Gene-by-Gene Report

Task Name: gene-by-gene

Suppose you have genomes and you want to characterize them. One of the ways to do that is to build a table of what genes are in each genome and what are not there.

1. Create a local BLAST db of your genome sequence/contigs. One db per one genome.
2. Create a file with sequences of genes you what to explore. This file will be the input file for the scheme
3. Setup location and name of BLAST db you created for the first genome.
4. Setup output files: report location and output file with annotated (with BLAST) sequence. You might want to delete the "Write Sequence" element if you do not need output sequences.
5. Run the scheme
5*. Run the scheme on the same input and output files changing BLAST db for each genome that you have.

As the result you will get the report file. With "Yes" and "No" field. "Yes" answer means that the gene is in the genome. "No" answer MIGHT mean that there is no gene in the genome. It is a good idea to analyze al
l the "No" sequences using annotated files. Just open a file and find a sequence with a name of a gene that has "No" result.

Parameters:

in - Input sequence file [Url datasets]

final-name - Annotation name used to compare genes and reference genomes (using 'blast_result' by dafault) [String]

exist-file - If a target report already exists you should specify how to handle that. Merge two table in one. Overwrite or Rename existing file (using 'Merge' by default) [String]

ident - Identity between gene sequence length and annotation length in per cent. BLAST identity (if specified) is checked after (using '90.0' percents by default) [Number]

out - Output report file [String]

blast-out - Location of BLAST output file [String]

search-type - Type of BLAST searches (using 'blastn' by default) [String]

db-name - Name of BLAST DB [String]

blast-path - Path to BLAST DB [String]

expected-value - This setting specifies the statistical significance threshold for reporting matches against database sequences (using '10.0' by default) [Number]

gapped-aln - Perform gapped alignment (using 'use' by default) [Boolean]

blast-name - Name for annotations (using 'blast_result' by default) [String]

tmpdir - Directory for temporary files (using UGENE temporary directory by default) [String]

toolpath - External tool path (using the path specified in UGENE by default) [String]

out-type - Type of BLAST output file (using 'XML (-m 7)' by default) [String]

Example:

```
ugene gene-by-gene --in=human_T1.fa --out=human_T1_report
```