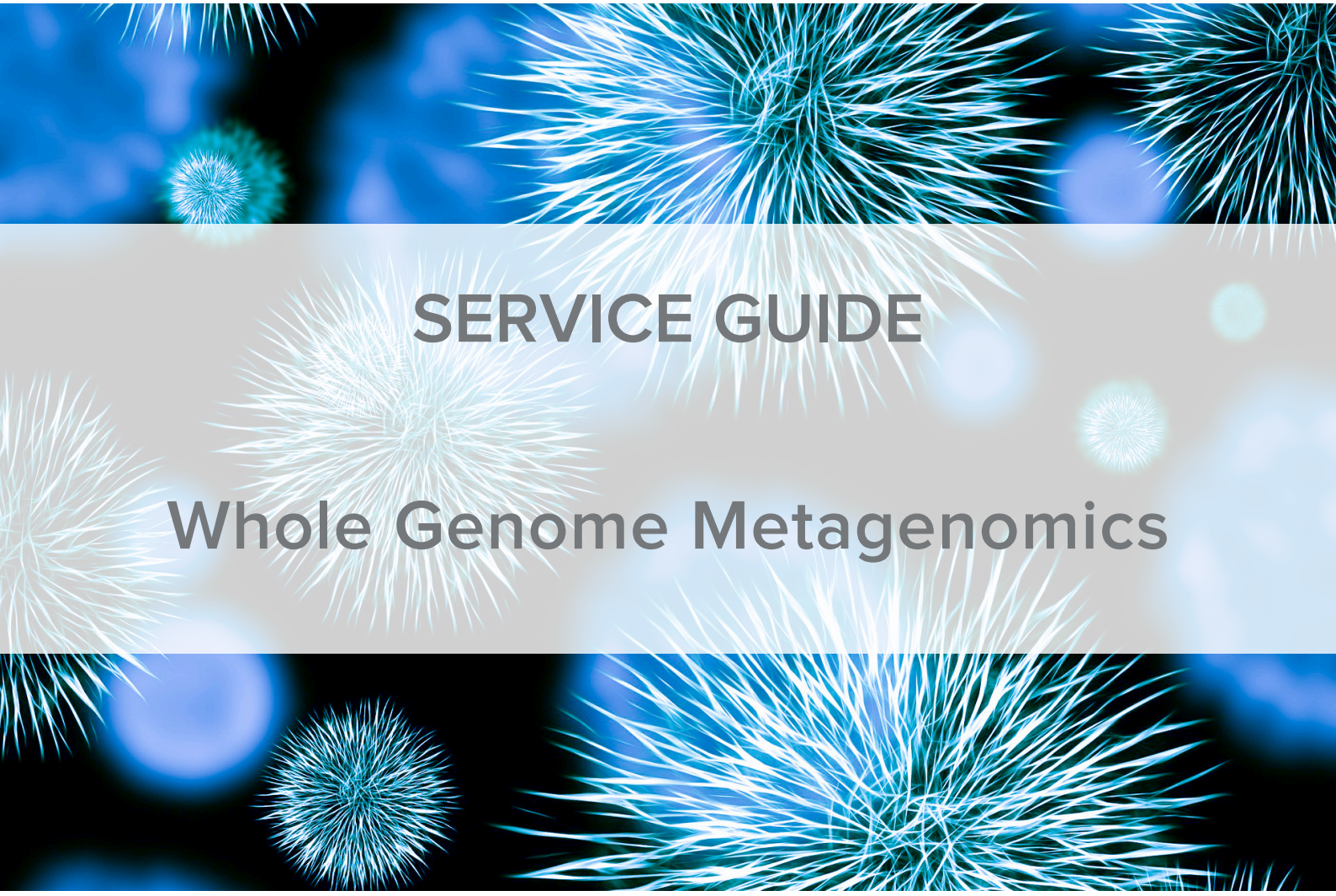


SERVICE GUIDE

Whole Genome Metagenomics

CustomerCare@agrf.org.au
1300 247 301
www.agrf.org.au
AGRF Ltd.
SVG2509WGM

Service Guide

Whole Genome Metagenomics

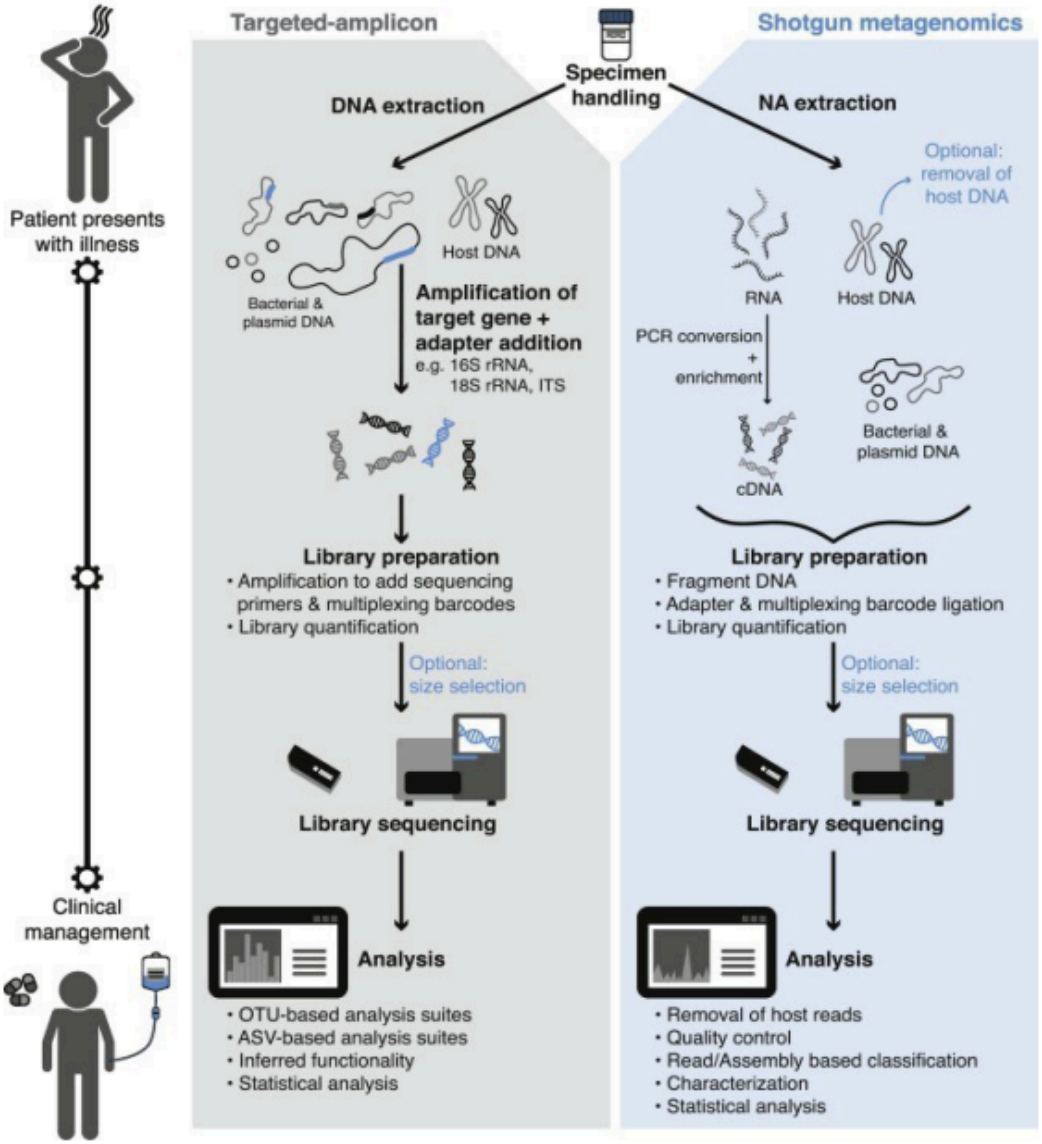
Service Guide
Whole Genome Metagenomics

1.0 Overview

AGRF's Whole Genome Metagenomics Service (MetaWGS) is an evolution of the Microbial Profiling service in which most, if not all microbes within a sample can be identified without needing to perform a targeted PCR approach. This new service does not PCR-amplify variable regions within bacteria or fungi, but instead uses a shotgun sequencing (Figure 1) approach in which all microbes are simultaneously identified. This powerful analysis removes amplicon-based biases from your data and gives you an almost complete picture of the community composition of your sample.

Figure 1: Comparison of target-amplicon approach to Whole Genome and whole-transcriptome metagenomics (Forbes, 2018).

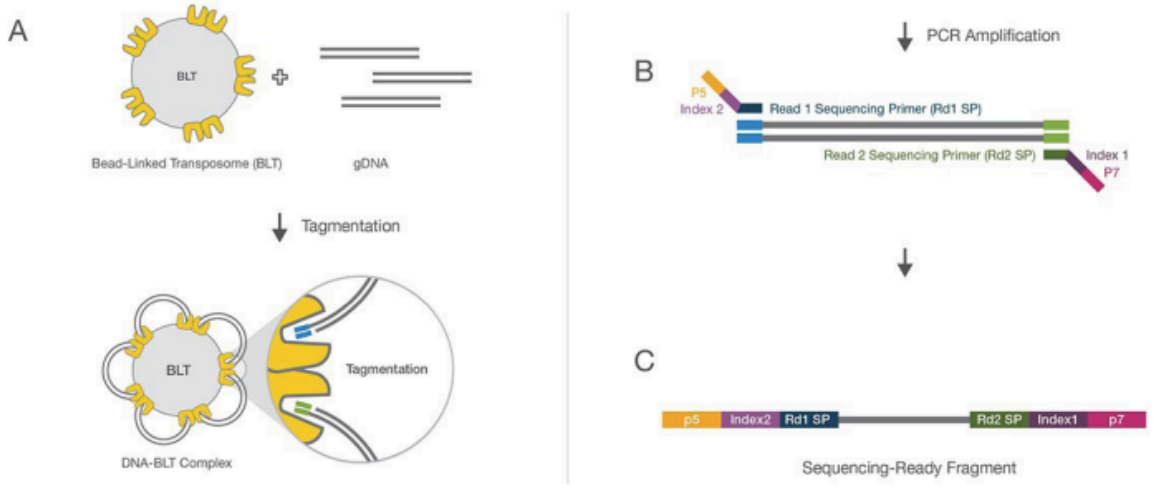
AGRF offers both targeted-amplicon and Whole Genome metagenomics.



2.0 Workflow

For this service, the Illumina DNA Preparation (M) will be utilised to prepare your samples as it accepts a vast range of sample inputs and produces a high-quality library (Figure 2). This is a transposase-based preparation which uses a bead-linked transposome (BLT) to generate a consistent library size (~350bp insert) and uniform coverage.

Figure 2: Graphical representation of the Illumina DNA Preparation (M) method. A) The BLT combines the fragmentation of the gDNA with addition of Illumina sequencing primers. B) PCR amplification of the fragmented material adds a unique-dual index (UDI) for sample multiplexing and to generate the final sequencing ready library. C) The final library ready for sequencing. (Illumina, 2020).



In addition, the service has two data output offerings using 150PE sequencing:

1. Guaranteed 5Gbp
   - Low-to-medium complexity material
   - Oral swabs, faecal, bioreactors, food
2. Guaranteed 10Gbp
   - Medium-to-high complexity material
   - Soil, wastewater, other environmental samples

3.0 Technical Consideration

In offering the MetaWGS service, we will combine your indexed library with other submissions, and run them together in the sequencing run. Due to inherent technical limitations and sequencing error rates, a very small number (<0.01%) of reads assigned to your sample may have originated from another sample. Likewise, very low levels of your samples may be present in another sample's data set. If this will affect your experimental interpretation, then you should consider processing samples in isolation as a custom project with AGRF (please contact us for details).

The MetaWGS service does not deplete host material as this is not currently possible for this offering. If your samples contain a large amount of host material such as nasal swabs or root extracts, most of the data we return will be the host. You may need to consider much deeper sequencing to analyse the microbial community, or you may need to consider enriching your samples prior to submission, but this can bias your data towards bacteria. Further to this, the MetaWGS service will not identify RNA viruses. If this is something you require, please contact us.

4.0 MetaWGS Submission Types

4.1 MetaWGS with Extraction

This service is for clients who wish to submit a sample that requires nucleic acid extraction, prior to library preparation. This service includes:

- Nucleic Acid Extraction
- Library Preparation using the Illumina DNA Preparation (M) kit.
- Sequencing on the Illumina NovaSeqX Plus platform with 150PE chemistry.
- Bioinformatics analysis and reporting of taxonomic distribution.

4.2 MetaWGS without Extraction

This service is for submitters who have are submitting gDNA. This service includes:

- Library Preparation using the Illumina DNA Preparation (M) kit.
- Sequencing on the Illumina NovaSeqX Plus platform with 150PE chemistry.
- Bioinformatics analysis and reporting of taxonomic distribution.

5.0 Sample Requirements for Submissions without Extraction

The standard input for any genome sequencing project is high quality genomic DNA.

Our recommendations for isolation of DNA are listed below:

- Standard silica column-based extractions generally produce DNA of suitable quality for Illumina library preparations.
- Please ensure the extraction protocol includes an RNase treatment to ensure removal of RNA.
- Purified DNA should be eluted/resuspended in 10mM TrisHCl, pH 8.5 or nuclease-free water.
- Quantification of gDNA by dsDNA assay such as PicoGreen or Qubit is highly recommended prior to submission.
- No AE buffer. No TE buffer. No buffer containing EDTA.

Table 1: Sample quantity required for the MetaWGS service.

Sample Type	DNA Quantity (ng)	DNA Concentration (ng/µL)	Volume (µL)
gDNA	≥ 100	≥7.5	25

Sample quality is a key factor for successful NGS experiments. AGRF will perform quality control prior to commencing a project using an agarose gel and fluorimetry, however we recommend you check the DNA prior to submission. It should be noted that for this service, the reported concentration that AGRF will provide will include all dsDNA within the sample, which also includes host material. It is highly recommended that the purity of the nucleic acid within your samples is assessed prior to submission and can be assessed by measuring the absorbance spectra via a spectrophotometer (e.g. Nanodrop). The ratio of absorbance values 260nm and 280nm or 230nm provides an estimate of sample purity or the presence of common contaminants. Purified DNA is expected to have a A260/280 ratio of ~1.8 (Table 2).

Table 2: Recommendations for sample purity as assessed by absorbance spectra.

Ratio	Target (DNA)	Low Ratio (<1.6) indications
A260/280	1.8	Residual phenol from extraction, proteins or very low conc. of nucleic acids (<1ng/µl)
A260/A230	≥2.0	Residual guanidine from the extraction protocol Carryover of carbohydrates (e.g. plant polysaccharides)

6.0 Sample and Data Storage

Samples are stored with AGRF for 1 month after you receive your data. If you wish for your samples to be returned, you must discuss this with your account manager during quoting or contact us after you receive your data. At the completion of your project, we can either:

- Return your samples by courier at ambient (please ask your account manager for a quote).
- Return samples by courier with dry ice (please ask your account manager for a quote).

If we are not notified within the specified time frame, samples will be automatically discarded.

Files will remain on the AGRF server for 1 month before being removed. A backup copy will be archived at AGRF for a further 2 months. Charges will apply for restoring archived files to the server.

7.0 Sample Submissions

7.1 Samples Requiring Nucleic Acid Extraction Online Submission

- Submit your sample details online.
- Select: “Extraction and Next-Generation Sequencing” as the Service Type.
- Please complete and upload the “Template File” excel template. Note: AGRF will use the “Sample Code” to name the sequences produced from the project.

Submit and print a paper copy of your sample submission, to be included with your sample package.

AGRF can organise dry ice shipment for your samples as part of your quoted services or you can use our free shipping between nodes once a week service. For information on this service go to Free Shipping.

Post/send/deliver samples to the following addresses;

Physical address (courier):

AGRF Adelaide
PLANT GENOMICS CENTRE
HARTLEY GROVE
URRBRAE
SA 5064

Postal address (mail):

AGRF Adelaide
PMB1 GLEN OSMOND
URRBRAE
SA 506

7.2 Extracted Samples (Purified Nucleic Acid) Online Submission

- Submit your sample details online.
- Select: “Next-Generation Sequencing” as the Service Type.
- Tube submissions - we require independent sets of tubes per target.
- Plate submissions - we require 1 plate only per 96 samples.
- Submission Format – by selecting tube or plate, the “Sample File” template link will appear. Click “Download Template” and enter your sample details:
- Each sample name must be unique and can only contain alphanumeric characters and underscores.
- Save completed Template File locally, select “Browse” to upload file.
- Submit and print a paper copy of your sample submission, to be included with your sample package.
- Note: Submission Format: • <24 samples – 1.5mL tube
- Tubes can be sent at room temperature using express post
- ≥24 samples:
- Please complete plate submissions, (an additional handling charge of $1.50 per sample will occur if tubes are used).

We recommend shipping plates that are heat-sealed, or strip-cap sealed on dry ice.

Shipping Address:

Melbourne MetaWGS service
AUSTRALIAN GENOME RESEARCH FACILITY
LEVEL 13, VICTORIAN COMPREHENSIVE CANCER CENTRE
305 GRATTAN STREET
MELBOURNE, VIC, 3000

*Note: our loading dock is open from 7am to 3.30pm weekdays.

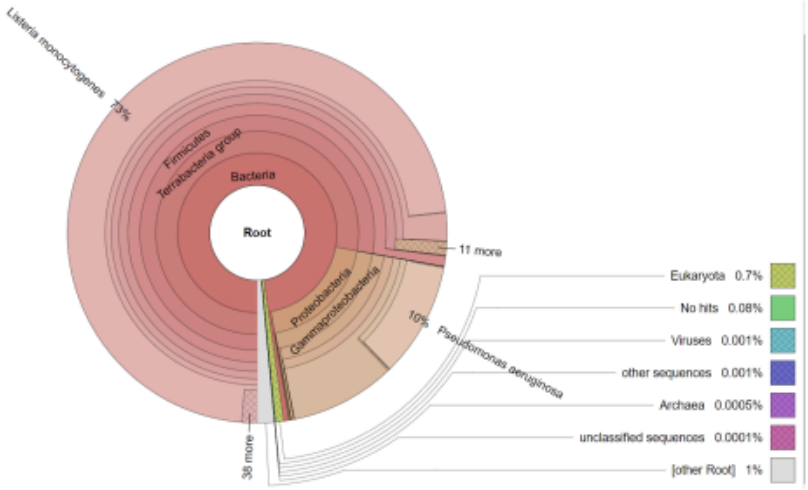
8.0 Results and Data Outputs

The standard MetaWGS service will provide read-based classification1 analysis results. All MetaWGS sequencing projects will undergo quality control to assess the quality of both the sequencing and the library preparation. The raw read sequences will be pre-processed and then processed using the latest version of Kraken and Bracken for profiling the composition of microbial communities using a custom-built database. The database is built using genomes and contigs from NCBI consisting of bacterial, fungal and viral sequences. Human, plant and other vector sequences are also included in the database for quality control. Further functional profiles will be generated using HUMAnN.

AGRF will provide the following data and results:

- FASTQ outputs for your individual samples
- Individual taxa abundance file for each sample, provided in .txt and .xls format
- Krona interactive plot for each sample, provided in .html file format (Figure 3).
- Bar plot provided in a.html file format containing all samples.
- .xls file containing absolute abundance from all samples.
- .xls file containing relative abundance from all samples.
- Gene families, pathway abundance and pathway coverage are provided for all samples combined as well as for all individual samples in text and excel format.

Figure 3: Example outputs of read-based classification analysis showing classification of test data.



1 Contig assembly classification analysis is a custom analysis which can provide host sequence filtering. If you require contig-based classification, please contact us.

9.0 Quality Statement

Non-clinical works are performed following the strict requirements of ISO17025: 2005. AGRF Ltd. is accredited in the field of Biological Testing (Scope: DNA Analysis) according to the ISO17025: 2005 standard by the National Association of Testing Authorities (NATA). Staff and analysis processes follow Standard Operating Procedures, which define responsibilities and quality checks to achieve reported standards. Compliance is monitored at regular reviews and during internal audits. All work is supervised by a person with relevant qualifications and is checked while in progress and upon completion to ensure that it meets the necessary ISO17025: 2005 standards.

10.0 Bibliography

Forbes, J. D. (2018). Highlighting clinical metagenomics for enhanced diagnostic decision-making: a step towards wider implementation. Computational and structural biotechnology journal, 16, 108-120. • Illumina. (2020). Illumina DNA Preperation (M). Retrieved from Illumina, link here