

# Package: lineagespot (via r-universe)

September 23, 2026

**Title** Detection of SARS-CoV-2 lineages in wastewater samples using next-generation sequencing

**Version** 1.17.1

**Date** 2026-08-20

**Description** Lineagespot is a framework written in R, and aims to identify SARS-CoV-2 related mutations based on a single (or a list) of variant(s) file(s) (i.e., variant calling format). The method can facilitate the detection of SARS-CoV-2 lineages in wastewater samples using next generation sequencing, and attempts to infer the potential distribution of the SARS-CoV-2 lineages.

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** false

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.1.2

**biocViews** VariantDetection, VariantAnnotation, Sequencing

**Imports** VariantAnnotation, MatrixGenerics, SummarizedExperiment, data.table, stringr, utils

**Suggests** BiocStyle, RefManageR, rmarkdown, knitr, testthat (>= 3.0.0)

**URL** <https://github.com/BiodataAnalysisGroup/lineagespot>

**BugReports** <https://github.com/BiodataAnalysisGroup/lineagespot/issues>

**BiocType** Software

**VignetteBuilder** knitr

**Config/testthat/edition** 3

**Config/pak/sysreqs**

make libbz2-dev libicu-dev liblzma-dev libpng-dev libxml2-dev libssl-dev xz-utils zlib1g-dev

**Repository** <https://bioc.r-universe.dev>

**Date/Publication** 2026-08-24 08:04:33 UTC

**RemoteUrl** <https://github.com/bioc/lineagespot>  
**RemoteRef** HEAD  
**RemoteSha** 1c3132ac4f90d93c26f53f0e17c2ee2858ff2545

**Contents**

|                            |          |
|----------------------------|----------|
| is_gff3 . . . . .          | 2        |
| lineagespot . . . . .      | 3        |
| lineagespot_hits . . . . . | 4        |
| list_input . . . . .       | 5        |
| list_vcf . . . . .         | 5        |
| merge_vcf . . . . .        | 6        |
| uniq_variants . . . . .    | 7        |
| <b>Index</b>               | <b>9</b> |

---

|         |                |
|---------|----------------|
| is_gff3 | <i>is_gff3</i> |
|---------|----------------|

---

**Description**

Identify whether a file is in GFF3 format.

**Usage**

is\_gff3(file)

**Arguments**

file                    Path to GFF3 file.

**Value**

result; TRUE if the input file is in GFF3 format, FALSE if not.

**Examples**

```
gff3_path <- system.file("extdata", "NC_045512.2_annot.gff3",  
  package = "lineagespot"  
)  
is_gff3(gff3_path)
```

---

|             |                    |
|-------------|--------------------|
| lineagespot | <i>lineagespot</i> |
|-------------|--------------------|

---

## Description

Identify SARS-CoV-2 related mutations based on a single (or a list) of variant(s) file(s)

## Usage

```
lineagespot(
  vcf_fls = NULL,
  vcf_folder = NULL,
  gff3_path = NULL,
  ref_folder = NULL,
  AF_threshold = 0.8
)
```

## Arguments

|                           |                                                                                         |
|---------------------------|-----------------------------------------------------------------------------------------|
| <code>vcf_fls</code>      | A character vector of paths to VCF files                                                |
| <code>vcf_folder</code>   | A path to a folder containing all VCF files that will be integrated into a single table |
| <code>gff3_path</code>    | Path to GFF3 file containing SARS-CoV-2 gene coordinates.                               |
| <code>ref_folder</code>   | A path to a folder containing lineage reports                                           |
| <code>AF_threshold</code> | A parameter indicating the AF threshold for identifying variants per sample             |

## Value

A list of three elements;

- Variants' table; A data table containing all variants that are included in the input VCF files
- Lineage hits; A data table containing identified hits between the input variants and the provided lineage reports
- Lineage report; A data table with computed metrics about the prevalence of the lineage of interest per sample.

## Examples

```
results <- lineagespot(
  vcf_folder = system.file("extdata", "vcf-files",
    package = "lineagespot"
  ),
  gff3_path = system.file("extdata",
    "NC_045512.2_annot.gff3",
    package = "lineagespot"
  ),
  ref_folder = system.file("extdata", "ref",
```

```

    package = "lineagespot"
  )
)

head(results$lineage.report)

```

---

|                  |                         |
|------------------|-------------------------|
| lineagespot_hits | <i>lineagespot_hits</i> |
|------------------|-------------------------|

---

### Description

Find overlapping variants with SARS-CoV-2 reference lineages

### Usage

```
lineagespot_hits(vcf_table = NULL, ref_folder = NULL)
```

### Arguments

|            |                                                                                                                   |
|------------|-------------------------------------------------------------------------------------------------------------------|
| vcf_table  | A tab-delimited table containing all variants for all samples. This input is generated by the merge_vcf function. |
| ref_folder | A path to lineages' reports                                                                                       |

### Value

A data table containing all identified SARS-CoV-2 variants based on the provided reference files

### Examples

```

variants_table <- merge_vcf(
  vcf_folder = system.file("extdata",
    "vcf-files",
    package = "lineagespot"
  ),
  gff3_path = system.file("extdata",
    "NC_045512.2_annot.gff3",
    package = "lineagespot"
  )
)

# use user-specified references
lineage_hits_table <- lineagespot_hits(
  vcf_table = variants_table,
  ref_folder = system.file("extdata", "ref",
    package = "lineagespot"
  )
)

```

---

|            |                   |
|------------|-------------------|
| list_input | <i>list_input</i> |
|------------|-------------------|

---

**Description**

Check the validity of input parameters from lineagespot function.

**Usage**

```
list_input(vcf_fls = NULL, vcf_folder = NULL, gff3_path = NULL)
```

**Arguments**

|            |                                                                                          |
|------------|------------------------------------------------------------------------------------------|
| vcf_fls    | A character vector of paths to VCF files.                                                |
| vcf_folder | A path to a folder containing all VCF files that will be integrated into a single table. |
| gff3_path  | Path to GFF3 file containing SARS-CoV-2 gene coordinates.                                |

**Value**

Return a character vector of paths to VCF files.

**Examples**

```
vcflist <- list_input(  
  vcf_folder = system.file("extdata", "vcf-files",  
    package = "lineagespot"  
  ),  
  gff3_path = system.file("extdata",  
    "NC_045512.2_annot.gff3",  
    package = "lineagespot"  
  )  
)
```

---

|          |                 |
|----------|-----------------|
| list_vcf | <i>list_vcf</i> |
|----------|-----------------|

---

**Description**

Identify VCF files from a group of files.

**Usage**

```
list_vcf(vcf_fls = NULL, vcf_folder = NULL, gff3_path = NULL)
```

**Arguments**

|            |                                                                                         |
|------------|-----------------------------------------------------------------------------------------|
| vcf_fls    | A character vector of paths to VCF files                                                |
| vcf_folder | A path to a folder containing all VCF files that will be integrated into a single table |
| gff3_path  | Path to GFF3 file containing SARS-CoV-2 gene coordinates.                               |

**Value**

- VCF list; A list where only VCF files are stored.

**Examples**

```
list_vcf_info <- list_vcf(
  vcf_folder = system.file("extdata", "vcf-files",
    package = "lineagespot"
  ),
  gff3_path = system.file("extdata",
    "NC_045512.2_annot.gff3",
    package = "lineagespot"
  )
)
print(list_vcf_info)
```

merge\_vcf

*merge\_vcf***Description**

Merge Variant Calling Format (VCF) files into a single tab-delimited table

**Usage**

```
merge_vcf(vcf_fls = NULL, vcf_folder = NULL, gff3_path = NULL)
```

**Arguments**

|            |                                                                                        |
|------------|----------------------------------------------------------------------------------------|
| vcf_fls    | A list of paths to VCF files                                                           |
| vcf_folder | A path to a folder containing all VCF file that will be integrated into a single table |
| gff3_path  | Path to GFF3 file                                                                      |

**Value**

A data table containiing all variants from each sample of the input VCF files

Examples

```
merge_vcf(  
  vcf_folder = system.file("extdata",  
    "vcf-files",  
    package = "lineagespot"  
  ),  
  gff3_path = system.file("extdata",  
    "NC_045512.2_annot.gff3",  
    package = "lineagespot"  
  )  
)
```

---

|               |                      |
|---------------|----------------------|
| uniq_variants | <i>uniq_variants</i> |
|---------------|----------------------|

---

Description

Lineage report for variants overlapping

Usage

```
uniq_variants(hits_table = NULL, AF_threshold = 0.8)
```

Arguments

- hits\_table      A tab-delimited table containing the identified overlaps/hits between the input files and the lineages' reports. This input is generated by the lineagespot\_hits function.
- AF\_threshold    A parameter indicating the AF threshold that is going to applied in order to identify the presence or not of a variant. This is used to compute the number of variants in a sample and eventually the proportion of a lineage.

Value

A data table with metrics assessing the abundance of every lineage in each samples

Examples

```
variants_table <- merge_vcf(  
  vcf_folder = system.file("extdata", "vcf-files",  
    package = "lineagespot"  
  ),  
  gff3_path = system.file("extdata",  
    "NC_045512.2_annot.gff3",  
    package = "lineagespot"  
  )  
)
```

```
lineage_hits_table <- lineagespot_hits(  
  vcf_table = variants_table,  
  ref_folder = system.file("extdata", "ref",  
    package = "lineagespot")  
)  
  
report <- uniq_variants(hits_table = lineage_hits_table)  
head(report)
```

# Index

`is_gff3`, [2](#)

`lineagespot`, [3](#)

`lineagespot_hits`, [4](#)

`list_input`, [5](#)

`list_vcf`, [5](#)

`merge_vcf`, [6](#)

`uniq_variants`, [7](#)