

Get Scientific Name from Genbank Taxon ID

- Fri 23 Feb 2018 02:15:56 PM CET
- Johan.Nylander@{nrm|nbis}.se

taxid2sci.sh

Description

Get taxon name from Genbank taxon ID.

Input

File with taxids

Output

Writes TaxID, rank, and taxon name to stdout.

Notes

Sends and retrieves queries in batches. Output order can be different than input order! To ensure the same input/output order, use sorting (see Examples).

Usage

```
taxid2sci.sh inputfile
```

Examples

```
$ ./taxid2sci.sh taxids.txt
3055    species Chlamydomonas reinhardtii
3052    genus   Chlamydomonas
3047    species Dunaliella tertiolecta
3046    species Dunaliella salina

$ ./taxid2sci.sh <(sort taxids.txt) | sort
3046    species Dunaliella salina
3047    species Dunaliella tertiolecta
3052    genus   Chlamydomonas
3055    species Chlamydomonas reinhardtii
```

Requirements

- Internet connection
- Entrez Direct suite of software

taxid2lineage.pl

Description

Get full taxonomic lineage from Genbank taxon ID.

Input

File with taxids

Output

Writes taxon lineage with ranks to stdout.

Notes

Sends and retrieves queries one by one (slow).

Usage

```
taxid2lineage.pl inputfile
```

Examples

```
$ ./taxid2lineage.pl taxids.txt
taxid:3046 no_rank:cellular organisms superkingdom:Eukaryota kingdom:Viridiplantae phylum:Chlorophyta
taxid:3047 no_rank:cellular organisms superkingdom:Eukaryota kingdom:Viridiplantae phylum:Chlorophyta
taxid:3052 no_rank:cellular organisms superkingdom:Eukaryota kingdom:Viridiplantae phylum:Chlorophyta
taxid:3055 no_rank:cellular organisms superkingdom:Eukaryota kingdom:Viridiplantae phylum:Chlorophyta

$ ./taxid2lineage.pl taxids.txt | cut -f5- | sed 's/\w\+://g'
Viridiplantae Chlorophyta Chlorophyceae Chlamydomonadales Dunaliellaceae Dunaliella
Viridiplantae Chlorophyta Chlorophyceae Chlamydomonadales Dunaliellaceae Dunaliella
Viridiplantae Chlorophyta Chlorophyceae Chlamydomonadales Chlamydomonadaceae
Viridiplantae Chlorophyta Chlorophyceae Chlamydomonadales Chlamydomonadaceae Chlamydomonas
```

Requirements

- Internet connection
- Perl
- BioPerl