Automated Simultaneous Analysis Phylogenetics (ASAP)

Quick-start tutorial for OS X users

This brief tutorial provides details for how to download, install, and use the ASAP software. The tutorial is done within the context of a customized set of FASTA formatted files, the most common data format used for ASAP analyses.

STEP 0: Install PAUP* and MUSCLE, and make sure that they are accessible from your \$PATH.

ASAP requires that the command-line version of PAUP* and MUSCLE are installed on your computer and are accessible from your path. You can verify this by typing 'which PAUP' or 'which muscle', the directory where PAUP and MUSCLE are installed should be returned. If this is not the case for either, contact your systems administrator for assistance.

Turing:ASAP sarkar\$ which PAUP /usr/local/bin/PAUP

Turing:ASAP sarkar\$ which muscle /usr/local/bin/muscle STEP 1: Download the ASAP installer for OS X

From the ASAP website (<u>http://www.ubio.org/ASAP</u>), download the latest ASAP installation package:



This will download the package, which you then double-click (on OS X 10.4 and above, it should automatically start the install program):



After you click on the Install button above, ASAP will be installed on your machine.

ASAP is a COMMAND-LINE SCRIPT; YOU WILL NOT SEE AN 'APPLICATION' in your Applications Folder.

You can verify installation by typing 'which ASAP' in the Terminal Application:



The terminal application should return '/usr/bin/ASAP'

STEP 2 Create Directory Structure and Populate '_partitions' directory

From the Method 1 section of the ASAP website, download the ASAP.dirstructure.zip file.

Method 1: Create a customized data set from FASTA formatted files

ASAP expects a particular directory structure wherein the data are found by the ASAP script and subjected to analyses. With a parent ASAP analysis directory, two sub-directories are required: "_partitions" and "_pendingAlignment". Take note that the prefixing '_' is required, and that the names are CASE-SENSITIVE. An example of the kind of directory structure that is required can be downloaded the (<u>ASAP.dirStructure.zip</u>). One approach might be to use this template directory structure as a place to start ASAP runs.

A separate FASTA file is required for each partition. The name of the partition in all the subsequent analyses will be derived from the FASTA file names, and the names of the taxa will be generated from the comment

AS. 0 0 0 Turing:- sarkar\$ cd Desktop Turing:Desktop sarkar\$ is ASAP.dirStructure Turing:Desktop sarkar\$ cd ASAP.dirStructure/ Turing:ASAP.dirStructure sarkar\$ is ASAP.settings ____partitions 17, 22 Turing:ASAP.dirStructure sarkar\$ __ug 20, 16 < 🕨 **DEVICES** Nam _____partitions 📑 Turing pendingAlignment pendingAlignme 🔮 iDisk ASAP.settings ► SHARED **PLACES** 🔜 Desktop 😭 sarkar Applications sarkarProjects manuscripts Documents Ownloads

Uncompress the directory onto your Desktop

The _partitions directory is where aligned files are placed, where each FASTA file contains a paritition. The _pendingAlignment directory is where unaligned files are places. ASAP.settings is a configuration file for customization of the ASAP script (it is recommended that this file only be manipulated by those familiar with Perl scripting).

Now, download the example _partitions directory:

An example of a _partition directory containing partitions that are correctly formated can be downloaded here <u>kere</u>.

The type and status of a given partition determines which directory you place it in:

- For partitions that *require sequence alignment*, place the respective FASTA files into the "_pendingAlignment" directory.
- For partitions that do not require sequence alignment (e.g., already aligned sequence or morphological data), place the respective FASTA files into the "_partitions" directory.

Unzip this file, and then copy the FASTA files into the _partitions directory within the ASAP.dirStructure folder created on your desktop.



The _partitions folder in your ASAP.dirStructure should now contain all the files copied:

	ASAP.dirStructure	(\bigcirc
		٩	\supset
	Name	Date Modified	
		Today, 16:45 EST	
	🗋 12S.fasta	Today, 16:45 EST	
1	🗋 16S.fasta	Today, 16:45 EST	
	ACRYS.fasta	Today, 16:45 EST	
	AHEM.fasta	Today, 16:45 EST	
	ALACI.fasta	Today, 16:45 EST	
	ALACX.fasta	Today, 16:45 EST	
	BCAS.fasta	Today, 16:45 EST	
	BCASI.fasta	Today, 16:45 EST	
	BHEM.fasta	Today, 16:45 EST	
	CYTB.fasta	Today, 16:45 EST	
	CYTB3rd.fasta	Today, 16:45 EST	
	CYTC.fasta	Today, 16:45 EST	
	GFIBI.fasta	Today, 16:45 EST	
	GFIBX.fasta	Today, 16:45 EST	
	📄 IRBP.fasta	Today, 16:45 EST	
	KCAS.fasta	Today, 16:45 EST	
	MORPH.fasta	Today, 16:45 EST	
	PANCR.fasta	Today, 16:45 EST	
	PRP1EX.fasta	Today, 16:45 EST	
	PRP1IN.fasta	Today, 16:45 EST	
	SINE.fasta	Today, 16:45 EST	
	VWF.fasta	Today, 16:45 EST	
	pendingAlignment	2007 Aug 17, 22:26 EDT	
	ASAP.settings	2007 Aug 20, 16:57 EDT	

STEP 3: Run ASAP

From a new Terminal window, go to your Desktop directory (using the cd command), then invoke the ASAP script giving it the directory to process:

```
Turing:~ sarkar$ cd
Turing:~ sarkar$ cd Desktop/
Turing:Desktop sarkar$ ASAP ASAP.dirStructure/
```

ASAP will then run. A number of new directories and files are generated as the script processes the data and runs the phylogenetic analyses through PAUP*:

- **______branchSupports.txt** this file contains the tab-delimited results of the branch supports at each node.
- _NEXUS.nex this file is the multi-partitioned dataset assembled from the _partition directory.
- **currList** this is an internal file name for ASAP to keep track of files that are in the _pendingAlignment or _partition directories. This prevents ASAP from repeating an already completed analysis. This can be overriden by entering "forceASAP" at the command line (e.g., ASAP asapDirName forceASAP).
- **getTrees.cmd** this file is a PAUP command file that is used to generate all the individual partition trees as well as the simultaneous analysis tree.
- **supportTests** this directory contains all the individual results from each of the performed tests
- **trees** this directory contains the results of the individual partition and simultaneous analysis trees (three files exist for each tree search performed: AllTrees, Bootstrap, and Consensus).