(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119032179 A
(43)申请公布日 2024.11.26

(21)申请号 202380034708.6
(22)申请日 2023.02.28
(30)优先权数据
63/315,510 2022.03.01 US
(85)PCT国际申请进入国家阶段日
2024.10.17
(86)PCT国际申请的申请数据
PCT/US2023/014125 2023.02.28
(87)PCT国际申请的公布数据
WO2023/167860 EN 2023.09.07
(71)申请人 64-X公司
地址 美国加利福尼亚州
(72)发明人 A·罗夫纳
(74)专利代理机构 北京律盟知识产权代理有限责任公司 11287
专利代理师 赵超
(51)Int.Cl.
C12N 15/864(2006.01)
C12N 5/07(2006.01)
C12N 9/22(2006.01)
C12N 15/113(2006.01)
C12Q 1/6869(2006.01)
C12N 5/10(2006.01)
C12N 15/90(2006.01)

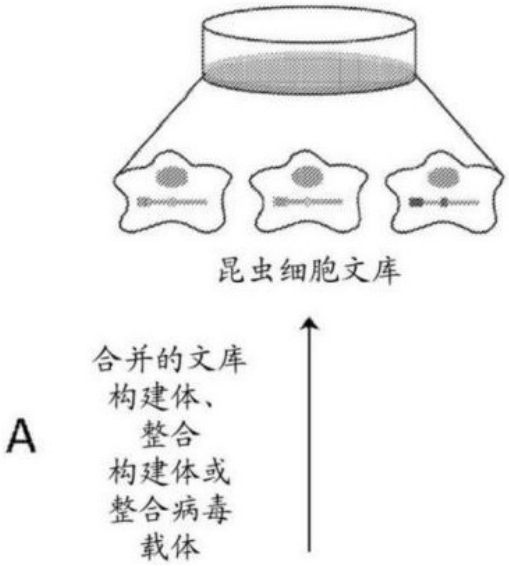
权利要求书3页 说明书98页 附图9页

(54)发明名称

昆虫细胞及其工程化的方法

(57)摘要

本公开提供了用于表达病毒载体的昆虫细胞系,及其制备和使用方法。所提供的方法采用:使用能够被包装到病毒载体中的标识符来选择和/或鉴定具有与病毒载体产生的有益特征相关的工程化序列的昆虫细胞系。示例性病毒载体包括AAV载体。



CN 119032179 A

1.一种方法,其包括:

a.从昆虫细胞文库产生多个腺相关病毒(AAV)载体,其中所述昆虫细胞文库包含多个昆虫细胞,其中每个昆虫细胞单独包含:

(i)包含位于两个功能性AAV ITR序列之间的条形码的核酸序列,

(ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

(iii)RNA引导的核酸酶,和

(iv)一个或多个产生AAV载体所必需的核酸序列,

其中所述多个昆虫细胞产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的标识符的标识符;以及

b.检测所述多个AAV载体的条形码。

2.一种选择用于产生腺相关病毒(AAV)载体的昆虫细胞的方法,所述方法包括:

a.用一个或多个产生AAV载体所必需的序列转染昆虫细胞文库,其中所述多个中的每个昆虫细胞单独包含:

(i)包含位于两个功能性AAV ITR序列之间的条形码的核酸序列,

(ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,和

(iii)RNA引导的核酸酶,

b.从所述昆虫细胞文库产生多个AAV载体,其中所述AAV载体各自包含对应于产生它的所述单独昆虫细胞的条形码;以及

c.检测所述多个AAV载体的条形码,其中所述条形码的丰度与所述昆虫细胞产生和/或分泌AAV载体相关。

3.如权利要求1或2所述的方法,其中所述一个或多个产生AAV载体所必需的序列是在杆状病毒载体的背景下。

4.如权利要求1至3中任一项所述的方法,其还包括确定特定条形码相对于存在于从所述昆虫细胞文库产生的所述多个AAV载体中的所有条形码的相对丰度的步骤。

5.如权利要求1至4中任一项所述的方法,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞,所述一个或多个扰动与AAV产生和/或AAV分泌的增加相关。

6.如权利要求1至5中任一项所述的方法,其中至少一个文库变体被整合到位于一对顺式作用整合序列之间的昆虫细胞基因组中。

7.如权利要求1至6中任一项所述的方法,其中所述工程化核酸序列的一个或两个拷贝被整合到所述昆虫细胞基因组中。

8.如权利要求1至7中任一项所述的方法,其中至少一个文库变体游离存在于所述昆虫细胞中。

9.如权利要求1至8中任一项所述的方法,其中所述两个功能性AAV ITR序列包括人AAV1 ITR、人AAV2 ITR、人AAV3b ITR、人AAV4 ITR、人AAV5 ITR、人AAV6 ITR、人AAV7 ITR、人AAV8 ITR、人AAV9 ITR、人AAV10 ITR、人AAV11 ITR、人AAV12 ITR或人AAV13 ITR。

10.如权利要求1至8中任一项所述的方法,其中所述两个功能性AAV ITR序列包括牛AAV(b-AAV)ITR、犬AAV(CAAV)ITR、小鼠AAV1 ITR、山羊AAV ITR、大鼠AAV ITR或禽AAV(AAAV)ITR。

11.如权利要求1至10中任一项所述的方法,其中所述一个或多个产生AAV载体所必需的序列包含(a)AAV Rep基因,(b)AAV Cap基因,(c)一个或多个AAV辅助基因;或(d)其组合。

12.如权利要求11所述的方法,其中所述一个或多个产生AAV载体所必需的序列包含编码人AAV1衣壳蛋白、人AAV2衣壳蛋白、人AAV3b衣壳蛋白、人AAV4衣壳蛋白、人AAV5衣壳蛋白、人AAV6衣壳蛋白、人AAV7衣壳蛋白、人AAV8衣壳蛋白、人AAV9衣壳蛋白、人AAV10衣壳蛋白、人AAV11衣壳蛋白、人AAV12衣壳蛋白,或人AAV13衣壳蛋白的AAV Cap基因。

13.如权利要求1至12中任一项所述的方法,其中所述一个或多个扰动包括内源性基因组编码序列的插入、缺失、取代、替换、表观遗传修饰和/或重排。

14.如权利要求1至13中任一项所述的方法,其中所述一个或多个文库变体包含至少两个文库变体,其中所述至少两个文库变体包含至少一个独特基因、至少一个独特ORF、至少一个独特gRNA序列,和/或至少一个独特非编码核酸,或其组合和/或多个。

15.如权利要求1至14中任一项所述的方法,其中所述方法还包括单细胞测序。

16.如权利要求1至15中任一项所述的方法,其中所述昆虫细胞文库包括Sf21细胞、Sf9细胞、BTI-TN-5B1-4(High Five)细胞、S2细胞、D.Mel2细胞或其任意的衍生物。

17.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

(i)包含位于两个功能性AAV ITR序列之间的条形码的工程化核酸序列,其中所述核酸序列游离存在于所述昆虫细胞内,

(ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

(iii)RNA引导的核酸酶,和

(iv)一个或多个产生AAV载体所必需的核酸序列,

其中所述昆虫细胞群产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的条形码的条形码。

18.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

(i)工程化核酸序列,其包含位于两个功能性AAV ITR序列之间的条形码,其中所述核酸序列被整合到位于一对顺式作用整合序列之间的昆虫细胞基因组中,

(ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

(iii)RNA引导的核酸酶,和

(iv)一个或多个产生AAV载体所必需的核酸序列,

其中所述昆虫细胞群产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的条形码的条形码。

19.如权利要求17或18所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包括Sf21细胞、Sf9细胞、BTI-TN-5B1-4(High Five)细胞、S2细胞、D.Mel2细胞或其任意的衍生物。

20.如权利要求17至19中任一项所述的昆虫细胞或昆虫细胞群,其中所述AAV载体包含人AAV1衣壳蛋白;人AAV2衣壳蛋白;人AAV3b衣壳蛋白;人AAV4衣壳蛋白;人AAV5衣壳蛋白;人AAV6衣壳蛋白;人AAV7衣壳蛋白;人AAV8衣壳蛋白;人AAV9衣壳蛋白;人AAV10衣壳蛋白;人AAV11衣壳蛋白;人AAV12衣壳蛋白;或人AAV13衣壳蛋白。

21.如权利要求17至19中任一项所述的昆虫细胞或昆虫细胞群,其中所述AAV载体包含人祖先AAV衣壳蛋白。

22.如权利要求17至21中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体包括AAV载体,其中所述AAV载体包含第一组两个病毒重复序列,所述第一组两个病毒重复序列包含一对反向末端重复序列(ITR),其是或包括人AAV1 ITR;人AAV2 ITR;人AAV3b ITR;人AAV4 ITR;人AAV5 ITR;人AAV6 ITR;人AAV7 ITR;人AAV8 ITR;人AAV9 ITR;人AAV10 ITR;人AAV11 ITR;人AAV12 ITR;或人AAV13 ITR。

23.如权利要求17至21中任一项所述的昆虫细胞或昆虫细胞群,其中所述AAV载体包含牛AAV(b-AAV)衣壳蛋白;犬AAV(CAAV)衣壳蛋白;小鼠AAV1衣壳蛋白;山羊AAV衣壳蛋白;大鼠AAV衣壳蛋白;或禽AAV(AAAV)衣壳蛋白。

24.如权利要求17至23中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体包括AAV载体,其中所述AAV载体包含一对ITR,其是或包括牛AAV(b-AAV)ITR;犬AAV(CAAV)ITR;小鼠AAV1 ITR;山羊AAV ITR;大鼠AAV ITR;或禽AAV(AAAV)ITR。

25.如权利要求17至24中任一项所述的昆虫细胞或昆虫细胞群,其中至少一个包含一个或多个形成病毒载体所必需的核酸序列的多核苷酸包含:

a.AAV Rep基因;

b.AAV Cap基因;

c.一种或多种AAV辅助基因;或

d.其组合。

26.如权利要求17至25中任一项所述的昆虫细胞群,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞群,所述一个或多个扰动与AAV产生和/或AAV分泌的增加相关。

27.如权利要求17至26中任一项所述的昆虫细胞或昆虫细胞群,其与缺乏所述一个或多个扰动的昆虫细胞产生的AAV群相比产生了包含至少一个改进的特征的AAV载体群。

28.如权利要求1至16中任一项所述的方法或如权利要求17至27中任一项所述的昆虫细胞或昆虫细胞群,其中所述至少一个条形码包含约5至约25个核苷酸的序列。

29.如权利要求1至16中任一项所述的方法或如权利要求17至27中任一项所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包含一个、两个、三个、四个、五个、六个、七个、八个、九个或十个文库变体。

30.如权利要求1至16中任一项所述的方法或如权利要求17至27中任一项所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包含多至100个文库变体。

昆虫细胞及其工程化的方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2022年3月1日提交的美国临时专利申请号63/315,510的优先权,所述申请的全部内容通过引用并入本文。

背景技术

[0003] 生物制剂市场诸如像基因和细胞疗法正在扩大,但目前承受着高制造成本和与扩大生产相关的困难的负担。持续需要改善的细胞系和分离它们的方法,用于包括但不限于生物制剂生产和病毒载体生产的应用。

发明内容

[0004] 本公开尤其提供了用于在昆虫细胞中产生和/或表达病毒载体的方法、系统和组合物。本公开认识到,用于在昆虫细胞中表达病毒载体的现有技术承受着低效的病毒生产和低效的分离优化的昆虫细胞系的筛选的负担。本公开提供了用于工程化昆虫细胞和/或病毒载体以获得改变的与病毒载体生产相关的特征和其他特征的平台技术。在一些实施方案中,提供的方法使产生和/或选择具有改善的病毒载体表达和/或生产特征(例如,增加的病毒载体表达、增加的表达持续时间、增加的稳定性等)的昆虫细胞系成为可能。在一些实施方案中,方法包括筛选由具有标识符的昆虫细胞的文库产生的病毒载体。提供的技术包括令人惊讶的见解,即使病毒载体获取表达其的昆虫细胞的标识符(例如,包括条形码序列和/或文库变体),从而使文库中细胞的病毒载体产生能力的有效评估、表征和/或鉴定成为可能。例如,在一些实施方案中,昆虫细胞群各自用文库构建体转化,所述文库构建体包含标识符,具有适合于将标识符包装到病毒载体中的结构(例如,对于AAV载体,标识符可以位于AAV ITR之间)。在一些实施方案中,可以通过丰富的独特标识符来筛选、选择和/或表征由此类昆虫细胞群产生的病毒载体池。

[0005] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。

[0006] 在一些实施方案中,由昆虫细胞产生的病毒载体包含与产生病毒载体的昆虫细胞的标识符相同的标识符。在一些实施方案中,病毒载体的标识符源自产生病毒载体的昆虫细胞的标识符。在一些实施方案中,病毒载体的标识符对应于产生病毒载体的昆虫细胞的标识符。在一些实施方案中,病毒载体的标识符与产生病毒载体的昆虫细胞的标识符至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多相同。

[0007] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,(ii)至少一个扰动,和(iii)至少一个包含一个或多个产生病毒载体所必需

的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。在一些实施方案中,昆虫细胞还包含有效载荷和/或至少一个文库变体。在一些实施方案中,昆虫细胞还包含至少一个扰动附加序列。

[0008] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,(ii)至少一个文库变体,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。在一些实施方案中,昆虫细胞还包含有效载荷和/或至少一个扰动。在一些实施方案中,昆虫细胞还包含至少一个扰动附加序列。

[0009] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,(ii)至少一个有效载荷,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。在一些实施方案中,昆虫细胞还包含至少一个文库变体和/或至少一个扰动。在一些实施方案中,昆虫细胞还包含至少一个扰动附加序列。

[0010] 在一些实施方案中,昆虫细胞还包含至少一个反式作用整合序列和/或至少一个顺式作用整合序列。

[0011] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符;(ii)至少一个文库变体和/或扰动;(iii)至少一个有效载荷;(iv)至少一个扰动附加序列;(v)至少一个反式作用整合序列和/或至少一个顺式作用整合序列;和(vi)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。

[0012] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符;(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸;以及任选的(iii)一个或多个包含至少一个扰动、至少一个文库变体、至少一个有效载荷、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列的工程化序列,并且其中昆虫细胞产生包含至少一个标识符的病毒载体。

[0013] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)文库构建体,和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中所述文库构建体包含标识符,其中所述标识符位于能够包装到病毒载体中的两个病毒重复序列之间,并且其中昆虫细胞产生包含所述标识符的病毒载体。在一些实施方案中,昆虫细胞任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动、至少一个文库变体、至少一个有效载荷、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列,并且其中昆虫细胞产生包含至少一个标识符的病毒载体。

[0014] 在一些实施方案中,由昆虫细胞产生的病毒载体包含与文库构建体相同的标识符。在一些实施方案中,病毒载体的标识符源自相关文库构建体的标识符。在一些实施方案

中,病毒载体的标识符对应于文库构建体的标识符。在一些实施方案中,病毒载体的标识符与文库构建体的标识符至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多相同。

[0015] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)文库构建体,和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中所述文库构建体包含标识符,其中所述标识符位于能够包装到病毒载体中的两个病毒重复序列之间,并且其中昆虫细胞产生包含所述标识符的病毒载体。在一些实施方案中,文库构建体任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含至少一个文库变体和/或至少一个有效载荷。在一些实施方案中,昆虫细胞任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列,并且其中昆虫细胞产生包含至少一个标识符的病毒载体。

[0016] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,所述多个昆虫细胞各自单独地包含一个或多个工程化序列,其中所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,并且其中昆虫细胞群产生单独地包含至少一个标识符的病毒载体。在一些实施方案中,昆虫细胞群任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动、至少一个文库变体、至少一个有效载荷、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列。

[0017] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,所述多个昆虫细胞各自单独地包含一个或多个工程化序列,其中所述一个或多个工程化序列共同包含:(i)至少一个位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,(ii)至少一个文库变体和/或扰动,(iii)至少一个有效载荷,和(iv)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞群产生单独地包含至少一个标识符的病毒载体。在一些实施方案中,昆虫细胞任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列。

[0018] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,所述多个昆虫细胞各自单独地包含一个或多个工程化序列,其中所述一个或多个工程化序列共同包含:(i)文库构建体和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中文库构建体包含位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,并且其中昆虫细胞群产生单独地包含标识符的病毒载体。在一些实施方案中,昆虫细胞群任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动、至少一个文库变体、至少一个有效载荷、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列。

[0019] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,所述多个昆虫细胞各自单独地包含一个或多个工程化序列,其中所述一个或多个工程化序列共同包含:(i)文库构建体和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其

中文库构建体包含位于能够包装到病毒载体中的两个病毒重复序列之间的标识符,并且其中昆虫细胞群产生单独地包含标识符的病毒载体。在一些实施方案中,文库构建体任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含至少一个文库变体和/或至少一个有效载荷。在一些实施方案中,昆虫细胞群任选地还包含一个或多个工程化序列,所述一个或多个工程化序列包含:至少一个扰动、至少一个扰动附加序列、至少一个反式作用整合序列,和/或至少一个顺式作用整合序列。

[0020] 在一些实施方案中,提供的昆虫细胞包含扰动(例如,一个或多个扰动)。在一些实施方案中,提供的昆虫细胞产生包含扰动(例如,一个或多个扰动)的病毒载体。在一些实施方案中,提供的昆虫细胞包含扰动(例如,一个或多个扰动),所述昆虫细胞也产生包含扰动(例如,一个或多个扰动)的病毒载体。在一些实施方案中,至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸包含扰动(例如,一个或多个扰动)。在一些实施方案中,提供的昆虫细胞包含扰动(例如,一个或多个扰动),并且至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸包含扰动(例如,一个或多个扰动)。在一些实施方案中,提供的昆虫细胞产生包含扰动(例如,一个或多个扰动)的病毒载体,并且至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸包含扰动(例如,一个或多个扰动)。在一些实施方案中,提供的昆虫细胞包含扰动(例如,一个或多个扰动),所述昆虫细胞也产生包含扰动(例如,一个或多个扰动)的病毒载体,并且至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸包含扰动(例如,一个或多个扰动)。

[0021] 昆虫细胞和/或昆虫细胞群可包括本领域适合于表达病毒载体的任何昆虫细胞。在一些实施方案中,昆虫细胞是来自或源自粘虫(例如,草地贪夜蛾(Spodoptera frugiperda))、果蝇(例如,果蝇属(Drosophila))或蚊子(例如,白纹伊蚊(Aedes albopictus))的细胞。

[0022] 在一些实施方案中,昆虫细胞包括Sf21细胞、Sf9细胞、BTI-TN-5B1-4(High Five)细胞、S2细胞、D.Mel2细胞或其任意的衍生物。在一些实施方案中,昆虫细胞包括Sf9细胞和/或其衍生物。

[0023] 在一些实施方案中,昆虫细胞适用于悬浮细胞培养。在一些实施方案中,昆虫细胞适用于贴壁细胞培养。在一些实施方案中,昆虫细胞群可以包含悬浮细胞和/或贴壁细胞中的细胞。

[0024] 本公开的昆虫细胞和/或方法可以用于产生任何病毒载体。在一些实施方案中,由本公开的昆虫细胞和/或方法产生的病毒载体包含扰动(例如,一个或多个扰动)。在一些实施方案中,扰动改变一个或多个与产生病毒载体相关的特征(例如,病毒载体稳定性等)或其他特征(例如,改变的治疗活性等)。

[0025] 在一些实施方案中,昆虫细胞文库的昆虫细胞已经被修饰以破坏或去除产生的病毒载体的受体。在一些实施方案中,昆虫细胞已经用阻断病毒载体感染的剂处理。

[0026] 在一些实施方案中,病毒载体是腺相关病毒(AAV)载体、杆状病毒载体、慢病毒载体、腺病毒载体、α病毒载体、辛德毕斯病毒载体、逆转录病毒载体(例如,γ逆转录病毒载体)、多瘤病毒载体(例如,猴病毒40(SV40)载体)、乳头瘤病毒载体(例如,牛乳头瘤病毒(BPV)载体)、痘苗病毒载体、单纯疱疹病毒(HSV)载体、麻疹病毒载体、棒状病毒载体、狂犬病病毒载体、水疱性口炎病毒(VSV)载体、小核糖核酸病毒载体(例如,脊髓灰质炎病毒载

体)、呼肠孤病毒载体、塞内卡病毒载体、艾柯病毒(echovirus)载体(例如,RIGVIR)、塞姆利基森林病毒(SFV)载体、黄病毒载体、指环病毒载体、新城疫病毒(NDV)载体、副粘病毒载体、仙台病毒载体、正粘病毒载体、流感病毒载体、冠状病毒载体和/或杂交病毒载体,和/或其衍生物、杂交体和/或工程化衍生物。

[0027] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达腺相关病毒(AAV)载体。在一些实施方案中,昆虫细胞包含(i)至少一个位于能够包装到AAV载体中的两个病毒重复序列之间的标识符和(ii)至少一个包含一个或多个产生AAV载体所必需的核酸序列的多核苷酸。在一些实施方案中,两个病毒重复序列各自是能够包装到AAV载体中的AAV ITR序列。在一些实施方案中,两个病毒重复序列各自是昆虫AAV ITR序列。在一些实施方案中,两个病毒重复序列各自是人AAV ITR序列。

[0028] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,其中多个中的每个昆虫细胞包含:(i)核酸序列,其包含位于两个功能性AAV ITR序列之间的条形码,其中所述核酸序列被整合至位于一对顺式作用整合序列之间的昆虫基因组中,(ii)一个或多个文库变体,其导致一个或多个扰动,和(iii)一个或多个产生AAV载体所必需的核酸序列,其中昆虫细胞群产生多个AAV载体,其中每个AAV载体包含对应于产生它的昆虫细胞的条形码的条形码。在一些实施方案中,至少一个文库变体包含gRNA,并且所述昆虫细胞还包含RNA引导的核酸酶。在一些实施方案中,由昆虫细胞产生的AAV载体还包含有效载荷。

[0029] 在一些实施方案中,昆虫细胞的顺式作用整合序列是源自无脊椎动物病毒(例如,鹿眼蛱蝶浓核病毒(Junonia coenia densovirus)(JcDNV))的病毒重复序列。在一些实施方案中,昆虫细胞的顺式作用整合序列是重组酶识别位点。

[0030] 在一些实施方案中,相对于缺乏一个或多个扰动的参考昆虫细胞群,昆虫细胞的一个或多个扰动与AAV产生和/或AAV分泌的增加相关。在一些实施方案中,相对于缺乏一个或多个扰动的参考昆虫细胞,包含一个或多个扰动的昆虫细胞具有至少10%的AAV产生和/或AAV分泌增加。

[0031] 在一些实施方案中,本文所述的昆虫细胞和/或文库构建体的两个病毒重复序列包含一对反向末端重复序列(ITR),其是或包括人AAV1 ITR;人AAV2 ITR;人AAV3bITR;人AAV4 ITR;人AAV5 ITR;人AAV6 ITR;人AAV7 ITR;人AAV8 ITR;人AAV9 ITR;人AAV10 ITR;人AAV11 ITR;人AAV12 ITR;人AAV13 ITR,或其任何组合。

[0032] 在一些实施方案中,本文所述的昆虫细胞和/或文库构建体的两个病毒重复序列包含一对反向末端重复序列(ITR),其是或包括牛AAV(b-AAV)ITR;犬AAV(CAAV)ITR;小鼠AAV1 ITR;山羊AAV ITR;大鼠AAV ITR;或禽AAV(AAAV)ITR。

[0033] 在一些实施方案中,昆虫细胞包含一个或多个多核苷酸,其编码一个或多个产生AAV载体所必需的蛋白质,诸如AAV衣壳蛋白。在一些实施方案中,AAV载体包含人AAV1衣壳蛋白;人AAV2衣壳蛋白;人AAV3b衣壳蛋白;人AAV4衣壳蛋白;人AAV5衣壳蛋白;人AAV6衣壳蛋白;人AAV7衣壳蛋白;人AAV8衣壳蛋白;人AAV9衣壳蛋白;人AAV10衣壳蛋白;人AAV11衣壳蛋白;人AAV12衣壳蛋白;或人AAV13衣壳蛋白。在一些实施方案中,AAV载体包含人祖先AAV衣壳蛋白。在一些实施方案中,AAV载体包含牛AAV(b-AAV)衣壳蛋白;犬AAV(CAAV)衣壳蛋白;小鼠AAV1衣壳蛋白;山羊AAV衣壳蛋白;大鼠AAV衣壳蛋白;或禽AAV(AAAV)衣壳蛋白。

[0034] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达AAV

载体,其中昆虫细胞包含(i)包含至少一个位于能够包装到AAV载体中的两个AAV ITR序列之间的标识符的文库构建体,和(ii)至少一个包含一个或多个产生AAV载体所必需的选自Rep基因、Cap基因、辅助基因或其组合的核酸序列的多核苷酸。在一些实施方案中,至少一个包含一个或多个形成AAV载体所必需的核酸序列的多核苷酸包含:AAV Rep基因、AAV Cap基因、一个或多个AAV辅助基因或其组合。在一些实施方案中,AAV载体具有复制能力。在一些实施方案中,AAV载体是条件复制的、复制缺乏的、无复制能力的和/或复制缺陷的。

[0035] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达杆状病毒载体(例如,苜蓿银纹夜蛾(Autographa californica)多核型多角体病毒(AcMNPV)载体)。

[0036] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达慢病毒载体。在一些实施方案中,昆虫细胞包含(i)至少一个位于能够包装到慢病毒载体中的两个病毒重复序列之间的标识符和(ii)至少一个包含一个或多个产生慢病毒载体所必需的核酸序列的多核苷酸。在一些实施方案中,两个病毒重复序列各自是能够包装到慢病毒载体中的慢病毒LTR序列。

[0037] 在一些实施方案中,本文所述的昆虫细胞和/或文库构建体的两个病毒重复序列包含一对LTR,其包括HIV LTR、SIV LTR、马传染性贫血病毒LTR、FIV LTR、维斯纳病毒LTR,或其衍生物或组合。在一些特定的实施方案中,慢病毒载体是HIV载体并且LTR是HIV LTR或其衍生物。在一些实施方案中,慢病毒载体包含慢病毒Psi序列。

[0038] 在一些实施方案中,提供了包含多个昆虫细胞的昆虫细胞群,其中多个中的每个昆虫细胞包含:(i)核酸序列,包含位于两个功能性慢病毒LTR序列之间的条形码,其中所述核酸序列被整合至位于一对顺式作用整合序列之间的昆虫基因组中,(ii)一个或多个文库变体,其导致一个或多个扰动,和(iii)一个或多个产生慢病毒载体所必需的核酸序列,其中昆虫细胞群产生多个慢病毒载体,其中每个慢病毒载体包含对应于产生它的昆虫细胞的条形码的条形码。在一些实施方案中,至少一个文库变体包含gRNA,并且所述昆虫细胞还包含RNA引导的核酸酶。在一些实施方案中,由昆虫细胞产生的慢病毒载体还包含有效载荷。

[0039] 在一些实施方案中,昆虫细胞的顺式作用整合序列是源自慢病毒的病毒重复序列。在一些实施方案中,昆虫细胞的顺式作用整合序列是重组酶识别位点。

[0040] 在一些实施方案中,慢病毒LTR包括HIV LTR、SIV LTR、马传染性贫血病毒LTR、FIV LTR、维斯纳病毒LTR或其衍生物或组合。在一些特定的实施方案中,慢病毒载体是HIV载体并且LTR是HIV LTR或其衍生物。在一些实施方案中,慢病毒载体包含慢病毒Psi序列。

[0041] 在一些实施方案中,相对于缺乏一个或多个扰动的参考昆虫细胞群,昆虫细胞的一个或多个扰动与慢病毒载体产生和/或慢病毒载体分泌的增加相关。在一些实施方案中,相对于缺乏一个或多个扰动的参考昆虫细胞,包含一个或多个扰动的昆虫细胞具有至少10%的慢病毒载体产生和/或慢病毒载体分泌增加。

[0042] 在一些实施方案中,由本公开的昆虫细胞和方法表达的慢病毒载体是人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒(FIV)载体、维斯纳病毒载体或其衍生物。

[0043] 在一些实施方案中,昆虫细胞包含一个或多个多核苷酸,其编码一个或多个产生慢病毒载体所必需的蛋白质,诸如慢病毒gag蛋白或其片段。在一些实施方案中,gag蛋白包

含一个或多个选自基质(MA)、衣壳(CA)和核衣壳(NC)结构域的结构域。在一些实施方案中,昆虫细胞包含一个或多个多核苷酸,其编码一个或多个产生慢病毒载体所必需的蛋白质,诸如慢病毒包膜蛋白或其片段。

[0044] 在一些实施方案中,慢病毒载体是包含源自不同病毒的gag蛋白和包膜蛋白的假型慢病毒载体。在一些实施方案中,昆虫细胞包含一个或多个编码一个或多个产生假型慢病毒载体所必需的蛋白质的多核苷酸,所述假型慢病毒载体包含源自人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒载体、维斯纳病毒载体或其衍生物的gag蛋白和/或env蛋白。

[0045] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达慢病毒载体,其中昆虫细胞包含(i)包含至少一个位于两个慢病毒LTR和/或Psi序列之间的标识符的文库构建体,所述序列能够包装到慢病毒载体中,和(ii)至少一个包含一个或多个产生慢病毒载体所必需的选自gag基因、env基因、pol基因或其组合的核酸序列的多核苷酸。在一些实施方案中,至少一个包含一个或多个形成慢病毒载体所必需的核酸序列的多核苷酸包含:慢病毒gag基因、慢病毒env基因、慢病毒pol基因或其组合。在一些实施方案中,慢病毒载体具有复制能力。在一些实施方案中,慢病毒载体是条件复制的、复制缺乏的、无复制能力的和/或复制缺陷的。

[0046] 在一些实施方案中,至少一个包含一个或多个形成慢病毒载体所必需的核酸序列的多核苷酸包含:HIV gag基因、HIV env基因、HIV pol基因或其组合。在一些实施方案中,慢病毒载体是具有复制能力的HIV载体。在一些实施方案中,慢病毒载体是条件复制的、复制缺乏的、无复制能力的和/或复制缺陷的HIV载体。

[0047] 在一些实施方案中,昆虫细胞已经先前或同时被遗传修饰以破坏或去除慢病毒的受体。在一些实施方案中,昆虫细胞已经用阻断慢病毒载体感染的剂处理。在此类实施方案中,整合载体和/或顺式作用整合序列不源自慢病毒。

[0048] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达单纯疱疹病毒(HSV)载体。在一些实施方案中,提供的昆虫细胞包含(i)至少一个位于能够包装到HSV载体中的两个病毒重复序列之间的标识符和(ii)至少一个包含一个或多个产生HSV载体所必需的核酸序列的多核苷酸。在一些实施方案中,HSV载体是或源自单纯疱疹病毒-1(HSV-1)、单纯疱疹病毒-2(HSV-2)、人巨细胞病毒(HCMV)、水痘-带状疱疹病毒(VZV)、爱泼斯坦-巴尔病毒(EBV)、卡波西肉瘤相关疱疹病毒(KSHV)、人疱疹病毒6和/或人疱疹病毒7和/或其衍生物。

[0049] 在一些实施方案中,本文所述的昆虫细胞和/或文库构建体的两个病毒重复序列包含末端a序列。

[0050] 在一些实施方案中,昆虫细胞包含一个或多个多核苷酸,其编码一个或多个产生HSV载体所必需的蛋白质,诸如HSV衣壳蛋白或其片段。在一些实施方案中,HSV衣壳包含VP5、VP19C、VP23、pre-VP22a和/或成熟的蛋白酶(UL26基因产物)。

[0051] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达HSV载体,其中昆虫细胞包含(i)包含至少一个位于HSV末端a序列之间的标识符的文库构建体,所述序列能够包装到HSV载体中,和(ii)至少一个包含一个或多个产生包含HSV衣壳蛋白的HSV载体所必需的核酸序列的多核苷酸。

[0052] 在一些实施方案中,本文提供的昆虫细胞、昆虫细胞群和/或方法可用于表达HSV-AAV杂交载体。在一些实施方案中,HSV-AAV杂交载体具有复制能力。在一些实施方案中,HSV-AAV杂交载体是条件复制的、复制缺乏的、无复制能力的和/或复制缺陷的。

[0053] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含一个或多个多核苷酸,所述一个或多个多核苷酸包含一个或多个产生病毒载体所必需的核酸序列。在一些实施方案中,包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸游离存在于昆虫细胞中。在一些实施方案中,包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸存在于昆虫细胞基因组中。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列包含异源性调控元件(例如,异源性启动子和/或异源性增强子)。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列包含异源性启动子序列,所述异源性启动子序列是或包括SV40启动子、延伸因子(EF)-1启动子、巨细胞病毒(CMV)启动子、磷酸甘油酸激酶(PGK)1启动子、泛素(Ubc)启动子、人β肌动蛋白启动子、四环素反应元件(TRE)启动子、脾灶形成病毒(SFFV)启动子、鼠干细胞病毒(MSCV)启动子、超级核心启动子(SCP)、CAG启动子或其衍生物。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列包含异源性增强子序列,其是或包括CMV早期增强子、cAMP反应元件(CRE)增强子或其衍生物。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列处于诱导型转录控制元件的控制之下。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列可以整合到昆虫细胞基因组中并处于诱导型转录控制元件(例如,诱导型启动子和/或诱导型增强子)的控制之下。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列可以游离存在于昆虫细胞中并处于诱导型转录控制元件(例如,诱导型启动子和/或诱导型增强子)的控制之下。

[0054] 本公开的昆虫细胞和/或昆虫细胞群的任何工程化序列可以游离存在和/或整合到昆虫细胞基因组中。在一些实施方案中,一个或多个工程化序列可以游离存在于昆虫细胞中,包括但不限于:标识符、扰动、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列。在一些实施方案中,一个或多个工程化序列可以整合到昆虫细胞基因组中,包括但不限于:标识符、扰动、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列。

[0055] 在一些实施方案中,一个或多个工程化序列存在于病毒载体中(例如,标识符、扰动、有效载荷等)。在一些实施方案中,昆虫细胞和/或昆虫细胞群表达各自包含标识符和扰动的病毒载体。在一些实施方案中,病毒载体包含改变与病毒载体产生相关的一个或多个特征和/或其他特征(例如,稳定性等)的扰动。在一些实施方案中,病毒载体还包含有效载荷。在一些实施方案中,病毒载体还包含报告基因和/或选择标记。

[0056] 在一些实施方案中,昆虫细胞包含文库构建体,其中构成文库构建体的一个或多个多核苷酸可以游离存在于昆虫细胞中,包括但不限于:标识符、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列。在一些实施方案中,昆虫细胞包含文库构建体,其中构成文库构建体的一个或多个多核苷酸可以整合到昆虫细胞中,包括但不限于:标识符、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列。在一些实施方案中,构成文库构建体的一个或多个多核苷酸可以以低拷贝数整合到昆虫细胞中,例如四个拷贝或更少、三个拷贝或更少、两个拷贝或更少,或单个拷贝。

[0057] 在一些实施方案中,一个或多个工程化序列包含异源性编码序列。在一些实施方

案中,一个或多个工程化序列包含异源性基因和/或异源性基因区段。在一些实施方案中,一个或多个工程化序列包含异源性调控元件(例如,异源性启动子和/或异源性增强子)。在一些实施方案中,一个或多个工程化序列包含异源性启动子序列,所述异源性启动子序列是或包括SV40启动子、延伸因子(EF)-1启动子、巨细胞病毒(CMV)启动子、磷酸甘油酸激酶(PGK)1启动子、泛素(Ubc)启动子、人β肌动蛋白启动子、四环素反应元件(TRE)启动子、脾灶形成病毒(SFFV)启动子、鼠干细胞病毒(MSCV)启动子、超级核心启动子(SCP)、CAG启动子或其衍生物。在一些实施方案中,一个或多个工程化序列包含异源性增强子序列,所述异源性增强子序列是或包括CMV早期增强子、cAMP反应元件(CRE)增强子或其衍生物。在一些实施方案中,一个或多个工程化序列包含诱导型转录控制元件。

[0058] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多于一个工程化序列(例如,标识符、扰动、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列)。在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个,或十个或更多个工程化序列。在一些实施方案中,昆虫细胞包含多至100个工程化序列。

[0059] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含有包含一个或多个工程化序列的文库构建体,所述一个或多个工程化序列可以包括例如:标识符、文库变体、有效载荷、扰动附加序列、反式作用整合序列,和/或顺式作用整合序列。在一些实施方案中,本公开的昆虫细胞包含至少一个文库构建体,其中所述至少一个文库构建体包含至少一个工程化序列。在一些实施方案中,文库构建体包含两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个,或十个或更多个工程化序列。在一些实施方案中,文库构建体包含多至100个工程化序列。

[0060] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群产生包含一个或多个工程化序列的病毒载体,所述一个或多个工程化序列可以包括例如:标识符、扰动、有效载荷,和/或顺式作用整合序列。

[0061] 在一些实施方案中,本公开的昆虫细胞包含至少一个文库构建体,所述文库构建体包含至少一个选自以下的工程化序列:至少一个条形码、至少一个标识符、至少一个文库变体、至少一个有效载荷、至少一个顺式作用整合序列,或其组合和/或其多个。

[0062] 在一些实施方案中,昆虫细胞包含至少一个包含条形码的工程化序列。在一些实施方案中,条形码包含约5至约25个核苷酸的序列。在一些实施方案中,提供的昆虫细胞包含多个独特的条形码,并且其中所述多个独特的条形码包含约5至约25个核苷酸的独特序列。

[0063] 在一些实施方案中,文库构建体包含至少一个条形码。在一些实施方案中,文库构建体包含标识符,所述标识符包含至少一个条形码。在一些实施方案中,文库构建体包含标识符,所述标识符包含至少一个条形码,其中条形码位于两个病毒重复序列之间。在一些实施方案中,条形码位于两个病毒重复序列之间并且不是标识符。在一些实施方案中,条形码不位于两个病毒重复序列之间。

[0064] 在一些实施方案中,昆虫细胞包含至少一个包含文库变体的工程化序列。文库变体可以包括但不限于包含基因、ORF、gRNA序列、非编码核酸或其组合的工程化序列。在一些

实施方案中,昆虫细胞包含一个、两个、三个、四个、五个、六个、七个、八个、九个或十个文库变体。在一些实施方案中,昆虫细胞包含多至100个文库变体。

[0065] 在一些实施方案中,昆虫细胞包含多个文库变体,其中所述多个文库变体包含至少一个工程化序列,所述至少一个工程化序列包含:至少一个独特基因、至少一个独特ORF、至少一个独特gRNA序列,和/或至少一个独特非编码核酸或其组合和/或多个。在一些实施方案中,昆虫细胞包含多个文库构建体,其中所述多个文库构建体包含:至少一个独特基因、至少一个独特ORF、至少一个独特gRNA序列、至少一个独特非编码核酸序列或其组合和/或多个。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含多个独特基因、多个独特ORF、多个独特gRNA序列、多个独特非编码核酸序列或其组合。

[0066] 在一些实施方案中,文库构建体包含gRNA序列。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少一个独特gRNA序列。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少100个独特gRNA序列。

[0067] 在一些实施方案中,文库构建体包含ORF。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少一个独特ORF。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少100个独特ORF。

[0068] 在一些实施方案中,文库构建体包含基因。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少一个独特基因。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少100个独特基因。

[0069] 在一些实施方案中,文库构建体包含非编码核酸序列。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体各自包含至少一个独特非编码核酸序列。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含至少100个独特非编码核酸序列。

[0070] 在一些实施方案中,文库构建体包含至少一个报告基因和/或选择标记。在一些实施方案中,一个或多个包含文库构建体的多核苷酸包含报告基因和/或选择标记。

[0071] 在一些实施方案中,文库构建体包含标识符。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中所述多个文库构建体包含多个标识符。在一些实施方案中,标识符包含至少一个条形码和/或至少一个文库变体。在一些实施方案中,昆虫细胞包含有包含标识符的文库构建体,所述标识符包含至少一个条形码和/或至少一个文库变体。在一些实施方案中,昆虫细胞群包含多个包含多个标识符的文库构建体,其中所述标识符包含多个条形码和/或多个文库变体。

[0072] 在一些实施方案中,文库构建体包含多个工程化序列,其中:多个工程化序列的第一子集位于两个病毒重复序列之间,并且多个工程化序列的第二子集位于两个病毒重复序列之外。

[0073] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多个包含至少一个文库变体和至少一个标识符的工程化序列,其中至少一个文库变体和至少一个标识符都位于两个病毒重复序列之间。在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多个包含以下的工程化序列:至少一个文库变体、至少一个标识符和至少一个有效载荷,其中所述至少一个文库变体、至少一个标识符和至少一个有效载荷位于两个病毒重复序列之间。

[0074] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多个包含至少一个文库变体和至少一个标识符的工程化序列,并且其中至少一个标识符位于两个病毒重复序列之间,并且其中至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多个包含至少一个文库变体、至少一个标识符和至少一个有效载荷的工程化序列,其中所述至少一个标识符和至少一个有效载荷位于两个病毒重复序列之间,并且其中所述至少一个文库变体位于两个病毒重复序列之外。

[0075] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含多个包含至少两个文库变体和至少一个标识符的工程化序列,其中所述至少一个标识符和至少两个文库变体中的至少一个文库变体位于两个病毒重复序列之间,并且其中所述至少两个文库变体中的至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,提供的昆虫细胞还包含位于两个病毒重复序列之间的有效载荷。

[0076] 在一些实施方案中,提供的文库构建体包含至少一个文库变体和至少一个标识符,其中所述至少一个文库变体和至少一个标识符都位于两个病毒重复序列之间。在一些实施方案中,提供的文库构建体包含:至少一个文库变体、至少一个标识符和至少一个有效载荷,其中所述至少一个文库变体、至少一个标识符和至少一个有效载荷位于两个病毒重复序列之间。

[0077] 在一些实施方案中,提供的文库构建体包含至少一个文库变体和至少一个标识符,并且其中至少一个标识符位于两个病毒重复序列之间,并且其中至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,提供的文库构建体包含至少一个文库变体、至少一个标识符和至少一个有效载荷,其中所述至少一个标识符和至少一个有效载荷位于两个病毒重复序列之间,并且其中所述至少一个文库变体位于两个病毒重复序列之外。

[0078] 在一些实施方案中,提供的文库构建体包含至少两个文库变体和至少一个标识符,其中所述至少一个标识符和至少两个文库变体中的至少一个文库变体位于两个病毒重复序列之间,并且其中所述至少两个文库变体中的至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,提供的文库构建体还包含位于两个病毒重复序列之间的有效载荷。

[0079] 在一些实施方案中,至少一个标识符包含条形码。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一个包含至少一个条形码的工程化序列,并且其中所述至少一个条形码位于两个病毒重复序列之间。在一些实施方案中,位于两个病毒重复序列之间的条形码是标识符。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一个包含至少一个条形码的工程化序列,并且其中所述至少一个条形码位于两个病毒重复序列之外。

[0080] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含多个条形码。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含多个条形码,其中至少一个条形码位于两个病毒重复序列之间。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含多个条形码,其中多个条形码位于两个病毒重复序列之间。在一些实施方案中,位于两个病毒重复序列之间的至少一个条形码是标识符。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含多个条形码,其中多个条形码的第一子集位于两个病毒重复序列之间,并且多个条形码的第二子集位于两个病毒重复序列之外。

[0081] 在一些实施方案中,提供的昆虫细胞包含文库构建体或其一部分的多于一个拷贝。在一些实施方案中,提供的昆虫细胞包含文库构建体或其一部分的1、2、3、4、5、6、7、8、9、10个或更多个拷贝。在一些实施方案中,提供的昆虫细胞包含文库构建体或其一部分的多个拷贝。在一些实施方案中,提供的昆虫细胞包含文库构建体或其一部分的一个至四个拷贝。在一些实施方案中,提供的昆虫细胞确切包含文库构建体或其一部分的两个拷贝。在一些实施方案中,提供的昆虫细胞确切包含文库构建体或其一部分的一个拷贝。

[0082] 在一些实施方案中,提供的昆虫细胞包含至少一个由单个连续核酸序列组成的文库构建体。在一些实施方案中,提供的昆虫细胞群包含多个文库构建体,其中每个单独的文库构建体由单个连续核酸序列组成,并且其中多个文库构建体包含多个独特核酸序列。在一些实施方案中,文库构建体包含多于一个不连续核酸序列。在一些实施方案中,文库构建体包含至少两个、三个、四个、五个、六个、七个、八个、九个或十个不连续核酸序列。在一些实施方案中,文库构建体包含多至100个不连续核酸序列。

[0083] 在一些实施方案中,提供的昆虫细胞群包含多个文库构建体,其中每个单独的文库构建体由不连续核酸序列组成,并且其中文库构建体包含多个独特核酸序列。在一些实施方案中,源自每个单独的昆虫细胞的核酸或其衍生物在单细胞测序方法过程中包含至少一个独特的细胞身份序列。在一些实施方案中,源自每个单独的昆虫细胞的多于一个核酸序列或其衍生物在单细胞测序方法过程中包含细胞身份序列。在一些实施方案中,源自每个单独的文库构建体的核酸或其衍生物在单细胞测序方法过程中包含至少一个独特的细胞身份序列。在一些实施方案中,源自每个单独的文库构建体的多于一个核酸序列或其衍生物在单细胞测序方法过程中包含细胞身份序列。

[0084] 在一些实施方案中,提供的昆虫细胞群包含多个文库构建体,其中每个昆虫细胞包含单个文库构建体,所述单个文库构建体由多个不连续核酸序列组成,并且其中文库构建体包含多个独特核酸序列,并且其中来自文库构建体的多于一个核酸序列(或其衍生物)在单细胞测序方法过程中包含细胞身份序列。在一些实施方案中,来自文库构建体的所有核酸序列(或其衍生物)在单细胞测序方法过程中包含细胞身份序列。

[0085] 在一些实施方案中,提供的昆虫细胞和/或病毒载体包含至少一个包含至少一个扰动的工程化序列。在一些实施方案中,提供的昆虫细胞包含至少一个游离存在于昆虫细胞中的扰动。在一些实施方案中,提供的昆虫细胞包含至少一个存在于昆虫细胞的基因组中的扰动。在一些实施方案中,提供的昆虫细胞包含至少两个扰动,其中至少一个扰动游离存在并且至少一个扰动存在于昆虫细胞的基因组中。

[0086] 在一些实施方案中,由本文提供的昆虫细胞表达的病毒载体包含扰动(例如,一个或多个扰动)。在一些实施方案中,病毒载体包含工程化序列,所述工程化序列包含存在于病毒核酸中的扰动。在一些实施方案中,病毒载体包含工程化序列,所述工程化序列包含存在于ITR中的扰动。在一些实施方案中,病毒载体包含工程化序列,所述工程化序列包含存在于ITR之间的扰动。

[0087] 在一些实施方案中,形成病毒载体所必需的一个或多个多核苷酸包含有包含扰动(例如,一个或多个扰动)的工程化序列。在一些实施方案中,形成病毒载体所必需的一个或多个多核苷酸游离存在并包含有包含扰动(例如,一个或多个扰动)的工程化序列。在一些实施方案中,形成病毒载体所必需的一个或多个多核苷酸存在于昆虫细胞的基因组中并包

含有包含扰动(例如,一个或多个扰动)的工程化序列。

[0088] 在一些实施方案中,提供的昆虫细胞和/或病毒载体包含多个独特的扰动。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少两个、三个、四个、五个、六个、七个、八个或九个独特的扰动。

[0089] 在一些实施方案中,提供的昆虫细胞和/或病毒载体包含至少一个扰动,所述扰动包括内源性基因组编码序列的插入、缺失、取代、替换、表观遗传修饰和/或重排。在一些实施方案中,内源性编码序列是或包含内源性基因或基因区段。

[0090] 在一些实施方案中,提供的昆虫细胞和/或病毒载体包含至少一个扰动,所述扰动包括内源性基因组调控元件的插入、缺失、取代、替换、表观遗传修饰和/或重排。在一些实施方案中,内源性调控元件是或包含内源性启动子序列和/或内源性增强子序列。

[0091] 在一些实施方案中,提供的昆虫细胞包含至少一个扰动附加序列。在一些实施方案中,扰动附加序列包含RNA引导的核酸酶或其衍生物。在一些实施方案中,RNA引导的核酸酶包含Cas9、Cpf1和/或CasZ或其衍生物,包括包含转录调控因子的融合蛋白(例如,Cas9-VPR或Cas9-KRAB-MeCP2融合物)、CRISPR蛋白与核酸酶结构域(例如,Fok1)的融合物、酶碱基编辑器(例如,BE和ABE融合物版本)、逆转录酶融合物(例如,引导编辑器)、CRISPR重组酶(例如,RecCas9)和CRISPR转座酶(例如,Tn7样转座酶系统Cas12k和与TniQ的级联复合物)。在一些实施方案中,RNA引导的核酸酶包含Cas9或其衍生物。

[0092] 所提供的文库构建体可以使用本领域已知的任何适当方法引入昆虫细胞。在一些实施方案中,通过转染和/或转导将文库构建体引入昆虫细胞中。在一些实施方案中,通过杆状病毒介导的感染将文库构建体引入昆虫细胞中。在一些实施方案中,通过慢病毒介导的转导将文库构建体引入昆虫细胞中。

[0093] 在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体的至少一个工程化序列游离存在。在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体包含至少一个包含至少一个文库变体的工程化序列,并且其中每个单独的细胞包含至少一个扰动附加序列,并且其中至少一个工程化序列游离存在。

[0094] 在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体的至少一个工程化序列包含至少一个文库变体,并且其中至少一个文库变体游离存在。在一些实施方案中,至少一个文库变体是效应物。在一些实施方案中,至少一个文库变体成为至少一个扰动。在一些实施方案中,至少一个文库变体包含至少一个ORF、至少一个基因、至少一个非编码核酸序列和/或至少一个gRNA,或其多个。

[0095] 在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体包含至少一个包含至少一个文库变体的工程化序列,所述至少一个文库变体包含至少一个gRNA,并且其中每个单独的细胞包含至少一个扰动附加序列,所述扰动附加序列包含RNA引导的核酸酶或非RNA引导的核酸酶或其衍生物,并且其中至少一个工程化序列游离存在。

[0096] 在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体包含至少一个包含至少一个文库变体的工程化序列,所述至少一个文库变体包含至少一个gRNA,其中每个单独的细胞包含至少一个扰动附加序列,所述扰动附加序列包含RNA引导的核酸酶或非RNA引导的核酸酶或其衍生物,并且其中至少一个工程化序列游离存在。

[0097] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包括至少一个扰动。在一些

实施方案中,提供的昆虫细胞和/或昆虫细胞群包含至少一个扰动,所述至少一个扰动包括基因组序列变化、游离型序列变化和/或表观遗传修饰。在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含至少一个扰动,所述至少一个扰动包括内源性基因组编码序列的插入、缺失、取代和/或重排。

[0098] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群产生单独地包含至少一个扰动的病毒载体。

[0099] 在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体的至少一个工程化序列存在于昆虫细胞的基因组中。在一些实施方案中,提供的昆虫细胞包含文库构建体,其中文库构建体包含至少一个包含至少一个文库变体的工程化序列,并且其中每个单独的细胞包含至少一个扰动附加序列,并且其中至少一个工程化序列存在于昆虫细胞的基因组中。

[0100] 在一些实施方案中,将文库构建体或其一部分在随机插入位点插入昆虫细胞的基因组中。在一些实施方案中,随机插入位点在预定的基因组位置子集中是随机的。在一些实施方案中,将文库构建体或其一部分在预定插入位点插入昆虫细胞的基因组中。

[0101] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一个反式作用整合序列,其中所述至少一个反式作用整合序列包含(i)至少一个整合构建体,(ii)至少一个重组酶,和/或其多肽、蛋白质、核酸或多核苷酸产物,(iii)至少一个核酸酶,和/或其多肽、蛋白质、核酸或多核苷酸产物,(iv)至少一个转座酶,和/或其多肽、蛋白质、核酸或多核苷酸产物,和/或(v)至少一个工程化序列。

[0102] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一个反式作用整合序列,所述反式作用整合序列包含整合构建体。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一对位于第一组病毒重复序列侧翼的顺式作用整合序列,其中顺式作用整合序列包含第二组病毒重复序列。在一些实施方案中,整合构建体是昆虫病毒载体或其衍生物。

[0103] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含第一组病毒重复序列(例如,用于将序列包装到病毒载体靶标中)和包含第二组病毒重复序列的顺式作用整合序列。

[0104] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一个反式作用整合序列,所述至少一个反式作用整合序列包含核酸酶和/或其多肽、蛋白质、核酸或多核苷酸产物。

[0105] 在一些实施方案中,至少一个反式作用整合序列包含至少一个核酸酶,并且其中所述至少一个核酸酶包括RNA引导的核酸酶或其融合物或衍生物。在一些实施方案中,至少一个反式作用整合序列包含至少一个核酸酶,并且其中所述至少一个核酸酶包括非RNA引导的核酸酶或其融合物或衍生物。在一些实施方案中,至少一个核酸酶包括Cas9、CasZ、Cpf1、与可编程DNA结合结构域(诸如TALE蛋白(TALEN)或锌指蛋白(ZFN))融合的工程化Fok1核酸酶结构域,和/或大范围核酸酶或其衍生物。在一些实施方案中,至少一个核酸酶包括Cas9。

[0106] 在一些实施方案中,至少一个反式作用整合序列还包含至少一个工程化序列。在一些实施方案中,至少一个反式作用整合序列包含RNA引导的核酸酶或其融合物或衍生物,

并且还包含至少一个gRNA。

[0107] 在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一对位于第一组病毒重复序列侧翼的顺式作用整合序列,并且其中顺式作用整合序列包含同源臂序列。

[0108] 在一些实施方案中,至少一个反式作用整合序列包含重组酶和/或其多肽、蛋白质、核酸或多核苷酸产物。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一对位于第一组病毒重复序列侧翼的顺式作用整合序列,并且其中顺式作用整合序列包含重组酶识别位点。在一些实施方案中,重组酶包括Cre、Flp、Dre、PhiC31和/或Bxb1,或其衍生物。

[0109] 在一些实施方案中,重组酶包括包含Cre的重组酶。在一些实施方案中,重组酶识别位点包括LoxP位点。

[0110] 在一些实施方案中,重组酶包括包含Bxb1的重组酶。在一些实施方案中,重组酶识别位点包括Att位点。

[0111] 在一些实施方案中,重组酶包括包含Flp的重组酶。在一些实施方案中,重组酶识别位点包括Frt位点。

[0112] 在一些实施方案中,至少一个反式作用整合序列包含转座酶和/或其多肽、蛋白质、核酸或多核苷酸产物。在一些实施方案中,提供的昆虫细胞和/或文库构建体包含至少一对位于第一组病毒重复序列侧翼的顺式作用整合序列,并且其中顺式作用整合序列包含转座酶识别位点。在一些实施方案中,转座酶包含Piggybac转座酶、睡美人转座酶和/或Tn5转座酶或其衍生物。在一些特定实施方案中,转座酶包含Piggybac转座酶或其衍生物。在一些特定实施方案中,转座酶包含睡美人转座酶或其衍生物。在一些特定实施方案中,转座酶包含Tn5转座酶或其衍生物。

[0113] 在一些实施方案中,相对于参考群,昆虫细胞群产生发生改变的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在将核酸转移至细胞的方式上发生改变的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在治疗上发生改变的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在预期应用上发生改变的病毒载体。

[0114] 在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在应用中功能性较低的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在应用中无功能性的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在将核酸转移至细胞方面功能性较低和/或无功能性的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在治疗上功能性较低和/或无功能性的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在预期应用中功能性较低和/或无功能性的病毒载体。

[0115] 在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在应用中功能性较高和/或增强的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在将核酸转移至细胞方面功能性较高和/或增强的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在治疗上功能性较高和/或增强的病毒载体。在一些实施方案中,相对于参考群,提供的昆虫细胞和/或昆虫细胞群产生在预期应用中功能性较高和/或增强的病毒载体。

[0116] 在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个在生产规范下改变病毒载体产生的工程化序列(例如,扰动)。在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个在生产规范下提供增加的病毒载体产生的工程化序列(例如,扰动)。在一些实施方案中,相对于参考细胞群,具有改变的病毒载体产生的昆虫细胞包含至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个或更多个扰动。

[0117] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含至少一个在当时现行的良好生产规范(cGMP)下提供增加的病毒载体产生的工程化序列(例如,扰动)。在一些实施方案中,至少一个工程化序列(例如,扰动)在良好生产规范(GMP)下提供增加的病毒载体产生。在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群包含至少一个在非良好生产规范(non-GMP)下提供增加的病毒载体产生的工程化序列(例如,扰动)。

[0118] 在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个提供增加的昆虫细胞群活力的工程化序列(例如,扰动)。在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个提供增加的昆虫细胞群产生病毒载体的持续时间的扰动。在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个提供增加的昆虫细胞群基因组稳定性的工程化序列(例如,扰动)。

[0119] 在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个在生产规范下提供减少的产生的病毒载体百分比的工程化序列(例如,扰动),所述病毒载体在应用中功能性较低。在一些实施方案中,相比于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个在生产规范下提供减少的产生的病毒载体百分比的扰动,所述病毒载体在应用中无功能性。在一些实施方案中,相对于参考细胞群,百分比减少的病毒载体在将核酸转移至细胞方面功能性较低和/或无功能性。在一些实施方案中,相对于参考细胞群,百分比减少的病毒载体在治疗上功能性较低和/或无功能性。在一些实施方案中,相对于参考细胞群,百分比减少的病毒载体在预期应用中功能性较低和/或无功能性。

[0120] 在一些实施方案中,相对于参考细胞群,提供的昆虫细胞和/或昆虫细胞群包含至少一个在生产规范下提供增加的产生的病毒载体百分比的工程化序列(例如,扰动),所述病毒载体在应用中功能性较高或增强。在一些实施方案中,相对于参考细胞群,百分比增加的病毒载体在将核酸转移至细胞方面功能性较高和/或增强。在一些实施方案中,相对于参考细胞群,百分比增加的病毒载体在治疗上功能性较高和/或增强。在一些实施方案中,相对于参考细胞群,百分比增加的病毒载体在预期应用中功能性较高和/或增强。

[0121] 在一些实施方案中,相对于参考细胞群,至少一个工程化序列(例如,扰动)在生产规范下提供增加的病毒载体百分比,所述病毒载体含有用于它们的预期应用的所有和/或必需的核酸序列和/或其他元件。

[0122] 在一些实施方案中,相对于参考细胞群,至少一个工程化序列(例如,扰动)在生产规范下提供减少的病毒载体百分比,所述病毒载体已失去用于它们的预期应用的所有和/或必需的核酸序列和/或其他元件,和/或使其发生突变。

[0123] 在一些实施方案中,提供的昆虫细胞和/或昆虫细胞群和/或病毒载体包含多个工程化序列,其中所述多个工程化序列包含多个扰动。

[0124] 在一些实施方案中,参考细胞群是:(a)不包含所述至少一个工程化序列的可比较的昆虫细胞群;和/或(b)能够产生病毒载体的标准细胞群。

[0125] 在一些实施方案中,相对于参考细胞群,由本文所述的昆虫细胞产生的病毒载体包含在生产规范下改变病毒载体产生的工程化序列(例如,扰动)。在一些实施方案中,病毒载体包含改变功能(例如,治疗功能、将核酸转移至细胞的能力)、基因组稳定性、昆虫细胞的产生产量、昆虫细胞产生持续时间、产生的含有所有必需核酸序列的病毒载体百分比,和/或与病毒载体产生、病毒载体活性和/或病毒载体应用相关的另一个特征的扰动。

[0126] 在一些实施方案中,昆虫细胞群通过以下方式产生:(a)将多个包含多个工程化序列的文库构建体引入到多个昆虫细胞中,其中单独的文库构建体包含至少一个位于第一组两个病毒重复序列之间的标识符,和(b)将至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸引入多个昆虫细胞中。

[0127] 在一些实施方案中,昆虫细胞群通过将多个包含多个工程化序列的文库构建体引入到多个昆虫细胞中产生,其中单独的文库构建体包含至少一个位于第一组两个病毒重复序列之间的标识符,其中多个昆虫细胞包含一个或多个产生病毒载体所必需的核酸序列。

[0128] 在一些实施方案中,提供了产生病毒载体的方法,其包括:在使得昆虫细胞产生病毒载体的条件下培养本公开的昆虫细胞群,并且其中每个产生的病毒载体包含至少一个源自产生病毒载体的昆虫细胞的至少一个标识符的标识符。在一些实施方案中,每个产生的病毒载体包含至少一个与产生病毒载体的昆虫细胞的至少一个标识符相同的标识符。

[0129] 在一些实施方案中,提供的方法包括:(a)从昆虫细胞文库产生病毒载体,其中昆虫细胞文库包含多个昆虫细胞,其中多个中的每个昆虫细胞单独地包含:(i)至少一个工程化序列,(ii)标识符,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,并且其中每个病毒载体包含源自产生病毒载体的昆虫细胞的标识符的标识符;以及(b)检测病毒载体中的标识符。在一些实施方案中,每个病毒载体包含与产生病毒载体的昆虫细胞的标识符相同的标识符。

[0130] 在一些实施方案中,提供的方法包括:(a)从昆虫细胞文库产生病毒载体,其中昆虫细胞文库包含多个昆虫细胞,其中多个中的每个昆虫细胞单独地包含:(i)至少一个工程化序列,(ii)至少一个标识符,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,并且其中每个病毒载体包含至少一个源自产生病毒载体的昆虫细胞的至少一个标识符的标识符;和(b)通过下一代测序检测病毒载体中的一个或多个标识符。在一些实施方案中,每个病毒载体包含至少一个与产生病毒载体的昆虫细胞的至少一个标识符相同的标识符。

[0131] 在一些实施方案中,提供的方法包括:(a)产生AAV文库,其中AAV文库包含来自昆虫细胞文库的多个AAV载体,其中昆虫细胞文库包含多个昆虫细胞,其中多个中的每个昆虫细胞单独地包含:(i)至少一个工程化序列,(ii)至少一个条形码序列,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中每个AAV病毒载体包含至少一个构建体,并且其中至少一个构建体包含源自产生AAV病毒载体的昆虫细胞的至少一个条形码序列的条形码序列;和(b)检测AAV文库中的一个或多个条形码序列。在一些实施方案中,每个AAV病毒载体包含至少一个构建体,并且其中至少一个构建体包含与产生AAV病毒载体的昆虫细胞的至少一个条形码序列相同的条形码序列。

[0132] 在一些实施方案中,提供的方法包括:(a)产生AAV文库,其中AAV文库包含来自昆虫细胞文库的多个构建体,其中昆虫细胞文库包含多个昆虫细胞,其中多个中的每个昆虫细胞单独地包含:(i)至少一个工程化序列,(ii)至少一个条形码序列,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中每个AAV病毒载体包含至少一个构建体,并且其中所述至少一个构建体包含源自产生AAV病毒载体的昆虫细胞的至少一个条形码序列的条形码序列;和(b)通过下一代测序检测AAV文库中的一个或多个条形码序列。在一些实施方案中,每个AAV病毒载体包含至少一个构建体,并且其中至少一个构建体包含与产生AAV病毒载体的昆虫细胞的至少一个条形码序列相同的条形码序列。在一些实施方案中,所提供的方法还包括对昆虫细胞文库的每个单独细胞内的至少一个或所有核酸序列或其衍生物进行单细胞测序的步骤,其中至少一个或所有核酸序列或其衍生物在单细胞测序方法过程中包含单细胞身份序列,并且其中至少一个或所有核酸序列或其衍生物包含至少一个文库构建体。

[0133] 在一些实施方案中,本公开提供的方法还包括对昆虫细胞文库的每个单独细胞内的至少一个或所有核酸序列或其衍生物进行单细胞测序,其中至少一个或所有核酸序列或其衍生物在单细胞测序方法过程中包含单细胞身份序列,并且其中至少一个或所有核酸序列或其衍生物包含至少一个包含至少一个标识符的文库变体。

[0134] 在一些实施方案中,本公开提供的方法还包括将第二文库构建体引入昆虫细胞中,其中昆虫细胞包含至少一个源自第一文库构建体的扰动,并且其中第二文库构建体包含至少一个包含位于第一组两个病毒重复序列之间的至少一个标识符的工程化序列。在一些实施方案中,将第二多个文库构建体引入多个昆虫细胞中,其中多个昆虫细胞包含至少一个源自第一多个文库构建体的扰动,并且其中第二多个文库构建体包含多个工程化序列,其中第二多个文库构建体中的每个单独的文库构建体包含至少一个位于第一组两个病毒重复序列之间的标识符。在一些实施方案中,提供的方法还包括通过单细胞测序和/或下一代测序检测一个或多个标识符和/或一个或多个工程化序列。在一些实施方案中,提供的方法包括多于两轮的文库构建体引入和一个或多个标识符和/或一个或多个工程化序列的检测。在一些实施方案中,提供的方法还包括使用机器学习方法来开发机器学习模型以鉴定靶扰动的期望组合。

[0135] 在一些实施方案中,本公开提供的方法还包括去除昆虫细胞的至少一个标识符的步骤。在一些实施方案中,在提供的昆虫细胞和/或文库构建体包含报告基因和/或选择标记的情况下,提供的方法可以包括去除报告基因和/或选择标记的步骤。

[0136] 在整个说明书中,在系统或组合物被描述为具有、包括或包含特定组件的情况下,或者在方法被描述为具有、包括或包含特定步骤的情况下,预期另外存在基本上由所述的组件组成或由所述的组件组成的本发明的系统或组合物,并且存在基本上由所述的步骤组成或由所述的步骤组成的根据本发明的方法。

[0137] 应当理解,步骤的次序或执行某些动作的次序并不重要,只要本发明保持可操作即可。另外,两个或更多个步骤或动作可以同时进行。

[0138] 以下描述仅用于说明和举例说明本公开,并不旨在将本发明限制于所描述的具体实施方案。

附图说明

[0139] 本文包括的附图由下图组成,仅用于说明目的,并不用于限制。

[0140] 图1描绘了用于鉴定或表征文库中昆虫细胞的病毒载体产生的技术平台的示意图,其中病毒载体获取与表达它们的昆虫细胞相关的标识符。

[0141] 图2描绘了用于生成表达包含标识符的AAV病毒载体的昆虫细胞文库的示例性方案。

[0142] 图3描绘了通过确定病毒载体池中标识符的相对富集和使用单细胞测序方法鉴定昆虫细胞中的工程化序列来鉴定和/或表征文库中昆虫细胞的病毒载体产生的示例性方案。

[0143] 图4描绘了用于生成表达AAV病毒载体的昆虫细胞文库的方法的示例性文库构建体的示意图。pSFX-AAV是用于游离表达的示例性AAV构建体,而pSFX-PB-AAV是示例性AAV-in-Transposase文库构建体。

[0144] 图5描绘了对病毒载体核酸(例如,来自病毒载体(例如,AAV病毒载体)或细胞基因组DNA的标识符(例如,条形码)序列)进行测序的示例性示意图概览。

[0145] 图6描绘了使用包含标识符的AAV-in-BAC构建体文库来生成表达AAV病毒载体的昆虫细胞文库的示例性方案。

[0146] 图7描绘了使用包含标识符的AAV-in-Transposase构建体文库来生成表达AAV病毒载体的昆虫细胞文库的示例性方案。

[0147] 图8描绘了使用包含标识符的游离构建体文库来生成表达AAV病毒载体的昆虫细胞文库的示例性方案。

[0148] 某些定义

[0149] 为了更容易理解本发明,下文首先定义了某些术语。在整个说明书中阐述了以下术语和其他术语的附加定义。本文中为了描述本发明的背景并提供关于其实践的额外细节所引用的出版物和其他参考材料特此以引用的方式并入。

[0150] 在本申请中,除非上下文另有明确说明,否则(i)术语“一(a/an)”在本文中用于指代一个或多于一个(即,指至少一个)文章的语法对象;(ii)术语“或”可以理解为意指“和/或”;(iii)术语“包含”和“包括”可理解为涵盖逐项列出的组分或步骤,无论是自身呈现还是与一个或多个额外组分或步骤一起呈现;并且(iv)在提供范围的情况下,端点包括在内。

[0151] 在整个说明书中,每当多核苷酸或多肽由字母序列表示时(例如,A、C、G和T,在多核苷酸的情况下分别代表腺苷、胞苷、鸟苷和胸苷),此类多核苷酸或多肽以从左到右的5’到3’或N-末端到C-末端的顺序呈现。

[0152] 约或大约:如本文所用,术语“约”或“大约”可以应用于一个或多个关注的值,包括与所陈述的参考值类似的值。一般来讲,熟悉上下文的本领域技术人员将理解所述上下文中“约”或“大约”所涵盖的相关差异程度。在一些实施方案中,术语“约”或“大约”是指落在规定参考值的±20%(大于或小于)内的值范围,除非另有说明或从上下文中另外明显看出。例如,在一些实施方案中,术语“约”或“大约”可以涵盖在参考值的20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更小之内的值范围。

[0153] 氨基酸:如本文所用,术语“氨基酸”是指可以并入多肽链中的任何化合物和/或物

质,例如,通过形成一个或多个肽键。在一些实施方案中,氨基酸具有通用结构,例如,H_2N-C(H)(R)-COOH。在一些实施方案中,氨基酸是天然存在的氨基酸。在一些实施方案中,氨基酸是非天然氨基酸;在一些实施方案中,氨基酸是D-氨基酸;在一些实施方案中,氨基酸是L-氨基酸。"标准氨基酸"是指天然存在的肽中常见的二十种标准L-氨基酸中的任一种。"非标准氨基酸"是指除标准氨基酸之外的任何氨基酸,无论其是合成制备的还是获得自天然来源的。从上下文可以清楚地看出,在一些实施方案中,术语"氨基酸"可以用于指代游离氨基酸;在一些实施方案中,它可以用于指代多肽的氨基酸残基。

[0154] 与...相关:如本文所用,如果一个事件或实体的存在、水平和/或形式与另一个事件或实体的存在、水平和/或形式有关、关联或相关联,则短语"与...相关"描述所述两个事件或实体。例如,如果第一实体(例如,标识符)的存在和/或水平与第二实体(例如,工程化序列)的存在和/或水平相关联(例如,在昆虫细胞中,例如,通过细胞分裂和/或遗传操作),则认为第一实体与第二实体相关。在一些实施方案中,如果两个或更多个实体存在于同一细胞、基因组、染色体或遗传区域中(例如,使得它们通过多代的昆虫细胞分裂一起遗传),则它们在物理上与彼此"相关"。在一些实施方案中,如果特定实体(例如,标识符和/或工程化序列)的存在、水平和/或形式例如与特定表型(例如,病毒载体的细胞产生)改变的水平或外显率相关联(例如,所述实体表现出增加的病毒载体产生的细胞),则认为所述实体与所述表型相关。在一些实施方案中,两个或更多个实体在物理上彼此相关,直接或间接地相互作用,使得它们在物理上彼此接近和/或保持在物理上彼此接近。在一些实施方案中,物理上彼此相关的两个或更多个实体彼此共价连接;在一些实施方案中,物理上彼此相关的两个或更多个实体不彼此共价连接,而是非共价相关,例如通过氢键、范德华相互作用、疏水相互作用、磁性及其组合。

[0155] 条形码:如本文所用,术语"条形码"是指能够在更大的异质性实体群中基本上区分实体(例如,分子、核酸、病毒、细胞等)或实体的组合的任何分子特征(例如,核酸序列)。在一些实施方案中,条形码是一种工程化序列。在一些实施方案中,条形码是一种工程化核酸序列。在一些实施方案中,条形码是文库构建体的一部分。在一些实施方案中,文库构建体可包含多于一个条形码。在文库构建体包含一个或多个条形码的情况下,一个或多个条形码可包含在(i)标识符内,(ii)标识符外,或(iii)标识符内外。在一些实施方案中,一个或多个条形码在检测(例如,通过下一代测序方法)时指示未被直接检测(例如,通过下一代测序方法)的一个或多个文库变体的身份。在一些实施方案中,条形码可以包含来自已知核酸序列池内的核酸序列。在一些实施方案中,条形码是外源性核酸序列。在一些实施方案中,条形码是内源性核酸序列。在一些实施方案中,条形码可以包含外源性序列和内源性序列。通常,条形码包含长度在3个核苷酸至50个核苷酸范围内的序列。例如,在一些实施方案中,条形码包含长度在由下限和上限(上限大于下限)界定的范围内的序列。在一些实施方案中,下限可以是约5个核苷酸、约6个核苷酸、约7个核苷酸、约8个核苷酸、约9个核苷酸、约10个核苷酸、约11个核苷酸、约12个核苷酸、约13个核苷酸、约14个核苷酸,或约15个核苷酸。在一些实施方案中,上限可以是约10个核苷酸、约15个核苷酸、约20个核苷酸、约25个核苷酸、约30个核苷酸、约35个核苷酸、约40个核苷酸、约45个核苷酸,或约50个核苷酸。在一些特定实施方案中,条形码包含长度在5个核苷酸至25个核苷酸范围内的序列。在一些特定实施方案中,条形码包含长度为约5个核苷酸、约10个核苷酸、约15个核苷酸、约20个核苷酸

或约25个核苷酸的序列。

[0156] 生物制剂:如本文所用,术语“生物制剂”在其最广泛的意义上是指由活生物体制成或源自活生物体(例如,在活系统中制造)的制品。根据本公开,生物制剂包括但不限于蛋白质、多肽、核酸、多核苷酸、病毒、病毒载体、治疗血清、毒素、抗毒素、疫苗、过敏原提取物、血液组分或衍生物、基因治疗产品、人体组织或细胞产品。在一些实施方案中,生物制剂是或包含可以用于治疗、治愈、减轻、预防或诊断严重或危及生命的疾病或医学病状的基于多肽的分子。在一些实施方案中,生物制剂是治疗性生物制剂。治疗性生物制剂是可以用于预防、治疗或治愈疾病或病状(例如,在哺乳动物中,例如,在人中)的那些。在一些实施方案中,生物制剂是诊断性生物制剂。在一些实施方案中,生物制剂用于制造(例如,制造基因治疗产品或细胞治疗产品)。在一些特定实施方案中,生物制剂是病毒载体。

[0157] 细胞身份序列:如本文所用,是指用于标记所有来自特定昆虫细胞的待测序的核酸以作为单细胞测序方法的一部分的核酸序列。本公开涵盖以下认识:使用细胞身份序列的单细胞测序可用于文库构建体包含多个不连续核酸序列的方法。在此类方法中,给定的细胞身份序列可以提供存在于相同细胞中的所有标识符和/或文库变体的关联。使用不连续文库构建体和单细胞测序的方法的示例性示意图描绘于图3。然而,本公开还涵盖以下认识:使用细胞身份序列的单细胞测序也可以在一些实施方案中用于文库构建体包含单个连续核酸序列的方法。本公开还涵盖以下认识:在单细胞测序方法过程中,表达的RNA的逆转录过程中特异性附加单细胞身份序列。应当理解,旨在用于单细胞测序的构建体或核酸应包含在表达的RNA中,使得在逆转录步骤过程中可以使用适当的引物用细胞身份序列对所有转录物进行单细胞标记。

[0158] 顺式作用整合序列:如本文所用,关于文库构建体的短语“顺式作用整合序列”是指通过一个或多个反式作用整合序列和/或其多肽、蛋白质、核酸或多核苷酸产物促进整合到细胞基因组(例如昆虫的细胞基因组)中的文库构建体本身上的核酸序列。在一些实施方案中,顺式作用整合序列包含在文库构建体中,并且在待整合到细胞基因组中的文库构建体的一部分的侧翼。在一些实施方案中,顺式作用整合序列包含同源臂、识别位点和/或病毒重复序列。

[0159] 可比较的:如本文所用,术语“可比较的”是指两种或更多种剂、实体、情况、条件集合、受试者、群体等,它们可能彼此不相同,但足够类似以允许它们之间的比较,使得本领域技术人员将理解,可以基于观察到的差异或类似性合理地得出结论。在一些实施方案中,可比较的剂的集合、实体、情况、条件集合、受试者、群体等的特征在于多个基本相同的特征和一个或少量不同的特征。本领域的普通技术人员将理解,在上下文中,在任何给定情况下,对于两种或更多种此类剂、实体、情况、条件集合、受试者、群体等需要怎样的同一性程度才被认为是可比较的。例如,本领域的普通技术人员将理解,当剂的集合、实体、情况、条件集合、受试者、群体等在特征为以下时彼此是可比较的:足够数目和类型的基本上相同的特征,以保证得出以下合理结论:在不同的环境集合、刺激、剂、实体、情况、条件集合、受试者、群体等下获得的结果或观察到的现象的差异是由那些不同特征的变化引起的或指示那些不同特征的变化。

[0160] 互补:如本文所用,术语“互补”是指根据标准沃森-克里克互补规则(Watson-Crick complementary rule)(腺嘌呤“A”碱基与胸腺嘧啶“T”配对,并且鸟嘌呤“G”碱基与

胞嘧啶“C”配对)进行碱基配对的核苷酸或核苷酸序列。“100％互补”的核苷酸序列或表现出“100％互补性”的核苷酸序列是在两个核苷酸序列的至少一个的整体上彼此碱基配对的核苷酸序列。寡核苷酸可以与比寡核苷酸长的模板多核苷酸“100％互补”(即,如果寡核苷酸的整个序列与模板多核苷酸的一部分碱基配对,则寡核苷酸与模板多核苷酸“100％互补”)。然而,“互补”的核酸序列不必是100％互补。通常,关于两个或更多个核酸序列的术语“互补”是指两个核酸序列之间存在足够的互补性,使得它们在严格条件下和/或在扩增方法(例如,PCR或LCR)的退火阶段期间使用的温度下杂交。

[0161] 构建体:如本文所用,术语“构建体”是指能够携带至少一个多核苷酸的实体。在一些实施方案中,构建体可以是质粒、转座子、粘粒或人工染色体(例如,人人工染色体(HAC)、酵母人工染色体(YAC)、细菌人工染色体或P1衍生的人工染色体(PAC))。在一些实施方案中,构建体有助于将多核苷酸转移至细胞。某些构建体能够在其所引入的宿主细胞中自主复制(例如,具有复制起点的质粒)。其他构建体可以整合到宿主细胞的基因组中,从而与宿主基因组一起复制。在一些实施方案中,提供了编码病毒载体的一个或多个元件的构建体。例如,在一些实施方案中,提供了编码用于在昆虫细胞中表达的病毒载体的一个或多个元件的质粒。在一些实施方案中,提供了编码病毒载体的一个或多个元件的构建体。在一些实施方案中,构建体是病毒载体。

[0162] 对照:如本文所用,术语“对照”是指“对照”的本领域理解的含义,是与结果进行比较的标准。通常,对照用于通过分离变量以便对此类变量做出结论来增强实验的完整性。在一些实施方案中,对照是与测试反应或测定同时进行以提供比较物的反应或测定。例如,在一个实验中,应用了“测试”(即,被测试的变量)。在第二个实验中,应用了“对照”,不应用被测试的变量。在一些实施方案中,对照是历史对照(例如,先前进行的测试或测定的历史对照,或先前已知的量或结果)。在一些实施方案中,对照是或包括打印的或以其他方式保存的记录。在一些实施方案中,对照是阳性对照。在一些实施方案中,对照是阴性对照。

[0163] 对应于:如本文所用,在多肽和核酸的上下文中,术语“对应于”通过与适当的参考多肽或核酸序列比较分别指定氨基酸残基或核苷酸残基的位置/身份。例如,在一些实施方案中,聚合物中的单体残基(例如,多肽中的氨基酸残基或多核苷酸中的核酸残基)可以被鉴定为“对应于”适当的参考聚合物中的残基。例如,本领域普通技术人员将理解,出于简单的目的,通常使用基于参考相关多肽的规范编号系统来指定多肽中的残基,使得“对应于”位置190处残基的氨基酸,例如,实际上不必是特定氨基酸链中的第190个氨基酸,而是对应于参考多肽中位置190处发现的残基;本领域普通技术人员容易理解如何鉴定“对应的”氨基酸(参见例如,Benson等人Nucl.Acids Res.(2013年1月1日)41(D1):D36-D42;Pearson等人PNAS第85卷,第2444-2448页,1988年4月)。本领域的技术人员将了解多种序列比对策略,包括可以用于例如根据本公开鉴定多肽和/或核酸中“对应的”残基的软件程序(例如,BLAST、FASTA等)。

[0164] 内源性:如本文所用,术语“内源性”是指存在于其自然环境中的任何事物。关于核酸,“内源性的”可以例如是指源自细胞内部的核酸或其序列。例如,内源性核酸序列是天然存在于昆虫细胞基因组中而无需人为操作(例如,在其天然位置并在天然表达元件的控制下)的核酸序列。

[0165] 工程化的:如本文所用,术语“工程化的”是指已由人为操作的方面。在核酸的上下

文中,工程化可以包括可以对核酸进行的任何类型的修饰。在一些实施方案中,当序列已经被人为操作以产生在自然环境中不存在的序列,例如,包括缺失、插入、取代、替换、重排和/或融合至第二序列时,多核苷酸可以被认为是“工程化的”。例如,工程化多核苷酸可以包含两个或更多个序列,所述两个或更多个序列在自然界中并未按此顺序连接在一起,而是因为它们被人为操作以在工程化多核苷酸中彼此直接连接。类似地,如果细胞或生物体已经被操作使得其遗传信息改变(例如,先前不存在的新遗传物质已经例如通过转化、配对、体细胞杂交、转染、转导或其他机制引入,或先前存在的遗传物质例如通过取代或缺失突变或通过配对方案改变或去除),则认为细胞或生物体是“工程化的”。在一些实施方案中,工程化细胞(例如,工程化昆虫细胞)是指已经经过操作,使得其遗传、表观遗传和/或表型特性相对于适当的参考细胞(诸如其他没有被如此操作的相同细胞)发生改变的细胞。在一些实施方案中,操作是或包括遗传操作。在一些实施方案中,工程化细胞是已经被操作,使得其含有和/或表达改变的量和/或根据相对于此类适当参考细胞改变的时序的关注的特定剂(例如,蛋白质、核酸和/或其特定形式)的细胞。按照惯例并且被本领域技术人员理解,工程化多核苷酸或细胞的后代典型地仍被称为“工程化的”,即使实际操作是在先前实体上进行的。

[0166] 工程化序列:如本文所用,短语“工程化序列”是指已经经过人为操作的序列(例如,核苷酸的序列)。在一些实施方案中,工程化序列是工程化核酸序列。在一些实施方案中,工程化核酸序列可以包括在5’至3’方向和/或3’至5’方向上读取的序列。工程化核酸序列可以包括可以对核酸进行的任何类型的修饰(例如,引入、取代、缺失、替换、重排、表观遗传修饰等)。在一些实施方案中,工程化序列是或包括昆虫细胞基因组序列的遗传修饰。在一些实施方案中,工程化序列是或包括昆虫细胞基因组DNA的表观遗传修饰。在一些实施方案中,工程化序列是或包括游离序列(例如,引入或修饰游离存在于细胞内的序列)。在一些实施方案中,工程化序列是或包括具有一种或多种表观遗传特征的游离序列。在一些实施方案中,工程化序列是或包括标识符。在一些实施方案中,工程化序列是或包括多核苷酸,所述多核苷酸包含一个或多个产生病毒载体所必需的核酸序列。在一些实施方案中,工程化序列是或包括扰动。在一些实施方案中,工程化序列是或包括文库变体。在一些实施方案中,工程化序列是或包括有效载荷。在一些实施方案中,工程化序列是或包括扰动附加序列。在一些实施方案中,工程化序列是或包括反式作用整合序列。在一些实施方案中,工程化序列是或包括顺式作用整合序列。在一些实施方案中,工程化序列是或包括文库构建体。在一些实施方案中,工程化序列是或包括条形码。

[0167] 游离基因:如本文所用,术语“游离基因”是指可以自主复制的染色体外遗传物质。在一些实施方案中,游离的遗传物质(例如,DNA)可以在质粒、粘粒、福斯质粒、人工染色体(例如,人人工染色体(HAC)、酵母人工染色体(YAC)、细菌人工染色体或P1衍生人工染色体(PAC))或病毒载体(例如,杆状病毒载体)的环境中。

[0168] 外源性:如本文所用,术语“外源性”是指被引入或已经被引入生物体或细胞中的任何实体。例如,“外源性核酸”是源自生物体或细胞外部的核酸。在一些实施方案中,昆虫细胞中的外源性核酸已经透过细胞膜被引入(例如,通过人为)。在一些实施方案中,外源性核酸可以是或包括存在于非天然环境中(例如,在不同位置和/或在非天然表达元件的控制下)的天然基因组中的核苷酸序列。在一些实施方案中,外源性核酸可以是或包括先前不存

在于生物体或细胞(例如,来自不同生物体)的基因组中的核苷酸序列。外源性核酸包括外源性基因。“外源性基因”是已经引入生物体或细胞(例如,通过转化/转染)的核酸或其序列,其编码RNA和/或蛋白质的表达,并且在本文中也被称为“转基因”。

[0169] 染色体外:如本文所用,术语“染色体外”是指不包含在染色体中的遗传物质。在昆虫细胞的情况下,染色体包括核染色体和线粒体染色体。染色体外遗传物质包括游离遗传物质。

[0170] 基因:如本文所用,术语“基因”是指在染色体中编码基因产物(例如,RNA产物,例如,多肽产物)的DNA序列。在一些实施方案中,基因包括编码序列(即,编码特定产物的序列)。在一些实施方案中,基因包括非编码序列。在一些特定实施方案中,基因可以包括编码(例如,外显子)序列和非编码(例如,内含子)序列两者。在一些实施方案中,基因可以包括一种或多种调控序列(例如启动子、增强子等)和/或内含子序列,所述序列例如可以控制或影响基因表达(例如,细胞类型特异性表达、诱导型表达等)的一个或多个方面。如本文所用,术语“基因”通常是指编码多肽或其片段的核酸的一部分;所述术语可以任选地涵盖调控序列,这对于本领域普通技术人员来说从上下文中将是清楚的。此定义并非旨在排除将术语“基因”应用于非蛋白质编码表达单位,而是为了澄清,在大多数情况下,如本文中使用的术语是指多肽编码核酸。在一些实施方案中,基因可以编码多肽,但所述多肽可能无功能,例如,基因变体可以编码相对于野生型基因不以相同方式发挥功能或根本不发挥功能的多肽。

[0171] 基因组:如本文所用,术语“基因组”是指个体生物体或细胞携带的总遗传信息,由其染色体的完整核酸序列表示。

[0172] 引导序列:如本文所用,术语“引导序列”是指对应于用于核酸酶介导的编辑(例如,用RNA引导的核酸酶)的引导RNA的核酸序列的核酸序列。术语“引导RNA”和“gRNA”是指促进RNA引导的核酸酶(诸如Cas9或Cpf1)与靶序列(诸如细胞中的基因组或游离序列)的特定缔合(或“靶向”)的任何核酸。gRNA可以是单分子的(包含单个RNA分子,称为嵌合的),或模块化的(包含超过一个,通常是两个独立的RNA分子,诸如crRNA和tracrRNA,它们通常彼此缔合,例如通过双链体化)。引导RNA,无论是单分子的还是模块化的,都包含“引导序列”,其与靶标内的序列(诸如需要编辑的细胞基因组中的DNA序列)完全或部分互补。引导序列在文献中被称为各种名称,包括但不限于“靶向结构域”、“互补区”(例如,Cotta-Ramusino等人的WO 2016/073990)、“间隔区”(例如,Briner等人,Molecular Cell 56(2),333-339,2014年10月23日)和一般地“crRNA”(例如,Jiang等人Nat Biotechnol.2013年3月;31(3):233-239)。不管它们的名称如何,引导序列的长度通常为约10至30个核苷酸。在一些实施方案中,引导序列的长度为15至25个核苷酸。在某些实施方案中,引导序列的长度为16至24个核苷酸(例如,长度为16、17、18、19、20、21、22、23或24个核苷酸)。在一些实施方案中,引导序列位于或接近gRNA的5’末端(例如,对于Cas9或从其衍生或获得的核酸酶)。在一些实施方案中,引导序列位于或接近gRNA的3’末端(例如,对于Cpf1或从其衍生或获得的核酸酶)。

[0173] 异源性:如本文所用,术语“异源性”是指在存在的特定环境(例如,细胞或生物体)中不天然存在的任何实体。在核酸的环境中,异源性是指不天然一起(例如,在相同多核苷酸或相同细胞中)出现的序列。在一些实施方案中,异源性序列可以是将序列重排、替换、插入、取代至非内源性环境(例如,来自不同的基因组位置或来自不同的生物体)中。例如,异

源性启动子序列或异源性增强子序列可以是与不同基因天然相关或来自不同生物体的序列。在一些实施方案中,异源性序列存在于昆虫细胞的基因组中。在一些实施方案中,异源性序列游离存在于昆虫细胞中。

[0174] 同源性:如本文所用,术语“同源性”或“同源物”是指寡核苷酸分子(例如,DNA分子和/或RNA分子)之间和/或多肽分子之间的总体关联性。在一些实施方案中,如果寡核苷酸分子(例如,DNA分子和/或RNA分子)和/或多肽分子的序列至少15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或99%相同,则认为它们彼此“同源”。在一些实施方案中,如果寡核苷酸分子(例如,DNA分子和/或RNA分子)和/或多肽分子的序列至少25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或99%相似(例如,在相应位置含有具有相关化学性质的残基),则认为它们彼此“同源”。例如,如本领域普通技术人员所熟知的,某些氨基酸通常彼此类似地被归类为“疏水”或“亲水”氨基酸,和/或具有“极性”或“非极性”的侧链。一种氨基酸取代为另一种相同类型的氨基酸通常可以被认为是“同源”取代。

[0175] 标识符:如本文所用,是指(i)可以被检测(例如,通过下一代测序)和(ii)使鉴定产生或衍生病毒载体的昆虫细胞或克隆细胞系成为可能的元件。本公开提供了在昆虫细胞和病毒载体之间共享和/或转移的标识符。在一些实施方案中,标识符包含核酸序列。在一些实施方案中,标识符包含一个或多个条形码、一个或多个文库变体或其组合。在一些实施方案中,标识符包含可以例如通过PCR(例如,定量PCR)、杂交(例如,使用探针)和/或测序(例如,下一代测序和/或Sanger测序)检测的序列。在一些实施方案中,标识符是或包含在昆虫细胞和病毒载体之间共享和/或转移的核酸序列。一些实施方案中,标识符包含在文库构建体中(例如用于引入到昆虫细胞中)。在一些实施方案中,标识符在适合于将标识符包装到病毒载体中的遗传结构(例如,病毒重复序列,例如,AAV ITR序列)之间包含在文库构建体中。在一些实施方案中,标识符存在于病毒载体的基因组中(例如,在病毒重复序列之间)。在一些实施方案中,病毒载体基因组中(例如,病毒重复序列之间)的标识符的检测与产生或衍生病毒载体的昆虫细胞或克隆细胞系相关。在一些实施方案中,病毒载体基因组中(例如,病毒重复序列之间)的标识符的检测与同产生或衍生病毒载体的昆虫细胞或克隆细胞系相关的一个或多个文库构建体和/或一个或多个文库变体的存在相关。在一些实施方案中,提供的方法可以包括对标识符进行测序(例如,使用下一代测序方法)以确定特定病毒载体在病毒载体池或样品中的相对丰度的步骤。

[0176] 同一性:如本文所用,术语“同一性”是指聚合物分子之间,例如核酸分子(例如,DNA分子和/或RNA分子)之间和/或多肽分子之间的总体关联性。例如,两个核酸或多肽序列的百分比同一性的计算可以通过出于最佳比较目的而比对两个序列来进行(例如,可以在第一序列和第二序列中的一者或二者中引入空位以进行最佳比对,并且可以出于比较目的忽略非同一性序列)。在一些实施方案中,出于比较目的比对的序列长度是参考序列长度的至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或基本上100%;然后比较对应位置处的核苷酸。当第一序列中的位置被与第二序列中的对应位置相同的残基(例如,核苷酸或氨基酸)占据时,两个分子(即,第一和第二)在所述位置处是相同的。两个序列之间的百分比同一性是将为了两个序列的最佳比对而需要引入的空位的数量和每个空位的长度考虑在内,由被比较的两个序列共享的相同位置的数量的函数。两个

序列之间的序列比较和百分比同一性的确定可以使用数学算法来完成。例如,可以使用Meyers和Miller的算法(CABIOS,1989,4:11-17,其以引用的方式整体并入本文)确定两个核苷酸序列之间的百分比同一性,所述算法已经并入ALIGN程序(2.0版)中。

[0177] 整合构建体:如本文所用,是指促进构建体诸如文库构建体的基因组整合的构建体。在一些实施方案中,整合构建体是昆虫病毒载体例如像鹿眼蛱蝶浓核病毒(JcDNV)或其衍生物。在一些实施方案中,整合构建体包括顺式作用整合序列和/或反式作用整合序列。

[0178] 分离的:如本文所用,是指物质和/或实体已经(1)与最初产生时(无论是在自然界和/或在实验环境中)与其相关的组分中的至少一些分离,且/或(2)由人为设计、产生、制备和/或制造。分离的物质和/或实体可以与最初与其相关的其他组分的约10%、约20%、约30%、约40%、约50%、约60%、约70%、约80%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或多于约99%分离。在一些实施方案中,分离的剂的纯度为约80%、约85%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或多于约99%。如本文所用,如果物质基本上不含其他组分,则所述物质是“纯的”。在一些实施方案中,如本领域技术人员将理解的,在已经与某些其他组分诸如像一种或多种载剂或赋形剂(例如,缓冲液、溶剂、水等)组合之后,物质仍可以被认为是“分离的”或甚至是“纯的”;在此类实施方案中,计算物质的百分比分离或纯度而不包括此类载剂或赋形剂。仅举一个实例,在一些实施方案中,自然界中存在的生物聚合物诸如多肽或多核苷酸在以下情况下被认为是“分离的”:a)凭借其来源或衍生源不与在自然界中以天然状态伴随它的一些或全部组分缔合;b)其基本上不含来自自然界中产生它的物种的相同物种的其他多肽或核酸;c)其由来自非自然界产生它的物种的细胞或其他表达系统的组分表达或以其他方式与所述组分相关。因此,例如,在一些实施方案中,化学合成的或在不同于自然产生它的细胞系统中合成的多肽被认为是“分离的”多肽。或者或另外,在一些实施方案中,已经经过一种或多种纯化技术的多肽可以被认为是“分离的”多肽,程度是其已经从其他组分分离,所述其他组分a)在自然界中与其相关;且/或b)在最初产生时与其相关。

[0179] 文库构建体:如本文所用,是指一个或多个用于产生昆虫细胞文库的核酸构建体。如本文所用,文库构建体包含(i)至少一个标识符,(ii)适合于将标识符包装到病毒载体中的遗传结构(例如,病毒重复序列,例如,AAV ITR序列),和(iii)任何文库变体(如果尚未包含在标识符中)。文库构建体还可以包含其他工程化序列,诸如但不限于任何顺式作用整合序列(如果文库构建体的基因组整合是相关联的),和/或任何条形码。在一些实施方案中,文库构建体的所有元件都包含在单个(即,一个)连续核酸(例如,DNA)片段上。在一些实施方案中,文库构建体是指多个、单独的且不连续核酸(例如,DNA)片段。当文库构建体包含多个核酸时(即,不连续文库构建体),它将包括至少一个包含位于用于将标识符包装到病毒载体中的序列(例如,病毒重复序列,例如,AAV ITR序列)之间的标识符的构建体。在文库构建体不连续的一些实施方案中,文库构建体还包含一个或多个构建体,其中每个单独的构建体包含一个或多个文库变体。所提供的采用不连续文库构建体的方法还将包括鉴定文库构建体的一个或多个文库变体的步骤(例如,昆虫细胞的单细胞测序)。例如,图3描绘了一种方法,其中文库构建体包含单独的DNA片段,其显示包含有包含标识符的构建体(例如,侧翼为ITR的条形码)和包含文库变体的三个其他构建体(例如,三个独立且独特的gRNA)的细胞,其中标识符(例如,条形码)被包装到病毒载体中,并通过单细胞测序确定与文库变体的

关联。可以设想,对于采用单个连续文库构建体的许多方法,单细胞测序并不总是需要的,但在一些情况下可能仍然有用。

[0180] 文库变体:如本文所用,是指文库构建体的元件,其会产生在细胞之间变化的扰动。例如,在细胞群中,每个细胞都会得到不同的文库构建体,所述文库构建体包含独特的文库变体或文库变体的组合。例如,文库变体可以包含基因、ORF、gRNA序列、非编码核酸或其多个和/或组合。在本公开的上下文中,文库变体不同于条形码,但是一个或多个文库变体可以与一个或多个条形码相关。在一些实施方案中,一个或多个文库变体和/或条形码可以包含在文库构建体内的标识符之内或在其之外。在一些实施方案中,可以(例如,在下一代测序方法中)直接检测一个或多个条形码本身而不是文库变体本身。在一些实施方案中,文库变体在本文中可以称为效应物,由此文库变体影响或引起在细胞间变化的扰动。在一些实施方案中,文库变体本身可以成为在细胞间变化的扰动。例如,在一些实施方案中,是gRNA的文库变体是效应物,其连同RNA引导的核酸酶(例如,扰动附加序列)导致细胞基因组DNA内的缺失。在其他实施方案中,是ORF或基因序列的文库变体在其转染到细胞中以及在一些情况下整合到基因组DNA中(例如,如通过反式作用和顺式作用整合序列进行)时,本身成为细胞的遗传物质的扰动或修饰。

[0181] 核酸:如本文所用,在其最广泛的意义上,是指被并入或可以被并入寡核苷酸链中的任何化合物和/或物质。在一些实施方案中,核酸是或是可以经由磷酸二酯键联并入寡核苷酸链中的化合物和/或物质。从上下文将清楚地看出,在一些实施方案中,“核酸”是指单独的核酸残基(例如,核苷酸和/或核苷);在一些实施方案中,“核酸”是指包含单独的核酸残基的寡核苷酸链。在一些实施方案中,“核酸”是或包括RNA;在一些实施方案中,“核酸”是或包括DNA。在一些实施方案中,核酸是、包含一个或多个天然核酸残基,或由其组成。在一些实施方案中,核酸是、包含一个或多个核酸类似物,或由其组成。

[0182] 有效载荷:如本文所用,术语“有效载荷”是指通过本公开的方法产生的病毒载体递送的任何关注的实体。例如,可以期望将此种有效载荷被引入细胞、器官、生物体和/或生物系统(例如,包括细胞)。在一些实施方案中,有效载荷序列是或包含通过本公开的病毒载体递送的异源性核酸序列。在一些实施方案中,有效载荷序列包含以下中的一项或多项:编码区、基因调控元件和转录终止子。基因调控元件的非限制性实例包括启动子、转录激活子、增强子和聚腺苷酸化信号。在一些实施方案中,有效载荷序列包含编码区、基因调控元件和转录终止子,它们相对于彼此定位,使得编码区位于基因调控元件和转录终止子之间。在一些实施方案中,编码区编码基因产物。在一些实施方案中,基因产物是RNA。在一些实施方案中,编码区编码多肽(诸如蛋白质,诸如糖蛋白)。在一些实施方案中,编码区编码融合多肽和/或嵌合多肽。在一些实施方案中,编码区编码一种基因产物。在一些实施方案中,编码区编码多于一种基因产物(例如,2、3、4、5、6、7或更多种基因产物)。在一些实施方案中,编码区编码调控RNA(例如,siRNA、微小RNA等)。在一些特定实施方案中,有效载荷编码一个或多个用于基因编辑的实体(例如,gRNA介导的编辑系统)。在一些实施方案中,有效载荷编码蛋白质产物。

[0183] 扰动:如本文所用,是指昆虫细胞、产生的病毒载体、产生病毒载体所必需的多核苷酸和/或其他构建体中的遗传修饰,所述遗传修饰由如本文所述的方法产生和/或鉴定。在一些实施方案中,昆虫细胞、产生的病毒载体、产生病毒载体所必需的多核苷酸和/或其

他构建体包含一个或多个扰动。这些是从本公开的用于表达和/或产生病毒载体的昆虫细胞文库中生成、产生、鉴定和/或选择的。在一些实施方案中,扰动是一个或多个文库变体的结果。在一些实施方案中,扰动是遗传修饰,其不是文库变体的结果,而是由如本文所述的方法产生和/或鉴定的遗传修饰。在一些实施方案中,扰动包含在至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸中的遗传修饰。在一些实施方案中,扰动与昆虫细胞(例如,用于表达病毒载体(例如,独立地和/或合成地))或病毒载体的一种或多种期望特征相关。在一些实施方案中,扰动包括基因组序列变化(例如,基因组插入、缺失、取代、重排等)、游离序列变化和/或表观遗传修饰。

[0184] 扰动附加序列:如本文所用,其包括有助于与文库构建体组合产生扰动的任何序列。例如,在一些实施方案中,文库构建体包含有包含gRNA的文库变体,并且扰动附加序列包含编码RNA引导的核酸酶或用于核酸酶介导的扰动的其他元件的序列。

[0185] 多肽:如本文所用,术语“多肽”是指通常通过肽键连接的残基(例如,氨基酸)的任何聚合链。在一些实施方案中,多肽具有天然存在的氨基酸序列。在一些实施方案中,多肽具有非天然存在的氨基酸序列。在一些实施方案中,多肽具有被工程化的氨基酸序列,因为它是通过人为的作用设计和/或产生的。在一些实施方案中,多肽可以包含天然氨基酸、非天然氨基酸或两者,或由其组成。在一些实施方案中,多肽可以包含一个或多个侧基或其他修饰,例如,在多肽的N-末端、在多肽的C-末端或其任何组合处修饰或附接到一个或多个氨基酸侧链。在一些实施方案中,此类侧基或修饰可以是乙酰化、酰胺化、脂化、甲基化、聚乙二醇化等,包括其组合。在一些实施方案中,多肽可以含有L-氨基酸、D-氨基酸或两者,并且可以含有本领域已知的多种氨基酸修饰或类似物中的任一种。在一些实施方案中,有用的修饰可以是或者包括例如末端乙酰化、酰胺化、甲基化等。在一些实施方案中,蛋白质可以包含天然氨基酸、非天然氨基酸、合成氨基酸及其组合。术语“肽”通常用于指代长度小于约100个氨基酸、小于约50个氨基酸、小于20个氨基酸或小于10个氨基酸的多肽。在一些实施方案中,蛋白质是抗体、抗体片段、其生物学活性部分和/或其特征部分。

[0186] 多核苷酸:如本文所用,术语“多核苷酸”是指任何核酸的聚合链。在一些实施方案中,多核苷酸是或包含RNA;在一些实施方案中,多核苷酸是或包含DNA。在一些实施方案中,多核苷酸是、包含一个或多个天然核酸残基,或由其组成。在一些实施方案中,多核苷酸是、包含一个或多个核酸类似物,或由其组成。在一些实施方案中,多核苷酸类似物与核酸的不同之处在于多核苷酸类似物不利用磷酸二酯主链。或者或另外,在一些实施方案中,多核苷酸具有一个或多个硫代磷酸和/或5’-N-亚磷酰胺键联,而不是磷酸二酯键。在一些实施方案中,多核苷酸是、包含一个或多个天然核苷(例如,腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷和脱氧胞苷),或由其组成。在一些实施方案中,多核苷酸是、包含一个或多个核苷类似物(例如,2-氨基腺苷、2-硫代胸苷、肌苷、吡咯并嘧啶、3-甲基腺苷、5-甲基胞苷、C-5丙炔基-胞苷、C-5丙炔基-尿苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮杂腺苷、7-脱氮杂鸟苷、8-氧代腺苷、8-氧代鸟苷、O(6)-甲基鸟嘌呤、2-硫代胞苷、甲基化碱基、插入的碱基及其组合),或由其组成。在一些实施方案中,多核苷酸包含一个或多个与天然核酸中的糖相比修饰的糖(例如,2’-氟核糖、核糖、2’-脱氧核糖、阿拉伯糖和己糖)。在一些实施方案中,多核苷酸具有编码功能性基因产物诸如RNA或蛋白质的核苷酸序列。在一些实施方案中,多

核苷酸包含一个或多个内含子。在一些实施方案中,通过以下中的一种或多种方法制备多核苷酸:从天然来源分离、通过基于互补模板的聚合进行酶促合成(体内或体外)、在重组细胞或系统中复制和化学合成。在一些实施方案中,多核苷酸的长度为至少3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190、20、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000或更多个残基。在一些实施方案中,多核苷酸部分或全部是单链;在一些实施方案中,多核苷酸部分或全部是双链。在一些实施方案中,多核苷酸具有包含至少一个编码多肽或是编码多肽的序列的补体的元件的核苷酸序列。在一些实施方案中,多核苷酸具有酶活性。

[0187] 预定的:如本文所用,术语"预定的"是指在实验和/或分析开始之前。例如,在一些实施方案中,当一组可能的结果(例如,插入位点)在引入工程化序列(例如,工程化序列可以被插入几个不同的基因组位置之一)的物理行为之前已知时,工程化序列的位置或特征可以被认为是预定的。

[0188] 蛋白质:如本文所用,术语"蛋白质"是指多肽(即,通过肽键彼此连接的至少两个氨基酸的串)。蛋白质可以包含除氨基酸之外的部分(例如,可以是糖蛋白、蛋白聚糖等)且/或可以以其他方式加工或修饰。本领域普通技术人员将理解,"蛋白质"可以是如由细胞产生的完整多肽链(具有或不具有信号序列),或可以是其特征部分。普通技术人员将理解,蛋白质有时可以包含多于一条例如通过一个或多个二硫键连接或者通过其他手段缔合的多肽链。

[0189] 参考:如本文所用,术语"参考"描述了标准或对照,比较是相对于所述标准或对照进行的。例如,在一些实施方案中,将关注的剂、动物、个体、群体、样品、序列或值与参考或对照剂、动物、个体、群体、样品、序列或值分别进行比较。在一些实施方案中,参考或对照与关注的测试或测定基本上同时进行测试和/或测定。在一些实施方案中,参考或对照是历史参考或对照,任选地体现在有形介质中。通常,如本领域技术人员将理解的,将参考或对照在与评估下的条件或环境可比较的条件或环境下测定或表征。本领域技术人员将理解,何时存在足够类似性以证明依赖于特定可能的参考或对照和/或与特定可能的参考或对照比较是合理的。在一些实施方案中,参考是阴性对照参考;在一些实施方案中,参考是阳性对照参考。在一些实施方案中,与参考细胞或参考细胞群进行比较,所述参考细胞或参考细胞群具有可比较的遗传特征并且已经在可比较的条件下培养(除了关于正在分析的变量之外)。在一些特定实施方案中,参考细胞在至少一个工程化序列和/或至少一个条形码序列的存在方面不同,但在其他方面是可比较的。

[0190] RNA引导的核酸酶:如本文所用,术语"RNA引导的核酸酶"是指以序列特异性方式结合特定靶核苷酸序列的多肽,并且所述多肽由与多肽复合并与靶序列杂交的引导RNA分子导向靶核苷酸序列。尽管RNA引导的核酸酶能够在结合后切割靶序列,但术语RNA引导的核酸酶还涵盖能够结合但不切割靶序列的无核酸酶活性的RNA引导的核酸酶。RNA引导的核酸酶对靶序列的切割可以导致单链或双链断裂。仅能够切割双链核酸分子的单链的RNA引导的核酸酶在本文中被称为切口酶。在一些实施方案中,RNA引导的核酸酶是或源自Cas9、Cas Z、CpfI和/或Fok1。

[0191] 反式作用整合序列:如本文所用,关于文库构建体的短语"反式作用整合序列"是

指不一定作为文库构建体本身的一部分包含的核酸序列,其是将文库构建体整合到昆虫细胞基因组中所必需的。反式作用整合序列和/或其多肽、蛋白质、核酸或多核苷酸产物(例如,重组酶)实现与顺式作用整合序列(例如,重组酶位点)协同将文库构建体或其一部分整合到细胞基因组(例如,昆虫细胞)中。在一些实施方案中,多于一个反式作用整合序列对于文库构建体的整合是必需的。

[0192] 转化或转染:如本文所用,术语“转化”或“转染”是指将外源性DNA引入到宿主细胞(例如,昆虫宿主细胞)中的任何过程。转化可以在自然或人工条件下使用本领域众所周知的各种方法进行。转化可以依赖任何已知方法将外源核酸序列插入真核宿主细胞。在一些实施方案中,特定转化方法是基于被转化的宿主细胞来选择的,并且可以包括但不限于病毒感染或转导、电穿孔、脂质体转染。在一些实施方案中,“转化的”细胞被稳定转化,因为插入的DNA能够作为自主复制质粒或作为宿主染色体的一部分进行复制。在一些实施方案中,转化的细胞在有限时间段内瞬时表达引入的核酸。

[0193] 病毒载体:如本文所用,术语“病毒载体”是指(i)能够携带至少一个多核苷酸,并且(ii)包含病毒蛋白(例如,衣壳蛋白,例如,病毒衣壳蛋白和/或其变体或衍生物)的实体。在一些实施方案中,病毒载体包含一个或多个核酸分子。在一些实施方案中,病毒载体可以有助于核酸转移至细胞。在一些实施方案中,病毒载体包含一个或多个核酸序列和一个或多个病毒衣壳蛋白。在一些实施方案中,病毒载体包含源自腺相关病毒(AAV)载体、腺病毒载体、杆状病毒载体、慢病毒载体和/或逆转录病毒载体的衣壳蛋白和/或核酸序列。在一些实施方案中,病毒载体包含包膜。

[0194] 本申请中引用的所有文献和类似材料,包括但不限于专利、专利申请、文章、书籍、论文和网页,无论此类文献和类似材料的格式如何,均以引用的方式明确地整体并入。如果一个或多个并入的文献和类似材料与本申请不同或矛盾,包括但不限于定义的术语、术语使用、描述的技术等,以本申请为准。本文所使用的小节标题仅出于组织性目的并且不解释为以任何方式限制所述的主题。

具体实施方式

[0195] 生物制剂包括细胞和基因疗法、疫苗和生物试剂,是一个快速增长的市场。昆虫细胞可以有效表达重组蛋白,并可用于产生多种生物产品,包括例如病毒、病毒样颗粒和/或疫苗。然而,用于产生生物制剂的供应链,尤其是那些采用病毒载体的供应链,效率非常低。即使是病毒载体或昆虫宿主细胞的细微变化也可以影响生产收率。例如,病毒载体有效载荷和血清型可以影响昆虫细胞代谢、活力、病毒载体组装、病毒载体产生和/或病毒载体表达。为了降低成本和遵守良好生产规范,病毒载体的生物生产(诸如在昆虫细胞培养中)非常重要。昆虫细胞包含哺乳动物系统中存在的大多数翻译后修饰途径,从而可以产生比在酵母细胞或原核细胞中表达的更类似(例如,抗原性、免疫原性和/或功能性)于天然哺乳动物蛋白质的生物产品。本公开认识到病毒载体生产的问题在于昆虫细胞系通常未优化且/或分离优化细胞系的筛选平台也是低效的。由于生物生产的高成本和长生产时间两者以及对更大量的病毒载体的市场需求不断增加,目前病毒载体的制造是不充足的。

[0196] 本公开尤其提供了用于在昆虫细胞中表达病毒载体的平台技术。本公开提供了使用基于文库的方法基于昆虫细胞或细胞系表达病毒载体的能力来筛选、评估和/或表征它

们的方法。所提供的技术涵盖以下见解:使病毒载体获取表达其的昆虫细胞的标识符(例如,包括条形码序列和/或文库变体),从而使对病毒载体基因组(例如,病毒重复序列之间)进行测序以提供用于鉴定具有更高病毒载体表达和/或病毒载体产生的其他特征的昆虫细胞的有效且稳健的方法成为可能。所提供的方法、系统和组合物为工程化具有有益病毒载体特征(例如,增加的病毒载体表达、增加的表达持续时间、增加的稳定性等)的昆虫细胞提供快速、稳健的平台。

[0197] 用于病毒载体表达的昆虫细胞平台

[0198] 本公开提供了用于产生、筛选、选择、工程化和/或鉴定用于表达病毒载体的昆虫细胞或细胞系的平台技术。本公开提供了使用基于文库的方法用标识符指示文库中的昆虫细胞来源的方法。本公开提供了新的平台技术,其中由昆虫细胞表达的病毒载体获取标识符(例如,包括条形码和/或文库变体)。可以将昆虫细胞文库产生的病毒载体合并并分析(例如,通过测序,例如,通过下一代测序)所需特征。病毒载体获取标识符使快速鉴定产生病毒载体的昆虫细胞成为可能,所述昆虫细胞具有所需特征,诸如但不限于增强或改善的病毒载体产生。本公开涵盖以下认识:可以通过对病毒载体的标识符(例如,包含条形码和/或文库变体)进行测序来筛选昆虫细胞文库以得到具有所需的病毒载体产生特征(例如,改善的病毒载体产生)的细胞系。例如,来自昆虫细胞文库表达的病毒载体的标识符池中特定标识符的相对丰度可以与文库中特定昆虫细胞系的相对病毒载体生产率相关(例如,高丰度的特定标识符可以与具有较高病毒载体生产率的昆虫细胞系相关)。

[0199] 图1提供了用于筛选昆虫细胞的病毒载体表达或其他病毒载体特征的示例性技术的示意图。所提供的技术采用昆虫细胞文库,所述昆虫细胞文库被工程化为包含标识符(例如,包含条形码和/或文库变体)和适当的用于将标识符包装到病毒载体中的遗传元件(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)(描绘于图1,步骤A)。在一些实施方案中,标识符在文库构建体的环境中,其可以是单个连续多核苷酸或两个或更多个不连续多核苷酸。在一些实施方案中,文库构建体还包含例如至少一个文库变体。在一些实施方案中,文库变体可以影响病毒载体产生或其他特征。在一些实施方案中,文库变体可以引起扰动,例如可以影响病毒载体产生或其他特征的一个或多个遗传修饰。在一些实施方案中,昆虫细胞被遗传工程化以包含一个或多个工程化序列,所述一个或多个工程化序列包括位于病毒包装序列之间的标志符和任选的以下中的任一者:扰动、文库变体、有效载荷、扰动附加序列、反式作用整合序列和/或顺式作用整合序列。

[0200] 昆虫细胞文库的昆虫细胞还可以包含(例如,已经游离存在和/或整合到昆虫细胞基因组内)或被工程化(例如,转染)以包含足以表达病毒载体的遗传物质(描绘于图1,步骤B)。用遗传物质(例如,一个或多个工程化序列)转染或转导昆虫细胞文库可以用于生成产生病毒载体的昆虫细胞文库。

[0201] 遗传物质可以包含例如用于复制的病毒基因和编码必需病毒蛋白例如衣壳蛋白的病毒基因。对于AAV载体,足以表达病毒载体的遗传物质可以包含(i)编码AAV Rep蛋白和AAV Cap蛋白的基因;和(ii)支持AAV复制所需的腺病毒基因(例如,E2、E4和VARNA)。在一些实施方案中,病毒载体还包含有效载荷(例如,编码有效载荷的核酸)。在一些实施方案中,有效载荷包括具有启动子、ORF和polyA信号的表达盒。在一些实施方案中,昆虫细胞文库的昆虫细胞用足以表达包含有效载荷和标识符的病毒载体的遗传物质转染。

[0202] 在一些实施方案中,所提供的技术包括一种独特的方法,借此病毒载体获取标识符(例如,在病毒载体基因组中,例如,在病毒重复序列之间)。因此,文库中昆虫细胞的每个克隆群都将表达具有独特标识符的病毒载体。这使得能够直接表征病毒载体并鉴定产生它的昆虫细胞。通常,本公开的产生病毒载体的昆虫细胞文库将包含至少一个工程化序列,所述至少一个工程化序列包含位于用于将标识符包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间的标识符和足以表达关注的病毒载体的遗传物质。产生病毒载体的昆虫细胞文库的昆虫细胞还可以包含一个或多个(例如,多至100个独特的)文库变体和/或扰动。一个或多个扰动可以存在于昆虫细胞和/或由昆虫细胞表达的病毒载体中。

[0203] 任何工程化序列(例如,编码病毒载体、标识符和/或其他工程化序列)可以游离存在(例如,在一个或多个质粒上)和/或被整合到昆虫细胞的基因组中。例如,在一些实施方案中,可以将一个或多个产生病毒载体所必需的核酸序列整合到昆虫细胞中。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列可以游离存在于昆虫细胞中。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列可以诱导表达(例如,在诱导型启动子的控制下)。

[0204] 在一些实施方案中,培养产生病毒载体的昆虫细胞文库,并使用本领域已知的任何适当的方法收获病毒载体(描绘于图1,步骤C)。在一些实施方案中,收获由文库的昆虫细胞产生的总病毒载体。在一些实施方案中,对应于时间间隔收获文库的昆虫细胞产生的病毒载体。例如,可以每天、每两天、每3天或以更长的间隔收获病毒载体,以评估一段时间内的病毒载体产生。在一些实施方案中,可以在长时间后洗涤昆虫细胞并收获病毒载体(例如,以评估病毒载体的持续产生)。在一些实施方案中,分离病毒载体(例如,包括标识符)的遗传物质。

[0205] 可以合并病毒载体并对遗传物质进行测序(例如,使用下一代测序)(描绘于图1,步骤D)。病毒载体池中标识符的存在表明昆虫细胞文库中的昆虫细胞或昆虫克隆细胞系(也具有标识符)可以产生病毒载体。在一些实施方案中,确定一个或多个标识符的相对丰度。在一些实施方案中,从病毒载体池确定的标识符的相对丰度与对应细胞系产生的相对病毒载体相关联。在一些实施方案中,标识符的相对丰度可以用于确定文库中不同昆虫细胞产生的特定病毒载体的数量(例如,相对数量)。在一些实施方案中,病毒载体池中标识符的丰度可以用于确定各种工程化序列(例如,文库变体和/或扰动)对文库中昆虫细胞的病毒载体产生能力的影响。

[0206] 本公开认识到病毒载体的量或丰度可以反映偏差(例如,特定昆虫细胞或昆虫克隆细胞系增加的细胞分裂或细胞数量),并且不一定是与昆虫细胞或克隆细胞系相关的病毒载体产生水平。因此,在一些实施方案中,病毒载体的量或丰度可以被归一化。在一些实施方案中,检测到的病毒载体的量或丰度可以例如针对细胞量和/或细胞DNA的数量进行归一化。在一些实施方案中,细胞文库在文库中的每个细胞样品中包含大约相同数量的细胞。

[0207] 可以鉴定与所需病毒载体特征(例如,高表达)相关的昆虫细胞,并确定它们的工程化序列(例如,通过基因组测序)。例如,可以选择与参考细胞相比在生产规范下(例如,在现行良好生产规范(cGMP)下)表现出增加的病毒载体产生的昆虫细胞。与参考细胞相比,增加的病毒载体产生可以是固定时间段内病毒载体数量的增加或延长的时间量的产生。在一

些实施方案中,可以选择相对于参考细胞产生病毒载体时间更长(例如,由于增加的活力、增加的基因组稳定性和/或增加的病毒载体产生的持续时间)的昆虫细胞。参考细胞可以包括不包含工程化序列的可比较的昆虫细胞和/或标准细胞(例如,能够产生病毒载体)。

[0208] 选定的昆虫细胞候选者通常可以用于产生病毒载体且/或对应的文库变体且/或所鉴定的扰动可以用于为新昆虫细胞文库的构建提供信息(描绘于图1,图F)。可以重复基于文库的平台方法,直到鉴定出表达具有所需特征和/或所需数量的病毒载体的工程化昆虫细胞。可以分析与所需特征相关的工程化序列,例如,使用机器学习(ML)方法来开发机器学习模型。经过训练的机器学习模型可以用于为未来的设计提供信息,并减少要筛选的昆虫细胞文库的数量和/或每个昆虫细胞文库的大小,从而减少时间和成本。在一些实施方案中,可以设计昆虫细胞文库且/或可以执行所述方法以鉴定在昆虫细胞中协同相互作用(例如,两个或更多个工程化序列组合)以具有所需特征(例如,一定水平的病毒载体产生)的工程化序列。在一些实施方案中,从本文所述的平台技术获得的所得昆虫细胞将具有一个、两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个工程化序列(例如,一个、两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个扰动)。在一些实施方案中,所提供的方法可以用于产生、筛选、选择、工程化和/或鉴定具有所需病毒载体产生特性(例如,一定水平的产生、所需持续时间的产生等)的昆虫细胞。在一些实施方案中,所提供的方法可以用于产生、筛选、选择、工程化和/或鉴定影响病毒载体产生的一个或多个扰动或扰动的组合。在一些实施方案中,选定的昆虫细胞候选者可以用于在额外修饰(例如,去除标识符)后产生病毒载体。在一些实施方案中,所鉴定的对应文库变体和/或扰动可以用于为类似昆虫细胞系或细胞系的构建提供信息(例如,扰动可以被引入相同的祖先细胞系、相似的祖先细胞系(其中一个或多个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸被整合到昆虫细胞基因组中)和/或非常不同的祖先细胞系中)。

[0209] 在一些实施方案中,昆虫细胞被工程化以包含病毒载体的一种或多种组分,所述一种或多种组分处于诱导型转录控制元件(例如,启动子)的控制之下。在一些实施方案中,对包含病毒载体的遗传元件的昆虫细胞进行操作以进一步包含文库构建体。在一些实施方案中,昆虫细胞包含多核苷酸,所述多核苷酸具有一个或多个在诱导型转录控制元件的控制下产生病毒载体所必需的元件(例如,游离和/或整合到昆虫基因组中),并且被进一步工程化以包含文库构建体。

[0210] 本领域技术人员将认识到,虽然图1的示意图描绘了对昆虫细胞文库进行工程化以包含一个或多个工程序列和一个或多个标识符,然后引入病毒载体的遗传元件,但是这些用于对表达病毒载体的昆虫细胞文库进行工程化的步骤可以按照任何顺序并使用本领域已知的任何方法进行。例如,对表达病毒载体的昆虫细胞文库进行工程化可以按照任何顺序并使用本领域已知的任何方法进行。例如,在一些实施方案中,同时或基本上同时引入工程化序列、标识符和病毒载体的元件。在一些实施方案中,生成昆虫细胞文库,其包含独特的工程化序列和形成病毒载体所必需的多核苷酸,并且随后引入标识符。在一些实施方案中,生成昆虫细胞文库,其包含独特的包含标识符的工程化序列,并且随后引入形成病毒载体所必需的多核苷酸。

[0211] 所提供的方法可以用于生成、鉴定和/或选择具有产生任何所需病毒载体的所需特征的昆虫细胞。通常,此种病毒载体对于其所需目的将是具有功能性的。例如,在一些实

施方案中,用于基因治疗的病毒载体将具有递送有效载荷(例如,核酸,例如,到靶细胞)的功能。在一些实施方案中,病毒载体是能够杀伤癌细胞的溶瘤病毒载体。在一些实施方案中,病毒载体有治疗活性。然而,在一些实施方案中,所提供的方法可以产生缺乏一种或多种功能性特征但保留其他关注的特征的非功能性病毒载体。在一些实施方案中,病毒载体是非功能性的或对于特定特征具有降低的功能性。例如,在一些实施方案中,病毒载体可能具有降低的转移有效载荷的能力或可能不能转移有效载荷。在一些实施方案中,病毒载体可能具有降低的杀伤癌细胞的能力。在一些实施方案中,病毒载体可能没有治疗活性。

[0212] 病毒载体

[0213] 在一些实施方案中,本公开的昆虫细胞系可以用于产生病毒载体。在一些实施方案中,所产生的病毒载体本身用作生物制剂和/或治疗。在一些实施方案中,所产生的病毒载体用于生成许多生物制剂和/或治疗的研究、生产和/或制造过程。例如,病毒载体可以以多种方式使用,包括但不限于疫苗、癌症治疗(例如,溶瘤治疗)和/或基因治疗(例如,体内基因和/或基因组编辑)。作为另一个实例,病毒载体可以以多种方式使用,包括但不限于研究、生产和/或制造:疫苗、癌症治疗(例如,溶瘤治疗)、基因治疗(例如,离体基因和/或基因组编辑)和/或细胞治疗(例如,离体基因和/或基因组编辑)。因此,有大量病毒载体可用于这些不同的应用。

[0214] 本公开的方法可以用于选择和/或鉴定具有表达病毒载体的有益特征的昆虫细胞,例如,本领域已知或本文描述的任何病毒载体;本公开不限于任何特定的病毒载体。

[0215] 病毒载体可以具有不同的大小、承载能力、具有不同的基因组结构(例如,DNA和/或RNA,以及单链和/或双链)。在一些实施方案中,病毒载体可以用于将核酸递送至细胞用于瞬时表达或长期表达。在一些实施方案中,病毒载体具有范围广泛的宿主细胞。在一些实施方案中,病毒载体可以具有有限和/或特定类型的宿主细胞。

[0216] 所提供的昆虫细胞和/或昆虫细胞群包含一个或多个多核苷酸,所述一个或多个多核苷酸包含一个或多个产生病毒载体所必需的核酸序列。在一些实施方案中,包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸游离存在于昆虫细胞中。在一些实施方案中,包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸存在于昆虫细胞基因组中。在一些实施方案中,一个或多个产生病毒载体所必需的核酸序列处于诱导型转录控制元件的控制之下。例如,在一些特定实施方案中,一个或多个产生病毒载体所必需的核酸序列可以整合到昆虫细胞基因组中并处于诱导型转录控制元件(例如,诱导型启动子和/或诱导型增强子)的控制之下。可以衍生出病毒载体的病毒包括但不限于腺相关病毒(AAV)、腺病毒、杆状病毒(例如,苜蓿银纹夜蛾多核型多角体病毒(AcMNPV))、慢病毒、α病毒(例如,辛德毕斯病毒)、逆转录病毒(例如,γ逆转录病毒)、多瘤病毒(例如,猴病毒40(SV40))、乳头瘤病毒(例如,牛乳头瘤病毒(BPV))、痘病毒(例如,痘苗病毒)、单纯疱疹病毒(HSV)、麻疹病毒、棒状病毒(例如,狂犬病病毒)、水疱性口炎病毒(VSV)、小核糖核酸病毒(例如,脊髓灰质炎病毒)、呼肠孤病毒、塞内卡病毒、艾柯病毒(例如,RIGVIR)、塞姆利基森林病毒(SFV)、黄病毒、指环病毒(https://www.ringtx.com)、新城疫病毒(NDV)、副粘病毒(例如,仙台病毒)、仙台病毒载体、正粘病毒(例如,流感病毒)、冠状病毒和杂交体和/或工程化病毒和/或病毒载体。

[0217] 在一些特定实施方案中,本公开上下文中的病毒载体源自腺相关病毒、腺病毒、杆

状病毒、慢病毒、逆转录病毒和/或单纯疱疹病毒。

[0218] 在一些实施方案中,本公开的昆虫细胞系可以用于产生病毒载体,诸如腺相关病毒载体、腺病毒载体、杆状病毒载体、慢病毒载体、逆转录病毒载体和/或单纯疱疹病毒载体。

[0219] 在一些实施方案中,本公开的产生的病毒载体适用于使用现行良好生产规范(cGMP)生成许多生物制剂和/或治疗的生产和/或制造过程。在一些实施方案中,病毒载体适用于生物制剂产品的工业规模制造。在一些实施方案中,病毒载体适用于符合当地监管标准(例如,FDA和/或EMA监管标准)的制造方法。

[0220] 病毒载体可以是活的和减毒的。在一些实施方案中,病毒载体可以是复制条件性的。在一些实施方案中,病毒载体可以是复制缺陷型的。在一些实施方案中,病毒载体可以是无复制能力的。在一些实施方案中,病毒载体是复制缺陷的。在一些实施方案中,病毒载体具有复制能力。在一些实施方案中,病毒载体是非致病性的。

[0221] 在一些实施方案中,本公开提供了一种独特的方法,病毒载体借此从表达它的昆虫细胞中获取标识符并将其并入其病毒核酸(例如,病毒基因组或构建体,例如病毒重复序列之间)。因此,由本公开的昆虫细胞和/或技术产生的病毒载体将在它们的病毒核酸(例如,病毒基因组或构建体,例如,在病毒重复序列之间)中具有标识符。在一些实施方案中,由本公开的昆虫细胞和/或方法产生的病毒载体将包含衣壳和工程化序列诸如核酸,其中核酸包含(i)有效载荷,(ii)标识符,和(iii)病毒基因组的一个或多个序列(例如,对于AAV可以在AAV ITR序列之间)。在一些实施方案中,随后将标识符和/或有效载荷或其一部分从病毒载体中去除。在一些实施方案中,随后将标识符和/或有效载荷或其一部分从病毒载体中去除,其中有效载荷或其一部分被另一有效载荷或其一部分替换。

[0222] 腺相关病毒(AAV)载体

[0223] 在一些实施方案中,由本公开的方法和昆虫细胞产生的病毒载体是腺相关病毒(AAV)载体。AAV是用于基因递送的常用病毒载体。在一些实施方案中,AAV载体具有低免疫原性(例如,在人中)。在一些实施方案中,AAV载体与范围广泛的宿主细胞相容。在一些实施方案中,AAV载体可以转导分裂细胞和静止细胞。在一些实施方案中,本公开的昆虫细胞产生如本文所述的AAV载体。

[0224] 在一些实施方案中,本公开提供了编码一个或多个产生AAV载体所必需的元件的核酸序列。AAV载体的基本元件可以包括Rep蛋白和/或衣壳(Cap)蛋白(例如,形成AAV衣壳的VP1、VP2和VP3)。在一些实施方案中,AAV载体的基本元件可以编码在一个或多个构建体上(例如,其可以被整合到或游离存在于昆虫细胞内)。在一些实施方案中,编码一个或多个产生AAV载体所必需的元件的核酸被整合到昆虫细胞的基因组中。在一些实施方案中,编码一个或多个产生AAV载体所必需的元件的核酸游离存在于昆虫细胞中。

[0225] 在一些实施方案中,本公开提供了包含衣壳和包含有效载荷的核酸的AAV载体。在一些实施方案中,AAV载体具有包含线性单链DNA核酸的二十面体蛋白衣壳。

[0226] 在一些实施方案中,由本公开的昆虫细胞和/或方法产生的病毒载体将包含AAV衣壳和核酸,其中核酸包含(i)有效载荷,(ii)标识符(例如,包含条形码和/或文库变体),和(iii)两个ITR序列(例如,源自AAV)。

[0227] 在一些实施方案中,AAV载体源自人AAV1;AAV2;AAV3b;AAV4;AAV5;AAV6;AAV7;

AAV8;AAV9;AAV10;AAV11;AAV12;AAV13,或其任何衍生物。在一些实施方案中,AAV载体是合成的和/或杂交的人AAV载体。在一些实施方案中,AAV载体源自牛AAV(b-AAV);犬AAV(CAAV);小鼠AAV1;山羊AAV;大鼠AAV;或禽AAV(AAAV)。

[0228] 在一些实施方案中,AAV载体可以被描述为具有血清型,其是对衣壳株和某些核酸序列(例如,ITR)的株的描述。例如,在一些实施方案中,AAV载体可以被描述为AAV2,其中载体具有AAV2衣壳和包含特征性AAV2反向末端重复序列(ITR)的核酸。在一些实施方案中,AAV载体可以被描述为假型,其中衣壳和ITR源自不同的AAV株,例如,AAV2/9将指代包含利用AAV2 ITR和AAV9衣壳的构建体的AAV载体。

[0229] 在一些实施方案中,AAV载体没有血清型和/或假型。在一些实施方案中,AAV载体包含工程化AAV衣壳和/或ITR(例如,与已知AAV血清型的衣壳和/或ITR没有显著同源性)。

[0230] 在一些实施方案中,本公开的AAV载体包含AAV衣壳。在一些实施方案中,AAV衣壳来自或源自AAV2、3、4、5、6、7、8、9、10、rh8、rh10、rh39、rh43或Anc80血清型的AAV衣壳,或其一种或多种杂交体。在一些实施方案中,AAV衣壳来自AAV祖先血清型。在一些实施方案中,AAV衣壳是祖先(Anc)AAV衣壳。Anc衣壳是从构建体序列创建的,所述构建体序列使用进化概率和进化模型构建以确定可能的祖先序列。因此,不知道自然界中是否存在Anc衣壳/构建体序列。在一些实施方案中,AAV衣壳是人工工程化序列(例如,与已知的AAV血清型衣壳没有显著同源性)。

[0231] 如本文所提供的,本公开的AAV载体可以包含AAV衣壳和AAV核酸的任何组合(例如,包含有效载荷和/或AAV ITR)。例如,野生型或变体AAV衣壳包裹AAV核酸,所述AAV核酸包含标识符和/或侧翼为AAV衍生的ITR的有效载荷。

[0232] 通常,AAV核酸由单链脱氧核糖核酸(ssDNA)组成。在一些实施方案中,AAV核酸包含一种或多种源自或修饰自天然存在的AAV基因组的组分。在一些实施方案中,AAV核酸包含源自或修饰自AAV的反向末端重复序列(ITR)。在一些实施方案中,AAV核酸包含有效载荷序列和两个ITR。在一些实施方案中,AAV载体包含衣壳和ssDNA,所述ssDNA包含有效载荷序列和两个病毒重复序列,例如ITR序列,在DNA链的每一端(5'和3')各一个。

[0233] 在一些实施方案中,所提供的AAV核酸包含有效载荷,所述有效载荷包括编码序列和一个或多个调控和/或控制序列,以及任选的5'和3'AAV衍生的反向末端重复序列(ITR)。在一些实施方案中,所提供的AAV核酸被包装到AAV衣壳中以形成AAV载体。

[0234] 在某些实施方案中,病毒载体包含有包含标识符和/或有效载荷序列的核酸以及侧翼为5'或"左"AAV ITR序列和3'或"右"AAV ITR序列的相关调控元件。5'和左的名称是指ITR序列相对于整个构建体的位置,在有义方向上从左到右读取。本领域普通技术人员将理解如何修饰给定的ITR序列以用作5'/左ITR或3'/右ITR,或其反义版本。

[0235] 在一些实施方案中,本文所述的AAV载体的AAV核酸通常包含顺式作用5'ITR序列和3'ITR序列(参见,例如,B.J.Carter,在"Handbook of Parvoviruses,"中,P.Tijsser编,CRC Press,第155-168页,1990,其以引用的方式整体并入本文)。在一些实施方案中,至少80%的典型ITR序列(例如,至少85%、至少90%或至少95%)并入本文提供的构建体中。修改这些ITR序列的能力属于本领域技术人员的能力。(参见,例如,texts such as Sambrook等人,"Molecular Cloning.A Laboratory Manual",第2版,Cold Spring Harbor Laboratory,New York,1989;和K.Fisher等人,J Virol.70:520-532,1996,所述文献中的

每一者均通过引用并入本文)。在一些实施方案中,标识符和/或有效载荷序列的侧翼为5'和3'AAV ITR序列。在一些实施方案中,AAV核酸包含标识符和侧翼为5'和3'AAV ITR序列的有效载荷。AAV ITR序列可以获得自任何已知的AAV,包括目前鉴定的AAV类型。

[0236] 在一些实施方案中,AAV载体核酸包含有效载荷、标识符和两个AAV ITR。在一些实施方案中,AAV载体包含衣壳和dsDNA,所述dsDNA包含(i)有效载荷和/或标识符,和(ii)两个AAV ITR序列,在DNA链的每一端(5'和3')各一个。

[0237] 通常,ITR能够形成发夹。形成发夹的能力可以有助于ITR的自启动(self-prime)能力,允许不依赖引物酶的第二DNA链的合成。ITR还可以帮助将AAV构建体有效封装在AAV载体中。AAV ITR序列可以获得自任何已知的AAV,包括昆虫AAV类型。

[0238] 在一些实施方案中,ITR包括天然存在的ITR序列的一种或多种修饰,例如截短、缺失、取代或插入。修改这些ITR序列的能力属于本领域技术人员的能力。(参见,例如,texts such as Sambrook等人"Molecular Cloning.A Laboratory Manual",第2版,Cold Spring Harbor Laboratory,New York(1989);和K.Fisher等人,J Virol.,70:520-532(1996),每一篇都以引用的方式整体并入本文)。例如,AAV2衍生的ITR序列的长度为约145个核苷酸。在一些实施方案中,ITR包含少于145个核苷酸,例如127、130、134或141个核苷酸。例如,在一些实施方案中,ITR包含110、111、112、113、114、115、116、117、118、119、120、121、122、123、124、125、126、127、128、129、130、131、132、133、134、135、136、137、138、139、140、141、142、143、144或145个核苷酸。

[0239] 在一些实施方案中,AAV载体有效载荷还包含常规控制元件,所述元件以允许其在转染有构建体或感染有由本公开产生的病毒载体的细胞中转录、翻译和/或表达的方式可操作地连接至编码序列。在一些实施方案中,AAV载体有效载荷任选地包含启动子、增强子、非翻译区(例如,5'UTR、3'UTR)、Kozak序列、内部核糖体进入位点(IRES)、剪接位点(例如,受体位点、供体位点)、聚腺苷酸化位点和/或其任何组合。

[0240] 在一些实施方案中,AAV载体有效载荷小于4kb。在一些实施方案中,AAV载体有效载荷可以包含至少500bp、至少1kb、至少1.5kb、至少2kb、至少2.5kb、至少3kb、至少3.5kb、至少4kb,或至少4.5kb的序列。在一些实施方案中,AAV载体有效载荷可以包含至多7.5kb、至多7kb、至多6.5kb、至多6kb、至多5.5kb、至多5kb、至多4.5kb、至多4kb、至多3.5kb、至多3kb,或至多2.5kb的序列。在一些实施方案中,AAV载体有效载荷可以包含约1kb至约2kb、约1kb至约3kb、约1kb至约4kb、约1kb至约5kb、约2kb至约3kb、约2kb至约4kb、约2kb至约5kb、约3kb至约4kb、约3kb至约5kb,或约4kb至约5kb的序列。

[0241] 在一些实施方案中,AAV载体可以指导有效载荷的长期表达。在一些实施方案中,AAV载体可以指导有效载荷的瞬时表达。

[0242] 在一些实施方案中,由本公开的昆虫细胞和/或方法产生的AAV载体将包含AAV衣壳和核酸,其中核酸包含(i)有效载荷,(ii)标识符(例如,条形码和/或文库变体),和(iii)两个病毒重复序列(例如,源自AAV的ITR)。

[0243] 在一些实施方案中,昆虫细胞文库的昆虫细胞被遗传修饰以包含一个或多个产生AAV载体所必需的核酸序列。在一些实施方案中,昆虫细胞文库的昆虫细胞可以具有所提供的一种或多种AAV载体组分,诸如像rep序列、cap序列和产生AAV载体所需的辅助功能。在一些实施方案中,AAV载体的一种或多种组分(例如,AAV Rep基因、AAV Cap基因、一种或多种

辅助基因或其组合)在诱导型转录控制元件(例如,启动子和/或增强子)的控制之下。在一些实施方案中,昆虫细胞(例如,昆虫细胞群的昆虫细胞)包含:(i)位于两个病毒重复序列之间的标识符,和(ii)一个或多个包含AAV Rep基因、AAV Cap基因、一种或多种辅助基因和/或其组合的多核苷酸,其中昆虫细胞产生包含标识符的AAV载体。在一些实施方案中,AAV载体包含有效载荷。在一些实施方案中,随后将标识符和/或有效载荷或其一部分从AAV载体中去除。在将标识符和/或有效载荷或其一部分从AAV载体中去除的一些实施方案中,有效载荷或其一部分被不同的有效载荷或其一部分替换。

[0244] 腺病毒载体

[0245] 在一些实施方案中,由本公开的方法和昆虫细胞产生的病毒载体是腺病毒载体。腺病毒是无包膜病毒,由于它们诱导的强烈免疫原性反应,它们通常用作疫苗。一些腺病毒因其优先感染和杀伤癌细胞的能力而被用于癌症治疗。

[0246] 在一些实施方案中,腺病毒载体源自人腺病毒。人腺病毒涵盖至少51种血清型的家族,所述血清型分类为几个亚组。例如,A亚组包括腺病毒血清型12、18和31。B亚组包括腺病毒血清型3、7、11a、11p、14、16、21、34、35和50。C亚组包括腺病毒血清型1、2、5和6。D亚组包括腺病毒血清型8、9、10、13、15、17、19、19p、20、22-30、32、33、36-39、42-49和51。E亚组包括腺病毒血清型4。F亚组包括腺病毒血清型40和41。本公开的腺病毒载体可以属于任何腺病毒组、亚组和/或血清型。

[0247] 在一些实施方案中,腺病毒载体源自任何血清型,诸如血清型1至血清型51(例如1、2、4、5...51)。例如,在一些实施方案中,腺病毒是2型腺病毒或5型腺病毒。在一些某些实施方案中,腺病毒载体至少部分源自5型腺病毒。

[0248] 在一些实施方案中,腺病毒载体是复制缺陷的(例如,某些必需的病毒基因被删除和/或被有效载荷序列替换)。本文所述方法中使用的任何腺病毒载体可以包括在E1、E2a、E2b、E3或E4编码区中的任何一个或多个中的缺失。在一些实施方案中,腺病毒载体是复制缺陷的并且缺少E4基因座(例如,E4编码区被删除)。在一些实施方案中,复制缺陷腺病毒载体可用作疫苗,用于癌症治疗和/或基因治疗。

[0249] 通常,腺病毒的特征在于高转导效率和直接高转基因表达。在一些实施方案中,腺病毒载体可以指导有效载荷的瞬时表达。在一些实施方案中,腺病毒载体指导靶细胞中有效载荷的瞬时表达。

[0250] 在一些实施方案中,腺病毒载体具有复制能力。例如,具有复制能力的腺病毒载体(例如,溶瘤载体)可以被工程化为优先在癌细胞中复制并通过裂解病毒复制的自然过程破坏癌细胞。

[0251] 在一些实施方案中,本公开提供了包含衣壳和包含有效载荷的核酸的腺病毒载体。在一些实施方案中,腺病毒载体具有包含线性双链体核酸的二十面体蛋白衣壳。在一些实施方案中,腺病毒载体的直径为约90-100nm。在一些实施方案中,腺病毒载体具有包含线性双链DNA的二十面体蛋白衣壳。

[0252] 通常,腺病毒载体核酸由双链DNA(dsDNA)组成,并具有一种或多种源自或修饰自天然存在的腺病毒基因组的组分。在一些实施方案中,腺病毒载体核酸包含源自或修饰自任何血清型的腺病毒(例如,5型腺病毒)的反向末端重复(ITR)序列。在一些实施方案中,腺病毒载体核酸包含有效载荷和两个病毒重复序列,诸如ITR。在一些实施方案中,腺病毒载

体核酸包含标识符和两个病毒重复序列,诸如ITR。在一些实施方案中,腺病毒载体核酸包含有效载荷、标识符和两个病毒重复序列,诸如ITR。在一些实施方案中,腺病毒载体包含衣壳和dsDNA,所述dsDNA包含(i)有效载荷和/或标识符,和(ii)两个ITR序列,在DNA链的每一端(5'和3')各一个。

[0253] 在一些实施方案中,腺病毒ITR具有长度为约40bp至约200bp的序列。例如,人5型腺病毒的ITR为103bp。在一些实施方案中,腺病毒ITR具有约30bp、约40bp、约50bp、约60bp、约70bp、约80bp、约90bp、约100bp、约110bp、约120bp、约130bp、约140bp、约150bp、约160bp、约170bp、约180bp、约190bp、约200bp、约220bp、约240bp、约260bp、约280bp、约300bp、约325bp、约350bp、约375bp,或约400bp的长度。在一些特定实施方案中,腺病毒ITR序列为约50至约250bp。

[0254] 在一些实施方案中,所提供的腺病毒载体包含腺病毒衍生的衣壳和核酸,所述核酸包含有效载荷和一个或多个腺病毒衍生的序列(诸如像ITR)。在一些实施方案中,所提供的腺病毒载体包含腺病毒衍生的衣壳和核酸,所述核酸包含标识符、有效载荷和至少两个病毒重复序列。在一些实施方案中,腺病毒载体包含有包含条形码和/或文库变体的标识符。在一些实施方案中,腺病毒载体有效载荷包括编码序列和一个或多个调控和/或控制序列。在一些实施方案中,腺病毒载体包含5'和3'腺病毒衍生的反向末端重复序列(ITR)。在一些特定的实施方案中,腺病毒载体包含衣壳和工程化腺病毒基因组,其中工程化基因组包含某些序列的缺失(例如,E1缺失和/或E3缺失)和有效载荷和/或标识符的插入。

[0255] 在一些实施方案中,腺病毒载体有效载荷小于7.5kb。在一些实施方案中,腺病毒载体有效载荷可以包含至少500bp、至少1kb、至少1.5kb、至少2kb、至少2.5kb、至少3kb、至少3.5kb、至少4kb、至少4.5kb、至少5kb、至少5.5kb、至少6kb、至少6.5kb、至少7kb、至少7.5kb的序列。在一些实施方案中,腺病毒载体有效载荷可以包含至多10kb、至多9.5kb、至多9kb、至多8.5kb、至多8kb、至多7.5kb、至多7kb、至多6.5kb、至多6kb、至多5.5kb、至多5kb、至多4.5kb,或至多4kb的序列。在一些实施方案中,腺病毒载体有效载荷可以包含约1kb至约2kb、约1kb至约3kb、约1kb至约4kb、约1kb至约5kb、约1kb至约6kb、约1kb至约7kb、约1kb至约8kb、约1kb至约9kb、约1kb至约10kb、约2kb至约4kb、约2kb至约6kb、约2kb至约8kb、约2kb至约10kb、约4kb至约6kb、约4kb至约8kb,或约5kb至约8kb的序列。

[0256] 在一些实施方案中,由本公开的昆虫细胞和/或方法产生的腺病毒载体将包含腺病毒衣壳和核酸,其中核酸包含(i)有效载荷,(ii)标识符(例如,条形码和/或文库变体),和(iii)两个病毒重复序列(例如,源自腺病毒的ITR)。

[0257] 在一些实施方案中,昆虫细胞文库的昆虫细胞被遗传修饰以包含一个或多个产生腺病毒载体所必需的核酸序列。在一些实施方案中,昆虫细胞文库的昆虫细胞可以具有所提供的一种或多种腺病毒载体组分,诸如像rep序列、cap序列和产生腺病毒载体所需的辅助功能。在一些实施方案中,腺病毒载体的一种或多种组分(例如,腺病毒Rep基因、腺病毒Cap基因、一种或多种辅助基因或其组合)在诱导型转录控制元件(例如,启动子和/或增强子)的控制之下。在一些实施方案中,昆虫细胞(例如,昆虫细胞群的昆虫细胞)包含:(i)位于两个病毒重复序列之间的标识符,和(ii)一个或多个包含腺病毒Rep基因、腺病毒Cap基因、一种或多种辅助基因和/或其组合的多核苷酸,其中昆虫细胞产生包含标识符的腺病毒载体。在一些实施方案中,腺病毒载体包含有效载荷。在一些实施方案中,随后将标识符和/

或有效载荷或其一部分从腺病毒载体中去除。在将标识符和/或有效载荷或其一部分从腺病毒载体中去除的一些实施方案中,有效载荷或其一部分被不同的有效载荷或其一部分替换。

[0258] 逆转录病毒和慢病毒载体

[0259] 在一些实施方案中,由本公开的方法和昆虫细胞产生的病毒载体是逆转录病毒载体。逆转录病毒是包膜病毒,其在宿主细胞中通过使用病毒逆转录酶将其RNA转录为DNA来复制。逆转录病毒DNA作为宿主基因组的一部分进行复制并且被称为前病毒。逆转录病毒载体可以包括但不限于基于或源自鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、亲嗜性逆转录病毒、猴免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)及其组合的那些。

[0260] 在一些实施方案中,由本公开的方法和昆虫细胞产生的病毒载体是慢病毒载体。慢病毒载体是多功能工具,因为它们具有转导非分裂细胞的能力。在一些实施方案中,慢病毒载体能够感染分裂细胞和非分裂细胞。通常,慢病毒载体使长期和/或稳定的基因表达成为可能,并被整合到宿主细胞基因组中。

[0261] 慢病毒是直径约80至约120nm的包膜颗粒。在一些实施方案中,本公开提供了具有在约50nm至约200nm范围内的直径的慢病毒载体。在一些实施方案中,本公开提供了具有在约80nm至约120nm范围内的直径的慢病毒载体。

[0262] 慢病毒可以含有几种结构蛋白,包括基质、衣壳、核衣壳、包膜和逆转录酶蛋白。在一些实施方案中,本公开提供了包含衣壳、包膜的慢病毒载体。在一些实施方案中,本公开提供了包含衣壳、包膜和核酸的慢病毒载体。在一些实施方案中,本公开提供了包含衣壳、包膜和包含有效载荷和/或标识符的核酸的慢病毒载体。

[0263] 在一些实施方案中,由本公开的昆虫细胞和/或方法产生的慢病毒载体将包含慢病毒衣壳(或其衍生物)、包膜和核酸,其中核酸包含(i)有效载荷,(ii)标识符,和(iii)两个长末端重复(LTR)序列(例如,源自慢病毒)。在一些实施方案中,两个LTR序列能够将核酸包装到慢病毒载体中。在一些实施方案中,慢病毒载体包含慢病毒Psi序列(或其衍生物或工程化变体)。

[0264] 在一些实施方案中,慢病毒载体包括基于人免疫缺陷病毒(HIV-1、HIV-2)、猫免疫缺陷病毒(FIV)、马传染性贫血病毒、猴免疫缺陷病毒(SIV)和梅迪-维斯纳病毒(MVV)的那些。在一些实施方案中,慢病毒核酸(即,工程化基因组)和包膜糖蛋白将基于不同的病毒,使得所得病毒载体是假型的。

[0265] 在一些实施方案中,慢病毒载体源自HIV-1。在一些实施方案中,慢病毒载体源自HIV-1并包含衣壳蛋白和核酸,所述核酸包含(i)有效载荷,(ii)标识符,和(iii)两个长末端重复(LTR)序列(例如,源自慢病毒)。在一些实施方案中,慢病毒载体是HIV载体并且其中昆虫细胞包含两个包含HIV LTR序列的病毒重复序列。在一些实施方案中,慢病毒载体是SIV载体并且其中昆虫细胞包含两个包含SIV LTR序列的病毒重复序列。在一些实施方案中,慢病毒载体是马传染性贫血病毒载体,并且其中昆虫细胞包含两个包含马传染性贫血病毒LTR序列的病毒重复序列。在一些实施方案中,慢病毒载体是FIV载体并且其中昆虫细胞包含两个包含FIV LTR序列的病毒重复序列。在一些实施方案中,慢病毒载体是维斯纳病毒载体并且其中昆虫细胞包含两个包含维斯纳病毒LTR序列的病毒重复序列。

[0266] 在一些实施方案中,慢病毒载体包含gag蛋白或其片段。在一些实施方案中,慢病

毒载体包含有包含一个或多个选自基质(MA)、衣壳(CA)和核衣壳(NC)结构域的结构域的gag蛋白。在一些实施方案中,慢病毒载体包含包膜蛋白或其片段。在一些实施方案中,慢病毒载体是假型慢病毒载体,其中gag蛋白和包膜蛋白源自不同的病毒。在一些实施方案中,慢病毒载体包含源自人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒载体、维斯纳病毒载体或其衍生物的gag蛋白和/或env蛋白。

[0267] 在一些实施方案中,慢病毒载体包含(i)慢病毒gag基因,(ii)慢病毒env基因,(iii)慢病毒pol基因,或(iv)其组合。在一些实施方案中,表达慢病毒载体的昆虫细胞包含以下中的一种或多种:(i)慢病毒gag基因,(ii)慢病毒env基因,和(iii)慢病毒pol基因。

[0268] 在一些实施方案中,安全特征被并入慢病毒载体中,其可以包括例如自失活长末端重复序列(LTR)和整合缺陷。在某些实施方案中,整合缺陷可能由载体基因组的元件赋予,但也可能源自包装系统的元件(例如,可能不是载体基因组的一部分但以反式提供的非功能性整合酶蛋白)。

[0269] 慢病毒具有单链RNA(ssRNA)基因组。在一些实施方案中,慢病毒载体包含为ssRNA的核酸,并且包含源自慢病毒诸如HIV-1和/或SIV的有效载荷和序列。在一些实施方案中,有效载荷的侧翼为长末端重复(LTR)序列。在一些实施方案中,慢病毒载体包含为ssRNA的核酸,并且包含标识符和病毒重复序列(例如HIV-1和/或SIV LTR)。在一些实施方案中,慢病毒载体包含为ssRNA的核酸,并且包含有效载荷、标识符和病毒重复序列(例如HIV-1和/或SIV LTR)。

[0270] 在一些实施方案中,慢病毒载体核酸可以包含来自慢病毒的5'LTR和3'LTR的序列,并且特别地可以包含来自慢病毒的5'LTR的R序列和U5序列和来自慢病毒的失活或自失活的3'LTR。LTR序列可以是来自任何物种的任何慢病毒的LTR序列。例如,它们可以是来自HIV、SIV、FIV或BIV的LTR序列。在一些实施方案中,LTR序列是HIV LTR序列。

[0271] 在一些实施方案中,慢病毒载体有效载荷小于8kb。在一些实施方案中,慢病毒载体有效载荷可以包含至少500bp、至少1kb、至少1.5kb、至少2kb、至少2.5kb、至少3kb、至少3.5kb、至少4kb、至少4.5kb、至少5kb、至少5.5kb、至少6kb、至少6.5kb、至少7kb、至少7.5kb,或至少8kb的序列。在一些实施方案中,慢病毒载体有效载荷可以包含至多10kb、至多9.5kb、至多9kb、至多8.5kb、至多8kb、至多7.5kb、至多7kb、至多6.5kb、至多6kb、至多5.5kb、至多5kb、至多4.5kb,或至多4kb的序列。在一些实施方案中,慢病毒载体有效载荷可以包含约1kb至约2kb、约1kb至约3kb、约1kb至约4kb、约1kb至约5kb、约1kb至约6kb、约1kb至约7kb、约1kb至约8kb、约1kb至约9kb、约1kb至约10kb、约2kb至约4kb、约2kb至约6kb、约2kb至约8kb、约2kb至约10kb、约4kb至约6kb、约4kb至约8kb,或约5kb至约8kb的序列。

[0272] 在一些实施方案中,昆虫细胞文库的昆虫细胞被遗传修饰以包含一个或多个产生慢病毒载体所必需的核酸序列。在一些实施方案中,昆虫细胞文库的昆虫细胞可以具有提供的一种或多种慢病毒载体组分。在一些实施方案中,慢病毒载体的一种或多种组分处于诱导型转录控制元件(例如,启动子和/或增强剂)的控制之下。在一些实施方案中,昆虫细胞(例如,昆虫细胞群的昆虫细胞)包含:(i)位于两个病毒重复序列之间的标识符,和(ii)一个或多个产生慢病毒载体所必需的多核苷酸,其中昆虫细胞产生包含标识符的慢病毒载体。在一些实施方案中,慢病毒载体包含有效载荷。在一些实施方案中,随后将标识符和/或有效载荷或其一部分从慢病毒载体中去除。在将标识符和/或有效载荷或其一部分从慢病

毒载体中去除的一些特定实施方案中,有效载荷或其一部分被不同的有效载荷或其一部分替换。

[0273] HSV载体

[0274] 在一些实施方案中,由本公开的方法和昆虫细胞产生的病毒载体是单纯疱疹病毒(HSV)载体。HSV是大的包膜病毒,其具有含有环形dsDNA基因组的二十面体衣壳。HSV体内特征在于神经元的终生潜在感染。这一特征使HSV载体可用于长期转基因表达。在一些实施方案中,HSV载体是有复制能力的减毒载体、无复制能力的重组载体或复制缺陷的辅助依赖型载体。

[0275] 在一些实施方案中,HSV载体具有在约120nm至约200nm之间的范围内的直径。在一些实施方案中,HSV载体是直径为约120至约200nm的包膜颗粒。在一些实施方案中,HSV载体具有在约100nm至约200nm之间的范围内的直径。

[0276] 在一些实施方案中,HSV载体来自或源自单纯疱疹病毒-1(HSV-1)、单纯疱疹病毒-2(HSV-2)、人巨细胞病毒(HCMV)、水痘-带状疱疹病毒(VZV)、爱泼斯坦-巴尔病毒(EBV)、卡波西肉瘤相关疱疹病毒(KSHV)、人疱疹病毒6和/或人疱疹病毒7和/或其衍生物。在一些实施方案中,HSV载体来自或源自HSV-1、HSV-2或其组合(例如,包含来自HSV-1的衣壳并包含源自HSV-2的核酸序列)。

[0277] 在一些实施方案中,病毒载体是HSV-AAV杂交载体。

[0278] 天然HSV-1基因组由两段基因组编码区组成,称为长独特区段和短独特区段(U_L和U_S),每个片段的侧翼均为反向重复序列(分别为TR_L/IR_L和IR_S/TR_S)。在一些实施方案中,HSV载体被工程化以缺少或缺失ICP0、ICP4、ICP22、ICP27和/或ICP47以降低毒性。设计HSV载体的方法和注意事项是本领域已知的,例如Manservigi等人,Open Virol J.2010;4:123-156。

[0279] 在一些实施方案中,本公开提供了包含衣壳蛋白和包含有效载荷的核酸的HSV载体。在一些实施方案中,HSV载体包含有包含VP5、VP19C、VP23、pre-VP22a和/或成熟的蛋白酶(UL26基因产物)的衣壳。

[0280] 在一些实施方案中,HSV核酸还包含获得自或源自HSV病毒的序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)。在一些实施方案中,本公开提供了HSV载体,其包含衣壳蛋白和包含HSV病毒重复序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。在一些实施方案中,本公开提供了包含衣壳、包膜和包含有效载荷的核酸的HSV载体。在一些实施方案中,本公开提供了HSV载体,其包含衣壳、包膜和包含有效载荷和HSV病毒重复序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。

[0281] 在一些实施方案中,HSV载体有效载荷小于100kb。在一些实施方案中,HSV载体有效载荷可以包含至少1kb、至少2kb、至少3kb、至少4kb、至少5kb、至少6kb、至少7kb、至少8kb、至少9kb、至少10kb、至少15kb、至少20kb、至少25kb、至少30kb、至少40kb,或至少50kb的序列。在一些实施方案中,HSV载体有效载荷可以包含至多150kb、至多140kb、至多130kb、至多120kb、至多110kb、至多100kb、至多90kb、至多80kb、至多70kb、至多60kb、至多50kb、至多40kb、至多30kb、至多25kb、至多20kb、至多15kb,或至多10kb的序列。在一些实施方案中,慢病毒载体有效载荷可以包含约1kb至约150kb、约1kb至约100kb、约1kb至约50kb、约1kb至约25kb、约5kb至约100kb、约5kb至约90kb、约5kb至约80kb、约5kb至约70kb、约5kb至约

60kb、约5kb至约50kb、约5kb至约40kb、约5kb至约30kb、约5kb至约25kb、约5kb至约20kb、约5kb至约15kb、约5kb至约10kb、约10kb至约100kb、约10kb至约50kb,或约10kb至约25kb的序列。

[0282] 在一些实施方案中,本公开提供了包含衣壳蛋白和包含标识符的核酸的HSV载体。在一些实施方案中,本公开提供了HSV载体,其包含衣壳蛋白和包含标识符和HSV病毒重复序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。在一些实施方案中,本公开提供了包含衣壳、包膜和包含标识符的核酸的HSV载体。在一些实施方案中,本公开提供了HSV载体,其包含衣壳、包膜和包含标识符和HSV病毒重复序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。在一些实施方案中,HSV载体包含两个病毒重复序列,所述两个病毒重复序列包含末端a序列。

[0283] 在一些实施方案中,HSV核酸包含有效载荷、标识符和获得自或源自HSV病毒的一个或多个序列。在一些实施方案中,通过本公开的方法和/或昆虫细胞产生的HSV载体包含HSV衣壳、包膜和包含有效载荷、标识符和获得自或源自HSV病毒的一个或多个序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。在一些实施方案中,HSV载体包含HSV衣壳、包膜和包含有效载荷和/或标识符、侧翼为HSV病毒重复序列(例如,TR_L/IR_L和/或IR_S/TR_S序列)的核酸。

[0284] 在一些实施方案中,昆虫细胞文库的昆虫细胞被遗传修饰以包含一个或多个产生HSV载体所必需的核酸序列。在一些实施方案中,昆虫细胞文库的昆虫细胞可以具有提供的一种或多种HSV载体组分。在一些实施方案中,HSV载体的一种或多种组分处于诱导型转录控制元件(例如,启动子和/或增强剂)的控制之下。在一些实施方案中,昆虫细胞(例如,昆虫细胞群的昆虫细胞)包含:(i)位于两个病毒重复序列之间的标识符,和(ii)一个或多个产生HSV载体所必需的多核苷酸,其中昆虫细胞产生包含标识符的HSV载体。在一些实施方案中,HSV载体包含有效载荷。在一些实施方案中,随后将标识符和/或有效载荷或其一部分从HSV载体中去除。在将标识符和/或有效载荷或其一部分从HSV载体中去除的一些特定实施方案中,有效载荷或其一部分被不同的有效载荷或其一部分替换。

[0285] 杆状病毒载体

[0286] 杆状病毒是已知对昆虫群影响最显著的病毒。它们是位于杆状衣壳中的双链、环状、超螺旋DNA分子。杆状病毒具有双链环状DNA基因组,长度通常为80kb至200kb之间。

[0287] 传统上,杆状病毒分为两个形态不同的属:核型多角体病毒(NPV)和颗粒体病毒。现代分类通常将杆状病毒分为四个属:α杆状病毒(鳞翅目特有的NPV)、β杆状病毒(鳞翅目特有的颗粒体病毒)、γ杆状病毒(膜翅目特有的NPV)和Δ杆状病毒(双翅目特有的NPV)。在一些实施方案中,杆状病毒是苜蓿银纹夜蛾多核型多角体病毒(AcMNPV)或其衍生物。

[0288] 在一些实施方案中,杆状病毒表达载体包含有效载荷。在一些实施方案中,将有效载荷插入杆状病毒基因组的非必需基因组基因座中,例如像,egt基因座和/或多角体蛋白基因座。

[0289] 在一些实施方案中,使用杆状病毒载体将病毒载体的一个或多个组分转移到昆虫细胞中。例如,在一些实施方案中,杆状病毒载体可以在其基因组中包含来自AAV载体的一个或多个序列。在一些实施方案中,杆状病毒载体包含AAV ITR序列和标识符。在一些实施方案中,包含AAV ITR序列和标识符(例如,位于ITR序列之间)的杆状病毒载体还可以包含有效载荷和/或一个或多个文库变体。

[0290] 文库构建体

[0291] 在本公开的上下文中,本公开提供了用于工程化昆虫细胞的文库构建体。在一些实施方案中,文库构建体可以包含两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个,或十个或更多个工程化序列。在一些实施方案中,文库构建体可以包含多至100个工程化序列。

[0292] 文库构建体至少包含标识符和适合于将标识符包装到病毒载体中的遗传结构。如本文所用的文库构建体还将包含任何文库变体。在一些实施方案中,文库构建体还包含用于包装到病毒载体中的有效载荷(例如,在病毒重复序列之间)。例如,在一些实施方案中,文库构建体包含(i)标识符和有效载荷,其位于病毒重复序列之间,(ii)至少一个包含至少一个文库变体的工程化序列。

[0293] 在一些实施方案中,文库构建体还包含一个或多个构建体,其包含顺式作用整合序列(例如,同源臂、识别位点和/或病毒重复序列)。例如,在一些实施方案中,文库构建体包含(i)标识符和有效载荷,其位于病毒重复序列之间,(ii)至少一个包含至少一个文库变体的工程化序列,和(iii)用于将文库构建体或其一部分整合到昆虫细胞基因组中的顺式作用整合序列。

[0294] 在一些实施方案中,文库构建体还包含至少一个条形码。例如,在一些实施方案中,文库构建体包含(i)标识符和有效载荷,其位于病毒重复序列之间,(ii)至少一个包含至少一个文库变体的工程化序列,和(iii)至少一个条形码。作为另一个实例,在一些实施方案中,文库构建体包含(i)标识符和有效载荷,其位于病毒重复序列之间,(ii)至少一个包含至少一个文库变体的工程化序列,(iii)至少一个条形码,和(iv)用于将文库构建体或其一部分整合到昆虫基因组中的顺式作用整合序列。

[0295] 在一些实施方案中,文库构建体还包含至少一个工程化序列,所述至少一个工程化序列包含至少一个报告基因和/或选择标记。在一些实施方案中,一个或多个包含文库构建体的多核苷酸包含报告基因和/或选择标记。任何合适的报告基因(例如,GFP、RFP、YFP、lacZ等)或选择标记(例如,其赋予可以人工选择的性状,例如,抗性盒等)都可以用在本公开的上下文中。

[0296] 文库构建体邻接

[0297] 文库构建体可以是单个连续构建体或多个不连续构建体。在一些实施方案中,文库构建体是单个(即一个)连续构建体。在具有单个连续文库构建体的此类实施方案中,所得病毒载体的表征将直接提供关于任何文库变体的信息(例如,病毒载体标识符的确定可以直接与任何文库变体相关联)。在一些实施方案中,文库构建体包含多个不连续构建体。在文库构建体是不连续文库构建体的此类实施方案中,所提供的方法还将包括鉴定昆虫细胞中的文库变体的步骤(例如,通过单细胞测序)。

[0298] 在一些实施方案中,昆虫细胞群包含多个文库构建体,其中每个单独的文库构建体由单个连续核酸序列组成,并且其中多个文库构建体包含多个独特核酸序列。在一些实施方案中,昆虫细胞群包含多个文库构建体,其中每个单独的文库构建体由不连续核酸序列组成,并且其中多个文库构建体包含多个独特核酸序列。

[0299] 在一些实施方案中,所提供的昆虫细胞包含有包含多个多核苷酸的文库构建体,其中每个单独的昆虫细胞确切包含构成文库构建体的多个多核苷酸的第一子集中的一个

独特的多核苷酸和构成文库构建体的多个多核苷酸的第二子集中的两个或更多个独特的多核苷酸。例如,在一些实施方案中,每个单独的昆虫细胞确切包含一个独特的标识符和两个或更多个独特的文库变体。在一些实施方案中,所提供的昆虫细胞包含有包含多个多核苷酸的文库构建体,其中每个单独的昆虫细胞确切包含构成文库构建体的多个多核苷酸的第一子集中的两个独特的多核苷酸和构成文库构建体的多个多核苷酸的第二子集中的多个独特的多核苷酸。例如,在一些实施方案中,每个单独的昆虫细胞确切包含两个独特的标识符和多个独特的文库变体。

[0300] 连续文库构建体

[0301] 在一些实施方案中,文库构建体是单个连续构建体,其包含至少一个标识符,其侧翼为适合于将标识符包装到病毒载体中的遗传结构(例如,病毒重复序列,例如,AAV ITR序列),以及任何文库变体。在一些实施方案中,单个连续文库构建体包含标识符和一个或多个文库变体,其中标识符和文库变体都位于病毒重复序列之间。在一些实施方案中,单个连续文库构建体包含位于病毒重复序列之间的标识符和位于病毒重复序列之外的一个或多个文库变体。在一些实施方案中,单个连续文库构建体包含位于病毒重复序列之间的标识符和一个或多个文库变体,以及位于病毒重复序列之外的一个或多个其他文库变体。在一些实施方案中,文库构建体是连续文库构建体,并且包含报告基因和/或选择标记。下表1中提供了单个连续文库构建体的一些示例性配置。

[0302] 表1-作为单个连续构建体的示例性文库构建体

[0303]

标识符	文库变体 (例如，gRNA、ORF、基因(昆虫和/或病毒)、非编码核酸或其一部分、多个或组合)	扰动	有效载荷
条形码	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
条形码	位于病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
条形码	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - gRNA	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - gRNA	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - ORF	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - ORF	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - 基因或其一部分	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - 基因或其一部分	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
文库变体 - 非编码核酸	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间

[0304]

文库变体 - 非编码核酸	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
条形码和文库变体	位于病毒重复序列之间的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间
条形码和文库变体	位于病毒重复序列之间和病毒重复序列之外的文库变体	KO、SNP、激活、抑制、插入	如果有的话，在病毒重复序列之间

[0305] 在细胞中的位置和拷贝数

[0306] 在一些实施方案中,文库构建体是游离的和/或被整合到昆虫细胞基因组中。在一些实施方案中,文库构建体是游离的单个连续构建体。在一些实施方案中,文库构建体是整合到昆虫细胞基因组中的单个连续构建体。

[0307] 在一些实施方案中,文库构建体还包含用于整合到昆虫细胞基因组中的顺式作用整合序列。例如,表1的示例性实施方案中的任一个还可以包含顺式作用整合序列(例如,同源臂、识别位点和/或病毒重复序列)。

[0308] 在一些实施方案中,单个连续文库构建体被整合到昆虫细胞基因组中。在一些实施方案中,单个连续文库构建体包含顺式作用整合序列。在一些实施方案中,单个连续文库构建体包含位于文库构建体的3'和5'端的顺式作用整合序列。在一些实施方案中,单个连续文库构建体以低拷贝数(例如,10个或更少拷贝的文库构建体)被整合到昆虫细胞基因组中。在一些实施方案中,四个或更少拷贝的单个连续文库构建体被整合到昆虫细胞基因组中。在一些实施方案中,三个或更少拷贝的单个连续文库构建体被整合到昆虫细胞基因组中。在一些实施方案中,两个或更少拷贝的单个连续文库构建体被整合到昆虫细胞基因组中。在一些实施方案中,单个拷贝的单个连续文库构建体被整合到昆虫细胞基因组中。

[0309] 在一些实施方案中,单个连续文库构建体游离存在于昆虫细胞中。在一些实施方案中,单个连续文库构建体以低拷贝数(例如,10个或更少拷贝的文库构建体,例如,4个或更少拷贝的文库构建体,例如,3个或更少拷贝的文库构建体、2个或更少拷贝的文库构建体,例如,单个(一个)拷贝的文库构建体)游离存在于昆虫细胞中。

[0310] 不连续文库构建体

[0311] 在一些实施方案中,提供的方法和细胞包括不连续文库构建体,其使例如同时筛选多个文库变体成为可能。在一些实施方案中,文库构建体包含多个不连续构建体,其中至少一个构建体包含标识符和适合于将标识符包装到病毒载体中的遗传结构。在一些实施方案中,不连续文库构建体包含第一构建体,所述第一构建体包含位于病毒重复序列之间的标识符和一个或多个其他构建体。在一些实施方案中,不连续文库构建体包含第一构建体,所述第一构建体包含位于病毒重复序列之间的标识符和有效载荷和一个或多个包含一个或多个文库变体的其他构建体。在一些实施方案中,不连续文库构建体包含第一构建体,所述第一构建体包含例如表1中描述的任何单个连续文库构建体和一个或多个其他构建体

(例如,包含其他文库变体)。

[0312] 在一些实施方案中,其他不连续文库构建体包含病毒载体的组分,例如病毒Cap基因和条形码。在一些实施方案中,文库包括具有工程化或不同血清型的病毒载体组分(例如,Cap基因)的文库构建体。例如,可以同时筛选具有不同血清型的Cap基因的病毒Cap基因文库(例如,以选择具有针对多种不同血清型的病毒载体的改进特征的昆虫细胞)。

[0313] 在一些特定的实施方案中,一个或多个其他构建体包含文库变体,所述文库变体是变体病毒载体组分(例如,工程化的或不同血清型的)。在一些实施方案中,文库构建体包含表1中描述的文库构建体和一个或多个其他构建体,其中至少一个其他构建体包含病毒Cap基因和条形码。

[0314] 在一些实施方案中,筛选了多个不同的文库变体文库,每个文库与不同的标识符相关。例如,可以同时筛选gRNA文库变体文库和Cap基因变体文库;这将使选择富含不同Cap基因文库变体的gRNA变体成为可能(例如,以选择那些对不同病毒载体血清型具有有益影响的gRNA介导的扰动)。

[0315] 在一些实施方案中,文库构建体还包含一个、两个、三个、四个、五个或更多个构建体,每个构建体包含一个或多个文库变体。在一些实施方案中,一个或多个包含一个或多个文库变体的其他构建体和/或包含标识符的构建体还包含一个或多个条形码。

[0316] 在一些实施方案中,文库构建体是不连续文库构建体,其包含两个、三个、四个、五个、六个、七个、八个、九个或十个不连续核酸序列(例如,单独的构建体)。在一些实施方案中,文库构建体是不连续文库构建体,其包含多至20个不连续核酸序列、多至30个不连续核酸序列、多至40个不连续核酸序列、多至50个不连续核酸序列、多至60个不连续核酸序列、多至70个不连续核酸序列、多至80个不连续核酸序列、多至90个不连续核酸序列或多至100个不连续核酸序列。

[0317] 在一些实施方案中,文库构建体是不连续的并且文库构建体的一个或多个单独的多核苷酸包含报告基因和/或选择标记。在一些实施方案中,文库构建体是不连续的并且文库构建体的多个单独的多核苷酸包含报告基因和/或选择标记。在一些实施方案中,文库构建体是不连续的并且文库构建体的每一个单独的多核苷酸包含报告基因和/或选择标记。

[0318] 表2-示例性不连续文库构建体

[0319]

第一构建体	其他构建体
位于病毒重复序列(用于包装到病毒载体中)之间的条形码	一个或多个其他构建体各自包含至少一个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的条形码	一个或多个其他构建体各自包含两个或更多个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的条形码和有效载荷	一个或多个其他构建体各自包含至少一个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的条形码和有效载荷	一个或多个其他构建体各自包含两个或更多个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的文库变体(作为标识符)和有效载荷	一个或多个其他构建体各自包含至少一个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的文库变体(作为标识符)和有效载荷	一个或多个其他构建体各自包含两个或更多个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的条形码、有效载荷和文库变体	一个或多个其他构建体各自包含至少一个文库变体，并且任选地每个文库变体与条形码相关。
位于病毒重复序列(用于包装到病毒载体中)之间的条形码、有效载荷和文库变体	一个或多个其他构建体各自包含两个或更多个文库变体，并且任选地每个文库变体与条形码相关。

[0320] 在细胞中的位置和拷贝数

[0321] 在一些实施方案中,文库构建体是不连续的并且一个或多个单独的构建体是游离的。在一些实施方案中,文库构建体是不连续的并且一个或多个单独的构建体被整合到昆虫细胞基因组中。在一些实施方案中,文库构建体是不连续的并且至少一个构建体是游离的并且至少一个构建体被整合到昆虫细胞基因组中。

[0322] 在一些实施方案中,不连续文库构建体的一个或多个单独的构建体被整合到昆虫细胞基因组中。在一些实施方案中,不连续文库构建体的一个或多个单独的构建体包含顺式作用整合序列。在一些实施方案中,不连续文库构建体的一个或多个单独的构建体包含位于每个构建体的3'和5'端的顺式作用整合序列。在一些实施方案中,顺式作用整合序列包含病毒重复序列(例如,位于用于包装到病毒载体中的任何病毒重复序列之外)。

[0323] 在文库构建体包含多个不连续构建体的一些实施方案中,一个或多个单独的构建体游离存在于昆虫细胞中。在一些实施方案中,一个或多个单独的构建体以低拷贝数(例如,10个或更少拷贝的文库构建体,例如,4个或更少拷贝的文库构建体,例如,3个或更少拷贝的文库构建体、2个或更少拷贝的文库构建体,例如,单个(一个)拷贝的文库构建体)游离存在于昆虫细胞中。

[0324] 标识符

[0325] 本公开提供了可以被检测并使鉴定产生和/或衍生病毒载体的昆虫细胞或克隆细胞系成为可能的标识符。在一些实施方案中,特定标识符的相对丰度可以在病毒载体(例如,每个包含标识符)的合并样品中被表征、检测和/或量化。在一些实施方案中,检测标识符的至少一部分。在一些实施方案中,标识符通过测序(例如,通过下一代测序)检测。例如,在一些实施方案中,标识符的至少一部分通过下一代测序和/或单细胞测序和/或Sanger测序检测。

[0326] 在一些实施方案中,标识符存在于昆虫细胞中并且还存在于昆虫细胞表达的病毒载体中。在一些实施方案中,本公开提供了存在于文库构建体的上下文中的标识符。在一些实施方案中,标识符位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如用于表达AAV载体的AAV ITR)之间。在一些实施方案中,如本文所述的昆虫细胞表达的病毒载体包含标识符,例如,在病毒载体的核酸中。在一些实施方案中,如本文所述的昆虫细胞表达的病毒载体包含核酸,所述核酸包含位于两个病毒重复序列(例如,AAV载体的ITR)之间的标识符。在一些实施方案中,病毒载体核酸内的标识符的至少一部分通过下一代测序和/或单细胞测序和/或Sanger测序检测。

[0327] 在一些实施方案中,表达病毒载体的昆虫细胞和表达的病毒载体均包含相同的标识符。在一些实施方案中,病毒载体的标识符对应于产生病毒载体的昆虫细胞的标识符。在一些实施方案中,病毒载体的标识符源自产生病毒载体的昆虫细胞的标识符。在一些实施方案中,病毒载体的标识符与产生病毒载体的昆虫细胞的标识符至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多相同。

[0328] 在一些实施方案中,标识符包含条形码。在一些实施方案中,标识符是或包含条形码,并且存在于位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如用于表达AAV载体的AAV ITR)之间的文库构建体中。在一些实施方案中,标识符是或包含存在于病毒载体核酸中的条形码。在一些实施方案中,所提供的方法包括检测包含来自病毒载体样品的条形码的标识符。

[0329] 在一些实施方案中,标识符包含文库变体。在一些实施方案中,标识符包含文库变体,并且存在于位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如用于表达AAV载体的AAV ITR)之间的文库构建体中。在一些实施方案中,标识符包含存在于病毒载体核酸中的文库变体。在一些实施方案中,所提供的方法包括检测包含来自病毒载体样品的文库变体的标识符。

[0330] 在一些实施方案中,标识符包含条形码和文库变体。在一些实施方案中,标识符包含条形码和文库变体,并且存在于位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如用于表达AAV载体的AAV ITR)之间的文库构建体中。在一些实施方案中,标识符包含存在于病毒载体核酸中的条形码和文库变体。在一些实施方案中,所提供的方法包括检测包含来自病毒载体样品的条形码和文库变体的标识符的至少一部分。

[0331] 条形码

[0332] 本公开涵盖以下认识:条形码可以用于跟踪不同组分之间的关联。在一些实施方案中,条形码是一种工程化序列。在一些实施方案中,条形码是一种工程化核酸序列。在一

些实施方案中,条形码是文库构建体的一部分。

[0333] 在一些实施方案中,文库构建体包含一个或多个条形码,所述条形码在检测时(例如,通过下一代测序方法)指示一个或多个文库变体和/或未直接检测的其他工程化序列的身份。在一些实施方案中,一个条形码与一个或多个工程化序列相关。例如,在一些实施方案中,一个条形码与一个或多个文库变体相关。在一些实施方案中,一个条形码与一个工程化序列相关。例如,在一些实施方案中,一个条形码与一个文库变体相关。

[0334] 在一些实施方案中,条形码不组成标识符。在一些实施方案中,条形码组成标识符。在一些实施方案中,条形码组成包含核酸序列的标识符。在一些实施方案中,条形码组成包含工程化序列的标识符。在一些实施方案中,标识符不包含条形码。在一些实施方案中,标识符包含条形码。在一些实施方案中,标识符包含有包含核酸序列的条形码。在一些实施方案中,标识符包含有包含工程化序列的条形码。在一些实施方案中,昆虫细胞包含多个条形码,其中至少一个条形码是标识符并且至少一个条形码不是标识符。

[0335] 在一些实施方案中,文库构建体包含条形码。在一些实施方案中,条形码用作标识符。在一些实施方案中,条形码用作标识符,并且位于两个病毒重复序列之间。在一些实施方案中,条形码不用作标识符。在一些实施方案中,条形码不用作标识符,并且位于两个病毒重复序列之间,但未被检测到(例如,通过下一代测序)。在一些实施方案中,文库构建体包含位于两个病毒重复序列之外的条形码序列,因此不被包装到病毒载体中。

[0336] 在一些实施方案中,条形码用作标识符,其中条形码存在于位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间的文库构建体中。在一些实施方案中,条形码用作使鉴定产生和/或衍生病毒载体的昆虫细胞或克隆细胞系成为可能的标识符。在一些实施方案中,条形码的相对丰度指示它所源自的昆虫细胞的相对生产率。在一些实施方案中,条形码用作标识符,并且在检测(例如,通过下一代测序方法)时也指示一个或多个文库变体和/或未直接检测的工程化序列的身份。

[0337] 在一些实施方案中,条形码不用作标识符,但在检测(例如,通过下一代测序方法和/或单细胞测序方法)时指示一个或多个文库变体和/或未直接检测的工程化序列的身份。例如,在文库构建体不连续的一些实施方案中,条形码可以用于追踪一个或多个文库变体。在一些实施方案中,不连续文库构建体包含第一构建体和一个或多个各自包含条形码的其他构建体,所述第一构建体包含位于病毒重复序列之间的标识符。在一些实施方案中,文库构建体还包含一个、两个、三个、四个、五个或更多个构建体,每个构建体包含一个或多个文库变体,其中每个单独的构建体还包含条形码。在一些实施方案中,每个文库变体均与独特的条形码相关。作为另一个实例,在文库构建体连续的一些实施方案中,条形码也可以用于追踪一个或多个文库变体。在一些实施方案中,连续文库构建体包含位于病毒重复序列之间的标识符。在一些实施方案中,文库构建体还包含一个或多个文库变体和一个或多个条形码。在一些实施方案中,每个文库变体均与独特的条形码相关。

[0338] 在一些实施方案中,条形码包含长度在3个核苷酸至50个核苷酸范围内的核酸序列。在一些实施方案中,条形码包含长度在由下限和上限(上限大于下限)界定的范围内的核酸序列。在一些实施方案中,下限可以是约5个核苷酸、约6个核苷酸、约7个核苷酸、约8个核苷酸、约9个核苷酸、约10个核苷酸、约11个核苷酸、约12个核苷酸、约13个核苷酸、约14个核苷酸,或约15个核苷酸。在一些实施方案中,上限可以是约10个核苷酸、约15个核苷酸、约

20个核苷酸、约25个核苷酸、约30个核苷酸、约35个核苷酸、约40个核苷酸、约45个核苷酸，或约50个核苷酸。在一些特定实施方案中，条形码包含长度在5个核苷酸至25个核苷酸范围内的序列。在一些特定实施方案中，条形码包含约5个核苷酸、约10个核苷酸、约15个核苷酸、约20个核苷酸或约25个核苷酸。在一些实施方案中，条形码包含DNA和/或RNA。在一些实施方案中，条形码包含长度在3个核苷酸至50个核苷酸或约5个核苷酸至约25个核苷酸范围内的DNA序列。在一些实施方案中，条形码包含长度在3个核苷酸至50个核苷酸或约5个核苷酸至约25个核苷酸范围内的RNA序列。

[0339] 在一些实施方案中，所提供的方法包括检测条形码(例如，通过测序，例如通过下一代测序和/或单细胞测序和/或Sanger测序)。

[0340] 文库变体

[0341] 本公开提供了包含一个或多个文库变体的昆虫细胞，所述一个或多个文库变体引起在文库的昆虫细胞之间变化的扰动。在一些实施方案中，文库变体引起可能影响病毒载体产生的某些特征的扰动。在一些实施方案中，文库变体包含引起扰动的工程化序列。在一些实施方案中，文库变体是序列变化。在一些实施方案中，文库变体是表观遗传变化。在一些实施方案中，文库变体在效应物中，由此文库变体影响或引起在细胞之间变化的扰动。在一些实施方案中，文库变体本身可以成为在细胞间变化的扰动。例如，在一些实施方案中，是gRNA的文库变体是效应物，其连同RNA引导的核酸酶(例如，扰动附加序列)导致细胞基因组DNA内的缺失。在其他实施方案中，文库变体是ORF或基因序列，在其转染到细胞中以及在一些情况下整合到基因组DNA中(例如，如通过反式作用和顺式作用整合序列进行)时，本身成为细胞的遗传物质的扰动或修饰。

[0342] 在一些实施方案中，文库变体包含引导RNA序列。在一些实施方案中，包含引导RNA序列的文库变体也可以是标识符(例如，使病毒载体与产生它的昆虫细胞相关的独特gRNA序列)。在一些实施方案中，昆虫细胞包含有包含引导序列的文库变体，其可以被病毒载体获取。在一些实施方案中，引导序列的长度为约10至30个核苷酸。在一些实施方案中，引导序列的长度为15至25个核苷酸。在某些实施方案中，引导序列的长度为16至24个核苷酸(例如，长度为16、17、18、19、20、21、22、23或24个核苷酸)。在一些实施方案中，引导序列位于或接近gRNA的5'末端(例如，对于Cas9或从其衍生或获得的核酸酶)。在一些实施方案中，引导序列位于或接近gRNA的3'末端(例如，对于Cpf1或从其衍生或获得的核酸酶)。

[0343] 在一些实施方案中，文库变体编码与引入基因组缺失相关的引导RNA序列。在一些实施方案中，文库变体编码与引入基因组突变(例如，SNP)相关的引导RNA序列。在一些实施方案中，文库变体编码与引入基因组重排相关的引导RNA序列。在一些实施方案中，文库变体编码与改变基因的表达(例如，激活和/或抑制)相关的引导RNA序列。

[0344] 在一些实施方案中，文库变体包含一个或多个ORF。在一些实施方案中，文库变体包含ORF。在一些实施方案中，文库变体包含编码RNA序列的ORF。在一些实施方案中，ORF编码多肽(诸如蛋白质，诸如糖蛋白)。在一些实施方案中，ORF编码融合多肽和/或嵌合多肽。

[0345] 在一些实施方案中，文库变体包含一个或多个基因。在一些实施方案中，文库变体包含基因。在一些实施方案中，文库变体包含编码RNA序列的基因。在一些实施方案中，基因编码多肽(诸如蛋白质，诸如糖蛋白)。在一些实施方案中，基因编码融合多肽和/或嵌合多肽。在一些实施方案中，文库变体包含昆虫基因。在一些实施方案中，文库变体包含病毒基

因(例如,Cap基因)。

[0346] 在一些实施方案中,文库变体编码非编码核酸序列。在一些实施方案中,文库变体编码调控RNA序列(例如,siRNA、微小RNA等)。

[0347] 在一些实施方案中,文库变体或其一部分也是标识符,但文库构建体可以包含一个或多个不是标识符的其他文库变体。在一些实施方案中,文库变体或其一部分是标识符,其中文库变体位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间。在一些实施方案中,文库变体不是标识符(例如,未被包装到病毒载体中)。在文库变体或其一部分是标识符的一些实施方案中,文库变体或其一部分对于此特定昆虫细胞或细胞系(以及由其表达的病毒载体)将是独特的。

[0348] 在一些实施方案中,文库变体包含病毒载体的组分。例如病毒载体的工程化组分和/或不同血清型的成分。以这种方式可以筛选不同工程化或血清型的病毒载体。在一些实施方案中,文库变体包含Cap基因(例如,不同血清型的Cap基因)。

[0349] 在一些实施方案中,所提供的昆虫细胞单独的包含至少一个文库变体,其中所述至少一个文库变体包含至少一个工程化序列,所述至少一个工程化序列包含至少一个基因、至少一个ORF、至少一个gRNA序列、至少一个独特非编码核酸,或其组合和/或多个。在一些实施方案中,昆虫细胞或昆虫细胞群包含多个文库变体,其中所述多个文库变体包含至少一个工程化序列,所述至少一个工程化序列包含:至少一个独特基因、至少一个独特ORF、至少一个独特gRNA序列,和/或至少一个独特非编码核酸或其组合和/或多个。

[0350] 在一些实施方案中,所提供的昆虫细胞包含一个或多个文库变体。在一些实施方案中,所提供的昆虫细胞包含两个、三个、四个、五个、六个、七个、八个、九个或十个文库变体。在一些实施方案中,所提供的昆虫细胞包含至少100个文库变体。在一些实施方案中,所提供的昆虫细胞包含约2至约100个文库变体、约2至约20个文库变体、约3至约30个文库变体、约4至约40个文库变体、约5至约50个文库变体。在一些实施方案中,所提供的昆虫细胞包含不超过10个文库变体、不超过20个文库变体、不超过30个文库变体、不超过40个文库变体、不超过50个文库变体、不超过60个文库变体、不超过70个文库变体、不超过80个文库变体、不超过90个文库变体,或不超过100个文库变体。

[0351] 在一些实施方案中,文库构建体包含至少一个文库变体和至少一个标识符,其中所述至少一个文库变体和至少一个标识符都位于两个病毒重复序列之间。在一些实施方案中,文库构建体包含至少一个文库变体和至少一个标识符,其中所述至少一个标识符位于两个病毒重复序列之间,并且所述至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,文库构建体包含至少两个文库变体和一个标识符,其中标识符和至少一个文库变体位于两个病毒重复序列之间,并且至少一个文库变体位于两个病毒重复序列之外。在一些实施方案中,文库构建体还包含位于两个病毒重复序列之间和/或之外的一个或多个其他工程化序列。例如,在一些实施方案中,文库构建体还包含位于两个病毒重复序列之间的有效载荷。作为另一个实例,在一些实施方案中,文库构建体还包含位于两个病毒重复序列之间和/或之外的一个或多个其他条形码。

[0352] 在一些实施方案中,文库构建体包含为文库变体的标识符和一个或多个其他文库变体。在一些实施方案中,文库构建体是单个连续文库构建体,其包含标识符,所述标识符是位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV

ITR)之间的文库变体,和位于用于包装到病毒载体中的序列之外的一个或多个其他文库变体。在一些实施方案中,文库构建体是不连续文库构建体,其包含标识符,所述标识符是位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间的文库变体,并且文库构建体是一个或多个包含一个或多个其他文库变体的其他构建体。在一些特定实施方案中,一个或多个其他文库变体包含病毒载体的组分(例如,Cap基因)。

[0353] 在一些实施方案中,文库构建体包含为条形码的标识符和一个或多个文库变体。在一些实施方案中,文库构建体是单个连续文库构建体,其包含标识符,所述标识符是位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间的条形码,并且单个连续文库构建体包含位于用于包装到病毒载体中的序列之外的一个或多个文库变体。在一些实施方案中,文库构建体是不连续文库构建体,其包含标识符,所述标识符是位于用于包装到病毒载体中的序列(例如,病毒重复序列,例如,用于表达AAV载体的AAV ITR)之间的条形码,并且文库构建体是一个或多个包含一个或多个文库变体的其他构建体。

[0354] 下表3中提供了文库构建体中的文库变体的一些示例性配置。

[0355] 表3-文库构建体中的示例性文库变体

[0356]

型式	标识符	文库变体	扰动
单个连续文库构建体	病毒重复序列之间的文库变体	gRNA，整合到细胞基因组中	基因组KO
单个连续文库构建体	病毒重复序列之间的条形码	gRNA，整合到细胞基因组中	基因组SNP
单个连续文库构建体	病毒重复序列之间的文库变体	gRNA，整合到细胞基因组中	基因组激活
单个连续文库构建体	病毒重复序列之间的条形码	gRNA，整合到细胞基因组中	基因组抑制
单个连续文库构建体	病毒重复序列之间的文库变体	ORF，整合到细胞基因组中	基因组插入
单个连续文库构建体	病毒重复序列之间的文库变体	gRNA，游离	基因组KO
单个连续文库构建体	病毒重复序列之间的条形码	gRNA，游离	基因组SNP
单个连续文库构建体	病毒重复序列之间的文库变体	gRNA，游离	基因组激活
单个连续文库构建体	病毒重复序列之间的条形码	gRNA，游离	基因组抑制
单个连续文库构建体	病毒重复序列之间的文库变体	ORF，游离	游离新基因或增加的拷贝数
不连续文库构建体	病毒重复序列之间的文库变体	gRNA，整合到细胞基因组中	基因组KO
不连续文库构建体	病毒重复序列之间的条形码	gRNA，整合到细胞基因组中	基因组SNP
不连续文库构建体	病毒重复序列之间的文库变体	gRNA，整合到细胞基因组中	基因组激活
不连续文库构建体	病毒重复序列之间的条形码	gRNA，整合到细胞基因组中	基因组抑制

[0357]

不连续文库构建体	病毒重复序列之间的文库变体	ORF，整合到细胞基因组中	基因组插入
不连续文库构建体	病毒重复序列之间的条形码	gRNA，游离	基因组KO
不连续文库构建体	病毒重复序列之间的文库变体	gRNA，游离	基因组SNP
不连续文库构建体	病毒重复序列之间的条形码	gRNA，游离	基因组激活
不连续文库构建体	病毒重复序列之间的文库变体	gRNA，游离	基因组抑制
不连续文库构建体	病毒重复序列之间的条形码	ORF，游离	游离新基因或增加的拷贝数

[0358] 在一些实施方案中,所提供的方法包括检测一个或多个文库变体(例如,通过测序,例如,通过下一代测序)。

[0359] 病毒重复序列

[0360] 本公开提供了适合于将标识符包装到病毒载体中的结构,诸如像病毒重复序列。在一些实施方案中,病毒重复序列是DNA和/或RNA序列。在一些实施方案中,病毒重复序列是DNA序列。在一些实施方案中,病毒重复序列是RNA序列。在一些实施方案中,位于两个病毒重复序列之间的任何核酸序列将被包装到病毒载体中。

[0361] 在一些实施方案中,病毒重复序列源自与靶病毒载体相同类型的病毒。例如,在一些实施方案中,靶病毒载体是AAV载体并且病毒重复序列源自AAV。在一些实施方案中,病毒重复序列源自与靶病毒载体相同的病毒株。例如,在一些实施方案中,靶病毒载体是AAV5载体并且病毒重复序列是源自AAV5的ITR。在一些实施方案中,病毒重复序列源自与靶病毒载体不同的病毒株,但仍能够包装到靶病毒载体中。在一些实施方案中,病毒重复序列是工程化病毒重复序列(例如,包括源自两种或更多种病毒的序列)。在一些实施方案中,病毒重复序列是合成病毒重复序列(例如,基于共有病毒重复序列设计)。

[0362] 在一些实施方案中,靶病毒载体是AAV载体并且病毒重复序列包含AAV ITR序列或其衍生物。在一些实施方案中,病毒重复序列源自与AAV载体不同的AAV株,但仍能够被靶病毒载体获取。在一些实施方案中,病毒重复序列是工程化AAV ITR序列。在一些实施方案中,病毒重复序列是合成AAV ITR序列。

[0363] 在一些实施方案中,靶病毒载体是腺病毒载体并且病毒重复序列包含腺病毒ITR序列或其衍生物。在一些实施方案中,病毒重复序列源自与腺病毒载体不同的腺病毒株,但仍能够被靶病毒载体获取。在一些实施方案中,病毒重复序列是工程化腺病毒ITR序列。在一些实施方案中,病毒重复序列是合成腺病毒ITR序列。

[0364] 在一些实施方案中,靶病毒载体是慢病毒载体并且病毒重复序列包含慢病毒ITR序列或其衍生物。在一些实施方案中,靶病毒载体是HIV-1载体并且病毒重复序列包含HIV-1ITR序列或其衍生物。在一些实施方案中,病毒重复序列源自与靶慢病毒载体不同的慢病毒株,但仍能够被靶慢病毒载体获取。在一些实施方案中,病毒重复序列是工程化腺病毒ITR序列。在一些实施方案中,病毒重复序列是合成腺病毒ITR序列。

[0365] 在一些实施方案中,靶病毒载体是HSV载体并且病毒重复序列包含HSV TR_L/IR_L和/或HSV IR_S/TR_S序列和/或其衍生物。在一些实施方案中,靶病毒载体是HSV-1载体和/或HSV-2载体,并且病毒重复序列包含HSV TR_L/IR_L和/或HSV IR_S/TR_S序列和/或其衍生物。在一些实施方案中,病毒重复序列源自与靶HSV载体不同的HSV株,但仍能够被靶HSV载体获取。在一些实施方案中,病毒重复序列是工程化HSV TR_L/IR_L和/或HSV IR_S/TR_S序列。在一些实施方案中,病毒重复序列是合成HSV TR_L/IR_L和/或HSV IR_S/TR_S序列。

[0366] 有效载荷

[0367] 本公开尤其提供了包含有效载荷的病毒载体。有效载荷通常是期望引入细胞、器官、生物体和/或生物系统(例如,包含细胞)中的任何关注的序列。例如,在一些实施方案中,病毒载体可以包含可以用于编辑细胞(例如,编码基因组编辑工具,例如,用于基因治疗和/或细胞治疗)的有效载荷。在一些实施方案中,有效载荷包含在文库构建体中。在一些实施方案中,有效载荷包含在文库构建体中并且位于用于包装到病毒载体中的序列(例如,病

毒重复序列,例如,用于表达AAV载体的AAV ITR)之间。

[0368] 在一些实施方案中,有效载荷序列包含以下中的一项或多项:编码区、基因调控元件和转录终止子。基因调控元件的非限制性实例包括启动子、转录激活子、增强子和聚腺苷酸化信号。在一些实施方案中,有效载荷序列包含编码区、基因调控元件和转录终止子,它们相对于彼此定位,使得编码区位于基因调控元件和转录终止子之间。

[0369] 在一些实施方案中,编码区编码基因产物。在一些实施方案中,基因产物是RNA。在一些实施方案中,编码区编码多肽(诸如蛋白质,诸如糖蛋白)。在一些实施方案中,编码区编码融合多肽和/或嵌合多肽。在一些实施方案中,编码区编码一种基因产物。在一些实施方案中,编码区编码多于一种基因产物(例如,2、3、4、5、6、7或更多种基因产物)。在一些实施方案中,有效载荷序列包含调控核酸,诸如tRNA、rRNA、tmRNA、miRNA、siRNA、piRNA、shRNA、反义RNA、双链RNA、snRNA、snoRNA和/或长非编码RNA(IncRNA)。在一些实施方案中,编码区编码调控RNA(例如,siRNA、微小RNA等)。

[0370] 在一些实施方案中,本文所述的病毒载体的有效载荷可以是基因治疗有效载荷并且可以编码对受试者(诸如患有疾病或病症的受试者)有益的任何蛋白质或其一部分。蛋白质可以是细胞外蛋白、细胞内蛋白或膜结合蛋白。在一些实施方案中,有效载荷序列编码的蛋白质包括但不限于哺乳动物蛋白质,例如人蛋白质。

[0371] 在一些实施方案中,编码的蛋白质是或包括糖蛋白。在一些实施方案中,编码的蛋白质是或包括抗体和/或Fc融合蛋白。在一些实施方案中,编码的蛋白质是或包括抗原。

[0372] 蛋白质可以是治疗性蛋白质。在一些实施方案中,被施用基因治疗的受试者患有疾病或病症,由此受试者的内源性版本的蛋白质有缺陷或产生的量有限或根本不产生。在一些此类实施方案中,有效载荷编码无缺陷版本的蛋白质。在一些实施方案中,被施用基因治疗的受试者患有由靶基因(例如,由靶基因的表达水平和/或靶多肽的活性水平)介导的疾病或病症,并且有效载荷编码靶基因或靶多肽的抑制剂。治疗性蛋白质的实例包括但不限于可输注或可注射的治疗性蛋白质、酶、酶辅因子、激素、血液或血液凝固因子、细胞因子和干扰素、生长因子、脂肪素等。

[0373] 在一些实施方案中,有效载荷可以包含基因编辑组分。在一些实施方案中,有效载荷可以包含切除核酸(例如,其中病毒载体将RNA引导的核酸酶递送至靶细胞)。在一些实施方案中,包含基因编辑组分的有效载荷可以编码本领域已知的任何合适的核酸内切酶。例如,有效载荷可以编码或包含CRISPR(成簇规律间隔短回文重复序列)/Cas系统的一个或多个组分。

[0374] 有效载荷序列可以是与相关病毒载体相容的任何长度。在一些实施方案中,有效载荷序列的侧翼为一个或多个获得自或源自病毒的序列(例如,AAV的ITR序列)。在一些实施方案中,有效载荷序列位于用于包装到病毒载体中的序列之间。在一些实施方案中,有效载荷序列位于病毒重复序列(例如,AAV的ITR序列)之间。

[0375] 在一些实施方案中,有效载荷包含启动子。在一些情况下,有效载荷可以包含增强子序列。

[0376] 在一些实施方案中,本文所述的任何有效载荷可以包含非翻译区(UTR),诸如5'UTR或3'UTR。基因的UTR被转录但不被翻译。5'UTR从转录起始位点开始,并持续到起始密码子,但不包含起始密码子。3'UTR紧随终止密码子开始,并持续直至转录终止信号。

[0377] 在一些实施方案中,编码本文提供的蛋白质的有效载荷可以包含聚腺苷酸化(poly(A))信号序列。大多数新生真核mRNA在其3'端具有poly(A)尾,这是在一个复杂的过程中添加的,所述过程包括初级转录物的切割和由poly(A)信号序列驱动的偶联聚腺苷酸化反应(参见,例如,Proudfoot等人,Cell 108:501-512,2002,其以引用的方式整体并入本文)。poly(A)尾赋予mRNA稳定性和可转移性(B.Alberts等人的Molecular Biology of the Cell,第三版,Garland Publishing,1994,其以引用的方式整体并入本文)。在一些实施方案中,poly(A)信号序列位于编码序列的3'。

[0378] 在一些实施方案中,编码蛋白质的有效载荷可以包含内部核糖体进入位点(IRES)。IRES形成复杂的二级结构,允许从任何位置发生翻译起始,mRNA紧接在IRES所在位置的下游(参见,例如,Pelletier和Sonenberg,Mal.Cell.Biol.8(3):1103-1112,1988)。本领域技术人员已知有几种IRES序列,包括来自例如口蹄疫病毒(FMDV)、脑心肌炎病毒(EMCV)、人鼻病毒(HRV)、蟋蟀麻痹病毒、人免疫缺陷病毒(HIV)、甲型肝炎病毒(HAV)、丙型肝炎病毒(HCV)和脊髓灰质炎病毒(PV)的那些。

[0379] 在一些实施方案中,本文提供的任何构建体可以包含剪接供体和/或剪接受体序列,其在转录过程发生的RNA加工过程中是有功能性的。在一些实施方案中,剪接位点参与反式剪接。

[0380] 在一些实施方案中,本文提供的有效载荷可以任选地包含编码报告多肽和/或蛋白质的序列("报告序列")和/或编码选择标记的序列(例如,其赋予可以人工选择的性状,例如,抗性盒等)。报告序列的非限制性实例包括编码以下的DNA序列:β-内酰胺酶、β-半乳糖苷酶(LacZ)、碱性磷酸酶、胸苷激酶、绿色荧光蛋白(GFP)、红色荧光蛋白、mCherry荧光蛋白、黄色荧光蛋白、氯霉素乙酰转移酶(CAT)和荧光素酶。报告序列的其他实例是本领域已知的。当与驱动其表达的控制元件相关时,报告序列可以提供通过常规方式可检测的信号,所述常规方式包括酶促、射线照相、比色、荧光或其他光谱测定;荧光激活细胞分选(FACS)测定;免疫学测定(例如,酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和免疫组织化学)。类似地,当与驱动其表达的控制元件相关时,选择标记序列可以赋予也可以通过常规方式人工选择的性状。

[0381] 启动子

[0382] 术语"启动子"是指被酶/蛋白质识别的DNA序列,其可以启动和/或发起可操作连接的编码序列(例如,基因)的转录。例如,启动子通常是指例如RNA聚合酶和/或任何相关因子结合的并且启动子可以自其发起转录的核苷酸序列。因此,在一些实施方案中,有效载荷包含可操作地连接至本文所述的非限制性实例启动子中的一个的编码序列。

[0383] 在一些实施方案中,启动子在昆虫细胞中具有功能。在一些实施方案中,启动子是杆状病毒ie1启动子。在一些实施方案中,启动子是源自杆状病毒多角体蛋白(polh)基因、p6.9基因和/或p10基因的启动子。

[0384] 在一些实施方案中,启动子是诱导型启动子、组成型启动子、昆虫细胞启动子、病毒启动子、嵌合启动子、工程化启动子、组织特异性启动子或本领域已知的任何其他类型的启动子。

[0385] 术语"组成型"启动子是指当与编码序列(例如,蛋白质编码序列)可操作地连接时,导致在大多数或所有生理条件下从细胞中的核酸转录RNA的核苷酸序列。

[0386] 组成型启动子的实例包括但不限于逆转录病毒Rous肉瘤病毒(RSV)LTR启动子、巨细胞病毒(CMV)启动子(参见,例如,Boshart等人,Cell 41:521-530,1985,其以引用的方式整体并入本文)、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EF1-α启动子(Invitrogen)。

[0387] 诱导型启动子允许调控基因表达并且可以通过外源提供的化合物、环境因素诸如温度或特定生理状态例如急性期、细胞的特定分化状态或仅在复制细胞中的存在来调控。诱导型启动子和诱导型系统可以从各种商业来源获得,包括但不限于Invitrogen、Clontech和Ariad。诱导型启动子的其他实例是本领域已知的。

[0388] 受外源提供的化合物调控的诱导型启动子的实例包括锌诱导型绵羊金属硫蛋白(MT)启动子、地塞米松(Dex)诱导型小鼠乳腺肿瘤病毒(MMTV)启动子、T7聚合酶启动子系统、蜕皮激素昆虫启动子、四环素-抑制型系统、四环素诱导型系统、RU486诱导型系统和雷帕霉素诱导型系统。

[0389] 术语“组织特异性”启动子是指仅在某些特定细胞类型和/或组织中有活性的启动子(例如,特定基因的转录仅发生在表达与组织特异性启动子结合的转录调控和/或控制蛋白的细胞内)。

[0390] 在一些实施方案中,调控和/或控制序列赋予组织特异性基因表达能力。在一些情况下,组织特异性调控和/或控制序列结合以组织特异性方式诱导转录的组织特异性转录因子。

[0391] 增强子

[0392] 术语“增强子”是指可以增加核酸编码序列(例如,蛋白质)的转录水平的核苷酸序列。增强子序列(长度通常为50-1500bp)通常通过为转录相关蛋白(例如转录因子)提供其他结合位点来提高转录水平。在一些实施方案中,在内含子序列中存在增强子序列。与启动子序列不同,增强子序列可以在距离转录起始位点远得多的位置(例如,与启动子相比)起作用。增强子的非限制性实例包括RSV增强子、CMV增强子、CMV早期增强子、cAMP反应元件(CRE)增强子和/或SV40增强子。

[0393] 在一些实施方案中,增强子包括杆状病毒增强子序列(hr1)。

[0394] 文库构建体元件相对于病毒重复序列的位置

[0395] 文库构建体可以包含一个或多个工程化序列。在一些实施方案中,文库构建体包含多个工程化序列。在一些实施方案中,文库构建体包含用于包装病毒载体的病毒重复序列。

[0396] 文库构建体上的一个或多个和/或多个工程化序列可以位于相对于彼此和相对于文库构建体上的病毒重复序列的许多位置和/或位置组合。例如,在一些实施方案中,昆虫细胞或昆虫细胞群包含至少一个包含多个工程化序列的文库构建体,其中多个工程化序列的第一子集位于病毒重复序列之间,并且多个工程化序列的第二子集位于病毒重复序列之外。

[0397] 在一些实施方案中,文库构建体上的多个工程化序列包含至少一个文库变体和至少一个标识符。在一些实施方案中,所有文库变体和标识符都位于病毒重复序列之间。在一些实施方案中,标识符位于病毒重复序列之间并且所有文库变体都位于病毒重复序列之外。在一些实施方案中,多个工程化序列包含至少两个文库变体和至少一个标识符。在一些

实施方案中,标识符和至少一个文库变体位于病毒重复序列之间,并且至少一个文库变体位于病毒重复序列之外。

[0398] 在一些实施方案中,多个工程化序列包含至少一个文库变体、至少一个标识符和至少一个有效载荷。在一些实施方案中,所有文库变体、标识符和有效载荷都位于病毒重复序列之间。在一些实施方案中,所有标识符和有效载荷都位于病毒重复序列之间,并且所有文库变体都位于病毒重复序列之外。在一些实施方案中,多个工程化序列包含至少两个文库变体、至少一个有效载荷和至少一个标识符。在一些实施方案中,所有标识符、有效载荷和至少一个文库变体位于病毒重复序列之间,并且至少一个文库变体位于病毒重复序列之外。

[0399] 在一些实施方案中,文库构建体还包含至少一个工程化序列,所述至少一个工程化序列包含至少一个报告基因和/或选择标记。在一些实施方案中,所有报告基因和/或选择标记都位于病毒重复序列之间。在一些实施方案中,所有报告基因和/或选择标记都位于病毒重复序列之外。在一些实施方案中,文库构建体包含至少一个报告基因和至少一个选择标记。在一些实施方案中,至少一个报告基因和至少一个选择标记都位于病毒重复序列之间。在一些实施方案中,至少一个报告基因和至少一个选择标记都位于病毒重复序列之外。在一些实施方案中,至少一个报告基因位于病毒重复序列之间并且至少一个选择标记位于病毒重复序列之外。在一些实施方案中,至少一个选择标记位于病毒重复序列之间,并且至少一个报告基因位于病毒重复序列之外。

[0400] 在一些实施方案中,文库构建体还包含至少一个包含至少一个条形码的工程化序列。在一些实施方案中,所有条形码都位于病毒重复序列之间。在一些实施方案中,所有条形码都位于病毒重复序列之外。在一些实施方案中,文库构建体包含多个条形码,其中第一子集位于病毒重复序列之间,并且第二子集位于病毒重复序列之外。

[0401] 在一些实施方案中,文库构建体包含至少一个标识符和至少一个条形码。在一些实施方案中,至少一个条形码用作标识符,并且在一些实施方案中,至少一个条形码不用作标识符。在一些实施方案中,至少一个条形码位于病毒重复序列之间,并且至少一个条形码或其一部分用作标识符。在一些实施方案中,至少一个条形码位于病毒重复序列之间,并且至少一个条形码不用作标识符。在一些实施方案中,文库构建体包含至少一个标识符和至少两个条形码。在一些实施方案中,所有标识符和至少一个条形码位于病毒重复序列之间,并且至少一个条形码位于病毒重复序列之外,其中位于病毒重复序列之间的至少一个条形码(或其一部分)用作标识符。在一些实施方案中,所有标识符和至少一个条形码位于病毒重复序列之间,并且至少一个条形码位于病毒重复序列之外,其中至少两个条形码均不用作标识符。

[0402] 在一些实施方案中,文库构建体包含至少两个条形码并且还包含至少一个其他工程化序列,其中没有条形码用作标识符并且所有条形码都用于追踪其他工程化序列。例如,在一些实施方案中,文库构建体包含至少两个条形码并且还包含至少一个文库变体,其中没有条形码用作标识符并且所有条形码都用于追踪文库变体。在一些实施方案中,一个条形码用作标识符并且其他条形码用于追踪文库变体。

[0403] 在一些实施方案中,一个条形码指示单个工程化序列(例如,单个文库变体)。在一些实施方案中,一个条形码指示多于一个工程化序列(例如,多于一个文库变体)。在一些实

施方案中,多于一个条形码指示一个工程化序列(例如,一个文库变体)。在一些实施方案中,多于一个条形码指示多于一个工程化序列(例如,多于一个文库变体)。在一些实施方案中,没有条形码用作标识符并且所有条形码都用于追踪包含条形码的其他工程化序列。在一些实施方案中,一个条形码用作标识符并且其他条形码用于追踪包含条形码的其他工程化序列。

[0404] 昆虫细胞中构建体的使用

[0405] 构建体和工程化序列的其他特征

[0406] 启动子

[0407] 在一些实施方案中,构建体包含启动子。术语“启动子”是指被酶/蛋白质识别的DNA序列,其可以启动和/或发起可操作连接的编码序列(例如,基因)的转录。例如,启动子通常是指例如RNA聚合酶和/或任何相关因子结合的并且启动子可以自其发起转录的核苷酸序列。因此,在一些实施方案中,工程化序列包含可操作地连接至启动子的编码序列。多种启动子是本领域已知的,其可以用于本文。

[0408] 在一些实施方案中,启动子是诱导型启动子、组成型启动子、昆虫细胞启动子、病毒启动子、嵌合启动子、工程化启动子、组织特异性启动子或本领域已知的任何其他类型的启动子。

[0409] 术语“组成型”启动子是指当与编码序列(例如,蛋白质编码序列)可操作地连接时,导致在大多数或所有生理条件下从细胞中的核酸转录RNA的核苷酸序列。

[0410] 组成型启动子的实例包括但不限于逆转录病毒Rous肉瘤病毒(RSV)LTR启动子、巨细胞病毒(CMV)启动子(参见,例如,Boshart等人,Cell 41:521-530,1985,其以引用的方式整体并入本文)、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EF1-α启动子(Invitrogen)。

[0411] 诱导型启动子允许调控基因表达并且可以通过外源提供的化合物、环境因素诸如温度或特定生理状态例如急性期、细胞的特定分化状态或仅在复制细胞中的存在来调控。诱导型启动子和诱导型系统可以从各种商业来源获得,包括但不限于Invitrogen、Clontech和Ariad。诱导型启动子的其他实例是本领域已知的。

[0412] 受外源提供的化合物调控的诱导型启动子的实例包括锌诱导型绵羊金属硫蛋白(MT)启动子、地塞米松(Dex)诱导型小鼠乳腺肿瘤病毒(MMTV)启动子、T7聚合酶启动子系统、蜕皮激素昆虫启动子、四环素-抑制型系统、四环素诱导型系统、RU486诱导型系统和雷帕霉素诱导型系统。

[0413] 术语“组织特异性”启动子是指仅在某些特定细胞类型和/或组织中有活性的启动子(例如,特定基因的转录仅发生在表达与组织特异性启动子结合的转录调控和/或控制蛋白的细胞内)。

[0414] 在一些实施方案中,调控和/或控制序列赋予组织特异性基因表达能力。在一些情况下,组织特异性调控和/或控制序列结合以组织特异性方式诱导转录的组织特异性转录因子。

[0415] 增强子

[0416] 在一些情况下,构建体可以包含增强子序列。术语“增强子”是指可以增加核酸编码序列(例如,蛋白质)的转录水平的核苷酸序列。增强子序列(长度通常为50-1500bp)通常

通过为转录相关蛋白(例如转录因子)提供其他结合位点来提高转录水平。在一些实施方案中,在内含子序列中存在增强子序列。与启动子序列不同,增强子序列可以在距离转录起始位点远得多的位置(例如,与启动子相比)起作用。增强子的非限制性实例包括RSV增强子、CMV增强子、CMV早期增强子、cAMP反应元件(CRE)增强子和/或SV40增强子。

[0417] 其他序列

[0418] 在一些实施方案中,本文所述的任何构建体可以包含非翻译区(UTR),诸如5'UTR或3'UTR。基因的UTR被转录但不被翻译。5'UTR从转录起始位点开始,并持续到起始密码子,但不包含起始密码子。3'UTR紧随终止密码子开始,并持续直至转录终止信号。

[0419] 在一些实施方案中,包含编码本文提供的蛋白质的工程化序列的构建体可以包含聚腺苷酸化(poly(A))信号序列。大多数新生真核mRNA在其3'端具有poly(A)尾,这是在一个复杂的过程中添加的,所述过程包括初级转录物的切割和由poly(A)信号序列驱动的偶联聚腺苷酸化反应(参见,例如,Proudfoot等人,Cell 108:501-512,2002,其以引用的方式整体并入本文)。poly(A)尾赋予mRNA稳定性和可转移性(B.Alberts等人的Molecular Biology of the Cell,第三版,Garland Publishing,1994,其以引用的方式整体并入本文)。在一些实施方案中,poly(A)信号序列位于编码序列的3'。

[0420] 在一些实施方案中,包含编码蛋白质的工程化序列的构建体可以包含内部核糖体进入位点(IRES)。IRES形成复杂的二级结构,允许从任何位置发生翻译起始,mRNA紧接在IRES所在位置的下游(参见,例如,Pelletier和Sonenberg,Mal.Cell.Biol.8(3):1103-1112,1988)。本领域技术人员已知有几种IRES序列,包括来自例如口蹄疫病毒(FMDV)、脑心肌炎病毒(EMCV)、人鼻病毒(HRV)、蟋蟀麻痹病毒、人免疫缺陷病毒(HIV)、甲型肝炎病毒(HAV)、丙型肝炎病毒(HCV)和脊髓灰质炎病毒(PV)的那些。

[0421] 在一些实施方案中,本文提供的任何构建体可以包含剪接供体和/或剪接受体序列,其在转录过程发生的RNA加工过程中是有功能性的。在一些实施方案中,剪接位点参与反式剪接。

[0422] 在一些实施方案中,本文提供的任何构建体可以任选地包含编码报告多肽和/或蛋白质的序列("报告序列")和/或编码选择标记的序列(例如,其赋予可以人工选择的性状,例如,抗性盒等)。报告序列的非限制性实例包括编码以下的DNA序列:β-内酰胺酶、β-半乳糖苷酶(LacZ)、碱性磷酸酶、胸苷激酶、绿色荧光蛋白(GFP)、红色荧光蛋白、mCherry荧光蛋白、黄色荧光蛋白、氯霉素乙酰转移酶(CAT)和荧光素酶。报告序列的其他实例是本领域已知的。当与驱动其表达的控制元件相关时,报告序列可以提供通过常规方式可检测的信号,所述常规方式包括酶促、射线照相、比色、荧光或其他光谱测定;荧光激活细胞分选(FACS)测定;免疫学测定(例如,酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和免疫组织化学)。类似地,当与驱动其表达的控制元件相关时,选择标记序列可以赋予也可以通过常规方式人工选择的性状。

[0423] 昆虫细胞工程化和细胞构建体进入

[0424] 在一些实施方案中,本公开提供了包括制备或获得昆虫细胞文库的方法。适用于遗传修饰的任何方法都可以用于在本公开的上下文中操作昆虫细胞,例如以引入文库构建体、引入另一构建体和/或引入扰动。例如,本领域已知许多用于将外源性核酸(例如,DNA)引入真核细胞的方法,包括转染、感染(例如,病毒转导)或电穿孔。

[0425] 用于将构建体递送至昆虫细胞的方法可以根据需要而变化。在某些实施方案中，某些构建体可以作为一个或多个质粒中的核酸(例如,DNA)构建体递送。递送方法包括但不限于电穿孔、显微注射、基因枪、穿刺转染、静水压、连续输注、超声处理、磁转染、化学媒介物(例如,寡核苷酸、脂质复合物、聚合物囊泡、复合物、树枝状聚合物、无机纳米颗粒和细胞穿膜肽)、病毒载体(例如,有复制能力的病毒载体、无复制能力的病毒载体、复制缺乏的病毒载体、复制缺陷的病毒载体、有复制能力的病毒载体,和/或条件复制病毒载体)。

[0426] 在一些实施方案中,昆虫细胞文库被遗传修饰以包含文库构建体和/或其他构建体。文库构建体和/或其他构建体可以通过例如对于核酸操作技术人员而言本领域已知的任何合适的方法来递送并且包括遗传工程化、重组工程化和合成技术(参见,例如,Sambrook等人,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Press,Cold Spring Harbor,N.Y,其以引用的方式整体并入本文)。在一些实施方案中,所提供的方法包括将组成文库构建体的一个或多个核酸转染到文库的昆虫细胞中。

[0427] 在一些实施方案中,昆虫细胞文库被遗传修饰以包含一个或多个产生病毒载体所必需的核酸序列。在一些实施方案中,表达和/或产生病毒载体的昆虫细胞文库可以具有一种或多种以反式形式提供给昆虫细胞的病毒载体组分。例如,可以使用任何适当的构建体将重组AAV构建体和/或工程化序列、rep序列、cap序列和产生本公开的AAV载体所需的辅助功能递送至包装宿主细胞。

[0428] 编码病毒元件的构建体可以通过例如对于核酸操作技术人员而言本领域已知的任何合适的方法来递送并且包括遗传工程化、重组工程化和合成技术(参见,例如,Sambrook等人,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Press,Cold Spring Harbor,N.Y,其以引用的方式整体并入本文)。在一些实施方案中,所提供的方法包括将一个或多个组成病毒构建体或编码病毒载体组分的核酸转染到文库的昆虫细胞中。在一些实施方案中,所提供的方法包括转染一个或多个含有编码病毒载体的核酸分子或基本上由其组成的质粒。在一些实施方案中,病毒载体组分包含在一个或多个转染到细胞中的质粒上。在一些实施方案中,病毒载体组分包含在一个、两个、三个或更多个各自转染到细胞中的质粒上。

[0429] 在一些实施方案中,使用杆状病毒将病毒载体组分感染到昆虫细胞中。在一些实施方案中,病毒载体组分包含在一个或多个转染到昆虫细胞中的杆状病毒载体上。

[0430] 在一些实施方案中,病毒载体是腺相关病毒(AAV),并且病毒转染包括将含有编码AAV载体的核酸分子或基本上由其组成的质粒转染到昆虫细胞中,并提供AAV载体复制和包装必要的AAV载体rep和/或cap。在一些实施方案中,将各自包含AAV载体组分的一个、两个、三个或更多个质粒转染到细胞中。在一些特定实施方案中,组合提供各种AAV组分(包括辅助病毒、有效载荷(例如,治疗基因)、rep和/或cap)的质粒(例如,1至5个质粒,例如,3个质粒)被转染到细胞中。

[0431] 在一些特定实施方案中,AAV构建体的病毒转染可以使用三重转染方法完成(例如,如美国专利号6,001,650中所述,其以引用的方式整体并入本文)。在一些实施方案中,AAV载体通过用一个或多个构建体转染宿主细胞而产生,所述一个或多个构建体包含一个或多个产生AAV载体所必需的核酸序列,包括但不限于rep序列和/或cap序列,和/或包含辅助功能的构建体。在一些实施方案中,rep和cap序列以反式形式发挥作用,用于生产性AAV

载体复制和衣壳化。在一些实施方案中,包含rep和/或cap序列的构建体支持高效AAV构建体产生而不生成任何可检测的野生型AAV载体(即,含有功能性rep和cap基因的AAV载体)。辅助功能构建体编码AAV载体复制所依赖的非AAV衍生病毒和/或细胞功能的核苷酸序列,所述序列可以具有AAV构建体复制所需的那些功能,包括但不限于参与AAV载体基因转录的激活、阶段特异性AAV载体mRNA剪接、AAV DNA复制、cap表达产物的合成和/或AAV载体衣壳组装的那些部分。这些基于病毒的功能可以源自任何已知的辅助病毒,诸如像腺病毒、疱疹病毒(除了1型单纯疱疹病毒之外)和痘苗病毒。

[0432] 在一些实施方案中,防止产生病毒载体再感染昆虫细胞文库的昆虫细胞可能是有益的。在一些实施方案中,昆虫细胞文库的昆虫细胞已经被修饰以破坏或去除产生的病毒载体的受体。在一些实施方案中,昆虫细胞已经用感染/阻断剂处理。

[0433] 构建体和工程化序列在昆虫细胞中的位置

[0434] 本文所述的任何构建体可以以在染色体外的形式存在于昆虫细胞中和/或被整合到昆虫细胞基因组中。

[0435] 在一些实施方案中,文库构建体是游离的和/或被整合到昆虫细胞基因组中。在一些实施方案中,文库构建体是游离的单个连续构建体。在一些实施方案中,文库构建体是整合到昆虫细胞基因组中的单个连续构建体。在一些实施方案中,文库构建体是不连续的并且一个或多个单独的构建体是游离的。在一些实施方案中,文库构建体是不连续的并且一个或多个单独的构建体被整合到昆虫细胞基因组中。在一些实施方案中,文库构建体是不连续的并且至少一个构建体是游离的并且至少一个构建体被整合到昆虫细胞基因组中。

[0436] 在一些实施方案中,一个或多个产生病毒载体所必需的序列以在染色体外的形式(例如,游离)存在于昆虫细胞内。在一些实施方案中,一个或多个产生病毒载体所必需的序列被整合到昆虫细胞的基因组中。

[0437] 在一些实施方案中,一个或多个产生病毒载体所必需的序列被整合到昆虫细胞的基因组中,其中一个或多个序列是诱导表达的。在一些实施方案中,所有产生病毒载体所必需的序列被整合到昆虫细胞的基因组中,其中一个或多个序列是诱导表达的。在一些实施方案中,所有产生病毒载体所必需的序列被整合到昆虫细胞的基因组中,其中两个或更多个序列是诱导表达的。在一些实施方案中,所有产生病毒载体所必需的序列被整合到昆虫细胞的基因组中,其中所有序列是诱导表达的。

[0438] 可以使用本领域已知的用于整合序列的任何方法。在一些实施方案中,整合是靶向整合(例如,在预定位点)。在一些实施方案中,整合是随机整合(例如,在随机位点)。在一些实施方案中,整合在预定的基因组位置子集中是随机的。

[0439] 在一些实施方案中,使用本领域已知的方法将一个或多个工程序列(例如,文库构建体和/或其他构建体)预先整合到昆虫文库的细胞中。在一些实施方案中,将具有一个或多个整合序列的昆虫细胞用编码病毒载体的序列转染。在一些实施方案中,一个或多个病毒载体序列也被整合到基因组中。

[0440] 在一些实施方案中,可以使用本领域技术人员已知的方法将用于培养病毒载体的昆虫细胞文库稳定地工程化为含有一种或多种此类病毒组分(例如,重组AAV构建体、rep序列、cap序列和/或辅助功能)。在一些实施方案中,此类稳定的昆虫细胞含有在诱导型启动子控制下的此类病毒载体组分。在一些实施方案中,此类病毒载体组分可以处于组成型启

动子的控制之下。在一些实施方案中,选定的稳定昆虫细胞可以含有处于组成型启动子控制之下的选定组分和处于一个或多个诱导型启动子控制之下的其他选定组分。例如,例如含有处于组成型启动子控制之下的E1辅助功能,以及处于诱导型启动子控制之下的rep和/或cap基因的稳定昆虫细胞。其他稳定的昆虫细胞可以由本领域技术人员使用常规方法作为产生昆虫细胞文库的基础来生成。然后可以操作此类稳定表达昆虫细胞系以包含多种不同的工程化序列,包括标识符序列,从而生成产生病毒载体的昆虫细胞文库。

[0441] 在一些实施方案中,外源性DNA(例如文库构建体)被整合到基因组DNA中,使得外源性DNA与昆虫细胞的基因组DNA邻接。在一些实施方案中,整合由细胞内源性的天然DNA修复机制介导。例如,通过将外源性DNA引入细胞,使位点特异性核酸酶产生整合位点,并使供体DNA被整合,整合即可发生。细胞可以在培养物中保持足够的时间以使DNA被整合。这通常将产生混合细胞群,包括(i)外源性DNA在由位点特异性核酸酶产生的整合位点整合到其中的重组细胞,以及任选的(ii)外源性DNA在除了所需的整合位点之外的位点整合到其中的细胞和/或任选的(iii)外源性DNA未整合到其中的细胞。本领域已知的选择方法可以用于富集文库中具有遗传组合物的细胞。

[0442] 游离文库构建体

[0443] 在一些实施方案中,本公开提供了昆虫细胞,其中一个或多个构建体(例如,文库构建体和/或编码产生病毒载体所必需的序列的构建体)没有被整合到昆虫细胞基因组中而是游离存在。可以使用本领域已知的用于将游离构建体引入昆虫细胞的任何方法。

[0444] 在一些实施方案中,一个或多个构建体是杆状病毒构建体。在一些实施方案中,文库构建体包括杆状病毒构建体。

[0445] 在一些实施方案中,单个连续文库构建体游离存在于昆虫细胞中。在一些实施方案中,不连续文库构建体的一个或多个单独的构建体游离存在于昆虫细胞中(例如,包含标识符和病毒包装序列的构建体)。在一些实施方案中,不连续文库构建体的所有单独的构建体游离存在于昆虫细胞中。

[0446] 在一些实施方案中,可以通过标准转染或电穿孔方法将外源性核酸(例如,DNA)引入昆虫细胞中。在一些实施方案中,引入昆虫细胞中的外源性核酸可以以单拷贝数、低拷贝数或高拷贝数游离存在于细胞中。在一些实施方案中,引入昆虫细胞的外源性核酸可以以可变的拷贝数以基因组形式存在和/或游离存在。本领域已知的选择和/或筛选方法可以用于富集具有游离构建体(例如,文库构建体)的细胞。

[0447] 在一些实施方案中,使用杆状病毒表达系统将外源核酸引入昆虫细胞。在一些实施方案中,杆状病毒构建体有利于将核酸转移到昆虫细胞中,在昆虫细胞中游离表达。在某些实施方案中,使用杆状病毒表达系统将编码AAV基因组(例如,侧翼有ITR序列的序列)的外源核酸(例如,DNA)引入昆虫细胞。

[0448] 在一些实施方案中,由第一昆虫细胞文库生成的病毒载体池用于转染昆虫细胞以生成第二昆虫细胞文库。在一些实施方案中,第二昆虫细胞文库以游离形式表达文库变体。在一些实施方案中,由第一昆虫细胞文库生成的病毒载体池是AAV载体池。在一些实施方案中,由第一昆虫细胞文库生成的AAV载体池用于转导昆虫细胞以生成以游离形式包含标识符的第二昆虫细胞文库。

[0449] 整合的文库构建体

[0450] 在一些实施方案中,本公开提供了昆虫细胞,其中一个或多个构建体(例如,文库构建体和/或编码产生病毒载体所必需的序列的构建体)被整合到昆虫细胞基因组中。可以使用本领域已知的用于整合序列的任何方法。

[0451] 在一些实施方案中,将构建体(例如,文库构建体)整合到昆虫细胞基因组内是由至少一个反式作用整合序列介导的。在一些实施方案中,至少一个反式作用整合序列包含(i)整合构建体,(ii)重组酶,(iii)核酸酶,(iv)转座酶,和/或其衍生物和/或融合物。在一些实施方案中,将构建体(例如,文库构建体)整合到昆虫细胞基因组内是由至少一个与顺式作用整合序列相协调的反式作用整合序列介导的。在一些实施方案中,顺式作用整合序列包含(i)病毒重复序列,(ii)重组酶识别位点,(iii)同源臂,和/或(iv)转座酶识别位点和/或其衍生物。

[0452] 在一些实施方案中,待整合到昆虫基因组中的构建体(例如,文库构建体)包含至少一对位于待整合序列侧翼的顺式作用整合序列。例如,提供了一种文库构建体,其包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列,其中第一组病毒重复序列侧翼为顺式作用整合序列。

[0453] 在一些实施方案中,将构建体整合到昆虫细胞基因组中是靶向整合。在一些实施方案中,将构建体整合到昆虫细胞基因组中是随机整合。在一些实施方案中,构建体在随机插入位点处被整合到昆虫细胞基因组中。在一些实施方案中,随机插入位点在预定的基因组位置子集中是随机的。在一些实施方案中,文库构建体在昆虫细胞基因组中的插入包括预定的插入位点。

[0454] 在一些实施方案中,文库构建体(例如,一个或多个与文库构建体相关的单独核酸)被整合到昆虫基因组中。在一些实施方案中,单个连续文库构建体被整合到昆虫基因组中。在一些实施方案中,不连续文库构建体的一个或多个单独的构建体(例如,包含标识符和病毒包装序列的构建体)游被整合到昆虫基因组中。在一些实施方案中,不连续文库构建体的所有单独的构建体被整合到昆虫基因组中。

[0455] 在一些实施方案中,可以通过标准转染或电穿孔方法将外源性核酸(例如,DNA)引入昆虫细胞中。在一些实施方案中,引入昆虫细胞中的外源性核酸可以以单拷贝数、低拷贝数或高拷贝数整合到基因组中或以单拷贝数、低拷贝数或高拷贝数游离存在于细胞中。在一些实施方案中,引入昆虫细胞的外源性核酸可以以可变的拷贝数以基因组形式存在和/或游离存在。本领域已知的选择和/或筛选方法可以用于富集具有整合的构建体(例如,文库构建体)的细胞。

[0456] 病毒转导介导的整合

[0457] 在一些实施方案中,文库构建体通过病毒转导介导的整合被整合到昆虫基因组中。在一些实施方案中,例如,提供了一种文库构建体,其包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列,其中第一组病毒重复序列侧翼为包含第二组病毒重复序列的顺式作用整合序列。在一些实施方案中,至少一个反式作用整合序列包含整合构建体,其中整合构建体是无脊椎细小病毒载体,例如像鹿眼蛱蝶浓核病毒(JcDNV)或其衍生物。例如,参见Bossin H,Fournier P,Royer C,等人Junonia coenia densovirus-based vectors for stable transgene expression in Sf9 cells:influence of the densovirus sequences on genomic integration.J Virol.2003;77(20):11060-11071,其特此通过引

用整体并入。

[0458] 核酸酶介导的整合

[0459] 在一些实施方案中,文库构建体通过核酸酶介导的整合被整合到昆虫基因组中。在一些实施方案中,例如,提供了一种文库构建体,其包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列,其中第一组病毒重复序列侧翼为包含同源臂的顺式作用整合序列。

[0460] 在一些实施方案中,至少一个反式作用整合序列包含核酸酶和/或其融合物和/或衍生物,其包含Cas9、CasZ、Cpf1、与可编程DNA结合结构域(诸如TALE蛋白(TALEN)或锌指蛋白(ZFN))融合的工程化Fok1核酸酶结构域,和/或大范围核酸酶和/或其衍生物。在一些实施方案中,核酸酶是或包含Cas9。

[0461] 在一些实施方案中,至少一个反式作用整合序列包含锌指核酸酶。锌指核酸酶(ZFN)是人工限制性酶,其是通过将锌指DNA结合结构域与DNA切割结构域组合生成的。ZFN中最常见的切割结构域是来自IIs型限制性核酸内切酶FokI的非特异性切割结构域,可以使用引导序列靶向所述结构域。使用ZFN的方法描述于例如WO 2009146179A1、WO 2008060510 A2和CN 102174576 A中,其以引用的方式整体并入。

[0462] 在一些实施方案中,至少一个反式作用整合序列包含转录激活因子样效应物核酸酶(TALEN)。TALEN是限制性酶,其可以被工程化以切割特定的DNA序列。使用TALEN的方法描述于例如WO 2014134412 A1、WO 2013163628 A2和WO 2014040370A1中,其以引用的方式整体并入。

[0463] 在一些实施方案中,至少一个反式作用整合序列包含非RNA引导的核酸酶和/或其融合物和/或衍生物。在一些实施方案中,文库构建体在昆虫细胞基因组内的插入是由至少两个反式作用整合序列介导的,其中至少一个反式作用整合序列包含RNA引导的核酸酶和/或其融合物和/或衍生物,并且至少一个反式作用整合序列包含gRNA序列。

[0464] 在一些实施方案中,文库构建体在昆虫细胞基因组中的插入是由至少一个反式作用整合序列介导的,并且其中至少一个反式作用整合序列包含核酸酶(例如,RNA引导的核酸酶或非RNA引导的核酸酶)和/或其融合物和/或衍生物,并且其中文库构建体包含至少一对为同源臂的顺式作用整合序列。在一些实施方案中,文库构建体包含两个病毒包装序列(例如,病毒重复序列)之间的标识符和/或有效载荷,所有这些都位于包含同源臂序列的顺式作用整合序列之间。

[0465] 在一些实施方案中,文库构建体在昆虫细胞基因组中的插入是由至少两个反式作用整合序列介导的,其中至少一个反式作用整合序列包含核酸酶(例如,RNA引导的核酸酶)和/或其融合物和/或衍生物,并且至少一个反式作用整合序列包含gRNA序列。在一些实施方案中,文库构建体包含两个病毒包装序列(例如,病毒重复序列)之间的标识符和/或有效载荷,所有这些都位于包含同源臂序列的顺式作用整合序列之间。

[0466] 在一些实施方案中,文库构建体通过核酸酶介导的整合(例如,使用CRISPR/Cas9)被整合到昆虫细胞的基因组中。在一些实施方案中,文库构建体的整合由细胞内源性的天然DNA修复机制介导。例如,通过将外源性DNA引入细胞,使位点特异性核酸酶产生整合位点,并使供体DNA被整合,整合即可发生。

[0467] 在一些实施方案中,文库构建体在昆虫细胞基因组内的插入是由Cas9介导的。在

一些实施方案中,文库构建体包含AAV ITR序列之间的标识符和/或有效载荷,所有这些都位于包含同源臂序列的顺式作用整合序列之间。

[0468] 下表5中提供了核酸酶介导的文库构建体整合的一些示例性设想。表4的每个实施方案都可以在单个连续文库构建体或不连续文库构建体的上下文中。在一些实施方案中,文库构建体是不连续文库构建体并且包含标识符的单独的构建体被整合到基因组中。在一些实施方案中,文库构建体是不连续文库构建体,并且文库构建体的两个或更多个单独的构建体通过核酸酶介导的整合被整合到昆虫基因组中。在一些实施方案中,不连续文库构建体的所有单独的构建体通过核酸酶介导的整合被整合到昆虫基因组中。

[0469] 表4-核酸酶介导的整合的示例性设想

[0470]

设想	文库变体	位置	顺式作用整合序列	反式作用整合序列
CRISPR介导的gRNA KO文库整合	gRNA，通过转染引入	靶向	同源臂	核酸酶(例如，Cas)和gRNA
CRISPR介导的gRNA SNP文库整合	gRNA，通过转染引入	靶向	同源臂	核酸酶(例如，Cas)和gRNA
CRISPR介导的gRNA激活文库整合	gRNA，通过转染引入	靶向	同源臂	核酸酶(例如，Cas)和gRNA
CRISPR介导的gRNA抑制文库整合	gRNA，通过转染引入	靶向	同源臂	核酸酶(例如，Cas)和gRNA
CRISPR介导的插入文库整合	ORF，通过转染引入	靶向	同源臂	核酸酶(例如，Cas)和gRNA

[0471] 重组酶介导的整合

[0472] 在一些实施方案中,文库构建体通过重组酶介导的整合被整合到昆虫基因组中。在一些实施方案中,例如,提供了一种文库构建体,其包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列,其中第一组病毒重复序列侧翼为包含重组酶识别位点的顺式作用整合序列。

[0473] 在一些实施方案中,至少一个反式作用整合序列包含重组酶。在一些实施方案中,重组酶包含Cre、Flp、Dre、PhiC31和/或Bxb1,和/或其衍生物和/或融合物。

[0474] 在一些实施方案中,反式作用整合序列包含有包含Cre的重组酶。在一些实施方案中,Cre介导文库构建体或其一部分整合到昆虫细胞基因组中。在一些实施方案中,文库构建体包含位于标识符和/或有效载荷侧翼的第一组病毒重复序列(例如,用于包装AAV载体的AAV ITR),其中第一组病毒重复序列侧翼为包含LoxP位点的顺式作用整合序列。

[0475] 在一些实施方案中,反式作用整合序列包含有包含Bxb1的重组酶。在一些实施方案中,Bxb1介导文库构建体或其一部分整合到昆虫细胞基因组中。在一些实施方案中,文库构建体包含位于标识符和/或有效载荷侧翼的第一组病毒重复序列(例如,用于包装AAV载体的AAV ITR),其中第一组病毒重复序列侧翼为包含Att位点的顺式作用整合序列。

[0476] 在一些实施方案中,反式作用整合序列包含有包含Flp的重组酶。在一些实施方案中,Flp介导文库构建体或其一部分整合到昆虫细胞基因组中。在一些实施方案中,文库构建体包含位于标识符和/或有效载荷侧翼的第一组病毒重复序列(例如,用于包装AAV载体的AAV ITR),其中第一组病毒重复序列侧翼为包含Frt位点的顺式作用整合序列。

[0477] 下表6中提供了重组酶介导的文库构建体整合的一些示例性设想。表5的每个实施方案都可以在单个连续文库构建体或不连续文库构建体的上下文中。在一些实施方案中，文库构建体是不连续文库构建体并且包含标识符的单独的构建体被整合到基因组中。在一些实施方案中，文库构建体是不连续文库构建体，并且文库构建体的两个或更多个单独的构建体通过重组酶介导的整合被整合到昆虫基因组中。在一些实施方案中，不连续文库构建体的所有单独的构建体通过重组酶介导的整合被整合到昆虫基因组中。

[0478] 表5-重组酶介导的整合的示例性设想

[0479]

设想	文库变体	位置	顺式作用整合序列	反式作用整合序列
重组酶介导的gRNA KO文库整合	gRNA，通过转染引入	靶向	识别位点诸如LoxP、Att、Frt	重组酶诸如Cre、Bxb1、Flp
重组酶介导的gRNA SNP文库整合	gRNA，通过转染引入	靶向	识别位点诸如LoxP、Att、Frt	重组酶诸如Cre、Bxb1、Flp
重组酶介导的gRNA激活文库整合	gRNA，通过转染引入	靶向	识别位点诸如LoxP、Att、Frt	重组酶诸如Cre、Bxb1、Flp
重组酶介导的gRNA抑制文库整合	gRNA，通过转染引入	靶向	识别位点诸如LoxP、Att、Frt	重组酶诸如Cre、Bxb1、Flp
重组酶介导的插入文库整合	ORF，通过转染引入	靶向	识别位点诸如LoxP、Att、Frt	重组酶诸如Cre、Bxb1、Flp

[0480] 转座酶介导的整合

[0481] 在一些实施方案中，文库构建体通过转座酶介导的整合被整合到昆虫基因组中。在一些实施方案中，例如，提供了一种文库构建体，其包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列，其中第一组病毒重复序列侧翼为包含转座酶识别位点的顺式作用整合序列。

[0482] 在一些实施方案中，至少一个反式作用整合序列包含转座酶。在一些实施方案中，转座酶包含睡美人转座酶和/或Piggybac转座酶和/或其衍生物和/或融合物。

[0483] 在一些实施方案中，反式作用整合序列包含有包含睡美人转座酶的转座酶。在一些实施方案中，睡美人转座酶介导文库构建体或其一部分整合到昆虫细胞基因组中。在一些实施方案中，文库构建体包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列(例如，用于包装AAV载体的AAV ITR)，其中第一组病毒重复序列侧翼为包含转座酶识别位点的顺式作用整合序列。

[0484] 在一些实施方案中，反式作用整合序列包含有包含Piggybac转座酶的转座酶。在一些实施方案中，Piggybac转座酶介导文库构建体或其一部分整合到昆虫细胞基因组中。在一些实施方案中，文库构建体包含位于标识符和/或有效载荷的侧翼的第一组病毒重复序列(例如，用于包装AAV载体的AAV ITR)，其中第一组病毒重复序列侧翼为包含转座酶识别位点的顺式作用整合序列。

[0485] 下表6中提供了转座酶介导的文库构建体整合的一些示例性设想。表7的每个实施方案都可以在单个连续文库构建体或不连续文库构建体的上下文中。在一些实施方案中，文库构建体是不连续文库构建体并且包含标识符的单独的构建体被整合到基因组中。在一些实施方案中，文库构建体是不连续文库构建体，并且文库构建体的两个或更多个单独的构建体通过转座酶介导的整合被整合到昆虫基因组中。在一些实施方案中，不连续文库构建体的所有单独的构建体通过转座酶介导的整合被整合到昆虫基因组中。

[0486] 表6-转座酶介导的整合的示例性设想

[0487]

设想	文库变体	位置	顺式作用整合序列	反式作用整合序列
转座酶介导的gRNA KO文库整合	gRNA，通过转染引入	随机	识别位点	转座酶诸如Piggybac或睡美人
转座酶介导的gRNA SNP文库整合	gRNA，通过转染引入	随机	识别位点	转座酶诸如Piggybac或睡美人
转座酶介导的gRNA激活文库整合	gRNA，通过转染引入	随机	识别位点	转座酶诸如Piggybac或睡美人
转座酶介导的gRNA抑制文库整合	gRNA，通过转染引入	随机	识别位点	转座酶诸如Piggybac或睡美人
转座酶介导的插入文库整合	ORF，通过转染引入	随机	识别位点	转座酶诸如Piggybac或睡美人

[0488] 昆虫细胞

[0489] 本公开提供了使用本文所述的方法鉴定和/或产生的昆虫细胞。在一些实施方案中，所提供的昆虫细胞表达病毒载体。在一些实施方案中，所提供的昆虫细胞含有一个或多个产生病毒载体所必需的多核苷酸。

[0490] 本公开还提供了包含本文所述的任何病毒载体、构建体和/或组合物的昆虫细胞。在一些实施方案中，本公开提供了昆虫细胞，其包含：(i)标识符，(ii)工程化序列，和(iii)一个或多个产生病毒载体所必需的元件。

[0491] 在一些实施方案中，所提供的昆虫细胞包含：(i)位于病毒包装序列(例如，病毒重复序列，例如，AAV ITR)之间的标识符，(ii)包含扰动和/或扰动附加序列的工程化序列，和(iii)一个或多个产生病毒载体所必需的元件。在一些实施方案中，所提供的昆虫细胞包含：(i)文库构建体，和(ii)一个或多个产生病毒载体所必需的元件，其中文库构建体包含位于病毒包装序列(例如，病毒重复序列，例如，AAV ITRs)之间的标识符，以及任选的文库变体和/或顺式作用整合序列。

[0492] 在一些实施方案中，所提供的技术包括一种独特的方法，借此病毒载体获取标识符(例如，在病毒载体基因组中，例如，在病毒重复序列之间)。因此，由文库的昆虫细胞表达的病毒载体将各自包含标识符。这使得能够直接表征病毒载体并鉴定产生它的昆虫细胞。

[0493] 在一些实施方案中，昆虫细胞产生包含有效载荷和标识符的病毒载体(例如，在病毒载体核酸中)。在一些实施方案中，病毒载体核酸的测序可以鉴定病毒载体池中标识符

(例如,条形码和/或文库变体)的丰度。不希望受理论的束缚,标识符的丰度可以将工程化序列与表型变化诸如病毒载体表达的变化联系起来。例如,病毒载体池中的丰富标识符可以鉴定包含可以具有改进的病毒载体产生的文库变体(和所得扰动)(例如,对应于所述标识符)的细胞。

[0494] 在一些实施方案中,昆虫细胞包含:(i)位于两个能够包装到病毒载体中的病毒重复序列之间的标识符,(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中昆虫细胞产生包含至少一个标识符的病毒载体。在一些实施方案中,标识符包括独特的文库变体(例如,引导序列、ORF等)。在一些实施方案中,标识符包括条形码并且昆虫细胞还包含一个或多个文库变体。在一些实施方案中,昆虫细胞包含文库构建体和一个或多个产生病毒载体所必需的核酸序列。

[0495] 在一些实施方案中,所提供的昆虫细胞是通过向昆虫细胞中引入(i)文库构建体和(ii)一个或多个产生病毒载体所必需的核酸序列而产生的。在一些实施方案中,所提供的昆虫细胞是通过用作为杆状病毒构建体的文库构建体转染来产生的。

[0496] 在一些实施方案中,本公开提供了昆虫细胞,其已被工程化以包含使用本文描述的方法鉴定的扰动。在一些实施方案中,昆虫细胞包含使用本公开的方法鉴定的一个、两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个扰动。在一些实施方案中,昆虫细胞包含一个或多个扰动(例如,其影响病毒载体产生)和产生病毒载体所必需的核酸序列。在一些实施方案中,昆虫细胞(i)包含一个或多个扰动并且(ii)产生递送有效载荷的病毒载体。在一些实施方案中,昆虫细胞不包含标识符(例如,标识符已被去除且/或具有鉴定的扰动的昆虫细胞已被工程化)。

[0497] 在一些实施方案中,提供了包含一个或多个工程化序列的昆虫细胞,所述一个或多个工程化序列共同包含:(i)包含位于能够包装到病毒载体中的第一组两个病毒重复序列之间的标识符的文库构建体,和(ii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,其中文库构建体包含标识符。在一些实施方案中,文库构建体还包含至少一个扰动和/或至少一个文库变体。在一些实施方案中,文库构建体还包含至少一个有效载荷。在一些实施方案中,文库构建体还包含至少一个扰动附加序列。在一些实施方案中,文库构建体还包含至少一个反式作用整合序列和/或至少一个顺式作用整合序列。

[0498] 可以在本公开的上下文中工程化或筛选本领域已知的任何合适的昆虫细胞系。用于表达病毒载体的昆虫细胞可以包括本领域已知的任何昆虫细胞类型。代表性昆虫细胞包括但不限于Sf21细胞、家蚕(Bombyx mori)Bm5细胞、草地贪夜蛾Sf9细胞、BTI-TN-5B1-4(High Five或Hi5)细胞、S2细胞、D.Mel2细胞、Se301细胞或其任意的衍生物。在一些实施方案中,昆虫细胞包括Sf9细胞和/或其任何衍生物。

[0499] 在一些实施方案中,昆虫细胞可以支持病毒生命周期。在一些实施方案中,本公开的昆虫细胞是高度可转染的。

[0500] 用于病毒生产的昆虫细胞是本领域中已知的。

[0501] 在一些实施方案中,本公开的昆虫细胞适合于贴壁细胞培养。在一些实施方案中,昆虫细胞在贴壁细胞培养基中培养。在一些实施方案中,昆虫细胞可以在无血清条件下生长。

[0502] 在一些实施方案中,本公开的昆虫细胞适合于悬浮细胞培养。

[0503] 在一些实施方案中,昆虫细胞在悬浮细胞培养基中培养。在一些实施方案中,悬浮细胞培养的昆虫细胞适合于大量培养(例如,≥1L容量、≥2L容量、≥3L容量、≥4L容量、≥5L容量、≥10L容量、≥20L容量、≥30L容量、≥40L容量、≥50L容量、≥60L容量、≥70L容量、≥80L容量、≥90L容量、≥100L容量、≥200L容量、≥300L容量、≥400L容量,或≥500L容量)。

[0504] 在一些实施方案中,本公开的昆虫细胞系适用于制造生物制剂(例如,病毒载体)。在一些实施方案中,昆虫细胞系适用于生物产品的工业规模制造。在一些实施方案中,昆虫细胞系适用于符合当地监管标准(例如,FDA和/或EMA监管标准)的制造方法。在一些实施方案中,昆虫细胞系适用于使用现行良好生产规范(cGMP)制造生物制剂(例如,病毒载体)。在一些实施方案中,昆虫细胞系适用于使用良好生产规范(GMP)制造生物制剂(例如,病毒载体)。在一些实施方案中,昆虫细胞系适用于使用非良好生产规范(non-GMP)制造生物制剂(例如,病毒载体)。

[0505] 扰动

[0506] 本公开提供了包含一个或多个扰动的昆虫细胞、产生的病毒载体、产生病毒载体所必需的多核苷酸和/或其他构建体。这些是从本公开的用于表达和/或产生病毒载体的昆虫细胞文库中生成、产生、鉴定和/或选择的。如本文所用,扰动包括昆虫细胞、产生的病毒载体、产生病毒载体所必需的多核苷酸和/或其他构建体中的遗传修饰,所述遗传修饰由如本文所述的方法产生和/或鉴定。

[0507] 在一些实施方案中,扰动是一个或多个文库变体的结果。在一些实施方案中,扰动是遗传修饰,其不是文库变体的结果,而是从如本文所述的方法产生和/或鉴定的遗传修饰。在一些实施方案中,扰动包含在至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸中的遗传修饰。

[0508] 在一些实施方案中,在昆虫细胞和/或病毒载体群中,昆虫细胞和/或病毒载体各自单独地包含至少一个包含多个独特扰动的工程化序列。在一些实施方案中,在昆虫细胞和/或病毒载体群中,昆虫细胞和/或病毒载体各自单独地包含至少一个包含至少两个、三个、四个、五个、六个、七个、八个或九个独特扰动的工程化序列。

[0509] 在一些实施方案中,扰动包括昆虫细胞、产生的病毒载体、产生病毒载体所必需的多核苷酸和/或其他构建体中的工程化序列,所述工程化序列可以包含基因组序列变化(例如,基因组插入、缺失或敲除、取代(例如,SNP)、替换、重排等)、游离序列变化(例如,插入、缺失或敲除、SNP(取代)、替换、重排等)和/或表观遗传修饰(例如,激活、抑制等)。在一些实施方案中,扰动包括病毒载体中的工程化序列。在一些实施方案中,扰动包括在至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸中的工程化序列。在一些实施方案中,由本文所述的昆虫细胞和/或方法产生的病毒载体包含扰动。

[0510] 在一些实施方案中,扰动包括对内源性编码序列的修饰。在一些实施方案中,内源性编码序列包含内源性基因或基因区段。在一些实施方案中,扰动包括对内源性调控元件的修饰,其中调控元件包含至少一个启动子序列和/或至少一个增强子序列。

[0511] 扰动的类型

[0512] 在一些实施方案中,扰动包括一个或多个基因组和/或游离修饰。在一些实施方案中,扰动包括一个或多个缺失或敲除、SNP(取代)、替换、插入、重排和/或表观遗传修饰(例

如,激活(例如,一个或多个基因的激活)和/或抑制(例如,一个或多个基因的抑制))。

[0513] 在一些实施方案中,扰动包括基因组或游离缺失或敲除。在一些实施方案中,包括基因组或游离缺失或敲除的扰动由为gRNA缺失或敲除文库的一部分的一个或多个文库变体的表达引起。在一些实施方案中,包括基因组或游离缺失或敲除的扰动不由一个或多个文库变体的表达引起。在一些实施方案中,通过任何合适的方式将文库变体引入昆虫细胞。在一些实施方案中,文库变体包含在文库构建体中,所述文库构建体是gRNA缺失或KO文库的一部分。

[0514] 在一些实施方案中,扰动包括基因组或游离SNP。在一些实施方案中,包括基因组或游离SNP的扰动由为gRNA SNP文库的一部分的一个或多个文库变体的表达引起。在一些实施方案中,包括基因组或游离SNP的扰动不由一个或多个文库变体的表达引起。在一些实施方案中,通过任何合适的方式将文库变体引入昆虫细胞。在一些实施方案中,文库变体包含在文库构建体中,所述文库构建体是gRNA SNP文库的一部分。

[0515] 在一些实施方案中,扰动包括基因组或游离激活(例如,一个或多个基因的激活)。在一些实施方案中,包括基因组或游离激活的扰动由为gRNA激活文库的一部分的一个或多个文库变体的表达引起。在一些实施方案中,包括基因组或游离激活的扰动不由一个或多个文库变体的表达引起。在一些实施方案中,通过任何合适的方式将文库变体引入昆虫细胞。在一些实施方案中,文库变体包含在文库构建体中,所述文库构建体是gRNA激活文库的一部分。

[0516] 在一些实施方案中,扰动包括基因组或游离抑制(例如,一个或多个基因的抑制)。在一些实施方案中,包括基因组或游离抑制的扰动由为gRNA抑制文库的一部分的一个或多个文库变体的表达引起。在一些实施方案中,包括基因组或游离抑制的扰动不由一个或多个文库变体的表达引起。在一些实施方案中,通过任何合适的方式将文库变体引入昆虫细胞。在一些实施方案中,文库变体包含在文库构建体中,所述文库构建体是gRNA抑制文库的一部分。

[0517] 在一些实施方案中,扰动包括基因组或游离插入(例如,一个或多个基因的插入)。在一些实施方案中,包括基因组或游离插入的扰动由为gRNA插入文库的一部分的一个或多个文库变体的表达引起。在一些实施方案中,包括基因组或游离插入的扰动不由一个或多个文库变体的表达引起。在一些实施方案中,通过任何合适的方式将文库变体引入昆虫细胞。在一些实施方案中,文库变体包含在文库构建体中,所述文库构建体是插入文库的一部分。

[0518] 在一些实施方案中,可以使用基因组编辑将包含一个或多个工程化序列的一个或多个扰动引入昆虫细胞文库的昆虫细胞中。下表7中提供了一些示例性类别的扰动和相关文库设想以及扰动附加序列。

[0519] 表7-昆虫细胞中的示例性扰动

[0520]

扰动	设想	扰动附加序列
基因组KO	病毒介导的gRNA KO文库整合 CRISPR介导的gRNA KO文库整合 重组酶介导的gRNA KO文库整合 转座子介导的gRNA KO文库整合 杆状病毒表达的gRNA KO文库 游离gRNA KO文库	Cas
基因组SNP	病毒介导的gRNA SNP文库整合 CRISPR介导的gRNA SNP文库整合 重组酶介导的gRNA SNP文库整合 转座子介导的gRNA SNP文库整合 杆状病毒表达的gRNA SNP文库 游离gRNA SNP文库	Cas
基因组激活	病毒介导的gRNA激活文库整合 CRISPR介导的gRNA激活文库整合 重组酶介导的gRNA激活文库整合 转座子介导的gRNA激活文库整合 杆状病毒表达的gRNA激活文库 游离gRNA激活文库	Cas
基因组抑制	病毒介导的gRNA抑制文库整合 CRISPR介导的gRNA抑制文库整合 重组酶介导的gRNA抑制文库整合 转座子介导的gRNA抑制文库整合 杆状病毒表达的gRNA抑制文库 游离gRNA抑制文库	Cas
基因组插入	病毒介导的插入文库整合 CRISPR介导的插入文库整合 重组酶介导的插入文库整合 转座子介导的插入文库整合	N/A
游离插入	游离插入文库(病毒重复侧翼序列包含在闭合环状质粒上) 杆状病毒表达的插入文库	N/A

[0521] 扰动附加序列

[0522] 在一些实施方案中,所提供的昆虫细胞包含扰动附加序列,其有助于与文库构建体组合而产生扰动。例如,在一些实施方案中,昆虫细胞包含(i)文库构建体,其包含一个或多个包含gRNA的文库变体,(ii)扰动附加序列,其包含编码RNA引导的核酸酶和/或其衍生物和/或融合物的序列,和/或(iii)核酸酶介导的扰动的其他元件。在一些实施方案中,扰动附加序列包含源自Cas9、CasZ、CpfI和/或Fok1的RNA引导的核酸酶。

[0523] 在一些实施方案中,扰动附加序列包含RNA引导的核酸酶,其包含Cas9、Cpf1和/或CasZ或其衍生物,包括包含转录调控因子的融合蛋白(例如,Cas9-VPR或Cas9-KRAB-MeCP2融合物)、CRISPR蛋白与核酸酶结构域(例如,Fok1)的融合物、酶碱基编辑器(例如,BE和ABE

融合物版本)、逆转录酶融合物(例如,引导编辑器)、CRISPR重组酶(例如,RecCas9)和CRISPR转座酶(例如,Tn7样转座酶系统Cas12k和与TniQ的级联复合物)。

[0524] 在一些实施方案中,扰动附加元件包括Cpf1核酸内切酶。在一些实施方案中,Cpf1包括Zetsche等人(2015)Cell 163:759-771中公开的Cpf1多肽以及美国2016/0208243中公开的Cpf1多肽的Cpf1同源物和直系同源物。其他工程化Cpf1变体是本领域普通技术人员已知的并且包括在当前公开的范围内(参见,例如,WO/2017/184768)。

[0525] 在一些实施方案中,所提供的昆虫细胞是通过向昆虫细胞中引入(i)文库构建体和(ii)扰动附加序列,和(iii)一个或多个产生病毒载体所必需的核酸序列而产生的。在一些实施方案中,本公开的表达病毒载体的细胞文库是通过向每个细胞(例如,昆虫细胞)中引入:(i)文库构建体,(ii)扰动附加序列,和(iii)一个或多个产生病毒载体所必需的核酸序列而生成的。

[0526] 在一些实施方案中,所提供的方法包括在昆虫细胞中表达扰动附加序列。在一些实施方案中,所提供的方法包括筛选由昆虫细胞文库产生的病毒载体,其中文库的每个细胞包含:(i)文库构建体,(ii)扰动附加序列,和(iii)一个或多个产生病毒载体所必需的核酸序列。

[0527] 例如,昆虫细胞可以使用基因组编辑以引入一个或多个工程化序列(例如,文库变体)。

[0528] 在一些实施方案中,文库变体和扰动附加序列对应于由天然存在的CRISPR系统调整的组分:引导RNA(作为文库变体)和RNA引导的核酸酶(作为扰动附加元件)。在CRISPR/Cas系统中,引导RNA(gRNA)与核酸内切酶(诸如Cas9核酸内切酶)形成复合物。然后,所述复合物被gRNA引导至DNA靶序列,所述DNA靶序列通常位于靶细胞的基因组中。Cas9或Cas9核酸内切酶是指包含Cas9蛋白或其片段的RNA引导的核酸内切酶(例如,包含Cas9的活性或非活性DNA切割结构域或部分非活性DNA切割结构域的蛋白质(例如,Cas9切口酶),和/或Cas9的gRNA结合结构域)。Cas9识别CRISPR重复序列中的短基序(PAM或原型间隔区相邻基序),以帮助区分自身和非自身。Cas9内切酶和引导RNA(例如,单引导RNA)序列和结构是本领域技术人员所熟知的(参见,例如,Ferretti等人,Proc.Natl.Acad.Sci.U.S.A.98:4658-4663(2001);Deltcheva等人,Nature471:602-607(2011);和Jinek等人,Science 337:816-821(2012))。

[0529] 筛选的昆虫细胞和病毒载体的潜在特征

[0530] 在一些实施方案中,扰动与病毒载体和/或用于病毒载体的表达和/或产生(例如,独立地和/或合成地)的一种或多种特征(例如,所需特征)相关。在一些实施方案中,单个(一个)扰动与病毒载体和/或用于病毒载体的表达和/或产生的一种或多种特征(例如,所需特征)相关。在一些实施方案中,两个或更多个扰动一起与病毒载体和/或用于病毒载体的表达和/或产生的一种或多种特征(例如,所需特征)相关。

[0531] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在应用和/或预期应用中以某种方式改变的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如以它们将核酸转移至细胞的方式改变的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在治疗上改变的病毒载体的产生相关。

[0532] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如以某种方式没有功能性和/或功能性较低的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如以某种方式功能性较高和/或增强的病毒载体的产生相关。

[0533] 在一些实施方案中,至少一个扰动与改变的(例如,增加的)病毒载体效力或感染细胞的能力相关。在一些实施方案中,至少一个扰动与改变的(例如,增加的)转导宿主细胞的能力相关。

[0534] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在将核酸转移至细胞方面没有功能性和/或功能性较低的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在将核酸转移至细胞方面功能性较高和/或增强的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在治疗上没有功能性和/或功能性较低的病毒载体的产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在治疗上功能性较高和/或增强的病毒载体的产生相关。

[0535] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的病毒载体的表达和/或产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与增加的病毒载体的表达和/或产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与增加的病毒载体的分泌相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在当时现行的良好生产规范(cGMP)下改变的(例如,增加的)病毒载体的表达和/或产生相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在良好生产规范(GMP)下改变的(例如,增加的)病毒载体的表达、产生和/或分泌相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如在非良好生产规范(non-GMP)下改变的(例如,增加的)病毒载体的表达、产生和/或分泌相关。

[0536] 在一些实施方案中,昆虫细胞或昆虫细胞群包含至少一个与病毒载体产生增加相关的扰动。在一些实施方案中,与缺乏至少一个扰动的可比较的昆虫细胞相比,包含至少一个扰动的昆虫细胞在病毒载体产生方面具有至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少12倍、至少15倍、至少20倍、至少25倍或至少50倍的增加。

[0537] 在一些实施方案中,与缺乏至少一个扰动的可比较的昆虫细胞相比,包含至少一个扰动的昆虫细胞在病毒载体表达方面具有至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少12倍、至少15倍、至少20倍、至少25倍或至少50倍的增加。

[0538] 在一些实施方案中,病毒载体包含至少一个与增加的病毒载体的产生和/或表达相关的扰动。在一些实施方案中,昆虫细胞以比表达缺乏扰动的病毒载体的可比较的昆虫细胞更高的水平表达包含至少一个扰动的病毒载体。在一些实施方案中,包含至少一个扰动的病毒载体以比缺乏扰动的可比较的病毒载体的表达高至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少12倍、至少15倍、至少20倍、至少25倍或至少50倍的水平产生和/或表达。

[0539] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的(例如,增加的)病毒载体的表达和/或产生的持续时间相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的(例如,增加的)昆虫细胞的活力相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的(例如,增加的)昆虫细胞的稳定性(例如,基因组稳定性)相关。

[0540] 在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的(例如,增加的)病毒载体的稳定性(例如,基因组稳定性)相关。在一些实施方案中,昆虫细胞和/或病毒载体包含至少一个扰动,并且与例如改变的(例如,降低的)产生的空病毒载体的百分比相关。

[0541] 在一些实施方案中,昆虫细胞和/或病毒载体需要一个或多个或两个或更多个扰动来产生任何上述关联、效应或表型。例如,在一些实施方案中,具有一个或多个或两个或更多个扰动的昆虫细胞和/或病毒载体具有改变的病毒载体产生水平(例如,增加的或减少的)。在一些实施方案中,昆虫细胞和/或病毒载体需要两个或更多个扰动,所述扰动在昆虫细胞和/或病毒载体中以合成方式相互作用以产生任何上述关联、效应或表型。例如,在一些实施方案中,两个或更多个扰动在昆虫细胞和/或病毒载体中以合成方式相互作用以导致病毒载体产生水平改变(例如,增加或减少)。

[0542] 在一些实施方案中,任何上述关联、效应或表型相对于参考群进行比较,其中参考群是不包含至少一个扰动的可比较的或标准的昆虫细胞群和/或病毒载体群。

[0543] 细胞工程化平台方法

[0544] 本公开提供了针对病毒载体表达和/或产生的特征和/或其他特征工程化和/或筛选昆虫细胞文库的方法。

[0545] 在一些实施方案中,本公开提供了从昆虫细胞文库产生和/或制造病毒载体的方法,其中文库的每个昆虫细胞单独地包含一个或多个工程化序列,所述一个或多个工程化序列包含(i)标识符和(ii)至少一个表达形成病毒载体所必需的一个或多个元件的核酸序列,并且其中表达的每个病毒载体都包含标识符。

[0546] 在一些实施方案中,本公开提供了筛选昆虫细胞文库的方法,其中文库的每个昆虫细胞单独地包含一个或多个工程化序列,所述一个或多个工程化序列包含(i)标识符和(ii)至少一个表达形成病毒载体所必需的一个或多个元件的核酸序列,并且其中表达的每个病毒载体都包含标识符。在一些实施方案中,所述方法包括检测病毒载体中的标识符的步骤(例如,通过下一代测序和/或单细胞测序)。

[0547] 在一些实施方案中,本公开提供了从昆虫细胞文库产生AAV载体的方法,其中文库的每个昆虫细胞单独地包含一个或多个工程化序列,所述一个或多个工程化序列包含(i)标识符和(ii)至少一个表达形成AAV载体所必需的一个或多个元件的核酸序列,并且其中表达的每个AAV载体都包含标识符。在一些实施方案中,至少一个工程化序列包含文库变体。在一些实施方案中,至少一个工程化序列包含文库变体和条形码。在一些实施方案中,标识符包含条形码。在一些实施方案中,文库变体在昆虫细胞和/或病毒载体中产生至少一个扰动(例如,改变AAV载体产生的一个方面的扰动)。

[0548] 在一些实施方案中,本公开提供了筛选昆虫细胞文库与AAV载体产生相关的特征的方法,其中文库的每个昆虫细胞单独地包含一个或多个工程化序列,所述一个或多个工

程化序列包含(i)标识符和(ii)至少一个表达形成AAV载体所必需的一个或多个元件的核酸序列,并且其中表达的每个AAV载体都包含标识符。在一些实施方案中,所述方法包括检测AAV载体中的标识符的步骤(例如,通过下一代测序)。在一些实施方案中,特定标识符的相对丰度是相对于AAV载体池中的所有标识符确定的。

[0549] 在一些实施方案中,至少一个工程化序列还包含有效载荷、报告基因、选择标记、扰动附加序列、反式作用整合序列和/或顺式作用整合序列。

[0550] 在一些实施方案中,本公开提供了包括以下的方法:(i)表达和/或产生病毒载体的昆虫细胞文库,其中文库的每个昆虫细胞包含(a)文库构建体,其包含位于病毒包装序列之间的标识符,和(b)一个或多个产生病毒载体所必需的多核苷酸;(ii)培养文库的昆虫细胞以产生包含标识符的病毒载体,以及(iii)检测病毒载体池中的标识符。

[0551] 在一些实施方案中,本公开提供了包括以下的方法:(i)表达和/或产生AAV载体的昆虫细胞文库,其中文库的每个昆虫细胞包含(a)文库构建体,其包含位于AAV ITR序列之间的标识符,和(b)一个或多个产生AAV载体所必需的多核苷酸;(ii)培养文库的昆虫细胞以产生包含标识符的AAV载体,以及(iii)检测AAV载体池中的标识符。

[0552] 图2提供了使用用于表达具有标识符的AAV病毒载体的示例性单个连续文库构建体来工程化表达和/或产生病毒载体的昆虫细胞文库的示例性平台方法的示意图。昆虫细胞被工程化以包含(i)单个连续文库构建体和(ii)具有AAV表达所必需的序列(例如,辅助基因和rep/cap)的AAV构建体,描绘为图2,步骤B。在一些实施方案中,示例性示意图利用整合构建体将单个文库构建体整合到文库的昆虫细胞中。在一些实施方案中,示例性示意图在文库的昆虫细胞中游离表达单个文库构建体。

[0553] 在一些实施方案中,文库构建体(例如,用于产生AAV的昆虫细胞文库)可以通过其他方式(例如,核酸酶介导的整合、重组酶介导的整合、转座酶介导的整合等)被整合到昆虫细胞中或在昆虫细胞中游离表达。在各种实施方案中,昆虫细胞文库的昆虫细胞应包含一致且低拷贝数的文库构建体(例如,单个拷贝)。图2,步骤C描绘了来自所得昆虫细胞文库的AAV载体的表达,其可以根据本文提供的方法进行筛选。

[0554] 图3提供了使用用于表达和/或产生具有标识符的AAV病毒载体的示例性不连续文库构建体来工程化表达和/或产生病毒载体的昆虫细胞文库的示例性平台方法的示意图。在一些实施方案中,不连续文库构建体被提供为一起构成文库构建体的一系列构建体。在一些实施方案中,不连续文库构建体包括标识符(例如,条形码)和一个或多个其他构建体,所述一个或多个其他构建体包含一个或多个文库变体,如图3,步骤A所描绘。在一些实施方案中,这些其他构建体中的每一个也含有一个或多个条形码。在一些实施方案中,文库构建体包含用于(例如,核酸酶介导的整合的同源臂、重组酶介导的整合的重组位点、转座酶介导的整合的转座酶位点等)进行基因组整合的顺式作用整合序列或者可以用于昆虫细胞中的游离表达。

[0555] 昆虫细胞被工程化以包含AAV表达所必需的序列(例如,辅助基因和rep/cap),描绘为图3,步骤C。在采用不连续文库构建体的方法中,对文库的昆虫细胞的至少一部分和/或AAV载体的至少一部分进行测序。图3,步骤D描绘了昆虫细胞的单细胞测序,并且图3,步骤E描绘了AAV载体中标识符的测序。值得注意的是,这些测序步骤可以以任何顺序进行。例如,在一些实施方案中,可以首先对AAV载体进行测序以确定与昆虫细胞和/或病毒载体的

所需特征相关的标识符(例如,用于病毒载体产生),然后可以通过单细胞测序对细胞进行测序以将特定标识符(例如,先前鉴定的标识符)与其潜在的致病(causative)文库变体相关联。在一些实施方案中,只有那些与选定的标识符相关的昆虫细胞可以通过单细胞测序进行测序。在一些实施方案中,AAV载体的测序和昆虫细胞的单细胞测序基本上同时进行。

[0556] 本领域技术人员将认识到,虽然图2和图3中的示意图描绘了筛选昆虫细胞文库的病毒载体表达和/或产生的特征,但步骤的顺序可以适当调整。

[0557] 在一些实施方案中,由第一昆虫细胞文库生成的病毒载体池用于转导昆虫细胞以生成包含文库变体的第二昆虫细胞文库。在一些实施方案中,第一昆虫细胞文库包含整合的文库构建体。在一些实施方案中,第一昆虫细胞文库包含游离文库构建体。在一些实施方案中,第二昆虫细胞文库包含整合的文库构建体。在一些实施方案中,第二昆虫细胞文库包含游离文库构建体。

[0558] 在一些实施方案中,病毒载体是从第一昆虫细胞文库中收获和/或合并的。在一些实施方案中,病毒载体是AAV病毒载体并且第一昆虫细胞文库是使用AAV-in-Transposase文库、AAV-in-Baculovirus文库和/或游离AAV文库生成的。在一些实施方案中,从第一昆虫细胞文库收获和/或合并的病毒载体用于转染昆虫细胞并生成第二昆虫细胞文库。在一些实施方案中,第二昆虫细胞文库以游离形式表达文库变体。在一些实施方案中,由第一昆虫细胞文库生成的AAV载体池用于转导昆虫细胞以生成以游离形式包含标识符的第二昆虫细胞文库。

[0559] 病毒载体的收获和合并

[0560] 使用本领域已知的任何适当的方法培养表达病毒载体的昆虫细胞文库,并收获产生的病毒载体。

[0561] 可以通过分批培养、分批补料培养或连续培养来培养表达病毒载体的昆虫细胞。表达病毒载体的昆虫细胞可以悬浮培养或附着于摇瓶、发酵罐或生物反应器中的固体载剂。培养后,可以收获昆虫细胞和/或上清液,并且可以使用本领域已知的方法从适当的级分中分离并纯化核酸。

[0562] 在一些实施方案中,病毒载体是从昆虫细胞文库中收获的。在一些实施方案中,在昆虫细胞表达足够的时间后收获病毒载体,这可以根据昆虫细胞类型和培养条件而变化。

[0563] 在一些实施方案中,收获由表达病毒载体的昆虫细胞文库的昆虫细胞产生的总病毒载体。在一些实施方案中,对应于时间间隔收获表达病毒载体的细胞文库的昆虫细胞产生的病毒载体。例如,可以每天、每两天、每3天或间隔更长的时间收获病毒载体,以评估随时间变化的病毒载体产生。

[0564] 在一些实施方案中,当昆虫细胞达到特定范围内的细胞密度时收获病毒载体。在一些实施方案中,病毒载体在特定时间后收获。在一些实施方案中,病毒载体在病毒转染后的12小时和2周之间收获。在一些实施方案中,病毒载体在病毒转染后的24和144小时之间收获。在一些实施方案中,病毒载体是从细胞培养基中收获的。在一些实施方案中,昆虫细胞在收获病毒载体的过程中被裂解。在一些实施方案中,当昆虫细胞产生至少阈值水平的病毒载体(例如,在纯化之前每个昆虫细胞平均至少约$1x\ 10^3$个病毒载体)时收获病毒载体。

[0565] 在一些实施方案中,可以在长时间后洗涤昆虫细胞并收获病毒载体(例如,以评估

病毒载体的持续产生)。在一些实施方案中,分离来自病毒载体(例如,其包含标识符)的病毒载体核酸。

[0566] 在一些实施方案中,昆虫细胞在收获前被合并。在一些实施方案中,病毒载体在病毒载体核酸(例如,DNA和/或RNA)测序之前合并。

[0567] 病毒载体测序

[0568] 在一些实施方案中,所提供的方法和技术包括病毒载体核酸的测序。在一些实施方案中,病毒载体核酸在测序前量化。在一些实施方案中,病毒载体核酸在测序前不量化。可以使用本领域的任何合适的测序方法。图5提供了示例性测序方法的示意图。

[0569] 在一些实施方案中,确定纯化后的病毒载体滴度。在一些实施方案中,使用定量PCR确定滴度。在某些实施方案中,使用对构建体具有特异性的TaqMan探针来确定构建体水平。在各种实施方案中,所提供的方法和技术包括扩增步骤,其中病毒载体核酸物质(或其一部分,例如,标识符)被扩增。虽然任何适用于应用的扩增反应都被认为与一些实施方案相容,但作为具体实例,在一些实施方案中,扩增步骤可以是或包括聚合酶链式反应(PCR)、滚环扩增(RCA)、多重置换扩增(MDA)、等温扩增及其任何组合。

[0570] 在一些实施方案中,待在本方法的上下文中使用的测序包括下一代测序。有许多使用不同的测序技术的不同的NGS平台。通常,NGS平台对数百万个DNA小片段进行并行测序。在一些实施方案中,NGS是或包括Solexa测序,其同时鉴定DNA碱基,因为每个碱基发射独特的荧光信号,并将它们添加到核酸链中。在一些实施方案中,NGS是或包括454测序,其在核苷酸通过聚合酶并入新的DNA链后再次使用荧光检测焦磷酸盐释放。在一些实施方案中,NGS是或包括离子流(ion torrent):质子/PGM测序,其测量来自单独的碱基通过DNA聚合酶并入的H+(质子)的直接释放。

[0571] 可以分析病毒载体核酸池中标识符的丰度。标识符序列可以用于选择促进所需病毒载体特征(例如,高产量)的昆虫细胞。例如,与参考细胞相比,增加的病毒载体产生可以是固定时间段内病毒载体数量的增加或延长的时间量的产生。在一些实施方案中,可以分析在不同连接点处的病毒载体池样品(例如,以评估在之后时间点产生病毒载体的昆虫细胞系)。使用病毒载体中的标识符,可以鉴定对应的昆虫细胞并确定所述细胞中的工程化序列(例如,文库变体)。例如,这可以通过文库变体与标识符的直接和预先确定的关联来完成(例如,在具有连续文库构建体的情况下),或者通过单细胞测序的其他步骤,然后将标识符与潜在的致病文库变体相关联来完成(例如,在具有不连续文库构建体的情况下)。

[0572] 单细胞测序

[0573] 在一些实施方案中,本公开的方法包括单细胞测序的步骤。在一些实施方案中,其中昆虫细胞包含文库构建体(例如,不连续文库构建体),所提供的方法包括单细胞测序步骤(例如,以鉴定昆虫细胞中未包装到病毒载体中的一个或多个工程化序列)。在一些实施方案中,应当理解单细胞测序步骤也可以用于使用连续文库构建体的情况。

[0574] 在一些实施方案中,其中文库构建体包含多个不连续构建体,所提供的方法包括单细胞测序步骤。在一些实施方案中,待通过单细胞测序进行测序的核酸(例如,构建体的)在昆虫细胞中表达为RNA。在一些实施方案中,所提供的包括单细胞测序的方法包括用细胞身份序列标记一个或多个表达的序列(例如,RNA,例如,mRNA)的步骤。在一些实施方案中,所有表达的序列(例如,RNA,例如,mRNA)都用细胞身份序列标记。在一些实施方案中,文库

的昆虫细胞将表达包含存在于细胞中的任何标识符和/或文库变体的poly-A尾mRNA。

[0575] 在一些实施方案中,用细胞身份序列标记一个或多个文库变体。在一些实施方案中,用细胞身份序列标记所有文库变体。在一些实施方案中,用细胞身份序列标记所有文库变体和所有标识符。在一些实施方案中,昆虫细胞文库的每个昆虫细胞或细胞系包含独特的细胞身份序列。因此,细胞身份序列使来自包含在连续或不连续文库构建体(将包含已测序的核酸)上的构建体的核酸与其来源的单独的昆虫细胞或细胞系相关联。

[0576] 本公开还涵盖以下认识:在单细胞测序方法过程中,在表达的RNA转化为cDNA的逆转录过程中,特异性附加单细胞身份序列。应当理解,旨在用于单细胞测序的构建体或核酸应包含在表达的RNA中,使得在逆转录步骤过程中可以使用适当的引物用细胞身份序列对所有转录物进行单细胞标记。

[0577] 在一些实施方案中,所提供的方法包括昆虫细胞文库的昆虫细胞的单细胞测序步骤和病毒载体测序步骤(例如,使用下一代测序)。本公开涵盖以下认识:通过对病毒载体和昆虫细胞进行测序,可以确定昆虫细胞中病毒载体丰度和任何文库变体之间的关联。

[0578] 在一些实施方案中,所提供的方法包括病毒载体测序步骤(例如,使用下一代测序),然后是单细胞测序步骤。在一些实施方案中,对基于病毒载体测序选择的那些昆虫细胞系进行昆虫细胞文库的昆虫细胞的单细胞测序。

[0579] 在一些实施方案中,昆虫细胞文库的昆虫细胞的单细胞测序与病毒载体测序(例如,使用下一代测序)同时进行。在一些实施方案中,在病毒载体测序之前进行昆虫细胞文库的昆虫细胞的单细胞测序。

[0580] 分析

[0581] 在一些实施方案中,本公开涵盖以下认识:病毒池中特定标识符的丰度可以用于鉴定具有改进的特征(例如,病毒载体特征或病毒载体产生特征(例如,高表达和/或产生))的昆虫细胞(在文库中的细胞中)。可以确定昆虫细胞中对应的工程化序列(例如,文库变体)。例如,这可以通过文库变体与经测序的标识符的直接和预先确定的关联来完成(例如,在具有连续文库构建体的情况下),或者通过单细胞测序的其他步骤,然后将标识符与潜在的致病文库变体相关联来完成(例如,在具有不连续文库构建体的情况下)。

[0582] 在一些实施方案中,所提供的方法和技术包括在病毒载体测序之前进行选择或筛选的步骤。例如,可以针对病毒载体的功能性特征选择或筛选由昆虫细胞文库产生的病毒载体,所述特征诸如像病毒载体稳定性、病毒载体效力、病毒载体感染细胞的能力、病毒载体结合(例如,与受体结合)、病毒载体转移核酸的能力等。在一些实施方案中,将此类选定的病毒载体合并并对其进行测序。使用此类方法,可以鉴定具有多种有益特征的昆虫细胞和/或扰动。例如,产生高水平的病毒载体且稳定的细胞系。

[0583] 在一些实施方案中,所提供的方法和技术包括在病毒载体测序后进行选择或筛选步骤。例如,选择或筛选的候选病毒载体或鉴定的扰动(例如,遗传变化)可以用于提供病毒载体文库构建的信息,所述病毒载体文库可以针对各种特征进行分析。例如,可以针对它们转导昆虫细胞的能力来选择或筛选此类病毒载体文库。

[0584] 选择的昆虫细胞候选者可以用于产生病毒载体,或鉴定的扰动(例如,遗传变化)可以用于提供新昆虫细胞文库构建的信息。基于文库的平台方法(描绘于图1,图A至F)可以被重复直到鉴定出表达具有所需特征和/或所需数量的病毒载体的工程化昆虫细胞。可以

分析与所需特征相关的工程化序列,例如,使用机器学习(ML)方法来开发机器学习模型。经过训练的机器学习模型可以用于为未来的设计提供信息,并减少要筛选的昆虫细胞文库的数量,从而减少时间和成本。在一些实施方案中,可以设计昆虫细胞文库且/或可以执行所述方法以鉴定在昆虫细胞中协同相互作用(例如,两个或更多个工程化序列组合)以具有所需特征(例如,一定水平的病毒载体产生)的工程化序列。在一些实施方案中,获得自本文所述的平台技术的所得昆虫细胞将具有一个、两个、三个、四个、五个或更多个工程化序列(例如,文库变体和/或扰动),使得生成具有所需病毒载体产生特性(例如,一定水平的产生、所需持续时间的产生等)的昆虫细胞。

[0585] 在一些实施方案中,机器学习模型被训练以生成预测,所述预测指示在昆虫细胞和/或病毒载体中具有一个或多个其他扰动的工程化序列(例如,扰动)是否可能具有协同作用和/或进一步改进的病毒载体特征。

[0586] 在各种实施方案中,机器学习模型是回归模型(例如,线性回归、逻辑回归或多项式回归)、决策树、随机森林、支持向量机、朴素贝叶斯模型、k-均值聚类或神经网络(例如,前馈网络、卷积神经网络(CNN)或深度神经网络(DNN))中的任何一种。可以使用机器学习实现方法来训练机器学习模型,所述方法诸如线性回归算法、逻辑回归算法、决策树算法、支持向量机分类、朴素贝叶斯分类、K-最近邻分类、随机森林算法、深度学习算法、梯度提升算法和降维技术中的任何一种。在各种实施方案中,使用监督学习算法、非监督学习算法、半监督学习算法(例如,部分监督)、弱监督、转移、多任务学习或其任何组合来训练机器学习模型。在各种实施方案中,机器学习模型包括在机器学习模型的训练过程中调整的参数。例如,调整参数以最小化损失函数,从而提高机器学习模型的预测能力。

[0587] 通常,训练机器学习模型以区分导致病毒载体表达变化的一个或多个编辑。例如,训练机器学习模型以识别训练实例中有助于增加或减少病毒载体表达的模式。作为一个具体实例,训练机器学习模型以鉴定特定基因组位置,如果编辑所述位置,可能会导致昆虫细胞增加和/或延长病毒载体产生。作为另一个具体实例,可以训练机器学习模型以鉴定特定基因组位置,如果编辑所述位置,会产生具有增加和/或延长的病毒载体产生的昆虫细胞。

[0588] 在各种实施方案中,使用机器学习模型输出的预测分数对鉴定的编辑进行归类。作为一个实例,被分配了高于阈值的分数的鉴定的编辑被归类为用于进一步测试的候选编辑。在各种实施方案中,阈值分数是0.5、0.6、0.7、0.75、0.8、0.85、0.90、0.91、0.92、0.93、0.94、0.95、0.96、0.97、0.98或0.99。不满足阈值分数标准的鉴定的编辑被归类为非候选编辑。

[0589] 总而言之,机器学习模型的应用使可以快速筛选出的编辑的进行于电脑中的预测和归类成为可能。因此,在基因组设计中仅使用候选编辑进行进一步测试,而将非候选编辑从进一步考虑中去除。这消除了非常耗时且成本高昂的体外测试所有编辑组合的需要。

[0590] 应用

[0591] 本公开尤其提供了用于表达病毒载体的昆虫细胞、用于生成昆虫细胞的构建体和使用本文所述的方法生成的病毒载体。病毒载体和相关的昆虫细胞可以用于许多应用,包括但不限于疫苗、癌症治疗(例如,溶瘤治疗)和/或基因治疗。病毒载体和昆虫细胞可以用于生成生物制剂和/或治疗的研究和制造过程,或用作生物制剂本身。

[0592] 例如,在一些实施方案中,病毒载体可以以多种方式使用,包括但不限于疫苗、癌

症治疗(例如,溶瘤治疗)和/或基因治疗(例如,体内基因和/或基因组编辑)。作为另一个实例,在一些实施方案中,病毒载体可以以多种方式使用,包括但不限于研究、生产和/或制造:疫苗、癌症治疗(例如,溶瘤治疗)、基因治疗(例如,离体基因和/或基因组编辑)和/或细胞治疗(例如,离体基因和/或基因组编辑)。因此,有大量病毒载体可用于这些不同的应用。

[0593] 在一些实施方案中,提供了包含如本文所述的文库构建体、病毒载体和/或昆虫细胞的组合物。在一些实施方案中,提供了本文所述的方法产生昆虫细胞的用途。在一些实施方案中,提供了本文所述的方法产生病毒载体(例如,AAV载体)的用途。在一些实施方案中,提供了本文所述的方法产生文库构建体的用途。

[0594] 在一些实施方案中,提供了本文所述的文库构建体、病毒载体和/或方法产生工程化昆虫细胞的用途。在一些实施方案中,提供了一种制造表达病毒载体的昆虫细胞的方法,所述方法包括引入一个或多个使用本文所述的筛选方法鉴定的扰动。

[0595] 本公开还提供了用使用本文所述的病毒载体和/或昆虫细胞的组合物(例如,药物组合物)治疗受试者的方法。

[0596] 昆虫细胞和病毒载体的示例性应用

[0597] 在一些实施方案中,本公开的昆虫细胞产生所需水平的病毒载体。在一些实施方案中,本公开的昆虫细胞包含一个或多个影响病毒载体产生的扰动。

[0598] 在一些实施方案中,提供了一种制造病毒载体的方法,其包括培养本文所述的昆虫细胞。在一些实施方案中,提供了一种制造靶水平的病毒载体的方法,其包括培养本文所述的昆虫细胞。在一些实施方案中,提供了昆虫细胞产生高于阈值水平的水平的病毒载体(例如,AAV载体)的用途。在一些实施方案中,用于产生病毒载体的昆虫细胞包含一个多个扰动并且以比缺乏一个或多个扰动的对应昆虫细胞更高的水平产生病毒载体。

[0599] 在一些实施方案中,提供了用于在昆虫细胞中产生病毒载体的方法,然后所述病毒载体可以用于生成细胞治疗。

[0600] 在一些实施方案中,本公开提供了一种制造和/或生产疫苗的方法,其包括培养本公开的昆虫细胞,其中病毒载体包含有包含疫苗组分的有效载荷。在一些实施方案中,提供了一种制造和/或生产基于AAV的疫苗的方法。

[0601] 在一些实施方案中,本公开提供了一种制造和/或生产溶瘤病毒载体的方法,其包括培养本公开的昆虫细胞,其中病毒载体是溶瘤病毒载体。

[0602] 在一些实施方案中,提供了用于在昆虫细胞中产生病毒载体的方法,所述病毒载体可以作为基因治疗。在一些实施方案中,产生的病毒载体具有允许体内基因治疗的有效载荷,其中生成的病毒载体可以施用于受试者。

[0603] 在一些实施方案中,提供了用于在昆虫细胞中产生病毒载体的方法,然后所述病毒载体可以用于离体基因治疗。在一些实施方案中,产生的病毒载体具有允许离体基因治疗的有效载荷,其中生成的病毒载体可以用于生成治疗细胞,其然后可以施用于受试者。

[0604] 提供以下实施例以便向熟练技术人员描述如何制备并使用本文所述的方法和组合物,并且不旨在限制本公开的范围。

[0605] 示例性实施方案

[0606] 实施方案1.一种昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含一个或多个工程化核酸序列,所述一个或多个工程化核酸序列共同包含:

[0607] a.至少一个位于能够包装到病毒载体中的一组两个病毒重复序列之间的标识符,和

[0608] b.至少一个包含一个或多个产生所述病毒载体所必需的核酸序列的多核苷酸,并且

[0609] 其中所述昆虫细胞产生包含所述至少一个标识符的病毒载体。

[0610] 实施方案2.一种昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含一个或多个工程化核酸序列,所述一个或多个工程化核酸序列共同包含:

[0611] a.至少一个位于能够包装到病毒载体中的一组两个病毒重复序列之间的标识符,和

[0612] b.至少一个包含一个或多个产生所述病毒载体所必需的核酸序列的多核苷酸,和

[0613] c.至少一个导致至少一个扰动的文库变体,

[0614] 其中所述昆虫细胞产生包含所述至少一个标识符的病毒载体。

[0615] 实施方案3.一种昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含一个或多个工程化核酸序列,所述一个或多个工程化核酸序列共同包含:

[0616] a.至少一个位于能够包装到病毒载体中的一组两个病毒重复序列之间的标识符,和

[0617] b.至少一个包含一个或多个产生所述病毒载体所必需的核酸序列的多核苷酸,和

[0618] c.至少一个导致至少一个扰动的文库变体,和

[0619] d.以下各者中的一个或多个:

[0620] (i)至少一个扰动附加序列,

[0621] (ii)至少一个反式作用整合序列,

[0622] (iii)至少一个顺式作用整合序列,并且

[0623] 其中所述昆虫细胞产生包含所述至少一个标识符的病毒载体。

[0624] 实施方案4.如实施方案1至3中任一项所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包括Sf21细胞、Sf9细胞、Hi5细胞、S2细胞、D.Mel2细胞或其任意的衍生物。

[0625] 实施方案5.如实施方案1至4中任一项所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包括悬浮细胞和/或贴壁细胞。

[0626] 实施方案6.如实施方案1至5中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体是腺相关病毒(AAV)载体、慢病毒载体、腺病毒载体、α病毒载体、辛德毕斯病毒载体、逆转录病毒载体(例如,γ逆转录病毒载体)、多瘤病毒载体(例如,猴病毒40(SV40)载体)、乳头瘤病毒载体(例如,牛乳头瘤病毒(BPV)载体)、痘苗病毒载体、单纯疱疹病毒(HSV)载体、麻疹病毒载体、棒状病毒载体、狂犬病病毒载体、水疱性口炎病毒(VSV)载体、小核糖核酸病毒载体(例如,脊髓灰质炎病毒载体)、呼肠孤病毒载体、塞内卡病毒载体、艾柯病毒载体(例如,RIGVIR)、塞姆利基森林病毒(SFV)载体、黄病毒载体、指环病毒载体(https://www.ringtx.com)、新城疫病毒(NDV)载体、副粘病毒载体、仙台病毒载体、正粘病毒载体、流感病毒载体、冠状病毒载体和/或杂交病毒载体,和/或其衍生物、杂交体和/或工程化衍生物。

[0627] 实施方案7.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中所述病

毒载体是腺相关病毒(AAV)载体。

[0628] 实施方案8.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中所述一组两个病毒重复序列各自是能够包装到AAV载体中的AAV ITR序列。

[0629] 实施方案9.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

[0630] (i)包含位于两个功能性AAV ITR序列之间的标识符的工程化核酸序列,其中所述核酸序列游离存在于所述昆虫细胞内,

[0631] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

[0632] (iii)RNA引导的核酸酶,和

[0633] (iv)一个或多个产生AAV载体所必需的核酸序列,

[0634] 其中所述昆虫细胞群产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的条形码的条形码。

[0635] 实施方案10.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

[0636] (i)工程化核酸序列,其包含位于两个功能性AAV ITR序列之间的标识符,其中所述核酸序列被整合到位于一对顺式作用整合序列之间的昆虫基因组中,

[0637] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

[0638] (iii)RNA引导的核酸酶,和

[0639] (iv)一个或多个产生AAV载体所必需的核酸序列,

[0640] 其中所述昆虫细胞群产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的条形码的条形码。

[0641] 实施方案11.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中AAV载体包含人AAV1衣壳蛋白;人AAV2衣壳蛋白;人AAV3b衣壳蛋白;人AAV4衣壳蛋白;人AAV5衣壳蛋白;人AAV6衣壳蛋白;人AAV7衣壳蛋白;人AAV8衣壳蛋白;人AAV9衣壳蛋白;人AAV10衣壳蛋白;人AAV11衣壳蛋白;人AAV12衣壳蛋白;或人AAV13衣壳蛋白。

[0642] 实施方案12.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中AAV载体包含人祖先AAV衣壳蛋白。

[0643] 实施方案13.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体包括AAV载体,其中AAV载体包含第一组两个病毒重复序列,所述第一组两个病毒重复序列包含一对反向末端重复序列(ITR),其是或包括人AAV1 ITR;人AAV2 ITR;人AAV3b ITR;人AAV4 ITR;人AAV5 ITR;人AAV6 ITR;人AAV7 ITR;人AAV8ITR;人AAV9 ITR;人AAV10 ITR;人AAV11 ITR;人AAV12 ITR;或人AAV13 ITR。

[0644] 实施方案14.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中AAV载体包含牛AAV(b-AAV)衣壳蛋白;犬AAV(CAAV)衣壳蛋白;小鼠AAV1衣壳蛋白;山羊AAV衣壳蛋白;大鼠AAV衣壳蛋白;或禽AAV(AAAV)衣壳蛋白。

[0645] 实施方案15.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体包括AAV载体,其中AAV载体包含一对ITR,其是或包括牛AAV(b-AAV)ITR;犬AAV(CAAV)

ITR;小鼠AAV1 ITR;山羊AAV ITR;大鼠AAV ITR;或禽AAV(AAAV)ITR。

[0646] 实施方案16.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中至少一个包含一个或多个形成病毒载体所必需的核酸序列的多核苷酸包含:

[0647] a.AAV Rep基因;

[0648] b.AAV Cap基因;

[0649] c.一种或多种AAV辅助基因;或

[0650] d.其组合。

[0651] 实施方案17.如实施方案2至16中任一项所述的昆虫细胞群,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞群,所述一个或多个扰动与AAV产生和/或AAV分泌的增加相关。

[0652] 实施方案18.如实施方案17所述的昆虫细胞群,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞,包含所述一个或多个扰动的昆虫细胞具有至少10%的AAV产生和/或AAV分泌增加。

[0653] 实施方案19.如实施方案2至18中任一项所述的昆虫细胞或昆虫细胞群,其与缺乏所述一个或多个扰动的昆虫细胞产生的AAV群相比产生了包含至少一个改进的特征的AAV载体群。

[0654] 实施方案20.如实施方案1至6中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体是慢病毒载体。

[0655] 实施方案21.如实施方案20所述的昆虫细胞或昆虫细胞群,其中所述一组两个病毒重复序列各自是能够包装到慢病毒载体中的慢病毒LTR序列。

[0656] 实施方案22.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

[0657] (i)包含位于两个功能性慢病毒LTR序列之间的标识符的工程化核酸序列,其中所述核酸序列游离存在于所述昆虫细胞内,

[0658] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

[0659] (iii)RNA引导的核酸酶,和

[0660] (iv)一个或多个产生慢病毒载体所必需的核酸序列,

[0661] 其中所述昆虫细胞群产生多个慢病毒载体,其中每个慢病毒载体包含对应于产生它的所述昆虫细胞的条形码的条形码。

[0662] 实施方案23.一种包含多个昆虫细胞的昆虫细胞或昆虫细胞群,其中每个昆虫细胞包含:

[0663] (i)工程化核酸序列,其包含位于两个功能性慢病毒LTR序列之间的标识符,其中所述核酸序列被整合到位于一对顺式作用整合序列之间的昆虫基因组中,

[0664] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

[0665] (iii)RNA引导的核酸酶,和

[0666] (iv)一个或多个产生慢病毒载体所必需的核酸序列,

[0667] 其中所述昆虫细胞群产生多个慢病毒载体,其中每个慢病毒载体包含对应于产生

它的所述昆虫细胞的条形码的条形码。

[0668] 实施方案24.如实施方案20至23中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体是人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒载体、维斯纳病毒载体或其衍生物。

[0669] 实施方案25.如实施方案20至24中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体包含慢病毒Psi序列。

[0670] 实施方案26.如实施方案20至25中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体包含gag蛋白或其片段。

[0671] 实施方案27.如实施方案20至26中任一项所述的昆虫细胞或昆虫细胞群,其中gag蛋白包含一个或多个选自基质(MA)、衣壳(CA)和核衣壳(NC)结构域的结构域。

[0672] 实施方案28.如实施方案20至27中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体包含包膜蛋白或其片段。

[0673] 实施方案29.如实施方案20至28中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体是包含源自不同病毒的gag蛋白和包膜蛋白的假型慢病毒载体。

[0674] 实施方案30.如实施方案20至29中任一项所述的昆虫细胞或昆虫细胞群,其中慢病毒载体包含源自人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒载体、维斯纳病毒载体或其衍生物的gag蛋白和/或env蛋白。

[0675] 实施方案31.如实施方案20至30中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体包括慢病毒载体,其中第一组两个病毒重复序列包含源自人免疫缺陷病毒(HIV)载体、猴免疫缺陷病毒(SIV)载体、马传染性贫血病毒载体、猫免疫缺陷病毒载体、维斯纳病毒载体或其衍生物的慢病毒LTR和/或Psi序列。

[0676] 实施方案32.如实施方案20至31中任一项所述的昆虫细胞或昆虫细胞群,其中至少一个包含一个或多个形成病毒载体所必需的核酸序列的多核苷酸包含:

[0677] a.慢病毒gag基因;

[0678] b.慢病毒env基因;

[0679] c.慢病毒pol基因;或

[0680] d.其组合。

[0681] 实施方案33.如实施方案20至32中任一项所述的昆虫细胞群,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞群,所述一个或多个扰动与慢病毒产生和/或慢病毒分泌的增加相关。

[0682] 实施方案34.如实施方案33所述的昆虫细胞群,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞,包含所述一个或多个扰动的昆虫细胞具有至少10%的慢病毒产生和/或慢病毒分泌增加。

[0683] 实施方案35.如实施方案20至32中任一项所述的昆虫细胞或昆虫细胞群,其与缺乏所述一个或多个扰动的昆虫细胞产生的慢病毒群相比产生了包含至少一个改进的特征的慢病毒载体群。

[0684] 实施方案36.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中所述昆虫细胞各自单独地包括Sf21细胞、Sf9细胞、BTI-TN-5B1-4(High Five)细胞、S2细胞、D.Mel2细胞或其任意的衍生物。

[0685] 实施方案37.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中所述病毒载体是有复制能力的病毒载体。

[0686] 实施方案38.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中病毒载体是条件复制的、复制缺乏的、无复制能力的和/或复制缺陷的病毒载体。

[0687] 实施方案39.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中每个单独的昆虫细胞包含至少一个文库构建体,其中至少一个文库构建体包含至少一个工程化序列,所述至少一个工程化序列包含至少一个条形码、至少一个标识符、至少一个文库变体、至少一个有效载荷、至少一个顺式作用整合序列或其组合和/或多个。

[0688] 实施方案40.如实施方案39所述的昆虫细胞或昆虫细胞群,其中所述顺式作用整合序列是重组酶识别位点。

[0689] 实施方案41.如实施方案39或40所述的昆虫细胞或昆虫细胞群,其中至少一个条形码包含约5至约25个核苷酸的序列。

[0690] 实施方案42.如实施方案39至41中任一项所述的昆虫细胞或昆虫细胞群,其中昆虫细胞群包含多个独特的条形码,并且其中多个独特的条形码包含约5至约25个核苷酸的独特序列。

[0691] 实施方案43.如实施方案39至42中任一项所述的昆虫细胞或昆虫细胞群,其中昆虫细胞各自单独地包含一个、两个、三个、四个、五个、六个、七个、八个、九个或十个文库变体。

[0692] 实施方案44.如实施方案39至43中任一项所述的昆虫细胞或昆虫细胞群,其中昆虫细胞各自单独地包含多至100个文库变体。

[0693] 实施方案45.如前述实施方案中任一项所述的昆虫细胞或昆虫细胞群,其中所述工程化核酸序列包含至少一个文库构建体,所述至少一个文库构建体经由转染引入每个单独的细胞中。

[0694] 实施方案46.如实施方案45所述的昆虫细胞或昆虫细胞群,其中所述工程化核酸序列包含至少一个文库构建体,所述至少一个文库构建体经由杆状病毒转导引入每个单独的细胞中。

[0695] 实施方案47.如实施方案45或46所述的昆虫细胞或昆虫细胞群,其中所述文库构建体包含至少一个包含至少一个文库变体的工程化序列。

[0696] 实施方案48.如实施方案45至47中任一项所述的昆虫细胞或昆虫细胞群,其中至少一个文库变体包含至少一个ORF、至少一个基因、至少一个非编码核酸序列和/或至少一个gRNA或其多个。

[0697] 实施方案49.如实施方案45至47中任一项所述的昆虫细胞或昆虫细胞群,其中每个单独的昆虫细胞包含至少一个工程化序列,所述至少一个工程化序列包含至少一对位于所述一组病毒重复序列侧翼的顺式作用整合序列。

[0698] 实施方案50.如实施方案49所述的昆虫细胞或昆虫细胞群,并且其中所述顺式作用整合序列包含同源臂序列。

[0699] 实施方案51.如实施方案49所述的昆虫细胞或昆虫细胞群,其中所述顺式作用整合序列包含重组酶识别位点。

[0700] 实施方案52.如实施方案51所述的昆虫细胞或昆虫细胞群,其中所述重组酶包含

Cre、Flp、Dre、PhiC31和/或Bxb1,或其衍生物。

[0701] 实施方案53.如实施方案49所述的昆虫细胞或昆虫细胞群,其中所述顺式作用整合序列包含转座酶识别位点。

[0702] 实施方案54.如实施方案53所述的昆虫细胞或昆虫细胞群,其中所述转座酶包含Piggybac转座酶、睡美人转座酶和/或Tn5转座酶,或其衍生物。

[0703] 实施方案55.一种由如实施方案7至19中任一项所述的昆虫细胞或昆虫细胞群产生的AAV载体群,其中所述AAV载体群包括至少一个改进的特征,其中至少一个改进的特征包括转移病毒核酸的能力改变、治疗活性改变,和/或无功能AAV群百分比减少,和/或在生产规范下病毒载体的百分比增加,所述病毒载体含有所有和/或必需的核酸序列和/或用于它们的预期应用的其他元件。

[0704] 实施方案56.一种由如实施方案18至25中任一项所述的昆虫细胞或昆虫细胞群产生的慢病毒载体群,其中所述慢病毒载体群包括至少一个改进的特征,其中至少一个改进的特征包括转移病毒核酸的能力改变、治疗活性改变,和/或无功能慢病毒群百分比减少,和/或在生产规范下病毒载体的百分比增加,所述病毒载体含有所有和/或必需的核酸序列和/或用于它们的预期应用的其他元件。

[0705] 实施方案57.如实施方案1至56中任一项所述的昆虫细胞群,其通过将多个包含多个文库构建体的工程化核酸序列引入到多个昆虫细胞中的步骤产生,其中单独的文库构建体包含至少一个位于第一组两个病毒重复序列之间的标识符,其中多个昆虫细胞包含一个或多个产生病毒载体所必需的核酸序列。

[0706] 实施方案58.一种产生病毒载体的方法,其包括:在使得昆虫细胞产生病毒载体的条件下培养如实施方案1至56中任一项所述的昆虫细胞群,并且其中每个产生的病毒载体包含至少一个源自产生病毒载体的昆虫细胞的至少一个标识符的标识符。

[0707] 实施方案59.一种方法,其包括:

[0708] a.从昆虫细胞文库产生病毒载体,其中昆虫细胞文库包含多个昆虫细胞,其中多个中的每个昆虫细胞单独地包含:(i)至少一个工程化核酸序列,(ii)至少一个标识符,和(iii)至少一个包含一个或多个产生病毒载体所必需的核酸序列的多核苷酸,并且其中每个病毒载体包含至少一个源自产生病毒载体的昆虫细胞的至少一个标识符的标识符;

[0709] b.检测所述病毒载体中的所述一个或多个标识符。

[0710] 实施方案60.一种方法,其包括:

[0711] a.从昆虫细胞文库产生多个腺相关病毒(AAV)载体,其中所述昆虫细胞文库包含多个昆虫细胞,其中每个昆虫细胞单独包含:

[0712] (i)包含位于两个功能性AAV ITR序列之间的标识符的核酸序列,

[0713] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,

[0714] (iii)RNA引导的核酸酶,和

[0715] (iv)一个或多个产生AAV载体所必需的核酸序列,

[0716] 其中所述多个昆虫细胞产生多个AAV载体,其中每个AAV载体包含对应于产生它的所述昆虫细胞的标识符的标识符;以及

[0717] b.检测所述多个AAV载体的标识符。

[0718] 实施方案61.一种选择用于产生腺相关病毒(AAV)载体的昆虫细胞的方法,所述方法包括:

[0719] a.用一个或多个产生AAV载体所必需的序列转染昆虫细胞文库,其中所述多个中的每个昆虫细胞单独包含:

[0720] (i)包含位于两个功能性AAV ITR序列之间的标识符的核酸序列,

[0721] (ii)一个或多个文库变体,其导致一个或多个扰动,其中至少一个文库变体包含gRNA,和

[0722] (iii)RNA引导的核酸酶,

[0723] b.从所述昆虫细胞文库产生多个AAV载体,其中所述AAV载体各自包含对应于产生它的所述单独昆虫细胞的标识符;以及

[0724] c.检测所述多个AAV载体的标识符,其中所述条形码的丰度与所述昆虫细胞产生和/或分泌AAV载体相关。

[0725] 实施方案62.如实施方案60或61所述的方法,其中所述一个或多个产生AAV载体所必需的序列是在杆状病毒载体的背景下。

[0726] 实施方案63.如实施方案60或61所述的方法,其中所述标识符是条形码。

[0727] 实施方案64.如实施方案60至63中任一项所述的方法,其还包括确定特定标识符相对于存在于从所述昆虫细胞文库产生的所述多个AAV载体中的所有标识符的相对丰度的步骤。

[0728] 实施方案65.如实施方案60至64中任一项所述的方法,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞,所述一个或多个扰动与AAV产生和/或AAV分泌的增加相关。

[0729] 实施方案66.如实施方案65所述的方法,其中相对于缺乏所述一个或多个扰动的参考昆虫细胞,包含所述一个或多个扰动的所述昆虫细胞具有至少10%的AAV产生和/或AAV分泌增加。

[0730] 实施方案67.如实施方案60至66中任一项所述的方法,其中所述RNA引导的核酸酶是无核酸酶活性的RNA引导的核酸酶。

[0731] 实施方案68.如实施方案60至67中任一项所述的方法,其中至少一个文库变体被整合到位于一对顺式作用整合序列之间的昆虫基因组中。

[0732] 实施方案69.如实施方案60至68中任一项所述的方法,其中所述工程化核酸序列的一个或两个拷贝被整合到所述昆虫细胞基因组中。

[0733] 实施方案70.如实施方案60至67中任一项所述的方法,其中至少一个文库变体游离存在于所述昆虫细胞中。

[0734] 实施方案71.如实施方案60至70中任一项所述的方法,其中所述两个功能性AAV ITR序列包括人AAV1 ITR、人AAV2 ITR、人AAV3b ITR、人AAV4 ITR、人AAV5ITR、人AAV6 ITR、人AAV7 ITR、人AAV8 ITR、人AAV9 ITR、人AAV10 ITR、人AAV11 ITR、人AAV12 ITR或人AAV13 ITR。

[0735] 实施方案72.如实施方案60至70中任一项所述的方法,其中所述两个功能性AAV ITR序列包括牛AAV(b-AAV)ITR、犬AAV(CAAV)ITR、小鼠AAV1 ITR、山羊AAV ITR、大鼠AAV ITR或禽AAV(AAAV)ITR。

[0736] 实施方案73.如实施方案60至72中任一项所述的方法,其中所述一个或多个产生

AAV载体所必需的序列包含(a)AAV Rep基因,(b)AAV Cap基因,(c)一个或多个AAV辅助基因;或(d)其组合。

[0737] 实施方案74.如实施方案73所述的方法,其中所述一个或多个产生AAV载体所必需的序列包含编码人AAV1衣壳蛋白、人AAV2衣壳蛋白、人AAV3b衣壳蛋白、人AAV4衣壳蛋白、人AAV5衣壳蛋白、人AAV6衣壳蛋白、人AAV7衣壳蛋白、人AAV8衣壳蛋白、人AAV9衣壳蛋白、人AAV10衣壳蛋白、人AAV11衣壳蛋白、人AAV12衣壳蛋白,或人AAV13衣壳蛋白的AAV Cap基因。

[0738] 实施方案75.如实施方案60至74中任一项所述的方法,其中所述一个或多个扰动包括内源性基因组编码序列的插入、缺失、取代、替换、表观遗传修饰和/或重排。

[0739] 实施方案76.如实施方案60至75中任一项所述的方法,其中所述一个或多个文库变体包含至少两个文库变体,其中所述至少两个文库变体包含至少一个独特基因、至少一个独特ORF、至少一个独特gRNA序列,和/或至少一个独特非编码核酸,或其组合和/或多个。

[0740] 实施方案77.如实施方案59至76中任一项所述的方法,其中所述方法还包括单细胞测序。

[0741] 实施方案78.一种分离的核酸,其包含构建体,所述构建体包含:

[0742] (i)位于两个功能性腺相关病毒(AAV)ITR序列之间的条形码,

[0743] (ii)选择标记,和

[0744] (iii)一个或多个文库变体。

[0745] 实施方案79.如实施方案78所述的分离的核酸,其中所述至少一个文库变体包含至少一个工程化序列,所述工程化序列包含至少一个基因、至少一个ORF、至少一个gRNA序列、至少一个非编码核酸,或其组合和/或多个。

[0746] 实施方案80.如实施方案78或79所述的分离的核酸,其中所述构建体是杆状病毒构建体。

[0747] 实施例

[0748] 实施例1:使用杆状病毒载体表达的用于筛选AAV产生的文库变体技术

[0749] 本实施例描述了一种产生和筛选昆虫细胞文库的示例性方法,以确定通过文库内的特定细胞扰动(例如,在含有特定文库变体的细胞中)产生的示例性病毒载体(例如,在此实施例中,AAV载体)的水平。具体而言,此实施例描述了一种将从昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法,其中用于创建昆虫细胞文库的文库变体(例如,gRNA和/或ORF)是标识符。图2提供了此纯化、扩增和测序的示意图概览。

[0750] 本实施例描述了一种生成和筛选产生AAV载体的昆虫细胞文库的方法。示例性AAV载体宿主生产细胞系可以是例如Sf9。将合成针对编码蛋白质基因产物(文库变体)的基因组序列或ORF序列的gRNA序列文库,并克隆到两个AAV ITR序列之间的载体(例如,杆状病毒载体)中。在杆状病毒载体的上下文中具有AAV ITR的这种构建体在此将被称为AAV-in-BAC。AAV-in-BAC构建体还可以含有选择标记基因,诸如抗生素抗性(例如,嘌呤霉素抗性)或荧光蛋白(例如,GFP),使含有文库构建体的昆虫细胞的未来选择或鉴定成为可能。然后可以用产生和包装重组AAV病毒载体所必要的质粒构建体(例如,形成病毒载体所必需的多核苷酸)(例如,pHelper和pAAV Rep-Cap)转染AAV-in-BAC文库的纯化细胞群。这些质粒上存在的功能可以指导ITR限定的AAV载体基因组的复制,并且这些ITR病毒载体构建体将被包装并释放到AAV载体中。

[0751] 病毒载体是使用本领域已知的方法产生的。可以通过转导示例性哺乳动物细胞来评估从昆虫细胞纯化的重组AAV病毒载体的功能。

[0752] 可以对纯化的AAV载体内所含的DNA(特别是作为标识符的gRNA和/或ORF编码序列)进行测序。从纯化的AAV载体池(例如,整个AAV载体池、来自选定细胞子集的AAV载体池或选定的AAV载体)中分离DNA。可以使用本领域已知的方法,诸如碱裂解,来纯化DNA。使用PCR自侧翼引物序列扩增DNA,例如像图5所描绘。纯化扩增产物,将其与具有同源悬突的测序衔接子(例如,Illumina衔接子)混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用例如NextSeq平台进行测序。测量每个标识符的序列读段频率,并确定它们在所有标识符池内的相对丰度。宿主细胞产生株文库中含有的标识符序列(例如,gRNA和/或ORF序列)也将被扩增并以同样的方式制备用于测序。从AAV载体相关DNA扩增的标识符序列(例如,gRNA和/或ORF序列)的相对丰度将与原始宿主细胞群中的标识符序列丰度进行比较。通过这种方式,可以鉴定导致扰动的文库变体,所述扰动直接改变宿主细胞生物学并导致差异AAV载体产生,因为与参考宿主细胞群相比,所述gRNA文库变体在AAV载体DNA群中富集或去富集。

[0753] 本文描述的方法还允许迭代轮次的文库筛选,其中在AAV载体群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的AAV载体产生的标识符序列(例如,gRNA和/或ORF序列)可以与AAV-in-BAC gRNA或ORF文库独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认中“堆叠”。

[0754] 实施例2:使用杆状病毒载体表达的用于筛选AAV产生的条形码文库技术

[0755] 本实施例描述了一种产生和筛选昆虫细胞文库的示例性方法,以确定通过文库内的特定细胞扰动(例如,在含有特定文库变体的细胞中)产生的示例性病毒载体(例如,AAV载体)的水平。具体而言,此实施例描述了一种将从昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法,其中将条形码用作标识符。此方法的示意图概览提供为图6。

[0756] 将构建体的示例性文库克隆到载体(例如,杆状病毒载体)中,所述载体包含位于两个AAV ITR序列之间的条形码。每个文库构建体可以包括文库变体(例如,gRNA和/或ORF文库变体),其位于例如两个AAV ITR序列之外。所得文库构建体将包括位于BAC载体中两个AAV ITR序列之间的条形码标识符序列和位于AAV ITR序列之外的一个或多个文库变体。文库变体(例如,gRNA和/或ORF)与条形码序列配对的相关配对可以在构建体设计和合成时预先确定,或者可以随机关联,取决于所使用的特定克隆方法,其中文库变体:条形码配对由克隆文库DNA的NGS确定。

[0757] 在gRNA文库变体文库的情况下,示例性AAV载体宿主细胞(例如,Sf9)被修饰以表达CRISPR核酸酶表达构建体,例如Cas9。生成了在组成型启动子和/或诱导型启动子(例如,四环素诱导型启动子(TetON))控制下瞬时或稳定表达Cas9的Sf9细胞。

[0758] 本公开涵盖以下认识:条形码序列可以鉴定产生AAV载体的昆虫细胞,以及昆虫细胞的对应的文库变体。图6提供了此纯化、扩增和测序的示意图概览。

[0759] 病毒载体是使用本领域已知的方法产生的。可以通过转导示例性哺乳动物细胞来评估从昆虫细胞纯化的重组AAV病毒载体的功能。

[0760] DNA可使用本领域已知的方法分离。可以对纯化的AAV载体内所含的DNA(特别是条

形码)进行测序。从纯化的AAV载体池(例如,整个AAV载体池、来自选定细胞子集的AAV载体池或选定的AAV载体)中分离DNA。可以使用本领域已知的方法,诸如碱裂解,来纯化DNA。使用PCR从侧翼引物序列扩增DNA,并纯化扩增的产物,与具有同源悬突的测序衔接子(例如,Illumina衔接子)混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用例如NextSeq平台进行测序。测量每个条形码的序列读段频率,并确定它们在所有条形码池内的相对丰度。

[0761] 宿主细胞产生株文库中含有的文库变体序列(例如,gRNA和/或ORF序列)也可扩增并测序。从AAV载体相关DNA扩增的条形码序列的相对丰度将与原始宿主细胞群中的条形码序列丰度进行比较。通过这种方式,可以鉴定导致扰动的文库变体,所述扰动直接改变宿主细胞生物学并导致差异AAV载体产生,因为与参考宿主细胞群相比,所述gRNA文库变体在AAV载体DNA群中富集或去富集。

[0762] 在此实施例中描述的方法还允许迭代轮次的文库筛选,其中在AAV载体群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的AAV载体产生的gRNA序列可以与AAV-in-BAC文库独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认中“堆叠”。

[0763] 实施例3:使用杆状病毒载体表达的用于筛选慢病毒产生的条形码化文库技术

[0764] 本实施例描述了一种产生和筛选昆虫细胞文库的示例性方法,以确定通过文库内的特定细胞扰动(例如,在含有特定文库变体的细胞中)产生的示例性病毒载体(例如,慢病毒载体)的水平。具体而言,此实施例描述了一种将从昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法,其中将条形码用作标识符。

[0765] 产生宿主生产细胞系(例如Sf9)内的示例性重组慢病毒载体产生昆虫细胞文库。示例性杆状病毒构建体各自包含位于慢病毒5’LTR和相关慢病毒包装序列(包括Psi)下游以及位于慢病毒3’LTR上游的条形码标识符,适用于慢病毒包装由LTR界定的转录的RNA。

[0766] 每个文库构建体可以包含文库变体(例如,gRNA和/或ORF文库变体),其位于例如LTR序列之外。所得文库构建体将包括位于BAC载体中两个慢病毒LTR序列之间的条形码标识符序列和位于慢病毒LTR序列之外的一个或多个文库变体。文库变体(例如,gRNA和/或ORF)与条形码序列配对的相关配对可以在构建体设计和合成时预先确定,或者可以随机关联,取决于所使用的特定克隆方法,其中文库变体:条形码配对由克隆文库DNA的NGS确定。

[0767] 在gRNA文库变体文库的情况下,示例性昆虫宿主细胞(例如,Sf9)被修饰以表达CRISPR核酸酶表达构建体,例如Cas9。生成了在组成型启动子和/或诱导型启动子(例如,四环素诱导型启动子(TetON))控制下瞬时或稳定表达Cas9的Sf9细胞。

[0768] 本公开涵盖以下认识:条形码序列可以鉴定产生慢病毒载体的昆虫细胞,以及昆虫细胞的对应的文库变体。

[0769] 慢病毒载体是使用本领域已知的方法产生的。可以通过转导示例性哺乳动物细胞来评估从昆虫细胞纯化的重组慢病毒载体的功能。

[0770] 将对纯化的慢病毒载体RNA进行测序,特别是标识符序列(文库变体(例如,gRNA和/或ORF编码序列)。从整个纯化的慢病毒载体池中分离全体RNA。使用碱裂解纯化RNA。使用具有标识符(例如,gRNA和/或ORF)下游结合位点的通用引物反向序列将RNA逆转录成cDNA。然后使用标识符序列(例如,gRNA和/或ORF)侧翼的通用正向和反向引物对此cDNA进

行PCR。纯化PCR扩增产物,将其与具有同源悬突的Illumina衔接子混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用例如NextSeq平台进行测序。宿主细胞产生株文库中含有的为整合DNA的标识符序列也将通过PCR被扩增并以同样的方式制备用于测序。

[0771] 从慢病毒载体扩增的条形码序列的相对丰度将与原始宿主细胞群中的条形码序列丰度进行比较。通过这种方式,可以鉴定导致扰动的文库变体,所述扰动直接改变宿主细胞生物学并导致差异慢病毒载体产生,因为与参考宿主细胞群相比,所述文库变体富集或去富集慢病毒载体群。

[0772] 在此实施例中描述的方法还允许迭代轮次的文库筛选,其中在慢病毒载体群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的慢病毒载体产生的文库变体序列可以独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认中“堆叠”。

[0773] 实施例4:用于筛选慢病毒产生的文库变体技术,使用DNA转座进行基因组整合

[0774] 本实施例描述了一种产生和筛选昆虫细胞文库的示例性方法,以确定通过文库内的特定细胞变体(例如,含有特定文库变体)产生的示例性病毒载体(例如,在此实施例中,慢病毒载体)的水平。具体而言,此实施例描述了一种将从昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法,其中用于创建昆虫细胞文库的文库变体(例如,gRNA和/或ORF)是标识符。此方法的示意图概览提供为图7。

[0775] 本实施例描述了一种在宿主产生细胞系(例如Sf9)内生成产生重组慢病毒载体的昆虫细胞文库的方法。将合成针对编码蛋白质基因产物(文库变体)的基因组序列或ORF序列的gRNA序列文库,并将其克隆到慢病毒5’LTR和相关慢病毒包装序列(包括Psi)下游以及位于慢病毒3’LTR上游的质粒中,适用于慢病毒包装由LTR界定的转录的RNA。此LTR限定的区段本身将位于DNA序列之间,使DNA能够酶促整合到宿主细胞基因组中,例如通过使用piggyBac转座酶经由从piggyBac DNA转座子系统中提取的侧翼同源反向末端重复序列(ITR)。此类具有位于转座子ITR之间的慢病毒LTR的构建体将在本文中被称为Lenti-in-Transposon。Lenti-in-Transposon质粒还将含有选择标记基因,诸如嘌呤霉素抗性或绿色荧光蛋白,使含有最终整合文库构建体的昆虫细胞的未来选择或鉴定成为可能。如克隆到Lenti-in-Transposon质粒中的文库(例如,gRNA和/或ORF)文库将被制备为纯化的质粒池,用于转染到Sf9细胞中。此质粒文库将与转座酶表达质粒(表达反式作用整合序列)一起转染到Sf9细胞中,以驱动酶促整合到宿主基因组内的随机位置。在此种转染后产生的基因组修饰的文库细胞将经由连续暴露于选择性剂(诸如嘌呤霉素)或通过基于荧光标记基因(诸如GFP)的荧光细胞分选来分离。然后以此方式分离的稳定克隆的文库群将用产生和包装重组慢病毒载体所必需的质粒构建体(例如,形成病毒载体所必需的多核苷酸)转染,诸如编码病毒糖蛋白(例如VSV-G蛋白)的质粒,以及第二代或第三代慢病毒包装质粒,以提供最低限度的慢病毒Gag Pol、Rev基因功能。这些质粒上存在的功能将指导ITR限定的慢病毒载体基因组从转座子整合的Lenti-in-Transposon序列的产生,并且这些慢病毒载体RNA将被包装并释放到慢病毒载体中。使用此Lenti-in-Transposon方法将使每个细胞的文库成员实现低拷贝数整合,使得在选择稳定文库群的过程中耗尽大量过量的非整合转染质粒DNA。

[0776] 慢病毒载体是使用本领域已知的方法产生的。可以通过转导示例性哺乳动物细胞来评估从昆虫细胞纯化的重组慢病毒载体的功能。

[0777] 将对纯化的慢病毒载体RNA进行测序,特别是标识符序列(文库变体(例如,gRNA

和/或ORF编码序列)。从整个纯化的慢病毒载体池中分离全体RNA。使用碱裂解纯化RNA。使用具有标识符(例如,gRNA和/或ORF)下游结合位点的通用引物反向序列将RNA逆转录成cDNA。然后使用标识符序列(例如,gRNA和/或ORF)侧翼的通用正向和反向引物对此cDNA进行PCR。纯化PCR扩增产物,将其与具有同源悬突的Illumina衔接子混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用例如NextSeq平台进行测序。宿主细胞产生株文库中含有的为整合DNA的标识符序列也将通过PCR被扩增并以同样的方式制备用于测序。

[0778] 将通过获得自每次Illumina读段的部分末端标识符序列对它们已知的全长序列的参考来鉴定每个读段的标识符。因此测量每个ORF的频率,并确定它们在所有标识符池内的相对丰度。从慢病毒载体相关RNA扩增的标识符(例如,gRNA和/或ORF)的相对丰度将与原始宿主细胞群中的标识符序列丰度进行比较。通过这种方式,可以鉴定导致扰动的文库变体,所述扰动直接改变宿主细胞生物学并导致差异慢病毒载体产生,因为与宿主细胞群相比,所述文库变体在慢病毒群中富集或去富集。

[0779] 在此实施例中描述的方法还允许迭代轮次的文库筛选,其中在慢病毒群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的慢病毒载体产生的标识符序列可以与Lenti-in-Transposon ORF文库独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认中“堆叠”。

[0780] 实施例5:用于筛选AAV产生的条形码文库技术,使用DNA转座进行基因组整合

[0781] 本实施例描述了一种产生和筛选昆虫细胞文库的示例性方法,以确定通过文库内的特定文库变体产生的示例性病毒载体(例如,AAV载体)的水平。具体而言,此实施例描述了昆虫细胞示例性文库的构建,所述昆虫细胞使用条形码标识符与基于CRISPR gRNA的文库,所述文库通过DNA转座被整合到昆虫细胞基因组中。此实施例还提供了一种将从所述昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法。

[0782] 本实施例描述了一种在昆虫宿主产生细胞系(例如Sf9)内生成产生重组AAV载体的昆虫细胞文库的方法。示例性AAV载体产生昆虫细胞系例如Sf9被修饰以稳定表达CRISPR核酸酶表达构建体,例如Cas9。针对基因组序列的gRNA序列文库(文库变体)被合成并在两个AAV ITR序列之外的位置克隆到质粒中,其中对应的条形码标识符序列位于两个AAV ITR序列之间。此ITR限定的区段位于DNA序列之间,使DNA能够酶促整合到宿主细胞基因组中,例如通过使用piggyBac转座酶经由从piggyBac DNA转座子系统中提取的侧翼同源反向末端重复序列(ITR)。此种具有位于转座子ITR之间的AAV ITR的构建体将在本文中被称为AAV-in-Transposon;此种AAV-in-Transposon构建体的示例性示意图提供于图4,标记为pSFX-PB-AAV。AAV-in-Transposon质粒还含有示例性选择标记基因、抗生素抗性基因(例如,嘌呤霉素抗性)和示例性荧光蛋白基因(例如,绿色荧光蛋白)的AAV有效载荷,使含有整合文库构建体并产生AAV载体的昆虫细胞的未来选择或鉴定成为可能。这些质粒文库与转座酶表达质粒(表达反式作用整合序列)一起转染到Sf9细胞中,以驱动酶促整合到宿主基因组内的随机位置。

[0783] 在此种转染后产生的基因组修饰的文库细胞经由连续暴露于选择性剂(诸如嘌呤霉素)或通过基于荧光标记基因(诸如GFP)的荧光细胞分选来分离。用荧光显微镜观察荧光细胞,并在所有转座酶:转座子比率下观察到荧光细胞,表明整合成功。

[0784] 然后用产生和包装重组AAV病毒载体所必要的质粒构建体(例如,形成病毒载体所

必需的多核苷酸)(例如,pHelper和pAAV Rep-Cap)转染纯化的细胞文库群。这些质粒上存在的功能指导ITR限定的AAV载体基因组从基因组整合的AAV-in-Transposon序列的复制,并且这些ITR病毒载体构建体被包装并释放到AAV载体中。

[0785] 转座子拷贝数是通过qPCR相对于质粒标准来测量的。转染的转座酶:转座子质粒的比率得到优化,使得转座子整合细胞的拷贝数低于以前观察到的,使大多数细胞仅接受大约1个遗传扰动(1个gRNA文库变体/来自文库的条形码对)。将具有顺式作用转座子的质粒与含有编码piggyBac转座酶的反式作用序列的质粒以各种转座酶:转座子比率共转染,以获得较低的拷贝数。拷贝数可随着转座子质粒相对量的增加而增加,以鉴定优化的转座酶:转座子比率,以实现每个细胞大约1个拷贝。

[0786] 病毒载体是使用本领域已知的方法产生的。可以通过转导到示例性哺乳动物细胞中并用荧光显微镜可视化荧光标记(例如,GFP)的细胞来评估从转座酶整合、ITR侧接的模板DNA中纯化的AAV的功能。

[0787] 本实施例描述了一种对纯化的AAV载体中包含的DNA,特别是对应于文库gRNA的条形码序列进行测序的方法。DNA使用本领域已知的方法分离。使用PCR自侧翼引物序列扩增DNA。纯化扩增产物,将其与具有同源悬突的Illumina衔接子混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用NextSeq平台进行测序。宿主细胞产生株文库中含有的gRNA序列也将被扩增并以同样的方式制备用于测序。测量每个gRNA的序列读段频率,并确定它们在所有gRNA池内的相对丰度。从AAV载体相关DNA扩增的gRNA序列的相对丰度将与原始宿主细胞群中的gRNA序列丰度进行比较。通过这种方式,可以鉴定导致扰动的gRNA文库变体,所述扰动直接改变宿主细胞生物学并导致差异AAV载体产生,因为与宿主细胞群相比,所述gRNA文库变体在AAV载体DNA群中富集或去富集。

[0788] 在此实施例中描述的方法还允许迭代轮次的文库筛选,其中在AAV载体群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的AAV载体产生的gRNA序列可以与AAV-in-Transposon gRNA文库独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认中“堆叠”。

[0789] 实施例6:用于筛选游离质粒DNA上AAV产生的技术

[0790] 本实施例描述了一种产生和筛选昆虫细胞文库的方法,以确定通过文库内的特定细胞变体(例如,含有特定文库变体)产生的示例性病毒载体(例如,在此实施例中,AAV载体)的水平。具体而言,此实施例描述了一种将从昆虫细胞文库产生的单独病毒载体与它们所源自的特定细胞变体相关联的方法,其中用于创建昆虫细胞文库的每个转基因文库变体与标识符相关联,所述标识符可以是文库变体本身和/或DNA条形码。此方法的示意图概览提供为图8。

[0791] 本实施例描述了一种在宿主产生细胞系(例如,Sf9)内生成产生重组AAV载体的昆虫细胞文库的方法。将合成针对编码蛋白质基因产物(文库变体)的基因组序列或ORF序列的gRNA序列文库并克隆到质粒中。这些质粒可能还包含DNA条形码序列。此文库变体:标识符关联可以在构建体设计和合成时预先确定,或者可以根据所利用的特定克隆方法随机关联。此文库将构建成标识符序列位于两个AAV ITR序列之间,并且将被转录到poly-A尾mRNA转录物中,其中此转录物是侧翼为ITR的转录单位或跨越居间ITR序列的转录单位。此种具有AAV-on-plasmid的构建体还将含有选择标记基因,诸如嘌呤霉素抗性或绿色荧光蛋白,

使含有质粒构建体以及SV40复制起点的昆虫细胞的未来选择或鉴定成为可能,促进在后续的培养传代过程中保留转染细胞内的质粒。

[0792] 如克隆到AAV-on-plasmid中的文库将被制备为纯化的质粒池,用于转染到昆虫细胞(例如,Sf9)中。然后将用产生和包装重组AAV载体所必需的质粒构建体(例如,形成病毒载体所必需的多核苷酸)(例如,pHelper和pAAV Rep-Cap)转染以此方式转染的细胞文库群。这些质粒上存在的功能将指导ITR限定的AAV载体基因组从ITR-on-Plasmid序列的复制,并且这些ITR载体DNA将被包装并释放到AAV载体中。以低效率转染细胞将限制每个细胞的独特质粒数量,与Sf9细胞内的SV40起点依赖性复制一起,使得能够经由质粒携带的选择标记在文库细胞纯化步骤过程中保留低拷贝转染。此外,在表达mRNA转录物中包含标识符将促进每个细胞、多拷贝质粒文库成员身份的解卷积,这首先通过将可检测文库成员限制为在宿主细胞核内实际发挥作用且未保留在细胞其他地方(诸如胞内体或细胞质)的那些质粒,其次通过使用建立的单细胞条形码化和RNA测序方法(诸如10x Genomics)捕获标识符序列来促进每个细胞多重质粒文库成员的鉴定。使用此AAV-on-Plasmid方法将限制每个宿主细胞的多重文库成员数量的问题,同时促进每个细胞多拷贝身份的解卷积。然后可以将这些每个细胞的组合文库身份与AAV载体富集条形码相关联,以自信地鉴定多重文库群中的命中。

[0793] 病毒载体是使用本领域已知的方法产生的。可以通过转导示例性哺乳动物细胞来评估从昆虫细胞纯化的重组AAV病毒载体的功能。

[0794] DNA包含在纯化的AAV载体内,特别是标识符序列。从整个纯化的AAV病毒载体池中分离全体DNA。使用碱裂解纯化DNA。使用PCR自侧翼引物序列扩增DNA。纯化扩增产物,将其与具有同源悬突的Illumina衔接子混合。然后将此材料扩增以添加Illumina衔接子和索引,并使用NextSeq平台进行测序。宿主细胞产生株文库中含有的标识符序列也将被扩增并以同样的方式制备用于测序。测量每个标识符的序列读段频率,并确定它们在所有标识符池内的相对丰度。从AAV载体相关DNA扩增的标识符序列的相对丰度将与原始宿主细胞群中的标识符序列丰度进行比较。

[0795] 在标识符是条形码的情况下,每个条形码都将与特定的文库变体(例如,gRNA和/或ORF文库变体)具有已知的关联,可以鉴定导致扰动的/ORF文库变体,所述扰动直接改变宿主细胞生物学并导致差异AAV载体产生,因为与宿主细胞群相比,ORF文库变体的相关条形码在AAV载体DNA群中富集或去富集。

[0796] 有了足够高倍的细胞文库覆盖率和测序深度,真正富集的命中将在固有多重转染质粒文库中被鉴定。此外,为了简化多重转染文库中的命中鉴定问题,可以对生产者群进行单细胞测序。生产者细胞将在细胞mRNA的逆转录过程中被单独标记(10XGenomics),将细胞身份序列附加到细胞mRNA,从而标记细胞文库群中含有的所有标识符。然后可以通过PCR特异性扩增细胞身份序列标记的含有标识符的cDNA,并且这些扩增子可以如上所述制备以用于由Illumina进行NGS。这些扩增子的测序将揭示细胞文库群中存在的每个细胞文库成员组合,允许辨别单个或组合的命中与错误富集的‘搭便车者(hitchhiker)’。

[0797] 在此实施例中描述的方法还允许迭代轮次的文库筛选,其中在AAV载体群中发现显著富集,并确认在细胞中引入扰动以得到更高滴度的AAV载体产生的文库变体序列可以与AAV-on-Plasmid文库独立引入宿主细胞,使突变的组合能够在连续几轮筛选和命中确认

中“堆叠”。

[0798] 等效方案

[0799] 本领域技术人员应当理解,本领域技术人员将容易地想到对本公开内容的各种改变、修改和改进。此类改变、修改和改进旨在成为本公开内容的一部分,并且旨在在本发明的精神和范围之内。因此,前述描述和附图仅是作为例子的,并且本公开中描述的任何发明均由所附权利要求书进一步详细描述。

[0800] 本领域技术人员应理解可归因于如本文所述的测定或其他方法中获得的值的典型的标准偏差或误差。本文中为了描述本发明的背景并提供关于其实践的额外细节所引用的出版物、网址和其他参考材料由此以引用的方式整体并入。

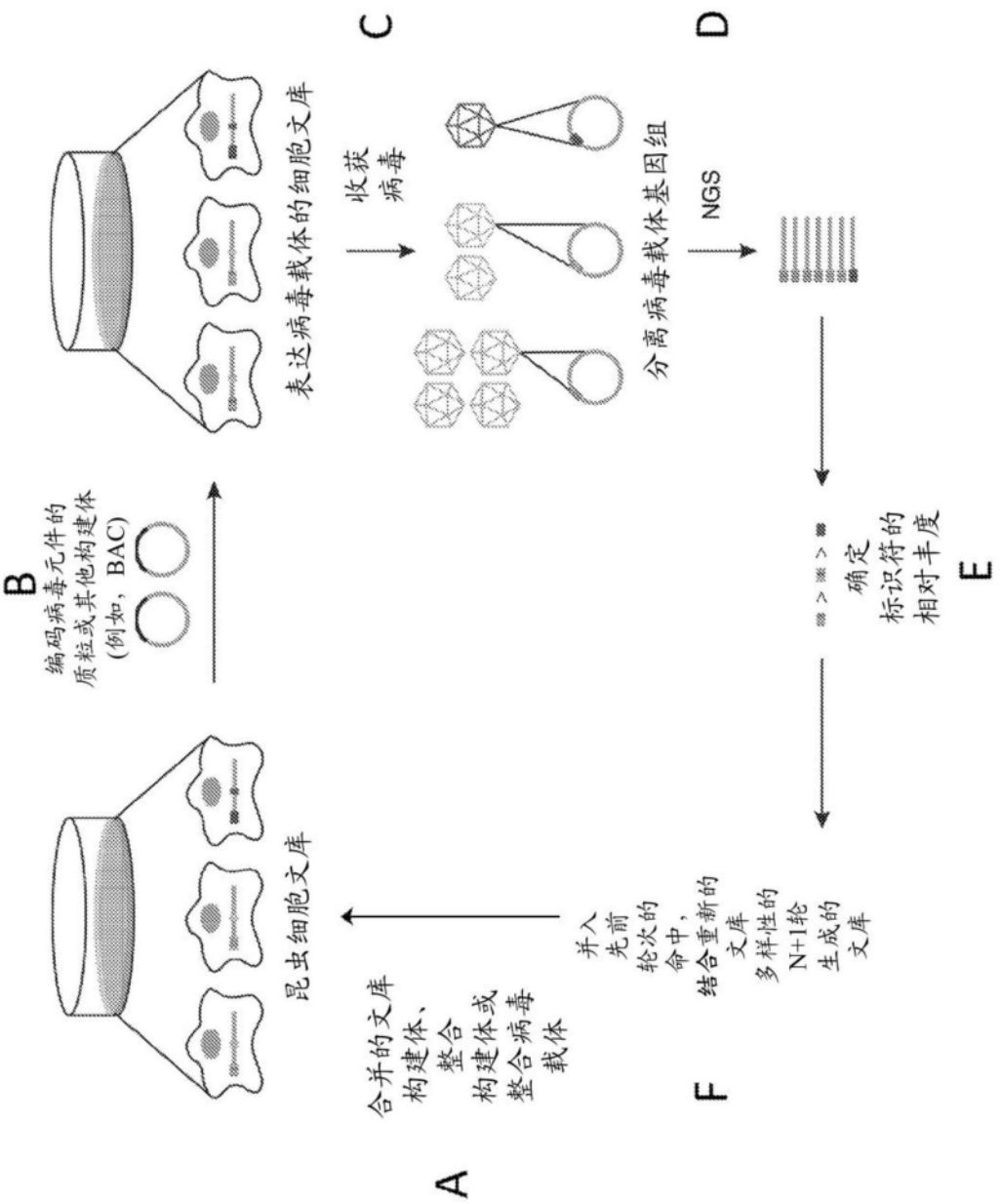

A
合并的文库
构建体、
整合
构建体或
整合病毒
载体
昆虫细胞文库
B
编码病毒元件的
质粒或其他构建体
(例如，BAC)
表达病毒载体的细胞文库
C
收获
病毒
分离病毒载体基因组
D
NGS
E
确定
标识符的
相对丰度
F
并入
先前
轮次的
命中，
结合重新的
文库
多样性的
N+1轮
生成的
文库

图1

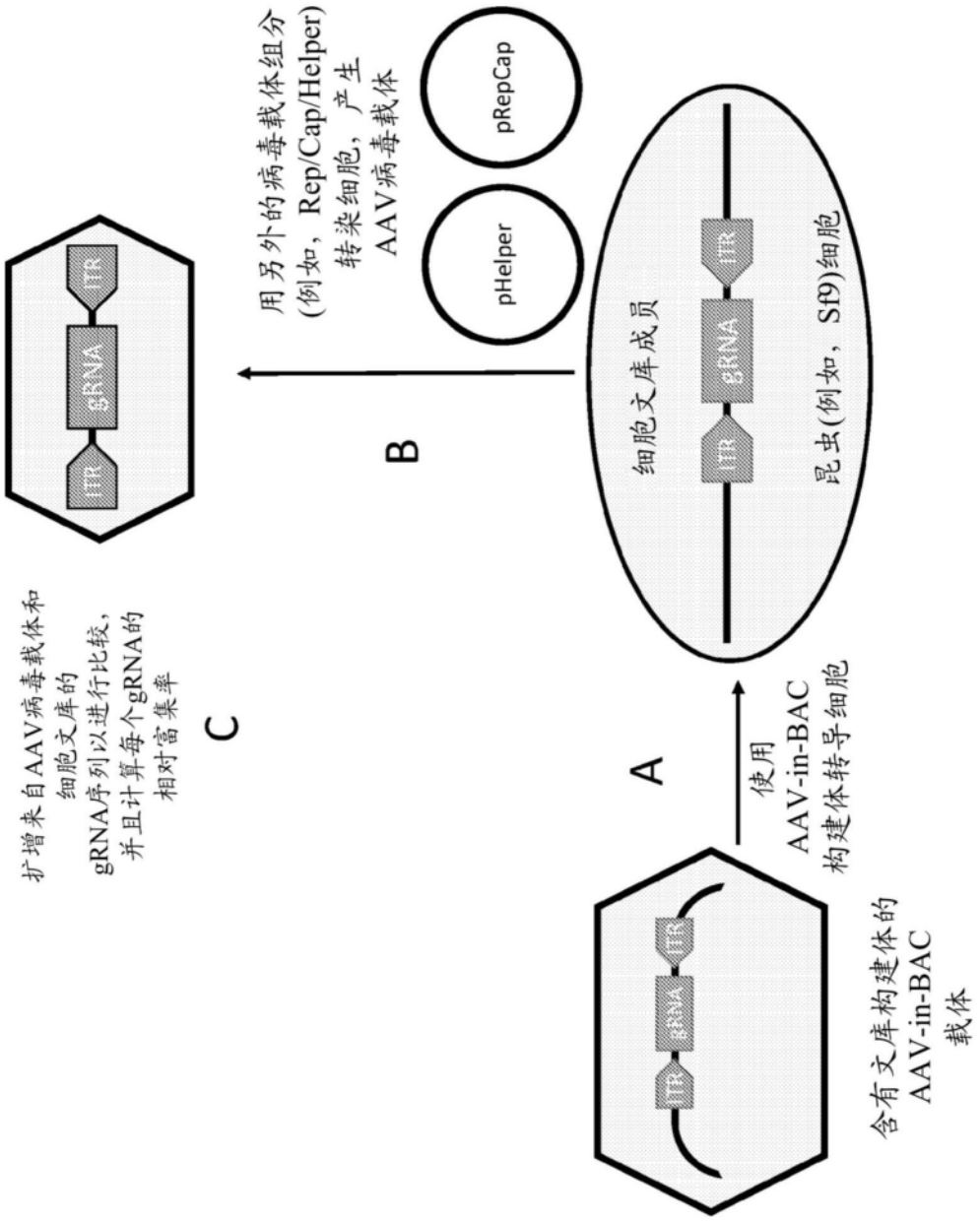

含有文库构建体的
AAV-in-BAC
载体
A
使用
AAV-in-BAC
构建体转导细胞
昆虫(例如，Sf9)细胞
细胞文库成员
ITR
gRNA
ITR
B
用另外的病毒载体组分
(例如，Rep/Cap/Helper)
转染细胞，产生
AAV病毒载体
pHelper
pRepCap
C
扩增来自AAV病毒载体和
细胞文库的
gRNA序列以进行比较，
并且计算每个gRNA的
相对富集率

图2

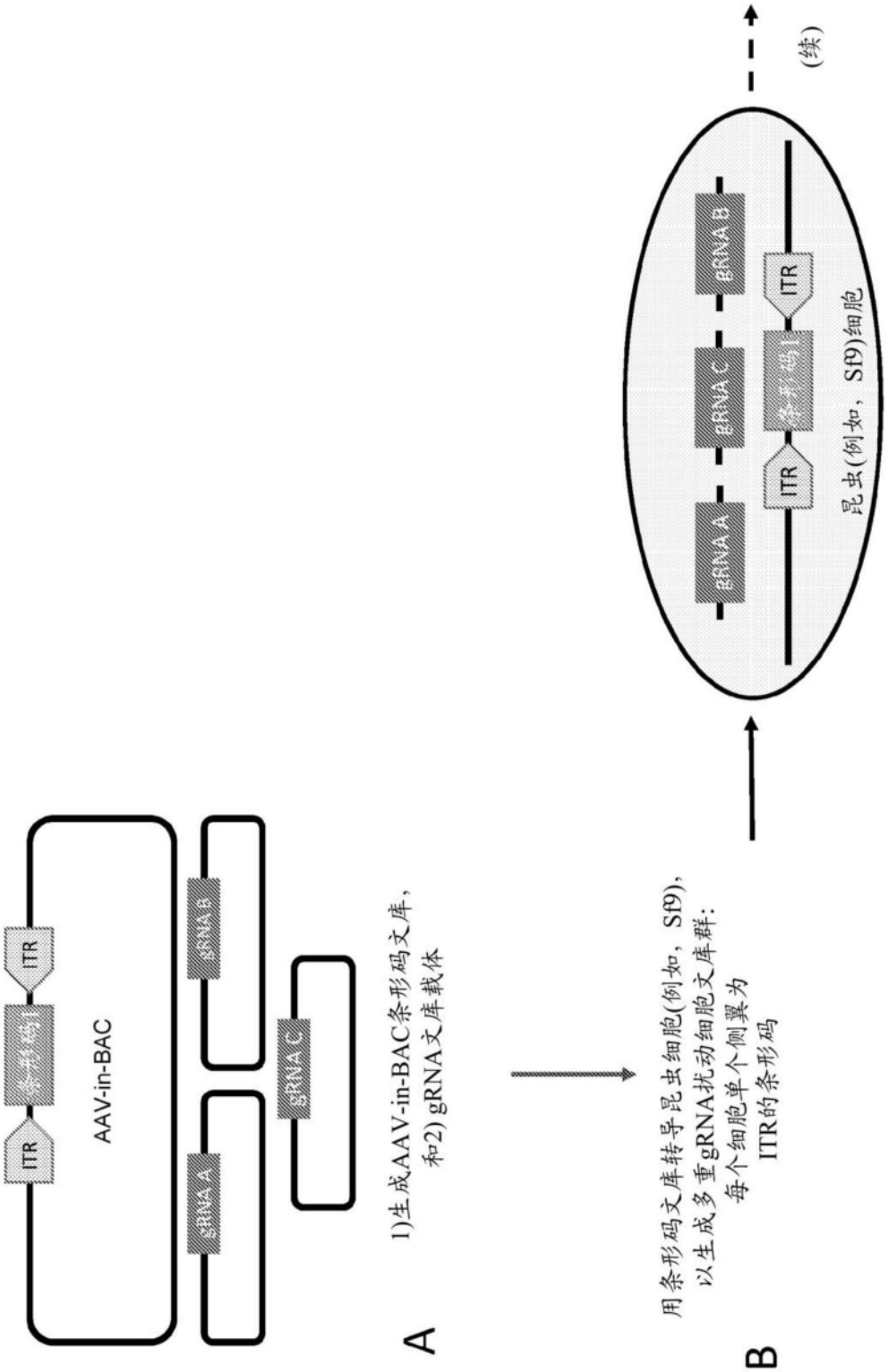

ITR
条形码1
ITR
AAV-in-BAC
gRNA A
gRNA B
gRNA C
A
1)生成AAV-in-BAC条形码文库，
和2) gRNA文库载体
用条形码文库转导昆虫细胞(例如，Sf9)，
以生成多重gRNA扰动细胞文库群：
每个细胞单个侧翼为
ITR的条形码
B
gRNA A
gRNA C
gRNA B
ITR
条形码1
ITR
昆虫(例如，Sf9)细胞
(续)

图3

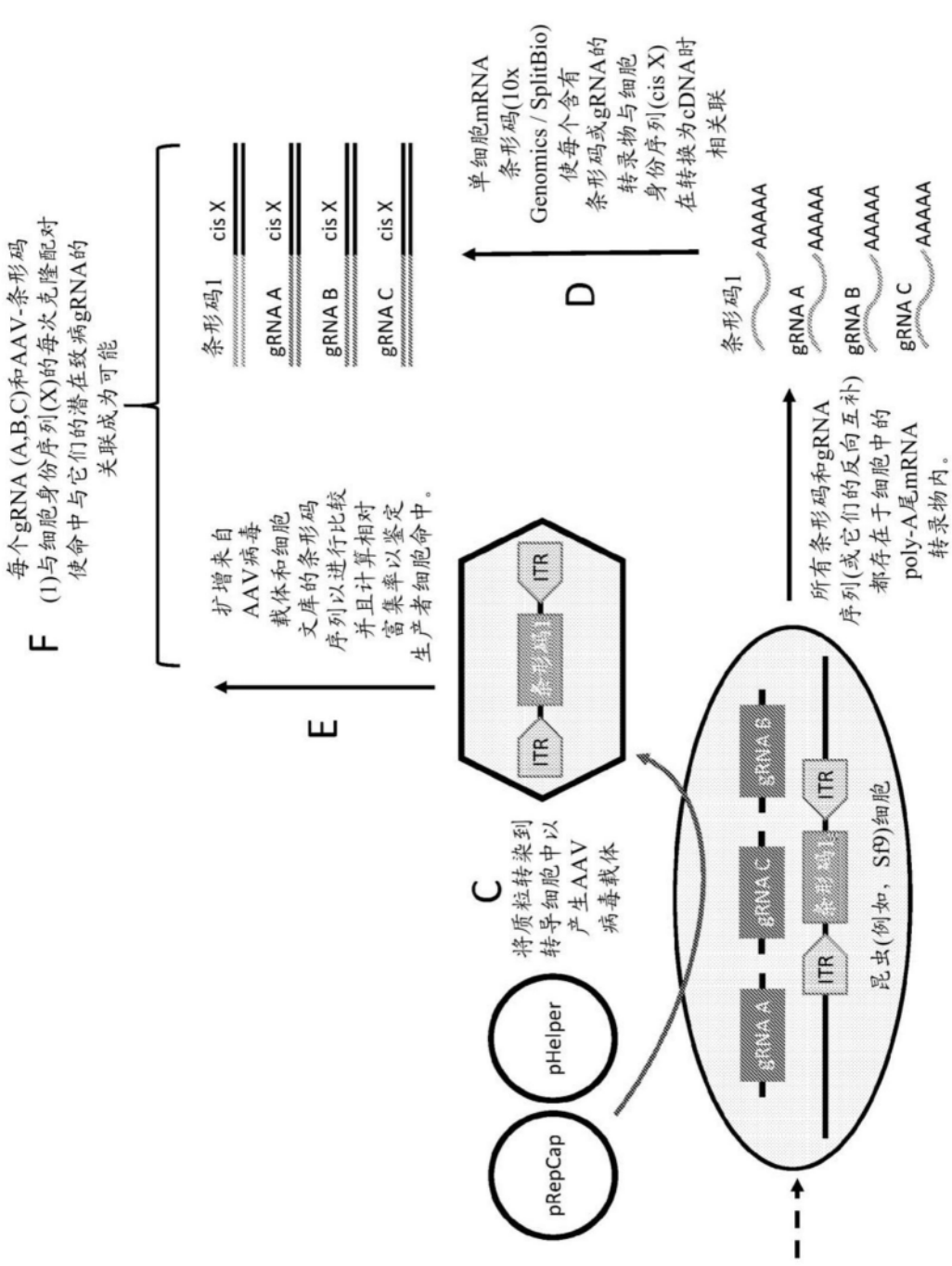

每个gRNA (A,B,C)和AAV-条形码
(1)与细胞身份序列(X)的每次克隆配对
使命中与它们的潜在致病gRNA的
关联成为可能
F
条形码1 cis X
gRNA A cis X
gRNA B cis X
gRNA C cis X
扩增来自
AAV病毒
载体和细胞
文库的条形码
序列以进行比较
并且计算相对
富集率以鉴定
生产者细胞命中。
E
单细胞mRNA
条形码(10x
Genomics / SplitBio)
使每个含有
条形码或gRNA的
转录物与细胞
身份序列(cis X)
在转换为cDNA时
相关联
D
条形码1 AAAAA
gRNA A AAAAA
gRNA B AAAAA
gRNA C AAAAA
所有条形码和gRNA
序列(或它们的反向互补)
都存在于细胞中的
poly-A尾mRNA
转录物内。
ITR
条形码1
ITR
C
将质粒转染到
转导细胞中以
产生AAV
病毒载体
pRepCap
pHelper
gRNA A
gRNA C
gRNA B
ITR
条形码1
ITR
昆虫(例如，Sf9)细胞

图3 (续)

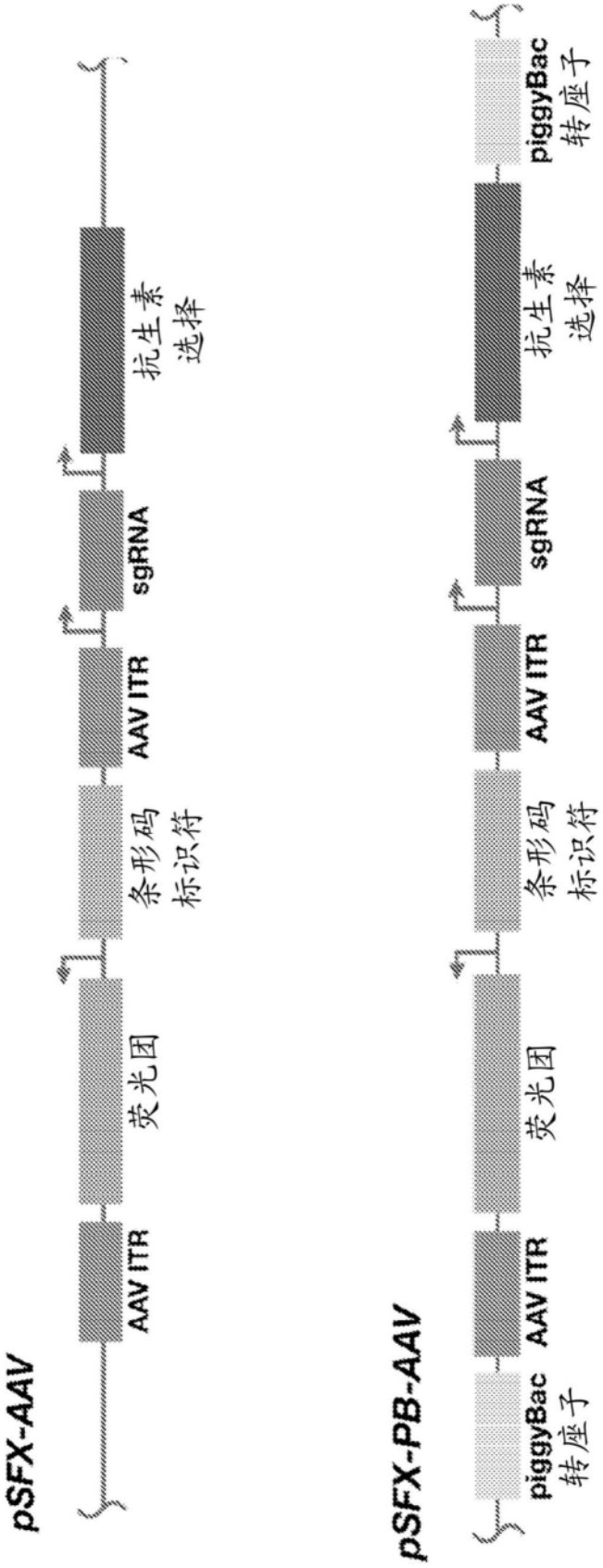

pSFX-AAV
AAV ITR
荧光团
条形码
标识符
AAV ITR
sgRNA
抗生素
选择
pSFX-PB-AAV
piggyBac
转座子
AAV ITR
荧光团
条形码
标识符
AAV ITR
sgRNA
抗生素
选择
piggyBac
转座子

图4

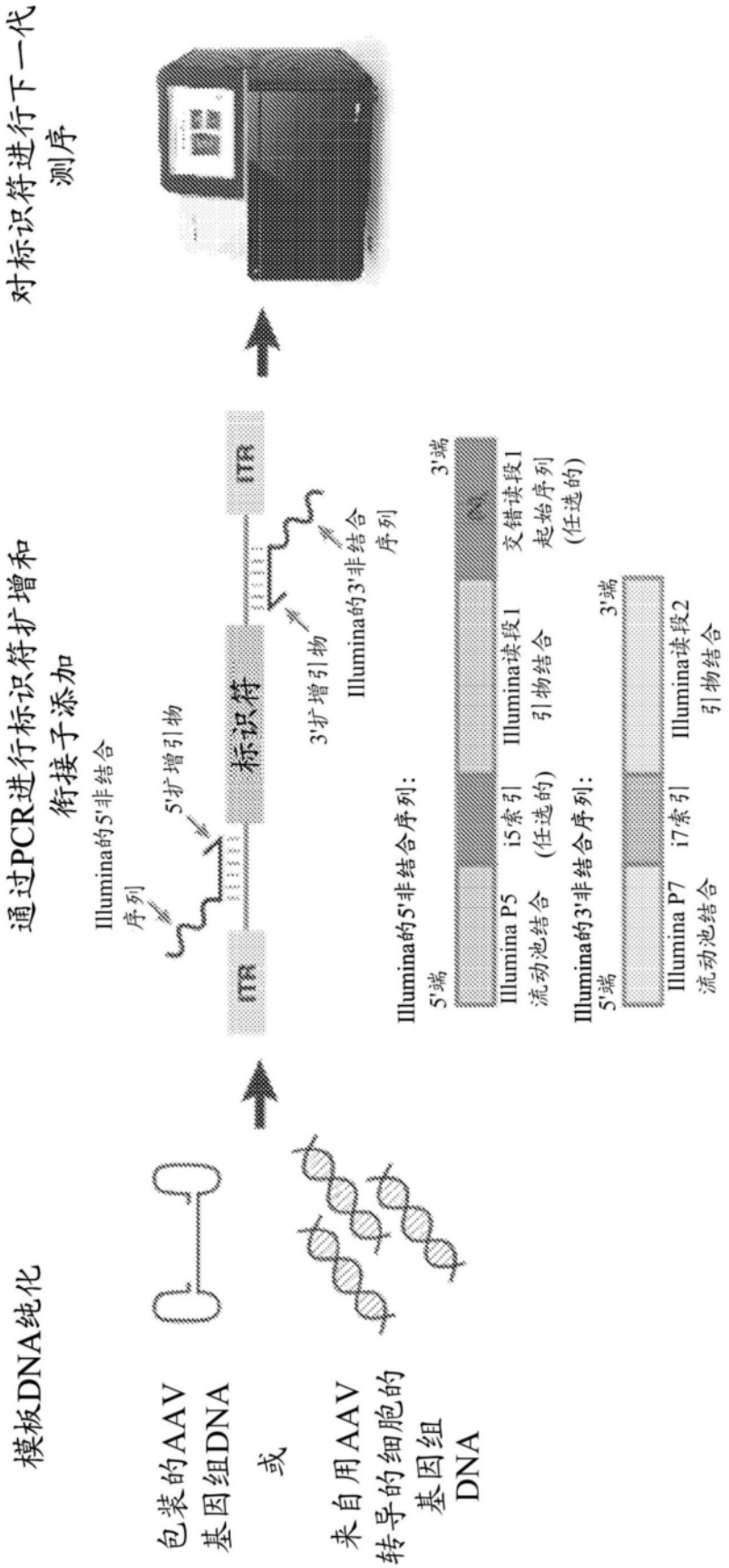

模板DNA纯化
包装的AAV基因组DNA
或
来自用AAV转导的细胞的基因组DNA
通过PCR进行标识符扩增和衔接子添加
Illumina的5'非结合序列
5'扩增引物
ITR
标识符
3'扩增引物
Illumina的3'非结合序列
ITR
Illumina的5'非结合序列:
5'端
3'端
Illumina P5 流动池结合
i5索引 (任选的)
Illumina读段1 引物结合
交错读段1起始序列 (任选的)
Illumina的3'非结合序列:
5'端
3'端
Illumina P7 流动池结合
i7索引
Illumina读段2 引物结合
对标识符进行下一代测序

图5

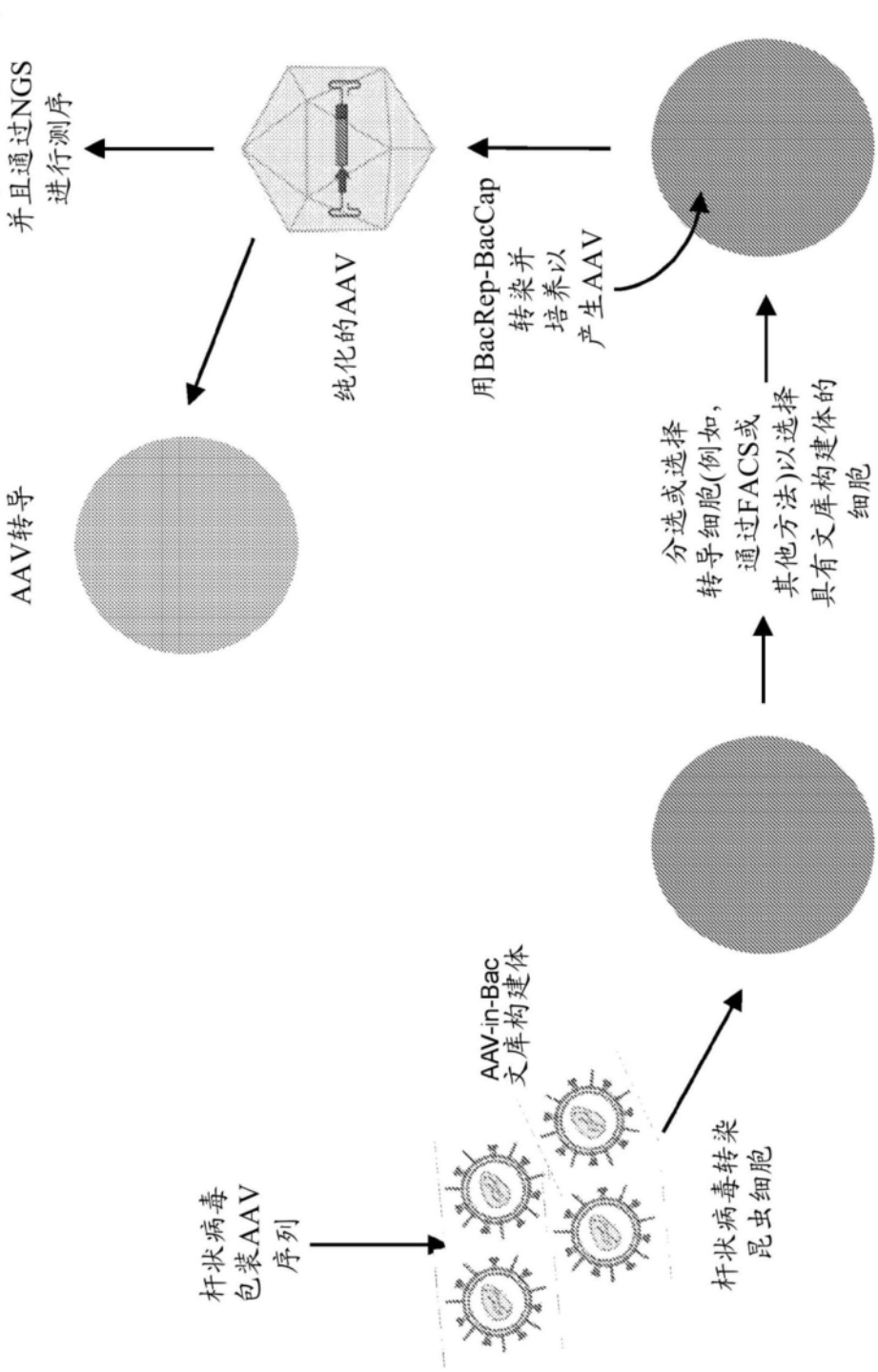

杆状病毒
包装AAV
序列
AAV-in-Bac
文库构建体
杆状病毒转染
昆虫细胞
分选或选择
转导细胞(例如，
通过FACS或
其他方法)以选择
具有文库构建体的
细胞
用BacRep-BacCap
转染并
培养以
产生AAV
纯化的AAV
AAV转导
从病毒核酸扩增
标识符(例如，条形码)
并且通过NGS
进行测序

图6

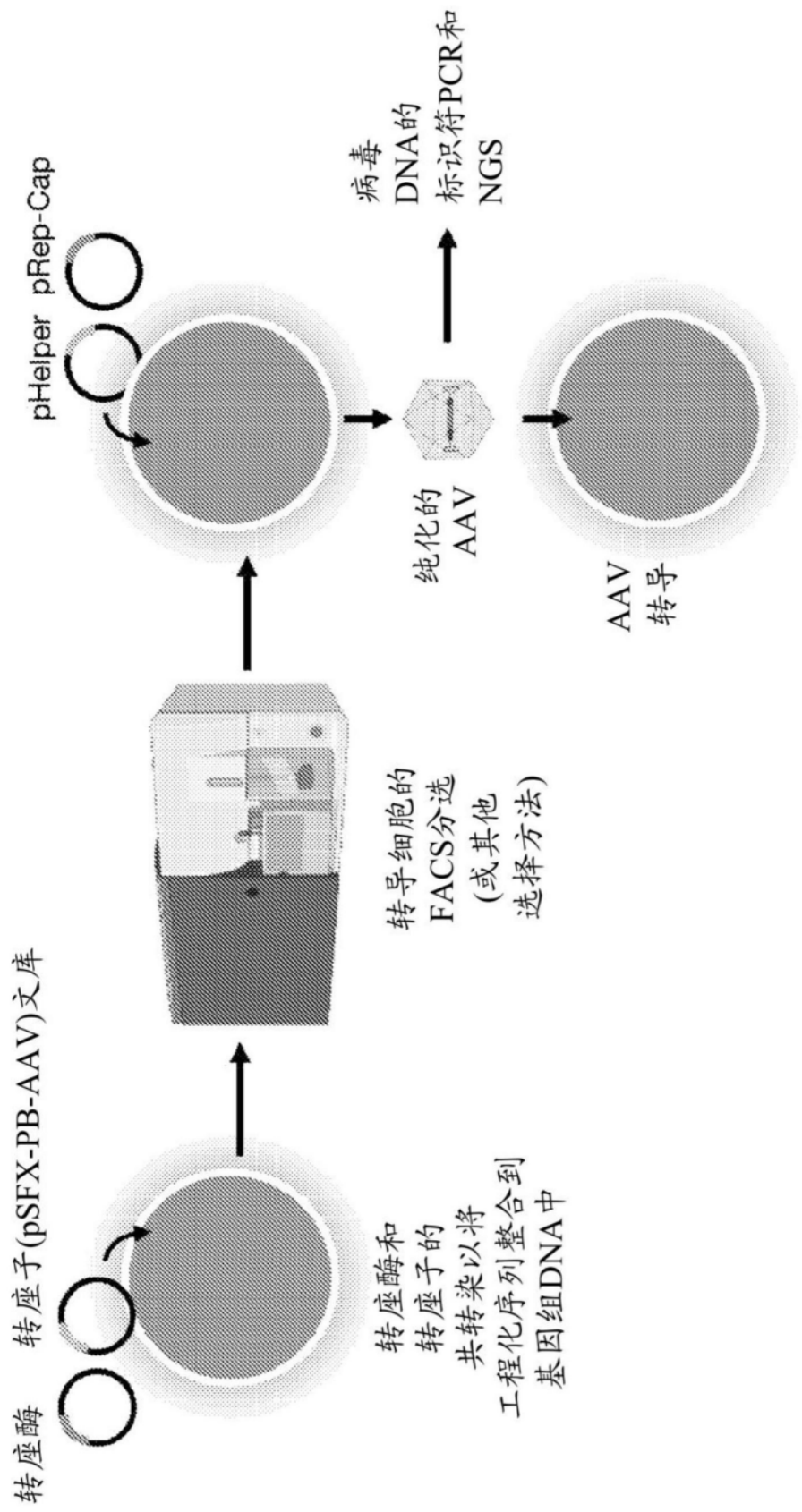

转座酶
转座子(pSFX-PB-AAV)文库
转座酶和
转座子的
共转染以将
工程化序列整合到
基因组DNA中
转导细胞的
FACS分选
(或其他
选择方法)
pHelper pRep-Cap
纯化的
AAV
病毒
DNA的
标识符PCR和
NGS
AAV
转导

图7

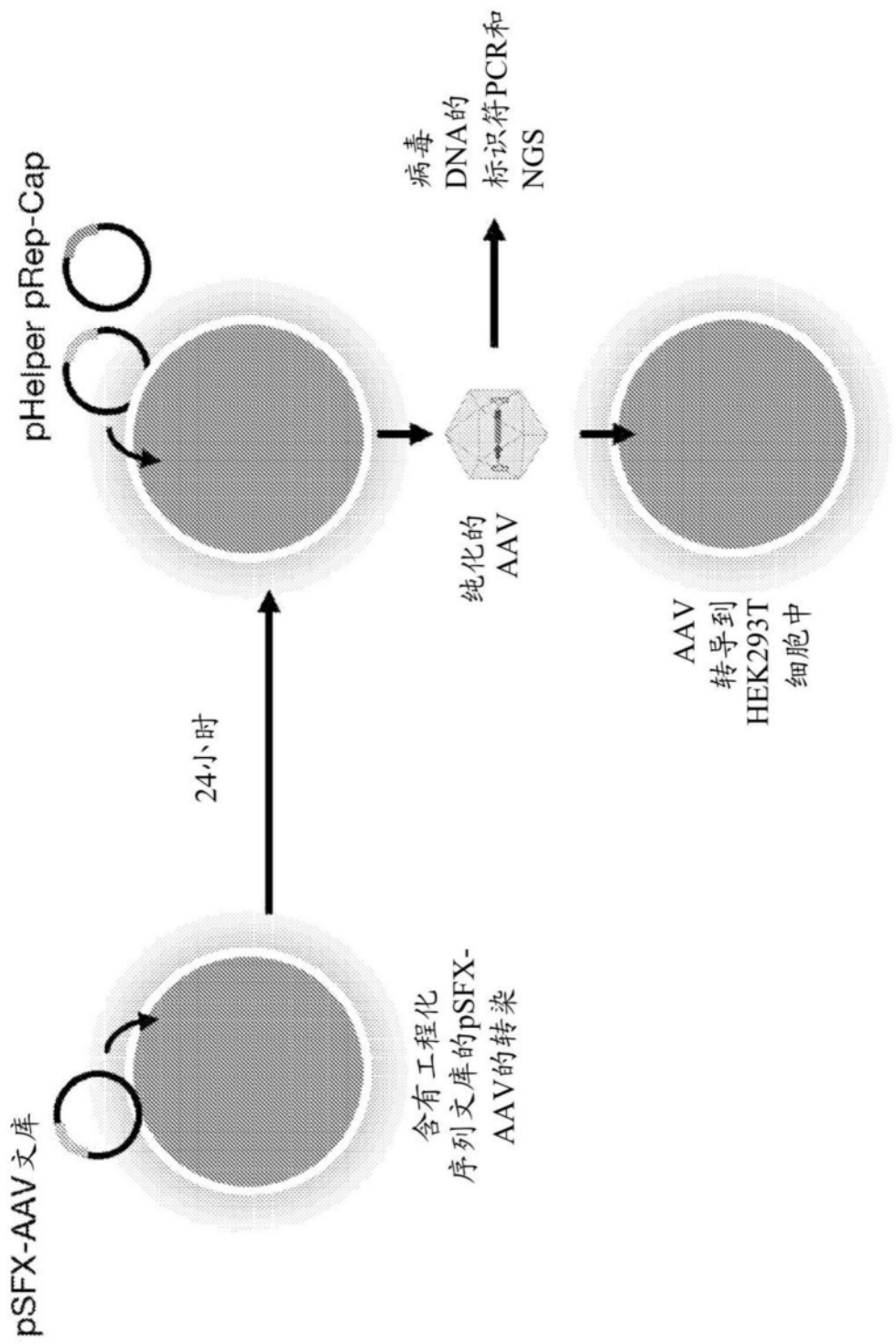

pSFX-AAV文库
含有工程化序列文库的pSFX-AAV的转染
24小时
pHelper pRep-Cap
纯化的AAV
病毒DNA的标识符PCR和NGS
AAV转导到HEK293T细胞中

图8