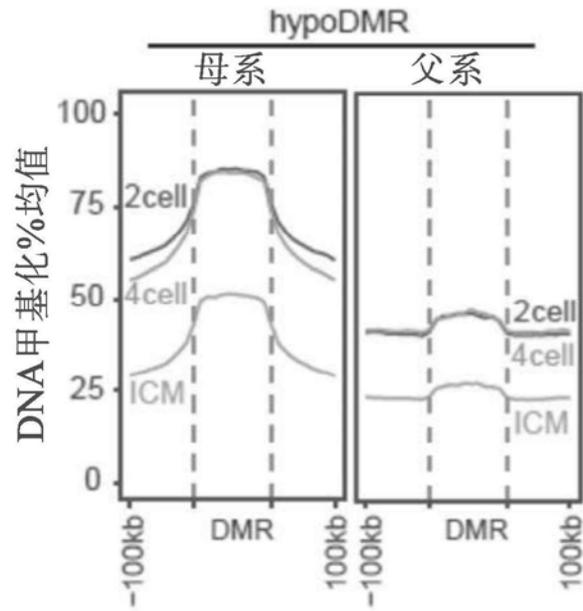


图5F

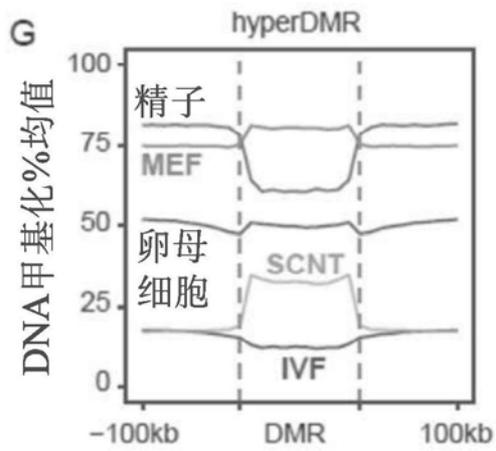


图5G

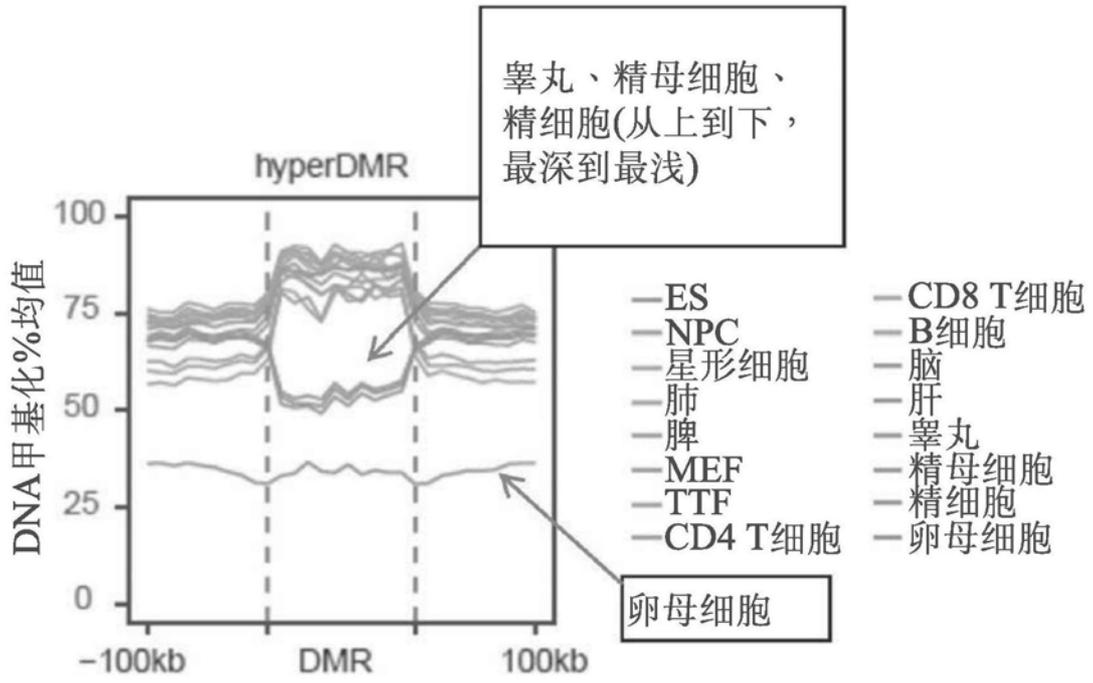


图5H

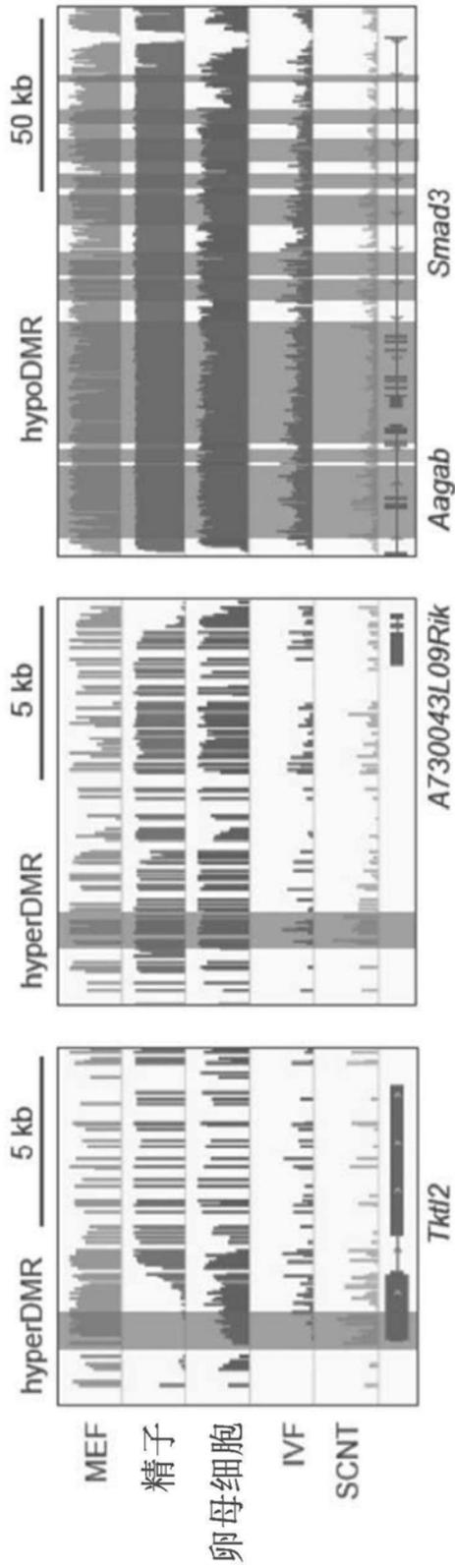


图6A

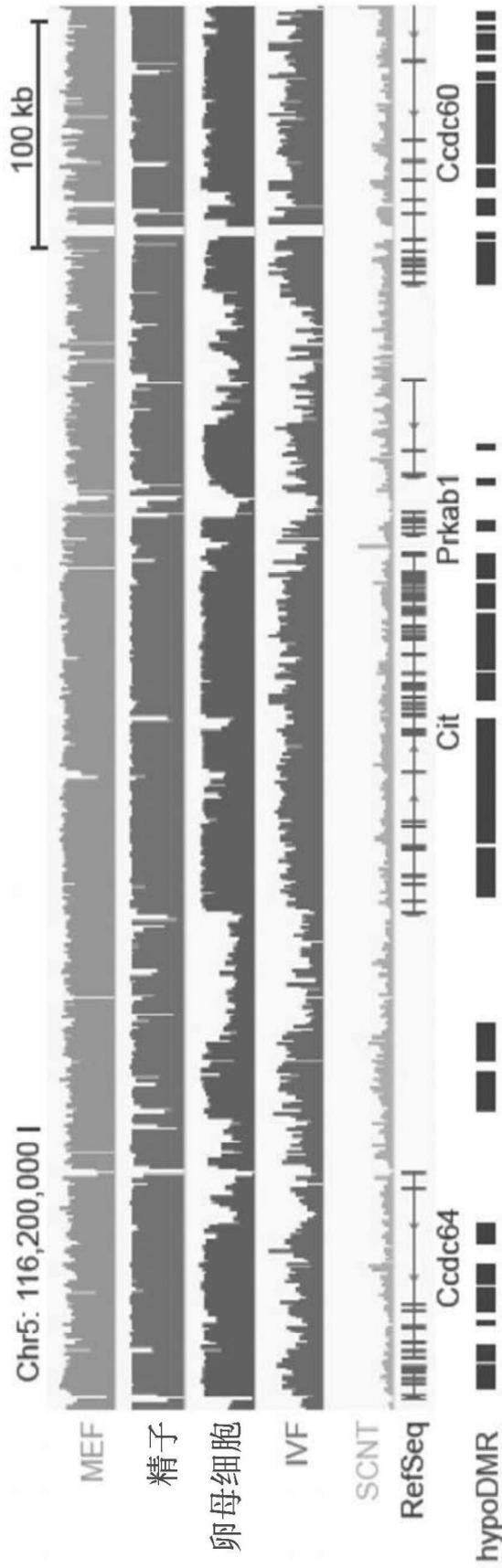


图6B

hyperDMR 基因(419)

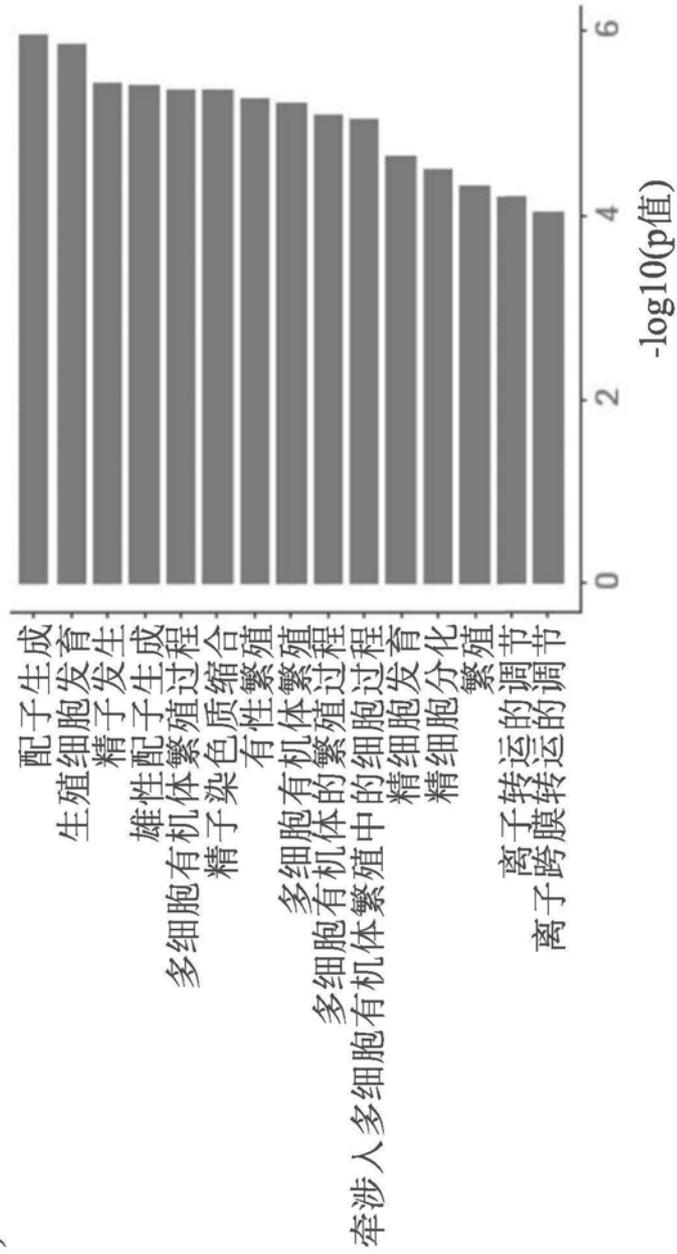


图6C

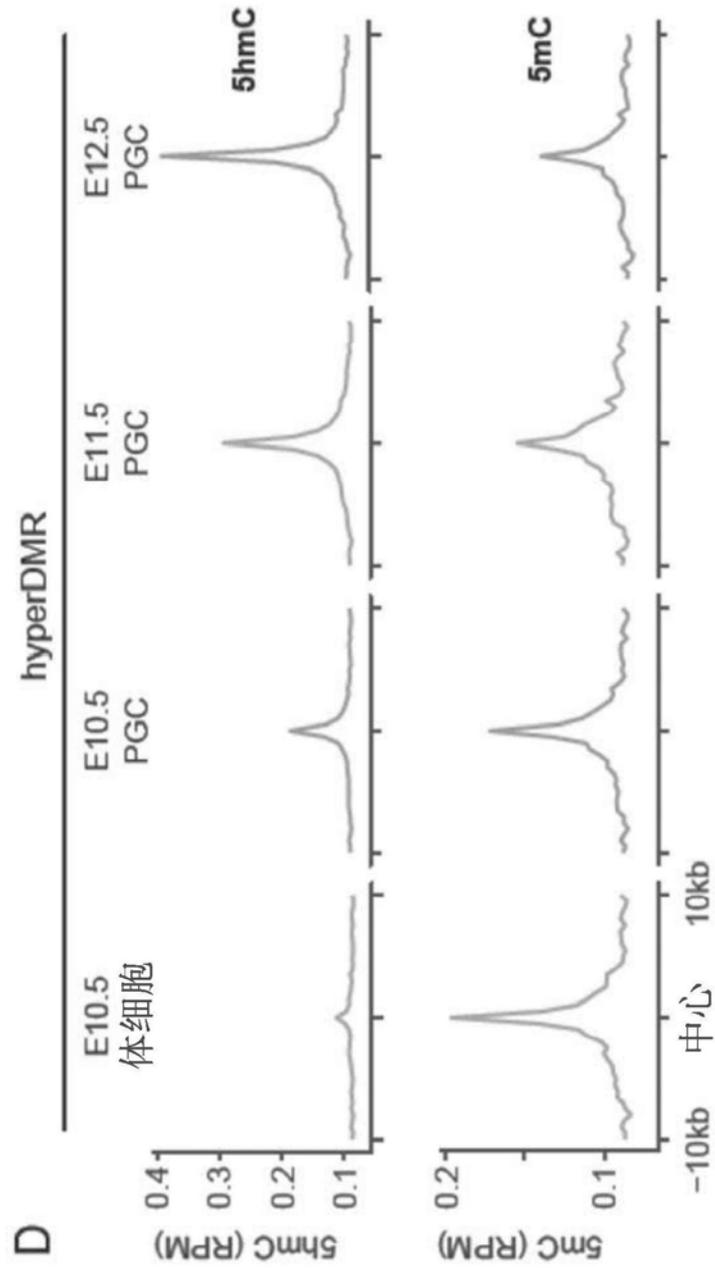


图6D

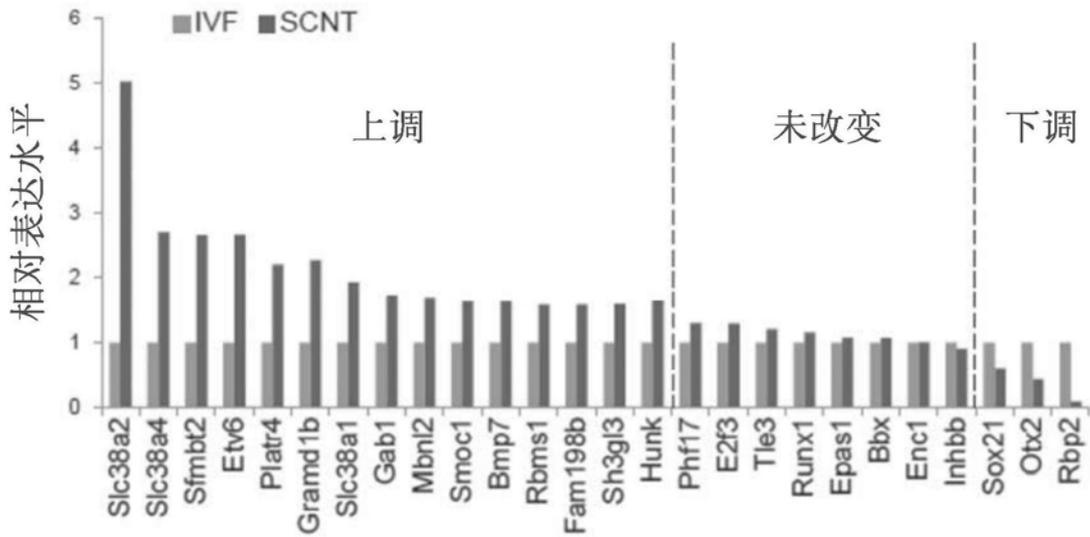


图7A

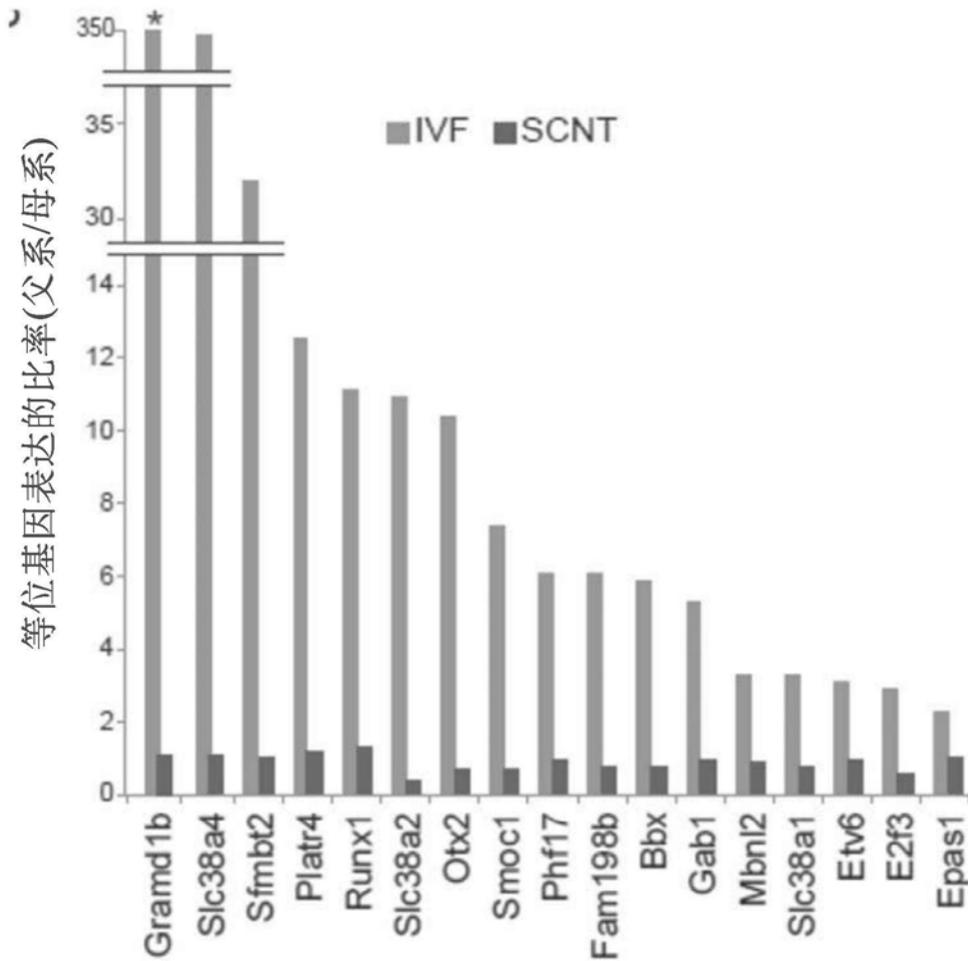


图7B

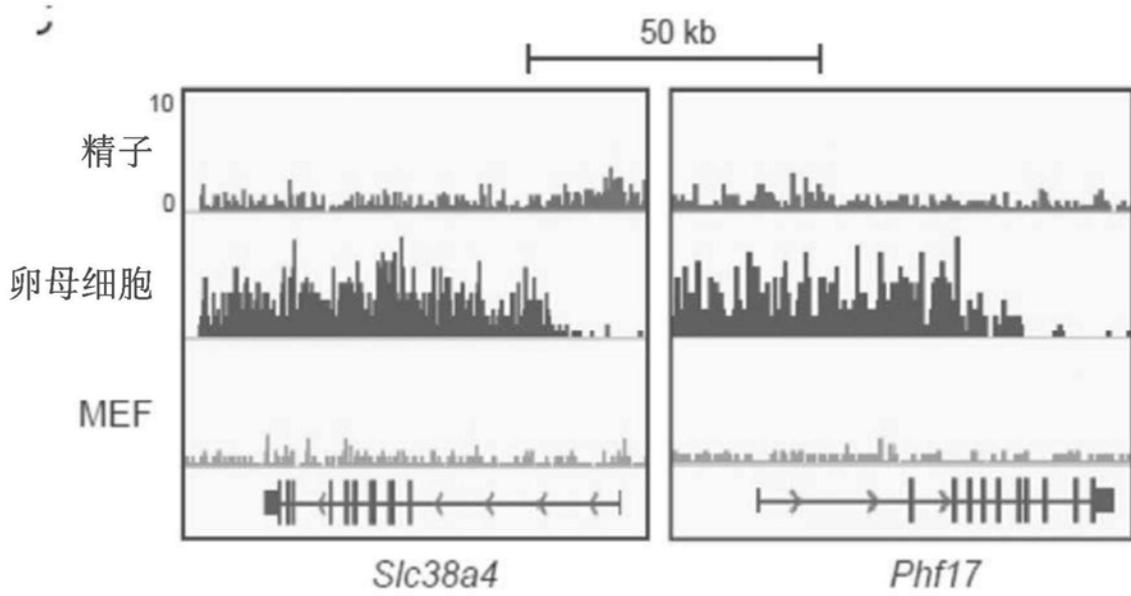


图7C

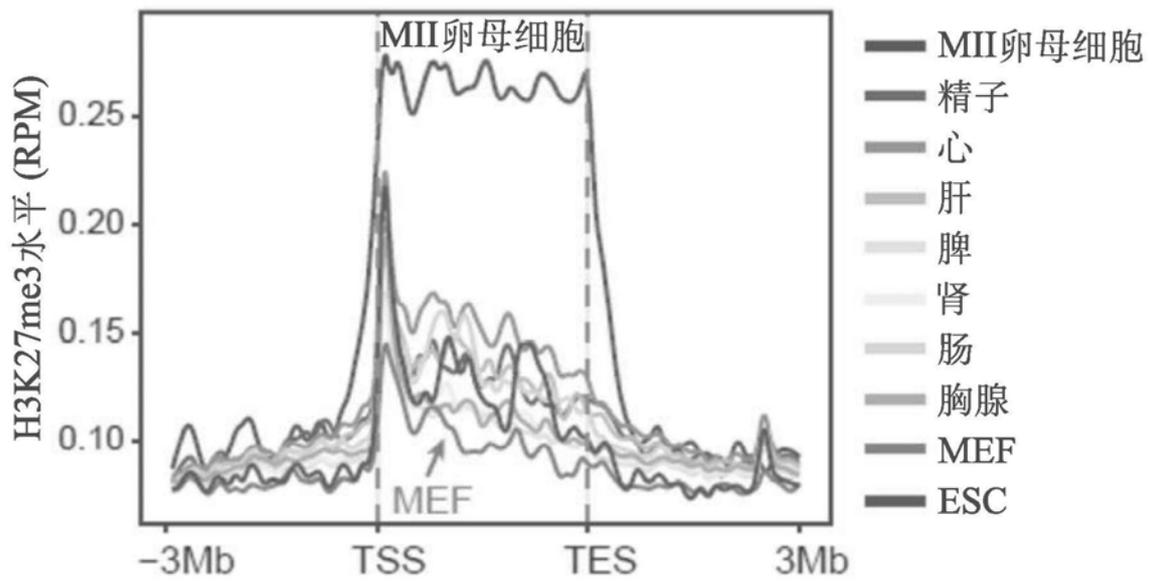


图7D

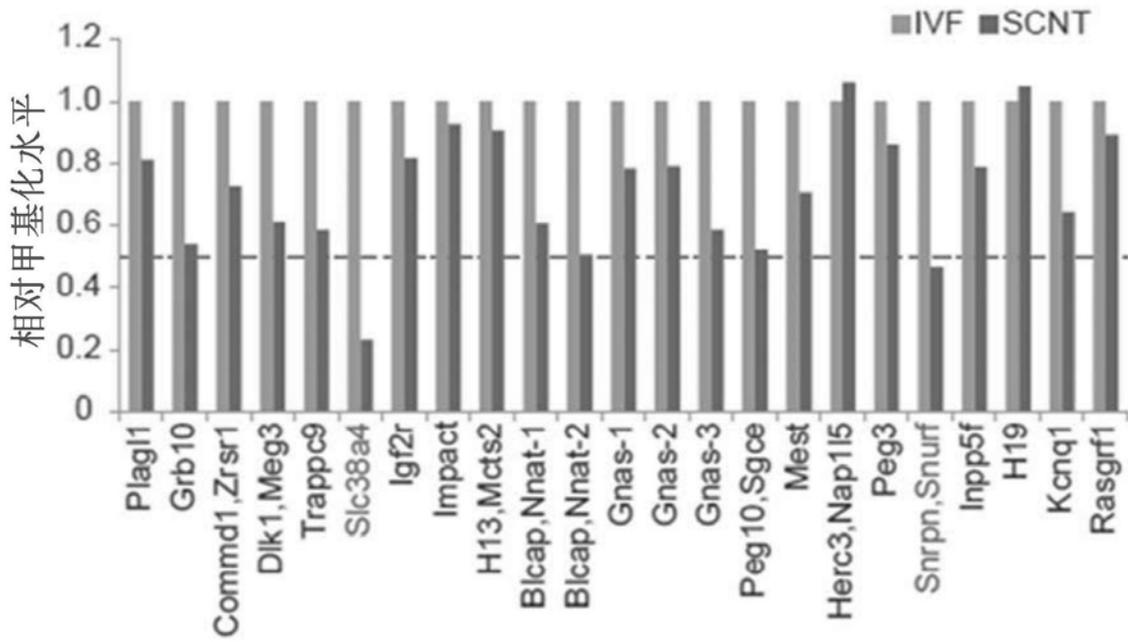


图8A

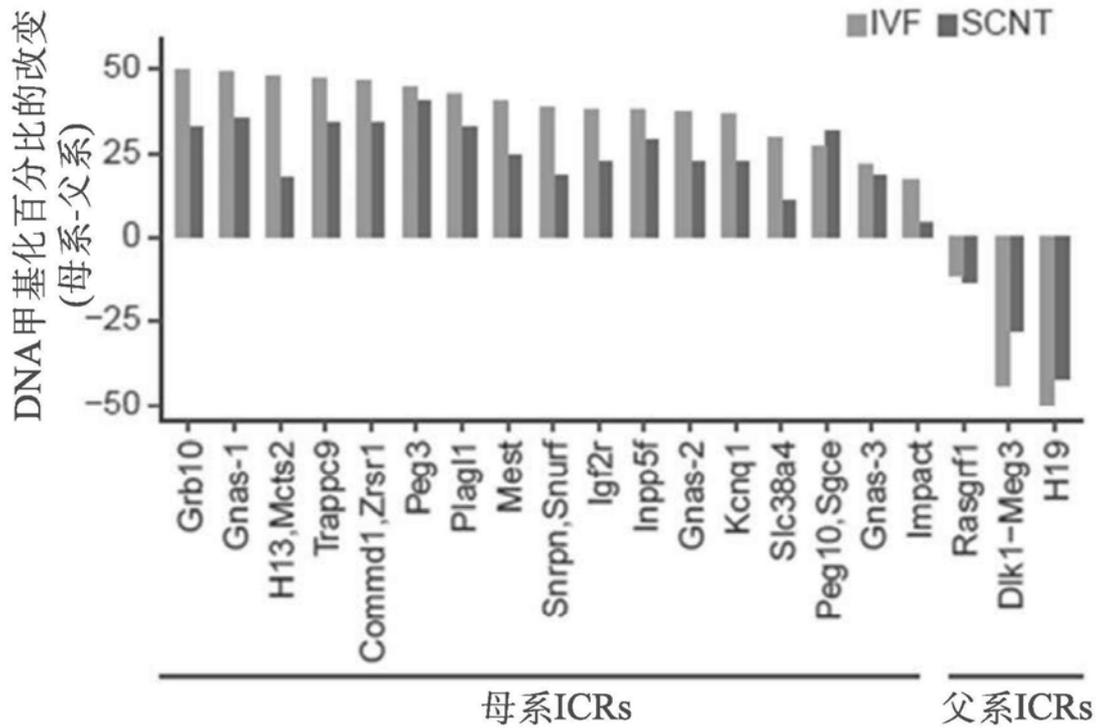


图8B

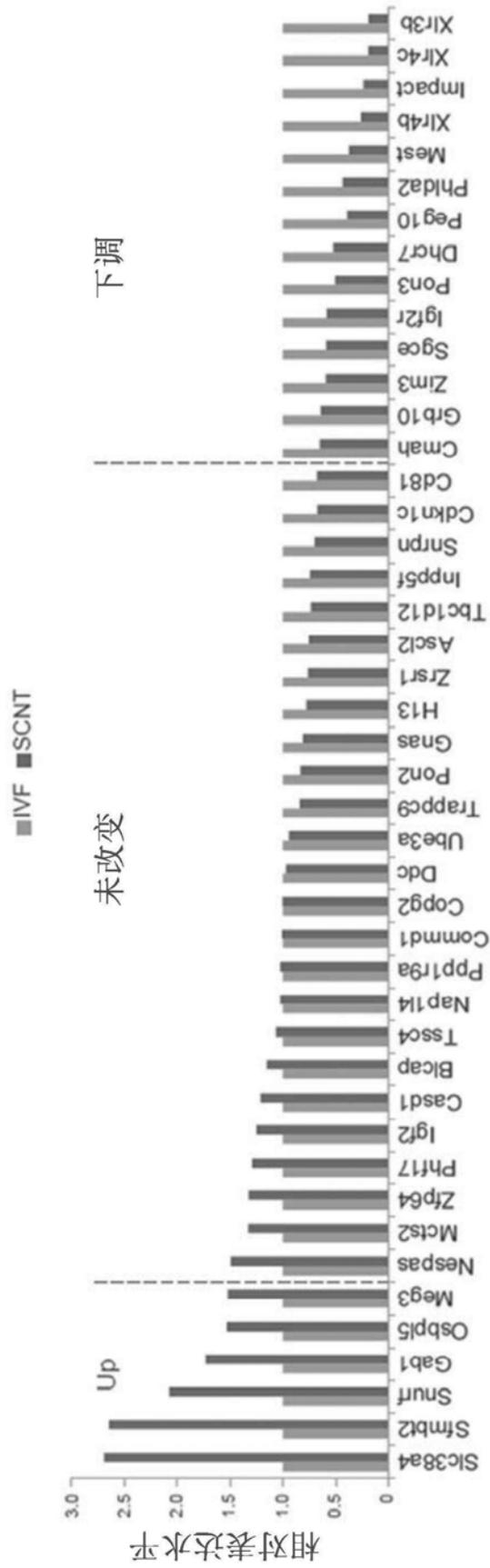


图8C

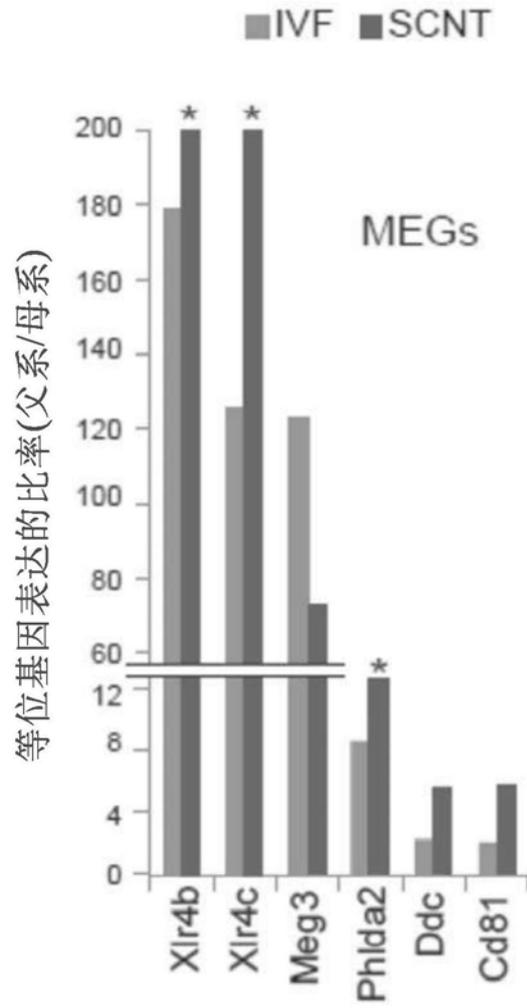


图8D

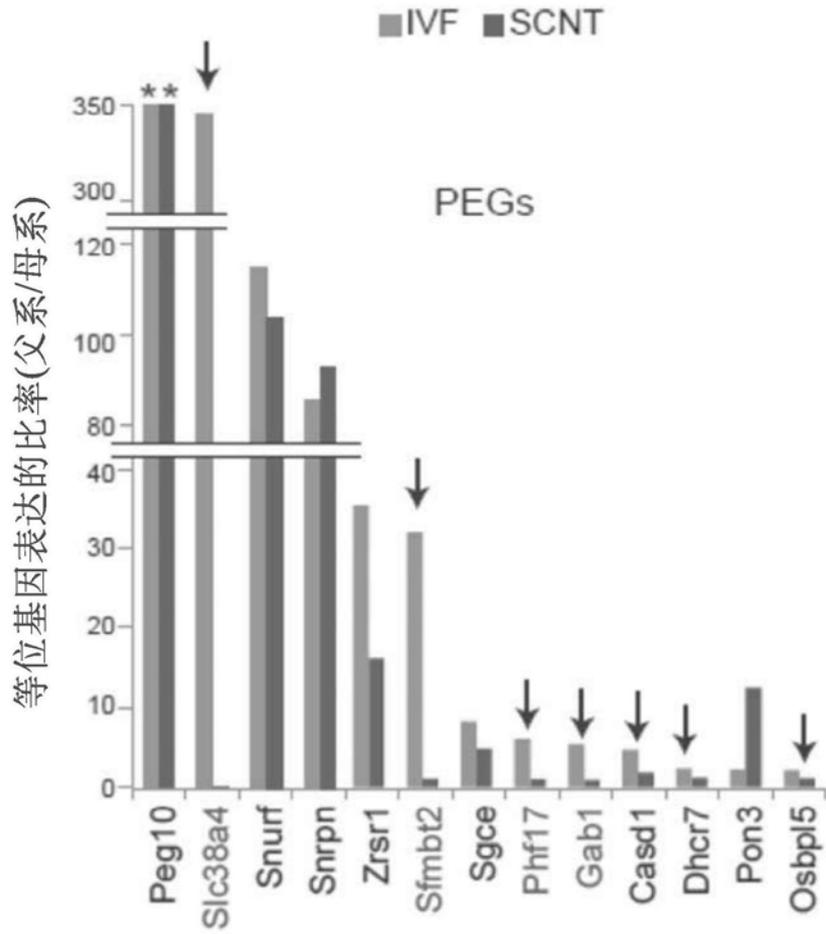


图8E

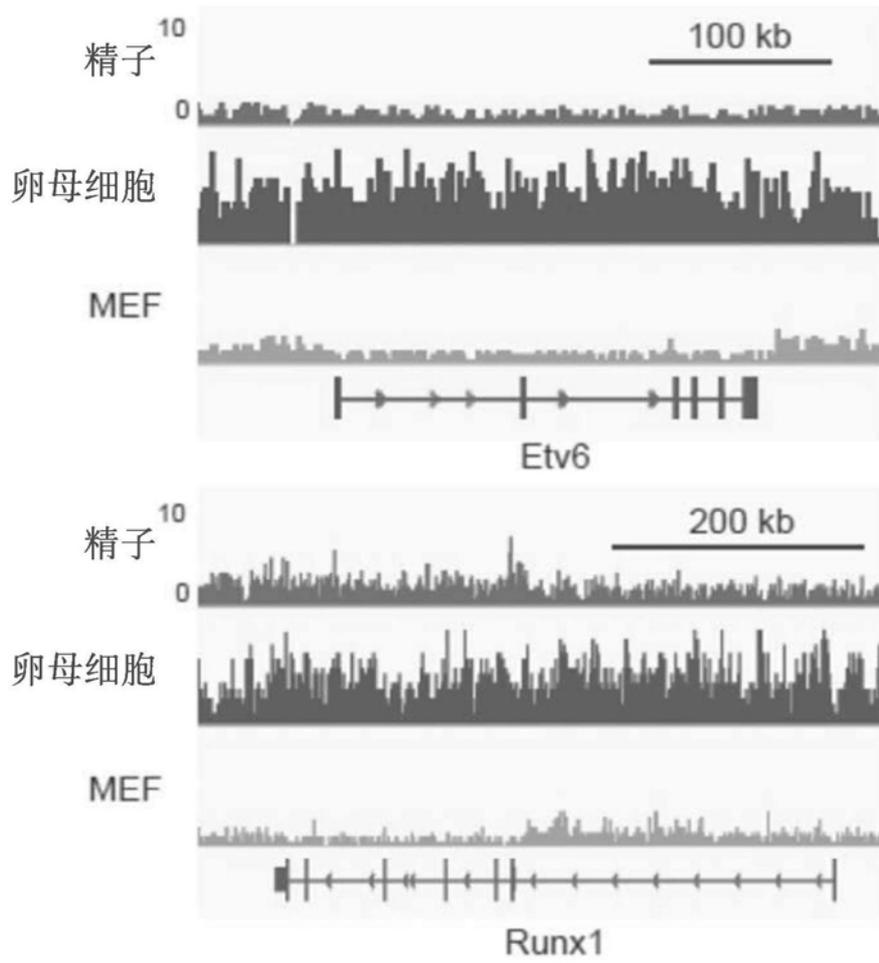


图8F