

Repository Setup Checklist

Use this checklist to ensure all files are properly organized before pushing to GitHub.

Downloaded Files Mapping

Main Files

- ☐ metagenome_pipeline.groovy → main.nf
- ☐ nextflow_config.groovy → nextflow.config

Configuration Files (conf/)

- ☐ base_config.groovy → conf/base.config
- ☐ slurm_config.groovy → conf/slurm.config
- ☐ aws_config.groovy → conf/awsbatch.config
- ☐ docker_singularity_config.groovy → conf/docker.config
- ☐ Create conf/singularity.config
- ☐ Create conf/conda.config
- ☐ Create conf/test.config

Module Files (modules/)

QC

- ☐ Extract FastQC process → modules/qc/fastqc.nf
- ☐ Extract MultiQC process → modules/qc/multiqc.nf

Preprocessing

- ☐ kneaddata_module.groovy → modules/preprocessing/kneaddata.nf

Taxonomy

- ☐ metaphlan_module.groovy → modules/taxonomy/metaphlan.nf

Functional

- ☐ humann_module.groovy → modules/functional/humann.nf

Assembly

- ☐ megahit_module.groovy → modules/assembly/megahit.nf

- ☐ Extract SPAdes process → `modules/assembly/spades.nf`
- ☐ Extract filter_contigs → `modules/assembly/filter_contigs.nf`

Annotation

- ☐ Extract from `annotation_modules.groovy`:
- ☐ `modules/annotation/prodigal.nf`
- ☐ `modules/annotation/kegg.nf`
- ☐ `modules/annotation/cazy.nf`
- ☐ `modules/annotation/ardb.nf`

Mapping

- ☐ Extract from `mapping_modules.groovy`:
- ☐ `modules/mapping/bwa.nf`
- ☐ `modules/mapping/bowtie2.nf`
- ☐ `modules/mapping/samtools.nf`

Binning

- ☐ Extract from `binning_modules.groovy`:
- ☐ `modules/binning/metabat2.nf`
- ☐ `modules/binning/maxbin2.nf`
- ☐ `modules/binning/concoct.nf`
- ☐ `modules/binning/dastool.nf`
- ☐ `modules/binning/checkm.nf`

Clustering

- ☐ Extract CD-HIT → `modules/clustering/cdhit.nf`

Growth

- ☐ Extract DEMIC → `modules/growth/demic.nf`

Bin Scripts (`bin/`)

- ☐ Create `bin/filter_bins_quality.py`
- ☐ Create `bin/parse_kegg_results.py`
- ☐ Create `bin/parse_cazy_results.py`
- ☐ Create `bin/summarize_taxonomy.R`
- ☐ Create `bin/summarize_functions.R`

- ☐ Create `bin/plot_growth_rates.R`
- ☐ Create `bin/merge_metaphlan_tables.py`
- ☐ Make all scripts executable (`chmod +x bin/*`)

Scripts (`scripts/`)

- ☐ `quick_start_script.sh` → `scripts/run_pipeline.sh`
- ☐ Create `scripts/setup_databases.sh`
- ☐ Create `scripts/validate_installation.sh`
- ☐ Create `scripts/generate_samplesheet.py`
- ☐ Create `scripts/clean_workdir.sh`
- ☐ Create `scripts/download_test_data.sh`
- ☐ Make all scripts executable (`chmod +x scripts/*`)

Assets (`assets/`)

- ☐ Create `assets/samplesheet_schema.json`
- ☐ Create `assets/multiqc_config.yaml`
- ☐ Optional: Add `assets/logo.png`

Documentation (`docs/`)

- ☐ Create `docs/usage.md`
- ☐ Create `docs/output.md`
- ☐ Create `docs/parameters.md`
- ☐ Create `docs/troubleshooting.md`
- ☐ Create `docs/databases.md`
- ☐ Create `docs/containers.md`
- ☐ Create `docs/citation.md`

Root Documentation

- ☐ `readme_file.txt` → `README.md`
- ☐ `setup_guide.txt` → `SETUP.md`
- ☐ `getting_started.txt` → `GETTING_STARTED.md`
- ☐ `usage_examples.txt` → `EXAMPLES.md`
- ☐ `project_structure.txt` → `PROJECT_STRUCTURE.md`
- ☐ Create `CHANGELOG.md`
- ☐ Create `LICENSE`
- ☐ Create `.gitignore`

Test Files (test/)

- ☐ Create test/test_samplesheet.csv
- ☐ Create test/README.md
- ☐ Add .gitkeep to