



(12) 发明专利申请

(10) 申请公布号 CN 112877472 A

(43) 申请公布日 2021.06.01

(21) 申请号 202110128389.8

(22) 申请日 2014.11.07

(30) 优先权数据

61/901,114 2013.11.07 US

61/901,857 2013.11.08 US

(62) 分案原申请数据

201480068555.8 2014.11.07

(71) 申请人 小利兰·斯坦福大学理事会

地址 美国加利福尼亚州

(72) 发明人 依维津·德弗拉芒克

迈克尔·凯尔泰斯

基兰·K·克霍什

马克·亚力克·科瓦尔斯基

兰斯·马丁 S·R·奎克

H·瓦朗坦

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 李敏春 郑霞

(51) Int.Cl.

C12Q 1/70 (2006.01)

C12Q 1/6883 (2018.01)

C12Q 1/6895 (2018.01)

C12Q 1/689 (2018.01)

C12Q 1/6869 (2018.01)

G16B 30/10 (2019.01)

G16B 40/20 (2019.01)

G16B 45/00 (2019.01)

G16B 50/00 (2019.01)

权利要求书1页 说明书33页 附图13页

(54) 发明名称

用于分析人体微生物组及其组分的无细胞核酸

(57) 摘要

本申请涉及用于分析人体微生物组及其组分的无细胞核酸。提供了用于分析个体中的微生物组或其个体组分的方法、装置、组合物和试剂盒。所述方法在感染的确定、微生物组结构的分析、个体的免疫活性的确定等中有用。在本发明的一些实施方案中，个体被用例如药物、膳食、放射疗法等的治疗方案治疗。

1. 一种确定来自非微生物宿主的无细胞核酸的样品中的微生物序列的存在和发生率的方法,所述方法包括:

- (i) 提供来自个体的无细胞核酸的样品;
- (ii) 对所述核酸进行高通量测序;
- (iii) 进行生物信息学分析以从分析扣除宿主序列;和
- (iv) 确定微生物序列的存在和发生率,以对非微生物宿主进行微生物组评价。

2. 根据权利要求1所述的方法,其中多种微生物的存在和发生率被确定。

3. 根据权利要求1所述的方法,其中对已通过无偏倚方法扩增的核酸样品进行高通量测序。

4. 根据权利要求1所述的方法,其中进行至少 10^6 个序列读段。

5. 根据权利要求1所述的方法,其中步骤(iv)包括将映射至微生物参考序列的序列的覆盖度与宿主参考序列的覆盖度比较。

6. 根据权利要求1所述的方法,其中步骤(iii)包括鉴定参考宿主序列和掩蔽在参考宿主基因组中存在的微生物序列或微生物模拟序列。

7. 根据权利要求1所述的方法,其中步骤(iii)包括鉴定参考微生物序列和掩蔽在参考微生物基因组中存在的宿主序列或宿主模拟序列。

8. 根据权利要求1所述的方法,其中一种或更多种病原微生物的存在被鉴定。

9. 根据权利要求1所述的方法,其中在2个或更多个时间点进行分析。

10. 根据权利要求1所述的方法,其中一种或更多种微生物核酸的量指示感染状态或治疗结果。

用于分析人体微生物组及其组分的无细胞核酸

[0001] 本申请是申请日为2014年11月07日,申请号为201480068555.8,发明名称为“用于分析人体微生物组及其组分的无细胞核酸”的申请的分案申请。

[0002] 发明背景

[0003] 人类微生物组现在被认为是人类健康的重要组成部分。社区水平的分析已经阐明了使微生物组的细菌和病毒组分的结构形成的因素,诸如年龄、膳食、地理位置、抗生素治疗和疾病。例如,个体的微生物组可通过感染病原生物而被改变,以使得该生物全身性地或在非期望的组织中的发生率 (prevalence) 增加。微生物组还可通过个体免疫活性的变化而被改变。

[0004] 出于多种目的,期望得到以下:用于快速鉴定特定的微生物组组分,例如在个体微生物组中共生 (commensal) 生物、共栖 (mutualistic) 生物、寄生生物、机会主义 (opportunistic) 生物和病原生物的存在或发生率的方法;和对整体微生物组结构的分析。本发明提供了监测临床样品中微生物组组成的灵敏、快速、非侵入性的方法。

[0005] 发明概述

[0006] 本发明提供了用于分析个体中的微生物组或其个体组分的方法、装置、组合物和试剂盒。所述方法在感染的确定、微生物组结构的分析、个体的免疫活性的确定等中有用。在一些实施方案中,本发明提供了确定个体中微生物的存在和发生率的方法,所述方法包括以下步骤:(i) 提供无细胞核酸的样品,即来自个体的DNA和/或RNA;(ii) 进行高通量测序,例如每个样品从约 10^5 起并且最多约 10^9 或更多个读段 (reads);(iii) 进行生物信息学分析以从分析扣除宿主序列,即人类、猫、狗等;以及(iv) 确定微生物序列的存在和发生率,例如通过将映射到微生物参考序列的序列的覆盖度与宿主参考序列的覆盖度比较来确定。

[0007] 宿主序列的扣除可包括鉴定参考宿主序列和掩蔽存在于参考宿主基因组中的微生物序列或微生物模拟序列的步骤。类似地,通过与微生物参考序列比较确定微生物序列的存在可包括鉴定参考微生物序列和掩蔽存在于参考微生物基因组中的宿主序列或宿主模拟序列的步骤。

[0008] 本发明的特征是来自个体的无细胞核酸的无偏倚分析。本发明的方法通常包括无偏倚扩增步骤,例如通过用通用引物进行PCR或通过接头连接到核酸并用接头特异性的引物扩增。本发明的方法通常在不存在微生物序列的序列特异性扩增下进行。该方法的益处是分析因而包括所有可得的微生物组序列,但是其需要生物信息学分析以鉴定宿主序列占优势的复杂数据集中的感兴趣的序列。

[0009] 本发明的方法的另外的益处是提供个体微生物组的快速评价的能力,例如分析可以在少于约3天、少于约2天、少于1天例如少于约24小时、少于约20小时、少于约18小时、少于约14小时、少于约12小时、少于约6小时、少于约2小时、少于约30分钟、少于约15分钟、少于约1分钟内完成。

[0010] 在一些实施方案中,无细胞核酸的分析被用于计算病原性评分,其中病原性评分是概括生物体的总体病原性以易于被例如保健医生理解的数值或字母值。存在于微生物组的不同微生物可以被分配不同的评分。

[0011] 微生物序列的存在和发生率的分析可以被用来提供：关于感染对疗法的响应的结论(determination),所述疗法包括抗微生物治疗诸如用抗生素、抗病毒剂、免疫、被动免疫疗法等治疗,膳食,免疫抑制等;关于感染在临床试验中的响应的结论等。从分析获得的信息可以被用来诊断状况、监测治疗、选择或修改治疗方案以及优化疗法。使用该方法,根据在治疗过程中在不同的时间获得的专属性数据,治疗和/或诊断方案可以是个性化和定制的,从而提供个体适用的方案。另外,患者样品可以在任何点获得用于分析:在治疗过程期间;暴露于病原体后;感染的过程等。微生物序列的存在和发生率的分析可以作为报告来提供。报告可以被提供给个体、医疗保健专业人员等。

[0012] 在一些实施方案中,无细胞核酸从选自由以下组成的组的生物样品获得:血液、血清、脑脊液、滑液、尿液和粪便。可使用从样品的无细胞部分例如血液的血清或血浆部分提取的核酸。在一些实施方案中,核酸选自由以下组成的组:双链DNA、单链DNA、单链DNA发夹、DNA/RNA杂交物、单链RNA、双链RNA和RNA发夹。在一些实施方案中,核酸选自由双链DNA、单链DNA和cDNA组成的组。在一些实施方案中,核酸是mRNA。在一些实施方案中,核酸是循环的无细胞DNA。

[0013] 在一些实施方案中,所述方法包括定量一种或更多种核酸以提供在样品中微生物的发生率的结论。在一些实施方案中,一种或更多种核酸高于预定的阈值的量指示感染或改变的发生率。在一些实施方案中,对于不同的微生物生物体存在不同的预定的阈值。在一些实施方案中,一种或更多种核酸的量的时间差异指示在感染、改变的发生率、对疗法的响应等方面的改变。

[0014] 在一些实施方案中,本发明提供了计算机可读介质,所述计算机可读介质包含记录于其上的一组指令,所述指令导致计算机进行以下步骤:(i)接收来自在来自受试者的无细胞核酸的样品中检测的一种或更多种核酸的高通量测序数据;(ii)进行生物信息学分析以从分析扣除宿主序列,即人类、猫、狗等;和(iii)确定微生物序列的存在和发生率,例如通过将映射到微生物参考序列的序列的覆盖度与宿主参考序列的覆盖度比较来确定。

[0015] 在一些实施方案中,本发明提供了用于实践本文描述的方法中的一种或更多种的试剂及其试剂盒。

[0016] 在一些实施方案中,提供了用于通过分析微生物组,例如通过分析病毒组,来评价个体、特别地人类个体的免疫活性的组合物和方法。在本发明的一些实施方案中,个体被用免疫抑制方案例如药物、放射疗法等治疗。在一些实施方案中,个体是用免疫抑制方案治疗的移植物接受者。在一些实施方案中,个体具有用免疫抑制方案治疗的自身免疫病。在其他的实施方案中,在免疫抑制方案不存在下评价个体的免疫活性。

[0017] 在一些实施方案中,来自个体的测量值取于两个或更多个时间点,其中病毒荷载的变化指示免疫活性的变化。个体可以根据免疫活性的评价而被治疗,例如其中在移植患者中不期望的增加的免疫活性的适应症用增加的免疫抑制剂水平来治疗;或其中免疫活性的不期望的降低用治疗试剂例如抗病毒剂等来治疗。

[0018] 核酸分析被用于鉴定和定量从患者收集的样品中的非人类无细胞核酸。微生物组的组分的组成如以上描述的进行。微生物组(病毒组)的病毒组分的结构允许预测免疫活性。在一些实施方案中,所述方法还包括在免疫抑制方案之前、在免疫抑制方案的开始或在免疫抑制方案的过程期间构建病毒组谱系(virome profile),其可被用作个体病毒组的变

化的参考。在一些实施方案中,循环的无细胞DNA是指环病毒属DNA。

[0019] 特别地,指环病毒科病毒的荷载是免疫强度的预测剂,其与器官移植物排斥的概率相关。虽然其他的病毒也可以是预测性的,但是对于将用影响此类病毒的荷载的抗病毒剂治疗的患者,此类病毒是常用的。

[0020] 在一些实施方案中,本发明提供了诊断或预测移植状态或结果的方法,所述方法包括以下步骤:(i) 提供来自受试者的样品,所述受试者已接受来自供体的移植物;(ii) 确定一种或更多种病毒组核酸的存在或不存在;和(iii) 基于病毒组荷载诊断或预测移植状态或结果。在一些实施方案中,移植状态或结果包括排斥、耐受、基于非排斥的同种异体移植物损伤、移植物功能、移植物存活、慢性移植物损伤或滴度药理学免疫抑制(titer pharmacological immunosuppression)。在一些实施方案中,一种或更多种核酸大于预定的阈值的量指示病毒载量和免疫活性。在一些实施方案中,阈值是无移植排斥或其他病理的迹象的临床稳定的移植后患者的标准值。在一些实施方案中,对于不同的移植结果或状态存在不同的预定的阈值。在一些实施方案中,一种或更多种核酸的量的时间差异指示免疫活性。

[0021] 在本文描述的任何实施方案中,移植移植物可以是任何实体器官、骨髓或皮肤移植物。在一些实施方案中,移植物选自由以下组成的组:肾移植物、心脏移植物、肝脏移植物、胰腺移植物、肺移植物、肠移植物和皮肤移植物。

[0022] 在一些实施方案中,本发明提供了用于实践本文描述的方法中的一种或更多种的试剂及其试剂盒。

[0023] 本申请提供了以下内容:

[0024] 1). 一种确定来自非微生物宿主的无细胞核酸的样品中的微生物序列的存在和发生率的方法,所述方法包括:

[0025] (i) 提供来自个体的无细胞核酸的样品;

[0026] (ii) 对所述核酸进行高通量测序;

[0027] (iii) 进行生物信息学分析以从分析扣除宿主序列;和

[0028] (iv) 确定微生物序列的存在和发生率,以对非微生物宿主进行微生物组评价。

[0029] 2). 根据1)所述的方法,其中多种微生物的存在和发生率被确定。

[0030] 3). 根据1)所述的方法,其中对已通过无偏倚方法扩增的核酸样品进行高通量测序。

[0031] 4). 根据1)所述的方法,其中进行至少 10^6 个序列读段。

[0032] 5). 根据1)所述的方法,其中步骤(iv)包括将映射至微生物参考序列的序列的覆盖度与宿主参考序列的覆盖度比较。

[0033] 6). 根据1)所述的方法,其中步骤(iii)包括鉴定参考宿主序列和掩蔽在参考宿主基因组中存在的微生物序列或微生物模拟序列。

[0034] 7). 根据1)所述的方法,其中步骤(iii)包括鉴定参考微生物序列和掩蔽在参考微生物基因组中存在的宿主序列或宿主模拟序列。

[0035] 8). 根据1)所述的方法,其中一种或更多种病原微生物的存在被鉴定。

[0036] 9). 根据1)所述的方法,其中在2个或更多个时间点进行分析。

[0037] 10). 根据1)所述的方法,其中一种或更多种微生物核酸的量指示感染状态或治疗

结果。

[0038] 11). 根据10)所述的方法,其中所述一种或更多种核酸的高于预定的阈值的量指示感染状态或治疗结果。

[0039] 12). 根据1)所述的方法,其中所述样品选自由血液、血清、尿液和粪便组成的组。

[0040] 13). 根据1)所述的方法,其中所述核酸选自由双链DNA、单链DNA、单链DNA发夹、DNA/RNA杂交物、单链RNA、双链RNA和RNA发夹组成的组。

[0041] 14). 根据1)所述的方法,其中所述核酸选自由双链DNA、单链DNA、和cDNA组成的组。

[0042] 15). 根据1)所述的方法,还包括提供所述个体的微生物组的评价。

[0043] 16). 根据1)所述的方法,其中所述微生物组的评价提供对疗法的响应的结论。

[0044] 17). 根据1)所述的方法,其中所述微生物组的评价提供对人体生理的测量。

[0045] 18). 根据1)所述的方法,其中所述微生物组的评价被用于计算非微生物宿主中存在的微生物的病原性评分。

[0046] 19). 一种计算机可读介质,包括:

[0047] 记录于其上的一组指令,所述一组指令导致计算机进行以下的步骤:

[0048] (i) 接受来自在来自受试者的样品中检测到的一种或更多种无细胞核酸的高通量数据;和

[0049] (ii) 进行生物信息学分析以从分析扣除宿主序列;和

[0050] (iii) 确定微生物序列的存在和发生率。

[0051] 20). 一种评价个体的免疫活性的方法,所述方法包括:

[0052] 提供来自受试者的样品;

[0053] 确定一种或更多种微生物组核酸在所述样品中的存在或不存在;和

[0054] 基于所述一种或更多种微生物组核酸的存在评价免疫活性。

[0055] 21). 根据20)所述的方法,其中分析微生物组的病毒组组分。

[0056] 22). 根据21)所述的方法,其中所述一种或更多种病毒组核酸的量的时间差异指示免疫活性状态。

[0057] 23). 根据21)所述的方法,包括定量个体中的病毒载量。

[0058] 24). 根据23)所述的方法,其中对所述病毒组分析指环病毒属的病毒载量。

[0059] 25). 根据20)所述的方法,其中所述个体已接受免疫抑制方案。

[0060] 26). 根据20)所述的方法,其中所述个体已接受移植。

[0061] 27). 根据26)所述的方法,其中所述移植选自由以下组成的组:骨髓移植、肾移植、心脏移植、肝移植、胰腺移植、肺移植、肠移植和皮肤移植。

[0062] 28). 根据20)所述的方法,其中所述核酸是循环的无细胞DNA。

[0063] 29). 根据20)所述的方法,其中所述一种或更多种核酸的存在或不存在通过选自由测序、核酸阵列和PCR组成的组的方法确定。

[0064] 30). 根据20)所述的方法,其中所述一种或更多种核酸的量指示移植状态或移植结果。

[0065] 31). 根据30)所述的方法,其中所述一种或更多种核酸的高于预定的阈值的量指示移植状态或移植结果。

[0066] 32).根据31)所述的方法,其中所述阈值是无移植排斥或其他病理的迹象的临床稳定的移植后患者的标准值。

[0067] 33).根据30)所述的方法,其中对于不同的移植结果或移植状态存在不同的预定的阈值。

[0068] 34).根据30)所述的方法,还包括根据免疫活性的评价治疗所述个体。

[0069] 在此说明书中提及的所有出版物和专利申请被通过引用并入本文,如同每个单独的出版物或专利申请被特别地且单独地指明通过引用并入一样。

[0070] 附图简述

[0071] 本发明的新颖特征在所附的权利要求中被特别地列出。通过参考以下列出使用本发明的原理的示例性实施方案的详细描述及其所附的附图,将获得对本发明的特征和优势的更好的理解:

[0072] 图1. 研究设计、读段统计和系统发育分布。A. 免疫抑制降低移植中排斥的风险,但增加感染的风险。B. 研究的设计。收集656个血浆样品,无细胞DNA被纯化并被测序至每个样品1.2Gbp的平均深度。C. 针对本研究的不同患者组部分,收集的样品的数目作为时间的函数。D. 研究群组中的患者的治疗方案,所有的患者用维持性免疫抑制治疗(对于成年心脏和肺移植接受者,基于他克莫司(TAC);以及对于儿科患者,环孢霉素(CYC))。用抗CMV预防、缬更昔洛韦(VAL)治疗CMV阳性(供体或接受者,CMV+)移植病例。在基于TAC的方案治疗的移植接受者的血液中测量他克莫司的平均水平(虚线是实际值,实线是窗平均过滤器)。E. 在过滤低质量和重复读段之后(平均86%,左)和在移出人类的和低复杂度读段之后(平均2%,右)剩余的读段的比例。F. 在移出人类读段之后,在不同系统分类水平上的相对基因组丰度(来自所有器官移植接受者的所有样品(n=656)的平均)。

[0073] 图2. 相对病毒基因组丰度作为药物剂量函数以及与健康参考的比较。A. 用免疫抑制剂他克莫司治疗的患者(47个患者,380个样品)的平均病毒组组成作为抗病毒药物剂量(缬更昔洛韦)和在血液中测量的他克莫司浓度的函数。为了解释病毒组组成对药物剂量的延长效果,药物剂量的数据被窗口平均过滤(窗口大小45天,参见图1中的C图)。当患者接受低剂量的免疫抑制剂和抗病毒药物时,疱疹病毒目和有尾噬菌体目在病毒组中占首要地位。相反的,当患者接受高剂量的这些药物时,指环病毒科在病毒组中占首要地位。B. 比较对应于以下的病毒组组成:健康参考(n=9)、移植后第1天具有低药物暴露的样品(n=13)、和相应于高药物暴露的样品(他克莫司 $\geq 9\text{ng/ml}$ 、缬更昔洛韦 $\geq 600\text{mg}$,n=68)。第一天样品的病毒组结构(I)和对一组健康个体测量的病毒组结构(H)不同于对相应于高药物剂量的样品(D)测量的指环病毒属占首要地位的分布。饼图显示了平均比例、盒形图中的p值基于曼-惠特尼检验。C. 所有样品、具有相同移植类型(心脏或肺)的患者中、受试者内(within subjects)、用相似的药物剂量(他克莫司水平 $\pm 0.5\text{ng/ml}$,缬更昔洛韦 $\pm 50\text{mg}$)治疗的患者和在一个月的时间跨度内从相同的受试者收集的样品的Bray-Curtis多样性。

[0074] 图3. 移植后微生物组组成的时间动态。A. 不同时间段的dsDNA和ssDNA病毒的相对丰度(所有样品的平均值)。在开始移植后药物治疗之后,ssDNA病毒的相对丰度快速增加。6个月后,观察到相反的趋势。B. 对于不同时间段,在系统分类的科和目水平上的病毒基因组丰度。指环病毒属的比例在移植后的最初几个月内快速扩展。疱疹病毒目、有尾噬菌体目和腺病毒科的比例在同一的时间段内下降。6个月后,观察到相反的趋势。C. 细菌门的相对丰

度的时间变化。与病毒丰度相比,不同细菌门的呈现 (representation) 在观察的移植后时期内相对不变。D. 为细菌属和病毒属的样品内 α 多样性的测量结果的香农熵作为时间的函数(按一个月的时间段分组的数据)。

[0075] 图4. 在抗病毒剂预防的存在和不存在下的病毒组组成和总病毒荷载。A. 绝对病毒载量作为时间的函数,该绝对病毒载量被测量为通过测序检测的病毒基因组拷贝数/人类基因组拷贝数。盒形图显示了不同的时间段,时间段的中心被标记在x轴上。对于所有的患者类别,在移植后的最初几周内总病毒载量增加(黑线为sigmoid拟合,载量的变化 7.4 ± 3)。B. 用免疫抑制剂和抗病毒药物二者治疗的CMV+病例的病毒载量和组成(78名患者,543个样品)。C. 仅用免疫抑制剂治疗的CMV-/-病例的病毒载量和组成(12名患者,75个样品)。

[0076] 图5. 罹患植入物排斥的患者中较低的指环病毒属荷载。A. 在罹患重度排斥事件的患者亚组(活组织检查等级 $\geq 2R/3A$,红色数据,20名患者,177个时间点)中和在移植后过程不罹患重度排斥的患者亚组(蓝色数据,40名患者,285个时间点)中的指环病毒属载量的时间依赖性。盒形图显示了不同的时间段,时间段的中心被标记在x轴上。实线是三次样条曲线(平滑参数0.75)。插图显示了排斥和感染的发生率与免疫活性的预期的负相关性的图。B. 在相同的时间点相对于对所有样品测量的平均载量的指环病毒属载量。将非排斥患者($N=208$)的时间归一化的载量与罹患中等排斥事件的患者(活组织检查等级 $1R$, $N=102$)和罹患重度排斥事件的患者(活组织检查等级 $\geq 2R/3A$, $N=22$)所测量的载量比较。p值反映出具有较大排斥风险的亚组的中值病毒载量较高的概率。p值通过对具有较大量的测量点的群体随机取样来计算。 N 次随机取样, $p = \text{sum}(\text{median}(A_{\text{排斥}}) > \text{median}(A_{\text{非排斥}})) / N$,其中 $N=10^4$ 并且 $A_{\text{排斥}}$ 和 $A_{\text{非排斥}}$ 分别是具有较大和较小的排斥和非排斥风险的群体的相对病毒载量。C. 在将患者分类成非排斥与重度排斥方面,相对的指环病毒属载量的表现的测试,接受者操作特征曲线,曲线下面积=0.72。

[0077] 图6. 基因组大小和命中统计、qPCR测定、以及在不同的系统分类水平上读段长度对测量的物种的相对丰度的影响。(A) 在具有1401个病毒基因组、32个真菌基因组和1980个细菌基因组的参考数据库中,基因组大小的分布。(B) 每一百万独特的测序分子中独特的blast命中的分布(命中的平均数目在x轴标注中说明)。C. 病毒、细菌和真菌的基因组当量(感染性病原物/二倍体人类)的分布(基因组当量的平均数目在x轴标注中说明)。(D) 将每一百万测序的总读段中发现的测序命中与利用qPCR检测到的病毒拷贝数比较。对于qPCR测定,从1ml血浆纯化DNA并且在 $100\mu\text{l}$ 体积中洗脱。(E) 对选择的病例测量CMV和细小病毒荷载。对所有样品测量的CMV病毒的最高载量(基因组当量,病毒/二倍体人类,G.E.)与临床上诊断为散播性CMV感染的两个病例相对应(a和b,阴影区域表示临床诊断的时间窗,*表示死亡的时间)。(c) 显示了罹患CMV病毒血症的儿科患者的时间轨迹。在一个儿科心脏移植患者中移植后立即检测细小病毒(d)。*(F) 在不同的系统分类水平上,读段长度对测量的物种相对丰度的影响($n=52$)。从50bp和100bp数据集得到的最丰富节点的丰度的斯皮尔曼样品-样品相关性,r值和p值,p(两样本曼-惠特尼U检验): $r=0.80, p=0.8$ (a); $r=0.86, p=0.4$ (b); $r=0.92, p=0.6$ (c); $r=0.84, p=0.5$ (d); $r=0.7, p=0.28$ (e); $r=0.99, p=1$ (f)。

[0078] 图7. 移植后针对成年心脏和肺移植患者的平均药物剂量和测量的水平以及药物剂量对病毒组组成的影响。(A-C) 对于本研究的成年心脏和肺移植患者,施用的缙更昔洛韦和强的松的平均剂量(A和C)以及测量的血液中的他克莫司的水平(B)。(D) 与病毒组分相

比,微生物组的细菌组分的组成对抗病毒剂和免疫抑制剂相对不敏感。(E)病毒组组成作为抗CMV药物(缙更昔洛韦)和免疫抑制剂(强的松)的剂量的函数。

[0079] 图8.移植后微生物组的细菌组分的时间动态,(A)细菌门的相对丰度作为时间的函数。(B)细菌属的相对丰度作为时间的函数。

[0080] 图9.不同患者类型的病毒组组成和总病毒荷载,(A和B)用免疫抑制剂和抗病毒剂两者治疗的CMV阳性成年心脏(A)、成年肺(B)、以及儿科心脏(C)移植接受者的病毒载量和组成。

[0081] 图10.CMV感染引发的同种异体植入物损伤。A.来自特定体液(BAL和血清)的CMV(人类疱疹病毒5,HHV-5)感染的临床报告与匹配临床测试数据(p值;曼-惠特尼U检验)的供体器官cfDNA信号之间相关性。B.对于具有多于一个临床阳性测试结果的感染,在感染的临床诊断和无细胞DNA水平(虚线指示bonferonni校正的显著性阈值)之间的相关性的P值。C.测试CMV阳性和CMV阴性患者中的CMV来源的无细胞DNA水平的表现的ROC曲线。

[0082] 图11.监测感染。A.与在测序中检测到的病毒感染的发生率比较的临床测试频率。B.相对于未测试的那些患者,被测试为特定的感染阳性(红色箭头)的患者的时间系列数据。(1)相对于未测试的患者(L34),在具有突显的临床阳性的L78中的腺病毒信号。(2)相对于未测试的患者(L57)中的持续信号,在具有一个阳性测试的L69中的多瘤病毒信号。(3)在具有突显的CMV(HHV-5)的阳性(红色)和阴性(黑色)测试二者的L58中的三种疱疹病毒感染(HHV-4、5和8)。(4)相对于在具有微孢子虫病的症状但未被测试的L78中观察到的信号,在具有显示的四个阳性测试的16中的微孢子虫信号。数据为相对于人的取对数的基因组当量,其中零值被替换为该测定的检测限(基因组当量的数目与被指定到靶基因组的单个序列读段相符)。

[0083] 发明详述

[0084] 将详细提及本发明的特别优选的实施方案。优选的实施方案的实施例被示例于下面的实施例部分。

[0085] 除非另外定义,本文使用的所有技术和科学术语具有与本发明所属领域技术人员通常理解的含义相同的含义。本文参考的所有专利和出版物被通过引用以其整体并入。

[0086] 当提供值的范围时,应理解为在该范围的上限和下限之间的每个中间值(到下限的单位的十分之一,除非上下文另外明确地说明),和在该表述的范围内的任何其他表述值或中间值被包含在本发明内。这些较小的范围的上限和下限可以单独地被包括在较小的范围内并且也被包含在本发明内,服从所表述的范围内的任何特别地排除性限制。当所表述的范围包括一个或两个限值时,排除了那些所包括的限值中的一个或两个的范围也被包含在本发明中。

[0087] 某些范围在本文中可以被前面加术语“约”的数值表示。在本文中使用术语“约”来提供对其后的精确数字以及接近于或近似于其后的数字的数字提供文字支持。在确定数字是否接近或近似于特别地引用的数字中,接近的或近似的未引用数字可以是在特别地引用的数字被展示的上下文中提供该特别地引用的数字的基本上等价值的数字。

[0088] 除非另外指明,本发明的实践采用本领域技术范围内的免疫学、生物化学、化学、分子生物学、微生物学、细胞生物学、基因组学和重组DNA的常规技术。参见Sambrook, Fritsch和Maniatis,MOLECULAR CLONING:A LABORATORY MANUAL,第二版(1989);CURRENT

PROTOCOLS IN MOLECULAR BIOLOGY (F.M.Ausubel等编辑., (1987));丛书METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J.MacPherson, B.D.Hames和G.R.Taylor编辑 (1995)), Harlow和Lane编辑 (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R.I.Freshney, 编辑 (1987))。

[0089] 本发明提供了用于分析个体中的微生物组或其个体组分的方法、装置、组合物和试剂盒。所述方法在感染的确定、微生物组结构的分析、个体的免疫活性的确定中 useful。在一些实施方案中,本发明提供了确定患者或受试者是否表现出免疫活性的方法。如本文所用的术语“个体”、“患者”或“受试者”包括人类和其他哺乳动物。

[0090] 定义

[0091] 如本文所用的术语状态或结果的“诊断 (diagnose)”或“诊断 (diagnosis)”包括预测或诊断状态或结果、确定对状态或结果的倾向、监测患者的治疗、诊断患者的治疗响应和状态或结果的预后、进程以及对特定治疗的响应。

[0092] 微生物群.如本文使用的,术语微生物群指在个体、通常地个体哺乳动物并且更通常地人类个体内存在的一组微生物。微生物群可包括:病原性物种;组成一种组织例如皮肤、口腔等的正常菌群,但在其他组织例如血液、肺等中是不期望的物种;在疾病不存在下发现的共生生物体;等。微生物组的一个亚组是病毒组,其组成微生物组的病毒组分。

[0093] 如本文所用的术语“微生物组组分”指个体菌株或物种。所述组分可以是病毒组分、细菌组分、真菌组分等。

[0094] 在健康的动物中,尽管内部组织例如脑、肌肉等一般被认为相对地无细菌物种,但是表面组织即皮肤和粘膜不断地与环境生物体接触并变得易于被微生物物种定殖。已知的或被认为在人类的任何解剖部位被发现的生物体的混合物被称为“原生 (indigenous) 微生物群”,包括原生微生物群的多个组分。除了原生微生物之外的是多种暂时性组分,诸如病原感染物或机会主义感染物。下文描述的生物体的参考序列是在例如Genbank数据库中公共可得并被已知的。

[0095] 人类的肠道微生物群以发现的在两个细菌门内的物种为主:拟杆菌门 (Bacteroidetes) 和厚壁菌门 (Firmicutes) 的成员占细菌群体的 >90%。在若干其他的门中放线菌门 (Actinobacteria) (例如双歧杆菌属 (Bifidobacterium) 的成员) 和变形菌门 (Proteobacteria) 是较不突出的代表。感兴趣的常见物种包括该群落的突出的或丰度较小的成员,并且可包括而限于:多形拟杆菌 (Bacteroides thetaiotaomicron); 粪拟杆菌 (Bacteroides caccae); 脆弱拟杆菌 (Bacteroides fragilis); 产黑色素拟杆菌 (Bacteroides melaninogenicus); 口腔拟杆菌 (Bacteroides oralis); 单形拟杆菌 (Bacteroides uniformis); 乳杆菌 (Lactobacillus); 产气荚膜梭菌 (Clostridium perfringens); 腐败梭菌 (Clostridium septicum); 破伤风梭菌 (Clostridium tetani); 两歧双歧杆菌 (Bifidobacterium bifidum); 金黄色葡萄球菌 (Staphylococcus aureus); 粪肠球菌 (Enterococcus faecalis); 大肠杆菌 (Escherichia coli); 肠炎沙门氏菌 (Salmonella enteritidis); 克雷伯氏菌 (Klebsiella) 属种; 肠杆菌 (Enterobacter) 属种; 奇异变形杆菌 (Proteus mirabilis); 铜绿假单胞菌 (Pseudomonas aeruginosa); 消化链球菌 (Peptostreptococcus) 属种; 消化球菌 (Peptococcus) 属种; Faecalibacterium 属种; 罗氏菌 (Roseburia) 属种; 瘤胃球菌 (Ruminococcus) 属种; Dorea 属种; Alistipes 属种等。

[0096] 在皮肤微生物组中大多数细菌属于四个不同的门:放线菌门、厚壁菌门、拟杆菌门和变形菌门。通常被认为是皮肤定殖物的微生物包括:放线菌门的棒状杆菌类(棒状杆菌属(*Corynebacterium*);丙酸菌属(*Propionibacterium*),诸如痤疮丙酸杆菌(*Propionibacterium acnes*);和短杆菌属(*Brevibacterium*))、微球菌属(*Micrococcus*)物种和葡萄球菌属(*Staphylococcus*)物种。最常被分离的真菌物种是马拉色氏霉菌(*Malassezia*)属种,其在分泌脂质的区域中特别普遍。蠕形螨属(*Demodex*)螨(诸如毛囊蠕形螨(*Demodex folliculorum*)和皮脂蠕形螨(*Demodex brevis*))也可存在。被认为在皮肤上生长的其他类型的真菌包括德巴利酵母(*Debaryomyces*)属种和隐球菌(*Cryptococcus*)属种。作为非共生体,烧伤往往被化脓性链球菌(*S. pyogenes*)、肠球菌(*Enterococcus*)属种或铜绿假单胞菌感染,并且还可被真菌和/或病毒感染。表皮葡萄球菌(*S. epidermidis*)是非常常见的皮肤共生体,但是其也是生长在留置医疗设备诸如导尿管或心脏瓣膜上的医院获得性感染的最常见原因。对于综述,参见*Nat Rev Microbiol.* (2011) 4月;9(4):244-53。

[0097] 病原性物种可以是细菌、病毒、原生动植物寄生虫、真菌物种等。细菌包括布鲁氏菌(*Brucella*)属种、密螺旋体(*Treponema*)属种、分枝杆菌(*Mycobacterium*)属种、李斯特菌(*Listeria*)属种、军团菌(*Legionella*)属种、螺杆菌(*Helicobacter*)属种、链球菌(*Streptococcus*)属种、奈瑟菌(*Neisseria*)属种、梭菌(*Clostridium*)属种、葡萄球菌(*Staphylococcus*)属种或杆菌(*Bacillus*)属种;包括但不限于梅毒螺旋体(*Treponema pallidum*)、结核分枝杆菌(*Mycobacterium tuberculosis*)、麻风分枝杆菌(*Mycobacterium leprae*)、单核细胞增生李斯特菌(*Listeria monocytogenes*)、嗜肺军团菌(*Legionella pneumophila*)、幽门螺杆菌(*Helicobacter pylori*)、肺炎链球菌(*Streptococcus pneumoniae*)、脑膜炎奈瑟菌(*Neisseria meningitidis*)、诺维氏梭菌(*Clostridium novyi*)、肉毒梭菌(*Clostridium botulinum*)、金黄色葡萄球菌(*Staphylococcus aureus*)、炭疽杆菌(*Bacillus anthracis*)等。

[0098] 寄生虫病原体包括毛滴虫属(*Trichomonas*)、弓形虫属(*Toxoplasma*)、贾第鞭毛虫属(*Giardia*)、隐孢子虫属(*Cryptosporidium*)、疟原虫属(*Plasmodium*)、利什曼虫属(*Leishmania*)、锥形虫属(*Trypanosoma*)、内变形虫属(*Entamoeba*)、血吸虫属(*Schistosoma*)、丝虫属(*Filariae*)、蛔虫属(*Ascaris*)、片血吸虫属(*Fasciola*);包括但不限于阴道毛滴虫(*Trichomonas vaginalis*)、刚地弓形虫(*Toxoplasma gondii*)、肠贾第虫(*Giardia intestinalis*)、*Cryptosporidium parva*、恶性疟原虫(*Plasmodium falciparum*)、克鲁斯氏锥体虫(*Trypanosoma cruzi*)、痢疾内变形虫(*Entamoeba histolytica*)、兰伯氏贾第虫(*Giardia lamblia*)、肝片吸虫(*Fasciola hepatica*)等。

[0099] 感染人类的病毒包括,例如,腺病毒相关病毒;爱知病毒(*Aichi virus*);澳大利亚蝙蝠狂犬病病毒;BK多瘤病毒;版纳病毒(*Banna virus*);巴马森林病毒(*Barmah forest virus*);布尼奥罗病毒(*Bunyamwera*);拉克罗斯布尼雅病毒(*Bunyavirus La Crosse*);北美雪兔布尼雅病毒(*Bunyavirus snowshoe hare*);猕猴疱疹病毒;金迪普拉病毒(*Chikungunya virus*);基孔肯雅病毒(*Chikungunya virus*);Cosavirus A;牛痘病毒;柯萨基病毒(*Coxsackievirus*);克里米亚-刚果出血热病毒(*Crimean-Congo hemorrhagic fever virus*)、登革病毒(*Dengue virus*);多理病毒(*Dugbe virus*);道格比病毒(*Dugbe virus*);杜梅海格病毒(*Duvenhage virus*);东部马脑炎病毒(*Eastern equine*

encephalitis virus);埃博拉病毒(Ebolavirus);艾柯病毒(Echovirus);脑心肌炎病毒; Epstein-Barr病毒;欧洲蝙蝠狂犬病毒;C型GB病毒/庚型肝炎病毒;汉坦病毒(Hantaan virus);亨得拉病毒(Hendra virus);甲型肝炎病毒;乙型肝炎病毒;丙型肝炎病毒;戊型肝炎病毒;丁型肝炎病毒;马痘病毒(Horsepox virus);人类腺病毒;人类星状病毒;人类冠状病毒;人类巨细胞病毒;人类肠道病毒68、70;人类疱疹病毒1;人类疱疹病毒2;人类疱疹病毒6;人类疱疹病毒7;人类疱疹病毒8;人类免疫缺陷病毒;人类乳头瘤病毒1;人类乳头瘤病毒2;人类乳头瘤病毒16、18;人类副流感病毒;人类细小病毒B19;人类呼吸道合胞体病毒;人类鼻病毒;人类SARS冠状病毒;人类spumaretrovirus;人类T淋巴细胞病毒;人类凸隆病毒(Human torovirus);甲型流感病毒;乙型流感病毒;丙型流感病毒;伊斯法罕病毒(Isfahan virus);JC多瘤病毒;日本脑炎病毒;呼宁砂粒样病毒(Junin arenavirus);KI多瘤病毒;昆津病毒(Kunjin virus);拉哥斯蝙蝠病毒(Lagos bat virus);维多利亚湖马尔堡病毒(Lake Victoria marburgvirus);兰加特病毒(Langat virus);拉沙病毒(Lassa virus);Lordsdale病毒;跳跃病病毒(Louping ill virus);淋巴细胞性脉络丛脑膜炎病毒;马丘波病毒(Machupo virus);马亚罗病毒(Mayaro virus);MERS冠状病毒;麻疹病毒;门戈脑心肌炎病毒(Mengo encephalomyocarditis virus);梅克尔细胞多瘤病毒(Merkel cell polyomavirus);莫科拉病毒(Mokola virus);接触传染性软疣病毒;猴痘病毒;腮腺炎病毒;墨累谷脑炎病毒;纽约病毒;尼帕病毒;诺沃克病毒;阿尼昂尼昂病毒(O'nyong-nyong virus);口疮病毒(Orf virus);奥罗普切病毒;皮秦德病毒(Pichinde virus);脊髓灰质炎病毒;庞塔托鲁静脉病毒(Punta toro phlebovirus);普马拉病毒(Puumala virus);狂犬病病毒;裂谷热病毒;Rosavirus A;罗斯河病毒;A组轮状病毒;B组轮状病毒;C组轮状病毒;风疹病毒;鹭山病毒;Salivirus A;白蛉热西西里病毒(Sandfly fever sicilian virus);札幌病毒;塞姆利基森林病毒(Semliki forest virus);汉城病毒;猴泡沫病毒;猴病毒5;辛德毕斯病毒(Sindbis virus);南安普顿病毒(Southampton virus);圣路易斯脑炎病毒(St.louis encephalitis virus);蜱传波瓦生病毒(Tick-borne powassan virus);细环病毒(Torque teno virus);托斯卡尼病毒(Toscana virus);尤尤库尼米病毒(Uukuniemi virus);牛痘病毒;水痘-带状疱疹病毒;天花病毒;委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus);水泡性口炎病毒;西方马脑炎病毒(Western equine encephalitis virus);Wu多瘤病毒(WU polyomavirus);西尼罗河病毒;亚巴猴肿瘤病毒;亚巴样病病毒;黄热病病毒;寨卡病毒(Zika virus);

[0100] 指环病毒科(Anelloviridae).指环病毒科由无包膜、环状、单链DNA病毒组成。已知指环病毒科的三个属感染人类,名为TTV、TTMDV和TTMV。

[0101] 细环病毒(TTV)是具有环状、反义基因组的无包膜单链DNA病毒。后来被命名为小细环样病毒(Torque Teno-like Mini Virus,TTMV)的较小的病毒也已被表征,并发现了具有在TTV和TTMV的基因组大小之间的基因组大小的第三种病毒,并且随后被命名为中细环样病毒(Torque Teno-like Midi Virus,TTMDV)。命名法的近期变化已将能够感染人类的三个指环病毒属(anelloviruses)分类成病毒的指环病毒科的甲型细环病毒属(Alphatorquevirus,TTV)、乙型细环病毒属(Betatorquevirus,TTMV)、和丙型细环病毒属(Gammatorquevirus,TTMDV)属。迄今,指环科病毒属仍被认为是待与人类疾病相关联的“孤儿”病毒。

[0102] 人类指环病毒属的基因组大小不同,TTV范围为从3.8-3.9kb,TTMDV3.2kb,TTMV范围为从2.8-2.9kb。指环病毒属的特有特征是在指环病毒属物种内和指环病毒属物种间均发现了极度的多样性;它们在核苷酸水平上可以表现出多达33%-50%的趋异度。不管核苷酸序列多样性如何,指环病毒属共有保守的基因组组成、转录谱、非编码GC富集区域和序列修饰,导致共有的病毒体结构和基因功能。

[0103] 指环病毒属感染在普通群体中高度流行。在日本的研究发现75%-100%的受测试的患者感染有三个人类指环病毒属中的至少一个,并且许多感染有多个物种。指环病毒属可感染年幼的儿童,记录的最早感染发生在生命的最初几个月内。这些病毒已经被发现于几乎每个被测试的身体部位、体液和组织,包括血浆、血清、外周血单核细胞(PBMC)、鼻咽抽出物、骨髓、唾液、母乳、粪便和多种组织,所述多种组织包括甲状腺、淋巴结、肺、肝、脾、胰腺和肾。指环病毒属的复制动力学实际上是未知的,因为无法在培养基中繁殖这些病毒。在肝脏细胞、骨髓细胞和循环的PBMC中已经描述了指示局部病毒复制的正链TTV DNA。

[0104] 指环病毒属主要通过粪-口传播扩散,尽管母婴传播和呼吸道传播也已被报道。关于TTV在脐带血样本中的存在存在矛盾的报道。

[0105] 指环病毒属的参考序列可以在Genbank访问,例如小细环病毒1型,登记号NC_014097.1;小细环病毒6型,登记号:NC_014095.1;中细环病毒2型,登记号:NC_014093.1;中细环病毒1型,登记号:NC_009225.1;细环病毒3型,登记号:NC_014081.1;细环病毒19型,登记号:NC_014078.1;小细环病毒8型,登记号:NC_014068.1。

[0106] 如本文使用的术语“抗生素”包括所有通常使用的抑菌抗生素和杀菌抗生素,通常是口服施用的那些抗生素。抗生素包括:氨基糖甙类,诸如阿米卡星、庆大霉素、卡那霉素、新霉素、链霉素和妥布霉素;头孢菌素,诸如头孢孟多、头孢唑啉、头孢氨苄、头孢来星、头孢利定、头孢噻吩、头孢匹林和头孢拉定;大环内酯类,诸如红霉素和醋竹桃霉素;青霉素类,诸如青霉素G、阿莫西林、氨苄西林、羧苄青霉素、氯唑西林、双氯西林、甲氧西林、萘夫西林、苯唑西林、非奈西林和替卡西林;多肽类抗生素,诸如杆菌肽、粘菌素甲磺酸(colistimethate)、粘菌素、多粘菌素B;四环素类,诸如氯四环素、地美环素、多西环素、美他环素、米诺环素、四环素和氧四环素;以及其他抗生素,诸如氯霉素、克林霉素、环丝氨酸、林可霉素、利福平、大观霉素、万古霉素和紫霉素。另外的抗生素被描述于“Remington's Pharmaceutical Sciences”第16版,(Mack Pub.Co.,1980),第1121-1178页。

[0107] 抗病毒剂.个体可接受抗病毒疗法,其将改变受该疗法影响的那些病毒的病毒载量。如此治疗的病毒感染的实例包括HIV、鲍恩样丘疹病、水痘、儿童HIV、人类牛痘、丙型肝炎、登革热、肠道病毒感染、疣状表皮发育不良、传染性红斑(第五疾病)、Buschke-Lowenstein巨大型尖锐湿疣、手足口病、单纯性疱疹、疱疹病毒6、带状疱疹、Kaposi水痘样疹、麻疹、挤奶者结节、接触传染性软疣、猴痘、口疮、玫瑰疹(Roseola Infantum)、风疹、天花、病毒性出血热、生殖器疣和非生殖器疣。

[0108] 抗病毒剂包括叠氮尿苷、安沙霉素、金刚烷胺、溴乙烯脱氧尿苷、氯乙烯脱氧尿苷、阿糖胞苷、去羟肌苷、脱氧野尻霉素(deoxynojirimycin)、双脱氧胞苷、双脱氧肌苷、双脱氧核苷、地昔洛韦、脱氧阿昔洛韦、碘苷(edoxuidine)、恩韦肟、非西他滨、膦甲酸盐(foscarnet)、非阿尿苷、氟胸苷、氟尿苷、金丝桃素、干扰素、白介素、羟乙基磺酸盐、奈韦拉平、喷他脒、利巴韦林、金刚乙胺、司他夫定、沙格莫丁、苏拉明、天花粉蛋白、三溴胸苷、三氯

胸苷、阿糖腺苷、齐多夫定 (zidoviridine)、扎西他滨、3-叠氮基-3-脱氧胸苷及其类似物、衍生物、药学上可接受的盐、酯、前体药物、共同药物 (codrug) 和被保护的形式。

[0109] 如本文所用的免疫抑制或免疫抑制方案指用剂治疗个体例如移植物接受者,以减少宿主免疫系统抵抗自身抗原或移植物的免疫响应。在下文更详细地描述了示例性免疫抑制方案。

[0110] 主要的免疫抑制剂包括钙调磷酸酶抑制剂,其与结合蛋白联合以抑制钙调磷酸酶活性,并且其包括例如他克莫司、环孢霉素A等。环孢霉素和他克莫司二者的水平必须被仔细监控。最初,水平可以维持在10-20ng/mL的范围,但3个月后,可使水平维持较低(5-10ng/mL)以减少肾毒性的风险。

[0111] 佐剂通常与钙调磷酸酶抑制剂联合,并且包括类固醇、咪唑硫嘌呤、霉酚酸酯和西罗莫司。感兴趣的方案包括钙调磷酸酶抑制剂和霉酚酸酯。佐剂的使用允许临床医生获得足够的免疫抑制,同时减少个体剂的剂量和毒性。在若干临床试验已经显示与咪唑硫嘌呤相比急性细胞排斥的发生率显著降低并且1年治疗失败减少后,肾移植接受者中的霉酚酸酯已被假定在免疫抑制中具有重要的作用。

[0112] 基于抗体的疗法可以使用单克隆抗体(例如鼠源单克隆抗体CD3)或多克隆抗体或抗CD25抗体(例如巴利昔单抗(basiliximab)、达利珠单抗(daclizumab)),并且在移植后早期时期被施用(最多8周)。基于抗体的疗法允许避免钙调磷酸酶抑制剂或钙调磷酸酶抑制剂剂量减少,可能减少肾毒性的风险。多克隆抗体和单克隆抗体的不利影响特征限制了其在一些患者中的使用。

[0113] 本文所用的术语“核酸”指包含两个或更多个核苷酸的多核苷酸。其可以是DNA或RNA。“变体”核酸是除了具有至少一个核苷酸修饰例如分别地缺失、插入或替换之外,具有与其原始核酸的序列一致的核苷酸序列的多核苷酸。变体可具有与原始核酸的核苷酸序列至少约80%、90%、95%或99%相同的核苷酸序列。

[0114] 在1948年首次在人类血浆中检测到循环DNA或无细胞DNA(Mandel,P.Métais,P.,C R Acad.Sci.Paris,142,241-243(1948))。自此,已在若干领域中建立了其与疾病的关联(Tong,Y.K.Lo,Y.M.,Clin Chim Acta,363,187-196(2006))。研究揭示了血液中的很多循环核酸由坏死的或凋亡的细胞产生(Giacona,M.B.,等Pancreas,17,89-97(1998))并且在诸如癌症的疾病中观察到来自凋亡的核酸的水平大大地提高(Giacona,M.B.等,Pancreas,17,89-97(1998);Fournie,G.J.,等,Cancer Lett,91,221-227(1995))。特别是对于癌症,其中循环DNA携带疾病的标志性标记,包括原癌基因中的突变、微卫星改变,并且对于某些癌症,血浆中的病毒基因组序列、DNA或RNA已经作为疾病的潜在生物标志物变的被越来越多地研究。例如,Diehl等最近阐明,对总的循环DNA中的低水平的循环肿瘤DNA的定量测定可充当与临床使用的标准生物标志物癌胚抗原相比更好的用于检测结肠直肠癌的复发的标志(Diehl,F.,等,Proc Natl Acad Sci,102,16368-16373(2005);Diehl,F.,等Nat Med,14,985-990(2008))。Maheswaran等报道了利用对血浆中的循环细胞基因分型,来检测肺癌患者中会影响药物治疗的表皮生长因子受体中的激活突变(Maheswaran,S.,等,N Engl J Med,359,366-377(2008))。这些结果共同地使血浆中游离的循环DNA成为癌症检测和治疗中的有用物质。循环DNA还可用于健康患者的胎儿诊断,在母体血液中循环的胎儿DNA作为性别、Rh D(rhesus D)状态、胎儿非整倍体、性连锁性紊乱的标志物。Fan等近期阐明了用于

通过对取自母体血液样品的无细胞DNA进行鸟枪法测序来检测胎儿非整倍体的策略,一种能够代替更有侵入性和风险的技术(诸如羊膜穿刺术或绒毛膜绒毛取样术)的方法(Fan, H.C., Blumenfeld, Y.J., Chitkara, U., Hudgins, L., Quake, S.R., Proc Natl Acad Sci, 105, 16266-16271 (2008))。

[0115] 如本文所用的术语“源自”指起源或来源,并且可包括天然地存在的、重组的、未纯化的或纯化的分子。源自原始核酸的核酸可以部分地或完全地包含原始核酸,并且可以是原始核酸的片段或变体。源自生物样品的核酸可以从该样品纯化的。

[0116] 根据本发明的方法中的“靶核酸”是待检测的核酸、DNA或RNA。源自生物体的靶核酸是具有源自所述生物体的序列的序列的多核苷酸并且对所述生物体是特定的。源自病原体的靶核酸指具有源自该特定病原体的多核苷酸序列的多核苷酸。

[0117] 在一些实施方案中,从样品获得少于1pg、5pg、10pg、20pg、30pg、40pg、50pg、100pg、200pg、500pg、1ng、5ng、10ng、20ng、30ng、40ng、50ng、100ng、200ng、500ng、1μg、5μg、10μg、20μg、30μg、40μg、50μg、100μg、200μg、500μg或1mg的核酸用于分析。在一些情况,从样品获得约1-5pg、5-10pg、10-100pg、100pg-1ng、1-5ng、5-10ng、10-100ng、100ng-1μg的核酸用于分析。

[0118] 在一些实施方案中,本文描述的方法被用于检测和/或定量相应于感兴趣的微生物或生物体的微生物组的核酸序列。本文描述的方法能够分析至少1、2、3、4、5、10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、 10^6 、 5×10^6 、 10^7 、 5×10^7 、 10^8 、 5×10^8 、 10^9 个或更多个序列读段。

[0119] 在一些实施方案中,本文描述的方法被用于例如通过确定来自微生物的mRNA相对于来自该微生物的DNA的存在来检测和/或定量基因表达。在一些实施方案中,本文描述的方法提供了多个基因的高度辨别的和定量的分析。本文描述的方法能够辨别并定量至少1、2、3、4、5、10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000个或更多个不同靶核酸的表达。

[0120] 包含无细胞核酸的样品从受试者获得。此类受试者可以是人类、驯养动物诸如牛、鸡、猪、马、兔、狗、猫、山羊等。在一些实施方案中,本发明中所用的细胞取自患者。样品包括,例如全血的非细胞级分、汗液、泪液、唾液、耳液(ear flow)、痰、淋巴液、骨髓悬浮液、淋巴液、尿液、唾液、精液、阴道液(vaginal fluid)、脑脊液、脑液、腹水、乳汁、呼吸道的分泌物、肠道或泌尿生殖道液体、组织或器官(例如肺)或已从器官(诸如乳腺、肺、肠、皮肤、子宫颈、前列腺、胰腺、心脏、肝和胃)移除的组织的灌洗物。此类样品可以通过离心、洗脱、密度梯度分离、单采血液成分术、亲和选择、淘选、FACS、用泛影钠离心等来分离。获得样品后,其可被直接使用、冷冻或在合适的培养基中维持短的时间段。

[0121] 可以使用本领域任何已知的技术获得血液样品,例如注射器或其他真空吸引器。使用前,血液样品可被任选地预处理或加工。样品,诸如血液样品可基于本文的任何方法和系统在从样品被获取时起4周、2周、1周、6天、5天、4天、3天、2天、1天、12小时、6小时、3小时、2小时或1小时内(或者如果冷冻,更长)被分析。当从受试者获得样品时(例如血液样品),量可以根据受试者尺寸和被筛选的条件而变化。在一些实施方案中,获得至少10ml、5ml、1ml、0.5ml、250μL、200μL、150μL、100μL、50μL、40μL、30μL、20μL、10μL、9μL、8μL、7μL、6μL、5μL、4μ

L、3 μ L、2 μ L或1 μ L的样品。在一些实施方案中,获得1-50 μ L、2-40 μ L、3-30 μ L或4-20 μ L的样品。在一些实施方案中,获得多于5 μ L、10 μ L、15 μ L、20 μ L、25 μ L、30 μ L、35 μ L、40 μ L、45 μ L、50 μ L、55 μ L、60 μ L、65 μ L、70 μ L、75 μ L、80 μ L、85 μ L、90 μ L、95 μ L或100 μ L的样品。

[0122] 无细胞级分优选地是血清或血浆。本文所用的术语生物样品的“无细胞级分”指生物样品的基本上无细胞的级分。本文所用的术语“基本上无细胞”指来自生物样品的制品包含少于约20,000个细胞/ml、优选地少于约2,000个细胞/ml、更优选地少于约200个细胞/ml、最优选地少于约20个细胞/ml。与某些现有技术方法相反,基因组DNA未被从非细胞样品排除,并且通常包含存在于样品中的从约50%到约90%的核酸。

[0123] 本发明的方法还可包括从生物样品制备无细胞级分。可以使用本领域已知的常规技术来制备无细胞级分。例如,血液样品的无细胞级分可通过将血液样品以约200-20,000g、优选地约200-10,000g、更优选地200-5,000g、最优选地约350-4,500g的低速离心持续约3-30分钟、优选地约3-15分钟、更优选地约3-10分钟、最优选地约3-5分钟来获得。生物样品可以通过超滤获得,以将细胞及其碎片与包含可溶性DNA或RNA的无细胞级分分离。常规地,使用0.22 μ m膜滤器进行超滤。

[0124] 本发明的方法还可包括浓缩(或富集)生物样品的无细胞级分中的靶核酸。可使用本领域已知的常规技术来浓缩靶核酸,所述常规技术诸如在高盐浓度的存在下的固相吸收、通过苯酚-氯仿有机萃取后用乙醇或异丙醇沉淀、或在高盐浓度或70-80%乙醇或异丙醇的存在下直接沉淀。浓缩的靶核酸可比无细胞级分中的靶核酸浓缩至少约2、5、10、20或100倍。根据本发明的方法,靶核酸无论是否被浓缩均可被用于扩增。

[0125] 在一些实施方案中,本发明提供了用于诊断或预测移植排斥的方法。术语“移植排斥”包括急性移植物排斥和慢性移植物排斥两者。急性排斥或AR”是当移植的组织是免疫学上的异物时,被组织移植物接受者的免疫系统的排斥。急性排斥的特征在于被移植的组织被接受者的免疫细胞侵袭,接受者的免疫细胞发挥其效应器功能并破坏被移植的组织。急性排斥的起始是快速的,并且通常在移植手术后几周内发生于人中。通常,急性排斥可用免疫抑制性药物诸如雷帕霉素、环孢霉素A、抗CD40L单克隆抗体等来抑制(inhibited)或阻抑(suppressed)。

[0126] “慢性移植物排斥或CR”通常在植入后几个月至几年内,甚至在急性排斥的成功免疫抑制的存在下发生于人中。纤维化是所有类型的器官移植的慢性排斥中的共同因素。慢性排斥通常可通过特定器官特有的一系列特定紊乱来描述。例如,在肺移植中,此类紊乱包括气道的纤维增生性破坏(闭塞性细支气管炎);在心脏移植或心脏组织的移植诸如瓣膜置换中,此类紊乱包括纤维化动脉粥样硬化;在肾移植中,此类紊乱包括梗阻性肾病、肾硬化、小管间质性肾病;以及在肝移植中,此类紊乱包括胆管消失综合征。慢性排斥还可以以缺血性损伤、被移植组织的去神经、与免疫抑制性药物有关的高血脂和高血压为特征。

[0127] 在一些实施方案中,本发明还包括用于确定免疫抑制方案对接受移植物如同种异体植入物的受试者的效果的方法。

[0128] 本发明的某些实施方案提供了预测已接受移植物的受试者中移植物存活的方法。本发明提供了诊断或预测移植患者或受试者中的移植物是将存活还是丢失的方法。在某些实施方案中,本发明提供了诊断或预测存在植入物长期存活的方法。植入物“长期”存活意指尽管发生一次或更多次前期的急性排斥发作,植入物存活持续在当前取样后至少约5年。

在某些实施方案中,移植物存活针对已发生至少一次急性排斥的发作的患者确定。如此,这些实施方案提供了确定或预测在急性排斥后移植物存活的方法。在某些实施方案中,移植物存活在移植疗法例如免疫抑制疗法的上下文中被确定或预测,其中免疫抑制疗法是本领域已知的。在又其他的实施方案中,提供了确定急性排斥的等级和/或严重度(并且不仅是其的存在)的方法。

[0129] 如在移植领域已知的,移植器官、组织或细胞可以是同种异体的或异种的,因此植入物可以是同种异体植入物或异种植入物。通过主题方法检测或鉴定的植入物耐受表型的特征是:它是在无免疫抑制疗法下出现的表型,即它存在于不经历免疫抑制疗法的宿主中,以使得免疫抑制剂不被施用至该宿主。移植植入物可以是任何实体器官和皮肤移植物。可通过本文描述的方法分析的器官移植的实例包括但不限于肾移植、胰腺移植、肝移植、心脏移植、肺移植、肠移植、肾移植后胰腺移植以及胰腺-肾同时移植。

[0130] 微生物组检测和分析

[0131] 本发明的方法包括对来自个体的无细胞核酸样品高通量测序,随后进行生物信息学分析以确定微生物序列的存在和发生率,所述序列可来自原生生物体,例如内脏、皮肤等的正常微生物组,或可以是非原生的,例如机会感染(opportunistic infection)、病理性感染等。可对全部的基因组或对其组分例如病毒组、细菌微生物组、真菌微生物组、原生生物微生物组等进行分析。核酸的实例包括但不限于双链DNA、单链DNA、单链DNA发夹、DNA/RNA杂交物、RNA(例如mRNA或miRNA)和RNA发夹。在一些实施方案中,核酸是DNA。在一些实施方案中,核酸是RNA。例如,无细胞RNA或DNA存在于人类血浆中。

[0132] 尽管可能存在其中足够的无细胞核酸是可得的并且可被直接测序的例子,但对微生物组核酸基因分型、和/或检测、鉴定和/或定量微生物组特异性核酸通常包括扩增样品的初始步骤。当核酸是RNA时,扩增步骤可在逆转录酶反应以将RNA转换成DNA后进行。优选的扩增是无偏倚的,即用于扩增的引物是通用引物,或将接头连接到被分析的核酸并且扩增引物是接头特异性的。PCR技术的实例包括但不限于,热启动PCR、巢式PCR、原位聚合酶克隆(polony)PCR、原位滚环扩增(RCA)、桥式PCR、picotiter PCR和乳液PCR。其他合适的扩增方法包括连接酶链式反应(LCR)、转录扩增、自主性序列复制、靶多核苷酸序列的选择性扩增、共有序列引发的聚合酶链式反应(CP-PCR)、随机引发的聚合酶链式反应(AP-PCR)、简并寡核苷酸引发的PCR(DOP-PCR)和基于核酸的序列扩增(NABSA)。其他可用于扩增特定的多态性基因座的扩增方法包括在美国专利号5,242,794、5,494,810、4,988,617和6,582,938中描述的方法。

[0133] 扩增后,扩增的核酸被测序。可使用高通量系统完成测序,一些高通量系统允许紧接着被测序的核苷酸被并入生长的链或在被测序的核苷酸被并入生长的链之后立刻检测该被测序的核苷酸,即序列的检测是实时的或基本实时的。在一些情况,高通量测序每小时产生至少1,000、至少5,000、至少10,000、至少20,000、至少30,000、至少40,000、至少50,000、至少100,000或至少500,000个序列读段;每个读段为至少50个碱基/读段、至少60个碱基/读段、至少70个碱基/读段、至少80个碱基/读段、至少90个碱基/读段、至少100个碱基/读段、至少120个碱基/读段或至少150个碱基/读段。可以使用本文描述的核酸诸如基因组DNA、源自RNA转录本的cDNA或RNA作为模板进行测序。

[0134] 在一些实施方案中,高通量测序涉及使用通过Helicos BioSciences

Corporation (Cambridge, Massachusetts) 可得的技术诸如通过合成的单分子测序 (SMSS) 方法。SMSS 是独特的, 因为其允许对整个基因组测序而不需要预扩增步骤。因此, 减少了核酸的测量中的失真和非线性。在 US 公布申请号 2006002471I、20060024678、20060012793、20060012784 和 20050100932 中部分地描述了 SMSS。

[0135] 在一些实施方案中, 高通量测序涉及使用通过 454 Lifesciences, Inc. (Branford, Connecticut) 可得的技术, 诸如 Pico Titer Plate 装置, 其包括传输由测序反应产生的化学发光信号以被仪器中的 CCD 摄像机记录的光纤板。光纤的使用允许在 4.5 小时内检测最少 2 千万个碱基。

[0136] 在 Marguiles, M., 等 “Genome sequencing in microfabricated high-density picolitre reactors”, Nature, doi:10.1038/nature03959 中以及在美国公布申请号 20020012930、20030058629、20030100102、20030148344、20040248161、20050079510、20050124022 和 20060078909 中描述了用于在光纤检测之后利用珠扩增的方法。

[0137] 在一些实施方案中, 使用 Clonal Single Molecule Array (Solexa, Inc.) 或利用可逆终止物化学的合成测序 (SBS) 进行高通量测序。这些技术被部分地描述于美国专利号 6,969,488、6,897,023、6,833,246、6,787,308, 和美国公布申请号 200401061 30、20030064398、20030022207, 以及 Constans, A., The Scientist 2003, 17 (13):36 中。

[0138] 在本方面的一些实施方案中, 可使用 AnyDot. 芯片 (Genovox, Germany) 进行 RNA 或 DNA 的高通量测量, AnyDot. 芯片允许监测生物过程 (例如, miRNA 表达或等位基因变异性 (SNP 检测))。特别地, AnyDot 芯片允许将核苷酸荧光信号检测增强 $10 \times - 50 \times$ 。在国际公布申请号 WO 02088382、WO 03020968、WO 0303 1947、WO 2005044836、PCTEP 05105657、PCMEP05105655; 以及德国专利申请号 DE 101 49 786、DE 102 14 395、DE 103 56 837、DE 10 2004 009 704、DE 10 2004 025 696、DE 10 2004 025 746、DE 10 2004 025 694、DE 10 2004 025 695、DE 10 2004 025 744、DE 10 2004 025 745、和 DE 10 2005 012 301 中部分地描述了 AnyDot 芯片和使用其的方法。

[0139] 其他高通量测序系统包括在 Venter, J., 等 Science 16 February 2001; Adams, M. 等 Science 24 March 2000; 和 M. J. Levene, 等 Science 299:682-686, 2003 年 1 月; 和美国公布申请号 20030044781 和 2006/0078937 中公开的那些系统。所有此类系统涉及通过经由聚合反应的碱基暂时添加来对具有多个碱基的靶核酸分子测序, 碱基暂时添加被在核酸的分子上测量, 即核酸聚合酶对待测序的模板核酸分子的活性被实时追踪。然后可通过鉴定在碱基添加的顺序中在每一步哪种碱基被通过核酸聚合酶的催化活性并入靶核酸的正在生长的互补链, 来推断序列。靶核酸分子复合物上的聚合酶被提供在适合于沿着靶核酸分子移动并在活性位点延伸寡核苷酸引物的位置。在邻近活性位点处提供多种标记类型的核苷酸类似物, 每种可区分类型的核苷酸类似物与靶核酸序列中的不同核苷酸互补。通过使用聚合酶在活性位点处将核苷酸类似物添加至核酸链, 正在生长的核酸链被延伸, 其中被添加的核苷酸类似物与靶核酸的活性位点处的核苷酸互补。作为聚合步骤的结果的被添加至寡核苷酸引物的核苷酸类似物被鉴定。重复提供带标记的核苷酸类似物、使正在生长的核酸链聚合、以及鉴定添加的核苷酸类似物的步骤, 以使得核酸链被进一步延伸并且靶核酸的序列被确定。

[0140] 在一些实施方案中, 进行鸟枪法测序。在鸟枪法测序中, DNA 被随机打断成许多小

片段,其被利用链终止法测序以获得读段。通过进行若干轮的该片段化和测序,获得靶DNA的多个重叠读段。然后计算机程序使用不同读段的重叠末端来将其组装成连续的序列。

[0141] 在一些实施方案中,本发明提供了用于利用测序检测并定量微生物序列的方法。在该情况中,技术人员可评价检测的灵敏度。灵敏度存在两个组成部分:(i)被分析的分子的数目(测序的深度)和(ii)测序过程的出错率。对于测序的深度,个体间差异的惯常评估值是每一千个约一个碱基不同。目前,测序仪诸如Illumina基因组分析仪具有超过36个碱基对的读长。尽管血液中宿主DNA的比例可以是根据个体的状态而变化的,技术人员可取90%作为基线评估值。以该比例的供体DNA,被分析的分子的约十分之一将是微生物的。在该基因组分析仪上,技术人员可获得每分析通道约1千万个分子,并且每次仪器运行存在8个分析通道。因此,如果每个通道装载一个样品,技术人员应该能够检测约 10^6 个分子,这些分子可被鉴定为是微生物的并且可提供关于微生物组的状态的信息。通过对更多的分子测序,即使用更多的通道,可以简单地获得更高的灵敏度。

[0142] 测序出错率也影响该技术的灵敏度。碱基替换的典型测序出错率在平台间变化,但在0.5-1.5%之间。这将0.16%的灵敏度的潜在限制推至0.50%。然而,通过对样品模板重测序多次系统地降低测序出错率是可能的,如已被Helicos BioSciences(Harris,T.D.,等Science,320,106-109(2008))阐明的。单次应用重测序会降低预期的出错率。

[0143] 测序后,序列的数据集被上传到数据处理器用于生物信息分析以从分析扣除宿主序列,即人类、猫、狗等;并例如通过将映射到微生物参考序列的序列的覆盖度与宿主参考序列的覆盖度比较来确定微生物序列的存在和发生率。宿主序列的扣除可包括鉴定参考宿主序列和掩蔽存在于参考宿主基因组中的微生物序列或微生物模拟序列的步骤。类似地,通过与微生物参考序列比较确定微生物序列的存在可包括鉴定参考微生物序列和掩蔽存在于参考微生物基因组中的宿主序列或宿主模拟序列的步骤。

[0144] 任选地清理数据集以检查序列质量,去除测序仪特异性核苷酸(接头序列)的残留物,并合并成对的重叠的末端读段以建立具有较少读取错误的较高质量的共有序列。重复的序列被鉴定为具有相同起始位点和长度的那些序列,并且可将重复(duplicates)从分析移出。

[0145] 本发明的重要特征是从分析扣除人类序列。由于扩增/测序步骤是无偏倚的,样品中的优势序列将是宿主序列。可以以若干种方式优化扣除过程以改进该过程的速度和准确度,例如通过进行多次扣除,其中初始对齐被设为粗过滤器,即用快速对齐器,并且用细过滤器即敏感的对齐器进行另外的对齐。

[0146] 读段的数据库首先与人类参考基因组对齐,以在生物信息学上扣除宿主DNA,人类参考基因组包括但不限于Genbank hg19参考序列。每个序列与人类参考序列中的最佳匹配序列对齐。被明确鉴定为人类的序列被在生物信息学上从分析移出。

[0147] 参考人类序列还可通过添加具有高命中率的重叠群来优化,所述重叠群包括但不限于存在于基因组中的在参考数据库中未被良好地代表的高度重复的序列。当使用包括一大组人类序列的数据库,例如整个NCBI NT数据库时,已经观察到在未与hg19对齐的读段中,大量的读段在流水线的稍后阶段被鉴定为人类的。在分析中较早地移出这些读段可通过建立扩展的人类参考来进行。该参考通过鉴定不为参考的人类序列数据库中,例如NCBI NT数据库,的人类重叠群,来创建,所述人类重叠群在初始的人类读段扣除之后具有高的覆

盖度。将那些重叠群添加至人类参考以创建更全面的参考集。另外,来自群组研究的新组装的人类重叠群可被用作对人类来源的读段的另外的掩蔽。

[0148] 人类基因组参考序列的包含非人类序列例如被整合入参考样品的基因组的病毒序列和细菌序列的区域可以被掩蔽(masked)。例如,Epstein-Barr病毒(EBV)具有其约80%的基因组被整合入hg19。

[0149] 然后将被鉴定为非人类的序列读段与微生物参考序列的核苷酸数据库比对。所述数据库可被选择为已知与宿主相关的那些微生物序列,例如人类共生微生物或病原性微生物的集合。

[0150] 微生物数据库可被优化以掩蔽或移出污染序列。例如已经观察到许多公共数据库条目包括并非源自微生物的人工序列,例如引物序列、宿主序列和其他污染物。对数据库进行初始比对或多重比对是期望的。当比对多个样品时,在读段覆盖方面显示出不齐的区域可以作为矫作物被掩蔽或移出。此类不齐的覆盖度的检测可通过多种度量标准来完成,诸如在特定核苷酸的覆盖度和于其中发现该核苷酸的整个重叠群的平均覆盖度之间的比率。通常,被呈现为比参考序列的平均覆盖度大约 $5\times$ 、约 $10\times$ 、约 $25\times$ 、约 $50\times$ 、约 $100\times$ 的序列是人工的。可选地,考虑到重叠群的整体覆盖度,可以应用二项测试以提供覆盖的每碱基可能性。从参考数据库移出污染物序列允许微生物的准确鉴定。本发明的方法的益处是利用样品的比对改进了数据库,例如在商业或临床使用之前,数据库可与1、10、20、50、100或更多个样品比对以改进数据库。

[0151] 每个高置信度读段可以与给定的微生物数据库中的多种生物体比对。为了基于该可能的映射冗余准确地对生物体丰度赋值,选择算法并利用其来计算最可能的生物体(例如参见Lindner等Nucl. Acids Res. (2013) 41(1):e10)。例如GRAMMy或GASiC算法可被用于计算给定的读段最可能来自的生物体。这些数据提供了关于微生物在无细胞核酸样品中的存在的信息。

[0152] 可根据本领域认可的方法进行与宿主序列或微生物序列的比对并指定到宿主序列或微生物序列。例如,如果遍布读段的长度存在不多于1个错配、不多于2个错配、不多于3个错配、不多于4个错配、不多于5个错配等,则50nt的读段可被指定为匹配给定的基因组。通常使用商用算法来比对和鉴定。此类比对算法的非限制性实例是bowtie2程序(Johns Hopkins University)。例如,可基于预期的比对速度选择末端到末端模式的预设选项,

[0153] --非常-快等同于:-D 5-R 1-N 0-L 22-i S,0,2.50

[0154] --快等同于:-D 10-R 2-N 0-L 22-i S,0,2.50

[0155] --灵敏等同于:-D 15-R 2-L 22-i S,1,1.15

[0156] --非常--灵敏等同于:-D 20-R 3-N 0-L 20-i S,1,0.50

[0157] 可比较的设定可被用于其他的比对算法或软件包。

[0158] 在确定无细胞核酸样品中生物体的发生率中,读段到生物体(即宿主或微生物组分)的这些指定然后被汇总并被用于计算被指定到给定样品中的每种微生物的读段的估计数目。该分析使对微生物基因组大小的计数归一化,以提供微生物覆盖度的计算值。将每个微生物的归一化的覆盖度与同一样品中的宿主序列覆盖度比较以解释样品间在测序深度上的差异。

[0159] 最终的确定提供由样品中的序列表示的微生物的数据集和那些微生物的发生率。

这些数据任选地被合计并展示用于即时可视化,例如以提供给个体或医疗保健提供者的报告的形式;或以带有超链接数据的浏览器格式书写的形式。对于每个样品或样品的群组,覆盖度评价可来自样品的元数据来合计并且被分类成表和图。

[0160] 任选地,被过滤掉的宿主序列可被用于其他目的,例如个性化医疗。例如,人类基因组中的某些SNP可允许医生鉴定给定患者的药物敏感性。人类来源的序列可揭示病毒到宿主基因组的整合(例如EBV、HPV、多瘤病毒)或被用于协同临床应用(例如,无细胞肿瘤DNA可被用于在患者中与感染监测平行地监测癌症进程,所述患者由于化疗而高度易受感染)。

[0161] 在一些实施方案中,无细胞核酸的分析被用于计算病原性评分,其中病原性评分是概括生物体的总体病原性以易于被例如医疗从业者理解的数值或字母值。存在于微生物组的不同微生物可被赋值为不同的评分。最终的“病原性评分”是许多不同因素的组合,并且通常被提供为例如范围为从0-1、0-10或0-100的任意单位;根据所有针对感兴趣的微生物观察到的病原性评分被提供为一个百分数,等。特定的参数和那些参数的权重可例如通过将函数与观察到的疾病严重度拟合来实验性地确定,或通过设定不同参数或标准的重要度来人工地确定。

[0162] 用于计算病原性评分的相关因素可包括但不限于:相对于参考受试者或受试者组(例如测试群体、已知感染、已知的未感染个体等)中的微生物的丰度的微生物丰度(例如如通过读段相对于人类读段的数目计算的)。在微生物基因组中发现的特定突变可参考与所述微生物相关的毒性、病原性、抗生素抗性等的数据库来做出,并且包括但不限于SNP、插入缺失、质粒等。特定微生物的同时存在包括但不限于特定比例和组的微生物。某些序列的表达例如mRNA的检测可与病原性评分相关,例如可提供微生物是否在活跃地复制或是在潜伏等的信息。地理特征也可被包括,其中地理(geography)指示暴露于感兴趣的微生物,例如宿主的旅行史;与被感染个体的相互作用等。

[0163] 还提供了用于实践一种或更多种以上描述的方法的试剂及其试剂盒。主体试剂及其试剂盒可很大地变化。感兴趣的试剂包括被特定地设计以用于在以上描述的产生中使用的试剂:(i)微生物组和个体的特征分析;(ii)鉴定微生物组特征;并且(iii)对来自从个体获得的样品中的微生物组的一种或更多种核酸检测和/或定量。所述试剂盒可包括使用本文所述的方法诸如PCR和测序进行核酸提取和/或核酸检测所需的试剂。所述试剂盒还包括用于数据分析的软件包,其可包含用于与测试配制文件(profile)比较的参考配制文件,并且特别地可包含如以上描述的优化的参考数据库。所述试剂盒可包含诸如缓冲液和H₂O的试剂。

[0164] 此类试剂盒还可包括诸如科技参考文献、包装插入材料、临床试验结果、和/或这些的概述等的信息,其指示或表明组合物的活性和/或优势,和/或其描述剂量、施用、副作用、药物相互作用或其他对医疗保健从业者有用的信息。此类试剂盒还可包括进入数据库的说明。此类信息可基于多种研究的结果,例如使用试验动物的涉及体内模型的研究和基于人类临床试验的研究。可向医疗保健提供者包括医师、护士、药剂师、处方官员(formulary officials)等提供、销售(marketed)和/或推销(promoted)本文描述的试剂盒。在一些实施方案中,还可直接向消费者销售试剂盒。

[0165] 上文的任何方法可由计算机程序产品进行,所述计算机程序产品包含记录在计算机可读介质上的计算机可执行逻辑。例如,计算机程序可执行以下功能中的一些或所有:

(i) 控制从样品分离核酸, (ii) 从样品预扩增核酸, (iii) 扩增、测序或排列样品中的特定区域, (iv) 鉴定并定量样品中的微生物序列, (v) 将从样品检测到的关于微生物的存在或发生率的数据与预定的阈值比较, (vi) 确定感染、微生物组健康、免疫活性状态或结果, (vi) 说明关于感染、微生物组健康、免疫活性等的样品状态。

[0166] 计算机可执行逻辑可在任何计算机中工作, 其可以是多种类型的通用计算机中的任一种, 所述通用计算机诸如个人计算机、网络服务器、工作站、或现在或以后开发的其他计算机平台。在一些实施方案中, 描述了计算机程序产品, 所述计算机程序产品包含具有计算机可执行逻辑 (计算机软件程序, 包括程序代码) 储存于其中的计算机可用介质。计算机可执行逻辑可以被处理器执行, 导致处理器进行本文描述的功能。在其他实施方案中, 一些功能使用例如硬件状态机主要在硬件中实施。实施硬件状态机以进行如本文描述的功能对于相关领域的技术人员将是明显的。

[0167] 所述程序可提供通过访问反映微生物组和个体的特征, 和/或在个体的循环中来自微生物组的一种或更多种核酸的定量的数据来评价个体中的微生物状态的方法。

[0168] 在一个实施方案中, 执行本发明的计算机逻辑的计算机还可包括数字输入设备诸如扫描仪。所述数字输入设备可提供关于核酸的信息, 例如存在或发生率。

[0169] 在一些实施方案中, 本发明提供了计算机可读介质, 所述计算机可读介质包含记录于其上的一组指令以导致计算机进行以下步骤: (i) 接收来自在样品中检测的一个或更多个核酸的数据; 和 (ii) 基于微生物组定量诊断或预测状态。

[0170] 还提供了微生物参考序列的数据库和人类参考序列的数据库。此类数据库通常将包含如以上描述的优化的数据集。

[0171] 在一些实施方案中, 本发明的方法提供了关于感染的个体状态。在一些此类实施方案中, 微生物感染是病原体, 其中病原体序列的任何存在指示临床上的相关感染。在其他的实施方案中, 发生率指示微生物载量, 其中预设的水平指示临床相关性。在一些此类实施方案中, 个体用或考虑用抗微生物疗法例如抗生素、被动或主动免疫疗法、抗病毒剂等来治疗。个体可在疗法之前、疗法期间和疗法之后被测试。

[0172] 微生物感染还可通过共生生物的载量来指示, 其中血液样品中的共生生物的水平指示肠道健康, 例如肠腔损坏。

[0173] 可单独地或相对于微生物DNA对微生物RNA进行比较, 其中微生物序列的RNA的过量, 例如约 $5\times$ 、 $10\times$ 、 $15\times$ 、 $20\times$ 、 $25\times$ 的微生物DNA的覆盖度指示活动性感染。在一些实施方案中, 如此分析的微生物是能够潜伏性感染的微生物, 例如疱疹病毒、肝炎病毒等。

[0174] 在其他的实施方案中, 微生物组的整体评价是感兴趣的, 其中微生物的种类的相对存在或发生率是感兴趣的。在本领域已知膳食和用药物例如他丁类、抗生素、免疫抑制剂等治疗可影响微生物组的整体健康, 并且因此, 确定微生物组的组成是感兴趣的。

[0175] 在一些实施方案中, 在来自微生物组的所述一种或更多种核酸的量上的时间差异可被用于监测抗微生物治疗的效果或被用于选择治疗。例如, 可在治疗之前或之后确定来自微生物组的一种或更多种核酸的量。来自微生物的一种或更多种核酸在治疗之后的减少可指示治疗是成功的。另外, 来自微生物组的一种或更多种核酸的量可被用于在多种治疗 (例如不同强度的治疗) 之间选择。

[0176] 在本发明的一方面, 提供了用于诊断或预测接受免疫抑制方案的受试者中的免疫

活性、移植状态或结果的方法。免疫抑制后,如上文描述的样品可从患者得到并被分析一个或多个微生物组包括病毒组核酸的存在或不存在。在一些实施方案中,样品是血液、血浆、血清或尿液。可随时间推移监测微生物核酸的比例和/或量,并且该比例的增加可被用于确定免疫活性。载量的定量可通过本领域已知的任何合适的方法确定,包括本文描述的那些方法,诸如测序、核酸阵列或PCR。

[0177] 在一些实施方案中,来自被免疫抑制的接受者的样品中的一种或更多种微生物组核酸的量被用于确定移植状态或结果。因此,在一些实施方案中,本发明的方法还包括对来自微生物组的一种或更多种核酸定量。在一些实施方案中,来自供体样品的一种或更多种核酸的量被确定为样品中的总核酸的百分比。在一些实施方案中,来自供体样品的一种或更多种核酸的量被确定为样品中的总核酸的比率。在一些实施方案中,来自供体样品的一种或更多种核酸的量被确定为与样品中的一种或更多种参考核酸相比的比率或百分比。例如,来自微生物组的一种或更多种核酸的量可被确定为样品中的总核酸的10%。可选地,来自微生物组的一种或更多种核酸的量可以是与样品中的总核酸相比比率为1:10。另外,来自微生物组的一种或更多种核酸的量可被确定为参考基因诸如 β 球蛋白的10%或与参考基因诸如 β 球蛋白的比率为1:10。在一些实施方案中,来自微生物组的一种或更多种核酸的量可被确定为浓度。例如,来自供体样品的一种或更多种核酸的量可被确定为 $1\mu\text{g}/\text{mL}$ 。

[0178] 在一些实施方案中,来自微生物组的一种或更多种核酸的高于预定的阈值的量指示免疫活性的状态。例如,无移植排斥或其他病理的迹象的临床稳定的患者的标准值可被确定。来自微生物组的一种或更多种核酸的低于临床上稳定的移植后患者的标准值的量的增加可指示稳定的结果。另一方面,来自微生物组的一种或更多种核酸的高于或为临床上稳定的移植后患者的标准值的量可指示增加的免疫活性和植入排斥的风险。

[0179] 在一些实施方案中,不同的预定的阈值指示不同的移植结果或状态。例如,如上文讨论的,来自微生物组的一种或更多种核酸的高于临床上稳定的移植后患者的标准值的量的增加可指示诸如移植物排斥或移植物损伤的移植状态或结果的变化。但是,来自微生物组的一种或更多种核酸的高于临床上稳定的移植后患者的标准值但低于预定的阈值水平的量的增加可指示较不严重的状况诸如病毒感染而不是移植物排斥。来自微生物组一种或更多种核酸的高于较高的阈值的量的增加可指示移植物排斥。

[0180] 在一些实施方案中,来自微生物组的所述一种或更多种核酸的量的时间差异指示免疫活性。例如,可随时间推移监测移植患者,以确定来自微生物组的一种或更多种核酸的量。来自微生物组的一种或更多种核酸的量的随后回归标准值的暂时性减少可指示较不严重的状况而不是移植排斥。在另一方面,来自微生物组的一种或更多种核酸的量的持续减少可指示严重的状况诸如缺乏有效的免疫抑制和植入排斥。

[0181] 在一些实施方案中,来自微生物组的所述一种或更多种核酸的量的时间差异可被用于监测免疫抑制剂治疗的效果或被用于选择免疫抑制剂治疗。例如,可在免疫抑制剂治疗之前和之后确定来自微生物组的一种或更多种核酸的量。在治疗之后,来自微生物组的一种或更多种核酸的减少可指示该治疗在预防移植物排斥方面是成功的。另外,来自微生物组的一种或更多种核酸的量可被用于在多种免疫抑制剂治疗之间选择,所述多种免疫抑制剂治疗例如不同强度的免疫抑制剂治疗。例如,来自微生物组的一种或更多种核酸的较低的量可指示需要非常有效的免疫抑制剂,而来自微生物组的一种或更多种核酸的较高的

量可指示可以使用较小效力的免疫抑制剂。

[0182] 本发明提供了灵敏且特异的方法。在一些实施方案中,本文描述的用于诊断或预测移植状态或结果的方法具有至少50%、60%、70%、80%、90%、95%或100%的灵敏度。在一些实施方案中,本文描述的方法具有至少50%的灵敏度。在一些实施方案中,本文描述的方法具有至少78%的灵敏度。在一些实施方案中,本文描述的方法具有约70%到约100%的特异性。在一些实施方案中,本文描述的方法具有约80%到约100%的特异性。在一些实施方案中,本文描述的方法具有约90%到约100%的特异性。在一些实施方案中,本文描述的方法具有约100%的特异性。

[0183] 本发明提供了通过监测来自非人类来源的无细胞DNA或RNA的序列的用于个体的非侵入性诊断,所述个体包括正在用免疫抑制方案治疗的个体、用抗微生物剂治疗过的个体等。例如,个体携带很多病毒,其中病毒载量在本文中被显示随个体的免疫活性而变化。用于监测免疫活性的优选的病毒是指环病毒属,其中病毒荷载在本文中被显示与个体的免疫活性相关。

[0184] 在一些实施方案中,本发明提供了用于检测和/或定量血浆中的或来自病毒颗粒的通常为游离的循环核酸的方法、装置、组合物和试剂盒,所述方法、装置、组合物和试剂盒用于诊断、预后、检测和/或治疗感染、免疫活性、移植状态或结果。

[0185] 在一些特定实施方案中,本发明提供了通过病毒组分析来非侵入性检测移植患者中的免疫活性的方法,所述方法避开了由来自其他外部来源的DNA所致的微嵌合的潜在问题并且对所有器官接受者是通用的而不考虑性别。在一些实施方案中,对个体的病毒组产生基因指纹。该方法允许可以以不依赖于供体和接受者的性别的方式进行可靠的序列鉴定。

[0186] 在例如与移植、自身免疫疾病的治疗等联合的免疫抑制方案之后,体液诸如血液可从患者获得并且被分析标志物。体液的实例包括但不限于涂片、痰、活检物、分泌物、脑脊液、胆汁、血液、淋巴液、唾液和尿液。病毒组序列的检测、鉴定和/或定量可利用以下来进行:实时PCR、芯片、循环核酸(例如无细胞DNA)的高通量鸟枪法测序和本领域已知的其他方法,包括本文描述的方法。可随着时间推移监测病毒载量并且该比例的增加可被用于确定免疫活性的状态或结果。

[0187] 在本文中描述的任何实施方案中,移植植入物可以是任何实体器官或皮肤移植。可通过本文描述的方法分析的器官移植的实例包括但不限于肾移植、胰腺移植、肝移植、心脏移植、肺移植、肠移植、肾移植后胰腺移植和同时的胰腺-肾移植。

[0188] 在一些其他实施方案中,本发明的方法被用于在个体水平上或在例如以临床试验的方式分析一组患者方面确定疗法治疗疾病(包括感染)的效果。此类实施方案通常包括比较患者或一组患者的两个时间点。预期患者状态由于治疗剂、治疗方案或对经受治疗的患者的疾病威胁的原因而在两个时间点之间不同。

[0189] 此类实施方案的模式实例可包括但不限于在两个或更多时间点分析微生物组,其中第一个时间点是经诊断但未被治疗的患者;并且第二个或另外的时间点是候选治疗剂或治疗方案治疗的患者。

[0190] 在另一种模式中,第一个时间点是候选治疗剂或治疗方案引起的,例如通过目前的临床标准查明的被诊断在疾病缓解期的患者。第二个或另外的时间点是候选治疗

剂或治疗方案治疗,并且受到疾病诱发剂威胁(例如在疫苗的背景中)的患者。

[0191] 在此类临床试验模式中,每组时间点可对应于单个患者、对应于患者组例如组群或对应于个体数据和组数据的混合。此类临床试验模式还可包括另外的对照数据,例如如本领域已知的安慰剂组、无疾病组等。感兴趣的模式包括交叉研究,随机、双盲、安慰剂对照、平行组试验也能够测试药物效果等。参见,例如,Clinical Trials:A Methodologic Perspective,第二版,S.Piantadosi,Wiley-Interscience;2005,ISBN-13:978-0471727811;和Design and Analysis of Clinical Trials:Concepts and Methodologies,S.Chow和J.Liu,Wiley-Interscience;2003;ISBN-13:978-0471249856,每个被特别地通过引用并入本文。

实施例

[0192] 人类病毒组对免疫抑制和抗病毒疗法的时间响应

[0193] 微生物组的病毒组分人类病毒组仍是相对地未被研究的(Wylie等(2012)Transl Res 160,283-290),并且关于免疫调节和抗病毒疗法对病毒组组成的影响知之甚少。尽管发现了膳食和病毒组组成之间的关联(Minot等(2011).Genome Research 21,1616-1625),以前显示健康的肠病毒组随时间推移维持非常稳定(Reyes等(2010)Nature 466,334-338),并且变化的主要来源是由于受试者之间的差异。

[0194] 免疫抑制疗法显著地减少了在器官移植中植入物排斥的风险但增加了接受者对感染的易感性。感染病毒病原体、特别是疱疹病毒巨细胞病毒(CMV)频繁发生并且增加接受者植入失败的风险。器官移植接受者因此频繁地经受针对CMV的抗病毒的预防性疗法或先天性疗法。

[0195] 在免疫抑制的水平和感染和排斥的风险之间的负相关关系仅留出狭窄的患者治疗可用的治疗窗。移植后护理被目前可用的用于诊断感染和排斥的方法的许多限制进一步复杂化。诊断排斥主要依赖于侵入性活组织检查,所述活组织检查具有观察者间变化性、高成本和患者不适的问题。鉴于免疫抑制后感染的症状减少,以及通常使用的诊断方法诸如抗原检测和基于PCR的分子测试依赖于特定的靶并因此依赖于对感染的来源先验假设的事实,感染的诊断是有挑战性的。

[0196] 作为最终的并发症,在对免疫抑制药物的敏感性方面患者-患者变化性可引起免疫抑制过度和免疫抑制不足,免疫抑制过度和免疫抑制不足分别增加感染或排斥的风险。

[0197] 存在很少的实质性方法测量免疫系统的健康,并且对免疫活性和微生物组的病毒组分之间的关联了解甚少。器官移植接受者用将免疫抑制药物和抗病毒药物组合的移植后疗法来治疗,为免疫调节对人类病毒组的影响提供了窗口。我们利用对血浆中的无细胞DNA测序来调查器官移植接受者的群组(656个样品,96名患者)中的药物-病毒组相互作用,并发现抗病毒剂和免疫抑制剂强烈地影响血浆中的病毒组的结构。我们在疗法开始时观察标记的病毒组组成动态,并发现总病毒载量随着免疫抑制增加,而微生物组的细菌组分在很大程度上维持不受影响。该数据提供了对在人类病毒组、免疫系统的状态和药物治疗的效果之间关系的洞察,并提供了用病毒组状态预测免疫活性的潜在应用。

[0198] 在本研究中,我们对在血浆中循环的无细胞DNA测序,以调查在器官移植后药物-微生物组的相互作用。我们研究了在经受了免疫抑制剂和抗病毒预防的联合的心脏和肺移

植接受者中的感染模式。我们发现免疫抑制剂和抗病毒剂对微生物组的病毒组分的结构具有强烈的影响,但对细菌组分无影响。在药物疗法的开始观察到强烈的组成动态,因为不同个体的病毒组组成收敛至相似的、药物决定的状态。总病毒载量响应于疗法显著地增加,因为病毒、特别是指环病毒属从免疫活性的降低得到好处。最终,我们证明测量指环病毒属荷载使排斥和非排斥接受者能够区分(stratification)。

[0199] 从96名实体器官移植接受者(41名成年心脏,24名儿科心脏,31名成年肺)纵向收集656个血浆样品。从血浆纯化无细胞DNA并测序。总计,我们获得8200亿个碱基(820Gbp)的测序数据,平均每个样品1.25Gbp(Illumina HiSeq,1×50bp读段,图1中的B图)。器官移植接受者持续地参加本研究,超过多于2年的过程,并且在移植后规定时间点从接受者收集样品,在移植后的最初几个月内样品收集的频率最高。图1中的C图显示了作为不同的患者类别移植后的时间的函数的被分析的样品数。

[0200] 作为标准移植后疗法的一部分,用抗病毒预防和免疫抑制治疗群组中的患者(图1中的D图)。对于成人心脏和肺移植接受者,维持性免疫抑制是基于他克莫司的,并用霉酚酸酯和强的松补充。儿科患者被用基于环孢霉素的抗排斥疗法治疗。CMV阳性移植接受者(接受者和/或供体以前感染CMV)用抗病毒预防来治疗,而CMV阴性接受者不用。方案设计在移植后的最初几个月内需要高剂量的免疫抑制剂和抗病毒药物,之后,随着排斥和感染的风险减少,剂量逐渐减少。考虑到免疫抑制可用的狭窄的治疗窗和他克莫司的药代动力学的大的患者-患者变化性,直接测量血液中的他克莫司的浓度并且调节剂量以维持靶药物水平。图1中的D图示出了在他克莫司治疗的患者的血液中测量的他克莫司的平均水平,并示出了药物治疗方案的设计。

[0201] DNA序列分析.在计算扣除人类来源的序列后,鉴定微生物来源的序列。为此,移出重复和低质量的读段并且将剩余的读段映射到人类参考基因组build hg19(BWA(Li和Durbin,2009),参见方法)。然后,收集未映射的读段,并且移出低复杂度的读短。图1中的E图示出了在应用重复和质量过滤器后剩余读段的比例(平均86%)的分布,以及在扣除人类读段后剩余的比例(平均2%)的分布。

[0202] 为了鉴定感染性病原(agents),利用BLAST将剩余的、高质量、独特的、非人类读段映射至病毒基因组(n=1401)、细菌基因组(n=1980)和真菌基因组(n=32)的参考数据库(下载自NCBI,图6中的A图)。将0.12%的独特地测序的读段与至少一个靶基因组比对(图6中的B图、C图)。我们利用定量PCR(qPCR)测定来验证通过基于测序的方法鉴定的阳性命中,该定量PCR(qPCR)测定靶向测序鉴定的靶的一个亚组(疱疹病毒4、5、6和细小病毒)。我们现在通过测序和qPCR测量的病毒计数之间数量一致(图6中的D图)。

[0203] 我们还发现用于检测疱疹病毒的测序测定的灵敏度与qPCR测量不分上下。测序测定可用的较大的捕获截面-完整的靶基因组对PCR扩增子靶区域-因此足以克服由测序文库制备的有限的效率和文库采样过疏导致的在测序中的信号丢失。横跨本研究中的所有样品,利用测序测量的最高CMV载量与罹患临床上诊断的散播性CMV感染的两名成年心脏移植患者相对应(参见图6中的E图)。

[0204] 为了测试用于DNA提取和测序文库制备的试剂中的潜在污染物的存在,我们进行了两个对照实验。首先,我们制备带有已知模板(λ gDNA, Pacbio产品编号:001-119-535)的2个样品,并纯化DNA用于利用上文描述的工作流(Illumina Miseq,3.4和3.5百万个读段)测

序。移出 λ 来源的序列,并将剩余的序列(0.4%)与上文描述的BLAST参考数据库比对。未发现在该研究中讨论的多种感染性病原的迹象,但我们确实检测到与肠杆菌科细菌科(变形菌门)(主要是大肠杆菌(>97%))和肠道细菌噬菌体(<1%)有关的序列,其可能是 λ DNA培养物的残留物。在第二个对照中,我们从无核酸酶的水制备样品用于测序。该样品连同与本研究无关的样品一起被包括在测序运行中,并且仅收集到映射到两个细菌物种的基因组的总共15个的有限数目的序列。再次,未发现以下讨论的感染性病原的迹象。

[0205] 我们利用Grammy,一种利用BLAST获得的序列相似性数据进行物种的相对丰度的最大可能性评价的工具,在不同的系统分类水平上研究了血浆中的微生物组组成。GRAMMY解决了靶基因组大小的差异和读段指定的不确定性。应注意,该方法仅允许评价其基因组数据在参考数据库中可得的物种的丰度。图1中的F图显示了在不同的系统分类水平上,物种的相对丰度(所有样品的平均)。我们发现病毒(73%)比细菌(25%)和真菌(2%)更丰富地被呈现(图1中的F图,a部分)。在病毒中,我们发现ssDNA病毒占据比dsDNA病毒(28%)更大的比例(72%)。发现了7个不同的病毒科(丰度>0.75%),有一个主要的科-指环病毒科,其占总群体的68%(图1中的F图,b部分)。指环病毒科的比例主要(97%)由来自甲型细环病毒属的病毒构成(图1中的F图,c部分)。甲型细环病毒属是细环病毒(TTV)的属,并且鉴定到与14个不同的细环病毒类型有关的序列(图1中的F图,d部分)。在人类群体中多瘤病毒的感染是普遍的,并且在实体器官移植后的第一年中多瘤病毒DNAemia并非不常见。在对应于当前群组中的36名患者的75个样品(11%)中发现了多瘤病毒来源的序列。发现了BK(41%)、JC(27%)、TS(4%)、Wu多瘤病毒(6%)、SV40(6%)和近期发现的HPyV6(13%)(Schowalter等2010)的存在的迹象(图1中的F图,e部分)。在细菌中,变形菌门(36%)、厚壁菌门(50%)、放线菌门(10%)、拟杆菌门(4%)是在样品中被最丰富地呈现的门(图1中的F图,f部分)。

[0206] 为了调查本研究可用的相对短的读段(50bp)的潜在的不正确指定,我们基于对一个亚组的样品($n=55$)收集的较长的、双末端读段($2\times 100\text{bp}$),检查了丰度评价对读段长度的依赖性。我们发现对于此处报告的所有系统分类水平,基于50bp亚读段(subreads)和100bp读段的丰度评价是相似的(图6中的F图)。

[0207] 病毒组组成对药物剂量的敏感性.使用关于药物剂量的可用临床数据分析药物-微生物组相互作用。此处,我们检查了用基于他克莫司的抗排斥方案治疗的成年心脏和肺移植患者的数据(47名患者和380个观察值),借以排除了用环孢霉素治疗的儿科患者和由于药物耐受问题而从他克莫司转换为环孢霉素免疫抑制的患者。从个体患者记录收集处方抗病毒药物剂量(缬更昔洛韦)和所测量的血液中他克莫司的水平的数据,并推导出对应于不同药物水平的样品的平均组成。为了解释微生物组组成对剂量变化的延迟效应,对药物水平和剂量数据进行滑动平均窗过滤(参见图1中的C图和图7中的A-C图;窗口大小45天)。

[0208] 我们发现微生物组的病毒组分是药物剂量的敏感性函数(47名患者,380个样品,图2中的A图)。但是,如下面进一步讨论的,微生物组的细菌组分的结构未被药物疗法显著地改变(图7中的D图)。当患者接受低剂量的缬更昔洛韦和他克莫司时,疱疹病毒目和有尾噬菌体目在病毒组占首要地位。相比之下,高剂量的免疫抑制剂和抗病毒剂引起由指环病毒科占首要地位的病毒组结构(在高药物水平占多达94%)。抗病毒预防意图预防CMV疾病,但是其他疱疹病毒也易受该药物影响,所以较高剂量的缬更昔洛韦引起较低比例的来自疱

疹病毒目的病毒并不出乎意料。指环病毒科利用宿主免疫系统的抑制的观察结果与来自文献的各种观察结果一致：之前显示在HIV患者中随着趋向AIDS的进程，指环病毒科的发生率增加，并且肝脏移植后指环病毒属TTV的总荷载增加。此外，近期报道了在发热的儿科患者中，指环病毒科的发生率增加。

[0209] 我们接下来将对器官移植接受者测量的病毒组组成和未使用免疫抑制剂或抗病毒剂的健康个体中观察到的组成 ($n=9$ ，测序数据从以前的研究可得) 比较。此处，我们将健康组成与在相应于最小药物暴露的药物疗法的开始对器官移植接受者测量的组成 (手术后第一天， $n=13$) 比较，并与对暴露于高药物水平的移植接受者测量的组成 (移植程序后，他克莫司 $\geq 9\text{ng/ml}$ 、缬更昔洛韦 $\geq 600\text{mg}$ ， $n=68$) 比较。我们发现健康参考样品的病毒组和对应于最小药物暴露的样品的病毒组的组成相似 (图2中的B图)。但是，健康参考样品的组成和对应于最小药物暴露的样品的组成与对高药物暴露样品测量的由指环病毒科占首要地位的组成不同。

[0210] 移植后最初3天内，对基于他克莫司的免疫抑制疗法补充诱导疗法 (用抗胸腺细胞球蛋白、达利珠单抗和巴利昔单抗)，并且患者贯穿移植后疗法另外接受皮质类固醇强的松。强的松和他克莫司的时间-剂量曲线是类似的：在疗法的开始高剂量，随后剂量逐渐减少 (图7中的A-C图)。因此，图2中的A图中的数据反映了强的松和他克莫司的联合效果。强的松和缬更昔洛韦对病毒组组成的不同效果的分析 (图7中的E图) 显示出在图2中的A图中观察到的相同趋势：较高的强的松剂量导致较大的指环科病毒呈现。最后，我们注意到一个亚组的患者未经抗病毒药物治疗。对应于该亚组的患者的数据允许我们如下文描述的进一步解释抗病毒药物和免疫抑制剂对病毒组组成的不同效果。

[0211] 微生物组多样性的划分。我们研究了微生物组的细菌组分和病毒组分的多样性。对于细菌和病毒二者，受试者内 (within-subject) 的多样性低于受试者间 (between-subject) 多样性 (Bray-Curtis β 多样性，细菌组成在门水平，病毒在科和目水平，图2中的C图)。根据移植类型、心脏或肺、或年龄对患者的数据划分不降低多样性。在受试者内，同样对于细菌和病毒两者，在一个月时间跨度内收集的样品的多样性较低。对于病毒，我们发现当比较在相似的药物剂量 (他克莫司水平 $\pm 0.5\text{ng/ml}$ ，缬更昔洛韦 $\pm 50\text{mg}$) 收集的样品时，多样性较低，但对于细菌并非如此。连同图2中的A图中的群体平均对药物剂量的敏感性一起，我们由此发现经受相同药物疗法的患者的病毒组的组成收敛至相似的状态。

[0212] 病毒组对药物剂量变化的动态响应。观察到病毒组对药物剂量变化的强的时间响应，与病毒组组成对药物剂量的敏感性一致。图3中的A图示出ssDNA和dsDNA病毒的相对基因组丰度的时间依赖性 (数据来自所有患者组和样品， $n=656$)。在移植后最初几个月期间，ssDNA病毒部分迅速扩展，随后在6个月后呈相反的趋势。图3中的B图示出了在科和目水平上分组的最丰富的病毒的时间依赖性相对组成，并且提供了关于病毒组组成动态的更多细节 (数据来自所有患者组和样品， $n=656$)。dsDNA部分由有尾噬菌体目、腺病毒科、多瘤病毒科和疱疹病毒目组成，在移植后的最初几周，它们总共占病毒组的95%。ssDNA病毒只占初始病毒组的5%，并且主要由指环病毒科的成员组成。由于腺病毒科、有尾噬菌体目和疱疹病毒目被抗病毒预防有效地靶向，由这些病毒类型所占的比例在最初几个月内剧烈下降。相比之下，由于指环病毒科很大程度地逃脱被抗病毒药物靶向并且利用患者降低的免疫活性，这些病毒类型的相对丰度快速增加 (在4.5-6个月期间最大值84%)。器官移植程序后6

个月,观察到相反的趋势,与由治疗方案指示的抗病毒药物和免疫抑制药物的减少一致。

[0213] 在门、目和属分类水平上进行观察,与病毒组分相比,微生物组的细菌组分随着时间保持相对稳定(图3中的C图, $n=656$,和图8)。图3中的D图显示细菌和病毒属的样品内 α 多样性作为时间的函数(香农熵,一个月的时间段,检查了590个细菌属、168个病毒属)。在疗法的开始,观察的病毒属的多样性下降(在第1个月内的 1.05 ± 0.5 到在第4-5个月内的 0.31 ± 0.33 , $p\ll 10^{-6}$,曼-惠特尼U检验),而在移植后疗法的过程期间,细菌的 α 多样性保持相对不变(在第1个月内的 2.2 ± 1.14 到在第4-5个月内的 2.6 ± 0.85 , $p=0.1$,曼-惠特尼U检验)。

[0214] 在移植后疗法的开始,总病毒载量增加。为了获得对于治疗药物对总病毒载量的影响的了解,我们通过将病毒靶的基因组覆盖度对人类基因组的覆盖度归一化,推导出所有病毒相对于人类基因组拷贝数的绝对基因组丰度。对于本研究的所有患者组部分,不论移植类型(心脏或肺)或年龄(成年或小儿)如何,在疗法的开始观察到总病毒载量的增加(图4中的A图)(载量的变化, 7.4 ± 3 ,sigmoid拟合,黑线)。结合相对丰度数据,总病毒载量数据显示在移植后最初3个月内,同时用抗病毒剂和免疫抑制剂治疗的患者的疱疹病毒目净减少且指环病毒科载量净增加。

[0215] 因此,所述数据显示出抗病毒剂和免疫抑制剂的联合对不同病毒类型的不同效果。所述数据还显示出总腺病毒科载量的减少,指示腺病毒科复制被缬更昔洛韦抑制,与以前的研究一致。图4中的B图归纳了所有移植类型的数据,但是当如下根据不同移植类型分类时观察到相同的趋势:成年心脏移植接受者($n=268$,图9中的A图),成年肺移植接受者($n=166$,图8中的B图)和用与他克莫司相对的环孢霉素治疗的儿科患者($n=99$,图9中的C图)。

[0216] 并非本研究群组中的所有患者都接受抗病毒药物和免疫抑制药物两者:对于其中供体和接受者二者在CMV抗体测定中均未显示出以前感染CMV的迹象的移植情况,认为由抗病毒预防引起的并发症的风险超过新获得CMV感染的潜在风险,并且患者因此不用抗病毒预防来治疗。因此,这些患者仅用免疫抑制剂来治疗。图9中的C图显示了CMV阴性病例($n=75$)的时间依赖性病毒载量和组成。仅免疫抑制剂的疗法的净效果是所有病毒类型包括疱疹病毒目和腺病毒科的扩增。逐渐减少免疫抑制导致总病毒载量的减少。

[0217] 罹患植入物排斥事件的患者中较低的指环病毒属荷载。考虑到指环病毒属荷载与免疫抑制程度的关联(参见图2中的A图和图4),并且考虑到免疫活性和排斥的风险之间的相关性,我们发问:指环病毒属荷载是否可被用于对排斥植入物接受者和非排斥植入物接受者分类。图5中的A图显示了对排斥患者和非排斥患者测量的指环病毒属载量作为移植后时间的函数。此处,在患者罹患至少一种活组织检查确定的中度或重度的排斥事件,活组织检查等级 $\geq 2R/3A$ (红色,20名患者,177个数据点)的情形中,该患者被分类为排斥。无排斥患者对应于贯穿其移植后过程未被诊断为罹患中度或重度的植入物损坏的患者(蓝色,活组织检查等级 $< 2R/3A$,40名患者,285个数据点)。

[0218] 图5中的A图显示了在几乎每个时间点排斥个体的指环病毒属荷载显著较低。我们接下来直接将排斥的患者的指环病毒属荷载与对不存在排斥的患者测量的荷载比较。为了解释以上描述的指环科病毒属载量的时间依赖性(图5中的A图),我们推导出相应于在同一时间点对所有样品测量的平均载量的指环科病毒属载量。图5中的B图显示了与对罹患中度排斥事件的患者(活组织检查等级 $1R$, $N=102$)和罹患重度排斥事件的患者(活组织检查等

级 $\geq 2R/3A$, $N=22$) 测量的载量相比, 非排斥患者的时间归一化的载量 ($N=208$)。该图显示对于较大排斥风险的患者, 时间归一化的载量显著较低。通过对具有较大的测量点的群体随机取样来计算 p 值, $p = \text{sum}(\text{median}(A_{\text{排斥}}) > \text{median}(A_{\text{非排斥}})) / N$, 其中 $N=10^4$ 并且 $A_{\text{排斥}}$ 和 $A_{\text{非排斥}}$ 分别是较大和较小的排斥和非排斥风险的群体的相对病毒载量 ($p=0.011$, $p=0.0002$ 和 $p=0.036$)。

[0219] 这些观察结果与排斥的风险和感染的发生率与患者的免疫活性具有相反的相关性的观点一致 (参见图5中的A图的插图)。因此, 对排斥患者观察到的较低的病毒载量指示该亚组的患者中的较高水平的免疫活性, 即使这些患者用相同的免疫抑制方案治疗。已知会存在在对免疫功能抑制的敏感性方面的患者-患者变化性, 并且免疫抑制的可预测性的缺乏是移植中的重要风险因素。尚未发现目前使用的用于测量免疫活性的商业测定对急性排斥或显著感染具有预测性。因此, 用于直接测量免疫活性的能代替或补充现有测定的方法的开发将是重要的。所记录的器官移植接受者中的总指环病毒属载量可作为可选的标记。图5中的C图显示了接受者操作特征并且测试了相对指环病毒属载量在对非排斥患者和排斥患者分类方面的表现 (曲线下面积=0.72)。

[0220] 我们已经通过对接受者血浆中的无细胞DNA测序研究了实体器官移植后的药物-微生物组相互作用。数据显示了关于血浆中的人类病毒组的基本结构和其如何响应于药物干扰的很多信息; 它们还显示了微生物组的细菌组分的组成对免疫抑制的相对不敏感性。这些数据可用于设计和优化移植后治疗方案。例如, 它们显示从初始的高剂量逐渐减少抗病毒预防导致疱疹病毒目比例的再起。以前已经显示CMV DNA载量预测CMV疾病复发和排斥, 提出了患者是否将从长期预防疗法受益的问题。

[0221] 免疫抑制后指环病毒科的丰度显著扩展也值得进一步考虑。指环病毒科在人类群体中是普遍存在的, 并且, 尽管尚未确立病原性, 指环病毒科目前作为致癌作用中的潜在共因子 (cofactor) 被研究。指环病毒科对免疫抑制的敏感性使器官移植成为用于研究指环病毒科的特性的理想设定, 特别是考虑到在移植接受者中观察到癌症的发生率增加。在罹患排斥事件的患者中低于平均值的指环病毒科荷载的观察结果指示在该亚组的患者中免疫抑制不足, 即使这些患者经受了按照方案规定的免疫抑制剂水平。这表明除了测量循环的药物水平之外, 在设计测定中允许直接测量患者的免疫活性的水平将是有价值的。在移植接受者的血液中鉴定的指环病毒科总荷载可作为个体患者的免疫抑制的总体状态的一个这样的标志。

[0222] 高通量DNA测序在感染的非假设诊断中 useful。鉴于感染在移植中频繁发生并且在免疫功能不全的个体中难以诊断的事实, 并且鉴于通过定量在血浆中循环的供体来源的人类DNA, 序列分析可另外提供关于植入物健康的信息, 该方法在移植的背景中是特别适宜的。在感染性疾病的其他领域, 开发扣除方法以消除人类DNA并富集病毒和微生物来源的DNA可以是有价值的。

[0223] 实验程序

[0224] 临床样品收集: 在Stanford University Hospital (SUH) 或Lucile Packard Children's Hospital (LPC) 招募患者, 并且如果他们是多器官移植的接受者, 则被排除。该研究被Stanford University Institutional Review Board批准 (方案#17666), 并且在2010年3月开始招募。关于患者招募和患者的移植后治疗的细节参见扩展实验程序部分。

[0225] 血浆处理和DNA提取:如以前描述的 (Fan等2008),在样品收集的三个小时内从全血样品提取血浆,并且储存在-80℃。当需要用于分析时,将血浆样品解冻,并且使用QIAamp循环核酸试剂盒 (Qiagen) 立即从0.5-1ml血浆提取循环DNA。

[0226] 测序文库的制备和测序:使用NEBNext DNA Library Prep Master Mix Set for Illumina和标准Illumina索引接头 (购自IDT),或使用基于微流体的自动化文库制备平台 (Mondrian ST,0vation SP Ultralow library system),从纯化的患者血浆DNA制备测序文库。使用Agilent 2100Bioanalyzer (高灵敏度DNA试剂盒) 表征文库并且通过qPCR定量。样品是26个不同的测序运行的一部分,并且在22个月的进程中被测序。每个通道测序平均6个样品。

[0227] 移植后监测和临床样品收集.该分析代表由National Institutes of Health资助以研究供体来源的无细胞DNA测定用于在胸部器官移植后诊断急性和慢性排斥以及同种异体植入失败的临床应用的前瞻性群组研究 (RC4 AI092673) 的子研究。如果患者在Stanford University Hospital (SUH) 或Lucile Packard Children's Hospital (LPCH) 接受了心脏或肺移植,则其被招募,并且如果他们是多器官移植的接受者,或如果他们在移植后在其他中心而不是SUH或LPCH被跟踪,则将他们排除。该研究被Stanford University Institutional Review Board批准 (方案#17666),并且在2010年3月开始招募。

[0228] 移植后治疗方案的细节、成年心脏移植接受者.移植后免疫抑制由术后立即施用甲基强的松500mg、随后每8小时125mg持续3个剂量组成。在术后第1、2和3天施用抗胸腺细胞球蛋白 (rATG) 1mg/kg。维持性免疫抑制由以下组成:在术后第1天开始每天两次20mg强的松,并且到术后第6个月逐渐减少到<0.1mg/kg/天,并且如果心内膜心肌活组织检查显示无细胞排斥的迹象,则进一步逐渐减少。在术后第1天开始他克莫司,并且进一步调节剂量以在第0-6个月期间维持10-15ng/ml的水平,在第6-12个月期间维持7-10ng/ml的水平,并且之后维持5-10ng/ml的水平。在术后第1天以每天两次1,000mg开始霉酚酸酯,并且如果需要,响应于白细胞减少,进行剂量调节。

[0229] 除非供体和接受者均是CMV阴性的,所有患者接受由以下组成的标准CMV (抗病毒) 预防:在术后第1天开始每12小时5mg/kg IV更昔洛韦 (根据肾功能调节)。当能够耐受口服药物时,接受者开始接受缙更昔洛韦,900mg,每天两次,持续2周,然后每天900mg直到移植后6个月,随后每天450mg直到移植后12个月,在该点处抗病毒预防不再继续。在白细胞减少的设定中,进行缙更昔洛韦剂量的减少。CMV+同种异体植入物的CMV-接受者还接受CMV超免疫球蛋白:在移植72小时内150mg/kg IV,在移植后第2、4、6和8周100mg/kg,并且在移植后第12和16周50mg/kg。

[0230] CMV-同种异体植入物的CMV-接受者不用抗病毒预防治疗,直到2012年5月;随后,这些接受者用阿昔洛韦来治疗,400mg,每天两次,持续一年。抗真菌预防由以下组成:伊曲康唑,每天300mg,持续手术后的前三个月;并且针对金罗维氏肺孢子虫 (pneumocystis jiroveci) 感染的预防由以下组成:甲氧苄啶/磺胺甲恶唑,每天80mg TMP组分。针对肺孢子虫属感染的预防无限期地继续,并且不耐受TMP-SMX的患者被用阿托伐醌、氨苯砞或吸入性喷他脒治疗。

[0231] 通过在移植后以以下排定的间隔进行监控心内膜心肌的活检物来监测所有心脏移植接受者的急性细胞排斥:在第一个月期间一周一次,两周一次直到第3个月,每月一次

直到第6个月,并且然后在第9、12、16、20和24个月。根据ISHLT 2004年修订的分级量表(0、1R、2R、3R) (29)对活检物分级。在移植后的以下的时间点从心脏移植接受者收集血液样品:第2、4和6周;第2、2.5、3、4、5、6、8、10、12、16、20和24个月。一个亚组的心脏移植接受者还在移植后第一天收集血液样品。如果在同一天进行血液取样和心内膜心肌活组织检查,需要注意确保在活组织检查程序之前收集血液。

[0232] 儿科心脏移植接受者. 诱导免疫抑制最初由以下组成:每两周达利珠单抗1mg/kg IV,持续总共5个剂量;并且于2011年8月开始在手术后第0天和第4天转换为巴利昔单抗10-20mg IV。接受者还立即用脉冲甲基强的松龙治疗,每8小时10mg/kg IV,持续3个剂量,随后用强的松,0.5mg/kg,一天两次,持续移植后的最初14天;在急性排斥的不存在下,随后在移植后的第一年期间逐渐减少皮质类固醇。

[0233] 钙调磷酸酶抑制主要由以下组成:环孢霉素,300-350ng/ml的目标水平持续第0-3个月,275-325ng/ml的目标水平持续第4-6个月,250-300ng/ml的目标水平持续第7-12个月并且在移植后第12个月之后200-250mg/ml的目标水平。不耐受环孢霉素的患者用他克莫司治疗。对于成年心脏移植接受者,用于预防机会主义感染和监控心内膜心肌活检物的方案是类似的。

[0234] 肺移植接受者. 移植后免疫抑制由以下组成:术后立即施用甲基强的松龙500-1000mg,随后每天两次0.5mg/kg IV。在第0天和第4天给予巴利昔单抗20mg IV用于诱导免疫抑制。维持性免疫抑制由以下组成:在术后第0-3天,甲基强的松龙0.5mg/kg IV,每天两次,随后强的松每天0.5mg/kg,直到第30天,并且随后在移植后第6-12个月期间每2-3个月逐渐减少至每天0.1mg/kg。在术后第0天开始他克莫司,并且调节剂量以在第0-6个月期间维持12-15ng/ml的水平,在第6-12个月期间维持10-15ng/ml的水平,其后维持5-10ng/ml的水平。在术后第0天以500mg每天两次开始霉酚酸酯,并如果需要,响应于白细胞减少,进行剂量调节。对于成年心脏移植群组,抗病毒、抗真菌和PCP预防是类似的。

[0235] 通过在移植后第1.5、3、6、12、18和24个月进行支气管活组织检查方案监测所有肺移植接受者的急性细胞排斥。如果基于症状或肺功能测试结果,临床原因指示,也进行活组织检查。以以下间隔从肺移植接受者收集血液样品用于研究目的:移植后第一天三次,第二天两次、和第三天一次,随后是第一周和第二周,以及第1.5、2、3、4.5、6、9、12、18和24个月。在进行符合方案和临床指示的活组织检查之前抽取血液样品。

[0236] 用于鉴定病原体来源的序列的工作流. 使用基于C的实用程序fastq.cpp移出准确的重复。使用为fastx包的一部分的质量过滤器(fastq_quality_filter-Q33-q21-p50)移出低质量的读段。然后利用BWA将剩余的读段与人类参考基因组build hg19(bwaaln-q25)比对。使用samtool(samtool view-f14)收集未映射的读段,并且使用Seqclean(seqclean-1 40-c1)移出低复杂度的数据。然后将读段与病毒参考基因组、细菌参考基因组和真菌参考基因组以及在ncbi_fungi下载的所有参考的选择比对。

[0237] 图6中的A图显示了基因组大小的分布。下列参数用于BLAST比对:奖励=1、惩罚=-3、字_大小=12、空位开放=5、空位延伸=2、e值= 10^{-4} 、perc_身份=90、剔除_限制=2。移出对齐长度短于45的blast命中。对于样品的一个亚组,较长的读段是可用的(2 3 100bp, n=55)。为了测试基因组丰度评价的稳健性,检查组成测量结果的长度依赖性。此处,读段被修剪为40、50、65、80和100bp的长度(fastx_trimmer),并利用以上描述的工作流

来分析。此处,对于40、50、65、80和100bp的读段,分别将具有短于37、45、59、72和80bp的对齐长度的blast命中移出。使用GRAMMy计算Genome Abundance Estimation Relative基因组丰度评价。该工具利用BLAST来源的核酸序列相似性数据来进行样品中的物种的相对丰度的最大可能性评价。GRAMMy过滤器通过BLAST比对度量标准(E-评分、对齐长度和相同比率)命中,并且解决了靶基因组大小以及在评价候选参考基因组的相对丰度中读段指定的不确定性。使用下列参数调用Grammy:python grammy_rdt.py;python grammy_pre.py-q "40,40,1"input set;python grammy_em.py-b 5-t 0.0001-n 100input.mtx;grammy_post.py input.est setinput.btp。

[0238] 自定义脚本被用于联合株水平丰度评价以从较高分类水平丰度获得丰度。此处,使用Taxtastic构建参考数据库的最小分类。

[0239] 定量绝对病毒载量。为了定量样品中的感染性病原的载量,收集blast命中结果并且利用自定义脚本(Bioperl)选择每个读段的最佳命中。图6中的B图显示了每一百万测序的独特分子独特病毒、细菌和真菌blast命中数目的分布。图6中的C图显示了与样品中存在的人类基因组拷贝的数目相应的病毒、细菌和真菌基因组拷贝的数目。感染性病原的基因组覆盖度被相对于人类基因组覆盖度归一化。

[0240] qPCR验证选择的病毒靶的测序结果。使用用于定量人类疱疹病毒4、5和6以及细小病毒的标准qPCR试剂盒(PrimerDesign,genesig),来验证一个亚组的无细胞DNA样品的测序结果。对从~1ml血浆提取并在100ml Tris缓冲液(50Mm[pH 8.1-8.2])中洗脱的cfDNA运行qPCR测定。血浆提取和PCR实验被在不同的设施上进行。运行无模板对照以核实PCR试剂被包括在每个实验中。图6中的D图将每一百万读段获得的blast命中的相对数目与利用qPCR确定的病毒基因组拷贝的浓度比较。

[0241] 无模板对照。进行无模板对照实验。从无核酸酶水(S01001,Nugen)产生测序文库。该文库连同7个另外的样品文库(无细胞人类DNA)一起被制备以测试在文库制备期间可能的样品-样品交叉干扰。为了确保在Illumina流动池上形成具有足够密度的簇,使该样品连同与本研究无关的样品一起被测序。其中与本研究无关的样品产生一千六百万个读段,无模板对照文库仅产生15个读段,所述15个读段映射到参考数据库中的两个物种methanocaldococcus janaschii基因组(9个命中)和枯草芽孢杆菌基因组(5个命中)。未发现人类相关序列的痕迹,指示样品-样品间的污染是低的。

[0242] 实施例2

[0243] 微生物组的临床监测

[0244] 使用如在实施例1中描述的方法,定量每个样品的映射到CMV基因组的读段。在临床上阳性感染的样品中观察到CMV丰度增加($p=7.10^{-9}$ 曼-惠特尼U检验,图10中的C图);在我们的样品中CMV来源的DNA的水平与具有0.91的AUC的CMV的临床报道匹配(图10中的C图)。该数据指示CMV监控可以使用相同的测序数据与排斥监测平行进行,并且引导我们检查其他病毒感染是否可被类似的监测。

[0245] 我们鉴定了被良好地表征的病原性和肿瘤病毒(图11中的A图)以及共生的细环病毒(TTV,甲型细环病毒属),其与以前的免疫抑制和TTV丰度之间的联系的结果一致。对这些病毒临床测试的频率相当不同,相对于其他病原体,CMV频繁监控(人类疱疹病毒5,HHV-5,在我们的群组中 $n=1082$ 个测试)(图11中的A图)。我们评价了相对于临床筛选频率

的感染的发生率(其中给定的病毒经由测序被检测到的样品的数目)。尽管CMV最频繁地被筛选(335个样品),如通过测序确定的其发生率(在22个样品中检测到)与不被常规筛选的其他病原体的发生率相似,该其他病原体包括腺病毒和多瘤病毒(分别在4个时机和1个时机被临床测试,图11中的A图)。

[0246] 腺病毒是社区获得性呼吸道感染,其可导致肺移植接受者中的植入物耗损,并且对于儿科患者具有特别高的风险。从一个被测试为腺病毒阳性的儿科患者(L78,图11中的B图的左上图)收集样品。该患者还在整个群组中具有最高的腺病毒来源的DNA载量。虽然测试通常限于儿科肺移植病例,在未被临床筛选的若干其他成年移植患者(例如L34,图11中的B图的左上图)中也观察到持续的腺病毒载量。

[0247] 多瘤病毒是肾移植后同种异体植入物排斥的主要原因,但通常并不被包括在肺移植后监控内。我们在未被测试该病原体的两名患者(L57和L15,图11中的B图的右上图)中检测多瘤病毒。在两个病例中,临床记录均指示持续的肾功能不全,其可能是由多瘤病毒感染所致。

[0248] 在大范围且非假设性筛选感染的益处的最近的实例中,我们检查了呈现出高载量的人类疱疹病毒(HHV)8(可在实体器官移植后导致并发症的肿瘤病毒)的患者(图11中的B图的左下图)。该患者(L58)被测试为对两种其他疱疹病毒(HHV-4a和HHV-5)阳性,所述两种其他疱疹病毒具有刺激HHV-8再激活的潜能。尽管只在特定的临床环境中建议移植后监测HHV-8,利用测序使非疑似病例中的该病毒能够被鉴定,该病毒以其他方式不被检测到。

[0249] 微生物组的临床监测.除了在血清中测量的病毒,我们还观察了在其他体液中检测的无细胞测量结果和真菌或细菌感染之间的关联,包括经由尿液培养检测的肺炎克雷伯氏菌(*Klebsiella pneumonia*)感染($ROC=0.98$)和在BAL中检测的真菌感染。细菌和真菌的关联性的表现对感染类型和检测的体液两者均是敏感。我们观察到血液具有更密切的关联的体液的较好表现,并且还观察到对背景信号的敏感性。例如,最常培养的细菌感染(假单胞菌属(*Pseudomonas*))在超过80%的我们的患者样品的无细胞测量中被检测到,其与仅在6%的我们的患者样品中检测到的、最常被检测的病毒病原性物种(CMV)形成鲜明对比。

[0250] 这突显了在为正常菌群的一部分的共生感染(包括假单胞菌属)和独自病原性且具有较低背景信号的非共生感染之间的重要区别。该区别可以解释对通常培养的共生感染测量的(例如,对于铜绿假单胞菌和大肠杆菌,分别是 $AUC=0.66$ 和 0.62),相应于非共生物(对于CMV $AUC=0.91$)在敏感性和特异性方面的不同。在共生细菌的情况中,临床问题不是存在或不存在,而是在不合适的身体部位存在或不存在。

[0251] 在我们的群组中,我们还检测到源自微孢子虫目(microsporidia)(可导致免疫抑制患者中的肠道感染的非共生真菌)的无细胞DNA。我们在表现出微孢子虫病的典型症状的患者L78中测量到持续性的微孢子虫目载量(图11中的B图的右下图)。腺病毒感染(L78,图11中的B图的左上图)是疑似原因,尽管内窥镜检查和乙状结肠镜检查结果是是非决定性的,并且粪便样品对艰难梭菌(*C.diff*)和腺病毒测试为阴性。基于我们的测序数据,微孢子虫病是最可能解释患者症状的,因为在该患者中测量的微孢子虫信号与来自被测试为微孢子虫目阳性的不相关群组的患者16的微孢子虫信号相似(图11中的B图的右下图)。

[0252] 由于每ml血浆具有超过一百亿个片段,循环无细胞DNA是了解人类生理的信息丰富的窗口,在癌症诊断和癌症治疗监测、遗传产前诊断和经由“基因组移植动态”(GTD)监测

心脏移植排斥中具有快速扩展的应用。在本研究中,我们采用对肺移植GTD的原理,肺移植是特别有挑战性的实体器官移植类型,受到差的存活率以及同种异体植入物排斥的错误和侵入性测试的限制。

[0253] 由于患有同种异体植入物排斥和急性排斥的肺移植接受者在临床上可表现出相似的症状,我们将GTD的范围扩展到感染性疾病监测。我们首先阐明了在临床测试结果和源自CMV(移植后植入物损伤的主要原因)的cfDNA的强烈相关性。我们还示出了:非假设性感染监测揭示了在与具有阳性临床测试结果和相关症状的患者比较具有相似的微生物cfDNA水平的患者中的很多未测试的病原体,包括未诊断的腺病毒、多瘤病毒、HHV-8、和微孢子虫的例子。这些实例表明了与病原物特定测试相对,大范围、基于测序监测感染的益处。该方法可直接被用作能协助确定感染的发生和来源的手段。在感染的发生率高、排斥和感染可共发生、以及感染和排斥症状难以辨别的移植的背景中,这可以是特别适用的。

[0254] 尽管本文已显示和描述了本发明的优选的实施方案,此类实施方案仅以示例的方式提供对于本领域技术人员来说将是明显的。本领域技术人员现在将想到许多变化形式、改变和替换而不偏离本发明。应理解,在实践本发明中可采用本文描述的本发明的实施方案的多种替代选择。意图所附权利要求限定本发明的范围并且在这些权利要求的范围内的方法和结构及其等价物藉此被涵盖。

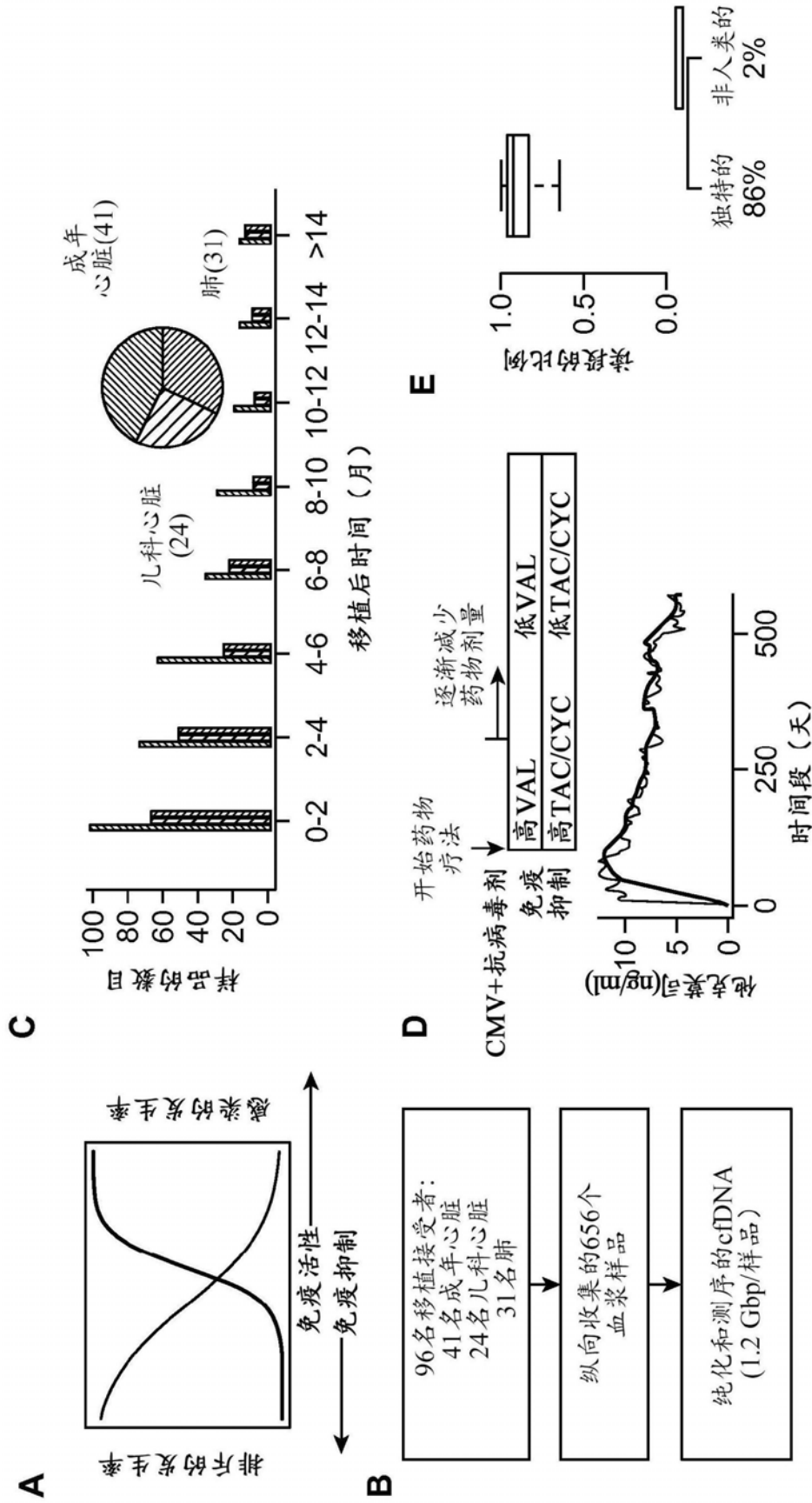


图1

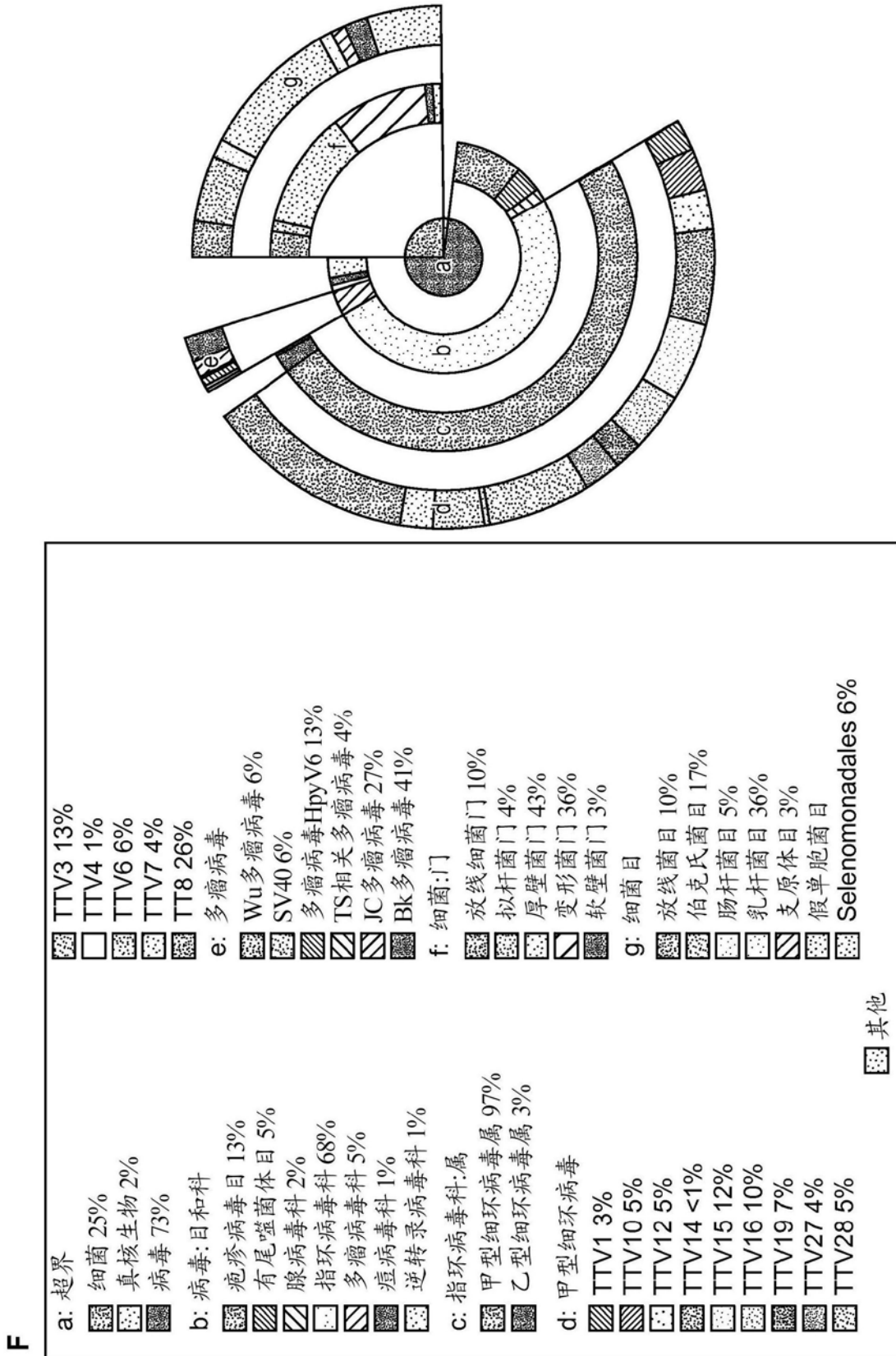


图1(续)

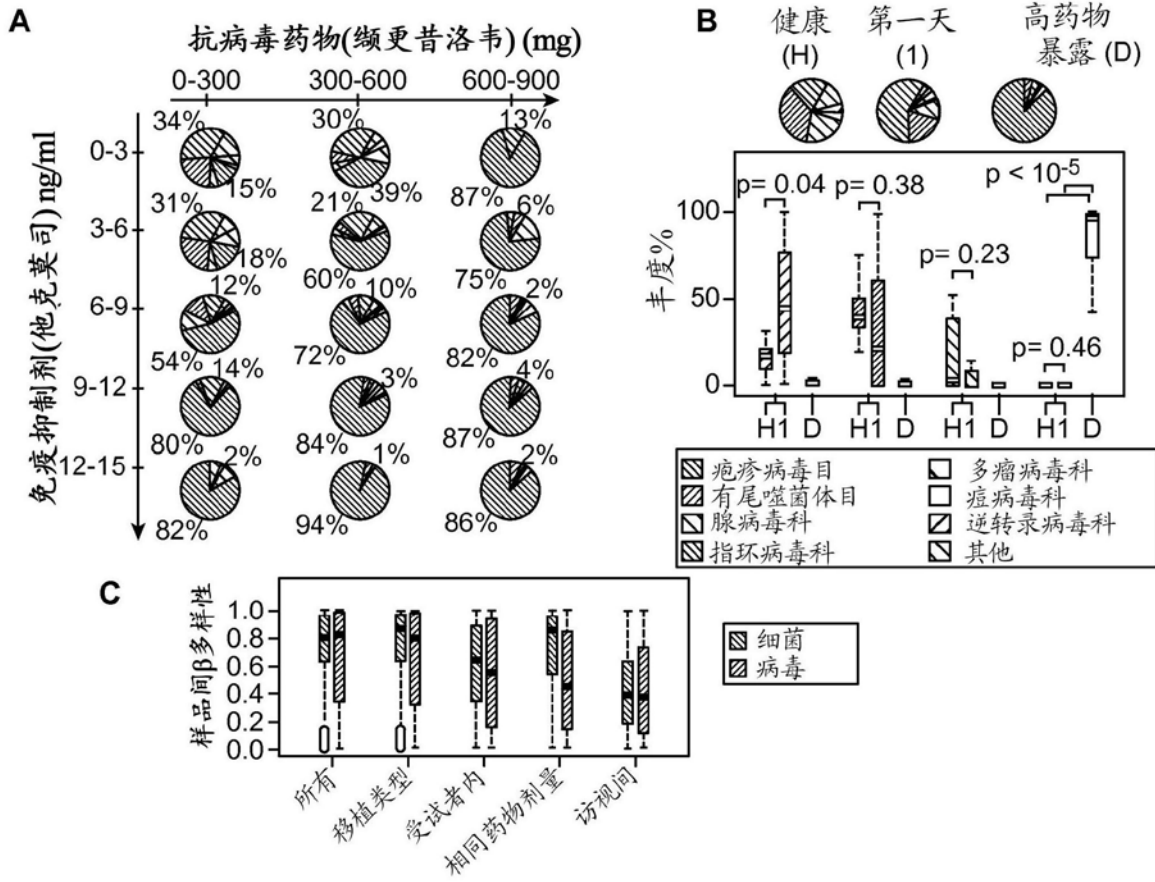


图2

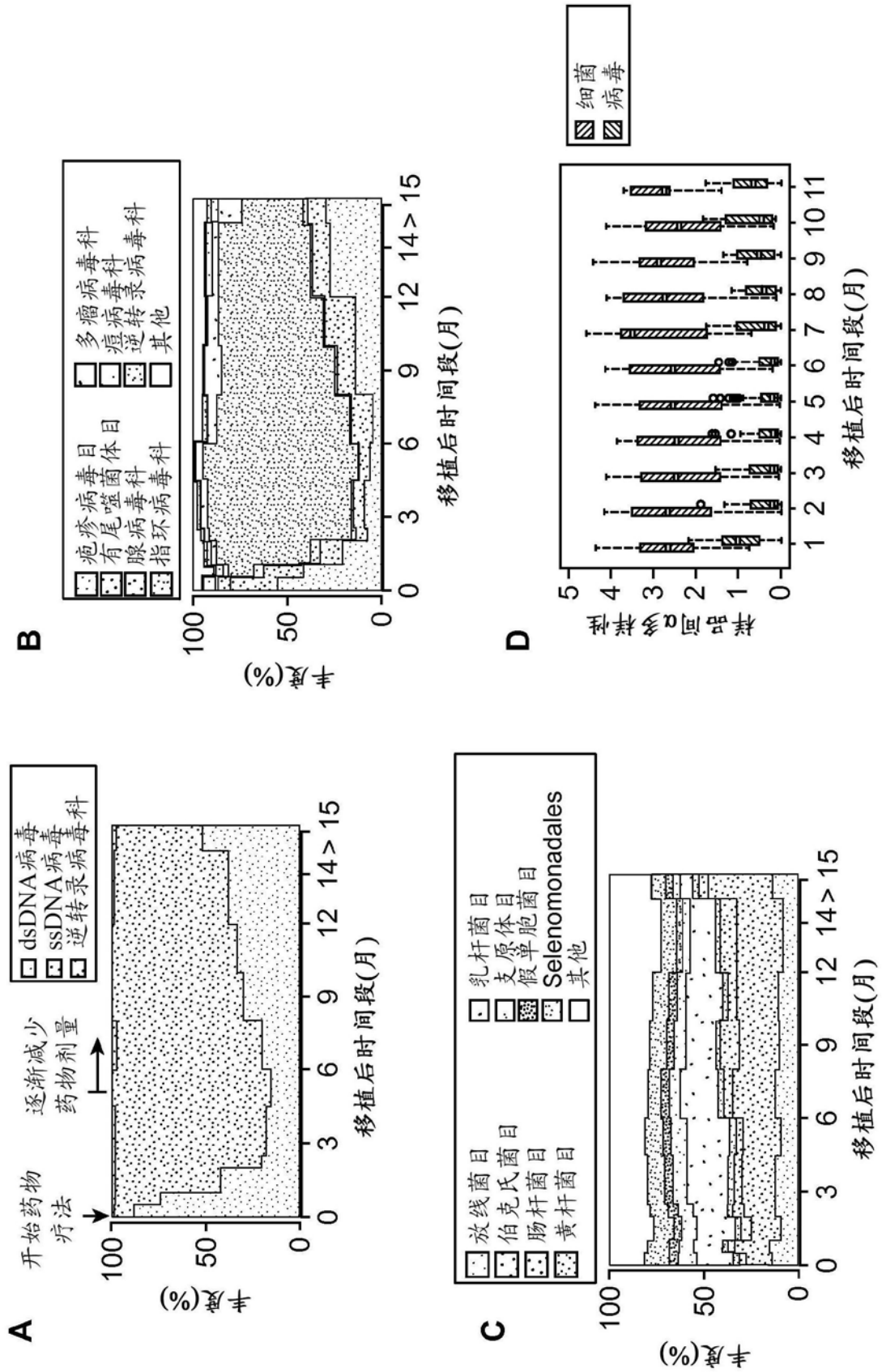


图3

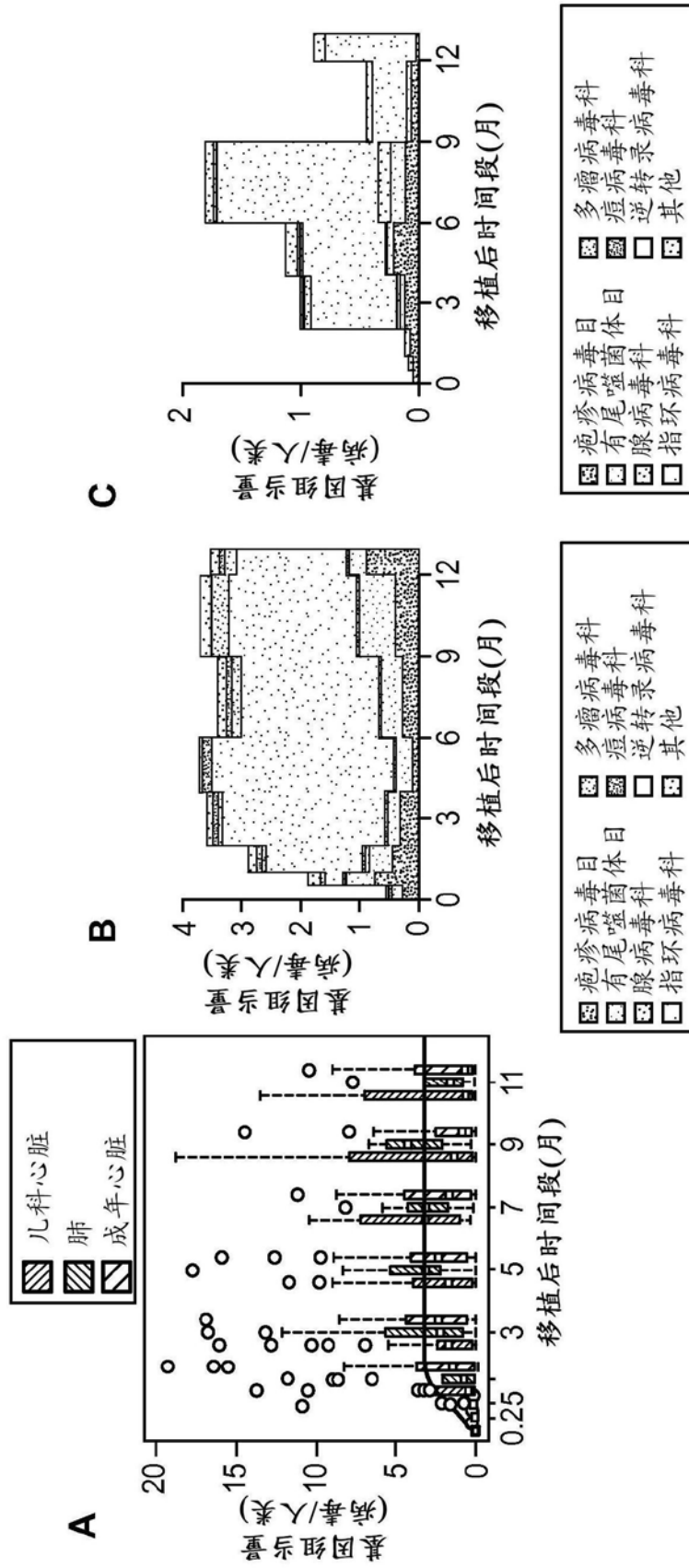


图4

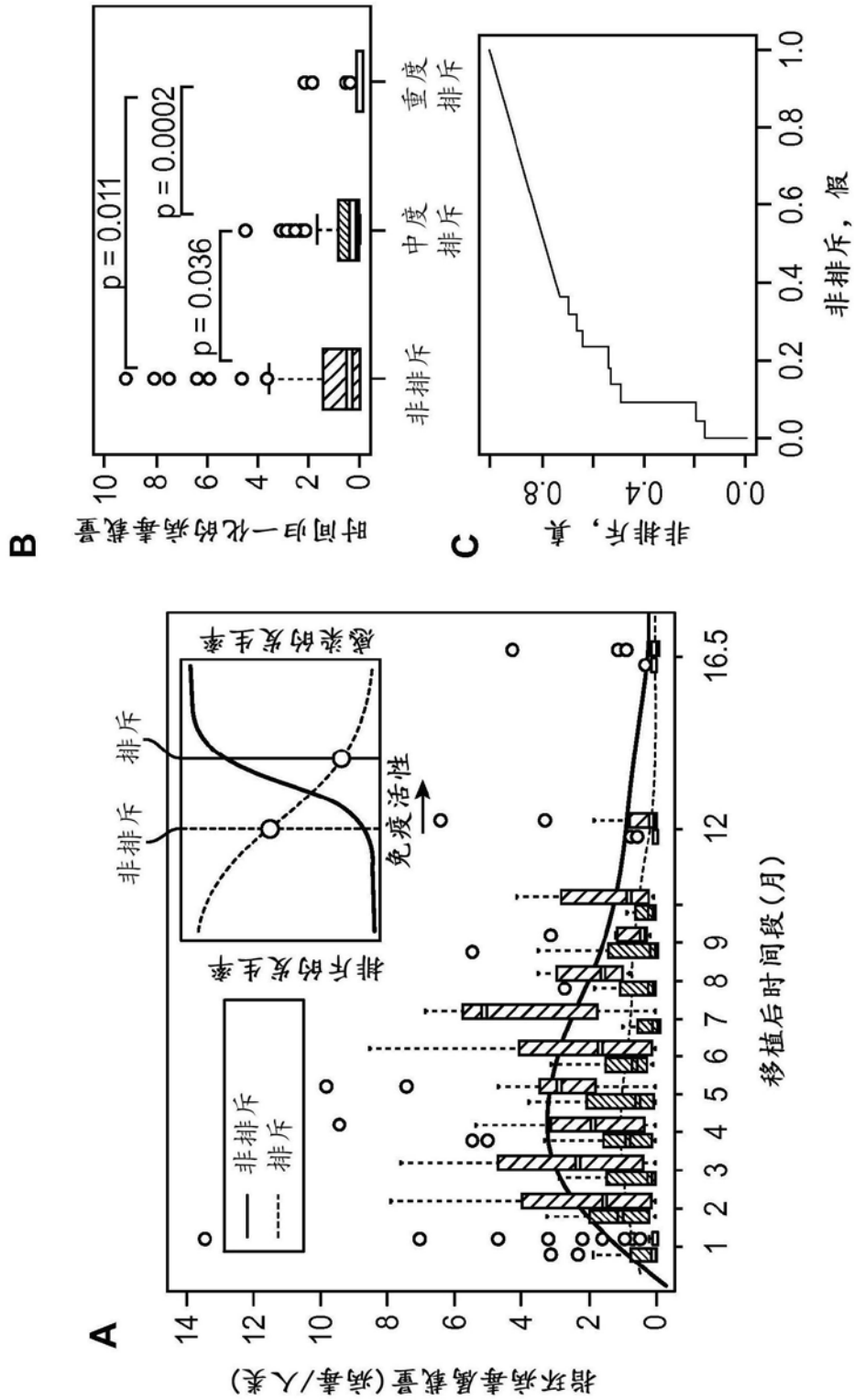


图5

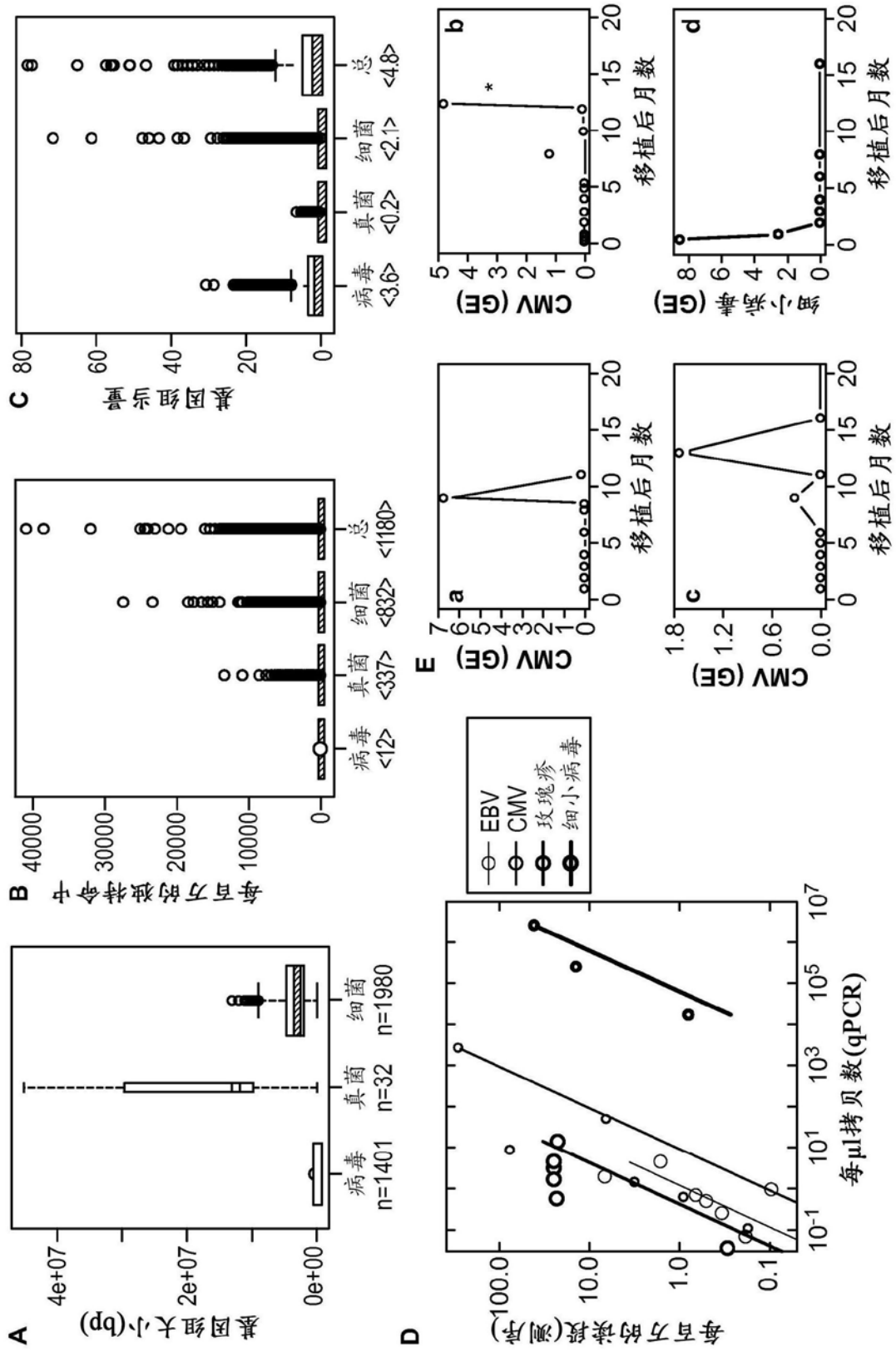


图6

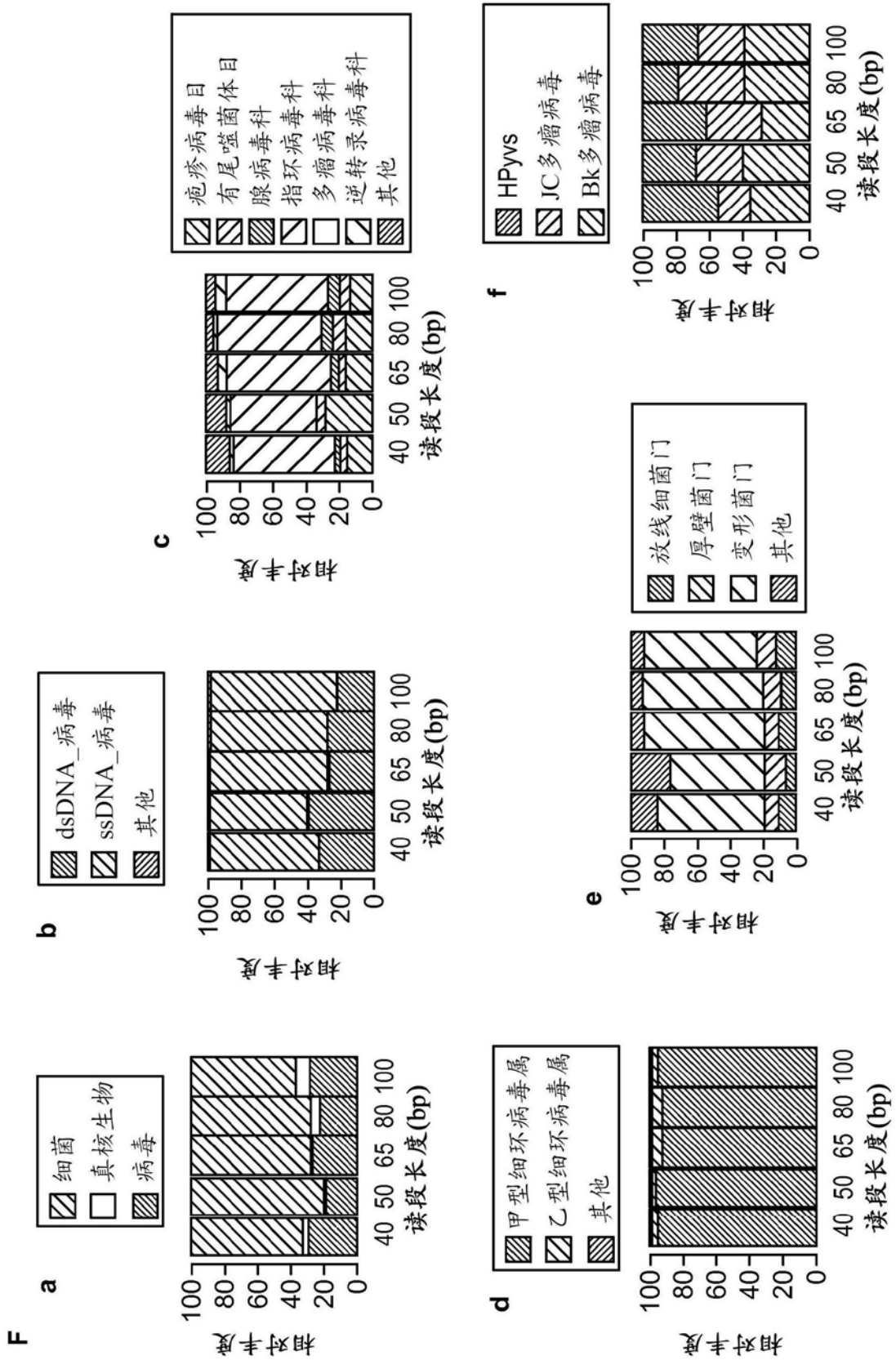


图6(续)

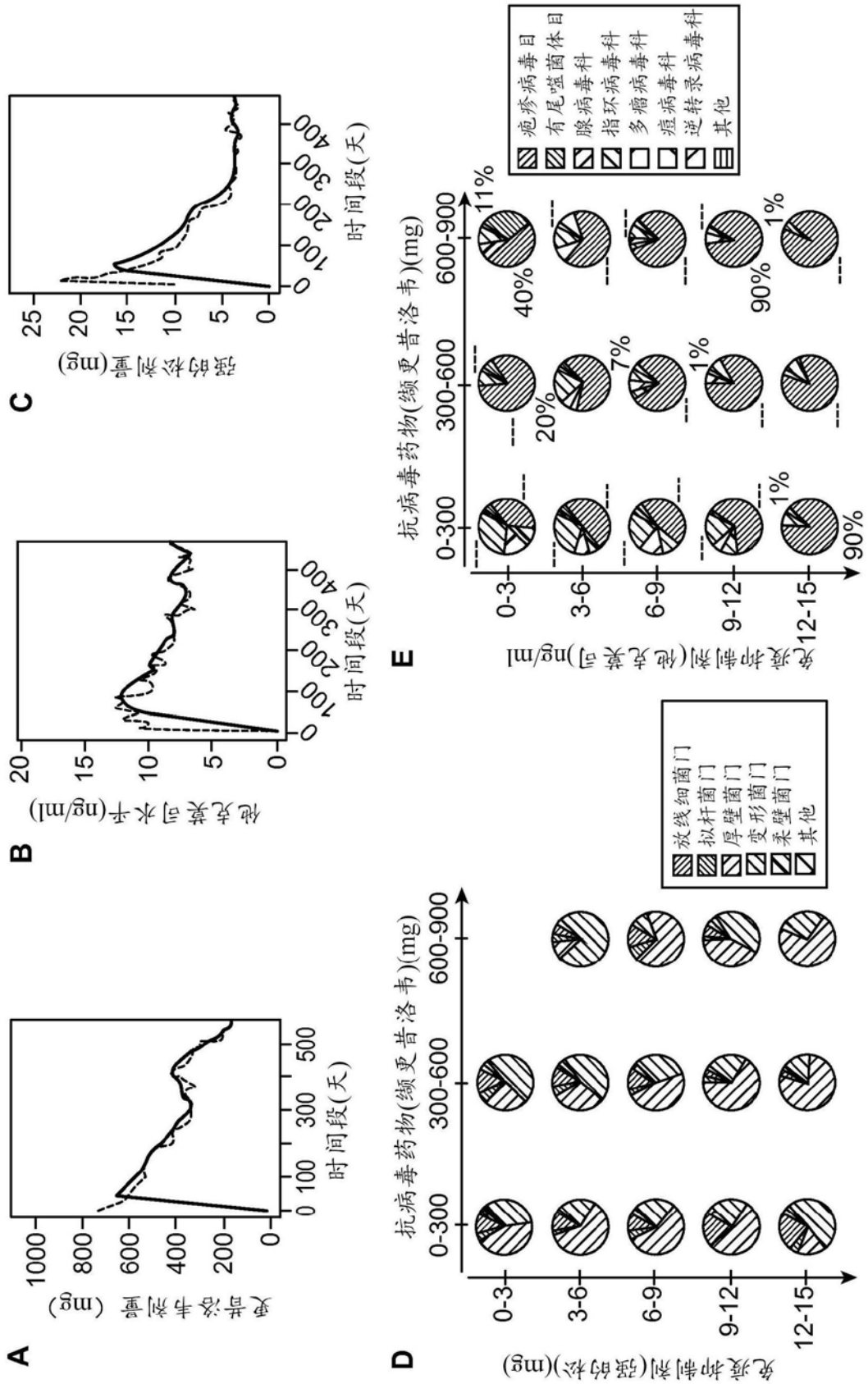


图7

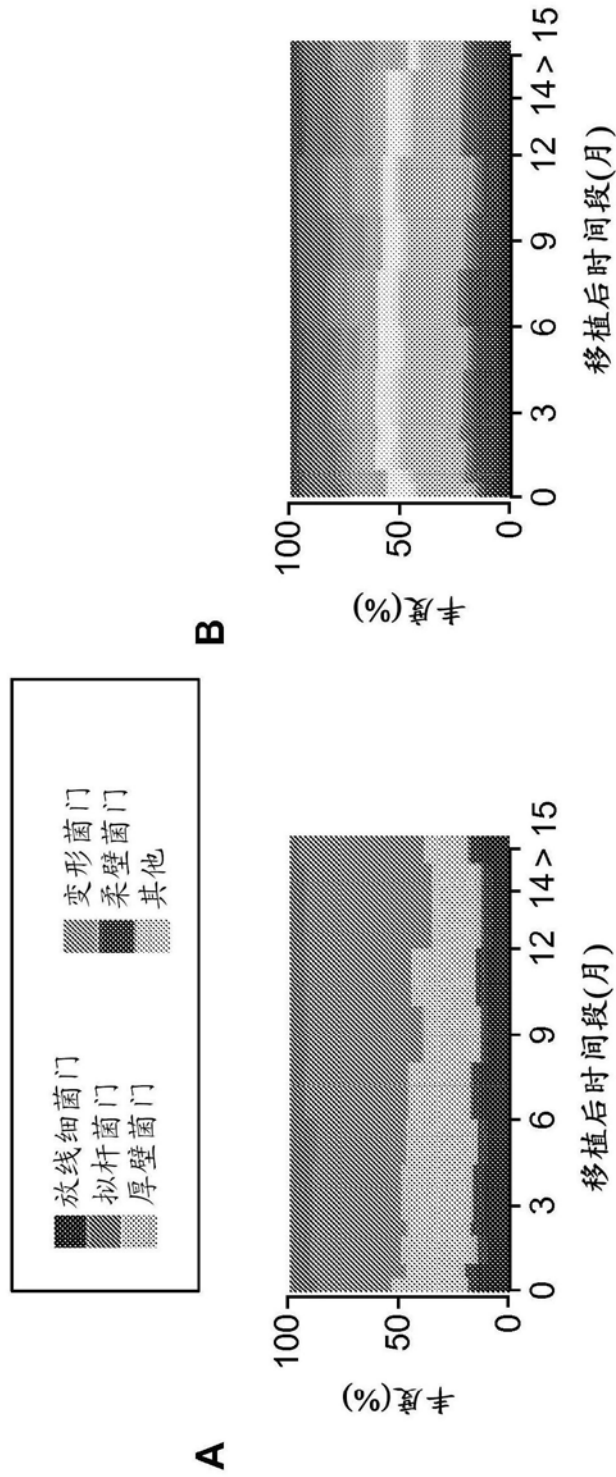


图8

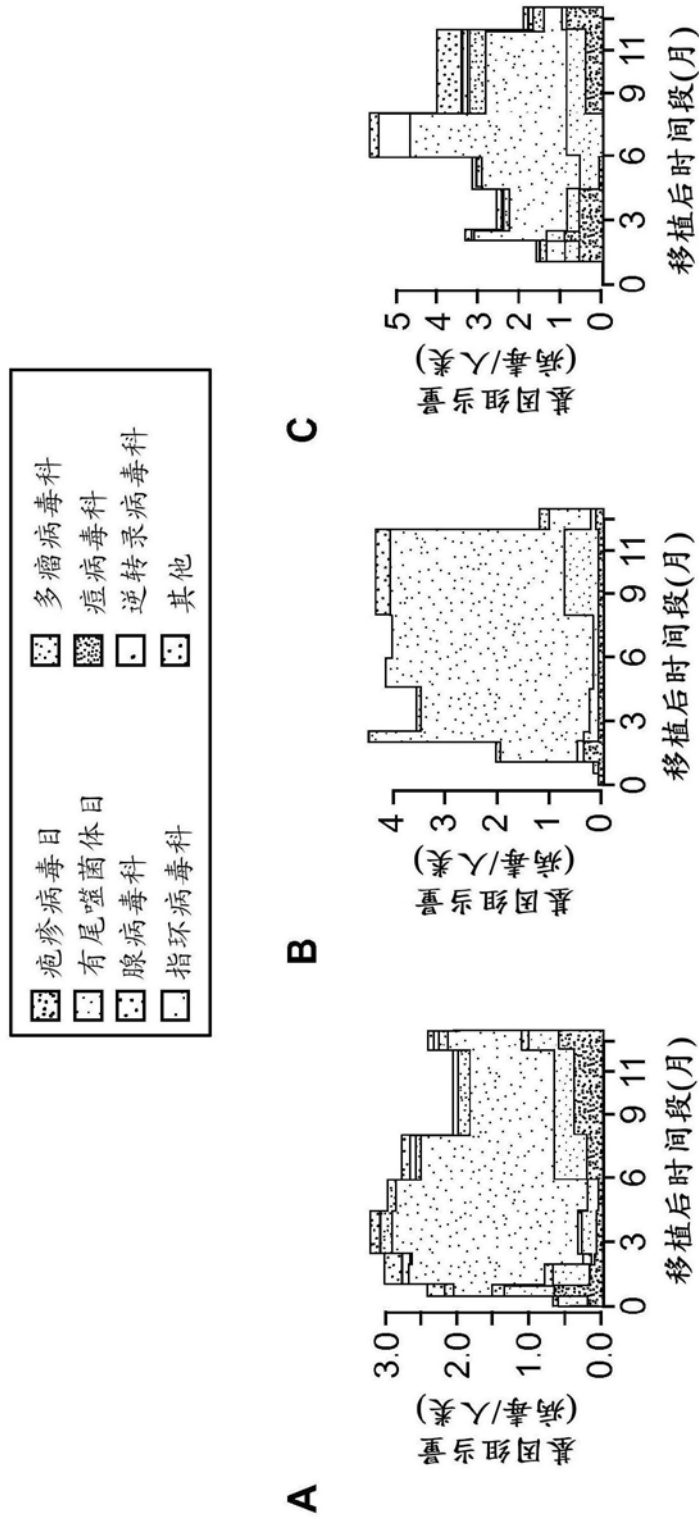


图9

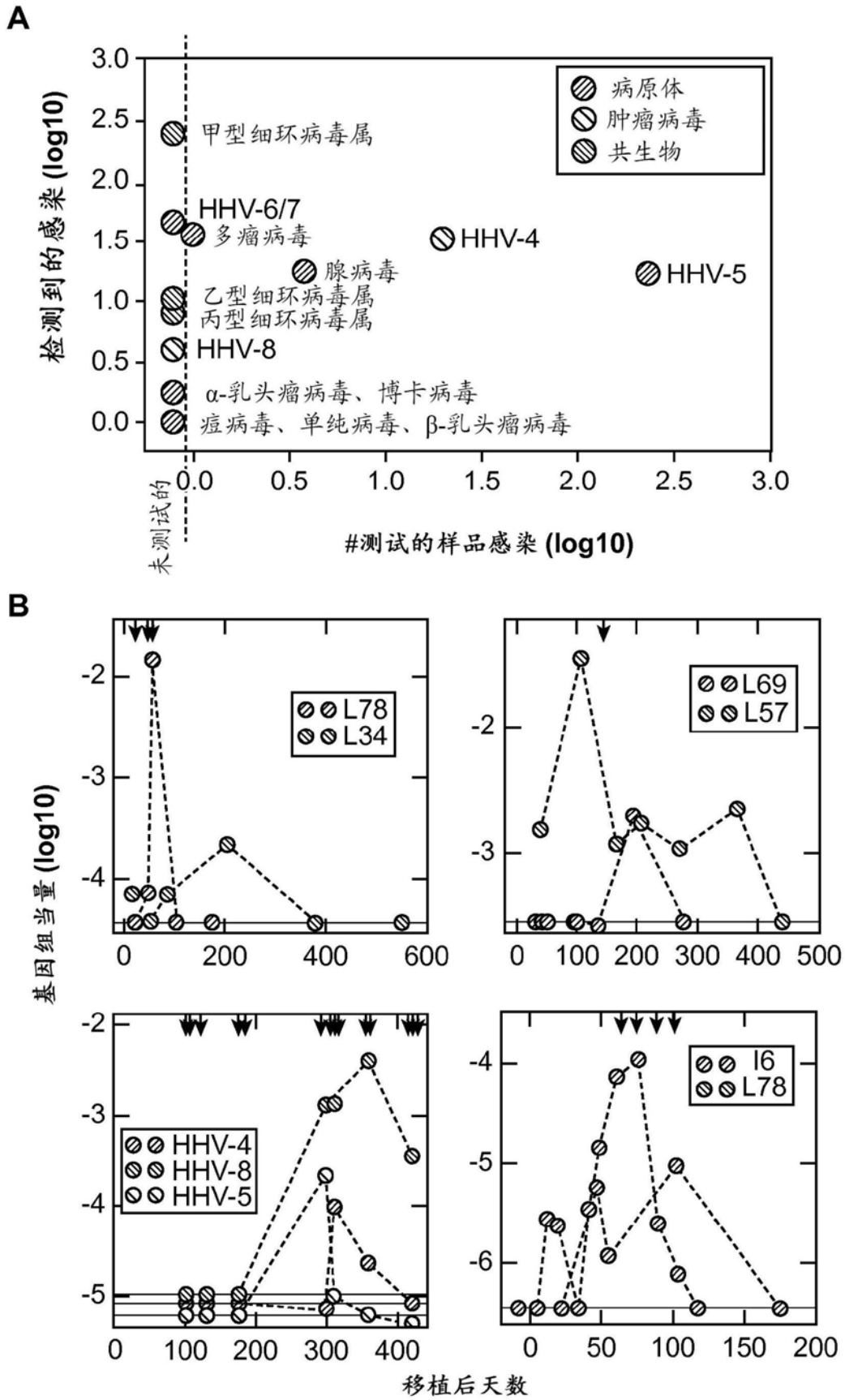


图11