



(12)发明专利

(10)授权公告号 CN 104404050 B

(45)授权公告日 2018.06.08

(21)申请号 201410535091.9

(22)申请日 2011.06.22

(65)同一申请的已公布的文献号

申请公布号 CN 104404050 A

(43)申请公布日 2015.03.11

(30)优先权数据

61/357,314 2010.06.22 US

61/357,317 2010.06.22 US

(62)分案原申请数据

201180038411.4 2011.06.22

(73)专利权人 瑞泽恩制药公司

地址 美国纽约州

(72)发明人 L·麦克唐纳 S·史蒂文斯

C·古雷尔 A·J·莫菲

K·A·霍希亚瓦

(74)专利代理机构 北京商专永信知识产权代理

事务所(普通合伙) 11400

代理人 郭玥 葛强

(51)Int.Cl.

C12N 15/13(2006.01)

C12N 15/85(2006.01)

A01K 67/027(2006.01)

C07K 16/00(2006.01)

C12N 5/10(2006.01)

C12N 5/0735(2010.01)

C12N 5/20(2006.01)

A61K 39/395(2006.01)

(56)对比文件

WO 0026373 A1,2000.05.11,全文.

WO 2008054606 A2,2008.05.08,全文.

WO 2009143472 A2,2009.11.26,全文.

Andrei V.Popov,et al.A human

Immunoglobulin λ Locus Is Similarly Well Expressed in Mice and Humans.《Journal of Experimental Medicine》.1999,第189卷(第10期),第1611-1619页.

审查员 陶冶

权利要求书2页 说明书38页

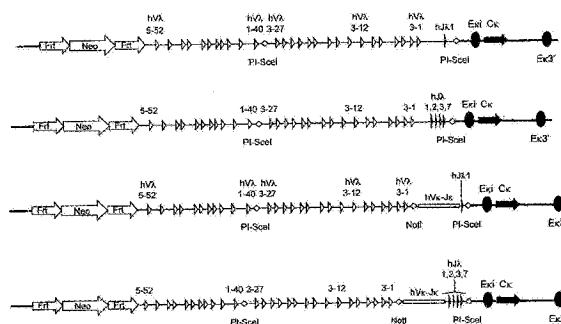
序列表32页 附图24页

(54)发明名称

杂交轻链小鼠

(57)摘要

提供了表达人 λ 可变(hV λ)序列的遗传修饰的小鼠,包括从内源小鼠 λ 轻链基因座表达hV λ 序列的小鼠,从内源小鼠 κ 轻链基因座表达hV λ 序列的小鼠,以及从其中hV λ 序列连接至小鼠恒定序列的转基因或附加体表达hV λ 序列的小鼠。提供了其为用于产生抗原结合蛋白的体细胞突变的人 λ 可变序列的来源的小鼠。提供了用于产生包含人 λ 可变序列的抗原结合蛋白包括人抗体的组合物和方法。



1. 一种轻链等位基因,所述轻链等位基因在内源性小鼠 κ 免疫球蛋白轻链基因座,所述轻链等位基因包含与小鼠C κ 基因区段可操作地连接的、未重排的人 λ 免疫球蛋白轻链可变基因区段(hV λ)和未重排的人 λ 连接基因区段(hJ λ),其中所述等位基因包括小鼠V κ 基因区段和小鼠J κ 基因区段的缺失。

2. 如权利要求1所述的轻链等位基因,其中,所述轻链等位基因包含多个未重排的hV λ 基因区段。

3. 如权利要求2所述的轻链等位基因,其中,所述轻链等位基因包含:

(a) 至少12个hV λ 基因区段;

(b) 13至28个hV λ 基因区段;或

(c) 28个或更多个hV λ 基因区段。

4. 如权利要求3所述的轻链等位基因,其中,所述hV λ 基因区段包含:hV λ 3-1、hV λ 4-3、hV λ 2-8、hV λ 3-9、hV λ 3-10、hV λ 2-11和hV λ 3-12。

5. 如权利要求3的轻链等位基因,其中,所述hV λ 基因区段包含:hV λ 2-14、hV λ 3-16、hV λ 2-18、hV λ 3-19、hV λ 3-21、hV λ 3-22、hV λ 2-23、hV λ 3-25和hV λ 3-27。

6. 如权利要求1所述的轻链等位基因,其中,所述人J λ 基因区段是hJ λ 1。

7. 如权利要求1所述的轻链等位基因,其中,所述轻链等位基因包含四个未重排的hJ λ 基因区段。

8. 如权利要求7所述的轻链等位基因,其中,所述四个未重排的hJ λ 基因区段是J λ 1、J λ 2、J λ 3和J λ 7。

9. 如权利要求1所述的轻链等位基因,其中,所述轻链等位基因包含:

(a) 横跨hV λ 3-12至hV λ 3-1的人 λ 轻链基因座的连续序列;

(b) 横跨hV λ 3-27至hV λ 3-1的人 λ 轻链基因座的连续序列;或

(c) 横跨hV λ 3-29至hV λ 3-1的人 λ 轻链基因座的连续序列,以及横跨hV λ 5-52至hV λ 1-40的人 λ 轻链基因座的连续序列。

10. 一种用于产生小鼠的方法,所述方法包括遗传修饰所述小鼠以使所述小鼠在其种系中包含如权利要求1-9中任一项所述的轻链等位基因。

11. 一种抗体,所述抗体是在如权利要求10所述方法产生的经遗传修饰的小鼠中生成的,其中所述抗体的轻链包括人 λ 免疫球蛋白轻链可变结构域和小鼠 κ 免疫球蛋白轻链恒定结构域。

12. 一种细胞,所述细胞包含如权利要求1-9中任一项所述的小鼠轻链等位基因,其中所述细胞不是胚胎干细胞。

13. 一种生成分离的小鼠胚胎干细胞(ES)的方法,所述方法包含遗传修饰所述ES细胞以使其在其种系中包含如权利要求1-9中任一项所述的轻链等位基因。

14. 一种编码重排的人 λ 轻链可变区和小鼠C κ 区的核酸,所述重排的人 λ 轻链可变区来源于所述人V λ 基因区段和所述人J λ 基因区段,所述核酸是从如权利要求10所述方法产生的经遗传修饰的小鼠中分离。

15. 一种分离的细胞,所述细胞来自如权利要求10所述方法产生的经遗传修饰的小鼠或者可从所述小鼠中获得,其中所述细胞不是胚胎干细胞,且包括人 λ 轻链可变区和小鼠C κ 区。

16. 如权利要求15所述的细胞,其中,所述细胞是B细胞。

17. 一种编码重排的人 λ 轻链可变区和小鼠C κ 区的核酸,所述重排的人 λ 轻链可变区来源于所述人V λ 基因区段和所述人J λ 基因区段,所述核酸是从权利要求16所述的细胞中分离。

18. 一种杂交瘤,所述杂交瘤由骨髓瘤细胞系与来自如权利要求10所述方法产生的经遗传修饰的小鼠的淋巴细胞融合获得,其中所述杂交瘤包括人 λ 轻链可变区和小鼠C κ 区。

19. 一种编码重排的人 λ 轻链可变区和小鼠C κ 区的核酸,所述重排的人 λ 轻链可变区来源于所述人V λ 基因区段和所述人J λ 基因区段,所述核酸是从权利要求18所述的杂交瘤中分离。

20. 使用如权利要求11所述的抗体在生产用作治疗剂的药物中的用途。

21. 一种产生体细胞突变的抗体的方法,所述抗体与目标抗原结合,所述方法包括:

(a) 将如权利要求10所述方法产生的经遗传修饰的小鼠暴露于目标抗原;

(b) 从(a)的小鼠中获得一种或多种B淋巴细胞,其中,所述一种或多种B淋巴细胞产生与目标抗原结合的抗体;和

(c) 鉴定编码(b)的抗体的免疫球蛋白轻链的核酸序列,其中,所述免疫球蛋白轻链包含人免疫球蛋白 λ 轻链可变结构域和小鼠 κ 免疫球蛋白恒定结构域;和

(d) 使用人免疫球蛋白恒定区核酸序列替换(c)的小鼠 κ 恒定区核酸序列来产生与目标抗原结合的人抗体;

其中所述体细胞突变是指未重排的基因区段经重排以形成表达可变结构域的基因。

杂交轻链小鼠

[0001] 本申请是申请号为201180038411.4、申请日为2011年6月22日、发明名称为“杂交轻链小鼠”的中国发明专利申请的分案申请,原申请为国际申请号为PCT/US2011/041370的国家阶段申请,该国际申请要求申请日分别为2010年6月22日和2010年6月22日,申请号分别为61/357,314和61/357,317的美国申请的优先权。

[0002] 领域

[0003] 包含与小鼠或人轻链恒定区(λ 或kappa(κ))有效连接的小鼠或人 λ 可变(V λ)轻链序列的遗传修饰的小鼠。表达包含含有来源于人 λ 可变(hV λ)基因区段的可变结构域、人 λ J(hJ λ)基因区段和小鼠轻链恒定(C λ)结构域的免疫球蛋白轻链的表位结合蛋白的遗传修饰的小鼠。遗传修饰的小鼠,其在内源小鼠轻链基因座上包含未重排的免疫球蛋白lambda(λ)轻链可变核酸序列。能够重排并且从内源轻链基因座表达嵌合人 λ /小鼠C λ 轻链的小鼠,所述内源轻链基因座包括,利用一个或多个hV λ 基因区段以及一个或多个hJ λ 基因区段对所有内源小鼠轻链可变区基因区段的替换。包含hV λ 结构域和小鼠C λ 结构域的体细胞突变的抗体。

[0004] 背景

[0005] 表达为完全人或者部分人及部分小鼠的抗体的小鼠在本领域是已知的。例如,已报导了,从包含人轻链和重链免疫球蛋白可变区基因的转基因表达完全人抗体的转基因小鼠。包括利用人HCVR和LCVR基因区段对内源小鼠重链可变区(HCVR)基因区段和 κ 轻链可变区(LCVR)基因区段的替换并且产生具有嵌合人/小鼠 κ 链的嵌合抗体的遗传修饰的小鼠也是已知的。

[0006] 抗体轻链由两个分离的基因座:kappa(κ)和lambda(λ)之一编码。小鼠抗体轻链主要为 κ 型。人中的 κ 对 λ 轻链使用的比率为约60:40,然而在小鼠中,该比率为约95:5。小鼠中 κ 轻链的偏爱性使用据报导在能够表达完全或部分人抗体的遗传修饰的小鼠中被保留。因此,表达完全或部分人抗体的小鼠显示在 λ 可变使用上受到限制。

[0007] 本领域需要产生用于制备表位结合蛋白的 λ 可变区(无论是小鼠的还是人的)。本领域需要表达完全或部分人抗体的小鼠,其中所述小鼠显示增加的 λ 可变(V λ)使用。

[0008] 本领域需要表达完全或部分人抗体的小鼠,其中所述小鼠显示增加的 λ 可变(V λ)使用。

[0009] 概述

[0010] 提供了遗传修饰的小鼠、胚胎、细胞、组织以及用于修饰小鼠的核酸构建体,和用于产生和使用它们的方法及组合物。提供了在kappa(κ)轻链的背景下产生lambda(λ)可变区(人或非人)的小鼠和细胞。还提供了在 κ 或 λ 轻链的背景下,例如从内源小鼠轻链基因座,产生人 λ 可变区的小鼠和细胞。还提供了,用于产生包含 λ 可变区的抗体的方法。还提供了,用于选择与关联(cognate) λ 可变区一起表达的重链的方法。

[0011] 提供了包括体细胞突变的可变区的嵌合和人抗原结合蛋白(例如,抗体)以及编码其的核酸,包括具有包含融合至小鼠轻链恒定结构域的、来源于人V λ 和人J λ 基因区段的的可变结构域的轻链的抗体。

[0012] 在一个方面,提供了在包含小鼠恒定区的轻链上表达人 λ 可变区序列的小鼠。在一个方面,提供了在包含 κ 恒定区的轻链上表达人 λ 可变区序列的小鼠。在一个方面,提供了从内源小鼠轻链基因座表达包含人 λ 可变区序列的轻链的小鼠。在一个方面,提供了包含重排的轻链基因的小鼠,所述轻链基因包含连接至小鼠恒定区序列的人 λ 可变序列;在一个实施方案中,小鼠恒定区序列为 λ 恒定序列;在一个实施方案中,小鼠恒定区序列为 κ 恒定序列。

[0013] 在一个方面,提供了遗传修饰的小鼠,其中所述小鼠包含未重排的人 λ 轻链可变基因区段(hV λ)和人 λ 连接基因区段(hJ λ)。在一个实施方案中,未重排的hV λ 和hJ λ 在小鼠轻链基因座上。在一个实施方案中,未重排的hV λ 和未重排的hJ λ 在转基因上并且有效地连接至人或小鼠恒定区序列。在一个实施方案中,未重排的hV λ 和未重排的hJ λ 在附加体(episome)上。在一个实施方案中,小鼠能够产生免疫球蛋白,所述免疫球蛋白包含从未重排的hV λ 序列和hJ λ 序列以及小鼠轻链恒定区(C λ)核酸序列衍生的轻链。还提供了用于产生和使用遗传修饰的小鼠的方法和组合物。提供了抗体,所述抗体包含(a)融合至小鼠重链恒定区的人重链可变结构域(hV μ),和(b)融合至小鼠C μ 结构域的人V λ ;包括其中,例如在本发明的小鼠的抗体或免疫细胞的选择过程中,对一个或多个可变结构域进行体细胞突变。在一个实施方案中,未重排的hV λ 和未重排的hJ λ 与人或小鼠 κ 恒定区(C κ)有效地连接。在一个实施方案中,未重排的hV λ 和未重排的hJ λ 与人或小鼠 λ 恒定区(C λ)有效地连接。

[0014] 在一个方面,提供了在其种系中在内源小鼠轻链基因座上包含人 λ 轻链可变区序列的小鼠,其中人 λ 可变区序列在包含小鼠免疫球蛋白恒定区基因序列的轻链中表达。

[0015] 在一个实施方案中,内源小鼠轻链基因座为 λ 基因座。在一个实施方案中,内源小鼠轻链基因座为 κ 基因座。

[0016] 在一个实施方案中,小鼠在内源小鼠轻链基因座上不存在内源轻链可变序列。

[0017] 在一个实施方案中,所有或基本上所有的内源小鼠轻链可变区基因区段被一个或多个人 λ 可变区基因区段替换。

[0018] 在一个实施方案中,人 λ 轻链可变区序列包含人J λ 序列。在一个实施方案中,人J λ 序列选自J λ 1、J λ 2、J λ 3、J λ 7及其组合。

[0019] 在一个实施方案中,人 λ 轻链可变区序列包含人轻链基因座的A簇的片段。在具体的实施方案中,人 λ 轻链基因座的A簇的片段从hV λ 3-27延伸至hV λ 3-1。

[0020] 在一个实施方案中,人 λ 轻链可变区序列包含人轻链基因座的B簇的片段。在具体的实施方案中,人 λ 轻链基因座的B簇的片段从hV λ 5-52延伸至hV λ 1-40。

[0021] 在一个实施方案中,人 λ 轻链可变区序列包含A簇的基因组片段和B簇的基因组片段。在一个实施方案中,人 λ 轻链可变区序列包含A簇的至少一个基因区段和B簇的至少一个基因区段。

[0022] 在一个实施方案中,超过10%的小鼠轻链天然库(**naïve** repertoire)来源于选自2-8、2-23、1-40、5-45和9-49的至少两个hV λ 基因区段。在一个实施方案中,超过20%的小鼠轻链天然库来源于选自2-8、2-23、1-40、5-45和9-49的至少三个hV λ 基因区段。在一个实施方案中,超过30%的小鼠轻链天然库来源于选自2-8、2-23、1-40、5-45和9-49的至少四个hV λ 基因区段。

[0023] 在一个方面,提供了表达免疫球蛋白轻链的小鼠,所述免疫球蛋白轻链包含与小鼠恒定区融合的人 λ 可变序列,其中所述小鼠显示约1:1的 κ 使用对 λ 使用的比率。

- [0024] 在一个实施方案中,从内源小鼠轻链基因座表达免疫球蛋白轻链。
- [0025] 在一个方面,提供了包含与小鼠 κ 轻链恒定区序列连续的 λ 轻链可变区序列(V λ)和至少一个J序列(J)的小鼠。
- [0026] 在一个实施方案中,小鼠缺乏功能性小鼠V κ 和/或小鼠J κ 基因区段。
- [0027] 在一个实施方案中,V λ 为人V λ (hV λ),并且J为人J λ (hJ λ)。在一个实施方案中,hV λ 和hJ λ 为未重排的基因区段。
- [0028] 在一个实施方案中,小鼠包含多个未重排的hV λ 基因区段和至少一个hJ λ 基因区段。在具体的实施方案中,多个未重排的hV λ 基因区段为至少12个基因区段、至少28个基因区段或至少40个基因区段。
- [0029] 在一个实施方案中,至少一个hJ λ 基因区段选自J λ 1、J λ 2、J λ 3、J λ 7及其组合。
- [0030] 在一个实施方案中,内源小鼠 λ 轻链基因座被完整或部分缺失。
- [0031] 在一个实施方案中,小鼠 κ 轻链恒定区序列在内源小鼠 κ 轻链基因座上。
- [0032] 在一个实施方案中,约10%至约45%的小鼠B细胞表达抗体,所述抗体包括包含人 λ 轻链可变(V λ)结构域和小鼠 κ 轻链恒定(C κ)结构域的轻链。
- [0033] 在一个实施方案中,人 λ 可变结构域来源于重排的hV λ /hJ λ 序列,所述hV λ /hJ λ 序列选自3-1/1、3-1/7、4-3/1、4-3/7、2-8/1、3-9/1、3-10/1、3-10/3、3-10/7、2-14/1、3-19/1、2-23/1、3-25/1、1-40/1、1-40/2、1-40/3、1-40/7、7-43/1、7-43/3、1-44/1、1-44/7、5-45/1、5-45/2、5-45/7、7-46/1、7-46/2、7-46/7、9-49/1、9-49/2、9-49/7和1-51/1。
- [0034] 在一个实施方案中,小鼠还包括来自人 κ 轻链基因座的人V κ -J κ 基因间区域,其中人V κ -J κ 基因间区域与V λ 序列和J序列连续。在具体的实施方案中,人V κ -J κ 基因间区域置于V λ 序列与J序列之间。
- [0035] 在一个方面,提供了小鼠,所述小鼠(a)在内源小鼠轻链基因座上包含至少12至至少40个未重排的人 λ 轻链可变区基因区段和至少一个人J λ 基因区段;(b)包含位于至少12至至少40个轻链可变区基因区段与至少一个人J λ 序列之间的人V κ -J κ 基因间序列;其中所述小鼠表达包含含有人V λ 结构域和小鼠C κ 结构域的轻链的抗体。
- [0036] 在一个方面,提供了表达包含含有 λ 可变序列和 κ 恒定序列的轻链的抗体的小鼠。
- [0037] 在一个实施方案中,小鼠显示了约1:1的 κ 使用对 λ 使用的比率。
- [0038] 在一个实施方案中,获自小鼠骨髓的未成熟B细胞群体显示约1:1的 κ 使用对 λ 使用的比率。
- [0039] 在一个方面,提供了遗传修饰的小鼠,其中所述小鼠包含与包含小鼠C λ 基因的小鼠轻链基因座有效连接的未重排的免疫球蛋白V λ 和J λ 基因区段。
- [0040] 在一个实施方案中,V λ 和/或J λ 基因区段为人基因区段。在一个实施方案中,V λ 和/或J λ 基因区段为小鼠基因区段,并且C λ 为小鼠C κ 。
- [0041] 在一个实施方案中,内源小鼠轻链基因座为 κ 轻链基因座。在一个实施方案中,内源小鼠轻链基因座为 λ 轻链基因座。
- [0042] 在一个实施方案中,未重排的V λ 和J λ 基因区段在内源小鼠轻链基因座上。
- [0043] 在一个实施方案中,未重排的免疫球蛋白V λ 和J λ 基因区段在转基因上。
- [0044] 在一个实施方案中,小鼠还在内源小鼠重链免疫球蛋白基因座上包括,一个或多个V、D和/或J基因区段对一个或多个重链V、D和/或J基因区段的替换。

[0045] 在一个实施方案中,小鼠在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段。

[0046] 在一个实施方案中,小鼠在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的人免疫球蛋白 λ 轻链可变基因区段(V λ)和 λ 连接基因区段(J λ)。

[0047] 在一个实施方案中,轻链可变基因基因座(“V λ 基因座”)包含至少一个人V λ (hV λ)基因区段。在一个实施方案中,V λ 基因座包含至少一个人J λ (hJ λ)基因区段。在另一个实施方案中,V λ 基因座包含多至4个hJ λ 基因区段。在一个实施方案中,V λ 基因座包含含有人 λ 和人 κ 基因组序列的连续序列。

[0048] 在一个实施方案中, κ 轻链可变基因基因座(“ κ 基因座”)包含至少一个人V λ (hV λ)基因区段。在一个实施方案中, κ 基因座包含至少一个人J λ (hJ λ)基因区段。在一个实施方案中, κ 基因座包含多至4个hJ λ 基因区段。在一个实施方案中, κ 基因座包含至少一个hV λ 和至少一个hJ λ ,并且不存在或基本上不存在功能性V κ 区基因区段以及不存在或基本上不存在功能性J κ 区基因区段。在一个实施方案中,小鼠不包含功能性V κ 区基因区段。在一个实施方案中,小鼠不包含功能性J κ 区基因区段。

[0049] 在一个实施方案中, λ 轻链可变基因基因座(“ λ 基因座”)包含至少一个hV λ 基因区段。在一个实施方案中, λ 基因座包含至少一个人J λ (hJ λ)基因区段。在另一个实施方案中, λ 基因座包含多至4个hJ λ 基因区段。

[0050] 在一个实施方案中,V λ 基因座包含多个hV λ 。在一个实施方案中,选择多个hV λ ,以导致反映人中观察到的V λ 使用的约10%、约20%、约30%、约40%、约50%、约60%、约70%、约80%或约90%或更多的 λ 轻链可变区库的表达。在一个实施方案中,V λ 基因座包含基因区段hV λ 1-40、1-44、2-8、2-14、3-21及其组合。

[0051] 在一个实施方案中,hV λ 包括3-1、4-3、2-8、3-9、3-10、2-11和3-12。在具体的实施方案中,V λ 基因座包括横跨V λ 3-12至V λ 3-1的人 λ 轻链基因座的连续序列。在一个实施方案中,V λ 基因座包含至少2、3、4、5、6、7、8、9、10、11或12个hV λ 。在具体的实施方案中,hV λ 包括3-1、4-3、2-8、3-9、3-10、2-11和3-12。在具体的实施方案中,V λ 基因座包含横跨V λ 3-12至V λ 3-1的人 λ 基因座的连续序列。在一个实施方案中,V λ 基因座在内源 κ 基因座上。在具体的实施方案中,V λ 基因座在内源 κ 基因座上并且内源 λ 轻链基因座被部分或完全缺失。在一个实施方案中,V λ 基因座在内源 λ 基因座上。在具体的实施方案中,V λ 基因座在内源 λ 基因座上并且内源 κ 基因座被部分或完全缺失。

[0052] 在一个实施方案中,V λ 基因座包括13至28个或更多的hV λ 。在具体的实施方案中,hV λ 包括2-14、3-16、2-18、3-19、3-21、3-22、2-23、3-25和3-27。在具体的实施方案中, κ 基因座包含横跨V λ 3-27至V λ 3-1的人 λ 基因座的连续序列。在一个实施方案中,V λ 基因座在内源 κ 基因座上。在具体的实施方案中,V λ 基因座在内源 κ 基因座上并且内源 λ 轻链基因座被部分或完全缺失。在另一个实施方案中,V λ 基因座在内源 λ 基因座上。在具体的实施方案中,V λ 基因座在内源 λ 基因座上并且内源 κ 基因座被部分或完全缺失。

[0053] 在一个实施方案中,V λ 基因座包含29至40个hV λ 。在具体的实施方案中, κ 基因座包含横跨V λ 3-29至V λ 3-1的人 λ 基因座的连续序列,横跨V λ 5-52至V λ 1-40的人 λ 基因座的连续序列。在具体的实施方案中,遗传修饰的小鼠中的hV λ 1-40与hV λ 3-29之间的所有或基本上所有序列基本上由下述组成:天然发现的(例如,在人群中)在hV λ 1-40基因区段下游(3'非

翻译部分的下游)的约959 bp的人 λ 序列、限制性酶位点(例如,PI-SceI)、紧接着天然发现的hV λ 3-29基因区段的上游的约3,431 bp的人 λ 序列。在一个实施方案中,V λ 基因座在内源小鼠 κ 基因座上。在具体的实施方案中,V λ 基因座在内源小鼠 κ 基因座上并且内源小鼠 λ 轻链基因座被部分或完全缺失。在另一个实施方案中,V λ 基因座在内源小鼠 λ 基因座上。在具体的实施方案中,V λ 基因座在内源小鼠 λ 基因座上并且内源小鼠 κ 基因座被部分或完全缺失。

[0054] 在一个实施方案中,V λ 基因座包含至少一个hJ λ 。在一个实施方案中,V λ 基因座包含多个hJ λ 。在一个实施方案中,V λ 基因座包含至少2、3、4、5、6或7个hJ λ 。在具体的实施方案中,V λ 基因座包含4个hJ λ 。在具体的实施方案中,4个hJ λ 为hJ λ 1、hJ λ 2、hJ λ 3和hJ λ 7。在一个实施方案中,V λ 基因座为 κ 基因座。在具体的实施方案中,V λ 基因座在内源 κ 基因座上并且内源 λ 轻链基因座被部分或完全缺失。在一个实施方案中,V λ 基因座包含一个hJ λ 。在具体的实施方案中,一个hJ λ 为hJ λ 1。在一个实施方案中,V λ 基因座在内源 κ 基因座上。在具体的实施方案中,V λ 基因座在内源 κ 基因座上并且内源 λ 轻链基因座被部分或完全缺失。在另一个实施方案中,V λ 基因座在内源 λ 基因座上。在具体的实施方案中,V λ 基因座在内源 λ 基因座上并且内源 κ 基因座被部分或完全缺失。

[0055] 在一个实施方案中,V λ 基因座包含至少一个hV λ 、至少一个hJ λ 和小鼠C κ 基因。在一个实施方案中,V λ 基因座包含至少一个hV λ 、至少一个hJ λ 和小鼠C λ 基因。在具体的实施方案中,小鼠C λ 基因为C λ 2。在具体的实施方案中,小鼠C λ 基因与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%、96%、97%、98%或至少99%的同一性。

[0056] 在一个实施方案中,小鼠在内源小鼠 κ 基因座上包括,利用一个或多个hV λ 基因区段对内源小鼠V κ 基因区段的替换,其中hV λ 基因区段有效地连接至内源小鼠C κ 区基因,以便小鼠重排人V λ 基因区段并且表达包含人V λ 结构域和小鼠C κ 的反向嵌合免疫球蛋白轻链。在一个实施方案中,90-100%的未重排的小鼠V κ 基因区段被至少一个未重排的hV λ 基因区段替换。在具体的实施方案中,所有或基本上所有内源小鼠V κ 基因区段被至少一个未重排的hV λ 基因区段替换。在一个实施方案中,利用至少12、至少28或至少40个未重排的hV λ 基因区段进行替换。在一个实施方案中,利用至少7个功能性未重排的hV λ 基因区段、至少16个功能性未重排的hV λ 基因区段或至少27个功能性未重排的hV λ 基因区段进行替换。在一个实施方案中,小鼠包括利用至少一个未重排的hJ λ 基因区段对所有小鼠J κ 基因区段的替换。在一个实施方案中,至少一个未重排的hJ λ 基因区段选自J λ 1、J λ 2、J λ 3、J λ 4、J λ 5、J λ 6、J λ 7及其组合。在具体的实施方案中,一个或多个hV λ 基因区段选自3-1、4-3、2-8、3-9、3-10、2-11、3-12、2-14、3-16、2-18、3-19、3-21、3-22、2-23、3-25、3-27、1-40、7-43、1-44、5-45、7-46、1-47、5-48、9-49、1-50、1-51、5-52 hV λ 基因区段及其组合。在具体的实施方案中,至少一个未重排的hJ λ 基因区段选自J λ 1、J λ 2、J λ 3、J λ 7及其组合。

[0057] 在一个实施方案中,小鼠在内源小鼠 λ 基因座上包括,利用一个或多个人V λ 基因区段对内源小鼠V λ 基因区段的替换,其中hV λ 基因区段有效地连接至小鼠C λ 区基因,以便小鼠重排hV λ 基因区段并且表达包含hV λ 结构域和小鼠C λ 的反向嵌合免疫球蛋白轻链。在具体的实施方案中,小鼠C λ 基因为C λ 2。在具体的实施方案中,小鼠C λ 基因与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性。在一个实施方案中,90-100%的未重排的小鼠V λ 基因区段被至少一个未重排的hV λ 基因区段替换。在具体的实施方案中,所有或基本上所有内源小鼠V λ 基因区段被至少一个未重排的hV λ 基因区段替换。

在一个实施方案中,利用至少12、至少28或至少40个未重排的hV λ 基因区段进行替换。在一个实施方案中,利用至少7个功能性未重排的hV λ 基因区段、至少16个功能性未重排的hV λ 基因区段或至少27个功能性未重排的hV λ 基因区段进行替换。在一个实施方案中,小鼠包括利用至少一个未重排的hJ λ 基因区段对所有小鼠J λ 基因区段的替换。在一个实施方案中,至少一个未重排的hJ λ 基因区段选自J λ 1、J λ 2、J λ 3、J λ 4、J λ 5、J λ 6、J λ 7及其组合。在具体的实施方案中,一个或多个hV λ 基因区段选自3-1、4-3、2-8、3-9、3-10、2-11、3-12、2-14、3-16、2-18、3-19、3-21、3-22、2-23、3-25、3-27、1-40、7-43、1-44、5-45、7-46、1-47、5-48、9-49、1-50、1-51、5-52 hV λ 基因区段及其组合。在具体的实施方案中,至少一个未重排的hJ λ 基因区段选自J λ 1、J λ 2、J λ 3、J λ 7及其组合。

[0058] 在一个方面,提供了包含位于内源小鼠 κ 轻链基因座上的人V κ -J κ 基因间区域序列的遗传修饰的小鼠。

[0059] 在一个实施方案中,人V κ -J κ 基因间区域序列在小鼠的内源 κ 轻链基因座上,所述基因座包含hV λ 和hJ λ 基因区段,并且人V κ -J κ 基因间区域序列位于hV λ 与hJ λ 基因区段之间。在具体的实施方案中,hV λ 和hJ λ 基因区段能够在小鼠中重组以形成功能性人 λ 轻链可变结构域。

[0060] 在一个实施方案中,提供了包含多个hV λ 和一个或多个hJ λ 的小鼠,并且人V κ -J κ 基因间区域序列在转录方向上被置于邻近的或最3'端的hV λ 序列的下游和第一个hJ λ 序列的上游或5'端。

[0061] 在一个实施方案中,人V κ -J κ 基因间区域为这样的区域,其位于人V κ 4-1基因区段的下游或3'端约130 bp,人V κ 4-1基因区段的3'非翻译区下游约130 bp,并且横跨至人J κ 1基因区段的上游或5'端约600 bp。在具体的实施方案中,人V κ -J κ 基因间区域在大小上为约22.8 kb。在一个实施方案中,V κ -J κ 基因间区域与从人V κ 4-1基因区段的3'非翻译区的末端延伸至人J κ 1基因区段的上游约600bp的人V κ -J κ 基因间区域具有约90%或更多、91%或更多、92%或更多、93%或更多、94%或更多或约95%或更多的同一性。在一个实施方案中,V κ -J κ 基因间区域包含SEQ ID NO:100。在具体的实施方案中,V κ -J κ 基因间区域包含SEQ ID NO:100的功能性片段。在具体的实施方案中,V κ -J κ 基因间区域为SEQ ID NO:100。

[0062] 在一个方面,提供了包含所述的人V κ -J κ 基因间区域序列的小鼠、小鼠细胞(例如,小鼠胚胎干细胞)、小鼠胚胎和小鼠组织,其中基因间区域序列是异位的。在具体的实施方案中,异位序列置于人源化内源小鼠免疫球蛋白基因座上。

[0063] 在一个方面,提供了包含所述的人V κ -J κ 基因间区域序列的分离的核酸构建体。在一个实施方案中,核酸构建体包含将人V κ -J κ 基因间区域序列靶向小鼠轻链基因座的靶向臂。在具体的实施方案中,小鼠轻链基因座为 κ 基因座。在具体的实施方案中,靶向臂将人V κ -J κ 基因间区域靶向修饰的内源小鼠 κ 基因座,其中靶向是至hV λ 序列与hJ λ 序列之间的位置。

[0064] 在一个方面,提供了遗传修饰的小鼠,其中小鼠包含不超过2个的轻链等位基因,其中轻链等位基因(a)在包含小鼠C μ 基因的内源小鼠轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段;以及,(b)在包含小鼠C μ 基因的内源小鼠轻链基因座上包含未重排的免疫球蛋白V μ 和J μ 基因区段。

[0065] 在一个实施方案中,内源小鼠轻链基因座为 κ 基因座。在另一个实施方案中,内源

小鼠轻链基因座为 λ 基因座。

[0066] 在一个实施方案中,不超过2个的轻链等位基因选自 κ 等位基因和 λ 等位基因、2个 κ 等位基因和2个 λ 等位基因。在具体的实施方案中,2个轻链等位基因之一为包含C λ 2基因的 λ 等位基因。

[0067] 在一个实施方案中,小鼠包含一个功能性免疫球蛋白轻链基因座和一个非功能性轻链基因座,其中所述功能性轻链基因座在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座上包含未重排的免疫球蛋白人V λ 和J λ 基因区段。

[0068] 在一个实施方案中,小鼠包含一个功能性免疫球蛋白轻链基因座和一个非功能性轻链基因座,其中所述功能性轻链基因座在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的免疫球蛋白人V λ 和J λ 基因区段。在一个实施方案中,C λ 基因为C λ 2。在具体的实施方案中,小鼠C λ 基因与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性。

[0069] 在一个实施方案中,小鼠还包含至少一个免疫球蛋白重链等位基因。在一个实施方案中,至少一个免疫球蛋白重链等位基因在表达人/小鼠重链的包含人重链基因的内源小鼠重链基因座上包含人V H 基因区段、人D H 基因区段和人J H 基因区段。在具体的实施方案中,小鼠包含2个免疫球蛋白重链等位基因,并且小鼠表达人/小鼠重链。

[0070] 在一个实施方案中,小鼠包含第一轻链等位基因,所述第一轻链等位基因在包含内源C κ 基因的内源小鼠 κ 基因座上包含未重排的hV κ 和未重排的hJ κ ;和第二轻链等位基因,所述第二轻链等位基因在包含内源C λ 基因的内源小鼠 λ 基因座上包含未重排的hV λ 和未重排的hJ λ 。在具体的实施方案中,第一和第二轻链等位基因为遗传修饰的小鼠的仅有的功能性轻链等位基因。在具体的实施方案中,小鼠包含非功能性 λ 基因座。在一个实施方案中,遗传修饰的小鼠不表达包含 λ 恒定区的轻链。

[0071] 在一个实施方案中,小鼠包含第一轻链等位基因,所述第一轻链等位基因在包含内源C κ 基因的内源小鼠 κ 基因座上包含未重排的hV κ 和未重排的hJ κ ;和第二轻链等位基因,所述第二轻链等位基因在包含内源C λ 基因的内源小鼠 λ 基因座上包含未重排的hV λ 和未重排的hJ λ 。在具体的实施方案中,第一和第二轻链等位基因是遗传修饰的小鼠的仅有的功能性轻链等位基因。在一个实施方案中,内源C λ 基因为C λ 2。在具体的实施方案中,小鼠C λ 基因与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性。

[0072] 在一个实施方案中,小鼠包含6个免疫球蛋白等位基因,其中第一等位基因在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座包含未重排的免疫球蛋白V λ 和J λ 基因区段,第二等位基因在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座上包含未重排的免疫球蛋白V κ 和J κ 基因区段,第三等位基因在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段,第四和第五等位基因各自独立地在包含小鼠重链基因的内源小鼠重链基因座上包含未重排的V H 和D H 以及J H 基因区段,并且第六等位基因(a)在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段,(b)包含无功能的 λ 基因座,或(c)完全或部分缺失 λ 基因座。

[0073] 在一个实施方案中,第一等位基因包含未重排的hV λ 和hJ λ 。在一个实施方案中,第二等位基因包含未重排的hV κ 和hJ κ 。在一个实施方案中,第三等位基因包含未重排的hV λ 和hJ λ 。在一个实施方案中,第四和第五等位基因各自独立地包含未重排的hV H 和hD H 以及hJ H 。

在一个实施方案中,第六等位基因包含被完全或部分缺失的内源小鼠 λ 基因座。

[0074] 在一个实施方案中,小鼠包含6个免疫球蛋白等位基因,其中第一等位基因在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段,第二等位基因在包含小鼠C λ 基因的内源小鼠 λ 轻链基因座上包含未重排的免疫球蛋白V λ 和J λ 基因区段,第三等位基因在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座上包含未重排的免疫球蛋白V κ 和J κ 基因区段,第四和第五等位基因各自独立地在包含小鼠重链基因的内源小鼠重链基因座上包含未重排的V H 和D H 以及J H 基因区段,并且第六等位基因(a)在包含小鼠C κ 基因的内源小鼠 κ 轻链基因座上包含未重排的免疫球蛋白V κ 和J κ 基因区段,(b)包含无功能的 κ 基因座,或(c)缺失 κ 基因座的一个或多个元件。

[0075] 在一个实施方案中,第一等位基因包含未重排的hV λ 和hJ λ 基因区段。在一个实施方案中,第二等位基因包含未重排的hV λ 和hJ λ 基因区段。在一个实施方案中,第三等位基因包含未重排的hV κ 和hJ κ 基因区段。在一个实施方案中,第四和第五等位基因各自独立地包含未重排的hV H 和hD H 以及hJ H 基因区段。在一个实施方案中,第六等位基因包含在功能上被沉默的内源小鼠 κ 基因座。

[0076] 在一个实施方案中,遗传修饰的小鼠包括包含重排的抗体基因的B细胞,所述抗体基因包含有效地连接至小鼠C L 结构域的重排的hV λ 结构域。在一个实施方案中,小鼠C L 结构域选自小鼠C κ 和小鼠C λ 结构域。在具体的实施方案中,小鼠C λ 结构域来源于C λ 2基因。在具体的实施方案中,小鼠C λ 结构域来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 结构域。

[0077] 在一个方面,提供了在为C κ 的C L 上表达V λ 区的遗传修饰的小鼠。在一个方面,提供了在选自人C κ 、人C λ 或小鼠C κ 的C L 上表达hV λ 区的遗传修饰的小鼠。在一个方面,提供了在小鼠C κ 上表达hV λ 区的遗传修饰的小鼠。

[0078] 在一个实施方案中,约10-50%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约9-28%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0079] 在具体的实施方案中,约23-34%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约9-11%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0080] 在具体的实施方案中,约19-31%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约9-17%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0081] 在具体的实施方案中,约21-38%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约24-27%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0082] 在具体的实施方案中,约10-14%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约9-13%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0083] 在具体的实施方案中,约31-48%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约15-21%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。在具体的实施方案中,约30-38%的小鼠脾细胞为B细胞(即,CD19-阳性),或其约33-48%表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0084] 在一个实施方案中,约52-70%的小鼠骨髓为B细胞(即,CD19-阳性),或其约31-47%的未成熟B细胞(即,CD19-阳性/B220-中间阳性(intermediate positive)/IgM-阳性)表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0085] 在一个实施方案中,约60%的小鼠骨髓为B细胞(即,CD19-阳性),或其约38.3%的未成熟B细胞(即,CD19-阳性/B220中间阳性/IgM-阳性)表达包含融合至小鼠C κ 结构域的hV λ 结构域的免疫球蛋白轻链。

[0086] 在一个实施方案中,小鼠表达包含轻链的抗体,所述轻链包含来源于人V和人J基因区段的可变结构域,以及来源于小鼠恒定区基因的恒定结构域。在一个实施方案中,小鼠恒定区基因为C κ 基因。在另一个实施方案中,小鼠恒定区基因为C λ 基因。在具体的实施方案中,C λ 区为C λ 2。在具体的实施方案中,小鼠C λ 基因来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 基因。在具体的实施方案中,抗体还包括包含来源于人V、人D和人J基因区段的可变结构域以及来源于小鼠重链恒定区基因的重链恒定结构域的重链。在一个实施方案中,小鼠重链恒定区基因包含重链恒定结构域的铰链-CH₂-CH₃序列。在另一个实施方案中,小鼠重链恒定区基因包含重链恒定结构域的CH₁-铰链-CH₂-CH₃序列。在另一个实施方案中,小鼠重链恒定区基因包含重链恒定结构域的CH₁-CH₂-CH₃-CH₄序列。在另一个实施方案中,小鼠重链恒定区基因包含重链恒定结构域的CH₂-CH₃-CH₄序列。

[0087] 在一个实施方案中,小鼠表达包含含有重排的人V λ -J λ 序列和小鼠C κ 序列的轻链的抗体。在一个实施方案中,重排的人V λ -J λ 序列来源于选自3-1、4-3、2-8、3-9、3-10、2-14、3-19、2-23、3-25、1-40、7-43、1-44、5-45、7-46、1-47、9-49和1-51基因区段的hV λ 基因区段的重排。在一个实施方案中,重排的人V λ -J λ 序列来源于选自J λ 1、J λ 2、J λ 3和J λ 7基因区段的hJ λ 基因区段的重排。

[0088] 在一个实施方案中,小鼠表达包含含有重排的免疫球蛋白 λ 轻链可变区的轻链的抗体,所述重排的免疫球蛋白 λ 轻链可变区包含选自3-1/1、3-1/7、4-3/1、4-3/7、2-8/1、3-9/1、3-10/1、3-10/3、3-10/7、2-14/1、3-19/1、2-23/1、3-25/1、1-40/1、1-40/2、1-40/3、1-40/7、7-43/1、7-43/3、1-44/1、1-44/7、5-45/1、5-45/2、5-45/7、7-46/1、7-46/2、7-46/7、9-49/1、9-49/2、9-49/7和1-51/1的人V λ /J λ 序列。在具体的实施方案中,B细胞表达抗体,所述抗体包含与小鼠重链恒定结构域融合的人免疫球蛋白重链可变结构域,以及与小鼠 κ 轻链恒定结构域融合的人免疫球蛋白 λ 轻链可变结构域。

[0089] 在一个方面,提供了表达抗体的小鼠,所述抗体包含,(a)含有来源于未重排的人重链可变区基因区段的重链可变结构域的重链,其中重链可变结构域融合至小鼠重链恒定(C_H)区;和(b)含有来源于未重排的hV λ 和hJ λ 的轻链可变结构域的轻链,其中轻链可变结构域融合至小鼠C_L区。

[0090] 在一个实施方案中,小鼠包含,(i)重链基因座,其包括用所有或基本上所有功能性人V、D和J基因区段对所有或基本上所有功能性内源小鼠V、D和J基因区段的替换,小鼠C_H基因,(ii)第一 κ 轻链基因座,其包括用所有、基本上所有或多个功能性hV λ 和hJ λ 基因区段对所有或基本上所有功能性内源小鼠V κ 和J κ 基因区段的替换,以及小鼠C κ 基因,(iii)第二 κ 轻链基因座,其包括用所有、基本上所有或多个功能性hV κ 和hJ κ 基因区段对所有或基本上所有功能性内源小鼠V κ 和J κ 基因区段的替换,以及小鼠C κ 基因。在一个实施方案中,小鼠不表达包含C λ 区的抗体。在一个实施方案中,小鼠包括C λ 基因和/或V λ 和/或J λ 基因区段的缺失。在一个实施方案中,小鼠包括非功能性 λ 轻链基因座。在具体的实施方案中, λ 轻链基因座被全部或部分缺失。

[0091] 在一个实施方案中,小鼠包含(i)重链基因座,其包括用所有或基本上所有功能性人V、D和J基因区段对所有或基本上所有功能性内源小鼠V、D和J基因区段的替换、小鼠C_H基因,(ii)第一λ轻链基因座,其包括用所有、基本上所有或多个功能性hVλ和hJλ基因区段对所有或基本上所有功能性内源小鼠Vλ和Jλ基因区段的替换,以及小鼠Cλ基因,(iii)第二λ轻链基因座,其包括用所有、基本上所有或多个功能性hVλ和hJλ基因区段对所有或基本上所有功能性内源小鼠Vλ和Jλ基因区段的替换,以及小鼠Cλ基因。在具体的实施方案中,小鼠Cλ基因为Cλ2。在具体的实施方案中,小鼠Cλ基因来源于与小鼠Cλ2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的Cλ基因。

[0092] 在一个实施方案中,小鼠包括C_κ基因和/或V_κ和/或J_κ基因区段的缺失。在一个实施方案中,小鼠包括非功能性κ轻链基因座。

[0093] 在一个方面,提供了表达抗体的遗传修饰的小鼠,其中超过10%、超过15%、超过20%、超过25%、超过30%、超过35%、超过40%、超过60%、超过70%、超过80%或超过90%的由小鼠产生的总IgG抗体包含λ-来源的可变结构域,并且其中小鼠表达包含与小鼠C_κ区融合的κ来源的可变结构域的抗体。在具体实施方案中,约15-40%、20-40%、25-40%、30-40%或35-40%的由小鼠产生的总抗体包含λ来源的可变结构域。

[0094] 在一个实施方案中,λ来源的可变结构域来源于hVλ和hJλ。在一个实施方案中,λ来源的可变结构域存在于包含小鼠C_κ区的轻链中。在具体的实施方案中,λ来源的可变区存在于包含小鼠Cλ区的轻链中。在另一个具体的实施方案中,Cλ区为Cλ2区。在一个实施方案中,κ来源的可变结构域来源于hV_κ和hJ_κ,并且在具体的实施方案中存在于包含小鼠C_κ区的轻链中。

[0095] 在一个方面,提供了包含上游同源臂和下游同源臂的分离的DNA构建体,其中上游和下游同源臂将所述构建体靶向小鼠κ基因座,并且构建体包含功能性未重排的hVλ区段和功能性未重排的hJλ区段,以及选择或标记序列。

[0096] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含用于靶向小鼠Vλ2上游的小鼠λ序列的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒,以及用于靶向小鼠Jλ2的3'端的小鼠λ序列的靶向臂。在一个实施方案中,选择盒为Frt'ed Hyg-TK盒。在一个实施方案中,3'靶向臂包含小鼠Cλ2、Jλ4、Cλ4和小鼠增强子2.4。

[0097] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含:用于靶向Vλ1的5'端的小鼠λ基因座的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒以及用于靶向小鼠Cλ1的3'端的小鼠λ序列的3'靶向臂。在一个实施方案中,选择盒为loxed新霉素盒。在一个实施方案中,3'靶向臂包含小鼠λ3'增强子和小鼠λ3'增强子3.1。

[0098] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含用于靶向Vλ2的5'端的小鼠λ基因座的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒以及用于靶向小鼠Jλ2的3'端与小鼠Cλ2的5'端的小鼠λ序列的3'靶向臂。在一个实施方案中,选择盒为Frt'ed潮霉素-TK盒。在一个实施方案中,3'靶向臂包含小鼠Cλ2-Jλ4-Cλ4基因区段和小鼠λ增强子2.4。

[0099] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含用于靶向Vλ2的5'端的小鼠λ基因座的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒、包含从hVλ3-12下游至hJλ1的末端的人λ轻链基因座的连续区域的人基因组片段,以及用于靶向小

鼠J λ 2的3'端的小鼠 λ 序列的3'靶向臂。在一个实施方案中,选择盒为Frt'ed新霉素盒。在一个实施方案中,3'靶向臂包含小鼠C λ 2-J λ 4-C λ 4基因区段和小鼠 λ 增强子2.4。

[0100] 在一个方面,提供了分离的DNA构建体,其包含从hV λ 3-12下游至hJ λ 1的末端的人 λ 轻链基因座的连续区域。

[0101] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含用于靶向V λ 2的5'端的小鼠 λ 基因座的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒,以及包含从hV λ 3-27下游至hV λ 2-8的末端的人 λ 轻链基因座的连续区域的人基因组片段。在一个实施方案中,选择盒为Frt'ed潮霉素盒。在一个实施方案中,人基因组片段包含3'靶向臂。在具体的实施方案中,3'靶向臂包含约53 kb的从hV λ 3-12下游至hV λ 2-8的末端的人 λ 轻链基因座。

[0102] 在一个方面,提供了分离的DNA构建体,其包含从hV λ 3-27下游至hV λ 3-12的末端的人 λ 轻链基因座的连续区域。

[0103] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含:用于靶向V λ 2的5'端的小鼠 λ 基因座的靶向臂、5'和3'侧翼连接有重组酶识别位点的选择盒、包含从hV λ 5-52下游至hV λ 1-40的末端的人 λ 轻链基因座的连续区域的第一人基因组片段、限制性酶位点以及包含从hV λ 3-29下游至hV λ 82K的末端的人 λ 轻链基因座的连续区域的第二人基因组片段。在一个实施方案中,选择盒为Frt'ed新霉素盒。在一个实施方案中,限制性酶位点为归巢内切核酸酶的位点。在具体的实施方案中,归巢内切核酸酶为PI-SceI。在一个实施方案中,第二人基因组片段为3'靶向臂。在具体的实施方案中,3'靶向臂包含约27 kb的从hV λ 3-29下游至hV λ 82K的末端的人 λ 轻链基因座。

[0104] 在一个方面,提供了分离的DNA构建体,其包含从hV λ 5-52下游至hV λ 1-40的末端的人 λ 轻链基因座的连续区域。

[0105] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含:用于靶向内源V κ 基因区段的5'端的小鼠 κ 基因座的靶向臂、两个并置的重组酶识别位点、并置的重组酶识别位点的3'端的选择盒以及用于靶向 κ 轻链可变基因区段的5'端的小鼠 κ 序列的3'靶向臂。在一个实施方案中,并置的重组酶识别位点彼此以相反的方向放置。在具体的实施方案中,重组酶识别位点是不同的。在另一个具体的实施方案中,重组酶识别位点为loxP位点和lox511位点。在一个实施方案中,选择盒为新霉素盒。

[0106] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含:用于靶向小鼠J κ 基因区段的5'端的小鼠 κ 基因座的靶向臂、选择盒、选择盒3'端的重组酶识别位点以及用于靶向小鼠J κ 基因区段的3'端与小鼠 κ 内含增强子(intronic enhancer)的5'端的小鼠 κ 序列的3'靶向臂。在一个实施方案中,选择盒为潮霉素-TK盒。在一个实施方案中,重组酶识别位点在转录上处于与选择盒相同的方向。在具体的实施方案中,重组酶识别位点为loxP位点。

[0107] 在一个方面,提供了分离的DNA构建体,其在转录的方向上从5'至3'包含:包含内源小鼠V κ 基因区段的5'端的序列的第一小鼠基因组片段、第一重组酶识别位点、第二重组酶识别位点以及包含内源小鼠J κ 基因区段的3'端与小鼠 κ 内含增强子的5'端之间的序列的第二小鼠基因组片段。

[0108] 在一个方面,提供了遗传修饰的小鼠,其中遗传修饰包括使用上述或本文中描述

的一个或多个DNA构建体进行的修饰。

[0109] 在一个方面,提供了分离的DNA构建体用于制备本文中描述的小鼠的用途。在一个方面,提供了本文中描述的分离的DNA构建体用于产生抗原结合蛋白的方法的用途。

[0110] 在一个方面,提供了非人干细胞,其包括包含上述和本文中描述的DNA构建体的靶向载体。在一个方面,提供了非人干细胞,其中非人干细胞来源于本文中描述的小鼠。

[0111] 在一个实施方案中,非人干细胞为胚胎干(ES)细胞。在具体的实施方案中,ES细胞为小鼠ES细胞。

[0112] 在一个方面,提供了本文中描述的非人干细胞用于产生本文中描述的小鼠的用途。在一个方面,提供了本文中描述的非人干细胞用于产生抗原结合蛋白的用途。

[0113] 在一个方面,提供了小鼠胚胎,其中小鼠胚胎包含本文中提供的遗传修饰。在一个实施方案中,提供了包含供体ES细胞的宿主小鼠胚胎,其中供体ES细胞包含本文中描述的遗传修饰。在一个实施方案中,小鼠胚胎为前桑椹胚期(pre-morula stage)胚胎。在具体的实施方案中,前桑椹胚期胚胎为4细胞期胚胎或8细胞期胚胎。在另一个具体实施方案中,小鼠胚胎为胚泡。

[0114] 在一个方面,提供了本文中描述的小鼠胚胎用于产生本文中描述的小鼠的用途。在一个方面,提供了本文中描述的小鼠胚胎用于产生抗原结合蛋白的用途。

[0115] 在一个方面,提供了非人细胞,其中非人细胞包含来源于本文中描述的遗传修饰的小鼠的重排的免疫球蛋白轻链基因序列。在一个实施方案中,细胞为B细胞。在一个实施方案中,细胞为杂交瘤。在一个实施方案中,所述细胞编码体细胞突变的免疫球蛋白轻链可变结构域和/或免疫球蛋白重链可变结构域。

[0116] 在一个方面,提供了非人细胞,其中非人细胞包含来源于本文中描述的遗传修饰的小鼠的重排的免疫球蛋白轻链基因序列。在一个实施方案中,细胞为B细胞。在一个实施方案中,细胞为杂交瘤。在一个实施方案中,所述细胞编码体细胞突变的免疫球蛋白轻链可变结构域和/或免疫球蛋白重链可变结构域。

[0117] 在一个方面,提供了本文中描述的非人细胞用于产生本文中描述的小鼠的用途。在一个方面,提供了本文中描述的非人细胞用于产生抗原结合蛋白的用途。

[0118] 在一个方面,提供了表达免疫球蛋白轻链的小鼠B细胞,所述免疫球蛋白轻链包含(a) 来源于hV λ 基因区段和hJ λ 基因区段的可变区;和(b) 小鼠C λ 基因。在一个实施方案中,小鼠C λ 基因选自C κ 和C λ 基因。在具体的实施方案中,C λ 基因为C λ 2。在具体的实施方案中,小鼠C λ 基因来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 基因。在一个实施方案中,小鼠B细胞还表达关联重链,所述重链包含(c) 来源于hV μ ,hD μ ,和(d) hJ μ 区段的可变区。在一个实施方案中,B细胞不包含重排的 λ 基因。在另一个实施方案中,B细胞不包含重排的 κ 基因。

[0119] 在一个方面,提供了用于在遗传修饰的小鼠中产生抗体的方法,包括:(a) 将遗传修饰的小鼠暴露于抗原,其中小鼠具有在内源轻链基因座上包含至少一个hV λ 和至少一个hJ λ 的基因组,其中内源轻链基因座包含小鼠C λ 基因;(b) 允许遗传修饰的小鼠产生对抗原的免疫应答;和(c) 从(b)的小鼠分离特异性识别抗原的抗体,或从(b)的小鼠分离包含特异性识别抗原的免疫球蛋白结构域的细胞,其中抗体包含来源于hV λ 、hJ λ 和小鼠C λ 基因的轻链。在具体的实施方案中,小鼠C λ 基因为小鼠C κ 基因。

[0120] 在一个实施方案中,提供了用于在遗传修饰的小鼠中产生抗体的方法,包括:(a)将遗传修饰的小鼠暴露于抗原,其中小鼠具有在内源 κ 基因座上包含至少一个hV λ 和在 κ 基因座上包含至少一个hJ λ 的基因组,其中 κ 基因座包含小鼠C κ 基因;(b)允许遗传修饰的小鼠产生对抗原的免疫应答;和(c)从(b)的小鼠分离特异性识别抗原的抗体,或从(b)的小鼠分离包含特异性识别抗原的免疫球蛋白结构域的细胞,其中抗体包含来源于hV λ 、hJ λ 和小鼠C κ 基因的轻链。

[0121] 在一个实施方案中, κ 轻链恒定基因选自人C κ 基因和小鼠C κ 基因。

[0122] 在一个实施方案中,提供了用于在遗传修饰的小鼠中产生抗体的方法,包括:(a)将遗传修饰的小鼠暴露于抗原,其中小鼠具有在 λ 轻链基因座上包含至少一个hV λ 和在 λ 轻链基因座上包含至少一个J λ 的基因组,其中 λ 轻链基因座包含小鼠C λ 基因;(b)允许遗传修饰的小鼠产生对抗原的免疫应答;和(c)从(b)的小鼠分离特异性识别抗原的抗体,或从(b)的小鼠分离包含特异性识别抗原的免疫球蛋白结构域的细胞,或在B的小鼠中鉴定编码结合抗原的重链和/或轻链可变结构域的核酸序列,其中抗体包含来源于hV λ 、hJ λ 和小鼠C λ 基因的轻链。

[0123] 在一个实施方案中, λ 轻链恒定基因选自人C λ 基因和小鼠C λ 基因。在一个实施方案中, λ 轻链恒定基因为人C λ 基因。在具体的实施方案中,人C λ 基因选自C λ 1、C λ 2、C λ 3和C λ 7。在一个实施方案中, λ 轻链恒定基因为小鼠C λ 基因。在具体的实施方案中,小鼠C λ 基因选自C λ 1、C λ 2和C λ 3。在具体的实施方案中,小鼠C λ 基因为C λ 2。在另一个具体的实施方案中,小鼠C λ 基因来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 基因。

[0124] 在一个方面,提供了用于在遗传修饰的小鼠中产生重排的抗体基因的方法,包括:(a)将遗传修饰的小鼠暴露于抗原,其中所述遗传修饰在内源轻链基因座上包含hV λ 和hJ λ ,其中内源轻链基因座包含小鼠C λ 基因或其功能性片段;和,(b)在所述小鼠中鉴定重排的免疫球蛋白基因,其中所述重排的免疫球蛋白基因包含 λ 轻链可变区基因区段和C λ 基因或其功能性片段。

[0125] 在一个实施方案中,该方法还包括从小鼠克隆编码重链和/或轻链可变区的核酸序列,其中重链和/或轻链可变区来自包含人V λ 和小鼠C λ 的抗体。

[0126] 在一个实施方案中,小鼠C λ 基因或其功能性片段选自人C λ 基因和小鼠C λ 基因,或其功能性片段。

[0127] 在一个实施方案中,提供了用于在遗传修饰的小鼠中产生重排的抗体基因的方法,包括:(a)将遗传修饰的小鼠暴露于抗原,其中所述遗传修饰包括在 κ 轻链基因座上包含hV λ 和hJ λ ,其中 κ 轻链基因座包含小鼠C κ 基因或其功能性片段;和(b)在所述小鼠中鉴定重排的免疫球蛋白基因,其中重排的免疫球蛋白基因包含 λ 轻链可变区基因区段和C κ 基因或其功能性片段。

[0128] 在一个实施方案中, κ 轻链恒定基因或其功能性片段选自人C κ 基因和小鼠C κ 基因或其功能性片段。

[0129] 在一个实施方案中,该方法还包括从小鼠克隆编码重链和/或轻链可变区的核酸序列,其中重链和/或轻链可变区来自包含人V λ 和小鼠C κ 的抗体。

[0130] 在一个实施方案中,提供了用于在遗传修饰的小鼠中产生重排的抗体基因的方

法,包括:(a)将遗传修饰的小鼠暴露于抗原,其中所述遗传修饰在小鼠 λ 轻链基因座上包含hV λ 和hJ λ ,其中 λ 轻链基因座包含小鼠C λ 基因或其功能性片段;和,(b)在所述小鼠中鉴定重排的免疫球蛋白基因,其中重排的免疫球蛋白基因包含 λ 轻链可变区基因区段和C λ 基因或其功能性片段。

[0131] 在一个实施方案中, λ 轻链恒定基因或其功能性片段选自人C λ 基因和小鼠C λ 基因或其功能性片段。在具体的实施方案中, λ 轻链恒定基因为小鼠C λ 基因或其功能性片段。

[0132] 在一个实施方案中,该方法还包括从小鼠克隆编码重链和/或轻链可变区的核酸序列,其中重链和/或轻链可变区来自包含人V λ 和小鼠C λ 的抗体。

[0133] 在一个方面,提供了用于产生抗体的方法,包括:将本文中描述的小鼠暴露于抗原,让小鼠产生包括产生特异性结合抗原的抗体的免疫应答,鉴定小鼠中的编码重链的重排的核酸序列和小鼠中的编码抗体的关联轻链可变结构域序列的重排的核酸序列,其中所述抗体特异性结合抗原,和利用融合至人恒定结构域的重链和轻链可变结构域的核酸序列来产生期望的抗体,其中期望的抗体包括包含融合至C μ 结构域的V λ 结构域的轻链。在一个实施方案中,V λ 结构域为人的并且C μ 结构域为人或小鼠C λ 结构域。在一个实施方案中,V λ 结构域为小鼠的并且C μ 结构域为人或小鼠C κ 结构域。

[0134] 在一个实施方案中,提供了用于产生抗体的方法,包括:将本文中描述的小鼠暴露于抗原,让小鼠产生包括产生特异性结合抗原的抗体的免疫应答,鉴定小鼠中的编码重链的重排的核酸序列和小鼠中的编码抗体的关联轻链可变结构域序列的重排的核酸序列,其中所述抗体特异性结合抗原,和利用融合至人恒定结构域的核酸序列的重链和轻链可变结构域的核酸序列来产生期望的抗体,其中期望的抗体包括包含融合至C κ 结构域的V λ 结构域的轻链。

[0135] 在一个实施方案中,提供了用于产生抗体的方法,包括:将本文中描述的小鼠暴露于抗原,让小鼠产生包括产生特异性结合抗原的抗体的免疫应答,鉴定小鼠中的编码重链可变结构域的重排的核酸序列和编码抗体的关联轻链可变结构域序列的重排的核酸序列,其中所述抗体特异性结合抗原,和利用融合至编码人重链恒定结构域和人轻链恒定结构域的核酸序列的所述核酸序列来产生来源于人序列的抗体,其中特异性结合抗原的抗体包括包含融合至小鼠C λ 结构域的人V λ 结构域的轻链。

[0136] 在一个实施方案中,小鼠C λ 区选自C λ 1、C λ 2和C λ 3。在具体的实施方案中,小鼠C λ 区为C λ 2。

[0137] 在一个方面,提供了用于产生重排的抗体轻链可变区基因序列的方法,包括(a)将本文中描述的小鼠暴露于抗原;(b)让小鼠产生免疫应答;(c)鉴定小鼠中的包含编码与小鼠C μ 结构域融合的重排的人V λ 结构域序列的核酸序列的细胞,其中细胞还编码包含人V μ 结构域和小鼠C μ 结构域的关联重链,并且其中细胞表达结合抗原的抗体;(d)从细胞克隆编码人V λ 结构域的核酸序列和编码关联人V μ 结构域的核酸序列;和,(e)使用编码人V λ 结构域的克隆的核酸序列和编码关联人V μ 结构域的克隆的核酸序列来产生完全人的抗体。

[0138] 在一个实施方案中,提供了用于产生重排的抗体轻链可变区基因序列的方法,包括(a)将本文中描述的小鼠暴露于抗原;(b)使小鼠产生免疫应答;(c)鉴定小鼠中的包含核酸序列的细胞,所述核酸序列编码在相同核酸分子上与编码小鼠C κ 结构域的核酸序列连续的重排的人V λ 结构域序列,其中细胞还编码包含人V μ 结构域和小鼠C μ 结构域的关联重链,并

且其中细胞表达结合抗原的抗体；(d) 从细胞克隆编码人V λ 结构域的核酸序列和编码关联人V μ 结构域的核酸序列；和(e) 使用编码人V λ 结构域的克隆的核酸序列和编码关联人V μ 结构域的克隆的核酸序列来产生完全人的抗体。

[0139] 在一个实施方案中，提供了用于产生重排的抗体轻链可变区基因序列的方法，包括(a) 将本文中描述的小鼠暴露于抗原；(b) 让小鼠产生针对抗原的免疫应答；(c) 鉴定小鼠中的包含编码与小鼠C λ 结构域融合的重排的人V λ 结构域序列的DNA的细胞，其中所述细胞还编码包含人V μ 结构域和小鼠C μ 结构域的关联重链，并且其中细胞表达结合抗原的抗体；(d) 从细胞克隆编码重排的人V λ 结构域的核酸序列和编码关联人V μ 结构域的核酸序列；和，(e) 使用编码人V λ 结构域的克隆的核酸序列和编码关联人V μ 结构域的克隆的核酸序列来产生完全人的抗体。在一个实施方案中，小鼠C λ 结构域为小鼠C λ 2。在具体的实施方案中，小鼠C λ 结构域来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 基因。

[0140] 在一个方面，提供了表达融合至内源轻链恒定区(C μ)的人 λ 来源的轻链的遗传修饰的小鼠，其中小鼠在用抗原免疫后，产生包含融合至小鼠C μ 结构域的人V λ 结构域的抗体。在一个实施方案中，小鼠C μ 结构域选自C κ 结构域和C λ 结构域。在一个实施方案中，小鼠C μ 结构域为C κ 结构域。在一个实施方案中，小鼠C μ 结构域为C λ 结构域。在具体的实施方案中，C λ 结构域为C λ 2。在具体的实施方案中，小鼠C λ 结构域来源于与小鼠C λ 2具有至少60%、至少70%、至少80%、至少90%、至少95%或至少98%的同一性的C λ 基因。

[0141] 在一个方面，提供了包含本文中描述的修饰的内源 κ 或 λ 轻链基因座的遗传修饰的小鼠，其表达与多个免疫球蛋白重链相关的多个免疫球蛋白 λ 轻链。在一个实施方案中，重链包括人序列。在多个实施方案中，人序列选自可变序列、C μ 1、铰链、C μ 2、C μ 3及其组合。在一个实施方案中，多个免疫球蛋白 λ 轻链包括人序列。在多个实施方案中，人序列选自可变序列、恒定序列及其组合。在一个实施方案中，小鼠包括失能的内源免疫球蛋白基因座，并且从转基因或染色体外附加体表达重链和/或 λ 轻链。在一个实施方案中，小鼠在内源小鼠基因座上包括，利用一个或多个人免疫球蛋白序列对一些或所有内源小鼠重链基因区段(即，V、D、J)和/或一些或所有内源小鼠重链恒定序列(例如，C μ 1、铰链、C μ 2、C μ 3或其组合)和/或一些或所有内源小鼠轻链序列(例如，V、J、恒定区或其组合)的替换。

[0142] 在一个方面，提供了适用于产生具有人 λ 来源的轻链的抗体的小鼠，其中在小鼠中产生的所有或基本上所有抗体经表达具有人 λ 来源的轻链。在一个实施方案中，从内源轻链基因座表达人 λ 来源的轻链。在一个实施方案中，内源轻链基因座为 κ 轻链基因座。在具体的实施方案中， κ 轻链基因座为小鼠 κ 轻链基因座。

[0143] 在一个方面，提供了用于产生人抗体的 λ 来源的轻链的方法，包括从本文中描述的小鼠获得轻链序列和重链序列，以及使用轻链序列和重链序列来产生人抗体。

[0144] 在一个方面，提供了用于产生抗原结合蛋白的方法，包括将本文中描述的小鼠暴露于抗原；让小鼠产生免疫应答；和从小鼠获得结合抗原的抗原结合蛋白，或从小鼠获得用于产生结合抗原的抗原结合蛋白的序列。

[0145] 在一个方面，提供了来源于本文中描述的小鼠的细胞。在一个实施方案中，细胞选自胚胎干细胞、多能细胞、诱导的多能细胞、B细胞和杂交瘤。

[0146] 在一个方面，提供了包含本文中描述的遗传修饰的细胞。在一个实施方案中，细胞

为小鼠细胞。在一个实施方案中,细胞选自杂交瘤和四源杂交瘤(quadroma)。在一个实施方案中,细胞表达包含与小鼠恒定序列融合的人 λ 可变序列的免疫球蛋白轻链。在具体的实施方案中,小鼠恒定序列为小鼠 κ 恒定序列。

[0147] 在一个方面,提供了来源于本文中描述的小鼠的组织。

[0148] 在一个方面,提供了本文中描述的小鼠或细胞的用于产生抗原结合蛋白的用途。在一个实施方案中,抗原结合蛋白为人蛋白质。在一个实施方案中,人蛋白质为人抗体。

[0149] 在一个方面,提供了由本文中描述的小鼠、细胞、组织或方法产生的抗原结合蛋白。在一个实施方案中,抗原结合蛋白为人蛋白质。在一个实施方案中,人蛋白质为人抗体。

[0150] 除非另外明确地指出或从上下文中很明显的看出,否则可将本文中描述的任何实施方案和方面彼此结合使用。通过参考随后的描述,其它实施方案对于本领域技术人员将变得很显然。

[0151] 附图概述

[0152] 图1显示了包括V λ 基因区段的簇(A、B和C)以及J λ 和C λ 区对(J-C对)的人 λ 轻链基因座的不按比例的比例的详细说明。

[0153] 图2显示了用于使内源小鼠 λ 轻链基因座失活的靶向策略的不按比例的比例的说明。

[0154] 图3显示了用于使内源小鼠 κ 轻链基因座失活的靶向策略的不按比例的比例的说明。

[0155] 图4A显示用于靶向内源小鼠 λ 轻链基因座的具有人 λ 轻链序列(包括12个hV λ 基因区段和hJ λ 1基因区段)的初始靶向载体(12/1- λ 靶向载体)的不按比例的比例的说明。

[0156] 图4B显示4个用于靶向内源小鼠 κ 轻链基因座的具有人 λ 轻链序列的初始靶向载体的不按比例的比例的说明,所述载体分别包括12个hV λ 基因区段以及hJ λ 1基因区段(12/1- κ 靶向载体)、12个hV λ 基因区段以及hJ λ 1、2、3和7基因区段(12/4- κ 靶向载体)、12个hV λ 基因区段、人V κ -J κ 基因组序列以及hJ λ 1基因区段(12(κ)1- κ 靶向载体)和12个hV λ 基因区段、人V κ -J κ 基因组序列以及hJ λ 1、2、3和7个基因区段(12(κ)4- κ 靶向载体)。

[0157] 图5A显示了用于将40个hV λ 基因区段和单个hJ λ 基因区段逐步插入小鼠 λ 轻链基因座内的靶向策略的不按比例的比例的说明。

[0158] 图5B显示了用于将40个hV λ 基因区段和单个hJ λ 基因区段逐步插入小鼠 κ 轻链基因座内的靶向策略的不按比例的比例的说明。

[0159] 图6显示了用于产生独特人 λ - κ 杂交靶向载体的靶向和分子工程步骤的不按比例的比例的说明,所述杂交靶向载体用于构建包含人 κ 基因间序列、多个hJ λ 基因区段或这两者的杂交轻链基因座。

[0160] 图7A显示包含有效地连接至内源C λ 2基因的40个hV λ 基因区段和单个hJ λ 基因区段的修饰的小鼠 λ 轻链基因座的基因座结构的不按比例的比例的说明。

[0161] 图7B显示4个独立的修饰的小鼠 κ 轻链基因座的基因座结构的不按比例的比例的说明,所述基因座包含有效地连接至内源C κ 基因的40个hV λ 基因区段和1或4个hJ λ 基因区段,并且具有或不具有连续的人V κ -J κ 基因组序列。

[0162] 图8A显示来自野生型小鼠(WT)、对于包括人V κ -J κ 基因组序列的12个hV λ 和4个hJ λ 基因区段是纯合的小鼠(12hV λ -V κ J κ -4hJ λ)以及对于40个hV λ 和1个hJ λ 基因区段是纯合的小鼠(40hV λ -1hJ λ)的针对CD19⁺门控的Ig λ ⁺和Ig κ ⁺脾细胞的等值线图。

[0163] 图8B显示从野生型(WT)、对于包括人V κ -J κ 基因组序列的12个hV λ 和4个hJ λ 基因区

段是纯合的小鼠 (12hV λ -V κ J κ -4hJ λ) 以及对于40个hV λ 和1个hJ λ 基因区段是纯合的小鼠 (40hV λ -1hJ λ) 收获的脾中的CD19⁺B细胞的总数。

[0164] 图9A, 在上图框中, 显示了来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对单线态 (singlet) 门控和针对B和T细胞 (分别地CD19⁺和CD3⁺) 染色的脾细胞的等值线图。下图框显示了来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对CD19⁺门控和针对Ig λ ⁺和Ig κ ⁺的表达染色的脾细胞的等值线图。

[0165] 图9B显示从野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 收获的脾中的CD19⁺、CD19⁺Ig κ ⁺和CD19⁺Ig λ ⁺B细胞的总数。

[0166] 图9C显示来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对CD19⁺门控和针对免疫球蛋白D (IgD) 和免疫球蛋白M (IgM) 染色的脾细胞的等值线图。在各个等值线图上标明了成熟 (对于WT, 72; 对于40hV λ -V κ J κ -4hJ λ , 51) 和过渡 (对于WT, 13; 对于40hV λ -V κ J κ -4hJ λ , 22) B细胞。

[0167] 图9D显示从野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 收获的脾中的CD19⁺B细胞、过渡B细胞 (CD19⁺IgM^{hi}IgD^{lo}) 和成熟B细胞 (CD19⁺IgM^{lo}IgD^{hi}) 的总数。

[0168] 图10A, 在上图框中, 显示了来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对B和T细胞 (分别地, CD19⁺和CD3⁺) 染色的骨髓的等值线图。下图框显示了来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对CD19⁺门控和针对ckit⁺和CD43⁺染色的骨髓的等值线图。祖B细胞 (Pro B cell) 和前B细胞标示于下图框的等值线图上。

[0169] 图10B显示从野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的股骨收获的骨髓中的祖B细胞 (CD19⁺CD43⁺ckit⁺) 和前B细胞 (CD19⁺CD43⁻ckit⁻) 的数目。

[0170] 图10C显示来自野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的针对单线态门控的、针对免疫球蛋白M (IgM) 和B220染色的骨髓的等值线图。未成熟的、成熟的和祖/前B细胞标示于各个等值线图上。

[0171] 图10D显示从野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的股骨分离的骨髓中的未成熟 (B220^{int}IgM⁺) 和成熟 (B220^{hi}IgM⁺) B细胞的总数。

[0172] 图10E显示从野生型小鼠 (WT) 和对于包括人V κ -J κ 基因组序列的40个hV λ 和4个J λ 基因区段是纯合的小鼠 (40hV λ -V κ J κ -4hJ λ) 的股骨分离的、针对未成熟 (B220^{int}IgM⁺) 和成熟 (B220^{hi}IgM⁺) B细胞门控的、针对Ig λ 和Ig κ 表达染色的骨髓的等值线图。

[0173] 图11显示从在内源小鼠 κ 轻链基因座上具有人 λ 轻链基因序列的小鼠的脾细胞RNA扩增的18个独立的RT-PCR克隆的V λ -J λ -C κ 连接的核苷酸序列比对。A6=SEQ ID NO:57; B6=SEQ ID NO:58; F6=SEQ ID NO:59; B7=SEQ ID NO:60; E7=SEQ ID NO:61; F7=SEQ ID NO:62; C8=SEQ ID NO:63; E12=SEQ ID NO:64; 1-4=SEQ ID NO:65; 1-20=SEQ ID NO:

66;3B43=SEQ ID NO:67;5-8=SEQ ID NO:68;5-19=SEQ ID NO:69;1010=SEQ ID NO:70;11A1=SEQ ID NO:71;7A8=SEQ ID NO:72;3A3=SEQ ID NO:73;2-7=SEQ ID NO:74。小写碱基表示因重组过程中的突变和/或N添加而产生的非种系碱基。由hJλ1和小鼠Cκ的核苷酸序列编码的构架4区域(FWR4)内的共有氨基酸示于序列比对的底部。

[0174] 图12显示从在内源小鼠κ轻链基因座上具有包括连续的人Vκ-Jκ基因组序列的人λ轻链基因序列的小鼠的脾细胞RNA扩增的12个独立的RT-PCR克隆的Vλ-Jλ-Cκ连接的核苷酸序列比对。5-2=SEQ ID NO:87;2-5=SEQ ID NO:88;1-3=SEQ ID NO:89;4B-1=SEQ ID NO:90;3B-5=SEQ ID NO:91;7A-1=SEQ ID NO:92;5-1=SEQ ID NO:93;4A-1=SEQ ID NO:94;11A-1=SEQ ID NO:95;5-7=SEQ ID NO:96;5-4=SEQ ID NO:97;2-3=SEQ ID NO:98。小写碱基表示因重组过程中的突变和/或N添加而产生的非种系碱基。由各个人Jλ和小鼠Cκ的核苷酸序列编码的构架4区域(FWR4)内的共有氨基酸示于序列比对的底部。

[0175] 图13显示从在内源小鼠λ轻链基因座上具有人λ轻链基因序列的小鼠的脾细胞RNA扩增的3个独立RT-PCR克隆的Vλ-Jλ-Cλ连接的核苷酸序列比对。2D1=SEQ ID NO:101;2D9=SEQ ID NO:102;3E15=SEQ ID NO:103。小写碱基表示因重组过程中的突变和/或N添加而产生的非种系碱基。由hJλ1和小鼠Cλ2的核苷酸序列编码的构架4区域(FWR4)内的共有氨基酸示于序列比对的底部。

[0176] 发明详述

[0177] 虽然详细地论述了不同实施方案的具体特征,但具体方面、实施方案和实施例的描述不限制权利要求的主题;权利要求描述了本发明的范围。本文中使用的术语和短语包括本领域中通常赋予它们的含义。

[0178] 术语“连续的”是指,在相同核酸分子上的存在,例如,如果两个核酸序列存在于相同核酸分子上但被另一个核酸序列中断,则它们是“连续的”。例如,重排的V(D)J序列与恒定区基因序列是“连续的”,虽然V(D)J序列的最后一个密码子之后并未紧接恒定区序列的第一个密码子。在另一个实例中,如果两个V基因区段序列存在于相同的基因组片段上,则它们是“连续的”,尽管它们可被不编码V区域的密码子的序列隔开,例如它们可被调控序列例如启动子或其它非编码序列隔开。在一个实施方案中,连续序列包括包含如在野生型基因组中发现的排列的基因组序列的基因组片段。

[0179] 短语“来源于”,当在“来源于”引述的基因或基因区段的可变区方面使用时,包括使序列追溯至特定未重排的基因区段的能力,所述未重排的基因区段经重排以形成表达可变结构域的基因(在适用的情况下,解释剪接差异和体细胞突变)。

[0180] 短语“功能性”,当在可变区基因区段或连接基因区段方面使用时,是指在表达的抗体库中的使用;例如,在人Vλ基因区段中3-1、4-3、2-8等是功能性的,而Vλ基因区段3-2、3-4、2-5等是非功能性的。

[0181] “重链基因座”包括染色体例如小鼠染色体上的位置,其中在野生型小鼠中发现重链可变区(V_H)、重链多样性区(D_H)、重链连接区(J_H)和重链恒定区(C_H) DNA序列。

[0182] “κ基因座”包括染色体例如小鼠染色体上的位置,其中在野生型小鼠中发现κ可变区(V_κ)、κ连接区(J_κ)和κ恒定区(C_κ) DNA序列。

[0183] “λ基因座”包括染色体例如小鼠染色体上的位置,其中在野生型小鼠中发现λ可变区(V_λ)、λ连接区(J_λ)和λ恒定区(C_λ) DNA序列。

[0184] 术语“未重排的”包括免疫球蛋白基因座的这样的状态,其中V基因区段和J基因区段(对于重链,还有D基因区段)分开地存在,但能够被连接以形成包含V(D)J库的单个V、(D)、J的重排的V(D)J基因。

[0185] 表达人 λ 可变结构域的小鼠

[0186] 先前已报导了表达完全人或部分人且部分小鼠的抗体的小鼠。**VELOCIMMUNE®**基因工程小鼠包括,利用人V(D)J基因区段对内源小鼠基因座上的未重排的V(D)J基因区段的替换。**VELOCIMMUNE®**小鼠表达具有人可变结构域和小鼠恒定结构域的嵌合抗体(参见,例如,美国专利7,605,237)。大部分其它报导涉及,在具有失能的内源免疫球蛋白基因座的小鼠中从完全人转基因表达完全人抗体的小鼠。

[0187] 抗体轻链由两个分开的基因座:kappa (κ) 和lambda (λ) 之一编码。小鼠抗体轻链主要为 κ 型。产生小鼠抗体的小鼠和产生完全人或嵌合人-小鼠抗体的修饰的小鼠在轻链使用上显示偏爱性。人也显示轻链偏爱性,但不像小鼠这样明显;小鼠的 κ 轻链对 λ 轻链的比率为约95:5,然而在人中,该比率为约60:40。小鼠中的更明显的偏爱性据认为不会严重地影响抗体多样性,因为在小鼠中, λ 可变基因座并不是如此多样的。在人中情况则不同。人 λ 轻链基因座是丰富多样的。

[0188] 人 λ 轻链基因座延伸超过1,000 kb并且包含超过80个编码可变(V)或连接(J)区段的基因(图1)。在人 λ 轻链基因座内,超过一半的所有观察到的V λ 结构域由基因区段1-40、1-44、2-8、2-14和3-21编码。总体上,约30个左右的人V λ 基因区段据认为是功能性的。存在7个J λ 基因区段,其中仅4个被认为是通常具有功能的J λ 基因区段--J λ 1、J λ 2、J λ 3和J λ 7。

[0189] 人中的 λ 轻链基因座在结构上与小鼠和人中的 κ 基因座相似,因为人 λ 轻链基因座具有几个能够重组以形成功能性轻链蛋白质的可变区基因区段。人 λ 轻链基因座包含约70个V基因区段和7个J λ -C λ 基因区段对。此类J λ -C λ 基因区段对中仅有4个显示是功能性的。在一些等位基因中,第五J λ -C λ 基因区段对据报导为假基因(C λ 6)。70 V λ 基因区段显示包含38个功能性基因区段。70个V λ 序列排列在3个簇中,其全都包含不同V基因家族组的不同成员(A、B和C簇;图1)。这是用于在非人动物中产生具有人V区域的抗体的相对未利用的多样性的潜在丰富资源。

[0190] 形成鲜明对比地,小鼠 λ 轻链基因座仅包含2个或3个(取决于品系)小鼠V λ 区基因区段(图2)。至少因为该原因,小鼠中的严重的 κ 偏爱性据认为对总的抗体多样性不是特别有害。

[0191] 根据小鼠 λ 轻链基因座的已公布的图谱,该基因座基本上由在约200 kb的跨度内的两个基因区段簇组成(图2)。这两个簇包含两组能够独立地重排的V、J和C基因:V λ 2-J λ 2-C λ 2-J λ 4-C λ 4和V λ 1-J λ 3-C λ 3-J λ 1-C λ 1。虽然已发现V λ 2与所有J λ 基因区段重组,但V λ 1显示只与C λ 1重组。C λ 4据认为是具有缺陷性剪接位点的假基因。

[0192] 小鼠 κ 轻链基因座明显不同。来自小鼠 κ 基因座的参与导致功能性轻链蛋白的重组事件的基因区段的结构和数目要复杂得多(图3)。因此,小鼠 λ 轻链未对一般小鼠中的抗体群体的多样性作出巨大贡献。

[0193] 在小鼠中利用人 λ 轻链基因座的丰富多样性,特别地,将很可能产生用于轻链V结构域的更完全的人库的来源。先前利用该多样性的尝试使用包含大量随机整合入小鼠基因组的人 λ 轻链基因座的人转基因(参见,例如,US 6,998,514和US 7,435,871)。包含此类随

机整合的转基因的小鼠据报导表达完全人 λ 轻链,然而,在一些情况下,一个或两个内源轻链基因座保持完整。该情形是不想要的,因为人 λ 轻链序列与小鼠轻链(κ 或 λ)在小鼠的表达的抗体库中相对抗。

[0194] 相比之下,本发明人描述了遗传修饰的小鼠,其能够从小鼠轻链基因座直接表达一个或多个 λ 轻链核酸序列,包括通过内源小鼠轻链基因座上的替换。能够从内源基因座表达人 λ 轻链序列的遗传修饰的小鼠可被进一步培育成包含人重链基因座,从而可用于表达包含完全人的V区(重链和轻链)的抗体的小鼠。在多种实施方案中,V区与小鼠恒定区一起表达。在多种实施方案中,内源小鼠免疫球蛋白基因区段不存在并且V区与人恒定区一起表达。此类抗体被证明在许多应用(诊断以及治疗)中是有用的。

[0195] 对于在小鼠中表达来源于人V λ 和J λ 基因区段的结合蛋白的各种实施方案,可实现许多有利方面。可通过将人 λ 序列置于内源轻链基因座例如小鼠 κ 或 λ 基因座上来实现有利方面。从此类小鼠产生的抗体可具有包含融合至小鼠C λ 区(特别地小鼠C κ 或C λ 区)的人V λ 结构域的轻链。小鼠还将表达人V λ 结构域,所述人V λ 结构域适于与人C λ 区(特别地C κ 和/或C λ 区)一起用于鉴定和克隆。因为此类小鼠中的B细胞发育在其它方面是正常的,所以能够在C λ 或C κ 区的背景中产生相容的V λ 结构域(包括体细胞突变的V λ 结构域)。

[0196] 描述了在免疫球蛋白 κ 或 λ 轻链基因座上包含未重排的V λ 基因区段的遗传修饰的小鼠。描述了表达包含具有融合至C κ 和/或C λ 区的人V λ 结构域的轻链的抗体的小鼠。

[0197] 免疫球蛋白 κ 轻链基因座的不表达转录物(Sterile Transcript)

[0198] 在小鼠中表达人免疫球蛋白 λ 序列的主题的变化反映在能够进行此类表达的遗传修饰的小鼠的不同实施方案上。因此,在一些实施方案中,遗传修饰的小鼠包含来自人基因座的某些非编码序列。在一个实施方案中,遗传修饰的小鼠在内源 κ 轻链基因座上包含人V λ 和J λ 基因区段,并且还包含人 κ 轻链基因组片段。在具体的实施方案中,人 κ 轻链基因组片段为在人V κ 基因区段与人J κ 基因区段之间天然发现的非编码序列。

[0199] 人和小鼠 κ 轻链基因座包含编码不表达转录物的序列,所述转录物不存在起始密码子或开放阅读框架,并且被当作是调控 κ 轻链基因座的转录的元件。此类不表达转录物产生自位于最接近的V κ 基因区段的下游或3'端与处于 κ 轻链恒定区基因(C κ)的上游的 κ 轻链内含增强子(E κ i)的上游或5'端之间的基因间序列。不表达转录物产生自基因间序列的重排(以形成融合至C κ 的V κ J κ 1区段)。

[0200] 处于C κ 基因的上游的 κ 轻链基因座的替换可除去编码不表达转录物的基因间区域。因此,在多种实施方案中,利用人 λ 轻链基因区段对小鼠C κ 基因上游的小鼠 κ 轻链序列的替换将导致人源化小鼠 κ 轻链基因座,所述基因座包含人V λ 和J λ 基因区段但不包含编码不表达转录物的 κ 轻链基因间区域。

[0201] 如本文中所描述的,利用人 λ 轻链基因区段对内源小鼠 κ 轻链基因座的人源化(其中人源化除去基因间区域)导致,与 λ 轻链使用的显著增加偶联的 κ 轻链基因座的使用的显著下降。因此,虽然不存在基因间区域的人源化小鼠是有用的(因为其可产生具有人轻链可变结构域(例如,人 λ 或 κ 结构域)的抗体),但是来自基因座的使用减少。

[0202] 还描述了,利用人V λ 和J λ 基因区段对内源小鼠 κ 轻链基因座的人源化,该人源化偶联人 κ 基因间区域的插入以产生V λ 基因座,所述基因座在转录上在最后一个人V λ 基因区段与第一个人J λ 基因区段之间包含 κ 基因间区域;所述基因座与不存在 κ 基因间区域的基因座

相比,在B细胞群体中显示了更高的表达。该观察结果与这样的假说相符:基因间区域--通过不表达转录物直接地,或间接地--抑制内源 λ 轻链基因座的使用。在这样的假说下,包括所述基因间区域将导致内源 λ 轻链基因座的使用减少,从而使得小鼠限制性选择,但利用修饰的(λ 至 κ)基因座来产生抗体。

[0203] 在不同的实施方案中,利用人 λ 轻链序列对小鼠C κ 基因上游的小鼠 κ 轻链序列的替换还在转录方向上包括,置于最3'端的V λ 基因区段的3'非翻译区与第一个人J λ 基因区段的5'端之间的人 κ 轻链基因间区域。或者,这样的基因间区域可通过在内源 λ 轻链基因座中产生缺失而从替换的内源 κ 轻链基因座(在小鼠C κ 基因的上游)省略。同样地,在该实施方案下,小鼠从包含人 λ 轻链序列的内源 κ 轻链基因座产生抗体。

[0204] 工程化小鼠以表达人V λ 结构域的方法

[0205] 描述了产生遗传修饰小鼠的各种方法,所述遗传修饰小鼠产生包含具有融合至内源C λ (例如C κ 或C λ)区的人V λ 结构域的轻链的抗体。描述了遗传修饰,在不同的实施方案中,所述遗传修饰包括一个或两个内源轻链基因座的缺失。例如,为了从内源抗体库消除小鼠 λ 轻链,可进行第一V λ -J λ -C λ 基因簇的缺失和利用人V λ -J λ 基因区段完全或部分地替换第二基因簇的V λ -J λ 基因区段。还提供了遗传修饰小鼠的胚胎、细胞,以及用于产生所述小鼠、小鼠胚胎和细胞的靶向构建体。

[0206] 一个内源V λ -J λ -C λ 基因簇的缺失和另一个内源V λ -J λ -C λ 基因簇的V λ -J λ 基因区段的替换,在不同的实施方案中,对动物的天然抗体恒定区的缔合和功能具有相对最小的破坏,因为内源C λ 基因保持完整并从而保持正常的功能性和与内源重链恒定区缔合的能力。因此,在此类实施方案中,取决于用于装配包含2条重链和2条轻链的功能性抗体分子的功能性轻链恒定区,修饰不影响其它内源重链恒定区。此外,在不同实施方案中,修饰不影响包括内源重链和轻链(例如连接至小鼠C λ 区的hV λ 结构域)的功能性膜结合的抗体分子的装配。因为至少一个功能性C λ 基因保留在内源基因座上,所以包括利用人V λ -J λ 基因区段对内源V λ -J λ -C λ 基因簇的V λ -J λ 基因区段的替换的动物应当能够产生正常 λ 轻链,所述轻链能够在免疫应答中通过存在于动物的表达的抗体库中的人V λ -J λ 基因区段结合抗原。

[0207] 在图2中提供了缺失的内源小鼠V λ -J λ -C λ 基因簇的示意性说明(未按比例)。如所说明的,小鼠 λ 轻链基因座被组织成2个基因簇,所述两个基因簇都包含能够重组以形成功能性小鼠 λ 轻链的功能性基因区段。内源小鼠V λ 1-J λ 3-C λ 3-J λ 1-C λ 1基因簇被具有侧翼连接重组位点的新霉素盒的靶向构建体(靶向载体1)缺失。另一个内源基因簇(V λ 2-V λ 3-J λ 2-C λ 2-J λ 4-C λ 4)被具有侧翼连接有重组位点的潮霉素-胸苷激酶盒的靶向构建体(靶向载体2)部分缺失。在该第二靶向事件中,C λ 2-J λ 4-C λ 4内源基因区段得到保留。使用与第一靶向构建体(靶向载体1)不同的重组位点构建第二靶向构建体(靶向载体2),从而允许在已实现成功靶向后选择性缺失选择盒。所得的双靶向基因座在功能上被沉默,因为不能产生内源 λ 轻链。该修饰的基因座可用于插入人V λ 和J λ 基因区段以产生包含人V λ 和J λ 基因区段的内源小鼠 λ 基因座,由此在修饰的基因座上重组后,动物产生包含连接至内源小鼠C λ 基因区段的重排的人V λ 和J λ 基因区段的 λ 轻链。

[0208] 在不同的实施方案中,遗传修饰小鼠以使得内源 λ 基因区段不具有功能,这导致在其抗体库中只显示 κ 轻链的小鼠,从而产生可用于评估 λ 轻链在免疫应答中的作用以及可用于产生包含V κ 结构域但不包含V λ 结构域的抗体库的小鼠。

[0209] 表达已在内源小鼠 λ 轻链基因座上重组的连接至小鼠C λ 基因的hV λ 的遗传修饰的小鼠可由本领域已知的任何方法产生。图4A中提供了,利用人V λ 和J λ 基因区段对内源小鼠V λ 2-V λ 3-J λ 2基因区段的替换的示意性说明(未按比例)。如所说明的,内源小鼠 λ 轻链基因座(已使其无功能)被包括侧翼连接有重组位点的新霉素盒的靶向构建体(12/1- λ 靶向载体)替换。V λ 2-V λ 3-J λ 2基因区段被包含含有12个hV λ 基因区段和单个hJ λ 基因区段的人 λ 序列的基因组片段置换。

[0210] 因此,该第一方法在内源 λ 轻链基因座上放置了与单个hJ λ 基因区段连续的一个或多个hV λ 基因区段(图4A)。

[0211] 可通过使用类似的技术来实现对修饰的内源 λ 轻链基因座的进一步修饰(以插入更多的hV λ 基因区段)。例如,在图5A中提供了,用于逐步插入额外的人hV λ 基因区段的两个另外的靶向构建体(+16- λ 和+12- λ 靶向载体)的示意性说明。如所说明的,使用由先前插入的人 λ 轻链序列提供的同源性,在连续的步骤中将包含特定人hV λ 基因区段的另外的基因组片段插入修饰的内源 λ 轻链基因座。通过依次与每一个说明的靶向构建体重组,将28个另外的hV λ 基因区段插入修饰的内源 λ 轻链基因座。这产生了嵌合基因座,所述基因座产生包含连接至小鼠C λ 基因的人V λ -J λ 基因区段的 λ 轻链蛋白质。

[0212] 在小鼠 λ 基因座上插入人 λ 轻链基因区段的上述方法保持位于C λ 2-J λ 4-C λ 4基因区段下游的增强子(命名为Enh 2.4、Enh和Enh3.1,图4A和图5A)。该方法在内源小鼠 λ 轻链基因座上导致单个修饰的等位基因(图7A)。

[0213] 提供了用于产生表达包含有效地连接至小鼠C λ 基因区段的hV λ 和J λ 基因区段的轻链的小鼠的组合物和方法,包括用于产生从内源小鼠 λ 轻链基因座表达此类基因的小鼠的组合物和方法。所述方法包括,选择性地使一个内源小鼠V λ -J λ -C λ 基因簇丧失功能(例如,通过靶向缺失),并在内源小鼠 λ 轻链基因座上使用hV λ 和J λ 基因区段,以在小鼠中表达hV λ 结构域。

[0214] 或者,在第二方法中,可将人 λ 轻链基因区段置于内源 κ 轻链基因座上。在不同的实施方案中,遗传修饰包括内源 κ 轻链基因座的缺失。例如,为了从内源抗体库消除小鼠 κ 轻链,可进行小鼠V κ 和J κ 基因区段的缺失。还提供了遗传修饰的小鼠胚胎、细胞,以及用于产生所述小鼠、小鼠胚胎和细胞的靶向构建体。

[0215] 由于上述原因,小鼠V κ 和J κ 基因区段的缺失使用相对最小的破坏。在图3中提供了,缺失的小鼠V κ 和J κ 基因区段的示意性说明(未按比例)。通过位于两个精确定位的靶向载体(各自使用位点特异性重组位点)之间的小鼠序列的重组酶介导的缺失,来缺失内源小鼠V κ 和J κ 基因区段。将第一靶向载体(J κ 靶向载体)用于第一靶向事件以缺失小鼠J κ 基因区段。将第二靶向载体(V κ 靶向载体)用于第二随后的靶向事件,以缺失位于最远端的小鼠V κ 基因区段的5'端的序列。两个靶向载体都包含位点特异性重组位点,从而允许在实现成功靶向后选择性缺失两个选择盒和所有其间的小鼠 κ 轻链序列。所得的经缺失的基因座在功能上被沉默,因为不能产生内源 κ 轻链。该修饰的基因座可用于插入hV λ 和J λ 基因区段以产生包含hV λ 和J λ 基因区段的内源小鼠 κ 基因座,从而在修饰的基因座上重组后,动物产生包含有效地连接至内源小鼠C κ 基因区段的重排的hV λ 和J λ 基因区段的 λ 轻链。可将不同的包含人 λ 轻链序列的靶向载体与该缺失的小鼠 κ 基因座结合使用,以产生包含与小鼠C κ 区有效连接的人 λ 基因区段的杂交轻链基因座。

[0216] 因此,第二方法将一个或多个 $\nu\lambda$ 基因区段置于小鼠 κ 轻链基因座上,与单个人 $J\lambda$ 基因区段连续(12/1- κ 靶向载体,图4B)。

[0217] 在不同的实施方案中,可改进该方法以添加基因区段和/或调控序列,从而优化小鼠抗体库内来自小鼠 κ 基因座的人 λ 轻链序列的使用。

[0218] 在第三方法中,将一个或多个 $h\nu\lambda$ 基因区段置于小鼠 κ 轻链基因座上,与4个 $hJ\lambda$ 基因序列连续(12/4- κ 靶向载体,图4B)。

[0219] 在第三方法中,将一个或多个 $h\nu\lambda$ 基因区段置于小鼠 κ 轻链基因座上,与人 κ 基因间序列和单个 $hJ\lambda$ 基因序列连续(12(κ)1- κ 靶向载体,图4B)。

[0220] 在第四方法中,将一个或多个 $h\nu\lambda$ 基因区段置于小鼠 κ 轻链基因座上,与人 κ 基因间序列和4个 $hJ\lambda$ 基因序列连续(12(κ)4- κ 靶向载体,图4B)。

[0221] 在小鼠 κ 基因座上插入人 λ 轻链基因区段的所有上述方法维持在 $C\kappa$ 基因上游的 κ 内含增强子元件(命名为 $E\kappa i$,图4B和图5B)和在 $C\kappa$ 基因下游的 $3'\kappa$ 增强子(命名为 $E\kappa 3'$,图4B和图5B)。所述方法在内源小鼠 κ 轻链基因座上导致4个分开的修饰的等位基因(图7B)。

[0222] 在不同的实施方案中,遗传修饰的小鼠包括内源小鼠 λ 轻链基因座的敲除。在一个实施方案中,通过缺失横跨 $\nu\lambda 2$ 至 $J\lambda 2$ 的区域和横跨 $\nu\lambda 1$ 至 $C\lambda 1$ 的区域的策略来敲除 λ 轻链基因座(图2)。减小或消除内源 λ 轻链基因座表达内源 λ 结构域的能力的任何策略适合于与本公开内容中的实施方案一起使用。

[0223] 来自遗传修饰的小鼠的 λ 结构域抗体

[0224] 在小鼠 κ 或 λ 轻链基因座上包含人 λ 序列的小鼠将表达包含融合至小鼠 C_L ($C\kappa$ 或 $C\lambda$)区的 $h\nu\lambda$ 区的轻链。将此类小鼠有利地培育成这样的小鼠,所述小鼠(a)包含在功能上被沉默的轻链基因座(例如,内源小鼠 κ 或 λ 轻链基因座的敲除);(b)包含含有有效地连接至内源小鼠 $C\lambda$ 基因的 hV 和 hJ 基因区段的内源小鼠 λ 轻链基因座;(c)包含含有有效地连接至内源小鼠 $C\kappa$ 基因的 $hV\kappa$ 和 $hJ\kappa$ 基因区段的内源小鼠 κ 轻链基因座;和,(d)这样的小鼠,其中一个 κ 等位基因包含 $hV\kappa$ 和 $hJ\kappa$;另一个 κ 等位基因包含 $h\nu\lambda$ 和 $hJ\lambda$;一个 λ 等位基因包含 $h\nu\lambda$ 和 $hJ\lambda$ 并且一个 λ 等位基因被沉默或敲除,或两个 λ 等位基因都包含 $h\nu\lambda$ 和 $hJ\lambda$;和,两个重链等位基因各自包含 hV_H 、 hD_H 和 hJ_H 。

[0225] 通过将编码 $h\nu\lambda$ 结构域的核酸克隆入具有编码人 $C\lambda$ 的基因的表达构建体,来将包含在 $C\kappa$ 或 $C\lambda$ 的背景中表达的 $h\nu\lambda$ 结构域的抗体用于产生完全人抗体。将所得的表达构建体转染入适当的宿主细胞,以表达显示融合至 $hC\lambda$ 的完全 $h\nu\lambda$ 结构域的抗体。

实施例

[0226] 下列实施例被提供用以描述如何产生和使用本发明的方法和组合物,并且无意限制本发明人所认定的本发明的范围。除非另外指出,否则温度以摄氏度表示,并且压力为大气压或接近大气压。

[0227] 实施例I

[0228] 小鼠免疫球蛋白轻链基因座的缺失

[0229] 使用 **VELOCIGENE®** 技术(参见,例如,美国专利6,586,251和Valenzuela等人(2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21 (6):652-659) 产生各种靶向构建

体来修饰小鼠基因组细菌人工染色体(BAC)文库,以使小鼠 κ 和 λ 轻链基因座失活。

[0230] 小鼠 λ 轻链基因座的缺失。通过同源重组修饰来自小鼠BAC克隆RP23-135k15 (Invitrogen)的DNA,以通过V λ -J λ -C λ 基因簇的靶向缺失来使内源小鼠 λ 轻链基因座失活(图2)。

[0231] 简而言之,使用靶向载体在单个靶向事件中缺失包含V λ 1-J λ 3-C λ 3-J λ 1-C λ 1基因区段的整个近端簇,所述靶向载体包含侧翼连接有loxP位点的新霉素盒,具有包含V λ 1基因区段5'端的序列的5'小鼠同源臂和包含C λ 1基因区段3'端的序列的3'小鼠同源臂(图2,靶向载体1)。

[0232] 制备第二靶向构建体以精确地缺失包含V λ 2-J λ 2-C λ 2-J λ 4-C λ 4的远端内源小鼠 λ 基因簇,除了靶向构建体包含含有V λ 2基因区段5'端的序列的5'小鼠同源臂和含有内源C λ 2基因区段5'端的序列的3'小鼠同源臂外(图2,靶向载体2)。因此,第二靶向构建体精确地缺失V λ 2-J λ 2,同时在内源小鼠 λ 基因座上留下完整的C λ 2-J λ 4-C λ 4。通过本领域已知的核型分析和筛选方法(例如, **Taqman®**) 确认包含灭活的内源 λ 基因座(如上文描述的)的ES细胞。随后从修饰的ES细胞分离DNA,将其经历利用CRE重组酶的处理,从而介导包含新霉素标记基因的近端靶向盒的缺失,仅在缺失点留下单个loxP位点(图2,下图)。

[0233] 小鼠 κ 轻链基因座的缺失。使用类似的上述方法产生几个靶向构建体,以通过同源重组修饰来自小鼠BAC克隆RP23-302g12和RP23-254m04 (Invitrogen)的DNA,从而在两步法中使小鼠 κ 轻链基因座失活(图3)。

[0234] 简而言之,使用靶向载体在单个靶向事件中缺失内源小鼠 κ 轻链基因座的J κ 基因区段(1-5),所述靶向载体包含潮霉素-胸苷激酶(hyg-TK)盒,且在hyg-TK盒的3'端含有单个loxP位点(图3, J κ 靶向载体)。用于产生该靶向载体的同源臂包含在内源小鼠J κ 基因区段的5'和3'端的小鼠基因组序列。在第二靶向事件中,制备第二靶向载体以缺失最远端的内源小鼠V κ 基因区段的上游(5')的小鼠基因组序列的部分(图3, V κ 靶向载体)。该靶向载体包含反向lox511位点、loxP位点和新霉素盒。用于产生该靶向载体的同源臂包含在最远端的小鼠V κ 基因区段上游的小鼠基因组序列。依次(即, J κ 然后V κ)将靶向载体用于靶向ES细胞中的DNA。通过本领域已知的核型分析和筛选方法(例如, Taqman™) 确认具有双靶向的染色体(即,被两个靶向载体靶向的单个内源小鼠 κ 基因座)的ES。随后从修饰的ES细胞分离DNA,将其经历利用Cre重组酶的处理,从而介导内源小鼠V κ 基因区段和两个选择盒的缺失,同时以彼此相反的方向留下两个并置的lox位点(图3,下图;SEQ ID NO:1)。

[0235] 因此,产生了两个包含完整增强子和恒定区的修饰的内源轻链基因座(κ 和 λ),用于以精确地方式使用下述靶向载体逐步插入未重排的人 λ 种系基因区段。

[0236] 实施例II

[0237] 用人 λ 轻链微小-基因座(Mini-Locus)对小鼠轻链基因座的替换

[0238] 工程化多个靶向载体以将人 λ 基因区段逐步插入内源小鼠 κ 和 λ 轻链基因座(使用类似的上述方法)。对内源轻链基因座进行多个独立的初始修饰,每一个修饰产生包含有效地连接至小鼠轻链恒定基因和增强子的hV λ 和J λ 基因区段的嵌合轻链基因座。

[0239] 包含12个人V λ 和1个人J λ 基因区段的人 λ 微小-基因座。使用称为RP11-729g4的BAC克隆(Invitrogen),对一系列初始靶向载体进行工程化以包含来自A簇的前12个连续的人V λ 基因区段和hJ λ 1基因区段或4个hJ λ 基因区段。图4A和4B分别显示了被构建用于产生人 λ 轻

链基因区段在小鼠 λ 和 κ 轻链基因座上的初始插入的靶向载体。

[0240] 对于第一组初始靶向载体,工程化来自729g4 BAC克隆的包含12个hV λ 基因区段和hJ λ 1基因区段的124,125 bp DNA片段,以在hJ λ 1基因区段的996 bp下游(3')包含PI-SceI位点,用于连接3'小鼠同源臂。不同的两组同源臂用于连接至该人片段;一组同源臂包含来自135k15 BAC克隆的内源小鼠 λ 序列(图4A),另一组包含分别来自小鼠BAC克隆RP23-302g12和RP23-254m04的小鼠V κ 和J κ 基因区段的5'和3'端的内源 κ 序列(图4B)。

[0241] 对于12/1- λ 靶向载体(图4A),在实施例1中描述的修饰的小鼠 λ 基因座的包含小鼠C λ 2-J λ 4-C λ 4和增强子2.4的27,847 bp DNA片段的5'末端上进行PI-SceI位点的工程化。通过将约28 kb小鼠片段连接至约124 kb人 λ 片段而将其用作3'同源臂,这产生了从5'至3'包含hJ λ 1基因区段、在hJ λ 1基因区段3'端的996 bp的人 λ 序列、在小鼠C λ 2基因5'端的1229 bp的小鼠 λ 序列、小鼠C λ 2基因和约28 kb小鼠片段的剩余部分的3'连接。人V λ 3-12基因区段的上游(5')是在5'小鼠同源臂的起点之前的另外1456 bp的人 λ 序列,所述5'小鼠同源臂包含相应于内源小鼠 λ 基因座5'端的序列的23,792bp的小鼠基因组DNA。在5'同源臂与人 λ 序列的起点之间的是侧翼连接有Frt位点的新霉素盒。

[0242] 因此,12/1- λ 靶向载体从5'至3'包括,5'同源臂(包含内源 λ 基因座5'端的约24 kb的小鼠 λ 基因组序列)、5'Frt位点、新霉素盒、3'Frt位点、约123 kb的人基因组 λ 序列(包含前12个连续的hV λ 基因区段和hJ λ 1基因区段)、PI-SceI位点和3'同源臂(包含约28 kb的小鼠基因组序列,包括内源C λ 2-J λ 4-C λ 4基因区段、小鼠增强子2.4序列和增强子2.4下游(3')的另外的小鼠基因组序列)(图4A)。

[0243] 以类似的方式,12/1- κ 靶向载体(图4B)使用相同的约124人 λ 片段,除了使用包含小鼠 κ 序列的小鼠同源臂,以便可通过同源重组实现对内源 κ 基因座的靶向。因此,12/1- κ 靶向载体从5'至3'包括,5'同源臂(包含内源 κ 基因座5'端的约23 kb的小鼠基因组序列)、I-CeuI位点、5'Frt位点、新霉素盒、3'Frt位点、约124 kb的人基因组 λ 序列(包含前12个连续的hV λ 基因区段和hJ λ 1基因区段)、PI-SceI位点和3'同源臂(包含约28 kb的小鼠基因组序列,包括内源小鼠C κ 基因、E κ i和E κ 3'以及E κ 3'下游(3')的另外的小鼠基因组序列)(图4B,12/1- κ 靶向载体)。

[0244] 与这两个初始靶向载体之一的同源重组产生了修饰的小鼠轻链基因座(κ 或 λ),其包含有效地连接至内源小鼠轻链恒定基因和增强子(C κ 或C λ 2以及E κ i/E κ 3'或Enh 2.4/Enh 3.1)基因的12个hV λ 基因区段和hJ λ 1基因区段,所述基因座在重组后导致嵌合 λ 轻链的形成。

[0245] 具有12个人V λ 和4个人J λ 基因区段的人 λ 微小基因座。在将多样性添加至嵌合 λ 轻链基因座的另一个方法中,将第三初始靶向载体工程化,以将来自A簇的前12个连续的人V λ 基因区段和hJ λ 1、2、3和7基因区段插入小鼠 κ 轻链基因座(图4B,12/4- κ 靶向载体)。通过从头DNA合成(Integrated DNA Technologies)产生包含hJ λ 1、J λ 2、J λ 3和J λ 7基因区段的DNA区段,其包含每一个J λ 基因区段和来自每一个J λ 基因区段的紧邻5'和3'区域的约100 bp的人基因组序列。将PI-SceI位点工程化至该约1 kb DNA片段的3'末端,并且连接至氯霉素盒。从相对于人BAC克隆729g4的hJ λ 1基因区段在5'和3'位置上的人 λ 序列PCR扩增同源臂。在修饰的729g4 BAC克隆上进行利用该中间靶向载体的同源重组,所述修饰的729g4 BAC克隆先前已利用侧翼连接有Frt位点的新霉素盒靶向人V λ 3-12基因区段的上游(5'),其还在

5'Frt位点的5'端包含I-CeuI位点。双靶向的729g4 BAC克隆从5'至3'包含I-CeuI位点、5'Frt位点、新霉素盒、3'Frt位点、约123 kb的片段(包含前12个hVλ基因区段)、约1 kb的片段(包含人Jλ1、2、3和7基因区段)、PI-SceI位点和氯霉素盒。用I-CeuI和PI-SceI一起降解该中间靶向载体,随后将其连接入修饰的小鼠BAC克隆(上文中描述的)以产生第三靶向载体。

[0246] 该连接导致用于将人λ序列插入内源κ轻链基因座的第三靶向载体,所述第三靶向载体从5'至3'包括,5'小鼠同源臂(包含内源小鼠κ基因座5'端的约23 kb的基因组序列)、I-CeuI位点、5'Frt位点、新霉素盒、3'Frt位点、约123 kb片段(包含前12个hVλ基因区段)、约1 kb的片段(包含hJλ1、2、3和7基因区段)、PI-SceI位点和3'同源臂(包含约28 kb的小鼠基因组序列,包括内源小鼠Cκ基因、Eκi和Eκ3'以及Eκ3'下游(3')的另外的小鼠基因组序列(图4B, 12/4-κ靶向载体)。利用该第三靶向载体的同源重组产生包含有效地连接至内源小鼠Cκ基因的12个hVλ基因区段和4个hJλ基因区段的修饰的小鼠κ轻链基因座,所述基因座在重组后导致嵌合人λ/小鼠κ轻链的形成。

[0247] 具有整合的人κ轻链序列的人λ微小基因座。以类似的方式将两个另外的靶向载体(与经工程化用于产生人λ基因区段至内源κ轻链基因座中的初始插入的那些靶向载体(图4B, 12/1-κ和12/4-κ靶向载体)相似)工程化,以通过使用独特地构建的靶向载体(包含连续的人λ和κ基因组序列)逐步地插入人λ轻链基因区段。此类靶向载体被构建来包括天然地位于人Vκ4-1与Jκ1基因区段之间的约23 kb人κ基因组序列。将该人κ基因组序列特别地置于这两个另外的靶向载体中的人Vλ与人Jλ基因区段之间(图4B, 12(κ) 1-κ和12(κ) 4-κ靶向载体)。

[0248] 使用上述修饰的RP11-729g4 BAC克隆产生包含人κ基因组序列的两个靶向载体(图6)。利用侧翼连接有NotI和AciSI限制位点的壮观霉素选择盒靶向该修饰的BAC克隆(图6, 左上)。利用壮观霉素盒的同源重组导致双靶向的729g4 BAC克隆,所述克隆从5'至3'包括,I-CeuI位点、5'Frt位点、新霉素盒、3'Frt位点、约123 kb片段(包含前12个hVλ基因区段)、位于hVλ3-1基因区段的九聚物序列下游(3')约200 bp的NotI位点、壮观霉素盒和AciSI位点。将包含人κ序列的单独的人BAC克隆(CTD-2366j12)以独立的两次靶向在hVκ4-1与hJκ1基因区段之间的位置上的工程化限制位点,以允许随后克隆约23 kb的片段,用于与双靶向的修饰的729g4 BAC克隆中包含的hVλ基因区段连接(图6, 上右)。

[0249] 简而言之,2366j12 BAC克隆在大小上为约132 kb并且包含hVκ基因区段1-6、1-5、2-4、7-3、5-2、4-1、在Vκ基因区段下游的人κ基因组序列、hJκ基因区段1-5、hCκ和约20 kb的人κ基因座的另外的基因组序列。首先用靶向载体靶向该克隆,所述靶向载体包含侧翼连接有Frt位点的潮霉素盒和在3'Frt位点下游(3')的NotI位点。用于该靶向载体的同源臂包含BAC克隆内的Vκ基因区段的5'和3'端的人基因组序列,从而在用该靶向载体进行同源重组后,Vκ基因区段被缺失并且NotI位点被工程化在hVκ4-1基因区段下游约133 bp(图6, 上右)。利用两个靶向载体在3'末端独立地靶向该修饰的2366j12 BAC克隆以用氯霉素盒使hJκ基因区段缺失,所述靶向载体还包含hJλ1基因区段、PI-SceI位点和AciSI位点或包含4个hJλ基因区段的人λ基因组片段(同上)、PI-SceI位点和AciSI位点(图6, 右上)。这两个相似的靶向载体的同源臂包含hJκ基因区段的5'和3'端的序列。利用此类第二靶向载体和修饰的2366j12 BAC克隆的同源重组产生双靶向的2366j12克隆,其从5'至3'包括5'Frt位点、潮霉素盒、3'Frt位点、NotI位点、人κ基因座的22,800 bp基因组片段(包含Vκ4-1与Jκ1基因区

段之间的基因间区域)、hJ λ 1基因区段或含有hJ λ 1、J λ 2、J λ 3和J λ 7的人 λ 基因组片段、PI-SceI位点和氯霉素盒(图6,右上)。使用双靶向的729g4和2366j12克隆通过两个连接步骤获得用于产生两个另外的修饰的两个最终的靶向载体。

[0250] 利用NotI和AseI降解双靶向的729g4和2366j12克隆,从而分别产生包含新霉素盒和hV λ 基因区段的一个片段,和包含人 κ 基因座的约23 kb基因组片段(包含在V κ 4-1与J κ 1基因区段之间的基因间区域)、hJ λ 1基因区段或包含hJ λ 1、J λ 2、J λ 3和J λ 7基因区段的基因组片段、PI-SceI位点和氯霉素盒的另一个片段。这些片段的连接产生两个独特的BAC克隆,其从5'至3'分别包含hV λ 基因区段、V κ 4-1与J κ 1基因区段之间的人 κ 基因组序列、hJ λ 1基因区段或包含hJ λ 1、J λ 2、J λ 3和J λ 7基因区段的基因组片段、PI-SceI位点和氯霉素盒(图6,下方)。随后用I-CeuI和PI-SceI降解这些新的BAC克隆以释放独特的片段(包含上游新霉素盒和连续的人 λ 及 κ 序列),将其连接入修饰的小鼠BAC克隆302g12,所述克隆从5'至3'包含内源 κ 基因座5'端的小鼠基因组序列、I-CeuI位点、5'Frt位点、新霉素盒、3'Frt位点、hV λ 基因区段(3-12至3-1)、V λ 3-1下游约200 bp的NotI位点、在人V κ 4-1与J κ 1基因区段之间天然发现的约23 kb的人 κ 序列、hJ λ 1基因区段或包含hJ λ 1、J λ 2、J λ 3和J λ 7基因区段的基因组片段、小鼠E κ i、小鼠C κ 基因和E κ 3'(图4,12hV λ -V κ J κ -hJ λ 1和12hV λ -V κ J κ -4hJ λ 靶向载体)。利用这两个靶向载体的同源重组产生了两个单独的修饰的小鼠 κ 轻链基因座,所述基因座包含与内源小鼠C κ 基因有效连接的12个hV λ 基因区段、人 κ 基因组序列以及1个或4个hJ λ 基因区段,所述基因座在重组后导致嵌合人 λ /小鼠 κ 轻链的形成。

[0251] 实施例III

[0252] 将另外的人V λ 基因区段工程化至人 λ 轻链微小基因座中

[0253] 使用类似的靶向载体和方法将另外的hV λ 基因区段独立地添加至实施例2中描述的每一个初始修饰中(图5A,+16- λ 靶向载体和图5B,+16- κ 靶向载体)。

[0254] 16个额外的人V λ 基因区段的引入。用于构建靶向载体(用于将16个另外的hV λ 基因区段添加至实施例2中描述的修饰的轻链基因座)的上游(5')同源臂包含内源 κ 或 λ 轻链基因座的5'端的小鼠基因组序列。3'同源臂与所有靶向载体一样,包含与实施例2中描述的修饰的人 λ 序列的5'末端重叠的人基因组序列。

[0255] 简而言之,工程化两个靶向载体(图5A和5B,+16- λ 或+16- κ 靶向载体)以用于将16个额外的hV λ 基因区段引入实施例2中描述的修饰的小鼠轻链基因座。工程化包含来自A簇的21个连续hV λ 基因区段的来自人BAC克隆RP11-761113(Invitrogen)的约172 kb DNA片段,以具有包含内源 κ 或 λ 轻链基因座的5'端的小鼠基因组序列的5'同源臂和包含人基因组 λ 序列的3'同源臂。此类靶向构建体中使用的5'小鼠 κ 或 λ 同源臂与实施例2中描述的5'同源臂相同(图5A和5B)。3'同源臂包括相应于实施例2中描述的人基因组 λ 序列的约123 kb片段的5'末端的等同物的人基因组 λ 序列的53,057 bp重叠。这两个靶向载体从5'至3'包括,5'小鼠同源臂(包含内源小鼠 κ 轻链基因座的5'端的约23 kb的基因组序列或内源 λ 轻链基因座的5'端的约24kb的小鼠基因组序列)、5'Frt位点、潮霉素盒、3'Frt位点和包含21个连续hV λ 基因区段的171,457 bp的人基因组 λ 序列,该人基因组 λ 序列的约53 kb与实施例3中描述的人 λ 序列的5'末端重叠,且用作该靶向构建体的3'同源臂(图5A和5B,+16- λ 或+16- κ 靶向载体)。利用此类靶向载体的同源重组独立地产生修饰的小鼠 κ 和 λ 轻链基因座,其各自包含有效地连接至内源小鼠恒定基因(C κ 或C λ 2)的28个hV λ 基因区段和hJ λ 1基因区段,所述基

因座在重组后导致嵌合轻链的形成。

[0256] 还以类似的方式将+16- κ 靶向载体用于将16个另外的hV λ 基因区段引入实施例2中描述的其它初始修饰,所述修饰整合了多个hJ λ 基因区段,且具有和不具有整合的人 κ 序列(图4B)。利用该靶向载体在包含其它初始修饰的内源小鼠 κ 基因座上进行的同源重组产生包含与内源小鼠C κ 基因有效连接的28个hV λ 基因区段和hJ λ 1、2、3和7基因区段(具有和不具有人V κ -J κ 基因组序列)的小鼠 κ 轻链基因座,所述基因座在重组后导致嵌合 λ - κ 轻链的形成。

[0257] 12个另外的人V λ 基因区段的引入。使用相似的靶向载体和方法将另外的hV λ 基因区段独立地添加至上文描述的每一个修饰中。与包含另外的hV λ 基因区段的靶向载体的同源重组所产生的最终的基因座结构示于图7A和7B中。

[0258] 简而言之,工程化靶向载体(图5A和5B,+12- λ 或12- κ 靶向载体)以将12个另外的hV λ 基因区段引入上述修饰的小鼠 κ 和 λ 轻链基因座。工程化来自人BAC克隆RP11-22I18(Invitrogen)的包含12个连续的来自B簇的hV λ 基因区段的93,674 bp DNA片段,以具有5'同源臂(包含内源小鼠 κ 或 λ 轻链基因座的5'端的小鼠基因组序列)和3'同源臂(包含人基因组 λ 序列)。用于该靶向构建体的5'同源臂与上述用于添加16个hV λ 基因区段的5'同源臂相同(图5A和5B)。通过将PI-SceI位点工程化至来自BAC克隆RP11-761I13的人 λ 序列的27,468 bp基因组片段中包含的人V λ 3-29P基因区段的5'端约3431bp来产生3'同源臂。该PI-SceI位点用作连接点,以使用+16- λ 或+16- κ 靶向载体(图5A和5B)将另外的约94 kb人 λ 序列的片段连接至约27 kb人 λ 序列片段(其与先前的修饰中的人 λ 序列的5'末端重叠)。这两个靶向载体从5'至3'包括,5'同源臂(包含内源 κ 轻链基因座的5'端的约23 kb的小鼠基因组序列或内源 λ 轻链基因座的5'端的约24kb的小鼠基因组序列)、5'Frt位点,新霉素盒、3'Frt位点和121,188bp的包含16个hV λ 基因区段和PI-SceI位点的人基因组 λ 序列,所述人基因组 λ 序列的约27 kb与获自16个另外的hV λ 基因区段的插入的人 λ 序列的5'末端重叠,且用作该靶向构建体的3'同源臂(图5A和5B,+12- λ 或12- κ 靶向载体)。利用此类靶向载体的同源重组独立地产生修饰的小鼠 κ 和 λ 轻链基因座(其包含有效地连接至内源小鼠恒定基因(C κ 或C λ 2)的40个hV λ 基因区段和人J λ 1),所述基因座在重组后导致嵌合轻链的形成(图5A和5B的下方)。

[0259] 还以类似的方式将+12- κ 靶向载体用于将12个另外的hV λ 基因区段引入其它初始修饰,所述初始修饰整合了多个hJ λ 基因区段,且具有和不具有整合的人 κ 序列(图4B)。利用该靶向载体在包含其它修饰的内源小鼠 κ 基因座上的进行同源重组产生小鼠 κ 轻链基因座,所述基因座包含有效地连接至内源小鼠C κ 基因的40个hV λ 基因区段和hJ λ 1、2、3和7基因区段(具有和不具有人V κ -J κ 基因组序列),所述基因座在重组后,导致嵌合 λ - κ 轻链的形成。

[0260] 实施例IV

[0261] 具有人 λ 轻链基因区段的靶向的ES细胞的鉴定

[0262] 按照上述实施例制备的靶向的BAC DNA被用于对小鼠ES细胞进行电穿孔,以产生修饰的ES细胞,所述细胞用于产生表达人 λ 轻链基因区段的嵌合小鼠。通过定量Taqman®测定来鉴定包含未重排的人 λ 轻链基因区段的插入的ES细胞。针对人 λ 序列和相关选择盒的插入(等位基因的获得,GOA)、内源小鼠序列和任何选择盒的丢失(等位基因的丢失,LOA)以及侧翼小鼠序列的保留(等位基因保留,AR),设计特异性引物组和探针。对于人 λ 序列的每一次额外插入,使用另外的引物组和探针来确认另外的人 λ 序列的存在,且使用先前的引物

组和探针来确认先前靶向的人序列的保留。表1显示了用于定量PCR测定的引物和相关探针。表2显示了用于确认人 λ 轻链基因区段的每一个部分在ES细胞克隆中的插入的组合。

[0263] 任选地用表达FLP的构建体转染具有人 λ 轻链基因区段的ES细胞,以除去通过插入包含人V λ 5-52-V λ 1-40基因区段的靶向构建体而引入的Frt'ed新霉素盒(图5A和5B)。新霉素盒可任选地通过与表达FLP重组酶的小鼠交配来除去(例如,US 6,774,279)。任选地,新霉素盒被保留在小鼠中。

[0264] 表1

引物	SEQ ID NO:	探针	SEQ ID NO:
hL2F	2	hL2P	24
hL2R	3		
hL3F	4	hL3P	25
hL3R	5		
NeoF	6	NeoP	26
NeoR	7		
61hJ1F	8	61hJ1P	27
61hJ1R	9		
67hT1F	10	67hT1P	28
67hT1R	11		

[0266]

67hT3F	12	67hT3P	29
67hT3R	13		
HygF	14	HygP	30
HygR	15		
MKD2F	16	MKD2P	31
MKD2R	17		
MKP8F	18	MKP8P	32
MKP8R	19		
MKP15F	20	MKP15P	33
MKP15R	21		
MK20F	22	—	—
MKP4R	23		
68h2F	34	68h2P	38
68h2R	35		
68h5F	36	68h5P	39
68h5R	37		
mL1F	75	mL1P	83
mL1R	76		
mL2F	77	mL2P	84
mL2R	78		
mL11F	79	mL11P	85
mL11R	80		
mL12F	81	mL12P	86
mL12R	82		

[0267] 表2

[0268]

修饰	测定	正向/反向引物组	探针	序列位置
12 个 hVλ & hJλ 1 的插入	GOA	hL2F/hL2R	hL2P	hVλ 3-12 - hVλ 3-1
		hL3F/hL3R	hL3P	
		61hJ1F/61hJ1R	61hJ1P	hJλ 序列
		NeoF/NeoR	NeoP	新霉素盒
	LOA	MK20F/MKP4R	—	失活的 κ 基因座的 lox511/loxP 序列
		HygF/HygR	HygP	来自失活的 λ 基因座的潮霉素盒
		mL1F/mL1R	mL1P	小鼠 Vλ 1-Cλ 1 簇

[0269]

		mL2F/mL2R	mL2P	小鼠 Vλ 2-Cλ 2 簇
		mL11F/mL11R	mL11P	
		mL12F/mL12R	mL12P	
	AR/LOA	MKD2F/MKD2R	MKD2P	5' Vκ 基因座中的小鼠序列
		MKP15F/MKP15R	MKP15P	3' Vκ 基因座中的小鼠序列
16 个 hVλ 的插入	GOA	67hT1F/67hT1R	67hT1P	hVλ 3-27 - hVλ 3-12
		67hT3F/67hT3R	67hT3P	
		HygF/HygR	HygP	潮霉素盒
	LOA	NeoF/NeoR	NeoP	新霉素盒
		mL1F/mL1R	mL1P	小鼠 Vλ 1-Cλ 1 簇
		mL2F/mL2R	mL2P	
		mL11F/mL11R	mL11P	小鼠 Vλ 2-Cλ 2 簇
		mL12F/mL12R	mL12P	
	AR	hL2F/hL2R	hL2P	hVλ 3-12 - hVλ 3-1
		hL3F/hL3R	hL3P	
	AR/LOA	MKD2F/MKD2R	MKD2P	5' Vκ 基因座中的小鼠序列
		MKP15F/MKP15R	MKP15P	3' Vκ 基因座中的小鼠序列
12 个 hVλ 的插入	GOA	68h2F/68h2R	68h2P	hVλ 5-52 - hVλ 1-40
		68h5F/68h5R	68h5P	
		NeoF/NeoR	NeoP	新霉素盒
	LOA	HygF/HygR	HygP	潮霉素盒
		mL1F/mL1R	mL1P	小鼠 Vλ 1-Cλ 1 簇
		mL2F/mL2R	mL2P	
		mL11F/mL11R	mL11P	小鼠 Vλ 2-Cλ 2 簇
		mL12F/mL12R	mL12P	
	AR	hL2F/hL2R	hL2P	hVλ 3-12 - hVλ 3-1
		hL3F/hL3R	hL3P	
		67hT1F/67hT1R	67hT1P	hVλ 3-27 - hVλ 3-12
		67hT3F/67hT3R	67hT3P	
	AR/LOA	MKD2F/MKD2R	MKD2P	5' Vκ 基因座中的小鼠序列
		MKP15F/MKP15R	MKP15P	3' Vκ 基因座中的小鼠序列

[0270] 实施例V

[0271] 从内源轻链基因座表达人λ轻链的小鼠的产生

[0272] 将上述靶向的ES细胞用作供体ES细胞并且通过 **VELOCIMOUSE®** 法(参见,例如,美国专利7,294,754和Poueymirou等人(2007) F0 generation mice that are essentially fully derived from the donor gene-targeted ES cell allowing immediate phenotypic analyses Nature Biotech.25(1):91-99) 将其引入8细胞期小鼠胚胎。独立地具有人 λ 基因区段的 **VELOCIMICE®** (F0小鼠完全来源于供体ES细胞) 通过使用检测独特的人 λ 基因区段(同上)的存在的等位基因测定的改进法(Valenzuela等人,同上)进行基因分型来鉴定。

[0273] 具有人 λ 轻链基因区段的小鼠的 κ : λ 轻链使用。使用流式细胞术分析对于与单个hJ λ 基因区段一起的hV λ 基因区段的3次连续插入的每一次插入(图5B)是纯合的小鼠和对于与单个hJ λ 基因区段或4个人J λ 基因区段(包括人V κ -J κ 基因组序列)一起的hV λ 基因区段的第一次插入(图4B)是纯合的小鼠的 κ 和 λ 轻链在脾细胞中的表达。

[0274] 简而言之,从成组的小鼠(每组3至7只动物)收获脾,使用载玻片将其磨碎。在用ACK裂解缓冲液(Lonza Walkersville)裂解红细胞(RBC)后,用特异于小鼠CD19(克隆1D3; BD Biosciences)、小鼠CD3(17A2; Biolegend)、小鼠Ig κ (187.1; BD Biosciences)和小鼠Ig λ (RML-42; Biolegend)的缀合有荧光染料的抗体对脾细胞进行染色。使用BD™ LSR II流式细胞仪(BD Biosciences)获得数据,使用FLOWJO™软件(Tree Star, Inc)分析数据。表3显示了在来自具有各个遗传修饰的成组动物的脾细胞中观察到的B细胞(CD19⁺)、 κ 轻链(CD19⁺Ig κ ⁺Ig λ ⁻)和 λ 轻链(CD19⁺Ig κ ⁻Ig λ ⁺)表达的平均百分数值。

[0275] 在类似的实验中,使用流式细胞术(如上所述)分析脾室(splenic compartment)的B细胞内容物的Ig κ 和Ig λ 表达,所述脾室来自对于有效地连接至小鼠C κ 基因的12个hV λ 和4个hJ λ 基因区段(包括人V κ -J κ 基因组序列)的第一次插入(图4B的下方)是纯合的小鼠和对于40个hV λ 和1个hJ λ 基因区段(图5B的下方或图7B的上方)是纯合的小鼠。图8A显示来自各组的代表性小鼠的CD19⁺B细胞的Ig λ 和Ig κ 表达。还记录了每一只小鼠的CD19⁺B细胞数目/脾(图8B)。

[0276] 在另一个实验中,使用不同细胞表面标记的流式细胞术,就通过B细胞发育的进展分析对于有效地连接至小鼠C κ 基因的40个hV λ 和4个hJ λ 基因区段(包括人V κ -J κ 基因组序列)(图7B的下方)是纯合的小鼠的脾室和骨髓室(bone marrow compartment)的B细胞内容物。

[0277] 简而言之,将两组小鼠(每一组N=3, 9-12周龄,雄性和雌性):野生型小鼠和对于有效地连接至小鼠C κ 基因的40个hV λ 和4个hJ λ 基因区段(包括人V κ -J κ 基因组序列)是纯合的小鼠处死,随后收获脾和骨髓。通过用完全RPMI培养基(补充有胎牛血清、丙酮酸钠、Hepes、2-巯基乙醇、非必需氨基酸和庆大霉素的RPMI培养基)冲洗从股骨收集骨髓。用ACK裂解缓冲液(Lonza Walkersville)裂解脾和骨髓制备物的RBC,随后用完全RPMI培养基进行洗涤。将 1×10^6 个细胞与抗-小鼠CD16/CD32(2.4G2, BD Biosciences)一起在冰上温育10分钟,然后用选择的抗体组进行标记,在冰上进行30分钟。

[0278] 骨髓组:抗-小鼠FITC-CD43(1B11, BioLegend)、PE-ckit(2B8, BioLegend)、PeCy7-IgM(II/41, eBioscience)、PerCP-Cy5.5-IgD(11-26c.2a, BioLegend)、APC-B220(RA3-6B2, eBioscience)、APC-H7-CD19(1D3, BD)和Pacific Blue-CD3(17A2, BioLegend)。

[0279] 骨髓和脾组:抗-小鼠FITC-Ig κ (187.1, BD)、PE-Ig λ (RML-42, BioLegend)、PeCy7-

IgM (II/41, ebioscience)、PerCP-Cy5.5-IgD (11-26c.2a, BioLegend)、Pacific Blue-CD3 (17A2, BioLegend)、APC-B220 (RA3-6B2, eBioscience)、APC-H7-CD19 (ID3, BD)。

[0280] 染色后, 洗涤细胞, 将其在2%甲醛中进行固定。在FACSCANTOII™流式细胞仪 (BD Biosciences) 上获取数据, 用FLOWJO™软件 (Tree Star, Inc.) 分析数据。图9A-9D显示来自各组的一只代表性小鼠的脾室的结果。图10A-10E显示来自各组的一只代表性小鼠的骨髓室的结果。表4显示了在来自具有不同遗传修饰的动物组的脾细胞中观察到的B细胞 (CD19⁺)、κ轻链 (CD19⁺Igκ⁺Igλ⁻) 和λ轻链 (CD19⁺Igκ⁻Igλ⁺) 表达的平均百分数值。表5显示了在野生型小鼠和对于有效地连接至小鼠Cκ基因的40个hVλ和4个hJλ基因区段 (包括人Vκ-Jκ基因组序列) 是纯合的小鼠的骨髓中观察到的B细胞 (CD19⁺)、成熟B细胞 (B220^{hi}IgM⁺)、未成熟B细胞 (B220^{int}IgM⁺)、表达κ轻链的未成熟B细胞 (B220^{int}IgM⁺Igκ⁺) 和表达λ轻链的未成熟B细胞 (B220^{int}IgM⁺Igλ⁺) 的平均百分数值。用其它组的上述小鼠重复本实验, 显示了相似的结果 (数据未显示)。

[0281] 表3

[0282]	基因型	B 细胞%	Ig κ ⁺ %	Ig λ ⁺ %
	野生型	46.2	91.0	3.6
	12 hVλ +hJλ 1	28.3	10.4	62.5
	12 hVλ -Vκ Jκ -hJλ 1	12.0	11.0	67.5
	12 hVλ -Vκ Jκ -4hJλ	41.8	17.2	68.4
	28 hVλ +hJλ 1	22.0	13.3	51.1
	40 hVλ +hJλ 1	28.2	24.3	53.0

[0283] 表4

[0284]	基因型	B 细胞%	Ig κ ⁺ %	Ig λ ⁺ %
	野生型	49.8	91.2	3.5
	40 hVλ -Vκ Jκ -4hJλ	33.3	41.6	43.1

[0285] 表5

[0286]

基因型	B 细胞%	成熟 B 细胞%	未成熟 B 细胞%	未成熟 Ig κ ⁺ B 细胞%	未成熟 Ig λ ⁺ B 细胞%
野生型	62.2	9.2	12.0	79.0	8.84
40hVλ -Vκ Jκ -4hJλ	60.43	2.59	7.69	38.29	43.29

[0287] 具有人λ轻链基因区段的小鼠的人Vλ基因使用。使用从脾细胞分离的RNA, 通过逆转录酶聚合酶链式反应 (RT-PCR), 就人λ轻链基因使用, 分析对于人λ序列的第一次插入 (hVλ3-12-hVλ3-1和hJλ1, 图5B) 是纯合的以及对于人λ序列的第三次插入 (hVλ5-52-hVλ3-1和

hJλ1,图5B)是纯合的小鼠。

[0288] 简而言之,收获脾,并在无菌一次性袋中用10 mL具有5%HI-FBS的RPMI-1640 (Sigma)灌注脾。随后将每一个包含单个脾的袋子置于STOMACHER™ (Seward)中,并在培养基中匀浆30秒。使用0.7μm细胞滤过滤器过滤匀浆的脾,随后通过离心(1000 rpm,10分钟)进行沉淀,并且在BD PHARM LYSER™ (BD Biosciences)中裂解RBC 3分钟。脾细胞用RPMI-1640稀释,再次离心,随后于1 mL PBS (Irvine Scientific)中重悬浮。使用本领域已知的标准技术从沉淀的脾细胞分离RNA。

[0289] 使用特异于人hVλ基因区段和小鼠Cκ基因的引物(表6)对脾细胞RNA进行RT-PCR。将PCR产物凝胶纯化,并克隆入pCR2.1-TOPO TA载体 (Invitrogen),然后用位于载体内的在侧翼连接克隆位点的位置上的引物M13正向 (GTAAAACGAC GGCCAG;SEQ ID NO:55)和M13反向 (CAGGAAACAG CTATGAC;SEQ ID NO:56)进行测序。对来源于人λ序列的第一次和第三次插入的84个克隆进行测序,以测定hVλ基因的使用(表7)。所选择的RT-PCR克隆的hVλ-hJλ1-mCκ连接的核苷酸序列示于图11中。

[0290] 以类似的方式,使用从脾细胞分离的RNA,通过RT-PCR就人λ轻链基因使用,分析对于有效地连接至内源小鼠Cκ基因的人λ轻链基因序列的第三次插入(即40个hVλ基因区段和4个hJλ基因区段,包括人Vκ-Jκ基因组序列,图7B的下方)是纯合的小鼠(如上文所述)。26个选择的RT-PCR克隆的人λ轻链基因区段的使用示于表8中。选择的RT-PCR克隆的hVλ-hJλ-mCκ连接的核苷酸序列示于图12中。

[0291] 以类似的方式,使用从脾细胞分离的RNA,通过RT-PCR就人λ轻链基因使用,分析对于有效地连接至内源小鼠Cλ2基因的人λ轻链基因区段的第一次插入(12个hVλ基因区段和hJλ1,图4A及图5A)是纯合的小鼠(如上文所述)。将特异于hVλ基因区段的引物(表6)与特异于小鼠Cλ2基因的两个引物之一(Cλ2-1 (SEQ ID NO:104)或Cλ2-2 (SEQ ID NO:105))配对。

[0292] 从在内源小鼠λ轻链基因座上具有人λ轻链基因区段的小鼠的RT-PCR克隆观察到,多个hVλ基因区段重排至hλ1。选择的RT-PCR克隆的hVλ-hJλ-mCλ2连接的核苷酸序列示于图13中。

[0293] 表6

[0294]

5' hV λ 引物

序列 (5'-3')

SEQ ID NO:

VLL-1	CCTCTCCTCC TCACCCTCCT	40
VLL-1n	ATGRCCDGST YYYCTCTCCT	41
VLL-2	CTCCTCACTC AGGGCACA	42
VLL-2n	ATGGCCTGGG CTCTGCTSCT	43
VLL-3	ATGGCCTGGA YCSTCTCC	44
VLL-4	TCACCATGGC YTGGRYCYCM YTC	45
VLL-4. 3	TCACCATGGC CTGGGTCTCC TT	46
VLL-5	TCACCATGGC CTGGAMTCYT CT	47
VLL-6	TCACCATGGC CTGGGCTCCA CTACTT	48
VLL-7	TCACCATGGC CTGGACTCCT	49
VLL-8	TCACCATGGC CTGGATGATG CTT	50
VLL-9	TAAATATGGC CTGGGCTCCT CT	51
VLL-10	TCACCATGCC CTGGGCTCTG CT	52
VLL-11	TCACCATGGC CCTGACTCCT CT	53

3' 小鼠 C κ 引物

序列 (5'-3')

SEQ ID NO:

mlgKC3'-1	CCCAAGCTTA CTGGATGGTG GGAAGATGGA	54
-----------	----------------------------------	----

[0295]

表7

表8

[0296]

hV λ	观察到的克隆的编号
3-1	2
4-3	3
2-8	7

克隆	hV λ	hJ λ
1-3	1-44	7
1-5	1-51	3
2-3	9-49	7

[0297]

3-9	4
3-10	3
2-14	1
3-19	1
2-23	7
3-25	1
1-40	9
7-43	2
1-44	2
5-45	8
7-46	3
9-49	6
1-51	3

2-5	1-40	1
2-6	1-40	7
3b-5	3-1	7
4a-1	4-3	7
4a-5	4-3	7
4b-1	1-47	3
5-1	3-10	3
5-2	1-40	7
5-3	1-40	7
5-4	7-46	2
5-6	1-40	7
5-7	7-43	3
6-1	1-40	1
6-2	1-40	2
6-7	1-40	3
7a-1	3-10	7
7a-2	9-49	2
7a-7	3-10	7
7b-2	7-43	3
7b-7	7-46	7
7b-8	7-43	3
11a-1	5-45	2
11a-2	5-45	7

[0298] 图11显示了,来自具有与单个hJ λ 基因区段一起的hV λ 基因区段的第一次和第三次插入的小鼠的RT-PCR克隆的hV λ -hJ λ 1-mC κ 连接的序列。图11中显示的序列说明,牵涉不同的hV λ 基因区段的独特重排,且hJ λ 1重组至小鼠C κ 基因。具有单个包含12个hV λ 基因区段和hJ λ 1的修饰的内源 κ 基因座的杂合小鼠和具有两个包含40个hV λ 基因区段和hJ λ 1的修饰的内源 κ 基因座的纯合小鼠都能产生有效地连接至小鼠C κ 基因的人 λ 基因区段和产生表达人 λ 轻链的B细胞。这些重排显示,嵌合基因座能够在此类小鼠的多个独立的B细胞中独立地重排人 λ 基因区段。此外,对内源 κ 轻链基因座的此类修饰并未使任何hV λ 基因区段不能操作或阻止嵌合基因座在B细胞发育过程中重组多个hV λ 和hJ λ (J λ 1) 基因区段,如通过观察到16个不同hV λ 基因区段与hJ λ 1重排(表7)所证明的。此外,此类小鼠产生功能性抗体,所述抗体包含有效地连接至小鼠C κ 基因的重排的人V λ -J λ 基因区段(作为内源免疫球蛋白轻链库的部分)。

[0299] 图12显示,从对于40个hV λ 和4个hJ λ 基因区段(包括人V κ -J κ 基因组序列)是纯合的小鼠选择的RT-PCR克隆的hV λ -hJ λ -mC κ 连接的序列。图12中显示的序列说明另外的独特重排,所述重排牵涉多个不同的hV λ 基因区段(横跨整个嵌合基因座),且多个不同的hJ λ 基因区段重排并且有效地连接至小鼠C κ 基因。具有包含40个hV λ 和4个hJ λ 基因区段的修饰的内源 κ 基因座的纯合小鼠也能够产生有效地连接至小鼠C κ 基因的人 λ 基因区段和产生表达人 λ

轻链的B细胞。此类重排还显示,所有时期的嵌合基因座都能够在小鼠的多个独立的B细胞中独立地重排人 λ 基因区段。此外,对内源 κ 轻链基因座的此类其它修饰显示,人 λ 基因区段的每一次插入并未使任何hV λ 和/或J λ 基因区段不能操作或阻止嵌合基因座在B细胞发育过程中重组hV λ 和J λ 基因区段,如通过在26个选择的RT-PCR克隆中观察到12个不同hV λ 基因区段与所有4个hJ λ 基因区段重排(表8)所证明的。此外,此类小鼠同样也产生功能性抗体,所述抗体包含有效地连接至小鼠C κ 区的人V λ -J λ 基因区段(作为内源免疫球蛋白轻链库的部分)。

[0300] 图13显示了来自对于12个hV λ 基因区段和hJ λ 1是纯合的小鼠的3个单独的RT-PCR克隆的hV λ -hJ λ -mC λ 2连接的序列。图13中显示的序列说明另外的独特重排,所述重排牵涉不同的hV λ 基因区段(横跨第一插入的长度),且hJ λ 1重排并且有效地连接至小鼠C λ 2基因(2D1=V λ 2-8J λ 1;2D9=V λ 3-10J λ 1;3E15=V λ 3-1J λ 1)。一个克隆显示,因hV λ -hJ λ 连接上的N添加而导致的非生产性重排(2D1,图13)。这在V(D)J组合中并不少见,因为已显示,重组过程中基因区段的连接是不精确的。虽然该克隆代表存在于此类小鼠的轻链库中的非生产性重组体,但这显示了,促成抗体基因之间的连接多样性的遗传机制在此类小鼠中正常地运作并且导致具有更大多样性的包含轻链的抗体库。

[0301] 具有包含12个hV λ 基因区段和hJ λ 1的修饰的内源 λ 基因座的纯合小鼠也能够产生有效地连接至内源小鼠C λ 基因的人 λ 基因区段和产生表达包含连接至小鼠C λ 区的人V λ 区的反向嵌合 λ 轻链的B细胞。此类重排还显示,置于其它轻链基因座(即, λ 基因座)上的人 λ 轻链基因区段能够在此类小鼠的多个独立的B细胞中独立地重排人 λ 基因区段。此外,对内源 λ 轻链基因座的修饰显示,人 λ 基因区段的插入并未使任何hV λ 和/或hJ λ 1基因区段不可操作或阻止嵌合基因座在B细胞发育过程中重组hV λ 和hJ λ 1基因区段。此外,此类小鼠也产生功能性抗体,所述抗体包含有效地连接至小鼠C λ 区的人V λ -J λ 基因区段(作为内源免疫球蛋白轻链库的部分)。

[0302] 如本实施例中所示,在内源 κ 和 λ 轻链基因座上具有人 λ 轻链基因区段的小鼠能够重排人 λ 轻链基因区段和在小鼠C κ 和/或C λ 区的背景中表达它们(作为小鼠的正常抗体库的部分),因为在脾和骨髓中在B细胞发育的不同检查点上都需要功能性轻链。此外,与野生型同窝出生仔相比较,此类小鼠中的B细胞的早期亚组(例如,前B细胞、祖B细胞和过渡B细胞)显示正常的表型(图9D、10A和10B)。观察到骨髓和外周B细胞群体中的小缺陷,这可归因于自体反应性未成熟B细胞的亚组的缺失和/或人 λ 轻链与小鼠重链的亚最佳缔合。然而,此类小鼠中观察到的Ig κ /Ig λ 使用显示了这样的情形:其与小鼠中观察到的相比,更接近人轻链表达。

[0303] 实施例VI

[0304] 从内源轻链基因座表达人 λ 轻链的小鼠的培育

[0305] 为了优化内源小鼠轻链基因座上的人 λ 基因区段的使用,将具有未重排的人 λ 基因区段的小鼠与在相对的内源轻链基因座(κ 或 λ)中包含缺失的另一种小鼠交配。例如,位于内源 κ 基因座上的人 λ 基因区段可以是存在于还在内源 λ 轻链基因座中具有缺失的小鼠中的唯一功能性轻链基因区段。这样,所获得的后代仅表达人 λ 轻链,如在上述实施例中描述的。通过本领域公认的标准技术或通过商业公司例如The Jackson Laboratory来进行育种。就独特的反向-嵌合(人-小鼠) λ 轻链的存在和内源小鼠 λ 轻链的不存在,筛选在内源 κ 基因座

上具有人 λ 轻链基因区段并且具有内源 λ 轻链基因座的缺失的小鼠品系。

[0306] 还将具有未重排的人 λ 轻链基因座的小鼠与包含利用人重链可变基因基因座对内源小鼠重链可变基因基因座的替换的小鼠(参见US6,596,541, Regeneron Pharmaceuticals, **VELOCIMMUNE®**基因工程小鼠)交配。**VELOCIMMUNE®**小鼠包括, 部分地, 具有包含有效地连接至内源小鼠恒定区基因座的人重链可变区的基因组, 以便所述小鼠响应抗原刺激产生包含人重链可变区和小鼠重链恒定区的抗体。可分离编码抗体重链可变区的DNA, 将其有效地连接至编码人重链恒定区的DNA。随后可将DNA在能够表达完全人抗体重链的细胞中进行表达。在进行适当的育种方案后, 获得这样的小鼠, 其具有利用人重链基因座对内源小鼠重链基因座的替换和在内源 κ 轻链基因座上具有未重排的人 λ 轻链基因座。可在用目标抗原免疫后分离包含体细胞突变的人重链可变区和人 λ 轻链可变区的抗体。

[0307] 实施例VII

[0308] 从表达人重链和人 λ 轻链的小鼠产生抗体

[0309] 在将包含未重排的人 λ 轻链基因座的小鼠与包含其它内源Ig基因座的修饰和缺失的各种期望的品系(如上文中所描述的)交配后, 用目标抗原免疫所选择的小鼠。

[0310] 一般而言, 用抗原攻击包含单个重排的人种系轻链区之一的**VELOCIMMUNE®**小鼠, 从动物的血清回收淋巴细胞(例如B细胞)。可将淋巴细胞与骨髓瘤细胞系融合以制备永生杂交瘤细胞系, 随后筛选和选择此类杂交瘤细胞系以鉴定产生特异于用于免疫的抗原的包含人重链和人 λ 轻链的抗体的杂交瘤细胞系。可分离编码重链和 λ 轻链的可变区的DNA, 将其连接至期望的重链和轻链的同种型恒定区。由于存在额外的hV λ 基因区段, 因此与内源小鼠 λ 基因座相比较, 轻链库的多样性急剧增加, 并且在免疫后赋予抗原特异性库更高的多样性。随后可在细胞例如CHO细胞中产生所得的克隆的抗体序列。可者, 可从抗原特异性淋巴细胞(例如, B细胞)直接分离编码抗原特异性嵌合抗体或轻链和重链的可变结构域的DNA。

[0311] 最初, 分离了具有人可变区和小鼠恒定区的高亲和力嵌合抗体。如上所述, 就期望的特征包括亲和力、选择性、表位等表征和选择抗体。用期望的人恒定区替换小鼠恒定区以产生包含体细胞突变的人重链和来源于本发明的未重排的人 λ 轻链基因座的人 λ 轻链的完全人抗体。适当的人恒定区包括例如, 野生型或经修饰的IgG1、IgG2、IgG3或IgG4。

序列表

<110> Regeneron Pharmaceuticals, Inc.

<120> 杂交轻链小鼠

<130> 795A-W0

<140> To be assigned

<141> Filed herewith

<150> 61/357,314

<151> 2010-06-22

<150> 61/357,317

<151> 2010-06-22

<160> 105

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 219

<212> DNA

<213> 人工序列

[0001]

<220>

<223> 合成的

<400> 1

```
acttttcagaa tgttcttgaa cagtctctga gaaacacgga agacggccgc ataacttcgt 60
atagtataca ttatacgaag ttattctaga cccccgggct cgataactat aacggtccta 120
aggtagcgac tcgagataac ttctgtataat gtatgctata cgaagtatc catggtaagc 180
ttacgtggca tacagtggtca gatcttctgt ttatcaagc 219
```

<210> 2

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 2

```
agctgaatgg aaacaaggca a
```

21

<210> 3

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 合成的

	<400> 3 ggagacaatg ccccagtga	19
	<210> 4 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 4 tcccataggg ctaggatttc c	21
	<210> 5 <211> 19 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 5 tcccctcaca ctgttcccc	19
[0002]	<210> 6 <211> 19 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 6 ggtggagagg ctattcggc	19
	<210> 7 <211> 17 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 7 gaacacggcg gcatcag	17
	<210> 8 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 8	

	tcaacctttc ccagcctgtc t	21
	<210> 9	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 9	
	ccccagagag agaaaacaga tttt	24
	<210> 10	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 10	
	ccctggtgaa gcatgtttgc	20
	<210> 11	
	<211> 20	
	<212> DNA	
[0003]	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 11	
	tgtggcctgt ctgccttacg	20
	<210> 12	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 12	
	cacacctaga ccccggaagt c	21
	<210> 13	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 13	
	tcgctttgcc agttgattct c	21

	<210> 14	
	<211> 17	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 14	
	tgcgccgat cttagcc	17
	<210> 15	
	<211> 18	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 15	
	ttgaccgatt ccttgccg	18
	<210> 16	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
[0004]	<220>	
	<223> 合成的	
	<400> 16	
	gcaaacaaaa accactggcc	20
	<210> 17	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 17	
	ggccacattc catgggttc	19
	<210> 18	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 18	
	ccatgactgg gcctctgtag ac	22

	<210> 19	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 19	
	caagtcaggg tgctaatgct gtatc	25
	<210> 20	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 20	
	cacagcttgt gcagcctcc	19
	<210> 21	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
[0005]	<220>	
	<223> 合成的	
	<400> 21	
	gggcactgga tacgatgtat gg	22
	<210> 22	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 22	
	tcataggtag gtctcagttt g	21
	<210> 23	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 23	
	tgatctgcgc tgtttcatcc t	21
	<210> 24	

	<211> 31	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 24	
	tgacatgaac catctgtttc tctctcgaca a	31
	<210> 25	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 25	
	agagacgctc cgaggtcaag gtgctctag	29
	<210> 26	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
[0006]	<223> 合成的	
	<400> 26	
	tgggcacaac agacaatcgg etg	23
	<210> 27	
	<211> 16	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 27	
	accctctgct gtccct	16
	<210> 28	
	<211> 26	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 28	
	ccaagcagga ggtgctcagt tcccaa	26
	<210> 29	
	<211> 24	

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 29	
	tccacactgt cggetgggag ctca	24
	<210> 30	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 30	
	acgagcgggt tcggcccatt c	21
	<210> 31	
	<211> 37	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
[0007]	<400> 31	
	ctgttcctct aaaactggac tccacagtaa atggaaa	37
	<210> 32	
	<211> 27	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 32	
	tgccgcttat acaacactgc catctgc	27
	<210> 33	
	<211> 37	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 33	
	agaagaagcc tgtactacag catccgtttt acagtca	37
	<210> 34	
	<211> 21	
	<212> DNA	

	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 34	
	gggetacttg aggaccttgc t	21
	<210> 35	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 35	
	gacagccctt acagagtttg gaa	23
	<210> 36	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
[0008]	<400> 36	
	aagaccagga gctetgccta agt	23
	<210> 37	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 37	
	cccatcacga actgaagttg ag	22
	<210> 38	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 38	
	cagggcctcc atcccaggca	20
	<210> 39	
	<211> 28	
	<212> DNA	
	<213> 人工序列	

	<220>		
	<223> 合成的		
	<400> 39		
	ccccagtgtg tgaatcactc taccctcc		28
	<210> 40		
	<211> 20		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的		
	<400> 40		
	cctctctctcc tcacctctct		20
	<210> 41		
	<211> 20		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的		
[0009]	<220>		
	<221> 变异		
	<222> (4)... (4)		
	<223> r=a 或 g		
	<220>		
	<221> 变异		
	<222> (9)... (9)		
	<223> s=c 或 g		
	<220>		
	<221> 变异		
	<222> 11, 12, 13		
	<223> y=c 或 t		
	<400> 41		
	atgrecdgst yyyctctect		20
	<210> 42		
	<211> 18		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的		
	<400> 42		
	ctcctcactc agggcaca		18
	<210> 43		

	<211> 20 <212> DNA <213> 人工序列 <220> <223> 合成的 <220> <221> 变异 <222> (18)... (18) <223> s=c 或 g <400> 43 atggcctggg ctctgctset	20
	<210> 44 <211> 19 <212> DNA <213> 人工序列 <220> <223> 合成的 <220> <221> 变异 <222> (11)... (11) <223> y=c 或 t	
[0010]	<220> <221> 变异 <222> (13)... (13) <223> s=c 或 g <400> 44 atggcctgga ycsctctec	19
	<210> 45 <211> 23 <212> DNA <213> 人工序列 <220> <223> 合成的 <220> <221> 变异 <222> 11, 16, 18, 21 <223> y=c 或 t <220> <221> 变异 <222> (15)... (15) <223> r=a 或 g <220> <221> 变异	

	<222> (20)... (20)	
	<223> m=a 或 c	
	<400> 45	
	tcaccatggc ytggrycycm ytc	23
	<210> 46	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 46	
	tcaccatggc ctgggtctcc tt	22
	<210> 47	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<220>	
	<221> 变异	
[0011]	<222> (16)... (16)	
	<223> m=a 或 c	
	<220>	
	<221> 变异	
	<222> (19)... (19)	
	<223> y=c 或 t	
	<400> 47	
	tcaccatggc ctggamtcyt ct	22
	<210> 48	
	<211> 26	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 48	
	tcaccatggc ctgggtcca ctactt	26
	<210> 49	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	

	<400> 49 tcaccatggc ctggactcct	20
	<210> 50 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 50 tcaccatggc ctggatgatg ctt	23
	<210> 51 <211> 22 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 51 taaatatggc ctgggctcct ct	22
[0012]	<210> 52 <211> 22 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 52 tcaccatgcc ctgggctctg ct	22
	<210> 53 <211> 22 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 53 tcaccatggc cctgactcct ct	22
	<210> 54 <211> 30 <212> DNA <213> 人工序列	
	<220> <223> 合成的	

[0013]	<400> 54	
	cccaagctta ctggatggtg ggaagatgga	30
	<210> 55	
	<211> 16	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 55	
	gtaaaacgac ggccag	16
	<210> 56	
	<211> 17	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 56	
	caggaaacag ctatgac	17
	<210> 57	
	<211> 440	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 57	
	gggcctgggc tctgtctctc ctcaccctcc tcactcaggg cacagggtcc tgggcccagt 60	
	ctgccctgac tcagcctccc tcgcgctcgg ggtctcctgg acagtcagtc accatctcct 120	
	gcactggaac cagcagtgac gttggtggtt ataactatgt ctcttggtac caacagcacc 180	
	caggcaaagc ccccaaactc atgatttatg aggtcagtaa gcggccctca ggggtccctg 240	
	atcgcttctc tggctccaag tctggcaaca cgccctccct gaccgtctct gggctccagg 300	
	ctgaggatga ggctgattat tactgcagct catatgcagg cagcaacaat ttegtcttcg 360	
	gaactgggac caaggtcacc gtctagggg ctgatgctgc accaactgta tccatcttcc 420	
	caccatccag taagcttggg	440
	<210> 58	
	<211> 441	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 58	
	atggcctggg ctctgtctgt cctcaccctc ctcaactcagg gcacagggtc ctgggcccag 60	
	tctgcctga ctacgctcc ctccgcgtcc ggtctcctg gacagtcagt caccatctcc 120	
	tgcactggaa ccagcagtga cgttggtggt tataactatg tctcctggta ccaacagcac 180	
	ccaggcaaag ccccaaact catgatttat gaggtcacta agcgccctc aggggtccct 240	

```

gategettct ctggetccaa gtctggcaac acggcctccc tgaccgtctc tgggctccag 300
gctgaggatg aggctgatta ttactgcage tcatatgcag gcagcaacaa ttatgtcttc 360
ggaactggga ccaaggtcac cgtcctaggg gctgatgctg caccaactgt atccatcttc 420
ccaccateca gtaagcttgg g 441

```

<210> 59
 <211> 441
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

```

<400> 59
atggcctggg ctctgetget cctcaccctc ctcactcagg gcacagggtc ctgggcccag 60
tetgccctga ctccagctcc ctcccgctcc gggctctctg gacagtcagt caccatctcc 120
tgcactggaa ccagcagtga cgttgggtgt tataactatg tctcctggta ccaacagcac 180
ccaggcaaaag ccccaaaact catgatttat gaggtcagta agcggccctc aggggtccct 240
gategettct ctggetccaa gtctggcaac acggcctccc tgaccgtctc tgggctccag 300
gctgaggatg aggctgatta ttactgcage tcatatgcag gcagcaacaa ttatgtcttc 360
ggaactggga ccaaggtcac cgtcctaggg gctgatgctg caccaactgt atccatcttc 420
ccaccateca gtaagcttgg g 441

```

<210> 60
 <211> 438
 <212> DNA
 <213> 人工序列

[0014]

<220>
 <223> 合成的

```

<400> 60
atggcctggg ctctgetctc caccctctc actcagggca cagggtctct ggcccagtct 60
gccctgaetc agcctccctc cgcgtccggg tctcctggac agtcagtcac catctctctg 120
actgaacca gcagtgaact tgggtgttat aactatgtct cctggtaaca acagcaccca 180
ggcaaaagccc ccaaaactcat gatttatgag gtcagtaagc ggccctcagg ggctccctgat 240
cgcttctctg gctccaagtc tggcaacacg gcctccctga ccgtctcttg gctccaggct 300
gaggatgagg ctgattatta ctgcagctca tatgcaggca gcaacaatta tgtcttcgga 360
actgggacca aggtcaccgt cctaggggct gatgtgcac caactgtatc catcttccca 420
ccatccagta agcttggg 438

```

<210> 61
 <211> 438
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

```

<400> 61
atggcctggg ctctgetget cctcaccctc ctcactcagg gcacagggtc ctgggcccag 60
tetgccctga ctccagctcc ctcccgctcc gggctctctg gacagtcagt caccatctcc 120
tgcactggaa ccagcagtga cgttgggtgt tataactatg tctcctggta ccaacagcac 180
ccaggcaaaag ccccaaaact catgatttat gaggtcagta agcggccctc aggggtccct 240
gategettct ctggetccaa gtctggcaac acggcctccc tgaccgtctc tgggctccag 300
gctgaggatg aggctgatta ttactgcage tcatatgcag gcagcaacaa tgtcttcgga 360

```

actgggacca aggtcaccgt cctaggggct gatgctgcac caactgtate catettccca 420
ccatccagta agcttggg 438

<210> 62
<211> 441
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 62
atggcctggg ctctgctect cctcacccte ctcactcagg gcacagggtc ctgggcccag 60
tctgcccga ctcagccctcc ctccgcgtcc gggctctctg gacagtcagt caccatctcc 120
tgcactggaa ccagcagtga cgttgggtgt tataactatg tctcctggta ccaacagcac 180
ccaggcaaaag cccccaaact catgatttat gaggtcagta agcggccctc aggggtccct 240
gatcgcttct ctggctccaa gtctggcaac acggcctccc tgaccgtctc tgggctccag 300
gctgaggatg aggttgatta ttactgcage tcatatgcag gcagcaacaa ttatgtcttc 360
ggaactggga ccaaggtcac cgtcctaggg gctgatgctg caccaactgt atccatcttc 420
ccaccatcca gtaagcttgg g 441

<210> 63
<211> 442
<212> DNA
<213> 人工序列

<220>
<223> 合成的

[0015]

<400> 63
atggcctggg ctctgctget cctcacccte ctcactcagg gcacagggtc ctgggcccag 60
tctgcccga ctcagccctcc ctccgcgtcc gggctctctg gacagtcagt caccatctcc 120
tgcactggaa ccagcagtga cgttgggtgt tataactatg tctcctggta ccaacagcac 180
ccaggcaaaag cccccaaact catgatttat gaggtcagta agcggccctc aggggtccct 240
gatcgcttct ctggctccaa gtctggcaac acggcctccc tgaccgtctc tgggctccag 300
gctgaggatg aggttgatta ttactgcage tcatatgcag gcagcaacaa ttatgtcttc 360
cggaactggg accaaggtca cgtcctaggg ggctgatgct gcaccaactg tatecatctt 420
cccaccatcc agtaagcttg gg 442

<210> 64
<211> 428
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 64
ccttcatttt ctccacaggt ctcigtgtc tgcctgtgct gactcagccc ccgtctgcat 60
ctgcccttgc gggagcctcg atcaagctca cctgcaccct aagcagttag cacagcact 120
acaccatcga atggtatcaa cagagaccag ggaggtcccc ccagtatata atgaaggtta 180
agagtgatgg cagccacagc aagggggacg ggatccccga tcgttctcat ggctccagtt 240
ctggggctga ccgtacctc accttctcca acctccagtc tgacgatgag gctgagtatc 300
actgtggaga gagccacacg attgatggcc aagtcggttg tgtcttcgga actgggacca 360
aggtcaccgt cctaggggct gatgctgcac caactgtatc catcttccca ccatccagta 420
agcttggg 428

<210> 65
 <211> 441
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

<400> 65
 atgacctgct cccctctcct cctcaccctt ctcattcact gcacagggtc ctgggcccag 60
 tctgtgttga cgcagecgcc ctcagtgtct geggcccag gacagaaggt caccatctcc 120
 tgcctctggaa gcagctccaa cattgggaat aattatgtat cctggtacca gcagctccca 180
 ggaacagccc ccaaactcct catTTatgac aataataagc gaccctcagg gattcctgac 240
 cgattctctg gctccaagtc tggcacgtca gccaccctgg gcatcacagg actccagact 300
 ggggacgagg ccgattatta ctgoggaaca tgggatagea gcctgagtgc ttatgtcttc 360
 ggaactggga ccaaggtcac cgtcctaggg gctgatgctg caccaactgt atccatcttc 420
 ccaccatcca gtgagcagtt a 441

<210> 66
 <211> 441
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

[0016] <400> 66
 atgacctgct cccctctcct cctcaccctt ctcattcact gcacagggtc ctgggcccag 60
 tctgtgttga cgcagecgcc ctcagtgtct geggcccag gacagaaggt caccatctcc 120
 tgcctctggaa gcagctccaa cattgggaat aattatgtat cctggtacca gcagctccca 180
 ggaacagccc ccaaactcct catTTatgac aataataage gaccctcagg gattcctgac 240
 cgattctctg gctccaagtc tggcacgtca gccaccctgg gcatcacagg actccagact 300
 ggggacgagg ccgattatta ctgoggaaca tgggatagea gcctgagtgc ggcttttttt 360
 ggaactggga ccaaggtcac cgtcctaggg gctgatgctg caccaactgt atccatcttc 420
 ccaccatcca gtgagcagtt a 441

<210> 67
 <211> 345
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

<400> 67
 cccgggcaga gggtcaccat ctcttgtttt ggaagcagct ccaacategg aagtaataact 60
 gtaaactggg accagcagct eccaggaacg gccccaaac tcctcateta tagtaataat 120
 cagcgccct caggggtccc tgaccgatte tctggctcca agtetggcac ctcagcctcc 180
 ctggccatca gtgggtcca gtctgaggat gaggetgatt attactgtgc agcatgggat 240
 gacagcctga atggttatgt ctteggaact gggaccaagg tcaccgtcct aggggctgat 300
 gctgcaccaa ctgtatccat ctteccacca tccagtgagc agtta 345

<210> 68
 <211> 432
 <212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 68

```
atggcctgga cccctctcct gctccccctc ctcactttct gcacagtctc tgaggcctcc 60
tatgagctga cacagccacc ctcggtgtea gtgtccccag gacaaacggc caggatcacc 120
tgctctggag atgcattgcc aaaaaaatat gcttattggg accagcagaa gtcaggccag 180
gcccctgtgc tggteatcta tgaggacage aaacgaccct ccgggatccc tgagagattc 240
cttggtctca gctcagggac aatggccacc ttgaactatca gtggggccca ggtggaggat 300
gaagctgaet actaetgtta ctcaacagac tacagtggta atcatgtctt cggaactggg 360
accaaggtea ccgtctagg ggctgatgct gcaccaactg tatccatctt cccaccatcc 420
agtgagcagt ta 432
```

<210> 69

<211> 426

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 69

```
atggcctgga ctcctctctt tetgttcttc ctcacttget gcccagggtc caattcccag 60
gctgtggtga ctccaggagc ctcactgaet gtgtccccag gagggacagt cactctcacc 120
tgttgctcca gcaactggagc tgteaccagt ggteattatc cctactgggt ccagcagaag 180
cctggccaag ccccccaggac actgatttat gatacaagca acaaacactc ctggacacct 240
gcccggttct caggctccct ccttgggggc aaagetgccc tgacctttc ggggtgcgcag 300
cctgaggatg aggcctgagta ttactgcttg ctctcctata gtggtgctta tgtcttcgga 360
actgggacca aggtcaccgt cctagggget gatgtgcac caactgtatc catcttccca 420
ccatcc 426
```

[0017]

<210> 70

<211> 331

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 70

```
agtggctctg ggacagacgg ccaggattac ctgtggggga aacaacattg gaagtaaaaa 60
tgtgcaetgg taccagcaga agccaggcca ggcctctgtg ctggctcatc atagggataa 120
caaccggccc tctgggatcc ctgagcgatt ctctggtctc aactcgggga acacggccac 180
cctgaccatc agcagagccc aagccgggga tgaggetgac tattactgtc aggtgtggga 240
cagcagcaet tatgtcttcg gaactgggac caagtcacc gtcttagggg ctgatgtgc 300
accaactgta tccatcttcc caccatccag t 331
```

<210> 71

<211> 417

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 71

```
actcctctcc tcctcctgtt cctctctcac tgcacagggt cctctctgca ggctgtgctg 60
actcagccgt ctctccctct tgcctctcct ggagcctcag ccagtctcac ctgcaccttg 120
cgcagtggca tcaatgttgg tacttacagg atatactggg accagcagaa gccagggagt 180
cctccccagt atctcctgag gtacaaatca gactcagata agcagcaggg ctctggagtc 240
cccagccgt tctctggatc caaagatgct tcggccaatg cagggatttt actcatctct 300
gggctccagt ctgaggatga ggctgactat tactgtatga ttiggcacag cagcgcttat 360
gtcttcggaa ctgggaccaa ggteaccgtc ctaggggctg atgtgcacc aactgta 417
```

<210> 72

<211> 393

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 72

```
tttctgttcc tcctcacttg ctgcccaggg tccaattctc agactgtggg gactcaggag 60
ccctcactga ctgtgtcccc aggagggaca gtcaactctc cctgtgcttc cagcacttga 120
gcagtcacca gtggttacta tccaaactgg ttccagcaga aacctggaca agcaccagg 180
gcactgattt atagtacaag caacaaacgc tcttgacccc ctgcccgggt ctccaggctcc 240
ctccttgggg gcaaagctgc cctgacactg tcagggtgtg agcctgagga cgaggctgag 300
tattactgcc tgcctacta tgggtgggtg tatgtcttcg gaactgggac caaggtcacc 360
gtcctagggg ctgatgtgc accaactgta tcc 393
```

[0018]

<210> 73

<211> 417

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 73

```
atggcctggg ctctgtctgt cctcactctc ctcactcagg acacagggtc ctgggcccag 60
cttgcccctga ctccagctgc ctccgtgtct ggggtctctg gacagtcgat caccatctcc 120
tgacttgga ccagcagtga tgttgggagt tataaccttg tctcctggta ccaacagcac 180
ccaggcaaag ccccaaact catgatttat gagggcagta agcggccctc aggggtttct 240
aatcgtttct ctggtccaa gtctggcaac acggcctccc tgacaatctc tgggctccag 300
gttgaggacg aggttgatta ttactgtctg tcatatgeag gtagtagcac ttatgtcttc 360
ggaactggga ccaaggctac cgtcctaggg gtgatgtctg caccaactgt atccate 417
```

<210> 74

<211> 348

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 74

```
cagtctgccc tgactcagcc tgcctccgtg tctgggtctc ctggacagtc gataccate 60
tcttgcactg gaaccagcag tgacgttggg gggtataact atgtctcctg gtaccaacag 120
caccaggga aagccccaa acicatgatt tatgagggtc gtaatcgccc ctccaggggtt 180
```

	tetaategct tctctggctc caagtctggc aacaeggeet ccctgaccat ctctgggctc	240
	caggctgagg acgaggctga ttattactgc agctcatata caagcagcag cacttatgtc	300
	ttcggaactg ggaccaaggt caccggcctg ggggctgatg ctgcacca	348
	<210> 75	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 75	
	aacaaccgag ctccaggtgt	20
	<210> 76	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 76	
	agggcagcct tgtctccaa	19
[0019]	<210> 77	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 77	
	cctgccagat tctcagctc	20
	<210> 78	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 78	
	catcacaggg gcacagactg	20
	<210> 79	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	

	<400> 79 gatttgctga gggcagggt	19
	<210> 80 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 80 ccccaagtct gatccttcct t	21
	<210> 81 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 81 gctgaccaac gatcgccctaa	20
[0020]	<210> 82 <211> 19 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 82 taagcgccac actgcacct	19
	<210> 83 <211> 24 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 83 cctgccagat tctcaggctc cctg	24
	<210> 84 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 84	

	ctgattggag acaaggctgc cct	23
	<210> 85	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 85	
	ccttcatact cttgcatect ceettctcca	30
	<210> 86	
	<211> 35	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 86	
	ttcettetct tctgtgactc aattatttgt ggaca	35
	<210> 87	
	<211> 159	
	<212> DNA	
[0021]	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 87	
	tctggcacct cagcctccct ggccatcaact gggctccagg ctgaggatga ggctgattat 60	
	tactgccagt cctatgacag cagcctgagt ggttctgtgt tcggaggagg cacccggtg 120	
	accgcccteg gggtgatgc tgcaccaact gtatccatc 159	
	<210> 88	
	<211> 159	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 88	
	tctggcacct cagcctccct ggccatcaact gggctccagg ctgaggatga ggctgattat 60	
	tactgccagt cctatgacag cagcctgagt ggttatgtct tcggaactgg gaccaaggte 120	
	accgtectag gggtgatgc tgcaccaact gtatccatc 159	
	<210> 89	
	<211> 159	
	<212> DNA	
	<213> 人工序列	
	<220>	

<223> 合成的

<400> 89

```
tctggcacct cagcctccct ggccatcagt gggtccagt ctgaggatga ggetgattat 60
tactgtgcag catgggatga cagcctgaat ggtgctgtgt teggaggagg caccagctg 120
accgcctcg gggtgatgc tgcaccaact gtatccatc 159
```

<210> 90

<211> 159

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 90

```
tctggcacct cagcctccct ggccatcagt gggtccagt cegaggatga ggetgattat 60
tactgtgcag catgggatga cagcctgagt ggtcggtgt teggcggagg gaccaagctg 120
accgtctag gggtgatgc tgcaccaact gtatccatc 159
```

<210> 91

<211> 153

<212> DNA

<213> 人工序列

<220>

<223> 合成的

[0022]

<400> 91

```
teggggaaca eggccacct gaccatcagc agagcccaag cgggggatga ggetgactat 60
tactgtcagg tgtgggacag cagcactget gtgttcggag gaggcacca getgaccgce 120
ctcggggetg atgetgcacc aactgtatcc atc 153
```

<210> 92

<211> 156

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 92

```
tcagggacaa tggccacctt gactatcagt ggggcccagg tggaggatga agctgactac 60
tactgttaet caacagacag cagtggtaat getgtgttcg gaggaggcac ccagctgacc 120
gacctcgggg ctgatgetgc accaactgta tccatc 156
```

<210> 93

<211> 159

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 93

```
tcagggacaa tggccacctt gactatcagt ggggcccagg tggaggatga agctgactac 60
```

tactgttact caacagacag cagtggtaat catagggtgt tcggcggagg gaccaagctg 120
accgtcctag gggetgatgc tgcaccaact gtatccatc 159

<210> 94
<211> 159
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 94
tetggcaect cagcctccct ggccatcaet gggetccagg ctgaggatga ggetgattat 60
tactgccagt cctatgacag cagcctgagt ggttatgtct tcggaactgg gaccaaggte 120
accgtcctag gggetgatgc tgcaccaact gtatccatc 159

<210> 95
<211> 159
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 95
gatgcttcgg ccaatgcagg gattttactc atctctgggc tccagtctga ggatgagget 60
gactattact gtatgatttg gcacagcage getgtggtat tcggcggagg gaccaagctg 120
accgtcctag gggetgatgc tgcaccaact gtatccatc 159

[0023]

<210> 96
<211> 153
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 96
cttgggggca aagctgccct gacactgtca ggtgtgcage ctgaggacga ggetgagtat 60
tactgcctgc tctactatgg tgggtctcgg gtgttcggcg gagggaccaa getgaccgte 120
ctaggggetg atgetgcace aactgtatcc atc 153

<210> 97
<211> 153
<212> DNA
<213> 人工序列

<220>
<223> 合成的

<400> 97
cttgggggca aagctgccct gacccttctg ggtgcgcage ctgaggatga ggetgagtat 60
tactgettgc tctectatag tgggtctcga gtattcggcg gagggaccaa getgaccgte 120
ctaggggetg atgetgcace aactgtatcc atc 153

<210> 98

<211> 165
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

<400> 98
 tcaggcctga atcggtacct gaccatcaag aacatecagg aagaggatga gagtgactac 60
 cactgtgggg cagaccatgg cagtgggagc aacttcgtgt ctgtgttcgg aggaggcacc 120
 cagctgaccg cctcgggggc tgatgctgea ccaactgtat ccatc 165

<210> 99
 <211> 164
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

<400> 99
 tctggcacgt cagccaccct gggcatcacc ggactccaga ctggggacga ggccgattat 60
 tactgcggaa catgggatag cagcctgagt gctggccccg ggtgttcggc ggagggacca 120
 agctgaccgt cctaggggct gatgctgcac caactgtatc catc 164

<210> 100
 <211> 22800
 <212> DNA
 <213> 人工序列

[0024]

<220>
 <223> 合成的

<400> 100
 aagctctaaa actacaaact gctgaaagat ctaatgacta ggacagccta gtaattttca 60
 taggggcata aatgtgaaac gccttgtgca tcgtagaaga aagcagaaga gaaagcattc 120
 ccaattttctt aactgccttt tacctatatt aatcagtaat atactggett ttacctctgt 180
 taatcataat aaacaaattc tcaataaatt ttatcgatac tcttcaatgc ctgctcagca 240
 acatttttcg aaggcagctc aagatattaa ataactcata agggccaacc tcctattgca 300
 gcattctttg ggatttaacc agtttcccaa gactcttttc acaatgttaa gatgttagaa 360
 atagatecaa aactagggtga tatatccctt agtaaaaactg tgagggtcaaa cttgtctggc 420
 taatgcttcc atttaaaaaat ttctctttct tgatccttca ttgtatgtac acaataaate 480
 aggggaaaaa tttaactgag tgaatcaaag tattctcatt attataatag gagcttcaca 540
 cacacacaaa aaaatcaatt ctattactct cagcctcagt tcctaaagcc aagttaaagt 600
 cctgtttetaa gatcattgtt gcatgaccat atgtattcca ggtctaactt aaactgtgga 660
 taaateccag caggacatta gagatttttg tgagagtaag catataggat tcagggttta 720
 tgagcttttag atttttcttg tcaaaatgaa tgagagttgc catatctaaa aattattccc 780
 agataaataa aattcactac ctagaattaa tttatgcata taagtagaaa tgctatctec 840
 ctttttacca tccaaagtgg aaagcctcat ggaactagaa attaatatta gaaaaatcag 900
 ttaataaaaag tatgtcattt catcaattea ataagttata atagcaaaaa accataataa 960
 attatcactt aaatgtcaat acatttataa actatggtac ataaatagga tattgaatag 1020
 ccattgatgc tcctgatgaa aattagcagg cagtataaaa tgataaatat gaagcacatg 1080
 tcaataaata aaataagttt tatgtaattt aggagaaaat ggtgataatg acacaaaatg 1140
 tgaattatgg atgcatctat aaaattcttt gtacatttgt gaattgtaaa tatttatctt 1200
 agagacatta ttactttgta tatgttccat ttgetcacct atatgtccca gtctccttac 1260
 aaatgctatg gccaaagaaa taggcataca tacatccttt gcaggctgag gcaggaaaaa 1320

[0025]

gatcttaacgg	aattttccag	tetatccctt	atctgtataa	gcaacttaag	aggccatgtg	1380
ctccaaatgg	tgcaaatata	agatggtaga	gcctctgtct	gcctggatcc	ttgagtggct	1440
gcatggagca	gagcaccttt	ctggccctgg	tgaagattgt	agcatgagca	agatataagc	1500
atttggttga	gctaggccat	gagatttggg	gcagtgggtat	aacctaccct	attatggaaa	1560
atataaatac	acaaaacaga	aaagagagag	agaagtgaga	gaagactgtg	agagaagtgc	1620
atgagagaag	actgtgtttt	gttcatttcc	tataatccca	tatcaccatg	ggatcctgtg	1680
ccctctgggtg	atcaaaactaa	tgttctacag	ctccaaagaa	gaatgctcgc	ctaactgtct	1740
cattccaatg	acctagagac	taaaagccaa	aaagaacctt	agaaattatc	tattgcattc	1800
tttgatgtaa	ggaaatatct	tagagggcac	agatagaaat	atcttaaccc	aggctactta	1860
gttcgtggca	gagctgagge	taaaaccagg	ccctttgaact	cctaattttg	tgtcttttac	1920
accttctcac	atcacttctc	caacccaaag	tctagcagaa	aaggctaaaa	taagatatat	1980
gcatagattt	gctattataa	gtccatgtac	tccctcagac	gctttaagat	ggggcttctc	2040
atggttcaca	ataagcagca	gagggagtg	aataactatc	ttcgtctccc	ctactgctat	2100
ttgtgcagtt	tgaagcttat	ctcttaaate	atgttttctt	ctcgtagtaa	atactacaac	2160
ttgtgccctt	tatgtgtgta	taaattttta	tataattttt	ttccatgaac	cattcaagta	2220
aaatggacac	tccaaaaaga	tgttcaataa	ggttacatgg	cttcacattg	ccccctctac	2280
accatcttgt	ggagctacac	attcacctca	cccaaatttg	agaaaaataa	tcaagaaaat	2340
gactctcact	agcagtgaga	ccaagtccat	aagcactaat	gteatcagtg	cacactgcag	2400
ccctcatgetg	ccaagcatgt	tttgggcgta	tccctggact	ggtttgggtg	catgatcaaa	2460
ggtacatttt	ccacctgcac	agcccccacc	tggatctata	gccttccttg	tgtctttgtg	2520
aacaacctag	tgtgaactca	aagtatgaga	cagatctcaa	ttaatttaga	aagtttattt	2580
tcccaagatt	aaggacaagc	ccatgataaa	gcctccagag	gtcctgatat	atgtgcccaa	2640
gggggtcggg	gcacagcttg	gtgttatata	ttttaggagg	acaagaaaca	tcaatcgata	2700
tgtagaagat	gtgcategct	ttgggtctga	aagggtgtgac	aactcaaggc	agggaagggg	2760
gcttcctgct	gggggttgcac	tgttttgagt	ctctgatcag	cccttcacat	gtgaaaggca	2820
ggtagagaaa	tagtcaattt	tgccttagtc	tggcttattg	aaacagtagg	gcagaagaag	2880
cattgcatat	gcatttgtct	gaagtgaaca	gagggatgac	tttgagctct	gtcctttctt	2940
tgtccacaag	gaattacctt	gtgggcaaat	tgtgagggag	gtatgtagct	tttttttctt	3000
tgtagctatc	ttatttagga	ataaaatggg	aggcaggttt	gcctgatgca	attcccagct	3060
tgaatttccc	ttttggctta	gtgatttttg	gggtcctgag	gtttattttt	tctttcacat	3120
tagtataact	acttttcttt	ttctaattcc	ttttctactt	gtatgtgtta	cagctgactt	3180
atgttaactg	caaaaagaat	tctgactaat	geaccatctg	actagaaggc	agggttcttc	3240
gatgataacg	aatctctccag	aatctagtaa	acagaattgc	ctgaaaaaga	ggtgggtgtc	3300
ttcttgggga	atttctcatg	gcaatgaatg	gcaactggcc	aaaggattta	tgaccagact	3360
gagctctctt	ttatctatct	tgttactcac	caagccctat	taggggtttg	gtcccacagg	3420
gacactgggt	tctaagttct	agggttaaac	agtcactctc	caggcccacc	acaccatacc	3480
ctcctgacat	ctgggtgaaca	geaataaaat	tgtttcttat	tctgaaaaatc	ctccaatact	3540
tccaccatcc	ccaaaaatgc	agtggaggag	gagagaaaat	gaattgttcc	attagagaac	3600
acaatatcca	ttatattatt	cttggccctt	gagatacctt	acaaaacaaa	tacaaaaaaa	3660
gtcccaattt	aacatctttt	aataatcttt	acaaaacaga	acacatctcc	tttcttgata	3720
atagtcaaga	ggctcagtggt	caactgtggt	gaaaagtgtc	agattctgggt	catgtttcaa	3780
aggtagaaaa	aatagaattt	gttaacatat	tggatgtgag	gcgtgggaga	aacgtgaaat	3840
caagggtggt	gcaagtgttt	aacctgagca	actagagaat	ttggaaggac	attttctgag	3900
atgggggaagg	caggcgggaa	tcagggatta	gagttgaaca	tattagacat	ttgagatgcc	3960
tgtagacct	ctaattggca	atateccttg	gacaggtgga	tgaatatgcg	tgattctgga	4020
gttcgggaaa	tagtccgggt	ggagatgcaa	atttgggaaa	cagggcgagg	ttactagcaa	4080
tgagttaaat	caatgaagge	aggctgggac	ctggcaggta	accecaacaag	tagaggtega	4140
agagatgaga	agaaaacagc	acaggagact	tagaagcagt	ggtcaggagg	aaggagtga	4200
accaagaaaag	tgatgtccca	gagccaacaa	aataaggatt	tcttttctgt	ttacaaatgt	4260
aaaattaaaa	ggtttaataa	aaagaaaatt	tacttttatg	gttgggtgtt	attaagtgggt	4320
ccaacactg	tctctatatt	gtagaatcag	aactctctca	tggcagtaga	aaatttggaa	4380
agttactttt	taaaagggtgt	gtgcactgct	gccttttgc	ggteaagttt	atgcactgca	4440
aattccaagg	acgattgctc	gtcagctttt	ctcctttaaa	atagctcagg	ctgtacaage	4500
tagaaaagaac	ctcgaagat	attccttcca	acatttgcac	ttgacttatg	ggaagtgcag	4560
gttcagecag	aaaagttgtg	tgcaaggccg	tttatgtaag	tttatcagac	ctgattctta	4620
eggctcttcc	cattgttttcg	agcctccctt	ccattcactt	cccgtctata	cgcgaccaag	4680

[0026]

tataggacag	gagtagttat	tetgcacttt	atagcagctc	cactgtcttg	cactctgatg	4740
ttctttaatt	acaagcttta	tgacagtgat	tetcaacctg	ctccactgcc	tecacctagt	4800
ggcagaaaga	agaaaatgtg	tgtaactcgg	gagtcctctg	tctgaaagct	cegggggtate	4860
atttcttcaa	agtccttgagc	ttgtttttgt	ttgtatttat	ttatttat	gttttagaga	4920
caaggctctg	cactgcactc	cagcctggga	gacagagcga	gacaattcag	gatctatcta	4980
gtgaataaag	agatatcagt	aatgactggt	ttatatgtg	gctgtagcgc	attcgaggga	5040
taattcgatt	ctgttctgct	ttcgaatgca	tggtcactg	taacctccaa	ctcccgggct	5100
caagcgatcc	tcctacctca	gcttctccag	tagttgagct	tgatttat	taaagtttca	5160
taaaattttg	gcatttcttt	ccacaatatg	gccatgtgtg	ctttactata	aaatattttc	5220
atcacaaaaat	ttacatcgct	ggaaatcccc	ataagccagt	ttgagaaaca	caaccaaga	5280
aagcagaaca	gactcaaatt	atcccttaaa	tecccttaaa	ccacaaatat	aaaacagtec	5340
gtgactgggc	gtgttggett	acacctgtaa	teccagcact	ttgggaggcc	aaggcgggtg	5400
gattacttga	gctcaggagt	tcaagaccag	cctggccaac	atggtgaaac	cccgtcccta	5460
ttaaaaatac	aaaattattc	aggagtgtgt	gcaggcagtt	gtaatcccag	ctacttggga	5520
ggctgaggca	ggagaatcac	ttgaaccag	gaggtggagg	ttgtagttag	ccaagattgt	5580
gccagtgcac	tcagcctgg	gcaacagagc	gagacttcca	tcctaaaaaa	aaaaaatata	5640
gtaaataaaa	tataaaaaaa	taaagcagtc	cctattgata	tctctttatt	cactaaatca	5700
acctggaatt	gacctgaatt	ctgatttttt	tttcatcatg	gattttttgc	attaattttg	5760
attgtttaaa	tattgcatta	aaatattatt	tatcttgact	actgagttag	egggacctec	5820
ttaaaattta	tgaccaaggc	aatgectcac	tcactcgcct	taccataate	tgggccacat	5880
atcaggggct	ccaatagcaa	gcaacatgac	ttttgaacag	ctaagacttc	tctcttcact	5940
gtgaagacca	gatggggcct	gcaaacagt	taacctctac	atgaaaatgc	acgagattcc	6000
aactacaacc	aggecacaaa	gactctgatg	gtgaagtccc	agccctccaa	gtcccaactt	6060
cctgaaggga	aagagcacc	caagttctga	ccagaggcca	gagtcataac	gaagatggaa	6120
tgtgagcttg	acatagaagg	ggtggtagea	cctggctcag	taatgaagag	gctttcggte	6180
ctgaagggaag	agctcagcac	attcaaagat	tagaaggag	gtcccagtc	taggagcagg	6240
gaaggagaga	aggeccaata	agaaacacag	acaggaggga	ggggtcagg	caagatcata	6300
ctggaaacaa	ctagagagct	aataaaagtc	acagtgccca	gtccccacat	ggaccagact	6360
cttcggaate	tctaggeate	aatttgggea	ccagtagitt	tcaaagttct	ccagaagatt	6420
ctatgcacac	cagccaagg	tgggaaccac	aggtgttgge	ctagggatca	tgacaatgag	6480
tttctaagtg	caataagaaa	cctccagaga	gtttaageag	gggaataatt	tgatttgttt	6540
cttgtttgtg	atttttaaag	atcagtctgg	ttactgtgtg	taagacaata	atccagaaaa	6600
tetgttgete	atgaaccaca	tatctgtaaa	tttgettccc	ctgtaactgg	atctaaccac	6660
caaaaattag	tacttactaa	gaaattacat	gccaggggac	tatgctaagt	aattcataaa	6720
cactatittta	tttactcttc	acagcaagtt	tataagagaa	acgttatatt	ttccacattt	6780
cggatgagaa	atttgaggct	tggggaaagt	taagtaattt	acctaattgc	acaccagttt	6840
cataagatgc	agagttaaga	ttctaattct	gtgtetaagt	tgatgetcca	tcaaacacac	6900
cacgcctcca	actaggaagc	aacatgctgg	ccagaggatg	ctgtcatcaa	gtttacagaa	6960
tggttagatt	tctaggcaca	gatgaataaa	tcaacatgtt	ggtttgcaat	agaatgaatc	7020
tatccagctc	tgaatttgca	tecaagggtt	tgtgagcaca	caagtetaaa	agtgtggcct	7080
cagctctget	aacttcacac	aggtgaatac	ctaggaggcc	accctctgag	accaccagat	7140
ggacagtcca	ccatctgttt	acagatggta	aagccacata	ccagctttgc	catctgatgt	7200
tctetattca	cattcaacat	ttatacaaga	aatagtcata	tggatccttt	tcaatagaca	7260
gtactgggga	aattgaattg	ccatatgcag	aagaatggaa	ctagacctct	atctctcacc	7320
aaatacaaaa	gttaactcaa	gacagattaa	agacttacat	ataagacctg	taactacaaa	7380
aacactagaa	gaaaacctag	ggaaaatgct	tctggaaatta	atctaggtga	agaactcagg	7440
actaagatat	caaaagcaca	agcaccacaa	caaaaataga	caaacaggac	ttaatataac	7500
tagaacgctt	ctgaacagca	agagaaataa	tcaatagagt	gaacagataa	tctgcagaat	7560
gggtgaaaaat	atttgcaaac	tatgcatect	acagggaat	aatgtccaga	atttagaagg	7620
aactcaaaaca	attcaacaac	aacagcaaaa	taacccacc	aaaaaagtgg	gcaagggaca	7680
tgaatagaca	tttttcaaaa	gaaggtatat	gatatgggtt	ggctctgtgt	ctccaccag	7740
atctcacctt	aaattgtaat	aateccacac	tatcatggga	gagaccgggt	gggaggtaat	7800
tgaatcatgg	gggcagggtt	gtcccatget	gttctcatga	tactgaataa	gtcctatgag	7860
atctgatgat	tttataaagg	ggagtteccc	tgcacacact	ctcttgccctg	cctecatgta	7920
atatgtgeet	ttgctttctc	tttgeettct	gccatgattg	tgaggcctct	ccagccatat	7980
ggaactgagt	caattaaacc	actttttctt	tgtaaattac	ccaatcttgg	gtatgtcttt	8040

	attagcagca	taagaacaga	ctaatacagt	gtacaaatgg	ccaagaagcg	tacaaaaaac	8100
	aaaatgctca	aatcactaat	cactagagaa	tcgcaagtta	aaaccacaat	gagatattat	8160
	cttacagcag	tcagaatgcc	tattattaaa	acaccaaaaa	ataacatgtt	ggcaaggatg	8220
	cagagaaaag	ggaatactta	cacattatta	gtgggaatgt	aaactagtac	agcttctgtg	8280
	gaaaacacta	tggagatttc	teaaagaact	agaaatagaa	ctaccatgtg	gttcagcaat	8340
	accacaactg	ggtatetacc	caaagggaaa	taaattatta	tataaaaaag	atatctgcac	8400
	tcacttgttt	attgcagcac	tattcacaat	agcaaagata	tggaatcaac	ccaagtgtcc	8460
	atcaacagat	gattggataa	agaaaacgtg	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	8520
	acacatacca	caatgaaata	ctattcaget	ataaagaaaa	gaatgaaate	atgtcttttg	8580
	cagcaatgtg	gttggaactg	gaggecatta	tettaagtgg	ataattcaaa	aacagaaggt	8640
	caaatgtcac	atgttctcac	ttataagtgg	gagctaaatg	atgtgtacac	atggacatag	8700
	agtgtggtat	gataaacact	ggagattgag	atgggtggaa	gggtggaagg	aggttgagtg	8760
	atgagaaaat	actaaatgga	tacaatatac	atgattcagg	cgatagatac	actaaaagcc	8820
	cagacttcac	cactacacag	tatagctatg	tagcaaaatt	gcacctgtat	tgtttaaatt	8880
	tatacaagta	aaaaaaagat	cgtaacgaatt	ctgtttttta	ttctctatga	aattactact	8940
	gagagtatta	tccaatgccg	tttctatgca	gtgcccccaa	tattatecat	ttagcagetc	9000
	ctatgcaatg	ccccagata	gaaattgtct	tcaactttta	tcccaggaaa	accttcagtc	9060
	acacgtagaa	actagaaatt	tttcccctag	atgaaagtta	tgtaacataa	cacattatct	9120
	tcatttagtc	ggtttccaag	aagctcagaa	ccagatttta	tgttcaatca	aaaactgett	9180
	attttaagtg	aggtttactg	aggataaat	tacaataaaa	gccacctttt	cgtgtatatt	9240
	tctataagtt	ttggcaaatg	catagctgtg	taaccacaa	cacattcaag	atataggaca	9300
	agtcctcat	cctttaaagt	tcctttatgc	cccttcttct	accccagccc	ttggcaacca	9360
	ctggtttttg	tctgatecaa	tcgtttgcct	cttcttgaat	gtcatgtaaa	tagagccatg	9420
	caatgtgaag	ccttttgagt	ctggctttgt	tcacttggtc	acttaggaga	atgcatttga	9480
	gattcatctt	tgtgttttcg	tgtagcacta	gttcaactgtc	tattgttgag	tagtattcca	9540
	ttgtgtggat	atgccacaga	ttgtttatct	agttaacaat	ttaaagccat	ttggtcattt	9600
	ctaattttta	gctgctaaga	ataaagtgtc	tgtaaagctt	ccaatgcagg	tttttgtgtg	9660
	aactcaggat	ttcatttcgc	ttgggtaaat	tcttagcttt	gggactgctg	agtcacttgg	9720
	taggtgtatg	ttgaacttta	taagaaactg	ccaaactgtt	ttccaaagtt	gctgtgctct	9780
	tttgcactcc	catcagcagt	gaatgagggt	tccacttget	cgagcctagt	attttaactt	9840
	cactatatac	cttctttgat	gacatatcct	ttcaaatttt	tggteaagtt	tttattgggg	9900
	tgttgttact	atggaactgtg	agagtctctt	gtatattctg	catatgattt	ttttctcaca	9960
	tttgtgtttt	atgaatatgt	tctcccaatg	tgtggcgcct	tttattttct	taacgtgcca	10020
	tgtgaagagc	agaagtttaa	ttttatgatg	tccaaattat	ctttttttct	tttctttttt	10080
	agatcaaaaat	aggggtctat	tttgattacc	actgttattt	tatctccatt	tgatttttca	10140
	tttttatttt	tatttttcta	atttcattgt	aaatttttaa	ttaaacccaa	atattctagg	10200
	ggaaagaggc	aagataaaaa	tagtctaact	tgggcataaa	ttttagagtc	atattctett	10260
	gccagaaaag	gaaactagct	ctettacatt	gattgtttta	tttcagacgt	cactacttta	10320
	tgaggatgcc	caaattatgg	gctttaaaaa	atatatatec	aaacaggggt	tcagaaagaa	10380
	taactaattt	gtccacaaca	acacaaaaaa	tgattecacc	ataagtttgc	ccagtgcacg	10440
	ggtctatatt	attttctata	tatcaaattc	tacaactggg	tcttaaagct	actgtacata	10500
	acctaaagta	aaatattagg	tattagtgtg	taagacattt	tatcatctat	gaaatgttgc	10560
	ctgttgtcat	agtttagagaa	tcttttaaaa	tatggagcta	ttttcataga	ttaaactatg	10620
	ccagttaaaa	gttgggtaaa	aagaactaca	gaataatatt	tatgtttatc	gtgtaaggtt	10680
	ttaaagcaaa	ctccaagtca	tttteatcaa	tgaatcaat	aaggttttgc	aaatatatat	10740
	gtatgaaaat	actgatttaa	aatgcaaata	aggggagagt	ttgagagaga	gagagagacc	10800
	aaatgatttt	ataattctag	taagtttata	ggtttatggg	gtttttacgt	acttttctac	10860
	ccaacttgtc	tataagactt	taatgaatca	cttagaattt	ttaaaataat	ttattattac	10920
	tctgtacctg	ttctttactc	tgeaatctt	accttgcctt	tttgtctaaa	agcaataaaa	10980
	tctgacctgg	tttatatcgt	atcattgatt	ttgttactta	gcaagcacag	tgatccatta	11040
	ggcctatgta	ggctcatggt	ttatacaaca	ctgccatctg	ctgacagagt	gtgacagtca	11100
	cagtcagcaa	cacgagacca	ctttattttc	atttttagtg	tttatagaaa	tatgaatata	11160
	cacaaatagt	ataatgaacc	ctaagcttca	caaattaaca	ttttgcta	cttgtttcaa	11220
	ctacegectc	ccccctcatc	caattactct	gttctctcac	ctcctcacac	acagacactg	11280
	gcagtatttt	tcagccaatc	attaatacgt	tgecaactga	taaggacttt	taaaaaacaa	11340
	ccaccattcc	atttatgattc	ccagcataat	tgagagtaat	tcctaatat	ccaataccca	11400

[0027]

	ttttctat	caatttce	gattgtct	aaactgtt	tacccta	ttgettaa	11460
	caaagtcc	gtcctgtt	acatatgt	aagtttt	caaacc	taaataaa	11520
	aaataaaa	taaataac	atttttt	attccagg	atagtga	agggttaa	11580
	ccattatt	gaaacata	tcacatca	ggactaga	tatcttga	tcaaaatt	11640
	agactgaa	tggaaaag	aggtatag	taaactta	taaaaact	aatgcaga	11700
	tctaagag	gatattag	agttgtac	gcattcata	ttcagtatt	atcagtatt	11760
	actcagct	atgtagtt	aatetaac	gaggagct	atcagata	gagatacat	11820
	tttctacca	aggcggac	tggaggcag	tggttcag	ctagacag	getgcagg	11880
	ccaggte	tccctgct	ctctcca	ctagcttg	actttcat	tgcagatg	11940
	gtgtttct	caagttcc	atagaaga	atagaaca	aaggagaa	aagcagtg	12000
	gcctctgt	atcaagca	atttttcc	aaatgcac	tagatttc	atgatgtc	12060
	aacagtc	actgcaa	agctgagg	ttagattt	ggctggga	ctgttgc	12120
	gtaaaaa	tgggattc	ttattaa	ataaggag	ggaagaa	ttgaaaac	12180
	ctatgca	gtgaaaa	taagaaac	aataaaaa	tgggcata	ttaaaaac	12240
	gcaattca	acagatga	ccccaat	caataaac	ttttaaat	tcaacctc	12300
	gagtgat	aaaacaca	tatgtatt	aaacaaaa	taaaatac	tgtattga	12360
	atttgagt	aaaaaaat	aaaagc	taatatca	tattggag	gatgtaga	12420
	gaggaaac	catggagg	ctatcatt	aaatgtgg	atgaaact	atacagaa	12480
	ttgagge	tttgtaaa	gaaaaat	cacacc	aaccaagt	cccttgca	12540
	atttttg	agacaaaa	gttatgta	tggatcat	caatatgt	cccttata	12600
	cagcata	cccctcag	ccattga	catgtgt	aacagctc	tatttttt	12660
	ttaat	ttagagac	agtctcac	tgctcac	gggtggag	agtggcg	12720
	tcataac	ctctagc	ctcgaa	tgggtcaa	catctc	ctcagc	12780
	ccaagtag	aggactac	gcattgg	caacacac	ctaatt	taaat	12840
	ttagagac	ggtctcac	tgttgcc	gctgggt	aactccta	tcaagc	12900
	ctcccac	tacttcac	agtctgt	gtatgtag	atggattg	ggtatga	12960
	accgtgcc	actcact	ttttatt	aatattcc	tgggatgg	gtaccgc	13020
	ttgtttac	attaatct	tgtaggac	ttgactgt	tccagttt	ttttaata	13080
	aataaaa	ctatgaat	ttgtgtat	tatacgtt	tgtgcta	tttcattt	13140
	ctgggata	ttttcatt	tttgggt	tactgtac	ttgatatt	aatatgtt	13200
	atcttcag	ttattct	caatatata	tctttatt	tccttgaa	ctcccatg	13260
	ttgtttag	gtgtgtgt	ttgtttcc	gggtttgg	ttttccca	tatttttt	13320
	ttatcgat	ccagtttg	tccagtg	cagagaac	acttcagt	atttcag	13380
	tattaaat	gttgaggt	gttacatg	ccagtata	gcaatttg	tatatgtt	13440
	atgagcact	gaaaaga	cgaattct	tgggtctg	tggagttt	cagcaatg	13500
	gatttatg	cttactca	gatggtg	ttgagttg	tgtgttct	cgatggc	13560
	tttaacat	ttgtcagg	attcta	ctctgtc	tcagtatt	cgcctctt	13620
	ctgtctca	aaagctga	tttctggt	tccctgg	ctgttggg	tgtgtggt	13680
	tcatttg	tctggact	ggagtatt	gttatgag	tttgatct	atttaaac	13740
	atctcagc	atttctct	ttgccact	ggaaggag	gttgggtg	tgaatgg	13800
	agagccgt	ctgccta	attgtttt	tgggctcc	ctttcttt	cccttgac	13860
	gagagagc	gtttttat	agggtttt	gtttttct	gcctgttg	gtttctgg	13920
	tgacaaa	ctccaga	aagtctgg	tggatgag	aaaaaga	cccggtga	13980
	gcactgtg	gtcgtct	gggtccca	gttccta	ggtctg	cttctctc	14040
	gcttccag	tcttcata	tttgettt	gtacaatg	cgggggtt	actttact	14100
	agagaaat	gtaaaagt	ttctactc	tctttcag	agcaaa	cccttgtg	14160
	ttttttta	ctttcaaaa	caaaacaaa	ggcagctg	acagtaa	agctagta	14220
	acccttgg	ggaaattc	gtccaaat	acatttta	tttgctag	cagtga	14280
	atcagaat	ttcaggtt	aaacaaat	atatttat	ttatgc	actaaa	14340
	gaaggcat	tatat	aagcact	ttgttct	ttaaaa	agaatt	14400
	tccctagg	atttgac	gcagctgg	ctgatcc	tggccact	agcacag	14460
	agggagag	gtagccct	cccagcc	cctcaaga	ggattctt	tctgggaa	14520
	gtaggta	ctaaatcg	ctggaaca	acaacgaa	aagaaagg	agagaa	14580
	agaaagga	aaagagag	agaagga	aagggagg	gggaagga	gaaggg	14640
	ggaaggga	ggaaggga	gaaggaga	aaaggaga	agggagg	agaggga	14700
	aggaagg	ggaaggga	gaaggaga	ggaaggaa	aaagaa	agaagaa	14760

[0028]

	aagaaagaca	agaaagaaag	aaagaaagaa	agaaagggga	aagaaaaagaa	agaggaaaga	14820
	aagagaaaga	aagaaaaagaa	agaaaggaaa	gaaagagaaa	gaaagaaaaaa	gaaagagaaa	14880
	gaaagagaaa	gacaagaaag	aaaaaggaaa	gaaaagaaaag	agaaagaaaaa	gaaagaaagg	14940
	aaagaaagag	aaagaaagaa	aaagaaagaa	agaaaagaag	aaagagaaag	aaagaaagaa	15000
	aaagaaagaa	agaaagaaag	aaagaaaaag	aaaaagaaaag	gagaaaaatga	cagcaattac	15060
	ttttgcaaca	acctaataata	agttttttta	aagttaaata	ttctgtttcca	tgcatgtgtg	15120
	gataccttat	aaataacagg	gcatectatg	acctgaattt	cccaaattat	gagttgaggg	15180
	tttgaactag	ttttaaaaaa	caaggaggcc	aggcgcaactg	gctcatgcct	gtaatcccag	15240
	cactttggga	ggctgaggca	ggtggatcac	gaggtcagga	gctcgagacc	agcettacca	15300
	acatagttaa	acaccgcctc	tactaaaaat	acaaaaatta	gccgggcgtg	atggtgcgca	15360
	cctgtaatct	cagctactca	gcaggctgag	gcaggagaat	cgtttgaacc	cagaaggcgg	15420
	aggttgcaat	gagccaagat	cacagcattg	cactccagcc	tgggcgacag	agggagactc	15480
	cgtcttcaaa	aaaaaaaaaa	aagacaagga	atctgtaaaa	caggcactgg	aagtatatgc	15540
	acttttattt	tcattctatg	ctatccgatg	cctactgcta	tttcccttca	tatttaacct	15600
	ccaacagctg	cattttgtct	cctccagacc	acctgattgg	agctcacgtg	ctcccacaca	15660
	gtacctccaa	ccagagagag	tcgagtccca	cagaaaggcg	taacaatcac	cagtaatttt	15720
	gcacttattt	tacattgtgc	cttgatacag	agtaactcaat	gaatgctctt	tgaatcatat	15780
	ttaataaata	tgtgtatttt	ggattgtage	atattgcage	tacctggata	tataatttta	15840
	ttagaaaaaa	aattttgtgt	ggetcaatec	acaaacgact	tttctctctc	tctctttctc	15900
	tttctccctc	tctctctctt	tcttctcagt	tgatgttget	ggagttcagt	gttgtgcaga	15960
	tggcagtgac	aaatgccatg	ggcacatgag	atatgataaa	aggteccetga	agaagggtga	16020
	gaaccagtta	tcttatgaaa	ttttccagag	tgggtactgg	atctctctctg	tctggcacea	16080
	tgttgccctc	agcccaaggg	gaatttcctt	ccagagacag	agggcagtg	ttgaggtggg	16140
	gagacagatc	gtaacactga	gacttacatg	aggacacca	acagaaaaaa	ggtggcaagt	16200
	atagaaaatt	ctttcttctg	gacagtcttc	tctgttctaa	cttcagcaaa	attctcccc	16260
	cagtggatgc	tattgcacaa	ccctacatat	gctatgtttt	ttctatata	cacttaccta	16320
	tgataaaatg	cattaattag	tcacagtaag	agggttaaca	caataactag	taataaaaa	16380
	gaacaattca	gtaaaataag	agttacttga	gcacaaacac	taggatata	tgacagtcaa	16440
	tctgatgacc	aagagggtga	ctaagcatct	aaacaggagg	gtaagtgtag	acagatgga	16500
	gacgttggac	aaagggtatg	ttcagttcca	ggctgggtatg	gagcgggaagg	cgatgatatg	16560
	tcattcaagct	actaaggcac	acaattttaa	atgagtaaat	tcttatttct	agaaatttct	16620
	ttttaatatt	ttcagactac	agttgcctac	aggtaactga	aaccccagaa	agcaaaattg	16680
	ttgataagga	ggtactactg	tacatgctcc	tttgaaccaa	ctttatcatt	tgctagtata	16740
	tacatatata	cctacatata	tacatatata	catacctgca	cacacctata	tgtatacgta	16800
	cacacacaca	cacgcacaca	cacacactca	catctactaa	tgttagaata	agtttgctaa	16860
	ataagatgca	caacttggtt	atgtcttaca	gagcaataaa	accataagca	ttgggggttat	16920
	cttttctact	agataaaaa	ccattatcat	tttcataaag	ttttctttac	attaacatct	16980
	aaacttttgc	atctagtgtt	taatcatcat	aaataggaag	caaatgaact	gtttctctag	17040
	tgaatcaaat	atccttgaaa	acatacatag	tcactttttt	ggttttattt	tattttttaga	17100
	taaattattt	aaagttttaa	ataatttaac	attcacataa	gtttgtgact	gtatattttg	17160
	acttggctct	tcaaacttaa	tttgtacttt	tatgtatcgt	gcttacctca	attttttatt	17220
	cacttttctc	aaactttgct	ggattggttt	attatttttg	tctatttctt	ttccttctag	17280
	tggtttggga	gggtttttta	aatcccatat	ctattgaatg	cctattaact	tgeccccctt	17340
	ttctttcaat	ctctatttcc	aeggcctgaa	gcattgagggc	caagctgtct	gtaaccagca	17400
	gagagatgac	ccaggtgtta	ttccactctc	cactgtccac	ctatcaccat	ttccagcccc	17460
	atagctctga	agtacggctt	ttctggggct	ctgtggggga	aactagaact	ggctgcttca	17520
	aggacacctc	ctgtttttgc	aatggaaaaa	atgtttctaa	attccagttt	ctctatgaat	17580
	tcaatgacat	ggttttaaat	tctgtggtgt	tcttcaaagt	tttttcttct	aataggacct	17640
	ctcatgattc	tccaaccacg	aaataaatte	attatcattt	ttatatctct	tctgtcattg	17700
	caaaggaggt	tttgaaagag	tggaggacgc	gtaaatgaac	tcaaaaatcc	acactattec	17760
	ttgtttccat	ctgtttgttc	ttcattgttt	ccattggcct	gtccgcctcc	tactctctct	17820
	cttagacttg	gagctctagc	ctcagccagg	atagggaaaa	gagagatcag	actgttaact	17880
	tgtctatgta	gaaaaggaag	acataagaaa	ctccattttg	atctgtatcc	tgaacaattg	17940
	ttttgccttg	agatgctgtt	aatctgtaac	tttagcccca	accttgtgct	cacagaaaca	18000
	tgtgttgtat	ggaatcaaga	tttaagggat	ctagggtgtg	gcagaatgtg	ccttgttaac	18060
	aacatgttta	caggcagtat	gcttggtaaa	agtcategc	attctccatt	ctegattaac	18120

[0029]

[0030]

taggggcaca	gtgcactgcg	gaaagccgca	gggacctctg	cccaggaaaa	ctgggtattg	18180
tccaagggtt	ctccccactg	agacagcctg	agatatggcc	ttgcgggatg	ggaaaagatct	18240
gaccgtccce	cagcctgaca	cccgtgaagg	gtctgegetg	aggaggatta	gtaaaagagg	18300
aaggectett	gcggttgaga	taagagggaag	ccctctgtct	cctgcatgcc	cctgggaacg	18360
gcatgtctea	gtgtaaaacc	tgattgtaca	ttcgtttctat	tctgagatag	gagaaaaccg	18420
ctctgtggct	ggaggcgaga	tatgctggcg	gcaatgetgc	tctgttggtc	tttactacac	18480
tgagatgttt	gggtgagaga	agcataaate	tggcctacgt	gcacatecag	gcatagtacc	18540
ttcccttgaa	tttacttggt	acacagattc	ctttgtctac	atgttttctt	gctgaccttc	18600
tccccactat	caccctgttc	tcctgccgca	ttcccttgcg	tgaggtagtg	aaaatagtaa	18660
tcaataaata	ctgagggaac	tcagagaccg	gtgccagegc	gggtccctccg	tatgctgagt	18720
gacggtccct	tgggcccact	gttccttctc	tatactttgt	ctctgtgtct	tatttctttt	18780
ctcagtctct	cgtcccacct	gacgagaaat	accacaggtt	gtggaggggc	tggacacccc	18840
ttcgagccag	gattatcagg	gcatttgggg	gtctgcaaaa	ctaagcccca	actcatcgat	18900
ttcacaactt	catccagagc	cagcctgaac	agtagttgcc	catgatttct	atgccttaat	18960
acgagaagag	aacatagggg	ctgggtgcc	agtaggtaga	cagggagggc	agggaactct	19020
aagacagage	ttgaggggct	cattcctctt	gcaaaatgaa	acaaaaacca	cagcactgaa	19080
tatgtaaate	tcggtggctg	aaccctctct	aggatagtaa	gccctgacac	aattgctget	19140
atcttctctt	tctctcaagg	aagtcaaaaa	acacctgcag	ccttactgtc	cccttggaaa	19200
caagatgaac	atctacattt	tctaaagtgg	gacaagaate	tctgttcata	tttatgtccc	19260
atgcatttgc	acgtggccgg	acaaaggact	ttgcttctgc	cagcacatct	gtcttcagat	19320
atgagaggaa	acagacacaa	cctggaggcg	gcaaagaage	agctctttct	caagtgcctt	19380
cctctatctc	cctaattcct	ggctaattgg	gcagccttga	tccttgggaa	tccaggacag	19440
atateccactc	gtgacaaact	agctggaaga	atgacaacca	atcaggttcc	aagcaccact	19500
ggatgtgaac	cacagaattt	cctcctctcc	ttgtggaatg	tcagettacg	tctgacaaaa	19560
aatgtaaaac	tgagagagtt	acaatcttaa	ggaggagtca	agctaaagca	gaaagaatca	19620
cctactctgg	actccagcat	gactgctgag	ctcaaatata	tatagagaga	gaaagaacca	19680
caaacttgaa	gatggatata	agctacagac	tttcttgagt	caggtaggga	aatggccatc	19740
cetcaaaect	tgcaaaaggc	aaacttatgc	cattgtgtcc	tctgacatac	tgggtgatgt	19800
actgtatggt	actgatgtga	ggggaaactc	ctaaaattgc	tagtaaatga	tgccaaataa	19860
aaagcaaaaa	tgataattct	tgaaatgtta	catctgagga	acattgctaa	aataatttat	19920
cagtagtttt	caggatgatt	tatagatgtg	cattgaagtg	tgtacttgtg	ctctctctct	19980
cctctctctc	tctctttctc	tcctctctct	cgtcttttct	ctccttgcce	ccctccctcc	20040
ctgactttcc	ttcctgtccc	ctccacagca	gtttatatatt	ttttcttgat	aatctaaact	20100
tgtctagggt	tcaatgtaaa	gcaccttcag	tgatgagtta	gttggaatgt	tccccaagaa	20160
attctatttc	cagcactctt	ttacatgaaa	tccaagaagc	tctcagacta	tcttactgac	20220
accttgcctt	tcctcaacag	atcaatctta	tcaatgtcca	tcacagatat	tttgtagaac	20280
ggtggatcct	ggcagagtct	cacagatget	tctgagacaa	catttgcctt	caaaaaatga	20340
accacacaca	tcctaaagat	ctcagccact	tcccatgttt	cattttgtgt	tacagcaaac	20400
atcacacaaa	tcattcctac	agatcaccac	tgcattgigt	caataaaata	gtttttgcaa	20460
caatggtaet	tatgataatc	atcttttatt	gtttacaaat	actgctttac	aatagttatt	20520
cggttgcact	gttcatatta	gatttccaat	tagctcactt	aggaacataa	gtccctcgaa	20580
cagctcagte	atctttttca	ttcctgttte	tatccctctc	atctctttcc	tttgagacag	20640
actatctect	acaetgaaac	aggaaagctt	ttaccttttt	ggcatgcttg	atttaaagat	20700
tatagaaaag	tatttgacaa	agaaaaactca	cacatgtgtg	tacatatctt	ttaaaaagtt	20760
atgtttatgc	attgcacagg	aatatcgaga	atgctaatag	gcaatgtcag	agtttactgt	20820
ttttcaaaat	tagtacagtt	ttattatttc	taaaaactat	aaaatgaata	tattcacatc	20880
accatacaga	agagtaggag	gagatggcat	aaagtgtcat	tgttctctct	ctgeaatccc	20940
aggagataac	taccaagcac	aatttatgtc	ttttaaaatt	cagcccgat	ttatatacat	21000
atataattcaa	tgtagatggg	atcatgatat	ctcaccacac	atactcttca	gtgacctgea	21060
ttttcacaaa	caccttccac	gtaactatat	agaagtctac	gtcttccctc	taatgtctgc	21120
tttgtctac	attgtaaagc	tctagcacag	tttaacccaa	ctcctattaa	tgaggatttt	21180
agttattttt	cactctttta	acaatatctc	catgtgtagt	cttatacata	cgtctgtaca	21240
cacttatccc	agtctaagga	gttcttttta	ccttccccca	tcccagcatt	ccctgtcacg	21300
cttgttgctt	cgtttgagtg	actttactcc	tggagtataa	tctgctgata	gttcagttaa	21360
aaacatggga	tctgagttta	ggtcacagct	ctgccactta	ctgccataag	ccagttcctt	21420
gacctctctg	ccctcaagtt	tttgcacctc	caaagtaggg	gataatatta	gttcttagtt	21480

catagagtct	tggaataat	taaatgtgat	gatccatgta	caatgtctgg	cacttagtaa	21540
gtgtcaata	aatgtcaccc	tttatgattg	gtattgeggtg	tatgtctgca	gagaaaatca	21600
ctttgtgtcc	cctttaaaaa	aggactatgc	ccttgggtcag	ctattttgca	cattaaattt	21660
cacttgccaa	tattaactct	ccacctctaa	cttgatccct	ctecttcctc	atcttctggt	21720
gagaccaaat	gctaattctg	ctattcaagg	caactageaa	agctgccagt	gacagaatca	21780
aataaaccta	cccctaactt	ttagaattgt	agttatgatt	tctgttgtaa	aagttactgt	21840
tgtggcagtc	agtattagtc	tttgggtctat	gatagcatct	ctgatctatt	attgaytttc	21900
aattakgtat	ttttttttat	ttattctgaa	aatgtttgtt	aagcatttgc	taagtaaaga	21960
tactggackg	agcctcccaa	atacagggca	aataaaacat	caaacagctt	ataatttaga	22020
agggtagaag	agaatctgaa	agcaggtaaa	aataaacagg	cactcggtcg	ggcgcggtgg	22080
ctcacgcctg	taatccagc	actttgggag	gccgaggtgg	gccgatcacg	aggtcaggag	22140
atcgagacca	tcctggctaa	cacggtgaaa	ccccgtctct	actaaaaata	caaaaaatta	22200
gcgaggcggtg	gtggcgggcg	ectttagtcc	cagctagtcg	ggaggctgag	gcaggagaat	22260
ggtgtgaacc	cgggagggcg	agcttgagct	gagccaagat	cgcaccactg	cactccagcc	22320
tgggygacag	agcgagactc	cgtctcaaaa	aaaataaata	aataaaataa	aaaataatta	22380
ggtactctag	gcccagtgac	ctgtctctgt	actctgtaaa	ttcaggtcac	ctgtctcagg	22440
ctaactctgag	agaaggtctc	tcttcagttg	aattttgaaa	gacaattagc	agttcacaag	22500
ctaaccagg	tggacaaaga	tgttcccaag	cagagggagt	gcttgtgaaa	gctggaggcc	22560
atagaaaaac	tctaaggagt	gtagggaggt	gggagtaatg	tatggaaggg	gtggagatgg	22620
aaggttaaga	gagatacaag	gctgcaaaaa	tggagctgga	ctcaaaagaa	aatactgaaa	22680
aggtcttcag	tgttggtgat	gagattaacta	tggaaacact	atggaacact	gggactecat	22740
ggcagcteca	aagatggcat	gcgcctggtc	cagctcagta	agagctgagc	tcttctgtg	22800

<210> 101
 <211> 154
 <212> DNA
 <213> 人工序列

[0031]

<220>
 <223> 合成的

<400> 101	
tctggcaaca	cgccctccct gaccgtctct gggctccagg ctgaggatga ggctgattat 60
tactgcagct	catatgcagg cagcaacaat ttaagtcttc ggaactggga ccaaggteac 120
cgctctaggt	cagcccaagt ccactccac tctc 154

<210> 102
 <211> 156
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

<400> 102	
tcagggacaa	tgccacctt gactatcagt ggggcccagg tggaggatga agctgactac 60
tactgttact	caacagacag cagtggtaat cattatgtct tcggaactgg gaccaaggtc 120
accgtcttag	gtcagcccaa gteactccc actctc 156

<210> 103
 <211> 150
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成的

	<400> 103 tctgggaaca cagccactct gaccatcagc gggaccacagg ctatggatga ggctgactat 60 tactgtcagg cgtgggacag cagcactgcc gtcttcggaa ctgggaccaa ggteaccgtc 120 ctaggteage ccaagtcac tccactctc 150	
	<210> 104 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
[0032]	<400> 104 aggtggaaac acggtgagag t 21	
	<210> 105 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成的	
	<400> 105 ccactcgggg aaaagttgga a 21	

人入轻链基因座

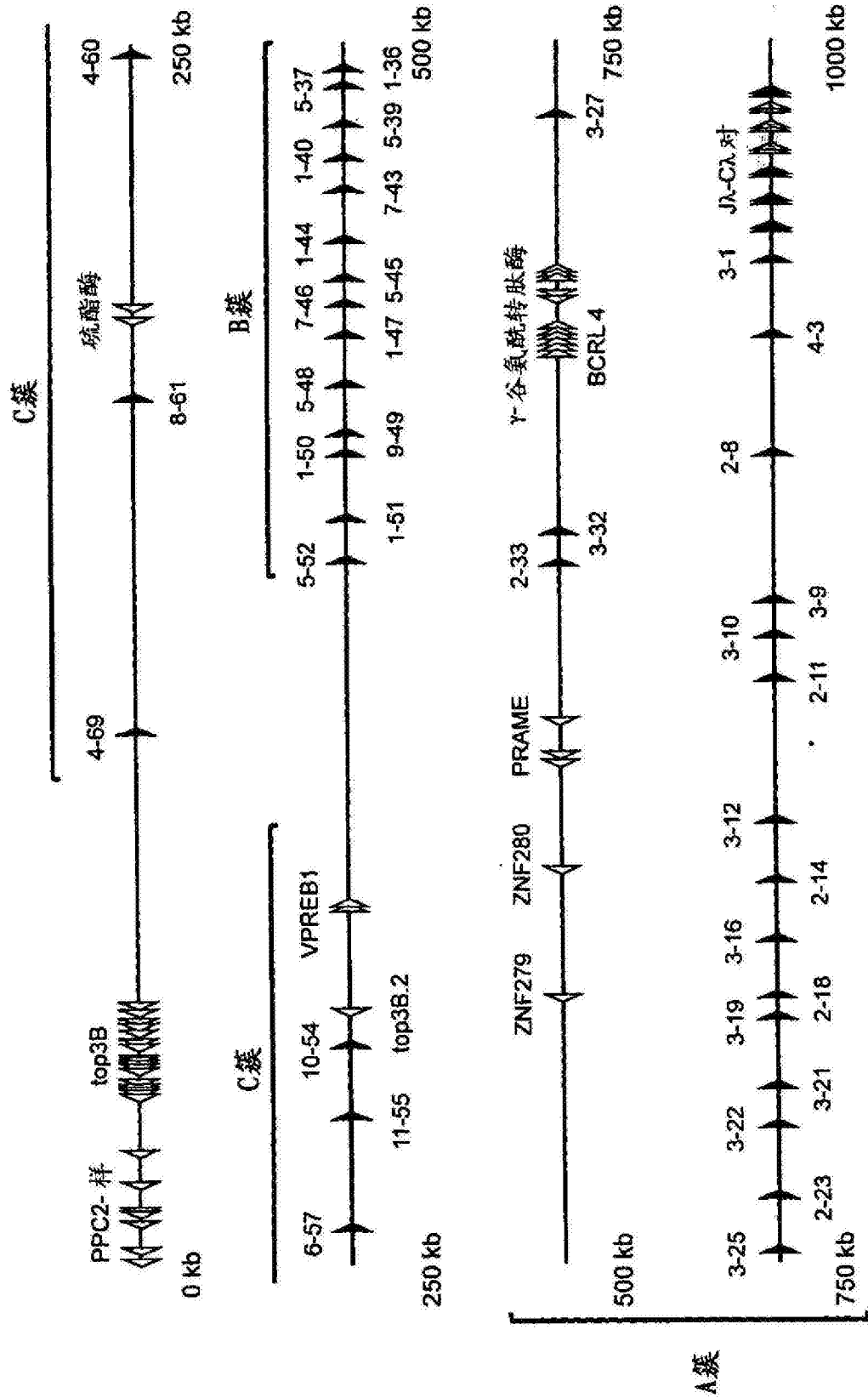


图1

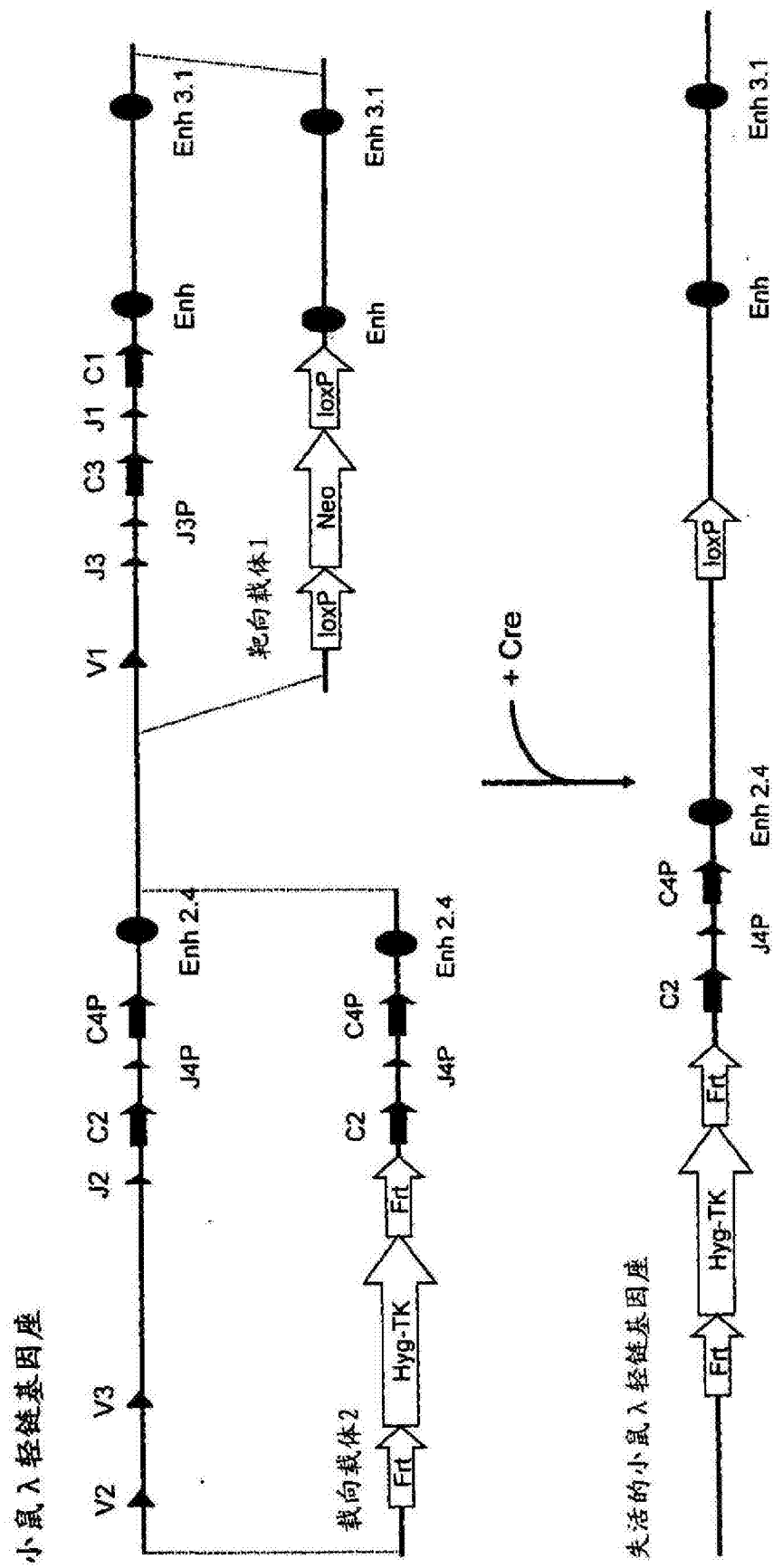


图2

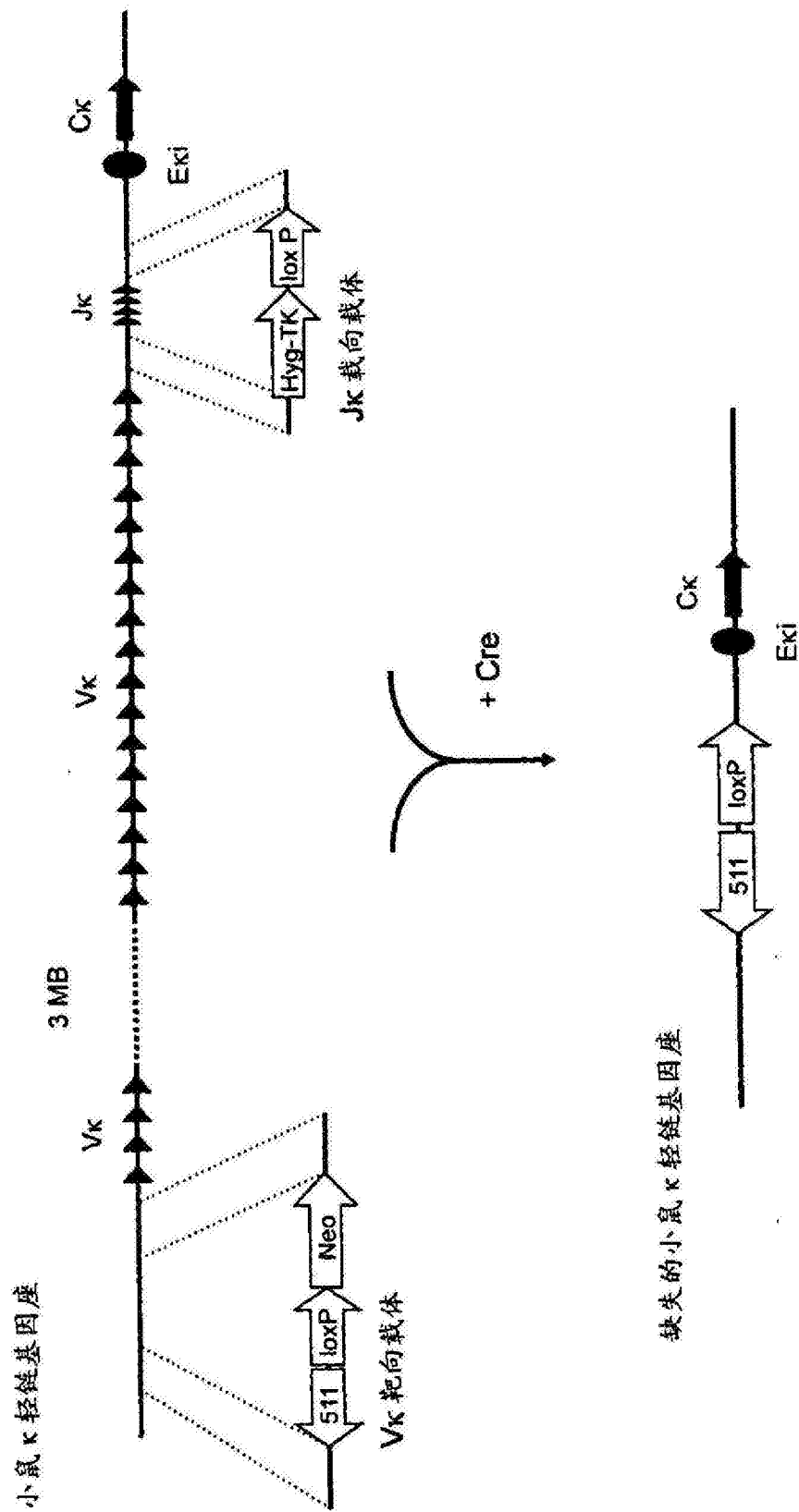


图3

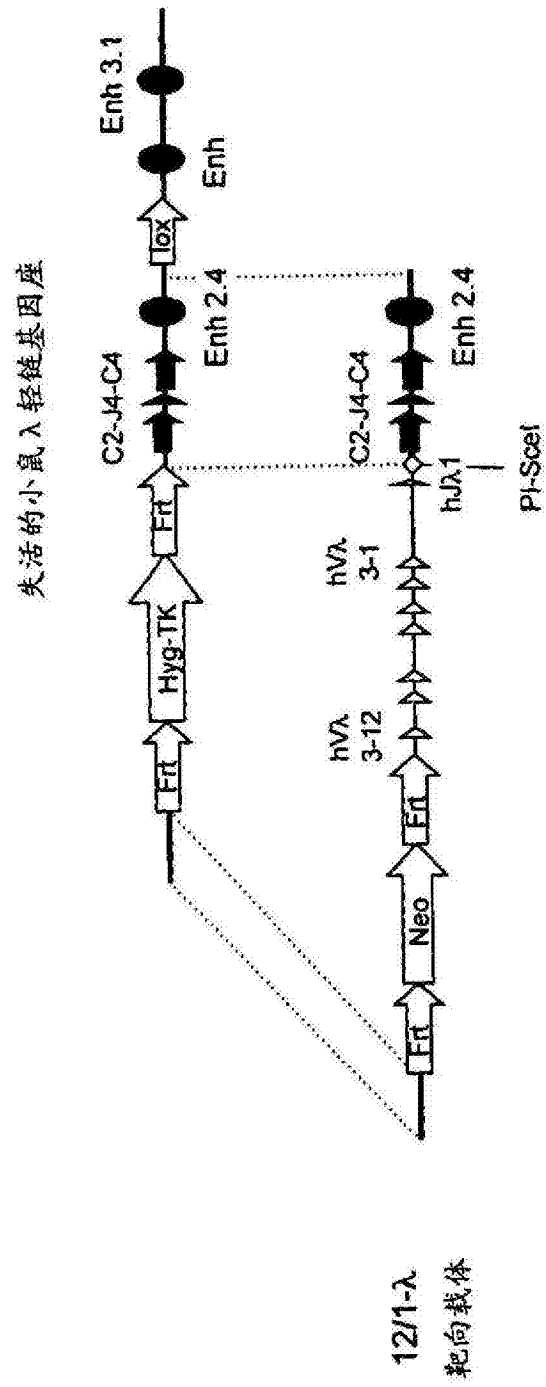


图4A

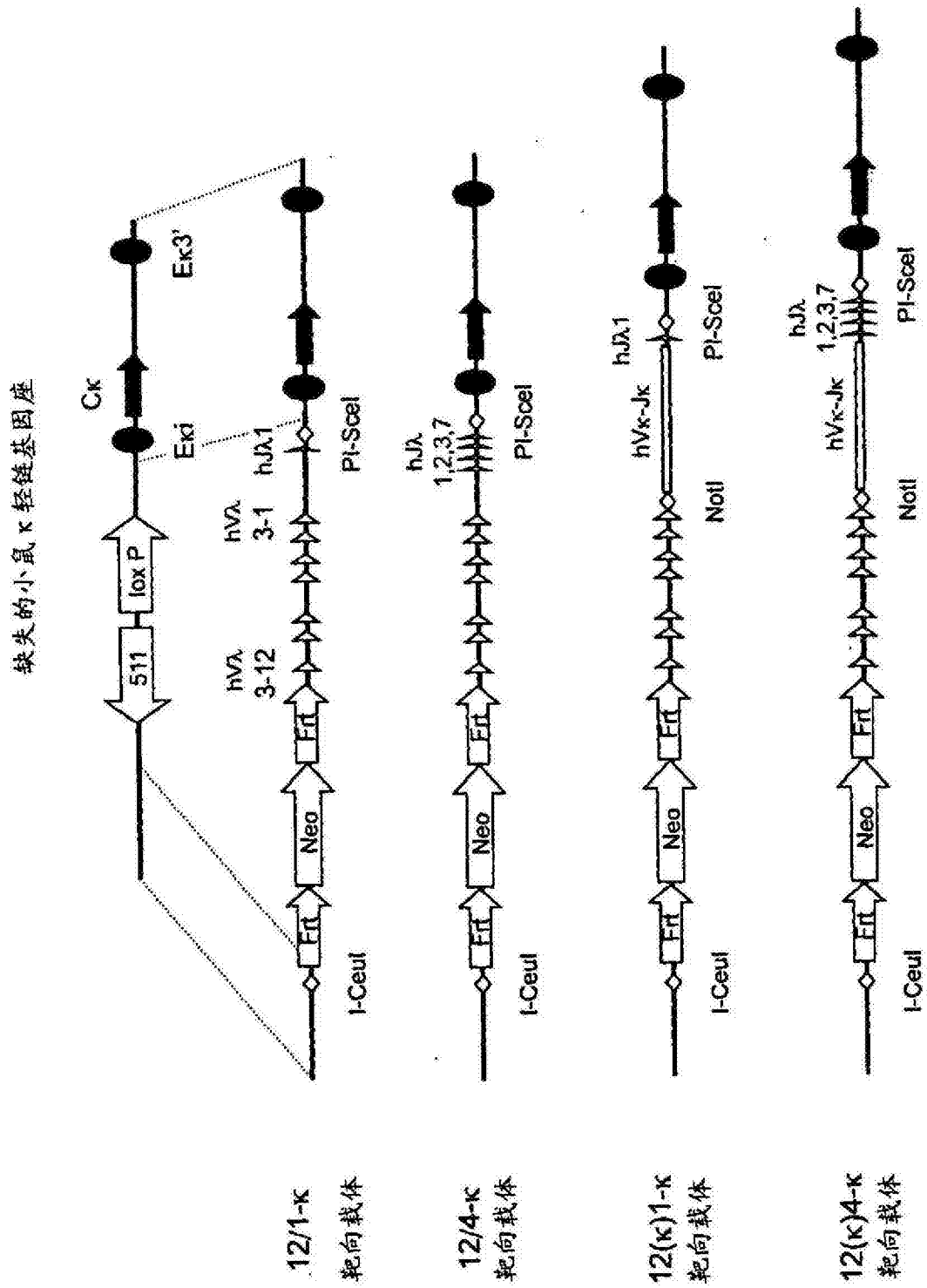


图4B

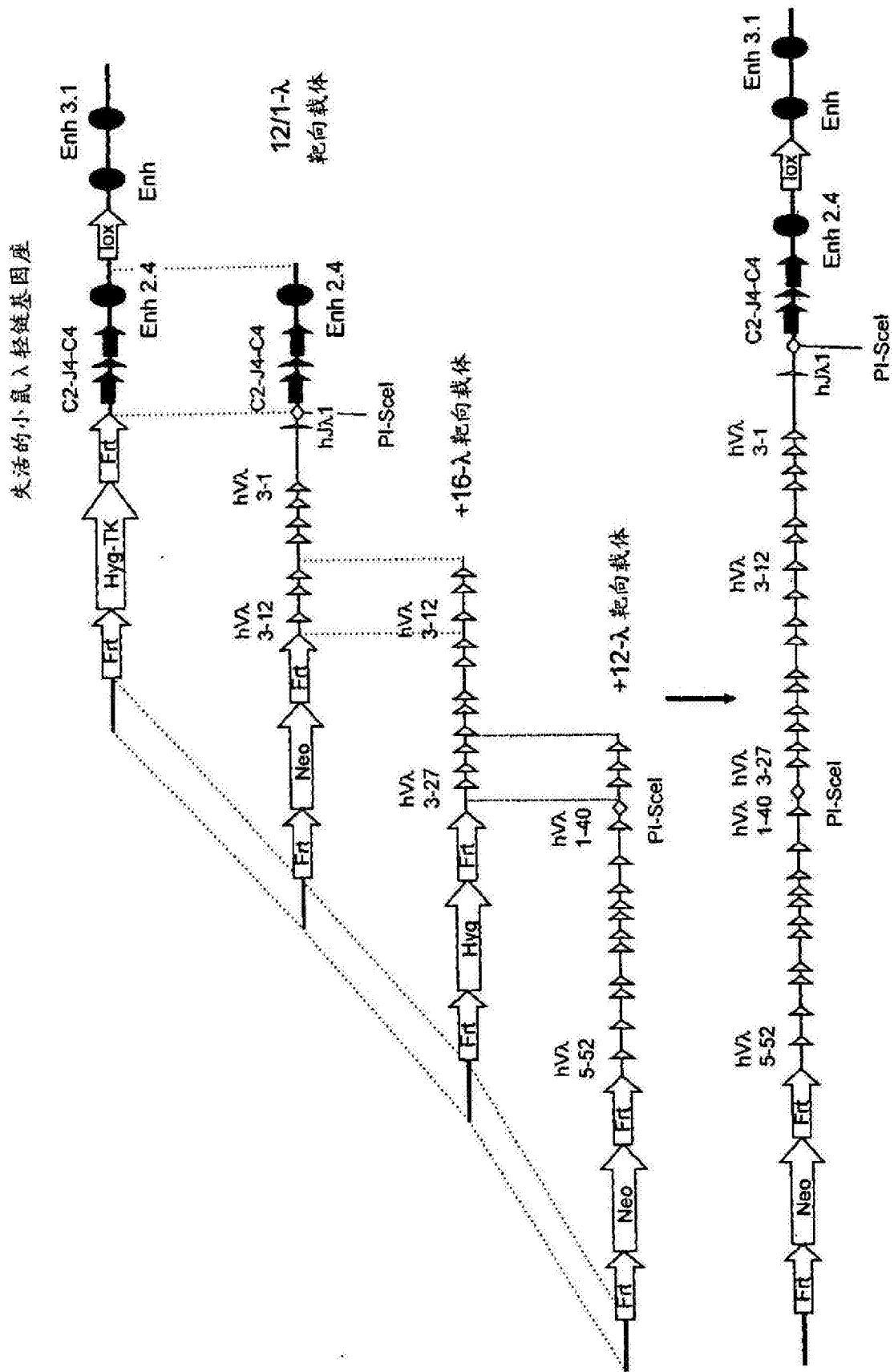


图5A

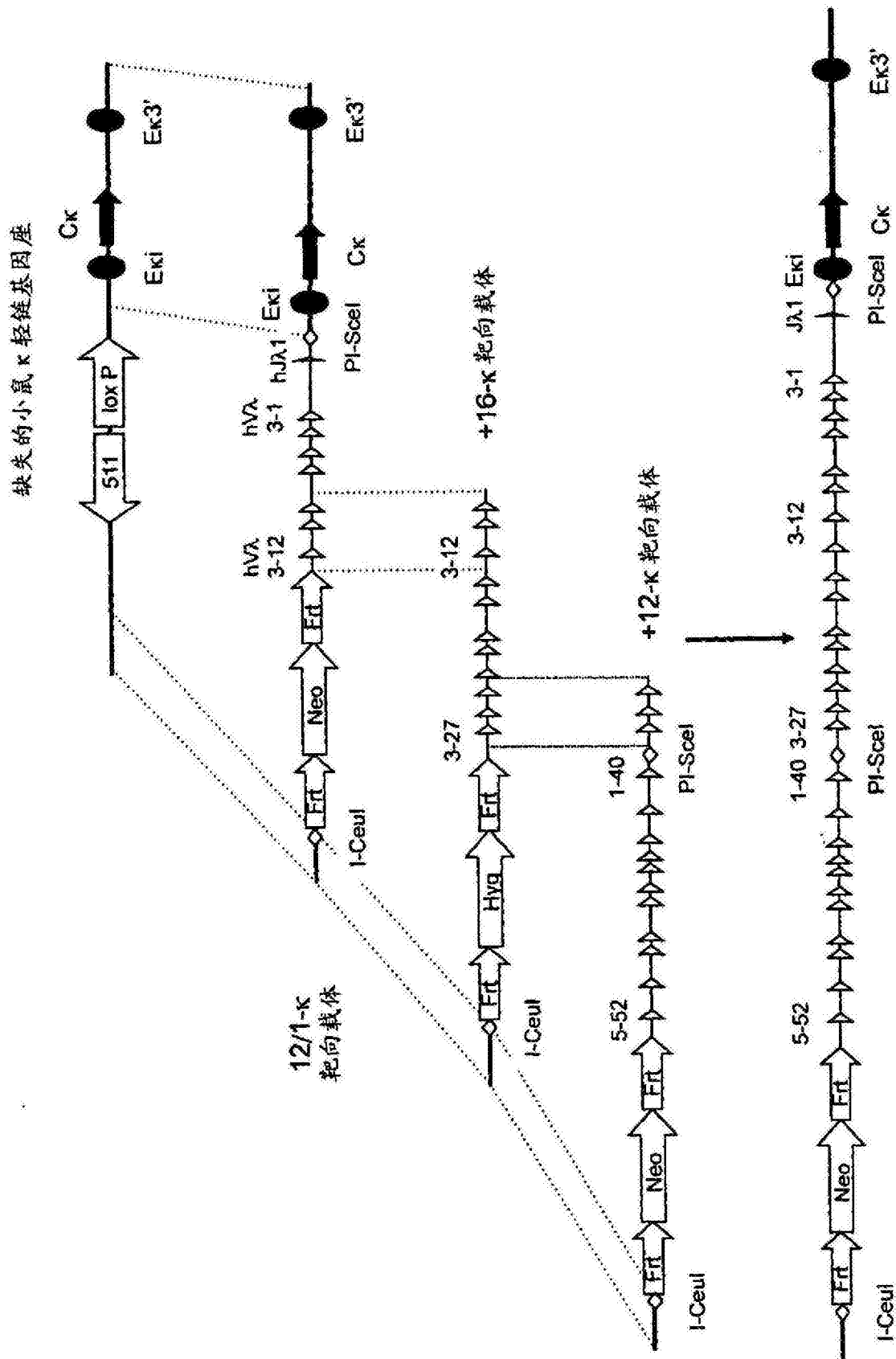


图5B

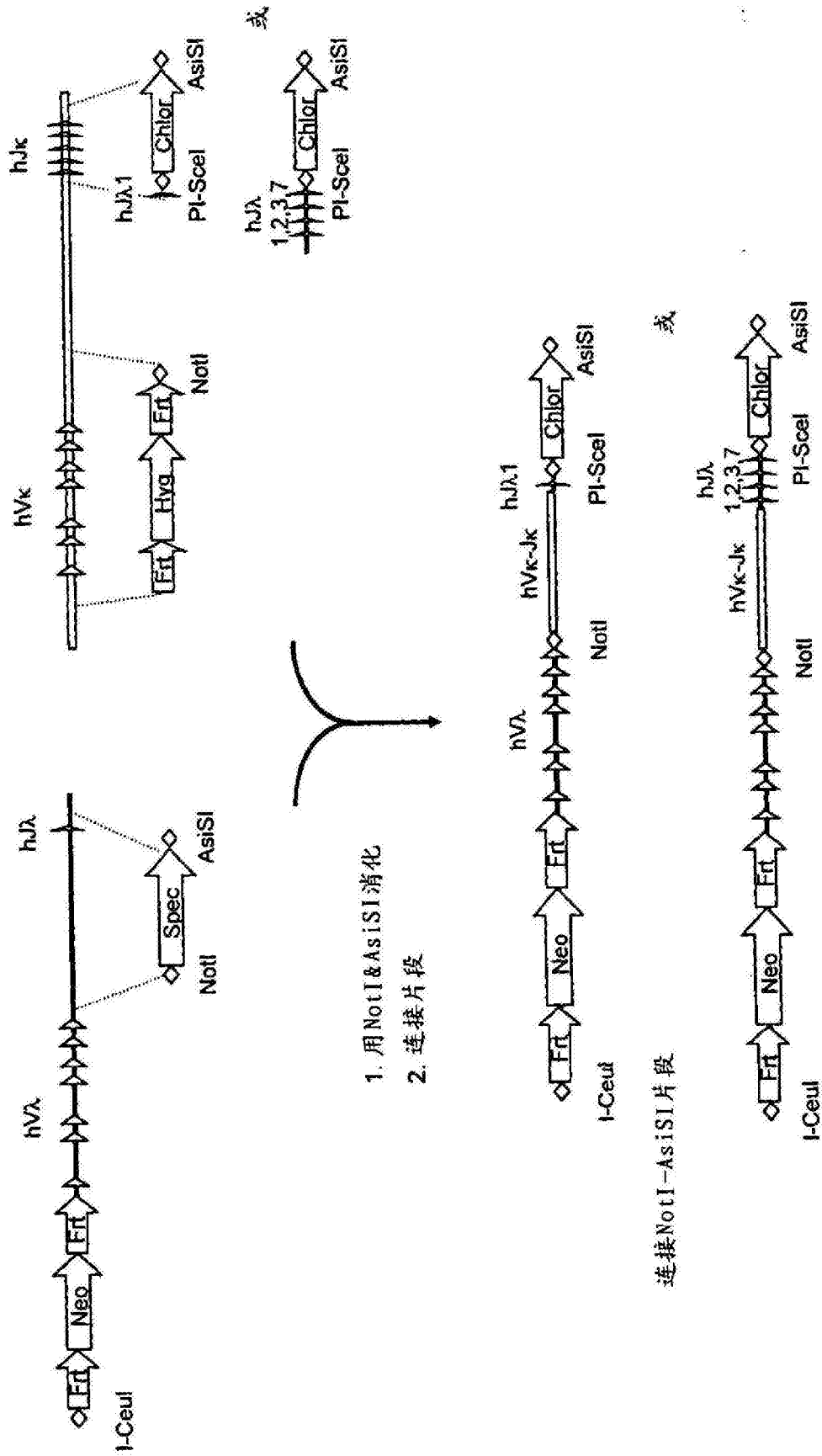


图6

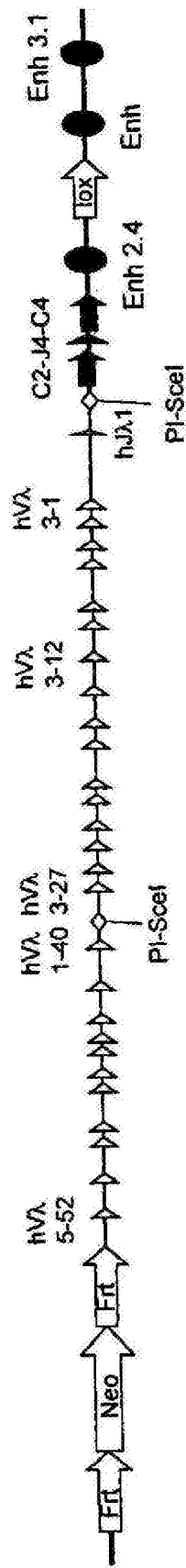


图7A

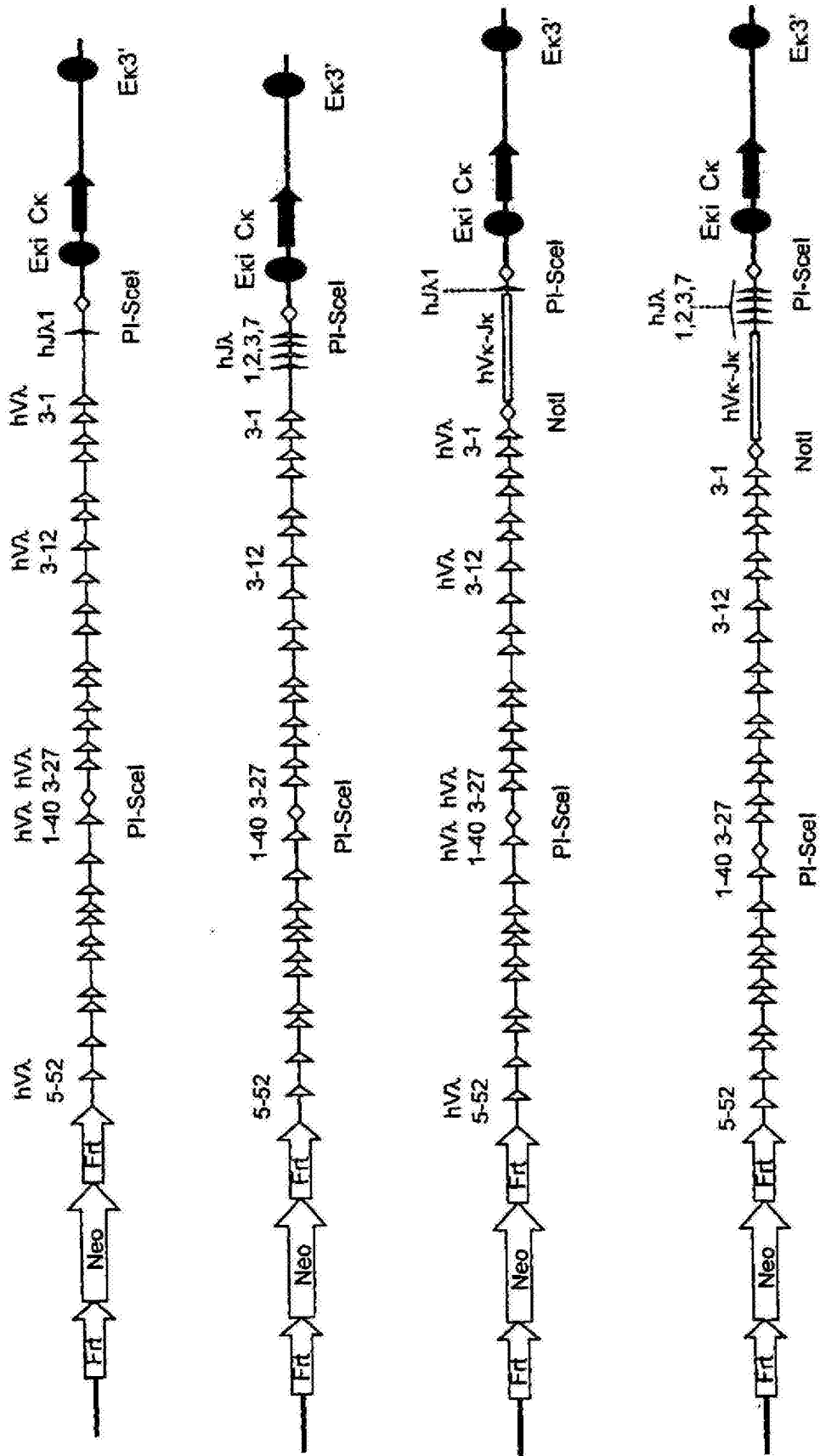


图7B

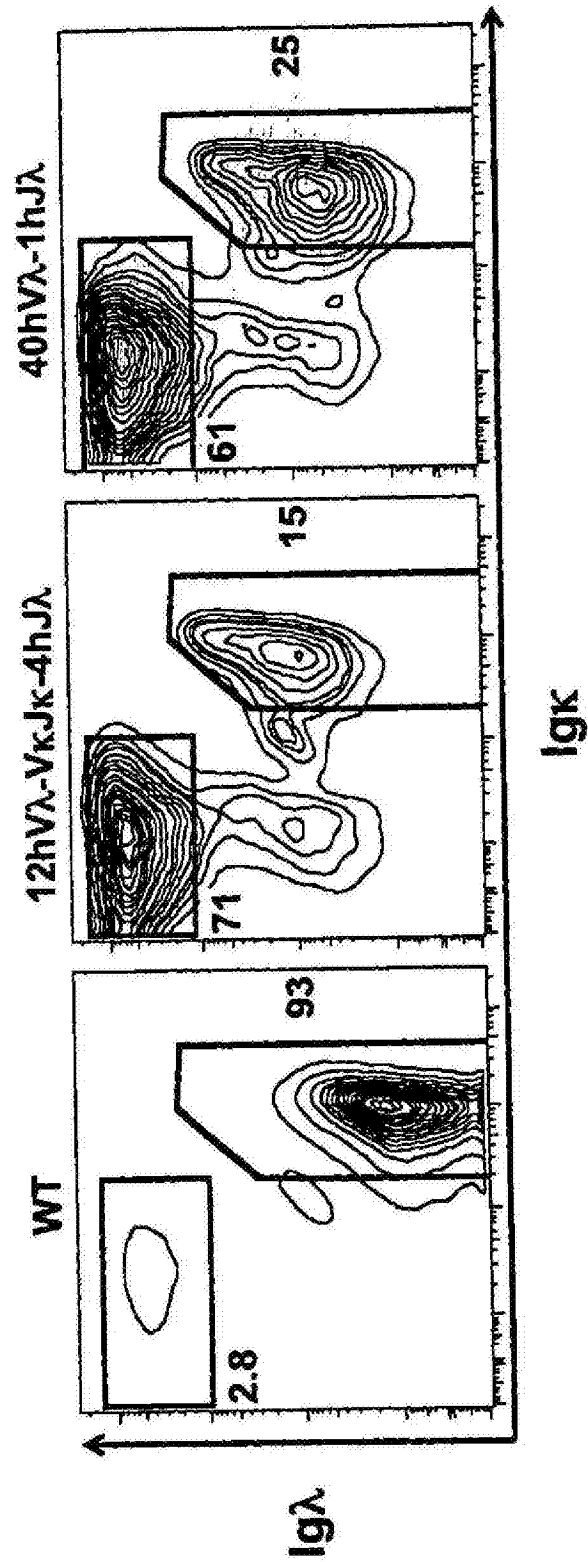


图8A

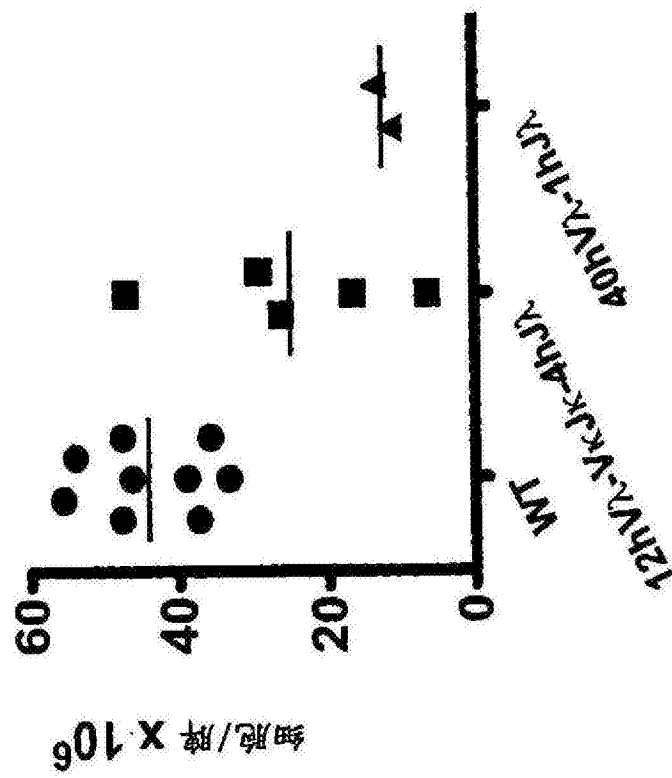


图8B

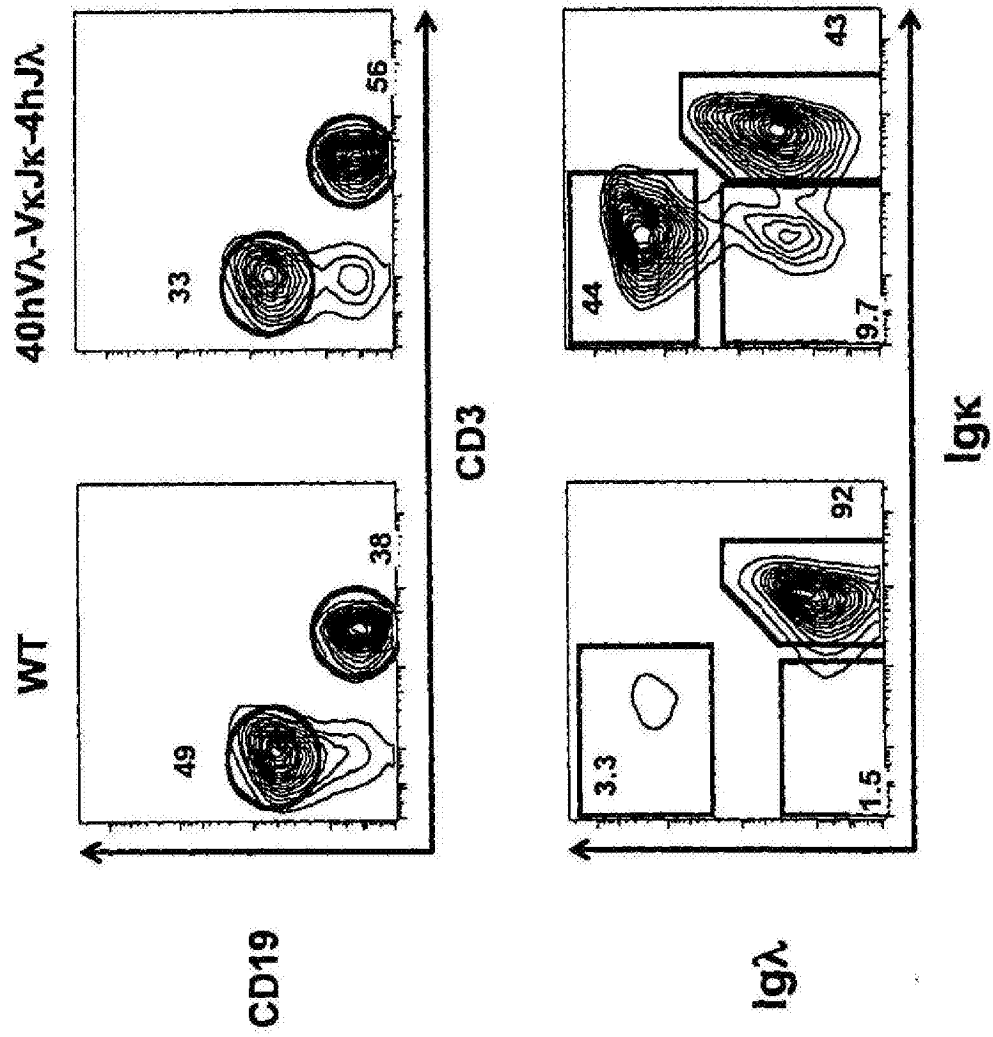


图9A

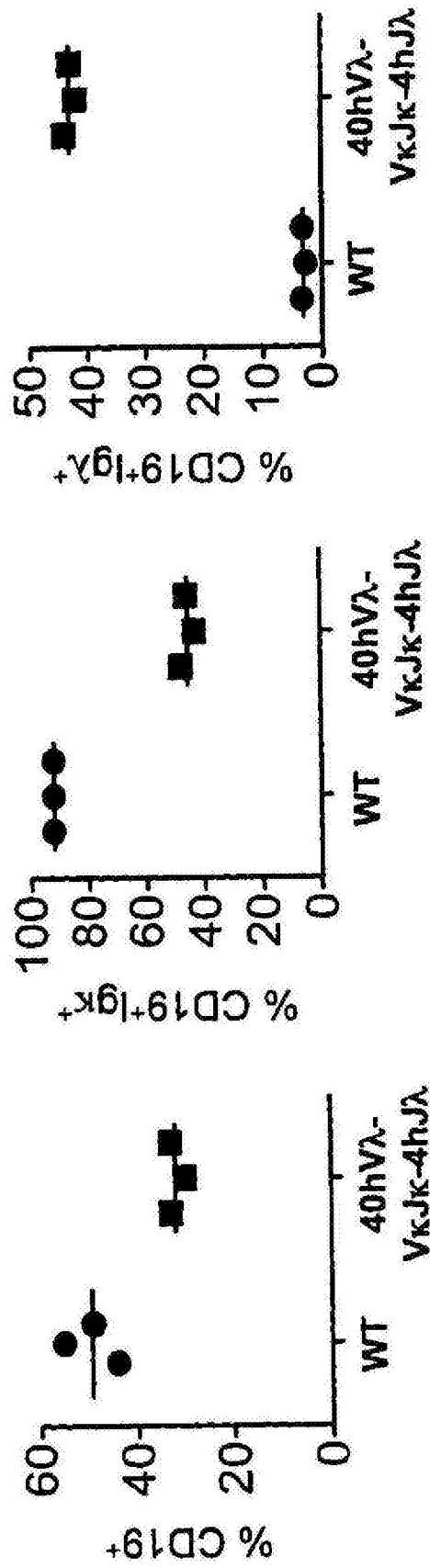


图9B

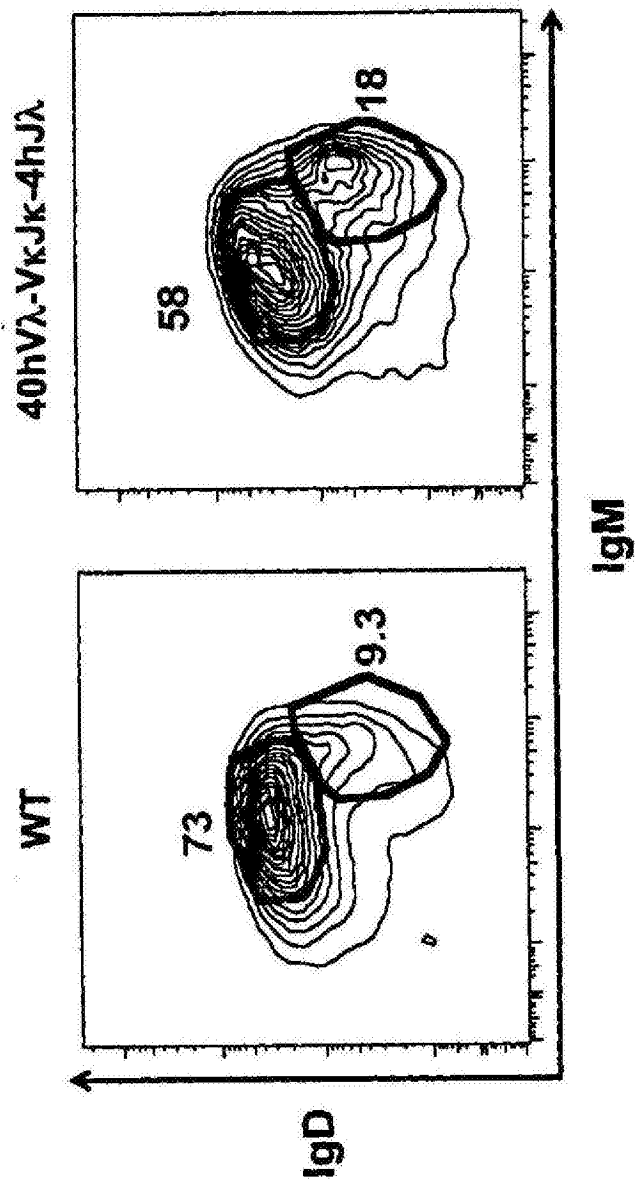


图9C

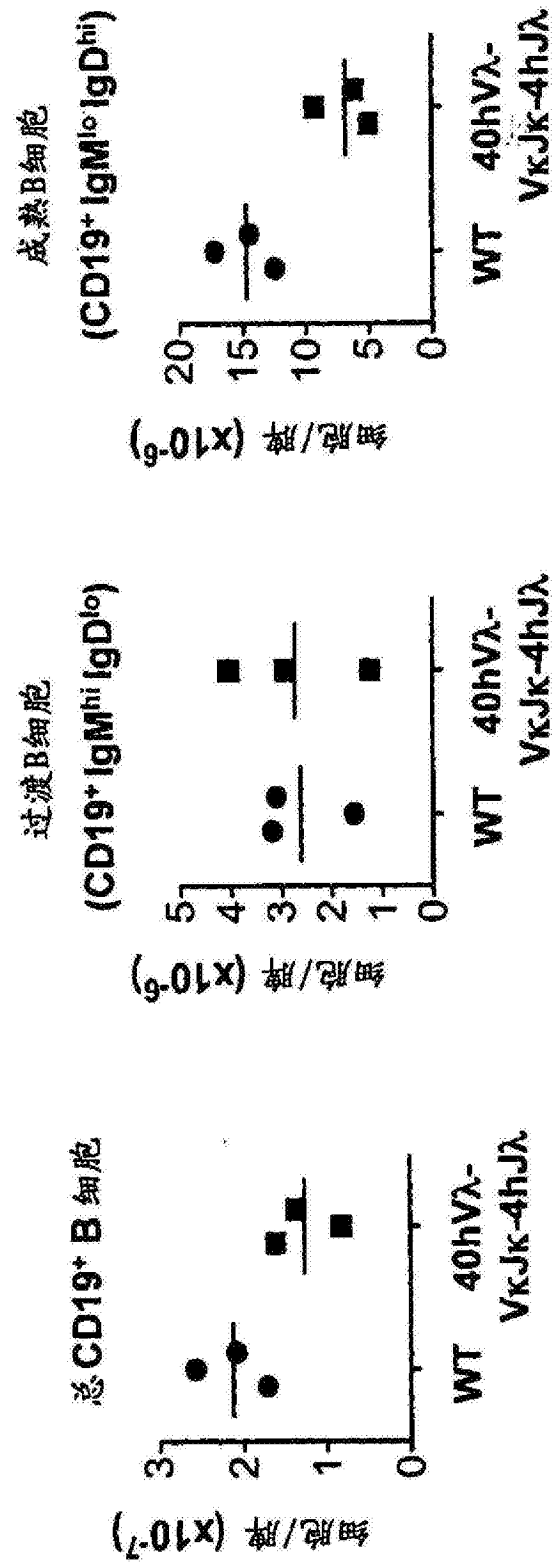


图9D

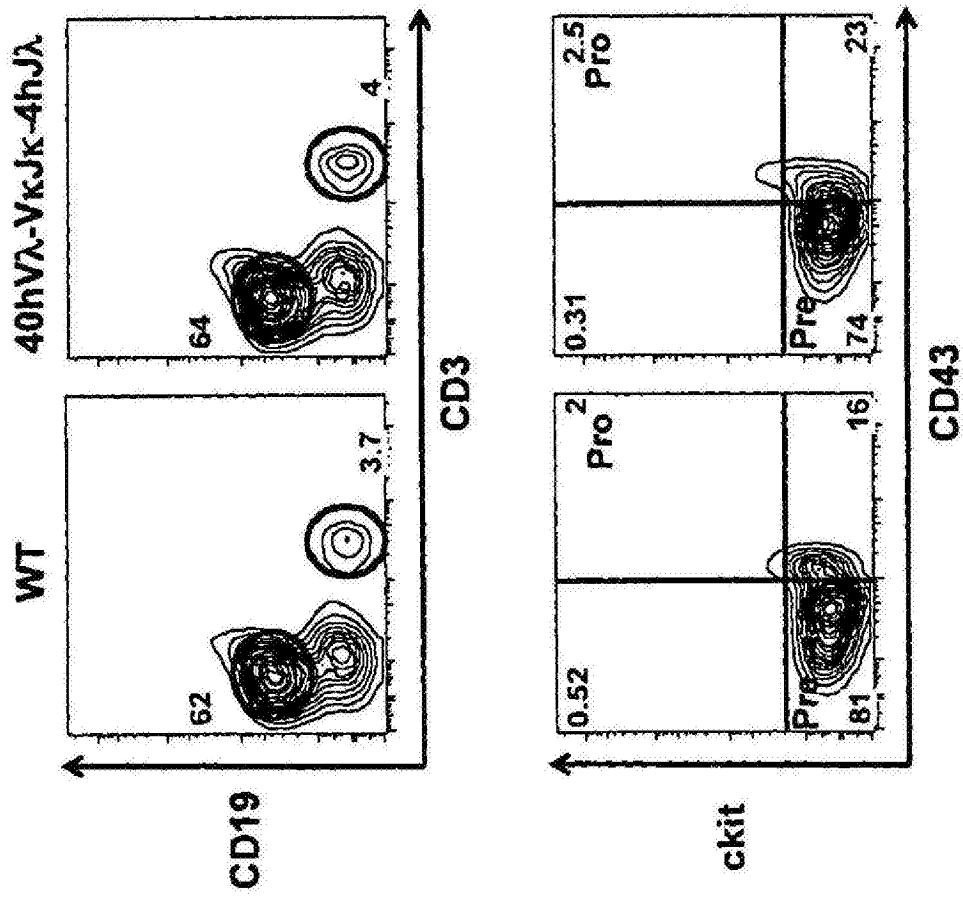


图10A

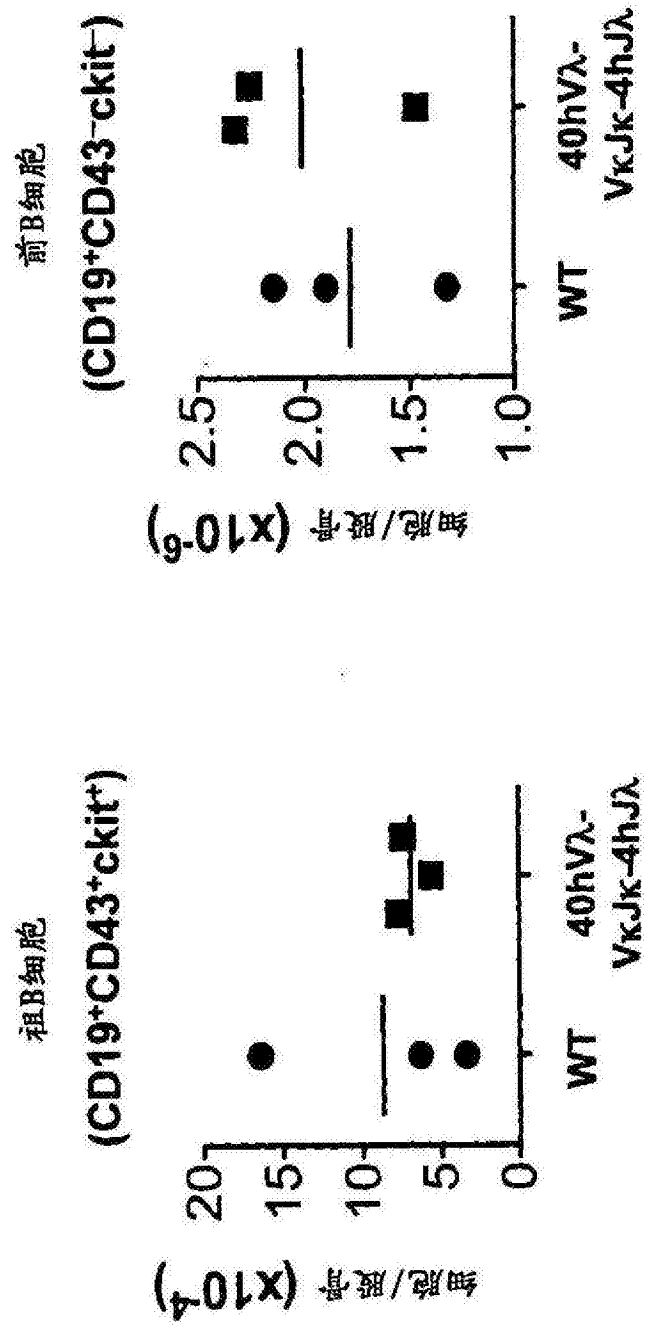


图10B

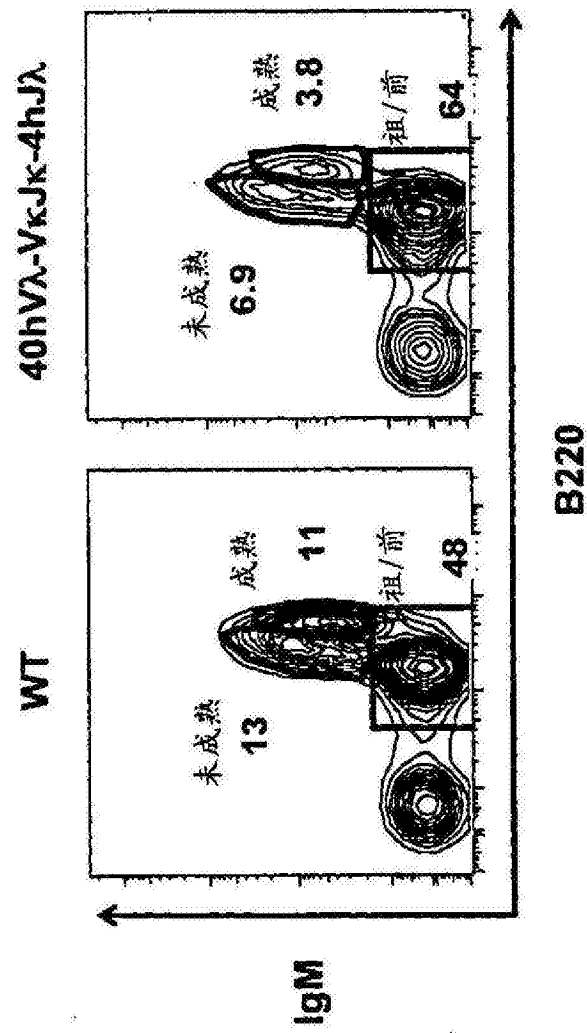


图10C

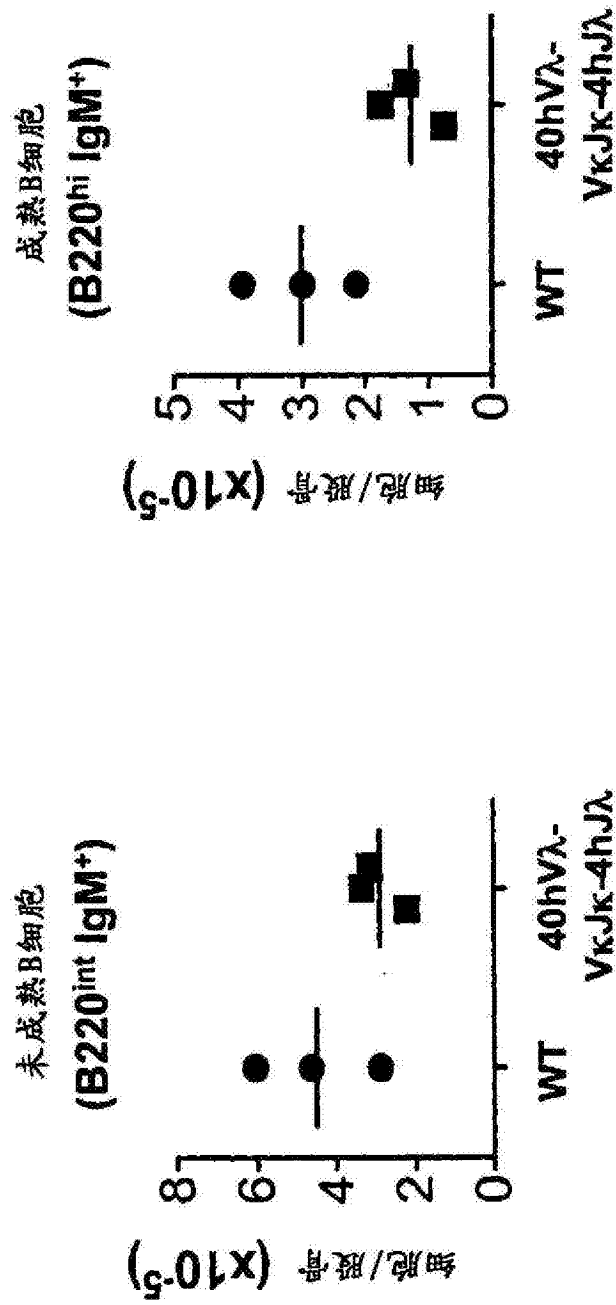


图10D

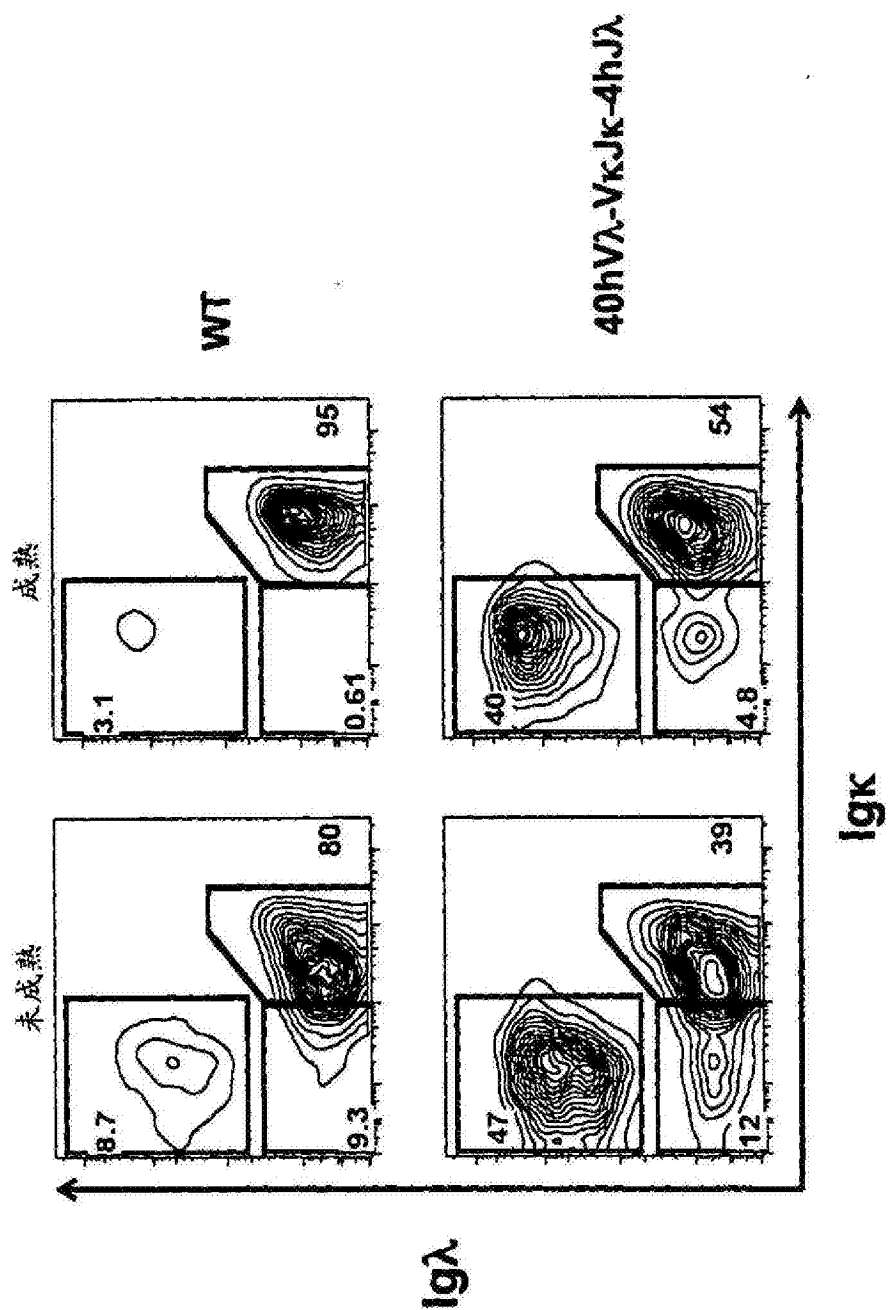


图10E

	3'人 Vλ	人 Jλ1	5'小鼠 Cκ
A6	GCAACAATT	tCGTCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
B6	GCAACAATT	ATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
F6	GCAACAATT	ATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
B7	GCAACAATT	ATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
E7	GCAACAAT	GTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
F7	GCAACAATT	ATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
C8	GCAACAATT	ATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
E12	CAAGTCGGTT	gTGTCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
1-4	TGAGTGCT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
1-20	TGAGTGCG	gccttcttTtGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
3B43	CTGAATGGT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
5-8	AGTGGTAAT	CATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
5-19	AGTGGTGCT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
1010	AGCAGCACT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
11A1	AGCAGCGCT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
7A8	GGTGGTGCT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
3A3	AGTAGCACT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
2-7	AGCAGCACT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GGGCTGATGCTGCACCAACTGTATCCATCTTC
FWR4		F G T G T K V T V L G A D A A P T V S I F	

图11

	3' 人 λ	人 λ	5' 小鼠 α
5-2	CAGCCTGAGTGGTTC	TGTGTTCCGGAGGAGGCACCCGGCTGACCGCCCTCG	GGGCTGATGCTGCACCAACTGTATCCATC
2-5	CAGCCTGAGTGGTT	ATGTCCTCGGAACATGGGACCAAGGTCAACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
1-3	CAGCCTGAATGGT	GCTGTGTTCCGGAGGAGGCACCCAGCTGACCGCCCTCG	GGGCTGATGCTGCACCAACTGTATCCATC
4B-1	CAGCCTGAGTGGTC	GGGTGTTCCGGAGGAGGCACCAAGCTGACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
3B-5	CAGCAGCACTGC	TGTGTTCCGGAGGAGGCACCCAGCTGACCGCCCTCG	GGGCTGATGCTGCACCAACTGTATCCATC
7A-1	CAGCAGTGGTAAT	GCTGTGTTCCGGAGGAGGCACCCAGCTGACCGCCCTCG	GGGCTGATGCTGCACCAACTGTATCCATC
5-1	CAGCAGTGGTAATCAATAG	GGTGTTCGGCGGAGGAGGCACCAAGCTGACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
4A-1	CAGCCTGAGTGGTT	ATGTCCTCCGGAACATGGGACCAAGGTCAACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
11A-1	CAGCAGCGCT	GTGGTAATCCGGCGGAGGAGGCACCAAGCTGACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
5-7	CTACTATGGTGGTGCTC	GGGTGTTCCGGCGGAGGAGGCACCAAGCTGACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
5-4	CTCCTATAGTGGTGCTCGa	GTATTCGGCGGAGGAGGCACCAAGCTGACCGTCCCTAG	GGGCTGATGCTGCACCAACTGTATCCATC
2-3	GAGCAACTTCGTGT	CTGTGTTCCGGAGGAGGCACCCAGCTGACCGCCCTCG	GGGCTGATGCTGCACCAACTGTATCCATC
FWR4		P G G G T K L T V L G A D A A P T V S I	

图12

	3'人Vλ	人JA1	5'小鼠Cλ2
2D1	GCAGGCAGCAACAATTa	aGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GTCAGCCCCAAGTCCCACTCCCACTCTC
2D9	GACAGCAGTGGTAATCAT	TATGTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GTCAGCCCCAAGTCCCACTCCCACTCTC
3E15	GACAGCAGCACTGCC	GTCCTTCGGAACTGGGACCAAGGTCACCGTCCTAG	GTCAGCCCCAAGTCCCACTCCCACTCTC
FWR4		F G T G T K V T V L G Q P K S T P T L	

图13