Applied Genomics Project: Compliance Verification

Confirmation of Alignment with AG Course Requirements

Martina Castellucci
University of Bologna

August 4, 2025

The project **"Hybrid Genome Assembly and Integrative Multi-Omics of *Purpureocillium lilacinum* PLA-C1"** is fully aligned with the key topics of the Applied Genomics course. It encompasses the entire genomic workflow, from environmental sample selection to functional validation and bioeconomic application. Below, a structured checklist connects the project steps to the course topics.

1. Choice of Organism and Sample Source

- **Project:** Isolation of *P. lilacinum* from PLA-enriched compost, selected for its potential to degrade bioplastics.
- **Course Topics:** Environmental genomics, selection of biologically relevant model organisms, linkage to industrial or ecological application.
- **Rationale:** Targeting organisms from PLA-rich environments increases the probability of detecting biodegradation pathways.

2. DNA Extraction and Quality Control

- **Project:** High-molecular-weight DNA extraction (CTAB-based), validated by Nanodrop, Qubit, and gel electrophoresis; purity $A_{260/280} \approx 1.85$.
- **Course Topics:** DNA/RNA quality assessment, importance of integrity and purity for next-generation sequencing (NGS).
- **Rationale:** High-quality DNA ensures accurate sequencing, reduces assembly fragmentation, and improves downstream analyses.

3. Sequencing Strategy: Hybrid Whole-Genome Sequencing

- **Project:** Illumina short reads (∼100X, 150 bp PE) + Oxford Nanopore long reads (∼30–50X).
- **Course Topics:** Short- vs. long-read sequencing, hybrid strategies, trade-offs in accuracy and contiguity.

- **Rationale:** Combining Illumina for error correction with Nanopore for structural resolution enables high-quality de novo fungal assembly.

4. Genome Assembly and Quality Assessment

- **Project:** Final assembly of 41.6 Mb, N50 = 1.32 Mb, BUSCO completeness = 98.3%, GC content = 53.1%.
- **Course Topics:** Genome assembly metrics (N50, GC%), completeness assessment (BUSCO), hybrid assembly validation.
- **Rationale:** Reliable assembly metrics ensure the genome is suitable for annotation and downstream multi-omics integration.

5. Functional Genome Annotation

- **Project:** Structural annotation with MAKER; functional annotation with InterProScan, CAZy, and AntiSMASH, highlighting esterases and PHA depolymerases.
- **Course Topics:** ORF prediction, domain search, biosynthetic gene cluster (BGC) detection, functional annotation pipelines.
- **Rationale:** Annotating gene content and enzyme-coding potential connects genome sequence to biotechnological applications.

6. Transcriptomic Validation (RNA-Seq)

- **Project:** RNA-Seq under PLA vs. control conditions revealed 84 differentially expressed genes, including candidate depolymerases.
- **Course Topics:** RNA-Seq workflows, transcript quantification, differential expression for functional validation.
- **Rationale:** Experimental validation links predicted gene content to actual functional response under bioplastic exposure.

7. Comparative and Phylogenomic Analysis

- **Project:** Orthogroup analysis and synteny mapping revealed 314 unique genes and distinct phylogenetic placement.
- **Course Topics:** Comparative genomics, orthology inference, genome evolution, and population-level diversity analysis.
- **Rationale:** Comparative data confirm the novelty of the isolate and support future functional or evolutionary studies.

8. Applied and Translational Perspective

- **Project:** Identification of enzymes relevant for PLA degradation and potential bioplastic waste valorization.
- **Course Topics:** Application of genomic data to industrial biotechnology, environmental genomics, and bioeconomy.
- **Rationale:** Project outcome provides actionable resources for enzyme discovery and circular bioeconomy strategies.

Integrated Checklist of Steps

- Selection of environmentally and industrially relevant organism
- Extraction of high-quality nucleic acids with NGS-compliant QC
- Hybrid short- and long-read sequencing
- Assembly with N50, GC%, and BUSCO completeness evaluation
- Structural and functional genome annotation
- Transcriptome-based functional gene validation
- Comparative genomics and phylogenomic context
- Translation to industrial and environmental applications