

# 团 体 标 准

T/SZGIA 1.2-2018

---

## 基于高通量测序的环境微生物检测 第 2 部分：人粪便微生物宏基因组检测方法

Detection of environmental microorganisms based on high-throughput sequencing  
Part 2: Methodology of detecting human fecal microbiota using metagenomic  
sequencing

2018-01-30 发布

2018-01-31 实施

---

深圳基因产学研资联盟 发布

# 前 言

本部分编写格式遵循了 GB/T 1.1-2009 给出的规则编写。

SZTT/SZGIA 1-2016 《基于高通量测序的环境微生物检测》分为四个部分：

- 第1部分：基本规程
- 第2部分：人粪便微生物宏基因组检测方法
- 第3部分：人肠道微生物 16s rRNA基因检测法
- 第4部分：临床样本病原微生物检测

本部分为SZTT/SZGIA 1-2016 的第2部分。

本部分代替SZTT/SZGIA 1.2-2016，与SZTT/SZGIA 1.2-2016相比，除编辑性修改外主要技术变化如下：

- 修订了“范围”部分描述（见“1 范围”）
- 增加了“术语和定义”部分内容（见3.2, 3.3, 3.4和3.5）
- 增加了“主要仪器和主要试剂”的内容（见“5 主要试剂”和“6 主要设备和材料”）
- 增加了“数据处理，结果与报告”部分内容（见7.4.5和7.4.8）
- 增加了“附录B、附录C”。

本部分由深圳基因产学研资联盟提出并归口。

本部分负责起草单位：深圳华大生命科学研究院、深圳基因产学研资联盟、深圳华大基因科技有限公司、深圳华大临床检验中心有限公司、南方医科大学、深圳龙岗区妇幼保健院、深圳市谱元基因研究院、深圳弘睿康生物科技有限公司、深圳市微健康基因研究院、华大精准营养（深圳）科技有限公司、深圳市恒创基因科技有限公司、上海锐翌生物科技有限公司、蓝色彩虹（深圳）科技有限公司、上海凡迪生物科技有限公司。

本部分主要起草人：陈冰、李俊桦、钟焕姿、杨芳明、林宇翔、韩默、蔡锴晔、唐美芳、周媛、蒋慧、罗佳波、谭晓梅、范钦、罗兰、魏凤香、李陶莎、程奇、谢强、杜佳婷、李倩一、李岱怡、覃俊杰、胡鹏、马振华、戴文魁、秦楠、刘靛、李骏、陈蕾。

# 基于高通量测序的环境微生物检测 第2部分:人粪便微生物宏基因组检测方法

## 1 范围

本部分规定了采用高通量测序技术进行人粪便微生物宏基因组检测的方法。  
本部分适用于利用宏基因组学方法进行人粪便微生物检测。

## 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅所注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本适用于本文件。

GB/T 6682 分析实验室用水规格和实验方法

GB/T 19495.2-2004 转基因产品检测实验室技术要求

GB/T 19495.3-2004 转基因产品检测 核酸提取纯化方法

GB/T 19495.5-2004 转基因产品检测核酸定量 PCR检测方法

GB 19489-2008 实验室生物安全通用要求

SNT 2497.21-2010 进出口危险化学品安全试验方法 第21部分:琼脂糖凝胶电泳试验

SZTT/SZGIA 1-2016 基于高通量测序的环境微生物检测 第1部分:基本规程

## 3 术语和定义

### 3.1

**微生物宏基因组** microbial metagenome

微生物宏基因组是以特定环境中整个微生物群落作为研究对象,无需分离培养,直接提取环境样本中全部微生物的DNA,分析其遗传组成、物种分类、系统进化、基因功能及代谢网络等。

### 3.2

**稳定剂** stabilizer

保持微生物核酸物质稳定的化学试剂。

### 3.3

**碱基识别质量** quality of base calling

衡量碱基正确识别的概率,通常以数字值直接表示。

碱基识别质量与碱基识别错误率的关系可用公式(1)表示:

$$Q = -10 \times \lg P \dots\dots\dots (1)$$

式中:

$Q$ ——碱基识别质量;

$P$ ——碱基识别错误率,即单次测序错误碱基数占碱基总数的比例,与碱基识别正确率相对。

### 3.4

#### Q20

测序数据中，碱基识别质量值大于20的碱基占有所有碱基的比例。

### 3.5

#### Q30

测序数据中，碱基识别质量值大于30的碱基占有所有碱基的比例。

注：碱基识别质量值为30时，表示该碱基的准确率为99.9%以上。Q30 ≥ 85%，则表示测序数据中85%以上的碱基质量值大于30。

### 3.6

#### 直接裂解法 direct lysis method

将样品直接悬浮在裂解缓冲液中处理，继而抽提纯化DNA。

### 3.7

#### 间接提取法 indirect extraction method

先用物理方法将微生物细胞从环境中分离出来，然后采用较温和的方法提取DNA。

## 4 缩略语

下列缩略语适合于本文件。

DNA——脱氧核糖核酸 (deoxyribonucleic acid)

bp——碱基对 (base pair)

## 5 主要试剂

除特别说明外，本标准所用试剂为分析纯或生化试剂。

### 5.1 DNA 提取试剂盒

### 5.2 琼脂糖

### 5.3 建库试剂盒

### 5.4 高通量测序试剂盒

注：分析纯：是化学试剂的一种纯度规格，属于二级品，标签为金光红，其纯度略低于优级纯，杂质含量略高于优级纯，适用于重要分析和一般性研究工作，如一般实验室、研究所等地方常用试剂。生化试剂：是指有关生命科学研究的生物材料或有机化合物，以及临床诊断、医学研究用的试剂。

## 6 主要设备和材料

### 6.1 高通量测序仪

### 6.2 紫外分光光度计

### 6.3 电泳仪

## 6.4 凝胶成像系统

## 6.5 低温冰箱：-20℃和-80℃

# 7 检测方法

## 7.1 检测对象

人粪便样品中的微生物。

## 7.2 检测要求

样品采集应符合国家法律法规及伦理规范。

## 7.3 样品采集、保存与运输

为有效地指导粪便样品的采集、保存与运输，使粪便样品中的待测微生物不受影响，保证检测结果准确可靠，开展宏基因组高通量测序所需粪便样品的采集、保存与运输应参考本标准。

粪便样品采集总则：使用洁净的器皿收集粪便，排便后即刻进行采样，采集时需避免器皿中粪便被水、尿液或其他分泌物（如经血）污染。

使用粪便杯和含稳定液的自采样套装两种方法对粪便样品的采集、保存与运输的详述指导标准请参考7.3.1与7.3.2。

### 7.3.1 粪便杯采集法

#### 7.3.1.1 采集

使用洁净的器皿收集粪便，采集时请确保器皿中无水、尿液或其他分泌物（如经血）污染。排便后即刻使用一次性粪便杯配套取样勺对粪便样品中部取样，将粪便从洁净器皿转移至粪便杯，立刻旋紧盖子。每个粪便杯中转移至少2 cm<sup>3</sup>体积大小的粪便样品。

#### 7.3.1.2 保存与运输

粪便杯采集粪便样品后，应立即置于干冰或-80℃保存。若不能即刻转移至上述环境保存，可暂存于4℃或4℃以下环境（如冰盒），30分钟内转移至干冰或-80℃保存。

干冰条件下运输。

### 7.3.2 含稳定液自采样套装采集法

#### 7.3.2.1 采集

使用洁净的器皿收集粪便，采集时请确保器皿中无水、尿液或其他分泌物（如经血）污染。排便后即刻按照自采样套装操作要求取中段粪便，将采集的粪便样品迅速转移至含有稳定剂的保存管中，使稳定液完全浸没粪便样品，盖紧管盖。

#### 7.3.2.2 保存与运输

含稳定液保存管采样后，该保存管按操作说明可于常温保存和运输，避免阳光直射。常温放置时间应不超过自采样套装操作说明限定的常温保存时间范围。对将超出自采样套装操作说明限定时间的常温粪便保存管需转移至-80℃冻存，在干冰条件下运输。

## 7.4 检测步骤

### 7.4.1 实验室条件

1) 用于人粪便微生物宏基因组检测的实验室设施和环境应符合 GB 19489-2008 和 GB/T

19495.2-2004 的规定。

2) 实验室用水应符合 GB/T 6682 的规定。

3) 实验室应按照不同工作内容划分独立区域，并有明显标志。如试剂储备和准备区，样品制备区，扩增区等，各区域间应避免交叉污染。

#### 7.4.2 样品接收与处理

1) 核对样品编号，记录样品信息，检查样品状态，应无密封不严导致泄漏或者污染，应无尿液、水、或其他分泌物（如经血）等杂质污染，否则该样品废弃，重新采样。准确记录相关信息（可参考附录 A 所示），保证样品的可追溯性。

2) 使用稳定高度可重复的方法提取粪便中微生物总 DNA。提取方法宜采取直接裂解法或间接提取法，宜使用规范 DNA 提取试剂盒进行粪便 DNA 提取。

#### 7.4.3 DNA 样品判定标准

粪便样品提取微生物总 DNA 后，依据 GB/T 19495.3-2004 和 GB/T 19495.5-2004 的规定检测 DNA 样品，并按表 1 所示的判定标准来区分其质量。

表1 DNA 样品质量判断标准

| 样品类型   | 检测结果                                       |                                    |               | 判定结论 |
|--------|--|------------------------------------|---------------|------|
| 基因组DNA | $m \geq 2.0 \mu\text{g}$                   | $c \geq 10 \text{ ng}/\mu\text{l}$ | 无降解           | A类   |
|        | $1.0 \mu\text{g} \leq m < 2.0 \mu\text{g}$ |                                    | 或轻微降解         | B类   |
|        | $m < 1.0 \mu\text{g}$                      | $c \leq 10 \text{ ng}/\mu\text{l}$ | 中度降解<br>或重度降解 | C类   |

注：1) m—总量(total mass)的缩写，此处指DNA总量；

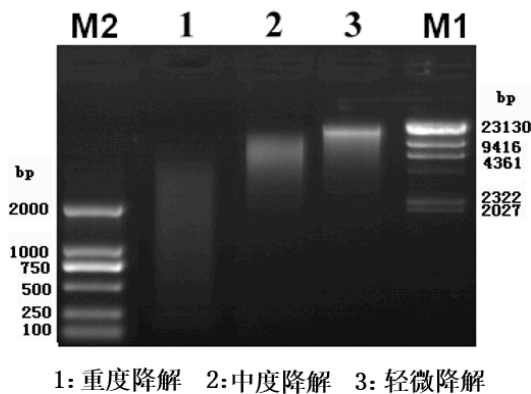
2) c—浓度(concentration)的缩写，此处指DNA浓度；

3) A类样品需同时满足 $m \geq 2.0 \mu\text{g}$ ， $c \geq 10.0 \text{ ng}/\mu\text{l}$ ，无降解或轻微降解三个条件，合格；

4) B类样品需同时满足 $1.0 \mu\text{g} \leq m < 2.0 \mu\text{g}$ ， $c \geq 10.0 \text{ ng}/\mu\text{l}$ ，无降解或轻度降解三个条件，合格；

5) C类样品需满足 $m < 1.0 \mu\text{g}$ ， $c \leq 10.0 \text{ ng}/\mu\text{l}$ ，中度降解或重度降解三个条件其中之一或以上，不完全合格。

常用荧光定量仪检测 DNA 样品浓度，根据 SNT 2497.21-2010 的规定检测 DNA 样品的完整性，DNA 样品完整性判断的标准如下图 1 所示。



1: 重度降解 2: 中度降解 3: 轻微降解

图1 基因组DNA样品降解程度示意图

注：1) 重度降解：电泳图中对应样品的泳道无明显主带，拖尾严重，如泳道1所示。

2) 中度降解：电泳图中对应样品的泳道可见1400bp左右主带，略有拖尾，如泳道2所示。

3) 轻微降解：电泳图中对应样品的泳道可见2000bp左右主带，略有拖尾，如泳道3所示。

#### 7.4.4 文库构建与高通量测序

使用相应建库和高通量测序试剂盒构建文库并测序，高通量平台测序过程参见SZTT/SZGIA 1-2016的规定。

根据表1的判断标准，建议如表2所示选择合适的DNA样品。

表2 建库测序选择标准

| DNA样品类别 | 建议                          |
|---------|-----------------------------|
| A类      | 满足建库测序要求                    |
| B类      | 满足建库测序要求                    |
| C类      | 不完全满足建库测序要求，宜重新采样和重新提取样品DNA |

#### 7.4.5 数据处理

样品经过高通量测序得到的原始数据，需完成质量控制，通过比对去除人基因组序列等基本步骤，生成非人源高质量短序列，再进行后续生物信息学分析过程。

#### 7.4.6 数据分析

分析过程按照SZTT/SZGIA 1-2016的规定执行。

#### 7.4.7 数据质控要求

##### 7.4.7.1 质量要求

测序完成后，每个DNA样品对应的有效可用数据质量值应符合如下要求：

- 1) Q20≥90%，该样品测序结果 90%以上的碱基质量值大于 20；
- 2) Q30≥80%，该样品测序结果 80%以上的碱基质量值大于 30。

注：本质量要求仅适用于读长不超过 300bp 的短序列。

##### 7.4.7.2 数量要求

为保证后续生物信息学分析的准确性和稳定性，要求每个样品生成的有效可用非人源高质量短序列数目（即单端测序的序列条数，或双端测序序列对数），至少大于 $1 \times 10^7$ 条。

#### 7.4.8 结果与报告

综合以上检测步骤和处理过程，报告样品中物种组成及各成分相对丰度。

##### 7.4.8.1 检测报告应包括下列信息：

- 1) 检测机构名称，联系方式和地址，检测员，样品接收时间，样品检测时间，报告时间；
- 2) 测序平台与测序策略；
- 3) 描述数据分析对应使用的注释方法和数据库版本情况；
- 4) 样品物种组成及各相应丰度，可选择适合的多元展现方式；
- 5) 其他个性化功能分析等内容（如有）。

##### 7.4.8.2 标准样品测试评估

使用已知菌株DNA组成的标准样品进行宏基因组检测，评估各检测机构检测方法的准确性，其检测操作规程及检测报告参考附录B和附录C。

各检测机构应能检出标准样品的全部物种组成；标准样品各物种相对丰度检测值波动范围小于或等于理论值的 $\pm 20\%$ ，检测流程判定“优秀”；波动范围大于理论值的 $\pm 20\%$ ，小于理论值的 $\pm 50\%$ ，检测流程判定“良好”；波动范围大于或等于理论值的 $\pm 50\%$ ，检测流程判定“待改进”。

#### 7.4.8.3 局限性描述

各检测机构在描述结果时应注明检测结果的局限性，也应说明非本标准规定或被认为是可供选择的所有操作细节以及可能影响检测结果的任何细节。



附 录 A  
(资料性附录)  
样品信息单

|            |  |   |  |
|------------|--|---|--|
| 运输信息       | 运输公司：_____ 快递单号：_____ 城市：_____   |   |  |
|            | 样品寄件人：_____ 样品寄件人联系电话：_____  |   |  |
|            | 运输状态： <input type="checkbox"/> 干冰 <input type="checkbox"/> 冰袋 <input type="checkbox"/> 常温 <input type="checkbox"/> 其他：_____                      |   |  |
| 样品联系人*     |  | 样品联系人电话   |  |
| 样品联系人邮箱*   |  | 样品联系人单位   |  |
| 有无传染性、致病性* | <input type="checkbox"/> 无 <input type="checkbox"/> 有（致病传染性说明：_____） <input type="checkbox"/> 未知   |   |  |
| 实验室环境要求*   | <input type="checkbox"/> 普通实验室 <input type="checkbox"/> P2 实验室（生物安全柜）<br>注：若粪便样本无传染性、致病性，宜在普通实验室或 P2 实验室开展相关实验；若粪便样本有传染性或传染性未知，宜在 P2 实验室中开展相关实验。 |   |  |
| 详细提取方案*    |  |   |  |
| 样品信息       | 信息栏  | 填写标准  |  |
| 样品名称*      |  | 1. 样品名称必需唯一化以区分其他样本；<br>2. 样品名称可由字母及数字组成；不能包含“_”、“-”、“#”、“%”、“/”，“&”和“*”等非常规符号。<br>3. 人体相关样品名称不能包含任何可识别个体身份的信息。 |  |
| 样品状态*      |  | 样本状态包括：DNA、速冻粪便、含稳定剂组织等   |  |
| 采集时间*      |  | 样本采集时间，按照年-月-日“yyyy-mm-dd”规范记录，对无法精确日期的样本可精确到月份或年份。例如：2017-01-23，2017-01 和 2017。                                |  |
| 采集地点*      |  | 采集地点信息需包括国家和省份信息。 例如：中国广东省。   |  |

注：

- 1) 粪便样品的采集及运输需遵守国家相关规定。
- 2) 样本信息采集标准参考国际基因组学标准联盟（Genomic standards consortium, <http://gensc.org/>）相关标准。
- 3) “\*”栏目为必填项。

**附录B**  
**(资料性附录)**  
**标准样品宏基因组检测**

**B.1 标准样品的选择**

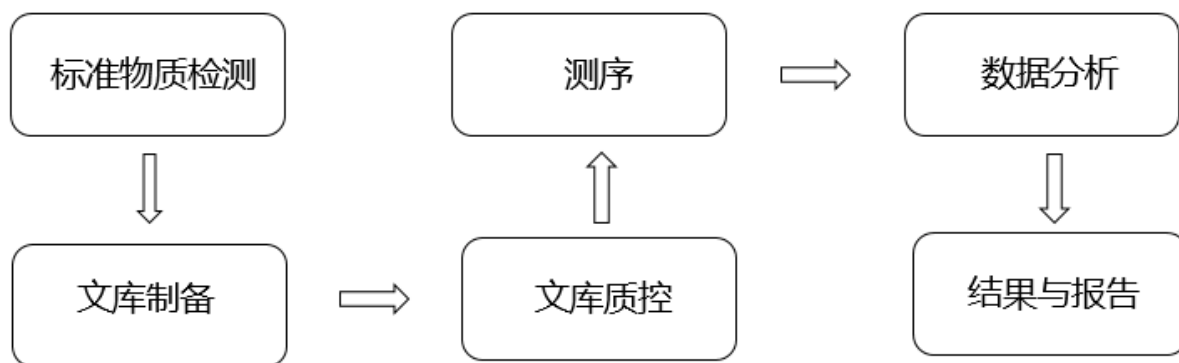
标准样品宜采用来源稳定、可溯源且参考序列信息已知的DNA混合物样品。本标准中所使用的测序标准样品为已知菌种提取基因组DNA、按特定比例混合而成的DNA混合物，菌种组成包含但不限于表B.1所示。

表B.1 测序标准样品组成

| 中文菌名    | 拉丁学名                             | 菌株号       | 革兰氏染色 | GC%  |
|---------|----------------------------------|-----------|-------|------|
| 罗伊氏乳酸杆菌 | <i>Lactobacillus reuteri</i>     | JCM 1112  | 阳性    | 38.9 |
| 暂无      | <i>Dorea longicatena</i>         | JCM 11232 | 阳性    | 41.4 |
| 暂无      | <i>Roseburia inulinivorans</i>   | DSM 16841 | 阴性    | 41.9 |
| 暂无      | <i>Roseburia hominis</i>         | DSM16839  | 阴性    | 48.5 |
| 大肠埃希氏菌  | <i>Escherichia coli</i>          | MG 1655   | 阴性    | 50.8 |
| 肠棕气单胞菌  | <i>Aeromonas enteropelogenes</i> | JCM 8355  | 阴性    | 59.7 |
| 柯林斯菌    | <i>Collinsella aerofaciens</i>   | JCM 10188 | 阳性    | 60.5 |

**B.2 标准物质宏基因组检测技术路线**

图B.1 标准样品宏基因组检测操作规程



注：数据分析时使用如下方法

- 1) 过滤低质量短序列：将片段全准确率在 90%以上且长度大于或等于 30bp 的短序列作为高质量短序列予以保留；
- 2) 去除人基因组序列：用软件 soap2(version 2.21)或 bowtie2(version 2.2.9) 将高质量短序列与人基因组参考序列 Hg19 比对，将与人基因组序列同源性≥90%以上的短序列作为人源序列去除，soap2(version 2.21) 软件关键参数：-M=4 -l=30 -v=7 -r=1 -c=0.9；bowtie2(version 2.2.9)软件使用--very-sensitive 选项，其余参数使用默认模式，每条短序列仅输出一条最优匹配；
- 3) 用软件 soap2(version 2.21) 或 bowtie2(version 2.2.9) 将去除人基因组序列后的非人源高质量短序列与标准物质 7 个物种的基因组参考序列比对，soap2(version 2.21)关键参数：-M=4, -l=30, -v=5, -r=1, -c=0.95；bowtie2(version 2.2.9) 软件参数使用默认模式，每条短序列仅输出一条最优匹配；
- 4) 相对丰度的计算方法：比对到单个物种的短序列数目占比对上标准物质中任一物种基因组的所有短序列数目的比例。

附录C  
(资料性附录)  
标准样品宏基因组检测报告

|   |        |  |  |        |  |  |
|---|--------|--|--|--------|--|--|
| 基本信息  | 检测机构   |  |  |        |  |  |
|   | 联系地址   |  |  |        |  |  |
|   | 联系方式   |  |  |        |  |  |
|   | 样品接收时间 |  |  | 样品检测时间 |  |  |
|   | 样品检测员  |  |  | 报告时间   |  |  |
| <div style="display: flex; justify-content: space-between;"> <span style="color: red;">检测指标</span> <span style="color: red;">样品名称</span> </div> |        |  |  |        |  |  |
| 测序平台  |        |  |  |        |  |  |
| 建库方法  |        |  |  |        |  |  |
| 测序策略  |        |  |  |        |  |  |
| 测序短序列数目 (M)   |        |  |  |        |  |  |
| 测序短序列碱基 (G)   |        |  |  |        |  |  |
| Q20   |        |  |  |        |  |  |
| Q30   |        |  |  |        |  |  |
| 比对率 <sup>1</sup>  |        |  |  |        |  |  |
| 覆盖率 <sup>2</sup>  |        |  |  |        |  |  |
| 分析策略  |        |  |  |        |  |  |

注 1: 去宿主后短序列数目比对到参考基因组上的比例;

注 2: 去宿主后短序列覆盖基因组区域的比例。

|   |                                  |      |      |      |      |      |      |
|---|----------------------------------|------|------|------|------|------|------|
| 标准样品<br><br>物种组成与定量结果                                     | 菌种 \ 样品名称                        |      |      |      |      |      |      |
|   |                                  | A(%) | δ(%) | A(%) | δ(%) | A(%) | δ(%) |
|   | <i>Lactobacillus reuteri</i>     |      |      |      |      |      |      |
|   | <i>Dorea longicatena</i>         |      |      |      |      |      |      |
|   | <i>Roseburia inulinivorans</i>   |      |      |      |      |      |      |
|   | <i>Roseburia hominis</i>         |      |      |      |      |      |      |
|   | <i>Escherichia coli</i>          |      |      |      |      |      |      |
|   | <i>Aeromonas enteropelogenes</i> |      |      |      |      |      |      |
|   | <i>Collinsella aerofaciens</i>   |      |      |      |      |      |      |
| 注：A-测序相对丰度；δ-测序相对丰度与理论相对丰度偏差，计算公式为（测序相对丰度-理论相对丰度）/理论相对丰度。 |                                  |      |      |      |      |      |      |
| 结果判定  |                                  |      |      |      |      |      |      |

## 参 考 文 献

1. Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. & Goodman, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **5**, R245–R249 (1998).
2. Lauber, C. L., Zhou, N., Gordon, J. I. & Knight, R. Effect of storage conditions on the assessment of bacterial community structure in soil and humann-associated samples. *FEMS Microbiol. Rev.* **307**, 80–86 (2010).
3. Balzola, F., Bernstein, C., Ho, G. T. & Lees, C. A human gut microbial gene catalogue established by metagenomic sequencing: Commentary. *Inflamm. Bowel Dis. Monit.* **11**, 28 (2010).
4. Junjie Qin, Yingrui Li, Zhiming Cai, Shenghui Li, Jianfeng Zhu, Fan Zhang, Suisha Liang, Wenwei Zhang, Yuanlin Guan, Dongqian Shen, Yangqing Peng, Dongya Zhang, Zhuye Jie, Wenxian Wu, Youwen Qin, Wenbin Xue, Junhua Li, Lingchuan Han, Donghui Lu, Peixian W, S. Z. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
5. Carroll, I. M., Ringel-Kulka, T., Siddle, J. P., Klaenhammer, T. R. & Ringel, Y. Characterization of the Fecal Microbiota Using High-Throughput Sequencing Reveals a Stable Microbial Community during Storage. *PLoS One* **7**, 1–7 (2012).
6. Hale, V. L. *et al.* Effects of field conditions on fecal microbiota. *J. Microbiol. Methods* **130**, 180–188 (2016).
7. Karatza, E., Vertzoni, M., Muenster, U. & Reppas, C. The Impact of Handling and Storage of Human Fecal Material on Bacterial Activity. *J. Pharm. Sci.* **105**, 3458–3461 (2016).
8. Song, S. J., Amir, A., Metcalf, J. L. & Amato, K. R. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems.* **1**, 1–12 (2016).
9. Blekhman, R. *et al.* Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data. *Sci. Rep.* 1–5 (2016).
10. Methé, B. a. *et al.* A framework for human microbiome research. *Nature* **486**, 215–221 (2012).