Read-based metagenome analysis pipeline PE metagenomic reads QC (fastQC) Trim adapters & Filtered reads quality filter (fastq) (bbduk.sh) Gene prediction rRNA gene filtering (FragGeneScane) (RNAselector, SortMeRNA, SINA) rRNA classification Gene annotation Gene annotation (targeted) (all) (SilvaNGS)

- specific protein domains (PFam) + HMMER

- proteins from a specific taxa AA sequences + blastx

- GenDB in-house
- MG-RAST web-based
- individual databases searches (PFam, KEGG, COG, TIGRFAM)

classification based on an alignment against a curated

reference database to exclude non-rRNA sequences