Annotation comparison Vetea Jacot

Generated by Doxygen 1.9.8

Chapter 1

Namespace Index

1.1 Namespace List

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

Chapter 2

Namespace Documentation

2.1 CDScompare Namespace Reference

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

Functions

- **format_mismatch_zones** (zones)

  Formate une liste de coordonnées de mismatch en chaîne lisible.
- **write_results** (all_results, alt_name, out_dir)
- build_cluster_list_from_Locus (ref_loci, alt_loci, dna_mol)

  Build clusters of loci from reference and alternative annotations.
- **annotation_comparison** (str ref_path, str alt_path, str out_dir, bool mode_align)

  Compare two GFF annotations and write results to a file.
- **usage** ()
- **main** ()

Variables

- **script_dir** = os.path.join(os.path.dirname(os.path.abspath(__file__)), "python_util")

2.1.1 Detailed Description

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

Author

: vetea, ranwez

2.1.2 Function Documentation

2.1.2.1 build_cluster_list_from_Locus()

```
CDScompare.build_cluster_list_from_Locus (
            ref_loci,
            alt_loci,
            dna_mol )
```

Build clusters of loci from reference and alternative annotations.

Args: read_ref: Dictionary mapping chromosome_strand to list of Locus objects from reference annotation read↩_alt: Dictionary mapping chromosome_strand to list of Locus objects from alternative annotation dna_mol↩: Chromosome_strand key to process

Returns: List of Cluster objects

2.2 CDScompR Namespace Reference

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

2.2.1 Detailed Description

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

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

2.3 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** ()

2.3.1 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

2.3.2 Function Documentation

2.3.2.1 multicomp()

```
CDSmulticompR.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.

See also

annotation_sort()

Parameters

ref_path	Path to the reference annotation file
alt_path	Path to the alternative annotation file
verbose	If True, triggers display of more information messages. Default is 'False'
create_strings	Boolean indicating wether to use the 'old' comparison function (old_compare_loci, 'True') or the new one (compare_loci, 'False')
exon_mode	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

2.3.2.2   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

result	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

2.3.2.3   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

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

Index