16S Metagenomics Frequently Asked Questions:

Q: What is 16s metagenomics workflow?

A: There are several steps involved in 16s metagenomics project. These include Sample collection, DNA extraction, library preparation, library quantification, sequencing, and statistical data analysis.

1. Sample Collection: A robust sample collection protocol that maintains the integrity of samples is both important and challenging in microbiome studies. The microbial composition of samples may be adversely affected by the collection process. Please visit the following link to find out what's the best collection method for your samples.

   Zymo Research: https://www.zymoresearch.com/pages/sample-collections

   Norgen BioTech: https://norgenbiotek.com/category/microbiome-collection-preservation-and-isolation

   BD: https://www.bd.com/en-us/offerings/capabilities/specimen-collection/swab-based-specimen-collection/bd-bbl-cultureswab-ez-collection-and-transport-systems

2. DNA isolation. The challenge here is to use a method that extracts DNA with "similar" efficiency for all microorganisms in order to reveal full diversity in the sample. Please check the following links to find out what extraction kit is best suited for your samples. You can always seek the information from the expert in the field you are working on to see which kit/method they use.

   PureLink™ Microbiome DNA Purification Kit: https://www.thermofisher.com/order/catalog/product/A29790

   Microbiome DNA Isolation Kit (Cat. 64100): https://norgenbiotek.com/product/microbiome-dna-isolation-kit

   QIAamp DNA Microbiome Kit: https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/microbial-dna/qiaamp-dna-microbiome-kit/#orderinginformation

   QIAamp DNA Stool Mini Kit: https://www.qiagen.com/us/spotlight-pages/ias/automated-qpcr-workflow/assay-setup/qiaamp-dna-stool-mini-kit/

   E.Z.N.A.® Stool DNA Kit: https://www.omegabiotek.com/product/e-z-n-a-stool-dna-kit/

   Meta-G-Nome™ DNA Isolation Kit for water: http://www.epibio.com/applications/nucleic-acid-purification-extraction-kits/dna-purification-genomic/meta-g-nome-dna-isolation-kit?details

3. Library construction (16S amplicon or any other target). ICBR Gene Expression & Genotyping Core currently offers 16S library construction services using either "quick-16S NGS Library Prep Kit" from Zymo or "Swift Amplicon 16S+ITS Panel" from Swift. If you are constructing own libraries, we recommend the method described by Illumina and the barcoding reagents provided in the Nextera XT kit (Illumina Inc.). (https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf). The method consists of two sequential PCR

reactions. You will need to order HPLC- purified primers and PCR reagents for the first reaction. The first PCR is exactly the same for every sample. In contrast, the second PCR reaction is unique for every sample. That's because you need a unique barcode for each sample. You can barcode up to 96 samples to be pooled for a single MiSeq run (2x300 cycles), using the Nextera XT kit. Other library construction protocols are also accepted. These include, for example, libraries constructed with the 16S ZymoResearch kit, or custom protocols as described in the literature.

4. Library quantification, sizing, and pooling. Use QUBIT for quantification, and the TapeStation (or Bioanalyzer) for sizing. If you have followed the Illumina recommended method for amplifying variable regions 3 and 4, the size of your library should be ~600 bp). Calculate molarity using Qubit concentration and average size on TapeStation, and pool equimolarly.

5. Sequencing run. For 16S (variable regions 3 and 4), we recommend MiSeq, version 3, 2x300 cycles with 20% PhiX spike-in and 10 pM loading. For a pool of 96 samples, this run format should yield ~100k reads per sample. Of course, it is your choice to request any other run configuration based on the goals of your experiment. Please keep in mind that version 2 configurations (e.g., 2x250 cycles) generate roughly half the number of reads as version 3 reagents (e.g., 2x300 cycles).

   a). For 16s projects started with library construction at ICBR GE core, the pooled library will be transferred directly to NextGen core for sequencing run. If you also need support with data analysis, please contact the Bioinformatics core (icbr-bioinformatics@ad.ufl.edu).

   b). For premade 16s libraries, please submit pool to the ICBR NextGen DNA sequencing core. We feel that it is best practice to request independent library pool evaluation services before sequencing, including quantification of "functional" library molecules by qPCR. You can choose to decline these services, in which case you will need to provide the core with all QC data for your pool, plus detailed conditions for sequencing run set up on the MiSeq. We need at least 15 ul at 10 nM (~3-5 ng/ul). To do so, your lab will need to have a CrossLab account (PI and lab members). From our website (www.biotech.ufl.edu), use your account credentials to login CrossLab and select UF-ICBR core. Navigate to the NextGen Sequencing core> Illumina sequencing services> MiSeq, etc. to fill out service request form.

6. Data analysis. You will need to do that on your own using your favorite method, or select data analysis services in your submission form. If you decide to request data analysis support, the ICBR Bioinformatics Core can do this for a fee and will contact you to discuss the process (ICBR-Bioinformatics@ad.ufl.edu).

Q: What are the DNA quality Requirements?

A: High sample quality is essential for successful 16s metagenomics experiments. Please ensure that your DNA samples**:**

- Is double-stranded. Single-stranded DNA will not be adapted in this template preparation process and can interfere with quantitation and polymerase binding.
- Has undergone a minimum of freeze-thaw cycles.
- Has not been exposed to high temperatures (> 65ºC for 1 hour can cause a detectable decrease in sequence quality), pH extremes (< 6 or > 9).
- Is intact during size-selection. If purifying DNA from a gel fragment, DO NOT use ethidium bromide for staining. We recommend SYBR Safe DNA Gel Stain with visualization on a blue light box (https://www.thermofisher.com/us/en/home/life-science/dna-rna-purification-

analysis/nucleic-acid-gel-electrophoresis/dna-stains/sybr-safe.html). UV will damage the DNA. Even a few seconds of UV irradiation appears to render DNA unsuitable for sequencing.

- Has an OD260/280 ratio of approximately 1.8 to 2.0.
- Does not contain insoluble material.
- Does not contain RNA contamination.
- Has not been exposed to intercalating fluorescent dyes or ultraviolet radiation.
- Does not contain chelation agents (ie. EDTA), divalent metal cations (i.e.,Mg^{2+}), denaturants (Guanidinium salts, Phenol), or detergents (SDS,Triton-X100).
- Does not contain carryover contamination from the starting organism/tissue (heme, humic acid, polysaccharides, polyphenols, etc.). Some DNA isolation procedures might contain some contaminant(s) that can affect the quality of the sequencing run.

Q: How should I quantify the DNA samples?

A: Use QUBIT (fluorescence-based method). Nanodrop (absorbance) is not adequate for quantification, although it does provide valuable information about sample's purity. Gel electrophoresis (or equivalent… TapeStation) may also be important if you're concerned about the integrity of the samples.

Q: How much DNA needed for 16S metagenomics library prep?

The standard protocol for Zymo kit requires between 5-20 ng of total DNA. For Swift kit, 1ng is recommended, and sufficient yields are attainable from as little as 10pg-50ng depending on sample type.

Q: Which kits are GE use for my 16s metagenomics project? What is the difference of those kits?

A: We use either "quick-16S NGS Library Prep Kit" from Zymo or "Swift Amplicon 16S+ITS Panel" from Swift. The Zymo kit uses two primer sets to amplify V1 and V2, or V3 and V4. The Swift kit utilizes five primer sets to amplify and create libraries for all V1-V9 and fungal ITS1 and ITS2 genes in a single primer pool.

Q: Will the GE Core send QC results before they starting my library prep?

A: We will perform a Qubit assay. We also random pick 4 samples to run a gDNA TapeStation to identify the integrity of gDNA. If you would like to run TapeStation for all your samples, per sample cost is $7.2. This is not included in the 16S library preparation service fee.

Q: How do customers submit replacement samples?

A: If after initial QC customer need to submit more sample to the core, please upload the sample info of replaced sample to iLab, and clearly indicate which samples replaced which samples.

Q: Can customers get back the samples?

A: Yes. Please make a note in the comment box in iLab submission form that you want the samples back upon the completion of the project. We keep customers' samples up to 6 months. It is customers' responsibility to retrieve the samples on time.

Q: How long does the core keep customer's leftover samples for?

A: Due to limited space in the GE cores freezer we only keep customer's samples for 6 months. If customer would like samples returned to them then they should contact the core to schedule a day and time to come and pick them up.

Q: What is the turnaround time for 16S metagenomics project?

A: Turnaround time depends largely on the volume of activity at the time of sample submission and the complexity of the project. Given typical workloads, library construction requires approximately 2-3 weeks, and sequencing-analysis often requires approximately 1 - 2 weeks after that. If data analysis services are requested, this would add 1-2 weeks to the process.