

## Sample Submission

Customers can either submit extracted DNA eluted in 10 mM Tris-HCl pH 8.0 or nuclease free water, or provide us with strains. For strain submissions, pure cultures of each strain are grown on plates or in broth. Cells are harvested and resuspended in a tube with DNA/RNA Shield (Zymo Research, USA) following MicrobesNG strain submission procedures, and sent to MicrobesNG for sequencing.

## DNA Processing

Five to forty microlitres of the cell suspension are lysed with 120  $\mu$ L of TE buffer containing lysozyme (final concentration 0.1 mg/mL) and RNase A (ITW Reagents, Barcelona, Spain) (final concentration 0.1 mg/mL), incubated for 25 min at 37°C.

Proteinase K (VWR Chemicals, Ohio, USA) (final concentration 0.1mg/mL) and SDS (Sigma-Aldrich, Missouri, USA) (final concentration 0.5% v/v) are added and incubated for 5 min at 65°C. Genomic DNA is purified using an equal volume of SPRI beads and resuspended in EB buffer (10mM Tris-HCl, pH 8.0).

DNA extracted by MicrobesNG, or sent directly by the customer is then quantified with the Quant-iT dsDNA HS kit (ThermoFisher Scientific) assay in an Eppendorf AF2200 plate reader (Eppendorf UK Ltd, United Kingdom) and diluted as appropriate.

## Illumina Sequencing (SGS and EGS)

Genomic DNA libraries are prepared using the Nextera XT Library Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: input DNA is increased 2-fold, and PCR elongation time is increased to 45 seconds.

DNA quantification and library preparation are carried out on a Hamilton Microlab STAR automated liquid handling system (Hamilton Bonaduz AG, Switzerland). Libraries are sequenced on an Illumina NovaSeq 6000 (Illumina, San Diego, USA) using a 250 bp paired end protocol.

Reads are adapter trimmed using Trimmomatic version 0.30 [1] with a sliding window quality cutoff of Q15. De novo assembly is performed on samples using SPAdes version 3.7 [2], and contigs are annotated using Prokka 1.11 [3].

## ONT Sequencing (EGS only)

Long read genomic DNA libraries are prepared with Oxford Nanopore SQK-LSK109 kit with Native Barcoding EXP-NBD104/114 (ONT, United Kingdom) using 400-500 ng of HMW DNA. Barcoded samples are pooled together into a single sequencing library and loaded in a FLO-MIN106 (R.9.4.1) or FLO-MIN111 (R10.3) flow cell in a GridION (ONT, United Kingdom).

Hybrid assembly is performed using Unicycler version 0.4.0 [4], and contigs are annotated using Prokka version 1.11 [3].

[1] Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

[2] Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455–477.

[3] Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 30(14):2068-9

[4] Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology* 13(6): e1005595.