`annmap` installation instructions

Tim Yates

April 30, 2024

Contents

1 Installation of the `annmap` database

The R package requires access to a MySQL database server running a current version of the `Annmap Database`. Database packs (complete with installation instructions) may be downloaded from `http://annmap.cruk.manchester.ac.uk/download`.

1.1 Pre-requisites

The following are required for the Annmap Database installation:

- MySQL Server v5.0.27+

- 20GB+ Free disk space

1.2 Installation

> **Important**
>
> **You do not need to pre-install Ensembl as you did before with the deprecated `exonmap` package**

First, you will need to download and unzip the `annmap` database for the species and version of Ensembl you are interested in studying. This can be downloaded from `http://annmap.cruk.manchester.ac.uk/download`

You will then need to create the database into which to load the data. To do this, simply log in to your MySQL server, and execute the following command[1]:

```
create database annmap_homo_sapiens_66 ;
```

Then, we have prepared 2 installation scripts to import the annmap database into your chosen MySQL server. These can be found in the root folder of the database package you first downloaded.

- Windows `importdb.bat`

- OS X/Linux `importdb.sh`

To utilise these files, you will need to edit them to set your username, select your password preferences (no password, prompt for password, or hardcoded password), and set the hostname of the server mysql is running on. All of these settings can be found towards the top of the installdb script.

Once run, the script will import all of the data into your MySQL server (this will take a few hours, due to the amount of data).

---

[1]This obviously assumes you are installing v56 of the homo_sapiens database. You'll need to replace the database name with the one you have downloaded.

2 Database configuration

2.1 Setting up the configuration folder

The `annmap` package requires a file called `databases.txt` to be placed inside a known folder. The package checks for this file in the following locations, and uses the first file it locates:

1. A location defined by the System Environment Variable `ANNMAP_HOME`.

2. A folder named `.annmap` inside the current user's home directory (On my OS X machine, this will be `/Users/tyates/.annmap/`)[2].

2.2 Setting up the `databases.txt` file

The `databases.txt` file is a tab-delimited file (a comma-delimited file is also accepted) in which each row defines a particular instance of an annmap database. The first row is always a list of tab-delimited column headings.

The columns; `port` and `password` may be left blank if you are using the default MySQL port (3306), or have a user account with no password required. Password can also be set to <ASK>, in which case, the system will ask you to enter your password before it connects to the server.

An example follows;

```
name    host            species       version  port  username   password
hs      localhost       homo_sapiens  66       4406  dbuser     dbpwd
mouse   anotherserver   mus_sapiens   66             userNoPwd
```

This example file (when put in the correct location) would set up two databases: one called 'hs' pointing to a v66 Homo Sapiens database running on the local machine (on a different port to the default); and another running v66 Mus Musculus on 'anotherserver'.

When you run `annmapConnect()` inside R, you should see both these appearing in a list for you to select which database to attach to. Or you can skip the menu by entering the database of choice into the command as a parameter: `annmapConnect('mouse')`.

---

[2]If you are not sure where your home directory is, you can find out through R by typing `Sys.getenv( "HOME" )` into an R console.