

(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
2021 年 9 月 23 日 (23.09.2021)



(10) 国际公布号
WO 2021/184413 A1

- (51) 国际专利分类号:
G16H 50/20 (2018.01) *G06N 3/00* (2006.01)
G16B 30/00 (2019.01)
- (21) 国际申请号: PCT/CN2020/081945
- (22) 国际申请日: 2020 年 3 月 28 日 (28.03.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010193593.3 2020年3月18日 (18.03.2020) CN
- (71) 申请人: 浙江大学(ZHEJIANG UNIVERSITY) [CN/CN]; 中国浙江省杭州市西湖区余杭塘路866号胡少华, Zhejiang 310058 (CN)。
- (72) 发明人: 胡少华(HU, Shaohua); 中国浙江省杭州市上城区城站路58号, Zhejiang 310009 (CN)。 张佩芬(ZHANG, Peifen); 中国浙江省杭州市上城区

城站路58号胡少华, Zhejiang 310009 (CN)。 来建波(LAI, Jianbo); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 蒋佳俊(JIANG, Jiajun); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 许毅(XU, Yi); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 奚彩曦(XI, Caixi); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 杜彦莉(DU, Yanli); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 吴玲玲(WU, Lingling); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 路静(LU, Jing); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。 牟婷婷(MOU, Tingting); 中国浙江省杭州市上城区城站路58号胡少华, Zhejiang 310009 (CN)。

(74) 代理人: 杭州求是专利事务所有限公司 (HANGZHOU QIUSHI PATENT OFFICE CO.,

(54) Title: GUT MICROBE-BASED BIOMARKERS FOR PREDICTING CURATIVE EFFECT ON BIPOLAR DISORDER, AND SCREENING AND APPLICATIONS THEREOF

(54) 发明名称: 双相情感障碍疗效预测的肠道微生物标志物及其筛选应用

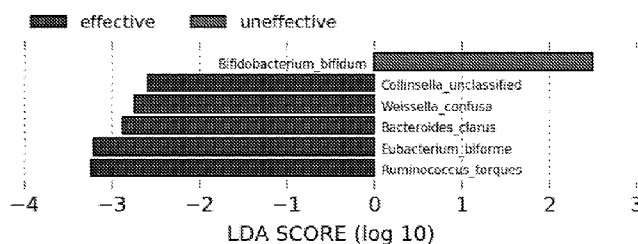


图 1

(57) Abstract: Provided are gut microbe-based biomarkers for predicting curative effect on bipolar disorder, and screening and applications thereof. Biomarker 1), i.e., Eubacterium_bifforme, biomarker 2), i.e., Ruminococcus_torques, and biomarker 3), i.e., Collinsella_unclassified, are comprised, and the relative abundance of a combination of the biomarkers in a treatment-effective group of patients with bipolar disorder is significantly increased. Also provided are an application of the biomarkers as test targets or test objects in preparation of a test kit, and an application of the biomarkers as targets in curative effect prediction.

(57) 摘要: 提供了双相情感障碍疗效预测的肠道微生物标志物及其筛选应用。其中, 生物标志物 1)Eubacterium_bifforme; 生物标志物 2)Ruminococcus_torques; 生物标志物 3)Collinsella_unclassified; 所述生物标志物组合在双相情感障碍患者治疗有效组中相对丰度均显著升高。还提供了根据所述生物标志物作为检测靶点或检测目标在制备检测试剂盒中的应用和作为靶点在疗效预测中的应用。

LTD); 中国浙江省杭州市玉古路 147 号, 黄鸿年科技综合大楼 215 室林松海, Zhejiang 310013 (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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, 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))。

双相情感障碍疗效预测的肠道微生物标志物及其筛选应用

技术领域

[0001] 本发明涉及双相情感障碍疗效预测的肠道微生物标志物及其筛选应用。

背景技术

[0002] 双相情感障碍是一类以抑郁、躁狂或者轻躁狂交替或混合发作为主要临床特征的慢性精神疾病，病因未明，发病年龄主要集中在成年早期或者青春期后期，全球发病率高，约为2~3%，与明显的职业、个人和社会功能等损害有关，并伴随着躯体疾病和早期死亡。目前针对双相情感障碍的治疗，主要是以药物治疗为基础，辅以物理治疗和心理治疗的综合治疗原则。即使经过治疗，残留的情绪症状往往仍然存在，大约37%的患者在1年内复发为抑郁症或躁狂症，60%的患者在2年内复发。多次反复发作，严重损害患者大脑功能，延长药物治疗时间，降低治疗有效性。

[0003] 尽管目前较为一致的观点认为双相情感障碍的发生是遗传因素和环境因素相互作用的结果，在神经、免疫以及内分泌等系统也表现出不同程度的变化，但针对双相情感障碍治疗疗效的早期预测，尚缺乏特异性的生物标志物。

发明内容

[0004] 为了克服现有技术的不足，本发明的目的是提供双相情感障碍疗效预测的肠道微生物标志物及其筛选应用。

[0005] 基于肠道微生物的双相情感障碍疗效预测生物标志物，包括以下6种中的多种：

"Bacteroides_clarus", "Eubacterium_biforme", "Weissella_confusa", "Ruminococcus_torques", "Bifidobacterium_dentium", "Collinsella_unclassified"在双相情感障碍患者治疗有效组与无效组中存在明显的差异。其中，"Bacteroides_clarus", "Eubacterium_biforme", "Weissella_confusa", "Ruminococcus_torques", "Collinsella_unclassified"在双相情感障碍患者治疗有效组相对丰度显著升高，而"Bifidobacterium_dentium"则下降。

[0006] 所述的生物标志物，为以下3种：

生物标志物1) Eubacterium_biforme; 重要性0.15259;

生物标志物2) Ruminococcus_torques; 重要性0.339395;

生物标志物 3) Collinsella_unclassified; 重要性 0.258467;

上述所述生物标志物组合在双相情感障碍患者治疗有效组中相对丰度均显著升高。

[0007] 所述的生物标志物是基于对其基因序列的计算所提供的。

[0008] 所述的生物标志物的相对丰度信息用于和参考值进行比较。

[0009] 一种根据所述的生物标志物作为检测靶点或检测目标在制备检测试剂盒中的应用。

[0010] 一种根据所述的生物标志物作为靶点在疗效预测中的应用。

[0011] 一种根据所述的生物标志物的筛选方法，步骤如下：

1) 样本收集：收集样本受试者包括双相情感障碍患者治疗前的粪便样品，在冰箱内-80℃条件下保存备用；

2) 对收集后储存的粪便样本进行核酸样本 DNA 的提取，宏基因组测序与组装、比对、筛选与质控之后，将高质量的测序片段输入到 Metaphlan2 软件，计算出物种的相对丰度；

3) 将上述所得双相情感障碍患者与健康对照物种的相对丰度信息输入到 LDA Effect Size (LEfSe) 系统，分析组间差异菌群。

[0012] 所述的筛选方法，进一步使用随机森林模型预测分析，步骤如下：

4.1) 双相情感障碍患者治疗有效组和治疗无效组为训练集，剩余样本作为测试集，计算训练集内每个样本中物种的相对丰度；

4.2) 将训练集中物种的相对丰度信息输入随机森林 (RF) 分类器中，并对分类器进行 5 次 10 折的交叉验证，对利用 RF 模型筛选出的每一个物种，依据其相对丰度信息计算双相情感障碍的患病风险、绘制 ROC 曲线，并计算其曲线下面积 (AUC)，将 AUC 作为判别模型效能评价的参数，在模型中输出每个物种的重要性指数，重要性指数越高，代表该标志物用来判别双相情感障碍和非双相情感障碍的重要性就越高。

[0013] 所述训练集中，样本受试者包括 27 个双相情感障碍治疗有效患者和 14 个治疗无效患者，测试集中，样本受试者包括 8 个双相情感障碍治疗有效患者和 5 个治疗无效患者。

[0014] 本发明的有益技术效果：

肠道微生物是存在于人体肠道中的微生物群落，通过“脑-肠”轴的神经、免疫以及内分泌等双向信息传递系统与大脑相互作用，参与影响人的情绪、行为和认知等功能，肠道菌群失调与多种神经精神疾病如阿尔茨海默病、孤独症、抑郁症以及双相情感障碍等的发生有关。肠道微生物的功能和组成受多种因素的影响，如饮食、运动、益生菌、抗生素以及粪菌移植等。

目前已有研究通过建造动物模型、调节肠道菌群变化的方法进一步探讨肠道菌群对大脑功能

状态的改善作用机制。双相情感障碍患者在使用第二代抗精神病药物如喹硫平稳定情绪、改善症状的同时，同样可使肠道菌群发生变化。

[0015] 因此，本发明通过对双相情感障碍患者治疗有效和治疗无效组肠道菌群以及基因序列进行分析，从而筛选出与双相情感障碍治疗疗效相关性高的生物标志物，并且利用该标志物能够准确地预测双相情感障碍疗效，监测治疗效果。

[0016] 本发明提出的双相情感障碍患者相关生物标记物对治疗疗效的反应是有价值的。第一，粪便样本提取便携且无创，可增加患者依从性。同时。第二，粪便样本具有可运输性，样本分析具有准确性和安全性。第三，本发明的标记物具有较高的特异性和灵敏性，可用于治疗疗效的预测。

附图说明

[0017] 图 1 为根据本发明一个实施例物种水平上双相情感障碍患者治疗有效组和治疗无效组的菌群相对丰度差异情况。图示表明，双相情感障碍患者治疗有效组和治疗无效组在不同种水平菌群相对丰度存在显著差异。

[0018] 图 2 为根据本发明的一个实施例对分类器进行 5 次 10 折交叉验证的错误率分布情况。

[0019] 图 3 为根据本发明的一个实施例基于随机森林模型(3 个肠道标志物)，由双相情感障碍患者治疗有效组和治疗无效组组成的训练集的接收者操作特征 (Receiver Operating Characteristic, ROC) 曲线和曲线下面积 (Area under Curve, AUC)。

[0020] 图 4 为根据本发明的一个实施例基于随机森林模型(3 个肠道微生物标志物)，由双相情感障碍患者治疗有效组和治疗无效组组成的测试集的接收者操作特征 ROC 曲线和曲线下面积 AUC。

具体实施方式

[0021] 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：

“双相情感障碍”，是一类既有抑郁发作又有躁狂发作的慢性精神疾病，呈反复发作性，患者在情绪处理以及认知功能等方面存在障碍，严重影响患者正常的社会生活等功能。全世界共患病人口约 1%，是导致青少年残疾的四大原因之一，但发病机制尚未完全阐明，目前的观点认为与遗传和环境的交互作用关系密切。

[0022] “疗效”，也即“治疗效果”，是指药物或者手术等手段治疗疾病后的效果，主要为四级评价标准，包括痊愈、显效、进步、无效。本研究中的“疗效”主要指“显效”和“无效”。

因为在本发明中,针对双相情感障碍的治疗采取的手段是单一药物治疗,且以 17 项汉密尔顿抑郁量表(HDRS-17)量表得分的变化率是否 $\geq 50\%$ 来评定药物治疗 1 月后是否有效。

[0023] “生物标志物”,是一类可用于判断疾病的病因、早期诊断以及评估疾病发生发展过程或者评价药物治疗在目标人群中药物疗效或者安全性的指示物。主要包括处于不同生物学水平(个体、细胞、分子)上的任何反映机体特定生物状态(如疾病)的物质,如评估糖尿病的血糖等小分子物质以及蛋白质、核酸、脂质等大分子物质。在本发明中,“生物标志物”也可以用“肠道微生物”、“肠道菌群”来表示,因为本研究所发现的与双相情感障碍相关的生物标志物均来自于经受试者肠道代谢后的粪便样本。

[0024] 本发明所述的生物标志物,通过运用高通量测序,批量分析双相情感障碍患者治疗有效组和治疗无效组的粪便样本。基于高通量测序数据,对双相情感障碍患者治疗有效组和治疗无效组进行比对,从而确定与双相情感障碍患者群疗效预测相关的特异性菌群。

实施例

[0025] 样品的收集与处理:收集 62 例双相情感障碍的粪便样本,采集粪便样本后分装至做好标记的冻存管内,冷冻条件下运输并转运至 -80°C 冰箱内保存备用。具体的,本发明受试者是基于符合《精神疾病的诊断和统计手册》(DSM-IV-TR)双相情感障碍诊断依据的住院和/或门诊患者。入组时受试者未服用药物治疗或者曾服用药物治疗但已停药至少 3 个月。同时满足无明显的自杀想法或既往自杀未遂、没有与其他精神障碍共病的患者。排除标准为: 1)慢性感染、严重的系统性疾病(如糖尿病)和自身免疫性疾病; 2)筛查前 4 周内食用抗生素、益生菌或益生菌; 3)目前怀孕、哺乳或月经不调的女性; 4)颅脑外伤史; 5)磁共振成像(MRI)的禁忌症,如金属植入物或幽闭恐惧症。在入组时,使用 17 项汉密尔顿抑郁量表(HDRS-17)和蒙哥马利-阿斯伯格抑郁量表(MADRS)评定抑郁严重程度,用青年躁狂量表(YMRS)评定躁狂程度。其中, HDRS-17 得分不低于 14 分被设定为当前抑郁发作的阈值。在患者入组后给予喹硫平单药标准化治疗 1 个月后,再次评估 HDRS-17 量表并计算得分,与入组时 HDRS-17 量表得分进行比较,计算变化率。根据双相情感障碍患者服用喹硫平单药治疗 1 个月与治疗前比较, HDRS-17 评分变化率是否 $\geq 50\%$,将双相情感障碍分为治疗有效组(HDRS-17 $\geq 50\%$)和治疗无效组(HDRS-17 $< 50\%$)。最终受试者分为双相情感障碍患者治疗有效组($n=35$)和治疗无效组($n=19$)。8 名双相情感障碍患者缺乏随访量表评分而未被分类。

[0026] 提取 DNA: 使用 QIAGEN DNA 试剂盒提取 DNA,得到核酸样本。

[0027] 构建文库并测序: 使用 NEBNextUltra TM DNA Library Prep Kit for Illumina(NEB, 美国)构

建 DNA 文库。用 Illumina NovaSeq 6000 进行双端 150bp DNA 文库测序。再通过比对、筛选、质控等步骤去除污染或者低质量的测序片段(reads)，最终得到高质量的测序片段。

[0028] 将上述所得高质量测序片段输入到 Metaphlan2 分析软件 (<http://segatalab.cibio.unitn.it/tools/metaphlan2/>)，经 1)与参考的标记基因比对；2)统计插入片段的数量；3)对标记基因的长度进行标准化等步骤得到对应菌群的相对丰度信息。

[0029] 根据上述得到的菌群相对丰度情况，利用 LDA Effect Size (LEfSe) 分析技术分析治疗有效组和治疗无效组两组间的菌群相对丰度差异情况。

[0030] 将上述“双相情感障碍患者治疗有效组和治疗无效组菌群的相对丰度输入到 LEfSe 在线分析网页 (<http://huttenhower.sph.harvard.edu/galaxy>) 进行两组之间菌群相对丰度信息的差异性比较。具体地，主要包括以下三个步骤：1) 首先，利用非参数因子 Kruskal-Wallis 秩和检验检测两组之间的物种相对丰度差异，获得显著性差异物种；2) 其次，利用 Wilcoxon 秩和检验检测上一步所获得的具有显著性差异的物种的所有亚种是否趋向于同一分类级别；3) 最后用线性判别分析 (LDA)，对数据进行降维、评估差异显著的菌群影响力 (即 LDA score) 得到最终的差异物种 (参照文献：Segata N, et al. Metagenomic biomarker discovery and explanation[J]. Genome Biol, 2011, 12(6): R60.)。

[0031] 在本发明中，将 54 个样本 (35 个治疗有效的双相情感障碍患者组和 19 个治疗无效的双相情感障碍患者组)经测序、序列比对后得到的菌群相对丰度信息输入 LEfSe 在线分析网页，结果显示 6 种菌群的相对丰度，包括 "Bacteroides_clarus", "Eubacterium_biforme", "Weissella_confusa", "Ruminococcus_torques", "Bifidobacterium_dentium", "Collinsella_unclassified" 在双相情感障碍患者治疗有效组与无效组中存在明显的差异。其中，"Bacteroides_clarus", "Eubacterium_biforme", "Weissella_confusa", "Ruminococcus_torques", "Collinsella_unclassified" 在双相情感障碍患者治疗有效组相对丰度显著升高，而 "Bifidobacterium_dentium" 则下降 (图 1)。图 1 采用 LEfSe 和 LDA 分析比较差异菌群相。以 $LDA \geq 2$ 作为差异有显著性的阈值。LDA 评分显示治疗无效组的 BD 患者 (右边部分，uneffective) 和治疗有效组 (左边，effective) 之间的细菌差异显著。

[0032] 利用随机森林分类器筛选双相情感障碍治疗过程中疗效预测的潜在生物标志物。

[0033] 为进一步筛选双相情感障碍肠道微生物标志物，以上述 LEfSe 分析技术筛选的双相情感障碍患者治疗有效组和无效组之间差异菌群相对丰度信息情况为基础，构建双相情感障碍治疗有效组和无效组肠道微生物标志物的训练集和测试集，并评估待测的测试集样本生物标

志物的含量值。其中，在本发明的实施方案中，训练集是指具有一定样本数的双相情感障碍治疗有效组和无效组受试者待测样本中各生物标志物含量的数据集合。

[0034] 具体包括如下步骤：

本发明从 54 个样品(35 个治疗有效的双相情感障碍患者组和 19 个治疗无效的双相情感障碍患者组)中，选取 27 个双相情感障碍患者和 14 个健康人作为训练集，剩余的 13 个样品 (8 个双相情感障碍患者和 5 个健康人) 作为测试集。

[0035] 利用 Python3 软件版本调用随机森林分类器(Random Forest Classifier, RF)。本发明对 RF 分类器进行 5 次 10 折交叉验证 (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)。

[0036] 本发明基于 5 次 10 折交叉验证结果，RF 分类器最终选择 3 个最优生物标记物组合，(训练集详细微生物标记物相对丰度信息见下表 1-1、1-2 所示，测试集详细微生物标记物相对丰度信息见下表 2 所示，表 3 显示出了 3 种生物标记物结合来预测训练集的疗效概率)。

[0037]

表 1-1 训练集差异菌群相对丰度信息

样本编号 (ID)	肠道微生物标志物相对丰度		
	Eubacterium_biform	Ruminococcus_torq	Collinsella_unclassifi
B1038	0	0.18418	0
s1B1069	1.2513	0.90922	0
BP23	0	0.04954	0.01088
HSH_104	0	0.0611	0
HSH_31	0.01577	0.69809	0
HSH_79	3.20858	2.63285	0
s1B1008	0.10844	0.15687	0
s1B2002	1.17929	0.79934	0
s1B1055	0	2.81496	0.33633
B1004	0	0	0
B1066	0	1.49858	0
B1103	1.27933	1.495	0
B1111	0	0	0
BP_2	0	0.26273	0
BP_9	0	0.09136	0
HSH_106	0	0.22034	0
HSH_109	0	0.13116	0
HSH_112	13.94316	8.34478	0
HSH_33	0	0.09347	0
HSH_39	0	0.16655	0
HSH_44	0.12433	0.26639	0
HSH_5	0	0	0
HSH_54	0	1.96458	0
HSH_62	0.24962	0.01857	0
HSH_92	0	0.31002	0
HSH_96	0	0.14818	0
s1B1065	0	0.32624	0
B1072	0	1.12409	0.01994

注释：B、BP、S1B、HSH：双相情感障碍

表 1-2 训练集差异菌群相对丰度信息

样本编号 (ID)	肠道微生物标志物相对丰度		
	Eubacterium_bifome	Ruminococcus_torques	Collinsella_unclassified
B2001	0	0.40142	0.02637
BP_3	0	0.13326	0.04488
BP_6	0	0.07184	0
HSH_157	0	0.27708	0.0739
HSH_21	0	1.35755	0
HSH_27	0.53173	0.0709	0
HSH_76	0	0.01317	0
HSH_88	0	1.04308	0
HSH_98	0	0.0982	0
s1B1063	0	0.27477	0
s1B1094	0	0.01742	0.00969
s1B2018	0	0.07185	0
s1B2008	0	0.0267	0

注释：B、BP、S1B、HSH：双相情感障碍

表 2 测试集差异菌群相对丰度信息

样本编号 (ID)	肠道微生物标志物相对丰度		
	Eubacterium_bifome	Ruminococcus_torques	Collinsella_unclassified
s1B1067	0	0.28133	0
s1B1098	0	0.78627	0
s1B2005	0	0.48294	0
s1B2020	0	0.3933	0
s1B2006	0	0.22309	0
s1B2007	0.05023	0.1966	0
s1B2010	0	6.144	0
HSH_120	0	0.34298	0
B1076	0	0.01986	0
BP21	0	0.08073	0
HSH_160	0	0.00896	0
s1B1050	0	0.03433	0
B1047	0	0.20299	0

注释：B、BP、S1B、HSH：双相情感障碍

表 3 训练集利用菌群标志物相对丰度信息预测治疗疗效的准确性

ID	Accuracy	ID	Accuracy
B1038	0.76579	HSH_5	0.609374
s1B1069	0.892749	HSH_54	0.745503
BP23	0.481324	HSH_62	0.712095

HSH_104	0.546565	HSH_92	0.714679
HSH_31	0.795832	HSH_96	0.761402
HSH_79	0.922007	s1B1065	0.714679
s1B1008	0.908255	B1072	0.292687
s1B2002	0.893349	B2001	0.306945
s1B1055	0.482749	BP_3	0.295577
B1004	0.609374	BP_6	0.53206
B1066	0.725082	HSH_157	0.292865
B1103	0.910725	HSH_21	0.627038
B1111	0.609374	HSH_27	0.567371
BP_2	0.757439	HSH_76	0.520719
BP_9	0.65124	HSH_88	0.598444
HSH_106	0.76572	HSH_98	0.649793
HSH_109	0.729001	s1B1063	0.707049
HSH_112	0.922245	s1B1094	0.420791
HSH_33	0.65124	s1B2018	0.53206
HSH_39	0.76579	s1B2008	0.514091
HSH_44	0.906386		

注释：B、BP、S1B、HSH：双相情感障碍。

[0038] 利用 RF 模型筛选训练集中菌群相对丰度对每一个体计算其双相情感障碍疗效预测风险，绘制 ROC 曲线，并计算出 AUC，作为判别模型效能评价参数。其中特异性表征的是对于无效判对的概率，敏感性指的是对于有效判对的概率，对训练集样本的判别效能为：AUC=91%，95%置信区间 CI=82.4%~99.6%。结果表明该模型所得代谢物组合可作为区分双相情感障碍与非双相情感障碍的潜在生物标志物（图 3）。

[0039] 基于随机森林模型(3 个生物标志物)由双相情感障碍有效和无效组组成的测试集的 ROC 曲线和 AUC，其中特异性表征的是对于无效判对的概率，敏感性指的是对于有效判对的概率，对训练集样本的判别效能为：AUC=82.5%，95%置信区间 CI=51.2~100%（图 4）。结果表明该模型所得标志物组合可作为区分双相情感障碍治疗有效和治疗无效疗效预测的潜在生物标志物。

[0040] 表 4 示出了 3 种生物标记物结合来预测测试集的患病概率。

[0041] 表 5 示出了 3 种生物标记物的详细信息。

[0042] 表 4 测试集利用菌群标志物相对丰度预测患病率的准确性

ID	Accuracy
s1B1067	0.707049
s1B1098	0.659275
s1B2005	0.688713

s1B2020	0.688713
s1B2006	0.76572
s1B2007	0.864901
s1B2010	0.74717
HSH_120	0.713914
B1076	0.51936
BP21	0.53206
HSH_160	0.520719
s1B1050	0.51756
B1047	0.766125

注释：B、BP、S1B、HSH：双相情感障碍

表 5

标志物名称	模型预测中的重要性	富集方向	测试集 AUC (%)	95%置信区间
Eubacterium_biforme	0.15259	effective	56.3	2.38~8.87
Ruminococcus_torques	0.339395	effective	80	4.49~10
Collinsella_unclassified	0.258467	effective	50	1.65~8.35

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，可以参照《分子克隆实验指南》第三版或者相关产品进行，所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，均与首次标明的内容相同。

[0043] 本发明采用宏基因组关联分析(Metagenome-Wide Association Study ,MWAS)的分析方法，经测序分析粪便样本的菌群组成及菌群相对丰度；用 Lefse 分析方法分析双相情感障碍患者治疗有效组和治疗无效组菌群相对丰度的差异情况；用随机森林判别模型判别双相情感障碍治疗有效组和治疗无效组群体，获得患病概率，用于双相情感障碍的疗效预测评估或者寻找潜在药物靶点。

[0044] 在本发明中，所述的测序(二代测序)和 MWAS 具有本领域所公知，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，可以依据文献(Jun Wang ,and Huijue Jia .Metagenome-wide association studies:fine-mining the microbiome .Nature Reviews Microbiology14 .8(2016):508-522 .)中记载的方法进行。

[0045] 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, et al. Untargeted Metabolic Profiling Identifies Altered Serum Metabolites of Type 2-Diabetes Mellitus in a Prospective ,Nested Case Control Study .Clin Chem 2015 ,61:487-497 .;)中记载的方法进行。

[0046] 在本发明中,构建了双相情感障碍患者治疗有效组和治疗无效组受试者生物标志物的训练集,在此基础上,对待测样本的生物标志物含量值进行评估。

[0047] 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,得到双相情感障碍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。

[0048] 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,为双相情感障碍患病治疗疗效的预测,寻找潜在药物靶点提供依据。

[0049] 因此,本发明提出以下应用:

所述的基于肠道菌群的双相情感障碍生物标志物组合作为检测靶点或检测目标在制备检测试剂盒中的应用。

[0050] 所述的基于肠道菌群的双相情感障碍治疗疗效生物标志物组合作为靶点在预测治疗中的应用。

[0051] 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. 基于肠道微生物的双相情感障碍疗效预测生物标志物，其特征在于，包括以下 6 种中的多种：

Bacteroides_clarus, *Eubacterium_biforme*, *Weissella_confusa*, *Ruminococcus_torques*, *Bifidobacterium_dentium*, *Collinsella_unclassified* 在双相情感障碍患者治疗有效组与无效组中存在明显的差异；

其中，*Bacteroides_clarus*, *Eubacterium_biforme*, *Weissella_confusa* , *Ruminococcus_torques*, *Collinsella_unclassified* 在双相情感障碍患者治疗有效组相对丰度显著升高，而 *Bifidobacterium_dentium* 则下降。

2. 根据权利要求 1 所述的生物标志物，其特征在于，为以下 3 种：

生物标志物 1) *Eubacterium_biforme*；

生物标志物 2) *Ruminococcus_torques*；

生物标志物 3) *Collinsella_unclassified*；

上述所述生物标志物组合在双相情感障碍患者治疗有效组中相对丰度均显著升高。

3. 根据权利要求 1 所述的生物标志物，其特征在于，所述的生物标志物是基于对其基因序列的计算所提供的。

4. 根据权利要求 1 或者 2 所述的生物标志物，其特征在于，所述的生物标志物的相对丰度信息用于和参考值进行比较。

5. 一种根据权利要求 1 或者 2 所述的生物标志物作为检测靶点或检测目标在制备检测试剂盒中的应用。

6. 一种根据权利要求 1 或者 2 所述的生物标志物作为靶点在疗效预测中的应用。

7. 一种根据权利要求 1 所述的生物标志物的筛选方法，其特征在于，步骤如下：

1) 样本收集：收集样本受试者包括双相情感障碍患者治疗前的粪便样品，在冰箱内 -80℃ 条件下保存备用；

2) 对收集后储存的粪便样本进行核酸样本 DNA 的提取，宏基因组测序与组装、比对、筛选与质控之后，将高质量的测序片段输入到 *Metaphlan2* 软件，计算出物种的相对丰度；

3) 将上述所得双相情感障碍患者与健康对照物种的相对丰度信息输入到 *LDA Effect Size (LEfSe)* 系统，分析组间差异菌群。

8. 根据权利要求 7 所述的筛选方法，进一步使用随机森林模型预测分析，步骤如下：

4.1) 双相情感障碍患者治疗有效组和治疗无效组为训练集，剩余样本作为测试集，计算训练集内每个样本中物种的相对丰度；

4.2) 将训练集中物种的相对丰度信息输入随机森林（RF）分类器中，并对分类器进行 5 次 10 折的交叉验证，对利用 RF 模型筛选出的每一个物种，依据其相对丰度信息计算双相情感障碍的患病风险、绘制 ROC 曲线，并计算其曲线下面积（AUC），将 AUC 作为判别模型效能评价的参数，在模型中输出每个物种的重要性指数，重要性指数越高，代表该标志物用来判别双相情感障碍和非双相情感障碍的重要性就越高。

9. 根据权利要求 8 所述的筛选方法，其特征在于，所述训练集中，样本受试者包括 27 个双相情感障碍治疗有效患者和 14 个治疗无效患者，测试集中，样本受试者包括 8 个双相情感障碍治疗有效患者和 5 个治疗无效患者。

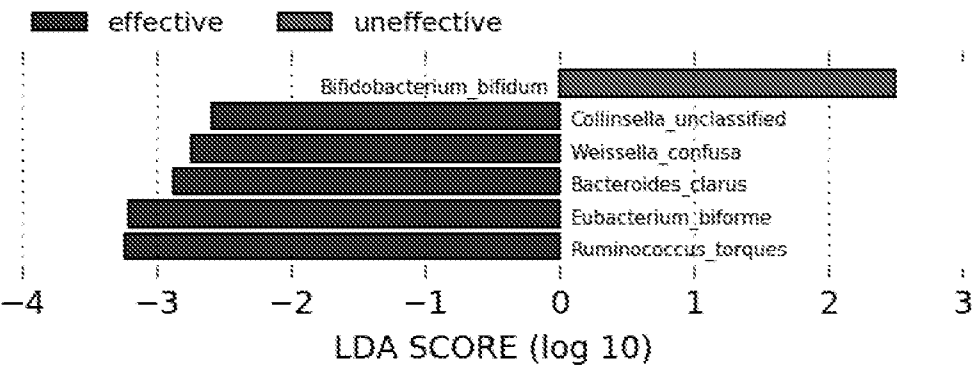


图 1

Cross-validation error diagram

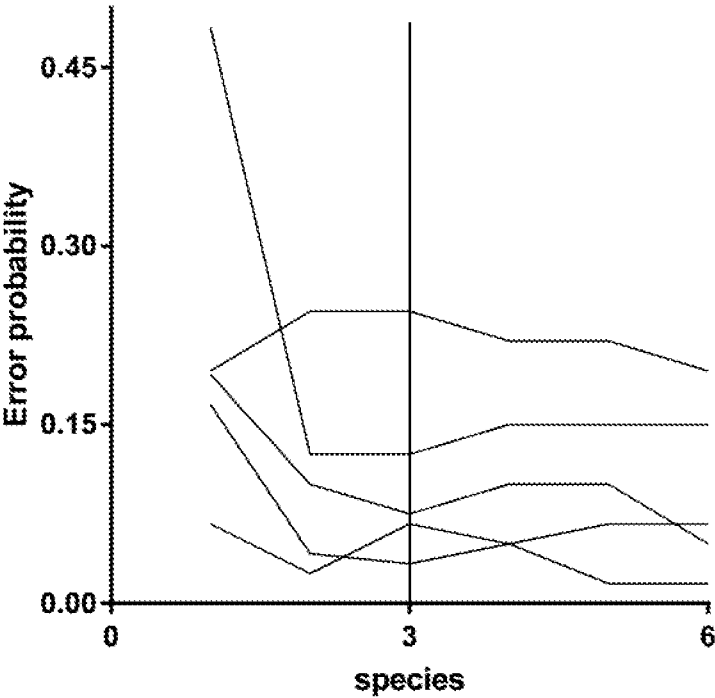


图 2

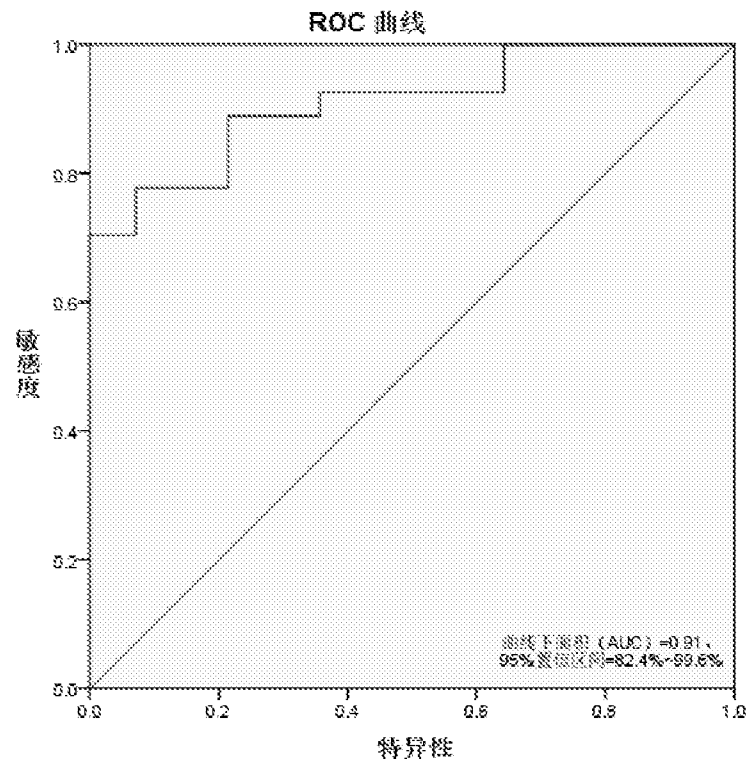


图 3

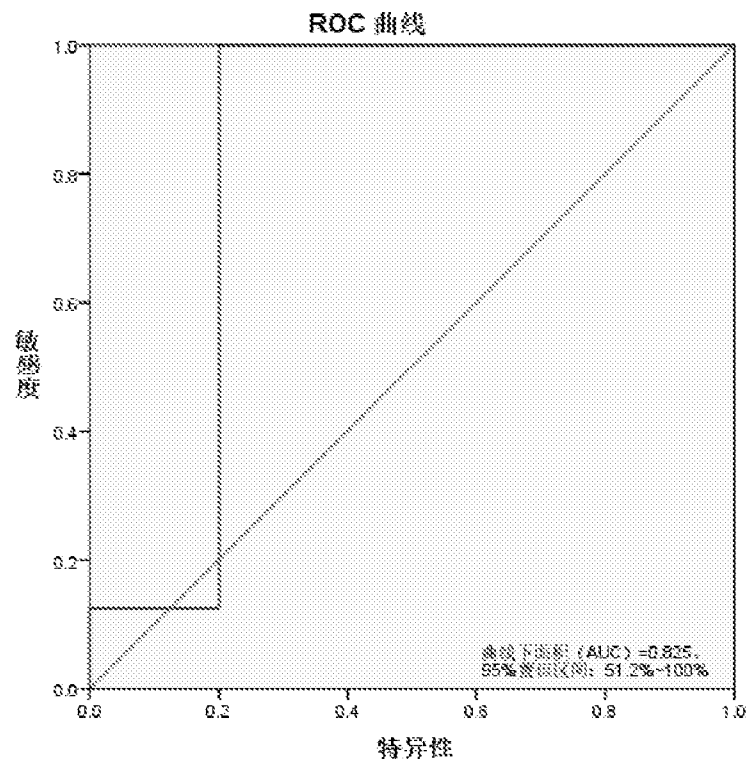


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/081945

A. CLASSIFICATION OF SUBJECT MATTER G16H 50/20(2018.01)i; G16B 30/00(2019.01)i; G06N 3/00(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) G16H; G16B; G06N 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, DWPI, CNTXT, WOTXT, USTXT, EPTXT, JPTXT, CNKI, Web of Science, 百度学术, Patents: applicant/inventor, 克拉鲁斯拟杆菌, Bacteroides clarus, 两形真杆菌, Eubacterium biforme, 融合魏斯氏菌, Weissella confusa, 扭链瘤胃球菌, Ruminococcus torques, 齿双歧杆菌, Bifidobacterium dentium, 柯林斯氏菌, Collinsella, 肠道微生物, gut microbiota, 双相情感障碍, 抑郁, bipolar disorder, depression				
C. DOCUMENTS CONSIDERED TO BE RELEVANT				
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
Y	黄婷婷 (HUANG, Tingting). "双相障碍抑郁相患者肠道菌群特征的研究 (Gut Microbiota Characteristics in Patients with Bipolar Depression)" 中国优秀硕士学位论文全文数据库医药卫生科技辑. (Chinese Master's Theses Full-Text Database, Medicine and Health Sciences), No. 2020, 03, 15 March 2020 (2020-03-15), article no. E071-62	1-9		
Y	CN 110838365 A (KANGMEIHUA BIG GENE TECHNOLOGY CO., LTD.) 25 February 2020 (2020-02-25) description paragraphs 2, 7-28	1-9		
Y	LU, Q.Q. et al. "Gut Microbiota in Bipolar Depression and Its Relationship to Brain Function: An Advanced Exploration." FRONTIERS IN PSYCHIATRY., Vol. 10, 29 October 2019 (2019-10-29), article number 784, pages 1-11	1-9		
Y	HU, S.H. et al. "Gut Microbiota Changes in Patients with Bipolar Depression." ADV.SCI., No. 6, 15 May 2019 (2019-05-15), article number 1900752, pages 1-11	1-9		
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.				
<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top; border: none;"> * 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 </td> <td style="width: 50%; vertical-align: top; border: none;"> "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 </td> </tr> </table>			* 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
* 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 01 December 2020		Date of mailing of the international search report 16 December 2020		
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China Facsimile No. (86-10)62019451		Authorized officer Telephone No.		

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/081945

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ZHENG, P. et al. "Gut Microbial Signatures Can Discriminate Unipolar from Bipolar Depression." <i>ADV.SCI.</i> , Vol. 7, No. 7, 05 February 2020 (2020-02-05), article number 1902862, pages 1-11	1-9
A	CN 105603066 A (JIN, Feng) 25 May 2016 (2016-05-25) entire document	1-9
A	HUANG, T.T. et al. "Current Understanding of Gut Microbiota in Mood Disorders: An Update of Human Studies." <i>FRONTIERS IN GENETICS.</i> , Vol. 10, 19 February 2019 (2019-02-19), article number 98, pages 1-12	1-9
A	郭丽阳 等 (GUO, Liyang et al.). "双相障碍不同临床相肠道菌群多样性及相对丰度的对照研究 (The Diversity and the Abundance of Gut Microbiome in Patients with Bipolar Disorder)" <i>中华精神科杂志 (Chinese Journal of Psychiatry)</i> , Vol. 51, No. 2, 30 April 2018 (2018-04-30), pp. 98-104	1-9

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/081945

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	110838365	A	25 February 2020	None	
CN	105603066	A	25 May 2016	CN 105603066	B 14 July 2020

国际检索报告

国际申请号

PCT/CN2020/081945

A. 主题的分类 G16H 50/20(2018.01)i; G16B 30/00(2019.01)i; G06N 3/00(2006.01)i 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		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G16H; G16B; G06N 包含在检索领域中的除最低限度文献以外的检索文献 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, DWPI, CNTXT, WOTXT, USTXT, EPTXT, JPTXT, CNKI, Web of Science, 百度学术, Patentics: 申请人/发明人, 克拉鲁斯拟杆菌, Bacteroides clarus, 两形真杆菌, Eubacterium biforme, 融合魏斯氏菌, Weissella confusa, 扭链瘤胃球菌, Ruminococcus torques, 齿双歧杆菌, Bifidobacterium dentium, 柯林斯氏菌, Collinsella, 肠道微生物, gut microbiota, 双相情感障碍, 抑郁, bipolar disorder, depression		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	黄婷婷. "双相障碍抑郁相患者肠道菌群特征的研究." 中国优秀硕士学位论文全文数据库医药卫生科技辑., 第2020年03期, 2020年 3月 15日 (2020 - 03 - 15), 文章号E071-62	1-9
Y	CN 110838365 A (康美华大基因技术有限公司) 2020年 2月 25日 (2020 - 02 - 25) 说明书第2、7-28段	1-9
Y	LU, Q.Q. 等. "Gut Microbiota in Bipolar Depression and Its Relationship to Brain Function: An Advanced Exploration." FRONTIERS IN PSYCHIATRY., 第10卷, 2019年 10月 29日 (2019 - 10 - 29), 文章号784, 第1-11页	1-9
Y	HU, S.H. 等. "Gut Microbiota Changes in Patients with Bipolar Depression." ADV. SCI., 第6期, 2019年 5月 15日 (2019 - 05 - 15), 文章号1900752, 第1-11页	1-9
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 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		
国际检索实际完成的日期 2020年 12月 1日		国际检索报告邮寄日期 2020年 12月 16日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 冯晓亮 电话号码 86-(10)-53961927

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
Y	ZHENG, P. 等. "Gut Microbial Signatures Can Discriminate Unipolar from Bipolar Depression." ADV. SCI., 第7卷, 第7期, 2020年 2月 5日 (2020 - 02 - 05), 文章号1902862, 第1-11页	1-9
A	CN 105603066 A (金锋) 2016年 5月 25日 (2016 - 05 - 25) 全文	1-9
A	HUANG, T. T. 等. "Current Understanding of Gut Microbiota in Mood Disorders: An Update of Human Studies." FRONTIERS IN GENETICS., 第10卷, 2019年 2月 19日 (2019 - 02 - 19), 文章号98, 第1-12页	1-9
A	郭丽阳 等. "双相障碍不同临床相肠道菌群多样性及相对丰度的对照研究." 中华精神科杂志., 第51卷, 第2期, 2018年 4月 30日 (2018 - 04 - 30), 第98-104页	1-9

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/081945

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	110838365	A	2020年 2月 25日	无	
CN	105603066	A	2016年 5月 25日	CN	105603066 B 2020年 7月 14日