SEEVA

(Spatial Evolutionary and Ecological Vicariance Analysis)

MANUAL version 1.00

Lena Struwe & Einar Heiberg

List of contents:

Legal Disclaimer

This SOFTWARE PRODUCT is provided by THE PROVIDERS "as is" and "with all faults." THE PROVIDERS make no representations or warranties of any kind concerning the safety, suitability, lack of viruses, inaccuracies, typographical errors, or other harmful components of this SOFTWARE PRODUCT. There are inherent dangers in the use of any software, and you are solely responsible for determining whether this SOFTWARE PRODUCT is compatible with your equipment and other software installed on your equipment. You are also solely responsible for the protection of your equipment and backup of your data, and THE PROVIDER will not be liable for any damages you may suffer in connection with using, modifying, or distributing this SOFTWARE PRODUCT.

(OK, that was the legal stuff, now on to the science.)

Updated: October 24, 2010

SOURCES AND DOCUMENTATION

How to cite this software and method

If you use this software you are required to cite the website http://seeva.heiberg.se as well as the publication describing the SEEVA methodology, Struwe et al., (MS), see below.

Publications

Heiberg, E. 20XX. SEEVA ver. X.XX. Software for Spatial Evolutionary and Ecological Vicariance Analysis. Available from the author at http://seeva.heiberg.se. [citation for the software]

Heiberg, E. & L. Struwe. 20XX. SEEVA manual, ver X.XX. On-line publication, Rutgers University. Available from the authors at http://www.rci.rutgers.edu/~struwe/seeva [citation for the manual]

Struwe, L., P. E. Smouse, E. Heiberg, S. Haag, & R. G. Lathrop. MS (2010). Spatial evolutionary and ecological vicariance analysis (SEEVA), a novel approach to biogeography and speciation research, with an example from Brazilian Gentianaceae. Journal of Biogeography, October 2010). [official citation for software and methodology]

[The SEEVA website lists additional publications and projects that have used this method.]

Websites

SEEVA software website: http://seeva.heiberg.se (download software here)
Website for general information about SEEVA: http://www.rci.rutgers.edu/~struwe/seeva

Contact information

Dr. Einar Heiberg, Lund University, Lund, Sweden. E-mail: einar@heiberg.se
Dr. Lena Struwe, Rutgers University, New Brunswick, NJ, USA. E-mail: struwe@aesop.rutgers.edu

SEEVA Listserve

The SEEVA mailing list/listserve can be used to ask questions, disseminate information, and announce new software versions and other developments. Lena Struwe is the administrator for the list. Only members of the list can post questions and these will reach everybody on the list. You can sign up for the SEEVA mailing list here:
https://email.rutgers.edu/mailman/listinfo/seeva_list

Feedback on methodology and software

This software and method is a work in progress. We appreciate any feedback and suggestions for improvement, see contact information above. If you have problems with running the program, please provide detailed information so we can try to help you.

INTRODUCTION AND METHODOLOGY

Please see Struwe et al. (submitted) for a detailed overview and justification of SEEVA methodology and mathematical formulas.

In contrast to vicariance biogeography, which assumes geographic separation of populations, the Spatial Evolutionary and Ecological Vicariance Analysis approach allows researchers also to look at ecological vicariance (differences) of sympatric and allopatric species and clades. This method can utilize GIS-derived dataset of collection-associated ecological and environmental data in combination with phylogenetic data to investigate trends in speciation using statistical methods with spatial interpretations. The method can also be used for other kinds of comparisons between groups and clades, in areas such as coevolution, diseases, morphological evolution, and niche comparisons.

Generally, SEEVA works by using measurements gathered from individuals of species or populations, and these measurements are then analyzed statistically for differences between groups (species) and/or clades. Two statistical test are being employed, the Divergence Index (to measure differences between groups or clades) and Fisher's Exact test (the latter to provide a *p*-value for tests with small sample sizes).

Environmental variables (data columns) are divided into categories either as non-ordered, qualitative categories (e.g., soil types) or ordered, quantitative sections representing subsets of the total variation in the variable (e.g., precipitation amounts). A statistical test is performed to investigate if the distribution of species or monophyletic clades in different categories of environmental variables shows a random or non-random pattern, with taxonomic group vs. counts for collections for the categories for one variable in an X x Y multi-way table.

Example of X x Y multiway table showing skewed character state distributions for two different groups using 4 categories (states) for one variable. The numbers inside the table are number of observations, i.e., collections.

number of observations	category 1	category 2	category 3	category 4
Group 1/clade A	0	10	16	21
Group 2/ clade B	10	15	9	0

A non-random (skewed) pattern is a stronger association for an environmental character state(s) with a specific group/clade, both historically (evolutionary) and presently. The classification of the environmental variables is rather coarse, but these tests provide a way of looking for broad patterns.

The Divergence Index (*D*) is a measure from 0 to 1 on how skewed a distribution is between two groups or clades, and *D* is independent of sample-size and can therefore be compared between groups/clades as well as variables. *D* is also independent from p-values, which indicates if the data distribution at a node is statistically significantly different between the two groups/clades, or not. We suggest that measures are taken to adjust the significance value of the *p*-value (usually

0.05) to adjust for multiple sampling from the same dataset, using for example the Bonferroni correction.

We have previously used another index of skewness, the Impact Index (*I*), but the Index of Divergence (*D*) is superior to *I* and replaces this. However, the Impact Index is still listed in the result outputs from the SEEVA software (as are chi-square statistics).

The SEEVA method can be enhanced when combined with biogeographic and phylogeographic spatial analysis, ancestral area analysis, dating methods, geographic mapping of populations, endangered species analysis, and ecological niche analysis. The method will work on any kind of data including absence/presence of diseases, morphological or phytochemical measurements, pollinator type, color morphs – as long as individuals are measured.

The species can be grouped in two ways for the SEEVA analysis: 1) species-by-species (Manual Analysis); and 2) by sister clades (Tree-based). The latter approach includes phylogenetic information, since species data from two (or more, if a polytomy) clades will be compared and analyzed, and environmental trends and reactions over the time of the evolution of a group can be assessed by comparing impact values between nodes.

Caveats and notes: The method works with a minimum of one record for each species, however, results based on such a small sample should be evaluated with caution. In general, correlations between divergence and environmental variables can be inferred as trends and tendencies within phylogenetic lineages, and not as the definite cause for the divergence until further research. It should not be assumed that environmental variables are independent, in fact, many of them are not, but an assumption of independency is not necessary for this analysis.

DOWNLOAD AND INSTALLATION

About the program
The program is written in Matlab, and is distributed as a stand-alone application, so it can be run on any Windows computer even if you do not have Matlab installed. If you want the SEEVA software as a regular Matlab program, please contact Einar Heiberg.

Computer requirements:
PC running Windows 7, Windows Vista, Windows XP, or Windows NT with at least 300 Mb free space on a hard drive and 256 MB RAM memory.

Step 1: Download the manual
Download the SEEVA software manual (this document) on one of these two websites: http://www.rci.rutgers.edu/~struwe/seeva or http://seeva.heiberg.se.

Step 2: Install Matlab Component Runtime
Matlab Component Runtime is a required program to run the SEEVA software as a standalone version. This is essentially a small Matlab installation, but it can only be used to run precompiled applications, and you need to have at least version 7.6.

If Matlab Component Runtime is already installed on your computer you can omit this step, but double check which version you have. In Windows XP, you can check this in the Control Panel, under Add/Remove programs. Look for MATLAB Component Runtime 7.8 on the list (size 466 MB). If your Matlab Component Runtime version is less than ver. 7.6, then you need to uninstall the old version first and then install the newer version. Uninstalling is done by clicking on the program name in the list and select "Remove". You need to be logged in as administrator to do this.

How to install the Matlab Component Runtime software on your computer:

1) Log in as administrator on your computer.
2) Download the file MCRInstaller_R2008a.exe (large file!!! 200+MB) from http://seeva.heiberg.se
3) Double-click on the downloaded file and install it. You might be required to restart your computer after installation.

Step 3: Download SEEVA software

1) Download the file seeva_vX.exe. [X stands for the version number of the software.] from http://seeva.heiberg.se
2) Create a new folder (we suggest the name SEEVA) on your computer and place the file seeva.exe in that folder.

Step 4: Download example and template files (optional)
One large (*Macrocarpaea*, +50 species) and one small taxon example (*Tachia*, 13 species) are provided as templates and can be downloaded on the SEEVA websites. Note that this data does not reflect complete scientific data and well-supported phylogenetic hypotheses. These data matrices are only templates for formatting, training, and testing by SEEVA users.

PREPARING THE INPUT FILES

Prepare the input file: Excel file with individualized data

SEEVA does not provide the data that you want to analyze, you need to provide this yourself in a spreadsheet (Excel, etc.) format. For each (geolocated) individual, add data to a spreadsheet either as quantitative measurements (e.g., rainfall amount) or qualitative (e.g., soil type). We use ArcGIS to do this. First we import the spreadsheet with all individuals, all with latitude and longitude, and map them onto baselayers in ArcGIS. Then we pull out environmental data from many baselayers of environmental data from sources such as USGS, WORLDCLIM, etc., and add this data for each collection back into the spreadsheet. The macro we used for this is available for this at the SEEVA website and was developed by Scott Haag: http://www.rci.rutgers.edu/~struwe/seeva

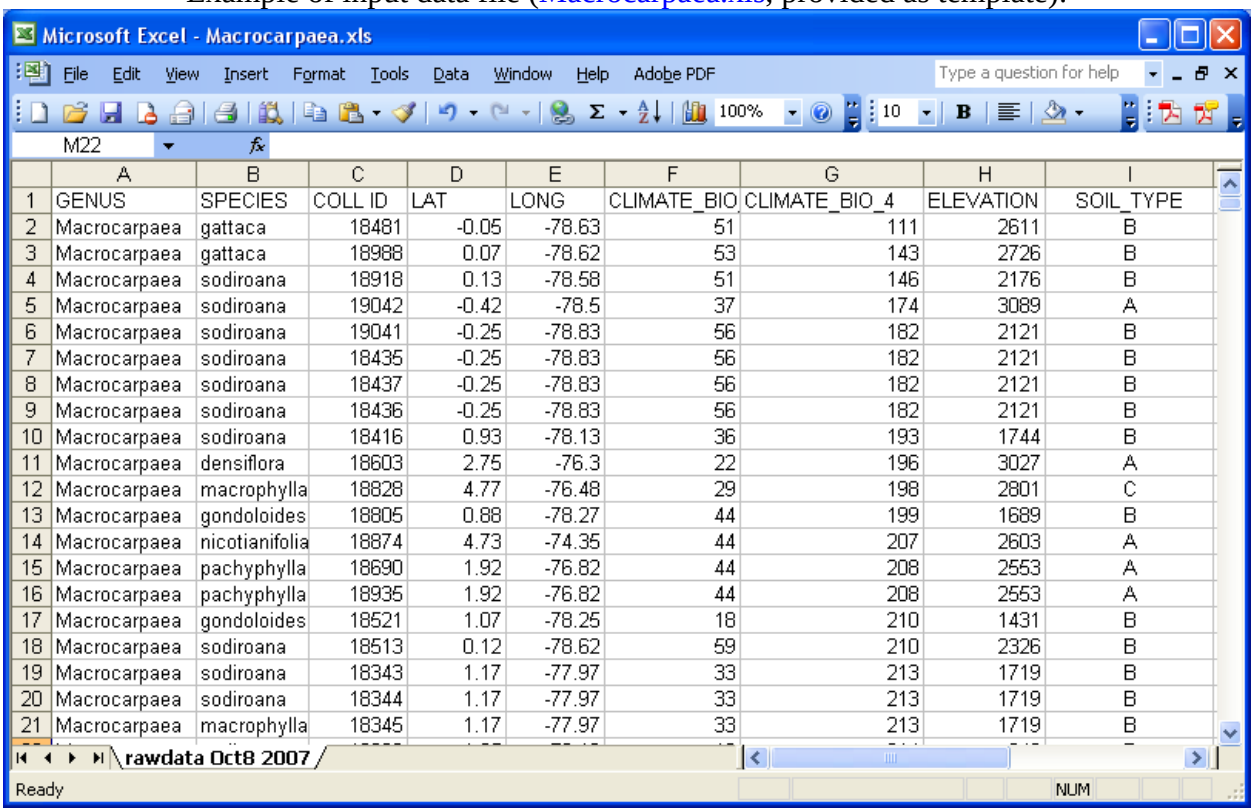
The Excel file should have one column named SPECIES, which will contain the species name for each collection (individual). The species column has to be to the left of the columns of variables you want to analyze. You can have additional columns with other data such as collector, year, country, collection ID, latitude, longitude, etc. included in the sheet, and this will not affect the SEEVA analysis, since you can pick and choose in the SEEVA software which columns to analyze.

The top (first) row is the name of each column for SEEVA, so make sure you name each column differently and have all columns named. The SEEVA program only reads the first worksheet in the Excel workbook.

Do not have empty rows or columns inside you spreadsheet. You can have empty cells for some individuals where you lack data for that particular individual for that particular variable.

Instead of environmental data you can use any type of data (toxicity, gene expression, chemical, morphological, or other measurements of traits, quantitative or qualitative). Instead of individuals you can use other groupings, but remember that you need a specific number in each variable cell (for example, 2.34 or A), not a range of numbers (2.3-2.5 = min-max) or polymorphisms (A, B).

Example of input data file (Macrocarpaea.xls, provided as template):



Microsoft Excel - Macrocarpaea.xls

	A	B	C	D	E	F	G	H	I
1	GENUS	SPECIES	COLL ID	LAT	LONG	CLIMATE_BIO	CLIMATE_BIO_4	ELEVATION	SOIL_TYPE
2	Macrocarpaea	gattaca	18481	-0.05	-78.63	51	111	2611	B
3	Macrocarpaea	gattaca	18988	0.07	-78.62	53	143	2726	B
4	Macrocarpaea	sodiroana	18918	0.13	-78.58	51	146	2176	B
5	Macrocarpaea	sodiroana	19042	-0.42	-78.5	37	174	3089	A
6	Macrocarpaea	sodiroana	19041	-0.25	-78.83	56	182	2121	B
7	Macrocarpaea	sodiroana	18435	-0.25	-78.83	56	182	2121	B
8	Macrocarpaea	sodiroana	18437	-0.25	-78.83	56	182	2121	B
9	Macrocarpaea	sodiroana	18436	-0.25	-78.83	56	182	2121	B
10	Macrocarpaea	sodiroana	18416	0.93	-78.13	36	193	1744	B
11	Macrocarpaea	densiflora	18603	2.75	-76.3	22	196	3027	A
12	Macrocarpaea	macrophylla	18828	4.77	-76.48	29	198	2801	C
13	Macrocarpaea	gondoloides	18805	0.88	-78.27	44	199	1689	B
14	Macrocarpaea	nicotianifolia	18874	4.73	-74.35	44	207	2603	A
15	Macrocarpaea	pachyphylla	18690	1.92	-76.82	44	208	2553	A
16	Macrocarpaea	pachyphylla	18935	1.92	-76.82	44	208	2553	A
17	Macrocarpaea	gondoloides	18521	1.07	-78.25	18	210	1431	B
18	Macrocarpaea	sodiroana	18513	0.12	-78.62	59	210	2326	B
19	Macrocarpaea	sodiroana	18343	1.17	-77.97	33	213	1719	B
20	Macrocarpaea	sodiroana	18344	1.17	-77.97	33	213	1719	B
21	Macrocarpaea	macrophylla	18345	1.17	-77.97	33	213	1719	B

rawdata Oct8 2007

Ready

NUM

Prepare the tree file (for phylogeny-based analyses)

A tree file is necessary if you want to base your analysis on evolutionary topologies and comparisons of node-by-node patterns (see below under Running SEEVA, step 2B). If you want to compare 2-5 groups/taxa without any particular topology, you can do a manual analysis without a tree file (see below under step 2A).

The tree file should be in the NEXUS format, and can include both a matrix and tree block (only the tree block will be read by SEEVA). The default is for the NEXUS tree file to end with .nex, but you can also open .txt files with SEEVA. Take a look at the example files if you are unsure about how to format your tree input file. Taxon names in the tree file must correspond exactly to the Excel sheet's species name column (watch for extra spaces that might be invisible).

You can have several trees in your treefile, but you can only analyze one at the time in SEEVA – the program will ask when you open the tree file which one you want to use. We do not recommend that you have treefiles with hundreds of trees, select a few only.

Phylogenetic trees can have polytomies in them, but remember that this will affect your degrees of freedom when you analyze a variable at that node (more groups and more classes will always increase the degrees of freedom). We are unsure what the upper limit is when it comes to number of possible branches in a polytomy, this has not yet been tested. However, the more resolved your tree is, the more you can tell from the results anyway.

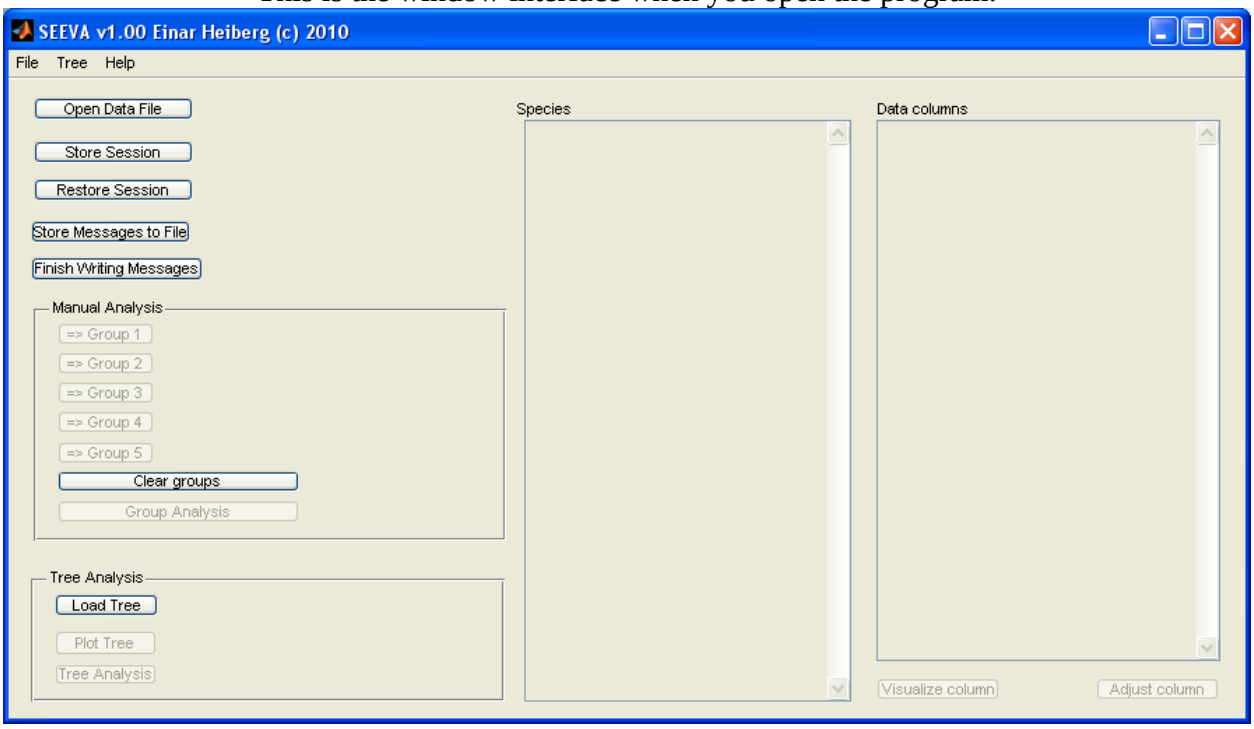
The tree file can have more species in it than the input data file, for example outgroups can be included. These will not be included in the SEEVA analysis unless you provide data for them, but will be included in the plotted trees.

RUNNING SEEVA

Start the program

Double click on seeva_vX.exe to start SEEVA, where X in the filename stands for the version you are running. If you get a warning about the publisher being unknown, ignore this and click OK]. The first time you open the program it will take some extra time because Matlab is loading in the background. Be patient. First a MS-DOS window opens (black background), which will contain the programs logfiles of actions, commands, and error messages. The second window that opens is the actual SEEVA program interface. Keep both windows open (you can minimize the MS-DOS window if you like).

This is the window interface when you open the program:



Overview of the window interface and options

Menu options:

File:

Open data file
Clear data and tree [=start over]
Quit

Tree:

Load tree [=load tree file]
Tree Analysis [run the SEEVA analysis]
Plot tree [=show tree]

Help:

Home page
About
Submit bug report

Main (Top Left) Buttons:

Open Data File: Opens Excel sheet with data.

Store Session: Saves your raw data and all settings in a Matlab log file for future reload into the SEEVA software.

Restore Session: Loads your Matlab log file for reloading a previous SEEVA analysis.

Store Messages to File: Opens a text log file of text shown in the MS-DOS window (good for troubleshooting and remembering commands)

Finish Writing Messages: Closes text log file

Columns (to the right):

Species: Species names pulled from data file under column name 'Species'

Data columns: Lists columns of variables pulled from the data file, listed in order from left to right in the data file. Type of data is listed after each name.

Visualize column: Shows the selected division of a variable into categories for the SEEVA analysis.

Adjust column: Gives you a choice of changing the categories of a variable (numerical, character, number of quintiles)

Manual Analysis

Group 1, 2, etc.

Clear Groups: Clears the selection you have made previously

Group Analysis: Runs the SEEVA analysis based on your selection of groups and variables.

Tree Analysis

Load Tree: Loads tree file

Plot Tree: Shows tree

Analyze Tree: Runs SEEVA analysis of all nodes in a tree for the variables that have been selected

Note: Buttons and commands might be grayed out if they are not applicable to the data you have loaded.

Saving a log file and storing your session for future use in Matlab

SEEVA provides two ways to store files that record your analysis in SEEVA.

The log file of the MS-DOS window, which contains commands, error messages, parsing results, etc., can be saved as a .txt file by using the *Store Messages to File* button. When you want to close this log output file, click the *Finish Writing Messages* button. You should open this logfile before you open your data file.

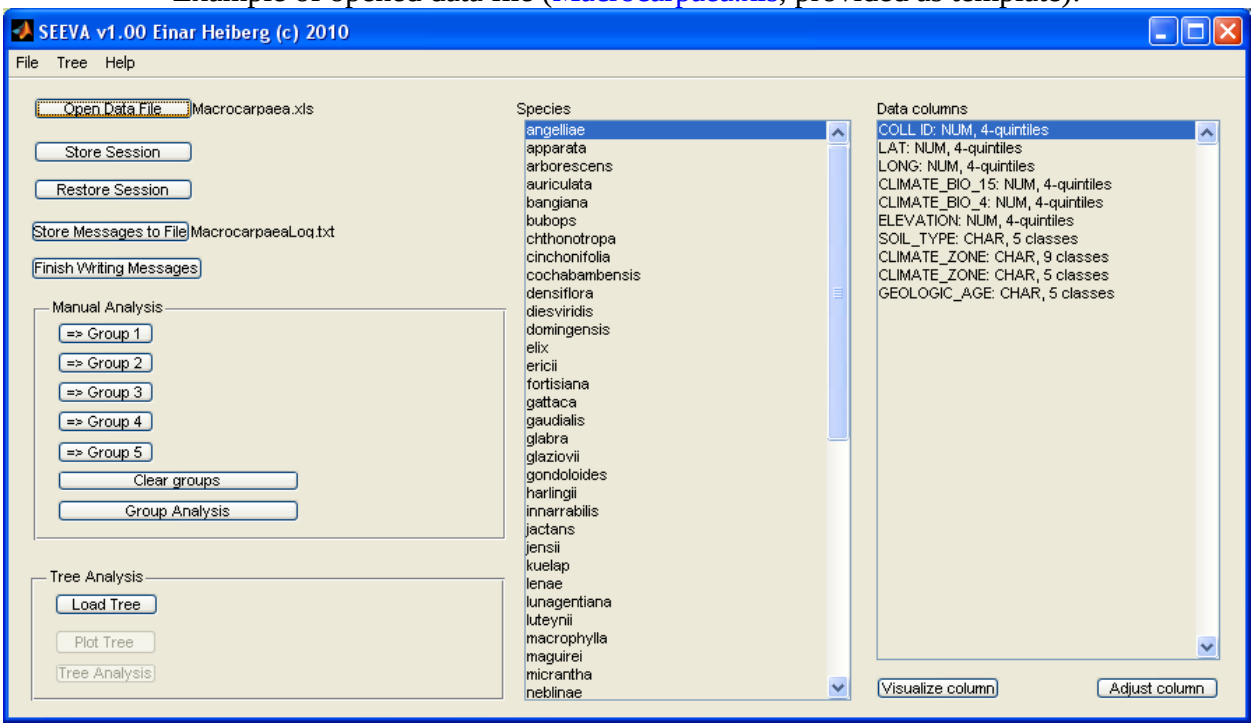
Your whole SEEVA session, including rawdata and settings for your columns, can be saved for future reload into SEEVA by clicking the *Store Session* button. This will save a Matlab file of your data and results. When you want to load a previous session file, click on the *Restore Session* button, and open your previously saved Matlab file. You should store the session at the end of a SEEVA run, before you close the program, to save your most recent settings.

Open Data File

Load your data by clicking on *Open Data File* and selecting your excel file with your input data (excel data sheet). Your species names should show in the SEEVA window in the *Species* window, and all of your data columns should be pulled in from the excel data sheet and shown in

the *Data Columns* window The data columns will list your variables, but also other columns in the Excel datasheet.

Example of opened data file (Macrocarpaea.xls, provided as template):



Options for data columns (variables)

In the data column window you see the list of columns with variables that can be analyzed by SEEVA. After each name (pulled from the top row of your Excel input file), you will see either NUM or CHAR, and the type of division that applies to that variable. NUM is numeric (quantitative data), and CHAR is character (qualitative) data. The data from each numeric variable can be divided up into states (categories) in different ways, for example as 4-quintiles (4 quartile states, each with the same number of observations).

Note: Shown in the *Data Column* window are the columns to the right of the Species column in your data analysis. If you have columns with information (genus, family, collection data, etc.) that you do not want to show up here, put those columns to the left of the species column in your input data file.

The default in SEEVA is 4 states, but you can change this for each variable by highlighting the column name and then clicking on *Adjust Column*. This will provide you with a new window where you can select the type of division you want for your numeric data (including treating it as qualitative character data, for example if your different soil types are indicated by numbers and not letters). Closing the window without clicking on a selection will keep the original setting (and give you a message that that particular command was aborted).

Note that these settings apply to your whole input file. If you want to do a partial analysis of just one subclade/subgroup and have your data categories apply only to that subclade's data, you will have to prepare a separate input file for that subclade. Since numeric categories are automatically calculated based on the total distribution of measurements, and the

total distribution will most likely differ between the whole group and a part of it, you will get different data splits if you look at the whole dataset or only a part of it.

The options for any number data are:

Character data (classes): Your data will be treated as if each number is a separate category, i.e., "3, 3, 4, 56, 56, 3, 3, 56" will be three categories, type 3, type 4, and type 56. This is useful for columns for characters such as such as vegetation type, soil type, and number of floral parts. For all quintile splits, the data will be split so that there are equal numbers of observations in each category (for example, for 4-quintiles it will be 4 categories with 25% of observations in each category). In cases of many repetitive numbers (3, 3, 3, 4, 4, 4, 4, 5, 6, 77, 78, etc.) this is not always possible, so the percentage distributions might sometimes be slightly different.

Numeric two levels (median split): Your data will be split into **two** categories, at the median split..

Numeric 3-quintiles (tertiles): Your data will be split into **three** categories.

Numeric 4-quintiles (quartiles): DEFAULT. Your data will be split into **four** categories.

Numeric 5-quintiles: Your data will be split into **five** categories.

Numeric 6-quintiles (sextiles): Your data will be split into **three** categories.

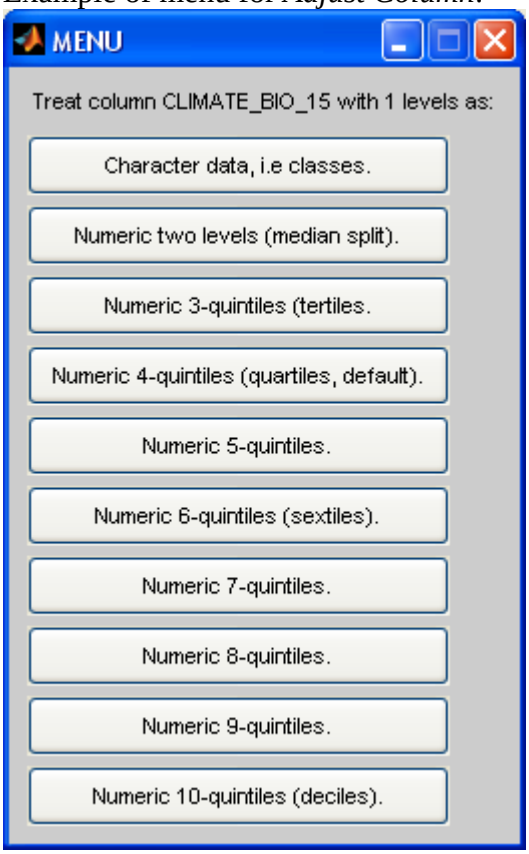
Numeric 7-quintiles: Your data will be split into **three** categories.

Numeric 8-quintiles: Your data will be split into **three** categories.

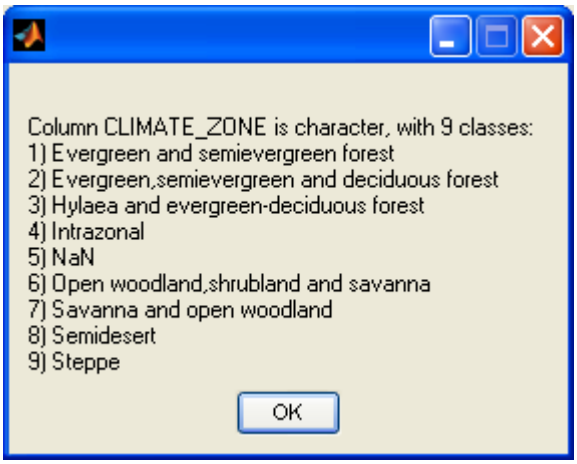
Numeric 9-quintiles: Your data will be split into **three** categories.

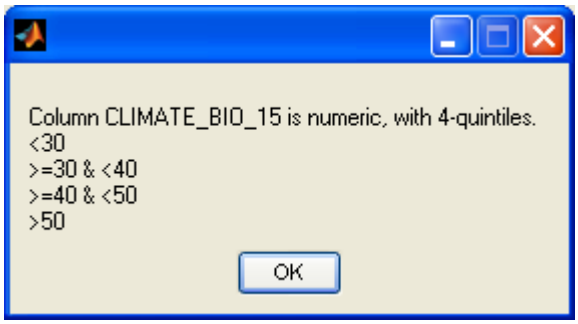
Numeric 10-quintiles (deciles): Your data will be split into **three** categories.

Example of menu for *Adjust Column*:



To check how your data is parsed into categories, highlight the data column name and then click on Visualize Column. A new window opens up and shows you the different categories applied to your data by SEEVA. Example of menus for *Visualize Column* (qualitative data to the right, quantitative data to the left):





NaN refers to empty data cells in your column, i.e, collections without any data for this variable. The "&" in the Visualize Column should be read as 'to'. The qualitative data categories are pulled directly from the text in the columns, so make sure you use exactly the same spelling and text for the same categories for qualitative data (no extra spaces, etc.).

Manual SEEVA analysis for comparison of groups (not based on phylogenetic trees) [MANUAL ANALYSIS]

This method is for comparing user-defined groups of species that are not based on the grouping of species into clades and subclades in a phylogenetic tree. This could be species from different habitats, different pollination or dispersal syndromes, different continents, etc.

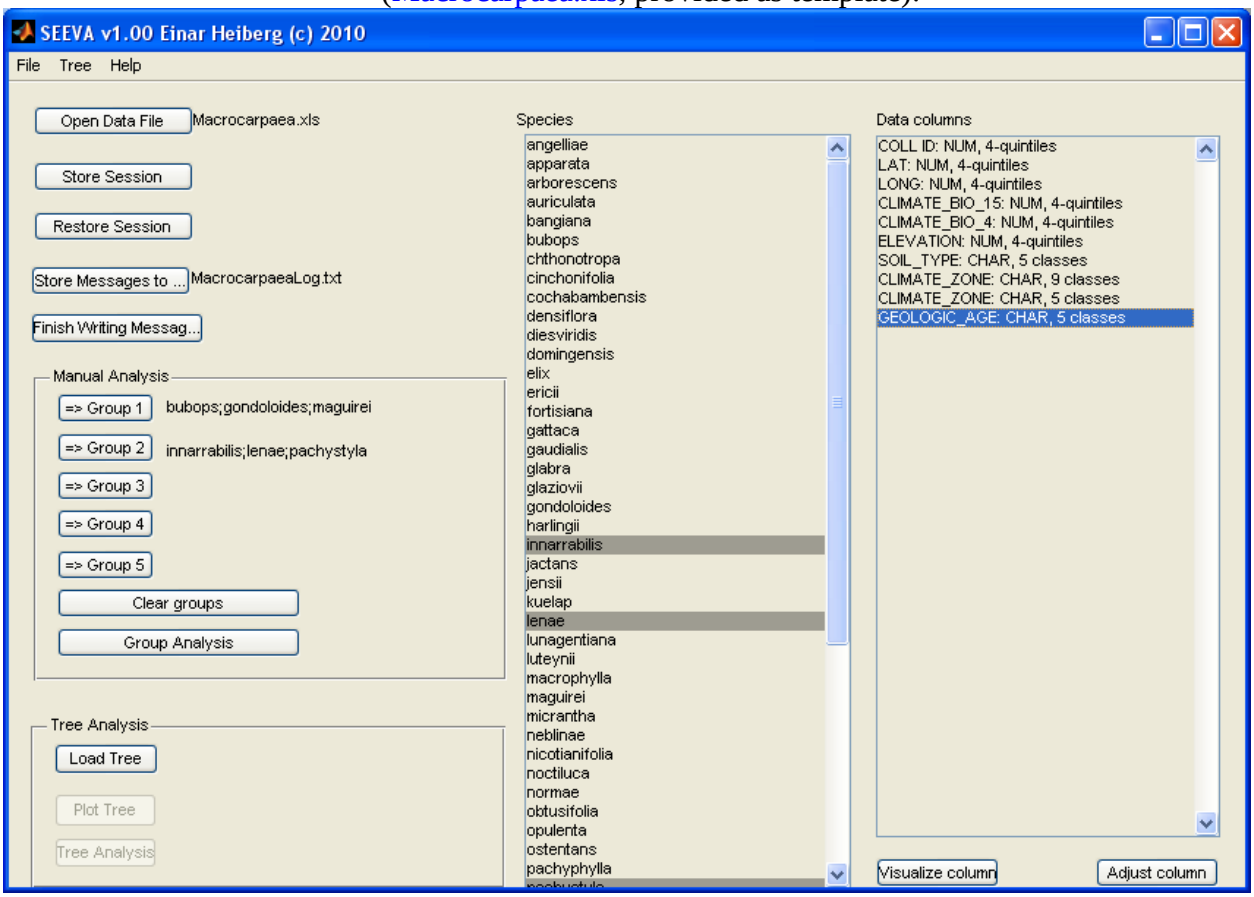
1) In the *Species* window select species to compare. Select several species for one group by holding down SHIFT (for many in a row) or CTRL (to select species not next to each other). All species for a group has to be selected at the same time. Click on *Group 1* to set the first grouping, on *group 2* to set the second grouping. The output result will list the species that are member of each group.

2) Select the *Data Column* name of the variable(s) to be analyzed by clicking on the name of the variable (you can select several, use SHIFT or CTRL).

3) Click on *Group Analysis* to run the SEEVA analysis.

4) The results will be copied to the Clipboard, so open Excel and paste (Ctrl+V) the text into the spreadsheet. (Future version of SEEVA will have a Save As option here, for direct saving into an Excel file).

Example of settings for Manual Analysis for three groups and three variables (Macrocarpaea.xls, provided as template):



Tree-based SEEVA analysis for comparison of clades [TREE ANALYSIS]

1) Select *Load tree* in the *Tree Analysis* box or *Tree* menu, and find and open the nexus file that contains the phylogenetic tree you want to use. If you have several trees in your nexus file, select the tree you want to use from the popup-menu that appears.

2) To view the loaded tree, use *Plot Tree* in the *Tree Analysis* box or *Tree* menu, which will open a new *Tree Visualization* window. This graphic tree can be printed or saved, resized, and you can also zoom in and out on the tree. Node numbers used in the SEEVA analysis results are indicated with numbers on the plotted tree. Closing the Tree Visualization window does not close the tree file you have loaded into SEEVA. [There are many better program to visualize trees, this is just to help you with the interpretation of the SEEVA results. Do not use this visualizer to prepare trees for publication.]

3) Select the data column(s) for the variable(s) you want to analyze.

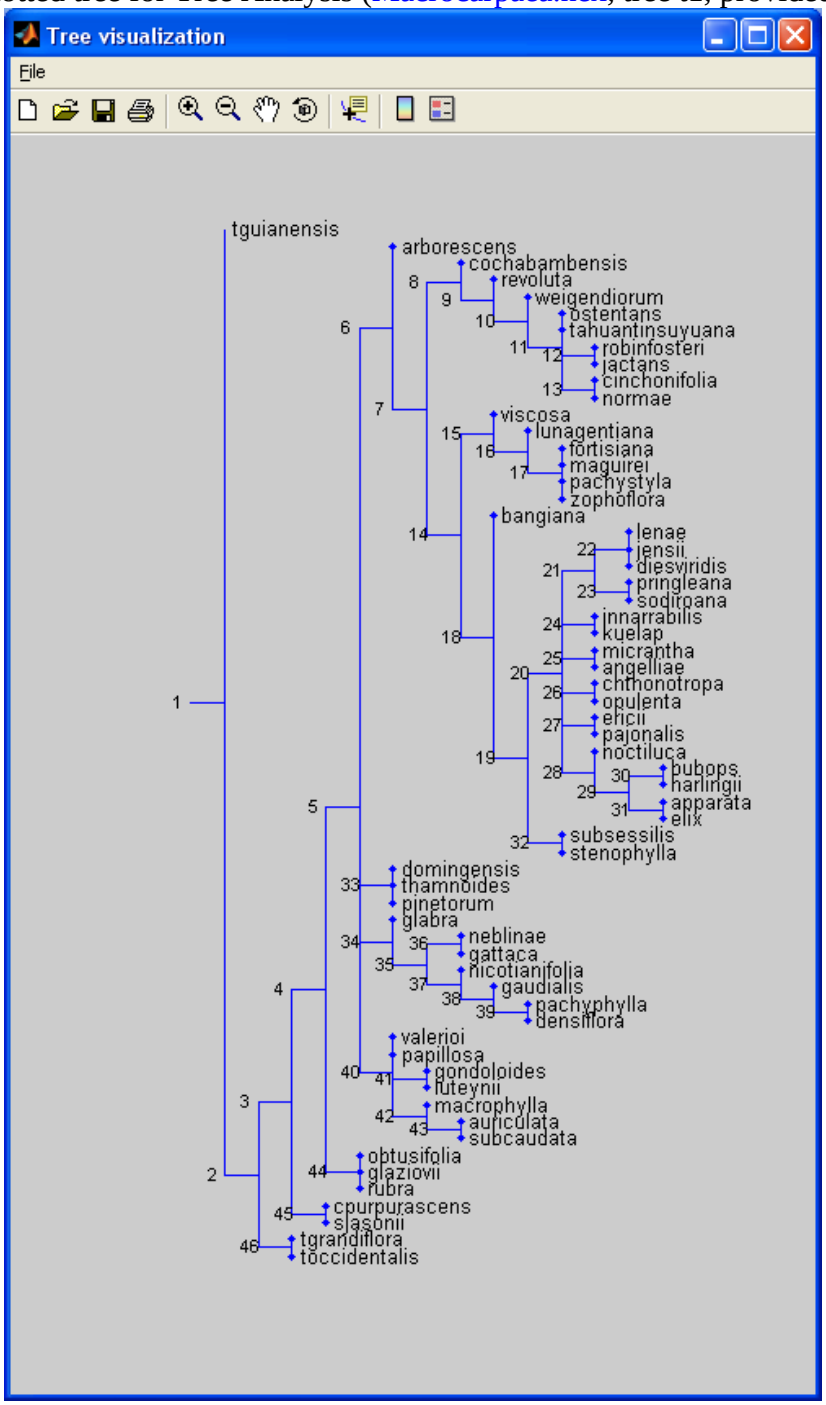
4) In the *Tree Analysis* box or in *Tree* menu, select *Tree Analysis* to run the SEEVA analysis. (If you get warnings that Fisher's exact test couldn't be run, that is fine. Click on OK and continue, see troubleshooting at the end of the manual.)

5) You will be prompted to save the results as a new .xls file

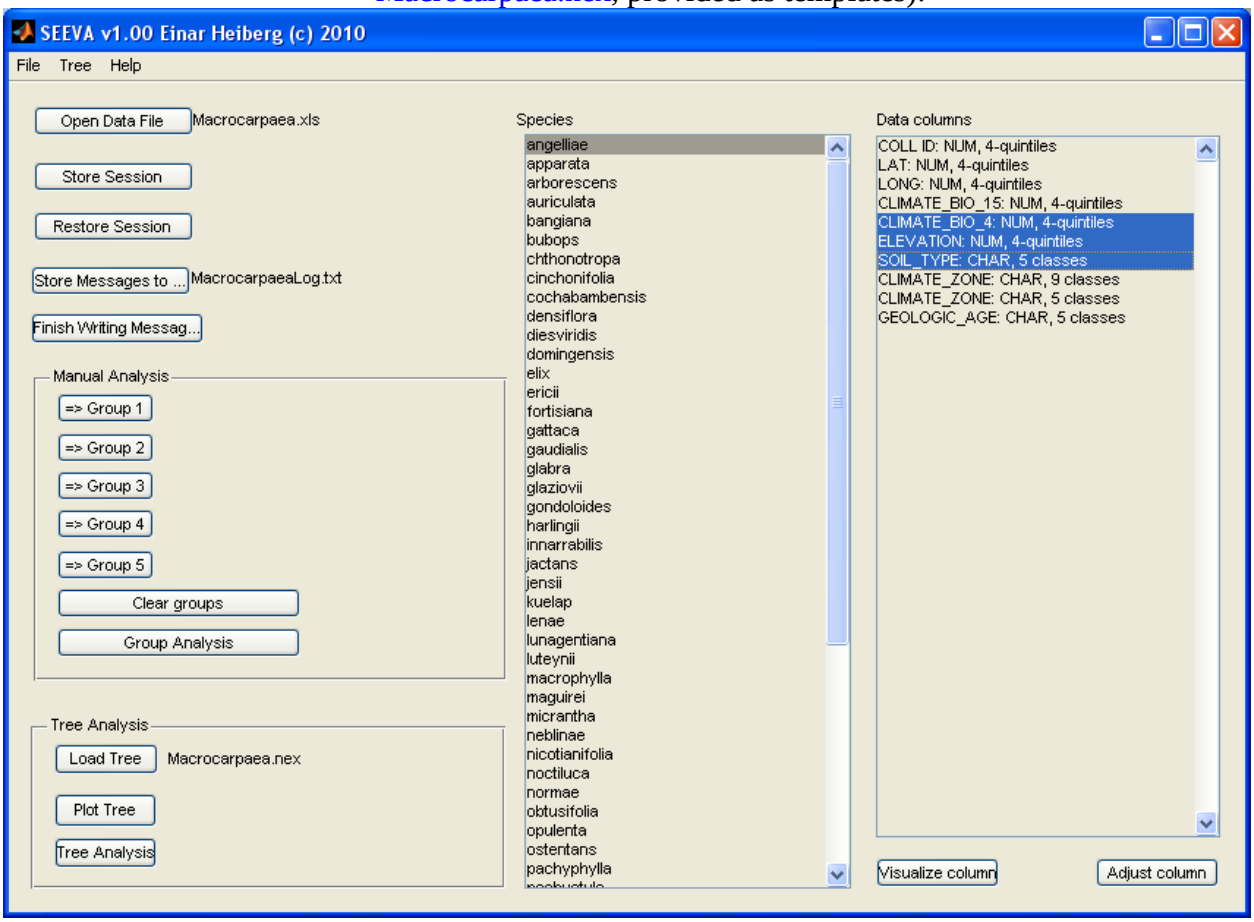
6) Tree visualization windows will pop up for each of the variables you analyzed, showing the index of divergence (*D*) at each node in the phylogenetic tree (shown as a red bar, the

longer the bar the higher *D* is). These tree graphs can be saved or printed. If you have large trees you might need to resize the window to see the species names and nodes well.

Example of plotted tree for Tree Analysis (Macrocarpaea.nex, tree t1, provided as template):

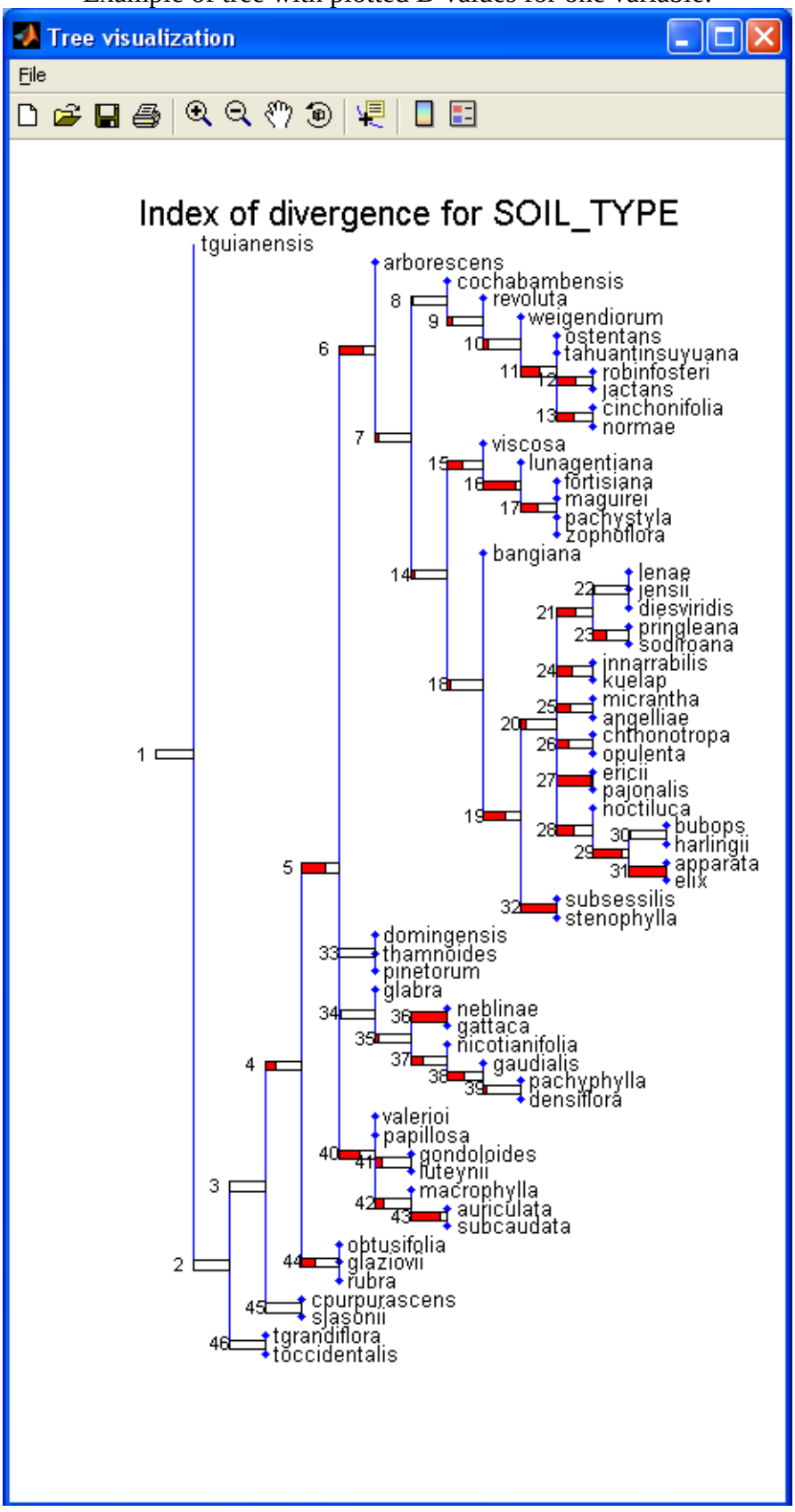


Example of settings for Tree Analysis for one tree and three variables (Macrocarpaea.xls and Macrocarpaea.nex, provided as templates):



After the SEEVA *Tree analysis*, windows will open up and show plotted a tree with Index of Divergence (*D*) values for each variable; node numbers are indicated with numbers on the tree. The longer the red bar is the higher is the D-value for that node. D-values go from 0 to 1, with 0 being no difference between clades, and 1 being maximally different between clades. Currently, the only way to see these plotted D values is after the Tree Analysis, they can't be save in any other way than image files. If you need to see them again later, just rerun that variable again.

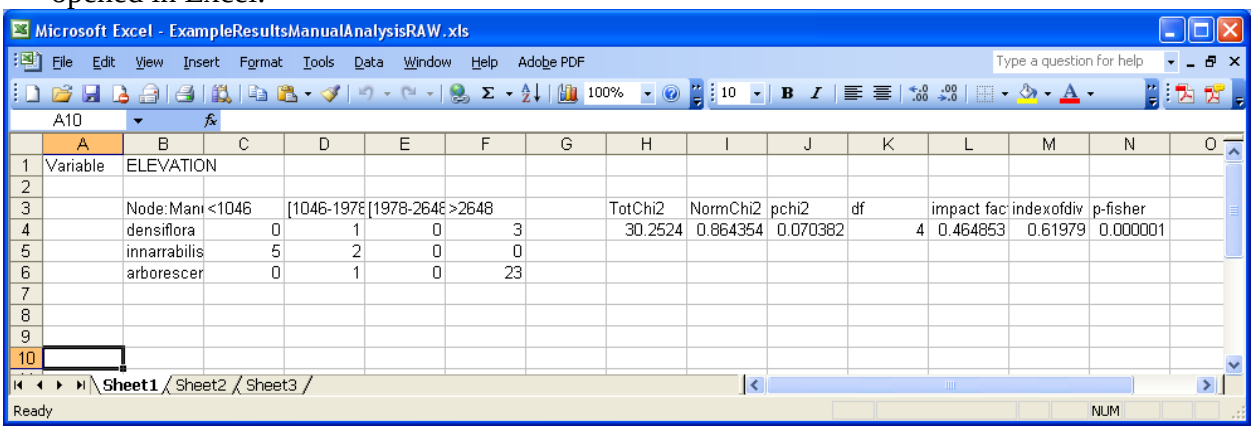
Example of tree with plotted D values for one variable:



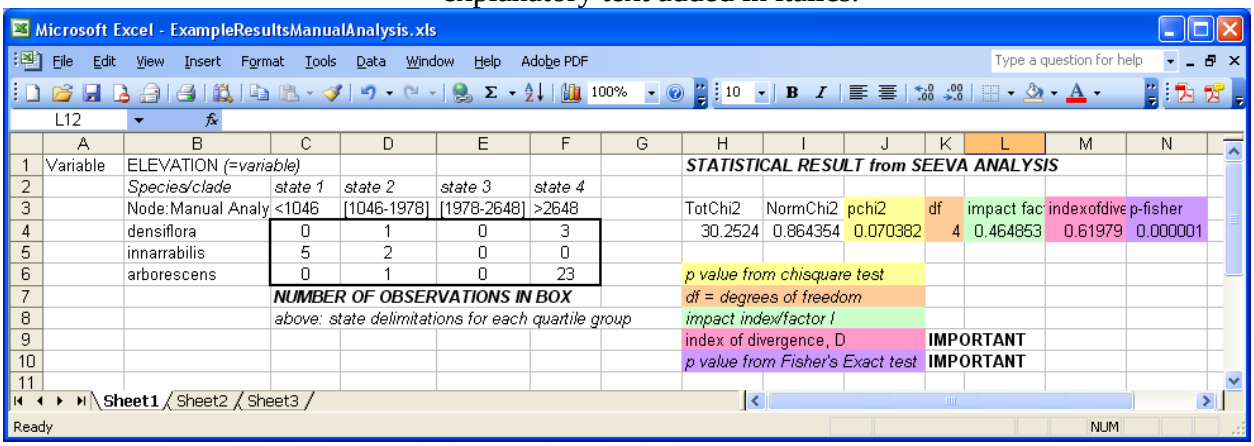
THE DATA OUTPUT FILE (RESULTS)

Open Excel and your results file. The setup of the results file is the same for *Manual Analysis* and *Tree Analysis*. Variables (Data Columns) are listed towards the right, so scroll to the right to see all the results if you analyzed more than one Data Column.

Example of raw result file, directly from SEEVA (ExampleResultsManualAnalysisRAW.xls), opened in Excel:

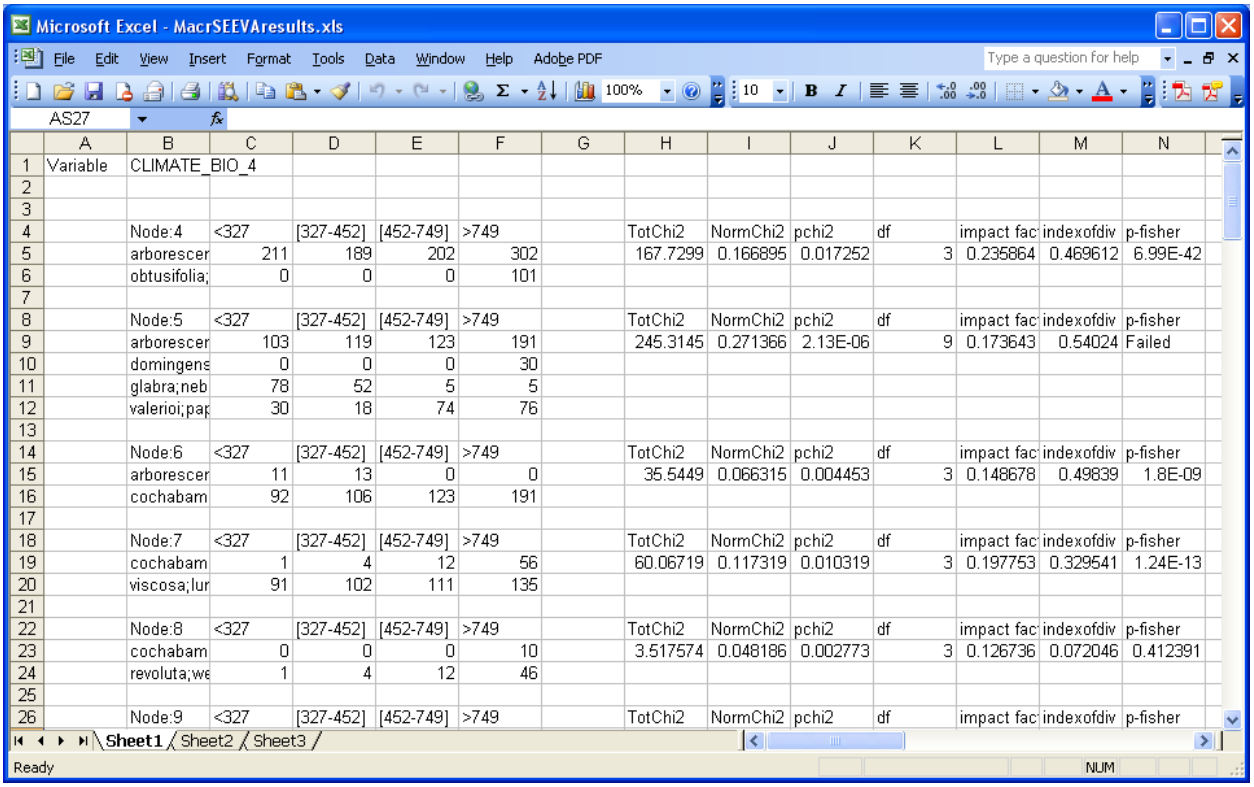


Example of results file from Manual Analysis (ExampleResultsManualAnalysis.xls) with explanatory text added in italics.



The Results file from a Tree Analysis looks the same, except the node numbers are listed and the results from each node is repeated in the spreadsheet.

Example of results file from Tree Analysis (ExampleResultsTreeAnalysisRAW.xls):



Microsoft Excel - MacrSEEVAresults.xls

File Edit View Insert Format Tools Data Window Help Adobe PDF

AS27

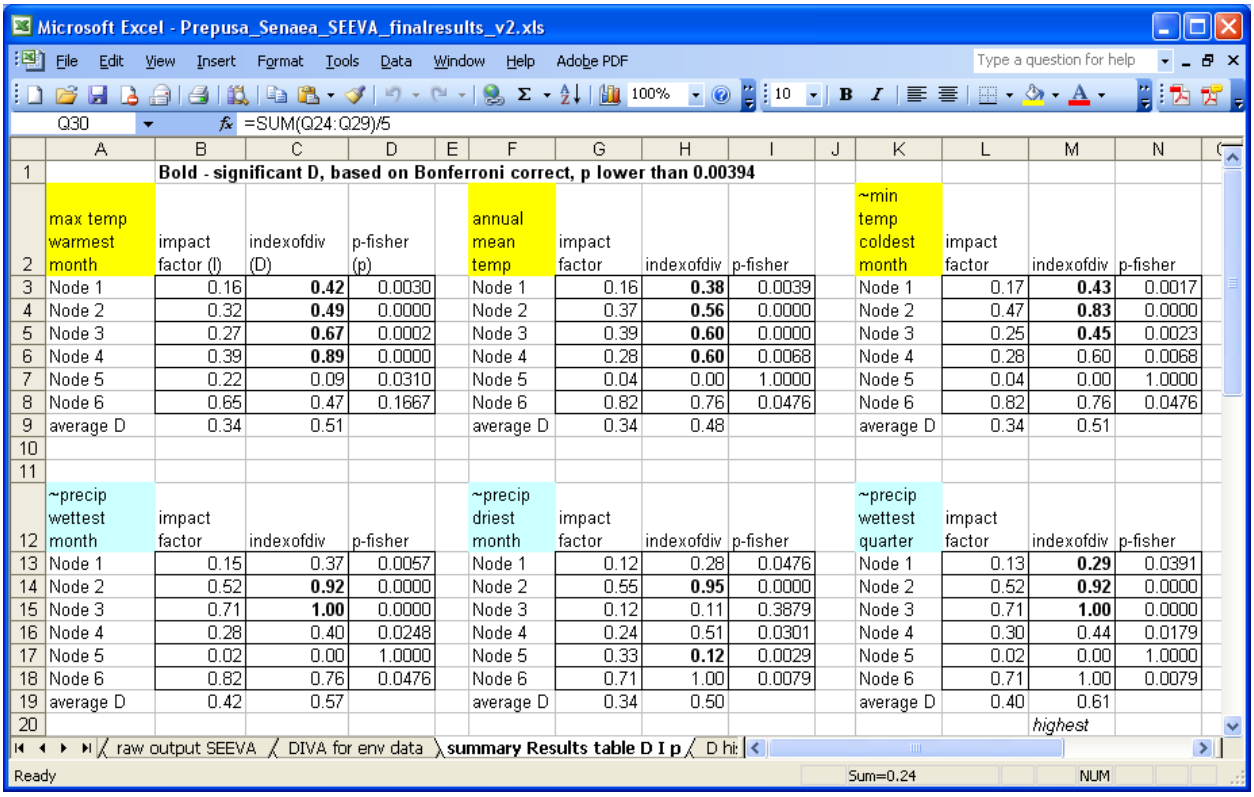
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Variable	CLIMATE_BIO_4												
2														
3														
4		Node:4	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher
5		arborescer	211	189	202	302		167.7299	0.166895	0.017252	3	0.235864	0.469612	6.99E-42
6		obtusifolia;	0	0	0	101								
7														
8		Node:5	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher
9		arborescer	103	119	123	191		245.3145	0.271366	2.13E-06	9	0.173643	0.54024	Failed
10		domingens	0	0	0	30								
11		glabra;neb	78	52	5	5								
12		valerioi;pap	30	18	74	76								
13														
14		Node:6	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher
15		arborescer	11	13	0	0		35.5449	0.066315	0.004453	3	0.148678	0.49839	1.8E-09
16		cochabam	92	106	123	191								
17														
18		Node:7	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher
19		cochabam	1	4	12	56		60.06719	0.117319	0.010319	3	0.197753	0.329541	1.24E-13
20		viscosa;lur	91	102	111	135								
21														
22		Node:8	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher
23		cochabam	0	0	0	10		3.517574	0.048186	0.002773	3	0.126736	0.072046	0.412391
24		revoluta;we	1	4	12	46								
25														
26		Node:9	<327	[327-452]	[452-749]	>749		TotChi2	NormChi2	pchi2	df	impact fac	indexofdiv	p-fisher

Sheet1 Sheet2 Sheet3

Ready NUM

These result files are not very pretty, and we are working on a better interface and summary of the results. Fo

r now, you will have to use Excel and make nice summary tables by hand, for example such as these:



Microsoft Excel - Prepusa_Senaea_SEEVA_finalresults_v2.xls

File Edit View Insert Format Tools Data Window Help Adobe PDF

Q30 =SUM(Q24:Q29)/5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		**Bold - significant D, based on Bonferroni correct, p lower than 0.00394**												
2	max temp warmest month	impact factor (I)	indexofdiv (D)	p-fisher (p)		annual mean temp	impact factor	indexofdiv	p-fisher		~min temp coldest month	impact factor	indexofdiv	p-fisher
3	Node 1	0.16	**0.42**	0.0030		Node 1	0.16	**0.38**	0.0039		Node 1	0.17	**0.43**	0.0017
4	Node 2	0.32	**0.49**	0.0000		Node 2	0.37	**0.56**	0.0000		Node 2	0.47	**0.83**	0.0000
5	Node 3	0.27	**0.67**	0.0002		Node 3	0.39	**0.60**	0.0000		Node 3	0.25	**0.45**	0.0023
6	Node 4	0.39	**0.89**	0.0000		Node 4	0.28	**0.60**	0.0068		Node 4	0.28	0.60	0.0068
7	Node 5	0.22	0.09	0.0310		Node 5	0.04	0.00	1.0000		Node 5	0.04	0.00	1.0000
8	Node 6	0.65	0.47	0.1667		Node 6	0.82	0.76	0.0476		Node 6	0.82	0.76	0.0476
9	average D	0.34	0.51			average D	0.34	0.48			average D	0.34	0.51	
10														
11														
12	~precip wettest month	impact factor	indexofdiv	p-fisher		~precip driest month	impact factor	indexofdiv	p-fisher		~precip wettest quarter	impact factor	indexofdiv	p-fisher
13	Node 1	0.15	0.37	0.0057		Node 1	0.12	0.28	0.0476		Node 1	0.13	**0.29**	0.0391
14	Node 2	0.52	**0.92**	0.0000		Node 2	0.55	**0.95**	0.0000		Node 2	0.52	**0.92**	0.0000
15	Node 3	0.71	**1.00**	0.0000		Node 3	0.12	0.11	0.3879		Node 3	0.71	**1.00**	0.0000
16	Node 4	0.28	0.40	0.0248		Node 4	0.24	0.51	0.0301		Node 4	0.30	0.44	0.0179
17	Node 5	0.02	0.00	1.0000		Node 5	0.33	**0.12**	0.0029		Node 5	0.02	0.00	1.0000
18	Node 6	0.82	0.76	0.0476		Node 6	0.71	1.00	0.0079		Node 6	0.71	1.00	0.0079
19	average D	0.42	0.57			average D	0.34	0.50			average D	0.40	0.61	
20													*highest*	

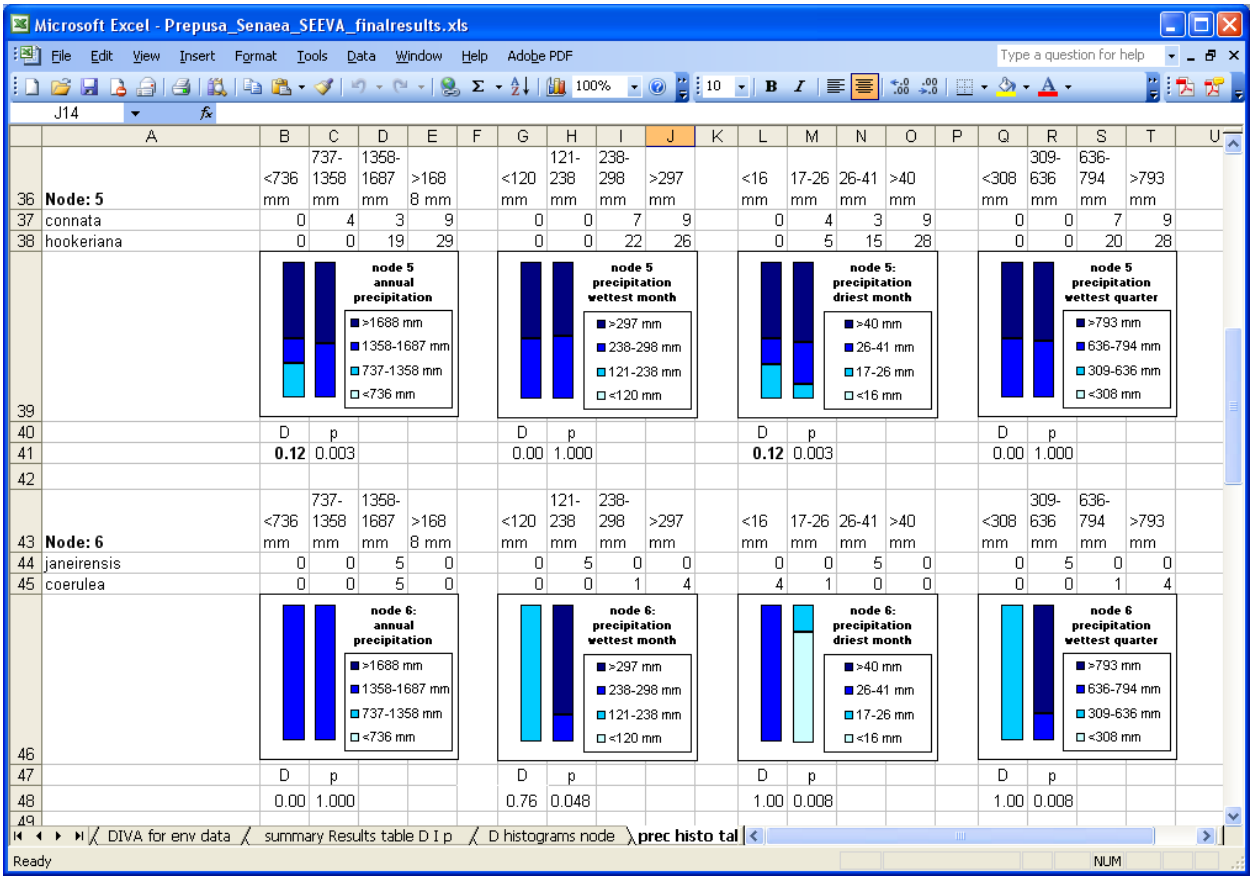
raw output SEEVA | DIVA for env data | **summary Results table D I p** | D hi

Ready Sum=0.24 NUM

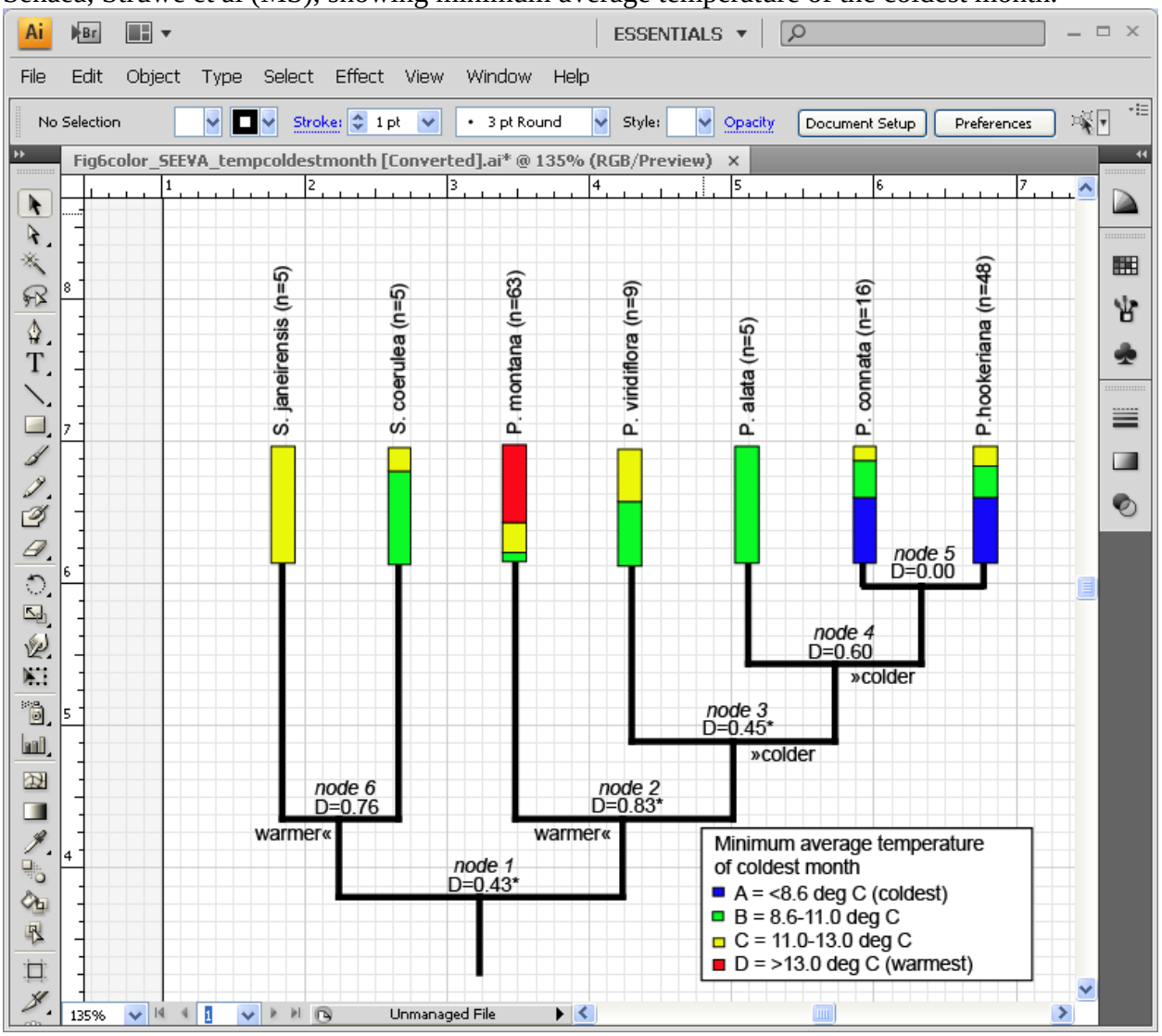
GRAPHIC PRESENTATION OF RESULTS

Currently SEEVA's result interface is extremely limited, and you will have to use tables and graphic features in Excel and other programs to summarize your data. We plan to have additional features in the next SEEVA. As examples, we have copied below some of the features in Excel we have used to make graphs that then have been imported into Illustrator to make more complex figures. These figures can either be in Color or Grayscale.

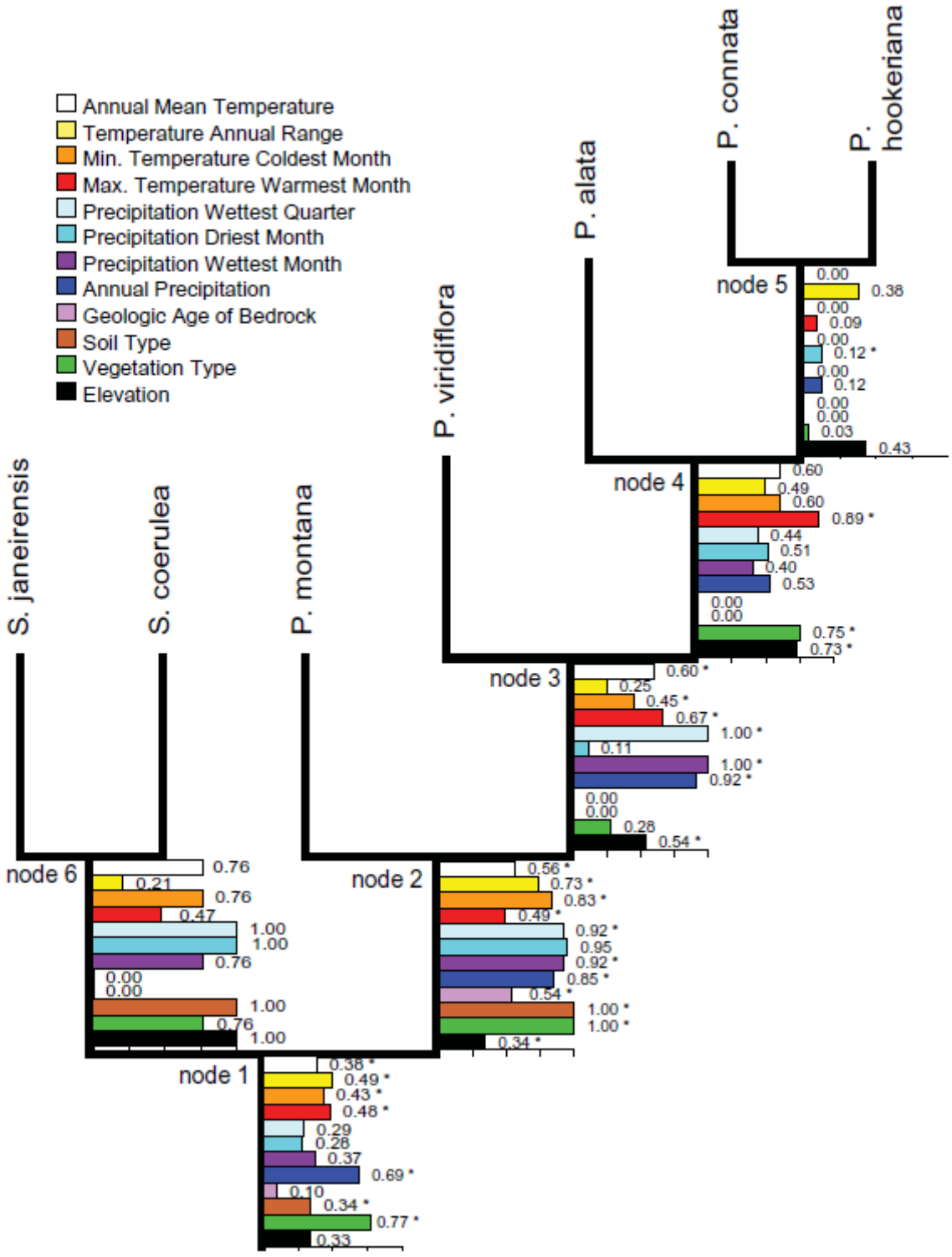
Histogram bars on distribution of categories of a variable within a species or at each node can be made in Excel (see example below from precipitation data for Prepusa and Senaea, Struwe et al. (MS)).



The histogram bars can be imported into Illustrator and used in a graphic design that illustrates the distribution of different categories of a variable (see example below from Prepusa and Senaea, Struwe et al (MS), showing minimum average temperature of the coldest month.



A range of divergence indices for nodes can also be illustrated, such as shown below (also made in Excel and Illustrator):



TROUBLESHOOTING AND KNOWN ISSUES

Fisher's Exact Test can't be run for some nodes:
Warnings will pop up if you have nodes that show no divergence at all for a variable at a node, i.e., only one variable category is present, and Fisher's exact test is then impossible to run for that node. Click OK on these and the analysis will continue for the rest of the nodes and variables. In the results file, there will be no p-values for the nodes where Fisher's exact test couldn't be run (and *D* will be 0).

SEEVA crashes due to too many variables analyzed simultaneously:
Warning – the number of variables that can be analyzed simultaneously are dependent on 1) how many columns that can be fit into a Excel Results sheet, and 2) your computer's RAM memory. We have successfully run analyses with up to 10 variables at the time. You can analyze your variables in batches and then combine the results for a complete overview.