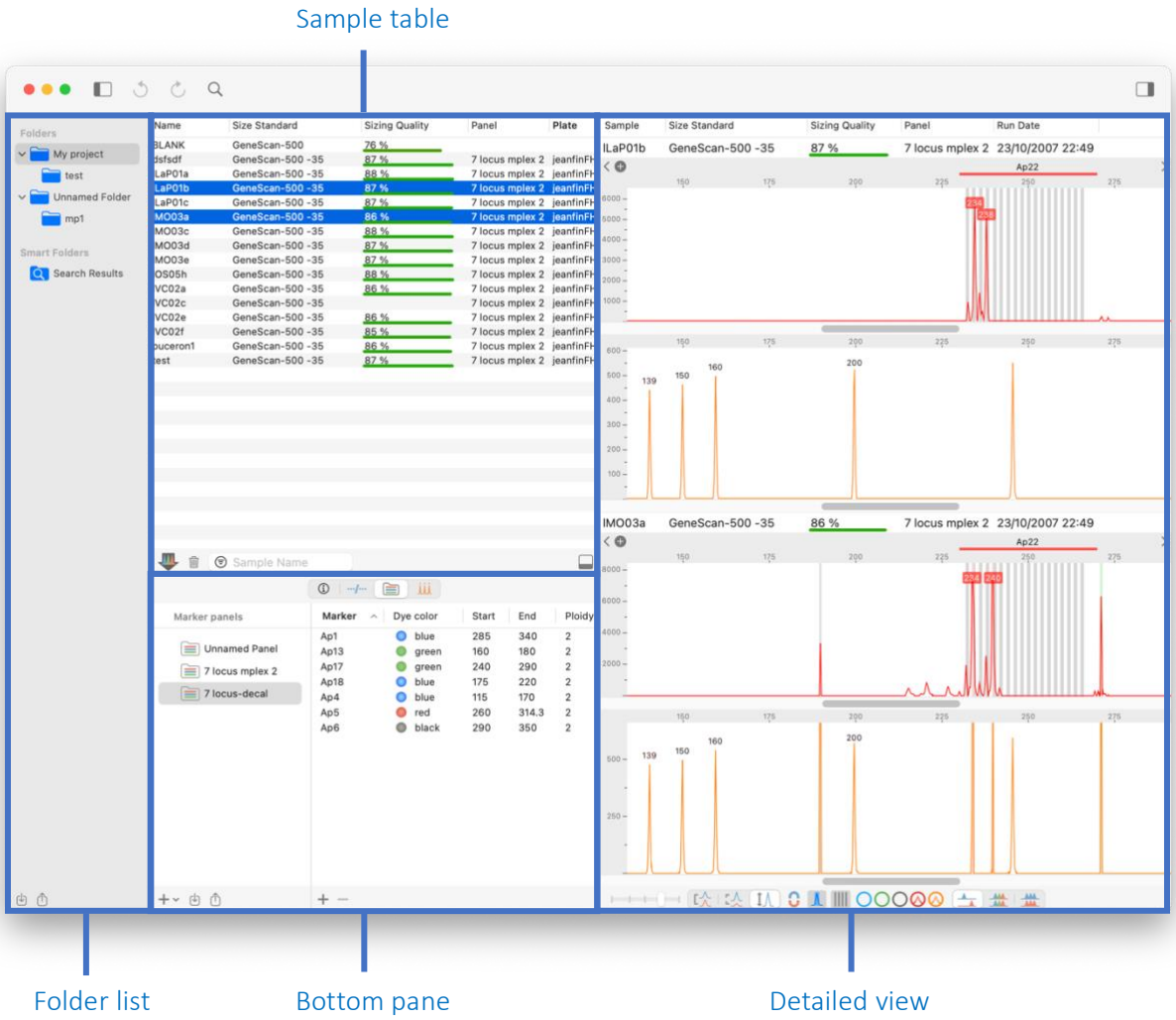
STRyper User Guide

Contents

STRyper interface

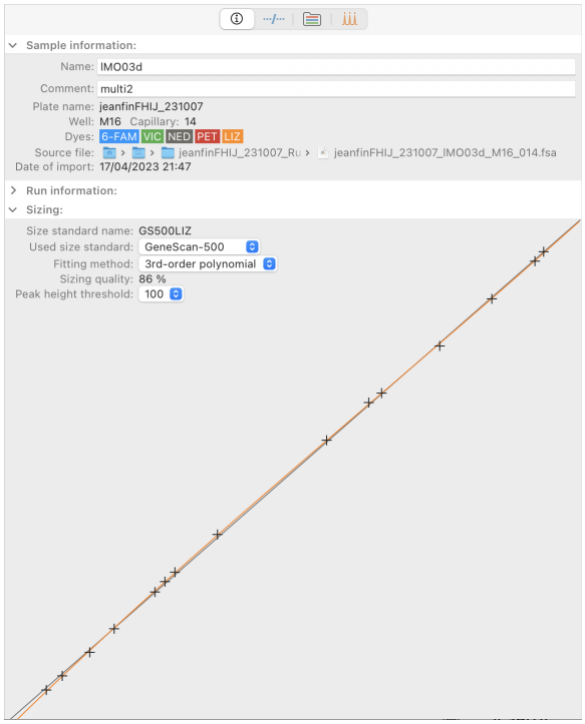
STRyper is an application that is designed to analyze files generated by capillary sequencers for microsatellite genotyping. The application main window comprises three sections.



The left sidebar contains the **folder list**.

The mid-section contains:

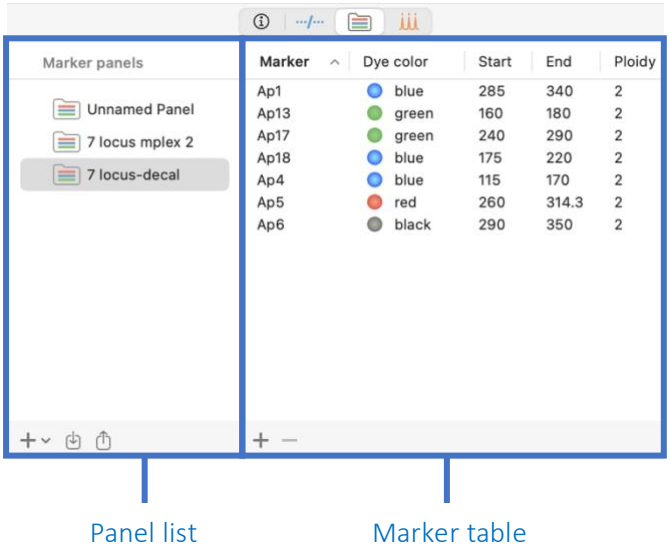
- **the sample table** listing the **samples** from the selected folder,
- **the bottom pane** with four **tabs**:
  - **the sample inspector** listing information about the selected sample(s),



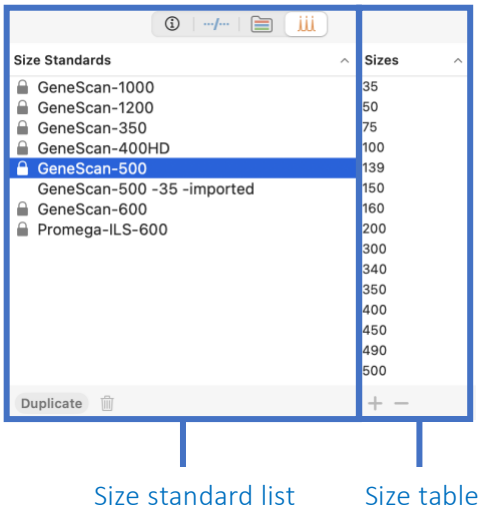
- **the genotype table**, listing genotypes for the sample shown in the sample table (if any):

Status	Sample	Panel	Marker	Size1	Size2	Allele1	Allele2	Offset
◉	4TV	7 locus mplex 2	Ap2	417.5	431.6	416	430	
◉	4TV	7 locus mplex 2	Ap2	417.4	431.6	416	430	
◉	Colmar1	7 locus mplex 2	Ap2	427.7	427.7	426	426	
◉	Colmar1	7 locus mplex 2	Ap2	427.7	427.7	426	426	
◉	JML06_1	7 locus mplex 2	Ap2	417.3	419.3	416	418	
◉	JML06_1	7 locus mplex 2	Ap2	417.4	419.4	416	418	
◉	JMLO06_2	7 locus mplex 2	Ap2	417.4	419.4	416	418	
◉	JMLO06_2	7 locus mplex 2	Ap2	417.4	419.4	416	418	
◉	L12	7 locus mplex 2	Ap2	417.4	427.7	416	426	
◉	L12	7 locus mplex 2	Ap2	417.5	427.7	416	426	
∅	L1_21_2	7 locus mplex 2	Ap2					
◉	L1_21_2	7 locus mplex 2	Ap2	429.6	431.3	428	430	
◉	L1_22	7 locus mplex 2	Ap2	417.4	427.8	416	426	
◉	L1_22	7 locus mplex 2	Ap2	417.5	427.8	416	426	
◉	L2_17_1	7 locus mplex 2	Ap2	417.5	429.7	416	428	
◉	L2_17_1	7 locus mplex 2	Ap2	417.2	429.5	416	428	
◉	L2_4V1_1	7 locus mplex 2	Ap2	417.3	417.3	416	416	
◉	L2_4V1_1	7 locus mplex 2	Ap2	417.4	417.4	416	416	

- **the marker tab**, listing **panels** and **markers** of the selected panel:



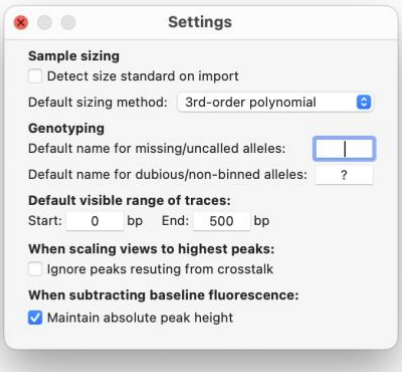
- **the size standard tab**, listing available **size standards** and the **sizes** of the selected size standard:



The right pane contains the detailed view of the selected samples, genotypes, or markers.

The folder list, the bottom pane and the right pane can be collapsed/expanded by pressing the collapse buttons , by dragging their separator, or via the View menu in the menu bar.

Additionally, STRyper has a **settings window** accessible from the menu bar via STRyper > Settings.



Managing samples

Sample folders

A **sample** refers to a chromatogram file that has been imported in the application.

A **folder** contains **samples** and/or other folders. You use folders to manage your samples into genotyping projects however you like. To create a folder, do one of the following:

- select File > New Folder from the menu bar,
- hover the mouse hover the "Folders" section of the folder list and click the ⊕ button:

To **add a subfolder to an existing folder** right-click the folder and choose Add Subfolder from the contextual menu.

Folders can be moved up or down, and into, or out of, parent folders.

To rename a folder, click its name or right-click the folder and select the **Rename** contextual menu.

Note: folders in the same parent folder cannot have the same name.

To delete a folder, you can:

- right-click it and select **Delete Folder** from the contextual menu,
- select it and choose **Edit > Delete Folder** from the menu bar.

Importing samples

STRyper imports files following the ABIF specification and ending with the ".fsa", ".FSA", ".hid" or ".HID" extensions. These files are imported as **samples** into **folders**.

Note: HID file support is experimental as the specifications of this format are not public.

To import samples, a folder must be selected. Then do one of the following:

- click the **Import Samples** button  at the bottom of the sample table,
- select **File > Import Samples…** from the menu bar,
- drag and drop FSA/HID files from the Finder into the sample table or onto a folder.

Note: you cannot import files by selecting or dropping their parent folder.

Copying samples and moving samples between folders

To copy samples, select samples from the sample table and select **Edit > Copy** from the menu bar, or right-click samples and select **Copy** from the contextual menu.

You can then **paste copied samples** via **Edit > Paste**.

Notes:

- Copying also copies the sample textual data shown in the table to the pasteboard. This data can be pasted to a text editor or a spreadsheet.
- Upon pasting, the copied samples are duplicated in their current state, not in the state they had when you copied them.

To move samples between folders, select samples and drag them to the desired folder in the folder list.

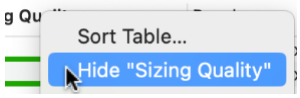
The sample table

Samples of the selected folder are listed in the sample table. The sample table has a default set of columns showing sample metadata, including sample name, plate, well, run date…

Name	Size Standard	Sizing Quality	Panel	Plate
A10				171218
LADDER	GeneScan-500	76 %		
IVC02f	GeneScan-500	84 %	7 locus mplex 2	jeanfinFHIJ_231007
IMO03a	GeneScan-500	85 %	7 locus mplex 2	jeanfinFHIJ_231007
IVC02a	GeneScan-500	85 %	7 locus mplex 2	jeanfinFHIJ_231007
IVC02e	GeneScan-500	85 %	7 locus mplex 2	jeanfinFHIJ_231007
IVC02b	GeneScan-500	85 %	7 locus mplex 2	jeanfinFHIJ_231007
IMO03d	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
ILaP01c	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
IMO03e	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
IVC02d	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
IOR03	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
ILaP01b	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
IMO03c	GeneScan-500	86 %	7 locus mplex 2	jeanfinFHIJ_231007
IOS05h	GeneScan-500	87 %	7 locus mplex 2	jeanfinFHIJ_231007
ILaP01a	GeneScan-500	87 %	7 locus mplex 2	jeanfinFHIJ_231007

To copy the text content of the table, simply select rows, then **Edit > Copy** from the menu bar. The copied content can be pasted to any text editor or spreadsheet application.

To show/hide columns, right click the table header and check/uncheck the columns you want to show/hide. Right-clicking a column header displays a contextual menu item allowing to hide that column.

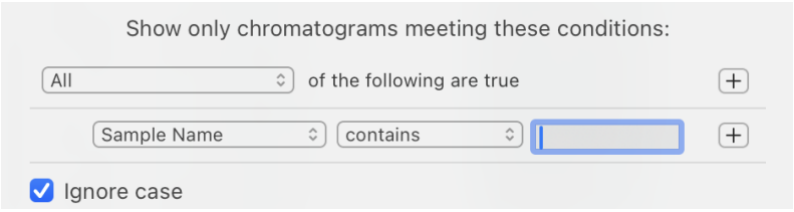


To sort the table by a column, simply click its header.

To sort the table by several columns, right-click the table header and select **Sort Table...** from the contextual menu. This brings a popover that lets you may change the priority of columns by dragging rows.

You may change the sorting order (ascending descending) by clicking the button with a chevron point down or up, and add remove sorting criteria using the "+" and "–" buttons: 

To filter samples to only show those meeting certain criteria, press the "filter" button below the sample table. A popover will allow you to define filter criteria.



Note:

- a filter on samples is associated to a folder. If a folder has a filter applied to it, the "filter" button changes appearance:
- a smart folder (see below) cannot be filtered. You may instead modify its search criteria.
- A filter on samples is not dynamic. If samples are modified so that they meet or no longer meet filter criteria, you must reapply the filter, or select another folder and go back to the previous selection, to update the sample table.

Sample search and smart folders

To find samples according to different criteria among all samples of the database, press the search button on the window toolbar or select **File > Find Samples...** from the menu bar.

Results appear in a special folder called a **Smart Folder**. **A new smart folder is created for each search**.

Smart folders in STRyper work similarly as smart folders in the Finder, which are created after a spotlight search is saved. This means that samples meeting the search criteria will automatically be listed in the smart folder. For some search criteria however (folder name, size standard, panel), the search results may not update immediately and will only update once the application saves its database (which occurs at least every 30 seconds).

You may also initiate a search by **adding a new smart folder**. To do so, hover the "Smart Folders" section in the folder list and click the ⊕ button: 

A smart folder can be renamed and moved just like a regular folder. However, you cannot drop any item (folder, samples) into a smart folder.

A smart folder is preserved between app launches, which means you **must delete a smart folder to discard search results**.

However, you need not create a new smart folder every time you want to find samples. **To modify the search criteria of an existing smart folder**, do one of the following:

- select the smart folder and click the button at the below the sample table,
- right-click the smart folder and select Modify Search Criteria from the contextual menu,
- select the smart folder and select Edit > Edit Smart Folder… from the menu bar.

A smart folder does *not* contain copies of samples present in other folders. It just gathers samples meeting the search criteria, which means that **deleting samples from a smart folder removes them from the database**.

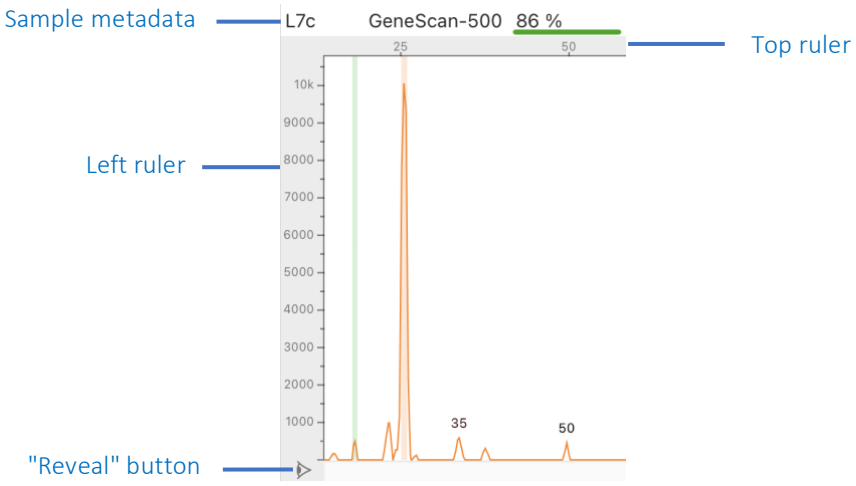
To reveal samples from smart folder in their original folder, right-click the sample(s) and select Reveal in Parent Folder from the contextual menu. **Note**: this item is disabled if clicked samples belong to multiple folders.

Viewing samples

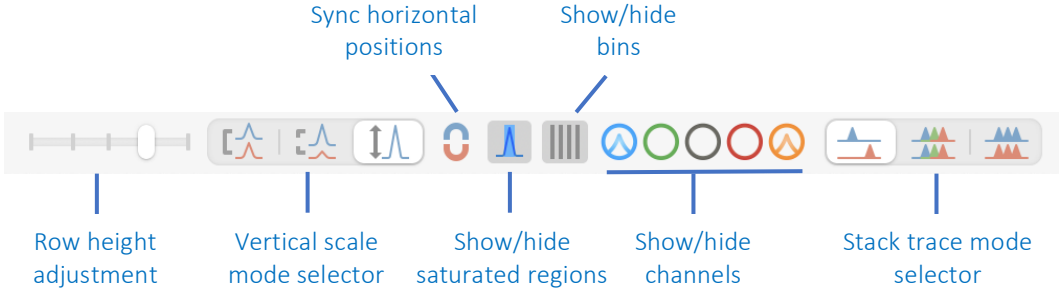
The detailed view

Selected samples are immediately displayed in the detailed view on the right pane.

If samples are not stacked, the detailed view shows sample metadata similarly to the sample table (with the same options for hiding columns), as well as rows showing **traces** (fluorescence curves).



When viewing samples, the detailed view shows an array of buttons at its bottom:



Sometimes, a sample or a genotype shown on the detail view may not be visible in the sample table or genotype table if the table has been scrolled. **To reveal a sample or genotype in its source table** click the Reveal Button ▷ at the bottom left of the traces.

Adjusting row height

The visible area of the detailed view can fit one to five rows showing traces. This number is set via row height adjustment slider.

Showing/hiding dyes

STRyper manages chromatogram file containing traces from four to five **channels** (colors).

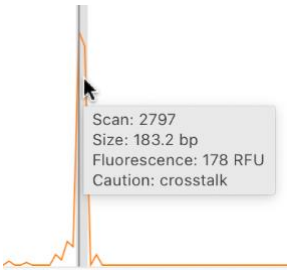
To show or hide a channel, click the buttons with the corresponding color:

To show a channel and hide all others at the same time, click the colored button with the option key (⎇) pressed.

Showing peak information

STRyper detects peaks in traces. If the detailed view shows a single trace per row, hovering a peak with the mouse brings a tooltip indicating basic information about the peak:

- "Scan" is the number of data points that were recorded by the sequencer up to the peak tip (represented by a vertical line). It represents the time since the start of the fluorescence recording.
- "Size" is the estimated size of the DNA fragment that caused the peak and is inferred from the molecular ladder,
- "Fluorescence" is the fluorescence intensity at the peak tip in Relative Fluorescence Units (RFU).
- "Caution: crosstalk" indicates that the peak likely results from crosstalk, i.e., it does not represent a DNA fragment from the channel that is visualized, but from another channel.



To show/hide peak information, select/deselect **Traces > Peak Information** from the menu bar.

Adding peaks manually

STRyper does not detect peaks whose fluorescence level is less than 100.

To add a peak to a trace in a region where no peak is detected, you can:

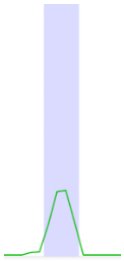
- right-click below the tip of an apparent peak and select **Add Peak here** from the contextual menu.
- If you use a trackpad with "force click" enabled, perform a force click (deep press) below the tip of the apparent peak.

Either action will add the peak at the clicked location. If the click location does not appear to correspond to a peak, or if there is already a peak at the clicked point, the action will have no effect.

Showing saturated regions

To show/hide **regions where the fluorescence has saturated the sequencer's camera**, press the Show/hide saturated regions button .

Saturated regions are represented by rectangles which helps to identify regions with a risk of **crosstalk** between dyes.



The color of the rectangle representing a saturated region corresponds to the dye that likely saturated the camera.

Note: the dye causing saturation is inferred by STRyper. This piece of information is not encoded in the chromatogram file.

Subtracting baseline fluorescence level

STRyper automatically subtract the baseline fluorescence level to facilitate the interpretation of peaks.

To show raw fluorescence data select **Traces > Subtract Baseline Fluorescence Level** from the menu bar such and make sure this menu item is disabled.

By default, the absolute height of peaks is not affected by this operation, which increases peak heights relative to the baseline. If you are not pleased with the result, uncheck the "Maintain absolute peak heights" checkbox in the application settings.

Note: curves are not smoothed by the application.

Stacking curves in the same row

To show all traces in separate rows (which is the default), press the left segment  of the stack trace mode selector.

To stack channels from each sample in a single row, press the middle segment  of the selector.

To stack traces from several samples in the same row, press the right segment  of the selector. In this mode, sample metadata is not shown since each row can show several samples, and the header of the detailed view is replaced by a text indicating the number of samples stacked.

Note: a maximum of 400 traces per channel can be shown stacked in the detailed view.

Adjusting vertical scale

The vertical scale, in relative fluorescence units (RFU), appears on the left ruler.

To adjust the vertical scale hence change the height of curves, click and drag on the left ruler.

To adjust the scale such that the tip of the highest visible peak appears near the top of the row, double-click the left ruler.

Note: if no peak was found in the trace(s), the vertical scale is adjusted to the highest fluorescence level of the trace(s) shown in the row.

There are three modes to adjust the vertical scale of rows:

To make all rows use the same vertical scale, activate the left button  of the vertical scale mode selector or select **Traces > Vertical Scales > Synchronized** from the menu bar.

To adjust the vertical independently for each row, activate the middle button  of the selector or select **Traces > Vertical Scales > Independent** from the menu bar.

To make vertical scale adjust automatically to the highest visible peak, activate the right button  of the selector or select **Traces > Vertical Scales > Scale to Highest Peaks** from the menu bar. In this mode, you can still adjust the vertical scale by clicking and dragging the left ruler, but the row will rescale

automatically upon scrolling. By default, this mode ignores peaks resulting from crosstalk (saturation in another channel). This can be changed in the application settings.

Horizontal positioning and zooming

For samples that have been sized with a size standard, a horizontal scale in base pairs (bp) is shown on the top ruler. The current position of the cursor also appears on the ruler. If no size standard is applied or if sample sizing failed, no scale is shown and the text “No size standard applied” or “Sample sizing failed” shows instead.

The default visible range of traces (by default, 0 to 550 base pairs) can be specified in the application settings.

Note: the range must be at least 2 base-pair wide and is constrained to 0 – 1200 base pairs.

To zoom in/out, do one of the following:

- Place the mouse pointer over the traces and either:
  - pinch or double-tap on the trackpad (double-tap zooms in and does not zoom out),
  - scroll vertically with the option key (⌥) pressed,
- cover the aera you wish to zoom to by clicking and dragging on the top ruler.

To zoom to the default range defined in the application settings**,** double-click the top ruler, or right-click it and select Zoom to Default Range from the contextual menu.

To synchronize the zoom scale and position across rows, press the Sync horizontal positions button such that it shows this symbol: or select Traces > Synchronize Positions from the menu bar.

Note: positions are not synchronized upon activating the setting, but upon adjusting the zoom or position of a particular row after the setting is active.

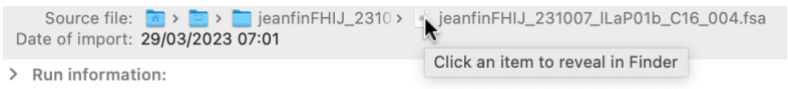
The sample inspector

The sample inspector displays information about selected samples, including sample names, comments, dye names, plate, well, run information and sizing information. These pieces of information are organized in sections.

To **show the sample inspector**, click the left button of the segmented control ⓘ of the bottom pane or select View > Sample Inspector from the menu bar.

To **collapse or expand a section**, click the triangle next to the section title.

Tip: to reveal the source file of a selected sample in the Finder, click the file icon in the “Sample information” section at the “Source file” field.



Sample sizing

Applying a size standard to samples

To compute sizes in base pairs, a sample must contain fluorescence data from a molecular ladder defined in a **size standard**. The name of this standard may be encoded in the FSA file if this piece of information was

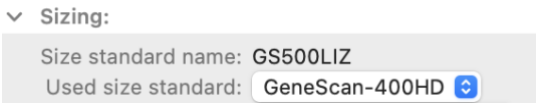
specified in the sequencing application. STRyper can use this name and apply a size standard whose name matches if the "Detect size standard on import" option is enabled in the application settings.

If no suitable size standard name was specified or found, an imported sample is not sized and cannot yet be genotyped. Sizes in base pairs do not show in the top ruler above traces. You must apply a suitable size standard manually.

Available size standards are listed in the size standard library, which you access by clicking the ⅉ button of segmented control of the bottom pane or by selecting View > Size Standard Library in the menu bar.

To apply a size standard to samples, do one of the following:

- From the size standard list, drag and drop a size standard onto the sample table. This will apply the size standard to all samples from the table.
- Right-click samples and chose the size standard from the **Apply Size Standard** contextual menu. This will apply the size standard to all clicked/selected samples.
- Go to the sample inspector, at the "Sizing" section and select an item from the "Used size standard" popup button. This will apply the size standard to all selected samples:

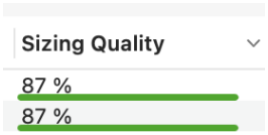


Applying a size standard triggers the detection of peaks in the appropriate dye and their assignment to fragment sizes listed in the standard. This occurs even if the size standard is the same as that already applied to samples.

The default fitting method specified in the application settings is used to compute the size in base pairs for each **scan number** (time when each fluorescence data point was recorded).

Checking sizing quality

After applying a size standard, and index of **sizing quality** from 0 to 100% appears in the Sizing Quality column of the sample table (and the detailed view) if this column is not hidden. Sizing quality is represented by a gauge whose length relative to the column is proportional to the quality. The color of the gauge varies from red to green depending on the quality.



This quality index is based on comparing the size computed for each ladder fragment to its true size, as defined in the size standard. The difference between both sizes is referred to as "offset". This offset is then compared between adjacent fragments. The absolute difference is offsets between nearby fragments is inversely proportional to the sizing quality.

Note: the sizing quality score is relative, note absolute. Lower quality scores than other samples of the same run may indicate errors in size assignment or electrophoretic problems, but **high scores do not guarantee the absence of error**. Molecular ladders must be checked visually even if scores are high.

Sizing quality can be evaluated by checking the fitting curve.

Why did sizing fail?

Sample sizing fails if not enough peaks of the molecular ladder could be assigned to sizes of the size standard. This may result from a fluorescence signal that is too weak for peak detection, electrophoretic problems, or application of an inappropriate size standard. This failure is indicated by:

- a sizing quality of 0%,
- a "Sizing failed" indication in the "Sizing" section of the sample inspector, in lieu of the fitting curve,
- "Sample sizing failed" indicated on the top ruler.

Note: no panel of markers can be visualized on a sample whose sizing failed.

If you suspect that sizing failed because the fluorescence signal is too weak, you may lower the **peak height threshold** in the Sizing section of the sample inspector. This threshold corresponds to the minimal fluorescence level a peak must have to be detected.

The default value is 100. Changing this setting will affect all selected samples and trigger the detection of peaks in the ladder for these samples.

The fitting curve

To show the **fitting curve** of a selected sample, go to the sample inspector and display the "Sizing" section. No curve is visible if no size standard is applied to a sample, if sizing failed or if several samples are selected.

The fitting curve shows the relationship between the time at which DNA fragments (peaks) of the molecular ladder were detected by the sequencer's camera (its scan number), representing the X axis, and their attributed sizes in base pairs, representing the Y axis. These peaks are represented by crosses.

The fitting curve shows the relationship between the scan number and the size in base pairs. This relationship is used to size the sample and is established by fitting a curve to the points (crosses) shown on the plot.

Ideally, the **curve should pass as close as possible to every point**. Points departing noticeably from the curve may indicate that a size was not attributed to the correct fragment of the ladder. In this case, ladder peaks may be corrected or ignored (see next section).

Choosing the appropriate fitting method

For each scan number (fluorescence data point), the size in base pairs is computed according to a **fitting method**. Three fitting methods are available:

- **The Linear regression** (1st order polynomial) method assumes a linear relationship between scan number (n) and size (s): $s = an + b$.
- **The 2nd order polynomial** method assumes a relationship of the form $s = an^2 + bn + c$.
- **The 3rd order polynomial** method assumes a relationship of the form = $an^3 + bn^2 + cn + d$.

The coefficients *a*, *b*, *c*, *d* are computed by solving a system of equations by the Cholesky decomposition method.

To apply a fitting method to samples, do one of the following:

- right-click samples from the sample tables and choose a fitting method in the **Fitting Method** contextual menu.
- Select samples from the sample table, go to the sample inspector, "Sizing" section and choose a value from the "Fitting method" popup button. 

The default fitting method can be defined in the application settings.

Which fitting method to choose depends on the user-evaluated consistency between sizing of different samples. To evaluate how much sample sizing deviates from the linear relationship between scan number and size you may inspect the fitting curve and check that it goes through all points.

Using higher-number polynomials is recommended as these better consider variations and anomalies in the electrophoresis efficiency. Anomalies have a more local influence on the sizing than with the linear regression method.

You must use the same fitting method for all samples analyzed at the same marker panel, regardless of the run. This is because different fitting methods yield sizes differing by several base pairs for the same allele.

A convenient way **to evaluate the consistency of a fitting method** is to stack samples of the same run/plate in the same row and to check how much peaks (from the ladder or from markers) superimpose. You may then choose the fitting method for which peaks superimpose the best.

If peaks of certain samples are off, this may indicate an error in the size attributed to peaks in these samples, in which case peaks of the ladder should be edited.

Checking and editing peaks in the ladder

It is highly recommended to check for errors in the assignment of peaks to sizes in the ladder, for each sample, regardless of the sizing quality index.

To **show sizes attributed to ladder peaks**, select samples and display the dye corresponding to the ladder. Make sure show a single trace per row. The size attributed to each ladder peak is shown as a label above the peak. Sizes that are not assigned to peaks (because no suitable peaks were found) appear in grey at the top of the row.

To assign a peak to a (different) size, grab the corresponding size label and drop it onto the peak. If the peak is already assigned to a size, this size will be replaced by the new size and will become unassigned.

Note: the range of peaks that can take a given size is restricted to ensures that sizes remain in ascending order from left to right.

To remove a size from the sizing, you can either:

- select the corresponding size label and hit the delete key,
- select the corresponding size label and **Edit > Remove Ladder Size** from the menu bar,
- double click the corresponding size label or the peak underneath,
- drag the corresponding size label above the row showing traces.

After each change, the sizing is automatically recomputed according to the fitting method applied to the sample. If the sample is the only one selected in the sample table, each change updates the fitting curve.

Tip: a quick way to check for errors in size assignments, or to assess if the chosen fitting method is the best, stack all samples in the same row and check whether peaks of the ladder superimpose neatly. If certain peaks appear off, as in the example on the right, this may indicate error in peak assignment for the corresponding sample(s).

Defining custom size standards

STRyper includes the specifications of several commercial size standards. These cannot be modified nor deleted.

To create a custom size standard, you must duplicate an existing one with the Duplicate button below the size standard library.

Then, sizes can be added, removed, or changed in the size table.

To remove a size standard size, select the size and click the − button below the table.

To add a new size standard size, select the desired size standard add click the + button below the size table.

To change the value of size standard size, select its cell and enter a new value.

Note: a size must be at least 20 bp and cannot exceed 1500 bp. A size standard must have at least four sizes and cannot have duplicate sizes.

Genotyping

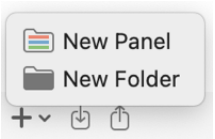
Genotyping requires defining **markers**, which are organized into **panels**. A panel corresponds to a multiplex of markers amplified in the same PCR.

Panels are listed in the marker library, which is accessible by clicking the appropriate button on the segmented control:  or selecting View > Marker Library in the menu bar.

Defining marker panels

If STRyper does not find any panel in the database, it creates one upon launch. This panel is called “New Panel” and contains no maker.

Like samples, panels can be organized into folders. **To create a new panel or folder**, click the + button below the panel library. Panels and folders can be renamed, deleted, and dragged like sample folders.



To apply a panel to samples, do one of the following:

- Drag the panel from the panel list onto the sample table. The panel will be applied to all samples.
- Right-click samples from the sample table and choose the panel from the Apply Marker Panel contextual menu. The panel will be applied to clicked/selected samples.

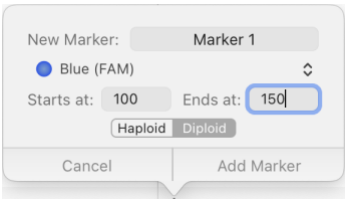
If the applied panel has **markers**, a blank **genotype** will be added to each sample for each marker upon applying the panel. The generated genotypes appear in the genotype table.

Defining markers

A **marker** defines the channel (color) of the fluorescent dye that labels the primers used to amplify the marker, and the range (in base pairs) where alleles are expected, as well as the ploidy of the locus. A marker also has a name.

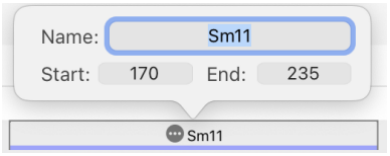
To add a marker to a panel, do one of the following:

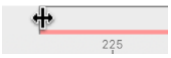
- select a panel from the panel list and click the + button below the marker table. A popover will let you define the name, the color (dye), the range and the ploidy of the new marker.
- On the detailed view showing the appropriate channel of a sample that has the panel applied to it, click the ⊕ button at the top left of the trace. Then, click and drag on the area where the mouse cursor shows a green '+' sign  to define the range of the new marker. Once you release the mouse, the popover will let you define the marker, except for the channel.



Note: **the ploidy and channel of a marker cannot be changed after its creation**. You may however **alter the range and name of a marker** at any time. To do so, perform one of the following:

- edit the corresponding cell for the "name,", "start" or "end" column of the marker table.
- double-click a marker label (rectangle) above a trace and enter new values in the popover.
- click the ••• button on the marker label, or right-click the label, and select **Edit Name and Range** from the contextual menu, which spawns the popover.



To resize a marker, you must first click it. You may then click and drag its edges. 

To navigate between markers of the same channel in the detailed view, you can:

- click a navigation button < > above the top ruler,
- place the cursor on the top ruler or the area showing markers and scroll left or right. Note: this requires a trackpad or an Apple mouse. This can be disabled by disabling **Traces > Allow Swipe between Markers** from the menu bar.
- perform a three-finger swipe, if this gesture is enabled in the trackpad preferences of macOS.

Note: the < button will move to the marker to the right of the visible area if there are no marker on the left, and the > button will move to the marker on the left if there are no marker on the right.

To zoom the view to the range of a marker, click the action button ••• on the marker label, or right-click it, and select **Zoom to Marker** from the contextual menu.

Defining bins

Once a marker is defined, STRyper can look for alleles by scanning peaks in the marker range. However, to attribute peaks to known alleles, **bins** must be defined. A bin corresponds to a size range where an allele is expected. A peak detected within a bin is attributed to the allele that has the bin's name.

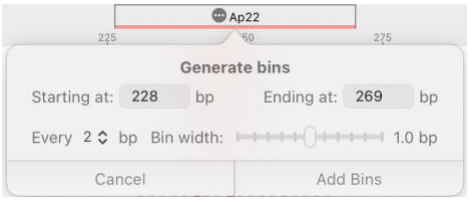
Like markers, bins have start and end coordinates in base pairs, and a name.

To generate bins for a marker, do one of the following:

- If the marker's panel is applied to samples, select one of these samples. Scroll the trace to show the range of the marker.

- Select a marker in the marker table. The detailed view shows the whole marker's range. However, this does not show traces and therefore does not suggest where to place bins appropriately.

Then, click the action button on the marker label, or right-click it, and select Generate Bins from the contextual menu. A popover will let you define the set of bins to add.



The spacing between bins should represent the motive length of the marker.

Note: generating bins removes existing bins for the marker.

Bins are represented as vertical rectangles whose top section shows their names.

Note: if bins are too narrow, their name is hidden. You may have to zoom in to see bin names.

Upon generating bins, the marker enters a mode allowing you to move bins.

Most of the time, bins generated this way do not coincide well with the marker's alleles. This is because an amplicon and a ladder DNA fragment of the same lengths can migrate at different speeds during electrophoresis. You may need to move the bins. **Note:** an alternative is to define an offset for the marker.

To move all bins of a marker together click the action button on the marker label, or right-click it, and select Move Bins from the contextual menu (this is not needed it you have just generated bins). The action button turns into a checkbox, the background of the trace in the marker's range takes a light pink color, and the mouse cursor becomes an open hand over this area to signify that bins can be moved. Genotypes cannot be edited, and saturated regions and peak information cannot be displayed while this area is visible. This affects all rows in which the marker shows. You may then move the bin set by dragging the pink area. A green area will appear underneath to represent the allowed range of the bin set, which ensures that bins remain in the marker's range.

Upon clicking the pink area, a red vertical line with two converging black triangles appears at the click location. This line represents the anchor position (in base pairs) around which you may shrink/expand the bin set (see below). This position will not move during this process.

You may **expand/shrink the bin set** by dragging the left or right edge of the pink area. A green area and a white vertical line will appear to denote the allowed range that the edge can take. This range ensures that bins remain in the marker's range and that the bin set is not expanded/shrunk excessively.

Note: expanding/shrinking the bin set changes the bin width proportionally to avoid overlap between bins.

Sometimes, you may need to add/edit bins individually. **To add/edit individual bins,** you can:

- click the action button on the marker label, or right-click it, and select Edit Individual Bins.
- If a bin in already present and if you use a trackpad with force touch enabled, perform a force click on the bin. This will select the clicked bin.

After either action, the action button is replaced by a checkbox, the background of the trace in the marker's range takes a light pink color, and the mouse cursor takes a green '+' sign over this area to signify that bins can be added. Genotypes cannot be edited, and saturated regions and peak information cannot be displayed while this area is visible. This affects all rows in which the marker shows.

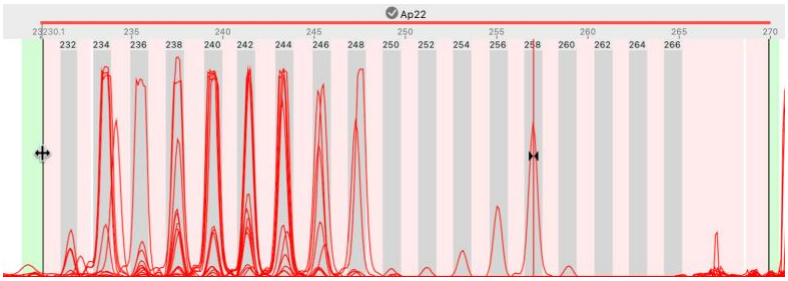
To add a bin, click and drag on the area where the mouse cursor shows a green '+' sign  to define the range of the new bin. The bin is effectively added upon releasing the mouse button, and takes a default name based on its size range.

To change the name and range of a bin, double-click it, which spawns a popover allowing to specify a name and range.

To resize a bin, you must first click the bin so its edges can be dragged. You may also **move a bin** by simply dragging it.

To exit bin editing, you can either: hit the escape key, click the checkbox on the marker label or anywhere on the trace outside the pink area.

Tip: to edit/move bins, you may stack many samples (preferably from the same run) in the same row and zoom to the marker's range. This way, you will see peaks in the same view, hence where bins should be. Peaks resulting from **stuttering** indicate the probable location of other alleles and can be used to add/adjust bins. You may then move the bin set such that the appropriate bin sits right behind the peak that represents the longest allele. If other bins are not located property with respect to peaks, position the anchor point at the middle of the bin that is behind the rightmost peak (simply by clicking at that location), and move the left edge of the pink area such that bin locations coincide with peak locations.



Note: certain alleles may have odd sizes due to mutations that do not correspond to indels of microsatellite motives. For such alleles, you may add specific bins.

Hiding bins

You may wish to **hide bins** to make peaks more visible. To do so, click the Show/hide bins button or deselect **Traces** > **Bins** from the menu bar.

Note: bins that are being edited or moved cannot be hidden.

Defining marker offsets

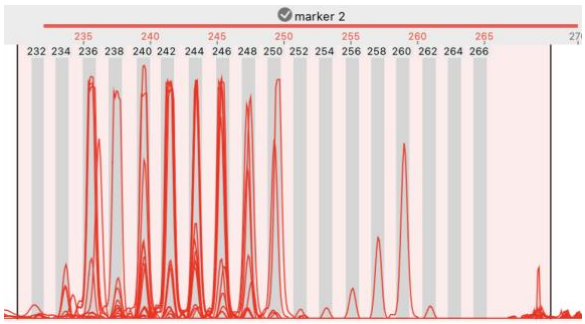
Amplicons and ladder DNA fragments often react differently to variations in electrophoretic conditions. Hence, the computed size of the same allele may slightly vary between runs, even on the same sequencer. Such variations can shift the location of alleles with respect to bins, defeating the purpose of bins. Moving bins as described in the previous section does not address the issue. A bin cannot be so wide as to cover the various locations of the same allele sequenced under different conditions.

STRyper's solution to this issue is to **adjust the offset of the marker** for specific samples, that is, to alter the computed sizes of peaks within the marker range for these samples without modifying the marker bins.

A **marker offset** is composed of two numbers that are terms a and b of the formula $y = a + bx$, where x is the size of a peak computed thanks to the molecular ladder, and y is the actual size used in the marker's range. If there is no offset, $a = 0$ and $b = 1$. b is constrained to the interval [0.9, 1.1]. Thankfully, you do not have to specify these numbers.

To **adjust the offset of a marker,** select one our several sample(s) of known genotype from the sample table. Then select **Adjust Offset** from the marker contextual menu (accessed via right-clicking the marker or by clicking the action button). A submenu will allow you to define the samples for which the offset will be applied. Ideally, these samples should be sequenced in similar conditions. Then, move bins such that their positions match those of peaks. Changing the offset does not modify bins. Indeed, sizes of the top ruler

move in sync with bins (and become red to denote the offset). The marker offset is represented by the difference between its range shown above traces and the pink area behind traces.



To exit offset adjustment, you can either hit the escape key, click the checkbox on the marker label or anywhere on the trace outside the pink area, or select another sample to show.

Because a marker offset is specific to certain samples, it is an attribute of the samples' *genotypes* for the marker. This offset therefore appears in a dedicated column of the genotype table. The numbers shown in parentheses are terms *a* and *b* of the offset.

Offset
(6.2, 0.978)
(6.2, 0.978)
(6.2, 0.978)
(6.2, 0.978)

To reset (remove) **offsets applied to genotypes**, select these genotypes on the genotype table, right-click the selection and select **Reset Offset(s)** from the contextual menu. You may alternatively select one or several samples in the sample table, click the action button on the marker label above traces, or right-click it, and select **Reset Offset.** A submenu will allow you to define the samples for which the offset will be reset.

Note:

- If samples with different offsets for the same marker are stacked in the same row, the offset that is represented by the position of bins and by sizes on the top ruler will be that of a random sample in the selection.
- Peaks that are outside the marker range (shown above traces) cannot be considered as alleles even if they are covered by the pink area representing the marker range with its offset. This is because this area may overlap that of another marker (see below).
- An offset may make bins appear outside their marker's range. Bins from different markers that are very close to each other may thus visually overlap. To avoid it, leave sufficient space between bins of different markers of the same dye. If overlapping bins do correspond to existing alleles, you should change amplicons sizes such that alleles can be unambiguously attributed to the correct marker.
- When you move bins of a marker that has an offset for the sample(s) shown, the marker temporarily resets its offset until you are done, so as to present its original state.

Copy markers across panels

To copy markers and all their bins from a panel to another, you may:

- drag markers from the marker table to another panel from the panel list,
- select markers from the marker table and select **Edit > Copy** from the menu bar. Then, select another panel from the panel list and do **Edit > Paste**.

Note: either action is prevented if exiting markers in the destination panel have the same names as, or overlap with, those that you want to copy.

Exporting and importing marker panels

You may wish **to export a panel to a text file** for archiving, editing some properties with a text editor, or transferring the panel to colleagues. For this, do one of the following:

- select the desired panel and click the button below the panel list,
- select the desired panel and select File > Export Panel…
- right-click the panel and select Export to Text File… from the contextual menu.

The exported file is a UTF-16 **tab-delimited** text file as in this example:

```
panel	new_panel
marker	marker1	154.9276	180.7307	green	2
bin	167	166.7578	167.4044
bin	164	163.5738	164.4571
bin	165	164.7016	165.4428
bin	163	162.6041	163.3247
bin	166	165.6051	166.4694
bin	168	167.647	168.6337
bin	162	161.5599	162.3508
bin	161	160.5403	161.3174
bin	176	176.033	176.9197
marker	marker2	242	275.6552	green	2
bin	253	252.3192	253.0365
bin	255	254.1426	255.0561
bin	270	269.4419	270.4216
```

The first line starts with the keyword “panel” followed by the panel name.

The second line defines the first marker. The line starts with the keyword “marker”. The following fields represent in this order: the marker name, the start coordinate, the end coordinate, the channel color (either “blue”, “green”, “red”, “black” or “orange”) and the ploidy (either 1 or 2).

The subsequent lines define the marker bins. Bins of a marker, and markers of a panel, can be described in any order, but all bins of a marker must be described right after the marker, and before the next marker. A row describing a bin must start with the keyword “bin”. The following fields specify its name, start and end coordinates.

Decimal separators must be periods.

Note: STRyper allows most Unicode characters (excluding tabs) in panel, marker and bin names. However, for compatibility with other applications, roman characters are recommended.

To import a panel from a text file conforming to the above specifications, do one of the following:

- click the import button below the panel list,
- right-click a panel folder and select Import Panel from File… from the contextual menu. The panel will be imported into that folder.

Upon importing, any blank line from the text file is skipped, leading and trailing spaces are removed from fields. If the imported file does not conform to the specifications, or if markers/bins overlap, are too short, or exceed the allowed range, an error message is thrown describing the issue.

Note: the text file to import can only describe a single panel. Import will fail if the keyword “panel” is found at the first column of any other line than the first one.

Allele calling

Alleles can be called for samples having a panel containing at least one marker. **Allele calling** consists in detecting peaks that represent alleles.

A **genotype** is a set of one or two alleles for a given sample at a marker. Each allele has a computed size in base pairs (which may not be an integer number), and a name. If the marker has no bin, alleles will get the names the “out of bin” alleles (by default, a question mark).

The genotype table lists the genotypes of all samples shown in the sample table. You can reorder, hide columns, and sort the genotype table like you can for the sample table.

If no genotype shows for a given sample at a marker, it means that the sample does not have a panel with this marker applied to it.

A genotype can have different statuses represented by icons in the **status column**:

○: the genotype has not been called nor edited. Its alleles have no size and have the name given to missing alleles, which is blank by default.
◉ : the alleles have been called by the application,
∅: no peak was found in the marker range during allele call,
: alleles have been edited by the user,
●: the marker has been modified (range, bins) since alleles were called/edited,
⚠ : sample sizing (marker offset, molecular ladder, fitting method) has been modified since alleles were called/edited. You may need to check the genotype.

Note: genotype status does not represent genotype quality.

To call alleles, do one of the following:

- Click the Call Alleles button ⅄⅄ below the genotype table. This calls alleles for all genotypes in the table for which alleles have not been called yet.
- Select and right-click genotypes from the genotype table and select **Call Alleles** from the contextual menu. This will call alleles for these genotypes, even those for which alleles have already been called or edited (that is, alleles will be re-called).
- Select and right-click samples from the sample table and select **Call Alleles** from the contextual menu (these samples must have panel applied to them). This action will call alleles for the genotypes at all markers applied to these samples.

During alleles calling, peaks are automatically attributed to alleles. Alleles from peaks whose tips are within bins take the bin names. Alleles outside bins take the name attributed to dubious our out-of-bin alleles. By default, this name is a question mark "?".

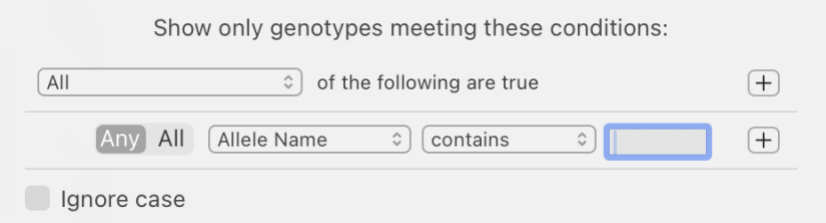
To modify the default names for non-binned or missing alleles, display the application settings and enter the character chain of your choice in the appropriate fields. These changes will not apply to non-bin or missing alleles for genotypes that have already been called.

Note: you need not call alleles automatically for genotyping. You may edit all genotypes manually (see next), should you wish to.

Viewing genotypes

To view alleles on the traces, select genotypes from the genotype table. This shows the corresponding traces at the marker's range in the detailed view. Allele names show above peaks that are assigned to alleles. These labels also appear when viewing samples, so long as traces are not stacked.

To filter genotypes to only show those meeting certain criteria, press the "filter" button ⚌ below the genotype table. A popover will allow you to define filter criteria. You may filter according to the name or size of alleles. If *both* alleles of a diploid genotype must meet the condition you define, select the **All** on the segmented button. If only one of the alleles must meet the condition, select **Any**. You selection has no effect on haploid genotypes.



Note:

- a filter on genotypes is associated to a folder and is recorded between app launches. If a folder has a filter applied to it, the "filter" button changes appearance:
- A filter on genotypes is not dynamic. If genotypes are modified so that they meet or no longer meet filter criteria, you must reapply the filter, or select another folder and go back to the previous selection, to update the genotype table.

To remove a filter on genotypes, press the "filter" button below the genotype table and select **Remove Filter** from the popover.

Tip: you can maintain a folder which shows all genotypes meeting certain criteria. For this, create a smart folder that would contain all samples (for instance, by setting an import date that is more recent that some old date) and apply a filter to the genotypes of that folder.

If a panel of markers is applied to samples, you can view their genotypes by right-clicking these samples on the sample table and selecting **Select Genotype(s)** from the contextual menu. If these samples have no genotype of if all their genotypes are filtered out, the menu item is disabled.

When genotypes are selected, you can go back to their source samples by right-clicking them on the genotypes and selecting **Show Source Sample(s)** from the contextual menu.

When viewing genotypes, the channel shown cannot be altered, since it corresponds to the marker's dye, and only one trace can be shown per row. Likewise, horizontal positions cannot be synchronized between rows, as this would cause rows to show ranges that do not correspond to the markers. Each view can be scrolled horizontally and zoomed, but its visible range will be reset to that of the marker each time the genotype is selected.

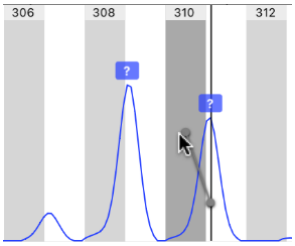
Double clicking the top ruler when the detailed view shows a genotype **moves the view to the range of the marker**.

Editing genotypes

The alleles of a genotypes are shown above peak as rectangular text label indicating the allele names. The color of these labels corresponds to the dye associated with the marker. The number of alleles is defined by the ploidy of the marker. When no peak has been found in the marker range, alleles are not visible and do not have a name.

To assign an allele to a peak in the marker range, you can:

- Double click the peak. If the peak has no allele, it will get one (possibly using an allele assigned to another peak). If the peak already has an allele, it will lose this allele, which may move to another peak if that other peak already has an allele (thereby creating a homozygote).
- Click the peak and drag. If the marker has bins, a handle starting at the position corresponding to the peak tip will appear. Drag the other end of the handle to the desired bin and release the mouse button. The peak will take an allele called after the bin name. If you release the mouse outside a bin, the allele will get the "out of bin" name (by default, a question mark).



- Drag an allele label that is above a peak to another peak. The name of the allele will become that of the bin at the peak, or the "out of bin" name if there is no bin behind the peak tip.

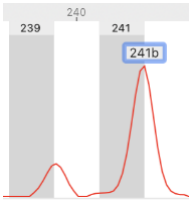
Note: none of the above is possible for peaks that are outside the marker range. If a peak outside the marker range should be considered as an allele, expand the marker range.

If a peak that should be an allele was not detected (probably because its fluorescence level is less than 100 units), right-click the peak and select **Add Peak Here** from the contextual menu. You may then add an allele to the peak as described above.

To remove an allele (without creating a homozygote if the marker is diploid) you can:

- select the allele label and hit the delete key,
- select the allele label and **Edit > Remove Allele** from the menu bar,
- drag the allele label above the top edge of the view showing traces.

This allele will have no size and will get the name for missing alleles (by default, an empty string).



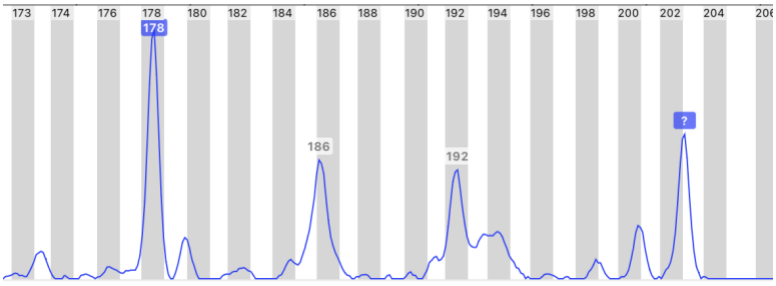
To give an allele an arbitrary name, do one of the following:

- Enter the name in the "allele1" or "allele2" column of the genotype table.
- Double-click the allele label above the peak and enter the desired name.

Tip: **to quickly review genotypes**, sort the genotype table by marker, allele1 and allele2 in ascending order (assuming allele names correspond to sizes). Then, view genotypes from top to bottom. This method helps detecting errors in allele call.

Supplementary peaks

In addition to alleles, STRyper may annotate other peaks found in a marker range if these peaks meet certain criteria. These peaks are not considered as alleles but are labelled like alleles except that their labels have a white background and a text in grey.



These labels can be edited like allele labels: they can be moved between peaks and renamed by double-clicking, and they can be deleted by the "delete" key. In addition, they can be deleted by double-clicking their peak. These actions have no effect on the sample's genotype as defined by the "Allele1" and "Allele2" columns of the genotype table.

Note: a supplementary peak label cannot be moved to a peak that has an allele. The allele must be removed first.

If you consider a supplementary peak as a real allele:

- Drag an allele label onto the peak. This will remove any supplementary peak label that is present.

- If no allele label is present at the marker, first remove the supplementary peak label from the peak (by double-clicking the peak or removing the label with the delete key), then double click the peak, which will attach an allele to it.

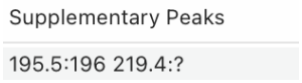
To annotate any peak as a supplementary peak (if it has no such label) click-click the peak and select **Add Supplementary Peak** from the contextual menu. This menu is not available if the peak has an allele.

To remove a supplementary peak, you can:

- select the peak label and hit the delete key,
- select the peak label and **Edit > Remove Peak** from the menu bar,
- drag the peak label above the top edge of the view showing traces,
- double-click the peak.

To remove all supplementary peaks from one or several genotypes, select these genotypes on the genotype table, right-click them and select **Remove Supplementary Peaks** from the contextual menu.

Information on supplementary peaks appear in the "Supplementary Peaks" column of the genotype table. For each peak, the column shows the size (in base pairs, rounded to the nearest decimal) of the supplementary peak and its name after a colon. Data from different peaks are separated by spaces and are sorted by increasing size. This column is not editable.



Exporting genotypes

To export genotypes as a tab-delimited text file, you can:

- Click the export button below the genotype table. This exports the whole table.
- Right-click selected genotypes and choose **Export Table to File…** from the contextual menu. This will export selected/clicked genotypes.

By default, the exported file comprises all the columns that are visible in the genotype table, and only these columns. You must change the column visibility if you want to change the content to be exported.

To add sample metadata to the exported genotypes, tick the **Add Sample-related Columns** checkbox in the export panel. Each visible field in the sample table will be added to each genotype.

You can copy selected genotypes to the pasteboard via **Edit > Copy**, and paste them to a spreadsheet or text editor. **Note**: this method does not copy sample metadata and column headers.

Exporting and importing folders

A folder, with all its their samples subfolders (including subfolders of subfolders) can be exported as a single file and imported afterward or into a different instance of the application. The exported archive contains all genotypes, size standards and panels (including markers and bins) applied to samples contained in the folder.

To export a folder, do one of the following:

- Select the folder and click the export button below the folder list.
- Select the folder and choose **File > Export Folder…** from the menu bar.
- Right click the folder and select **Export to Archive…** in the contextual menu.

If the exported folder is a smart folder, it is exported as a regular folder. Its search criteria are therefore not encoded in the file.

To import a folder from an archive, do one of the following:

- click the import button below the folder list,
- select File > Import Archived Folder… from the menu bar,

and chose the folder to import.

The folder will be imported at the bottom of the folder list. Any panel applied to samples in that folder that is not already present in the database is imported and placed in a new folder named after the imported folder. Any size standard not already present in the database is imported as well and the keyword "-imported" is appended to its name.

FAQ

Where is my data stored?

The data is stored in ~/Library/Application support/STRyper/. You may access this folder from the Finder by selecting Go > Go To Folder… and pasting this path into the field.

Can I change the location of the database?

No. STRyper looks for its database in ~/Library/Application support/STRyper/. Moving the database elsewhere will make the application recreate a blank database in this folder. However, you may replace the database folder with a symbolic link named identically and pointing to another folder.

Is my data backed up?

STRyper saves its database every 30 seconds, if necessary, but it does make backup copies. We recommend using a backup solution, such as Time Machine, to back up the database.

Does STRyper do database versioning?

No, STRyper does not maintain different versions of the database. For this, you must use your own backup/versioning solution.

How can I work on the same project on different computers?

You can export and import folders between computers. Transferring/synchronizing the whole database has not been tested. If you which to do so, you must make sure that the database is fully transferred before opening STRyper.

Is there an equivalent to Genemapper's bin sets?

STRyper has a more flexible solution: changing the offset of a marker for specific samples.

Why don't the size standards listed in the library have dyes names?

The size standard library lists, for instance, "GeneScan 500" rather than "GeneScan 500 **LIZ**", "LIZ" being the dye used. This is because the dye is irrelevant to sizing, only the size of fragments matter. Simply ignore the dye used by your size standard, just make sure to apply one with appropriate sizes.

Can I use allelic ladders?

STRyper does allow specifying allelic ladders for markers. An allelic ladder, also called "inter-lane standard", contains a mixture of known alleles for one or several markers. STRyper treats a chromatogram file obtained from such mixture as a regular sample. You may however use this sample to efficiently adjust the offset of a marker for all samples obtained in the same conditions.

Why was this peak not attributed to an allele?

This may be because the peak is considered by the application as resulting from crosstalk, i.e., it does not represent a DNA fragment from the channel that is visualized, but from another channel. You can check this by showing peak tooltips. You can still assign an allele manually to the peak.

Alternatively, the peak may be too faint to be detected if its fluorescence level is less than 100. In this case, hovering it with the mouse does not make a vertical line appear at its tip. You can still force the detection of the peak and assign it to an allele.

Can I use STRyper for human identification and forensics?

STRyper **must not** be used for human identification and forensics. STRyper comes with no guarantee whatsoever.

Does STRyper assess genotype quality?

No. Genotype quality must be assessed by the user. In our experience, genotyping quality computed by applications is poorly informative.

Is there a history of changes applied to genotypes?

STRyper does not maintain a history of manual changes applied to genotypes, as this is not important to population geneticists. You may still add notes to a genotype in the "Notes" column of the genotype table.

How can I export genotypes with one row per individual for all markers?

You cannot. This is because samples from the same folder can have different panels, hence markers (this is especially true for smart folders). However, if samples are analyzed at the same markers in the exported genotype file, you can achieve what you want with R:

```
requires(data.table)
genotypes <- fread("genotypes.txt", header = TRUE, sep ="\t") # use the appropriate path for "genotypes.txt"

# if you wish to paste allele names into genotype strings:

genotypes[,genotype := paste(Allele1, Allele2, sep = "")]
genotypesPerIndividual <- dcast(genotypes, Sample ~ Marker, value.var = "genotype")

# if not:
genotypesPerIndividual <- dcast(genotypes, Sample ~ Marker, value.var = c("Allele1","Allele2"))
# you may use c("Allele1","Allele2","Size1","Size2") if required

# You may include more columns in the formula, e.g.: Sample + Panel + Well ~ Marker, assuming these columns are present in the table
```

Can I define polyploid markers?

STRyper only manages haploid and diploid markers. At a polyploid marker, the genotype cannot directly be determined from the number of peaks. However, you may use supplementary peaks to annotate additional alleles. This could be a solution to genotype polyploid markers.

Can I use STRyper for other markers than microsatellites, such as AFLPs?

You should probably not. STRyper was designed for microsatellite markers.

How can STRyper be used in a multi-user environment?

As for any regular app, the user of STRyper is the user logged in the operating system. If several users must use the app on the same computer and work on different projects, the computer administrator must create a user account for each.

Different users can of course separate their projects into different folders, but these folders will be accessible to everyone using the app from the same user account.

Where can I report bugs?

Please, report bugs at https://github.com/jeanlain/STRyper/issues.

Will STRyper be ported to Windows?

The author of STRyper does not have the resources to port the application to other platforms. To make porting a reality, point a Windows developer to https://github.com/jeanlain/STRyper.