WALTER: web-based analyser of the length of telomeres

USER MANUAL

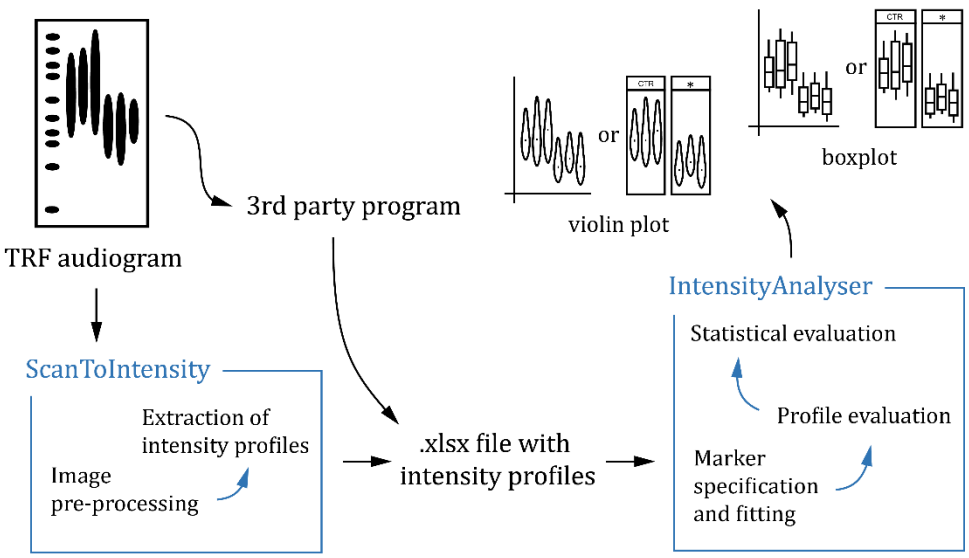
Martin Lyčka, Vratislav Peška, Martin Demko, Ioannis Spyroglou,
Agata Kilar, Jiří Fajkus and Miloslava Fojtová

Contents

About WALTER

WALTER is web-based toolset consisting of two tools: ScanToIntensity and IntensityAnalyser. Both tools are for the analysis of terminal restriction fragments (TRF) scans. ScanToIntensity tool is able to convert selected areas of a TRF scan into an intensity profiles while IntensityAnalyser tool analyses the resultant file from ScanToIntensity to provide a boxplot or a violin plot depiction of the profiles with a possible statistical analysis (picture of the WALTER workflow is shown below).

The whole WALTER toolset is written in an R programming language using the Shiny package and is available at: https://www.ceitec.eu/chromatin-molecular-complexes-jiri-fajkus/rg51/tab?tabId=125#WALTER.



Workflow of the WALTER toolset

1 ScanToIntensity tool

This tool transforms TRF scans and respective marker(s) into intensity profiles. However, this tool can be omitted from the WALTER toolset analysis and replaced by a third-party program which might be necessary in some cases as specified later in the text.

1.1 Upload of the TRF scan

In the tab "Upload picture"

1) Click on "Browse…" and select file of your interest.

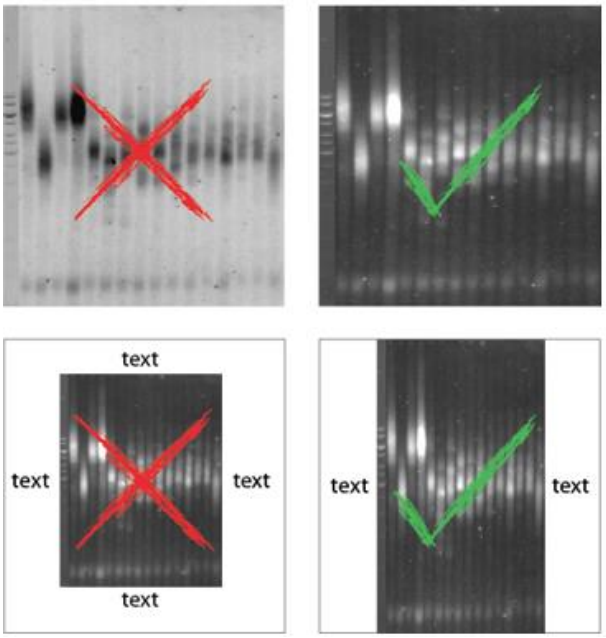
   Note: After confirmation you might need to wait for a while based on the size of the file.

   TIP: You can try to scale down your picture to a REASONABLE extent to reduce the time requirement for this step and later, the time required for transformation. It was proven that lowering the resolution of the TRF scan should not affect the outcome of the analysis significantly. However, retaining the original resolution is necessary if the statistical evaluation in IntensityAnalyser tool follows, because it will be affected by the change in resolution.

2) After a successful upload, the "Run program" button appears. Clicking it will take you to the tab "Sample transformation".

IMPORTANT: The TRF scan must be of maximum of **100 MB in size** and in **3 colour channel or grayscale**, **.bmp or**

.jpg format and the **higher signal must match the brighter colour.** Although the presence of text in the picture is a not an issue, as long as it is not present above or below the TRF scan itself; this would alter the analysis by the later tool (intensity profiles would include the intensity of the text and intensity of the background behind it).



TIP: To avoid the necessity to specify the telomere smear in intensity profiles later in the IntensityAnalyser tool, it is better to cut out the image in a way that technical defects (bubbles, blotches, smears) that are not related to telomeres are not included in the analysis.

1.2 Transformation to intensity profiles

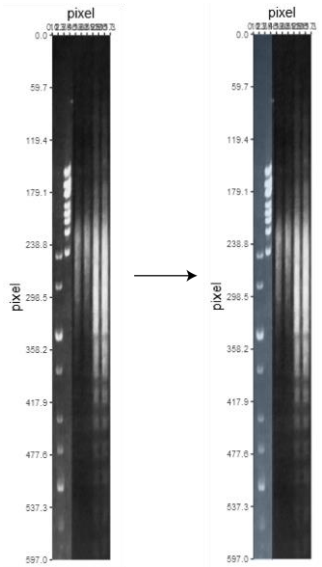
In the tab "Sample transformation"

1) Modify the uploaded picture by changing the plot size or by colour inversion to be able to easily distinguish each telomere smear.

2) Highlight the area of the picture corresponding to the marker by click, hold, and drag. Click on the button "Calculate profile".

Note: The profile is saved as selected sample 0. Consecutively, "Your last saved profile" area updates and the slider "Selected samples" will increase by one.

Note: Markers on the other side of the TRF scan (if present) can be used for marker correction if shift is expected. This marker can be saved as any sample number but better for it to be the last one.



TIP: More markers can be calculated as one profile if they are **adjacent samples** on the TRF scan. You can simply highlight the whole area corresponding to markers and calculate it as a single profile.

IMPORTANT: In case that more adjacent markers were used on one side and you want to use marker correction option in IntensityAnalyser tool, the markers on the other side of the TRF scan needs to be saved in the same manner so later there will be a profile with the same number of peaks.

3) Highlight the area corresponding to the first sample and click on the button "Calculate profile".

4) Repeat step 3 for all samples you want to analyse.

   Note: Profiles are saved as numbers in the increasing slider called "Selected sample number". The last number of the slider will always be empty, ready to be uploaded with data from the next sample.

5) Click on the button "Intensity profiles" to download excel (.xlsx) file with saved intensity profiles for further analysis by the IntensityAnalyser tool (2.1).

IMPORTANT: Do NOT try to calculate another profile while the program is busy! Just wait until the program finishes the task (busy bar will disappear as well as the "Your last saved profile" area updates). Calculation of one line can take approximately several seconds up to half-minute depending on the resolution of TRF scan and selected area.

IMPORTANT: If you need to recalculate some profile, move to the corresponding profile using a slider, highlight the area and click on the button "Calculate profile". This does not increase the number of the slider which increases only if samples are saved to the last number of the slider. If you want to continue with the calculation, move to the last number of the slider that is an empty slot.

1.3 Troubleshooting

Sometimes it may happen that the uploaded TRF scan is not of a standard quality and therefore there might be some issues with its analysis. Selected examples and ways to handle them are presented.

1.3.1 Lines on my TRF scan are curved

In this case, it is necessary to highlight the area of the width covering the whole length of the telomere signal. If the curvature is so pronounced that it could affect the calculation of the telomere length itself, it is better to use programs suited to handle the calculation of intensity profiles with curved samples (e.g., MultiGauge Ver. 3.0 (FujiFilm) or ImageJ).

Note: After processing the TRF scan with a different program, it is necessary to modify the resultant file for subsequent analysis by the IntensityAnalyser tool (specified in 2.1)

1.3.2 Lines on my TRF scan are stained with bubbles, blotches or smears not related to telomeres OUTSIDE of the telomere-specific signal

Technical defects outside of the telomere smear are generally not an issue in the following analysis by IntensityAnalyser tool. However, where such defects are of relatively high signal intensities (stronger than the telomere specific signal), it is beneficial to cut them off the TRF scan. Note that some background outside of the telomere signal is later used for the background correction in the IntensityAnalyser tool.

1.3.3 Lines on my TRF scan are stained with bubbles, blotches or smears not belonging to telomeres INSIDE of the telomere-specific signal

In this case it is necessary to select the area of maximum possible width covering the whole length of the telomere signal without any defects inside. If this is not possible, the sample is not suitable for analysis.

2 IntensityAnalyser tool

This tool facilitates recalculation of intensity profiles to a resulting violin plot or a boxplot with a possible statistical analysis to evaluate differences in telomere lengths between groups of samples.

2.1 Upload of the excel sheet with intensity profiles

In the tab "Upload file"

1) Click on "Browse…" and select the excel file of interest (1.2).
2) After successful upload, the "Run program" button appears. Click on it to move to the tab "Marker".

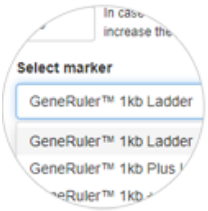
IMPORTANT: The Excel sheet must have a specific structure: the first column refers to the pixel number, the second to the intensity profile of the marker, and the subsequent columns to the intensity profiles of the samples (numbered). The first row refers to the names of columns and the second row contains pixel position of the marker(s)/samples in the TRF scan. If you have used a different program for to calculate the intensity profiles, it is necessary to modify the file accordingly.

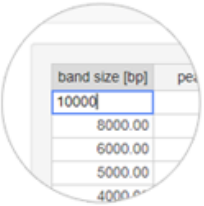
	A	B	C	D	E
1	pixel	marker	1	2	3
2	0	141	188	247	274
3	0	6.725216	10.25741	9.954118	10.10051
4	1	6.772275	10.31624	9.958039	10.09659
5	2	6.858549	10.20251	9.958039	10.10835
6	3	6.846784	10.19075	10.00902	10.14365
7	4	6.854627	10.35937	10.00118	10.16718
8	5	6.819333	10.36722	10.05608	10.20247
9	6	6.870314	10.28094	10.06392	10.21424

2.2 Specify Marker

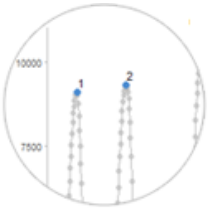
In the tab "Marker"

1) Select marker/combination of markers from the list. If the marker of your choice is not included, edit the band sizes accordingly.



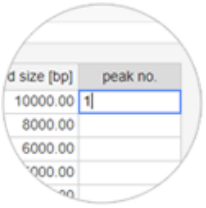


2) Check the graph if highlighted points on the highest peak correspond to the bands of the ladder. If they do not, change the sensitivity of found peak maxima in the respective box called "Sensitivity".



TIP: If there are too many pixel maxima, increase the sensitivity number to have a lower number of detected maxima.

3) Write in the adjacent column a peak number corresponding to the respective band.



Note: A smaller pixel number corresponds to a larger band size.

4) Click on the button "Proceed without marker correction" or "Proceed with marker correction" depending if you

want to correct for the possible shift and if the marker on the other side of the TRF scan was saved as intensity profile in the ScanToIntensity tool.

2.3 Marker correction

In the tab "Marker correction"

1) Select sample corresponding to the second marker intensity profile.
2) Check the graph if highlighted points on the highest peak correspond to the bands of the ladder. If they do not, change the sensitivity of found peak maxima in the respective box called "Sensitivity".

3) Write in the adjacent column a peak number corresponding to the respective band.

   Note: All bands filled in the tab "Marker" needs to be filled for the second marker too. Graph showing the first marker intensity profile with the highlighted points should help you to fill the second one accordingly.
4) Click on the button "Proceed".

2.4 Specify fitting of the marker(s)

In the tab "pixel/MW ratio"

2.4.1 Without marker correction (marker on one side)

1) Select the polynomial model that best fits to the points that were determined in the tab “Marker”.

   Note: The best fit is automatically recommended (and also mentioned in the text under the selection box) based on testing of the models by ANOVA.

   IMPORTANT: It is better to check the table and recommendation if there are any problems. Such a problem might be an increase in base pairs with increasing pixel number which cannot be correct. An example of what this might cause is provided in the chapter troubleshooting (2.9.1).

2) Click on the button “Choose”.

   Note: Clicking this button locks previous tabs and automatically moves you to the tab “Length calc.”. If any mistake has occurred and you want to change the specifications, click the button “Abort Analysis”.

2.4.2 With marker correction (markers on both sides)

1) Select the polynomial model that overall best fits to the points of all virtual markers calculated for each sample based on the points of the first and second marker specified in tabs “Marker” and “Marker correction”.

   Note: The best fit is automatically recommended (and also mentioned in the text under the selection box) based on testing of the models by ANOVA.

   IMPORTANT: It is better to check the table and recommendation of each virtual marker fit for every sample if

there are any problems. It is possible to move between samples using slider located under the graph. Such a problem might be an increase in base pairs with increasing pixel number which cannot be correct. An example of what this might cause is provided in the chapter troubleshooting (2.9.1).

2) Click on the button "Choose".

   Note: Clicking this button locks previous tabs and automatically moves you to the tab "Length calc.". If any mistake has occurred and you want to change the specifications, click the button "Abort Analysis".

2.5 Area selection for telomere length calculation

There are two options to analyse intensity profiles: **Automatic analysis** or **Manual analysis**.

Given that Automatic analysis can be easily combined with the background correction and allows easy reproducibility of the evaluation of TRF profiles by different people/labs, it is HIGHLY recommended as your first choice.

2.5.1 Automatic analysis

In the tab "Length calc."

2.5.1.1 Telomere smear IS covered in "red rectangle" (automatically highlighted)

1) Do a background correction of the sample by clicking on the button "BG correction".

IMPORTANT: After the background correction, the red rectangle can move to a different place out of the telomere smear. In that case, it is necessary to select the area of the signal again by changing coordinates in "Specification…area:" (details in 2.5.1.2).

Note: The background correction can be removed by clicking the button "Remove BGC".

2) Click on the button "Auto save".

   Note 1: The analysed intensity profile is now visible in the tab "Graphic result" and the graph under "Your selected area" is updated.

   Note 2: You can restart the saved area by clicking on the button "Restart". This also restarts the specifications of the first and last peak areas for this sample.

3) Move the sample slider (in the top left corner) to the next sample and repeat the process.

2.5.1.2 Telomere smear is NOT already covered in "red rectangle"

A peak belonging to the telomere-specific signal might not be of the highest intensity (e.g., a signal of interstitial telomeric repeats or other difficulties not removed during previous processing as described in 1.3.2), or distribution of the telomere-specific signal might not be unimodal. In these cases, it is necessary to specify the area of the telomere-specific signal.

1) Highlight the area of the first peak in the telomere-specific signal in the blue rectangle by click, hold and drag.

2) Click on the button "Save" under "Specification of the first peak area". The numbers update and the automatically highlighted area in the main graph changes (its left side).
3) Highlight the area of the last peak in the telomere-specific signal in the blue rectangle by click, hold and drag.
4) Click on the button "Save" under "Specification of the last peak area". The numbers update and the automatically highlighted area in the main graph changes (its right side).
5) Follow steps described in the 2.5.1.1.

   IMPORTANT: There is usually just one peak in the telomere-specific signal (unimodal distribution); in that case, in both steps 1) and 3), the same peak is highlighted. When the telomere-specific signal is fragmented, it is necessary to be aware of it and correctly highlight the first and the last peak of such a signal.

2.5.2 Manual analysis

In the tab "Length calc."

1) Highlight the area of the telomere smear in the blue rectangle by click, hold and drag.
2) Click on the button "Manual save".

   Note 1: The analysed intensity profile should now be visible in the tab "Graphic result" and the graph under "Your selected area" is updated.

Note 2: You can restart the saved area by clicking on the button called "Restart". This also restarts the specifications of the first and last peak area for the sample.

3) Move the sample slider in the top left corner to the next sample you want to analyse and repeat the process.

TIP: It is possible to combine Manual analysis with background correction if the approximate area of the telomere-specific signal is first highlighted in red (interval for the calculation of background correction is based on the left and right border of the automatically highlighted area). This can be achieved by changing specifications of the peak area and threshold values. However, it is still **highly recommended to use Automatic analysis with DEFAULT values** because of the repeatability of the evaluation process.

IMPORTANT: It is NOT recommended to manipulate the width of the telomere-specific signal by changing the threshold value; the default value of 28% is based on extensive empirical testing. The **only** appropriate time is when adjusting the area for background correction, and only in cases where there is a clear problem in close proximity to the telomere signal (e.g. a stain close to the telomere signal might be included in the selected area. In this case, manipulation with the threshold value might lower the impact of such a stain).

WARNING: Manipulation of the threshold value during background correction is **NOT** anyhow reflected in the report.

IMPORTANT: Even if changing the threshold value and specifications of the peak areas allows the use of Automatic analysis for any desired area (and, unlike in the background correction process, would be stated as such in the report), the reproducibility of such a result is NOT guaranteed!

IMPORTANT: If you feel that the intensity profiles of samples are not correct because they are suddenly cut or the value of the bp is not stably decreasing, you have probably incorrectly filled the tab "Marker" or selected the wrong polynomial order. In these cases, click the button "Abort Analysis". It will move you back to the tab "Marker" so that you can correct your mistakes.

Note: Clicking the button "Abort Analysis" will disable tabs "Length calc." and "Graphic result" and will restart all the work you have done in the tab "Length calc."

2.6 Graphical display of the results

In the tab "Graphic result"

The graph with samples analysed is shown on the right side from two columns with options. This is the low-resolution graph. If you want to download a ready to publish high-resolution graph, click the button "Show high-res plot". The picture of the high-resolution plot will show up and can be downloaded as a picture by the right click and "Save" option available in any browser.

Note: The calculation of the plot can take some time, especially for the violin plot. Until the calculation is done no plot

or an older plot (if you have previously opened the high-res plot option) is shown.

The graph can be presented either as a boxplot or a violin plot.

The boxplot visualises data through their quartiles showing the weighted median of telomere length, 1st and 3rd quartile based on this median and the minimum and maximum based on the selected area of the telomere-specific signal.

The violin plot visualisation shows all parameters of the boxplot and in addition includes information about the intensity of the signal throughout the boxplot area. Therefore, such plotting is useful in cases of non-standard telomere signals (e.g., fragmented signals).

TIP: Set all the configuration for the violin plot (e.g. range, ticks) in a boxplot mode and then switch to the violin plot to lower the time necessary for the violin plot calculation.

2.6.1 Variable parameters of the graph

2.6.1.1 Change of the width and the height of the graph

You can change the width and height of the graph by rewriting numbers in boxes “Width” and “Height”.

2.6.1.2 Change of the start and the end of the y-axis

You can change the start and end of the y-scale by rewriting numbers in boxes “Min” and “Max”.

2.6.1.3 Change of the range of the y-axis

You can change the range of the y-axis from auto option to several other pre-defined options with the slider “Range of y-axis”.

2.6.1.4 *Change of the type of the graph*

You can change the type of the graph by changing the parameters in the “Plot options” selection.

2.7 Statistical analysis

Groups can be compared in 4 different types of tests (2 groups are always compared by Welch's t-test). There is a possibility to compare data by multiple Welch's t-test either vs control group (CTR) or between each group. However, those tests do NOT adjust p-value for multiple testing. More stringent testing is provided by the Tamhane-Dunnett test that compares groups to the control group and by the Games-Howell test that compares groups between each other. Due to the possibility of a type II error, it is recommended to use the multiple Welch's t-test. For the same reason, significance is set as: * < 0.1, ** < 0.05, *** < 0.01 p-value.

In the tab “Graphic result”

1) To the box “Control group name”, write the name you want to assign to your control group (CTR, WT etc.)
2) To each sample number that is part of the statistical evaluation write the respective names of the groups that the samples belong to.
3) Click the button “Evaluate”.
4) Select the test that you want to use for the statistical evaluation.

Note: After clicking button "Evaluate", the box "Control group name" and table with sample numbers and groups will freeze until you click the button "Cancel".

IMPORTANT: There are several conditions that your table must meet so that the statistical evaluation is possible (otherwise an error with the problem will be raised).

- There **must be** a group with the name you have entered into the box "Control group name" (control group).
- There must be **at least 3 samples in** each **group** (however, it is STRONGLY recommended to have at least 5 samples).
- There must be **at least 2 groups**.
- Group name must be filled for **ALL** samples.

2.8 Report of the analysis

In the tab "Graphic result" + downloaded html file

1) Click on the button "Save report".
2) Open downloaded html file.

Note: The html file (name formatted as (current date)_report_IntensityAnalyser_WALTER) includes chosen parameters in each step of the analysis by the IntensityAnalyser tool.

2.9 Troubleshooting

2.9.1 Problems with marker and model fitting resulting in wrong results

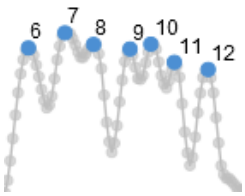
In some cases, the polynomial model is not decreasing for the whole pixel interval which results in the wrong recalculation of pixel-to-base pair length. This problem is connected to the filling of the table in the tab “Marker”. If there is a sudden drop in MW over a short distance, the recalculation will most likely be wrong. In these cases, it is better to delete such a marker band and let it be extrapolated.

IMPORTANT: Always have your telomere-specific signal inside bands of the marker you have used! Having it in the middle of the marker is preferable.

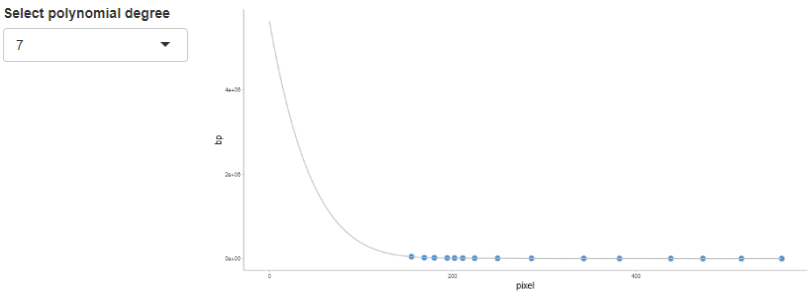
Here is the example of such a situation handled wrongly and correctly:

Tab “Marker”

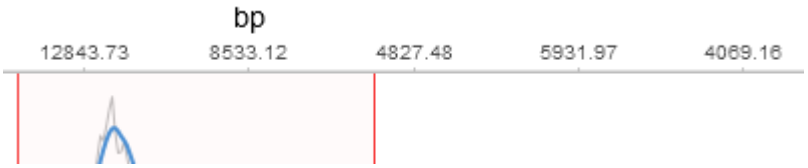
band size [bp]	peak no.
48502	6
24508	7
20555	8
17000	9



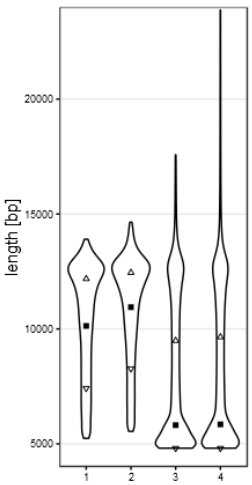
Tab “pixel/MW ratio”



Tab “Length calc.”

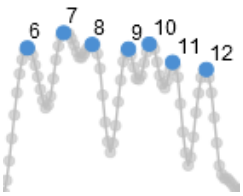


Tab “Graphic result”

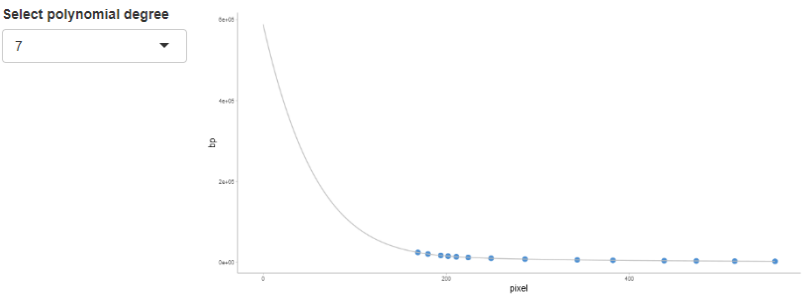


Tab “Marker”

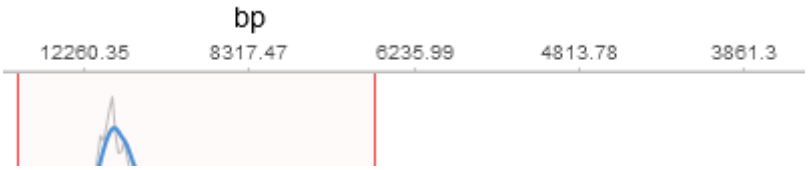
band size [bp]	peak no.
48502	
24508	7
20555	8
17000	9



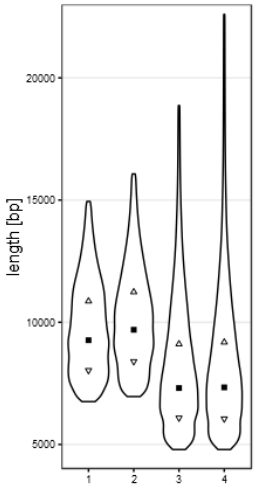
Tab “pixel/MW ratio”



Tab "Length calc."



Tab "Graphic result



2.9.2 There is a sample in the intensity profile file that I do NOT want to analyse

There are two ways how to exclude an unwanted sample from the analysis outcome:

Remove column with the number corresponding to the sample that you do not want to use in the evaluation and rename all the column numbers.

OR

Either skip the sample or click on restart button for said sample in the tab "Length calc." However, using this way will lead to missing number in the graph.

3 Example

From the website https://www.ceitec.eu/chromatin-molecular-complexes-jiri-fajkus/rg51/tab?tabId=125#WALTER it is possible to download a .zip file with an example of a TRF scan and its analysis for learning purposes.