

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019年10月31日 (31.10.2019)



(10) 国际公布号
WO 2019/204985 A1

- (51) 国际专利分类号:
C12Q 1/68 (2018.01) *C12N 1/00* (2006.01)
- (21) 国际申请号: PCT/CN2018/084276
- (22) 国际申请日: 2018年4月24日 (24.04.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 深圳华大生命科学研究院 (BGI SHENZHEN) [CN/CN]; 中国广东省深圳市盐田区盐田街道北山工业区综合楼, Guangdong 518083 (CN).
- (72) 发明人: 王奇(WANG, Qi); 中国广东省深圳市盐田区盐田街道北山工业区综合楼, Guangdong 518083 (CN)。 郭锐进(GUO, Ruijin); 中国广东省深圳市盐田区盐田街道北山工业区综合楼, Guangdong 518083 (CN)。 鞠艳梅(JU, Yanmei); 中国广东省深圳市盐田区盐田街道北山工业区综合楼, Guangdong 518083 (CN)。 贾慧珺(JIA, Huijue); 中国广东省深圳市盐田区盐田街道北山工业区综合楼, Guangdong 518083 (CN)。
- (74) 代理人: 北京清亦华知识产权代理事务所(普通合伙) (TSINGYIHUA INTELLECTUAL PROPERTY LLC); 中国北京市海淀区清华园清华大学照澜院商业楼301室, Beijing 100084 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

(54) Title: OSTEOPOROSIS BIOMARKER AND USE THEREOF

(54) 发明名称: 骨质疏松生物标志物及其用途

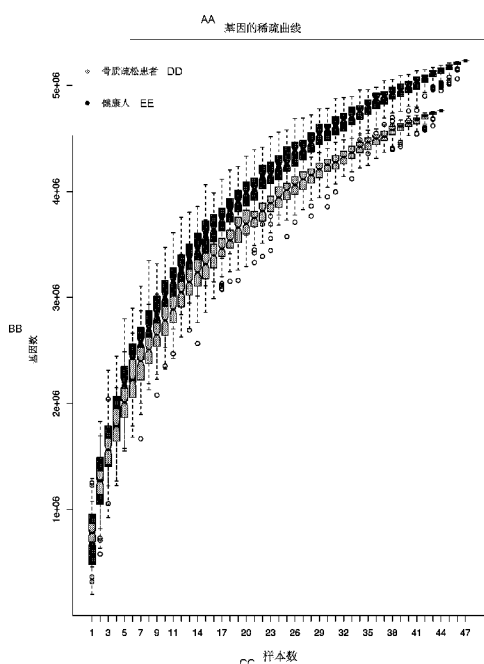


图1
AA Rarefaction curve of genes
BB Gene number
CC Sample number
DD Osteoporosis patient
EE Healthy person

(57) Abstract: The present invention relates to an osteoporosis biomarker and use thereof. The biomarker comprises *Bacteroides thetaiotaomicron*, *Bacteroides uniformis*, *Bacteroides intestinalis*, *Bacteroides dorei*, *Ruminococcus* sp., *Akkermansia muciniphila*, *Parabacteroides merdae*, *Ruminococcus torques*, *Dialister invisus*, and analogues thereof.

(57) 摘要: 本发明涉及骨质疏松生物标志物及其用途。所述生物标志物包括多形拟杆菌(*Bacteroides thetaiotaomicron*), 单形拟杆菌(*Bacteroides uniformis*), *Bacteroides intestinalis*, *Bacteroides dorei*, 瘤胃球菌(*Ruminococcus* sp.), 嗜粘蛋白-艾克曼菌 *Akkermansia muciniphila*, *Parabacteroides merdae*, 扭链瘤胃球菌 *Ruminococcus torques*, *Dialister invisus* 及它们的类似物。

WO 2019/204985 A1

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

骨质疏松生物标志物及其用途

技术领域

本发明涉及生物医药领域，具体地，本发明涉及骨质疏松生物标志物及其用途。尤其
5 地，本发明涉及骨质疏松或相关疾病的生物标志物、诊断或预测骨质疏松或相关疾病风险的方法、试剂盒及骨质疏松生物标志物在制备试剂盒中的用途。

背景技术

骨质疏松症（英文：osteoporosis，来自希腊文 porous bones，意为“多孔的骨头”）是一
10 种因骨质密度下降，而骨折风险提高的疾病。起因是矿物质大量流失，导致骨头中的钙质不断流失到血液中；骨质疏松症也是中高年龄族群最常见的骨折原因。易于因骨质疏松而骨折的骨骼部位有脊椎、前臂骨、髋关节骨。通常骨折前都不会有任何症状，一直到骨骼变得松软易折，稍微受压就会断裂；发生慢性疼痛及机能衰退后，就连日常活动都会导致再度骨折。

15 骨质疏松症会随着年纪增加而加重，约 15% 的白人 50 多岁起会出现症状，80 岁以上则会提高到 70%；骨质疏松亦多见于女性，甚于男性患者。发达国家中利用筛检发现 2% - 8% 男性及 9% - 38% 的女性确诊罹患骨质疏松；发展中国家的发病率则尚不明朗。2010 年，欧洲有将近 2200 万女性患者和 550 万左右男性患者，同年在美国，发现有 800 万左右女性和 100-200 万男性患者。骨质疏松症的危险因子包括性别（尤其是女性）、太早停
20 经、种族（尤其是白人和亚洲人）、骨头结构较细、身体质量指数过低、抽烟、酗酒、活动量不足、具有家族病史。

虽然骨质疏松症多数情况下并不会直接导致死亡，但骨质疏松症增加骨折机会，从而
影响病患者的健康和独立生活能力，更大大增加社会医疗负担。现有的骨密度筛查只是表面的了解骨密度的变化，但并不能从整体上对患者进行骨质疏松症状的评估。

25 有关骨质疏松的诊断和筛查还需要进一步改进。本领域迫切需要对骨质疏松生物标志物进行进一步的研究。

发明内容

本申请是基于发明人对以下事实和问题的发现和认识作出的：肠道微生物是存在于人
30 体肠道中的微生物群落，是人体的“第二基因组”。人体肠道菌群和宿主构成一个相互关联的整体，肠道微生物不仅能降解食物中消化的营养成分、宿主维生素以及其他的一些营养物质，还能促进肠上皮细胞的分化与成熟，从而激活肠道免疫系统以及调节宿主能量存储

与代谢，这些在人体的消化吸收、免疫反应、代谢活性等方面都发挥着重要的作用。因此，本发明发明人通过对骨质疏松患者以及健康人群的肠道菌群以及基因序列进行分析，从而筛选出与骨质疏松症相关性高的生物标志物，并且利用该标志物能够准确地诊断骨质疏松或相关疾病或者预测患病风险，并且可以用于监测治疗效果。

5 因此，本发明的目的在于提供用于评估骨质疏松风险或者早期诊断骨质疏松的生物标志物，以及骨质疏松的诊断和患病风险评估方法，可以解决现有骨质疏松诊断方法不能做到早期预警、不能预测骨质疏松发病以及发展的趋势等缺点。从而可以应用于预测骨质疏松发病以及发展的趋势，以及应用于疾病病理分型。

10 据认为，由于以下原因，骨质疏松相关的生物标记物对早期诊断是有价值的。第一，本发明的标记物具有特异性和灵敏性。第二，粪便的分析保证准确性、安全性、可负担性和患者依从性。并且粪便的样本是可运输的。基于聚合酶链反应（PCR）的试验舒适且无创，所以人们会更容易参与给定的筛选程序。第三，本发明的标记物还可以用作用于对骨质疏松患者进行治疗监测的工具以检测对治疗的响应。

15 根据本发明的第一方面，本发明提供了一种生物标志物。根据本发明的实施例，该生物标志物包括选自下列中的至少一种：

多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物，单形拟杆菌 (*Bacteroides uniformis*) 或其类似物，*Bacteroides intestinalis* 或其类似物，*Bacteroides dorei* 或其类似物，瘤胃球菌 (*Ruminococcus sp.*) 或其类似物，嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物，*Parabacteroides merdae* 或其类似物，扭链瘤胃球菌 (*Ruminococcus torques*) 或其类似物，*Dialister invisus* 或其类似物，所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 类似物与多形拟杆菌 (*Bacteroides thetaiotaomicron*) 的基因组序列相比，比对相似度在 85% 以上，所述单形拟杆菌 (*Bacteroides uniformis*) 类似物与单形拟杆菌 (*Bacteroides uniformis*) 的基因组序列相比，比对相似度在 85% 以上，所述 *Bacteroides intestinalis* 类似物与 *Bacteroides intestinalis* 的基因组序列相比，比对相似度在 85% 以上，所述 *Bacteroides dorei* 类似物与 *Bacteroides dorei* 的基因组序列相比，比对相似度在 85% 以上，所述瘤胃球菌 (*Ruminococcus sp.*) 类似物与瘤胃球菌 (*Ruminococcus sp.*) 的基因组序列相比，比对相似度在 85% 以上，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 类似物与嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 的基因组序列相比，比对相似度在 85% 以上，所述 *Parabacteroides merdae* 类似物与 *Parabacteroides merdae* 的基因组序列相比，比对相似度在 20 25 30 85% 以上，所述扭链瘤胃球菌 (*Ruminococcus torques*) 类似物与扭链瘤胃球菌 (*Ruminococcus torques*) 的基因组序列相比，比对相似度在 85% 以上，所述 *Dialister invisus* 类似物与 *Dialister invisus* 的基因组序列相比，比对相似度在 85% 以上。这些生物标志物均可以作为骨质疏松

症检测的生物学标记物，可以通过确定对象肠道菌群中是否存在这些标志物中的一种或者两种或者多种，从而有效地确定检测对象是否患有或者易感骨质疏松症（即预测患有骨质疏松症的风险），并且还可以进一步将这些生物标志物用于监控骨质疏松症患者的治疗效果。另外，当健康样本量足够多的时候，本领域技术人员还可以根据检验和计算方法，得到每个生物标志物在肠道中的正常值或者正常的范围，从而用来指示每种标志物在健康样本中的含量，由此，可以通过对样本中这些生物标志物的至少一种在肠道菌群中的含量进行检测，来确定对象是否患有或者易感骨质疏松症，同时可以用来监控骨质疏松症患者的治疗效果的效率。而且本领域技术人员可知的是，当某种未知的微生物或者某种核酸来源的某些基因序列与某种已知菌株的基因序列相比，比对相似度在 85% 以上的时候，即可认为该微生物与该菌株属于同一属，或者可以将基因序列归类到与该菌株同属，而同属的微生物通常具有相同或相似的功能，因此，也可以利用这些类似物作为骨质疏松症的标志物。

本发明中比对相似性，也可以称为比对相似度，是指序列比对过程中目标序列（待确定的序列）和参考序列（已知序列）之间相同碱基或氨基酸残基序列所占比例的大小。

根据本发明的实施例，所述 *Bacteroides intestinalis* 为 *Bacteroides intestinalis* DSM 17393，所述瘤胃球菌（*Ruminococcus sp.*）为 *Ruminococcus sp.* 5_1_39BFAA，所述嗜粘蛋白-阿克曼菌（*Akkermansia muciniphila*）为嗜粘蛋白-阿克曼菌 ATCC BAA-835（*Akkermansia muciniphila* ATCC BAA-835），所述 *Parabacteroides merdae* 为 *Parabacteroides merdae* ATCC 43184，所述扭链瘤胃球菌（*Ruminococcus torques*）为扭链瘤胃球菌 L2-14（*Ruminococcus torques* L2-14），所述 *Dialister invisus* 为 *Dialister invisus* DSM 15470。这些生物标志物作为相应菌种的代表菌株，均可以用来指示骨质疏松症或者骨质疏松症相关疾病的患病状态或者患病风险。

根据本发明的实施例，所述多形拟杆菌（*Bacteroides thetaiotaomicron*）类似物与多形拟杆菌（*Bacteroides thetaiotaomicron*）的基因组序列相比，比对相似度在 95% 以上，所述单形拟杆菌（*Bacteroides uniformis*）类似物与单形拟杆菌（*Bacteroides uniformis*）的基因组序列相比，比对相似度在 95% 以上，所述 *Bacteroides intestinalis* 类似物与 *Bacteroides intestinalis* 的基因组序列相比，比对相似度在 95% 以上，所述 *Bacteroides dorei* 类似物与 *Bacteroides dorei* 的基因组序列相比，比对相似度在 95% 以上，所述瘤胃球菌（*Ruminococcus sp.*）类似物与瘤胃球菌（*Ruminococcus sp.*）的基因组序列相比，比对相似度在 95% 以上，所述嗜粘蛋白-阿克曼菌（*Akkermansia muciniphila*）类似物与嗜粘蛋白-阿克曼菌（*Akkermansia muciniphila*）的基因组序列相比，比对相似度在 95% 以上，所述 *Parabacteroides merdae* 类似物与 *Parabacteroides merdae* 的基因组序列相比，比对相似度在 95% 以上，所述扭链瘤胃球菌（*Ruminococcus torques*）类似物与扭链瘤胃球菌（*Ruminococcus torques*）的基因组序

列相比，比对相似度在 95%以上，所述 *Dialister invisus* 类似物与 *Dialister invisus* 的基因组序列相比，比对相似度在 95%以上。本领域技术人员可知的是，当某种未知微生物或者某种核酸来源的基因序列与某种已知菌株相比，比对相似度在 95%以上的时候，即可以认为该微生物与该菌株同种，或者可以将基因序列归类到与该菌株同种。由此，本领域技术人员可以直接通过对检测对象中的核酸序列信息获取，然后将其与这些菌株的基因组序列进行比对，如有 95%以上的序列相似性，则就可以作为检测对象是否患有骨质疏松症或者易感骨质疏松症的标志。

根据本发明的实施例，当所述各菌类似物与相应的菌的基因组序列相比，比对覆盖度在 80%以上，且比对相似度在 85%以上时，均可以认为这些类似物与相应菌属于同一属，可以作为骨质疏松症的标志物。优选地，当这些类似物与相应的菌的比对覆盖度在 80%以上，且比对相似度在 95%以上时，均可以认为这些类似物与相应菌同种，可以作为骨质疏松症的标志物。

本发明中比对覆盖度，指的是对目标序列与参考序列比对的过程中，目标序列中拿来和参考序列进行比对的序列的长度占检测序列总长度的比例。

根据本发明的第二方面，本发明提出了一种诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险的方法。根据本发明的实施例，所述方法包括步骤：（1）从所述对象中采集样本；（2）确定步骤（1）中获得的所述样本中根据本发明第一方面所述的生物标志物的相对丰度信息；（3）将步骤（2）中所述的相对丰度信息与参考数据集或参考值进行比较。所述方法不仅仅用于专利法意义上的疾病诊断，同时可以用作科学研究或者其他个人遗传信息的丰富以及遗传信息库的丰富等非疾病诊断。利用检测对象中的各生物标志物的相对丰度信息与参考数据集或参考值进行比较，来确定对象是否患有骨质疏松症或相关疾病，或者预测其患有骨质疏松症或者相关疾病的风险。

本发明中所述参考数据集指的是对已确诊为患病个体和健康个体的样本进行操作，所获得的各生物标志物的相对丰度信息，用来作为每种生物标志物的相对丰度的参考。在本发明的一个实施方案中，参考数据集是指训练数据集。根据本发明，所述训练集是指和验证集具有本领域公知的含义。在本发明的一个实施方案中，所述训练集是指包含一定样本数的骨质疏松症受试者和非骨质疏松症受试者待测样本中的各生物标志物的含量的数据集。所述验证集是用来测试训练集性能的独立数据集。

本发明中所述参考值指的是健康对照的参考值或正常值。本领域技术人员已知，当样本容量足够大时，可利用本领域公知的检测和计算方法获得样品中每个生物标志物的正常值（绝对值）的范围。当采用测定方法检测生物标志物的水平时，可将样品中的生物标志物水平的绝对值直接与参考值进行比较，以评估患病风险以及诊断或早期诊断骨质疏松症

或相关疾病，任选地，可以包括统计方法。

本发明中所述骨质疏松症相关疾病，意指与骨质疏松症相互关联的疾病，包括可以引发骨质疏松症的前期的症状或疾病，以及由骨质疏松症引发的后续的或者并发的症状或疾病。

5 根据本发明的实施例，所述方法可以进一步附加如下技术特征：

根据本发明的实施例，所述参考数据集包括来自多个骨质疏松患者和多个健康对照的样本中的生物标志物的相对丰度信息，所述生物标志物为根据本发明第一方面所述的生物标志物。

10 根据本发明的实施例，在将步骤（2）中所述的相对丰度信息与参考数据集进行比较的步骤中，还包括执行多元统计模型以获得患病概率。利用多元统计模型可以实现快速高效检测。

根据本发明的实施例，所述多元统计模型为随机森林模型。

根据本发明的实施例，所述患病概率大于阈值表明所述对象患有骨质疏松或相关疾病或者有患有骨质疏松或相关疾病的风险。

15 根据本发明的实施例，所述阈值为 0.5。

20 根据本发明的实施例，当与参考值比较时，所述嗜粘蛋白-艾克曼菌（*Akkermansia muciniphila*）或其类似物、*Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的减少表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中；所述多形拟杆菌（*Bacteroides thetaiotaomicron*）或其类似物、单形拟杆菌（*Bacteroides uniformis*）或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌（*Ruminococcus sp.*）或其类似物、扭链瘤胃球菌（*Ruminococcus torques*）或其类似物的增加表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中。

25 根据本发明的实施例，步骤（2）中所述生物标志物的相对丰度信息是利用测序方法得到的，进一步包括：从所述对象的所述样本中分离得到核酸样本，基于所获得的所述核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果，以及基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物的相对丰度信息。根据本发明的一种实施例，可以利用 SOAP2 和 MAQ 的至少一种将测序结果与参考基因集进行比对，由此，可以提高比对的效率，进而可以提高骨质疏松症检测的效率。根据本发
30 明的实施例，可以同时多种（至少两种）生物标志物进行检测，可以提高骨质疏松症检测的效率。

根据本发明的实施例，所述参考基因集包括从多个骨质疏松患者和多个健康对照的样

本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集。本发明中的参考基因集可以是已有的基因集，如现有的已经公开的肠道微生物参考基因集；也可以是将多个骨质疏松症患者和多个健康对照的样品进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集，由此获得的参考基因集信息更全面，检测结果更可靠。

本发明中所述非冗余基因集作本领域技术人员通常的理解来解释，简单来说去除冗余基因后的剩余基因的集合。冗余基因通常指的是一条染色体上出现的一个基因的多个复份。

根据本发明的实施例，所述样本为粪便样本。

10 根据本发明的实施例，所述测序方法是通过第二代测序方法或第三代测序方法进行的。进行测序的手段并不受特别限制，通过二代或者三代测序的方法进行测序，可以实现快速高效的测序。

15 根据本发明的实施例，所述测序方法是通过选自 HiSeq2000、SOLiD、454、和单分子测序装置的至少一种进行的。由此，能够利用这些测序装置的高通量、深度测序的特点，从而有利于对后续测序数据进行分析，尤其是进行统计学检验时的精确性和准确度。

根据本发明的第三方面，本发明提出了一种试剂盒，包括用于检测生物标志物的试剂，所述生物标志物包括根据本发明的第一方面的生物标志物。利用该试剂盒，可以确定这些标志物在肠道菌群中的相对丰度，由此，可以通过所得到的相对丰度值，从而确定对象是否患有或者易感骨质疏松症，以及用于监控骨质疏松症患者的治疗效果。

20 根据本发明的实施例，所述试剂盒包括一组参考数据集或者参考值，用来作为每种生物标志物的相对丰度的参考。优选可以将参考数据集或者参考值附在物理载体上，例如光盘，如 CD-ROM 等。

25 根据本发明的实施例，所述试剂盒还包括第一计算机程序产品，该第一计算机程序产品用来执行获得所述的参考数据集或者参考值。即该第一计算机程序产品用来执行获得诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的一组参考数据集或者参考值。

根据本发明的实施例，所述试剂盒还包括第二计算机程序产品，该第二计算机程序产品还可以用来执行根据本发明第二方面所述的诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险的方法。

30 根据本发明的第四方面，本发明提出了生物标志物在制备试剂盒中的用途，所述试剂盒用于诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险。

根据本发明的实施例，所述诊断或预测包括以下步骤：1) 从所述对象中采集样本；2) 确定步骤 1) 中获得的所述样本中的生物标志物的相对丰度信息，所述生物标志物为根据本发明第一方面所述的生物标志物；3) 将步骤 2) 中所述的相对丰度信息与参考数据集或参考值进行比较。根据所述的试剂盒，可以确定这些标志物在肠道菌群中的相对丰度，由此
5 可以通过所得到的相对丰度值，从而确定对象是否患有或者易感骨质疏松症，以及用于监控骨质疏松症患者的治疗效果的效率。

根据发明的实施例，以上生物标志物在制备试剂盒中的用途，可以进一步附加如下技术特征：

10 根据本发明的实施例，所述参考数据集包括来自多个骨质疏松患者和多个健康对照的样本中的生物标志物的相对丰度信息，所述生物标志物为根据本发明的第一方面的所述生物标志物。

根据本发明的实施例，在将步骤 2) 中所述的相对丰度信息与参考数据集进行比较的步骤中，还包括执行多元统计模型以获得患病概率；优选地，所述多元统计模型为随机森林模型。

15 根据本发明的实施例，所述患病概率大于阈值表明所述对象患有骨质疏松或相关疾病或者有患有骨质疏松或相关疾病的风险；优选地，所述阈值为 0.5。

根据本发明的实施例，当与参考值比较时，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物、*Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的减少表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险
20 中；所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物、单形拟杆菌 (*Bacteroides uniformis*) 或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌 (*Ruminococcus sp.*) 或其类似物、扭链瘤胃球菌 (*Ruminococcus torques*) 或其类似物的增加表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中。

25 根据本发明的实施例，通过测序方法得到步骤 2) 中所述生物标志物的相对丰度信息，进一步包括：从所述对象的所述样本中分离得到核酸样本，基于所获得的所述核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果，以及基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物的相对丰度信息。

30 根据本发明的实施例，所述参考基因集包括从多个骨质疏松患者和多个健康对照的样本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集。

根据本发明的实施例，所述样本为粪便样本。

根据本发明的实施例，所述测序方法是通过第二代测序方法或第三代测序方法进行的。

根据本发明的实施例，所述测序方法是通过选自 HiSeq2000、SOLiD、454、和单分子测序装置的至少一种进行的。

5 根据本发明的第五方面，本发明提出了一种生物标志物作为靶点用于筛选治疗或者预防骨质疏松或相关疾病的药物的用途。根据本发明的实施例，所述生物标志物包括根据本发明的第一方面的生物标志物。根据本发明的实施例，可以利用候选药物使用前和使用后对这些生物标志物的影响，从而确定候选药物是否可以用于治疗或预防骨质疏松症。

10 根据本发明的第六方面，本发明提出了一种生物标志物在诊断对象是否患有骨质疏松症或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险中的用途。根据本发明的实施例，所述生物标志物包括根据本发明第一方面的生物标志物。

根据本发明的第七方面，本发明提出了一种药物，所述药物用于预防或治疗骨质疏松或相关疾病。根据本发明的实施例，所述药物能够使得检测对象中嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*)或其类似物、*Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的相对丰度值增加；或者所述药物能够使得多形拟杆菌 (*Bacteroides thetaiotaomicron*)或其类似物、单形拟杆菌 (*Bacteroides uniformis*)或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌 (*Ruminococcus sp.*)或其类似物、扭链瘤胃球菌 (*Ruminococcus torques*)或其类似物的相对丰度值减少。

15

本发明所取得的有益效果为：粪便是人体的代谢产物，其内不仅包含人体的代谢产物，还包括对我们的机体代谢和免疫以及机体其他功能的变化密切相关的肠道微生物，对粪便进行研究，发现在骨质疏松症患者和健康人群的肠道菌群的组成上存在明显的差异，可以准确地对骨质疏松症患者进行患病风险评估、早期诊断。本发明通过对骨质疏松症患者和健康人群的肠道菌群的比较和分析，得到多种相关的肠道菌株，结合高质量的骨质疏松症人群和非骨质疏松症人群 MLGs 作为训练集，能够准确地对骨质疏松症患者进行患病风险评估、早期诊断。该方法与目前常用的诊断方法相比，具有方便、快捷的特点。

20

25 本发明的附加方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

附图说明

30 本发明的上述和/或附加的方面和优点从结合下面附图对实施例的描述中将变得明显和容易理解，其中：

图 1 示出了根据本发明的一个实施例在基因水平上骨质疏松患者和健康对照的数量差异图，可以看出骨质疏松组和健康人组的肠道菌群组成存在差异情况。

图 2 示出了根据本发明的一个实施例随机森林分类器中 5 次 10 折交叉验证的错误率分布情况图。

图 3 示出了根据本发明的一个实施例基于随机森林模型（9 个肠道标志物）判别骨质疏松和健康对照组成的训练集的接收者操作曲线（ROC）和曲线下面积（AUC）。

5 图 4 示出了根据本发明的一个实施例基于随机森林模型（9 个肠道标志物），由健康对照和骨质疏松（健康：7 和 患病：7）组成的验证集的接收者操作曲线（ROC）和曲线下面积（AUC）。

10 图 5 示出了根据本发明的一个实施例中确定对象是否患有骨质疏松症或相关疾病或者预测对象是否患有骨质疏松症或相关疾病的设备的结构示意图，其中图 a 为所述设备的示意图，图 b 为设备中的生物标志物相对丰度确定装置的示意图。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

15 需要说明的是，本发明中所提供的术语的解释仅为了使本领域技术人员更好地理解本发明，并非对本发明限制。

针对现有骨质疏松诊断方法不能做到早期预警、不能预测骨质疏松发病以及发展的趋势等缺点，本发明提出一种用于评估骨质疏松症风险或者早期诊断骨质疏松症的生物标志物，以及骨质疏松症的诊断和患病风险评估方法，能预测骨质疏松发病以及发展的趋势，
20 应用于疾病病理分型。

生物标志物

根据本发明的第一方面，本发明提出了用于骨质疏松症的生物标志物。

本发明所用术语具有相关领域普通技术人员通常理解的含义。然而，为了更好地理解本发明，对一些定义和相关术语的解释如下：

25 根据本发明，术语“骨质疏松”，是一种因骨质密度下降，而骨折风险提高的疾病。骨质疏松可以经由临床诊断，也可以参考世界卫生组织（WHO）推荐的诊断标准，即基于 DXA 测定骨密度通常用 T-Score（T 值）表示， $T \text{ 值} = (\text{测定值} - \text{同性别同种族正常成人骨峰值}) / \text{正常成人骨密度标准差}$ ，例如当 T 值 ≤ -2.5 时，可以认为患有骨质疏松症。

根据本发明，生物标志物物质的水平通过相对丰度指示。

30 根据本发明，术语“生物标志物”，也称为“生物学标志物”，是指个体的生物状态的可测量指标。这样的生物标记物可以是在个体中的任何物质，只要它们与被检个体的特定生物状态（例如，疾病）有关系，例如，核酸标志物（也可以称为基因标志物，例如 DNA），

蛋白质标志物，细胞因子标志物，趋化因子标志物，碳水化合物标志物，抗原标志物，抗体标志物，物种标志物（种/属的标记）和功能标志物（KO/OG 标记）等。其中，核酸标志物的含义并不局限于现有可以表达为具有生物活性的蛋白质的基因，还包括任何核酸片段，可以为 DNA，也可以为 RNA，可以是经过修饰的 DNA 或者 RNA，也可以是未经修饰的 DNA 或者 RNA 或其组合。在本文中核酸标志物有时也可以称为特征片段。在本发明中，生物标志物也可以用“肠道标志物”来替代，因为本发明所发现的与骨质疏松症密切相关的几种生物标志物均存在于受试者的肠道内。生物标志物经过测量和评估，经常用以检查正常生物过程，致病过程，或治疗干预药理响应，而且在许多科学领域都是有用的。

根据本发明的实施例，可以运用高通量测序，批量分析健康人群和骨质疏松症患者的粪便样本。基于高通量测序数据，对健康人群与骨质疏松症患者群进行比对，从而确定与骨质疏松症患者群相关的特异性核酸序列。简言之，其步骤如下：

样品的收集与处理：收集健康人群与骨质疏松症患者群的粪便样本，使用试剂盒进行 DNA 提取，得到核酸样本；

文库构建和测序：DNA 文库构建和测序是利用高通量测序进行，以便得到粪便样品中所包含肠道微生物的核酸序列；

通过生物信息学的分析方法，确定与骨质疏松症患者相关的特异性肠道微生物核酸序列。首先，将测序序列（reads）与参照基因集（也称为参考基因集，可以为新构建的基因集或任何已知序列的数据库，例如，采用已知的人肠道微生物群落非冗余基因集）进行比对。接下来，基于比对结果，分别确定来自健康人群和骨质疏松症患者群粪便样品的核酸样本中各基因的相对丰度。通过将测序序列与参照基因集进行比对，可以将测序序列与参照基因集中的基因建立对应关系，从而针对核酸样本中的特定基因，与其相对应的测序序列的数目可以有效地反映该基因的相对丰度。由此，可以通过比对结果，按照常规的统计分析，确定在核酸样本中基因的相对丰度。最后，在确定核酸样本中各基因的相对丰度后，对来自健康人群和骨质疏松症患者群粪便的核酸样本中各基因的相对丰度进行统计检验，由此，可以判断在健康人群和骨质疏松症患者人群中是否存在相对丰度有显著差异的基因，如果存在基因是显著差异的，则该基因被当作是异常状态的生物标志物，即核酸标志物。

另外，对于已知或新构建的参照基因集，其通常包含基因物种信息和功能注释，由此，在确定基因相对丰度的基础上，可以进一步通过将基因的物种信息和功能注释进行分类，从而确定肠道菌群中各微生物的物种相对丰度和功能相对丰度，也就可以进一步确定异常状态的物种标志物和功能标志物。简言之，确定物种标志物和功能标志物的方法进一步包括：将健康人群和骨质疏松症患者群的测序序列与参照基因集进行比对；基于比对结果，分别确定健康人群和的骨质疏松症患者群的核酸样本中各基因的物种相对丰度和功能相

对丰度；对来自健康人群和骨质疏松症病人群的核酸样本中各基因的物种相对丰度和功能相对丰度进行统计学检验；以及分别确定在健康人群和骨质疏松症患者群的核酸样本之间相对丰度存在显著差异的物种标志物和功能标志物。根据本发明的实施例，可以采用对来自相同物种的基因的相对丰度和具有相同功能注释的基因的相对丰度进行统计检验，例如加和、取平均值、中位数值等，来确定功能相对丰度和物种相对丰度。

最后，确定了在健康人群和骨质疏松症患者群的粪便样品之间相对丰度存在显著差异的生物学标志物，即包括微生物物种：多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物，单形拟杆菌 (*Bacteroides uniformis*) 或其类似物，*Bacteroides intestinalis* 或其类似物，*Bacteroides dorei* 或其类似物，瘤胃球菌 (*Ruminococcus sp.*) 或其类似物，嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物，*Parabacteroides merdae* 或其类似物，扭链瘤胃球菌 (*Ruminococcus torques*) 或其类似物，*Dialister invisus* 或其类似物。由此，通过检测上述微生物至少一种是否存在，来有效地确定对象是否患有或者易感骨质疏松症病，并且可以用于监控骨质疏松症病患者的治疗效果。在本文中所使用的术语“存在”应做广义理解，既可以指的是定性分析样本中是否含有相应的目标物，也可以指对样本中的目标物进行定量分析，并且还可以进一步将所得到的定量分析结果与参照（例如通过对具有已知状态的样本进行平行试验所得到的定量分析结果）进行统计学分析或者任何已知数学运算所得到的结果。本领域技术人员可以根据需要和试验条件进行容易的选择。根据本发明的实施例，还可以通过确定这些微生物在肠道菌群中的相对丰度，从而能够确定对象是否患有或者易感骨质疏松症病，以及用于监控骨质疏松症病患者的治疗效果。

可以通过检测对象肠道菌群中是否存在上述微生物物种中的至少一种，也可以是检测对象肠道菌群中是否存在上述中的两种或者多种，即是否存在上述生物标志物组合，从而来有效地确定对象是否患有或者易感骨质疏松症，并且可以用于监控骨质疏松症病患者的治疗效果。在本文中，术语“生物标志物组合”是指两个或者更多个生物标记物组成的组合。

对于物种标志物和功能标志物本领域技术人员还可以通过常规的菌种鉴别手段和生物活性检验手段来确定在肠道菌群中是否存在所述物种和功能。例如，菌种鉴别可以通过进行 16s rRNA 进行。

根据本发明的一种具体实施方式，本发明提出了试剂在制备试剂盒中的用途，所述试剂盒用来诊断是否患有骨质疏松症或相关疾病或者预测是否患有骨质疏松症或相关疾病的风险，所述试剂用来检测本发明的生物标志物。

30 检测对象是否患有骨质疏松症或相关疾病或者预测对象是否患有骨质疏松症或相关疾病的设备

根据本发明的又一方面，本发明提出了一种检测对象中是否患有骨质疏松症或相关疾

病或者预测对象是否患有骨质疏松症或相关疾病的设备，如图 5 所示。根据本发明的实施例，所述设备包括样本采集装置 100、生物标志物相对丰度确定装置 200 以及患病概率确定装置 300（如图 5 中 a 所示）。其中，样本采集装置适于从所述对象中采集样本；生物标志物相对丰度确定装置与所述样本采集装置相连，其适于确定所获得的样本中的生物标志物的相对丰度信息，所述生物标志物为根据本发明的第一方面的生物标志物；所述患病概率确定装置与所述生物标志物相对丰度确定装置相连，所述患病概率确定装置用于将相对丰度确定装置中获得的生物标志物的相对丰度信息与参考数据集或参考值进行比对。

根据本发明的一种具体实施方式，所述参考数据集包括来自多个骨质疏松症患者和多个健康对照的样本中的根据本发明的第一方面的生物标志物的相对丰度信息。

10 根据本发明的一种具体实施方式，所述患病概率确定装置中还包括执行多元统计模型以获得患病概率；优选地，所述多元统计模型为随机森林模型。根据本发明的一种优选实施方式，所述患病概率大于阈值表明所述对象患有骨质疏松症或相关疾病或者有患有骨质疏松症或相关疾病的风险；优选地，所述阈值为 0.5。根据本发明的一种优选实施方式，当与参考值进行比较时，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物、
15 *Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的减少表明所述对象患有骨质疏松症或相关疾病或者处于患有骨质疏松症或相关疾病的风险中，所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物、单形拟杆菌 (*Bacteroides uniformis*) 或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌 (*Ruminococcus sp.*) 或其类似物、扭链瘤胃球菌 L2-14 (*Ruminococcus torques* L2-14) 或
20 其类似物的增加表明所述对象患有骨质疏松症或相关疾病或者处于患有骨质疏松症或相关疾病的风险中。

根据本发明的一种具体实施方式，所述生物标志物相对丰度确定装置进一步包括：核酸样本分离单元 210、测序单元 220 以及比对单元 230（如图 5 中 b 所示）。根据本发明的实施例，核酸样本分离单元适于从所述对象的所述样本中分离得到核酸样本，测序单元与核酸样本分离单元相连，并且基于所获得的核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果，比对单元与测序单元相连，并且基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物的相对丰度信息。

根据本发明的一种具体实施方式，所述参考基因集包括从多个骨质疏松症患者和多个健康对照的样本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道
30 微生物基因集合并，得到所述参考基因集。

根据本发明的实施例，测序单元并不受特别限制。优选地，所述测序单元利用第二代测序方法或第三代测序方法进行。优选地，所述测序单元为选自 HiSeq2000、SOLiD、454、

和单分子测序装置的至少一种。由此，能够利用这些测序装置的高通量、深度测序的特点，从而有利于对后续测序数据进行分析，尤其是进行统计学检验时的精确性和准确度。

根据本发明的一个实施例，所述比对单元利用选自 SOAP2 和 MAQ 的至少一种进行所述比对。由此，可以提高比对的效率，进而可以提高检测骨质疏松症的效率。

5 另外，根据本发明的实施例，本发明还提出了一种药物筛选方法。由此，根据本发明实施例，骨质疏松症密切相关的标志物作为药物设计靶点来进行药物的筛选，促进新的治疗骨质疏松症病的药物的发现。例如，可以通过检测与候选药物接触前后，生物标志物水平的变化，来确定候选药物是否可以作为治疗或预防骨质疏松症病的药物。例如，检测有害生物标志物水平在接触药物候选物之后是否有所降低，有益生物标志物水平在接触药物
10 候选物之后是否有所升高。另外，还可以通过确定药物对生物标志物的至少一种的生物活性的直接影响或间接影响来对候选化合物是否可以作为治疗或预防骨质疏松症的药物来进行筛选。由此，根据本发明的实施例，本发明还提出了根据骨质疏松症的生物标志物在筛选治疗或预防骨质疏松症的药物中的用途。

应理解，在本发明范围内中，本发明的上述各技术特征和在下文（如实施例）中具体
15 描述的各技术特征之间都可以互相组合，从而构成新的或优选的技术方案。限于篇幅，在此不再一一累述。

下面参考具体实施例，对本发明进行说明，需要说明的是，这些实施例仅仅是说明性的，而不能理解为对本发明的限制。

若未特别指明，实施例中所采用的技术手段为本领域技术人员所熟知的常规手段，所
20 采用的试剂和产品也均为可商业获得的。未详细描述的各种过程和方法是本领域中公知的常规方法，所用试剂的来源、商品名以及有必要列出其组成成分者，均在首次出现时标明，其后所用相同试剂如无特殊说明，均以首次标明的内容相同。

本发明采用宏基因组关联分析（MWAS）的分析方法，经测序分析粪便样本的菌群
25 组成，功能差异；用随机森林判别模型判别骨质疏松群体和非骨质疏松群体，获得患病概率，用于骨质疏松的患病风险评估、诊断、早期诊断或，寻找潜在药物靶点。

根据本发明，术语“MLG”是指操作分类单元(Metagenomic Linkage Group) (Qin J, Li Y, Cai Z, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes[J]. Nature, 2012, 490(7418): 55-60.)，是在系统发生学研究或群体遗传学研究中，
30 为了便于进行分析，人为给某一个分类单元（品系，种，属，分组等）设置的同一标志。通常按照相似性阈值将序列划分为不同的 MLG，每一个 MLG 通常被视为一个微生物物种。若一个 MLG 中有超过 50% 的序列以 95% 的碱基相似性比对上已知微生物物种，

则认为 MLG 为此已知物种；若一个 MLG 中有超过 50% 的序列以 85% 的碱基相似性比
对上已知微生物属水平，则认为 MLG 为此已知物种属水平注释。

根据本发明，术语“个体”指动物，特别是哺乳动物，如灵长类动物，最好是人。

5 根据本发明，术语如“一”、“一个”和“这”不仅指单数的个体，而是包括可以用
来说明特定实施方式的通常的一类。

在本发明中，所述的测序（二代测序）和 MWAS 具有本领域公知，本领域技术人员
可以根据具体情况进行调整。根据本发明的实施例，可以根据文献（Wang, Jun, and
Huijue Jia. "Metagenome-wide association studies: fine-mining the microbiome." *Nature
Reviews Microbiology* 14.8 (2016): 508-522.）中记载的方法进行。

10 在本发明中，随机森林模型和 ROC 曲线的使用方法为本领域所公知，本领域技术
人员可以根据具体情况进行参数设置和调整。根据本发明的实施例，可以根据文献
（Drogan D, Dunn WB, Lin W, Buijsse B, Schulze MB, Langenberg C, Brown M, Floegel
a., Dietrich S, Rolandsson O, Wedge DC, Goodacre R, Forouhi NG, Sharp SJ, Spranger J,
Wareham NJ, Boeing H: Untargeted Metabolic Profiling Identifies Altered Serum
15 Metabolites of Type 2-Diabetes Mellitus in a Prospective, Nested Case Control Study. *Clin
Chem* 2015, 61:487-497.; Mihalik SJ, Michaliszyn SF, de las Heras J, Bacha F, Lee S, Chace
DH, DeJesus VR, Vockley J, Arslanian SA: Metabolomic profiling of fatty acid and amino
acid metabolism in youth with obesity and type 2 diabetes: evidence for enhanced
mitochondrial oxidation. *Diabetes Care* 2012, 35:605-611., 通过引用全文并入此处），中记
20 载的方法进行。

在本发明中，构建了骨质疏松受试者和非骨质疏松受试者的生物标志物的训练集，
并以此为基准，对待测样本的生物标志物含量值进行评估。

在本发明中，非骨质疏松受试者为精神状态良好的受试者。

在本发明中，所述受试者可以为人或者模型动物。

25 本领域技术人员知晓，当进一步扩大样本量时，利用本领域公知的样本检测和计算
方法，可以得出每种生物标志物在样本中的正常含量值区间（绝对数值）。可以将检测
得到的生物标志物含量的绝对值与正常含量值进行比较，任选地，还可以结合统计学方
法，以得出骨质疏松患病风险评价、诊断等。

不希望受任何理论的限制，发明人指出这些生物标志物是存在于人体中的肠道菌

群。通过本发明所述的方法对受试者肠道菌群进行关联分析，得到骨质疏松群体的所述生物标志物在菌群检测中表现出一定的含量范围值。

实施例 1

1.1 样本收集：

5 参照文献 A metagenome-wide association study of gut microbiota in type 2 diabetes (Qin, J. et al. . Nature 490, 55–60 (2012)) 记载的方法，采集粪便样品后冷冻运输并迅速转移到-80℃保存，进行 DNA 提取，得到提取的 DNA 样本。

1.2 宏基因组测序与筛选

10 利用所提取的 DNA 样本构建测序文库，在 Illumina HiSeq2000 测序平台上进行单向 (Single-end) 宏基因组测序 (读长 100bp)。对测序产生的数据进行过滤 (quality-controlled, 去除 adapter 污染序列、去低质量序列和去宿主基因组污染序列)。

1.3 基因集比对

将上述“1.2 宏基因组测序与筛选”后的高质量测序片段 (reads) 与肠道参考基因集进行比对，从而得到基因的相对丰度 (参考文献 Qin, J. et al.2012)。

15 1.4 物种分类注释与丰度计算

通过与 IMG (v400) 数据库进行比对，参考文献 A metagenome-wide association study of gut microbiota in type 2 diabetes (Qin, J. et al. Nature 490, 55–60 (2012)) 记载的方法，对对预测的基因进行物种分类。对于门水平的物种分类，比对的相似度在 65%以上，比对覆盖度在 70%以上作为门水平的物种分类的临界值。对于属水平的物种分类，比对的相似度在 85%以上作为属水平的分类的临界值。对于比对的相似性在 95%以上作为种、株水平的物种分类的临界值。

25 然后参照文献 A metagenome-wide association study of gut microbiota in type 2 diabetes (Qin, J. et al. Nature 490, 55–60 (2012)) 中记载的方法，利用基因的相对丰度计算该物种的相对丰度，并用秩和检验 (Wilcoxon rank-sum test) 进行统计检验 ($p < 0.05$)，确定病例与对照之间的相对丰度存在显著差异的物种。

1.5 生物标志物丰度计算

30 根据基因丰度对基因进行聚类 (参考文献 Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 490, 55–60 (2012).)，选取聚类基因数大于 50 的 MLGs 进行物种注释；并取对应基因丰度中位数的办法，得到对应 MLGs 的丰度，并计算病例与对照之间的相对丰度存在显著差异的 MLGs。

1.6 利用随机森林 (ROC/AUC) 筛选骨质疏松发生发展的潜在生物标志物

为进一步筛选潜在疾病肠道生物标志物，本实施例构建了骨质疏松症受试者和非骨质疏松症受试者的生物标志物的训练集，并以此为基准，对待测样本的生物标志物含量值进行评估。其中，在本发明中，所述训练集和所述验证集具有本领域公知的含义。在本发明的实施方案中，训练集是指包含一定样本数的骨质疏松症受试者和非骨质疏松症受试者待测样本中的各生物标志物的含量的数据集合。验证集是用来测试训练集性能的独立数据集合。其中，非骨质疏松症受试者为精神状态良好的受试者，受试者可以为人或者模型动物，在本实施例中是以人为受试者进行实验的。

具体包括如下步骤：

10 本发明从 91 个样品（健康人：47 和骨质疏松病人：44 人）中，随机地从 44 个患病样本中选取 37 个样品，随机地从 47 个正常样本中抽取 40 正常样本，然后组成 77 个样本（37 个骨质疏松和 40 个正常人）作为训练集（表 1），其余样品作为验证集（包括 7 个骨质疏松病人和 7 个正常人）。

1.6.1 利用训练集数据验证筛选得到的生物标志物

15 首先，按照 1.4-1.5 描述的方法计算训练集中每个样本的相对丰度。然后将训练集基因数量大于 50 的 MLG 输入随机森林（randomForest 4.6-12 in R 3.2.5, RF）分类器。对分类器进行 5 次 10 折交叉验证，10 次重复，利用 RF 模型筛选的 MLG 相对丰度对每一个体计算其骨质疏松患病风险（图 2），并绘制受试者操作特征（receiver operation characteristic, ROC）曲线，并计算出曲线下面积（AUC）作为判别模型效能评价参数。

20 选取标志物组合数 < 30，且判别效能最佳的组合为本发明组合。在模型中输出每个 MLG 的选择频率，频率越高，代表该标志物用来判别骨质疏松和非骨质疏松的重要性越高。

结果显示，本发明所得 RF 分类器包含了 9 个代谢物（即 9 个生物标志物），这 9 个生物标志物对应的相对丰度如表 1 所示，其详细信息如表 2 所示。图 2 示出了随机森林分类器中 5 次 10 折交叉验证的错误率分布情况。图 2 中黑色粗曲线代表 5 次试验，

25 10 次重复的平均值，竖线代表所选最佳组合中 MLG 数目。图 3 示出了基于随机森林模型（9 个生物标志物）判断骨质疏松症患者和健康对照，训练集的操作曲线（ROC）和曲线下面积（AUC），其中特异性表征的是对不患病判对的概率，敏感性指的是对于患病判对的概率，其中对训练集样本的判别效能为：AUC=95.92%，95% 置信区间 CI=86.4%-100%（图 3），结果表明该模型所得代谢物组合可作为区分骨质疏松与非骨质疏松的潜在生物标志物。

30

其中，表 3 中，每种标志物基因集大小代表的是每种标志物中包括的核酸序列的个数；标志物基因集注释数代表的是：代表其中有多少基因注释到这个标志物上，标志物最优注释表征的是根据每种标志物包括的所有基因集与 IMG(v400)数据库进行比对，得到的相应的物种分类；最优注释基因比例表征的是：这个基因簇里面有多少比例的基因注释到那个物种；最优注释相似度表征的是：这些基因簇里注释到这个物种，所有基因的注释准确度的均值作为该标志物的最优注释相似度；富集方向代表的是，每种生物标志物在骨质疏松症患者和健康对照中的相对丰度的变化，其中 $OS \leq H$ 代表的是该生物标志物在骨质疏松症患者中的相对丰度小于在健康对照中的相对丰度， $H \leq OS$ 代表的是该生物标志物在骨质疏松症患者中的相对丰度大于在健康对照中的相对丰度；筛选频率代表的是：进行 5 折 10 次交叉验证，该生物标志物被选择的频率；验证集 AUC 代表的是：代表在训练集数据得到模型下，对验证集数据的判别程度；95%置信区间 (95%CI) 在 a 到 b 之间，代表的是对于给出的每种生物标志物，有相应的 95%的概率可以说样本介于给出的 a 到 b 之间，发生错误的概率为 5%。

从表 3 可以看出，富集方向一栏中相比较于健康对照，骨质疏松患者在嗜粘蛋白-艾克曼菌 ATCC BAA-835 (*Akkermansia muciniphila* ATCC BAA-835)、*Parabacteroides merdae* ATCC 43184、*Dialister invisus* DSM 15470 均表现为相对丰度减少，在多形拟杆菌(*Bacteroides thetaiotaomicron*)、单形拟杆菌 (*Bacteroides uniformis*)、*Bacteroides intestinalis* DSM 17393、*Bacteroides dorei*、瘤胃球菌 sp. 5_1_39BFAA (*Ruminococcus* sp. sp. 5_1_39BFAA)、扭链瘤胃球菌 L2-14 (*Ruminococcus torques* L2-14) 均表现为相对丰度增加。

1.6.2 利用验证集数据验证筛选得到的生物标志物

本发明，随即使用独立人群对该模型进行验证，患病概率 (RP) ≥ 0.5 预测个体具有骨质疏松患疾病风险或者患有骨质疏松。首先，按照 1.4-1.5 描述的方法计算验证集中每个样本中各生物标志物的相对丰度。然后按照 1.6.1 的方法利用随机森林模型对验证集数据进行验证。

基于该模型：

图 4 示出了基于随机森林模型 (9 个生物标志物) 判断骨质疏松症患者和健康对照，验证集接收者操作曲线 (ROC) 和曲线下面积 (AUC)。其中，基于 9 个标记物，对独立验证集 (骨质疏松=7 和健人对照=7)，模型的判别 AUC=94.33% (95% CI= 87.13%-99.4%) (图 4, 表 3、表 4-1、表 4-2、表 5)；基于 *Ruminococcus* sp. 5_1_39BFAA，验证集曲线下面积为 0.9082，特异性高。

在 3.2.5 版本 R 中使用 “randomForest 4.6-12 package” 进行随机森林模型分类和回归。输入包括训练集数据 (即训练样本中选定的 MLGs 标记物的相对丰度，见表 1)，样本疾病

状态(训练样本的样本疾病状态为矢量, '1'代表骨质疏松, '0'代表健康人), 以及一个验证集(验证集中所选 MLGs 标记物的相对丰度, 见表 4-1、4-2)。然后, 发明人利用 R 软件中随机森林包的随机森林函数建立分类和预测函数对验证集数据进行预测, 输出即为预测结果(患病概率); 阈值为 0.5, 如果疾病的概率 ≥ 0.5 , 则认为有患骨质疏松的风险或者患有骨质

5 疏松。

表 1 骨质疏松随机森林模型训练集肠道标志物(MLG)相对丰度数据

样本编号*	标记物编号													
	2273	2443	1902	2821	1032	367	536	1083	2787					
OS_9	6.4114E-06	2.826E-05	7.8341E-06	6.2069E-07	3.9534E-10	0	0	1.39E-09	2.1135E-10					
OS_8	1.365E-06	6.5831E-06	2.6034E-07	6.34E-06	8.7459E-09	9.9109E-10	1.6231E-06	5.4052E-10	1.8994E-09					
OS_7	1.7102E-06	2.1527E-05	4.872E-07	1.6618E-07	0	2.2621E-09	3.3303E-09	1.0497E-07	1.4521E-10					
OS_6	4.4509E-06	8.5332E-06	4.8249E-06	9.321E-06	2.8551E-07	0	4.1628E-09	1.0606E-07	0					
OS_5	8.6392E-06	5.2788E-06	6.9601E-06	2.1014E-05	1.2586E-09	0	2.8154E-10	0	0					
OS_4	6.6533E-07	8.5207E-06	1.0498E-07	2.8559E-06	0	0	7.0174E-07	6.5029E-07	0					
OS_37	7.549E-07	2.9126E-06	6.546E-08	1.2422E-06	0	0	1.5629E-09	1.2882E-09	0					
OS_36	5.9239E-06	9.111E-06	2.5848E-06	2.1767E-05	2.3102E-09	1.8814E-07	1.7208E-09	1.2451E-06	0					
OS_35	7.422E-07	2.3602E-05	1.2746E-06	1.5804E-06	8.8639E-09	0	4.398E-06	2.8499E-07	0					
OS_34	2.9534E-06	4.7854E-06	1.1607E-06	1.0762E-06	4.4776E-09	0	9.9469E-07	4.8523E-07	9.5148E-10					
OS_33	6.873E-07	2.345E-06	3.0248E-09	2.0878E-05	2.293E-06	0	1.6025E-10	2.6255E-06	2.169E-09					
OS_32	8.0481E-07	1.6076E-06	9.5428E-08	1.6402E-06	1.0543E-09	2.2162E-07	5.8456E-07	4.911E-09	1.0304E-09					
OS_31	6.6501E-07	5.8257E-06	5.0014E-06	5.1938E-07	8.9227E-09	0	2.9633E-06	1.729E-07	7.5398E-10					
OS_30	3.14E-07	2.8877E-06	2.546E-07	7.0665E-06	4.8939E-10	0	4.2256E-07	3.2093E-06	0					
OS_3	5.0134E-06	7.641E-06	1.3113E-05	1.3336E-07	6.341E-10	0	5.109E-10	1.8365E-07	0					
OS_29	3.4801E-06	1.0801E-05	2.3791E-06	7.252E-07	0	0	1.1692E-10	3.3892E-08	0					

OS_28	5.769E-07	3.5972E-06	7.7342E-08	1.2083E-06	3.2814E-07	2.6182E-09	4.8455E-06	1.3874E-09	0
OS_27	9.4263E-07	1.5663E-06	1.3111E-06	1.2569E-07	1.771E-10	0	2.2065E-07	1.7713E-07	0
OS_26	5.0588E-06	5.2326E-06	5.31E-06	5.51E-06	2.3331E-08	2.8413E-07	3.9085E-07	2.0304E-08	0
OS_25	3.7558E-07	4.4464E-06	8.928E-09	1.9078E-06	5.5387E-10	0	1.8164E-07	4.3716E-10	0
OS_24	7.4494E-07	1.3764E-06	1.0114E-07	1.2154E-07	5.081E-11	3.5942E-09	4.204E-09	3.1585E-07	0
OS_23	9.657E-07	1.4572E-06	3.9892E-07	7.6155E-08	2.3297E-08	0	5.5172E-07	1.4651E-07	0
OS_22	3.0549E-06	2.0918E-05	1.857E-07	5.1454E-06	0	0	4.4149E-06	6.4619E-08	9.7107E-07
OS_21	6.6791E-06	8.6859E-06	5.5727E-08	4.7276E-05	0	1.4736E-09	3.1341E-06	0	0
OS_20	1.2564E-06	4.7339E-06	2.5558E-07	4.8694E-07	9.1261E-11	0	2.9148E-09	8.5395E-10	0
OS_2	1.5526E-06	7.7189E-06	1.9553E-07	2.0597E-05	7.1589E-10	0	2.5573E-10	2.5866E-07	0
OS_19	2.7736E-06	8.6596E-06	5.6086E-07	7.1038E-06	1.2368E-09	0	3.8744E-09	2.1466E-06	0
OS_18	5.6799E-07	1.9377E-06	3.2802E-07	4.9949E-08	0	0	8.7792E-07	2.5266E-08	0
OS_17	1.5112E-06	9.9785E-06	1.7881E-09	8.997E-09	3.4682E-08	0	4.8095E-10	1.1505E-07	0
OS_16	2.6926E-08	5.1929E-07	4.7277E-07	5.9245E-07	1.8532E-08	1.931E-09	4.5084E-07	1.9885E-07	0
OS_15	1.4853E-06	1.6E-05	2.2886E-06	2.6727E-05	5.4478E-11	0	8.8555E-07	1.7135E-07	0
OS_14	2.2175E-06	1.2755E-05	6.0933E-07	5.408E-06	4.5213E-09	0	4.1042E-06	1.3176E-07	0
OS_13	6.4659E-07	1.5125E-06	2.1457E-07	3.024E-06	3.7338E-08	0	7.3494E-07	4.8448E-08	0
OS_12	1.2771E-05	1.3001E-05	3.9329E-06	4.2596E-07	0	0	3.5975E-10	0	0
OS_11	6.2147E-07	8.4137E-06	2.9906E-06	9.5142E-08	0	0	2.2498E-09	2.6514E-08	0
OS_10	4.5503E-06	1.1268E-06	1.5206E-06	3.53E-06	4.0215E-10	0	1.1974E-09	2.4652E-07	1.4344E-10
OS_1	1.0454E-06	1.4173E-06	6.4056E-07	3.4115E-06	3.7714E-08	4.7886E-09	2.1615E-09	4.0262E-08	1.0247E-08

H_9	2.776E-06	1.527E-05	1.1056E-07	2.2049E-07	0	0	2.4321E-06	0	1.1041E-09
H_8	6.6724E-08	3.0384E-07	8.7673E-10	4.4818E-06	0	5.679E-09	1.4837E-07	0	1.7706E-09
H_7	3.3578E-09	1.492E-08	1.4717E-09	7.4077E-07	0	2.0344E-09	2.4197E-09	2.7818E-07	2.4791E-06
H_6	4.4292E-07	2.1674E-06	1.4818E-06	1.1436E-07	0	8.291E-06	1.6657E-06	1.2754E-06	2.5094E-10
H_5	1.6168E-06	5.291E-06	6.1732E-07	6.7321E-07	5.0091E-09	3.2357E-09	1.4466E-06	7.6494E-08	5.9687E-10
H_40	1.8093E-07	8.2651E-06	5.765E-06	1.9339E-07	9.2306E-09	1.3826E-09	4.3475E-07	3.1085E-07	2.4234E-08
H_4	5.4755E-07	1.455E-06	4.9288E-07	8.3495E-08	1.0375E-08	3.8409E-08	9.981E-07	3.2909E-07	0
H_39	7.5892E-08	3.0717E-07	7.2791E-07	2.5786E-08	2.1282E-07	1.0246E-10	4.4362E-08	5.5415E-07	0
H_38	9.9491E-07	1.8403E-05	1.8841E-06	2.8627E-07	0	6.3766E-09	7.447E-09	0	0
H_37	1.5807E-07	5.0728E-06	6.2761E-08	2.1402E-06	2.8917E-09	9.8484E-09	1.1424E-06	1.0633E-06	3.9983E-09
H_36	0.00010165	2.5399E-08	1.7321E-09	3.6179E-08	0	0	1.2564E-08	0	0
H_35	4.0939E-07	1.1618E-06	6.012E-08	4.3613E-08	0	3.3518E-06	1.216E-06	4.6388E-10	3.3913E-09
H_34	4.0014E-09	3.4145E-08	4.5091E-09	4.1236E-09	0	0	4.6116E-09	1.5399E-09	2.5777E-08
H_33	4.4387E-07	8.0784E-06	9.9577E-08	1.7138E-06	0	1.7122E-09	2.3562E-07	9.7784E-08	2.3898E-09
H_32	2.4365E-06	2.2162E-05	1.2207E-06	1.3183E-07	0	0	9.4131E-09	2.8631E-07	0
H_31	3.7663E-07	1.736E-06	2.5521E-07	8.1244E-08	5.4664E-08	3.2293E-09	3.7305E-07	7.3253E-08	0
H_30	7.3318E-06	5.6717E-06	1.5975E-09	6.6515E-06	2.1668E-10	8.0359E-09	4.2037E-09	0	0
H_3	5.0949E-06	1.4768E-07	1.1926E-08	5.8092E-07	2.5721E-06	0	0	0	0
H_29	9.8238E-09	1.5265E-08	6.6553E-10	6.901E-07	2.1288E-10	2.3693E-07	2.5922E-06	0	3.562E-08
H_28	1.8977E-07	3.1723E-06	7.6098E-08	1.5601E-07	0	3.2716E-09	3.3887E-07	2.7214E-07	6.4672E-09
H_27	2.858E-07	3.439E-06	1.4821E-06	1.4818E-06	7.7575E-08	0	3.3482E-07	3.786E-07	4.1597E-10

H_26	4.9375E-07	2.7958E-07	8.8832E-09	9.9405E-08	0	7.3714E-09	4.5667E-09	1.4713E-07	1.25E-09
H_25	3.268E-06	1.1604E-05	2.0701E-07	4.6105E-06	0	7.544E-08	1.4073E-06	4.0015E-08	2.1123E-10
H_24	2.919E-06	3.4606E-08	2.5767E-06	2.3021E-07	0	5.001E-09	7.1201E-09	0	2.8801E-09
H_23	7.3326E-07	1.3829E-05	1.0597E-06	9.5982E-06	5.4375E-08	0	3.2465E-09	8.6251E-08	5.8254E-10
H_22	3.3213E-07	2.4934E-06	5.0431E-08	3.024E-06	6.8219E-10	0	0	5.0517E-10	0
H_21	2.4286E-07	7.2572E-06	4.1067E-08	1.2365E-05	5.0098E-10	0	1.1175E-06	3.6207E-09	0
H_20	2.1377E-06	5.5338E-06	2.5175E-09	9.6543E-06	0	0	4.6563E-09	1.8577E-10	2.0145E-10
H_2	1.3694E-06	7.3109E-06	7.761E-08	2.0676E-06	1.3173E-09	0	0	0	0
H_19	2.1248E-06	2.8271E-06	1.0032E-08	2.2935E-07	0	0	1.4364E-08	0	2.3255E-09
H_18	3.0455E-07	6.7712E-06	2.4364E-07	2.4956E-06	0	2.1925E-10	6.3552E-09	4.604E-08	0
H_17	1.4848E-06	6.1403E-06	6.066E-07	2.9309E-07	0	6.7396E-10	3.3622E-06	6.3384E-09	0
H_16	1.5981E-06	8.1303E-06	4.6208E-07	1.7019E-05	0	3.3152E-08	2.6984E-07	2.2773E-07	0
H_15	1.1106E-05	1.9227E-05	4.2867E-06	4.219E-07	0	7.2231E-07	9.8661E-07	1.8403E-07	0
H_14	2.4193E-07	7.9519E-06	9.2721E-07	1.6123E-07	0	8.7763E-10	7.7823E-07	6.5687E-08	0
H_13	2.728E-07	4.804E-07	3.7718E-09	4.159E-07	0	0	1.1908E-07	0	0
H_12	6.653E-07	1.3528E-05	3.7181E-06	7.5274E-07	2.7805E-10	0	1.2066E-06	0	1.1566E-09
H_11	1.7154E-06	2.1425E-05	3.7143E-06	1.425E-07	0	3.0277E-09	1.4234E-06	6.1299E-10	3.8999E-09
H_10	1.3344E-06	5.1196E-06	3.4075E-07	1.2329E-07	6.1017E-08	0	1.3407E-06	1.3011E-07	0
H_1	1.1782E-08	1.1909E-07	4.9934E-09	1.0575E-07	0	1.5513E-08	1.0027E-06	9.6061E-09	1.0699E-06

*样本编号 (OS: 骨质疏松患者; H: 健康对照)

表 2.9 种生物标记物结合预测训练集的患病概率 (患病概率 ≥ 0.5 确认个体具有患骨质疏松的风险或者患有骨质疏松)

样本编号 (OS: 骨质疏松患者; H: 健康对照)	患病概率		样本编号 (OS: 骨质疏松患者; H: 健康对照)	患病概率
H_1	0.052631579		OS_1	0.681319
H_10	0.578125		OS_10	0.897727
H_11	0.259887006		OS_11	0.579487
H_12	0.392857143		OS_12	0.373057
H_13	0.191489362		OS_13	0.828729
H_14	0.190217391		OS_14	0.708791
H_15	0.514792899		OS_15	0.553846
H_16	0.413978495		OS_16	0.356383
H_17	0.754491018		OS_17	0.544974
H_18	0.45212766		OS_18	0.634286
H_19	0.195121951		OS_19	0.716763
H_2	0.557894737		OS_2	0.857923
H_20	0.261904762		OS_20	0.511236
H_21	0.288888889		OS_21	0.279793
H_22	0.56122449		OS_22	0.570621
H_23	0.569767442		OS_23	0.862857

H_24	0.18125	OS_24	0.395722
H_25	0.503144654	OS_25	0.541667
H_26	0.244680851	OS_26	0.634921
H_27	0.461956522	OS_27	0.327778
H_28	0.180487805	OS_28	0.438889
H_29	0.103773585	OS_29	0.720207
H_3	0.546875	OS_3	0.785311
H_30	0.584699454	OS_30	0.645833
H_31	0.337278107	OS_31	0.539683
H_32	0.387878788	OS_32	0.206186
H_33	0.20212766	OS_33	0.55914
H_34	0.279069767	OS_34	0.482051
H_35	0.169398907	OS_35	0.903226
H_36	0.383783784	OS_36	0.770492
H_37	0.345744681	OS_37	0.672222
H_38	0.49197861	OS_4	0.673469
H_39	0.321428571	OS_5	0.601124
H_4	0.487309645	OS_6	0.433526
H_40	0.286585366	OS_7	0.394595
H_5	0.43814433	OS_8	0.687179
H_6	0.123529412	OS_9	0.647668

H_7	0.207253886		
H_8	0.123076923		
H_9	0.234375		

表 3.9 种生物标记物详细信息

标物编号	标物因素大小	标物因素注释	标志物最优注释 (生物标志物信息)	最优注释基 因比例	最优注释相似 度	富集方向 (OS: 骨质疏松患者; H: 健康对照)	筛选 频率*	验证 集 AUC (%)	95% CI(%)
2273	1099		<i>Bacteroides thetaiotaomicron</i>	0.984531392	98.65725508	H<OS	49	75.51	47.23-100
2443	523		<i>Bacteroides uniformis</i>	0.965583174	98.10289109	H < OS	40	51.02	13.5-88.54
1902	353		<i>Bacteroides intestinalis</i> DSM 17393	0.892351275	94.95568254	H < OS	30	69.39	32.58-93.95
2821	204		<i>Bacteroides dorei</i>	0.950980392	96.98175258	H < OS	21	65.31	33.73-96.89
1032	200		Ruminococcus sp. 5_1_39BFAA	0.56	85.74892857	H < OS	34	90.82	72.63-100
367	103		<i>Akkermansia</i>	0.747572816	88.07051948	OS < H	30	63.27	32.58-93.95

2787	0	3.20505E-09	1.2547E-07	0	5.0161E-10	1.1818E-09	0
------	---	-------------	------------	---	------------	------------	---

表 4-2. 随机森林模型验证集肠道标志物 (MLG) 相对丰度数据

标记物编号	样本编号 (OS: 骨质疏松患者; H: 健康对照)							
	H_47	H_43	OS_39	H_42	H_41	OS_43	OS_38	
2273	3.7556E-07	6.426E-07	2.5216E-06	4.6507E-06	1.5082E-06	1.1081E-06	1.0332E-06	
2443	1.477E-05	6.911E-08	3.6248E-06	1.6102E-05	1.489E-06	1.5954E-05	1.7421E-05	
1902	1.1852E-06	3.7621E-10	4.0354E-09	2.7004E-06	1.3818E-08	1.9641E-07	6.5955E-07	
2821	7.0599E-07	3.6146E-08	6.5889E-07	6.5494E-07	1.3119E-05	4.571E-07	3.9734E-07	
1032	0	0	9.4094E-11	0	0	5.6741E-07	4.1392E-08	
367	2.9032E-06	5.7886E-11	1.6944E-09	0	0	0	7.9515E-10	
536	2.3341E-06	4.8186E-07	3.2719E-09	7.4953E-07	8.4377E-07	6.5167E-10	2.6355E-10	
1083	0	9.1478E-09	9.4145E-10	2.5364E-10	1.1926E-07	2.7529E-07	2.7437E-08	
2787	1.7488E-08	0	1.7537E-10	4.8587E-09	0	5.5972E-09	0	

表 5. 随机森林模型(分别基于 9 个肠道标志物、*Ruminococcus sp.* 5_1_39BFAA)预测骨质疏松和健康对照的样品患有骨质疏松的风险或患有骨质疏松的概率(患病概率 \geq 0.5 确认个体具有患骨质疏松的风险或者患有骨质疏松。)

样本编号 (OS: 骨质疏松患者; H: 健康对照)	基于 9 个肠道标志物的患病概率	基于标志物 1032 (<i>Ruminococcus sp.</i> 5_1_39BFAA) 的患病概率
H_41	0.334	0.002
H_42	0.534	0.002
H_43	0.13	0.002
H_44	0.36	0.75
H_45	0.21	0.35
H_46	0.478	0.002
H_47	0.072	0.002
OS_38	0.638	0.94
OS_39	0.624	0.984
OS_40	0.548	0.002
OS_41	0.648	0.762
OS_42	0.472	0.956
OS_43	0.798	0.866
OS_44	0.634	0.756

以上结果表明, 本发明公开的生物标志物具有较高的准确度和特异性, 具有良好的开发为诊断方法的前景, 从而为骨质疏松的患病风险评估、诊断、早期诊断, 寻找潜在药物靶点提供依据。

在本发明中, 术语“第一”、“第二”仅用于描述目的, 而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此, 限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中, “多个”的含义是至少两个, 例如两个, 三个等, 除非另有明确具体的限定。

在本发明中, 除非另有明确的规定和限定, 术语“安装”、“相连”、“连接”、“固定”等术语应做广义理解, 例如, 可以是固定连接, 也可以是可拆卸连接, 或成一体; 可以是机械连接, 也可以是电连接或彼此可通讯; 可以是直接相连, 也可以通过中间媒介间接相连, 可以是两个元件内部的连通或两个元件的相互作用关系, 除非另有明确的限定。对于本领域的普通技术人员而言, 可以根据具体情况理解上述术语在本发明中的具体含义。

在本发明中，除非另有明确的规定和限定，第一特征在第二特征“上”或“下”可以是第一和第二特征直接接触，或第一和第二特征通过中间媒介间接接触。而且，第一特征在第二特征“之上”、“上方”和“上面”可以是第一特征在第二特征正上方或斜上方，或仅仅表示第一特征水平高度高于第二特征。第一特征在第二特征“之下”、“下方”和“下面”可以是第一特征在第二特征正下方或斜下方，或仅仅表示第一特征水平高度小于第二特征。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

权利要求书

1、一种骨质疏松生物标志物，其特征在于，包括选自下列中的至少一种：

5 多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物，单形拟杆菌 (*Bacteroides uniformis*) 或其类似物，*Bacteroides intestinalis* 或其类似物，*Bacteroides dorei* 或其类似物，瘤胃球菌 (*Ruminococcus sp.*) 或其类似物，嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物，*Parabacteroides merdae* 或其类似物，扭链瘤胃球菌 (*Ruminococcus torques*) 或其类似物，*Dialister invisus* 或其类似物，

10 所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 类似物与多形拟杆菌 (*Bacteroides thetaiotaomicron*) 的基因组序列相比，比对相似度在 85% 以上，

所述单形拟杆菌 (*Bacteroides uniformis*) 类似物与单形拟杆菌 (*Bacteroides uniformis*) 的基因组序列相比，比对相似度在 85% 以上，

所述 *Bacteroides intestinalis* 类似物与 *Bacteroides intestinalis* 的基因组序列相比，比对相似度在 85% 以上，

15 所述 *Bacteroides dorei* 类似物与 *Bacteroides dorei* 的基因组序列相比，比对相似度在 85% 以上，

所述瘤胃球菌 (*Ruminococcus sp.*) 类似物与瘤胃球菌 (*Ruminococcus sp.*) 的基因组序列相比，比对相似度在 85% 以上，

20 所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 类似物与嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 的基因组序列相比，比对相似度在 85% 以上，

所述 *Parabacteroides merdae* 类似物与 *Parabacteroides merdae* 的基因组序列相比，比对相似度在 85% 以上，

所述扭链瘤胃球菌 (*Ruminococcus torques*) 类似物与扭链瘤胃球菌 (*Ruminococcus torques*) 的基因组序列相比，比对相似度在 85% 以上，

25 所述 *Dialister invisus* 类似物与 *Dialister invisus* 的基因组序列相比，比对相似度在 85% 以上。

2、根据权利要求 1 所述的生物标志物，其特征在于，所述 *Bacteroides intestinalis* 为 *Bacteroides intestinalis* DSM 17393，所述瘤胃球菌 (*Ruminococcus sp.*) 为瘤胃球菌 5_1_39BFAA (*Ruminococcus sp.* 5_1_39BFAA)，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 为嗜粘蛋白-艾克曼菌 ATCC BAA-835 (*Akkermansia muciniphila* ATCC BAA-835)，所述 *Parabacteroides merdae* 为 *Parabacteroides merdae* ATCC 43184，所述扭链瘤胃球菌 (*Ruminococcus torques*) 为扭链瘤胃球菌 L2-14 (*Ruminococcus torques* L2-14)，

所述 *Dialister invisus* 为 *Dialister invisus* DSM 15470。

3、根据权利要求 1 所述的生物标志物，其特征在于，所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 类似物与多形拟杆菌 (*Bacteroides thetaiotaomicron*) 的基因组序列相比，比对相似度在 95% 以上，

5 所述单形拟杆菌 (*Bacteroides uniformis*) 类似物与单形拟杆菌 (*Bacteroides uniformis*) 的基因组序列相比，比对相似度在 95% 以上，

所述 *Bacteroides intestinalis* 类似物与 *Bacteroides intestinalis* 的基因组序列相比，比对相似度在 95% 以上，

10 所述 *Bacteroides dorei* 类似物与 *Bacteroides dorei* 的基因组序列相比，比对相似度在 95% 以上，

所述瘤胃球菌 (*Ruminococcus sp.*) 类似物与瘤胃球菌 (*Ruminococcus sp.*) 的基因组序列相比，比对相似度在 95% 以上，

所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 类似物与嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 的基因组序列相比，比对相似度在 95% 以上，

15 所述 *Parabacteroides merdae* 类似物与 *Parabacteroides merdae* 的基因组序列相比，比对相似度在 95% 以上，

所述扭链瘤胃球菌 (*Ruminococcus torques*) 类似物与扭链瘤胃球菌 (*Ruminococcus torques*) 的基因组序列相比，比对相似度在 95% 以上，

20 所述 *Dialister invisus* 类似物与 *Dialister invisus* 的基因组序列相比，比对相似度在 95% 以上。

4、一种诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险的方法，其特征在于，包括：

(1) 从所述对象中采集样本；

25 (2) 确定步骤 (1) 中获得的所述样本中根据权利要求 1~3 任一项所述的生物标志物的相对丰度信息；

(3) 将步骤 (2) 中所述的相对丰度信息与参考数据集或参考值进行比较；

优选地，所述参考数据集包括来自多个骨质疏松患者和多个健康对照的样本中的根据权利要求 1~3 中任一项所述的生物标志物的相对丰度信息。

30 5、根据权利要求 4 所述的方法，其特征在于，在将步骤 (2) 中所述的相对丰度信息与参考数据集进行比较的步骤中，还包括执行多元统计模型以获得患病概率。

6、根据权利要求 5 所述的方法，其特征在于，所述多元统计模型为随机森林模型。

7、根据权利要求 5 所述的方法，其特征在于，所述患病概率大于阈值表明所述对象患

有骨质疏松或相关疾病或者有患有骨质疏松或相关疾病的风险。

8、根据权利要求 7 所述的方法，其特征在于，所述阈值为 0.5。

9、根据权利要求 4 所述的方法，其特征在于，当与参考值比较时，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物、*Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的减少表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中；所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物、单形拟杆菌 (*Bacteroides uniformis*) 或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌 (*Ruminococcus sp.*) 或其类似物、扭链瘤胃球菌 (*Ruminococcus torques*) 或其类似物的增加表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中。

10、根据权利要求 4 所述的方法，其特征在于，步骤 (2) 中所述生物标志物的相对丰度信息是利用测序方法得到的，进一步包括：

从所述对象的所述样本中分离得到核酸样本，

15 基于所获得的所述核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果，

以及基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物的相对丰度信息，

20 所述参考基因集包括从多个骨质疏松患者和多个健康对照的样本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集。

11、根据权利要求 10 所述的方法，其特征在于，所述样本为粪便样本。

12、根据权利要求 10 所述的方法，其特征在于，所述测序方法是通过第二代测序方法或第三代测序方法进行的。

25 13、一种试剂盒，其特征在于，包括用于检测权利要求 1~3 任一项所述的生物标志物的试剂。

14、根据权利要求 13 所述的试剂盒，其特征在于，所述试剂盒包括以下中的至少一种：一组参考数据集或者参考值，所述参考数据集或者参考值用来作为每种生物标志物的相对丰度的参考。

30 15、根据权利要求 14 所述的试剂盒，其特征在于，所述试剂盒还包括第一计算机程序产品，所述第一计算机产品用来执行获得所述的参考数据集或者参考值。

16、根据权利要求 13 所述的试剂盒，其特征在于，所述试剂盒还包括第二计算机程序产品，所述第二计算机产品用来执行权利要求 4 所述的诊断对象是否患有骨质疏松或相关

疾病或者预测对象是否患有骨质疏松或相关疾病的风险的方法。

17、权利要求 1~3 任一项所述的生物标志物在制备试剂盒中的用途，所述试剂盒用于诊断对象是否患有骨质疏松或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险。

5 18、根据权利要求 17 所述的用途，其特征在于，所述诊断或预测包括如下步骤：

1) 从所述对象中采集样本；

2) 确定步骤 1) 中获得的所述样本中根据权利要求 1~3 中任一项所述的生物标志物的相对丰度信息；

3) 将步骤 2) 中所述的相对丰度信息与参考数据集或参考值进行比较。

10 19、根据权利要求 18 所述的用途，其特征在于，所述参考数据集包括来自多个骨质疏松患者和多个健康对照的样本中的生物标志物的相对丰度信息，所述生物标志物为根据权利要求 1~3 中任一项所述的生物标志物。

20、根据权利要求 17 所述的用途，其特征在于，在将步骤 2) 中所述的相对丰度信息与参考数据集进行比较的步骤中，还包括执行多元统计模型以获得患病概率。

15 21、根据权利要求 20 所述的用途，其特征在于，所述多元统计模型为随机森林模型。

22、根据权利要求 20 所述的用途，其特征在于，所述患病概率大于阈值表明所述对象患有骨质疏松或相关疾病或者有患有骨质疏松或相关疾病的风险。

23、根据权利要求 22 所述的用途，其特征在于，所述阈值为 0.5。

20 24、根据权利要求 18 所述的用途，其特征在于，当与参考值比较时，所述嗜粘蛋白-艾克曼菌 (*Akkermansia muciniphila*) 或其类似物、*Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的减少表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中；所述多形拟杆菌 (*Bacteroides thetaiotaomicron*) 或其类似物、单形拟杆菌 (*Bacteroides uniformis*) 或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌 (*Ruminococcus sp.*) 或其类似物、扭链瘤胃球菌
25 (*Ruminococcus torques*) 或其类似物的增加表明所述对象患有骨质疏松或相关疾病或者处于患有骨质疏松或相关疾病的风险中。

25、根据权利要求 18 所述的用途，其特征在于，通过测序方法得到步骤 2) 中所述生物标志物的相对丰度信息，进一步包括：

从所述对象的所述样本中分离得到核酸样本，

30 基于所获得的所述核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果，

以及基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物

的相对丰度信息。

26、根据权利要求 25 所述的用途，其特征在于，所述参考基因集包括从多个骨质疏松患者和多个健康对照的样本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集。

5 27、生物标志物作为靶点用于筛选治疗或者预防骨质疏松或相关疾病的药物的用途，其中所述生物标志物包括权利要求 1~3 任一项所述的生物标志物。

28、生物标志物在诊断对象是否患有骨质疏松症或相关疾病或者预测对象是否患有骨质疏松或相关疾病的风险中的用途，其中所述生物标志物包括权利要求 1~3 中任一项所述的生物标志物。

10 29、一种检测对象是否患有骨质疏松症或相关疾病或者预测对象是否患有骨质疏松症或相关疾病的设备，其特征在于，包括：

样本采集装置，所述样本采集装置适于从所述对象中采集样本；

生物标志物相对丰度确定装置，生物标志物相对丰度确定装置与所述样本采集装置相连，其适于确定所获得的样本中的生物标志物的相对丰度信息，所述生物标志物包括权利
15 要求 1~3 中任一项所述的生物标志物；

患病概率确定装置，所述患病概率确定装置与所述生物标志物相对丰度确定装置相连，所述患病概率确定装置用于将相对丰度确定装置中获得的生物标志物的相对丰度信息与参考数据集或参考值进行比对。

30、根据权利要求 29 所述的设备，其特征在于，所述参考数据集包括来自多个骨质疏
20 松患者和多个健康对照的样本中的根据权利要求 1~3 中任一项所述的生物标志物的相对丰度信息。

31、根据权利要求 29 所述的设备，其特征在于，所述患病概率确定装置中还包括执行多元统计模型以获得患病概率。

25 32、根据权利要求 29 所述的设备，其特征在于，所述生物标志物相对丰度确定装置进一步包括：

核酸样本分离单元，所述核酸样本分离单元适于从所述对象的所述样本中分离得到核酸样本；

测序单元，所述测序单元与所述核酸样本分离单元相连，并且基于所获得的核酸样本，构建 DNA 文库，对所述 DNA 文库进行测序，以便获得测序结果；

30 比对单元，所述比对单元与所述测序单元相连，并且基于所述测序结果，将测序结果与参考基因集进行比对，以确定所述生物标志物的相对丰度信息。

33、根据权利要求 32 所述的设备，其特征在于，参考基因集包括从多个骨质疏松患者

和多个健康对照的样本中进行宏基因组测序，获得非冗余基因集，然后将所述非冗余基因集与肠道微生物基因集合并，得到所述参考基因集。

- 34、一种药物，其特征在于，所述药物用于预防或治疗骨质疏松或相关疾病，所述药物能够使得检测对象中嗜粘蛋白-艾克曼菌（*Akkermansia muciniphila*）或其类似物、
5 *Parabacteroides merdae* 或其类似物、*Dialister invisus* 或其类似物的相对丰度值增加；或者所述药物能够使得多形拟杆菌（*Bacteroides thetaiotaomicron*）或其类似物、单形拟杆菌（*Bacteroides uniformis*）或其类似物、*Bacteroides intestinalis* 或其类似物、*Bacteroides dorei* 或其类似物、瘤胃球菌（*Ruminococcus sp.*）或其类似物、扭链瘤胃球菌（*Ruminococcus torques*）或其类似物的相对丰度值减少。

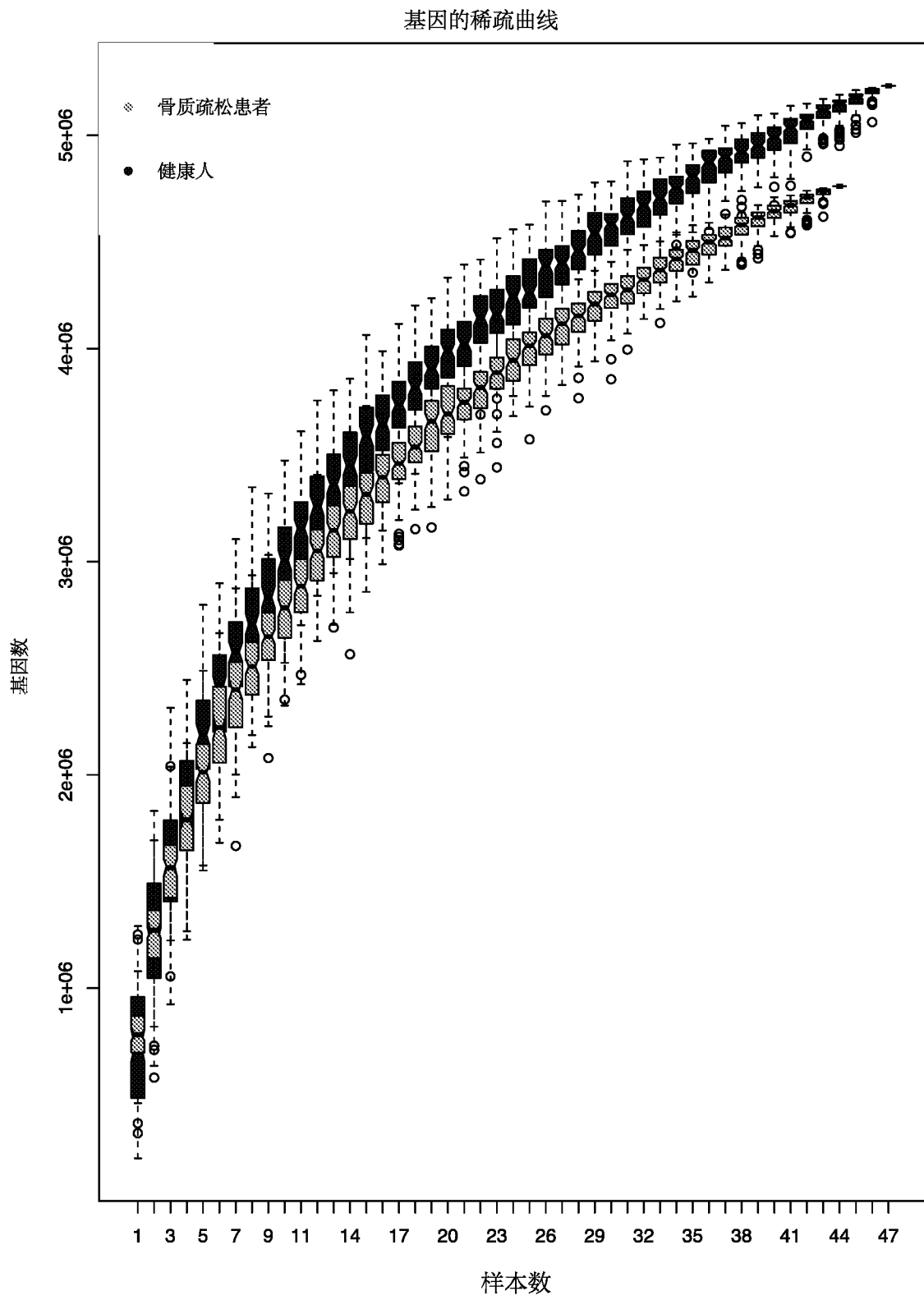


图 1

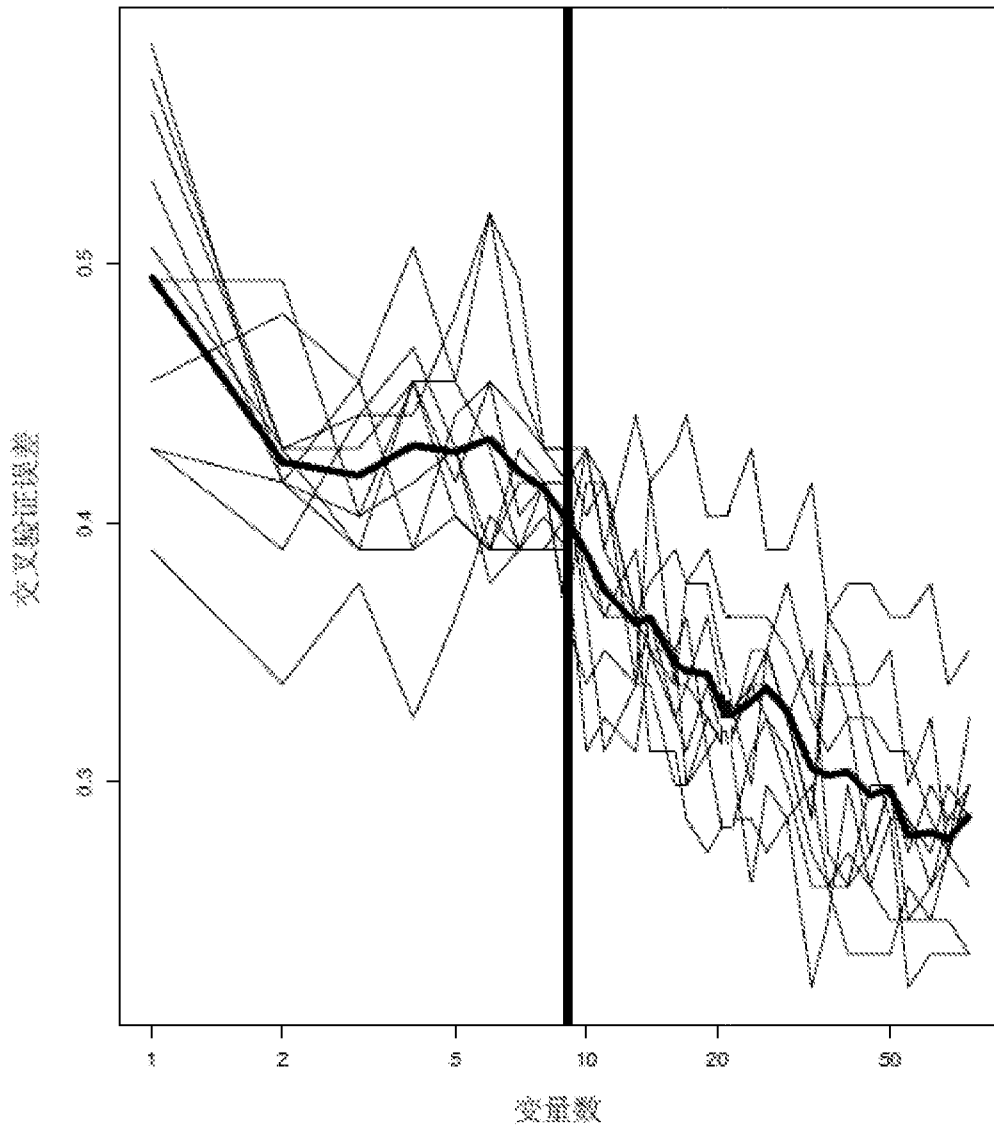


图 2

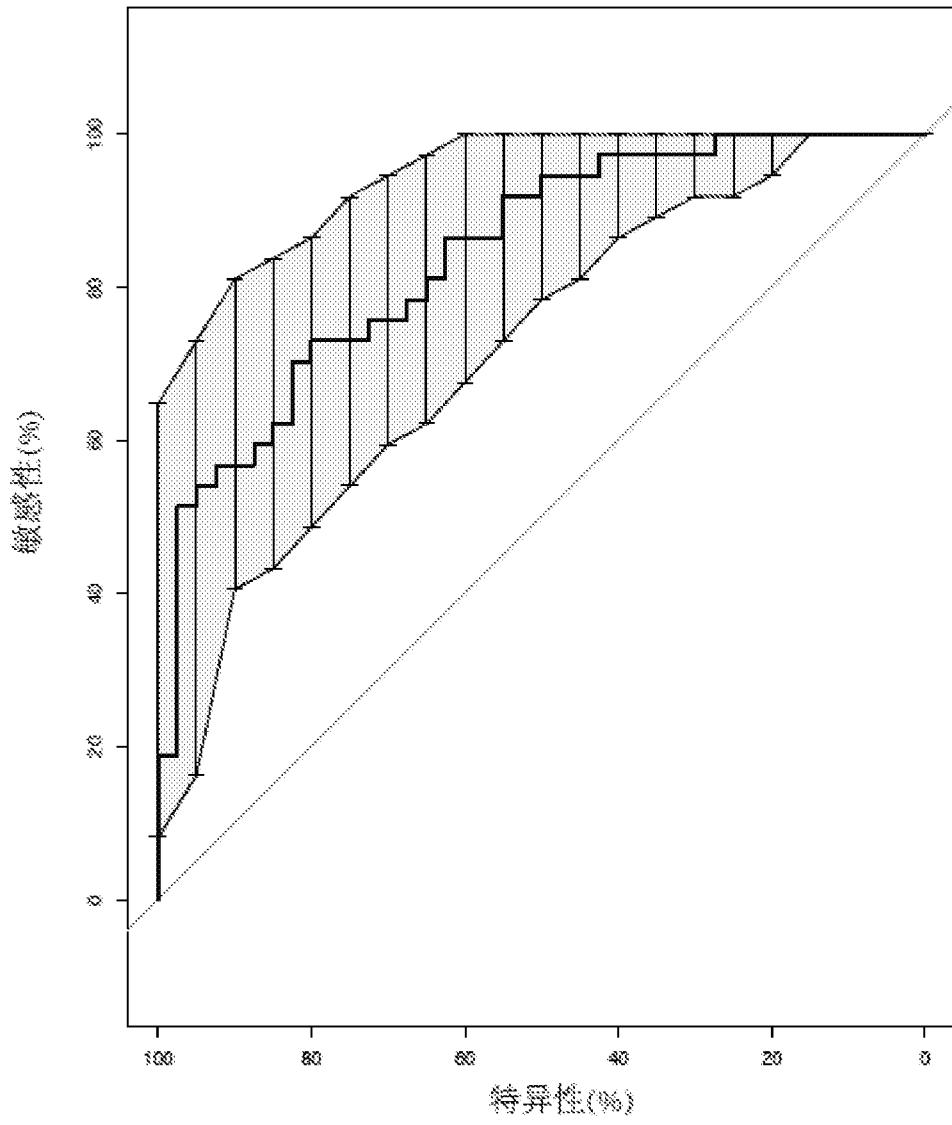


图 3

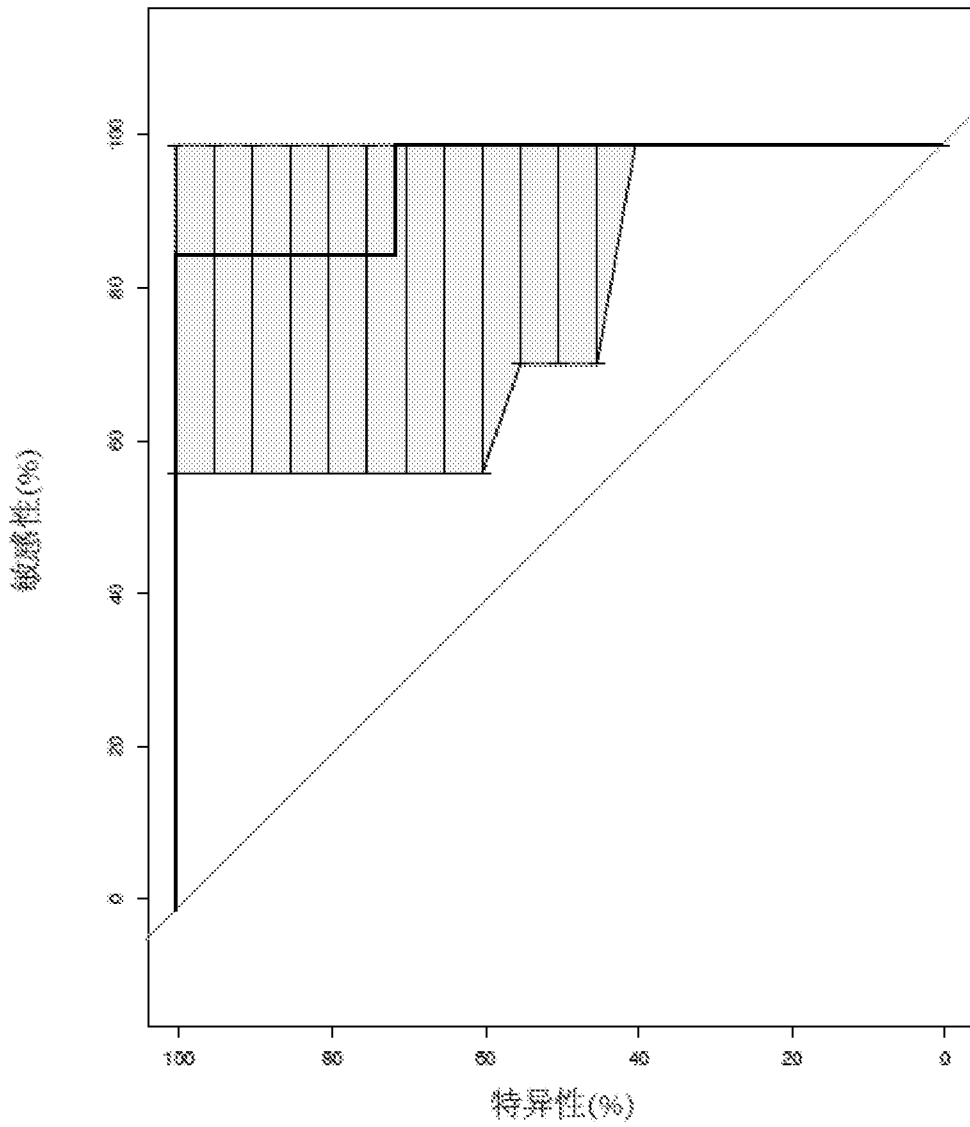


图 4

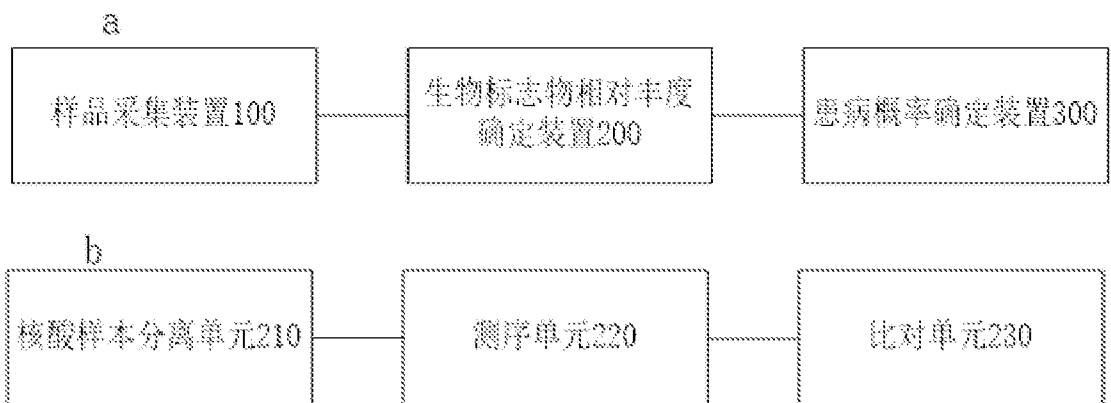


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/084276

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/68(2018.01)i; C12N 1/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q1/68 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; NCBI; JPABS; TWABS; CNTXT; WOTXT; USTXT; EPTXT; CJFD; CNKI; VEN; DWPI; SIPOABS; PUBMED; ELSEVIER; Patents: 多形拟杆菌, 骨质疏松, 肠道菌群, 丰度, 样本, Bacteroides thetaiotaomicron, osteoporosis, Gut microbiota, intestinal, abundance, sample

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 107075446 A (BGI SHENZHEN CO., LTD. ET AL.) 18 August 2017 (2017-08-18) description, paragraphs 13, 22-27, and 31, and table 1	1-3, 13-16, 29-34
A	黄永铨 等 (Huang, Yongquan et al.). "肠道菌群与骨质疏松关系的研究进展 (Gut Microbiota and Osteoporosis)" 南方医科大学学报 (<i>Journal of Southern Medical University</i>), Vol. 37, No. 2, 31 December 2017 (2017-12-31), pp. 278-280	1-34

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

06 August 2018

Date of mailing of the international search report

19 September 2018

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/084276

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

[1] See form PCT/ISA/206, 1(i)

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **Part of claims 1-34 relating to Bacteroides thetaiotaomicron or analogues thereof.**

- Remark on Protest**
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
 - The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
 - No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/084276

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 107075446 A	18 August 2017	WO 2016049932 A1	07 April 2016

国际检索报告

国际申请号

PCT/CN2018/084276

<p>A. 主题的分类 C12Q 1/68(2018.01) i; C12N 1/00(2006.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>											
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号) C12Q1/68 C12N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNABS;NCBI;JPABS;TWABS;CNTXT;WOTXT;USTXT;EPTXT;CJFD;CNKI;VEN;DWPI; SIPOABS;PUBMED;ELSEVIER;Patentics; 多形拟杆菌, 骨质疏松, 肠道菌群, 丰度, 样本, Bacteroides thetaiotaomicron, osteoporosis, Gut microbio- ta, intestinal, abundance, sample</p>											
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 107075446 A (深圳华大基因科技有限公司 等) 2017年 8月 18日 (2017 - 08 - 18) 说明书第13, 22-27, 31段, 表5-1</td> <td>1-3, 13-16, 29-34</td> </tr> <tr> <td>A</td> <td>黄永铨 等. "肠道菌群与骨质疏松关系的研究进展" 南方医科大学学报, 第37卷, 第2期, 2017年 12月 31日 (2017 - 12 - 31), 第278-280页</td> <td>1-34</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 107075446 A (深圳华大基因科技有限公司 等) 2017年 8月 18日 (2017 - 08 - 18) 说明书第13, 22-27, 31段, 表5-1	1-3, 13-16, 29-34	A	黄永铨 等. "肠道菌群与骨质疏松关系的研究进展" 南方医科大学学报, 第37卷, 第2期, 2017年 12月 31日 (2017 - 12 - 31), 第278-280页	1-34
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求									
X	CN 107075446 A (深圳华大基因科技有限公司 等) 2017年 8月 18日 (2017 - 08 - 18) 说明书第13, 22-27, 31段, 表5-1	1-3, 13-16, 29-34									
A	黄永铨 等. "肠道菌群与骨质疏松关系的研究进展" 南方医科大学学报, 第37卷, 第2期, 2017年 12月 31日 (2017 - 12 - 31), 第278-280页	1-34									
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>											
<p>* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件</p>											
国际检索实际完成的日期	国际检索报告邮寄日期										
2018年 8月 6日	2018年 9月 19日										
ISA/CN的名称和邮寄地址	受权官员										
中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088	陈典										
传真号 (86-10)62019451	电话号码 (86-512)88996499										

第III栏 缺乏发明单一性的意见(续第1页第3项)

本国际检索单位在该国际申请中发现多项发明，即：

[1] 参见PCT/ISA/206表中1(i)

1. 由于申请人按时缴纳了被要求缴纳的全部附加检索费，本国际检索报告涉及全部可作检索的权利要求。
2. 由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索，本单位未通知缴纳任何加费。
3. 由于申请人仅按时缴纳了部分被要求缴纳的附加检索费，本国际检索报告仅涉及已缴费的那些权利要求，具体地说，是权利要求：
4. 申请人未按时缴纳被要求缴纳的附加检索费。因此，本国际检索报告仅涉及权利要求书中首先提及的发明；包含该发明的权利要求是： 权利要求1-34中涉及多形拟杆菌(*Bacteroides thetaiotaomicron*)或其类似物的部分。

对异议的意见

- 申请人缴纳了附加检索费，同时提交了异议书，适用时，缴纳了异议费。
- 申请人缴纳了附加检索费，同时提交了异议书，但未在通知书规定的时间期限内缴纳异议费。
- 缴纳附加检索费时未提交异议书。

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/084276

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 107075446 A	2017年 8月 18日	WO 2016049932 A1	2016年 4月 7日