

Package: CentroidR (via r-universe)

October 17, 2026

Title CentroidR

Version 0.0.0.9004

Maintainer Adriano Rutz <adafede@gmail.com>

Description CentroidR provides the infrastructure to centroid profile spectra.

License AGPL (>= 3)

URL <https://github.com/adafede/CentroidR>,
<https://adafede.github.io/CentroidR>

BugReports <https://github.com/adafede/CentroidR/issues>

Depends R (>= 4.5.0)

Imports BiocParallel (>= 1.44.0), logger (>= 0.4.2), MsCoreUtils (>= 1.22.1), mzR (>= 2.44.0), optparse (>= 1.8.2), Spectra (>= 1.20.1)

Suggests tinytest (>= 1.4.3)

ByteCompile TRUE

Collate 'CentroidR-package.R' 'centroid_one_file.R'

Config/roxygen2/markdown TRUE

Config/roxygen2/version 8.1.0

Encoding UTF-8

Language en-US

LazyData TRUE

biocViews centroiding, spectra

X-schema.org-keywords centroiding, spectra

Config/pak/sysreqs cmake make libuv1-dev libnetcdf-dev libssl-dev

Repository <https://adafede.r-universe.dev>

Date/Publication 2026-10-17 07:48:51 UTC

RemoteUrl <https://github.com/adafede/CentroidR>

RemoteRef main

RemoteSha 1eaaabd66a44c798c3982584fb20c6d69699c837

Contents

centroid_one_file	2
-----------------------------	---

Index	5
--------------	----------

centroid_one_file	<i>Centroid an mzML file with configurable parameters</i>
-------------------	---

Description

This function processes an mzML file to apply centroiding with detailed controls for peak picking, smoothing, and noise estimation. It allows fine-tuning of MS1 and MS2 peak detection, optimizing spectral data analysis for various experimental needs.

Usage

```
centroid_one_file(
    file,
    pattern,
    replacement,
    min_datapoints_ms1 = 5L,
    min_datapoints_ms2 = 1L,
    mz_tol_da_ms1 = 0.0025,
    mz_tol_da_ms2 = 0.0025,
    mz_tol_ppm_ms1 = 5,
    mz_tol_ppm_ms2 = 5,
    mz_fun_ms1 = base::mean,
    mz_fun_ms2 = base::mean,
    int_fun_ms1 = base::max,
    int_fun_ms2 = base::max,
    mz_weighted = TRUE,
    time_domain = TRUE,
    intensity_exponent = 3
)
```

Arguments

file	character(1). Path to the input mzML file. Must be a valid, accessible mzML format file.
pattern	character(1). Regular expression pattern to match in the input file path, used for modifying the output file path.
replacement	character(1). Replacement string for altering the output file path based on the pattern.
min_datapoints_ms1	integer(1). Minimum datapoints to be considered for MS1 data. Must be at least 1. Default: 5L.

min_datapoints_ms2	integer(1). Minimum datapoints to be considered for MS2 data. Must be at least 1. Default: 1L.
mz_tol_da_ms1	numeric(1). m/z tolerance in Daltons for MS1 data. Default: 0.0025.
mz_tol_da_ms2	numeric(1). m/z tolerance in Daltons for MS2 data. Default: 0.0025.
mz_tol_ppm_ms1	numeric(1). m/z tolerance in parts per million (ppm) for MS1. Default: 5.
mz_tol_ppm_ms2	numeric(1). m/z tolerance in parts per million (ppm) for MS2. Default: 5.
mz_fun_ms1	function. Function to aggregate m/z values within each peak group (MS1). Ignored if mz_weighted = TRUE. Default: base::mean.
mz_fun_ms2	function. Function to aggregate m/z values within each peak group (MS2). Ignored if mz_weighted = TRUE. Default: base::mean.
int_fun_ms1	function. Function to aggregate peak intensities within each peak group (MS1). Default: base::max.
int_fun_ms2	function. Function to aggregate peak intensities within each peak group (MS2). Default: base::max.
mz_weighted	logical(1). If TRUE, uses intensity-weighted mean for m/z value aggregation. Default: TRUE.
time_domain	logical(1). If TRUE, uses square root for m/z weighting (TRUE for TOF). Default: TRUE.
intensity_exponent	numeric(1). If mz_weighted=TRUE, the exponent of the intensity for m/z weighting. Default: 3.

Details

MS1 and MS2 spectra are processed separately with their own tolerances, so the two levels can be tuned independently.

Within a spectrum, profile points are grouped by the m/z tolerances, a group is split back into separate peaks at deep valleys so that closely spaced peaks are not merged, and each group is reduced to a single peak. A group is only reduced when it holds at least min_datapoints_ms* points of non-zero intensity; below that it is reported as zero intensity and filtered out. Spectra left without any peak by this step are restored from the input, so no spectrum is lost.

File path modifications are supported via pattern and replacement. pattern is a regular expression, as in sub().

Value

logical(1). TRUE if the file was centroided, or TRUE if it was skipped because the output file already existed; FALSE if the input was missing or if centroiding failed. Invalid argument types raise an error instead of returning FALSE.

Author(s)

Johannes Rainer, Adriano Rutz

Examples

```
## Not run:  
# Example usage (requires an mzML file):  
# centroid_one_file(  
#   file = "raw_data.mzML",  
#   pattern = "\.mzML$",  
#   replacement = "_centroided.mzML",  
#   min_datapoints_ms1 = 5L,  
#   min_datapoints_ms2 = 1L  
# )  
  
## End(Not run)
```

Index

centroid_one_file, [2](#)