

Annotation comparison Vetea Jacot

Generated by Doxygen 1.9.8

1 Namespace Index	1
1.1 Namespace List	1
2 Namespace Documentation	3
2.1 CDScompR Namespace Reference	3
2.1.1 Detailed Description	3
2.1.2 Function Documentation	4
2.1.2.1 annotation_comparison()	4
2.1.2.2 write_results()	4
2.2 CDSmulticompR Namespace Reference	5
2.2.1 Detailed Description	5
2.2.2 Function Documentation	6
2.2.2.1 multicomp()	6
2.2.2.2 result_to_dict()	6
2.2.2.3 write_multi_results()	7
Index	9

Chapter 1

Namespace Index

1.1 Namespace List

Here is a list of all documented namespaces with brief descriptions:

CDScompR	Created on Thu May 2 11:57:29 2024	3
CDSmulticompR	Created on Fri Jul 5 13:18:44 2024	5

Chapter 2

Namespace Documentation

2.1 CDScompR Namespace Reference

Created on Thu May 2 11:57:29 2024.

Functions

- **write_results** (all_results, alt_name, out_dir, verbose=False)
This function writes to a new 'results.csv' file the results of the annotation comparison retrieved from the identities dictionary returned by the annotation_match function.
- **build_cluster_list** (read_ref, read_alt, dna_mol)
- **annotation_comparison** (ref_path, alt_path, out_dir, verbose=False, create_strings=False, exon_left_mode=False)
Main function of this program.
- **usage ()**
- **main ()**

Variables

- str **script_dir** = os.path.dirname(__file__) + "/python_util/"

2.1.1 Detailed Description

Created on Thu May 2 11:57:29 2024.

This script is used to compute the distance between two structural annotations of a same genome, one reference and one alternative annotation.

Author

: vetea, ranwez

It expects as input the paths to the annotation files (in GFF format), displays the computed distances between all annotation pairs, and creates a results CSV file detailing the loci comparisons between the annotations

2.1.2 Function Documentation

2.1.2.1 annotation_comparison()

```
CDScompR.annotation_comparison (
    ref_path,
    alt_path,
    out_dir,
    verbose = False,
    create_strings = False,
    exon_mode = False )
```

Main function of this program.

Given a reference and alternative path, gets the corresponding GFF files and compares the two annotations to return their information about their loci's comparison

Parameters

<i>ref_path</i>	Path of the GFF file describing the reference annotation
<i>alt_path</i>	Path of the GFF file describing the alternative annotation
<i>verbose</i>	If True, triggers display of more information messages. Default is 'False'
<i>create_strings</i>	Boolean indicating whether to use the 'old' comparison function (old_compare_loci, 'True') or the new one (compare_loci, 'False')
<i>exon_mode</i>	Boolean indicating if the main comparison structures read from the file should be coding sequences (CDS, False) or exons (True). Default is 'False' (CDS comparison)

See also

[compare_loci\(\)](#)
[old_compare_loci\(\)](#)

Returns

Returns a list of lists of dictionaries describing the comparison of the structure identity between the loci of each annotation

2.1.2.2 write_results()

```
CDScompR.write_results (
    all_results,
    alt_name,
    out_dir,
    verbose = False )
```

This function writes to a new 'results.csv' file the results of the annotation comparison retrieved from the identities dictionary returned by the annotation_match function.

See also

[annotation_match\(\)](#)

Parameters

<code>results</code>	A dictionary of list of list of dictionaries containing results of the annotation comparison, as returned by <code>annotation_match</code>
<code>alt_name</code>	String indicating the name of the original alternative annotation file
<code>verbose</code>	If True, triggers display of more information messages. Default is 'False'

See also

`annotation_match()`

Remarks

Results are written in CSV ('Comma-Separated Values') format

2.2 CDSmulticompR Namespace Reference

Created on Fri Jul 5 13:18:44 2024.

Functions

- [`result_to_dict` \(result\)](#)

Extracts the computed identity for each locus of the given result dictionary and add it to a new dictionary.
- [`write_multi_results` \(multi_results, ref_path, alt_paths\)](#)

Writes the results returned by the function `multicomp` into a results synthesis CSV file detailing the loci identity for each alternative.
- [`multicomp` \(ref_path, alt_paths, verbose, create_strings, exon_mode\)](#)

Compares all alternative annotations given (`alt_paths`) to the reference annotation (`ref_path`) by calling `annotation_sort` and appends the loci identities to a list.
- [`usage` \(\)](#)
- [`main` \(\)](#)

2.2.1 Detailed Description

Created on Fri Jul 5 13:18:44 2024.

This script is used to compute the distance between multiple structural annotations of a same genome, one reference and one or more alternative annotations.

Author

: vetea

It expects as input the paths to the annotation files (in GFF format), displays the computed distances between all annotation pairs, and creates results CSV files for each alternative and one global synthesis file

2.2.2 Function Documentation

2.2.2.1 multicompare()

```
CDSmulticompR.multicompare (
    ref_path,
    alt_paths,
    verbose,
    create_strings,
    exon_mode )
```

Compares all alternative annotations given (alt_paths) to the reference annotation (ref_path) by calling annotation_sort and appends the loci identities to a list.

See also

[annotation_sort\(\)](#)

Parameters

<i>ref_path</i>	Path to the reference annotation file
<i>alt_path</i>	Path to the alternative annotation file
<i>verbose</i>	If True, triggers display of more information messages. Default is 'False'
<i>create_strings</i>	Boolean indicating whether to use the 'old' comparison function (old_compare_loci, 'True') or the new one (compare_loci, 'False')
<i>exon_mode</i>	Boolean indicating if the main comparison structures read from the file should be coding sequences (CDS, False) or exons (True). Default is 'False' (CDS comparison)

See also

[compare_loci\(\)](#)
[old_compare_loci\(\)](#)

Returns

Returns the list of all loci identities (dictionaries) of all alternatives

2.2.2.2 result_to_dict()

```
CDSmulticompR.result_to_dict (
    result )
```

Extracts the computed identity for each locus of the given result dictionary and add it to a new dictionary.

Parameters

<i>result</i>	Result dictionary, as returned by CDScompR's annotation_compare() function
---------------	--

Returns

Returns a dictionary with the locus name as key, and the comparison identity as value

2.2.2.3 write_multi_results()

```
CDSmulticompR.write_multi_results (
  multi_results,
  ref_path,
  alt_paths )
```

Writes the results returned by the function multicomp into a results synthesis CSV file detailing the loci identity for each alternative.

See also

[multicomp\(\)](#)

Parameters

<i>multi_results</i>	List of loci identities (dictionaries), as returned by multicomp
<i>ref_path</i>	Path to the reference annotation GFF file

Index

annotation_comparison
CDScompR, [4](#)

CDScompR, [3](#)
annotation_comparison, [4](#)
write_results, [4](#)

CDSmulticompR, [5](#)
multicomp, [6](#)
result_to_dict, [6](#)
write_multi_results, [7](#)

multicomp
CDSmulticompR, [6](#)

result_to_dict
CDSmulticompR, [6](#)

write_multi_results
CDSmulticompR, [7](#)

write_results
CDScompR, [4](#)