HaploPainter Manual
Version 1.0

This document is located at the HaploPainter web site documentation
http://haplopainter.sourceforge.net
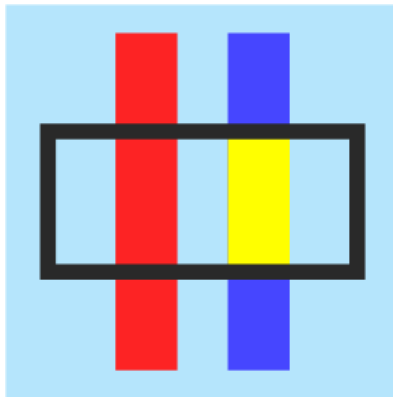
Please be kind to your environment and
<u>do not print</u> this document.

Index

How to start

A good way to start is to prepare a pedigree file in linkage format. This format is very simple and can easily be generated by any text editor. Alternatively you may prepare data in a HaploPainter's specific input format (CSV) or query a database. They allow you to import further individual specific details. After you imported one or more pedigrees, HaploPainter will draw them - coincidental placing individuals while minimizing line crosses.

During the process of pedigree calculation HaploPainter not accepts key input. In the right part of the window some information about the current calculation status is displayed. The example figure below tells you that HaploPainter has finished the calculation No.12, found 5 crosses there and that the best pedigree consists of 1 cross.

Try #12   Crosses #5   min crosses #1

If you don't like the result, you always can HaploPainter let redraw the pedigree by pressing F5. The most changes at the pedigree will be retained until you press this button.

In order to attach haplotype data from other linkage programs (Merlin, Simwalk2, Genehunter, Allegro) you can import those files after you have been loaded pedigree data. Again the haplotype file can consist of data of more then one family. The last step would be the import of mapping data - replacing marker names and coordinates.

Since version 1.0l, HaploPainter changed the way how imported data are handled. If you choose to import pedigree data, HaploPainter will attach new pedigrees to existing one and replace pedigrees with the same family identifier. Imported haplotype data are scanned for family identifiers and are only imported if they match existing pedigrees. Mapping data can be imported for specific families or for all families depending of the "Global map import" check button.

HaploPainter will give you plenty of options to modify the drawing style. You can change it in the "Configuration" menu, available form the "Option" menu or context menu (right mouse bottom). One of the interesting options in respect of haplotype drawing is the ability to manually select which marker should be drawn or not. This is very useful for "long" haplotypes as obtained from genome scans with high-density marker panels of many thousands of single nucleotide polymorphisms (SNPs), which now can be sized to a human readable length. Any part of the haplotype can be surrounded at any individual by one ore more boxes, for instance if you wish to draw attention to regions of interest like recombination events.

If you choose "Save" or "Save as" from the File menu, HaploPainter saves the complete current drawing status of all families and all modifications you made after import of data. You can restore this state by chose "Open" from the "File" Menu. The Save function overlaps with other export functions which produce ether images (PS, PDF, PNG, SVG) or LINKAGE like text files (CSV).

Algorithms, best practice and limitations

The problem with cross free node placing is that even if you find an intelligent way to build on the pedigree de novo, the calculation time for all possible combinations increases rapidly with the number of nodes. HaploPainter is programmed in Perl – a computer language which is not known to be the fastest, but even if you would do it in other languages, big families would tax your patience rapidly.

For this reason HaploPainter pursues a heuristic strategy. Instead of trying every possible combination, it places individuals of a sibling group by chance for a fixed number of times. The very first step is that HaploPainter looks for founder couples and the generation they occur. Those founder couples again are placed by chance between sibling groups of that generation.

For the most common used pedigrees in the field of linkage this approach works well so far. While large but loop less pedigrees make no problems, big families with many loops will bring HaploPainter to its knees. Because of this limitation HaploPainter decides to break all loops during import of pedigrees - if your pedigree consists of more of three of them. You still can change this behavior later with the "Breaking loop" option allowing you to decide yourself which loop should be broken or not.

How ever, for the majority of pedigrees HaploPainter can find cross free solutions or at least good starting points to correct them manually with the graphical user interface. The GUI has almost no limitations in the placement of individual nodes. Connecting lines are still drawn automatically, so you can do it in a short time.

Installation

For some systems HaploPainter is available as perlapp generated executable. As HaploPainter depends of numerous libraries (GTK+, Cairo and Glib) and packing of all this libraries in one file would lead to significant increase in start time, they are outsourced and come together with HaploPainter. The only thing to consider is to keep the libraries and HaploPainter together or to put the libraries in the system PATH. If those executables are not available from the download page for your system, you still could start the program from the source code file. This is a file ending with ".pl" The program language Perl is available for nearby all computer systems and with a bit of knowledge, patience and expert help you can run HaploPainter from source code files on every computer system. The procedures distinguish between systems – here shown for Windows.

In order to connect HaploPainter with an Oracle database you also must install the appropriate drivers or the oracle client. It also may be necessary to set the ORACLE_HOME and NLS_LANG environment variables.

Execute from source files - Windows

Installing Perl

Download and install the latest Perl Version for Windows from ActiveState
www.activestate.com

Installing GTK+ runtime libraries

HaploPainter uses Cairo and Pango for the creation of exported image formats. All of the precompiled libraries you will need are included in the GTK+ package. In order to use this libraries download and install the windows all-in-one-bundle from http://www.gtk.org/ and don't forget to set the system path variable at the bin folder from GTK+ or place the Perl program inside the bin order!

Installing compulsory missing Perl modules

Perl modules from ActiveState's Perl are added via Perl Package Manager - ppm. Type the following commands into the command console.

➤ ppm install Sort-Naturally
➤ ppm install http://www.lostmind.de/gtk2-perl/ppm/MSWin32-x86-multi-thread-5.8/Glib.ppd
➤ ppm install http://www.lostmind.de/gtk2-perl/ppm/MSWin32-x86-multi-thread-5.8/ExtUtils-Depends.ppd
➤ ppm install http://www.lostmind.de/gtk2-perl/ppm/MSWin32-x86-multi-thread-5.8/ExtUtils-PkgConfig.ppd
➤ ppm install http://www.lostmind.de/gtk2-perl/ppm/MSWin32-x86-multi-thread-5.8/Cairo.ppd
➤ ppm install http://www.lostmind.de/gtk2-perl/ppm/MSWin32-x86-multi-thread-5.8/Gtk2.ppd

(The perl binding modules for the GTK stuff may be available for newer versions and be stored at other places!)

Installing facultative missing Perl modules for database support

In order to use HaploPainter's capabilities to query databases you must install some more modules. Without these modules HaploPainter will start anyway, but will produce warnings when you try to use this functionality.

➤ ppm install DBD-PgPP
➤ ppm install DBD-mysql
➤ ppm install DBD-Oracle

Other Systems

For other systems it could be necessary to compile libraries yourself. Field reports are welcome!

Data Import

For data import there are Import Pedigree …, Import Haplotypes … and Import Map File … options available from the File menu. It is the best to follow this order. Don't be confused with the Open command! That's for the HaploPainter specific data storage. Those files become the file extension '.hp' by default and are not human readable.

Import Pedigrees

There are two kinds of text input files that HaploPainter accepts to define pedigree structures – LINKAGE and CSV. The CSV format is designed in a way, that the first 6 columns have identical meaning comparing to LINKAGE format. If you start to collect

pedigree data in a database table you must prepare a table or view with the same structure as used for CSV format. Per default HaploPainter handles text data encoded as UTF-8. You can change it via File → Encoding. HaploPainter deals with Unicode in a way that it looks for the byte-order mark (BOM) sequence 'U+FEFF' at the beginning of documents, and only if it did not found this sequence take care of the setting. In that way it is for instance possible to read Excel sheets which are saved as *Unicode Text format* (UTF-16LE) although the setting in HaploPainter is different. The encoding is much important for saving/exporting pedigrees in CSV format. In this case HaploPainter take care of it. (See also the chapter Encoding and Unicode support)

LINKAGE

Well, there are actually many different LINKAGE pedigree files used by statistical software. The format used here is also called pre-MAKEPED LINKAGE format.

Example

Family 2002
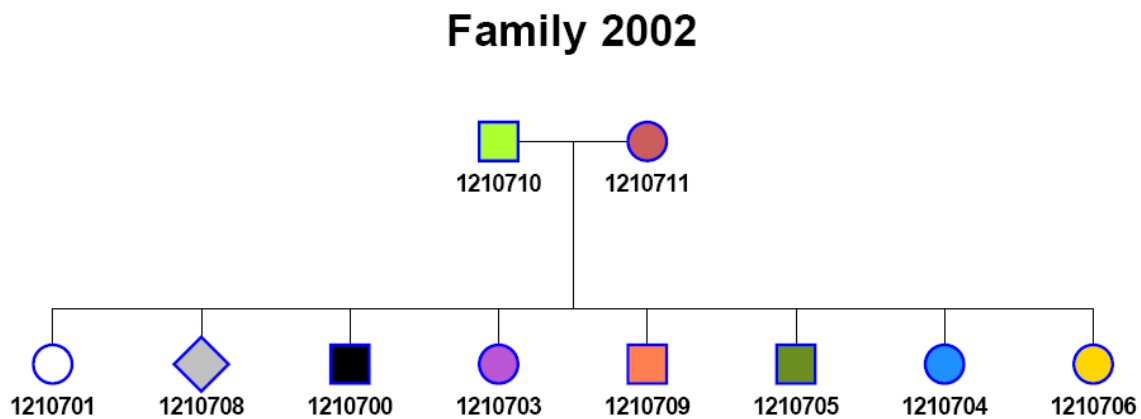
#FAMILY	PERSON	FATHER	MOTHER	GENDER	AFF_STATUS
2002	1210708	1210710	1210711	0	0
2002	1210701	1210710	1210711	2	1
2002	1210700	1210710	1210711	1	2
2002	1210709	1210710	1210711	1	3
2002	1210706	1210710	1210711	2	4
2002	1210705	1210710	1210711	1	5
2002	1210704	1210710	1210711	2	6
2002	1210703	1210710	1210711	2	7
2002	1210710	0	0	1	8
2002	1210711	0	0	2	9

HaploPainter is tolerant in the definition of the exact format. Empty lines and those beginning with the signs '#', '!' '*' or '\' are ignored and may be used to integrate comments. HaploPainter tries to find out the used column delimiter in a way that it first looks for semicolons, followed by tabulators and at least other white space. Only lines resulting in 6 or more not empty columns are further processed, and only columns 1-6 will be taken into account. White space at the beginning and end of all values is deleted.
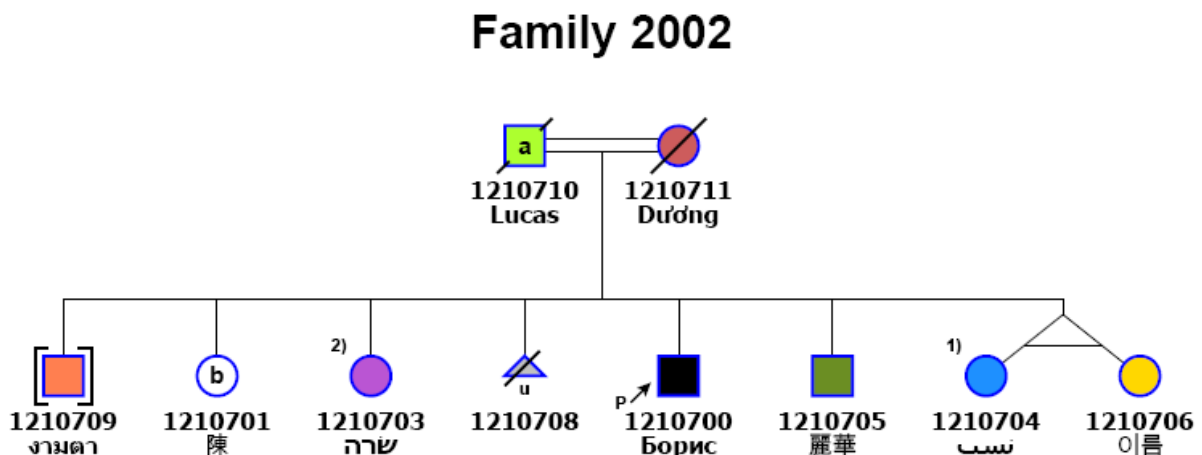
LINKAGE format

FAMILY	not NULL
PERSON	not NULL
FATHER	not NULL
MOTHER	not NULL
GENDER	not NULL, 0 or x for unknown, 1 or m for male, 2 or f for female
AFFECTION	not NULL, 0 or x, 1, 2, 3, 4, 5, 6, 7, 8, 9

As you can see, all columns must have defined values! The used signs x/m/f for gender and affection status also can be in upper case. The affection status field in the originally LINKAGE format only can hold the values 0, 1 and 2. This can be extended in HaploPainter by up to 10 possible values. Every affection status is assigned to a defined color with what symbols are filled later.

CSV

The CSV format is influenced by the LINKAGE format but differs in the number of possible columns and the way it is parsed. As in LINKAGE format, empty lines and those beginning with the signs '#', '!' '*' or '\' are ignored and may be used to integrate comments or header rows. Only a single tabulator sign is allowed as column separator! Columns 1-6 apply the same rules like LINKAGE format (See also there). Field values again are case insensitive. The following figure shows the same pedigree from the sample above but extended by HaploPainter specific columns. It illustrates many of HaploPainters additionally features like Unicode support, twins, termination of pregnancy symbols, deceased individuals, adopted individuals or consanguineous individuals.

Family 2002

CSV format

FAMILY	not NULL
PERSON	not NULL
FATHER	not NULL
MOTHER	not NULL
GENDER	not NULL, 0 or x, 1 or m, 2 or f

AFFECTION	not NULL, 0 or x, 1, 2, 3, 4, 5, 6, 7, 8, 9
IS_DECEASED	NULL or 0 or n, 1 or y
IS_SAB_OR_TOP	NULL or 0 or n, 1 or y
IS_PROBAND	NULL or 0 or n, 1 or y
IS_ADOPTED	NULL or 0 or n, 1 or y
ARE_TWINS	NULL, [m or d]_text
ARE_CONSANGUINEOUS	NULL, text
TEXT_INSIDE_SYMBOL	NULL, char(3)
TEXT_BESIDE_SYMBOL	NULL, text
TEXT1_BELOW_SYMBOL	NULL, text
TEXT2_BELOW_SYMBOL	NULL, text
TEXT3_BELOW_SYMBOL	NULL, text
TEXT4_BELOW_SYMBOL	NULL, text

Meaning of additionally fields

IS_DECEASED: Set to 1, y or Y if the individual is deceased. In that case a diagonal line through the symbol representing that individual is drawn. This line is spitted into two lines if text in the center of the symbol is placed

IS_SAB_OR_TOP: Set to 1, y or Y if individual is the result of a spontaneous abortion or termination of pregnancy. The individual symbol is drawn as a triangle now. Later you can define in configurations if the gender is placed as text below the symbol or not.

IS_PROBAND: Set to 1, y or Y if the individual is declared as proband. (The first affected family member coming to medical attention) This result in a 'P' sign attached to an arrow pointing to that individual.

IS_ADOPTED: Set to 1, y or Y if the individual has been adopted. This individual is drawn in brackets.

ARE_TWINS: Individuals sharing the same string in this column are recognized as twins. The first character in that string tells HaploPainter if the twins are monozygotic 'm' or dizygotic 'd' followed by an underscore followed by an unique identifier. Examples for allowed strings are: m_1, d_gruppe1. Twins are drawn according to Bennett et al. 1995. A twin group must contain two individuals or more.

ARE_ CONSANGUINEOUS: Individuals sharing the same string in this column are recognized as consanguineous (individuals are related to each other). This only makes sense if individuals have offspring. In that case a double line between couples is drawn. Although HaploPainter automatically detect consanguinities, for instance if it find loops, you additionally can define any couple as consanguineous. This is good for situations when the real degree of consanguinity is unknown or you want to draw more compact pedigrees. A consanguineous group must contain two individuals. In the consanguineous menu you also can define that.

TEXT_INSIDE_SYMBOLS: Place text inside symbols.

TEXT_BESIDE_SYMBOL: Place text left upper hand to symbols

TEXT<n>_BELOW_SYMBOL: HaploPainter gives you the opportunity to place text below symbols. Per default the individual identifier is one of those text lines, but you can attach up to four further text lines below. In the configuration menu you can arrange the combination of columns you want to show in your pedigree (See Ped & Case Info).

Database

Import of pedigree data from data base tables is supported – to date for MySQL, PostgreSQL and Oracle. To use this feature the proper Perl modules must be installed (Only if you run HaploPainter from source file! HaploPainter will work without DBD modules because they are loaded at runtime but will result in warnings). For Oracle you also need the Oracle client, NLS_LANG and ORACLE_HOME environment variables and an entry in the tnsnames.ora file. You reach the data base connection window by pressing F4 or via File->Import Pedigrees … -> Database. For databases that are located at your computer you also can say 'localhost' instead of using the host name. After you successfully are connected to the database a "Choose Pedigrees" menu will appear – here is an example:
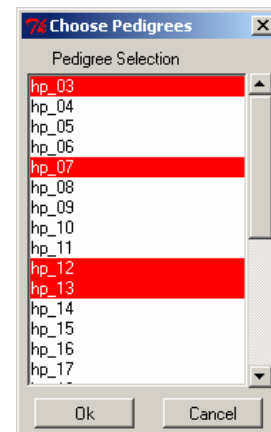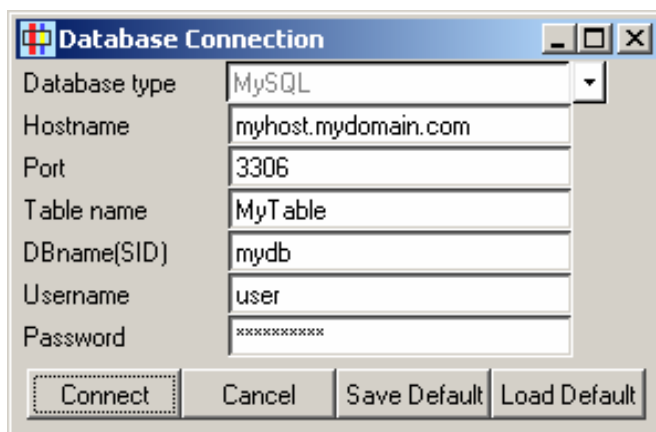
Inside this box all families are listed and can be marked by use of the left mouse bottom in combination with the shift and control keys. All marked pedigrees are imported after pressing 'Ok'.

The Database Connection window has two other functions to save the connection data as default (Save default) and reload them (Load Default).

Import Haplotypes

A haplotype describes the linear sequence of alleles at multiple loci in the genome that are transmitted on the same chromosome. In diploid cells a genotype is composed at two alleles. Commonly used methods for measurement of genotypes are not able to distinguish which allele belongs to which chromosome, but in many cases likely haplotypes can be reconstructed through statistical methods. Haplotypes are nice instruments in the interpretation of linkage results. They are able to map a disease causing gene to a region in the genome. HaploPainter supports four statistical software packages that are usual in the calculation of haplotypes:

Genehunter, Allegro, Simwalk2 and Merlin. Specific output files produced by this software can be imported and visualized together with pedigrees. Import haplotype files via File → Import Haplotypes and choose the import format:

- Genehunter → haplo.dump
- SimWalk2 → HAPLO.??? (horizontal orientation)
- Merlin → merlin.chr (vertical orientation)
- Allegro → haplo.out, ihaplo.out, founder.out

SimWalk2 controls the way it outputs haplotype results through a batch file. In batch number #46 line 1 you can switch between default horizontal (N) or vertical (Y) output orientation. HaploPainter expects to keep the default value (N). Merlin also has two possible output modes. Per default Merlin produces files with haplotypes in vertical orientation – the format that HaploPainter supports. You control that in Merlin over the command console parameter '--horizontal'.

Supplementary information enclosed with those output files such recombination events are ignored by HaploPainter. It uses the following rules for haplotype coloring.

- Try to find out the first marker beginning from p-telomere from which the phase can be derived and then trace back the haplotype with the color from that phase up to the chromosomal starting point.
- The first marker showing differences is declared as the point of recombination and the color is changed to the recombinant haplotype until the next recombination event occurs.
- Missing genotypes are interpreted as uninformative.
- Other uninformative genotypes are drawn in special thin blocks when set in options.

After you have imported haplotypes, if you are not satisfied with the colors HaploPainter have been choosing for you, you can press F6 to mix all colors new. The configuration menu gives you the opportunity to manually modify colors of haplotypes or fix some colors so that pressing F6 has no effect for those. You also can prevent markers for being shown.

Some tips for interpreting haplotypes

Points of recombination are often surrounded by uninformative markers. Due to the way HaploPainter deals with analysis results it is possible that your region of interest is shown too small or shifted. That should be checked and manually corrected! To do this, just double click at alleles from those individuals. The menu which appears gives you an instrument to change the phase or declare markers as uninformative. You may find out uninformative markers if you move the mouse cursor over alleles which became flashing in red in that case. In general you should orientate at the results of your multipoint linkage. Also be distrustful if you see double recombinations (two neighbored recombinations). They mostly are hints for genotyping errors. Some software packages tend to complete haplotypes for not genotyped individuals. You never should thrust those haplotypes – see it at as a suggestion.

Import marker maps

To replace dummy marker names which are automatically created after import of haplotype data you need to import a map file too. There are only two columns of interest: marker name and marker position. These columns are expected to be at positions 2 and 3 of the input files. The expected column separator is white space. As in LINKAGE format, empty lines and those beginning with the signs '#', '!' '*' or '\' are ignored and may be used to integrate comments or header rows. There are two options for map files concerning the column positions:

- Chromosome   Marker Name   Marker Position   (CHR-POS-MARKER)
- Chromosome   Marker Position Marker Name   (CHR-MARKER-POS)

As you can see, the only difference is a switch between columns 2 and 3.

A map file can be loaded specific for a family or for all families. This depends on the setting of the 'Global map import' flag in the File menu. This flag is off per default, with other words in that case an imported map file only belongs to the current drawn family.

Data Export and Publishing Tips

Since Version 1.0 HaploPainter uses internally Cairo, Pango and GTK+ for rendering exported graphics. If you use the stand alone versions you don't need additionally programs to convert images into other formats. Only if you plan to run HaploPainter in your own Perl environment you have to take care about that (see installation part). Within HaploPainter there are different formats available in **vector describing languages** (Postscript, PDF, SVG), **pixel graphic** (to date just PNG) and **flat file** (CSV). Keep in mind, that images produced by HaploPainter will cover the RGB color space. If your plans are images produced for the web this doesn't matter but if not you must take car of this. Some journals like to have graphic in CMYK colors, and probably you don't have much experience with transformations from RGB to CMYK. Unfortunately this is not as easy as you might think, because you must choose the right color profile that should be used for transformation of color spaces. It is not possible to describe every RGB color in CMYK. Furthermore the end result and color impressions depend of the calibration of your monitor the used paper and other printing dependent stuff. So, whenever possible let the journal do this job for you or ask them which profiles give the best results. In the case that you want to do it yourself you should take a look at Adobe Acrobat and Adobe Photoshop. It also is helpfully to print the image on a CMYK color laser printer to have a rough idea about the color conversion effect. Many journals prefer TIFF images with embedded ICC profiles and color resolution greater 300 dpi. Sending of images in other vector graphic formats, which are accepted (like AI/Adobe Illustrator, EPS, PS, PDF, CoralDRAW/CDR, FreeHand/FH) would be of advantage (at least for you) as the layout and printer professionals could transform the drawing to the best resolution and colors herself. Those images send by you also are much smaller in size.

Vector describing formats

Vector images describe drawing elements as mathematically formula. The main advantage is its scalability – you can create every possible image resolution you want! The other advantage is you can tailor images within common desktop publishing software.

SVG

Scalable **V**ector **G**raphics is based on XML and has been developed as open standard from the World Wide Web Consortium (W3C) since 1999. Since that time most modern web browsers integrated native support (Firefox, Opera, Safari, Chrome). It seems that Adobe and Microsoft try to subvert this standard, because Adobe as most known provider for SVG browser plug-ins loosed their interest to give further support after buying Macromedia and Microsoft itself interferes with its own proprietary vector format VML (which again is ignored by other browsers). How ever, SVG is a good choice if you plan to integrate images into web pages. Probably Microsoft's internet explorer will support SVG sometimes if the SVG diffusion rate increases over the web. You can import SVG into a plenty of desktop publishing software like the free available Inkscape.

Postscript

Adobes Postscript is something like a program language for printers. Most printers today are equipped with Postscript interpreters – so you may use Postscript to deal with high quality printer jobs. Programs like Ghostview/Ghostscript are able to view them or convert Postscript into PDF, EPS or other graphic formats.

PDF

Adobe's PDF is the de facto standard for platform independent document exchange. The format is very compact and viewers for every platform are free available – so my recommendation is to take this format when ever possible.

Pixel formats

PNG

PNG is a good choice for displaying pixel graphics particularly in web browsers. The format is free for use, has better performance then GIF and even Microsoft's Internet Explorer will show it …

Flat file

CSV

Exported CSV format is identical with the imported counterpart. So you may import files in LINKAGE format, make your modifications within the GUI and export it as

extended CSV/tab delimited version. This also would be an easy way to produce templates for other pedigrees.

Encoding and Unicode support

Since HaploPainter Version 1.0 reading and writing files in Unicode is supported. There are a lot of Unicode coding types, but UTF-8 and UTF-16LE seem to be the most relevant do date. UTF-8 is the most common type and going to be the web standard. So why UTF-16LE? The answer is Microsoft Excel. Using Excel to create CSV files ready for import into HaploPainter is a good idea! Although most text editors would do their job well, the handling with tables in Excel is quite comfortable. If you plan to work with strange characters (for you they are surely not strange) you must save data in Excels Unicode Text format and this is in fact UTF-16LE. The encoding type setting is located inside the file menu. There are three kinds available: ASCII, UTF-8 and UTF-16LE – the default being UTF-8.

May be you wondered about the check button "Write BOM" available from the File menu. This is the special Unicode character 'U+FEFF' (byte-order mark) represented by a series of 2-4 bytes and is used as marker to indicate the encoding type. I don't want to comment the actual discussion if this is good or not! HaploPainter's default is to write a BOM sequence for Unicode files – and you can switch it off, if you want! HaploPainter is able to read Unicode files which are encoded in ASCII, UTF-8, UTF-16LE, UTF-16BE, UTF-32LE and UTF-32BE. If a BOM sequence is found HaploPainter take care of it – whatever encoding type is preset.

Working with Pedigrees

Navigation

Once you imported the pedigrees, haplotypes and mapping data some functions to navigate through the pedigrees are available for instance the navigation menu.

The navigation menu

The arrow symbol is activated by default. In this mode you can activate symbols ether by clicking single symbols or clicking the mouse button while moving. This force HaploPainter to draw a rectangle with dashed line. Every symbol inside this rectangle is activated – indicated by a change to red. Now you can move symbols around by drag and drop. Note that you must place symbols at positions that are not occupied by other symbols. Those symbols will not be moved! Zooming in or out the drawing is available in different ways – via plus/minus navigation menu symbol, plus/minus keys, Edit menu and Context menu. Use the hand symbol to shift the drawing by clicking the mouse button to have a fix point while moving the mouse. In case of many imported pedigrees you can change the current pedigree by clicking the green left/right arrows (resulting in a shift of +/- next one pedigree) or alternatively use the View → Draw Pedigree menu. This menu also is able to separate from the main menu to easy navigate through the pedigrees. Two other useful functions to center or fit the pedigree into the current window – Center View/Fit View are accessible from the Edit and Context menu.
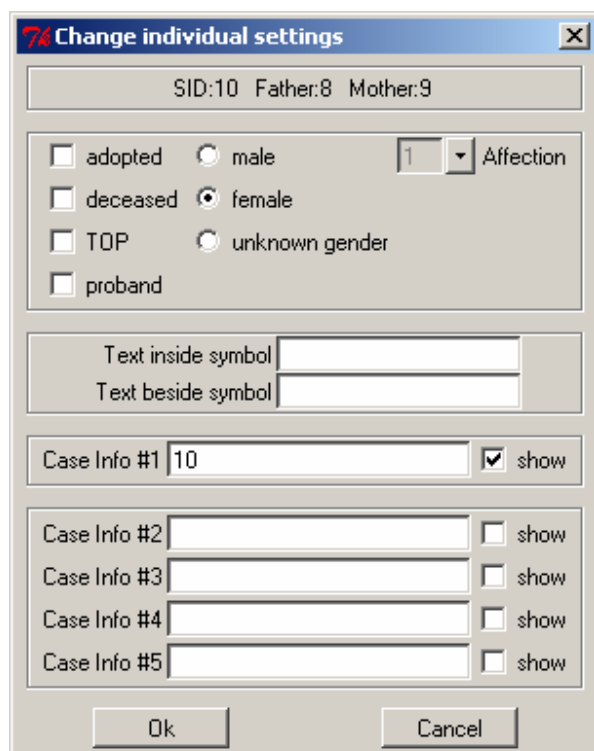
Redrawing Pedigrees

Most modifications you made during a pedigree live will be conserved until you do some of the following things:

- carry out the Redraw Ped (F5) function
- make changes in the way how loops are broken
- make changes in the Inter symbol distance value from the Configuration/Ped Style menu
- extending pedigrees with the GUI functions

Redrawing a pedigree will force HaploPainter to draw the pedigree according to the settings of the Options → *Sort by SID* and *Sort couples by gender* flag. If the *Sort couples by gender* flag is not set, the individuals belonging to a couple group are assembled by chance. If the flag is set, the **male individuals are placed first**! That's the only difference. The second modifier in respect of drawing order - *Sort by SID* reflects the handling of sibling groups. By default the order of siblings is chosen by chance. If the flag *Sort by SID* is set HaploPainter will change the drawing order of the members of a sibling group and applies a special name sorting mechanism (nsort from the module Sort::Naturally). The algorithm works in a way human normally would think about sort and is an intelligent combination of numerically and lexically sort. By using both flags you can HaploPainter force to draw pedigrees in a predetermined way taking individual naming into account.

How to change individual settings

If you *double click at a symbol* the following menu will appear that allows you to change individual settings and displays some useful information.

14
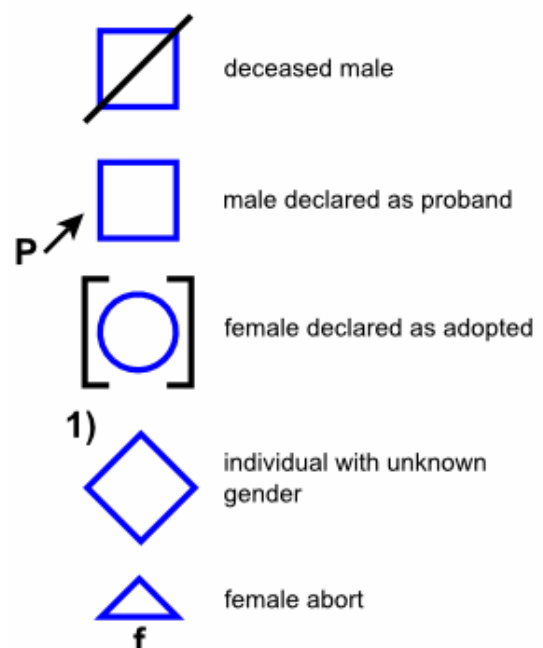
Change individual settings by double click at symbols   **Examples for symbols**

Affection	A numerical value ranging from 0-9. Dependent of the settings in Options menu → Ped Color the symbol is filled with that color
Male	Males are drawn as squares
Female	Females are drawn as circles
Unknown	Individuals with unknown gender are drawn as rotated squares
Adopted	Adopted individuals are drawn with brackets around the symbol
Deceases	Deceased individuals are drawn with a diagonal line through the individual. The line is divided in case of text inside symbols
TOP	**T**ermination **o**f **p**regnancy and aborts are drawn as triangle symbol. The gender is drawn as one letter (m/f) below the symbol
Proband	Attaches an arrow with the letter P left to symbols and is used to indicate persons that primarily came to medical attention
Text inside symbol	Place text in the middle of symbols
Text beside symbol	To put some text to the left upper corner of symbols
Case Info #1	This is the individual identifier which is placed below symbols. It is not allowed to duplicate this name. The show check button left to this field controls the display for all individuals
Case Info #2-5	You may put further information text lines below symbols by using these fields. The show check button left to this field again controls the display for all individuals. You can fine tune the settings for Case Info from the Options menu → Ped & Case Info

Building and extending pedigrees with the graphical user interface

The functions to create new or modify existing pedigrees are available from the Edit menu or from the mouse Context menu (here exemplary shown).

The Context menu

New Pedigrees

To create a new pedigree you also can press the key combination "Contr+N". A window will appear that asks for a pedigree name. After saying "Ok" HaploPainter draws a core family consisting of a couple and one offspring with preformed individual identifiers. To change the identifiers, gender or other individual settings double click at the respective symbol.

Delete Individuals

To delete individuals mark one or more symbols and chose the "Delete Individuals" function. You also can do this by pressing the keys "Contr+D" or "Del". Remaining childless founders will be automatically deleted too!

Add Siblings

To add one or up to 8 siblings to another one mark one individual and execute the "Add Sibling" function. A window will appear that asks for sibling identifiers and gender. After saying "Ok" HaploPainter attaches all siblings for which you entered a valid identifier to the pedigree. To change the identifiers, gender or other individual settings double click at the respective symbol.

Add Parents

Parents can be added only to founder individuals. Mark the founder and execute the "Add Parents" function. A window will appear that asks for the mother and father identifiers. After saying "Ok" HaploPainter attaches a new founder couple to the pedigree. To change the identifiers or other individual settings double click at the respective symbol.

Add Mate & Offspring

You can attach a new mate and offspring to any individual. Mark one individual and execute the "Add Mate & Offspring" function. A window will appear that asks for the mate and offspring identifiers. After saying "Ok" HaploPainter attaches a new founder (only if that mate not already exists) and a new offspring with unknown gender to the pedigree. To change the identifiers, gender or other individual settings double click at the respective symbol.

Consanguinity Settings

Consanguineous individuals are automatically detected if the pedigree consists of the according loop structure. How ever, you can declare any couple as consanguineous too. The function to change consanguineous for all couples at once can be achieved from the Options and Context menu → Consanguinity Settings. A new window displaying all couples will appear. Here you can mark any combination of entries, preview the changes and permanently change the new setting by pressing "Ok".
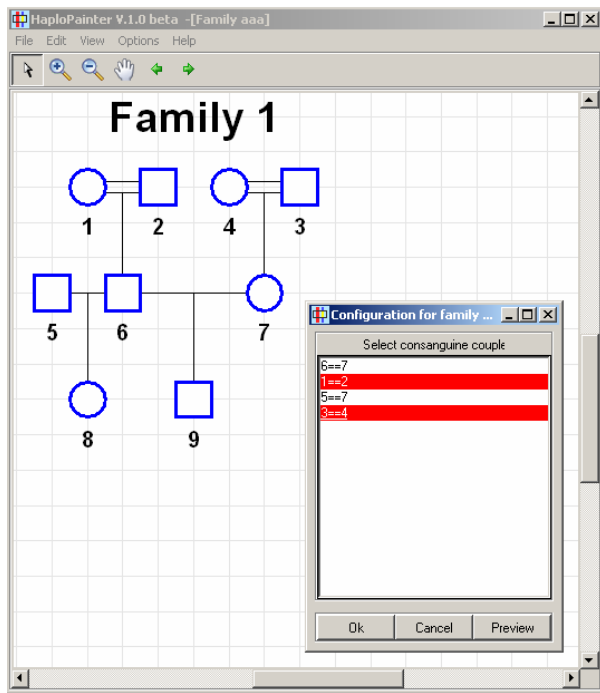
Consanguinity Setting menu

The other way to quickly change the setting is to mark the two symbols of one couple and choose "Set/Unset Consanguinity" from the context menu or Edit menu.

Twin Setting

To declare/undeclare two or more individuals as twins mark the symbols of a sibling group and choose "Set/Unset Twins" from the context menu or Edit menu. A new submenu gives you the opportunity to mark twins either to be monozygotic or dizygotic. The differences are shown in the following figure where the individuals 3+4 and 5+6+7 are declared as monozygotic twins and the other twin groups as dizygotic.
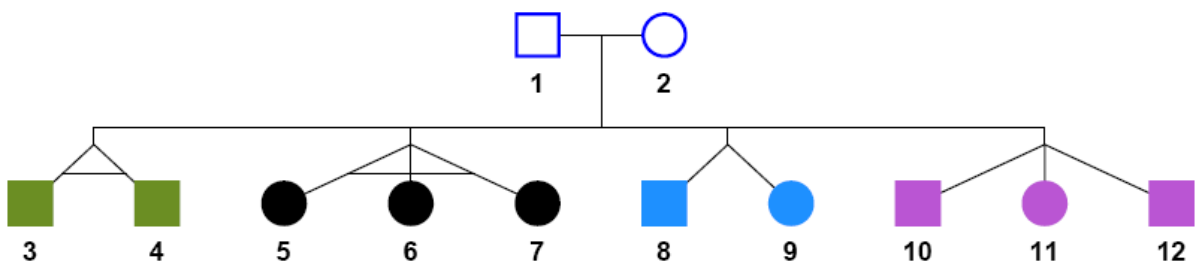

In the example above the "**Inter siblings Y distance**" parameter from the Configuration → Ped Style menu was increased to a value of 45 to better separate the twin groups from the horizontal sibling connecting line. This value is set per default to a lower value to have more space for haplotypes and case information.
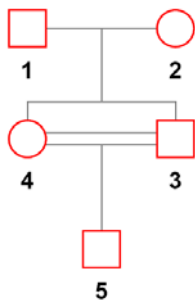
Loop breaking

In respect of finding a good drawing solution, loops are the major challenge! Breaking a loop can be the only way to draw the pedigree without line crosses. Loops are broken by a duplication of an individual that is located and the "end" node of a loop.

compulsory broken loop
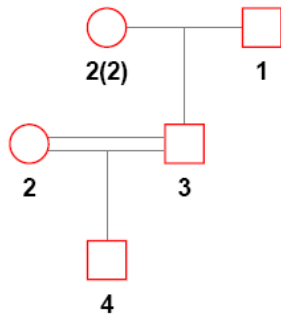
Due to HaploPainter's internal algorithms some loops are always broken, for instance the pedigree on the left that belongs to the category parent-offspring mate situation.

offspring mating

On the other hand there are situations when breaking of loops makes no sense - like the offspring mating situation shown on the left. Those unproblematic loops would not lead to line crosses and therefore will never be broken or selectable for breaking.

As more loops a pedigree consists as more time it takes to find a good drawing solution. For this reason HaploPainter breaks loops automatically when it finds more then 3 of them. Later you can decide which loops should be broken or not. To select them there is a "Breaking loops …" window, available from the Options or Context menu. The following figures show the test pedigree hp_19 which consists of 3 loops. After you followed the "Breaking loops …" function a new window appears that gives you the possibilities to choose either select breaking no loops, all loops or to manually select loops for breaking. Here it was selected to manually break the loop ending with the individuals 8==9.
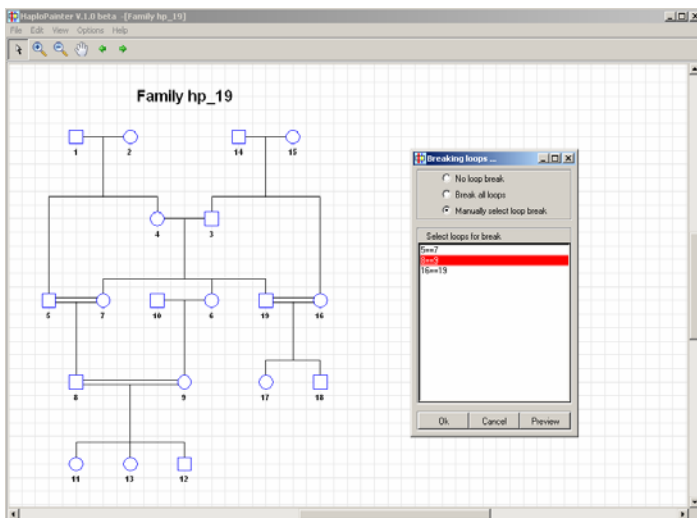
Loop breaking selection window
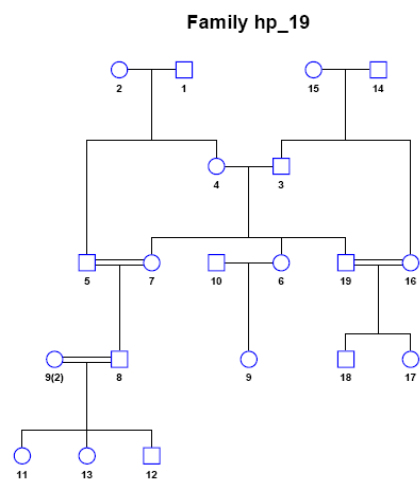
Result of breaking the loop: 8==9

Clear families

There are two functions to completely remove the current drawn ('Clear current') and all families ('Clear all') from HaploPainter. They are available from the File menu.

The configuration menu

The configuration menu is the most important place to modify the way haplotypes are drawn and the drawing style of pedigrees. To illustrate the effects of all parameters would be a hard job! A practicable way to find out what a parameter really does would be to change one parameter and **press the preview button to see what happens!** You start the menu either from the Options or Context menu. The configuration menu consists of 7 tabs.

IMPORTANT: any changes in the settings will only affect the currently drawn pedigree!

Hap Style

The settings in this menu will mainly affect the haplotype drawing style. Parameters collected there have quantitative values and are changed via scroll bars.
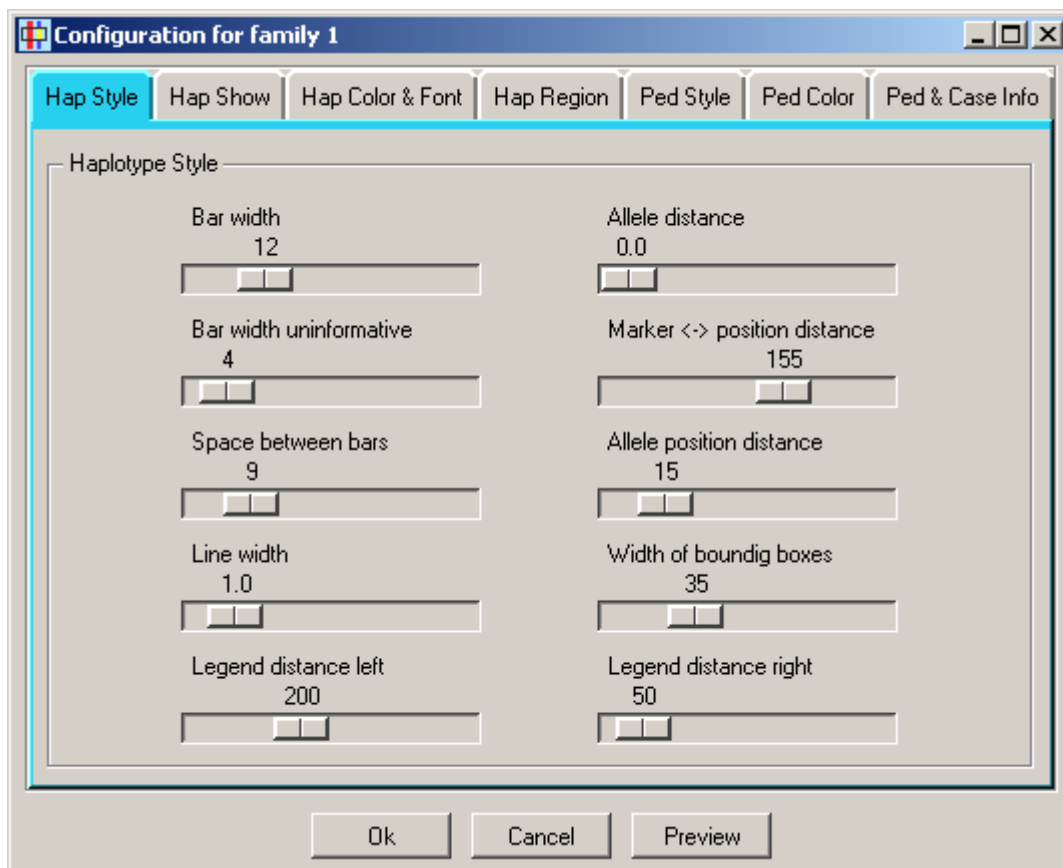

Bar width	Size of haplotype bars in X- dimension
Bar width uninformative	Size of bars from uninformative alleles in X - dimension

Space between bars	Space between maternal and paternal haplotype from one person
Line width	Bar border line width, makes sense without 'fill out bars' option
Legend distance left	Distance between left marker legend and pedigree
Allele distance	Space between alleles in Y - dimension
Marker position distance	Distance between left marker legend and pedigree
Allele position distance	Distance between alleles and bars in X- dimension
Width of bounding box	Effects bounding boxes of marked haplotype blocks
Legend distance right	Distance between right marker legend and pedigree

Hap Show

The settings in this menu also will mainly affect the haplotype and legend drawing style. Parameters collected here have qualitative aspects and are changed via check buttons.
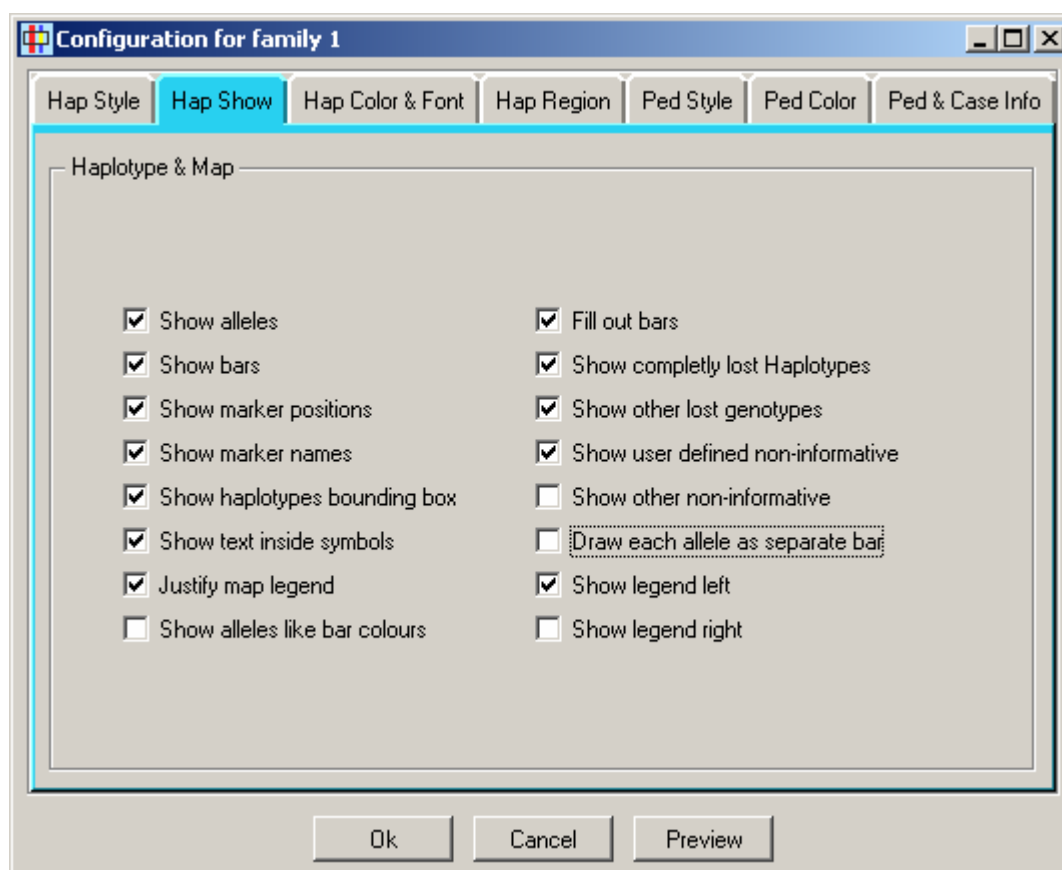

Show alleles	Set if allele numbers are shown
Show Bars	Set if haplotype bars are shown
Show marker positions	Show marker coordinates on the left side of each generation
Show marker names	Show marker names on the left side of each generation
Show question mark	Show question mark inside person symbols for unknown affection status
Fill out bars	Fill out bars at given color

| Show completely lost Haplotypes	Mark haplotypes without information
Show other lost genotypes	Mark unique lost genotypes
Show user defined non-informative	Mark manually, as not informative declared genotypes
Show other non-informative	Mark genotyped but uninformative genotypes
Draw each allele as separate bar	Good idea only when 'Fill out bars' option is not set
Show legend left	Show the left haplotype marker legend
Show legend right	Show the right haplotype marker legend

Hap Color & Font

In this menu you can change the colors and font attributes of different types of information content like alleles from haplotypes, case information, pedigree header or phase unknown color. If you want to show all haplotypes, except of one in a unique color, press the 'Color of all haplotypes' button. Before you do so, you should anchor one or more haplotypes of course! To do that chose the founder from where the haplotype derives in the browse entry box and activate 'Anchor paternal/maternal haplotype' in the check button. You also can fade out haplotypes by activation of 'Hide paternal/maternal' check buttons.
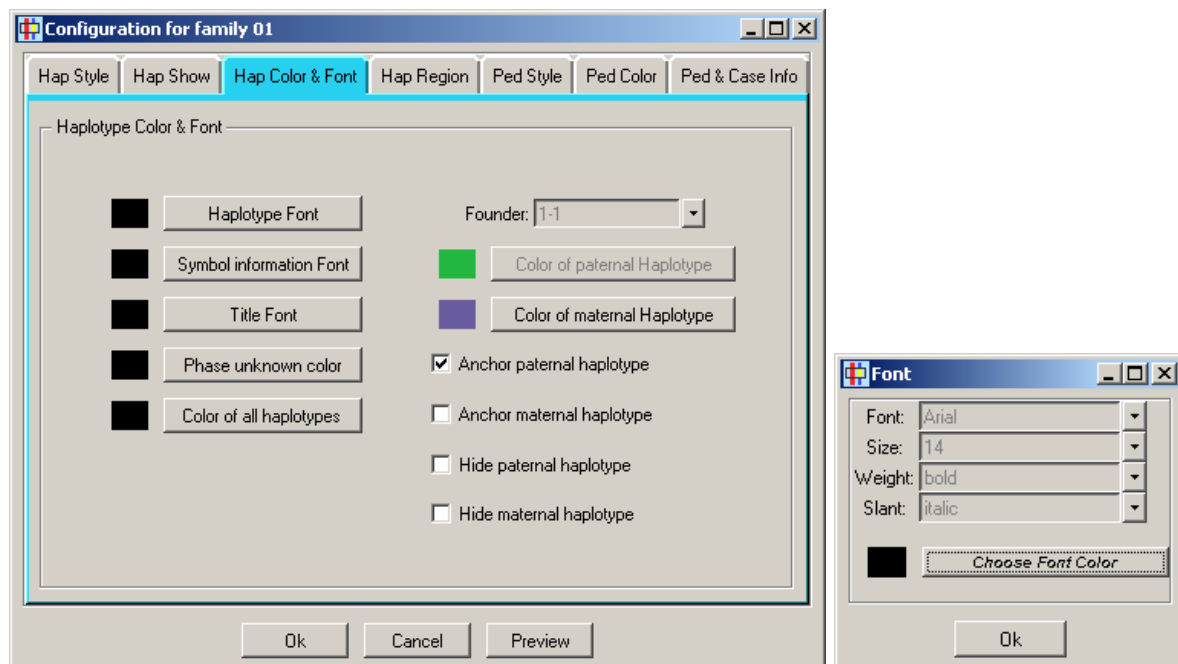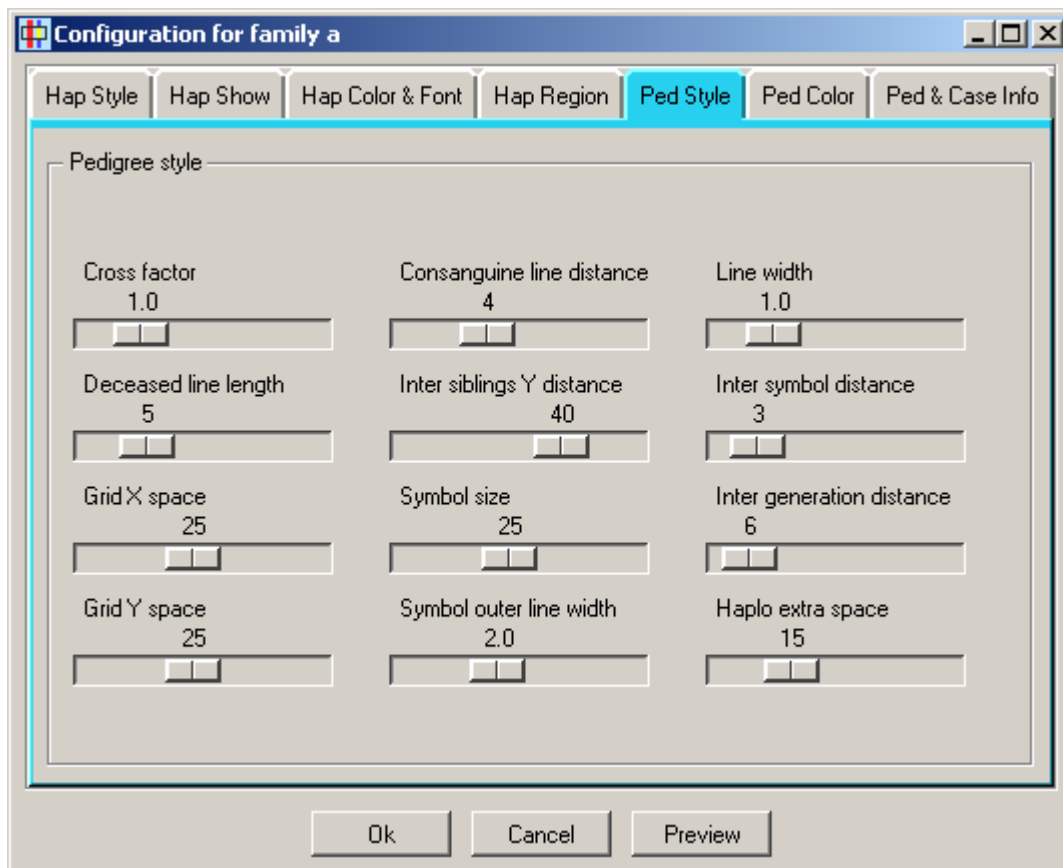

Hap Region

The list box on the left side shows marker positions and names. Only red marked markers will be shown. To select/unselect multiple markers keep Control key pressed while clicking with the left mouse button over rows. Press "control-a" to select all markers.

The list box on the right side allows drawing of black unfilled boxes around markers. You can define as many boxes as you want for any individual which has to be selected with the browse entry field above.
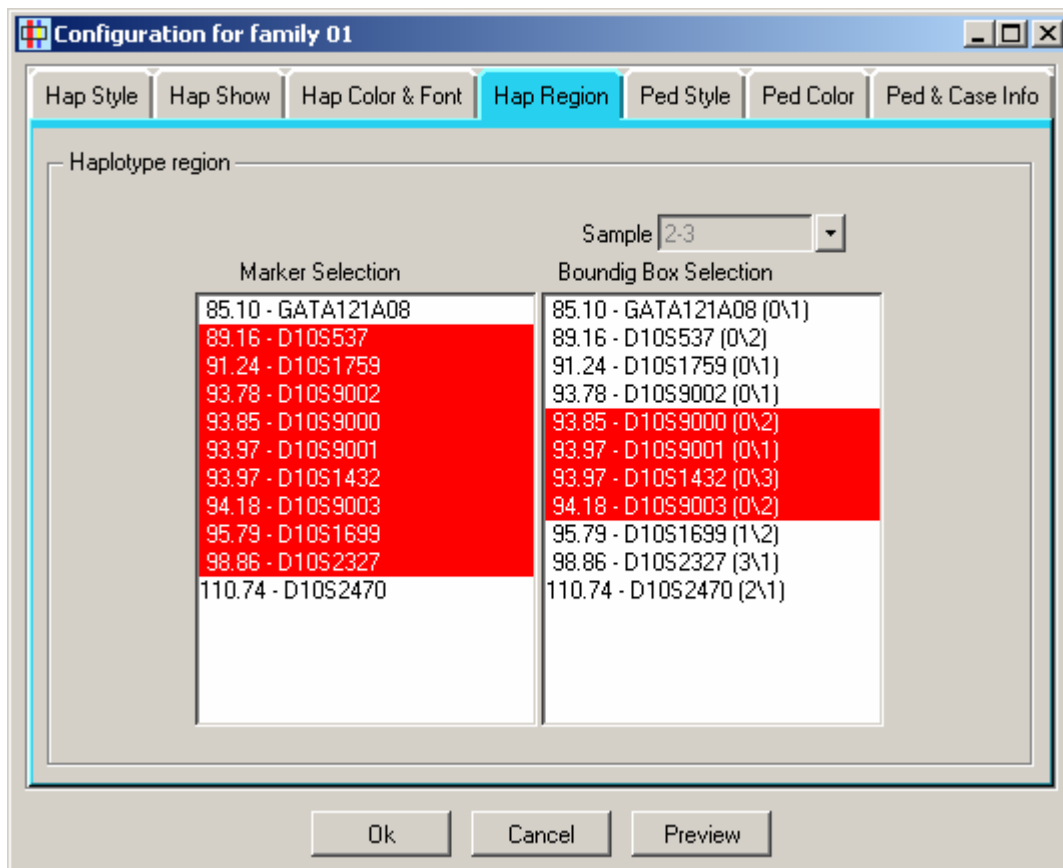
Ped Style

The settings in this menu will mainly affect the pedigree drawing style. Parameters collected there have quantitative values and are changed via scroll bars.

| Cross factor	Affects the lines connecting parents with their offspring group
Deceased line length	Size of the diagonal line indicating that individual is deceased
Grid X space	Width of grid raster in X - dimension
Grid Y space	Width of grid raster in Y - dimension
Consanguinity line distance	Distance between the two lines indicating that individuals are consanguineous
Intersib Y distance	Affects the lines connecting members of a offspring group with each other
Symbol size	Symbol size
Symbol outer line width	Width of the outer line of symbols if it is shown
Line width	Width of symbols connecting lines
Inter symbol distance	Number of grid units that are placed between symbols in X - dimension
Inter generation distance	Number of grid units that are placed between symbols in Y - dimension
Haplo extra space	Fine tune the length of haplotype bars. It only works for haplotypes that are drawn without alleles and legends when bars are compressed/elongated to the available space according to the inter generation distance

Ped Color

The Ped Color submenu allows changes for the most color relevant parameters.
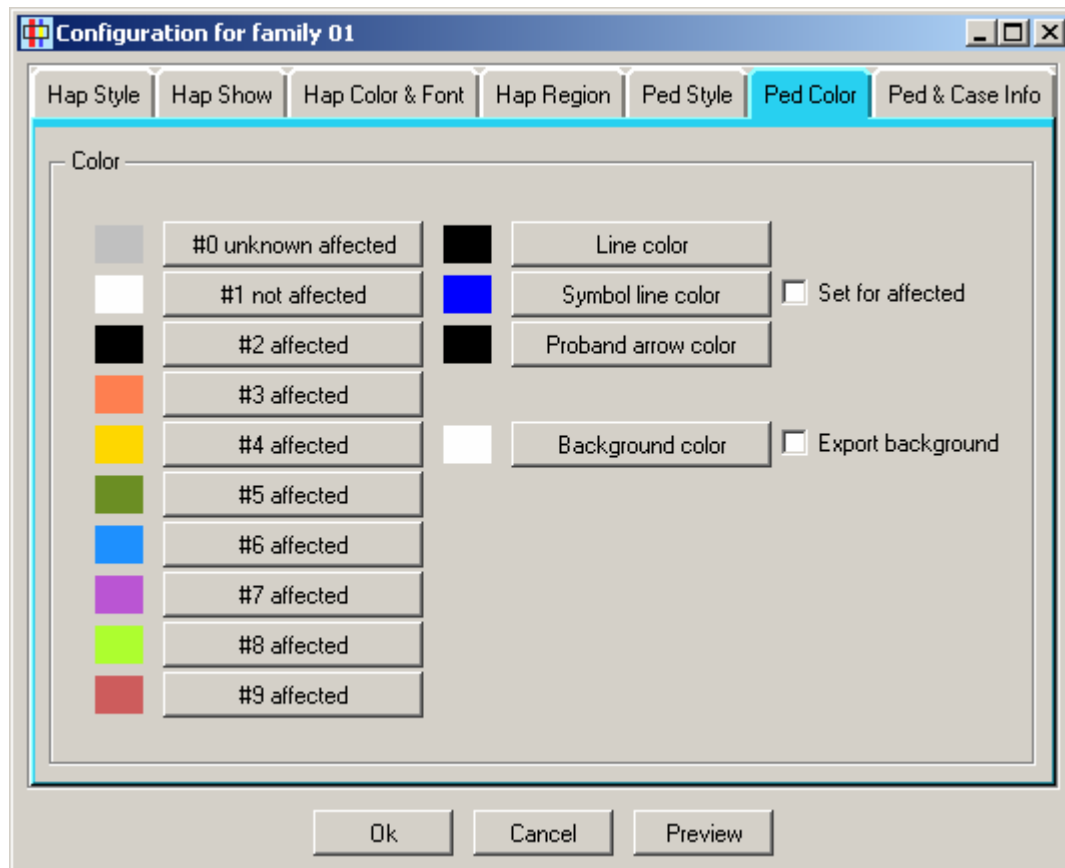
| Affection colors #0-9	Symbols are filled with colors according to their affection status value ranging from 0-9
Line color	The color of lines that connect symbols with each other
Symbol line color	The color of outer lines of symbols. The check box "Set for affected" right hand of the button controls if this line is visible for individuals with affection status other then 1
Proband arrow color	The color of the proband arrow
Background color	Set the background color. It is only exported into other graphic formats if the Export background checkbox flag is switched on.

Ped & Case Info

In this menu there are settings for the title and case information. The rows from case 1-5 are shown below symbols if the check button is switched on. Furthermore you can choose in which order the fields are drawn.
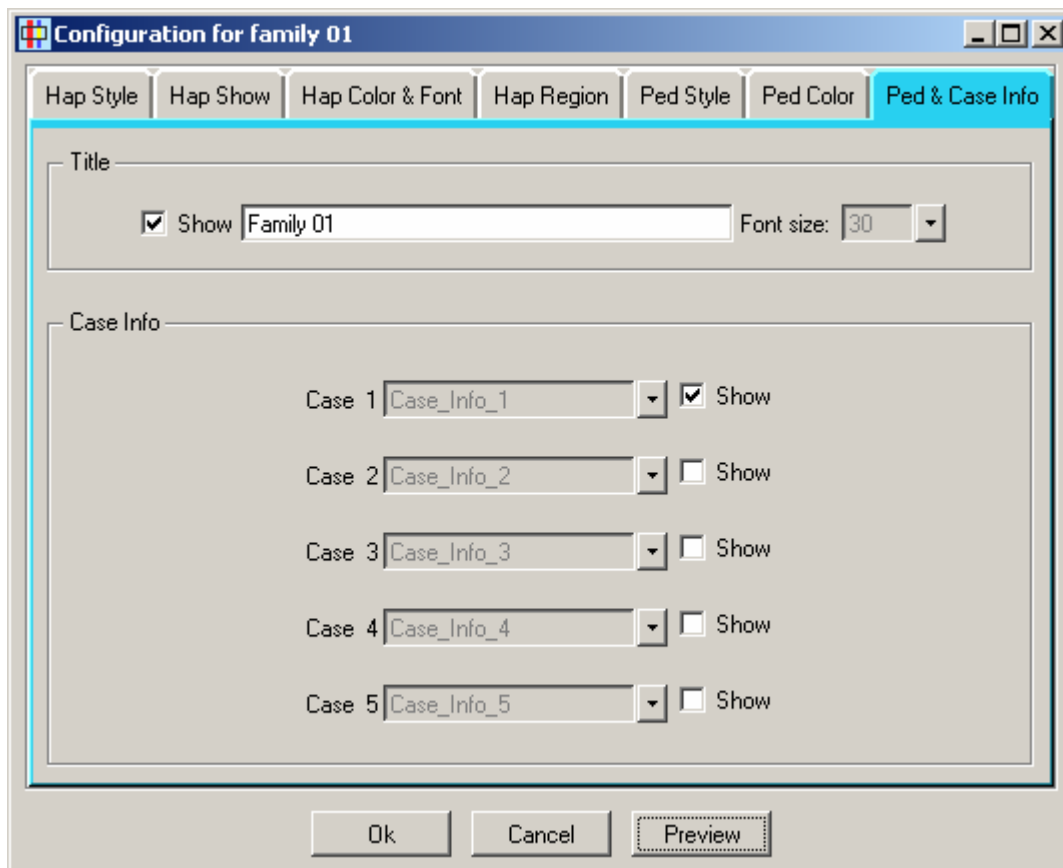
Page settings

The page settings are mostly relevant for exporting pedigrees as images. The graphic resolution is very important for the pixel formats PNG. For vector graphic formats this parameter is handled in different ways. SVG graphics uses it to define the scaling factor to which the graphic should be expanded. Postscript has a default resolution of 72dpi. HaploPainter tries to extend the pedigrees according to the Margin, Paper Size and Paper Orientation settings. This may not be what you want and cutting the image with other software is necessary.
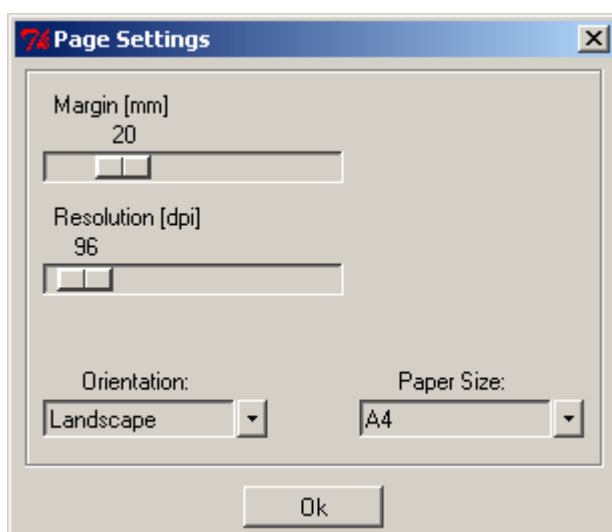
Show Grid

The grid is visible only in the graphical user interface. It never is exported into images! The grid (visible or not) is internally used as an anchor point for the positioning of symbols. The configuration parameters Grid X space and Grid Y space affect the size of it.

You can switch it on/off by the check button from the File menu or Context menu.

Storing and Restoring defaults

There is no special file format for storing of your favorite default parameters - instead HaploPainter is able to extract drawing style features from every regularly stored file in HAPLOPAINTER format! One way to go is to prepare a HAPLOPAINTER file with some small families and modify/store them. Now you can attach the default parameters from a certain family from that file to others by choosing the Open default … function from the File menu.

A window will appear asking you for a family and if you want to import the parameters from that family only to the current drawn family or to all families.

Command line mode

HaploPainter can be executed in a command line mode without usage of the graphical interface. The first argument must be **–b** to tell HaploPainter being in command line (**b**atch) mode or **–h** to output a list of supported parameters.

All following parameters have to be in a key – value manner. Allowed keys, their meaning and allowed values illustrates the following table.

-pedfile	Name of the imported pedigree file in ASCII or UTF-8 format
-pedformat	Format of the imported pedigree file **[linkage, csv]**
-outfile	Name of the exported file
-outformat	Format of the exported file **[ps, pdf, svg, png]**
-hapfile	Name of the imported haplotype file
-hapformat	Format of the imported haplotype file **[allegro, Genehunter, merlin, simwalk]**
-mapfile	Name of the imported map file
-mapformat	Format of the imported map file **[1, 2]**
-family	Name of one family identifier from the pedigree file
-resolution	Image resolution in pixels (max 4800)
-paper	Paper format **[A0-A5,B4,B5,Letter,Ledger,Legal,11x17]**
-orientation	Paper orientation **[landscape, portrait]**

-border	Margin around the image in pixels (max 999)
-bgcolor	Background color in hex notation (e.g. #ffcc00)
-dbtype	Database type **[oracle, postgresql, mysql]**
-dbport	Database port
-dbsid	Database Name/SID
-dbhost	Database server host name (localhost is valid)
-dbtable	Database Table/Relation Name
-dbuname	Database user/login name
-dbpasswd	Database password
-breakloop	Controls loop breaking **[0=No loop break, 2=Break all loops]**
-sortbyid	Drawing order of individuals is sorted by ID
-sortbygender	Drawing order of couples is sorted by gender
-h	Displaying a list of supported parameters