



Jun 18, 2025

AMAS to Concatenate Sequences

DOI

<https://dx.doi.org/10.17504/protocols.io.j8nlk45ewg5r/v1>

Dakota Betz¹

¹ucsd

Rouse Lab



Dakota Betz

ucsd

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.j8nlk45ewg5r/v1>

Protocol Citation: Dakota Betz 2025. AMAS to Concatenate Sequences. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.j8nlk45ewg5r/v1>

License: This is an open access protocol distributed under the terms of the **[Creative Commons Attribution License](#)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited



Protocol status: Working

We use this protocol and it's working

Created: September 14, 2020

Last Modified: June 18, 2025

Protocol Integer ID: 42101

Keywords: concatenate sequences brief instruction, sequencematrix, sequences in the command line, sequence, amas package, ama, java

Abstract

Brief instructions on how to use the AMAS Package to concatenate sequences in the command line. Especially useful if your SequenceMatrix is no longer compatible with your version of Java.



Installation

- 1 Prerequisites: **Python 3** previously installed.
Download the AMAS package here:

Software

AMAS

NAME

Borowiec, M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660.

DEVELOPER

<https://pypi.org/project/amas/#files>

REPOSITORY

<https://pypi.org/project/amas/>

SOURCE LINK

Once downloaded, drag the AMAS folder into your Applications folder. You will need the path to this folder in order to proceed.

Concatenate

- 2 Use the command line to navigate to the folder containing your aligned files (one alignment per gene).

Note

Make sure the names for each terminal (sequence) are **IDENTICAL** for every gene, or the concatenation will not work.

- 3 Use the **concat** command to concatenate your files:
 - **-i** indicates the input files you want to concatenate
 - *gene1.fas* is your first aligned gene file (drag the file to the terminal to automatically include the path to it)
 - *gene2.fas* is your second aligned gene file (drag the file to the terminal to automatically include the path to it)

Note

Note: There are two genes listed here, but you can add as many as you want. Simply separate the paths/filenames by spaces.



- **-f** indicates the format of your files (in this case fasta, but you can also use nexus, phylip, etc)
- **-d** indicates the data type (DNA in most cases, use aa for amino acids)

Command

new command name

```
python3 /Applications/amas-1.0/amas/AMAS.py concat -i gene1.fas  
gene2.fas -f fasta -d dna
```

This will generate two files:

- "concatenated.out" is your concatenated file. Add the file extension ".fas" so you can open it in AliView or Mesquite to check that it worked (recommend AliView if you're staying in FASTA format).
- "partitions.txt" lists the partitions for your new concatenated file

Note

Recommendation: rename your output and partitions files so that they are specific to the current analysis you're running and/or match the input file names (e.g. if you're running a 6-gene concatenated analysis based on mafft aligned files on July 12, try something like "6mftCCT_Jul12_out.fas" and "6mftCCT_Jul12_partitions.txt").

Additional Options

- 4 If you would like to change the output format (something other than FASTA) use the argument **-u**, as in:

Command**new command name**

```
python3 /Applications/amas-1.0/amas/AMAS.py concat -i gene1.fas  
gene2.fas -f fasta -d dna -u phylip
```

- 5 You can specify the format of the partitions file, depending on the program you want to use. For RAxML:

Command**new command name**

```
python3 /Applications/amas-1.0/amas/AMAS.py concat -f fasta -d dna -i  
gene1.fas gene2.fas -u phylip --part-format raxml
```

If you open the partitions file, you can check that it is properly formatted. For RAxML, it should look something like this:

DNA, 16S = 1-605

DNA, COX1 = 606-1200

DNA, 28S = 1201-1800