

# PaleoTax/Graph

Database management system to record, process, analyse,  
and output taxonomic, geographic and stratigraphic data in  
palaeontology

Data visualisation tool

**Version 2.0 • Manual**

Version 2.0.2

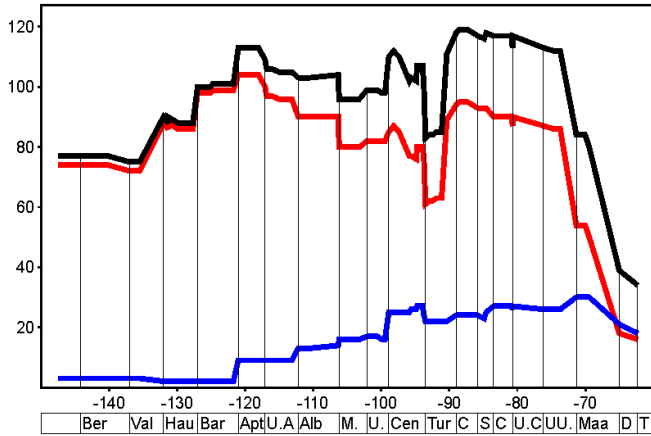
# Contents

<b>1. Introduction.....</b>	<b>3</b>
1.1. <i>What is PaleoTax/Graph, what can it do?</i> .....	3
1.2. <i>Conditions of use</i> .....	4
1.3. <i>The program</i> .....	4
1.4. <i>Hardware requirements</i> .....	5
1.5. <i>The manual</i> .....	5
<b>2. Installation.....</b>	<b>5</b>
<b>3. Start PaleoTax/Graph.....</b>	<b>5</b>
<b>4. Applications.....</b>	<b>7</b>
4.1. <i>Correlation</i> .....	7
4.2. <i>Dendrogram</i> .....	9
4.3. <i>Chart</i> .....	10
Time raster .....	10
File format .....	11
PaleoTax output.....	11
Range and abundance .....	12
Only the ranges.....	14

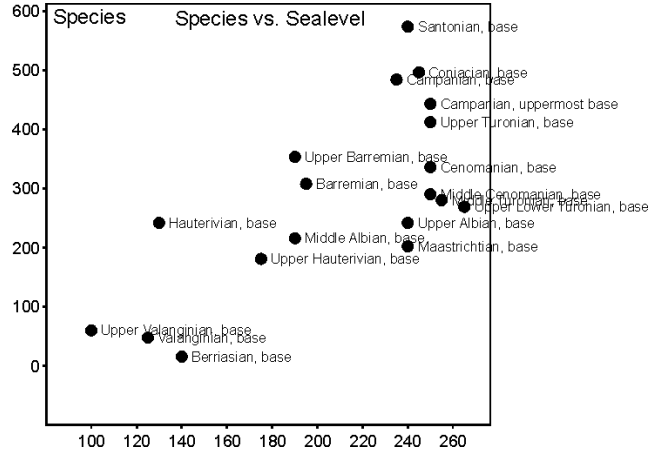
# 1. Introduction

## 1.1. What is PaleoTax/Graph, what can it do?

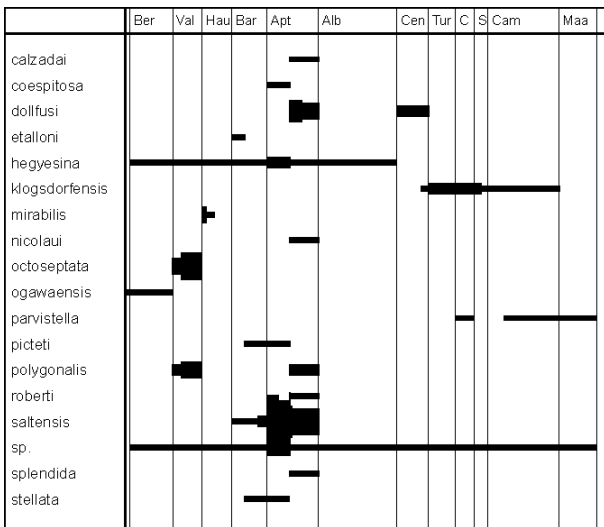
PaleoTax/Graph is an additional tool to visualize the result of analysis or estimations carried out with PaleoTax. PaleoTax/Graph offers a wide range of tools, including the following,



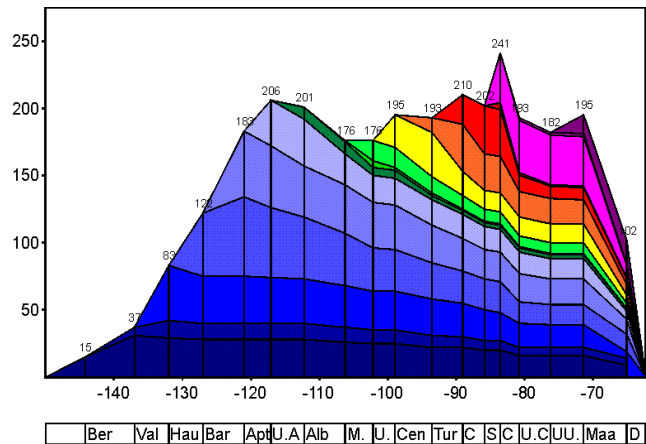
line diagrams



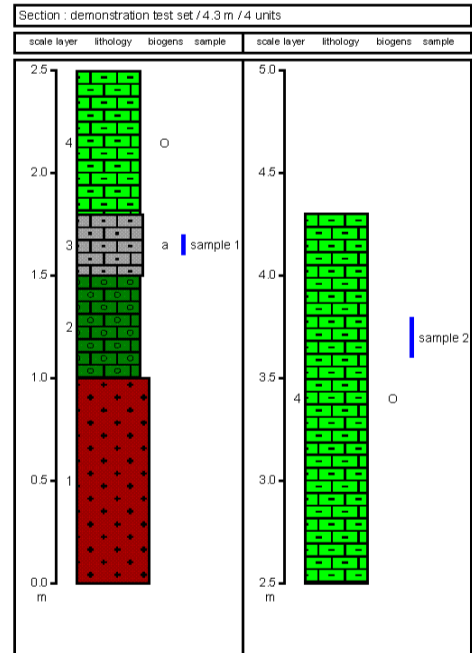
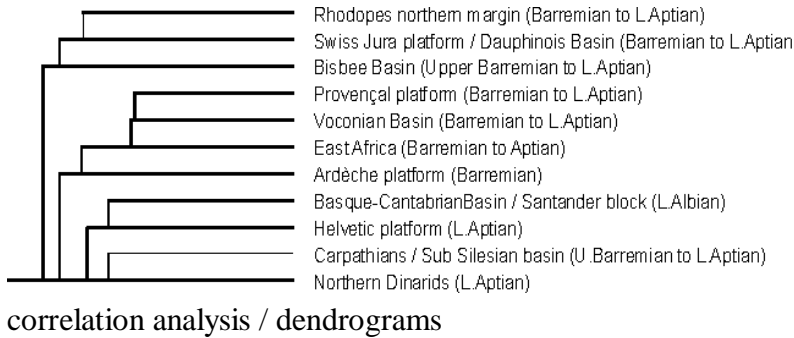
xy-diagrams



range charts (generally for taxa)



polygons (origination/extinction diagrams)



Graphs are exported to files of the Windows Meta File format (WMF). This format can easily be imported by many applications and modified by vector programs (such as CorelDraw).

PaleoTax/Graph is only a by-product and the Windows version of a MS DOS program which has never been released to public. The reason to release it now, is to have a simple tool for visualizing the results of data analysis in PaleoTax. Analysing data became with the increasing amount of data more and more interesting. PaleoTax and PaleoTax/Graph cannot replace sophisticated Paleobiology style data analysis, but it can help to give a rough idea about the data. The used methods are simple but robust.

As for PaleoTax valid, this tool can only grow in its functionality when critical evaluated. Suggestions are therefore very welcome. Therefore check regularly for updates.

## 1.2. Conditions of use

PaleoTax/Graph is released as freeware. You may use the software without paying any fee, but you bear the risks involved: the author will not be responsible for their correctness or for agreement of the results of analysis with your expectations. Although it cannot be ruled out that updated versions will be sold or service contracts concluded. You will always have the right to work with the free version. At present you cannot advance any claims for support, for the elimination of faults, the improvement of the program, or training, but the author will endeavour to improve the program, to speedily remove faults and to advise its users.

## 1.3. The program

PaleoTax/Graph does not form part of the PaleoTax program (as PaleoTax/Map) and is delivered as an independent program which will be installed in the Hdb2Win program folder. PaleoTax/Graph expect input files with the extension PGR in the ASCII format. These files are normally produced by PaleoTax, but can be also produced by any other program. The first lines of the files contain information about the data and the way how to process by PaleoTax/Graph.

## 1.4. Hardware requirements

The following are the minimum hardware requirements for PaleoTax/Graph:

Parameter	
Processor speed	500 MHz
Main memory	128 MB
Free hard disk space (for the program and data)	1 MB

Operation system: Windows 95, 98, ME, 2000, XP. Not tested under Windows Vista.

## 1.5. The manual

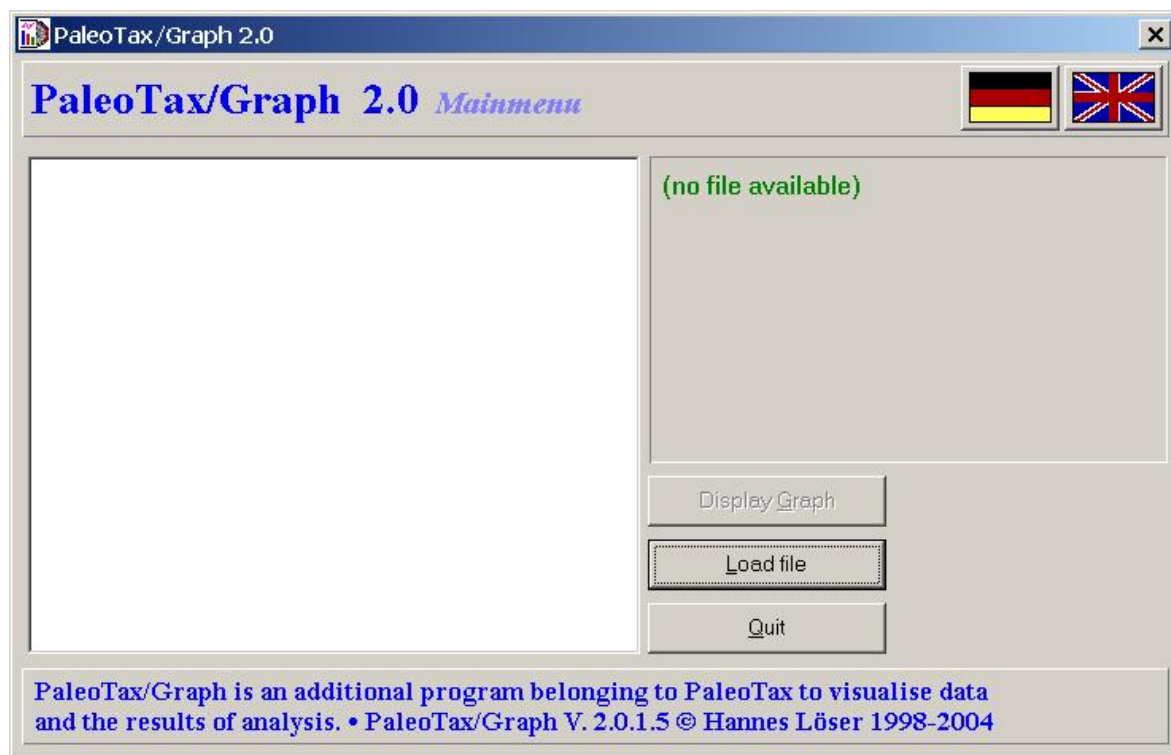
This manual is only a preliminary version since it does encompass only the description of a small part of the functions of PaleoTax/Graph. Check regularly not only for program updates but also for updates of the manual. New versions of the program and the documentation are released with new versions of Hdb2Win.

## 2. Installation

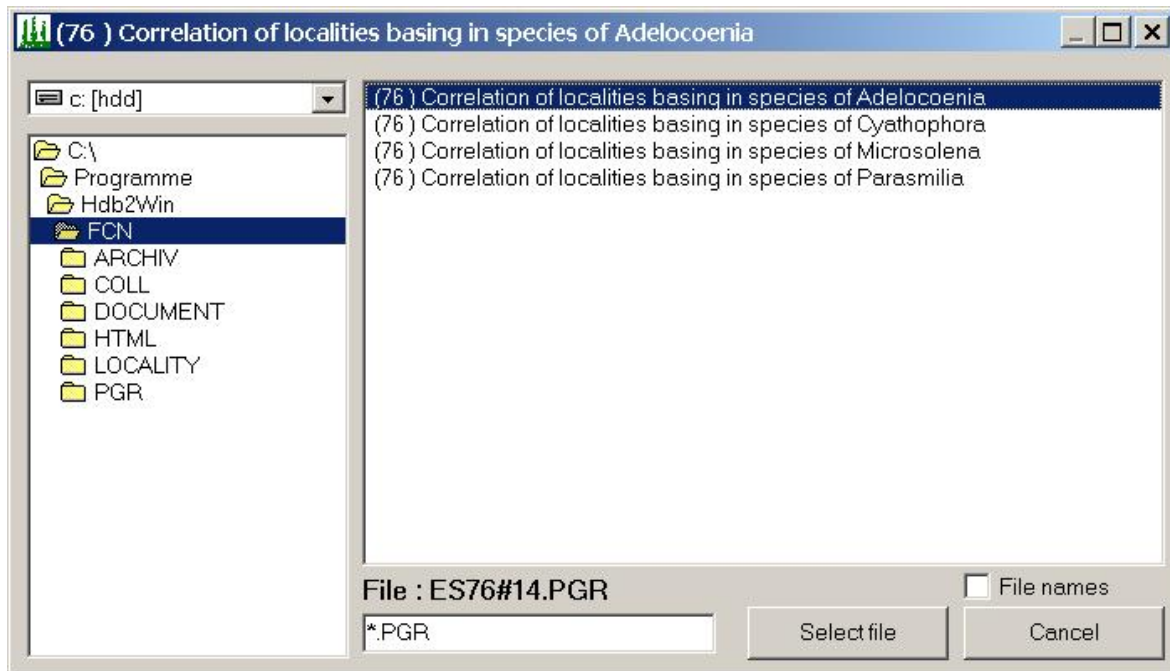
PaleoTax/Graph 2.0 will be installed together with PaleoTax. Both are available as a download from the PaleoTax web page. Copy the downloaded ZIP archive into a temporary directory on your hard disc and unpack the ZIP archive. It contains several archives and an installation program INSTALL.EXE. Start it. Click on the British banner to choose the English language. Follow the instruction.

## 3. Start PaleoTax/Graph

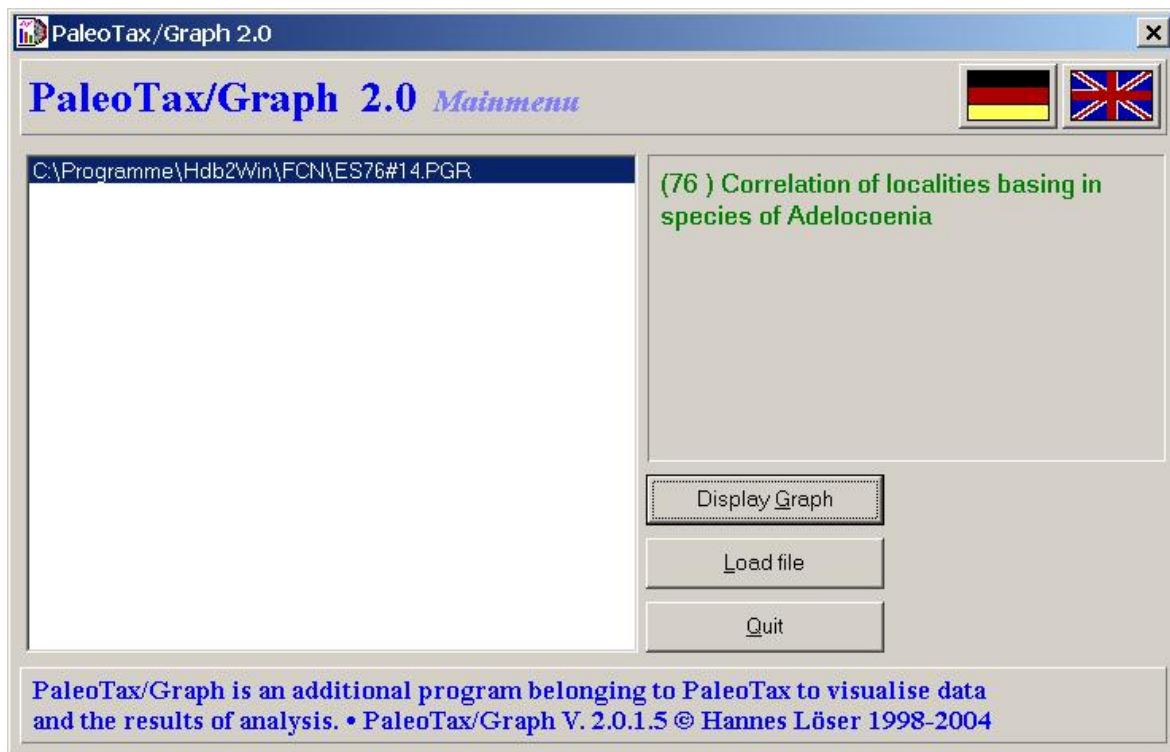
The program is located in the HDB2WIN directory and its name is PGRAPH2.EXE. It may run in English and German. After starting, the following menu will appear:



You may switch the language clicking on one of the banners. When you click on Load File, you may select a file :



After selection a file, the corresponding application of PaleoTax/Graph will be started. After terminating, the previously selected file will be kept in the file window of PaleoTax/Graph:



When terminating the program and starting again, the list will be reload (as far as the corresponding files are not deleted). Display Graph would now display the graph again, Load file opens again the load file dialog.

## 4. Applications

PaleoTax/Graph includes various applications as mentioned above. You cannot select the applications yourself since the program start the application according to the information in the PGR files. The applications have not got a traditional menu, normally the windows is divided into a (left) option and a (right) preview part. Change the options and you see immediately the changes in the right portion of the window.

### 4.1. Correlation

Correlation means to evaluate the similarity or dissimilarity of objects on the basis of their characteristics. For instance, localities of fossils can be more or less similar under the aspect of the fossils which are found there. Normally, localities in one and the same facies, which are geographically or stratigraphically closely related, may have also a similar fauna and will show a high correlation. In this example, the localities are the examined individuals (n), and the species are the characteristics (m).

The correlation bases on 1/0 values (species m exists on locality n or not). If two localities are compared, there is a certain number of joint species ("C"), and a certain number of species which do not occur at no one of the two localities ("A"). The easiest way ("simple correlation") is to look for a positive correlation, say, only to value the number of joint species. This may give a coarse idea, but it ignores that the absence of certain species may also deliver important information, and it ignores when small to large faunas are compared.

Therefore a high number of various Correlation Coefficients (a good overview give Cheetham & Hazel 1969) exists. PaleoTax/Graph offers only a small number. If your favourite correlation coefficient lacks, please do not hesitate to contact me. It is easy to incorporate it.

The following abbreviations are used:

C	number of species which occur in both localities
A	number of species which does not occur in no one of the both locality
N1	number of species present in the first locality
N2	number of species present in the second locality
Nt	number of involved species at all (N1+N2-C)
E1	N1-C
E2	N2-C

Of course, instead of localities, genera may be used and instead of species, regions. Any correlation can be applied.

The module expect a file (extension PGR) with the following data.

First line of the file :

;C: [Title in German]

;C: [Title in English]

;D: 4,[configuration file],[list file],0

e.g.

;C: Korrelation der Gattung X

;C: Correlation of genus X

;D: 4,AW29.CFG,AW29#11.LST,0

In the description line (;D:) the first number stands for the routine. In the configuration file options for the dendrogram (see below) are saved (and reloaded automatically the next time). The list file contains the

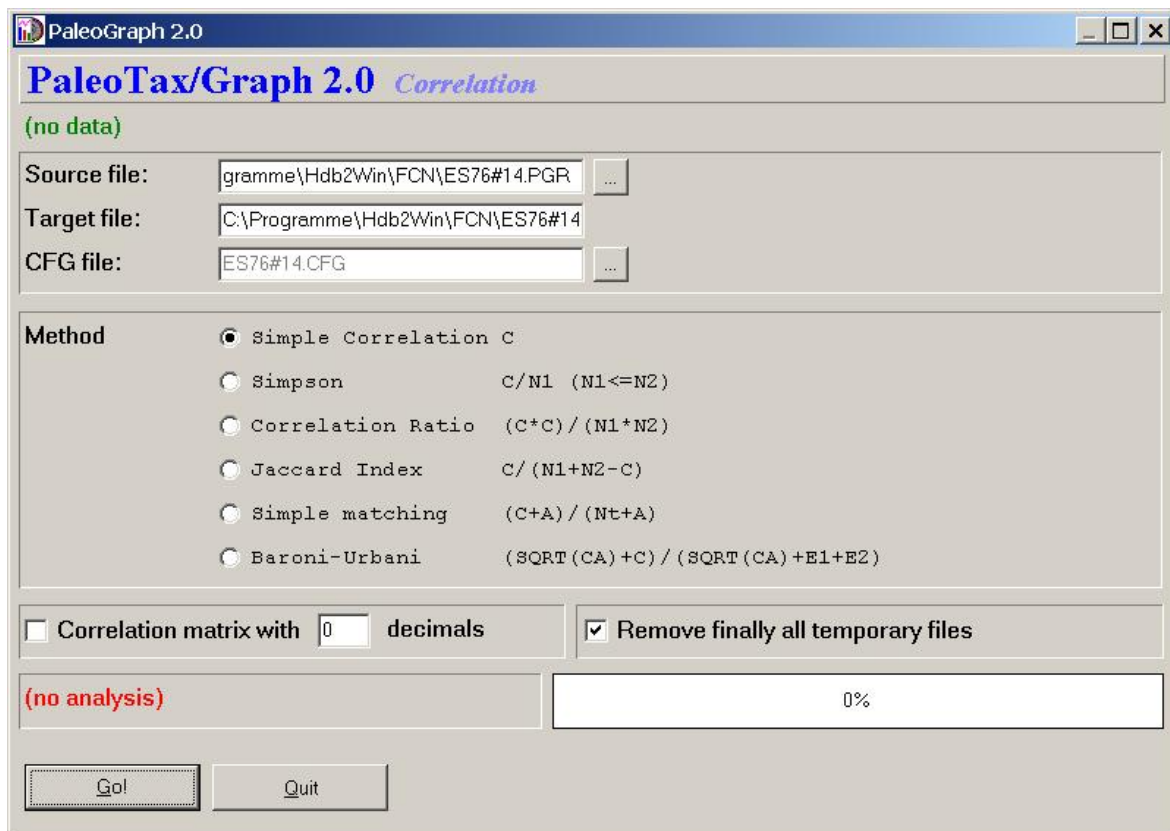
names of the individuals (in the example above, the localities). The last number is for further extensions and should be always zero.

After the three heading description lines follow the data:

```
1 1155
2 1155
```

The first digit (five characters long) refers to the individuals in the list file; the item with the number 1 refers to the item in the first line of the list file. The second digit (5 characters long) is the characteristic (here, the number of the species). Bother digits are not separated ! The format of the line is 1111122222 where 11111 stands for the first digit and 22222 for the second digit.

The maximal number of pairs is restricted to 15,000, the number of individuals to 1,000. The value of individuals and characteristics at all may exceed this value up to 10,000, say, they may not be current. But the maximum of items which can be correlated, is limited to 1,000.



The processing is easy: select one of the methods, and click on Go!. The program offers also the output of a correlation matrix. You have to tick the box and select the number of decimals.

The program does not only the correlation, it calculates also a cluster diagram using the agglomerative single linkage method (Shi 1993). If successfully processed, the program outputs a ASCII file with the extension CLT which contains the description of the cluster diagram with the following format:

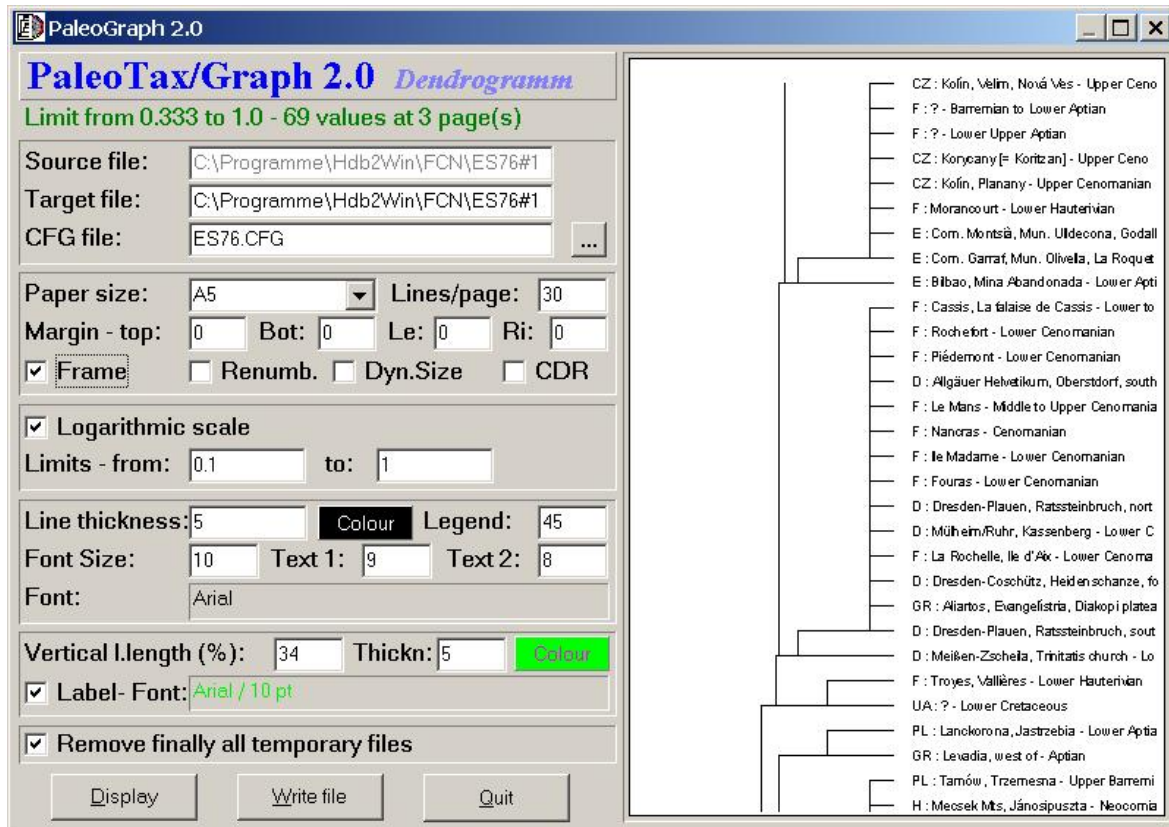
```
1 13 1.000000 1
```

The first number (5 digits) indicate the first individual, the second number (5 digits) the second individual, which form a group (cluster) at the value indicated as third number (12 digits) resulting in the cluster indicated as fourth number (5 digits). The individuals are successively added to the diagram until it remains only one item.

The correlation module calls automatically the dendrogram module which opens the CLT file and calculates the dendrogram. At the present moment you have no influence on the methods used for building the tree.

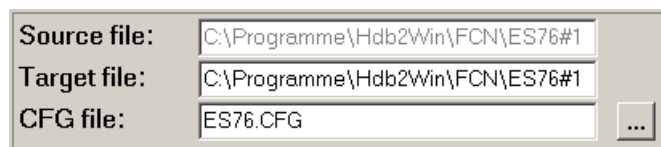
## 4.2. Dendrogram

The dendrogram expects a file with the extension CLT which format is explained above, and the list file. The program interprets the CLT file and constructs a cluster diagram (dendrogram). The diagram can have a maximum of 4096 individuals and can be distributed on 255 pages.

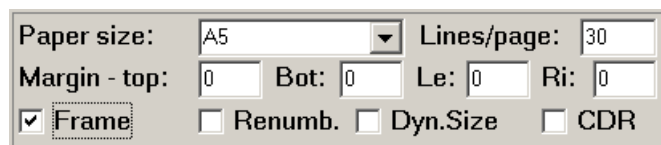


The window is divided into a left hand sided option part and a right hand sided preview part. When first called, the program reads the input file and displays the dendrogram using default values.

The left hand option panel is subdivided in six sub panels. Above the first panel it shows the number of values, their range, and the number of pages.



The first panel contains the source file (which normally cannot be modified since it comes from the dendrogram module), the target file and the configuration file.



The second panel defines the paper size, the lines per page, the margins (top, bottom, left, right). The first checkbox in the bottom line defines whether a frame should be placed around the graph, the second checkbox changes the labels and number the items currently, the third box adapt the frame size to the number of items (if there are less items on the page than in "Lines/page" indicated). The last checkbox should be marked if the file will be imported later by CorelDraw, because this program handles some objects from a WMF file other than e.g. WinWord or other Microsoft programs.

<input checked="" type="checkbox"/> <b>Logarithmic scale</b>
Limits - from: <input type="text" value="0.1"/> to: <input type="text" value="1"/>

Normally the scale of the graph is linear, but by ticking this checkbox, it can be changed to logarithmic. Sometimes it results in a clearer graph. The limits are defined automatically, but can be set manually as well.

Line thickness: <input type="text" value="5"/>	<b>Colour</b>	Legend: <input type="text" value="45"/>
Font Size: <input type="text" value="10"/>	Text 1: <input type="text" value="9"/>	Text 2: <input type="text" value="8"/>
Font: <input type="text" value="Arial"/>		

The fourth panel is dedicated to the lines and labels of the graph. The "Line thickness" is given in 0.1mm units. The colour can be selected in a dialog. The number behind "Legend" gives the percentage of the text on the right hand side of the graph. A low value gives more space to the dendrogram. The "Font size" applies when the text is undivided. When the text is separated by a tabulator into two parts, the font size in "Text 1" and

When clicking on Display or hitting SPACE, the program displays the first page. Via PgDn and PgUp, other pages can be displayed. Clicking on Write file force the program to write the graph in WMF files. For each page, a separate (numbered) file will be created. Quit terminates the application.

### 4.3. Chart

The Chart modul displays the ranges of taxa in time. It offers a wide range of functions of which only a small part is used by PaleoTax outputs. The chart has the form of a table: on the left hand side is a list of taxa, on the right hand side is a field for graphs which indicates an extend in time. Everything can be adapted: the paper size, the margins, the portion of the legend, the thickness of lines etc.

#### Time raster

An essential element of the chart program is a time frame, which is supplied in a separate ASCII file with the extension PSC. The standard PSC file (STANDARD.PSC) looks like this:

```
Pli   -5.1
Mio   -22.5
Oli   -27
Eoc   -54.9
Tha   -62.3
Dan   -65
Maa   -71.3
Cam   -83.5
S     -85.79
C     -89
Tur   -93.5
...
```

"Text 2" are used, respectively. By clicking in the field behind "Font: ", the font can be selected.

Vertical l.length (%): <input type="text" value="34"/>	Thickn: <input type="text" value="5"/>	<b>Colour</b>
<input checked="" type="checkbox"/> Label- Font: <input type="text" value="Arial / 10 pt"/>		

For orientation, vertical lines can be drawn. The "Vertical line length" gives the percentage from the whole page length, measured from above. The thickness is as well in units of 0.1 mm. The colour may be select in a dialog. The checkbox "Label" should be ticked when labels for the lines are desired; the font can be selected by clicking in the field on the right.

<input checked="" type="checkbox"/> <b>Remove finally all temporary files</b>
---

The program creates a number of temporary files which normally can be removed.

It follows the format <name><tab><-ma>, where the name of the stage or period is followed by a tabulator and the numeric value. The value always corresponds to the base (lower limit) of the unit. So in the above example, the base of the Danian (Dan) is -65 ma. Stages and periods are abbreviated; later you will see, why. You may modify, add or delete values, you can even create your own PSC (**P**aleo**T**ax/**G**raph **S**cale) time raster file and load it when you are working with the program. The standard scale will be installed with the program. It is very important that you compare your own time frame (as recorded in the table AGEIUGS) with the PSC file. It is very probable that you need to adapt the PSC file to the time frame you use. Not doing so may looking the charts strange.

### File format

The PGR file has as a header and data part. The contains contains the description of the file:

```
;C: Chart Verbreitung der Gattungen auf der Basis der Arten
;C: Chart Stratigraphical distribution of the genera based on species
;D: 5,GENER2.CFG,STANDARD.PSC,0,0,0
```

The first lines described the file in German and English, below follows the instruction line for the program: The first value describes the mode (1 for standard chart, 5 for abundance chart), the two file names the configuration and raster file (see below), and the three values at the end are only for internal use.

Afterwards follow the data part:

```
Aulastraeoporidae   Aulastraeopora   -117.07   -112.2
Aulastraeoporidae   Aulastraeopora   -121   -112.2
Aulastraeoporidae   Aulastraeopora   -121   -110.6
Aulastraeoporidae   Aulastraeopora   -117.07   -112.2
Aulastraeoporidae   Aulastraeopora   -117.07   -112.2
Aulastraeoporidae   Aulastraeopora   -98.9 -93.5
Aulastraeoporidae   Aulastraeopora   -112.2   -106.18
Aulastraeoporidae   Blothrocyathus   -127   -121
Aulastraeoporidae   Budaia          -121   -117.07
Aulastraeoporidae   Budaia          -121   -112.2
Aulastraeoporidae   Budaia          -115.11   -97.39
...
```

The data line has the following format:

```
text1[<tab>text2]<tab><value from><tab><value to>[<tab>color[<tab>bar thickness]]
```

Text1 and text2 can be genus and species or family and genus. The second part may be absent. Both value (from - to) consists of two digits (always negative values) separated by a <tab> or comma, giving the range. Bar colour and bar thickness are optional (not applied by PaleoTax output). The values can be separated by tabulators or commas.

### PaleoTax output

Actually exist with SR9 only four profiles which export PGR files showing charts.

### *Species*

Stratigraphy from indications (Chart) [Stratigraphie aus Indikationen (Chart)] : Calculates the **range and abundance** for each species from the indications, say from the localities which are assigned to the citations in the literature. - The output is automatically sorted according to current genus and species name.

Stratigraphical distribution (Chart) [Stratigraphische Verbreitung (Chart)] : Displays **only the ranges** of each species as recorded in the species table. The user may decide whether the standard range (as calculated for each species based on the indications) or critical range is used. - The output is not automatically sorted and it makes sense, the sort the table according to your need before starting the output.

The difference between both is, that the first profile gives more information because it values the abundance (chart bars have a differing thickness), whereas the second profile just shows the range and not more.

### Genera

Stratigraphy based on species range (Chart) [Stratigraphie auf Arten basierend (Chart)] : Equivalent to above, the **range and abundance** for each genera is calculated on the base of the species. Also here, the user may select: the standard range (as calculated for each species based on the indications) or critical species range is used. Note that both may differ depending on data recorded. - The output is automatically sorted according to current family and genus name.

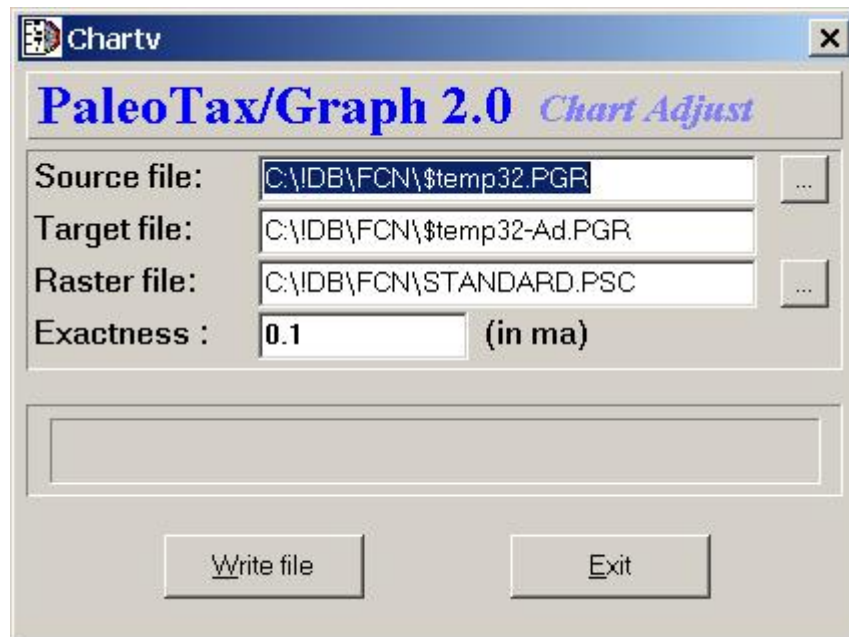
Stratigraphical distribution (Chart) [Stratigraphische Verbreitung (Chart)] : Displays **only the ranges** of genera as recorded in the genera table. The user may decide whether the standard range (as calculated for each genus based on the indications) or critical range is used. - The output is not automatically sorted and it makes sense, the sort the table according to your need before starting the output.

The difference between both is, that the first profile gives more information because it values the abundance (chart bars have a differing thickness), whereas the second profile just shows the range and not more.

Depending on both methods (with and without abundance data) different procedures are applied.

### Range and abundance

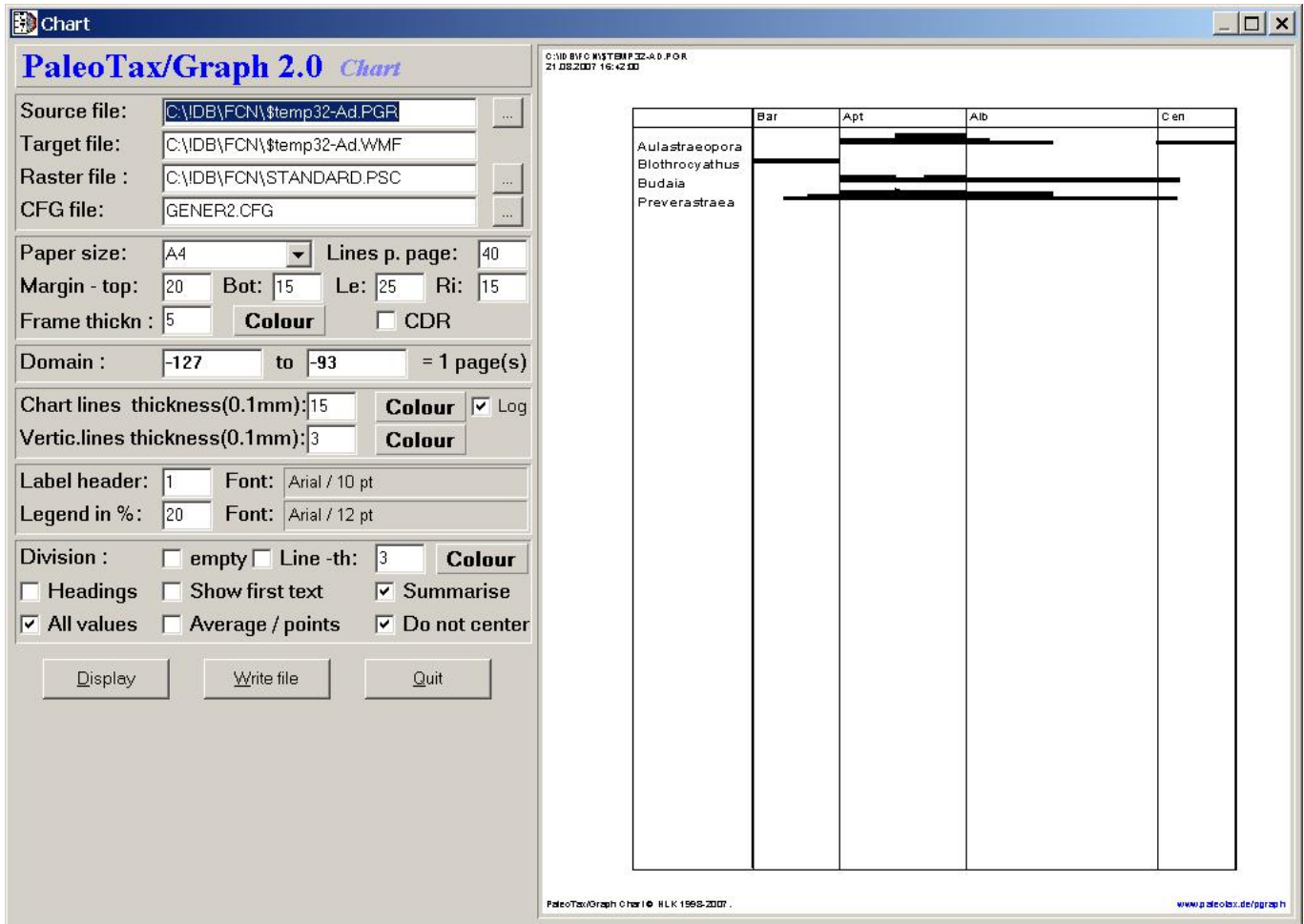
To calculate the abundance, a small program (Chart Adjust) is started which looks like this:



The source file and target file are automatically set by the program. You may change the raster file to get a coarser or finer raster for the display. The value Exactness should be as fine as the values are in your scale. A small value gives finer results but produces much data, a higher value produces coarser output but less data. Click on Write file, and the target file will be created and the chart program will be opened:



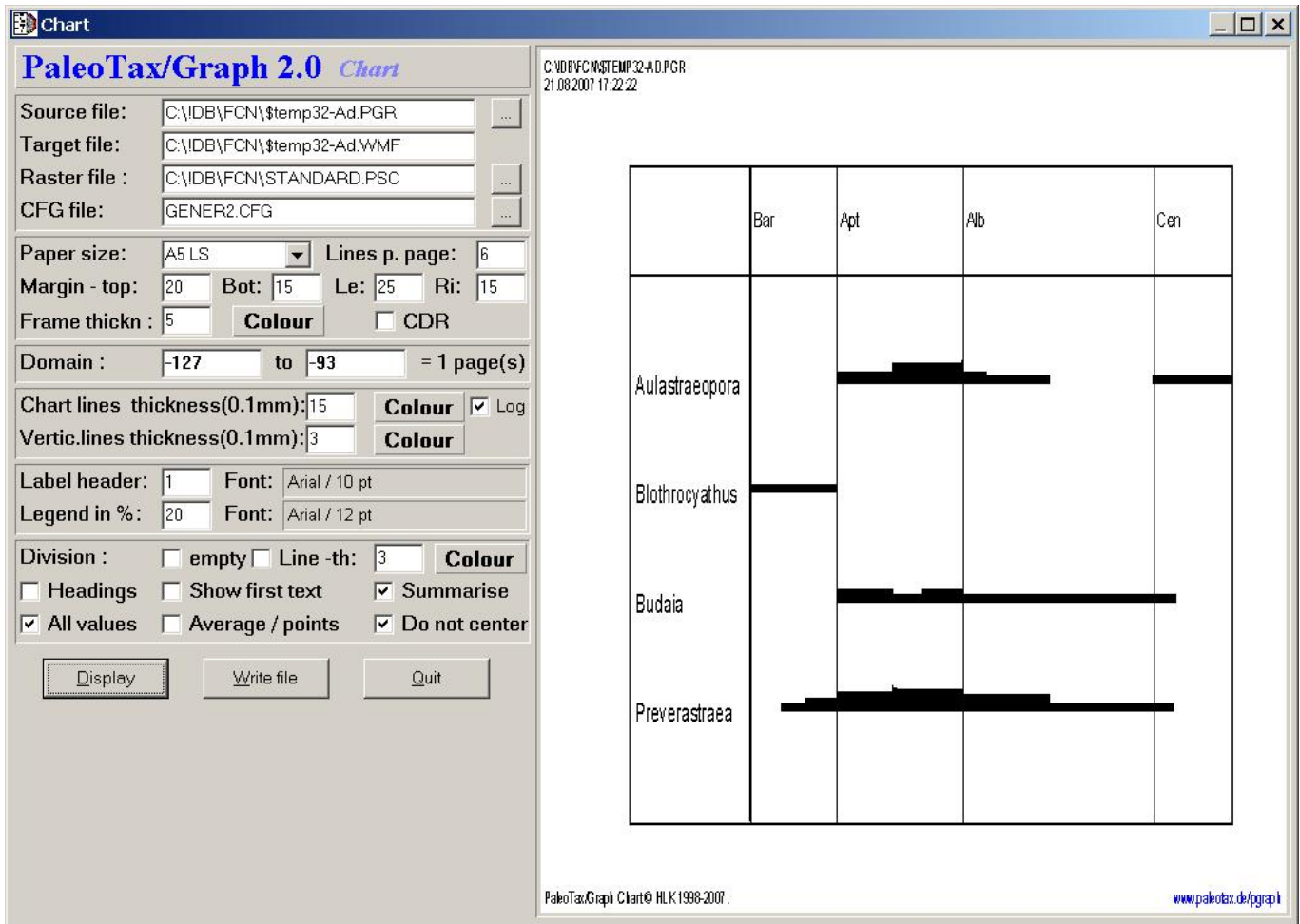
This looks very strange - it is because values are not yet summarised ! Tag the box Summarize below right:



Because the raster file was quite coarse for this chart, you can see some artefacts in the limit between Early and Late Aptian and in the limit between Aptian and Albian. Values in the raster file and the exactness should therefore correspond. It is also recommended, to make the raster file a little bit coarser as it would be the truth: what is the difference (the base of the Santonian) of -85.79 and -85.8 ma ? What you can also see in the chart: the width of the columns correspond directly to the duration of the time span. Please compare directly below for details of the chart program.

### Only the ranges

The chart offers a wide range of options and the best way to use it, is to just try it. The result is immediately shown on the right hand side of the program window.



The left hand option panel is subdivided in six sub panels. Above the first panel it shows the number of values, their range, and the number of pages.



The first panel contains the source file, the target file, the raster file and the configuration file.



The second panel defines the paper size, the chart lines per page, the margins (top, bottom, left, right). The CDR checkbox should be tagged if the file will be imported later by CorelDraw, because this program handles some objects from a WMF file other than e.g. WinWord or other Microsoft programs.

<b>Domain :</b>	<input type="text" value="-127"/>	to	<input type="text" value="-93"/>	= 1 page(s)
-----------------	-----------------------------------	----	----------------------------------	-------------

The third panel defines the time range of the chart and indicates the number of pages resulting from page size, lines per page and number of chart lines in the source file. You may modify the time range.

<b>Chart lines thickness(0.1mm):</b>	<input type="text" value="15"/>	<b>Colour</b>	<input checked="" type="checkbox"/> Log
<b>Vertic.lines thickness(0.1mm):</b>	<input type="text" value="3"/>	<b>Colour</b>	

You may change the colour and thickness of chart lines (in 0.1 mm) and vertical leading lines (in 0.1 mm). When the abundance function is used, the abundance can be linear (abundance multiplied with the base chart line thickness) or logarithmic. Tag the box Log to get the thickness logarithmically calculated.

<b>Label header:</b>	<input type="text" value="1"/>	<b>Font:</b>	<input type="text" value="Arial / 10 pt"/>
<b>Legend in %:</b>	<input type="text" value="20"/>	<b>Font:</b>	<input type="text" value="Arial / 12 pt"/>

‘Label header’ indicates the number of headlines, ‘Legend in %’ the percentage of space which will be occupied by the legend. Font, font size etc. may be modified by clicking in the boxes.

<b>Division :</b>	<input type="checkbox"/> empty	<input type="checkbox"/> Line -th:	<input type="text" value="3"/>	<b>Colour</b>
<input type="checkbox"/> Headings	<input type="checkbox"/> Show first text	<input checked="" type="checkbox"/> Summarise		
<input checked="" type="checkbox"/> All values	<input type="checkbox"/> Average / points	<input checked="" type="checkbox"/> Do not center		

This panel contains offers various options.

‘Division’ defines how groups of taxa should be separated. This does not apply if the second text is empty. But normally, the first part is the family and the second the genus, or the first part is the genus and the second one the species. Let’s take the second case: species can be grouped by genera since the output is normally alphabetically classified. The separation between genera can be by a empty line (‘empty’) or by a horizontal line (‘Line’). In the latter case, line thickness (x 0.1 mm) and color can be defined.

If first and second part are present, headings can separated the groups (‘Headings’). If first and second part are present, the first part can be suppressed in the list (this makes sense if heading is tagged).

In the case of an abundance chart, ‘Summarise’ has to be tagged.

‘All values’ means, that all text lines printed, independently whether the bar is visible or not. This might be the case, when the range was manually adapted and one or various bars become out of range.

‘Average / points’ is only valid for non abundance charts: the average range is calculated and plotted as a dot.

‘Do not center’ is only useful for abundance charts: normally the bars are centred, but varying thickness are more obvious when the have a unique baseline, e.g. the bars are not centered.

---

References	<p>Cheetham, A.H. &amp; Hazel, J.E. 1969. Binary (presence-absence) similarity coefficients. -- <i>Journal of Paleontology</i>, 43, 5: 1130-1136; Tulsa, Okla.</p> <p>Löser, H. 2001. PaleoTax - Datenbanksystem zur Erfassung, Verarbeitung und Ausgabe taxonomischer, geographischer und stratigraphischer Daten in der Paläontologie. Version 2.0. Handbuch und Referenz. -- 136 pp.; Dresden (CPress).</p> <p>Löser, H. 2004a. PaleoTax. Database management system to record, process, analyse and output taxonomic, geographic and stratigraphic data in palaeontology. Version 2.1. Manual. -- <i>Publicaciones ocasionales</i>, 3: 1-153; Hermosillo (UNAM).</p> <p>Löser, H. 2004b. PaleoTax - a database program for palaeontological data. -- <i>Computer &amp; Geosciences</i>, 30, 5: 513-521; Amsterdam (Elsevier Scientific Publishing Company).</p> <p>Shi, G.R. 1993a. Multivariate data analysis in palaeoecology and palaeobiogeography - a review. -- <i>Palaeogeography, Palaeoclimatology, Palaeoecology</i>, 105: 199-234; Amsterdam.</p>
Internet site	<a href="http://www.paleotax.de">http://www.paleotax.de</a>
Internet site for P/Graph	<a href="http://www.paleotax.de/pgraph/">http://www.paleotax.de/pgraph/</a>
E-Mail	<a href="mailto:info@paleotax.de">info@paleotax.de</a>
List server registration	<a href="https://www.link-up.de/mailman/listinfo/ls">https://www.link-up.de/mailman/listinfo/ls</a>
Mail to the list server	<a href="mailto:ls@paleotax.de">ls@paleotax.de</a>
Documentation history	V.2 / January 2008 • V.1 / August 2004

CorelDraw, WinWord, Windows, and WMF are registered trademarks.