Sub-workflows

This directory contains modular sub-workflows that can be run independently or as part of the main pipeline.

Available Sub-workflows

1. Taxonomy Profiling (`taxonomy_profiling.nf`)

Performs taxonomic profiling using MetaPhlAn.

Usage:

```bash
nextflow run workflows/taxonomy_profiling.nf \
    --input samplesheet.csv \
    --outdir results_taxonomy \
    --metaphlan_db /path/to/metaphlan_db \
    -profile docker
```

Outputs:

- Taxonomic profiles for each sample

- Bowtie2 alignment files

- Merged taxonomic abundance table

2. Functional Profiling (`functional_profiling.nf`)

Performs functional profiling using HUMAnN.

Usage:

```bash
nextflow run workflows/functional_profiling.nf \
    --input samplesheet.csv \
    --outdir results_functional \
    --humann_nucleotide_db /path/to/chocophlan \
    --humann_protein_db /path/to/uniref \
    -profile docker
```

Outputs:

- Gene family abundance tables

- Pathway abundance tables

- Pathway coverage tables

3. Assembly and Binning (`assembly_binning.nf`)

Performs metagenomic assembly, genome binning, and quality assessment.

Usage:

```bash
nextflow run workflows/assembly_binning.nf \
    --input samplesheet.csv \
    --outdir results_assembly \
    --checkm_db /path/to/checkm_data \
    --assembler megahit \
    --binning_tools metabat2,maxbin2 \
    --coassembly \
    -profile docker
```

Outputs:

- Assembled contigs

- Genome bins (MAGs)

- Bin quality reports (CheckM)

- Read mapping files

Using Sub-workflows in Main Pipeline

These sub-workflows are automatically integrated into the main pipeline. They are separated here for:

- Modularity and reusability

- Independent testing

- Easier maintenance

- Flexible execution

Parameters

Each sub-workflow accepts standard pipeline parameters. See the main `README.md` for complete parameter documentation.

Examples

Run Only Taxonomy and Function

```bash
bash

# Use main pipeline with skip flags
nextflow run main.nf \
    --input samples.csv \
    --outdir results \
    --skip_assembly \
    --skip_binning \
    --skip_growth_rates \
    -profile docker
```

Run Only Assembly and Binning

```bash
bash

# Use assembly_binning sub-workflow
nextflow run workflows/assembly_binning.nf \
    --input samples.csv \
    --outdir results_mags \
    --checkm_db /path/to/checkm_data \
    -profile docker
```

Chain Multiple Sub-workflows

```bash
bash

# Run workflows sequentially
nextflow run workflows/taxonomy_profiling.nf \
    --input samples.csv \
    --outdir results_step1 \
    -profile docker
```

```
nextflow run workflows/functional_profiling.nf \
    --input samples.csv \
    --outdir results_step2 \
    -profile docker

nextflow run workflows/assembly_binning.nf \
    --input samples.csv \
    --outdir results_step3 \
    -profile docker
```

Development

To add a new sub-workflow:

1. Create a new `.nf` file in this directory

2. Define the workflow with appropriate inputs and outputs

3. Include necessary modules from `modules/`

4. Document usage and outputs in this README

5. Optionally integrate into `main.nf`

Testing

Test sub-workflows with small datasets:

```bash
# Test with test data
nextflow run workflows/taxonomy_profiling.nf \
    --input test/test_samplesheet.csv \
    --outdir test_results \
    --metaphlan_db ~/databases/metaphlan_db \
    -profile test,docker
```