Identification of factors impacting reproducibility and quality of microbiome profile analysis

Evgueni Doukhanine, Anne Bouevitch, Lindsay Pozza, Carlos Merino and Rafal M. Iwasiow
DNA Genotek Inc., Ottawa, Canada

Introduction

The evolved characterization of human-associated microbial communities has given scientists a new perspective on diseases of the GI tract. Metagenomic profiles associated with diabetes, inflammatory bowel disorders and obesity are showing diagnostic, prognostic and monitoring potential. These microbial communities are dynamic systems that have been shown to respond to dietary, environmental, and even seasonal changes. Considering this intrinsic complexity, successful identification and validation of biomarkers requires the execution of large, longitudinal population studies. The proper execution of such studies relies on capturing an accurate representation or 'taking a snapshot' of the donor's biology at the point-of-collection and reliably reproducing it through the analytical pipeline.

Microbiome-Wide Association Studies (MWAS) require the establishment of scalable and reliable methods for the *pre-analytical* (collection and stabilization), *in vitro* (processing and assay), and *in silico* (computational analysis and interpretation) handling of biological material and resultant data. In the present study, the variability introduced in metagenomic profiles by pre-analytical and *in vitro* conditions was quantified and compared under different collection conditions including unstabilized samples or samples stabilized with OMNIgene®•GUT (OMR-200). OMNIgene•GUT is a device for fecal self-collection, stabilization and homogenization for microbiome analysis.

While driving variability in a metagenomics workflow to zero may be unachievable, our results illustrate that pre-analytical conditions—by far—have the largest impact on accuracy and reproducibility of microbiome profiling. Clinically relevant taxa, (such as *Prevotella copri*) when left vulnerable to time and temperature can overgrow, leading to spurious results. Furthermore, collection and proper stabilization with OMNIgene•GUT can greatly reduce this source of variability.

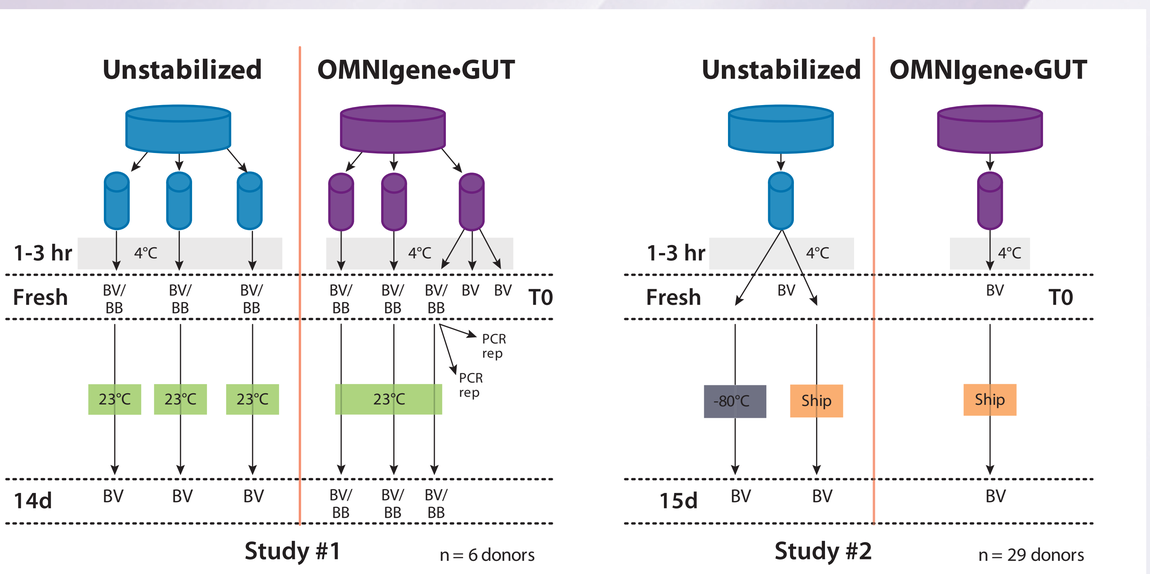
If a goal for the scientific community is to produce more informative, reproducible and scalable studies at the population-level, we propose OMNIgene•GUT as a pre-analytical standard for fecal sample collection and stabilization for metagenomic analysis.

Materials and methods

Sample collection and DNA extraction

- Donors each collected one stool specimen and sampled equal amounts into tubes without a chemistry preservative and into OMR-200 which includes a stabilizer.
- Samples were extracted or held at varying temperatures according to the study designs (below).
- DNA was extracted from samples using the PowerFecal® DNA Isolation Kit (MoBio® Laboratories) using either Bead Vortexer (BV) (MoBio® Laboratories) or Mini-Beadbeater (BB) (BioSpec®).

Study designs



16S rRNA Sequencing

- Study #1 16S rRNA V4 hypervariable region paired-end amplicon sequencing.
- Study #2 16S rRNA V3/V4 hypervariable region paired-end amplicon sequencing.
- For both studies used Illumina® MiSeq® instrument, sequences were quality filtered using QIIME and custom scripts. Paired-end reads were merged and searched against the Greengenes reference database, clustered at 97% by UCLUST. After data normalization, sample-to-sample distance was measured using Bray-Curtis distance on operational taxonomic unit (OTU) abundance data (utilizes taxon abundance differences across samples, employing a pair-wise normalization by dividing the sum of differences by the sum of all abundances).

Statistical methods utilized

- Two-sample Kolmogorov–Smirnov test was used as a nonparametric evaluation of the equality of continuous, one-dimensional probability distributions.
- Two-way ANOVA test evaluated influence of two different categorical independent variables on one continuous dependent variable.

Results – Sources of variability

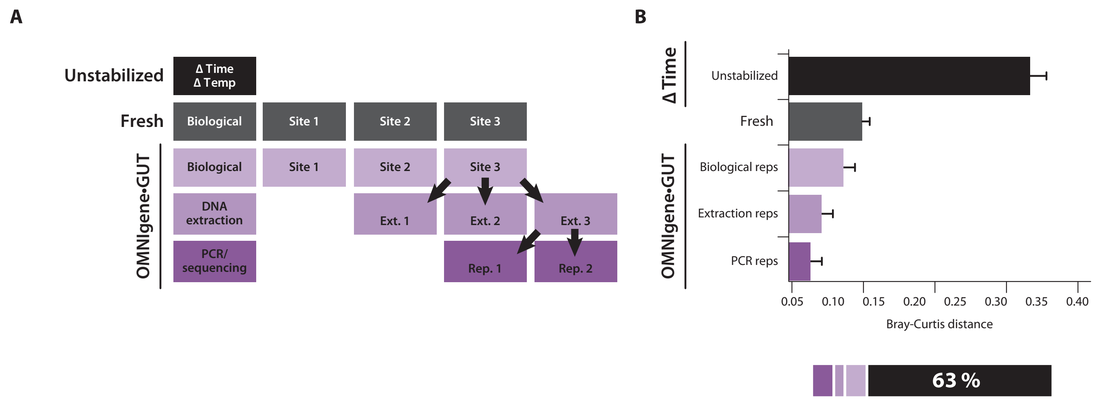


Figure 1: Sources of variability during microbiome evaluation and impact of stabilization/homogenization of samples. *(A) Diagram describes the method in attaining replicates at each potential stage of variability during a microbiome study. (B) Bray-Curtis distance comparison within each of the source of variability shows that the largest contributor is improper stabilization of collected samples. Δ Time, 14 days at 23°C.*

Results – Impact of extraction

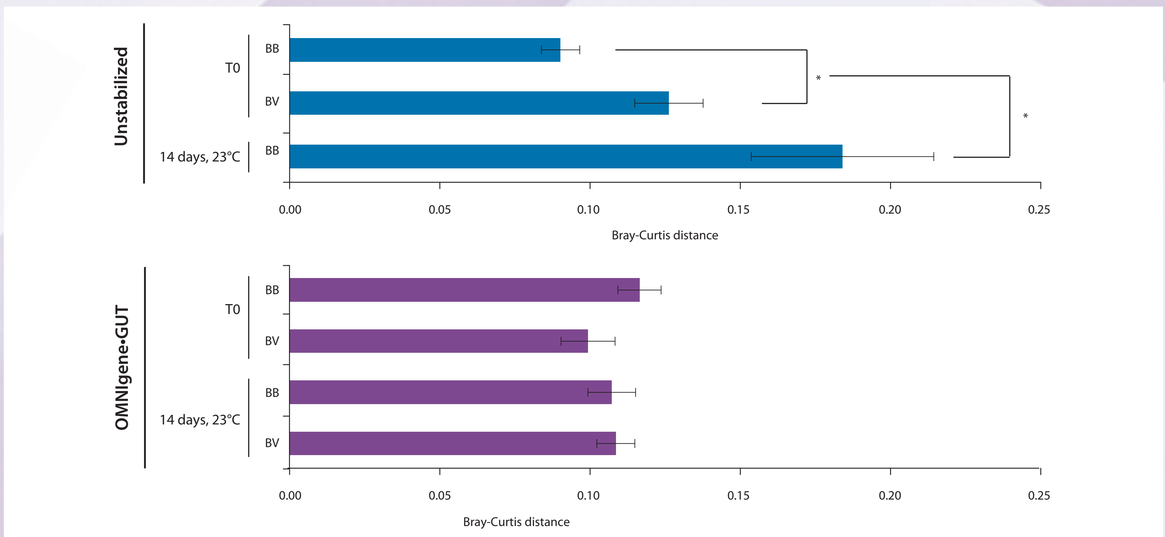


Figure 2: Differences in mechanical lysis during sample extraction have a significant impact on the microbiome profile. *Proper stabilization/ homogenization technology with OMNIgene•GUT can significantly reduce differences seen between mechanical lysis methods. Significant difference was evaluated by two-way ANOVA test with no rank matching (*p-value<0.015).*

Results – Pre-analytical changes during transport

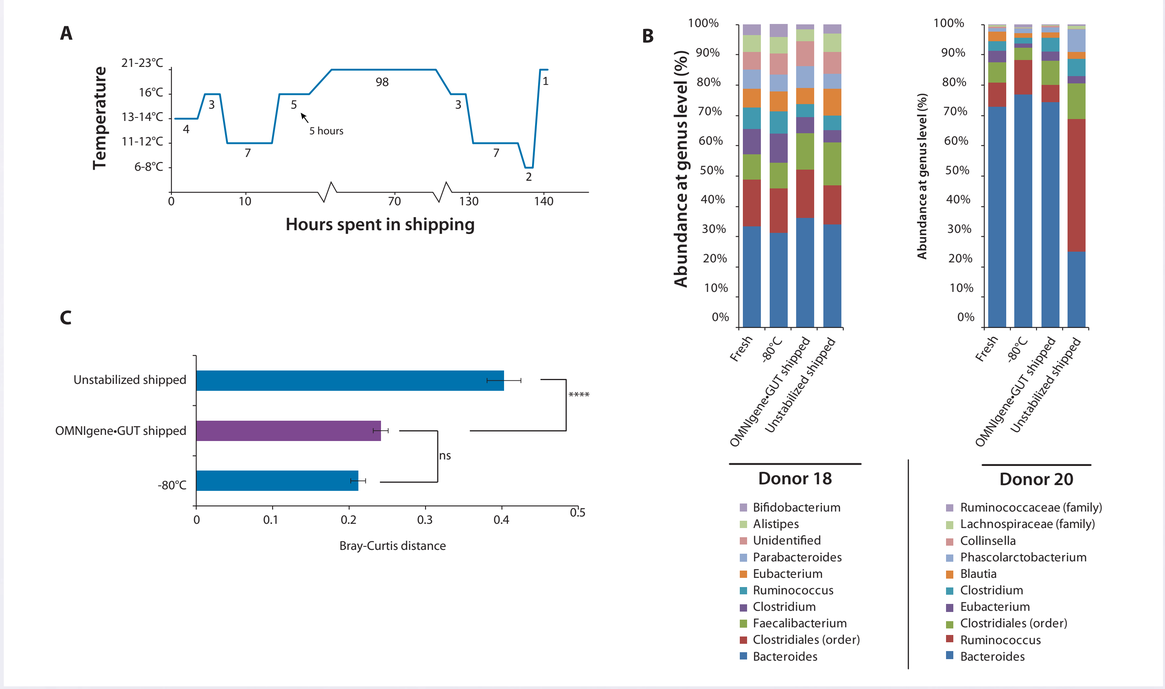


Figure 3: Exposure of fecal samples to environmental conditions has a significant impact on the microbiome profile. *(A) Temperatures experienced during shipping of collected stool samples collected in Ottawa and sent to Vancouver (Canada) by UPS with a return trip. Samples were extracted immediately upon return. (B) Genus level differences in % abundance (top ten by abundance are shown) between shipped, frozen and fresh samples from representative donors showcase donor dependent profile changes during unstabilized transport. OTUs that were not resolved to genus level are classified in brackets. (C) Samples collected in OMNIgene•GUT and shipped did not significantly differ in their microbiome profiles compared to -80°C samples. Unstabilized shipped samples significantly differed in their microbiome profile compared to OMNIgene•GUT shipped samples as measured by 2 sample Kolmogorov–Smirnov test (p-value < 0.0001).*

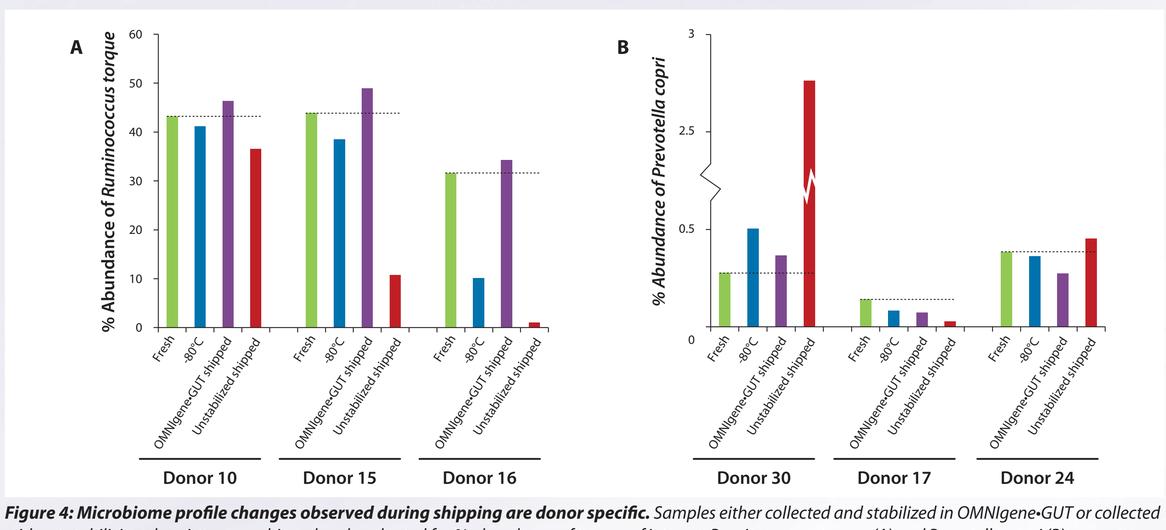


Figure 4: Microbiome profile changes observed during shipping are donor specific. *Samples either collected and stabilized in OMNIgene•GUT or collected without stabilizing chemistry were shipped and evaluated for % abundance of taxons of interest, Ruminococcus torque (A) and Prevotella copri (B). Sample transport without proper stabilization resulted in large taxon changes that are not consistent between donors. Samples collected in OMNIgene•GUT and shipped did not significantly differ in their microbiome compared to fresh samples.*

Discussion

- Gut microbiome studies have inherent sources of variability at the point-of-collection, extraction and sequencing which leads to requirements of increased replicates to build confidence in data.
- Choice of extraction methodology for gut samples leads to differences in data quality. Sample homogenization and stabilization reduce the variability introduced by extraction.
- Sample transport leads to unpredictable taxonomy level changes that can make significant changes to data interpretation. Clinically relevant microbiome results can be misinterpreted due to artificial changes of the microbiome. In our example, *Prevotella copri*, which has been correlated with rheumatoid arthritis[1], only shows consistent and accurate results compared to fresh when the OMNIgene•GUT stabilizing chemistry is utilized.

Conclusions

- Accurate measurement of a host microbiome requires standardization in methods of sample collection, stabilization, handling and extraction.
- Our data illustrates that pre-analytical and analytical factors contribute to microbiome variability, however, biospecimen stabilization is by far the most critical.
- OMNIgene•GUT collection, stabilization and homogenization technology effectively mitigates environmental impacts, enabling accurate representation of the in vivo biology.

Citations:

[1] Scher JU, et al. Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife*. 2013;2

DNA genotek

Superior samples
Proven performance

www.dnagenotek.com • support@dnagenotek.com