

Manual for MUMmer_filter

MUMmer_filter includes two scripts that are implemented in PERL. *mummer_strict_filter.pl* is designed to identify the orthologous regions between the reference and the query genome based on the comparison with the program NUCmer in the MUMmer package. *mummer_aln_extract.pl* is designed to write down the mapping of query genome against the reference in the FASTA format.

You may need to change the access permission to the scripts in Linux systems by:

```
Chmod 755 mummer_strict_filter.pl
```

```
Chmod 755 mummer_aln_extract.pl
```

The following commands are methods to compare one query genome with a reference and generate the mapping result.

```
nucmer --mum -p <prefix> <reference> <query>
```

```
delta-filter -l <prefix>.delta > <prefix>.delta2
```

```
mummer_strict_filter.pl <prefix>.delta2 > <prefix>.map
```

```
mummer_aln_extract.pl <prefix>.map > <prefix>.map.fas
```

The <prefix>.map file is in the same format as the delta files generated by NUCmer and can be handled by all tools from the MUMmer package. <prefix>.map.fas is a sequence file in FASTA format and represents the mapping of query against the reference. Gaps and missing regions in the query genome are represented as '-', while the small deletions in the query genome are presented as '.'. The region that are absent in the reference are not shown in the FASTA sequence. Thus, <prefix>.map.fas keeps the same length as the reference, such that reference and mappings from multiple queries can be easily compared.