

Inputs

Assembly
(.fasta)

Annotation
(.gff)

Length
threshold

Workflow

Search candidates
BLAST

Filtering

$90\% < \text{similarity} < 100\%$

$|\text{length BLAST match} - \text{query length}| < 0.2 \times \text{query length}$

Determine region in assembly

1. define upstream & downstream region
2. BLAST

Filtering

similarity = 100%

Update sequence

1. define start/stop in assembly
2. define new sequence (case 3 & 4: reverse complement)

Outputs

Corrected
assembly
(.fasta)

Log file
(.tsv)

