

# SCIENTIFIC REPORTS

**OPEN**

## Sample storage conditions significantly influence faecal microbiome profiles

Received: 24 July 2015

Accepted: 13 October 2015

Published: 17 November 2015

Jocelyn M Choo<sup>1,\*</sup>, Lex EX Leong<sup>1,\*</sup> & Geraint B Rogers<sup>1,2</sup>

Sequencing-based studies of the human faecal microbiota are increasingly common. Appropriate storage of sample material is essential to avoid the introduction of post-collection bias in microbial community composition. Rapid freezing to  $-80^{\circ}\text{C}$  is commonly considered to be best-practice. However, this is not feasible in many studies, particularly those involving sample collection in participants' homes. We determined the extent to which a range of stabilisation and storage strategies maintained the composition of faecal microbial community structure relative to freezing to  $-80^{\circ}\text{C}$ . Refrigeration at  $4^{\circ}\text{C}$ , storage at ambient temperature, and the use of several common preservative buffers (RNAlater, OMNIGene.GUT, Tris-EDTA) were assessed relative to freezing. Following 72 hours of storage, faecal microbial composition was assessed by 16S rRNA amplicon sequencing. Refrigeration was associated with no significant alteration in faecal microbiota diversity or composition. However, samples stored using other conditions showed substantial divergence compared to  $-80^{\circ}\text{C}$  control samples. Aside from refrigeration, the use of OMNIGene.GUT resulted in the least alteration, while the greatest change was seen in samples stored in Tris-EDTA buffer. The commercially available OMNIGene.GUT kit may provide an important alternative where refrigeration and cold chain transportation is not available.

The importance of the gut microbiome in human health and disease is increasingly clear. Compelling evidence now exists that bacterial colonisation of the gut plays a central role in the development and regulation of the host immunity<sup>1,2</sup>, metabolism<sup>3</sup>, and even the gut-brain axis<sup>4</sup>. Disruption of these homeostatic roles through perturbation of the normal microbial balance, a phenomenon referred to as dysbiosis, has been associated with a wide range of pathological conditions; including obesity<sup>5,6</sup>, autoimmune diseases<sup>7</sup>, chronic gastrointestinal inflammatory diseases<sup>8,9</sup>, type I and II diabetes<sup>10,11</sup> and carcinogenesis<sup>12,13</sup>.

Much of the recent insight gained into host-microbiome interactions has been derived through the application of next generation sequencing approaches. In particular, amplicon sequencing targeted to the 16S ribosomal RNA gene provides a relatively inexpensive and rapid characterisation of microbial communities<sup>14</sup>. When applied to faecal material, this approach represents an effective and non-invasive means to assess the gut microbiology, and has been applied across a wide range of research contexts. The ability to detect fine-scale changes in microbiota composition makes 16S gene sequencing useful in assessing changes in longitudinal sample collections, or relatively minor differences between clinically-defined groups. However, the inherent sensitivity of the approach also makes amplicon sequencing particularly susceptible to biases introduced through inappropriate sample handling or storage. For example, faecal microbiota sequencing profiles have been shown to change significantly during ambient temperature storage after 48 hours<sup>15,16</sup>. Where performing nucleic acid extraction on fresh samples immediately after collection is impractical, freezing and storage at  $-80^{\circ}\text{C}$  is widely considered to be best practice when preserving microbial composition for sequence-based analysis<sup>17-19</sup>.

<sup>1</sup>Infection and Immunity Theme, South Australia Health and Medical Research Institute, North Terrace, Adelaide 5000, SA, Australia. <sup>2</sup>School of Medicine, Flinders University, Bedford Park, Adelaide 5042, SA, Australia. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to G.R. (email: geraint.rogers@sahmri.com)

Rapid freezing and storage may not always be possible. In many cases, study subjects are requested to collect samples at home. In others, sampling may take place at remote sites where low temperature storage is unavailable. Even where cold temperature storage is available in the form of domestic freezers, the placement of faecal samples in appliances used for food storage is unacceptable to many, and may represent a potential health risk. The cold-chain transfer of samples to processing sites presents additional challenges particularly in the case of large-scale studies, where costs can be prohibitive.

In response, several studies have set out to assess the influence of different sample stabilisation strategies other than ultra-low temperature freezing on the microbial composition of faecal samples. Some effective approaches, such as the immersion of faecal material in ethanol<sup>20</sup>, may preclude many modes of sample transportation, including post and air-freight. Commercial reagents designed for other purposes, such as the RNA preservative, RNAlater, are commonly used as a general preservative, but have been shown to perform relatively poorly in maintaining microbiota composition<sup>20,21</sup>. Buffers containing EDTA can inhibit the growth of certain bacteria at concentrations used for sample preservation<sup>22,23</sup>, but their impact on sequencing-based microbiota profiling is not well characterised. In addition, OMNIgene.GUT DNA Stabilisation Kit (DNA Genotek), a commercially available approach to temporary preservation of samples at ambient temperature, has recently become available, but its effectiveness relative to other approaches has not been independently assessed. The current absence of data supporting an approach that is appropriate for home faecal sample collection, and non-specialist sample transportation, is therefore a problem for researchers when designing clinical studies with remote collection components.

Our study assessed the extent to which a number of approaches to faecal sample stabilisation could maintain microbiota composition relative to immediate freezing at  $-80^{\circ}\text{C}$ , as determined by 16S rRNA gene sequencing. The approaches evaluated in this study include dry storage at  $4^{\circ}\text{C}$  (refrigeration) and at ambient temperature, and in suspension in Tris-EDTA (TE) buffer, RNAlater, and OMNIgene.GUT. To minimise the potential of inter-subject variation as a confounding factor<sup>24</sup>, repeat faecal samples obtained from a single subject were analysed. Samples were incubated for 72 hours, a period typical of that associated with the transfer of collected samples to a processing laboratory.

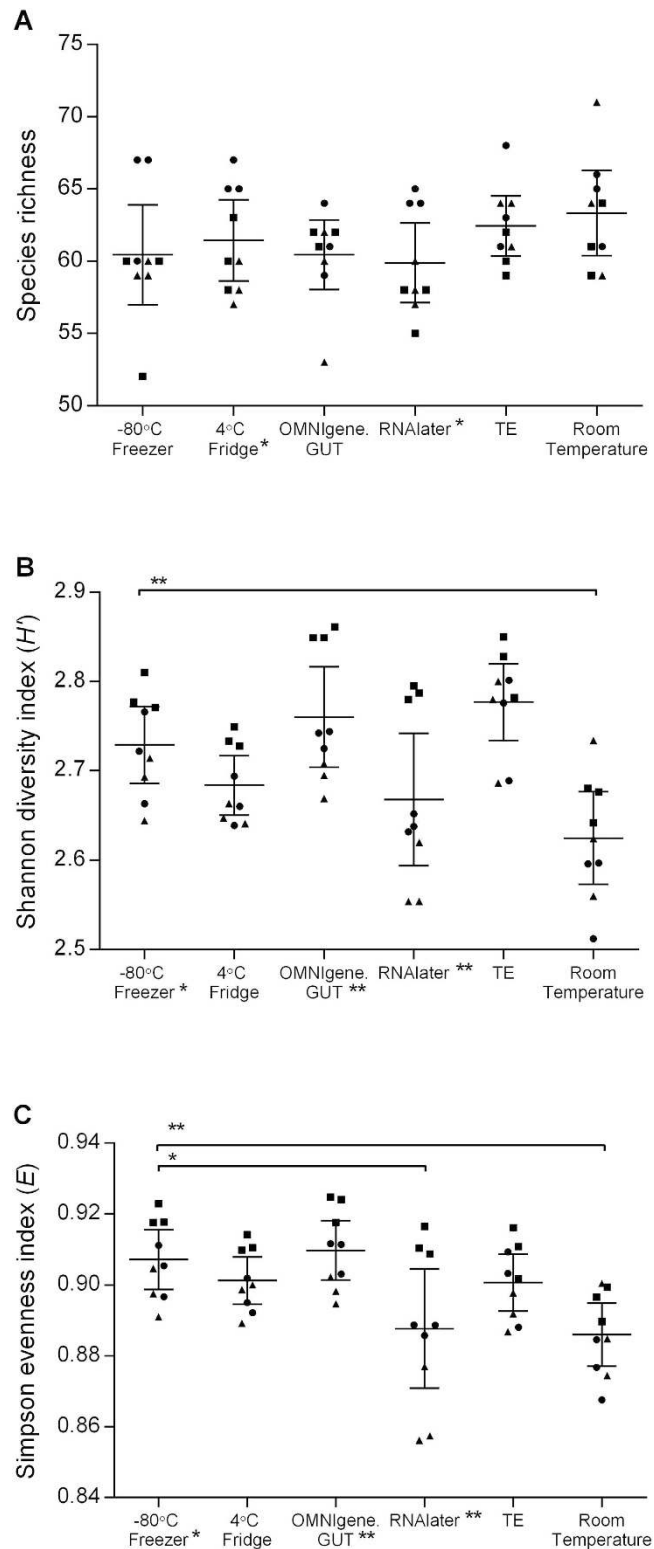
## Results

**Faecal samples alpha diversity.** A total of 1,183,513 16S rRNA sequence reads were obtained following quality filtering, equating to  $21,917 \pm 745$  (mean  $\pm$  SEM) reads per sample. After the removal of chimeras (4.83%), the number of mapped sequence reads ranged from 10,409 to 33,943 (mean  $\pm$  SEM of  $20,612 \pm 701$ ) (Supplementary Table 1).

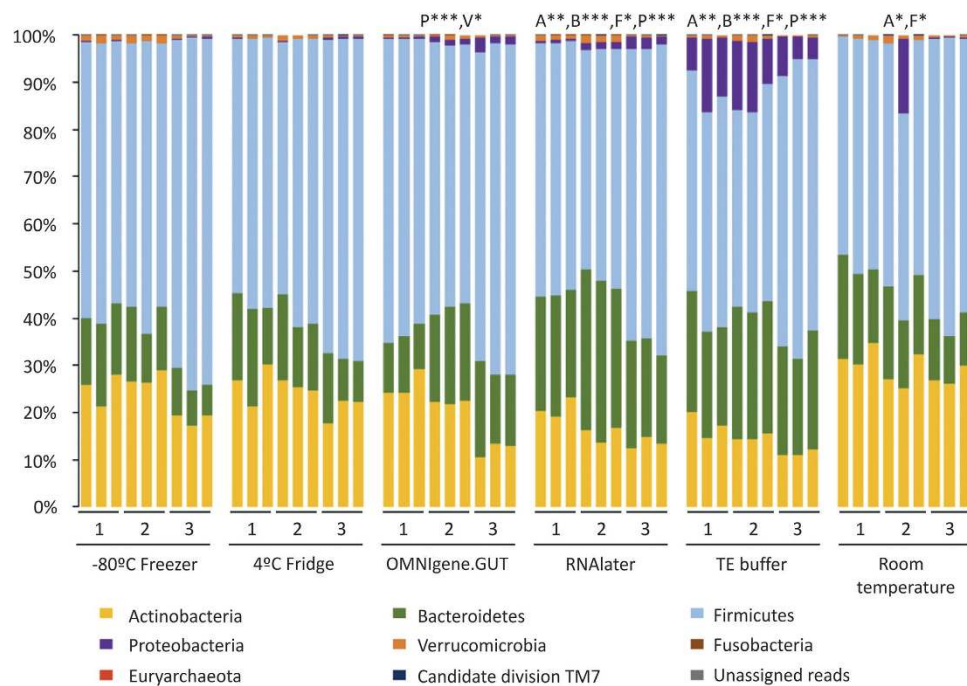
The effect of storage conditions on faecal microbiota alpha diversity was assessed based on OTU richness (measured based on the absolute number of taxa), diversity (Shannon  $H'$ ) and evenness (Simpson  $E$ ). OTU richness did not differ significantly between storage conditions (Fig. 1A). However, samples stored at room temperature had significantly lower Shannon diversity ( $p = 0.004$ , Mann-Whitney  $U$  test) and evenness ( $p = 0.002$ ) compared to those stored at  $-80^{\circ}\text{C}$  (Fig. 1B,C). Additionally, samples stored in RNAlater had lower evenness compared to  $-80^{\circ}\text{C}$  controls ( $p = 0.031$ ), whereby a considerable decrease in evenness was observed, particularly for samples in two out of the three collections. Therefore, variation between samples stored using each methodology was also assessed between each of the three collection groups. Significant differences in OTU richness were observed between different collections in RNAlater and at room temperature ( $p = 0.025$ , rank-based Kruskal-Wallis test). The diversity and evenness also significantly varied across sample groups that were stored at  $-80^{\circ}\text{C}$  ( $p = 0.025$ ), in OMNIgene.GUT ( $p = 0.004$ ) and in RNAlater ( $p = 0.004$ ).

**Phylum relative abundance.** The dominant phyla across all storage conditions were Firmicutes (median relative abundance, 57.2%; IQR 50.6%, 72.9%), Actinobacteria (21.4%; IQR 15.4%, 26.7%) and Bacteroidetes (18.1%; IQR 13.0%, 22.6%), while a substantial proportion of Proteobacteria was also observed in samples stored in TE solution (9.42%; IQR 5.9%, 1.5%) (Fig. 2). The relative abundance of these four phyla differed significantly when samples were stored in either TE buffer ( $p < 0.05$ , Mann-Whitney  $U$  test) or in RNAlater as compared to those at  $-80^{\circ}\text{C}$  (Table 1). A significant increase in the relative abundances of Actinobacteria ( $5.6\% \pm 2.5\%$ ) ( $p = 0.014$ ) and Firmicutes ( $10.5\% \pm 0.9\%$ ) ( $p = 0.019$ ) were also observed in room temperature samples, and Proteobacteria in OMNIgene.GUT samples ( $1.0\% \pm 0.4\%$ ) ( $p = 0.0002$ ). No significant variation was detected between the samples stored at  $4^{\circ}\text{C}$  and  $-80^{\circ}\text{C}$ .

**Microbiota ordination and taxon distribution.** Principal Coordinate Analysis (PCoA) of Bray-Curtis distances revealed that the samples formed two distinct clusters according to sample collection, with samples in collection 1 and 2 forming one cluster and samples in collection 3 in the other (Fig. 3). Within the sample cluster of collection 1 and 2, samples were separated along the  $y$ -axis based on the storage conditions, and a similar observation was noted in collection 3. Consistent with these findings, a PERMANOVA (Permutational multivariate analysis of variance) test on the Bray-Curtis dissimilarity matrices indicated that both the collection group ( $p < 0.0001$ ) and the storage conditions ( $p < 0.0001$ ) contributed significantly to the differences in microbial composition of the samples. However, the higher pseudo- $F$  score of the collection group suggested that differences between the samples has greater effect on the changes observed than differences due to storage conditions (Pseudo- $F$ ,



**Figure 1. Species diversity following incubation under six different storage conditions.** The extent of microbiota structural and composition diversities were measured using (A) Taxa S (species richness), (B) Shannon-Weiner diversity index, (C) Simpson's evenness index. Each point represents the diversity score for a replicate from collection 1 (●), collection 2 (▲) or collection 3 (■). Error bars represent SEM. Within-group and between-group variations were measured using Kruskal-Wallis one-way ANOVA and Mann-Whitney *U*-test, respectively. Significant variance is indicated by asterisks; single asterisk (\*) indicates  $p \leq 0.05$ , double asterisk (\*\*) represents  $p \leq 0.01$ .



**Figure 2. Relative abundance at phylum level for each sample incubated under six different storage conditions.** Storage conditions that differed significantly from the control ( $-80^{\circ}\text{C}$ ) are indicated with respective phylum abbreviation and asterisks. The respective phyla were abbreviated as follow: Actinobacteria (A), Bacteroidetes (B), Firmicutes (F), Proteobacteria (P) and Verrucomicrobia (V). Statistical significance was assessed by Mann-Whitney  $U$ -test and indicated by asterisks; single asterisk (\*) represents  $p \leq 0.05$ , double asterisk (\*\*) represents  $p \leq 0.01$ , and triple asterisk (\*\*\*) represents  $p \leq 0.001$ .

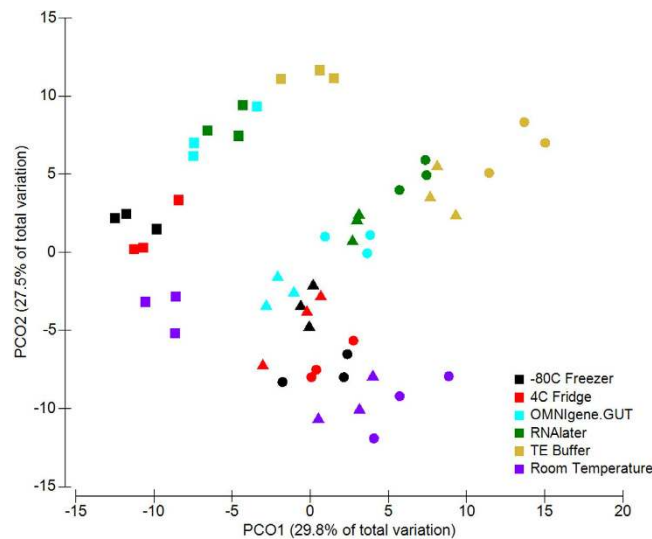
Phylum	Difference in mean relative abundance $\pm$ standard error of mean				
	$-80^{\circ}\text{C}$ vs $4^{\circ}\text{C}$ fridge	$-80^{\circ}\text{C}$ vs OMNIgene.GUT	$-80^{\circ}\text{C}$ vs RNA later	$-80^{\circ}\text{C}$ vs TE buffer	$-80^{\circ}\text{C}$ vs RT
Firmicutes	$2.4 \pm 1.4$	$3.6 \pm 0.8$	$7.8 \pm 1.6$	$12.7 \pm 1.2$	$10.5 \pm 0.9$
Bacteroidetes	$2.0 \pm 0.5$	$6.9 \pm 1.1$	$13.6 \pm 3.2$	$12.2 \pm 2.4$	$3.6 \pm 0.2$
Actinobacteria	$1.7 \pm 0.3$	$4.1 \pm 1.7$	$7.0 \pm 2.4$	$9.2 \pm 1.7$	$5.6 \pm 2.5$
Proteobacteria	$0.04 \pm 0.02$	$1.0 \pm 0.4$	$1.3 \pm 0.4$	$10.0 \pm 2.3$	$1.8 \pm 1.8$

**Table 1. Mean difference in the relative abundance of the phyla Firmicutes Bacteroidetes, Actinobacteria and Proteobacteria in different storage conditions compared to  $-80^{\circ}\text{C}$ .**

collection = 34.65; preservation method = 18.24) (PERMANOVA  $p = 0.0001$ ) (Table 2). These results indicated that the faecal microbiota of the different collection groups were highly heterogeneous; variation that could influence the observed differences between the storage conditions.

Pairwise PERMANOVA tests between each of the assessed storage conditions revealed that only samples stored at  $4^{\circ}\text{C}$  did not differ significantly in composition compared to those stored at  $-80^{\circ}\text{C}$  (Table 2). Samples from the remaining storage approaches each showed significant divergence to  $-80^{\circ}\text{C}$  controls (PERMANOVA  $p = 0.0001$ ), with OMNIgene.GUT ( $t = 2.9592$ ) having the least variation among these methods ( $t > 3.3$ ). In addition, microbial composition also differed significantly between samples stored at  $4^{\circ}\text{C}$ , using OMNIgene.GUT, in RNeasy, in TE buffer, and dry at room temperature (PERMANOVA  $p < 0.0003$ ) (Supplementary Table 2). Samples stored dry at room temperature ( $t = 3.0168$ ) and using OMNIgene.GUT ( $t = 3.1518$ ) differed least from those stored at  $4^{\circ}\text{C}$ , while the greatest shifts were observed in samples stored in TE buffer ( $t > 4.4647$ ), consistent with the trend observed when samples were compared to  $-80^{\circ}\text{C}$  storage.

A heat map based on the OTU abundance of the major genera (contributing approximately 86% of the variation as determined by SIMPER analysis) was generated to identify microbial composition that contributed to the variation between samples (Fig. 4). Consistent with the pairwise comparison analysis, changes were observed in the abundance of only a single taxa for the refrigerated samples (*Anaerostipes*,



**Figure 3. Clustering of samples due to storage conditions by PCoA, based on Bray-Curtis similarity distance.** The first two principal coordinates are plotted on the  $x$ - and  $y$ -axes, respectively (representing 57.3% of the total variation). Faecal collections sampled at three different time points are represented by circle (●) for collection 1, triangle (▲) for collection 2, and square (■) for collection 3. Storage conditions are indicated by colour.

Source	df	SS	MS	Pseudo- $F$	P(perm)
Storage	5	3184.9	636.99	18.236	0.0001
Collections	2	2420.9	1210.4	34.653	0.0001
Storage x Collections	10	948.13	94.813	2.7144	0.0001
Residual	36	1257.5	34.93		
Total	53	7811.4			
Comparison				$t$	P(perm)
−80°C Freezer vs 4°C Fridge				1.7381	0.0491
−80°C Freezer vs OMNIgene.GUT				2.9592	0.001
−80°C Freezer vs RNAlater				4.6425	0.001
−80°C Freezer vs TE Buffer				5.6022	0.001
−80°C Freezer Room Temperature				3.3297	0.001

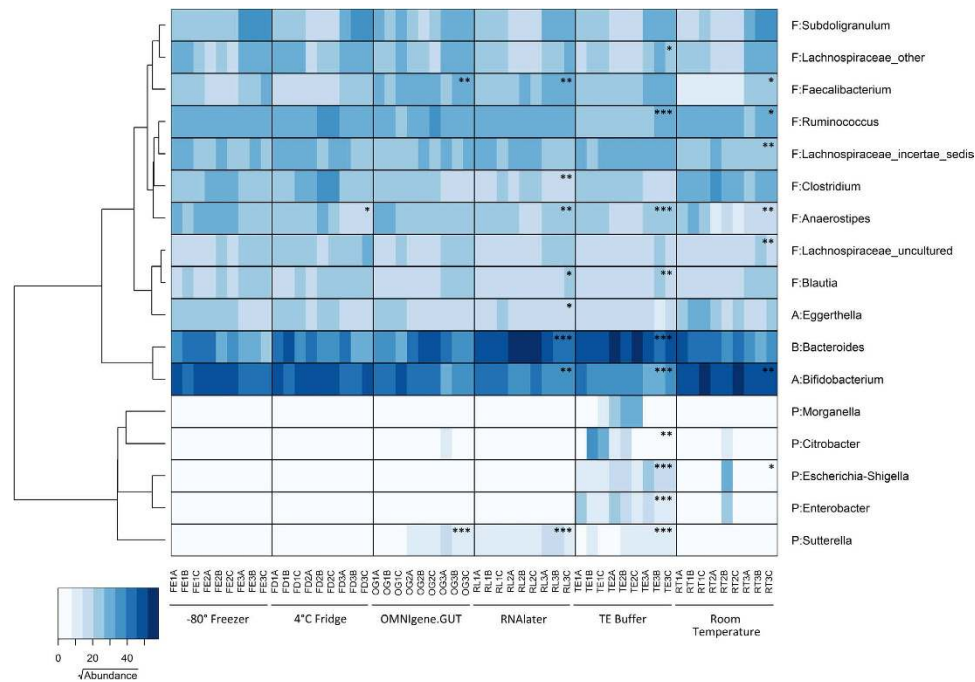
**Table 2. PERMANOVA analysis on Bray-Curtis distance for faecal microbiota structure on the genera level from three different collections and six different storage conditions (permutations = 9999).**

df = degree of freedom; SS = sum of squares; MS = mean sum of squares; Pseudo- $F$  =  $F$  value by permutation;  $t$  =  $t$ -statistic.

$p = 0.024$ ) and two taxa for the OMNIgene.GUT samples (*Sutterella*,  $p < 0.0001$ ; *Faecalibacterium*,  $p = 0.001$ ). Several taxa in the samples stored in RNAlater, in TE buffer and at room temperature were also significantly different in abundance to the  $-80^\circ\text{C}$  samples. Specifically, a decreased abundance of the genus *Anaerostipes* ( $p < 0.001$ ) was observed in all three conditions. The abundance of *Bifidobacterium* was lower in the RNAlater ( $p = 0.004$ ) and TE buffer ( $p < 0.001$ ) sample groups, but higher in room temperature samples ( $p = 0.004$ ). The RNAlater and TE buffer samples also had an increased abundance of *Bacteroides* ( $p < 0.001$ ).

## Discussion

In recent years, there has been increasing interest in understanding the relationship between the gut microbiota and human health. Investigating this relationship has been greatly facilitated by next-generation sequencing<sup>4,11</sup>. However, the application of these gene-based approaches can only yield useful insight if the material analysed accurately represents the microbial community at the site sampled. By assessing the impact of a number of commonly used storage conditions on sequencing-based, faecal microbial profiles, we provide an empirical basis for future study design. We used storage at  $-80^\circ\text{C}$  as a control based on



**Figure 4. The distribution of the major genera in samples incubated under six different storage conditions.** The heatmap shows square root-transformed read counts for the 17 taxa identified as contributing most to variance (86% across all samples) as determined by SIMPER analysis. The dendrogram shows the similarity relationship of genera based on Bray-Curtis distance and Ward's hierarchical clustering method. Phyla are abbreviated as follows: Actinobacteria (A), Bacteroidetes (B), Firmicutes (F) and Proteobacteria (P). Storage conditions that differ significantly from the control ( $-80^{\circ}\text{C}$ ) by Mann-Whitney  $U$ -test are indicated with asterisks; single asterisk (\*) represents  $p \leq 0.05$ , double asterisk (\*\*) represents  $p \leq 0.01$ , and triple asterisk (\*\*\*) represents  $p \leq 0.001$ .

previous studies that reported no significant variation between fresh samples and those stabilised in this way<sup>18,20</sup>.

Substantial interpersonal variation in gut microbiota composition results from differences in factors such as diet<sup>25</sup> and genetic background<sup>26</sup>. This variation makes discerning the impact of sample storage more difficult when material from multiple individuals is included. For this reason, repeat samples collected from a single subject were used. However, intra-individual variation in sample composition also occurs, and for this reason, we assessed variation between serial samples that were stored at  $-80^{\circ}\text{C}$ . These samples were found to have significant variation in microbial diversity and evenness. PERMANOVA analysis suggested that these variations resulted from differences in sample composition at the different collection timepoints. However, since storage conditions were assessed based on portions of the same starting material, such variation would not be a source of bias.

Freezing has been previously shown to result in a higher ratio of Firmicutes to Bacteroidetes due to changes to the cellular structure of the Gram-positive bacteria when frozen<sup>18,19</sup>; a potential confounder given the use of these conditions in our control group. However, all samples in this study were frozen at  $-80^{\circ}\text{C}$  following incubation at the specified conditions. Further, samples were subjected to stringent lysis conditions, including physical disruption, to prevent bias arising from any effects of freeze-thawing.

Our finding that the microbial composition of refrigerated ( $4^{\circ}\text{C}$ ) samples did not differ significantly to that of controls stored at  $-80^{\circ}\text{C}$  is in keeping with a previous study performed over a 24 hour period<sup>27</sup>. However, perhaps unsurprisingly, the composition of samples stored at room temperature changed substantially. These changes principally represented increases in the relative abundance of *Bifidobacterium* species, belonging to the phylum Actinobacteria, and decreases in the relative abundance of *Anaerostipes*, *Ruminococcus*, *Faecalibacterium* and Lachnospiraceae species, belonging to the phylum Firmicutes. These findings are concordant with a report by Roesch and colleagues<sup>16</sup> in which they identified a similar decline in Firmicutes bacteria in samples stored at room temperature for 72 hours. The observed shifts in bacterial community structure were also reflected in measures of diversity compared to the  $-80^{\circ}\text{C}$  controls, consistent with the overgrowth of a limited number of species. A number of other prior studies have reported no significant variation in the composition of faecal samples stored at room temperature<sup>15,21</sup>. However, it is noteworthy that these studies did not report substantial actinobacterial populations in their subject groups (<5% relative abundance), with greater relative abundances of Bacteroidetes or Firmicutes.

Among the samples stored in stabilisation reagents, the microbiota profiles of those in OMNIgene.GUT diverged least from the  $-80^{\circ}\text{C}$  controls, followed by those in RNAlater, and those in TE buffer (as reflected in the t-statistic of the pair-wise PERMANOVA analysis). The two taxa whose abundance changed significantly in samples stored in OMNIgene.GUT were the *Sutterella* and *Faecalibacterium* genera. This is concerning given the many reports of associations between these genera and a range of gastrointestinal diseases<sup>28–31</sup>. In samples stored in RNAlater and TE buffer, there was an increase in the Bacteroides population, and a decrease in Firmicutes (*Blautia*), and Actinobacteria (*Bifidobacterium*). In addition, except for the OMNIgene.GUT group, the relative abundance of *Anaerostipes* differed significantly in all storage conditions compared to the  $-80^{\circ}\text{C}$  controls, suggesting that this taxon is particularly sensitive to changes in storage conditions.

Interestingly, we noted an increase in OTU relative abundance for facultative anaerobic Proteobacteria, including *Escherichia-Shigella*, *Citrobacter* and *Enterobacter*, in samples stored in TE buffer. The increase in abundance of the facultative anaerobe, Proteobacteria, may be due to exposure of the faecal material to aerobic conditions when suspended in the buffer. Furthermore, these changes (an increase abundance of Proteobacteria and a reduction of Firmicutes and Actinobacteria) resemble those associated with some types of diet-induced gut dysbiosis<sup>32–34</sup>, and as such, we suggest that the use of TE buffer for sample stabilisation should only be considered in combination with low temperature storage to prevent bacterial proliferation.

Samples stored in RNAlater have been shown previously to have decreased DNA purity<sup>21</sup> and lower DNA extraction yields<sup>20</sup>, potentially affecting downstream analysis. In addition, these studies have reported storage in RNAlater results in significantly lower<sup>21</sup> or higher<sup>20</sup> microbial diversity than  $-80^{\circ}\text{C}$  samples (a discrepancy that may be due to differences in analyses being performed on human and non-human primate samples, respectively). Our findings are in keeping with these previous investigations, with room temperature storage in RNAlater resulting in significantly reduced Shannon diversity.

Limitations of our study included the analysis of samples from a single subject, a strategy used to control for inter-subject variation. However, as the impact of storage conditions may vary between individuals, our work could be extended to samples from a larger subject population, with differing gut microbial communities. Further, as new faecal sample stabilisation strategies emerge, it will be important to determine their efficacy relative to the approaches described here.

## Conclusions

Our study highlights several important considerations for studies involving microbiome-wide analysis of faecal samples. Storage conditions have the potential to introduce substantial alterations to microbial community profiling based on 16S rRNA gene sequencing. Where ultra-low temperature storage is unavailable, changes to microbial composition can be minimized by storage and transportation of samples at  $4^{\circ}\text{C}$ . Changes in microbiota composition in unstabilised samples may explain incongruent findings from previous sequencing-based faecal studies. Significant intra-individual variation occurs in the composition of faecal samples from individual subjects, which is greater than the variation observed between the different storage methods. These findings underline the importance of high frequency, longitudinal faecal collection to understand the temporal dynamics of microbial communities.

## Methods

**Faecal sample collection and processing.** Faecal samples were collected from a single healthy volunteer ( $n=1$ ) on three different occasions over a three week period with written informed consent. The study protocol was approved by, and carried out in accordance with the guidelines of, the South Australian Health and Medical Research Institute Internal Review Committee. There had been no antibiotic exposure during the collection period, or in the 30 days prior to collection. Each sample was homogenized, and three aliquots stored using each of the six test conditions: freezing at  $-80^{\circ}\text{C}$ , refrigeration at  $4^{\circ}\text{C}$ , resuspension in a  $7\times$  volume of TE buffer (10 mM Tris-HCl pH 8.0, 100 mM EDTA), resuspension in a  $7\times$  volume of RNAlater (Ambion, Austin, USA), use of an OMNIgene.GUT stabilisation kit (DNA Genotek, Ontario, Canada), or storage in a dry collection tube. With the exception of samples stored using the OMNIgene.GUT stabilisation kit, all samples were placed in a sterile 2 mL screwcap tube. TE buffer, RNA later and OMNIgene.GUT samples were stored at ambient temperature. All samples were stored in their respective conditions for 72 hours and then frozen at  $-80^{\circ}\text{C}$  prior to DNA extraction, to control for any impact that freezing might have on bacterial community composition.

**DNA extraction and quantification.** DNA extraction was performed using MoBio Powerlyzer Powersoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, USA). Samples that were immersed in solution were centrifuged at  $13,000\times g$  for 5 min and the supernatant discarded. Samples were washed with 1 mL of cold  $1\times$  PBS (pH7.2) (Life Technologies, Melbourne, Australia) and centrifuged at  $13,000\times g$  for 10 mins. DNA extraction was performed according to the manufacturer's instructions with the following modifications. Samples were placed into bead tubes with solution C1 and heated at  $65^{\circ}\text{C}$  for 10 min, prior to two cycles of bead beating at 6.5 m/s for 1 min using a FastPrep-24 bead beater (MP Biomedicals, Santa Ana, USA). Total DNA was eluted in 100  $\mu\text{L}$  of sterile water. DNA concentration was quantified fluorometrically with a Qubit dsDNA HS Assay kit (Life Technologies).

**16S rRNA gene sequencing.** The V4 hypervariable region of the bacterial 16S rRNA gene was amplified from faecal DNA extracts using modified universal bacterial primer pairs 515F (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACHVGGGTWTCTAAT-3'), with Illumina adapter overhang sequences as indicated by underline. Amplicons were generated, cleaned, indexed and sequenced according to the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation protocol ([http://support.illumina.com/downloads/16s\\_metagenomic\\_sequencing\\_library\\_preparation.html](http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html)) with certain modifications. Briefly, an initial PCR reaction contained at least 12.5 ng of DNA, 5 µL of forward primer (1 µM), 5 µL of reverse primer (1 µM) and 12.5 µL of 2× KAPA HiFi Hotstart ReadyMix (KAPA Biosystems, Wilmington, MA, USA) in a total volume of 25 µL. The PCR reaction was performed on a Veriti 96-well Thermal Cycler (Life Technologies) using the following program: 95 °C for 3 min, followed by 25 cycles of 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec and a final extension step at 72 °C for 5 min. Samples were multiplexed using a dual-index approach with the Nextera XT Index kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer's instructions. The final library was paired-end sequenced at 2 × 300 bp using a MiSeq Reagent Kit v3 on the Illumina MiSeq platform. Sequencing was performed at the David R Gunn Genomics Facility, South Australian Health and Medical Research Institute.

**Bioinformatic analysis.** The Quantitative Insights Into Microbial Ecology (QIIME, v1.8.0)<sup>35</sup> software was used to analyse the 16S rRNA sequence generated from paired-end amplicon sequencing using bioinformatics pipeline as previously described<sup>36</sup>. Briefly, barcoded forward and reverse sequencing reads were quality filtered and merged using Paired-End reAd mergeR (PEAR v0.9.6)<sup>37</sup>. Chimeras were detected and filtered from the paired-end reads using USEARCH (v6.1)<sup>38</sup> against the 97% clustered representative sequences from the Greengenes database (v13.8)<sup>39</sup>. Operational taxonomic units (OTUs) were assigned to the reads using an open reference approach with UCLUST algorithm (v1.2.22q) against the SILVA database release 111 (July 2012)<sup>40</sup> that was clustered at 97% identity. During the OTU assignment, sequences were pre-clustered at 80% against the reference prior to *de novo* clustering. A minimum subsampling depth of 10,409 reads was selected based on the lowest read depth of samples. QIIME was used to generate an alpha rarefaction curve to confirm that the relationship between read depth and new taxon detection approached an asymptote in all samples at this depth. No samples were eliminated following subsampling.

**Diversity measurements and statistical analyses.** PAST (v3.04)<sup>41</sup> was used to calculate taxa abundance (Taxa\_S), Shannon's diversity index and Simpson's evenness index for each sample. Beta diversity (Bray-Curtis similarity) was determined using PRIMER (v6, PRIMER-E Ltd, Plymouth, UK). The similarity matrices based on Bray-Curtis distance were calculated using the sample-normalised, square root transformed relative OTU abundance. Principal coordinate analysis (PCoA) was used to visualize clustering of samples based on their similarity matrices. The two-factor permutational multivariate analysis of variance (PERMANOVA)<sup>42</sup> on the Bray-Curtis matrices was used to test the null hypothesis of no difference amongst a *priori*-defined groups using PERMANOVA + add-on package for PRIMER. The test was computed using unrestricted permutation of raw data with 9,999 random permutations and at a significance level of 0.01.

Variation in microbiota composition at the genus-level was assessed using a two-step approach. First, OTUs that contributed to the overall variation between all six storage conditions were determined using SIMilarity of PERcentages (SIMPER) analysis in PAST. Subsequently, the OTU abundance of the top 17 taxa (accounting for 86% of the dissimilarity between groups) were used to generate a heatmap using gplots package of R statistical software<sup>43</sup>. Hierarchical clustering of the taxa was performed on Bray-Curtis dissimilarity and clustered using Ward's method. The Kruskal-Wallis one-way ANOVA was used to evaluate the variation among samples in each storage condition. The Mann-Whitney *U* test (two-tailed) was used to compare the variation between the respective storage conditions and -80 °C. Non-parametric statistical analyses were performed using the GraphPad PRISM v6.05 (GraphPad Software Inc., California, USA).

## References

- Hansen, C. H. *et al.* Mode of delivery shapes gut colonization pattern and modulates regulatory immunity in mice. *Journal of immunology* **193**, 1213–1222, doi: 10.4049/jimmunol.1400085 (2014).
- Clarke, T. B. *et al.* Recognition of peptidoglycan from the microbiota by Nod1 enhances systemic innate immunity. *Nature medicine* **16**, 228–231, doi: 10.1038/nm.2087 (2010).
- Yano, J. M. *et al.* Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264–276, doi: 10.1016/j.cell.2015.02.047 (2015).
- Carabotti, M., Scirocco, A., Maselli, M. A. & Severi, C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of gastroenterology: quarterly publication of the Hellenic Society of Gastroenterology* **28**, 203–209 (2015).
- Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023, doi: 10.1038/4441022a (2006).
- Backhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15718–15723, doi: 10.1073/pnas.0407076101 (2004).



7. Scher, J. U. *et al.* Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife* **2**, e01202, doi: 10.7554/eLife.01202 (2013).
8. Schaubeck, M. *et al.* Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence. *Gut*, doi: 10.1136/gutjnl-2015-309333 (2015).
9. Hedin, C. *et al.* Siblings of patients with Crohn's disease exhibit a biologically relevant dysbiosis in mucosal microbial metacommunities. *Gut*, doi: 10.1136/gutjnl-2014-308896 (2015).
10. Kostic, A. D. *et al.* The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell host & microbe* **17**, 260–273, doi: 10.1016/j.chom.2015.01.001 (2015).
11. Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60, doi: 10.1038/nature11450 (2012).
12. Schulz, M. D. *et al.* High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. *Nature* **514**, 508–512, doi: 10.1038/nature13398 (2014).
13. Feng, Q. *et al.* Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nature communications* **6**, 6528, doi: 10.1038/ncomms7528 (2015).
14. Prakash, S., Rodes, L., Coussa-Charley, M. & Tomaro-Duchesneau, C. Gut microbiota: next frontier in understanding human health and development of biotherapeutics. *Biologics: targets & therapy* **5**, 71–86, doi: 10.2147/BTT.S19099 (2011).
15. 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**, e46953, doi: 10.1371/journal.pone.0046953 (2012).
16. Roesch, L. F. *et al.* Influence of fecal sample storage on bacterial community diversity. *The open microbiology journal* **3**, 40–46, doi: 10.2174/1874285800903010040 (2009).
17. Wu, G. D. *et al.* Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC microbiology* **10**, 206, doi: 10.1186/1471-2180-10-206 (2010).
18. Fouhy, F. *et al.* The effects of freezing on faecal microbiota as determined using MiSeq sequencing and culture-based investigations. *PLoS one* **10**, e0119355, doi: 10.1371/journal.pone.0119355 (2015).
19. Bahl, M. I., Bergstrom, A. & Licht, T. R. Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. *FEMS microbiology letters* **329**, 193–197, doi: 10.1111/j.1574-6968.2012.02523.x (2012).
20. Hale, V. L., Tan, C. L., Knight, R. & Amato, K. R. Effect of preservation method on spider monkey (*Ateles geoffroyi*) fecal microbiota over 8 weeks. *Journal of microbiological methods* **113**, 16–26, doi: 10.1016/j.mimet.2015.03.021 (2015).
21. Dominianni, C., Wu, J., Hayes, R. B. & Ahn, J. Comparison of methods for fecal microbiome biospecimen collection. *BMC microbiology* **14**, 103, doi: 10.1186/1471-2180-14-103 (2014).
22. Carozzi, F. M. & Sani, C. Fecal Collection and Stabilization Methods for Improved Fecal DNA Test for Colorectal Cancer in a Screening Setting. *Journal of Cancer Research*, 818675 (2013).
23. Reidmiller, J. S. *et al.* Antimicrobial properties of the chelating agent EDTA on Streptococcal bovine mastitis isolates. *Journal of food protection* **69**, 1460–1462 (2006).
24. Mackenzie, B.W., Waite, D.W. & Taylor, M. W. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Frontiers in Microbiology* **6**, 130 (2015).
25. David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome biology* **15**, R89, doi: 10.1186/gb-2014-15-7-r89 (2014).
26. Ericsson, A. C. *et al.* Effects of vendor and genetic background on the composition of the fecal microbiota of inbred mice. *PLoS one* **10**, e0116704, doi: 10.1371/journal.pone.0116704 (2015).
27. Tedjo, D. I. *et al.* The effect of sampling and storage on the fecal microbiota composition in healthy and diseased subjects. *PLoS one* **10**, e0126685, doi: 10.1371/journal.pone.0126685 (2015).
28. Sokol, H. *et al.* Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflammatory Bowel Diseases* **15**, 1183–1189 (2009).
29. Wang, L. *et al.* Increase abundance of *Sutterella* spp. and *Ruminococcus torques* in feces of children with autism spectrum disorder. *Molecular Autism* **4**, 42 (2013).
30. Willing, B. *et al.* Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflammatory Bowel Diseases* **15**, 653–660 (2009).
31. Williams, B. L., Hornig, M., Parekh, T. & Lipkin, W. I. Application of novel PCR-based methods for detection, quantification and phylogenetic characterization of *Sutterella* species in intestinal biopsy samples from children with autism and gastrointestinal disturbances. *mBio* **3** (2012).
32. Lupp, C. *et al.* Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell host & microbe* **2**, 119–129 (2007).
33. Murphy, E. F. *et al.* Divergent metabolic outcomes arising from targeted manipulation of the gut microbiota in diet-induced obesity. *Gut* **62**, 220–226 (2013).
34. Zhu, L. *et al.* Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. *Hepatology* **57**, 601–609 (2013).
35. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
36. Jervis-Bardy, J. *et al.* Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* **3**, 19, doi: 10.1186/s40168-015-0083-8 (2015).
37. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620, doi: 10.1093/bioinformatics/btt593 (2014).
38. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461, doi: 10.1093/bioinformatics/btq461 (2010).
39. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME journal* **6**, 610–618, doi: 10.1038/ismej.2011.139 (2012).
40. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590–D596 (2013).
41. Hammer, O., Harper, D. A. T. & Ryan, P. D. PAST: paleontological statistics software package for education and data analysis. *Palaentologia Electronica* **4**, 9 (2001).
42. Andersen, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**, 32–46 (2001).
43. Warnes, G. R. *et al.* Gplots: various R programming tools for plotting data, v2.12.1. <http://cran.r-project.org/web/packages/gplots/index.html> (2013).

### Author Contributions

J.C., L.L. and G.R. conceived of the study. J.C. and L.L. performed all analyses. J.C., L.L. and G.R. interpreted the data and wrote the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Choo, J. M. *et al.* Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350; doi: 10.1038/srep16350 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>