

Metaxa FAQ

This FAQ intends to answer some of the questions we have gotten about the Metaxa software package. Please look through this document before contacting me with a question. If the FAQ does not answer your question, you are very welcome to send an e-mail with the question to metaxa@microbiology.se

Does Metaxa work on Windows?

Short answer is no. I do not know anyone who has gotten Metaxa to work under Cygwin, so that is *not* a good option. However, Metaxa works very fine under virtualization software such as Virtual Box running a virtual Linux environment. Another good option is to find a colleague with a Mac and kindly ask him or her to run your sequences for you.

There are no people running either Mac or Linux on my department [or: I am afraid of the Linux people/Mac fanboys], what do I do?

I would recommend downloading Virtual Box (<http://www.virtualbox.org/>) and Bio-Linux (<http://envgen.nox.ac.uk/tools/bio-linux>). When you have gotten Virtual Box to run with Bio-Linux it should be relatively straightforward to install HMMER3 and Metaxa on the virtual machine using the Metaxa manual.

Metaxa runs with my specified input file but the output file says that no sequences are detected. What goes on?

Probably you have specified a file that does not exist. Currently, Metaxa does not warn the user that the input file is non-existent (which of course it will in a later version). It may also be that you are executing Metaxa in another directory than you intend. Check that the input file you desire to use is located in the current directory by typing “ls” on the command line.

Why is Metaxa using the old BLAST engine? Is there any plan to use BLAST+?

Metaxa currently uses the old BLAST engine (blastall) instead of the newer BLAST+ variant, primarily of time constraints. The BLAST+ output format requires a new parser to be written, which we have not gotten around to yet. The intention is to add BLAST+ support in a coming version of Metaxa; however, this feature is currently not of top priority. It is one of the smaller improvements that could be made to the software, but it requires a substantial amount of work to implement and make fully reliable.

The Metaxa installer script (install_metaxa) does not work. What can I do?

Under some operating system environments (primarily on some Linux distributions and on Mac OS X versions prior to 10.4) the installer script that comes with Metaxa might not work. Usually this happens because of lack of permissions, but it can also be caused by differing login file structures on various Linux distribution. The first thing that should be tested if Metaxa refuses to run after install, is to logout of the system completely, and then re-login again. Sometimes this solves the issue.

If this does not help, the installer script actually failed, and Metaxa must be installed manually. This can be done by moving into the Metaxa directory and copy all files starting

with “metaxa” into the preferred bin-directory, e.g. by typing “cp -r metaxa* ~/bin/”. After this has been done, you may need to edit your login script so Metaxa is added to your PATH. If your bin-directory is already present in your PATH, this should not be a problem, and you should be able to run Metaxa immediately after copying it into the bin-directory, e.g. by typing “metaxa -h”.

I have a multicore CPU, why is Metaxa so slow?

Most likely you have not specified Metaxa to use more than one CPU. Using the “--cpu [number]” option the speed of Metaxa improves dramatically.

Metaxa spawns a lot of HMMER processes that use up more CPUs than I specified using the --cpu option. What goes on?

Metaxa’s multi-threading system only partially takes into account the number of CPUs you specified using the --cpu option. If it is critical for you that Metaxa does not eat more CPU power than specified you should use the “--multi_thread F” option, which forces Metaxa to run the HMMER searches sequentially within the CPU limit.

I started Metaxa and it stopped early on in the progress with the following message: “Checking and handling input sequence data (should not take long)...” Why isn’t the software continuing?

You have not specified any input file (using the -i option). Therefore Metaxa has stopped, waiting for standard input. If this was unintended press control-C, and try again adding the input file option.

Can I use Metaxa to find SSU sequences in my newly sequenced and assembled genome?

This is not a good idea, as the intention of Metaxa never was to find SSU sequences in entire genomes, there are other software packages, such as RNAmmer, that do this much better. Instead, Metaxa is aimed at finding SSU sequences in fragmentary data, such as that generated in metagenomics or by community PCR.

Can I use Metaxa to determine the number of copies of SSU sequences in a certain species?

No. Metaxa will have problems with multiple SSUs on the same input query, and will therefore not be suitable for any kind of copy number detection approach.

What makes Metaxa 1.1 much faster than previous Metaxa versions?

A clever trick of black magic inspired by a blog post by HMMER’s creator Sean Eddy on hmmscan vs hmmsearch (<http://selab.janelia.org/people/eddys/blog/?p=424>). As the speedup comes from the extraction step, the speed increase will be largest for huge data sets with only a small proportion of actual SSU sequences (typically large 454 metagenomes).