



(12)发明专利

(10)授权公告号 CN 105209918 B

(45)授权公告日 2017.09.29

(21)申请号 201380076431.X

(22)申请日 2013.05.09

(65)同一申请的已公布的文献号

申请公布号 CN 105209918 A

(43)申请公布日 2015.12.30

(85)PCT国际申请进入国家阶段日

2015.11.09

(86)PCT国际申请的申请数据

PCT/CN2013/075406 2013.05.09

(87)PCT国际申请的公布数据

W02014/179965 EN 2014.11.13

(73)专利权人 宝洁公司

地址 美国俄亥俄州

(72)发明人 李睿 黄适 T·何 刘吉泉

徐健

(74)专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 罗文锋 李炳爱

(51)Int.Cl.

G01N 33/569(2006.01)

(56)对比文件

CN 101833613 A,2010.09.15,

Xiao Zhu et al.Possible variation of the human oral bacterial community after wearing removable partial dentures by DGGE.《World J Microbiol Biotechnol》.2012,第28卷

Colombo et al..Impact of Periodontal Therapy on the Subgingival Microbiota of Severe Periodontitis: Comparison between Good Responders and “Refractory” Subjects by the Human Oral Microbe Identification Microarray (HOMIM).《J Periodontol》.2012,第83卷(第10期),

Laksmana et al..Metagenomic Analysis of Subgingival Microbiota Following Non-Surgical Periodontal Therapy: a Pilot Study.《The Open Dentistry Journal》.2012,第6卷

审查员 王在竹

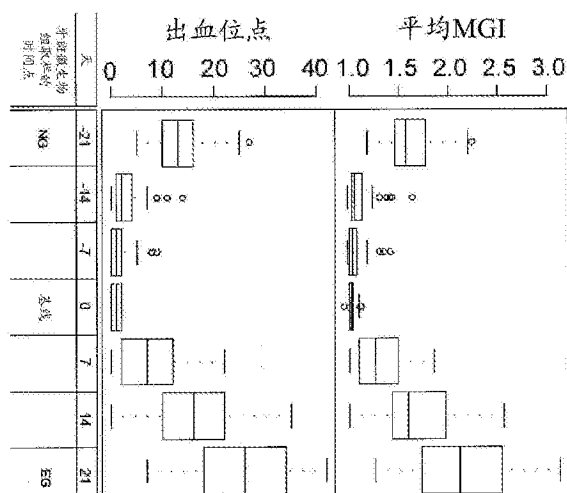
权利要求书5页 说明书25页 附图20页

(54)发明名称

生物标记鉴定方法和系统

(57)摘要

本发明通过设计与口腔微生物群落分析结合的一种减退-发展模型提供了一种鉴定指示受试哺乳动物的状况的生物标记的方法,其中所述状况选自第一疾病的存在、第一疾病的严重性、对第一疾病的敏感性以及它们的组合。还提供了一种鉴定指示受试哺乳动物的状况的生物标记的计算机辅助系统。



1. 一种鉴定指示受试哺乳动物的状况的生物标记的方法,所述方法包括以下步骤:

a) 选择具有第一疾病的第一组测试哺乳动物;其中所述第一疾病是齿龈炎;

b) 获取第一口腔样品,所述第一口腔样品包含来自所述具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型;

c) 治疗已进行第一口腔取样的所述具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物达21天,以便消除或减少所述第一疾病;

d) 获取第二口腔样品,所述第二口腔样品包含来自已进行治疗的第一组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型;

e) 通过避免口腔卫生操作达21天使所述第一疾病在已进行第二口腔取样的所述第一组测试哺乳动物中的每一个哺乳动物中复发;

f) 获取第三口腔样品,所述第三口腔样品包含来自已复发所述第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

g) 通过16S rRNA分析测量所述第一口腔样品、第二口腔样品和第三口腔样品以分别获得在所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;

h) 跨所述第一组测试哺乳动物对所获得的在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,以鉴定其丰度与所述第一组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为第一组微生物类型,其中所述状况选自:所述第一疾病的存在、所述第一疾病的严重性、对所述第一疾病的敏感性以及它们的组合;

i) 从所述第一组微生物类型中选择一种或多种微生物类型作为指示所述受试哺乳动物的状况的生物标记。

2. 根据权利要求1所述的方法,其中在步骤h)中,用成对比较分析或多变量分析对所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析。

3. 根据权利要求2所述的方法,其中所述多变量分析选自主成分分析、主坐标分析、对应分析、降趋对应分析、聚类分析、判别分析、典型判别分析以及它们的组合。

4. 根据权利要求2所述的方法,其中在步骤h)中,用成对比较分析对所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,所述成对比较分析包括以下步骤:

1) 将所述第一组测试哺乳动物中的每一个哺乳动物的所述第一微生物群落和所述第二微生物群落进行比较,以测定在所述第一微生物群落和所述第二微生物群落之间所获得的每种微生物类型的丰度的变化;

2) 跨所述第一组测试哺乳动物将来自步骤1)的所获得的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的初级组;

3) 将所述第一组测试哺乳动物中的每一个哺乳动物的所述第二微生物群落和所述第三微生物群落进行比较,以测定在所述第二微生物群落和所述第三微生物群落之间所获得的每种微生物类型的丰度的变化;

4) 跨所述第一组测试哺乳动物将来自步骤3)的所获得的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的次级组;以及

5) 将所述微生物类型的初级组和所述微生物类型的次级组进行比较,以鉴定出那些重叠的微生物类型作为第一组微生物类型。

5. 根据权利要求2所述的方法,其中在步骤h)中,用多变量分析对所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,所述多变量分析包括以下步骤:

1) 将所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所获得的丰度中方差最大的向量;以及

2) 鉴定出具有所获得的表现出与导出的向量有统计意义上的显著的相关性的丰度的那些微生物类型作为第一组微生物类型。

6. 根据权利要求2所述的方法,其中在步骤h)中,用多变量分析对所获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,所述多变量分析包括以下步骤:

1) 将所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所获得的丰度中方差最大的向量;

2) 将所获得的第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落中的每一种的一种或多种微生物类型的丰度投影至所导出的向量上,以获得针对第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落的每一种的投影值;

3) 针对所述第一组测试哺乳动物的每一个跨第一微生物群落、第二微生物群落和第三微生物群落计算所述投影值的变化速率;

4) 基于所计算出的变化速率,将所述第一组测试哺乳动物分成测试哺乳动物的第一子组和测试哺乳动物的第二子组,其中所述测试哺乳动物的第一子组与所述测试哺乳动物的第二子组相比,表现出更大的变化速率;以及

5) 分别将所述测试哺乳动物的第一子组的第一微生物群落、第二微生物群落和第三微生物群落与所述测试哺乳动物的第二子组的第一微生物群落、第二微生物群落和第三微生物群落进行比较,以鉴定其在第一微生物群落、第二微生物群落和第三微生物群落的每一种中的丰度在测试哺乳动物的第一子组和测试哺乳动物的第二子组间具有统计意义上的显著差异的那些微生物类型作为第一组微生物类型。

7. 根据权利要求1所述的方法,所述方法还包括以下步骤:

1) 选择具有第二微生物相关的疾病的第二组测试哺乳动物;

2) 重复步骤b)至h)以鉴定第二组微生物类型;

3) 将所述第一组微生物类型和所述第二组微生物类型进行比较,以鉴定出那些重叠的微生物类型作为微生物类型的子组;以及

4) 从所述微生物类型的子组选择一种或多种微生物类型作为指示所述受试哺乳动物状况的生物标记,其中所述状况选自:所述第一疾病和所述第二微生物相关的疾病的存在、所述第一疾病和所述第二微生物相关的疾病的严重性、对所述第一疾病和所述第二微生物相关的疾病的敏感性以及它们的组合。

8. 根据权利要求1所述的方法,其中所述微生物类型选自细菌的分类种类、微生物的功能种类以及它们的组合。

9. 根据权利要求8所述的方法,其中所述微生物类型选自细菌门、细菌目、细菌科、细菌纲、细菌属、细菌种、微生物的功能基因、微生物基因同源组、微生物的肽或蛋白质的基序、微生物的保守的肽或蛋白质域、微生物的非编码核苷酸序列以及它们的组合。

10. 根据权利要求1所述的方法,其中通过采集并合并来自选自象限1和3以及象限2和4的两对象限中所有牙齿的龈上牙斑而获得所述第一口腔样品、第二口腔样品和第三口腔样品。

11. 根据权利要求1所述的方法,其中所述统计显著性具有 $p < 0.05$ 的水平。

12. 一种鉴定指示受试哺乳动物的状况的生物标记的计算机辅助系统,其包括:

a) 用于取样的取样部分:

1) 第一口腔样品,所述第一口腔样品包含来自具有微生物相关的疾病的一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型,

2) 第二口腔样品,所述第二口腔样品包含来自已进行治疗以消除或减少所述微生物相关的疾病的该组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型,和

3) 第三口腔样品,所述第三口腔样品包含来自已复发所述微生物相关的疾病的该组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

b) 与取样部分相联的测量部分,其中所述测量部分被配置用于测量第一口腔样品、第二口腔样品和第三口腔样品,以分别获得在所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;和

c) 与所述测量部分相联的计算部分,其中所述计算部分被配置用于跨该组测试哺乳动物接收并统计学分析所获得的在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度,以鉴定其丰度与该组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为指示所述受试哺乳动物的状况的生物标记,

其中所述状况选自:所述微生物相关的疾病的存在、所述微生物相关的疾病的严重性、对所述微生物相关的疾病的敏感性以及它们的组合。

13. 根据权利要求12所述的计算机辅助系统,其中所述计算部分包括:

1) 与所述测量部分相联的输入模块,其中所述输入模块用于输入所获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;

2) 与所述输入模块相联的数据处理模块,其中所述数据处理模块被配置用于跨该组测试哺乳动物对所输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,以鉴定其丰度与所述状况具有统计显著性的相关

性的那些微生物类型;以及

3) 与所述数据处理模块相联的输出模块,其中所述输出模块用于显示那些经鉴定的、作为指示所述受试哺乳动物的状况的生物标记的微生物类型。

14. 根据权利要求13所述的计算机辅助系统,其中所述数据处理模块被配置用于对所输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行成对比较分析或多变量分析。

15. 根据权利要求14所述的计算机辅助系统,其中所述多变量分析选自主成分分析、主坐标分析、对应分析、降趋对应分析、聚类分析、判别分析、典型判别分析以及它们的组合。

16. 根据权利要求14所述的计算机辅助系统,其中所述数据处理模块被配置用于对所输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行成对比较分析,所述成对比较分析包括:

1) 将该组测试哺乳动物中的每一个哺乳动物的所述第一微生物群落和所述第二微生物群落进行比较,以测定在所述第一微生物群落和所述第二微生物群落之间所输入的每种微生物类型的丰度的变化;

2) 跨该组测试哺乳动物将来自步骤1)的所输入的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的初级组;

3) 将该组测试哺乳动物中的每一个哺乳动物的所述第二微生物群落和所述第三微生物群落进行比较,以测定在所述第二微生物群落和所述第三微生物群落之间所输入的每种微生物类型的丰度的变化;

4) 跨该组测试哺乳动物将来自步骤3)的所输入的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的次级组;以及

5) 将所述微生物类型的初级组和所述微生物类型的次级组进行比较,以鉴定出那些重叠的微生物类型。

17. 根据权利要求14所述的计算机辅助系统,其中所述数据处理模块被配置用于对所输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行多变量分析,所述多变量分析包括:

1) 将所输入的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所输入的丰度中方差最大的向量;以及

2) 鉴定出具有表现出与所导出的向量有统计意义上的显著的相关性的所输入的丰度的那些微生物类型。

18. 根据权利要求14所述的计算机辅助系统,其中所述数据处理模块被配置用于对输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行多变量分析,所述多变量分析包括:

1) 将所输入的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所输入的丰度中方差最大的向量;

2) 将所输入的该组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落中的每一种的一种或多种微生物类型的丰度投影至导出的向量

上,以获得针对该组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落的中的每一种的投影值;

3) 针对该组测试哺乳动物中的每一个哺乳动物跨第一微生物群落、第二微生物群落和第三微生物群落计算所述投影值的变化速率;

4) 基于所计算出的变化速率,将该组测试哺乳动物分成测试哺乳动物的第一子组和测试哺乳动物的第二子组,其中所述测试哺乳动物的第一子组与所述测试哺乳动物的第二子组相比,表现出更大的变化速率;以及

5) 分别将所述测试哺乳动物的第一子组的第一微生物群落、第二微生物群落和第三微生物群落与所述测试哺乳动物的第二子组的第一微生物群落、第二微生物群落和第三微生物群落进行比较,以鉴定其在所述第一微生物群落、第二微生物群落和第三微生物群落中的每一种中的丰度在所述测试哺乳动物的第一子组和所述测试哺乳动物的第二子组间具有有统计意义上的显著差异的那些微生物类型。

生物标记鉴定方法和系统

技术领域

[0001] 本发明涉及鉴定指示受试的哺乳动物状况的生物标记的方法。本发明另外涉及鉴定指示受试的哺乳动物状况的生物标记的计算机辅助系统。

背景技术

[0002] 通常根据多种症状评估受试者的健康状况。然而，今天所用的许多症状由于其主观性描述和与疾病状态的不确定关系而令人误解。

[0003] 术语“生物标记(生物学标记)”是在1989年作为医学主题词(MeSH)术语引入的，被定义为一种客观的测量和评价的特性，作为正常的生物过程、致病过程，或治疗干预的药物反应的指标。在过去数十年间生物标记的发现有了巨大的发展。生物标记在药用生物学中起到了重要作用。生物标记有助于早期诊断、疾病预防、药物靶点识别、药物反应等。已经识别出针对许多疾病的若干生物标记，诸如针对胆固醇的血清LDL、血压、针对癌症的P53基因(Loukopoulos P, Thornton JR, Robinson WF (2003年5月)。“Clinical and pathologic relevance of p53 index in canine osseous tumors”。Vet. Pathol. 40 (3): 237-48. doi: 10.1354/vp.40-3-237) 和 MMPs (Loukopoulos P, Mungall BA, Straw RC, Thornton JR, Robinson WF (2003年7月)。“Matrix metalloproteinase-2 and-9 involvement in canine tumors”。Vet. Pathol. 40 (4): 382-94. doi: 10.1354/vp.40-4-382) 等。在20世纪90年代中期DNA微阵列的引入产生了转录组学的革命并引发生命科学家研究方式的重大转变。随后，代谢组学和代谢物组学，起初主要应用于安全相关的生物标记，开始转向疾病相关的生物标记。

[0004] 科学家持续努力以发现与健康或疾病的深层原因密切关联的新的生物标记。他们的发现是通过给医生们一个更客观和可量化的临床决策的基础 来改变医学实践。口腔内发现的微生物群落，其结构或功能随病情的进展而变化，提供最有价值的线索中的一个。

[0005] 口腔是微生物定植的主要位点。口腔微生物群落在不同个体间、相同口腔内的不同位点间或相同位点的不同时间点间变动。微生物群落的差异决定了口腔微生物生态系统的平衡，其与口腔健康状态和甚至整个系统健康状态直接相关。然而发现生物标记必须从微生物群落中数以万计的微生物类型进行选择，这提出了挑战。

[0006] 齿龈炎(其涉及牙周软组织的炎症)是最普遍的感染和最常见的人类口腔疾病中的一种。作为一个世界性的健康问题，它影响了大多数儿童和青少年。这种疾病据信是牙斑的积累和随后的牙斑微生物和宿主组织间相互作用所引起的。虽然没有发生结合上皮的根向移动，在探查时这些组织变成结节状红斑并且出血。此外，慢性齿龈炎可发展成牙周炎，这是不可逆的牙周感染，其特征不在于齿槽骨的损耗、附件损耗、形成牙周腔洞和最终牙缺失。因此，针对齿龈炎的预防性测量，以及用于其预后和早期诊断的改善的工具具有特别重要的临床意义。

[0007] 若干因素阻碍了齿龈炎的病原性调查，这仍然知之甚少。在实际人群中，齿龈炎症状可为可逆的和不稳定的，因为多个内部或外部因素，包括口腔卫生操作(个人的或职业

的)、免疫系统损伤、受伤、饮食和口腔状态,可潜在地影响疾病发展从而混淆疾病监测。此外,齿龈炎的临床诊断是基于单个观察并且是由人类检查者进行判断。因此,很难在不同病人和不同检查者之间比较结果。此外,虽有口腔微生物群落和慢性口腔感染的疑似多重细菌性质的复杂性,齿龈炎相关联的微生物群的族群调查通常限于仅少数可培养的细菌(例如“红色复合物”包括牙龈卟啉单胞菌(*Porphyromonas gingivalis*)、福赛斯坦纳菌(*Tannerella forsythia*)和齿垢密螺旋体(*Treponema denticola*)),这些细菌提供的数据点对于可能引起齿龈炎的各种微生物的彻底分析是不够的。

[0008] 因此,仍然需要改进诊断方法用于评估受试者的健康状况。仍然需要调查疾病的病源学。仍然需要识别能够充当更敏感、更可靠和更客观的疾病测量的生物标记。仍然需要准确测定受试者对于疾病的易感性以便阻止和控制不期望的状况和疾病。

发明内容

[0009] 为了应对这些挑战和/或需求,设计了与口腔微生物群落的分析结合的减退-发展模型(RPM)以调查疾病的病源学。

[0010] 在一个方面,本发明涉及一种鉴定指示受试哺乳动物的状况的生物标记的方法,所述方法包括以下步骤:

[0011] a) 选择具有第一疾病的第一组测试哺乳动物;

[0012] b) 获取第一口腔样品,所述第一口腔样品包含来自具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型;

[0013] c) 治疗已进行第一口腔取样的具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物,以便消除或减少第一疾病;

[0014] d) 获取第二口腔样品,所述第二口腔样品包含来自经过治疗的第一组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型;

[0015] e) 使第一疾病在已进行第二口腔取样的第一组测试哺乳动物中的每一个哺乳动物中复发;

[0016] f) 获取第三口腔样品,所述第三口腔样品包含已复发第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

[0017] g) 测量所述第一口腔样品、第二口腔样品和第三口腔样品以分别获得在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;

[0018] h) 跨第一组测试哺乳动物对获得的在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,以鉴定其丰度与第一组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为第一组微生物类型,其中所述状况选自:第一疾病的存在、所述疾病的严重性、对所述疾病的敏感性以及它们的组合;

[0019] i) 从所述第一组微生物类型中选择一种或多种微生物类型作为指示所述受试哺乳动物的状况的生物标记。

[0020] 在另一方面,本发明涉及鉴定指示受试哺乳动物的状况的生物标记的计算机辅助系统,其包括:

[0021] a) 用于取样的取样部分:

[0022] 1) 第一口腔样品,所述第一口腔样品包含来自具有第一疾病的一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型,

[0023] 2) 第二口腔样品,所述第二口腔样品包含来自经过治疗以消除或减少所述疾病的该组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型,和

[0024] 3) 第三口腔样品,所述第三口腔样品包含已复发所述疾病的该组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

[0025] b) 与取样部分相联的测量部分,其中所述测量部分被配置用于测量第一口腔样品、第二口腔样品和第三口腔样品,以分别获得在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;和

[0026] c) 与测量部分相联的计算部分,其中所述计算部分被配置用于跨该组测试哺乳动物接收并统计学分析所获得的在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度,以鉴定其丰度与该组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为指示所述受试哺乳动物的状况的生物标记,

[0027] 其中所述状况选自:第一疾病的存在、所述疾病的严重性、对所述疾病的敏感性以及它们的组合。

[0028] 本方法和系统是基于以下概念:平衡的口腔环境是口腔健康和卫生状态的指示灯,特别是就微生物群落的平衡而言。通过将RPM和口腔微生物群落分析相结合实现了本发明。RPM是减退-发展模型,其包括两个区段,即,从疾病状态至健康状态的第一区段和从健康状态至疾病复发状态的第二区段。微生物群落分析在过去被限制在仅一个区段中,例如从疾病状态至健康状态或从健康状态至疾病状态。通过将RPM和口腔微生物群落分析相结合,本发明提供了鉴定生物标记的有效方法,所述生物标记能用作健康状态的客观的、可再现的和敏感的度量。

[0029] 对于本领域的技术人员来说,通过下面的详细描述,本发明的这些和其它的特征、方面和优点将变得显而易见。

附图说明

[0030] 虽然在说明书之后提供了特别限定和清楚地要求保护本发明的权利要求书,但据信通过以下附图说明可更好地理解本发明。在这些附图中,

[0031] 图1A示出了根据本发明的具体实施例模拟人群中齿龈炎发展的纵向研究的设计。图1B示出了针对50个受试者在整个研究的不同时间点某些临床参数的值。

[0032] 图2示出了根据本发明的具体实施例所述方法的研究流水线。

[0033] 图3示出了根据本发明的具体实施例,27个属水平上的生物标记丰度,所述生物标记区分50个受试者中健康状态和齿龈炎状态(同时包括天然存在的齿龈炎状态和实验诱导

的齿龈炎状态)。

[0034] 图4A和4B示出了根据本发明的具体实施例,针对从50个受试者在三种不同状态下(即,天然存在的齿龈炎状态(“NG”)、基线状态(“基线”)和实验诱导的齿龈炎状态(“EG”))收集的150个口腔样品,来自测得的属级上的细菌数据的主成分分析(PCA)的主成分1和2(PC1和PC2)的曲线图。

[0035] 图5A和5B示出了根据本发明的具体实施例,牙斑微生物群的生物体结构的主坐标分析(PCoA)。每个点对应于微生物群落。图5A是基于UniFrac距离的;并且图5B是基于thetaYC距离的。

[0036] 图6A、6B和6C是根据本发明的具体实施例,示出了通过PCA分析鉴定的15个齿龈炎驱动属之间的相互作用的相关性网络。

[0037] 图7A示出了健康状态和齿龈炎微生物群之间的功能差异;图7B示出了16S rRNA基因序列对相邻类的聚簇(Clusters of Orthologous Groups, COG)的普式叠合分析(Procrustes analysis);图7C示出了33个齿龈炎富集的相邻类(orthologous groups, OG, 其编码鞭毛生物合成途径的组分)。

[0038] 图8A、8B、8C和8D示出了根据本发明的具体实施例,两种具有不同齿龈炎敏感性的宿主类型的鉴定。图8A示出了微生物群结构(即PC1-值)变化的模式和沿RPM的Mazza齿龈指数变化。图8B示出了50个受试者沿PCA的主成份1和2(PC1和PC2)的分布,其中垂直的虚线将50名受试者分成I型和II型宿主。图8C示出了在I型和II型宿主间的齿龈炎敏感性的差异。图8D示出了区分I型和II型宿主的8个属级细菌生物标记的丰度。

[0039] 图9示出了基于齿龈炎的存在、使用齿龈炎微生物指数MiG27的试用分类,所述MiG27是由基于根据本发明的具体实施例鉴定的27个生物标记的丰度的函数计算出来的。

[0040] 图10示出了基于齿龈炎的严重程度、使用齿龈炎微生物指数MiG15的试用分类,所述MiG15是由基于根据本发明的另一个具体实施例鉴定的15个生物标记的丰度的函数计算出来的。

[0041] 图11示出了基于对齿龈炎的敏感性、使用齿龈炎微生物指数MiG-S的试用分类,所述MiG-S是由基于根据本发明的另外的具体实施例鉴定的8个生物标记的丰度的函数计算出来的。如左图所示通过基于牙斑微生物群的(即基于MiG-S的)齿龈炎敏感的宿主类型分类的ROC(受试者操作特性)曲线下方的面积来测量MiG-S的准确性。

具体实施方式

[0042] 定义

[0043] 如本文所用,术语“哺乳动物”是指各种哺乳动物纲的温血脊椎动物的任一种,包括人类。在本文上下文中,哺乳动物也可被称为“受试者”或“宿主”。

[0044] 如本文所用,术语“一组哺乳动物”是指出于研究目的将多个哺乳动物聚成一组。组内的数量可为任一大于1的可数数目。根据具体研究的目的准确度的需要,组内哺乳动物的数目可多达1000,10000或甚至更大。

[0045] 如本文所用,术语“微生物群落”、“微生物群”、“微生物区系”、“微生物菌丛”和“群落”在本文互换使用,是指一组各种不同的通常栖息于身体器官或部分的微生物。术语“微生物”意为微观或亚微观尺寸的有机体,特别是细菌或原生动物,更优选地为细菌。

[0046] 如本文所用,术语“微生物相关的疾病”包括由微生物导致的或影响的或相关联的疾病。

[0047] 如本文所用,术语“样品”、“口腔样品”、或“生物样品”是从受试者分离的生物材料,用于按照本方法进行分析,诸如唾液、龈沟液(GCF)、龈上菌斑、龈下菌斑、吸入或呼出的空气、口腔灌洗、舌刮片、口腔组织拭子或活检和血清。理想的样品能够包含微生物群落。

[0048] 如本文所用,术语“统计显著性”是一种数学工具,用于确定实验的结果是具体因素间关系的结果或仅为随机的结果。统计显著性被用于拒绝或接受所谓的零假设。假设是研究人员试图证明的解释。零假设通常认为,研究者所研究的因素对数据上的差异并没有影响,或这些因素之间没有联系。统计显著性通常写作例如 $t=0.02$ 、 $p<0.05$ 。这里,“t”表示检验分数,并且“ $p<0.05$ ”意为事件偶然发生的概率小于5%。这些数字会使得零假设被拒绝。

[0049] 如本文所用,参考疾病或状况,术语“敏感性”和它的形容词形式“敏感的”可与“易感性”和其形容词形式“易感的”互换使用,表示在暴露于有害刺激或病原体时患疾病或病症的可能性。

[0050] 如本文所用,当用于权利要求中时,冠词“一个”和“一种”被理解为是指一种或多种受权利要求书保护的或根据权利要求书所述的物质。

[0051] 如本文所用,术语“包含”、“包括”、“含有”、“具有”的含义是非限制性的,即可加入不影响结果的其它步骤和其它节段。以上术语涵盖术语“由……组成”和“基本上由……组成”。

[0052] 减退-发展模型(RPM)

[0053] 本发明是基于减退-发展模型(RPM)的,其被设计来模拟哺乳动物病害的退化和复发。在代表自然发病状态、健康状态和疾病复发状态的三个不同时间点获取口腔样品。因此,RPM可以用于显示出受试者内和自然种群中两者的微生物群的异质性源。

[0054] 该RPM显示出在受试者内如上文所述的不同时间点的和对疾病具有不同敏感性的受试者之间微生物群的异质性源。在任一种情况下,如它们的微生物属性所反映的,宿主中健康状态和疾病状态间没有清晰的界限:它们(微生物)的分布以及它们的退行性或渐进模式不是分立的而是一个梯度样(渐变)的过程。不受理论的束缚,据信从相对健康的状态发展至疾病状态主要是由某些细菌驱动的,在此类发展过程中其大多数丰度增加而其中一些丰度降低。因此,RPM可用于模拟微生物相关疾病的减退和复发,其优选地但不一定是口腔疾病。

[0055] 根据具体的实施例,所述生物相关疾病选自:齿龈炎、牙周炎、龋齿、口臭、口腔溃疡、早产、低出生体重、糖尿病、呼吸系统疾病、心脏病、脑卒中、菌血症、全身健康问题以及它们的组合。

[0056] 样品收集和储存

[0057] 根据具体的状况,所述口腔样品,优选地为牙齿、假牙(当存在时)、齿龈和舌表面的生物膜形式,可选自唾液样品、牙斑样品、舌背样品、舌苔样品、粘膜样品以及它们的组合。牙斑样品可来自各种位置。例如,所述牙斑样品可选自龈上牙斑样品、龈下牙斑样品、牙齿牙斑样品以及它们的任何组合。所述样品的选择对于鉴定生物标记的准确性是至关重要的。例如,据信与唾液微生物群相比,牙斑微生物群对齿龈炎更敏感。因此,就齿龈炎而言,

口腔样品优选地是牙斑样品。

[0058] 只要疾病可以被消除或减少,就可通过任何方法实现所述疾病的治疗。例如,在治疗学上有效的方案后,可对受试者施用治疗学上有效量的药剂和/或治疗剂。在具体的实施例中,为受试哺乳动物提供了能够消除或减少病害或状况的优质牙膏和特别的牙刷。受试者被要求特定周期(例如两周)内每天刷牙两次。

[0059] 只要所述疾病可复发,就可通过任何方法实现所述疾病的复发。在大多数情况下,对受试者身体的病害相关部分什么都不做就能使所述疾病复发。例如,就齿龈炎而言,受试者可简单地遵循以下方案:受试者不进行任何口腔卫生操作,其包括刷牙、用任何产品漱口,使用牙线和牙病预防。任选地,受试者可在就寝时间使用糖类或其它合适的细菌食物。口腔细菌过夜利用所述糖类并产生升高水平的细菌代谢物。

[0060] 一旦采集,所述样品可立即用于后续步骤。另选地,所述样品可置于冷冻机中供今后使用。在一些情况下,将新采集的样品立即低温冷冻,通常低于 -20°C ,优选地低于 -50°C ,更优选地低于 -70°C ,并且最优选地低于 -90°C 。样品保持冷冻状态,直到准备分析。

[0061] 微生物群落分析

[0062] 口腔存在着各种不同的、丰富而复杂微生物群落。高度多样的微生物区系栖息于正常口腔的各种表面。细菌同时聚积在硬质和软质口腔组织上的生物膜中。对于口腔细菌,细菌粘附是尤其重要的。

[0063] 口腔细菌已经进化出机制以感测其环境并逃避或修饰宿主。细菌占据的齿面和齿龈的上皮提供的生态位。直到最近,根据多个单一菌株相互作用来考虑宿主和口腔细菌之间缔合。然而,越来越明显的是口腔微生物包含复杂的群落,口腔健康或疾病取决于宿主和作为一个整体的微生物群落之间的相互作用。虽然继续研究特定微生物的致病特性是重要的,但只有在它们驻留其中的群落特性的语境下这些才是相关的。理解驱动疾病或健康的微生物群落是防治微生物相关疾病的关键。

[0064] 用于跟踪和诊断宿主状况(疾病、饮食等等)的人微生物群的可能性依赖于并受制于在群体水平上微生物群和状况之间联系的异质性程度。在肠道中,宿主间微生物群结构的变化似乎主宰了状况间的变化(例如瘦或肥胖,或采用正常或高脂肪饮食)。然而,现在本发明的发明人惊讶地发现对于口腔微生物群相反的情况看起来是对的,在受试者内健康的和疾病的口腔微生物群之前的差异大于个体间差异。虽然响应的在不同的身体内栖息的微生物群落大小差异的机制未知,但发明人的发现表明作为针对口腔以及或许甚至系统疾病的生物标记,口腔微生物群可能具有某些优势。

[0065] 因此,根据本发明,对在三个不同时间点的口腔样品(其代表自然疾病状态、健康状态和疾病复发状态)进行测量和对比以鉴定与状况具有统计显著性的丰度相关性的那些微生物类型,其中所述状况选自:疾病的存在、疾病的严重性、疾病的敏感性以及它们的组合。

[0066] 可使用许多技术来测量所述口腔样品以获取口腔微生物群落结构、功能和动态数据。在一方面,通过选择特定微生物群体,可使用基于培养的方法来调查自然和人类活动影响的环境中的微生物生态学。表征微生物生态学的标准培养技术涉及分离和使用商业生长培养基(诸如Luria-Bertani培养基、Nutrient Agar和Tryptic Soy Agar)表征微生物。基于培养的技术的主要限制条件是通过显微镜观察到的任何环境中 $>99\%$ 的微生物不能通过

标准培养技术培养。另一方面,随着近来基因组学和测序技术的进展,已经发现了多种不依赖培养的基于对来自样品的核酸、蛋白和脂类直接分离和分析的分子方法,揭示了微生物群落的结构和功能信息。分子方法(诸如遗传指纹分析、宏基因组学、宏蛋白组学、宏转录组学和蛋白质基因组学)对于发现和表征庞大的微生物的多样性和理解它们与生物的和非生物的环境因素的相互作用是至关重要的。

[0067] 根据具体的实施例,通过一种或多种方法测量所述口腔样品,所述方法选自16S rRNA(核糖核酸)分析、遗传指纹分析、克隆文库方法、变性或温度梯度凝胶电泳、随机扩增多态性DNA(脱氧核糖核苷酸)、DNA扩增指纹图谱、扩增核糖体核酸限制性分析、DNA芯片、荧光原位杂交、DNA-DNA杂交、宏基因组学、宏蛋白组学、宏转录组学、蛋白基因组学、Luria-Bertani培养基分离技术、Nutrient Agar分离技术、Tryptic Soy Agar分离技术以及它们的任何组合。微生物群落的分子分析显示出可培养的级分代表了任何给定样品中存在的原核菌种总数的<1%。这些分析方法的组合可提供更全面的微生物多样性评估。优选地,本发明中使用了选自16S rRNA分析、宏基因组学和它们的组合的一种方法来测量所述口腔样品以获取在所述口腔微生物群落中一种或多种微生物类型的丰度。最优选地,使用16S rRNA分析来研究所述口腔样品的微生物群落。

[0068] 根据具体的实施例,通过上述一种或多种方法获取所述口腔样品的口腔微生物群落中一种或多种微生物类型的丰度。

[0069] 根据具体的实施例,所述微生物类型选自细菌的分类种类、微生物的功能种类以及它们的组合。更具体地和优选地,所述微生物类型选自细菌门、细菌目、细菌科、细菌纲、细菌属、细菌种、微生物的功能基因、微生物基因同源组、微生物的肽或蛋白质的基序、微生物的保守的肽或蛋白质域、微生物的非编码核苷酸序列以及它们的组合,优选细菌属。

[0070] 可启动本发明通过尝试尽可能多的微生物类型以确定哪种微生物类型最能满足目的。例如,可分别鉴定细菌门、细菌属和细菌种,并分别对其丰度定量。在三个不同时间点的样品间能根据所述微生物类型的丰度的显著差异进行鉴定,使得对所述微生物群落变化进行鉴定以研究疾病的病因。因此,本文出于实现本发明的目的将选择据信在三个不同时间点的丰度显著变化的一种或多种微生物类型。

[0071] 根据本发明,通过成对比较分析或多变量分析、跨测试哺乳动物组统计分析所述微生物类型的丰度,以鉴定其丰度与该组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为第一组微生物类型,其中所述状况选自:疾病的存在、疾病的严重性、对疾病的敏感性以及它们的组合。

[0072] 多变量分析选自主成分分析、主坐标分析、对应分析、降趋对应分析、聚类分析、判别分析、典型判别分析以及它们的组合,优选主成分分析。主成分分析(PCA)是一种数学过程,其使用一个正交转换,将一组可能相关的变量的观测值转化为一组线性相非关变量(称为主成分,PC)的值。主成分的数量小于或等于初始变量的数量。这种变换是这样定义的:第一主成分(PC1)具有最大的方差(即,占尽可能多的数据的可变性),每一个后续成分在其正交于(即与之不相关)之前的组分的约束下依次具有最高可能的方差。只有当数据集是共同垂直分布的情况下,主成分保证是独立的。主成分分析法对初始变量的相对比例敏感。

[0073] 在具体的实施例中,统计显著性水平为 $p < 0.05$,优选地 $p < 0.01$,并且更优选地 $p < 0.001$ 。

[0074] 鉴定生物标记的方法

[0075] 本发明的一个方面提供了鉴定指示受试哺乳动物的状况的生物标记的方法,其包括以下步骤:

[0076] a) 选择具有第一疾病的第一组测试哺乳动物;

[0077] b) 获取第一口腔样品,所述第一口腔样品包含来自具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型;

[0078] c) 治疗已进行第一口腔取样的具有第一疾病的第一组测试哺乳动物中的每一个哺乳动物,以便消除或减少第一疾病;

[0079] d) 获取第二口腔样品,所述第二口腔样品包含来自已进行治疗的第一组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型;

[0080] e) 使第一疾病在已进行第二口腔取样的第一组测试哺乳动物中的每一个哺乳动物中复发;

[0081] f) 获取第三口腔样品,所述第三口腔样品包含已复发第一疾病的第一组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

[0082] g) 测量所述第一口腔样品、第二口腔样品和第三口腔样品以分别获得在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;

[0083] h) 跨第一组测试哺乳动物对获得的在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,以鉴定其丰度与第一组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型为第一组微生物类型,其中所述状况选自:第一疾病的存在、所述疾病的严重性、对所述疾病的敏感性以及它们的组合;

[0084] i) 从所述第一组微生物类型中选择一种或多种微生物类型作为指示所述受试哺乳动物的状况的生物标记。

[0085] 根据具体的实施例,在步骤h)中,用成对比较分析对获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,其包括以下步骤:

[0086] 1) 将所述第一组测试哺乳动物中的每一个哺乳动物的所述第一微生物群落和所述第二微生物群落进行比较,以测定在所述第一微生物群落和所述第二微生物群落之间所获得的每种微生物类型的丰度的变化;

[0087] 2) 跨第一组测试哺乳动物将来自步骤1)的所获得的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的初级组;

[0088] 3) 将所述第一组测试哺乳动物中的每一个哺乳动物的所述第二微生物群落和所述第三微生物群落进行比较,以测定在所述第二微生物群落和所述第三微生物群落之间所获得的每种微生物类型的丰度的变化;

[0089] 4) 跨第一组测试哺乳动物将来自步骤3)的所获得的每种微生物类型的丰度的变

化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的次级组;以及

[0090] 5) 将微生物类型的初级组和微生物类型的次级组进行比较,以鉴定出那些重叠的微生物类型作为第一组微生物类型。

[0091] 根据另一个具体的实施例,在步骤h)中,用多变量分析对获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,其包括以下步骤:

[0092] 1) 将所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所获得的丰度中方差最大的向量;以及

[0093] 2) 鉴定出具有表现出与导出的向量有统计意义上的显著的相关性的获得的丰度的那些微生物类型作为第一组微生物类型。

[0094] 有趣的是,本发明人的研究也阐明了微生物与哺乳动物群体中疾病结果的异质性的联系,其使得有可能基于微生物生态学将所述哺乳动物群体分成疾病敏感的哺乳动物和疾病较不敏感的哺乳动物。不受任何理论的束缚,发现疾病敏感的哺乳动物的特征在于:与疾病较不敏感的哺乳动物相比,从健康状态到疾病状态微生物群落结构的变化更剧烈。

[0095] 因此,根据另一个具体的实施例,在步骤h)中,用多变量分析对获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,其包括以下步骤:

[0096] 1) 将所获得的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所获得的丰度中方差最大的向量;

[0097] 2) 将获得的第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落中的每一种的一种或多种微生物类型的丰度投影至导出的变量上,以获得针对第一组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落的每一种的投影值;

[0098] 3) 针对第一组测试哺乳动物的每一个跨第一微生物群落、第二微生物群落和第三微生物群落计算所述投影值的变化速率;

[0099] 4) 基于所计算出的变化速率,将第一组测试哺乳动物分成测试哺乳动物的第一子组和测试哺乳动物的第二子组,其中所述测试哺乳动物的第一子组与测试哺乳动物的第二子组相比,表现出更大的变化速率;以及

[0100] 5) 分别将测试哺乳动物的第一子组的第一微生物群落、第二微生物群落和第三微生物群落与测试哺乳动物的第二子组的第一微生物群落、第二微生物群落和第三微生物群落进行比较,以鉴定第一微生物群落、第二微生物群落和第三微生物群落中的每一种中丰度在测试哺乳动物的第一子组和测试哺乳动物的第二子组间具有有统计意义上的显著差异的那些微生物类型,作为第一组微生物类型。

[0101] 根据另一个实施例,本发明方法还包括以下步骤:

[0102] 1) 选择具有第二疾病的第二组测试哺乳动物;

[0103] 2) 重复步骤b)至h)以鉴定第二组微生物类型;

[0104] 3) 将第一组微生物类型和第二组微生物类型进行比较,以鉴定出那些重叠的微生物类型作为微生物类型的子组;以及

[0105] 4) 从所述微生物类型的子组选择一种或多种微生物类型作为指示所述受试哺乳动物状况的生物标记,其中所述状况选自:第一疾病和第二疾病的存在、第一疾病和第二疾病的严重性、对第一疾病

[0106] 和第二疾病的敏感性以及它们的组合。

[0107] 用于鉴定生物标记的计算机辅助系统

[0108] 根据本发明,提供了有助于实施本发明的所述方法的计算机辅助系统。鉴定指示受试哺乳动物的状况的生物标记的本计算机辅助系统包括:

[0109] a) 用于取样的取样部分:

[0110] 1) 第一口腔样品,所述第一口腔样品包含来自具有疾病的一组测试哺乳动物中的每一个哺乳动物的第一微生物群落,其中所述第一微生物群落包含一种或多种微生物类型,

[0111] 2) 第二口腔样品,所述第二口腔样品包含来自已进行治疗以消除或减少所述疾病的该组测试哺乳动物中的每一个哺乳动物的第二微生物群落,其中所述第二微生物群落包含一种或多种微生物类型,和

[0112] 3) 第三口腔样品,所述第三口腔样品包含已复发所述疾病的该组测试哺乳动物中的每一个哺乳动物的第三微生物群落,其中所述第三微生物群落包含一种或多种微生物类型;

[0113] b) 与取样部分相联的测量部分,其中所述测量部分被配置用于测量第一口腔样品、第二口腔样品和第三口腔样品,以分别获得在第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;和

[0114] c) 与测量部分相联的计算部分,其中所述计算部分被配置用于在跨该组测试哺乳动物接收并统计学分析所获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度,以鉴定其丰度与该组测试哺乳动物的状况具有统计显著性的相关性的那些微生物类型作为指示所述受试哺乳动物的状况的生物标记,其中所述状况选自:所述疾病的存在、所述疾病的严重性、对所述疾病的敏感性以及它们的组合。

[0115] 所述取样部分可包括一种或多种器件,其形式选自勺、棉签、刀片、刷、探头、以及它们的任何组合。在具体的实施例中,所述取样部分包括无菌棉签,通过用所述无菌棉签轻轻擦涂暴露的牙齿表面实现取样。

[0116] 在具体的实施例中,本系统可包括样品储存部分用于储存样品。如果从所述取样部分采集的样品不立即使用,推荐将其储存在样品储存部分中。在另外的具体的实施例中,所述样品储存具有温度调节单元,其可向样品储存部分提供宽泛的储存温度范围,优选地低于30℃,更优选地低于0℃。在一个优选的实施例中,所述样品储存部分提供的储存温度低于-20℃,优选地低于-50℃,更优选地低于-70℃,并且最优选地低于-90℃。

[0117] 所述测量部分可以包括一个执行一种或多种方法的子部分,所述方法选自16S rRNA分析、遗传指纹分析、克隆文库方法、变性或温度梯度凝胶电泳、随机扩增多态性DNA、DNA扩增指纹图谱、扩增核糖体核酸限制性分析、DNA芯片、荧光原位杂交、DNA-DNA杂交、宏基因组学、宏蛋白组学、宏转录组学、蛋白基因组学、Luria-Bertani培养基分离技术、Nutrient Agar分离技术、Tryptic Soy Agar分离技术以及它们的任何组合。微生物群落结构或功能数据可能随测量部分中实施的具体方法而变化。

[0118] 所述计算部分可为任何形式。例如,它可为个人计算机或包含计算程序的便携式设备。根据具体的实施例,所述计算部分包括:

[0119] 1) 与测量部分相联的输入模块,其中所述输入模块用于输入获得的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度;

[0120] 2) 与输入模块相联的数据处理模块,其中所述数据处理模块被配置用于跨该组测试哺乳动物对第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行统计学分析,以鉴定其丰度与所述状况具有统计显著性的相关性的那些微生物类型;以及

[0121] 3) 与数据处理模块相联的输出模块,其中所述输出模块用于显示那些经鉴定的、作为指示所述受试哺乳动物的状况的生物标记的微生物类型。

[0122] 根据另外的具体的实施例,所述数据处理模块包含用于对输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行成对比较分析或多变量分析的程序。

[0123] 根据另外的具体的实施例,所述数据处理模块包含用于对输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行成对比较分析的程序,所述程序包含指令用于:

[0124] 1) 将该组测试哺乳动物中的每一个哺乳动物的所述第一微生物群落和所述第二微生物群落进行比较,以测定在所述第一微生物群落和所述第二微生物群落之间所输入的每种微生物类型的丰度的变化;

[0125] 2) 跨该组测试哺乳动物将来自步骤1)的所输入的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的初级组;

[0126] 3) 将该组测试哺乳动物中的每一个哺乳动物的所述第二微生物群落和所述第三微生物群落进行比较,以测定在所述第二微生物群落和所述第三微生物群落之间所输入的每种微生物类型的丰度的变化;

[0127] 4) 跨该组测试哺乳动物将来自步骤3)的所输入的每种微生物类型的丰度的变化进行比较,以选择表现出统计意义上的显著性的丰度变化的那些微生物类型作为微生物类型的次级组;以及

[0128] 5) 将微生物类型的初级组和微生物类型的次级组进行比较,以鉴定出那些重叠的微生物类型。

[0129] 根据另外的具体的实施例,所述数据处理模块包含用于对输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行多变量分析的程序,所述程序包含指令用于:

[0130] 1) 将所输入的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所输入的丰度中方差最大的向量;以及

[0131] 2) 鉴定出具有表现出与导出的向量有统计意义上的显著的相关性的微生物群落的那些微生物类型。

[0132] 根据另外的具体的实施例,所述数据处理模块包含用于对输入的第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度进行多变量分析

的程序,所述程序包含指令用于:

[0133] 1) 将所输入的所述第一微生物群落、第二微生物群落和第三微生物群落中的一种或多种微生物类型的丰度正交转换,以导出占所输入的丰度中方差最大的向量;

[0134] 2) 将所输入的该组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落中的每一种的一种或多种微生物类型的丰度投影至导出的变量上,以获得针对该组测试哺乳动物中的每一个哺乳动物的第一微生物群落、第二微生物群落和第三微生物群落中的每一种的投影值;

[0135] 3) 针对第一组测试哺乳动物的每一个跨第一微生物群落、第二微生物群落和第三微生物群落计算所述投影值的变化速率;

[0136] 4) 基于所计算出的变化速率,将第一组测试哺乳动物分成测试哺乳动物的第一子组和测试哺乳动物的第二子组,其中所述测试哺乳动物的第一子组与测试哺乳动物的第二子组相比,表现出更大的变化速率;以及

[0137] 5) 分别将测试哺乳动物的第一子组的第一微生物群落、第二微生物群落和第三微生物群落与测试哺乳动物的第二子组的第一微生物群落、第二微生物群落和第三微生物群落进行比较,以鉴定第一微生物群落、第二微生物群落和第三微生物群落中的每一种中丰度在测试哺乳动物的第一子组和测试哺乳动物的第二子组间具有有统计意义上的显著差异的那些微生物类型。

[0138] 在具体的实施例中,所述取样部分、测量部分和计算部分单独或以任何组合形式可以实现为计算机程序产品,其包含在计算机可读介质中的计算机可执行指令。示例性计算机可读介质包括芯片存储设备、磁盘存储设备、闪存装置、可编程逻辑器件、专用集成电路、可下载电信号等等。此外,适用于本发明的计算机程序产品可定位于一个单一的设备或计算平台,或可跨多个设备或计算平台分布。

[0139] 必要时,上述部分中的一者或多者可压缩成一个大型装置或小型便携式装置。

[0140] 实例

[0141] 本文实例旨在例示本发明,却未必用于限制或换句话讲限定本发明的范畴。

[0142] 缩略语表:

[0143] NG:天然存在的齿龈炎

[0144] EG:实验性齿龈炎

[0145] MGI:改性的/Mazza齿龈指数

[0146] BOP:探诊后出血

[0147] MiGs:牙龈炎的微生物指数

[0148] RPM:减退-发展模型

[0149] PAM:围绕中心点的分割算法(聚类算法)

[0150] PCA:主成分分析

[0151] PCoA:主坐标分析

[0152] FDR:错误发现率

[0153] COG:同源群的聚类

[0154] MiG:牙龈炎的微生物指数

[0155] MiG:牙龈炎敏感性的微生物指数

[0156] ROC:受试者操作特性

[0157] CI:置信区间

[0158] KO:KEGG直系同源

[0159] RPM设计图1A示出了模拟人群中齿龈炎发展的纵向研究的设计。为了理解齿龈炎的发病,建立了齿龈炎的实验模型来用作人的非侵入式模型。实验是经宝洁北京技术中心(中国)机构审查委员会批准并根据赫尔辛基世界医学协会声明(1996修订),在宝洁(北京)技术有限公司口腔护理部进行的。遵循针对优良临床试验规范(GCP)的ICH指南。从北京区域募集了91个受试者。提供了自愿知情同意书。

[0160] 包括了符合以下条件的个人:年满18岁;拥有至少12颗天然牙;在启动前访问中(-21天)经Mazza牙龈指数测量(MGI)有至少5个出血点的;有齿龈炎但不是牙周炎;基于对参与本研究的医学史/更新的审查由研究员/助理确定总体健康良好的。个人的排除标准包括:严重的牙周病,表现为脓性分泌物,广泛的牙齿松动,和/或严重退化;任何针对牙病预防治疗的给药需要抗生素术前用药的状况;在所述研究期间自我报告怀孕的或意图怀孕的和哺乳中的女性;牙龈组织中典型的变色或色素沉着;面部固定矫治;牙龈组织中典型的变色或色素沉着;在研究过程中的任何时候使用抗生素;可能会干扰受试者安全完成所述研究的任何疾病或状况。在整个研究中每周测量每个受试者的临床参数。将在任何时间点落入排除标准的个体从研究参与中排除。

[0161] 所述RPM包括三个阶段。

[0162] 阶段I,口腔卫生阶段(-21天至0天):在-21天、-14天、-7天和0天采用Mazza齿龈指数进行齿龈炎检查。接受牙病预防(龈上和龈下预防)和牙齿抛光后,指示每个受试者每日两次回到这个地点,在这个时间,他们在监督下使用魅力亮洁型手动牙刷(佳洁士,中国制造)、目前市售的没有任何明显的抗微生物活性的防蛀牙膏刷牙三分钟,然后用牙线清洁牙齿的邻面面积。在接下来的21天遵循这个刷牙方案,同时对每个受试者每次访视记录MGI。在口腔卫生阶段期间,如果受试者的出血点大于一处,受试者接受至多三次牙病预防。

[0163] 阶段II,实验性齿龈炎阶段(0天至21天):在这个阶段期间,受试者没有任何口腔卫生操作,其包括刷牙、用任何产品漱口、牙线洁齿和牙病预防。在实验性齿龈炎阶段的7天、14天和21天受试者也接受齿龈炎检查。

[0164] 阶段III,恢复阶段:指示每个受试者每日两次回到这个地点,在这个时间他们在监督下使用阶段I的产品和技术刷牙。受试者在恢复阶段接受一次牙病预防,并且受试者也接受齿龈炎检查,包括测量出血点,以记录和确认他们已经回到与他们进入该研究时同等或优选的更好的健康。如果需要的话,受试者接受一次附加的预防,并监测直到被认为是健康的。

[0165] 用探诊后出血(BOP)和Mazza齿龈指数(MGI)作为临床测量来评估齿龈炎。针对每个受试者记录BOP频率和平均MGI。MGI同时测量了炎症的病症和严重出血的程度。具体地,由牙医在每颗牙齿的近中颊[侧]和远中舌角进行探查,最多56个点。得分在0至5的范围内,0为正常表现和健康的,齿龈炎最高为5分:自发出血(无缘无故情况下)。由同一受过良好训练的牙医测量所有受试者的MGI以减少技术变量。

[0166] 图1B示出了在整个研究中50个受试者的上述临床参数的变化。在图1B中,方框代表四分位差(IQR),其中的直线代表中值。晶须代表 $1.5 \times IQR$ 内的最低值和最高值。在-21天,

所有受试者表现出一定程度的齿龈炎症,其代表天然存在的齿龈炎(“NG”),BOP在5至27的范围内,并且平均MGI为1.18至2.24。然后这些受试者进行严格的口腔卫生操作三周,其导致在0天(“基线”)BOP和MGI大幅降低(BOP和MGI中值分别为1.00和1.02,代表健康齿龈状态)。然后宿主另外进行针对诱导齿龈炎的口腔卫生计划三周,导致BOP(中值23)和MGI(中值2.11)大幅增加,其代表实验性齿龈炎(“EG”)的状态。

[0167] 龈上牙斑取样在-21天、0天和21天按照下述步骤从每个受试者采集龈上牙斑样品。在取样前受试者不进行口腔卫生操作,其包括刷牙、牙线洁齿和漱口。在摄入食物或饮料(水除外)后2小时采集样品。在MGI检查后,每个受试者用50ml无菌水漱口。在MGI检查15分钟后,由有资质的牙医用格雷西刮匙(Gracey curette)采集沿齿龈线2mm厚的牙斑。对于每个受试者,对所有牙齿在两对不同象限(1和3或2和4)采集牙斑样品并合并至一管。通过用无菌棉签擦拭收集格雷西刮匙(Gracey curette)上的牙斑。将拭子端部放入0.6ml TE20缓冲液(20mM Tris-HCl PH 8.0,2mM EDTA(乙二胺四乙酸))。在分离DNA前,所有样品储存在-70℃下。

[0168] 牙斑DNA提取规程按照稍作修改的Dr.Larry Fernery的规程从人齿牙斑提取总DNA(Ravel J,等人(2011)Vaginal microbiome of reproductive-age women.Proc.Natl.Acad.Sci.U.S.A.108:4680-4687)。一般来讲,在DNA分离实验前将冰冻的样品在冰上解冻。将初始样品(250μl)转移至干净的Bead-Beating-Tube(2ml Eppendorf管)中。在制备Lytic-Enzyme Cocktail(裂解酶混合物)时将样品悬液置于冰上。所有样品中加入新鲜制备的Lytic-Enzyme-Cocktail Master-Mix(100ul;在20mM乙酸钠中包含50μl Lysozyme~500KU=10mg/ml,6μl变溶菌素,25KU/ml,3μl溶葡萄球菌酶,4000U/ml和41μl TE缓冲液)在37℃下温育45分钟。向溶胞产物混合物中加入750mg干净的和干燥的0.1mm直径的Zirconia-Silica-Bead。在室温下、样品在Qiagen TissueLyser LT(36次振荡/秒)中经受珠子振动2分钟。将180μl的粗制溶胞产物转移至新管中,用DNeasy[®] Blood&Tissue Mini Kit通过Qiacube分离DNA。

[0169] 细菌16S rRNA基因扩增子测序从50个个体中获取150个牙斑样品并进行分析,每个个体提供在NG(-21天),基线(0天)和EG(+21天)的三个时间点的样品。采用454Titanium,条形码的16S rDNA扩增子测序产生了总共3,181,659个初始序列,得到总共1,093,922个处理后的序列(即,质量评价和控制措施后的序列)。每个样品的处理后的序列的数量在437至28,456的范围内,平均每个样品为7293个序列。所有序列存放在序列读取档案登录ID SRA058763名下。

[0170] 比较牙斑微生物群的系统发育结构首先在R中用ade4包(Dray S&Dufour AB(2007)The ade4package:Implementing the duality diagram for ecologists.Journal of Statistical Software 22(4):1-20)进行PCA分析,以将不同时间点间微生物群落结构的差异可视化。普式叠合分析法试图将一个矩阵中的点(诸如通过PCA获得的点)伸展和旋转,以尽可能靠近其它矩阵的点,从而保留每个矩阵内点之间的相对距离。在R中使用ade4包进行两个转换后数据的子组(即NG-基线,EG-基线和NG-EG的前四个主成分的数据矩阵)的简单普式旋转,以测试对于所述组群的微生物群在不同时间点间差异的程度。

[0171] 也进行主坐标分析(PCoA)已确认在齿龈炎人群和健康人群间微生物群结构的差异。在每个样品中,通过选择基于UCLUST(Edgar RC(2010)Search and clustering

orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460-2461) 的最长序列从每个OTU(操作分类单元)选择代表性序列。使用BLAST的megablast,每个序列被分配给CORE (Griffen AL,等人(2011)CORE:a phylogenetically-curated 16S rDNA database of the core oral microbiome. *PLoS One* 6(4):e19051) 中的系统发育的最接近的近亲。将所得的样品ID(标识)作图文件和分类作图文件用于输入FastUniFrac (Hamady M, Lozupone C,&Knight R(2010)Fast UniFrac:facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing和PhyloChip data. *ISME J* 4(1):17-27),其允许对基于区别所述生物体进化历史的分数的群落间距离进行成对比较。然后使用PCoA对这些距离聚类以降维,其中所述主坐标(PC)以降序描述了在新空间中每个轴说明的变形程度。此外,使用MOTHUR (Schloss PD,Gevers D,&Westcott SL(2011)Reducing the effects of PCR(Polymerase Chain Reaction)amplification和sequencing artifacts on 16S rRNA-based studies. *PLoS One* 6(12):e27310)进行基于ThetaYC的群落结构比较。ThetaYC测量了两个群落间的结构相异性。创建在所有样品间基于成对thetaYC距离的矩阵用于聚类和PCoA分析。

[0172] 统计分析为测试微生物群的结构异质性,通过围绕中心点的分割算法(PAM),使用归一化的属(或OTU)的丰度的Jensen-Shannon散度(JSD)进行牙斑微生物群间的聚类。基于silhouette指数的最大值来选择聚类的最佳数量。

[0173] 然后在R中使用ade4包进行PCA分析以将基于PAM的聚类可视化。在分析之前,将数据进行样品量归一化,去除丰度非常低的属(如果在所有样品中它们的平均丰度低于0.1%)以降低噪音。鉴定表现出与PC1最高的相关性的细菌属并将其加亮。

[0174] 使用加权相关性网络分析(WGCNA)以研究微生物缔合和相互作用。该方法应用于构建牙斑中细菌相互作用网络。在这些网络中,一个节点对应于一种给定微生物的微生物丰度分布。如果节点在整个环境扰动中具有显著的成对相关性的,这些节点就是有联系的。首先计算所有受试者中所有属之间的成对Pearson相关性。然后鉴定所述相关性的软阈值功率以按照近似无尺度拓扑结构的规则构建一个稳健的网络。计算属的拓扑重叠以反映它们相对的相互关联性。最后,输出数据并通过Cytoscape (<http://www.cytoscape.org>)可视化。所述成对Pearson相关性的功率是在EG中 $\beta=3$,无尺度拓扑拟合指数=0.7。在所述网络中绘制了平均相对丰度高于0.1%的和两种细菌间连接强度>0.05的口腔细菌(属级)。

[0175] 为了评估齿龈炎的牙斑微生物群的效应,本发明人基于所选的系统发育标记(生物标记)、通过成对t检验或spearman相关性方法、针对每个个体定义和计算了齿龈炎的微生物指数。针对每个个体样品,由 $f(A_i, A_j)$ 代表的微生物齿龈炎指数用下式计算:

$$[0176] \quad f(A_i, A_j) = b \left(\frac{\sum_{i \in N} A_i}{N} - \frac{\sum_{j \in M} A_j}{M} \right)$$

[0177] 其中N是在这些所选择的系统发育的标记中齿龈炎富集的标记的总数,M是在这些所选择的系统发育的标记中健康富集的标记的总数, A_i 是每种齿龈炎富集的标记i的丰度, A_j 是每种健康富集的标记j的丰度, $\sum_{i \in N} A_i$ 是 A_i 所有齿龈炎富集的标记i的总和, $\sum_{j \in M} A_j$ 是 A_j 所有健康富集的标记j的总和,并且b是可为10或任何其它数字的常数。

[0178] 牙斑宏基因组测序针对18个牙斑,分别提取宏基因组DNA并测序。所述样品是在基

线和EG两者、来源于9个受试者,其包括来自齿龈炎聚类I的5个受试者和来自齿龈炎聚类II的4个受试者。用NEXTflex™技术(BIO Scientific Corp.,USA)制备末端配对测序文库。首先用液氮将宏基因组DNA片段化。然后将包括标记序列的测序接头连接至尺寸经过选择的片段上。引入10个PCR循环以富集正确连接的片段。然后在HiSeq(Illumina,USA)上以2×150bp的测序长度对所富集产物进行测序。这些测序结果经过质量过滤,然后对其进行人工识别和分别存档。所有序列存放在序列读取档案登录ID SRA058763名下。

[0179] 基因的功能分类为了探查所编码的功能,使用带默认参数的IDBA(<http://i.cs.hku.hk/~alse/hkubrg/projects/idba/>)将微生物测序结果组装成重叠群。然后将组装好的重叠群提交至MetaGeneMark用于使用默认参数进行基因调用(gene calling)。然后使用BLAST和perl脚本将所述基因片段与COG数据库进行功能性赋值。通过COG对超过60%的基因进行注释。用R(2.15.1)生成了基于COG赋值的功能基因分布的PCA。

[0180] 图2示出了汇总如上所述的研究流水线的流程图。

[0181] 结果

[0182] 对于150个牙斑微生物群中的每一个,鉴定出细菌门、细菌属和细菌种并且通过针对参考数据库分类赋值对它们的相对丰度进行定量(CORE(Griffen AL,等人(2011) CORE:a phylogenetically-curated 16S rDNA database of the core oral microbiome.PLoS One 6(4):e19051))。

[0183] 一种齿龈炎减退和发展的实验性易控制的模型如图1B中所示,在群体水平上,基于成对t检验,在EG(平均BOP为 26.00 ± 9.59 ,平均MGI为 2.12 ± 0.48)下MGI($p < 0.001$)和BOP($p = 0.026$)显著高于在NG(平均BOP为 13.5 ± 5.12 ,平均MGI为 1.61 ± 0.24)下的。此外,对于单个受试者,在NG和EG之间的临床参数是显著相关的,诸如BOP(Pearson相关性: $r = 0.31$, $p = 0.03$)和平均MGI(Pearson相关性: $r = 0.35$, $p = 0.01$)。

[0184] 齿龈炎相关联的微生物群的结构特征和功能特征为鉴定与齿龈炎相关联的微生物群的结构特征,将150个健康的和疾病的微生物群通过PCA、基于属级分类群的相对丰度进行聚类,观察到在健康的牙斑微生物群(基线,三角形)和齿龈炎相关联的牙斑微生物群(NG,方块型;EG,点状)之间生物体结构的区别(参见图4A)。所述健康的和患病的微生物群大部分沿基线和NG/EG的交界处聚集,表明微生物群结构和疾病状态之间的联系。在EG下在NG和EG之间受试者结构内MGI更高在很大程度上是一致的,表明在相同受试者中微生物群落扰动与齿龈炎复发的方式相同。基于UniFrac和ThetaYC距离(参见图5A和5B)的PCoA也支持这些结果。因此在每个受试者体内可能存在微生物群-疾病的关联。

[0185] 本发明人通过所有细菌分类单元的相对丰度与宿主状态的相关性来检查所述微生物群-疾病的关联。

[0186] 在门级上,几乎所有的序列是来自13个细菌门,其包括在口腔中常见的6个主要细菌门:厚壁菌门(Firmicutes)、变形细菌门(Proteobacteria)、拟杆菌门(Bacteroidetes)、放线菌门(Actinobacteria)、梭杆菌门(Fusobacteria)和TM7(每一种在至少一个时间点具有的平均相对丰度 $>1\%$)。在齿龈炎状态(NG和EG)和健康齿龈状态(基线)之间,发现5个主要的门有显著差异($p < 0.01$;成对t检验):放线菌门(Actinobacteria)、厚壁菌门(Firmicutes)、TM7、拟杆菌门(Bacteroidetes)和梭杆菌门(Fusobacteria)。沿NG-基线-EG发展的群落结构的时间偏移是明显的,其特征在于在基线处放线菌门(Actinobacteria)

和厚壁菌门 (Firmicutes) 以及在NG和EG下TM7、拟杆菌门 (Bacteroidetes) 和梭杆菌门 (Fusobacteria) 的相对丰度升高。

[0187] 在属级上,27个细菌属(每一个在至少一个时间点的平均相对丰度 $>0.1\%$) 在基线和齿龈炎 (NG和EG两者) 之间是差异分布的 ($p<0.05$, 成对t检验; FDR (错误发现率) $q<0.2$)。在它们中,5种(链球菌属 (Streptococcus)、罗氏菌属 (Rothia)、放线菌属 (Actinomyces)、嗜血杆菌属 (Haemophilus) 和劳特罗普氏菌属 (Lautropia)) 在基线处显示出升高的丰度、而22个属(纤毛菌属 (Leptotrichia)、普雷沃菌属 (Prevotella)、梭杆菌属 (Fusobacterium)、TM7、卟啉菌属 (Porphyromonass)、坦纳菌属 (Tannerella)、月形单胞菌属 (Selenomonas)、未培养的毛螺旋菌属 (Uncultured_Lachnospiraceae)、未分类的丛毛单胞菌属 (unclassified_Comamonadaceae)、消化球菌属 (Peptococcus)、杆菌属 (Aggregatibacter)、卡托氏菌属 (Catonella)、密螺旋体属 (Treponema)、SR1、弯杆菌属 (Campylobacter)、真杆菌属 (Eubacterium)、消化链球菌属 (Peptostreptococcus)、未分类的拟杆菌属 (unclassified_Bacteroidaceae)、Solobacterium、约翰森氏菌属 (Johnsonella)、Oribacterium、和未分类的韦荣球菌属 (unclassified_Veillonellaceae)) 在NG和EG中均富集。图3示出了据信代表齿龈健康和齿龈炎 (针对自然发生的齿龈炎和实验性的齿龈炎两者) 的27个属级细菌生物标记。也显示了在不同阶段的微生物群落中鉴定的口腔细菌的相对丰度。这些细菌分类单元可能作为疾病标志物。

[0188] 目前将宿主分成患病组和健康组的临床实践是基于任意的MGI截止值1.10~1.12。然而,此类对疾病和健康的双峰定义与所观察到的宿主和微生物群的特性相反。为将样品间平均MGI值的分布可视化,用图4B中每个样品的平均MGI值缩放PCA的数据点。平均MGI和PC1值示出了显著的相关性 ($p<0.05$)。因此,事实上,临床参数(例如MGI) 在单个宿主和人群中的分布均是连续的。此外,PCA分析表明所述微生物群在NG、基线和EG之间的过渡不是离散的过程,而是梯度样的(参见图4A)。因此需要一种新的临床模型,所述模型考虑疾病表型和微生物群结构两者沿梯度的分布,其将有助于提供更客观的疾病状态测量并且允许对微生物群和疾病之间的联系做更合适的统计检验。

[0189] 一个给定的微生物在PC1的投影坐标看起来捕获了微生物群结构沿疾病减退和发展的梯度样的异质性和发展,因为在受试者内和跨群组的PC1中的变化与健康的和患病的微生物群之间的结构分离在很大程度上是一致的(参见图4B)。此外,使用所有150个样品沿PC1定义的微生物群的相对次序与单用仅健康的、仅NG的或仅EG的微生物群定义的那些相似 (Spearman相关性; 全部与仅健康的: $\rho=0.95$, $p<0.001$; 全部与NG: $\rho=0.97$, $p<0.001$; 全部与EG: $\rho=0.97$, $p<0.001$)。因此PC1看起来是主要描述符,并且因此是用于定量测量RPM-片段(例如NG-至-基线和基线-至-EG) 两者中微生物群的发展的良好代理。

[0190] 针对沿RPM的50个宿主,发现了15种细菌属驱动了沿PC1的微生物群异质化,因为它们的丰度梯度与它们对应的样品在PC1上的坐标显著相关 (Spearman $\rho>0.7$, FDR $q<0.2$), 如下面表1所示。

[0191] 表1: 显示出与PC1的显著相关性的口腔细菌

属	Rho 值
罗氏菌属 (<i>Rothia</i>)	-0.76
嗜血杆菌属 (<i>Haemophilus</i>)	-0.7
普雷沃菌属 (<i>Prevotella</i>)	0.85
纤毛菌属 (<i>Leptotrichia</i>)	0.81
梭杆菌属 (<i>Fusobacterium</i>)	0.71
月形单胞菌属 (<i>Selenomonas</i>)	0.85
未培养的毛螺旋菌属 (未培养的毛螺旋菌属 (<i>uncultured Lachnospiraceae</i>))	0.83
TM7	0.81
坦纳菌属 (<i>Tannerella</i>)	0.74
消化球菌属 (<i>Peptococcus</i>)	0.82
消化链球菌属 (<i>Peptostreptococcus</i>)	0.73
卡托氏菌属 (<i>Catonella</i>)	0.73
密螺旋体属 (<i>Treponema</i>)	0.82
<i>Solobacterium</i>	0.72
未分类的拟杆菌属 (<i>unclassified Bacteroidaceae</i>)	0.72

[0193] 这些驱动器包括罗氏菌属 (*Rothia*) ,嗜血杆菌属 (*Haemophilus*) ,普雷沃菌属 (*Prevotella*) ,纤毛菌属 (*Leptotrichia*) ,梭杆菌属 (*Fusobacterium*) ,月形单胞菌属 (*Selenomonas*) ,未培养的毛螺旋菌属 (未培养的毛螺旋菌属 (*uncultured Lachnospiraceae*)) ,TM7,坦纳菌属 (*Tannerella*) ,消化球菌属 (*Peptococcus*) ,消化链球菌属 (*Peptostreptococcus*) ,卡托氏菌属 (*Catonella*) ,密螺旋体属 (*Treponema*) ,*Solobacterium*和未分类的拟杆菌属 (*unclassified Bacteroidaceae*)。15个属中的两属,罗氏菌属 (*Rothia*) 和嗜血杆菌属 (*Haemophilus*) ,沿PC1降低了相对丰度(“负驱动器”),而其它13个属沿PC1增加(“正驱动器”)。为了理解它们在宿主群体的齿龈炎发展中的生态学作用,使用50个宿主中细菌属的相对丰度来分别创建在NG下、基线处和EG下的细菌相关性网络。有趣的是,在EG网络中,驱动器普雷沃菌属 (*Prevotella*) ,月形单胞菌属 (*Selenomonas*) ,未培养的毛螺旋菌属 (*Uncultured_Lachnospiraceae*) ,卡托氏菌属 (*Catonella*) ,消化链球菌属 (*Peptostreptococcus*) ,密螺旋体属 (*Treponema*) ,嗜血杆菌属 (*Haemophilus*) 和纤毛菌属 (*Leptotrichia*) 是前8种最多连接的节点,表明它们是潜在的齿龈炎细菌相互作用的主要枢纽。在三个不同时间点上15个PC1驱动的细菌属间的相互作用是类似的:在NG,基线和EG网络中的每个中,13个正驱动器彼此正向相互作用,与2个负驱动器负向相互作用。有趣的是,在15个PC1驱动器中在EG和基线处的网络连接比在NG下的多(参见图6A、6B和6C),在NG网络中仅有24个连接(14个节点),尽管在基线网络中存在65个连接(14个节点)以及在EG网络中存在54个连接(15个节点)。图6A、6B和6C中,节点的尺寸与所述分类单元的相对丰度成比例。其相对丰度显著性地相关联的那些分类单元通过‘连接’联结(实线:正相关性;虚线:负相关性)。因此,不受控制的环境因素可能会掩盖某些微生物间的联系,这突出了实验性疾病模型诸如RPM的价值。

[0194] 为测试齿龈炎微生物群的功能特征,分别对来自18个牙斑(来自9个受试者,每个人同时在基线和EG下取样)的基因组DNA进行鸟枪法测序,测序平均深度为3.94Gb/样品(参

见表2)。

[0195]

表 2: 针对 150 个牙斑微生物群制备的宏基因组鸟枪法序列的特征

宿主 id	年龄	性别	基线				EG (实验诱导的牙龈炎)			
			样品 ID	平均 MGI	出血 位点数	宏基因组样品 尺寸 (Gb)	宏基因组 序列	样品 ID	平均 MGI	出血位 点数
9066	32	F	9066B	1.00	0	3.92	26110020	9066E	1.93	26
9174	36	F	9174B	1.05	2	3.75	25005760	9174E	2.63	39
9183	26	M	9183B	1.02	1	4.15	27644386	9183E	1.77	19
9439	27	F	9439B	1.00	0	3.73	24860734	9439E	1.68	19
9445	27	M	9445B	1.05	2	3.25	21661610	9445E	1.98	26
9147	28	M	9147B	1.04	2	3.99	26585026	9147E	2.34	29
9148	41	F	9148B	1.07	2	4.02	26815920	9148E	3.13	41
9307	34	F	9307B	1.04	2	4.14	27567214	9307E	2.33	32
9325	32	F	9325B	1.00	0	4.14	27590228	9325E	2.52	36

[0196] 选取这9个受试者以最大化被采样微生物群的系统发育多样性。基于直系同源聚类分析 (COG) 数据库对所述微生物群中编码的功能基因进行分析,并基于指定的直系同源

组 (OG) 的相对丰度进行比较。有趣的是, 普式叠合分析指出在18个样品中, 系统发育的和功能的测量 (COG) 之间的一致性优异的 ($p < 0.001$, 通过10000蒙特卡洛标签置换)。如图7A所示, 疾病状态显著地影响了微生物群落功能。疾病状态对微生物群落功能的影响似乎被第一轴很好地隔离了。此外, 如图7B所示, 基于编码的功能基因对18个微生物群的聚类几乎与基于生物体结构的聚类相同, 表明齿龈炎微生物群在功能基因结构上与健康的 (微生物群) 有区别。图7B中, 将在第一四维上每个普式转换的拟合报告为通过10000蒙特卡洛标签置换的p值。总的来说, 涉及24个功能分类的1205个OG (从25个分类中的4873个OG选出) 是齿龈炎正相关或负相关的 ($p < 0.01$)。例如, 在功能分类N (细胞运动) 中, 齿龈炎中富集了与鞭毛生物合成途径最相关的33个OG, 而在健康宿主中富集的仅3个OG (全部与菌毛组装蛋白有关)。图7C示出了33个齿龈炎富集的OG, 其编码鞭毛生物合成途径的组分。示意图改编自京都基因与基因组百科全书 (KEGG), 对应的KO (KEGG直系同源) 的基因名加亮。另一方面, 在功能分类P中 (无机离子转运和代谢), 在健康的微生物群中富集了32个OG, 同时仅19个被废弃。因此齿龈炎微生物群在结构和功能上均区别于健康的微生物群。

[0197] PC1和疾病表型间的关联根据基于16S分类法的PCA的健康和患病微生物群的分类与根据基于功能基因的PCA的 (健康和患病微生物群的分类) 相同, 表明沿PC1轴的每个样品的值对于齿龈炎微生物群的结构和功能特点均是有用的描述符。

[0198] PC1值看起来提供了临床上有用的信息。表3和4示出了微生物群结构的改变 (Δ PC1) 与MGI的变化 (Δ MGI) 之间的相关性。表2是指在NG-基线、基线-EG和NG-EG三个阶段的50个宿主组群。表3是指在NG-基线的单个阶段的另外41个宿主组群。仅显示了带有其对应的 $p < 0.05$ (即显著相关性) 的rho值。

[0199] 表3

[0200]

Spearman 相关性 (rho)		NG 至基线*				基线至 EG				NG 至 EG			
		NG.PC1	Δ PC1	NG.MGI	Δ MGI	B.PC1	Δ PC1	B.MGI	Δ MGI	EG.PC1	Δ PC1	EG.MGI	Δ MGI
NG 至基 线*	NG.PC1	不适用	-	-	-	-	-	-	-	-	-	-	-
	Δ PC1	-0.86	不适用	-	-	-	-	-	-	-	-	-	-
	NG.MGI	0.37	-0.42	不适用	-	-	-	-	-	-	-	-	-
	Δ MGI	-0.36	<u>0.40</u>	-0.98	不适用	-	-	-	-	-	-	-	-
基线 至 EG	B.PC1	不适用				不适用	-	-	-	-	-	-	-
	Δ PC1	0.28	<u>-0.51</u>		不适用		-	-	-	-	-	-	-
	B.MGI					不适用		-	-	-	-	-	-
	Δ MGI			0.39	<u>-0.38</u>	<u>0.53</u>		不适用		-	-	-	-
NG 至 EG	EG.PC1	0.4	-0.44			0.89		0.48	不适用	-	-	-	-
	Δ PC1	-0.56	<u>0.43</u>			0.49		0.43	0.43	不适用	-	-	-
	EG.MGI			0.39	-0.37	0.53		1.00	0.48	0.43	不适用	-	-
	Δ MGI					0.44		0.86	0.41	<u>0.56</u>	0.86	不适用	-

[0201] 表4

[0202]

Spearman 相关性 (rho)		NG 至基线*			
		NG.PC1	Δ PC1	NG.MGI	Δ MGI
NG 至基线*	NG.PC1	不适用	-	-	-
	Δ PC1	-0.43	不适用	-	-
	NG.MGI			不适用	-
	Δ MGI		<u>0.37</u>		不适用

[0203] 在NG下(并且也在EG下),所述50个受试者中MGI和PC1值之间有显著的相关性(Spearman相关性NG:rho=0.37,p<0.01;EG:rho=0.48,p<0.001)。此外,在NG和基线之间(并且也在基线和EG之间),所述100个微生物群的PC1值与MGI正相关(Spearman相关性;所有:rho=0.74,p<0.001;NG-基线:rho=0.77,p<0.001;基线-EG:rho=0.79,p<0.001)。

[0204] PC1的变化(Δ PC1)也是临床上相关的。在50个宿主中,在两个RPM区段(NG-至-基线和基线-至-EG)中的每个中,在受试者内的 Δ PC1和 Δ MGI(即MGI的变化)是显著相关的(标记为下划线并加粗),同时区段间 Δ PC1- Δ PC1之间的相关性和阶段 Δ MGI- Δ MGI之间的相关性也是显著的(标记为下划线并斜体)。此外,在NG和EG之间,在受试者内的 Δ PC1与在受试者内的 Δ MGI显著相关(Spearman相关性rho=0.56,p=0)。有趣的是,对于具有相对稳定的MGI NG-至-EG的10个下五分位数受试者, Δ PC1与 Δ MGI的无显著相关性(Spearman相关性rho=0.25, p=0.48)。然而,对于MGI变化最大NG-至-EG的10个上五分位数受试者, Δ MGI与 Δ PC1显著相关(Spearman相关性rho=0.64,p=0.05),表明 Δ PC1可定量示范齿龈炎中症状变化的程度。

[0205] 具有不同齿龈炎敏感性的两类宿主在50个受试者中,大多数宿主在从NG至EG的疾病发展过程中表现出大体上一致的微生物群结构(参见图8A)。虽然NG-基线(或基线-EG)PC1-变化在所述50个宿主组群中差别很大,微生物群改变NG-基线的速率和微生物群变化基线-EG的速率在每个受试者内是大体上相似的。MGI变化的速率符合类似的模式。此外,由于微生物群结构(即PC1),在EG下齿龈炎严重程度(即,MGI)与在NG下高度相关。对于在EG下(与NG相比)的大多数宿主,疾病结果以及微生物群结构的持续表明存在宿主依赖的(以及可能是个人的)因素决定了自然人群中齿龈炎复发的易感性。

[0206] 基于微生物群在受试者内的变化(在NG-至-基线的 Δ PC1和在基线-至-EG的 Δ PC1)和沿RPM的临床症状(在NG-至-基线的 Δ MGI和在基线-至-EG的 Δ MGI)的PCA显示出在所述50个宿主中疾病易感性的趋异。

[0207] 如图8B所示,将50个成员的组群中所有宿主绘制在基于微生物群和MGI的分布变化的PCA的前两种主成分上。示出了描述所述50个宿主沿PCA的主成分分布的直方图和核密度图(实线)。竖直的虚线将50个宿主分成I型(点)和II型(三角形)。对作为对于这些聚类的主要因素在这4个变量进行测定并通过它们加载至在这2种主成分来绘制。“a”代表 Δ PC1(NG-基线);“b”代表 Δ MGI(NG-基线);“c”代表 Δ PC1(基线-EG);并且“d”代表 Δ MGI(基线-EG)。

[0208] 所述50个宿主的分布模式提出了双峰分布(p=0.74,对于基于针对单峰性的Hartigans'dip检验法的非双峰分布假设),其中可绘制一条有识别能力的线将宿主分成两类,其被本发明人命名为齿龈炎较不敏感的I型(17个个体)和齿龈炎敏感的II型(33个个

体)。与I型宿主相比,II型宿主的特征在于微生物群结构和MGI均更剧烈的变化。(参见图8A和图8C)。对于平均的II型宿主,沿RPM的PC1-变化速率是0.33/天,并且沿RPM的MGI-变化速率是0.05/天,其分别为平均的I型宿主的2.21倍和1.89倍(参见图8C)。

[0209] 在NG和EG下,在齿龈炎敏感的类型和某些分类单元的相对丰度之间存在显著的相关性($p < 0.05$, Wilcoxon秩和检验法),其包括营养缺陷菌属(Abiotrophia),月形单胞菌属(Selenomonas),未培养的毛螺旋菌属(uncultured Lachnospiraceae),消化球菌属(Peptococcus),未分类的拟杆菌属(unclassified Bacteroidaceae),消化链球菌属(Peptostreptococcus),Oribacterium和韦荣球菌属(Veillonellaceae);与I型宿主相比,这些分类单元全富集在II型宿主中,除了在I型中富集的营养菌缺陷属(Abiotrophia)(参见图8D)。这些II型宿主相关联的属的大多数(5个)在15个PC1-驱动器中。

[0210] 有趣的是,与I型宿主相比,在NG和EG下的II型宿主中富集的那些属在它们的基线下的II型宿主中的丰度也更高。因此在宿主中牙斑微生物群的异质性可能至少部分地解释,作为原因或作为结果,宿主间齿龈炎敏感性的表型变化和对于人群中疾病复发的可能易感性。

[0211] 齿龈炎的微生物指标在受试者之间和在受试者内PC1和疾病症状(MGI)的强相关性表明PC1可能用于建疾病发展模型。为测试是否可基于牙斑微生物群构建齿龈炎的预测模型,将50个宿主的组群用作模型构建的训练组,同时募集另外41个带有天然存在的齿龈炎的受试者,然后每人在NG和基线均取样(因此对82个附加的微生物群样品进行测序)用于模型验证。

[0212] (1) MiG27: 本发明人基于27种细菌标记物(其在50个宿主的组群中基线阶段和齿龈炎阶段之间不同)的相对丰度,通过下列方程式得出“齿龈炎的微生物指标(MiG)” (MiG27):

$$[0213] \quad MiG27 = \left(\frac{\sum_{i=22} \text{丰度}(g_{\text{齿龈炎富集的}})_i}{22} - \frac{\sum_{j=5} \text{丰度}(g_{\text{健康状态富集的}})_j}{5} \right) \times 10$$

[0214] 在所述50个宿主的组群中,在NG-至-基线($p < 0.001$, 学生t检验)和基线-至-EG($p < 0.001$, 学生t检验)下该指标与MGI高度相关: 受试者工作特征(ROC)曲线下的面积在NG-至-基线下为99.52% (95%的置信区间: 98.77%–99.52%), 在基线-至-EG下为99.84% (95%的置信区间: 99.53%–99.84%)。

[0215] 为验证MiG27的预测能力,本发明人使用了其NG微生物群预测在新组群中41个宿主的齿龈炎状态。图9示出了另外41个宿主的组群的MiG27指标。方块代表IQR,其中直线代表中值。晶须代表1.5x IQR内的最低值和最高值。在NG ($MGI > 1.18$) 和基线 ($MGI < 1.12$) 之间的MiG27有显著差异($p < 0.001$, 成对t检验, t统计量=22.3),例如具有最高MiG27的前27个样品全部正确分类为齿龈炎。针对疾病状态与健康状态的预测总准确率(基于线性判别分析)是94% (即,错误率为6.1%) (参见下表5)。这些数据表明该MiG27对于在临床环境下筛查患病牙龈有价值。

[0216] (2) MiG15: 虽然MiG27辨别健康状态和齿龈炎具有高准确率,针对齿龈炎群体的患病严重性的分类系统将是具有价值的。因此得出了MiG15,其基于驱动微生物群沿PC1的结构异质性的15个细菌属的相对丰度。通过下列公式计算给定微生物群的MiG15:

$$[0217] \quad MiG15 = \left(\frac{\sum_{i=1}^{13} \text{丰度}(g_{\text{高PC1富集的}})_i}{13} - \frac{\sum_{j=2}^2 \text{丰度}(g_{\text{低PC1富集的}})_j}{2} \right) \times 10$$

[0218] 然后本发明人使用线性回归分析将MiG15 (X) 上的相对PC1值 (Y: 齿龈炎的发展) 回归。用于预测的公式为: $Y = -0.97 - 4.62X$ 。该修订过的模型能够解释在50个宿主的组群中60%的PC1位置变化。基于41个宿主的组群中NG微生物群测试了该模型对疾病严重性的预测能力。图10示出了另外的41个宿主的组群中的MiG15指标。方框代表四分位差 (IQR), 其中直线代表中值。晶须代表1.5x IQR内的最低值和最高值。热图表明MiG15区分宿主的健康和齿龈炎状态的能力。MiG15示出了与在NG下每个受试者平均MGI的显著相关性 ($p < 0.05$, Spearman相关性)。预测值和测试值 (PC1的) 归类成三个分位数显示出预测错误率为24.4% (参见下表5)。因此, 基于MiG15的模型能够预测由PC1定义的人宿主中齿龈炎的严重性, 准确度为约75%。

[0219] (3) MiG-敏感性 (MiG-S): 本发明人还基于8种细菌标记物 (其在NG下、50个宿主的组群中I型和II型间不同) (MiG-S) 的相对丰度, 通过下列方程式得出“齿龈炎敏感性的微生物指标 (MiG-S)”:

$$[0220] \quad MiG-S = \left(\frac{\sum_{i=1}^7 \text{丰度}(g_{\text{II型富集的}})_i}{7} - \frac{\sum_{j=1}^1 \text{丰度}(g_{\text{I型富集的}})_j}{1} \right) \times 10$$

[0221] 在所述50个宿主的组群中, 该指标与类型高度相关 ($p < 0.05$, Wilcoxon秩和检验法): 在ROC曲线下的面积为74.0% (95%的置信区间: 60.2%–74.0%) (参见图11), 表明预测齿龈炎敏感性宿主类型的准确性最高至74.0%。

[0222] 表5: 基于牙斑微生物群的人齿龈炎的预测模型

[0223]

错误率		MiG27	MiG15	MiG-S
临床状态	健康状态对齿龈炎	6.1%	6.1%	-
归类的齿龈炎状态	基于 MGI	41.5%	41.5%	-
	基于 PC1	24.4%	24.4%	-
宿主的齿龈炎敏感性	基于 PC1 和 MGI 的变化模式	-	-	26.0%

[0224] 讨论

[0225] 本齿龈炎的减退-发展模型显示出在受试者内和自然群体中齿龈微生物群的异质性源。在更健康的状态和更患病的状态之间的发展程序主要是由15个细菌属所驱动的, 其中大多数在发育过程中相对丰度增加 (除了两种减少)。微生物群的分选学偏移伴有功能的偏移: 观察到的齿龈炎富集的功能诸如鞭毛生物合成可能追溯到细菌口腔运动, 因为鞭毛能辅助侵入宿主组织并逃脱吞噬作用。值得注意的是, 这些疾病驱动的分类单元, 在正常人中大部分主要的口腔共生体在自然齿龈炎和实验性齿龈炎群体中均充当主要的微生物相互作用的枢纽。因此该证据支持了齿龈炎的多微生物本质, 其不是由一种特定病原体而是由我们口腔微生物群落的特定成员驱动的所有分类学的和功能的变化所驱动的。

[0226] 存在两种具有不同的齿龈炎敏感性/易感性的宿主类型 (I型和II型), II型宿主与I型宿主相比具有平均超过两倍的急性疾病发展的特征。此外, 齿龈炎复发看起来是个人化的, 因为在EG下的齿龈炎严重性 (例如, MGI) 与在NG下的 (齿龈炎严重性) 是高度相关的, 同

时疾病进展 速率(即基线-至-EG)与疾病减退速率(即NG-至-基线)是高度相关的。本发明人已鉴定出与所述两种宿主类型的微生物关联,8种细菌分类单元与在NG、基线和EG中的每个下的II型宿主特异性相关联(7种富集,一种耗尽)。然而,由于在分类单元和宿主类型间的此种关联甚至在基线下(即“健康”状态)实际持续,微生物因素在宿主类型形成过程中可能起到主要的作用,并且II型宿主可能因为其在NG下栖息的微生物群而倾向于复发齿龈炎。

[0227] 此外,揭示齿龈的微生物群中这些主要变化来源可能会影响牙周病。齿龈炎可发展成牙周炎,牙周炎是成人中牙缺失的主要原因。然而,齿龈炎在牙周炎的发病机制中的作用仍存在争议:它们之间的一种病原学连接已被假定但尚未证实。一个混杂的因素是并非所有的齿龈炎案例发展为牙周炎:流行病学研究显示约50%的成人在超过6颗牙齿周围有齿龈炎,而仅有15%的成人患牙周炎。在所鉴定的“齿龈炎-驱动器”属中,若干种(例如,福赛斯坦纳菌(*Tannerella forsythensis*),微小消化链球菌(*Peptostreptococcus micros*),核梭杆菌亚种(*Fusobacterium nucleatum subsp.*),副嗜沫嗜血杆菌(*Haemophilus paraphrophilus*)和碳酸噬胞菌种(*Capnocytophaga sp.*)口腔克隆CZ006等)据报道与牙周炎相关。此外,本发明人鉴定的那些潜在的严重齿龈炎标记物(例如坦纳菌属(*Tannerella*),密螺旋体菌种(*Treponema*)和TM7门)据报道在牙周炎中富集。此外,已经发现II型宿主的若干潜在的标记物(例如月形单胞菌属(*Selenomonas*),消化链球菌属(*Peptostreptococcus*),未分类的毛螺旋菌属(*unclassified Lachnospiraceae*),韦荣球菌属(*Veillonellaceae*)和*Oribacterium*属),其表现出更高的疾病急性和对复发的易感性,在牙周炎中富集。集合的证据强烈支持严重齿龈炎与牙周炎的联系,并且还提供了对人类群体中牙周炎易感性的变化的一个可能的解释。

[0228] 最后,经鉴定的齿龈炎发展和易感性的微生物驱动器提供了改善临床实践的新机会。在齿龈炎中,齿龈组织表现出颜色变化、轮廓蚀变、沟内分泌物增加和在刺激时出血。基于一种或多种此类宿主症状,目前提出或实行的牙龈指数是主观的,容易出现人为偏见和错误,难以复制,因为此类指数是严重依赖于人类检查员的视觉观察和个人判断。例如,尽管它在临床实践中普遍使用,作为齿龈炎严重性的主观测量的MGI在不同检查员之间重复性很差。此外,由于齿龈炎的症状在不同牙齿(和甚至探查点)之间可能变化非常大,对每个患者的28颗牙的每一颗测试两个探查位点可为时间密集型和劳动密集型的。这些弊端已经共同混淆了交叉检查员和齿龈炎的跨患者分析。本发明人开发和验证了一种针对齿龈炎的替代的和可能互补的量度,其基于牙斑微生物群的定量分析。本发明的基于MiG的预测模型能够以95%的准确度预测患病的微生物群,以75%的准确度区分不同的疾病阶段,并且可能预测疾病敏感性。因此MiG能充当齿龈健康状态和齿龈炎易感性的更敏感、更可靠和更客观的度量,并且因此有助于齿龈疾病的诊断、预后和干预。

[0229] 人类微生物群作为跟踪和诊断宿主状况(疾病、饮食等)的位点的潜力依赖于并受限于在群体水平上微生物群-状况间关联的异质性。在肠中,宿主间微生物群结构的变化看起来主宰了状况间的变化(例如瘦或肥胖,或采用正常或高脂肪饮食)。然而,这里的结果显示出对于口腔微生物群相反的情况看起来是对的,在受试者内健康的和疾病的口腔微生物群之前的差异大于个体间差异。虽然响应的在不同的身体内栖息的微生物群落大小差异的机制未知,但本发明表明作为针对口腔以及或许甚至系统疾病的生物标记,口腔微生物群

可能具有某些优势。

[0230] 除非另外指明,所有百分比、比率和比例均以总组合物的重量计。除非另外指明,所有温度均以摄氏度(°C)为单位。除非另外指明,所有的测量均在25°C下进行。所有组分或组合物含量是指该组分或组合物的活性物质含量,并且不包括市售来源中可能存在的杂质,例如残余溶剂或副产物。

[0231] 应该理解,在本说明书中给出的每一上限值包括每一个下限值,如同该下限值在本文中也明确地表示。在本说明书中给出的每一下限值将包括每一个上限值,如同该上限值在本文中也明确地表示。在本说明书中给出的每一数值范围将包括包含于该较大数值范围内的所有较小的数值范围,如同该较小的数值范围在本文中也明确地表示。

[0232] 应当了解,本文所公开的量纲和值不旨在严格限于所引用的精确值。相反,除非另外指明,每个这样的量纲旨在表示所述值以及围绕该值功能上等同的范围。例如,所公开的量纲“40mm”旨在表示“约40mm”。

[0233] 除非明确排除或限制,将本文引用的每篇文献,包括任何交叉引用或相关专利或专利申请,全文以引用方式并入本文。任何文献的引用不是对其作为本文所公开的或受权利要求书保护的发明的现有技术,或其单独地或与任何其它参考文献的任何组合,或者参考、提出、建议或公开任何此类发明的认可。此外,当本发明中术语的任何含义或定义与以引用方式并入的文件中相同术语的任何含义或定义矛盾时,应当服从在本发明中赋予该术语的含义或定义。

[0234] 虽然已经举例说明和描述了本发明的特定实施方案,但是对于本领域的技术人员来说显而易见的是,在不背离本发明的实质和范围的情况下可做出多个其它改变和变型。因此,本文旨在所附权利要求中涵盖属于本发明范围内的所有这些改变和变型。

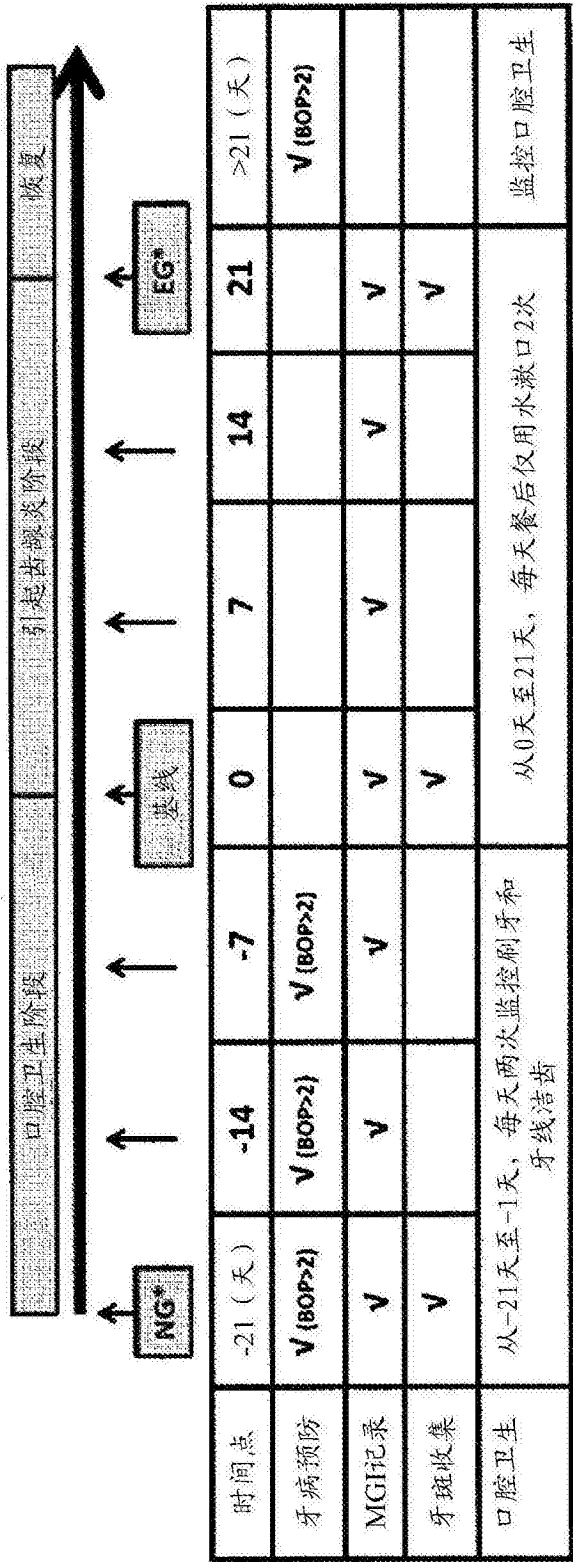


图1A

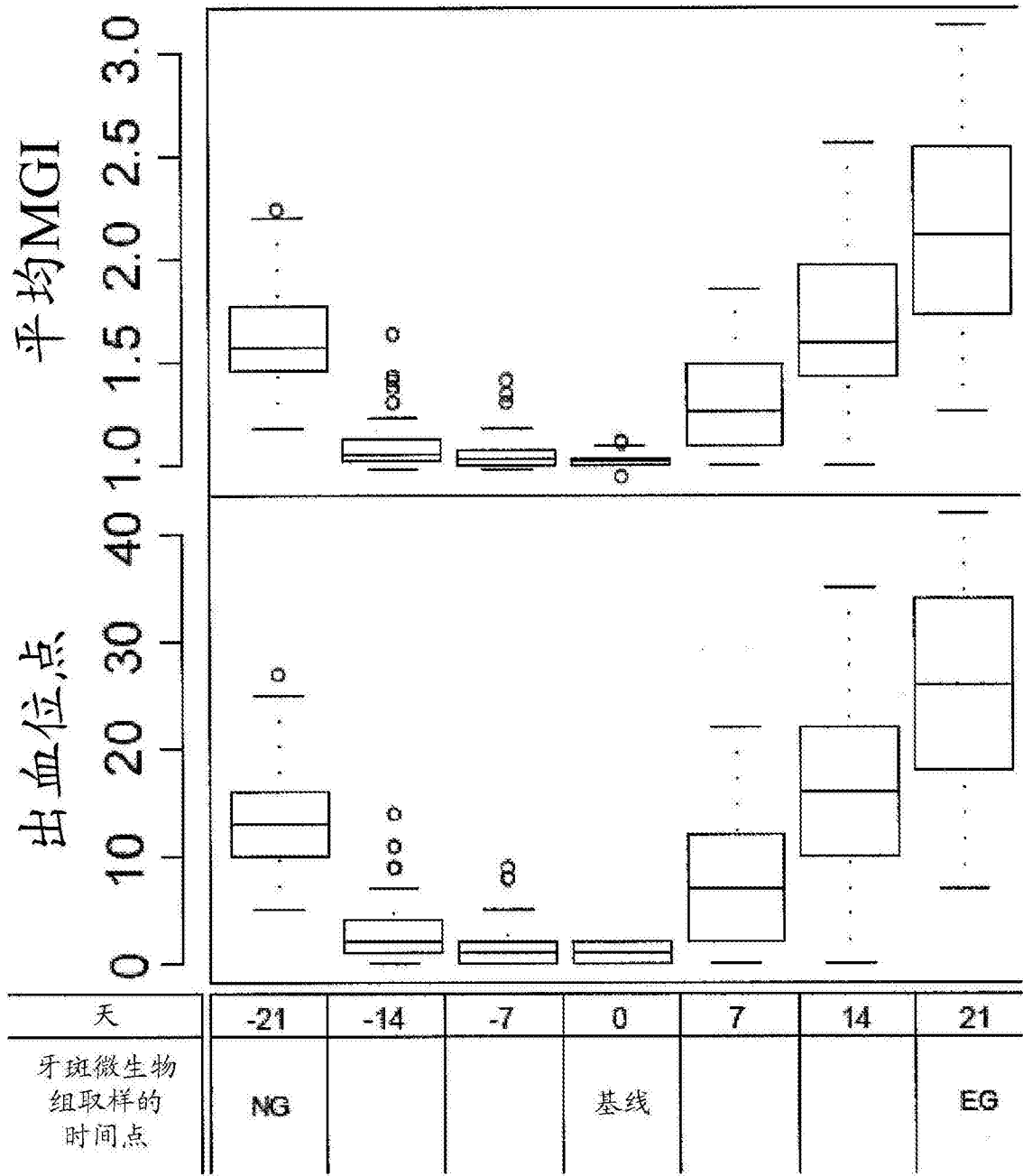


图1B

研究流水线

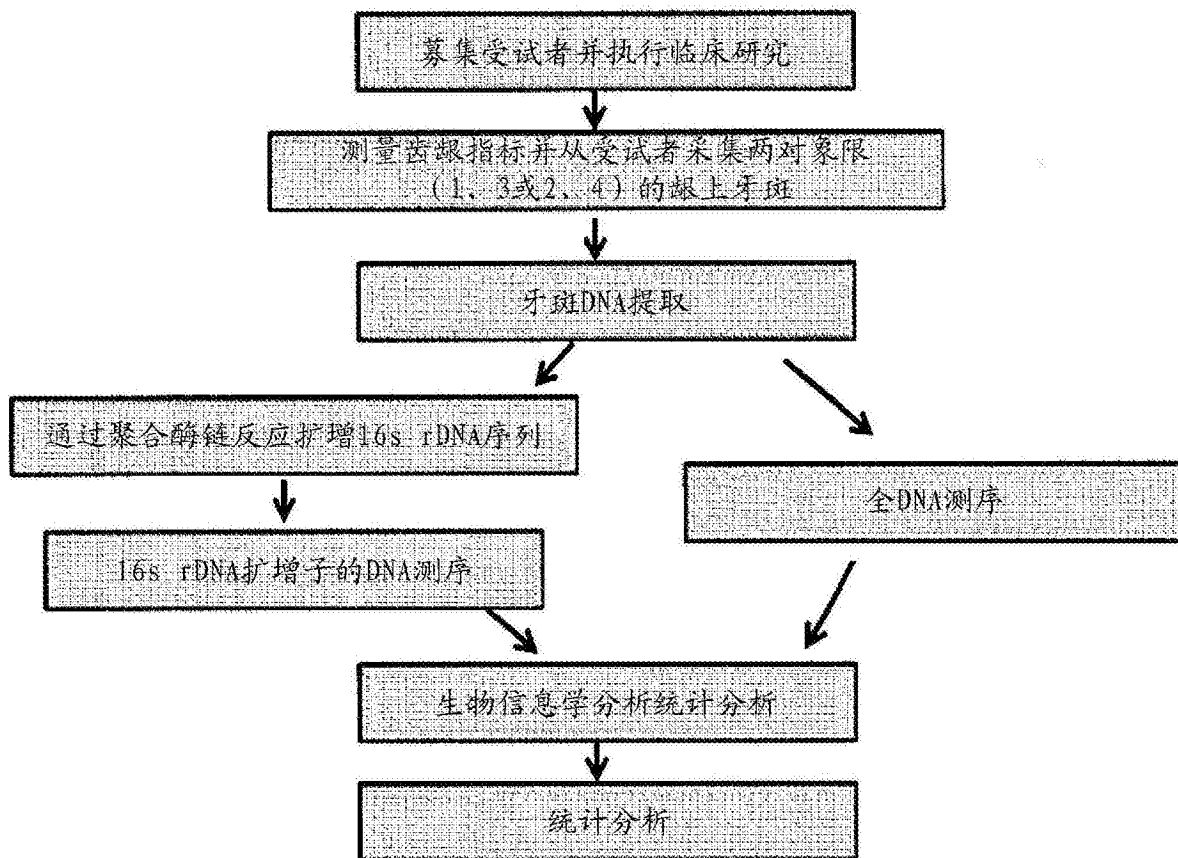


图2

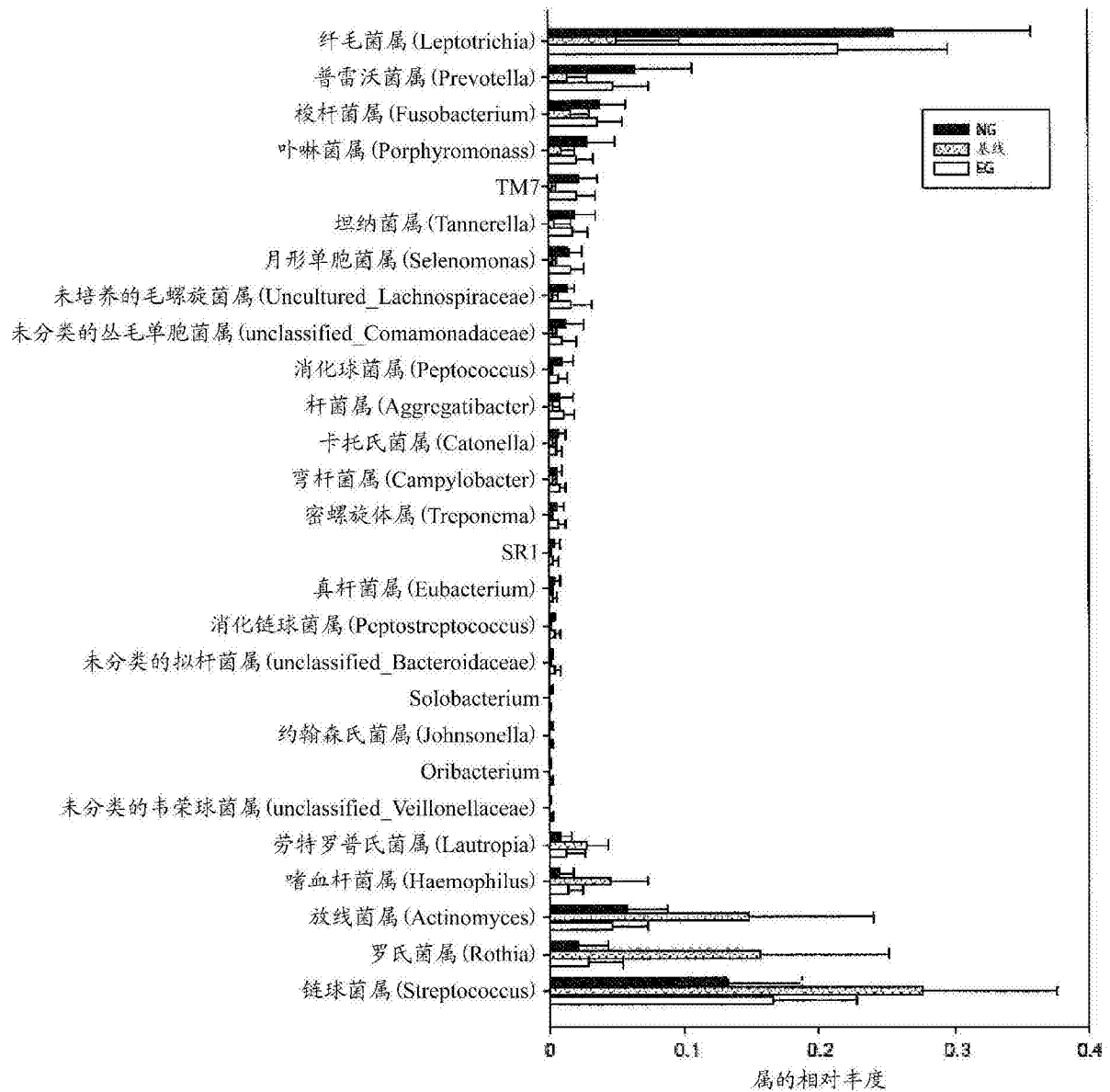


图3

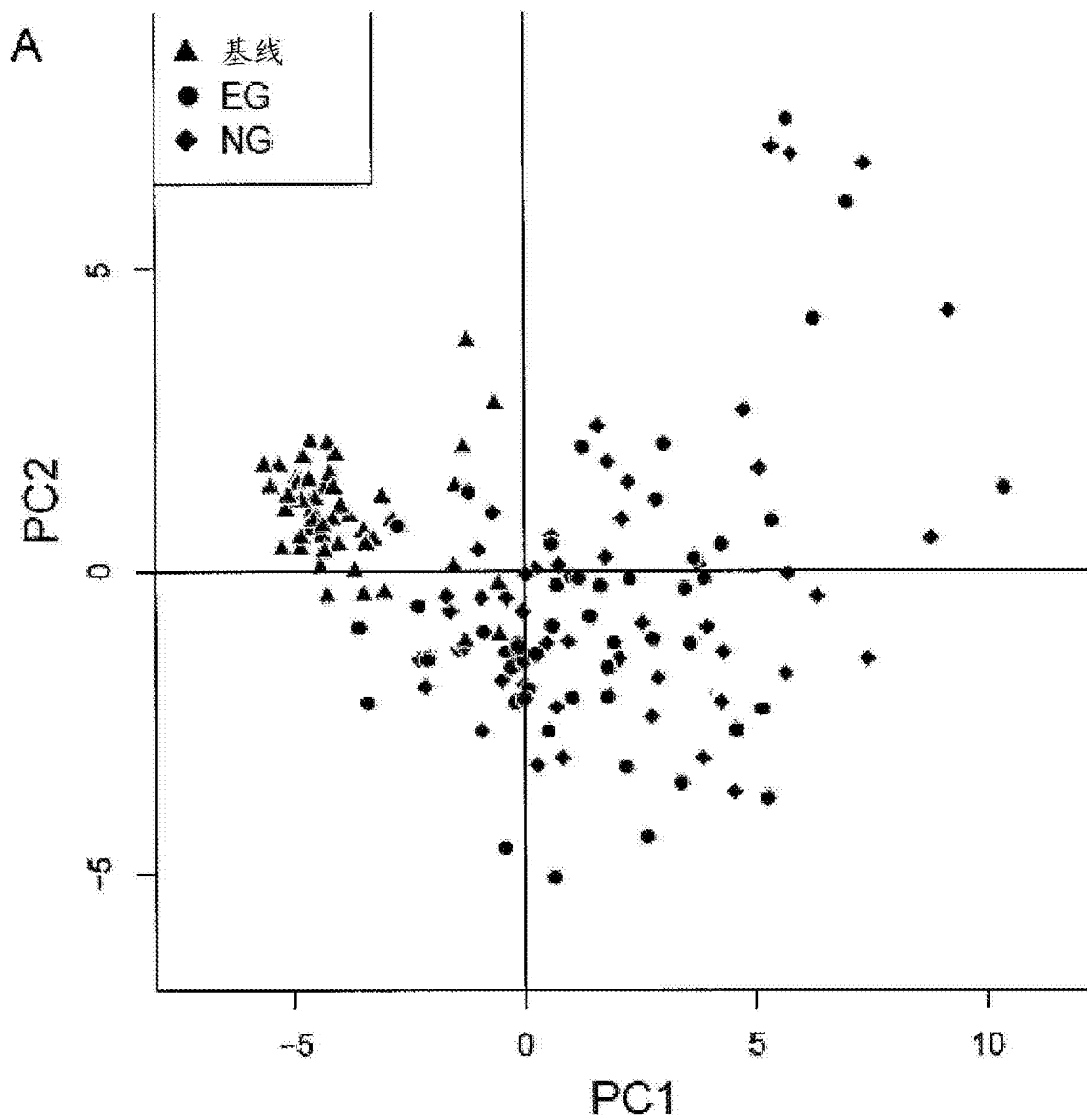


图4A

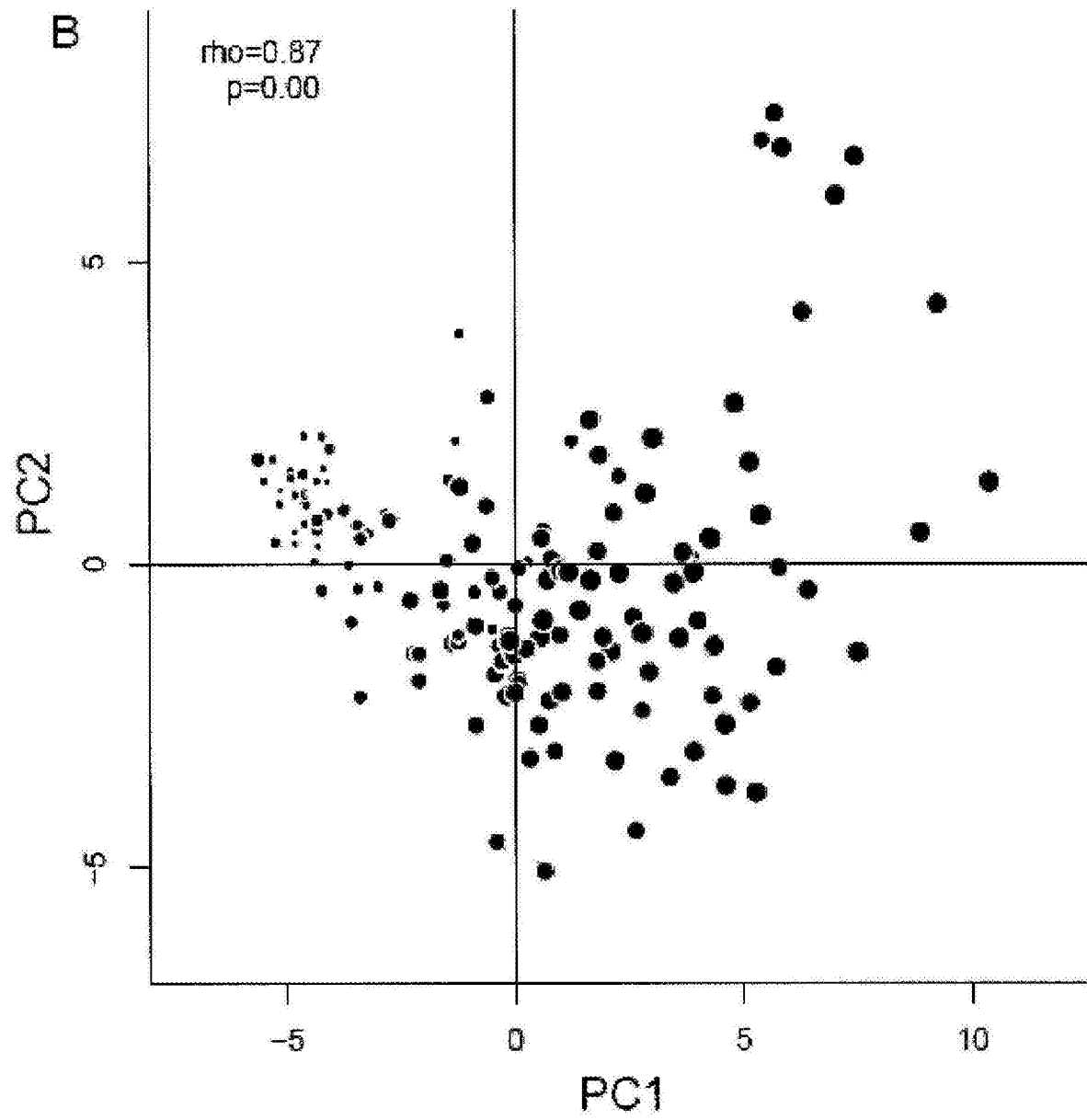


图4B

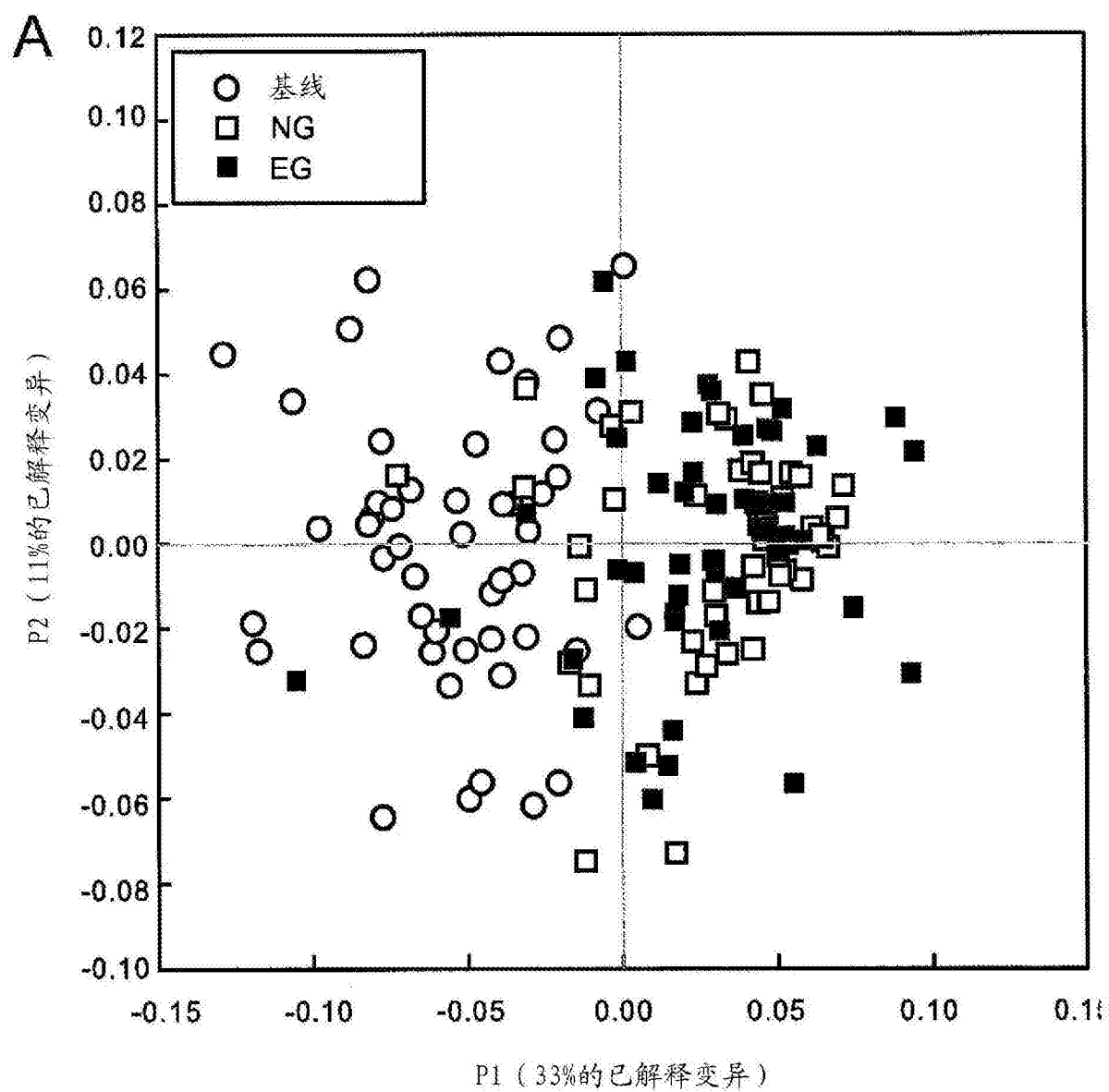


图5A

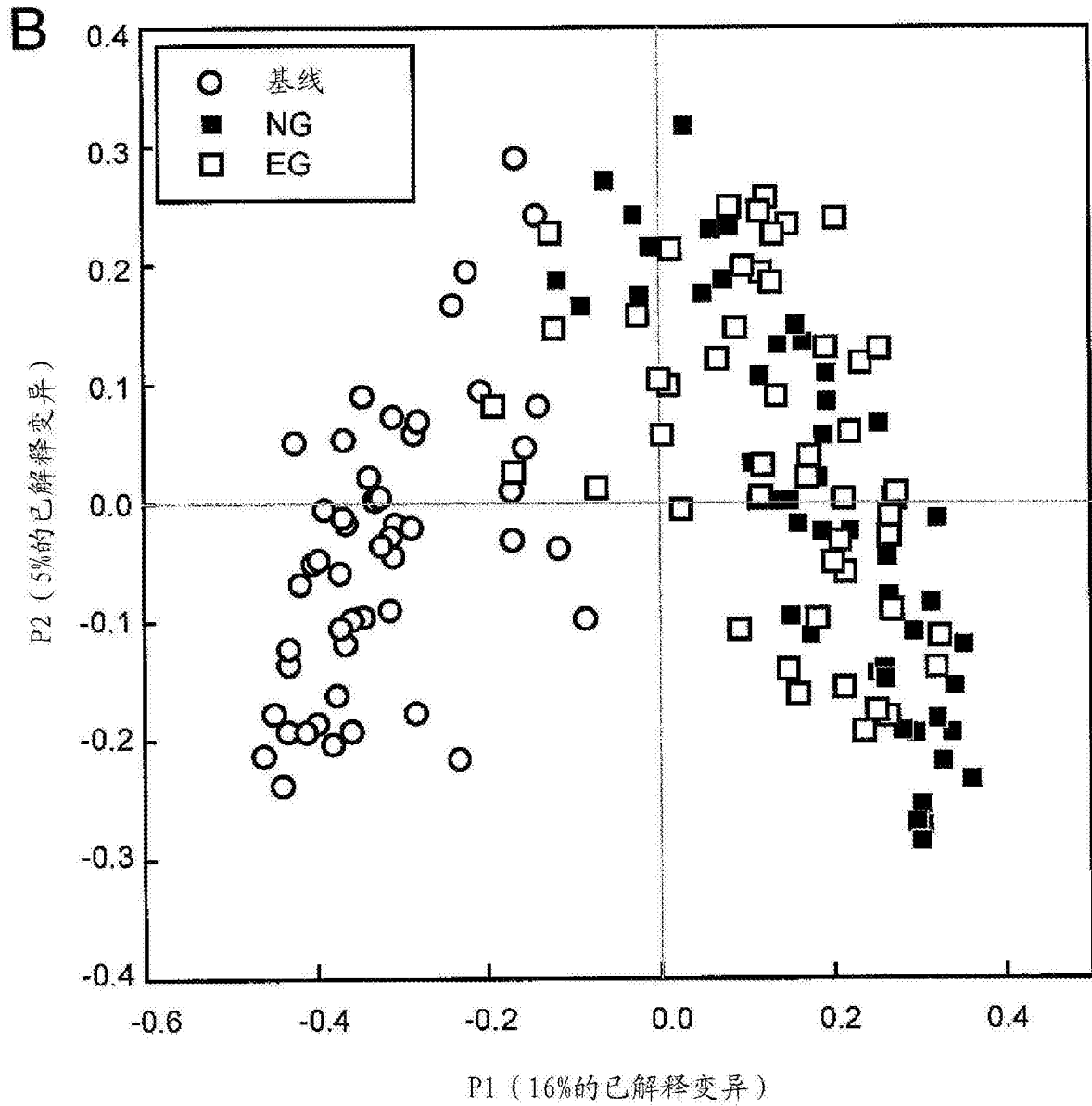


图5B

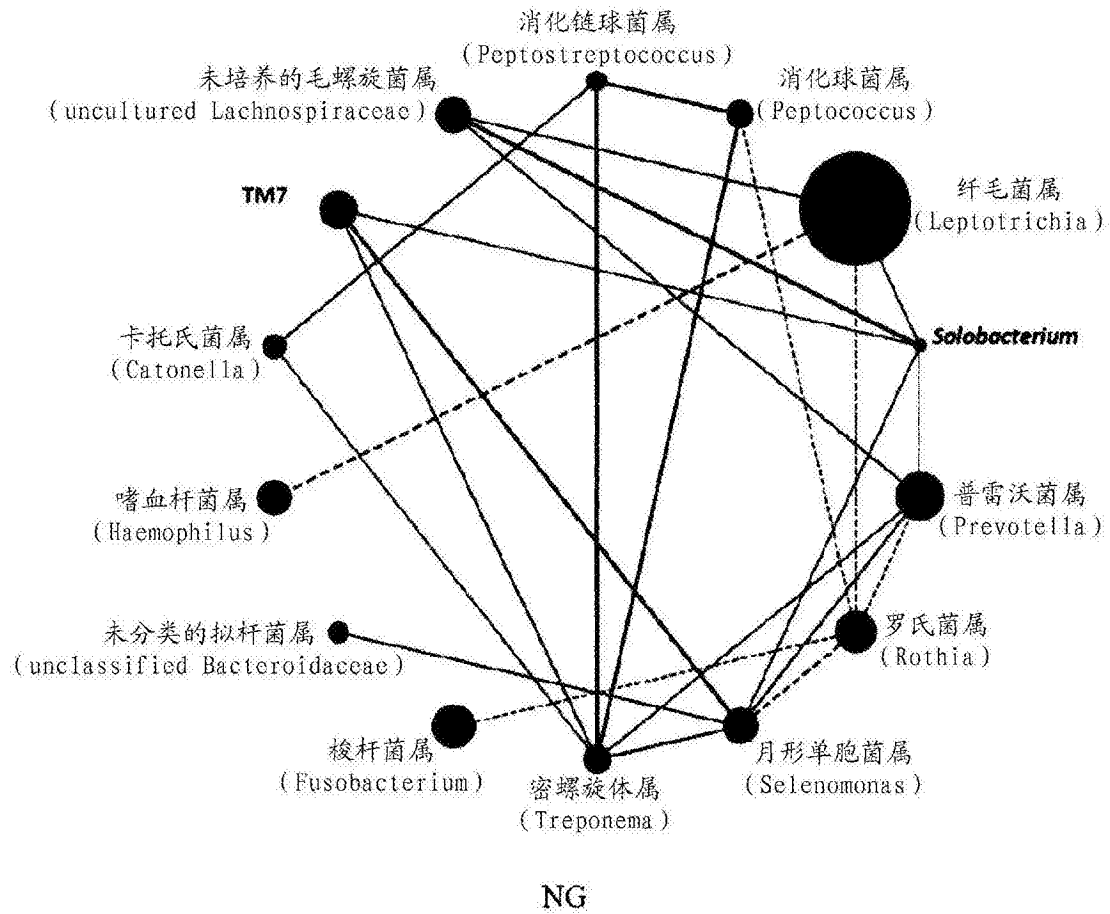


图6A

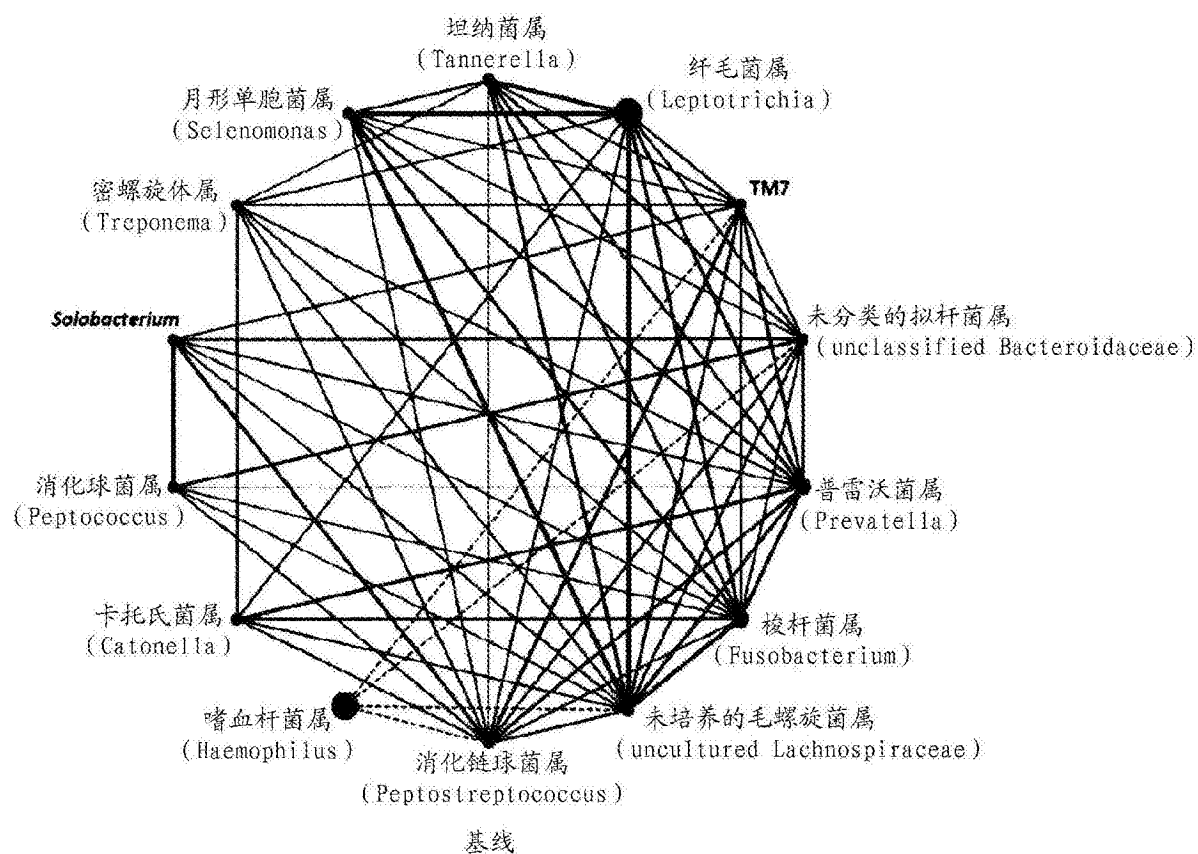


图6B

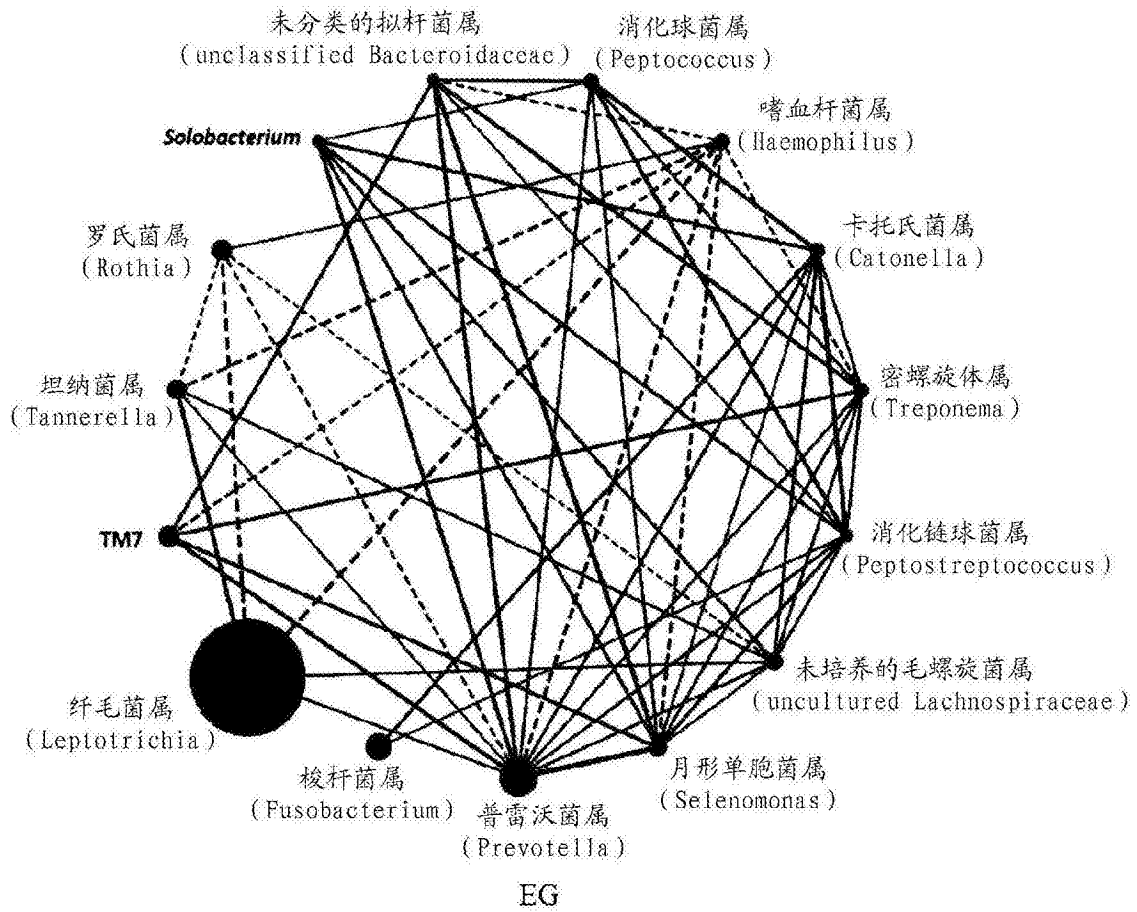


图6C

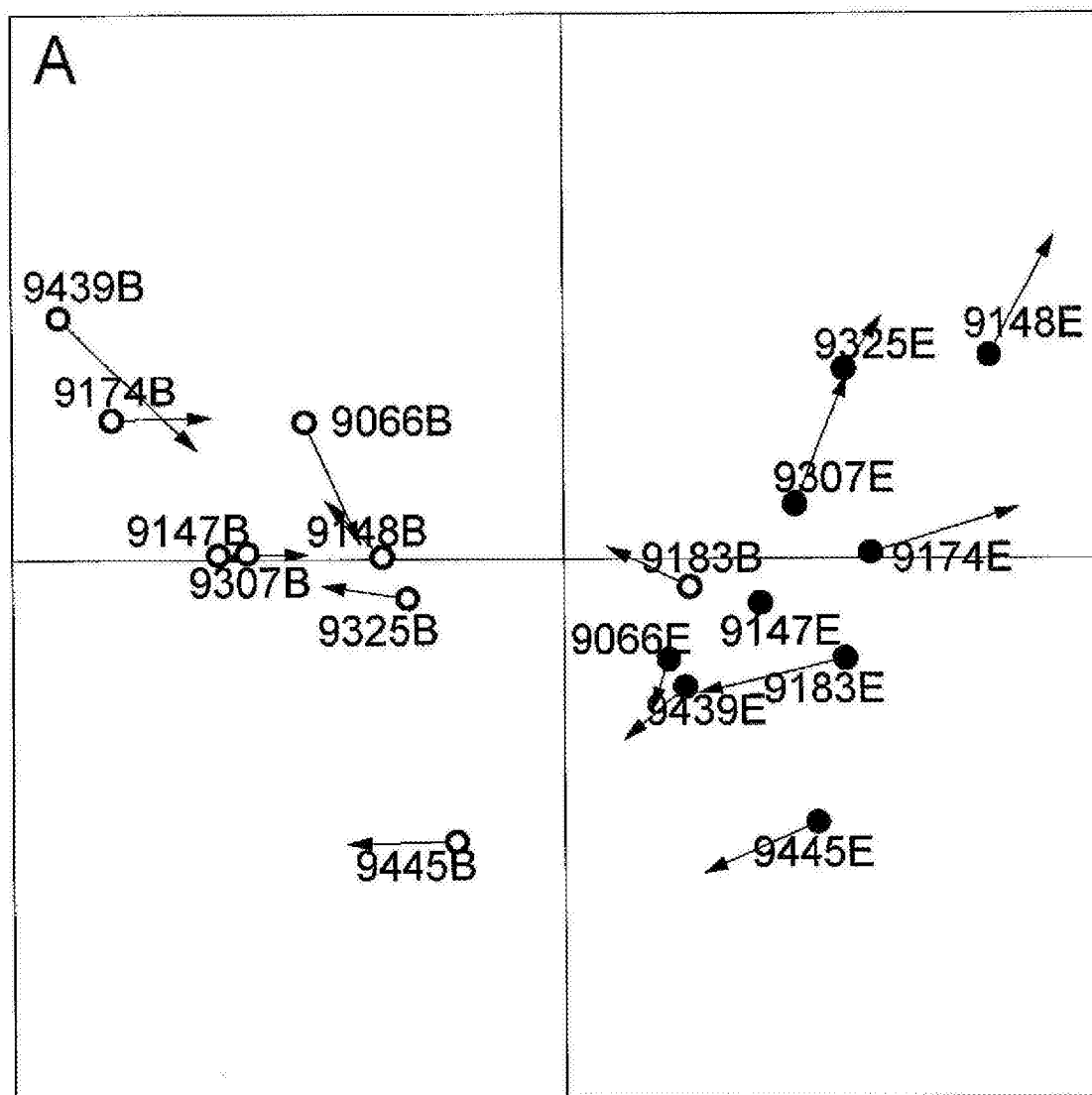


图7A

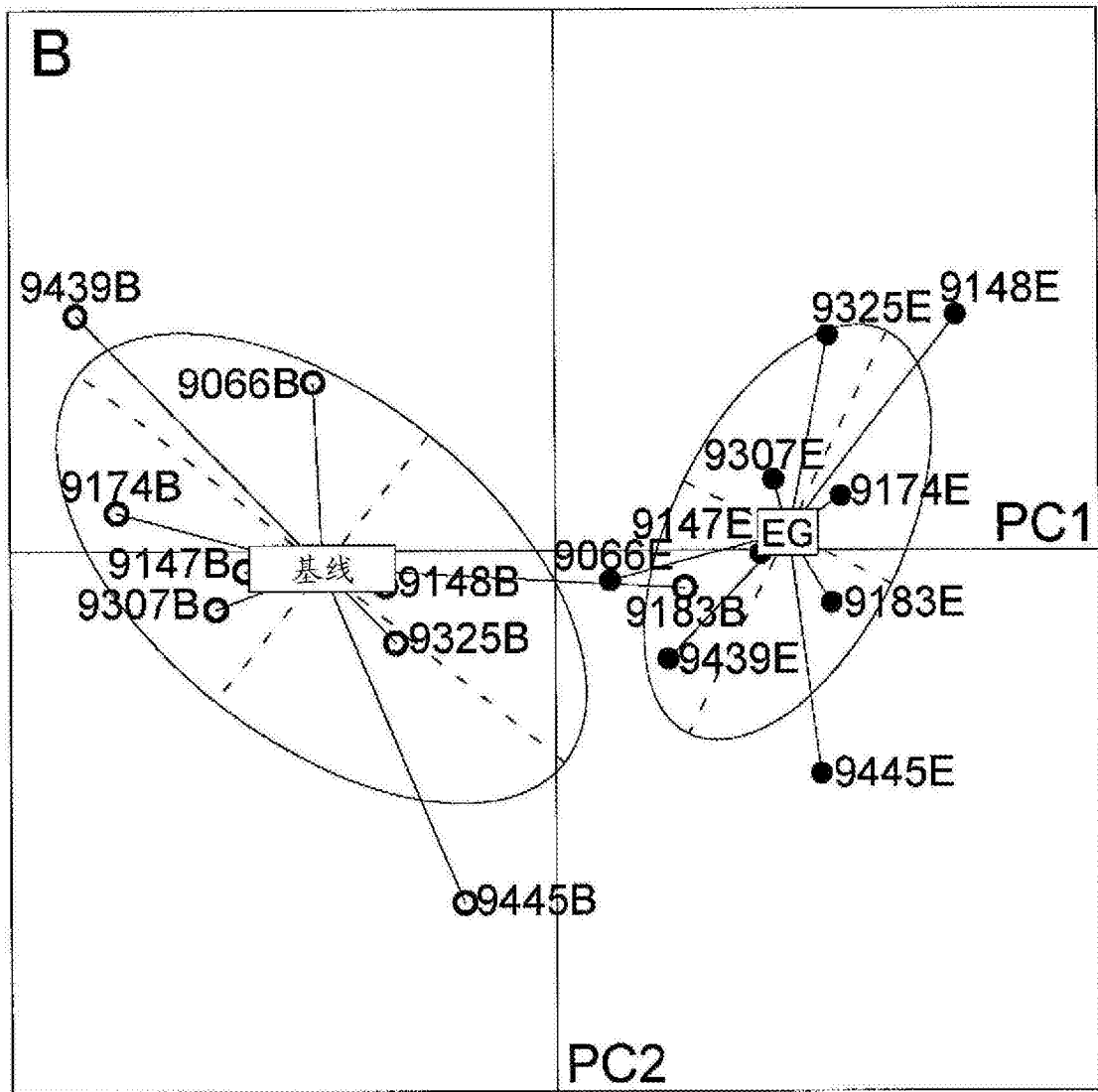


图7B

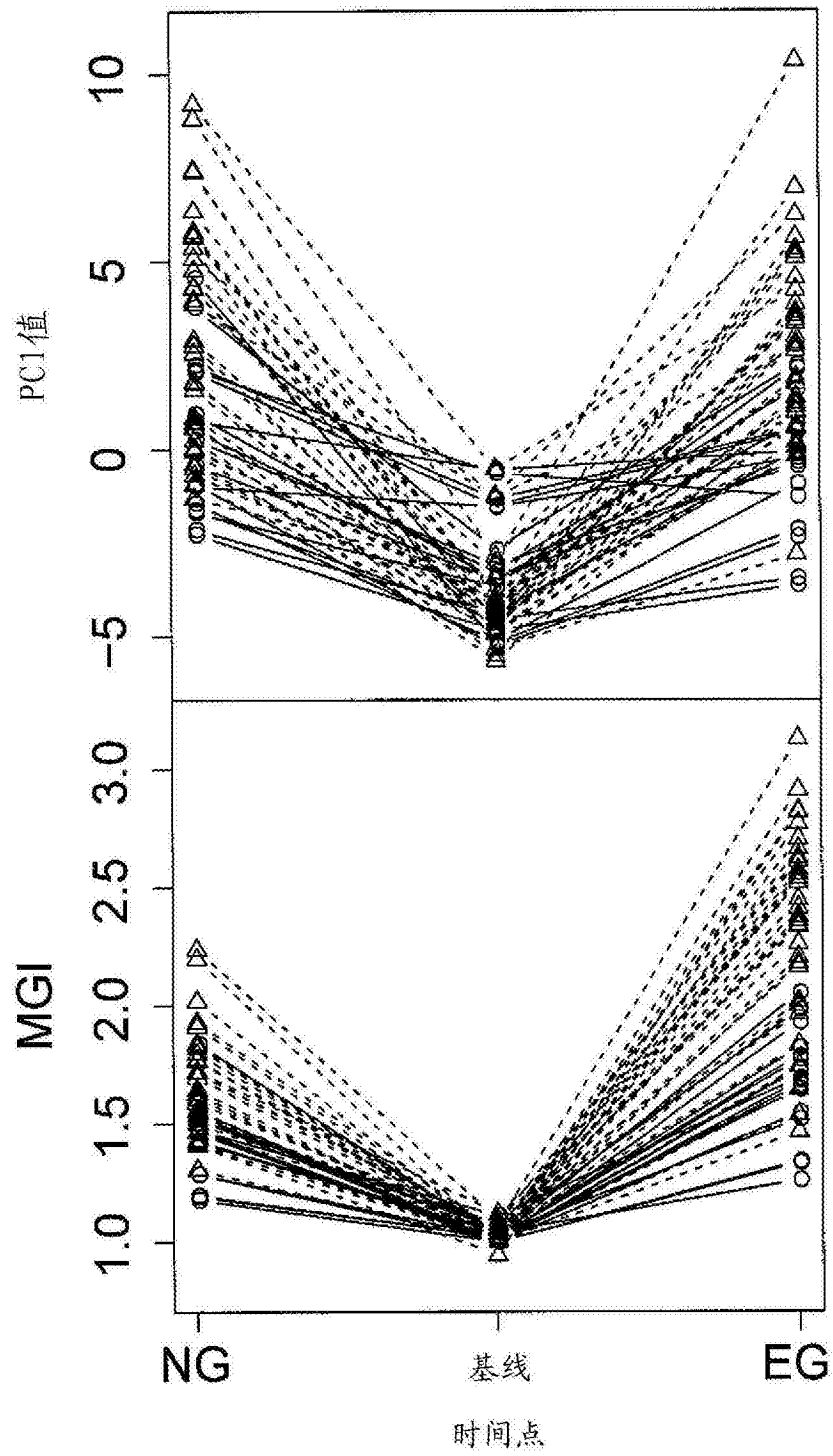


图8A

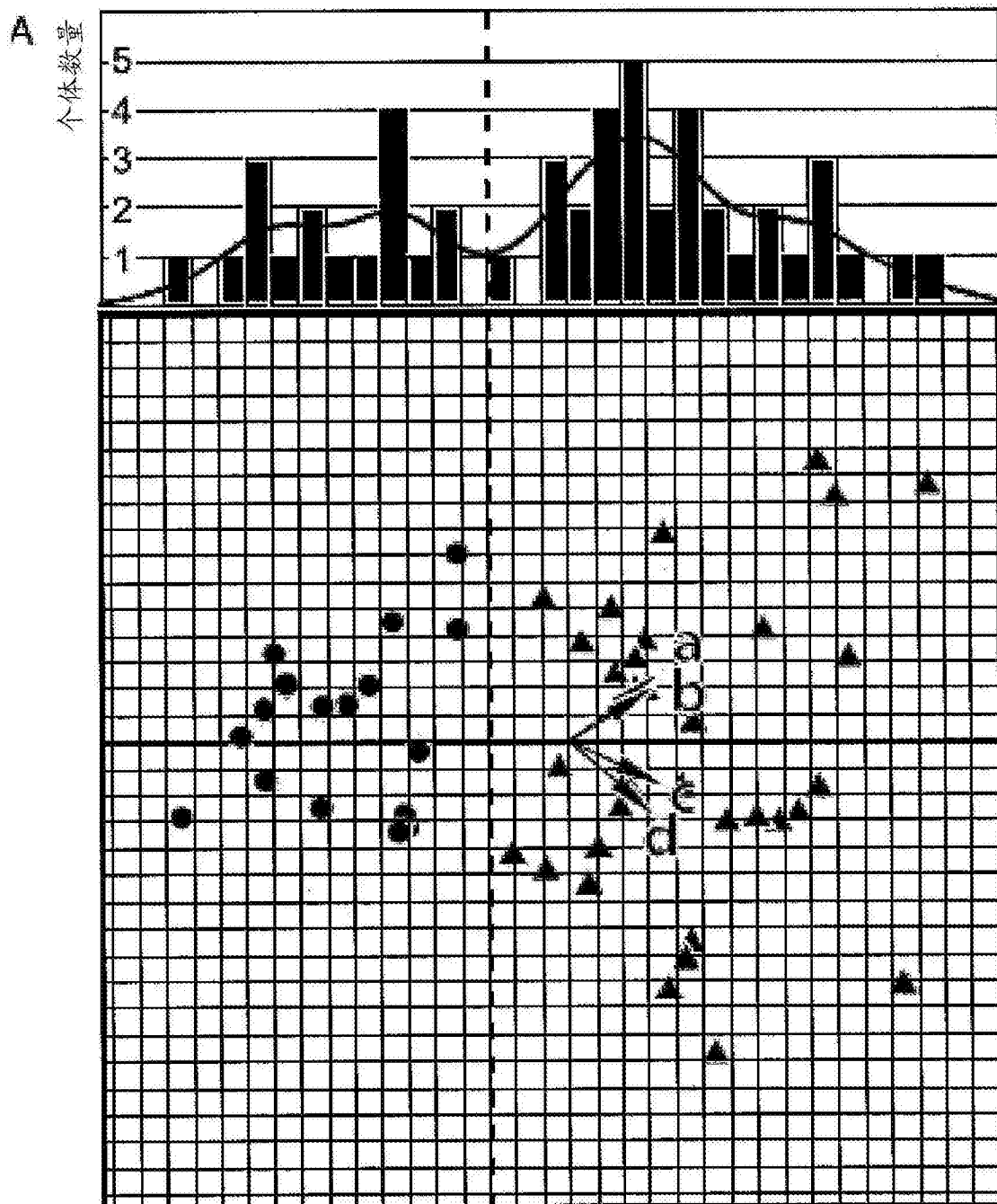


图8B

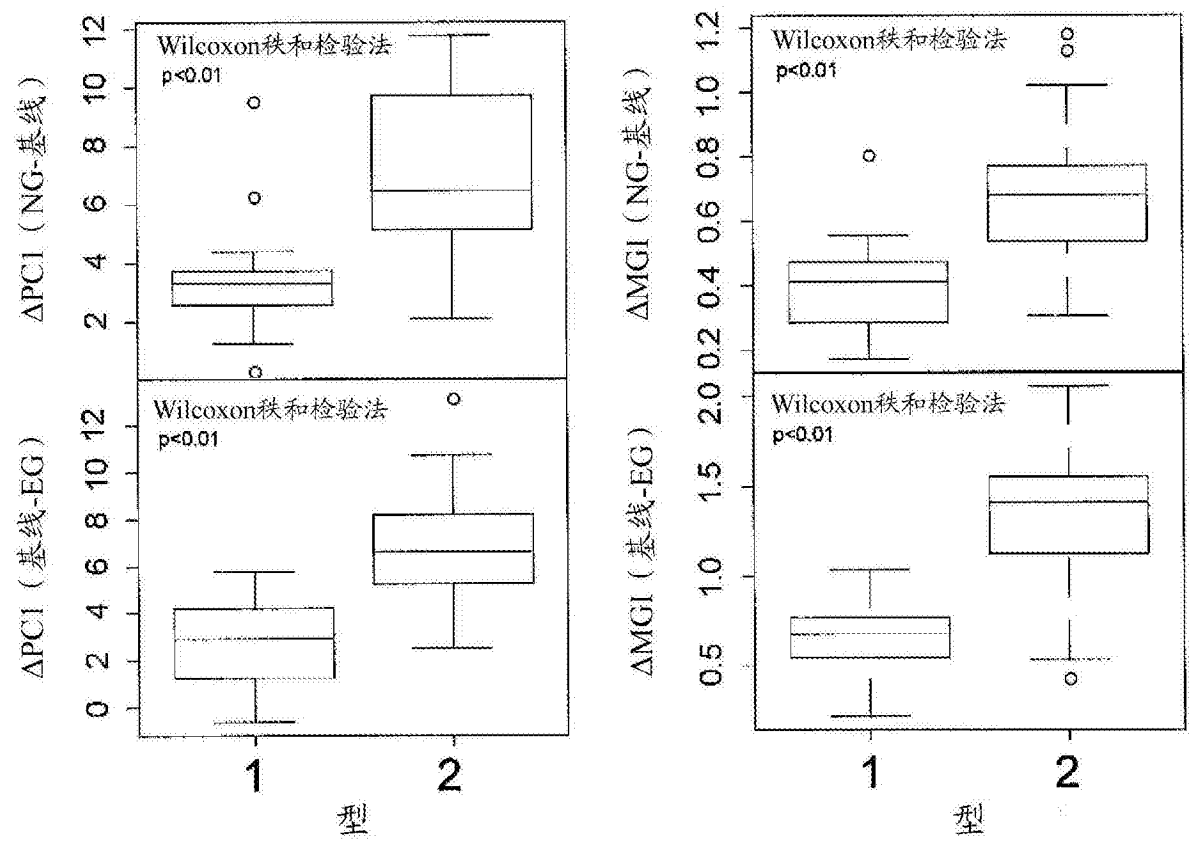


图8C

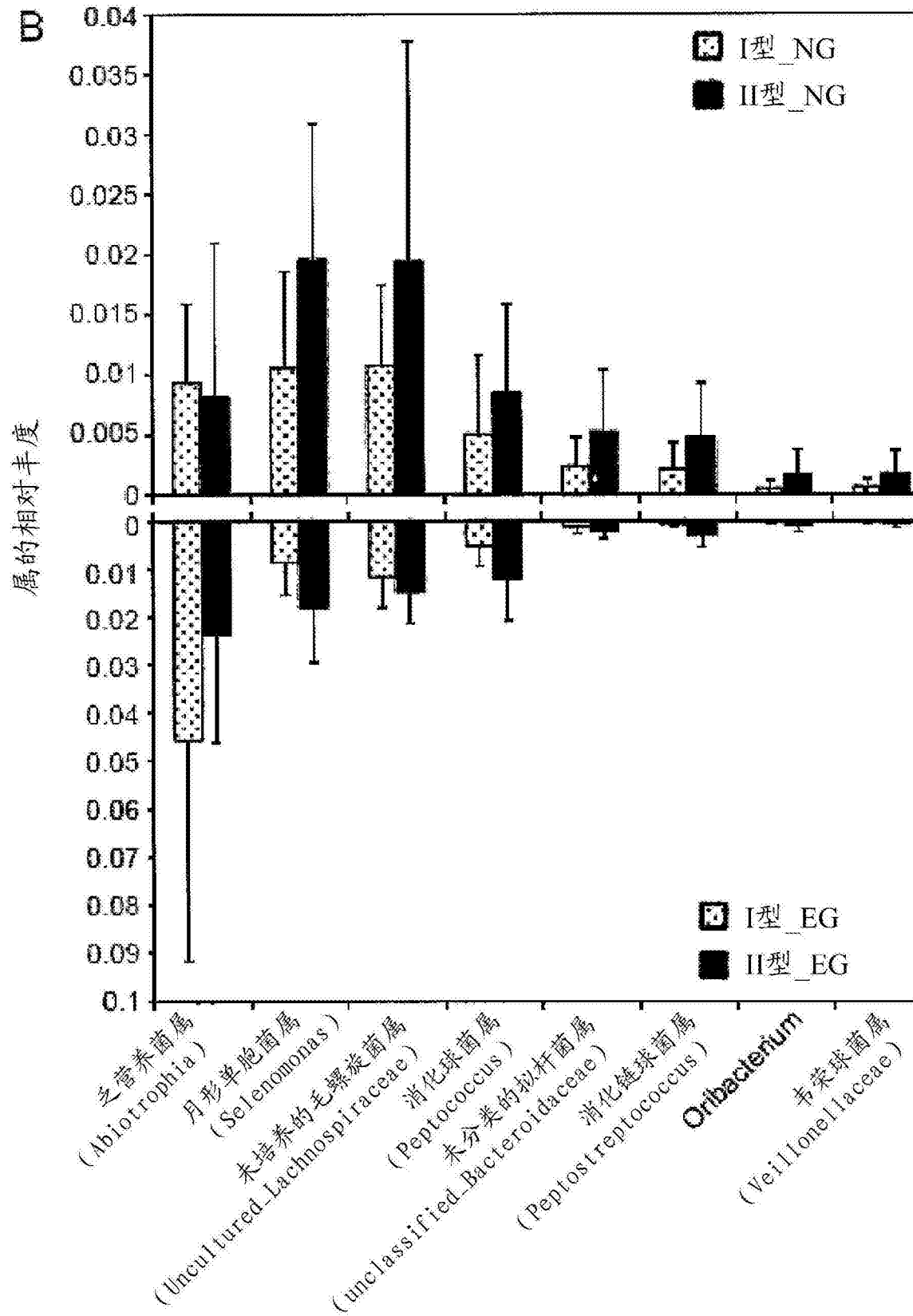


图8D

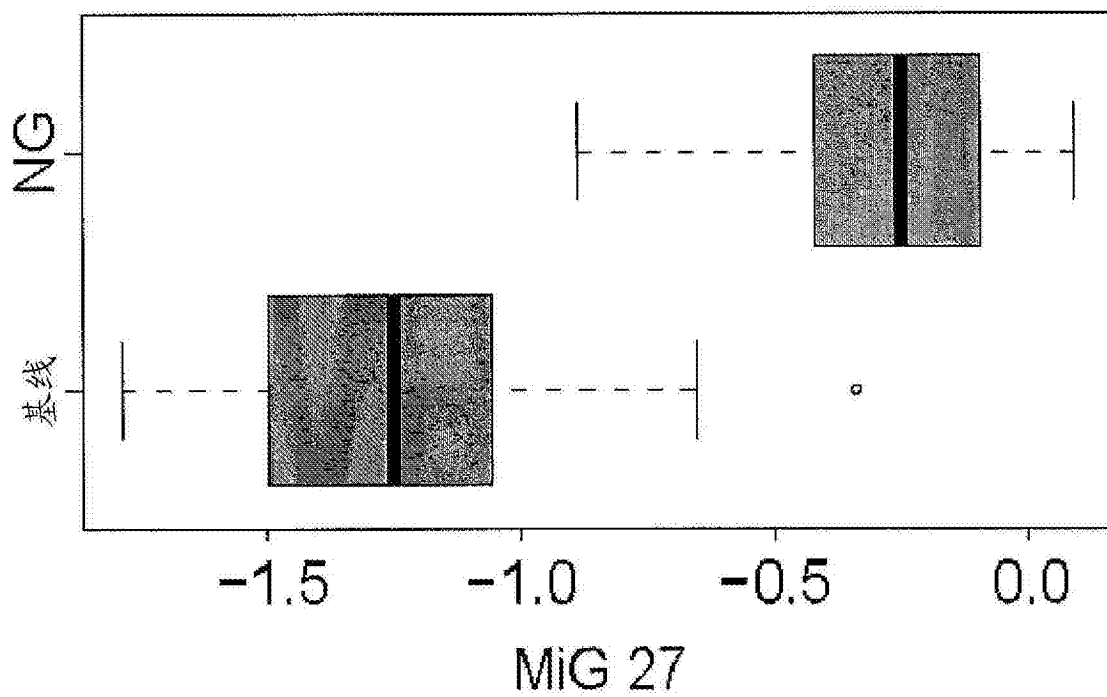


图9

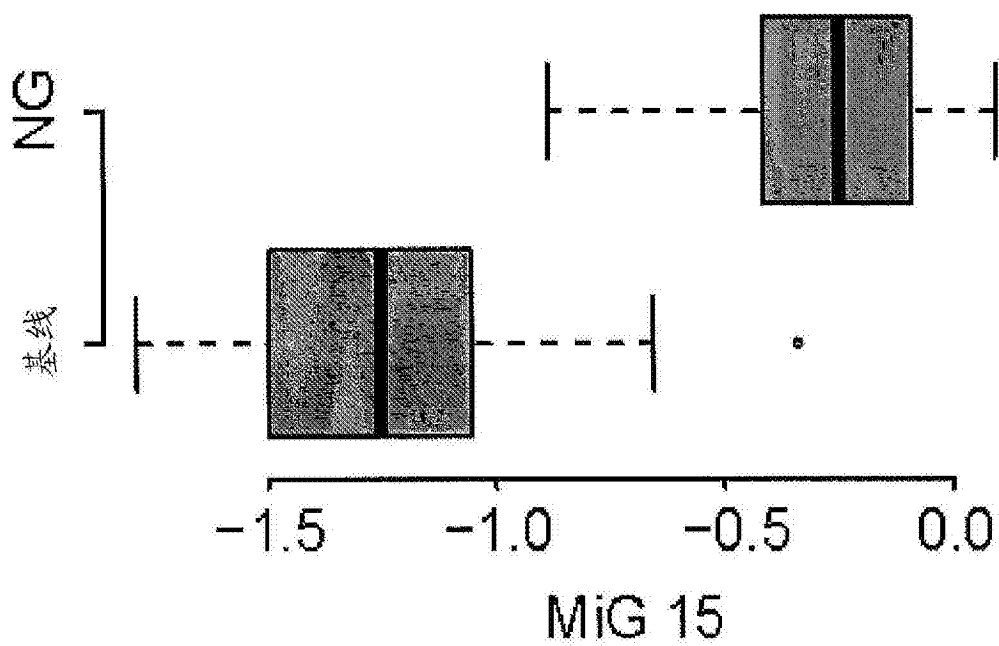


图10

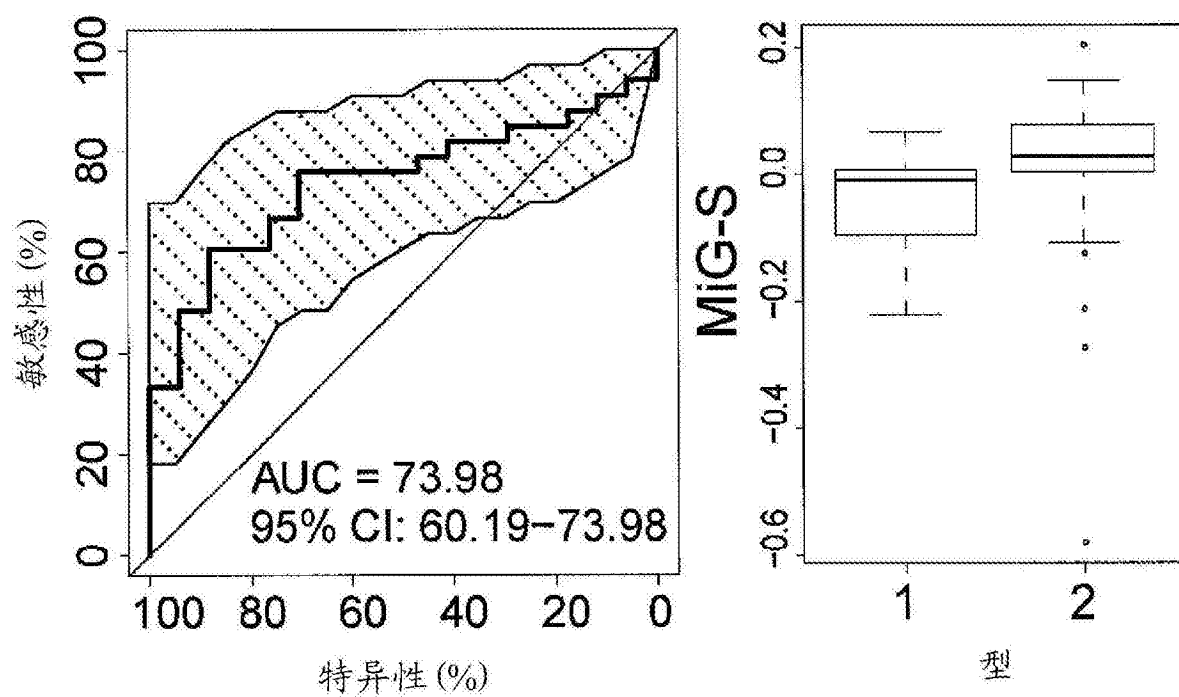


图11