**Apscale manual**

Last updated: 11.07.2025, apscale version 4.0.5

Author: Dominik Buchner

**Table of contents**

[1. Introduction 3](#_Toc205477309)

[2. Installation 3](#_Toc205477310)

[3. Usage 4](#_Toc205477311)

[3.1. Creating a project 4](#_Toc205477312)

[3.2. Configuring the settings 4](#_Toc205477313)

[3.3. Adding data 4](#_Toc205477314)

[3.4. Running a module – Linear workflow 4](#_Toc205477315)

[3.4.1. Paired-end merging 4](#_Toc205477316)

[3.4.2. Primer trimming 4](#_Toc205477317)

[3.4.3. Quality filtering 4](#_Toc205477318)

[3.4.4. Dereplication 4](#_Toc205477319)

[3.5. Running a module – Modular workflow 4](#_Toc205477320)

[3.5.1. Denoising 4](#_Toc205477321)

[3.5.2. Swarm clustering 4](#_Toc205477322)

[3.5.3. Replicate merging 4](#_Toc205477323)

[3.5.4. Negative control filtering 4](#_Toc205477324)

[3.5.5. Read table generation 4](#_Toc205477325)

[3.6. Working with the read data store 4](#_Toc205477326)

[3.6.1. Apscale analyze 4](#_Toc205477327)

[3.6.2. Adding sample metadata 4](#_Toc205477328)

[3.6.3. Adding sequence metadata 4](#_Toc205477329)

[3.6.4. Correcting species names via GBIF 4](#_Toc205477330)

[3.6.5. Validating species occurrence via GBIF 4](#_Toc205477331)

[3.6.6. Uploading datasets to ENA 4](#_Toc205477332)

[3.6.7. Exporting read tables 4](#_Toc205477333)

# Introduction

Apscale is a metabarcoding pipeline that handles the most common tasks in metabarcoding pipelines like paired-end merging, primer trimming, quality filtering, denoising, swarm and threshold-based clustering as well as basic data handling operations such as replicate merging and the removal of reads found in the negative controls. It uses a simple command line interface and is configured via a single configuration file. To add metadata to the dataset, a simple, browser-based interface has been introduced in version 4.0. Apscale automatically uses the available resources on the machine it runs on while still providing the option to use less if desired. All modules can be run on their own or as a comprehensive workflow.

Several different programs are called within the workflow. These include vsearch (Rognes et al. 2016), cutadapt (Martin 2011) and swarm (Mahé et al. 2021). Please cite those accordingly, when using apscale.

Apscale (Buchner et al. 2022) has also been published and we are happy if we are cited too.

# Installation

# Usage

## Creating a project

## Configuring the settings

## Adding data

## Running a module – Linear workflow

## Running the full pipeline

## Paired-end merging

## Primer trimming

## Quality filtering

## Dereplication

## Running a module – Modular workflow

## Denoising

## Swarm clustering

## Replicate merging

## Negative control filtering

## Read table generation

## Working with the read data store

## Apscale analyze

## Adding sample metadata

## Adding sequence metadata

## Correcting species names via GBIF

## Validating species occurrence via GBIF

## Uploading datasets to ENA

## Exporting read tables