WinClada ver. 1.00.08 (updated 24 Oct 2003)
Program copyright 1999-2003 by Kevin C. Nixon

A basic manual comprised of contributions by Christopher Hardy, Diana Lipscomb, Kevin Nixon, and Helga Ochoterena

Contents

(i) General Information about WinClada ................................................................. 3
(ii) Technical Information .......................................................................................... 3
   Setting up NONA to run with WinClada ............................................................... 3
   Suggested file extensions ....................................................................................... 3
   Opening and Saving Nexus files ........................................................................... 3
   Opening old NONA and Hennig86 files ............................................................... 4
I. Working with matrices (winDADA mode) ............................................................. 5
   A. Creating data matrices ...................................................................................... 5
      1. Creating morphology data matrices ............................................................... 5
         Creating a new file ......................................................................................... 5
         Entering or changing taxon names ............................................................... 5
         Long taxon names ....................................................................................... 6
         Entering character descriptions ................................................................... 6
         Entering character data ............................................................................. 6
         Saving the matrix ...................................................................................... 6
      2. Creating DNA data matrices .......................................................................... 9
         Downloading sequences from GenBank ......................................................... 9
         Directly from Sequencher™ projects ............................................................ 9
         From Clustal™ alignments .......................................................................... 9
         Manually in MS Word™ .............................................................................10
      3. Conversion from PAUP™ or MacClade™ matrix ........................................... 10
   B. Saving a matrix .................................................................................................. 10
   C. Compatibility with PAUP and MacClade .......................................................... 10
   D. Alignment of DNA sequences ......................................................................... 10
   E. Editing matrices .................................................................................................. 11
      Deactivating or deleting uninformative characters ........................................... 11
   F. Merging datasets ................................................................................................ 11
   G. Manual random matrix perturbations ............................................................... 12
   H. Summary of Key & Mouse Commands in winDADA mode .............................. 13
      Moving cursor & selecting/unselecting taxa or characters .............................. 13
      Allocating view space for shorter or longer terminal names ............................ 13
      Incremental right scroll through matrix ......................................................... 13
      Switch between open matrices ...................................................................... 13
      Adjusting column width ............................................................................ 13
I. Other useful tools, features and notes about winDADA ....................................... 13
   A. Heuristics ......................................................................................................... 15
   B. Ratchet ............................................................................................................ 16
   C. Incongruence Test ........................................................................................... 18
D. Bootstrap or Jackknife

Trouble making the bootstrap support values go away from tree? .......................................................... 21
To get back to mode where you can view hashmarks on branches? ................................................................ 21

III. Viewing, manipulating, and printing your trees (winCLADOS mode) .......................................................... 22

A. Viewing and working with trees .................................................................................................................. 22
1. To see trees previously saved in a tree file ........................................................................................................ 22
2. Tree styles ......................................................................................................................................................... 22
3. Viewing optimizations / Tracing characters ..................................................................................................... 23
4. Tree fonts ......................................................................................................................................................... 25
5. Manipulating taxa and branches on a tree ........................................................................................................ 25

B. Ancestor Reconstructions .......................................................................................................................... 26

C. Consensus Trees ............................................................................................................................................... 27

D. Printing or Exporting Trees .......................................................................................................................... 27

E. Summary of Key Commands for Trees (in winCLADOS mode) ................................................................. 27
   Moving tree in view window .......................................................................................................................... 27
   Thinning or Thickening branches .................................................................................................................. 27
   Resizing & Reproportioning tree .................................................................................................................. 27
   Adjusting size of hashmarks ......................................................................................................................... 27
   Scrolling through different trees in memory .................................................................................................. 28
   Selecting & Unselecting trees ....................................................................................................................... 28

F. Functions of Icons on tool bar at top of tree view ..................................................................................... 28

G. Final Notes and Comments about winCLADOS mode ............................................................................... 28
(i) General Information about WinClada

**WHAT:** WinClada was written by Kevin Nixon of Cornell University. It is a data editor, tree viewer/printer, and analytical tool for use with NONA, PIWE, and HENNIG86. WinClada replaces the older DOS programs DADA and CLADOS, and combines capabilities of both with additional new capabilities. Hence, the WinClada modes that perform the functions of the old DADA and CLADOS are referred to below as winDADA and winCLADOS modes, respectively.

**AVAILABILITY:** WinClada is available as a shareware download from the website [cladistics.com](http://cladistics.com). Information regarding payment for the program can be found at this website.

**CITATION:** Nixon, K. C. 2002. WinClada ver. 1.00.08. Published by the author, Ithaca, NY.

**REQUIREMENTS:** WinClada is a Win32 program that will run on Intel compatible machines running Windows 95, Windows 98, NT ver. 4.0+, 2000 and XP.

**CAPABILITIES:** WinClada reads DADA, CLADOS, NONA, PIWE or HENNIG86 files. It will also read most NEXUS format files that contain cladistic data matrices.

(ii) Technical Information

**SETTING UP NONA TO RUN WITH WINCLADA:**

NONA ➔ (by Pablo Goloboff, available at [cladistics.com](http://cladistics.com)) can be set up to run as a daughter process from WinClada. To do this:

WinClada and NONA should be in same directory folder (e.g., “Cladistics” on the C: drive). Open WinClada. Tell WinClada where to find NONA (Under ANALYZE > SPAWN > NONA > ENTER PATH). Then browse to locate NONA version (e.g., “NONA.exe”). After setting this the first time, you should not have to do it again unless you move the location of the program files.

**SUGGESTED FILE EXTENSIONS:**
Note: Using the default file extensions supported by WinClada makes loading files easier. Test (*.txt) files can also be used if they are formatted correctly; however, for WinClada to find these files in the browser window, they should have the following extensions. These extensions can be changed in Windows prior to opening the files.

- .winc – WinClada data file. The most useful format, in which much information about characters and terminals may be recorded, but this not compatible with other phylogenetic software.
- .ss - Hennig86/NONA/DADA compatible data file. The ss refers to the name of the original executable for Hennig86. This stood for “SuperStar”.
- .rat - A Hennig86/NONA/CLADOS compatible tree file that contains output trees from a ratchet run. These trees will have various lengths.

NEXUS, FASTA, and GDE files:

- .gde - A file with GDE format matrix.
- .fst - A file with a FASTA format matrix. However, simple modifications often need to be made in order for WinClada to open it (as described below).
- .nex, .NEXUS - A NEXUS format matrix.

**OPENING AND SAVING NEXUS FILES IN WINCLADA:**

WinClada 1.0 will read NEXUS files that were generated in PAUP or MacClade and contain a matrix. WinClada may have problems opening a Nexus tree file that does not contain a matrix, even if the matrix to which the tree corresponds is open in WinClada. If you wish to view a Nexus tree file that was generated in PAUP or MacClade, it is best to save that tree in a Nexus file with the appropriate matrix in PAUP or MacClade. Then open this matrix plus tree file in WinClada.

Some earlier versions of WinClada are not able to read interleaved NEXUS files.

WinClada will also save a matrix plus, if desired, trees in Nexus format, which can be used in PAUP or MacClade. However, notes and special character weighting schemes that are part of the WinClada file will not be saved in this nexus file.
OPENING OLD NONA AND HENNIG86 FILES IN WINCLADA:

WinClada does not read all files that NONA and HENNIG86 will read. Some common problems occur when there are not spaces following commands; for example, the command cc]10;p/; works fine in Hennig86 or NONA but will create difficulties for some versions of WinClada.

    Change cc]10;p/; to cc ] 10;p /;

Also, semicolons may be necessary to end some functions in WinClada that are not required in NONA or HENNIG86.
I. Working with matrices (winDADA mode)

A. Creating data matrices:

1. Creating morphology data matrices
   
a. Creating a new file:
      
      Choose “MATRIX > New matrix (create)”. A window (below, left) appears that prompts for the number of taxa and the number of characters. These can be modified later (as explained below) so one can create a matrix with approximate numbers of taxa and characters.

      After choosing numbers of taxa and characters, press “OK! Resize!” to create the matrix.

      This takes you immediately to the winDADA data editor window (below, right):

   
   [Insert image of winDADA data editor window]

   b. Entering or changing taxon names:
      
      Open the terminal dialog via "TERMS > Terminal dialog". The window below appears, and allows one to scroll through (with "Next" and "Previous") and edit taxon names. winDADA allows spaces and periods in a name. Notes may also be entered about the delimitation of the terminal, citation, and descriptions or comments about taxon sampling (e.g., collection numbers from which the DNA used for sequencing were taken).

   [Insert image of terminal dialog window]

   **Auto apply ON or OFF:**
   Numerous missing or inapplicable characters may cause a terminal to float in an analysis, resulting in poor consensus resolution. Often, it may desirable to have these potentially problematic terminals labeled. WinClada allows for automatic tagging of terminals with more than a specified number of missing characters (set under "TERMS > Ambiguity filter"). This feature is in effect when "Auto apply ON" is marked (default). Automatic tagging can be turned off by choosing "Auto apply OFF."
Using Long Taxon Names:
The ends of long terminal names may be hidden behind the characters block. Fix this by holding down [SHIFT + right arrow] until the names can be easily read. To reduce the space assigned for the terminal names, [SHIFT + left arrow] are pressed.

c. Entering character descriptions (“Character dialog”)
Choose “CHARS > Character dialog” and the following window opens:

- Enter character names (not under “Char ID”!), and state assignments.
- Choose how you want the character treated (additive or nonadditive, etc.).
- Apply (or choose autoapply) and use the “Next” button to advance to the next character.

NOTE: Matrices and lists of characters and states may be output to text files. Given the output file format we suggest to use capitals for the character name and lower case for the states names. This makes the output files much easier to read.

d. Entering character data
1.) First unlock the matrix (this is a safety feature that prevents you from accidentally overwriting data).
- Choose “EDIT > Unlocked - data entry allowed”
- Alternatively, enter a state in any cell and the program will ask if you wish to unlock the matrix.
- It is a good idea to relock your matrix when you are finished with editing.

2.) Character States may be entered for the terminals by several ways. Below are two:

a.) Simply typing over the dashes “-” in the matrix.
Character states must be indicated by numbers 0 to 9 (default), or nucleic acid bases.
DNA codings are obtained via “VIEW >DNA (IUPAC) code”

To facilitate entering data, the cursor can be set to move automatically when after a character state is entered. The default is for the cursor to automatically jump one space to the right in a row, such that all character data for a given taxon may be entered in succession.

"VIEW >Cursor Settings” opens a window that allows one to change this (see below). The top three buttons determine which direction the cursor moves (if at all).

The next four buttons (which can be used together or separately) change the way the cursor looks during data entry (see below, left):

The “Display data” buttons allow you to view taxon names, character names and state names while editing or viewing the data (see above, right).

By default, the background color of the character at the cursor position is black, and the character number itself is white. This can be changed using the two color buttons at the bottom of the “Cursor Settings” window.

b.) Entering character data via the Character panel
Useful for entering or checking character states or viewing the distribution of character states among the taxa.

Choose “INTERFACE >Submode Cpanel and the following window (below, left) opens.

- Use the “Prev Char” and “Next Char” to scroll through the characters
- “Mode” changes to a table format for the information. You can also enter/edit the matrix in this mode (after unlocking it) just by clicking the corresponding taxon state(s) for the current character. All the changes that you do here will automatically apply to the matrix.
- Click “Done“ to return to the normal winDADA interface.
c.) Entering character data via the Taxon panel

Score taxa for characters here

If many terminals have the same state, then it is possible to score them as so all at once. Just select them in the matrix, then use TPanel to score them.
d.) Entering polymorphisms

Polymorphisms can be entered using Character or Taxon Panels, as shown above, or as follows:

[Image of polymorphic entry dialog]

Enter polymorphisms using this window.
("EDIT >Enter Polymorphisms")

e.) Saving the matrix:

Select “MATRIX >Save as” the first time you save a file. Select to save as a WinClada or NONA file. Name and then save the file.

NOTE: NONA files do not allow for the saving of all of your cursor settings or notes entered under character and terminal dialogs. If you wish to conserve this information, save it as a .winc file.

2. Creating DNA data matrices

a. Downloading sequences from GenBank:

Select all sequences you want in GenBank and then select "FASTA" format and "save" as a text document. Then all of your sequences will be in one text document file in FASTA format.

WinClada does not currently read the type of FASTA file you now have, but the file can be modified in a word processor (e.g., MS Word) easily. All that is required is to put the appropriate header and ending text onto the file (as below for item d) and then by removing the “>” prefix to all taxon names and take out any other extraneous information that FASTA format puts into the taxon/sequence names using the “Find & Replace” function under the EDIT menu in MS Word. This modified FASTA file is then saved as a text file and opened in WinClada.

Alternatively, the FASTA text file may be opened in Sequencher and the sequences may now be exported from Sequencher to WinClada-readable format automatically according to the procedure described below.

b. Directly from Sequencher™ projects or Sequencher “alignments”:

Having opened a multi-sequence FASTA file in Sequencher (as downloaded from GenBank above in item 2) or having loaded multiple individual sequence text files into a Sequencher project window, the procedure from here is the same:

Select all sequences. Assemble a “contig” (either with or without aligning), then open the contig and, in the upper window, select all component sequences (do not select the “consensus sequence”) and then, under FILE, export “Consensus” as a NEXUS file (in DOS format if on a MAC). This file should have all sequences in it in an interleaved fashion. Some earlier versions of WinClada are not able to read interleaved NEXUS files. Thus, these files may be reopened in PAUP and resaved as non-interleaved NEXUS files.

NOTES: Sequencher inserts a colon (":") for gaps and missing data. WinClada does not currently recognize the colon and has problems with it. Therefore, in Word, open the NEXUS file and do a “Find and Replace” under the “Edit” menu to replace all "::" with "-".

This file can now be opened in WinClada.

Sequencher may also insert hard returns within a sequence that may cause some problems when reading the file. This is not usually the case, but if having problems opening the file, use “Find” in MS Word to check that the sequences are not interrupted buy a hard return.

Alternatively, a Sequencher NEXUS file may be opened in MacClade (which seems to accept Sequencher’s colons, etc.), resaved, and then opened in WinClada.

c. From Clustal™ alignments:
A Clustal alignment can be saved as a “.gde” file and this can be opened in WinClada. Other file save options in Clustal (e.g., nexus, fasta, nbrf, etc.) are not wholly compatible with the latest versions of WinClada.

d. **Manually in MS Word™:**

Follow the NONA format. An example is provided below. Simply cut and paste your sequences into a text file and format as below.

The file must start with “dread” or “dnaread” and the next line(s) may contain notes about the matrix as seen below.

The next line has first the number of bases (characters; e.g., 10 in the example below) followed by the number of terminals (OTUs).

This is followed by the taxon names and their sequences as shown. It is important that all sequences be of the same length (“n”s or “-”s may be inserted anywhere in sequence or at the ends to achieve this). Preparing the matrix in a proportional block font such as “Courier” facilitates this. The “word count” tool in Word can also be used to confirm the number of characters each taxon has.

The matrix file is terminated as shown below. This file should be saved as a text file (a .txt file can be opened in WinClada). Once in WinClada, the matrix can be manipulated and saved as a “.winc” file or “.ss” file.

dread
`example matrix; supplementary notes about matrix may be put between single quotes`

10 3

```plaintext
taxon_A
acctggtacg

taxonomy_B
acgtggtacg

taxonomy_C
acctgcttcg
```

```plaintext
proc /;
;
```

3. **Conversion from PAUP™ or MacClade™ matrix.**

Latest versions of WinClada read PAUP or MacClade NEXUS files. However, for WinClada to find the file in its browser window, they must have the extension “.nex”. Extensions may be changed in Windows before attempting to open it in WinClada.

However, MacClade will also export a matrix as a “NONA” file, which is read by WinClada. This option will create a file with a “.non” extension. This is not official NONA extension but it can be read anyway, as it is a text file with the appropriate NONA format and readable by WinClada.

**B. Saving a matrix:**

Select “MATRIX > Save as” the first time you save a file. Select to save as a WinClada or NONA file. Name and then save the file.

**NOTE:** NONA (.ss) files do not allow for the saving of all of your extra character descriptions, cursor settings or notes entered under character and terminal dialogs. If you wish to conserve this information, save it as a WinClada (.winc) file.

**C. Compatibility with PAUP and MacClade:**

WinClada recognizes and will open nexus files created in PAUP and MacClade that contain cladistic data matrices.

WinClada will also write a file to nexus format, which can then be opened in PAUP or MacClade (OUTPUT > Export nexus file). However, character weights and deactivations recorded in WinClada are not saved in this nexus file.

**D. Alignment of DNA sequences:**

DNA sequences can be manually aligned in winDADA mode. Under ANALYZE, choose “Moleculoid” then “Manual Alignment”. This feature will automatically add characters onto the end of the matrix to allow for the insertion of gaps. Gaps can then be inserted or removed using the [Insert] and [Delete] buttons, respectively.
NOTE: WinClada has a security mode that alerts the user in the event that they were to accidentally delete bases from the original sequence.

E. Editing matrices: Any matrix cell entry (bases or character state codings) may be edited (e.g., if there is change in the electropgerogram interpretation) by typing the new base on top of the old one. In order to prevent accidental changes, WinClada has a security routine that will ask if the matrix is really to be modified (even when one is using the alignment function). Simply select to "unlock" the matrix and then changes can be made.

F. Merging datasets

1. With “FILE >Open”, open the data sets you wish to combine.
2. Select matrices under “MATRIX >Select all matrices”.
3. Open the matrix merge dialog by choosing “MATRIX >New matrix merge”. The following window will appear. This window will prompt you with decisions, such as whether or not you want to match terminals by order or name.

   One can also select which matrix to be the “Primary” and “Secondary” respectively. Then, under “Taxon Orphan Control”, one can choose to include (“keep orphans”) or discard (“discard orphans”) taxa that are present in the “secondary” matrix and not the “primary” matrix, if relevant.

   The “Canned Heat” box in this window provides an alternative, automatic way of doing what’s listed below. It is not necessary to use these functions; thus, “Canned Heat” will not be described.
4. Below are several, but not all, of the available options to choose in the marge matrix dialog.

a. **Merging matrices when the terminals are the same but there are two different sets of characters:**
   Choose “Match terminals by name” and “Don’t match characters (not equal)”. Select “Merge matrices” and a new matrix is created:

<table>
<thead>
<tr>
<th>Primary Matrix:</th>
<th>Secondary Matrix:</th>
<th>Merged Matrix:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxon A 000</td>
<td>Taxon A 111</td>
<td>Taxon A 000111</td>
</tr>
<tr>
<td>Taxon B 110</td>
<td>Taxon B 011</td>
<td>Taxon B 110011</td>
</tr>
<tr>
<td>Taxon C 111</td>
<td>Taxon C 011</td>
<td>Taxon C 111011</td>
</tr>
</tbody>
</table>

b. **Merging matrices when the characters are the same but there are two different sets of terminals:**
   Choose “Don’t match terminals (not equal)” and “Match characters by Name”. Then select “Merge matrices” and a new matrix is created:

<table>
<thead>
<tr>
<th>Primary Matrix:</th>
<th>Secondary Matrix:</th>
<th>Merged Matrix:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxon A 000</td>
<td>Taxon D 111</td>
<td>Taxon A 000</td>
</tr>
<tr>
<td>Taxon B 110</td>
<td>Taxon E 011</td>
<td>Taxon B 110</td>
</tr>
<tr>
<td>Taxon C 111</td>
<td>Taxon F 011</td>
<td>Taxon C 111</td>
</tr>
<tr>
<td>Taxon D 111</td>
<td></td>
<td>Taxon D 111</td>
</tr>
<tr>
<td>Taxon E 011</td>
<td></td>
<td>Taxon E 011</td>
</tr>
<tr>
<td>Taxon F 011</td>
<td></td>
<td>Taxon F 011</td>
</tr>
</tbody>
</table>

c. **Merging matrices when both characters and terminals are to be kept separate:**
   Choose “Don’t match terminals” and “Don’t match characters”. “Merge matrices”:

<table>
<thead>
<tr>
<th>Primary Matrix:</th>
<th>Secondary Matrix:</th>
<th>Merged Matrix:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxon A 000</td>
<td>Taxon D 111</td>
<td>Taxon A 000---</td>
</tr>
<tr>
<td>Taxon B 110</td>
<td>Taxon E 011</td>
<td>Taxon B 110---</td>
</tr>
<tr>
<td>Taxon C 111</td>
<td>Taxon F 011</td>
<td>Taxon C 111---</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Taxon D ---111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Taxon E ---011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Taxon F ---011</td>
</tr>
</tbody>
</table>

**G. Manual Random Matrix Perturbations**
Characters can be randomly reweighted ("CHARS> Randomly select chars", followed by reweighting) and taxa can be randomly selected or unselected for later deletion, etc., in order to create new and randomly perturbed matrices for various applications.
H. Summary of Key & Mouse Commands in winDADA mode

Most keys work more or less as they would in a text editor/word processor. Below are some keys with special function.

The LEFT mouse button: Single click places the cursor in that position. Double click on a taxon name or character column selects and highlights that taxon or character. A “click and drag” from a selection allows one to select multiple items.

The RIGHT mouse button unselects a taxon by clicking on the taxon name, and characters by clicking on one of the cells of that character. A "click and drag" with this mouse button unselects more than one.

[arrow keys]: moving between taxa and/or characters in matrix.

[shift] + [left or right arrow]: moves character block display on the matrix left or right, respectively, to allow one to see less or more, respectively, of the terminal names.

[F3]: incremental right scroll through matrix, by one column at a time.

[F6]: incremental left scroll through matrix

[F12]: go to next open matrix (i.e., switch between open matrices)

[F2]: Repeated pressing of this will taken matrix column width to its minimum. Often, after resizing and manipulating a tree in the winCLADOS interface, the matrix columns in winDADA interface will be quite wide. You can return the columns to their normal width with [F3].

I. Other useful tools, features and notes about winDADA

Note: The WinClada menu is common sensical in that all functions relating to the matrix, characters, terminals, and analysis options may be found as separate items by the same names on the top menu bar.

Adding taxa: Go to the terms drop-down menu and choose the last item, "add terminals.” Indicate the number of taxa you wish to add in the pop-up window, and they will be placed at the bottom of your matrix. (This is a subset of the “expand matrix” function discussed above.)

Changing default settings: Many other attributes and settings be controlled through the winDADA menu system (esp. under the VIEW menu). Once set, these settings may be saved as described above under “Character state colors”.

Character properties: A number of character properties are listed at the top of the character dialog box. The default properties are “additive,” “activated,” and “unselected.” These can be changed at any time if the default properties are not those desired for the analysis of a particular character.

Character state colors: Character state and matrix colors may be personalized. These and other options are found under the VIEW menu as exemplified below. To see the personalized color selection for the character states (and their backgrounds), one needs to select the option “VIEW >Color code states”. Once set, these and other settings may be saved for the matrix (if the file is saved with *.winc format) or for general use of the program (if the settings are saved using the option “FILE >Save current settings as default”. When opening a matrix, the settings with which the particular matrix was saved last will be read automatically. The last saved “default” settings may also be read via “FILE >Read default settings”.
Cursor settings: Cursor display attributes can be modified to facilitate adding and editing data. In order to implement different cursor display modes, select “cursor settings” from the view drop-down menu. The “display data” option will allow for the display of taxon names and character names and states no matter what the location of the cursor in the matrix (see section 4b for details).

Expanding matrix: Select "expand matrix" from the matrix drop-down menu. By default, when using this feature, 10 terminals and 10 characters will be added to the matrix but one can modify the values to add virtually any number of taxa and characters to an existing matrix.

Exporting & creating files with information about the matrix, characters, &/or taxa: In winDADA mode, these options are found under the OUTPUT menu.

Outgroup selection: When organizing taxa in a matrix, remember that the first taxon (usually 0) will, by default, be treated as the outgroup. Other outgroups (including multiple terminals) may be selected later by rerooting in winCLADOS mode as discussed later under viewing trees.

Matrix fonts: Fonts may be controlled through “VIEW >Fonts”. Once set, these settings may be saved as described above under “Character state colors”.

Numbering terminals and characters: The default of WinClada is to number terminals and characters from zero. To number taxa and characters from “0” or “1”, select the appropriate option under the VIEW menu. Alternatively, both characters and taxa can be independently numbered from zero or one using the corresponding options under "Chars" or "Terms".

Reorder, delete, or deactivate taxa and characters: If you wish to move, delete, or deactivate characters or taxa, the most effective method is to first select those taxa/characters and then perform the desired action on them. All of these actions can be performed by making selections from the taxa and character drop-down menus.

Show character statistics: The statistics for each character and for all characters in the matrix can be displayed above the matrix under “VIEW >Show character statistics”. This option will display the amount of information contained per character and in the matrix, which is calculated by subtracting the minimum number of steps from the maximum number of steps (as also shown). One can also identify if and which characters are weighted, deactivated or coded as additive. For example, if the statistics for the matrix (at the left part of the section) show "WTS 145 / 153", this indicates that there are some characters weighted to zero (i.e., excluded from the analysis). If one observes the values in the columns of this display, above each character column in the matrix, one will see that some of the characters will have a value of "0". Also, if the statistics for the matrix show, for example, "ACTIV 0 : [153]", this indicates that all 153 characters in the matrix are activated, which can be corroborated by the symbol "y" present in all columns. Last, if the statistics for the matrix show, for example, "ADDIT -153 : +0", one can infer that all 153 characters were coded as non additive, which can be corroborated by the presence of the symbol "+" in all columns. If one does not wish to display these statistics, one can hide them by just unselecting the option under the VIEW menu.
II. Analyzing data (using NONA)

Simple tree searches are conveniently launched from the "ANALYZE" menu, which will submit a string of commands to NONA to analyze the data and return tree(s) to WinClada’s tree viewer.

The following four options from the ANALYZE menu are described below:

A. Heuristics
B. Ratchet (Island Hopper)
C. Incongruence test
D. Bootstrap/Jackknife

If more complicated tree searches are desired, then NONA (or various other programs) can be "spawned" and these more complicated searches can be orchestrated directly in NONA, using command lines, as detailed in the documentation available separately for NONA.

A. Heuristics (For data sets that don’t have many island problems or of less than ca. 120 taxa):

Choosing heuristics brings up the following window:

A standard strategy for an initial search is:

10000  Maximum trees to hold
25-100  Number of replications
1-10   Starting trees per rep
0      Random Seed

Then use default settings elsewhere in Heuristics window. Search. NONA will then find all the most parsimonious trees or as many as the "Maximum trees to keep" value was defined.

Below is information to facilitate an understanding what NONA is doing and selecting search strategies.

a. Maximum trees to hold: Lets you set the number of trees to be kept in memory. The default is 100. This is equivalent to the “hold” command in NONA. The logically consistent value for this box should be AT LEAST the product of the "number of replications" times the number of trees held in each replication (“starting trees per rep”) since
this will be the number of trees that NONA will keep when performing an additional TBR search (see "Multiple TBR + TBR" below).

b. **Number of replications**: Set the number of times you want the program to randomize the order of the terminals, create a Wagner tree, and submit it to branch-swapping (storing in memory as many trees as set with "starting trees per rep").

c. **Starting trees per rep**: Determines the maximum number of trees to keep in each replication of branch-swapping. The smaller the number, the less time spent swapping on one island, and therefore the search will be less extensive within that island. An appropriate balance should be found for each particular matrix and computer in which a reasonable amount of time is spent when exploring each potential island. This varies with the size and complexity of the matrices and can be adjusted by running a pilot search using relatively few replications and trees held in each of them. For large matrices (e.g., with hundreds of terminals), it is recommended to hold one or a few trees. For average sized matrices and computers, reasonable values for the pilot search may be 100 replications with 20 trees held on each. According to the speed of the analyses and the variation in length of the trees found on each replication, adjustments can be made in these parameters.

d. **Random seed**: The default value of the program ("0") is to use a pseudo-random number generator based on the time in order to select the sequence of addition of the replications. If this default is changed, NONA will use a particular random addition sequence of taxa that is defined by the number selected by the user. This is not recommended for common use purposes, since the general goal of an analysis is to find the most parsimonious trees, which is more likely by starting from random points in the tree space.

e. **Name of Stem**: enter a name of a file where the output will be written. Two different files are created. The first (with the .out extension) records the details of the search. The second (with the .tre extension) records the trees obtained by the search.

f. **The Search Strategy sub-box**: allows for fine tuning search strategies.
   1. **Multiple TBR** - searches for trees using tree bisection-reconnection method of branch-swapping for as many replications as indicated.

   2. **Multiple TBR+TBR** - searches for trees using tree bisection-reconnection method of branch-swapping for each of the replicates, then repeats this process the number of times indicated in the number of replications box. Then after reps, takes all equally most parsimonious trees found during reps, and TBR swaps those trees until the memory limit defined with “Maximum trees to keep (hold)” is reached.

   3. **Treefile+TBR** – takes a file containing a cladogram(s) already saved and branch swaps on it (them) using TBR until the memory limit defined with “Maximum trees to keep (hold)” is reached.

**B. Ratchet** (or island hopper) (for large and complex data sets; e.g., greater than ca. 120-150 taxa or with many islands):

One's matrix should be prepared properly with two steps, before performing the Ratchet.

1. Because of the reweighting step in the Ratchet, the matrix must include only potentially informative characters (i.e., mop up the matrix using the corresponding option under the "CHARS" menu and then delete the selected characters under "CHARS >Delete selected characters").

2. Under "VIEW >Numeric mode" switch all characters to numeric mode (i.e., states as 0, 1, 2 and not a, g, c, t, etc.). This is because WinClada submits the matrix to NONA for searches in Numeric format. If in DNA format, then some versions of WinClada will send up an alert window informing the user of this, which must be closed with the mouse before the procedure is performed. If you wish to have the ratchet runs performed sequentially and automatically (e.g., over night), then you only submit numeric matrices for Ratchet runs.

The Ratchet itself is initiated through the window below, opened under “ANALYZE >Ratchet (Island Hopper)”.

16
Explanation of "Options settings" box:

The default settings are "amb-, poly=". "poly=" allows polytomies to occur in trees; "poly-" does not. "amb-" resolves dichotomies, rather than a polytomies, only when support for the resolved branch is unambiguous; i.e., ambiguous support is insufficient. Support for a branch is considered unambiguous only when its length is greater than zero under all possible character optimizations. "amb-" will resolve a polytomy even if the resolution is ambiguously supported, i.e., is not supported under all possible optimizations.

If ambiguously supported nodes are treated as polytomous, swapping takes longer (per tree) than simply comparing the different unsupported dichotomous resolutions, but when branches are collapsed, fewer trees may have to be swapped if the data produce unresolved clades.

"Random constraint level": Nixon (1999) suggested that the Ratchet is more effective in finding islands by randomly constraining a subset of groups during each iteration. Usually when between 10 and 20% of the nodes are constrained during both the equal and weighted searches, releasing the constraints for the final search of most parsimonious trees.

The basics of the Ratchet and the parameters are as follows:

1. An initial (Wagner) tree is generated using a random taxon entry sequence, followed by TBR branch swapping holding as many trees as indicated under “# Trees to hold/iteration” (generally only 1-few trees should be held). The best 1-few tree(s) found are then held.
2. A random subset of the characters is selected and more heavily weighted (the default is currently set at ca. 10% of the characters, as you’ll see under “# Characters to sample”). Five to 25% reweighting generally works well, but this can be adjusted at will.
3. With the reweighted matrix, the shortest tree(s) found by swapping on the Wagner tree generated in step 1 are used as a starting point for a next round of TBR branch swapping.
4. Weights are then reset to original and the best trees found by swapping on the reweighted matrix are now swapped with the original weights to find the optimal tree on a new island.
5. The routine then goes back and repeats steps 2—4 for as many cycles as set under “# of iterations/rep” (200 is the default value), holding as many trees each round as set under “# Trees to hold/iteration” as set above.

The default of 200 generally works well and only rarely is it better to increase/decrease this number than changing the percentage of characters to be reweighted. Too few iterations will not get
Steps 1—5 constitute a single Ratchet run. The Ratchet maximizes finding new starting points and reduces the amount of time spent during each new search, retaining the most parsimonious trees. The reweighting strategy allows for more islands to be explored in shorter amounts of time than with conventional search strategies. However, as steps 1—5 effectively started from just a single Wagner tree, even the Ratchet may find suboptimal trees due to island problems. Therefore, it is recommended that one do sequential ratchet runs (e.g., 5-10; set under “Number of sequential ratchet runs” in the “Mult-ratchet” sub-box). WinClada will collect all of the shortest trees obtained from these sequential runs.

Ratchet output:

When the “Island Hop” button is pressed, you will see NONA begin the analysis. When complete the program automatically defaults to winDADA and the shortest trees found are shown. A file that shows the commands executed by NONA is written and called <filename>.pro. Another file containing all the trees corresponding to the original matrix search on each iteration is saved using the *.rat extension. The trees on this file will therefore be as many as the number of iterations and will have different lengths.

C. Incongruence Length Difference (ILD) Test

Homogeneity between two or more data sets may be measured using the Incongruence Length Difference (ILD) test (explained in Farris et al. 1994. Cladistics 10: 315-319; based on Mickevich and Farris, 1981):

For data sets $X$ and $Y$, the incongruence length difference is

$$D_{XY} = L_{(X+Y)} - (L_X + L_Y)$$

Where $L_X$ is the length of the most parsimonious tree from data set $X$

$L_Y$ is the length of the most parsimonious tree from data set $Y$

$L_{(X+Y)}$ is the length of the combined analysis

$D_{XY}$ is 0 when the two data sets agree on the same tree; it is large when minimizing homoplasy in one data set increases homoplasy in the other.

The ILD test uses this metric to quantify incongruence. Calculating $D_{XY}$ between a series of random partitions from the combined pool of data provides the null distribution with which the real $D_{XY}$ can be tested for significance. If there is significant disagreement between the two data sets, the added lengths of the resampled matrices will be significantly longer than the tree lengths of the original matrices. See Farris et al. (1994) for more information.

To run an ILD test:
- Open the different data sets in WinClada (e.g., rbcL vs. 18S genes). Select each matrix under “MATRIX >Select all matrices”.
- Choose “ANALYZE >Farris et al. Incongruence test (ILD)”. The following window opens:
- Fill in the parameters that you wish NONA to use for the searches. When “Run ILD test” is pressed you will see NONA run, then the output is written into a file, and a window will appear reporting the results of the test as follows:

```
S = 50, W = 50, Significantly incongruent at P = 0.0196 (RESULTS IN ARNRES)
```

**D. Bootstrap or Jackknife**

1. Bootstrap or Jackknife:
   a. Bootstraps and Jackknifes, as these are resampling procedures, work faster if one removes all non-informative characters from your matrix and save this as a new matrix (SAVE THIS UNDER A DIFFERENT NAME), with only informative characters, for bootstrap or jackknife analysis.
      --In winDADA mode, go to "CHARS >Mop uninformative chars" to select all uninformative characters.
      --Then "CHARS >delete selected chars".
      --Save this matrix under a different name (e.g., "matrix01_bootstrap.winc").
   
   b. Open this modified matrix.
   
   c. Usually, you will want to display support values on a most parsimonious tree or a consensus tree. Open this tree (below, left).

   d. Under "ANALYZE >Bootstrap/Jackknife/CR with NONA", the following window opens:

   ![Bootstrap/Jackknife window](image)

   Indicates that 1 tree (e.g., the MPT) is open.

   e. "Number of replications" refers to the number of bootstrap or jackknife replicates. Usually this is done with a minimum of 100 to 1000 replicates, such that robust support percentages are obtained.
f. "Number of search reps (mult*N)", "Starting trees per rep hold/", "Random seed (0=TIME)" are analogous to the search parameters under a conventional search, so see under "Heuristics" above for a description of these.

g. Additional parameters in this window are:
   1.) Max* option: The default is to not do a max*, which would exhaustively swap after each rep on the shortest trees found. To do otherwise would provide for a more thorough search, but is more time consuming. One may decide that time is better spent doing additional reps to find more distinct islands, as trees on these are more different from each other than those trees on the same island. However, doing a max* is an option.

   2.) "Save Consensus" is the default option, this being that bootstrap or jackknife frequencies are calculated on the basis of the strict consensus tree obtained from each replicate. This "strict consensus" jackknife or bootstrap yields frequencies that are equal to or lower than those obtained with the "frequency-within-replicates" method, because a group that occurs in some but not all trees within any particular replicate receives a score of zero for that replicate (signifying absence from the consensus tree), rather than a score greater than zero, signifying its frequency of occurrence among optimum trees from that replicate (Soreng and Davis 1998; Grass Phylogeny Working Group 2001).

h. Upon pressing "bootstrap" or "jackknife", NONA will open and the search is conducted. Following the procedure, trees obtained are deposited in winCLADOS (WinClada's tree viewer). If 100 replicates are chosen and the strict consensus method is chosen, then 101 new trees will appear and frequencies may be viewed. The first 100 of these new trees are those from the 100 "consensus" replicates. The 101st (i.e., last) of these new trees is a majority rule consensus of all of the 100 bootstrap or jackknife trees.

i. To view the support percentages obtained above on your most parsimonious tree or consensus tree, then return to this tree ("TREES> Go to tree" opens the following window). If you opened this tree before the bootstrap or jackknife analysis, then this tree is the first tree (number 1). Ignore the title of this little pop-up window. Despite its "New Matrix/Resize matrix" title, it is not the window to create or resize a matrix.
Trouble making the bootstrap support values go away from tree? Under "NODES >show support(toggle)", toggle on and off to make support values go away and reappear.

To get back to mode where you can view hashmarks on branches? WinClada has a bug here. The only way to return from bootstrap frequency viewing mode to that where hashmarks can be seen is to exit the program and reopen the matrix with tree.
III. Viewing, manipulating, and printing your trees (winCLADOS mode)

All trees viewed in WINCLADA must be related to an open data matrix with corresponding terminal taxa. If no trees are in ram, a pectinate tree is generated using the order of taxa as in the matrix. This tree will be produced simply by opening a data matrix and then selecting WinCLADOS from the "interface" menu. THIS DEFAULT TREE HAS NO REAL MEANING. Do NOT publish it.

A. Viewing and working with trees

As already stated, trees following tree searches from the “ANALYZE” options will automatically be shown under winCLADOS mode following a search.

You can switch between views of the Matrix and the Trees by toggling between “winDADA” (matrix view) and “winCLADOS” (tree view) mode under “INTERFACE” menu.

1. To see trees previously saved in a tree (.tre) file:

Choose “FILE >Open tree file”. In the browser, choose the desired tree file (e.g., the one created during a previous analysis of the data set.

2. Tree styles

Along the top just above the tree are a series of buttons for changing the way the tree looks.

a. The first two allow one to zoom in (the green button with a “z”) and out (the yellow button with a “Z”). One can also zoom up or down on a tree by toggling between the “z” key and the “shift+z” key.

b. The position of the tree on the screen can be changed with the arrow keys.

c. The tree can be compressed or spread out using the F3, F4, Shift+F3, Shift+F4 keys, or these buttons:
d. The style in which the tree is drawn can be changed by clicking these buttons:

For more tree styles, select "tree styles" under VIEW menu.

**Note:** The use of "rectangular" or "smooth rectangular" branching allows for the easiest placement of character data on the trees. All of the following discussion is predicated on the use of this tree style.

e. The font and size of the taxon names can be changed under "VIEW >Fonts".

f. Use F2 and shift-F2 to adjust the width of the lines in the trees.

3. Viewing optimizations / Tracing characters on tree

To trace characters &/or character states on a tree, use the “HASHMARKS” menu or click the button with a picture of a character state on a branch:
Using the default options, black hashmarks indicate nonhomoplasious changes (synapomorphies or autapomorphies). White circles show homoplasious STATES. This means that a character that has a reversal will have one of the states (the one that is not reversing) as nonhomoplasious, while the one that reverses will be homoplasious. This feature, as well as the appearance of the hashmarks can be changed using the options under "HASHMARKS >Hasmark properties".

Labeling the character state changes can be accomplished by clicking these buttons:

- Clicking on the state button (B) a second time will show state transformations (e.g., "1 > 2")
- OR: Choose “HASHMARKS >Number hashmarks (characters)” and the number representing each character will appear above the hashmarks.
- OR Select “Number hashmarks (states)” and the state for each character appears below the hashmark.

If character and state names have been entered in the matrix, words instead of numbers can be used to label branches:

- Choose “HASHMARKS >Label hashmarks with Char-Names” and the name of each character will appear above the hashmarks.

The character state changes can also be mapped on as colors (“MacClade Style”) by clicking the blue button marked with an “X” (this is equal to applying the “Diagnoser toggle” under the DIAGNOSER drop down menu). When this happens a dialog box opens that shows what state corresponds to what color. Arrow keys allow you to click and scroll through all the different characters:
5. Manipulating taxa and branches on a tree:

Using the "mouse mode" functions under the EDIT/MOUSE drop-down menu will allow you to manipulate the position of taxa on the tree. These features may be useful for creating particular tree topologies and fitting characters to them. Five particularly useful features are:

The default setting is to label unambiguous optimizations only. One can use this dialog box to look at the alternative optimizations such as "fast" and "slow" as shown here.

Alternatively, one can use the "OPTIMIZATION" menu to choose "Fast" or "Slow" optimizations.

**Note:** The optimizer drop-down menu will allow for the choice of optimizations (unambiguous, fast, slow) for all characters on the tree. If one wishes to use different types of optimization for different parts of the tree one can use the "force optimization" option.
a. **Move branch mode:** Select the branch you wish to move by clicking on that taxon or node with the left mouse button. Move the cursor to the destination node or terminal. Click the right mouse button. The selected taxon (node) will be moved to that point.

![Select Mouse Mode](image)

b. **Reroot mode:** Note: Default rooting is with taxon zero (0) in matrix. Use reroot mode to assess the effects of placing the root of the tree in different positions. Select "reroot mode" on the pop-up. Then, to move a node to the root position click on the node; to move a terminal to the root position, double-click on the terminal.

c. **Hide/Alias mode:** This feature is particularly useful for analyses with numerous taxa. When you apply this radio-button, one can rename a node that contains many taxa and summarize the information. If one double clicks on a node a dialogue box will appear in which one can type the desired substitute name for the node. For example, one can rename a clade that contains only species of the coffee family as Rubiaceae. After entering the new name the clade will be reduced to that name and in between brackets there will be the number of species contained in the clade. A right mouse click on the node will show again the original clade in addition to the summary information. The hide/alias feature complements with the next one.

d. **Subtree mode:** With this option one can create a subtree that includes only part of the complete cladogram. For example, if one used the previouse option to indicate where the members of the Rubiaceae family fall in the cladogram, but one desires to also show the relationships among the members of this family, one can left click on the mouse to only show that clade. Again, a right click on the mouse will reshow the original cladogram.

e. **Select node mode/Select clade mode:** With these two options, one can select and highlight even combination of nodes or clades on a tree. Printing with these selected with print their highlighting too. Once selected, one can opt to create text files (under Output menu) describing various aspects of these nodes or clades (such as state summaries).

### B. Ancestor Reconstructions

A text file can be created with the reconstructions for all "activated" characters for any number of selected nodes or clades. One can select any combination of clades with the cursor (EDIT/MOUSE> Mouse modes> Select clade mode), nodes (Mouse modes> Select node mode), or both. In the appropriate mouse mode, click once (to select clades) or twice (to select nodes). Then under OUTPUT menu, select "DESCRIBE" to create the file. Open this file and view or print all states optimized to the nodes that were selected when creating the file. An example of a nodal reconstruction is given below (in this example for ecological characters):

```
NODE 39: (Thamnochortus erectus, Thamnochortus insignis)
Altitude 0-300m; Avg. Ann. Rainfall 400-600 mm; Rain Seasonality winter rain; Fire Survival re-sprouting; Ground Water Availability none; Bedrock TMS; Soil_Rockiness pebbles_to_boulders; SE_Clouds absent.
```
C. Consensus & Compromise Trees

Consensus trees can be easily and quickly calculated when in tree viewing (winCLADOS) mode. With multiple trees open in the winCLADOS treeviewer, go to "TREES> Consensus Compromise". Choose from the following:

1. Consensus (Strict)
   Creates a tree in which only clades that are present in ALL most parsimonious trees (MPTs) are resolved (i.e., all clades not present in all MPTs are collapsed).

2. Nelsen (Collapse + consensus)
   This option is clarified by Nixon & Carpenter (1996. Cladistics 12: 305-321).

3. Majority Fools
   Otherwise known as “Majority Rules”, this creates a consensus tree in which clades present in a majority (but not all) of the MPTs are resolved.

Note: The consensus tree you calculate is added to the end of the trees you currently have open in WinClada, but only in ram memory. For example, if your analysis returned 100 MPTs, then your consensus tree will be the 101th tree. If you wish to save it in a different file you can use "TREES >Save current tree to file".

D. Printing or Exporting Trees

1. Printing
   Choose “FILE >Print preview” and use arrow keys, zoom keys, etc. to make the tree have the desired appearance on the page. Note that WinClada uses standard USA paper size (8.5 X 11 inch).
   Then choose “FILE >Print”.

2. Exporting as graphics file
   A tree can be saved in a “metafile” (*.emf). This makes it easy to prepare trees for publication or on-screen presentations by transforming the trees into a format that can be edited and manipulated in various graphics programs (e.g., Adobe photoshop, Adobe Illustrator, Powerpoint, etc.). After inserting the metafile image into, for example, a Powerpoint page one can double click on the image (or "ungroup") and it will be decomposed onto individual elements that can be independently edited. Adobe Illustrator is the most powerful program for editing tree metafiles.

E. Summary of Key Commands for Trees (in winCLADOS mode):

[arrow keys]: moving tree in the view window.

[F2]: thin tree branches.

[shift] + [F2]: thicken tree branches.

[F3 or F9]: compact tree horizontally.

[shift] + [F3 or F9]: expand tree horizontally.

[F4]: compact tree vertically.

[shift] + [F4]: expand tree vertically.

SHIFT-Z and z: Zoom tree in and out (scales tree). This not only affects the size of the tree on the screen, but also has a direct effect on the size of the tree printed.

SHIFT-X and x: (In tree print view only) Zooms out and in on page view.

[F5]: decreases width of hasmarks.
[shift]+[F5]: increases width of hasmarks.

[F6]: decreases height of hashmarks.

[shift]+[F6]: increases height of hashmarks.

[F7]: scroll down through trees in ram.

[F8]: scroll up through trees in ram.

CTL-HOME, HOME. CTL-HOME repositions the tree to the upper left-hand corner of the winCLADOS view window (i.e., the "home" position). HOME will return the tree to the left side of the view window.

[S] and [U]: Selects and unselects the current tree.

**F. Functions of Icons on tool bar at top of tree view:**

- switches to matrix interface
- move to next tree.
- move to previous tree.
- Expand and contract tree horizontally and/or vertically.
- Various tree forms.
- Place lollipop, lollipop around terminal name. May change color under VIEW menu.
- Show character state below hashmark. Pressing this twice will show state transformation.
- Show character number (or name) above hashmark.
- Show character hashmarks.
- Character Diagnoser toggle; to look at optimization of each character, one at a time, on tree. States are color coded.
- Return tree to normal size.
- Increase tree size proportionally.
- Decrease tree size proportionally.

**G. Final Notes and Comments about winCLADOS mode:**

**Note:**
Unlike CLADOS and NONA, tree ram space is automatically allocated when trees are read in from files. The initial value is 100 trees. When reading a file, if more trees are encountered, the ram space is increased to a value based on current trees + new trees + 100. You will receive a message if this reallocation of ram occurs.
If you need more space, you can also directly increase tree space (= hold in NONA) under “TREES > Rezise maxtre [ram space]”.

**READING TREES:**
WinClada checks for a match in number of taxa in the tree and in the matrix and generates a message if there is a discrepancy. If the tree file contains more taxa than in the current matrix, the read will be aborted without reading any trees. KCN is working on ways to automatically deal with mismatches that will cause fewer difficulties. In the meantime, if you get such a message, check that the tree file was generated with the same taxa.

**WARNINGS, ADVICE ABOUT TREES:**
Confusion still persists about the way in which NONA stores trees. By default, when using NONA, if you save trees with the “sv” command, trees are saved as fully dichotomous - even if some branches are only ambiguously supported (i.e., the minimum length under some optimization is 0 steps). WinClada simply reads a tree file - and if the trees are dichotomous, it displays them as such. Thus, “unsupported” branches may be displayed. However, under the default optimization (unambiguous changes only) if characters are mapped there will be no character changes at those nodes.

To display ONLY branches that have a minimum length > 0, set the “VIEW” to collapse unsupported branches. Note that in version 0.9, this may not display correctly with jetson (curved) trees. We suggest using cladus, jetclad, or gothic trees if the option is set.

Note, that under FAST or SLOW optimization, branches that have a minimum length 0 may still have length under that particular optimization, and will be shown as such.