



(12)发明专利申请

(10)申请公布号 CN 110366413 A

(43)申请公布日 2019.10.22

(21)申请号 201880014459.3

(22)申请日 2018.02.02

(30)优先权数据

62/453,929 2017.02.02 US

62/479,878 2017.03.31 US

(85)PCT国际申请进入国家阶段日
2019.08.28

(86)PCT国际申请的申请数据

PCT/US2018/016562 2018.02.02

(87)PCT国际申请的公布数据

W02018/144798 EN 2018.08.09

(71)申请人 EPIZYME股份有限公司

地址 美国马萨诸塞州

(72)发明人 斯科特·理比希

(74)专利代理机构 北京派特恩知识产权代理有限公司 11270

代理人 薛恒 王琳

(51)Int.Cl.

A61K 31/33(2006.01)

A61K 31/35(2006.01)

A61K 31/40(2006.01)

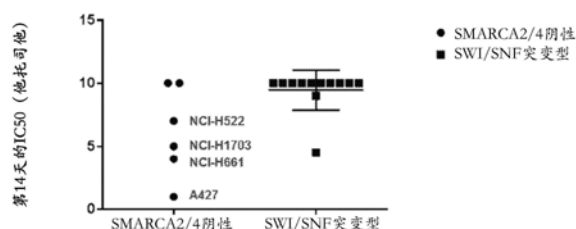
权利要求书9页 说明书96页
序列表74页 附图3页

(54)发明名称

癌症治疗形式

(57)摘要

本披露提供了可用于治疗病症,例如增生性病症,诸如某一癌症的治疗形式,例如,策略、治疗方法、患者分层方法、组合和组合物。本披露的一些方面提供了用于用EZH2抑制剂治疗依赖于EZH2(zeste增强子2多梳蛋白抑制复合物2)功能的细胞增生性病症例如癌症的治疗形式、方法、策略、组合物、组合和剂型。



1. 一种方法,该方法包括将zeste增强子同源物2 (EZH2) 抑制剂给予至患有或被诊断为患有细胞增生性病症的受试者,该细胞增生性病症的特征在于展现出SMARCA2和/或SMARCA4的功能丧失的细胞或细胞群。

2. 一种在有需要的受试者中治疗细胞增生性病症的方法,该方法包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂,其中该细胞增生性病症的特征在于展现出SMARCA2和/或SMARCA4的功能丧失的细胞或细胞群。

3. 如权利要求1或2中任一项所述的方法,其中,该细胞增生性病症是肺的细胞增生性病症。

4. 一种在有需要的受试者中治疗肺的细胞增生性病症的方法,该方法包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂。

5. 如权利要求4所述的方法,其中,该细胞增生性病症包括或其特征在于展现出SMARCA2的功能丧失和/或SMARCA4的功能丧失的细胞或细胞群。

6. 如权利要求1、2或4中任一项所述的方法,其中,该细胞增生性病症包括或其特征在于展现出SMARCA2和SMARCA4的功能丧失的细胞或细胞群。

7. 如权利要求1、2或4中任一项所述的方法,其中,该细胞增生性病症的特征在于干细胞、干细胞样细胞或祖细胞起源。

8. 如权利要求1、2或4中任一项所述的方法,其中,该肺的细胞增生性病症的特征在于肺中的恶性生长或病灶。

9. 如权利要求8所述的方法,其中,该恶性生长或病灶是原发性病灶。

10. 如权利要求8所述的方法,其中,该恶性生长或病灶是、或其特征在于继发性或转移性病灶。

11. 如权利要求8所述的方法,其中,该恶性生长是恶性肿瘤、癌或类癌肿瘤。

12. 如权利要求1、2或4中任一项所述的方法,其中,该肺的细胞增生性病症是石棉诱导的增生、鳞状化生、和良性反应性间皮化生。

13. 如权利要求1、2或4中任一项所述的方法,其中,该肺的细胞增生性病症是肺癌。

14. 如权利要求13所述的方法,其中,该肺癌是小细胞肺癌。

15. 如权利要求13所述的方法,其中,该肺癌是非小细胞肺癌。

16. 如权利要求13所述的方法,其中,该肺癌是鳞状细胞癌。

17. 如权利要求13所述的方法,其中,该肺癌是腺癌。

18. 如权利要求13所述的方法,其中,该肺癌是小细胞癌。

19. 如权利要求13所述的方法,其中,该肺癌是大细胞癌。

20. 如权利要求13所述的方法,其中,该肺癌是腺鳞状细胞癌。

21. 如权利要求13所述的方法,其中,该肺癌是间皮瘤。

22. 如权利要求1、2或4中任一项所述的方法,其中,该细胞增生性疾病的特征在于原发性肿瘤,其中该原发性肿瘤

(A) 展现出SMARCA2/SMARCA4双重丧失;并且

(B) 是低分化的和/或展现出上皮间充质转化(EMT)特征。

23. 如权利要求22所述的方法,其中,该原发性肿瘤展现出低的E-钙粘蛋白表达水平和高的波形蛋白表达水平。

24. 如权利要求1、2或4中任一项所述的方法,其中,该受试者已经或正在被给予另外的治疗剂,该治疗剂的给予与该EZH2抑制剂的给予同时或在时间上接近。

25. 如权利要求24所述的方法,其中,该另外的治疗剂是标准护理剂。

26. 如权利要求25所述的方法,其中,该另外的药剂是或包括在示意图1中列出的药剂,或者是或包括在示意图1中列出的两种或更多种药剂的组合。

27. 如权利要求24所述的方法,其中,该另外的治疗剂是免疫检查点抑制剂。

28. 如权利要求27所述的方法,其中,该免疫检查点抑制剂是CTLA4抑制剂、PD-1抑制剂和/或PD-L1抑制剂、LAG3抑制剂、B7-H3抑制剂或Tim3抑制剂。

29. 如权利要求28所述的方法,其中,该免疫检查点抑制剂包括伊匹单抗、替西木单抗、AGEN-1884、尼沃鲁单抗、派姆单抗、阿特珠单抗、德瓦鲁单抗、阿维单抗、BMS-936559、AMP-224、MEDI-0680、TSR-042、BGB-108、STI-1014、KY-1003、ALN-PDL、BGB-A317、KD-033、REGN-2810、PDR-001、SHR-1210、MGD-013、PF-06801591、CX-072、IMP-731、LAG-525、BMS-986016、GSK-2831781、依诺妥珠单抗、1241-8H9、DS-5573、MBG-453、或其组合。

30. 如权利要求24所述的方法,其中,该EZH2抑制剂和该另外的治疗剂依序给予至该受试者。

31. 如权利要求24所述的方法,其中,该EZH2抑制剂和该另外的治疗剂经由不同的给予途径且以不同的间隔给予。

32. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂每天两次口服给予。

33. 如权利要求1、2或4中任一项所述的方法,其中,该方法进一步包括检测SMARCA2和/或SMARCA4蛋白表达和/或SMARCA2和/或SMARCA4蛋白的功能。

34. 如权利要求33所述的方法,其中,该SMARCA2和/或该SMARCA4蛋白的表达和/或功能通过包括以下的方法评价:

(a) 从该受试者获得生物样品;

(b) 将该生物样品或其部分与特异性结合SMARCA2或SMARCA4的抗体接触;并且

(c) 检测与SMARCA2或SMARCA4结合的该抗体的量。

35. 如权利要求1、2或4中任一项所述的方法,其中,该方法进一步包括检测获自该受试者的生物样品中的编码该SMARCA2的基因和/或编码该SMARCA4蛋白的基因中的基因组突变。

36. 如权利要求35所述的方法,其中,该基因组突变通过包括以下的方法检测:

(a) 从该受试者获得生物样品;

(b) 对该生物样品中的编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或编码SMARCA4蛋白或其部分的至少一种DNA序列进行测序;并且

(c) 确定该编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或该编码SMARCA4蛋白或其部分的至少一种DNA序列是否包含影响该SMARCA2蛋白或该SMARCA4蛋白的表达和/或功能的突变。

37. 如权利要求1或2中任一项所述的方法,其中,该EZH2抑制剂抑制组蛋白3的赖氨酸27 (H3K27) 的三甲基化。

38. 一种方法,该方法包括检测获自受试者的样品中的SMARCA2和/或SMARCA4功能丧失。

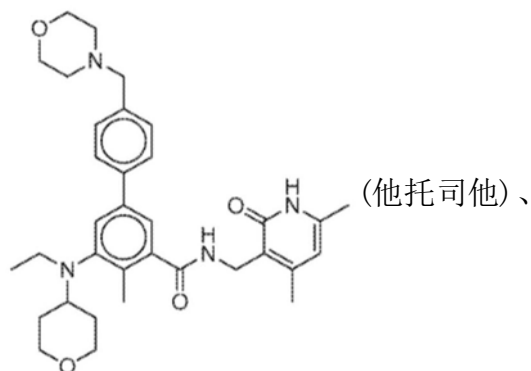
39. 如权利要求38所述的方法, 其中, 该受试者患有癌症。

40. 如权利要求38或39中任一项所述的方法, 其中, 如果在该受试者中检测到SMARCA2和/或SMARCA4功能丧失, 则该方法进一步包括向该受试者给予EZH2抑制剂。

41. 如权利要求40所述的方法, 其中, 该SMARCA2功能丧失与编码SMARCA2蛋白的基因中的基因组突变不相关, 和/或其中该SMARCA4功能丧失与编码SMARCA4的基因中的基因组突变相关。

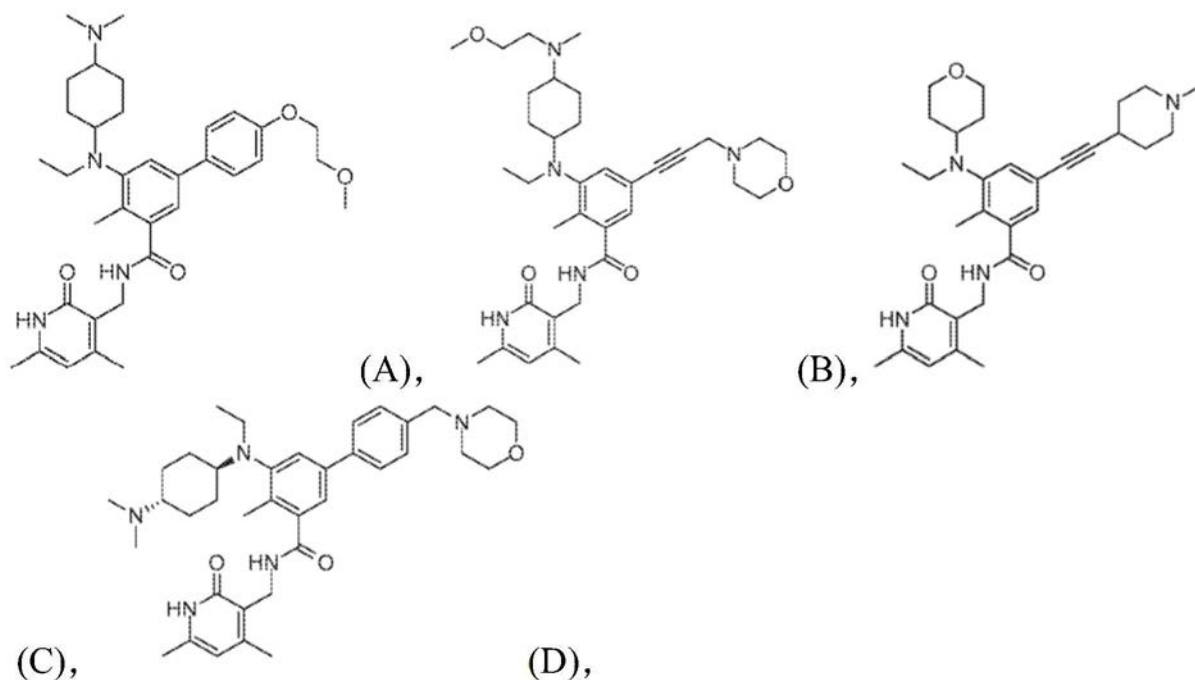
42. 如权利要求38所述的方法, 其中, 该受试者患有NSCLC。

43. 如权利要求1、2或4中任一项所述的方法, 其中, 该EZH2抑制剂是



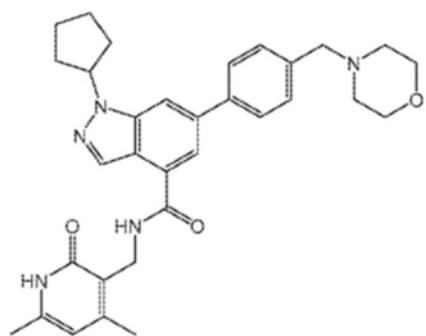
或其药学上可接受的盐。

44. 如权利要求1、2或4中任一项所述的方法, 其中, 该EZH2抑制剂是



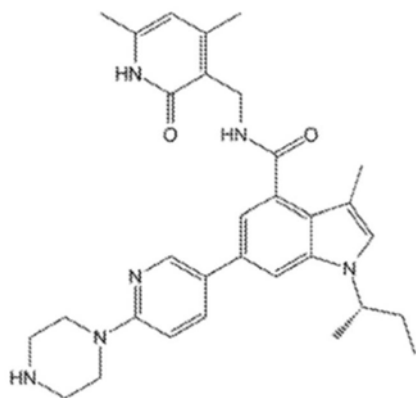
其立体异构体、药学上可接受的盐和/或溶剂化物。

45. 如权利要求1、2或4中任一项所述的方法, 其中, 该EZH2抑制剂是



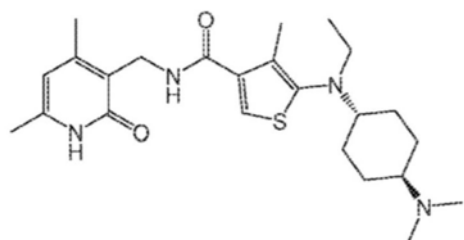
或其药学上可接受的盐。

46.如权利要求1、2或4中任一项所述的方法，其中，该EZH2抑制剂是



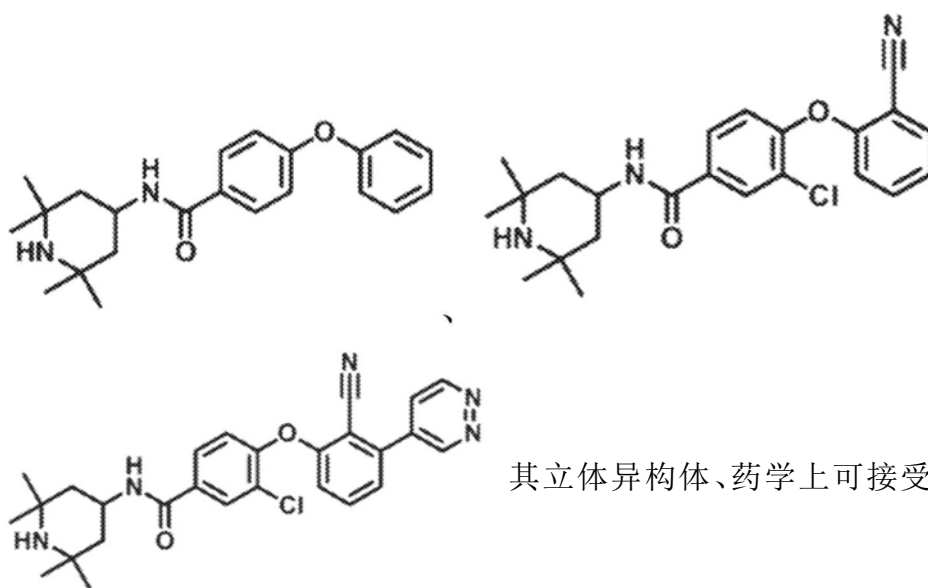
其立体异构体、药学上可接受的盐和/或溶剂化物。

47.如权利要求1、2或4中任一项所述的方法，其中，该EZH2抑制剂是



其立体异构体、药学上可接受的盐和/或溶剂化物。

48.如权利要求1、2或4中任一项所述的方法，其中，该EZH2抑制剂是



其立体异构体、药学上可接受的盐和/或溶剂化

物。

49. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂是口服给予的。
50. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂被配制为口服片剂。
51. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂以10mg/kg/天与1600mg/kg/天之间的剂量给予。
52. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂以约100、200、400、800、或1600mg的剂量给予。
53. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂以约800mg的剂量给予。
54. 如权利要求1、2或4中任一项所述的方法,其中,该EZH2抑制剂每天两次(BID)给予。
55. zeste增强子同源物2 (EZH2) 抑制剂用于在有需要的受试者中治疗细胞增生性病征的用途,该用途包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂,其中该细胞增生性病征的特征在于展现出SMARCA2和/或SMARCA4的功能丧失的细胞或细胞群。
56. 如权利要求55所述的用途,其中,该细胞增生性病征是肺的细胞增生性病征。
57. zeste增强子同源物2 (EZH2) 抑制剂用于在有需要的受试者中治疗肺的细胞增生性病征的用途,该用途包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂。
58. 如权利要求57所述的用途,其中,该细胞增生性病征包括或其特征在于展现出SMARCA2的功能丧失和/或SMARCA4的功能丧失的细胞或细胞群。
59. 如权利要求55-58中任一项所述的用途,其中,该细胞增生性病征包括或其特征在于展现出SMARCA2和SMARCA4的功能丧失的细胞或细胞群。
60. 如权利要求55-59中任一项所述的用途,其中,该细胞增生性病征的特征在于干细胞、干细胞样细胞或祖细胞起源。
61. 如权利要求55-60中任一项所述的用途,其中,该肺的细胞增生性病征的特征在于肺中的恶性生长或病灶。
62. 如权利要求55-61中任一项所述的用途,其中,该恶性生长或病灶是原发性病灶。
63. 如权利要求55-62中任一项所述的用途,其中,该恶性生长或病灶是、或其特征在于继发性或转移性病灶。
64. 如权利要求55-63中任一项所述的用途,其中,该恶性生长是恶性肺肿瘤、癌或类癌肿瘤。
65. 如权利要求55-64中任一项所述的用途,其中,该肺的细胞增生性病征是石棉诱导的增生、鳞状化生、和良性反应性间皮化生。
66. 如权利要求55-65中任一项所述的用途,其中,该肺的细胞增生性病征是肺癌。
67. 如权利要求66所述的用途,其中,该肺癌是小细胞肺癌。
68. 如权利要求66所述的用途,其中,该肺癌是非小细胞肺癌。
69. 如权利要求66所述的用途,其中,该肺癌是鳞状细胞癌。
70. 如权利要求66所述的用途,其中,该肺癌是腺癌。
71. 如权利要求66所述的用途,其中,该肺癌是小细胞癌。
72. 如权利要求66所述的用途,其中,该肺癌是大细胞癌。

73. 如权利要求66所述的用途,其中,该肺癌是腺鳞状细胞癌。
74. 如权利要求66所述的用途,其中,该肺癌是间皮瘤。
75. 如权利要求55-74中任一项所述的用途,其中,该细胞增生性疾病的特征在于原发性肿瘤,其中该原发性肿瘤
- (A) 展现出SMARCA2/SMARCA4双重丧失;并且
 - (B) 是低分化的和/或展现出上皮间充质转化(EMT)特征。
76. 如权利要求75所述的用途,其中,该原发性肿瘤展现出低的E-钙粘蛋白表达水平和高的波形蛋白表达水平。
77. 如权利要求55-76中任一项所述的用途,其中,该受试者已经或正在被给予另外的治疗剂,该治疗剂的给予与该EZH2抑制剂的给予同时或在时间上接近。
78. 如权利要求77所述的用途,其中,该另外的治疗剂是标准护理剂。
79. 如权利要求78所述的用途,其中,该另外的药剂是或包括在示意图1中列出的药剂,或者是或包括在示意图1中列出的两种或更多种药剂的组合。
80. 如权利要求79所述的用途,其中,该另外的治疗剂是免疫检查点抑制剂。
81. 如权利要求80所述的用途,其中,该免疫检查点抑制剂是CTLA4抑制剂、PD-1抑制剂和/或PD-L1抑制剂、LAG3抑制剂、B7-H3抑制剂或Tim3抑制剂。
82. 如权利要求81所述的用途,其中,该免疫检查点抑制剂包括伊匹单抗、替西木单抗、AGEN-1884、尼沃鲁单抗、派姆单抗、阿特殊单抗、德瓦鲁单抗、阿维单抗、BMS-936559、AMP-224、MEDI-0680、TSR-042、BGB-108、STI-1014、KY-1003、ALN-PDL、BGB-A317、KD-033、REGN-2810、PDR-001、SHR-1210、MGD-013、PF-06801591、CX-072、IMP-731、LAG-525、BMS-986016、GSK-2831781、依诺妥珠单抗、1241-8H9、DS-5573、MBG-453、或其组合。
83. 如权利要求77-82中任一项所述的用途,其中,该EZH2抑制剂和该另外的治疗剂依序给予至该受试者。
84. 如权利要求77-83中任一项所述的用途,其中,该EZH2抑制剂和该另外的治疗剂经由不同的给予途径且以不同的间隔给予。
85. 如权利要求55-84中任一项所述的用途,其中,该EZH2抑制剂每天两次口服给予。
86. 如权利要求55-85中任一项所述的用途,其中,该用途进一步包括检测SMARCA2和/或SMARCA4蛋白表达和/或SMARCA2和/或SMARCA4蛋白的功能。
87. 如权利要求86所述的用途,其中,该SMARCA2和/或该SMARCA4蛋白的表达和/或功能通过包括以下的步骤评价:
- (a) 从该受试者获得生物样品;
 - (b) 将该生物样品或其部分与特异性结合SMARCA2或SMARCA4的抗体接触;并且
 - (c) 检测与SMARCA2或SMARCA4结合的该抗体的量。
88. 如权利要求55-87中任一项所述的用途,其中,该用途进一步包括检测获自该受试者的生物样品中的编码该SMARCA2的基因和/或编码该SMARCA4蛋白的基因中的基因组突变。
89. 如权利要求88所述的用途,其中,该基因组突变通过包括以下的步骤检测:
- (a) 从该受试者获得生物样品;
 - (b) 对该生物样品中的编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或编码

SMARCA4蛋白或其部分的至少一种DNA序列进行测序;并且

(c) 确定该编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或该编码SMARCA4蛋白或其部分的至少一种DNA序列是否包含影响该SMARCA2蛋白或该SMARCA4蛋白的表达和/或功能的突变。

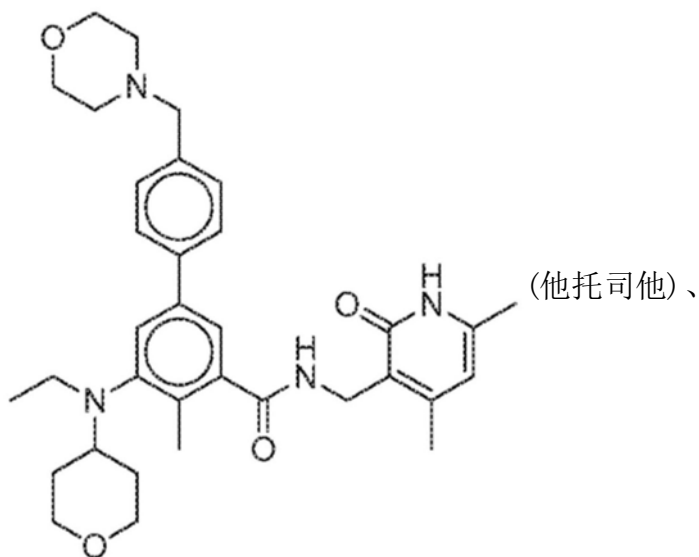
90. 如权利要求55所述的用途, 其中, 该EZH2抑制剂抑制组蛋白3的赖氨酸27 (H3K27) 的三甲基化。

91. 如权利要求88-89中任一项所述的用途, 其中, 该用途进一步包括检测获自受试者的样品中的SMARCA2和/或SMARCA4功能丧失。

92. 如权利要求91所述的用途, 其中, 该SMARCA2功能丧失与编码SMARCA2蛋白的基因中的基因组突变不相关, 和/或其中该SMARCA4功能丧失与编码SMARCA4的基因中的基因组突变相关。

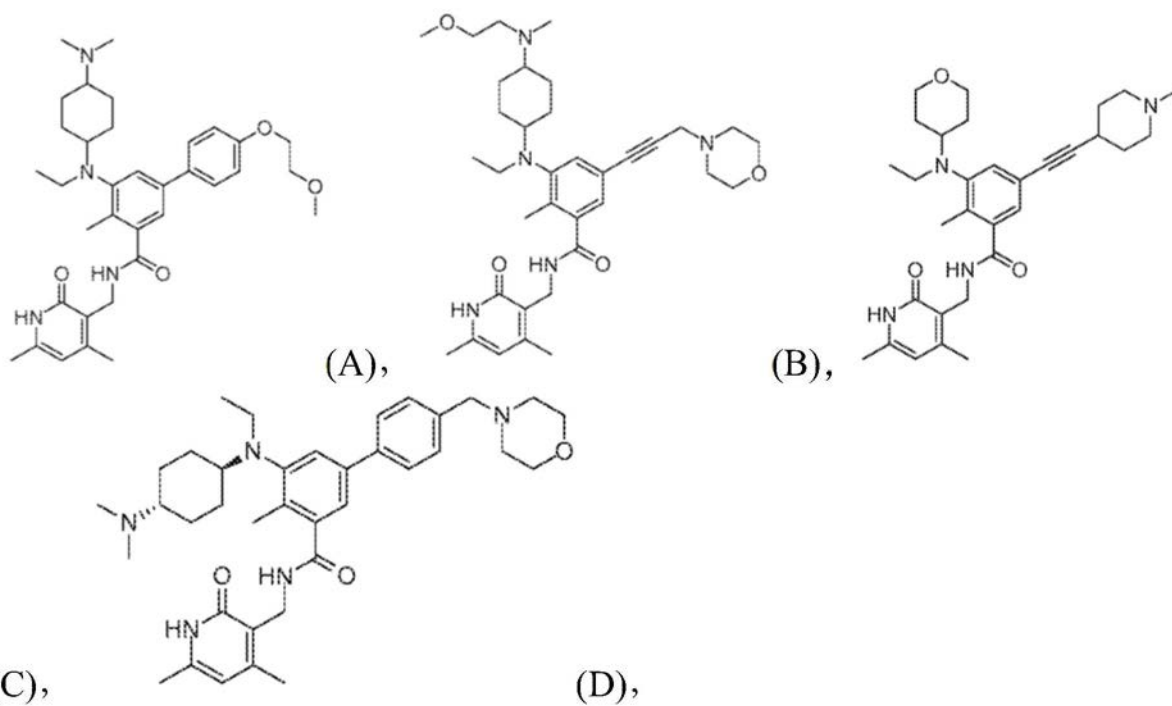
93. 如权利要求91-92中任一项所述的用途, 其中, 该受试者患有NSCLC。

94. 如权利要求55-93中任一项所述的用途, 其中, 该EZH2抑制剂是



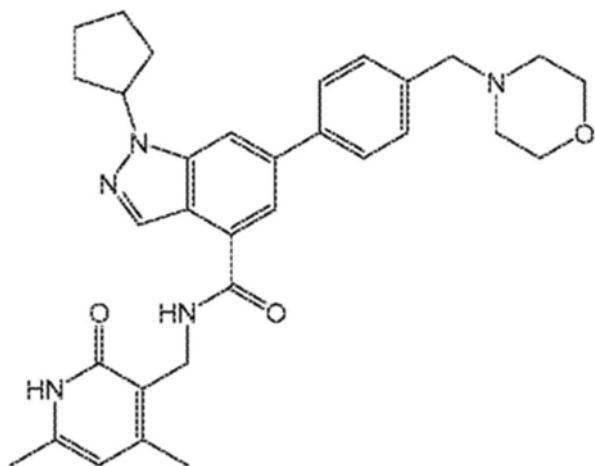
或其药学上可接受的盐。

95. 如权利要求55-93中任一项所述的用途, 其中, 该EZH2抑制剂是



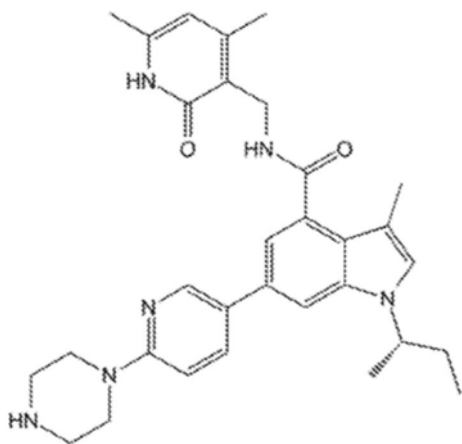
其立体异构体、药学上可接受的盐和/或溶剂化物。

96. 如权利要求55-93中任一项所述的用途, 其中, 该EZH2抑制剂是



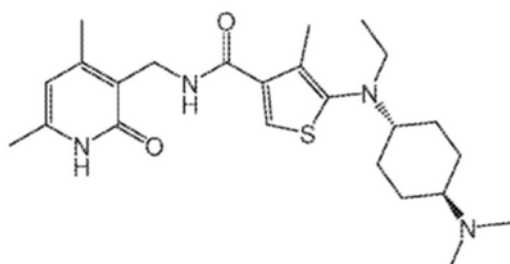
或其药学上可接受的盐。

97. 如权利要求55-93中任一项所述的用途, 其中, 该EZH2抑制剂是



其立体异构体、药学上可接受的盐和/或溶剂化物。

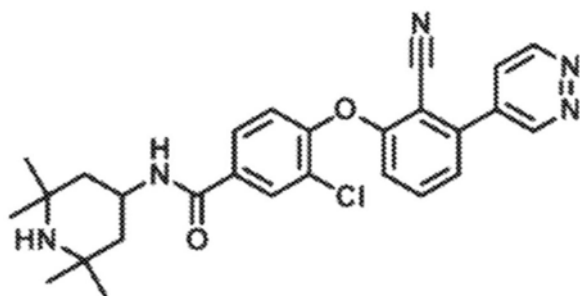
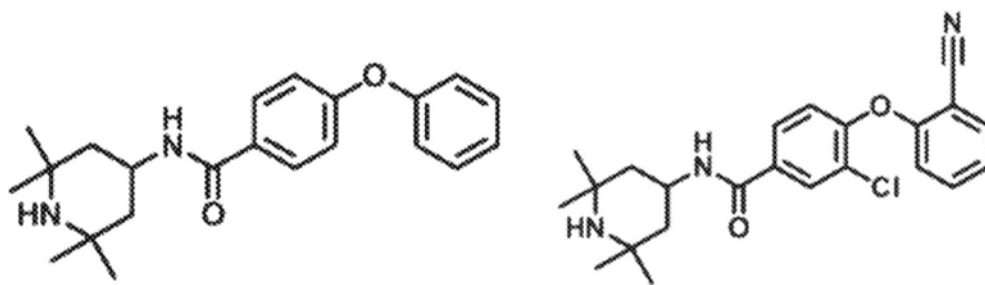
98. 如权利要求55-93中任一项所述的用途, 其中, 该EZH2抑制剂是



其立体异构体、药学上可接受的盐和/或溶剂化

物。

99. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂是



其立体异构体、药学上可接受的盐和/或溶

剂化物。

100. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂是口服给予的。

101. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂被配制为口服片剂。

102. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂以10mg/kg/天与1600mg/kg/天之间的剂量给予。

103. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂以约100、200、400、800、或1600mg的剂量给予。

104. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂以约800mg的剂量给予。

105. 如权利要求55-93中任一项所述的用途,其中,该EZH2抑制剂每天两次 (BID) 给予。

癌症治疗形式

[0001] 相关申请的交叉引用

[0002] 本申请要求2017年2月2日提交的美国临时专利申请序列号62/453,929和2017年3月31日提交的美国临时专利申请序列号62/479,878的权益。以上提及的申请中的每一个的全部内容通过援引以其全文并入本文。

[0003] 序列表

[0004] 本申请包含经由EFS-Web以ASCII格式提交并特此通过援引以其全文并入的序列表。创建于2018年1月17日的所述ASCII副本名为EPIZ-074001W0_ST25.txt且大小为196,906字节。

发明内容

[0005] 本披露提供了可用于治疗病症,例如增生性病症,诸如某一癌症的治疗形式,例如,策略、治疗方法、患者分层方法、组合和组合物。本披露的一些方面提供了用于用EZH2抑制剂治疗依赖于EZH2(zeste增强子2多梳蛋白抑制复合物2)功能的细胞增生性病症例如癌症的治疗形式、方法、策略、组合物、组合和剂型。

[0006] 本披露的一些方面提供了用于治疗细胞增生性病症的治疗形式,这些细胞增生性病症的特征在于存在源自干细胞、干细胞样细胞、祖细胞或未成熟细胞的过度增生性细胞或细胞群,例如癌细胞或癌细胞群,其中该过度增生性细胞或细胞群包含赋予该癌细胞对EZH2功能的依赖性的遗传和/或表观遗传损伤。在一些实施例中,该细胞增生性病症例如癌症的特征在于干细胞、干细胞样细胞或祖细胞起源,和在至少一种调控多梳蛋白信号传导的基因中的一种或多种遗传和/或表观遗传损伤的组合。在一些实施例中,该细胞增生性病症例如癌症的特征在于导致一个或多个SWI/SNF复合物成员的功能丧失的一种或多种遗传和/或表观遗传损伤,该一个或多个SWI/SNF复合物成员例如INI-1(也称为SMARCB1,SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子,亚家族b,成员1)、SMARCA2(SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子,亚家族a,成员2)、和/或SMARCA4(SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子,亚家族a,成员4)。例如,在一些实施例中,该细胞增生性病症的特征在于导致SMARCA2和/或SMARCA4的功能丧失的一种或多种遗传和/或表观遗传损伤。在一些实施例中,该细胞增生性病症是肺的细胞增生性病症,例如肺癌。在某些实施例中,该EZH2抑制剂是他托司他(tazemetostat)。在一些实施例中,该细胞增生性病症是癌症。在一些实施例中,该细胞增生性病症的特征在于实体瘤。在一些实施例中,该细胞增生性病症是肺的细胞增生性病症,例如肺癌,例如像非小细胞肺癌、小细胞肺癌或间皮瘤。在某些实施例中,本文提供的治疗形式,例如某些策略、治疗方法和患者分层方法包括将该EZH2抑制剂与一种或多种另外的治疗剂的给予在时间上接近地给予至有需要的受试者,例如患有本文所述的细胞增生性病症的受试者。在一些实施例中,该一种或多种另外的治疗剂包括标准护理剂,例如临床上常用于该细胞增生性病症的一线、二线或三线治疗的药剂。在一些实施例中,该一种或多种另外的药剂包括免疫检查点抑制剂,例如PD-1或PDL-1抑制剂。

[0007] 本披露的一些方面提供了方法,这些方法包括将EZH2抑制剂给予至患有或被诊断为患有细胞增生性病症的受试者,该细胞增生性病症的特征在于展现出SMARCA2和/或SMARCA4的功能丧失的细胞或细胞群。本披露的一些方面提供了在有需要的受试者中治疗细胞增生性病症的方法,这些方法包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂,其中该细胞增生性病症的特征在于展现出SMARCA2和/或SMARCA4的功能丧失的细胞或细胞群。

[0008] 在一些实施例中,该细胞增生性病症是肺的细胞增生性病症。本披露的一些方面提供了在有需要的受试者中治疗肺的细胞增生性病症的方法,这些方法包括向该受试者给予治疗有效量的zeste增强子同源物2 (EZH2) 抑制剂。在一些实施例中,该细胞增生性病症包括或其特征在于展现出SMARCA2的功能丧失和/或SMARCA4的功能丧失的细胞或细胞群。在一些实施例中,该细胞增生性病症包括或其特征在于展现出SMARCA2和SMARCA4的功能丧失的细胞或细胞群。在一些实施例中,该细胞增生性病症的特征在于干细胞、干细胞样细胞或祖细胞起源。在一些实施例中,该肺的细胞增生性病症的特征在于肺中的恶性生长或病灶。在一些实施例中,该恶性生长或病灶是原发性病灶。在一些实施例中,该恶性生长或病灶是、或其特征在于继发性或转移性病灶。在一些实施例中,该肺癌是恶性肿瘤、癌或类癌肿瘤。在一些实施例中,该肺的细胞增生性病症是石棉诱导的增生、鳞状化生、和良性反应性间皮化生。在一些实施例中,该肺的细胞增生性病症是肺癌。在一些实施例中,该肺癌是小细胞肺癌。在一些实施例中,该肺癌是非小细胞肺癌。在一些实施例中,该肺癌是鳞状细胞癌。在一些实施例中,该肺癌是腺癌。在一些实施例中,该肺癌是小细胞癌。在一些实施例中,该肺癌是大细胞癌。在一些实施例中,该肺癌是腺鳞状细胞癌。在一些实施例中,该肺癌是间皮瘤。

[0009] 在一些实施例中,该细胞增生性疾病的特征在于原发性肿瘤,其中该原发性肿瘤(A)展现出SMARCA2/SMARCA4双重丧失;并且(B)是低分化的和/或展现出上皮间充质转化(EMT)特征。在一些实施例中,该原发性肿瘤展现出低的E-钙粘蛋白表达水平和高的波形蛋白表达水平。

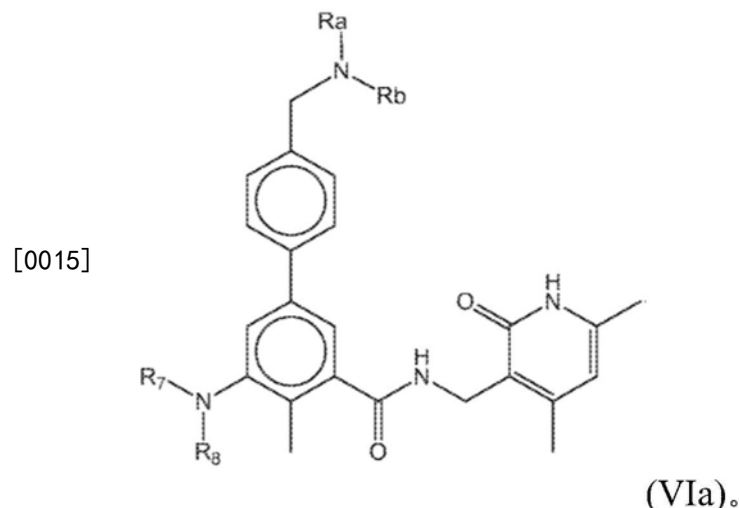
[0010] 在一些实施例中,该受试者已经或正在被给予另外的治疗剂,该治疗剂的给予与该EZH2抑制剂的给予同时或在时间上接近。在一些实施例中,该另外的治疗剂是标准护理剂。在一些实施例中,该另外的药剂是或包括在示意图1中列出的药剂,或者是或包括在示意图1中列出的两种或更多种药剂的组合。在一些实施例中,该另外的治疗剂是免疫检查点抑制剂。在一些实施例中,该免疫检查点抑制剂是CTLA4抑制剂、PD-1抑制剂和/或PD-L1抑制剂、LAG3抑制剂、B7-H3抑制剂或Tim3抑制剂。在一些实施例中,该免疫检查点抑制剂包括伊匹单抗 (Ipilimumab)、替西木单抗 (Ticilimumab)、AGEN-1884、尼沃鲁单抗 (Nivolumab)、派姆单抗 (Pembrolizumab)、阿特珠单抗 (Atezolizumab)、德瓦鲁单抗 (Durvalumab)、阿维单抗 (Avelumab)、BMS-936559、AMP-224、MEDI-0680、TSR-042、BGB-108、STI-1014、KY-1003、ALN-PDL、BGB-A317、KD-033、REGN-2810、PDR-001、SHR-1210、MGD-013、PF-06801591、CX-072、IMP-731、LAG-525、BMS-986016、GSK-2831781、依诺妥珠单抗 (Enoblituzumab)、1241-8H9、DS-5573、MBG-453、或其组合。在一些实施例中,该EZH2抑制剂和该另外的治疗剂依序给予至该受试者。在一些实施例中,该EZH2抑制剂和该另外的治疗剂经由不同的给予途径且以不同的间隔给予。在一些实施例中,该EZH2抑制剂每天两次口服给予。

[0011] 在一些实施例中,该方法进一步包括检测SMARCA2和/或SMARCA4蛋白表达和/或SMARCA2和/或SMARCA4蛋白的功能。在一些实施例中,该SMARCA2和/或该SMARCA4蛋白的表达和/或功能通过包括以下的方法评价:(a)从该受试者获得生物样品;(b)将该生物样品或其部分与特异性结合SMARCA2或SMARCA4的抗体接触;并且(c)检测与SMARCA2或SMARCA4结合的该抗体的量。

[0012] 在一些实施例中,该方法进一步包括检测获自该受试者的生物样品中编码该SMARCA2的基因和/或编码该SMARCA4蛋白的基因中的基因组突变。在一些实施例中,该基因组突变通过包括以下的方法检测:(a)从该受试者获得生物样品;(b)对该生物样品中的编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或编码SMARCA4蛋白或其部分的至少一种DNA序列进行测序;并且(c)确定该编码SMARCA2蛋白或其部分的至少一种DNA序列、和/或该编码SMARCA4蛋白或其部分的至少一种DNA序列是否包含影响该SMARCA2蛋白或该SMARCA4蛋白的表达和/或功能的突变。

[0013] 在一些实施例中,该EZH2抑制剂抑制组蛋白3的赖氨酸27 (H3K27) 的三甲基化。

[0014] 在一些实施例中,这些治疗形式,例如,治疗方法、组合物或组合包括或使用具有下式 (VIa) 的小分子EZH2抑制剂或其药学上可接受的盐或酯。



[0016] 在一些实施例中,这些具有式 (VIa) 的化合物可以包括以下特征中的一个或多个:

[0017] R_a 和 R_b 中的每一个独立地是H或任选地被一个或多个 $-Q_3-T_3$ 取代的 C_1-C_6 烷基。

[0018] R_a 和 R_b 与其所附接的N原子一起形成具有0或1个另外的杂原子的4至12元杂环烷基环,其中该4至12元杂环烷基环任选地被一个或多个 $-Q_3-T_3$ 取代。

[0019] R_a 和 R_b 与其所附接的N原子一起是具有0或1个另外的杂原子的4至7元杂环烷基环,其中该4至7元杂环烷基环任选地被一个或多个 $-Q_3-T_3$ 取代。

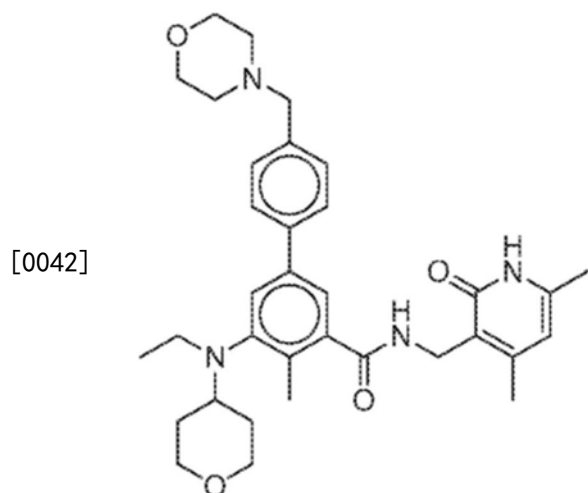
[0020] 每个 Q_3 独立地是键或未取代的或取代的 C_1-C_3 烷基接头。

[0021] 每个 T_3 独立地是H、卤代、 C_1-C_3 烷基、 OR_d 、 $COOR_d$ 、 $S(O)_2R_d$ 、 NR_dR_e 或4至7元杂环烷基,其中 R_d 和 R_e 中的每一个独立地是H或 C_1-C_6 烷基。

[0022] R_7 是 C_1-C_6 烷基、 C_3-C_8 环烷基或4至12元杂环烷基,这些基团各自任选地被一个或多个 $-Q_5-T_5$ 取代。

[0023] R_7 是 C_1-C_6 烷基、 C_3-C_8 环烷基或4至12元(例如,4至7元)杂环烷基,这些基团各自任选地被一个或多个 $-Q_5-T_5$ 取代。例如, R_7 不是H。

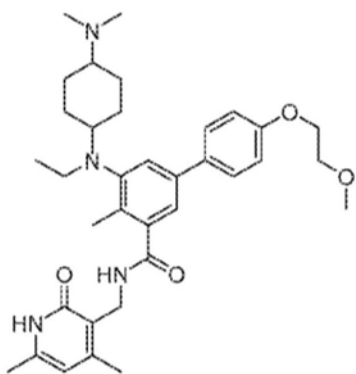
- [0024] R₇是任选地被一个或多个-Q₅-T₅取代的4至7元杂环烷基。
- [0025] R₇是哌啶基、四氢吡喃、环戊基、或环己基,这些基团各自任选地被一个Q₅-T₅取代。
- [0026] 每个Q₅独立地是键、CO、S(O)₂、NHC(O)或C₁-C₃烷基接头。
- [0027] 每个T₅独立地是H、卤代、S(O)_qR_q、C₁-C₆烷基、C₁-C₆烷氧基、C₃-C₈环烷基、4至12元杂环烷基、或C₆-C₁₀芳基,其中q是0、1或2并且R_q是C₁-C₆烷基、C₂-C₆烯基、C₂-C₆炔基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基、或5元或6元杂芳基。
- [0028] 每个T₅独立地是H、卤代、C₁-C₆烷基、C₁-C₆烷氧基、C₃-C₈环烷基、C₆-C₁₀芳基、或4至12元(例如,4至7元)杂环烷基。
- [0029] Q₅是键,并且T₅是C₁-C₆烷基、C₃-C₈环烷基、或4至12元(例如,4至7元)杂环烷基。
- [0030] Q₅是CO、S(O)₂、或NHC(O);并且T₅是C₁-C₆烷基、C₁-C₆烷氧基、C₃-C₈环烷基、或4至12元(例如,4至7元)杂环烷基。
- [0031] Q₅是C₁-C₃烷基接头并且T₅是H或C₆-C₁₀芳基。
- [0032] Q₅是C₁-C₃烷基接头并且T₅是C₃-C₈环烷基、4至7元杂环烷基、或S(O)_qR_q。
- [0033] R₇是环戊基或环己基,这些基团各自任选地被一个Q₅-T₅取代。
- [0034] Q₅是NHC(O)并且T₅是C₁-C₆烷基或C₁-C₆烷氧基。
- [0035] R₇是异丙基。
- [0036] R₈是H、C₁-C₆烷基、或4至7元杂环烷基,其中C₁-C₆烷基和杂环烷基各自任选地被选自下组的一个或多个取代基取代,该组由以下各项组成:卤代、羟基、COOH、C(O)O-C₁-C₆烷基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、和二-C₁-C₆烷基氨基。
- [0037] R₈是H、甲基、或乙基。
- [0038] R₈是甲基。
- [0039] R₈是乙基。
- [0040] R₈是4至7-杂环烷基,例如,四氢吡喃。
- [0041] 在一些实施例中,该EZH2抑制剂是



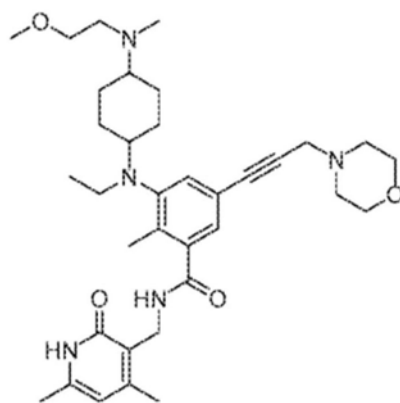
(他托司他)、

- [0043] 或其药学上可接受的盐。
- [0044] 在一些实施例中,该EZH2抑制剂是

[0045]

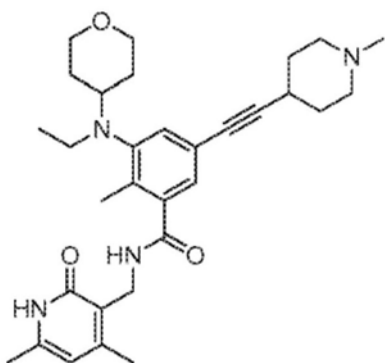


(A),

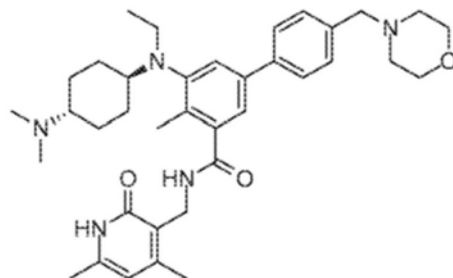


(B),

[0046]



(C),



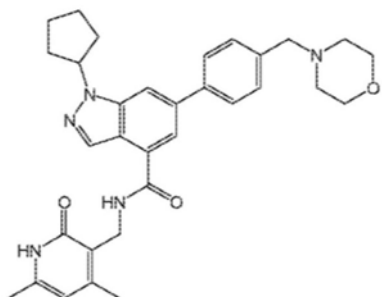
(D),

或其立

体异构体、药学上可接受的盐和/或溶剂化物。

[0047] 在一些实施例中,该EZH2抑制剂是

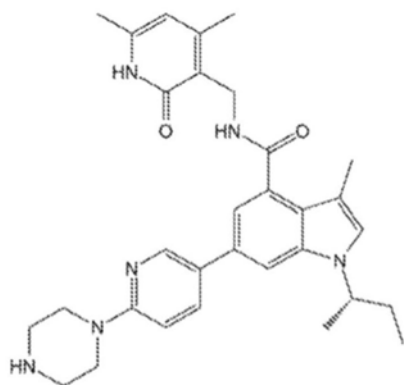
[0048]



或其药学上可接受的盐。

[0049] 在一些实施例中,该EZH2抑制剂是

[0050]

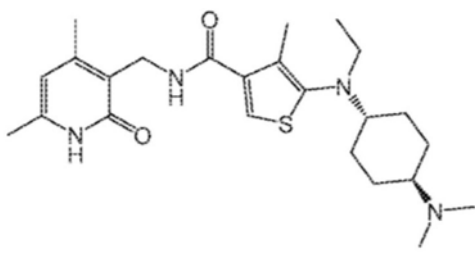


或其立体异构体、药学上可接受的盐和/或溶剂化

物。

[0051] 在一些实施例中,该EZH2抑制剂是

[0052]

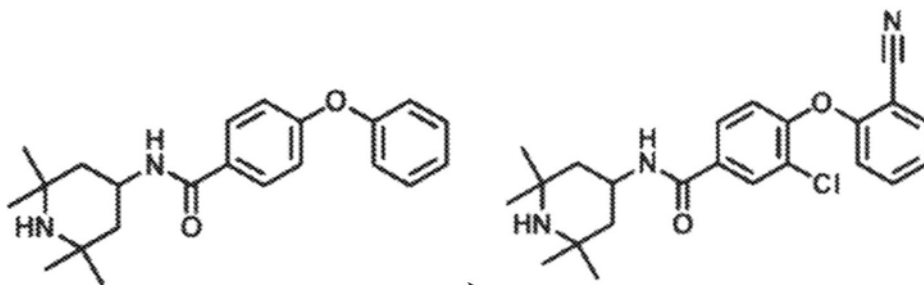


或其立体异构体、药学上可接受的盐

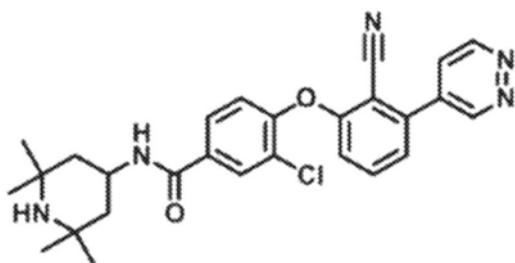
[0053] 和/或溶剂化物。

[0054] 在一些实施例中,该EZH2抑制剂是

[0055]



[0056]



或其立体异构体、药学上可接受的盐和/或

溶剂化物。

[0057] 在一些实施例中,该EZH2抑制剂可以包括以下、基本上由以下组成或由以下组成: CPI-1205或GSK343。

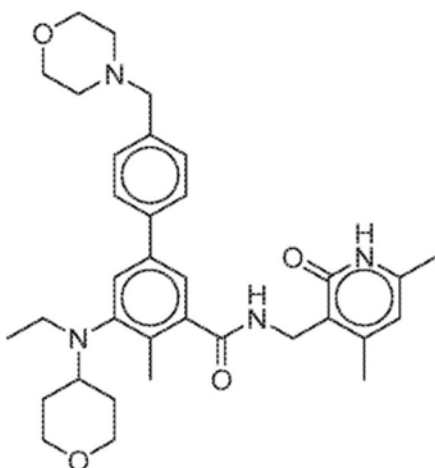
[0058] 在一些实施例中,该EZH2抑制剂是口服给予的。在一些实施例中,该EZH2抑制剂被配制为口服片剂。在一些实施例中,该EZH2抑制剂以10mg/kg/天与1600mg/kg/天之间的剂量给予。在一些实施例中,该EZH2抑制剂以约100、200、400、800、或1600mg的剂量给予。在一些实施例中,该EZH2抑制剂以约800mg的剂量给予。在一些实施例中,该EZH2抑制剂每天两次(BID)给予。

[0059] 本披露的一些方面提供了方法,这些方法包括检测获自受试者的样品中的SMARCA2和/或SMARCA4功能丧失。在一些实施例中,受试者患有癌症。在一些实施例中,如果在该受试者中检测到SMARCA2和/或SMARCA4功能丧失,则该方法进一步包括向该受试者给予EZH2抑制剂。在一些实施例中,该SMARCA2功能丧失与编码SMARCA2蛋白的基因中的基因组突变不相关,和/或其中该SMARCA4功能丧失与编码SMARCA4的基因中的基因组突变相关。在一些实施例中,其中该受试者患有NSCLC。

[0060] 在一些实施例中,本文提供的治疗形式包括或使用选自表1的化合物或其药学上可接受的盐或酯和一种或多种其他治疗剂。

[0061] 在一些实施例中,本文提供的治疗形式包括或使用下面提供的化合物:

[0062]



[0063] 或其药学上可接受的盐或酯和一种或多种其他治疗剂。

[0064] 以上概述以非限制性的方式意欲说明本文所披露的技术的一些实施例、优势、特征、和用途。本文所披露的技术的其他实施例、优势、特征、和用途将由于详细描述、附图、实例和权利要求书而是清楚的。

附图说明

[0065] 本专利或申请文件含有至少一幅彩色附图。在请求并支付必要的费用后，官方将会提供带有一幅或多幅彩色附图的本专利或专利申请公开物的副本。

[0066] 以上和另外特征将从结合附图时进行的以下详细描述而得以更清楚理解。

[0067] 图1. SWI/SNF复合物的亚基在各种适应症上进行突变。

[0068] 图2. SMARCA2/SMARCA4和SWI/SNF突变型肺癌细胞在体外对EZH2抑制的敏感性。

[0069] 图3. EZH2抑制对SMARCA4单一丧失NSCLC细胞系异种移植物的体内影响。

[0070] 图4. EZH2抑制对SMARCA2/SMARCA4双重丧失NSCLC细胞系异种移植物的体内影响。

具体实施方式

[0071] 本披露的一些方面提供了可用于用EZH2抑制剂治疗依赖于EZH2 (zeste增强子2多梳蛋白抑制复合物2) 功能的细胞增生性病症例如癌症的背景下的治疗形式，例如，方法、策略、组合物、组合和剂型。本披露的一些方面是基于以下认识：细胞增生性病症病状的亚型（例如，某些癌症的亚型）依赖于EZH2功能，并且因此可以有效地用EZH2抑制剂治疗。在一些实施例中，EZH2依赖性亚型的特征在于存在源自干细胞、干细胞样细胞、祖细胞或未成熟细胞的过度增生性细胞或细胞群，例如癌细胞或癌细胞群，其中至少一种过度增生性细胞或细胞群，例如至少一种癌细胞包含赋予癌细胞对EZH2功能的依赖性的遗传和/或表观遗传损伤。在一些实施例中，遗传或表观遗传损伤导致一个或多个SWI/SNF复合物成员的功能丧失，该一个或多个SWI/SNF复合物成员例如INI-1（也称为SMARCB1，SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，亚家族b，成员1）、SMARCA2（SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，亚家族a，成员2；有时也称为BRM、SNF2L2或SNF2LA）、和/或SMARCA4（SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，亚家族a，成员4；有时也称为brahma同源物BRG1、CSS4、MRD16、RTPS2、SNF2L4或SNF2LB）。例如，在一些实施例中，细胞增生性病症的特征在于导致SMARCA2和/或SMARCA4的功能丧失的遗传或表观遗传损

伤。

[0072] 本披露的一些方面是基于以下认识：展现出SMARCA2和/或SMARCA4功能丧失的某些细胞增生性病症，例如一些癌症依赖于EZH2功能并且因此对用EZH2抑制剂的治疗敏感。例如，本披露的一些方面提供了用于治疗实体瘤的治疗形式，例如，方法、策略、组合物、组合和剂型，这些实体瘤的特征在于干细胞、干细胞样细胞或祖细胞起源和SMARCA2或SMARCA4的功能丧失。

[0073] SWI/SNF复合物成员的基因组、mRNA和蛋白质序列，包括不与功能丧失或疾病或病症的状态相关的序列变体和同种型，是本领域技术人员已知的。在本文中，例如在下面的“示例性序列”部分中提供了SMARCA2和SMARCA4的示例性非限制性序列。其他合适的序列，例如，其他物种的序列以及功能序列变体是本领域技术人员已知的，并且本披露在此方面不受限制。

[0074] 本披露的一些方面是基于以下认识：在特征在于SMARCA4和SMARCA2的功能丧失的某些细胞增生性病症中，SMARCA4功能由于基因突变（典型地是SMARCA4基因的双等位基因突变）而丧失，而SMARCA2的功能丧失与基因突变不相关，但与表观遗传沉默相关。因此，本披露的一些方面提供了：在对用EZH2抑制剂治疗敏感的细胞增生性病症的一些实施例中，SMARCA2功能的丧失是例如通过SMARCA2调控序列的高甲基化而表观遗传下调或沉默SMARCA2基因表达的结果。本披露的一些方面提供了方法，这些方法包括再激活过度增生性细胞中，例如还展现出由基因突变介导的SMARCA4的功能丧失的恶性细胞中的表观遗传学抑制的SMARCA2表达，通过使这些细胞与EZH2抑制剂（例如他托司他）接触来进行。典型地，此类过度增生性细胞中的EZH2抑制和SMARCA2再激活导致对细胞存活和/或增生的抑制。在一些临床实施例中，用EZH2抑制剂治疗患有特征在于SMARCA2和SMARCA4的功能丧失的过度增生性疾病的患者导致抑制过度增生和/或消除过度增生性细胞。

[0075] 编码SWI/SNF复合物的成员的基因中的损伤先前在多种癌症中已报道。图1列出了其中报道了此类损伤的一些示例性恶性适应症。已经在各种细胞增生性疾病中观察到例如基于遗传损伤的SMARCA2和/或SMARCA4的丧失，这些细胞增生性疾病包括例如一些实体瘤适应症，例如像某些恶性横纹肌样瘤（例如，卵巢恶性横纹肌样瘤（MRT0）、卵巢高钙血症型小细胞癌（SCCOHT）；参见例如，2016年9月26日提交的PCT申请PCT/US 2016/053673，其全部内容通过援引并入本文），和某些肺癌亚型（例如，非小细胞肺癌、小细胞肺癌、腺肉瘤、鳞状细胞肉瘤）。特征在于SMARCA2和/或SMARCA4功能丧失的其他细胞增生性病症将是本领域技术人员已知的，或者将是技术人员仅仅通过常规实验基于本披露可确定的。本披露在此方面不受限制。

[0076] 下表1A提供了对NSCLC原发性肿瘤中的SMARCA2/SMARCA4丧失频率的汇总。

[0077] 表1A：

[0078]

参考文献	亚型	双重丧失 (IHC)	SMARCA4单一 丧失 (IHC)	SMARCA4 突变	SMARCA2 突变
Matsubara等人 2013	腺样	5/93 (5%)	11/93 (12%)	无数据	无数据
Resimann等人 2003	鳞状和腺样	6/60 (10%)	6/60 (10%)	无数据	无数据
Oike等人2013	鳞状和腺样	6/103 (6%)	16/103 (16%)	1/6	0/6
国家卫生研究院阿特 拉斯 (NIH Atlas) 数据库	鳞状和腺样		15-30%	0.56-3.31%	0%
Epizyme公司内 部数据*	NSCLC	6/226 (3%)	19/272 (7%)	无数据	无数据

[0079] 观察到的3%-10%的双重SMARCA2/SMARCA4丧失频率相当于仅在美国每年有7,000-23,000例的NSCLC。本披露的一些方面是基于以下令人惊讶的发现:SMARCA4和SMARCA2蛋白丧失在某些癌症中,例如NSCLC中比包含功能丧失突变的编码基因的频率显著更高。

[0080] 通过基因组序列分析不能检测到没有潜在基因组突变的蛋白质功能丧失。因此,用于基于DNA序列分析分类与SMARCA2和/或SMARCA4功能丧失相关的过度增生性疾病的常规方法易于产生假阴性结果,并且典型地低估双重SMARCA2/SMARCA4功能丧失的频率。本披露的一些方面提供了用于通过分析SMARCA2和/或SMARCA4的蛋白表达水平或蛋白功能来准确测定例如获自患有癌症的受试者的肿瘤活检中的过度增生性细胞或细胞群中的SMARCA2和SMARCA4状态的方法。例如,在一些实施例中,提供了患者分层方法,该患者分层方法包括检测获自患有癌症例如肺癌的受试者的生物样品中SMARCA2和/或SMARCA4蛋白的水平,并且将该水平与参考或对照水平,例如在健康的非恶性细胞中观察到或预期的水平进行比较。

[0081] 在一些实施例中,该方法包括通过基于免疫学的方法,例如通过免疫组织化学、蛋白质印迹、ELISA或其他合适的测定法检测获自受试者的样品中SMARCA2和/或SMARCA4蛋白的水平。在一些实施例中,该方法包括基于蛋白质动力学测定法,例如通过测定样品中SMARCA2和/或SMARCA4的酶活性的测定法检测SMARCA2和/或SMARCA4活性的水平。在一些实施例中,本文提供的方法可以以比常规的基于DNA测序的方法更高的准确性检测例如获自受试者的恶性细胞中展现出SMARCA2/SMARCA4双重丧失的过度增生性细胞或细胞群。

[0082] 在一些实施例中,该方法包括:如果SMARCA2和/或SMARCA4的蛋白水平相比于参考或对照水平降低,则将癌症,例如肺癌,诸如NSCLC分类为对用EZH2抑制剂的治疗敏感的。例如,在一些实例中,该方法包括:如果SMARCA2和/或SMARCA4蛋白的蛋白水平相比于参考或对照水平降低,则将癌症分类为对用EZH2抑制剂的治疗敏感的。在一些实施例中,如果癌症展现出双重SMARCA2/SMARCA4丧失,并且如果SMARCA2功能或SMARCA4功能、或二者丧失而在对应的编码基因中无功能丧失突变,则将该癌症分类为对用EZH2抑制剂的治疗敏感的。例如,在一些实施例中,该方法包括:将特征在于基于SMARCA4基因中的基因组突变的SMARCA4

功能丧失,和与SMARCA2基因中的基因组突变不相关的SMARCA2功能丧失的癌症分类为对用EZH2抑制剂的治疗敏感的。再例如,在一些实施例中,该方法包括:将特征在于基于SMARCA2基因中的基因组突变的SMARCA2功能丧失,和与SMARCA4基因中的基因组突变不相关的SMARCA4功能丧失的癌症分类为对用EZH2抑制剂的治疗敏感的。

[0083] 在一些实施例中,提供了一种方法,该方法包括将EZH2抑制剂(例如他托司他)给予至携带展现出SMARCA2/SMARCA4双重丧失的过度增生性细胞的受试者。在一些实施例中,受试者携带具有干细胞、干细胞样细胞或祖细胞起源、并且展现出SMARCA2/SMARCA4双重丧失的实体瘤,其中SMARCA2和/或SMARCA4的丧失与对应的编码基因中的功能丧失突变不相关。例如,在一些实施例中,该方法包括向患有癌症,例如肺癌,例如像NSCLC的受试者给予EZH2抑制剂,该癌症的特征在于基于SMARCA4基因中的基因组突变的SMARCA4功能丧失,和与SMARCA2基因中的基因组突变不相关的SMARCA2功能丧失。再例如,在一些实施例中,该方法包括向患有癌症的受试者给予EZH2抑制剂,该癌症的特征在于基于SMARCA2基因中的基因组突变的SMARCA2功能丧失,和与SMARCA4基因中的基因组突变不相关的SMARCA4功能丧失。

[0084] 本披露的一些方面提供了:EZH2抑制可以抑制或消除特征在于SMARCA2和/或SMARCA4的功能丧失,例如SMARCA2和SMARCA4的双重丧失的细胞的过度增生性状态,其中细胞中的至少一种功能丧失损伤是表观遗传损伤。受试者中细胞的过度增生性状态典型地与细胞增生性病症相关,例如与癌病状或癌前病状相关。可用本文提供的治疗形式治疗的细胞增生性病症包括所有形式的细胞增生性病症,例如癌症、癌前期或癌前病状、良性生长或病灶、恶性生长或病灶、和转移性病灶。在一些实施例中,细胞增生性病症的特征在于增生、化生或发育异常。在一些实施例中,细胞增生性疾病的特征在于原发性肿瘤。在一些实施例中,原发性肿瘤是实体瘤。在一些实施例中,原发性肿瘤是液体瘤。在一些实施例中,细胞增生性疾病的特征在于恶性生长或肿瘤。在一些实施例中,细胞增生性疾病的特征在于继发性或转移性肿瘤。

[0085] 本披露的一些方面提供了适于治疗肺的细胞增生性病症的治疗形式,该细胞增生性病症的特征在于SMARCA2和/或SMARCA4的功能丧失,例如SMARCA2和SMARCA4的双重功能丧失,其中细胞中的至少一种功能丧失损伤是表观遗传损伤。肺的细胞增生性病症是涉及肺细胞的细胞增生性病症。肺的细胞增生性病症可以包括影响肺细胞的所有形式的细胞增生性病症。肺的细胞增生性病症可以包括肺癌、肺的癌前期或癌前病状、肺的良性生长或病灶、和肺的恶性生长或病灶,以及身体除肺外的组织和器官中的转移性病灶。在一个方面,本披露的组合物可以用于治疗肺癌或肺的细胞增生性病症、或用于鉴定用于此类目的的合适候选物。肺癌可以包括所有形式的肺的癌症。肺癌可以包括恶性肺肿瘤、原位癌、典型类癌肿瘤、和非典型类癌肿瘤。肺癌可以包括小细胞肺癌(“SCLC”)、非小细胞肺癌(“NSCLC”)、鳞状细胞癌、腺癌、小细胞癌、大细胞癌、腺鳞状细胞癌、和间皮瘤。肺癌可以包括“瘢痕癌”、细支气管肺泡癌、巨细胞癌、梭形细胞癌、和大细胞神经内分泌癌。肺癌可以包括具有组织学和超微结构异质性的肺肿瘤(例如,混合细胞类型)。

[0086] 肺的细胞增生性病症可以包括影响肺细胞的所有形式的细胞增生性病症。肺的细胞增生性病症可以包括肺癌、肺的癌前病状。肺的细胞增生性病症可以包括肺的增生、化生、和发育异常。肺的细胞增生性病症可以包括石棉诱导的增生、鳞状化生、和良性反应性

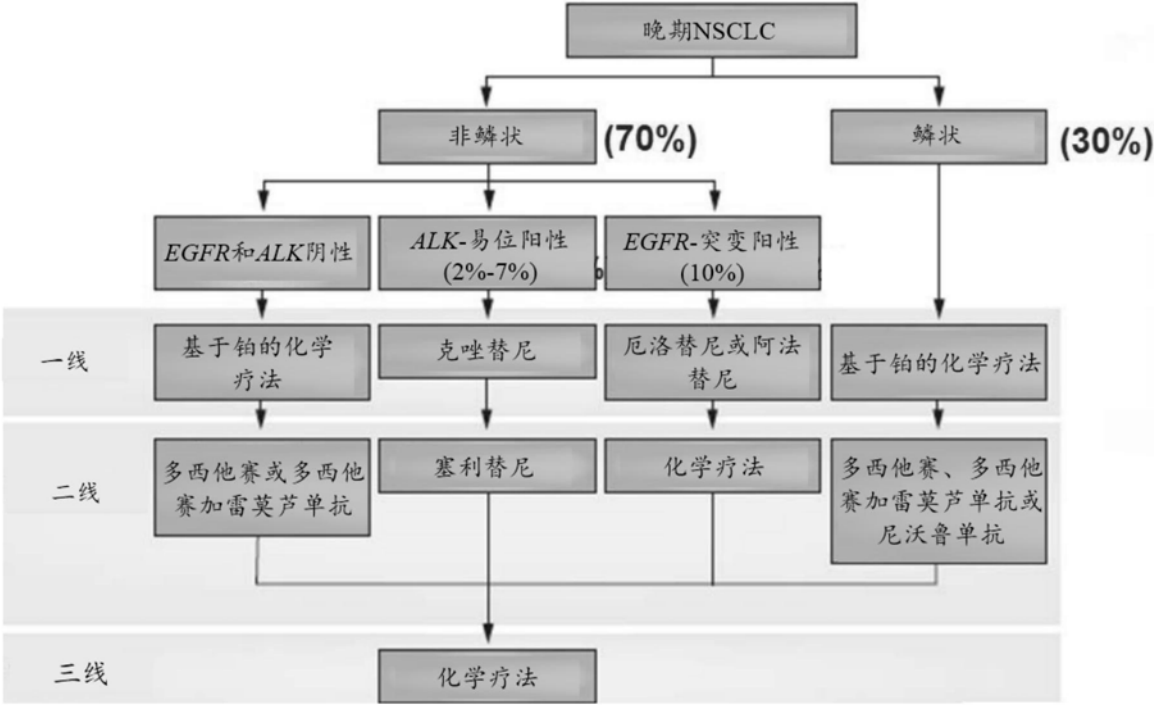
间皮化生。肺的细胞增生性病征可以包括柱状上皮被复层鳞状上皮置换和粘膜发育异常。暴露于吸入性有害环境因素(诸如香烟烟雾和石棉)的个体可能会增加发展肺的细胞增生性病征的风险。可使个体易于发展肺的细胞增生性病征的先前的肺部疾病可以包括慢性间质性肺病、坏死性肺病、硬皮病、类风湿病、结节病、间质性肺炎、肺结核、反复性肺炎、特发性肺纤维化、肉芽肿、石棉肺、纤维性肺泡炎、和霍奇金病。

[0087] 本披露的一些方面提供了适于治疗肺癌的治疗形式,该肺癌例如是特征在于SMARCA2和/或SMARCA4的功能丧失,例如SMARCA2和SMARCA4功能的双重丧失,其中细胞中的至少一种功能丧失损伤是表观遗传损伤的肺癌。肺癌是全球癌症相关死亡的最常见原因。仅在美国每年有约225,000个新例的肺癌诊断。约85%-90%的肺癌被表征为非小细胞肺癌(NSCLC),该非小细胞肺癌显示出各种各样的遗传驱动突变。肺癌的治疗已从化学疗法进化为靶向疗法。然而,对于新的治疗形式(例如方法、策略、组合物、组合和剂型)以及高效的患者分层仍存在大量未满足的临床需求。对于接受晚期化学疗法的患者尤其如此。最近开发的分子靶向疗法最适于治疗腺癌(例如,非鳞状癌),而有效的靶向治疗无法用于其他肺癌亚型。

[0088] 对患者分层和NSCLC临床管理的示例性范例的综述在Thomas等人Nature Reviews [自然综述]2016中进行了描述,其全部内容通过援引并入本文。下面的示意图1是改编自Thomas等人以概述一线、二线和三线治疗中的一些示例性治疗形式。应理解,在此包括下面的示意图以说明由临床医生使用的某些示例性治疗形式,它不限制本披露的范围,并且其他合适的患者分层和治疗形式是本领域技术人员已知的。

[0089] 示意图1:

[0090]



[0091] 虽然在常规且靶向的治疗形式的初始治疗方案中经常观察到良好的应答,但在大多数情况下最终发展了对此类治疗剂的抗性,并且对于那些发展抗性或难治性疾病的患者的治疗选择是有限的。正在开发针对某些肺癌适应症的新靶向治疗形式,例如免疫检查

点抑制剂,但仍然需要有效的治疗选择用于一线治疗和对标准护理治疗策略有抗性的肺癌的治疗。

[0092] 本披露的一些方面提供了适于治疗血液系统的细胞增生性病症的治疗形式,该细胞增生性病症的特征在于SMARCA2和/或SMARCA4的功能丧失,例如SMARCA2和SMARCA4的双重丧失,其中细胞中的至少一种功能丧失损伤是表观遗传损伤。血液系统的细胞增生性病症是涉及血液系统细胞的细胞增生性病症。适于本文提供的策略、治疗形式、方法、组合和组合物的血液系统的细胞增生性病症可以包括淋巴瘤、白血病、髓样肿瘤、肥大细胞肿瘤、脊髓发育不良、良性单克隆丙种球蛋白病、淋巴瘤样肉芽肿、淋巴瘤样丘疹病、真性红细胞增多症、慢性髓细胞白血病、特发性髓样化生、和特发性血小板增多症。血液系统的细胞增生性病症可以包括血液系统细胞的增生、发育异常、和化生。在一些实施例中,本文提供的策略、治疗形式、方法、组合和组合物用于治疗选自本披露的血液学癌症或本披露的血液学细胞增生性病症组成的组的癌症。本披露的血液学癌症可以包括多发性骨髓瘤、淋巴瘤(包括霍奇金淋巴瘤、非霍奇金淋巴瘤、儿童期淋巴瘤、和淋巴细胞和皮肤起源的淋巴瘤)、白血病(包括儿童期白血病、毛细胞白血病、急性淋巴细胞性白血病、急性髓细胞白血病、慢性淋巴细胞性白血病、慢性髓细胞白血病、慢性骨髓性白血病、和肥大细胞白血病)、髓样肿瘤以及肥大细胞肿瘤。

[0093] 本披露的一些方面提供了适于治疗癌症的治疗形式。在一些实施例中,癌症的特征在于SMARCA2和/或SMARCA4的功能丧失,例如SMARCA2和SMARCA4的双重丧失,其中细胞中的至少一种功能丧失损伤是表观遗传损伤。在一些实施例中,癌症的特征在于作为干细胞、干细胞样细胞或祖细胞的细胞起源。在一些实施例中,癌症是低分化癌症。在一些实施例中,癌症的特征在于实体瘤。在一些实施例中,癌症的特征在于继发性或转移性肿瘤。在一些实施例中,癌症是化学疗法抗性 or 难治性的。在一些实施例中,癌症是第一线、第二线和/或第三线治疗抗性 or 难治性的。在一些实施例中,癌症源自免疫细胞。在一些实施例中,癌症是呈淋巴瘤,例如B细胞淋巴瘤、非霍奇金淋巴瘤或弥漫性大B细胞淋巴瘤(DLBCL)的形式。在一些实施例中,癌症是肾上腺皮质癌、AIDS相关癌症、AIDS相关淋巴瘤、肛门癌、肛门直肠癌症、肛管癌症、阑尾癌、儿童小脑星形细胞瘤、儿童大脑星形细胞瘤、基底细胞癌、皮肤癌(非黑色素瘤)、胆道癌、肝外胆管癌、肝内胆管癌、膀胱癌(bladder cancer)、膀胱癌(urinary bladder cancer)、骨关节癌、骨肉瘤和恶性纤维组织细胞瘤、脑癌、脑肿瘤、脑干胶质瘤、小脑星形细胞瘤、脑星形细胞瘤/恶性胶质瘤、室管膜瘤、成神经管细胞瘤、幕上原始神经外胚层肿瘤、视觉通路和下丘脑胶质瘤、乳腺癌、支气管腺瘤/类癌、类癌肿瘤、胃肠道神经系统癌、神经系统淋巴瘤、中枢神经系统癌、中枢神经系统淋巴瘤、宫颈癌、儿童癌症、慢性淋巴细胞性白血病、慢性骨髓性白血病、慢性骨髓增生性病症、结肠癌、结肠直肠癌、皮肤T细胞淋巴瘤、淋巴肿瘤、蕈样真菌病、塞扎里(Seziary)综合征、子宫内膜癌、食管癌、颅外生殖细胞肿瘤、性腺外生殖细胞肿瘤、肝外胆管癌、眼癌、眼内黑色素瘤、视网膜母细胞瘤、胆囊癌、胃部(胃)癌、胃肠道类癌肿瘤、胃肠道间质瘤(GIST)、生殖细胞瘤、卵巢生殖细胞肿瘤、妊娠滋养细胞肿瘤胶质瘤、头颈癌、肝细胞(肝)癌、霍奇金淋巴瘤、下咽癌、眼内黑色素瘤、眼癌、胰岛细胞瘤(内分泌胰腺)、卡波西肉瘤、肾癌、肾脏癌、肾癌、喉癌、急性成淋巴细胞性白血病、急性髓样白血病、慢性淋巴细胞性白血病、慢性骨髓性白血病、毛细胞白血病、嘴唇和口腔癌症、肝癌、肺癌、非小细胞肺癌、小细胞肺癌、AIDS相关淋巴瘤、非霍

奇金淋巴瘤、原发性中枢神经系统淋巴瘤、瓦尔登斯特伦 (Waldenstroem) 巨球蛋白血症、成神经管细胞瘤、黑色素瘤、眼内(眼)黑色素瘤、默克(merkel)细胞癌、间皮瘤恶性、间皮瘤、转移性鳞状颈部癌、口腔癌、舌癌、多发性内分泌腺瘤综合征、蕈样真菌病、骨髓发育不良综合征、骨髓发育不良/骨髓增生性疾病、慢性骨髓性白血病、急性髓样白血病、多发性骨髓瘤、慢性骨髓增生性疾病、鼻咽癌、神经母细胞瘤、口腔癌(oral cancer)、口腔癌(oral cavity cancer)、口咽癌、卵巢癌、卵巢上皮癌、卵巢低度恶性潜能肿瘤、胰腺癌、胰岛细胞胰腺癌、鼻窦和鼻腔癌、甲状旁腺癌、阴茎癌、咽癌、嗜铬细胞瘤、松果体母细胞瘤和幕上原始神经外胚层肿瘤、垂体瘤、浆细胞肿瘤/多发性骨髓瘤、胸膜肺母细胞瘤、前列腺癌、直肠癌、肾盂和输尿管移行细胞癌、视网膜母细胞瘤、横纹肌肉瘤、唾液腺癌、尤文肉瘤家族肿瘤、卡波西肉瘤、软组织肉瘤、上皮样肉瘤、滑膜肉瘤、子宫癌、子宫肉瘤、皮肤癌(非黑色素瘤)、皮肤癌(黑色素瘤)、默克细胞皮肤癌、小肠癌、软组织肉瘤、鳞状细胞癌、胃(胃部)癌、幕上原始神经外胚层肿瘤、睾丸癌、咽喉癌、胸腺瘤、胸腺瘤和胸腺癌、甲状腺癌、肾盂和输尿管以及其他泌尿器官的移行细胞癌、妊娠滋养细胞肿瘤、尿道癌、子宫内膜子宫癌、子宫肉瘤、子宫体癌、阴道癌、外阴癌或威尔姆氏(Wilm's)肿瘤。

[0094] 在一些实施例中,可以用本披露的策略、治疗形式、方法、组合和组合物治疗的癌症包括实体瘤。在一些实施例中,可以用本披露的策略、治疗形式、方法、组合和组合物治疗的癌症包括或源自上皮起源的细胞。在一些实施例中,可以用本披露的策略、治疗形式、方法、组合和组合物治疗的癌症是原发性肿瘤。在一些实施例中,可以用本披露的策略、治疗形式、方法、组合和组合物治疗的癌症是继发性肿瘤。在一些实施例中,癌症是转移性的。

[0095] 本披露的一些方面提供了适于治疗癌症的治疗形式,该癌症根据美国癌症联合委员会(AJCC)TNM分类系统来分期,其中肿瘤(T)已被指定为TX、T1、T1mic、T1a、T1b、T1c、T2、T3、T4、T4a、T4b、T4c、或T4d期;并且其中局部淋巴结(N)已被指定为NX、N0、N1、N2、N2a、N2b、N3、N3a、N3b、或N3c期;并且其中远端转移(M)可以被指定为MX、M0、或M1期。在一些实施例中,适于用本文提供的形式治疗的癌症是根据美国癌症联合委员会(AJCC)分类而分期为I期、IIA期、IIB期、IIIA期、IIIB期、IIIC、或IV期的癌症。在一些实施例中,适于用本文提供的形式治疗的癌症可以根据AJCC分类而指定级别为GX级(例如,无法评估的级别)、级别1、级别2、级别3或级别4。在一些实施例中,待治疗的癌症根据AJCC病理分类(pN)分期为pNX、pN0、PN0(I-)、PN0(I+)、PN0(mol-)、PN0(mol+)、PN1、PN1(mi)、PN1a、PN1b、PN1c、pN2、pN2a、pN2b、pN3、pN3a、pN3b、或pN3c。

[0096] 本披露的一些方面提供了适于治疗癌症的治疗形式,该癌症包括已被确定为直径小于或等于2厘米的肿瘤。在一些实施例中,待治疗的癌症可以包括已被确定为直径是从约2厘米至约5厘米的肿瘤。在一些实施例中,待治疗的癌症可以包括已被确定为直径大于或等于约3厘米的肿瘤。在一些实施例中,待治疗的癌症可以包括已被确定为直径大于5厘米的肿瘤。在一些实施例中,待治疗的癌症可以通过微观外观分类为高度分化、中度分化、低分化、或未分化。在一些实施例中,待治疗的癌症可以关于有丝分裂计数(例如,细胞分裂量)或核多态性(例如,细胞变化)通过微观外观分类。在一些实施例中,待治疗的癌症可以通过微观外观被分类为与坏死区域(例如,死亡或退化细胞的区域)相关。在一些实施例中,待治疗的癌症可以被分类为具有异常核型、具有染色体异常数目、或具有外观异常的一个或多个染色体。在一些实施例中,待治疗的癌症可以被分类为非整倍体、三倍体、四倍体、或

具有改变的倍数性。在一些实施例中，待治疗的癌症可以被分类为具有染色体易位、或整个染色体的缺失或复制、或染色体的一部分的缺失、复制或扩增区域。

[0097] 在一些实施例中，待治疗的癌症可以通过DNA细胞术、流式细胞术或图像细胞术来评价。在一些实施例中，待治疗的癌症可以分型为具有10%、20%、30%、40%、50%、60%、70%、80%、或90%的在细胞分裂的合成阶段(如在细胞分裂的S期)的细胞。在一些实施例中，待治疗的癌症可以分型为具有低S期分数或高S期分数。

[0098] 在一些实施例中，本披露提供了可用于治疗癌症的治疗形式。治疗癌症可以导致肿瘤大小的减小。肿瘤大小的减小也可被称为“肿瘤消退”。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，肿瘤大小相对于其治疗前的大小减小5%或更多；更优选地，肿瘤大小减小10%或更多；更优选地，减小20%或更多；更优选地，减小30%或更多；更优选地，减小40%或更多；甚至更优选地，减小50%或更多；并且最优选地，减小多于75%或更多。肿瘤大小可以通过任何可重复测量方式来测量。肿瘤的大小可以测量为肿瘤的直径。

[0099] 治疗癌症可以导致肿瘤体积的减小。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，肿瘤体积相对于其治疗前的大小减小5%或更多；更优选地，肿瘤体积减小10%或更多；更优选地，减小20%或更多；更优选地，减小30%或更多；更优选地，减小40%或更多；甚至更优选地，减小50%或更多；并且最优选地，减小多于75%或更多。肿瘤体积可以通过任何可重复测量方式来测量。

[0100] 在一些实施例中，治疗癌症导致肿瘤数量的减少。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，肿瘤数量相对于治疗前的数量减少5%或更多；更优选地，肿瘤数量减少10%或更多；更优选地，减少20%或更多；更优选地，减少30%或更多；更优选地，减少40%或更多；甚至更优选地，减少50%或更多；并且最优选地，减少多于75%。肿瘤数量可以通过任何可重复测量方式来测量。肿瘤数量可以通过计数肉眼或在指定的放大率下可见的肿瘤来测量。优选地，指定的放大率是2x、3x、4x、5x、10x、或50x。

[0101] 在一些实施例中，治疗癌症可以导致远离原发性肿瘤位点的其他组织或器官中的转移性病灶数量的减少。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，转移性病灶数量相对于治疗前的数量减少5%或更多；更优选地，转移性病灶数量减少10%或更多；更优选地，减少20%或更多；更优选地，减少30%或更多；更优选地，减少40%或更多；甚至更优选地，减少50%或更多；并且最优选地，减少多于75%。转移性病灶数量可以通过任何可重复测量方式来测量。转移性病灶数量可以通过计数肉眼或在指定的放大率下可见的肿瘤来测量。优选地，指定的放大率是2x、3x、4x、5x、10x、或50x。

[0102] 在一些实施例中，相比于仅接受载体的群体，治疗癌症可以导致经治疗的受试者群体的平均存活时间延长。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，平均存活时间延长超过30天；更优选地，超过60天；更优选地，超过90天；并且最优选地，超过120天。群体的平均存活时间的延长可以通过任何可重复的方式来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗开始后计算群体的平均存活长度来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗第一轮完成后计算群体的平均存活长度来测量。

[0103] 在一些实施例中，相比于未经治疗的受试者群体，治疗癌症可以导致经治疗的受

试者群体的平均存活时间延长。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,平均存活时间延长超过30天;更优选地,超过60天;更优选地,超过90天;并且最优选地,超过120天。群体的平均存活时间的延长可以通过任何可重复的方式来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗开始后计算群体的平均存活长度来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗第一轮完成后计算群体的平均存活长度来测量。

[0104] 在一些实施例中,相比于接受单一疗法(使用的药物不是本披露的化合物或其药学上可接受的盐、溶剂化物、类似物或衍生物)的群体,治疗癌症可以导致经治疗的受试者群体的平均存活时间延长。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,平均存活时间延长超过30天;更优选地,超过60天;更优选地,超过90天;并且最优选地,超过120天。群体的平均存活时间的延长可以通过任何可重复的方式来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗开始后计算群体的平均存活长度来测量。群体的平均存活时间的延长可以例如通过使用活性化合物治疗第一轮完成后计算群体的平均存活长度来测量。

[0105] 在一些实施例中,相比于仅接受载体的群体,治疗癌症可以导致经治疗的受试者群体的死亡率降低。相比于未经治疗的群体,治疗癌症可以导致经治疗的受试者群体的死亡率降低。相比于接受单一疗法(使用的药物不是本披露的化合物或其药学上可接受的盐、溶剂化物、类似物或衍生物)的群体,治疗癌症可以导致经治疗的受试者群体的死亡率降低。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,死亡率降低超过2%;更优选地,超过5%;更优选地,超过10%;并且最优选地,超过25%。经治疗的受试者群体的死亡率的降低可以通过任何可重复的方式来测量。群体的死亡率的降低可以例如通过使用活性化合物治疗开始后计算每个单位时间内群体的疾病相关死亡的平均数量来测量。群体的死亡率的降低可以例如通过使用活性化合物治疗第一轮完成后计算每个单位时间内群体的疾病相关死亡的平均数量来测量。

[0106] 在一些实施例中,治疗癌症可以导致肿瘤生长速率的降低。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,在治疗之后,肿瘤生长速率相对于治疗前的数量降低至少5%;更优选地,肿瘤生长速率降低至少10%;更优选地,降低至少20%;更优选地,降低至少30%;更优选地,降低至少40%;更优选地,降低至少50%;甚至更优选地,降低至少50%;并且最优选地,降低至少75%。肿瘤生长速率可以通过任何可重复测量方式来测量。肿瘤生长速率可以根据每个单位时间内肿瘤直径的变化来测量。

[0107] 在一些实施例中,治疗癌症可以导致肿瘤再生长的减少。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,在治疗之后,肿瘤再生长少于5%;更优选地,肿瘤再生长少于10%;更优选地,少于20%;更优选地,少于30%;更优选地,少于40%;更优选地,少于50%;甚至更优选地,少于50%;并且最优选地,少于75%。肿瘤再生长可以通过任何可重复测量方式来测量。肿瘤再生长的测量例如是通过测量治疗后的居前肿瘤收缩后肿瘤直径的增加。治疗停止后肿瘤无法重新出现指示肿瘤再生长的减少。

[0108] 在一些实施例中,治疗细胞增生性病症可以导致细胞增生速率的降低。优选地,在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后,在治疗之后,细胞增生速率降低至少5%;更优选地,至少10%;更优选地,至少20%;更优选地,至少30%;更优选地,至少

40%；更优选地，至少50%；甚至更优选地，至少50%；并且最优选地，至少75%。细胞增生速率可以通过任何可重复测量方式来测量。细胞增生速率的测量例如是通过测量每个单位时间内组织样品中分裂细胞的数量。

[0109] 在一些实施例中，治疗细胞增生性病症可以导致增生细胞的比例的降低。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，在治疗之后，增生细胞的比例降低至少5%；更优选地，至少10%；更优选地，至少20%；更优选地，至少30%；更优选地，至少40%；更优选地，至少50%；甚至更优选地，至少50%；并且最优选地，至少75%。增生细胞的比例可以通过任何可重复测量方式来测量。优选地，增生细胞的比例例如是通过定量相对于非分裂细胞的数量而言组织样品中分裂细胞的数量来测量。增生细胞的比例可以是等同于细胞有丝分裂指数。

[0110] 在一些实施例中，治疗或预防细胞增生性病症可以导致细胞增生面积或区域大小的减小。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，在治疗之后，细胞增生面积或区域大小相对于其治疗前的大小减小至少5%；更优选地，减小至少10%；更优选地，减小至少20%；更优选地，减小至少30%；更优选地，减小至少40%；更优选地，减小至少50%；甚至更优选地，减小至少50%；并且最优选地，减小至少75%。细胞增生面积或区域大小可以通过任何可重复测量方式来测量。细胞增生面积或区域大小可以通过细胞增殖面积或区域的直径或宽度来测量。

[0111] 在一些实施例中，治疗或预防细胞增生性病症可以导致具有异常外观或形态的细胞的数量或比例的降低。优选地，在用本文提供的策略、治疗形式、方法、组合和组合物治疗之后，在治疗之后，具有异常形态的细胞的数量相对于其治疗前的大小减少至少5%；更优选地，减少至少10%；更优选地，减少至少20%；更优选地，减少至少30%；更优选地，减少至少40%；更优选地，减少至少50%；甚至更优选地，减少至少50%；并且最优选地，减少至少75%。异常细胞的外观或形态可以通过任何可重复测量方式来测量。异常细胞形态可以通过显微镜检查，例如使用倒置的组织培养显微镜来测量。异常细胞形态可以采取核多形性的形式。

[0112] 在一些实施例中，治疗细胞增生性病症可以导致过度增生性细胞死亡，并且优选地，细胞死亡导致过度增生性细胞群体中细胞数量减少至少10%。更优选地，细胞死亡意指减少至少20%；更优选地，减少至少30%；更优选地，减少至少40%；更优选地，减少至少50%；最优选地，减少至少75%。群体中细胞数量可以通过任何可重复的方式来测量。群体中细胞数量可以通过荧光激活细胞分选 (FACS)、免疫荧光显微镜和光学显微镜来测量。测量细胞死亡的方法示于Li等人, Proc Natl Acad Sci U S A. [美国国家科学院院刊] 100 (5): 2674-8, 2003中。在一些实施例中，细胞死亡通过细胞凋亡发生。

[0113] 在一些实施例中，通过向有需要的受试者给予EZH2抑制剂治疗细胞增生性病症，例如癌症导致以下项中的一个或多个：通过使细胞累积在细胞周期的一个或多个时期（例如G1、G1/S、G2/M）、或诱导细胞衰老、或促进肿瘤细胞分化来防止癌细胞增生；经由细胞毒性、坏死或细胞凋亡促进癌细胞的细胞死亡，但优选不引起大量正常细胞的细胞死亡。

[0114] 在本披露的方法的某些实施例中，将本文提供的治疗形式，例如治疗策略、治疗方法、分子测定法、组合物和组合应用至或给予至有需要的受试者，例如患有细胞增生性病症的受试者。在一些实施例中，受试者已被诊断为患有癌症。在一些实施例中，受试者是成人。

在一些实施例中,受试者是儿科受试者。在一些实施例中,受试者是人。

[0115] 在某些实施例中,受试者是成人,并且向受试者给予治疗有效量的EZH2抑制剂,例如他托司他,其中治疗有效量是约100mg至约1600mg。在某些实施例中,受试者是成人,并且EZH2抑制剂的治疗有效量是约100mg、200mg、400mg、800mg、或约1600mg。在某些实施例中,受试者是成人,并且EZH2抑制剂的治疗有效量是约800mg,例如以400mg的剂量每天口服两次来给予800mg/天。

[0116] 在某些实施例中,受试者是儿科的,并且EZH2抑制剂,例如他托司他可以以230mg/m²与600mg/m²之间、包括端值在内的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以230mg/m²与305mg/m²之间、包括端值在内的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以240mg/m²的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以300mg/m²的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂在向成人受试者每天两次给予1600mg后以稳态时曲线下面积(AUC)(AUCSS)的约60%的剂量给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂每天以约600mg/m²的剂量给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂每天以至少600mg/m²的剂量给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂在向成人受试者每天两次给予800mg后以稳态时曲线下面积(AUC)(AUCSS)的约80%的剂量给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以约390mg/m²的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以至少390mg/m²的剂量每天两次(BID)给予。在某些实施例中,受试者是儿科的,并且EZH2抑制剂以300mg/m²与600mg/m²之间的剂量每天两次(BID)给予。

[0117] 在一些实施例中,例如在受试者是儿科的一些实施例中,EZH2抑制剂被配制为口服混悬液。

[0118] 本披露的一些方面提供了适于通过向有需要的受试者给予治疗有效剂量的EZH2抑制剂来治疗细胞增生性病症,例如本文所述的癌症的组合治疗形式。在一些此类组合治疗实施例中,本文提供的治疗形式包括方法,这些方法包括向有需要的受试者,例如患有细胞增生性病症的受试者给予EZH2抑制剂,其中受试者已经或正在被给予另外的治疗剂,该治疗剂的给予与EZH2抑制剂的给予在时间上接近。在一些实施例中,提供了治疗形式,这些治疗形式包括向受试者给予EZH2抑制剂和另外的治疗剂。在一些实施例中,在时间上接近的给予是指EZH2抑制剂和另外的治疗剂以任意顺序在彼此的数小时或数天内的连续给予、或者是指EZH2抑制剂(例如,每日两次)和另外的治疗剂(例如,每周一次)的给予方案中的重叠,持续一定时间段(例如,至少一周、至少一个月、至少一轮治疗等)。

[0119] 在一些实施例中,本披露提供了可用于相比于单一治疗剂方式改善患有特征在于SMARCA2和/或SMARCA4的丧失的细胞增生性疾病例如癌症的受试者的临床结果和/或预后的组合疗法策略、治疗形式、方法、组合和组合物。在一些实施例中,本文提供的组合疗法方式导致相比于单一疗法需要更短的时间段来实现所期望的临床结果(例如,部分或完全疾病缓解、抑制肿瘤生长、疾病稳定)。在一些实施例中,本文提供的组合疗法方式导致相比于单一疗法更佳的临床效果(例如,完全缓解相比于部分缓解、稳定性疾病相比于渐进性疾病、再发风险低)。

[0120] 如本文所用,术语“组合治疗”,“组合疗法”和“共同疗法”可互换使用,并且通常是

指特征在于如本文提供的EZH2抑制剂和另外的治疗剂的治疗形式。典型地,组合治疗形式是旨在由治疗剂组合的并行作用提供有益效果的特定治疗方案的一部分。组合的有益效果可以包括但不限于由治疗剂的组合产生的药代动力学或药效动力学共同作用。组合给予这些治疗剂典型地在限定的时间段(通常数分钟、数小时、数天或数周,这取决于选择的组合)内进行。在一些实施例中,组合治疗包括以依序方式给予两种或更多种治疗剂,其中在不同时间给予每种治疗剂,以及以基本上同时的方式给予这些治疗剂或这些治疗剂中的至少两种。可以例如通过向受试者给予具有固定比率的每种治疗剂的单一剂型或以针对治疗剂的多个独立剂型来完成基本上同时的给予。依序或基本上同时给予每种治疗剂可以通过任何适当的途径来实现,这些途径包括但不限于口服途径、静脉内途径、肌内途径、和通过粘膜组织的直接吸收。可以通过相同途径或通过不同途径给予这些治疗剂。治疗剂可以根据相同或不同的给予间隔给予。例如,可以通过静脉内注射给予选择的组合的第一治疗剂,同时可以口服地给予组合的其他治疗剂。可替代地,例如,可以口服地给予所有治疗剂或可以通过静脉内注射给予所有治疗剂。

[0121] 在一些实施例中,组合疗法还包括进一步组合其他生物活性成分和非药物治疗法(例如,手术或放射治疗)来给予如上所述的治疗剂。当组合疗法进一步包括非药物治疗时,非药物治疗可在任何合适时间进行,只要实现治疗剂与非药物治疗的组合的共同作用的有益效果即可。例如,在适当的情况下,当非药物治疗暂时从治疗剂的给予中移除时(也许是数天甚至数个星期),仍然能实现有益的效果。

[0122] 在一些实施例中,另外的治疗剂是化学治疗剂(也称为抗肿瘤剂或抗增生剂),例如烷基化剂;抗生素;抗代谢药;解毒剂;干扰素;多克隆或单克隆抗体;EGFR抑制剂;HER2抑制剂;组蛋白脱乙酰酶抑制剂;激素;有丝分裂抑制剂;MTOR抑制剂;多激酶抑制剂;丝氨酸/苏氨酸激酶抑制剂;酪氨酸激酶抑制剂;VEGF/VEGFR抑制剂;紫杉烷或紫杉烷衍生物、芳香酶抑制剂、蒽环霉素、微管靶向药物、拓扑异构酶有毒药物、分子靶标或酶的抑制剂(例如,激酶或蛋白质甲基转移酶)、胞苷类似物药物或本领域技术人员已知的任何化学治疗剂、免疫检查点抑制剂、或任何抗肿瘤剂或抗增生剂。

[0123] 适于根据本文提供的组合治疗形式使用的示例性烷基化剂包括但不限于环磷酰胺(Cytosan;Neosar);苯丁酸氮芥(Leukeran);美法仑(Alkeran);卡莫司汀(BiCNU);白消安(Busulfex);洛莫司汀(CeeNU);达卡巴嗪(DTIC-Dome);奥沙利铂(Eloxatin);卡莫司汀(Gliadel);异环磷酰胺(Ifex);氮芥(Mustargen);白消安(Myleran);卡铂(Paraplatin);顺铂(CDDP;Platinol);替莫唑胺(Temodar);噻替哌(Thiopex);苯达莫司汀(Treanda);或链脲霉素(Zanosar)。

[0124] 示例性合适抗生素包括但不限于多柔比星(阿霉素);多柔比星脂质体(Doxil);米托蒽醌(Novantrone);博来霉素(Blenoxane);柔红霉素(Cerubidine);柔红霉素脂质体(DaunoXome);更生霉素(Cosmegen);表柔比星(Ellence);伊达比星(Idamycin);普卡霉素(Mithracin);丝裂霉素(Mutamycin);喷司他丁(Nipent);或戊柔比星(Valstar)。

[0125] 示例性抗代谢药包括但不限于氟尿嘧啶(Adrucil);卡培他滨(Xeloda);羟基脲(Hydrea);巯基嘌呤(Purinethol);培美曲塞(Alimta);氟达拉滨(Fludara);奈拉滨(Arranon);克拉屈滨(Cladribine Novaplus);氯法拉滨(Clolar);阿糖胞苷(Cytosar-U);地西他滨(Dacogen);阿糖胞苷脂质体(DepoCyt);羟基脲(Droxia);普拉曲沙(Folotyn);氟

尿昔 (FUDR) ; 吉西他滨 (Gemzar) ; 克拉屈滨 (Leustatin) ; 氟达拉滨 (Oforta) ; 甲氨蝶呤 (MTX; Rheumatrex) ; 甲氨蝶呤 (Trexall) ; 硫鸟嘌呤 (Tabloid) ; TS-1 或阿糖胞苷 (Tarabine PFS) 。

[0126] 示例性解毒剂包括但不限于氨磷汀 (Ethyol) 或美司钠 (Mesnex) 。

[0127] 示例性干扰素包括但不限于干扰素 α -2b (Intron A) 或干扰素 α -2a (Roferon-A) 。

[0128] 示例性多克隆或单克隆抗体包括但不限于曲妥珠单抗 (Herceptin) ; 奥法木单抗 (Arzerra) ; 贝伐单抗 (Avastin) ; 利妥昔单抗 (Rituxan) ; 西妥昔单抗 (Erbix) ; 帕尼单抗 (Vectibix) ; 托西莫单抗/碘-131托西莫单抗 (Bexxar) ; 阿仑单抗 (Campath) ; 替伊莫单抗 (Zevalin; In-111; Y-90Zevalin) ; 吉妥珠单抗 (Mylotarg) ; 依库珠单抗 (Soliris) 或狄诺塞麦。

[0129] 示例性EGFR抑制剂包括但不限于吉非替尼 (Iressa) ; 拉帕替尼 (Tykerb) ; 西妥昔单抗 (Erbix) ; 埃罗替尼 (Tarceva) ; 帕尼单抗 (Vectibix) ; PKI-166; 卡纳替尼 (CI-1033) ; 马妥珠单抗 (EMD-72000) 或EKB-569。

[0130] 示例性HER2抑制剂包括但不限于曲妥珠单抗 (Herceptin) ; 拉帕替尼 (Tykerb) 或 AC-480。

[0131] 组蛋白脱乙酰酶抑制剂包括但不限于伏立诺他 (Zolinza) 。

[0132] 示例性激素包括但不限于它莫西芬 (Soltamox; Nolvadex) ; 雷洛昔芬 (Evista) ; 甲地孕酮 (Megace) ; 亮丙瑞林 (Lupron; Lupron Depot; EligaRd; Viadur) ; 氟维司群 (Faslodex) ; 来曲唑 (Femara) ; 曲普瑞林 (Trelstar LA; Trelstar Depot) ; 依西美坦 (Aromasin) ; 戈舍瑞林 (Zoladex) ; 比卡鲁胺 (Casodex) ; 阿那曲唑 (Arimidex) ; 氟甲睾酮 (Androxy; Halotestin) ; 甲羟孕酮 (Provera; Depo-Provera) ; 雌氮芥 (Emcyt) ; 氟他胺 (Eulexin) ; 托瑞米芬 (Fareston) ; 地加瑞克 (Firmagon) ; 尼鲁米特 (Nilandron) ; 阿巴瑞克 (Plenaxis) ; 或睾内酯 (Teslac) 。

[0133] 示例性有丝分裂抑制剂包括但不限于紫杉醇 (Taxol; Onxol; Abraxane) ; 多西他赛 (Taxotere) ; 长春新碱 (Oncovin; Vincasar PFS) ; 长春碱 (Velban) ; 依托泊苷 (Toposar; Etopophos; VePesid) ; 替尼泊苷 (Vumon) ; 伊沙匹隆 (Ixempra) ; 诺考达唑; 埃博霉素; 长春瑞滨 (Navelbine) ; 喜树碱 (CPT) ; 伊立替康 (Camptosar) ; 托泊替康 (Hycamtin) ; 安吡啶或片螺素D (LAM-D) 。

[0134] 示例性MTOR抑制剂包括但不限于依维莫司 (Afinitor) 或替西罗莫司 (Torisel) ; 雷帕鸣、地磷莫司; 或AP23573。

[0135] 示例性多激酶抑制剂包括但不限于索拉非尼 (Nexavar) ; 舒尼替尼 (Sutent) ; BIBW 2992; E7080; ZD6474; PKC-412; 莫特沙尼 (motesanib) ; 或AP24534。

[0136] 示例性丝氨酸/苏氨酸激酶抑制剂包括但不限于鲁伯斯塔 (ruboxistaurin) ; eril/ 盐酸法舒地尔; 夫拉平度 (flavopiridol) ; 塞利西利 (seliciclib) (CYC202; Roscovitine) ; SNS-032 (BMS-387032) ; Pkc412; 苔藓抑素; KAI-9803; SF1126; VX-680; AZD1152;ARRY-142886 (AZD-6244) ; SCI0-469; GW681323; CC-401; CEP-1347或PD 332991。

[0137] 示例性酪氨酸激酶抑制剂包括但不限于埃罗替尼 (Tarceva) ; 吉非替尼 (Iressa) ; 伊马替尼 (Gleevec) ; 索拉非尼 (Nexavar) ; 舒尼替尼 (Sutent) ; 曲妥珠单抗 (Herceptin) ; 贝伐单抗 (Avastin) ; 利妥昔单抗 (Rituxan) ; 拉帕替尼 (Tykerb) ; 西妥昔单抗 (Erbix) ; 帕尼

单抗 (Vectibix); 依维莫司 (Afinitor); 阿仑单抗 (Campath); 吉妥珠单抗 (Mylotarg); 坦罗莫司 (Torisel); 帕唑帕尼 (Votrient); 达沙替尼 (Sprycel); 尼洛替尼 (Tasigna); 伐他拉尼 (Ptk787; ZK222584); CEP-701; SU5614; MLN518; XL999; VX-322; Azd0530; BMS-354825; SKI-606CP-690; AG-490; WHI-P154; WHI-P131; AC-220; 或 AMG888。

[0138] 示例性 VEGF/VEGFR 抑制剂包括但不限于贝伐单抗 (Avastin); 索拉非尼 (Nexavar); 舒尼替尼 (Sutent); 兰尼单抗; 哌加他尼; 或凡德他尼。

[0139] 示例性微管靶向药物包括但不限于紫杉醇、多西他赛、长春新碱、长春碱、诺考达唑、埃博霉素和诺维本。

[0140] 示例性拓扑异构酶有毒药物包括但不限于替尼泊苷、依托泊苷、阿霉素、喜树碱、柔红霉素、更生霉素、米托蒽醌、安吡啶、表柔比星和伊达比星。

[0141] 示例性紫杉烷或紫杉烷衍生物包括但不限于紫杉醇和多西他赛。

[0142] 示例性一般化学治疗剂、抗肿瘤剂、抗增生剂包括但不限于六甲蜜胺 (Hexalen); 异维甲酸 (Accutane; Amnesteem; Claravis; Sotret); 维甲酸 (Vesanoid); 阿扎胞苷 (Vidaza); 硼替佐米 (Velcade); 天冬酰胺酶 (Elspar); 左旋咪唑 (Ergamisol); 米托坦 (Lysodren); 丙卡巴肼 (Matulane); 培门冬酶 (Oncaspar); 地尼白介素-毒素连接物 (Ontak); 卟吩姆 (Photofrin); 阿地白介素 (Proleukin); 来那度胺 (Revlimid); 贝沙罗汀 (Targretin); 沙利度胺 (Thalomid); 坦罗莫司 (Torisel); 三氧化二砷 (Trisenox); 维替泊芬 (Visudyne); 含羞草氨酸 (Leucenol); (1M 替加氟-0.4M 5-氯-2,4-二羟基嘧啶-1M 氧嗪酸钾)、或洛伐他汀。

[0143] 在一些实施例中, 提供了组合治疗形式, 其中另外的治疗剂是细胞因子, 例如 G-CSF (粒细胞集落刺激因子)。在另一个方面, 本文提供的 EZH2 抑制剂可以与放射疗法组合给予。放射疗法也可以与本文提供的 EZH2 抑制剂和本文所述的作为多药剂疗法的一部分的其他化学治疗剂组合给予。在又一个方面, 本文提供的 EZH2 抑制剂可以与标准化学疗法组合组合给予, 这些标准化学疗法组合诸如但不限于 CMF (环磷酰胺、甲氨蝶呤和 5-氟尿嘧啶)、CAF (环磷酰胺、阿霉素和 5-氟尿嘧啶)、AC (阿霉素和环磷酰胺)、FEC (5-氟尿嘧啶、表柔比星和环磷酰胺)、ACT 或 ATC (阿霉素、环磷酰胺和紫杉醇)、利妥昔单抗、Xeloda (卡培他滨)、顺铂 (CDDP)、卡铂、TS-1 (摩尔比为 1:0.4:1 的替加氟、吉莫斯特和奥斯他特钾)、喜树碱-11 (CPT-11、伊立替康或 CamptosarTM)、CHOP (环磷酰胺、羟基柔红霉素、oncovin、和强的松或泼尼松龙)、R-CHOP (利妥昔单抗、环磷酰胺、羟基柔红霉素、oncovin、强的松或泼尼松龙)、或 CMFP (环磷酰胺、甲氨蝶呤、5-氟尿嘧啶和强的松)。

[0144] 在一些优选的实施例中, 本文提供的 EZH2 抑制剂可以与酶, 诸如受体或非受体激酶的抑制剂一起给予。受体或非受体激酶是例如, 酪氨酸激酶或丝氨酸/苏氨酸激酶。本文所述的激酶抑制剂是小分子、多核酸、多肽、或抗体。

[0145] 示例性激酶抑制剂包括但不限于贝伐单抗 (靶向 VEGF)、BIBW 2992 (靶向 EGFR 和 Erb2)、西妥昔单抗/Erbitux (靶向 Erb1)、伊马替尼/Gleevec (靶向 Bcr-Abl)、曲妥珠单抗 (靶向 Erb2)、吉非替尼/Iressa (靶向 EGFR)、兰尼单抗 (靶向 VEGF)、哌加他尼 (靶向 VEGF)、埃罗替尼/Tarceva (靶向 Erb1)、尼洛替尼 (靶向 Bcr-Abl)、拉帕替尼 (靶向 Erb1 和 Erb2/Her2)、GW-572016/二甲苯磺酸拉帕替尼 (靶向 HER2/Erb2)、帕尼单抗/Vectibix (靶向 EGFR)、凡德他尼 (靶向 RET/VEGFR)、E7080 (多个靶标, 包括 RET 和 VEGFR)、Herceptin (靶向 HER2/Erb2)、

PKI-166 (靶向EGFR)、卡纳替尼/CI-1033 (靶向EGFR)、舒尼替尼/SU-11464/Sutent (靶向EGFR和FLT3)、马妥珠单抗/Emd7200 (靶向EGFR)、EKB-569 (靶向EGFR)、Zd6474 (靶向EGFR和VEGFR)、PKC-412 (靶向VEGR和FLT3)、伐他拉尼/Ptk787/ZK222584 (靶向VEGR)、CEP-701 (靶向FLT3)、SU5614 (靶向FLT3)、MLN518 (靶向FLT3)、XL999 (靶向FLT3)、VX-322 (靶向FLT3)、Azd0530 (靶向SRC)、BMS-354825 (靶向SRC)、SKI-606 (靶向SRC)、CP-690 (靶向JAK)、AG-490 (靶向JAK)、WHI-P154 (靶向JAK)、WHI-P131 (靶向JAK)、索拉非尼/Nexavar (靶向RAF激酶、VEGFR-1、VEGFR-2、VEGFR-3、PDGFR- β 、KIT、FLT-3和RET)、达沙替尼/Sprycel (BCR/ABL和Src)、AC-220 (靶向F1t3)、AC-480 (靶向所有HER蛋白、“panHER”)、二磷酸莫特沙尼 (靶向VEGF1-3、PDGFR和c-kit)、狄诺塞麦 (靶向RANKL, 抑制SRC)、AMG888 (靶向HER3) 和AP24534 (多个靶标, 包括F1t3)。

[0146] 示例性丝氨酸/苏氨酸激酶抑制剂包括但不限于雷帕吗 (靶向mTOR/FRAP1)、地磷莫司(Deforolimus) (靶向mTOR)、Certican/依维莫司 (靶向mTOR/FRAP1)、AP23573 (靶向mTOR/FRAP1)、Eril/盐酸法舒地尔 (靶向RHO)、夫拉平度 (靶向CDK)、塞利西利/CYC202/Roscovitine (靶向CDK)、SNS-032/BMS-387032 (靶向CDK)、鲁伯斯塔 (靶向PKC)、Pkc412 (靶向PKC)、苔藓抑素 (靶向PKC)、KAI-9803 (靶向PKC)、SF1126 (靶向PI3K)、VX-680 (靶向极光激酶)、Azd1152 (靶向极光激酶)、Arry-142886/AZD-6244 (靶向MAP/MEK)、SCIO-469 (靶向MAP/MEK)、GW681323 (靶向MAP/MEK)、CC-401 (靶向JNK)、CEP-1347 (靶向JNK)、和PD 332991 (靶向CDK)。

[0147] 在一些实施例中, 提供了组合治疗形式, 这些组合治疗形式包括如本文提供的EZH2抑制剂和免疫检查点抑制剂。免疫检查点蛋白抑制免疫细胞 (例如, T细胞) 对某些细胞的作用。免疫检查点信号传导在平衡受试者的针对免疫系统所靶向的细胞 (例如, 感染或恶性细胞) 和免疫系统效应子未靶向的细胞 (例如, 健康细胞) 的免疫应答中起重要作用。不希望受任何特定理论的束缚, 据信一些癌细胞从免疫系统监视和破坏中逃避是由异常免疫检查点信号传导介导的, 其中癌细胞通过激活宿主的免疫细胞中的一个或多个免疫检查点信号传导通路来调节或消除宿主的免疫应答。已经鉴定了各种免疫检查点信号传导蛋白, 例如但不限于CTLA4、PD-1、PD-L1、LAG3、B7-H3和Tim3, 并且已开发了靶向此类免疫检查点蛋白的免疫检查点抑制剂。此类免疫检查点抑制剂降低或消除它们靶向的免疫检查点信号传导通路的活性, 并且因此可以增强受试者的例如针对原本逃避适当的免疫系统监视的病理细胞的免疫应答。例如, 已报道一些免疫检查点抑制剂有效地抑制阻止T细胞介导的对感染细胞或癌细胞的攻击的免疫检查点信号传导。因此, 本文所述的免疫检查点抑制剂能够或支持靶向恶性或感染性细胞的免疫系统监测和效应子功能 (例如, 呈T细胞攻击的形式)。本文提及的一些免疫检查点抑制剂包括特异性结合免疫细胞 (例如, T细胞) 上的一种或多种检查点蛋白并抑制该一种或多种检查点蛋白活性的单克隆抗体。本披露的免疫检查点抑制剂可以用于增强受试者的针对任何类型的癌细胞的免疫应答。

[0148] 虽然可以靶向任何检查点蛋白, 但是本披露的示例性免疫检查点抑制剂可以靶向、结合以下蛋白、并且/或者抑制以下蛋白活性: 该蛋白包括但不限于CTLA4、PD-1、PD-L1、LAG3、B7-H3、Tim3或其任何组合。靶向、结合CTLA4、并且/或者抑制CTLA4活性的免疫检查点抑制剂可以包括伊匹单抗、替西木单抗、AGEN-1884或其组合。靶向、结合PD-1和/或PD-L1、并且/或者抑制PD-1和/或PD-L1活性的免疫检查点抑制剂可以包括尼沃鲁单抗、派姆单抗、

阿特殊单抗、德瓦鲁单抗、阿维单抗、BMS-936559、AMP-224、MEDI-0680、TSR-042、BGB-108、STI-1014、KY-1003、ALN-PDL、BGB-A317、KD-033、REGN-2810、PDR-001、SHR-1210、MGD-013、PF-06801591、CX-072或其组合。靶向、结合LAG3、并且/或者抑制LAG3活性的免疫检查点抑制剂可以包括IMP-731、LAG-525、BMS-986016、GSK-2831781或其组合。靶向、结合B7-H3、并且/或者抑制B7-H3活性的免疫检查点抑制剂可以包括依诺妥珠单抗、1241-8H9、DS-5573或其组合。靶向、结合Tim3、并且/或者抑制Tim3活性的免疫检查点抑制剂可以包括MBG-453。

[0149] 适用于本文提供的组合治疗形式中的示例性免疫检查点抑制剂包括但不限于伊匹单抗、替西木单抗、AGEN-1884、尼沃鲁单抗、派姆单抗、阿特殊单抗、德瓦鲁单抗、阿维单抗、BMS-936559、AMP-224、MEDI-0680、TSR-042、BGB-108、STI-1014、KY-1003、ALN-PDL、BGB-A317、KD-033、REGN-2810、PDR-001、SHR-1210、MGD-013、PF-06801591、CX-072、IMP-731、LAG-525、BMS-986016、GSK-2831781、依诺妥珠单抗、1241-8H9、DS-5573、或其组合。

[0150] 例如，在一些实施例中，提供了用于治疗细胞增生性疾病，例如某些癌症的组合疗法策略、治疗形式和方法，其中EZH2抑制剂是他托司他、或其药学上可接受的盐，并且免疫检查点抑制剂是阿特殊单抗。例如，在一些实施例中，提供了一种方法，该方法包括向有需要的受试者，例如患有或被诊断为患有增生性疾病（例如，癌症）的受试者给予治疗有效量的他托司他或药学上其可接受的盐、和治疗有效量的阿特殊单抗。在一些实施例中，细胞增生性疾病是肺的细胞增生性疾病。在一些实施例中，肺的细胞增生性疾病是肺癌。在一些实施例中，肺癌是NSCLC。在一些实施例中，肺癌是SCLC。在一些实施例中，肺癌是转移性肺癌。在一些实施例中，肺癌是一线、二线、或三线肺癌治疗抗性或难治性的，例如，如本文所述的或本领域另外已知或使用的。在一些实施例中，肺癌的特征在于SMARCA2和/或SMARCA4功能丧失。在一些实施例中，肺癌的特征在于由表观遗传损伤介导的SMARCA2功能丧失。在一些实施例中，肺癌的特征在于由遗传损伤介导的SMARCA4功能丧失。在一些实施例中，肺癌的特征在于由表观遗传损伤介导的SMARCA2功能丧失和由遗传损伤介导的SMARCA4功能丧失。在一些实施例中，肺癌的特征在于低分化的肿瘤或病灶。在一些实施例中，肺癌的特征在于上皮间充质转化的特征。

[0151] 在某些实施例中，本披露提供了一种在有需要的受试者中治疗细胞增生性疾病，例如癌症的方法，该方法包括向受试者给予本文提供的EZH2抑制剂和免疫检查点抑制剂的组合。在一些实施例中，EZH2抑制剂是他托司他。在一些实施例中，EZH2抑制剂以800mg的口服剂量每天两次给予。在一些实施例中，免疫检查点抑制剂是阿特殊单抗（TECENTRIQ™）。在一些实施例中，免疫检查点抑制剂，例如阿特殊单抗作为静脉内输注以1200mg的剂量在约60分钟内给予，每3周一次（对于关于阿特殊单抗的另外信息，参见accessdata.fda.gov/drugsatfda_docs/label/2016/761034s0001b1.pdf，将其内容并入本文）。

[0152] 在某些实施例中，本披露提供了在有需要的受试者中治疗肺癌，例如NSCLC、SCLC、间皮瘤或任何其他形式的肺癌的方法，该方法包括向受试者给予他托司他和阿特殊单抗（TECENTRIQ™）的组合，该他托司他以800mg的口服剂量每天两次给予，该阿特殊单抗作为静脉内输注以1200mg的剂量在约60分钟内给予，每3周一次。

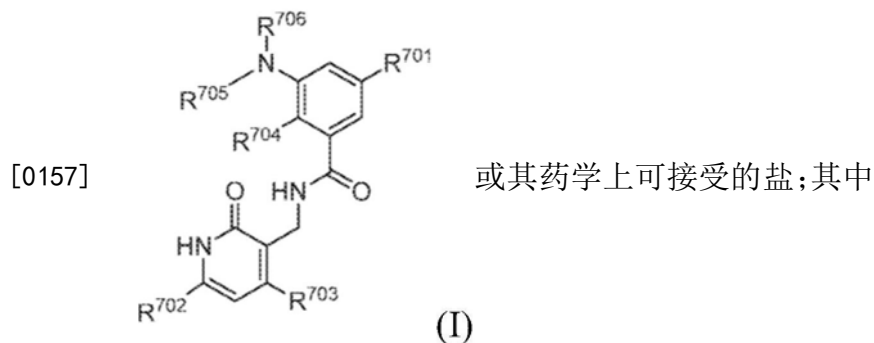
[0153] 在某些实施例中，本披露提供了在有需要的受试者中治疗非霍奇金淋巴瘤（或任何其他形式的血基质癌症）的方法，该方法包括向受试者给予他托司他和阿特殊单抗（TECENTRIQ™）的组合，该他托司他以800mg的口服剂量每天两次给予，该阿特殊单抗作为静

脉内输注以1200mg的剂量在60分钟内给予,每3周一次。

[0154] 在一些实施例中,本文提供的治疗形式包括监测受试者中靶细胞或组织的甲基化状态,例如通过本文所述的或本领域技术人员另外已知的方法进行,例如通过本文所述的或本领域技术人员另外已知的方法进行。在一些实施例中,本文提供的治疗形式包括监测受试者中靶细胞或组织中的SMARCA2和/或SMARCA4蛋白表达或蛋白功能的状态,例如通过本文所述的或本领域技术人员另外已知的方法进行。在一些实施例中,本文提供的治疗形式包括监测受试者中的免疫应答状态,例如通过本文所述的或本领域技术人员另外已知的方法进行。

[0155] 适于与本文提供的治疗形式一起使用的各种小分子EZH2抑制剂先前已描述。适用于本文提供的治疗形式中的EZH2抑制剂的一些非限制性实例是以下专利中描述的那些:US 8,410,088、US 8,765,732、US 9,090,562、US 8,598,167、US 8,962,620、US-2015/0065483、US 9,206,157、US 9,006,242、US 9,089,575、US 2015-0352119、WO 2014/062733、US-2015/0065503、WO 2015/057859、US 8,536,179、WO 2011/140324、公布为WO/2014/124418的PCT/US 2014/015706、公布为WO/2013/120104的PCT/US2013/025639、以及公布为US 2015/0368229的US 14/839,273,每篇专利的全部内容通过援引并入本文。

[0156] 在一些实施例中,适用于本文所述的策略、治疗形式、方法、组合和组合物中的EZH2抑制剂具有下式(I):



[0158] R^{701} 是H、F、 OR^{707} 、 NHR^{707} 、 $-(C=C)-(CH_2)_{n7}-R^{708}$ 、苯基、5元或6元杂芳基、 C_{3-8} 环烷基、含有1-3个杂原子的4-7元杂环烷基,其中苯基、5元或6元杂芳基、 C_{3-8} 环烷基或4-7元杂环烷基各自独立地任选地被一个或多个选自以下的基团取代:卤代、 C_{1-3} 烷基、OH、 $O-C_{1-6}$ 烷基、 $NH-C_{1-6}$ 烷基、和被 C_{3-8} 环烷基或含有1-3个杂原子的4-7元杂环烷基取代的 C_{1-3} 烷基,其中 $O-C_{1-6}$ 烷基和 $NH-C_{1-6}$ 烷基中的每一个任选地被羟基、 $O-C_{1-3}$ 烷基或 $NH-C_{1-3}$ 烷基取代,并且 $O-C_{1-3}$ 烷基和 $NH-C_{1-3}$ 烷基中的每一个任选地进一步被 $O-C_{1-3}$ 烷基或 $NH-C_{1-3}$ 烷基取代;

[0159] R^{702} 和 R^{703} 中的每一个独立地是H、卤代、 C_{1-4} 烷基、 C_{1-6} 烷氧基或 C_6-C_{10} 芳氧基,这些基团各自任选地被一个或多个卤代取代;

[0160] R^{704} 和 R^{705} 中的每一个独立地是 C_{1-4} 烷基;

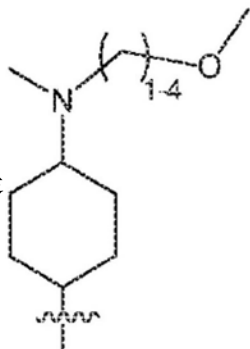
[0161] R^{706} 是被N(C_{1-4} 烷基)₂取代的环己基,其中一个或两个 C_{1-4} 烷基任选地被 C_{1-6} 烷氧基取代;或者 R^{706} 是四氢吡喃基;

[0162] R^{707} 是任选地被一个或多个选自以下的基团取代的 C_{1-4} 烷基:羟基、 C_{1-4} 烷氧基、氨基、单-或二- C_{1-4} 烷基氨基、 C_{3-8} 环烷基和含有1-3个杂原子的4-7元杂环烷基,其中 C_{3-8} 环烷基或4-7元杂环烷基各自独立地进一步任选地被 C_{1-3} 烷基取代;

[0163] R^{708} 是任选地被一个或多个选自以下的基团取代的 C_{1-4} 烷基:OH、卤代和 C_{1-4} 烷氧基、含有1-3个杂原子的4-7元杂环烷基、或 $O-C_{1-6}$ 烷基,其中4-7元杂环烷基可以任选地进一步被OH或 C_{1-6} 烷基取代;并且

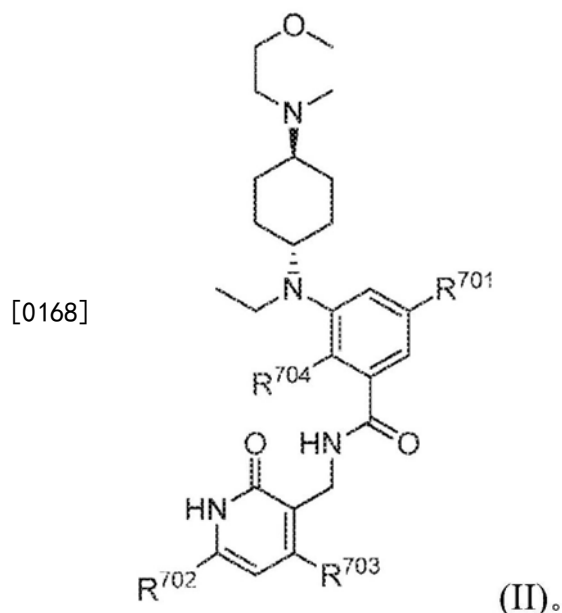
[0164] n_7 是0、1或2。

[0165] 在一些实施例中, R^{706} 是被 $N(C_{1-4}\text{烷基})_2$ 取代的环己基,其中一个 C_{1-4} 烷基是未取代的并且另一个被甲氧基取代。



[0166] 在一些实施例中, R^{706} 是

[0167] 在一些实施例中,化合物具有式II:



[0169] 在一些实施例中, R^{702} 是甲基或异丙基并且 R^{703} 是甲基或甲氧基。

[0170] 在一些实施例中, R^{704} 是甲基。

[0171] 在一些实施例中, R^{701} 是 OR^{707} 并且 R^{707} 是任选地被 OCH_3 或吗啉取代的 C_{1-3} 烷基。

[0172] 在一些实施例中, R^{701} 是H或F。

[0173] 在一些实施例中, R^{701} 是四氢吡喃基、苯基、吡啶基、嘧啶基、吡嗪基、咪唑基或吡唑基,这些基团中的每一个任选地被以下基团取代:甲基、甲氧基、被吗啉取代的乙基或 $-OCH_2CH_2OCH_3$ 。

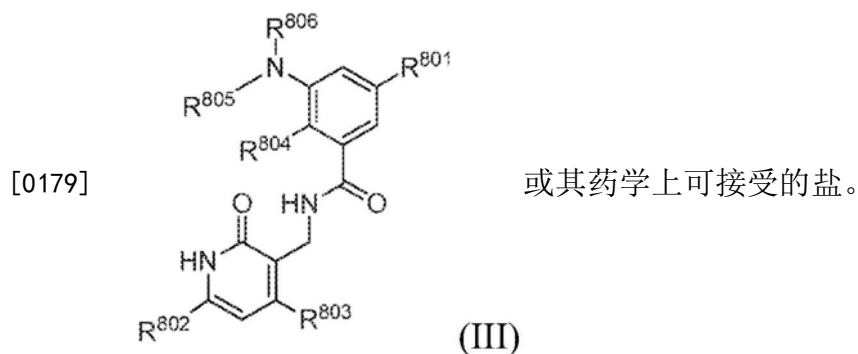
[0174] 在一些实施例中, R^{708} 是吗啉、哌啶、哌嗪、吡咯烷、二氮杂环庚烷或氮杂环丁烷,这些基团中的每一个任选地被以下基团取代:OH或 C_{1-6} 烷基。

[0175] 在一些实施例中, R^{708} 是吗啉。

[0176] 在一些实施例中, R^{708} 是被 C_{1-6} 烷基取代的哌嗪。

[0177] 在一些实施例中, R^{708} 是甲基、叔丁基或 $C(CH_3)_2OH$ 。

[0178] 在一些实施例中, 可以用于本文所述的策略、治疗形式、方法、组合和组合物中的 EZH2 抑制剂可以具有下式 III:



[0180] 在该式中:

[0181] R^{801} 是 C_{1-6} 烷基、 C_{2-6} 烯基、 C_{2-6} 炔基、 C_{3-8} 环烷基、含有 1-3 个杂原子的 4-7 元杂环烷基、苯基或 5 元或 6 元杂芳基, 这些基团中的每一个被 $O-C_{1-6}$ 烷基- R_x 或 $NH-C_{1-6}$ 烷基- R_x 取代, 其中 R_x 是羟基、 $O-C_{1-3}$ 烷基或 $NH-C_{1-3}$ 烷基, 并且 R_x 任选地进一步被 $O-C_{1-3}$ 烷基或 $NH-C_{1-3}$ 烷基取代, 除了当 R_x 是羟基时; 或者 R^{801} 是被 $-Q_2-T_2$ 取代的苯基, 其中 Q_2 是键或任选地被卤代、氰基、羟基或 C_{1-6} 烷氧基取代的 C_{1-3} 烷基接头, 并且 T_2 是任选取代的 4 至 12 元杂环烷基; 并且 R^{801} 任选地进一步被取代;

[0182] R^{802} 和 R^{803} 中的每一个独立地是 H、卤代、 C_{1-4} 烷基、 C_{1-6} 烷氧基或 C_6-C_{10} 芳氧基, 这些基团各自任选地被一个或多个卤代取代;

[0183] R^{804} 和 R^{805} 中的每一个独立地是 C_{1-4} 烷基; 并且

[0184] R^{806} 是 $-Q_x-T_x$, 其中 Q_x 是键或 C_{1-4} 烷基接头, T_x 是 H、任选取代的 C_{1-4} 烷基、任选取代的 C_3-C_8 环烷基或任选取代的 4 至 14 元杂环烷基。

[0185] 在一些实施例中, Q_x 和 Q_2 中的每一个独立地是键或甲基接头, 并且 T_x 和 T_2 中的每一个独立地是四氢吡喃基、被 1、2 或 3 个 C_{1-4} 烷基基团取代的哌啶基或被 $N(C_{1-4} \text{ 烷基})_2$ 取代的环己基, 其中一个或两个 C_{1-4} 烷基任选地被 C_{1-6} 烷氧基取代。

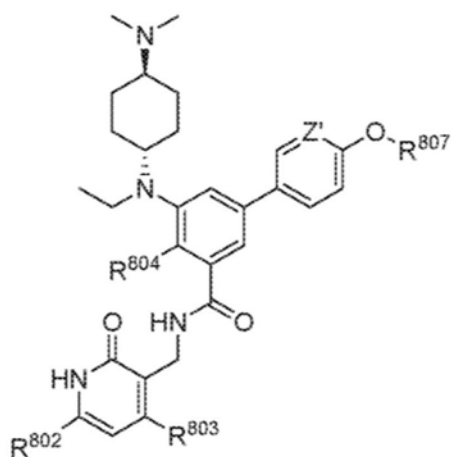
[0186] 在一些实施例中, R^{806} 是被 $N(C_{1-4} \text{ 烷基})_2$ 取代的环己基或者 R^{806} 是四氢吡喃基。



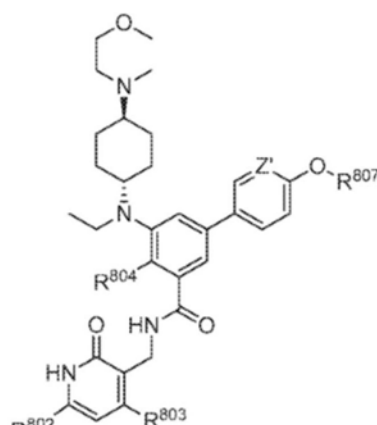
[0188] 在一些实施例中, R^{801} 是被 $O-C_{1-6}$ 烷基- R_x 取代的苯基或 5 元或 6 元杂芳基, 或者 R^{801} 是被 CH_2 -四氢吡喃基取代的苯基。

[0189] 在一些实施例中, 在一些实施例中, 根据本披露的一些方面的化合物具有式 IVa 或 IVb:

[0190]



(IVa) 或



(IVb),

其

中Z' 是CH或N,并且R⁸⁰⁷是C₂₋₃烷基-R_x。

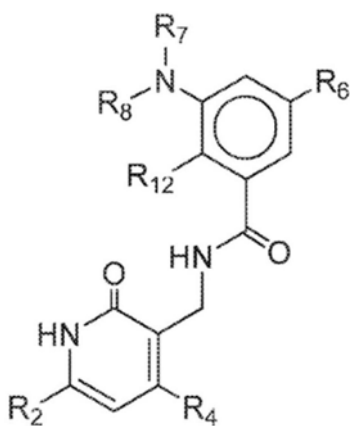
[0191] 在一些实施例中,R⁸⁰⁷是-CH₂CH₂OH、-CH₂CH₂OCH₃或-CH₂CH₂OCH₂CH₂OCH₃。

[0192] 在一些实施例中,R⁸⁰²是甲基或异丙基并且R⁸⁰³是甲基或甲氧基。

[0193] 在一些实施例中,R⁸⁰⁴是甲基。

[0194] 在一些实施例中,本披露的化合物可以具有下式(V):

[0195]



或其药学上可接受的盐或酯。

(V),

[0196] 在该式中:

[0197] R₂、R₄和R₁₂各自独立地是C₁₋₆烷基;

[0198] R₆是C₆-C₁₀芳基或5元或6元杂芳基,这些基团中的每一个任选地被一个或多个-Q₂-T₂取代,其中Q₂是键或任选地被卤代、氰基、羟基或C₁-C₆烷氧基取代的C₁-C₃烷基接头,并且T₂是H、卤代、氰基、-OR_a、-NR_aR_b、-(NR_aR_bR_c)⁺A⁻、-C(O)R_a、-C(O)OR_a、-C(O)NR_aR_b、-NR_bC(O)R_a、-NR_bC(O)OR_a、-S(O)₂R_a、-S(O)₂NR_aR_b或R_{S2},其中R_a、R_b和R_c中的每一个独立地是H或R_{S3},A⁻是药学上可接受的阴离子,R_{S2}和R_{S3}中的每一个独立地是C₁-C₆烷基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基或5元或6元杂芳基,或者R_a和R_b与其所附接的N原子一起形成具有0或1个另外的杂原子的4至12元杂环烷基环,并且R_{S2}、R_{S3}、和由R_a和R_b形成的4至12元杂环烷基环任选地被一个或多个-Q₃-T₃取代,其中Q₃是键或C₁-C₃烷基接头,各自任选地被卤代、氰基、羟基或C₁-C₆烷氧基取代,并且T₃选自下组,该组由以下各项组成:卤代、氰基、C₁-C₆烷基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基、5元或6元杂芳基、OR_d、COOR_d、-S(O)₂R_d、-NR_dR_e和-C(O)NR_dR_e,R_d和R_e中的每一个独立地是H或C₁-C₆烷基,或者-Q₃-T₃是氧代;或者任何两个相邻-Q₂-T₂与其所附接的原子一起形成任选地含有1-4个选自N、O和S的杂原子并且任选地被一个或

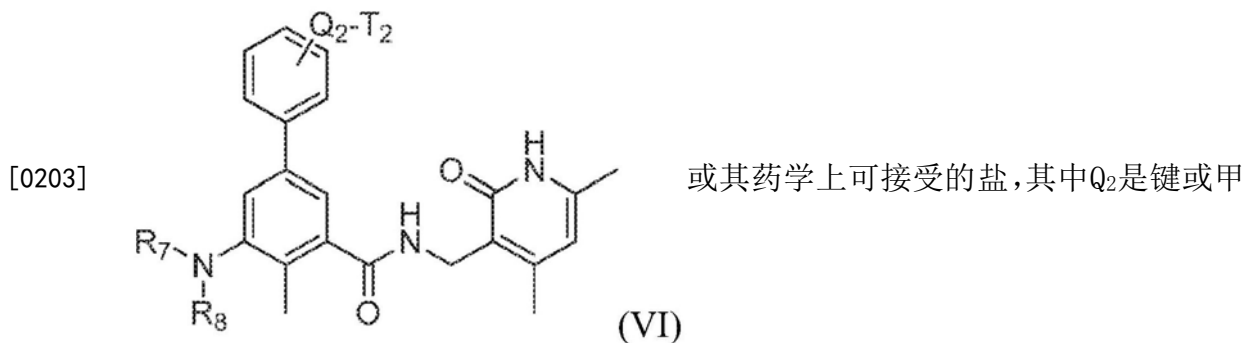
多个选自下组的取代基取代的5元或6元环,该组由以下各项组成:卤代、羟基、COOH、C(O)O-C₁-C₆烷基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、二-C₁-C₆烷基氨基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基和5元或6元杂芳基;

[0199] R₇是-Q₄-T₄,其中Q₄是键、C₁-C₄烷基接头、或C₂-C₄烯基接头,每个接头任选地被卤代、氰基、羟基或C₁-C₆烷氧基取代,并且T₄是H、卤代、氰基、NR_fR_g、-OR_f、-C(O)R_f、-C(O)OR_f、-C(O)NR_fR_g、-C(O)NR_fOR_g、-NR_fC(O)R_g、-S(O)₂R_f或RS₄,其中R_f和R_g中的每一个独立地是H或RS₅,RS₄和RS₅中的每一个独立地是C₁-C₆烷基、C₂-C₆烯基、C₂-C₆炔基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基或5元或6元杂芳基,并且RS₄和RS₅中的每一个任选地被一个或多个-Q₅-T₅取代,其中Q₅是键、C(O)、C(O)NR_k、NR_kC(O)、S(O)₂或C₁-C₃烷基接头,R_k是H或C₁-C₆烷基,并且T₅是H、卤代、C₁-C₆烷基、羟基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、二-C₁-C₆烷基氨基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基、5元或6元杂芳基或S(O)_qR_q,其中q是0、1、或2并且R_q是C₁-C₆烷基、C₂-C₆烯基、C₂-C₆炔基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基或5元或6元杂芳基,并且T₅任选地被一个或多个选自下组的取代基取代,该组由以下各项组成:卤代、C₁-C₆烷基、羟基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、二-C₁-C₆烷基氨基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基和5元或6元杂芳基,除了当T₅是H、卤代、羟基、或氰基时;或者-Q₅-T₅是氧代;并且

[0200] R₈是H、卤代、羟基、COOH、氰基、RS₆、OR_{S6}或COOR_{S6},其中RS₆是C₁-C₆烷基、C₂-C₆烯基、C₂-C₆炔基、C₃-C₈环烷基、4至12元杂环烷基、氨基、单-C₁-C₆烷基氨基、或二-C₁-C₆烷基氨基,并且RS₆任选地被一个或多个选自下组的取代基取代,该组由以下各项组成:卤代、羟基、COOH、C(O)O-C₁-C₆烷基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基和二-C₁-C₆烷基氨基;或者R₇和R₈与其所附接的N原子一起形成具有0至2个另外的杂原子的4至11元杂环烷基环,并且由R₇和R₈形成的4至11元杂环烷基环任选地被一个或多个-Q₆-T₆取代,其中Q₆是键、C(O)、C(O)NR_m、NR_mC(O)、S(O)₂或C₁-C₃烷基接头,R_m是H或C₁-C₆烷基,并且T₆是H、卤代、C₁-C₆烷基、羟基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、二-C₁-C₆烷基氨基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基、5元或6元杂芳基或S(O)_pR_p,其中p是0、1、或2并且R_p是C₁-C₆烷基、C₂-C₆烯基、C₂-C₆炔基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基、或5元或6元杂芳基,并且T₆任选地被一个或多个选自下组的取代基取代,该组由以下各项组成:卤代、C₁-C₆烷基、羟基、氰基、C₁-C₆烷氧基、氨基、单-C₁-C₆烷基氨基、二-C₁-C₆烷基氨基、C₃-C₈环烷基、C₆-C₁₀芳基、4至12元杂环烷基和5元或6元杂芳基,除了当T₆是H、卤代、羟基、或氰基时;或者-Q₆-T₆是氧代。

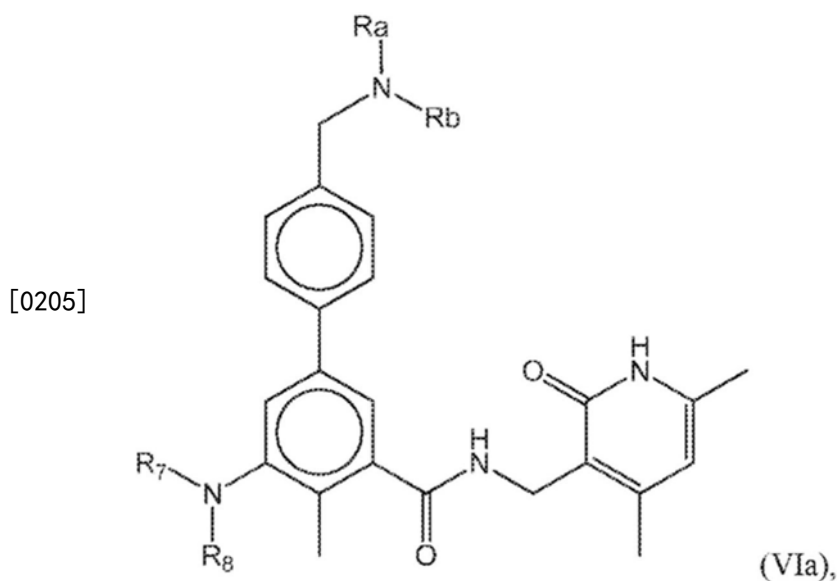
[0201] 在一些实施例中,R₆是C₆-C₁₀芳基或5元或6元杂芳基,这些基团中的每一个任选地独立地被一个或多个-Q₂-T₂取代,其中Q₂是键或C₁-C₃烷基接头,并且T₂是H、卤代、氰基、-OR_a、-NR_aR_b、-(NR_aR_bR_c)⁺A⁻、-C(O)NR_aR_b、-NR_bC(O)R_a、-S(O)₂R_a或RS₂,其中R_a和R_b中的每一个独立地是H或RS₃,RS₂和RS₃中的每一个独立地是C₁-C₆烷基,或者R_a和R_b与其所附接的N原子一起形成具有0或1个另外的杂原子的4至7元杂环烷基环,并且RS₂、RS₃和由R_a和R_b形成的4至7元杂环烷基环中的每一个任选地独立地被一个或多个-Q₃-T₃取代,其中Q₃是键或C₁-C₃烷基接头并且T₃选自下组,该组由以下各项组成:卤代、C₁-C₆烷基、4至7元杂环烷基、OR_d、-S(O)₂R_d和-NR_dR_e,R_d和R_e中的每一个独立地是H或C₁-C₆烷基,或者-Q₃-T₃是氧代;或者任何两个相邻-Q₂-T₂与其所附接的原子一起形成任选地含有1-4个选自N、O和S的杂原子的5元或6元环。

[0202] 在一些实施例中,化合物具有式(VI):



基接头,T₂是H、卤代、-OR_a、-NR_aR_b、-(NR_aR_bR_c)⁺A⁻或-S(O)₂NR_aR_b,R₇是哌啶基、四氢吡喃基、环戊基或环己基,这些基团各自任选地被一个-Q₅-T₅取代,并且R₈是乙基。

[0204] 本披露的一些方面提供了具有式(VIa)的化合物:



[0206] 及其药学上可接受的盐或酯,其中R₇、R₈、R_a和R_b是本文所定义的。

[0207] 具有式(VIa)的化合物可以包括以下特征中的一个或多个:

[0208] 在一些实施例中,R_a和R_b中的每一个独立地是H或任选地被一个或多个-Q₃-T₃取代的C₁-C₆烷基。

[0209] 在一些实施例中,R_a和R_b中之一是H。

[0210] 在一些实施例中,R_a和R_b与其所附接的N原子一起形成具有0或1个相对于N原子另外的杂原子的4至7元杂环烷基环(例如,氮杂环丁烷基、吡咯烷基、咪唑烷基、吡唑烷基、噁唑烷基、异噁唑烷基、三唑烷基、哌啶基、1,2,3,6-四氢吡啶基、哌嗪基、吗啉基、1,4-二氮杂环庚烷基、1,4-氧杂氮杂环庚烷基、2-氧杂-5-氮杂二环[2.2.1]庚烷基、2,5-二氮杂二环[2.2.1]庚烷基等),并且该环任选地被一个或多个-Q₃-T₃取代。

[0211] 在一些实施例中,R_a和R_b与其所附接的N原子一起形成氮杂环丁烷基、吡咯烷基、咪唑烷基、吡唑烷基、噁唑烷基、异噁唑烷基、三唑烷基、四氢呋喃基、哌啶基、1,2,3,6-四氢吡啶基、哌嗪基或吗啉基,并且该环任选地被一个或多个-Q₃-T₃取代。

[0212] 在一些实施例中,一个或多个-Q₃-T₃是氧代。

[0213] 在一些实施例中,Q₃是键或未取代的或取代的C₁-C₃烷基接头。

[0214] 在一些实施例中, T_3 是 H、卤代、4 至 7 元杂环烷基、 C_1 - C_3 烷基、 OR_d 、 $COOR_d$ 、 $-S(O)_2R_d$ 或 $-NR_dR_e$ 。

[0215] 在一些实施例中, R_d 和 R_e 中的每一个独立地是 H 或 C_1 - C_6 烷基。

[0216] 在一些实施例中, R_7 是 C_3 - C_8 环烷基或 4 至 7 元杂环烷基, 这些基团各自任选地被一个或多个 $-Q_5-T_5$ 取代。

[0217] 在一些实施例中, R_7 是哌啶基、四氢吡喃、四氢-2H-噻喃基、环戊基、环己基、吡咯烷基或环庚基, 这些基团各自任选地被一个或多个 $-Q_5-T_5$ 取代。

[0218] 在一些实施例中, R_7 是环戊基、环己基或四氢-2H-噻喃基, 这些基团中的每一个任选地被一个或多个 $-Q_5-T_5$ 取代。

[0219] 在一些实施例中, Q_5 是 $NHC(O)$ 并且 T_5 是 C_1 - C_6 烷基或 C_1 - C_6 烷氧基, 每个在一些实施例中, 一个或多个 $-Q_5-T_5$ 是氧代。

[0220] 在一些实施例中, R_7 是 1-氧化-四氢-2H-噻喃基或 1,1-二氧化-四氢-2H-噻喃基。

[0221] 在一些实施例中, Q_5 是键并且 T_5 是氨基、单- C_1 - C_6 烷基氨基、二- C_1 - C_6 烷基氨基。

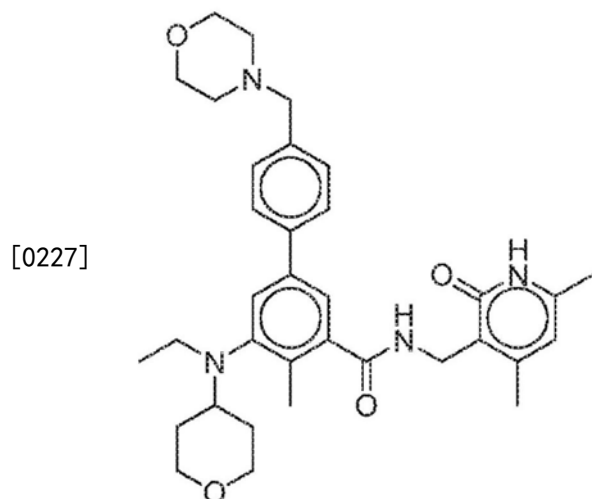
[0222] 在一些实施例中, Q_5 是 CO 、 $S(O)_2$ 、或 $NHC(O)$; 并且 T_5 是 C_1 - C_6 烷基、 C_1 - C_6 烷氧基、 C_3 - C_8 环烷基、或 4 至 7 元杂环烷基。

[0223] 在一些实施例中, R_8 是 H 或 C_1 - C_6 烷基, 该 C_1 - C_6 烷基任选地被一个或多个选自下组的取代基取代, 该组由以下各项组成: 卤代、羟基、 $COOH$ 、 $C(O)O$ - C_1 - C_6 烷基、氰基、 C_1 - C_6 烷氧基、氨基、单- C_1 - C_6 烷基氨基、和二- C_1 - C_6 烷基氨基。

[0224] 在一些实施例中, R_8 是 H、甲基或乙基。

[0225] 适用于本文提供的策略、治疗形式、方法、组合和组合物中的具有式 (I)-(VIa) 的其他化合物描述于美国公开 20120264734 中, 将其内容据此通过援引以其全文并入。具有式 (I)-(VIa) 的化合物适于与一种或多种其他治疗剂, 例如与如本文提供的免疫检查点抑制剂作为组合疗法的一部分给予。

[0226] 在本文提供的策略、治疗形式、方法、组合和组合物的一些实施例中, EZH2 抑制剂是化合物 44



[0228] 或其药学上可接受的盐。化合物 44 还可以称之为他托司他、EPZ 006438 或 6438。

[0229] 如本文所述的化合物 44 或其药学上可接受的盐有效地靶向野生型和突变型 EZH2 两者。与其他组蛋白甲基转移酶相比, 化合物 44 是口服生物可利用性的并且对 EZH2 具有高

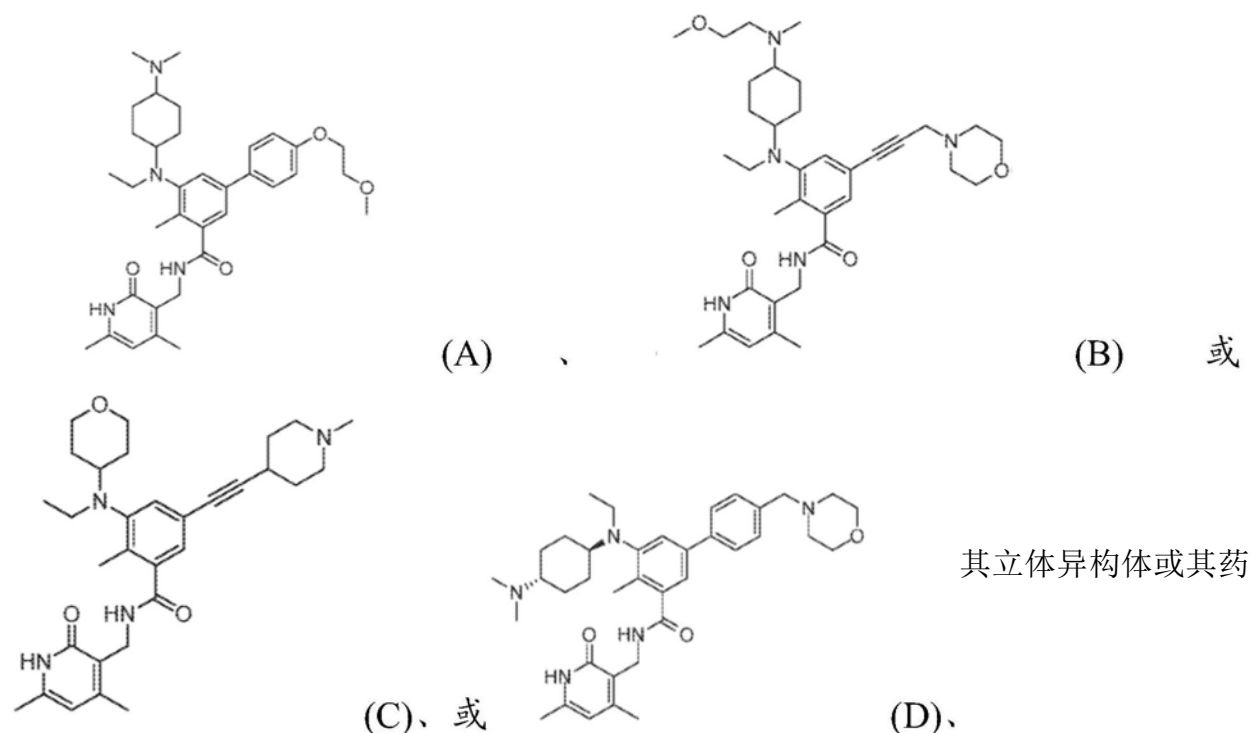
选择性(即,通过 K_i 计算的 $>20,000$ 倍选择性)。重要的是,化合物44具有靶甲基标记抑制作用,该抑制作用导致在体外杀伤遗传上确定的癌细胞。动物模型在抑制靶甲基标记后也显示了持续的体内功效。

[0230] 在一些实施例中,以每日大约100mg至大约3200mg,诸如约100mg BID至约1600mg BID(例如,100mg BID、200mg BID、400mg BID、800mg BID、或1600mg BID)的剂量向受试者给予化合物44或其药学上可接受的盐,用于治疗生发中心来源的淋巴瘤。

[0231] 在一些实施例中,将化合物44或其药学上可接受的盐与本文提供的免疫检查点抑制剂组合(同时或依序地)给予至受试者。

[0232] 在一些实施例中,可以用于在此呈现的策略、治疗形式、方法、组合和组合物中的化合物是:

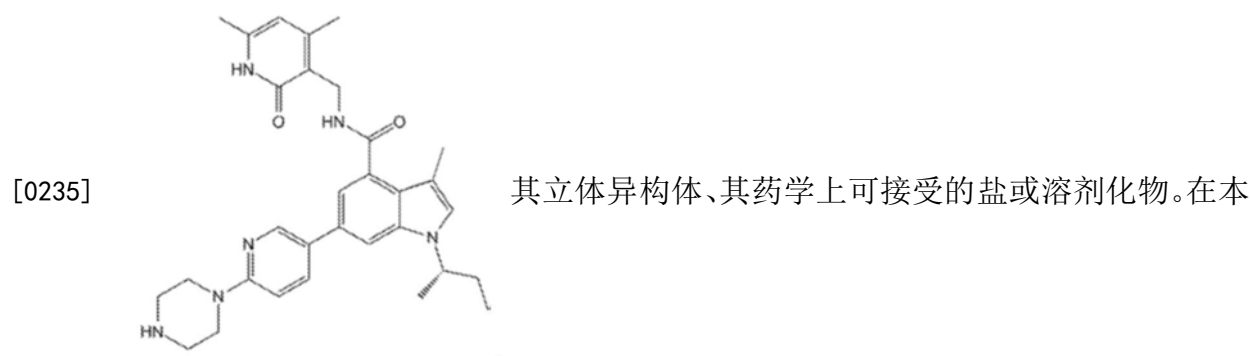
[0233]



其立体异构体或其药

学上可接受的盐和溶剂化物。

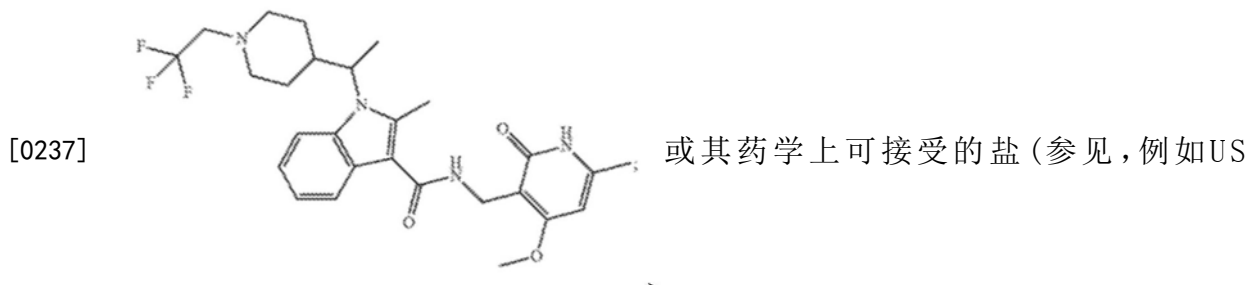
[0234] 在一些实施例中,EZH2抑制剂可以包括以下、基本上由以下组成或由以下组成:具有下式的GSK-126:



文提供的策略、治疗形式、方法、组合和组合物的一些实施例中,EZH2抑制剂是US 8,536,179(描述了GSK-126等化合物并且对应于WO 2011/140324)中所述的EZH2抑制剂,将每篇专

利的全部内容通过援引以其全文并入本文。

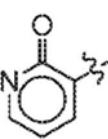
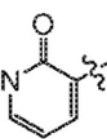
[0236] 在本文提供的策略、治疗形式、方法、组合和组合物的一些实施例中,EZH2抑制剂是公布为W0/2014/124418的PCT/US 2014/015706、公布为W0/2013/120104的PCT/US 2013/025639以及公布为US 2015/0368229的US 14/839,273中所述的EZH2抑制剂,将每篇专利的全部内容通过引用以其全文并入本文。在本文提供的策略、治疗形式、方法、组合和组合物的一些实施例中,EZH2抑制剂是具有下式的化合物:



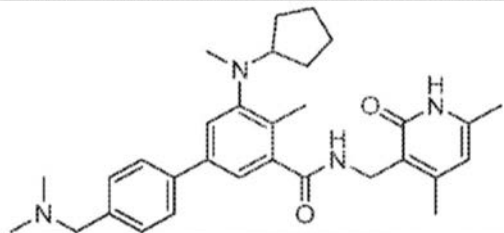
2015/0368229,将其内容并入本文)。

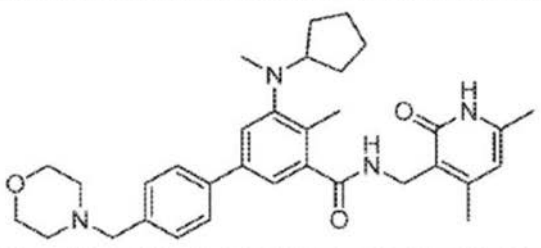
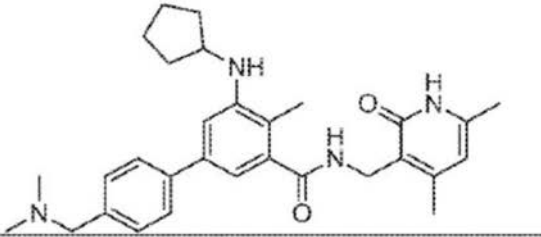
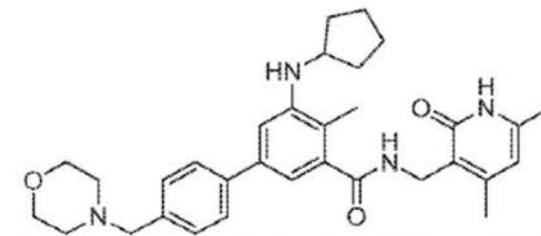
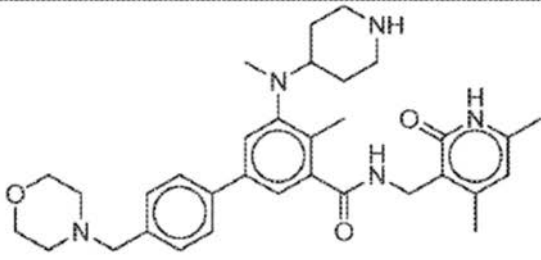
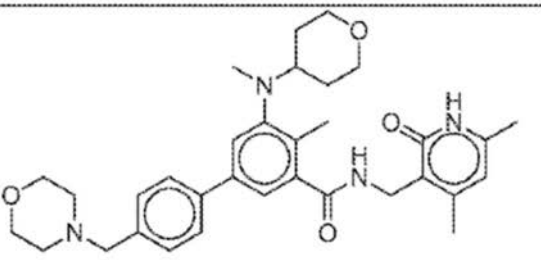
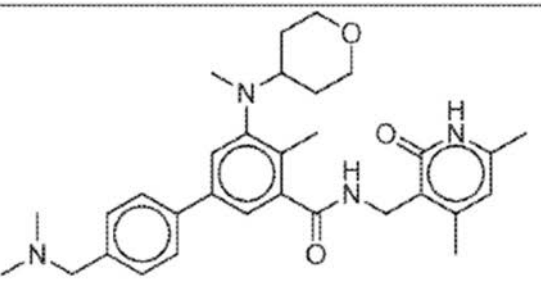
[0238] 在一些实施例中,EZH2抑制剂是作为化合物本身,即作为游离碱或“裸”分子使用的小分子。在一些实施例中,EZH2抑制剂是其盐,例如,裸分子的单-HCl盐或三-HCl盐、单-HBr盐或三-HBr盐。

[0239] 适于本文提供的策略、治疗形式、方法、组合和组合物的代表性化合物包括列于表

1中的化合物。在下表中,  的每次出现应解释为  。

[0240] 表1

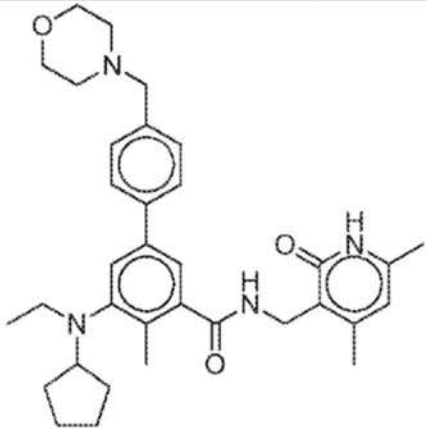
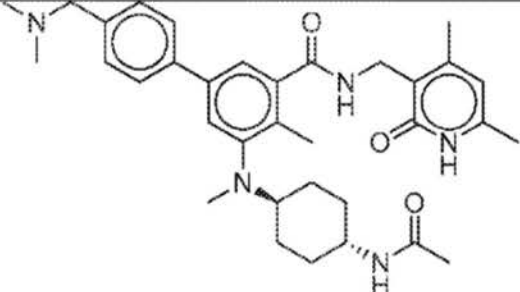
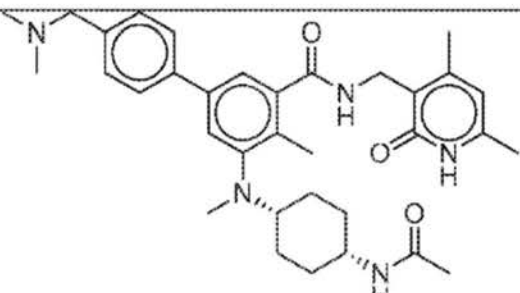
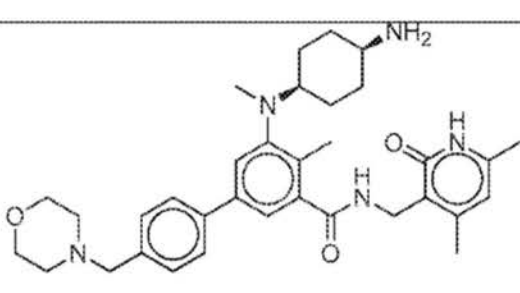
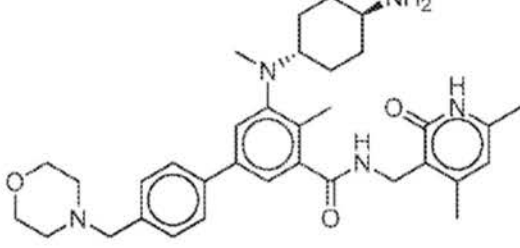
化合物 编号	结构	MS (M+1) ⁺
1		501.39

化合物 编号	结构	MS (M+1) ⁺
2		543.22
3		486.21
4		529.30
11		558.45
12		559.35
13		517.3

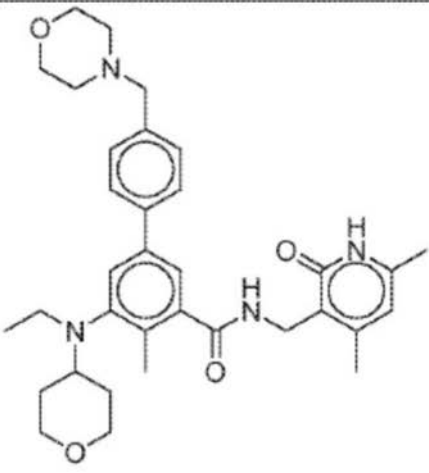
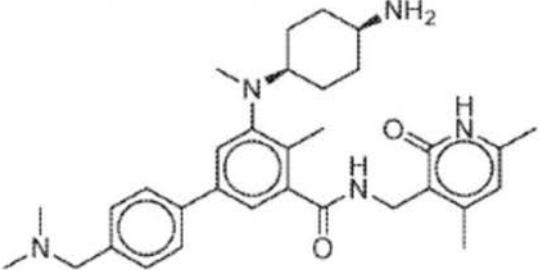
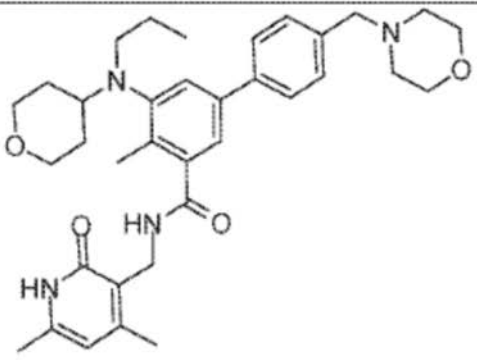
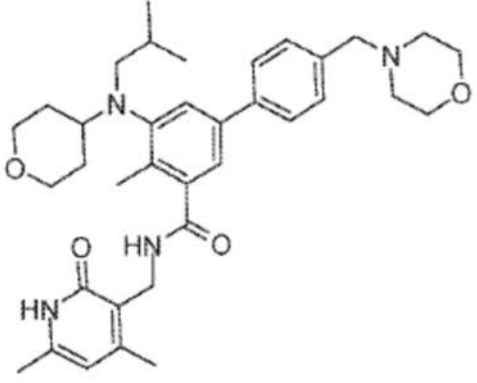
[0242]

化合物 编号	结构	MS (M+1) ⁺
14		557.4
16		515.4
20		614.4
21		614.4
27		516.35

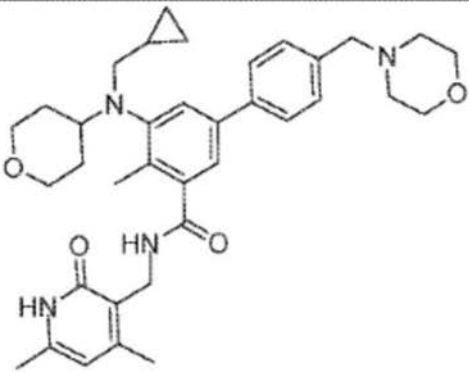
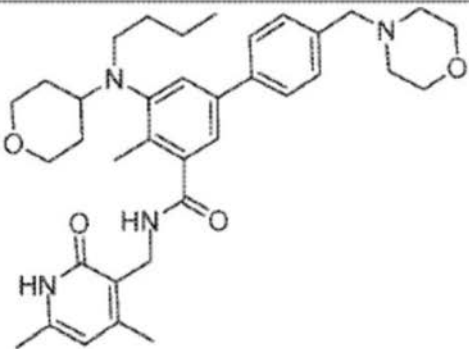
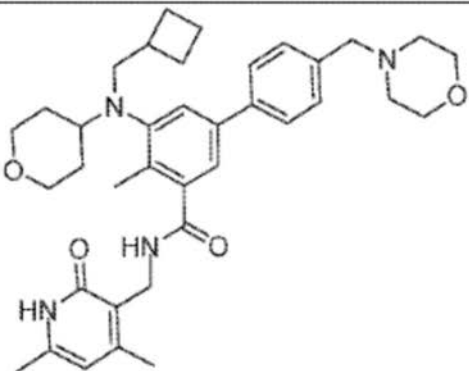
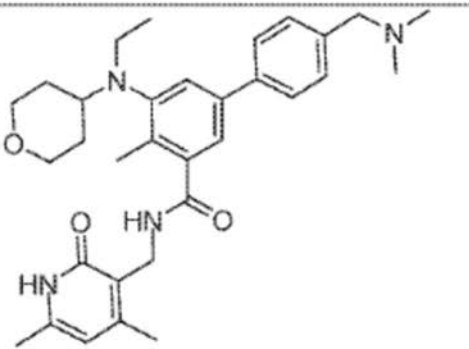
[0243]

化合物 编号	结构	MS (M+1) ⁺
36		557.35
39		572.35
40		572.35
42		572.4
43		572.6

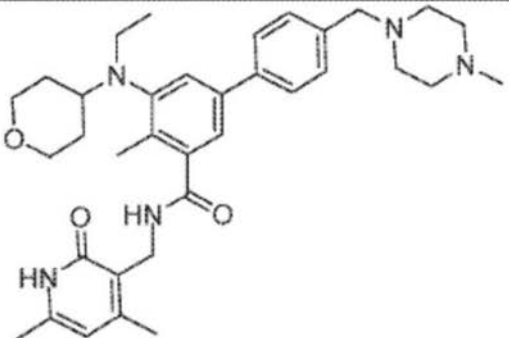
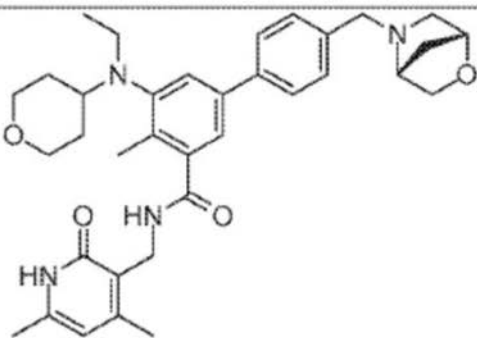
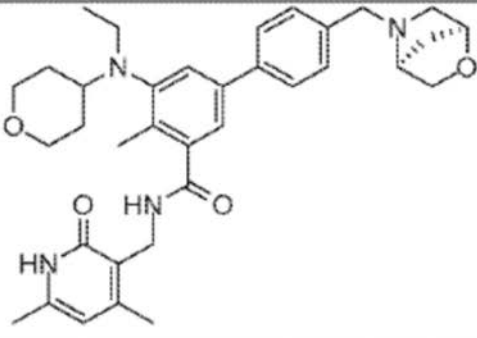
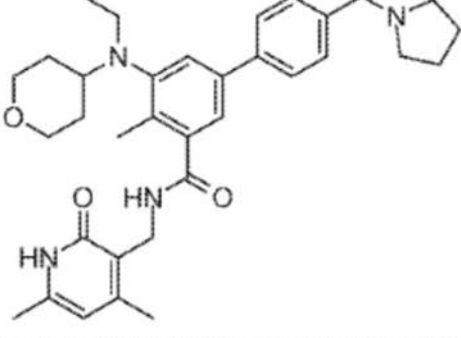
[0244]

化合物 编号	结构	MS (M+1) ⁺
44		573.40
47		530.35
59		587.40
60		601.30

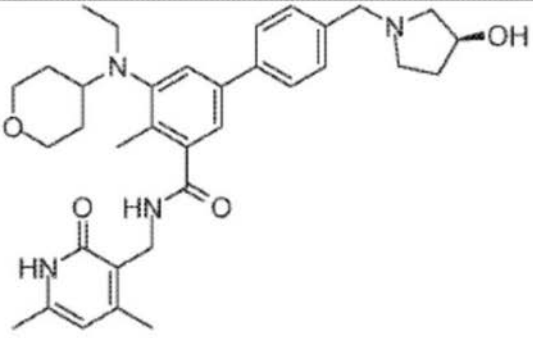
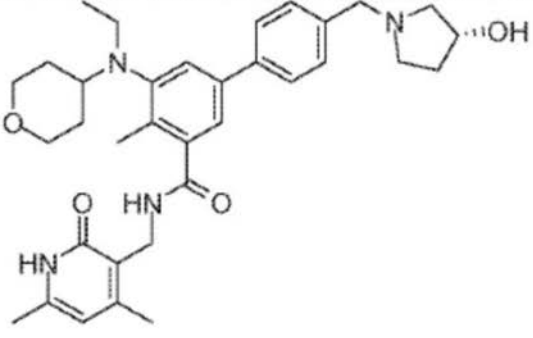
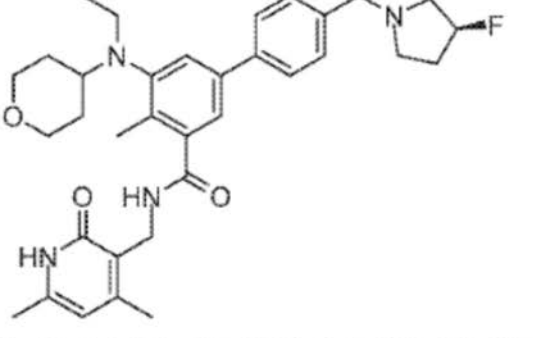
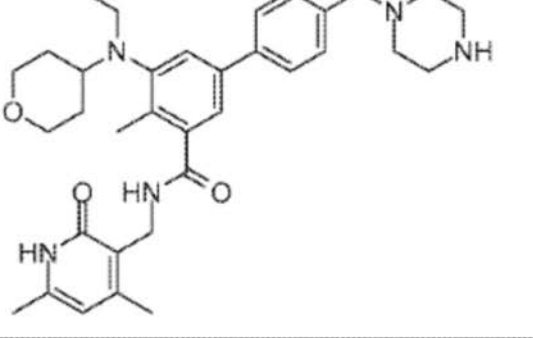
[0245]

化合物 编号	结构	MS (M+1) ⁺
61		599.35
62		601.35
63		613.35
65		531.30

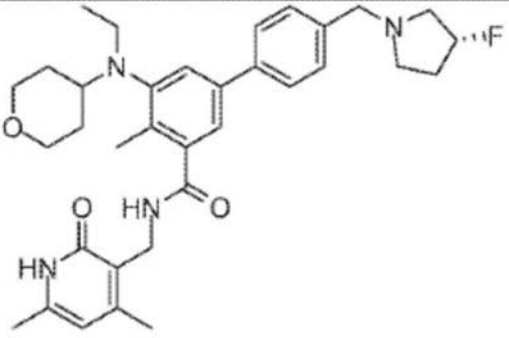
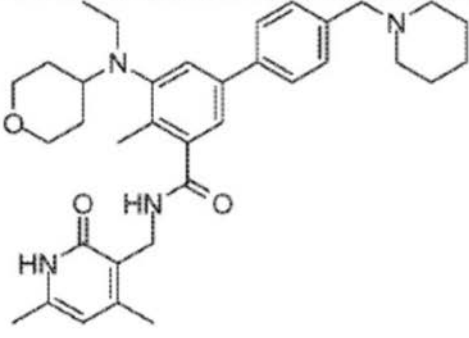
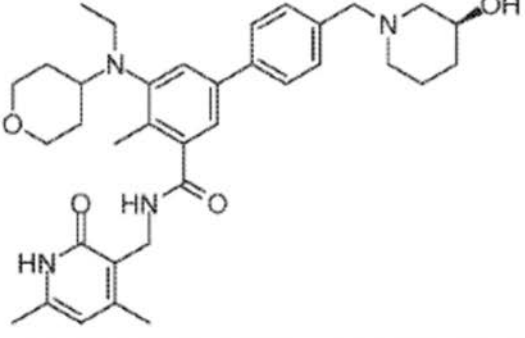
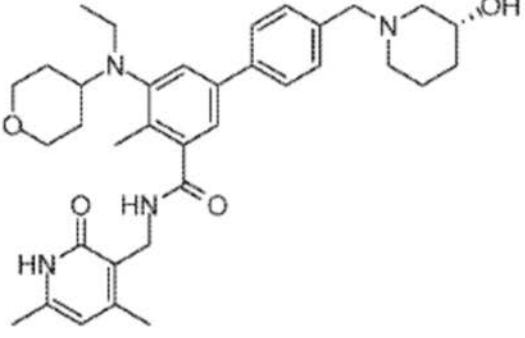
[0246]

化合物 编号	结构	MS (M+1) ⁺
66		586.40
67		585.25
68		585.35
69		557.25

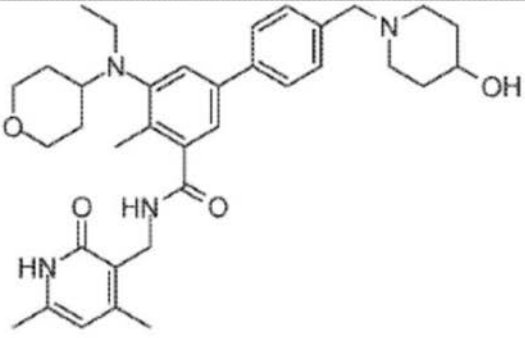
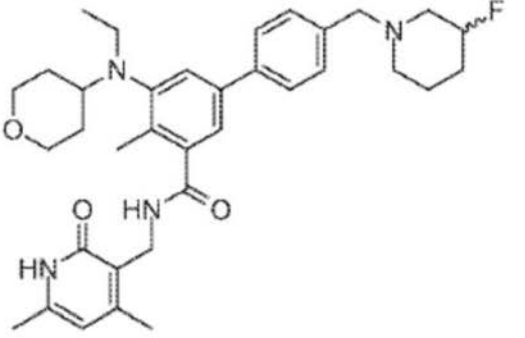
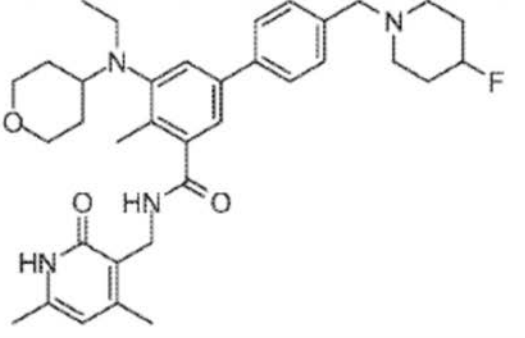
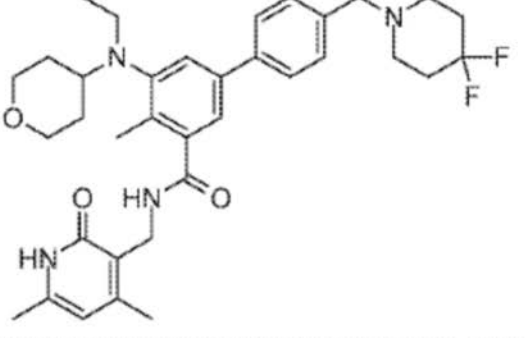
[0247]

化合物 编号	结构	MS (M+1) ⁺
70		573.40
71		573.40
72		575.35
73		572.10

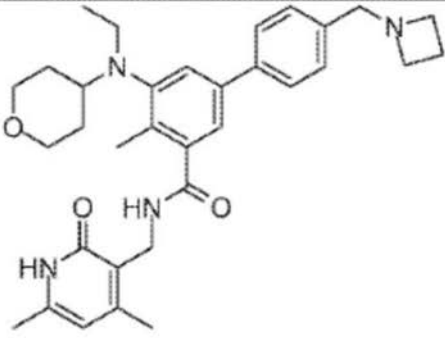
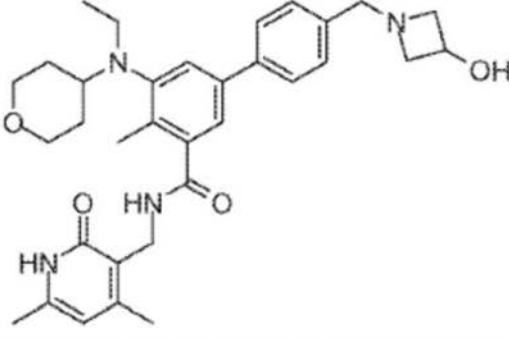
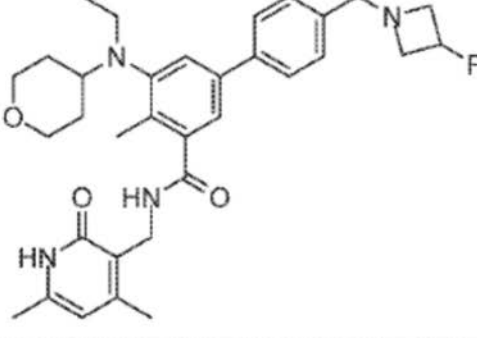
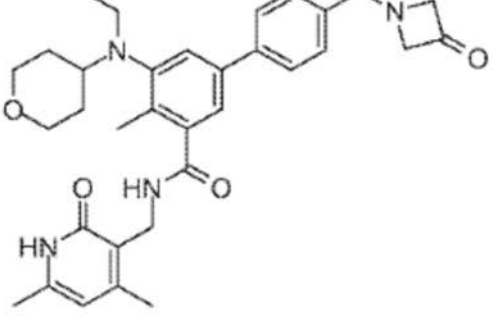
[0248]

化合物 编号	结构	MS (M+1) ⁺
74		575.35
75		571.25
76		587.40
77		587.45

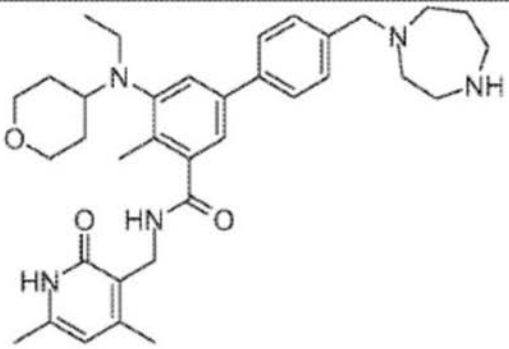
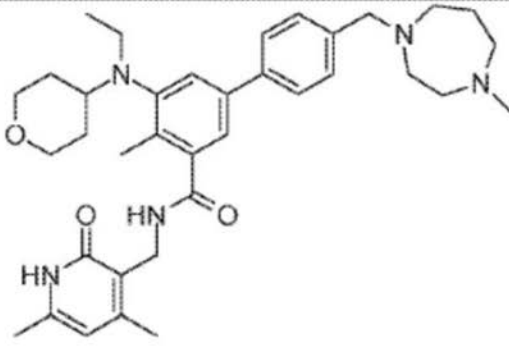
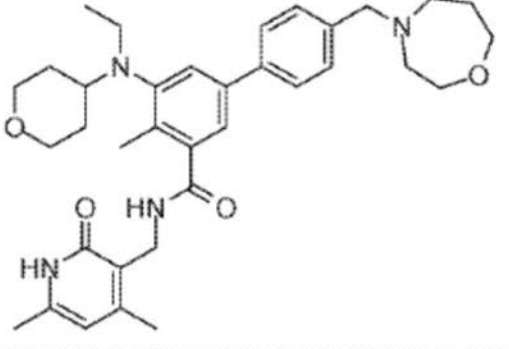
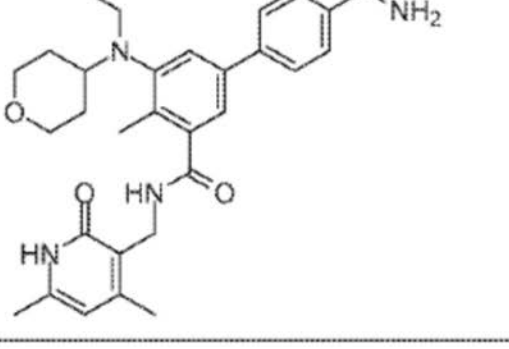
[0249]

化合物 编号	结构	MS (M+1) ⁺
78		587.20
79		589.35
80		589.30
81		607.35

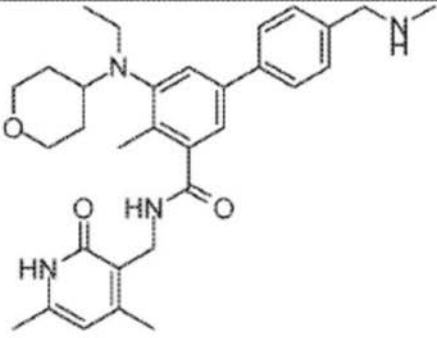
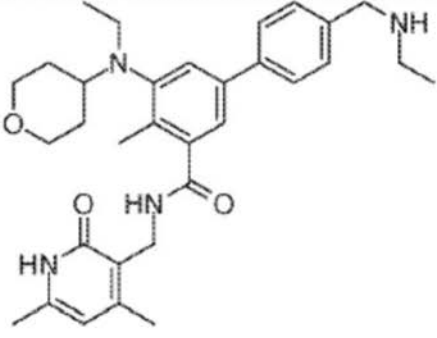
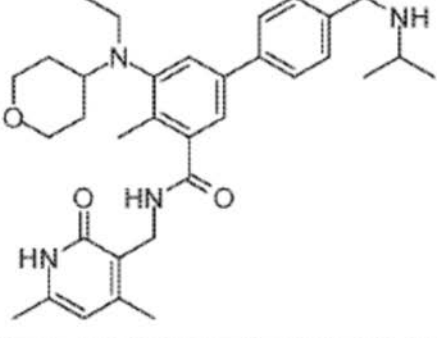
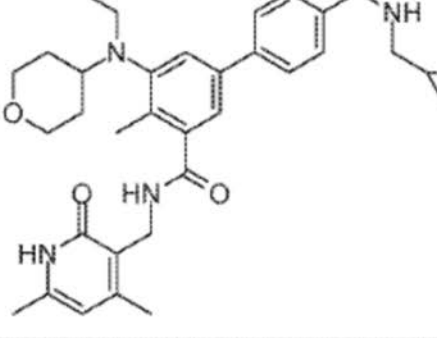
[0250]

化合物 编号	结构	MS (M+1) ⁺
82		543.40
83		559.80
84		561.25
85		

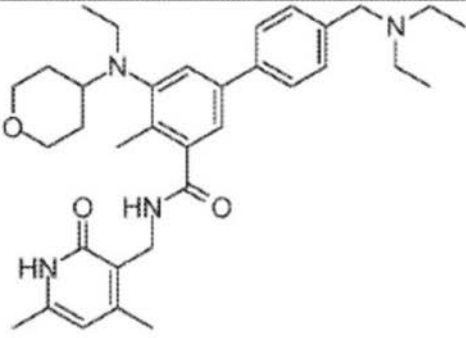
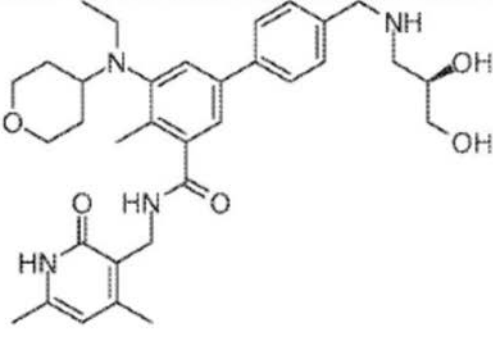
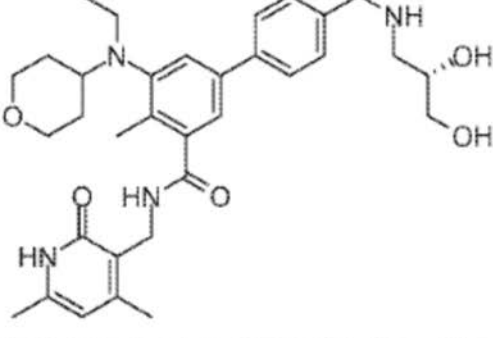
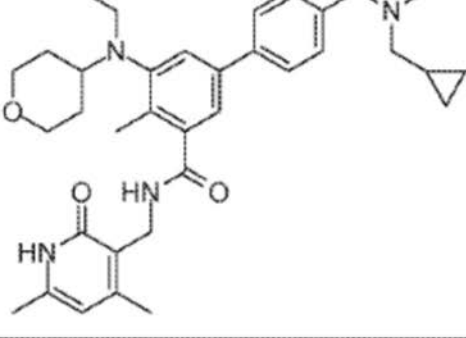
[0251]

化合物 编号	结构	MS (M+1) ⁺
86		585.37
87		600.30
88		587.40
89		503.40

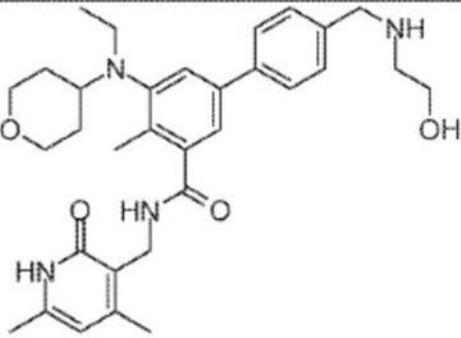
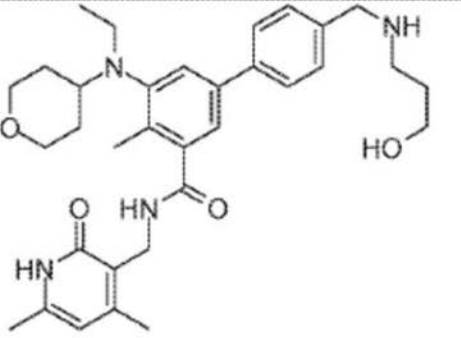
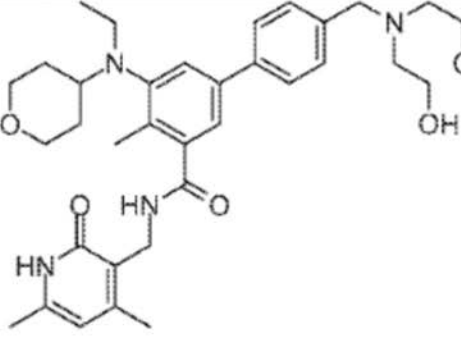
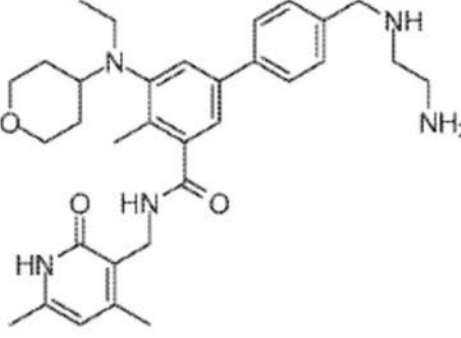
[0252]

化合物 编号	结构	MS (M+1) ⁺
90		517.30
91		531.35
92		545.40
93		557.35

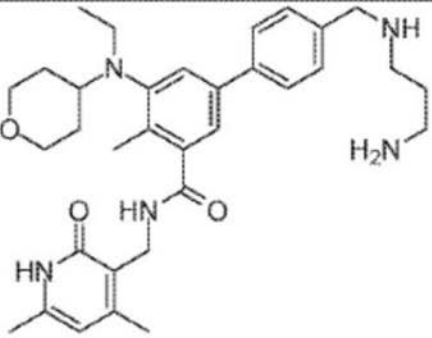
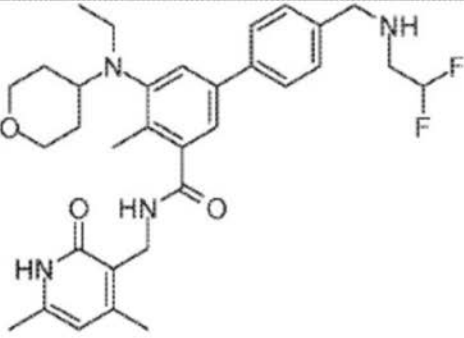
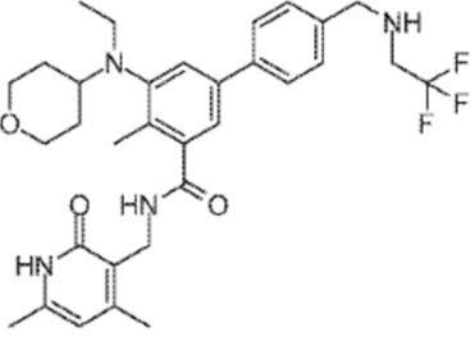
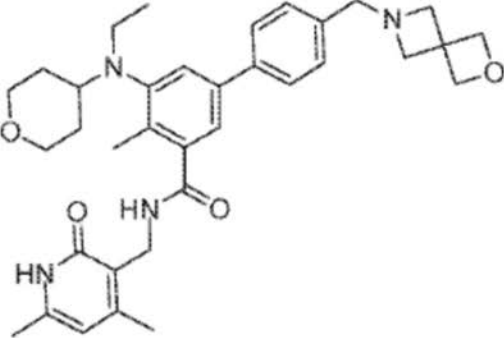
[0253]

化合物 编号	结构	MS (M+1) ⁺
94		559.20
95		599.35 (M+Na)
96		577.25
97		571.40

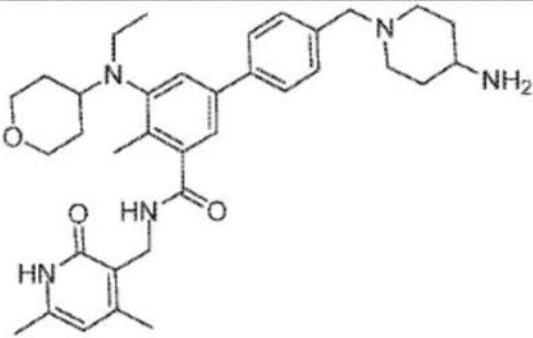
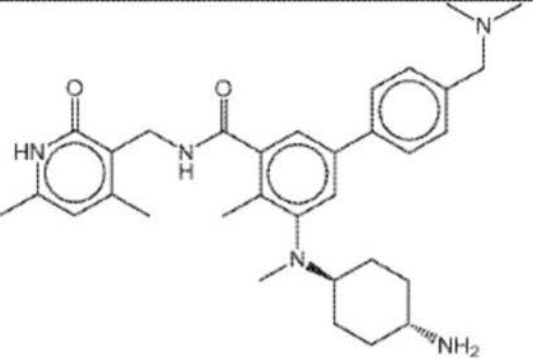
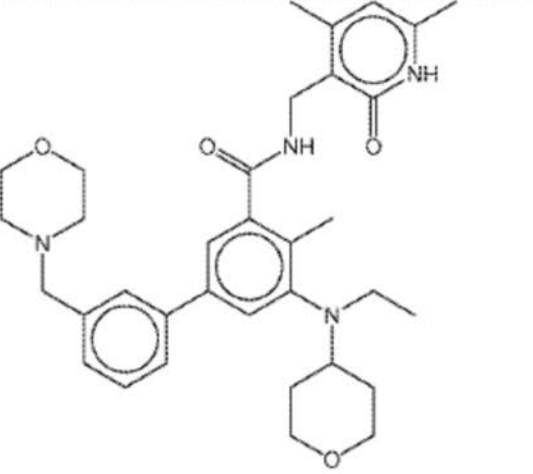
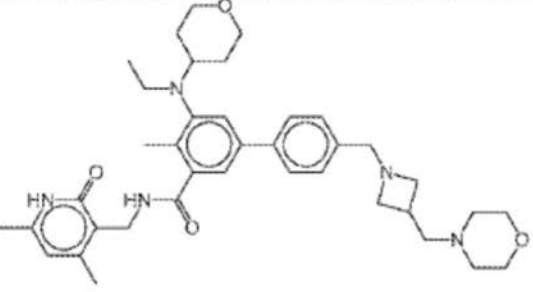
[0254]

化合物 编号	结构	MS (M+1) ⁺
98		547.35
99		561.30
100		591.25
101		546.35

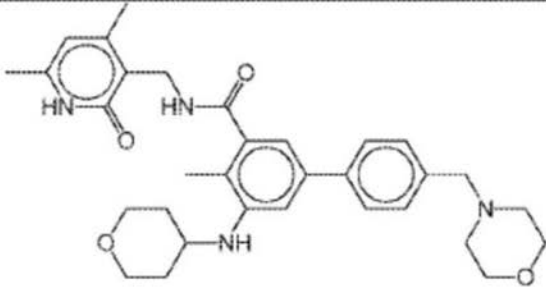
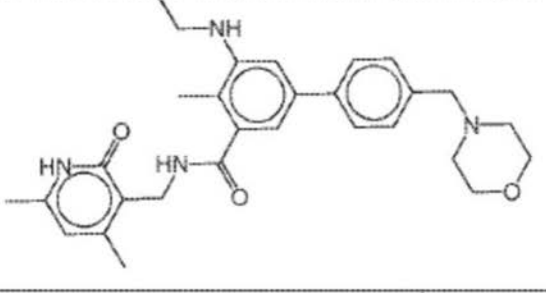
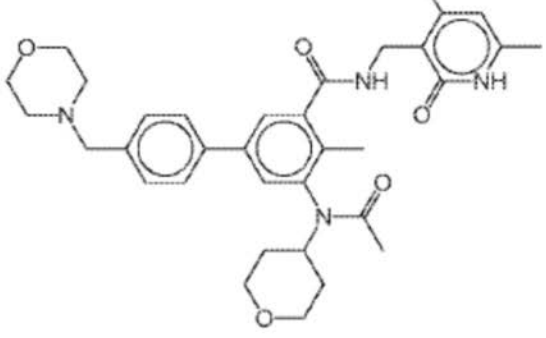
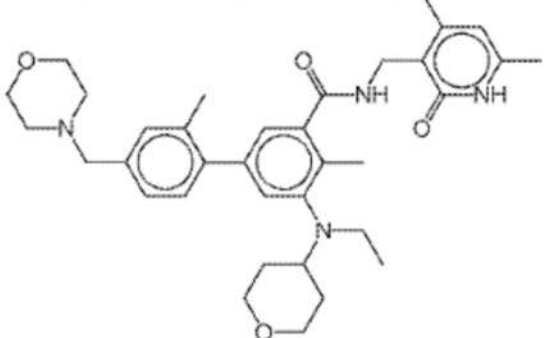
[0255]

化合物 编号	结构	MS (M+1) ⁺
102		560.20
103		567.30
104		585.25
105		585.40

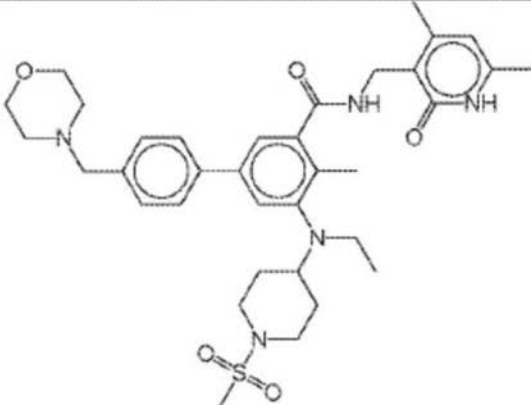
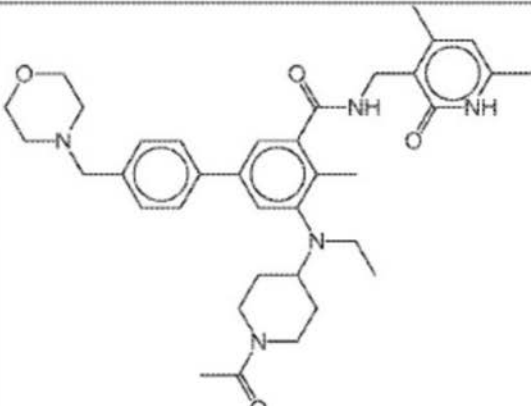
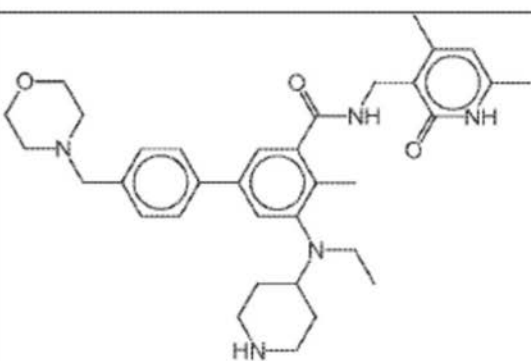
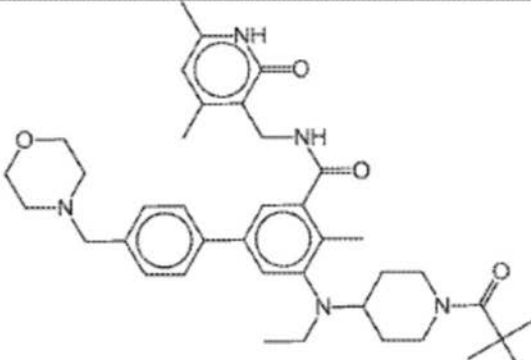
[0256]

化合物 编号	结构	MS (M+1) ⁺
107		
108		530.35
114		573.25
115		642.45

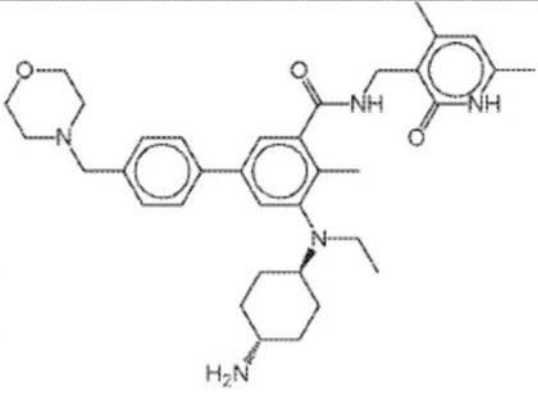
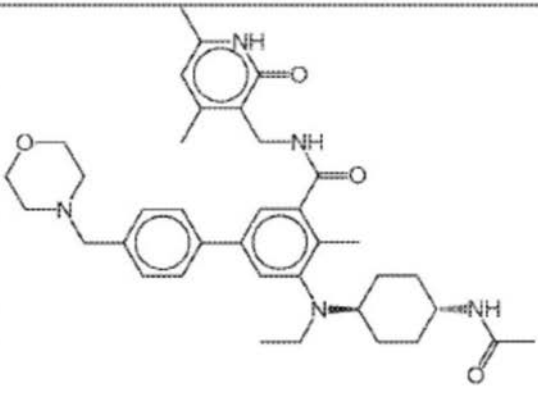
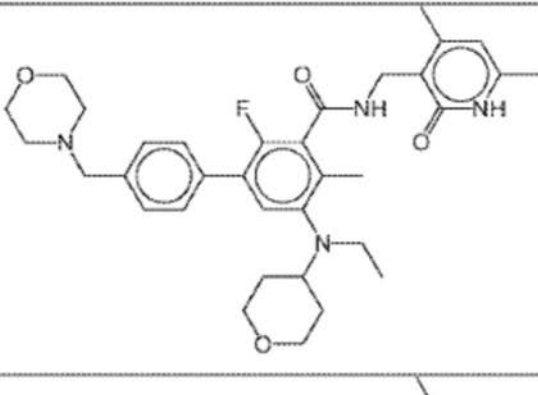
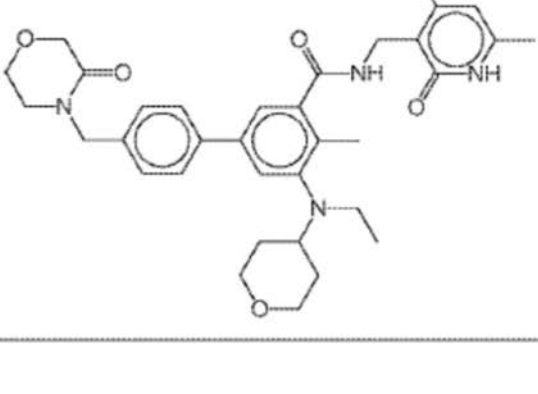
[0257]

化合物 编号	结构	MS (M+1) ⁺
116		545.15
117		489.20
119		609.35
122		587.55

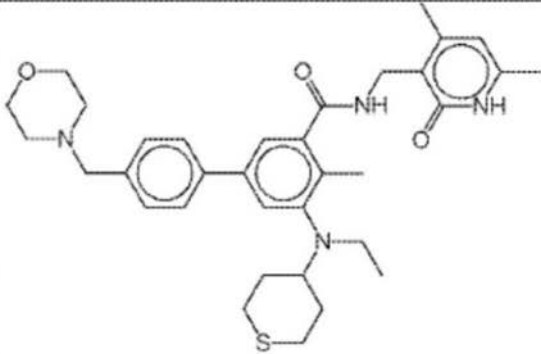
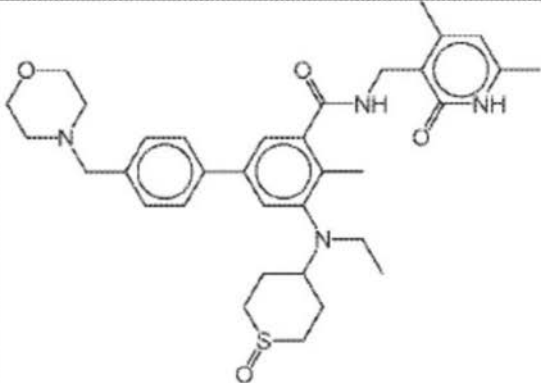
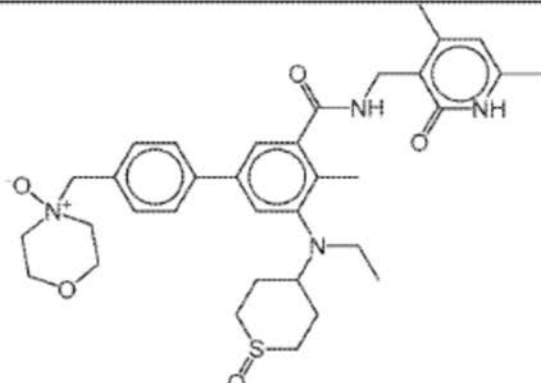
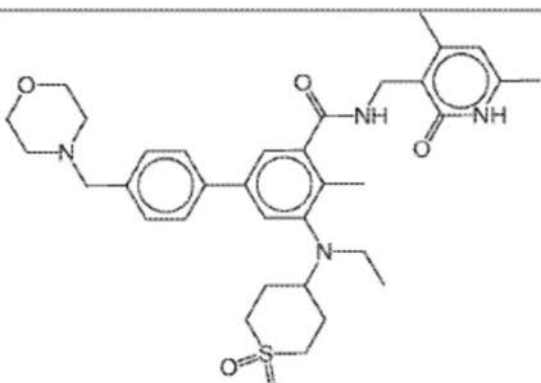
[0258]

化合物 编号	结构	MS (M+1) ⁺
124		650.85
125		614.75
126		572.35
127		656.65

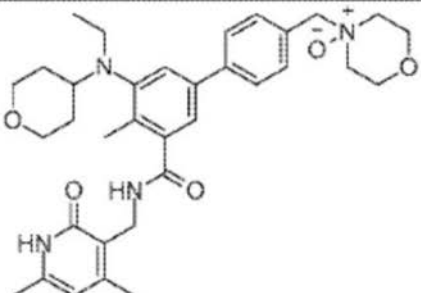
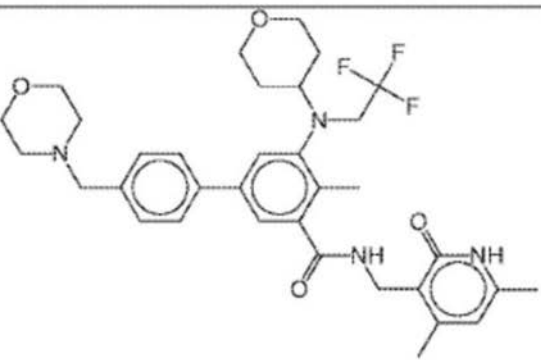
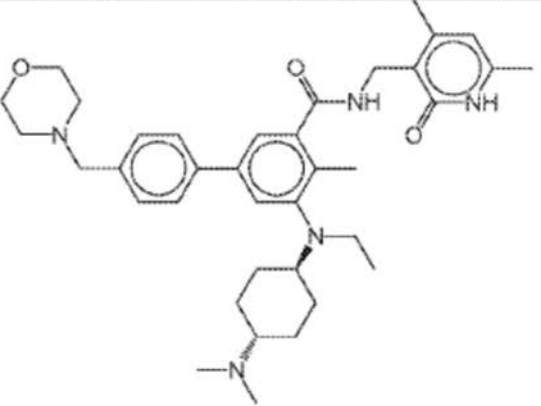
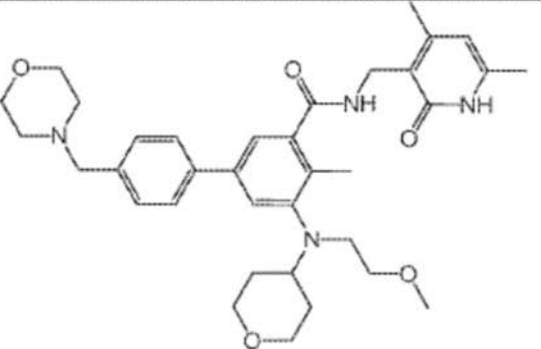
[0259]

化合物 编号	结构	MS (M+1) ⁺
128		586.45
129		628.35
130		591.2
131		587.35

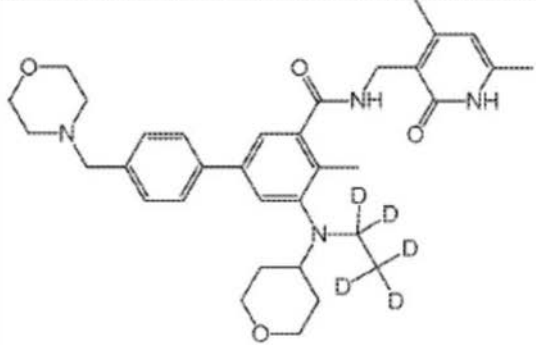
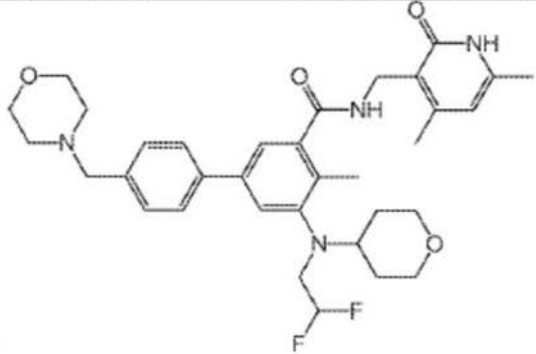
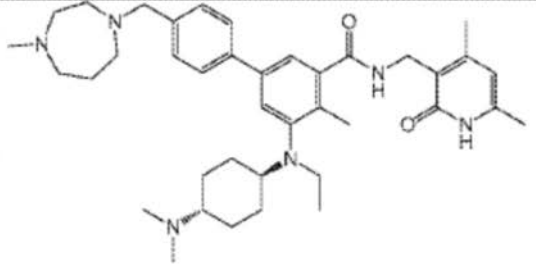
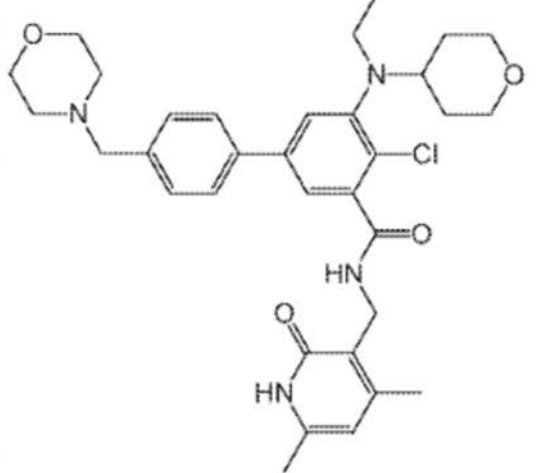
[0260]

化合物 编号	结构	MS (M+1) ⁺
132		589.25
133		605.25
135		621.40
136		621.45

[0261]

化合物 编号	结构	MS (M+1) ⁺
137		589.35
138		627.5
141		614.65
142		603.45

[0262]

化合物 编号	结构	MS (M+1) ⁺
143		578.35
144		609.15
146		641.50
178		593.60

[0264] 如本文所用,“烷基”、“C₁、C₂、C₃、C₄、C₅或C₆烷基”或“C₁-C₆烷基”旨在包括C₁、C₂、C₃、C₄、C₅或C₆直链(线性)饱和脂肪族烃基和C₃、C₄、C₅或C₆支链饱和脂肪族烃基。例如,C₁-C₆烷基旨在包括C₁、C₂、C₃、C₄、C₅和C₆烷基基团。烷基的实例包括具有从一至六个碳原子的部分,诸

如但不限于甲基、乙基、正丙基、异丙基、正丁基、仲丁基、叔丁基、正戊基、仲戊基或正己基。

[0265] 在某些实施例中,直链或支链烷基具有六个或更少碳原子(例如,对于直链为 C_1-C_6 ,对于支链为 C_3-C_6),并且在一些实施例中,直链或支链烷基具有四个或更少碳原子。

[0266] 如本文所用,术语“环烷基”是指具有3至30个碳原子(例如, C_3-C_{10})的饱和或不饱和和非芳香族烃单环或多环(例如,稠合环、桥接环或螺环)体系。环烷基的实例包括但不限于环丙基、环丁基、环戊基、环己基、环庚基、环辛基、环戊烯基、环己烯基、环庚烯基和金刚烷基。除非另外指定,否则术语“杂环烷基”是指具有一个或多个杂原子(诸如O、N、S或Se)的饱和或不饱和和非芳香族3-8元单环、7-12元二环(稠合环、桥接环或螺环)或11-14元三环体系(稠合环、桥接环或螺环)。杂环烷基基团的实例包括但不限于哌啶基、哌嗪基、吡咯烷基、二噁烷基、四氢呋喃基、异吡啶基、二氢吡啶基、咪唑烷基、吡唑烷基、噁唑烷基、异噁唑烷基、三唑烷基、四氢呋喃基、环氧乙烷基、氮杂环丁烷基、氧杂环丁烷基、硫杂环丁烷基、1,2,3,6-四氢吡啶基、四氢吡喃基、二氢吡喃基、吡喃基、吗啉基、1,4-二氮杂环庚烷基、1,4-氧杂氮杂环庚烷基、2-氧杂-5-氮杂二环[2.2.1]庚烷基、2,5-二氮杂二环[2.2.1]庚烷基、2-氧杂-6-氮杂螺[3.3]庚烷基、2,6-二氮杂螺[3.3]庚烷基、1,4-二氮杂8-氮杂螺[4.5]癸烷基等。

[0267] 术语“任选取代的烷基”是指未取代的烷基或具有指定取代基置换烃主链的一个或多个碳上的一个或多个氢原子的烷基。此类取代基可以包括例如烷基、烯基、炔基、卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、磷酸酯基、膦酸基、次膦酸基、氨基(包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基)、酰基氨基(包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基)、脒基、亚氨基、硫基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。

[0268] “芳基烷基”或“芳烷基”部分是被芳基取代的烷基(例如,苯基甲基(苄基))。“烷基芳基”部分是被烷基取代的芳基(例如,甲基苯基)。

[0269] 如本文所用,“烷基接头”旨在包括 C_1 、 C_2 、 C_3 、 C_4 、 C_5 或 C_6 直链(线性)饱和和二价脂肪族烃基和 C_3 、 C_4 、 C_5 或 C_6 支链饱和脂肪族烃基。例如, C_1 - C_6 烷基接头旨在包括 C_1 、 C_2 、 C_3 、 C_4 、 C_5 和 C_6 烷基接头基团。烷基接头的实例包括具有从一至六个原子的部分,诸如但不限于甲基($-CH_2-$)、乙基($-CH_2CH_2-$)、正丙基($-CH_2CH_2CH_2-$)、异丙基($-CHCH_3CH_2-$)、正丁基($-CH_2CH_2CH_2CH_2-$)、仲丁基($-CHCH_3CH_2CH_2-$)、异丁基($-C(CH_3)_2CH_2-$)、正戊基($-CH_2CH_2CH_2CH_2CH_2-$)、仲戊基($-CHCH_3CH_2CH_2CH_2-$)或正己基($-CH_2CH_2CH_2CH_2CH_2CH_2-$)。

[0270] “烯基”包括在长度和可能的取代方面与上述烷基相似但含有至少一个双键的不饱和脂肪族基团。例如,术语“烯基”包括直链烯基基团(例如,乙烯基、丙烯基、丁烯基、戊烯基、己烯基、庚烯基、辛烯基、壬烯基、癸烯基)和支链烯基基团。在某些实施例中,直链或支链烯基基团在其主链中具有六个或更少碳原子(例如,对于直链为 C_2-C_6 ,对于支链为 C_3-C_6)。术语“ C_2-C_6 ”包括含有二至六个碳原子的烯基基团。术语“ C_3-C_6 ”包括含有三至六个碳原子的烯基基团。

[0271] 术语“任选取代的烯基”是指未取代的烯基或具有置换一个或多个烃主链碳原子上的一个或多个氢原子的指定取代基的烯基。此类取代基可以包括例如烷基、烯基、炔基、

卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、磷酸酯基、膦酸基、次膦酸基、氨基(包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基)、酰基氨基(包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基)、脒基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。

[0272] “炔基”包括在长度和可能的取代方面与上述烷基相似但含有至少一个三键的不饱和脂肪族基团。例如,“炔基”包括直链炔基基团(例如,乙炔基、丙炔基、丁炔基、戊炔基、己炔基、庚炔基、辛炔基、壬炔基、癸炔基)和支链炔基基团。在某些实施例中,直链或支链炔基基团在其主链中具有六个或更少碳原子(例如,对于直链为 C_2-C_6 ,对于支链为 C_3-C_6)。术语“ C_2-C_6 ”包括含有二至六个碳原子的炔基基团。术语“ C_3-C_6 ”包括含有三至六个碳原子的炔基基团。

[0273] 术语“任选取代的炔基”是指未取代的炔基或具有置换一个或多个烃主链碳原子上的一个或多个氢原子的指定取代基的炔基。此类取代基可以包括例如烷基、烯基、炔基、卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、磷酸酯基、膦酸基、次膦酸基、氨基(包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基)、酰基氨基(包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基)、脒基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。

[0274] 其他任选取代的部分(例如任选取代的环烷基、杂环烷基、芳基或杂芳基)包括未取代的部分和具有指定取代基中的一个或多个的部分二者。例如,取代的杂环烷基包括被一个或多个烷基基团取代的那些,诸如2,2,6,6-四甲基-哌啶基和2,2,6,6-四甲基-1,2,3,6-四氢吡啶基。

[0275] “芳基”包括具有芳香性的基团,包括具有至少一个芳香族环的“共轭的”或多环体系并且在环结构中不含任何杂原子。实例包括苯基、苄基、1,2,3,4-四氢萘基等。

[0276] “杂芳基”基团是如上文所定义的芳基基团,但在环结构中具有从一至四个杂原子,并且还可称为“芳基杂环”或“杂芳香族基团”。如本文所用,术语“杂芳基”旨在包括由碳原子以及一个或多个杂原子例如1或1-2或1-3或1-4或1-5或1-6个杂原子,或例如1、2、3、4、5或6个杂原子组成的稳定的5元、6元或7元单环或7元、8元、9元、10元、11元或12元双环芳香族杂环,这些杂原子独立地选自氮、氧和硫组成的组。氮原子可以是取代的或未取代的(即, N 或 NR ,其中 R 是 H 或如所定义的其他取代基)。氮和硫杂原子可以任选地被氧化(即, $N \rightarrow O$ 和 $S(O)_p$,其中 $p=1$ 或 2)。应注意,芳香族杂环中 S 和 O 原子的总数不大于1。

[0277] 杂芳基基团的实例包括吡咯、呋喃、噻吩、噻唑、异噻唑、咪唑、三唑、四唑、吡唑、噁唑、异噁唑、吡啶、吡嗪、哒嗪、嘧啶等。

[0278] 此外,术语“芳基”和“杂芳基”包括多环芳基和杂芳基基团,例如三环、二环,例如萘、苯并噁唑、苯并二噁唑、苯并噻唑、苯并咪唑、苯并噻吩、亚甲基二氧苯基、喹啉、异喹啉、茶啉、吲哚、苯并呋喃、嘌呤、苯并呋喃、脱氮杂嘌呤、中氮茚。

[0279] 在多环芳香族环的情况中,仅需要环中的一个芳香族(例如,2,3-二氢吲哚),但

所有环都可以是芳香族(例如,喹啉)。第二环也可以是稠合或桥接的。

[0280] 环烷基、杂环烷基、芳基或杂芳基环可以在一个或多个环位置(例如,成环碳或杂原子,诸如N)被如上所述的取代基取代,这些取代基是例如烷基、烯基、炔基、卤素、羟基、烷氧基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、烷基氨基羰基、芳烷基氨基羰基、烯基氨基羰基、烷基羰基、芳基羰基、芳烷基羰基、烯基羰基、烷氧基羰基、氨基羰基、烷基硫代羰基、磷酸酯基、膦酸基、次膦酸基、氨基(包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基)、酰基氨基(包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基)、脒基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、亚磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。芳基和杂芳基基团也可以与不是芳香族的脂环族环或杂环稠合或桥接,以形成多环体系(例如,四氢萘、亚甲基二氧苯基)。

[0281] 如本文所用,“碳环(carbocycle或carbocyclic ring)”旨在包括具有指定碳数的任何稳定的单环、二环或三环,这些环中的任一种可以是饱和的、不饱和的或芳香族的。碳环包括环烷基和芳基。例如,C₃-C₁₄碳环旨在包括具有3、4、5、6、7、8、9、10、11、12、13或14个碳原子的单环、二环或三环。碳环的实例包括但不限于环丙基、环丁基、环丁烯基、环戊基、环戊烯基、环己基、环庚烯基、环庚基、环庚烯基、金刚烷基、环辛基、环辛烯基、环辛二烯基、苄基、苯基、萘基、茛满基、金刚烷基和四氢萘基。桥接环也包括在碳环的定义中,包括例如[3.3.0]二环辛烷、[4.3.0]二环壬烷、[4.4.0]二环癸烷和[2.2.2]二环辛烷。在一个或多个碳原子连接两个非相邻碳原子时出现桥接环。在一些实施例中,桥环是一个或两个碳原子。应注意,桥总是将单环转化为三环。在环发生桥接时,针对环所列举的取代基也可以存在于桥上。还包括稠合(例如,萘基、四氢萘基)环和螺环。

[0282] 如本文所用,“杂环”或“杂环基团”包括含有至少一个环杂原子(例如,N、O或S)的任何环结构(饱和、不饱和或芳香族)。杂环包括杂环烷基和杂芳基。杂环的实例包括但不限于吗啉、吡咯烷、四氢噻吩、哌啶、哌嗪、氧杂环丁烷、吡喃、四氢吡喃、氮杂环丁烷和四氢呋喃。

[0283] 杂环基团的实例包括但不限于吡啶基、吡啶基、苯并咪唑基、苯并呋喃基、苯并硫代呋喃基、苯并苯硫基、苯并噁唑基、苯并噁唑啉基、苯并噻唑基、苯并三唑基、苯并四唑基、苯并异噁唑基、苯并异噻唑基、苯并咪唑啉基、咪唑基、4aH-咪唑基、咪唑基、色烷基、色烯基、噌啉基、十氢喹啉基、2H,6H-1,5,2-二噻嗪基、二氢氟[2,3-b]四氢呋喃、呋喃基、呋喃基、咪唑烷基、咪唑啉基、咪唑基、1H-吡唑基、吡唑基、二氢吡唑基、吡唑基、3H-吡唑基、靛红酰基、异苯并呋喃基、异色烷基、异吡唑基、异吡唑啉基、异吡唑基、异喹啉基、异噻唑基、异噁唑基、亚甲基二氧苯基、吗啉基、萘基、八氢异喹啉基、噁二唑基、1,2,3-噁二唑基、1,2,4-噁二唑基、1,2,5-噁二唑基、1,3,4-噁二唑基、1,2,4-噁二唑基5(4H)-酮、噁唑烷基、噁唑基、羟吡唑基、嘧啶基、菲啶基、菲咯啉基、吩嗪基、吩噻嗪基、吩噻噻基、吩噻噻基(phenoxathinyl)、吩噻噻基、酞嗪基、哌嗪基、哌啶基、哌啶酮基、4-哌啶酮基、胡椒基、蝶啶基、嘌呤基、吡喃基、吡嗪基、吡唑烷基、吡唑啉基、吡唑基、哒嗪基、吡啶并噁唑、吡啶并咪唑、吡啶并噻唑、吡啶基(pyridinyl)、吡啶基(pyridyl)、嘧啶基、吡咯烷基、吡咯啉基、2H-吡咯基、吡咯基、喹啉基、喹啉基、4H-喹啉基、喹噻啉基、奎宁环基、四氢呋喃基、四氢异喹啉基、四氢喹啉基、四唑基、6H-1,2,5-噻二嗪基、1,2,3-噻二唑基、1,2,4-噻二唑基、1,2,5-

噻二唑基、1,3,4-噻二唑基、噻蒎基、噻唑基、噻吩基、噻吩并噻唑基、噻吩并噁唑基、噻吩并咪唑基、苯硫基、三嗪基、1,2,3-三唑基、1,2,4-三唑基、1,2,5-三唑基、1,3,4-三唑基和咕吨基。

[0284] 如本文所用,术语“被取代”意指指定原子上的任何一个或多个氢被来自所指示基团中的选择物置换,其条件是不超过指定原子的正常化合价,并且该取代产生稳定化合物。在取代基是氧代或酮基(即, $=O$)时,原子上的2个氢原子被置换。酮基取代基不存在于芳香族部分上。如本文所用,环双键是在两个相邻环原子之间形成的双键(例如, $C=C$ 、 $C=N$ 或 $N=N$)。“稳定化合物”和“稳定结构”意在指示足够稳健以便在从反应混合物中分离至可用纯度并且配制为有效治疗剂后继续存在的化合物。

[0285] 在与取代基的键显示与连接环中两个原子的键交叉时,则此类取代基可以与该环中的任何原子键合。在列出取代基但不指明此类取代基通过哪个原子与具有给定式的化合物的剩余部分键合时,则此类取代基可以经验这个式中的任何原子键合。取代基和/或变量的组合是允许的,但仅在此类组合产生稳定化合物时允许。

[0286] 在任何变量(例如, R_1)在化合物的任何成分或式中出现多于一次时,其在每次出现时的定义与其在其他每次出现时的定义无关。因此,例如,如果显示一基团被0-2个 R_1 部分取代,则该基团可以任选地被多至两个 R_1 部分取代,并且 R_1 在每次出现时独立地选自 R_1 的定义。另外,取代基和/或变量的组合是允许的,但仅在此类组合产生稳定化合物时才允许。

[0287] 术语“羟基(hydroxy或hydroxyl)”包括具有 $-OH$ 或 $-O^-$ 的基团。

[0288] 如本文所用,“卤代”或“卤素”是指氟、氯、溴和碘。术语“全卤化”通常是指其中所有氢原子都被卤素原子置换的部分。术语“卤代烷基”或“卤代烷氧基”是指被一个或多个卤素原子取代的烷基或烷氧基。

[0289] 术语“羰基”包括含有以双键连接到氧原子的碳的化合物和部分。含有羰基的部分的实例包括但不限于醛、酮、羧酸、酰胺、酯、酸酐等。

[0290] 术语“羧基”是指 $-COOH$ 或其 C_1 - C_6 烷基酯。

[0291] “酰基”包括含有酰基($R-C(O)-$)或羰基的部分。“取代的酰基”包括其中一个或多个氢原子被例如以下基团置换的酰基基团:烷基基团、炔基基团、卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、磷酸酯基、膦酸基、次膦酸基、氨基(包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基)、酰基氨基(包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基)、脞基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。

[0292] “芳酰基”包括具有与羰基基团结合的芳基或杂芳香族部分的部分。芳酰基基团的实例包括苯基羧基、萘基羧基等。

[0293] “烷氧基烷基”、“烷基氨基烷基”和“硫代烷氧基烷基”包括如上文所述的烷基基团,其中氧、氮或硫原子置换一个或多个烃主链碳原子。

[0294] 术语“烷氧基(alkoxy或alkoxyl)”包括共价连接到氧原子的取代的和未取代的烷基、烯基和炔基基团。烷氧基基团或烷氧基的实例包括但不限于甲氧基、乙氧基、异丙氧基、丙氧基、丁氧基和戊氧基。取代的烷氧基基团的实例包括卤化烷氧基基团。烷氧基基团可以

被诸如以下的基团取代：烯基、炔基、卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、磷酸酯基、膦酸基、次膦酸基、氨基（包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基）、酰基氨基（包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基）、脒基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。卤素取代的烷氧基基团的实例包括但不限于氟甲氧基、二氟甲氧基、三氟甲氧基、氯甲氧基、二氯甲氧基和三氯甲氧基。

[0295] 术语“醚”或“烷氧基”包括含有与两个碳原子或杂原子键合的氧的化合物或部分。例如，该术语包括“烷氧基烷基”，该烷氧基烷基是指共价键合到与烷基基团共价键合的氧原子的烷基、烯基或炔基基团。

[0296] 术语“酯”包括含有结合到与羰基基团的碳键合的氧原子的碳或杂原子的化合物或部分。术语“酯”包括烷氧基羧基基团，诸如甲氧基羰基、乙氧基羰基、丙氧基羰基、丁氧基羰基、戊氧基羰基等。

[0297] 术语“硫代烷基”包括含有与硫原子连接的烷基基团的化合物或部分。硫代烷基基团可以被诸如以下的基团取代：烷基、烯基、炔基、卤素、羟基、烷基羰氧基、芳基羰氧基、烷氧基羰氧基、芳氧基羰氧基、羧酸酯基、羧酸、烷基羰基、芳基羰基、烷氧基羰基、氨基羰基、烷基氨基羰基、二烷基氨基羰基、烷基硫代羰基、烷氧基、氨基（包括烷基氨基、二烷基氨基、芳基氨基、二芳基氨基和烷基芳基氨基）、酰基氨基（包括烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基）、脒基、亚氨基、巯基、烷硫基、芳硫基、硫代羧酸酯基、硫酸酯基、烷基亚磺酰基、磺酸基、氨磺酰基、磺酰胺基、硝基、三氟甲基、氰基、叠氮基、杂环基、烷基芳基、或者芳香族或杂芳香族部分。

[0298] 术语“硫代羰基”或“硫代羧基”包括含有以双键连接到硫原子的碳的化合物和部分。

[0299] 术语“硫醚”包括含有与两个碳原子或杂原子键合的硫原子的部分。硫醚的实例包括但不限于烷硫代烷基、烷硫代烯基、和烷硫代炔基。术语“烷硫代烷基”包括具有键合到与烷基基团键合的硫原子的烷基、烯基或炔基基团的部分。类似地，术语“烷硫代烯基”是指其中烷基、烯基或炔基基团键合到与烯基基团共价键合的硫原子的部分；并且“烷硫代炔基”是指其中烷基、烯基或炔基基团键合到与炔基基团共价键合的硫原子的部分。

[0300] 如本文所用，“胺”或“氨基”是指未取代的或取代的 $-NH_2$ 。“烷基氨基”包括其中 $-NH_2$ 的氮与至少一个烷基基团结合的化合物的基团。烷基氨基基团的实例包括苄基氨基、甲基氨基、乙基氨基、苯乙基氨基等。“二烷基氨基”包括其中 $-NH_2$ 的氮与至少两个另外的烷基基团结合的基团。二烷基氨基基团的实例包括但不限于二甲基氨基和二乙基氨基。“芳基氨基”和“二芳基氨基”包括其中氮分别与至少一个或两个芳基基团结合的基团。“氨基芳基”和“氨基芳氧基”是指被氨基取代的芳基和芳氧基。“烷基芳基氨基”、“烷基氨基芳基”或“芳基氨基烷基”是指与至少一个烷基基团和至少一个芳基基团结合的氨基基团。“烷氨基烷基”是指结合到还与烷基基团结合的氮原子的烷基、烯基或炔基基团。“酰基氨基”包括其中氮与酰基基团结合的基团。酰基氨基的实例包括但不限于烷基羰基氨基、芳基羰基氨基、氨基甲酰基和脲基基团。

[0301] 术语“酰胺”或“氨基羧基”包括含有与羰基或硫代羰基基团的碳结合的氮原子的化合物或部分。该术语包括“烷基氨基羧基”基团,该烷基氨基羧基基团包括结合到与羰基或硫代羰基基团的碳结合的氨基基团上的烷基、烯基或炔基基团。它还包括“芳基氨基羧基”基团,该芳基氨基羧基基团包括结合到与羰基或硫代羰基基团的碳结合的氨基基团上的芳基或杂芳基部分。术语“烷基氨基羧基”、“烯基氨基羧基”、“炔基氨基羧基”和“芳基氨基羧基”包括其中烷基、烯基、炔基和芳基部分分别与氮原子结合的部分,该氮原子又与羰基基团的碳结合。酰胺可以被诸如以下的取代基取代:直链烷基、支链烷基、环烷基、芳基、杂芳基或杂环。酰胺基团上的取代基可以进一步被取代。

[0302] 可以通过用氧化剂(例如,3-氯过氧苯甲酸(mCPBA)和/或过氧化氢)处理将含有氮的本披露化合物转化成N-氧化物以得到本披露的其他化合物。因此,当化合价和结构允许时,所有示出和要求保护的含氮化合物被认为包括如所示的化合物及其N-氧化物衍生物(可以被指定为 $N \rightarrow O$ 或 N^+-O^-)两者。此外,在其他情况下,本披露的化合物中的氮可以被转化成N-羟基或N-烷氧基化合物。例如,可以通过由诸如m-CPBA的氧化剂将母体胺氧化来制备N-羟基化合物。当化合价和结构允许时,所有示出和要求保护的含氮化合物被认为涵盖如所示的化合物及其N-羟基(即N-OH)和N-烷氧基(即N-OR,其中R是取代的或未取代的 C_1 - C_6 烷基、 C_1 - C_6 烯基、 C_1 - C_6 炔基、3-14元碳环或3-14元杂环)衍生物。

[0303] “异构现象”意指具有相同分子式但其原子的键合顺序或其原子的空间排列不同的化合物。原子空间排列不同的异构体称为“立体异构体”。并非彼此的镜像的立体异构体称为“非对映异构体”,并且是彼此的不可重叠的镜像的立体异构体称为“对映异构体”或有时称为光学异构体。含有等量的具有相反手性的单独对映异构形式的混合物称为“外消旋混合物”。

[0304] 键合到四个不相同取代基的碳原子称为“手性中心”。

[0305] “手性异构体”意指具有至少一个手性中心的化合物。具有多于一个手性中心的化合物可以作为单独非对映异构体或作为非对映异构体的混合物存在,称为“非对映异构混合物”。在存在一个手性中心时,立体异构体可以通过该手性中心的绝对构型(R或S)来表征。绝对构型是指与手性中心附接的取代基的空间排列。所考虑的与手性中心附接的取代基按照卡恩-英戈尔德-普雷洛格(Cahn, Ingold和Prelog)的顺序规则排序。(Cahn等人, Angew. Chem. Inter. Edit. [应用化学国际版]1966, 5, 385; 勘误表511; Cahn等人, Angew. Chem. [应用化学]1966, 78, 413; Cahn和Ingold, J. Chem. Soc. [化学学会会刊]1951(伦敦), 612; Cahn等人, Experientia [实验]1956, 12, 81; Cahn, J. Chem. Educ. [化学教育期刊]1964, 41, 116)。

[0306] “几何异构体”意指因围绕双键或环烷基接头(例如,1,3-环丁基)的旋转受阻而存在的非对映异构体。这些构型的名称通过前缀顺式和反式或Z和E来区分,这些前缀指示根据卡恩-英戈尔德-普雷洛格规则,基团位于分子中双键的同侧或对侧。

[0307] 应理解,本文提供的小分子EZH2抑制剂可以被描绘为不同手性异构体或几何异构体。还应理解,在化合物具有手性异构体或几何异构体形式时,所有异构体形式均旨在被包括于本披露的范围内,并且化合物的命名不排除任何异构体形式。

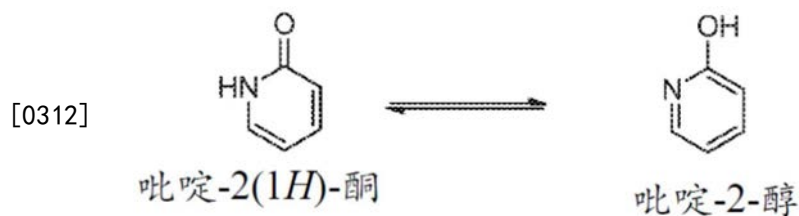
[0308] 此外,本披露中讨论的结构和其他化合物包括其所有阻转异构体。“阻转异构体”是其中两个异构体的原子空间排列不同的立体异构体类型。阻转异构体是因大基团围绕中

心键的旋转受阻引起的旋转受限而存在。此类阻转异构体典型地作为混合物存在,但由于色谱技术的最新发展,已可能在所选情况下分离两种阻转异构体的混合物。

[0309] “互变异构体”是平衡地存在并且易于从一种异构体形式转化成另一种形式的两种或更多种结构异构体之一。这种转化导致氢原子的形式迁移,伴随着相邻共轭双键的转化。互变异构体在溶液中作为互变异构组的混合物存在。在可能发生互变异构的溶液中,互变异构体将达到化学平衡。互变异构体的确切比率取决于若干因素,包括温度、溶剂和pH。可通过互变异构作用相互转化的互变异构体的概念被称为互变异构现象。

[0310] 在可能的多种类型的互变异构现象中,通常观察到两种。在酮-烯醇互变异构现象中,电子和氢原子发生同时转移。环-链互变异构现象由于糖链分子中的醛基基团(-CHO)与同一分子中的一个羟基基团(-OH)反应而赋予其如葡萄糖所展现的环状(环形状)形式而发生。

[0311] 常见的互变异构对是:杂环中(例如,在核碱基中,诸如鸟嘌呤、胸腺嘧啶和胞嘧啶)的酮-烯醇、酰胺-腈、内酰胺-内酰亚胺、酰胺-亚胺酸互变异构现象、亚胺-烯胺和烯胺-烯胺。如下文所示的,酮-烯醇平衡的实例在吡啶-2(1H)-酮与相应的吡啶-2-醇之间。



[0313] 应理解,本披露的化合物可以描绘为不同互变异构体。还应理解,在化合物具有互变异构形式时,所有互变异构形式都旨在被包括在本披露的范围内,并且化合物的命名不排除任何互变异构体形式。

[0314] 本文所披露的具有式(I)-(VIa)的EZH2抑制剂包括化合物本身,以及其盐和其溶剂化物(若适用)。例如,盐可在阴离子与芳基或杂芳基取代的苯化合物上的带正电基团(例如,氨基)之间形成。合适的阴离子包括氯离子、溴离子、碘离子、硫酸根、硫酸氢根、氨基磺酸根、硝酸根、磷酸根、柠檬酸根、甲磺酸根、三氟乙酸根、谷氨酸根、葡萄糖醛酸根、戊二酸根、苹果酸根、马来酸根、琥珀酸根、富马酸根、酒石酸根、甲苯磺酸根、水杨酸根、乳酸根、萘磺酸根、和乙酸根(例如,三氟乙酸根)。术语“药学上可接受的阴离子”是指适于形成药学上可接受的盐的阴离子。同样,盐也可在阳离子与芳基或杂芳基取代的苯化合物上的带负电基团(例如,羧酸根)之间形成。合适的阳离子包括钠离子、钾离子、镁离子、钙离子、和铵阳离子(诸如四甲基铵离子)。芳基-或杂芳基-取代的苯化合物还包括那些含有季氮原子的盐。在盐形式中,应理解,化合物与盐的阳离子或阴离子的比率可以为1:1,或不同于1:1的任何比率,例如3:1、2:1、1:2或1:3。

[0315] 另外,本披露的EZH2抑制性化合物(例如,化合物的盐)可以以水合或非水合(无水)形式或作为与其他溶剂分子的溶剂化物而存在。水合物的非限制性实例包括一水合物、二水合物等。溶剂化物的非限制性实例包括乙醇溶剂化物、丙酮溶剂化物等。

[0316] “溶剂化物”意指含有化学计量量或非化学计量量的溶剂的溶剂加成形式。一些化合物倾向于将固定摩尔比的溶剂分子捕获在结晶固态中,从而形成溶剂化物。如果溶剂是水,则所形成的溶剂化物是水合物;并且如果溶剂是醇,则所形成的溶剂化物是醇化物。水

合物通过将一分子或多分子水与一分子物质组合而形成,其中水保持其分子状态为H₂O。

[0317] 如本文所用,术语“类似物”是指在结构上与另一化合物类似、但在组成上稍微不同的化合物(如由不同元素的原子置换一个原子或在特定官能团存在下、或由另一官能团置换一个官能团)。因此,类似物是与参考化合物在功能和外观上,但不在结构或来源上类似或相当的化合物。

[0318] 如本文所用,术语“衍生物”是指具有共同的核心结构,并且被如本文所述的不同基团取代的化合物。例如,由式(I)表示的所有化合物均是芳基或杂芳基取代的苯化合物,并且具有式(I)作为共同核心。

[0319] 本披露的一些实施例涵盖本发明的EZH2抑制性化合物中存在的原子的一些或所有同位素。同位素包括原子数相同但质量数不同的那些原子。借助一般实例且不受限制,氢同位素包括氕和氘,并且碳同位素包括C-13和C-14。

[0320] 在本披露的某些方面,当EZH2的抑制剂抑制突变型EZH2的组蛋白甲基转移酶活性比它抑制野生型EZH2的组蛋白甲基转移酶活性更有效时,EZH2的抑制剂“选择性地抑制”突变型EZH2的组蛋白甲基转移酶活性。例如,在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少40%。在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少50%。在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少60%。在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少70%。在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少80%。在一些实施例中,选择性抑制剂具有的对于突变型EZH2的IC₅₀比对于野生型EZH2的IC₅₀低至少90%。

[0321] 在一些实施例中,突变型EZH2的选择性抑制剂对野生型EZH2基本上没有抑制作用。

[0322] 在本披露的某些方面,抑制剂(例如本文所披露的化合物)抑制H3-K27me₂向H3-K27me₃的转化。在一些实施例中,抑制剂据说是抑制H3-K27的三甲基化。因为H3-K27me₁向H3-K27me₂的转化先于H3-K27me₂向H3-K27me₃的转化,H3-K27me₁向H3-K27me₂的转化天然地也抑制H3-K27me₂向H3-K27me₃的转化,即,它抑制H3-K27的三甲基化。也有可能抑制H3-K27me₂向H3-K27me₃的转化而不会抑制H3-K27me₁向H3-K27me₂的转化。这种类型的抑制尽管不会抑制H3-K27的二甲基化,但也将导致对H3-K27的三甲基化的抑制。

[0323] 在一些实施例中,抑制剂(例如本文所披露的化合物)抑制H3-K27me₁向H3-K27me₂的转化和H3-K27me₂向H3-K27me₃的转化。这种抑制剂可以直接单独抑制H3-K27me₁向H3-K27me₂的转化。可替换地,这种抑制剂可以直接抑制H3-K27me₁向H3-K27me₂的转化和H3-K27me₂向H3-K27me₃的转化。

[0324] 在本披露的某些方面,EZH2抑制剂(例如本文所披露的化合物)抑制组蛋白甲基转移酶活性。组蛋白甲基转移酶活性的抑制可以使用任何适合的方法来检测。该抑制可以被测量,例如,依据组蛋白甲基转移酶的活性比率或者依照组蛋白甲基转移酶活性的产物来测量。

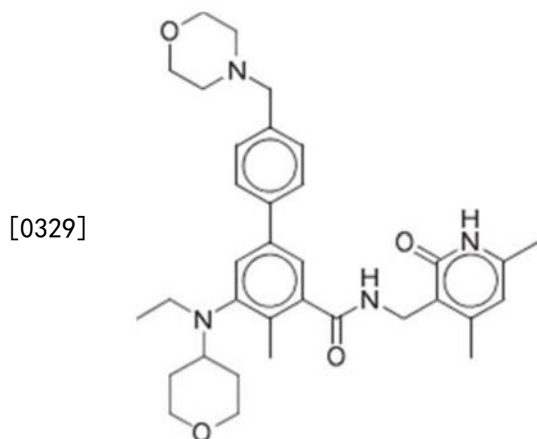
[0325] 在一些实施例中,提供了策略、治疗形式、方法、组合和组合物,其特征在于对EZH2活性的可测量抑制,例如相比于合适对照的可测量EZH2抑制。在一些实施例中,EZH2抑制是

相比于合适对照,例如未经治疗的对照细胞、组织或受试者中观察到或预期的EZH2活性的至少10%抑制。在一些实施例中,存在EZH2抑制剂时的EZH2酶活性的比率小于或等于不存在EZH2抑制剂时的相应酶活性的90%。在一些实施例中,存在EZH2抑制剂时的EZH2抑制是相比于合适对照,例如不存在抑制剂时的活性的至少20%、25%、30%、40%、50%、60%、70%、75%、80%、90%或95%抑制。在一些实施例中,抑制是相比于合适对照的至少99%抑制。即,存在抑制剂时的酶活性的比率小于或等于不存在抑制剂时的相应活性的1%。

[0326] 在一些实施例中,将本文提供的治疗剂(例如EZH2抑制剂)和(在适用情况下)任何另外的治疗剂(例如免疫检查点抑制剂)提供在适于给予至人受试者的药物配制品中。在使用多于一种治疗剂的实施例中,可以将每种治疗剂分别配制成药物配制品,并且独立地,例如依序地给予至受试者。在一些此类实施例中,不同的药物组合物可以经由相同的途径,例如肠胃外途径给予,或者可替代地,经由不同的途径,例如肠内和肠胃外途径给予。例如,在本文提供的组合治疗形式的一些实施例中,可以将EZH2抑制剂配制成用于口服给予并且将另外的治疗剂,例如免疫检查点抑制剂配制成用于肠胃外给予。

[0327] 合适的包含EZH2抑制剂的药物组合物先前已得到描述,并且包括,例如但不限于以下中列出的那些:US 8,410,088、US 8,765,732、US 9,090,562、US 8,598,167、US 8,962,620、US-2015/0065483、US 9,206,157、US 9,006,242、US 9,089,575、US 2015-0352119、WO 2014/062733、US-2015/0065503、WO 2015/057859、US 8,536,179、WO 2011/140324、公布为WO/2014/124418的PCT/US2014/015706、公布为WO/2013/120104的PCT/US 2013/025639、以及公布为US 2015/0368229的US 14/839,273,每篇专利的全部内容通过援引并入本文。基于本披露和本领域的一般知识,另外的合适药物组合物对于本领域技术人员是清楚的。

[0328] 本披露还提供了药物组合物和组合,这些药物组合物和组合包含具有式(I)-(VIa)的化合物或其药学上可接受的盐以及本文所披露的一种或多种其他治疗剂,例如一种或多种免疫检查点抑制剂,它们以治疗或预防如本文所述的疾病或病状的剂量与药学上合适的载体或一种或多种赋形剂混合。在一个方面,本披露还提供了药物组合物,这些药物组合物包含表I的任何化合物或其药学上可接受的盐和一种或多种治疗剂,它们以治疗或预防如本文所述的疾病或病状的剂量与药学上合适的载体或一种或多种赋形剂混合。在另一个方面,本披露还提供了药物组合物,这些药物组合物包含化合物44



[0330] 或其药学上可接受的盐和一种或多种治疗剂,它们以治疗或预防如本文所述的疾

病或病状的剂量与药学上合适的载体或一种或多种赋形剂混合。本披露的药物组合物也可以与其他治疗剂或治疗形式同时地、依序地、或交替地组合给予。

[0331] 本披露的组合物的混合物或组合也可以作为简单的混合物或以合适的配制药物组合物给予至患者。例如，本披露的一个方面涉及一种药物组合物或组合，该药物组合物或组合包含治疗有效剂量的具有式 (I) – (VIa) 的EZH2抑制剂或其药学上可接受的盐、水合物、对映体或立体异构体；一种或多种其他治疗剂、和药学上可接受的稀释剂或载体。

[0332] “药物组合物”是呈适于给予至受试者的形式的含有本披露化合物的配制品。本文所述的具有式 (I) – (VIa) 的化合物和 (在适用情况下) 一种或多种其他治疗剂各自可以单独配制或以活性成分的任何组合配制成药物组合物。因此，可以基于每种药物组合物的剂型适当地选择一种或多种给予途径。可替代地，本文所述的具有式 (I) – (VIa) 的化合物和一种或多种其他治疗剂可以配制成一种药物组合物。

[0333] 在一些实施例中，药物组合物呈散装形式或呈单位剂型。单位剂型是多种形式中的任一种，包括例如胶囊、IV袋、片剂、气溶胶吸入器上的单泵或小瓶。单位剂量组合物中的活性成分 (例如，所披露的化合物或其盐、水合物、溶剂化物或异构体的配制品) 的量是有效量，并且根据所涉及的具体治疗而变化。本领域技术人员应认识到，有时有必要依据患者的年龄和病状对剂量进行常规改变。剂量还将取决于给予途径。考虑了多种途径，包括口服、经肺、经直肠、肠胃外、经皮、皮下、静脉内、肌内、腹膜内、吸入、经颊、舌下、胸膜内、鞘内、鼻内等。用于局部或经皮给予本披露的化合物的剂型包括散剂、喷雾剂、软膏剂、糊剂、乳膏剂、洗剂、凝胶剂、溶液、贴剂和吸入剂。在一些实施例中，将活性化合物在无菌条件下与药学上可接受的载体以及与所需要的任何防腐剂、缓冲剂或推进剂混合。

[0334] 如本文所用，短语“药学上可接受的”是指在合理医学判断的范围内适于与人和动物的组织接触而无过度毒性、刺激、过敏应答或其他问题或并发症，与合理益处/风险比相称的那些化合物、阴离子、阳离子、材料、组合物、载体和/或剂型。

[0335] 术语“药学上可接受的赋形剂”是指可用于制备药物组合物、通常是安全、无毒的并且不是生物学上或其他方面不希望的赋形剂，并且包括对于兽用连同人类药用是可接受的赋形剂。如本说明书和权利要求书中所用的“药学上可接受的赋形剂”包括一种和超过一种的这种赋形剂。

[0336] 本披露的药物组合物被配制成与其预期的给予途径相容。给予途径的实例包括肠胃外，例如静脉内、皮内、皮下、口服 (例如，吸入)、经皮 (局部) 和经粘膜给予。用于肠胃外、皮内或皮下施用的溶液或混悬液可以包含以下组分：无菌稀释剂，诸如注射用水、盐水溶液、固定油、聚乙二醇、甘油、丙二醇或其他合成溶剂；抗菌剂，诸如苄基醇或对羟苯甲酸甲酯；抗氧化剂，诸如抗坏血酸或亚硫酸氢钠；螯合剂，诸如乙二胺四乙酸；缓冲剂，诸如乙酸盐、柠檬酸盐或磷酸盐，以及用于调节张力的试剂，诸如氯化钠或右旋糖。可以用酸或碱 (诸如盐酸或氢氧化钠) 调节pH。肠胃外制剂可以封装在由玻璃或塑料制成的安瓿、一次性注射器或多剂量小瓶中。

[0337] 本披露的组合物，例如包含EZH2抑制剂的配制品，可以以许多目前用于化学治疗剂治疗的熟知方法给予至受试者。例如，对于治疗癌症，可以将包含EZH2抑制剂的配制品直接注射进肿瘤、注射进血流或体腔或口服或使用贴剂应用于皮肤。针对EZH2抑制剂和 (在适用情况下) 任何另外的治疗剂选择的剂量应足以构成有效治疗，但不能高到引起不可接受

的副作用。疾病病状的状态(例如,癌症、癌前病变等)和患者的健康状况在治疗期间和治疗后的合理期限内应优选地密切监测。

[0338] 如本文所用,术语“治疗有效量”是指药剂的用于治疗、改善或预防所鉴定疾病或病状、或者展现出所期望的临床效果,例如可检测的治疗或抑制作用的量。本文提供了EZH2抑制剂和一些示例性另外的治疗剂的示例性、非限制性有效量和有效剂量范围。在一些实施例中,所期望的临床效果可以例如通过本领域已知的任何合适的测定方法直接检测。在一些实施例中,所期望的临床效果可以通过代用测量法来测量。例如,在一些实施例中,可以监测表观遗传学抑制的SMARCA2和/或SMARCA4表达的再激活以确定EZH2抑制剂的合适的治疗有效量。用于受试者的精确有效量将取决于受试者的体重、身材、和健康状况;病状的性质和程度;以及选择用于给予的治疗剂或治疗剂组合。针对给定情况的治疗有效量可以通过临床医师的技术和判断范围内的常规实验来确定。在一个优选方面,待治疗的疾病或病状是癌症。在另一个方面,待治疗的疾病或病状是细胞增殖性病症。

[0339] 在某些实施例中,相比单独使用每种药剂的单一疗法,当组合使用每种药剂时每种药剂的治疗有效量会更低。这种更低的治疗有效量能负担的治疗方案的毒性更低。

[0340] 对于本文所述的许多化合物,例如各种EZH2抑制剂和各种另外的治疗剂,已报道治疗有效量或有效剂量范围。在一些实施例中,有效量最初可以在细胞培养测定(例如,肿瘤细胞的细胞培养测定)中或者在动物模型(通常是大鼠、小鼠、兔、狗或猪)中估算。动物模型还可以用于确定给予的适当浓度范围和途径。然后此类信息可以用于确定用于在人中给予的可用剂量和途径。治疗/预防功效和毒性可以通过标准制药程序在细胞培养物或实验动物中来确定,例如ED₅₀(在50%群体中治疗有效的剂量)和LD₅₀(对50%群体具有致死性的剂量)。毒性与治疗效果之间的剂量比是治疗指数,并且它可以表达为比率LD₅₀/ED₅₀。展现大的治疗指数的药物组合物是优选的。剂量可以在此范围内变化,这取决于所采用的剂型、患者的敏感性和给予途径。

[0341] 调整剂量和给予以提供一种或多种活性药剂的足够水平或维持所期望的效果。可以考虑的因素包括疾病状态的严重性,受试者的一般健康状况,受试者的年龄、体重和性别,饮食,给予时间和频率,一种或多种药物组合,反应灵敏度,以及对疗法的耐受性/应答。长效药物组合物可以每3至4天、每周、或每两周一次给予,这取决于特定配制品的半衰期和清除率。

[0342] 含有本披露活性化合物的药物组合物可以以通常已知的方式制备,例如借助常规混合、溶解、制粒、制糖衣、磨粉、乳化、胶囊化、包埋或冻干工艺的方式。药物组合物可以以常规方式使用一种或多种药学上可接受的载体(包括赋形剂和/或助剂)来配制,这些载体促进将活性化合物加工成可以在药学上使用的制剂。当然,适当的配制品取决于所选的给予途径。

[0343] 适于可注射用途的药物组合物包含无菌水溶液(在水溶性的情况下)或分散体以及用于临时制备无菌可注射溶液或分散液的无菌粉末。对于静脉内给予,适合的载体包括生理盐水、抑菌水、Cremophor EL™(巴斯夫,新泽西州派西派尼市(Parsippany, N.J.))或磷酸盐缓冲盐水(PBS)。在所有情况下,组合物必须是无菌的并且应该在存在可易注射性的程度上是流动的。它在制造和存储的条件下必须是稳定的并且必须抗微生物(诸如细菌和真菌)的污染作用而保存。载体可以是含有例如以下物质的溶剂或分散介质:水、乙醇、多元

醇(例如,甘油、丙二醇和液体聚乙二醇等)及其合适的混合物。恰当的流动性可以例如通过使用诸如卵磷脂的包衣来维持,在分散液的情况下通过维持所需粒径来维持,以及通过使用表面活性剂来维持。防止微生物的作用可以通过各种抗菌剂和抗真菌剂(例如对羟苯甲酸酯、三氯叔丁醇、苯酚、抗坏血酸、硫柳汞等)来实现。在许多情况下,将优选的是将等渗剂(例如糖、多元醇(诸如甘露醇和山梨醇)和氯化钠)包含在组合物中。可以通过在组合物中包含延迟吸收的试剂(例如,单硬脂酸铝和明胶)来实现可注射组合物的延长吸收。

[0344] 无菌可注射溶液可以通过将所需量的活性化合物根据需要与以上列举的成分中的一种或其组合一起掺入适当溶剂中,随后过滤灭菌来制备。通常,通过将活性化合物掺入无菌媒介物中来制备分散液,该无菌媒介物含有基础分散介质和来自以上列举的那些的其他所需成分。在用于制备无菌可注射溶液的无菌粉末的情况下,制备方法为真空干燥和冷冻干燥,这些干燥产生了活性成分加来自其先前经无菌过滤溶液的任何另外的所希望成分的粉末。

[0345] 口服组合物通常包含惰性稀释剂或可食用的药学上可接受的载体。它们可以封装在明胶胶囊中或压成片剂。出于口服治疗性给予的目的,可以将活性化合物与赋形剂一起掺入并且以片剂、锭剂或胶囊剂的形式使用。口服组合物也可以使用用作漱口水的流体载体制备,其中将流体载体中的化合物口服施用并漱口(swished)并且吐出或咽下。可以包含药学上相容的粘合剂和/或佐剂材料作为组合物的一部分。片剂、丸剂、胶囊剂、锭剂等可以含有以下成分中的任一种或具有类似性质的化合物:粘合剂,诸如微晶纤维素、黄蓍胶或明胶;赋形剂,诸如淀粉或乳糖;崩解剂,诸如海藻酸、Primogel或玉米淀粉;润滑剂,诸如硬脂酸镁或Sterotes;助流剂,诸如胶体二氧化硅;甜味剂,诸如蔗糖或糖精;或调味剂,诸如薄荷、水杨酸甲酯或橙味剂。

[0346] 对于通过吸入给予,化合物从加压容器或分配器(其含有合适的推进剂,例如气体(诸如二氧化碳))或喷雾器以气溶胶喷雾形式递送。

[0347] 全身给予也可以通过经粘膜或经皮的方式。对于经粘膜或经皮给予,在配制品中使用适于待渗透的障碍的渗透剂。此类渗透剂是本领域通常已知的,并且对于经粘膜给予,包括例如洗涤剂、胆汁盐、和夫西地酸衍生物。经粘膜给予可以通过使用鼻喷雾剂或栓剂来实现。对于经皮给予,将活性化合物配制成软膏剂、油膏剂、凝胶剂或乳膏剂,如本领域通常已知的。

[0348] 可以将活性化合物与将保护该化合物免于从体内快速消除的药学上可接受的载体一起制备,诸如控制释放配制品,包括植入物和微囊化的递送系统。可以使用可生物降解、生物相容的聚合物,诸如乙烯乙酸乙烯酯、聚酞类、聚乙醇酸、胶原、聚原酸酯类、和聚乳酸。用于制备此类配制品的方法对本领域技术人员而言是清楚的。材料也可以从阿尔扎公司(Alza Corporation)和新星制药有限公司(Nova Pharmaceuticals, Inc.)商购获得。脂质体混悬液(包括靶向受感染细胞、具有针对病毒抗原的单克隆抗体的脂质体)也可以用作药学上可接受的载体。这些可以根据本领域技术人员已知的方法来制备,例如,如在美国专利号4,522,811中所描述的。

[0349] 特别有利的是以单位剂型配制口服或肠胃外组合物以便于给予和剂量统一。如本文所用的单位剂型是指适合作为单一剂量用于待治疗受试者的物理离散单位;每单位含有经计算以与所需的药物载体的联合产生所期望的治疗效果的预定量的活性化合物。本披露

的单位剂型的规格由活性化合物的独特特征和待实现的特定治疗效果决定并且直接取决于它们。

[0350] 在治疗性应用的一些实施例中,本文提供的治疗剂的剂量根据使用的一种或多种特定药剂,接受患者的年龄、体重和临床病状和给予疗法的临床医师或执业医师的经验和判断,以及影响所选剂量的其他因素而变化。通常,一种或多种活性成分的剂量应足以导致肿瘤的生长减慢、并且优选消退,并且还优选地使癌症完全消退。在一些实施例中,剂量的范围可以是每天约0.01mg/kg至每天约5000mg/kg。在优选方面,剂量的范围可以是每天约1mg/kg至每天约1000mg/kg。在一个方面,剂量的范围将为约0.1mg/天至约50g/天;约0.1mg/天至约25g/天;约0.1mg/天至约10g/天;约0.1mg至约3g/天;或约0.1mg至约1g/天,呈单一、分开或连续剂量(该剂量可以针对患者的以kg计体重、以m²计体表面积、和以岁计年龄来调整)。在本文别处提供了另外的合适剂量。例如,患者肿瘤的消退可以参考肿瘤的直径来测量。肿瘤直径的减小指示消退。肿瘤在停止治疗后无法重新出现也指示消退。如本文所用,术语“剂量有效的方式”是指活性化合物在受试者或细胞中产生所期望的生物效应的量。

[0351] 如本文所用,“药学上可接受的盐”是指本披露的化合物,例如本文所述的小分子EZH2抑制剂的衍生物,其中母体化合物通过制备其酸或碱的盐来修饰。例如本文提供的EZH2抑制剂的药学上可接受的盐的实例包括但不限于碱性残基(诸如胺)的矿物盐或有机酸盐、酸性残基(诸如羧酸)的碱盐或有机盐等。药学上可接受的盐包括例如由无毒的无机酸或有机酸形成的母体化合物的常规无毒盐或季铵盐。例如,此类常规无毒盐包括但不限于衍生于无机酸和有机酸的那些,这些无机酸和有机酸选自2-乙酰氧基苯甲酸、2-羟基乙烷磺酸、乙酸、抗坏血酸、苯磺酸、苯甲酸、重碳酸、碳酸、柠檬酸、依地酸、乙烷二磺酸、1,2-乙烷磺酸、富马酸、葡庚糖酸、葡糖酸、谷氨酸、乙醇酸、乙二醇阿散酸、己基间苯二酚酸(hexylresorcinic)、水巴米克酸(hydrabamic)、氢溴酸、盐酸、氢碘酸、羟基马来酸、羟萘甲酸、羟乙基磺酸、乳酸、乳糖酸、月桂基磺酸、马来酸、苹果酸、扁桃酸、甲烷磺酸、萘磺酸、硝酸、草酸、扑酸、泛酸、苯乙酸、磷酸、聚半乳糖醛酸、丙酸、水杨酸、硬脂酸、亚乙酸(subacetic)、琥珀酸、氨基磺酸、对氨基苯磺酸、硫酸、鞣酸、酒石酸、甲苯磺酸,以及常见的氨基酸,例如甘氨酸、丙氨酸、苯丙氨酸、精氨酸等。

[0352] 药学上可接受的盐的其他实例包括己酸、环戊烷丙酸、丙酮酸、丙二酸、3-(4-羟基苯甲酰基)苯甲酸、肉桂酸、4-氯苯磺酸、2-萘磺酸、4-甲苯磺酸、樟脑磺酸、4-甲基二环-[2.2.2]-辛-2-烯-1-甲酸、3-苯基丙酸、三甲基乙酸、叔丁基乙酸、粘糠酸等。本披露还涵盖在以下发生时形成的盐:母体化合物中存在的酸性质子被金属离子(例如,碱金属离子、碱土金属离子或铝离子)置换;或与有机碱(诸如乙醇胺、二乙醇胺、三乙醇胺、氨丁三醇、N-甲基葡糖胺等)配位。

[0353] 应理解,对药学上可接受的盐的所有提及包括相同盐的溶剂加成形式(溶剂化物)。

[0354] 本披露的组合物还可以被制备为酯,例如,药学上可接受的酯。例如,化合物中的羧酸官能团可以转化成其相应的酯,例如甲酯、乙酯或其他酯。而且,化合物中的醇基可以转化成其相应的酯,例如乙酸酯、丙酸酯或其他酯。组合物或其药学上可接受的盐或溶剂化物以口服、经鼻、经皮、经肺、吸入、经颊、舌下、腹腔内、皮下、肌内、静脉内、直肠、胸膜内、鞘

内和肠胃外的形式给予。在一些实施例中,化合物是口服给予的。本领域技术人员应认识到某些给予途径的优点。

[0355] 利用化合物的剂量方案是根据多种因素来选择的,包括患者的类型、物种、年龄、体重、性别和医学病状;待治疗病状的严重性;给予途径;患者的肾功能和肝功能;以及所采用的具体化合物或其盐。普通技术的医师或兽医可以容易地确定并且开出预防、对抗病状、或阻止病状的进展所需的药物的有效量。

[0356] 用于配制和给予本披露所披露的化合物的技术可见于remington:the Science and Practice of Pharmacy[雷明顿:药物科学与实践],第19版,Mack Publishing Co.[马克出版有限公司],Easton[伊斯顿],PA[宾夕法尼亚州](1995)中。在一些实施例中,本文所述的化合物及其药学上可接受的盐与药学上可接受的载体或稀释剂组合用于药物制剂中。合适的药学上可接受的载体包括惰性固体填充剂或稀释剂和无菌水性或有机溶液。化合物将以足以提供在本文所述的范围内的所希望剂量的量存在于此类药物组合物中。

[0357] 除非另外指明,否则本文所用的所有百分比和比率都是按重量计。根据不同实例,本披露的其他特征和优点是清楚的。所提供的实例说明了在实践本披露中有用的不同组分和方法。实例并不限制要求保护的披露。基于本披露,技术人员可以鉴定并采用可用于实践本披露的其他组分和方法。

[0358] 在一些实施例中,“有需要的受试者”是患有EZH2介导的蛋白质甲基化起作用的病症的受试者,或者相对于大群体具有增加的发展这种病症的风险的受试者。在一些实施例中,有需要的受试者患有细胞增生性疾病,例如癌症。在一些实施例中,受试者患有特征在于SMARCA2和/或SMARCA4功能丧失的癌症。在一些实施例中,受试者患有特征在于SMARCA2/SMARCA4双重功能丧失,其中SMARCA2功能丧失由表观遗传损伤介导的癌症。在一些实施例中,受试者患有其中免疫系统逃避,例如癌细胞经由免疫检查点信号传导的免疫系统逃避也起作用的病症。“受试者”包括哺乳动物。哺乳动物可以是例如任何哺乳动物,例如人、灵长类、鸟、小鼠、大鼠、家禽、狗、猫、牛、马、山羊、骆驼、绵羊或猪。优选地,哺乳动物是人。

[0359] 在一些实施例中,受试者是已经被诊断为癌症或癌前病状、具有癌症或癌前病状的症状、或具有发展癌症或癌前病状的风险的人受试者。在一些实施例中,受试者表达突变型EZH2蛋白。例如,突变型EZH2包含一个或多个突变,其中该突变是取代、点突变、无义突变、错义突变、缺失、或插入或任何其他本文所述的EZH2突变。在一些实施例中,受试者表达野生型EZH2蛋白。

[0360] 有需要的受试者可以患有难治性或抗性癌症。“难治性或抗性癌症”意指对治疗,例如用单一疗法(例如仅用化学治疗剂的单一疗法)的治疗没有应答的癌症。在一些实施例中,癌症可以是用于该特定类型的癌症的标准护理治疗难治性或抗性的。癌症可以在治疗开始时就有抗性,或在治疗期间变得有抗性。在一些实施例中,有需要的受试者的癌症在最近的疗法缓解后再发。在一些实施例中,有需要的受试者接受了所有已知用于癌症治疗的有效疗法但都失败了。在一些实施例中,有需要的受试者接受了至少一种先前疗法。在某些实施例中,先前疗法是单一疗法。在某些实施例中,先前疗法是组合疗法。

[0361] 在一些实施例中,有需要的受试者可以由于以前的疗法而具有继发性癌症。“继发性癌症”意指由于或因为以前的致癌疗法(诸如化学疗法)发生的癌症。

[0362] 受试者也可以对EZH2组蛋白甲基转移酶抑制剂或任何其他治疗剂展现出抗性。

[0363] 如本文所用,术语“应答性”是与术语“应答”、“敏感”、和“敏感性”可互换,并且它意指当给予本披露的组合物时受试者表现出治疗性应答,例如,受试者的肿瘤细胞或肿瘤组织经受细胞凋亡和/或坏死,和/或显示降低的生长、分裂或增生。该术语也意指当给予本披露的组合物时,受试者相对于大的群体会或具有更高的概率表现出治疗性应答,例如,受试者的肿瘤细胞或肿瘤组织经受细胞凋亡和/或坏死和/或显示降低的生长、分裂或增生。

[0364] 术语“样品”是指任何源自受试者的生物样品,包括但不限于细胞、组织样品、体液(包括但不限于粘液、血液、血浆、血清、尿液、唾液、和精液)、肿瘤细胞和肿瘤组织。优选地,样品选自骨髓、外周血细胞、血液、血浆和血清。样品可以由正在治疗或测试的受试者提供。可替代地,样品可以根据本领域的常规实践由医师获得。

[0365] 如本文所用,“正常细胞”是不能被分类为“细胞增生性病症”的一部分的细胞。正常细胞缺乏可导致不希望的病状或疾病发展的非常规生长或异常生长、或两者。优选地,正常细胞具有正常功能的细胞周期检查点控制机制。

[0366] 如本文所用,“接触细胞”是指一种状态,其中化合物或其他物质组合物直接接触细胞,或足够接近以在细胞内诱导所期望的生物效应。

[0367] 如本文所用,“治疗(treating或treat)”描述了出于对抗疾病、病状或病症的目的而对患者进行的管理和护理,并且包括给予EZH2抑制剂和/或免疫检查点抑制剂,以减轻疾病、病状或病症的症状或并发症,或者消除疾病、病状或病症。

[0368] 从以下实例将更全面地理解本文所披露的技术的一些实施例、优点、特征和用途。这些实例旨在说明本披露的一些益处并描述特定实施例,但是并不旨在例示本披露的全部范围,并且因此不限制本披露的范围。

[0369] 实例

[0370] 在体外癌细胞系中对EZH2抑制的敏感性

[0371] 在各种肺癌细胞系中测定SWI/SNF复合物蛋白的状态。所有测试的肺癌细胞系中约1/3展现出SWI/SNF成员蛋白畸变。下表2A示出鉴定为携带一个或多个SWI-SNF改变的31种肺癌细胞系中的SMARCA2和SMARCA4蛋白状态。深灰色指示功能丧失,浅灰色指示正常功能。如表2A中所示,31种SWI/SNF功能丧失肺癌细胞系中有10种展现出单一SMARCA4丧失,而列出在表的顶部的31种系中的8种展现出双重SMARCA2/SMARCA4丧失。

[0372] 表2A:

[0373]

亚型	细胞系	SMARCA2	SMARCA4	ARID1A/B	KRAS
腺癌	NCI-H23				
大细胞	NCIH1581				
腺癌	A427				
腺癌	NCI-H522				
大细胞	NCI-H661				
NSCLC	NCI-H1299				
腺癌	NCIH1693				
腺癌	NCIH1703				
腺癌	NCIH2023				
腺癌	NCIH1568				
腺癌	NCIH838				
腺癌	NCIH1793				
鳞状	HCC15				
腺癌	NCIH2030				
腺癌	A549				
腺癌	NCIH1573				
腺癌	NCIH1944				
腺癌	NCIH2085				
腺癌	NCIH2347				
腺癌	NCIH2291				
腺癌	NCIH1648				
腺癌	NCIH1734				
未分化的	Calu-6				
腺癌	NCIH1650				
大细胞	NCI-H460				
鳞状	NCIH1869				
大细胞	NCIH810				
NSCLC	NCIH2172				
	NCIH1993				
NSCLC	NCIH2110				
腺癌	NCIH1563				

[0374] 下表2B示出鉴定为携带一个或多个SWI-SNF改变的33种肺癌细胞系中的SMARCA2和SMARCA4蛋白状态。深灰色指示突变,浅灰色指示功能丧失,并且空白指示正常功能。

[0375] 表2B:

[0376]

亚型	细胞系	SMARCA2	SMARCA4	KRAS	ARID1A/B	他托司他IC50
腺癌	NCI-H23					10
腺癌	A427					1
腺癌	NCI-H522					7
大细胞	NCI-H661					4
NSCLC	NCI-H1299					10
腺癌	NCIH1703					5.4
腺癌	NCIH2023					10
腺癌	NCIH1568					10
腺癌	NCIH838					10
腺癌	NCIH1793					9
鳞状	HCC15					
腺癌	NCIH2030					10
腺癌	A549					10
腺癌	NCIH1573					10
腺癌	NCIH1944					10
腺癌	NCIH2085					10
腺癌	NCIH1693					
大细胞	NCIH1581					10
腺癌	NCIH2347					
腺癌	NCIH2291					10
腺癌	NCIH1648					
腺癌	NCIH1734					10
未分化的	Calu-6					4.5
大细胞	NCI-H460					10
鳞状	NCIH1869					
大细胞	NCIH810					2.4
NSCLC	NCIH2172					10
	NCIH1993					10
NSCLC	NCIH2110					10
腺癌	NCIH1563					
腺癌	NCI-H441					10
腺癌	NCIH2122					10
腺癌	NCIH1373					10

	突变
	低至没有mRNA/蛋白质丧失

[0377] 将SWI/SNF改变的细胞系在体外用EZH2抑制剂他托司他处理,并且在处理14天后评估细胞增殖(参见图2)。发现SMARCA2/SMARCA4双重丧失肺癌细胞系比具有其他SWI/SNF畸变的肺癌细胞系对EZH2抑制更敏感。

[0378] 在体内肺癌异种移植物中对EZH2抑制的敏感性

[0379] 在NSCLC异种移植物模型的背景下将SMARCA4单一丧失和SMARCA2/SMARCA4双重丧失NSCLC细胞系用EZH2抑制剂他托司他以临床上可实现的剂量(约250mg/kg的体重)体内处理(图3和4)。与体外数据一致,肿瘤生长抑制在SMARCA2/SMARCA4双重丧失异种移植物中比在SMARCA4单一丧失异种移植物中更显著。在四种SMARCA2/SMARCA4双重丧失细胞系中的两

种中,观察到肿瘤消退(图3)。

[0380] 讨论

[0381] 本文提供的数据证明,肺癌的亚型,SMARCA2/SMARCA4双重丧失NSCLC可以有效地通过EZH2抑制来治疗。原发性NSCLC肿瘤(包括展现出SMARCA2/SMARCA4双重丧失的那些)典型地是低分化的腺癌类型(例如,实体腺癌),并且很多情况下展现出上皮间充质转化(EMT)特征(例如,低的E-钙粘蛋白表达水平和高是波形蛋白表达水平)。这些特性与横纹肌样瘤(例如,低分化和间充质样的)的特征一致,并且因此指明先前未认识的特征在于SMARCA2/SMARCA4双重丧失的NSCLC的横纹肌样亚型。SMARCA2和SMARCA4双重丧失与NSCLC患者存活率降低相关(参见,例如reisman等人Cancer res[癌症研究]2003,通过援引并入本文)。此外,双重丧失肿瘤很多情况下对与NSCLC相关的其他突变(例如,EGFR、KRAS、ALK融合)呈阴性,因此限制了可用的疗法选择。因此,SMARCA2/SMARCA4双重丧失NSCLC肿瘤种类代表有高度未满足的医疗需求的肺癌亚型。本披露证明,在此类肿瘤中EZH2抑制有效地抑制肿瘤生长和/或引起所期望的临床结果。

[0382] 示例性序列

[0383] SMARCA2

[0384] >NM_001289396.1智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子,亚家族a,成员2(SMARCA2),转录物变体3,mRNA

[0385] TCAGAAGAAAGCCCCGAGATCACAGAGACCCGGCGAGATCACAGAGACCCGGCCTGAAGGAACGTGGA
AAGACCAATGTACCTGTTTTGACCGGTTGCCTGGAGCAAGAAGTTCCAGTTGGGGAGAATTTTCAGAAGATAAAGT
CGGAGATTGTGAAAGACTTGACTTGCAGCATTACTCTACTGACTGGCAGAGACAGGAGAGGTAGATGTCCACGCC
CACAGACCCTGGTGCATGCCCCACCCAGGGCCTTCGCCGGGGCCTGGGCCTTCCCCTGGGCCAATTCTTGGGCCT
AGTCCAGGACCAGGACCATCCCCAGGTTCCGTCCACAGCATGATGGGGCCAAGTCCTGGACCTCCAAGTGTCTCCC
ATCCTATGCCGACGATGGGGTCCACAGACTTCCCACAGGAAGGCATGCATCAAATGCATAAGCCCATCGATGGTAT
ACATGACAAGGGGATTGTAGAAGACATCCATTGTGGATCCATGAAGGGCACTGGTATGCGACCACCTCACCCAGGC
ATGGGGCCCTCCCCAGAGTCCAATGGATCAACACAGCCAAGGTTATATGTCAACCACACCCATCTCCATTAGGAGCCC
CAGAGCACGTCTCCAGCCCTATGTCTGGAGGAGGCCCAACTCCACCTCAGATGCCACCAAGCCAGCCGGGGGCCCT
CATCCCAGGTGATCCGCAGGCCATGAGCCAGCCCAACAGAGGTCCCTCACCTTTCAGTCCTGTCCAGCTGCATCAG
CTTCGAGCTCAGATTTTAGCTTATAAAATGCTGGCCCGAGGCCAGCCCCCTCCCCGAAACGCTGCAGCTTGCAGTCC
AGGGGAAAAGGACGTTGCCTGGCTTGCAGCAACAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GCAGCAGCAGCAACAGCAGCCGAGCAGCAGCCGCCGCAACCACAGACGCAGCAACAACAGCAGCCGGGCCCTTGTT
AACTACAACAGACCATCTGGCCCGGGGCCGGAGCTGAGCGGCCCGAGCACCCCGCAGAAGCTGCCGGTGCCGCGC
CCGGCGGCCCGGCCCTCGCCCGCGCCCCCGCAGCCGCGCAGCCGCCCGCGGCCGCGAGTGCCCGGGCCCTCAGTGCC
GCAGCCGGCCCCGGGGCAGCCCTCGCCCGTCTCCAGCTGCAGCAGAAGCAGAGCCGCATCAGCCCCATCCAGAAA
CCGCAAGGCCTGGACCCCGTGGAATTCTGCAAGAGCGGAATACAGACTTCAGGCCCGCATAGCTCATAGGATAC
AAGAACTGGAAAATCTGCCTGGCTCTTTGCCACCAGATTTAAGAACCAAGCAACCGTGGAATAAAAGCACTTCG
GTTACTCAATTTCCAGCGTCAGCTGAGACAGGAGGTGGTGGCCTGCATGCGCAGGGACACGACCCTGGAGACGGCT
CTCAACTCCAAAGCATACAAACGGAGCAAGCGCCAGACTCTGAGAGAAGCTCGCATGACCGAGAAGCTGGAGAAGC
AGCAGAAGATTGAGCAGGAGAGGAAACGCCGTGAGAAACACCAGGAATACCTGAACAGTATTTTGCAACATGCAAA
AGATTTTAAGGAATATCATCGGTCTGTGGCCGGAAGATCCAGAAGCTCTCCAAAGCAGTGGCAACTTGGCATGCC

AACACTGAAAGAGAGCAGAAGAAGGAGACAGAGCGGATTGAAAAGGAGAGAATGCGGGCGACTGATGGCTGAAGATG
AGGAGGGTTATAGAAAAGTATTGATCAAAAAGAAAGACAGGCGTTTAGCTTACCTTTTGCAGCAGACCGATGAGTA
TG TAGCCAATCTGACCAATCTGGTTTGGGAGCACAAGCAAGCCCAGGCAGCCAAAGAGAAGAAGAAGAGGAGGAGG
AGGAAGAAGAAGGCTGAGGAGAATGCAGAGGGTGGGGAGTCTGCCCTGGGACCGGATGGAGAGCCCATAGATGAGA
GCAGCCAGATGAGTGACCTCCCTGTCAAAGTGACTCACACAGAAACCGGCAAGGTTCTGTTCCGACCAGAAGCACC
CAAAGCAAGTCAGCTGGACGCCTGGCTGGAAATGAATCCTGGTTATGAAGTTGCCCTAGATCTGACAGTGAAGAG
AGTGATTCTGATTATGAGGAAGAGGATGAGGAAGAAGAGTCCAGTAGGCAGGAAACCGAAGAGAAAATACTCCTGG
ATCCAAATAGCGAAGAAGTTTCTGAGAAGGATGCTAAGCAGATCATTGAGACAGCTAAGCAAGACGTGGATGATGA
ATACAGCATGCAGTACAGTGCCAGGGGCTCCCAGTCTACTACACCGTGGCTCATGCCATCTCGGAGAGGGTGGAG
AAACAGTCTGCCCTCCTAATTAATGGGACCCTAAAGCATTACCAGCTCCAGGGCCTGGAATGGATGGTTTCCCTGT
ATAATAACAACCTTGAACGGAATCTTAGCCGATGAAATGGGGCTTGGAAGACCATACAGACCATTGCACTCATCAC
TTATCTGATGGAGCAGAAAAGACTCAATGGCCCCATCTCATCATTGTTCCCTTTGACTCTATCTAACTGGACA
TATGAATTTGACAAATGGGCTCCTTCTGTGGTGAAGATTTCTTACAAGGGTACTCTGCCATGCGTCGCTCCCTTG
TCCCCAGCTACGGAGTGGCAAATTCATGTCCTCTTGACTACTTATGAGTATATTATAAAAGACAAGCACATTCT
TGCAAAGATTTCGGTGAAAATACATGATAGTGGACGAAGGCCACCGAATGAAGAATCACCAGTGAAGCTGACTCAG
GTCTTGAACACTCACTATGTGGCCCCCAGAAGGATCCTCTTGACTGGGACCCCGCTGCAGAATAAGCTCCCTGAAC
TCTGGGCCCTCCTCAACTTCCTCCTCCCAACAATTTTAAAGAGCTGCAGCACATTTGAACAATGGTTCAATGCTCC
ATTTGCCATGACTGGTGAAAGGGTGGACTTAAATGAAGAAGAACTATATTGATCATCAGGCGTCTACATAAGGTG
TTAAGACCATTTTTACTAAGGAGACTGAAGAAAGAAGTTGAATCCCAGCTTCCCGAAAAAGTGAATATGTGATCA
AGTGTGACATGTCAGCTCTGCAGAAGATTCTGTATCGCCATATGCAAGCCAAGGGGATCCTTCTCACAGATGGTTC
TGAGAAAGATAAGAAGGGGAAAGGAGGTGCTAAGACACTTATGAACACTATTATGCAGTTGAGAAAAATCTGCAAC
CACCACATATATGTTTCAGCACATTGAGGAATCCTTTGCTGAACACCTAGGCTATTCAAATGGGGTCATCAATGGGG
CTGAACTGTATCGGGCTCAGGGAAGTTGAGCTGCTTGATCGTATTCTGCCAAAATTGAGAGCGACTAATCACCG
AGTGCTGCTTTTCTGCCAGATGACATCTCTCATGACCATCATGGAGGATTATTTGCTTTTCGGAACCTCCTTTAC
CTACGCCTTGATGGCACCACCAAGTCTGAAGATCGTGCTGCTTTGCTGAAGAAATTCAATGAACCTGGATCCCAGT
ATTTCAATTTTCTTGCTGAGCACAAGAGCTGGTGGCCTGGGCTTAAATCTTCAGGCAGCTGATACAGTGGTCATCTT
TGACAGCGACTGGAATCCTCATCAGGATCTGCAGGCCAAGACCGAGCTCACCGCATCGGGCAGCAGAACGAGGTC
CGGGTACTGAGGCTCTGTACCGTGAACAGCGTGGAGGAAAAGATCCTCGCGGCCGAAAATACAAGCTGAACGTGG
ATCAGAAAGTGATCCAGGCGGGCATGTTTGACCAAAGTCTTCAAGCCACGAGCGGAGGGCATTCTGCAGGCCAT
CTTGAGCATGAGGAGGAAAAATGAGGAAGAAGATGAAGTACCGGACGATGAGACTCTGAACCAAATGATTGCTCGA
CGAGAAGAAGAATTTGACCTTTTTATGCGGATGGACATGGACCGGCGGAGGGAAGATGCCCGGAACCCGAAACGGA
AGCCCCGTTTAATGGAGGAGGATGAGCTGCCCTCCTGGATCATTAAGGATGACGCTGAAGTAGAAAGGCTCACCTG
TGAAGAAGAGGAGGAGAAAAATATTTGGGAGGGGTCCCGCCAGCGCCGTGACGTGGACTACAGTGACGCCCTCACG
GAGAAGCAGTGGCTAAGGGCCATCGAAGACGGCAATTTGGAGGAAATGGAAGAGGAAGTACGGCTTAAGAAGCGAA
AAAGACGAAGAAATGTGGATAAAGATCCTGCAAAAGAAGATGTGGAAAAAGCTAAGAAGAGAAGAGGCCGCCCTCC
CGCTGAGAAACTGTACCAAATCCCCCAAAGTACAAAGCAGATGAACGCTATCATCGATACTGTGATAAACTAC
AAAGATAGGTGTAACGTGGAGAAGGTGCCAGTAATTCTCAGTTGGAAATAGAAGGAAACAGTTCAGGGCGACAGC
TCAGTGAAGTCTTCATTACGTTACCTTCAAGGAAGAATTACCAGAATACTATGAATTAATTAGGAAGCCAGTGA
TTTCAAAAAATAAAGGAAAGGATTCGTAATCATAAGTACCGGAGCCTAGGCGACCTGGAGAAGGATGTCATGCTT

CTCTGTCACAACGCTCAGACGTTCAACCTGGAGGGATCCCAGATCTATGAAGACTCCATCGTCTTACAGTCAGTGT
TTAAGAGTGGCCGGCAGAAAATTGCCAAAGAGGAAGAGAGTGAGGATGAAAGCAATGAAGAGGAGGAAGAGGAAGA
TGAAGAAGAGTCAGAGTCCGAGGCAAAATCAGTCAAGGTGAAAATTAAGCTCAATAAAAAAGATGACAAAGGCCGG
GACAAAGGGAAAGGCAAGAAAAGGCCAAATCGAGGAAAAGCCAAACCTGTAGTGAGCGATTTTGACAGCGATGAGG
AGCAGGATGAACGTGAACAGTCAGAAGGAAGTGGGACGGATGATGAGTGATCAGTATGGACCTTTTTCCTTGGTAG
AACTGAATTCCTTCCTCCCCTGTCTCATTTCTACCCAGTGAGTTCATTTGTCATATAGGCACTGGGTGTTTCTAT
ATCATCATCGTCTATAAACTAGCTTTAGGATAGTGCCAGACAAACATATGATATCATGGTGTAACAAACACACACA
TACACAAATATTTGTAACATATTGTGACCAAATGGGCCTCAAAGATTCAGATTGAAACAAACAAAAAGCTTTTGAT
GGAAAATATGTGGGTGGATAGTATATTTCTATGGGTGGGTCTAATTTGGTAACGGTTTGATTGTGCCTGGTTTTAT
CACCTGTTTCAGATGAGAAGATTTTGTCTTTTGTAGCACTGATAACCAGGAGAAGCCATTAAAAGCCACTGGTTAT
TTTATTTTTCATCAGGCAATTTTCGAGGTTTTTATTTGTTTCGGTATTGTTTTTTTACACTGTGGTACATATAAGCA
ACTTTAATAGGTGATAAATGTACAGTAGTTAGATTTACCTGCATATACATTTTCCATTTTATGCTCTATGATCT
GAACAAAAGCTTTTTGAATTGTATAAGATTTATGTCTACTGTAAACATTGCTTAATTTTTTTGCTCTTGATTTAA
AAAAAGTTTTGTTGAAAGCGCTATTGAATATTGCAATCTATATAGTGTATTGGATGGCTTCTTTTGTACCCTGAT
CTCCTATGTTACCAATGTGTATCGTCTCCTTCTCCCTAAAGTGTACTTAATCTTTGCTTTCTTTGCACAATGTCTT
TGGTTGCAAGTCATAAGCCTGAGGCAAATAAAATTCCAGTAATTTTGAAGAATGTGGTGTGGTGCTTTCCTAATA
AAGAAATAATTTAGCTTGACAAAAA(AAA) (SEQ ID NO:1)。

[0386] >NM_139045.3智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调节子,亚
家族a,成员2(SMARCA2),转录物变体2,mRNAGCGTCTTCCGGCGCCCGCGGAGGAGGCGAGGGTGGGACG
CTGGGCGGAGCCCGAGTTTAGGAAGAGGAGGGGACGGCTGTCATCAATGAAGTCATATTCATAATCTAGTCCTCTC
TCCCTCTGTTTCTGTACTCTGGGTGACTCAGAGAGGGAAGAGATTTCAGCCAGCACACTCCTCGCGAGCAAGCATT
CTCTACTGACTGGCAGAGACAGGAGAGGTAGATGTCCACGCCCACAGACCCTGGTGCGATGCCCCACCCAGGGCCT
TCGCCGGGGCCTGGGCCTTCCCCTGGGCCAATTCTTGGGCCTAGTCCAGGACCAGGACCATCCCCAGGTTCCGTCC
ACAGCATGATGGGGCCAAGTCCTGGACCTCCAAGTGTCTCCCATCCTATGCCGACGATGGGGTCCACAGACTTCCC
ACAGGAAGGCATGCATCAAATGCATAAGCCCATCGATGGTATACATGACAAGGGGATTGTAGAAGACATCCATTGT
GGATCCATGAAGGGCACTGGTATGCGACCACCTCACCCAGGCATGGGCCCTCCCCAGAGTCCAATGGATCAACACA
GCCAAGGTTATATGTCACCACACCCATCTCCATTAGGAGCCCCAGAGCACGTCTCCAGCCCTATGTCTGGAGGAGG
CCCAACTCCACCTCAGATGCCACCAAGCCAGCCGGGGGCCCTCATCCCAGGTGATCCGCAGGCCATGAGCCAGCCC
AACAGAGGTCCCTCACCTTTCAGTCCTGTCCAGCTGCATCAGCTTCGAGCTCAGATTTTAGCTTATAAAATGCTGG
CCCAGGGCCAGCCCTCCCCGAAACGCTGCAGCTTGAGTCCAGGGGAAAAGGACGTTGCCTGGCTTGACAGCAACA
ACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCCGCAGCAGCAGCCG
CCGCAACCACAGACGCAGCAACAACAGCAGCCGGCCCTTGTTAACTACAACAGACCATCTGGCCCCGGGGCCGGAGC
TGAGCGGCCCCGAGACCCCCGAGAAGCTGCCGGTGCCCGCGCCCCGGCGGCCCTCGCCCCGCCCCCGCAGC
CGCGCAGCCGCCCGCGGCCGAGTGCCCGGGCCCTCAGTGCCGAGCCGGCCCCGGGGCAGCCCTCGCCCGTCCTC
CAGCTGCAGCAGAAGCAGAGCCGCATCAGCCCCATCCAGAAACCGCAAGGCCTGGACCCCGTGAAATTCTGCAAG
AGCGGGAATACAGACTTCAGGCCGCATAGCTCATAGGATACAAGAACTGGAAAATCTGCCTGGCTCTTTGCCACC
AGATTTAAGAACCAAAGCAACCGTGGAATAAAAGCACTTCGGTTACTCAATTTCCAGCGTCAGCTGAGACAGGAG
GTGGTGGCCTGCATGCGCAGGGACACGACCCTGGAGACGGCTCTCAACTCCAAAGCATACAAACGGAGCAAGCGCC
AGACTCTGAGAGAAGCTCGCATGACCGAGAAGCTGGAGAAGCAGCAGAAGATTGAGCAGGAGAGGAAACGCCGTCA

GAAACACCAGGAATACCTGAACAGTATTTTGAACATGCAAAAGATTTTAAGGAATATCATCGGTCTGTGGCCGGA
AAGATCCAGAAGCTCTCCAAAGCAGTGGCAACTTGGCATGCCAACTGAAAGAGAGCAGAAGAAGGAGACAGAGC
GGATTGAAAAGGAGAGAATGCGGCGACTGATGGCTGAAGATGAGGAGGGTTATAGAAAAGTATTGATCAAAAGAA
AGACAGGCGTTTAGCTTACCTTTTGCAGCAGACCGATGAGTATGTAGCCAATCTGACCAATCTGGTTTGGGAGCAC
AAGCAAGCCCAGGCAGCCAAAGAGAAGAAGAAGAGGAGGAGGAGGAAGAAGAAGGCTGAGGAGAATGCAGAGGGTG
GGGAGTCTGCCCTGGGACCGGATGGAGAGCCCATAGATGAGAGCAGCCAGATGAGTGACCTCCCTGTCAAAGTGAC
TCACACAGAAACCGGCAAGGTTCTGTTTCGGACCAGAAGCACCCAAAGCAAGTCAGCTGGACGCCTGGCTGGAAATG
AATCCTGGTTATGAAGTTGCCCCTAGATCTGACAGTGAAGAGAGTGATTCTGATTATGAGGAAGAGGATGAGGAAG
AAGAGTCCAGTAGGCAGGAAACCGAAGAGAAAATACTCCTGGATCCAAATAGCGAAGAAGTTTCTGAGAAGGATGC
TAAGCAGATCATTGAGACAGCTAAGCAAGACGTGGATGATGAATACAGCATGCAGTACAGTGCCAGGGGCTCCCAG
TCCTACTACACCGTGGCTCATGCCATCTCGGAGAGGGTGGAGAAACAGTCTGCCCTCCTAATTAATGGGACCCTAA
AGCATTACCAGCTCCAGGGCCTGGAATGGATGGTTTCCCTGTATAATAACAAGTGAACGGAATCTTAGCCGATGA
AATGGGGCTTGAAAGACCATAACAGACCATTGCACTCATCACTTATCTGATGGAGCACAAAAGACTCAATGGCCCC
TATCTCATCATTGTTCCCTTTCGACTCTATCTAACTGGACATATGAATTTGACAAATGGGCTCCTTCTGTGGTGA
AGATTTCTTACAAGGGTACTCCTGCCATGCGTCGCTCCCTTGTCCCCAGCTACGGAGTGGCAAATTCATGTCTCT
CTTGACTACTTATGAGTATATTATAAAAGACAAGCACATTCTTGCAAAGATTCGGTGGAAATACATGATAGTGGAC
GAAGGCCACCGAATGAAGAATCACCAGTGAAGCTGACTCAGGTCTTGAACACTCACTATGTGGCCCCCAGAAGGA
TCCTCTTGACTGGGACCCCGCTGCAGAATAAGCTCCCTGAACTCTGGGCCCTCCTCAACTTCCTCCTCCCAACAAT
TTTTAAGAGCTGCAGCACATTTGAACAATGGTTCAATGCTCCATTTGCCATGACTGGTGAAGGGTGGACTTAAAT
GAAGAAGAACTATATTGATCATCAGGCGTCTACATAAGGTGTTAAGACCATTTTACTAAGGAGACTGAAGAAAG
AAGTTGAATCCCAGCTTCCCGAAAAAGTGAATATGTGATCAAGTGTGACATGTCAGCTCTGCAGAAGATTCTGTA
TCGCCATATGCAAGCCAAGGGGATCCTTCTCACAGATGGTTCTGAGAAAAGATAAGAAGGGGAAAGGAGGTGCTAAG
ACACTTATGAACACTATTATGCAGTTGAGAAAAATCTGCAACCAACCATATATGTTTCAGCACATTGAGGAATCCT
TTGCTGAACACCTAGGCTATTCAAATGGGGTCATCAATGGGGCTGAACTGTATCGGGCCTCAGGGAAGTTTGAGCT
GCTTGATCGTATTCTGCCAAAATTGAGAGCGACTAATCACCGAGTGCTGCTTTTCTGCCAGATGACATCTCTCATG
ACCATCATGGAGGATTATTTTGCTTTTCGGAACCTTCCTTTACCTACGCCTTGATGGCACCACCAAGTCTGAAGATC
GTGCTGCTTTGCTGAAGAAATTCAATGAACCTGGATCCCAGTATTTCAATTTCTTGCTGAGCACAAGAGCTGGTGG
CCTGGGCTTAAATCTTCAGGCAGCTGATACAGTGGTCACTTTGACAGCGACTGGAATCCTCATCAGGATCTGCAG
GCCAAGACCGAGCTCACCGCATCGGGCAGCAGAACGAGGTCCGGGTACTGAGGCTCTGTACCGTGAACAGCGTGG
AGGAAAAGATCCTCGCGGCCGCAAAATACAAGCTGAACGTGGATCAGAAAAGTGATCCAGGCGGGCATGTTTGACCA
AAAGTCTTCAAGCCACGAGCGGAGGGCATTCCTGCAGGCCATCTTGAGCATGAGGAGGAAAATGAGGAAGAAGAT
GAAGTACCGGACGATGAGACTCTGAACCAAATGATTGCTCGACGAGAAGAAGAATTTGACCTTTTTATGCGGATGG
ACATGGACCGGCGGAGGGAAGATGCCCCGAACCCGAAACGGAAGCCCCGTTAATGGAGGAGGATGAGCTGCCCTC
CTGGATCATTAAGGATGACGCTGAAGTAGAAAGGCTCACCTGTGAAGAAGAGGAGGAGAAAATATTTGGGAGGGGG
TCCCGCCAGCGCCGTGACGTGGACTACAGTGACGCCCTCACGGAGAAGCAGTGGCTAAGGGCCATCGAAGACGGCA
ATTTGGAGGAAATGGAAGAGGAAGTACGGCTTAAGAAGCGAAAAAGACGAAGAAATGTGGATAAAGATCCTGCAAA
AGAAGATGTGGAAGAAAGCTAAGAAGAGAAGAGGCCGCCCTCCCGCTGAGAACTGTCACCAAATCCCCCAAAGT
ACAAAGCAGATGAACGCTATCATCGATACTGTGATAAACTACAAAGATAGTTCAGGGCGACAGCTCAGTGAAGTCT
TCATTCAGTTACCTTCAAGGAAAGAATTACCAGAATACTATGAATTAATTAGGAAGCCAGTGGATTTCAAAAAAT

AAAGGAAAGGATTCGTAATCATAAGTACCGGAGCCTAGGCGACCTGGAGAAGGATGTCATGCTTCTCTGTCACAAC
GCTCAGACGTTCAACCTGGAGGGATCCCAGATCTATGAAGACTCCATCGTCTTACAGTCAGTGTTTAAGAGTGCCC
GGCAGAAAATTGCCAAAGAGGAAGAGAGTGAGGATGAAAGCAATGAAGAGGAGGAAGAGGAAGATGAAGAAGAGTC
AGAGTCCGAGGCAAAATCAGTCAAGGTGAAAATTAAGCTCAATAAAAAAGATGACAAAGGCCGGGACAAAGGGAAA
GGCAAGAAAAGGCCAAATCGAGGAAAAGCCAAACCTGTAGTGAGCGATTTTGACAGCGATGAGGAGCAGGATGAAC
GTGAACAGTCAGAAGGAAGTGGGACGGATGATGAGTGATCAGTATGGACCTTTTTCCTTGGTAGAACTGAATTCCT
TCCTCCCCTGTCTCATTTCTACCCAGTGAGTTCATTTGTTCATATAGGCACTGGGTTGTTTCTATATCATCATCGTC
TATAAACTAGCTTTAGGATAGTGCCAGACAAACATATGATATCATGGTGTAACCAACACACACATACACAAATATT
TGTAACATATTGTGACCAAAATGGGCCCTCAAAGATTGAGATTGAAACAAACAAAAAGCTTTTGATGGAAAATATGTG
GGTGGATAGTATATTTCTATGGGTGGGTCTAATTTGGTAACGGTTTGATTGTGCCTGGTTTTATCACCTGTTTCTAGA
TGAGAAGATTTTTGTCTTTTGTAGCACTGATAACCAGGAGAAGCCATTAAGGCCACTGGTTATTTTTATTTTTTCAT
CAGGCAATTTTCGAGGTTTTTATTTGTTTCGGTATTGTTTTTTTACACTGTGGTACATATAAGCAACTTTAATAGGT
GATAAATGTACAGTAGTTAGATTTCACCTGCATATACATTTTTTCCATTTTATGCTCTATGATCTGAACAAAAGCTT
TTTGAATTGTATAAGATTTATGTCTACTGTAAACATTGCTTAATTTTTTTTGTCTTGATTTAAAAAAGTTTTGT
TGAAAGCGCTATTGAATATTGCAATCTATATAGTGTATTGGATGGCTTCTTTTGTACCCCTGATCTCCTATGTTAC
CAATGTGTATCGTCTCCTTCTCCCTAAAGTGACTTAATCTTTGCTTTCTTTGCACAATGTCTTTGGTTGCAAGTC
ATAAGCCTGAGGCAAAATAAAATTCAGTAATTTGGAAGAATGTGGTGTGGTGCTTTCCTAATAAAGAAATAATTT
AGCTTGACAAAAA (SEQ ID NO:2)。

[0387] >NM_001289397.1 智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员2 (SMARCA2)，转录物变体4，mRNA

[0388] GCGTCTTCCGGCGCCCGCGGAGGAGGCGAGGGTGGGACGCTGGGCGGAGCCCGAGTTAGGAAGAGGA
GGGGACGGCTGTCATCAATGAAGTCATATTCATAATCTAGTCCTCTCTCCCTCTGTTTCTGTACTCTGGGTGACTC
AGAGAGGGAAGAGATTGAGCCAGCACACTCCTCGCGAGCAAGCATTACTCTACTGACTGGCAGAGACAGGAGAGGT
AGATGTCCACGCCCACAGACCCTGGTGCGATGCCCCACCCAGGGCCTTCGCCGGGGCCTGGGCCTTCCCCTGGGCC
AATTCTTGGGCCTAGTCCAGGACCAGGACCATCCCCAGGTTCCGTCCACAGCATGATGGGGCCAAGTCCTGGACCT
CCAAGTGTCTCCCATCCTATGCCGACGATGGGGTCCACAGACTTCCCACAGGAAGGCATGCATCAAATGCATAAGC
CCATCGATGGTATACATGACAAGGGGATTGTAGAAGACATCCATTGTGGATCCATGAAGGGCACTGGTATGCGACC
ACCTCACCCAGGCATGGGCCCTCCCCAGAGTCCAATGGATCAACACAGCCAAGGTTATATGTCACCACACCCATCT
CCATTAGGAGCCCCAGAGCACGTCTCCAGCCCTATGTCTGGAGGAGGCCAACTCCACCTCAGATGCCACCAAGCC
AGCCGGGGGCCCTCATCCCAGGTGATCCGCAGGCCATGAGCCAGCCCAACAGAGGTCCCTCACCTTTTCAGTCTGT
CCAGCTGCATCAGCTTCGAGCTCAGATTTTAGCTTATAAAATGCTGGCCCCGAGGCCAGCCCCTCCCCGAAACGCTG
CAGCTTGCACTCCAGGGGAAAAGGACGTTGCCTGGCTTGCAAGCAACAACAGCAGCAGCAACAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCCGAGCAGCAGCCGCCGCAACCACAGACGCAGCAACAACAGCA
GCCGGCCCTTGTTAACTACAACAGACCATCTGGCCCCGGGGCCGAGCTGAGCGGCCCGAGACCCCGCAGAAGCTG
CCGGTGCCCGCGCCCGGGCGGCCGCTCGCCCGCGCCCCCGCAGCCGCGCAGCCCGCGCGCCGCGCAGTGCCCG
GGCCCTCAGTGCCGAGCCGGCCCCGGGGCAGCCCTCGCCCGTCTCCAGCTGCAGCAGAAGCAGAGCCGCATCAG
CCCCATCCAGAAACCGCAAGGCCTGGACCCCGTGAAATTCTGCAAGAGCGGGAATACAGACTTCAGGCCCGCATA
GCTCATAGGATACAAGAACTGGAATCTGCCTGGCTCTTTGCCACCAGATTTAAGAACCAAGCAACCGTGGAAC
TAAAGCACTTCGGTTACTCAATTTCCAGCGTCAGCTGAGACAGGAGGTGGTGGCCTGCATGCGCAGGGACACGAC

CCTGGAGACGGCTCTCAACTCCAAAGCATACAAACGGAGCAAGCGCCAGACTCTGAGAGAAGCTCGCATGACCGAG
AAGCTGGAGAAGCAGCAGAAGATTGAGCAGGAGAGGAAACGCCGTCAGAAACACCAGGAATACCTGAACAGTATTT
TGCAACATGCAAAAGATTTTAAGGAATATCATCGGTCTGTGGCCGAAAAGATCCAGAAGCTCTCCAAAGCAGTGGC
AACTTGGCATGCCAACACTGAAAGAGAGCAGAAGAAGGAGACAGAGCGGATTGAAAAGGAGAGAATGCGGCGACTG
ATGGCTGAAGATGAGGAGGGTTATAGAAAACCTGATTGATCAAAAAGAAAGACAGGCGTTTAGCTTACCTTTTGCAGC
AGACCGATGAGTATGTAGCCAATCTGACCAATCTGGTTTGGGAGCACAAGCAAGCCCAGGCAGCCAAAGAGAAGAA
GAAGAGGAGGAGGAGGAAGAAGAAGGCTGAGGAGAATGCAGAGGGTGGGGAGTCTGCCCTGGGACCGGATGGAGAG
CCCATAGATGAGAGCAGCCAGATGAGTGACCTCCCTGTCAAAGTGACTCACACAGAAACCGGCAAGGTTCTGTTTCG
GACCAGAAGCACCCAAAGCAAGTCAGCTGGACGCCTGGCTGGAAATGAATCCTGGTTATGAAGTTGCCCTAGATC
TGACAGTGAAGAGAGTGATTCTGATTATGAGGAAGAGGATGAGGAAGAAGAGTCCAGTAGGCAGGAAACCGAAGAG
AAAATACTCCTGGATCCAAATAGCGAAGAAGTTTCTGAGAAGGATGCTAAGCAGATCATTGAGACAGCTAAGCAAG
ACGTGGATGATGAATACAGCATGCAGTACAGTGCCAGGGGCTCCAGTCCCTACTACACCGTGGCTCATGCCATCTC
GGAGAGGGTGGAGAAACAGTCTGCCCTCCTAATTAATGGGACCCTAAAGCATTACCAGCTCCAGGGCCTGGAATGG
ATGGTTTCCCTGTATAATAACAACTTGAACGGAATCTTAGCCGATGAAATGGGGCTTGGAAGACCATACAGACCA
TTGCACTCATCACTTATCTGATGGAGCACAAGAACTCAATGGCCCCCTATCTCATCATTGTTCCCCCTTCGACTCT
ATCTAACTGGACATATGAATTTGACAAATGGGCTCCTTCTGTGGTGAAGATTTCTTACAAGGGTACTCCTGCCATG
CGTCGCTCCCTTGTCCCCCAGCTACGGAGTGGCAAATTCATGTCTCTTGACTACTTATGAGTATATTATAAAAG
ACAAGCACATTCTTGCAAAGATTTCGGTGGAAATACATGATAGTGGACGAAGGCCACCGAATGAAGAATCACCCTG
CAAGCTGACTCAGGTGGACTTAAATGAAGAAGAACTATATTGATCATCAGGCGTCTACATAAGGTGTTAAGACCA
TTTTTACTAAGGAGACTGAAGAAAGAAGTTGAATCCAGCTTCCCGAAAAAGTGGAATATGTGATCAAGTGTGACA
TGTCAGCTCTGCAGAAGATTCTGTATCGCCATATGCAAGCCAAGGGGATCCTTCTCACAGATGGTTCTGAGAAAGA
TAAGAAGGGGAAAGGAGGTGCTAAGACACTTATGAACACTATTATGCAGTTGAGAAAAATCTGCAACCACCCATAT
ATGTTTCAGCACATTGAGGAATCCTTTGCTGAACACCTAGGCTATTCAAATGGGGTCATCAATGGGGCTGAACTGT
ATCGGGCCTCAGGGAAGTTTGAGCTGCTTGATCGTATTCTGCCAAAATTGAGAGCGACTAATCACCGAGTGCTGCT
TTTCTGCCAGATGACATCTCTCATGACCATCATGGAGGATTATTTTGCTTTTCGGAACCTTCCTTTACCTACGCCTT
GATGGCACCACCAAGTCTGAAGATCGTGCTGCTTTGCTGAAGAAATTCAATGAACCTGGATCCCAGTATTTTCATTT
TCTTGCTGAGCACAAGAGCTGGTGGCCTGGGCTTAAATCTTCAGGCAGCTGATACAGTGGTCATCTTTGACAGCGA
CTGGAATCCTCATCAGGATCTGCAGGCCCAAGACCGAGCTCACCGCATCGGGCAGCAGAACGAGGTCCGGGTACTG
AGGCTCTGTACCGTGAACAGCGTGGAGGAAAAGATCCTCGCGCCGCAAAAATACAAGCTGAACGTGGATCAGAAAG
TGATCCAGGCGGGCATGTTTGACCAAAAAGTCTTCAAGCCACGAGCGGAGGGCATTCTGCAGGCCATCTTGAGCA
TGAGGAGGAAAATGAGGAAGAAGATGAAGTACCGGACGATGAGACTCTGAACCAAATGATTGCTCGACGAGAAGAA
GAATTTGACCTTTTTATGCGGATGGACATGGACCGGCGGAGGGAAGATGCCCGGAACCCGAAACGGAAGCCCCGTT
TAATGGAGGAGGATGAGCTGCCCTCCTGGATCATTAAGGATGACGCTGAAGTAGAAAGGCTCACCTGTGAAGAAGA
GGAGGAGAAAATATTTGGGAGGGGGTCCCGCCAGCGCCGTGACGTGGACTACAGTGACGCCCTCACGGAGAAGCAG
TGGCTAAGGGCCATCGAAGACGGCAATTTGGAGGAAATGGAAGAGGAAGTACGGCTTAAGAAGCGAAAAAGACGAA
GAAATGTGGATAAAGATCCTGCAAAAGAAGATGTGGAAGAAAGCTAAGAAGAGAAGAGGCCGCCCTCCCGCTGAGAA
ACTGTCACCAAATCCCCCAAACCTGACAAAGCAGATGAACGCTATCATCGATACTGTGATAAACTACAAAGATAGT
TCAGGGCGACAGCTCAGTGAAGTCTTCATTTCAGTTACCTTCAAGGAAAGAATTACCAGAATACTATGAATTAATTA
GGAAGCCAGTGGATTTCAAAAAAATAAAGGAAAGGATTCGTAATCATAAGTACCGGAGCCTAGGCGACCTGGAGAA

GGATGTCATGCTTCTCTGTCACAACGCTCAGACGTTCAACCTGGAGGGATCCCAGATCTATGAAGACTCCATCGTC
TTACAGTCAGTGTTTAAAGAGTGCCCGGCAGAAAATTGCCAAAGAGGAAGAGAGTGAGGATGAAAGCAATGAAGAGG
AGGAAGAGGAAGATGAAGAAGAGTCAGAGTCCGAGGCAAAATCAGTCAAGGTGAAAATTAAGCTCAATAAAAAAGA
TGACAAAGGCCGGGACAAAGGGAAGGCAAGAAAAGGCCAAATCGAGGAAAAGCCAAACCTGTAGTGAGCGATTTT
GACAGCGATGAGGAGCAGGATGAACGTGAACAGTCAGAAGGAAGTGGGACGGATGATGAGTGATCAGTATGGACCT
TTTTCTTGGTAGAACTGAATTCCTTCCCTCCCCTGTCTCATTTCTACCCAGTGAGTTCATTTGTCATATAGGCACT
GGGTTGTTTCTATATCATCATCGTCTATAAACTAGCTTTAGGATAGTGCCAGACAAACATATGATATCATGGTGTA
AAAAACACACACATACACAAATATTTGTAACATATTGTGACCAAATGGGCCTCAAAGATTCAGATTGAAACAAACA
AAAAGCTTTTGATGGAAAATATGTGGGTGGATAGTATATTTCTATGGGTGGGTCTAATTTGGTAACGGTTTGATTG
TGCCTGGTTTTATCACCTGTTTCAGATGAGAAGATTTTTGTCTTTTGTAGCACTGATAACCAGGAGAAGCCATTAAA
AGCCACTGGTTATTTTATTTTCATCAGGCAATTTTCGAGGTTTTTATTTGTTTCGGTATTGTTTTTTTACACTGTG
GTACATATAAGCAACTTTAATAGGTGATAAATGTACAGTAGTTAGATTTACCTGCATATACATTTTCCATTTTA
TGCTCTATGATCTGAACAAAAGCTTTTGAATTGTATAAGATTTATGTCTACTGTAAACATTGCTTAATTTTTTTG
CTCTTGATTTAAAAAAAAGTTTTGTTGAAAGCGCTATTGAATATTGCAATCTATATAGTGTATTGGATGGCTTCTT
TTGTACCCCTGATCTCCTATGTTACCAATGTGTATCGTCTCCTTCTCCCTAAAGGTACTTAATCTTTGCTTTCTT
TGCACAATGTCTTTGGTTGCAAGTCATAAGCCTGAGGCAAATAAAATTCCAGTAATTTCGAAGAATGTGGTGTTGG
TGCTTTCCTAATAAAGAAATAATTTAGCTTGACAAAAA(AAA) (SEQ ID NO:3)。

[0389] >NM_001289398.1智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员2 (SMARCA2)，转录物变体5，mRNA

[0390] CTTGGAGAGGCGGAGGTGGAAACGATGCGCAGGAGTTGGCTTGGGGCTTTTTGTTTGGTGTCCCTGT
TTACCTATTATAATCATGGATCCCCTCTGCTTTGTGATACTGTGAACCACGCATAACAGCAATTCCTTACACCAC
CGGGTTGAGAAGAAGGCGCCTGAGGCTGACTTCTGGACCTGCCGTCACGCAGTAAAGATGTGGTTGGCCATCGAA
GACGGCAATTTGGAGGAAATGGAAGAGGAAGTACGGCTTAAGAAGCGAAAAAGACGAAGAAATGTGGATAAAGATC
CTGCAAAAGAAGATGTGAAAAAGCTAAGAAGAGAAGAGGCCGCCCTCCCGCTGAGAACTGTCACCAAATCCCC
CAAAGTACAAAGCAGATGAACGCTATCATCGATACTGTGATAAACTACAAAGATAGTTTCAGGGCGACAGCTCAGT
GAAGTCTTCATTTCAGTTACCTTCAAGGAAAGAATTACCAGAATACTATGAATTAATTAGGAAGCCAGTGGAATTC
AAAAATAAAGGAAAGGATTCGTAATCATAAGTACCGGAGCCTAGGCGACCTGGAGAAGGATGTCATGCTTCTCTG
TCACAACGCTCAGACGTTCAACCTGGAGGGATCCCAGATCTATGAAGACTCCATCGTCTTACAGTCAGTGTTTAAAG
AGTGCCCGGCAGAAAATTGCCAAAGAGGAAGAGAGTGAGGATGAAAGCAATGAAGAGGAGGAAGAGGAAGATGAAG
AAGAGTCAGAGTCCGAGGCAAAATCAGTCAAGGTGAAAATTAAGCTCAATAAAAAAGATGACAAAGGCCGGGACAA
AGGGAAGGCAAGAAAAGGCCAAATCGAGGAAAAGCCAAACCTGTAGTGAGCGATTTTGACAGCGATGAGGAGCAG
GATGAACGTGAACAGTCAGAAGGAAGTGGGACGGATGATGAGTGATCAGTATGGACCTTTTTCTTGGTAGAACTG
AATTCCTTCCCTCCCCTGTCTCATTTCTACCCAGTGAGTTCATTTGTCATATAGGCACTGGGTTGTTTCTATATCAT
CATCGTCTATAAACTAGCTTTAGGATAGTGCCAGACAAACATATGATATCATGGTGTAACAAACACACACATACAC
AAATATTTGTAACATATTGTGACCAAATGGGCCTCAAAGATTCAGATTGAAACAAACAAAAAGCTTTTGATGGAAA
ATATGTGGGTGGATAGTATATTTCTATGGGTGGGTCTAATTTGGTAACGGTTTGATTGTGCCTGGTTTTATCACCT
GTTTCAGATGAGAAGATTTTTGTCTTTTGTAGCACTGATAACCAGGAGAAGCCATTAAAAGCCACTGGTTATTTTAT
TTTTCATCAGGCAATTTTCGAGGTTTTTATTTGTTTCGGTATTGTTTTTTTACACTGTGGTACATATAAGCAACTTT
AATAGGTGATAAATGTACAGTAGTTAGATTTACCTGCATATACATTTTCCATTTTATGCTCTATGATCTGAACA

AAAGCTTTTGAATTGTATAAGATTTATGTCTACTGTAAACATTGCTTAATTTTTTTTGCTCTTGATTAAAAAA
GTTTTGTTGAAAGCGCTATTGAATATTGCAATCTATATAGTGATTGGATGGCTTCTTTGTACCCCTGATCTCT
ATGTTACCAATGTGTATCGTCTCCTTCTCCCTAAAGTGTACTTAATCTTTGCTTTCTTTGCACAATGTCTTTGGTT
GCAAGTCATAAGCCTGAGGCAAATAAAATTCAGTAATTTCAAGAATGTGGTGTGGTGCTTTCCCTAATAAAGAA
ATAATTTAGCTTGACAAAAA(AAA) (SEQ ID NO:4)。

[0391] >NP_001276325.1可能的全局转录激活因子SNF2L2同种型a[智人]

[0392] MSTPTDPGAMPHPGPSPGPGSPGPILGSPGPGSPGSGVHSMGSPGPPSVSHPMPTMGSTDFPQE
GMHQMHPKIDGIDHDKGIVEDIHCGSMKGTGMRPPHPGMGPPQSPMDQHSQGYMSPHPSPLGAPEHVSSPMSGGGPT
PPQMPPSQPGALIPGDPQAMSQPNRGSPFSPVQLHQLRAQILAYKMLARGQPLPETLQLAVQGKRTLPLGLQQQQQ
QQQQQQQQQQQQQQQQQQPPQPQTQQQQPALVNYNRPSPGPELSGPSTPQKLVPAPGGRSPAPPAAAQ
PPAAAVPGSPVPQAPAGQSPVLQLQQKQSRISPIQKPQGLDPEILQEREYRLQARIAHRIQELNLPGLPPDL
RTKATVELKALRLNFRQLRQEVVACMRDRTTLETALNSKAYKRSKRQTLREARMTEKLEKQQKIEQERKRRQKH
QEYLNLSILQHAKDFKEYHRSVAGKIQKLSKAVATWHANTEREQKKETERIEKERMRLMAEDEEGYRKLIDQKKDR
RLAYLLQQTDEYVANLTNLVWEHKQAQAAKEKKRRRRKKKAEENAEGGESALGPDGEPIDESSQMSDLPVKVTH
ETGKVLFGPEAPKASQLDAWLEMNPGYEVAPRSDSEESDSYEEDEEEESSRQETEEKILLDPNSEEVSEKDAKQ
I IETAKQDVDDEYSMQYSARGSQSYTVAHAISERVEKQSALLINGTLKHYQLQGLEWMVSLYNNNLNGILADEMG
LGKTIQTIALITYLMEHKRLNGPYLIIVPLSTLSNWTYEFDKWAPSVKISYKGPAMRRSLVPQLRSGKFNLLT
TYEYIIKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVLNTHYVAPRRILLTGTPQLNKLPELWALLNLLPTIFK
SCSTFEQWFNAPFAMTGERVDLNEETILIRRLHKVLRPFLRLKKEVESQLPEKVEYVIKCDMSALQKILYRH
MQAKGILLTDGSEKDKKGGGAKTLMNTIMQLRKICNHPYMFQHIIEESFAEHLGYSNGVINGAELYRASGKFELLD
RILPKLRATNHRVLLFCQMTSLMTIMEDYFAFRNFLYLRLDGTTKSEDRAALLKKFNPGSYQIFLLSTRAGGLG
LNLQAADTVVIFDSDWNPHQDLQAQDRAHRIGQNEVRVLRRLCTVNSVEEKILAAAKYKLNVDQKVIQAGMFDQKS
SSHERRAFLQAILEHEEENEEDEVPDETNLQMIARREEEFDFMRMDMDRRREDARNPKRKPRLMEEDELPSWI
IKDDAEVERLTCEEEEEKIFGRGSRQRDDVDSALTEKQWLRAIEDGNLEEMEEVRLKKRKRNRNVDKDPKED
VEKAKRRRGRPPAEKLSNPPKLTQMNAI IDTVINYKDRCNVEKVPNSQLEIEGNSSGRQLSEVFIQLPSRKEL
PEYYELIRKPVDFKKIKERIRNHKYRSLGDLEKDVMLLCHNAQTFNLEGSQIYEDSIVLQSVFKSARQKIAKEES
EDESNEEEEEDEEESSEAKSVKVIKLNKKDDKGRDKGKGRPNRGKAKPVVSDFDSDDEEQDEREQSEGSGTD
DE (SEQ ID NO:5)。

[0393] >NP_620614.2可能的全局转录激活因子SNF2L2同种型b[智人]

[0394] MSTPTDPGAMPHPGPSPGPGSPGPILGSPGPGSPGSGVHSMGSPGPPSVSHPMPTMGSTDFPQE
GMHQMHPKIDGIDHDKGIVEDIHCGSMKGTGMRPPHPGMGPPQSPMDQHSQGYMSPHPSPLGAPEHVSSPMSGGGPT
PPQMPPSQPGALIPGDPQAMSQPNRGSPFSPVQLHQLRAQILAYKMLARGQPLPETLQLAVQGKRTLPLGLQQQQQ
QQQQQQQQQQQQQQQQQQPPQPQTQQQQPALVNYNRPSPGPELSGPSTPQKLVPAPGGRSPAPPAAAQ
PPAAAVPGSPVPQAPAGQSPVLQLQQKQSRISPIQKPQGLDPEILQEREYRLQARIAHRIQELNLPGLPPDL
RTKATVELKALRLNFRQLRQEVVACMRDRTTLETALNSKAYKRSKRQTLREARMTEKLEKQQKIEQERKRRQKH
QEYLNLSILQHAKDFKEYHRSVAGKIQKLSKAVATWHANTEREQKKETERIEKERMRLMAEDEEGYRKLIDQKKDR
RLAYLLQQTDEYVANLTNLVWEHKQAQAAKEKKRRRRKKKAEENAEGGESALGPDGEPIDESSQMSDLPVKVTH
ETGKVLFGPEAPKASQLDAWLEMNPGYEVAPRSDSEESDSYEEDEEEESSRQETEEKILLDPNSEEVSEKDAKQ
I IETAKQDVDDEYSMQYSARGSQSYTVAHAISERVEKQSALLINGTLKHYQLQGLEWMVSLYNNNLNGILADEMG

LGKTIQTIALITYLMEHKRLNGPYLIIVPLSTLSNWTYEFDKWAPSVVKISYKGPAMRRSLVPQLRSGKFNVLTT
TYEYIIKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVLNTHYVAPRRILLTGTPQLNKLPELWALLNLLPTIFK
SCSTFEQWFNAPFAMTGERVDLNEEETILIRRLHKVLRPFLRLRLKKEVESQLPEKVEYVIKCDMSALQKILYRH
MQAKGILLTDGSEKDKKGKGAKTLMNTIMQLRKICNHPYMFQHIEESFAEHLGYSNGVINGAELYRASGKFELLD
RILPKLRATNHRVLLFCQMTSLMTIMEDYFAFRNFLYLRLDGTTSKEDRAALLKKFNPGSQYFIFLLSTRAGGLG
LNLQAADTVVIFDSDWNPHQDLQAQDRAHRIGQQNEVRVLRRLCTVNSVEEKILAAAKYKLNVDQKVIQAGMFDQKS
SSHERRAFLQAILEHEEENEEDVPDDETNLNQMIAARREEFDLFMRMDMDRRREDARNPKRKPRLMEEDELPSWI
IKDDAEVERLTCEEEEEKIFGRGSRQRRDVEDYSDALTEKQWLRAIEDGNLEEMEEVRLKKRKRNRNVDKDPKED
VEKAKRRGRPPAEKLSNPPKLTQMNAIIDTVINYKDSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKE
RIRNHKYRSLGDLEKDVMLLCHNAQTFNLEGSQIYEDSIVLQSVFKSARQKIAKEEESSEDESNEEEEEDEESES
EAKSVKVIKLNKKDDKGRDKGKGKRPNRGKAKPVVSDFSDDEEQDEREQSESGTDDE (SEQ ID NO:6)。

[0395] >NP_001276326.1可能的全局转录激活因子SNF2L2同种型c[智人]

[0396] MSTPTDGPAMPHPGSPGPGSPGPILGSPGPGSPGSHMMGSPGPPSVHPMTMGSTDFPQE
GMHQMHPIDGIDHKGIVEDIHCGSMKGTGMRPPHPGMPPQSPMDQHSQGYMSPHPSPLGAPEHVSSPMSGGGPT
PPQMPPSQPGALIPGDPQAMSQPNRGSPSPVQLHQLRAQILAYKMLARGQPLPETLQLAVQGKRTLPLGLQQQQQ
QQQQQQQQQQQQQQQQQQPPQPPQPTQQQQQPALVNYNRPSGPGPELSGPSTPQKLVPAPGGRPSAPPAAAQ
PPAAAVPGPSVPQAPAGQSPVLQLQKQSRISPIQKPQGLDVEILQEREYRLQARIAHRIQELENLPGSLPPDL
RTKATVELKALRLNFQRQLRQEVVACMRDRTLETALNSKAYKRSKRQTLREARMTEKLEKQKQIEQERKRRQKH
QEYLSILQHAKDFKEYHRSVAGKIQKLSKAVATWHANTEREQKKETERIEKERMRLMAEDEEGYRKLIDQKKDR
RLAYLLQQTDEYVANLTNLVWEHKQAQAAKEKKRRRRKKKAEENAEGGESALGPDGEPIDESSQMSDLPVKVTHT
ETGKVLFGPEAPKASQLDAWLEMNPGYEVAPRSDSEESDSYEEEDEEESSRQETEEKILLDPNSEEVSEKDAKQ
IETAKQDVDDEYSMQYSARGSQSYTTVAHAISERVEKQSALLINGTLKHYQLQGLEWMVSLYNNNLNGILADEMG
LGKTIQTIALITYLMEHKRLNGPYLIIVPLSTLSNWTYEFDKWAPSVVKISYKGPAMRRSLVPQLRSGKFNVLTT
TYEYIIKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVDLNEEETILIRRLHKVLRPFLRLRLKKEVESQLPEKV
EYVIKCDMSALQKILYRHMMAKGIILLTDGSEKDKKGKGAKTLMNTIMQLRKICNHPYMFQHIEESFAEHLGYSNG
VINGAELYRASGKFELLDRIPLKLRATNHRVLLFCQMTSLMTIMEDYFAFRNFLYLRLDGTTSKEDRAALLKKFNE
PGSQYFIFLLSTRAGGLGLNLQAADTVVIFDSDWNPHQDLQAQDRAHRIGQQNEVRVLRRLCTVNSVEEKILAAAKY
KLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEENEEDVPDDETNLNQMIAARREEFDLFMRMDMDRRREDAR
NPKRKPRLMEEDELPSWIIKDDAEVERLTCEEEEEKIFGRGSRQRRDVEDYSDALTEKQWLRAIEDGNLEEMEEVRL
LKKRKRNRNVDKDPKEDVEKAKRRGRPPAEKLSNPPKLTQMNAIIDTVINYKDSSGRQLSEVFIQLPSRKEL
PEYYELIRKPVDFKKIKERIRNHKYRSLGDLEKDVMLLCHNAQTFNLEGSQIYEDSIVLQSVFKSARQKIAKEEES
EDESNEEEEEDEESESSEAKSVKVIKLNKKDDKGRDKGKGKRPNRGKAKPVVSDFSDDEEQDEREQSESGTD
DE (SEQ ID NO:7)。

[0397] >NP_001276327.1可能的全局转录激活因子SNF2L2同种型d[智人]

[0398] MWLAIEDGNLEEMEEVRLKKRKRNRNVDKDPKEDVEKAKRRGRPPAEKLSNPPKLTQMNAIID
TVINYKDSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKERIRNHKYRSLGDLEKDVMLLCHNAQTFNLEGS
QIYEDSIVLQSVFKSARQKIAKEEESSEDESNEEEEEDEESESSEAKSVKVIKLNKKDDKGRDKGKGKRPNRGK
AKPVVSDFSDDEEQDEREQSESGTDDE (SEQ ID NO:8)。

[0399] SMARCA4

[0400] >NM_001128849.1 智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，亚家族a，成员4 (SMARCA4)，转录物变体1，mRNA

[0401] GGC GGGG GAGG CGCC GGAAGTCGACGGCGCGCGGCTCCTGCAGGAGGCCACTGTCTGCAGCTCCC
GTGAAGATGTCCACTCCAGACCCACCCCTGGGCGGAACCTCCTCGGCCAGGTCCCTCCCCGGGCCCTGGCCCTTCCC
CTGGAGCCATGCTGGGCCCTAGCCCGGTCCCTCGCCGGGCTCCGCCACAGCATGATGGGGCCCAGCCAGGGCC
GCCCTCAGCAGGACACCCCATCCCCACCCAGGGGCTGGAGGGTACCCTCAGGACAACATGCACCAGATGCACAAG
CCCATGGAGTCCATGCATGAGAAGGGCATGTCGGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCGGT
CAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCTGGG
TGGCTCTGAGCATGCCTCTAGTCCAGTTCCAGCCAGTGGCCCGTCTTCGGGGCCCCAGATGTCTTCCGGGCCAGGA
GGTGCCCCGCTGGATGGTGCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCCAACCCCATTTAACCAGAACC
AGCTGCACCAGCTCAGAGCTCAGATCATGGCCTACAAGATGTGGCCAGGGGGCAGCCCCCTCCCCGACCACCTGCA
GATGGCGGTGCAGGGCAAGCGGCCGATGCCCGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGTCC
GCAACAGGACCCGGCCCTGGCCCTGGCCCGGCCCGGTCCCGGCCCGGCACCTCCAAATTACAGCAGGC
CTCATGGTATGGGAGGGCCCAACATGCCTCCCCCAGGACCCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCTCC
TGGAGGGCCTCCCAAGCCCTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGCACCCCTCAGAAGCTG
ATTCCCCCGCAGCCAACGGGCCGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCCGCCGCTCGCCCGTGATGCCAC
CGCAGACCCAGTCCCCCGGCAGCCGGCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATCAC
CCCCATCCAGAAGCCGCGGGGCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCATC
GCACACCGAATTACAGGAACCTGAAAACCTTCCCGGTCCCTGGCCGGGGATTGCGAACCAAGCGACCATTGAGC
TCAAGGCCCTCAGGCTGCTGAACTTCCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACAGC
GCTGGAGACAGCCCTCAATGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTGAG
AAGCTGGAGAAGCAGCAGAAGATCGAGCAGGAGCGCAAGCGCCGCGCAGAAGCACCAGGAATACCTCAATAGCATTC
TCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTGGC
CACGTACCATGCCAACACGGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGCTC
ATGGCTGAAGATGAGGAGGGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCAGC
AGACAGACGAGTACGTGGCTAACCTCACGGAGCTGGTGCGGCAGCACAAGGCTGCCAGGTGCGCAAGGAGAAAA
GAAGAAAAAGAAAAAGAAGAAGGCAGAAAATGCAGAAGGACAGACGCCTGCCATTGGGGCCGATGGCGAGCCTCTG
GACGAGACCAGCCAGATGAGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGGAAGATCCTCACAGGCACAG
ATGCCCCCAAAGCCGGGCAGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGATAG
TGAAGAAAGTGGCTCAGAAGAAGAGGAAGAGGAGGAGGAAGAGCAGCCGCAGGCAGCACAGCCTCCACCCTG
CCCGTGGAGGAGAAGAAGAAGATTCCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCATTG
AGAATGCCAAGCAAGATGTCGATGATGAATATGGCGTGTCCAGGCCCTTGACGTGGCCTGCAGTCCTACTATGC
CGTGGCCCATGCTGTCACTGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTTCAAACAGTACCAG
ATCAAAGGTTTGGAGTGGCTGGTGTCCCTGTACAACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCTGG
GGAAGACCATCCAGACCATCGCGCTCATCACGTACCTCATGGAGCACAAACGCATCAATGGGCCCTTCCTCATCAT
CGTGCCTCTCTCAACGCTGTCCAACCTGGGCGTACGAGTTTGACAAGTGGGGCCCCCTCCGTGGTGAAGGTGTCTTAC
AAGGGATCCCCAGCAGCAAGACGGGCCTTTGTCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGACGT
ACGAGTACATCATCAAAGACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTACCG
CATGAAGAACCACCACTGCAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCCGCCGCTGCTGCTGACG

GGCACACCGCTGCAGAACAAAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCCTGCTGCCACCACATCTTCAAGAGCT
GCAGCACCTTCGAGCAGTGGTTTAAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAAAC
CATTCTCATCATCCGGCGTCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGGCC
CAGTTGCCCCAAAAGGTGGAGTACGTATCAAGTGCGACATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACATGC
AGGCCAAGGGCGTGCTGCTGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATGAA
CACCATCATGCAGCTGCGGAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGCAC
TTGGGGTTCACTGGCGGCATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAGAA
TTCTTCCCAAACCTCCGAGCAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATGGA
AGATTACTTTGCGTATCGCGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGCTG
CTGAAAACCTTCAACGAGCCCGGCTCTGAGTACTTCATCTTCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCTGA
ACCTCCAGTCGGCAGACACTGTGATCATTTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGACCG
AGCCCACCGCATCGGGCAGCAGAACGAGGTGCGTGTGCTCCGCCTCTGCACCGTCAACAGCGTGGAGGAGAAGATC
CTAGCTGCAGCCAAGTACAAGCTCAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTCCGACCAGAAGTCCTCCA
GCCATGAGCGGCGCGCCTTCTGCAAGGCCATCCTGGAGCACGAGGAGCAGGATGAGAGCAGACACTGCAGCACGGG
CAGCGGCAGTGCCAGCTTCGCCCACACTGCCCTCCGCCAGCGGGCGTCAACCCCGACTTGAGGAGCCACCTCTA
AAGGAGGAAGACGAGGTGCCCCGACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGT
TCATGCGCATGGACCTGGACCGCAGGCGGAGGAGGCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGA
CGAGCTCCCCTCGTGGATCATCAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATG
TTCGGCCGTGGCTCCCGCCACCGCAAGGAGGTGGACTACAGCGACTCACTGACGGAGAAGCAGTGGCTCAAGAAAA
TTACAGGAAAAGATATCCATGACACAGCCAGCAGTGTGGCACGTGGGCTACAATTCCAGCGTGGCCTTCAGTTCTG
CACACGTGCGTCAAAGGCCATCGAGGAGGGCACGCTGGAGGAGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCA
CGGAAGCGCAAGCGAGACAGCGACCGCGCTCCTCCACCCGACCACCAGCACCCGCAGCCGCGACAAGGACGACG
AGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCCTGCCGAGAACTCTCCCCTAACCCACCCAACCTCACCAAGAA
GATGAAGAAGATTGTGGATGCCGTGATCAAGTACAAGGACAGCAGCAGTGGACGTCAGCTCAGCGAGGTCTTCATC
CAGCTGCCCTCGCGAAAGGAGCTGCCCCGAGTACTACGAGCTCATCCGAAGCCCGTGGACTTCAAGAAGATAAAGG
AGCGCATTCGCAACCACAAGTACCGCAGCCTCAACGACCTAGAGAAGGACGTCATGCTCCTGTGCCAGAACGCACA
GACCTTCAACCTGGAGGGCTCCCTGATCTATGAAGACTCCATCGTCTTGACAGTCGGTCTTACCAGCGTGCGGCAG
AAAATCGAGAAGGAGGATGACAGTGAAGGCGAGGAGAGTGAGGAGGAGGAAGAGGGCGAGGAGGAAGGCTCCGAAT
CCGAATCTCGGTCCGTCAAAGTGAAGATCAAGCTTGCCGGAAGGAGAAGGCACAGGACCGGCTGAAGGGCGGCCG
GCGGCGGCCGAGCCGAGGGTCCCGAGCCAAGCCGGTCTGTAGTGACGATGACAGTGAGGAGGAACAAGAGGAGGAC
CGCTCAGGAAGTGGCAGCGAAGAAGACTGAGCCCCGACATTCCAGTCTCGACCCCGAGCCCCTCGTTCCAGAGCTG
AGATGGCATAGGCCTTAGCAGTAACGGGTAGCAGCAGATGTAGTTTCAGACTTGAGTAAAAGTGTATAAACAAAA
GAATCTTCCATATTTATACAGCAGAGAAGCTGTAGGACTGTTTGTGACTGGCCCTGTCCTGGCATCAGTAGCATCT
GTAACAGCATTAAGTGTCTTAAAGAGAGAGAGAGAATTCCGAATTGGGGAACACACGATACCTGTTTTTCTTTT
CCGTTGCTGGCAGTACTGTTGCGCCGAGTTTGGAGTCACTGTAGTTAAGTGTGGATGCATGTGCGTCACCGTCCA
CTCCTCCTACTGTATTTTATTGGACAGGTCAGACTCGCCGGGGGGCCCGGCGAGGGTATGTCAAGTGTCACTGGATGT
CAAACAGTAATAAATTAACCAACAACAAAACGCACAGCCAAAAA (SEQ ID NO:9)。

[0402] >NM_001128844.1 智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员4 (SMARCA4)，转录物变体2，mRNA

[0403] GGAGAGGCCGCCGCGGTGCTGAGGGGGAGGGGAGCCGGCGAGCGCGCGCAGCGGGGGCGCGGGTGG
CGCGCGTGTGTGTGAAGGGGGGGCGGTGGCCGAGGCGGGCGGGCGCGCGCGAGGCTTCCCCTCGTTTGGCGGCG
GCGGCGGCTTCTTTGTTTCGTGAAGAGAAGCGAGACGCCATTCTGCCCCCGCCCCGCGCGGAGGGGGCGGGGAG
GCGCCGGGAAGTCGACGGCGCCGGCGGCTCCTGCGTCTCGCCCTTTTGCCAGGCTAGAGTGCAGTGGTGCAGTCA
TGGTTCACTGCAGCCTCAACCTCCTGGACTCAGCAGGAGGCCACTGTCTGCAGTCCCGTGAAGATGTCCACTCCA
GACCCACCCCTGGGCGGAACTCCTCGGCCAGGTCCTTCCCCGGGCCCTGGCCCTTCCCCTGGAGCCATGCTGGGCC
CTAGCCCGGGTCCCTCGCCGGGCTCCGCCCACAGCATGATGGGGCCAGCCAGGGCCGCCCTCAGCAGGACACCC
CATCCCCACCCAGGGGCCTGGAGGGTACCCTCAGGACAACATGCACCAGATGCACAAGCCCATGGAGTCCATGCAT
GAGAAGGGCATGTGCGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCGGTGAGGGGGCCATGCTGGGA
TGGGGCCCCCGCCAGCCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCCTGGGTGGCTCTGAGCATGCCTC
TAGTCCAGTTCCAGCCAGTGGCCCGTCTTCGGGGCCCCAGATGTCTTCCGGGCCAGGAGTGCCCCGCTGGATGGT
GCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCAACCCCATTTAACCAGAACCAGCTGCACCAGCTCAGAG
CTCAGATCATGGCCTACAAGATGCTGGCCAGGGGGCAGCCCTCCCCGACCACCTGCAGATGGCGGTGCAGGGCAA
GCGGCCGATGCCCCGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGTCCGCAACAGGACCCGGCCCT
GGCCCTGGCCCTGGCCCCGGCCCGGTCCCGGCCCGGCACCTCCAAATTACAGCAGGCCTCATGGTATGGGAGGGC
CCAACATGCCTCCCCCAGGACCCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCTCCTGGAGGGCCTCCCAAGCC
CTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGCACCCCTCAGAAGCTGATCCCCCGCAGCCAACG
GGCCGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCGCCGCTCGCCCGTGATGCCACCGCAGACCCAGTCCCCCG
GGCAGCCGGCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATCACCCCATCCAGAAGCCGCG
GGGCCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCATCGCACACCGAATTCAGGAA
CTTGAAAACCTTCCCCGGTCCCTGGCCGGGATTTGCGAACCAAGCGACCATTGAGCTCAAGGCCCTCAGGCTGC
TGAACCTCCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACAGCGCTGGAGACAGCCCTCAA
TGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTGAGAAGCTGGAGAAGCAGCAG
AAGATCGAGCAGGAGCGCAAGCGCCGCGAGAAGCACCAGGAATACCTCAATAGCATTCTCCAGCATGCCAAGGATT
TCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTGGCCACGTACCATGCCAACAC
GGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGCTCATGGCTGAAGATGAGGAG
GGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCAGCAGACAGACGAGTACGTGG
CTAACCTCACGGAGCTGGTGCAGGACACAAGGCTGCCAGGTGCGCAAGGAGAAAAAGAAGAAAAAGAAAAAGAA
GAAGGCAGAAAATGCAGAAGGACAGACGCCTGCCATTGGGCCGGATGGCGAGCCTCTGGACGAGACCAGCCAGATG
AGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGGAAGATCCTCACAGGCACAGATGCCCCCAAAGCCGGGC
AGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGATAGTGAAGAAAGTGGCTCAGA
AGAAGAGGAAGAGGAGGAGGAGGAAGAGCAGCCGAGGCAGCACAGCCTCCCACCCTGCCCGTGGAGGAGAAGAAG
AAGATTCCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCATTGAGAATGCCAAGCAAGATG
TCGATGATGAATATGGCGTGTCCCAGGCCCTTGACGTGGCCTGCAGTCTACTATGCCGTGGCCCATGCTGTACAC
TGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTCAAACAGTACCAGATCAAAGGTTTGGAGTGG
CTGGTGTCCCTGTACAACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCTGGGGAAGACCATCCAGACCA
TCGCGCTCATCACGTACCTCATGGAGCACAACGCATCAATGGGCCCTTCCCTCATCATCGTGCCTCTCTCAACGCT
GTCCAACCTGGGCGTACGAGTTTGACAAGTGGGGCCCCCTCCGTGGTGAAGGTGTCTTACAAGGGATCCCCAGCAGCA
AGACGGGCCTTTGTCCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGACGTACGAGTACATCATCAAAG

ACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTCACCGCATGAAGAACCACCACTG
CAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCCGCCGCTGCTGCTGACGGGCACACCGCTGCAGAAC
AAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCCTGCTGCCCACCATCTTCAAGAGCTGCAGCACCTTCGAGCAGT
GGTTTAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAAACCATTCTCATCATCCGGCG
TCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGGCCAGTTGCCGAAAAGGTG
GAGTACGTCATCAAGTGCAGCATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACATGCAGGCCAAGGGCGTGCTGC
TGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATGAACACCATCATGCAGCTGCG
GAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGCACTTGGGGTTCACTGGCGGC
ATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAGAATTCTTCCCAAACCTCCGAG
CAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATGGAAGATTACTTTGCGTATCG
CGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGCTGCTGAAAACCTTCAACGAG
CCCGGCTCTGAGTACTTCATCTTCCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCTGAACCTCCAGTCGGCAGACA
CTGTGATCATTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGACCGAGCCCACCGCATCGGGCA
GCAGAACGAGGTGCGTGTGCTCCGCCTCTGCACCGTCAACAGCGTGAGGAGAAGATCCTAGCTGCAGCCAAGTAC
AAGCTCAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTTCGACCAGAAGTCCTCCAGCCATGAGCGGCGCGCCT
TCCTGCAGGCCATCCTGGAGCACGAGGAGCAGGATGAGAGCAGACACTGCAGCACGGGCAGCGGCAGTGCCAGCTT
CGCCACACTGCCCCCTCCGCCAGCGGGCGTCAACCCGACTTGAGGAGGCCACCTCTAAAGGAGGAAGACGAGGTG
CCCGACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGTTTCATGCGCATGGACCTGG
ACCGCAGGCGCGAGGAGGCCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGACGAGCTCCCCTCGTGGAT
CATCAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATGTTTCGGCCGTGGCTCCCGC
CACCGCAAGGAGGTGGACTACAGCGACTCACTGACGGAGAAGCAGTGGCTCAAGGCCATCGAGGAGGGCAGCTGG
AGGAGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCACGAAGCGCAAGCGAGACAGCGACCGCGGCTCCTCCAC
CCCGACCACCAGCACCCGAGCCGCGACAAGGACGACGAGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCGCTGCC
GAGAACTCTCCCCTAACCCACCCAACCTCACCAAGAAGATGAAGAAGATTGTGGATGCCGTGATCAAGTACAAGG
ACAGCAGCAGTGGACGTCAGCTCAGCGAGGTCTTCATCCAGCTGCCCTCGCGAAAGGAGTGCCCGAGTACTACGA
GCTCATCCGCAAGCCCGTGGACTTCAAGAAGATAAAGGAGCGCATTCGCAACCACAAGTACCGCAGCCTCAACGAC
CTAGAGAAGGACGTCATGCTCCTGTGCCAGAACGCACAGACCTTCAACCTGGAGGGCTCCCTGATCTATGAAGACT
CCATCGTCTTGCACTCGGTCTTACCAGCGTGCGGCAGAAAATCGAGAAGGAGGATGACAGTGAAGGCGAGGAGAG
TGAGGAGGAGGAAGAGGGCGAGGAGGAAGGCTCCGAATCCGAATCTCGGTCCGTCAAAGTGAAGATCAAGCTTGGC
CGGAAGGAGAAGGCACAGGACCGGTGAAGGGCGGCCGCGGCGGCGGAGCCGAGGGTCCCCGAGCCAAGCCGGTGC
TGAGTGACGATGACAGTGAGGAGGAACAAGAGGAGGACCGCTCAGGAAGTGGCAGCGAAGAAGACTGAGCCCCGAC
ATTCCAGTCTCGACCCCGAGCCCCTCGTTCCAGAGCTGAGATGGCATAGGCCTTAGCAGTAACGGGTAGCAGCAGA
TGTAGTTTCAGACTTGGAGTAAACTGTATAAACAAGAATCTTCCATATTTATACAGCAGAGAAGCTGTAGGAC
TGTTTTGTGACTGGCCCTGTCTTGGCATCAGTAGCATCTGTAACAGCATTAAGTGTCTTAAAGAGAGAGAGAGAA
TTCCGAATTGGGGAACACACGATACCTGTTTTTCTTTTCCGTTGCTGGCAGTACTGTTGCGCCGAGTTTGGAGTC
ACTGTAGTTAAGTGTGGATGCATGTGCGTCACCGTCCACTCCTCTACTGTATTTTATTGGACAGGTCAGACTCGC
CGGGGGCCCGGCGAGGGTATGTCAGTGTCACTGGATGTCAAACAGTAATAAATTAAACCAACAACAAAACGCACAG
CCAAAAA (SEQ ID NO:10)。

[0404] >NM_001128845.1智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，

亚家族a,成员4 (SMARCA4),转录物变体4,mRNA

[0405] ATGTCCACTCCAGACCCACCCCTGGGCGGAACCTCCTCGGCCAGGTCCTTCCCCGGGCCCTGGCCCTTC
CCCTGGAGCCATGCTGGGCCCTAGCCCGGTCCCTCGCCGGGCTCCGCCACAGCATGATGGGGCCAGCCAGGG
CCGCCCTCAGCAGGACACCCCATCCCCACCCAGGGGCTGGAGGGTACCCTCAGGACAACATGCACCAGATGCACA
AGCCCATGGAGTCCATGCATGAGAAGGGCATGTGCGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCG
GTCAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCCTG
GGTGGCTCTGAGCATGCCTCTAGTCCAGTTCCAGCCAGTGGCCCGTCTTCGGGGCCCCAGATGTCTTCCGGGCCAG
GAGGTGCCCCGCTGGATGGTGCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCCCAACCCCATTTAACCAGAA
CCAGCTGCACCAGCTCAGAGCTCAGATCATGGCCTACAAGATGCTGGCCAGGGGGCAGCCCCCTCCCCGACCACCTG
CAGATGGCGGTGCAGGGCAAGCGGCCGATGCCCGGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGT
CCGCAACAGGACCCGGCCCTGGCCCTGGCCCCGGCCCGGTCCCGGCCCGGCACCTCCAAATTACAGCAG
GCCTCATGGTATGGGAGGGCCCAACATGCCTCCCCAGGACCCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCT
CCTGGAGGGCTCCCAAGCCCTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGACCCCTCAGAAGC
TGATTCCCCCGCAGCCAACGGGCCGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCCGCCGCTCGCCCGTGATGCC
ACCGCAGACCCAGTCCCCCGGCAGCCGGCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATC
ACCCCATCCAGAAGCCGCGGGGCCCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCA
TCGCACACCGAATTACAGAACTTGAACCTTCCCGGTCCCTGGCCGGGGATTGCGAACCAAGCGACCATTGA
GCTCAAGGCCCTCAGGCTGCTGAACTTCCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACA
GCGCTGGAGACAGCCCTCAATGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTG
AGAAGCTGGAGAAGCAGCAGAAGATCGAGCAGGAGCGCAAGCGCCGCGCAGAAGCACCAGGAATACCTCAATAGCAT
TCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTG
GCCACGTACCATGCCAACACGGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGC
TCATGGCTGAAGATGAGGAGGGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCA
GCAGACAGACGAGTACGTGGCTAACCTCACGGAGCTGGTGCAGCAGCACAAGGCTGCCAGGTCGCCAAGGAGAAA
AAGAAGAAAAAGAAAAAGAAGGCAGAAAATGCAGAAGGACAGACGCTGCCATTGGGCGCGATGGCGAGCCTC
TGGACGAGACCAGCCAGATGAGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGGAAGATCCTCACAGGCAC
AGATGCCCCCAAAGCCGGGCAGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGAT
AGTGAAGAAAGTGGCTCAGAAGAAGAGGAAGAGGAGGAGGAAGAGCAGCCGCAGGCAGCACAGCCTCCACCC
TGCCCGTGGAGGAGAAGAAGAAGATTCCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCAT
TGAGAATGCCAAGCAAGATGTCGATGATGAATATGGCGTGTCCAGGCCCTTGACGTGGCCTGCAGTCCTACTAT
GCCGTGGCCCATGCTGTCACTGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTTCAAACAGTACC
AGATCAAAGGTTTGGAGTGGCTGGTGTCCCTGTACAACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCT
GGGGAAGACCATCCAGACCATCGCGCTCATCACGTACCTCATGGAGCACAACGCATCAATGGGCCCTTCTCATC
ATCGTGCCTCTCTCAACGCTGTCCAACCTGGGCGTACGAGTTTGACAAGTGGGCCCCCTCCGTGGTGAAGGTGTCTT
ACAAGGGATCCCCAGCAGCAAGACGGGCCTTTGTCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGAC
GTACGAGTACATCATAAAGACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTACAC
CGCATGAAGAACCACCACTGCAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCCGCCGCTGCTGCTGA
CGGGCACACCGCTGCAGAACAAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCTGCTGCCACCATCTTCAAGAG
CTGCAGCACCTTCGAGCAGTGGTTTAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAA

ACCATTCTCATCATCCGGCGTCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGG
CCCAGTTGCCCCAAAAGGTGGAGTACGTCATCAAGTGCGACATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACAT
GCAGGCCAAGGGCGTGCTGCTGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATG
AACACCATCATGCAGCTGCGGAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGC
ACTTGGGGTTCACTGGCGGCATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAG
AATTCTTCCCAAACCTCCGAGCAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATG
GAAGATTACTTTGCGTATCGCGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGC
TGCTGAAAACCTTCAACGAGCCCGGCTCTGAGTACTTCATCTTCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCT
GAACCTCCAGTCGGCAGACACTGTGATCATTTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGAC
CGAGCCCACCGCATCGGGCAGCAGAACGAGGTGCGTGTGCTCCGCCTCTGCACCGTCAACAGCGTGGAGGAGAAGA
TCCTAGCTGCAGCCAAGTACAAGCTCAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTCCAGCAGAAGTCCTC
CAGCCATGAGCGGCGCGCCTTCTGCAAGGCCATCCTGGAGCACGAGGAGCAGGATGAGGAGGAAGACGAGGTGCCC
GACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGTTTCATGCGCATGGACCTGGACC
GCAGGCGCGAGGAGGCCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGACGAGCTCCCCCTCGTGGATCAT
CAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATGTTCCGGCCGTGGCTCCCCGCCAC
CGCAAGGAGGTGGACTACAGCGACTCACTGACGGAGAAGCAGTGCGTCAAGACCCTGAAGGCCATCGAGGAGGGCA
CGCTGGAGGAGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCACGGAAGCGCAAGCGAGACAGCGACGCCGGCTC
CTCCACCCCGACCACCAGCACCCGCAGCCGCGACAAGGACGACGAGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCG
CCTGCCGAGAACTCTCCCCTAACCACCCAACCTCACCAAGAAGATGAAGAAGATTGTGGATGCCGTGATCAAGT
ACAAGGACAGCAGCAGTGGACGTCAGCTCAGCGAGGTCTTCATCCAGCTGCCCTCGCGAAAGGAGCTGCCCGAGTA
CTACGAGCTCATCCGCAAGCCCCTGGACTTCAAGAAGATAAAGGAGCGCATTCGCAACCACAAGTACCGCAGCCTC
AACGACCTAGAGAAGGACGTCATGCTCCTGTGCCAGAACGCACAGACCTTCAACCTGGAGGGCTCCCTGATCTATG
AAGACTCCATCGTCTTGCAGTCGGTCTTACCAGCGTGCGGCAGAAAATCGAGAAGGAGGATGACAGTGAAGGCGA
GGAGAGTGAGGAGGAGGAAGAGGGCGAGGAGGAAGGCTCCGAATCCGAATCTCGGTCCGTCAAAGTGAAGATCAAG
CTTGCCCGGAAGGAGAAGGCACAGGACCGGCTGAAGGGCGGCCGGCGGCCGAGCCGAGGGTCCCAGCCAAGC
CGGTGCTGAGTGACGATGACAGTGAGGAGGAACAAGAGGAGGACCGCTCAGGAAGTGGCAGCGAAGAAGACTGAGC
CCCACATTCCAGTCTCGACCCCGAGCCCCTCGTTCCAGAGCTGAGATGGCATAGGCCCTTAGCAGTAACGGGTAGC
AGCAGATGTAGTTTCAGACTTGAGTAAACTGTATAAACAAAAGAATCTTCCATATTTATACAGCAGAGAAGCTG
TAGGACTGTTTGTGACTGGCCCTGTCCTGGCATCAGTAGCATCTGTAACAGCATTAAGTGTCTTAAAGAGAGAGAG
AGAGAATTCCGAATTGGGGAACACACGATACCTGTTTTTCTTTTCCGTTGCTGGCAGTACTGTTGCGCCGAGTTT
GGAGTCACTGTAGTTAAGTGTGGATGCATGTGCGTCACCGTCCACTCCTCCTACTGTATTTTATTGGACAGGTCAG
ACTCGCCGGGGGCCCGGCGAGGGTATGTCAGTGTCACTGGATGTCAAACAGTAATAAATTAACCAACAACAAAAC
GCACAGCCAAAAAAAAA (SEQ ID NO:11)。

[0406] >NM_001128846.1 智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员4 (SMARCA4)，转录物变体5，mRNA

[0407] ATGTCCACTCCAGACCCACCCCTGGGCGAACTCCTCGGCCAGGTCCTTCCCCGGGCCCTGGCCCTTC
CCCTGGAGCCATGCTGGGCCCTAGCCCGGCTCCCTCGCCGGGCTCCGCCACAGCATGATGGGGCCCAGCCAGGG
CCGCCCTCAGCAGGACACCCCATCCCCACCCAGGGGCTGGAGGGTACCCTCAGGACAACATGCACCAGATGCACA
AGCCCATGGAGTCCATGCATGAGAAGGGCATGTCGGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCG

GTCAGGGGGCCATGCTGGGATGGGGCCCCCGCCCAGCCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCTG
GGTGGCTCTGAGCATGCCTCTAGTCCAGTTCAGCCAGTGGCCCGTCTTCGGGGCCCCAGATGTCTTCGGGGCCAG
GAGGTGCCCCGCTGGATGGTGCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCCAACCCCATTTAACCAGAA
CCAGCTGCACCAGCTCAGAGCTCAGATCATGGCCTACAAGATGCTGGCCAGGGGGCAGCCCCCTCCCCGACCACCTG
CAGATGGCGGTGCAGGGCAAGCGGCCGATGCCCGGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGT
CCGCAACAGGACCCGGCCCTGGCCCTGGCCCTGGCCCCGGCCCGGTCCCGGCCCGGCACCTCCAAATTACAGCAG
GCCTCATGGTATGGGAGGGCCCAACATGCCTCCCCAGGACCCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCT
CCTGGAGGGCCTCCCAAGCCCTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGCACCCCTCAGAAGC
TGATTCCCCCGCAGCCAACGGGGCCGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCCGCCCTCGCCCGTGATGCC
ACCGCAGACCCAGTCCCCCGGGCAGCCGGCCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATC
ACCCCATCCAGAAGCCGCGGGGCCCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCA
TCGCACACCGAATTCAGGAACTTGAAAACCTTCCCGGTCCCTGGCCGGGGATTGCGAACCAGCGACCATTGA
GCTCAAGGCCCTCAGGCTGCTGAACTTCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACA
GCGCTGGAGACAGCCCTCAATGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTG
AGAAGCTGGAGAAGCAGCAGAAGATCGAGCAGGAGCGCAAGCGCCGGCAGAAGCACCAGGAATACCTCAATAGCAT
TCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTG
GCCACGTACCATGCCAACACGGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGC
TCATGGCTGAAGATGAGGAGGGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCA
GCAGACAGACGAGTACGTGGCTAACCTCACGGAGCTGGTGCAGCAGCACAAGGCTGCCAGGTGCCAAGGAGAAA
AAGAAGAAAAAGAAAAAGAAAGGCAGAAAATGCAGAAGGACAGACGCCTGCCATTGGGCCGGATGGCGAGCCTC
TGGACGAGACCAGCCAGATGAGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGGAAGATCCTCACAGGCAC
AGATGCCCCCAAAGCCGGGCAGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGAT
AGTGAAGAAAGTGGCTCAGAAGAAGAGGAAGAGGAGGAGGAAGAGCAGCCGCAGGCAGCACAGCCTCCCACCC
TGCCCGTGGAGGAGAAGAAGAAGATTCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCAT
TGAGAATGCCAAGCAAGATGTCGATGATGAATATGGCGTGTCAGGCCCTTGACGTGGCCTGCAGTCCTACTAT
GCCGTGGCCCATGCTGTCACTGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTCAAACAGTACC
AGATCAAAGGTTTGGAGTGGCTGGTGTCCCTGTACAACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCT
GGGGAAGACCATCCAGACCATCGCGCTCATCACGTACCTCATGGAGCACAACGCATCAATGGGCCCTTCCTCATC
ATCGTGCCTCTCTCAACGCTGTCCAACCTGGGCGTACGAGTTTGACAAGTGGGCCCCCTCCGTGGTGAAGGTGTCTT
ACAAGGGATCCCCAGCAGCAAGACGGGCCTTTGTCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGAC
GTACGAGTACATCATCAAAGACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTCAC
CGCATGAAGAACCACCACTGCAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCCGCCGCTGCTGCTGA
CGGGCACACCGCTGCAGAACAAAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCCTGCTGCCCACCATCTTCAAGAG
CTGCAGCACCTTCGAGCAGTGGTTTAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAA
ACCATTCTCATCATCCGGCGTCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGG
CCCAGTTGCCCCAAAAGGTGGAGTACGTCATCAAGTGCGACATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACAT
GCAGGCCAAGGGCGTGCTGCTGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATG
AACACCATCATGCAGCTGCGGAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGC
ACTTGGGGTTCACTGGCGGCATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAG

AATTCTTCCCAAACCTCCGAGCAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATG
GAAGATTACTTTGCGTATCGCGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGC
TGCTGAAAACCTTCAACGAGCCCGGCTCTGAGTACTTCATCTTCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCT
GAACCTCCAGTCGGCAGACACTGTGATCATTTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGAC
CGAGCCCACCGCATCGGGCAGCAGAACGAGGTGCGTGTGCTCCGCTCTGCACCGTCAACAGCGTGGAGGAGAAGA
TCCTAGCTGCAGCCAAGTACAAGCTCAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTGACCCAGAAGTCCTC
CAGCCATGAGCGGCGCGCTTCTGTCAGGCCATCCTGGAGCACGAGGAGCAGGATGAGGAGGAAGACGAGGTGCCC
GACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGTTCATGCGCATGGACCTGGACC
GCAGGCGCGAGGAGGCCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGACGAGCTCCCTCGTGGATCAT
CAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATGTTGCGCCGTGGCTCCCGCCAC
CGCAAGGAGGTGGACTACAGCGACTACTGACGGAGAAGCAGTGCGTCAAGACCCTGAAGGCCATCGAGGAGGGCA
CGCTGGAGGAGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCACGGAAGCGCAAGCGAGACAGCGACGCCGGCTC
CTCCACCCCGACCACCAGCACCCGCAGCCGCGACAAGGACGACGAGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCG
CCTGCCGAGAACTCTCCCCTAACCCACCCAACCTCACCAAGAAGATGAAGAAGATTGTGGATGCCGTGATCAAGT
ACAAGGACAGCAGTGACGTCAGCTCAGCGAGGTCTTCATCCAGCTGCCCTCGCGAAAGGAGCTGCCCGAGTACTA
CGAGCTCATCCGAAGCCCGTGGACTTCAAGAAGATAAAGGAGCGCATTTCGAACCACAAGTACCGCAGCCTCAAC
GACCTAGAGAAGGACGTCATGCTCCTGTGCCAGAACGCACAGACCTTCAACCTGGAGGGCTCCCTGATCTATGAAG
ACTCCATCGTCTTGACGTCGGTCTTACCAGCGTGCGGCAGAAAATCGAGAAGGAGGATGACAGTGAAGGCGAGGA
GAGTGAGGAGGAGGAAGAGGGCGAGGAGGAAGGCTCCGAATCCGAATCTCGGTCCGTCAAAGTGAAGATCAAGCTT
GGCCGGAAGGAGAAGGCACAGGACCGGCTGAAGGGCGGCCGCGCGGCCGAGCCGAGGGTCCCGAGCCAAGCCGG
TCGTGAGTGACGATGACAGTGAGGAGGAACAAGAGGAGGACCGCTCAGGAAGTGGCAGCGAAGAAGACTGAGCCCC
GACATTCCAGTCTCGACCCCGAGCCCCTCGTTCCAGAGCTGAGATGGCATAGGCCTTAGCAGTAACGGGTAGCAGC
AGATGTAGTTTCAGACTTGGAGTAAACTGTATAAAACAAAAGAATCTTCCATATTTATACAGCAGAGAAGCTGTAG
GACTGTTTGTGACTGGCCCTGTCCTGGCATCAGTAGCATCTGTAACAGCATTAAGTGTCTTAAAGAGAGAGAGAGA
GAATTCGAATTGGGGAACACACGATACCTGTTTTTCTTTTCCGTTGCTGGCAGTACTGTTGCGCCGAGTTTGA
GTCAGTGTAGTTAAGTGTGGATGCATGTGCGTCACCGTCCACTCCTCCTACTGTATTTTATTGGACAGGTCAGACT
CGCCGGGGGCCCCGGCGAGGGTATGTCAGTGTCACTGGATGTCAAACAGTAATAAATTAAACCAACAACAAAACGCA
CAGCCAAAAA (SEQ ID NO:12)。

[0408] >NM_001128847.1智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员4 (SMARCA4)，转录物变体6，mRNA

[0409] ATGTCCACTCCAGACCCACCCCTGGGCGGAACCTCCTCGGCCAGGTCCTTCCCCGGGCCCTGGCCCTTC
CCCTGGAGCCATGCTGGGCCCTAGCCCGGTCCCTCGCCGGGCTCCGCCACAGCATGATGGGGCCAGCCAGGG
CCGCCCTCAGCAGGACACCCCATCCCCACCCAGGGGCTGGAGGGTACCCTCAGGACAACATGCACCAGATGCACA
AGCCCATGGAGTCCATGCATGAGAAGGGCATGTGCGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCG
GTCAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCTG
GGTGGCTCTGAGCATGCCTCTAGTCCAGTTCCAGCCAGTGCCCCGTCTTCGGGGCCCCAGATGTCTTCCGGGCCAG
GAGGTGCCCCGCTGGATGGTGCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCCAACCCATTTAACCAGAA
CCAGCTGCACCAGCTCAGAGCTCAGATCATGGCCTACAAGATGCTGGCCAGGGGGCAGCCCCTCCCCGACCACCTG
CAGATGGCGGTGCAGGGCAAGCGGCCGATGCCCGGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGT

CCGCAACAGGACCCGGCCCTGGCCCTGGCCCTGGCCCCGGCCCGGGTCCCGGCCCGGCACCTCCAAATTACAGCAG
GCCTCATGGTATGGGAGGGCCCAACATGCCTCCCCAGGACCCCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCT
CCTGGAGGGCTCCCAAGCCCTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGCACCCCTCAGAAGC
TGATTCCCCCGCAGCCAACGGGCGGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCCGCGCCTCGCCCGTGATGCC
ACCGCAGACCCAGTCCCCCGGGCAGCCGGCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATC
ACCCCCATCCAGAAGCCGCGGGGCCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCA
TCGCACACCGAATTCAAGAACTTGAACCTTCCCGGGTCCCTGGCCGGGGATTGCGAACCAGCGACCATTTGA
GCTCAAGGCCCTCAGGCTGCTGAACTTCCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACA
GCGCTGGAGACAGCCCTCAATGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTG
AGAAGCTGGAGAAGCAGCAGAAGATCGAGCAGGAGCGCAAGCGCCGGCAGAAGCACCAGGAATACCTCAATAGCAT
TCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTG
GCCACGTACCATGCCAACACGGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGC
TCATGGCTGAAGATGAGGAGGGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCA
GCAGACAGACGAGTACGTGGCTAACCTCACGGAGCTGGTGCGGCAGCACAAGGCTGCCAGGTCGCCAAGGAGAAA
AAGAAGAAAAAGAAAAAGAAGGCAGAAAATGCAGAAGGACAGACGCTGCCATTGGGCGGATGGCGAGCCTC
TGGACGAGACCAGCCAGATGAGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGGAAGATCCTCACAGGCAC
AGATGCCCCCAAAGCCGGGCAGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGAT
AGTGAAGAAAGTGGCTCAGAAGAAGAGGAAGAGGAGGAGGAAGAGCAGCCGCAGGCAGCACAGCCTCCCACCC
TGCCCGTGGAGGAGAAGAAGATTCCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCAT
TGAGAATGCCAAGCAAGATGTCGATGATGAATATGGCGTGTCCCAGGCCCTTGACGTGGCCTGCAGTCCTACTAT
GCCGTGGCCCATGCTGTCACTGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTCAAACAGTACC
AGATCAAAGGTTTGGAGTGGCTGGTGTCCCTGTACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCT
GGGGAAGACCATCCAGACCATCGCGCTCATCACGTACCTCATGGAGCACAACGCATCAATGGGCCCTTCTCATC
ATCGTGCCTCTCTCAACGCTGTCCAACCTGGGCGTACGAGTTTGACAAGTGGGCCCCCTCCGTGGTGAAGGTGTCTT
ACAAGGGATCCCCAGCAGCAAGACGGGCCTTTGTCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGAC
GTACGAGTACATCATCAAAGACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTCAC
CGCATGAAGAACCACCACTGCAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCTCGCCCTGCTGCTGA
CGGGCACACCGCTGCAGAACAAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCTGCTGCCCACCATCTTCAAGAG
CTGCAGCACCTTCGAGCAGTGGTTTAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAA
ACCATTCTCATCATCCGGCGTCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGG
CCCAGTTGCCCCAAAAGGTGGAGTACGTCATCAAGTGCGACATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACAT
GCAGGCCAAGGGCGTGCTGCTGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATG
AACACCATCATGCAGCTGCGGAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGC
ACTTGGGGTTCACTGGCGGCATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAG
AATTCTTCCCAAACCTCCGAGCAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATG
GAAGATTACTTTGCGTATCGCGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGC
TGCTGAAAACCTTCAACGAGCCCGGCTCTGAGTACTTCATCTTCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCT
GAACCTCCAGTCGGCAGACACTGTGATCATTTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGAC
CGAGCCCACCGCATCGGGCAGCAGAACGAGGTGCGTGTGCTCCGCCTCTGCACCGTCAACAGCGTGGAGGAGAAGA

TCCTAGCTGCAGCCAAGTACAAGCTCAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTCCGACCAGAAGTCCTC
CAGCCATGAGCGGCGCGCCTTCCTGCAGGCCATCCTGGAGCACGAGGAGCAGGATGAGGAGGAAGACGAGGTGCCC
GACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGTTTCATGCGCATGGACCTGGACC
GCAGGCGCGAGGAGGCCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGACGAGCTCCCCTCGTGGATCAT
CAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATGTTCCGGCCGTGGCTCCCGCCAC
CGCAAGGAGGTGGACTACAGCGACTCACTGACGGAGAAGCAGTGGCTCAAGGCCATCGAGGAGGGCACGCTGGAGG
AGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCACGGAAGCGCAAGCGAGACAGCGACGCCGGCTCCTCCACCCC
GACCACCAGCACCCGCAGCCGCGACAAGGACGACGAGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCCGCTGCCGAG
AAACTCTCCCCTAACCCACCCAACTCACCAGAAGATGAAGAAGATTGTGGATGCCGTGATCAAGTACAAGGACA
GCAGCAGTGGACGTCAGCTCAGCGAGGTCTTCATCCAGCTGCCCTCGCGAAAGGAGCTGCCCCAGTACTACGAGCT
CATCCGCAAGCCCGTGGACTTCAAGAAGATAAAGGAGCGCATTCGCAACCACAAGTACCGCAGCCTCAACGACCTA
GAGAAGGACGTCATGCTCCTGTGCCAGAACGCACAGACCTTCAACCTGGAGGGCTCCCTGATCTATGAAGACTCCA
TCGTCTTGACGTCGGTCTTACCAGCGTGGCGCAGAAAATCGAGAAGGAGGATGACAGTGAAGGCGAGGAGAGTGA
GGAGGAGGAAGAGGGCGAGGAGGAAGGCTCCGAATCCGAATCTCGGTCCGTCAAAGTGAAGATCAAGCTTGGCCGG
AAGGAGAAGGCACAGGACCGGCTGAAGGGCGGCCGGCGGCCGAGCCGAGGGTCCCGAGCCAAGCCGGTCTGTGA
GTGACGATGACAGTGAGGAGGAACAAGAGGAGGACCGCTCAGGAAGTGGCAGCGAAGAAGACTGAGCCCCGACATT
CCAGTCTCGACCCCGAGCCCCCTCGTTCCAGAGCTGAGATGGCATAGGCCTTAGCAGTAACGGGTAGCAGCAGATGT
AGTTTCAGACTTGGAGTAAACTGTATAAACAAAAGAATCTTCCATATTTATACAGCAGAGAAGCTGTAGGACTGT
TTGTGACTGGCCCTGTCTTGGCATCAGTAGCATCTGTAACAGCATTAACTGTCTTAAAGAGAGAGAGAGAATTC
CGAATTGGGGAACACACGATACCTGTTTTCTTTCCGTTGCTGGCAGTACTGTTGCGCCGAGTTTGGAGTCACT
GTAGTTAAGTGTGGATGCATGTGCGTCACCGTCCACTCCTCTACTGTATTTTATTGGACAGGTCAGACTCGCCGG
GGGCCCCGGCAGGGTATGTCAGTGTCACTGGATGTCAAACAGTAATAAATTAAACCAACAACAAAACGCACAGCCA
AAAAAAA (SEQ ID NO:13)。

[0410] >NM_001128848.1智人SWI/SNF相关、基质相关、肌动蛋白依赖性的染色质调控子，
亚家族a，成员4 (SMARCA4)，转录物变体7，mRNA

[0411] ATGTCCACTCCAGACCCACCCCTGGGCGGAACCTCCTCGGCCAGGTCCTTCCCCGGGCCCTGGCCCTTC
CCCTGGAGCCATGCTGGGCCCTAGCCCGGTCCCTCGCCGGGCTCCGCCACAGCATGATGGGGCCAGCCAGGG
CCGCCCTCAGCAGGACACCCCATCCCCACCCAGGGGCTGGAGGGTACCTCAGGACAACATGCACCAGATGCACA
AGCCCATGGAGTCCATGCATGAGAAGGGCATGTGCGACGACCCGCGCTACAACCAGATGAAAGGAATGGGGATGCG
GTCAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCATGGACCAGCACTCCCAAGGTTACCCCTCGCCCCTG
GGTGGCTCTGAGCATGCCTCTAGTCCAGTTCCAGCCAGTGGCCCGTCTTCGGGGCCCCAGATGTCTTCCGGGCCAG
GAGGTGCCCCGCTGGATGGTGCTGACCCCCAGGCCTTGGGGCAGCAGAACCGGGGCCCAACCCATTTAACCAGAA
CCAGCTGCACCAGCTCAGAGCTCAGATCATGGCCTACAAGATGCTGGCCAGGGGGCAGCCCCCTCCCGACCACCTG
CAGATGGCGGTGCAGGGCAAGCGGCCGATGCCCGGATGCAGCAGCAGATGCCAACGCTACCTCCACCCTCGGTGT
CCGCAACAGGACCCGGCCCTGGCCCTGGCCCGGCCCGGGTCCCGGCCCGGCACCTCCAAATTACAGCAG
GCCTCATGGTATGGGAGGGCCCAACATGCCTCCCCAGGACCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCCT
CCTGGAGGGCCTCCCAAGCCCTGGCCTGAAGGACCCATGGCGAATGCTGCTGCCCCACGAGCACCCCTCAGAAGC
TGATTCCCCCGCAGCCAACGGGCCGCCCTTCCCCCGCGCCCCCTGCCGTCCCACCCGCCGCTCGCCCGTGATGCC
ACCGCAGACCCAGTCCCCCGGCAGCCGGCCAGCCCGCGCCCATGGTGCCACTGCACCAGAAGCAGAGCCGCATC

ACCCCCATCCAGAAGCCGCGGGGCCTCGACCCTGTGGAGATCCTGCAGGAGCGCGAGTACAGGCTGCAGGCTCGCA
TCGCACACCGAATTTCAGGAACCTGAAAACCTTCCCGGGTCCCTGGCCGGGGATTGCGAACCAGCGACCATTGA
GCTCAAGGCCCTCAGGCTGCTGAACTTCCAGAGGCAGCTGCGCCAGGAGGTGGTGGTGTGCATGCGGAGGGACACA
GCGCTGGAGACAGCCCTCAATGCTAAGGCCTACAAGCGCAGCAAGCGCCAGTCCCTGCGCGAGGCCCGCATCACTG
AGAAGCTGGAGAAGCAGCAGAAGATCGAGCAGGAGCGCAAGCGCCGCGCAGAAGCACCAGGAATACCTCAATAGCAT
TCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTACAGGCAAAATCCAGAAGCTGACCAAGGCAGTG
GCCACGTACCATGCCAACACGGAGCGGGAGCAGAAGAAAGAGAACGAGCGGATCGAGAAGGAGCGCATGCGGAGGC
TCATGGCTGAAGATGAGGAGGGGTACCGCAAGCTCATCGACCAGAAGAAGGACAAGCGCCTGGCCTACCTCTTGCA
GCAGACAGACGAGTACGTGGCTAACCTCACGGAGCTGGTGCAGCAGCACAAGGCTGCCAGGTCGCCAAGGAGAAA
AAGAAGAAAAAGAAAAAGAAGGCAGAAAATGCAGAAGGACAGACGCCTGCCATTGGGCCGGATGGCGAGCCTC
TGGACGAGACCAGCCAGATGAGCGACCTCCCGGTGAAGGTGATCCACGTGGAGAGTGGAAGATCCTCACAGGCAC
AGATGCCCCCAAAGCCGGGCAGCTGGAGGCCTGGCTCGAGATGAACCCGGGGTATGAAGTAGCTCCGAGGTCTGAT
AGTGAAGAAAGTGGCTCAGAAGAAGAGGAAGAGGAGGAGGAAGAGCAGCCGCAGGCAGCACAGCCTCCACCC
TGCCCGTGGAGGAGAAGAAGAAGATTCCAGATCCAGACAGCGATGACGTCTCTGAGGTGGACGCGCGGCACATCAT
TGAGAATGCCAAGCAAGATGTCGATGATGAATATGGCGTGTCCAGGCCCTTGACGTGGCCTGCAGTCCTACTAT
GCCGTGGCCCATGCTGTCACTGAGAGAGTGGACAAGCAGTCAGCGCTTATGGTCAATGGTGTCTCAAACAGTACC
AGATCAAAGGTTTGGAGTGGCTGGTGTCCCTGTACAACAACAACCTGAACGGCATCCTGGCCGACGAGATGGGCCT
GGGGAAGACCATCCAGACCATCGCGCTCATCACGTACCTCATGGAGCACAACGCATCAATGGGCCCTTCTCATC
ATCGTGCCTCTCTCAACGCTGTCCAACCTGGGCGTACGAGTTTGACAAGTGGGCCCCCTCCGTGGTGAAGGTGTCTT
ACAAGGGATCCCCAGCAGCAAGACGGGCCTTTGTCCCCAGCTCCGGAGTGGGAAGTTCAACGTCTTGCTGACGAC
GTACGAGTACATCATAAAGACAAGCACATCCTCGCCAAGATCCGTTGGAAGTACATGATTGTGGACGAAGGTCAC
CGCATGAAGAACCACCACTGCAAGCTGACGCAGGTGCTCAACACGCACTATGTGGCACCCCGCCGCTGCTGCTGA
CGGGCACACCGCTGCAGAACAAGCTTCCCGAGCTCTGGGCGCTGCTCAACTTCTGCTGCCACCATCTTCAAGAG
CTGCAGCACCTTCGAGCAGTGGTTTAACGCACCCTTTGCCATGACCGGGGAAAAGGTGGACCTGAATGAGGAGGAA
ACCATTCTCATCATCCGGCGTCTCCACAAAGTGCTGCGGCCCTTCTTGCTCCGACGACTCAAGAAGGAAGTCGAGG
CCCAGTTGCCCCAAAAGGTGGAGTACGTATCAAGTGCGACATGTCTGCGCTGCAGCGAGTGCTCTACCGCCACAT
GCAGGCCAAGGGCGTGCTGCTGACTGATGGCTCCGAGAAGGACAAGAAGGGCAAAGGCGGCACCAAGACCCTGATG
AACACCATCATGCAGCTGCGGAAGATCTGCAACCACCCCTACATGTTCCAGCACATCGAGGAGTCCTTTTCCGAGC
ACTTGGGGTTCACTGGCGGCATTGTCCAAGGGCTGGACCTGTACCGAGCCTCGGGTAAATTTGAGCTTCTTGATAG
AATTCTTCCCAAACCTCCGAGCAACCAACCACAAAGTGCTGCTGTTCTGCCAAATGACCTCCCTCATGACCATCATG
GAAGATTACTTTGCGTATCGCGGCTTTAAATACCTCAGGCTTGATGGAACCACGAAGGCGGAGGACCGGGGCATGC
TGCTGAAAACCTTCAACGAGCCCGGCTCTGAGTACTTCATCTTCTGCTCAGCACCCGGGCTGGGGGGCTCGGCCT
GAACCTCCAGTCGGCAGACACTGTGATCATTTTTTGACAGCGACTGGAATCCTCACCAGGACCTGCAAGCGCAGGAC
CGAGCCCACCGCATCGGGCAGCAGAACGAGGTGCGTGTGCTCCGCTCTGCACCGTCAACAGCGTGGAGGAGAAGA
TCCTAGCTGCAGCCAAGTACAAGCTAACGTGGACCAGAAGGTGATCCAGGCCGGCATGTTGACCAGAAGTCCTC
CAGCCATGAGCGGCGCGCCTTCTGAGGCCATCCTGGAGCACGAGGAGCAGGATGAGGAGGAAGACGAGGTGCCC
GACGACGAGACCGTCAACCAGATGATCGCCCGGCACGAGGAGGAGTTTGATCTGTTTCATGCGCATGGACCTGGACC
GCAGGCGCGAGGAGGCCCGCAACCCCAAGCGGAAGCCGCGCCTCATGGAGGAGGACGAGCTCCCTCGTGGATCAT
CAAGGACGACGCGGAGGTGGAGCGGCTGACCTGTGAGGAGGAGGAGGAGAAGATGTTGCGCCGTGGCTCCCGCCAC

CGCAAGGAGGTGGACTACAGCGACTACTGACGGAGAAGCAGTGGCTCAAGGCCATCGAGGAGGGCACGCTGGAGG
AGATCGAAGAGGAGGTCCGGCAGAAGAAATCATCACGGAAGCGCAAGCGAGACAGCGACGCCGGCTCCTCCACCCC
GACCACCAGCACCCCGCAGCCGCGACAAGGACGACGAGAGCAAGAAGCAGAAGAAGCGCGGGCGGCCGCTGCCGAG
AAACTCTCCCCTAACCCACCCAACCTCACCAAGAAGATGAAGAAGATTGTGGATGCCGTGATCAAGTACAAGGACA
GCAGTGGACGTCAGCTCAGCGAGGTCTTCATCCAGCTGCCCTCGCGAAAGGAGCTGCCCGAGTACTACGAGCTCAT
CCGCAAGCCCGTGGACTTCAAGAAGATAAAGGAGCGCATTCGCAACCACAAGTACCGCAGCCTCAACGACCTAGAG
AAGGACGTCATGCTCCTGTGCCAGAACGCACAGACCTTCAACCTGGAGGGCTCCCTGATCTATGAAGACTCCATCG
TCTTGCAGTCGGTCTTCACCAGCGTGCAGCAGAAAATCGAGAAGGAGGATGACAGTGAAGGCGAGGAGAGTGAGGA
GGAGGAAGAGGGCGAGGAGGAAGGCTCCGAATCCGAATCTCGGTCCGTCAAAGTGAAGATCAAGCTTGGCCGGAAG
GAGAAGGCACAGGACCGGTGAAGGGCGGCCGCGCGGCCGAGCCGAGGGTCCCGAGCCAAGCCGGTCGTGAGTG
ACGATGACAGTGAGGAGGAACAAGAGGAGGACCGCTCAGGAAGTGGCAGCGAAGAAGACTGAGCCCCGACATTCCA
GTCTCGACCCCGAGCCCCTCGTTCCAGAGCTGAGATGGCATAGGCCTTAGCAGTAACGGGTAGCAGCAGATGTAGT
TTCAGACTTGAGTAAAACTGTATAAAACAAAAGAATCTTCCATATTTATACAGCAGAGAAGCTGTAGGACTGTTTG
TGACTGGCCCTGTCTTGGCATCAGTAGCATCTGTAACAGCATTAACTGTCTTAAAGAGAGAGAGAGAGAATTCCGA
ATTGGGGAACACACGATACCTGTTTTTCTTTTCCGTTGCTGGCAGTACTGTTGCGCCGCAGTTTGGAGTCACTGTA
GTTAAGTGTGGATGCATGTGCGTCACCGTCCACTCCTCTACTGTATTTTATTGGACAGGTCAGACTCGCCGGGGG
CCCGGCGAGGGTATGTCAGTGTCACTGGATGTCAAACAGTAATAAATTAACCAACAACAAAACGCACAGCCAAAA
AAAAA (SEQ ID NO:14)。

[0412] >NP_001122321.1转录激活因子BRG1同种型A[智人]

[0413] MSTPDPLGGTPRPGSPGPGSPGAMLGSPGPGSPGSAHSMMPSPGPPSAGHP IPTQGGPGYPQDN
MHQMHKPMESMHEKMSDDPRYNQMKGMGRSGGHAGMGPPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHQLRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWEGPMANAAAPT
STPQKLIPPQTGRPSAPPAPVPPAASPVMPPTQSPGQPAQPAPMVPLHQKQSRTPIQKPRGLDPVEILQEREY
RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLLNFQRQLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQKQIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQLTKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKLIDQKKDKRLAYLLQQTDEYVANLTELVRQHKAQVAKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEQPQA
AQPPTLPVEKKKIPDPDSDDVSEVDARHI IENAKQDVDDEYGVSAQALARGLQSYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVS LYNNLNGLADEMGLGKTIQTIALITYLMEHKRINGPFLIIVPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLLTTEYI I IKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVLNTHYVAP
RRLLLTGTP LQNKLP ELWALLNFLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEETIL IIRRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAQGVLLTDGSEKDKKGKGGTKLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRI LPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLGLNLQSADTVI IFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAKYKLNVDQKVIQAGMFDQSSSHERRAFLQAILEHEEQDESRHCSTGSGSASFHTAPPPAGVNP
LEEPPLKEEDEVPPDETQMIARHEEEFDLFMRMDLDRRREEARNPKRKPRLMEEDELPSWI IKDDAEVERLTCE
EEEEKMFGRGSRHRKEVDYSDSLTEKQWLKKITGKDIHDTASSVARGLQFQRGLQFCTRASKAIEEGTLEEIEEEV
RQKKSSRKRKRDS DAGSSTPTTSTRSRDKDDESKQKKRGRPPAEKLSNPPNLTCKMKKIVDAVIKYKDS SSGRQ

LSEVFIQLPSRKELPEYYELIRKPVDFKKIKERIRNHKYRSLNDLEKDVMLLCQNAQTFNLEGLSIYEDSIVLQSV
FTSVRQKIEKEDDSEGESEEEEEEEEESESESRSVKVKIKLGRKEKAQDRLKGGRRRPSRGSRAKPVVSDDDSE
EEQEEDRSGSGSEED (SEQ ID NO:15)。

[0414] >NP_001122316.1转录激活因子BRG1同种型B[智人]

[0415] MSTDPPLGGTPRPGSPGPGSPGAMLGSPGSPGSAHSMGSPGPPSAGHIPITQGPGGYPDN
MHQMHKPMESMHEKGMSSDDPRYNQMKGMGRSGGHAGMGPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHLQRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQMQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWPEGPMANAAAPT
STPQKLIPPQPTGRPSPAPPVPPAASVMPPTQTSQGQAQAPAMVPLHQKQSRITPIQKPRGLDPVEILQEREY
RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLLNFQRLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQQKIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQKLTAKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKIDQKKDKRLAYLLQQTDEYVANLTELVRQHAAQVAKEKKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEQQA
AQPPTLPVEEKKKIPDPDSDDVSEVDARHI IENAKQDVDEYGVSAQALARGLSYYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVS LYNNNLNGILADEMGLGKTIQTIALITYLMEHKRINGPFLIIVPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLTTYEYIIKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVLNTHYVAP
RRLLLTGTPQLNKLPELWALLNLLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEETILIRRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAKGVLLTDGSEKDKKGKGGTKTLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRLPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLGLNLQSADTVIIFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAKYKLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEQDESRHCSTGSGSASFAHTAPPPAGVNP
LEEPPLKEEDEVPDDETQMIARHEEEFDLFMRMDLDRRREARNPKRKPRLMEDELPWI IKDDAEVERLTCE
EEEEKMFGRGSRHRKEVDYSDSLTEKQWLKAIIEGTLEEIEEEVRQKKSSRKRKRDSDAGSSTPTTSTRSRDKDE
SKKQKKRGRPPAEKLSNPNNLTKKMKKIVDAVIKYKSSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKE
RIRNHKYRSLNDLEKDVMLLCQNAQTFNLEGLSIYEDSIVLQSVFTSVRQKIEKEDDSEGESEEEEEEEEESESE
ESRSVKVKIKLGRKEKAQDRLKGGRRRPSRGSRAKPVVSDDDSEEEQEEDRSGSGSEED (SEQ ID NO:16)。

[0416] >NP_001122317.1转录激活因子BRG1同种型C[智人]

[0417] MSTDPPLGGTPRPGSPGPGSPGAMLGSPGSPGSAHSMGSPGPPSAGHIPITQGPGGYPDN
MHQMHKPMESMHEKGMSSDDPRYNQMKGMGRSGGHAGMGPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHLQRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQMQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWPEGPMANAAAPT
STPQKLIPPQPTGRPSPAPPVPPAASVMPPTQTSQGQAQAPAMVPLHQKQSRITPIQKPRGLDPVEILQEREY
RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLLNFQRLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQQKIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQKLTAKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKIDQKKDKRLAYLLQQTDEYVANLTELVRQHAAQVAKEKKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEQQA
AQPPTLPVEEKKKIPDPDSDDVSEVDARHI IENAKQDVDEYGVSAQALARGLSYYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVS LYNNNLNGILADEMGLGKTIQTIALITYLMEHKRINGPFLIIVPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLTTYEYIIKDKHILAKIRWKYMIVDEGHRMKNHHCKLTQVLNTHYVAP

RRLLLTGTPLQNKLPPELLWALLNLLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEEETILIIRRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAQGVLLTDGSEKDKKGGTKTLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRLPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLGLNLQSADTVIIFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAAKYKLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEQDEEEDVPDDETQNMIAHHEEFDLFLM
RMDLDRRREARNPKRKPRLMEDELPSWIIKDDAEVERLTCEEEEEKMFGRGSRHRKEVDYSDSLTEKQWLKTLK
AIEEGTLEEIEEEVRQKSSSRKRKRSDAGSSTPTTSTRSRDKDDESKKQKKRGRPPAEKLSNPPNLTKKMKKIV
DAVIKYKDSSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKERIRNHKYRSLNDLEKDVMLLCQNAQTFNLE
GSLIYEDSIVLQSVFTSVRQKIEKEDDSEGESEEEEEEGEESSESRSVKVVIKLRKEKAQDRLKGGRRRPSR
GSRAPVVSDDDSEEEQEEDRSGSGSEED (SEQ ID NO:17)。

[0418] >NP_001122318.1转录激活因子BRG1同种型D[智人]

[0419] MSTDPPLGGTTPRPGSPGPGSPGAMLGSPGSPGSAHSMGSPGPPSAGHPIPTQGGGYPQDN
MHQMHKPMESMHEKGMDDPRYNQMKGMGRSGGHAGMGPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHQLRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQQQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWPEGPMANAAAPT
STPQKLIPPQPTGRPSPAPPVPPAASVMPPTQSPGQPAQAPMVPLHQKQSRITPIQKPRGLDPVEILQEREY
RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLNLFQRQLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQKQIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQLTKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKLIDQKKDKRLAYLLQQTDEYVANLTELVRQHKAQVAKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEPQA
AQPTPLVEEKKKIPDPDSDDVSEVDARHIENAKQDQDDEYGVSAQALARGLQSYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVSLYNNNLNGILADEMGLGKTIQTIALITYLMEHKRINGPFLIIVPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLTTYEYIIKDKHILAKIRWKYMIVDEGHRMKNHCKLTQVLNTHYVAP
RRLLLTGTPLQNKLPPELLWALLNLLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEEETILIIRRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAQGVLLTDGSEKDKKGGTKTLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRLPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLGLNLQSADTVIIFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAAKYKLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEQDEEEDVPDDETQNMIAHHEEFDLFLM
RMDLDRRREARNPKRKPRLMEDELPSWIIKDDAEVERLTCEEEEEKMFGRGSRHRKEVDYSDSLTEKQWLKTLK
AIEEGTLEEIEEEVRQKSSSRKRKRSDAGSSTPTTSTRSRDKDDESKKQKKRGRPPAEKLSNPPNLTKKMKKIV
DAVIKYKDSSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKERIRNHKYRSLNDLEKDVMLLCQNAQTFNLE
SLIYEDSIVLQSVFTSVRQKIEKEDDSEGESEEEEEEGEESSESRSVKVVIKLRKEKAQDRLKGGRRRPSRG
SRAKPVVSDDDSEEEQEEDRSGSGSEED (SEQ ID NO:18)。

[0420] >NP_001122319.1转录激活因子BRG1同种型E[智人]

[0421] MSTDPPLGGTTPRPGSPGPGSPGAMLGSPGSPGSAHSMGSPGPPSAGHPIPTQGGGYPQDN
MHQMHKPMESMHEKGMDDPRYNQMKGMGRSGGHAGMGPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHQLRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQQQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWPEGPMANAAAPT
STPQKLIPPQPTGRPSPAPPVPPAASVMPPTQSPGQPAQAPMVPLHQKQSRITPIQKPRGLDPVEILQEREY

RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLLNFQRQLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQQKIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQKLTAKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKLIDQKKDKRLAYLLQQTDEYVANLTELVRQHKAQVAKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEQPQA
AQPPTLPVEEKKKIPDPDSDDVSEVDARHI IENAKQDVDDEYGVSAQALARGLSYYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVS LYNNLNGILADEMGLGKTIQTIALITYLMEHKRINGPFLI I VPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLTTYEYI IKDKHILAKIRWKYMI VDEGHRMKNHHCKLTQVLNTHYVAP
RRLLLTGTPQLNKLPELWALLNLLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEEETILI I RRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAQGVLLTDGSEKDKKGKGGTKLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRI LPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLG LNLQSADTVI IFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAAKYKLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEQDEEEDEV PDDETVNQMIARHEEEFDLFM
RMDLDRRREARNPKRKPRLMEDELPSWI IKDDAEVERLTCEEEEEKMFGRGSRHRKEVDYSDSLTEKQWLKAIE
EGTLEEIEEEVRQKKSSRKRKRDSAGSSTPTTSTRSRDKDES KKKQKKRGRPPAEKLSNPPNLTKMKKIVDAV
IKYKDS SGRQLSEVFIQLPSRKELPEYYELIRKPVDFFKKIKERIRNHKYRSLNDLEKDVMLLCQNAQTFNLEGLS
IYEDSIVLQSVFTSVRQKIEKEDDSEGESEEEEEEGESESRSVKV KIKLGRKEKAQDRLKGGRRRPSRGSRA
AKPVVSDDDSEEEQEEDRSGSGSEED (SEQ ID NO:19)。

[0422] >NP_001122320.1转录激活因子BRG1同种型F[智人]

[0423] MSTPDPLGGTPRPGSPGPGSPGAMLGSPGSPGSAHSMGSPGPPSAGHIPITQGGGYPQDN
MHQMHPMESMHEKGM SDDPRYNQMKGMGRSGGHAGMGPPPSPMDQHSQGYPSPLGGSEHASSPVPASGPSSGPQ
MSSGPGGAPLDGADPQALGQQNRGPTPFNQNLHQLRAQIMAYKMLARGQPLPDHLQMAVQGKRPMGPMQMQMPTL
PPPSVSATGPGPGPGPGPGPGPAPPNYSRPHGMGGPNMPPPGPSGVPPGMPGQPPGGPPKPWPEGPMANAAAPT
STPQKLIPPQTGRPSPAPPAPPAASPVMPPTQSPGQPAQAPMVPLHQKQSRTPIQKPRGLDPVEILQEREY
RLQARIAHRIQELENLPGSLAGDLRTKATIELKALRLLNFQRQLRQEVVCMRRDTALETALNAKAYKRSKRQSLR
EARITEKLEKQQKIEQERKRRQKHQEYLSILQHAKDFKEYHRSVTGKIQKLTAKAVATYHANTEREQKKENERIEK
ERMRLMAEDEEGYRKLIDQKKDKRLAYLLQQTDEYVANLTELVRQHKAQVAKKKKKKKKAENAEGQTPAIG
PDGEPLDETSQMSDLPVKVIHVESGKILTGTDAPKAGQLEAWLEMNPGYEVAPRSDSEESGSEEEEEEEEEEQPQA
AQPPTLPVEEKKKIPDPDSDDVSEVDARHI IENAKQDVDDEYGVSAQALARGLSYYAVAHAVTERVDKQSALMVNG
VLKQYQIKGLEWLVS LYNNLNGILADEMGLGKTIQTIALITYLMEHKRINGPFLI I VPLSTLSNWAYEFDKWAPS
VVKVSYKGSPAARRAFVPQLRSGKFNVLTTYEYI IKDKHILAKIRWKYMI VDEGHRMKNHHCKLTQVLNTHYVAP
RRLLLTGTPQLNKLPELWALLNLLPTIFKSCSTFEQWFNAPFAMTGEKVDLNEEETILI I RRLHKVLRPFLLRRL
KKEVEAQLPEKVEYVIKCDMSALQRVLYRHMQAQGVLLTDGSEKDKKGKGGTKLMNTIMQLRKICNHPYMFQHIE
ESFSEHLGFTGGIVQGLDLYRASGKFELLDRI LPKLRATNHKVLLFCQMTSLMTIMEDYFAYRGFKYLRLDGTTKA
EDRGMLLKTFNEPGSEYFIFLLSTRAGGLG LNLQSADTVI IFDSDWNPHQDLQAQDRAHRIGQQNEVRVRLCTVN
SVEEKILAAAKYKLNVDQKVIQAGMFDQKSSSHERRAFLQAILEHEEQDEEEDEV PDDETVNQMIARHEEEFDLFM
RMDLDRRREARNPKRKPRLMEDELPSWI IKDDAEVERLTCEEEEEKMFGRGSRHRKEVDYSDSLTEKQWLKAIE
EGTLEEIEEEVRQKKSSRKRKRDSAGSSTPTTSTRSRDKDES KKKQKKRGRPPAEKLSNPPNLTKMKKIVDAV
IKYKDS SGRQLSEVFIQLPSRKELPEYYELIRKPVDFFKKIKERIRNHKYRSLNDLEKDVMLLCQNAQTFNLEGLS
IYEDSIVLQSVFTSVRQKIEKEDDSEGESEEEEEEGESESRSVKV KIKLGRKEKAQDRLKGGRRRPSRGSRA

KPVVSDDDDSEEEQEEDRSGSGSEED (SEQ ID NO:20)。

[0424] 参考文献

[0425] 本文(例如在背景技术、发明内容、具体实施方式、实例和/或参考文献部分中)提及的所有出版物、专利、专利申请、专利公开和数据库条目(例如,序列数据库条目)特此通过援引以其全文并入,如同每个单独的出版物、专利、专利申请、专利公开和数据库条目被专门地且单独地通过援引并入本文一样。在冲突的情况下,将以本申请(包括本文的任何定义)为准。

[0426] 等效物和范围

[0427] 除非另有定义,否则本文所用的所有技术和科学术语都具有与本披露所属领域的普通技术人员通常所理解的相同的含义。在说明书中,单数形式也包括复数形式,除非上下文另有明确规定。将本文提及的所有公开、专利申请、专利以及其他参考文献通过援引并入本文。不承认本文所引用的参考文献是所要求保护的发明的现有技术。在发生冲突的情况下,将以包括定义的本说明书为准。此外,这些材料、方法和实例仅是说明性的并且不旨在是限制性的。

[0428] 本领域技术人员将认识到、或能够仅使用常规实验来确定本文所述的实施例的许多等效物。本披露的范围不旨在限于以上描述,而是如所附权利要求中所阐述。

[0429] 除非相反地指出或另外从上下文明显看出,否则冠词诸如“一个/种(a/an)”和“该/所述(the)”可以意指一个/种或多于一个/种,。除非相反地指出或另外从上下文明显看出,否则如果存在一个、多于一个或所有组成员,则认为满足了在组的两个或更多个成员之间包括“或”的权利要求或描述。在两个或更多个组成员之间包括“或”的组的披露内容提供其中恰好存在该组的一个成员的实施例、其中存在该组的超过一个成员的实施例以及其中存在所以组成员的实施例。出于简洁的目的,那些实施例未在本文中单独说明,但应理解,这些实施例中的每一个在本文中提供并且可以具体地要求保护或放弃保护。

[0430] 应理解,本披露涵盖其中将一个或多个限制、要素、条款或描述性术语从一个或多个权利要求或从说明书的一个或多个相关部分引入到另一权利要求中的所有变型、组合和排列。例如,可以对附属于另一权利要求的权利要求加以修改,以使其包括一个或多个在附属于同一基础权利要求的任何其他权利要求中所见的限制。此外,在权利要求叙述组合物的情况下,应理解,除非另外指示或者除非对于本领域的普通技术人员来说明显将出现矛盾或不一致,否则包括根据本文所披露的任何制备或使用方法的根据本领域已知的方法来制备或使用组合物的方法。

[0431] 在要素是以例如马库西组(Markush group)格式等列表呈现的情况下,应理解,也披露了这些要素的每个可能的亚组,并且可以从该组中除去任何要素或要素的亚组。还应注意,术语“包含(comprising)”旨在是开放性的并且允许包括另外的要素或步骤。通常,应理解,当实施例、产物或方法被称为包括特定要素、特征或步骤时,还提供了由此类要素、特征或步骤组成或基本上由其组成的实施例、产物或方法。出于简洁的目的,那些实施例未在本文中单独说明,但应理解,这些实施例中的每一个在本文中提供并且可以具体地要求保护或放弃保护。

[0432] 在给出范围时,端点被包括在内。此外,应理解,除非另外指出或另外从上下文和/或本领域普通技术人员的理解明显看出,则表示为范围的值可假定为所陈述范围内的任

何特定值(例如在一些实施例中)到该范围的下限的单位的十分之一,除非上下文另有明确规定。出于简洁的目的,在每个范围中的值未在本文中单独说明,但应理解,这些值中的每一个在本文中提供并且可以具体地要求保护或放弃保护。还应理解,除非另外指出或另外从上下文和/或本领域普通技术人员的理解明显看出,否则表示为范围的值可以假定给定范围内的任何子范围,其中子范围的端点表示至与该范围的下限单位的十分之一相同程度的精确度。

[0433] 此外,应理解,本披露的任何特定实施例可以从任何一个或多个权利要求中明确排除。在给出范围的情况下,范围内的任何值可以从任何一个或多个权利要求中明确排除。本发明的组合物和/或方法的任何实施例、要素、特征、应用或方面可以从任何一个或多个权利要求中排除。出于简洁的目的,本文中未明确阐述排除一个或多个要素、特征、目的或方面的所有实施例。

序列表

<110> EPIZYME股份有限公司

<120> 癌症治疗形式

<130> EPIZ-074/001W0

<150> 62/453,929

<151> 2017-02-02

<150> 62/479,878

<151> 2017-03-31

<160> 20

<170> PatentIn 3.5版

<210> 1

<211> 5879

<212> DNA

<213> 智人

<400> 1

```

tcagaagaaa gccccgagat cacagagacc cggcgagatc acagagaccc ggcctgaagg 60
aacgtggaaa gaccaatgta cctgttttga ccggttgccct ggagcaagaa gttccagttg 120
gggagaattt tcagaagata aagtcggaga ttgtggaaag acttgacttg cagcattact 180
ctactgactg gcagagacag gagaggtaga tgtccacgcc cacagaccct ggtgcgatgc 240
cccacccagg gccttcgccg gggcctgggc cttcccctgg gccaattctt gggcctagtc 300
caggaccagg accatcccca ggttccgtcc acagcatgat ggggcccaagt cctggacctc 360
caagtgtctc ccatcctatg ccgacgatgg ggtccacaga cttcccacag gaaggcatgc 420
atcaaagtca taagcccatc gatggtatac atgacaaggg gattgtagaa gacatccatt 480
gtggatccat gaagggcact ggtatgacac cacctcacc aggcattgggc cttcccaga 540
gtccaatgga tcaacacagc caaggttata tgtcaccaca cccatctcca ttaggagccc 600
cagagcacgt ctccagccct atgtctggag gaggcccaac tccacctcag atgccaccaa 660
gccagccggg ggccctcatc ccaggtgatc cgcaggccat gagccagccc aacagaggtc 720
cctcaccttt cagtcctgtc cagctgcac agcttcgagc tcagatttta gcttataaaa 780
tgctggcccc aggccagccc ctccccgaaa cgctgcagct tgcagtccag gggaaaagga 840
cgttgccctg cttgcagcaa caacagcagc agcaacagca gcagcagcag cagcagcagc 900
agcagcagca gcagcaacag cagccgcagc agcagccgcc gcaaccacag acgcagcaac 960
aacagcagcc ggcccttggt aactacaaca gaccatctgg ccgggggccg gagctgagcg 1020
gcccagcac cccgcagaag ctgccggtgc ccgcgcccgg cggccggccc tcgcccgcgc 1080
ccccgcagc cgcgcagccg cccgcggccg cagtgcgccg gccctcagtg ccgcagccgg 1140
ccccggggca gccctcgccc gtccctccagc tgcagcagaa gcagagccgc atcagcccca 1200
tccagaaacc gcaaggcctg gaccccgctg aaattctgca agagcgggaa tacagacttc 1260
aggcccgcat agctcatagg atacaagaac tggaaaatct gcctggctct ttgccaccag 1320
atttaagaac caaagcaacc gtggaactaa aagcatttc gttactcaat ttccagcgtc 1380

```

agctgagaca ggaggtggtg gcctgcatgc gcaggacac gaccctggag acggctctca 1440
 actccaaagc atacaaacgg agcaagcgcc agactctgag agaagctcgc atgaccgaga 1500
 agctggagaa gcagcagaag attgagcagg agaggaaacg ccgtcagaaa caccaggaat 1560
 acctgaacag tattttgcaa catgcaaaag attttaagga atatcatcgg tctgtggccg 1620
 gaaagatcca gaagctctcc aaagcagtgg caacttggca tgccaacact gaaagagagc 1680
 agaagaagga gacagagcgg attgaaaagg agagaatgcg gcgactgatg gctgaagatg 1740
 aggagggtta tagaaaactg attgatcaaa agaaagacag gcgttttagct taccttttgc 1800
 agcagaccga tgagtatgta gccaatctga ccaatctggt ttgggagcac aagcaagccc 1860
 aggcagccaa agagaagaag aagaggagga ggaggaagaa gaaggctgag gagaatgcag 1920
 aggggtgggga gtctgccctg ggaccggatg gagagcccat agatgagagc agccagatga 1980
 gtgacctccc tgtcaaagtg actcacacag aaaccggcaa ggttctgttc ggaccagaag 2040
 cacccaaagc aagtcagctg gacgcctggc tggaaatgaa tcctggttat gaagttgccc 2100
 ctagatctga cagtgaagag agtgattctg attatgagga agaggatgag gaagaagagt 2160
 ccagtaggca ggaaaccgaa gagaaaatac tcctggatcc aaatagcgaa gaagtttctg 2220
 agaaggatgc taagcagatc attgagacag ctaagcaaga cgtggatgat gaatacagca 2280
 tgcagtacag tgccaggggc tcccagtcct actacaccgt ggctcatgcc atctcggaga 2340
 ggggtggagaa acagtctgcc ctccctaatta atgggaccct aaagcattac cagctccagg 2400
 gcctggaatg gatggtttcc ctgtataata acaacttgaa cggaatctta gccgatgaaa 2460
 tggggcttgg aaagaccata cagaccattg cactcatcac ttatctgatg gagcacaana 2520
 gactcaatgg cccctatctc atcattgttc ccctttcgac tctatctaac tggacatatg 2580
 aatttgacaa atgggctcct tctgtggtga agatttctta caagggtact cctgccatgc 2640
 gtcgctccct tgtccccag ctacggagtg gcaaattcaa tgtcctcttg actacttatg 2700
 agtatattat aaaagacaag cacattcttg caaagattcg gtggaaatac atgatagtgg 2760
 acgaaggcca ccgaatgaag aatcaccact gcaagctgac tcaggtcttg aacactcact 2820
 atgtggcccc cagaaggatc ctcttgactg ggaccccgt gcagaataag ctccctgaac 2880
 tctgggccct cctcaacttc ctccctccaa caatttttaa gagctgcagc acatttgaac 2940
 aatggttcaa tgctccattt gccatgactg gtgaaagggt ggacttaaata gaagaagaaa 3000
 ctatatattg catcaggcgt ctacataagg tgtaagacc atttttacta aggagactga 3060
 agaaagaagt tgaatcccag cttcccgaag aagtggaaata tgtgatcaag tgtgacatgt 3120
 cagctctgca gaagattctg tatcgccata tgcaagccaa ggggatcctt ctacagatg 3180
 gttctgagaa agataagaag gggaaaggag gtgctaagac acttatgaac actattatgc 3240
 agttgagaaa aatctgcaac cacccatata tgtttcagca cattgaggaa tcctttgctg 3300
 aacacctagg ctattcaaat ggggtcatca atggggctga actgtatcgg gcctcaggga 3360
 agtttgagct gcttgatcgt attctgccaa aattgagagc gactaatcac cgagtgtgc 3420
 ttttctgcca gatgacatct ctcatgacca tcatggagga ttattttgct tttcggaact 3480
 tcctttacct acgccttgat ggcaccacca agtctgaaga tcgtgctgct ttgctgaaga 3540
 aattcaatga acctggatcc cagtatttca ttttcttgct gagcacaaga gctggtggcc 3600
 tgggcttaaa tcttcaggca gctgatacag tggatcatct tgacagcgac tggaatcctc 3660
 atcaggatct gcaggcccaa gaccgagctc accgcatcgg gcagcagaac gaggtccggg 3720

```

tactgaggct ctgtaccgtg aacagcgtgg aggaaaagat cctcgcggcc gcaaaataca 3780
agctgaacgt ggatcagaaa gtgatccagg cgggcatgtt tgaccaaaag tcttcaagcc 3840
acgagcggag ggcatcctg caggccatct tggagcatga ggaggaaaat gaggaagaag 3900
atgaagtacc ggacgatgag actctgaacc aatgattgc tcgacgagaa gaagaatttg 3960
acctttttat gcggatggac atggaccggc ggagggaaga tgcccggaac ccgaaacgga 4020
agccccgttt aatggaggag gatgagctgc ctcctggat cattaaggat gacgctgaag 4080
tagaaaaggct cacctgtgaa gaagaggagg agaaaatatt tgggaggggg tcccgcagc 4140
gccgtgacgt ggactacagt gacgccctca cggagaagca gtggctaagg gccatcgaag 4200
acggcaattt ggaggaaaat gaagaggaag tacggcttaa gaagcgaata agacgaagaa 4260
atgtggataa agatcctgca aaagaagatg tggaaaaagc taagaagaga agaggccgcc 4320
ctcccgtga gaaactgtca ccaaactccc ccaaactgac aaagcagatg aacgctatca 4380
tcgatactgt gataaactac aaagataggt gtaacgtgga gaagtgccc agtaattctc 4440
agttggaaat agaaggaaac agttcagggc gacagctcag tgaagtcttc attcagttac 4500
cttcaaggaa agaattacca gaatactatg aattaattag gaagccagtg gatttcaaaa 4560
aaataaagga aaggattcgt aatcataagt accggagcct aggcgacctg gagaaggatg 4620
tcatgcttct ctgtcacaac gctcagacgt tcaacctgga gggatcccag atctatgaag 4680
actccatcgt cttacagtca gtgtttaaga gtgcccggca gaaaattgcc aaagaggaag 4740
agagtgagga tgaaagcaat gaagaggagg aagaggaaga tgaagaagag tcagagtccg 4800
aggcaaaatc agtcaagggt aaaattaagc tcaataaaaa agatgacaaa ggccgggaca 4860
aagggaagga caagaaaagg ccaaactcag gaaaagccaa acctgtagtg agcgattttg 4920
acagcgatga ggagcaggat gaacgtgaac agtcagaagg aagtgggacg gatgatgagt 4980
gatcagtatg gaccttttc cttggtagaa ctgaattcct tcctcccctg tctcatttct 5040
accagtgag ttcatttgtc atataggcac tgggttggtt ctatatcatc atcgtctata 5100
aactagcttt aggatagtgc cagacaaaca tatgatata tggtgtaaaa aacacacaca 5160
tacacaaata tttgtaacat attgtgacca aatgggcctc aaagattcag attgaaacaa 5220
acaaaaagct tttgatggaa aatatgtggg tggatagtat atttctatgg gtgggtctaa 5280
tttggtaacg gtttgattgt gcctggtttt atcacctgtt cagatgagaa gatttttgtc 5340
ttttgtagca ctgataacca ggagaagcca ttaaaagcca ctggttattt tatttttcat 5400
caggcaattt tcgaggtttt tatttgttcg gtattgtttt tttacactgt ggtacatata 5460
agcaacttta atagtgata aatgtacagt agttagattt cacctgcata tacatttttc 5520
cattttatgc tctatgatct gaacaaaagc ttttgaatt gtataagatt tatgtctact 5580
gtaaacattg ctttaatttt ttgctcttga tttaaaaaaa agttttgttg aaagcgctat 5640
tgaatattgc aatctatata gtgtattgga tggtctcttt tgtcacctg atctcctatg 5700
ttaccaatgt gtatcgtctc cttctcccta aagtgtactt aatctttgct tctttgcac 5760
aatgtctttg gttgcaagtc ataagcctga ggcaaataaa attccagtaa tttcgaagaa 5820
tgtggtgttg gtgctttcct aataaagaaa taatttagct tgacaaaaaa aaaaaaaaaa 5879

```

<210> 2

<211> 5838

<212> DNA

<213> 智人

<400> 2

gcgtcttccg gcgccccg cgaggaggcgag ggtgggacgc tgggcggagc ccgagtttag 60
 gaagaggagg ggacggctgt catcaatgaa gtcatattca taatctagtc ctctctccct 120
 ctgtttctgt actctgggtg actcagagag ggaagagatt cagccagcac actcctcgcg 180
 agcaagcatt actctactga ctggcagaga caggagaggt agatgtccac gcccacagac 240
 cctgggtgca tgccccaccc agggccttcg ccggggcctg ggccttcccc tgggccaatt 300
 cttgggccta gtccaggacc aggaccatcc ccaggttccg tccacagcat gatggggcca 360
 agtcctggac ctccaagtgt ctcccatcct atgccgacga tgggggtccac agacttccca 420
 caggaaggca tgcatcaaat gcataagccc atcgatggta tacatgacaa ggggattgta 480
 gaagacatcc attgtggatc catgaagggc actggtatgc gaccacctca cccaggcatg 540
 ggccctcccc agagtccaat ggatcaacac agccaagggt atatgtcacc acacctatct 600
 ccattaggag cccagagca cgtctccagc cctatgtctg gaggaggccc aactccacct 660
 cagatgccac caagccagcc gggggccctc atcccagggt atccgcaggc catgagccag 720
 cccaacagag gtccctcacc tttcagtcct gtccagctgc atcagcttcg agctcagatt 780
 ttagcttata aaatgctggc ccgaggccag cccctccccg aaacgtgca gcttgagtc 840
 caggggaaaa ggacgttgcc tggtttgag caacaacagc agcagcaaca gcagcagcag 900
 cagcagcagc agcagcagca gcagcagca cagcagccgc agcagcagcc gccgcaacca 960
 cagacgcagc aacaacagca gccggccctt gttaactaca acagaccatc tggcccgggg 1020
 ccggagctga gcggccccgag cccccgcag aagctgccgg tgcccgcgcc cggcggccgg 1080
 ccctcgcccc cgccccccgc agccgcgcag ccgcccgcgg ccgcagtgcc cgggccctca 1140
 gtgccgcagc cggccccggg gcagccctcg cccgtcctcc agctgcagca gaagcagagc 1200
 cgcatcagcc ccatccagaa accgcaaggc ctggaccccc tggaaattct gcaagagcgg 1260
 gaatacagac ttcaggcccc catagctcat aggatacaag aactggaaaa tctgcctggc 1320
 tctttgccac cagatttaag aaccaaagca accgtggaac taaaagcact tcggttactc 1380
 aatttccagc gtcagctgag acaggaggtg gtggcctgca tgcgcaggga cagcacctg 1440
 gagacggctc tcaactcaa agcatacaaa cggagcaagc gccagactct gagagaagct 1500
 cgcatgaccg agaagctgga gaagcagcag aagattgagc aggagaggaa acgccgtcag 1560
 aaacaccagg aatacctgaa cagtattttg caacatgcaa aagattttta ggaatatcat 1620
 cgtctgttg ccggaagat ccagaagctc tccaaagcag tggcaacttg gcatgccaac 1680
 actgaaaagag agcagaagaa ggagacagag cggattgaaa aggagagaat gcggcgactg 1740
 atggctgaag atgaggagg ttatagaaaa ctgattgac aaaagaaaga caggcgttta 1800
 gcttaccttt tgcagcagac cgatgagtat gtagccaatc tgaccaatct ggtttgggag 1860
 cacaagcaag cccaggcagc caaagagaag aagaagagga ggaggaggaa gaagaaggct 1920
 gaggagaatg cagagggtgg ggagtctgcc ctgggaccgg atggagagcc catagatgag 1980
 agcagccaga tgagtacct ccctgtcaaa gtgactcaca cagaaaccgg caaggttctg 2040
 ttcggaccag aagcacccaa agcaagtcag ctggacgcct ggctggaaat gaatcctggt 2100
 tatgaagttg cccctagatc tgacagtga gagagtgatt ctgattatga ggaagaggat 2160
 gaggaagaag agtccagtag gcaggaaacc gaagagaaaa tactcctgga tccaaatagc 2220

```

gaagaagttt ctgagaagga tgctaagcag atcattgaga cagctaagca agacgtggat 2280
gatgaataca gcatgcagta cagtgccagg ggctcccagt cctactacac cgtggctcat 2340
gccatctcgg agaggggtgga gaaacagtct gccctcctaa ttaatgggac cctaaagcat 2400
taccagctcc agggcctgga atggatgggt tccctgtata ataacaactt gaacggaatc 2460
ttagccgatg aaatggggct tggaaagacc atacagacca ttgcactcat cacttatctg 2520
atggagcaca aaagactcaa tggcccctat ctcatcattg ttcccctttc gactctatct 2580
aactggacat atgaatttga caaatgggct cttctgtgg tgaagatttc ttacaagggt 2640
actcctgcca tgcgtcgtc cttgtcccc cagctacgga gtggcaaatt caatgtcctc 2700
ttgactactt atgagtatat tataaaagac aagcacattc ttgcaaagat tcggtggaaa 2760
tacaatgatg tggacgaagg ccaccgaatg agaatacacc actgcaagct gactcaggtc 2820
ttgaacactc actatgtggc ccccgaaagg atcctcttga ctgggacccc gctgcagaat 2880
aagctccctg aactctgggc cctcctcaac ttctcctcc caacaatttt taagagctgc 2940
agcacatttg aacaatgggt caatgtcca tttgcatga ctggtgaaag ggtggactta 3000
aatgaagaag aaactatatt gatcatcagg cgtctacata aggtgttaag accattttta 3060
ctaaggagac tgaagaaaga agttgaatcc cagcttcccg aaaaagtgga atatgtgatc 3120
aagtgtgaca tgtcagctct gcagaagatt ctgtatcgcc atatgcaagc caaggggatc 3180
cttctcacag atggttctga gaaagataag aaggggaaag gaggtgctaa gacacttatg 3240
aacactatta tgcagttgag aaaaatctgc aaccacccat atatgtttca gcacattgag 3300
gaatcctttg ctgaacacct aggctattca aatggggctc tcaatggggc tgaactgtat 3360
cgggcctcag ggaagtttga gctgcttgat cgtattctgc caaaattgag agcgactaat 3420
caccgagtg tcttttctg ccagatgaca tctctcatga ccatcatgga ggattatttt 3480
gcttttcgga acttccttta cctacgcctt gatggcacca ccaagtctga agatcgtgct 3540
gctttgctga agaaattcaa tgaacctgga tcccagtatt tcattttctt gctgagcaca 3600
agagctgggtg gcctgggctt aaatcttcag gcagctgata cagtggctcat ctttgacagc 3660
gactggaatc ctcatcagga tctgcaggcc caagaccgag ctcaccgcat cgggcagcag 3720
aacgaggtcc ggggtactgag gctctgtacc gtgaacagcg tggaggaaaa gatcctcgcg 3780
gccgcaaaat acaagctgaa cgtggatcag aaagtgatcc aggcgggcat gtttgaccaa 3840
aagtcttcaa gccacgagcg gagggcattc ctgcaggcca tcttgagca tgaaggaggaa 3900
aatgaggaag aagatgaagt accggacgat gagactctga accaaatgat tgctcgacga 3960
gaagaagaat ttgacctttt tatgcggatg gacatggacc ggcggaggga agatgcccg 4020
aaccgaaac ggaagccccg ttaaatggag gaggatgagc tgccctcctg gatcattaag 4080
gatgacgctg aagtagaaaag gctcacctgt gaagaagagg aggagaaaat atttgggagg 4140
gggtcccgcc agcgccgtga cgtggactac agtgacgcc tcacggagaa gcagtggcta 4200
agggccatcg aagacggcaa tttggaggaa atggaagagg aagtacggct taagaagcga 4260
aaaagacgaa gaaatgtgga taaagatcct gcaaaagaag atgtggaaaa agctaagaag 4320
agaagaggcc gccctccgc tgagaaactg tcaccaaatc ccccaaact gacaaagcag 4380
atgaacgcta tcatcgatac tgtgataaac taaaagata gttcaggcg acagctcagt 4440
gaagtcttca ttcagttacc ttcaaggaaa gaattaccag aatactatga attaattagg 4500
aagccagtgg atttcaaaaa aataaaggaa aggattcgta atcataagta ccggagccta 4560

```

ggcgacctgg agaaggatgt catgcttctc tgtcacaacg ctcagacgtt caacctggag 4620
 ggatcccaga tctatgaaga ctccatcgtc ttacagtcag tgtttaagag tgccccggcag 4680
 aaaattgccca aagaggaaga gaggtaggat gaaagcaatg aagaggagga agaggaagat 4740
 gaagaagagt cagagtccga ggcaaaatca gtcaagggtga aaattaagct caataaaaaa 4800
 gatgacaaaag gccgggacaa agggaaaggc aagaaaaggc caaatcgagg aaaagccaaa 4860
 cctgtagtga gcgattttga cagcgatgag gaggcaggat aacgtgaaca gtcagaagga 4920
 agtgggacgg atgatgagt atcagtatgg acctttttcc ttggtagaac tgaattcctt 4980
 cctccccctgt ctcatattcta cccagtgagt tcatttgtca tataggcact gggttgtttc 5040
 tatatcatca tcgtctataa actagcttta ggatagtgcc agacaaacat atgatatcat 5100
 ggtgtaaaaa acacacacat acacaaatat ttgtaacata ttgtgaccaa atgggcctca 5160
 aagattcaga ttgaaacaaa caaaaagctt ttgatggaaa atatgtgggt ggatagtata 5220
 tttctatggg tgggtctaata ttggtaacgg tttgattgtg cctggtttta tcacctgttc 5280
 agatgagaag atttttgtct tttgtagcac tgataaccag gagaagccat taaaagccac 5340
 tggttatttt atttttcatc aggcaatttt cgaggttttt atttgttcgg tattgttttt 5400
 ttacactgtg gtacatataa gcaactttaa taggtgataa atgtacagta gttagatttc 5460
 acctgcatat acatttttcc attttatgct ctatgatctg aacaaaagct ttttgaattg 5520
 tataagattt atgtctactg taaacattgc ttaatttttt tgctcttgat ttaaaaaaaa 5580
 gttttgttga aagcgctatt gaatattgca atctatatag tgtattggat ggcttctttt 5640
 gtcacctga tctcctatgt taccaatgtg tatcgtctcc ttctccctaa agtgtactta 5700
 atctttgctt tctttgcaca atgtctttgg ttgcaagtca taagcctgag gcaataaaaa 5760
 ttccagtaat ttcgagaat gtggtgttgg tgctttccta ataaagaaat aatttagctt 5820
 gacaaaaaaa aaaaaaaa 5838

<210> 3

<211> 5664

<212> DNA

<213> 智人

<400> 3

gcgtcttccg gcgcccgcgg aggaggcgag ggtgggacgc tgggcggagc ccgagtttag 60
 gaagaggagg ggacggctgt catcaatgaa gtcatattca taatctagtc ctctctccct 120
 ctgtttctgt actctgggtg actcagagag ggaagagatt cagccagcac actcctcgcg 180
 agcaagcatt actctactga ctggcagaga caggagaggt agatgtccac gccacagac 240
 cctggtgcga tgccccaccc agggccttcg ccggggcctg ggcttccccc tgggccaatt 300
 cttgggccta gtccaggacc aggaccatcc ccaggttccg tccacagcat gatggggcca 360
 agtcctggac ctccaagtgt ctcccatcct atgccgacga tggggtccac agacttccca 420
 caggaaggca tgcatcaaat gcataagccc atcgatggta tacatgacaa ggggattgta 480
 gaagacatcc attgtggatc catgaagggc actggtatgc gaccacctca cccaggcatg 540
 ggccctcccc agagtccaat ggatcaacac agccaaggtt atatgtcacc acaccatct 600
 ccattaggag cccagagca cgtctccagc cctatgtctg gaggaggccc aactccacct 660
 cagatgccac caagccagcc gggggccctc atcccaggtg atccgcaggc catgagccag 720

```

cccaacagag gtccctcacc ttctcagtcct gtccagctgc atcagcttcg agctcagatt 780
ttagcttata aaatgctggc ccgaggccag cccctccccg aaacgctgca gcttgagctc 840
caggggaaaa ggacgttgcc tggcttgag caacaacagc agcagcaaca gcagcagcag 900
cagcagcagc agcagcagca gcagcagcaa cagcagccgc agcagcagcc gccgcaacca 960
cagacgcagc aacaacagca gccggccctt gttaactaca acagaccatc tggcccgggg 1020
ccggagctga gcggccccgag caccgccgag aagctgccgg tgcccgcgcc cggcgccgg 1080
ccctcgcccc cgccccccgc agccgcgcag ccgcccgcgg ccgcagtgcc cgggccctca 1140
gtgccgcagc cggccccggg gcagccctcg cccgtcctcc agctgcagca gaagcagagc 1200
cgcatcagcc ccatccagaa accgcaaggc ctggaccccc tggaaattct gcaagagcgg 1260
gaatacagac ttcaggcccc catagctcat aggatacaag aactggaaaa tctgcctggc 1320
tctttgccac cagatttaag aaccaaaagc accgtggaac taaaagcact tcggttactc 1380
aatttccagc gtcagctgag acaggaggtg gtggcctgca tgcgcaggga cagcacctg 1440
gagacggctc tcaactccaa agcatacaaa cggagcaagc gccagactct gagagaagct 1500
cgcatgaccg agaagctgga gaagcagcag aagattgagc aggagaggaa acgccgtcag 1560
aaacaccagg aatacctgaa cagtattttg caacatgcaa aagattttta ggaatatcat 1620
cggctctgtg ccgaaaagat ccagaagctc tccaaagcag tggcaacttg gcatgccaac 1680
actgaaagag agcagaagaa ggagacagag cggattgaaa aggagagaat gcggcgactg 1740
atggctgaag atgaggaggg ttatagaaaa ctgattgatc aaaagaaaga caggcgttta 1800
gcttaccttt tgcagcagac cgatgagtat gtagccaatc tgaccaatct ggtttgggag 1860
cacaagcaag cccaggcagc caaagagaag aagaagagga ggaggaggaa gaagaaggct 1920
gaggagaatg cagagggtgg ggagtctgcc ctgggaccgg atggagagcc catagatgag 1980
agcagccaga tgagtacct ccctgtcaaa gtgactcaca cagaaaccgg caaggttctg 2040
ttcggaccag aagcacccaa agcaagtcag ctggacgcct ggctggaaat gaatcctggt 2100
tatgaagttg cccctagatc tgacagtga gagagtgatt ctgattatga ggaagaggat 2160
gaggaagaag agtccagtag gcaggaaacc gaagagaaaa tactcctgga tccaaatagc 2220
gaagaagttt ctgagaagga tgctaagcag atcattgaga cagctaagca agacgtggat 2280
gatgaataca gcatgcagta cagtgccagg ggctcccagt cctactacac cgtggctcat 2340
gccatctcgg agagggtgga gaaacagtct gccctcctaa ttaatgggac cctaaagcat 2400
taccagctcc agggcctgga atggatggtt tccctgtata ataacaactt gaacggaatc 2460
ttagccgatg aaatggggct tggaaagacc atacagacca ttgcactcat cacttatctg 2520
atggagcaca aaagactcaa tggcccctat ctcatcattg ttcccctttc gactctatct 2580
aactggacat atgaatttga caaatgggct cttctgtgg tgaagatttc ttacaagggt 2640
actcctgcca tgcgtcgctc cttgtcccc cagctacgga gtggcaaatt caatgtcctc 2700
ttgactactt atgagtatat tataaaagac aagcacattc ttgcaaagat tcggtgga 2760
tacatgatag tggacgaagg ccaccgaatg aagaatcacc actgcaagct gactcaggtg 2820
gacttaaatg aagaagaaac tatattgatc atcaggcgtc tacataaggt gttaagacca 2880
tttttactaa ggagactgaa gaaagaagtt gaatcccagc ttcccgaaaa agtggaaat 2940
gtgatcaagt gtgacatgtc agctctgcag aagattctgt atcgccatat gcaagccaag 3000
gggatccttc tcacagatgg ttctgagaaa gataagaagg ggaaaggagg tgctaagaca 3060

```

cttatgaaca ctattatgca gttgagaaaa atctgcaacc acccatatat gtttcagcac 3120
 attgaggaat cctttgctga acacctaggc tattcaaatg ggggtcatcaa tggggctgaa 3180
 ctgtatcggg cctcagggaa gtttgagctg cttgatcgta ttctgccaaa attgagagcg 3240
 actaatcacc gagtgctgct tttctgccag atgacatctc tcatgaccat catggaggat 3300
 tatttttgctt ttcggaactt cctttacctt cgccttgatg gcaccaccaa gtctgaagat 3360
 cgtgctgctt tgctgaagaa attcaatgaa cctggatccc agtatttcat tttcttgctg 3420
 agcacaagag ctggtggcct gggcttaaat cttcaggcag ctgatacagt ggtcatcttt 3480
 gacagcgact ggaatcctca tcaggatctg caggcccaag accgagctca ccgcatcggg 3540
 cagcagaacg aggtccgggt actgaggctc tgtaccgtga acagcgtgga ggaaaagatc 3600
 ctcgcggccg caaaaatacaa gctgaacgtg gatcagaaag tgatccaggc gggcatgttt 3660
 gacaaaaagt cttcaagcca cgagcggagg gcattcctgc aggccatctt ggagcatgag 3720
 gaggaatatg aggaagaaga tgaagtaccg gacgatgaga ctctgaacca aatgattgct 3780
 cgacgagaag aagaatttga cctttttatg cggatggaca tggaccggcg gagggaagat 3840
 gcccggaaacc cgaaacggaa gccccgttta atggaggagg atgagctgcc ctcttgatc 3900
 attaaggatg acgctgaagt agaaaggctc acctgtgaag aagaggagga gaaaatat 3960
 gggaggggggt cccgccagcg ccgtgacgtg gactacagtg acgccctcac ggagaagcag 4020
 tggctaaggg ccatcgaaga cggcaatttg gaggaatgg aagaggaagt acggcttaag 4080
 aagcgaaaaa gacgaagaaa tgtggataaa gatcctgcaa aagaagatgt ggaaaaagct 4140
 aagaagagaa gaggccgccc tcccgtgag aaactgtcac caaatcccc caaactgaca 4200
 aagcagatga acgctatcat cgatactgtg ataaactaca aagatagttc agggcgacag 4260
 ctcagtgaag tcttcattca gttacctca aggaaagaat taccagaata ctatgaatta 4320
 attaggaagc cagtggattt caaaaaata aaggaaagga ttcgtaatca taagtaccgg 4380
 agcctaggcg acctggagaa ggatgtcatg cttctctgtc acaacgtca gacgttcaac 4440
 ctggagggat cccagatcta tgaagactcc atcgtcttac agtcagtgtt taagagtgcc 4500
 cggcagaaaa ttgccaaaga ggaagagagt gaggatgaaa gcaatgaaga ggaggaagag 4560
 gaagatgaag aagagtcaga gtccgaggca aaatcagtca aggtgaaaat taagctcaat 4620
 aaaaaagatg acaaaggccg ggacaaaggg aaaggcaaga aaaggccaaa tcgaggaaaa 4680
 gccaaacctg tagtgagcga ttttgacagc gatgaggagc aggatgaacg tgaacagtca 4740
 gaaggaagtg ggacggatga tgagtgatca gtatggacct ttttccttgg tagaactgaa 4800
 ttccttcctc ccctgtctca tttctacca gtgagttcat ttgtcatata ggcactgggt 4860
 tgttttctata tcatcatcgt ctataaacta gcttttaggat agtgccagac aaacatatga 4920
 tatcatggtg taaaaaacac acacatacac aaatatgttg aacatatgtg gaccaaattg 4980
 gcctcaaaga ttcagattga aacaaacaaa aagcttttga tggaaaatat gtgggtggat 5040
 agtatatttc tatgggtggg tctaatttgg taacggttgg attgtgcctg gttttatcac 5100
 ctgttcagat gagaagattt ttgtcttttg tagcactgat aaccaggaga agccattaaa 5160
 agccactggt tattttat 11 ttcacaggc aattttcgag gtttttattt gttegggtatt 5220
 gtttttttac actgtggtac atataagcaa ctttaatagg tgataaatgt acagtagtta 5280
 gatttcacct gcatatacat ttttccattt tatgtcttat gatctgaaca aaagcttttt 5340
 gaattgtata agatttatgt ctactgtaaa cattgcttaa tttttttgct cttgatttaa 5400

aaaaaagttt tgttgaaagc gctattgaat attgcaatct atatatgtgta ttggatggct 5460
 tcttttgtca ccctgatctc ctatgttacc aatgtgtatc gtctccttct ccctaaagtg 5520
 tacttaatct ttgcttttct tgcacaatgt ctttggttgc aagtcataag cctgaggcaa 5580
 ataaaattcc agtaatttcg aagaatgtgg tgttggtgct ttcctaataa agaaataatt 5640
 tagcttgaca aaaaaaaaaa aaaa 5664

<210> 4

<211> 1846

<212> DNA

<213> 智人

<400> 4

cttgagagg cggagggtga aacgatgcgc aggagtggc ttggggcttt ttgtttgcgt 60
 gtccctgttt acctattcat aatcatggat cccctctgct ttgtgatact gtgaaccacg 120
 cataacagca attctttaca ccaccgggtt gagaagaagg cgcctgaggc tgactttctg 180
 gacctgccgt cacgcagtaa agatgtggtt ggccatcgaa gacggcaatt tggaggaaat 240
 ggaagaggaa gtacggctta agaagcgaag aagacgaaga aatgtggata aagatcctgc 300
 aaaagaagat gtggaaaaag ctaagaagag aagaggccgc cctcccgtg agaaactgtc 360
 accaaatccc cccaaactga caaagcagat gaacgctatc atcgatactg tgataaacta 420
 caaagatagt tcagggcgac agctcagtga agtcttcatt cagttacctt caaggaaaga 480
 attaccagaa tactatgaat taattaggaa gccagtggat ttcaaaaaaa taaaggaaag 540
 gattcgtaat cataagtacc ggagcctagg cgacctggag aaggatgtca tgcttctctg 600
 tcacaacgct cagacgttca acctggaggg atcccagatc tatgaagact ccatcgtctt 660
 acagtcagtg ttttaagagtg cccggcagaa aattgccaaa gaggaagaga gtgaggatga 720
 aagcaatgaa gaggaggaag aggaagatga agaagagtca gagtccgagg caaaatcagt 780
 caaggtgaaa attaagctca ataaaaaaga tgacaaaggc cgggacaaag ggaaaggcaa 840
 gaaaaggcca aatcgaggaa aagccaaacc tgtagtgagc gattttgaca gcatgagga 900
 gcaggatgaa cgtgaacagt cagaaggaag tgggacggat gatgagtgat cagtatggac 960
 ctttttcctt ggtagaactg aattccttcc tcccctgtct cttttctacc cagttagttc 1020
 atttgtcata taggcactgg gttgtttcta tatcatcatc gtctataaac tagctttagg 1080
 atagtgccag acaaacatat gatatcatgg tgtaaaaaac acacacatac acaaataattt 1140
 gtaacatatt gtgaccaaag gggcctcaaa gattcagatt gaaacaaaca aaaagctttt 1200
 gatggaaaat atgtgggtgg atagtatatt tctatgggtg ggtctaattt ggtaacgggt 1260
 tgattgtgcc tggttttatc acctgttcag atgagaagat ttttgtcttt tgtagcactg 1320
 ataaccagga gaagccatta aaagccactg gttattttat ttttcatcag gcaattttcg 1380
 aggtttttat ttgttcggta ttgttttttt acactgtggt acatataagc aactttaata 1440
 ggtgataaat gtacagtagt tagatttcac ctgcatatac atttttccat tttatgctct 1500
 atgatctgaa caaaagcttt ttgaattgta taagatttat gtctactgta aacattgctt 1560
 aatttttttg ctcttgattt aaaaaaaagt tttgttgaaa gcgctattga atattgcaat 1620
 ctatatagtg tattggatgg cttcttttgt caccctgatc tcctatgtta ccaatgtgta 1680
 tcgtctcctt ctccctaaag tgtacttaat ctttgctttc tttgcacaat gtctttggtt 1740

```

gcaagtcata agcctgaggc aaataaaatt ccagtaattt cgaagaatgt ggtgttggtg 1800
ctttcctaataa aaagaaataa tttagcttga caaaaaaaaa aaaaaa 1846
<210> 5
<211> 1590
<212> PRT
<213> 智人
<400> 5
Met Ser Thr Pro Thr Asp Pro Gly Ala Met Pro His Pro Gly Pro Ser
1          5          10          15
Pro Gly Pro Gly Pro Ser Pro Gly Pro Ile Leu Gly Pro Ser Pro Gly
          20          25          30
Pro Gly Pro Ser Pro Gly Ser Val His Ser Met Met Gly Pro Ser Pro
          35          40          45
Gly Pro Pro Ser Val Ser His Pro Met Pro Thr Met Gly Ser Thr Asp
          50          55          60
Phe Pro Gln Glu Gly Met His Gln Met His Lys Pro Ile Asp Gly Ile
65          70          75          80
His Asp Lys Gly Ile Val Glu Asp Ile His Cys Gly Ser Met Lys Gly
          85          90          95
Thr Gly Met Arg Pro Pro His Pro Gly Met Gly Pro Pro Gln Ser Pro
          100          105          110
Met Asp Gln His Ser Gln Gly Tyr Met Ser Pro His Pro Ser Pro Leu
          115          120          125
Gly Ala Pro Glu His Val Ser Ser Pro Met Ser Gly Gly Gly Pro Thr
          130          135          140
Pro Pro Gln Met Pro Pro Ser Gln Pro Gly Ala Leu Ile Pro Gly Asp
145          150          155          160
Pro Gln Ala Met Ser Gln Pro Asn Arg Gly Pro Ser Pro Phe Ser Pro
          165          170          175
Val Gln Leu His Gln Leu Arg Ala Gln Ile Leu Ala Tyr Lys Met Leu
          180          185          190
Ala Arg Gly Gln Pro Leu Pro Glu Thr Leu Gln Leu Ala Val Gln Gly
          195          200          205
Lys Arg Thr Leu Pro Gly Leu Gln Gln Gln Gln Gln Gln Gln Gln Gln
210          215          220
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Gln
225          230          235          240
Gln Gln Pro Pro Gln Pro Gln Thr Gln Gln Gln Gln Gln Pro Ala Leu
          245          250          255

```

Val Asn Tyr Asn Arg Pro Ser Gly Pro Gly Pro Glu Leu Ser Gly Pro		
260	265	270
Ser Thr Pro Gln Lys Leu Pro Val Pro Ala Pro Gly Gly Arg Pro Ser		
275	280	285
Pro Ala Pro Pro Ala Ala Ala Gln Pro Pro Ala Ala Val Pro Gly		
290	295	300
Pro Ser Val Pro Gln Pro Ala Pro Gly Gln Pro Ser Pro Val Leu Gln		
305	310	315
Leu Gln Gln Lys Gln Ser Arg Ile Ser Pro Ile Gln Lys Pro Gln Gly		
325	330	335
Leu Asp Pro Val Glu Ile Leu Gln Glu Arg Glu Tyr Arg Leu Gln Ala		
340	345	350
Arg Ile Ala His Arg Ile Gln Glu Leu Glu Asn Leu Pro Gly Ser Leu		
355	360	365
Pro Pro Asp Leu Arg Thr Lys Ala Thr Val Glu Leu Lys Ala Leu Arg		
370	375	380
Leu Leu Asn Phe Gln Arg Gln Leu Arg Gln Glu Val Val Ala Cys Met		
385	390	395
Arg Arg Asp Thr Thr Leu Glu Thr Ala Leu Asn Ser Lys Ala Tyr Lys		
405	410	415
Arg Ser Lys Arg Gln Thr Leu Arg Glu Ala Arg Met Thr Glu Lys Leu		
420	425	430
Glu Lys Gln Gln Lys Ile Glu Gln Glu Arg Lys Arg Arg Gln Lys His		
435	440	445
Gln Glu Tyr Leu Asn Ser Ile Leu Gln His Ala Lys Asp Phe Lys Glu		
450	455	460
Tyr His Arg Ser Val Ala Gly Lys Ile Gln Lys Leu Ser Lys Ala Val		
465	470	475
Ala Thr Trp His Ala Asn Thr Glu Arg Glu Gln Lys Lys Glu Thr Glu		
485	490	495
Arg Ile Glu Lys Glu Arg Met Arg Arg Leu Met Ala Glu Asp Glu Glu		
500	505	510
Gly Tyr Arg Lys Leu Ile Asp Gln Lys Lys Asp Arg Arg Leu Ala Tyr		
515	520	525
Leu Leu Gln Gln Thr Asp Glu Tyr Val Ala Asn Leu Thr Asn Leu Val		
530	535	540
Trp Glu His Lys Gln Ala Gln Ala Ala Lys Glu Lys Lys Lys Arg Arg		
545	550	555
Arg Arg Lys Lys Lys Ala Glu Glu Asn Ala Glu Gly Gly Glu Ser Ala		

	565		570		575
Leu Gly Pro Asp Gly Glu Pro Ile Asp Glu Ser Ser Gln Met Ser Asp					
	580		585		590
Leu Pro Val Lys Val Thr His Thr Glu Thr Gly Lys Val Leu Phe Gly					
	595		600		605
Pro Glu Ala Pro Lys Ala Ser Gln Leu Asp Ala Trp Leu Glu Met Asn					
	610		615		620
Pro Gly Tyr Glu Val Ala Pro Arg Ser Asp Ser Glu Glu Ser Asp Ser					
625		630		635	640
Asp Tyr Glu Glu Glu Asp Glu Glu Glu Glu Ser Ser Arg Gln Glu Thr					
	645		650		655
Glu Glu Lys Ile Leu Leu Asp Pro Asn Ser Glu Glu Val Ser Glu Lys					
	660		665		670
Asp Ala Lys Gln Ile Ile Glu Thr Ala Lys Gln Asp Val Asp Asp Glu					
	675		680		685
Tyr Ser Met Gln Tyr Ser Ala Arg Gly Ser Gln Ser Tyr Tyr Thr Val					
	690		695		700
Ala His Ala Ile Ser Glu Arg Val Glu Lys Gln Ser Ala Leu Leu Ile					
705		710		715	720
Asn Gly Thr Leu Lys His Tyr Gln Leu Gln Gly Leu Glu Trp Met Val					
	725		730		735
Ser Leu Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly					
	740		745		750
Leu Gly Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu					
	755		760		765
His Lys Arg Leu Asn Gly Pro Tyr Leu Ile Ile Val Pro Leu Ser Thr					
	770		775		780
Leu Ser Asn Trp Thr Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val					
785		790		795	800
Lys Ile Ser Tyr Lys Gly Thr Pro Ala Met Arg Arg Ser Leu Val Pro					
	805		810		815
Gln Leu Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr					
	820		825		830
Ile Ile Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met					
	835		840		845
Ile Val Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr					
	850		855		860
Gln Val Leu Asn Thr His Tyr Val Ala Pro Arg Arg Ile Leu Leu Thr					
865		870		875	880

Gly Thr Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn		
885	890	895
Phe Leu Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp		
900	905	910
Phe Asn Ala Pro Phe Ala Met Thr Gly Glu Arg Val Asp Leu Asn Glu		
915	920	925
Glu Glu Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro		
930	935	940
Phe Leu Leu Arg Arg Leu Lys Lys Glu Val Glu Ser Gln Leu Pro Glu		
945	950	955
Lys Val Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Lys Ile		
965	970	975
Leu Tyr Arg His Met Gln Ala Lys Gly Ile Leu Leu Thr Asp Gly Ser		
980	985	990
Glu Lys Asp Lys Lys Gly Lys Gly Gly Ala Lys Thr Leu Met Asn Thr		
995	1000	1005
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1010	1015	1020
His Ile Glu Glu Ser Phe Ala Glu His Leu Gly Tyr Ser Asn Gly		
1025	1030	1035
Val Ile Asn Gly Ala Glu Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1040	1045	1050
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Arg		
1055	1060	1065
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
1070	1075	1080
Asp Tyr Phe Ala Phe Arg Asn Phe Leu Tyr Leu Arg Leu Asp Gly		
1085	1090	1095
Thr Thr Lys Ser Glu Asp Arg Ala Ala Leu Leu Lys Lys Phe Asn		
1100	1105	1110
Glu Pro Gly Ser Gln Tyr Phe Ile Phe Leu Leu Ser Thr Arg Ala		
1115	1120	1125
Gly Gly Leu Gly Leu Asn Leu Gln Ala Ala Asp Thr Val Val Ile		
1130	1135	1140
Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala Gln Asp		
1145	1150	1155
Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val Leu Arg		
1160	1165	1170
Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala Ala Ala		

1175	1180	1185
Lys Tyr Lys Leu Asn Val	Asp Gln Lys Val Ile	Gln Ala Gly Met
1190	1195	1200
Phe Asp Gln Lys Ser Ser	Ser His Glu Arg Arg	Ala Phe Leu Gln
1205	1210	1215
Ala Ile Leu Glu His Glu	Glu Glu Asn Glu Glu	Glu Asp Glu Val
1220	1225	1230
Pro Asp Asp Glu Thr Leu	Asn Gln Met Ile Ala	Arg Arg Glu Glu
1235	1240	1245
Glu Phe Asp Leu Phe Met	Arg Met Asp Met Asp	Arg Arg Arg Glu
1250	1255	1260
Asp Ala Arg Asn Pro Lys	Arg Lys Pro Arg Leu	Met Glu Glu Asp
1265	1270	1275
Glu Leu Pro Ser Trp Ile	Ile Lys Asp Asp Ala	Glu Val Glu Arg
1280	1285	1290
Leu Thr Cys Glu Glu Glu	Glu Glu Lys Ile Phe	Gly Arg Gly Ser
1295	1300	1305
Arg Gln Arg Arg Asp Val	Asp Tyr Ser Asp Ala	Leu Thr Glu Lys
1310	1315	1320
Gln Trp Leu Arg Ala Ile	Glu Asp Gly Asn Leu	Glu Glu Met Glu
1325	1330	1335
Glu Glu Val Arg Leu Lys	Lys Arg Lys Arg Arg	Arg Asn Val Asp
1340	1345	1350
Lys Asp Pro Ala Lys Glu	Asp Val Glu Lys Ala	Lys Lys Arg Arg
1355	1360	1365
Gly Arg Pro Pro Ala Glu	Lys Leu Ser Pro Asn	Pro Pro Lys Leu
1370	1375	1380
Thr Lys Gln Met Asn Ala	Ile Ile Asp Thr Val	Ile Asn Tyr Lys
1385	1390	1395
Asp Arg Cys Asn Val Glu	Lys Val Pro Ser Asn	Ser Gln Leu Glu
1400	1405	1410
Ile Glu Gly Asn Ser Ser	Gly Arg Gln Leu Ser	Glu Val Phe Ile
1415	1420	1425
Gln Leu Pro Ser Arg Lys	Glu Leu Pro Glu Tyr	Tyr Glu Leu Ile
1430	1435	1440
Arg Lys Pro Val Asp Phe	Lys Lys Ile Lys Glu	Arg Ile Arg Asn
1445	1450	1455
His Lys Tyr Arg Ser Leu	Gly Asp Leu Glu Lys	Asp Val Met Leu
1460	1465	1470

Leu Cys His Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Gln Ile		
1475	1480	1485
Tyr Glu Asp Ser Ile Val Leu Gln Ser Val Phe Lys Ser Ala Arg		
1490	1495	1500
Gln Lys Ile Ala Lys Glu Glu Glu Ser Glu Asp Glu Ser Asn Glu		
1505	1510	1515
Glu Glu Glu Glu Glu Asp Glu Glu Glu Ser Glu Ser Glu Ala Lys		
1520	1525	1530
Ser Val Lys Val Lys Ile Lys Leu Asn Lys Lys Asp Asp Lys Gly		
1535	1540	1545
Arg Asp Lys Gly Lys Gly Lys Lys Arg Pro Asn Arg Gly Lys Ala		
1550	1555	1560
Lys Pro Val Val Ser Asp Phe Asp Ser Asp Glu Glu Gln Asp Glu		
1565	1570	1575
Arg Glu Gln Ser Glu Gly Ser Gly Thr Asp Asp Glu		
1580	1585	1590

<210> 6

<211> 1572

<212> PRT

<213> 智人

<400> 6

Met Ser Thr Pro Thr Asp Pro Gly Ala Met Pro His Pro Gly Pro Ser		
1	5	10
Pro Gly Pro Gly Pro Ser Pro Gly Pro Ile Leu Gly Pro Ser Pro Gly		
20	25	30
Pro Gly Pro Ser Pro Gly Ser Val His Ser Met Met Gly Pro Ser Pro		
35	40	45
Gly Pro Pro Ser Val Ser His Pro Met Pro Thr Met Gly Ser Thr Asp		
50	55	60
Phe Pro Gln Glu Gly Met His Gln Met His Lys Pro Ile Asp Gly Ile		
65	70	75
His Asp Lys Gly Ile Val Glu Asp Ile His Cys Gly Ser Met Lys Gly		
85	90	95
Thr Gly Met Arg Pro Pro His Pro Gly Met Gly Pro Pro Gln Ser Pro		
100	105	110
Met Asp Gln His Ser Gln Gly Tyr Met Ser Pro His Pro Ser Pro Leu		
115	120	125
Gly Ala Pro Glu His Val Ser Ser Pro Met Ser Gly Gly Gly Pro Thr		
130	135	140

Pro	Pro	Gln	Met	Pro	Pro	Ser	Gln	Pro	Gly	Ala	Leu	Ile	Pro	Gly	Asp
145						150				155					160
Pro	Gln	Ala	Met	Ser	Gln	Pro	Asn	Arg	Gly	Pro	Ser	Pro	Phe	Ser	Pro
				165					170					175	
Val	Gln	Leu	His	Gln	Leu	Arg	Ala	Gln	Ile	Leu	Ala	Tyr	Lys	Met	Leu
			180					185					190		
Ala	Arg	Gly	Gln	Pro	Leu	Pro	Glu	Thr	Leu	Gln	Leu	Ala	Val	Gln	Gly
		195					200					205			
Lys	Arg	Thr	Leu	Pro	Gly	Leu	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln
	210					215					220				
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Pro	Gln
225					230					235					240
Gln	Gln	Pro	Pro	Gln	Pro	Gln	Thr	Gln	Gln	Gln	Gln	Gln	Pro	Ala	Leu
				245					250					255	
Val	Asn	Tyr	Asn	Arg	Pro	Ser	Gly	Pro	Gly	Pro	Glu	Leu	Ser	Gly	Pro
			260					265					270		
Ser	Thr	Pro	Gln	Lys	Leu	Pro	Val	Pro	Ala	Pro	Gly	Gly	Arg	Pro	Ser
		275					280					285			
Pro	Ala	Pro	Pro	Ala	Ala	Ala	Gln	Pro	Pro	Ala	Ala	Ala	Val	Pro	Gly
	290					295					300				
Pro	Ser	Val	Pro	Gln	Pro	Ala	Pro	Gly	Gln	Pro	Ser	Pro	Val	Leu	Gln
305				310					315					320	
Leu	Gln	Gln	Lys	Gln	Ser	Arg	Ile	Ser	Pro	Ile	Gln	Lys	Pro	Gln	Gly
			325						330				335		
Leu	Asp	Pro	Val	Glu	Ile	Leu	Gln	Glu	Arg	Glu	Tyr	Arg	Leu	Gln	Ala
		340						345					350		
Arg	Ile	Ala	His	Arg	Ile	Gln	Glu	Leu	Glu	Asn	Leu	Pro	Gly	Ser	Leu
	355					360						365			
Pro	Pro	Asp	Leu	Arg	Thr	Lys	Ala	Thr	Val	Glu	Leu	Lys	Ala	Leu	Arg
	370					375					380				
Leu	Leu	Asn	Phe	Gln	Arg	Gln	Leu	Arg	Gln	Glu	Val	Val	Ala	Cys	Met
385				390						395				400	
Arg	Arg	Asp	Thr	Thr	Leu	Glu	Thr	Ala	Leu	Asn	Ser	Lys	Ala	Tyr	Lys
			405						410				415		
Arg	Ser	Lys	Arg	Gln	Thr	Leu	Arg	Glu	Ala	Arg	Met	Thr	Glu	Lys	Leu
		420						425				430			
Glu	Lys	Gln	Gln	Lys	Ile	Glu	Gln	Glu	Arg	Lys	Arg	Arg	Gln	Lys	His
	435					440						445			
Gln	Glu	Tyr	Leu	Asn	Ser	Ile	Leu	Gln	His	Ala	Lys	Asp	Phe	Lys	Glu

450	455	460
Tyr His Arg Ser Val	Ala Gly Lys Ile Gln Lys Leu Ser Lys Ala Val	
465	470	475
Ala Thr Trp His Ala	Asn Thr Glu Arg Glu Gln Lys Lys Glu Thr Glu	480
	485	490
Arg Ile Glu Lys Glu	Arg Met Arg Arg Leu Met Ala Glu Asp Glu Glu	495
	500	505
Gly Tyr Arg Lys Leu	Ile Asp Gln Lys Lys Asp Arg Arg Leu Ala Tyr	510
	515	520
Leu Leu Gln Gln Thr	Asp Glu Tyr Val Ala Asn Leu Thr Asn Leu Val	525
	530	535
Trp Glu His Lys Gln	Ala Gln Ala Ala Lys Glu Lys Lys Lys Arg Arg	540
545	550	555
Arg Arg Lys Lys Lys	Ala Glu Glu Asn Ala Glu Gly Gly Glu Ser Ala	560
	565	570
Leu Gly Pro Asp Gly	Glu Pro Ile Asp Glu Ser Ser Gln Met Ser Asp	575
	580	585
Leu Pro Val Lys Val	Thr His Thr Glu Thr Gly Lys Val Leu Phe Gly	590
	595	600
Pro Glu Ala Pro Lys	Ala Ser Gln Leu Asp Ala Trp Leu Glu Met Asn	605
	610	615
Pro Gly Tyr Glu Val	Ala Pro Arg Ser Asp Ser Glu Glu Ser Asp Ser	620
625	630	635
Asp Tyr Glu Glu Glu	Asp Glu Glu Glu Glu Ser Ser Arg Gln Glu Thr	640
	645	650
Glu Glu Lys Ile Leu	Leu Asp Pro Asn Ser Glu Glu Val Ser Glu Lys	655
	660	665
Asp Ala Lys Gln Ile	Ile Glu Thr Ala Lys Gln Asp Val Asp Asp Glu	670
	675	680
Tyr Ser Met Gln Tyr	Ser Ala Arg Gly Ser Gln Ser Tyr Tyr Thr Val	685
	690	695
Ala His Ala Ile Ser	Glu Arg Val Glu Lys Gln Ser Ala Leu Leu Ile	700
705	710	715
Asn Gly Thr Leu Lys	His Tyr Gln Leu Gln Gly Leu Glu Trp Met Val	720
	725	730
Ser Leu Tyr Asn Asn	Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly	735
	740	745
Leu Gly Lys Thr Ile	Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu	750
	755	760
		765

His Lys Arg Leu Asn Gly Pro Tyr Leu Ile Ile Val Pro Leu Ser Thr		
770	775	780
Leu Ser Asn Trp Thr Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val		
785	790	795 800
Lys Ile Ser Tyr Lys Gly Thr Pro Ala Met Arg Arg Ser Leu Val Pro		
805	810	815
Gln Leu Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr		
820	825	830
Ile Ile Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met		
835	840	845
Ile Val Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr		
850	855	860
Gln Val Leu Asn Thr His Tyr Val Ala Pro Arg Arg Ile Leu Leu Thr		
865	870	875 880
Gly Thr Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn		
885	890	895
Phe Leu Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp		
900	905	910
Phe Asn Ala Pro Phe Ala Met Thr Gly Glu Arg Val Asp Leu Asn Glu		
915	920	925
Glu Glu Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro		
930	935	940
Phe Leu Leu Arg Arg Leu Lys Lys Glu Val Glu Ser Gln Leu Pro Glu		
945	950	955 960
Lys Val Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Lys Ile		
965	970	975
Leu Tyr Arg His Met Gln Ala Lys Gly Ile Leu Leu Thr Asp Gly Ser		
980	985	990
Glu Lys Asp Lys Lys Gly Lys Gly Gly Ala Lys Thr Leu Met Asn Thr		
995	1000	1005
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1010	1015	1020
His Ile Glu Glu Ser Phe Ala Glu His Leu Gly Tyr Ser Asn Gly		
1025	1030	1035
Val Ile Asn Gly Ala Glu Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1040	1045	1050
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Arg		
1055	1060	1065
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		

1070	1075	1080
Asp Tyr Phe Ala Phe Arg	Asn Phe Leu Tyr Leu	Arg Leu Asp Gly
1085	1090	1095
Thr Thr Lys Ser Glu Asp	Arg Ala Ala Leu Leu	Lys Lys Phe Asn
1100	1105	1110
Glu Pro Gly Ser Gln Tyr	Phe Ile Phe Leu Leu	Ser Thr Arg Ala
1115	1120	1125
Gly Gly Leu Gly Leu Asn	Leu Gln Ala Ala Asp	Thr Val Val Ile
1130	1135	1140
Phe Asp Ser Asp Trp Asn	Pro His Gln Asp Leu	Gln Ala Gln Asp
1145	1150	1155
Arg Ala His Arg Ile Gly	Gln Gln Asn Glu Val	Arg Val Leu Arg
1160	1165	1170
Leu Cys Thr Val Asn Ser	Val Glu Glu Lys Ile	Leu Ala Ala Ala
1175	1180	1185
Lys Tyr Lys Leu Asn Val	Asp Gln Lys Val Ile	Gln Ala Gly Met
1190	1195	1200
Phe Asp Gln Lys Ser Ser	Ser His Glu Arg Arg	Ala Phe Leu Gln
1205	1210	1215
Ala Ile Leu Glu His Glu	Glu Glu Asn Glu Glu	Glu Asp Glu Val
1220	1225	1230
Pro Asp Asp Glu Thr Leu	Asn Gln Met Ile Ala	Arg Arg Glu Glu
1235	1240	1245
Glu Phe Asp Leu Phe Met	Arg Met Asp Met Asp	Arg Arg Arg Glu
1250	1255	1260
Asp Ala Arg Asn Pro Lys	Arg Lys Pro Arg Leu	Met Glu Glu Asp
1265	1270	1275
Glu Leu Pro Ser Trp Ile	Ile Lys Asp Asp Ala	Glu Val Glu Arg
1280	1285	1290
Leu Thr Cys Glu Glu Glu	Glu Glu Lys Ile Phe	Gly Arg Gly Ser
1295	1300	1305
Arg Gln Arg Arg Asp Val	Asp Tyr Ser Asp Ala	Leu Thr Glu Lys
1310	1315	1320
Gln Trp Leu Arg Ala Ile	Glu Asp Gly Asn Leu	Glu Glu Met Glu
1325	1330	1335
Glu Glu Val Arg Leu Lys	Lys Arg Lys Arg Arg	Arg Asn Val Asp
1340	1345	1350
Lys Asp Pro Ala Lys Glu	Asp Val Glu Lys Ala	Lys Lys Arg Arg
1355	1360	1365

Gly Arg Pro Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro Lys Leu		
1370	1375	1380
Thr Lys Gln Met Asn Ala Ile Ile Asp Thr Val Ile Asn Tyr Lys		
1385	1390	1395
Asp Ser Ser Gly Arg Gln Leu Ser Glu Val Phe Ile Gln Leu Pro		
1400	1405	1410
Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile Arg Lys Pro		
1415	1420	1425
Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn His Lys Tyr		
1430	1435	1440
Arg Ser Leu Gly Asp Leu Glu Lys Asp Val Met Leu Leu Cys His		
1445	1450	1455
Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Gln Ile Tyr Glu Asp		
1460	1465	1470
Ser Ile Val Leu Gln Ser Val Phe Lys Ser Ala Arg Gln Lys Ile		
1475	1480	1485
Ala Lys Glu Glu Glu Ser Glu Asp Glu Ser Asn Glu Glu Glu Glu		
1490	1495	1500
Glu Glu Asp Glu Glu Glu Ser Glu Ser Glu Ala Lys Ser Val Lys		
1505	1510	1515
Val Lys Ile Lys Leu Asn Lys Lys Asp Asp Lys Gly Arg Asp Lys		
1520	1525	1530
Gly Lys Gly Lys Lys Arg Pro Asn Arg Gly Lys Ala Lys Pro Val		
1535	1540	1545
Val Ser Asp Phe Asp Ser Asp Glu Glu Gln Asp Glu Arg Glu Gln		
1550	1555	1560
Ser Glu Gly Ser Gly Thr Asp Asp Glu		
1565	1570	

<210> 7

<211> 1514

<212> PRT

<213> 智人

<400> 7

Met Ser Thr Pro Thr Asp Pro Gly Ala Met Pro His Pro Gly Pro Ser		
1	5	10
Pro Gly Pro Gly Pro Ser Pro Gly Pro Ile Leu Gly Pro Ser Pro Gly		
20	25	30
Pro Gly Pro Ser Pro Gly Ser Val His Ser Met Met Gly Pro Ser Pro		
35	40	45

Gly Pro Pro Ser Val Ser His Pro Met Pro Thr Met Gly Ser Thr Asp		
50	55	60
Phe Pro Gln Glu Gly Met His Gln Met His Lys Pro Ile Asp Gly Ile		
65	70	75
His Asp Lys Gly Ile Val Glu Asp Ile His Cys Gly Ser Met Lys Gly		
85	90	95
Thr Gly Met Arg Pro Pro His Pro Gly Met Gly Pro Pro Gln Ser Pro		
100	105	110
Met Asp Gln His Ser Gln Gly Tyr Met Ser Pro His Pro Ser Pro Leu		
115	120	125
Gly Ala Pro Glu His Val Ser Ser Pro Met Ser Gly Gly Gly Pro Thr		
130	135	140
Pro Pro Gln Met Pro Pro Ser Gln Pro Gly Ala Leu Ile Pro Gly Asp		
145	150	155
Pro Gln Ala Met Ser Gln Pro Asn Arg Gly Pro Ser Pro Phe Ser Pro		
165	170	175
Val Gln Leu His Gln Leu Arg Ala Gln Ile Leu Ala Tyr Lys Met Leu		
180	185	190
Ala Arg Gly Gln Pro Leu Pro Glu Thr Leu Gln Leu Ala Val Gln Gly		
195	200	205
Lys Arg Thr Leu Pro Gly Leu Gln Gln Gln Gln Gln Gln Gln Gln Gln		
210	215	220
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Gln		
225	230	235
Gln Gln Pro Pro Gln Pro Gln Thr Gln Gln Gln Gln Gln Pro Ala Leu		
245	250	255
Val Asn Tyr Asn Arg Pro Ser Gly Pro Gly Pro Glu Leu Ser Gly Pro		
260	265	270
Ser Thr Pro Gln Lys Leu Pro Val Pro Ala Pro Gly Gly Arg Pro Ser		
275	280	285
Pro Ala Pro Pro Ala Ala Ala Gln Pro Pro Ala Ala Ala Val Pro Gly		
290	295	300
Pro Ser Val Pro Gln Pro Ala Pro Gly Gln Pro Ser Pro Val Leu Gln		
305	310	315
Leu Gln Gln Lys Gln Ser Arg Ile Ser Pro Ile Gln Lys Pro Gln Gly		
325	330	335
Leu Asp Pro Val Glu Ile Leu Gln Glu Arg Glu Tyr Arg Leu Gln Ala		
340	345	350
Arg Ile Ala His Arg Ile Gln Glu Leu Glu Asn Leu Pro Gly Ser Leu		

355	360	365
Pro Pro Asp Leu Arg Thr Lys	Ala Thr Val Glu Leu Lys	Ala Leu Arg
370	375	380
Leu Leu Asn Phe Gln Arg Gln	Leu Arg Gln Glu Val Val	Ala Cys Met
385	390	395
Arg Arg Asp Thr Thr Leu Glu	Thr Ala Leu Asn Ser Lys	Ala Tyr Lys
405	410	415
Arg Ser Lys Arg Gln Thr Leu	Arg Glu Ala Arg Met Thr	Glu Lys Leu
420	425	430
Glu Lys Gln Gln Lys Ile Glu	Gln Glu Arg Lys Arg Arg	Gln Lys His
435	440	445
Gln Glu Tyr Leu Asn Ser Ile	Leu Gln His Ala Lys Asp	Phe Lys Glu
450	455	460
Tyr His Arg Ser Val Ala Gly	Lys Ile Gln Lys Leu Ser	Lys Ala Val
465	470	475
Ala Thr Trp His Ala Asn Thr	Glu Arg Glu Gln Lys Lys	Glu Thr Glu
485	490	495
Arg Ile Glu Lys Glu Arg Met	Arg Arg Leu Met Ala Glu	Asp Glu Glu
500	505	510
Gly Tyr Arg Lys Leu Ile Asp	Gln Lys Lys Asp Arg Arg	Leu Ala Tyr
515	520	525
Leu Leu Gln Gln Thr Asp Glu	Tyr Val Ala Asn Leu Thr	Asn Leu Val
530	535	540
Trp Glu His Lys Gln Ala Gln	Ala Ala Lys Glu Lys Lys	Lys Arg Arg
545	550	555
Arg Arg Lys Lys Lys Ala Glu	Glu Asn Ala Glu Gly Gly	Glu Ser Ala
565	570	575
Leu Gly Pro Asp Gly Glu Pro	Ile Asp Glu Ser Ser Gln	Met Ser Asp
580	585	590
Leu Pro Val Lys Val Thr His	Thr Glu Thr Gly Lys Val	Leu Phe Gly
595	600	605
Pro Glu Ala Pro Lys Ala Ser	Gln Leu Asp Ala Trp Leu	Glu Met Asn
610	615	620
Pro Gly Tyr Glu Val Ala Pro	Arg Ser Asp Ser Glu Glu	Ser Asp Ser
625	630	635
Asp Tyr Glu Glu Glu Asp Glu	Glu Glu Glu Ser Ser Arg	Gln Glu Thr
645	650	655
Glu Glu Lys Ile Leu Leu Asp	Pro Asn Ser Glu Glu Val	Ser Glu Lys
660	665	670

Asp Ala Lys Gln Ile Ile Glu Thr Ala Lys Gln Asp Val Asp Asp Glu		
675	680	685
Tyr Ser Met Gln Tyr Ser Ala Arg Gly Ser Gln Ser Tyr Tyr Thr Val		
690	695	700
Ala His Ala Ile Ser Glu Arg Val Glu Lys Gln Ser Ala Leu Leu Ile		
705	710	715
Asn Gly Thr Leu Lys His Tyr Gln Leu Gln Gly Leu Glu Trp Met Val		
725	730	735
Ser Leu Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly		
740	745	750
Leu Gly Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu		
755	760	765
His Lys Arg Leu Asn Gly Pro Tyr Leu Ile Ile Val Pro Leu Ser Thr		
770	775	780
Leu Ser Asn Trp Thr Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val		
785	790	795
Lys Ile Ser Tyr Lys Gly Thr Pro Ala Met Arg Arg Ser Leu Val Pro		
805	810	815
Gln Leu Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr		
820	825	830
Ile Ile Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met		
835	840	845
Ile Val Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr		
850	855	860
Gln Val Asp Leu Asn Glu Glu Glu Thr Ile Leu Ile Ile Arg Arg Leu		
865	870	875
His Lys Val Leu Arg Pro Phe Leu Leu Arg Arg Leu Lys Lys Glu Val		
885	890	895
Glu Ser Gln Leu Pro Glu Lys Val Glu Tyr Val Ile Lys Cys Asp Met		
900	905	910
Ser Ala Leu Gln Lys Ile Leu Tyr Arg His Met Gln Ala Lys Gly Ile		
915	920	925
Leu Leu Thr Asp Gly Ser Glu Lys Asp Lys Lys Gly Lys Gly Gly Ala		
930	935	940
Lys Thr Leu Met Asn Thr Ile Met Gln Leu Arg Lys Ile Cys Asn His		
945	950	955
Pro Tyr Met Phe Gln His Ile Glu Glu Ser Phe Ala Glu His Leu Gly		
965	970	975
Tyr Ser Asn Gly Val Ile Asn Gly Ala Glu Leu Tyr Arg Ala Ser Gly		

980	985	990
Lys Phe Glu Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn		
995	1000	1005
His Arg Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile		
1010	1015	1020
Met Glu Asp Tyr Phe Ala Phe Arg Asn Phe Leu Tyr Leu Arg Leu		
1025	1030	1035
Asp Gly Thr Thr Lys Ser Glu Asp Arg Ala Ala Leu Leu Lys Lys		
1040	1045	1050
Phe Asn Glu Pro Gly Ser Gln Tyr Phe Ile Phe Leu Leu Ser Thr		
1055	1060	1065
Arg Ala Gly Gly Leu Gly Leu Asn Leu Gln Ala Ala Asp Thr Val		
1070	1075	1080
Val Ile Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala		
1085	1090	1095
Gln Asp Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val		
1100	1105	1110
Leu Arg Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala		
1115	1120	1125
Ala Ala Lys Tyr Lys Leu Asn Val Asp Gln Lys Val Ile Gln Ala		
1130	1135	1140
Gly Met Phe Asp Gln Lys Ser Ser Ser His Glu Arg Arg Ala Phe		
1145	1150	1155
Leu Gln Ala Ile Leu Glu His Glu Glu Glu Asn Glu Glu Glu Asp		
1160	1165	1170
Glu Val Pro Asp Asp Glu Thr Leu Asn Gln Met Ile Ala Arg Arg		
1175	1180	1185
Glu Glu Glu Phe Asp Leu Phe Met Arg Met Asp Met Asp Arg Arg		
1190	1195	1200
Arg Glu Asp Ala Arg Asn Pro Lys Arg Lys Pro Arg Leu Met Glu		
1205	1210	1215
Glu Asp Glu Leu Pro Ser Trp Ile Ile Lys Asp Asp Ala Glu Val		
1220	1225	1230
Glu Arg Leu Thr Cys Glu Glu Glu Glu Glu Lys Ile Phe Gly Arg		
1235	1240	1245
Gly Ser Arg Gln Arg Arg Asp Val Asp Tyr Ser Asp Ala Leu Thr		
1250	1255	1260
Glu Lys Gln Trp Leu Arg Ala Ile Glu Asp Gly Asn Leu Glu Glu		
1265	1270	1275

Met Glu Glu Glu Val Arg Leu Lys Lys Arg Lys Arg Arg Arg Asn		
1280	1285	1290
Val Asp Lys Asp Pro Ala Lys Glu Asp Val Glu Lys Ala Lys Lys		
1295	1300	1305
Arg Arg Gly Arg Pro Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro		
1310	1315	1320
Lys Leu Thr Lys Gln Met Asn Ala Ile Ile Asp Thr Val Ile Asn		
1325	1330	1335
Tyr Lys Asp Ser Ser Gly Arg Gln Leu Ser Glu Val Phe Ile Gln		
1340	1345	1350
Leu Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile Arg		
1355	1360	1365
Lys Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn His		
1370	1375	1380
Lys Tyr Arg Ser Leu Gly Asp Leu Glu Lys Asp Val Met Leu Leu		
1385	1390	1395
Cys His Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Gln Ile Tyr		
1400	1405	1410
Glu Asp Ser Ile Val Leu Gln Ser Val Phe Lys Ser Ala Arg Gln		
1415	1420	1425
Lys Ile Ala Lys Glu Glu Glu Ser Glu Asp Glu Ser Asn Glu Glu		
1430	1435	1440
Glu Glu Glu Glu Asp Glu Glu Glu Ser Glu Ser Glu Ala Lys Ser		
1445	1450	1455
Val Lys Val Lys Ile Lys Leu Asn Lys Lys Asp Asp Lys Gly Arg		
1460	1465	1470
Asp Lys Gly Lys Gly Lys Lys Arg Pro Asn Arg Gly Lys Ala Lys		
1475	1480	1485
Pro Val Val Ser Asp Phe Asp Ser Asp Glu Glu Gln Asp Glu Arg		
1490	1495	1500
Glu Gln Ser Glu Gly Ser Gly Thr Asp Asp Glu		
1505	1510	
<210> 8		
<211> 248		
<212> PRT		
<213> 智人		
<400> 8		
Met Trp Leu Ala Ile Glu Asp Gly Asn Leu Glu Glu Met Glu Glu Glu		
1	5	10
		15

Val Arg Leu Lys Lys Arg Lys Arg Arg Arg Asn Val Asp Lys Asp Pro
 20 25 30
 Ala Lys Glu Asp Val Glu Lys Ala Lys Lys Arg Arg Gly Arg Pro Pro
 35 40 45
 Ala Glu Lys Leu Ser Pro Asn Pro Pro Lys Leu Thr Lys Gln Met Asn
 50 55 60
 Ala Ile Ile Asp Thr Val Ile Asn Tyr Lys Asp Ser Ser Gly Arg Gln
 65 70 75 80
 Leu Ser Glu Val Phe Ile Gln Leu Pro Ser Arg Lys Glu Leu Pro Glu
 85 90 95
 Tyr Tyr Glu Leu Ile Arg Lys Pro Val Asp Phe Lys Lys Ile Lys Glu
 100 105 110
 Arg Ile Arg Asn His Lys Tyr Arg Ser Leu Gly Asp Leu Glu Lys Asp
 115 120 125
 Val Met Leu Leu Cys His Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser
 130 135 140
 Gln Ile Tyr Glu Asp Ser Ile Val Leu Gln Ser Val Phe Lys Ser Ala
 145 150 155 160
 Arg Gln Lys Ile Ala Lys Glu Glu Glu Ser Glu Asp Glu Ser Asn Glu
 165 170 175
 Glu Glu Glu Glu Glu Asp Glu Glu Glu Ser Glu Ser Glu Ala Lys Ser
 180 185 190
 Val Lys Val Lys Ile Lys Leu Asn Lys Lys Asp Asp Lys Gly Arg Asp
 195 200 205
 Lys Gly Lys Gly Lys Lys Arg Pro Asn Arg Gly Lys Ala Lys Pro Val
 210 215 220
 Val Ser Asp Phe Asp Ser Asp Glu Glu Gln Asp Glu Arg Glu Gln Ser
 225 230 235 240
 Glu Gly Ser Gly Thr Asp Asp Glu
 245

<210> 9

<211> 5589

<212> DNA

<213> 智人

<400> 9

ggcgggggag gcgccgggaa gtcgacggcg ccggcgggctc ctgcaggagg ccaactgtctg 60
 cagctcccggt gaagatgtcc actccagacc caccctggg cggaactcct cggccaggtc 120
 cttccccggg ccctggccct tcccctggag ccatgctggg ccctagcccg ggctccctcgc 180
 cgggctccgc ccacagcatg atggggccca gcccagggcc gccctcagca ggacacccca 240

tccccaccca ggggcctgga ggggtaccctc aggacaacat gcaccagatg cacaagccca 300
 tggagtccat gcatgagaag ggcatgtcgg acgacccgcg ctacaaccag atgaaaggaa 360
 tggggatgcg gtcagggggc catgctggga tggggccccc gccagcccc atggaccagc 420
 actcccaagg ttaccctcgc cccctgggtg gctctgagca tgcctctagt ccagttccag 480
 ccagtggccc gtcttcgggg ccccagatgt cttccgggcc aggaggtgcc ccgctggatg 540
 gtgctgaccc ccaggccttg gggcagcaga accggggccc aacccattt aaccagaacc 600
 agctgcacca gctcagagct cagatcatgg cctacaagat gctggccagg gggcagcccc 660
 tccccgacca cctgcagatg gcggtgcagg gcaagcggcc gatgcccggg atgcagcagc 720
 agatgccaac gctacctcca ccctcggtgt ccgcaacagg acccgccctt ggccctggcc 780
 ctggccccgg cccgggtccc ggcccggcac ctccaaatta cagcaggcct catggtatgg 840
 gagggcccaa catgcctccc ccaggacct cgggcgtgcc ccccgggatg ccaggccagc 900
 ctcttgaggg gcctcccaag ccctggcctg aaggacctt ggccaatgct gctgccccca 960
 cgagcacccc tcagaagctg attccccgcg agccaacggg ccgcccttcc cccgcgcccc 1020
 ctgccgtccc acccgccgcc tcgcccgtga tgccaccgca gaccagtc cccgggcagc 1080
 cggcccagcc cgcgcccatt gtgccactgc accagaagca gagccgcac acccccatcc 1140
 agaagccgcg gggcctcgac cctgtggaga tcctgcagga gcgcgagtag aggctgcagg 1200
 ctgcacatgc acaccgaatt caggaaattg aaaaccttcc cgggtccctg gccggggatt 1260
 tgcgaaccaa agcgaccatt gagtcaagg ccctcaggct gctgaacttc cagaggcagc 1320
 tgcgccagga ggtggtggtg tgcatgcgga gggacacagc gctggagaca gccctcaatg 1380
 ctaaggccta caagcgcagc aagcgccagt ccctgcgcga ggcccgcac actgagaagc 1440
 tggagaagca gcagaagatc gagcaggagc gcaagcggcg gcagaagcac caggaatacc 1500
 tcaatagcat tctccagcat gccaaggatt tcaaggaata tcacagatcc gtcacaggca 1560
 aaatccagaa gctgaccaag gcagtggcca cgtaccatgc caacacggag cgggagcaga 1620
 agaaagagaa cgagcggatc gagaaggagc gcatgcggag gctcatggct gaagatgagg 1680
 aggggtaccg caagctcatc gaccagaaga aggacaagcg cctggcctac ctcttgacgc 1740
 agacagacga gtacgtggct aacctcacgg agctggtgcg gcagcacaag gctgcccagg 1800
 tcgccaagga gaaaaagaag aaaaagaaaa agaagaaggc agaaaatgca gaaggacaga 1860
 cgcctgccat tgggccggat ggcgagcctc tggacgagac cagccagatg agcgacctcc 1920
 cgggtgaagg gatccacgtg gagagtggga agatcctcac aggcacagat gcccccaaag 1980
 ccgggcagct ggaggcctgg ctcgagatga acccggggta tgaagtagct ccgaggtctg 2040
 atagtgaaga aagtggctca gaagaagagg aagaggagga ggaggaagag cagccgcagg 2100
 cagcacagcc tcccacctg cccgtggagg agaagaagaa gattccagat ccagacagcg 2160
 atgacgtctc tgaggtggac gcgcggcaca tcattgagaa tgccaagcaa gatgtcgatg 2220
 atgaatatgg cgtgtcccag gcccttgac gtggcctgca gtctactat gccgtggccc 2280
 atgctgtcac tgagagagt gacaagcagt cagcgcttat ggtcaatggt gtctcaaac 2340
 agtaccagat caaaggtttg gattggctgg tgtccctgta caacaacaac ctgaacggca 2400
 tcctggccga cgagatgggc ctggggaaga ccatccagac catcgcgtc atcacgtacc 2460
 tcatggagca caaacgcac aatgggcctt tcctcatcat cgtgcctctc tcaacgtgt 2520
 ccaactgggc gtacgagttt gacaagtggg cccctccgt ggtgaaggtg tcttacaagg 2580

gatccccagc agcaagacgg gcctttgtcc cccagctccg gagtgggaag ttcaacgtct 2640
 tgctgacgac gtacgagtac atcatcaaag acaagcacat cctcgccaag atccgttgga 2700
 agtacatgat tgtggacgaa ggtcaccgca tgaagaacca ccaactgcaag ctgacgcagg 2760
 tgctcaacac gcactatgtg gcaccccgcc gcctgctgct gacgggcaca ccgctgcaga 2820
 acaagcttcc cgagctctgg gcgctgctca acttcctgct gccaccatc ttcaagagct 2880
 gcagcacctt cgagcagtgg tttaacgcac cttttgccat gaccggggaa aaggtggacc 2940
 tgaatgagga ggaaaccatt ctcatcatcc ggcgtctcca caaagtgtg cggcccttct 3000
 tgctccgacg actcaagaag gaagtcgagg ccagttgcc cgaaaagggtg gagtacgtca 3060
 tcaagtgcga catgtctgcg ctgcagcgag tgctctaccg ccacatgcag gccaaaggcg 3120
 tgctgctgac tgatggctcc gagaaggaca agaagggcaa aggcggcacc aagaccctga 3180
 tgaacaccat catgcagctg cggaagatct gcaaccaccc ctacatgttc cagcacatcg 3240
 aggagtcctt ttccgagcac ttggggttca ctggcgccat tgtccaaggg ctggacctgt 3300
 accgagcctc gggtaaattt gagcttcttg atagaattct tcccaaactc cgagcaacca 3360
 accacaaagt gctgctgttc tgccaaatga cctccctcat gaccatcatg gaagattact 3420
 ttgcgtatcg cggttttaaa tacctcaggc ttgatggaac cacgaaggcg gaggaccggg 3480
 gcatgctgct gaaaaccttc aacgagcccc gctctgagta cttcatcttc ctgctcagca 3540
 cccgggctgg ggggctcggc ctgaacctcc agtcggcaga cactgtgatc atttttgaca 3600
 gcgactggaa tcctcaccag gacctgcaag cgcaggaccg agcccaccgc atcgggcagc 3660
 agaacgaggt gcgtgtgctc cgcctctgca ccgtcaacag cgtggaggag aagatcctag 3720
 ctgcagccaa gtacaagctc aacgtggacc agaagtgat ccaggccggc atgttcgacc 3780
 agaagtcctc cagccatgag cggcgcgcct tcctgcaggc catcctggag cagcaggagc 3840
 aggatgagag cagacactgc agcacgggca gcggcagtgc cagcttcgcc cacactgccc 3900
 ctccgccagc gggcgtcaac cccgacttgg aggagccacc tctaaaggag gaagacgagg 3960
 tgccccgacga cgagaccgtc aaccagatga tcgccccgca cgaggaggag tttgatctgt 4020
 tcatgcgcat ggacctggac cgcaggcgcg aggaggcccc caaccccaag cggaagccgc 4080
 gcctcatgga ggaggacgag ctcccctcgt ggatcatcaa ggacgacgcg gaggtggagc 4140
 ggctgacctg tgaggaggag gaggagaaga tgttcggccg tggctcccgc caccgcaagg 4200
 aggtggacta cagcgactca ctgacggaga agcagtggct caagaaaatt acaggaagag 4260
 atatccatga cacagccagc agtgtggcac gtgggctaca attccagcgt ggccttcagt 4320
 tctgcacacg tgcgtcaaaag gccatcgagg agggcacgct ggaggagatc gaagaggagg 4380
 tccggcagaa gaaatcatca cggaagcgca agcgagacag cgacgccggc tcctccaccc 4440
 cgaccaccag caccgcagc cgcgacaagg acgacgagag caagaagcag aagaagcgcg 4500
 ggcggccgcc tgccgagaaa ctctccccta acccaccbaa cctcaccaag aagatgaaga 4560
 agattgtgga tgccgtgatc aagtacaagg acagcagcag tggacgtcag ctacgcagg 4620
 tcttcatcca gctgccctcg cgaaaggagc tgccccagta ctacgagctc atccgcaagc 4680
 ccgtggactt caagaagata aaggagcgca ttgcgaacca caagtaccgc agcctcaacg 4740
 acctagagaa ggacgtcatg ctctgtgcc agaacgcaca gaccttcaac ctggagggtc 4800
 ccctgatcta tgaagactcc atcgtcttgc agtcggtctt caccagcgtg cggcagaaaa 4860
 tcgagaagga ggatgacagt gaaggcgagg agagttagga ggaggaagag ggcgaggagg 4920

aaggctccga atccgaatct cggctccgtca aagtgaagat caagcttggc cggaaggaga 4980
 aggcacagga cgggctgaag ggcggccggc ggcggccgag ccgagggtcc cgagccaagc 5040
 cggtcgtgag tgacgatgac agtgaggagg aacaagagga ggaccgctca ggaagtggca 5100
 gcgaagaaga ctgagccccg acattccagt ctcgaccccc agccccctcgt tccagagctg 5160
 agatggcata ggccttagca gtaacgggta gcagcagatg tagtttcaga cttggagtaa 5220
 aactgtataa acaaaaagaat cttccatatt tatacagcag agaagctgta ggactgtttg 5280
 tgactggccc tgtcctggca tcagtagcat ctgtaacagc attaactgtc ttaaagagag 5340
 agagagagaa ttccgaattg gggaacacac gatactgtt tttcttttcc gttgctggca 5400
 gtactgttgc gccgcagttt ggagtcactg tagttaagtg tggatgcatg tgcgtcaccg 5460
 tccactcctc ctactgtatt ttattggaca ggtcagactc gccggggggc cggcgagggt 5520
 atgtcagtg cactggatgt caaacagtaa taaattaaac caacaacaaa acgcacagcc 5580
 aaaaaaaaa 5589

<210> 10

<211> 5779

<212> DNA

<213> 智人

<400> 10

ggagaggccg ccgcggtgct gagggggagg ggagccggcg agcgcgcgcg cagcgggggc 60
 gcgggtggcg cgcgtgtgtg tgaagggggg gcggtggccg aggcggggcg gcgcgcgcgc 120
 gaggcttccc ctcgtttggc ggcggcggcg gcttctttgt ttcgtgaaga gaagcgagac 180
 gccattctg cccccggccc cgcgcggagg ggcgggggag gcgccgggaa gtcgacggcg 240
 ccggcggctc ctgcgtctcg cccttttgcc caggctagag tgcaagtgtg cggtcatggt 300
 tcaactgcagc ctcaacctcc tggactcagc aggaggccac tgtctgcagc tcccgtgaag 360
 atgtccactc cagaccacc cctgggcgga actcctcggc caggtccttc cccgggcct 420
 ggcccttccc ctggagccat gctgggcctt agcccgggtc cctcgccggg ctccgcccac 480
 agcatgatgg ggcccagccc agggccgccc tcagcaggac accccatccc caccagggg 540
 cctggagggt accctcagga caacatgcac cagatgcaca agcccatgga gtccatgcat 600
 gagaagggca tgtcggacga cccgcgtac aaccagatga aaggaatggg gatgcggtca 660
 gggggccatg ctgggatggg gccccgccc agcccatgg accagcactc ccaaggttac 720
 ccctcgcccc tgggtggctc tgagcatgcc tctagtccag ttccagccag tggcccgtct 780
 tcggggcccc agatgtcttc cgggccagga ggtgccccgc tggatggtgc tgacccccag 840
 gccttggggc agcagaaccg gggcccaacc ccatttaacc agaaccagct gcaccagctc 900
 agagctcaga tcatggccta caagatgctg gccagggggc agcccctccc cgaccacctg 960
 cagatggcgg tgcaggcaa gcggccgatg cccgggatgc agcagcagat gccaacgcta 1020
 cctccaccct cgggtgtccg aacaggaccg ggccctggcc ctggccctgg ccccgccccg 1080
 ggtcccggcc cggcacctcc aaattacagc aggcctcatg gtatgggagg gcccaacatg 1140
 cctccccag gaccctcggg cgtgcccccc gggatgccag gccagcctcc tggagggcct 1200
 cccaagccct ggctgaagg acccatggcg aatgctgctg cccccacag caccctcag 1260
 aagctgattc cccgcagcc aacgggccgc cttcccccg cggccctgc cgtcccacc 1320

gccgcctcgc ccgtgatgcc accgcagacc cagtcccccg ggcagccggc ccagcccgcg 1380
 cccatgggtgc cactgcacca gaagcagagc cgcataccccc ccatccagaa gccgcggggc 1440
 ctcgaccctg tggagatcct gcaggagcgc gactacaggc tgcaggctcg catcgcacac 1500
 cgaattcagg aacttgaaaa ccttccccggg tccctggccg gggatttgcg aaccaaagcg 1560
 accattgagc tcaaggccct caggctgctg aacttccaga ggcagctgcg ccaggaggtg 1620
 gtggtgtgca tgcggaggga cacagcgtg gagacagccc tcaatgctaa ggcctacaag 1680
 cgcagcaagc gccagtccct gcgcgaggcc cgcatactg agaagctgga gaagcagcag 1740
 aagatcgagc aggagcgcaa gcgccggcag aagcaccagg aatacctcaa tagcattctc 1800
 cagcatgcca aggatttcaa ggaatatcac agatccgtca caggcaaaat ccagaagctg 1860
 accaaggcag tggccacgta ccatgccaac acggagcggg agcagaagaa agagaacgag 1920
 cggatcgaga aggagcgcag gcggaggtc atggctgaag atgaggagg gtaccgcaag 1980
 ctcacgcacc agaagaagga caagcgctg gcctacctt tgcagcagac agacgagtag 2040
 gtggctaacc tcacggagct ggtgcggcag cacaaggctg cccaggctgc caaggagaaa 2100
 aagaagaaaa agaaaaagaa gaaggcagaa aatgcagaag gacagacgcc tgccattggg 2160
 ccgcatggcg agcctctgga cgagaccagc cagatgagcg acctcccggt gaaggtgatc 2220
 cacgtggaga gtgggaagat cctcacaggc acagatgccc ccaaagccgg gcagctggag 2280
 gcctggctcg agatgaaccc ggggtatgaa gtagctccga ggtctgatag tgaagaaagt 2340
 ggctcagaag aagaggaaga ggaggaggag gaagagcagc cgcaggcagc acagcctccc 2400
 accctgcccc tggaggagaa gaagaagatt ccagatccag acagcgatga cgtctctgag 2460
 gtggacgcgc ggcacatcat tgagaatgcc aagcaagatg tcgatgatga atatggcgtg 2520
 tcccaggccc ttgcacgtgg cctgcagtcc tactatgccg tggcccatgc tgtcactgag 2580
 agagtggaca agcagtcagc gcttatggtc aatggtgtcc tcaaacagta ccagatcaaa 2640
 ggtttgaggt ggctgggtgc cctgtacaac aacaacctga acggcatcct ggccgacgag 2700
 atgggcctgg ggaagaccat ccagaccatc gcgctcatca cgtacctcat ggagcacaaa 2760
 cgcataaatg ggcccttcct catcatcgtg cctctctcaa cgctgtccaa ctgggcgtac 2820
 gagtttgaca agtgggcccc ctccgtgggtg aaggtgtctt acaagggatc cccagcagca 2880
 agacgggcct ttgtccccc gctccggagt gggaagttca acgtcttgct gacgacgtac 2940
 gagtacatca tcaaagacaa gcacatcctc gccaaagatc gttggaagta catgattgtg 3000
 gacgaaggct accgcatgaa gaaccaccac tgcaagctga cgcagggtgct caacacgcac 3060
 tatgtggcac cccgccgcct gctgctgacg ggcacaccgc tgcagaacaa gcttcccag 3120
 ctctgggcgc tgctcaactt cctgctgccc accatcttca agagctgcag caccttcgag 3180
 cagtggttta acgcaccctt tgccatgacc ggggaaaagg tggacctgaa tgaggaggaa 3240
 accattctca tcatccggcg tctccacaaa gtgctgcggc cttctttgct ccgacgactc 3300
 aagaaggaag tcgaggccca gttgcccgaa aaggtggagt acgtcatcaa gtgcgacatg 3360
 tctgcgctgc agcgagtgt ctaccgccac atgcaggcca agggcgtgct gctgactgat 3420
 ggctccgaga aggacaagaa gggcaaaggc ggcaccaaga ccctgatgaa caccatcatg 3480
 cagctgcgga agatctgcaa ccaccctac atgttccagc acatcgagga gtccttttcc 3540
 gagcacttgg ggttacttgg cggcattgtc caagggtggt acctgtaccg agcctcgggt 3600
 aaatttgagc ttcttgatag aattcttccc aaactccgag caaccaacca caaagtgtg 3660

ctgttctgcc aaatgacctc cctcatgacc atcatggaag attactttgc gtatcgcggc 3720
 tttaaatacc tcaggcttga tggaaccacg aaggcggagg accggggcat gctgctgaaa 3780
 accttcaacg agcccggctc tgagtacttc atcttctctg tcagcacccg ggctgggggg 3840
 ctcggcctga acctccagtc ggcagacact gtgatcattt ttgacagcga ctggaatcct 3900
 caccaggacc tgcaagcgca ggaccgagcc caccgcatcg ggcagcagaa cgaggtgcgt 3960
 gtgctccgcc tctgcaccgt caacagcgtg gaggagaaga tcctagctgc agccaagtac 4020
 aagctcaacg tggaccagaa ggtgatccag gccggcatgt tcgaccagaa gtcctccagc 4080
 catgagcggc gcgccttcct gcaggccatc ctggagcacg aggagcagga tgagagcaga 4140
 cactgcagca cgggcagcgg cagtgccagc ttcgcccaca ctgcccctcc gccagcgggc 4200
 gtcaaccccc acttgaggga gccacctcta aaggaggaag acgaggtgcc cgacgacgag 4260
 accgtcaacc agatgatcgc ccggcacgag gaggagtttg atctgttcat gcgcatggac 4320
 ctggaccgca ggcgcgagga ggcccgaac cccaagcgga agccgcgcct catggaggag 4380
 gacgagctcc cctcgtggat catcaaggac gacgcggagg tggagcggct gacctgtgag 4440
 gaggaggagg agaagatgtt cggccgtggc tcccgccacc gcaaggaggt ggactacagc 4500
 gactcactga cggagaagca gtggctcaag gccatcgagg agggcacgct ggaggagatc 4560
 gaagaggagg tccggcagaa gaaatcatca cggaagcgca agcgagacag cgacgccggc 4620
 tcctccaccc cgaccaccag caccgcgagc cgcgacaagg acgacgagag caagaagcag 4680
 aagaagcgcg ggcggccgcc tgccgagaaa ctctccccta acccacccaa cctcaccaag 4740
 aagatgaaga agattgtgga tgccgtgatc aagtacaagg acagcagcag tggacgtcag 4800
 ctcagcgagg tcttcatcca gctgccctcg cgaaaggagc tgcccagta ctacgagctc 4860
 atccgcaagc ccgtggactt caagaagata aaggagcgca ttcgcaacca caagtaccgc 4920
 agcctcaacg acctagagaa ggacgtcatg ctctgtgtcc agaacgcaca gaccttcaac 4980
 ctggagggct ccctgatcta tgaagactcc atcgtcttgc agtcggtctt caccagcgtg 5040
 cggcagaaaa tcgagaagga ggatgacagt gaaggcgagg agagtgagga ggaggaagag 5100
 ggcgaggagg aaggctccga atccgaatct cggctccgtca aagtgaagat caagcttggc 5160
 cggaaggaga aggcacagga ccggctgaag ggcgggccggc ggcgggccgag ccgagggtcc 5220
 cgagccaagc cggctcgtgag tgacgatgac agtgaggagg aacaagagga ggaccgctca 5280
 ggaagtggca gcgaagaaga ctgagccccg acattccagt ctcgaccccc agccccctgt 5340
 tccagagctg agatggcata ggccttagca gtaacgggta gcagcagatg tagtttcaga 5400
 cttggagtaa aactgtataa acaaaagaat cttccatatt tatacagcag agaagctgta 5460
 ggactgtttg tgactggccc tgtcctggca tcagtagcat ctgtaacagc attaaactgtc 5520
 ttaaagagag agagagagaa ttccgaattg gggaacacac gatacctgtt tttcttttcc 5580
 gttgctggca gtactgttgc gccgcagttt ggagtcactg tagttaagtg tggatgcatg 5640
 tgcgtcaccg tccactcctc ctactgtatt ttattggaca ggtcagactc gccgggggcc 5700
 cggcgagggt atgtcagtgt cactggatgt caaacagtaa taaattaaac caacaacaaa 5760
 acgcacagcc aaaaaaaaaa 5779

<210> 11

<211> 5329

<212> DNA

<213> 智人

<400> 11

```

atgtccactc cagacccacc cctgggcgga actcctcggc caggtccttc cccgggcccct 60
ggcccttccc ctggagccat gctggggcct agccccgggtc cctcgccggg ctccgcccac 120
agcatgatgg ggcccagccc agggccgccc tcagcaggac accccatccc caccagggg 180
cctggagggt accctcagga caacatgcac cagatgcaca agcccatgga gtccatgcat 240
gagaagggca tgtcggacga cccgcgtac aaccagatga aaggaatggg gatgcggtca 300
gggggccatg ctgggatggg gccccgccc agcccatgg accagcactc ccaaggttac 360
ccctcgcccc tgggtggctc tgagcatgcc tctagtccag ttccagccag tggcccgtct 420
tcggggcccc agatgtcttc cgggccagga ggtgccccgc tggatggtgc tgacccccag 480
gccttggggc agcagaaccg gggcccaacc ccatttaacc agaaccagct gcaccagctc 540
agagctcaga tcatggccta caagatgctg gccagggggc agccctccc cgaccacctg 600
cagatggcgg tgcaggcaa gcggccgatg cccgggatgc agcagcagat gccaacgcta 660
cctccaccct cgggtgtccgc aacaggaccc ggccctggcc ctggccctgg ccccgccccg 720
gggtccggcc cggcacctcc aaattacagc aggcctcatg gtatgggagg gccaacatg 780
cctccccag gaccctcggg cgtgcccccc gggatgccag gccagcctcc tggaggcct 840
cccaagccct ggctgaagg acccatggcg aatgctgctg cccccacgag caccctcag 900
aagctgattc ccccgagcc aacgggcccgc cttcccccg cggccctgc cgtcccacc 960
gccgcctgc ccgtgatgcc accgcagacc cagtccccg ggcagccggc ccagcccgcg 1020
cccatggtgc cactgcacca gaagcagagc cgcacaccc ccatccagaa gccgcggggc 1080
ctcgaccctg tggagatcct gcaggagcgc gactacaggc tgcaggctcg catcgcacac 1140
cgaattcagg aacttgaaaa cttcccggg tccctggccg gggatttgcg aaccaaagcg 1200
accattgagc tcaaggccct caggctgctg aacttcaga ggcagctgcg ccaggaggtg 1260
gtggtgtgca tgcggaggga cacagcgtg gagacagccc tcaatgctaa ggcctacaag 1320
cgcagcaagc gccagtccct gcgcgaggcc cgcactactg agaagctgga gaagcagcag 1380
aagatcgagc aggagcgcaa gcgccggcag aagcaccagg aatacctcaa tagcattctc 1440
cagcatgcca aggatttcaa ggaatatcac agatccgtca caggcaaat ccagaagctg 1500
accaaggcag tggccacgta ccatgccaac acggagcggg agcagaagaa agagaacgag 1560
cggatcgaga aggagcgcag gcggaggctc atggctgaag atgaggagg gtaccgcaag 1620
ctcatcgacc agaagaagga caagcgctg gcctacctt tgcagcagac agacgagtac 1680
gtggctaacc tcacggagct ggtgcggcag cacaaggctg cccaggtcgc caaggagaaa 1740
aagaagaaaa agaaaaagaa gaaggcagaa aatgcagaag gacagacgcc tgccattggg 1800
ccgatggcg agcctctgga cgagaccagc cagatgagcg acctcccggt gaaggtgatc 1860
cacgtggaga gtgggaagat cctcacaggc acagatgcc ccaaagccgg gcagctggag 1920
gcctggctcg agatgaaccc ggggtatgaa gtagctccga ggtctgatag tgaagaaagt 1980
ggctcagaag aagaggaaga ggaggaggag gaagagcagc cgcaggcagc acagcctccc 2040
accctgcccc tggaggagaa gaagaagatt ccagatccag acagcgatga cgtctctgag 2100
gtggacgcgc ggcacatcat tgagaatgcc aagcaagatg tcgatgatga atatggcgtg 2160
tcccaggccc ttgcacgtgg cctgcagtcc tactatgccg tggcccatgc tgtcactgag 2220

```

```

agagtggaca agcagtcagc gcttatggtc aatgggtgtcc tcaaacagta ccagatcaaa 2280
ggtttggagt ggctgggtgc cctgtacaac aacaacctga acggcatcct ggccgacgag 2340
atgggccttg ggaagaccat ccagaccatc gcgctcatca cgtacctcat ggagcacaaa 2400
cgcatcaatg ggcccttcct catcatcgtg cctctctcaa cgctgtccaa ctgggcgtac 2460
gagtttgaca agtggggccc ctccgtgggtg aaggtgtctt acaagggatc cccagcagca 2520
agacggggcct ttgtccccc gctccggagt gggaagtcca acgtcttgct gacgacgtac 2580
gagtacatca tcaaagacaa gcacatcctc gccaaagtcc gttggaagta catgattgtg 2640
gacgaaggtc accgcatgaa gaaccaccac tgcaagctga cgcaggtgct caacacgcac 2700
tatgtggcac cccgccgcct gctgctgacg ggcacaccgc tgcagaacaa gcttcccag 2760
ctctggggcg tgctcaactt cctgctgccc accatcttca agagctgcag caccttcgag 2820
cagtggttta acgcaccctt tgccatgacc ggggaaaagg tggacctgaa tgaggaggaa 2880
accattctca tcatccggcg tctccacaaa gtgctgcggc cttctttgct ccgacgactc 2940
aagaagggaag tcgaggccca gttgcccga aaggtggagt acgtcatcaa gtgcgacatg 3000
tctgcgtgc agcgagtgt ctaccgccac atgcaggcca agggcgtgct gctgactgat 3060
ggctccgaga aggacaagaa gggcaaaggc ggcaccaaga ccctgatgaa caccatcatg 3120
cagctgcgga agatctgcaa ccaccctac atgttccagc acatcgagga gtccttttcc 3180
gagcacttgg ggttcactgg cggcattgtc caagggtggt acctgtaccg agcctcgggt 3240
aaatttgagc ttcttgatag aattcttccc aaactccgag caaccaacca caaagtgtgt 3300
ctgttctgcc aaatgacct cctcatgacc atcatggaag attactttgc gtatcgcggc 3360
tttaaatacc tcaggcttga tggaaccacg aaggcggagg accggggcat gctgctgaaa 3420
accttcaacg agcccggtc tgagtacttc atcttctgc tcagcaccgc ggctgggggg 3480
ctcggcctga acctccagtc ggcagacact gtgatcattt ttgacagcga ctggaatcct 3540
caccaggacc tgcaagcgca ggaccgagcc caccgcatcg ggcagcagaa cgaggtgcgt 3600
gtgctccgcc tctgcaccgt caacagcgtg gaggagaaga tcctagctgc agccaagtac 3660
aagctcaacg tggaccagaa ggtgatccag gccggcatgt tcgaccagaa gtcctccagc 3720
catgagcggc gcgccttcct gcaggccatc ctggagcacg aggagcagga tgaggaggaa 3780
gacgaggtgc ccgacgacga gaccgtcaac cagatgatcg cccggcacga ggaggagttt 3840
gatctgttca tgcgcatgga cctggaccgc aggcgcgagg agggccgcaa cccaagcgg 3900
aagccgcgcc tcatggagga ggacgagctc ccctcgtgga tcatcaagga cgacgcggag 3960
gtggagcggc tgacctgtga ggaggaggag gagaagatgt tcggccgtgg ctcccgcac 4020
cgcaaggagg tggactacag cgactcactg acggagaagc agtggctcaa gacctgaag 4080
gccatcgagg agggcacgct ggaggagatc gaaggaggag tccggcagaa gaaatcatca 4140
cggaagcgca agcgagacag cgacgccggc tctccacc ccgaccaccag caccgcgac 4200
cgcgacaagg acgacgagag caagaagcag aagaagcgcg ggcggccgcc tgccgagaaa 4260
ctctccccta acccacccaa cctcaccag aagatgaaga agattgtgga tgccgtgatc 4320
aagtacaagg acagcagcag tggacgtcag ctacgcgagg tcttcatcca gctgcctcg 4380
cgaaaggagc tgcccagata ctacgagctc atccgcaagc ccgtggactt caagaagata 4440
aaggagcgca ttcgcaacca caagtaccgc agcctcaacg acctagagaa ggacgtcatg 4500
ctcctgtgcc agaacgcaca gaccttcaac ctggagggtc ccctgatcta tgaagactcc 4560

```

atcgtcttgc agtcggtctt caccagcgtg cggcagaaaa tcgagaagga ggatgacagt 4620
 gaaggcgagg agagtgagga ggaggaagag ggcgaggagg aaggctccga atccgaatct 4680
 cggctccgtca aagtgaagat caagcttggc cggaaggaga aggcacagga ccggtgaag 4740
 ggcggccggc ggcggccgag ccgaggggtcc cgagccaagc cggtcgtgag tgacgatgac 4800
 agtgaggagg aacaagagga ggaccgctca ggaagtggca gcgaagaaga ctgagccccg 4860
 acattccagt ctcgaccccc agccccctcgt tccagagctg agatggcata ggccttagca 4920
 gtaacgggta gcagcagatg tagtttcaga cttggagtaa aactgtataa aaaaaagaat 4980
 cttccatatt tatacagcag agaagctgta ggactgtttg tgactggccc tgtcctggca 5040
 tcagtagcat ctgtaacagc attaactgtc ttaaagagag agagagagaa ttccgaattg 5100
 gggaacacac gatacctgtt tttcttttcc gttgctggca gtactgttgc gccgcagttt 5160
 ggagtcaactg tagttaagtg tggatgcatg tgcgtcaccg tccactctc ctactgtatt 5220
 ttattggaca ggtcagactc gccggggggc cggcgagggt atgtcagtgt cactggatgt 5280
 caaacagtaa taaattaaac caacaacaaa acgcacagcc aaaaaaaaa 5329

<210> 12

<211> 5326

<212> DNA

<213> 智人

<400> 12

atgtccactc cagacccacc cctgggcgga actcctcggc caggtccttc cccgggacct 60
 ggcccttccc ctggagccat gctgggccct agcccgggtc cctcgccggg ctccgcccac 120
 agcatgatgg ggcccagccc agggccgccc tcagcaggac accccatccc caccagggg 180
 cctggagggt accctcagga caacatgcac cagatgcaca agcccatgga gtccatgcat 240
 gagaagggca tgtcggacga cccgcgctac aaccagatga aaggaatggg gatgcggtca 300
 gggggccatg ctgggatggg gccccgccc agcccatgg accagcactc ccaaggttac 360
 cctcgcgcc tgggtggctc tgagcatgcc tctagtccag ttccagccag tggccctct 420
 tcggggcccc agatgtcttc cgggccagga ggtgccccgc tggatggtgc tgacccccag 480
 gccttggggc agcagaaccg gggcccaacc ccatttaacc agaaccagct gcaccagctc 540
 agagctcaga tcatggccta caagatgctg gccagggggc agcccctccc cgaccacctg 600
 cagatggcgg tgcagggcaa gcggccgatg cccgggatgc agcagcagat gccaacgcta 660
 cctccaccct cgggtgtccg aacaggaccg ggccctggcc ctggccctgg ccccgccccg 720
 ggtcccggcc cggcacctcc aaattacagc aggcctcatg gtatgggagg gcccaacatg 780
 cctccccag gaccctcggg cgtgcccccc gggatgccag gccagcctcc tggagggcct 840
 cccaagccct ggccctgaagg acccatggcg aatgctgctg ccccccagag caccctcag 900
 aagctgattc cccgcagcc aacgggccc ccttcccccg cgcacctgc cgtccaccc 960
 gccgcctgc cgtgatgcc accgcagacc cagtcccccg ggcagccggc ccagcccgcg 1020
 cccatggtgc cactgcacca gaagcagagc cgcacaccc ccatccagaa gcccgggggc 1080
 ctcgaccctg tggagatcct gcaggagcgc gactacaggc tgcaggctcg catcgcacac 1140
 cgaattcagg aacttgaaaa ccttcccggg tccttgccg gggatttgcg aaccaaagcg 1200
 accattgagc tcaaggccct caggctgctg aacttccaga ggcagctgcg ccaggaggtg 1260

gtggtgtgca tgcggaggga cacagcgctg gagacagccc tcaatgctaa ggcctacaag 1320
 cgcagcaagc gccagtccct gcgcgaggcc cgcactactg agaagctgga gaagcagcag 1380
 aagatcgagc aggagcgcaa gcgccggcag aagcaccagg aatacctcaa tagcattctc 1440
 cagcatgcca aggatttcaa ggaatatcac agatccgtca caggcaaaat ccagaagctg 1500
 accaaggcag tggccacgta ccatgccaac acggagcggg agcagaagaa agagaacgag 1560
 cggatcgaga aggagcgcat gcggaggctc atggctgaag atgaggaggg gtaccgcaag 1620
 ctcatcgacc agaagaagga caagcgctg gcctacctct tgcagcagac agacgagtac 1680
 gtggctaacc tcacggagct ggtgcggcag cacaaggctg cccaggctgc caaggagaaa 1740
 aagaagaaaa agaaaaagaa gaaggcagaa aatgcagaag gacagacgcc tgccattggg 1800
 ccggatggcg agcctctgga cgagaccagc cagatgagcg acctcccggg gaaggtgatc 1860
 cacgtggaga gtgggaagat cctcacaggc acagatgccc ccaaagccgg gcagctggag 1920
 gcctggctcg agatgaaccc ggggtatgaa gtagctccga ggtctgatag tgaagaaagt 1980
 ggctcagaag aagaggaaga ggaggaggag gaagagcagc cgcaggcagc acagcctccc 2040
 accctgcccc tggaggagaa gaagaagatt ccagatccag acagcgatga cgtctctgag 2100
 gtggacgcgc ggcacatcat tgagaatgcc aagcaagatg tcgatgatga atatggcgtg 2160
 tcccaggccc ttgcacgtgg cctgcagtcc tactatgccg tggcccatgc tgtcactgag 2220
 agagtggaca agcagtcagc gcttatggtc aatggtgtcc tcaaacagta ccagatcaaa 2280
 ggtttggagt ggctgggtgc cctgtacaac aacaacctga acggcatcct ggccgacgag 2340
 atgggcctgg ggaagaccat ccagaccatc gcgctcatca cgtacctcat ggagcacaaa 2400
 cgcataatg ggcccttcct catcatcgtg cctctctcaa cgctgtccaa ctgggcgtac 2460
 gagtttgaca agtggggccc ctccgtgggtg aaggtgtctt acaagggatc cccagcagca 2520
 agacgggcct ttgtcccca gctccggagt gggaagtcca acgtcttgct gacgacgtac 2580
 gagtacatca tcaaagacaa gcacatcctc gccaaagatc gttggaagta catgattgtg 2640
 gacgaaggtc accgcatgaa gaaccaccac tgcaagctga cgcagggtgct caacacgcac 2700
 tatgtggcac cccgccgcct gctgctgacg ggcacaccgc tgcagaacaa gcttcccag 2760
 ctctgggcgc tgctcaactt cctgctgccc accatcttca agagctgcag caccttcgag 2820
 cagtggttta acgcaccctt tgccatgacc ggggaaaagg tggacctgaa tgaggaggaa 2880
 accattctca tcatccggcg tctccacaaa gtgctgcggc ctttcttgct cgcagactc 2940
 aagaaggaag tcgaggccca gttgcccga aaggtggagt acgtcatcaa gtgcgacatg 3000
 tctgcgtgc agcgagtgt ctaccgccac atgcaggcca agggcgtgct gctgactgat 3060
 ggctccgaga aggacaagaa gggcaaaggc ggcaccaaga ccctgatgaa caccatcatg 3120
 cagctgcgga agatctgcaa ccaccctac atgttccagc acatcgagga gtccctttcc 3180
 gagcacttgg ggttcaactg cggcattgtc caagggtgg acctgtaccg agcctcgggt 3240
 aaatttgagc ttcttgatag aattcttccc aaactccgag caaccaacca caaagtgtg 3300
 ctgttctgcc aaatgacctc cctcatgacc atcatggaag attactttgc gtatcgcggc 3360
 tttaaatacc tcaggcttga tggaaaccag aaggcggagg accggggcat gctgctgaaa 3420
 acctcaacg agcccggctc tgagtacttc atcttctgc tcagcaccg ggctgggggg 3480
 ctggcctga acctccagtc ggcagacact gtgatcattt ttgacagcga ctggaatcct 3540
 caccaggacc tgcaagcgca ggaccgagcc caccgcatcg ggcagcagaa cgaggtgcgt 3600

gtgctccgcc tctgcaccgt caacagcgtg gaggagaaga tcctagctgc agccaagtac 3660
 aagctcaacg tggaccagaa ggtgatccag gccggcatgt tcgaccagaa gtcctccagc 3720
 catgagcggc ggccttcct gcaggccatc ctggagcacg aggagcagga tgaggaggaa 3780
 gacgaggtgc ccgacgacga gaccgtcaac cagatgatcg cccggcacga ggaggagttt 3840
 gatctgttca tgcgcatgga cctggaccgc aggcgcgagg aggcccgcaa cccaagcgg 3900
 aagccgcgcc tcatggagga ggacgagctc ccctcgtgga tcatcaagga cgacgcggag 3960
 gtggagcggc tgacctgtga ggaggaggag gagaagatgt tcggccgtgg ctcccggcac 4020
 cgcaaggagg tggactacag cgactactg acggagaagc agtggctcaa gacctgaag 4080
 gccatcgagg agggcacgct ggaggagatc gaagaggagg tccggcagaa gaaatcatca 4140
 cggaagcgca agcgagacag cgacgccggc tcctccacc cgaccaccag caccgcgagc 4200
 cgcgacaagg acgacgagag caagaagcag aagaagcgcg ggcgccgcc tgccgagaaa 4260
 ctctccccta acccacccaa cctcaccaag aagatgaaga agattgtgga tgccgtgatc 4320
 aagtacaagg acagcagtgg acgtcagctc agcgaggtct tcatccagct gccctcgca 4380
 aaggagctgc ccgagtacta cgagctcatc cgcaagccc tggacttcaa gaagataaag 4440
 gagcgcatte gcaaccacaa gtaccgcagc ctcaacgacc tagagaagga cgtcatgctc 4500
 ctgtgccaga acgcacagac cttcaacctg gagggctccc tgatctatga agactccatc 4560
 gtcttgccagt cggctcttcac cagcgtgcgg cagaaaatcg agaaggagga tgacagtga 4620
 ggcgaggaga gtgaggagga ggaagagggc gaggaggaag gctccgaatc cgaatctcgg 4680
 tccgtcaaag tgaagatcaa gcttgccgg aaggagaagg cacaggaccg gctgaagggc 4740
 ggccggcggc ggccgagccg aggggtcccga gccaagccgg tcgtgagtga cgatgacagt 4800
 gaggaggaac aaggaggagga ccgctcagga agtggcagcg aagaagactg agccccgaca 4860
 ttccagtctc gaccccgagc ccctcgttcc agagctgaga tggcataggc cttagcagta 4920
 acgggtagca gcagatgtag tttcagactt ggagtaaaac tgtataaaca aaagaatctt 4980
 ccatatttat acagcagaga agctgtagga ctgtttgtga ctggccctgt cctggcatca 5040
 gtagcatctg taacagcatt aactgtctta aagagagaga gagagaattc cgaattgggg 5100
 aacacacgat acctgttttt cttttccgtt gctggcagta ctgttgccg gcagtttgga 5160
 gtcactgtag ttaagtgtgg atgcatgtgc gtcaccgtcc actcctccta ctgtatttta 5220
 ttggacaggt cagactcgcc gggggcccgg cgagggtatg tcagtgtcac tggatgtcaa 5280
 acagtaataa attaaaccaa caacaaaacg cacagccaaa aaaaaa 5326

<210> 13

<211> 5320

<212> DNA

<213> 智人

<400> 13

atgtccactc cagaccacc cctgggcgga actcctcggc caggtccttc cccgggcect 60
 ggcccttccc ctggagccat gctgggcect agcccggtc cctcgccggg ctccgcccac 120
 agcatgatgg ggcccagccc agggccgccc tcagcaggac accccatccc caccagggg 180
 cctggagggg accctcagga caacatgcac cagatgcaca agcccatgga gtccatgcat 240
 gagaagggca tgtcggacga cccgcgtac aaccagatga aaggaatggg gatgcggtca 300

```

gggggccatg ctgggatggg gccccgccc agccccatgg accagcactc ccaaggttac 360
ccctcgcccc tgggtggctc tgagcatgcc tctagtccag ttccagccag tggcccgtct 420
tcggggcccc agatgtcttc cgggccagga ggtgccccgc tggatggtgc tgacccccag 480
gccttggggc agcagaaccg gggcccaacc ccatttaacc agaaccagct gcaccagctc 540
agagctcaga tcatggccta caagatgctg gccagggggc agcccctccc cgaccacctg 600
cagatggcgg tgcagggcaa gcggccgatg cccgggatgc agcagcagat gccaacgcta 660
cctccaccct cgggtgccgc aacaggaccc ggccctggcc ctggccctgg ccccgccccg 720
ggtcccggcc cggcacctcc aaattacagc aggcctcatg gtatgggagg gcccaacatg 780
cctccccag gaccctcggg cgtgcccccc gggatgccag gccagcctcc tggagggcct 840
cccaagccct ggccctgaagg acccatggcg aatgctgctg ccccccagag caccctcag 900
aagctgattc ccccgagcc aacggggcgc cttcccccg cgcacctgc cgtccaccc 960
gccgcctcgc ccgtgatgcc accgcagacc cagtccccg ggcagccggc ccagcccgcg 1020
cccatggtgc cactgcacca gaagcagagc cgcacaccc ccatccagaa gcccgggggc 1080
ctcgaccctg tggagatcct gcaggagcgc gactacaggc tgcaggctcg catcgcacac 1140
cgaattcagg aacttgaaaa cttccccggg tccttgccg gggatttgcg aaccaaagcg 1200
accattgagc tcaaggccct caggctgctg aacttccaga ggcagctcgc ccaggaggtg 1260
gtgggtgtgca tgcggaggga cacagcgtg gagacagccc tcaatgctaa ggcctacaag 1320
cgcagcaagc gccagtccct gcgcgaggcc cgcactactg agaagctgga gaagcagcag 1380
aagatcgagc aggagcgcaa gcgccggcag aagcaccagg aatacctcaa tagcattctc 1440
cagcatgcca aggatttcaa ggaatatcac agatccgtca caggcaaaat ccagaagctg 1500
accaaggcag tggccacgta ccatgccaac acggagcggg agcagaagaa agagaacag 1560
cggatcgaga aggagcgcag gcggaggctc atggctgaag atgaggagg gtaccgcaag 1620
ctcatcgacc agaagaagga caagcgcctg gcctacctt tgcagcagac agacaggtac 1680
gtggctaacc tcacggagct ggtgcggcag cacaaggctg ccaggtcgc caaggagaaa 1740
aagaagaaaa agaaaaagaa gaaggcagaa aatgcagaag gacagacgcc tgccattggg 1800
ccgatggcg agcctctgga cgagaccagc cagatgagcg acctcccggt gaaggtgatc 1860
cacgtggaga gtgggaagat cctcacaggc acagatgccc ccaaagccgg gcagctggag 1920
gcctggctcg agatgaaccc ggggtatgaa gtagctccga ggtctgatag tgaagaaagt 1980
ggctcagaag aagaggaaga ggaggaggag gaagagcagc cgcaggcagc acagcctccc 2040
accctgcccc tggaggagaa gaagaagatt ccagatccag acagcgatga cgtctctgag 2100
gtggacgcgc ggcacatcat tgagaatgcc aagcaagatg tcgatgatga atatggcgtg 2160
tcccaggccc ttgcacgtgg cctgcagtc tactatgccg tggcccatgc tgtcactgag 2220
agagtggaca agcagtcagc gcttatggtc aatggtgtcc tcaaacagta ccagatcaaa 2280
ggtttggagt ggctgggtgc cctgtacaac aacaacctga acggcatcct ggccgacgag 2340
atgggcctgg ggaagaccat ccagaccatc gcgctcatca cgtacctcat ggagcacaaa 2400
cgcacatgat ggcccttct catcatcgtg cctctctcaa cgctgtccaa ctgggcgtac 2460
gagtttgaca agtgggcccc ctccgtggtg aaggtgtctt acaagggatc cccagcagca 2520
agacgggcct ttgtcccca gctccggagt gggaagtcca acgtcttgct gacgacgtac 2580
gagtacatca tcaaagacaa gcacatcctc gccaatgcc gttggaagta catgattgtg 2640

```

gacgaaggct accgcatgaa gaaccaccac tgcaagctga cgcaggtgct caacacgcac 2700
 tatgtggcac cccgccgcct gctgctgacg ggcacaccgc tgcagaacaa gtttcccag 2760
 ctctgggctg tgctcaactt cctgctgccc accatcttca agagctgcag caccttcgag 2820
 cagtgggttta acgcaccctt tgccatgacc ggggaaaagg tggacctgaa tgaggaggaa 2880
 accatttctca tcatccggcg tctccacaaa gtgctgcggc ctttcttgct cgcagactc 2940
 aagaaggaag tcgaggccca gttgcccga aaggtggagt acgtcatcaa gtgcgacatg 3000
 tctgcgctgc agcgagtgc ctaccgccac atgcaggcca agggcgtgct gctgactgat 3060
 ggctccgaga aggacaagaa gggcaaaggc ggcaccaaga ccctgatgaa caccatcatg 3120
 cagctgcgga agatctgcaa ccaccctac atgttccagc acatcgagga gtccttttcc 3180
 gagcacttgg ggttcaactg cggcattgtc caagggtgag acctgtaccg agcctcgggt 3240
 aaatttgagc ttcttgatag aattcttccc aaactccgag caaccaacca caaagtgtctg 3300
 ctgttctgcc aaatgacctc cctcatgacc atcatggaag attactttgc gtatcgcggc 3360
 tttaaatacc tcaggcttga tggaaaccag aaggcggagg accggggcat gctgctgaaa 3420
 acctcaacg agcccggtc tgagtacttc atcttctgc tcagcaccg ggctgggggg 3480
 ctggcctga acctccagtc ggcagacact gtgatcattt ttgacagcga ctggaatcct 3540
 caccaggacc tgcaagcgca ggaccgagcc caccgcatcg ggcagcagaa cgaggtgcgt 3600
 gtgctccgcc tctgcaccgt caacagcgtg gaggagaaga tcctagctgc agccaagtac 3660
 aagctcaacg tggaccagaa ggtgatccag gccggcatgt tcgaccagaa gtcctccagc 3720
 catgagcggc gcgccttctc gcaggccatc ctggagcacg aggagcagga tgaggaggaa 3780
 gacgaggtgc ccgacgacga gaccgtcaac cagatgatcg cccggcacga ggaggagttt 3840
 gatctgttca tgcgcatgga cctggaccgc aggcgcgagg aggcccgcaa cccaagcgg 3900
 aagccgcgcc tcatggagga ggacgagctc ccctcgtgga tcatcaagga cgacgcggag 3960
 gtggagcggc tgacctgtga ggaggaggag gagaagatgt tcggccgtgg ctcccgccac 4020
 cgcaaggagg tggactacag cgactcactg acggagaagc agtgggtcaa ggccatcgag 4080
 gagggcacgc tggaggagat cgaagaggag gtccggcaga agaaatcatc acggaagcgc 4140
 aagcgagaca gcgacgccgg ctctccacc ccgaccacca gcaccgcag ccgcgacaag 4200
 gacgacgaga gcaagaagca gaagaagcgc gggcgggccg ctgccgagaa actctcccct 4260
 aaccacacca acctaccaa gaagatgaag aagattgtgg atgccgtgat caagtacaag 4320
 gacagcagca gtggacgtca gctcagcgag gtcttcatcc agctgccctc gcgaaaggag 4380
 ctgcccagat actacgagct catccgcaag cccgtggact tcaagaagat aaaggagcgc 4440
 attcgcaacc acaagtaccg cagcctcaac gacctagaga aggacgtcat gtcctgtgc 4500
 cagaacgcac agaccttcaa cctggagggc tcctgatct atgaagactc catcgtcttg 4560
 cagtcggtct tcaccagcgt gcggcagaaa atcgagaagg aggatgacag tgaaggcgag 4620
 gagagtgagg aggaggaaga gggcgaggag gaaggctccg aatccgaatc tcggtccgtc 4680
 aaagtgaaga tcaagcttgg ccggaaggag aaggcacagg accggctgaa gggcgggcgg 4740
 cggcgggcga gccgagggtc ccgagccaag ccggtcgtga gtgacgatga cagtgaggag 4800
 gaacaagagg aggaccgtc aggaagtggc agcgaagaag actgagcccc gacattccag 4860
 tctcgacccc gagcccctcg ttccagagct gagatggcat aggccttagc agtaacgggt 4920
 agcagcagat gtagtttcag acttgagta aaactgtata aacaaaagaa tcttccatat 4980

ttatacagca gagaagctgt aggactgttt gtgactggcc ctgtcctggc atcagtagca 5040
 tctgtaacag cattaactgt cttaaagaga gagagagaga attccgaatt ggggaacaca 5100
 cgatacctgt ttttcttttc cgttgctggc agtactgttg cgccgcagtt tggagtcact 5160
 gtagttaagt gtggatgcat gtgcgtcacc gtccactcct cctactgtat tttattggac 5220
 aggtcagact cgccgggggc ccggcgaggg tatgtcagtg tctactggatg tcaaacagta 5280
 ataaattaaa ccaacaacaa aacgcacagc caaaaaaaaa 5320

<210> 14

<211> 5317

<212> DNA

<213> 智人

<400> 14

atgtccactc cagacccacc cctgggcgga actcctcggc caggtccttc cccgggcccct 60
 ggcccttccc ctggagccat gctgggcccct agccccgggtc cctcgccggg ctccgcccac 120
 agcatgatgg ggcccagccc agggccgccc tcagcaggac accccatccc caccagggg 180
 cctggagggt accctcagga caacatgcac cagatgcaca agcccatgga gtccatgcat 240
 gagaagggca tgtcggacga cccgcgtac aaccagatga aaggaatggg gatgcggtca 300
 gggggccatg ctgggatggg gccccgccc agcccatgg accagcactc ccaaggttac 360
 ccctcgcccc tgggtggctc tgagcatgcc tctagtccag ttccagccag tggcccgtct 420
 tcggggcccc agatgtcttc cgggccagga ggtgccccgc tggatggtgc tgacccccag 480
 gccttggggc agcagaaccg gggcccaacc ccatttaacc agaaccagct gcaccagctc 540
 agagctcaga tcatggccta caagatgctg gccagggggc agcccctccc cgaccacctg 600
 cagatggcgg tgcagggcaa gcggccgatg cccgggatgc agcagcagat gccaacgcta 660
 cctccaccct cgggtgtccg aacaggaccg ggccctggcc ctggccctgg ccccgccccg 720
 ggtcccggcc cggcacctcc aaattacagc aggcctcatg gtatgggagg gcccaacatg 780
 cctccccag gaccctcggg cgtgcccccc gggatgccag gccagcctcc tggagggcct 840
 cccaagccct ggccgaagg acccatggcg aatgctgctg ccccccagag caccctcag 900
 aagctgattc cccgcagcc aacgggccc cttcccccg cggccctgc cgtcccacc 960
 gccgcctgc ccgtgatgcc accgcagacc cagtcccccg ggcagccggc ccagcccgcg 1020
 cccatggtgc cactgcacca gaagcagagc cgcacaccc ccatccagaa gccgcggggc 1080
 ctcgaccctg tggagatcct gcaggagcgc gactacaggc tgcaggctcg catcgcacac 1140
 cgaattcagg aacttgaaaa cttcccggg tccttgggcg gggatttgcg aaccaaagcg 1200
 accattgagc tcaaggccct caggtgctg aacttcaga ggcagctgcg ccaggaggtg 1260
 gtggtgtgca tgcggaggga cacagcgtg gagacagccc tcaatgctaa ggcctacaag 1320
 cgcagcaagc gccagtccct gcgcgaggcc cgcactactg agaagctgga gaagcagcag 1380
 aagatcgagc aggagcgcaa gcgccggcag aagcaccagg aatacctcaa tagcattctc 1440
 cagcatgcca aggatttcaa ggaatatcac agatccgtca caggcaaat ccagaagctg 1500
 accaaggcag tggccacgta ccatgccaac acggagcggg agcagaagaa agagaacgag 1560
 cggatcgaga aggagcgcac gcggaggtc atggctgaag atgaggagg gtaccgcaag 1620
 ctcatcgacc agaagaagga caagcgctg gcctacctt tgcagcagac agacgagtac 1680

gtggctaacc tcacggagct ggtgcggcag cacaaggctg cccaggtcgc caaggagaaa 1740
 aagaagaaaa agaaaaagaa gaaggcagaa aatgcagaag gacagacgcc tgccattggg 1800
 ccggatggcg agcctctgga cgagaccagc cagatgagcg acctcccggt gaaggtgatc 1860
 cacgtggaga gtgggaagat cctcacaggc acagatgccc ccaaagccgg gcagctggag 1920
 gcctggctcg agatgaaccc ggggtatgaa gtagctccga ggtctgatag tgaagaaagt 1980
 ggctcagaag aagaggaaga ggaggaggag gaagagcagc cgcaggcagc acagcctccc 2040
 accctgcccc tggaggagaa gaagaagatt ccagatccag acagcgatga cgtctctgag 2100
 gtggacgcgc ggcacatcat tgagaatgcc aagcaagatg tcgatgatga atatggcgtg 2160
 tcccaggccc ttgcacgtgg cctgcagtcc tactatgccg tggcccatgc tgtcactgag 2220
 agagtggaca agcagtcagc gcttatggtc aatgggtgtc tcaaacagta ccagatcaaa 2280
 ggtttggagt ggctgggtgc cctgtacaac aacaacctga acggcatcct ggccgacgag 2340
 atgggccttg ggaagaccat ccagaccatc gcgctcatca cgtacctcat ggagcacaaa 2400
 cgcataatg ggcccttcct catcatcgtg cctctctcaa cgtgttccaa ctgggcgtac 2460
 gagtttgaca agtgggcccc ctccgtggtg aaggtgtctt acaagggatc cccagcagca 2520
 agacgggcct ttgtccccc gctccggagt gggaagtcca acgtcttgct gacgacgtac 2580
 gagtacatca tcaaagacaa gcacatcctc gccaaagatc gttggaagta catgattgtg 2640
 gacgaaggtc accgcatgaa gaaccaccac tgcaagctga cgcagggtgt caacacgcac 2700
 tatgtggcac cccgccgcct gctgctgacg ggcacaccgc tgcagaacaa gcttcccag 2760
 ctctgggcgc tgtctaacct cctgctgccc accatcttca agagctgcag caccttcgag 2820
 cagtggttta acgcaccctt tgccatgacc ggggaaaagg tggacctgaa tgaggaggaa 2880
 accattctca tcatccggcg tctccacaaa gtgctgcggc cttctttgct ccgacgactc 2940
 aagaaggaag tcgaggccca gttgcccga aaggtggagt acgtcatcaa gtgcgacatg 3000
 tctgcgtgc agcgagtgt ctaccgccac atgcaggcca agggcgtgtg gctgactgat 3060
 ggctccgaga aggacaagaa gggcaaaggc ggcaccaaga ccctgatgaa caccatcatg 3120
 cagctgcgga agatctgcaa ccaccctac atgttccagc acatcgagga gtccttttcc 3180
 gagcacttgg ggttacttgg cggcattgtc caagggtgtg acctgtaccg agcctcgggt 3240
 aaatttgagc ttcttgatag aattcttccc aaactccgag caaccaacca caaagtgtg 3300
 ctgtttctgcc aaatgacctc cctcatgacc atcatggaag attactttgc gtatcgcggc 3360
 tttaaataacc tcaggcttga tggaaccacg aaggcggagg accggggcat gctgctgaaa 3420
 accttcaacg agcccggctc tgagtacttc atcttctgc tcagcaccgc ggctgggggg 3480
 ctcggcctga acctccagtc ggcagacact gtgatcattt ttgacagcga ctggaatcct 3540
 caccaggacc tgcaagcgca ggaccgagcc caccgcatcg ggcagcagaa cgagggtcgt 3600
 gtgctccgcc tctgcaccgt caacagcgtg gaggagaaga tcctagctgc agccaagtac 3660
 aagctcaacg tggaccagaa ggtgatccag gccggcatgt tcgaccagaa gtctctccagc 3720
 catgagcggc gcgccttcct gcaggccatc ctggagcacg aggagcagga tgaggaggaa 3780
 gacgaggtgc ccgacgacga gaccgtcaac cagatgatcg cccggcacga ggaggagttt 3840
 gatctgttca tgcgcatgga cctggaccgc aggcgcgagg agggccgcaa cccaagcgg 3900
 aagccgcgcc tcatggagga ggacgagctc ccctcgtgga tcatcaagga cgacgcggag 3960
 gtggagcggc tgacctgtga ggaggaggag gagaagatgt tcggccgtgg ctcccgcac 4020

cgcaaggagg tggactacag cgactcactg acggagaagc agtgggtcaa ggccatcgag 4080
 gagggcacgc tggaggagat cgaagaggag gtccggcaga agaaatcatc acggaagcgc 4140
 aagcgagaca gcgacgccgg ctctctccacc ccgaccacca gcaccccgag ccgcgacaag 4200
 gacgacgaga gcaagaagca gaagaagcgc gggcgggccgc ctgccgagaa actctcccct 4260
 aacccaccca acctcaccaa gaagatgaag aagattgtgg atgccgtgat caagtacaag 4320
 gacagcagtg gacgtcagct cagcgaggtc ttcatccagc tgccctcgcg aaaggagctg 4380
 cccgagtact acgagctcat ccgcaagccc gtggacttca agaagataaa ggagcgcatt 4440
 cgcaaccaca agtaccgcag cctcaacgac ctagagaagg acgtcatgct cctgtgccag 4500
 aacgcacaga ccttcaacct ggagggctcc ctgatctatg aagactccat cgtcttgcag 4560
 tcggtcttca ccagcgtgcg gcagaaaatc gagaaggagg atgacagtga aggcgaggag 4620
 agtgaggagg aggaagaggg cgaggaggaa ggctccgaat ccgaatctcg gtccgtcaaa 4680
 gtgaagatca agcttgccg gaaggagaag gcacaggacc ggctgaaggc cgcccgccgg 4740
 cggccgagcc gaggggtccc agccaagccg gtcgtgagtg acgatgacag tgaggaggaa 4800
 caagaggagg accgctcagg aagtggcagc gaagaagact gagccccgac attccagtct 4860
 cgacccccgag cccctcgttc cagagctgag atggcatagg ccttagcagt aacgggtagc 4920
 agcagatgta gtttcagact tggagtaaaa ctgtataaac aaaagaatct tccatattta 4980
 tacagcagag aagctgtagg actgtttgtg actggccctg tcctggcatc agtagcatct 5040
 gtaacagcat taactgtctt aaagagagag agagagaatt ccgaattggg gaacacacga 5100
 tacctgtttt tcttttccgt tgctggcagt actgttgcgc cgcagtttgg agtcaactgta 5160
 gttaagtgtg gatgcatgtg cgtcaccgtc cactcctcct actgtatttt attggacagg 5220
 tcagactcgc cggggggccc gcgagggtat gtcagtgtca ctggatgtca aacagtaata 5280
 aattaaacca acaacaaaac gcacagccaa aaaaaaa 5317

<210> 15

<211> 1679

<212> PRT

<213> 智人

<400> 15

Met	Ser	Thr	Pro	Asp	Pro	Pro	Leu	Gly	Gly	Thr	Pro	Arg	Pro	Gly	Pro
1				5				10					15		
Ser	Pro	Gly	Pro	Gly	Pro	Ser	Pro	Gly	Ala	Met	Leu	Gly	Pro	Ser	Pro
			20					25					30		
Gly	Pro	Ser	Pro	Gly	Ser	Ala	His	Ser	Met	Met	Gly	Pro	Ser	Pro	Gly
			35				40						45		
Pro	Pro	Ser	Ala	Gly	His	Pro	Ile	Pro	Thr	Gln	Gly	Pro	Gly	Gly	Tyr
			50				55				60				
Pro	Gln	Asp	Asn	Met	His	Gln	Met	His	Lys	Pro	Met	Glu	Ser	Met	His
65				70					75					80	
Glu	Lys	Gly	Met	Ser	Asp	Asp	Pro	Arg	Tyr	Asn	Gln	Met	Lys	Gly	Met
				85					90					95	

Gly Met Arg Ser Gly Gly His Ala Gly Met Gly Pro Pro Pro Ser Pro		
100	105	110
Met Asp Gln His Ser Gln Gly Tyr Pro Ser Pro Leu Gly Gly Ser Glu		
115	120	125
His Ala Ser Ser Pro Val Pro Ala Ser Gly Pro Ser Ser Gly Pro Gln		
130	135	140
Met Ser Ser Gly Pro Gly Gly Ala Pro Leu Asp Gly Ala Asp Pro Gln		
145	150	155
Ala Leu Gly Gln Gln Asn Arg Gly Pro Thr Pro Phe Asn Gln Asn Gln		
165	170	175
Leu His Gln Leu Arg Ala Gln Ile Met Ala Tyr Lys Met Leu Ala Arg		
180	185	190
Gly Gln Pro Leu Pro Asp His Leu Gln Met Ala Val Gln Gly Lys Arg		
195	200	205
Pro Met Pro Gly Met Gln Gln Gln Met Pro Thr Leu Pro Pro Pro Ser		
210	215	220
Val Ser Ala Thr Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro		
225	230	235
Gly Pro Gly Pro Ala Pro Pro Asn Tyr Ser Arg Pro His Gly Met Gly		
245	250	255
Gly Pro Asn Met Pro Pro Pro Gly Pro Ser Gly Val Pro Pro Gly Met		
260	265	270
Pro Gly Gln Pro Pro Gly Gly Pro Pro Lys Pro Trp Pro Glu Gly Pro		
275	280	285
Met Ala Asn Ala Ala Ala Pro Thr Ser Thr Pro Gln Lys Leu Ile Pro		
290	295	300
Pro Gln Pro Thr Gly Arg Pro Ser Pro Ala Pro Pro Ala Val Pro Pro		
305	310	315
Ala Ala Ser Pro Val Met Pro Pro Gln Thr Gln Ser Pro Gly Gln Pro		
325	330	335
Ala Gln Pro Ala Pro Met Val Pro Leu His Gln Lys Gln Ser Arg Ile		
340	345	350
Thr Pro Ile Gln Lys Pro Arg Gly Leu Asp Pro Val Glu Ile Leu Gln		
355	360	365
Glu Arg Glu Tyr Arg Leu Gln Ala Arg Ile Ala His Arg Ile Gln Glu		
370	375	380
Leu Glu Asn Leu Pro Gly Ser Leu Ala Gly Asp Leu Arg Thr Lys Ala		
385	390	395
Thr Ile Glu Leu Lys Ala Leu Arg Leu Leu Asn Phe Gln Arg Gln Leu		

405								410				415					
Arg	Gln	Glu	Val	Val	Val	Cys	Met	Arg	Arg	Asp	Thr	Ala	Leu	Glu	Thr		
420								425				430					
Ala	Leu	Asn	Ala	Lys	Ala	Tyr	Lys	Arg	Ser	Lys	Arg	Gln	Ser	Leu	Arg		
435								440				445					
Glu	Ala	Arg	Ile	Thr	Glu	Lys	Leu	Glu	Lys	Gln	Gln	Lys	Ile	Glu	Gln		
450								455				460					
Glu	Arg	Lys	Arg	Arg	Gln	Lys	His	Gln	Glu	Tyr	Leu	Asn	Ser	Ile	Leu		
465								470				475				480	
Gln	His	Ala	Lys	Asp	Phe	Lys	Glu	Tyr	His	Arg	Ser	Val	Thr	Gly	Lys		
485								490				495					
Ile	Gln	Lys	Leu	Thr	Lys	Ala	Val	Ala	Thr	Tyr	His	Ala	Asn	Thr	Glu		
500								505				510					
Arg	Glu	Gln	Lys	Lys	Glu	Asn	Glu	Arg	Ile	Glu	Lys	Glu	Arg	Met	Arg		
515								520				525					
Arg	Leu	Met	Ala	Glu	Asp	Glu	Glu	Gly	Tyr	Arg	Lys	Leu	Ile	Asp	Gln		
530								535				540					
Lys	Lys	Asp	Lys	Arg	Leu	Ala	Tyr	Leu	Leu	Gln	Gln	Thr	Asp	Glu	Tyr		
545								550				555				560	
Val	Ala	Asn	Leu	Thr	Glu	Leu	Val	Arg	Gln	His	Lys	Ala	Ala	Gln	Val		
565								570				575					
Ala	Lys	Glu	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Ala	Glu	Asn	Ala		
580								585				590					
Glu	Gly	Gln	Thr	Pro	Ala	Ile	Gly	Pro	Asp	Gly	Glu	Pro	Leu	Asp	Glu		
595								600				605					
Thr	Ser	Gln	Met	Ser	Asp	Leu	Pro	Val	Lys	Val	Ile	His	Val	Glu	Ser		
610								615				620					
Gly	Lys	Ile	Leu	Thr	Gly	Thr	Asp	Ala	Pro	Lys	Ala	Gly	Gln	Leu	Glu		
625								630				635				640	
Ala	Trp	Leu	Glu	Met	Asn	Pro	Gly	Tyr	Glu	Val	Ala	Pro	Arg	Ser	Asp		
645								650				655					
Ser	Glu	Glu	Ser	Gly	Ser	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu		
660								665				670					
Gln	Pro	Gln	Ala	Ala	Gln	Pro	Pro	Thr	Leu	Pro	Val	Glu	Glu	Lys	Lys		
675								680				685					
Lys	Ile	Pro	Asp	Pro	Asp	Ser	Asp	Asp	Val	Ser	Glu	Val	Asp	Ala	Arg		
690								695				700					
His	Ile	Ile	Glu	Asn	Ala	Lys	Gln	Asp	Val	Asp	Asp	Glu	Tyr	Gly	Val		
705								710				715				720	

Ser Gln Ala Leu Ala Arg Gly Leu Gln Ser Tyr Tyr Ala Val Ala His		
725	730	735
Ala Val Thr Glu Arg Val Asp Lys Gln Ser Ala Leu Met Val Asn Gly		
740	745	750
Val Leu Lys Gln Tyr Gln Ile Lys Gly Leu Glu Trp Leu Val Ser Leu		
755	760	765
Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly Leu Gly		
770	775	780
Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu His Lys		
785	790	800
Arg Ile Asn Gly Pro Phe Leu Ile Ile Val Pro Leu Ser Thr Leu Ser		
805	810	815
Asn Trp Ala Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val Lys Val		
820	825	830
Ser Tyr Lys Gly Ser Pro Ala Ala Arg Arg Ala Phe Val Pro Gln Leu		
835	840	845
Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr Ile Ile		
850	855	860
Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met Ile Val		
865	870	880
Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr Gln Val		
885	890	895
Leu Asn Thr His Tyr Val Ala Pro Arg Arg Leu Leu Leu Thr Gly Thr		
900	905	910
Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn Phe Leu		
915	920	925
Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp Phe Asn		
930	935	940
Ala Pro Phe Ala Met Thr Gly Glu Lys Val Asp Leu Asn Glu Glu Glu		
945	950	960
Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
965	970	975
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		
980	985	990
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
995	1000	1005
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
1010	1015	1020
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		

1025	1030	1035
Ile Met Gln Leu Arg Lys	Ile Cys Asn His Pro	Tyr Met Phe Gln
1040	1045	1050
His Ile Glu Glu Ser Phe	Ser Glu His Leu Gly	Phe Thr Gly Gly
1055	1060	1065
Ile Val Gln Gly Leu Asp	Leu Tyr Arg Ala Ser	Gly Lys Phe Glu
1070	1075	1080
Leu Leu Asp Arg Ile Leu	Pro Lys Leu Arg Ala	Thr Asn His Lys
1085	1090	1095
Val Leu Leu Phe Cys Gln	Met Thr Ser Leu Met	Thr Ile Met Glu
1100	1105	1110
Asp Tyr Phe Ala Tyr Arg	Gly Phe Lys Tyr Leu	Arg Leu Asp Gly
1115	1120	1125
Thr Thr Lys Ala Glu Asp	Arg Gly Met Leu Leu	Lys Thr Phe Asn
1130	1135	1140
Glu Pro Gly Ser Glu Tyr	Phe Ile Phe Leu Leu	Ser Thr Arg Ala
1145	1150	1155
Gly Gly Leu Gly Leu Asn	Leu Gln Ser Ala Asp	Thr Val Ile Ile
1160	1165	1170
Phe Asp Ser Asp Trp Asn	Pro His Gln Asp Leu	Gln Ala Gln Asp
1175	1180	1185
Arg Ala His Arg Ile Gly	Gln Gln Asn Glu Val	Arg Val Leu Arg
1190	1195	1200
Leu Cys Thr Val Asn Ser	Val Glu Glu Lys Ile	Leu Ala Ala Ala
1205	1210	1215
Lys Tyr Lys Leu Asn Val	Asp Gln Lys Val Ile	Gln Ala Gly Met
1220	1225	1230
Phe Asp Gln Lys Ser Ser	Ser His Glu Arg Arg	Ala Phe Leu Gln
1235	1240	1245
Ala Ile Leu Glu His Glu	Glu Gln Asp Glu Ser	Arg His Cys Ser
1250	1255	1260
Thr Gly Ser Gly Ser Ala	Ser Phe Ala His Thr	Ala Pro Pro Pro
1265	1270	1275
Ala Gly Val Asn Pro Asp	Leu Glu Glu Pro Pro	Leu Lys Glu Glu
1280	1285	1290
Asp Glu Val Pro Asp Asp	Glu Thr Val Asn Gln	Met Ile Ala Arg
1295	1300	1305
His Glu Glu Glu Phe Asp	Leu Phe Met Arg Met	Asp Leu Asp Arg
1310	1315	1320

Arg Arg Glu Glu Ala Arg Asn Pro Lys Arg Lys Pro Arg Leu Met		
1325	1330	1335
Glu Glu Asp Glu Leu Pro Ser Trp Ile Ile Lys Asp Asp Ala Glu		
1340	1345	1350
Val Glu Arg Leu Thr Cys Glu Glu Glu Glu Lys Met Phe Gly		
1355	1360	1365
Arg Gly Ser Arg His Arg Lys Glu Val Asp Tyr Ser Asp Ser Leu		
1370	1375	1380
Thr Glu Lys Gln Trp Leu Lys Lys Ile Thr Gly Lys Asp Ile His		
1385	1390	1395
Asp Thr Ala Ser Ser Val Ala Arg Gly Leu Gln Phe Gln Arg Gly		
1400	1405	1410
Leu Gln Phe Cys Thr Arg Ala Ser Lys Ala Ile Glu Glu Gly Thr		
1415	1420	1425
Leu Glu Glu Ile Glu Glu Glu Val Arg Gln Lys Lys Ser Ser Arg		
1430	1435	1440
Lys Arg Lys Arg Asp Ser Asp Ala Gly Ser Ser Thr Pro Thr Thr		
1445	1450	1455
Ser Thr Arg Ser Arg Asp Lys Asp Asp Glu Ser Lys Lys Gln Lys		
1460	1465	1470
Lys Arg Gly Arg Pro Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro		
1475	1480	1485
Asn Leu Thr Lys Lys Met Lys Lys Ile Val Asp Ala Val Ile Lys		
1490	1495	1500
Tyr Lys Asp Ser Ser Ser Gly Arg Gln Leu Ser Glu Val Phe Ile		
1505	1510	1515
Gln Leu Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile		
1520	1525	1530
Arg Lys Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn		
1535	1540	1545
His Lys Tyr Arg Ser Leu Asn Asp Leu Glu Lys Asp Val Met Leu		
1550	1555	1560
Leu Cys Gln Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Leu Ile		
1565	1570	1575
Tyr Glu Asp Ser Ile Val Leu Gln Ser Val Phe Thr Ser Val Arg		
1580	1585	1590
Gln Lys Ile Glu Lys Glu Asp Asp Ser Glu Gly Glu Glu Ser Glu		
1595	1600	1605
Glu Glu Glu Glu Gly Glu Glu Glu Gly Ser Glu Ser Glu Ser Arg		

1610	1615	1620
Ser Val Lys Val Lys Ile	Lys Leu Gly Arg Lys	Glu Lys Ala Gln
1625	1630	1635
Asp Arg Leu Lys Gly Gly	Arg Arg Arg Pro Ser	Arg Gly Ser Arg
1640	1645	1650
Ala Lys Pro Val Val Ser	Asp Asp Asp Ser Glu	Glu Glu Gln Glu
1655	1660	1665
Glu Asp Arg Ser Gly Ser	Gly Ser Glu Glu Asp	
1670	1675	
<210> 16		
<211> 1647		
<212> PRT		
<213> 智人		
<400> 16		
Met Ser Thr Pro Asp Pro Pro	Leu Gly Gly Thr Pro Arg Pro	Gly Pro
1	5	10
Ser Pro Gly Pro Gly Pro Ser	Pro Gly Ala Met Leu Gly Pro	Ser Pro
20	25	30
Gly Pro Ser Pro Gly Ser Ala	His Ser Met Met Gly Pro Ser	Pro Gly
35	40	45
Pro Pro Ser Ala Gly His Pro	Ile Pro Thr Gln Gly Pro Gly	Gly Tyr
50	55	60
Pro Gln Asp Asn Met His Gln	Met His Lys Pro Met Glu Ser	Met His
65	70	75
Glu Lys Gly Met Ser Asp Asp	Pro Arg Tyr Asn Gln Met Lys	Gly Met
85	90	95
Gly Met Arg Ser Gly Gly His	Ala Gly Met Gly Pro Pro Pro	Ser Pro
100	105	110
Met Asp Gln His Ser Gln Gly	Tyr Pro Ser Pro Leu Gly Gly	Ser Glu
115	120	125
His Ala Ser Ser Pro Val Pro	Ala Ser Gly Pro Ser Ser Gly	Pro Gln
130	135	140
Met Ser Ser Gly Pro Gly Gly	Ala Pro Leu Asp Gly Ala Asp	Pro Gln
145	150	155
Ala Leu Gly Gln Gln Asn Arg	Gly Pro Thr Pro Phe Asn Gln	Asn Gln
165	170	175
Leu His Gln Leu Arg Ala Gln	Ile Met Ala Tyr Lys Met Leu	Ala Arg
180	185	190
Gly Gln Pro Leu Pro Asp His	Leu Gln Met Ala Val Gln Gly	Lys Arg

195	200	205
Pro Met Pro Gly Met Gln Gln Gln Met Pro Thr Leu Pro Pro Pro Ser		
210	215	220
Val Ser Ala Thr Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro		
225	230	235
Gly Pro Gly Pro Ala Pro Pro Asn Tyr Ser Arg Pro His Gly Met Gly		240
245	250	255
Gly Pro Asn Met Pro Pro Pro Gly Pro Ser Gly Val Pro Pro Gly Met		
260	265	270
Pro Gly Gln Pro Pro Gly Gly Pro Pro Lys Pro Trp Pro Glu Gly Pro		
275	280	285
Met Ala Asn Ala Ala Ala Pro Thr Ser Thr Pro Gln Lys Leu Ile Pro		
290	295	300
Pro Gln Pro Thr Gly Arg Pro Ser Pro Ala Pro Pro Ala Val Pro Pro		
305	310	315
Ala Ala Ser Pro Val Met Pro Pro Gln Thr Gln Ser Pro Gly Gln Pro		
325	330	335
Ala Gln Pro Ala Pro Met Val Pro Leu His Gln Lys Gln Ser Arg Ile		
340	345	350
Thr Pro Ile Gln Lys Pro Arg Gly Leu Asp Pro Val Glu Ile Leu Gln		
355	360	365
Glu Arg Glu Tyr Arg Leu Gln Ala Arg Ile Ala His Arg Ile Gln Glu		
370	375	380
Leu Glu Asn Leu Pro Gly Ser Leu Ala Gly Asp Leu Arg Thr Lys Ala		
385	390	395
Thr Ile Glu Leu Lys Ala Leu Arg Leu Leu Asn Phe Gln Arg Gln Leu		
405	410	415
Arg Gln Glu Val Val Val Cys Met Arg Arg Asp Thr Ala Leu Glu Thr		
420	425	430
Ala Leu Asn Ala Lys Ala Tyr Lys Arg Ser Lys Arg Gln Ser Leu Arg		
435	440	445
Glu Ala Arg Ile Thr Glu Lys Leu Glu Lys Gln Gln Lys Ile Glu Gln		
450	455	460
Glu Arg Lys Arg Arg Gln Lys His Gln Glu Tyr Leu Asn Ser Ile Leu		
465	470	475
Gln His Ala Lys Asp Phe Lys Glu Tyr His Arg Ser Val Thr Gly Lys		480
485	490	495
Ile Gln Lys Leu Thr Lys Ala Val Ala Thr Tyr His Ala Asn Thr Glu		
500	505	510

Arg Glu Gln Lys Lys Glu Asn Glu Arg Ile Glu Lys Glu Arg Met Arg		
515	520	525
Arg Leu Met Ala Glu Asp Glu Gly Tyr Arg Lys Leu Ile Asp Gln		
530	535	540
Lys Lys Asp Lys Arg Leu Ala Tyr Leu Leu Gln Gln Thr Asp Glu Tyr		
545	550	555
Val Ala Asn Leu Thr Glu Leu Val Arg Gln His Lys Ala Ala Gln Val		
565	570	575
Ala Lys Glu Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Ala Glu Asn Ala		
580	585	590
Glu Gly Gln Thr Pro Ala Ile Gly Pro Asp Gly Glu Pro Leu Asp Glu		
595	600	605
Thr Ser Gln Met Ser Asp Leu Pro Val Lys Val Ile His Val Glu Ser		
610	615	620
Gly Lys Ile Leu Thr Gly Thr Asp Ala Pro Lys Ala Gly Gln Leu Glu		
625	630	635
Ala Trp Leu Glu Met Asn Pro Gly Tyr Glu Val Ala Pro Arg Ser Asp		
645	650	655
Ser Glu Glu Ser Gly Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu		
660	665	670
Gln Pro Gln Ala Ala Gln Pro Pro Thr Leu Pro Val Glu Glu Lys Lys		
675	680	685
Lys Ile Pro Asp Pro Asp Ser Asp Asp Val Ser Glu Val Asp Ala Arg		
690	695	700
His Ile Ile Glu Asn Ala Lys Gln Asp Val Asp Asp Glu Tyr Gly Val		
705	710	715
Ser Gln Ala Leu Ala Arg Gly Leu Gln Ser Tyr Tyr Ala Val Ala His		
725	730	735
Ala Val Thr Glu Arg Val Asp Lys Gln Ser Ala Leu Met Val Asn Gly		
740	745	750
Val Leu Lys Gln Tyr Gln Ile Lys Gly Leu Glu Trp Leu Val Ser Leu		
755	760	765
Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly Leu Gly		
770	775	780
Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu His Lys		
785	790	795
Arg Ile Asn Gly Pro Phe Leu Ile Ile Val Pro Leu Ser Thr Leu Ser		
805	810	815
Asn Trp Ala Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val Lys Val		

820	825	830
Ser Tyr Lys Gly Ser Pro Ala Ala Arg Arg Ala Phe Val Pro Gln Leu		
835	840	845
Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr Ile Ile		
850	855	860
Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met Ile Val		
865	870	875
Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr Gln Val		880
885	890	895
Leu Asn Thr His Tyr Val Ala Pro Arg Arg Leu Leu Leu Thr Gly Thr		
900	905	910
Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn Phe Leu		
915	920	925
Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp Phe Asn		
930	935	940
Ala Pro Phe Ala Met Thr Gly Glu Lys Val Asp Leu Asn Glu Glu Glu		
945	950	955
Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
965	970	975
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		
980	985	990
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
995	1000	1005
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
1010	1015	1020
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		
1025	1030	1035
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1040	1045	1050
His Ile Glu Glu Ser Phe Ser Glu His Leu Gly Phe Thr Gly Gly		
1055	1060	1065
Ile Val Gln Gly Leu Asp Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1070	1075	1080
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Lys		
1085	1090	1095
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
1100	1105	1110
Asp Tyr Phe Ala Tyr Arg Gly Phe Lys Tyr Leu Arg Leu Asp Gly		
1115	1120	1125

Thr Thr Lys Ala Glu Asp Arg Gly Met Leu Leu Lys Thr Phe Asn		
1130	1135	1140
Glu Pro Gly Ser Glu Tyr Phe Ile Phe Leu Leu Ser Thr Arg Ala		
1145	1150	1155
Gly Gly Leu Gly Leu Asn Leu Gln Ser Ala Asp Thr Val Ile Ile		
1160	1165	1170
Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala Gln Asp		
1175	1180	1185
Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val Leu Arg		
1190	1195	1200
Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala Ala Ala		
1205	1210	1215
Lys Tyr Lys Leu Asn Val Asp Gln Lys Val Ile Gln Ala Gly Met		
1220	1225	1230
Phe Asp Gln Lys Ser Ser Ser His Glu Arg Arg Ala Phe Leu Gln		
1235	1240	1245
Ala Ile Leu Glu His Glu Glu Gln Asp Glu Ser Arg His Cys Ser		
1250	1255	1260
Thr Gly Ser Gly Ser Ala Ser Phe Ala His Thr Ala Pro Pro Pro		
1265	1270	1275
Ala Gly Val Asn Pro Asp Leu Glu Glu Pro Pro Leu Lys Glu Glu		
1280	1285	1290
Asp Glu Val Pro Asp Asp Glu Thr Val Asn Gln Met Ile Ala Arg		
1295	1300	1305
His Glu Glu Glu Phe Asp Leu Phe Met Arg Met Asp Leu Asp Arg		
1310	1315	1320
Arg Arg Glu Glu Ala Arg Asn Pro Lys Arg Lys Pro Arg Leu Met		
1325	1330	1335
Glu Glu Asp Glu Leu Pro Ser Trp Ile Ile Lys Asp Asp Ala Glu		
1340	1345	1350
Val Glu Arg Leu Thr Cys Glu Glu Glu Glu Glu Lys Met Phe Gly		
1355	1360	1365
Arg Gly Ser Arg His Arg Lys Glu Val Asp Tyr Ser Asp Ser Leu		
1370	1375	1380
Thr Glu Lys Gln Trp Leu Lys Ala Ile Glu Glu Gly Thr Leu Glu		
1385	1390	1395
Glu Ile Glu Glu Glu Val Arg Gln Lys Lys Ser Ser Arg Lys Arg		
1400	1405	1410
Lys Arg Asp Ser Asp Ala Gly Ser Ser Thr Pro Thr Thr Ser Thr		

1415	1420	1425
Arg Ser Arg Asp Lys Asp Asp Glu Ser Lys Lys Gln Lys Lys Arg		
1430	1435	1440
Gly Arg Pro Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro Asn Leu		
1445	1450	1455
Thr Lys Lys Met Lys Lys Ile Val Asp Ala Val Ile Lys Tyr Lys		
1460	1465	1470
Asp Ser Ser Ser Gly Arg Gln Leu Ser Glu Val Phe Ile Gln Leu		
1475	1480	1485
Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile Arg Lys		
1490	1495	1500
Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn His Lys		
1505	1510	1515
Tyr Arg Ser Leu Asn Asp Leu Glu Lys Asp Val Met Leu Leu Cys		
1520	1525	1530
Gln Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Leu Ile Tyr Glu		
1535	1540	1545
Asp Ser Ile Val Leu Gln Ser Val Phe Thr Ser Val Arg Gln Lys		
1550	1555	1560
Ile Glu Lys Glu Asp Asp Ser Glu Gly Glu Glu Ser Glu Glu Glu		
1565	1570	1575
Glu Glu Gly Glu Glu Glu Gly Ser Glu Ser Glu Ser Arg Ser Val		
1580	1585	1590
Lys Val Lys Ile Lys Leu Gly Arg Lys Glu Lys Ala Gln Asp Arg		
1595	1600	1605
Leu Lys Gly Gly Arg Arg Arg Pro Ser Arg Gly Ser Arg Ala Lys		
1610	1615	1620
Pro Val Val Ser Asp Asp Asp Ser Glu Glu Glu Gln Glu Glu Asp		
1625	1630	1635
Arg Ser Gly Ser Gly Ser Glu Glu Asp		
1640	1645	
<210> 17		
<211> 1617		
<212> PRT		
<213> 智人		
<400> 17		
Met Ser Thr Pro Asp Pro Pro Leu Gly Gly Thr Pro Arg Pro Gly Pro		
1	5	10
Ser Pro Gly Pro Gly Pro Ser Pro Gly Ala Met Leu Gly Pro Ser Pro		

20	25	30
Gly Pro Ser Pro Gly Ser Ala His	Ser Met Met Gly Pro Ser Pro Gly	
35	40	45
Pro Pro Ser Ala Gly His Pro Ile	Pro Thr Gln Gly Pro Gly Gly Tyr	
50	55	60
Pro Gln Asp Asn Met His Gln Met	His Lys Pro Met Glu Ser Met His	
65	70	75
Glu Lys Gly Met Ser Asp Asp Pro	Arg Tyr Asn Gln Met Lys Gly Met	
85	90	95
Gly Met Arg Ser Gly Gly His Ala	Gly Met Gly Pro Pro Pro Ser Pro	
100	105	110
Met Asp Gln His Ser Gln Gly Tyr	Pro Ser Pro Leu Gly Gly Ser Glu	
115	120	125
His Ala Ser Ser Pro Val Pro Ala	Ser Gly Pro Ser Ser Gly Pro Gln	
130	135	140
Met Ser Ser Gly Pro Gly Gly Ala	Pro Leu Asp Gly Ala Asp Pro Gln	
145	150	155
Ala Leu Gly Gln Gln Asn Arg Gly	Pro Thr Pro Phe Asn Gln Asn Gln	
165	170	175
Leu His Gln Leu Arg Ala Gln Ile	Met Ala Tyr Lys Met Leu Ala Arg	
180	185	190
Gly Gln Pro Leu Pro Asp His Leu	Gln Met Ala Val Gln Gly Lys Arg	
195	200	205
Pro Met Pro Gly Met Gln Gln Gln	Met Pro Thr Leu Pro Pro Pro Ser	
210	215	220
Val Ser Ala Thr Gly Pro Gly Pro	Gly Pro Gly Pro Gly Pro Gly Pro	
225	230	235
Gly Pro Gly Pro Ala Pro Pro Asn	Tyr Ser Arg Pro His Gly Met Gly	
245	250	255
Gly Pro Asn Met Pro Pro Pro Gly	Pro Ser Gly Val Pro Pro Gly Met	
260	265	270
Pro Gly Gln Pro Pro Gly Gly Pro	Pro Lys Pro Trp Pro Glu Gly Pro	
275	280	285
Met Ala Asn Ala Ala Ala Pro Thr	Ser Thr Pro Gln Lys Leu Ile Pro	
290	295	300
Pro Gln Pro Thr Gly Arg Pro Ser	Pro Ala Pro Pro Ala Val Pro Pro	
305	310	315
Ala Ala Ser Pro Val Met Pro Pro	Gln Thr Gln Ser Pro Gly Gln Pro	
325	330	335

Ala Gln Pro Ala Pro Met Val Pro Leu His Gln Lys Gln Ser Arg Ile		
340	345	350
Thr Pro Ile Gln Lys Pro Arg Gly Leu Asp Pro Val Glu Ile Leu Gln		
355	360	365
Glu Arg Glu Tyr Arg Leu Gln Ala Arg Ile Ala His Arg Ile Gln Glu		
370	375	380
Leu Glu Asn Leu Pro Gly Ser Leu Ala Gly Asp Leu Arg Thr Lys Ala		
385	390	395
Thr Ile Glu Leu Lys Ala Leu Arg Leu Leu Asn Phe Gln Arg Gln Leu		
405	410	415
Arg Gln Glu Val Val Val Cys Met Arg Arg Asp Thr Ala Leu Glu Thr		
420	425	430
Ala Leu Asn Ala Lys Ala Tyr Lys Arg Ser Lys Arg Gln Ser Leu Arg		
435	440	445
Glu Ala Arg Ile Thr Glu Lys Leu Glu Lys Gln Gln Lys Ile Glu Gln		
450	455	460
Glu Arg Lys Arg Arg Gln Lys His Gln Glu Tyr Leu Asn Ser Ile Leu		
465	470	475
Gln His Ala Lys Asp Phe Lys Glu Tyr His Arg Ser Val Thr Gly Lys		
485	490	495
Ile Gln Lys Leu Thr Lys Ala Val Ala Thr Tyr His Ala Asn Thr Glu		
500	505	510
Arg Glu Gln Lys Lys Glu Asn Glu Arg Ile Glu Lys Glu Arg Met Arg		
515	520	525
Arg Leu Met Ala Glu Asp Glu Glu Gly Tyr Arg Lys Leu Ile Asp Gln		
530	535	540
Lys Lys Asp Lys Arg Leu Ala Tyr Leu Leu Gln Gln Thr Asp Glu Tyr		
545	550	555
Val Ala Asn Leu Thr Glu Leu Val Arg Gln His Lys Ala Ala Gln Val		
565	570	575
Ala Lys Glu Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Ala Glu Asn Ala		
580	585	590
Glu Gly Gln Thr Pro Ala Ile Gly Pro Asp Gly Glu Pro Leu Asp Glu		
595	600	605
Thr Ser Gln Met Ser Asp Leu Pro Val Lys Val Ile His Val Glu Ser		
610	615	620
Gly Lys Ile Leu Thr Gly Thr Asp Ala Pro Lys Ala Gly Gln Leu Glu		
625	630	635
Ala Trp Leu Glu Met Asn Pro Gly Tyr Glu Val Ala Pro Arg Ser Asp		

			645						650						655			
Ser	Glu	Glu	Ser	Gly	Ser	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu		
			660						665						670			
Gln	Pro	Gln	Ala	Ala	Gln	Pro	Pro	Thr	Leu	Pro	Val	Glu	Glu	Lys	Lys			
			675						680						685			
Lys	Ile	Pro	Asp	Pro	Asp	Ser	Asp	Asp	Val	Ser	Glu	Val	Asp	Ala	Arg			
			690						695						700			
His	Ile	Ile	Glu	Asn	Ala	Lys	Gln	Asp	Val	Asp	Asp	Glu	Tyr	Gly	Val			
705				710						715						720		
Ser	Gln	Ala	Leu	Ala	Arg	Gly	Leu	Gln	Ser	Tyr	Tyr	Ala	Val	Ala	His			
			725						730						735			
Ala	Val	Thr	Glu	Arg	Val	Asp	Lys	Gln	Ser	Ala	Leu	Met	Val	Asn	Gly			
			740						745						750			
Val	Leu	Lys	Gln	Tyr	Gln	Ile	Lys	Gly	Leu	Glu	Trp	Leu	Val	Ser	Leu			
			755						760						765			
Tyr	Asn	Asn	Asn	Leu	Asn	Gly	Ile	Leu	Ala	Asp	Glu	Met	Gly	Leu	Gly			
			770						775						780			
Lys	Thr	Ile	Gln	Thr	Ile	Ala	Leu	Ile	Thr	Tyr	Leu	Met	Glu	His	Lys			
785				790						795						800		
Arg	Ile	Asn	Gly	Pro	Phe	Leu	Ile	Ile	Val	Pro	Leu	Ser	Thr	Leu	Ser			
			805						810						815			
Asn	Trp	Ala	Tyr	Glu	Phe	Asp	Lys	Trp	Ala	Pro	Ser	Val	Val	Lys	Val			
			820						825						830			
Ser	Tyr	Lys	Gly	Ser	Pro	Ala	Ala	Arg	Arg	Ala	Phe	Val	Pro	Gln	Leu			
			835						840						845			
Arg	Ser	Gly	Lys	Phe	Asn	Val	Leu	Leu	Thr	Thr	Tyr	Glu	Tyr	Ile	Ile			
			850						855						860			
Lys	Asp	Lys	His	Ile	Leu	Ala	Lys	Ile	Arg	Trp	Lys	Tyr	Met	Ile	Val			
865				870						875						880		
Asp	Glu	Gly	His	Arg	Met	Lys	Asn	His	His	Cys	Lys	Leu	Thr	Gln	Val			
			885						890						895			
Leu	Asn	Thr	His	Tyr	Val	Ala	Pro	Arg	Arg	Leu	Leu	Leu	Thr	Gly	Thr			
			900						905						910			
Pro	Leu	Gln	Asn	Lys	Leu	Pro	Glu	Leu	Trp	Ala	Leu	Leu	Asn	Phe	Leu			
			915						920						925			
Leu	Pro	Thr	Ile	Phe	Lys	Ser	Cys	Ser	Thr	Phe	Glu	Gln	Trp	Phe	Asn			
			930						935						940			
Ala	Pro	Phe	Ala	Met	Thr	Gly	Glu	Lys	Val	Asp	Leu	Asn	Glu	Glu	Glu			
945				950						955						960		

Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
965	970	975
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		
980	985	990
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
995	1000	1005
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
1010	1015	1020
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		
1025	1030	1035
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1040	1045	1050
His Ile Glu Glu Ser Phe Ser Glu His Leu Gly Phe Thr Gly Gly		
1055	1060	1065
Ile Val Gln Gly Leu Asp Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1070	1075	1080
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Lys		
1085	1090	1095
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
1100	1105	1110
Asp Tyr Phe Ala Tyr Arg Gly Phe Lys Tyr Leu Arg Leu Asp Gly		
1115	1120	1125
Thr Thr Lys Ala Glu Asp Arg Gly Met Leu Leu Lys Thr Phe Asn		
1130	1135	1140
Glu Pro Gly Ser Glu Tyr Phe Ile Phe Leu Leu Ser Thr Arg Ala		
1145	1150	1155
Gly Gly Leu Gly Leu Asn Leu Gln Ser Ala Asp Thr Val Ile Ile		
1160	1165	1170
Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala Gln Asp		
1175	1180	1185
Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val Leu Arg		
1190	1195	1200
Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala Ala Ala		
1205	1210	1215
Lys Tyr Lys Leu Asn Val Asp Gln Lys Val Ile Gln Ala Gly Met		
1220	1225	1230
Phe Asp Gln Lys Ser Ser Ser His Glu Arg Arg Ala Phe Leu Gln		
1235	1240	1245
Ala Ile Leu Glu His Glu Glu Gln Asp Glu Glu Glu Asp Glu Val		

1250	1255	1260
Pro Asp Asp Glu Thr Val	Asn Gln Met Ile Ala	Arg His Glu Glu
1265	1270	1275
Glu Phe Asp Leu Phe Met	Arg Met Asp Leu Asp	Arg Arg Arg Glu
1280	1285	1290
Glu Ala Arg Asn Pro Lys	Arg Lys Pro Arg Leu	Met Glu Glu Asp
1295	1300	1305
Glu Leu Pro Ser Trp Ile	Ile Lys Asp Asp Ala	Glu Val Glu Arg
1310	1315	1320
Leu Thr Cys Glu Glu Glu	Glu Glu Lys Met Phe	Gly Arg Gly Ser
1325	1330	1335
Arg His Arg Lys Glu Val	Asp Tyr Ser Asp Ser	Leu Thr Glu Lys
1340	1345	1350
Gln Trp Leu Lys Thr Leu	Lys Ala Ile Glu Glu	Gly Thr Leu Glu
1355	1360	1365
Glu Ile Glu Glu Glu Val	Arg Gln Lys Lys Ser	Ser Arg Lys Arg
1370	1375	1380
Lys Arg Asp Ser Asp Ala	Gly Ser Ser Thr Pro	Thr Thr Ser Thr
1385	1390	1395
Arg Ser Arg Asp Lys Asp	Asp Glu Ser Lys Lys	Gln Lys Lys Arg
1400	1405	1410
Gly Arg Pro Pro Ala Glu	Lys Leu Ser Pro Asn	Pro Pro Asn Leu
1415	1420	1425
Thr Lys Lys Met Lys Lys	Ile Val Asp Ala Val	Ile Lys Tyr Lys
1430	1435	1440
Asp Ser Ser Ser Gly Arg	Gln Leu Ser Glu Val	Phe Ile Gln Leu
1445	1450	1455
Pro Ser Arg Lys Glu Leu	Pro Glu Tyr Tyr Glu	Leu Ile Arg Lys
1460	1465	1470
Pro Val Asp Phe Lys Lys	Ile Lys Glu Arg Ile	Arg Asn His Lys
1475	1480	1485
Tyr Arg Ser Leu Asn Asp	Leu Glu Lys Asp Val	Met Leu Leu Cys
1490	1495	1500
Gln Asn Ala Gln Thr Phe	Asn Leu Glu Gly Ser	Leu Ile Tyr Glu
1505	1510	1515
Asp Ser Ile Val Leu Gln	Ser Val Phe Thr Ser	Val Arg Gln Lys
1520	1525	1530
Ile Glu Lys Glu Asp Asp	Ser Glu Gly Glu Glu	Ser Glu Glu Glu
1535	1540	1545

Glu	Glu	Gly	Glu	Glu	Glu	Gly	Ser	Glu	Ser	Glu	Ser	Arg	Ser	Val	
1550						1555					1560				
Lys	Val	Lys	Ile	Lys	Leu	Gly	Arg	Lys	Glu	Lys	Ala	Gln	Asp	Arg	
1565						1570					1575				
Leu	Lys	Gly	Gly	Arg	Arg	Arg	Pro	Ser	Arg	Gly	Ser	Arg	Ala	Lys	
1580						1585					1590				
Pro	Val	Val	Ser	Asp	Asp	Asp	Ser	Glu	Glu	Glu	Gln	Glu	Glu	Asp	
1595						1600					1605				
Arg	Ser	Gly	Ser	Gly	Ser	Glu	Glu	Asp							
1610						1615									
<210>	18														
<211>	1616														
<212>	PRT														
<213>	智人														
<400>	18														
Met	Ser	Thr	Pro	Asp	Pro	Pro	Leu	Gly	Gly	Thr	Pro	Arg	Pro	Gly	Pro
1			5					10					15		
Ser	Pro	Gly	Pro	Gly	Pro	Ser	Pro	Gly	Ala	Met	Leu	Gly	Pro	Ser	Pro
			20					25					30		
Gly	Pro	Ser	Pro	Gly	Ser	Ala	His	Ser	Met	Met	Gly	Pro	Ser	Pro	Gly
			35				40						45		
Pro	Pro	Ser	Ala	Gly	His	Pro	Ile	Pro	Thr	Gln	Gly	Pro	Gly	Gly	Tyr
			50				55				60				
Pro	Gln	Asp	Asn	Met	His	Gln	Met	His	Lys	Pro	Met	Glu	Ser	Met	His
65				70						75				80	
Glu	Lys	Gly	Met	Ser	Asp	Asp	Pro	Arg	Tyr	Asn	Gln	Met	Lys	Gly	Met
				85					90					95	
Gly	Met	Arg	Ser	Gly	Gly	His	Ala	Gly	Met	Gly	Pro	Pro	Pro	Ser	Pro
			100					105						110	
Met	Asp	Gln	His	Ser	Gln	Gly	Tyr	Pro	Ser	Pro	Leu	Gly	Gly	Ser	Glu
			115				120						125		
His	Ala	Ser	Ser	Pro	Val	Pro	Ala	Ser	Gly	Pro	Ser	Ser	Gly	Pro	Gln
			130				135						140		
Met	Ser	Ser	Gly	Pro	Gly	Gly	Ala	Pro	Leu	Asp	Gly	Ala	Asp	Pro	Gln
145					150					155					160
Ala	Leu	Gly	Gln	Gln	Asn	Arg	Gly	Pro	Thr	Pro	Phe	Asn	Gln	Asn	Gln
					165				170					175	
Leu	His	Gln	Leu	Arg	Ala	Gln	Ile	Met	Ala	Tyr	Lys	Met	Leu	Ala	Arg
			180					185						190	

Gly Gln Pro Leu Pro Asp His Leu Gln Met Ala Val Gln Gly Lys Arg		
195	200	205
Pro Met Pro Gly Met Gln Gln Gln Met Pro Thr Leu Pro Pro Pro Ser		
210	215	220
Val Ser Ala Thr Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro		
225	230	235
Gly Pro Gly Pro Ala Pro Pro Asn Tyr Ser Arg Pro His Gly Met Gly		
	245	250
		255
Gly Pro Asn Met Pro Pro Pro Gly Pro Ser Gly Val Pro Pro Gly Met		
	260	265
		270
Pro Gly Gln Pro Pro Gly Gly Pro Pro Lys Pro Trp Pro Glu Gly Pro		
	275	280
		285
Met Ala Asn Ala Ala Ala Pro Thr Ser Thr Pro Gln Lys Leu Ile Pro		
	290	295
		300
Pro Gln Pro Thr Gly Arg Pro Ser Pro Ala Pro Pro Ala Val Pro Pro		
305	310	315
		320
Ala Ala Ser Pro Val Met Pro Pro Gln Thr Gln Ser Pro Gly Gln Pro		
	325	330
		335
Ala Gln Pro Ala Pro Met Val Pro Leu His Gln Lys Gln Ser Arg Ile		
	340	345
		350
Thr Pro Ile Gln Lys Pro Arg Gly Leu Asp Pro Val Glu Ile Leu Gln		
	355	360
		365
Glu Arg Glu Tyr Arg Leu Gln Ala Arg Ile Ala His Arg Ile Gln Glu		
	370	375
		380
Leu Glu Asn Leu Pro Gly Ser Leu Ala Gly Asp Leu Arg Thr Lys Ala		
385	390	395
		400
Thr Ile Glu Leu Lys Ala Leu Arg Leu Leu Asn Phe Gln Arg Gln Leu		
	405	410
		415
Arg Gln Glu Val Val Val Cys Met Arg Arg Asp Thr Ala Leu Glu Thr		
	420	425
		430
Ala Leu Asn Ala Lys Ala Tyr Lys Arg Ser Lys Arg Gln Ser Leu Arg		
	435	440
		445
Glu Ala Arg Ile Thr Glu Lys Leu Glu Lys Gln Gln Lys Ile Glu Gln		
	450	455
		460
Glu Arg Lys Arg Arg Gln Lys His Gln Glu Tyr Leu Asn Ser Ile Leu		
465	470	475
		480
Gln His Ala Lys Asp Phe Lys Glu Tyr His Arg Ser Val Thr Gly Lys		
	485	490
		495
Ile Gln Lys Leu Thr Lys Ala Val Ala Thr Tyr His Ala Asn Thr Glu		

500	505	510
Arg Glu Gln Lys Lys Glu Asn Glu Arg Ile Glu Lys Glu Arg Met Arg		
515	520	525
Arg Leu Met Ala Glu Asp Glu Glu Gly Tyr Arg Lys Leu Ile Asp Gln		
530	535	540
Lys Lys Asp Lys Arg Leu Ala Tyr Leu Leu Gln Gln Thr Asp Glu Tyr		
545	550	555
Val Ala Asn Leu Thr Glu Leu Val Arg Gln His Lys Ala Ala Gln Val		
565	570	575
Ala Lys Glu Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Ala Glu Asn Ala		
580	585	590
Glu Gly Gln Thr Pro Ala Ile Gly Pro Asp Gly Glu Pro Leu Asp Glu		
595	600	605
Thr Ser Gln Met Ser Asp Leu Pro Val Lys Val Ile His Val Glu Ser		
610	615	620
Gly Lys Ile Leu Thr Gly Thr Asp Ala Pro Lys Ala Gly Gln Leu Glu		
625	630	635
Ala Trp Leu Glu Met Asn Pro Gly Tyr Glu Val Ala Pro Arg Ser Asp		
645	650	655
Ser Glu Glu Ser Gly Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu		
660	665	670
Gln Pro Gln Ala Ala Gln Pro Pro Thr Leu Pro Val Glu Glu Lys Lys		
675	680	685
Lys Ile Pro Asp Pro Asp Ser Asp Asp Val Ser Glu Val Asp Ala Arg		
690	695	700
His Ile Ile Glu Asn Ala Lys Gln Asp Val Asp Asp Glu Tyr Gly Val		
705	710	715
Ser Gln Ala Leu Ala Arg Gly Leu Gln Ser Tyr Tyr Ala Val Ala His		
725	730	735
Ala Val Thr Glu Arg Val Asp Lys Gln Ser Ala Leu Met Val Asn Gly		
740	745	750
Val Leu Lys Gln Tyr Gln Ile Lys Gly Leu Glu Trp Leu Val Ser Leu		
755	760	765
Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly Leu Gly		
770	775	780
Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu His Lys		
785	790	795
Arg Ile Asn Gly Pro Phe Leu Ile Ile Val Pro Leu Ser Thr Leu Ser		
805	810	815

Asn Trp Ala Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val Lys Val		
820	825	830
Ser Tyr Lys Gly Ser Pro Ala Ala Arg Arg Ala Phe Val Pro Gln Leu		
835	840	845
Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr Ile Ile		
850	855	860
Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met Ile Val		
865	870	875
Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr Gln Val		
	885	890
Leu Asn Thr His Tyr Val Ala Pro Arg Arg Leu Leu Leu Thr Gly Thr		
	900	905
Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn Phe Leu		
	915	920
Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp Phe Asn		
	930	935
Ala Pro Phe Ala Met Thr Gly Glu Lys Val Asp Leu Asn Glu Glu Glu		
945	950	955
Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
	965	970
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		
	980	985
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
	995	1000
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
	1010	1015
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		
	1025	1030
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
	1040	1045
His Ile Glu Glu Ser Phe Ser Glu His Leu Gly Phe Thr Gly Gly		
	1055	1060
Ile Val Gln Gly Leu Asp Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
	1070	1075
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Lys		
	1085	1090
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
	1100	1105
Asp Tyr Phe Ala Tyr Arg Gly Phe Lys Tyr Leu Arg Leu Asp Gly		

1115	1120	1125
Thr Thr Lys Ala Glu Asp	Arg Gly Met Leu Leu	Lys Thr Phe Asn
1130	1135	1140
Glu Pro Gly Ser Glu Tyr	Phe Ile Phe Leu Leu	Ser Thr Arg Ala
1145	1150	1155
Gly Gly Leu Gly Leu Asn	Leu Gln Ser Ala Asp	Thr Val Ile Ile
1160	1165	1170
Phe Asp Ser Asp Trp Asn	Pro His Gln Asp Leu	Gln Ala Gln Asp
1175	1180	1185
Arg Ala His Arg Ile Gly	Gln Gln Asn Glu Val	Arg Val Leu Arg
1190	1195	1200
Leu Cys Thr Val Asn Ser	Val Glu Glu Lys Ile	Leu Ala Ala Ala
1205	1210	1215
Lys Tyr Lys Leu Asn Val	Asp Gln Lys Val Ile	Gln Ala Gly Met
1220	1225	1230
Phe Asp Gln Lys Ser Ser	Ser His Glu Arg Arg	Ala Phe Leu Gln
1235	1240	1245
Ala Ile Leu Glu His Glu	Glu Gln Asp Glu Glu	Glu Asp Glu Val
1250	1255	1260
Pro Asp Asp Glu Thr Val	Asn Gln Met Ile Ala	Arg His Glu Glu
1265	1270	1275
Glu Phe Asp Leu Phe Met	Arg Met Asp Leu Asp	Arg Arg Arg Glu
1280	1285	1290
Glu Ala Arg Asn Pro Lys	Arg Lys Pro Arg Leu	Met Glu Glu Asp
1295	1300	1305
Glu Leu Pro Ser Trp Ile	Ile Lys Asp Asp Ala	Glu Val Glu Arg
1310	1315	1320
Leu Thr Cys Glu Glu Glu	Glu Glu Lys Met Phe	Gly Arg Gly Ser
1325	1330	1335
Arg His Arg Lys Glu Val	Asp Tyr Ser Asp Ser	Leu Thr Glu Lys
1340	1345	1350
Gln Trp Leu Lys Thr Leu	Lys Ala Ile Glu Glu	Gly Thr Leu Glu
1355	1360	1365
Glu Ile Glu Glu Glu Val	Arg Gln Lys Lys Ser	Ser Arg Lys Arg
1370	1375	1380
Lys Arg Asp Ser Asp Ala	Gly Ser Ser Thr Pro	Thr Thr Ser Thr
1385	1390	1395
Arg Ser Arg Asp Lys Asp	Asp Glu Ser Lys Lys	Gln Lys Lys Arg
1400	1405	1410

Gly Arg Pro Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro Asn Leu		
1415	1420	1425
Thr Lys Lys Met Lys Lys Ile Val Asp Ala Val Ile Lys Tyr Lys		
1430	1435	1440
Asp Ser Ser Gly Arg Gln Leu Ser Glu Val Phe Ile Gln Leu Pro		
1445	1450	1455
Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile Arg Lys Pro		
1460	1465	1470
Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn His Lys Tyr		
1475	1480	1485
Arg Ser Leu Asn Asp Leu Glu Lys Asp Val Met Leu Leu Cys Gln		
1490	1495	1500
Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Leu Ile Tyr Glu Asp		
1505	1510	1515
Ser Ile Val Leu Gln Ser Val Phe Thr Ser Val Arg Gln Lys Ile		
1520	1525	1530
Glu Lys Glu Asp Asp Ser Glu Gly Glu Glu Ser Glu Glu Glu Glu		
1535	1540	1545
Glu Gly Glu Glu Glu Gly Ser Glu Ser Glu Ser Arg Ser Val Lys		
1550	1555	1560
Val Lys Ile Lys Leu Gly Arg Lys Glu Lys Ala Gln Asp Arg Leu		
1565	1570	1575
Lys Gly Gly Arg Arg Arg Pro Ser Arg Gly Ser Arg Ala Lys Pro		
1580	1585	1590
Val Val Ser Asp Asp Asp Ser Glu Glu Glu Gln Glu Glu Asp Arg		
1595	1600	1605
Ser Gly Ser Gly Ser Glu Glu Asp		
1610	1615	

<210> 19

<211> 1614

<212> PRT

<213> 智人

<400> 19

Met Ser Thr Pro Asp Pro Pro Leu Gly Gly Thr Pro Arg Pro Gly Pro		
1	5	10
Ser Pro Gly Pro Gly Pro Ser Pro Gly Ala Met Leu Gly Pro Ser Pro		
20	25	30
Gly Pro Ser Pro Gly Ser Ala His Ser Met Met Gly Pro Ser Pro Gly		
35	40	45

Pro	Pro	Ser	Ala	Gly	His	Pro	Ile	Pro	Thr	Gln	Gly	Pro	Gly	Gly	Tyr
50						55					60				
Pro	Gln	Asp	Asn	Met	His	Gln	Met	His	Lys	Pro	Met	Glu	Ser	Met	His
65					70				75						80
Glu	Lys	Gly	Met	Ser	Asp	Asp	Pro	Arg	Tyr	Asn	Gln	Met	Lys	Gly	Met
				85					90					95	
Gly	Met	Arg	Ser	Gly	Gly	His	Ala	Gly	Met	Gly	Pro	Pro	Pro	Ser	Pro
			100					105						110	
Met	Asp	Gln	His	Ser	Gln	Gly	Tyr	Pro	Ser	Pro	Leu	Gly	Gly	Ser	Glu
		115					120					125			
His	Ala	Ser	Ser	Pro	Val	Pro	Ala	Ser	Gly	Pro	Ser	Ser	Gly	Pro	Gln
		130					135					140			
Met	Ser	Ser	Gly	Pro	Gly	Gly	Ala	Pro	Leu	Asp	Gly	Ala	Asp	Pro	Gln
145					150					155					160
Ala	Leu	Gly	Gln	Gln	Asn	Arg	Gly	Pro	Thr	Pro	Phe	Asn	Gln	Asn	Gln
					165					170					175
Leu	His	Gln	Leu	Arg	Ala	Gln	Ile	Met	Ala	Tyr	Lys	Met	Leu	Ala	Arg
				180					185					190	
Gly	Gln	Pro	Leu	Pro	Asp	His	Leu	Gln	Met	Ala	Val	Gln	Gly	Lys	Arg
		195						200					205		
Pro	Met	Pro	Gly	Met	Gln	Gln	Gln	Met	Pro	Thr	Leu	Pro	Pro	Pro	Ser
		210					215					220			
Val	Ser	Ala	Thr	Gly	Pro	Gly	Pro	Gly	Pro	Gly	Pro	Gly	Pro	Gly	Pro
225					230					235					240
Gly	Pro	Gly	Pro	Ala	Pro	Pro	Asn	Tyr	Ser	Arg	Pro	His	Gly	Met	Gly
				245						250				255	
Gly	Pro	Asn	Met	Pro	Pro	Pro	Gly	Pro	Ser	Gly	Val	Pro	Pro	Gly	Met
			260					265						270	
Pro	Gly	Gln	Pro	Pro	Gly	Gly	Pro	Pro	Lys	Pro	Trp	Pro	Glu	Gly	Pro
		275						280					285		
Met	Ala	Asn	Ala	Ala	Ala	Pro	Thr	Ser	Thr	Pro	Gln	Lys	Leu	Ile	Pro
		290					295						300		
Pro	Gln	Pro	Thr	Gly	Arg	Pro	Ser	Pro	Ala	Pro	Pro	Ala	Val	Pro	Pro
305					310					315					320
Ala	Ala	Ser	Pro	Val	Met	Pro	Pro	Gln	Thr	Gln	Ser	Pro	Gly	Gln	Pro
				325						330				335	
Ala	Gln	Pro	Ala	Pro	Met	Val	Pro	Leu	His	Gln	Lys	Gln	Ser	Arg	Ile
			340						345					350	
Thr	Pro	Ile	Gln	Lys	Pro	Arg	Gly	Leu	Asp	Pro	Val	Glu	Ile	Leu	Gln

355	360	365
Glu Arg Glu Tyr Arg Leu Gln	Ala Arg Ile Ala His Arg	Ile Gln Glu
370	375	380
Leu Glu Asn Leu Pro Gly Ser	Leu Ala Gly Asp Leu Arg	Thr Lys Ala
385	390	395
Thr Ile Glu Leu Lys Ala Leu	Arg Leu Leu Asn Phe Gln	Arg Gln Leu
405	410	415
Arg Gln Glu Val Val Val Cys	Met Arg Arg Asp Thr Ala	Leu Glu Thr
420	425	430
Ala Leu Asn Ala Lys Ala Tyr	Lys Arg Ser Lys Arg Gln	Ser Leu Arg
435	440	445
Glu Ala Arg Ile Thr Glu Lys	Leu Glu Lys Gln Gln Lys	Ile Glu Gln
450	455	460
Glu Arg Lys Arg Arg Gln Lys	His Gln Glu Tyr Leu Asn	Ser Ile Leu
465	470	475
Gln His Ala Lys Asp Phe Lys	Glu Tyr His Arg Ser Val	Thr Gly Lys
485	490	495
Ile Gln Lys Leu Thr Lys Ala	Val Ala Thr Tyr His Ala	Asn Thr Glu
500	505	510
Arg Glu Gln Lys Lys Glu Asn	Glu Arg Ile Glu Lys Glu	Arg Met Arg
515	520	525
Arg Leu Met Ala Glu Asp Glu	Glu Gly Tyr Arg Lys Leu	Ile Asp Gln
530	535	540
Lys Lys Asp Lys Arg Leu Ala	Tyr Leu Leu Gln Gln Thr	Asp Glu Tyr
545	550	555
Val Ala Asn Leu Thr Glu Leu	Val Arg Gln His Lys Ala	Ala Gln Val
565	570	575
Ala Lys Glu Lys Lys Lys Lys	Lys Lys Lys Lys Lys Ala	Glu Asn Ala
580	585	590
Glu Gly Gln Thr Pro Ala Ile	Gly Pro Asp Gly Glu Pro	Leu Asp Glu
595	600	605
Thr Ser Gln Met Ser Asp Leu	Pro Val Lys Val Ile His	Val Glu Ser
610	615	620
Gly Lys Ile Leu Thr Gly Thr	Asp Ala Pro Lys Ala Gly	Gln Leu Glu
625	630	635
Ala Trp Leu Glu Met Asn Pro	Gly Tyr Glu Val Ala Pro	Arg Ser Asp
645	650	655
Ser Glu Glu Ser Gly Ser Glu	Glu Glu Glu Glu Glu Glu	Glu Glu Glu
660	665	670

Gln Pro Gln Ala Ala Gln Pro Pro Thr Leu Pro Val Glu Glu Lys Lys		
675	680	685
Lys Ile Pro Asp Pro Asp Ser Asp Asp Val Ser Glu Val Asp Ala Arg		
690	695	700
His Ile Ile Glu Asn Ala Lys Gln Asp Val Asp Asp Glu Tyr Gly Val		
705	710	715
Ser Gln Ala Leu Ala Arg Gly Leu Gln Ser Tyr Tyr Ala Val Ala His		
	725	730
Ala Val Thr Glu Arg Val Asp Lys Gln Ser Ala Leu Met Val Asn Gly		
	740	745
Val Leu Lys Gln Tyr Gln Ile Lys Gly Leu Glu Trp Leu Val Ser Leu		
	755	760
Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly Leu Gly		
	770	775
Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu His Lys		
785	790	795
Arg Ile Asn Gly Pro Phe Leu Ile Ile Val Pro Leu Ser Thr Leu Ser		
	805	810
Asn Trp Ala Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val Lys Val		
	820	825
Ser Tyr Lys Gly Ser Pro Ala Ala Arg Arg Ala Phe Val Pro Gln Leu		
	835	840
Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr Ile Ile		
	850	855
Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met Ile Val		
865	870	875
Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr Gln Val		
	885	890
Leu Asn Thr His Tyr Val Ala Pro Arg Arg Leu Leu Leu Thr Gly Thr		
	900	905
Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn Phe Leu		
	915	920
Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp Phe Asn		
	930	935
Ala Pro Phe Ala Met Thr Gly Glu Lys Val Asp Leu Asn Glu Glu Glu		
945	950	955
Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
	965	970
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		

980	985	990
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
995	1000	1005
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
1010	1015	1020
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		
1025	1030	1035
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1040	1045	1050
His Ile Glu Glu Ser Phe Ser Glu His Leu Gly Phe Thr Gly Gly		
1055	1060	1065
Ile Val Gln Gly Leu Asp Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1070	1075	1080
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Lys		
1085	1090	1095
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
1100	1105	1110
Asp Tyr Phe Ala Tyr Arg Gly Phe Lys Tyr Leu Arg Leu Asp Gly		
1115	1120	1125
Thr Thr Lys Ala Glu Asp Arg Gly Met Leu Leu Lys Thr Phe Asn		
1130	1135	1140
Glu Pro Gly Ser Glu Tyr Phe Ile Phe Leu Leu Ser Thr Arg Ala		
1145	1150	1155
Gly Gly Leu Gly Leu Asn Leu Gln Ser Ala Asp Thr Val Ile Ile		
1160	1165	1170
Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala Gln Asp		
1175	1180	1185
Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val Leu Arg		
1190	1195	1200
Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala Ala Ala		
1205	1210	1215
Lys Tyr Lys Leu Asn Val Asp Gln Lys Val Ile Gln Ala Gly Met		
1220	1225	1230
Phe Asp Gln Lys Ser Ser Ser His Glu Arg Arg Ala Phe Leu Gln		
1235	1240	1245
Ala Ile Leu Glu His Glu Glu Gln Asp Glu Glu Glu Asp Glu Val		
1250	1255	1260
Pro Asp Asp Glu Thr Val Asn Gln Met Ile Ala Arg His Glu Glu		
1265	1270	1275

Glu Phe Asp Leu Phe Met Arg Met Asp Leu Asp Arg Arg Arg Glu		
1280	1285	1290
Glu Ala Arg Asn Pro Lys Arg Lys Pro Arg Leu Met Glu Glu Asp		
1295	1300	1305
Glu Leu Pro Ser Trp Ile Ile Lys Asp Asp Ala Glu Val Glu Arg		
1310	1315	1320
Leu Thr Cys Glu Glu Glu Glu Glu Lys Met Phe Gly Arg Gly Ser		
1325	1330	1335
Arg His Arg Lys Glu Val Asp Tyr Ser Asp Ser Leu Thr Glu Lys		
1340	1345	1350
Gln Trp Leu Lys Ala Ile Glu Glu Gly Thr Leu Glu Glu Ile Glu		
1355	1360	1365
Glu Glu Val Arg Gln Lys Lys Ser Ser Arg Lys Arg Lys Arg Asp		
1370	1375	1380
Ser Asp Ala Gly Ser Ser Thr Pro Thr Thr Ser Thr Arg Ser Arg		
1385	1390	1395
Asp Lys Asp Asp Glu Ser Lys Lys Gln Lys Lys Arg Gly Arg Pro		
1400	1405	1410
Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro Asn Leu Thr Lys Lys		
1415	1420	1425
Met Lys Lys Ile Val Asp Ala Val Ile Lys Tyr Lys Asp Ser Ser		
1430	1435	1440
Ser Gly Arg Gln Leu Ser Glu Val Phe Ile Gln Leu Pro Ser Arg		
1445	1450	1455
Lys Glu Leu Pro Glu Tyr Tyr Glu Leu Ile Arg Lys Pro Val Asp		
1460	1465	1470
Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn His Lys Tyr Arg Ser		
1475	1480	1485
Leu Asn Asp Leu Glu Lys Asp Val Met Leu Leu Cys Gln Asn Ala		
1490	1495	1500
Gln Thr Phe Asn Leu Glu Gly Ser Leu Ile Tyr Glu Asp Ser Ile		
1505	1510	1515
Val Leu Gln Ser Val Phe Thr Ser Val Arg Gln Lys Ile Glu Lys		
1520	1525	1530
Glu Asp Asp Ser Glu Gly Glu Glu Ser Glu Glu Glu Glu Glu Gly		
1535	1540	1545
Glu Glu Glu Gly Ser Glu Ser Glu Ser Arg Ser Val Lys Val Lys		
1550	1555	1560
Ile Lys Leu Gly Arg Lys Glu Lys Ala Gln Asp Arg Leu Lys Gly		

1565	1570	1575
Gly Arg Arg Arg Pro Ser	Arg Gly Ser Arg Ala	Lys Pro Val Val
1580	1585	1590
Ser Asp Asp Asp Ser Glu	Glu Glu Gln Glu Glu	Asp Arg Ser Gly
1595	1600	1605
Ser Gly Ser Glu Glu Asp		
1610		
<210> 20		
<211> 1613		
<212> PRT		
<213> 智人		
<400> 20		
Met Ser Thr Pro Asp Pro Pro Leu Gly Gly Thr Pro Arg Pro Gly Pro		
1	5	10
Ser Pro Gly Pro Gly Pro Ser Pro Gly Ala Met Leu Gly Pro Ser Pro		
20	25	30
Gly Pro Ser Pro Gly Ser Ala His Ser Met Met Gly Pro Ser Pro Gly		
35	40	45
Pro Pro Ser Ala Gly His Pro Ile Pro Thr Gln Gly Pro Gly Gly Tyr		
50	55	60
Pro Gln Asp Asn Met His Gln Met His Lys Pro Met Glu Ser Met His		
65	70	75
Glu Lys Gly Met Ser Asp Asp Pro Arg Tyr Asn Gln Met Lys Gly Met		
85	90	95
Gly Met Arg Ser Gly Gly His Ala Gly Met Gly Pro Pro Pro Ser Pro		
100	105	110
Met Asp Gln His Ser Gln Gly Tyr Pro Ser Pro Leu Gly Gly Ser Glu		
115	120	125
His Ala Ser Ser Pro Val Pro Ala Ser Gly Pro Ser Ser Gly Pro Gln		
130	135	140
Met Ser Ser Gly Pro Gly Gly Ala Pro Leu Asp Gly Ala Asp Pro Gln		
145	150	155
Ala Leu Gly Gln Gln Asn Arg Gly Pro Thr Pro Phe Asn Gln Asn Gln		
165	170	175
Leu His Gln Leu Arg Ala Gln Ile Met Ala Tyr Lys Met Leu Ala Arg		
180	185	190
Gly Gln Pro Leu Pro Asp His Leu Gln Met Ala Val Gln Gly Lys Arg		
195	200	205
Pro Met Pro Gly Met Gln Gln Gln Met Pro Thr Leu Pro Pro Pro Ser		

210	215	220
Val Ser Ala Thr Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro		
225	230	235
Gly Pro Gly Pro Ala Pro Pro Asn Tyr Ser Arg Pro His Gly Met Gly		240
245	250	255
Gly Pro Asn Met Pro Pro Pro Gly Pro Ser Gly Val Pro Pro Gly Met		
260	265	270
Pro Gly Gln Pro Pro Gly Gly Pro Pro Lys Pro Trp Pro Glu Gly Pro		
275	280	285
Met Ala Asn Ala Ala Ala Pro Thr Ser Thr Pro Gln Lys Leu Ile Pro		
290	295	300
Pro Gln Pro Thr Gly Arg Pro Ser Pro Ala Pro Pro Ala Val Pro Pro		
305	310	315
Ala Ala Ser Pro Val Met Pro Pro Gln Thr Gln Ser Pro Gly Gln Pro		320
325	330	335
Ala Gln Pro Ala Pro Met Val Pro Leu His Gln Lys Gln Ser Arg Ile		
340	345	350
Thr Pro Ile Gln Lys Pro Arg Gly Leu Asp Pro Val Glu Ile Leu Gln		
355	360	365
Glu Arg Glu Tyr Arg Leu Gln Ala Arg Ile Ala His Arg Ile Gln Glu		
370	375	380
Leu Glu Asn Leu Pro Gly Ser Leu Ala Gly Asp Leu Arg Thr Lys Ala		
385	390	395
Thr Ile Glu Leu Lys Ala Leu Arg Leu Leu Asn Phe Gln Arg Gln Leu		
405	410	415
Arg Gln Glu Val Val Val Cys Met Arg Arg Asp Thr Ala Leu Glu Thr		
420	425	430
Ala Leu Asn Ala Lys Ala Tyr Lys Arg Ser Lys Arg Gln Ser Leu Arg		
435	440	445
Glu Ala Arg Ile Thr Glu Lys Leu Glu Lys Gln Gln Lys Ile Glu Gln		
450	455	460
Glu Arg Lys Arg Arg Gln Lys His Gln Glu Tyr Leu Asn Ser Ile Leu		
465	470	475
Gln His Ala Lys Asp Phe Lys Glu Tyr His Arg Ser Val Thr Gly Lys		480
485	490	495
Ile Gln Lys Leu Thr Lys Ala Val Ala Thr Tyr His Ala Asn Thr Glu		
500	505	510
Arg Glu Gln Lys Lys Glu Asn Glu Arg Ile Glu Lys Glu Arg Met Arg		
515	520	525

Arg Leu Met Ala Glu Asp Glu Glu Gly Tyr Arg Lys Leu Ile Asp Gln		
530	535	540
Lys Lys Asp Lys Arg Leu Ala Tyr Leu Leu Gln Gln Thr Asp Glu Tyr		
545	550	555
Val Ala Asn Leu Thr Glu Leu Val Arg Gln His Lys Ala Ala Gln Val		
	565	570
Ala Lys Glu Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Ala Glu Asn Ala		
	580	585
Glu Gly Gln Thr Pro Ala Ile Gly Pro Asp Gly Glu Pro Leu Asp Glu		
	595	600
Thr Ser Gln Met Ser Asp Leu Pro Val Lys Val Ile His Val Glu Ser		
	610	615
Gly Lys Ile Leu Thr Gly Thr Asp Ala Pro Lys Ala Gly Gln Leu Glu		
625	630	635
Ala Trp Leu Glu Met Asn Pro Gly Tyr Glu Val Ala Pro Arg Ser Asp		
	645	650
Ser Glu Glu Ser Gly Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu		
	660	665
Gln Pro Gln Ala Ala Gln Pro Pro Thr Leu Pro Val Glu Glu Lys Lys		
	675	680
Lys Ile Pro Asp Pro Asp Ser Asp Asp Val Ser Glu Val Asp Ala Arg		
690	695	700
His Ile Ile Glu Asn Ala Lys Gln Asp Val Asp Asp Glu Tyr Gly Val		
705	710	715
Ser Gln Ala Leu Ala Arg Gly Leu Gln Ser Tyr Tyr Ala Val Ala His		
	725	730
Ala Val Thr Glu Arg Val Asp Lys Gln Ser Ala Leu Met Val Asn Gly		
	740	745
Val Leu Lys Gln Tyr Gln Ile Lys Gly Leu Glu Trp Leu Val Ser Leu		
	755	760
Tyr Asn Asn Asn Leu Asn Gly Ile Leu Ala Asp Glu Met Gly Leu Gly		
	770	775
Lys Thr Ile Gln Thr Ile Ala Leu Ile Thr Tyr Leu Met Glu His Lys		
785	790	795
Arg Ile Asn Gly Pro Phe Leu Ile Ile Val Pro Leu Ser Thr Leu Ser		
	805	810
Asn Trp Ala Tyr Glu Phe Asp Lys Trp Ala Pro Ser Val Val Lys Val		
	820	825
Ser Tyr Lys Gly Ser Pro Ala Ala Arg Arg Ala Phe Val Pro Gln Leu		

835	840	845
Arg Ser Gly Lys Phe Asn Val Leu Leu Thr Thr Tyr Glu Tyr Ile Ile		
850	855	860
Lys Asp Lys His Ile Leu Ala Lys Ile Arg Trp Lys Tyr Met Ile Val		
865	870	875
Asp Glu Gly His Arg Met Lys Asn His His Cys Lys Leu Thr Gln Val		
885	890	895
Leu Asn Thr His Tyr Val Ala Pro Arg Arg Leu Leu Leu Thr Gly Thr		
900	905	910
Pro Leu Gln Asn Lys Leu Pro Glu Leu Trp Ala Leu Leu Asn Phe Leu		
915	920	925
Leu Pro Thr Ile Phe Lys Ser Cys Ser Thr Phe Glu Gln Trp Phe Asn		
930	935	940
Ala Pro Phe Ala Met Thr Gly Glu Lys Val Asp Leu Asn Glu Glu Glu		
945	950	955
Thr Ile Leu Ile Ile Arg Arg Leu His Lys Val Leu Arg Pro Phe Leu		
965	970	975
Leu Arg Arg Leu Lys Lys Glu Val Glu Ala Gln Leu Pro Glu Lys Val		
980	985	990
Glu Tyr Val Ile Lys Cys Asp Met Ser Ala Leu Gln Arg Val Leu Tyr		
995	1000	1005
Arg His Met Gln Ala Lys Gly Val Leu Leu Thr Asp Gly Ser Glu		
1010	1015	1020
Lys Asp Lys Lys Gly Lys Gly Gly Thr Lys Thr Leu Met Asn Thr		
1025	1030	1035
Ile Met Gln Leu Arg Lys Ile Cys Asn His Pro Tyr Met Phe Gln		
1040	1045	1050
His Ile Glu Glu Ser Phe Ser Glu His Leu Gly Phe Thr Gly Gly		
1055	1060	1065
Ile Val Gln Gly Leu Asp Leu Tyr Arg Ala Ser Gly Lys Phe Glu		
1070	1075	1080
Leu Leu Asp Arg Ile Leu Pro Lys Leu Arg Ala Thr Asn His Lys		
1085	1090	1095
Val Leu Leu Phe Cys Gln Met Thr Ser Leu Met Thr Ile Met Glu		
1100	1105	1110
Asp Tyr Phe Ala Tyr Arg Gly Phe Lys Tyr Leu Arg Leu Asp Gly		
1115	1120	1125
Thr Thr Lys Ala Glu Asp Arg Gly Met Leu Leu Lys Thr Phe Asn		
1130	1135	1140

Glu Pro Gly Ser Glu Tyr Phe Ile Phe Leu Leu Ser Thr Arg Ala		
1145	1150	1155
Gly Gly Leu Gly Leu Asn Leu Gln Ser Ala Asp Thr Val Ile Ile		
1160	1165	1170
Phe Asp Ser Asp Trp Asn Pro His Gln Asp Leu Gln Ala Gln Asp		
1175	1180	1185
Arg Ala His Arg Ile Gly Gln Gln Asn Glu Val Arg Val Leu Arg		
1190	1195	1200
Leu Cys Thr Val Asn Ser Val Glu Glu Lys Ile Leu Ala Ala Ala		
1205	1210	1215
Lys Tyr Lys Leu Asn Val Asp Gln Lys Val Ile Gln Ala Gly Met		
1220	1225	1230
Phe Asp Gln Lys Ser Ser Ser His Glu Arg Arg Ala Phe Leu Gln		
1235	1240	1245
Ala Ile Leu Glu His Glu Glu Gln Asp Glu Glu Glu Asp Glu Val		
1250	1255	1260
Pro Asp Asp Glu Thr Val Asn Gln Met Ile Ala Arg His Glu Glu		
1265	1270	1275
Glu Phe Asp Leu Phe Met Arg Met Asp Leu Asp Arg Arg Arg Glu		
1280	1285	1290
Glu Ala Arg Asn Pro Lys Arg Lys Pro Arg Leu Met Glu Glu Asp		
1295	1300	1305
Glu Leu Pro Ser Trp Ile Ile Lys Asp Asp Ala Glu Val Glu Arg		
1310	1315	1320
Leu Thr Cys Glu Glu Glu Glu Glu Lys Met Phe Gly Arg Gly Ser		
1325	1330	1335
Arg His Arg Lys Glu Val Asp Tyr Ser Asp Ser Leu Thr Glu Lys		
1340	1345	1350
Gln Trp Leu Lys Ala Ile Glu Glu Gly Thr Leu Glu Glu Ile Glu		
1355	1360	1365
Glu Glu Val Arg Gln Lys Lys Ser Ser Arg Lys Arg Lys Arg Asp		
1370	1375	1380
Ser Asp Ala Gly Ser Ser Thr Pro Thr Thr Ser Thr Arg Ser Arg		
1385	1390	1395
Asp Lys Asp Asp Glu Ser Lys Lys Gln Lys Lys Arg Gly Arg Pro		
1400	1405	1410
Pro Ala Glu Lys Leu Ser Pro Asn Pro Pro Asn Leu Thr Lys Lys		
1415	1420	1425
Met Lys Lys Ile Val Asp Ala Val Ile Lys Tyr Lys Asp Ser Ser		

1430	1435	1440
Gly Arg Gln Leu Ser Glu	Val Phe Ile Gln Leu	Pro Ser Arg Lys
1445	1450	1455
Glu Leu Pro Glu Tyr Tyr	Glu Leu Ile Arg Lys	Pro Val Asp Phe
1460	1465	1470
Lys Lys Ile Lys Glu Arg	Ile Arg Asn His Lys	Tyr Arg Ser Leu
1475	1480	1485
Asn Asp Leu Glu Lys Asp	Val Met Leu Leu Cys	Gln Asn Ala Gln
1490	1495	1500
Thr Phe Asn Leu Glu Gly	Ser Leu Ile Tyr Glu	Asp Ser Ile Val
1505	1510	1515
Leu Gln Ser Val Phe Thr	Ser Val Arg Gln Lys	Ile Glu Lys Glu
1520	1525	1530
Asp Asp Ser Glu Gly Glu	Glu Ser Glu Glu Glu	Glu Glu Gly Glu
1535	1540	1545
Glu Glu Gly Ser Glu Ser	Glu Ser Arg Ser Val	Lys Val Lys Ile
1550	1555	1560
Lys Leu Gly Arg Lys Glu	Lys Ala Gln Asp Arg	Leu Lys Gly Gly
1565	1570	1575
Arg Arg Arg Pro Ser Arg	Gly Ser Arg Ala Lys	Pro Val Val Ser
1580	1585	1590
Asp Asp Asp Ser Glu Glu	Glu Gln Glu Glu Asp	Arg Ser Gly Ser
1595	1600	1605
Gly Ser Glu Glu Asp		
1610		

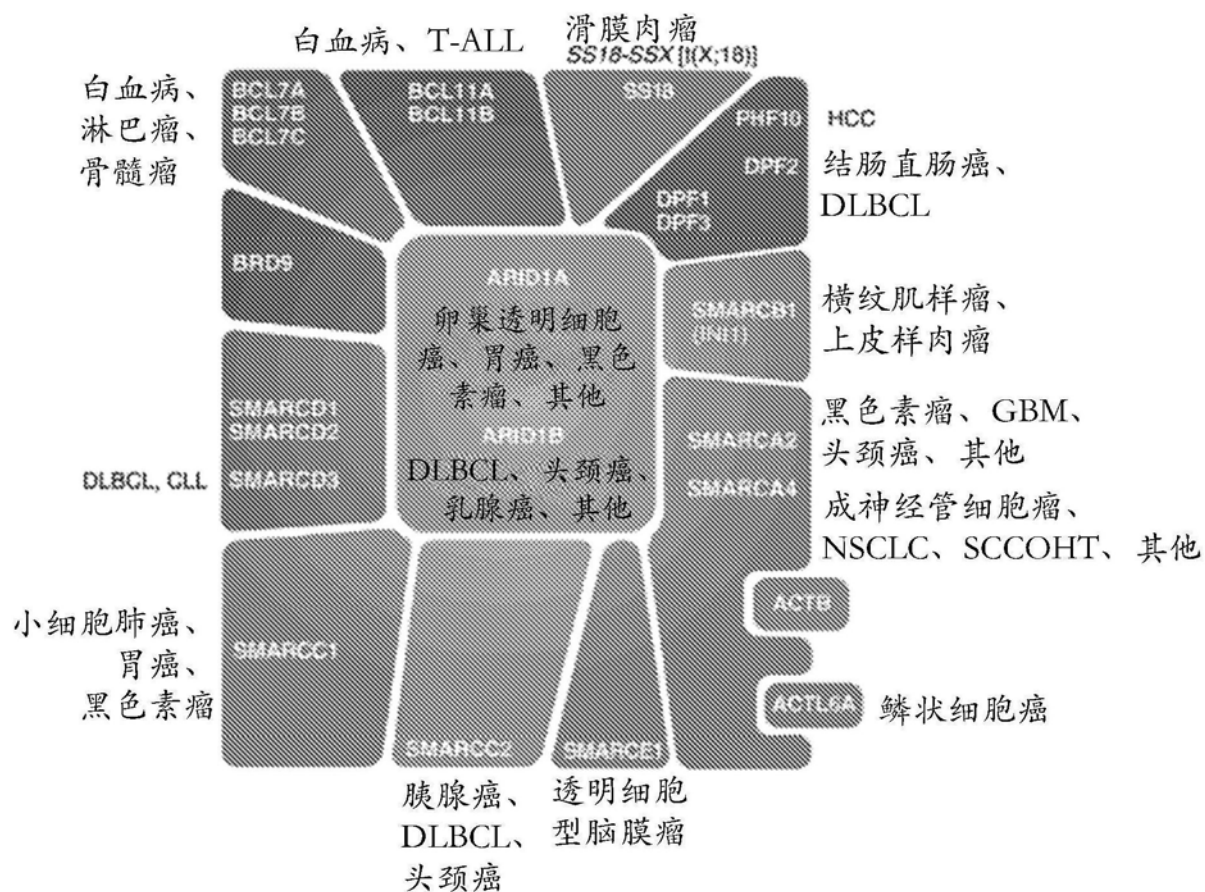


图1

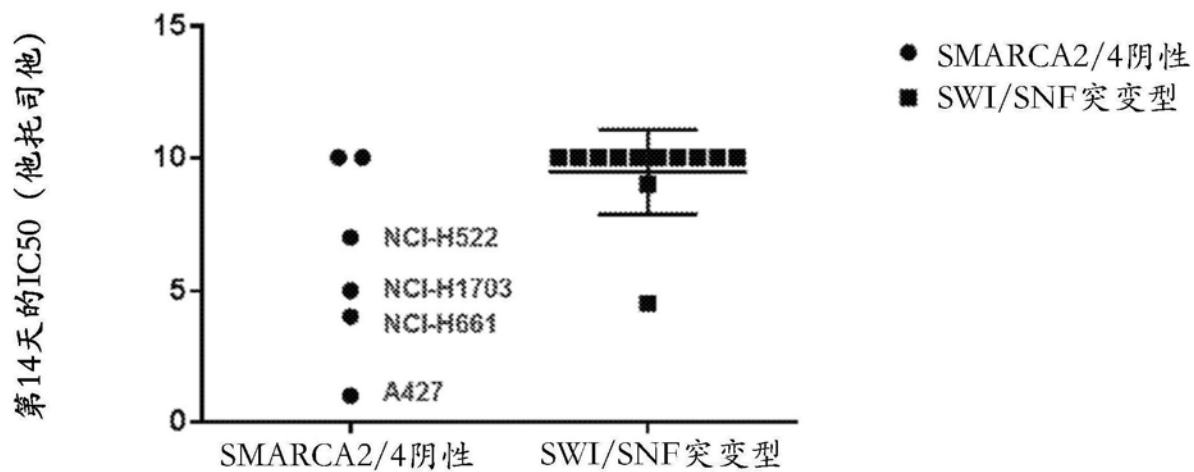


图2

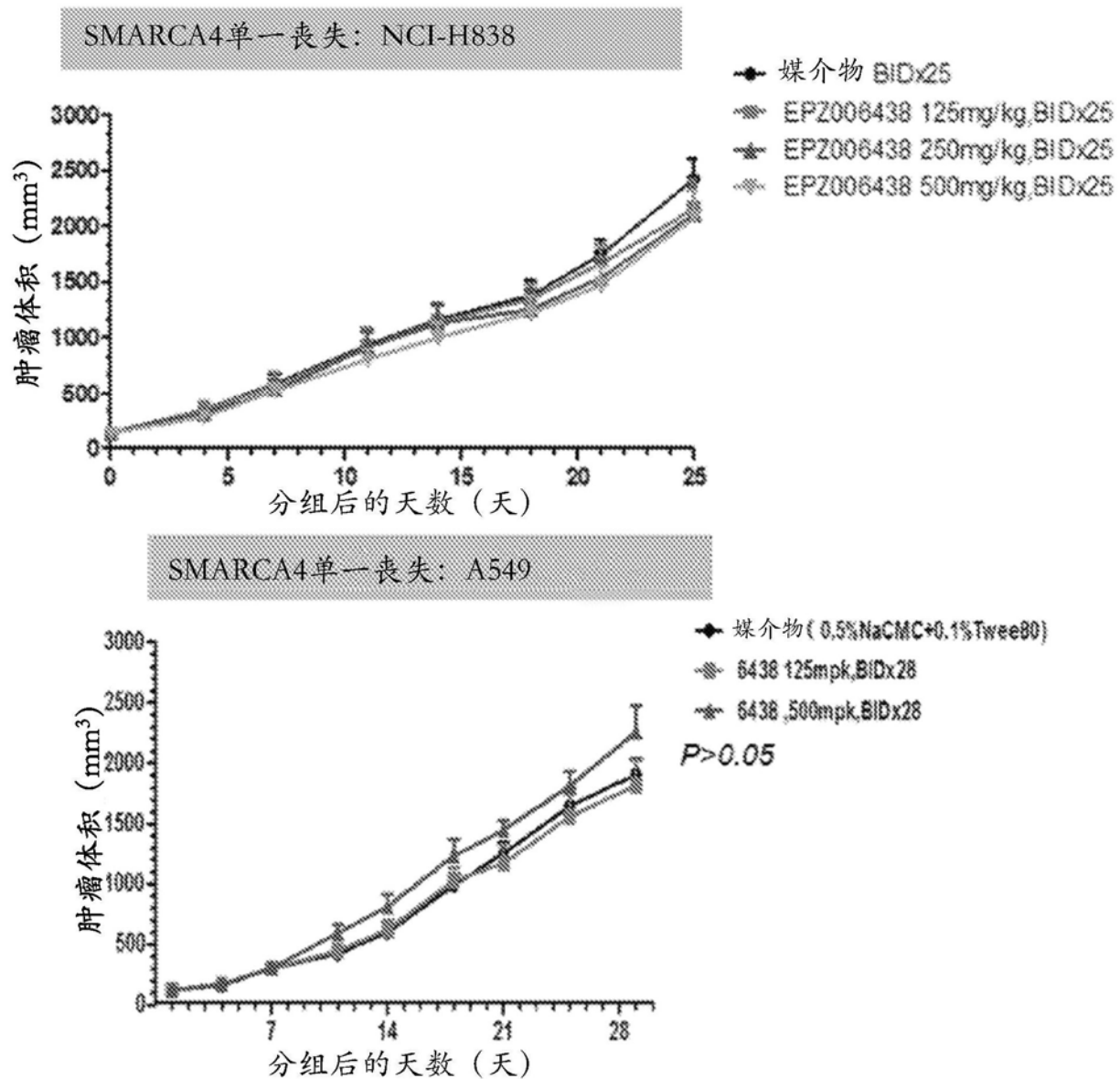


图3

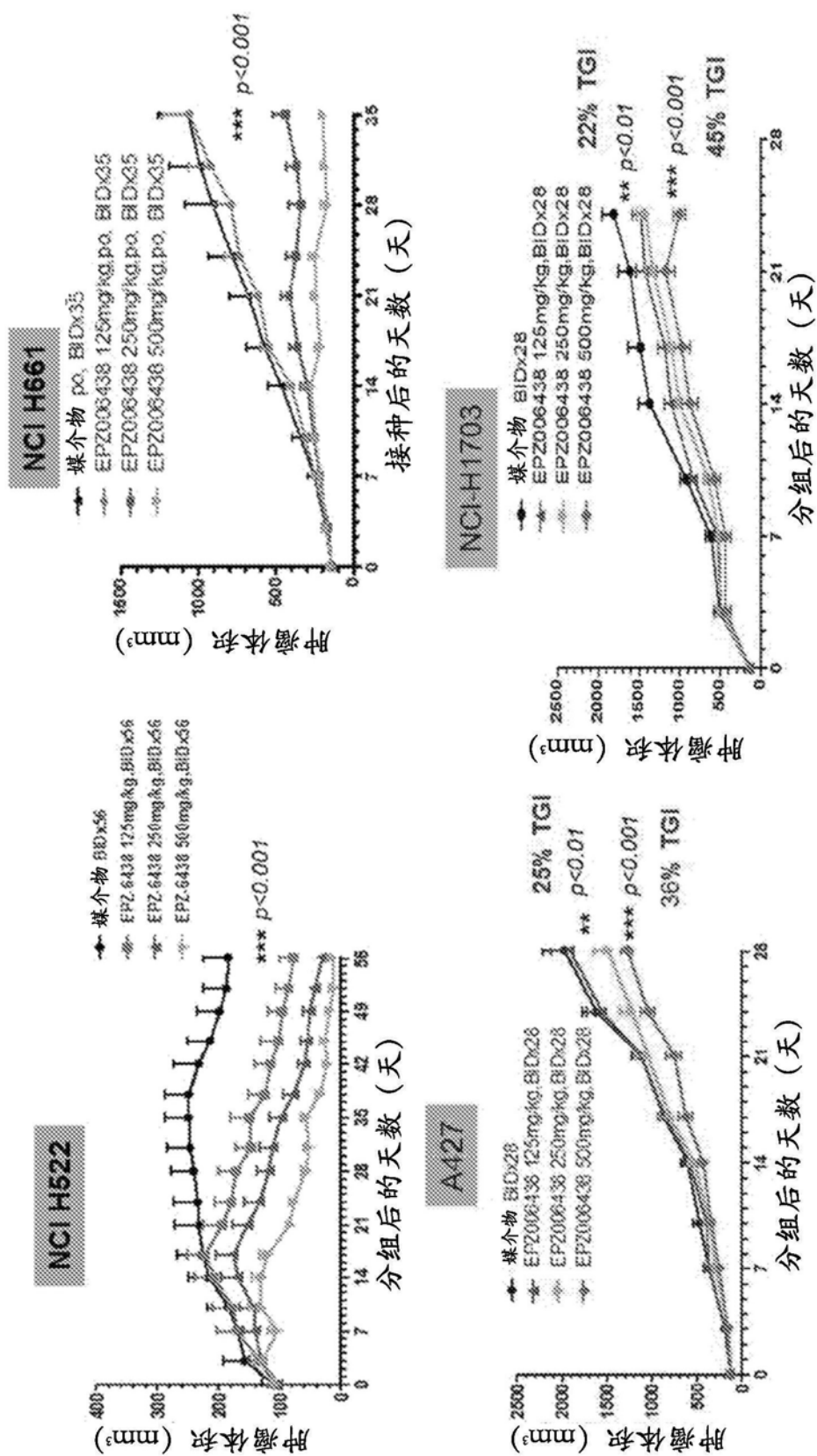


图4