

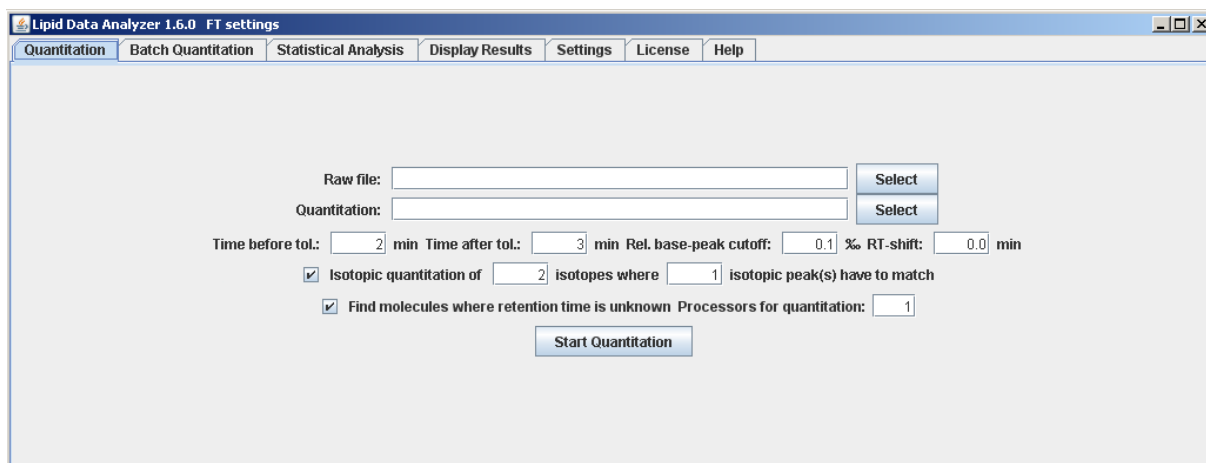
Lipid Data Analyzer 1.6

User Manual

1	General overview	1
2	Quantitation.....	1
3	Batch quantitation	3
4	Manual verification – visualization.....	4
4.1	Visualization – left menu.....	4
4.2	Visualization – 3D viewer	5
4.3	Visualization – Chromatogram viewer.....	6
5	Statistics	7
5.1	File selection and settings.....	7
5.2	Heat map and visualization settings	9
5.3	Bar charts.....	14
5.4	Overview bar chart	16
6	Settings.....	17
7	Licensing.....	17
8	Help.....	18
	Appendix A - Preparation of the mass list Excel file	21
	Appendix B - Configuration settings	22
	Path settings	22
	Look & feel settings	22
	Default quantitation settings.....	22
	Excel result file settings	22
	3D viewer default settings.....	22
	Standards prefix settings	22
	mzXML → chrom settings.....	23
	General algorithm settings	23
	Memory settings.....	23
	Appendix C – mzTab settings	23
	MS instrument specific settings	23
	Other mzTab settings	24
	Appendix D - Trouble shooting	24

1 General overview

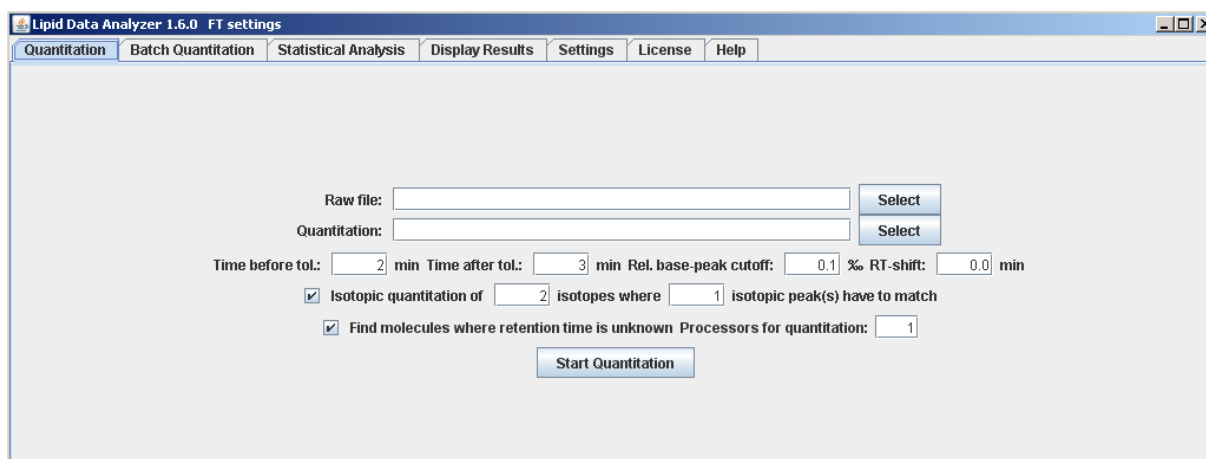
There are two ways to start the program, one as plain executable and one with a console. If the console is used, a command line window (black one) appears, which is the output of the program log.



In the LDA window six tabs are available. For “Quantitation” and “Batch Quantitation” see chapter 2 and chapter 3 the “Quantitation” and “Batch quantitation” section; for “Results Analysis” see chapter 5 the “Statistics” section; for “Display Results” see chapter 4 “Manual verification – visualization”; for licensing visit chapter 6; and for help chapter 8.

2 Quantitation

In order to start a single quantitation a raw file (Raw file:) and an Excel file (Quantitation) containing the mass list is required. The default settings for the input fields can be entered in the LipidDataAnalyzer.properties file (see Appendix B - Configuration settings).



Raw file: File in mzXML or chrom format is required. However, it is possible to use Thermo RAW format and the Waters Mass Lynx raw directories directly if XCalibur or MassLynx is installed.

Quant. files: Here the name of an Excel file has to be entered, containing the mass list. The format of the file is described in Appendix A - Preparation of the mass list Excel file.

Time before tolerance: This field specifies the retention time tolerance in minutes before the entered retention time in the Excel file. Example is given after “Time after tolerance”. This field is just used if a retention time has been entered in the mass-list Excel file.

Time after tolerance: This field specifies the retention time tolerance in minutes after the entered retention time in the Excel file. Example is given in the next paragraph. This field is just used if a retention time has been entered in the mass-list Excel file.

Example for time tolerance setting: E.g. in the excel file the retention time for a specific analyte is specified with 24 minutes, in the time before tolerance 2min and in the time after tolerance 3min is entered. Then, the algorithm will look for the analyte in the time range from 22-27 minutes.

Rel. base-peak cutoff: Peaks that are less intense than this per mille value of the highest identified are discarded by the algorithm. For the determination of the highest peak, just found analytes are taken into account (irrespective of the analyte class). This threshold can accelerate the quantitation, since intensities that are too small are discarded before the 3D quantitation is performed.

RT-shift: The retention time could shift from batch to batch. The entered value will be added to the retention time defined in the mass list file. Thus, the mass list file has not to be changed every time.

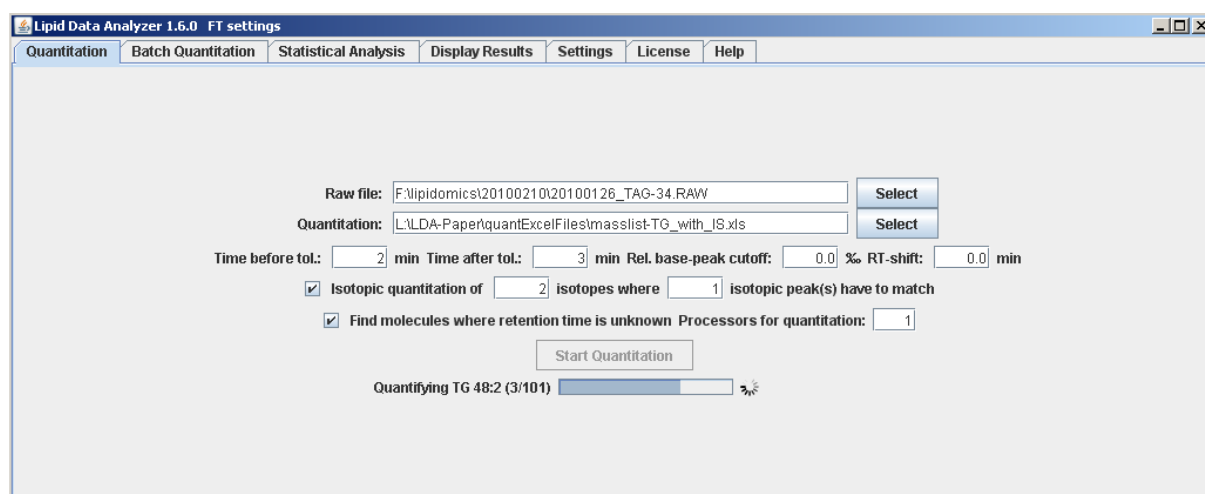
Isotopic quantitation of ... isotopes where ... isotopic peaks have to match: The checkbox indicates if in addition to the +0 peak other isotopic peaks should be quantified. The first value determines the amount of additional isotopes to be quantified (e.g. 2 would mean that +0, +1, and +2 isotopes will be quantified). The second value determines how many isotopes have to conform the theoretical isotopic distribution (e.g. 1: just the +1 peak has to conform).

Find molecules where the retention time is unknown: If this checkbox is selected, molecules in the mass list without retention time entry are quantified, discarded otherwise. However, the quantitation of molecules without retention time consumes more time.

Processors for quantitation: The amount of processors that shall be used for the quantitation. LDA detects the amount of processors available, automatically. The default value is always n-1, and the software does not allow for more than n processors.

Start Quantitation:

This button starts the quantitation.



After the quantitation has been started, first Thermo RAW format and the Waters Mass Lynx raw directories are translated in mzXML (if the quantitation is based on the raw file formats). Second, mzXML files are translated in the internal chrom file format. If the raw file is Thermo RAW or Waters

Mass Lynx and not mzXML, the intermediate mzXML is deleted automatically after conversion to chrom format. Then, on the chrom files, the quantitation is started. The results of the quantitation will be stored in an Excel file, whereby the name of the results file will be \$path_of_the_chrom_file_without_suffix\$_\$quant_file_name\$ (e.g. raw file: F:\lipidomics\20100210\20100126_TAG-34.chrom quant file: F:\lipidomics\TG_NH4_ACN.xls → result file: F:\lipidomics\20100210\20100126_TAG-34_TG_NH4_ACN.xls).

3 Batch quantitation

The quantitation pages for “Quantitation” and “Batch Quantitation” look similar. The difference is that for “Quantitation” the file chooser requires a single file, whereas in “Batch Quantitation” a directory containing the files to quantify and a directory containing the Excel mass list files is required.

Lipid Data Analyzer 1.6.0 FT settings

Quantitation Batch Quantitation Statistical Analysis Display Results Settings License Help

Raw files: Select

Quant. files: Select

Time before tol.: 2 min Time after tol.: 3 min Rel. base-peak cutoff: 0.1 % RT-shift: 0.0 min

☒ Isotopic quantitation of 2 isotopes where 1 isotopic peak(s) have to match

☒ Find molecules where retention time is unknown Processors for quantitation: 1

Start Quantitation

In contrast to the normal quantitation, in batch mode, a list will appear that shows the progress of the quantitation. The files receive a green check mark, if they are quantified successfully. If not, a red X.

Lipid Data Analyzer 1.6.0 FT settings

Quantitation Batch Quantitation Statistical Analysis Display Results Settings License Help

Raw files: F:\lipidomics\Niki\20100901\Eichkurven_20100827_BW Select

Quant. files: F:\lipidomics\Niki\massList Select

Time before tol.: 2 min Time after tol.: 3 min Rel. base-peak cutoff: 0.1 % RT-shift: 0.0 min

☒ Isotopic quantitation of 2 isotopes where 1 isotopic peak(s) have to match

☒ Find molecules where retention time is unknown Processors for quantitation: 1

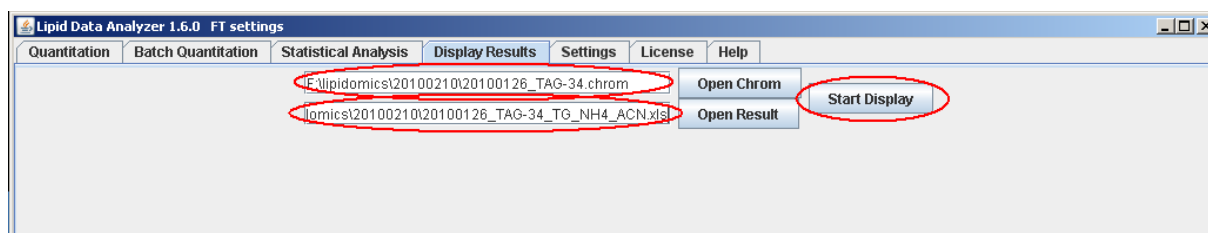
Start Quantitation

Data	Quant-File	Progress
20100827_WB_eichreihe_4a.chrom	20100831_Massenliste_EICHKURVE.xls	Finished
20100827_WB_eichreihe_7c.chrom	20100831_Massenliste_EICHKURVE.xls	Quantifying PC 32:2 (2/19)
20100827_WB_eichreihe_7b.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_12c.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_7a.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_12b.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_12a.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_3c.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_3b.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_3a.chrom	20100831_Massenliste_EICHKURVE.xls	
20100827_WB_eichreihe_6c.chrom	20100831_Massenliste_EICHKURVE.xls	

Quantifying 20100827_WB_eichreihe_7c.chrom with 20100831_Massenliste_EICHKURVE.xls

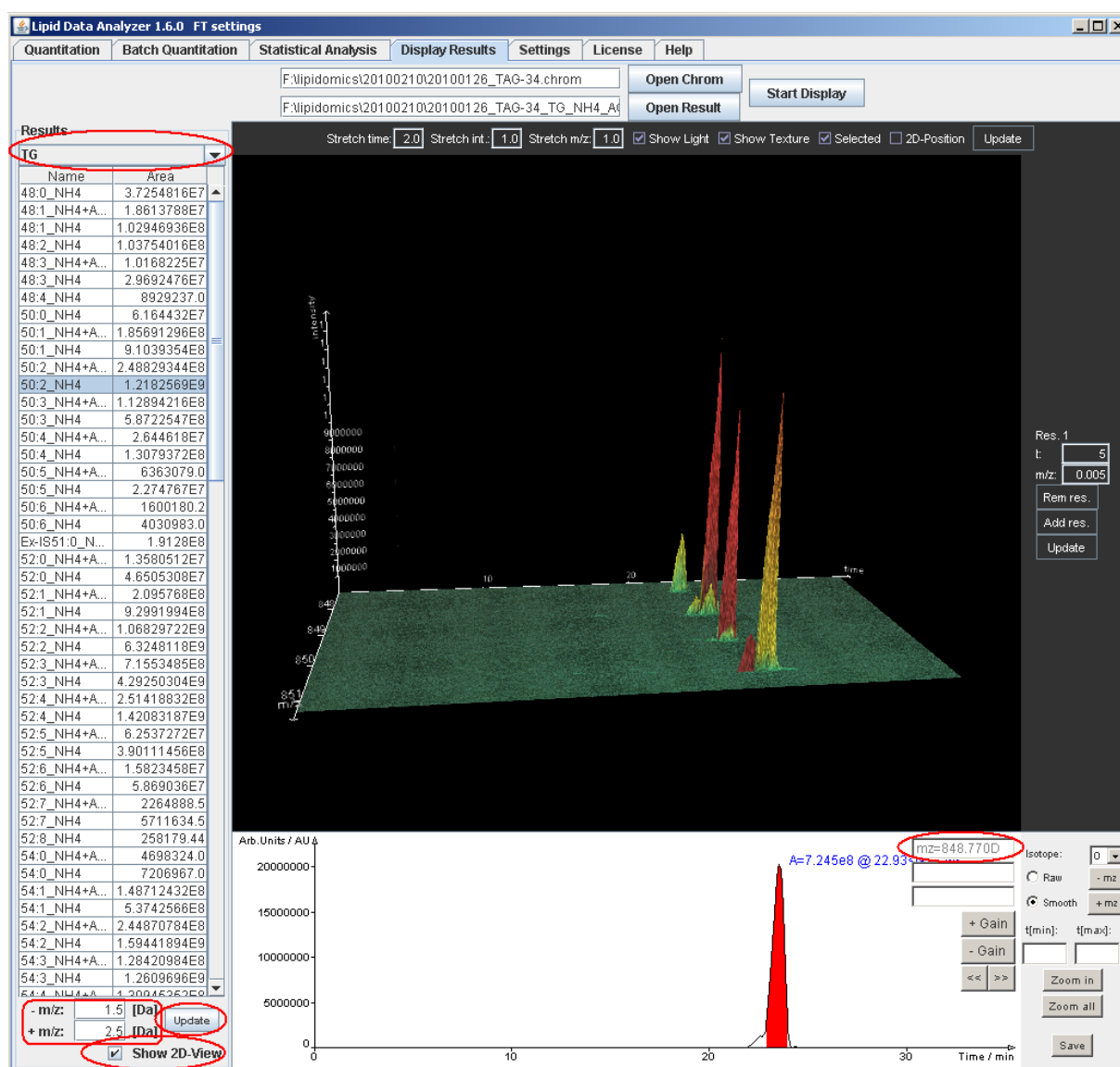
4 Manual verification – visualization

In order to visualize the results 2 files are required: the chrom file and the results Excel file:



After the files have been selected press “Start Display”. The now appearing display can be divided into two parts: 1) a menu on the left side (see 4.1); 2) a display frame, whereupon the upper part shows a 3D view (see 4.2) and the lower part displays the 2D chromatogram (see 4.3).

4.1 Visualization – left menu



The menu has at the top a selection box, where you can switch between the different lipid classes. Then a table follows, whereupon the name of the analyte is in the first column and the area in the second. The display name consists in principle of “\$name\$:\$double_bonds\$” (e.g. 50:2). If there

exists more than one adduct/modification, the modification name is added with `_$modification_name$` (e.g. 50:2_NH4). At the bottom of the menu, the range for the 3D viewer can be set. The current center m/z value of the extraction is displayed in the 2D viewer (see chapter 4.3). After the m/z range has been changed, the “Update” button has to be pressed to change the display. The “Show 2D-View” select box can fade out the 2D viewer.

If a row of the table is clicked the peaks are displayed in the 3D and the chromatogram viewer. If the row is clicked with the right mouse button a popup appears:

50:0	6.164432E7
50:1	9.1039354E8
50:2	1.2182569E9
50:3	Add analyte before
50:4	Add analyte after
50:5	Delete analyte
Ex-IS51:0	1.9128E8
52:0	4.651642F7

This allows to add an additional analyte or to delete the selected one. To delete several analytes at once, keep Ctrl or Shift pressed and select the analytes, this will prevent refreshing of the 3D viewer and save time. As soon as you release Ctrl or Shift, the 3D viewer will be refreshed by the analyte that was selected at last. Before/after just determines the position where the analyte will be added. The adding of a new analyte requires the name of the analyte, its chemical formula, the name and the formula of the modification if any, the m/z tolerance of the chromatogram and the exact mass.

Enter new analyte

Add a new analyte before 50:2

Name: 50:1

Formula: C53H100O6

modification: NH4

Formula: NH4

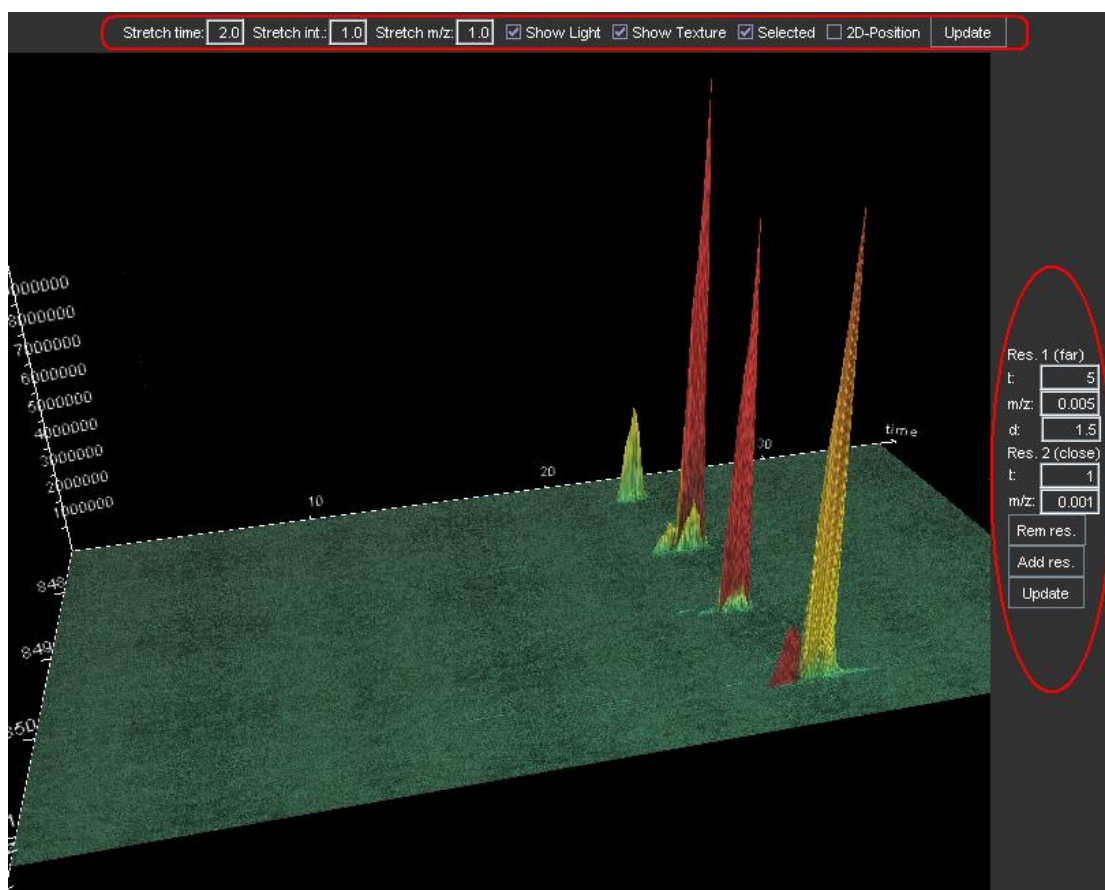
m/z tolerance: 0.02

exact mass: 850.78578

Cancel

OK

4.2 Visualization – 3D viewer



The 3D viewer provides a menu bar at the top which can change the appearance of the viewer, and a menu bar on the right side which is responsible for the resolution of the viewer. The view on the plane can be changed when the mouse is clicked inside the drawing canvas.

Top menu bar:

- Stretch time/int./m/z: The 3D viewer itself treats every dimension equally, so that every dimension should have the same length. However, for the chromatography it makes more sense to stretch the time axis. Thus, the “stretch time” is automatically set to 2.0. Any of the three dimensions can be stretched or shrunk with these three input fields. The settings are applied after the “Update” button has been pressed.
- Show Light: The 3D viewer has a light source falling on the plane, resulting in bright or shadowy spots. The light effects can be switched off with this check box.
- Show Texture: In order to give the picture more contrast a texture is rendered over the plane. The texture can be turned off with this check box.
- Selected: If selected, the quantified and stored peaks are colored in red; the ones that are just quantified and not stored are colored in green. If not selected, the coloring depends on the signal intensity (green: low intensity; yellow: medium intensity; red high intensity).
- 2D-Position: This shows the stripe that has been taken to extract the chromatogram in gold.

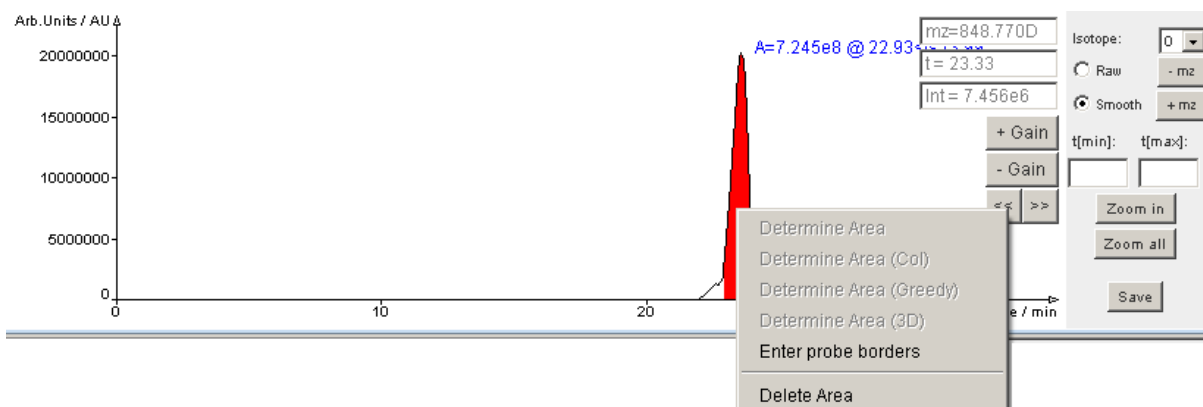
Right menu bar:

- t: Is the time resolution in seconds.
- m/z: Is the m/z resolution in Daltons.
- Add res.: The viewer has the ability to change the resolution level depending on how close the view is on the plane. Several resolution levels can be entered. If an additional resolution level is required just press this button. The point where the resolution is changed is defined by the d field.
- d: Is the distance field; it appears if several resolution levels are set. The distance is always measured from the center of the plane. 1 corresponds to a total distance of the m/z axis to the center of the plane. This value is rather empirical, but due to general applicability to different resolutions this approach had been chosen.
- Rem res.: Removes a resolution level
- Update: This button has to be pressed when the resolution settings should take effect for the 3D view.

Mouse buttons:

- Left mouse button: Changes the viewing angle on the canvas.
- Middle mouse button: Zooming in and out.
- Right mouse button: Moving the plane.

4.3 Visualization – Chromatogram viewer



If the mouse is hovered over the painting canvas a crosshairs cursor is rendered, whereupon the current time and intensity is displayed in the two text fields on the right side of the canvas (e.g. “t = 23.33” and “Int = 7.456e6”). The third text field indicates the m/z value of the chromatogram. With +/- Gain the displayed intensity range can be zoomed, and with “<<” and “>>” it is possible to shift the currently zoomed time range. If a peak is quantified and stored it is displayed in red, if it is just quantified and not stored it is displayed in green. If the right mouse button is clicked under a curve the displayed popup menu appears with the following options:

- Determine Area: detects a peak area with the standard ASAPRatio algorithm.
- Determine Area (Col): detects a peak area with the MASPECTRAS algorithm that integrates over local maxima.
- Determine Area (Greedy): sets the peak borders at sudden changes in the steepness of the curve.
- Determine Area (3D): 3D algorithm for more accurate peak border confinement.
- Delete Area: deletes a quantified area.
- Enter probe borders: manually define the borders of the peak:

The peak can be defined in 2D or 3D mode. In 2D mode the chromatogram is used and just the start- and stop time has to be entered. For the 3D mode, the m/z values have to be entered, and an ellipse is fitted through these values.

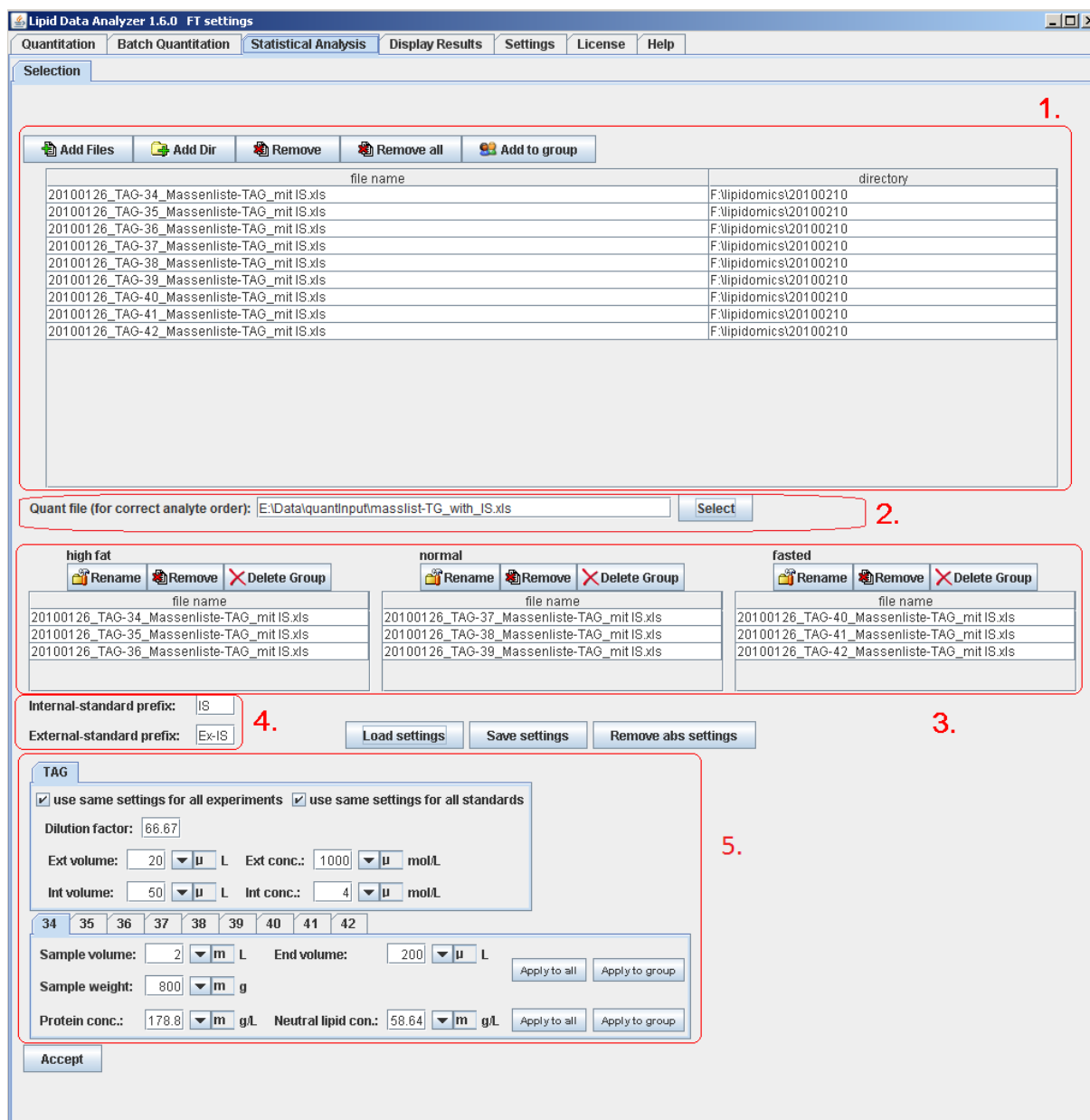
In the right menu of the chromatogram viewer the following options are available:

- Isotope: Switches between the chromatograms of the different isotopes.
- Raw/Smooth: Toggles between the raw and the smoothed version of the chromatogram (smoothed is the default one).
- t[min], t[max] and Zoom in: This is for zooming the time axis. With t[min] and t[max] the display range is defined and with “Zoom in” the settings are used for the viewer. After zooming the << and >> buttons can be used.
- Zoom all: zooms the time axis out again.

5 Statistics

5.1 File selection and settings

The page for the selection of the files can be split into for 4 parts, whereby changes in this section will just take effect in the statistics section, after the “Accept” button had been pressed.



Number 1 is for the selection of files.

- Add Files: This button allows selecting result files (Excel) for the analysis. The files appear then in the table below.
- Add Dir: Does the same like “Add files” but reads automatically the content of a whole directory and adds all of the Excel files in the directory to the table.
- Remove: Removes one or several results from the table. The files have to be selected before in the table.
- Remove all: Removes all files in the table.
- Add to group: Is for the grouping of results. First, a set of files has to be selected in the table, and then this button has to be pressed. Then, a dialog box asks for the name of the group and the group is added to number 3, the grouping part.

Number 2 is for the sorting of the analytes.

If, in the quantitation step, several analytes cannot be detected, not all of the analytes appear in the Excel results file. Then, it is hard for the algorithm to specify the correct order. At number 2, the original quantitation file can be provided. The analytes will then be displayed in the order they occur in the Excel file. The provision of the original Excel file is not mandatory, but can be helpful if several analytes were not found.

Number 3 displays the groups (they have to be selected at number 1). The grouping is not mandatory, but it can help in the interpretation of the results. Each group is displayed in a small table and the following buttons are available:

- **Rename:** Renames the group.
- **Remove:** Removes members of one group after they had been selected in the table where the group members are shown.
- **Delete Group:** Deletes the entire group.

Number 4 is for the selection of molecules as standard. It is assumed that the standards in the results carry a certain prefix (e.g. IS or Ex-IS). This prefix has to be entered in the two input fields. The usage of standards is not mandatory, so these fields can remain empty. If there are prefixes entered and no molecule carries these prefix, the standardization is neglected.

Number 5 is for the standardization on certain values like the protein content or the calculation of absolute values (just valid if there are standards and the ionization efficiencies of the analyte can be assumed to be similar like the one of the standards). This section is not mandatory. Before this section can be used, the results have to be selected first in number 1 and the prefixes for the standards have to be defined in number 4. Then the “Add absolute settings” button has to be pressed. The application reads now the result files and looks for standards.

The upper box of this section is for the assessment of absolute values for the standard. The box has a tab for each lipid class, so the settings can be made specifically for each class.

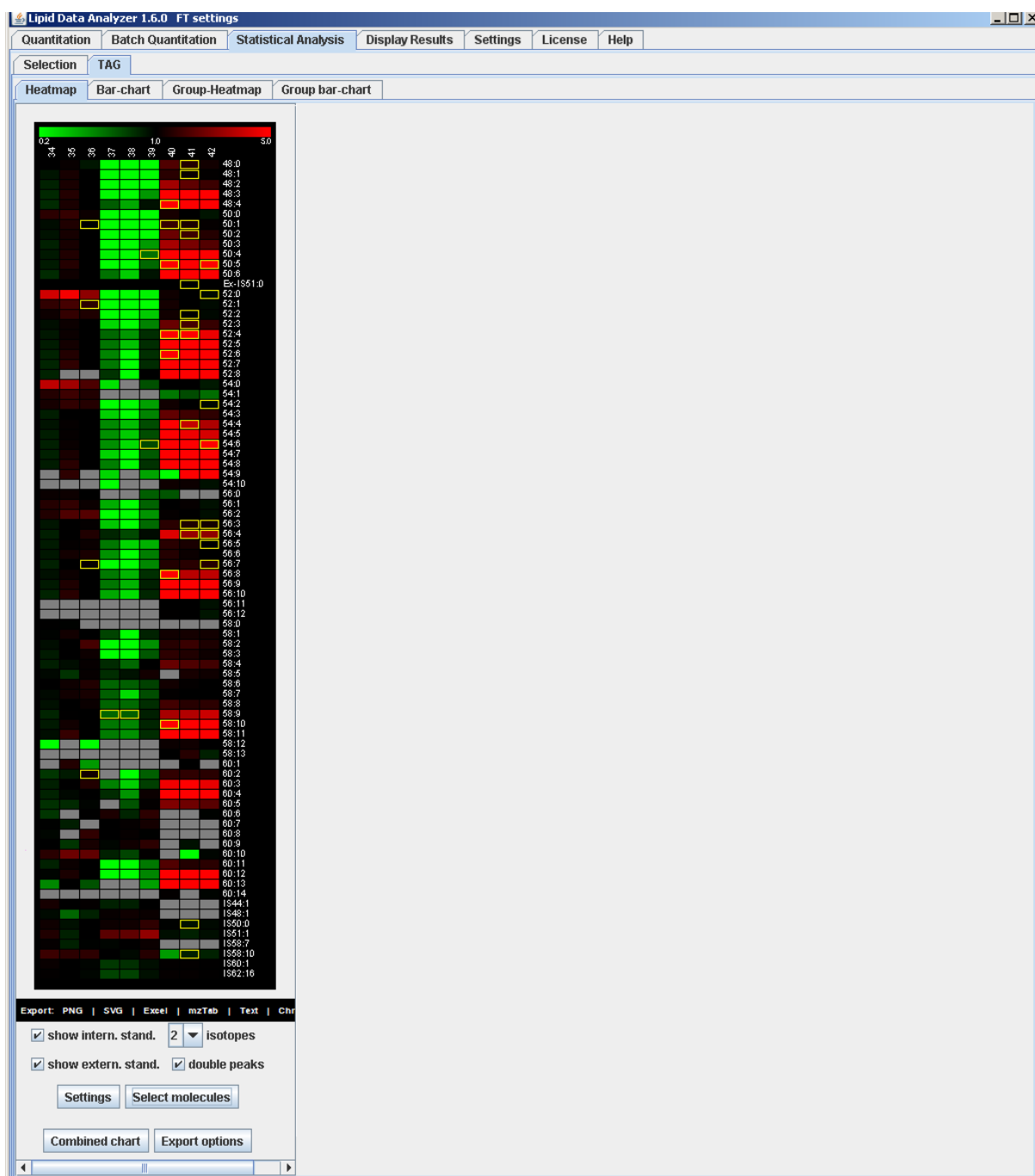
- **use same settings for all experiments:** If this checkbox is selected, it is assumed that the standards are added to each experiment at the same amounts. So a specific setting for each experiment is possible.
- **use same settings for all standards:** If this checkbox is selected, it is assumed that all of the external standards (“Ext volume” and “Ext conc.” has to be entered) and internal standards (“Int volume” and “Int conc.” has to be entered) are added at the same amounts. If not, a list of the standard appears and each standard can be entered separately.

The dilution factor corresponds to the dilution from the adding of the external standard until the internal standards are added. If the settings are entered for each experiment separately the buttons “Apply to all” and “Apply to group” appear. These buttons permit the propagation of the entered values to all experiments or to the members of a group.

The lower box is for the specification of sample specific settings like sample volume (before sample preparation) and end volume (after sample preparation). The sample volume and the end volume are mandatory, whereas the “Sample weight”, the “Protein conc.” and the “Neutral lipid conc.” are not. If latter values are entered, standardization on these values is possible. The “Apply to all” and “Apply to group” buttons permit the propagation of the entered values to all experiments or to the members of a group.

5.2 Heat map and visualization settings

After the “Accept” button had been pressed, a tab for each lipid class appears. This tab again contains a tabbed pane with the entries “Heatmap” and “Bar-chart”, and if groups are selected “Group-Heatmap” and “Group bar-chart”. The “Heatmap” tab is selected by default. The tab itself contains the heat map at the top, followed by an export bar, and some control elements at the bottom.



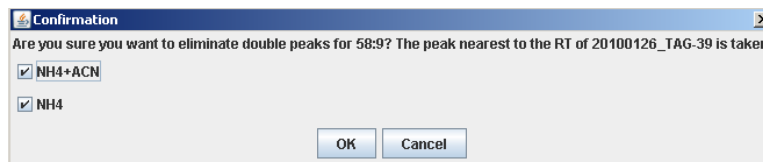
Heat map:

At the top, the heat map contains a color legend. The values in the map are calculated relative to the median of one molecule over all samples/groups. If the intensity of a molecule of one sample/group is lower than the median, it is colored green, the ones that are higher are red, and the ones around the median are black. If one analyte could not be quantified it is gray. In the heat map the samples/groups are organized horizontally and the lipids vertically. If a quantitation is based on more than one peak (could be an ambiguous identification; each isotope separately), a yellow rectangle is around the peak. If the mouse is hovered over one heat map cell, a white rectangle is rendered and at the bottom of the application a line with the name of the lipid, the name of the sample, the value relative to the median, a value depending on the “Settings” option (can be standardized), the original value of the quantitation, and the amount of available isotopes is displayed.

Mouse clicks on cells in the heat map (not valid for the group heat map):

- Left mouse button (does not work for gray fields): The LDA visualizes the quantitation (see chapter 4). If the quantitation is based on several peaks due to adducts/modifications, the LDA jumps to the first one it finds.

- Right mouse button (does not work for gray fields and the ones with a yellow rectangle): A popup menu appears with 2 options: 1) “Choose just one peak for doubles”; this option is for the automated removal of double peak identifications (yellow rectangle). It automatically selects in the other samples the one peak as correct, that is closer to the retention time of the sample where the right mouse button has been clicked. After this option has been selected the application asks for which adducts/modifications this operation should be performed.



2) “Quant. anal. at not found”; this option tries to automatically quantify analytes which are grey at the retention time of the selected analyte, whereby it does not take the theoretical isotopic distribution into account. This procedure tries out all of the available quantitation methods, starting with the 3D method, then the “greedy” method followed by the MASPECTRAS and ASAPRatio method (see chapter 4.3). For this option the application asks as well for which adducts/modifications this operation should be performed like in “Choose just one peak for doubles”.

Mouse clicks on the “sample/group” name:

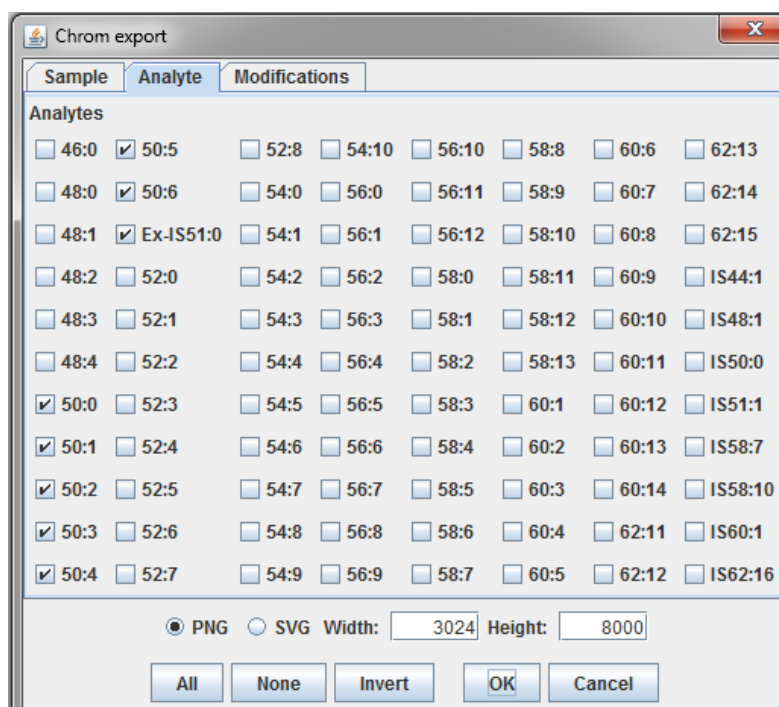
- Left mouse button: a bar chart for the sample/group is displayed, containing all the available analytes.
- Right mouse button (just for sample): Renaming of sample name.

Mouse clicks on the molecule name:

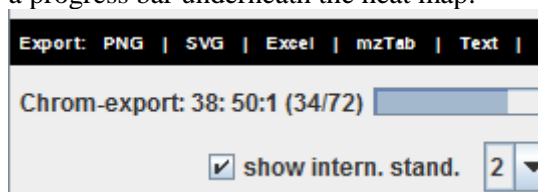
- Left mouse button: a bar chart for the molecule is displayed, containing all available samples/groups.
- Right mouse button (just valid for sample): A popup menu appears containing 3 options: 1) “Remove analyte in all probes”: deletes one lipid in all result files. 2) “Select analyte”: If there are a lot of lipids and a lot of result files, the removal of a lipid can be quite time consuming. However, if there are several analytes deleted, it consumes the same time. With this option analytes can be selected for the removal; if selected, the name of the analyte is surrounded by a blue rectangle. The removal takes effect on all of the selected analytes as soon as the “Choose just one peak for doubles” option is pressed. Here again a confirmation box appears asking for the adducts/modification on which the removal should be performed.

Export bar:

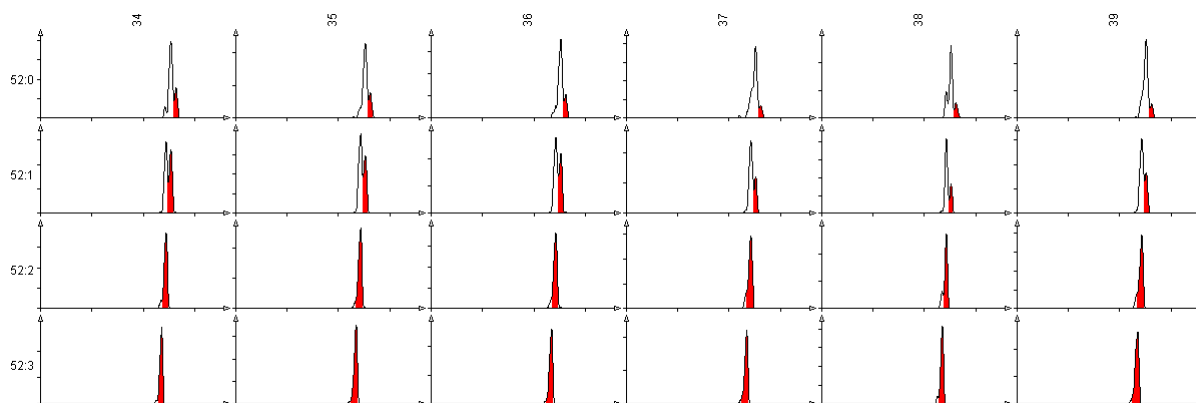
The picture of the heat map or the data respectively can be exported in the following formats: PNG = portable network graphics; SVG = scalable vector graphics; Excel = Microsoft Excel; Text = text based format (tab delimited). Excel and Text export the values in the type that are selected by the user (see next paragraph – Control elements). Furthermore, there is the option to export results in mzTab format, in order to submit them to a public repository. For mzTab, the raw peak areas are exported, these can be affected by standardization and isotope setting only. The mzTab button exports all lipid classes at once and generates a file containing all experiments; the data can be exported to chromatograms with the “Chroms” button. Here, the exported analytes, experiments and modifications can be selected.



It is recommended to export just one or two lipid series at once. The chroms export can take some time, depending on the data and the amount of selected hits. The progress of the export is presented in a progress bar underneath the heat map.



At the end of the chroms export, the user is informed that the export is finished.



This picture is a chroms export of TG52 of the LDA. It can be easily seen that analytes with more double bound elute slightly earlier. This is a good quality criterion that the algorithm selected the correct hit.

Control elements:

- show intern. stand.: should the internal standards be displayed in the heat map
- show extern. stand.: should the external standards be displayed in the heat map
- isotopes: what is the highest isotope number that should be used for the heat map for all of the analytes. However, the highest isotope number is determined for each analyte separately (e.g. if

we have three samples 1, 2 and 3 and for 1 and 2 the isotopes +0,+1, and+2 are found, whereas in sample 3 just +0 and +1 are found; just +0 and +1 are taken for the heat map).

- double peaks: should duplicate peak identifications be flagged with the yellow rectangle
- Settings button: This is the major settings panel for the display.

The screenshot shows a settings window with the following fields and controls:

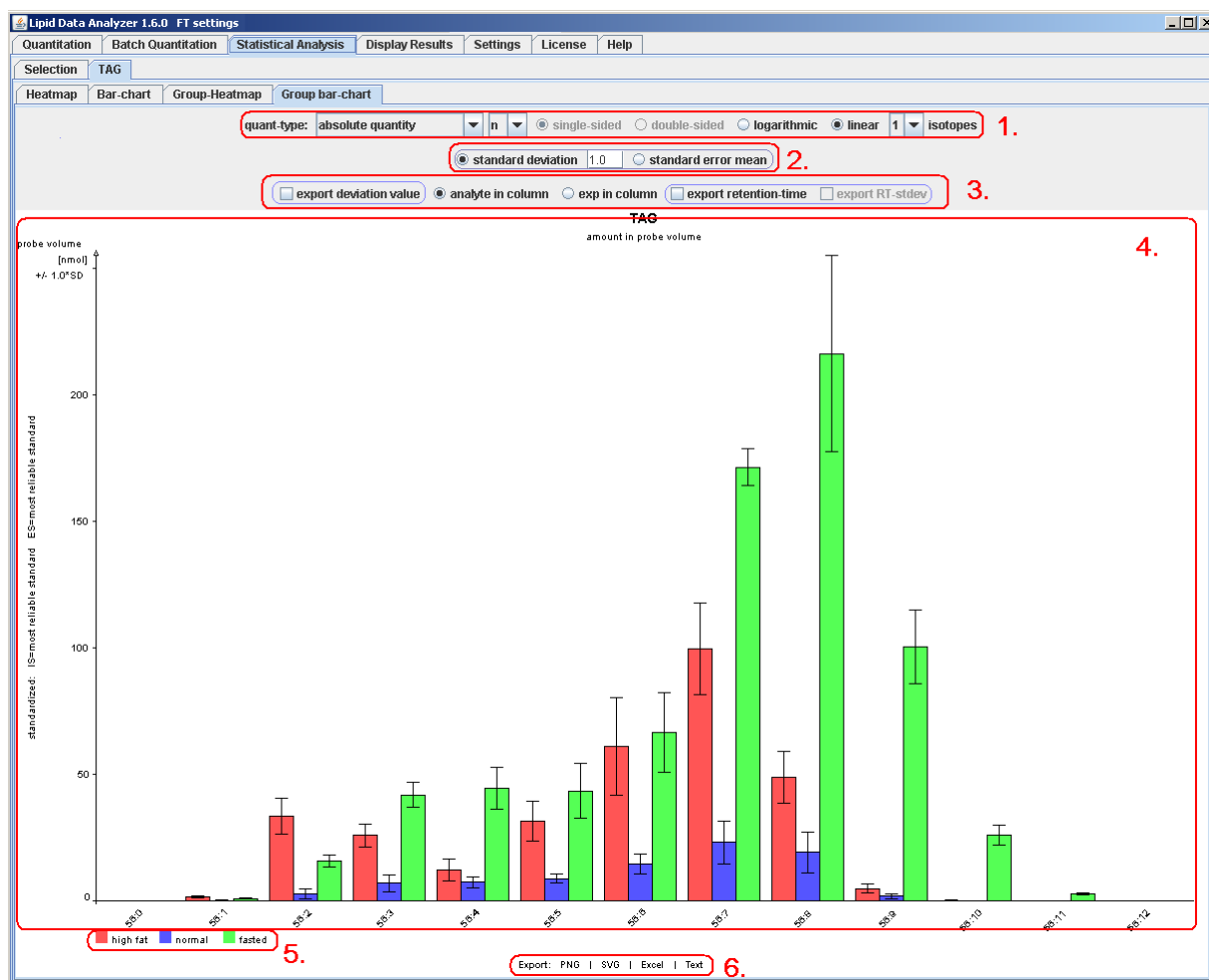
- value type:** A dropdown menu currently showing 'relative value'.
- internal standard correction:** A checkbox that is checked.
- external standard correction:** A checkbox that is checked.
- consider dilution:** A checkbox that is checked.
- divisor unit:** A dropdown menu currently showing 'm'.
- use AU:** A checkbox that is checked.
- Two additional dropdown menus, both showing 'most reliable standard'.
- An **OK** button at the bottom center.

- value type: which value should be displayed and used for the heat map or the bar chart.
 - relative value: Just a value in arbitrary units. This value could be the calculated quantity for the analyte itself, or it could be standardized on standards, depending on the other settings.
 - relative to base peak: The values are calculated relative to the highest found peak of this lipid class.
 - relative to measured class amount: Here percentual values are determined relative to the sum of all analytes for one sample (standards are not considered in this sum) of one lipid class.
 - relative to highest total peak: The values are calculated relative to the highest found peak of all available lipid classes.
 - relative to total amount: The values are calculated relative to the highest found peak of all available lipid classes.
 - amount end-volume: Returns the amount [mol] in the end volume, before the MS measurement. For this value standardization on an internal standard and absolute quantities (see chapter 5.1 number 5) are required.
 - conc. end volume: Returns the concentration [mol/L] in the end volume, before the MS measurement. For this value standardization on an internal standard and absolute quantities (see chapter 5.1 number 5) are required.
 - weight end-volume: returns the weight [gram] in the end volume, before the MS measurement. For this value standardization on an internal standard and absolute quantities (see chapter 5.1 number 5) are required.
 - amount sample-volume: Returns the amount [mol] in the sample volume, before the sample preparation steps (dilution). For this value standardization on an internal and/or external standard and absolute quantities (see chapter 5.1 number 5) are required.
 - conc. sample-volume: Returns the concentration [mol/L] in the sample volume, before the sample preparation steps (dilution). For this value standardization on an internal and/or external standard and absolute quantities (see chapter 5.1 number 5) are required.
 - weight sample-volume: Returns the amount [gram] in the sample volume, before the sample preparation steps (dilution). For this value standardization on an internal and/or external standard and absolute quantities (see chapter 5.1 number 5) are required.
 - relation to measured neutral lipid: This is for the standardization on the total measured lipid mass[g] of the current lipid class measured by MS. For this value standardization

- on an internal and/or external standard and absolute quantities (see chapter 5.1 number 5) are required.
 - relative to sample weight: This is for the standardization on the total weight of the sample mass[g] of the current lipid class measured by MS. For this value standardization on an internal and/or external standard and absolute quantities (see chapter 5.1 number 5) are required.
 - relation to protein content: This is for the standardization on a differently measured protein mass (e.g. measured by a kit). The values for the standardization have to be entered at the absolute quantities (see chapter 5.1 number 5). If there is a standard available, this value is presented in mol/g, else in AU/g.
 - relation to neutral lipid content: This is for the standardization on a differently measured lipid mass (e.g. measured by a kit). The values for the standardization have to be entered at the absolute quantities (see chapter 5.1 number 5). If there is a standard available, this value is presented in mol/g, else in AU/g.
- internal standard correction: is the standardization on an internal standard desired. If yes, it is possible to choose between 3 standardization methods.
 - most reliable standard: Internally developed method, that detects trustworthy standards and out of these a ratio to the other groups is calculated.
 - median: The median of the standards of one sample is taken as reference value for standardization
 - single standards (several ones are available e.g. IS50:0): A standardization on every entered standard is possible.
- consider dilution: Consider the dilution in the calculation.
- divisor unit: If the value is standardized by some value, this value defines the magnitude of the divisor (e.g. pmol/mL; the m for milli is defined by this setting). This setting is applicable for “conc. end-volume”, “conc. sample-volume”, “relation to measured neutral lipid”, “relative to sample weight”, “relation to protein content”, and “relation to neutral lipid content”.
- use AU: use arbitrary units instead of SI units.
- Select molecules: To fade out specific molecules of the heat map.
- Combined chart: Shows selected molecules in one chart over all samples/groups (see chapter 5.3).
- Export options: Defines the information that is exported to Excel or text format. This dialog box consists of a radio button with the options “analytes in column” and “experiments in column”, and additionally a checkbox for the export of the retention time. Additionally for groups, the standard deviation or the standard error of the values and the standard deviation of the retention time can be exported.

5.3 Bar charts

The bar charts are accessible via the heat map or the “Combined chart” button (see chapter 5.2).



The parts encircled in violet are just visible for the grouped view. In general the bar chart painter can be separated into 6 parts:

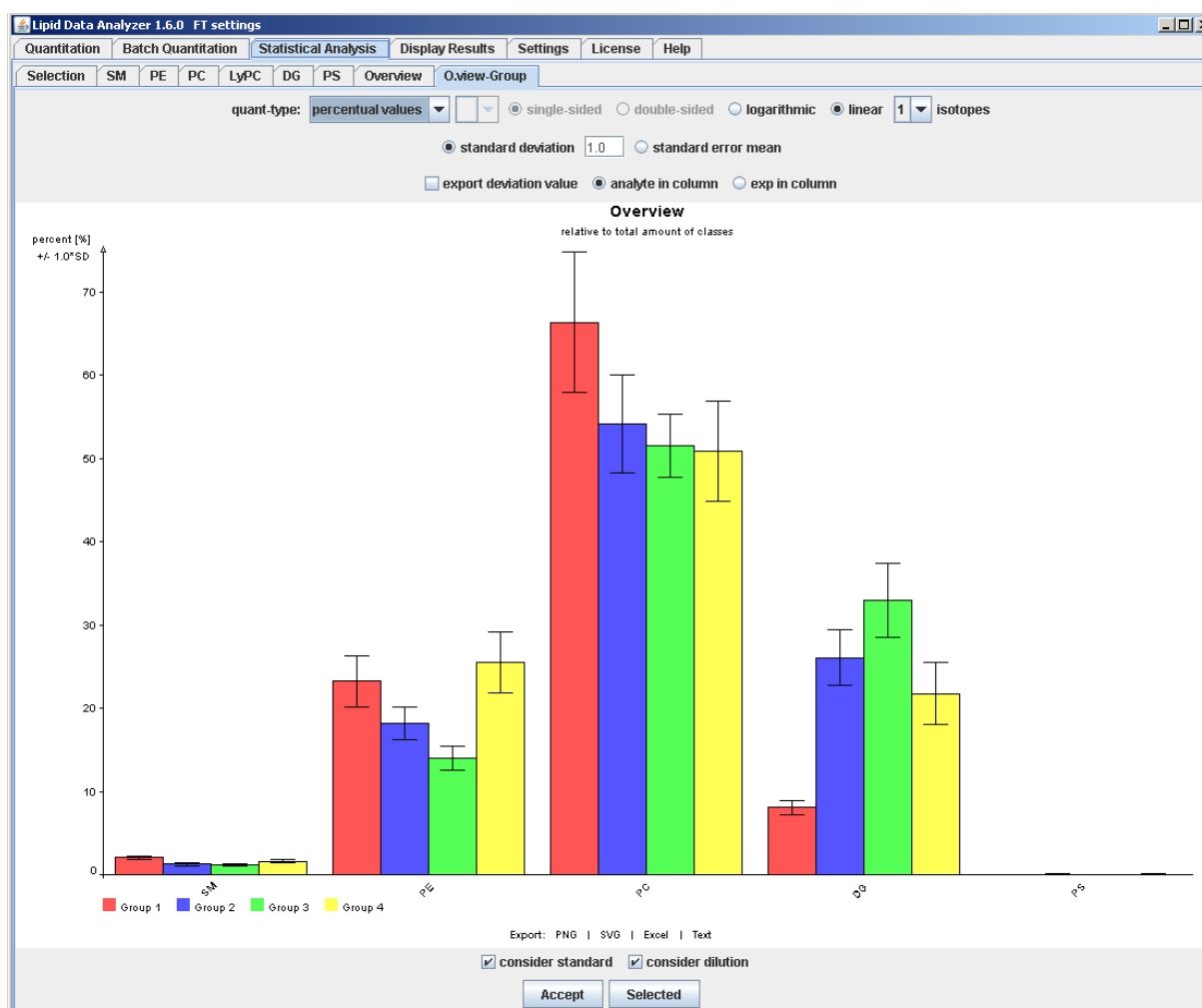
1. Display settings: This is for changing the bar chart display.

- **quant-type**: Determines the type of values that have to be displayed.
 - a. **area absolute**: The absolute area values as they are returned from the algorithm, or if there is a standardization on an internal or external standard these values are the standardized ones.
 - b. **percentual values**: The selected molecules/group/sample correspond to 100% and the values are displayed relative to this amount.
 - c. **area relative to molecule**: The values are displayed relative to the median.
 - d. **area relative to standard**: The values are displayed in relation to the calculated standard (just available if a standardization is selected).
 - e. **absolute quantity**: The values are displayed in the quantity that are selected at 5.2 “Settings button”. The bar chart will always appear in this setting by default, except if “relative value” is selected in the settings.
- **a small select box after the quant-type**: This box is just visible for the “absolute quantity” settings. The input specifies the magnifier (available are none, m, μ , n, p, f, a, %, and ‰).
- **radio button single-sided double-sided**: just enabled for “area relative to molecule” and “area relative to standard”. For single sided the values always start with 0. For double sided the area is relative to the median/standard, and lower values are painted in negative direction and higher values in positive direction.
- **radio button logarithmic linear**: Available for “area absolute”, “percentual values”, and “absolute quantity”. This option switches between linear and logarithmic (log10) display.
- **radio button log10 log2**: Available for “area relative to molecule” and “area relative to standard”. Toggles between decadic- and dual logarithm.
- **isotopes**: Determines the amount of isotopes that should be used for the quantitative value.

2. Deviation settings: Available just for groups. This value defines how the error bars should be displayed. The deviation can be displayed at any multiplicative value (1.0 is default) and the standard error mean can be displayed.
3. Export settings: Settings for the export in Excel and Text format. These are the same like in chapter 5.2 “Export options”.
4. Bar char: The main painting area. At the top in the center, the lipid class and the value type is displayed. On the left, written vertically, the used standardization procedure is displayed. At the top of the vertical axis, the current unit is displayed. In the horizontal axis, the molecule, samples or respectively groups are displayed.
5. Color selection: In this legend the used color and the corresponding item is displayed. If the mouse is clicked on this legend, the color can be changed.
6. Export bar: The picture of the heat map or the data respectively can be exported in the following formats: PNG = portable network graphics; SVG = scalable vector graphics; Excel = Microsoft Excel; Text = text based format (tab delimited).

5.4 Overview bar chart

If there are several lipid classes available, a bar chart for the overview of the contribution of the separate lipid classes is available.

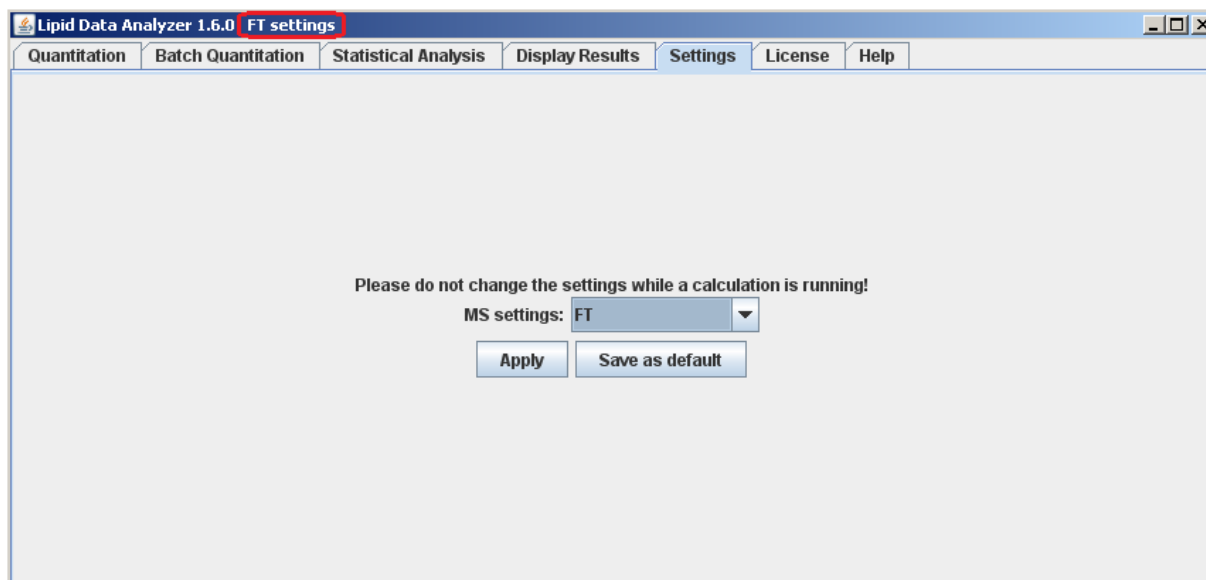


This overview has the option to “consider standard” (the current standard settings are taken) and to “consider dilution”. If the “consider standard” is selected, an overview is displayed containing the standards used. If for one lipid class no standard is available, the class is neglected (like here LyPC). Without the “consider standard” these classes would have been displayed. Attention: without the “consider standard” it is assumed that all classes ionize equally!

As quant-type, just “area absolute” and “percentual values” is available.

6 Settings

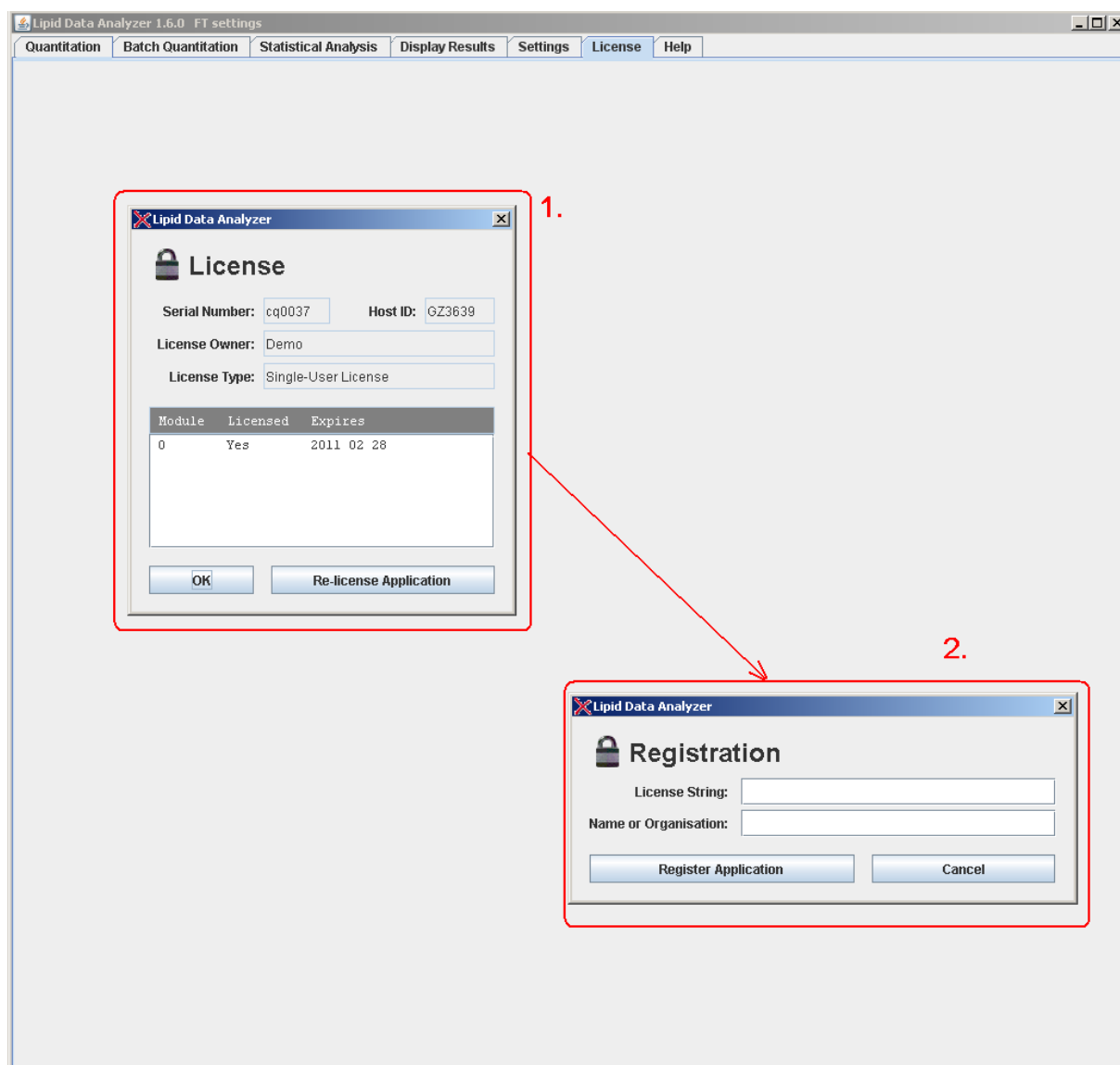
The LDA can be used for several MS machines. The Settings tab has been introduced to change these settings easily. **The settings must not be changed while a calculation is running!**



The current settings are displayed in the title of the application. The “Apply” button applies the settings to the current session. The “Save as default” button sets the selected settings as default, which will be loaded automatically at the next start of the LDA.

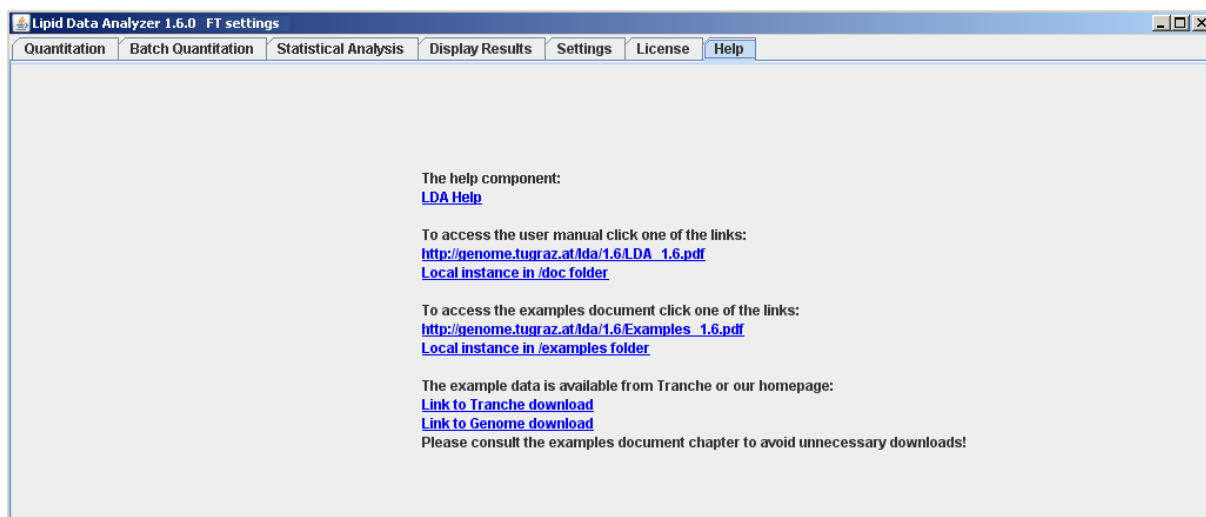
7 Licensing

In order to view the license, the tab “License” has to be clicked. Then a dialog window shows the details about the current license (1. in next figure). If you acquired a new license, click on “Re-license” and a dialog window appears where the license hash plus the name of the institution has to be entered.

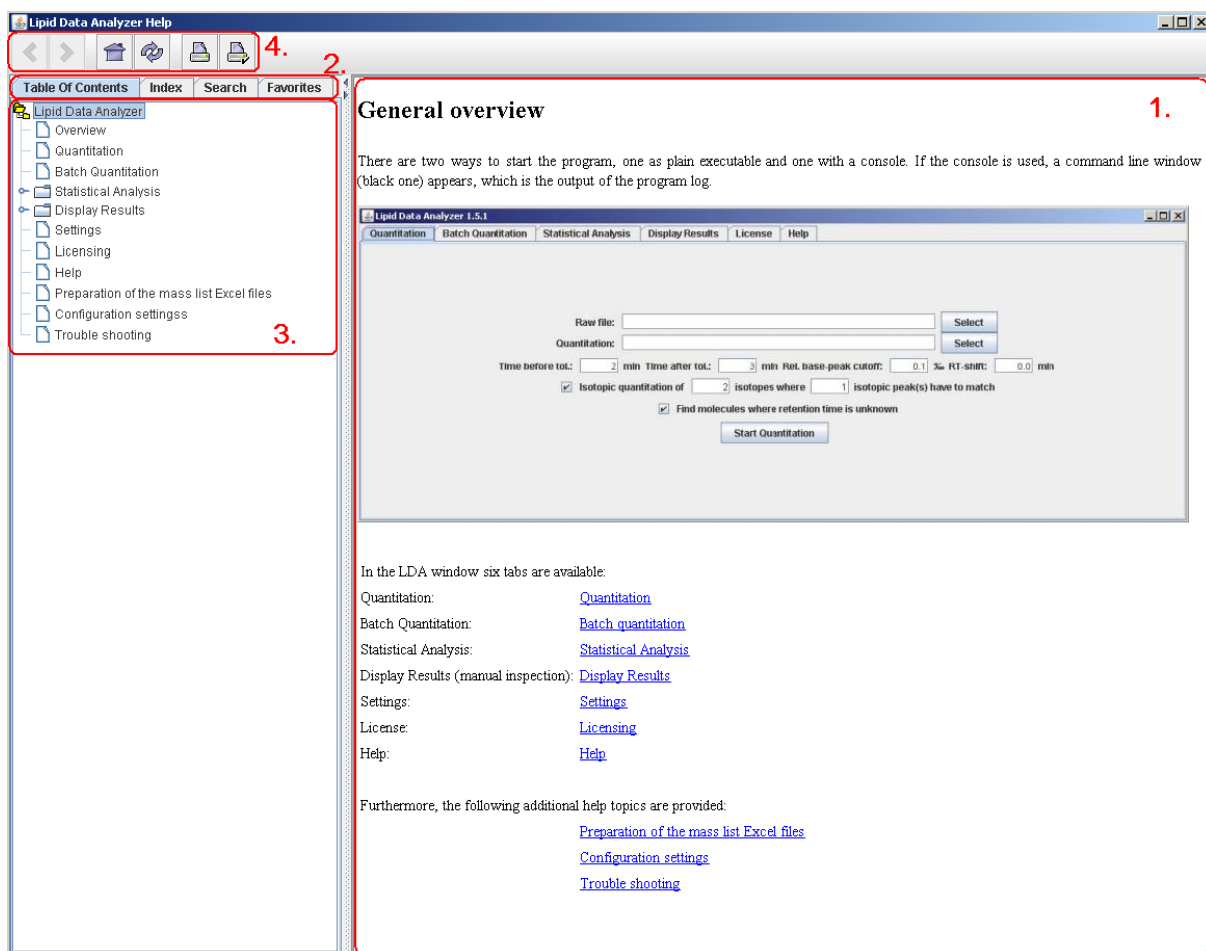


8 Help

The help page contains links to several resources. The first one is to an online help component (LDA help, explained in detail after the next figure). Furthermore, it provides links to the user manual (this file), the Examples.pdf (file that shows the functionality of the software with the aid of the published example data), and the example data that can be downloaded from the Tranche repository. The links to PDFs are provided to the home page and to local files that come with the installation package.



After the link LDA help had been clicked, a window containing the help component pops up.



The help component contains 4 parts:

1. Information panel: It contains information about the current help topic. Cross links to other help topics are inside these pages.
2. Tabs containing different ways to search for the desired help topic:
 - Table of contents: The topics are organized like in a table of contents in a document.
 - Index: The help topics are ordered alphabetically.
 - Search: Enables to search the help for words.

- Favorites: Organization of favorite pages of the help. The user has to right click with the mouse in the tabbed pane of favorites and can add help pages to this menu.
3. Navigation tree for the different tabs explained before.
 4. Control bar. This bar contains buttons for “previous page”, “next page”, “home”, “reload page”, “print”, and “print page setup”

Appendix A - Preparation of the mass list Excel file

The Excel file containing the mass lists must follow some conventions.

Name	dbs	C	H	O	P	N	D	M	MH+	mass(form[*NH4] name[NH4])	mass(form[*N2C2H7] name[NH4+ACN])	tR (min)
48	0	51	98	6	0	0	0	0	806,736313	807,743589	824,77013	865,79668
48	1	51	96	6	0	0	0	0	804,720663	805,727939	822,75448	863,78103
48	2	51	94	6	0	0	0	0	802,705013	803,712289	820,73883	861,76538
48	3	51	92	6	0	0	0	0	800,689363	801,696639	818,72318	859,74973
48	4	51	90	6	0	0	0	0	798,673713	799,680989	816,70753	857,73408
50	0	53	102	6	0	0	0	0	834,767613	835,774889	852,80143	893,82798
50	1	53	100	6	0	0	0	0	832,751963	833,759224	850,78578	891,81233
50	2	53	98	6	0	0	0	0	830,736313	831,743589	848,77013	889,79668
50	3	53	96	6	0	0	0	0	828,720663	829,727939	846,75448	887,78103
50	4	53	94	6	0	0	0	0	826,705013	827,712289	844,73883	885,76538
50	5	53	92	6	0	0	0	0	824,689363	825,696639	842,72318	883,74973
50	6	53	90	6	0	0	0	0	822,673713	823,680989	840,70753	881,73408
Ex-151	0	54	104	6	0	0	0	0	848,783263	849,79054	866,81708	907,84363
52	0	55	106	6	0	0	0	0	862,798913	863,80619	880,83273	921,85928
52	1	55	104	6	0	0	0	0	860,783263	861,79054	878,81708	919,84363
52	2	55	102	6	0	0	0	0	858,767613	859,774889	876,80143	917,82798
52	3	55	100	6	0	0	0	0	856,751963	857,759224	874,78578	915,81233
52	4	55	98	6	0	0	0	0	854,736313	855,743589	872,77013	913,79668
52	5	55	96	6	0	0	0	0	852,720663	853,727939	870,75448	911,78103
52	6	55	94	6	0	0	0	0	850,705013	851,712289	868,73883	909,76538
52	7	55	92	6	0	0	0	0	848,689363	849,696639	866,72318	907,74973
52	8	55	90	6	0	0	0	0	846,673713	847,680989	864,70753	905,73408
54	0	57	110	6	0	0	0	0	890,830213	891,83749	908,86403	949,89058
54	1	57	108	6	0	0	0	0	888,814563	889,82184	906,84838	947,87493
54	2	57	106	6	0	0	0	0	886,798913	887,80619	904,83273	945,85928

The file consists of several sheets, each representing a lipid (molecule) class. If you want to analyze several lipid classes, just make several sheets. Second, there must be a header column. The header column must contain the following keywords, otherwise it won't be accepted: Name, tR (min), at least one mass(...), and at least one element column. Except of the "Name" column, which must be the first column, the order of the columns does not matter for the Excel file. Detailed description of the example in the figure (columns from right to left):

- "Name": The name of the analyte
- "dbs": This column is optional. If the analyte name contains double bonds in it, they should be entered in this column. The values of this column must be integer format. If the name of the analyte does not contain any double bonds, just remove this column. Generally, it is possible to enter the double bonds in the "Name" column directly; however, the sorting of the analytes in the results is improved if this column is used.
- Element columns ("C", "H", "O", "P", "N", "D"): In this column the elemental composition (chemical formula) of the analyte has to be entered. The values of this column must be in integer format. If there is a value in float format (e.g. the M column), the column will be ignored as elemental column (a warning at the beginning of the quantitation appears). A column is an elemental column if it contains just one or two characters, beginning in upper case and ending lower case (e.g. H, Na, He, P, etc.). If you require new elements, just add an additional column and enter the probable occurrences in nature in the elementconfig.xml file (see Appendix B - Configuration settings where to find this file).
- Additional calculation columns (e.g. "M", "MH+"): These columns are not read by the application, and are just a help for the user to calculate the masses. Do not mind that the comma in the float value is a "," and not a ".", the Excel I used for the screen shot is using the German notation.
- Mass columns ("mass(...)"): A mass column contains the mass values to be quantified. There can be several mass columns because of different modifications/adducts. A mass column must contain the keyword "mass" followed by round brackets. Within the round brackets, there must be the keywords "name" and "form" followed by square brackets [...]. In the square brackets after "name" the display name has to be entered; after "form" the chemical formula. The chemical formula of the modification/adduct must be in the notation \$element_symbol\$element_amount\$(e.g. H4). If no element amount is entered, it is assumed to be one (e.g. NH4=N1H4). If elements are lost by the modification enter a "-" in front (e.g. -NH4 would correspond to a loss in ammonium). **A space is interpreted as +**, that means if you enter "-N H4" this would correspond to -N+H4. Additionally, the charge state can be entered by charge=\$charge_state\$. If this information is missing in the brackets, a charge state of 1 is assumed.

- tR (min): This is the retention time column. The input of the retention time is optional. If it has not been entered, the whole chromatogram is searched for a peak (if the checkbox “Find molecules where retention time is unknown” is checked).

Third, it is possible to enter a general start and stop retention time for a lipid class. The notation is “Start-RT: ...” and “Stop-RT: ...”. These cells can be anywhere before the row with the header row.

Appendix B - Configuration settings

This chapter covers the manual adaption of LDA setting parameters. General settings can be found in the .settings file, and machine specific ones in the LipidDataAnalyzer.properties file. For each machine such a properties file is present in the properties folder in the installation directory. It is recommended to change the settings for the machine in the corresponding file in the properties folder and, after saving this file, to select the machine and click on “Save as default” (see chapter 6). If the settings are changed directly in the LipidDataAnalyzer.properties, these settings are overwritten the next time somebody clicks on “Save as default”. The .settings file can be edited directly. In the LipidDataAnalyzer.properties file, not only algorithmic parameters are stored, but default settings for the display as well. The changes take effect after the properties file had been stored and the LDA is restarted.

Path settings (.settings)

This part contains path settings to executables or required files.

- ReadWPath: path to ReadW executable that is required if Thermo Finnegan RAW files are used directly.
- MassWolfPath: path to the MassWolf executable if the Waters .raw directories should be used directly.
- ElementConfig: path to the required elementconfig.xml file that contains the used chemical elements and the occurrence of their isotopes in nature.

Look & feel settings (.settings)

- LookAndFeel: This value is by default “system”, which corresponds to your system specific look & feel (e.g. Windows). If you prefer Java look & feel change this value to “java”, then you have the same look & feel like in this manual.

Default quantitation settings (.properties)

This stores default settings for the quantitation input page.

- basePeakCutoff: The default value for the “Rel. base-peak cutoff” (see chapter 2).

Excel result file settings (.settings)

- OverviewExcelWorkbook: If “true”, an overview tab is created in the Excel file.

3D viewer default settings (.properties)

For the resolution of the 3D viewer, default settings can be entered.

- threeDViewerDefaultTimeResolution: The default time resolution of the viewer.
- threeDViewerDefaultMZResolution: The default m/z resolution of the viewer.

Standards prefix settings (.settings)

The field for the prefixes of the standard can be set by default (see chapter 5.1).

- ISDefaultInput: Default prefix for the internal standard
- ESDefaultInput: Default prefix for the external standard

mzXML → chrom settings(.properties)

These are settings for the translation from the mzXML to the chrom format.

- `maxFileSizeForChromTranslationAtOnce`: An approximate value how many MB can be read at once into the memory. The mzXML is translated into the chrom in several rounds. If there is e.g. a mzXML file with 200MB and an m/z range from 600-1400, then this file will be translated in two rounds if this parameter is 100: first round: masses from 600-1000; second round masses from 1000-1400. Thus this parameter does not directly correspond to the required memory, the required memory is normally more. If this parameter would have been 200, the whole file would have been translated in one round. The reason for this parameter is to permit the translation of huge files on machines with low RAM. If this parameter is increased, the translation time is decreased. If this parameter is too high a `java.lang.OutOfMemoryException` will be thrown and the file cannot be translated.
- `chromMultiplicationFactorForInt` and `chromMultiplicationFactorForInt`: This parameter define the highest resolution of the chrom file (e.g. `chromMultiplicationFactorForInt=1000` and `chromMultiplicationFactorForInt=1` would mean a max resolution of 0.001Da; `chromMultiplicationFactorForInt=1000` and `chromMultiplicationFactorForInt=5` would mean a max resolution of 0.005Da; `chromMultiplicationFactorForInt=100` and `chromMultiplicationFactorForInt=1` would mean a max resolution of 0.01Da).

General algorithm settings(.properties)

This section contains the most general settings of the algorithm. A detailed description of the parameters of the 3D algorithm is beyond the scope of a user manual. The parameters are described in the `LipidDataAnalyzer.properties` file.

- `coarseChromMzTolerance`: the +/-m/z range that is used for the extraction of the first chromatogram of the algorithm which is displayed in the 2D viewer (see chapter 4.3).
- `chromSmoothRange`: the smooth range in seconds for Savitzky-Golay filter of the first chromatogram.
- `chromSmoothRepeats`: the smooth repeats for Savitzky-Golay filter of the first chromatogram.
- `use3D`: If yes, the 3D algorithm is used by default for the quantitation, else the 2D algorithm.

Memory settings(.bat, .sh, or. vmoptions)

The memory settings for the executables can be changed in the `Lipid Data Analyzer.vmoptions`, and for the console in the `.bat` or `.sh`, respectively. The `-Xmx` parameter corresponds to the maximally reservable memory, and `-Xms` the immediately reserved memory. The `-Xms` value should be approximately half of the `-Xmx` value.

Appendix C – mzTab settings

LDA provides the possibility to export metadata in the mzTab format. None of this information is mandatory for the mzTab export, but helps for the reproducibility of your data. All changes in these settings require a restart of the LDA application.

The metadata is split in two configuration files: i) `.properties` file containing information about the mass spectrometer ii) `mzTab.properties` containing experiment specific information such as affiliations and sample origin information.

MS instrument specific settings (.properties)

In the current implementation, the following metadata can be exported in mzTab format:

- `mzTabInstrumentName`: the name of the instrumentation
- `mzTabInstrumentIonSource`: the ion source of the instrumentation (e.g. ESI, or MALDI)
- `mzTabInstrumentAnalyzer`: the analyzer of the instrument (e.g. ion trap, FT-ICR, etc.)
- `mzTabInstrumentDetector`: the ion detector (e.g. electron multiplier)

Examples are provided within the files. Be sure for all parameters that the entries contain always three “,” otherwise this setting will be neglected. If you are using (as recommended) controlled vocabulary, please crosscheck its validity (<http://www.ebi.ac.uk/ontology-lookup/>).

Other mzTab settings (mzTab.properties)

All the other metadata for mzTab is stored in this file. The entries here must contain three commas such as in the MS instrument specific settings. The exception is the contact settings, which must contain at least two commas. The first comma separates the name, the second one the email address, and the rest of the commas are included in the affiliation. The contacts have to start with one, and MUST have consecutive numbers, e.g.

contact_1=\$name_first_contact\$, \$email_first_contact\$, \$affiliation_first_contact\$

contact_2=\$name_second_contact\$, \$email_second_contact\$, \$affiliation_second_contact\$.

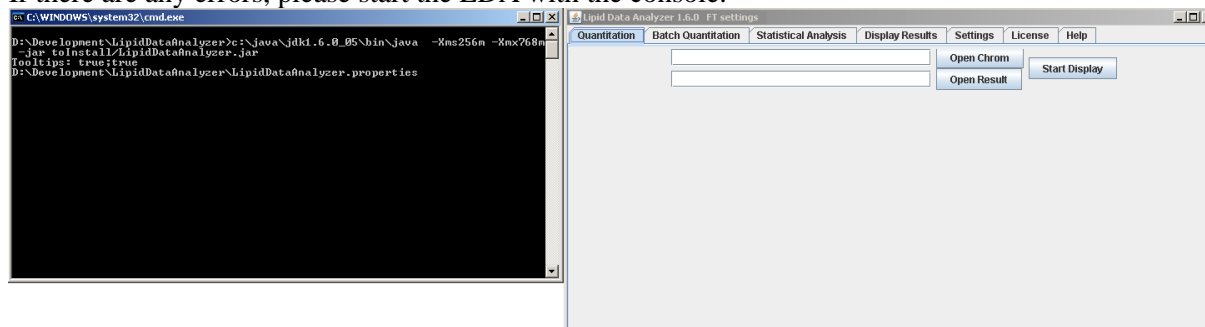
Other supported metadata settings of this file are:

- species: from which species originates the sample
- tissue: from which tissue originates the sample
- celltype: the cell type of the sample

If any of the values are unknown or cannot be assigned, they should be commented or left out. The export to mzTab is possible with missing values.

Appendix D - Trouble shooting

If there are any errors, please start the LDA with the console.



Then, please try to reproduce the error! Afterwards, please make a copy of the console output and send it with a detailed description of the preceding user interactions to the developers (lda@genome.tugraz.at).