

LIFE AND EARTH HAVE EVOLVED TOGETHER

LEON CROIZAT

ANALYSES AND SOFTWARE HAVE TO EVOLVE TOGETHER

THE AUTHORS



DANIEL RAFAEL MIRANDA-ESQUIVEL & SUSY ECHEVERRÍA

# MARTITRACKS MANUAL

PUBLISHED BY THE AUTHORS

Copyright © 2010 Daniel Rafael Miranda-Esquivel & Susy Echeverría

based on tufte- $\text{\LaTeX}$ , please refer to [tufte-latex.googlecode.com](http://tufte-latex.googlecode.com) for further information

Licensed under the Apache License, Version 2.0 (the “License”); you may not use this file except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “as is” basis, without warranties or conditions of any kind, either express or implied. See the License for the specific language governing permissions and limitations under the License.

# *Contents*

*Forewords*      13

*Introduction*      15

*MartiTracks, the program*      23

*Worked example: testo.dat*      29

*MartiTracks commands*      33

*Bibliography*      37

*Index*      43



## *List of Figures*

1	Individual tracks from testo.dat	32
2	Generalized tracks from testo.dat	32





## *List of Tables*



*Dedicated to those who appreciate good  
Biogeography and love dogs.*



## Forewords

This manual is intended to be a clear and compressive introduction to Panbiogeography and geometric track analyzes, using the program *MartiTracks*. The book is divided into four main chapters and some appendices. First chapter is a general introduction to Panbiogeography and related terms, second chapter presents *MartiTracks*' algorithm while third chapter is devoted to commands used. The fourth and final chapter presents some theoretical and empirical data and its analyzes to explore *MartiTracks* options.

The program was developed using Free-Pascal, as Pascal is an intuitive and easy to handle language for programming, the debate is open to consider whether C/C++ or Python/Perl/Java are more suitable languages for this kind of problem. As a free software project, feel free to criticize the code or the algorithm, and if you improve it, we will be more than delighted and thankful.

This manual has been developed using Tufte-L<sup>A</sup>T<sub>E</sub>X, maybe (in our humble opinion), the best environment to write a book or a handout.

We are grateful to many people, .....

We are open to new ideas or suggestions to improve our analyzes, the program, and this manual itself.

° [http://ciencias.uis.edu.co/  
labsist/pantrack](http://ciencias.uis.edu.co/labsist/pantrack)

° <http://www.freepascal.org>

° [http://code.google.com/p/  
tufte-latex](http://code.google.com/p/tufte-latex)



# *Introduction*

## *Panbiogeography*

Panbiogeography is a unique and very useful approach to capture any distributional pattern/structure in studies that use geographical data. Mapping biogeographic tracks is a cost efficient way to reduce the initial complexity that we find in the data sets (Craw, 1989b), as Croizat's tracks analysis (Croizat, 1952, 1958) provides a conceptual framework for understanding the biogeographic structure and the relationships of areas.

Panbiogeography was developed by Croizat in his books, "Manual of phytogeography"(Croizat, 1952), "Panbiogeography"(Croizat, 1958) and "Space, Time, Form: The Biological Synthesis"(Croizat, 1964). Panbiogeography was inspired by Croizat's idea of connecting two essential elements in evolution: Space and Time: Earth and Life evolve together, which means that biotas evolve together with the geographic barriers (Morrone and Crisci, 1995; Morrone, 2000).

Croizat's panbiogeography constituted a unique critique to Darwin's biogeography ideas about natural selection, means of dispersal and geographic distribution (Craw, 1987). This approach challenged the idea that means of dispersal are the principal factors responsible for the evolution of the distribution (Craw, 1989a; Craw *et al.*, 1999).

Panbiogeography assumes that taxa distribution evolves through two stages: mobilism and immobilism. The interaction of these two states is called vicariant form-making model. Within the mobil-

ism stage, an ancestor taxon expands to establish on new territory through its means of dispersal ("means of survival") when the geographic and climatic factors are favorable. Later, when the range geographic is established (immobilism) the appearance of barriers allows the isolation and differentiation of taxa (vicariant form-making) (Grehan, 1988; Michaux, 1989; Craw *et al.*, 1999; Crisci and Morrone, 1992).

Panbiogeography has been considered an independent research program because is the only approach that focuses on the spatial or geographical sector as a fundamental pre-condition to any analysis of the patterns and processes of evolutionary change (Craw *et al.*, 1999; Grehan, 1994, 2001a).

The assumptions that make the Panbiogeography different from other historical biogeography approaches are (Craw *et al.*, 1999, :19): "1. Distribution patterns constitute an empirical databases for biogeographical analysis.

2. Distribution patterns provide information about where, when, and how animals and plants evolve;

3. The spatial and temporal component of these distribution patterns can be graphically represented;

4. Testable hypothesis about historical relationship between the evolution of distributions and Earth history can be derived from geographic correlations between distribution graphs as geological/geomorphic features."

### *Track analysis*

Panbiogeography as a biogeographic method is called track analysis. There are four principal concepts within this method: track, node, main massing, and baseline (Crisci and Morrone, 1992; Craw *et al.*, 1999; Grehan, 2001a; Espinos-Organista *et al.*, 2002).

Individual tracks are the basic units of panbiogeography. These



are lines drawn on a map that connect different localities or distribution points of a particular taxon or group of taxa, so that the sum of the segment lengths that connect all the distribution points is the smallest possible. In the graph theory words, the individual track is a minimum spanning tree (Craw *et al.*, 1999; Morrone, 2004; Page, 1987).

Generalized tracks or standard tracks are called repetitive patterns because summarized distributions of diverse individual taxa (Michaux, 1989). These are lines on a map resulting from the superimposition of the individual tracks. Generalized tracks are interpreted as distributional patterns of ancestral biota that had been fragmented by tectonic and climatic events (Craw, 1988b). These panbiogeographic element is considered as a conjecture (=hypothesis) of a common biogeography history or primary biogeography homology sensu Morrone (2001), which means that analyzed taxa are spatiotemporally integrated in a biota (Craw, 1983; Morrone, 2001, 2004).

Nodes are areas or localities where two or more generalized tracks are overlapped. These are complex areas or tectonic and biotic convergence zones (Craw, 1988a; Crisci and Morrone, 1992; Morrone, 2004; Page, 1987). The nodes are considered priority areas for conservation or sites of biological endemism because represented localities of high diversity, distribution boundaries, disjunction, incongruence and recombination, specimens that are difficult to identify and unusual hybrids (Craw *et al.*, 1999; Morrone and Espinosa, 1998; Grehan, 1993).

The Panbiogeographic method involves basically main three steps (Morrone, 2004): Firstly, construction of two or more taxon individual tracks (minimum spanning tree from distributional localities). Secondly, delimitation of generalized tracks through geographic congruence between individual tracks. Finally, determination of nodes

within the intersection areas between generalized tracks.

There are different approaches within panbiogeographic method. Croizat's manual reconstruction (Croizat, 1958, 1964), Page's Spanning graphs (Page, 1987), Craw's Track compatibility (Craw, 1989b) and PAE ("Parsimony Analysis of Endemicity") (Rosen, 1984; Craw *et al.*, 1999; Luna-Vega *et al.*, 2000; Morrone and Marquez, 2001)

### *Croizat's manual reconstruction*

Croizat's manual reconstruction is done by drawing on a map taxon's individual tracks and then overlapping them to determinate generalized tracks. Basically the method consists of the following steps (Morrone, 2004):

1. Connect the disjunct areas of distribution by lines forming individual tracks,
2. The individual tracks are oriented using baselines,
3. Overlap individual tracks to determinate generalized tracks,
4. Recognize the nodes where two or more generalized tracks are intersected,
5. Indicate on a map the generalized tracks, baselines and nodes.

### *Page's Spanning graphs*

Page's Spanning graphs was the first attempt to quantify the panbiogeographic analysis using graph theory. Its main methodological steps are (Page, 1987):

1. connect the disjunct areas of distribution through a MST (Minimum Spanning Trees) forming an individual track for each species,
2. Construct incidence and connectivity matrices for individual tracks to recognized shared elements and track congruence,
3. Construct a connectivity matrix for all tracks together to search circuits that indicate incongruence tracks,
4. Indicate on a map the generalized tracks, baselines and nodes.

### *Craw's tracks compatibility analysis*

Craw's tracks compatibility analysis (Craw, 1989b) involves finding the largest clique of compatible tracks through a locality/distribution per track matrix (Craw *et al.*, 1999). In general, its algorithm involves the following steps (Craw, 1989b; Grehan, 2001b; Morrone, 2004):

1. Construct individual tracks for different taxa from distribution localities using a MST,
2. Generate a matrix (areas x individual tracks), where the presence of an area is represented by "1" and its absence by "0".
3. Find the largest clique for compatible individual tracks which is considered the generalized track.
4. Evaluate statistically the generalized track,
5. Indicate on a map the generalized tracks, baselines, and nodes.

### *Parsimony Analysis of Endemicity*

PAE ("Parsimony Analysis of Endemicity") has been considered a panbiogeographic method (Craw *et al.*, 1999; Luna-Vega *et al.*, 2000; Morrone and Marquez, 2001; Morrone, 2004). Its algorithm includes the following steps (Morrone, 2004):

1. Construct individual tracks for different taxa from distribution localities using a MST,
2. Generate a matrix (areas x individual tracks), where the presence of a species is represented by "1" and the absence by "0". Within this step, an extra area is added with "0" (all species absence) to polarize the topology (incorrectly called an 'area cladogram' (Garzón-Orduña *et al.*, 2008)),
3. Analyze the matrix using a parsimony algorithm. Each clade supported by at least two individual tracks (as "synapomorphies") is considered a generalized track,
4. "Disconnect" or delete the species that support the different clades and reanalyze the matrix to search clades that are supported

by other taxa (Garcia-Barros *et al.*, 2002),

5. Indicate on a map the generalized tracks, baselines and nodes.

Since Croizat (1958, 1964) presented his panbiogeographic method, the above mentioned quantitative techniques had been proposed, Page's graph theory method (Page, 1987), Henderson's analysis (Henderson, 1989), which is similar to Page's analysis, and Craw's compatibility track analysis (Craw, 1989b). These techniques have emerged taking into account Croizat's framework for analysis: individual tracks congruence to define generalized tracks (Page, 1987; Henderson, 1989; Craw, 1989a). Page (1987) suggested a quantitative approach using graph theory by incidence and connectivity matrices, but this method has not been applied to real data, perhaps because of its difficult calculation and computational complexity. Craw (1989b) proposed the compatibility track analysis based on distributional compatibility which is analogous with Meacham's character compatibility approach (Meacham, 1984) to phylogenetic systematics (Craw *et al.*, 1999). Within the latter method, two or more individual tracks are considered compatible if either one is included within or replicated by the other. Morrone and Crisci (1995) claimed that the compatibility in Panbiogeography is used in a restricted way depending on the method. "Non-overlapping tracks" are incompatible according to compatibility track analysis and should be compatible in Croizat's track analysis by geographic proximity.

Within the quantitative techniques proposed in the history of the panbiogeography (Page, 1987; Craw, 1989b; Henderson, 1989), only Craw's compatibility tracks analysis has been implemented and now part of the algorithm, the minimal spanning tree, has been automated in the software package "Croizat" (Cavalcanti, 2009).

Grid analysis such as PAE (Parsimony Analysis of Endemicity) has been considered a panbiogeographic approach (Craw *et al.*, 1999; Luna-Vega *et al.*, 2000; Morrone and Marquez, 2001; Morrone,

2004), but some authors do not consider it a historical biogeography method (Humphries, 2000; Garcia-Barros *et al.*, 2002; Garzón-Orduña *et al.*, 2008). Although this method has been widely used (Luna-Vega *et al.*, 2000; Morrone and Marquez, 2001; Huidobro *et al.*, 2006; Mihoc *et al.*, 2006; Espinosa-Perez *et al.*, 2009; Contreras-Medina *et al.*, 2007), its implementation to the panbiogeographic analysis was secondary. Its direct objective is the analysis of endemism and its relation to the track analysis sensu Croizat could become ambiguous.

Despite of these methods, Croizat's manual reconstruction has been the most used in panbiogeographic studies because its simplicity and the alleged direct connection to Croizat's panbiogeographic concept. Nevertheless, the debate about the reliability of the manual reconstruction is still valid. The method turns to be ambiguous and generates subjective results with large data set due to overcrowded points (Franco-Rosselli, 2001; Liria, 2008).

### *Suggested literature*

- *Craw et al.* (1999) is a good summary of theories and methods, while the empirical approach could be flawed.
- *Crisci* (2001) is an historical account of biogeography, and is illustrative of the so called forces that shaped present biogeographic thought.



# *MartiTracks, the program*

## *Definitions*

*MartiTracks* algorithm has been presented in our paper (cite), therefore we suggest you to read the paper first.

The program tracks MST or segments, therefore you can conduct your analysis in either of those two ways.

## *The program and its calculations*

If you use MST, the species' MST will be compared to all species MST, comparing each segment but outputting the answer by species, so you can expect at most  $nsp$  possible tracks, where  $nsp$  is the number of initial species. As all segments or MST are not congruent, the real number of answers lies in a number smaller than that.

° If you are not familiar with graph theory, you might find these links useful:

[http://en.wikipedia.org/wiki/Minimum\\_spanning\\_tree](http://en.wikipedia.org/wiki/Minimum_spanning_tree)

[http://en.wikipedia.org/wiki/Glossary\\_of\\_graph\\_theory](http://en.wikipedia.org/wiki/Glossary_of_graph_theory)

## *MartiTracks files. General information*

The program uses two input files,

1. -the input file with the distributional data. (Only the input file is mandatory).
2. -the parameters/orders file with the values for parameters and the commands to be executed in bash  
and creates a kml output file

*MartiTracks input file*

mt05 accepts as an input, a text file comprised of three columns, labels lat long separated by tabs or spaces but not commas (,), most spreadsheet and SIG programs can output such text files, or you can use awk/bash to reshape them.

As an example we include

- a bash script file that converts a comma delimited file in a mt05 input file (csv2mt.sh)
- an awk script file that converts a GlobalMapper output in a mt05 input file (gm2mt.sh)

*MartiTracks output file*

mt05 writes the generalized tracks (and in some cases, points) to the kml output file, including or not the individual tracks. The output is readable using google-earth or GIS programs as QGIS. As some programs may or may not read the output, we only keep compatibility with google-earth/QGIS. If your beloved GIS program does not read the output, please consider read it with google-earth/QGIS.

*Download the program and a data set*

You must download a binary for your platform of choice, from:

<http://tux.uis.edu.co/labsist/martitracks> or

<http://code.google.com/p/martitracks/>

Binaries are provided for Linux 64, Win 32 and 64

and the data sets from:

[http://tux.uis.edu.co/labsist/martitracks/data-example.](http://tux.uis.edu.co/labsist/martitracks/data-example.zip)

zip

° a valid data set = test0.dat

```
sp1  1  9
sp1  3 11
sp1  6 12
sp2  1  9
sp2  1 10
sp2  3 10
sp2  4 11
sp2  5 12
sp2  6 12
sp3  4 13
sp3  5 11
sp3  8  8
sp4  4 12
sp4  6 11
sp4  7  8
sp5  8  8
sp5  7  6
sp5  8  2
sp6  8  8
sp6  7  5
sp6  8  3
```

° note: if your label field includes inner tabs/spaces, these will be considered as field separator therefore are NOT allowed, you can replace such characters by an underscore or delete them all (mt05 converts the dash to underscore)

° A common mistake is different labels for the same species, although mt05 is not case sensitive, it is (quite) sensitive to misspellings therefore Simulium and simulium are the same, while Simulium, Simuliun or Simuliun are not.

° with Linux version, you might need to convert the file mt05-xx to an execute file by typing in a command-line window:

```
$ chmod +x mt05
```



## Command modes

*MartiTracks* has two uses-interfaces:

1. A text user interface, and
2. a command-line (bash like) interface.

in the text user interface (TUI), the user can choose among different options including: changes of parameters for analysis, track a pair, groups or the whole data, find the index of congruence (IC) between pairs of tracks, or among several tracks. print kml file, etc. The command line interface was created for search strategies previously defined. The input file, output file and parameters files must be defined.

<sup>o</sup>[http://en.wikipedia.org/wiki/Text\\_user\\_interface#TUI\\_under\\_Unix-like\\_systems](http://en.wikipedia.org/wiki/Text_user_interface#TUI_under_Unix-like_systems)

### Text user interface

You can use the text interface by simply typing at the prompt:

```
$ ./mt05
```

for Linux, and

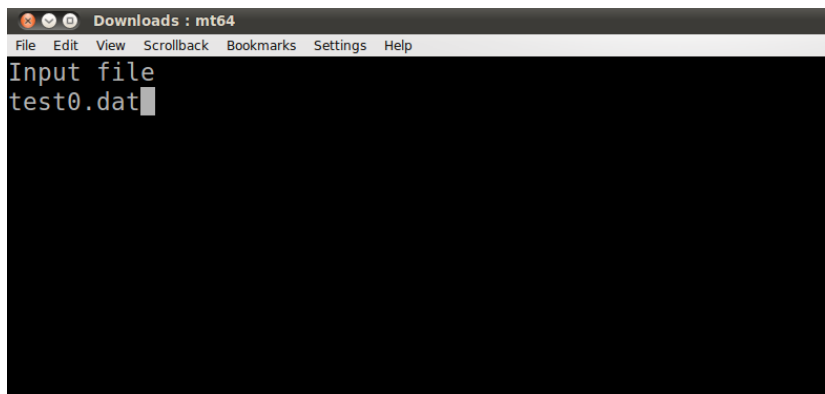
```
$ mt05-winXX.exe
```

for Windows in a command-line window

or by clicking on the *MartiTracks* icon (win. only).

**1. Enter the input name:** As soon as you open *mt05*, it asks for the input file name:

<sup>o</sup> TUI typography: commands will be [C]ut, to indicate that the instruction is cut and the letter to be pressed is c]



**2. Set up the analysis:** Once you specified the input name, a list of commands help the users to set up the analysis.

```

File Edit View Scrollback Bookmarks Settings Help
MARTI TRACKS

Press a key
[ ]

TRACK =
[a]ll [g]roup [p]air
[t]rackpair track[s]all
[j]oin two tracks
join a gro[u]p of tracks
[l]ist tracks
[v]italstatistixs
[w]e write to kml

Congruence =
gr[o]up co[n]gru pair

Decision rules =
[c]ut [r]ules
[m]inimal congruence

Original tracks= 6 Total tracks= 6
Repeated= 0 New tracks= 0
Cut=0.25 / R1=0.50 / R2=0.25 / R3=0.75
print kml FALSE / min congruence= 0.85
Beta I 04-a / 16 07 2010 \ Input:test0.dat \ Output:test0.dat.kml

* q=close+quit
* z=delete prev. analyses

```

Then we must define the values of the parameters, to be used in the analysis.

There are five parameters:

[c]ut value,

three [r]ules of decision, and

the value for [m]inimal congruence.

Possible **parameters** values are:

[c]ut		<i>set cut &lt;real value 0-360&gt;</i>
[r]ules	lmax	<i>set lmax &lt;real value 0-360&gt;</i>
	lmin	<i>set lmin &lt;real value 0-360&gt;</i>
	maxline	<i>set maxline &lt;real value 0-360&gt;</i>
[m]inimal congruence		<i>set ci &lt;real value 0-1&gt;</i>

These values are used to calculate the similarity among individual tracks in any analysis, joint or track.

You might use the predefined valuer or set up them at your choice. Press the appropriate letter to change a parameter, and enter the value. After this, we are ready to conduct our analysis.

\* cut value is a real <0 - 360> value expressed in degrees  
 [0.0] = no cut - [360.0] = all points will be collapsed to one.  
 \* rules are real <0 - 360> values expressed in degrees  
 \* minimal congruence is a real <0 - 1> value.  
 [0.0] = no congruence - [1.0] = totally congruent.  
 please refer to our paper (cite) for further information

As we might have similar initial MSTs, we need to joint those MST that are the same according to the minimum value of **[m]**inimal congruence fixed. We must press **[u]**, to join the gro**[u]**p. The program will ask the number of the initial and final tracks to be joint. Now we must use **[a]ll**, to find the congruent segments and define the generalized tracks or distributional patterns of species.

Finally, we need to eliminate those redundant generalized tracks, typing again **[u]**, but joint from the track number number of species + 1 to the track number of total tracks, because first tracks are individual tracks.

If, we want to write our results in a kml file, we must type **[k]** and **[+]** to activate the output to the kml file. The kml info must change from FALSE to TRUE. Then, we will use **[w]** if we want to write the whole information of the analysis including: individual tracks, and generalized tracks, or we can type **[e]** to write only the generalized tracks into the kml file.

For the command-line user interface we need to specified: the input file, output file, and parameters file (see page 33).

```
$ ./mt05-64 <input file> <output file> <parameters file>
```

° please remember:

when you press a letter a command is executed

when you are asked for a value you must finish the input using [enter] or [return]



° You must download the program and the example data set, please see page 24.

## Worked example: *testo.dat*

First of all, we must open the program (see page 25), read the data set and define the values of the parameters, to be used in this analysis. For *testo.dat* we will use the following values:

```
[c]ut          0.25
[r]ules        0.75 0.5 1
[m]inimal congruence  0.8
```

**Search options:** Once the parameters have been defined, we will search for the general patterns of distribution.

As we might have similar initial MSTs, we need to reduce them taking into account the similarity among the individual tracks to join (gro[u]p of tracks) and the value of [m]inimal congruence fixed.

In the case of *testo.dat*, we gro[u]p from the track number 1 [first] to the track number 6 [last]. Thus, we find whether there are similar individual tracks that can be consider as the same track, and those tracks will be joint.

Later, we might need to redefine the values of the paramaters, in the case of *testo.dat* we will set:

```
[c]ut          0.5
[r]ules        1.5 1 2
[m]inimal congruence  0.8
```

Then, Typing [a]ll, we will find the congruent segments for each individual tracks in order to delimitate the generalized tracks or distributional patterns of species.

° please remember:  
to set each of these parameters you have to type the letters inside the square brackets [ ] and write the corresponding value followed by and [enter] or [return]

Finally, we need to eliminate those redundant generalized tracks, typing again [**u**], but joint from the track number 7 to the track 9, because first six tracks are individual tracks.

### Write a kml file

Now, we need to write our results (the generalized tracks) in a kml file. To do that, first we must type [**k**] and [**+**] to activate the output to the kml file. The kml info must change from FALSE to TRUE. Then, we will use [**w**] if we want to write the whole information of the analysis including: individual tracks, and generalized tracks, or we can type [**e**] to write only the generalized tracks into the kml file.

° .kml file extension is mandatory with googleeearth, otherwise googleeearth will complain and will not open the file

### Command line interface

For the command-line user interface we need to specified:

the input file,  
the output file,  
and the parameters file.

For linux 64 bits:

```
$ ./mt05-64 <input file> <output file> <parameters file>
```

For Windows 64 bits:

```
$ mt05-64.exe <input file> <output file> <parameters file>
```

For Windows 32 bits:

```
$ mt05-32.exe <input file> <output file> <parameters file>
```

**Output file:** it must have an extension .kml.

**parameters file:** This interface was designed for parameters/searching strategies defined previously. These search strategies have to be set in the parameters file. There are two main predefined strategies, *croizato*, and *croizat1*.

The framework of *croizato* is:

1. Find similar individual tracks [= [**u**]] from the first, to the last individual track

° CL typography: commands will be *croizato*, to indicate that the instruction is named Croizato and you must type *croizato* in your parameter file, the command could be written using lower or upper case.

2. Calculate the congruent segments among individual tracks (delimitate generalized tracks) [= **[a]**]
3. Find similar generalized tracks [= **[u]**] from the first, to the last generalized tracks

Therefore, our analysis made with the TUI is a *croizato* analysis.

The framework of *croizat1* is:

1. Calculate the congruent segments among individual tracks (delimitate generalized tracks) by **[a]**
2. Find similar generalized tracks by **[u]** from the first, to the last generalized tracks

Thus, the commands for the analysis of testo.dat, using par1-search1.txt are:

For linux 64 bits:

```
$ ./mt05-64 test0.dat test0.kml par1-search1.txt
```

For Windows 64 bits:

```
$ mt05-64.exe test0.dat test0.kml par1-search1.txt
```

For Windows 32 bits:

```
$ mt05-32.exe test0.dat test0.kml par1-search1.txt
```

*Results of testo.dat analysis*

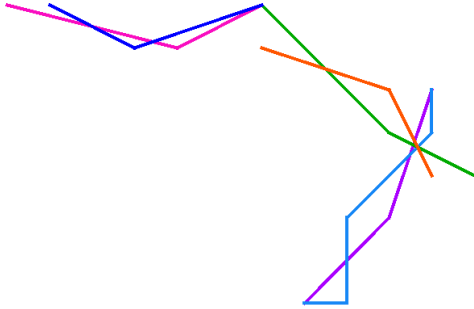


Figure 1: Individual tracks from `testo.dat`

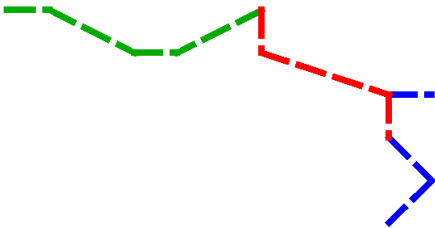


Figure 2: Generalized tracks from `testo.dat`



## *MartiTracks commands*

### *Text interface mode*

general

**[l]** == list List all tracks

comparing \*tracks\*

**[a]** == all Track all the species using MSTs to compare

**[g]** == group Track a group of contiguous species

**[p]** == pair Track a pair of species (contiguous or not)

In all the three cases a new MST=track will be created if the species compared are congruent, while in a **[p]**air comparison only a single track could be created. In an **[a]**ll/**[g]**roup comparison, there could be more than one track.

## comparing \*segments\*

<b>[s]</b>	== all	Track all the species using segments to compare
<b>[t]</b>	== pair	Track a pair of species (contiguous or not)

## join

<b>[u]</b>	== join group	Join a group of contiguous tracks if a given pair has a minimal congruence equal or higher than the minimal congruence inputted in the value file or the default value used in the TUI. If you want to join a group regardless of the minimal congruence value, reduce the min. cong. value to 0.0. After this late analysis, replace the original min. cong. value.
<b>[j]</b>	== join pair	Join two, contiguous or not, tracks independent of the minimal congruence value.

## congruence calculation

<b>[o]</b>	== Congr. group	Calculate the congruence by pairs of a contiguous group, as the output is a list, is easier to read a small group than a large group.
<b>[n]</b>	== Congr. pair	Calculate the congruence of a single pair.

## modify decision rules

<b>[c]</b>	== cut	Change the minimal cut value.
<b>[r]</b>	== rules	Modify rules.
<b>[m]</b>	== min. cong.	Change the minimal congruence value.

### *Bash command mode*

A bash command is an instruction that will perform an analysis equal to the analysis made in the text user interface [see page 33]. This mode is more appropriated to medium to huge data sets, when the analysis could last more than a few minutes or to test different parameter values.

The commands are given in the equivalence of the TUI, with a short explanation.

```
kmlgen      == write to the kml file ONLY the generalized tracks
              (default). Must be in the file before the analysis.
kmlall      == write to the kml file all, individual and generalized
              tracks. Must be in the file before the analysis.
croizat0    == join (individual tracks), track (individual tracks), and
              join (generalized tracks).
              = [u] (1 - # individual tracks), [a], [u] ( individual tracks
              + 1) total tracks
croizat1    == track (individual tracks) and join (generalized tracks)
              = [a], [u] (individual tracks +1) total tracks
```

A sample command file is

```
set cv 1.5
set lmax 2.0
set lmin 0.6
set maxline 3.0
set ci 0.80
kmlall
croizat0
```

So, these seven lines in a file will change the default parameters to

```
cv = 1.5  
lmax = 2.0  
lmin = 0.6  
maxline = 3.0  
ci = 0.80
```

and will use these values to join the individual tracks if the congruence value of the two MSTs compared is larger than 0.80, then *MartiTracks* will track those resulting MST to obtain the generalized tracks and will eliminate (joint) those redundant tracks using the same rule used before ( $ci = 0.80$ ), and will output the individual and the generalized tracks.

# Bibliography

Cavalcanti, M. (2009) *Croizat: A Software Package for Quantitative Analysis in Panbiogeography* Privately published, Rio de Janeiro., <http://www.panbiog.infobio.net/croizat>.

Contreras-Medina, R., Luna Vega, I., & Morrone, J. (2007) Application of parsimony analysis of endemism to Mexican gymnosperm distributions: grid-cells, biogeographical provinces and track analysis. *Biological Journal of the Linnean Society*, 92(3), 405–417.

Craw, R. (1983) *Biogeography of New Zealand: a methodological and conceptual approach* PhD thesis, Wellington.

Craw, R. C. (1989a) New Zealand biogeography: a panbiogeographic approach. *New Zealand Journal of Zoology*, 16(4), 527–547.

Craw, R. (1988a) Continuing the synthesis between panbiogeography, phylogenetic systematics and geology as illustrated by empirical studies on the biogeography of New Zealand and the Chatham Islands. *Systematic Zoology*, 37(3), 291–310.

Craw, R. C. (1987) Panbiogeography and Structuralist Biology. *Rivista di Biologia - Biology Forum*, 80, 226–229.

Craw, R. C. (1988b) Continuing the Synthesis Between Panbiogeography, Phylogenetic Systematics and Geology as Illustrated by Empirical Studies on the Biogeography of New Zealand and the Chatham Islands. *Systematic Zoology*, 37(3), 291–310.

Craw, R. C. (1989b) Quantitative panbiogeography: Introduction to methodsb. *New Zealand Journal of Zoology*, 16(4), 485–494.

Craw, R. C., Grehan, J. R., & Heads, M. J. (1999) *Panbiogeography: Tracking the History of Life*. Oxford University Press.

Crisci, J. V. & Morrone, J. J. (1992) Panbiogeografía y biogeografía cladística: paradigmas actuales de la biogeografía histórica. *Ciencias*, 6, 87–97.

Crisci, J. (2001) The voice of historical biogeography. *Journal of Biogeography*, 28(2), 157–168.

Croizat, L. (1952) *Manual of Phytogeography*. The Hague: Junk.

Croizat, L. (1958) *Panbiogeography*. Caracas: Published by the author.

Croizat, L. (1964) *Space, Time, Form: The Biological Synthesis*. Caracas: Published by the author.

Espinos-Organista, D., Morrone, J. J., Llorente-Bousquets, J., & Villela-Flores, O. (2002) *Patrones de homología espacial*, chapter 4, (pp. 53–64). Las Prensas de Ciencias, Universidad Nacional Autónoma de México.

Espinosa-Perez, M., E., H. M., & J., M. J. (2009) Identification of generalized tracks for the species of Isopoda (Peracarida) from the Eastern Pacific. *Journal of Crustacean Biology*, 29, 224–231.

Franco-Rosselli, P. (2001) *Estudios panbiogeográficos en Colombia*, (pp. 221–224). Las Prensas de Ciencias, Universidad Nacional Autónoma de México: México.

García-Barros, E., Gurrea, P., Lucianez, M. J., Cano, J. M., Munguira, M. L., Moreno, J. C., Sainz, H., Sanz, M. J., & Simon, J. C. (2002) Parsimony analysis of endemism and its application to animal and plant geographical distributions in the Ibero-Balearic region (western Mediterranean). *Journal of Biogeography*, 29(1), 109–124.

Garzón-Orduña, I. J., Miranda-Esquivel, D. R., & Donato, M. (2008) Parsimony analysis of endemism describes but does not explain: an illustrated critique. *Journal of Biogeography*, 35(5), 903–913.

Grehan, J. R. (1993) Conservation biogeography and the biodiversity crisis: a global problem in space/time. *Biodiversity Letters*, 1, 134–140.

Grehan, J. R. (1994) The Beginning and End of Dispersal: The Representation of Panbiogeography. *Journal of Biogeography*, 21(5), 451–462.

Grehan, J. R. (2001a) *Introducción a la Biogeografía en Latinoamérica: Teorías, Conceptos, Métodos y Aplicaciones*, chapter Panbiogeografía y la geografía de la vida, (pp. 181–195). Las Prensas de Ciencias, Universidad Nacional Autónoma de México: México.

Grehan, J. R. (2001b) *Islas Galapagos: biogeografía, tectónica y evolución en un archipiélago oceánico*, (pp. 153–160). Las Prensas de Ciencias, Universidad Nacional Autónoma de México: México.

Grehan, J. (1988) Panbiogeography: evolution in space and time. *Rivista di Biologia - Biology Forum*, 81(4), 469–485.

Henderson, I. (1989) Quantitative panbiogeography: an investigation into concepts and methods. *New Zealand Journal of Zoology*, 16(4), 495–510.

Huidobro, L., Morrone, J. J., Villalobos, J. L., & Alvarez, F. (2006) Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican Transition Zone. *Journal of Biogeography*, 33(4), 731–741.

Humphries, C. J. (2000) Form, Space and Time; Which Comes First?. *Journal of Biogeography*, 27(1), 11–15.

Liria, J. (2008) Geographic information systems and spatial analysis: a combined method for panbiogeographic studies. *Revista Mexicana de Biodiversidad*, 79(1), 281–284.

Luna-Vega, I., Ayala, O. A., Morrone, J. J., & Espinosa-Organista, D. (2000) Track analysis and conservation priorities in the cloud forests of Hidalgo, Mexico. *Diversity and Distributions*, 6(3), 137–143.

Meacham, C. A. (1984) The role of hypothesized direction of characters. *Taxon*, 33, 26–38.

Michaux, B. (1989) Generalized Tracks and Geology. *Systematic Zoology*, 38(4), 390–398.

Mihoc, M. A. K., Morrone, J. J., Negritto, M. A., & Cavieres, L. A. (2006) Evolution of the series Microphyllae (Adesmia, Fabaceae) in the Andean Cordillera: a biogeographic approach. *Revista Chilena de Historia Natural*, 79(3), 389–404.

Morrone, J. J. (2000) A new regional biogeography of the Amazonian subregion, mainly based on animal taxa. *Anales del Instituto de Biología Universidad Nacional Autónoma de México, Serie Zoología*, 71(2), 99–123.

Morrone, J. J. (2001) Homology, biogeography and areas of endemism. *Diversity and Distributions*, 7, 297–300.

Morrone, J. J. & Crisci, J. V. (1995) Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics*, 26, 373–401.

Morrone, J. J. & Marquez, J. (2001) Halffter's Mexican Transition Zone, beetle generalized tracks, and geographical homology. *Journal of Biogeography*, 28(5), 635–650.

Morrone, J. & Espinosa, D. (1998) La relevancia de los atlas biogeográficos para la conservación de la biodiversidad mexicana. *Ciencia*, 49(3), 12–16.



Morrone, J. J. (2004) *Homología Biogeográfica: Las Coordenadas Espaciales de la Vida*, volume 37 of *Cuadernos del Instituto de Biología*. Mexico: Instituto de Biología, Universidad Nacional Autónoma de México.

Page, R. D. M. (1987) Graphs and generalized tracks: quantifying Croizat's panbiogeography. *Systematic Zoology*, 36(1), 1–17.

Rosen, B. R. (1984) *Fossils and climate*, chapter Reef coral biogeography and climate through the late Cainozoic: just islands in the sum or a critical pattern of islands?, (pp. 201–262). Wiley, Chichester, UK.



# *Index*

- cli:command line interface, [30](#)
- command modes, [25](#)
- commands, [33](#)
- commands:bash, [35](#)
- commands:text interface, [33](#)
- Craw's tracks compatibility analysis, [18](#)
- Croizat:books, [15](#)
- download the program, [24](#)
- license, [4](#)
- PAE, [17](#)
- PAE:algorithm, [19](#)
- Page's Spanning graphs, [18](#)
- Panbiogeography:assumptions, [16](#)
- parameters:modify, [29](#)
- parameters:values, [26](#)
- testo.dat: command line run, [31](#)
- Track analysis, [16](#)
- Track analysis:generalized tracks, [16](#)
- Track analysis:individual tracks, [16](#)
- Track analysis:manual reconstruction, [17](#)
- Track analysis:nodes, [17](#)
- tui:text user interface, [25](#)
- valid input data set, [24](#)