



(12)发明专利申请

(10)申请公布号 CN 110691792 A

(43)申请公布日 2020.01.14

(21)申请号 201880016199.3

T·G·约翰斯顿

(22)申请日 2018.01.10

小罗纳德·J·豪斯

(30)优先权数据

62/444,802 2017.01.10 US

62/551,752 2017.08.29 US

62/596,662 2017.12.08 US

(74)专利代理机构 北京坤瑞律师事务所 11494

代理人 封新琴

(85)PCT国际申请进入国家阶段日

2019.09.05

(51)Int.Cl.

C07K 16/28(2006.01)

C12Q 1/6869(2006.01)

A61P 35/00(2006.01)

A61K 31/454(2006.01)

A61K 39/00(2006.01)

(86)PCT国际申请的申请数据

PCT/US2018/013227 2018.01.10

(87)PCT国际申请的公布数据

W02018/132518 EN 2018.07.19

(71)申请人 朱诺治疗学股份有限公司

地址 美国华盛顿州

权利要求书9页 说明书125页

序列表24页 附图47页

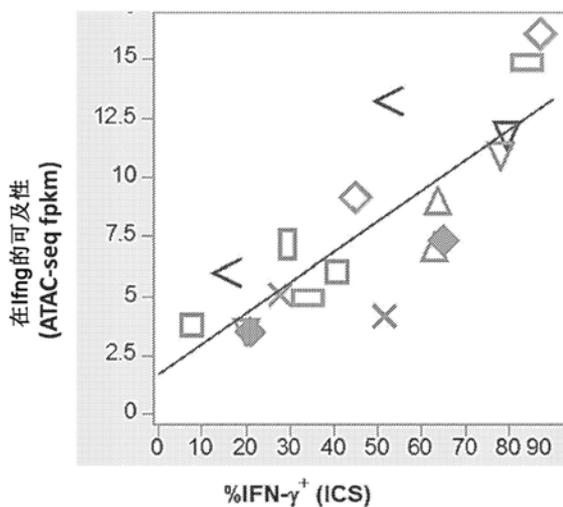
(72)发明人 M·L·博尼哈德 D·G·库格勒

(54)发明名称

细胞疗法的表观遗传学分析及相关方法

(57)摘要

本文中提供了鉴定基因组区域的方法,所述基因组区域预测细胞疗法的治疗结果和/或细胞的表型或功能。在一些实施方案中,所述方法包括细胞(例如用于细胞疗法的工程化细胞)的表观遗传学和/或表观基因组学分析,其与制备用于细胞疗法的工程化细胞和/或预测对细胞疗法的反应的方法有关。在一些实施方案中,所述方法包括评估、表征和分析一个或多个基因区域的表观遗传学特性的变化或修饰的步骤,所述特性比如染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和/或DNA甲基化。在一些实施方案中,表观遗传学和/或表观基因组学分析包括确定细胞例如用于细胞疗法的工程化细胞的表观遗传学特性。



1. 一种鉴定与细胞疗法的治疗结果相关的一个或多个基因组区域的方法,所述方法包括:

(a) 分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞或群包含在下列组合物中:(i) 第一细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二组合物,或(ii) 包含所述重组受体的第二细胞组合物;和

(b) 鉴定所述一个或多个基因组区域中的一个或多个,其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关连细胞疗法的结果,所述细胞疗法包括对一位受试者或一组受试者施用包含所述重组受体的第二细胞组合物。

2. 权利要求1的方法,其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

3. 权利要求2的方法,其中所述结果是反应,并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

4. 权利要求2的方法,其中所述结果是毒性,并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

5. 权利要求2或权利要求4的方法,其中毒性是剂量限制性毒性(DLT)。

6. 权利要求1-5中任一项的方法,其中:

所述第一组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞;并且/或者

所述第二组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞。

7. 一种确定细胞组合物的一个或多个特性或特征的方法,所述方法包括分析或确定T细胞组合物的一个或多个基因组区域的表观遗传学特性,所述T细胞组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞。

8. 权利要求7的方法,其中细胞组合物是(i) 第一T细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二T细胞组合物,或(ii) 包含所述重组受体的细胞的第二T细胞组合物。

9. 权利要求6-8中任一项的方法,其中:

所述细胞组合物包含大于或大于约70%、大于或大于约75%、大于或大于约80%、大于或大于约85%、大于或大于约90%、大于或大于约95%或大于或大于约98%的CD4+和/或CD8+原代人T细胞;并且/或者

所述细胞组合物基本上由CD4+和/或CD8+原代人T细胞组成。

10. 权利要求7或权利要求8的方法,进一步包括将一个或多个基因组区域中的每一个的表观遗传学特性单独地与来自不同的细胞组合物的细胞的相应表观遗传学特性和/或与参考谱进行比较,所述参考谱任选地为已知指示或关联细胞组合物的属性或特征的参考谱。

11. 权利要求10的方法,其中所述比较指示或关联细胞组合物内的细胞的状态、表型或功能,任选地为活化、效应子或记忆状态;细胞组合物内的细胞的一致性或均匀性;当施用于一位受试者或一组受试者时,细胞组合物是否或是否可能表现出或产生结果;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;和/或细胞组合物中的工程化细胞的比例或频率。

12. 一种用于确定或鉴定与细胞组合物的属性或特征相关的表观遗传学特性的方法，所述方法包括：

(a) 确定或测量在第一细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度；

(b) 确定或测量在第二细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度；和

(c) 比较(a)中的水平或程度和(b)中的水平或程度，其中一个或多个基因组区域的表观遗传学特性的水平或程度的差异(任选地显著差异)鉴定或确定表观遗传学特性的存在，所述表观遗传学特性指示或关联在第一和第二组合物之一的细胞中存在但在另一者中不存在的属性或特征。

13. 权利要求12的方法，其中：

所述第一组合物和第二组合物之一包含有待用重组受体基因工程改造的细胞，并且所述第一组合物和第二组合物中的另一者包含经工程改造以表达重组受体的细胞；

所述第一组合物和第二组合物包含来自不同供体的原代细胞，所述供体任选基于疾病状态、疾病严重性或疾病类型而不同的供体；

所述第一组合物和第二组合物包含处于工程化细胞的制造过程的不同阶段或步骤的细胞；

所述第一组合物和第二组合物之一包含与药剂接触以调节细胞的活性、表型或功能的细胞，并且所述第一和第二组合物中的另一者包含并不如此接触的相似细胞；或者

所述第一组合物和第二组合物之一包含细胞组合物样品，所述细胞组合物样品与所述第一和第二组合物之一施用于受试者后发生或已经发生但另一者则不然的结果相关。

14. 权利要求13的方法，其中所述药剂是多肽或蛋白质、肽、抗体、核酸、病毒载体或病毒制剂或小分子化合物。

15. 权利要求13或权利要求14的方法，其中所述药剂是刺激性试剂，任选抗CD3/抗CD28；免疫调节剂、对CAR特异的抗独特型抗体或其抗原结合片段、免疫检查点抑制剂、代谢途径调节剂、腺苷受体拮抗剂、激酶抑制剂、抗TGFβ抗体或抗TGFβR抗体或细胞因子。

16. 权利要求12-15中任一项的方法，其中第一组合物的属性或特征指示状态、表型或功能，任选地活化、效应子或记忆状态、表型或功能；外源核酸整合的位置、丰度或频率；细胞组合物内的细胞的克隆性；细胞组合物中的工程化细胞的比例或频率；和/或当施用于一位受试者或一组受试者组时，细胞组合物是否或是否可能表现出或产生结果。

17. 权利要求12或权利要求16的方法，其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

18. 权利要求17的方法，其中所述结果是反应，并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

19. 权利要求18的方法，其中所述结果是毒性，并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

20. 权利要求18或权利要求19的方法，其中毒性是剂量限制性毒性(DLT)。

21. 权利要求12-20中任一项的方法，其被重复多次。

22. 权利要求21的方法,其中鉴定或确定存在于所述第一和第二组合物之一的大多数中但不存在于另一者中的表观遗传学特性。

23. 一种评估细胞组合物的属性或特征的方法,其包括:

(a) 分析在细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞组合物包含用重组受体工程改造的细胞和/或有待用重组受体基因工程改造的细胞;和

(b) 将一个或多个基因组区域的表观遗传学特性单独地与参考谱比较,其中所述比较指示细胞组合物是否或是否可能表现出所述属性或特征。

24. 权利要求23的方法,其中所述属性或特征指示状态、表型或功能,任选地活化、效应子或记忆状态、表型或功能;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;细胞组合物中的工程化细胞的比例或频率;和/或当施用于一位受试者或一组受试者组时,细胞组合物是否或是否可能表现出或产生结果。

25. 权利要求23或权利要求24的方法,其中所述属性或特征是当施用于一位受试者或一组受试者时细胞组合物是否或是否可能表现出或产生结果,并且所述方法是为了评估将施用于受试者的细胞组合物。

26. 权利要求24或权利要求25的方法,其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

27. 权利要求26的方法,其中所述结果是反应,并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

28. 权利要求27的方法,其中所述结果是毒性,并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

29. 权利要求27或权利要求28的方法,其中毒性是剂量限制性毒性(DLT)。

30. 权利要求23-29中任一项的方法,其中如果所述比较指示细胞组合物将要或可能将要表现出期望的结果,则将细胞组合物施用于受试者。

31. 权利要求23-30中任一项的方法,其中如果所述比较指示细胞组合物不会或可能不会表现出期望的结果,则任选下列之一:

- (i) 施用其中细胞组合物被改变的细胞组合物;
- (ii) 施用其中细胞剂量被改变的细胞组合物;
- (iii) 施用其中施用于受试者的细胞的剂量方案被改变的细胞组合物;
- (iv) 将细胞组合物与一种或多种其他治疗剂联合施用;或
- (v) 不对受试者施用所述细胞组合物。

32. 权利要求30或权利要求31的方法,其中所述期望的结果是完全缓解、部分缓解或持续缓解和/或是3级或更低等级的神经毒性或1级或2级神经毒性,是3级或更低等级的CRS,或是1级或2级CRS,或者不包括任何等级的神经毒性或任何等级的CRS。

33. 权利要求31或权利要求32的方法,其中通过在工程改造细胞组合物中的细胞的一个或多个步骤中改变一种或多种作用剂或条件来改变细胞组合物。

34. 权利要求33的方法,其中所述一种或多种作用剂或条件选自血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;细胞组合物的

来源或细胞类型;输入组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌注;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的细胞组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。

35. 权利要求33或权利要求34的方法,其中在施用改变的细胞组合物之前,对包含在改变的细胞组合物中的细胞重复步骤(a)和(b)。

36. 权利要求35的方法,其中改变细胞的给药方案包括在对受试者施用第一剂量的细胞之后对受试者施用第二剂量的细胞。

37. 权利要求36的方法,其中在施用第一剂量的细胞之后至少1周、2周、3周、4周、2个月、3个月、4个月、5个月、6个月、9个月或12个月施用后续剂量的细胞。

38. 权利要求23-37中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

39. 权利要求38的方法,其中所述阈值:

是与当施用于具有相同或相似疾病或病症的受试者时显示表现出期望的结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平;

是与来自已经单独施用于一组受试者的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的平均、中位或平均的值或水平,或者在所述平均、中位或平均的值或水平的标准偏差之内,其中该组受试者的每一位在施用后继续表现出期望的结果;或者

是与来自正常或健康受试者的相似细胞组合物中的表观遗传学特性相关或指示所述表观遗传学特性的值或水平。

40. 权利要求23-37中任一项的方法,其中:

所述比较包括差异性可及性分析;并且/或者

所述参考谱包括参考表观遗传学图谱,其包含一个或多个基因组区域内的序列读段的峰。

41. 权利要求40的方法,其中:

所述参考表观遗传学图谱是根据细胞组合物的可及性分析(任选地染色质可及性)确定的,所述细胞组合物在将所述细胞或细胞组合物施用于具有相同或相似疾病或病症的受试者后显示表现出期望的结果;

所述参考表观遗传学图谱是根据来自已经单独施用于一组受试者的多个细胞组合物中的可及性分析(任选地染色质可及性)的序列读段的共同峰确定的,其中该组受试者的每一位在施用后继续表现出期望的结果;或者

所述参考表观遗传学图谱是根据来自正常或健康受试者的相似细胞组合物的可及性分析(任选地染色质可及性)确定的。

42. 一种评估细胞组合物方法,其包括:

(a) 分析包含在输出细胞组合物中的细胞和/或包含在输入组合物中的细胞的一个或多个基因组区域的表观遗传学特性,所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生;和

(b) 将一个或多个基因组区域的表观遗传学特性单独地与参考谱比较,其中所述比较指示所述细胞是否或是否可能表现出预定的特征或属性。

43. 权利要求42的方法,其中预定的特征或属性是组合物内的细胞的状态、表型或功能;细胞组合物内的细胞的一致性 or 均匀性;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;和/或细胞组合物中的工程化细胞的比例或频率。

44. 权利要求43的方法,其中预定的表型或属性是指示细胞的效应子功能或活化状态的状态、表型或功能并且/或者指示细胞表现出幼稚表型或长寿记忆表型。

45. 权利要求43或权利要求44的方法,其中一种或多种测试剂或条件包括血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;输入组合物的来源或细胞类型;输入组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌流;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的输入组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。

46. 权利要求43-45中任一项的方法,其中一种或多种测试剂或条件包括来自测试化合物文库的一种或多种化合物。

47. 权利要求42-46中任一项的方法,其包括如果所述比较指示细胞组合物具有或可能具有期望的特征或属性,则选择用于培养细胞的一种或多种测试剂或条件和/或选择将施用于受试者的细胞组合物。

48. 权利要求42-46中任一项的方法,其包括如果所述比较指示细胞组合物不具有或可能不具有期望的特征或属性,则用一种或多种另外的测试剂或条件重复步骤(a)和(b)。

49. 权利要求42-48中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

50. 权利要求49的方法,其中所述阈值:

是与已知表现出期望的属性或特征的细胞组合物的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平;或者

是与来自已知表现出期望的属性或特征的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的平均、中位或平均的值或水平,或者在所述平均、中位或平均的值或水平的标准偏差之内。

51. 权利要求42-48中任一项的方法,其中:

所述比较包括差异性可及性分析;并且/或者

所述参考谱包括参考表观遗传学图谱,其包含一个或多个基因组区域内的序列读段的峰。

52. 权利要求51的方法,其中:

所述参考表观遗传学图谱是根据已知表现出期望的属性或特征的细胞组合物的可及性分析(任选地染色质可及性)确定的;

所述参考表观遗传学图谱是根据来自已知表现出期望的属性或特征的多个细胞组合物的可及性分析(任选地染色质可及性)的序列读段的共有峰确定的。

53. 权利要求42-52中任一项的方法,其中期望的结果或特征是指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型或功能。

54. 权利要求1-53中任一项的方法,其中所述基因组区域包含基因组基因座或基因。

55. 权利要求1-54中任一项的方法,其中基因组区域包含基因的编码区、开放阅读框、非编码区、基因间区或调控元件。

56. 权利要求1-55中任一项的方法,其中所述基因组区域包含基因的开放阅读框。

57. 权利要求1-55中任一项的方法,其中所述基因组区域包含基因间区或调控元件。

58. 权利要求1-55和57中任一项的方法,其中基因组区域包含内含子、外显子、顺式调控元件、启动子、增强子、上游激活序列(UAS)、3'非翻译区(UTR)、5'UTR、非编码RNA产生区、非编码RNA(ncRNA)基因、miRNA基因、siRNA基因、piRNA基因、snoRNA基因、lncRNA基因、核糖体RNA(rRNA)基因、小RNA结合位点、非编码RNA结合位点、假基因、转录终止位点(TTS)、重复序列、端粒区、可及的染色质区、不可及的染色质区、开放染色质区和/或异染色质区。

59. 权利要求1-58中任一项的方法,其中表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。

60. 权利要求1-59中任一项的方法,其中表观遗传学特性是染色质可及性。

61. 权利要求1-60中任一项的方法,其中:

所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者

所述表观遗传学特性包括基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

62. 权利要求59-61中任一项的方法,其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

63. 权利要求59-62中任一项的方法,其中通过ATAC-seq确定染色质可及性。

64. 权利要求59-63中任一项的方法,其中评估表观遗传学特性包括:

- (1) 从细胞或细胞群分离染色质,
- (2) 用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,
- (3) 对所有或部分带有标签的片段测序以产生多个序列读段;
- (4) 比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和
- (5) 确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

65. 权利要求64的方法,其中分析或评估表观遗传学特性进一步包括比较在来自两个或更多个细胞或细胞组合物的样品之间不同的序列读段的峰,并且任选地鉴定序列读段的峰。

66. 权利要求64或权利要求65的方法,其中序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。

67. 权利要求64-66中任一项的方法,其中分析或评估表观遗传学特性进一步包括进行基因组区域的基序分析、转录因子占据分析和/或生物途径分析,所述基因组区域被鉴定为含有在来自两个或更多个细胞群的样品之间不同的序列读段的峰。

68. 权利要求64-67中任一项的方法,其中分析或评估表观遗传学特性进一步包括确定

含有序列读段峰的基因组区域内的核小体的位置。

69. 权利要求1-68中任一项的方法,其中分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与沿着一个或多个基因组区域或其子集的每一个的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者

包括,对于沿着基因组区域长度的多个位点或部分中的每一个,生成指示在所述位点或部分的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述位点或部分的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

70. 权利要求69的方法,其中所述分析任选地进一步包括确定所述表观遗传学特性的总体程度或水平,任选地确定基因组区域上的可及性的总体程度或水平。

71. 权利要求1-70中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。

72. 权利要求1-71中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化与跨越一个或多个基因组区域或其子集的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

73. 权利要求39、50、71或权利要求72的方法,其中所述值或水平是或包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。

74. 权利要求39、50和71-73中任一项的方法,其中所述值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进行总计或求和。

75. 权利要求1-74中任一项的方法,其中分析包括去除线粒体读段和/或另外的污染序列的步骤,其根据是所述读段的序列同一性、质量、映射位置或其他测序特性。

76. 权利要求1-75中任一项的方法,其中分析包括去除重复读段以提高定量准确性的步骤。

77. 权利要求1-76中任一项的方法,其中分析包括将序列读段分离成代表特定的表观遗传学特性(任选地染色质可及性或染色质占据)的子集的步骤,其中使用测序的片段的大小来确定其代表所述表观遗传学特性的程度或水平。

78. 权利要求1-6和54-77中任一项的方法,其中对来自多位各自独立地被施用第二细胞组合物的受试者的细胞组合物进行步骤(a)和(b),所述第二细胞组合物包含用重组受体工程改造的细胞。

79. 权利要求1-6和54-78中任一项的方法,其中,对于每个基因组区域或其子集,准备一个显示器,所述显示器包括针对多位受试者的每一位的映射到细胞疗法结果的每个基因组基因座的序列读段的值或水平。

80. 权利要求79的方法,其中显示器包括热图、散点图、层次聚类和/或星座图。

81. 权利要求79或权利要求80的方法,其中所述鉴定所述一个或多个基因组区域包括基于细胞疗法的结果进行聚类分析。

82. 权利要求79或权利要求80的方法,其中所述鉴定指示或关联细胞疗法的结果的所述一个或多个基因组区域包括确定具有相同或相似结果的至少大多数受试者是否在显示

器中聚集在一起。

83. 权利要求82的方法,其中如果具有相同或相似结果的至少55%、60%、70%、80%、90%、95%或更多的受试者在显示器中聚集在一起,则鉴定基因组区域。

84. 权利要求1-83中任一项的方法,其中分析了细胞的全基因组。

85. 权利要求1-84中任一项的方法,其中分析了细胞的基因组的一部分。

86. 权利要求85的方法,其中基因组的一部分包括一个或多个基因组区域,任选地为一个或多个基因组基因座,其关联或指示或可能关联或指示细胞的表型、活化状态、活化信号的强度或效应子功能。

87. 权利要求1-86中任一项的方法,其中所述分析进一步包括进行主成分分析(PCA)、生物途径分析、基因本体论(GO)分析和/或基序分析。

88. 权利要求87的方法,其中所述分析包括与T细胞记忆表型、T细胞活化状态、效应子功能、细胞因子反应、运输、持久性或衰竭相关的一个或多个基因组区域的生物途径分析和/或基因子集分析。

89. 权利要求86-88中任一项的方法,其中所述一个或多个基因组区域包含与细胞的效应子样功能或活化状态相关或指示所述效应子样功能或活化状态的一个或多个基因组基因座。

90. 权利要求86-89中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因座,该组的组成为:Nr4a1、Cblb、Irf4、Tbx21、Eomes、Ifng、Il2ra、Il2、Csf2、Gzmb、Tnfsf10、Gata3、Mir155、Sox21、Ctla4、Lag3和Pdc1。

91. 权利要求86-90中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因组基因座,该组的组成为:Ctla4、Il2ra、Il2、Ifng和Gzmb。

92. 权利要求1-91中任一项的方法,其中分析了(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的表观遗传学特性。

93. 权利要求1-92中任一项的方法,其中鉴定了包含两个或更多个基因组区域的集合。

94. 一种评估转基因整合的方法,所述方法包括:确定在用重组受体基因工程改造的细胞或细胞组合物中的包含转基因核酸序列的一个或多个基因组区域的表观遗传学特性。

95. 权利要求94的方法,其中通过将编码重组受体的核酸引入到细胞组合物中的一个或多个细胞中来进行基因工程。

96. 权利要求95的方法,其中所述引入是通过用包含核酸的病毒载体转导。

97. 权利要求94-96中任一项的方法,其中表观遗传学特性是染色质可及性。

98. 权利要求94-97中任一项的方法,其中:

所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者

所述表观遗传学特性包括基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

99. 权利要求94-98中任一项的方法,其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

100. 权利要求94-99中任一项的方法,其中通过ATAC-seq确定染色质可及性。

101. 权利要求100中任一项的方法,其中评估表观遗传学特性包括:

- (1) 从细胞或细胞群分离染色质,
- (2) 用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,
- (3) 对所有或部分带有标签的片段测序以产生多个序列读段;
- (4) 比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和
- (5) 确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

102. 权利要求101的方法,其中分析或评估表观遗传学特性进一步包括确定映射到或对应于转基因核酸序列的序列读段的峰。

103. 权利要求94-102中任一项的方法,其中序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。

104. 权利要求94-103中任一项的方法,其中分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与包含转基因核酸序列的基因组区域的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者

包括,对于包含沿着基因组区域长度的转基因核酸序列的基因组区域,生成指示在所述区域的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述区域的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

105. 权利要求94-104中任一项的方法,其中确定表观遗传学特性包括确定、测量或量化跨越包含转基因核酸序列的基因组区域的染色质可及性的值或水平。

106. 权利要求94-105中任一项的方法,其中确定表观遗传学特性包括确定、测量或量化与跨越包含转基因核酸序列的基因组区域的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

107. 权利要求1-106中任一项的方法,其中细胞组合物,任选地第一细胞组合物和/或第二细胞组合物包含获自来自受试者的样品和/或从受试者选择或分离的原代细胞。

108. 权利要求1-107中任一项的方法,其中所述细胞是免疫细胞。

109. 权利要求1-108中任一项的方法,其中免疫细胞是T细胞或NK细胞。

110. 权利要求1-109中任一项的方法,其中T细胞是CD4⁺和/或CD8⁺T细胞。

111. 权利要求1-110中任一项的方法,其中:

重组受体结合、识别或靶向与疾病或病症相关的抗原;并且/或者重组受体是T细胞受体或功能性非T细胞受体;并且/或者

重组受体是嵌合抗原受体(CAR)。

112. 权利要求111的方法,其中:

CAR包含与抗原特异性地结合的细胞外抗原识别结构域和包含ITAM的细胞内信号传导结构域,其中任选地,细胞内信号传导结构域包含CD3- ζ (CD3 ζ) 链的细胞内结构域;并且/或者其中CAR进一步包含共刺激信号传导区,其任选地包含CD28或4-1BB的信号传导结构域。

细胞疗法的表观遗传学分析及相关方法

相关申请的交叉引用

[0001] 本申请要求2017年1月10日提交的名称为“细胞疗法的表观遗传学分析及相关方法”的美国临时申请号62/444,802、2017年8月29日提交的名称为“细胞疗法的表观遗传学分析及相关方法”的美国临时申请号62/551,752和2017年12月8日提交的名称为“细胞疗法的表观遗传学分析及相关方法”的美国临时申请号62/596,662的优先权,将这些申请的内容通过提述完整并入本文。

通过提述并入序列表

[0002] 本申请与电子格式的序列表一起提交。提供的序列表为2018年1月10日创建的名称为735042009440SeqList.TXT的文件,其大小为35,537字节。将序列表电子格式信息通过提述完整并入。

领域

[0003] 在一些方面,本公开涉及鉴定基因组区域的方法,所述基因组区域预测细胞疗法的治疗结果和/或细胞的表型或功能。在一些实施方案中,所述方法包括细胞(例如用于细胞疗法的工程化细胞)的表观遗传学和/或表观基因组学分析,其与制备用于细胞疗法的工程化细胞和/或预测对细胞疗法的反应的方法有关。在一些实施方案中,所述方法包括评估、表征和分析一个或多个基因区域的表观遗传学特性的变化或修饰的步骤,所述特性比如染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和/或DNA甲基化。在一些实施方案中,表观遗传学和/或表观基因组学分析包括确定细胞例如用于细胞疗法的工程化细胞的表观遗传学特性。

背景

[0004] 有多种策略可用于制备和施用与过继细胞疗法结合使用的细胞,包括用于制备基因工程化T细胞或涉及施用基因工程化T细胞的方法,所述基因工程化T细胞比如用抗原受体(比如CAR)工程改造的T细胞。在一些方面,可用的方法可能不完全令人满意。需要制备与过继细胞疗法有关的细胞和施用所述细胞的另外的策略。提供了满足这种需要的方法。

概述

[0005] 本文中提供了鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法,所述方法包括:(a)分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞或群包含在下列组合物中:(i)第一细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二组合物,或(ii)包含所述重组受体的第二细胞组合物;和(b)鉴定所述一个或多个基因组区域中的一个或多个,其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关连细胞疗法的结果,所述细胞疗法包括施用包含所述重组受体的第二细胞组合物。在一些实施方案中,所述结果任选地是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

[0006] 本文中提供了鉴定与细胞疗法的治疗结果相关的一个或多个基因组区域的方法,所述方法包括:(a)分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,

所述细胞或群包含在下列组合中：(i) 第一细胞组合物，其有待用重组受体基因工程改造以产生包含所述重组受体的第二组合物，或(ii) 包含所述重组受体的第二细胞组合物；和(b) 鉴定所述一个或多个基因组区域中的一个或多个，其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关联细胞疗法的结果，所述细胞疗法包括对一位受试者或一组受试者施用包含所述重组受体的第二细胞组合物。

[0007] 在一些实施方案中，结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

[0008] 还提供了确定细胞组合物的一个或多个特性或特征的方法，所述方法包括分析或确定T细胞组合物中的一个或多个基因组区域的表观遗传学特性，所述T细胞组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞。

[0009] 在一些实施方案中，细胞组合物是(i) 第一T细胞组合物，其有待用重组受体基因工程改造以产生包含所述重组受体的第二T细胞组合物，或(ii) 包含所述重组受体的细胞的第二T细胞组合物。

[0010] 在一些实施方案中，所述方法进一步包括将一个或多个基因组区域中的每一个的表观遗传学特性单独地与来自不同的细胞组合物的细胞的相应表观遗传学特性和/或与参考谱进行比较，所述参考谱任选地为已知指示或关联细胞组合物的属性或特征的参考谱。

[0011] 在一些实施方案中，所述比较指示或关联细胞组合物内的细胞的状态、表型或功能，任选地为活化、效应子或记忆状态；细胞组合物内的细胞的一致性 or 均匀性；当施用于一位受试者或一组受试者时，细胞组合物是否或是否可能表现出或产生结果；外源核酸整合的位置、丰度或频率；细胞组合物内的细胞的克隆性；和/或细胞组合物中的工程化细胞的比例或频率。

[0012] 本文中提供了用于确定或鉴定与细胞组合物的属性或特征相关的表观遗传学特性的方法，所述方法包括：(a) 确定或测量在第一细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度；(b) 确定或测量在第二细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度；和(c) 比较(a)中的水平或程度和(b)中的水平或程度，其中一个或多个基因组区域的表观遗传学特性的水平或程度的差异(任选地显著差异) 鉴定或确定表观遗传学特性的存在，所述表观遗传学特性指示或关联在第一和第二组合物之一的细胞中存在但在另一者中不存在的属性或特征。

[0013] 在一些实施方案中，第一组合物和第二组合物之一包含用重组受体基因工程改造的细胞，并且第一组合物和第二组合物的另一者包含经工程改造以表达所述重组受体的细胞；第一组合物和第二组合物包含来自不同供体的原代细胞，任选地所述供体为基于疾病状态、疾病严重性或疾病类型而不同的供体；第一组合物和第二组合物包含用于工程改造细胞的制造过程的不同阶段或步骤的细胞；第一组合物和第二组合物之一包含与药剂接触以调节细胞活性、表型或功能的细胞，并且第一和第二组合物的另一者包含不被如此接触的相似细胞；或者，第一组合物和第二组合物之一包含细胞组合物的样品，所述样品与施用于受试者后发生或已经发生于第一和第二组合物之一但未发生于另一者的结果相关。在一些实施方案中，所述药剂是多肽或蛋白质、肽、抗体、核酸、病毒载体或病毒制剂或小分子化合物。在一些实施方案中，所述药剂是刺激性试剂，任选抗CD3/抗CD28；免疫调节剂、对CAR

特异的抗独特型抗体或其抗原结合片段、免疫检查点抑制剂、代谢途径调节剂、腺苷受体拮抗剂、激酶抑制剂、抗TGF β 抗体或抗TGF β R抗体或细胞因子。

[0014] 在一些实施方案中,第一组合物的属性或特征指示状态、表型或功能,任选地活化、效应子或记忆状态、表型或功能;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;细胞组合物中的工程化细胞的比例或频率;和/或当施用于一位受试者或一组受试者组时,细胞组合物是否或是否可能表现出或产生结果。

[0015] 本文中提供了评估细胞组合物的属性或特征的方法,所述方法包括:(a)分析在细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞组合物包含用重组受体工程改造的细胞和/或有待用重组受体基因工程改造的细胞;和(b)将一个或多个基因组区域的表观遗传学特性单独地与参考谱比较,其中所述比较指示细胞组合物是否或是否可能表现出所述属性或特征。

[0016] 在一些实施方案中,所述属性或特征指示状态、表型或功能,任选地活化、效应子或记忆状态、表型或功能;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;细胞组合物中的工程化细胞的比例或频率;和/或当施用于一位受试者或一组受试者组时,细胞组合物是否或是否可能表现出或产生结果。在一些实施方案中,所述属性或特征是当施用于一位受试者或一组受试者时细胞组合物是否或是否可能表现出或产生结果,并且所述方法是为了评估将施用于受试者的细胞组合物。

[0017] 本文中还提供了鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法,所述方法包括:(a)确定或测量在第一治疗组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度;(b)确定或测量在第二治疗组合物中包含的细胞或细胞群的所述一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度;(c)将一个或多个基因组区域的(a)中的水平或程度和(b)中的水平或程度进行比较。

[0018] 在一些实施方案中,所述方法进一步包括鉴定一个或多个基因组区域中的一个或多个,其中与在(b)中确定或测量的水平或程度相比,在(a)中确定或测量的水平或程度不同,任选地显著不同。

[0019] 在一些实施方案中,对于多个基因组区域中的每一个,在(a)中检测或测量的水平或程度与在(b)中检测或测量的水平或程度之间的差异或显著差异指示,所述表观遗传学特性或其程度或水平关联发生或已经发生于第一和第二治疗组合物之一但未发生于另一者的结果、预测所述结果或预测所述结果的可能性或风险,其中所述结果任选地是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

[0020] 在一些实施方案中,基因组区域包括基因组基因座或基因。在一些实施方案中,基因组区域包括基因的开放阅读框。在一些实施方案中,表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。在一些实施方案中,表观遗传学特性是染色质可及性。在一些实施方案中,所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者所述表观遗传学特性包括跨越基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

[0021] 在一些实施方案中,通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。在一些实施方案中,通过ATAC-seq确定染色质可及性。

[0022] 在一些实施方案中,分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与沿着一个或多个基因组区域或其子集的每一个的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者包括,对于沿着基因组区域长度的多个位点或部分中的每一个,生成指示在所述位点或部分的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述位点或部分的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

[0023] 在一些实施方案中,所述分析任选地进一步包括确定所述表观遗传学读出的总体程度或水平,任选地确定基因组区域上的可及性的总体程度或水平。在一些实施方案中,分析表观遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。在一些实施方案中,分析表观遗传学特性包括确定、测量或量化与跨越一个或多个基因组区域或其子集的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

[0024] 在一些实施方案中,所述值或水平是或包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。在一些实施方案中,所述值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进行总计或求和。

[0025] 在一些实施方案中,对多位各自独立地被施用第二细胞组合物的受试者进行步骤(a)和(b),所述第二细胞组合物包含用重组受体工程改造的细胞。

[0026] 在一些实施方案中,对于每个基因组区域或其子集,准备一个显示器,所述显示器包括针对多位受试者的每一位的映射到细胞疗法结果的每个基因组基因座的序列读段的值或水平。

[0027] 在一些实施方案中,显示器包括热图、散点图、层次聚类和/或星座图。在一些实施方案中,所述鉴定所述一个或多个基因组区域包括基于细胞疗法的结果进行聚类分析。在一些实施方案中,所述鉴定指示或关联细胞疗法的结果的所述一个或多个基因组区域包括确定具有相同或相似结果的至少大多数受试者是否在显示器中聚集在一起。在一些实施方案中,如果具有相同或相似结果的至少55%、60%、70%、80%、90%、95%或更多的受试者在显示器中聚集在一起,则鉴定基因组区域。

[0028] 在一些实施方案中,分析细胞的全基因组。在一些实施方案中,分析细胞的基因组的一部分。在一些实施方案中,基因组的一部分包括一个或多个基因组区域,任选地为一个或多个基因组基因座,其关联或指示或可能关联或指示细胞的表型、活化状态、活化信号的强度或效应子功能。

[0029] 在一些实施方案中,细胞疗法的结果是缓解、毒性、细胞疗法的免疫原性或表型或功能、完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

[0030] 在一些实施方案中,缓解是完全缓解、部分缓解、疾病进展或可分子检测的疾病。

[0031] 在一些实施方案中,毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。在一些实施方案中,毒性是剂量限制性毒性(DLT)。

[0032] 在一些实施方案中,分析了(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的表观遗传学特性。在一些实施方案中,鉴定了包含两个或更多个基因组区域的集合。

[0033] 在一些实施方案中,第一细胞组合物和第二细胞组合物包含从受试者选择或分离的原代细胞。在一些实施方案中,所述细胞是免疫细胞。在一些实施方案中,所述免疫细胞是T细胞或NK细胞。在一些实施方案中,T细胞是CD4⁺和/或CD8⁺ T细胞。

[0034] 在一些实施方案中,分析了第二细胞组合物。在一些实施方案中,第二细胞组合物包含编码重组受体的核酸。

[0035] 在一些实施方案中,核酸分子被包含在病毒载体中。在一些实施方案中,病毒载体是腺病毒、慢病毒、逆转录病毒、疱疹病毒或腺相关病毒载体。

[0036] 在一些实施方案中,通过在一种或多种条件或试剂的存在下培养输入组合物来产生第一细胞组合物和/或第二细胞组合物。在一些实施方案中,一个或多个基因组区域包含涉及或可能涉及细胞的活化状态或效应子状态的基因。

[0037] 本文中还提供了评估将施用于受试者的细胞组合物的方法,所述方法包括:(a) 分析在细胞组合物中包含的细胞的一个或多个基因组区域的表观遗传学谱,所述细胞组合物包含用重组受体工程改造的细胞;和(b) 将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示当施用于受试者时细胞群是否或是否可能表现出或产生结果。

[0038] 在一些实施方案中,细胞疗法的结果是缓解、毒性、细胞疗法的免疫原性或表型或功能、完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的持久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。在一些实施方案中,缓解是完全缓解或部分缓解。

[0039] 在一些实施方案中,如果所述比较指示细胞组合物将要或可能将要表现出结果,则将细胞组合物施用于受试者。

[0040] 在一些实施方案中,如果所述比较指示细胞组合物不会或可能不会表现出结果,则任选下列之一:(i) 施用其中细胞组合物被改变的细胞组合物;(ii) 施用其中细胞剂量被改变的细胞组合物;(iii) 施用其中施用于受试者的细胞的剂量方案被改变的细胞组合物;(iv) 将细胞组合物与一种或多种其他治疗剂联合施用;或(v) 不对受试者施用所述细胞组合物。

[0041] 在一些实施方案中,在施用改变的细胞组合物之前,对包含在改变的细胞组合物中的细胞重复步骤(a)和(b)。

[0042] 在一些实施方案中,改变细胞的给药方案包括在对受试者施用第一剂量的细胞之后对受试者施用第二剂量的细胞。在一些实施方案中,在施用第一剂量的细胞之后至少1周、2周、3周、4周、2个月、3个月、4个月、5个月、6个月、9个月或12个月施用后续剂量的细胞。

[0043] 在一些实施方案中,一个或多个基因组区域关联或指示对细胞疗法的反应。

[0044] 在一些实施方案中,参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

[0045] 在一些实施方案中, 阈值: 是当施用于具有相同或相似疾病或病症的受试者时显示表现出结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者是来自当施用于受试者时显示表现出结果的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。在一些实施方案中, 阈值包括来自正常或健康受试者的细胞中表观遗传学特性的值或水平。在一些实施方案中, 阈值包括表现出幼稚或长寿记忆表型的细胞中的表观遗传学特性的值或水平。

[0046] 在一些实施方案中, 阈值: 是与当施用于具有相同或相似疾病或病症的受试者时显示表现出期望的结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平;或者是与来自已经单独施用于一组受试者的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的平均、中位或平均的值或水平, 或者在所述平均、中位或平均的值或水平的标准偏差之内, 其中该组受试者的每一位在施用于受试者后继续显示表现出期望的结果;或者是与来自正常或健康受试者的相似细胞组合物中的表观遗传学特性相关或指示所述表观遗传学特性的值或水平。

[0047] 本文中还提供了评估细胞培养物的方法, 所述方法包括: (a) 分析包含在输出细胞组合物中的细胞的一个或多个基因组区域的表观遗传学谱, 所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生; 和 (b) 将每个基因组区域的表观遗传学谱单独地与参考谱比较, 其中所述比较指示所述细胞是否或是否可能表现出预定的表型或功能。

[0048] 在一些实施方案中, 预定的表型或功能指示细胞的效应子功能或活化状态和/或指示细胞表现出幼稚表型或长寿记忆表型。

[0049] 在一些实施方案中, 一种或多种测试剂或条件包括血清的存在或浓度; 培养时间; 刺激剂的存在或量; 刺激剂的类型或程度; 氨基酸的存在或量; 温度; 输入组合物的来源或细胞类型; 输入组合物中细胞类型的比率或百分比, 任选地CD4+/CD8+细胞的比率; 珠子的存在或量; 细胞密度; 静置培养; 摇摆培养; 灌注; 病毒载体的类型; 载体拷贝数; 转导佐剂的存在; 冷冻保存中的输入组合物的细胞密度; 重组受体的表达程度; 或调节细胞表型的化合物的存在。在一些实施方案中, 一种或多种测试剂或条件包括来自测试化合物文库的一种或多种化合物。

[0050] 在一些实施方案中, 所述方法包括如果所述比较指示细胞组合物具有或可能具有所述表型或功能, 则选择用于培养细胞的一种或多种测试剂或条件。在一些实施方案中, 所述方法包括如果所述比较指示细胞组合物具有或可能不具有所述表型或功能, 则用一种或多种另外的测试剂或条件重复步骤 (a) 和 (b)。

[0051] 在一些实施方案中, 参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

[0052] 在一些实施方案中, 阈值: 是显示表现出所述表型或功能的参考细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者是来自显示表现出所述表型或功能的多个参考细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水

平。

[0053] 在一些实施方案中,参考细胞组合物具有指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型。

[0054] 在一些实施方案中,分析表观遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。在一些实施方案中,分析表观遗传学特性包括确定、测量或量化与一个或多个基因组区域或其子集中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的序列读段的值或水平。在一些实施方案中,确定、测量或量化值或水平是或包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。在一些实施方案中,确定、测量或量化值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进行总计或求和。

[0055] 在一些实施方案中,所述一个或多个基因组区域包括包含至少2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的集合。

[0056] 在一些实施方案中,所述一个或多个基因组区域包含与细胞的效应子样功能或活化状态相关或指示所述效应子样功能或活化状态的一个或多个基因组基因座。

[0057] 在一些实施方案中,所述一个或多个基因组区域包括选自下组的基因座,该组的组成为:Nr4a1、Cblb、Irf4、Tbx21、Eomes、Ifng、Il2ra、Il2、Csf2、Gzmb、Tnfsf10、Gata3、Mir155、Sox21、Ctla4、Lag3和Pdcd1。在一些实施方案中,所述一个或多个基因组区域包括选自下组的基因组基因座,该组的组成为:Ctla4、Il2ra、Il2、Ifng和Gzmb。

[0058] 在一些实施方案中,基因组区域包括基因组基因座或基因。在一些实施方案中,基因组区域包括基因的开放阅读框。在一些实施方案中,基因组区域包含基因间区或调控元件。在一些实施方案中,基因组区域包含内含子、外显子、顺式调控元件、启动子、增强子、上游激活序列(UAS)、3'非翻译区(UTR)、5'UTR、非编码RNA产生区、非编码RNA(ncRNA)基因、miRNA基因、siRNA基因、piRNA基因、snoRNA基因、lncRNA基因、核糖体RNA(rRNA)基因、小RNA结合位点、非编码RNA结合位点、假基因、转录终止位点(TTS)、重复序列、端粒区、可及的染色质区、不可及的染色质区、开放染色质区和/或异染色质区。

[0059] 在一些实施方案中,表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。在一些实施方案中,表观遗传学特性是染色质可及性。在一些实施方案中,通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。在一些实施方案中,通过ATAC-seq确定染色质可及性。

[0060] 在一些实施方案中,评估表观遗传学特性包括:(1)从细胞或细胞群分离染色质,(2)用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,(3)对所有或部分带有标签的片段测序以产生多个序列读段;(4)比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和(5)确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。在一些实施方案中,分析或评估表观遗传学特性进一步包括比较在来自两个或更多个细胞或组合物的样品之间不同的序列读段的峰,并且任选地鉴定序列读段的峰。在一些实施方案中,序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。在一些实施方案中,分析或评估表

观遗传学特性进一步包括进行基因组区域的基序分析、转录因子占据分析和/或生物途径分析,所述基因组区域被鉴定为含有在来自两个或更多个细胞群的样品之间不同的序列读段的峰。在一些实施方案中,分析或评估表观遗传学特性进一步包括确定含有序列读段峰的基因组区域内的核小体的位置。

[0061] 在一些实施方案中,分析包括去除线粒体读段和/或另外的污染序列的步骤,其根据是所述读段的序列同一性、质量、映射位置或其他测序特性。在一些实施方案中,分析包括去除重复读段以提高定量准确性的步骤。在一些实施方案中,分析包括将序列读段分离成代表特定的表观遗传学特性(任选地染色质可及性或染色质占据)的子集的步骤,其中使用测序的片段的大小来确定其代表所述表观遗传学特性的程度或水平。

[0062] 在一些实施方案中,分析进一步包括进行主成分分析(PCA)、生物途径分析、基因本体论(GO)分析和/或基序分析。

[0063] 在一些实施方案中,所述细胞获自来自受试者的样品。在一些实施方案中,所述细胞是免疫细胞,任选T细胞,任选CD4⁺和/或CD8⁺ T细胞。

[0064] 在一些实施方案中,重组受体结合、识别或靶向与疾病或病症相关的抗原;并且/或者重组受体是T细胞受体或功能性非T细胞受体;并且/或者重组受体是嵌合抗原受体(CAR)。在一些实施方案中,CAR包括与抗原特异性地结合的细胞外抗原识别结构域和包含ITAM的细胞内信号传导结构域,其中任选地,细胞内信号传导结构域包括CD3- ζ (CD3 ζ)链的细胞内结构域;并且/或者其中CAR进一步包括共刺激信号传导区,其任选地包括CD28或4-1BB的信号传导结构域。

[0065] 本文中还提供了包含多个细胞的细胞组合物,其中一个集合中的一个或多个基因的表观遗传学特性的水平或值高于或低于组合物中至少50%细胞中的阈值。

[0066] 在一些实施方案中,所述水平或值高于或低于组合物中至少60%、70%、75%、80%、85%、90%、95%或更多细胞中的阈值。在一些实施方案中,阈值:是显示表现出所述表型或功能的参考细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者是来自显示表现出所述表型或功能的多个参考细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。

[0067] 在一些实施方案中,参考细胞组合物具有指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型。

[0068] 在一些实施方案中,所述集合包括(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域。

[0069] 本文中还提供了评估转基因整合的方法,所述方法包括:确定在用重组受体基因工程改造的细胞或细胞组合物中的包含转基因核酸序列的一个或多个基因组区域的表观遗传学特性。在一些实施方案中,通过将编码重组受体的核酸引入到细胞组合物中的一个或多个细胞中来进行基因工程。

[0070] 在一些实施方案中,引入是通过用包含核酸的病毒载体转导。在一些实施方案中,表观遗传学特性是染色质可及性。在一些实施方案中,表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者表观遗传学特性包括基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

[0071] 在一些实施方案中,通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。在一些实施方案中,通过ATAC-seq确定染色质可及性。

[0072] 在一些实施方案中,评估表观遗传学特性包括:(1)从细胞或细胞群分离染色质,(2)用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,(3)对所有或部分带有标签的片段测序以产生多个序列读段;(4)比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和(5)确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

[0073] 在一些实施方案中,分析或评估表观遗传学特性进一步包括确定映射到或对应于转基因核酸序列的序列读段的峰。在一些实施方案中,序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。在一些实施方案中,分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与包含转基因核酸序列的基因组区域的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者包括,对于包含沿着基因组区域长度的转基因核酸序列的基因组区域,生成指示在所述区域的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述区域的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。在一些实施方案中,确定表观遗传学特性包括确定、测量或量化跨越包含转基因核酸序列的基因组区域的染色质可及性的值或水平。在一些实施方案中,确定表观遗传学特性包括确定、测量或量化与跨越包含转基因核酸序列的基因组区域的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。在一些实施方案中,细胞组合物,任选地第一细胞组合物和/或第二细胞组合物包含获自来自受试者的样品和/或从受试者选择或分离的原代细胞。

[0074] 在一些实施方案中,所述细胞是免疫细胞。在一些实施方案中,所述免疫细胞是T细胞或NK细胞。在一些实施方案中,T细胞是CD4⁺和/或CD8⁺ T细胞。

[0075] 在一些实施方案中,重组受体结合、识别或靶向与疾病或病症相关的抗原;并且/或者重组受体是T细胞受体或功能性非T细胞受体;并且/或者重组受体是嵌合抗原受体(CAR)。在一些实施方案中,CAR包含与抗原特异性地结合的细胞外抗原识别结构域和包含ITAM的细胞内信号传导结构域,其中任选地,细胞内信号传导结构域包含CD3- ζ (CD3 ζ)链的细胞内结构域;并且/或者其中CAR进一步包含共刺激信号传导区,其任选地包含CD28或4-1BB的信号传导结构域。

附图简述

[0076] 图1A显示了通过细胞内细胞因子染色(ICS)测量的干扰素- γ (IFN γ)产生(在x轴上显示)与通过ATAC-seq确定的编码IFN γ (Ifng)的基因的可及性(在y轴上显示)的相关性。

[0077] 图1B显示了通过ICS测量的程序性细胞死亡蛋白1(PD-1)产生(在x轴上显示)与通过ATAC-seq确定的编码PD1(Pdcd1)的基因的可及性(在y轴上显示)的相关性。

[0078] 图1C显示了通过ICS测量的再刺激后产生IFN γ 的细胞的%(在x轴上显示)与通过ATAC-seq确定的编码IFN γ (Ifng)的基因的可及性(在y轴上显示)的相关性。

[0079] 图1D显示了通过ICS测量的再刺激后产生白细胞介素2(IL-2)的细胞的%(在x轴

上显示)与通过ATAC-seq确定的编码IL-2 (I12)的基因的可及性(在y轴上显示)的相关性。

[0080] 图2通过工程化T细胞组合物的CD4⁺或CD8⁺ T细胞中的选择基因的细胞内细胞因子染色(ICS)(左图)或通过ATAC-seq测量的在相同的工程化T细胞组合物中的CD4⁺或CD8⁺中的基因的染色质可接近性(右图)显示了蛋白质表达,其各自与细胞疗法施用于受试者后的反应结果相关。

[0081] 图3A显示了使用ATAC-seq对来自含有表达抗CD19 CAR的基因工程化人T细胞的T细胞组合物的CD8⁺细胞进行的全基因组分析的结果。每列代表基于每个基因的染色质可及性差异的层次聚类,如通过每个基因的基因体上FPKM的总和计算的,其被显示为低(蓝色)或高(红色)。将受试者按缓解组分类,包括显示完全缓解(CR)、疾病进展(PD)或部分缓解(PR)证据的受试者。还根据来自正常供体(ND)的CD8⁺细胞显示了通过ATAC-seq进行的染色质可及性结果。星号指示在三个月时从CR转换为PD的受试者。

[0082] 图3B显示了展示按照缓解组在全基因组ATAC-seq数据上观察到的CD8⁺CDP聚类的聚类决策树(星座图)。

[0083] 图4A显示了使用ATAC-seq对来自全基因组测序数据的基因的靶向集合的子集的分析结果,所述全基因组测序数据在来自含有表达抗CD19CAR的基因工程化人T细胞的CDP的CD8⁺细胞上进行。

[0084] 图4B显示了基于按照缓解组获自CDP的CD8⁺细胞中的特定基因的ATAC-seq数据的染色质可及性的聚类决策树(星座图)。

[0085] 图5A显示了借助于ATAC-seq的基因的染色质可及性(表示为每个基因的基因体上的FPKM总和的相对值),所述基因指示在没有经历基因工程改造的冷冻保存的FACS纯化的CD4⁺/CD8⁺细胞组合物(CMAT)中的效应细胞表型。

[0086] 图5B显示了借助于ATAC-seq的基因的染色质可及性(表示为每个基因的基因体上的FPKM总和的相对值),所述基因指示CD8⁺CDP中的效应细胞表型。

[0087] 图6A和6B显示了按照缓解分组的受试者的每个基因的基因体上FPKM的总和测量的与信号强度和效应子功能相关的基因座的染色质可及性。

[0088] 图7A和7B显示了在表达抗BCMA CAR的T细胞中的基因表达的主成分分析(PCA)(基于RNA-seq结果;图7A)和染色质可及性(基于ATAC-seq结果;图7B)的结果,所述表达抗BCMA CAR的T细胞从4个不同的供体(供体1-4)生成,其经过用BCMA缀合的珠子刺激24小时(24hr+stim)或7天(d7+stim),或者在来那度胺存在或不存在下在没有刺激的情况下培养24小时(24hr)。

[0089] 图8A和8B显示了描绘在CAR⁺ T细胞中随着基因表达的log₂倍数变化的表达的统计显著性(调整的p值的log₁₀)的火山图,所述CAR⁺ T细胞在来那度胺存在或不存在下用BCMA缀合的珠子刺激24小时(24hr+stim,图8A)或7天(d7+stim,图8B)。表格指示了显示表达上升(上)或下降(下)的基因或峰的数目。

[0090] 图8C和8D显示了描绘在CAR⁺ T细胞中随着染色质可及性的log₂倍数变化的染色质可及性变化的统计显著性(调整的p值的log₁₀)的火山图,所述CAR⁺ T细胞经过用BCMA缀合的珠子刺激24小时(24hr+stim,图8C)或7天(d7+stim,图8D)。表格指示了显示可及性上升(上)或下降(下)的基因或峰的数目。

[0091] 图9A和9B描绘了在用BCMA缀合的珠子刺激24小时(24hr+stim,图9A)或7天(d7+

stim,图9B)的CAR+ T细胞中对生物途径的影响的方向性和显著性。

[0092] 图10显示了比较每个基因的单体的染色质可及性峰(菱形)和平均染色质可及性变化(圆形)的图,所述比较针对涉及T细胞活化和信号传导的选定基因的基因表达变化而言。

[0093] 图11显示了基序富集分析、富集对数p值、预计结合峰基序的流行与转录因子,所述峰在第7天培养物中在来那度胺存在下具有上升的可及性。

[0094] 图12A和12B显示了在不同的冷冻保存的CD4+或CD8+工程化细胞组合物(CDP)或未经工程改造的匹配样品(CMAT)中的示例性免疫基因(例如,细胞表面标志物CD3 ϵ 、CD8a、CD8b和CD4)上的示例性染色质可及性谱;在一些情况下,通过表型将CMAT样品分离为幼稚T细胞(T_N)、中枢记忆T细胞(T_{CM})、效应和效应记忆T细胞(T_{E+EM})或效应记忆RA(T_{EMRA})。

[0095] 图13A显示了在CD27+CCR7+、CD27+CCR7-和CD27-CCR7-细胞以及大量CD8+细胞中比较的在编码区和CCR7基因附近的基因间区中的示例性染色质可及性峰谱。图13B显示了在CD27+CCR7+、CD27+CCR7-和CD27-CCR7-细胞以及大量CD8+细胞的各个基因组位置(包括基因的基因间区、内含子区和启动子区)内的可及性峰的分布。

[0096] 图14A显示了在CD4+和CD8+CDP细胞中使用MACS2调用的总染色质可及性峰的数目,以及使用NucleoATAC确定的无核小体区域的数目,所述CDP细胞来自按反应结果分组(1个月CR、3个月CR、1个月PD、3个月PD或PR)的已经被施用工程化CAR-T细胞的受试者。

[0097] 图14B-14D显示了与在3个月时实现CR的受试者比较的在3个月时实现PD的受试者中的两个示例性免疫相关基因(基因1:图14B和基因2:图14C)附近的基因组区域的示例性差异性可及性峰谱和可及性峰的定量。图14D显示了来自按反应结果分组(1个月CR、3个月CR、1个月PD、3个月PD或PR)的已经被施用工程化CAR-T细胞的受试者的峰的定量。

[0098] 图15A显示了在经工程改造以表达CAR的冷冻保存的工程化细胞组合物(CDP)或未经工程改造的匹配样品(CMAT)中编码抗CD19 CAR的病毒载体序列的整合体的比例数(计算为(比对读段x读长)/(构建体大小))。图15B显示了通过针对用编码抗CD19 CAR的载体(CAR整合体)转导的细胞和用空病毒载体(空整合体)转导的细胞将真阳性率与假阳性率作图而生成的接受者操作特征(ROC)曲线。

[0099] 图16A显示了CDP和CMAT样品中映射到CAR序列的ATACseq读段的数目。图16B显示了在工程化细胞中使用ATAC-seq确定的整合体数与通过定量聚合酶链反应(qPCR)确定的载体拷贝数(VCN)相比较的图。

[0100] 图17A和17B显示了在按反应结果分组(1个月CR、3个月CR、1个月PD、3个月PD或PR)(排除正常供体样品)的已经被施用工程化CAR-T细胞的抗CD19 CAR+CD4+和CD8+ T细胞受试者中的VCN(图17A)和使用ATAC-seq确定的整合体数(图17B)。

[0101] 图18显示了通过在来自按反应结果分组(1个月CR、3个月CR、1个月PD、3个月PD或PR)(排除正常供体样品)的受试者的CD4+和CD8+CDP细胞中的具有未知克隆性的50,000个细胞的不一致读段对的映射而评估的独特整合位点的数目。

[0102] 图19A显示了评估的在跨越来自7位示例性受试者的不同CD8+CMAT细胞样品中的编码T细胞受体 β 可变(TRBV)区的基因座的示例性染色质可及性峰谱。图19B显示了ND或实现CR或PD的受试者中的CD8+CDP样品中的总TCR可及性和变异系数(CV)%。图19C显示了在来自已经接受施用抗CD19 CAR+ T细胞而实现CR、PR或PD的受试者的CD8+抗CD19 CAR+ T细

胞中的总体相对TCR可及性。

[0103] 图20A显示了根据来自分离自三(3)位健康供体的T细胞的CDP和CMAT的差异性可及性分析描绘随着染色质可及性的 \log_2 倍数变化的染色质可及性的统计显著性(调整的p值的 \log_{10})的火山图。表格指示了显示可及性上升(上)或下降(下)的基因或峰的数目。

[0104] 图20B和20C显示了来自免疫相关基因的基因模块分析的“细胞因子”模块(图20B)和“耗尽”模块(图20C)中的基因的示例性个体启动子的启动子可及性(计数_frips)。

[0105] 图21显示了在CD4⁺和CD8⁺细胞群中CD8A启动子的可及性与CD4启动子的可及性之比的相对可及性。

[0106] 图22A显示了展示在与实现完全缓解(CR)的受试者相比具有疾病进展(PD)的受试者中具有更高或更低可及性的峰的 \log_2 倍数变化和调整的p值的火山图,所述完全缓解为最佳总缓解(BOR)、3个月持续缓解(3MO)或6个月持续缓解(6MO)。表格指示了显示可及性上升(上)或下降(下)的基因或峰的数目。CR或PD组中的受试者数目显示在每个图的左上角。

[0107] 图22B显示了在与具有0-2Ntx等级的受试者相比已发生3-5级神经毒性(Ntx)、或者与具有0-1级CRS的受试者相比具有2-5级细胞因子释放综合征的受试者中具有更高或更低可及性的峰的倍数变化和调整的p值,其中相应数目的峰差异性地存在于样品中。表格指示了显示可及性上升(上)或下降(下)的基因或峰的数目。每个毒性等级的受试者数目显示在每个图的左上角。

[0108] 图23A和23B显示了展示对于CDP(图23A)或CMAT(图23B)在基于缓解组的受试者中具有更高或更低可及性的峰的火山图。图23C和23D显示了展示对于CDP(图23C)或CMAT(图23D)在基于神经毒性(Ntx)或细胞因子释放综合征(CRS)的受试者中具有更高或更低可及性的峰的火山图。

[0109] 图24A显示了在CD8⁺CDP和CMAT样品中具有FDR<0.1的鉴定峰的数目,并且图24B显示了以FRiPs指示的富集,其中使用改良或标准的ATAC-seq,具有3个技术重复。

详细说明

评估细胞的表观遗传学状态的方法

[0110] 提供了用于确定细胞或细胞组合物的一个或多个基因组区域的表观遗传学谱或特征的方法,所述细胞组合物包括含有与过继细胞疗法结合使用的源自受试者的原代细胞例如T细胞的细胞组合物。在一些实施方案中,细胞组合物包括与制造或工程化细胞疗法有关的组合物,包括在用重组受体例如嵌合抗原受体(CAR)使细胞工程化之前和之后的组合物。在一些方面,细胞或细胞组合物的基因组区域的表观遗传学谱或特征和/或一种或多种表观遗传学特性可提供关于细胞的某些特征或特性的信息,包括关联细胞或细胞组合物的表型、活化状态和/或功能和/或指示、关联、关连和/或预测细胞疗法的一种或多种治疗结果的那些信息。这些方法基于以下观察:某些基因或基因组区域的表观遗传学特性(比如染色质可及性)允许详细、精细和/或全面地跟踪细胞组合物的特征和特性,这对于诸如依赖于转录的方法(例如RNA测序)的其他系统是不可能的。在一些方面,发现这样的表观遗传学特性与诸如疾病结果、反应结果、毒性结果之类的结果和/或细胞的表型、持久性、活性和/或功能相关。使用现有分析方法,比如评估转录、蛋白质表达的方法和/或通过细胞内染色,在某些方面未观察到这样的相关性。

[0111] 提供了基于基因组区域的表观遗传学特性和/或根据一个或多个基因组区域的表

观遗传学谱所确定来鉴定指示、关联、关连和/或预测结果(比如细胞疗法的治疗结果)的基因组区域的方法。还提供了基于特定基因组区域的表观遗传学特性来评估用于过继细胞疗法的细胞组合物或细胞培养组合物的方法。

[0112] 所提供的方法可用于鉴定在施用过继细胞疗法之前预测治疗结果的基因组区域。本文中提供了鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法。所提供的方法可用于鉴定指示、关联、关连和/或预测过继细胞疗法的特定结果或特性(例如关于期望的反应和/或安全性结果)的表观遗传学谱。在一些实施方案中,所提供的方法可用于在施用之前评估细胞或细胞组合物的一种或多种特性或特征,例如通过将表观遗传学谱与另一样品的表观遗传学谱和/或与参考谱比较。在一些实施方案中,所述方法包括分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞或群包含在下列组合物中:第一细胞组合物,其有待用重组受体基因工程改造以产生含有所述重组受体的第二组合物,或含有所述重组受体的第二细胞组合物;并且鉴定所述一个或多个基因组区域中的一个或多个,其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关连细胞疗法的结果,所述细胞疗法包括施用含有所述重组受体的第二细胞组合物。

[0113] 在一些实施方案中,所提供的方法可用于测量或量化基因组区域例如一个或多个基因组基因座和/或一个或多个基因(比如基因的集合)的表观遗传学特性,比如染色质可及性,以提供关于与细胞疗法例如CAR+ T细胞疗法结合使用的细胞的特征或特点的信息,所述特征或特点包括预测对细胞疗法的反应或指示期望的细胞表型或功能的特征或特点。在一些方面,所提供的方法可以与优化或改善细胞疗法结合使用,所述改善细胞疗法包括改善疗法的结果,例如在施用细胞疗法之后的反应和/或安全性结果和/或细胞疗法的质量。

[0114] 在一些情况下,这优于现有的方法和细胞疗法,因为现有的方法和细胞疗法可能难以预测反应,可能难以确定最佳剂量,并且/或者细胞疗法的质量是可变的。所提供的方法可用于在施用前提供更好的关于用于过继细胞疗法的工程化细胞的特征和特点的信息,使得可以容易且快速地确定最佳剂量,以提高细胞疗法的功效和安全性。所提供的方法还可用于鉴定或表征与制造或工程化细胞疗法有关的细胞,包括与各种过程参数(例如温度、培养条件和其他参数)的影响有关的细胞,所述参数可以或可能影响细胞的表型活性、持久性或功能。

[0115] 在一些实施方案中,所提供的方法可用于鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法,所述方法包括:(a) 确定或测量在第一治疗组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度;(b) 确定或测量在第二治疗组合物中包含的细胞或细胞群的所述一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度;和(c) 将一个或多个基因组区域的(a)中的水平或程度和(b)中的水平或程度进行比较。

[0116] 在一些实施方案中,所提供的方法可用于评估将施用于受试者的细胞组合物,所述方法包括:分析在细胞组合物中的细胞的一个或多个基因组区域的表观遗传学谱,所述细胞组合物含有用重组受体工程改造的细胞;并且将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示当施用于受试者时细胞群是否或是否可能表现出或产生结果。在一些实施方案中,所提供的评估细胞培养物的方法包括分析包含在输出细胞组

合物中的细胞的一个或多个基因组区域的表观遗传学谱,所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生;并且将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示所述细胞是否或是否可能表现出预定的表型、持久性、活性和/或功能。

[0117] 在一些实施方案中,所提供的方法可用于评估用于细胞疗法的细胞的状态、质量、一致性、表型、克隆性、均匀性、特点和/或特性;选择指示、关联、关连和/或预测细胞疗法的特定结果或特性(例如,具有期望的反应和/或安全性结果)的细胞或细胞组合物;和/或修改或改变细胞的剂量、类型和/或工程化过程的一个或多个步骤或参数,使得可以基于表观遗传学特性的评估来优化或改进供施用的细胞组合物。

[0118] 在一些方面,所提供的方法可用于确定细胞例如用于过继细胞疗法的细胞的状态、质量、一致性、表型、克隆性、均匀性、特点和/或特性。在一些方面,所提供的方法可用于一个或多个工程化阶段的细胞、或在工程化之前从受试者获得的细胞组合物,用于不同阶段的纯化或选择的细胞亚群、或在施用工程化细胞之后从受试者获得的细胞。在一些方面,所述方法可用于在施用于受试者之前评估细胞的状态、质量、一致性、表型、克隆性、均匀性、特点和/或特性,并且来自分析方法的结果可用于选择将治疗的受试者,确定治疗方案,包括剂量和频率和/或另外的治疗,和/或修改或改变工程化或制造过程以获得更合乎需要的供施用的细胞组合物。在一些实施方案中,所述方法可用于获得更均匀且预测有效的供施用的细胞组合物,以增强功效和/或减少副作用。在一些方面,所提供的方法可以与其他方法或测定法联合使用或结合使用以表征细胞群中的细胞,所述测定法例如确定细胞表面标志物表达、细胞的持久性、活力和/或扩增以确定与这样的特点和/或表型的任何相关性的测定法。

[0119] 在一些实施方案中,可以做出各种改变,并且在各种实施方案中可以取代等同物。另外,可以进行许多修改以使特定的情况、材料、物质组成、过程、过程管理条例或步骤适应于各种实施方案的目的、精神或范围。在一些实施方案中,本文中描述和展示每个单独的变化形式具有离散的部件和特征,其可以容易地与任何其他几个实施方案的特征分离或组合,而不脱离各种实施方案的范围或精神。所有这样的修改预期落入与本公开相关的权利要求书的范围内。

表观遗传学/表观基因组学分析

[0120] 在任何提供的方法中,表观遗传学和/或表观基因组学分析可包括评估、表征和分析在基因座、多个基因座或基因组基因座、基因组区和/或基因组中的变化或修饰的步骤,所述变化或修饰比如染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和/或DNA甲基化。

[0121] 在一些方面,一个或多个基因组区域包括基因组基因座或基因。在一些方面,基因组基因座包括基因组中的固定位置,并且可包括基因的编码区、开放阅读框、非编码区、基因间区或调控元件。在一些实施方案中,一个或多个基因组区域、基因座、元件或区间包括编码区、非编码区、基因间区、内含子、外显子、近端和远端顺式调控区、启动子区、增强子区、上游激活序列(UAS)、转录物的非翻译区(UTR,例如3'UTR或5'UTR)、非编码RNA产生区、非编码RNA(ncRNA)基因(例如,miRNA、siRNA、piRNA、snoRNA或lncRNA)、核糖体RNA(rRNA)基因、小RNA结合位点、非编码RNA结合位点、假基因、转录终止位点(TTS)、重复序列、端粒区

和/或可及或不可及的区域(例如,开放染色质和/或异染色质)。

[0122] 在一些实施方案中,所提供的方法涉及一个或多个表观遗传学和/或表观基因组学分析步骤。在一些实施方案中,分析包括大规模分析,例如,分析多个基因组区域、基因组基因座、基因座或全基因组分析。在一些实施方案中,表观遗传学和/或表观基因组学分析包括确定细胞例如用于细胞疗法的工程化细胞的表观遗传学特性、状态和/或谱。在一些实施方案中,所提供的方法涉及确定细胞或细胞群或组合物的表观遗传学/表观基因组学特性、状态和/或谱。在一些实施方案中,所述方法可涉及使用测序,比如大规模测序,例如高通量测序或下一代测序,以评估细胞或细胞组合物的表观遗传学/表观基因组学特性、状态和/或谱。在一些实施方案中,所述方法涉及比对和/或过滤从测定法获得的序列和/或将序列映射到基因组,例如参考基因组。在一些方面,所述方法涉及在序列被映射的情况下确定基因组区域、基因座和/或区间,比如确定映射到基因组的特定区域、基因座和/或区间的序列读段的峰。

[0123] 在一些实施方案中,所提供的方法还涉及,例如,通过比较特定细胞或细胞组合物与另一细胞或细胞组合物的表观遗传学/表观基因组学特性和/或谱来分析表观遗传学/表观基因组学特性和/或谱。在一些实施方案中,所提供的方法涉及用于分析或比较的各个下游步骤或过程,例如计算实施的步骤和/或方法的应用,用于评估细胞或细胞组合物的一种或多种特性或特点。

[0124] 在一些方面,用于确定或评估表观遗传学/表观基因组学特性、状态和/或谱的示例性方法(例如使用诸如ATAC-seq的测定法)及其分析和/或应用可涉及以下一个或多个步骤:1)生成ATACseq文库;2)修剪和映射读段;3)去除重复读段;4)过滤线粒体污染;5)过滤非核小体片段;6)调用可及性峰;7)组装共有峰集;8)将峰中的读段计数;9)将样品聚类;和/或10)执行差异性可及性分析。在一些实施方案中,示例性表观遗传学和/或表观基因组学分析包括评估染色质的状态,例如染色质可及性、开放性或压缩。在所提供方法的一些实施方案中,表观遗传学和/或表观基因组学分析包括评估染色质可及性。在一些实施方案中,染色质可及性分析与用于大规模测序例如高通量测序或下一代测序的步骤偶联。在一些实施方案中,用于评估染色质可及性的方法包括评估核小体占据、组蛋白修饰和/或转录因子占据。在一些实施方案中,使用DNA插入元件、DNA修饰酶(例如, DNase或MNase)和/或抗体(包括其片段)确定染色质可及性。用于表观遗传学和/或表观基因组学分析的示例性测定法包括用于鉴定基因组的蛋白质结合位点的与高通量测序偶联的染色质免疫沉淀(ChIP-seq)、用于以碱基对分辨率确定DNA甲基化的亚硫酸氢盐测序、DNaseI-Seq、用于评估开放染色质的具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)、以及用于确定染色体的空间组织的染色体构象捕获(3C)和相关方法,比如染色体构象捕获芯片(4C)、染色体构象捕获碳拷贝(5C)、Hi-C、3C-Seq(具有高通量测序的3C)、4C-Seq、5C-Seq和HiC-Seq。在一些实施方案中,表观遗传学和/或表观基因组学分析包括具有高通量测序的甲醛辅助的调控元件分离(FAIRE-seq)。

确定表观遗传学/表观基因组学谱

[0125] 在所提供方法的一些实施方案中,表观遗传学和/或表观基因组学分析,比如确定表观遗传学/表观基因组学特性、状态和/或谱,包括评估基因组区域、基因座和/或区间和/或全基因组水平(例如,遍及大部分或整个基因组)的染色质可及性。在一些方面,基于特定

的细胞或细胞组合物的一个或多个基因组区域、基因座和/或区间和/或全基因组水平的表观遗传学/表观基因组学特性,确定表观遗传学/表观基因组学谱。

[0126] 在本文中提供的任何分析方法的一些实施方案中,使用计算机系统来执行一个或多个步骤、功能、过程或脚本。在一些实施方案中,计算机系统被整合到分析系统中并且是分析系统的一部分,所述分析系统例如液体处理器、桥式扩增系统(例如Illumina cBot)和/或测序系统(例如Illumina Genome Analyzer、HiSeq或MiSeq系统)。在一些实施方案中,计算机系统被连接到或移植到分析系统。

使用ATAC-seq的染色质可及性评估

[0127] 在一些实施方案中,使用与高通量测序方法偶联的DNA插入元件评估染色质可及性。在一些实施方案中,使用ATAC-seq(比如在US20160060691中描述的方法,将其通过提述完整并入本文,以及其中的方法的任何变化形式)评估染色质可及性。在一些实施方案中,用于表观遗传学和/或表观基因组学分析的示例性测定法包括在例如W02015102536、W02015159295、W02015130968、W0201692070,Dirks等人,Clinical Epigenetics(2016) 8:122,Buenroostro等人,Nat Methods.(2013) 10(12):1213-1218,Sung等人,Nat Methods.(2016) 13(3):222-228中描述的那些,将以上文献通过提述完整并入本文。在一些实施方案中,染色质可及性测定法比如ATAC-seq具有诸如下列优势:文库制备简单,测定时间短(与某些测定法需要3-4天相比,可在数小时内获得结果),不需要超声处理或苯酚-氯仿提取,不需要抗体(消除可由抗体引入的限制和偏差),不需要敏感的酶消化(消除费力且敏感的酶滴定),只需要极少量的细胞作为输入,以及宽的动态范围。在一些实施方案中,染色质可及性测定法比如ATAC-seq具有诸如下列优势:在单次测定中和/或在单个时间点提供对组合物中的细胞的状态、质量、一致性、表型、特点和/或特性(例如,遍及基因组的染色质可及性的状态)的更准确的预测性评估,不需要单独评估诸如RNA或蛋白质表达水平的其他参数。

[0128] 在一些方面,确定或评估表观遗传学/表观基因组学特性、状态和/或谱(例如使用诸如ATAC-seq的测定法)涉及以下一个或多个步骤:(a)染色质分离;(b)标签化;(c)测序;(d)序列映射;和/或(e)峰确定。在一些方面,确定或评估表观遗传学/表观基因组学特性、状态和/或谱(例如使用诸如ATAC-seq的测定法)涉及以下一个或多个步骤:(1)从细胞或细胞群分离染色质,(2)用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,(3)对所有或部分带有标签的片段测序以产生多个序列读段;(4)比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和/或(5)确定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

[0129] 在一些实施方案中,用于评估染色质可及性的测定法(例如ATAC-seq)涉及以下一个或多个步骤:(i)洗涤和裂解细胞;(ii)标签化;(iii)DNA纯化;(iv)PCR预扩增;(v)定量PCR(qPCR)扩增;(vi)PCR n-扩增;(vii)DNA纯化;(viii)文库生成和(ix)测序。在一些实施方案中,还进行质量控制的测定法或步骤。用于文库生成和/或质量控制的示例性步骤包括以下中的一个或多个:(i)基因组DNA分离和评估;(ii)大小选择以去除残留引物(例如,使用2%琼脂糖);(iii)DNA纯化;(iv)qPCR;(v)DNA定量;和(vi)文库的稀释和汇集。在一些实施方案中,使用高通量或下一代测序对从样品生成的文库进行测序。在一些实施方案中,使用一系列细胞稀释液例如含有不同细胞浓度或细胞数的样品来测试本文中提供的方法所

需的细胞数。

[0130] 在一些实施方案中,当执行所述方法的一个或多个步骤时,可以使用其他参数或度量来确定样品和数据的质量。在一些实施方案中,用于确定样品和数据的质量的参数或度量包括每样品映射读段数、与受试者基因组(例如,人类基因组)比对的百分比、非冗余读段的百分比、独特读段数、已知峰(阳性对照)的回收、评估适当标签化的qPCR扩增、基因组DNA的核小体带评估和/或重复或不同稀释度之间的相关性。

[0131] 在一些实施方案中,用于评估染色质可及性的测定法(例如ATAC-seq)包括:用DNA插入元件(例如插入酶复合物)处理从细胞群分离的染色质,以产生带有标签的基因组DNA片段。在这个步骤中,使用诸如Tn5或MuA的插入酶将染色质标签化(即,在同一个反应中切割并且加标签),所述插入酶切割染色质开放区域中的基因组DNA,并且向片段的两端添加衔接子。用于将分离的基因组DNA标签化的方法是本领域已知的(参见,例如,Caruccio *Methods Mol. Biol.* 2011 733:241-55; Kaper等人, *Proc. Natl. Acad. Sci.* 2013 110:5552-7; Marine等人, *Appl. Environ. Microbiol.* 2011 77:8071-9和US20100120098)并且可从例如Illumina (San Diego, Calif.) 商购获得。这样的系统可以容易地适用于本文。在一些情况下,可以调整条件以获得期望的染色质中的插入水平(例如,在开放区域中平均每50至200个碱基对发生插入)。

[0132] 可通过任何适合的方法制作在测定法中使用的染色质,例如可含有来自细胞(例如获自受试者的细胞)的基因组DNA、组蛋白和/或其他染色质相关因子的染色质。在一些实施方案中,可以分离、裂解细胞核,并且可以例如从核被膜进一步纯化染色质。在一些实施方案中,可以通过使分离的细胞核与反应缓冲液接触来分离染色质。在一些实施方案中,分离的细胞核可以在其与反应缓冲液(其包含插入酶复合物和其他必需的试剂)接触时裂解,这允许插入酶复合物进入染色质。在一些实施方案中,所述测定法包括从细胞群分离细胞核;将分离的细胞核与转座酶和衔接子合并,其中所述合并导致细胞核的裂解而释放所述染色质并产生带有衔接子标签的基因组DNA片段。

[0133] 在将染色质片段化并且加标签以产生带有标签的基因组DNA片段之后,对至少一些带有衔接子标签的片段进行测序以产生多个序列读段。可以使用任何已知的测序方法例如下一代或高通量测序方法将片段测序。例如,可以使用Illumina的可逆终止子法、Roche的焦磷酸测序法(454)、Life Technologies的边连接边测序(SOLiD平台)或Life Technologies的Ion Torrent平台将片段测序。这样的方法的实例描述于例如Margulies等人, *Nature* 2005437:376-80; Ronaghi等人, *Analytical Biochemistry* 1996 242:84-9; Shendure等人, *Science* 2005 309:1728-32; Imelfort等人, *Brief Bioinform.* 200910:609-18; Fox等人, *Methods Mol Biol.* 2009;553:79-108; Appleby等人, *Methods Mol Biol.* 2009;513:19-39和Morozova等人, *Genomics.* 200892:255-64,将以上文献通过提述并入本文。在扩增步骤期间,可以将与选定的下一代测序平台相容的正向和反向测序引物位点添加到片段的末端。在一些实施方案中,可以使用PCR引物扩增片段,所述PCR引物与已经添加到所述片段的标签杂交,其中用于PCR的引物具有与特定测序平台相容的5'尾。在一些情况下,所使用的引物可以含有分子条形码(“索引”),以便在测序之前可以将不同的池汇集在一起,并且可以使用条形码序列将序列读段追踪到特定样品。

[0134] 在一些方面,所述测定法包括确定核酸在位点的可及性,其中所述核酸来自细胞

样品,所述测定法包括:将具有插入酶的多个分子标签插入到核酸中并且使用所述分子标签确定所述位点的可及性。细胞样品可以来自主要来源,例如来自受试者的细胞。在一些实施方案中,细胞样品包括来自受试者的细胞,所述细胞经过选择和/或工程改造或修饰,例如,经过工程改造以表达重组受体。细胞样品可以由单细胞组成。细胞样品可以由有限数量的细胞组成(例如小于约500,000个细胞)。

[0135] 在一些实施方案中,所述测定法进一步包括确定可及性,以鉴定在特定基因座处与核酸和/或染色质结合的一种或多种蛋白质。在一些实施方案中,至少一种蛋白质是转录因子。另外,所述测定法可包括使用分子标签来生成核酸的可及性图谱。

[0136] 在插入分子标签的过程中,可以使核酸片段化成多个片段。在一些实施方案中,可以扩增所述片段。在一些实施方案中,可以将片段测序以生成多个测序读段。这可用于确定核酸在任何给定位点的可及性。可以使用高通量测序技术将片段测序。在一些实施方案中,可以基于插入酶的序列插入偏好将测序读段归一化。可以使用经过测序的读段的长度来确定染色质状态注释。

[0137] 核酸可以与多个缔合分子结合。缔合分子可以是例如蛋白质、核酸或糖类。在一些实施方案中,缔合分子可包含组蛋白。在其他情况下,缔合分子可包含适体。

[0138] 插入酶可以是能够将核酸序列插入到核酸中的任何酶。在一些实施方案中,插入酶可以基本上非序列依赖性的方式将核酸序列插入到核酸中。插入酶可以是原核的或真核的。插入酶的实例包括但不限于转座酶、HERMES和HIV整合酶。转座酶可以是Tn转座酶(例如Tn3、Tn5、Tn7、Tn10、Tn552、Tn903)、MuA转座酶、Vibhar转座酶(例如来自哈维氏弧菌)、Ac-Ds、Ascot-1、Bs1、Cin4、Copia、En/Spm、F元件、hobo、Hsmar1、Hsmar2、IN(HIV)、IS1、IS2、IS3、IS4、IS5、IS6、IS10、IS21、IS30、IS50、IS51、IS150、IS256、IS407、IS427、IS630、IS903、IS911、IS982、IS1031、ISL2、L1、Mariner、P元件、Tam3、Tc1、Tc3、Tel、THE-1、Tn/O、TnA、Tn3、Tn5、Tn7、Tn10、Tn552、Tn903、To11、To12、Tn10、Ty1、任何原核转座酶,或与以上列出的那些转座酶有关和/或从其衍生的任何转座酶。在一些实施方案中,转座酶是Tn5转座酶或其衍生物。

[0139] 在一些情况下,与亲本转座酶相关和/或从其衍生的转座酶可包含与亲本转座酶的相应肽片段具有至少约50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%氨基酸序列同源性的肽片段。肽片段的长度可以是至少约10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、400、或500个氨基酸。例如,衍生自Tn5的转座酶可包含具有50个氨基酸长度并且与亲本Tn5转座酶中的相应片段80%同源的肽片段。在一些实施方案中,可以通过添加一种或多种阳离子来促进和/或触发插入。阳离子可以是二价阳离子,例如像Ca²⁺、Mg²⁺和Mn²⁺。

[0140] 分子标签可包括测序衔接子、锁核酸(LNA)、拉链核酸(ZNA)、RNA、亲和反应性分子(例如生物素、地高辛)、自身互补分子、硫代磷酸酯修饰、叠氮基或炔基。在一些实施方案中,测序衔接子可进一步包含条形码标记。此外,条形码标记可包括独特序列。独特序列可用于鉴定各个插入事件。任何标签可进一步包含荧光标签(例如荧光素、罗丹明、Cy3、Cy5、噻唑橙)。

[0141] 另外,插入酶可进一步包含亲和标签。在一些实施方案中,亲和标签可以是抗体。抗体可以与例如转录因子、修饰的核小体或修饰的核酸结合。修饰的核酸的实例包括但不

限于甲基化或羟甲基化DNA。在其他情况下,亲和标签可以是单链核酸(例如ssDNA、ssRNA)。在一些实例中,单链核酸可以与靶核酸结合。在其他情况下,插入酶可以进一步包含核定位信号。

[0142] 在一些实施方案中,可以使细胞(例如衍生自受试者的细胞)透化以允许插入酶进入。可以最小程度扰动细胞样品中的细胞核的方式进行透化。在一些情况下,可以使用透化剂使细胞样品透化。透化剂的实例包括但不限于NP40、洋地黄皂苷、吐温、链球菌溶血素和阳离子脂质。在其他情况下,可以使用低渗休克和/或超声波处理使细胞样品透化。在其他情况下,插入酶可以是高度带电的,这可以使其透化穿过细胞膜。

[0143] 在一些实施方案中,所述方法包括比对、映射和/或分析从测定法生成的大规模数据(例如高通量测序数据)的步骤。在一些实施方案中,所生成的数据的分析包括比对、固定读段配对、去除PCR重复、过滤到映射的读段和质量读段和/或过滤出线粒体序列和/或峰调用步骤中的一个或多个;并且可以进一步包括进一步的分析或应用步骤,包括核小体定位、转座子插入位点、基因组浏览器可视化、差异可及性分析、基序富集、基因本体论(GO)富集和/或确定转录因子占据。在一些实施方案中,其他处理步骤包括解复用原始数据、对齐和过滤,以及评估质量度量。

[0144] 在一些方面,所述方法包括比对、过滤和将序列读段映射到基因组的基因组区域的步骤。在一些方面,所述方法包括确定每个细胞或细胞群的多个基因组区域中的序列读段的峰的步骤。在一些实施方案中,本文中提供的任何方法进一步包括用于比对、过滤和/或映射序列结果的一个或多个步骤、功能、过程或脚本,所述序列结果从本文中提供的方法的一个或多个步骤、功能、过程或脚本获得(例如,从表观遗传学分析获得的序列)。在一些实施方案中,所述方法包括计算执行的步骤、功能、过程或脚本,例如,使用一个或多个计算机程序和/或通过使用计算算法来执行。在一些实施方案中,还提供了用于进行或执行本文中提供的方法的一个或多个步骤的计算机系统、计算机可读指令、软件、系统和/或装置。在一些实施方案中,可以计算执行任何进一步的分析或应用步骤,比如本文所述的任何步骤,例如,使用一个或多个计算机程序和/或通过使用计算算法来执行。

[0145] 在一些实施方案中,所述方法涉及例如使用ATAC-seq鉴定、比对、过滤、处理、映射和/或分析从本文中提供的方法的一个或多个步骤获得的序列,例如从染色质可及性分析获得的序列数据。在一些实施方案中,序列的鉴定、比对、过滤、处理、映射和/或分析包括用于获得信号峰的序列操作和比对程序和/或用于执行进一步分析和/或应用的程序,例如,鉴定本文所述方法中的一个或多个基因组区域。在一些实施方案中,序列的鉴定、比对、过滤、处理、映射和/或分析和/或执行进一步分析和/或应用包括以任何顺序依次或同时进行本文所述的示例性步骤、功能、过程或脚本中的任何一者或多者。在一些方面,可以使用计算脚本、工具和/过程来执行用于序列的鉴定、比对、过滤、处理、映射和/或分析和/或用于执行进一步分析和/或应用的任何步骤和/或程序,并且可以例如通过连接各个计算步骤、工具和/或过程而形成分析流水线,例如,一系列连接的步骤和/或程序或其集合。在一些方面,步骤、功能、过程或脚本中的一者或多者可以被实现或执行相似功能的相似算法、步骤、功能、过程或脚本取代。在一些实施方案中,一个或多个步骤的顺序次序可以是任何次序,或者可以并行执行任何一个或多个步骤、功能、过程或脚本。

[0146] 在一些实施方案中,序列(例如,从ATAC-seq获得)的鉴定、比对、过滤、处理、映射

和/或分析的方法由一组例如处于YAML(一种人可读数据序列化语言)格式的配置文件驱动。在一些实施方案中,配置文件指定了一般流水线运行参数,例如,在处理中使用的CPU核心的数目和上载输出文件的位置,以及每个样品的信息,例如样品名称和样品中测序读段的类型。在一些实施方案中,用于序列的鉴定、比对、过滤、处理、映射和/或分析的步骤和/或方法根据所提供的配置推断出需要生成的文件、文件的顺序以及处理脚本和工具。在一些实施方案中,可以将这些步骤、功能、过程或脚本中的每一者添加到队列,所述队列被分发到一组计算节点。在一些实施方案中,在运行期间,输入和输出可以被组织成公共文件夹和文件命名结构,并且可以通过记录在每次运行期间的事件而生成日志。在一些实施方案中,分析方法包括在设置或命令中的步骤、功能、过程或脚本。

[0147] 在一些实施方案中,例如根据ATAC-seq数据确定的表观遗传学和/或表观基因组学特性、状态和/或谱的分析方法由例如人可读数据序列化语言的一组配置文件驱动。在一些实施方案中,配置文件指定了一般流水线运行参数,例如,在处理中使用的CPU核心的数目和上载输出文件的位置,以及每个样品的信息,例如样品名称和样品中测序读段的类型。在一些实施方案中,流水线根据所提供的配置推断出需要生成的文件、文件的顺序以及处理脚本和工具。在一些实施方案中,可以将这些步骤、功能、过程或脚本中的每一者添加到队列,所述队列被分发到一组计算节点。在一些实施方案中,在流水线运行期间,输入和输出可以被组织成公共文件夹和文件命名结构,并且可以通过记录在每次运行期间的事件而生成日志。在一些实施方案中,分析方法包括在设置或命令中的步骤、功能、过程或脚本。

[0148] 在某些情况下,步骤、功能、过程或脚本是分支的,并且以非线性方式连接。在一些实施方案中,步骤、功能、过程或脚本包括涉及下一代测序结果的一般计算处理或一般分析的步骤(例如将压缩文件解压缩)或针对所使用的特定数据生成平台(例如,特定的下一代测序平台)的具体步骤。在一些实施方案中,将配置文件读入流水线并生成有向无环图(DAG),其包含处理步骤、功能、过程或脚本。

[0149] 分析设置的示例性方法包括下文描述的步骤。在一些实施方案中,所述步骤、功能、过程或脚本包括所述的步骤、功能、过程或脚本中的任何一者或多者。

[0150] 在一些实施方案中,用于序列的鉴定、比对、过滤、处理、映射和/或分析或进一步分析和/或应用的步骤和/或方法中的步骤、功能、过程或脚本中的一者或多者是自动化的。在一些实施方案中,可以使用脚本(例如,Perl脚本或shell脚本)来调用本文所述的各种示例性步骤、功能、过程或脚本中的任一者(参见,例如,Tisdall,Mastering Perl for Bioinformatics, O'Reilly&Associates, Inc., Sebastopol, Calif 2003; Michael, R., Mastering Unix Shell Scripting, Wiley Publishing, Inc., Indianapolis, Ind. 2003)。在一些实施方案中,分析方法可以全部或部分地体现在一个或多个专用程序中,例如,各自任选地用诸如C++的编译语言编写,然后以二进制形式编译和分发。在一些实施方案中,可以全部或部分地以现有序列分析平台内的模块的形式或通过调用所述平台内的功能性来实施分析方法。在一些实施方案中,分析方法包括全部响应于单个起始队列而被自动调用的多个步骤、功能、过程或脚本(例如,源自人类活动、另一个计算机程序或机器的触发事件的一个或组合)。

[0151] 在一些实施方案中,分析方法中的步骤、功能、过程或脚本或步骤、功能、过程或脚本的任何组合可以响应于队列而自动发生。可以计算机文件的格式提供输出。在一些实施

方案中,输出是FASTA文件、VCF文件、文本文件、.bedGraph文件或含有序列数据(比如与参考基因组的序列比对的核酸序列)的XML文件。

[0152] 在一些实施方案中,可计算分析序列读段以鉴定片段的末端(从其可推断转座子切割位点)。在一些实施方案中,片段的一端可由测序读段开始处的序列定义,并且片段的另一端可由第二测序读段开始处的序列定义,其中通过配对末端测序(例如,使用Illumina的测序平台)获得第一和第二测序读段,例如产生配对末端读段R1和R2。通过检查较长序列读段的开始和结束(例如,具有两个衔接子的序列;一个在一端,另一个在另一端)可以获得相同的信息。在一些实施方案中,单个序列读段可含有两个衔接子序列,在这种情况下,可以从单个序列读段推断出片段的两个末端(其对应于两个单独的转座酶的两个切割位点)。通过例如将片段末端映射到感兴趣的区域的核苷酸序列上,并且对那些位置之间的碱基对的数目进行计数,可以计算出片段的长度。可以利用在序列读段开始和/或结束处的核苷酸序列获得所使用的信息。

[0153] 在一些实施方案中,对应于基于大规模测序例如高通量测序或下一代测序的表观遗传学或表观基因组学数据的序列信息,例如来自ATAC-seq的结果,其可包含正向(R1)和/或反向(R2)读段,可含有对于反应中的每个样品的约或至少约 10^6 、 10^7 、 10^8 、 2×10^8 、 3×10^8 、 4×10^8 、 5×10^8 、 10^9 、 10^{10} 或更多个靶多核苷酸读段或簇,例如,可包含约、少于约或多于约 2×10^6 、 3×10^6 、 4×10^6 、 5×10^6 、 10^7 、 10^8 或更多个靶多核苷酸或簇。

[0154] 在一些实施方案中,可以检索来自高通量序列的原始数据,该原始数据可以呈压缩形式,比如来自任何存储位置,例如基于云的存储器,或者可以将其下载到计算集群上。用于检索数据的示例性脚本或工具包括例如get_bcl,例如呈.bcl文件形式的碱基调用。在一些实施方案中,来自定序器运行的数据可以被解压缩,以便使用下游工具例如使用脚本untar_bcls进行处理。输出测序数据可以是各种输出数据类型或格式中的任何一种,包括但不限于*.bcl、*.fasta、*.csfasta、*.seq.txt、*.qseq.txt、*.fastq、*.sff、*.prb.txt、*.sms、*.srs和/或*.qv。在一些实施方案中,一个或多个处理步骤、功能、过程或脚本包括将测序数据输出的格式转换成可以在分析方法的进一步步骤中使用或用于显示的格式,例如,特别是用于图形用户界面(GUI)中。例如,一个名为bcl2fastq的程序可用于将.bcl原始碱基调用文件转换为压缩的FASTQ文件。在涉及配对末端运行的实施方案中,每个样品具有两个文件R1和R2,分别对应于正向和反向读段(每个DNA片段的开始和结束)。

[0155] 在一些实施方案中,针对下游步骤将原始测序计数归一化。在一些实施方案中,归一化包括估计尺寸因子和离差,以及执行负二项广义线性模型拟合。在一些实施方案中,可以使用脚本或工具(比如DESeq2)来评估这样的步骤。在一些实施方案中,归一化还包括基于测量值的归一化,所述测量值比如峰中读段的分数(FRiP;显示信号的富集,计算为(峰中读段数)/(总读段数))。在一些实施方案中,可以确定FRiP归一化的计数(例如,变量betaPrior设置为真)。

[0156] 在一些实施方案中,处理步骤检索由任何步骤生成的压缩文件并且将其解压缩。在一些实施方案中,检索压缩文件(例如,压缩的FASTQ文件)并将文件移动到按样品分类的分析目录中的示例性步骤包括被称为get_data的程序。在一些实施方案中,将文件解压缩(例如将压缩的FASTQ文件解压缩)使得它们可以由下游工具处理的示例性步骤包括被称为解压缩(unzip)的程序。

[0157] 在一些方面,可以使用脚本或工具(例如,fastqc)生成指示序列质量的统计数据,例如,测序运行、碱基调用质量、污染、过度聚类的整体统计数据。

[0158] 在一些实施方案中,对于多个簇的一个或多个测序读段中鉴定的序列,将其在位置上映射到参考序列,例如参考基因组。通常,映射涉及将一个序列沿着另一个序列放置,沿着每个序列迭代地引入空位,对两个序列匹配的程度进行评分,并且在一些方面,在沿着参考序列的各个位置上进行重复。最佳评分匹配被认为是映射并且代表关于序列之间的关系的程度的推断。在一些实施方案中,测序读段与之相比的参考序列是参考基因组,比如与受试者相同物种的成员的基因组。参考基因组可以是完全或不完全的。在一些实施方案中,参考基因组仅含有含靶多核苷酸的区域。在一些实施方案中,参考序列包含人基因组或由其组成。在一些实施方案中,参考序列包含除受试者之外或从中取得样品的一种或多种生物的多核苷酸的序列或由其组成。在一些实施方案中,参考序列包含多个已知序列或由其组成,所述已知序列比如用于扩增靶多核苷酸序列的所有探针序列(例如,每个不同的靶多核苷酸的每个序列B和/或序列B')。由一个引物的延伸生成的测序数据(例如来自正向引物的R1序列)可以被映射到与由另一个引物的延伸生成的测序数据(例如来自反向引物的R2序列)相同或不同的参考序列。由一个引物的延伸生成的测序数据可以两次或更多次被映射到参考序列,其中每次映射使用不同的映射算法。R1序列的映射可以独立于R2序列。

[0159] 在一些实施方案中,对来自测序读段的读段的位置映射的示例性步骤包括例如被称为get_genome、build_bowtie2_index、Index_genome_fasta和map_atac_reads的步骤、功能、过程或脚本。位置映射的示例性步骤例如包括,取得指定生物的适当基因组文件和来自其在线来源的版本(比如get_genome)的步骤;使用bowtie2_build和/或build_bowtie2_index索引基因组文件使得读段在位置上被映射到基因组的步骤;为下载的基因组文件建立FASTA索引以允许由下游工具(比如Index_genome_fasta)进行位置访问的步骤;以及使用bowtie2和/或map_atac_reads将ATAC-seq序列读段映射到基因组以便确定其位置的步骤。

[0160] 在一些实施方案中,基因组参考序列是人基因组。在一些实施方案中,参考基因组序列是NCBI36/hg18序列。或者,参考基因组序列是GRCh37/hg19。公共序列信息的其他来源包括GenBank、dbEST、dbSTS、EMBL(欧洲分子生物学实验室)和DDBJ(日本DNA数据库)。许多计算机算法可用于比对序列,包括但不限于BLAST(Altschul等人,1990)、BLITZ(MPsrch)(Sturrock&Collins,1993)、FASTA(Person&Lipman,1988)或BOWTIE(Langmead等人,(2009) Genome Biology 10:R25.1-R25.10)。

[0161] 映射或比对程序的其他实例包括:BLAT,其来自Kent Informatics(Santa Cruz, Calif.)(Kent,W.J.,Genome Research 4:656-664(2002));SOAP2,其来自Beijing Genomics Institute(Beijing,Conn.)或BGI Americas Corporation(Cambridge,Mass.);Bowtie(Langmead等人,Genome Biology,10:R25(2009));核苷酸数据库的有效大规模比对(Efficient Large-Scale Alignment of Nucleotide Databases)(ELAND)或序列和变异的一致性评估(Consensus Assessment of Sequence and Variation)(CASAVA)软件的ELANDv2组件(Illumina,San Diego,Calif.);RTG Investigator,其来自Real Time Genomics,Inc.(San Francisco,Calif.);Novoalign,其来自Novocraft(Selangor,Malaysia);Exonerate,European Bioinformatics Institute(Hinxton,UK)(Slater,G.和

Birney, E., *BMC Bioinformatics* 6:31 (2005)), Clustal Omega, 其来自 University College Dublin (Dublin, Ireland) (Sievers F. 等人, *Mol Syst Biol* 7, article 539 (2011)); ClustalW 或 ClustalX, 其来自 University College Dublin (Dublin, Ireland) (Larkin M.A. 等人, *Bioinformatics*, 23, 2947-2948 (2007)); 和 FASTA, European Bioinformatics Institute (Hinxton, UK) (Pearson W.R. 等人, *PNAS* 85 (8):2444-8 (1988); Lipman, D.J., *Science* 227 (4693):1435-41 (1985))。在一些实施方案中, 通过使用 ELAND 软件的 Illumina 基因组分析仪的生物信息学比对分析来处理序列的克隆扩增拷贝的一端。在一些方面, 分析可包括基因组浏览器可视化的步骤。

[0162] 在一些实施方案中, 生成含有序列数据的输出文件, 所述序列数据比如与参考基因组的序列比对的核酸序列。在其他实施方案中, 输出包含描述主题核酸中相对于参考基因组的一个或多个突变和/或其他表观遗传学或表观基因组学数据比对的坐标或字符串。已知的比对字符串包括简单的无空位比对报告 (Simple UnGapped Alignment Report) (SUGAR)、详细有用的标记的有空位比对报告 (Verbose Useful Labeled Gapped Alignment Report) (VULGAR) 和紧凑异质的有空位比对报告 (Compact Idiosyncratic Gapped Alignment Report) (CIGAR) (Ning, Z. 等人, *Genome Research* 11 (10):1725-9 (2001))。在一些实施方案中, 输出是序列比对, 例如序列比对图谱 (SAM) 或二进制比对图谱 (BAM) 文件, 其包含 CIGAR 字符串 (SAM 格式描述于例如 Li 等人, *The Sequence Alignment/Map format and SAMtools*, *Bioinformatics*, 2009, 25 (16):2078-9)。在一些实施方案中, CIGAR 显示或包括每行一个的空位比对。CIGAR 是一种压缩的成对比对格式, 其被报告为 CIGAR 字符串。

[0163] 在一些实施方案中, 可以使用下游工具将比对的序列文件转换为不同的文件格式用于分析。在一些实施方案中, 这样的转换步骤包括将 BAM 比对文件转换为下游 HOMER 工具可使用 (例如使用 `make_homer_tagdir`) 的文件格式的步骤。

[0164] 在一些实施方案中, 来自簇的 R1 序列包含来自多个不同的靶多核苷酸的正向序列, 并且来自簇的 R2 序列包含反向序列。当每个反向序列被选择为靶向特定的靶多核苷酸时, 其在参考序列 (例如参考基因组) 内的序列和位置通常是已知的, 并且可以预期来自同一个簇的 R1 序列落入预计的核苷酸距离内。预计的核苷酸距离可以基于包含片段化样品多核苷酸的样品的平均或中值片段长度, 或基于这样的中值或平均片段长度代表不太可能的片段长度的上阈值距离。因此, 在一些实施方案中, 与距离来自同一个簇的 R2 序列的阈值距离更远的位置对准的 R1 序列可能是错误的并且被丢弃。在一些实施方案中, 沿着在来自同一个簇的比对的 R1 和 R2 序列之间的参考序列的上阈值距离 (高于它的针对簇的序列读段被丢弃) 约为或大于约 1,000、2,500、5,000、7,500、10,000、12,500、15,000、20,000 个或更多个碱基对。在一些实施方案中, 丢弃 R1 序列与参考序列 (例如参考基因组) 的非独特区域的比对, 并且将所述序列与参考序列内的较小子集的独特序列重新比对。

[0165] 在一些实施方案中, 测序读段可以是重复的, 并且可以在初始位置映射和/或比对后去除重复序列。当映射测序读段时, 可以通过比对算法将重复读段标记为重复。例如, 比对算法内的标记重复子程序检查比对序列文件 (例如, *.BAM 文件) 中的所有记录, 并判定哪些读段是其他读段的重复。在一些实施方案中, 可以存在两种一般类型的重复中的一种或多种: 光学重复, 其通常可由初级分析软件中的缺陷引起, 和 PCR 重复, 其可由重复 PCR 反应

引起。

[0166] 在一些实施方案中,分析方法涉及去除重复读段,例如以提高定量准确性。在分析方法的一些实施方案中,步骤、功能、过程或脚本用于去除重复序列,例如标记和去除由PCR扩增偏倚和集群错误调用引起的重复读段。在一些实施方案中,这样的步骤可以减少数据集中的噪声量。在一些实施方案中,所述步骤可包括用于将映射后的读段的命名和标引正确排序的算法或程序,并输出给出每染色体的映射计数的idxstats文件。在一些实施方案中,去除重复和进一步标引的示例性步骤包括,例如使用被称为Picard_remove_duplicates的脚本或工具,标记和去除由PCR扩增偏倚和集群错误调用引起的重复读段以减少数据集中的噪声量的步骤。在一些方面,可以使用Picard或相似的脚本或工具,例如使用被称为atac_insert_size_metrics的脚本或工具来计算和可视化恢复的片段大小的分布,并且例如使用被称为atac_alignment_summary的脚本或工具来计算关于读段在各种类型的基因组特征中的位置分布的统计数据。

[0167] 在一些实施方案中,分析涉及去除线粒体读段和/或另外的污染序列,其根据是读段的序列同一性、质量、映射位置或其他测序特性。在一些实施方案中,分析方法涉及去除映射至线粒体基因组的读段、线粒体读段和/或另外的污染序列。在一些方面,线粒体读段是ATAC-seq文库中的主要污染物,其可贡献大部分例如高达80%的序列读段,并且可以降低下游步骤的准确性。在一些实施方案中,去除的根据是测序读段的序列同一性、质量、映射位置或其他测序特性。在一些实施方案中,用于去除线粒体读段和/或另外的污染序列的示例性步骤包括filter_mtDNA_reads的步骤。在一些实施方案中,分析方法涉及保留和比较线粒体读段。

[0168] 在一些实施方案中,分析涉及将序列读段分离成代表特定表观遗传学特性(例如染色质可及性或染色质占据)的子集。在一些实施方案中,测序片段的大小可用于确定其代表所述特性的程度或水平。在一些实施方案中,一个或多个基因组区域、基因座、元件或区间内的读段数目和/或归一化读段数目可用于确定其代表所述特性的程度或水平。在一些实施方案中,一个或多个基因组区域、基因座、元件或区间中的信号的特定数量或质量的差异可用于确定其代表所述特性的程度或水平。

[0169] 在一些实施方案中,分析涉及在特定的链方向上将映射读段的固定数目的位置移位例如约3、4或5个碱基对,以便基于转座酶例如Tn5转座酶的性质导致序列读段的移位,并且涉及文库制备。在一些实施方案中,用于移位读段位置的示例性步骤包括shift_atac_alignments。

[0170] 在一些实施方案中,分析方法涉及评估或测量ATAC-seq片段的序列读段。将DNA的可及性程度与片段的更多序列读段相关联,使得来自序列读段的信号形成例如通过使用峰调用工具、脚本或算法可检测或测量的峰。在一些实施方案中,所提供的方法包括评估、描绘或确定信号峰的步骤,所述信号峰代表可及的基因组的一个或多个区域中的DNA。在一些方面,序列读段的峰信号是具有富集信号、高于背景的信号或与周围区域相比更高信号的区域。在一些实施方案中,ATAC-seq峰可以与基因组基因座区域图谱或基因组图谱重叠。正如通过量化ATAC-seq片段测量的富集或耗竭可及性和/或占据信号(例如峰信号)的区域的身份可以根据其在参考基因组上映射的位置加以确定。通过整合进一步的基因组和表观基因组学数据,比如关于组蛋白修饰的信息或活性转录的证据,可以将这些区域进一步分类

为不同的调控元件类型-启动子、增强子、隔离子等 (Buenrostro等人 (2013) Nature Methods, 10:1213-1218)。

[0171] 在一些实施方案中,可及的区域,例如ATAC-seq片段的大小可以为约5至约20,000bp,比如大小约10至约10,000bp、约50至约5,000bp、约100至约1,000bp、约200至约900bp、约300至约800bp、约400至约700bp、约500至约600bp、约10至约500bp、约20至约400bp、约30至约300bp、约40至约200bp、约40至约100bp、约50至约100bp、约60至约100bp、约70至约100bp、约80至约100bp、约90至约100bp、约100至约500bp、约200至约500bp、约300至约500bp、约400至约500bp、约100至约400bp、约100至约300bp或约100至约200bp。

[0172] 在一些实施方案中,所述方法包括例如通过使用峰调用工具、脚本或算法可以找到、确定和/或注释信号峰的步骤、功能、过程或脚本。在一些实施方案中,峰调用可在特定的样品上进行和/或存在于样品的一个或多个生物学重复中和/或存在于一个或多个不同的样品中。在一些实施方案中,峰调用允许区分信号和噪声。在一些方面,峰调用可包括一个或多个步骤,比如估计片段长度和调整读段位置、鉴定局部噪声、鉴定富集区域或峰区域和/或估计错误发现率(FDR)。在一些方面,被调用的峰可以是宽的。在一些方面,被调用的峰可以是窄的。可得到并且可以使用各种峰调用脚本中的任何一种。在一些实施方案中,示例性峰调用脚本、算法或软件包括MACS、MACS2、Epic、SICER、BayesPEak、homer_findPeak、Jmosaics、T-PIC、EDD、GEM或SPP。在一些方面,使用MACS或MACS2执行峰值调用。在一些实施方案中,示例性峰调用步骤、功能、过程或脚本可以包括用于调用可及性峰的基于模型的ChIP-Seq (MACS) 或MACS2分析,比如本文所述的macs_callpeaks步骤。在一些实施方案中,MACS或MACS2峰调用使用样品交换策略通过调用对照上的样品峰和样品上的对照峰这两者来估计错误发现率(FDR)。在一些实施方案中,这样的峰调用步骤可以使用总文库大小针对数据集中的较小者标度数据集中的较大者,以进行归一化。在一些实施方案中,峰包括通过量化ATAC-seq片段测量的富集或耗竭可及性和/或占据信号的基因组区域。在一些实施方案中,峰区域的大小可以在10和10,000bp之间。样品之间的峰的比较可以在下游加以使用以鉴定条件之间的活性调控元件(例如,与特定基因或转录因子相关),和/或将其用于鉴定特定细胞状态或预测细胞疗法的结果、细胞疗法的效力、细胞组合物或培养物的毒性和/或其他特点的标记。

[0173] 其他示例性峰调用步骤、功能、过程或脚本可以包括使用HOMER找到峰的步骤,在一些实施方案中包括标识为homer_findPeaks的脚本和/或工具。在一些实施方案中,可以采用例如使用被称为homer_find_motifs的脚本或工具搜索基序(例如,在通过MACS和HOMER发现的峰集内共享的富集的转录因子结合基序)的步骤。

[0174] 在峰调用之后,可以使用其他步骤、功能、过程或脚本,比如峰注释步骤,比如利用包括可能与其相关的最近基因的附加元数据来注释峰,例如使用homer_annotate_peaks或其他注释策略,以便于进一步分析。在一些实施方案中,用于注释的其他步骤包括检索与生物体相关的转录组注释文件和流水线正在运行的基因组版本(例如,使用get_gtf_annotation)和/或检索注释的子集,例如仅包括蛋白质编码转录物的注释转录组(例如使用gtf_coding_transcripts_only)。

[0175] 在一些方面,所述方法包括从存在于多个生物学重复中和/或两个或更多个样品中的共同或重叠峰生成一个或多个共有峰集,例如共有峰可及性。在一些实施方案中,用于

从两个或更多个样品中的共同或重叠峰生成共有峰的示例性峰调用脚本、算法或软件是 DiffBind。

[0176] 在一些实施方案中,所提供的方法和测定法包括制作细胞基因组区域的表观遗传学图谱(也称为表观遗传学谱)。在一些方面,表观遗传学图谱描绘了整个基因组、基因组的一部分或一个或多个特定基因附近或周围或内部的跨越多个不同区域例如编码序列、基因间间隔区、调节区(例如启动子)等的表观遗传学状态。可以通过将获自序列读段的信息映射到该区域来完成这个步骤。在一些实施方案中,计算分析序列读段以产生多个数字输出,将其映射到感兴趣区域的表示(例如,图形表示)。用于映射的示例性信息包括但不限于:(i) 转座酶的切割位点;(ii) 所产生的片段的大小;(iii) 片段长度;(iii) 具有限定长度范围的序列读段的位置;和(iv) 序列读段丰度,例如峰信号。

[0177] 在一些实施方案中,可以计算序列读段丰度,即在序列读段中代表基因组区域中的特定序列的次数。在某些情况下,可以生成例如使用峰调用工具确定的描绘序列读段的峰信号的表观遗传学谱或图谱。得到的表观遗传学图谱可以提供感兴趣区域的染色质分析。例如,根据映射的信息,所述图谱可以显示以下一项或多项:沿着所述区域的染色质可及性;DNA结合蛋白(例如,转录因子)对所述区域中的位点的占据;所述区域中的无核小体DNA;沿着所述区域的核小体定位;和沿着所述区域的染色质状态。在一些实施方案中,所述测定法进一步涉及通过例如一种DNA结合蛋白在该蛋白质与之结合的多个位点上的聚集数据来测量DNA结合蛋白的结合位点的全局占据。在一些情况下,所述图谱也可以用序列信息和关于序列的信息(例如,启动子、内含子、外显子、已知的增强子、转录起始位点、非翻译区、终止子等的位置)加以注释,以便在注释的背景下可以查看表观遗传学信息。在一些实施方案中,可以使用计算实施的脚本或工具来生成表观遗传学/表观基因组学图谱。示例性步骤包括创建全基因组片段计数并且在基因组浏览器上将计数可视化。可以利用的示例性脚本或工具包括make_homer_ucsc_file,其可以创建允许全基因组的片段计数堆积的.bedGraph文件;和homer_bedgraph_to_bigwig,其可以将bedGraph文件转换为被大多数基因组浏览器使用的将跨基因组的片段覆盖可视化的二进制压缩的bigWig文件。

[0178] 在一些方面,分析包括生成与基因的特定元件相关的度量。在一些方面,这样的度量包括在注释基因的启动子上、或在注释基因的编码区上的可及性。在一些方面,分析包括生成诸如每个基因的启动子区的归一化的可及性计数度量(启动子可及性)之类的度量。在一些实施方案中,度量的注释和生成可用于进一步的下游分析,例如,比较表观遗传学谱、聚类和/或生物途径分析。在一些实施方案中,可以使用脚本或工具计算实施这样的步骤,例如,用于注释基因和/或调控元件。在一些实施方案中,可以使用诸如ChIPpeakAnno或Homer的工具。在一些实施方案中,可以定义调控元件,比如启动子区,或包括转录起始位点(TSS)的区域。在一些方面,TSS可以定义为例如启动子上游的近端500bp、1000bp、1500bp或2000bp和下游500bp、1000bp、1500bp或2000bp。在一些方面,可以例如基于FRiP将计数归一化。在一些方面,启动子可及性可以用作与基因相关的度量。

[0179] 在一些实施方案中,表观遗传学图谱可以提供关于活性调控区和/或与调控区结合的转录因子的信息。例如,可以根据生成的测序读段的长度推断核小体位置。在一些实施方案中,可以根据生成的测序读段的大小、分布和/或位置推断转录因子结合位点。在一些实施方案中,可以根据生成的测序读段推断新的转录因子结合位点。在一些实施方案中,可

以根据生成的测序读段推断新的转录因子。

[0180] 在一些方面,如果进行生物学或技术重复,则可以评价重复以评估方法的质量和保真度。在一些实施方案中,可以使用区间分析评价重复。在一些实施方案中,可以使用共同峰和独特峰的Venn图来评估重复。在一些实施方案中,可以将每个重复的计数例如log₂归一化计数绘制在X-Y图上,并且可以计算相关系数(例如斯皮尔曼相关性)来评估重复。

[0181] 在一些方面,可以评估质量控制度量来确定来自表观遗传学/表观基因组学分析的序列信息和数据的质量。示例性控制度量包括峰中读段的分数(FRiP)、非冗余分数(NRF)、PCR瓶颈系数(PBC)、相对链交叉相关(RSC)、归一化链交叉相关(NSC)和不可重现的发现率(IDR)、确定未映射的未配对的读段、重复百分比、mtDNA污染、独特峰的数量、总映射读段(测序深度)和有效序列深度(冗余与非冗余)。在一些实施方案中,来自样品的序列读段含有小于15%、小于10%、小于8%或小于6%的未配对或未映射的读段。在一些方面,重复读段的分数可取决于输入和mtDNA污染。在一些方面,低mtDNA也可以是富集不良的指示或与吸出的沉淀一致。在一些方面,总映射读段的数目(测序深度)是至少10⁵、10⁶、10⁷或10⁸个或更多个读段。在一些方面,有效序列深度是至少10⁵、10⁶、10⁷或10⁸个或更多个读段。在一些方面,低FRiP可指示样品的富集不良。在一些方面,样品的FRiP为至少0.05、0.1、0.2、0.3或更高。

样品评估

[0182] 在一些实施方案中,测量、评估和/或确定了在样品比如含有细胞或细胞组合物的样品中的表观遗传学和/或表观基因组学状态。在一些实施方案中,样品是从受试者取得、采集和/或获得的生物样品和/或含有从受试者取得、采集和/或获得的细胞。在某些实施方案中,受试者患有疾病或病症和/或疑似患有疾病或病症。在一些实施方案中,受试者已经接受、将接受治疗或为接受治疗的候选者。在一些实施方案中,所述治疗是施用细胞疗法和/或免疫疗法。在某些实施方案中,细胞疗法治疗和/或能够治疗所述疾病或病症。在一些实施方案中,样品中的细胞或细胞组合物含有用于细胞疗法的细胞。在一些实施方案中,所述治疗是含有一种或多种工程化细胞的细胞疗法。在一些实施方案中,工程化细胞表达重组受体。在一些实施方案中,重组受体是嵌合抗原受体(CAR)或重组T细胞受体(TCR)。

[0183] 在一些实施方案中,从已经、将要被施用疗法或为施用疗法的候选者的受试者中取得、采集和/或获得样品。在一些实施方案中,在治疗或施用疗法(例如细胞疗法)之前取得、采集和/或获得样品。在一些实施方案中,在治疗或施用疗法(例如细胞疗法)之后取得、采集和/或获得样品。

[0184] 在某些实施方案中,样品含有、取自、衍生自和/或源自细胞或细胞组合物。在一些实施方案中,细胞被包含在样品中。在一些实施方案中,样品含有取自、源自和/或衍生自生物样品和/或与生物样品相同的来源的细胞。在一些实施方案中,细胞,例如来自生物样品和/或来自与生物样品相同的来源的细胞,是已经与处理和制备工程化细胞相结合产生的细胞,比如用于过继细胞疗法的细胞和/或为这种用途配制的那些细胞,例如,在包含药学上可接受的赋形剂(recipient)或冷冻保存剂的药物组合物中的细胞。在一些实施方案中,通过所述方法评估的细胞,例如来自生物样品和/或来自与生物样品相同的来源的细胞,其在一些方面可以是一个或多个对照细胞或一个或多个参考细胞。

[0185] 在一些实施方案中,通过所提供的方法和/或组合物评估的生物样品的细胞和/或

与生物样品相同的来源的细胞已被转导或有待转导为含有异源核酸和/或编码异源蛋白质的核酸或其他核酸或多肽产物(例如重组蛋白)。在一些实施方案中,异源核酸编码结合分子,比如重组受体,比如嵌合抗原受体(CAR)或转基因T细胞受体(TCR)。在一些实施方案中,所述细胞构成这种细胞的群、含有这种细胞和/或富集这种细胞的组合物,例如其中表达结合分子的细胞构成组合物中的总细胞或某种类型的细胞(比如T细胞,如CD8⁺和/或CD4⁺细胞)的至少15%、20%、25%、30%、35%、40%、50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多百分比。在一些实施方案中,所述细胞是原代T细胞。其中这些组合物是药物组合物和给药制剂,比如用于过继细胞疗法。

[0186] 在一些方面,样品(例如含有细胞的样品)衍生或分离自血液或血液衍生的样品,或者是或衍生自血液成分单采术或白细胞单采术产品。示例性样品包括全血、外周血单核细胞(PBMC)、白细胞、骨髓、胸腺、组织活检、肿瘤、白血病、淋巴瘤、淋巴结、肠相关淋巴组织、粘膜相关淋巴组织、脾、其他淋巴组织、肝脏、肺、胃、肠、结肠、肾、胰腺、乳房、骨、前列腺、子宫颈、睾丸、卵巢、扁桃体或其他器官、和/或由其衍生的细胞。在细胞疗法例如过继细胞疗法的背景下,样品包括来自自体 and 同种异体来源的样品。

[0187] 在一些实施方案中,样品包括来自基因工程化细胞的染色质制备物,所述细胞表达异源核酸和/或编码异源蛋白质的核酸或其他核酸或多肽产物(例如人或人源重组蛋白)。在一些实施方案中,细胞通常是真核细胞,比如哺乳动物细胞,并且典型地是人细胞。在一些实施方案中,细胞,例如样品和/或生物样品的细胞,来源于血液、骨髓、淋巴或淋巴器官,是免疫系统的细胞,比如先天性或适应性免疫的细胞,例如髓样细胞或淋巴样细胞,包括淋巴细胞,通常是T细胞和/或NK细胞。其他示例性细胞包括干细胞,例如多能和多能性干细胞,包括诱导的多能性干细胞(iPSC)。在一些实施方案中,细胞是原代细胞,比如直接从受试者分离和/或从受试者分离并冷冻的细胞。

[0188] 在一些实施方案中,细胞是T细胞(比如CD4⁺ T细胞和/或CD8⁺ T细胞)和/或自然杀伤(NK)细胞。在一些实施方案中,细胞是单核细胞或粒细胞,例如髓样细胞、巨噬细胞、嗜中性粒细胞、树突细胞、肥大细胞、嗜酸性粒细胞和/或嗜碱性粒细胞。在一些实施方案中,细胞或细胞组合物包含T细胞。在一些实施方案中,细胞或细胞组合物包含CD4⁺细胞。在一些实施方案中,细胞或细胞组合物包含CD8⁺细胞。

[0189] 在一些实施方案中,细胞和/或参考细胞可以来自任何来源。在一些实施方案中,细胞分离或获自受试者,例如人受试者。在一些实施方案中,细胞包括多个细胞或细胞群。细胞可以分离自软组织或体液或体外培养的细胞培养物。在一些实施方案中,染色质可以分离自来自软组织的细胞,所述软组织比如脑、肾上腺、皮肤、肺、脾、肾、肝、脾、淋巴结、骨髓、膀胱、胃、小肠、大肠或肌肉。在一些实施方案中,细胞来自受试者的体液,例如血液、血浆、唾液、粘液、痰、脑脊液、胸膜液、泪液、乳管液、淋巴、痰液、脑脊液、滑液、尿液、羊水或精液。在一些实施方案中,可以从细胞培养物例如细胞系获得细胞。

[0190] 在一些实施方案中,细胞包括先前经过修饰(例如基因工程)的细胞。在一些实施方案中,细胞包括获自受试者的细胞,所述细胞经工程改造或修饰,例如,经过工程改造以表达重组受体。在一些实施方案中,细胞包括获自受试者但未曾经历基因工程的细胞。在一些实施方案中,细胞包括获自受试者的细胞,并且本文中提供的方法可以在细胞工程化或修饰之前或之后在细胞样品中进行。在一些实施方案中,细胞包括已来自受试者的生物

样品(例如血液)选择或纯化的细胞。在一些实施方案中,细胞包括表达特定的细胞表面表达标志物的细胞子集。

[0191] 在一些实施方案中,待评估的核酸(例如基因组DNA、染色体DNA)来自血细胞,例如来自全血样品或全血中的细胞亚群的血细胞。全血中的细胞亚群包括血小板、红细胞(红细胞)、血小板和白细胞(即外周血白细胞,其由嗜中性粒细胞、淋巴细胞、嗜酸性粒细胞、嗜碱性粒细胞和单核细胞组成)。白细胞可以进一步分为两个组,粒细胞(也称为多形核白细胞,并包括嗜中性粒细胞、嗜酸性粒细胞和嗜碱性粒细胞)和单核白细胞(包括单核细胞和淋巴细胞)。淋巴细胞可以进一步分为T细胞、B细胞和NK细胞。外周血细胞存在于循环血液池中,而不是被隔离在淋巴系统、脾、肝或骨髓内。

[0192] 在一些实施方案中,通过使用针对细胞表面标志物的标记抗体的已知方法,可通过诸如荧光激活细胞分选(FACS)或磁激活细胞分选(MACS)的选择方法从异质细胞群(例如血液)中选择细胞群。可以使用这些方法分离多种细胞,包括干细胞、癌症干细胞和血细胞亚群。可以通过FACS或MACS从血液中选择示例性细胞和使用的示例性标志物包括:T细胞(CD3+CD4+CD8+)、B细胞(CD19+CD20+)、树突细胞(CD11c+CD20+)、NK细胞(CD56+)、干细胞/前体细胞(CD34+;仅限造血干细胞)、巨噬细胞/单核细胞(CD14+CD33+)、粒细胞(CD66b+)、血小板(CD41+CD61+CD62+)、红细胞(CD235a+)、内皮细胞(CD146+)和上皮细胞(CD326+)。可以使用针对其他细胞表面标志物的抗体分离这些细胞的亚群。

[0193] 在一些实施方案中,细胞,例如样品中的细胞,比如生物样品和/或来自与生物样品相同的来源的细胞,包含通过基因工程引入的一个或多个核酸,并由此表达这种核酸的重组或基因工程产物。在一些实施方案中,核酸是异源的,即通常在细胞或从细胞获得的样品中不存在,比如从另一生物或细胞获得的核酸,例如,通常在被工程化的细胞中和/或从中获得这种细胞的生物中未发现。在一些实施方案中,核酸不是天然存在的,比如在自然界未发现的核酸,包括包含编码来自多种不同细胞类型的各种结构域的核酸的嵌合组合的核酸。

[0194] 在一些实施方案中,可以在转导细胞或工程化细胞的制备、生产或制造中的任何观点评估表观遗传学和/或表观基因组学特性、状态和/或谱,所述细胞包括被转导或将要或已经被转导用于过继细胞疗法和受试者的治疗后监测的细胞。用于处理细胞的示例性步骤包括涉及细胞的分离、分开、选择、培养(例如,刺激细胞,例如诱导其增殖和/或活化)、转导、洗涤、悬浮、稀释、浓缩和/或配制的步骤,包括已知的和/或本文所述的那些步骤。在具体的实施方案中,处理步骤包括用病毒载体颗粒转导细胞,其中与病毒载体颗粒温育的至少一部分在启动转导的封闭系统或封闭室中进行。所述方法可以进一步和/或替代地包括其他处理步骤,例如用于分离、分开、选择、培养(例如,刺激细胞,例如诱导其增殖和/或活化)、洗涤、悬浮、稀释、浓缩和/或配制细胞的步骤。在一些实施方案中,细胞或细胞组合物包括已经受转导并且然后例如在37°C下培养多日或约1天、2天或3天,比如通常多日或约4天、5天、6天、7天、8天、9天、10天、11天、12天、13天、14天或更多时间。

[0195] 在一些实施方案中,细胞或细胞组合物包含在基因工程制造过程的任何阶段的细胞。在一些实施方案中,样品含有衍生自在基因工程制造过程的任何阶段的细胞的核酸和/或染色质制备物。例如,样品可包含来自已经用编码重组和/或异源分子的病毒载体颗粒转导的细胞的核酸和/或染色质制备物。在一些实施方案中,样品含有细胞,例如自体或同种

异体细胞,所述细胞通过用编码重组抗原受体(例如CAR)的异源核酸转导而被工程化并经过培养或扩增,例如以供与过继细胞疗法结合使用。在一些情况下,样品含有来自这样的已经冷冻保存的转导细胞的核酸和/或染色质制备物,所述制备物在一些方面被称为冷冻保存的药物产品(CDP)。在一些情况下,样品含有来自这样的已经配制为用于施用于受试者的转导细胞的核酸和/或染色质制备物,所述制备物在一些方面被称为配制药物产品(FDP)。

[0196] 在一些实施方案中,从受试者中获得样品,所述获得在这样的受试者接受了包含已经例如用编码重组和/或异源分子(例如CAR)的病毒载体颗粒转导的细胞的治疗之后。在一些实施方案中,作为对照,对于未经受转导和/或基因工程的患者匹配的对照样品,可以进行所提供的方法,所述对照样品可以是含有将要用于转导的选定细胞或富集细胞的样品。在一些实施方案中,这种患者匹配的对照样品可以是冷冻保存的样品,其在一些情况下被称为冷冻保存的材料(CMAT)。在一些实施方案中,核酸和/或染色质制备物是从细胞分离的。在一些情况下,从约 1×10^4 个、 1×10^5 个、 1×10^6 个、 1×10^7 个或更多个细胞分离核酸和/或染色质制备物。在一些实施方案中,核酸和/或染色质制备物是从 1×10^6 个细胞分离的。在一些情况下,从构成样品或其选定部分的所有或基本上所有细胞分离核酸和/或染色质制备物。

[0197] 在一些实施方案中,将细胞与一种或多种细胞刺激剂一起温育,所述细胞刺激剂是细胞结合剂,比如抗原结合试剂,比如抗体,其能够诱导细胞内信号传导和/或细胞增殖。在一些实施方案中,将细胞与抗CD3/抗CD28珠一起温育,包括与其混合。

[0198] 根据本文所述的方法,可以评估样品的一种或多种表观遗传学和/或表观基因组学特性,例如通过ATAC-seq确定的染色质可及性谱。在一些方面,所提供的方法可用于在工程化的各个阶段(例如在工程化之前或在工程化之后)评估、评价和/或表征样品中的细胞和/或细胞组合物,以便表达重组受体。在一些方面,表观遗传学和/或表观基因组学特性可用于评估将要施用的细胞或组合物的质量、一致性和/或特点,和/或在接受细胞疗法之前鉴定可能表现出期望的结果(例如对细胞疗法作出反应)和/或可能具有发生毒性的风险的受试者。

[0199] 在一些实施方案中,在细胞工程化之前取得、采集和/或获得样品。在一些实施方案中,在细胞工程化之后取得、采集和/或获得样品。在一些实施方案中,在治疗或施用疗法(例如细胞疗法)之后取得、采集和/或获得样品。根据本文所述的方法、试剂盒和制品,可以在接受免疫疗法之前或之后评估样品的一种或多种表观遗传学和/或表观基因组学特性。在一些实施方案中,在将细胞工程化之前30分钟、1小时、2小时、3小时、4小时、5小时、6小时、8小时、10小时、12小时、14小时或更长时间内(或在约上述时间内或约在上述时间)采集样品,例如,以便表达重组受体。在一些实施方案中,在将细胞工程化之后30分钟、1小时、2小时、3小时、4小时、5小时、6小时、8小时、10小时、12小时、14小时或更长时间内(或在约上述时间内或约在上述时间)采集样品,例如,以便表达重组受体。在一些实施方案中,在从受试者获得样品例如血液样品之后30分钟、1小时、2小时、3小时、4小时、5小时、6小时、8小时、10小时、12小时、14小时或更长时间内(或在约上述时间内)采集样品。在一些实施方案中,在免疫疗法例如细胞疗法施用开始之前30分钟、1小时、2小时、3小时、4小时、5小时、6小时、8小时、10小时、12小时、14小时或更长时间内(或在约上述时间内或约在上述时间)采集样品。在一些实施方案中,在免疫疗法例如细胞疗法施用开始之后30分钟、1小时、2小时、3小

时、4小时、5小时、6小时、8小时、10小时、12小时、14小时或更长时间内(或在约上述时间内或约在上述时间)采集样品。

[0200] 在一些方面,可以使用本文中提供的任何方法来确定细胞或细胞组合物的特性、质量和/或一致性。在一些方面,可以使用任何方法来评估在所述时间点中的任何一个或多个时间点的细胞或细胞组合物的特性、质量和/或一致性,所述时间点例如在工程化之前、在工程化中间过程、在工程化之后,在将细胞施用于受试者之前、施用期间和/或在施用之后。

[0201] 在一些实施方案中,在测定法中使用的细胞群可以由任何数目的细胞组成,例如,约500至约 10^6 个或更多个细胞,约500至约100,000个细胞,约500至约50,000个细胞,约500至约10,000个细胞,约50至1000个细胞,约1至500个细胞,约1至100个细胞,约1至50个细胞或单个细胞。在一些实施方案中,细胞样品包括少于约1000、2000、3000、4000、5000、6000、7000、8000、9000、10,000、15,000、20,000、25,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、120,000、140,000、160,000、180,000、200,000、250,000、300,000、350,000、400,000、450,000、500,000、600,000、700,000、800,000、900,000或1,000,000个细胞。在一些实施方案中,细胞样品包括多于约1000、2000、3000、4000、5000、6000、7000、8000、9000、10,000、15,000、20,000、25,000、30,000、40,000、50,000、60,000、70,000、80,000、90,000、100,000、120,000、140,000、160,000、180,000、200,000、250,000、300,000、350,000、400,000、450,000、500,000、600,000、700,000、800,000、900,000或1,000,000个细胞。在一些实施方案中,与其他方法相比,本文中提供的方法需要相对较少的细胞。

表观遗传学/表观基因组学谱的处理和进一步分析

[0202] 在一些实施方案中,所提供的方法还涉及表观遗传学/表观基因组学特性和/或谱的一个或多个方面的附加处理和/或分析。在一些实施方案中,所提供的方法包括比较特定的细胞或细胞组合物与另一细胞或细胞组合物的表观遗传学/表观基因组学特性和/或谱,例如比较和/或差异分析,或生成用于归一化或进一步的分析步骤的度量和计数。在一些方面,进一步的处理和/或分析包括使用另外的步骤来分析序列结果,比如核小体占据分析。

[0203] 在一些方面,可以计算实施各种附加处理和/或进一步分析中的任何一者或多者。在一些实施方案中,处理和/或进一步分析可包括例如差异性可及性(差异峰)分析、归一化测量和评估核小体占据/定位。在一些实施方案中,用于分析和鉴定一个或多个基因组区域的方法(例如,处理和/或进一步分析)可包括以任何顺序依次或同时进行本文所述的步骤、功能、过程或脚本中的任何一者或多者。步骤、功能、过程或脚本中的一者或多者可以被实现或执行相似功能的相似算法、步骤、功能、过程或脚本取代。在一些实施方案中,一个或多个步骤的顺序次序可以是任何次序,或者可以并行执行任何一个或多个步骤、功能、过程或脚本。

[0204] 在一些实施方案中,各种处理和/或进一步分析中的一者或多者可涉及鉴定一个或多个基因组区域,例如编码区、非编码区、基因间区、非翻译区、内含子、外显子、顺式调控区、小RNA结合位点、重复序列、端粒区和/或可及或不可及的区域(例如,开放染色质和/或异染色质),所述区域与细胞疗法治疗的结果相关和/或可用于评估例如将要施用于受试者的细胞组合物或细胞培养物。在一些实施方案中,使用两个或更多个样品中的一个或多个

基因组区域的表观遗传学和/或表观基因组学谱的差异来鉴定一个或多个基因组区域,所述样品例如来自不同受试者、来自具有不同结果的受试者、来自接受不同治疗的受试者、来自细胞工程化的不同阶段和/或来自经受不同条件的细胞的样品。在一些实施方案中,一个或多个基因组区域在表观遗传学和/或表观基因组学分析中包括不同的信号峰。在一些实施方案中,所述一个或多个基因组区域包括显示不同统计参数或度量的区域,所述参数或度量例如在两个或更多个不同样品之间的区域上的平均信号、中值信号、总和信号或总信号。在一些实施方案中,可以在所述方法的其他任选步骤的框架内使用处理和/或进一步分析,包括阈值化和聚类方法、预测建模和/或差异性峰调用。

差异性峰分析

[0205] 在一些实施方案中,可以使用测定法来比较两个或更多个样品。在一些实施方案中,测定法可包括使用本文中提供的测定法或分析方法分析第一细胞群或第一细胞组合物以产生第一表观遗传学图谱;使用本文中提供的测定法或分析方法分析第二细胞群或第二细胞组合物以产生第二表观遗传学图谱;并且将第一表观遗传学图谱与第二表观遗传学图谱进行比较,例如,用于检测染色质可及性或转录因子占据的任何差异或变化。

[0206] 在一些方面,分析包括差异分析,例如差异性可及性或差异性峰分析。在一些方面,分析包括比较两个或更多个样品中存在的序列或信号峰、和/或两个或更多个样品之间的特定序列或信号峰的差异性富集。所述两个或更多个样品可包括来自第一细胞群或细胞组合物的细胞和来自第二细胞群或细胞组合物的细胞。在一些方面,分析包括确定在两个或更多个样品中存在的峰,并比较两个或更多个样品之间的峰的宽度和幅度或峰的存在和不存在。

[0207] 在一些实施方案中,可以使用两个或更多个样品例如细胞群或细胞组合物之间的峰比较来鉴定例如含有活性调控元件和/或任选地与特定的基因或转录因子相关的在样品中差异地可及的区域、基因座、元件或区间,并且/或者鉴定标记,例如在特定细胞状态或关联或预测特定结果(例如细胞疗法的结果)的特定基因组区域、元件或区间处或附近的峰的标记。

[0208] 在一些实施方案中,可以使用诸如DESeq2、DiffBind、MAnorm、csaw、DPChIP、BADS、diffReps、DIME、HMCan-diff、ChIPDiff、MMDiff、THOR、POLYPHEMUS、GenoGAM、normR、chromstaR、PePr、ChIPComp、EpiCentr、ODIN、histoneHMM、ImpulseDE2、QChIPat、SICER、MACS2、unique Peaks、ODIN、RSEG、MAnorm、Homer、DBChIP、multiGPS、edgeR以及描述于Steinhauser等人,(2016) Briefings in Bioinformatics, 17 (6) :953-966或<https://omictools.com/peak-calling-category>中的差异性峰调用步骤、功能、过程、工具或脚本来确定峰的差异。在一些方面,所述方法可用于鉴定在两个或更多个样品之间表现差异性可及性的峰。在一些方面,所述方法包括使用定量可及性数据计算多个样品的差异性结合位点。在一些方面,可用于差异分析的示例性样品包括不同的细胞组合物或细胞类型(例如CD4+或CD8+细胞)、或处于工程化过程的不同阶段的细胞(例如CMAT或CDP)、来自不同供体的细胞或经受不同处理(例如用药物处理或用药剂温育)的细胞。

[0209] 在一些方面,可以使用差异性峰调用步骤、功能、过程或脚本(比如DiffBind)来确定峰的差异。在一些方面,分析步骤鉴定在两个或更多个样品之间不同地存在的基因组区域、基因座和/或区间,并且包括峰集的处理,包括重叠和合并峰集,将峰集中的测序读段重

叠区间计数,和基于读段密度差异的证据鉴定统计学上显著差异地呈现的峰位点。在一些实施方案中,分析步骤鉴定峰重叠或共有峰,例如,在分析的2个或更多个样品中存在的可及性峰。在一些实施方案中,在确定重叠后,过滤在多于两个的文库中发现的峰,并提取测序计数。在一些实施方案中,计算了测量值,比如峰中读段的分数(FRiP;显示信号的富集,计算为(峰中读段数)/(总读段数))。

[0210] 在一些实施方案中,可以确定两个或更多个样品之间的峰信号的差异或变化。在一些实施方案中,确定了一个样品与另一个样品相比的峰信号的倍数变化,所述另一个样品比如对照样品或参考样品,例如未经处理的样品或表现出期望的特点或结果的样品,例如本文所述的任何参考样品。在一些情况下,将数据转换并以对数标度绘制。在某些实施方案中,对数变换是常用对数($\log_{10}(x)$)、自然对数($\ln(x)$)或二进制对数($\log_2(x)$)。在一些实施方案中,将数据绘制为描绘x轴倍数变化例如 \log_2 倍数变化以及在y轴上的统计显著性例如p值的 $-\log_{10}$ 的火山图。

[0211] 在一些方面,两个或更多个样品之间的峰的比较可用于鉴定例如样品或条件之间的活性调控元件(任选地与特定基因或转录因子相关),和/或将其用于鉴定特定细胞状态或预测结果(比如反应或安全性结果)或细胞或细胞组合物的其他特点的标记。

[0212] 在一些实施方案中,可以使用包括特征选择算法的分析步骤来鉴定在样品中差异地可及的区域、基因座、元件或区间,并且/或者鉴定标记,例如在特定细胞状态或关联或预测特定结果(例如细胞疗法的结果)的特定基因组区域、元件或区间处或附近的峰的标记。

[0213] 在一些方面,可以在全基因组范围上,或者在整个基因组全局,或者在感兴趣的特定基因组区域、基因座和/或区间(例如,在感兴趣的特定基因或基因集合处或附近)进行或评估差异峰分析。在一些方面,可以在特定的基因集合中进行或评估差异峰分析,所述基因例如与细胞类型规范、T细胞分化和/或发育、免疫细胞功能、免疫表型和/或转录因子相关的基因,包括本文所述的任何基因或基因的集合、集或模块。

[0214] 可以通过差异性可及性分析来评估任何两个样品,例如细胞群或细胞组合物。在一些方面,两个或更多个细胞样品(例如细胞组合物)不同或可能不同或可能在一个或多个特性、属性或特征上不同,所述特性、属性或特征比如与细胞的表型或功能有关,包括T细胞的一个或多个表型或功能,比如T细胞活化状态、记忆表型、分化状态、效应子功能;细胞因子反应、产生或分泌;细胞溶解活性、运输或迁移能力、持久性或衰竭;或与例如由基于病毒的转导介导的转基因的存在或不存在有关。

[0215] 在一些实施方案中,已知第一和第二细胞样品之一(例如细胞组合物)富集细胞,例如超过75%的此类细胞,具有或可能具有幼稚表型或长寿记忆表型,而另一个样品(例如其他细胞组合物)具有未知或未确定的记忆表型。在一些实施方案中,已知第一和第二细胞样品之一(例如细胞组合物)富集细胞,例如超过75%的被活化的此类细胞,而另一个样品(例如其他细胞组合物)具有未知或未确定的活化状态。在一些实施方案中,已知第一和第二细胞样品之一(例如细胞组合物)是均匀的或同质的或相对均匀或同质的,例如在组合物或样品中大于85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的细胞表现出一个或多个特性、属性或特征,例如CD4⁺细胞、CD8⁺细胞的百分比;一种或多种活化标志物、记忆标志物、凋亡或细胞健康标志物或指示细胞的表型或功能的其他标志物的水平。在这样的例子中,另一个细胞样品(例如细胞组合物)的所述特

性、属性或特征具有未知或未确定的均匀性分布。在一些实施方案中,已知第一和第二样品之一(例如细胞组合物)具有每细胞特定的平均或均值转基因(例如病毒载体)拷贝数(载体拷贝数(VCN)),并且另一个细胞样品(例如细胞组合物)具有未知或未确定的VCN。可以将细胞组合物的任何数目的期望特征、属性或特性与差异性可及性分析结合使用。

[0216] 第一和第二组合物的其他示例性特性或特征包括与细胞组合物的工程化、细胞来源、疾病类型或状态的差异有关的任何特性或特征。在一些实施方案中,第一组合物和第二组合物之一包含有待用重组受体基因工程改造的细胞,并且第一组合物和第二组合物中的另一者包含经工程改造以表达重组受体的细胞。在一些实施方案中,不同的样品,例如第一细胞组合物和第二细胞组合物,包含来自不同供体的原代细胞,所述供体任选基于疾病状态、疾病严重性或疾病类型而不同的供体。在一些实施方案中,不同的样品,例如第一细胞组合物和第二细胞组合物,包含处于工程化细胞的制造过程的不同阶段或步骤的细胞。在一些实施方案中,不同的样品,例如第一细胞组合物和第二细胞组合物,包含与药剂接触以调节细胞的活性、表型或功能的细胞,并且第一和第二组合物中的另一者包含并不如此接触的相似细胞。在这样的实例中,这样的药剂可以包括多肽或蛋白质、肽、抗体、核酸、病毒载体或病毒制剂,或小分子化合物,包括例如刺激性试剂,任选抗CD3/抗CD28;免疫调节剂、对重组受体例如CAR特异的抗独特型抗体或其抗原结合片段、免疫检查点抑制剂、代谢途径调节剂、腺苷受体拮抗剂、激酶抑制剂、抗TGFβ抗体或抗TGFβR抗体或细胞因子。

[0217] 在一些实施方案中,在不同的时间从同一个体获得第一细胞群和第二细胞群或另外的细胞群。在其他实施方案中,第一细胞群和第二细胞群或另外的细胞群是从组织或不同的个体获得的不同细胞群。在一些实施方案中,第一细胞群和第二细胞群或另外的细胞群获自不同的个体和/或个体组,例如表现出不同的疾病严重性、治疗反应、治疗结果、毒性和/或副作用、生理反应、细胞的分子和/或功能活性和/或表型的个体或个体组。在一些实施方案中,不同的样品,例如第一细胞组合物和第二细胞组合物,包括细胞组合物样品,所述细胞组合物样品与第一和第二组合物之一施用于受试者后发生或已经发生但另一者则不然的结果相关。

[0218] 可用于所述方法的用于比较的示例性细胞包括,例如,从正常受试者分离的细胞、从受试者分离的未经历工程化的细胞、从显示完全缓解的受试者分离的细胞、从显示持续缓解的受试者分离的细胞和/或在特定治疗之前或之后从受试者分离的细胞。在一些实施方案中,细胞可以经历其他过程或处理,例如,基于细胞表面标志物表达的选择或基因工程。

[0219] 在一些实施方案中,可以进行差异峰分析,比较两个或更多个样品例如两个或更多个测试样品或一组样品的表观遗传学谱。在一些实施方案中,可以进行差异峰分析,将一个或多个测试样品(例如待测细胞组合物)的可及性谱与一个或多个参考样品的可及性谱或参考谱(例如本文在例如第II.C.1节中所述的任何参考样品或参考谱)进行比较。

归一化测量和度量

[0220] 在一些实施方案中,来自表观遗传学/表观基因组学谱的序列和/或峰的分析包括进一步处理或分析序列和/或峰结果,例如将结果归一化或量化。在一些方面,分析包括采取归一化测量。在一些情况下,归一化测量可包括生成与基因相关的度量,例如每注释基因的归一化计数。在一些实施方案中,进一步处理或分析来自测定法例如染色质可及性测定

法(比如ATAC-seq)的数据,以确定阈值和/或值,鉴定与特定结果相关的基因座和/或基因,并且/或者与其他测定法的结果和分析相比较。例如,在一些实施方案中,可以通过绘制与基因组基因座区域图谱或基因组图谱重叠的频率(比如归一化频率)来鉴定ATAC-seq峰。在一些实施方案中,归一化频率值包括FPKM或RPKM。在一些实施方案中,如本文中使用的术语“FPKM”是指每百万映射片段数的每千碱基片段数。在一些实施方案中,术语“RPKM”可以指每百万映射读段数的每千碱基读段数。FPKM和RPKM是量化任何基因组特征的丰度的单位,其通过与基因组特征比对的测序读段的丰度来确定,所述基因组特征例如外显子、转录物、转座子的插入、组蛋白修饰、核小体占据、DNA甲基化或蛋白质(比如转录因子)的占据、或任何基因组坐标。FPKM和RPKM测量通过基因组单元的相对长度以及映射到它的读段的总数将丰度归一化,以便于比较样品内和样品之间的丰度水平。在一些方面,归一化频率值包括上四分位数(UQ)、大小因子(SF)、每百万计数(CPM)、每百万映射读段数的每千碱基读段数(RPKM)、每百万映射片段数的每千碱基片段数(FPKM)或每百万映射的每千碱基转录物(TPM)。

[0221] 在一些实施方案中,选择一个或多个基因组区域、基因座、元件或区间进行分析。例如,在一些实施方案中,评估和/或比较来自基因组区域(例如启动子或增强子区)的数据。在一些实施方案中,评估和/或比较来自基因例如基因体的编码区内或周围的数据。例如,在一些实施方案中,评估和/或比较基因的基因体的FPKM或RPKM值的总和。在一些实施方案中,针对一个或多个基因组区域例如基因组基因座评估和/或比较FPKM或RPKM值的总和。

[0222] 在一些实施方案中,确定与参考值和/或归一化值相比的差异模式,例如,与参考值或归一化值相比更高或更低的染色质可及性。在一些实施方案中,测定法包括确定值的差异,例如通过归一化值确定的染色质可及性,例如确定并比较在一个或多个基因组区域(例如基因组基因座,例如基因组中的多个基因座或基因座集合)的FPKM或RPKM。在一些实施方案中,多个基因座或基因座的集合、集或模块包括本文所述的任何一种,或使用本文中提供的方法鉴定的任何一种。

[0223] 在一些实施方案中,确定一个或多个基因组区域(例如基因组基因座)的值的差异涉及确定一个或多个基因组区域例如获自癌性生物样品的核酸中的基因组基因座的FPKM或RPKM值:i)相对于一个或多个基因组区域例如获自参考生物样品的参考核酸中的基因组基因座的FPKM或RPKM值,平均或总FPKM或RPKM值的变化大于1至20倍,例如约1倍、2倍、3倍、4倍或5倍;并且ii)相对于一个或多个基因组区域例如获自参考生物样品的参考核酸中的基因组基因座的FPKM或RPKM值,FPKM或RPKM范围降低大于约10倍。例如,在一些实施方案中,分析包括确定一个或多个基因组区域(例如基因组基因座)的值,例如FPKM或RPKM值,基于本文所述的用于表观遗传学和/或表观基因组学分析的任何测定法生成一个或多个基因组区域例如基因组基因座的测序标签计数的矩阵,并确定与参考值或归一化值相比的值的差异。然后,分析进一步包括聚类分析,比如层次聚类或主成分分析(PCA),其基于一个或多个细胞群中的值的变化或差异,所述细胞群例如来自一位或多位受试者或经历一种或多种处理的细胞,比如本文所述的任何其他或另外的分析。

[0224] 在一些实施方案中,测定法包括确定值的差异,例如通过归一化值确定的染色质可及性,其可以使用针对有效文库大小调整的归一化值来确定。在一些实施方案中,可以使

用负二项模型来测试差异性可及性。在一些实施方案中,在一个或多个基因组区域、基因座、元件或区间(例如基因组基因座)中的峰(例如来自一种或多种不同条件的样品之间的序列读段峰)的差异,可用于确定表观基因组学和/或表观遗传学谱的差异,例如,染色质可及性的差异。

[0225] 在一些实施方案中,这样的归一化测量包括使用计算实施的脚本或工具(例如, `homer_calculate_per_gene_accessibility_norm`)来计算代表基因的一般可及性的归一化的每基因可及性值(例如,归一化的FPKM值)。这样计算的度量或评分可用于进一步的下游分析,例如,与其他样品比较或与细胞疗法的结果相关联。

[0226] 在一些实施方案中,归一化还可以包括基于测量值的归一化,所述测量值比如峰中读段的分数(FRiP;显示信号的富集,计算为(峰中读段数)/(总读段数))。在一些实施方案中,可以确定一个或多个特定的基因组区域或区间的FRiP归一化计数。

核小体占据分析

[0227] 在一些方面,进一步的处理步骤包括评估核小体占据或定位或鉴定表观遗传学/表观基因组学谱中的无核小体区域(NFR)。

[0228] 在一些实施方案中,峰含有NFR和/或与NFR相关(参见,例如, Schep等人, (2015) *Genome Research* 25:1757-1770)。在一些方面,无核小体区域可与转录起始位点、启动子或转录终止位点(参见,例如, Yadon等人, (2010) *Mol. Cell. Biol.* 30 (21) :5110-5122)和/或更激活或更分化的状态相关。在一些方面, NFR与高的基因组可及性相关。在一些实施方案中,峰调用和差异峰分析可包括一个或多个步骤,例如过滤亚核小体读段,找到峰,将每峰的插入位点计数,聚集每基因的启动子和/或转录起始位点峰信号和/或检查稳定共享峰中的差异性可及性。在一些方面,表观遗传学谱包括鉴定差异性可及性的峰,例如,鉴定或确定在一个或多个基因组区域、基因座、元件或区间(例如基因组基因座)中的峰(例如来自一种或多种不同条件的样品之间的序列读段峰)的差异,可用于确定表观基因组学和/或表观遗传学谱的差异,例如,来自一种或多种不同条件的样品之间的染色质可及性的差异。

[0229] 在一些实施方案中,序列读段可以按长度分组。在一些实施方案中,一些序列可以基于其大小被注释为无核小体序列(即,来自被预测将在核小体之间的片段的序列)(参见,例如, Schep等人, (2015) *Genome Research* 25:1757-1770)。还可以鉴定与单核小体、双核小体和三核小体相关的读段。

[0230] 在一些实施方案中,可以使用步骤、功能、过程、工具或脚本来确定核小体占据、核小体定位和/或无核小体区域,例如,用于分离出可以代表核小体结合的染色质的特定大小的片段。在一些方面,可以分离或过滤出特定大小的片段,例如大于100bp、150bp、200bp、250bp或更大的片段。在一些方面,一些所述片段可以代表核小体结合的染色质。在一些方面,用于过滤出特定大小(例如,大于100bp、150bp、200bp、250bp或更大)的片段的示例性步骤包括 `Filter_nucleosomal_fragments`。在一些实施方案中,这样的步骤可以富集信号。在一些方面,过滤和未过滤的读段两者都可用于进一步或下游分析以分别推断开放染色质和核小体占据的染色质。在一些实施方案中,可以使用步骤、功能、过程、工具或脚本来确定核小体占据,例如,使用ATAC-Seq数据(比如NucleoATAC)调用核小体位置和占据(参见,例如, Schep等人, (2015) *Genome Res.* 25 (11) :1757-1770)。

分析和应用

[0231] 在一些实施方案中,所提供的方法还涉及表观遗传学/表观基因组学特性和/或谱的一个或多个方面的附加应用和/或分析。在一些实施方案中,所提供的方法包括进一步的分析或应用步骤,比如将特定细胞或细胞组合物的表观遗传学/表观基因组特性和/或谱与参考样品和/或参考谱的那些进行比较。在一些方面,进一步的分析或应用包括确定另外的特点、特性和/或状态;使用另外的标准或方法来分析序列结果;进行另外的下游分析、基因子集分析、生物途径分析、转录占据分析、或基序分析、聚类分析和/或附加应用,例如,与细胞疗法的结果关联或相关,评估转基因序列的整合,与参考样品和/或参考谱比较,和/或评估细胞或细胞组合物的表型或状态或其他特性。

[0232] 在一些方面,可以计算实施各种附加分析和/或应用中的任何一者或多者。在一些实施方案中,所提供的方法涉及用于分析或比较的各个下游步骤或过程,例如计算实施的步骤和/或方法的应用,用于评估细胞或细胞组合物的一种或多种特性或特点。在一些实施方案中,附加分析和/或应用可以包括例如生物途径分析、基因本体论(GO)富集、基序富集、基因子集分析、转录因子占据、与参考谱比较和/或整合分析,例如确定转座子插入位点。在一些实施方案中,用于分析和鉴定一个或多个基因组区域的方法(例如,进一步的分析或应用)可包括以任何顺序依次或同时进行本文所述的示例性步骤、功能、过程或脚本中的任何一者或多者。步骤、功能、过程或脚本中的一者或多者可以被实现或执行相似功能的相似算法、步骤、功能、过程或脚本取代。在一些实施方案中,一个或多个步骤的顺序次序可以是任何次序,或者可以并行执行任何一个或多个步骤、功能、过程或脚本。

参考表观遗传学特性或谱

[0233] 在一些实施方案中,本文所述的测定法、步骤和/或程序中一者或多者可用于确定特定样品例如特定的细胞、细胞群或细胞组合物的表观遗传学和/或表观基因组学特性、状态和/或谱。在一些方面,所述方法用于确定参考细胞或参考细胞组合物或对照细胞和/或对照细胞组合物的表观遗传学和/或表观基因组学特性、状态和/或谱。在一些方面,本文所述的测定法、步骤和/或程序中一者或多者可用于确定参考表观遗传学特性,例如参考表观遗传学和/或表观基因组学谱。在一些方面,任何这样的特性或谱可用于与测试样品或测试条件或程序比较和/或评估测试样品或测试条件或程序,所述测试样品例如其他细胞或细胞组合物,所述测试条件或程序例如用于生成、制造和/或操作细胞或细胞组合物例如用于过继细胞疗法的细胞组合物的条件或程序。在一些方面,进一步的分析或应用方法或程序涉及将测试样品的表观遗传学/表观基因组学谱与参考样品和/或参考谱进行比较。

[0234] 在一些实施方案中,可以通过鉴定表观遗传学特性,例如通过参考或对照细胞样品、群体或组合物的染色质可及性分析来确定参考表观遗传学特性。在一些实施方案中,参考或对照细胞样品表现出已知的属性或特征,或疑似或可能具有已知的属性或特征,所述属性或特征例如与细胞的表型或功能有关,包括T细胞的一个或多个表型或功能,比如T细胞活化状态、记忆表型、分化状态、效应子功能;细胞因子反应、产生或分泌;细胞溶解活性、运输或迁移能力、持久性或衰竭;或与例如由基于病毒的转导介导的转基因的存在或不存在有关。在一些实施方案中,参考或对照样品来自健康或正常受试者。

[0235] 在一些实施方案中,参考样品包括源自施用细胞疗法之后继续实现期望的结果(例如期望的反应或安全性结果)的受试者的样品,例如细胞或细胞组合物。在一些实施方案中,参考样品包括源自用细胞疗法治疗后继续实现部分缓解(PR)或完全缓解(CR)和/或

持续缓解的受试者的样品。在一些实施方案中,参考样品是源自继续实现期望的安全性结果的受试者的样品,所述安全性结果例如没有发生毒性(例如细胞因子释放综合征(CRS)或神经毒性(NT))或者没有发生重度CRS或重度NT。在一些实施方案中,期望的治疗结果,例如反应或安全性结果,包括下文第III节中描述的任何结果。

[0236] 在一些方面,参考样品可包括具有期望的状态、质量、一致性、表型、特点和/或特性的细胞或细胞组合物。例如,在一些方面,可能期望含有均匀或一致的细胞的细胞组合物。在一些实施方案中,参考样品可包括已知表现出一致性和/或均匀性和/或特定表型、状态、质量、特点和/或特性的细胞或细胞组合物。在一些方面,参考样品可含有特定比例的表现出特定表型的细胞。在一些方面,参考表观遗传学和/或表观基因组学谱可获自所描述的任何参考样品。

[0237] 在一些实施方案中,参考谱包括参考样品的一个或多个基因组区域、基因座和/或区间的表观遗传学特性或状态的集合。在一些实施方案中,参考谱是参考表观遗传学和/或表观基因组学图谱。在一些实施方案中,参考谱是根据来自多个细胞组合物(例如,参考细胞组合物或样品)中的可及性分析(任选地染色质可及性)的序列读段的共同峰确定的表观遗传学图谱。在一些实施方案中,参考谱是一组度和/或归一化值,例如FPKM值或启动子可及性度量。

[0238] 在一些实施方案中,参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

[0239] 在一些实施方案中,阈值:是当施用于具有相同或相似疾病或病症的受试者时显示表现出结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者是来自当施用于受试者时显示表现出一个或多个期望的结果的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。在一些实施方案中,阈值包括来自正常或健康受试者的细胞中表观遗传学特性的值或水平。在一些实施方案中,阈值包括表现出特定表型,例如幼稚或长寿记忆表型的细胞中的表观遗传学特性的值或水平。

[0240] 在一些实施方案中,所提供的方法可用于评估将施用于受试者的细胞组合物,所述方法包括:分析在细胞组合物中的细胞的一个或多个基因组区域的表观遗传学谱,所述细胞组合物含有用重组受体工程改造的细胞;并且将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示当施用于受试者时细胞群是否或是否可能表现出或产生结果(例如期望的结果)。在一些实施方案中,所提供的评估细胞培养物的方法包括分析包含在输出细胞组合物中的细胞的一个或多个基因组区域的表观遗传学谱,所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生;并且将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示所述细胞是否或是否可能表现出预定的表型、持久性、活性和/或功能。

[0241] 在一些实施方案中,所述比较是在一个样品(例如测试样品)的表观遗传学/表观基因组学特性、状态和/或谱与参考样品的表观遗传学/表观基因组学特性、状态和/或谱之间进行的。在一些方面,与参考谱进行比较。在一些实施方案中,参考样品和/或参考谱可以是本文所述的任何一种。在一些方面,参考样品可以是来自表现出期望的结果(例如反应结果或安全性结果)的受试者的样品和/或包含期望的特性和/或特点的产品。

[0242] 本文中提供的方法还可用于评估细胞组合物的一个或多个特性或特征,比如细胞组合物的一致性、组成、表型、功能和/或活性。在一些方面,可以将细胞组合物与参考样品(例如包含参考细胞组合物)的表观遗传学和/或表观基因组学谱进行比较。在一些方面,所述方法可用于选择或生成用于治疗适当的细胞组合物。在一些方面,本文所述的测定法和分析可用于选择用于治疗的细胞,例如用于过继细胞疗法的工程化细胞,所述细胞被预测为更一致和/或均匀,和/或表现出特定的表型或组成,和/或与施用后更一致、有效和/或可预测的结果相关。在一些方面,通过选择表现出与功效和/或安全性和/或更明确、一致和/或均匀的细胞组成相关的表观遗传学和/或表观基因组学特性的细胞,本文中提供的方法可用于针对用于治疗的细胞的功效和/或安全性提供可靠的预测指标。

[0243] 在一些实施方案中,所提供的方法可用于评估将施用于受试者的细胞组合物,所述方法包括:分析在细胞组合物中的细胞的一个或多个基因组区域的表观遗传学谱,所述细胞组合物含有用重组受体工程改造的细胞;并且将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示当施用于受试者时细胞群是否或是否可能表现出或产生结果和/或产生与细胞疗法的结果相关的一致、明确和/或均匀的细胞组成。

[0244] 在一些实施方案中,如果所述比较指示细胞组合物将要或可能将要表现出结果(例如期望的反应或安全性结果),则可以将细胞组合物施用于受试者。

[0245] 在一些实施方案中,如果所述比较指示细胞组合物不会或可能不会表现出结果(例如期望的反应或安全性结果),则任选下列之一:(i)可以在施用前改变细胞组合物;(ii)可以在施用前改变细胞组合物的剂量;(iii)可以在施用前改变细胞组合物的剂量方案;(iv)可以与一种或多种其他治疗剂联合施用细胞组合物;或(v)不对受试者施用所述细胞组合物。

[0246] 在一些方面,基于一种或多种表观遗传学/表观基因组学分析的结果,可以在施用前改变细胞组合物。例如,在一些方面,如果特定的细胞组合物的表观遗传学/表观基因组学谱的比较是与参考细胞组合物比较,并且所述比较指示所述细胞组合物不会或可能不会表现出结果(例如期望的反应或安全性结果),则可以在细胞治疗之前修改细胞工程化的一个或多个步骤或程序。在一些方面,这样的修改可包括在一种或多种测试剂或条件存在下培养用重组受体工程改造的细胞和/或将要被基因工程改造的细胞。在一些实施方案中,一种或多种测试剂或条件包括血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;细胞组合物的来源或细胞类型;细胞组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌流;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的细胞组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。在一些实施方案中,一种或多种测试剂或条件包括来自测试化合物文库的一种或多种化合物。在一些实施方案中,如果所述比较指示细胞组合物具有或可能具有所述表型或功能,则所述方法还包括选择用于培养细胞的一种或多种测试剂或条件。在一些实施方案中,如果所述比较指示细胞组合物具有或可能不具有所述表型或功能,则所述方法还可以包括用一种或多种其他测试剂或条件重复与其他样品的表观遗传学和/或表观基因组学分析和/或比较。

[0247] 本文中提供的方法还可以用作诊断(例如,提供诊断的方法以及提供预后的方法)和/或用于选择适合的受试者和/或细胞用于治疗。所述方法可包括,例如,使用本文所述的

测定法分析细胞例如从受试者获得的细胞中的表观遗传学和/或表观基因组学特征,以产生表观遗传学图谱;并基于表观遗传学图谱和/或进一步的分析和/或定量提供诊断或预后。例如,本文所述的测定法和分析可用于预测功效、完全缓解、分子学缓解、缓解的持久性、细胞持久性、细胞扩增安全性(毒性或不良反应不存在或降低)和/或治疗(例如过继细胞疗法)的免疫原性不存在。在一些实施方案中,本文所述的测定法和分析可用于选择用于治疗的药剂,例如用于过继细胞疗法的工程化细胞,其被预测为更有效和/或更安全,例如与完全缓解或分子学缓解相关。通过鉴定具有改变的表观遗传学和/或表观基因组学特征(例如改变的染色质可及性)的基因组区域(例如基因组基因座),并且选择表现出与功效和/或安全性相关的表观遗传学和/或表观基因组学特性的细胞,本文中提供的方法可用于针对用于治疗的细胞的功效和/或安全性提供可靠的预测指标。

[0248] 在一些实施方案中,所提供的方法可用于预后,例如,用于确定治疗的受试者是否具有复发的风险。所述方法可用于鉴定用细胞(例如工程化细胞)治疗的可能经历疾病或病症(例如癌症)复发的受试者,这样可以为受试者提供另外的治疗选项,包括另外的或改变的剂量、改变的频率和/或用另外的药剂(例如表达重组受体的其他细胞)治疗、化学疗法、放射、生物修饰剂、抑制剂和其他适合的疗法。所述方法可用于预测受试者对治疗的长期反应。

[0249] 在一些实施方案中,所提供的方法可用于确定对于患有疾病或病症(例如癌症或肿瘤)的受试者的适合剂量、频率、剂量方案和/或疗程。在一些实施方案中,治疗(例如用于过继细胞疗法的细胞)的剂量可以基于预测的待施用细胞的功效和/或安全性来确定。疗程是指在诊断后或在治疗后对受试者采取的治疗措施。例如,确定缓解、毒性、复发、散布或受试者存活的可能性可用于确定治疗的剂量和程度和/或待施用的任何联合治疗。

[0250] 在一些实施方案中,所提供的方法可用于疾病或病症的表征、分类、区分、分级、分期、诊断或预后,或选择用于治疗例如过继细胞疗法的适合细胞,所述细胞的特征在于表观遗传学模式(例如,染色质可及性或DNA结合蛋白占据的模式)。例如,所述方法可用于确定,与来自正常受试者或在过继细胞疗法之后表现出稳健和/或持续缓解的受试者(例如表现出完全缓解的受试者)的相同类型的细胞的表观遗传学图谱相比,从受试者获得的用于工程化和施用的细胞的表观遗传学图谱是相似还是不同。所述方法还可用于预测个体对治疗(例如过继细胞疗法)的易感性。

生物途径、PCA和其他下游分析

[0251] 在一些方面,所提供的测定法和分析方法可以包括另外的、进一步的或下游的分析,包括联合或应用其他基因组分析或功能分析的步骤。在一些方面,进一步或下游的分析可用于与分析或应用步骤中的任何一个或多个联合。在一些方面,示例性的进一步或下游分析步骤可包括阈值化和聚类方法、预测建模、基因本体论(GO)分析、基序分析和/或主成分分析(PCA)。在一些方面,可以计算实施任何下游分析方法。

[0252] 在一些方面,进一步或下游分析是PCA。在一些方面,PCA可用于降低数据的维度并检查数据中的广义方差,并且鉴定促成变异的关键驱动因素。PCA利用正交变换将可能相关变量的一组观察值转换为被称为主成分的线性不相关变量的一组值。在一些方面,将变换定义为使得第一主成分具有最大可能方差(例如,尽可能多地解释数据中的变异性),并且每个后续成分依次具有在约束下的最高可能方差,其与前面的成分正交。

[0253] 在一些方面,进一步或下游分析可包括生物途径分析,例如包括基因网络、途径、生物过程和分子功能的分析。在一些方面,基于生物、分子和/或功能关系,可以使鉴定的表观遗传学特性、谱和/或一个或多个基因组区域、基因座区域和/或区间类似于特定的生物过程或途径。在一些方面,这些关系对于基于途径或过程的折叠关联(collapsed association)方法或推断特定的表观遗传学特性和/或谱的表型影响是有用的。示例性的生物途径分析包括但不限于:提供生物关系的最近基因的反应组学途径和基因本体论(GO)生物过程;用于提供表型关系的疾病本体论;以及提供分子和功能关系的蛋白质结构域信息和分子功能(通过基因本体论注释)。其他示例性生物途径分析包括Ingenuity途径分析(IPA)、Pathway Studio途径分析、基于KEGG的分析、基因集富集分析(GSEA)、信号传导途径影响分析(SPIA)、EnrichNet、GGEA、信号传导途径影响分析(SPIA)和TopoGSA。可用于生物途径分析的示例性程序和方法包括过代表分析或富集分析(ORA)、功能归类评分(FCS)和途径拓扑(PT)。

[0254] 在一些方面,进一步或下游分析可包括聚类分析,例如,层次聚类。在一些方面,层次聚类涉及创建具有从顶部到底部或从底部到顶部的预定次序的聚类。在一些方面,层次聚类可以是凝聚的(自底向上),其中每个观察在其自身的聚类中开始,并且当一个聚类向上移动层次时合并聚类对;或者是分裂的(自顶向下),其中所有观察在一个聚类中开始,并且当一个聚类向下移动层次时,递归地执行分割。可以在来自三个或更多个不同样品的一个或多个基因组位置的表观遗传学/表观基因组学谱上使用层次聚类。在一些方面,可以基于基因的集合、集或模块的表观遗传学谱(例如使用每个基因的度量值)来执行层次聚类。在一些实施方案中,层次聚类可以鉴定与聚类外的样品相比显示出彼此更相似的表观遗传学谱的样品的组或聚类。在一些实施方案中,可以使用计算实施的脚本、工具或软件包(比如JMP或Pheatmap R)来执行层次聚类。

[0255] 在一些方面,进一步或下游分析可包括基序分析,例如,确定重复或共有序列基序。在一些方面,进一步或下游分析可包括转录因子结合基序分析和/或转录因子占据分析。在一些实施方案中,可以使用诸如homer_find_motifs的脚本或工具来执行转录因子结合基序分析。在一些方面,脚本或工具可以搜索使用本文中提供的方法确定的峰集内共享的富集的转录因子结合基序。

[0256] 在一些实施方案中,可以使用聚类分析连同—个或多个参考样品和/或—个或多个不同的样品来鉴定—个或多个基因组区域,例如可以在表征和/或评估特定细胞和/或将治疗的受试者中在基因座的集合中使用的基因组基因座。例如,在一些实施方案中,对于来自表现出不同结果的受试者(例如,表现出完全缓解的一些受试者和表现出部分缓解、疾病进展或无缓解的一些受试者)的细胞的表观遗传学和/或表观基因组学分析的数据,可以进行层次聚类或主成分分析(PCA),以鉴定与期望的结果(例如,完全缓解、毒性降低、扩增增加)相关的特定的基因座集合。在一些实施方案中,来自具有期望的结果(例如,完全缓解、毒性降低、扩增增加)的受试者的细胞是用于聚类分析和/或确定用于比较的参考值的参考样品。在这些情况的一些中,任何类型的表观遗传学和/或表观基因组学分析都可用于鉴定与期望的结果(例如,完全缓解、毒性降低、扩增增加)相关的特定基因座。一旦鉴定基因座和/或基因座的集合,就可以确定测试细胞群的表观遗传学和/或表观基因组学特性或模式,并将其与与期望的结果相关的样品或组中的相应特性或模式进行比较,并且可以选择

细胞和/或受试者以便进行治疗,例如使用过继细胞疗法进行治疗。在一些实施方案中,可以确定一个或多个基因组区域(例如基因组基因座)的参考阈值(例如,FPKM或RPKM阈值)并将其用作选择测试细胞和/或患者的阈值。在一些实施方案中,对测试细胞和/或受试者中的一个或多个基因组区域(例如,基因组基因座)的表观遗传学和/或表观基因组学特性进行进一步的聚类分析,并且如果所述一个或多个基因组区域(例如基因组基因座)的表观遗传学和/或表观基因组学特性与具有期望的结果(例如,完全缓解、毒性降低、扩增增加)的参考细胞或组聚类,则所述细胞和/或受试者被选择为用于进行治疗。

[0257] 在一些实施方案中,可以基于一个或多个基因组区域(例如基因组基因座)的基因组区间的平均和/或总FPKM或RPKM值来执行聚类分析。在一些实施方案中,基于一个或多个基因组区域(例如基因组基因座)的基因体(例如,编码区)上的FPKM或RPKM值的总和来执行聚类。在一些实施方案中,聚类分析被描绘为树状图、星座图或决策树。在一些实施方案中,基于特定点或特定区间(例如,与编码区相关或与非编码区相关的那些)处的值(例如,FPKM或RPKM)来执行聚类分析。

[0258] 在一些实施方案中,可以使用各种绘图或图表来表示或可视化本文中提供的分析结果和方法。示例性绘图或图表包括venn图、MA图(M(对数比)和A(平均值))、主成分分析(PCA)图、箱形图和热图。

[0259] 在一些实施方案中,可以对细胞进行另外的分析,例如分子、表型或功能分析。例如,可以使用荧光激活细胞分选(FACS)和/或激光捕获显微切割(LCM)分析细胞样品,进行表型分析。在一些实施方案中,可以将细胞样品和/或核酸分成多个部分。可以基于分子标签(例如荧光标签)来划分所述部分。在一些实施方案中,可以分选细胞样品和/或核酸。可以在将分子标签插入到核酸中之后进行分选。可以在将片段测序之前进行分选。还可以使用诸如荧光原位杂交(FISH)的技术分析细胞样品的基因转录。染色质可及性可与表型、转录或翻译分析相关联。

[0260] 在一些实施方案中,表观遗传学和/或表观基因组学分析可以与其他检测和/或分析方法联合使用,比如转录分析、转录组分析、转录因子占据测定法、RNAseq、蛋白质表达、蛋白质组学分析、蛋白质修饰分析、功能活性测定法、流式细胞术和/或细胞内细胞因子染色。在一些实施方案中,确定在来自一种或多种表观遗传学和/或表观基因组学分析(例如ATAC-seq)和其他分析方法(例如ChIPseq、RNAseq分析或细胞内细胞因子染色)的结果之间的相关性。在一些实施方案中,来自各种测定法或分析方法的结果是强相关的,例如具有高的相关系数。在一些实施方案中,来自各种测定法或分析方法的结果不是强相关的,例如具有低的相关系数。在一些实施方案中,相关性取决于基因和/或基因座和/或实验条件。在一些实施方案中,分析方法包括比较、关联和/或进一步分析所获得的数据与其他数据集,例如其他表观遗传学和/或表观基因组学数据集和/或其他表型或生理数据集。例如,分析方法包括比较、关联和/或进一步分析所获得的数据与全基因组表达数据(例如RNA-seq数据)或全基因组核小体定位数据。

[0261] 在一些实施方案中,表观遗传学和/或表观基因组学分析包括确定翻译后组蛋白修饰。在一些实施方案中,表观遗传学分析包括检测一种或多种组蛋白修饰,比如组蛋白H1、H2A、H2B、H3和/或H4中特定残基的乙酰化、甲基化、磷酸化、泛素化、小泛素化(sumoylation)和生物素化。

基因集合分析

[0262] 在一些方面,可以在一个或多个选定的基因组区域或基因座上进行任何分析方法,包括另外的、进一步的或下游的分析。在一些方面,可以基于一种或多种选择的基因比如基因子集、基因集合或基因模块来进行所述方法。在一些方面,任何分析方法可用于评估在特定基因或在基因的特定集合、模块或集(例如,与相似活性和/或功能相关和/或被共调节或共表达的基因的集合、模块或集)处或附近的基因组区域、基因组基因座中的表观遗传学/表观基因组学状态。在一些方面,可以对预先选择的基因或基因的集合、模块或集进行分析。在一些方面,可以进行分析,使得表现出相似的表观遗传学和/或表观基因组学特性、状态和/或模式的基因可以用于鉴定感兴趣的基因或基因的集合、模块或集,所述基因例如关联、关连或指示结果或特点(例如,期望的结果或特点)的那些基因。在一些方面,这样的基因的集合、模块或集可用于评估测试样品。

[0263] 在一些实施方案中,在评估基因的集合、模块或集的表观遗传学/表观基因组学状态中用于分析的基因组区域、基因座和/或区间可包括在基因的特定元件或部分、或与基因相关的基因组区域或元件处的信号,其包括内含子、外显子、顺式调控元件、启动子、增强子、上游激活序列(UAS)、3'非翻译区(UTR)、5'UTR、非编码RNA产生区域、非编码RNA(ncRNA)基因、miRNA基因、siRNA基因、piRNA基因、snoRNA基因、lncRNA基因、核糖体RNA(rRNA)基因、小RNA结合位点、非编码RNA结合位点、假基因或转录终止位点(TTS)。

[0264] 在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,可位于与细胞类型规范、T细胞分化和/或发育或转录因子相关的基因内或其附近。在一些实施方案中,一个或多个基因组区域,例如基因组基因座,是编码与特定的细胞表型相关的生物标志物或指示物(例如特定的免疫细胞表型或免疫细胞分化标志物)的基因座。例如,在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,包括与T细胞分化和/或功能相关的标志物,例如细胞因子基因、免疫调节蛋白基因、细胞表面标志物、细胞凋亡标志物、细胞死亡标志物和/或转录调节因子。在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,与特定的免疫细胞表型相关,比如T细胞表型,例如幼稚T细胞表型、初始效应子表型、短期或中期/晚期效应子表型,例如在Best等人,Nature Immunology (2013) 14, 404-412中描述的任何表型,将其通过提述并入本文。在一些实施方案中,受试者是人,并且鉴定的基因座的根据是人基因组。

[0265] 在一些方面,可以基于基因的集合、模块或集进行分析,所述基因例如被描述为与已知途径相关或受到共调节的基因或基因的组。在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,包括一个或多个选择的基因座,其与T细胞或T细胞(例如CD8⁺ T细胞)的活性、刺激、功能和/或特定表型相关,例如在Best等人,Nature Immunology (2013) 14, 404-412中描述的任何一种,将其通过提述并入本文。例如,在一些方面,可以评估基因的集合、模块或集,其被鉴定为涉及或指示T细胞分化状态、效应子功能、耗尽或特异性记忆功能和/或特异性T细胞亚群。

[0266] 在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,包括选自以下的一个或多个免疫调节基因座:Ifng、Tbet、Lag3、Pd1、Cd25、Foxp3、Il2、Ki67、Tnf、Il13、Il17、Il17a、Foxp和Foxp3。在一些实施方案中,一个或多个基因组区域,例如基因组基因座,例如基因座的集合,包括选自以下的一个或多个基因座:Ifng、Tbx21、

Tnf、Il13、Il2ra、Lag3、Il2、Mki67、Il17a、Pdc1和Foxp3。在一些实施方案中，基因座的集合包括Ifng和Pdc1。在一些实施方案中，基因座的集合包括与初始效应子应答相关的一个或多个基因座，比如Ctla4、IL2ra、Ifng、Gzmb和Il2。在一些实施方案中，基因座的集合包括与细胞分裂相关的一个或多个基因座，比如Myc、Id3、Egr2、Tnf、Cd69和Pkm2。在一些实施方案中，基因座的集合包括与细胞周期和细胞分裂相关的一个或多个基因座，比如Myb、Hist1h3a、Cdk1和Ccd45。在一些实施方案中，基因座的集合包括与幼稚和晚期记忆应答相关的一个或多个基因座，比如Sell、Nsg2、S1fn5和Cnr2。在一些实施方案中，基因座的集合包括与早期效应子和晚期记忆相关的一个或多个基因座，比如Ly6a、Rpl和Snora。在一些实施方案中，基因座的集合包括与短期效应子应答相关的一个或多个基因座，比如Id2、Cxcr3、Zeb2、Cx3cr1、Klrg1、S1pr5和Itgam。在一些实施方案中，基因座的集合包括与记忆前体应答相关的一个或多个基因座，比如Bcl2、Tcf7、Il17r和Foxo3。在一些实施方案中，基因座的集合包括与幼稚细胞或晚期效应子应答相关的一个或多个基因座：Cxcr6、S1pr4、S1pr1、Klf2和Klf3。在一些实施方案中，基因座的集合包括与短期效应子或记忆应答相关的一个或多个基因座，比如Prdm1和Hif2a。在一些实施方案中，基因座的集合包括与晚期效应子应答相关的一个或多个基因座，比如Tbx21、Prf1、Bhlhe40、Cd44、Klrc2和Il12rb1。在一些实施方案中，基因座的集合包括与中期或晚期效应子应答相关的一个或多个基因座，比如Dmrta1、Bcl2、Edaradd、Prss12、Cnrip1和Aqp9。在一些实施方案中，基因座的集合包括与晚期效应子应答相关的一个或多个基因座，比如Unc5a、Xcl1、Yes1、Cdh1、Myo3b、Dock9和Atn1。在一些实施方案中，基因座的集合包括选自以下的一个或多个基因座：Nr4a1、Cblb、Irf4、Tbx21、Eomes、Ifng、Csf2、Gzmb、Tnfsf10、Gata3、Mir155、Sox21、Ctla4、Lag3、Pdc1 Actb和/或或者Gapdh。在一些实施方案中，基因座的集合包括管家基因，例如Actb或Gapdh，其可以指示细胞的活化状态。

[0267] 在一些实施方案中，如果将要施用的细胞（例如用于过继细胞疗法的工程化细胞）表现出与效应子样T细胞不同的表观遗传学和/或表观基因组学特性，则基于本文中提供的方法选择所述细胞用于治疗。在一些实施方案中，如果来自受试者的细胞表现出与效应子样T细胞不同的表观遗传学和/或表观基因组学特性，则选择所述受试者用工程化细胞治疗。例如，如果与T细胞的效应子功能或效应子样表型相关的一个或多个基因座表现出相对较低的染色质可及性水平，则可以选择细胞和/或受试者进行治疗。

转基因整合分析

[0268] 在一些实施方案中，所述方法可以包括鉴定和/或表征递送至细胞或引入到细胞中的核酸分子或构建体或转基因（例如引入到细胞中用于工程化例如用于表达重组受体的构建体）的整合位点。在一些实施方案中，基于核酸序列，例如含有与本文所述方法获得的整合构建体序列邻近的核酸序列，可以确定一个或多个整合位点。在一些实施方案中，可以确定整合的数目和位置。在一些实施方案中，可以确定在核酸整合位点处的表观遗传学和/或表观基因组学谱。在一些实施方案中，用于整合的核酸分子或构建体包括载体，例如病毒载体、核酸分子、转座子和/或编码转基因例如重组受体的核酸构建体。在一些实施方案中，所述方法可包括确定整合的核酸分子或构建体的数目和/或丰度。例如，在一些实施方案中，可以确定在已经被工程改造的细胞中例如在被工程改造为表达重组受体的细胞中的核酸（例如载体，比如病毒载体，或构建体）的拷贝数。在一些实施方案中，所述方法包括确定

编码的基因例如编码的重组受体的表达水平。

[0269] 在一些实施方案中,所述方法可包括确定多个T细胞或T细胞群(例如多个工程化T细胞和/或含有多个T细胞或T细胞群例如多个工程化T细胞的组合物)中的T细胞克隆性。在一些方面,所述方法可用于通过确定在基因座处或附近的表观遗传学和/或表观基因组学特性和/或表观遗传学和/或表观基因组学谱(比如染色质可及性)来确定多个T细胞或T细胞群中的克隆性,所述基因座编码T细胞受体(TCR)基因,例如编码TCR α 、 β 、 γ 和/或 δ 链的可变区和/或恒定区的基因。在一些方面,所述方法可用于通过确定核酸分子(例如编码重组受体的核酸序列)的插入位点来确定多个T细胞或T细胞群中的克隆性。在一些实施方案中,克隆性评估可用于表征多个T细胞或T细胞群,例如组合物中的多个T细胞或T细胞群。在一些方面,这样的确定可用于确定治疗方案,例如,包括工程化细胞的施用剂量、时间和/或频率,和/或用于确定培养或工艺条件的必要或有利的修改。

示例性分析脚本

[0270] 不希望受命名法的约束,提供了示例性序列操作和比对程序的描述,例如在本文中提述的脚本。这个列表并不意图是穷尽的或限定的,而是此处列出的脚本用作可能对提供的分析有用的脚本类型的例子。

[0271] 在一些实施方案中,这些脚本和/或步骤可以并行执行和/或以分支方式加以组织,使得可以在相同的时间以任何顺序执行多个步骤。

[0272] `get_bcl`:从中间存储位置(在一些实施方案中,在基于云的存储上)检索来自高通量测序器的压缩原始数据(处于**.bcl**文件形式的碱基调用)并下载到计算集群。

[0273] `untar_bcls`:来自定序器运行的数据被解压缩,以便使用下游工具处理。

[0274] `bcl2fastq`:软件工具**bcl2fastq**用于将原始碱基调用文件转换为**gzip**压缩的**FASTQ**文件,所述**FASTQ**文件包含测序运行中每个读段的序列和元数据。在示例性实施方案中,对于来自**ATAC-Seq**的配对末端运行,每个样品具有两个文件**R1**和**R2**,分别对应于正向和反向读段(每个DNA片段的开始和结束)。

[0275] `get_data`:检索来自前一步骤的压缩的**FASTQ**文件并将其移动到根据样品的分析目录中。

[0276] `unzip`:将压缩的**FASTQ**文件解压缩,以便下游工具可以处理它们。

[0277] `fastqc`:生成一组关于测序运行和碱基调用质量、以及污染或过度聚类的潜在迹象的整体统计数据报告。

[0278] `get_genome`:获取指定生物和版本的适当基因组文件。在一些实施方案中,这可以包括在线来源和/或数据库。

[0279] `build_bowtie2_index`:使用被称为**bowtie2_build**的软件工具标引基因组文件,使得读段可以在位置上被映射到基因组。

[0280] `Index_genome_fasta`:为下载的基因组文件建立**FASTA**索引以允许由下游工具进行位置访问。

[0281] `map_atac_reads`:使用映射软件**bowtie2**将**ATAC-seq**读段的序列映射回基因组以确定其位置。

[0282] `Picard_remove_duplicates`:使用软件套件(例如**Picard**)来标记和去除由**PCR**扩增偏倚和集群错误调用引起的重复读段以减少数据集中的噪声量。在一些情况下,这个步

骤可确保对映射后读段的正确排序命名和标引,并输出给出每染色体的映射计数的idxstats文件。

[0283] `atac_insert_size_metrics`:使用软件(例如Picard)来计算和可视化从数据集中恢复的片段大小的分布。

[0284] `atac_alignment_summary`:使用软件(例如Picard)来计算与读段在各种类型的基因组特征中的位置分布有关的统计数据。

[0285] `filter_mtDNA_reads`:过滤出线粒体DNA。

[0286] `shift_atac_alignments`:在ATAC-seq文库制备过程中,将映射片段的位置移位,以便通过Tn5转座酶导致4个或5个碱基对插入。

[0287] `Filter_nucleosomal_fragments`:过滤出大于100bp的片段。在一些实施方案中,预期这些片段可代表核小体结合的染色质而不是游离染色质。这个步骤能够富集信号,并且过滤和未过滤的读段都将用于下游分析以分别推断开放染色质和核小体占据的染色质。

[0288] `make_homer_tagdir`:将BAM比对文件转换为下游HOMER工具可使用的文件格式。

[0289] `homer_findPeaks`:可以被采用的使用HOMER找到峰的步骤。

[0290] `homer_find_motifs`:搜索在通过MACS和HOMER发现的峰集内共享的富集的转录因子结合基序。

[0291] `homer_annotate_peaks`:利用包括可能与其相关的最近基因的附加元数据来注释峰。在一些实施方案中,这个步骤可以被另一个注释策略代替。

[0292] `get_gtf_annotation`:检索与生物体相关的转录组注释文件和流水线正在运行的基因组版本。

[0293] `gtf_coding_transcripts_only`:仅有蛋白质编码转录物的转录组注释的子集

[0294] `homer_calculate_per_gene_accessibility_norm`:计算代表基因的一般可及性的归一化的每基因可及性值(例如,归一化的FPKM值)。

[0295] `make_homer_ucsc_file`:创建.bedGraph文件。在一些实施方案中,这允许全基因组的片段计数堆积。

[0296] `homer_bedgraph_to_bigwig`:将bedGraph文件转换为二进制压缩的bigWig文件。在一些情况下,这个文件类型被基因组浏览器使用,以便将跨基因组的片段覆盖可视化。

[0297] `macs_callpeaks`:使用MACS2调用可及性峰。在一些方面,峰包括通过量化ATAC-seq片段测量的富集或耗竭可及性和/或占据信号的基因组区域。在一些实施方案中,峰区域的大小可以在10和10,000bp之间。样品之间的峰的比较可以在下游加以使用以鉴定条件之间的活性调控元件(任选地,与特定基因或转录因子相关),和/或将其用于鉴定特定细胞状态或预测细胞疗法的结果、细胞疗法的性能、细胞组合物或培养物的毒性和/或其他特点标记。

[0298] `nucleoatac_run`:运行NucleoATAC,其为一种能够调用峰区内转座酶插入位点并将其计数并且输出关于来自前面步骤的数据的汇总表的工具。

[0299] 可以在分析流水线的框架内使用另外的下游分析,包括阈值化和聚类方法、预测建模和/或差异性峰调用。

治疗结果

[0300] 在所提供的方法的一些实施方案中,细胞群的基因组的表观遗传学特性(例如染

色质可及性)与(或者可以与)根据所提供的方法的细胞疗法的治疗结果的发生相关,所述细胞群例如被包含在用重组受体基因工程改造或已经用重组受体基因工程改造的细胞的组合物中。

[0301] 在所提供的方法的一些实施方案中,细胞群的基因组的表观遗传学特性(例如染色质可及性)可以预测和/或诊断或检测根据所提供的方法的细胞疗法的治疗结果发生的可能性,所述细胞群例如被包含在用重组受体基因工程改造或已经用重组受体基因工程改造的细胞的组合物中。

[0302] 在一些实施方案中,受试者已经、正在接受或将接受治疗,比如细胞疗法,例如用于治疗受试者的疾病或病症。例如,在一些实施方案中,细胞疗法是过继细胞疗法,包括涉及施用表达对感兴趣的疾病或疾患特异的嵌合受体比如嵌合抗原受体(CAR)和/或其他重组抗原受体的细胞的疗法,以及其他过继免疫细胞和过继T细胞疗法。在一些实施方案中,过继细胞疗法包括施用一定剂量的表达重组受体比如CAR或其他重组抗原受体的细胞。在一些实施方案中,嵌合受体(比如嵌合抗原受体)含有一个或多个结构域,其将配体结合结构域(例如抗体或抗体片段)与细胞内信号传导结构域相结合,所述配体结合结构域提供针对期望的抗原(例如肿瘤抗原)的特异性。在一些实施方案中,细胞内信号传导结构域是活化细胞内结构域部分,比如T细胞活化结构域,其提供初级活化信号。在一些实施方案中,细胞内信号传导结构域含有或另外含有促进效应子功能的共刺激信号传导结构域。在一些实施方案中,当经基因工程化进入免疫细胞中时,嵌合受体可调节T细胞活性,并且,在一些情况下,可调节T细胞分化或稳态,由此产生具有改进的体内寿命、存活和/或持久性的基因工程化细胞,以供例如在过继细胞治疗方法中使用。

毒性结果

[0303] 在一些实施方案中,可以评估或监测在受试者中施用治疗剂(例如CAR T细胞)的毒性结果。在一些实施方案中,毒性结果是毒性事件的存在或与之相关,所述毒性事件比如细胞因子释放综合征(CRS)、重度CRS(sCRS)、巨噬细胞活化综合征、肿瘤溶解综合征、至少或约38摄氏度的持续三天或更多天的发热、至少或约20mg/dL的血浆C反应蛋白(CRP)水平、神经毒性和/或重度神经毒性。在一些实施方案中,毒性结果是体征或症状、特定体征和症状和/或其数量或程度,其存在或不存在的可以指定受试者中的毒性的具体程度、严重性或水平。指定或确定与治疗剂(例如CAR-T细胞)的不期望的毒性结果有关的特定体征、症状和/或其数量或程度在本领域技术人员的水平之内。

[0304] 在一些实施方案中,毒性结果是与毒性事件相关的指标。在一些实施方案中,毒性结果是一种或多种生物标志物的存在或不存在的存在,或一种或多种生物标志物水平的存在或不存在的存在。在一些实施方案中,生物标志物是存在于血清或其他体液或组织中的指示细胞因子释放综合征(CRS)、重度CRS或CRS相关结果的分子。在一些实施方案中,生物标志物是存在于血清或其他体液或组织中的指示神经毒性或重度神经毒性的分子。

[0305] 在一些实施方案中,如果毒性事件,比如CRS相关结果,例如,如果CRS指标或其他毒性生化指标的血清水平相比于基线或治疗前水平(比如刚好在施用第一剂量的治疗剂之前的指标的血清水平)高或高约10倍、高或高约15倍、高或高约20倍、高或高约25倍、高或高约50倍、高或高约75倍、高或高约100倍、高或高约125倍、高或高约150倍、高或高约200倍、或高或高约250倍,则受试者表现出毒性或毒性结果。

[0306] 在一些方面,毒性结果是或关联或指示细胞因子释放综合征(CRS)或重度CRS(sCRS)。在过继T细胞疗法和对受试者施用其他生物制品之后的一些情况下,可能发生CRS,例如sCRS。参见Davila等人,Sci Transl Med 6, 224ra25 (2014); Brentjens等人,Sci. Transl. Med. 5, 177ra38 (2013); Grupp等人,N. Engl. J. Med. 368, 1509-1518 (2013); 和 Kochenderfer等人,Blood 119, 2709-2720 (2012); Xu等人,Cancer Letters 343 (2014) 172-78。

[0307] 通常,CRS由例如由T细胞、B细胞、NK细胞、单核细胞和/或巨噬细胞介导的全身性免疫应答扩大引起。这样的细胞可释放大量炎症介质,比如细胞因子和趋化因子。细胞因子可激发急性炎症反应和/或诱导内皮器官损伤,这可导致微血管渗漏、心力衰竭或死亡。重度、危及生命的CRS可导致肺浸润和肺损伤、肾衰竭或弥散性血管内凝血。其他重度、危及生命的毒性可包括心脏毒性、呼吸窘迫、神经毒性和/或肝衰竭。

[0308] 在施用表达CAR的细胞的背景下,CRS通常在输注表达CAR的细胞之后6-20天发生。参见Xu等人,Cancer Letters 343 (2014) 172-78。在一些情况下,CRS在CAR T细胞输注之后不到6天或超过20天发生。CRS的发生率和时间可能与输注时的基线细胞因子水平或肿瘤负荷有关。通常,CRS涉及干扰素(IFN)- γ 、肿瘤坏死因子(TNF)- α 和/或白细胞介素(IL)-2的血清水平升高。在CRS中可快速诱导的其他细胞因子是IL-1 β 、IL-6、IL-8和IL-10。

[0309] 与CRS相关的示例性体征或症状包括发热、寒战、恶寒、低血压、呼吸困难、急性呼吸窘迫综合征(ARDS)、脑病、天冬氨酸转氨酶(AST)/丙氨酸转氨酶(ALT)升高、肾衰竭、心脏疾患、缺氧、神经系统障碍和死亡。神经系统并发症包括谵妄、癫痫发作样活动、意识模糊、找词困难、失语和/或变得迟钝。其他的CRS相关体征或结果包括疲劳、恶心、头痛、癫痫发作、心动过速、肌痛、皮疹、急性血管渗漏综合征、肝功能损害和肾衰竭。在一些方面,CRS与诸如血清-铁蛋白、d-二聚体、转氨酶、乳酸脱氢酶和甘油三酯的一种或多种因子的增加相关,或伴有低纤维蛋白原血症或肝脾肿大。

[0310] 在一些实施方案中,与CRS相关的体征或症状包括以下一种或多种:持续性发热,例如指定温度,例如高于或高于约38摄氏度持续两天或更多天(例如三天或更多天,例如四天或更多天或持续至少连续三天)的发热;高于或高于约38摄氏度的发热;细胞因子(例如IFN γ 或IL-6)的升高;和/或毒性的至少一种临床症状,比如低血压(例如,通过至少一种静脉内血管加压器测量);缺氧(例如,低于或低于约90%的血浆氧(PO_2)水平);和/或一种或多种神经系统疾患(包括精神状态改变、迟钝和癫痫发作)。

[0311] 示例性CRS相关结果包括一种或多种因子(包括细胞因子和趋化因子以及与CRS相关的其他因子)的血清水平升高或高血清水平。示例性结果进一步包括一种或多种这样的因子的合成或分泌的增加。这样的合成或分泌可以通过T细胞或与T细胞相互作用的细胞(比如先天免疫细胞或B细胞)来执行。

[0312] 在一些实施方案中,在CAR治疗之前、期间或之后监测一种或多种炎性标志物,例如细胞因子或趋化因子。在一些方面,一种或多种细胞因子或趋化因子包括IFN- γ 、TNF- α 、IL-2、IL-1 β 、IL-6、IL-7、IL-8、IL-10、IL-12、sIL-2R α 、粒细胞巨噬细胞集落刺激因子(GM-CSF)或巨噬细胞炎性蛋白(MIP)。在一些实施方案中,监测IFN- γ 、TNF- α 和IL-6。

[0313] 在一些实施方案中,一种或多种生物标志物的存在指示毒性事件比如CRS或神经毒性的等级、严重性或程度。在一些实施方案中,毒性结果是毒性事件的特定等级、严重性

或程度,比如CRS或神经毒性的特定等级、严重性或程度。在一些实施方案中,关于一定等级、严重性或程度的毒性事件的存在可以是剂量限制性毒性。在一些实施方案中,毒性事件的不存在或低于一定等级、严重性或程度的毒性事件的存在可以指示剂量限制性毒性的不存在。

[0314] 已经开发了显得与CRS的发作相关的CRS标准,从而预测哪些患者更可能具有发生sCRS的风险(参见Davilla等人,Science translational medicine.2014;6(224):224ra25)。因素包括发热、缺氧、低血压、神经系统改变、炎性细胞因子的血清水平升高,其治疗诱导的升高可与治疗前肿瘤负荷和sCRS症状良好相关。关于CRS的诊断和管理的其他指南是已知的(参见例如Lee等人,Blood.2014;124(2):188-95)。在一些实施方案中,反映CRS等级的标准是下表1中详述的标准。

等级	症状描述
1 轻度	未危及生命,只需要对症治疗,比如解热药和止吐药(比如发热、恶心、疲劳、头痛、肌痛、不适)
2 中度	需要适度干预并且对干预有反应: 需氧量<40%, 或 对流体或低剂量的单一血管加压药有反应的低血压, 或 2级器官毒性(根据 CTCAE v4.0)
3 重度	需要积极干预并且对干预有反应: 需氧量≥40%, 或 需要高剂量的单一血管加压药(例如,去甲肾上腺素≥20μg/kg/min,多巴胺≥10μg/kg/min,去氧肾上腺素≥200μg/kg/min,或肾上腺素≥10μg/kg/min)的低血压, 或 需要多种血管加压药(例如,加压素+上述药物之一,或相当于≥20μg/kg/min去甲肾上腺素的联合血管加压药)的低血压, 或 3级器官毒性或4级转氨酶升高(根据 CTCAE)

	v4.0)
4 危及生命	危及生命： 需要呼吸机支持，或 4级器官毒性(不包括转氨酶升高)
5 致命的	死亡

[0315] 在一些实施方案中，毒性结果是重度CRS。在一些实施方案中，毒性结果是不存在重度CRS (例如中度或轻度CRS)。在一些实施方案中，重度CRS包括等级为3或更高的CRS，如表1中列出。在一些实施方案中，重度CRS包括等级为2或更高的CRS，如2级、3级、4级或5级CRS。

[0316] 在一些实施方案中，通过ELISA测量毒性结果例如CRS相关结果的水平，例如CRS指标的血清水平。在一些实施方案中，可以测量发热和/或C反应蛋白 (CRP) 水平。在一些实施方案中，发热且CRP \geq 15mg/dL的受试者可被认为具有发生重度CRS的高风险。在一些实施方案中，CRS相关血清因子或CRS相关结果包括炎性细胞因子和/或趋化因子的水平和/或浓度的升高，所述细胞因子和/或趋化因子包括Flt-3L、不规则趋化因子 (fracktalkine)、粒细胞巨噬细胞集落刺激因子 (GM-CSF)、白细胞介素-1 β (IL-1 β)、IL-2、IL-5、IL-6、IL-7、IL-8、IL-10、IL-12、干扰素 γ (IFN- γ)、巨噬细胞炎性蛋白 (MIP)-1、MIP-1、sIL-2R α 或肿瘤坏死因子 α (TNF α)。在一些实施方案中，因子或结果包括C反应蛋白 (CRP)。除了作为CRS的早期且易于测量的风险因子之外，CRP还是细胞扩增的标志物。在一些实施方案中，测量出具有高水平CRP (如 \geq 15mg/dL) 的受试者具有CRS。在一些实施方案中，测量出具有高水平CRP的受试者没有CRS。在一些实施方案中，CRS的量度包括CRP和指示CRS的另一因子的量度。

[0317] 在一些方面，毒性结果是神经毒性或与之相关。在一些实施方案中，与神经毒性的临床风险相关的体征或症状包括意识模糊、谵妄、表达性失语、迟钝、肌阵挛、嗜睡、精神状态改变、惊厥、癫痫发作样活动、癫痫发作 (任选地通过脑电图 (EEG) 证实)、 β 淀粉样蛋白 (A β) 水平升高、谷氨酸水平升高和氧自由基水平升高。在一些实施方案中，基于严重性对神经毒性进行分级 (例如，使用1-5级量表 (参见例如Guido Cavaletti&Paola Marmioli Nature Reviews Neurology 6,657-666 (2010年12月);National Cancer Institute—Common Toxicity Criteria版本4.03 (NCI-CTCAE v4.03))。在一些实施方案中，如果在施用后受试者显示出限制自理 (例如，洗澡、穿衣和脱衣、吃饭、如厕、服药) 的症状，则认为受试者响应于或继发于施用细胞疗法或其细胞剂量而产生了“重度神经毒性”，所述症状为：1) 外周运动神经病的症状，包括外周运动神经的炎症或变性；2) 外周感觉神经病的症状，包括外周感觉神经的炎症或变性、感觉迟钝，比如感觉失真，导致异常和不愉快的感觉，神经痛，比如沿着神经或一组神经的剧烈疼痛感，和/或感觉异常，比如感觉神经元的功能障碍，导致在没有刺激的情况下的刺痛、麻木、压力、寒冷和温暖的异常皮肤感觉。在一些实施方案中，重度神经毒性包括等级为3或更高的神经毒性，如表2中列出。在一些实施方案中，重度神经毒性包括等级为2或更高的神经毒性，如2级、3级、4级或5级神经毒性。

等级	症状描述
1 无症状或轻度	轻度或无症状
2 中度	限制诸如准备饭菜、购买杂货或衣服、使用电话、管理钱财等工具性日常生活活动(ADL)的症状的存在
3 重度	限制诸如洗澡、穿衣和脱衣、自己吃饭、如厕、服药等自理性 ADL 的症状的存在
4 危及生命	危及生命的症状, 需要紧急干预
5 致命的	死亡

[0318] 在一些实施方案中,毒性结果是剂量限制性毒性。在一些实施方案中,毒性结果是不存在剂量限制性毒性。在一些实施方案中,剂量限制性毒性 (DLT) 被定义为任何3级或更高等级的毒性,如通过任何已知或公布的用于评估特定毒性(比如本文所述的任何毒性)的指南所评估的,所述指南包括国家癌症研究所 (NCI) 的不良事件常用术语标准 (CTCAE) 4.0 版。

反应结果

[0319] 在一些实施方案中,治疗结果是对细胞疗法的反应或功效结果、细胞疗法的毒性结果、对细胞疗法的免疫原性应答或是细胞疗法的另一特点或特征(例如,细胞在受试者中的持久性或扩增)。在某些实施方案中,治疗结果是细胞疗法的治疗或预防功效结果。在一些实施方案中,细胞疗法的反应结果包括在特定剂量的功效结果。在一些方面,细胞疗法的治疗结果包括实现反应所需的细胞疗法的剂量。在一些方面,评估了细胞的持久性和/或扩增。

[0320] 在一些实施方案中,可以监测或评估对受试者施用细胞疗法的反应结果。在一些实施方案中,细胞疗法的反应结果是完全缓解 (CR)。在一些实施方案中,通过监测受试者中的疾病负荷来评估反应结果。在一些实施方案中,可以评估无反应、部分缓解或临床或完全缓解的存在。在一些实施方案中,细胞疗法的反应结果是反应的耐久性。

[0321] 在一些实施方案中,部分缓解 (PR) 或完全缓解 (CR) 是其中治疗剂降低或阻止受试者中疾病或病症的扩张或负荷的缓解。例如,在疾病或病症是肿瘤的情况下,与用治疗剂(例如CAR T细胞)治疗之前相比,如果在骨髓或可分子检测的癌症中的肿瘤大小、体积、转移、胚细胞的百分比有降低和/或有预后或存活或与肿瘤负荷相关的或其他症状的改善,则降低的疾病负荷存在或是存在的。

[0322] 在一些方面,受试者(比如患有NHL的受试者)的缓解率的根据是Lugano标准。

(Cheson等人, (2014) JCO 32 (27) :3059-3067; Johnson等人, (2015) Radiology 2:323-338; Cheson, B.D. (2015) Chin Clin Oncol 4 (1) :5)。在一些方面, 缓解评估利用临床、血液学和/或分子方法中的任何一种。在一些方面, 使用Lugano标准评估的缓解涉及在适当时使用正电子发射断层摄影 (PET) - 计算机断层摄影 (CT) 和/或CT。对于FDG浓聚的淋巴瘤, PET-CT评估可进一步包括使用氟脱氧葡萄糖 (FDG)。在一些方面, 在PET-CT将用于评估在FDG浓聚组织学方面的缓解的情况下, 可以使用5点量表。在一些方面, 5点量表包括以下标准: 1、在背景以上没有摄取; 2、摄取 \leq 纵隔; 3、摄取 $>$ 纵隔但 \leq 肝脏; 4、适度摄取 $>$ 肝脏; 5、摄取明显高于肝脏和/或新病灶; X、不太可能与淋巴瘤有关的新的摄取区域。

[0323] 在一些方面, 使用Lugano标准描述的完全缓解涉及在各个可测量部位处的完全代谢缓解和完全放射学缓解。在一些方面, 这些部位包括淋巴结和淋巴外部位, 其中当使用PET-CT时, CR被描述为具有评分1、2或3, 在5点量表上具有或不具有残留肿块。在一些方面, 在脾脏或骨髓内具有高生理摄取或激活的Waldeyer环或结外部位 (例如, 用化学疗法或髓细胞样集落刺激因子), 摄取可能大于正常纵隔和/或肝脏。在这种情况下, 如果初始受累部位的摄取不大于周围正常组织, 即使所述组织具有高生理摄取, 也可以推断完全代谢缓解。在一些方面, 使用CT评估淋巴结中的缓解, 其中CR被描述为没有淋巴外疾病部位, 并且靶淋巴结/淋巴结肿块必须复原到病灶最长横径 ≤ 1.5 cm。进一步的评估部位包括骨髓, 其中基于PET-CT的评估应指示缺乏骨髓中FDG浓聚疾病的证据, 并且基于CT的评估应指示正常形态, 如果不确定应该是IHC阴性的话。其他部位可包括应复原到正常的器官肿大的评估。在一些方面, 评估了不可测量的病灶和新病灶, 其在CR的情况下应该不存在 (Cheson等人, (2014) JCO 32 (27) :3059-3067; Johnson等人, (2015) Radiology 2:323-338; Cheson, B.D. (2015) Chin Clin Oncol 4 (1) :5)。

[0324] 在一些方面, 使用Lugano标准描述的部分缓解 (PR) 涉及在各个可测量部位处的部分代谢和/或放射学缓解。在一些方面, 这些部位包括淋巴结和淋巴外部位, 其中当使用PET-CT时, PR被描述为具有评分4或5, 具有与基线相比的降低的摄取和任何大小的残留肿块。这样的发现暂时可以指示疾病有反应。在治疗结束时, 这样的发现可以指示残留疾病。在一些方面, 使用CT评估淋巴结中的缓解, 其中PR被描述为多达6个可测量的靶淋巴结和淋巴结外部位的产物维度总和 (SPD) 减小 $\geq 50\%$ 。如果病灶太小而不能在CT上测量, 则指定5mm x 5mm为默认值; 如果病灶不再可见, 则该值为0mm x 0mm; 对于 >5 mm x 5mm但小于正常的淋巴结, 使用实际测量结果用于计算。进一步的评估部位包括骨髓, 其中基于PET-CT的评估应指示残余摄取, 其高于正常骨髓中的摄取, 但与基线 (与来自允许的化疗疗法的反应性变化相容的弥漫性摄取) 相比降低。在一些方面, 如果在淋巴结缓解的情况下在骨髓中存在持久的局灶性改变, 则应考虑用MRI或活组织检查或间隔扫描进一步评价。在一些方面, 其他部位可以包括器官肿大的评估, 其中, 在超过正常的长度方面, 脾脏必须已经复原 $>50\%$ 。在一些方面, 评估不可测量的病灶和新病灶, 其在PR的情况下应该不存在/正常、复原, 而没有增加。也可以使用基于PET-CT和/或CT的评估来测量无反应/病情稳定 (SD) 或疾病进展 (PD)。(Cheson等人, (2014) JCO 32 (27) :3059-3067; Johnson等人, (2015) Radiology 2:323-338; Cheson, B.D. (2015) Chin Clin Oncol 4 (1) :5)。

[0325] 在一些方面, 无进展生存期 (PFS) 被描述为在疾病 (比如癌症) 的治疗期间和之后受试者带病生活但不会加重的时间长度。在一些方面, 客观缓解 (OR) 被描述为可测量的缓

解。在一些方面,客观缓解率(ORR)被描述为达到CR或PR的患者的比例。在一些方面,总生存期(OS)被描述为从疾病(例如癌症)的诊断或治疗开始日期起始的被诊断患有该疾病的受试者仍然活着的时间长度。在一些方面,无事件生存期(EFS)被描述为在癌症治疗结束之后受试者保持没有某些并发症或事件的预期避免或延迟治疗的时间长度。这些事件可包括癌症的复发或某些症状的发作,例如已扩散到骨骼的癌症所致的骨痛或死亡。

[0326] 在一些实施方案中,缓解持续时间(DOR)的量度包括从记录肿瘤缓解到疾病进展的时间。在一些实施方案中,用于评估缓解的参数可以包括持续缓解,例如,在从疗法开始的一段时间之后持续的缓解和/或对于疗法的持久正向缓解。在一些实施方案中,通过在疗法开始后大约1、2、3、4、5、6、7、8、9、10、11、12、18或24个月的缓解率来指示持续缓解。在一些实施方案中,缓解持续超过3个月或超过6个月。在一些实施方案中,持续缓解是在施用疗法之后第3个月测量的缓解,例如3个月缓解。在一些实施方案中,持续缓解是在施用疗法之后第6个月测量的缓解,例如6个月缓解。

[0327] 在一些方面,使用RECIST标准来确定客观肿瘤缓解;在一些方面,在实体瘤中。(Eisenhauer等人,European Journal of Cancer 45(2009)228-247.) 在一些方面,使用RECIST标准来确定靶病灶的客观肿瘤缓解。在一些方面,使用RECIST标准确定的完全缓解被描述为所有靶病灶的消失,并且任何病理性淋巴结(无论是靶还是非靶的)在短轴上必须减小至 $<10\text{mm}$ 。在其他方面,使用RECIST标准确定的部分缓解被描述为靶病灶的直径总和至少减小30%(以基线总和直径为参考)。在其他方面,以研究中的最小总和为参考(如果基线总和在研究中是最小的话,这包括基线总和),疾病进展(PD)被描述为靶病灶的直径总和增加至少20%。除了相对增加20%之外,所述总和还必须显示至少5mm的绝对增加(在一些方面,一个或多个新病灶的出现也被视为进展)。在其他方面,以研究期间的最小直径总和为参考,病情稳定(SD)被描述为既没有足够的缩小定性为PR,也没有足够的增加定性为PD。

[0328] 在一些实施方案中,疾病或病症是肿瘤,并且疾病负荷的降低是肿瘤尺寸的减小。在一些实施方案中,通过一个或多个因素的降低来指示疾病负荷的降低,所述因素比如受试者或其体液或器官或组织中的疾病细胞的负荷或数目、肿瘤的质量或体积、或转移的程度或范围。在一些实施方案中,可以针对形态疾病和/或最小残留病的程度评估或监测疾病负荷,例如肿瘤负荷。

[0329] 在一些实施方案中,检测、评估或测量受试者中的疾病或病症的负荷。通过检测受试者的器官、组织或体液(比如血液或血清)中的疾病细胞或疾病相关细胞(例如受试者中的肿瘤细胞)的总数,可以在一些方面检测疾病负荷。在一些实施方案中,通过测量实体瘤的质量和/或转移的数量或程度来评估疾病负荷例如肿瘤负荷。在一些方面,评估受试者的生存、某一段时间内的生存率、生存率范围、无事件或无症状生存期的存在或持续时间、或无复发生存率。在一些实施方案中,评估了疾病或病症的任何症状。在一些实施方案中,指定了疾病或病症负荷的量度。

[0330] 在一些实施方案中,疾病负荷可以包括受试者或受试者的器官、组织或体液中(比如肿瘤器官或组织或例如将指示转移的其他位置)的疾病的总细胞数。例如,在某些血液恶性肿瘤的背景下,可以检测和/或量化在血液或骨髓中的肿瘤细胞。

[0331] 在一些实施方案中,疾病负荷可包括肿瘤的质量、转移的数量或程度和/或骨髓中存在的胚细胞的百分比。

[0332] 在一些实施方案中,受试者患有白血病。疾病负荷的程度可以通过评估血液或骨髓中的残留白血病来确定。

[0333] 在一些方面,受试者比如慢性淋巴细胞白血病 (CLL) 受试者的缓解率的根据是国际慢性淋巴细胞白血病工作组 (IWCLL) 的缓解标准 (Hallek等人, Blood 2008, Jun 15; 111 (12): 5446-5456)。在一些方面,这些标准描述如下:完全缓解 (CR), 其在一些方面要求:通过免疫表型分型没有外周血克隆淋巴细胞、没有淋巴结病,没有肝肿大或脾肿大,没有全身症状,以及令人满意的血细胞计数;伴有不完全骨髓恢复的完全缓解 (CRi), 其在一些方面被描述为以上CR,但没有正常的血细胞计数;部分缓解 (PR), 其在一些方面被描述为淋巴细胞计数下降 $\geq 50\%$ 、淋巴结病减少 $\geq 50\%$ 或肝脏或脾脏减小 $\geq 50\%$,并且外周血细胞计数改善;疾病进展 (PD), 其在一些方面被描述为淋巴细胞计数上升 $\geq 50\%$,达到 $>5 \times 10^9/L$,淋巴结病增加 $\geq 50\%$,肝脏或脾脏的大小增加 $\geq 50\%$,Richter转化或由于CLL所致的新的血细胞减少;和病情稳定,其在一些方面被描述为不满足CR、CRi、PR或PD的标准。

[0334] 在一些实施方案中,如果在施用细胞剂量的1个月内,受试者中的淋巴结的尺寸小于或小于约20mm、小于或小于约10mm或小于或等于约10mm,则受试者表现出CR或OR。

[0335] 在一些实施方案中,在受试者的骨髓中(或在根据所述方法治疗的大于50%、60%、70%、80%、90%或更多的受试者的骨髓中)未检测到CLL的指数克隆。在一些实施方案中,通过IgH深度测序评估CLL的指数克隆。在一些实施方案中,在施用细胞之后或之后约或之后至少或之后至少约1、2、3、4、5、6、12、18或24个月时未检出指数克隆。

[0336] 在一些实施方案中,与用治疗剂治疗之前骨髓中的母细胞的百分比相比,如果骨髓中胚细胞的百分比有降低,则存在反应结果。在一些实施方案中,与治疗前骨髓中胚细胞的数目或百分比相比,如果骨髓中胚细胞的数目或百分比有至少或至少约20%、30%、40%、50%、60%、70%、80%、90%、95%或更多的减少或降低,则存在疾病负荷的减轻。

[0337] 在一些实施方案中,如果受试者未表现出形态学疾病(非形态学疾病)或未表现出显著的形态学疾病,则受试者表现出缓解。在一些实施方案中,例如如通过光学显微镜术检测的,如果在骨髓中有大于或等于5%的胚细胞,则受试者表现出形态学疾病。在一些实施方案中,如果骨髓中有少于5%的胚细胞,则受试者表现出完全缓解或临床缓解。

[0338] 在一些实施方案中,如果受试者在治疗前表现出形态学疾病并且在具有或没有分子疾病(例如,最小残留病 (MRD) 的情况下表现出完全缓解(例如,骨髓中少于5%的胚细胞),则受试者表现出疾病负荷的减轻或降低,所述分子疾病是可分子检测的,例如,在治疗后通过流式细胞术或定量PCR检测。在一些实施方案中,如果受试者在治疗前表现出分子疾病并且在治疗后不表现出分子疾病,则受试者表现出疾病负荷的减轻或降低。

[0339] 在一些实施方案中,受试者可表现出完全缓解,但存在小比例的形态学上不可检测的(通过光学显微镜技术)残留白血病细胞。如果受试者在骨髓中表现出小于5%的胚细胞并且表现出可分子检测的癌症,则称受试者表现出最小残留病 (MRD)。在一些实施方案中,可以使用允许少量细胞的灵敏检测的各种分子技术中的任何一种来评估可分子检测的癌症。

[0340] 在一些实施方案中,细胞疗法的反应结果是分子缓解结果。在一些方面,这样的技术包括PCR测定,其可确定由染色体易位产生的独特的Ig/T细胞受体基因重排或融合转录物。在一些实施方案中,可使用流式细胞术来鉴定基于白血病特异性免疫表型的癌细胞。在

一些实施方案中,癌症的分子检测可检出100,000个正常细胞中的少至1个白血病细胞或胚细胞或10,000个正常细胞中的1个白血病细胞或胚细胞。在一些实施方案中,如果例如通过PCR或流式细胞术检出100,000个细胞中的至少或多于1个白血病细胞,则受试者表现出可分子检测的最小残留病(MRD)。

[0341] 在一些实施方案中,受试者的疾病负荷是不可分子检测的或为MRD⁻,使得在一些情况下,使用PCR或流式细胞技术不能在受试者中检出白血病细胞。

[0342] 在一些实施方案中,反应结果是不存在CR或存在完全缓解(CR),其中受试者实现或表现出最小残留病或可分子检测的疾病状态。在一些实施方案中,反应结果是存在具有可分子检测的疾病的CR或存在没有可分子检测的疾病的CR。在一些实施方案中,使用如本文所述的方法评估受试者的疾病负荷,所述方法例如通过流式细胞术或qPCR方法评估骨髓中的胚细胞或分子疾病的方法。

[0343] 在本文中提供的方法的一些实施方案中,缓解被确定为完全缓解(CR)和/或客观缓解(OR);和/或受试者在施用细胞剂量的1个月之内表现出CR、OR、淋巴结的尺寸小于或小于约20mm;和/或任选地在施用细胞剂量之后或之后约或之后至少或之后至少约1、2、3、4、5、6、12、18或24个月时任选地通过IgH深度测序评估在受试者的骨髓中(或在超过50%的根据所述方法治疗的受试者的骨髓中)未检出疾病或病症(比如CLL或NHL)的指数克隆。

[0344] 在一些方面,评估了毒性结果和/或宿主免疫应答的存在或不存在。在一些实施方案中,细胞疗法的反应结果是缺乏免疫应答。在一些实施方案中,关于毒性结果和反应结果的信息可以在受试者中联合地加以评估,比如并行评估或在大约相同的时间或基本上相同的时间评估,并将其用于告知受试者的给药决定或适应性治疗。

[0345] 在一些实施方案中,毒性结果或反应结果是存在的并且/或者可加以评估或监测。在一些实施方案中,在毒性结果和反应结果存在时监测毒性结果和反应结果。在一些实施方案中,评估毒性结果或反应结果的时间为:在(或大约在)其中在受试者中可检出毒性症状或功效的时间段内或在这样的其中在受试者中未检出与无反应或毒性相关的不利结果的时间。在一些实施方案中,所述时间段接近或基本接近于当毒性结果和/或反应结果在受试者中达到峰值时。

[0346] 在一些实施方案中,毒性结果或反应结果是存在的,或者可以在这样的仅施用单剂量治疗剂的时间段评估或监测毒性结果或反应结果。在过继细胞疗法的背景下,施用给定“剂量”包括以单一组合物和/或单次不间断给药的形式例如以单次注射或连续输注的形式施用给定量或数目的细胞,并且还包括在指定的时间段内(不超过3天)以分次剂量的形式施用提供在多个单独的组合物或输液中的给定量或数目的细胞。因此,在一些情况下,第一剂量是在单个时间点给予或开始的指定数目的细胞的单次或连续施用。然而,在一些情况下,在不超过三天的时间内以多次注射或输注的形式施用第一剂量,比如每天一次,持续三天或两天,或在单日时段内多次输注。

[0347] 术语“分次剂量”是指拆分的剂量,使得在超过一天的时间内施用所述剂量。本方法涵盖这种类型的给药并且将其视为单剂量。

[0348] 如本文中使用的,“第一剂量”用于描述给定剂量的时间,在一些情况下,其可以是唯一的剂量,或者其后可以有一个或多个重复或另外的剂量。该术语不一定意味着受试者之前从未接受过一定剂量的治疗剂,甚至受试者之前未接受过一定剂量的相同或基本上相

同的治疗剂。

[0349] 在一些实施方案中,可以通过监测与毒性结果相关的一种或多种症状或事件以及与反应结果相关的一种或多种症状或事件来评估毒性结果和反应结果。

用于过继细胞疗法的工程化细胞

[0350] 在一些实施方案中,所提供的方法可用于评估过继细胞疗法的细胞。在任何提供的方法中,用于细胞疗法的细胞的表观遗传学和/或表观基因组学分析可包括评估和分析基因组基因座或基因组中的变化或修饰的步骤,所述变化或修饰比如染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和/或DNA甲基化。在一些实施方案中,所提供的方法涉及细胞的一个或多个表观遗传学和/或表观基因组学分析步骤。在一些方面,表观遗传学和/或表观基因组学分析在细胞的基因工程之前进行。在一些情况下,在将细胞已经用重组受体基因工程改造之后进行表观遗传学和/或表观基因组学分析。

[0351] 在一些实施方案中,分析包括大规模分析,例如,细胞的多个基因座的分析或全基因组分析。在一些实施方案中,表观遗传学和/或表观基因组学分析包括确定细胞例如用于细胞疗法的工程化细胞的表观遗传学特性。在一些实施方案中,细胞疗法是T细胞疗法,例如肿瘤浸润淋巴细胞(TIL)疗法、转基因TCR疗法或表达嵌合抗原受体(CAR)的T细胞疗法。

细胞

[0352] 细胞通常是真核细胞,比如哺乳动物细胞,并且通常是人细胞,例如源自人受试者并且例如被工程改造以表达重组受体的细胞。在一些实施方案中,细胞来源于血液、骨髓、淋巴或淋巴器官,是免疫系统的细胞,比如先天性或适应性免疫的细胞,例如髓样细胞或淋巴样细胞,包括淋巴细胞,通常是T细胞和/或NK细胞。其他示例性细胞包括干细胞,例如多能和多能性干细胞,包括诱导的多能性干细胞(iPSC)。细胞通常是原代细胞,比如直接从受试者分离和/或从受试者分离并冷冻的细胞。在一些实施方案中,细胞包括T细胞的一个或多个亚群或其他细胞类型,比如整个T细胞群、CD4⁺细胞、CD8⁺细胞及其亚群,比如由功能、活化状态、成熟度、分化潜力、扩增、再循环、定位和/或持久能力、抗原特异性、抗原受体类型、在特定器官或区室中存在、标志物或细胞因子分泌特性、和/或分化程度定义的那些。参考待治疗的受试者,细胞可以是同种异体的和/或自体的。其中这些方法包括现成的方法。在一些方面,比如对于现成技术,细胞是多能性和/或多能的,比如干细胞,比如诱导的多能性干细胞(iPSC)。在一些实施方案中,所述方法包括从受试者分离细胞,制备、处理、培养和/或工程改造它们,并在冷冻保存之前或之后将它们重新引入到同一受试者中。

[0353] 其中T细胞和/或CD4⁺和/或CD8⁺ T细胞的亚型和亚群是幼稚T(T_N)细胞、效应T细胞(T_{EFF})、记忆T细胞及其亚型,比如干细胞记忆T(T_{SCM})、中枢记忆T(T_{CM})、效应记忆T(T_{EM})或终末分化效应记忆T细胞、肿瘤浸润淋巴细胞(TIL)、未成熟T细胞、成熟T细胞、辅助T细胞、细胞毒性T细胞、粘膜相关不变T(MAIT)细胞、天然存在和适应性调节T(Treg)细胞、辅助T细胞,比如TH1细胞、TH2细胞、TH3细胞、TH17细胞、TH9细胞、TH22细胞、滤泡辅助性T细胞、 α/β T细胞和 δ/γ T细胞。

[0354] 在一些实施方案中,细胞是自然杀伤(NK)细胞。在一些实施方案中,细胞是单核细胞或粒细胞,例如髓样细胞、巨噬细胞、嗜中性粒细胞、树突细胞、肥大细胞、嗜酸性粒细胞和/或嗜碱性粒细胞。

在一些实施方案中,细胞包含通过基因工程引入的一种或多种核酸,由此表达这样的

核酸的重组或基因工程产物。在一些实施方案中,核酸是异源的,即通常在细胞或从细胞获得的样品中不存在,比如从另一生物或细胞获得的核酸,例如,通常在被工程化的细胞中和/或从中获得这种细胞的生物中未发现。在一些实施方案中,核酸不是天然存在的,比如在自然界未发现的核酸,包括编码来自多种不同细胞类型的各种结构域的核酸的嵌合组合的核酸。

用于工程化的细胞的制备

[0355] 在一些实施方案中,工程化细胞的制备包括一个或多个培养和/或制备步骤。可以从样品中分离用于引入编码转基因受体(比如CAR)的核酸的细胞,所述样品比如生物样品,例如从受试者获得或衍生的样品。在一些实施方案中,从其分离细胞的受试者是患有疾病或病症或需要细胞疗法或将其施用细胞疗法的受试者。在一些实施方案中,受试者是需要特定治疗干预的人,所述治疗干预比如过继细胞疗法,其中细胞被分离、处理和/或工程化。

[0356] 因此,在一些实施方案中,细胞是原代细胞,例如原代人细胞。样品包括直接取自受试者的组织、液体和其他样品,以及由一个或多个处理步骤产生的样品,所述处理步骤比如分离、离心、基因工程(例如用病毒载体转导)、洗涤和/或温育。生物样品可以是直接从生物来源获得的样品或经过处理的样品。生物样品包括但不限于体液,比如血液、血浆、血清、脑脊液、滑液、尿液和汗液、组织和器官样品,包括由其衍生的处理样品。

[0357] 在一些方面,从其衍生或分离细胞的样品是血液或血液衍生的样品,或者是或衍生自血液成分单采术或白细胞单采术产品。示例性样品包括全血、外周血单核细胞(PBMC)、白细胞、骨髓、胸腺、组织活检、肿瘤、白血病、淋巴瘤、淋巴结、肠相关淋巴组织、粘膜相关淋巴组织、脾、其他淋巴组织、肝脏、肺、胃、肠、结肠、肾、胰腺、乳房、骨、前列腺、子宫颈、睾丸、卵巢、扁桃体或其他器官、和/或由其衍生的细胞。在细胞疗法例如过继细胞疗法的背景下,样品包括来自自体 and 同种异体来源的样品。

[0358] 在一些方面,第二剂量的细胞衍生自与第一剂量的细胞相同的血液成分单采术产品。在一些实施方案中,多个剂量(例如第一、第二、第三等等)的细胞衍生自相同的血液成分单采术产品。

[0359] 在其他实施方案中,第二(或其他后续)剂量的细胞衍生自血液成分单采术产品,其不同于第一(或其他先前)剂量的细胞来源的血液成分单采术产品。

[0360] 在一些实施方案中,细胞衍生自细胞系,例如T细胞系。在一些实施方案中,细胞获自异种来源,例如,来自小鼠、大鼠、非人灵长类动物和猪。

[0361] 在一些实施方案中,细胞的分离包括一个或多个制备步骤和/或基于非亲和力的细胞分离步骤。在一些实例中,在一种或多种试剂的存在下将细胞洗涤、离心和/或温育,例如,以去除不需要的组分,富集期望的组分,裂解或去除对特定试剂敏感的细胞。在一些实例中,基于一种或多种性质分离细胞,所述性质比如密度、粘附性质、大小、对特定组分的敏感性和/或抗性。

[0362] 在一些实例中,例如通过血液成分单采术或白细胞单采术获得来自受试者的循环血液的细胞。在一些方面,样品含有淋巴细胞,包括T细胞、单核细胞、粒细胞、B细胞、其他有核白细胞、红细胞和/或血小板,并且在一些方面含有除红细胞和血小板之外的细胞。

[0363] 在一些实施方案中,洗涤从受试者收集的血细胞,例如以去除血浆部分,并将所述

细胞置于适当的缓冲剂或介质中,以备随后的处理步骤。在一些实施方案中,将细胞用磷酸盐缓冲盐水(PBS)洗涤。在一些实施方案中,洗涤溶液缺乏钙和/或镁和/或许多或所有二价阳离子。在一些方面,使用半自动“直流”离心机(例如,Cobe 2991细胞处理器,Baxter)根据制造商的说明完成洗涤步骤。在一些方面,通过切向流过滤(TFF)根据制造商的说明完成洗涤步骤。在一些实施方案中,洗涤后将细胞重悬于各种生物相容性缓冲剂例如不含Ca⁺⁺/Mg⁺⁺的PBS中。在某些实施方案中,去除血细胞样品的组分并将细胞直接重悬于培养基中。

[0364] 在一些实施方案中,所述方法包括基于密度的细胞分离方法,例如通过裂解红细胞并通过Percoll或Ficoll梯度离心从外周血制备白细胞。

[0365] 在一些实施方案中,分离方法包括基于细胞中一种或多种特异性分子(比如表面标志物,例如表面蛋白、细胞内标志物或核酸)的表达或存在来分离不同的细胞类型。在一些实施方案中,可以使用任何已知的基于这样的标志物的分离方法。在一些实施方案中,分离是基于亲和力或基于免疫亲和力的分离。例如,在一些方面,分离包括基于一种或多种标志物(通常是细胞表面标志物)的细胞表达或表达水平分离细胞和细胞群,例如通过用与这样的标志物特异性地结合的抗体或结合配偶体温育,随后通常为洗涤步骤以及将已经结合抗体或结合配偶体的细胞与未结合抗体或结合配偶体的细胞分离。

[0366] 这样的分离步骤可以基于其中保留已经结合试剂的细胞用于进一步使用的阳性选择、和/或其中保留未与抗体或结合配偶体结合的细胞的阴性选择。在一些实例中,两个部分都被保留以供进一步使用。在一些方面,在没有可用于特异性地鉴定异质群体中的细胞类型的抗体的情况下,阴性选择可能特别有用,使得基于由除期望群体之外的细胞表达的标志物最佳地进行分离。

[0367] 分离不必导致100%富集或去除特定细胞群或表达特定标志物的细胞。例如,特定类型的细胞(比如表达标志物的细胞)的阳性选择或富集是指增加这样的细胞的数量或百分比,但不必导致不表达所述标志物的细胞的完全不存在。同样,特定类型的细胞(比如表达标志物的细胞)的阴性选择、去除或耗竭是指减少这样的细胞的数量或百分比,但不必导致所有这样的细胞的完全去除。

[0368] 在一些实例中,进行了多轮分离步骤,其中使来自一个步骤的阳性或阴性选择的级分经历另一个分离步骤,比如随后的阳性或阴性选择。在一些实例中,比如通过将细胞与多种抗体或结合配偶体一起温育,单个分离步骤可以同时耗竭表达多种标志物的细胞,其中每种抗体或结合配偶体对于靶向阴性选择的标志物是特异的。同样,通过将细胞与各种细胞类型上表达的多种抗体或结合配偶体一起温育,可以同时进行多种细胞类型的阳性选择。

[0369] 例如,在一些方面,通过阳性或阴性选择技术分离特定的T细胞亚群,比如阳性或表达高水平的一种或多种表面标志物的细胞,所述细胞例如CD28⁺、CD62L⁺、CCR7⁺、CD27⁺、CD127⁺、CD4⁺、CD8⁺、CD45RA⁺和/或CD45RO⁺ T细胞。

[0370] 例如,可以使用抗CD3/抗CD28缀合的磁珠(例如,DYNABEADS® M-450CD3/CD28 T Cell Expander)进行CD3⁺、CD28⁺ T细胞的阳性选择。

[0371] 在一些实施方案中,通过阳性选择富集特定的细胞群、或通过阴性选择耗竭特定的细胞群来进行分离。在一些实施方案中,通过将细胞与一种或多种抗体或其他结合剂一起温育来完成阳性或阴性选择,所述抗体或其他结合剂与分别在阳性或阴性选择的细胞上

表达(标志物⁺)或以较高水平表达(标志物^高)的一种或多种表面标志物特异性地结合。

[0372] 在一些实施方案中,通过非T细胞(比如B细胞、单核细胞或其他白细胞)上表达的标志物(比如CD14)的阴性选择,将T细胞从PBMC样品分离。在一些方面,使用CD4⁺或CD8⁺选择步骤来分离CD4⁺辅助细胞和CD8⁺细胞毒性T细胞。通过对一个或多个幼稚、记忆和/或效应T细胞亚群上表达或以相对较高程度表达的标志物的阳性或阴性选择,可以将这样的CD4⁺和CD8⁺群体进一步分类成亚群。

[0373] 在一些实施方案中,例如通过基于与对应的亚群相关的表面抗原的阳性或阴性选择,从CD8⁺细胞进一步富集或耗竭幼稚、中枢记忆、效应记忆和/或中枢记忆干细胞。在一些实施方案中,进行中枢记忆T(T_{CM})细胞的富集以增加效力,比如改善施用后的长期存活、扩增和/或植入,这在一些方面在这样的亚群中特别稳健。参见Terakura等人(2012) Blood. 1: 72-82; Wang等人(2012) J Immunother. 35(9):689-701。在一些实施方案中,联合富含T_{CM}的CD8⁺T细胞和CD4⁺T细胞进一步增强效力。

[0374] 在实施方案中,记忆T细胞存在于CD8⁺外周血淋巴细胞的CD62L⁺和CD62L⁻亚群中。比如使用抗CD8和抗CD62L抗体,可以从PBMC富集或耗竭CD62L⁻CD8⁺和/或CD62L⁺CD8⁺级分。

[0375] 在一些实施方案中,中枢记忆T(T_{CM})细胞的富集基于CD45R0、CD62L、CCR7、CD28、CD3和/或CD127的阳性表达或高表面表达;在一些方面,它基于对表达或高度表达CD45RA和/或颗粒酶B的细胞的阴性选择。在一些方面,通过表达CD4、CD14、CD45RA的细胞的耗竭以及表达CD62L的细胞的阳性选择或富集,进行富含T_{CM}细胞的CD8⁺群体的分离。在一个方面,以基于CD4表达选择的阴性细胞级分开始进行中枢记忆T(T_{CM})细胞的富集,将所述阴性细胞级分进行基于CD14和CD45RA表达的阴性选择和基于CD62L的阳性选择。在一些方面,同时进行这样的选择,并且在其他方面,以任一顺序依次进行。在一些方面,用于制备CD8⁺细胞群或亚群的相同的基于CD4表达的选择步骤也用于生成CD4⁺细胞群或亚群,使得来自基于CD4的分离的阳性级分和阴性级分都被保留并用于这些方法的后续步骤中,任选地随后为一个或多个另外的阳性或阴性选择步骤。

[0376] 在特定的实例中,对PBMC样品或其他白细胞样品进行CD4⁺细胞选择,其中阴性级分和阳性级分两者都被保留。然后基于CD14和CD45RA或CD19的表达对阴性级分进行阴性选择,并基于中枢记忆T细胞的标志物特征(如CD62L或CCR7)进行阳性选择,其中以任一顺序进行阳性选择和阴性选择。

[0377] 通过鉴定具有细胞表面抗原的细胞群,将CD4⁺T辅助细胞分类为幼稚、中枢记忆和效应细胞。可通过标准方法获得CD4⁺淋巴细胞。在一些实施方案中,幼稚CD4⁺T淋巴细胞是CD45R0⁻、CD45RA⁺、CD62L⁺、CD4⁺T细胞。在一些实施方案中,中枢记忆CD4⁺细胞是CD62L⁺和CD45R0⁺的。在一些实施方案中,效应CD4⁺细胞是CD62L⁻和CD45R0⁻的。

[0378] 在一个实例中,为了通过阴性选择富集CD4⁺细胞,单克隆抗体混合物通常包括针对CD14、CD20、CD11b、CD16、HLA-DR和CD8的抗体。在一些实施方案中,抗体或结合配偶体与固体支持物或基质(比如磁珠或顺磁珠)结合,以允许分离用于阳性和/或阴性选择的细胞。例如,在一些实施方案中,使用免疫磁性(或亲和磁性)分离技术(在Methods in Molecular Medicine, vol.58:Metastasis Research Protocols, Vol.2:Cell Behavior In Vitro and In Vivo,第17-25页,编辑:S.A.Brooks和U.Schumacher©Humana Press Inc., Totowa,NJ中综述)来分开或分离细胞和细胞群。

[0379] 在一些方面,将待分离的细胞的样品或组合物与小的可磁化或磁响应材料一起温育,所述材料比如磁响应颗粒或微粒,比如顺磁珠(例如DynaBeads或MACS珠)。磁响应材料(例如颗粒)通常直接或间接地附着于结合配偶体(例如抗体),所述结合配偶体与存在于期望分离(例如,期望被阴性或阳性选择)的一个细胞、多个细胞或细胞群上的分子(例如表面标志物)特异性地结合。

[0380] 在一些实施方案中,磁性颗粒或珠子包含与特异性结合成员(比如抗体或其他结合配偶体)结合的磁响应材料。有许多熟知的在磁分离方法中使用的磁响应材料。适合的磁性颗粒包括在Molday的美国专利No. 4,452,773和欧洲专利说明书EP 452342B中描述的那些磁性颗粒,将其通过提述并入本文。胶体大小颗粒(如在Owen的美国专利No. 4,795,698和Liberti等人的美国专利No. 5,200,084中所述的那些)是其他的实例。

[0381] 温育通常在这样的条件下进行,抗体或结合配偶体或分子(比如二级抗体或其他试剂)借此条件与细胞表面分子(如果存在于样品中的细胞上)特异性地结合,所述二级抗体或其他试剂与附着于磁性颗粒或珠子的这样的抗体或结合配偶体特异性地结合。

[0382] 在一些方面,将样品置于磁场中,具有附着于其上的磁响应或可磁化颗粒的那些细胞将被吸引到磁体上并与未标记的细胞分离。对于阳性选择,保留被吸引到磁体上的细胞;对于阴性选择,保留未被吸引的细胞(未标记的细胞)。在一些方面,在相同的选择步骤期间进行阳性和阴性选择的组合,其中阳性和阴性级分被保留并进一步处理或经受进一步的分离步骤。

[0383] 在某些实施方案中,磁响应颗粒被包被在一级抗体或其他结合配偶体、二级抗体、凝集素、酶或链霉亲和素中。在某些实施方案中,磁性颗粒通过对一种或多种标志物特异的一级抗体的包被而附着于细胞。在某些实施方案中,用一级抗体或结合配偶体标记细胞而不是珠子,然后加入细胞类型特异性二级抗体包被的或其他结合配偶体(例如链霉亲和素)包被的磁性颗粒。在某些实施方案中,将链霉亲和素包被的磁性颗粒与生物素化的一级抗体或二级抗体结合使用。

[0384] 在一些实施方案中,磁响应颗粒保持附着于细胞,所述细胞随后将被温育、培养和/或工程化;在一些方面,将颗粒附着于细胞,用于施用于患者。在一些实施方案中,从细胞去除可磁化或磁响应颗粒。从细胞去除可磁化颗粒的方法是已知的,包括例如使用竞争性非标记抗体、和可磁化颗粒或与可切割接头缀合的抗体。在一些实施方案中,可磁化颗粒是可生物降解的。

[0385] 在一些实施方案中,基于亲和力的选择通过磁激活细胞分选(MACS)(Miltenyi Biotec, Auburn, CA)来进行。磁激活细胞分选(MACS)系统能够以高纯度选择其上具有附着的磁化颗粒的细胞。在某些实施方案中,MACS以这样的模式操作,其中在施加外部磁场之后依次洗脱非目标物质和目标物质。也就是说,附着在磁化颗粒上的细胞保持在适当的位置,而未附着的物质被洗脱。然后,在完成该第一个洗脱步骤之后,以某种方式释放被捕获在磁场中并被免受洗脱的物质,使得它们可以被洗脱和回收。在某些实施方案中,将非靶细胞标记并从异质细胞群中去除。

[0386] 在某些实施方案中,使用系统、装置或设备进行分离或分开,所述系统、装置或设备进行分离、细胞制备、分开、处理、温育、培养和/或制剂方法步骤中的一个或多个。在一些方面,所述系统用于在封闭或无菌环境中进行这些步骤中的每一个,例如,以便将错误、用

户操作和/或污染降到最低。在一个实例中,所述系统是如国际专利申请公开号W02009/072003或US20110003380A1中描述的系统。

[0387] 在一些实施方案中,所述系统或设备在一体化或自含式系统中和/或以自动或可编程方式进行分离、处理、工程化和制剂步骤中的一个或多个(例如全部)。在一些方面,所述系统或设备包括与所述系统或设备通信的计算机和/或计算机程序,其允许用户对处理、分离、工程化和制剂步骤的各个方面进行编程、控制、评估结果和/或调整。

[0388] 在一些方面,使用CliniMACS系统(Miltenyi Biotec)进行分离和/或其他步骤,例如,用于在封闭和无菌系统中在临床规模水平上自动分离细胞。组件可包括集成微型计算机、磁分离单元、蠕动泵和各种夹管阀。在一些方面,集成计算机控制该仪器的所有组件并指导该系统以标准化顺序执行重复程序。在一些方面,磁分离单元包括可移动永磁体和用于选择柱的支架。蠕动泵控制管组各处的流速,并与夹管阀一起确保缓冲剂通过该系统的流动受到控制和细胞的连续悬浮。

[0389] 在一些方面,CliniMACS系统使用提供在无菌、无热原溶液中的抗体偶联的可磁化颗粒。在一些实施方案中,在用磁性颗粒标记细胞之后,洗涤细胞以除去过量的颗粒。然后将细胞制备袋连接到管组,管组进而连接到含有缓冲剂的袋和细胞收集袋。管组包括预组装的无菌管,包括预柱和分离柱,仅供一次性使用。在启动分离程序之后,系统自动将细胞样品施加到分离柱上。标记的细胞保留在柱内,而未标记的细胞通过一系列洗涤步骤去除。在一些实施方案中,与本文所述的方法一起使用的细胞群是未标记的并且未保留在柱中。在一些实施方案中,与本文所述的方法一起使用的细胞群是标记的并且保留在柱中。在一些实施方案中,在去除磁场后,将与本文所述方法一起使用的细胞群从柱中洗脱,并收集在细胞收集袋内。

[0390] 在某些实施方案中,使用CliniMACS Prodigy系统(Miltenyi Biotec)进行分离和/或其他步骤。在一些方面,CliniMACS Prodigy系统配备有细胞处理单元,其允许通过离心将细胞自动洗涤和分级分离。CliniMACS Prodigy系统还可以包括机载相机和图像识别软件,其通过辨别源细胞产品的宏观层来确定最佳细胞分级分离终点。例如,外周血自动分离成红细胞、白细胞和血浆层。CliniMACS Prodigy系统还可以包括集成细胞培养室,其完成细胞培养方案,例如细胞分化和扩增、抗原加载和长期细胞培养。输入端口可以允许无菌移除和补充培养基,并且可以使用集成显微镜监测细胞。参见,例如Klebanoff等人(2012) *J Immunother.* 35 (9) :651-660, Terakura等人(2012) *Blood.* 1:72-82和Wang等人(2012) *J Immunother.* 35 (9) :689-701。

[0391] 在一些实施方案中,通过流式细胞术收集和富集(或耗竭)本文所述的细胞群,其中在流体流中携带针对多个细胞表面标志物染色的细胞。在一些实施方案中,通过制备规模的(FACS)-分选收集和富集(或耗竭)本文所述的细胞群。在某些实施方案中,通过使用机电系统(MEMS)芯片结合基于FACS的检测系统来收集和富集(或耗竭)本文所述的细胞群(参见例如W02010/033140, Cho等人(2010) *Lab Chip* 10, 1567-1573; 和Godin等人(2008) *J Biophoton.* 1 (5) :355-376。在两种情况下,可以用多种标志物标记细胞,从而允许以高纯度分离定义明确的T细胞亚群。

[0392] 在一些实施方案中,用一种或多种可检测标志物标记抗体或结合配偶体,以促进针对阳性和/或阴性选择的分离。例如,分离可以基于与荧光标记抗体的结合。在一些实例

中,比如通过荧光激活细胞分选(FACS),包括例如与流式细胞检测系统联合的制备规模(FACS)和/或微机电系统(MEMS)芯片,基于对一种或多种细胞表面标志物特异的抗体或其他结合配偶体的结合来分离细胞,所述细胞表面标志物被携带在流体流中。这样的方法允许同时基于多种标志物进行阳性和阴性选择。

[0393] 在一些实施方案中,制备方法包括在分离、温育和/或工程化之前或之后冷冻例如冷冻保存细胞的步骤。在一些实施方案中,冷冻和随后的解冻步骤去除细胞群中的粒细胞,并在一定程度上去除单核细胞。在一些实施方案中,将细胞悬浮在冷冻溶液中,例如随后为洗涤步骤以去除血浆和血小板。在一些方面,可以使用各种已知的冷冻溶液和参数中的任何一种。一个实例涉及使用含有20% DMSO和8%人血清白蛋白(HSA)的PBS或其他适合的细胞冷冻培养基。然后用培养基1:1稀释,使得DMSO和HSA的终浓度分别为10%和4%。然后将细胞通常以1°/分钟的速率冷冻至-80°C并储存在液氮储罐的气相中。

[0394] 在一些实施方案中,在基因工程之前(或与基因工程相结合)温育和/或培养细胞。温育步骤可包括培养、培育、刺激、活化和/或增殖。在一些实施方案中,在刺激条件或刺激剂的存在下温育组合物或细胞。这样的条件包括设计为诱导群体中细胞的增殖、扩增、活化和/或存活,模拟抗原暴露和/或引发细胞用于基因工程(比如用于引入重组抗原受体)的那些条件。

[0395] 所述条件可包括下列一种或多种:特定的介质、温度、氧含量、二氧化碳含量、时间、作用剂,例如营养素、氨基酸、抗生素、离子和/或刺激因子,如细胞因子、趋化因子、抗原、结合配偶体、融合蛋白、重组可溶性受体和任何其他旨在激活细胞的作用剂。

[0396] 在一些实施方案中,刺激条件或作用剂包括一种或多种作用剂,例如配体,其能够激活TCR复合物的细胞内信号传导结构域。在一些方面,所述作用剂在T细胞中开启或启动TCR/CD3细胞内信号级联。这样的作用剂可以包括抗体,比如对TCR特异的抗体,例如抗CD3。在一些实施方案中,刺激条件包括一种或多种能够刺激共刺激受体的作用剂,例如配体,例如抗CD28。在一些实施方案中,这样的作用剂和/或配体可以与固体支持物如珠子和/或一种或多种细胞因子结合。任选地,扩增方法可以进一步包括向培养基中加入抗CD3和/或抗CD28抗体的步骤(例如,浓度至少约0.5ng/ml)。在一些实施方案中,刺激剂包括IL-2、IL-15和/或IL-7。在一些方面,IL-2浓度为至少约10单位/mL。

[0397] 在一些方面,根据例如在Riddell等人的美国专利No.6,040,177,Klebanoff等人(2012) J Immunother. 35 (9):651-660, Terakura等人(2012) Blood. 1:72-82和/或Wang等人(2012) J Immunother. 35 (9):689-701中描述的技术进行温育。

[0398] 在一些实施方案中,通过向培养起始组合物添加饲养细胞(比如非分裂外周血单核细胞(PBMC))来扩增T细胞(例如,使得所得细胞群针对待扩增的初始群体中的每个T淋巴细胞含有至少约5个、10个、20个或40个或更多个PBMC饲养细胞);并温育培养物(例如,持续足以扩增T细胞数量的时间)。在一些方面,非分裂饲养细胞可包括 γ 照射的PBMC饲养细胞。在一些实施方案中,用约3000至3600拉德范围的 γ 射线照射PBMC以阻止细胞分裂。在一些方面,在添加T细胞群之前将饲养细胞添加到培养基中。

[0399] 在一些实施方案中,刺激条件包括适合人T淋巴细胞生长的温度,例如,至少约25摄氏度,通常至少约30度,并且通常为或约为37摄氏度。任选地,温育可以进一步包括添加非分裂的EBV转化的类淋巴母细胞(LCL)作为饲养细胞。可以用约6000至10,000拉德的

γ 射线照射LCL。在一些方面,以任何适合的量提供LCL饲养细胞,比如LCL饲养细胞与初始T淋巴细胞的比例为至少约10:1。

[0400] 在实施方案中,通过用抗原刺激幼稚或抗原特异性T淋巴细胞而获得抗原特异性T细胞,比如抗原特异性CD4⁺和/或CD8⁺ T细胞。例如,针对巨细胞病毒抗原的抗原特异性T细胞系或克隆可以通过从感染的受试者中分离T细胞并用相同的抗原体外刺激所述细胞而生成。

[0401] 在一些实施方案中,可以改变或测试与培养或处理细胞有关的任何一种或多种条件或作用剂,并根据所提供的方法评估它们的如通过表观遗传学分析(例如染色质可及性)确定的对细胞的表型或功能或特点的影响。在一些实施方案中,可以将作用剂或条件附加至细胞培养物,并且可以评估基因组区域或指示细胞的表型或功能的区域的表观遗传学特性。在一些实施方案中,评估了指示鉴定幼稚细胞或长寿记忆细胞的基因或基因的集合。在一些实施方案中,评估了指示细胞(比如效应细胞或效应记忆细胞)的效应子样功能的基因或基因的集合。提供了这样的基因的示例。

通过细胞表达的重组受体

[0402] 细胞通常表达重组受体。所述受体可包括抗原受体,比如功能性非TCR抗原受体,包括嵌合抗原受体(CAR),和其他抗原结合受体,比如转基因T细胞受体(TCR)。受体还可以包括其他嵌合受体,比如与特定配体结合并具有与CAR中存在的那些相似的跨膜和/或细胞内信号传导结构域的受体。其中受体是抗原受体和含有它的一种或多种组分的受体。重组受体可包括嵌合受体,比如含有配体结合结构域或其结合片段和细胞内信号传导结构域或区域的嵌合受体、功能性非TCR抗原受体、嵌合抗原受体(CAR)和T细胞受体(TCR),比如重组或转基因TCR、嵌合自身抗体受体(CAAR)和任何前述受体的组分。重组受体(比如CAR)通常包括与一种或多种细胞内信号传导组分连接(在一些方面,经由接头和/或跨膜结构域)的细胞外抗原(或配体)结合结构域。

嵌合抗原受体(CAR)

[0403] 在一些实施方案中,工程化细胞(比如T细胞)表达对特定抗原(或标志物或配体)具有特异性的CAR,所述抗原比如在特定细胞类型表面上表达的抗原。在一些实施方案中,抗原是多肽。在一些实施方案中,它是碳水化合物或其他分子。在一些实施方案中,与正常或非靶向细胞或组织相比,抗原在疾病或病症的细胞(例如肿瘤或病原细胞)上选择性地表达或过表达。在其他实施方案中,抗原在正常细胞上表达和/或在工程化细胞上表达。

[0404] 在特定的实施方案中,重组受体(比如嵌合受体)含有细胞内信号传导区,所述细胞内信号传导区包括胞质信号传导结构域或区域(也可互换地称为细胞内信号传导结构域或区域),比如能够在T细胞中诱导初级活化信号的胞质(细胞内)区,例如T细胞受体(TCR)组分的胞质信号传导结构域或区域(例如CD3-zeta(CD3 ζ)链的 ζ 链的胞质信号传导结构域或区域或其功能变体或信号传导部分),并且/或者其包含免疫受体酪氨酸活化基序(ITAM)。

[0405] 在一些实施方案中,嵌合受体进一步含有与配体(例如抗原)抗原特异性地结合的细胞外配体结合结构域。在一些实施方案中,嵌合受体是CAR,其含有与抗原特异性地结合的细胞外抗原识别结构域。在一些实施方案中,配体(比如抗原)是在细胞表面上表达的蛋白质。在一些实施方案中,CAR是TCR样CAR,并且抗原是加工的肽抗原,比如细胞内蛋白的肽

抗原,其与TCR一样,在主要组织相容性复合体(MHC)分子的背景下在细胞表面上被识别。

[0406] 示例性抗原受体(包括CAR)、以及用于将这样的受体工程化和引入细胞中的方法包括例如在以下文献中描述的那些:国际专利申请公开号W0200014257、W02013126726、W02012/129514、W02014031687、W02013/166321、W02013/071154、W02013/123061,美国专利申请公开号US2002131960、US2013287748、US20130149337,美国专利No.6,451,995、7,446,190、8,252,592、8,339,645、8,398,282、7,446,179、6,410,319、7,070,995、7,265,209、7,354,762、7,446,191、8,324,353和8,479,118,和欧洲专利申请号EP2537416,和/或Sadelain等人,Cancer Discov.2013April;3(4):388-398;Davila等人(2013)PLoS ONE 8(4):e61338;Turtle等人,Curr.Opin.Immunol.,2012October;24(5):633-39;Wu等人,Cancer,2012March 18(2):160-75。在一些方面,抗原受体包括如美国专利No.7,446,190中描述的CAR、和国际专利申请公开号W0/2014055668A1中描述的那些。CAR的实例包括任何上述出版物中公开的CAR,比如W02014031687、US 8,339,645、US 7,446,179、US 2013/0149337、美国专利No.7,446,190、美国专利No.8,389,282,Kochenderfer等人,2013,Nature Reviews Clinical Oncology,10,267-276(2013);Wang等人(2012)J.Immunother.35(9):689-701;和Brentjens等人,Sci Transl Med.2013 5(177)。还参见国际专利公开号W02014031687,美国专利No.8,339,645、7,446,179、7,446,190和8,389,282,以及美国专利申请公开号US2013/0149337。其中嵌合受体是嵌合抗原受体(CAR)。嵌合受体(比如CAR)通常包括细胞外抗原结合结构域,比如抗体分子的一部分,通常是抗体的可变重(V_H)链区和/或可变轻(V_L)链区,例如scFv抗体片段。

[0407] 在一些实施方案中,CAR构建为对特定抗原(或标志物或配体)具有特异性,比如在过继治疗所靶向的特定细胞类型中表达的抗原,例如,旨在诱导抑制(dampening)反应的癌症标志物和/或抗原,比如在正常或非病变细胞类型上表达的抗原。因此,CAR通常在其细胞外部分包括一个或多个抗原结合分子,比如一个或多个抗原结合片段、结构域或部分,或一个或多个抗体可变结构域和/或抗体分子。在一些实施方案中,CAR包括抗体分子的一个或多个抗原结合部分,比如衍生自单克隆抗体(mAb)的可变重链(VH)和可变轻链(VL)的单链抗体片段(scFv)。

[0408] 在一些实施方案中,抗体或其抗原结合部分在细胞上表达为重组受体(比如抗原受体)的一部分。其中抗原受体为功能性非TCR抗原受体,比如嵌合抗原受体(CAR)。通常,含有表现出针对肽-MHC复合物的TCR样特异性的抗体或抗原结合片段的CAR也可称为TCR样CAR。在一些实施方案中,对TCR样CAR的MHC-肽复合物特异的细胞外抗原结合结构域在一些方面经由接头和/或跨膜结构域与一种或多种细胞内信号传导组分连接。在一些实施方案中,这样的分子通常可通过天然抗原受体(比如TCR)模拟或接近信号,并且,任选地通过这种受体与共刺激受体联合模拟或接近信号。

[0409] 在一些实施方案中,重组受体,比如嵌合受体(例如CAR),包括与抗原(或配体)结合(例如特异性地结合)的配体结合结构域。其中嵌合受体所靶向的抗原是在经由过继细胞疗法靶向的疾病、病症或细胞类型的背景下表达的抗原。其中疾病和病症是增殖性、肿瘤性以及恶性疾病和疾患,包括癌症和肿瘤,包括血液癌症、免疫系统的癌症,比如淋巴瘤、白血病和/或骨髓瘤,比如B、T和髓性白血病、淋巴瘤和多发性骨髓瘤。

[0410] 在一些实施方案中,抗原(或配体)是多肽。在一些实施方案中,它是碳水化合物或

其他分子。在一些实施方案中,与正常或非靶向细胞或组织相比,抗原(或配体)在疾病或病症的细胞(例如肿瘤或病原细胞)上选择性地表达或过表达。在其他实施方案中,抗原在正常细胞上表达和/或在工程化细胞上表达。

[0411] 在一些实施方案中,CAR含有特异性识别抗原的抗体或抗原结合片段(例如scFv),所述抗原比如在细胞表面上表达的完整抗原。

[0412] 在一些实施方案中,抗原(或配体)是肿瘤抗原或癌症标志物。在一些实施方案中,抗原(或配体)是或包括 α v β 6整合素(avb6整合素)、B细胞成熟抗原(BCMA)、B7-H3、B7-H6、碳酸酐酶9(CA9,也称为CAIX或G250)、癌-睾丸抗原、癌/睾丸抗原_{1B}(CTAG,也称为NY-ES0-1和LAGE-2)、癌胚抗原(CEA)、细胞周期蛋白、细胞周期蛋白A2、C-C基序趋化因子配体1(CCL-1)、CD19、CD20、CD22、CD23、CD24、CD30、CD33、CD38、CD44、CD44v6、CD44v7/8、CD123、CD138、CD171、表皮生长因子蛋白(EGFR)、截短的表皮生长因子蛋白(tEGFR)、III型表皮生长因子受体突变(EGFR vIII)、上皮糖蛋白2(EPG-2)、上皮糖蛋白40(EPG-40)、肝配蛋白B2、肝配蛋白受体A2(EPha2)、雌激素受体、Fc受体样5(FCRL5;也称为Fc受体同源物5或FCRH5)、胎儿乙酰胆碱受体(胎儿AChR)、叶酸结合蛋白(FBP)、叶酸受体 α 、神经节苷脂GD2、O-乙酰化GD2(OGD2)、神经节苷脂GD3、糖蛋白100(gp100)、G蛋白偶联受体5D(GPCR5D)、Her2/neu(受体酪氨酸激酶erb-B2)、Her3(erb-B3)、Her4(erb-B4)、erbB二聚体、人类高分子量黑色素瘤相关抗原(HMW-MAA)、乙型肝炎表面抗原、人白细胞抗原A1(HLA-A1)、人白细胞抗原A2(HLA-A2)、IL-22受体 α (IL-22Ra)、IL-13受体 α 2(IL-13Ra2)、激酶插入结构域受体(kdr)、 κ 轻链、L1细胞粘附分子(L1-CAM)、L1-CAM的CE7表位、富含亮氨酸重复序列蛋白家族成员8A(LRRC8A)、Lewis Y、黑色素瘤相关抗原(MAGE)-A1、MAGE-A3、MAGE-A6、间皮素、c-Met、鼠巨细胞病毒(CMV)、粘蛋白1(MUC1)、MUC16、自然杀伤组2成员D(NKG2D)配体、黑色素A(MART-1)、神经细胞粘附分子(NCAM)、癌胚胎抗原、黑色素瘤优先表达抗原(PRAME)、孕酮受体、前列腺特异性抗原、前列腺干细胞抗原(PSCA)、前列腺特异性膜抗原(PSMA)、受体酪氨酸激酶样孤儿受体1(ROR1)、生存素、滋养层细胞糖蛋白(TPBG也称为5T4)、肿瘤相关糖蛋白72(TAG72)、血管内皮生长因子受体(VEGFR)、血管内皮生长因子受体2(VEGFR2)、维尔姆斯瘤1(WT-1)、病原体特异性抗原、或通用标签相关抗原、和/或生物素化分子,和/或由HIV、HCV、HBV或其他病原体表达的分子。在一些实施方案中,受体所靶向的抗原包括与B细胞恶性肿瘤相关的抗原,比如许多已知的B细胞标志物的任何一种。在一些实施方案中,所述抗原是或包括CD20、CD19、CD22、ROR1、CD45、CD21、CD5、CD33、Ig κ 、Ig λ 、CD79a、CD79b或CD30。

[0413] 在一些实施方案中,所述抗原是或包括病原体特异性或病原体表达的抗原。在一些实施方案中,抗原是病毒抗原(例如来自HIV、HCV、HBV等的病毒抗原)、细菌抗原和/或寄生虫抗原。

[0414] 在一些实施方案中,抗体或抗原结合片段(例如scFv)特异性地识别抗原,比如CD19。

[0415] 在一些实施方案中,scFv和/或V_H结构域衍生自FMC63。FMC63通常是指针对表达人源CD19的Nalm-1和-16细胞产生的小鼠单克隆IgG1抗体(Ling,N.R.等人(1987).Leucocyte typing III.302)。FMC63抗体包含分别在SEQ ID NO:38、39中列出的CDRH1和H2;在SEQ ID NO:40或54中列出的CDRH3;以及在SEQ ID NO:35中列出的CDRL1,以及CDR L2 36或55和CDR-L3序列37或56。FMC63抗体包含含有氨基酸序列SEQ ID NO:41的重链可变区(V_H)和含

有氨基酸序列SEQ ID NO:42的轻链可变区(V_L)。在一些实施方案中,svFv包含含有CDRL1序列SEQ ID NO:35、CDRL2序列SEQ ID NO:36和CDRL3序列SEQ ID NO:37的可变轻链和/或含有CDRH1序列SEQ ID NO:38、CDRH2序列SEQ ID NO:39和CDRH3序列SEQ ID NO:40的可变重链。在一些实施方案中,svFv包含在SEQ ID NO:41中列出的FMC63的重链可变区和在SEQ ID NO:42中列出的FMC63的轻链可变区。在一些实施方案中,可变重链和可变轻链通过接头连接。在一些实施方案中,接头在SEQ ID NO:59中列出。在一些实施方案中,scFv依次包含V_H、接头和V_L。在一些实施方案中,scFv依次包含V_L、接头和V_H。在一些实施方案中,svFc由SEQ ID NO:57中列出的核苷酸序列或表现出与SEQ ID NO:57具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列编码。在一些实施方案中,svFc包含在SEQ ID NO:43中列出的氨基酸序列或表现出与SEQ ID NO:43具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0416] 在一些实施方案中,svFc衍生自SJ25C1。SJ25C1是针对表达人源CD19的Na1m-1和-16细胞产生的小鼠单克隆IgG1抗体(Ling,N.R.等人(1987).Leucocyte typing III.302)。SJ25C1抗体包含分别在SEQ ID NO:47-49中列出的CDRH1、H2和H3,以及分别在SEQ ID NO:44-46中列出的CDRL1、L2和L3序列。SJ25C1抗体包含含有氨基酸序列SEQ ID NO:50的重链可变区(V_H)和含有氨基酸序列SEQ ID NO:51的轻链可变区(V_L)。在一些实施方案中,svFv包含含有CDRL1序列SEQ ID NO:44、CDRL2序列SEQ ID NO:45和CDRL3序列SEQ ID NO:46的可变轻链和/或含有CDRH1序列SEQ ID NO:47、CDRH2序列SEQ ID NO:48和CDRH3序列SEQ ID NO:49的可变重链。在一些实施方案中,svFv包含在SEQ ID NO:50中列出的SJ25C1的重链可变区和在SEQ ID NO:51中列出的SJ25C1的轻链可变区。在一些实施方案中,可变重链和可变轻链通过接头连接。在一些实施方案中,接头在SEQ ID NO:52中列出。在一些实施方案中,scFv依次包含V_H、接头和V_L。在一些实施方案中,scFv依次包含V_L、接头和V_H。在一些实施方案中,svFc包含在SEQ ID NO:53中列出的氨基酸序列或表现出与SEQ ID NO:53具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0417] 在一些实施方案中,CAR含有特异性识别细胞内抗原的TCR样抗体,比如抗体或抗原结合片段(例如scFv),所述细胞内抗原比如肿瘤相关抗原,在细胞表面上呈现为MHC-肽复合物。在一些实施方案中,识别MHC-肽复合物的抗体或其抗原结合部分可以作为重组受体(比如抗原受体)的一部分在细胞上表达。其中抗原受体为功能性非TCR抗原受体,比如嵌合抗原受体(CAR)。通常,含有表现出针对肽-MHC复合物的TCR样特异性的抗体或抗原结合片段的CAR也可称为TCR样CAR。

[0418] 关于“主要组织相容性复合体”(MHC)是指一种蛋白质,通常为糖蛋白,其含有多态性肽结合位点或结合沟,在一些情况下,所述结合位点或结合沟可与多肽的肽抗原(包括由细胞机器加工的肽抗原)复合。在一些情况下,MHC分子可以在细胞表面上展示或表达,包括具有肽的复合物,即MHC-肽复合物,其用于以可被T细胞上的抗原受体识别的构象呈递抗原,所述抗原受体比如TCR或TCR样抗体。通常,MHC I类分子是异二聚体,其具有跨膜 α 链,在一些情况下具有三个 α 结构域和非共价结合的 β 2微球蛋白。通常,MHC II类分子由两种跨膜糖蛋白 α 和 β 组成,这两种糖蛋白通常都跨膜。MHC分子可包括MHC的有效部分,所述有效部分

含有抗原结合位点或结合肽的位点以及被适当抗原受体识别所必需的序列。在一些实施方案中，MHC I类分子将起源于胞质溶胶的肽递送至细胞表面，MHC-肽复合物在这里被T细胞（比如通常为CD8⁺ T细胞，但在一些情况下为CD4⁺ T细胞）识别。在一些实施方案中，MHC II类分子将起源于囊泡系统的肽递送至细胞表面，它们在这里通常被CD4⁺ T细胞识别。通常，MHC分子由一组连锁基因座编码，其在小鼠中统称为H-2，并且在人类中统称为人白细胞抗原（HLA）。因此，通常人MHC也可称为人白细胞抗原（HLA）。

[0419] 术语“MHC-肽复合物”或“肽-MHC复合物”或其变体是指肽抗原和MHC分子的复合物或结合，例如其通常借助于MHC分子的结合沟或缝中的肽的非共价相互作用。在一些实施方案中，MHC-肽复合物呈现或展示在细胞表面上。在一些实施方案中，MHC-肽复合物可以被抗原受体特异性地识别，所述抗原受体比如TCR、TCR样CAR或其抗原结合部分。

[0420] 在一些实施方案中，多肽的肽（比如肽抗原或表位）可与MHC分子结合，例如以便被抗原受体识别。通常，肽衍生自或基于诸如多肽或蛋白质的较长生物分子的片段。在一些实施方案中，肽的长度通常为约8个至约24个氨基酸。在一些实施方案中，为了识别MHC II类复合物，肽具有（或具有大约）9至22个氨基酸的长度。在一些实施方案中，为了识别MHC I类复合物，肽具有（或具有大约）8至13个氨基酸的长度。在一些实施方案中，在MHC分子（比如MHC-肽复合物）的背景下识别肽时，抗原受体（比如TCR或TCR样CAR）产生或触发诱导T细胞活化信号，所述活化信号诱导T细胞应答，比如T细胞增殖、细胞因子产生、细胞毒性T细胞应答或其他应答。

[0421] 在一些实施方案中，TCR样抗体或抗原结合部分是已知的或可通过已知的方法产生（参见例如美国公开申请号US 2002/0150914；US2003/0223994；US 2004/0191260；US 2006/0034850；US 2007/00992530；US20090226474；US20090304679；和国际PCT公开号WO 03/068201）。

[0422] 在一些实施方案中，可以通过用有效量的含有特异性MHC-肽复合物的免疫原免疫宿主来产生与MHC-肽复合物特异性地结合的抗体或其抗原结合部分。在一些情况下，MHC-肽复合物的肽是能够与MHC结合的抗原的表位，所述抗原比如肿瘤抗原，例如通用肿瘤抗原、骨髓瘤抗原或如下所述的其他抗原。在一些实施方案中，然后将有效量的免疫原施用于宿主以引出免疫应答，其中免疫原保持其三维形式持续一段时间，所述时间足以引出针对MHC分子结合沟中的肽的三维呈现的免疫应答。然后测定从宿主收集的血清，以确定是否产生了识别MHC分子结合沟中的肽的三维呈现的期望抗体。在一些实施方案中，可以评估产生的抗体，以证实抗体能够区分MHC-肽复合物与单独的MHC分子、单独的感兴趣的肽、以及MHC和无关肽的复合物。然后可以分离期望抗体。

[0423] 在一些实施方案中，可以通过采用抗体文库展示方法（比如噬菌体抗体文库）产生与MHC-肽复合物特异性地结合的抗体或其抗原结合部分。在一些实施方案中，可以生成突变体Fab、scFv或其他抗体形式的噬菌体展示文库，例如，其中文库的成员在一个或多个CDR或CDR的一个或多个残基处被突变。参见例如美国公开申请号US20020150914、US2014/0294841；和Cohen CJ.等人（2003）*J Mol. Recogn.* 16:324-332。

[0424] 本文中的术语“抗体”以最广义的方式加以使用，包括多克隆和单克隆抗体，包括完整抗体和功能性（抗原结合）抗体片段，包括片段抗原结合（Fab）片段、F(ab')₂片段、Fab'片段、Fv片段、重组IgG（rIgG）片段、能够特异性结合抗原的可变重链（V_H）区、单链抗体片

段,包括单链可变片段(scFv)和单结构域抗体(例如sdAb、sdFv、纳米抗体)片段。该术语涵盖基因工程和/或其他修饰形式的免疫球蛋白,比如胞内抗体、肽体、嵌合抗体、全人抗体、人源化抗体和异源偶联抗体、多特异性(例如双特异性)抗体、双抗体、三抗体和四抗体、串联di-scFv、串联tri-scFv。除非另外说明,否则术语“抗体”应理解为涵盖其功能性抗体片段。该术语还涵盖完整或全长抗体,包括任何类型或亚型的抗体,包括IgG及其亚型、IgM、IgE、IgA和IgD。

[0425] 在一些实施方案中,抗原结合蛋白、抗体及其抗原结合片段特异性识别全长抗体的抗原。在一些实施方案中,抗体的重链和轻链可以是全长的或者可以是抗原结合部分(Fab、F(ab')₂、Fv或单链Fv片段(scFv))。在其他实施方案中,抗体重链恒定区选自例如IgG1、IgG2、IgG3、IgG4、IgM、IgA1、IgA2、IgD和IgE,特别是选自例如IgG1、IgG2、IgG3和IgG4,更特别地为IgG1(例如,人IgG1)。在另一个实施方案中,抗体轻链恒定区选自例如κ或λ,特别是κ。

[0426] 其中提供的抗体为抗体片段。“抗体片段”指的是除了完整抗体之外的分子,其包含与完整抗体所结合的抗原结合的所述完整抗体的一部分。抗体片段的实例包括但不限于Fv、Fab、Fab'、Fab'-SH、F(ab')₂;双抗体;线性抗体;可变重链(V_H)区、单链抗体分子如scFv和单结构域V_H单一抗体;和由抗体片段形成的多特异性抗体。在具体的实施方案中,抗体是包含可变重链区和/或可变轻链区的单链抗体片段,比如scFv。

[0427] 术语“可变区”或“可变结构域”是指涉及抗体与抗原结合的抗体重链或轻链的结构域。天然抗体的重链和轻链的可变结构域(分别为V_H和V_L)通常具有相似的结构,每个结构域包含四个保守的框架区(FR)和三个CDR。(参见例如Kindt等人,Kuby Immunology,6th ed.,W.H.Freeman和Co.,第91页(2007)。单个V_H或V_L结构域可足以赋予抗原结合特异性。此外,可以使用来自结合抗原的抗体的V_H或V_L结构域来分离结合特定抗原的抗体,以分别筛选互补的V_L或V_H结构域的文库。参见,例如,Portolano等人,J.Immunol.150:880-887(1993);Clarkson等人,Nature 352:624-628(1991)。

[0428] 单结构域抗体是包含抗体的全部或部分重链可变结构域或全部或部分轻链可变结构域的抗体片段。在某些实施方案中,单结构域抗体是人单结构域抗体。在一些实施方案中,CAR包含特异性结合抗原的抗体重链结构域,所述抗原比如癌症标志物或被靶向的细胞(比如肿瘤细胞或癌细胞)或疾病的细胞表面抗原,比如本文所述或已知的任何靶抗原。

[0429] 可以通过各种技术制造抗体片段,包括但不限于完整抗体的蛋白水解消化以及通过重组宿主细胞产生。在一些实施方案中,抗体是重组产生的片段,比如包含非天然存在的排列的片段,比如具有通过合成接头(例如肽接头)连接的两个或更多个抗体区或链的片段,和/或不能通过酶消化天然存在的完整抗体来产生的片段。在一些实施方案中,抗体片段是scFv。

[0430] “人源化”抗体是其中所有或基本上所有CDR氨基酸残基衍生自非人CDR并且所有或基本上所有FR氨基酸残基衍生自人FR的抗体。人源化抗体任选地可以包括衍生自人抗体的抗体恒定区的至少一部分。非人抗体的“人源化形式”是指经历了人源化的非人抗体的变体,通常为了降低对人的免疫原性,同时保留亲本非人抗体的特异性和亲和力。在一些实施方案中,用来自非人抗体(例如,自其衍生CDR残基的抗体)的相应残基取代人源化抗体中的一些FR残基,例如,用于恢复或提高抗体特异性或亲和力。

[0431] 因此,在一些实施方案中,嵌合抗原受体(包括TCR样CAR)包括含有抗体或抗体片段的细胞外部分。在一些实施方案中,抗体或片段包括scFv。在一些方面,嵌合抗原受体包括含有抗体或片段的细胞外部分和细胞内信号传导区。在一些实施方案中,细胞内信号传导区包含细胞内信号传导结构域。在一些实施方案中,细胞内信号传导结构域是或包含初级信号传导结构域、能够在T细胞中诱导初级活化信号的信号传导结构域、T细胞受体(TCR)组分的信号传导结构域、和/或包含免疫受体酪氨酸活化基序(ITAM)的信号传导结构域。

[0432] 在一些实施方案中,重组受体,比如CAR,比如其抗体模块,进一步包括间隔物,其可以是或包括免疫球蛋白恒定区的至少一部分或其变体或修饰形式,比如铰链区,例如IgG4铰链区和/或CH1/CL和/或Fc区。在一些实施方案中,重组受体进一步包含间隔物和/或铰链区。在一些实施方案中,恒定区或部分是人IgG(比如IgG4或IgG1)的恒定区或部分。在一些方面,恒定区的部分用作抗原识别组分(例如scFv)和跨膜结构域之间的间隔区。间隔物的长度与不存在间隔物相比可以在抗原结合后提供增加的细胞反应性。示例性间隔物,例如铰链区,包括在国际专利申请公开号W02014031687中描述的那些。在一些实例中,间隔物的长度为或约为12个氨基酸或其长度不超过12个氨基酸。示例性间隔物包括具有至少约10至229个氨基酸、约10至200个氨基酸、约10至175个氨基酸、约10至150个氨基酸、约10至125个氨基酸、约10至100个氨基酸、约10至75个氨基酸、约10至50个氨基酸、约10至40个氨基酸、约10至30个氨基酸、约10至20个氨基酸、或约10至15个氨基酸(并且包括任何列出范围的端点之间的任何整数)的那些间隔物。在一些实施方案中,间隔区具有约12个或更少的氨基酸,约119个或更少的氨基酸,或约229个或更少的氨基酸。示例性间隔物包括单独的IgG4铰链、与CH2和CH3结构域连接的IgG4铰链、或与CH3结构域连接的IgG4铰链。示例性间隔物包括但不限于在Hudecek等人(2013)Clin.Cancer Res.,19:3153、Hudecek等人(2015)Cancer Immunol Res.3(2):125-135或国际专利申请公开号W02014031687、美国专利No.8,822,647或公开的申请号US2014/0271635中描述的那些。在一些实施方案中,恒定区或部分是人IgG(比如IgG4或IgG1)的恒定区或部分。在一些实施方案中,间隔物具有在SEQ ID NO:1列出的序列,并且由SEQ ID NO:2中列出的序列编码。在一些实施方案中,间隔物具有在SEQ ID NO:3中列出的序列。在一些实施方案中,间隔物具有在SEQ ID NO:4中列出的序列。

[0433] 在一些方面,间隔物是多肽间隔物,其(a)包含免疫球蛋白铰链的全部或部分或其修饰形式或由其组成,或包含约15个或更少的氨基酸,并且不包含CD28细胞外区域或CD8细胞外区域,(b)包含免疫球蛋白铰链(任选地IgG4铰链)的全部或部分或其修饰形式或由其组成,和/或包含约15个或更少的氨基酸,并且不包含CD28细胞外区域或CD8细胞外区域,或(c)长度为或约为12个氨基酸和/或包含免疫球蛋白铰链(任选地IgG4)的全部或部分或其修饰形式或由其组成;或(d)由SEQ ID NO:1,3-5,27-34或58中列出的氨基酸序列或与其具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的任何前述变体组成或包含所述氨基酸序列,或(e)包含式X₁PPX₂P或由其组成,其中X₁是甘氨酸、半胱氨酸或精氨酸,并且X₂是半胱氨酸或苏氨酸。

[0434] 在一些实施方案中,恒定区或部分是人IgD的恒定区或部分。在一些实施方案中,间隔物具有在SEQ ID NO:5中列出的序列。在一些实施方案中,间隔物具有表现出与SEQ ID NO:1,3,4和5的任一者具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列。

[0435] 这个抗原识别结构域通常与一种或多种细胞内信号传导组分连接,所述细胞内信号传导组分例如在CAR的情况下通过抗原受体复合物(比如TCR复合物)模拟激活和/或经由另一种细胞表面受体模拟信号的信号传导组分。在一些实施方案中,所述信号可以是免疫刺激和/或共刺激的。在一些实施方案中,它可以是抑制性的,例如免疫抑制性的。因此,在一些实施方案中,抗原结合组分(例如,抗体)与一个或多个跨膜和细胞内信号传导结构域和/或区域连接。在一些实施方案中,跨膜结构域与细胞外结构域融合。在一个实施方案中,使用与受体(例如CAR)中的结构域之一天然相关的跨膜结构域。在一些情况下,通过氨基酸取代选择或修饰跨膜结构域,以避免这样的结构域与相同或不同表面膜蛋白的跨膜结构域结合,从而将与受体复合物的其他成员的相互作用最小化。

[0436] 在一些实施方案中,跨膜结构域衍生自天然或合成来源。在来源是天然的情况下,在一些方面,结构域衍生自任何膜结合蛋白或跨膜蛋白。跨膜区包括衍生自T细胞受体的 α 、 β 或 ζ 链、CD28、CD3 ϵ 、CD45、CD4、CD5、CD8、CD9、CD16、CD22、CD33、CD37、CD64、CD80、CD86、CD134、CD137、CD154(即至少包含其跨膜区)的那些跨膜区和/或含有其功能变体的跨膜区,所述功能变体比如保留其结构(例如跨膜)特性的实质性部分的变体。在一些实施方案中,跨膜结构域是衍生自CD4、CD28或CD8(例如CD8 α)或其功能变体的跨膜结构域。一些实施方案中的跨膜结构域是合成的。在一些方面,合成跨膜结构域主要包含疏水残基,比如亮氨酸和缬氨酸。在一些方面,苯丙氨酸、色氨酸和缬氨酸的三联体将见于合成跨膜结构域的每个末端。在一些实施方案中,通过接头、间隔物和/或跨膜结构域加以连接。

[0437] 其中细胞内信号传导结构域是通过天然抗原受体模拟或接近信号、通过这种受体与共刺激受体联合模拟或接近信号、和/或仅通过共刺激受体模拟或接近信号的那些细胞内信号传导结构域。在一些实施方案中,存在短的寡肽或多肽接头,例如长度为2至10个氨基酸的接头,比如含有甘氨酸和丝氨酸(例如甘氨酸-丝氨酸双联体)的接头,并且所述接头形成在CAR的跨膜结构域和细胞质信号传导结构域之间的连接。

[0438] 受体(例如CAR)通常包括至少一种细胞内信号传导组分。在一些实施方案中,受体包括TCR复合物的细胞内组分,比如介导T细胞活化和细胞毒性的TCR CD3链,例如CD3 ζ 链。因此,在一些方面,抗原结合部分与一个或多个细胞信号传导模块连接。在一些实施方案中,细胞信号传导模块包括CD3跨膜结构域、CD3细胞内信号传导结构域和/或其他CD跨膜结构域。在一些实施方案中,受体(例如CAR)进一步包括一种或多种另外的分子的一部分,所述分子比如Fc受体 γ 、CD8、CD4、CD25或CD16。例如,在一些方面,CAR或其他嵌合受体包括在CD3-zeta(CD3- ζ)或Fc受体 γ 与CD8、CD4、CD25或CD16之间的嵌合分子。

[0439] 在一些实施方案中,在连接CAR或其他嵌合受体时,受体的细胞质结构域或细胞内信号传导结构域和/或区域激活免疫细胞的正常效应子功能或应答中的至少一种,所述免疫细胞例如经工程化以表达CAR的T细胞。例如,在一些情况下,CAR诱导T细胞的功能,比如细胞溶解活性或T辅助活性,比如细胞因子或其他因子的分泌。在一些实施方案中,使用抗原受体组分或共刺激分子的细胞内信号传导结构域的截短部分代替完整的免疫刺激链,例如,条件是它转导效应子功能信号。在一些实施方案中,一个或多个细胞内信号传导结构域和/或区域包括T细胞受体(TCR)的细胞质序列,并且在一些方面还包括在自然情况下与这种受体协同作用以在抗原受体结合后启动信号转导的共受体的那些序列、和/或这样的分子的任何衍生物或变体、和/或具有相同功能能力的任何合成序列。

[0440] 在天然TCR的情况下,完全激活通常不仅需要通过TCR传导信号,还需要共刺激信号。因此,在一些实施方案中,为了促进完全激活,用于生成二级信号或共刺激信号的组分也包括在CAR中。在其他实施方案中,CAR不包括用于生成共刺激信号的组分。在一些方面,另外的CAR在同一细胞中表达,并提供用于生成二级信号或共刺激信号的组分。

[0441] T细胞活化在一些方面被描述为由两类细胞质信号传导序列介导:通过TCR启动抗原依赖性初级活化的那些序列(初级细胞质信号传导序列),以及以抗原非依赖性方式起作用以提供二级信号或共刺激信号的那些序列(二级细胞质信号传导序列)。在一些方面,CAR包括这样的信号传导组分中的一种或两种。

[0442] 在一些方面,CAR包括源自信号传导分子或结构域的初级细胞质信号传导序列,其在自然环境中促进TCR复合物的初级活化。以刺激方式起作用的初级细胞质信号传导序列可含有信号传导基序,其被称为免疫受体酪氨酸活化基序或ITAM。含有初级细胞质信号传导序列的ITAM的实例包括衍生自TCR或CD3 ζ 链、FcR γ 、CD3 γ 、CD3 δ 和CD3 ϵ 或FcR β 的那些。在一些实施方案中,CAR中的细胞质信号传导分子含有细胞质信号传导结构域、其部分或源自CD3 ζ 的序列。

[0443] 在一些实施方案中,CAR包括共刺激受体的信号传导结构域和/或跨膜部分,所述共刺激受体比如CD28、4-1BB、OX40、DAP10和ICOS。在一些方面,同一个CAR包括活化组分和共刺激组分两者。

[0444] 在一些实施方案中,活化结构域包含在一个CAR内,而共刺激组分由识别呈递在同一个细胞上的另一种抗原的另一个CAR提供。在一些实施方案中,CAR包括活化或刺激性CAR、共刺激CAR,两者都在同一个细胞上表达(参见WO2014/055668)。在一些方面,细胞包含一种或多种刺激或活化性CAR和/或共刺激CAR。在一些实施方案中,细胞进一步包含抑制性CAR(iCAR,参见Fedorov等人,Sci.Transl.Medicine,5(215)(2013年12月),比如识别除与疾病或病症相关和/或对疾病或病症特异的抗原之外的抗原的CAR,由此通过抑制性CAR与其配体的结合来减少或抑制通过疾病靶向的CAR递送的活化信号,以便例如减少脱靶效应。

[0445] 在一些实施方案中,重组受体(比如CAR)的细胞内信号传导组分包含CD3 ζ 细胞内结构域和共刺激信号传导区。在某些实施方案中,细胞内信号传导结构域包含与CD3(例如CD3- ζ)细胞内结构域连接的CD28跨膜和信号传导结构域。在一些实施方案中,细胞内信号传导结构域包含与CD3 ζ 细胞内结构域连接的嵌合的CD28和CD137(4-1BB, TNFRSF9)共刺激结构域。

[0446] 在一些实施方案中,CAR在细胞质部分中包括一个或多个,例如两个或更多个共刺激结构域和活化结构域,例如初级活化结构域。示例性CAR包括CD3- ζ 、CD28和4-1BB的细胞内组分。

[0447] 在一些实施方案中,CAR或其他受体进一步包括标志物,比如细胞表面标志物,所述标志物可用于证实表达受体的细胞转导或工程化,所述受体比如截短形式的细胞表面受体,比如截短的EGFR(tEGFR)。在一些方面,标志物包括CD34、NGFR或表皮生长因子受体(例如,tEGFR)的全部或部分(例如,截短形式)。在一些实施方案中,编码标志物的核酸与编码接头序列的多核苷酸可操作连接,所述接头序列比如可切割的接头序列,例如T2A。例如,标志物和任选的接头序列可以是在公开的专利申请号WO2014031687中公开的任何一种。例如,标志物可以是任选地与接头序列(比如T2A可切割的接头序列)连接的截短的EGFR

(tEGFR)。在一些实施方案中,标志物是非天然存在于T细胞上或非天然存在于T细胞表面上的分子(例如细胞表面蛋白)或其部分。

[0448] 在一些实施方案中,所述分子是非自身分子,例如非自身蛋白质,即未被其中有待过继转移细胞的宿主的免疫系统识别为“自身”的蛋白质。

[0449] 在一些实施方案中,除了用作基因工程的标志物(例如,用于选择成功工程化的细胞)之外,标志物没有治疗功能和/或不产生作用。在其他实施方案中,标志物可以是治疗分子或以别的方式发挥某一期望作用的分子,比如在体内将会遇到的细胞的配体,比如在过继转移和遇到配体后增强和/或抑制细胞应答的共刺激或免疫检查点分子。

[0450] 在一些情况下,CAR被称为第一代、第二代和/或第三代CAR。在一些方面,第一代CAR是在抗原结合后只提供CD3链诱导的信号信号的CAR;在一些方面,第二代CAR是提供这种信号和共刺激信号的CAR,比如包括来自共刺激受体如CD28或CD137的细胞内信号传导结构域的CAR;在一些方面,第三代CAR是包括不同的共刺激受体的多个共刺激结构域的CAR。

[0451] 在一些实施方案中,嵌合抗原受体包括细胞外部分,所述细胞外部分含有抗原结合结构域,比如抗体或抗原结合抗体片段,比如scFv或Fv。在一些方面,嵌合抗原受体包括含有抗体或片段的细胞外部分和细胞内信号传导结构域。在一些实施方案中,抗体或片段包括scFv或单结构域V_H抗体,并且细胞内结构域含有ITAM。在一些方面,细胞内信号传导结构域包括CD3-zeta (CD3 ζ) 链的 ζ 链的信号传导结构域。在一些实施方案中,嵌合抗原受体包括连接细胞外结构域和细胞内信号传导结构域的跨膜结构域。在一些方面,跨膜结构域含有CD28的跨膜部分。在一些实施方案中,嵌合抗原受体含有T细胞共刺激分子的细胞内结构域。细胞外结构域和跨膜结构域可以直接或间接连接。在一些实施方案中,细胞外结构域和跨膜结构域通过间隔物连接,比如本文所述的任何间隔物。在一些实施方案中,受体含有衍生跨膜结构域的分子的细胞外部分,比如CD28细胞外部分。在一些实施方案中,嵌合抗原受体含有衍生自T细胞共刺激分子或其功能变体的细胞内结构域,比如在跨膜结构域和细胞内信号传导结构域之间。在一些方面,T细胞共刺激分子是CD28或41BB。

[0452] 例如,在一些实施方案中,CAR含有抗体(例如抗体片段)、作为或含有CD28的跨膜部分或其功能变体的跨膜结构域、和细胞内信号传导结构域,所述细胞内信号传导结构域含有CD28的信号传导部分或其功能变体和CD3 ζ 的信号传导部分或其功能变体。在一些实施方案中,CAR含有抗体(例如抗体片段)、作为或含有CD28的跨膜部分或其功能变体的跨膜结构域、和细胞内信号传导结构域,所述细胞内信号传导结构域含有4-1BB的信号传导部分或其功能变体和CD3 ζ 的信号传导部分或其功能变体。在一些这样的实施方案中,受体进一步包括含有Ig分子(比如人Ig分子)的一部分(比如Ig铰链,例如IgG4铰链)的间隔物,比如仅有铰链的间隔物。

[0453] 在一些实施方案中,受体(例如CAR)的跨膜结构域是人CD28的跨膜结构域或其变体,例如人CD28(登录号:P10747.1)的27-氨基酸跨膜结构域,或是这样的跨膜结构域,其包含在SEQ ID NO:8中列出的氨基酸序列或表现出与SEQ ID NO:8具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列;在一些实施方案中,重组受体的含有跨膜结构域的部分包含在SEQ ID NO:9列出的氨基酸序列或与其具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列。

[0454] 在一些实施方案中,细胞内信号传导区包含人CD28的细胞内共刺激信号传导结构域或其功能变体或部分,比如其41个氨基酸结构域和/或比如在天然CD28蛋白的位置186-187具有LL至GG取代的结构域。在一些实施方案中,细胞内信号传导结构域可包含在SEQ ID NO:10或11中列出的氨基酸序列或表现出与SEQ ID NO:10或11具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列。在一些实施方案中,细胞内区包含4-1BB的细胞内共刺激信号传导结构域或其功能变体或部分,比如人4-1BB(登录号Q07011.1)的42-氨基酸细胞质结构域或其功能变体或部分,比如在SEQ ID NO:12中列出的氨基酸序列或表现出与SEQ ID NO:12具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列。

[0455] 在一些实施方案中,细胞内信号传导区包含人CD3链任选地人CD3 ζ 刺激信号传导结构域或其功能变体,比如人CD3 ζ (登录号:P20963.2)的同工型3的112AA细胞质结构域或CD3 ζ 信号传导结构域,如在美国专利No.7,446,190或美国专利No.8,911,993中描述。在一些实施方案中,细胞内信号传导区包含在SEQ ID NO:13、14或15中列出的氨基酸序列或表现出与SEQ ID NO:13、14或15具有至少或至少约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多序列同一性的氨基酸序列。

[0456] 在一些方面,间隔物仅含有IgG的铰链区,比如仅有IgG4或IgG1的铰链,比如在SEQ ID NO:1中列出的仅有铰链的间隔物。在其他实施方案中,间隔物是Ig铰链,例如,与C_H2和/或C_H3结构域连接的IgG4铰链。在一些实施方案中,间隔物是Ig铰链,例如与C_H2和C_H3结构域连接的IgG4铰链,比如在SEQ ID NO:3列出。在一些实施方案中,间隔物是Ig铰链,例如仅与C_H3结构域连接的IgG4铰链,比如在SEQ ID NO:4中列出。在一些实施方案中,间隔物是或包含富含甘氨酸-丝氨酸序列或其他柔性接头,比如已知的柔性接头。

[0457] 在一些实施方案中,嵌合抗原受体含有T细胞共刺激分子的细胞内结构域。在一些方面,T细胞共刺激分子是CD28或41BB。

[0458] 例如,在一些实施方案中,CAR包括抗体(比如抗体片段),包括scFv;间隔物,比如含有免疫球蛋白分子的一部分(比如重链分子的铰链区和/或一个或多个恒定区)的间隔物,比如含有Ig-铰链的间隔物;含有CD28衍生的跨膜结构域的全部或部分的跨膜结构域;CD28衍生的细胞内信号传导结构域;和CD3 ζ 信号传导结构域。在一些实施方案中,CAR包括抗体或片段(比如scFv)、间隔物(比如任何含有Ig铰链的间隔物)、CD28衍生的跨膜结构域、4-1BB衍生的细胞内信号传导结构域和CD3 ζ 衍生的信号传导结构域。

[0459] 在一些实施方案中,编码这样的CAR构建体的核酸分子进一步包括编码T2A核糖体跳跃元件和/或tEGFR序列的序列,例如在编码CAR的序列的下游。在一些实施方案中,还可以生成表达抗原受体(例如CAR)的T细胞,以表达截短的EGFR(EGFRt)作为非免疫原性选择表位(例如通过引入编码由T2A核糖体开关分离的CAR和EGFRt的构建体,从同一个构建体表达两种蛋白质),然后其可用作检测这样的细胞的标志物(参见例如美国专利No.8,802,374)。

[0460] 术语“多肽”和“蛋白质”可互换使用,是指氨基酸残基的聚合物,并且不限于最小长度。包括所提供的受体及其他多肽(例如接头或肽)在内的多肽,可包含包括天然和/或非天然氨基酸残基在内的氨基酸残基。这些术语还包括多肽的表达后修饰,例如糖基化、唾液

酸化、乙酰化和磷酸化。在一些方面,多肽可以含有相对于自然或天然序列的修饰,只要蛋白质保持期望的活性即可。这些修饰可以是有意的,比如通过定点诱变,或者可以是偶然的,比如通过产生蛋白质的宿主的突变或由于PCR扩增引起的错误。

T细胞受体 (TCR)

[0461] 在一些实施方案中,提供了表达识别靶多肽的肽表位或T细胞表位的T细胞受体(TCR)或其抗原结合部分的工程化细胞(比如T细胞),所述靶多肽比如肿瘤抗原、病毒蛋白或自身免疫蛋白。

[0462] 在一些实施方案中,“T细胞受体”或“TCR”是含有可变 α 和 β 链(也分别称为TCR α 和TCR β)或可变 γ 和 δ 链(也分别称为TCR α 和TCR β)的分子或其抗原结合部分,其能够与结合MHC分子的肽特异性地结合。在一些实施方案中,TCR呈 $\alpha\beta$ 形式。通常,以 $\alpha\beta$ 和 $\gamma\delta$ 形式存在的TCR通常在结构上相似,但表达它们的T细胞可具有不同的解剖位置或功能。TCR可见于细胞表面上或呈可溶形式。通常,TCR见于T细胞(或T淋巴细胞)表面上,它在这里通常负责识别与主要组织相容性复合体(MHC)分子结合的抗原。

[0463] 除非另外说明,否则术语“TCR”应理解为涵盖完整的TCR以及其抗原结合部分或其抗原结合片段。在一些实施方案中,TCR是完整或全长TCR,包括 $\alpha\beta$ 形式或 $\gamma\delta$ 形式的TCR。在一些实施方案中,TCR是抗原结合部分,其小于全长TCR但与MHC分子中结合的特异性肽结合,例如与MHC-肽复合物结合。在一些情况下,TCR的抗原结合部分或片段可仅含全长或完整TCR的结构域的一部分,但仍然能够结合肽表位,比如完整TCR所结合的MHC-肽复合物。在一些情况下,抗原结合部分含有TCR的可变结构域,比如TCR的可变 α 链和可变 β 链,其足以形成与特异性MHC-肽复合物的结合位点。通常,TCR的可变链含有涉及识别肽、MHC和/或MHC-肽复合物的互补决定区。

[0464] 在一些实施方案中,TCR的可变结构域含有高变环或互补决定区(CDR),所述高变环或互补决定区通常是抗原识别和结合能力和特异性的主要贡献者。在一些实施方案中,TCR的CDR或其组合形成了给定TCR分子的全部或基本上全部的抗原结合位点。TCR链的可变区内的各个CDR通常被框架区(FR)分开,框架区(FR)与CDR相比通常在TCR分子中展示较小的变异性(参见,例如Jores等人,Proc.Nat'l Acad.Sci.U.S.A.87:9138,1990;Chothia等人,EMBO J.7:3745,1988;还参见Lefranc等人,Dev.Comp.Immunol.27:55,2003)。在一些实施方案中,CDR3是负责抗原结合或特异性的主要CDR,或者是在用于抗原识别和/或与肽-MHC复合物的加工肽部分相互作用的给定TCR可变区的三个CDR中最重要CDR。在一些情况下, α 链的CDR1可与某些抗原肽的N端部分相互作用。在一些情况下, β 链的CDR1可与肽的C端部分相互作用。在一些情况下,CDR2对与MHC-肽复合物的MHC部分的相互作用或对其识别起作用最强,或是主要负责的CDR。在一些实施方案中, β 链的可变区可含有另外的高变区(CDR4或HVR4),其通常参与超抗原结合而不是抗原识别(Kotb(1995)Clinical Microbiology Reviews,8:411-42)。

[0465] 在一些实施方案中,TCR还可含有恒定结构域、跨膜结构域和/或短胞质尾(参见,例如,Janeway等人,Immunobiology:The Immune System in Health and Disease,3rd Ed.,Current Biology Publications,p.4:33,1997)。在一些方面,TCR的每条链可具有一个N末端免疫球蛋白可变结构域、一个免疫球蛋白恒定结构域、跨膜区和在C末端的短胞质尾。在一些实施方案中,TCR与涉及介导信号转导的CD3复合物的不变蛋白质相关。

[0466] 在一些实施方案中,TCR链含有一个或多个恒定结构域。例如,给定TCR链(例如 α 链或 β 链)的细胞外部分可以含有两个免疫球蛋白样结构域,比如可变结构域(例如, $V\alpha$ 或 $V\beta$;通常氨基酸1至116基于Kabat编号,Kabat等人,“Sequences of Proteins of Immunological Interest,US Dept.Health and Human Services,Public Health Service National Institutes of Health,1991,5th ed.)和邻近细胞膜的恒定结构域(例如, α 链恒定结构域或 $C\alpha$,通常链的位置117至259基于Kabat编号,或 β 链恒定结构域或 $C\beta$,通常链的位置117至295基于Kabat)。例如,在一些情况下,由两条链形成的TCR的细胞外部分含有两个膜近端恒定结构域和两个膜远端可变结构域,所述可变结构域各自含有CDR。TCR的恒定结构域可以含有短连接序列,其中半胱氨酸残基形成二硫键,从而连接TCR的两条链。在一些实施方案中,TCR可以在 α 链和 β 链中各自具有另一个半胱氨酸残基,使得TCR在恒定结构域中含有两个二硫键。

[0467] 在一些实施方案中,TCR链含有跨膜结构域。在一些实施方案中,跨膜结构域带正电荷。在一些情况下,TCR链含有胞质尾。在一些情况下,该结构允许TCR与CD3及其亚基等其他分子结合。例如,含有具有跨膜区的恒定结构域的TCR可以将蛋白质锚定在细胞膜中并与CD3信号传导装置或复合物的不变亚基结合。CD3信号传导亚基(例如CD3 γ 、CD3 δ 、CD3 ϵ 和CD3 ζ 链)的胞内尾含有一个或多个涉及TCR复合物的信号传导能力的免疫受体酪氨酸活化基序或ITAM。

[0468] 在一些实施方案中,TCR可以是两条链 α 和 β (或任选地 γ 和 δ)的异二聚体,或者它可以是单链TCR构建体。在一些实施方案中,TCR是含有两条独立链(α 和 β 链或 γ 和 δ 链)的异二聚体,所述两条链例如通过一个或多个二硫键连接。

[0469] 在一些实施方案中,TCR可以从已知的TCR序列比如 $V\alpha$ 、 $V\beta$ 链序列生成,针对所述序列的基本上全长的编码序列是容易获得的。从细胞来源获得全长TCR序列(包括V链序列)的方法是熟知的。在一些实施方案中,编码TCR的核酸可以从多种来源获得,例如通过一个或多个给定细胞内的或从其分离的TCR编码核酸的聚合酶链反应(PCR)扩增,或可公开获得的TCR DNA序列的合成。

[0470] 在一些实施方案中,TCR获自生物来源,比如来自细胞,比如来自T细胞(例如细胞毒性T细胞)、T细胞杂交瘤或其他可公开获得的来源。在一些实施方案中,可以从体内分离的细胞获得T细胞。在一些实施方案中,TCR是胸腺选择的TCR。在一些实施方案中,TCR是新表位限制的(neoepitope-restricted)TCR。在一些实施方案中,T细胞可以是培养的T细胞杂交瘤或克隆。在一些实施方案中,可以根据TCR序列的知识来合成产生TCR或其抗原结合部分或其抗原结合片段。

[0471] 在一些实施方案中,从鉴定自或选自文库筛选的TCR生成所述TCR,所述文库为针对靶多肽抗原或其靶T细胞表位的候选TCR的文库。通过扩增来自分离自受试者的T细胞(包括存在于PBMC、脾或其他淋巴器官中的细胞)的 $V\alpha$ 和 $V\beta$ 库,可以生成TCR文库。在一些情况下,T细胞可以从肿瘤浸润淋巴细胞(TIL)扩增而来。在一些实施方案中,可以从CD4⁺或CD8⁺细胞生成TCR文库。在一些实施方案中,TCR可以从健康受试者的正常T细胞来源即正常TCR文库扩增而来。在一些实施方案中,TCR可以从患病受试者的T细胞来源即患病的TCR文库扩增而来。在一些实施方案中,例如通过在从人获得的样品(比如T细胞)中的RT-PCR,使用简并引物来扩增 $V\alpha$ 和 $V\beta$ 的基因库。在一些实施方案中,scTv文库可以从幼稚的 $V\alpha$ 和 $V\beta$ 文库组

装而来,其中扩增产物被克隆或组装为通过接头分开。取决于受试者和细胞的来源,文库可以是HLA等位基因特异性的。或者,在一些实施方案中,可通过亲本或支架TCR分子的诱变或多样化生成TCR文库。在一些方面,比如通过例如 α 或 β 链的诱变,TCR经历定向进化。在一些方面,TCR的CDR内的特定残基被改变。在一些实施方案中,可以通过亲和力成熟来修饰选择的TCR。在一些实施方案中,例如通过筛选以评估针对肽的CTL活性,可以选择抗原特异性T细胞。在一些方面,比如通过结合活性,例如对抗原的特定亲和力或亲合力,可以选择例如存在于抗原特异性T细胞上的TCR。

[0472] 在一些实施方案中,基因工程化抗原受体包括重组T细胞受体(TCR)和/或从天然存在的T细胞克隆的TCR。在一些实施方案中,从患者鉴定、分离针对靶抗原(例如,癌抗原)的高亲和力T细胞克隆,并且将其引入到细胞中。在一些实施方案中,已经在用人免疫系统基因(例如人白细胞抗原系统或HLA)工程化的转基因小鼠中生成了针对靶抗原的TCR克隆。参见,例如,肿瘤抗原(参见,例如,Parkhurst等人(2009) *Clin Cancer Res.* 15:169-180和Cohen等人(2005) *J Immunol.* 175:5799-5808。在一些实施方案中,使用噬菌体展示来分离针对靶抗原的TCR(参见,例如,Varela-Rohena等人(2008) *Nat Med.* 14:1390-1395和Li(2005) *Nat Biotechnol.* 23:349-354。

[0473] 在一些实施方案中,TCR或其抗原结合部分是经过修饰或工程改造的TCR或其抗原结合部分。在一些实施方案中,使用定向进化方法来生成具有改变的性质的TCR,比如其对特异性MHC-肽复合物具有更高的亲和力。在一些实施方案中,通过展示方法实现定向进化,所述展示方法包括但不限于酵母展示(Holler等人(2003) *Nat Immunol.* 4,55-62;Holler等人(2000) *Proc Natl Acad Sci U S A.* 97,5387-92)、噬菌体展示(Li等人(2005) *Nat Biotechnol.* 23,349-54)、或T细胞展示(Chervin等人(2008) *J Immunol Methods.* 339,175-84)。在一些实施方案中,展示方法涉及工程改造或修饰已知的亲本或参考TCR。例如,在一些情况下,野生型TCR可以用作模板,用于产生其中CDR的一个或多个残基被突变的诱变TCR,并且选择了具有期望的改变的性质的突变体,所述改变的性质比如对期望靶抗原的更高亲和力。

[0474] 在一些实施方案中,用于产生或生成感兴趣的TCR的靶多肽的肽是已知的或者可由技术人员容易地鉴定。在一些实施方案中,基于感兴趣的靶多肽(比如下文描述的靶多肽)中的HLA限制性基序的存在,可以确定适合用于生成TCR或抗原结合部分的肽。在一些实施方案中,使用可用的计算机预测模型来鉴定肽。在一些实施方案中,关于预测MHC I类结合位点,这样的模型包括但不限于ProPred1(Singh和Raghava(2001) *Bioinformatics* 17(12):1236-1237,和SYFPEITHI(参见Schuler等人,(2007) *Immunoinformatics Methods in Molecular Biology*, 409(1):75-93 2007)。在一些实施方案中,MHC限制性表位是HLA-A0201,其在所有高加索人的大约39%-46%中表达,因此代表用于制备TCR或其他MHC-肽结合分子的MHC抗原的适当选择。

[0475] 使用计算机预测模型的HLA-A0201结合基序和蛋白酶体和免疫蛋白酶体的切割位点是已知的。关于预测MHC I类结合位点,这样的模型包括但不限于ProPred1(更详细地描述于Singh和Raghava,ProPred:prediction of HLA-DR binding sites.BIOINFORMATICS 17(12):1236-1237 2001)和SYFPEITHI(参见Schuler等人,SYFPEITHI,Database for Searching and T-Cell Epitope Prediction.in *Immunoinformatics Methods in*

Molecular Biology, vol409(1):75-93 2007)。

[0476] 在一些实施方案中,TCR或其抗原结合部分可以是重组产生的天然蛋白质或其突变形式,其中一种或多种性质比如结合特性已被改变。在一些实施方案中,TCR可以衍生自各种动物物种之一,比如人、小鼠、大鼠或其他哺乳动物。TCR可以是细胞结合形式或呈可溶形式。在一些实施方案中,为了所提供的方法的目的,TCR是在细胞表面上表达的细胞结合形式。

[0477] 在一些实施方案中,TCR是全长TCR。在一些实施方案中,TCR是抗原结合部分。在一些实施方案中,TCR是二聚体TCR(dTCR)。在一些实施方案中,TCR是单链TCR(sc-TCR)。在一些实施方案中,dTCR或scTCR具有如WO 03/020763、WO 04/033685、WO2011/044186中描述的结构。

[0478] 在一些实施方案中,TCR含有对应于跨膜序列的序列。在一些实施方案中,TCR确实含有对应于细胞质序列的序列。在一些实施方案中,TCR能够与CD3形成TCR复合物。在一些实施方案中,任何TCR(包括dTCR或scTCR)可以与在T细胞表面上产生活性TCR的信号传导结构域连接。在一些实施方案中,TCR在细胞表面上表达。

[0479] 在一些实施方案中,dTCR含有第一多肽和第二多肽,其中对应于TCR α 链可变区序列的序列与对应于TCR α 链恒定区细胞外序列的序列的N末端融合,其中对应于TCR β 链可变区序列的序列与对应于TCR β 链恒定区细胞外序列的序列的N末端融合,所述第一和第二多肽通过二硫键连接。在一些实施方案中,所述键可以对应于天然二聚体 $\alpha\beta$ TCR中存在的天然链间二硫键。在一些实施方案中,在天然TCR中不存在链间二硫键。例如,在一些实施方案中,可以将一个或多个半胱氨酸掺入dTCR多肽对的恒定区细胞外序列中。在一些情况下,天然和非天然二硫键两者都可能是期望的。在一些实施方案中,TCR含有锚定到膜上的跨膜序列。

[0480] 在一些实施方案中,dTCR含有TCR α 链和TCR β 链,所述TCR α 链含有可变 α 结构域、恒定 α 结构域和衔接至恒定 α 结构域的C末端的第一二聚化基序,所述TCR β 链包含可变 β 结构域、恒定 β 结构域和衔接至恒定 β 结构域的C末端的第二二聚化基序,其中所述第一和第二二聚化基序容易相互作用而在第一二聚化基序中的氨基酸和第二二聚化基序中的氨基酸之间形成共价键,所述共价键将TCR α 链和TCR β 链连接在一起。

[0481] 在一些实施方案中,TCR是scTCR。通常,可以使用已知的方法生成scTCR,参见例如Soo Hoo, W.F.等人,PNAS(USA) 89, 4759(1992); Wülfing, C.和Plückthun, A., J.Mol.Biol.242, 655(1994); Kurucz, I.等人,PNAS(USA) 90 3830(1993); 国际公开PCT号WO 96/13593、WO 96/18105、WO99/60120、WO99/18129、WO 03/020763、WO2011/044186; 和Schlueter, C.J.等人, J.Mol.Biol.256, 859(1996)。在一些实施方案中,scTCR含有引入的非天然链间二硫键以促进TCR链的结合(参见例如国际公开PCT号WO 03/020763)。在一些实施方案中,scTCR是非二硫键连接的截短的TCR,其中与其C末端融合的异源亮氨酸拉链促进链结合(参见例如国际公开PCT号WO99/60120)。在一些实施方案中,scTCR含有通过肽接头与TCR β 可变结构域共价连接的TCR α 可变结构域(参见例如国际公开PCT号WO99/18129)。

[0482] 在一些实施方案中,scTCR含有由对应于TCR α 链可变区的氨基酸序列构成的第一区段、由对应于TCR β 链可变区序列的氨基酸序列(与对应于TCR β 链恒定结构域细胞外序列的氨基酸序列的N末端融合)构成的第二区段、和连接所述第一区段的C末端与所述第二区

段的N末端的接头序列。

[0483] 在一些实施方案中,scTCR含有由与 α 链细胞外恒定结构域序列的N末端融合的 α 链可变区序列构成的第一区段、和由与 β 链细胞外恒定区序列和跨膜序列的N末端融合的 β 链可变区序列构成的第二区段,以及任选地,连接所述第一区段的C末端与所述第二区段的N末端的接头序列。

[0484] 在一些实施方案中,scTCR含有由与 β 链细胞外恒定结构域序列的N末端融合的TCR β 链可变区序列构成的第一区段、和由与 α 链细胞外恒定区序列和跨膜序列的N末端融合的所提供的 α 链可变区序列构成的第二区段,以及任选地,连接所述第一区段的C末端与所述第二区段的N末端的接头序列。

[0485] 在一些实施方案中,连接第一和第二TCR区段的scTCR的接头可以是能够形成单个多肽链同时保留TCR结合特异性的任何接头。在一些实施方案中,接头序列可以例如具有式-P-AA-P-,其中P是脯氨酸,AA代表氨基酸序列,其中氨基酸是甘氨酸和丝氨酸。在一些实施方案中,第一和第二区段配对,使得其可变区序列被定向为用于这种结合。因此,在一些情况下,接头具有足够的长度以跨越第一区段的C末端和第二区段的N末端之间的距离,或反之亦然,但是不能太长以至于阻断或减少scTCR与靶配体的结合。在一些实施方案中,接头可含有或含有约10至45个氨基酸,比如10至30个氨基酸或26至41个氨基酸残基,例如29、30、31或32个氨基酸。在一些实施方案中,接头具有式-PGGG-(SGGG)₅-P-,其中P是脯氨酸,G是甘氨酸,并且S是丝氨酸(SEQ ID NO:22)。在一些实施方案中,接头具有序列GSADDAKKDAAKKDGKS(SEQ ID NO:23)。

[0486] 在一些实施方案中,scTCR含有共价二硫键,其将 α 链恒定结构域的免疫球蛋白区的残基与 β 链恒定结构域的免疫球蛋白区的残基连接。在一些实施方案中,在天然TCR中不存在链间二硫键。例如,在一些实施方案中,可以将一个或多个半胱氨酸掺入到scTCR多肽的第一和第二区段的恒定区细胞外序列中。在一些情况下,天然和非天然二硫键两者都可能是期望的。

[0487] 在含有引入的链间二硫键的dTCR或scTCR的一些实施方案中,不存在天然二硫键。在一些实施方案中,形成天然链间二硫键的一个或多个天然半胱氨酸被另一个残基取代,例如被丝氨酸或丙氨酸取代。在一些实施方案中,通过将第一和第二区段上的非半胱氨酸残基突变为半胱氨酸,可以形成引入的二硫键。TCR的示例性非天然二硫键描述于公开的国际PCT号W02006/000830中。

[0488] 在一些实施方案中,TCR或其抗原结合片段对靶抗原表现出这样的亲和力,其平衡结合常数在或在约 10^{-5} 到 10^{-12} M之间,并且包括其中的所有个体值和范围。在一些实施方案中,靶抗原是MHC-肽复合物或配体。

[0489] 在一些实施方案中,可以通过PCR、克隆或其他适合的手段将编码TCR比如 α 和 β 链的一个或多个核酸扩增,并克隆到适合的一种或多种表达载体中。表达载体可以是任何适合的重组表达载体,并且可以用于转化或转染任何适合的宿主。适合的载体包括设计为用于增殖和扩增或用于表达或用于这两者的载体,比如质粒和病毒。

[0490] 在一些实施方案中,载体可以是pUC系列(Fermentas Life Sciences)、pBluescript系列(Stratagene,LaJolla,Calif.)、pET系列(Novagen,Madison,Wis.)、pGEX系列(Pharmacia Biotech,Uppsala,Sweden)或pEX系列(Clontech,Palo Alto,Calif.)的

载体。在一些情况下,也可以使用噬菌体载体,比如 λ G10、 λ GT11、 λ ZapII (Stratagene)、 λ EMBL4和 λ NM1149。在一些实施方案中,可以使用植物表达载体,其包括pBI01、pBI101.2、pBI101.3、pBI121和pBIN19 (Clontech)。在一些实施方案中,动物表达载体包括pEUK-C1、pMAM和pMAMneo (Clontech)。在一些实施方案中,使用病毒载体,比如逆转录病毒载体。

[0491] 在一些实施方案中,可以使用标准重组DNA技术制备重组表达载体。在一些实施方案中,在适当情况下并且考虑载体是基于DNA还是基于RNA,载体可含有调控序列,例如转录和翻译起始和终止密码子,所述调控序列对其中有待引入该载体的宿主(例如,细菌、真菌、植物或动物)的类型是特异的。在一些实施方案中,载体可含有与编码TCR或抗原结合部分(或其他MHC-肽结合分子)的核苷酸序列可操作连接的非天然启动子。在一些实施方案中,启动子可以是非病毒启动子或病毒启动子,比如巨细胞病毒(CMV)启动子、SV40启动子、RSV启动子和在鼠干细胞病毒的长末端重复序列中发现的启动子。还考虑了其他已知的启动子。

[0492] 在一些实施方案中,在获得T细胞克隆之后,分离TCR α 和 β 链并将其克隆到基因表达载体中。在一些实施方案中,TCR α 和 β 基因经由小核糖核酸病毒2A核糖体跳跃肽连接,使得两条链共表达。在一些实施方案中,通过逆转录病毒或慢病毒载体或通过转座子完成TCR的基因转移(参见,例如,Baum等人(2006)Molecular Therapy:The Journal of the American Society of Gene Therapy.13:1050-1063;Frecha等人(2010)Molecular Therapy:The Journal of the American Society of Gene Therapy.18:1748-1757;和Hackett等人(2010)Molecular Therapy:The Journal of the American Society of Gene Therapy.18:674-683。

[0493] 在一些实施方案中,为了生成编码TCR的载体,将分离自表达感兴趣的TCR的T细胞克隆的总cDNA进行 α 链和 β 链PCR扩增,并将其克隆到表达载体中。在一些实施方案中,将 α 链和 β 链克隆到同一载体中。在一些实施方案中,将 α 链和 β 链克隆到不同载体中。在一些实施方案中,将生成的 α 链和 β 链组入逆转录病毒载体例如慢病毒载体中。

嵌合自身抗体受体(CAAR)

[0494] 在一些实施方案中,重组受体是嵌合自身抗体受体(CAR)。在一些实施方案中,CAAR对于自身抗体是特异的。在一些实施方案中,表达CAAR的细胞,例如经工程改造以表达CAAR的T细胞,可用于特异性地结合并杀伤表达自身抗体的细胞,而不杀伤表达正常抗体的细胞。在一些实施方案中,表达CAAR的细胞可用于治疗与自身抗原表达相关的自身免疫性疾病,比如自身免疫性疾病。在一些实施方案中,表达CAAR的细胞可靶向B细胞,所述B细胞最终产生自身抗体并在其细胞表面上展示自身抗体,将这些B细胞标记为用于治疗干预的疾病特异性靶标。在一些实施方案中,通过使用抗原特异性嵌合自身抗体受体靶向引起疾病的B细胞,表达CAAR的细胞可用于有效靶向和杀伤自身免疫性疾病中的致病性B细胞。在一些实施方案中,重组受体是CAAR,比如描述于美国专利申请公开号US 2017/0051035中的任一种。

[0495] 在一些实施方案中,CAAR包含自身抗体结合结构域、跨膜结构域和细胞内信号传导区。在一些实施方案中,细胞内信号传导区包含细胞内信号传导结构域。在一些实施方案中,细胞内信号传导结构域是或包含初级信号传导结构域、能够在T细胞中诱导初级活化信号的信号传导结构域、T细胞受体(TCR)组分的信号传导结构域、和/或包含免疫受体酪氨酸

活化基序 (ITAM) 的信号传导结构域。在一些实施方案中,细胞内信号传导区包含次级或共刺激信号传导区(次级细胞内信号传导区)。

[0496] 在一些实施方案中,自身抗体结合结构域包含自身抗原或其片段。自身抗原的选择可取决于被靶向的自身抗体的类型。例如,可以选择自身抗原,因为它识别与特定疾病状态例如自身免疫性疾病(比如自身抗体介导的自身免疫性疾病)相关的靶细胞(例如B细胞)上的自身抗体。在一些实施方案中,自身免疫性疾病包括寻常型天疱疮(PV)。示例性自身抗原包括桥粒芯糖蛋白1(Dsg1)和Dsg3。

多靶向

[0497] 在一些实施方案中,细胞和方法包括多靶向策略,比如在细胞上表达两种或更多种基因工程受体,每种受体识别相同或不同抗原,并且通常各自包括不同的细胞内信号传导组分。这样的多靶向策略描述于例如国际专利申请公开号WO 2014055668A1(描述活化性和共刺激CAR的组合,例如,靶向在脱靶例如正常细胞上单独存在但仅一起存在于待治疗的疾病或病症的细胞上的两种不同的抗原)和Fedorov等人,Sci.Transl.Medicine,5(215)(2013年12月)(描述表达活化性和抑制性CAR的细胞,比如其中活化性CAR与在正常或非患病细胞和待治疗的疾病或病症的细胞上表达的一种抗原结合,并且抑制性CAR与仅在正常细胞或不希望被治疗的细胞上表达的另一抗原结合)。

[0498] 例如,在一些实施方案中,细胞包括表达第一基因工程抗原受体(例如,CAR或TCR)的受体,其能够通常在与被所述第一受体识别的抗原(例如第一抗原)特异性地结合后诱导针对细胞的活化或刺激信号。在一些实施方案中,细胞进一步包括第二基因工程抗原受体(例如,CAR或TCR),例如嵌合共刺激受体,其能够通常在与被所述第二受体识别的第二抗原特异性地结合后诱导针对免疫细胞的共刺激信号。在一些实施方案中,第一抗原和第二抗原是相同的。在一些实施方案中,第一抗原和第二抗原是不同的。

[0499] 在一些实施方案中,第一和/或第二基因工程抗原受体(例如CAR或TCR)能够诱导针对细胞的活化或刺激信号。在一些实施方案中,受体包括含有ITAM或ITAM样基序的细胞内信号传导组分。在一些实施方案中,由第一受体诱导的活化涉及细胞中导致免疫应答启动的信号转导或蛋白质表达的变化,比如ITAM磷酸化和/或ITAM介导的信号转导级联的启动、在结合受体附近的免疫突触和/或分子(例如CD4或CD8等)簇的形成、一种或多种转录因子(例如NF- κ B和/或AP-1)的活化和/或诸如细胞因子等因子的基因表达的诱导、增殖和/或存活。

[0500] 在一些实施方案中,第一和/或第二受体包括共刺激受体的细胞内信号传导结构域,所述共刺激受体比如CD28、CD137(4-1BB)、OX40和/或ICOS。在一些实施方案中,第一和第二受体包括不同的共刺激受体的细胞内信号传导结构域。在一个实施方案中,第一受体含有CD28共刺激信号传导区,并且第二受体含有4-1BB共刺激信号传导区,反之亦然。

[0501] 在一些实施方案中,第一和/或第二受体包括含有ITAM或ITAM样基序的细胞内信号传导结构域和共刺激受体的细胞内信号传导结构域两者。

[0502] 在一些实施方案中,第一受体含有包含ITAM或ITAM样基序的细胞内信号传导结构域,并且第二受体含有共刺激受体的细胞内信号传导结构域。与同一细胞中诱导的活化或刺激信号组合的共刺激信号是导致免疫应答的信号,所述免疫应答比如稳健和持续的免疫应答,比如基因表达增加、细胞因子和其他因子的分泌、以及T细胞介导的效应子功能(比如

细胞杀伤)。

[0503] 在一些实施方案中,单独的第一受体连接和单独的第二受体连接都没有诱导稳健的免疫应答。在一些方面,如果仅一种受体被连接,则细胞变得对抗原耐受或无反应或被抑制和/或不被诱导增殖或分泌因子或执行效应子功能。然而,在一些这样的实施方案中,当多种受体被连接时,例如在遇到表达第一和第二抗原的细胞后,实现了期望的反应,比如完全免疫激活或刺激,例如指示为一种或多种细胞因子的分泌、增殖、持久性和/或执行免疫效应子功能(比如靶细胞的细胞毒性杀伤)。

[0504] 在一些实施方案中,所述两种受体分别诱导针对细胞的活化性信号和抑制性信号,使得受体之一与其抗原的结合激活细胞或诱导反应,而第二抑制性受体与其抗原的结合诱导遏制或抑制该反应的信号。实例为活化性CAR和抑制性CAR或iCAR的组合。例如,可以使用这样的策略,其中活化性CAR结合在疾病或病症中表达但在正常细胞上也表达的抗原,并且抑制性受体结合在正常细胞上表达但在疾病或病症的细胞上不表达的单独抗原。

[0505] 在一些实施方案中,在这样的情况下采用多靶向策略:与特定疾病或病症相关的抗原短暂地(例如,在与基因工程相关的刺激后)或持久地在非患病细胞上表达和/或在工程化细胞自身上表达。在这样的情况下,通过要求连接两种独立的和单独的特异性抗原受体,可以改善特异性、选择性和/或有效性。

[0506] 在一些实施方案中,多种抗原(例如第一和第二抗原)在被靶向的细胞、组织或疾病或病症上比如在癌细胞上表达。在一些方面,所述细胞、组织、疾病或病症是多发性骨髓瘤或多发性骨髓瘤细胞。在一些实施方案中,多种抗原中的一种或多种通常也在不希望用细胞疗法靶向的细胞比如正常或非患病细胞或组织和/或工程化细胞自身上表达。在这样的实施方案中,通过要求连接多种受体以实现细胞的反应,实现了特异性和/或有效性。

组合物和制剂

[0507] 在一些实施方案中,通过本文所述的方法提供工程化细胞,所述方法包括表观遗传学和/或表观基因组学分析。在一些实施方案中,将细胞例如用重组受体加以基因工程改造的细胞(例如CAR-T细胞)提供为组合物,包括药物组合物和制剂,比如包含以给定剂量或其部分施用的细胞数的单位剂型组合物。药物组合物和制剂通常包含一种或多种任选的药学上可接受的载体或赋形剂。在一些实施方案中,组合物包含至少一种另外的治疗剂。

[0508] 术语“药物制剂”是指其为这样的形式的制剂,所述形式允许其中所含的活性成分的生物活性是有效的,并且所述制剂不含另外的对该制剂所施用的受试者具有不可接受的毒性的组分。

[0509] “药学上可接受的载体”是指药物制剂中除了活性成分以外的对受试者无毒的成分。药学上可接受的载体包括但不限于缓冲剂、赋形剂、稳定剂或防腐剂。

[0510] 在一些方面,载体的选择部分地由特定的细胞和/或由施用方法决定。因此,有多种适合的制剂。例如,药物组合物可含有防腐剂。适合的防腐剂可包括,例如,对羟基苯甲酸甲酯、对羟基苯甲酸丙酯、苯甲酸钠和苯扎氯铵。在一些方面,使用两种或更多种防腐剂的混合物。防腐剂或其混合物通常以总组合物重量的约0.0001%至约2%的量存在。载体描述于例如Remington's Pharmaceutical Sciences第16版,0sol,A.Ed.(1980)。药学上可接受的载体通常在采用的剂量和浓度下对于接受者是无毒的,并且包括但不限于:诸如磷酸盐、柠檬酸盐、和其他有机酸的缓冲剂;包括抗坏血酸和甲硫氨酸在内的抗氧化剂;防腐剂(如

十八烷基二甲基苄基氯化铵；氯化六甲双铵；苯扎氯铵；苄索氯铵；苯酚、丁醇或苄醇；烷基对羟基苯甲酸酯，如对羟基苯甲酸甲酯或对羟基苯甲酸丙酯；儿茶酚；间苯二酚；环己醇；3-戊醇；和间甲酚）；低分子量（少于约10个残基）多肽；蛋白质，如血清白蛋白、明胶、或免疫球蛋白；亲水性聚合物，如聚乙烯吡咯烷酮；氨基酸，如甘氨酸、谷氨酰胺、天冬酰胺、组氨酸、精氨酸、或赖氨酸；单糖类、二糖类、和其他碳水化合物类，包括葡萄糖、甘露糖、或糊精；螯合剂，如EDTA；糖类，如蔗糖、甘露醇、海藻糖或山梨糖醇；成盐反离子，如钠；金属复合物（例如，Zn-蛋白质复合物）；和/或非离子型表面活性剂，如聚乙二醇（PEG）。

[0511] 在一些方面，缓冲剂被包括在组合物中。适合的缓冲剂包括，例如，柠檬酸、柠檬酸钠、磷酸、磷酸钾和各种其他酸和盐。在一些方面，使用两种或更多种缓冲剂的混合物。缓冲剂或其混合物通常以总组合物重量的约0.001%至约4%的量存在。用于制备可施用的药物组合物的方法是已知的。示例性方法更详细地描述于例如Remington: The Science and Practice of Pharmacy, Lippincott Williams&Wilkins; 第21版(2005年5月1日)。

[0512] 制剂可包括水溶液。制剂或组合物还可含有多于一种对细胞治疗的特定适应症、疾病或病症有用的活性成分，优选那些具有与细胞互补的活性的活性成分，其中各自的活性彼此无不良影响。这样的活性成分以对预期目的有效的量适当地组合存在。因此，在一些实施方案中，药物组合物进一步包含其他药物活性剂或药物，比如化疗剂，例如天冬酰胺酶、白消安、卡铂、顺铂、柔红霉素、多柔比星、氟尿嘧啶、吉西他滨、羟基脲、甲氨蝶呤、紫杉醇、利妥昔单抗、长春碱和/或长春新碱。

[0513] 在一些实施方案中，药物组合物含有有效治疗或预防疾病或病症的量（比如治疗有效量或预防有效量）的细胞。一些实施方案中，通过定期评估被治疗的受试者来监测治疗或预防功效。可通过单次推注施用细胞、通过多次推注施用细胞、或通过连续输注施用细胞来递送期望的剂量。

[0514] 可以使用标准施用技术、制剂和/或装置施用细胞和组合物。细胞的施用可以是自体的或异源的。例如，免疫应答细胞或祖细胞可以从一位受试者获得，并施用于同一位受试者或不同的相容性受试者。可以通过局部注射（包括导管给药）、全身注射、局部注射、静脉注射或肠胃外给药来施用外周血衍生（例如，体内、离体或体外衍生）的免疫应答细胞或其后代。当施用治疗组合物（例如，含有遗传修饰的免疫应答细胞的药物组合物）时，通常将其配制成单位剂量可注射形式（溶液、悬浮液、乳液）。

[0515] 制剂包括经口服、静脉内、腹膜内、皮下、肺、透皮、肌内、鼻内、口腔、舌下或栓剂给药的制剂。在一些实施方案中，经肠胃外施用细胞群。如本文中使用的术语“肠胃外”包括静脉内、肌内、皮下、直肠、阴道和腹膜内给药。在一些实施方案中，使用通过静脉内、腹膜内或皮下注射的外周全身递送将细胞施用于受试者。

[0516] 在一些实施方案中，将组合物提供为无菌液体制剂，例如等渗水溶液、悬浮液、乳液、分散体或粘性组合物，其在一些方面可以缓冲至选定的pH。液体制剂通常比凝胶、其他粘性组合物和固体组合物更容易制备。另外，液体组合物稍微更便于施用，特别是通过注射。另一方面，可以将粘性组合物配制在适当的粘度范围内，以提供更长的与特异性组织接触的时间。液体或粘性组合物可包含载体，其可以是溶剂或分散介质，所述溶剂或分散介质含有例如水、盐水、磷酸盐缓冲盐水、多元醇（例如甘油、丙二醇、液体聚乙二醇）及其适合的混合物。

[0517] 可以通过将细胞掺入溶剂中来制备无菌注射溶液,所述溶剂比如具有适合的载体、稀释剂或赋形剂(比如无菌水、生理盐水、葡萄糖、右旋糖等)的混合物。组合物可含有辅助物质,比如润湿剂、分散剂或乳化剂(例如甲基纤维素)、pH缓冲剂、胶凝或粘度增强添加剂、防腐剂、调味剂和/或着色剂,这取决于给药途径和期望的制剂。在一些方面可以查阅标准文本来制备适合的制剂。

[0518] 可以加入各种增强组合物稳定性和无菌性的添加剂,包括抗微生物防腐剂、抗氧化剂、螯合剂和缓冲剂。可以通过各种抗细菌剂和抗真菌剂例如对羟基苯甲酸、三氯叔丁醇、苯酚和抗坏血酸确保预防微生物作用。可通过使用吸收延迟剂(例如单硬脂酸铝和明胶)来实现注射药物形式的延长吸收。

[0519] 用于体内给药的制剂通常是无菌的。可以例如通过无菌过滤膜进行过滤而容易地实现无菌性。

给药方法

[0520] 在一些实施方案中,所提供的方法通常涉及对患有疾病或病症的受试者施用表达重组分子的细胞的剂量,所述重组分子比如重组受体,比如CAR、其他嵌合受体或其他抗原受体(如转基因TCR),所述疾病或病症比如其组分被所述重组分子例如受体特异性地识别的疾病或病症和/或通过所述重组分子进行治疗的疾病或病症。在一些实施方案中,使用所提供的用于评估表观遗传学和/或表观基因组学的方法来分析细胞。在一些方面,可以在基因工程之前和/或之后对细胞进行表观遗传学和/或表观基因组学分析。给药通常实现疾病或病症的一种或多种症状的改善和/或治疗或预防疾病或病症或其症状。

[0521] 所治疗的疾病或病症可以是其中抗原表达与疾病状况或疾患的病因学相关和/或参与疾病状况或疾患的病因学(例如引起、加剧或以其他方式参与这种疾病、病症或疾患)的任何疾病或病症。示例性疾病和病症可包括与恶性肿瘤或细胞转化(例如癌症)、自身免疫性疾病或炎性疾病或传染性疾病(由细菌、病毒或其他病原体引起)相关的疾病或病症。本文中描述了示例性抗原,其包括与可以治疗的各种疾病和病症相关的抗原。在具体的实施方案中,嵌合抗原受体或转基因TCR特异性地结合与疾病或病症相关的抗原。

[0522] 其中,疾病、病症和疾患是肿瘤,包括实体瘤、恶性血液肿瘤和黑色素瘤,并且包括局部和转移性肿瘤、感染性疾病,比如感染病毒或其他病原体例如HIV、HCV、HBV、CMV、HPV和寄生虫病,以及自身免疫性和炎性疾病。在一些实施方案中,疾病、疾患或病症是肿瘤、癌症、恶性肿瘤、新生物或其他增殖性疾病或疾患。这样的疾病包括但不限于白血病、淋巴瘤,例如慢性淋巴细胞白血病(CLL)、ALL、非霍奇金淋巴瘤、急性髓性白血病、多发性骨髓瘤、难治性滤泡性淋巴瘤、套细胞淋巴瘤、惰性B细胞淋巴瘤、B细胞恶性肿瘤、结肠癌、肺癌、肝癌、乳腺癌、前列腺癌、卵巢癌、皮肤癌、黑色素瘤、骨癌和脑癌、卵巢癌、上皮癌、肾细胞癌、胰腺癌、霍奇金淋巴瘤、宫颈癌、结直肠癌、胶质母细胞瘤、神经母细胞瘤、尤文肉瘤、髓母细胞瘤、骨肉瘤、滑膜肉瘤和/或间皮瘤。

[0523] 在一些实施方案中,所述疾病或病症是肿瘤,并且所述受试者在施用第一剂量之前具有大的肿瘤负荷,比如大的实体瘤或大的疾病相关(例如肿瘤)细胞数目或体积。在一些方面,受试者具有大量转移和/或广泛的转移灶定位。在一些方面,受试者中的肿瘤负荷低并且受试者具有很少的转移。在一些实施方案中,通过受试者中的初始疾病负荷确定剂量的大小或时间。例如,尽管在一些方面,可以在第一剂量中对受试者施用较低数目的细

胞,但在疾病负荷较低的情况下,剂量可以更高。

[0524] 在一些实施方案中,疾病或病症是感染性疾病或病症,例如但不限于病毒、逆转录病毒、细菌和原生动物感染,免疫缺陷,巨细胞病毒(CMV)、EB病毒(Epstein-Barr virus (EBV))、腺病毒、BK多瘤病毒感染。在一些实施方案中,疾病或病症是自身免疫性或炎性疾病或病症,比如关节炎,例如类风湿性关节炎(RA)、I型糖尿病、系统性红斑狼疮(SLE)、炎性肠病、银屑病、硬皮病、自身免疫性甲状腺疾病、格雷夫斯病、克罗恩病、多发性硬化、哮喘和/或与移植相关的疾病或病症。

[0525] 在一些实施方案中,与疾病或疾患相关的抗原是或包括选自下组的抗原,该组的组成为: α v β 6整合素(avb6整合素)、B细胞成熟抗原(BCMA)、B7-H3、B7-H6、碳酸酐酶9(CA9,也称为CAIX或G250)、癌-睾丸抗原、癌/睾丸抗原1B(CTAG,也称为NY-ESO-1和LAGE-2)、癌胚抗原(CEA)、细胞周期蛋白、细胞周期蛋白A2、C-C基序趋化因子配体1(CCL-1)、CD19、CD20、CD22、CD23、CD24、CD30、CD33、CD38、CD44、CD44v6、CD44v7/8、CD123、CD138、CD171、表皮生长因子蛋白(EGFR)、截短的表皮生长因子蛋白(tEGFR)、III型表皮生长因子受体突变(EGFR vIII)、上皮糖蛋白2(EPG-2)、上皮糖蛋白40(EPG-40)、肝配蛋白B2、肝配蛋白受体(ephrine)A2(EPHa2)、雌激素受体、Fc受体样5(FCRL5;也称为Fc受体同源物5或FCRH5)、胎儿乙酰胆碱受体(胎儿AChR)、叶酸结合蛋白(FBP)、叶酸受体 α 、胎儿乙酰胆碱受体、神经节苷脂GD2、O-乙酰化GD2(OGD2)、神经节苷脂GD3、糖蛋白100(gp100)、G蛋白偶联受体5D(GPCR5D)、Her2/neu(受体酪氨酸激酶erbB2)、Her3(erb-B3)、Her4(erb-B4)、erbB二聚体、人类高分子量黑色素瘤相关抗原(HMW-MAA)、乙型肝炎表面抗原、人白细胞抗原A1(HLA-A1)、人白细胞抗原A2(HLA-A2)、IL-22受体 α (IL-22Ra)、IL-13受体 α 2(IL-13Ra2)、激酶插入结构域受体(kdr)、 κ 轻链、L1细胞粘附分子(L1CAM)、L1-CAM的CE7表位、富含亮氨酸重复序列蛋白家族成员8A(LRRC8A)、Lewis Y、黑色素瘤相关抗原(MAGE)-A1、MAGE-A3、MAGE-A6、间皮素、c-Met、鼠巨细胞病毒(CMV)、粘蛋白1(MUC1)、MUC16、自然杀伤组2成员D(NKG2D)配体、黑色素A(MART-1)、神经细胞粘附分子(NCAM)、癌胚胎抗原、黑色素瘤优先表达抗原(PRAME)、孕酮受体、前列腺特异性抗原、前列腺干细胞抗原(PSCA)、前列腺特异性膜抗原(PSMA)、受体酪氨酸激酶样孤儿受体1(ROR1)、生存素、滋养层细胞糖蛋白(TPBG也称为5T4)、肿瘤相关糖蛋白72(TAG72)、血管内皮生长因子受体(VEGFR)、血管内皮生长因子受体2(VEGFR2)、维尔姆斯瘤1(WT-1)、病原体特异性或病原体表达的抗原、或通用标签相关抗原、和/或生物素化分子,和/或由HIV、HCV、HBV或其他病原体表达的分子。

[0526] 如本文中使用的,“治疗”(“treatment”)及其语法变体,比如“治疗”(“treat”)或“治疗”(“treating”)是指疾病或病症或疾患或与之相关的症状、不良反应或结果或表型的完全或部分改善或减轻。期望的治疗效果包括但不限于预防疾病的发生或复发、症状减轻、疾病的任何直接或间接病理结果的削减、预防转移、降低疾病进展速度、改善或减轻疾病状态、以及缓解或改善预后。所述术语不一定意味着完全治愈疾病或完全消除任何症状或对所有症状或结果的影响。

[0527] 如本文中使用的,“延迟疾病的发展”意味着推迟、阻碍、减慢、延缓、稳定、抑制和/或推延疾病(比如癌症)的发展。这种延迟可以具有不同的时间长度,取决于病史和/或被治疗的个体。对于本领域技术人员显而易见的是,足够或显著的延迟实际上可以包括预防,因为个体不发生所述疾病。例如,可以延迟晚期癌症比如转移的发展。

[0528] 如本文中使用的,“预防”包括提供关于受试者中疾病的发生或复发的预防,所述受试者可能易患所述疾病但尚未被诊断患有所述疾病。在一些实施方案中,提供的细胞和组合物用于延迟疾病的发展或减慢疾病的进展。

[0529] 如本文中使用的,“抑制”功能或活性是当与除了感兴趣的条件或参数之外的其他相同的条件相比时或者与另一条件相比时降低了所述功能或活性。例如,抑制肿瘤生长的细胞与不存在所述细胞的肿瘤生长速率相比降低了肿瘤的生长速率。

[0530] 在给药的情况下,药剂(例如药物制剂、细胞或组合物)的“有效量”是指在剂量/量和时间段方面有效实现期望的结果(比如为治疗或预防结果)所需的量。

[0531] 药剂(例如药物制剂或细胞)的“治疗有效量”是指在剂量和时间段方面有效实现期望的治疗结果(比如用于治疗疾病、病症或疾患)和/或治疗的药代动力学或药效学作用所需的量。治疗有效量可根据诸如受试者的疾病状态、年龄、性别和体重等因素以及施用的细胞群而变化。在一些实施方案中,所提供的方法包括以有效量(例如治疗有效量)施用细胞和/或组合物。

[0532] “预防有效量”是指在剂量和时间段方面有效实现期望的预防结果所需的量。通常(但不一定),由于预防剂量是在疾病之前或早期的受试者中使用,因此预防有效量将小于治疗有效量。在肿瘤负荷较低的情况下,在一些方面的预防有效量将高于治疗有效量。

[0533] 用于过继细胞疗法的细胞的施用方法是已知的,并且可以与所提供的方法和组合物结合使用。例如,过继T细胞治疗方法描述于例如Gruenberg等人的美国专利申请公开号2003/0170238;Rosenberg的美国专利No.4,690,915;Rosenberg(2011)Nat Rev Clin Oncol.8(10):577-85)。参见,例如Themeli等人(2013)Nat Biotechnol.31(10):928-933;Tsukahara等人(2013)Biochem Biophys Res Commun 438(1):84-9;Davila等人(2013)PLoS ONE8(4):e61338。

[0534] 在一些实施方案中,通过自体转移进行细胞疗法,例如过继细胞疗法,例如过继T细胞毒疗法,其中细胞是从将要接受所述细胞疗法的受试者中分离和/或以其他方式制备的,或来自这样的受试者来源的样品。因此,在一些方面,细胞来源于需要治疗的受试者,例如患者,并且在分离和处理后将细胞施用于同一位受试者。

[0535] 在一些实施方案中,通过同种异体转移进行细胞疗法,例如过继细胞疗法,例如过继T细胞疗法,其中细胞是从将要接受或最终接受所述细胞疗法的受试者以外的受试者(例如,第一受试者)分离和/或以其他方式制备的。在这样的实施方案中,然后将细胞施用于相同物种的不同受试者,例如第二受试者。在一些实施方案中,第一和第二受试者在遗传上是相同的。在一些实施方案中,第一和第二受试者在遗传上是相似的。在一些实施方案中,第二受试者表达与第一受试者相同的HLA类或超类型。

[0536] 所述细胞可以通过任何适合的手段施用,例如通过推注输注,通过注射,例如静脉内或皮下注射、眼内注射、眼周注射、视网膜下注射、玻璃体内注射、经中隔注射、巩膜下注射、脉络膜内注射、前房注射、结膜下(subconjunctival)注射、结膜下注射、眼球筋膜下注射、球后注射、球周注射或后部近巩膜递送。在一些实施方案中,将它们通过肠胃外、肺内和鼻内施用,并且如果需要局部治疗,则通过病灶内施用。肠胃外输注包括肌内、静脉内、动脉内、腹膜内、胸腔内、颅内或皮下施用。在一些实施方案中,通过单次推注施用细胞来施用给定剂量。在一些实施方案中,其施用为通过例如在不超过3天的时间内多次推注施用细胞,

或通过连续输注施用细胞。

[0537] 为了预防或治疗疾病,适当的剂量可取决于有待治疗的疾病的类型、细胞或重组受体的类型、疾病的严重性和病程、所述细胞是为预防目的还是治疗目的而施用、先前的治疗、受试者的临床病史和对细胞的反应、以及主治医师的判定。在一些实施方案中将组合物和细胞一次性地或以一系列治疗的方式适当地施用于受试者。

[0538] 在一些实施方案中,将细胞作为联合治疗的一部分施用,比如与另一治疗干预同时或以任何顺序依次施用,所述另一治疗干预比如抗体或工程化细胞或受体或其他药剂,比如细胞毒性剂或治疗剂。因此,在一些实施方案中,同时或以任何顺序依次将细胞与一种或多种另外的治疗剂共同施用或结合另一治疗干预共同施用。在一些情况下,将细胞与另一种疗法在时间上足够接近地共同施用,使得细胞群增强一种或多种另外的治疗剂的作用,反之亦然。在一些实施方案中,在一种或多种另外的治疗剂之前施用细胞。在一些实施方案中,在一种或多种另外的治疗剂之后施用细胞。在一些实施方案中,一种或多种另外的药剂包括细胞因子,比如IL-2或其他细胞因子,以便例如增强持久性。在一些实施方案中,所述方法包括施用化学治疗剂,例如调节性化学治疗剂,例如用于在剂量施用之前减轻肿瘤负荷。

[0539] 用免疫耗竭(例如,淋巴细胞清除)疗法预处理受试者在一些方面可以改善过继细胞疗法(ACT)的效果。

[0540] 因此,在一些实施方案中,所述方法包括在细胞疗法开始之前对受试者施用预处理剂,比如淋巴细胞清除剂或化学治疗剂,比如环磷酰胺、氟达拉滨或其组合。例如,可以在细胞疗法开始之前至少2天(比如之前至少3天、4天、5天、6天或7天)对受试者施用预处理剂。在一些实施方案中,在细胞疗法开始之前不超过7天(比如之前不超过6天、5天、4天、3天或2天)对受试者施用预处理剂。

[0541] 在一些实施方案中,将受试者用(约)20mg/kg到100mg/kg之间,比如(约)40mg/kg到80mg/kg之间的剂量的环磷酰胺预处理。在一些方面,将受试者用(约)60mg/kg的环磷酰胺预处理。在一些实施方案中,环磷酰胺可以单剂量形式施用或者可以多个剂量的形式施用,比如每日、隔日或每三日给予。在一些实施方案中,施用环磷酰胺每日一次,持续一天或两天。在一些实施方案中,当淋巴细胞清除剂包含环磷酰胺时,对受试者施用(约)100mg/m²到500mg/m²之间,比如(约)200mg/m²到400mg/m²或250mg/m²到350mg/m²之间的剂量的环磷酰胺,包括各端值。在一些情况下,对受试者施用约300mg/m²的环磷酰胺。在一些实施方案中,环磷酰胺可以单剂量形式施用或者可以多个剂量的形式施用,比如每日、隔日或每三日给予。在一些实施方案中,每日施用环磷酰胺,比如持续如1-5天,例如3至5天。在一些情况下,在细胞疗法开始之前,对受试者每日施用约300mg/m²的环磷酰胺,持续3天。

[0542] 在一些实施方案中,当淋巴细胞清除剂包含氟达拉滨时,对受试者施用(约)1mg/m²到100mg/m²之间,比如(约)10mg/m²到75mg/m²、15mg/m²到50mg/m²、20mg/m²到40mg/m²或24mg/m²到35mg/m²之间的剂量的氟达拉滨,包括各端值。在一些情况下,对受试者施用约30mg/m²的氟达拉滨。在一些实施方案中,氟达拉滨可以单剂量形式施用或者可以多个剂量的形式施用,比如每日、隔日或每三日给予。在一些实施方案中,每日施用氟达拉滨,比如持续如1-5天,例如3至5天。在一些情况下,在细胞疗法开始之前,对受试者每日施用约30mg/m²的氟达拉滨,持续3天。

[0543] 在一些实施方案中,淋巴细胞清除剂包括药剂的联合,比如环磷酰胺和氟达拉滨的联合。因此,药剂的联合可包括任何剂量或给药方案(比如本文所述的那些)的环磷酰胺,以及任何剂量或给药方案(比如本文所述的那些)的氟达拉滨。例如,在一些方面,在第一或后续剂量之前,对受试者施用60mg/kg(约2g/m²)的环磷酰胺和3至5个剂量的25mg/m²的氟达拉滨。

[0544] 一旦将细胞施用于受试者(例如人),在一些方面通过许多已知方法中的任何一种来测量工程化细胞群的生物活性。要评估的参数包括工程化或天然T细胞或其他免疫细胞与抗原的特异性结合,例如采用体内成像,或例如采用离体ELISA或流式细胞术。在某些实施方案中,可以使用本领域已知的任何适合的方法测量工程化细胞破坏靶细胞的能力,例如在Kochenderfer等人,J.Immunotherapy,32(7):689-702(2009)和Herman等人J.Immunological Methods,285(1):25-40(2004)中描述的细胞毒性测定。在某些实施方案中,还可以通过测定某些细胞因子(比如CD107a、IFN γ 、IL-2和TNF)的表达和/或分泌来测量细胞的生物活性。在一些方面,通过评估临床结果(比如肿瘤负荷或负载的降低)来测量生物活性。在一些方面,评估了毒性结果、细胞的持久性和/或扩增、和/或宿主免疫应答的存在或不存在。

[0545] 在某些实施方案中,以任何数量的方式修饰工程化细胞,使得其治疗或预防功效增加。例如,由细胞群表达的工程化CAR或TCR可以直接地或通过接头间接地与靶向模块缀合。将化合物(例如CAR或TCR)与靶向模块缀合的实践是本领域已知的。参见,例如,Wadwa等人,J.Drug Targeting3:1 1 1(1995)和美国专利5,087,616。

[0546] 在一些实施方案中,将细胞作为联合治疗的一部分施用,比如与另一治疗干预同时或以任何顺序依次施用,所述另一治疗干预比如抗体或工程化细胞或受体或药剂,比如细胞毒性剂或治疗剂。在一些实施方案中,同时或以任何顺序依次将细胞与一种或多种另外的治疗剂共同施用或结合另一治疗干预共同施用。在一些情况下,将细胞与另一种疗法在时间上足够接近地共同施用,使得细胞群增强一种或多种另外的治疗剂的作用,反之亦然。在一些实施方案中,在一种或多种另外的治疗剂之前施用细胞。在一些实施方案中,在一种或多种另外的治疗剂之后施用细胞。在一些实施方案中,一种或多种另外的药剂包括细胞因子,比如IL-2,以便例如增强持久性。

给药

[0547] 在一些实施方案中,根据提供的方法和/或利用所提供的制品或组合物将一定剂量的细胞施用于受试者。在一些实施方案中,根据受试者中的特定疾病或病症确定剂量的大小或时间。在一些情况下,鉴于所提供的描述,可以凭经验确定针对特定疾病的剂量的大小或时间。

[0548] 在一些实施方案中,细胞剂量包含(约)2x 10⁵个细胞/kg到(约)2x10⁶个细胞/kg之间,比如(约)4x 10⁵个细胞/kg到(约)1x 10⁶个细胞/kg之间,或(约)6x 10⁵个细胞/kg到(约)8x 10⁵个细胞/kg之间。在一些实施方案中,细胞剂量包含不超过2x 10⁵个细胞(例如表达抗原的细胞,比如表达CAR的细胞)每千克受试者体重(细胞/kg),比如不超过(约)3x 10⁵个细胞/kg,不超过(约)4x 10⁵个细胞/kg,不超过(约)5x 10⁵个细胞/kg,不超过(约)6x 10⁵个细胞/kg,不超过(约)7x 10⁵个细胞/kg,不超过(约)8x 10⁵个细胞/kg,不超过(约)9x 10⁵个细胞/kg,不超过(约)1x 10⁶个细胞/kg,或不超过(约)2x 10⁶个细胞/kg。在一些实施方案

中,细胞剂量包含至少或至少约或者为或约为 2×10^5 个细胞(例如表达抗原的细胞,例如表达CAR的细胞)每千克受试者体重(细胞/kg),例如至少或至少约或者为或约为 3×10^5 个细胞/kg,至少或至少约或者为或约为 4×10^5 个细胞/kg,至少或至少约或者为或约为 5×10^5 个细胞/kg,至少或至少约或者为或约为 6×10^5 个细胞/kg,至少或至少约或者为或约为 7×10^5 个细胞/kg,至少或至少约或者为或约为 8×10^5 个细胞/kg,至少或至少约或者为或约为 9×10^5 个细胞/kg,至少或至少约或者为或约为 1×10^6 个细胞/kg,或至少或至少约或者为或约为 2×10^6 个细胞/kg。

[0549] 在过继细胞疗法的背景下,施用给定“剂量”包括以单一组合物和/或单次不间断给药的形式例如以单次注射或连续输注的形式施用给定量或数目的细胞,并且还指在指定的时间段内(不超过3天)以分次剂量的形式施用提供在多个单独的组合物或输液中的给定量或数目的细胞。因此,在一些情况下,剂量是在单个时间点给予或开始的指定数目的细胞的单次或连续施用。然而,在一些情况下,在不超过三天的时间内以多次注射或输注的形式施用所述剂量,比如每天一次,持续三天或两天,或在单时段内多次输注。

[0550] 因此,在一些方面,以单一药物组合物的形式施用细胞。

[0551] 在一些实施方案中,以多个组合物的形式施用细胞,所述组合物共同含有单剂量的细胞。

[0552] 因此,在一些方面可以分次剂量的形式施用一个或多个剂量。例如,在一些实施方案中,可以将剂量经过2天或3天施用于受试者。用于分次给药的示例性方法包括在第一天施用25%的剂量并且在第二天施用剩余的75%的剂量。在其他实施方案中,可以在第一天施用33%的剂量,并且在第二天施用剩余的67%。在一些方面,在第一天施用10%的剂量,在第二天施用30%的剂量,并且在第三天施用60%的剂量。在一些实施方案中,分次剂量的分布不超过3天。

[0553] 在一些实施方案中,在一些方面使用与关于第一剂量和第二剂量之间的时间的相同的时间指南,例如通过施用第一和多个后续剂量给予多个剂量,其中在施用第一或先前剂量之后大于约28天的时间点给予每个后续剂量。

[0010] 在某些实施方案中,细胞或细胞亚型的单独群体以范围约1百万个至约1000亿个细胞和/或每千克体重的该细胞量施用于受试者,例如像,100万个至约500亿个细胞(例如,约500万个细胞、约2500万个细胞、约5亿个细胞、约10亿个细胞、约50亿个细胞、约200亿个细胞、约300亿个细胞、约400亿个细胞,或由前述任何两个值定义的范围),比如约1000万个至约1000亿个细胞(例如,约2000万个细胞、约3000万个细胞、约4000万个细胞、约6000万个细胞、约7000万个细胞、约8000万个细胞、约9000万个细胞、约100亿个细胞、约250亿个细胞、约500亿个细胞、约750亿个细胞、约900亿个细胞,或由前述任何两个值定义的范围),并且在一些情况下约1亿个细胞至约500亿个细胞(例如,约1.2亿个细胞、约2.5亿个细胞、约3.5亿个细胞、约4.5亿个细胞、约6.5亿个细胞、约8亿个细胞、约9亿个细胞、约30亿个细胞、约300亿个细胞、约450亿个细胞)或这些范围之间的任何值和/或每千克体重的所述值。剂量可以根据疾病或疾患和/或患者和/或其他治疗特有的属性而变化。

[0554] 在一些实施方案中,基因工程细胞的剂量包括(约) 1×10^5 至 5×10^8 个表达CAR的总T细胞、 1×10^5 至 2.5×10^8 个表达CAR的总T细胞、 1×10^5 至 1×10^8 个表达CAR的总T细胞、 1×10^5 至 5×10^7 个表达CAR的总T细胞、 1×10^5 至 2.5×10^7 个表达CAR的总T细胞、 1×10^5 至 1×10^7

个表达CAR的总T细胞、 1×10^5 至 5×10^6 个表达CAR的总T细胞、 1×10^5 至 2.5×10^6 个表达CAR的总T细胞、 1×10^5 至 1×10^6 个表达CAR的总T细胞、 1×10^6 至 5×10^8 个表达CAR的总T细胞、 1×10^6 至 2.5×10^8 个表达CAR的总T细胞、 1×10^6 至 1×10^8 个表达CAR的总T细胞、 1×10^6 至 5×10^7 个表达CAR的总T细胞、 1×10^6 至 2.5×10^7 个表达CAR的T细胞、 1×10^6 至 1×10^7 个表达CAR的总T细胞、 1×10^6 至 5×10^6 个表达CAR的总T细胞、 1×10^6 至 2.5×10^6 个表达CAR的总T细胞、 2.5×10^6 至 5×10^8 个表达CAR的总T细胞、 2.5×10^6 个至 2.5×10^8 个表达CAR的总T细胞、 2.5×10^6 至 1×10^8 个表达CAR的总T细胞、 2.5×10^6 至 5×10^7 个表达CAR的总T细胞、 2.5×10^6 到 2.5×10^7 个表达CAR的总T细胞、 2.5×10^6 至 1×10^7 个表达CAR的总T细胞、 2.5×10^6 至 5×10^6 个表达CAR的总T细胞、 5×10^6 至 5×10^8 个表达CAR的总T细胞、 5×10^6 至 2.5×10^8 个表达CAR的总T细胞、 5×10^6 至 1×10^8 个表达CAR的总T细胞、 5×10^6 至 5×10^7 个表达CAR的总T细胞、 5×10^6 至 2.5×10^7 个表达CAR的总T细胞、 5×10^6 至 1×10^7 个表达CAR的总T细胞、 1×10^7 至 5×10^8 个表达CAR的总T细胞、 1×10^7 至 2.5×10^8 个表达CAR的总T细胞、 1×10^7 至 1×10^8 个表达CAR的总T细胞、 1×10^7 至 5×10^7 个表达CAR的总T细胞、 1×10^7 至 2.5×10^7 个表达CAR的总T细胞、 2.5×10^7 至 5×10^8 个表达CAR的总T细胞、 2.5×10^7 至 2.5×10^8 个表达CAR的总T细胞、 2.5×10^7 个至 1×10^8 个表达CAR的总T细胞、 2.5×10^7 至 5×10^7 个表达CAR的总T细胞、 5×10^7 至 5×10^8 个表达CAR的总T细胞、 5×10^7 至 2.5×10^8 个表达CAR的总T细胞、 5×10^7 至 1×10^8 个表达CAR的总T细胞、 1×10^8 至 5×10^8 个表达CAR的总T细胞、 1×10^8 至 2.5×10^8 个表达CAR的总T细胞、或 2.5×10^8 至 5×10^8 个表达CAR的总T细胞。

[0555] 在一些实施方案中，基因工程细胞的剂量包括至少或至少约 1×10^5 个表达CAR的细胞、至少或至少约 2.5×10^5 个表达CAR的细胞、至少或至少约 5×10^5 个表达CAR的细胞、至少或至少约 1×10^6 个表达CAR的细胞、至少或至少约 2.5×10^6 个表达CAR的细胞、至少或至少约 5×10^6 个表达CAR的细胞、至少或至少约 1×10^7 个表达CAR的细胞、至少或至少约 2.5×10^7 个表达CAR的细胞、至少或至少约 5×10^7 个表达CAR的细胞、至少或至少约 1×10^8 个表达CAR的细胞、至少或至少约 2.5×10^8 个表达CAR的细胞、或至少或至少约 5×10^8 个表达CAR的细胞。

[0556] 在一些实施方案中，细胞疗法包括施用包含一定细胞数的剂量，所述细胞数为(约) 1×10^5 至 5×10^8 个表达重组受体的总细胞、总T细胞或总外周血单核细胞(PBMC)，(约) 5×10^5 至 1×10^7 个表达重组受体的总细胞、总T细胞或总外周血单核细胞(PBMC)，或(约) 1×10^6 至 1×10^7 个表达重组受体的总细胞、总T细胞或总外周血单核细胞(PBMC)，包括各端值。在一些实施方案中，细胞疗法包括施用包含一定细胞数的细胞剂量，所述细胞数为至少或至少约 1×10^5 个表达重组受体的总细胞、总T细胞或总外周血单核细胞(PBMC)，例如至少或至少约 1×10^6 、至少或至少约 1×10^7 、至少或至少约 1×10^8 个这样的细胞。在一些实施方案中，所述细胞数是关于CD3+或CD8+的总数，在一些情况下也是表达重组受体的(例如CAR+)细胞。在一些实施方案中，细胞疗法包括施用包含一定细胞数的剂量，所述细胞数为(约) 1×10^5 至 5×10^8 个CD3+或CD8+总T细胞或表达重组受体的CD3+或CD8+细胞，(约) 5×10^5 至 1×10^7 个CD3+或CD8+总T细胞或表达重组受体的CD3+或CD8+细胞，或(约) 1×10^6 至 1×10^7 个CD3+或CD8+总T细胞或表达重组受体的CD3+或CD8+细胞，包括各端值。在一些实施方案中，细胞疗法包括施用包含一定细胞数的剂量，所述细胞数为(约) 1×10^5 至 5×10^8 个总的CD3+/CAR+或CD8+/CAR+细胞，(约) 5×10^5 至 1×10^7 个总的CD3+/CAR+或CD8+/CAR+细胞，或(约) 1×10^6

至 1×10^7 个总的CD3+/CAR+或CD8+/CAR+细胞,包括各端值。

[0557] 在一些实施方案中,所述剂量的T细胞包括CD4+ T细胞、CD8+ T细胞或CD4+和CD8+ T细胞。

[0558] 在一些实施方案中,例如,当受试者是人时,CD8+ T细胞的剂量(包括在包含CD4+和CD8+ T细胞的剂量中)包括约 1×10^6 到 5×10^8 个总的表达重组受体(例如,CAR)的细胞,例如,范围为约 5×10^6 个至 1×10^8 个这种细胞,比如 1×10^7 、 2.5×10^7 、 5×10^7 、 7.5×10^7 、 1×10^8 或 5×10^8 个总的这种细胞,或前述任何两个值之间的范围。在一些实施方案中,对患者施用多个剂量,并且每个剂量或总剂量可以在任何前述值内。在一些实施方案中,细胞剂量包括施用(约) 1×10^7 至 0.75×10^8 个总的表达重组受体的CD8+ T细胞、 1×10^7 至 2.5×10^7 个总的表达重组受体的CD8+ T细胞、(约) 1×10^7 至 0.75×10^8 个总的表达重组受体的CD8+ T细胞,包括各端值。在一些实施方案中,细胞剂量包括施用(约) 1×10^7 、 2.5×10^7 、 5×10^7 、 7.5×10^7 、 1×10^8 或 5×10^8 个总的表达重组受体的CD8+ T细胞。

[0559] 在一些实施方案中,将细胞(例如表达重组受体的T细胞)的剂量以单剂量形式施用于受试者,或在两周、一个月、三个月、六个月、一年或更长的时间内仅施用一次。

[0560] 在一些实施方案中,所述剂量包含一定数目的细胞、一定数目的表达重组受体(例如,CAR)的细胞、一定数目的T细胞或一定数目的外周血单核细胞(PBMC),其范围为每千克受试者体重约 10^5 至约 10^6 个这种细胞,和/或这种细胞的数目为不多于过每千克受试者体重约 10^5 或约 10^6 个这种细胞。例如,在一些实施方案中,第一剂量或后续剂量包括少于或不多于每千克受试者体重(约) 1×10^5 、(约) 2×10^5 、(约) 5×10^5 或(约) 1×10^6 个这种细胞。在一些实施方案中,第一剂量包括每千克受试者体重(约) 1×10^5 、(约) 2×10^5 、(约) 5×10^5 或(约) 1×10^6 个这种细胞,或在前述任何两个值之间的范围内的值。在具体的实施方案中,细胞的数目和/或浓度是指表达重组受体(例如CAR)的细胞的数目。在其他实施方案中,细胞的数目和/或浓度是指所施用的所有细胞、T细胞或外周血单核细胞(PBMC)的数目或浓度。

[0561] 在一些实施方案中,例如,当受试者是人时,剂量包括少于约 1×10^8 个总的表达重组受体(例如,CAR)的细胞、T细胞或外周血单核细胞(PBMC),例如,范围为约 1×10^6 个至 1×10^8 个这种细胞,比如 2×10^6 个、 5×10^6 个、 1×10^7 个、 5×10^7 个或 1×10^8 个总的这种细胞,或前述任何两个值之间的范围。

[0562] 在一些实施方案中,所述剂量含有每 m^2 受试者少于约 1×10^8 个总的表达重组受体(例如,CAR)的细胞、T细胞或外周血单核细胞(PBMC),例如,范围为每 m^2 受试者约 1×10^6 个至 1×10^8 个这种细胞,比如每 m^2 受试者 2×10^6 个、 5×10^6 个、 1×10^7 个、 5×10^7 个或 1×10^8 个这种细胞,或前述任何两个值之间的范围。

[0563] 在某些实施方案中,剂量中的细胞、表达重组受体(例如,CAR)的细胞、T细胞或外周血单核细胞(PBMC)的数目大于每千克受试者体重约 1×10^6 个这种细胞,例如每千克受试者体重 2×10^6 、 3×10^6 、 5×10^6 、 1×10^7 、 5×10^7 、 1×10^8 、 1×10^9 或 1×10^{10} 个这种细胞,和/或每 m^2 受试者或总计 1×10^8 、 1×10^9 或 1×10^{10} 个这种细胞,或前述任何两个值之间的范围。

[0564] 在一些方面,基于一个或多个标准来确定剂量的大小,所述标准比如受试者对先前治疗例如化学疗法的反应;受试者的疾病负荷,比如肿瘤负荷、体积、大小或程度;转移的范围或类型、阶段;和/或受试者发生毒性结果的可能性或发生率,所述毒性结果例如CRS、巨噬细胞活化综合征、肿瘤溶解综合征、神经毒性和/或针对所施用的细胞和/或重组受体

的宿主免疫应答。

[0565] 在一些方面,通过受试者中的疾病或病症的负荷确定剂量的大小。例如,在一些方面,基于在施用细胞剂量开始之前即刻存在于受试者中的肿瘤负荷来确定在剂量中施用的细胞数。在一些实施方案中,第一和/或后续剂量的大小与疾病负荷呈反相关。在一些方面,如在大疾病负荷的情况下,将少量细胞,例如每千克受试者体重少于约 1×10^6 个细胞施用于受试者。在其他实施方案中,如在疾病负荷较低的情况下,将大量的细胞,比如每千克受试者体重多于约 1×10^6 个细胞施用于受试者。

[0566] 在一些方面,可以使用表观遗传学特性的评估结果(例如来自ATAC-seq分析的谱)来确定治疗方案,例如包括工程化细胞的施用剂量、时间和/或频率。在一些方面,可以使用如使用本文所述的方法确定的细胞组合物(比如工程化细胞组合物或从受试者获得的预工程化细胞组合物)的特点和/或特性来选择将治疗的受试者和/或确定适当的治疗剂量。在一些实施方案中,可以使用本文所述的方法来评估施用的组合物中的细胞的临床剂量。

试剂盒和制品

[0567] 还提供了试剂盒和制品,比如含有用于执行本文中提供的方法的试剂的那些试剂盒和制品,所述试剂例如用于评估细胞(例如用于工程化的细胞)中的一个或多个基因组区域(例如基因组基因座)的表观遗传学特性的试剂。在一些实施方案中,试剂盒还包括用于评估一个或多个基因组区域(例如基因组基因座)的其他参数或进行另外的表观遗传学和/或表观基因组学分析的试剂。

[0568] 在一些实施方案中,提供了包括核酸、插入酶和插入元件的试剂盒,其中:插入元件可包含含有预定序列的核酸,并且插入酶可进一步包含亲和标签。在一些实施方案中,试剂盒进一步包括缔合分子,例如与核酸缔合的蛋白质(例如组蛋白)或核酸(例如适体)。在一些实施方案中,亲和标签可以是抗体。在一些实施方案中,抗体可以结合转录因子、修饰的核小体和/或修饰的核酸。修饰的核酸的实例包括但不限于甲基化或羟甲基化DNA。亲和标签也可以是单链核酸(例如ssDNA、ssRNA)。在一些实施方案中,单链核酸可以与靶核酸结合。在一些情况下,插入酶可以进一步包含核定位信号。

[0569] 在一些实施方案中,试剂盒进一步包括:(a)用于从细胞群分离细胞核的试剂;(b)转座酶和转座子标签,和(c)转座酶反应缓冲液,其中试剂盒组分被配置为使得反应缓冲液、转座酶和衔接子与细胞核的体外组合导致细胞核裂解以释放染色质并且产生带有衔接子标签的基因组DNA片段。

[0570] 在一些实施方案中,试剂盒可包括:(a)细胞裂解缓冲液;(b)包含亲和标签的插入酶;和(c)包含核酸的插入元件,其中所述核酸包含预定序列。插入酶可以是例如转座酶。插入酶还可以包含两个或更多个连接在一起的酶模块。

[0571] 在一些实施方案中,试剂盒任选地包括其他组分,例如:PCR引物、诸如聚合酶、缓冲液、核苷酸的PCR试剂,用于另外的测定法的试剂,所述另外的测定法例如细胞内细胞因子染色、流式细胞术、染色质免疫沉淀和/或另外的表观遗传学和/或表观基因组学分析。在一些实施方案中,用于另外的测定法的试剂包括用于进行体外测定法以测量特定分子的表达或水平的组分。在一些情况下,体外测定法是免疫测定、基于适体的测定、组织学或细胞学测定或mRNA表达水平测定。在一些实施方案中,体外测定法选自酶联免疫吸附测定(ELISA)、免疫印迹、免疫沉淀、放射免疫测定(RIA)、免疫染色、流式细胞术测定、表面等离

子体共振 (SPR)、化学发光测定、侧向流动免疫测定、抑制测定和亲合力测定。在一些方面,试剂是特异性地结合分子的结合试剂。在一些情况下,结合试剂是抗体或其抗原结合片段、适体或核酸探针。试剂盒的各种组分可以存在于分开的容器中,或者可以将某些相容的组分预组合到单个容器中。在一些实施方案中,试剂盒进一步包括关于使用试剂盒的组分来实践所提供的方法的说明书。

定义

[0572] 除非另外定义,本文中使用的的所有术语、符号和其他技术和科学术语或用辞预期具有与所要求保护的主体所属领域的普通技术人员通常所理解的相同的含义。在一些情况下,为了清楚和/或为了便于参考而在本文中定义了具有通常理解的含义的术语,并且本文中包含的这些定义不应被解释为表示与本领域通常所理解的实质性差异。

[0573] 出于所有目的将本申请中提及的所有出版物,包括专利文献、科学论文和数据库在内,通过提述完整并入本文,其程度如同每一单独的出版物通过提述单独并入一样。如果本文中阐述的定义与通过提述并入本文的专利、申请、公开申请和其他出版物中阐述的定义相反或以别的方式不一致,则本文中阐述的定义优先于通过提述并入本文的定义。

[0574] 在本文中使用的章节标题仅仅是出于组织的目的,而不应当解释为限制所描述的主题。

[0575] 如本文中使用的,单数形式“一个(a)”、“一个(an)”和“该(the)”包括复数指示物,除非上下文另外明确规定。例如,“一个(a)”或“一个(an)”表示“至少一个”或“一个或多个”。应当理解的是,在本文中描述的方面和变体包括“由...组成”和/或“基本上由...组成”的方面和变体。

[0576] 如本文中使用的术语“约”是指本技术领域的技术人员容易知道的对应值的通常误差范围。在本文中提及的“大约”值或参数包括(并描述)针对该值或参数本身的实施方案。例如,提及“约X”的描述包括“X”的描述。

[0577] 贯穿本公开内容,将要求保护的主体各个方面以范围格式呈现。应当理解的是,范围形式的描述仅仅是为了方便和简洁,并且不应该被解释为对要求保护的主体范围的硬性限制。因此,应该认为范围的描述具体公开了所有可能的子范围以及该范围内的各个数值。例如,在提供一系列值的情况下,应当理解的是,在该范围的上限和下限之间的每个中间值以及在该陈述范围内的任何其他陈述值或中间值都被包括在所要求保护的主体内。这些较小范围的上限和下限可以被独立地包括在所述较小范围内,并且还被包括在要求保护的主体内,受制于所陈述范围内的任何特别排除的极限值。当陈述范围包括极限值之一或两者时,排除这些被包括的极限值的任一者或两者的范围也被包括在要求保护的主体内。无论范围的宽度如何,这都是适用的。

[0578] 如本文中使用的,组合物是指包括细胞在内的两种或更多种产物、物质或化合物的任何混合物。它可以是溶液、悬浮液、液体、粉末、糊剂、水性的、非水性的或其任何组合。

[0579] 如本文中使用的,细胞或细胞群对于特定标志物是“阳性”的陈述是指特定标志物(通常是表面标志物)在细胞上或细胞中的可检测存在。当提及表面标志物时,该术语是指通过流式细胞术检测的表面表达的存在,例如,通过用与标志物特异性地结合的抗体染色并检测所述抗体,其中在基本上高于在其他条件都相同时用同种型匹配的对照进行相同的程序检测到的染色的水平和/或在与已知对标志物呈阳性的细胞基本上相似的水平、和/或

在比已知对标志物呈阴性的细胞显著更高的水平可通过流式细胞术检测染色。

[0580] 如本文中使用的,细胞或细胞群对于特定标志物是“阴性”的陈述是指特定标志物(通常是表面标志物)在细胞上或细胞中的实质性可检测存在的缺乏。当提及表面标志物时,该术语是指通过流式细胞术检测的表面表达的缺乏,例如,通过用与标志物特异性地结合的抗体染色并检测所述抗体,其中在基本上高于在其他条件都相同时用同种型匹配的对照进行相同的程序检测到的染色的水平、和/或在比已知对标志物呈阳性的细胞显著更低的水平、和/或在与已知对标志物呈阴性的细胞相比基本上相似的水平通过流式细胞术未检测到染色。

[0581] 如本文中使用的术语“载体”是指能够传送与其连接的另一个核酸的核酸分子。该术语包括作为自身复制核酸结构的载体以及组入其已被引入的宿主细胞中的基因组中的载体。某些载体能够指导与它们可操作连接的核酸的表达。这样的载体在本文中被称为“表达载体”。

[0582] 如本文中使用的,“受试者”是哺乳动物,比如人或其他动物,并且通常是人。

示例性实施方案

[0583] 其中提供的实施方案为:

1. 一种鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法,所述方法包括:

(a) 分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞或群包含在下列组合物中:(i) 第一细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二组合物,或(ii) 包含所述重组受体的第二细胞组合物;和

(b) 鉴定所述一个或多个基因组区域中的一个或多个,其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关连细胞疗法的结果,所述细胞疗法包括施用包含所述重组受体的第二细胞组合物。

2. 实施方案1的方法,其中所述结果任选地是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

3. 一种鉴定预测细胞疗法的治疗结果的一个或多个基因组区域的方法,所述方法包括:

(a) 确定或测量在第一治疗组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度;

(b) 确定或测量在第二治疗组合物中包含的细胞或细胞群的所述一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度;

(c) 将一个或多个基因组区域的(a)中的水平或程度和(b)中的水平或程度进行比较。

4. 实施方案3的方法,进一步包括鉴定一个或多个基因组区域中的一个或多个,其中与在(b)中确定或测量的水平或程度相比,在(a)中确定或测量的水平或程度是不同的,任选地显著不同。

5. 实施方案3或实施方案4的方法,其中对于多个基因组区域中的每一个,在(a)中检测或测量的水平或程度与在(b)中检测或测量的水平或程度之间的差异或显著差异指示,所述表观遗传学特性或其程度或水平关联发生或已经发生于第一和第二治疗组合物之一但

未发生于另一者的结果、预测所述结果或预测所述结果的可能性或风险,其中所述结果任选地是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

6. 实施方案1-5中任一项的方法,其中所述基因组区域包含基因组基因座或基因。

7. 实施方案1或实施方案2的方法,其中所述基因组区域包含基因的开放阅读框。

8. 实施方案1-7中任一项的方法,其中表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。

9. 实施方案1-8中任一项的方法,其中表观遗传学特性是染色质可及性。

10. 实施方案1-9中任一项的方法,其中:

所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者

所述表观遗传学特性包括跨越基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

11. 实施方案8-10中任一项的方法,其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

12. 实施方案8-11中任一项的方法,其中通过ATAC-seq确定染色质可及性。

13. 实施方案1-12中任一项的方法,其中分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与沿着一个或多个基因组区域或其子集的每一个的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者

包括,对于沿着基因组区域长度的多个位点或部分中的每一个,生成指示在所述位点或部分的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述位点或部分的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

14. 实施方案13的方法,其中所述分析任选地进一步包括确定所述表观遗传学读出的总体程度或水平,任选地确定基因组区域上的可及性的总体程度或水平。

15. 实施方案1-14中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。

16. 实施方案1-14中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化与跨越一个或多个基因组区域或其子集的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

17. 实施方案15或实施方案16的方法,其中所述值或水平是或包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。

18. 实施方案15-17中任一项的方法,其中所述值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进行总计或求和。

19. 实施方案1-18中任一项的方法,其中对多位各自独立地被施用第二细胞组合物的受试者进行步骤(a)和(b),所述第二细胞组合物包含用重组受体工程改造的细胞。

20. 实施方案16-19中任一项的方法,其中,对于每个基因组区域或其子集,准备一个显

示器,所述显示器包括针对多位受试者的每一位的映射到细胞疗法结果的每个基因组基因座的序列读段的值或水平。

21. 实施方案20的方法,其中显示器包括热图、散点图、层次聚类和/或星座图。

22. 实施方案20或实施方案21的方法,其中所述鉴定所述一个或多个基因组区域包括基于细胞疗法的结果进行聚类分析。

23. 实施方案20或实施方案21的方法,其中所述鉴定指示或关联细胞疗法的结果的所述一个或多个基因组区域包括确定具有相同或相似结果的至少大多数受试者是否在显示器中聚集在一起。

24. 实施方案23的方法,其中如果具有相同或相似结果的至少55%、60%、70%、80%、90%、95%或更多的受试者在显示器中聚集在一起,则鉴定基因组区域。

25. 实施方案1-24中任一项的方法,其中分析了细胞的全基因组。

26. 实施方案1-24中任一项的方法,其中分析了细胞的基因组的一部分。

27. 实施方案26的方法,其中基因组的一部分包括一个或多个基因组区域,任选地为一个或多个基因组基因座,其关联或指示或可能关联或指示细胞的表型、活化状态、活化信号的强度或效应子功能。

28. 实施方案1-27中任一项的方法,其中细胞疗法的结果是缓解、毒性、细胞疗法的免疫原性或表型或功能、完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

29. 实施方案28的方法,其中缓解是完全缓解、部分缓解、疾病进展或可分子检测的疾病。

30. 实施方案28的方法,其中毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

31. 实施方案28或实施方案30的方法,其中毒性是剂量限制性毒性(DLT)。

32. 实施方案1-31中任一项的方法,其中分析了(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的表观遗传学特性。

33. 实施方案1-32中任一项的方法,其中鉴定了包含两个或更多个基因组区域的集合。

34. 实施方案1-33中任一项的方法,其中第一细胞组合物和第二细胞组合物包含从受试者选择或分离的原代细胞。

35. 实施方案1-34中任一项的方法,其中所述细胞是免疫细胞。

36. 实施方案1-35中任一项的方法,其中免疫细胞是T细胞或NK细胞。

37. 实施方案1-36中任一项的方法,其中T细胞是CD4⁺和/或CD8⁺ T细胞。

38. 实施方案1-37中任一项的方法,其中分析了第二细胞组合物。

39. 实施方案1-38中任一项的方法,其中第二细胞组合物包含编码重组受体的核酸。

40. 实施方案39的方法,其中核酸分子被包含在病毒载体中。

41. 实施方案40的方法,其中病毒载体是腺病毒、慢病毒、逆转录病毒、疱疹病毒或腺相关病毒载体。

42. 实施方案1-41中任一项的方法,其中通过在一种或多种条件或试剂的存在下培养输入组合物来产生第一细胞组合物和/或第二细胞组合物。

43. 实施方案1-42中任一项的方法,其中一个或多个基因组区域包含涉及或可能涉及

细胞的活化状态或效应子状态的基因。

44. 一种评估将施用于受试者的细胞组合物的方法,所述方法包括:

(a) 分析在细胞组合物中包含的细胞的一个或多个基因组区域的表观遗传学谱,所述细胞组合物包含用重组受体工程改造的细胞;和

(b) 将每个基因组区域的表观遗传学谱单独地与参考谱比较,其中所述比较指示当施用于受试者时细胞群是否或是否可能表现出或产生结果。

45. 实施方案144的方法,其中细胞疗法的结果是缓解、毒性、细胞疗法的免疫原性或表型或功能、完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发、缓解的耐久性、与功效相关或指示功效的结果、或与毒性相关或指示毒性的结果。

46. 实施方案45的方法,其中缓解是完全缓解或部分缓解。

47. 实施方案44-46中任一项的方法,其中如果所述比较指示细胞组合物将要或可能将要表现出结果,则将细胞组合物施用于受试者。

48. 实施方案47的方法,其中如果所述比较指示细胞组合物不会或可能不会表现出结果,则任选下列之一:

(i) 施用其中细胞组合物被改变的细胞组合物;

(ii) 施用其中细胞剂量被改变的细胞组合物;

(iii) 施用其中施用于受试者的细胞的剂量方案被改变的细胞组合物;

(iv) 将细胞组合物与一种或多种其他治疗剂联合施用;或

(v) 不对受试者施用所述细胞组合物。

49. 实施方案48的方法,其中在施用改变的细胞组合物之前,对包含在改变的细胞组合物中的细胞重复步骤(a)和(b)。

50. 实施方案49的方法,其中改变细胞的给药方案包括在对受试者施用第一剂量的细胞之后对受试者施用第二剂量的细胞。

51. 实施方案50的方法,其中在施用第一剂量的细胞之后至少1周、2周、3周、4周、2个月、3个月、4个月、5个月、6个月、9个月或12个月施用后续剂量的细胞。

52. 实施方案44-51中任一项的方法,其中一个或多个基因组区域关联或指示对细胞疗法的反应。

53. 实施方案44-52中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

54. 实施方案53的方法,其中所述阈值:

是当施用于具有相同或相似疾病或病症的受试者时显示表现出结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者

是来自当施用于受试者时显示表现出结果的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。

55. 实施方案44-54中任一项的方法,其中阈值包括来自正常或健康受试者的细胞中表观遗传学特性的值或水平。

56. 实施方案44-55中任一项的方法,其中阈值包括表现出幼稚或长寿记忆表型的细胞

中表观遗传学特性的值或水平。

57. 一种评估细胞培养物的方法,其包括:

(a) 分析包含在输出细胞组合物中的细胞的一个或多个基因组区域的表现遗传学谱,所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生;和

(b) 将每个基因组区域的表现遗传学谱单独地与参考谱比较,其中所述比较指示所述细胞是否或是否可能表现出预定的表型或功能。

58. 实施方案57的方法,其中预定的表型或功能指示细胞的效应子功能或活化状态和/或指示细胞表现出幼稚表型或长寿记忆表型。

59. 实施方案57或实施方案58的方法,其中一种或多种测试剂或条件包括血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;输入组合物的来源或细胞类型;输入组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌注;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的输入组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。

60. 实施方案58或实施方案59的方法,其中一种或多种测试剂或条件包括来自测试化合物文库的一种或多种化合物。

61. 实施方案57-60中任一项的方法,其包括如果所述比较指示细胞组合物具有或可能具有所述表型或功能,则选择用于培养细胞的一种或多种测试剂或条件。

62. 实施方案57-60中任一项的方法,其包括如果所述比较指示细胞组合物具有或可能不具有所述表型或功能,则用一种或多种另外的测试剂或条件重复步骤(a)和(b)。

63. 实施方案57-62中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表现遗传学特性的阈值或一个或多个基因组区域内的总表现遗传学特性的阈值。

64. 实施方案63的方法,其中所述阈值:

是显示表现出所述表型或功能的参考细胞组合物的细胞中的一个或多个基因组区域中的表现遗传学特性(任选地染色质可及性)的值或水平;或者

是来自显示表现出所述表型或功能的多个参考细胞组合物的每一个的细胞的一个或多个基因组区域中的表现遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。

65. 实施方案64的方法,其中参考细胞组合物具有指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型。

66. 实施方案44-65中任一项的方法,其中分析表现遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。

67. 实施方案44-66中任一项的方法,其中分析表现遗传学特性包括确定、测量或量化与一个或多个基因组区域或其子集的表现遗传学特性(任选地染色质可及性)相关或指示所述表现遗传学特性的序列读段的值或水平。

68. 实施方案66或实施方案67的方法,其中确定、测量或量化值或水平包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。

69. 实施方案66-68中任一项的方法,其中确定、测量或量化值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进

行总计或求和。

70. 实施方案1-69中任一项的方法,其中所述一个或多个基因组区域包括包含至少2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的集合。

71. 实施方案27-70中任一项的方法,其中所述一个或多个基因组区域包含与细胞的效应子样功能或活化状态相关或指示所述效应子样功能或活化状态的一个或多个基因组基因座。

72. 实施方案27-71中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因座,该组的组成为:Nr4a1、Cblb、Irf4、Tbx21、Eomes、Ifng、Il2ra、Il2、Csf2、Gzmb、Tnfsf10、Gata3、Mir155、Sox21、Ctla4、Lag3和Pdcd1。

73. 实施方案27-72中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因组基因座,该组的组成为:Ctla4、Il2ra、Il2、Ifng和Gzmb。

74. 实施方案27-73中任一项的方法,其中所述基因组区域包含基因组基因座或基因。

75. 实施方案27-74中任一项的方法,其中所述基因组区域包含基因的开放阅读框。

76. 实施方案27-75中任一项的方法,其中表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。

77. 实施方案27-76中任一项的方法,其中表观遗传学特性是染色质可及性。

78. 实施方案77的方法,其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

79. 实施方案77或实施方案78的方法,其中通过ATAC-seq确定染色质可及性。

80. 实施方案1-79中任一项的方法,其中所述细胞获自来自受试者的样品。

81. 实施方案80的方法,其中所述细胞是免疫细胞,任选T细胞,任选CD4+和/或CD8+ T细胞。

82. 实施方案1-81中任一项的方法,其中:

重组受体结合、识别或靶向与疾病或病症相关的抗原;并且/或者重组受体是T细胞受体或功能性非T细胞受体;并且/或者

重组受体是嵌合抗原受体(CAR)。

83. 实施方案82的方法,其中:

CAR包含与抗原特异性地结合的细胞外抗原识别结构域和包含ITAM的细胞内信号传导结构域,其中任选地,细胞内信号传导结构域包含CD3- ζ (CD3 ζ) 链的细胞内结构域;并且/或者其中CAR进一步包含共刺激信号传导区,其任选地包含CD28或4-1BB的信号传导结构域。

84. 一种细胞组合物,其包含多个细胞,其中一个集合中的一个或多个基因的表现遗传学特性的水平或值高于或低于组合物中至少50%细胞中的阈值。

85. 实施方案84的细胞组合物,其中所述水平或值高于或低于组合物中至少60%、70%、75%、80%、85%、90%、95%或更多细胞中的阈值。

86. 实施方案84或实施方案85的细胞组合物,其中所述阈值:

是显示表现出所述表型或功能的参考细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)的值或水平;或者

是来自显示表现出所述表型或功能的多个参考细胞组合物的每一个的细胞的一个或

多个基因组区域中的表观遗传学特性(任选地染色质可及性)的平均、中位或平均的值或水平。

87. 实施方案84-86中任一项的细胞组合物,其中参考细胞组合物具有指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型。

88. 实施方案84-87中任一项的组合物,其中所述集合包含(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域。

89. 一种鉴定与细胞疗法的治疗结果相关的一个或多个基因组区域的方法,所述方法包括:

(a) 分析或确定细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞或群包含在下列组合物中:(i) 第一细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二组合物,或(ii) 包含所述重组受体的第二细胞组合物;和

(b) 鉴定所述一个或多个基因组区域中的一个或多个,其中总体上跨越所述一个或多个基因组区域的表观遗传学特性预测、指示或关连细胞疗法的结果,所述细胞疗法包括对一位受试者或一组受试者施用包含所述重组受体的第二细胞组合物。

90. 实施方案89的方法,其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

91. 实施方案90的方法,其中所述结果是反应,并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

92. 实施方案90的方法,其中所述结果是毒性,并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

93. 实施方案90或实施方案92的方法,其中毒性是剂量限制性毒性(DLT)。

94. 实施方案89-93中任一项的方法,其中:

所述第一组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞;并且/或者

所述第二组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞。

95. 一种确定细胞组合物的一个或多个特性或特征的方法,所述方法包括分析或确定T细胞组合物中的一个或多个基因组区域的表观遗传学特性,所述T细胞组合物富含CD4+原代人T细胞和/或CD8+原代人T细胞。

96. 实施方案95的方法,其中细胞组合物是(i) 第一T细胞组合物,其有待用重组受体基因工程改造以产生包含所述重组受体的第二T细胞组合物,或(ii) 包含所述重组受体的细胞的第二T细胞组合物。

97. 实施方案84-96中任一项的方法,其中:

所述细胞组合物包含大于或大于约70%、大于或大于约75%、大于或大于约80%、大于或大于约85%、大于或大于约90%、大于或大于约95%或大于或大于约98%的CD4+和/或CD8+原代人T细胞;并且/或者

所述细胞组合物基本上由CD4+和/或CD8+原代人T细胞组成。

98. 实施方案95-97的方法,进一步包括将一个或多个基因组区域中的每一个的表观遗传学特性单独地与来自不同的细胞组合物的细胞的相应表观遗传学特性和/或与参考谱进行比较,所述参考谱任选地为已知指示或关联细胞组合物的属性或特征的参考谱。

99. 实施方案98的方法,其中所述比较指示或关联细胞组合物内的细胞的状态、表型或功能,任选地为活化、效应子或记忆状态;细胞组合物内的细胞的一致性 or 均匀性;当施用于一位受试者或一组受试者时,细胞组合物是否或是否可能表现出或产生结果;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;和/或细胞组合物中的工程化细胞的比例或频率。

100. 一种用于确定或鉴定与细胞组合物的属性或特征相关的表观遗传学特性的方法,所述方法包括:

(a) 确定或测量在第一细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性的水平或程度或相对水平或程度;

(b) 确定或测量在第二细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的所述表观遗传学特性的水平或程度或相对水平或程度;和

(c) 比较(a)中的水平或程度和(b)中的水平或程度,其中一个或多个基因组区域的表观遗传学特性的水平或程度的差异(任选地显著差异)鉴定或确定表观遗传学特性的存在,所述表观遗传学特性指示或关联在第一和第二组合物之一的细胞中存在但在另一者中不存在的属性或特征。

101. 实施方案100的方法,其中:

所述第一组合物和第二组合物之一包含有待用重组受体基因工程改造的细胞,并且所述第一组合物和第二组合物中的另一者包含经工程改造以表达重组受体的细胞;

所述第一组合物和第二组合物包含来自不同供体的原代细胞,所述供体任选基于疾病状态、疾病严重性或疾病类型而不同的供体;

所述第一组合物和第二组合物包含处于工程化细胞的制造过程的不同阶段或步骤的细胞;

所述第一组合物和第二组合物之一包含与药剂接触以调节细胞的活性、表型或功能的细胞,并且所述第一和第二组合物中的另一者包含并不如此接触的相似细胞;或者

所述第一组合物和第二组合物之一包含细胞组合物样品,所述细胞组合物样品与所述第一和第二组合物之一施用于受试者后发生或已经发生但另一者则不然的结果相关。

102. 实施方案101的方法,其中所述药剂是多肽或蛋白质、肽、抗体、核酸、病毒载体或病毒制剂或小分子化合物。

103. 实施方案101或实施方案102的方法,其中所述药剂是刺激性试剂,任选抗CD3/抗CD28;免疫调节剂、对CAR特异的抗独特型抗体或其抗原结合片段、免疫检查点抑制剂、代谢途径调节剂、腺苷受体拮抗剂、激酶抑制剂、抗TGF β 抗体或抗TGF β R抗体或细胞因子。

104. 实施方案100-103中任一项的方法,其中第一组合物的属性或特征指示状态、表型或功能,任选地活化、效应子或记忆状态、表型或功能;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;细胞组合物中的工程化细胞的比例或频率;和/或当施用于一位受试者或一组受试者组时,细胞组合物是否或是否可能表现出或产生结果。

105. 实施方案100或实施方案104的方法,其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

106. 实施方案105的方法,其中所述结果是反应,并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

107. 实施方案106的方法,其中所述结果是毒性,并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

108. 实施方案106或实施方案107的方法,其中毒性是剂量限制性毒性(DLT)。

109. 实施方案100-108中任一项的方法,其被重复多次。

110. 实施方案109的方法,其中鉴定或确定存在于所述第一和第二组合物之一的大多数中但不存在于另一者中的表观遗传学特性。

111. 一种评估细胞组合物的属性或特征的方法,其包括:

(a) 分析在细胞组合物中包含的细胞或细胞群的一个或多个基因组区域的表观遗传学特性,所述细胞组合物包含用重组受体工程改造的细胞和/或有待用重组受体基因工程改造的细胞;和

(b) 将一个或多个基因组区域的表观遗传学特性单独地与参考谱比较,其中所述比较指示细胞组合物是否或是否可能表现出所述属性或特征。

112. 实施方案111的方法,其中所述属性或特征指示状态、表型或功能,任选地活化、效应子或记忆状态、表型或功能;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;细胞组合物中的工程化细胞的比例或频率;和/或当施用于一位受试者或一组受试者组时,细胞组合物是否或是否可能表现出或产生结果。

113. 实施方案113或实施方案112的方法,其中所述属性或特征是当施用于一位受试者或一组受试者时细胞组合物是否或是否可能表现出或产生结果,并且所述方法是为了评估将施用于受试者的细胞组合物。

114. 实施方案112或实施方案113的方法,其中所述结果是与功效、缓解、持久性、毒性或免疫原性相关或指示它们的结果。

115. 实施方案114的方法,其中所述结果是反应,并且所述反应是完全缓解、部分缓解、疾病进展、可分子检测的疾病、复发或缓解的耐久性。

116. 实施方案115的方法,其中所述结果是毒性,并且所述毒性是细胞因子释放综合征(CRS)、重度CRS、3级或更高等级的CRS、神经毒性、重度神经毒性、3级或更高等级的神经毒性和/或脑水肿。

117. 实施方案115或实施方案116的方法,其中毒性是剂量限制性毒性(DLT)。

118. 实施方案111-117中任一项的方法,其中如果所述比较指示细胞组合物将要或可能将要表现出期望的结果,则将细胞组合物施用于受试者。

119. 实施方案111-118中任一项的方法,其中如果所述比较指示细胞组合物不会或可能不会表现出期望的结果,则任选下列之一:

(i) 施用其中细胞组合物被改变的细胞组合物;

(ii) 施用其中细胞剂量被改变的细胞组合物;

(iii) 施用其中施用于受试者的细胞的剂量方案被改变的细胞组合物;

(iv) 将细胞组合物与一种或多种其他治疗剂联合施用;或

(v) 不对受试者施用所述细胞组合物。

120. 实施方案118或实施方案119的方法,其中所述期望的结果是完全缓解、部分缓解或持续缓解和/或是3级或更低等级的神经毒性或1级或2级神经毒性,是3级或更低等级的

CRS,或是1级或2级CRS,或者不包括任何等级的神经毒性或任何等级的CRS。

121.实施方案119或实施方案120的方法,其中通过在工程改造细胞组合物中的细胞的一个或多个步骤中改变一种或多种作用剂或条件来改变细胞组合物。

122.实施方案121的方法,其中所述一种或多种作用剂或条件选自血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;细胞组合物的来源或细胞类型;输入组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌注;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的细胞组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。

123.实施方案121或实施方案122的方法,其中在施用改变的细胞组合物之前,对包含在改变的细胞组合物中的细胞重复步骤(a)和(b)。

124.实施方案123的方法,其中改变细胞的给药方案包括在对受试者施用第一剂量的细胞之后对受试者施用第二剂量的细胞。

125.实施方案124的方法,其中在施用第一剂量的细胞之后至少1周、2周、3周、4周、2个月、3个月、4个月、5个月、6个月、9个月或12个月施用后续剂量的细胞。

126.实施方案111-125中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

127.实施方案126的方法,其中所述阈值:

是与当施用于具有相同或相似疾病或病症的受试者时显示表现出期望的结果的细胞组合物的细胞中的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平;

是与来自已经单独施用于一组受试者的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的平均、中位或平均的值或水平,或者在所述平均、中位或平均的值或水平的标准偏差之内,其中该组受试者的每一位在施用后继续表现出期望的结果;或者

是与来自正常或健康受试者的相似细胞组合物中的表观遗传学特性相关或指示所述表观遗传学特性的值或水平。

128.实施方案111-125中任一项的方法,其中:

所述比较包括差异性可及性分析;并且/或者

所述参考谱包括参考表观遗传学图谱,其包含一个或多个基因组区域内的序列读段的峰。

129.实施方案128的方法,其中:

所述参考表观遗传学图谱是根据细胞组合物的可及性分析(任选地染色质可及性)确定的,所述细胞组合物在将所述细胞或细胞组合物施用于具有相同或相似疾病或病症的受试者后显示表现出期望的结果;

所述参考表观遗传学图谱是根据来自已经单独施用于一组受试者的多个细胞组合物中的可及性分析(任选地染色质可及性)的序列读段的共同峰确定的,其中该组受试者的每一位在施用后继续表现出期望的结果;或者

所述参考表观遗传学图谱是根据来自正常或健康受试者的相似细胞组合物的可及性

分析(任选地染色质可及性)确定的。

130. 一种评估细胞组合物的方法,其包括:

(a) 分析包含在输出细胞组合物中的细胞和/或包含在输入组合物中的细胞的一个或多个基因组区域的表观遗传学特性,所述输出组合物通过在一种或多种测试剂或条件存在下培养输入组合物而产生;和

(b) 将一个或多个基因组区域的表观遗传学特性单独地与参考谱比较,其中所述比较指示所述细胞是否或是否可能表现出预定的特征或属性。

131. 实施方案130的方法,其中预定的特征或属性是组合物内的细胞的状态、表型或功能;细胞组合物内的细胞的一致性 or 均匀性;外源核酸整合的位置、丰度或频率;细胞组合物内的细胞的克隆性;和/或细胞组合物中的工程化细胞的比例或频率。

132. 实施方案131的方法,其中预定的表型或属性是指示细胞的效应子功能或活化状态的状态、表型或功能并且/或者指示细胞表现出幼稚表型或长寿记忆表型。

133. 实施方案131或实施方案132的方法,其中一种或多种测试剂或条件包括血清的存在或浓度;培养时间;刺激剂的存在或量;刺激剂的类型或程度;氨基酸的存在或量;温度;输入组合物的来源或细胞类型;输入组合物中细胞类型的比率或百分比,任选地CD4+/CD8+细胞的比率;珠子的存在或量;细胞密度;静置培养;摇摆培养;灌注;病毒载体的类型;载体拷贝数;转导佐剂的存在;冷冻保存中的输入组合物的细胞密度;重组受体的表达程度;或调节细胞表型的化合物的存在。

134. 实施方案131-133中任一项的方法,其中一种或多种测试剂或条件包括来自测试化合物文库的一种或多种化合物。

135. 实施方案131-134中任一项的方法,其包括如果所述比较指示细胞组合物具有或可能具有期望的特征或属性,则选择用于培养细胞的一种或多种测试剂或条件和/或选择将施用于受试者的细胞组合物。

136. 实施方案131-134中任一项的方法,其包括如果所述比较指示细胞组合物不具有或可能不具有期望的特征或属性,则用一种或多种另外的测试剂或条件重复步骤(a)和(b)。

137. 实施方案131-136中任一项的方法,其中参考谱包括一个或多个基因组区域的每一个的表观遗传学特性的阈值或一个或多个基因组区域内的总表观遗传学特性的阈值。

138. 实施方案137的方法,其中所述阈值:

是与已知表现出期望的属性或特征的细胞组合物的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平;或者

是与来自已知表现出期望的属性或特征的多个细胞组合物的每一个的细胞的一个或多个基因组区域中的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的平均、中位或平均的值或水平,或者在所述平均、中位或平均的值或水平的标准偏差之内。

139. 实施方案131-136中任一项的方法,其中:

所述比较包括差异性可及性分析;并且/或者

所述参考谱包括参考表观遗传学图谱,其包含一个或多个基因组区域内的序列读段的

峰。

140. 实施方案139的方法, 其中:

所述参考表观遗传学图谱是根据已知表现出期望的属性或特征的细胞组合物的可及性分析(任选地染色质可及性)确定的;

所述参考表观遗传学图谱是根据来自已知表现出期望的属性或特征的多个细胞组合物的可及性分析(任选地染色质可及性)的序列读段的共有峰确定的。

141. 实施方案131-140中任一项的方法, 其中期望的结果或特征是指示幼稚T细胞、长寿记忆T细胞、中枢记忆T细胞(Tcm)或干细胞样记忆T细胞(Tcsm)的表型或功能。

142. 实施方案89-141中任一项的方法, 其中所述基因组区域包含基因组基因座或基因。

143. 实施方案89-142中任一项的方法, 其中基因组区域包含基因的编码区、开放阅读框、非编码区、基因间区或调控元件。

144. 实施方案89-143中任一项的方法, 其中所述基因组区域包含基因的开放阅读框。

145. 实施方案89-143中任一项的方法, 其中所述基因组区域包含基因间区或调控元件。

146. 实施方案89-143和145中任一项的方法, 其中基因组区域包含内含子、外显子、顺式调控元件、启动子、增强子、上游激活序列(UAS)、3'非翻译区(UTR)、5'UTR、非编码RNA产生区、非编码RNA(ncRNA)基因、miRNA基因、siRNA基因、piRNA基因、snoRNA基因、lncRNA基因、核糖体RNA(rRNA)基因、小RNA结合位点、非编码RNA结合位点、假基因、转录终止位点(TTS)、重复序列、端粒区、可及的染色质区、不可及的染色质区、开放染色质区和/或异染色质区。

147. 实施方案89-146中任一项的方法, 其中表观遗传学特性选自染色质可及性、核小体占据、组蛋白修饰、空间染色体构象、转录因子占据和DNA甲基化。

148. 实施方案89-147中任一项的方法, 其中表观遗传学特性是染色质可及性。

149. 实施方案89-148中任一项的方法, 其中:

所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度, 并且/或者

所述表观遗传学特性包括基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

150. 实施方案147-149中任一项的方法, 其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

151. 实施方案147-150中任一项的方法, 其中通过ATAC-seq确定染色质可及性。

152. 实施方案147-151中任一项的方法, 其中评估表观遗传学特性包括:

- (1) 从细胞或细胞群分离染色质,
- (2) 用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,
- (3) 对所有或部分带有标签的片段测序以产生多个序列读段;
- (4) 比对、过滤所述序列读段并且将其映射到基因组的基因组区域; 和
- (5) 确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

153. 实施方案152的方法,其中分析或评估表观遗传学特性进一步包括比较在来自两个或更多个细胞或细胞组合物的样品之间不同的序列读段的峰,并且任选地鉴定序列读段的峰。

154. 实施方案152或实施方案153的方法,其中序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。

155. 实施方案152-154中任一项的方法,其中分析或评估表观遗传学特性进一步包括进行基因组区域的基序分析、转录因子占据分析和/或生物途径分析,所述基因组区域被鉴定为含有在来自两个或更多个细胞群的样品之间不同的序列读段的峰。

156. 实施方案152-155中任一项的方法,其中分析或评估表观遗传学特性进一步包括确定含有序列读段峰的基因组区域内的核小体的位置。

157. 实施方案89-156中任一项的方法,其中分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与沿着一个或多个基因组区域或其子集的每一个的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者

包括,对于沿着基因组区域长度的多个位点或部分中的每一个,生成指示在所述位点或部分的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述位点或部分的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

158. 实施方案157的方法,其中所述分析任选地进一步包括确定所述表观遗传学特性的总体程度或水平,任选地确定基因组区域上的可及性的总体程度或水平。

159. 实施方案89-158中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化跨越一个或多个基因组区域的染色质可及性的值或水平。

160. 实施方案89-159中任一项的方法,其中分析表观遗传学特性包括确定、测量或量化与跨越一个或多个基因组区域或其子集的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

161. 实施方案127、138、159或实施方案160的方法,其中所述值或水平是或包括确定一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值。

162. 实施方案127、138和159-161中任一项的方法,其中所述值或水平是或包括对一个或多个基因组区域或其子集的每一个内的每百万映射读段的每千碱基片段数(FPKM)值进行总计或求和。

163. 实施方案89-162中任一项的方法,其中分析包括去除线粒体读段和/或另外的污染序列的步骤,其根据是所述读段的序列同一性、质量、映射位置或其他测序特性。

164. 实施方案89-163中任一项的方法,其中分析包括去除重复读段以提高定量准确性的步骤。

165. 实施方案89-164中任一项的方法,其中分析包括将序列读段分离成代表特定的表观遗传学特性(任选地染色质可及性或染色质占据)的子集的步骤,其中使用测序的片段的大小来确定其代表所述表观遗传学特性的程度或水平。

166. 实施方案89-94和142-165中任一项的方法,其中对来自多位各自独立地被施用第

二细胞组合物的受试者的细胞组合物进行步骤(a)和(b),所述第二细胞组合物包含用重组受体工程改造的细胞。

167. 实施方案89-94和142-166中任一项的方法,其中,对于每个基因组区域或其子集,准备一个显示器,所述显示器包括针对多位受试者的每一位的映射到细胞疗法结果的每个基因组基因座的序列读段的值或水平。

168. 实施方案167的方法,其中显示器包括热图、散点图、层次聚类和/或星座图。

169. 实施方案89-169和142-169中任一项的方法,其中所述鉴定所述一个或多个基因组区域包括基于细胞疗法的结果进行聚类分析。

170. 实施方案89-169和142-169中任一项的方法,其中所述鉴定指示或关联细胞疗法的结果的所述一个或多个基因组区域包括确定具有相同或相似结果的至少大多数受试者是否在显示器中聚集在一起。

171. 实施方案170的方法,其中如果具有相同或相似结果的至少55%、60%、70%、80%、90%、95%或更多的受试者在显示器中聚集在一起,则鉴定基因组区域。

172. 实施方案89-171中任一项的方法,其中分析了细胞的全基因组。

173. 实施方案89-172中任一项的方法,其中分析了细胞的基因组的一部分。

174. 实施方案173的方法,其中基因组的一部分包括一个或多个基因组区域,任选地为一个或多个基因组基因座,其关联或指示或可能关联或指示细胞的表型、活化状态、活化信号的强度或效应子功能。

175. 实施方案89-174中任一项的方法,其中所述分析进一步包括进行主成分分析(PCA)、生物途径分析、基因本体论(GO)分析和/或基序分析。

176. 实施方案175的方法,其中所述分析包括与T细胞记忆表型、T细胞活化状态、效应子功能、细胞因子反应、运输、持久性或衰竭相关的一个或多个基因组区域的生物途径分析和/或基因子集分析。

177. 实施方案174-176中任一项的方法,其中所述一个或多个基因组区域包含与细胞的效应子样功能或活化状态相关或指示所述效应子样功能或活化状态的一个或多个基因组基因座。

178. 实施方案174-177中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因座,该组的组成为:Nr4a1、Cblb、Irf4、Tbx21、Eomes、Ifng、Il2ra、Il2、Csf2、Gzmb、Tnfsf10、Gata3、Mir155、Sox21、Ctla4、Lag3和Pdcd1。

179. 实施方案174-178中任一项的方法,其中所述一个或多个基因组区域包含选自下组的基因组基因座,该组的组成为:Ctla4、Il2ra、Il2、Ifng和Gzmb。

180. 实施方案89-180中任一项的方法,其中:

分析了(约)2至50、2至20、2至10、2至5、5至50、5至20、5至10、10至50、10至20或20至50个基因组区域的表观遗传学特性;或者

分析了至少2、3、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、200、300、400、500个或更多个基因组区域的表观遗传学特性;或者

分析了仅一个基因组区域的表观遗传学特性。

181. 实施方案89-180中任一项的方法,其中鉴定了包含两个或更多个基因组区域的集合。

182. 一种评估转基因整合的方法,所述方法包括:确定在用重组受体基因工程改造的细胞或细胞组合物中的包含转基因核酸序列的一个或多个基因组区域的表观遗传学特性。

183. 实施方案182的方法,其中通过将编码重组受体的核酸引入到细胞组合物中的一个或多个细胞中来进行基因工程。

184. 实施方案183的方法,其中所述引入是通过用包含核酸的病毒载体转导。

185. 实施方案182-184中任一项的方法,其中表观遗传学特性是染色质可及性。

186. 实施方案182-185中任一项的方法,其中:

所述表观遗传学特性包括染色质可及性、染色质可及性的水平或程度、染色质可及性的相对水平或程度,并且/或者

所述表观遗传学特性包括基因组区域的染色质可及性的程度或水平、相对程度或水平、或谱或图谱。

187. 实施方案182-186中任一项的方法,其中通过具有高通量测序的转座酶可及的染色质测定法(ATAC-seq)或与高通量测序偶联的染色质免疫沉淀(ChIP-seq)来确定染色质可及性。

188. 实施方案182-187中任一项的方法,其中通过ATAC-seq确定染色质可及性。

189. 实施方案188中任一项的方法,其中评估表观遗传学特性包括:

- (1) 从细胞或细胞群分离染色质,
- (2) 用插入酶复合物处理染色质以生成带有标签的基因组DNA片段,
- (3) 对所有或部分带有标签的片段测序以产生多个序列读段;
- (4) 比对、过滤所述序列读段并且将其映射到基因组的基因组区域;和
- (5) 确定或鉴定每个细胞或细胞群的多个基因组区域中的序列读段的峰。

190. 实施方案189的方法,其中分析或评估表观遗传学特性进一步包括确定映射到或对应于转基因核酸序列的序列读段的峰。

191. 实施方案182-190中任一项的方法,其中序列读段的峰包含具有峰信号、水平或值的序列读段,所述峰信号、水平或值被富集、高于背景和/或与周围区域的序列读段相比更高。

192. 实施方案182-191中任一项的方法,其中分析表观遗传学特性包括生成表观遗传学图谱,所述图谱显示与包含转基因核酸序列的基因组区域的表观遗传学特性相关或指示所述表观遗传学特性的序列读段(任选地与染色质可及性相关或指示染色质可及性的序列读段)的谱,并且/或者

包括,对于包含沿着基因组区域长度的转基因核酸序列的基因组区域,生成指示在所述区域的表观遗传学读出(任选地染色质可及性)的一个或多个序列读段,其中所述一个或多个序列读段的数量指示在所述区域的所述表观遗传学特性(任选地所述染色质可及性)的程度或水平。

193. 实施方案182-192中任一项的方法,其中确定表观遗传学特性包括确定、测量或量化跨越包含转基因核酸序列的基因组区域的染色质可及性的值或水平。

194. 实施方案182-193中任一项的方法,其中确定表观遗传学特性包括确定、测量或量化与跨越包含转基因核酸序列的基因组区域的表观遗传学特性(任选地染色质可及性)相关或指示所述表观遗传学特性的值或水平。

195. 实施方案89-194中任一项的方法,其中细胞组合物,任选地第一细胞组合物和/或第二细胞组合物包含获自来自受试者的样品和/或从受试者选择或分离的原代细胞。

196. 实施方案89-195中任一项的方法,其中所述细胞是免疫细胞。

197. 实施方案89-196中任一项的方法,其中免疫细胞是T细胞或NK细胞。

198. 实施方案89-197中任一项的方法,其中T细胞是CD4⁺和/或CD8⁺T细胞。

199. 实施方案89-198中任一项的方法,其中:

重组受体结合、识别或靶向与疾病或病症相关的抗原;并且/或者重组受体是T细胞受体或功能性非T细胞受体;并且/或者

重组受体是嵌合抗原受体(CAR)。

200. 实施方案199的方法,其中:

CAR包含与抗原特异性地结合的细胞外抗原识别结构域和包含ITAM的细胞内信号传导结构域,其中任选地,细胞内信号传导结构域包含CD3- ζ (CD3 ζ) 链的细胞内结构域;并且/或者其中CAR进一步包含共刺激信号传导区,其任选地包含CD28或4-1BB的信号传导结构域。

实施例

[0584] 下面的实施例仅仅是为了说明的目的而被包括在内,并非意在限制本发明的范围。

实施例1: 样品制备和通过ATAC-seq分析CART细胞染色质可及性

[0585] 使用利用测序的转座酶可及的染色质测定法(ATAC-Seq),评估用嵌合抗原受体进行基因工程改造的CD4⁺/CD8⁺ T细胞组合物的染色质可及性。

[0586] 根据人外周血单核细胞(PBMC)的白细胞单采术,通过基于免疫亲和力的富集来分离CD4⁺和/或CD8⁺ T细胞。将分离的CD4⁺/CD8⁺ T细胞激活并用编码抗CD19 CAR的病毒载体转导。所述病毒载体构建体进一步编码截短的EGFR(EGFRt),所述EGFRt用作CAR表达的替代标志物;EGFRt编码区通过T2A跳跃序列与CAR序列分开。转导后,使细胞在培养基中扩增,通过冷冻保存进行冷冻。制备了来自19位受试者的总共43个冷冻保存的工程化细胞组合物(CDP)。还冷冻保存(CMAT)和评估了与这些受试者中的18位匹配但未经历基因工程的CD4⁺/CD8⁺细胞组合物。在一些情况下,通过表型将CMAT样品分离为幼稚T细胞(T_N)、中枢记忆T细胞(T_{CM})、效应与效应记忆T细胞(T_{E+EM})或效应记忆RA(T_{EMRA})用于分析。为了生成用于ATAC-Seq分析的文库,将细胞解冻、洗涤并裂解。然后使用酶(Tn5转座酶)将DNA片段化并加标签(“标签化”),所述酶介导双链DNA的片段化并将合成的寡核苷酸连接至DNA片段。然后使用柱净化样品,接着进行五个PCR扩增循环和柱纯化。对样品进行qPCR扩增,并在所有样品上观察到正常扩增,其指示成功标签化的DNA。基于核小体显带法(banding)定性地确定了成功的文库生成。针对用作定量的校正因子的文库大小分布,在大小选择的DNA上运行Agilent D1000电泳图谱。大小选择去除了可能干扰测序的残留引物和引物二聚体。然后制作qPCR稀释液并合并文库用于测序。在测序之后,将DNA与参考基因组比对(Langmead等人, Genome Biol. (2009) 10:R25)。使用各种测序度量分析了序列数据的质量,所述度量包括针对总映射读段、与基因组比对的%、非冗余部分(冗余)、线粒体DNA污染、有效序列深度以及全基因组可及性峰和峰中的读段片段比例(FRiP)的度量。观察到跨越使用上述方法生成的文库的平均>97%的读段与参考基因组比对。

[0587] 执行了另外的处理步骤以确保数据质量,包括过滤出质量检查失败的读段、去除重复以及固定读段配对。鉴定了核小体定位、转座子插入位点和转录因子占据,并使用基因组浏览器使数据可视化(美国专利申请公开号US 20160060691)。

[0588] 执行峰调用以鉴定具有可及染色质的序列读段。为了量化染色质可及性,在一些研究中确定了在每个基因体之内的每百万映射片段的每千碱基片段数 (FPKM),并且使用DESeq2软件根据FRiP归一化的测序标签计数来计算其他的峰可及性。在所有情况下,通过区间分析鉴定了共同和独特的峰,然后通过基因体FPKM值或FRIP归一化的峰计数的下游询问执行定量和/或差异性可及性。这些计数度量用于与其他测定法直接关连,输入到基因模块分析中或被询问在各个感兴趣的组之间的差异。包括后续研究在内,已经对大约450个ATAC-seq文库进行了测序。

实施例2:通过流式细胞术和ATAC-seq评估T细胞标志物

[0589] 使用ATAC-Seq来评估与含有基因工程细胞的细胞组合物中的T细胞活化状态或反应相关的基因的染色质可及性,并将其与通过细胞内细胞因子染色 (ICS; 在一些情况下被称为细胞内流式细胞术) 的相同基因的蛋白质表达进行比较。将如实施例1中所述产生的表达抗CD19 CAR的基因工程人T细胞解冻。为了通过流式细胞术评估蛋白质标志物,在Golgi抑制剂存在下用佛波醇肉豆蔻酸乙酸酯 (PMA) /离子霉素再刺激在培养物中的细胞,并通过流式细胞术评估。为了测量基因的染色质可及性,通过如实施例1中概述的ATAC-seq评估基因工程细胞,无需进一步的再刺激。

[0590] 表E1提供了使用多变量相关性确定的每个标志物的相关性的p值。图1A和1B分别显示了通过ICS测量的干扰素- γ (IFN γ) 和程序性细胞死亡蛋白1 (PD-1) 产生(在x轴上显示)与通过ATAC-seq确定的编码每种蛋白质(分别为Ifng和Pdc1)的基因的可及性(在y轴上显示)的代表性相关性。如表1所示,在染色质可及性程度和许多评估的T细胞标志物的蛋白质表达之间没有统计学上显著的相关性,尽管某些细胞因子标志物的表达确实表现出与染色质可及性的显著相关性。

表E1. ATAC-seq与ICS初始 多变量相关性	
细胞因子	P值
IFN-g	p<0.0001
IL-13	p=0.0002
PD-1	p=0.018
Tbet	p=0.05
CD25	p=0.106
IL-2	p=0.104
Lag3	p=0.171
KI-67	p=0.123
TNF	p=0.249
膜联蛋白 V	p=0.342
Foxp3	p=0.491

[0591] 在另一项研究中,总体上如上所述进行ATAC-seq以评估在编码IFN γ 和IL-2的基因(分别为Ifng和I12)处的染色质可及性。在用佛波醇肉豆蔻酸乙酸酯(PMA)/离子霉素再刺激6小时后,通过ICS评估IFN γ 和IL-2的产生。如图1C所示,在编码IFN γ 的基因处的染色质可及性与通过ICS测量的IFN γ 的产生相关,其中 R^2 值为0.579,斯皮尔曼等级相关(ρ)为0.7536,并且 $\text{prob}>|\rho|$ 为0.0012。如图1D所示,在编码IL-2的基因处的染色质可及性与IL-2的产生相关,其中 R^2 值为0.229,斯皮尔曼等级相关(ρ)为0.5679,并且 $\text{prob}>|\rho|$ 为0.0272。

[0592] 表E1和图1A-1D中的结果与以下发现一致:在一些情况下,给定基因的染色质可及性可以预测和/或关联细胞活化后的基因表达或其他结果。

[0593] ICS和ATAC-seq用于评估CD4+细胞或CD8+细胞中的示例性T细胞标志物的免疫表型(分别为蛋白质表达或染色质可及性),所述CD4+细胞或CD8+细胞获自含有工程化T细胞的CDP样品,所述工程化T细胞已被施用于使用上述方法治疗B细胞恶性肿瘤的自体受试者。ICS和ATAC-seq的结果独立地进一步与受试者对用自体工程化细胞治疗的反应结果(完全缓解(CR)、疾病进展(PD)或部分缓解(PR))相关。确定了代表性基因的细胞因子可及性水平或借助于ICS的蛋白质表达水平,将其显示在层次聚类中并通过反应结果进行注释。作为对照,还通过ICS和ATAC-Seq评估了正常供体(ND)细胞中的T细胞标志物的水平。图2显示了对CD4+CDP样品和CD8+CDP样品的免疫表型分型数据的反应聚类。如图2所示,通过ATAC-seq测量的CD8+CDP样品中的细胞因子产生与反应结果相关。通过ICS未观察到与反应结果的相关性。

实施例3:通过ATAC-seq对染色质可及性的全基因组评估

[0594] 使用ATAC-seq在来自CDP的解冻CD8+细胞上进行全基因组分析,所述CDP含有如上

文实施例1中所述产生的表达抗CD19 CAR的基因工程人T细胞。基于治疗是否产生完全缓解(CR)、疾病进展(PD)或部分缓解(PR),将ATAC-seq的结果独立地进一步与受试者对用自体工程化细胞治疗的反应结果相关联。星号指示在细胞剂量开始之后三个月时从CR转换为PD的受试者。同样通过ATAC-seq评估了正常供体(ND) CD8+细胞的全基因组分析。基于每个基因的染色质可及性差异进行层次聚类,如通过每个基因的基因体上FPKM的总和计算的,其被标度为针对每个基因的低(蓝色)至高(红色)。

[0595] 来自已经被施用自体CDP的9位受试者中的每一位的CDP的细胞中多个基因的染色质可及性的代表性全基因组分析示于图3A中。图3B显示了展示按照缓解组(在3个月时转换为PD的以PR和PD样品聚类的CR患者)在ATAC-seq数据上观察到的CD8+CDP聚类的聚类决策树(星座图)。

[0596] 进一步分析了来自上面的基因的靶向集合的选择子集,将其作为层次聚类呈现在图4A中,并且作为星座图呈现在图4B中。与上述类似,按照缓解组在ATAC-seq数据上观察到所指示的特定基因的CD8+CDP聚类。

实施例4:通过ATAC-seq评估CD8+CAR T细胞中的染色质可及性

[0597] 基于表型将含有从受试者获得但未用抗CD19 CAR工程改造的T细胞的冷冻保存的CMAT样品分离为幼稚T细胞(T_N)、中枢记忆T细胞(T_{CM})、效应细胞与效应记忆T细胞(T_{E+EM})或效应记忆RA(T_{EMRA})。针对基因的六个集合中的每一个的染色质可及性评估来自每个子集的细胞,所述基因代表由Immgen联盟鉴定的来自记忆CD8+ T细胞模块的选定基因子集(Best等人, *Nature Immunology* (2013) 14:404-412)。基于每个基因的染色质可及性差异进行层次聚类,如通过每个基因的基因体上FPKM的总和计算的,其被标度为针对每个基因的低(蓝色)至高(红色)。针对评估的细胞类型的每个基因集合的染色质可及性的特征谱显示在图5A中。结果证明,基因子集的这些集合的染色质可及性指示细胞的表型效应子样状态。

[0598] 使用相同的基因集合来评估来自CDP的解冻CD8+细胞中的染色质可及性,所述CDP含有如上文实施例1中所述产生的表达抗CD19 CAR的基因工程人T细胞。同样通过ATAC-seq评估了正常供体(ND) CD8+细胞。基于治疗是否产生完全缓解(CR)、疾病进展(PD)或部分缓解(PR),将ATAC-seq的结果独立地进一步与受试者对用自体工程化细胞治疗的反应结果相关联。如图5B所示,与显示CR或ND证据的受试者相比,来自显示PR和PD证据的受试者的CD8+CDP显得具有更多的效应子样表型,正如通过与效应子样表型相关的基因的染色质可及性所确定的。

实施例5:通过ATAC-seq评估在选定基因座的CD8+染色质可及性

[0599] 正如在与T细胞的信号强度和效应子功能相关的基因座处的每个基因的基因体上FPKM的总和所计算的,进一步分析了来自CDP的解冻的CD8+细胞的染色质可及性,所述CDP含有如上文实施例1中所述产生的表达抗CD19 CAR的基因工程人T细胞。基于治疗是否产生完全缓解(CR)、疾病进展(PD)或部分缓解(PR),将ATAC-seq的结果独立地进一步与受试者对用自体工程化细胞治疗的反应结果相关联。同样通过ATAC-seq评估了正常供体(ND) CD8+细胞。

[0600] 如图6A和6B所示,除了pdcd1之外,所有测试基因的染色质可及性的较高fpkm水平与发生部分缓解或疾病进展相关。pdcd1基因座的染色质可及性的较低fpkm水平与发生部分缓解或疾病进展相关。跟踪T细胞的活化状态的管家基因Actb和Gapdh在PR和PD中增加,

表明来自CDP样品的施用于继续发生PR或PD的受试者的细胞可能处于更为活化的状态。这些结果证明,这些基因或基因的集合可以是用于预测细胞疗法治疗的反应结果的表现遗传学标志物。

实施例6:在来那度胺存在或不存在下在CAR T细胞中的基因表达和染色质可及性分析

[0601] 评估了在来那度胺存在或不存在下在刺激后的CAR T细胞中的基因表达和染色质可及性。

[0602] 将由四(4)个不同的独立供体生成的表达抗BCMA CAR的T细胞在来那度胺存在或不存在下用50 μ g/mL BCMA缀合的珠子刺激24小时(24hr+stim)或7天(d7+stim),或在无刺激下培养24小时(24hr)。通过RNA测序(RNA-seq)评估表达CAR的细胞的基因表达,并且通过利用测序的转座酶可及的染色质测定法(ATAC-Seq)进行染色质可及性分析。

[0603] 对从RNA制备的互补DNA(cDNA)样品进行RNA-seq,所述RNA分离自培养的表达式抗BCMA CAR的细胞。通常如Buenrostro等人在Nat Methods. (2013) 10(12):1213-1218中所述进行ATAC-seq。使用MACS2调用ATAC-seq可及性峰($q < 0.01$),并使用DiffBind从2个或多个样品中存在的重叠峰生成共有集。

[0604] 对从DESeq2归一化计数生成的RNA-seq和ATAC-seq数据集进行主成分分析(PCA)。计算了差异性表达(DE,对于RNA-seq)或共有峰可及性(DA,对于ATAC-seq),模拟在24小时和第7天的供体效应(供体1-4)和处理效应(来那度胺对载体)。对于RNA-seq,差异性基因座选择截止值为 $q \leq 0.05$ 且 \log_2 倍数变化 ≥ 0.5 ,或对于ATAC-seq, $q \leq 0.1$ 。考虑到在每种处理条件之内的供体效应,进行基因本体论(GO)富集分析,并使用Ingenuity Pathway Analysis软件(Qiagen, Inc.)对在 $q < 0.1$ 下差异表达的基因子集确定活化z评分。对于第7天刺激(d7+stim)的ATAC-seq数据,使用共有峰集作为背景,利用HOMER软件对在来那度胺存在下显示更可及的峰进行基序富集分析。

[0605] 代表基因表达的总体多样性或基因组上的染色质可及性的PCA结果显示在图7A(基因表达;基于RNA-seq结果)和图7B(染色质可及性;基于ATAC-seq结果)中。绘制椭圆来指示这些组,因为观察到导致基因表达或染色质可及性变化的主要因素是培养时间和刺激的存在。与在来那度胺不存在下培养的细胞(三角形,载体)相比,在来那度胺存在下培养的细胞(圆形)表现出不同的总体基因表达和染色质可及性,显示出在各个供体和培养条件下的来那度胺处理效应。对于来那度胺处理,总体变化方向(通过在三角形和圆形之间的虚线显示)在每个供体中是相似的,并且与在有刺激或无刺激下培养24小时的细胞中的变化相比,在刺激下培养7天的细胞中的变化程度通常更高。

[0606] 图8A-8D显示了在来那度胺存在下基因表达(图8A和8B,分别在刺激下培养24小时和7天后)或染色质可及性(图8C和8D,分别在刺激下培养24小时和7天后)的变化。如图所示,与培养24小时相比,来那度胺对基因表达和染色质可及性的影响在培养7天时更大。在来那度胺存在下培养7天后,如基因表达变化所示,总共583个基因被改变(图8B),而在2804个峰处的染色质可及性被改变(图8D)。这些结果表明来那度胺处理改变了CAR-T细胞的转录谱和表现遗传学谱两者。

[0607] 鉴定了被差异性地富集表达的基因的生物信号传导途径(图9A和9B)。显示了在24小时(图9A)或7天(图9B)时对生物途径的影响的方向性和显著性。结果显示,来那度胺的存在导致涉及T细胞活化和信号传导的基因的表达增加。结果显示,在来那度胺存在和不存在

下受到差异性调节的途径显示出免疫突触相关基因、涉及细胞因子信号传导的基因和涉及T细胞活化途径的基因的富集。

[0608] 对于选定的基因子集,包括涉及T细胞活化和信号传导的基因在内,对在刺激下培养7天的细胞比较了在来那度胺存在下的基因表达和染色质可及性变化。图10显示了针对通过RNA-seq测量的相应基因表达变化绘制的各个染色质可及性峰(菱形)和每个基因的平均染色质可及性变化(圆形),显示了在两种方法之间的信号的一致性。

[0609] 在来那度胺存在下培养7天时具有增加的可及性的峰的基序富集分析的结果显示在图11中。在来那度胺存在下,预计结合各种转录因子的基序(被理解为涉及T细胞活化和信号传导)在具有增加的可及性的峰中富集。

[0610] 所述结果与在来那度胺存在下在表达CAR的T细胞中的功能活性的增强一致。

实施例7:在细胞工程不同阶段的细胞的染色质可及性谱分析

[0611] 评估了在示例性免疫基因(例如细胞因子、趋化因子、细胞表面标志物)处或其附近的染色质可及性,并在来自用实施例1中描述的示例性抗CD19 CAR进行基因工程之前和之后的细胞组合物的样品之间进行了比较。基于冷冻保存的CD4⁺或CD8⁺工程化细胞组合物(CDP)或未经工程改造的匹配样品(CMAT),将样品解冻。在一些情况下,通过表型将CMAT样品分离为幼稚T细胞(T_N)、中枢记忆T细胞(T_{CM})、效应与效应记忆T细胞(T_{E+EM})或效应记忆RA(T_{EMRA})用于分析。基本上如实施例1中所述生成文库。

[0612] 例如通过使用bowtie(Langmead等人,(2009) Genome Biology 10:R25.1-R25.10)或bowtie2将ATAC-seq读段与参考基因组比对并映射回参考基因组以确定它们的位置,并且进行分析。进行了另外的处理步骤,包括去除重复,过滤出线粒体DNA,在ATAC-seq文库制备过程中移动映射片段的位置以便通过Tn5转座酶导致4个或5个碱基对插入,并过滤出大于100个碱基对的片段(bp),在一些情况下,这些大于100个碱基对的片段片段可以代表核小体结合的染色质而不是无核小体的染色质。还评估了核小体定位(例如使用NucleoATAC)、转座子插入位点和转录因子占据。使用MACS2进行可及性峰的鉴定,所述可及性峰包括通过量化ATAC-seq片段测量的富集或耗竭可及性和/或占据信号的基因组区域。

[0613] 细胞表面标志物基因的示例性谱显示在12A和12B中。如所示,染色质可及性峰与特定的表面标志物的表达相关(例如,在编码CD8⁺细胞中的CD3 ϵ 、CD8a和CD8b的基因处或其附近存在可及性峰,并且在编码CD4⁺细胞中的CD3 ϵ 和CD4的基因处或其附近存在可及性峰)。观察到在匹配的CDP和CMAT细胞样品中的其他免疫基因处或其附近的染色质可及性峰的一些差异,这与制造过程中细胞的状态和/或表型的变化一致。

实施例8:细胞亚群中峰谱的基因组区间分析

[0614] 基于表面标志物表达和表型,将如实施例1中所述的来自相同供体的CD8⁺CDP细胞和CD8⁺CMAT细胞分离成如下亚群:CD27⁺CCR7⁺、CD27⁺CCR7⁻、CD27⁻CCR7⁻、幼稚T细胞(T_N)、中枢记忆T细胞(T_{CM})、效应细胞与效应记忆T细胞(T_{E+EM})和效应记忆RA(T_{EMRA})。总体上如实施例7中所述对亚群进行ATAC-seq。通过评估细胞亚群之间和之中的共同或重叠峰和独特峰,确定可及性峰并进行基因组区间分析。

[0615] 对CDP和CMAT样品的不同亚群中的共同和独特峰的分析显示,CDP样品含有更多独特的可及性峰,表明已经受刺激和基因工程的CDP样品中的染色质可及性增加。大多数CMAT可及性峰与CDP样品峰相同。CD27⁺CCR7⁺CMAT样品的可及性谱显示与T_N、T_{CM}、T_{E+EM}和T_{EMRA}样品

的可及性谱有很大重叠,而CD27+CCR7+CDP含有更多独特的可及性峰。

[0616] 比较了在CD27+CCR7+、CD27+CCR7-和CD27-CCR7-细胞以及大量CD8+细胞中的在编码区和CCR7基因附近的基因间区中的峰谱。如图13A所示,CCR7+细胞在CCR7基因的编码区上显示出与CCR7-细胞相似的峰谱,但在CCR7+细胞中的编码区上游的基因间区中观察到峰,指示了与CCR7在细胞中的表达相关的上游增强子位点。

[0617] 将CD27+CCR7+CDP细胞、CD27+CCR7-CDP细胞和CD27-CCR7-CDP细胞的总峰谱进行基因可及性分析。对于CD27+CCR7+CDP群体,观察到最大数目的独特峰。图13B显示了在细胞群体的各个基因组位置(包括基因的基因间区、内含子区和启动子区)内的可及性峰的分布。

[0618] 这些结果与以下结论一致:在基因编码区周围而不仅仅是在基因编码序列内的可及性可以提供关于细胞状态的信息。

实施例9:基因组可及性和峰谱以及在施用工程化T细胞的受试者中的反应结果

[0619] 评估了在CDP细胞组合物中的基因组可及性与峰谱之间的关系以及已被施用工程化T细胞的受试者中的反应结果。总体上如上文实施例1和3中所述,对于已被施用自体工程化细胞组合物的受试者,通过ATAC-seq确定了CD4+和CD8+CDP细胞可及性峰谱。

[0620] 评估了CD4+和CD8+CDP细胞中的基因组可及性峰和无核小体区的总数以及反应结果(1个月CR、3个月CR、1个月PD、3个月PD或PR)。使用MACS2调用ATAC-seq可及性峰,并使用NucleoATAC确定了无核小体区,如实施例7中所述。如图14A所示,在CD8+CDP细胞中,与表现出3个月CR的受试者相比,在表现出PD或PR的受试者中,观察到总基因组可及性峰和无核小体区的数目是更高的,表明更大的全基因组可及性。测量峰无核小体区产生了与测量染色质可及性峰值的分析相似的结果。在样品之间在具有不同反应结果的受试者中的两个示例性免疫相关基因(基因1和基因2)附近的基因组区域展示不同的可及性峰的示例性差异可及性峰谱显示在图14B-14D中。如所示,与在3个月时表现出CR的受试者相比,在3个月时表现出PD的受试者中的示例性基因座附近的可及性是更高的。

实施例10:评估编码嵌合抗原受体的载体的整合

[0621] 使用ATAC-seq来评估与编码嵌合抗原受体(CAR)的病毒载体在工程化细胞中的整合相关的各种参数。

[0622] 通常如实施例1和7中所述对CDP细胞组合物进行ATAC-seq,所述CDP细胞组合物通过用编码抗CD19 CAR的病毒载体(CAR整合体)转导而被工程化以表达抗CD19 CAR。将用空病毒载体(空整合体)转导的CMAT细胞组合物和细胞用作对照。通过在比对步骤过程中将构建体作为人工染色体处理,将来自ATAC-seq读段的序列与编码CAR的核酸序列比对。整合体(ATAC-seq CAR整合体)的比例数被计算为(比对读段x读长)/(构建体大小)。通过将真阳性率与假阳性率作图而生成接受者操作特征(ROC)曲线,并确定曲线下面积(AUC)。

[0623] 如图15A和15B所示,分析表明编码CAR的核酸序列仅整合在工程化CDP细胞中而不整合在CMAT细胞中,并且整合到基因组中如ATACseq信号指示的可及区中。CAR整合体的AUC(95%置信区间)为1,其中CI低二项式为0.61847,CI高二项式为0.6397;空整合体的AUC为0.9766,其中CI低二项式为0.056865,CI高二项式为0.63939。所述结果与ATAC-seq方法用于评估重组载体(例如病毒载体,比如编码CAR的载体)整合到工程化细胞基因组中的效用一致。

[0624] 在另一项研究中,将在CMAT和抗CD19 CAR T细胞CDP组合中通过使用ATAC-seq确定的整合体数和通过定量聚合酶链反应(qPCR)确定的载体拷贝数(VCN)进行了比较。如图16A所示,ATAC-seq读段映射到CDP细胞中的CAR编码序列,但在CMAT细胞中则不然。通过ATAC-seq确定的整合体与通过qPCR确定的VCN的比例数之间的相关性显示在图16B中(非参数斯皮尔曼 ρ :0.3121,概率 $>|\rho|$:0.2073)。

[0625] 在来自已被施用自体抗CD19 CAR+CDP工程化细胞组合物的受试者的抗CD19 CAR+CD4+和CD8+ T细胞中,也评估了借助于ATAC-seq的整合体和VCN。如图17A和17B所示(不包括正常供体样品),在来自实现PD或PR的受试者的CD8+CDP细胞中观察到更高的整合体数,这与以下观察结果一致,即,CAR构建体在某些个体中的表达可能取决于染色质可及性而不同,即使当总VCN可能相似时也如此。

[0626] 通过映射CD4+和CD8+CDP细胞中具有未知克隆性的50,000个细胞的不一致读段对,评估了独特整合位点,所述CD4+和CD8+CDP细胞来自实现不同的反应结果(1个月CR、3个月CR、1个月PD、3个月PD或PR)的受试者。结果显示在图18中。所述结果与以下观察结果一致,即,在具有相对相似的VCN的不同受试者中,CAR构建体的表达可由于表观遗传学状态和细胞的健康而变化,导致不同的反应或持久的反应结果。

实施例11:T细胞受体(TCR)基因座的可及性

[0627] 评估了跨越不同的抗CD19 CAR+ T细胞CDP T细胞组合物和CMAT T细胞组合物中或来自已经接受抗CD19 CAR+ T细胞施用的受试者的CD8+抗CD19 CAR+ T细胞中的编码T细胞受体(TCR)链的基因座的染色质可及性峰谱。

[0628] 如图19A所示,观察到编码TCR β 链的基因座上的可及性峰在来自7位示例性受试者的CD8+CMAT样品中是不同的。在ND或实现CR或PD的受试者中的CD8+CDP样品中的总TCR可及性和变异系数(CV)%显示在图19B中。如图19C所示,实现不同的反应结果的受试者显示出跨越TCR基因座的可变的相对TCR可及性,其中来自PR或PD受试者的样品是更加寡克隆的。所述结果与ATAC-seq分析用于评估TCR基因座区域的可及性差异并且通常用于评估细胞组合物内的T细胞的克隆性的效用一致。

实施例12:来自细胞工程不同阶段的不同供体的细胞中的染色质可及性谱及其他分析

[0629] 使用ATAC-seq评估来自三位不同的健康供体的CD4+/CD8+ T细胞组合物在用不同的嵌合抗原受体进行基因工程之前和之后的染色质可及性,并且进行另外的下游分析,包括主成分分析(PCA)、差异性可及性分析、生物途径分析和在选定基因子集处的分析。

[0630] 根据PBMC的白细胞单采术,通过基于免疫亲和力的富集从三(3)位健康供体(供体1、2和3)分离CD4+和/或CD8+ T细胞,并冷冻保存(CMAT,在工程化之前)。将分离的CD4+/CD8+ T细胞活化24小时,并用编码两种不同的抗CD19 CAR之一或抗BCMA CAR的病毒载体转导,或模拟转导作为对照,并冷冻保存(CDP,冷冻保存的工程化细胞组合物)。总体上如上文实施例1中所述对细胞进行ATAC-seq分析。

[0631] 映射获得的序列并使用MACS2调用ATAC-seq可及性峰。用DESeq2包处理原始测序计数,以估计尺寸因子,估计离差,并执行负二项广义线性模型拟合。用设置为真的betaPrior提取峰中读段的分数(FRiP;显示信号富集,计算为(峰中读段数)/(总读段数))归一化的计数。在将原始数据解复用、比对并针对如上文的质量规格过滤之后,将它们输入,并且在R中完成以下分析步骤。用ChIPpeakAnno包或Homer软件注释峰。转录起始位点

(TSS)区被定义为启动子的上游近端2000bp到下游500bp。用DiffBind包进行峰重叠和共有峰的组分析。计算重叠后,过滤在多于两个的文库中发现的峰,并提取测序计数。使用各种测序质量控制度量来分析获得的序列和峰,所述度量包括未映射的未配对的和重复的读段、线粒体DNA的部分、序列深度、射到编码CAR的构建体的读段、具有0.1或更小的错误发现率(FDR)的MACS2峰的数目、FRiP和独特峰的数目,以便确保数据的一致性和保真度。

[0632] 基于各种样品的总表观遗传学谱进行聚类分析。结果表明,样品的聚类倾向于首先基于样品的类型(例如CMAT或CDP),然后基于供体,并且(对于CDP)最后基于构建体的类型。

[0633] 针对维数约减进行主成分分析(PCA),以检查数据中的总体方差,并解析共有峰的关键驱动因素,所述共有峰例如在2个或更多个样品中、在来自不同供体的CMAT和各种CDP样品上存在的峰(PC1,导致33.996%的变异;与CDP和CMAT之间的差异相关;PC2,导致10.13%的变异;与T细胞状态相关)。在这个步骤中计算了31个主成分。所述结果显示了来自不同供体的三种不同CMAT的模式,其在工程化后之产生与上述聚类数据一致的三种不同的CDP。用于工程的各种不同的CAR构建体产生相似的PCA分析谱。

[0634] 使用以DESeq2包构建的模型进行差异性可及性分析。提取差异性结果,并将显著性峰指定为通过0.1的错误发现率(FDR或 q)的那些峰。来自所有供体的CDP和CMAT的示例性差异性可及性分析的结果显示在图20A中。使用Ingenuity Canonical Pathway、Biofunctions和Predicted Upstream Regulators分析,基于基因集富集分析(Gene Set Enrichment Analysis)(GSEA),使用差异性峰进行生物途径分析。所述结果显示在涉及T细胞应答的各种途径、与免疫细胞活性和/或功能相关的信号传导途径、与免疫细胞相关的生物功能和与免疫功能相关的基因的预测的上游调控因子中的差异性峰的富集。

[0635] 使用以下选择的基因模块进行基因模块分析:在“T细胞模块”中的基因(Chaussabel等人,(2008),Immunity 29(1):150-164);在“细胞因子”模块中的基因(Immgen Consortium(Best等人,Nature Immunology(2013)14:404-412));在Th1.细胞模块、CD8pos. T_{EM} 模块和CD4pos. T_{EM} 模块中的基因(xCell(Aran等人Genome Biology(2017)18:220;<http://xcell.ucsf.edu/>));“记忆”模块(Weng等人,(2012)Nat Rev Immunol.12(4):306-15)、包括趋化因子受体基因的“运输”模块和“衰竭”模块(Martinez等人,(2015)Immunity42(2):265-278)。FRiP归一化的计数或FPKM计数用于基因模块分析。在FRiP归一化计数的情况下,将启动子处的峰用作基因水平度量的代用指标。将计数进行 \log_2 变换并以图形方式与色标对齐。用JMP软件或Pheatmap R包执行层次聚类。

[0636] 基于每个基因的启动子区的归一化可及性计数度量(启动子可及性)的聚类分析显示了与CDP样品分开的CMAT样品的聚类,并且CMAT和CDP谱的分析显示工程化过程似乎降低了记忆T细胞(T_{MEM})和效应细胞特征中的可检测变异。结果还表明,可以从不同的供体产生具有不同表观遗传状态的各种工程化细胞组合物(CDP)。

[0637] 针对“T细胞模块”、“记忆T细胞模块”和“运输模块”基因的CDP样品的聚类分析显示,来自不同供体的CDP通常按区分的供体聚集在一起。在“细胞因子”模块和“衰竭”模块中的基因的示例性个体启动子处的启动子可及性的结果显示在图20B(“细胞因子”)和20C(“衰竭”)中。

[0638] 所述结果与启动子可及性作为例如在制造过程中区分T细胞的不同状态的度量的

效用一致。

实施例13:在施用工程化T细胞的受试者中的染色质可及性分析和治疗结果

[0639] 评估了来自临床研究中的已被施用表达嵌合抗原受体 (CAR) 的自体工程化细胞的受试者的CMAT和CDP样品中的染色质可及性谱。确定并比较了在实现反应结果的受试者中或在发生毒性的受试者中的可及性谱。

A. 受试者和细胞组合物

[0640] 患有复发或难治性 (R/R) 侵袭性非霍奇金淋巴瘤 (NHL) 的成年人受试者被施用表达抗CD19 CAR的自体工程化CD4+和/或CD8+ T细胞。为了生成工程化细胞,根据PBMC的白细胞单采术,通过基于免疫亲和力的富集分离CD4+/CD8+ T细胞,并冷冻保存 (CMAT,在工程化之前)。将分离的CD4+/CD8+T细胞活化并且用编码抗CD19 CAR的病毒载体转导,所述抗CD19 CAR含有抗CD19 scFv、免疫球蛋白衍生的间隔物、衍生自CD28的跨膜结构域、衍生自4-1BB的共刺激区和CD3 ζ 细胞内信号传导结构域。所述病毒载体进一步含有编码截短的受体的序列,所述截短的受体用作CAR表达的替代标志物;所述序列通过T2A核糖体跳跃序列与CAR序列分开。将得到的工程化细胞冷冻保存 (CDP,冷冻保存的工程化细胞组合物)。

B. ATAC-seq和质量度量

[0641] 总体上如实施例1中所述,通过ATAC-seq测定来自24位受试者的总共82个CD4+或CD8+CMAT或CDP样品。映射获得的序列并使用MACS2调用ATAC-seq可及性峰。使用各种测序质量控制度量来分析获得的序列和峰,所述度量包括未映射的未配对的和重复的读段、线粒体DNA的部分、有效序列深度、具有0.1或更小的错误发现率 (FDR) 的MACS2峰的数目、峰中读段的分数 (FRiP) 和独特峰的数目,以便确保数据的一致性和保真度。由于低富集和/或数据保真度,排除了七 (7) 个样品。

[0642] 在一些情况下,获得并测定了技术重复样品。针对技术重复评估了区间分析、PCA、峰重叠和峰中的Log2归一化计数,并计算Spearman相关性。结果表明,这些重复具有高度相似的谱,具有较大的重叠和较高的相关系数。

C. CD4+和CD8+CDP和CMAT的分析

[0643] 基于CD4+和CD8+细胞群中每个基因座的染色质可及性,在CDP和CMAT细胞群 (作为细胞类型的确认) 中的CD4和CD8A启动子处评估了每个基因的启动子区的归一化可及性计数度量 (启动子可及性)。确定CD8A启动子的可及性与CD4启动子的可及性的比率。如图21所示,CD8A:CD4可及性比率反映了CDP和CMAT样品两者的每种组合物的CD4+或CD8+状态。

[0644] 对所有样品的共有峰 (144,591个峰) 进行主成分分析 (PCA)。结果显示,在PCA图上的不同聚类中,通常CMAT CD4+样品聚集在一起,并且CMAT CD8+细胞聚集在一起。CD4+CDP样品和CD8+CDP样品倾向于聚集在同一组内,其中CD4+CDP样品和CD8+CDP样品基于CD4或CD8表达形成亚簇。结果表明,工程化过程导致CD4+和CD8+细胞总谱的变异减小。相比之下,在工程化之前,CD4+和CD8+样品显示出变化更大的谱。

[0645] 使用总共有峰集 (CDP为122,495个峰,CMAT为106,867个峰) 的聚类分析在CD4+和CD8+CDP样品中进行。结果显示,对于CMAT和CDP样品两者,CD4+或CD8+细胞类型通常聚集在一起。

D. 与临床结果相关的分析

[0646] 在来自在施用工程化细胞组合物之后继续实现不同的反应结果或毒性结果的受

试者的样品中,对具有不同可及性的峰进行差异性可及性分析。

[0647] 图22A显示了在与实现完全缓解 (CR) 的受试者相比具有疾病进展 (PD) 的受试者中具有更高或更低可及性的峰的log2倍数变化和调整的p值,所述完全缓解为最佳总缓解 (BOR)、3个月持续缓解 (3MO) 或6个月持续缓解 (6MO),其中相应数目的峰差异性地存在于样品中。在3个月和6个月的反应分析中存在大量的差异性可及峰,表明实现3或6个月CR的受试者的细胞组合物中的表观遗传学状态与具有PD的受试者相比是不同的。峰重叠分析显示,在6个月反应分析中显示差异可及的许多峰与3个月反应分析中的峰重叠,并且在BOR分析中的一些峰也与3个月反应分析中的峰重叠。

[0648] 图22B显示了在与具有0-2Ntx等级的受试者相比已发生3-5级神经毒性 (Ntx)、或者与具有0-1级CRS的受试者相比具有2-5级细胞因子释放综合征的受试者中具有更高或更低可及性的峰的log2倍数变化和调整的p值,其中相应数目的峰差异性地存在于样品中。重叠分析显示,只有一些峰在Ntx和CRS分析之间重叠。结果显示在针对Ntx和CRS两者的组之间的大量差异性可及性峰。

E. 与临床结果相关的CMAT和CDP分析

[0649] 对共有峰集的归一化计数的PCA、以及来自实现不同反应结果或毒性结果的受试者的CDP或CMAT样品中的差异性可及性分析总体上如上文所述。针对CDP样品 (共有集:112,495个峰)的PCA显示,CD4+和CD8+样品通常基于细胞类型 (CD4+或CD8+细胞) 而聚集。CMAT样品 (共有集:106,867个峰)的PCA也表明,样品通常基于细胞类型 (CD4+或CD8+) 而聚集,但是这些样品不像CDP样品那样紧密聚集,表明与CDP样品内的变异性相比,在不同的CMAT样品内的变异性是更大的。

[0650] 针对BOR、3个月和6个月缓解,分别在CDP样品和CMAT样品中进行反应结果的差异性可及性分析。如图23A (CDP) 和23B (CMAT) 所示,对于CDP样品中的3个月和6个月反应分析,大量的峰是差异性地可及的,但在CMAT样品中很少的峰是差异性地可及的。重叠分析显示,对于CDP,许多差异性可及的峰在3个月和6个月反应分析中重叠,但每个峰与BOR峰的重叠是低的,而对于CMAT,存在和重叠的峰很少。

[0651] 针对Ntx和CRS的差异性可及性分析分别在CDP和CMAT样品中进行。如图23C (CDP) 所示,对于Ntx,很少的峰显示是差异性可及的,而对于具有2-5级CRS的受试者,更多数目的峰是差异性地可及的。如图23D (CMAT) 所示,对于Ntx,观察到大量的峰是差异性地可及的,但对于CRS而言很少是差异性地可及的。针对CDP和CMAT样品两者的重叠分析显示在Ntx和CRS分析之间的重叠峰很少。

[0652] 结果显示,单独检查CDP降低了模型用于分配与临床结果相关的差异性可及性峰的总效能,并且仅检查CMAT降低了模型用于分配与反应和CRS相关的差异性可及性峰的效能,这与以下观察结果一致:即,询问CDP状态可提供关于CRS的潜在发生的信息。对于Ntx,检查CMAT保留了大量的差异性可及性峰,这与以下观察结果一致:即,起始T细胞和疾病状态可提供关于神经毒性的潜在发生的信息。

[0653] 结果证明了ATACseq分析的效用:其用于鉴定可能与临床结果 (例如反应或安全性结果) 相关的细胞类型的潜在差异,以及鉴定与临床结果相关的表观遗传学特性或特征。

实施例14:改良的用于CAR T细胞分析中的染色质可及性分析的ATAC-seq方法

[0654] 使用改良的ATAC-seq方法来评估用嵌合抗原受体进行基因工程改造的CD8+ T细

胞中的染色质可及性谱,并与总体上在实施例1中描述的ATAC-seq方法进行比较。

[0655] 对来自患有NHL的受试者的CD8+CDP样品和来自健康供体的和CD8+CMAT样品进行总体上如Corces等人(2017) Nature Methods 14:959-962所述的改良的ATAC-seq分析,或进行总体上如实施例1中所述的ATAC-seq方法(标准ATAC-seq)。改良包括向洗涤缓冲液中添加PBS,向裂解缓冲液中添加吐温-20和洋地黄皂苷,以及向转座反应添加PBS、吐温-20和洋地黄皂苷,并且使用不同的DNA纯化柱。

[0656] 映射获得的序列并使用MACS2调用ATAC-seq可及性峰。使用各种测序质量控制度量来分析从改良或标准的ATAC-seq获得的序列和峰,所述度量包括未映射的未配对的和重复的读段、线粒体DNA的部分、有效序列深度、具有0.1或更小的错误发现率(FDR)的MACS2峰的数量、峰中读段的分数(FRiP)和独特峰的数量,以确保数据的一致性和保真度。在使用改良的ATACseq获得的一些样品中,线粒体DNA部分是较低的。

[0657] 图24A显示了在CD8+CDP和CMAT样品中具有FDR<0.1的鉴定峰的数量,并且图24B显示了以FRiPs指示的富集,其中使用改良或标准的ATAC-seq,具有3个技术重复。结果显示,使用改良的ATAC-seq的鉴定峰的数量是更高的,并且使用改良的ATACseq制备的样品中的富集信号是显著更高的。

[0658] 峰重叠分析显示,使用标准和改良的ATACseq方法两者,在样品的技术重复之间的峰重叠都是高的。在标准和改良的ATACseq峰之间的重叠显示,几乎所有标准ATACseq峰都存在于来自改良的ATACseq的峰中,其中在使用改良的ATACseq处理的样品中存在更多的峰。在技术重复之间和跨越不同方法的样品间相关性是高的。差异性可及性分析、基因组浏览器可视化、使用与T细胞相关的基因模块的聚类分析、衰竭、细胞因子和短期效应子与记忆模块、以及基因集富集分析(GSEA)显示了在使用标准和改良的ATACseq方法获得的数据之间的一致结果。

[0659] 结果显示,改良的ATACseq方法生成的数据具有降低的线粒体读段部分、增加的复杂性和效率、富集的信号和增加的跨越样品的技术再现性。在来自标准和改良的ATACseq的结果的差异性分析和生物途径分析中观察到一致的模式。

[0660] 本发明并不意图限于特定公开的实施方案的范围,所提供的实施方案例如用于说明本发明的各个方面。根据本文中的描述和教导,对所述组合物和方法的各种修改将变得显而易见。可以在不脱离本公开的真实范围和精神的情况下实施这样的变化,并且这些变化预期落入本公开的范围。

序列

#	序列	注释
1	ESKYGPPCPPCP	间隔物(IgG4 铰链)(aa)
2	GAATCTAAGTACGGACCGCCCTGCCCCCTTGCCCT	间隔物(IgG4 铰链)(nt)
3	ESKYGPPCPPCPGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMEALHNHYTQKSLSLGLGK	铰链-CH3 间隔物
4	ESKYGPPCPPCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMEALHNHYTQKSLSLGLGK	铰链-CH2-CH3 间隔物
5	RWPESPKAQASSVPTAQPQAEGSLAKATTAPATTRNTGRGGEEKKKEKEKEEQEERETKTPECPSHTQPLGVYLLTPAVQDLWLRDKATFTCFVVGSDLKDAHLTWEVAGKVPTGGVEEGLLERHSNGSQSQHSRLTLPRSLWNAGTSVTCTLNHPSLPPQRLMALREPAAQAPVKLSLNLASSDPPEAASWLLCEVSGFSPPNILLMWLEDQREVNTSGFAPARPPPQPGSTTFWAWSVLRVPAPPSPQPATYTCVVSHEDSRTLLNASRSLEVSIVTDH	IgD-铰链-Fc
6	LEGGGEGRGSLLTCGDVEENPGPR	T2A
7	MLLLVTSLLLCELPHPAFLIPRKVCNGIGIGEFKDSLSINATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSPEGCWGPEPR	tEGFR

#	序列	注释
	DCVSCRNVSRGRECVDKCNLLEGEPRFVENSECIQCHPEC LPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGE NNTLVWKYADAGHVCHLCHPNCTYGCTGPGLEGCPNTP KIPSIATGMVVGALLLLLVALGIGLFM	
8	FWVLVVVGGVLACYSLLVTVAFIIFWV	CD28(登 录 号 P10747 的 氨 基 酸 153-179)
9	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPFPGPSKP FWVLVVVGGVLACYSLLVTVAFIIFWV	CD28(登 录 号 P10747 的 氨 基 酸 114-179)
10	RSKRSRLLHSDYMNMTPRRPGPTRKHYPYAPPRDFAAYR S	CD28(P10747 的 氨 基 酸 180-220)
11	RSKRSRGGHSDYMNMTPRRPGPTRKHYPYAPPRDFAAY RS	CD28(LL 到 GG)
12	KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCE L	4-1BB(Q07011.1 的 氨 基 酸 214-255)
13	RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRR GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKGHDGLYQGLSTATKDTYDALHMQUALPPR	CD3 ζ
14	RVKFSRSAEPPAYQQGQNQLYNELNLGRREEYDVLDKRR GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKGHDGLYQGLSTATKDTYDALHMQUALPPR	CD3 ζ
15	RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRR GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKGHDGLYQGLSTATKDTYDALHMQUALPPR	CD3 ζ
16	RKVCNGIGIGEFKDSLSINATNIKHFKNCTSISGDLHILPVAF RGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLH AFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGD VIISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENSCKA	tEGFR

#	序列	注释
	TGQVCHALCSPEGCWGPEPRDCVSCRNVSRGRECVDKCN LLEGEPREFVENSECIQCHPECLPQAMNITCTGRGPDNCIQC AHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLC HPNCTYGCTGPGLEGCP TNGPKIPSIATGMV GALLLLL VVA LGIGLFM	
17	EGRGSLTTCGDVEENPGP	T2A
18	GSGATNFSLLKQAGDVEENPGP	P2A
19	ATNFSLLKQAGDVEENPGP	P2A
20	QCTNYALLKLAGDVESNPGP	E2A
21	VKQTLNFDLLKLAGDVESNPGP	F2A
22	-PGGG-(SGGGG)5-P-, 其中 P 是脯氨酸, G 是甘氨酸, 并且 S 是丝氨酸	接头
23	GSADDAKKDAAKKGKS	接头
24	atgcttctcctggtgacaagccttctgctctgtgagttaccacaccagcattcctcctgatcca	GMCSFR α 链信 号序列
25	MLLLVTSLLLCELPHPAFLIP	GMCSFR α 链信 号序列
26	MALPVTALLLPLALLHA	CD8 α 信号肽
27	Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro	铰链
28	Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro	铰链
29	ELKTPLGDTHTCPRCPEPKSCDTPPPCPRCPEPKSCDTPPPC PRCPEPKSCDTPPPCPRCP	铰链
30	Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro	铰链
31	Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro	铰链
32	Tyr Gly Pro Pro Cys Pro Pro Cys Pro	铰链
33	Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro	铰链
34	Glu Val Val Val Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro	铰链
35	RASQDISKYLN	FMC63 CDR L1
36	SRLHSGV	FMC63 CDR L2

#	序列	注释
37	GNTLPYTFG	FMC63 CDR L3
38	DYGVS	FMC63 CDR H1
39	VIWGSETTYNSALKS	FMC63 CDR H2
40	YAMDYWG	FMC63 CDR H3
41	EVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPP RKGLEWLGVWGSETTYNSALKSRLTIKDNSKSKVFLK MNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSVTVS S	FMC63 VH
42	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPD GTVKLLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDI ATYFCQQGNTLPYTFGGGKLEIT	FMC63 VL
43	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPD GTVKLLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDI ATYFCQQGNTLPYTFGGGKLEITGSTSGSGKPGSGEGSTK GEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQP PRKGLEWLGVWGSETTYNSALKSRLTIKDNSKSKVFLK MNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSVTVS S	FMC63 scFv
44	KASQNVGTNVA	SJ25C1 CDR L1
45	SATYRNS	SJ25C1 CDR L2
46	QQYNRYPYT	SJ25C1 CDR L3
47	SYWMN	SJ25C1 CDR H1
48	QIYPGDGDTNYNGKFKG	SJ25C1 CDR H2
49	KTISSVDFYFDY	SJ25C1 CDR H3
50	EVKLQQSGAELVRPGSSVKISCKASGYAFSSYWMNWVKQ RPGQGLEWIGQIYPGDGDTNYNGKFKGQATLTADKSSSTA YMQLSGLTSEDSAVYFCARKTISSVDFYFDYWGQGTTVT VSS	SJ25C1 VH
51	DIELTQSPKFMSTSVGDRVSVTCKASQNVGTNVAWYQQK	SJ25C1 VL

#	序列	注释
	PGQSPKPLIYSATYRNSGVPDRFTGSGSGTDFTLTITNVQSK DLADYFCQQYNRYPYTSGGGTKLEIKR	
52	GGGSGGGSGGGGS	接头
53	EVKLQQSGAELVRPGSSVKISCKASGYAFSSYWMNWVKQ RPGQGLEWIGQIYPGDGDTNYNGKFKGQATLTADKSSSTA YMQLSGLTSEDSAVYFCARKTISSVDFYFDYWGQGTTVT VSSGGGSGGGGSGGGSDIELTQSPKFMSTSVGDRVSVT CKASQNVGTNVAWYQQKPGQSPKPLIYSATYRNSGVPDRF TGSGSGTDFTLTITNVQSKDLADYFCQQYNRYPYTSGGGT KLEIKR	SJ25C1 scFv
54	HYYYGGSYAMDY	FMC63 HC-CDR3
55	HTSRLHS	FMC63 LC-CDR2
56	QQGNTLPYT	FMC63 LC-CDR3
57	gacatccagatgaccagaccacctccagcctgagcgcaccctggcgaccgggtgacc atcagctgccgggcccagccaggacatcagcaagtacctaactggtatcagcagaagcccg acggcaccgtcaagctgctgatctaccacaccagccggctgcacagcggcgtgccagcc ggtttagcggcagcggclccggcaccgaactacagcctgaccatctccaacctggaacagga agatatagccactactttgcccagcagggaacacactgccctacacctttggcggcggaac aaagctggaatcaccggcagcacctccggcagcggcaagcctggcagcggcgaggga gcaccaaggcgaggtgaagctgcaggaaagcggcctggcctggtggccccagccag agcctgagcgtgacctgcaccgtgagcggcgtgagcctgcccgaactacggcgtgagctgg atccggcagccccaggaaggcctggaatggctggcgtgatctggggcagcagac cacctactacaacagcgcctgaagagccggctgacctcatcaaggacaacagcaagag ccaggtgttctgaagatgaacagcctgcagaccgacgacaccgccatctactactgcgcca agcactactactacggcggcagctacgccatggactactggggccagggcaccagcgtga ccgtgagcagc	scFv 编码序列
58	X ₁ PPX ₂ P	铰链

#	序列	注释
	X ₁ 是甘氨酸、半胱氨酸或精氨酸 X ₂ 是半胱氨酸或苏氨酸	
59	GSTSGSGKPGSGEGSTKG	接头

序列表

<110> 朱诺治疗学股份有限公司

BONYHADI, Mark L.

KUGLER, David

<120> 细胞疗法的表观遗传学及相关方法

<130> 735042009440

<140> 尚未指定

<141> 与此同时提交

<150> 62/444,802

<151> 2017-01-10

<150> 62/551,752

<151> 2017-08-29

<150> 62/596,662

<151> 2017-12-08

<160> 59

<170> 针对Windows版本4.0的FastSEQ

<210> 1

<211> 12

<212> PRT

<213> 智人

<220>

<223> 间隔物 (IgG4铰链)

<400> 1

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro

1

5

10

<210> 2

<211> 36

<212> DNA

<213> 智人

<220>

<223> 间隔物 (IgG4铰链)

<400> 2

gaatctaagt acggaccgcc ctgccccct tgccect 36

<210> 3

<211> 119

<212> PRT

<213> 智人

<220>

Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
 115 120 125
 Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
 130 135 140
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 145 150 155 160
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 165 170 175
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 180 185 190
 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 195 200 205
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 210 215 220
 Leu Ser Leu Gly Lys
 225
 <210> 5
 <211> 282
 <212> PRT
 <213> 智人
 <220>
 <223> IgD-铰链-Fc
 <400> 5
 Arg Trp Pro Glu Ser Pro Lys Ala Gln Ala Ser Ser Val Pro Thr Ala
 1 5 10 15
 Gln Pro Gln Ala Glu Gly Ser Leu Ala Lys Ala Thr Thr Ala Pro Ala
 20 25 30
 Thr Thr Arg Asn Thr Gly Arg Gly Gly Glu Glu Lys Lys Lys Glu Lys
 35 40 45
 Glu Lys Glu Glu Gln Glu Glu Arg Glu Thr Lys Thr Pro Glu Cys Pro
 50 55 60
 Ser His Thr Gln Pro Leu Gly Val Tyr Leu Leu Thr Pro Ala Val Gln
 65 70 75 80
 Asp Leu Trp Leu Arg Asp Lys Ala Thr Phe Thr Cys Phe Val Val Gly
 85 90 95
 Ser Asp Leu Lys Asp Ala His Leu Thr Trp Glu Val Ala Gly Lys Val
 100 105 110
 Pro Thr Gly Gly Val Glu Glu Gly Leu Leu Glu Arg His Ser Asn Gly
 115 120 125

Ser Gln Ser Gln His Ser Arg Leu Thr Leu Pro Arg Ser Leu Trp Asn
 130 135 140
 Ala Gly Thr Ser Val Thr Cys Thr Leu Asn His Pro Ser Leu Pro Pro
 145 150 155 160
 Gln Arg Leu Met Ala Leu Arg Glu Pro Ala Ala Gln Ala Pro Val Lys
 165 170 175
 Leu Ser Leu Asn Leu Leu Ala Ser Ser Asp Pro Pro Glu Ala Ala Ser
 180 185 190
 Trp Leu Leu Cys Glu Val Ser Gly Phe Ser Pro Pro Asn Ile Leu Leu
 195 200 205
 Met Trp Leu Glu Asp Gln Arg Glu Val Asn Thr Ser Gly Phe Ala Pro
 210 215 220
 Ala Arg Pro Pro Pro Gln Pro Gly Ser Thr Thr Phe Trp Ala Trp Ser
 225 230 235 240
 Val Leu Arg Val Pro Ala Pro Pro Ser Pro Gln Pro Ala Thr Tyr Thr
 245 250 255
 Cys Val Val Ser His Glu Asp Ser Arg Thr Leu Leu Asn Ala Ser Arg
 260 265 270
 Ser Leu Glu Val Ser Tyr Val Thr Asp His
 275 280

<210> 6

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> T2A

<400> 6

Leu Glu Gly Gly Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp
 1 5 10 15
 Val Glu Glu Asn Pro Gly Pro Arg
 20

<210> 7

<211> 357

<212> PRT

<213> 人工序列

<220>

<223> tEGFR

<400> 7

Met Leu Leu Leu Val Thr Ser Leu Leu Leu Cys Glu Leu Pro His Pro

1	5	10	15
Ala Phe Leu Leu Ile Pro Arg Lys Val Cys Asn Gly Ile Gly Ile Gly			
	20	25	30
Glu Phe Lys Asp Ser Leu Ser Ile Asn Ala Thr Asn Ile Lys His Phe			
	35	40	45
Lys Asn Cys Thr Ser Ile Ser Gly Asp Leu His Ile Leu Pro Val Ala			
	50	55	60
Phe Arg Gly Asp Ser Phe Thr His Thr Pro Pro Leu Asp Pro Gln Glu			
65	70	75	80
Leu Asp Ile Leu Lys Thr Val Lys Glu Ile Thr Gly Phe Leu Leu Ile			
	85	90	95
Gln Ala Trp Pro Glu Asn Arg Thr Asp Leu His Ala Phe Glu Asn Leu			
	100	105	110
Glu Ile Ile Arg Gly Arg Thr Lys Gln His Gly Gln Phe Ser Leu Ala			
	115	120	125
Val Val Ser Leu Asn Ile Thr Ser Leu Gly Leu Arg Ser Leu Lys Glu			
	130	135	140
Ile Ser Asp Gly Asp Val Ile Ile Ser Gly Asn Lys Asn Leu Cys Tyr			
145	150	155	160
Ala Asn Thr Ile Asn Trp Lys Lys Leu Phe Gly Thr Ser Gly Gln Lys			
	165	170	175
Thr Lys Ile Ile Ser Asn Arg Gly Glu Asn Ser Cys Lys Ala Thr Gly			
	180	185	190
Gln Val Cys His Ala Leu Cys Ser Pro Glu Gly Cys Trp Gly Pro Glu			
	195	200	205
Pro Arg Asp Cys Val Ser Cys Arg Asn Val Ser Arg Gly Arg Glu Cys			
	210	215	220
Val Asp Lys Cys Asn Leu Leu Glu Gly Glu Pro Arg Glu Phe Val Glu			
225	230	235	240
Asn Ser Glu Cys Ile Gln Cys His Pro Glu Cys Leu Pro Gln Ala Met			
	245	250	255
Asn Ile Thr Cys Thr Gly Arg Gly Pro Asp Asn Cys Ile Gln Cys Ala			
	260	265	270
His Tyr Ile Asp Gly Pro His Cys Val Lys Thr Cys Pro Ala Gly Val			
	275	280	285
Met Gly Glu Asn Asn Thr Leu Val Trp Lys Tyr Ala Asp Ala Gly His			
	290	295	300
Val Cys His Leu Cys His Pro Asn Cys Thr Tyr Gly Cys Thr Gly Pro			
305	310	315	320

Gly Leu Glu Gly Cys Pro Thr Asn Gly Pro Lys Ile Pro Ser Ile Ala
 325 330 335
 Thr Gly Met Val Gly Ala Leu Leu Leu Leu Val Val Ala Leu Gly
 340 345 350
 Ile Gly Leu Phe Met
 355

<210> 8

<211> 27

<212> PRT

<213> 智人

<220>

<223> CD28

<300>

<308> UniProt P10747

<309> 1989-07-01

<400> 8

Phe Trp Val Leu Val Val Val Gly Gly Val Leu Ala Cys Tyr Ser Leu
 1 5 10 15
 Leu Val Thr Val Ala Phe Ile Ile Phe Trp Val
 20 25

<210> 9

<211> 66

<212> PRT

<213> 智人

<220>

<223> CD28

<300>

<308> UniProt P10747

<309> 1989-07-01

<400> 9

Ile Glu Val Met Tyr Pro Pro Pro Tyr Leu Asp Asn Glu Lys Ser Asn
 1 5 10 15
 Gly Thr Ile Ile His Val Lys Gly Lys His Leu Cys Pro Ser Pro Leu
 20 25 30
 Phe Pro Gly Pro Ser Lys Pro Phe Trp Val Leu Val Val Val Gly Gly
 35 40 45
 Val Leu Ala Cys Tyr Ser Leu Leu Val Thr Val Ala Phe Ile Ile Phe
 50 55 60
 Trp Val

65

<210> 10

<211> 41

<212> PRT

<213> 智人

<220>

<223> CD28

<300>

<308> UniProt P10747

<309> 1989-07-01

<400> 10

Arg Ser Lys Arg Ser Arg Leu Leu His Ser Asp Tyr Met Asn Met Thr

1 5 10 15

Pro Arg Arg Pro Gly Pro Thr Arg Lys His Tyr Gln Pro Tyr Ala Pro

20 25 30

Pro Arg Asp Phe Ala Ala Tyr Arg Ser

35 40

<210> 11

<211> 41

<212> PRT

<213> 智人

<220>

<223> CD28

<400> 11

Arg Ser Lys Arg Ser Arg Gly Gly His Ser Asp Tyr Met Asn Met Thr

1 5 10 15

Pro Arg Arg Pro Gly Pro Thr Arg Lys His Tyr Gln Pro Tyr Ala Pro

20 25 30

Pro Arg Asp Phe Ala Ala Tyr Arg Ser

35 40

<210> 12

<211> 42

<212> PRT

<213> 智人

<220>

<223> 4-1BB

<300>

<308> UniProt Q07011.1

<309> 1995-02-01

<400> 12

Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
 1 5 10 15
 Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
 20 25 30
 Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
 35 40

<210> 13

<211> 112

<212> PRT

<213> 智人

<220>

<223> CD3ζ

<400> 13

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly
 1 5 10 15
 Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
 20 25 30
 Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
 35 40 45
 Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
 50 55 60
 Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
 65 70 75 80
 Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
 85 90 95
 Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
 100 105 110

<210> 14

<211> 112

<212> PRT

<213> 智人

<220>

<223> CD3ζ

<400> 14

Arg Val Lys Phe Ser Arg Ser Ala Glu Pro Pro Ala Tyr Gln Gln Gly
 1 5 10 15
 Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
 20 25 30

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
 35 40 45
 Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
 50 55 60
 Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
 65 70 75 80
 Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
 85 90 95
 Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
 100 105 110

<210> 15

<211> 112

<212> PRT

<213> 智人

<220>

<223> CD3ζ

<400> 15

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
 1 5 10 15
 Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
 20 25 30
 Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
 35 40 45
 Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
 50 55 60
 Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
 65 70 75 80
 Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
 85 90 95
 Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
 100 105 110

<210> 16

<211> 335

<212> PRT

<213> 人工序列

<220>

<223> tEGFR

<400> 16

Arg Lys Val Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp Ser Leu

1	5	10	15
Ser Ile Asn Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr Ser Ile			
	20	25	30
Ser Gly Asp Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp Ser Phe			
	35	40	45
Thr His Thr Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu Lys Thr			
	50	55	60
Val Lys Glu Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro Glu Asn			
65	70	75	80
Arg Thr Asp Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg Gly Arg			
	85	90	95
Thr Lys Gln His Gly Gln Phe Ser Leu Ala Val Val Ser Leu Asn Ile			
	100	105	110
Thr Ser Leu Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly Asp Val			
	115	120	125
Ile Ile Ser Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile Asn Trp			
	130	135	140
Lys Lys Leu Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile Ser Asn			
145	150	155	160
Arg Gly Glu Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His Ala Leu			
	165	170	175
Cys Ser Pro Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys Val Ser			
	180	185	190
Cys Arg Asn Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys Asn Leu			
	195	200	205
Leu Glu Gly Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys Ile Gln			
	210	215	220
Cys His Pro Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys Thr Gly			
225	230	235	240
Arg Gly Pro Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp Gly Pro			
	245	250	255
His Cys Val Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn Asn Thr			
	260	265	270
Leu Val Trp Lys Tyr Ala Asp Ala Gly His Val Cys His Leu Cys His			
	275	280	285
Pro Asn Cys Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly Cys Pro			
	290	295	300
Thr Asn Gly Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala			
305	310	315	320

Leu Leu Leu Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met
 325 330 335

<210> 17

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> T2A

<400> 17

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
 1 5 10 15

Gly Pro

<210> 18

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> P2A

<400> 18

Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val
 1 5 10 15

Glu Glu Asn Pro Gly Pro
 20

<210> 19

<211> 19

<212> PRT

<213> 人工序列

<220>

<223> P2A

<400> 19

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
 1 5 10 15

Pro Gly Pro

<210> 20

<211> 20

<212> PRT

<213> 人工序列

<220>

<223> E2A

<400> 20

Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

<210> 21

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> F2A

<400> 21

Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val

1 5 10 15

Glu Ser Asn Pro Gly Pro

20

<210> 22

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> 接头

<220>

<221> 重复序列

<222> (5) ... (9)

<223> SGGGG重复5次

<400> 22

Pro Gly Gly Gly Ser Gly Gly Gly Gly Pro

1 5 10

<210> 23

<211> 17

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 23

Gly Ser Ala Asp Asp Ala Lys Lys Asp Ala Ala Lys Lys Asp Gly Lys

1 5 10 15

Ser

<210> 24

<211> 66

<212> DNA

<213> 人工序列

<220>

<223> GMCSFR α 链信号序列

<400> 24

atgcttctcc tggtgacaag cttctgctc tgtgagttac cacaccagc attcctctg 60

atccca 66

<210> 25

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> GMCSFR α 链信号序列

<400> 25

Met Leu Leu Leu Val Thr Ser Leu Leu Leu Cys Glu Leu Pro His Pro

1 5 10 15

Ala Phe Leu Leu Ile Pro

20

<210> 26

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> CD8 α 链信号肽

<400> 26

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala

<210> 27

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 27

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro

1 5 10 15

<210> 28

<211> 12

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 28

Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro

1 5 10

<210> 29

<211> 61

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 29

Glu Leu Lys Thr Pro Leu Gly Asp Thr His Thr Cys Pro Arg Cys Pro

1 5 10 15

Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro Glu

 20 25 30

Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro Glu Pro

 35 40 45

Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro

 50 55 60

<210> 30

<211> 12

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 30

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro

1 5 10

<210> 31

<211> 12

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 31

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro
1 5 10

<210> 32

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 32

Tyr Gly Pro Pro Cys Pro Pro Cys Pro
1 5

<210> 33

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 33

Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro
1 5 10

<210> 34

<211> 14

<212> PRT

<213> 人工序列

<220>

<223> 铰链

<400> 34

Glu Val Val Val Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro
1 5 10

<210> 35

<211> 11

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR L1

<400> 35

Arg Ala Ser Gln Asp Ile Ser Lys Tyr Leu Asn
1 5 10

<210> 36

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR L2

<400> 36

Ser Arg Leu His Ser Gly Val

1 5

<210> 37

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR L3

<400> 37

Gly Asn Thr Leu Pro Tyr Thr Phe Gly

1 5

<210> 38

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR H1

<400> 38

Asp Tyr Gly Val Ser

1 5

<210> 39

<211> 16

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR H2

<400> 39

Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser Ala Leu Lys Ser

1 5 10 15

<210> 40

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR H3

<400> 40

Tyr Ala Met Asp Tyr Trp Gly

1 5

<210> 41

<211> 120

<212> PRT

<213> 人工序列

<220>

<223> FMC63 VH

<400> 41

Glu Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln

1 5 10 15

Ser Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu Pro Asp Tyr

20 25 30

Gly Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu Glu Trp Leu

35 40 45

Gly Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser Ala Leu Lys

50 55 60

Ser Arg Leu Thr Ile Ile Lys Asp Asn Ser Lys Ser Gln Val Phe Leu

65 70 75 80

Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Ile Tyr Tyr Cys Ala

85 90 95

Lys His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr Trp Gly Gln

100 105 110

Gly Thr Ser Val Thr Val Ser Ser

115 120

<210> 42

<211> 107

<212> PRT

<213> 人工序列

<220>

<223> FMC63 VL

<400> 42

Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly

1 5 10 15

Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Ser Lys Tyr

	20		25		30														
Leu	Asn	Trp	Tyr	Gln	Gln	Lys	Pro	Asp	Gly	Thr	Val	Lys	Leu	Leu	Ile				
	35						40					45							
Tyr	His	Thr	Ser	Arg	Leu	His	Ser	Gly	Val	Pro	Ser	Arg	Phe	Ser	Gly				
	50						55					60							
Ser	Gly	Ser	Gly	Thr	Asp	Tyr	Ser	Leu	Thr	Ile	Ser	Asn	Leu	Glu	Gln				
65					70					75					80				
Glu	Asp	Ile	Ala	Thr	Tyr	Phe	Cys	Gln	Gln	Gly	Asn	Thr	Leu	Pro	Tyr				
				85						90					95				
Thr	Phe	Gly	Gly	Gly	Thr	Lys	Leu	Glu	Ile	Thr									
				100						105									
<210>	43																		
<211>	245																		
<212>	PRT																		
<213>	人工序列																		
<220>																			
<223>	FMC63	scFv																	
<400>	43																		
Asp	Ile	Gln	Met	Thr	Gln	Thr	Thr	Ser	Ser	Leu	Ser	Ala	Ser	Leu	Gly				
1				5						10				15					
Asp	Arg	Val	Thr	Ile	Ser	Cys	Arg	Ala	Ser	Gln	Asp	Ile	Ser	Lys	Tyr				
				20						25				30					
Leu	Asn	Trp	Tyr	Gln	Gln	Lys	Pro	Asp	Gly	Thr	Val	Lys	Leu	Leu	Ile				
	35						40					45							
Tyr	His	Thr	Ser	Arg	Leu	His	Ser	Gly	Val	Pro	Ser	Arg	Phe	Ser	Gly				
	50						55					60							
Ser	Gly	Ser	Gly	Thr	Asp	Tyr	Ser	Leu	Thr	Ile	Ser	Asn	Leu	Glu	Gln				
65					70					75					80				
Glu	Asp	Ile	Ala	Thr	Tyr	Phe	Cys	Gln	Gln	Gly	Asn	Thr	Leu	Pro	Tyr				
				85						90					95				
Thr	Phe	Gly	Gly	Gly	Thr	Lys	Leu	Glu	Ile	Thr	Gly	Ser	Thr	Ser	Gly				
				100						105				110					
Ser	Gly	Lys	Pro	Gly	Ser	Gly	Glu	Gly	Ser	Thr	Lys	Gly	Glu	Val	Lys				
				115						120				125					
Leu	Gln	Glu	Ser	Gly	Pro	Gly	Leu	Val	Ala	Pro	Ser	Gln	Ser	Leu	Ser				
	130						135					140							
Val	Thr	Cys	Thr	Val	Ser	Gly	Val	Ser	Leu	Pro	Asp	Tyr	Gly	Val	Ser				
145						150					155				160				
Trp	Ile	Arg	Gln	Pro	Pro	Arg	Lys	Gly	Leu	Glu	Trp	Leu	Gly	Val	Ile				

<211> 5
 <212> PRT
 <213> 人工序列
 <220>
 <223> SJ25C1 CDR H1
 <400> 47
 Ser Tyr Trp Met Asn
 1 5
 <210> 48
 <211> 17
 <212> PRT
 <213> 人工序列
 <220>
 <223> SJ25C1 CDR H2
 <400> 48
 Gln Ile Tyr Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys Phe Lys
 1 5 10 15
 Gly
 <210> 49
 <211> 13
 <212> PRT
 <213> 人工序列
 <220>
 <223> SJ25C1 CDR H3
 <400> 49
 Lys Thr Ile Ser Ser Val Val Asp Phe Tyr Phe Asp Tyr
 1 5 10
 <210> 50
 <211> 122
 <212> PRT
 <213> 人工序列
 <220>
 <223> SJ25C1 VH
 <400> 50
 Glu Val Lys Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ser
 1 5 10 15
 Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser Tyr
 20 25 30
 Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile

	35		40		45												
	Gly	Gln	Ile	Tyr	Pro	Gly	Asp	Gly	Asp	Thr	Asn	Tyr	Asn	Gly	Lys	Phe	
	50					55						60					
	Lys	Gly	Gln	Ala	Thr	Leu	Thr	Ala	Asp	Lys	Ser	Ser	Ser	Thr	Ala	Tyr	
	65					70						75				80	
	Met	Gln	Leu	Ser	Gly	Leu	Thr	Ser	Glu	Asp	Ser	Ala	Val	Tyr	Phe	Cys	
					85						90				95		
	Ala	Arg	Lys	Thr	Ile	Ser	Ser	Val	Val	Asp	Phe	Tyr	Phe	Asp	Tyr	Trp	
				100						105				110			
	Gly	Gln	Gly	Thr	Thr	Val	Thr	Val	Ser	Ser							
			115					120									

<210> 51

<211> 108

<212> PRT

<213> 人工序列

<220>

<223> SJ25C1 VL

<400> 51

	Asp	Ile	Glu	Leu	Thr	Gln	Ser	Pro	Lys	Phe	Met	Ser	Thr	Ser	Val	Gly	
	1				5					10					15		
	Asp	Arg	Val	Ser	Val	Thr	Cys	Lys	Ala	Ser	Gln	Asn	Val	Gly	Thr	Asn	
				20					25					30			
	Val	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Ser	Pro	Lys	Pro	Leu	Ile	
				35					40					45			
	Tyr	Ser	Ala	Thr	Tyr	Arg	Asn	Ser	Gly	Val	Pro	Asp	Arg	Phe	Thr	Gly	
				50					55					60			
	Ser	Gly	Ser	Gly	Thr	Asp	Phe	Thr	Leu	Thr	Ile	Thr	Asn	Val	Gln	Ser	
	65					70						75			80		
	Lys	Asp	Leu	Ala	Asp	Tyr	Phe	Cys	Gln	Gln	Tyr	Asn	Arg	Tyr	Pro	Tyr	
					85						90				95		
	Thr	Ser	Gly	Gly	Gly	Thr	Lys	Leu	Glu	Ile	Lys	Arg					
				100							105						

<210> 52

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 52

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15
 <210> 53
 <211> 245
 <212> PRT
 <213> 人工序列
 <220>
 <223> SJ25C1 scFv
 <400> 53
 Glu Val Lys Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ser
 1 5 10 15
 Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser Tyr
 20 25 30
 Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile
 35 40 45
 Gly Gln Ile Tyr Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys Phe
 50 55 60
 Lys Gly Gln Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser Thr Ala Tyr
 65 70 75 80
 Met Gln Leu Ser Gly Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
 85 90 95
 Ala Arg Lys Thr Ile Ser Ser Val Val Asp Phe Tyr Phe Asp Tyr Trp
 100 105 110
 Gly Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
 115 120 125
 Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Glu Leu Thr Gln Ser
 130 135 140
 Pro Lys Phe Met Ser Thr Ser Val Gly Asp Arg Val Ser Val Thr Cys
 145 150 155 160
 Lys Ala Ser Gln Asn Val Gly Thr Asn Val Ala Trp Tyr Gln Gln Lys
 165 170 175
 Pro Gly Gln Ser Pro Lys Pro Leu Ile Tyr Ser Ala Thr Tyr Arg Asn
 180 185 190
 Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe
 195 200 205
 Thr Leu Thr Ile Thr Asn Val Gln Ser Lys Asp Leu Ala Asp Tyr Phe
 210 215 220
 Cys Gln Gln Tyr Asn Arg Tyr Pro Tyr Thr Ser Gly Gly Gly Thr Lys
 225 230 235 240

Leu Glu Ile Lys Arg

245

<210> 54

<211> 12

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR H3

<400> 54

His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr

1

5

10

<210> 55

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR L2

<400> 55

His Thr Ser Arg Leu His Ser

1

5

<210> 56

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> FMC63 CDR L3

<400> 56

Gln Gln Gly Asn Thr Leu Pro Tyr Thr

1

5

<210> 57

<211> 735

<212> DNA

<213> 人工序列

<220>

<223> 编码scFv的序列

<400> 57

gacatccaga tgaccagac cacctccagc ctgagcgcca gcctgggcca ccgggtgacc 60
atcagctgcc gggccagcca ggacatcagc aagtacctga actggtatca gcagaagccc 120
gacggcaccg tcaagctgct gatctaccac accagccggc tgcacagcgg cgtgcccagc 180

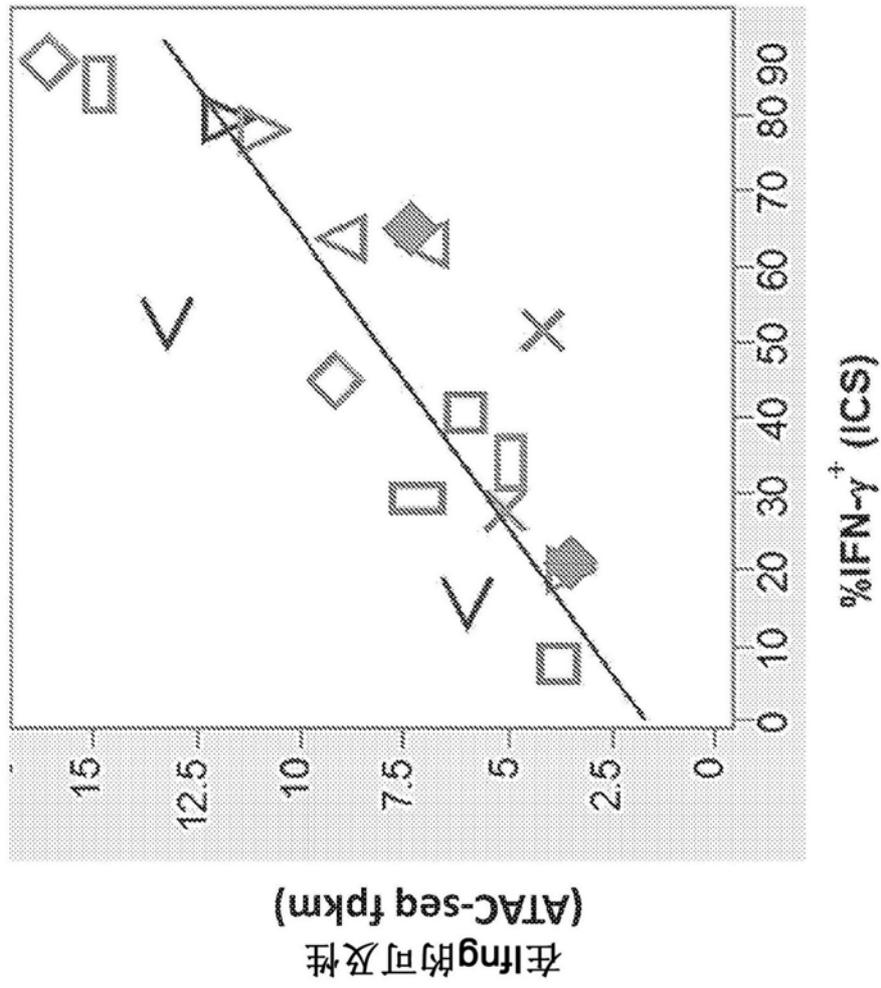


图1A

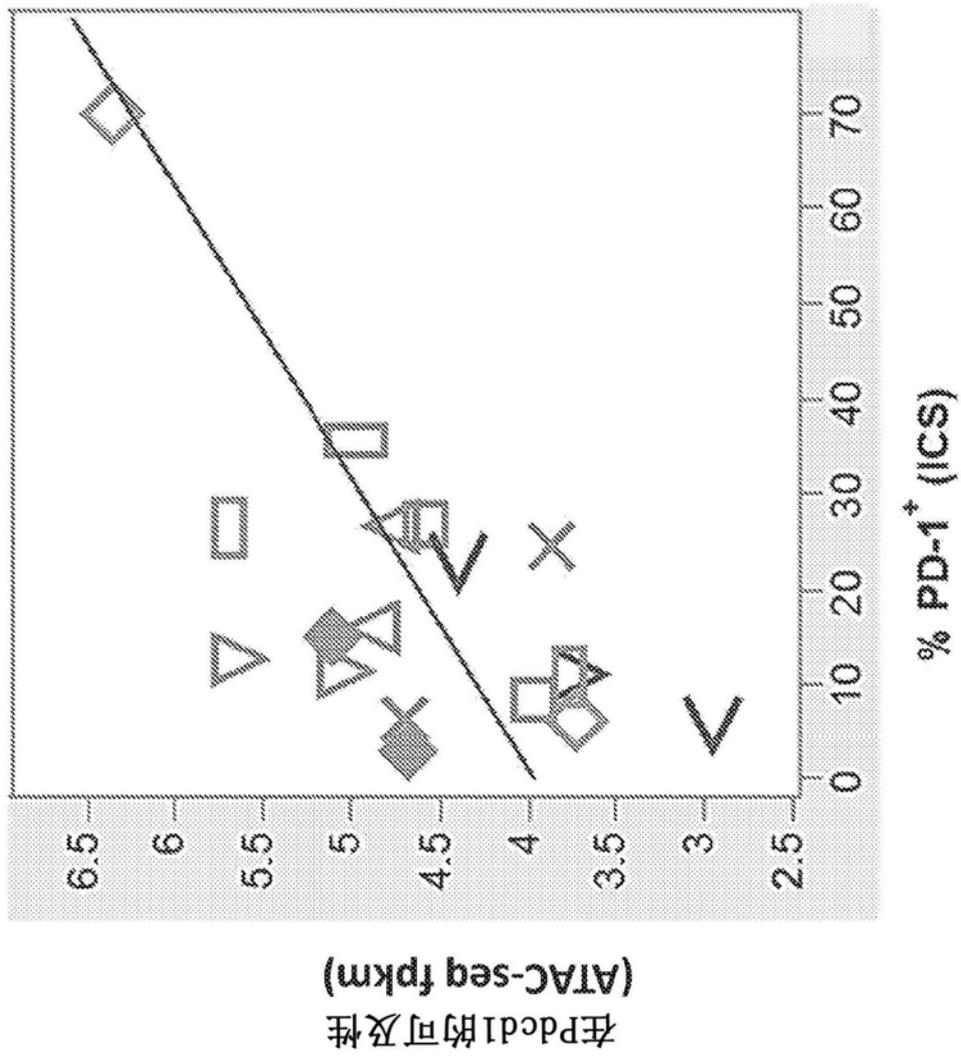


图1B

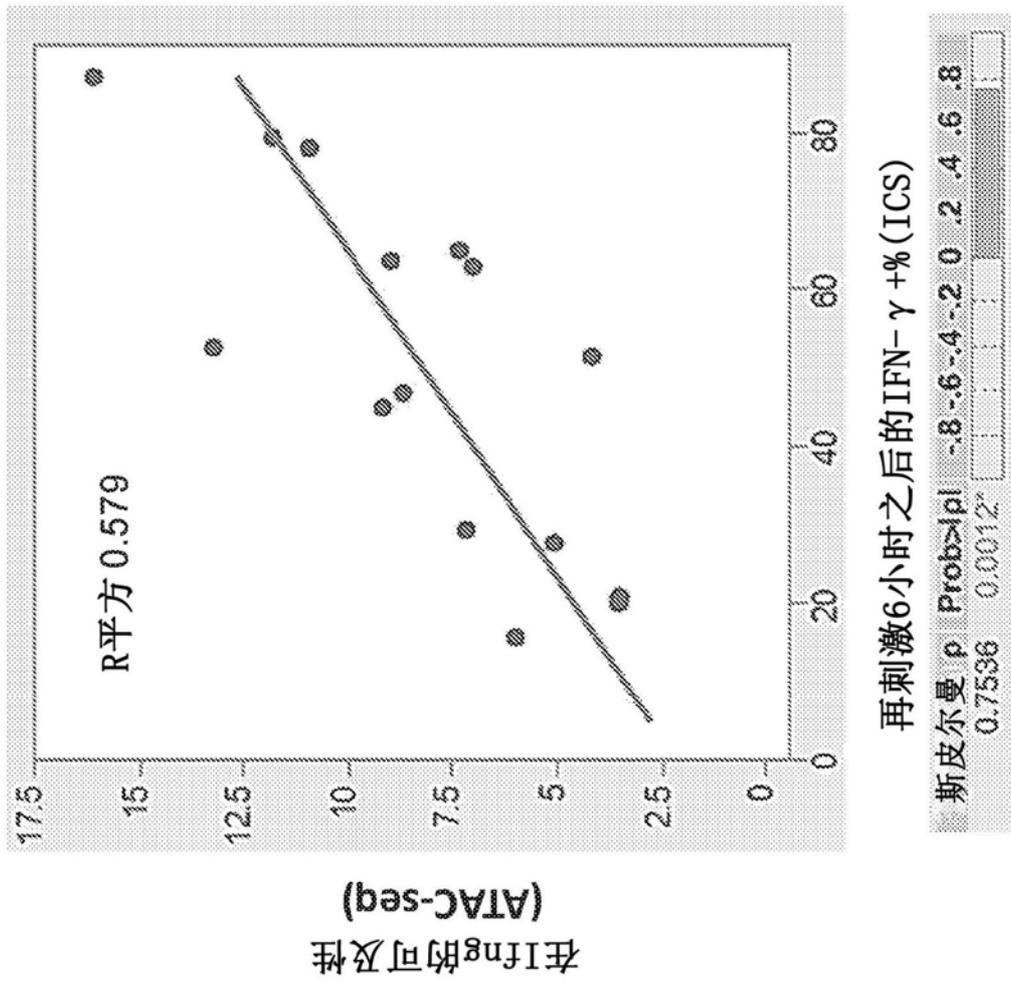


图1C

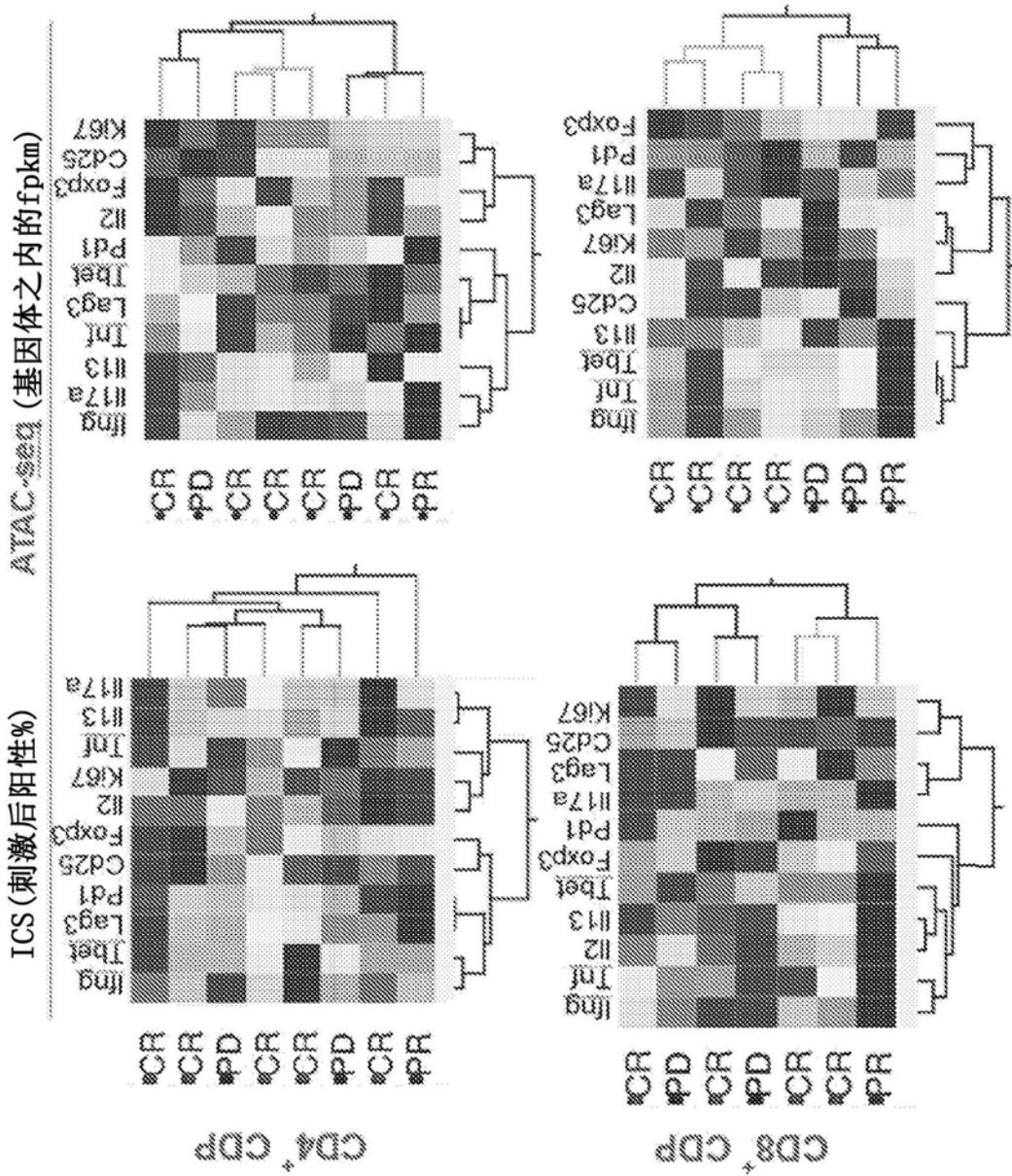


图2

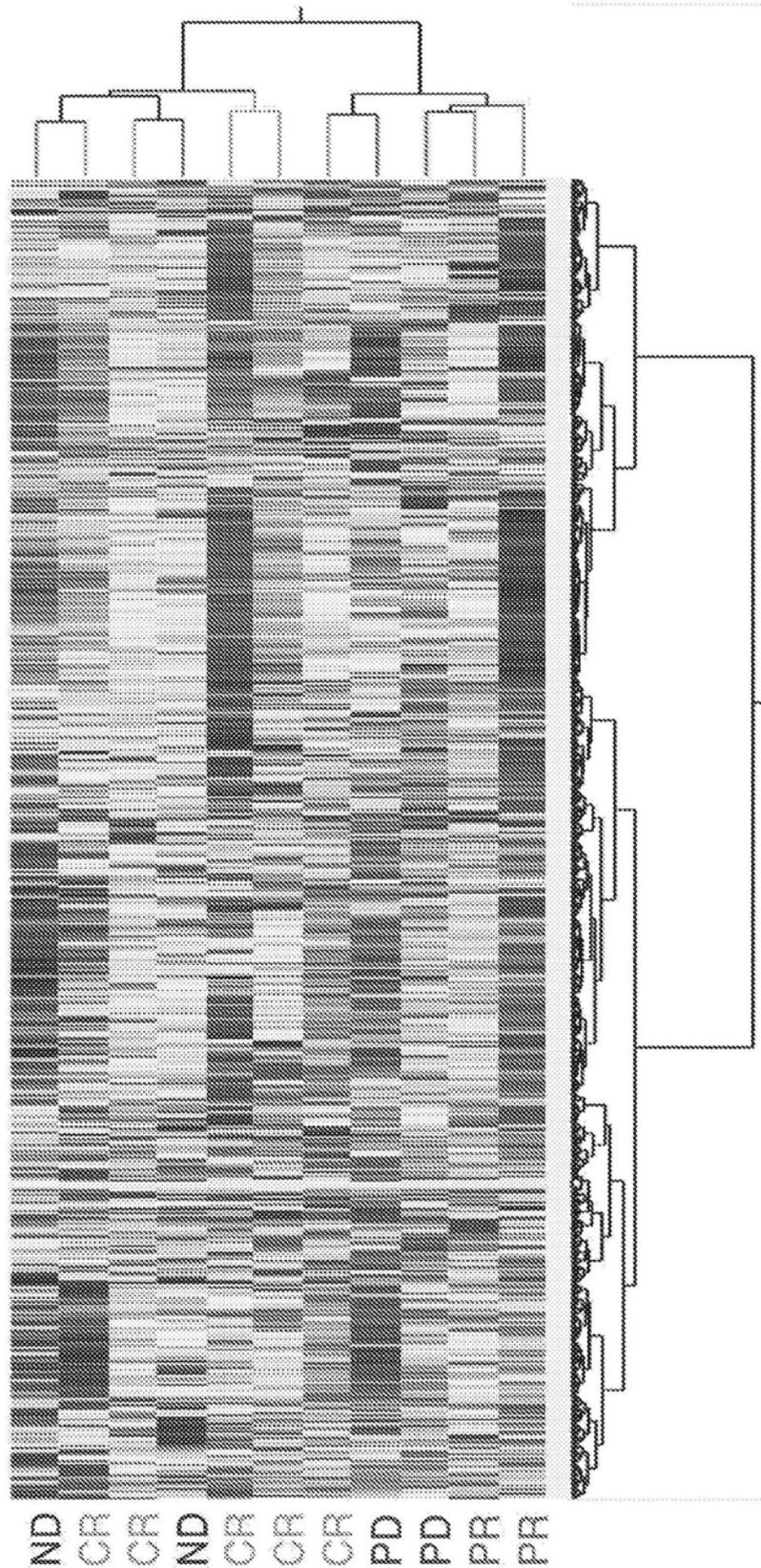


图3A

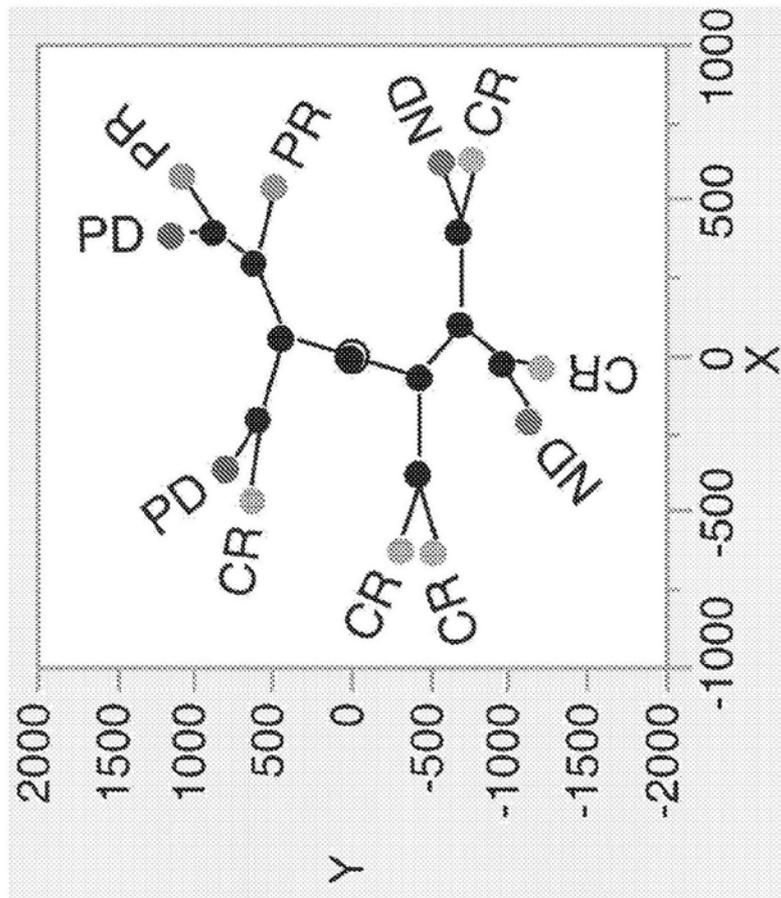


图3B

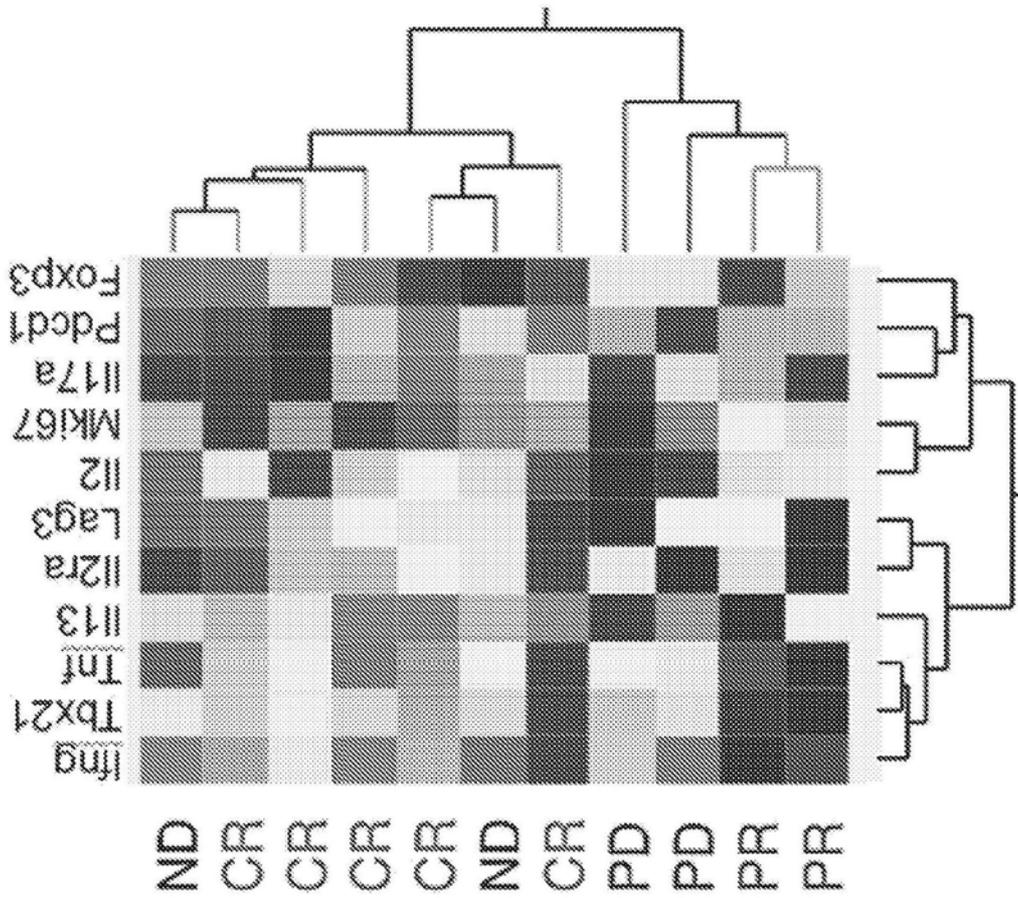


图4A

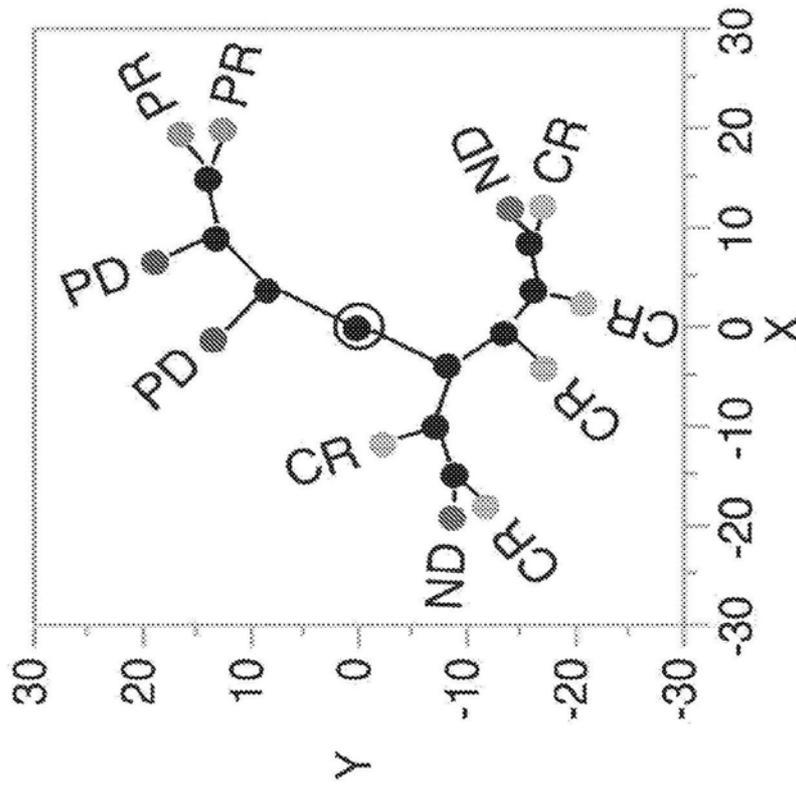


图4B

CMAT

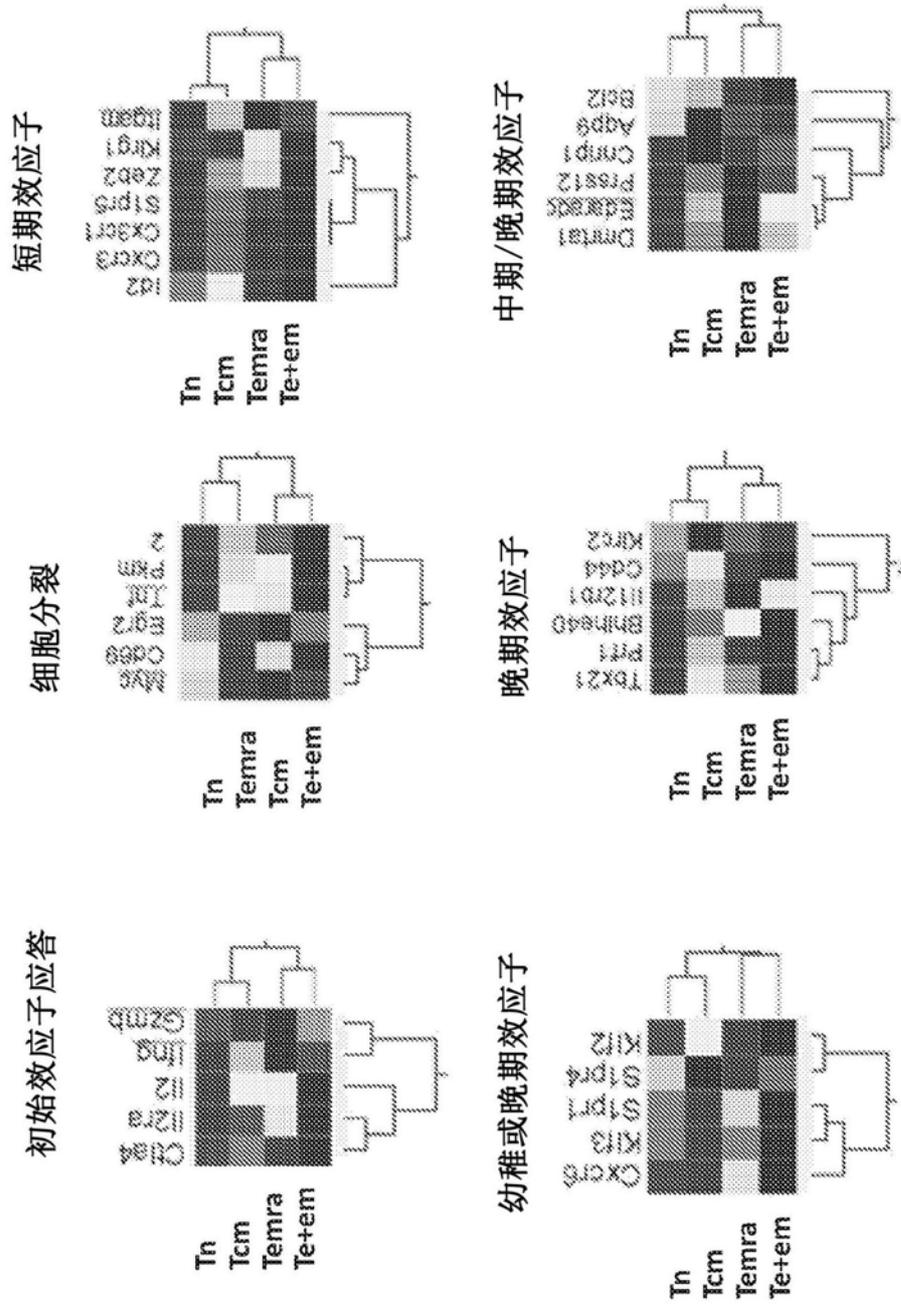


图5A

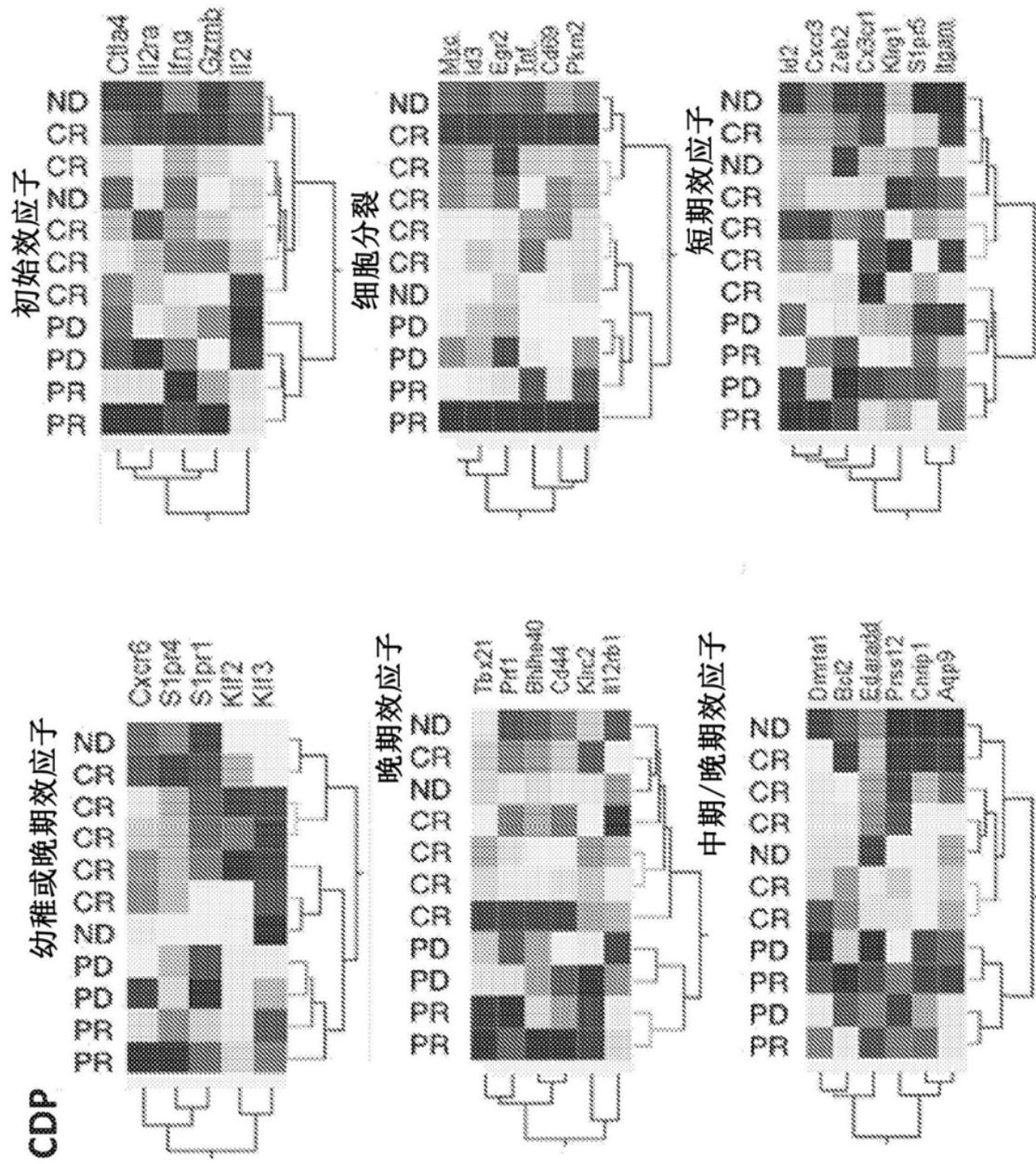


图5B

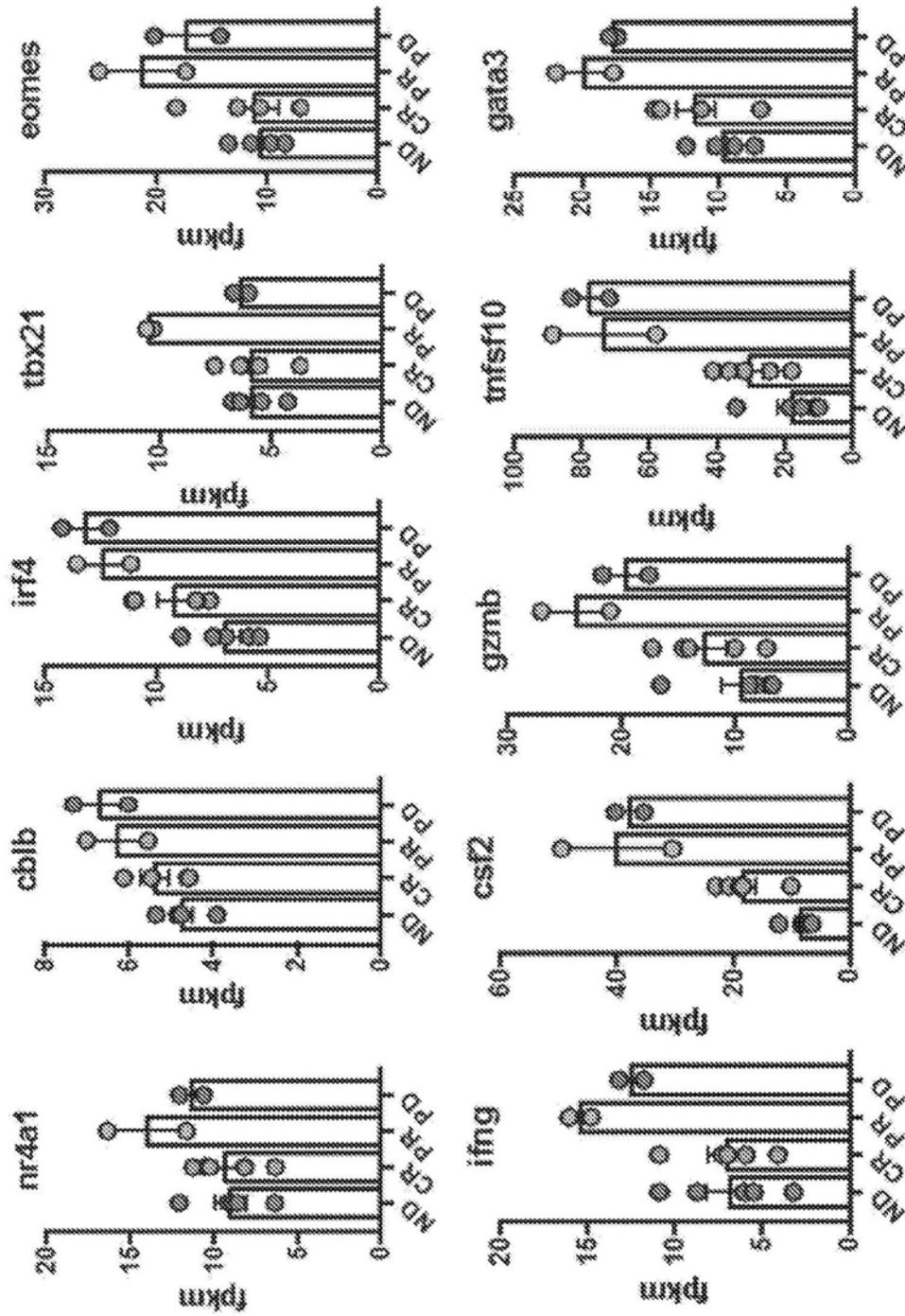


图6A

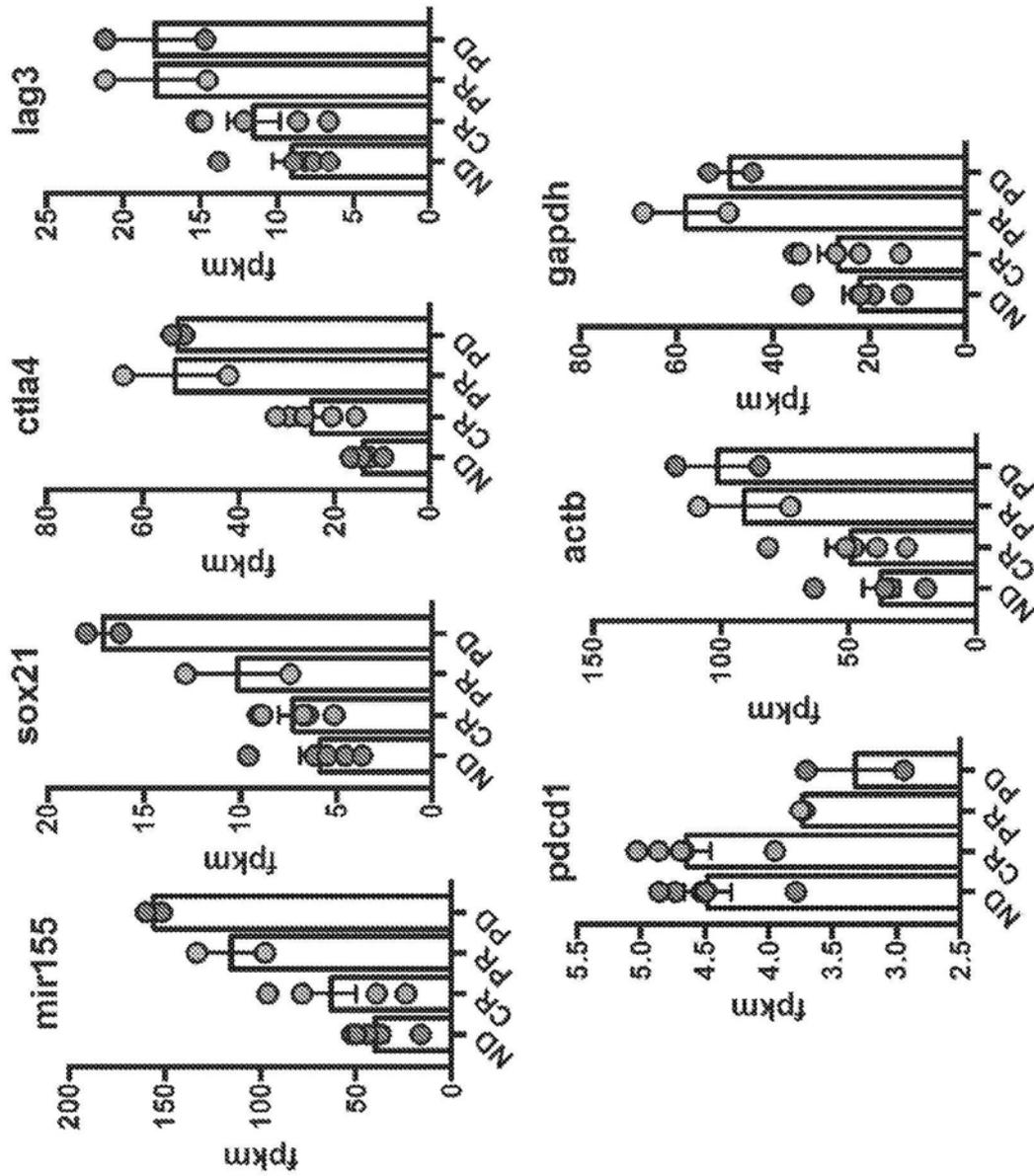


图6B

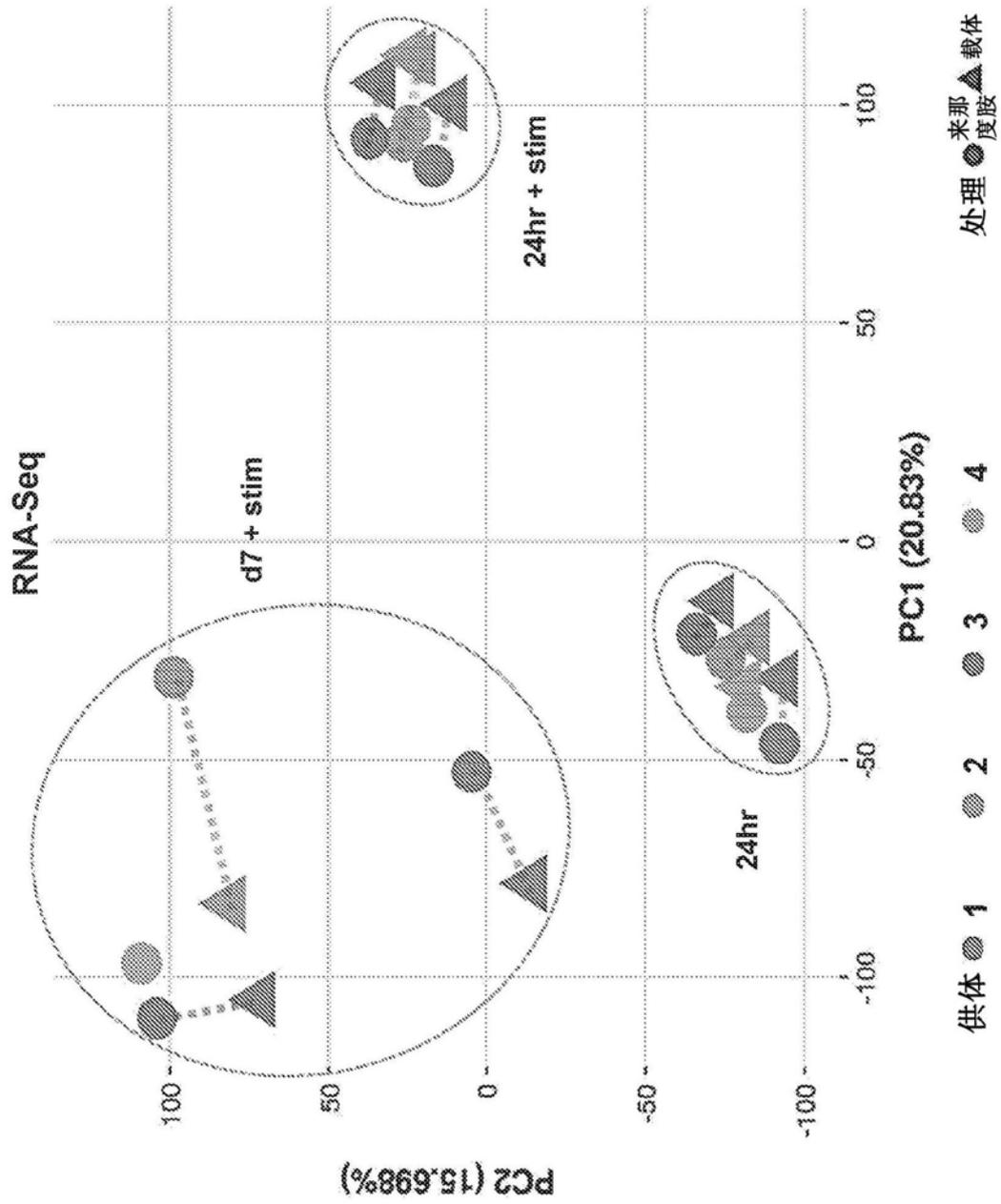


图7A

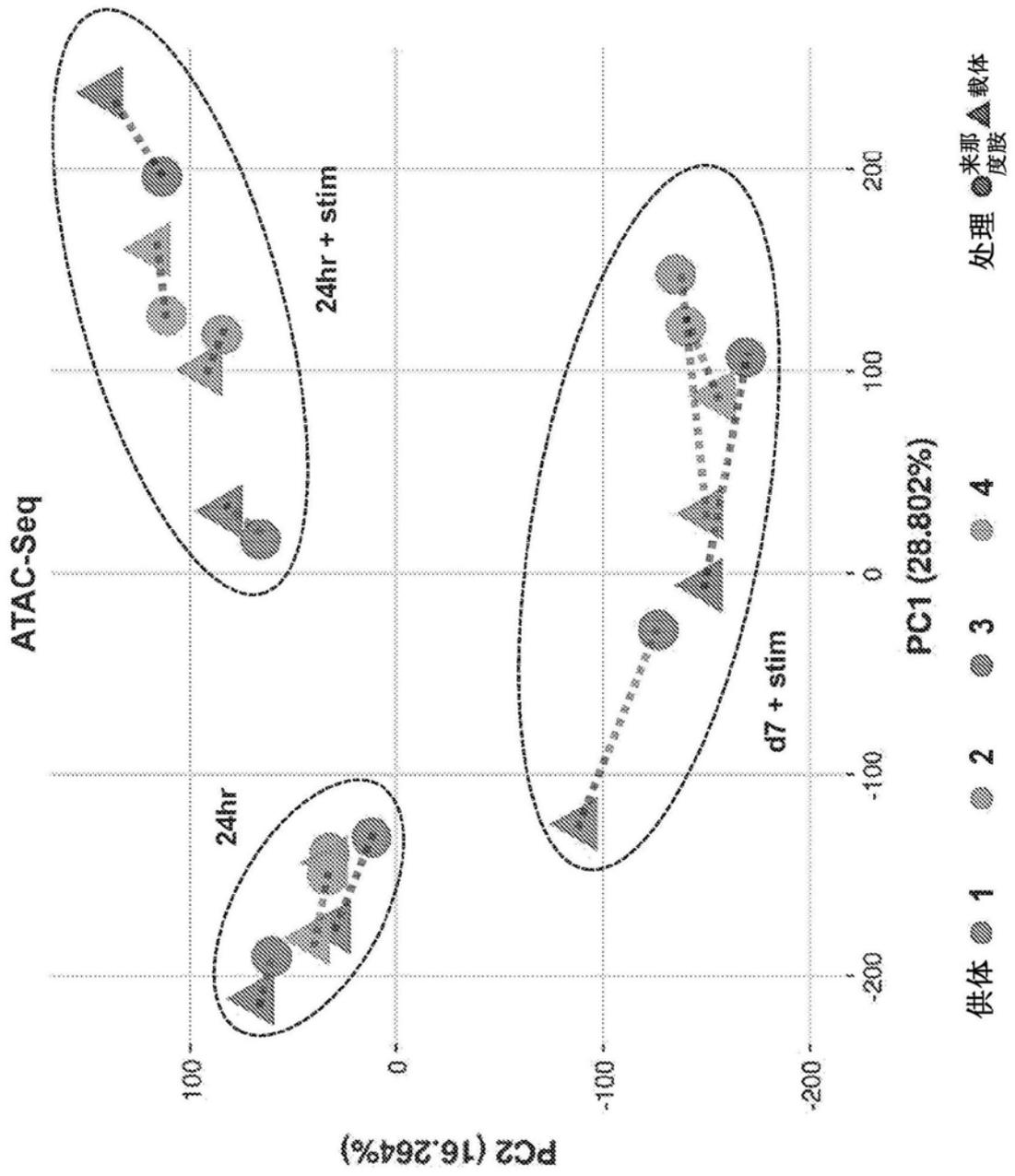


图7B

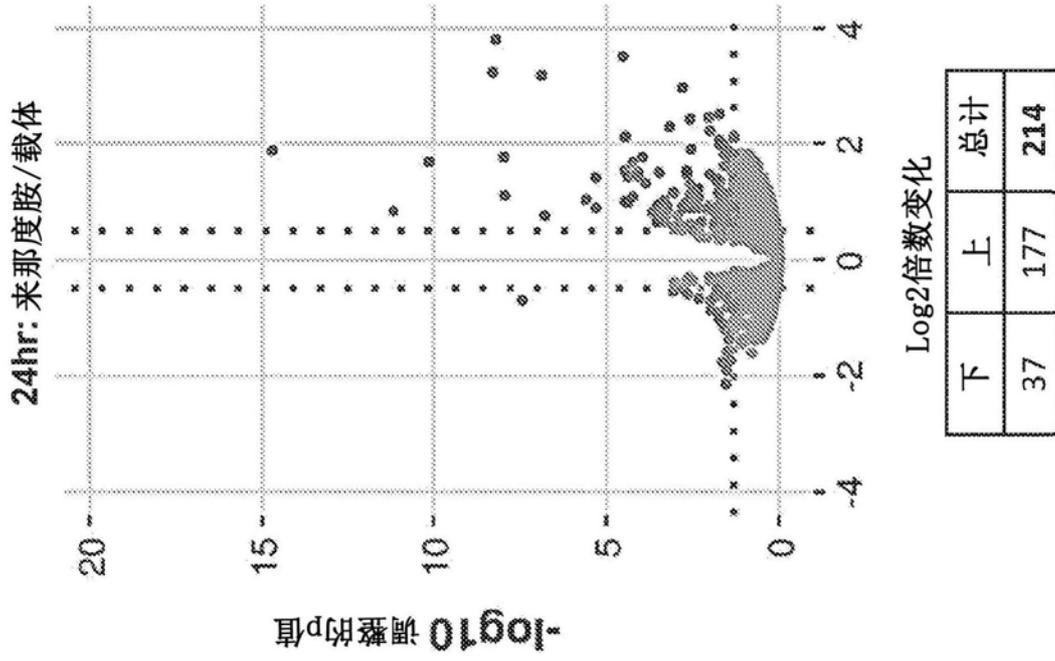


图8A

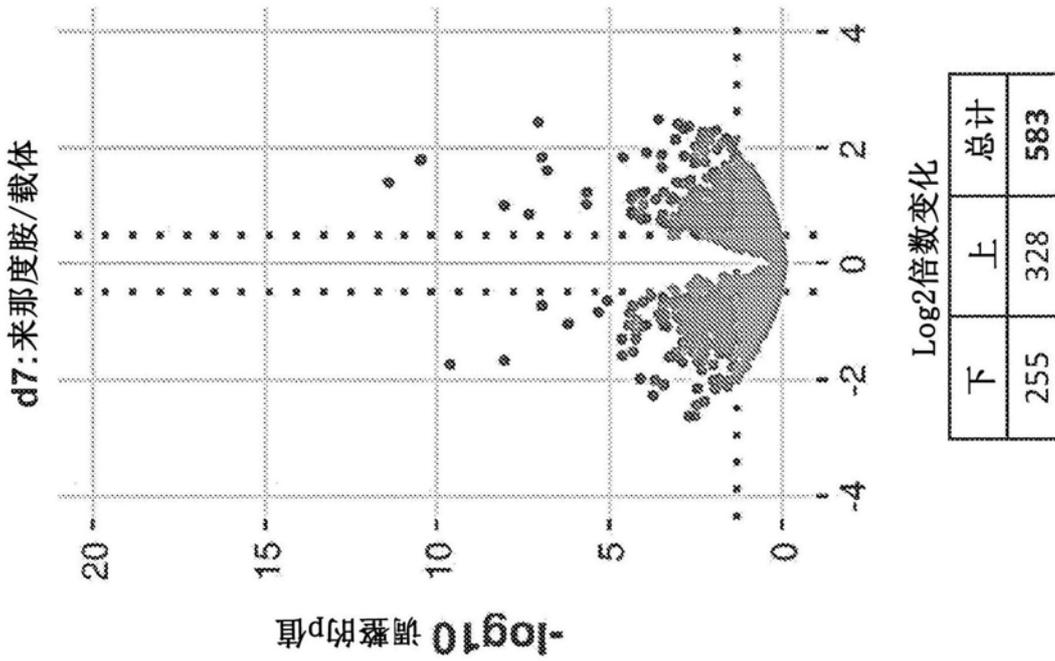


图8B

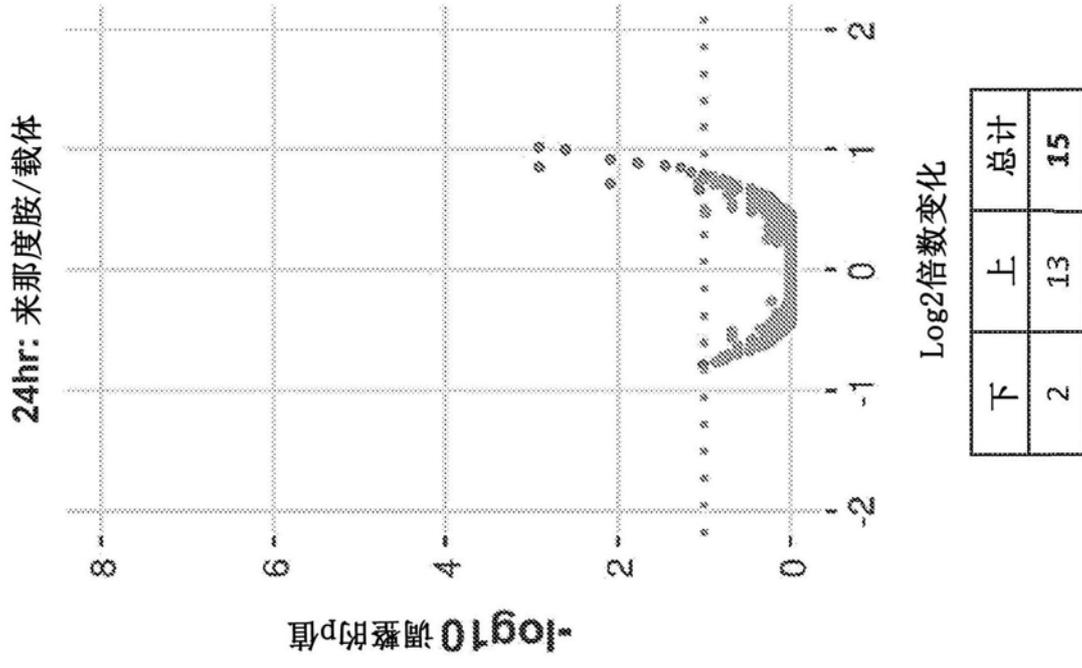


图8C

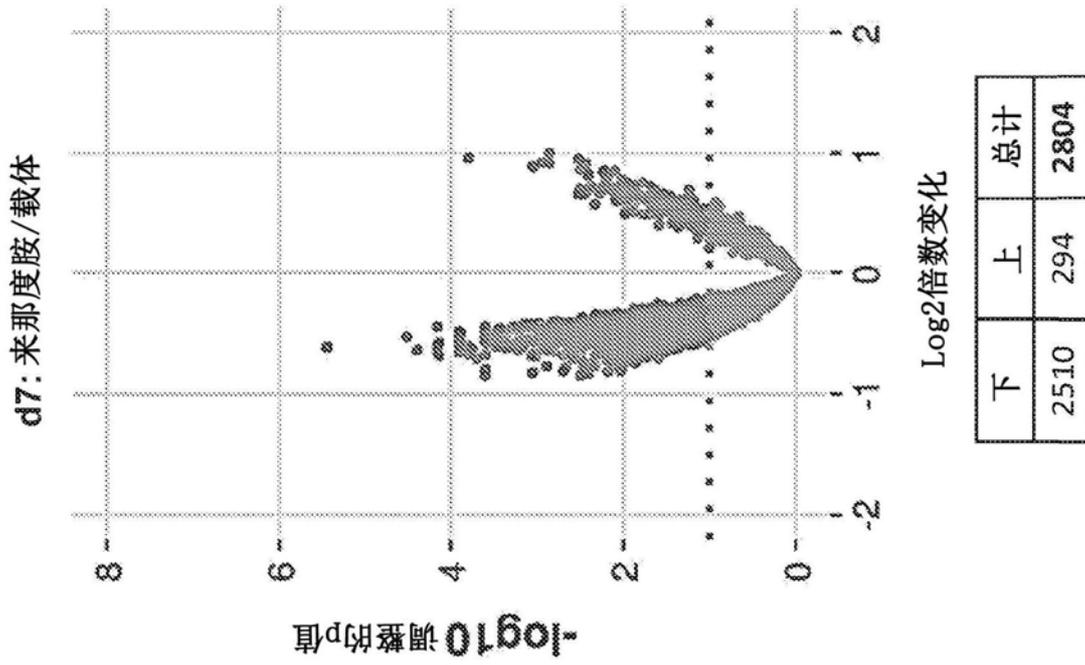


图8D

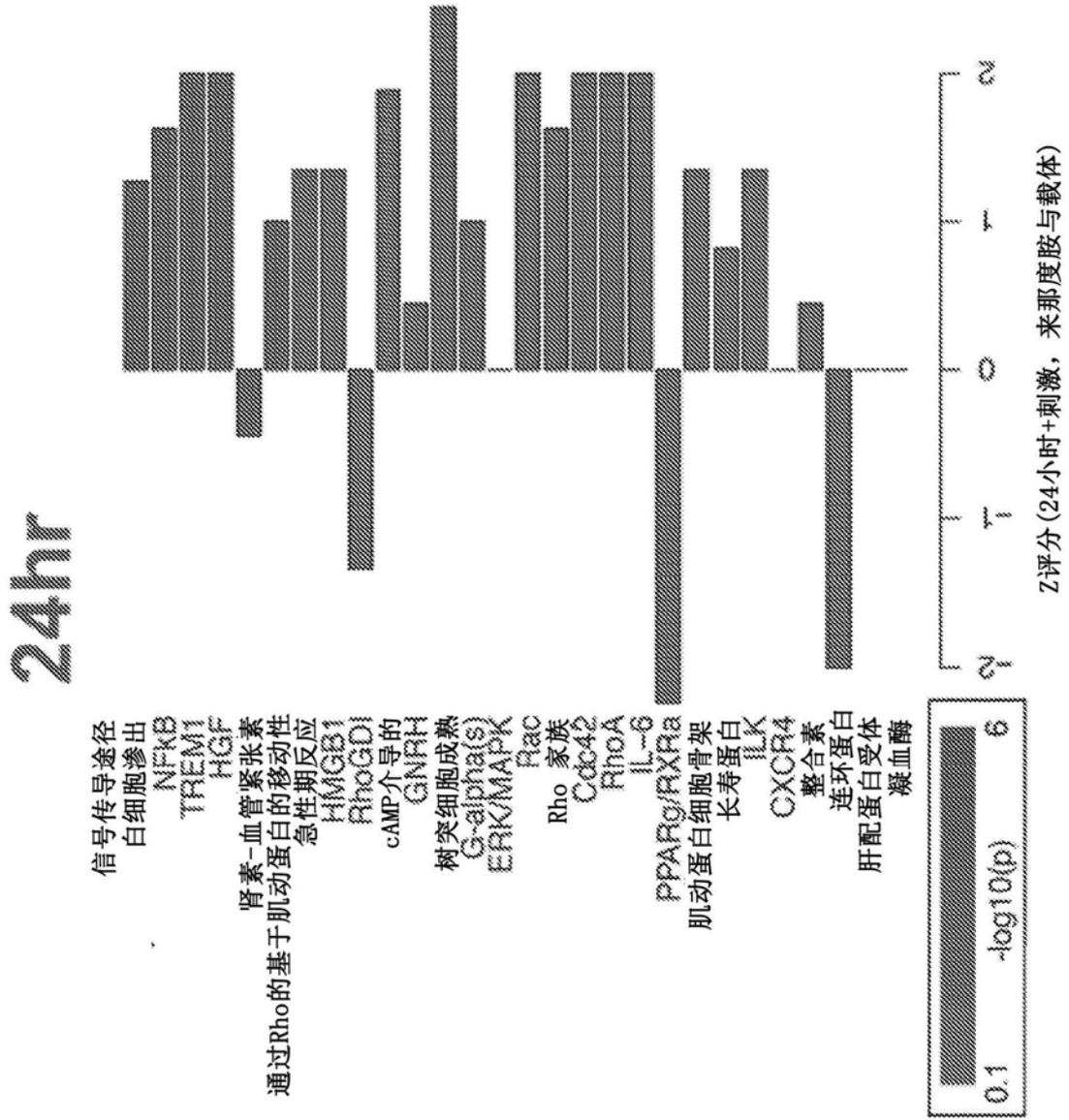


图9A

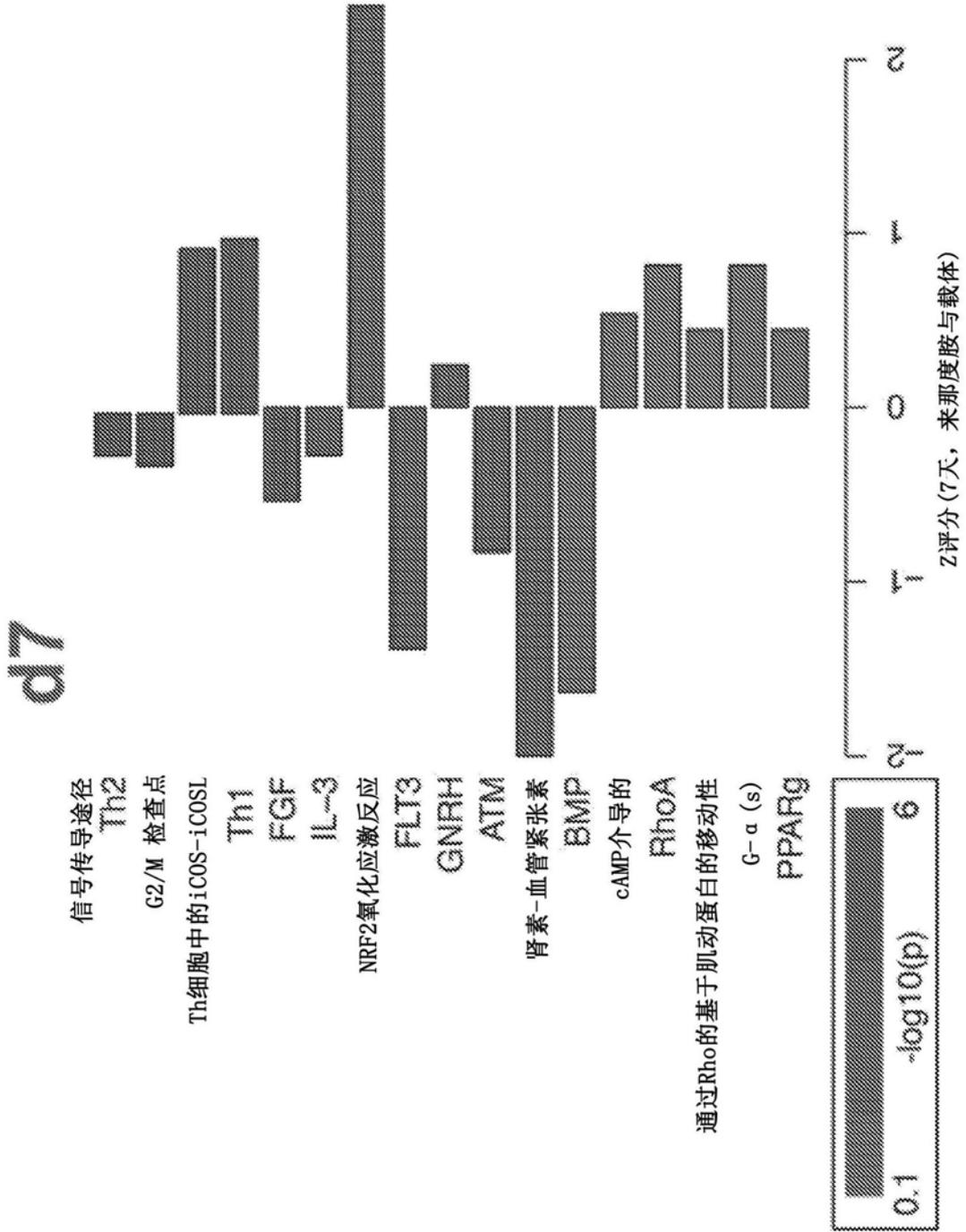


图9B

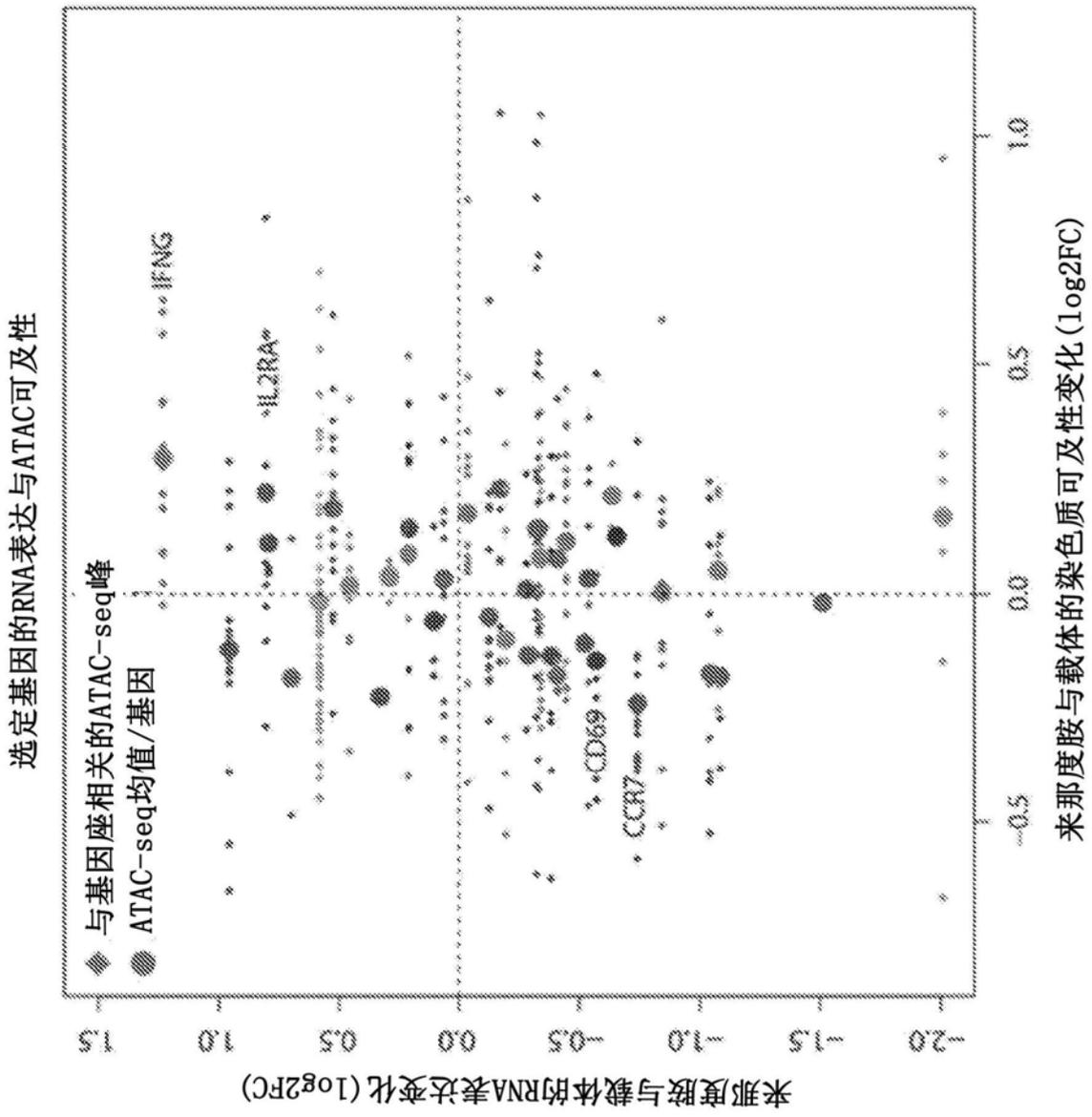


图10

基序名称	基序	Log P值	具有基序的 靶序列的%
Atf3(bZIP)/GBM-ATF3	ATGAGTCAAT	-7.66E+01	61.90%
BATF(bZIP)/Th17-BATF	ATGAGTCAAT	-7.54E+01	61.56%
Fra1(bZIP)/BT549-Fra1	ATGAGTCAAT	-7.39E+01	58.50%
AP-1(bZIP)/ThioMac-PU.1	ATGAGTCAAT	-7.33E+01	62.24%
JunB(bZIP)/DendriticCells-Junb	ATGAGTCAAT	-7.17E+01	58.16%
FosI2(bZIP)/3T3L1-FosI2	ATGAGTCAAT	-6.18E+01	46.94%
Jun-AP1(bZIP)/K562-cJun	ATGAGTCAAT	-5.14E+01	39.12%
Smad3(MAD)/NPC-Smad3	ATGAGTCAAT	-4.50E+01	56.46%
NFkB-p65(RHD)/GM12787-p65	ATGAGTCAAT	-4.41E+01	29.59%
RUNX1(Runt)/Jurkat-RUNX1	ATGAGTCAAT	-4.04E+01	50.00%
RUNX(Runt)/HPC7-Runx1	ATGAGTCAAT	-3.77E+01	41.16%
Bach2(bZIP)/OCILy7-Bach2	ATGAGTCAAT	-3.62E+01	26.53%

图11

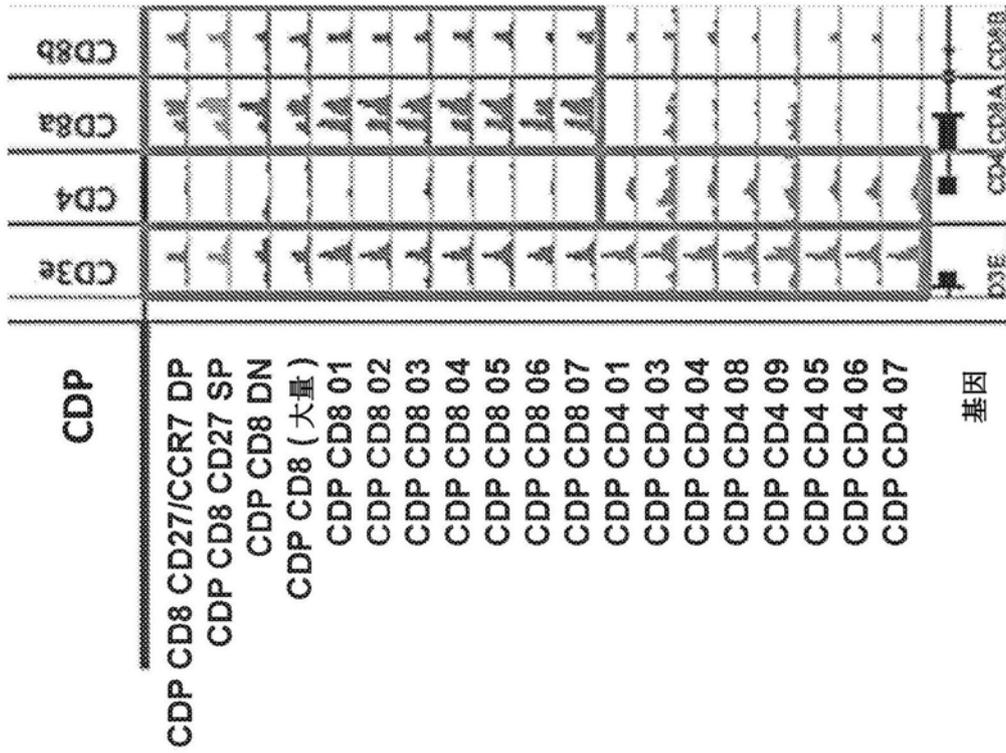


图12A

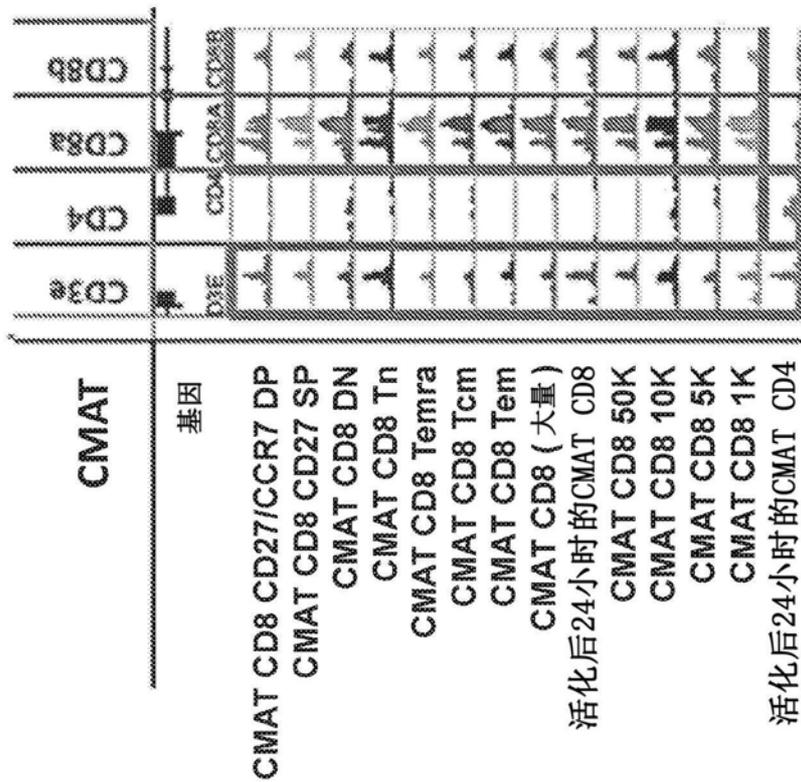


图12B

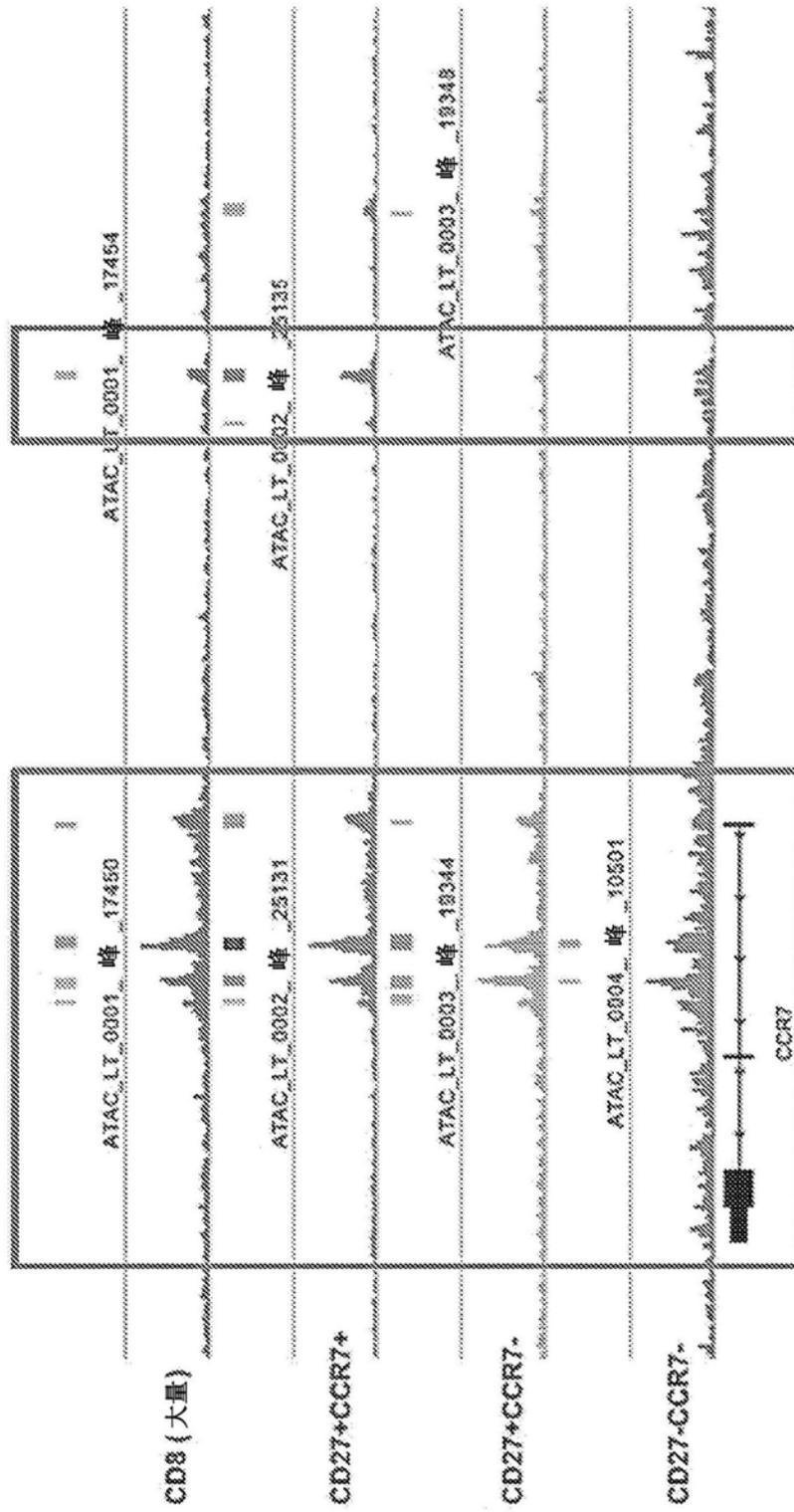


图13A

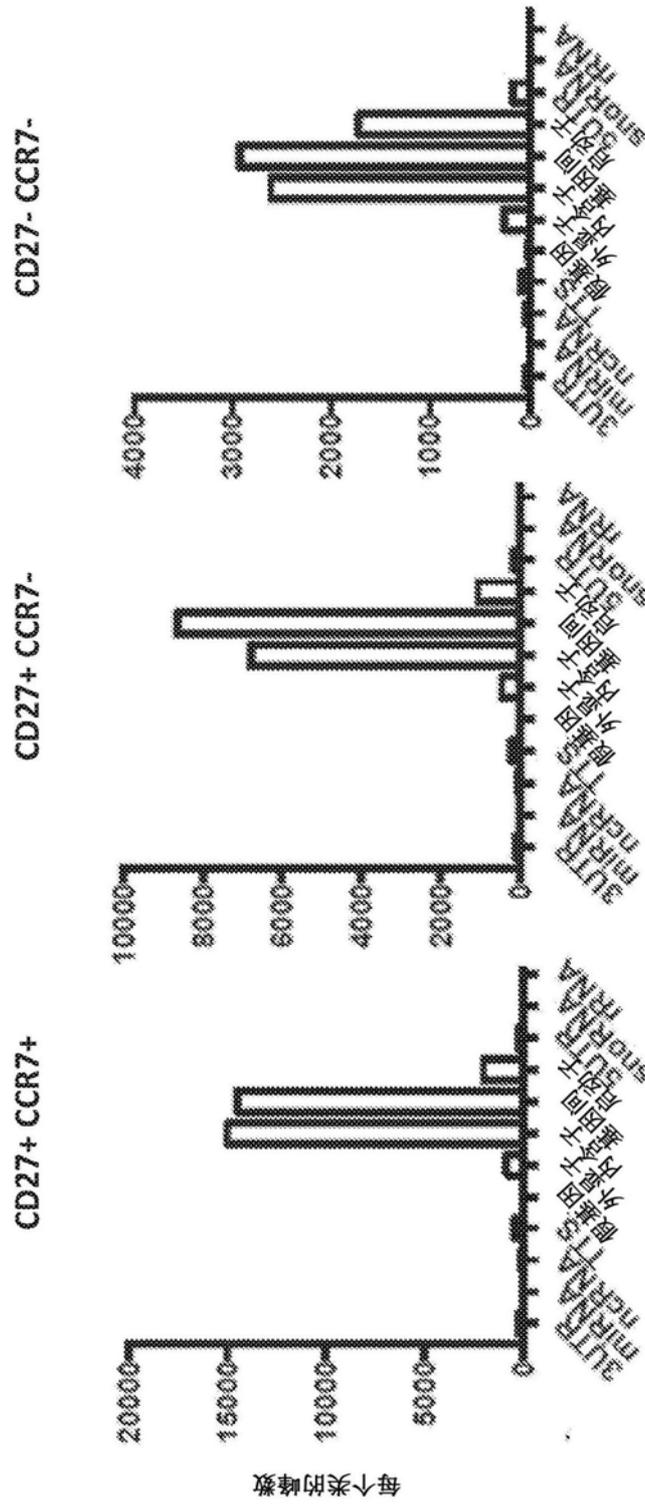


图13B

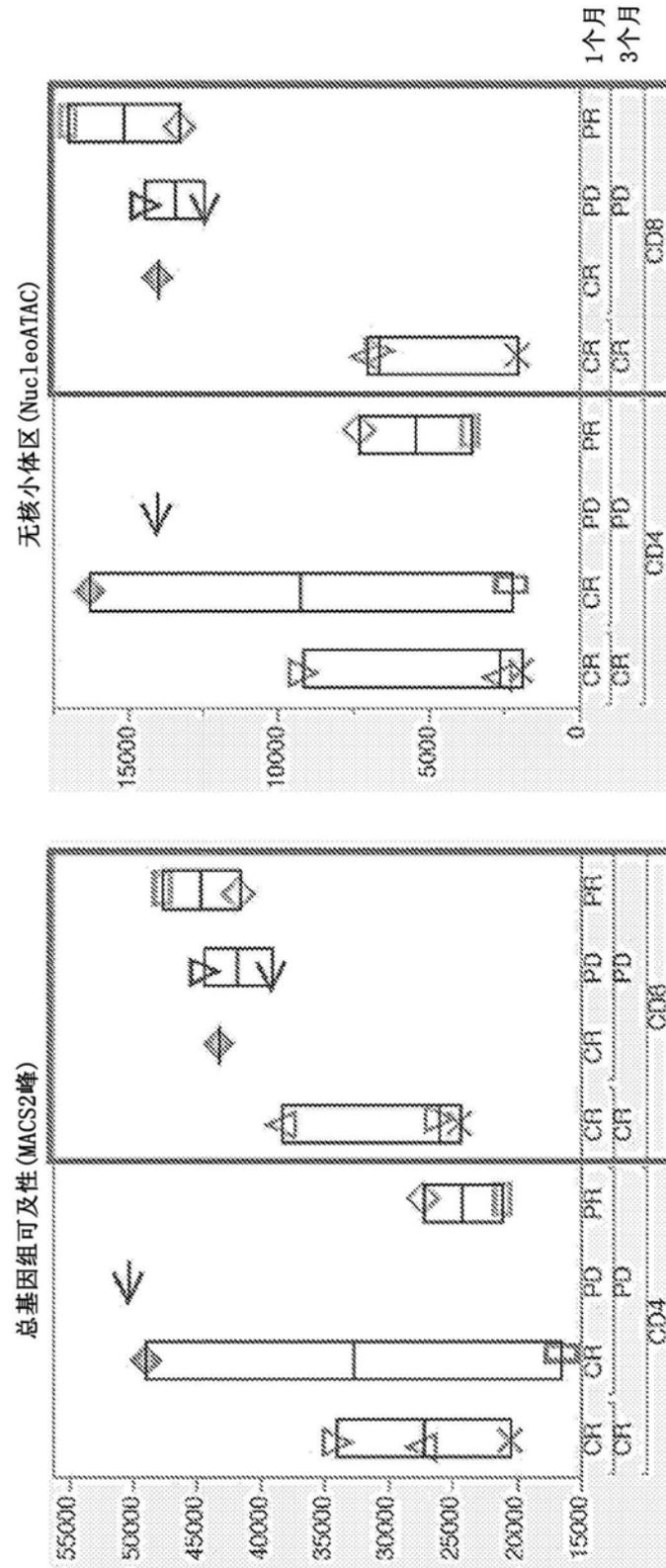


图14A

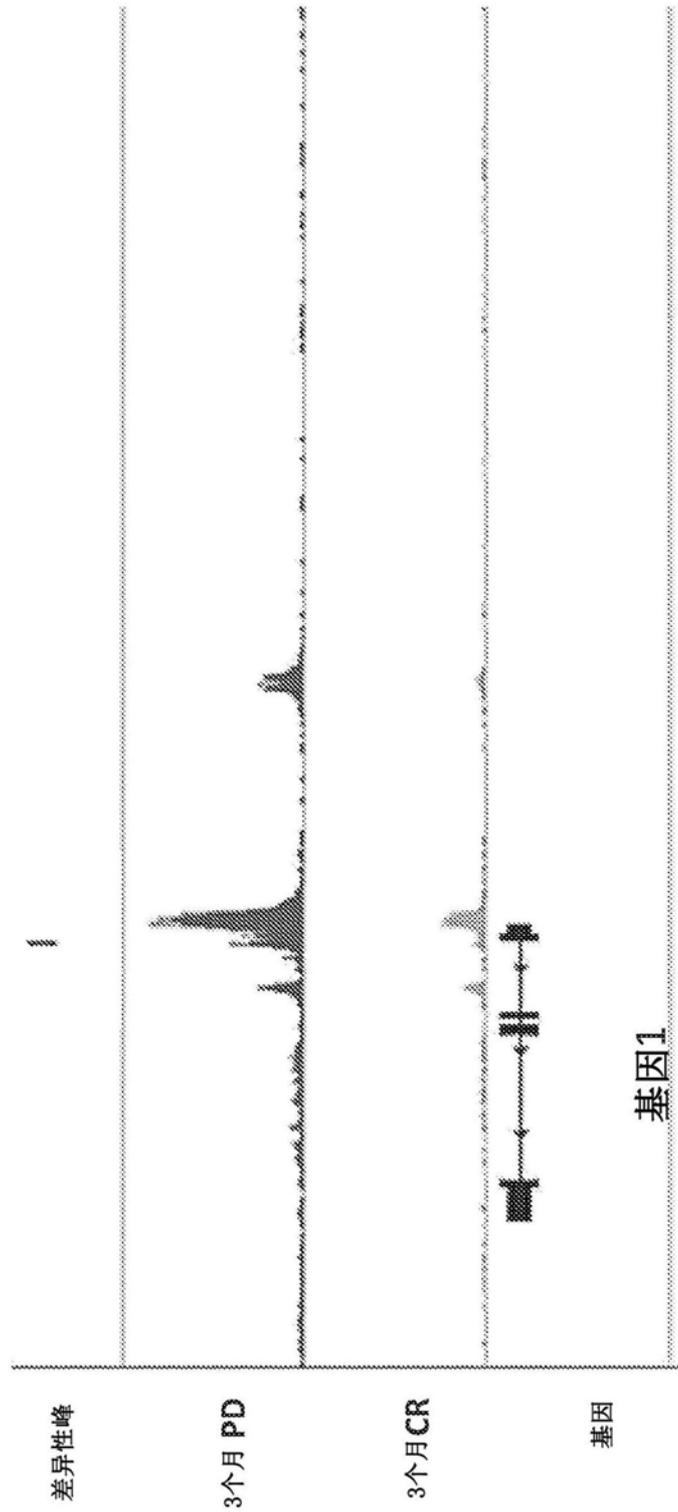


图14B

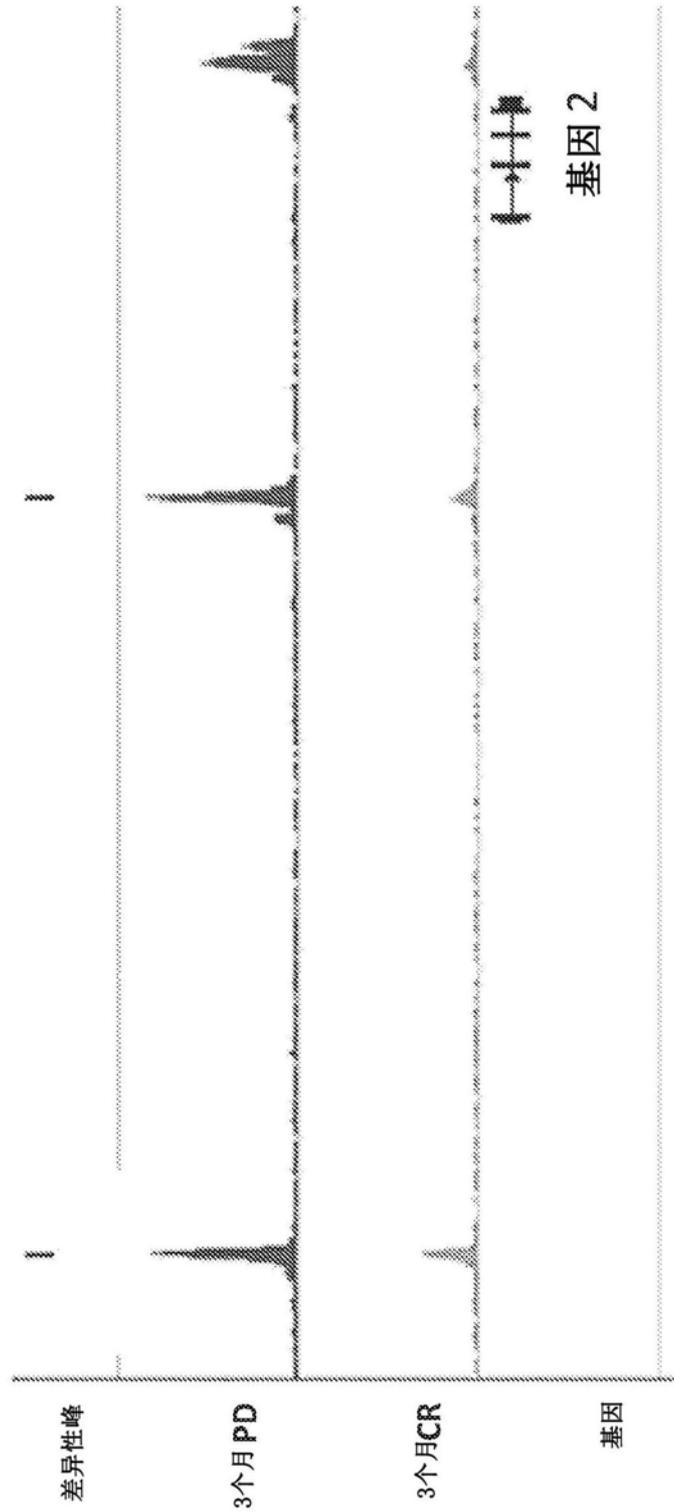


图14C

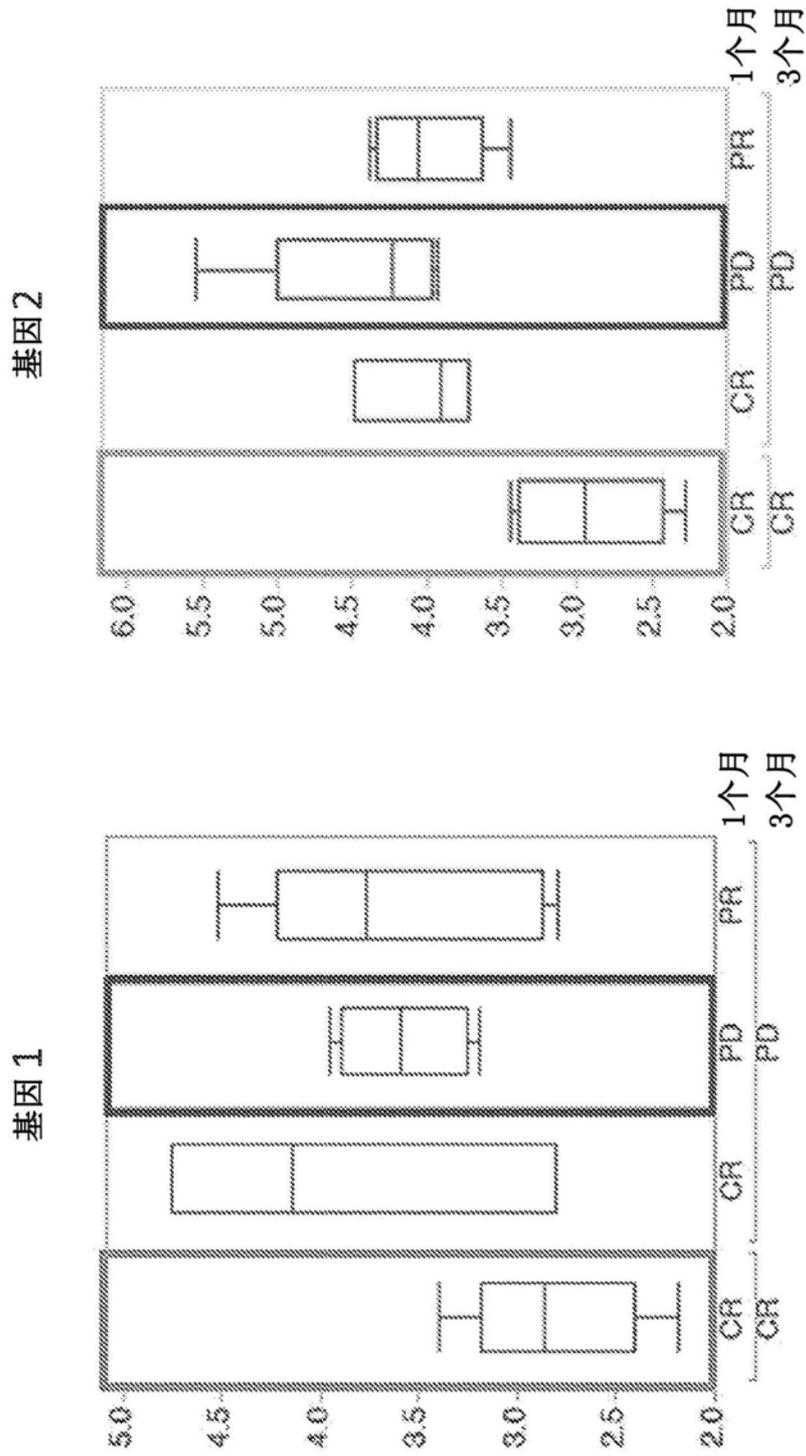


图14D

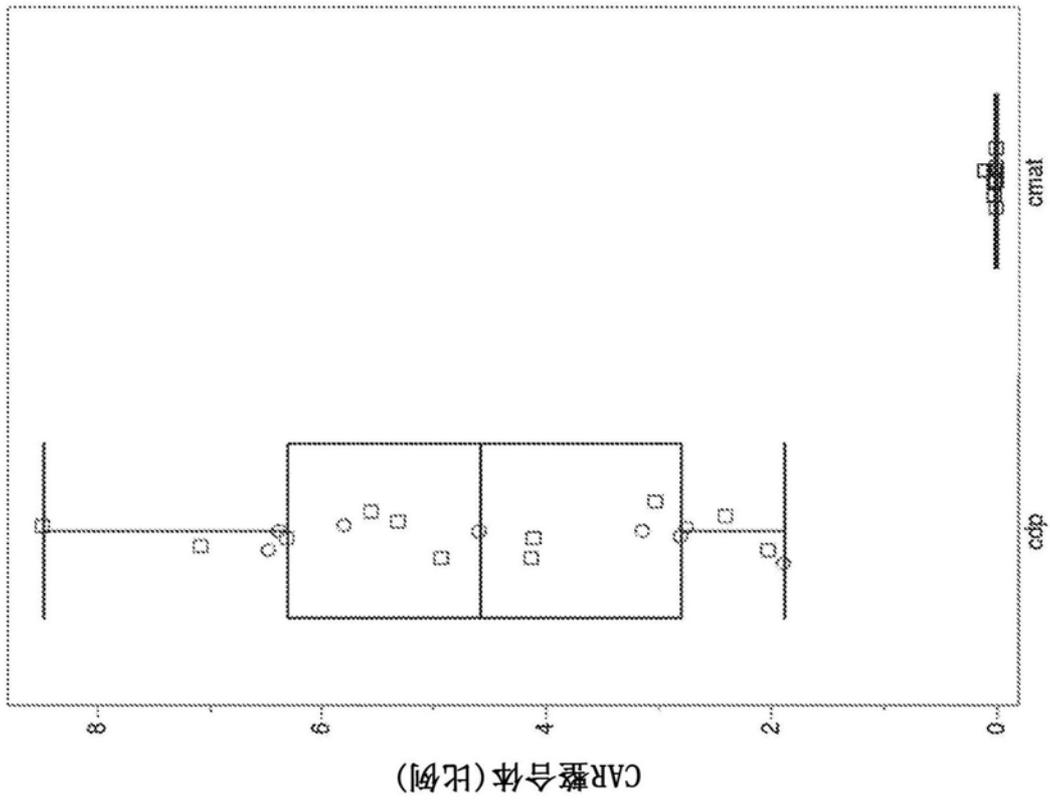


图15A

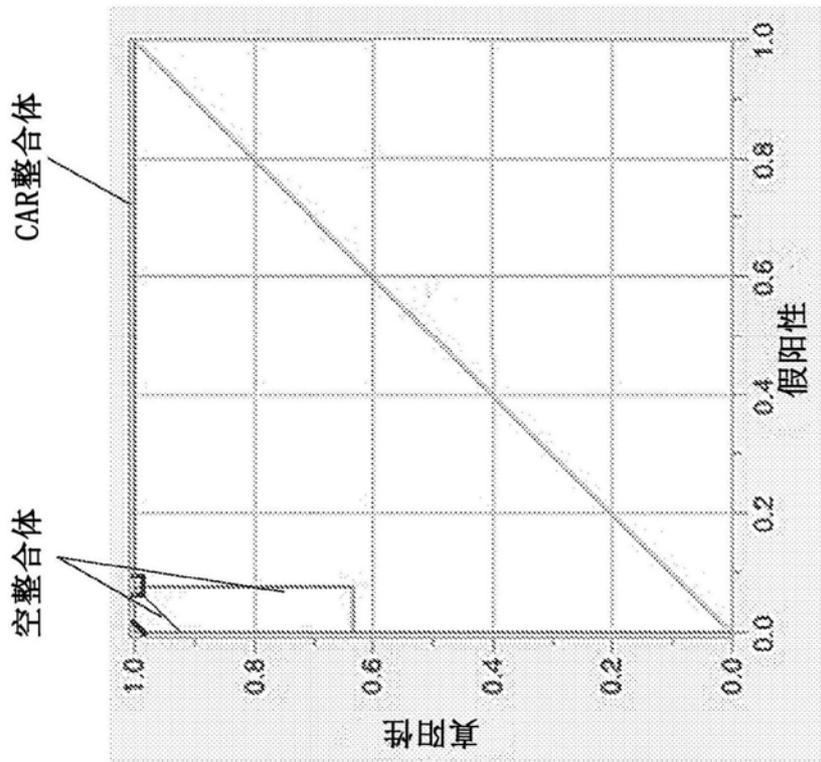


图15B

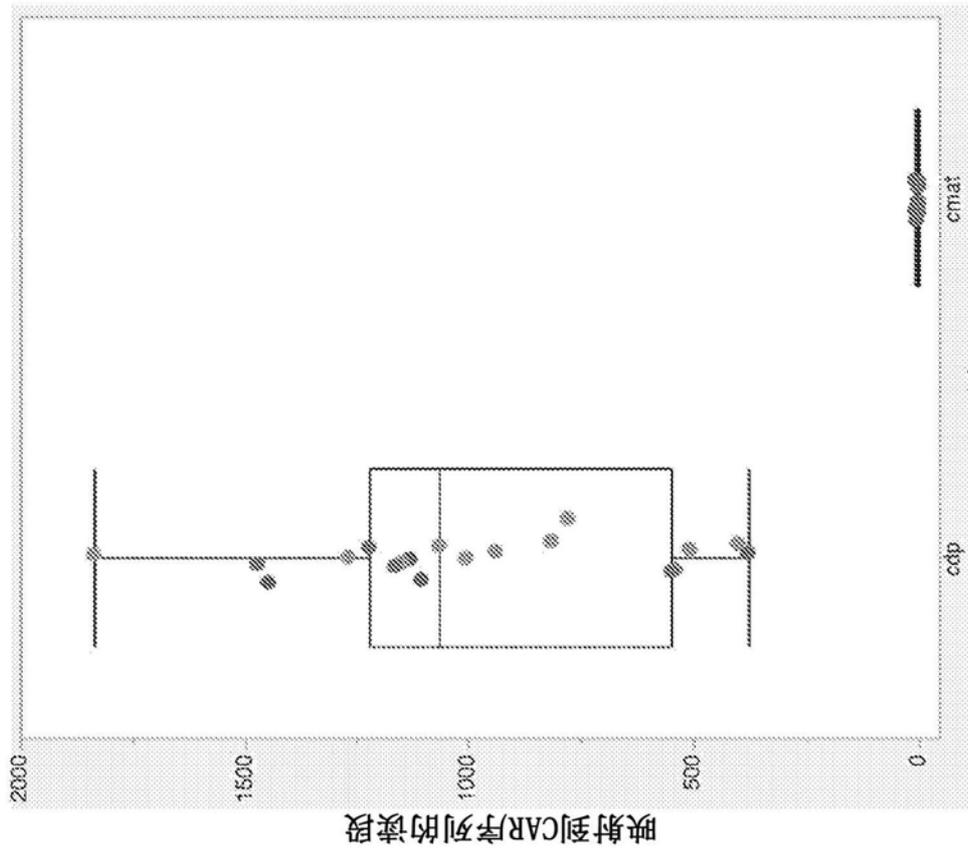


图16A

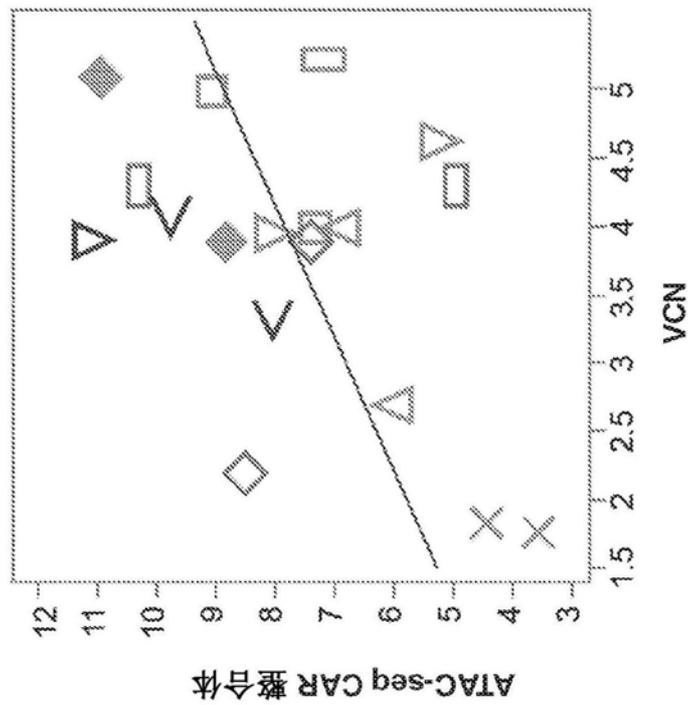


图16B

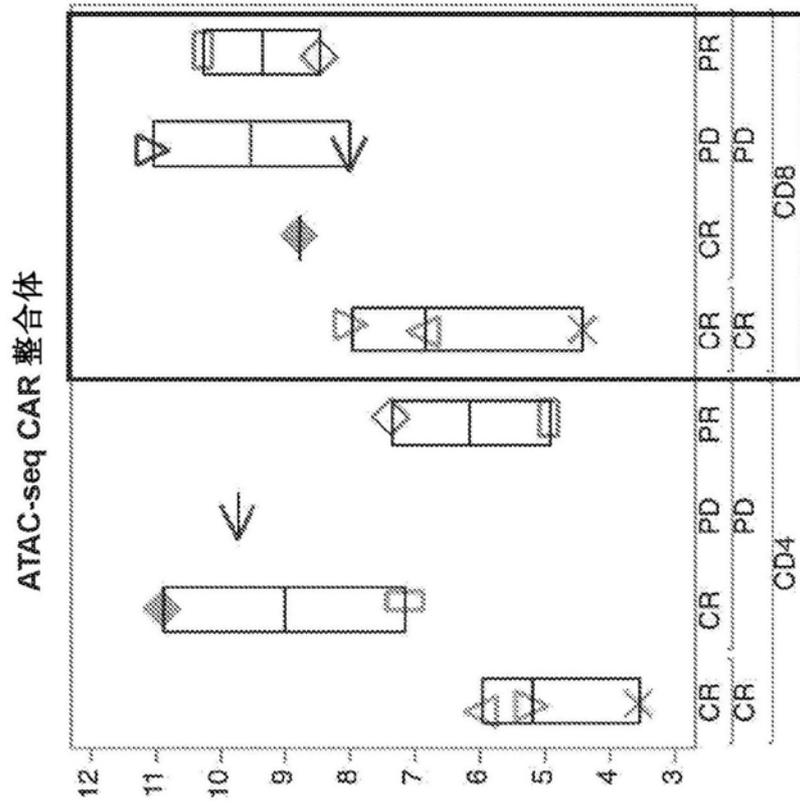


图17B

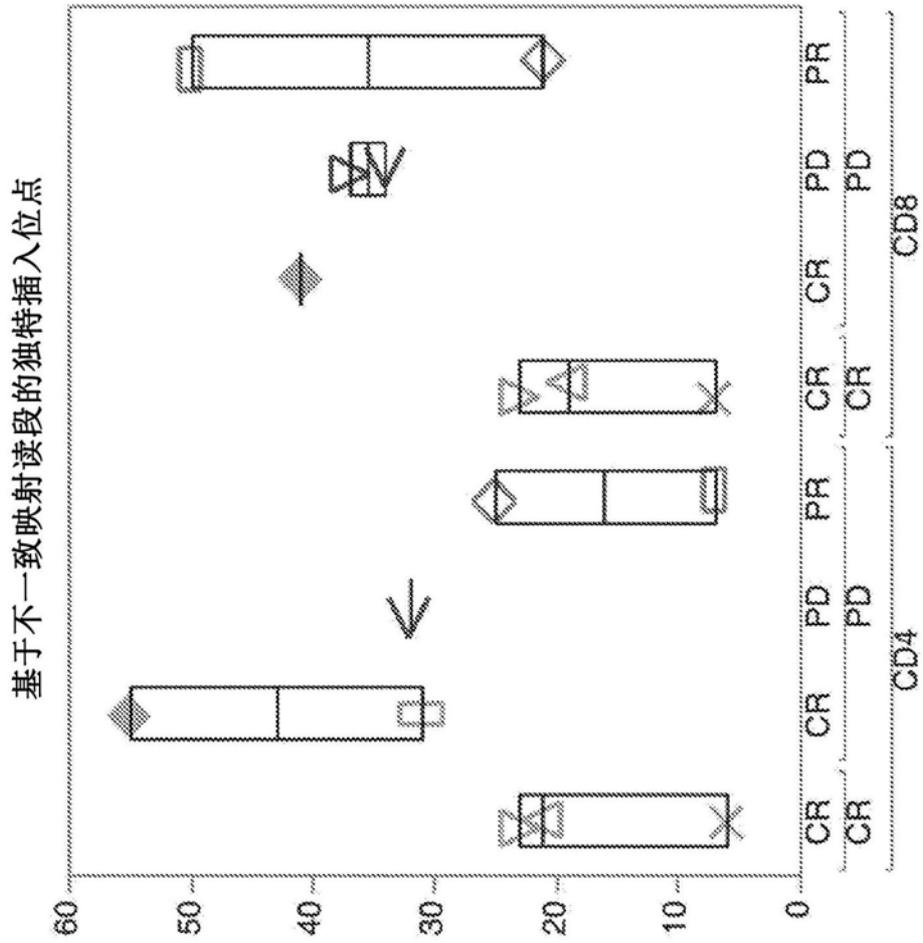


图18

跨越7个不同的CMAT CD8+样品的差异性峰

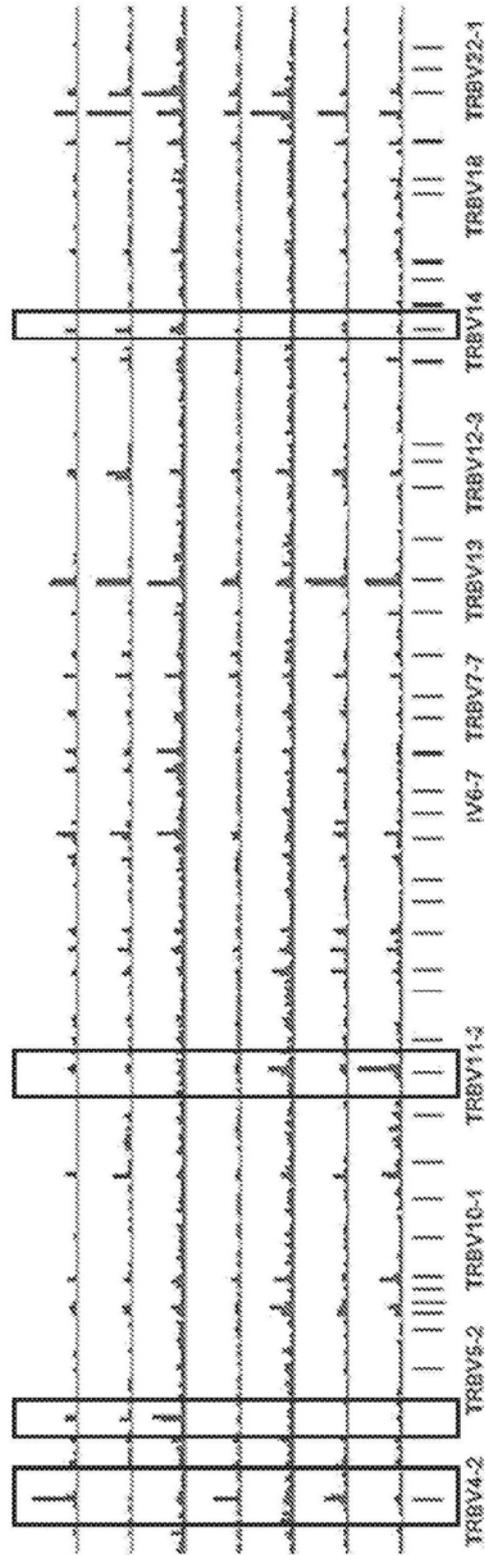


图19A

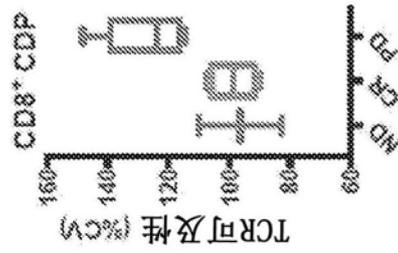


图19B

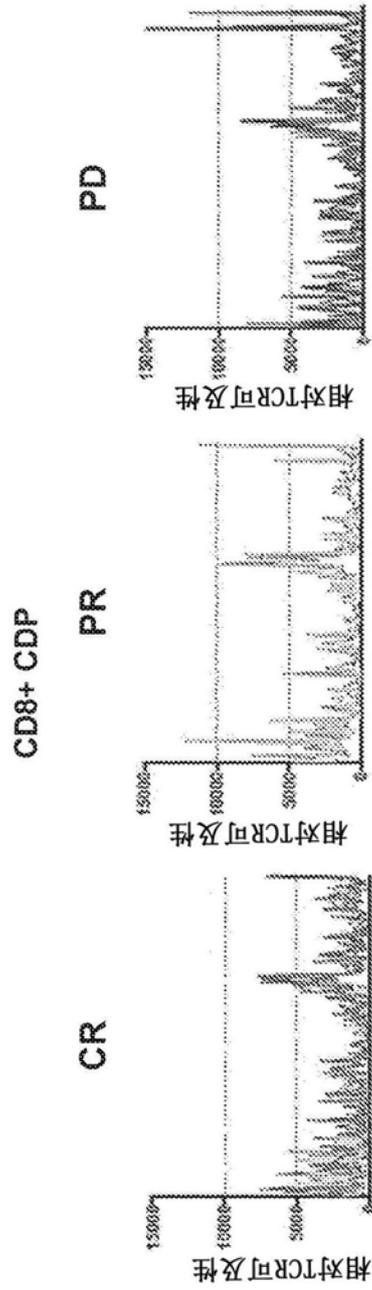


图19C

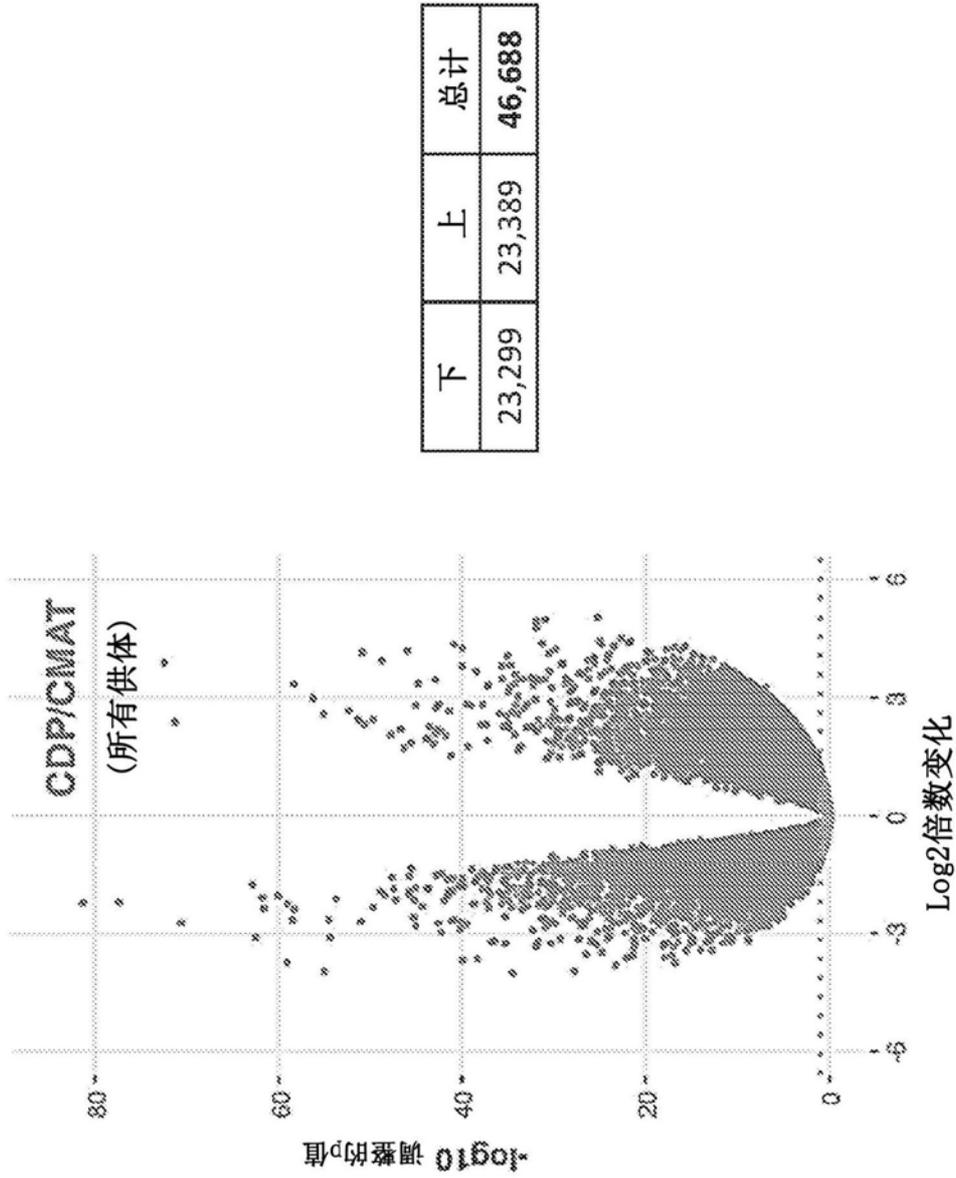


图20A

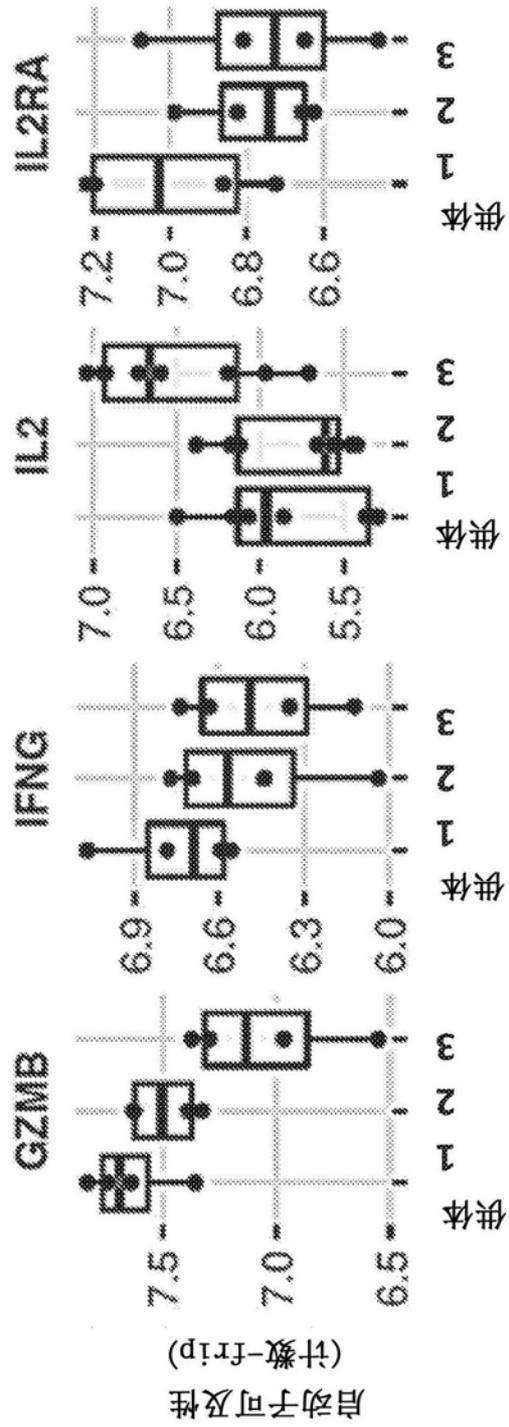


图20B

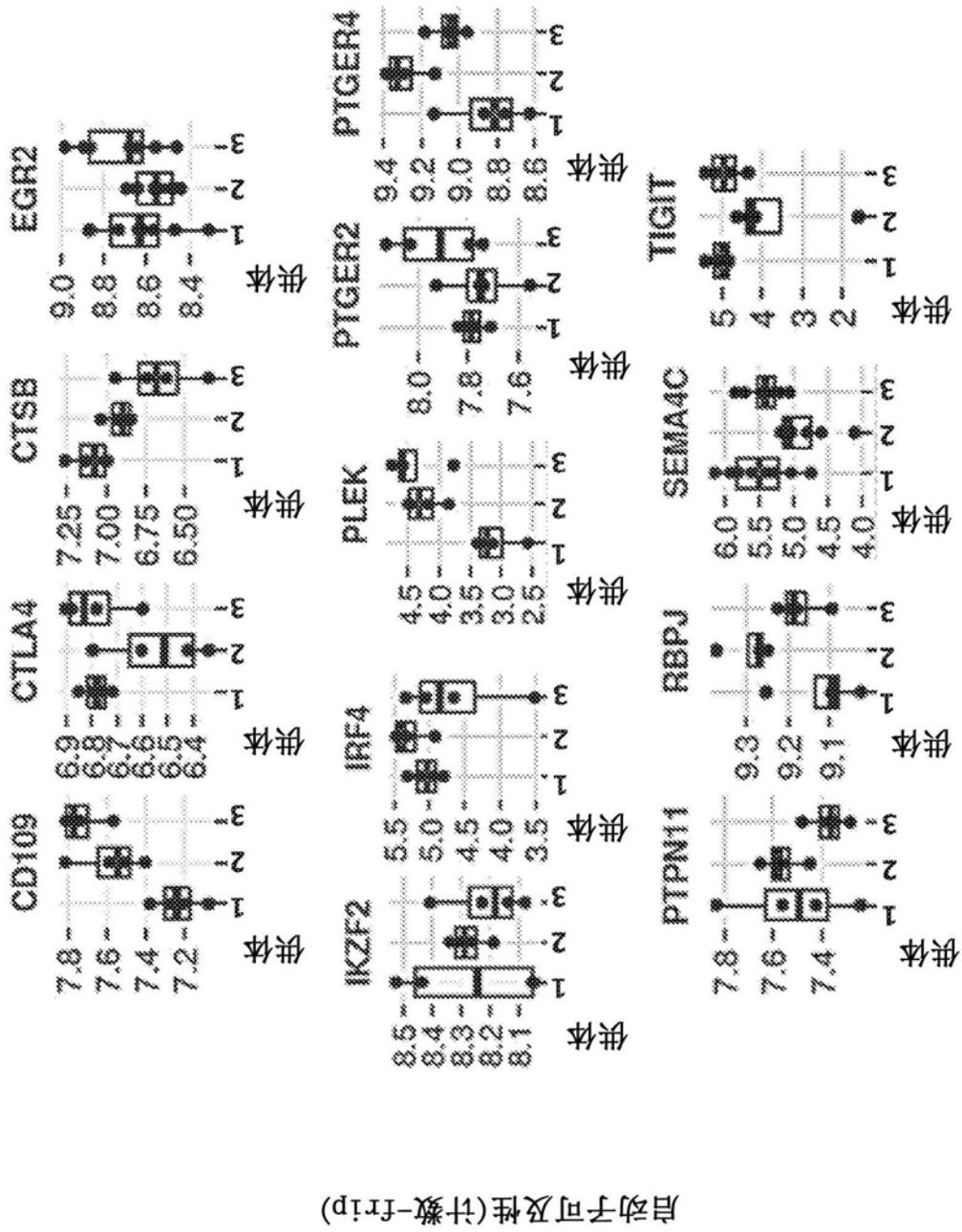


图20C

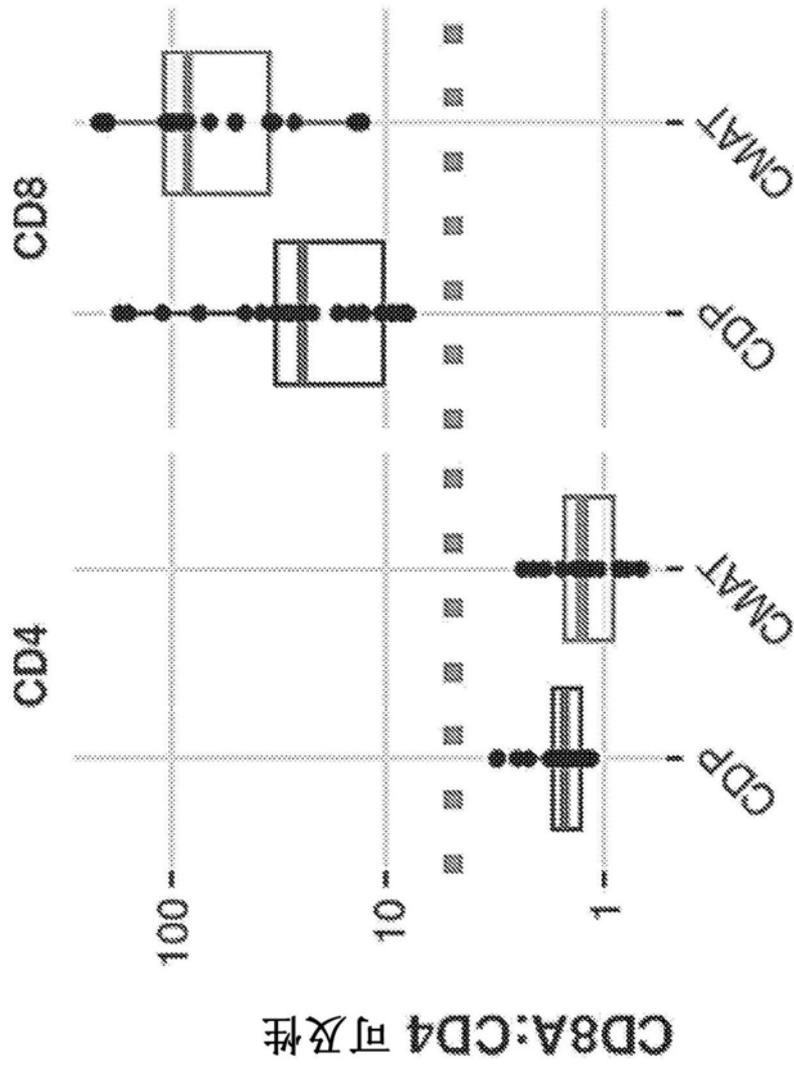


图21

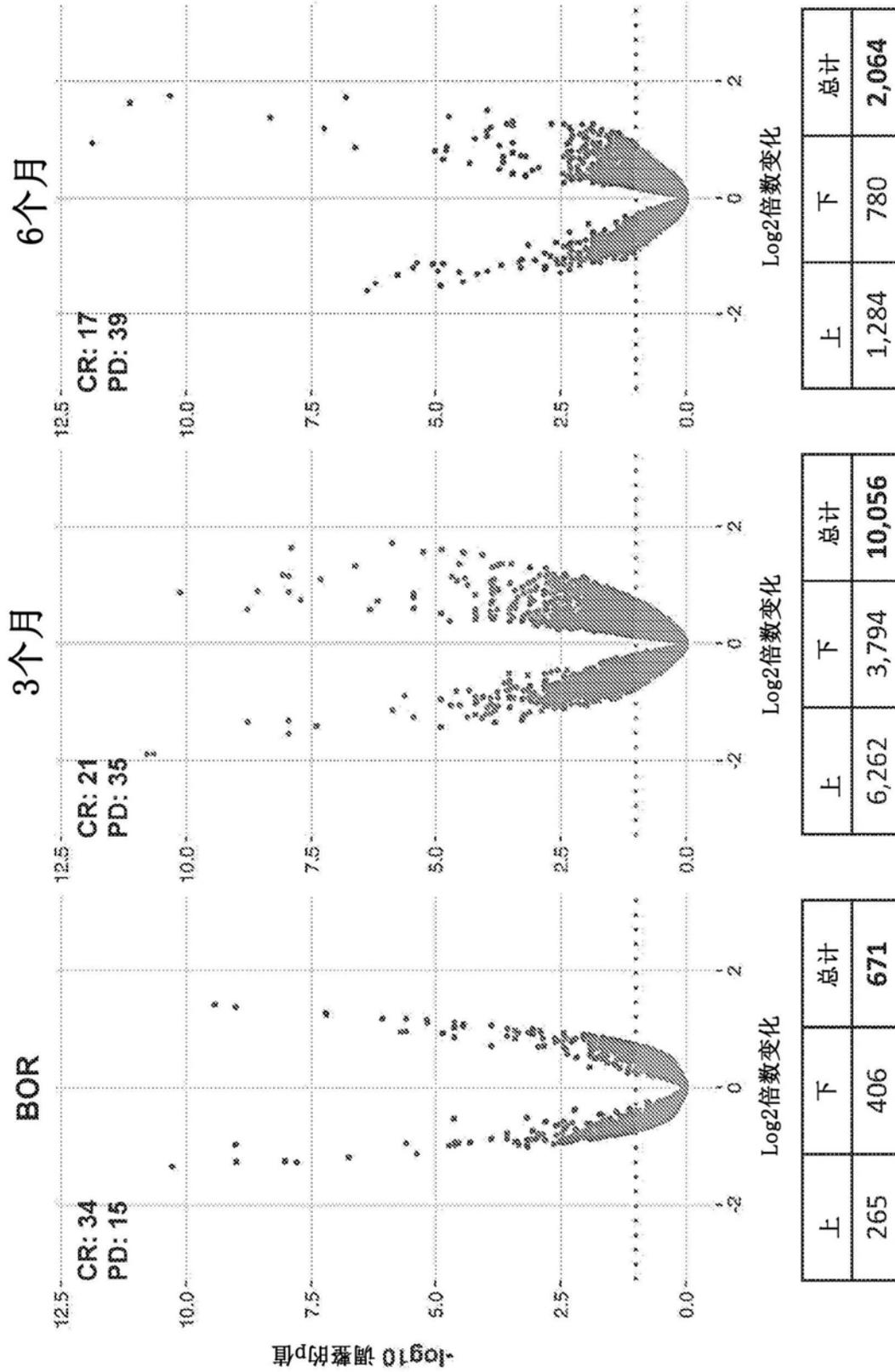


图22A

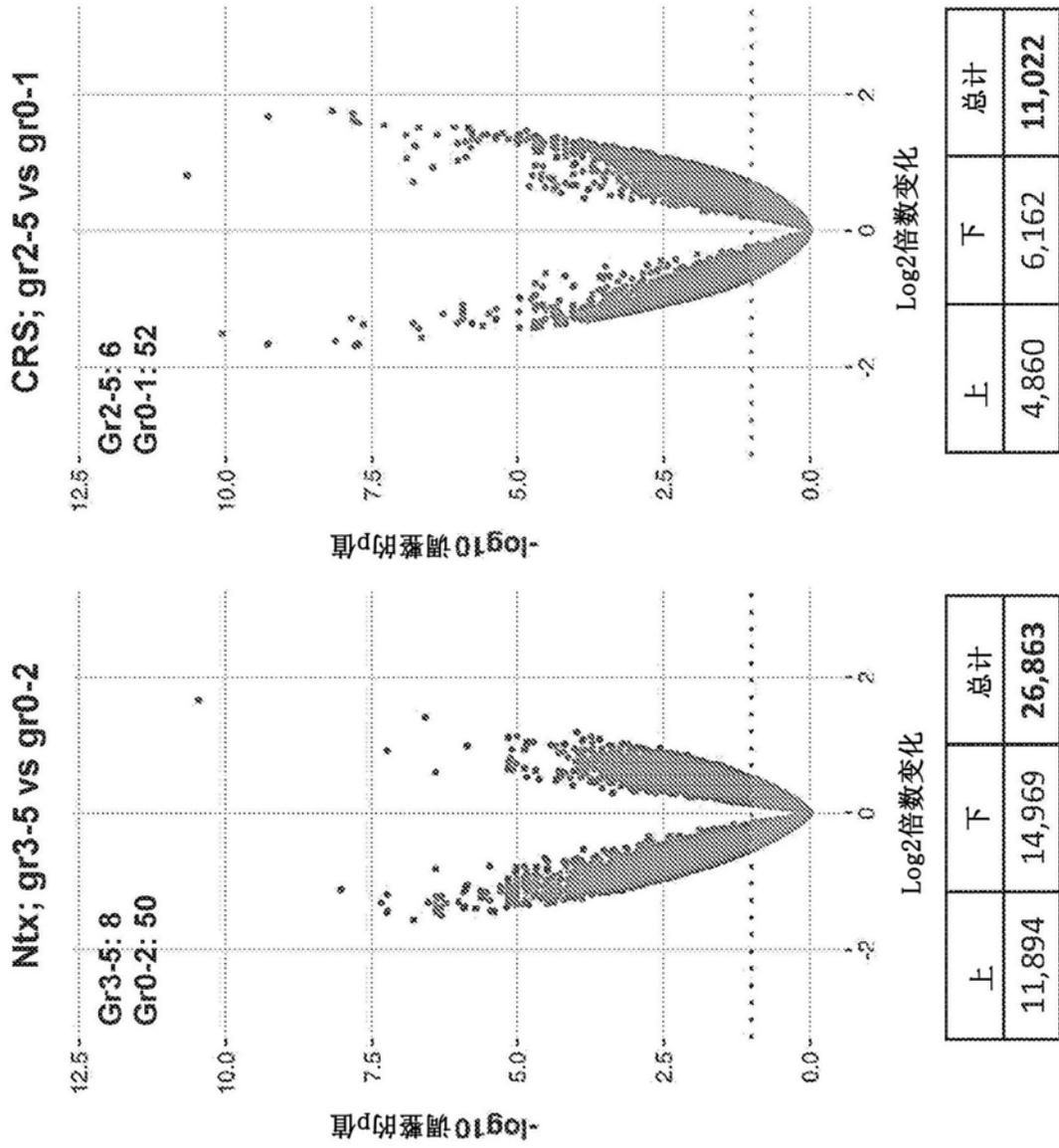


图22B

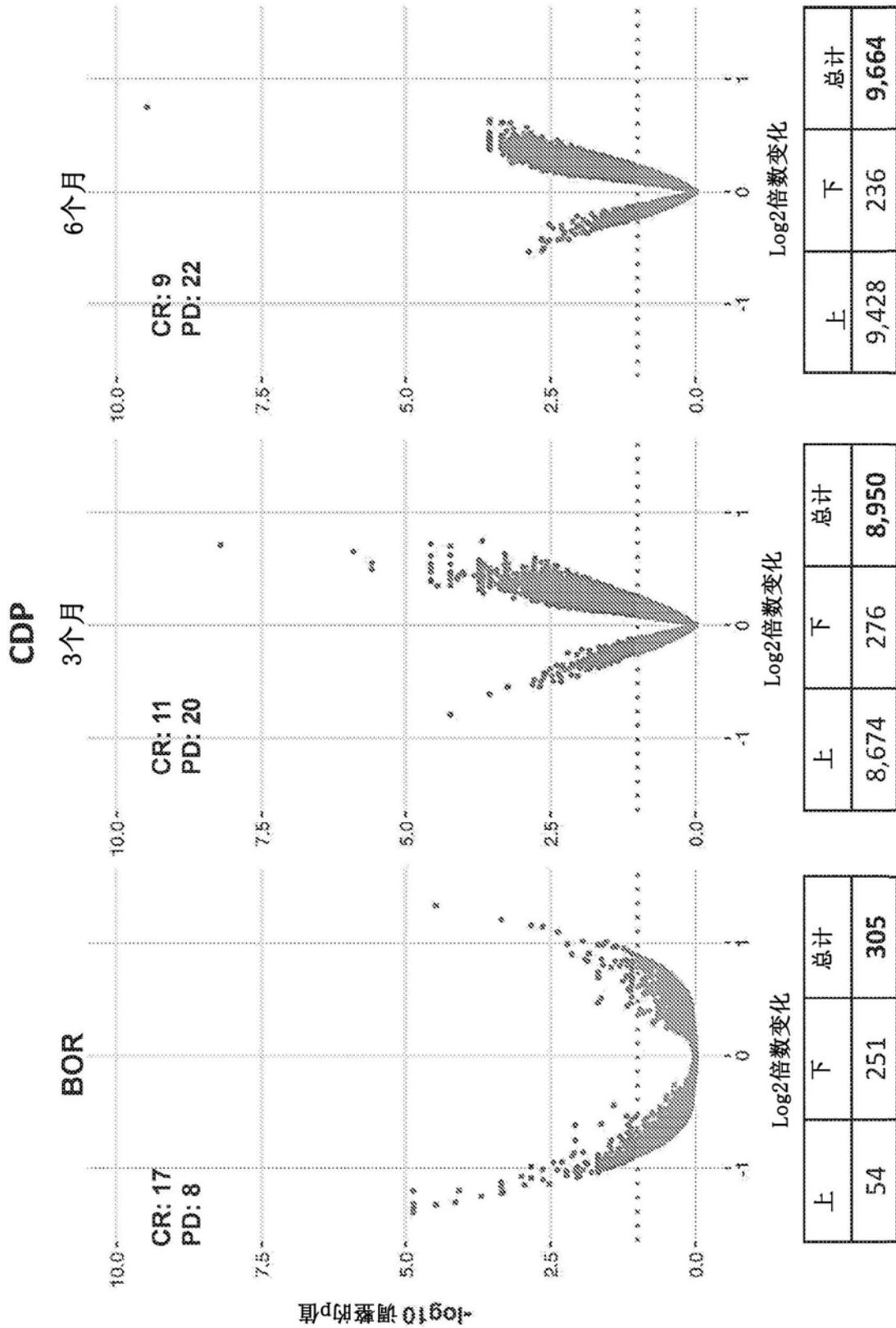


图23A

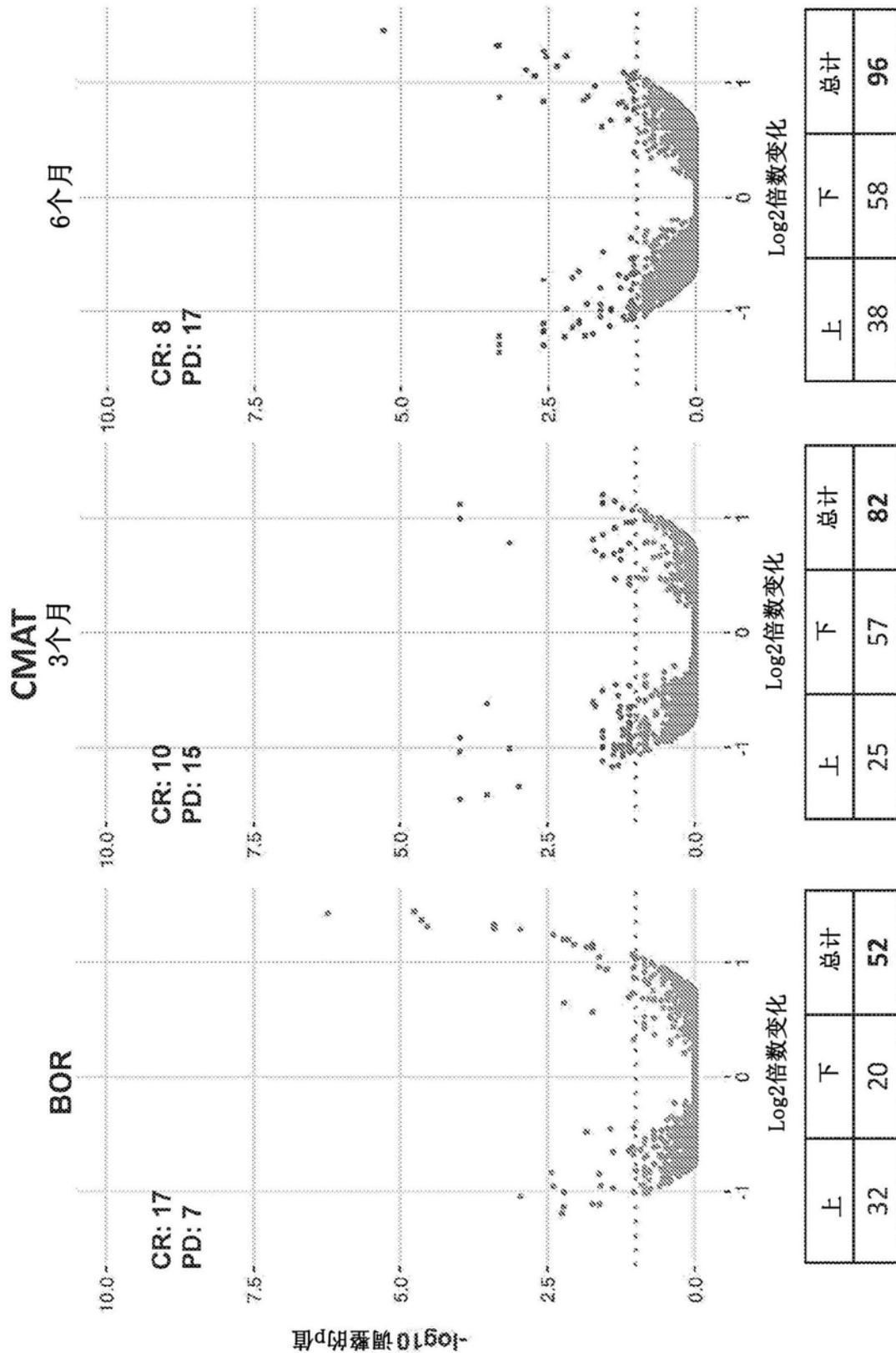


图23B

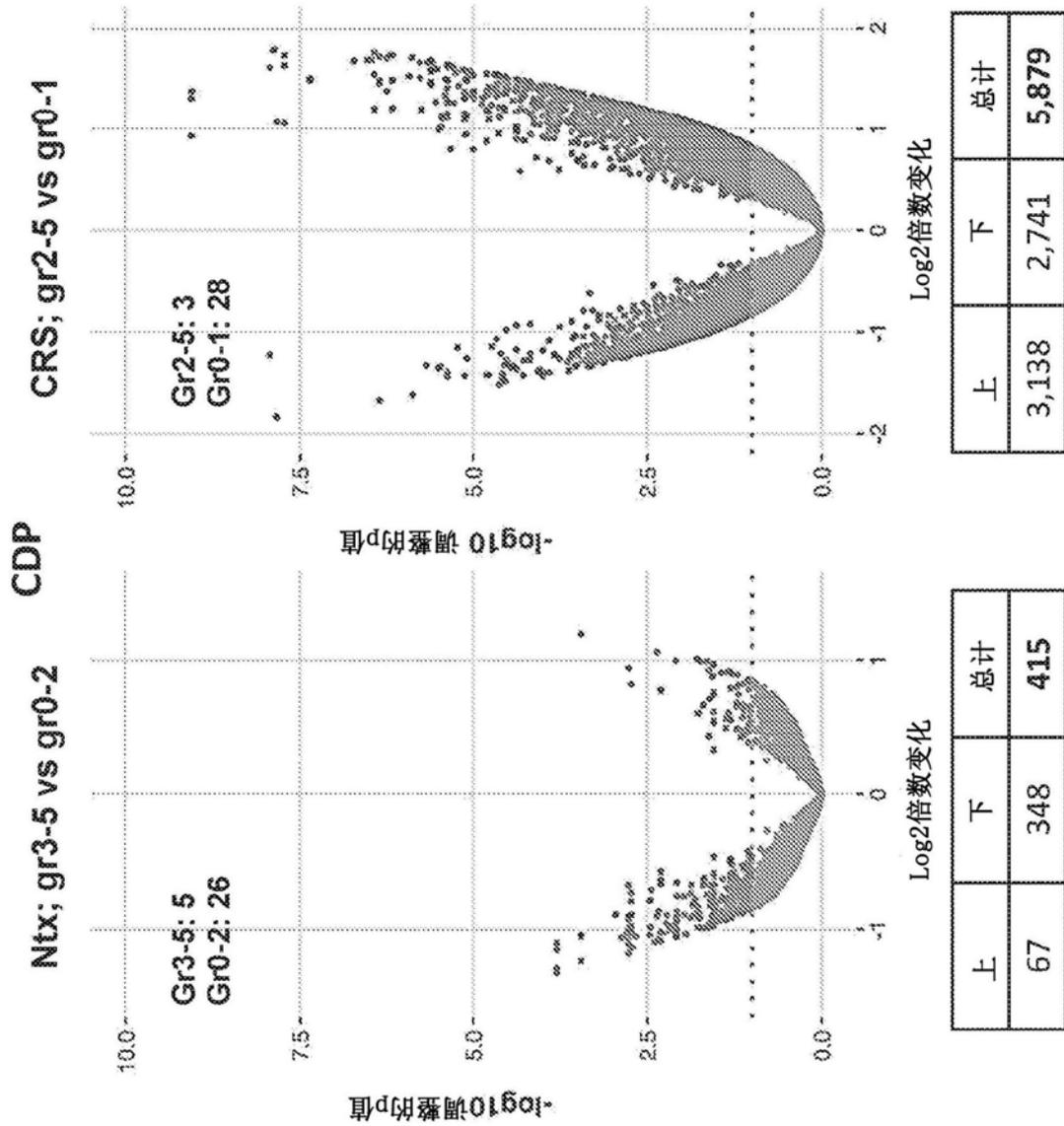


图23C

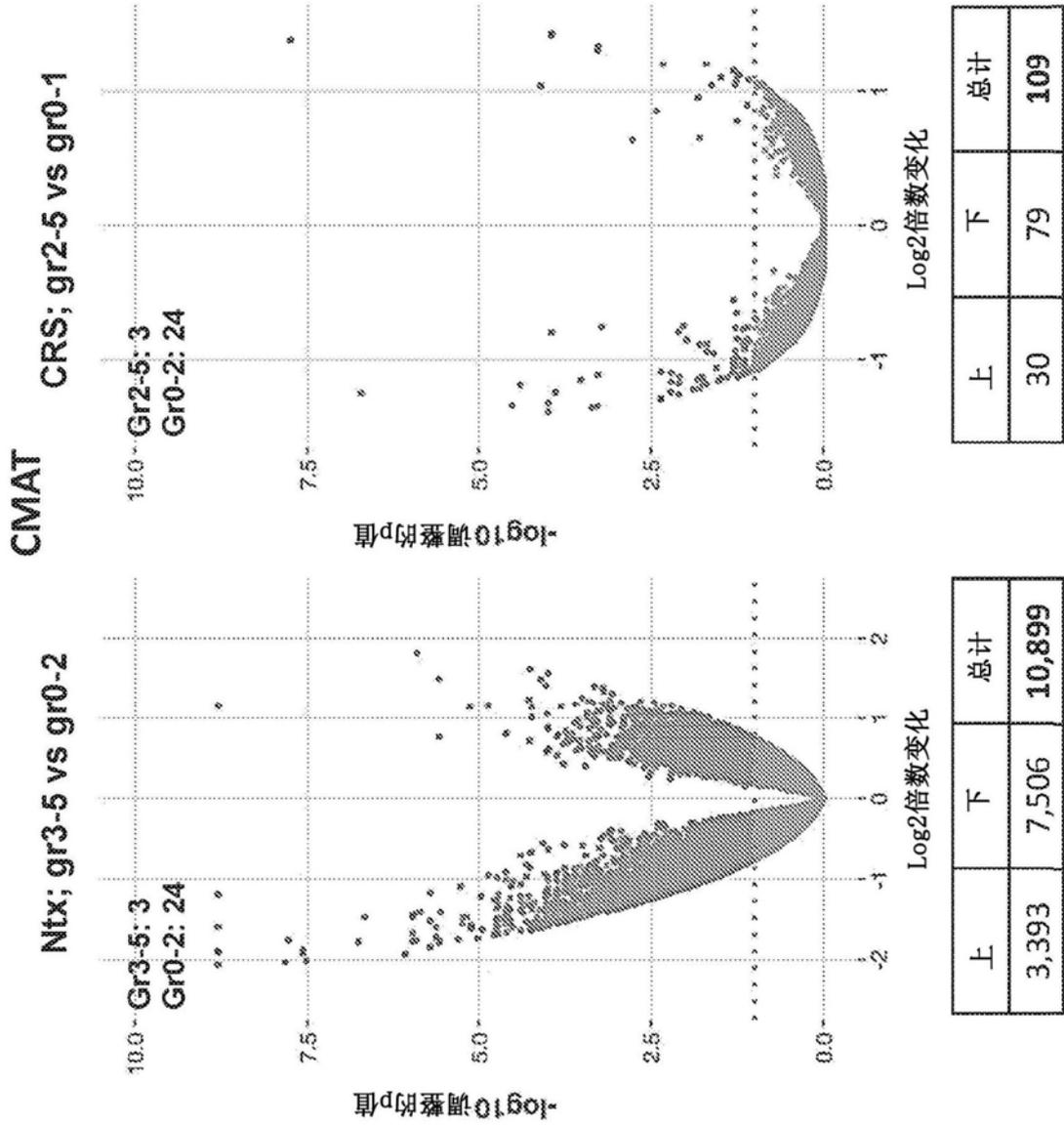


图23D

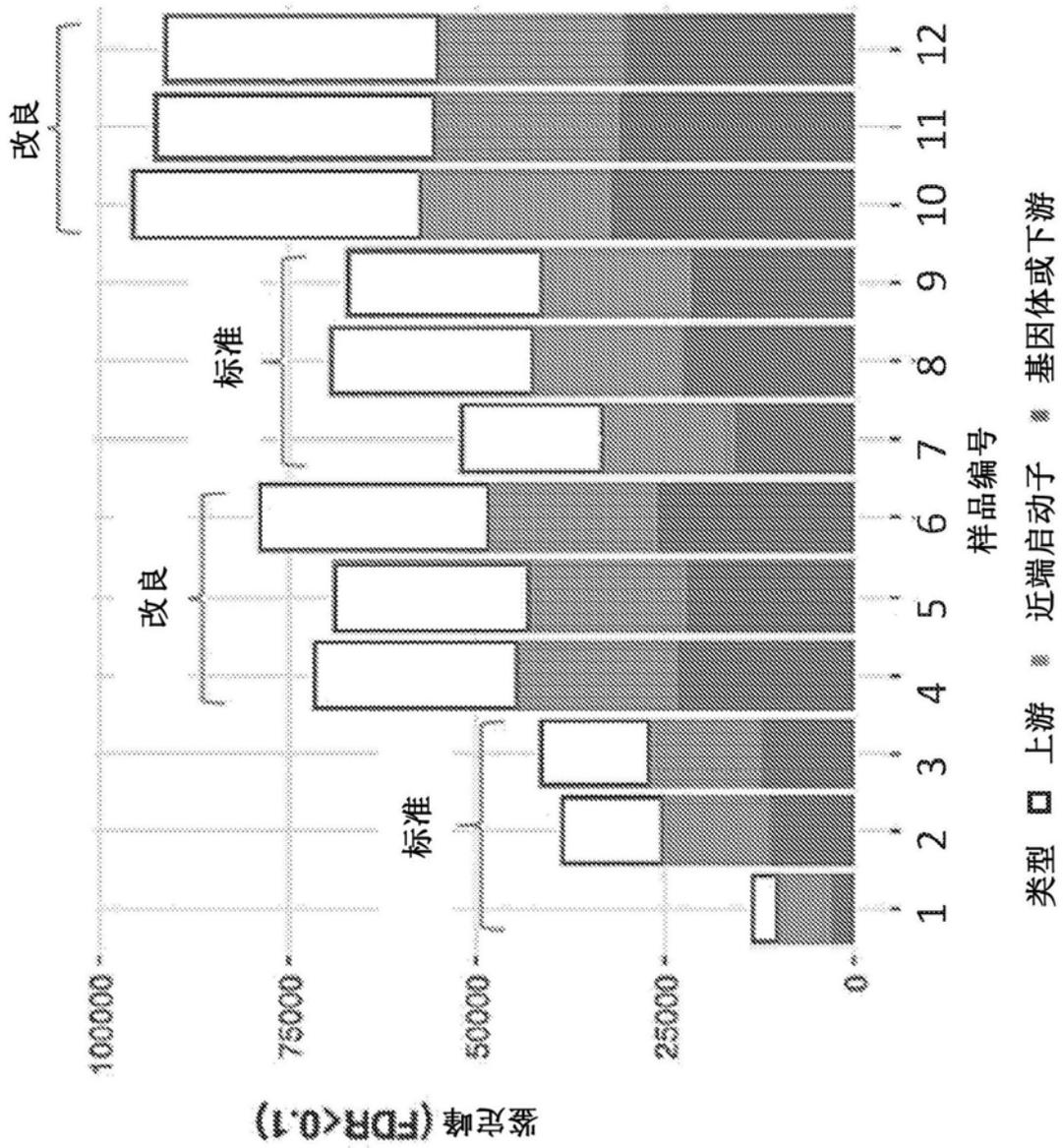


图24A

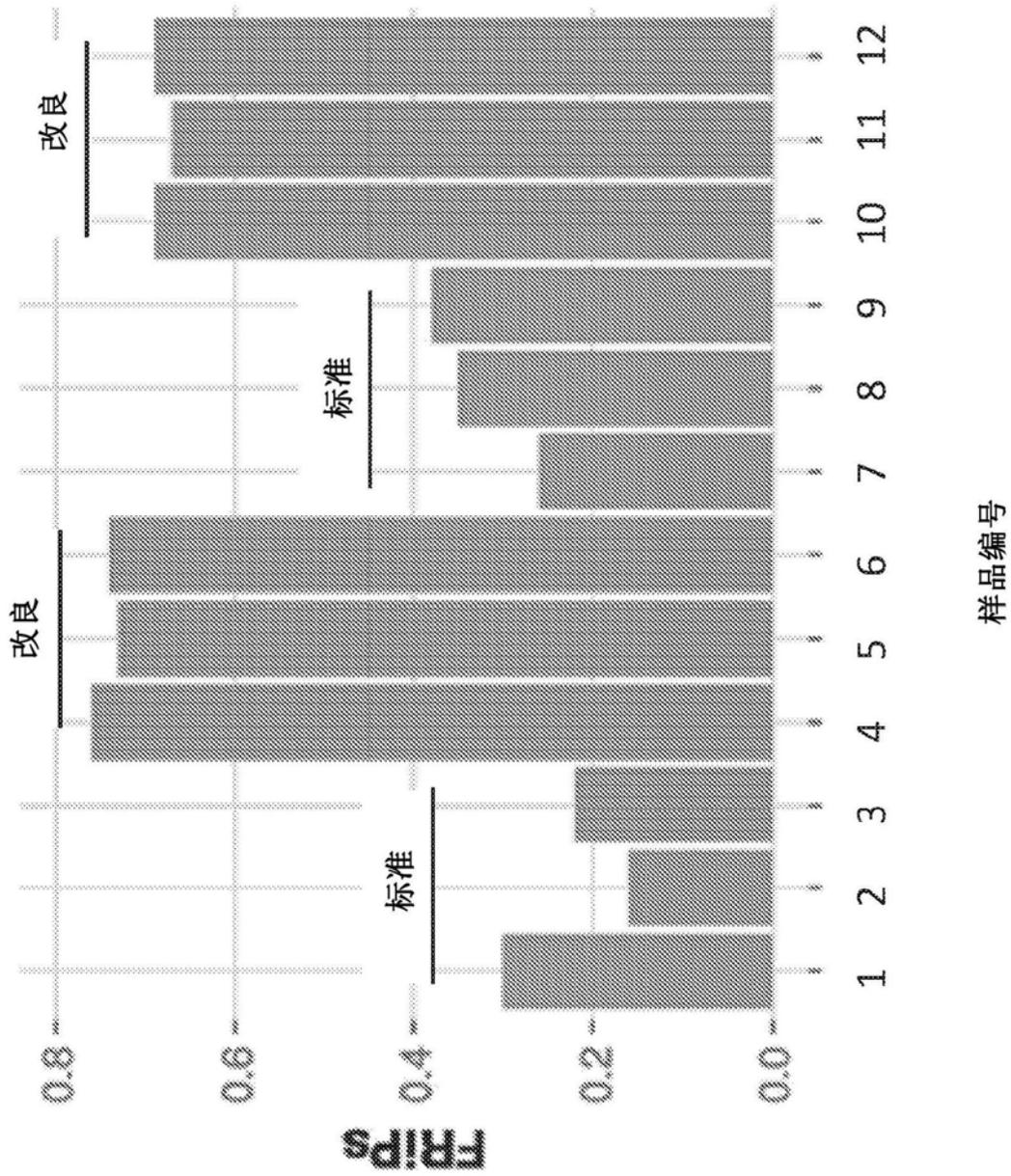


图24B