

MySQL

Trabalhando com os dados do NCBI Taxonomy

Prof. Dr. Dieval Guizelini
PPG em Bioinformática da UFPR

Criando o banco de dados

- Esses procedimentos devem ser feitos apenas uma vez:

```
create database taxonomy;  
use taxonomy;
```

Names

```
CREATE TABLE `ncbi_names` (  
  `tax_id` mediumint(11) unsigned NOT NULL default '0',  
  `name_txt` varchar(255) NOT NULL default '',  
  `unique_name` varchar(255) default NULL,  
  `name_class` varchar(32) NOT NULL default '',  
  KEY `tax_id` (`tax_id`),  
  KEY `name_class` (`name_class`),  
  KEY `name_txt` (`name_txt`)  
);
```

Importando os dados...

```
LOAD DATA INFILE 'C:\\projetos\\taxonomy\\names.dmp'  
INTO TABLE ncbi_names  
FIELDS TERMINATED BY '\\t|\\t'  
LINES TERMINATED BY '\\t|\\n'  
(tax_id, name_txt, unique_name, name_class);
```

nodes

```
CREATE TABLE `ncbi_nodes` (  
  `tax_id` mediumint(11) unsigned NOT NULL default '0',  
  `parent_tax_id` mediumint(8) unsigned NOT NULL default '0',  
  `rank` varchar(32) default NULL,  
  `embl_code` varchar(16) default NULL,  
  `division_id` smallint(6) NOT NULL default '0',  
  `inherited_div_flag` tinyint(4) NOT NULL default '0',  
  `genetic_code_id` smallint(6) NOT NULL default '0',  
  `inherited_GC_flag` tinyint(4) NOT NULL default '0',  
  `mitochondrial_genetic_code_id` smallint(4) NOT NULL default '0',  
  `inherited_MGC_flag` tinyint(4) NOT NULL default '0',  
  `GenBank_hidden_flag` smallint(4) NOT NULL default '0',  
  `hidden_subtree_root_flag` tinyint(4) NOT NULL default '0',  
  `comments` varchar(255) default NULL,  
  PRIMARY KEY (`tax_id`),  
  KEY `parent_tax_id` (`parent_tax_id`)  
);
```

Nodes

```
LOAD DATA INFILE 'nodes.dmp'  
INTO TABLE ncbi_nodes  
FIELDS TERMINATED BY '\t|\t'  
LINES TERMINATED BY '\t|\n'  
(tax_id, parent_tax_id, rank, embl_code, division_id,  
inherited_div_flag, genetic_code_id, inherited_GC_flag,  
mitochondrial_genetic_code_id, inherited_MGC_flag,  
GenBank_hidden_flag, hidden_subtree_root_flag, comments);
```

Referências

- NCBI Taxonomy
ftp://ftp.ncbi.nlm.nih.gov/pub/taxonomy/new_taxdump/taxdump_readme.txt
- Exemplo