Bacterial genomics: Sequencing and Bioinformatics Training

14 – 18 July 2025

Sequencing and Bioinformatics Training Course

Course Overview

This one-week intensive training course provides participants with a comprehensive introduction to sequencing workflows and bioinformatics analysis, with a focus on bacterial genomics. The course is designed for laboratory scientists, researchers, and students who wish to understand the end-of-the-end process of Bacterial DNA sequencing and Bioinformatics analysis. Participants will engage in both hands-on wet-lab activities and practical bioinformatics sessions.

Course Contents

Lectures

- DNA Extraction: Theory
- Bacterial Genomics and Applications

Wet-Lab

- DNA Extraction: Laboratory Protocols
- DNA Quality Control and Assessment
- Library Preparation for Sequencing
- Illumina Next-Generation Sequencing

Bioinformatics analysis

- Introduction to Unix Command Line and Conda Environments
- Sequence Data Retrieval and Preprocessing
- Genome Assembly and Quality Metrics
- Species Identification and Typing
- AMR and Virulence Gene Screening
- Phylogenetic Analysis and Interpretation
- Overview of Online GUI Bioinformatics tools

Learning Outcomes

1. Understand the key steps in DNA extraction, library preparation, and sequencing.
2. Navigate and use Unix command line tools and manage software using Conda.
3. Retrieve, assess, and preprocess sequencing data for downstream analysis.
4. Assemble bacterial genomes and interpret assembly metrics.
5. Identify bacterial species and determine clonal complex, strain types, spa types, SCCmec types etc
6. Detect antimicrobial resistance and virulence genes
7. Construct and interpret phylogenetic trees to assess genomic relatedness.

Training Timetable

Time	Monday	Tuesday	Wednesday	Thursday	Friday
8:00 - 9:00 AM	Welcome & Introduction DNA Extraction Lecture	Library Preparation		Lecture: Bacterial Genomics	AMR & Virulence Gene Screening
9:00 - 10:30 AM	DNA Extraction Lab	Library Preparation		Bacterial Genomics	AMR & Virulence Gene Screening
10:30 - 11:00 AM	**Tea Break**				
11:00 - 1:00 PM	DNA Extraction Lab	Library Preparation	Finish Library Prep & Start Sequencing	Dataset Retrieval, QC & Preprocessing	Phylogenetics
1:00 - 2:00 PM	**Lunch**				
2:00 - 3:30 PM	DNA Quality Control Check	Library Preparation	Intro to Unix & Conda	Genome Assembly, Species ID & Typing	Online Tools Overview
3:30 - 5:00 PM		Library Preparation	Intro to Unix & Conda	Genome Assembly, Species ID & Typing	Feedback & Closing

DNA Extraction Using Zymo BashingBead™ Lysis Kit

Materials Needed

- Fungal or bacterial cells (50–100 mg wet weight)
- Water or isotonic buffer (e.g., PBS)
- ZR BashingBead™ Lysis Tube (0.1 mm & 0.5 mm beads)
- BashingBead™ Buffer
- Bead beater (e.g., TerraLyzer™, FastPrep®-24, Disruptor Genie™, vortex)
- Microcentrifuge
- Zymo-Spin™ III-F Filter
- Zymo-Spin™ IICR Column
- Collection tubes
- Genomic Lysis Buffer
- DNA Pre-Wash Buffer
- g-DNA Wash Buffer
- DNA Elution Buffer
- 1.5 ml microcentrifuge tube

Procedure

Sample Preparation

1. Add 50–100 mg of fungal or bacterial cells (wet weight) resuspended in up to 200 µl of water or isotonic buffer (e.g., PBS) into a ZR BashingBead™ Lysis Tube (0.1 mm & 0.5 mm).
2. Add 750 µl BashingBead™ Buffer to the tube.
3. Cap tightly to prevent leakage.

Cell Disruption

4. Secure the tube in a bead beater fitted with a 2 ml tube holder.
5. Process at maximum speed for ≥ 5 minutes.
6. Processing time varies by device:
   - High-speed disrupters: ~3 minutes (e.g., TerraLyzer™, FastPrep®-24)
   - Lower-speed devices: up to 20 minutes (e.g., Disruptor Genie™, vortex)

Centrifugation

7. Centrifuge the lysis tube at 10,000 × g for 1 minute.

Filtration

8. Transfer up to 400 µl of supernatant to a Zymo-Spin™ III-F Filter in a Collection Tube.
9. Centrifuge at 8,000 × g for 1 minute.

Lysis Buffer Addition

10. Add 1,200 µl Genomic Lysis Buffer to the filtrate in the Collection Tube.

DNA Binding

11. Transfer 800 µl of the mixture to a Zymo-Spin™ IICR Column in a Collection Tube.
12. Centrifuge at 10,000 × g for 1 minute.
13. Repeat with the remaining mixture.
14. Note: Max column capacity is 800 µl.

Pre-Wash

15. Discard the flow-through and place the column in a new Collection Tube.
16. Add 200 µl DNA Pre-Wash Buffer, then centrifuge at 10,000 × g for 1 minute.

Wash

17. Add 500 µl g-DNA Wash Buffer to the column.
18. Centrifuge at 10,000 × g for 1 minute.

DNA Elution

19. Transfer the column to a clean 1.5 ml microcentrifuge tube.
20. Add 100 µl (minimum 35 µl) of DNA Elution Buffer directly to the column matrix.
21. Centrifuge at 10,000 × g for 30 seconds to elute the DNA.

DNA Extraction Worksheet

SNo.	Sample ID	A260/280	A260/230	Nanodrop Conc (ng/ul)	Qubit Conc (ng/ul)
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					

Principles of select Next Generation Sequencing Methods

Illumina sequencing principle

In short-read sequencing by Illumina, DNA library is loaded onto a flow cell, and the adapters from the modified DNA hybridize to the oligonucleotides that coat the surface of the flow cell. Once the fragments have attached, cluster generation begins, where thousands of copies of each fragment are generated through a process known as bridge amplification. In this process, the strand folds over, and the adapter on the end of the molecule hybridizes to another oligonucleotide in the flow cell. DNA polymerase incorporates nucleotides to build double-stranded bridges of the DNA molecules, which are subsequently denatured to leave single-stranded DNA fragments tethered to the flow cell. This process is repeated over and over, generating several million dense clusters of double-stranded DNA. After bridge amplification, the reverse DNA strands are cleaved and washed away, leaving only the forward strands. Then, sequencing by synthesis begins, in which fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) are incorporated into the newly synthesized DNA strand at each cycle. After incorporation, a laser excites the fluorophore on the strand, which emits a characteristic fluorescence emission signal that corresponds to the base. The phosphate-linked fluorophore is cleaved out by an enzyme and released into the buffer. Another set of dNTPs are added in the subsequent cycle.

PacBio sequencing principle

In single molecule real time (SMRT) sequencing, DNA is fragmented and ligated to hairpin adapters to form a topologically circular molecule known as a SMRTbell. Once the SMRTbell is generated, it is bound by a DNA polymerase and loaded onto a SMRT cell for sequencing. Each SMRT cell contains up to 8 million zero-mode waveguides (ZMWs) which are chambers that hold picoliter volumes. As the DNA mixture floods the ZMW, the SMRTbell template and polymerase become immobilized on the bottom of the chamber. Fluorescently labelled dNTPs are added to begin the sequencing reaction. As the polymerase begins to synthesize the new DNA strand, a fluorescent dNTP is briefly held in the detection volume, and a light pulse from the bottom of the well excites the fluorophore. The emission is detected by a camera which records the wavelength and relative position of the base in the strand. The phosphate-linked fluorophore is cleaved from the nucleotide and released into the buffer preventing it from fluorescing anymore.

The DNA sequence is determined by changing fluorescence emissions that are recorded within each ZMW with a different color corresponding to each DNA base.

Nanopore sequencing principle

In Oxford Nanopore Technologies (ONT) sequencing, arbitrarily long DNA are tagged with sequencing adapters preloaded with a motor protein on one or both ends. The DNA is combined with tethering proteins and loaded onto the flow cell for sequencing. The flow cell contains thousands of protein nanopores embedded in a synthetic membrane, and the tethering proteins bring the DNA molecules toward these nanopores. Then, the sequencing adapter inserts into the opening of the nanopore, and the motor protein begins to unwind the double-stranded DNA. An electric current is applied, which, in concert with the motor protein, drives the negatively charged DNA through the pore at a rate of about 450 bases per second. As the DNA moves through the pore, it causes characteristic disruptions to the current, known as a 'squiggle'. Changes in current within the pore correspond to a particular k-mer (i.e., a string of DNA bases of length k) which is used to identify the DNA sequence.

Reference

Ebertz, D. A. (2020, November 2). A Journey Through The History Of DNA Sequencing. *The DNA Universe BLOG*. https://the-dna-universe.com/2020/11/02/a-journey-through-the-history-of-dna-sequencing/

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., & Law, M. (2012). Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology*, *2012*. https://doi.org/10.1155/2012/251364

Logsdon, G. A., Vollger, M. R., & Eichler, E. E. (2020). Long-read human genome sequencing and its applications. *Nature Reviews. Genetics*, *21*(10), 597–614. https://doi.org/10.1038/s41576-020-0236-x

Sereika, M., Kirkegaard, R. H., Karst, S. M., Michaelsen, T. Y., Sørensen, E. A., Wollenberg, R. D., & Albertsen, M. (2022). Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nature Methods*, *19*(7), Article 7. https://doi.org/10.1038/s41592-022-01539-7

Zhou, X., Ren, L., Meng, Q., Li, Y., Yu, Y., & Yu, J. (2010). The next-generation sequencing technology and application. *Protein & Cell*, *1*(6), 520–536. https://doi.org/10.1007/s13238-010-0065-3

Links to YouTube videos of the Principles of Next Generation Sequencing Technologies

Next Generation Sequencing (NGS) - An Introduction

Sanger Sequencing/Chain Termination Method

Introduction to SMRT Sequencing

Single Molecule Real Time Sequencing - Pacific Biosciences

Illumina Sequencing by Synthesis

Nanopore DNA sequencing

Introduction to Nanopore Sequencing