MAIA - MicroArray Image Analysis
Version 2.75
User Manual

MAIA download page: http://bioinfo.curie.fr/projects/maia/

Author(s): Eugene Novikov (Institut Curie)
E-mail: Eugene.Novikov@curie.fr

Installation

MAIA can be downloaded from the MAIA download page http://bioinfo.curie.fr/projects/maia/

MAIA runs on Windows platforms 95/98/Me/NT/2000/XP and needs the Java Runtime Environment (JRE) to be installed: (http://www.java.com/en/download/)

Click MAIA Setup 2.75.exe to start the MAIA 2.75 installer and follow the instructions*.

MAIA 2.75 installation creates a "Curie/MAIA 2.75" folder in the list of Programs of the Windows Start menu. This new folder contains the following entries:

- MAIA 2.75 starts Microarray image analysis software;
- User Manual is a user manual pdf file;
- Uninstall MAIA will remove MAIA from your computer.

Installation procedure may also create a "MAIA" icon on your Desktop.

*) Installation procedure asks about the default size of the JVM (Java Virtual Machine) memory allocation pool. It is recommended to set it as large as possible, but not larger than the amount of available RAM.

From MAIA 2.7 to MAIA 2.75

- *Output File Format:*
  - The order of the outputted fields can be defined using Drag-and-Drop of the table rows;
  - Double click on the selected row appends the corresponding field to the list of the outputted fields;
  - Any field can be set at the top or bottom of the list of the outputted fields using the corresponding items of the popup menu: {“Move to First”, “Move to Last”};
  - The definition of the *Output File Format* can be imported from a previously exported analysis results file.
- *Diagnostic Plots:*
  - Switch between the linear and log scales is performed via the popup menu;
  - Number of graph panels in the rows, and the height of the graph panels can be different for the *Quality Histograms* and *Results Plots* panels;
  - Lowess normalization has been implemented for the *Scatter plot* and *M-A plot*. It is available via the Menu Item “Lowess Fit” from the popup menu at the *Scatter Plot* and *M-A plot* panels. The normalized ratio values can be set for output using the Menu Item “Set Normalized” from the same popup menu.

- The popup menu {“All to Run” “All to Skip”} has been added to the *Batch Processing Window* to include or to not include all files in the batch run.
- Binary Experimental files are *Zip* compressed.

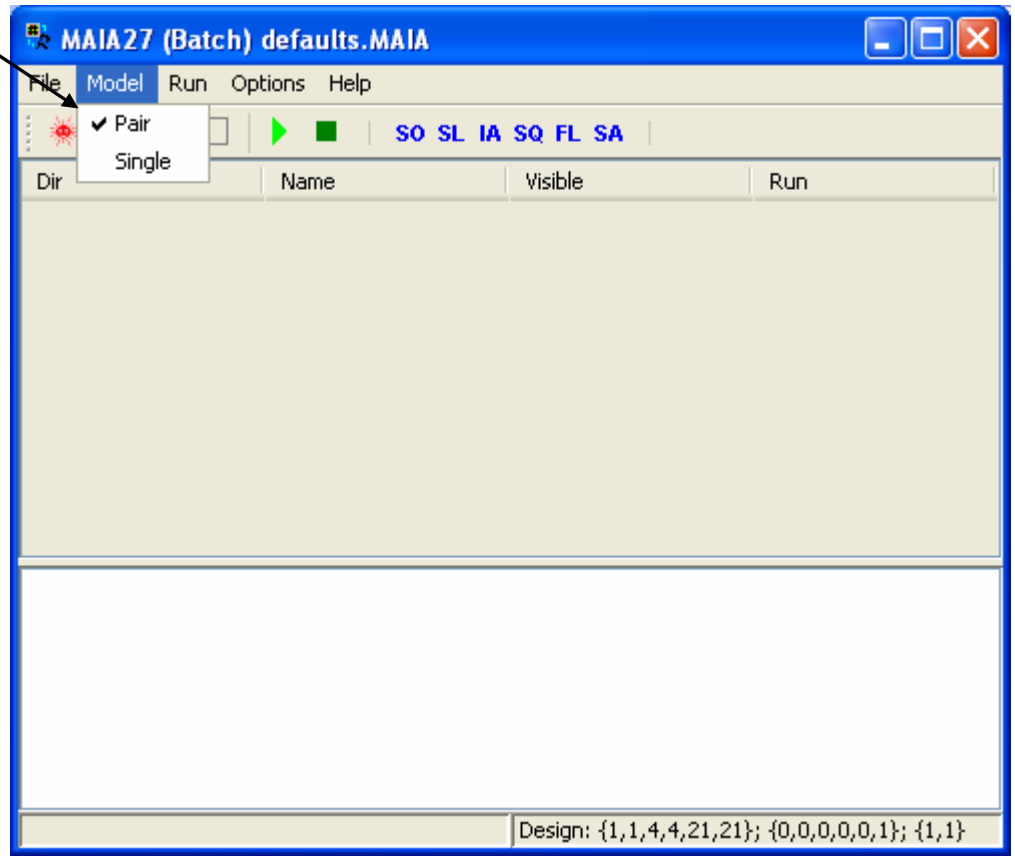
Batch Processing Window

The Menu “Model” allows one to select the image analysis model:

…

“Model|MAIAPair”
Two Color Image Analysis

“Model|MAIASingle”
One Color Image Analysis

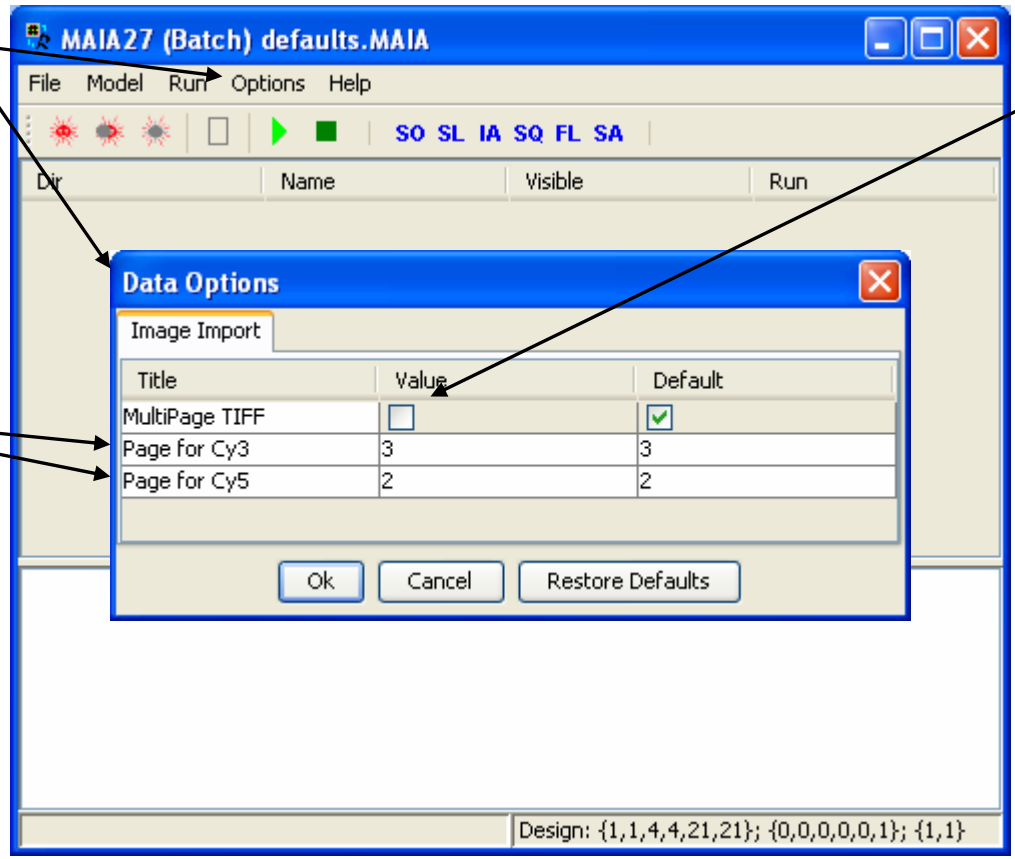


Two Color Image Analysis

Data Import Settings

To define the format of the microarray image files select the Menu Item "Options|Data Options" (Alt+D).

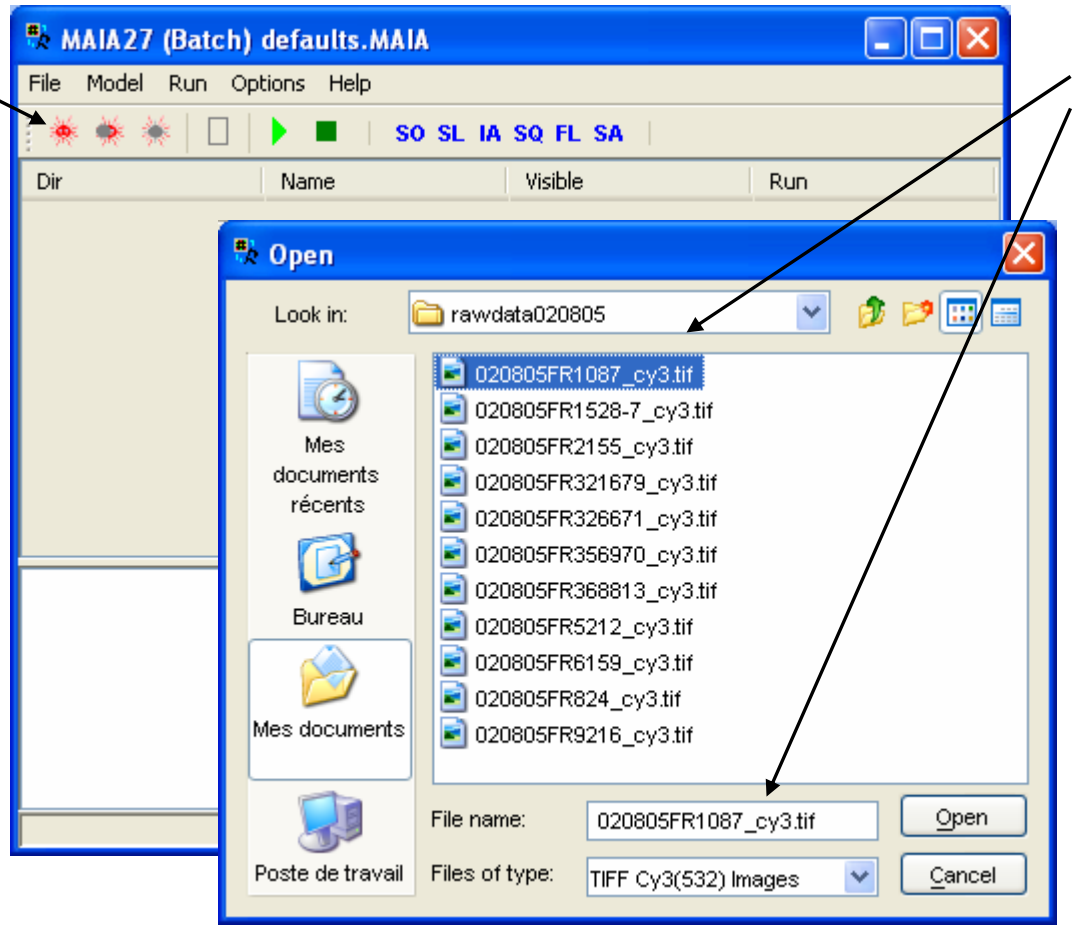
For multi-page TIFF, specify the pages for the Cy3 and Cy5 images.



Two options are available:
(i) Cy3 and Cy5 TIFF images are packed into one multi-page TIFF file (checked);
(ii) Cy3 and Cy5 TIFF images are stored in separate files (unchecked).

File Name Selection

Use the Toolbar button "New Experiment" or the Menu Item "File|New Experiment" (Ctrl+I) to select microarray images.

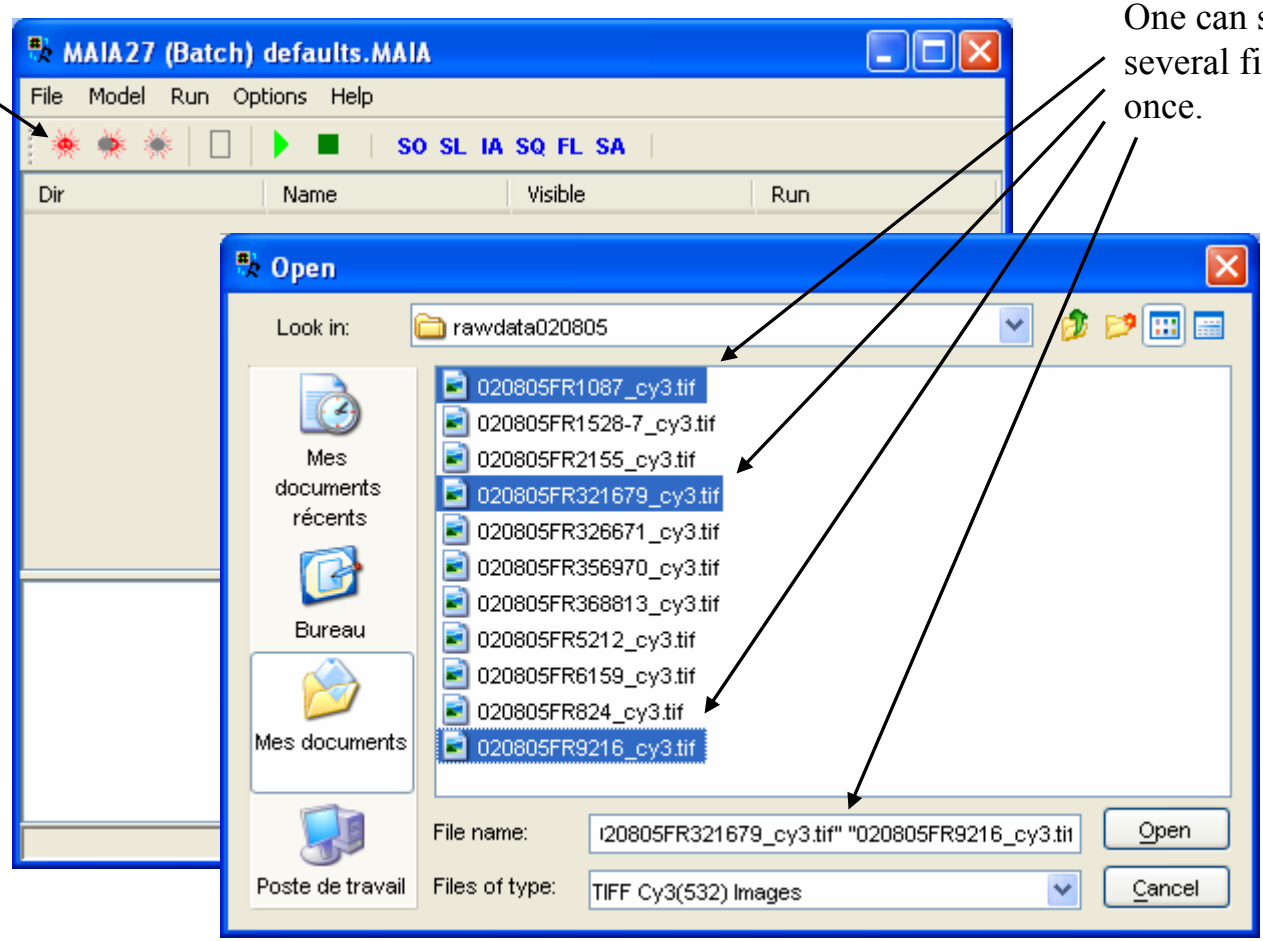


When single-page TIFF files are used, File Browser shows up only Cy3 file names. The correspondent Cy5 file name will be downloaded automatically.
In this case filenames for the pair of Cy3 and Cy5 images must differ only by the suffix: "cy3" or "532" for Cy3 images, and "cy5" or "635" for Cy5 images.

For multi-page TIFF, filenames can be arbitrary.

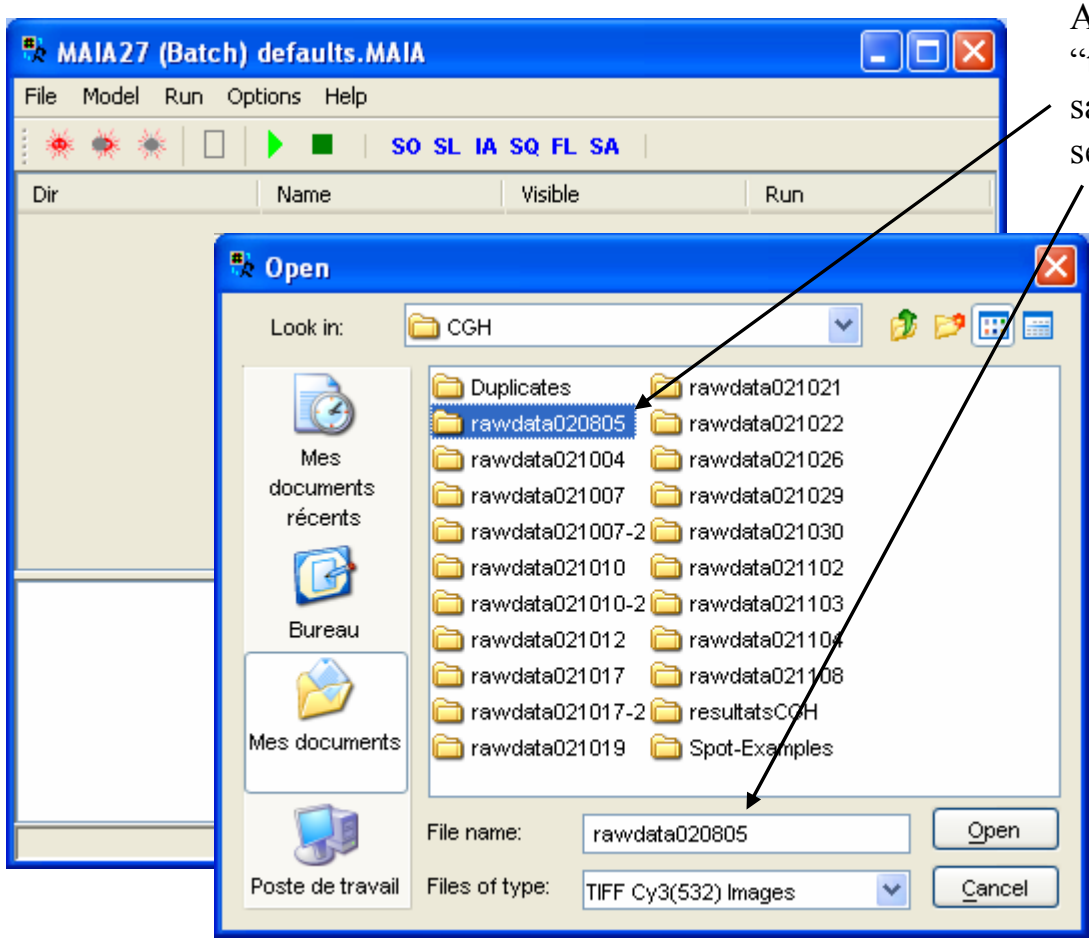
Multiple File Name Selection

Using the Toolbar button “New Experiment” or the Menu Item “File|New Experiment” (Ctrl+N) more files can be added into the table.



One can select several filenames at once.

Directory Selection

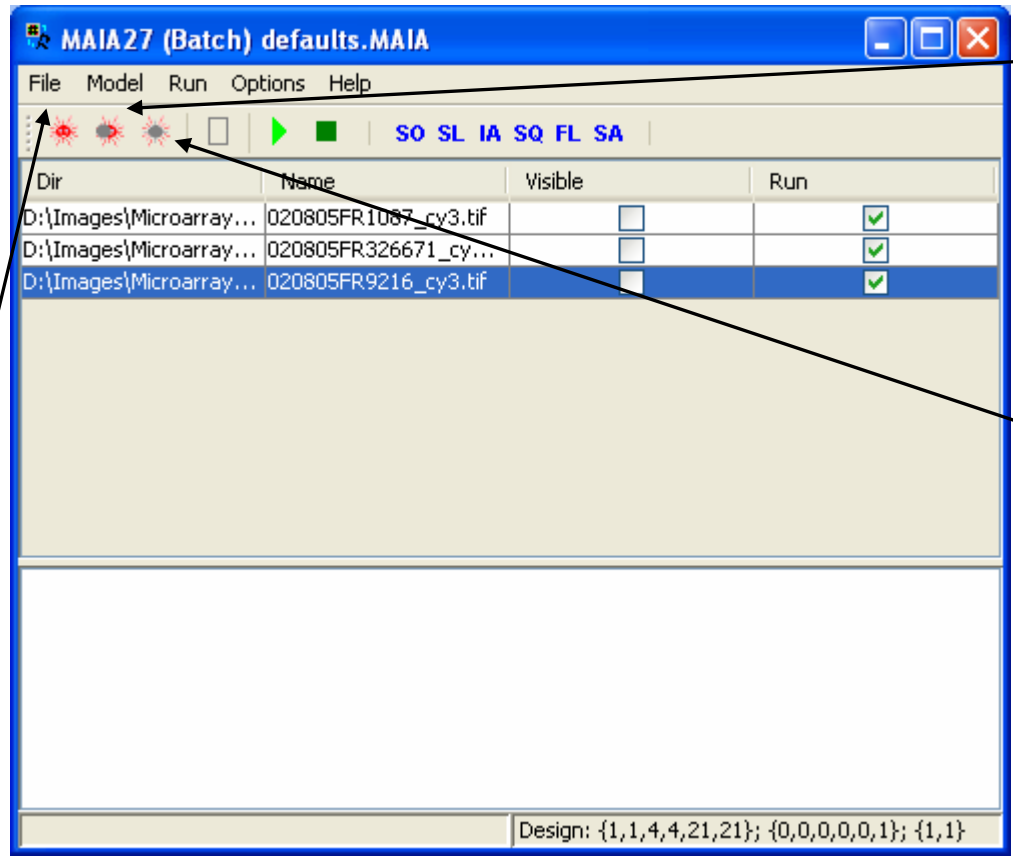


All TIFF files (extension "tif"/"tiff") from the same directory can be selected.

Batch of File Names

The selected filenames appear in the table.

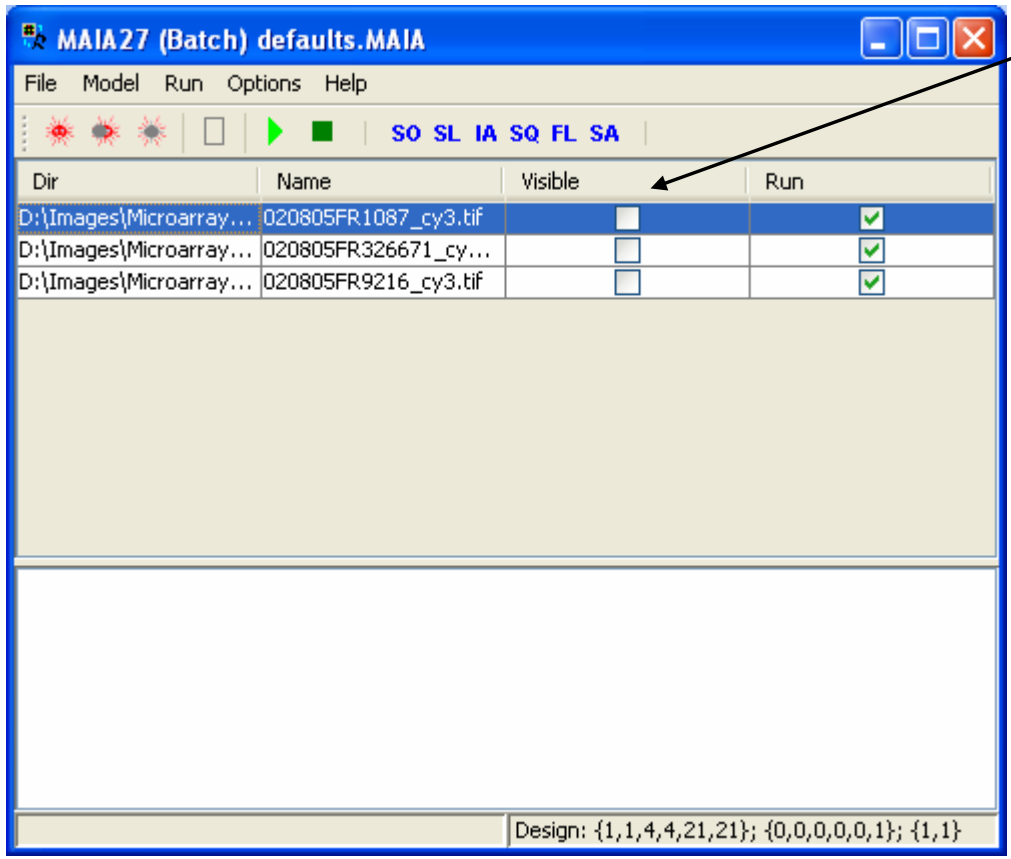
The whole batch (a list of files and accompanying options) can be saved on the disk (using the Menu Item "File|Save Group …" (Ctrl+S)) to be able to restore it (using the Menu Item "File|Load Group …" (Ctrl+O)) to reanalyze the batch.



To remove filenames from the batch one may use the Toolbar button "Remove Experiment" or the Menu Item "File|Remove Experiment" (Ctrl+E).

The toolbar button "Remove All Experiments" or the Menu Item "File|Remove All Experiments" (Ctrl+Alt+E) will remove all filenames from the batch.

Ready for Analysis



Check the "Visible" field to open (download) an image .

Main Processing Window

Three panels are created: Ratio image, Cy3 and Cy5 channel images.

Another pair of images (Cy3/Cy5) can be downloaded using the "Load Data …" button from the Toolbar or the Menu Item "File|Load|Data …" (Ctrl+O).

For the new images, image file format (i.e. multi-page TIFF versus single-page TIFF) can be changed using the Menu Item "Options|Data Options" (Alt+D).

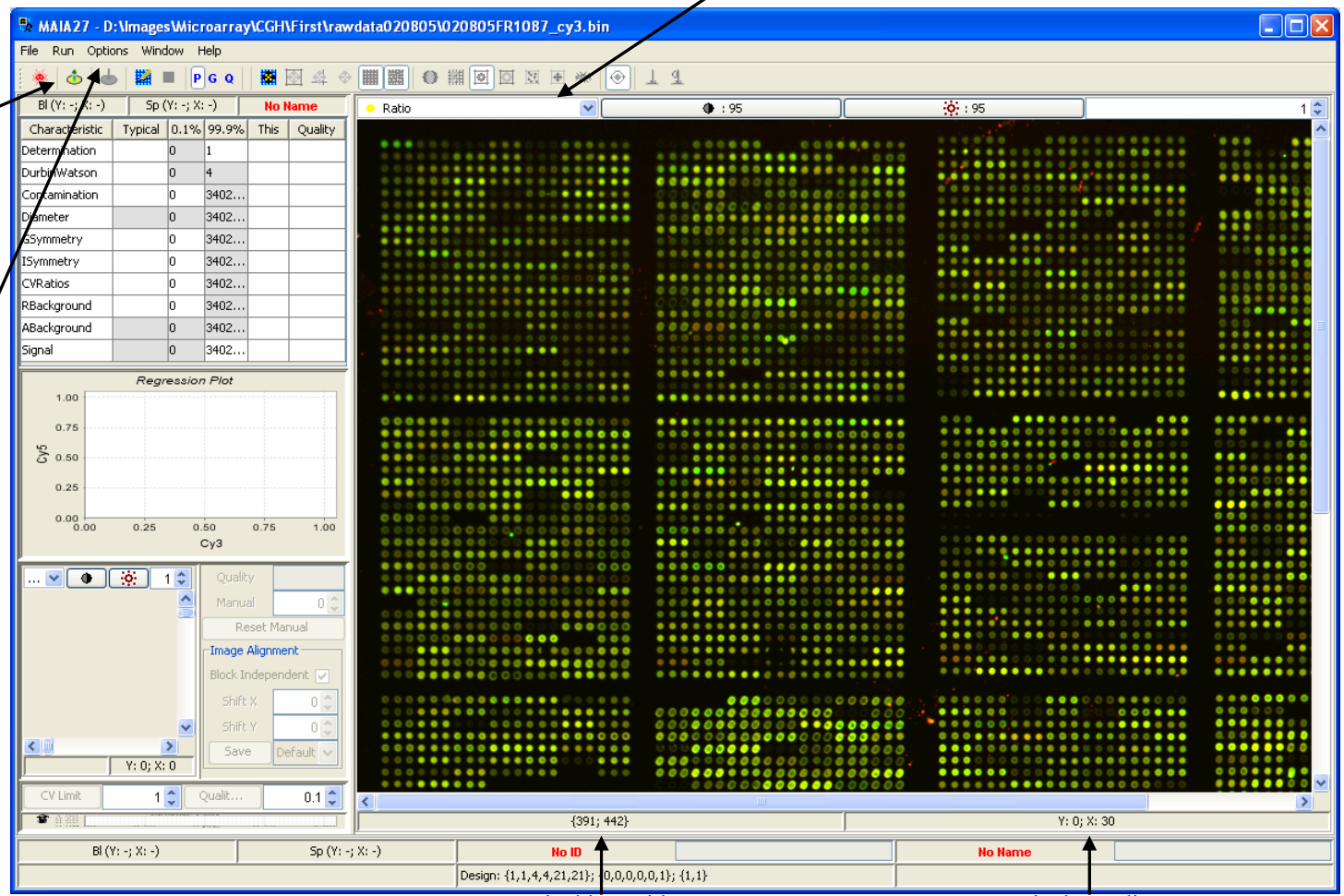
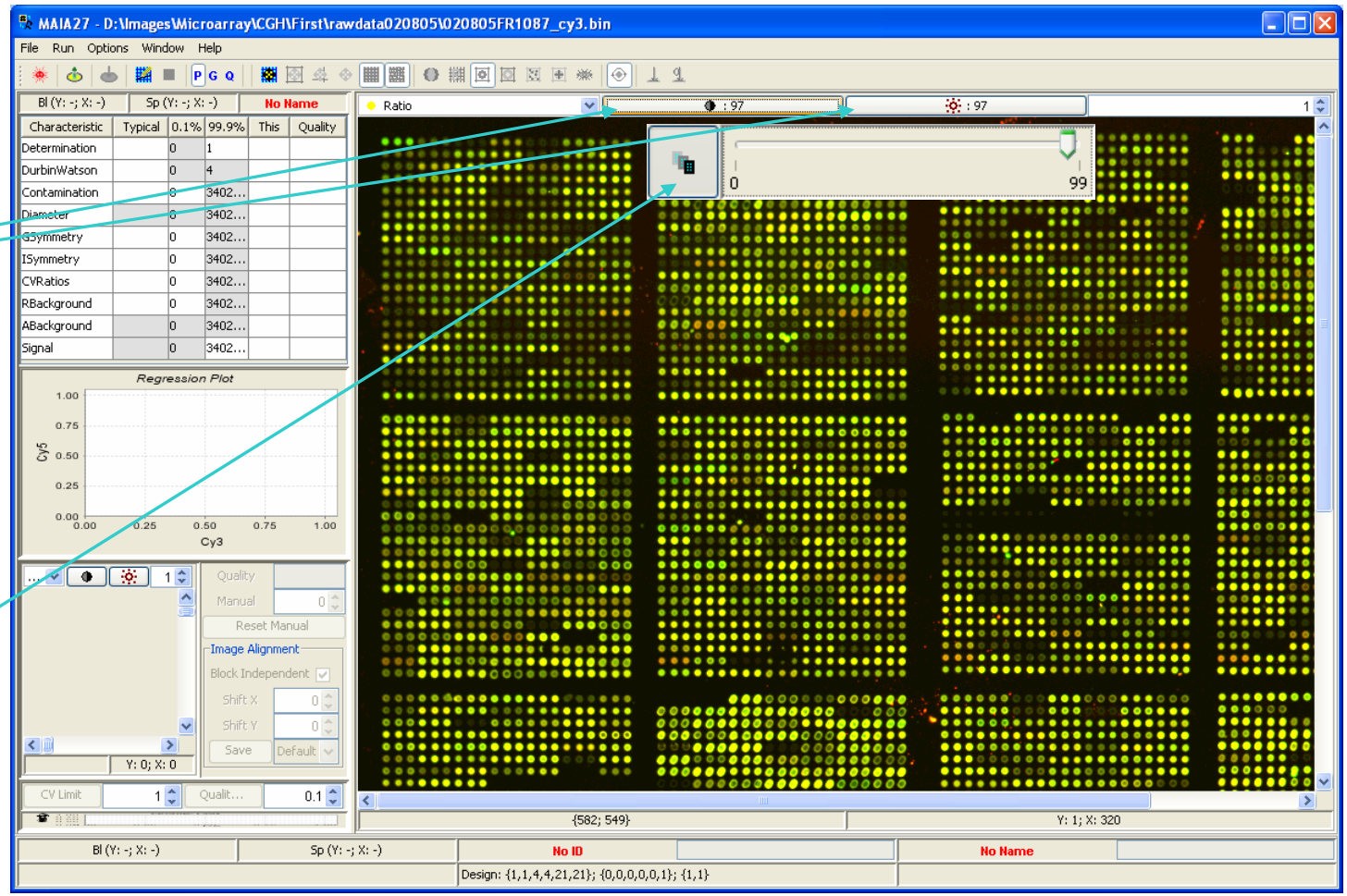


Image Visualization Settings

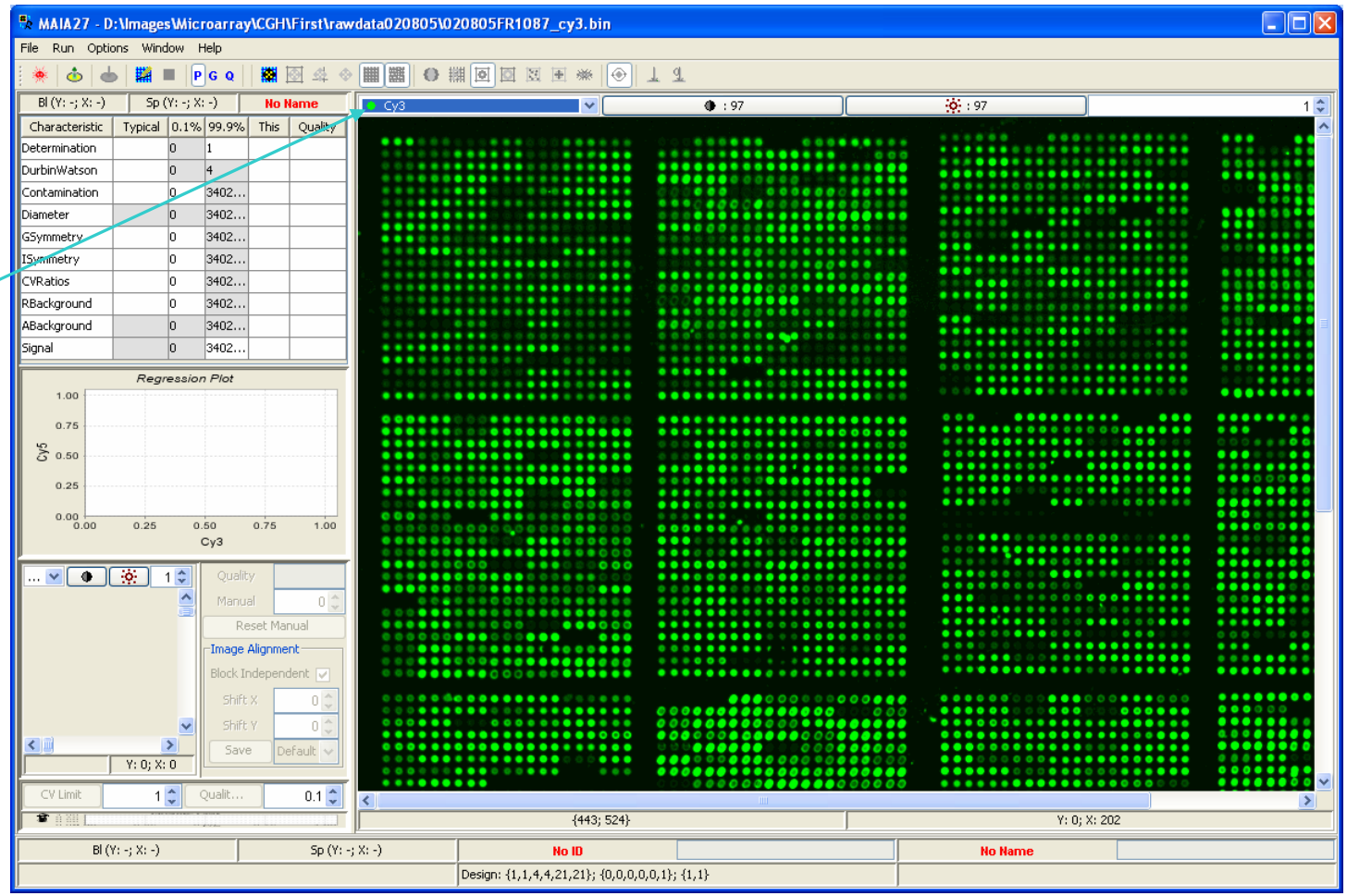
"Contrast" and "Brightness" controls can be used to adjust brightness and contrast of the images.

Brightness and contrast can be adjusted either independently for each color channel (the button "All Images" is off) or simultaneously for all channels (the button "All Images" is on).



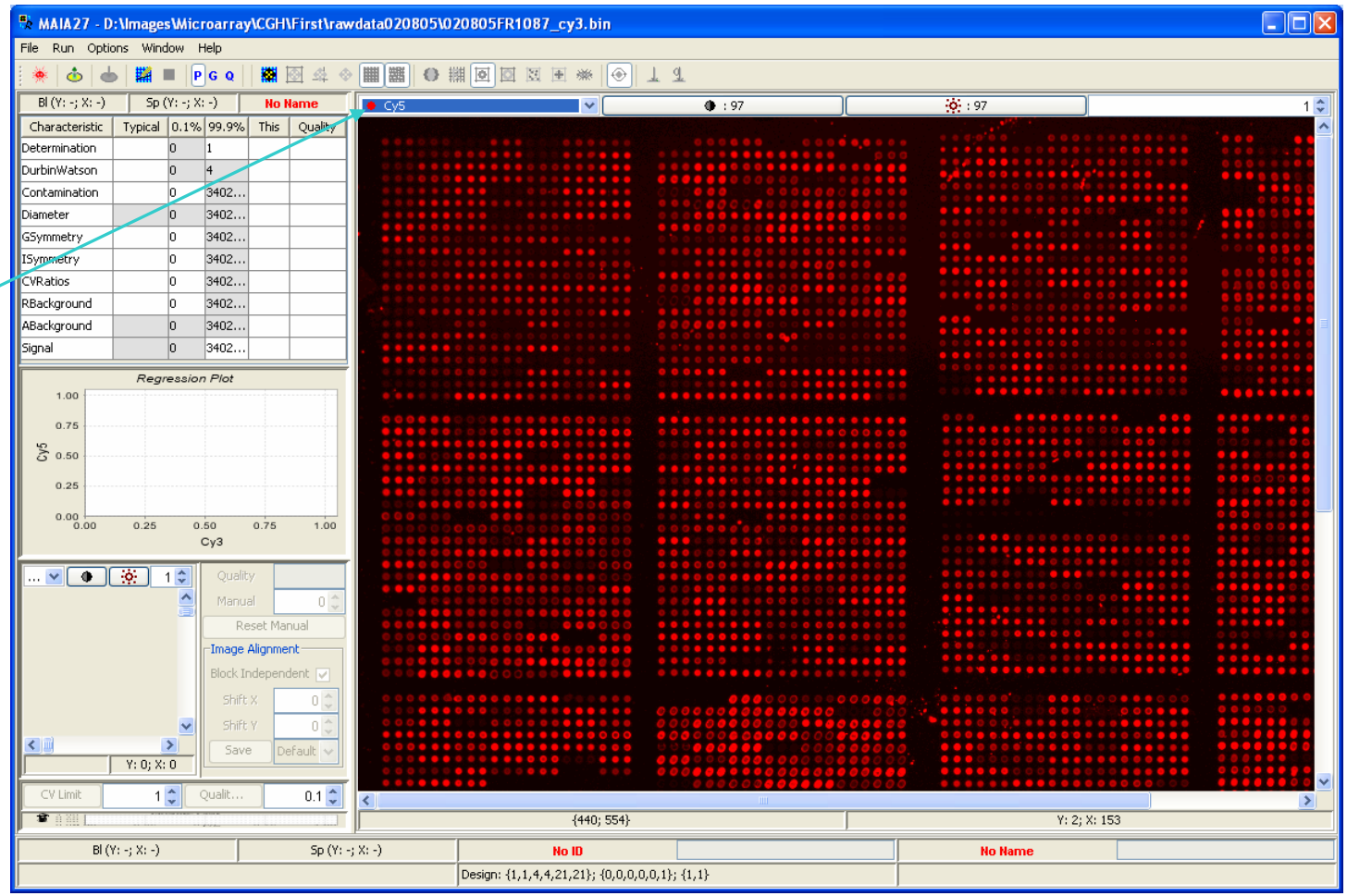
Green Channel

Select the green-dot (Cy3) to visualize the image colored in green.



Red Channel

Select the red-dot (Cy5) to visualize the image colored in red.



Color Swap

By default, green color is used for the Cy3 image and red color – for the Cy5 image. This assignment can be inverted by the Menu Item “File|Swap Colors”.

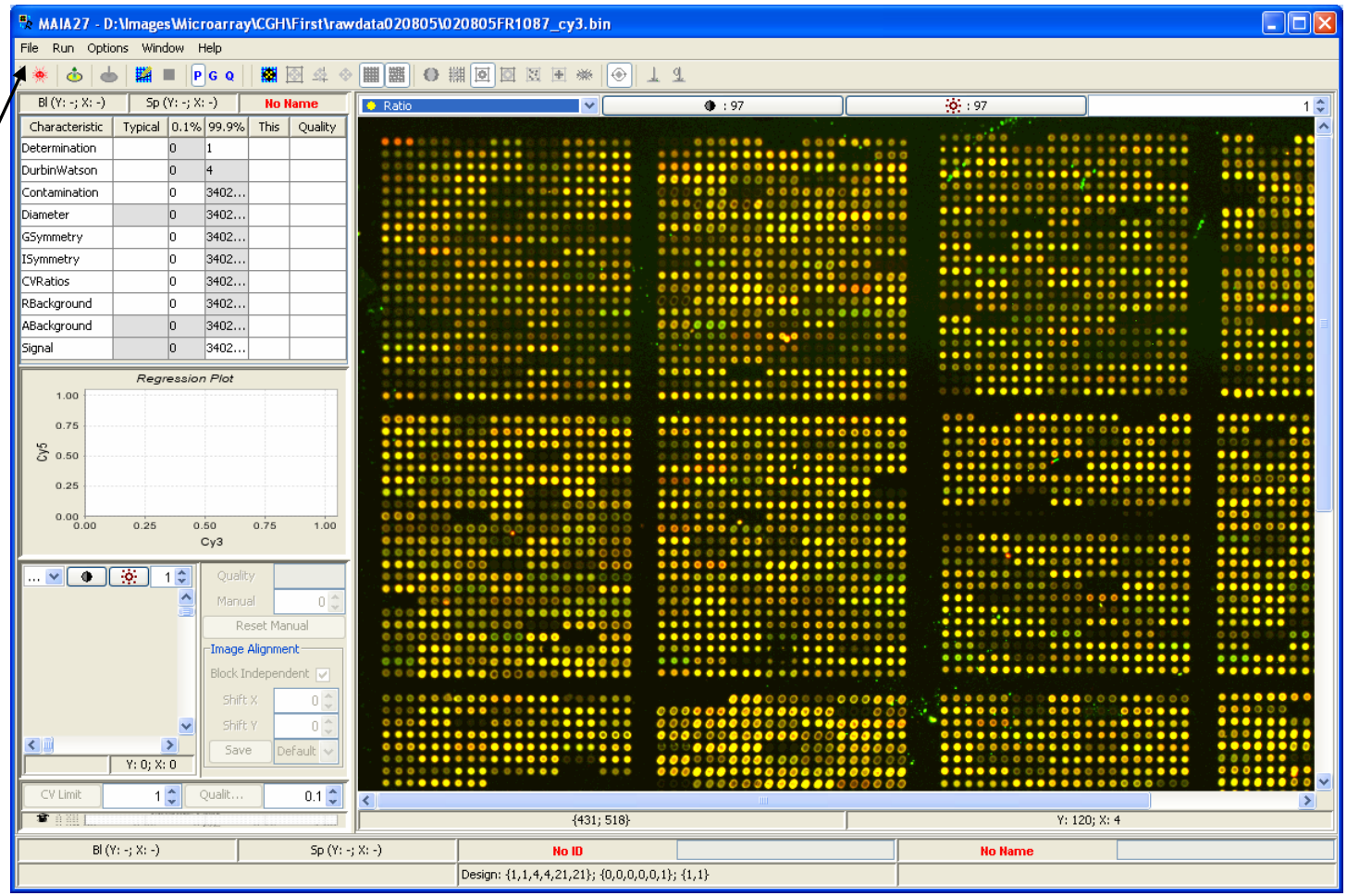
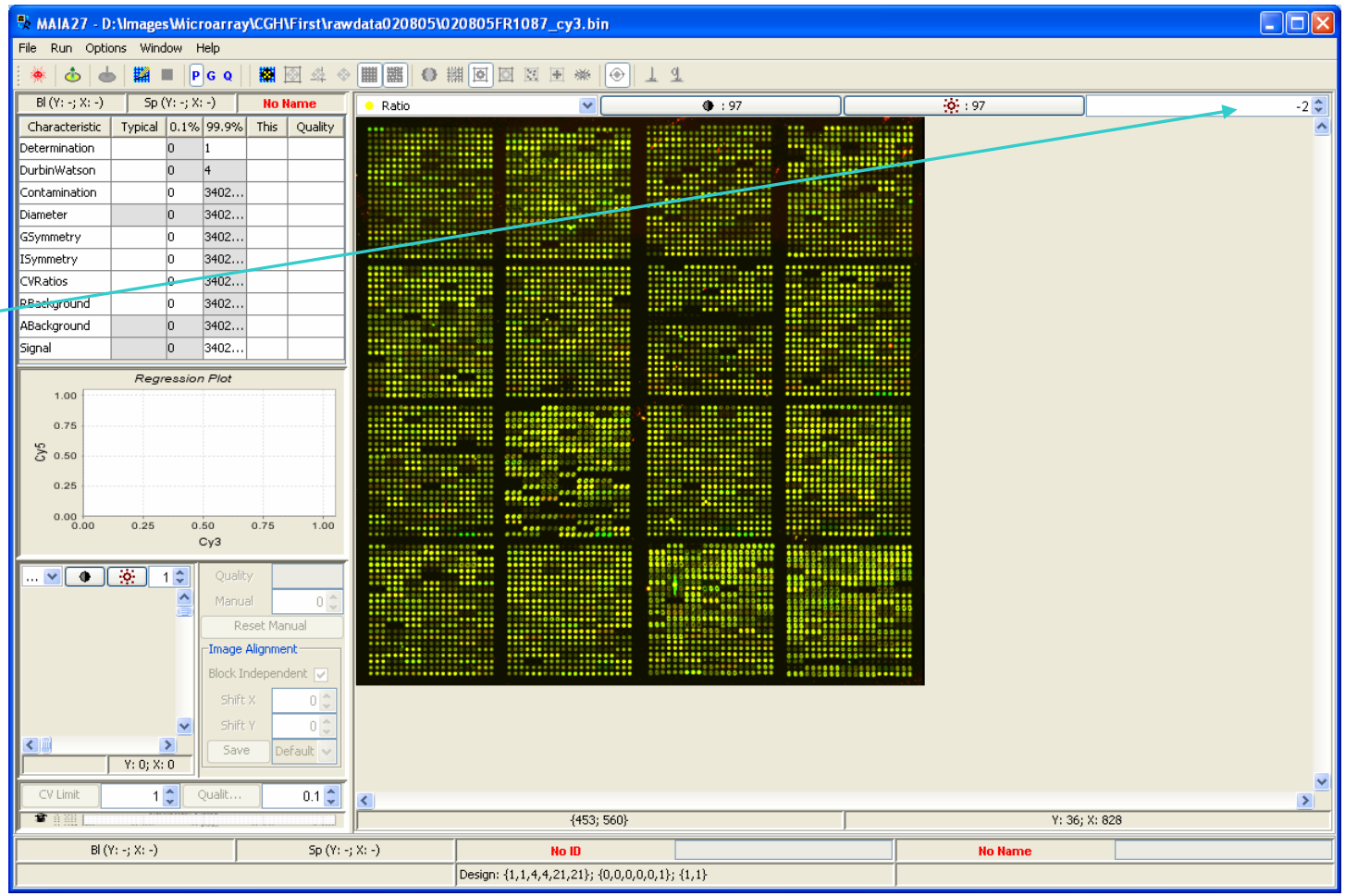


Image Zoom

Image can be zoomed using either the “Zoom” spinner box or the mouse wheel.

Negative values of the zoom indicate contraction; positive values indicate stretching. Original image is obtained with either 1 or -1 zoom. (Zoom does not influence the analysis.)



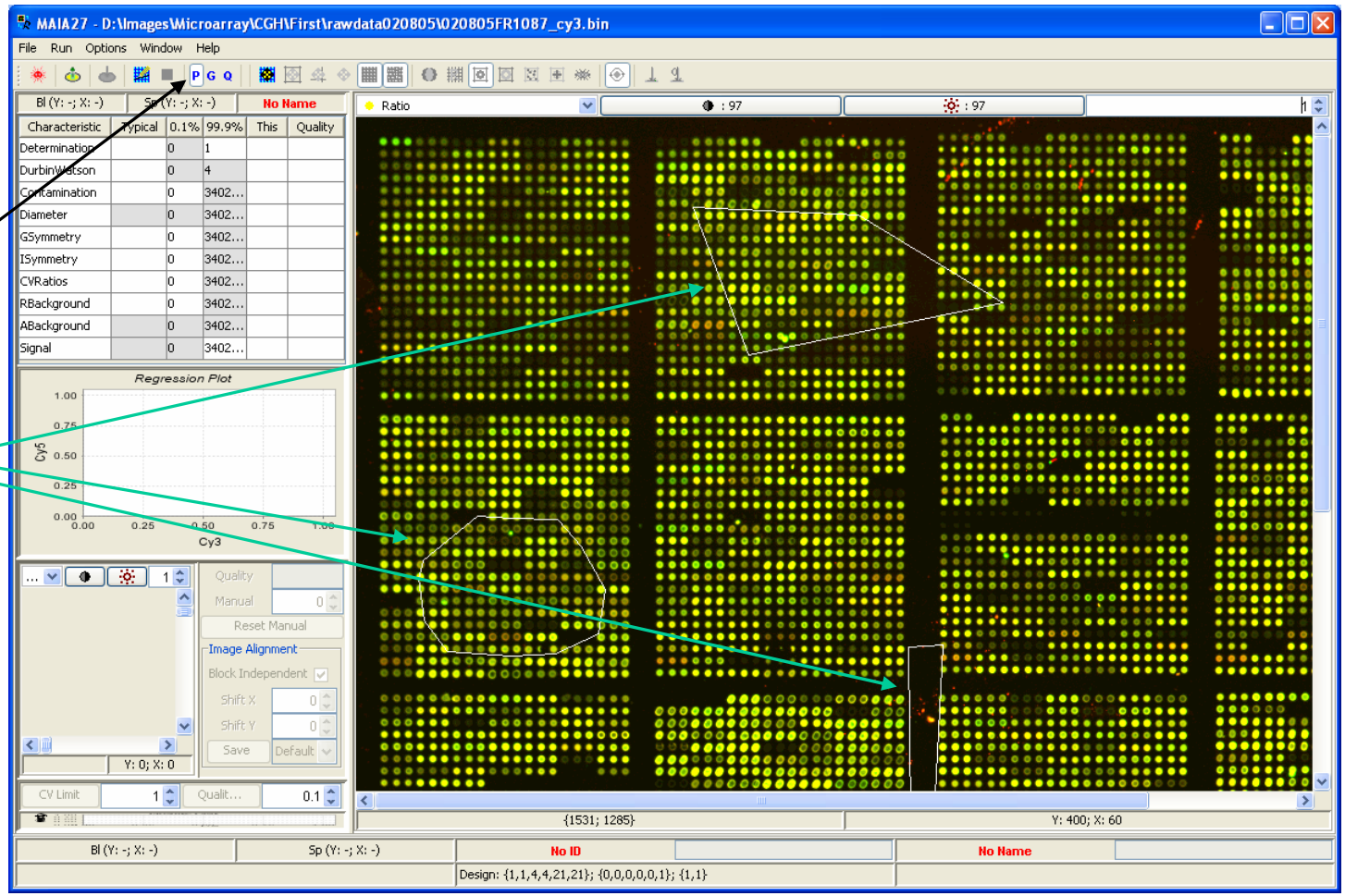
Manual Pixel Flagging (I)

Groups of (bad) pixels can be flagged out using the “Lasso selection” tool.

The “Manual Pixel Flagging” toggle button should be selected.

Ctrl+Left Clicks create the contour. Ctrl+Right Click closes the contour.

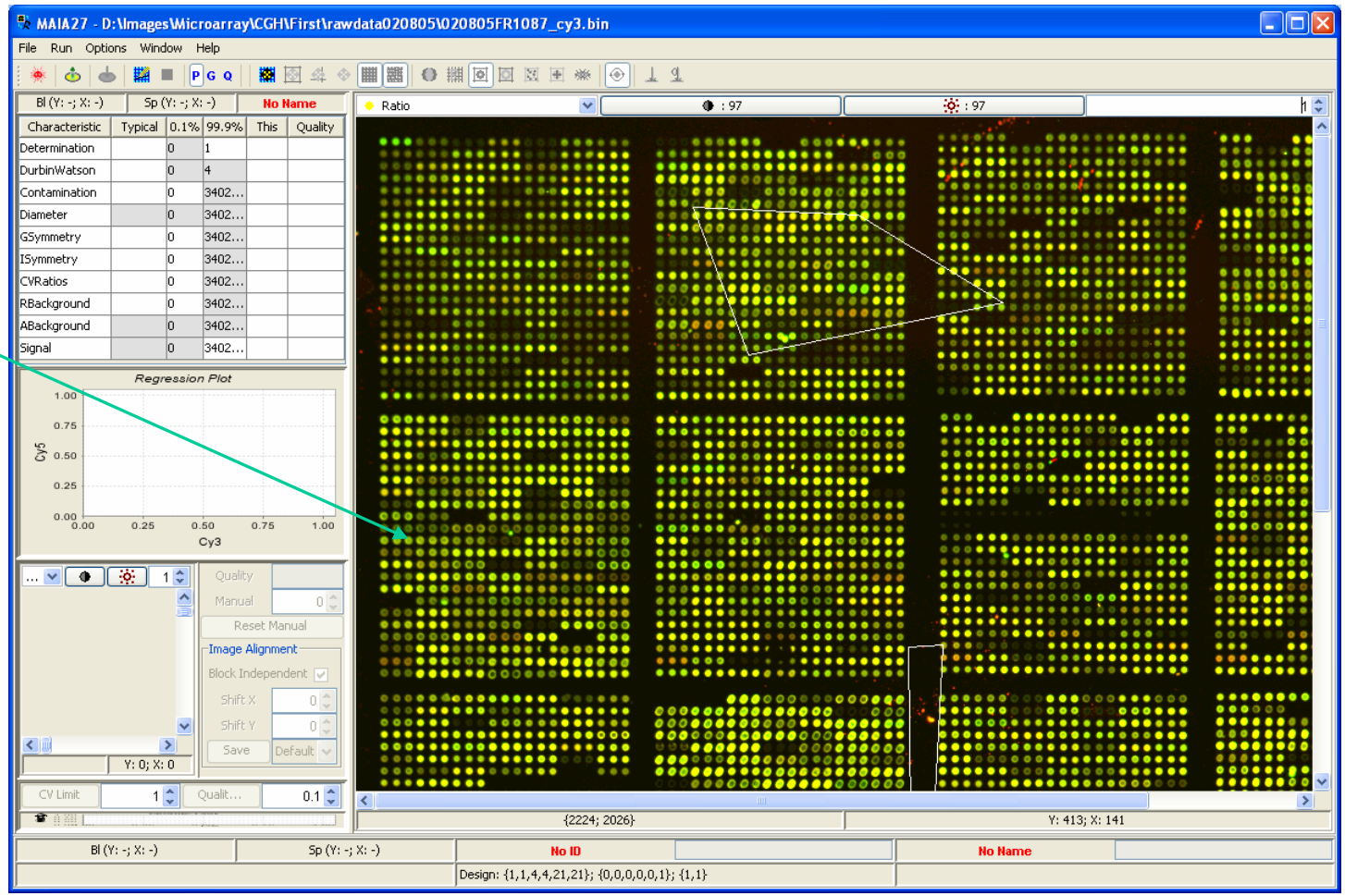
Flagged pixels are converted into the background pixels in spot localization and into the saturated pixels in spot quantification.



Manual Pixel Flagging (II)

Ctrl+Left Click within a contour effaces this contour.

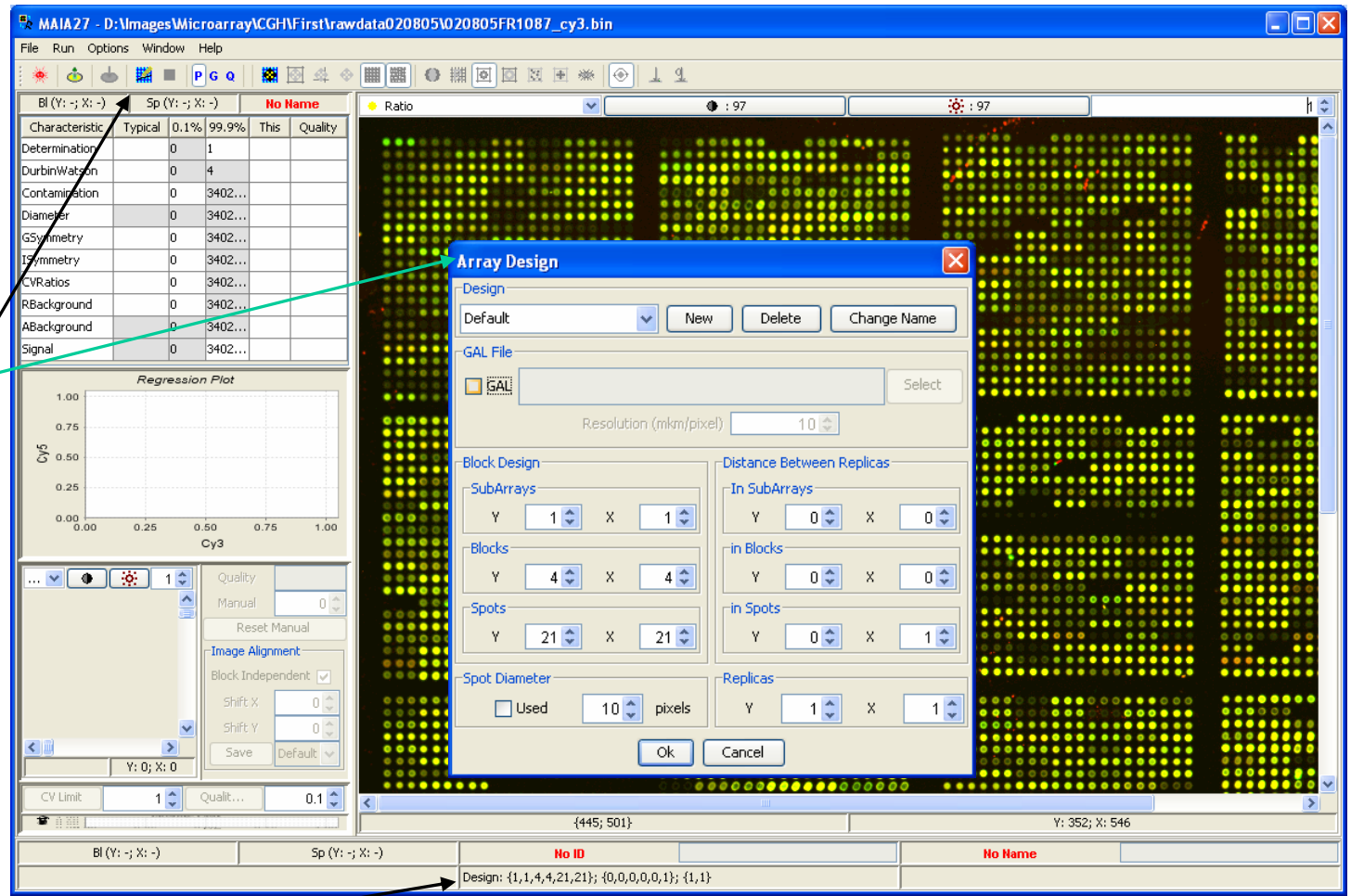
Double click on the image effaces all contours.



Array Design

To start image processing, array design should be properly defined: use the “Array Design” button from the Toolbar or select the Menu Item “Options|Array Design” (Alt+A).

See next page for details.



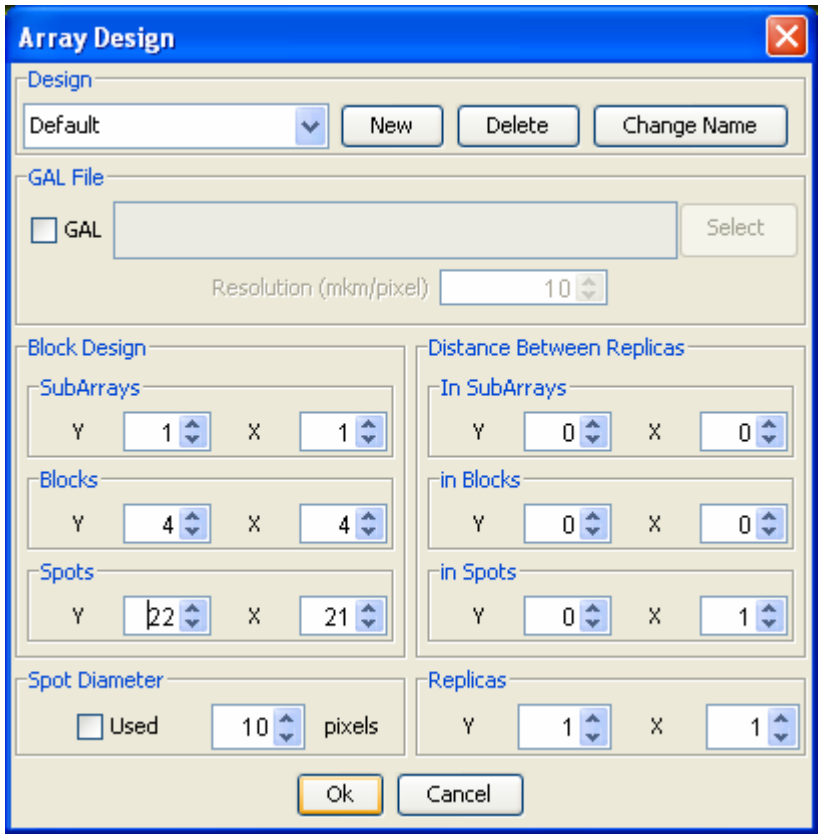
Currently used Array Design

Array Design in Detail

One may use several microarray designs under different names to be able to switch quickly from one design to another.

Amount of sub-arrays, blocks (per sub-array) and spots (per block) in Y and X directions of the array.

Spot diameter may be used as a prior value in spot localization and spot quantification procedures.



Array Design can be completely specified using GAL files (Axon Instruments, Inc. http://www.axon.com). Correct image resolution (mkm/pixel) should also be provided.

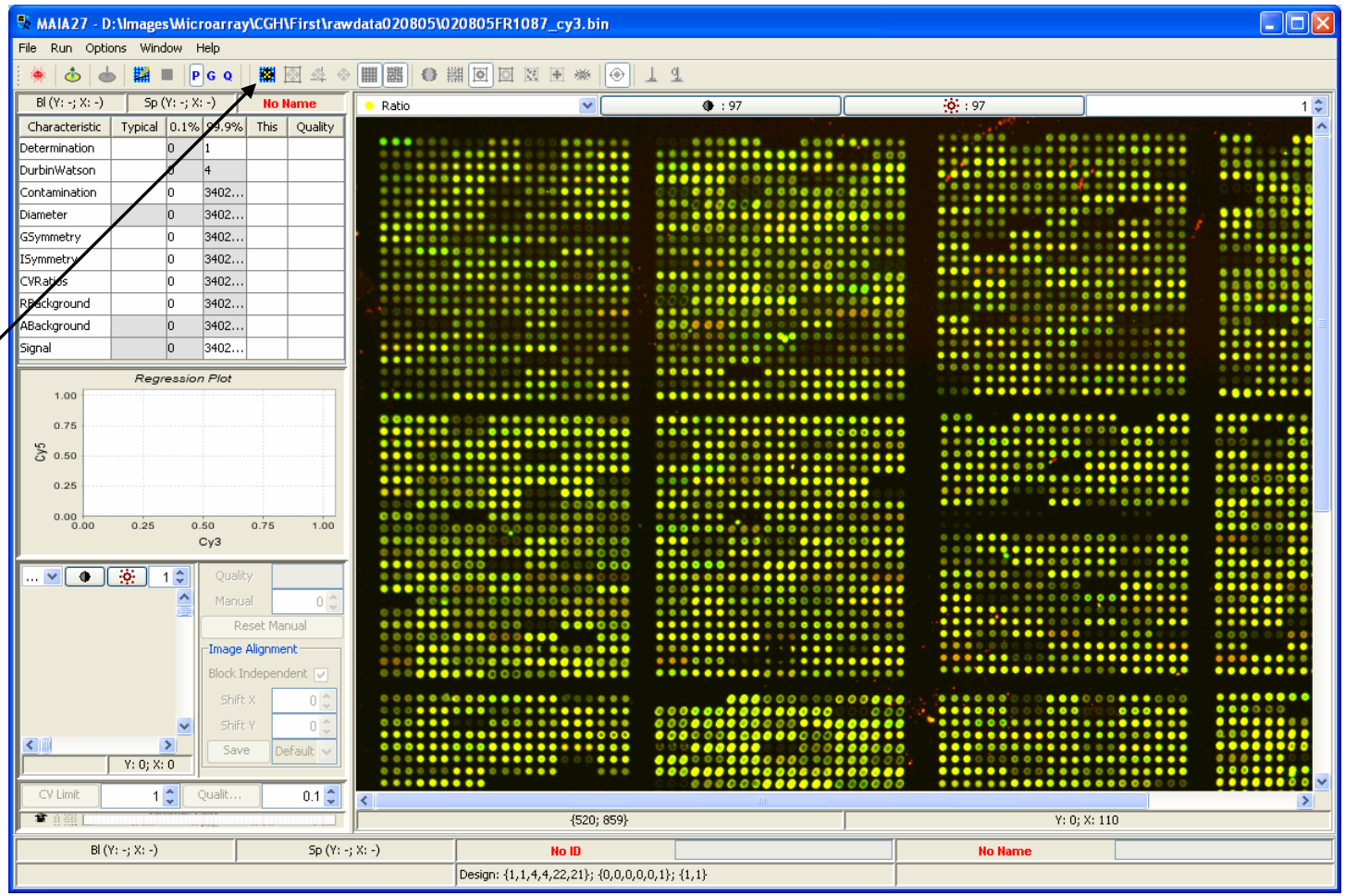
Relative coordinates of the replicated spots: it defines the position of the replicated spot with respect to the current one.

Amount of the replicated spots in the Y and X directions.

Spot Localization

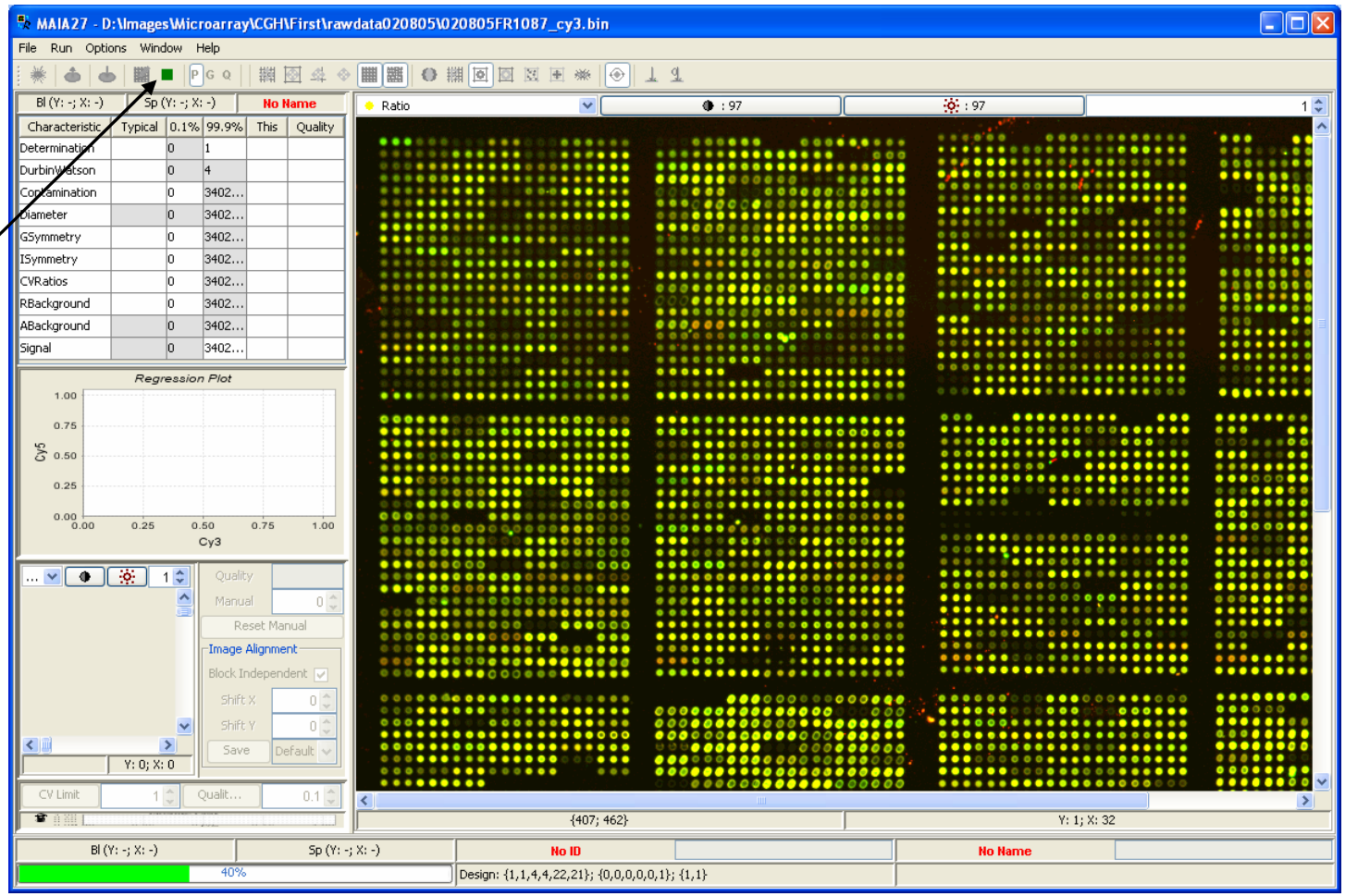
To start spot localization (or grid finding) use the "Spot Localization" button from the Toolbar or select the Menu Item "Run|Spot Localization" (Ctrl+F6).

For automatic grid generation it is advisable to ensure relatively broad external margins – distances from the edges of the array to the spotting area.



Terminate Processing

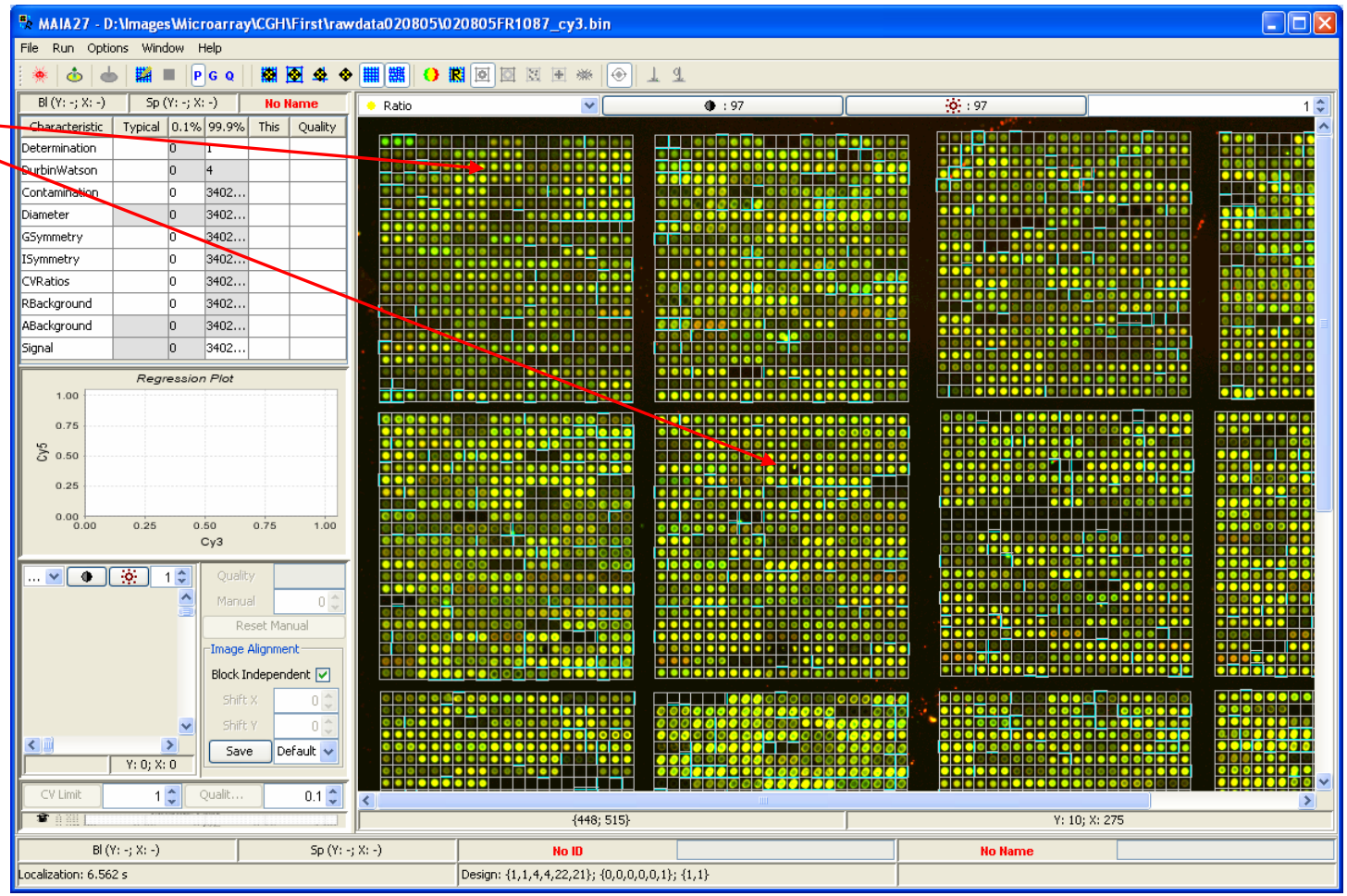
Any processing can be stopped by pressing the "Stop" button on the Toolbar or selecting the Menu Item "Run|Stop" (Ctrl+F5).



Spot Localization Output

Typical result of the spot localization: two grids are imposed over the image:

•*Main Grid* is composed of the straight lines separating neighborhood spot rows or columns;
•*Adjusted Grid* is composed of the piecewise lines – refined borders between the neighborhood spots.

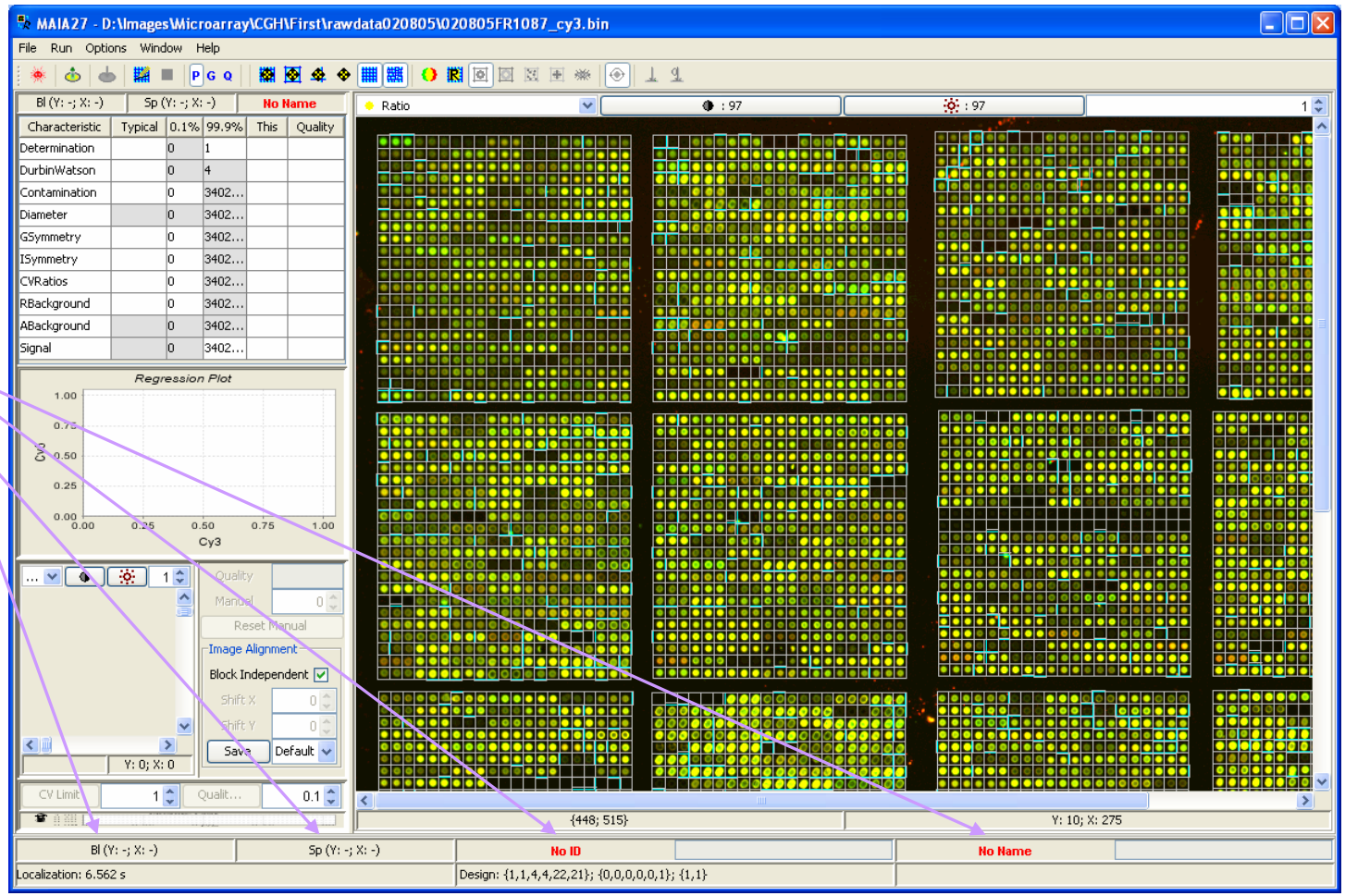


E. Novikov and E. Barillot, A noise-resistant algorithm for grid finding in microarray image analysis. *Machine Vision and Applications*, 2006, 17, 337-345.

Spot Localization Output: Spot Identification

"Under-mouse" coordinates of the block (Bl), spot (Sp), clone ID and clone Name.

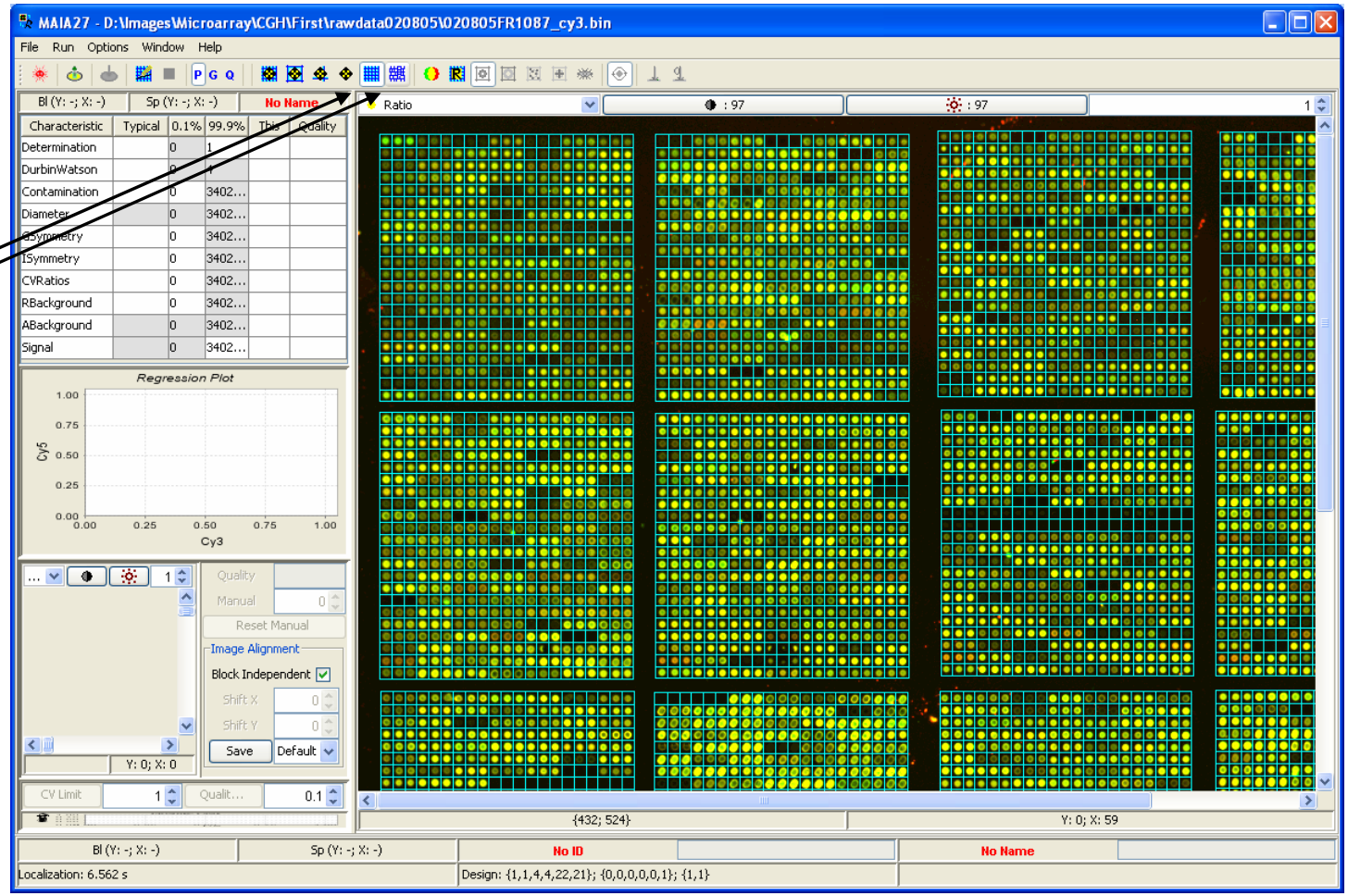
Clone IDs and clone names are available from GAL files.



Spot Localization Output: Main Grid

Using the Toolbar buttons "Show/Hide Main grid" or "Show/Hide Adjusted grid" one can mask either of two spot localization grids.

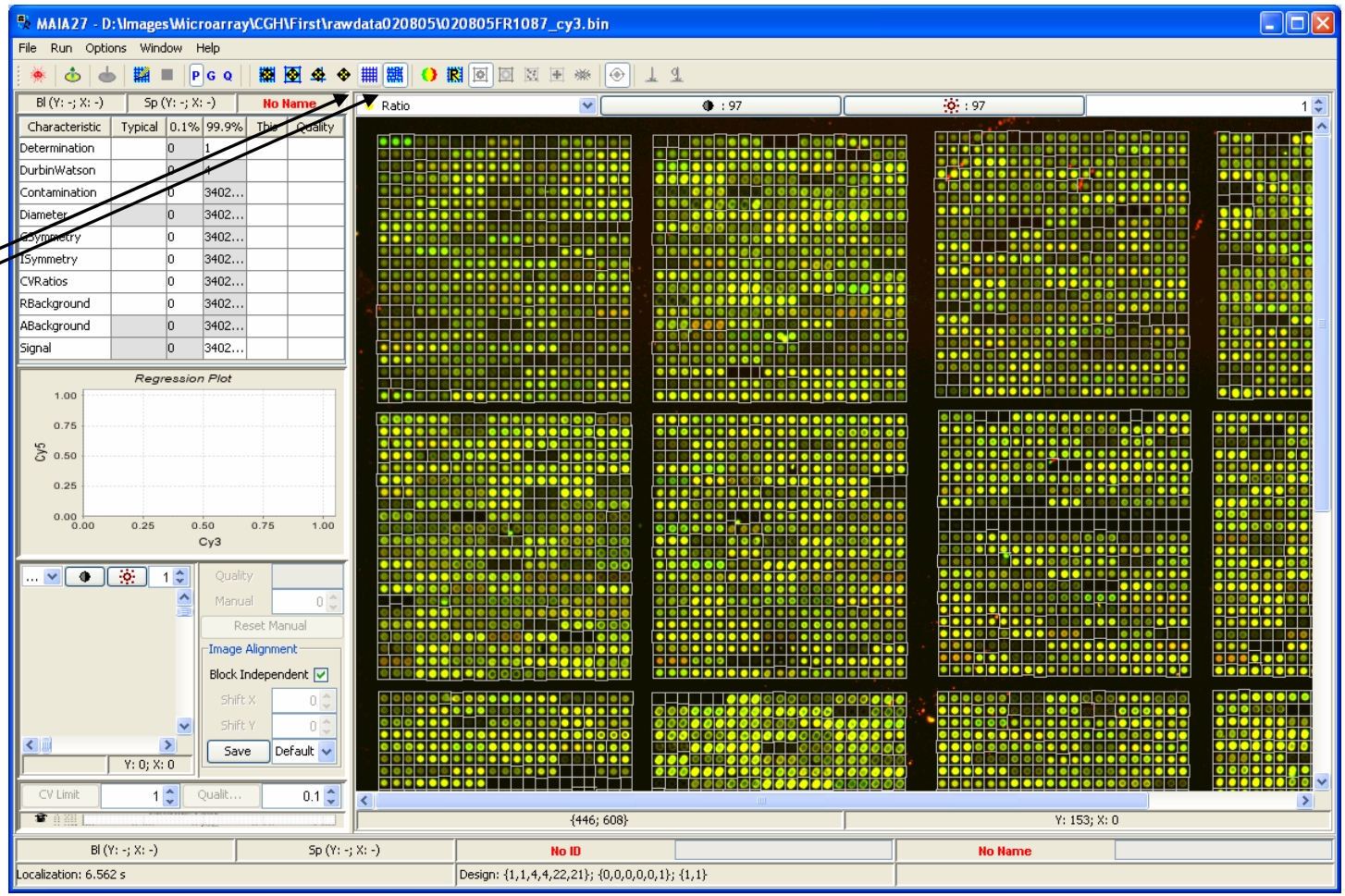
Main Grid is shown.



Spot Localization Output: Adjusted Grid

Using the Toolbar buttons "Show/Hide Main grid" or "Show/Hide Adjusted grid" one can mask either of two spot localization grids.

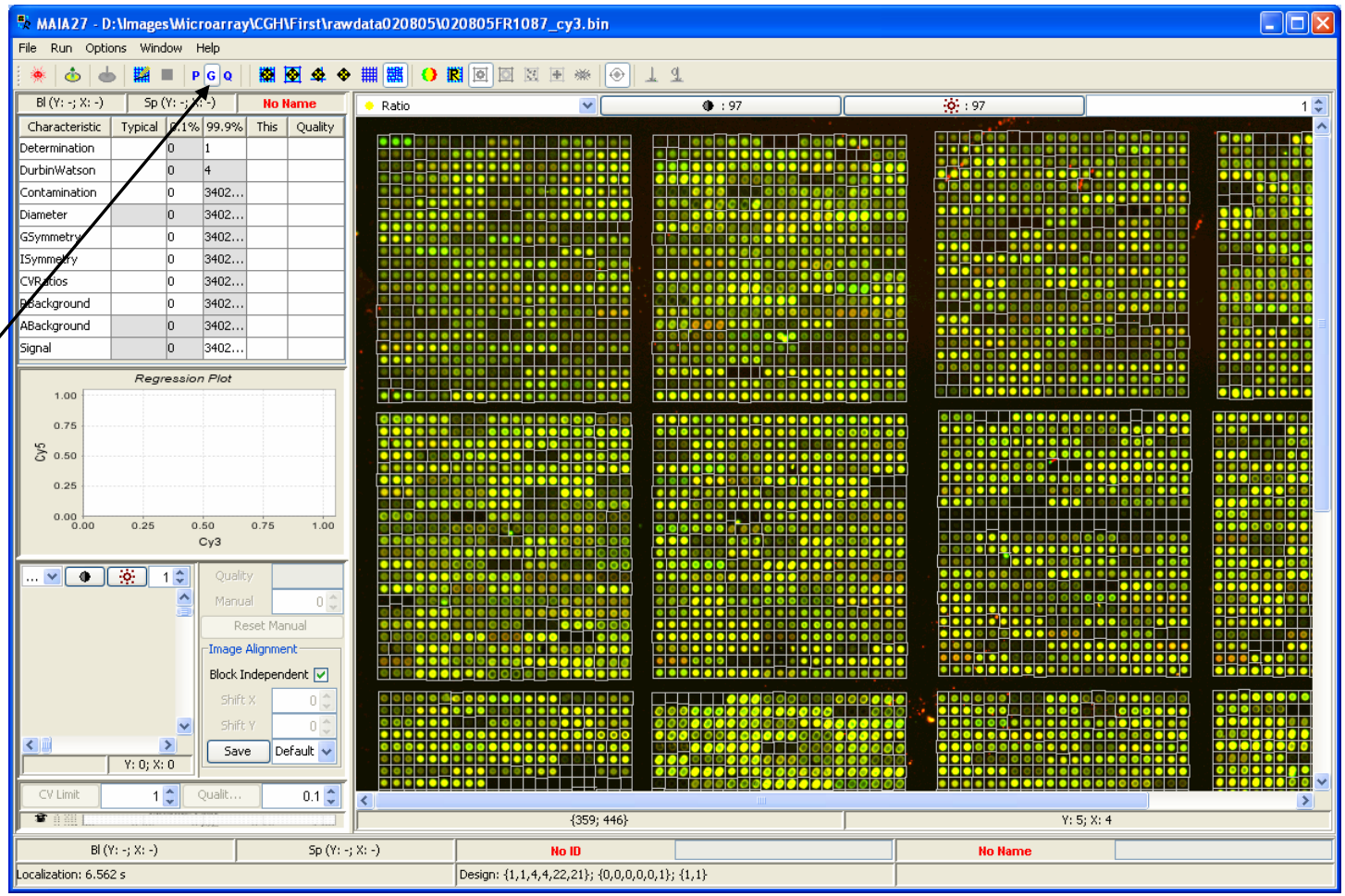
Adjusted Grid is shown.



Manual Correction of the Grids

If the generated grids are corrupted, manual correction can be applied: select the toggle button "Manual Grid Correction".

All manual corrections of the grids can be "undone". Ctrl-Z implements step-by-step "UnDo" and Ctrl-Shft-Z – step-by-step "ReDo".



Manual Correction of the Main Grid: Grid Movements

The selected grid can be shifted on the discrete number of spot rows/columns or moved smoothly over the image.

Select a grid and iterate through the grids:

Shift + Left Click
Shift + Home
Shift + End
Shift + PgUp
Shift + PgDn

Move Selection by Pixel:

Shift + Drag
Shift + {↑,↓,→,←}

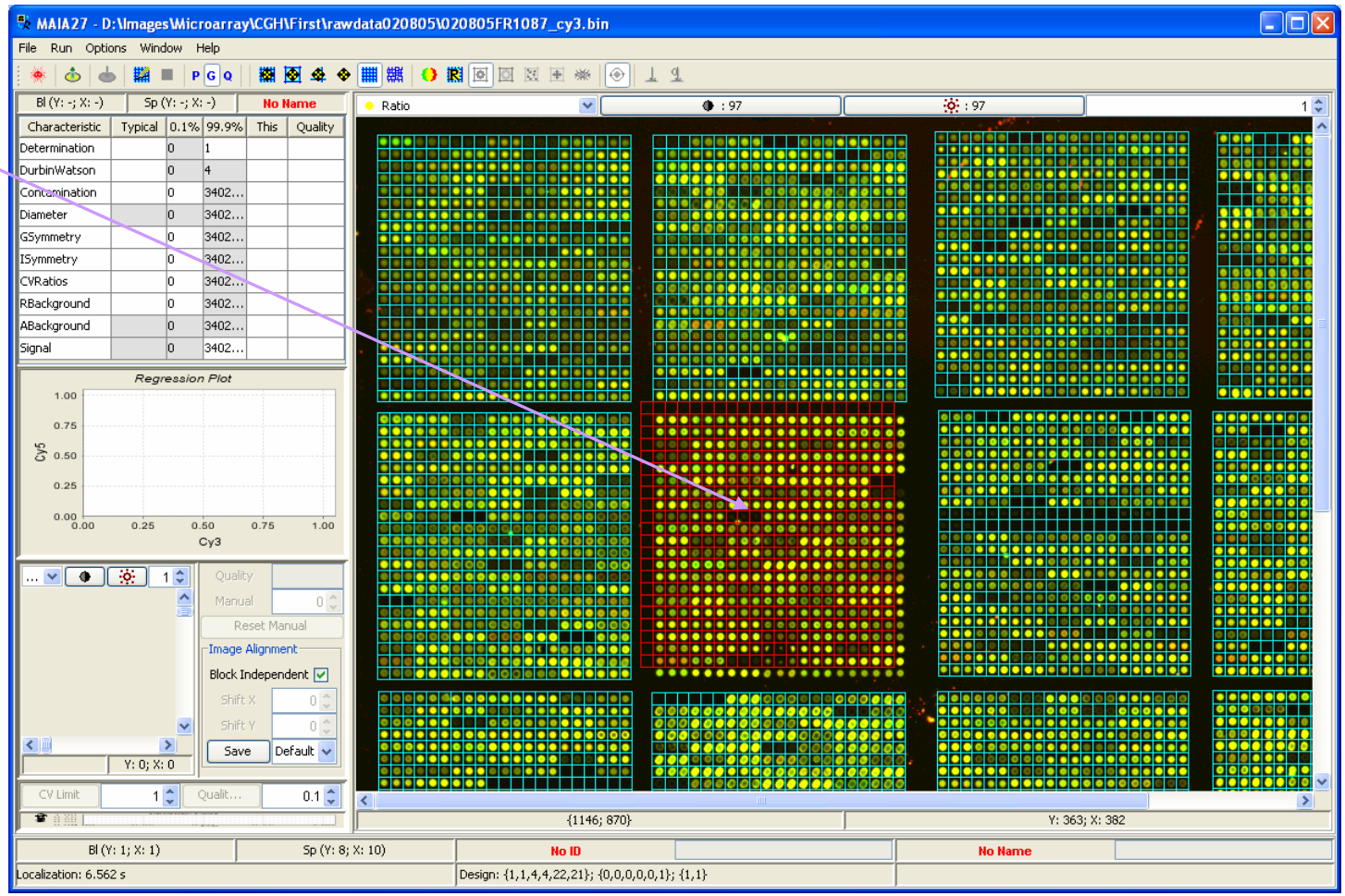
Move Selection by Spot:

Shift + Ctrl + {↑,↓,→,←}

Clear Selection:

Mouse Click
Ctrl + Del

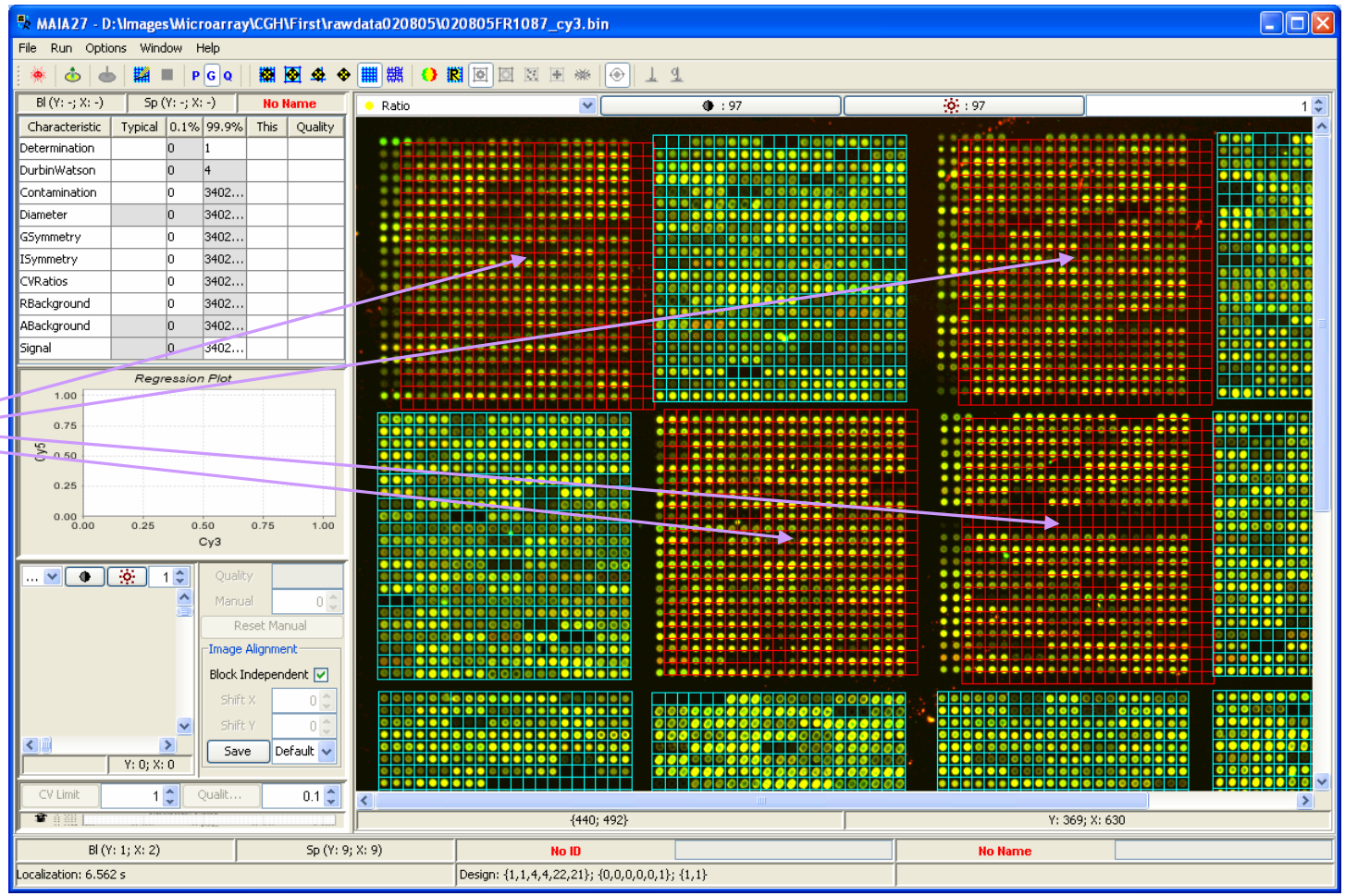
Upon selection the grid changes the color.



Manual Correction of the Main Grid: Multiple Grids Selection

Several grids can be selected using Shift+Left Click.

Shift+Double Left Click selects all grids on the image.



Manual Correction of the Main Grid : Line Movements

The line separations can be corrected in the main grid.

Select a line and iterate through the lines:

Ctrl + Left Click
Ctrl + Home
Ctrl + End
Ctrl + PgUp
Ctrl + PgDn

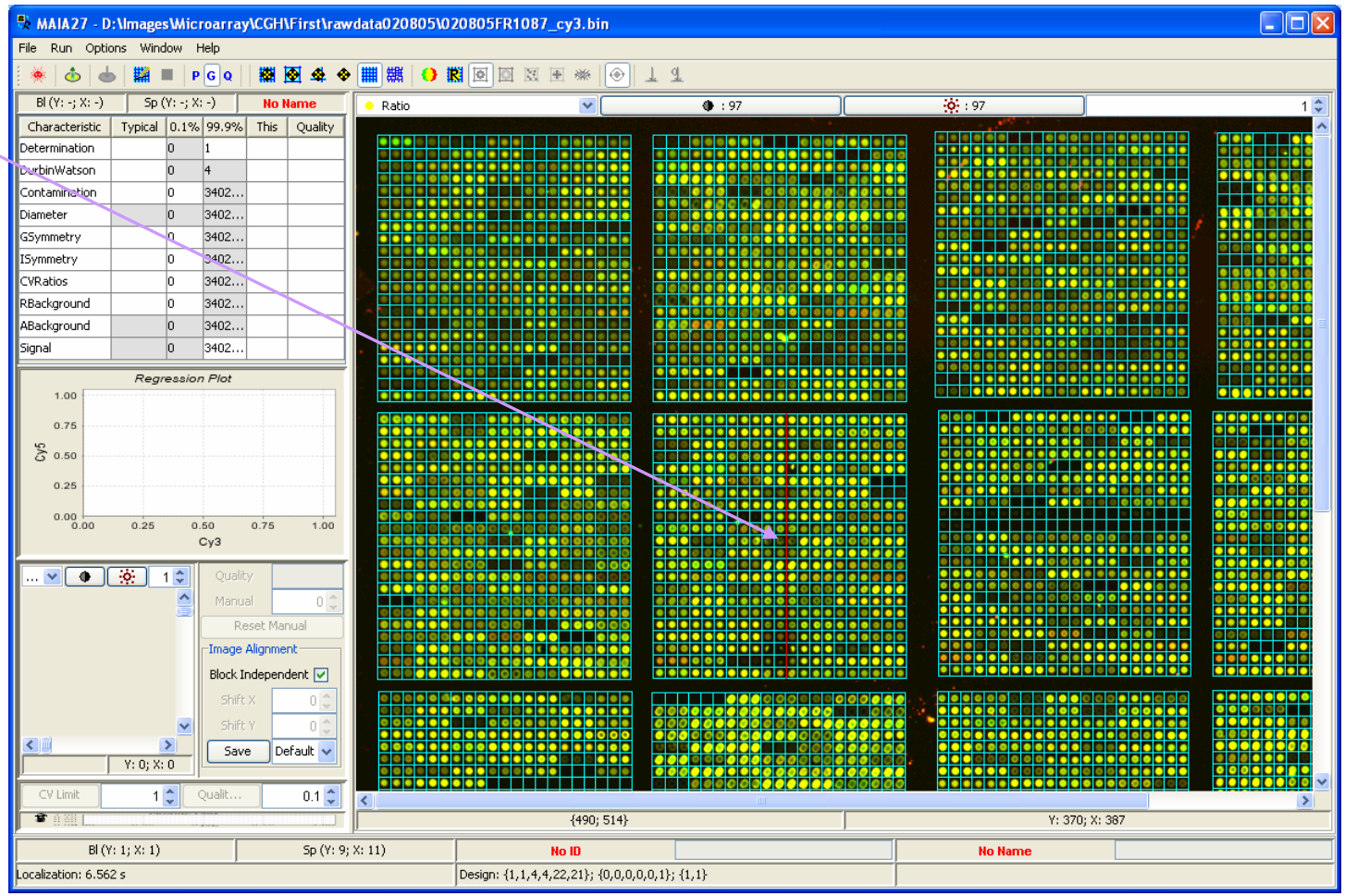
Move Selection by Pixel:

Ctrl + Drag
Ctrl + {↑,↓,→,←}

Clear Selection:

Mouse Click
Ctrl + Del

Upon selection the line changes the color.



Manual Correction of the Adjusted Grid

If a separation (cut) between the neighborhood spots is erroneous, one can perform manual correction of the selected cut position.

Select a cut and iterate through the cuts:

Alt + Left Click
Alt + Home
Alt + End
Alt + PgUp
Alt + PgDn

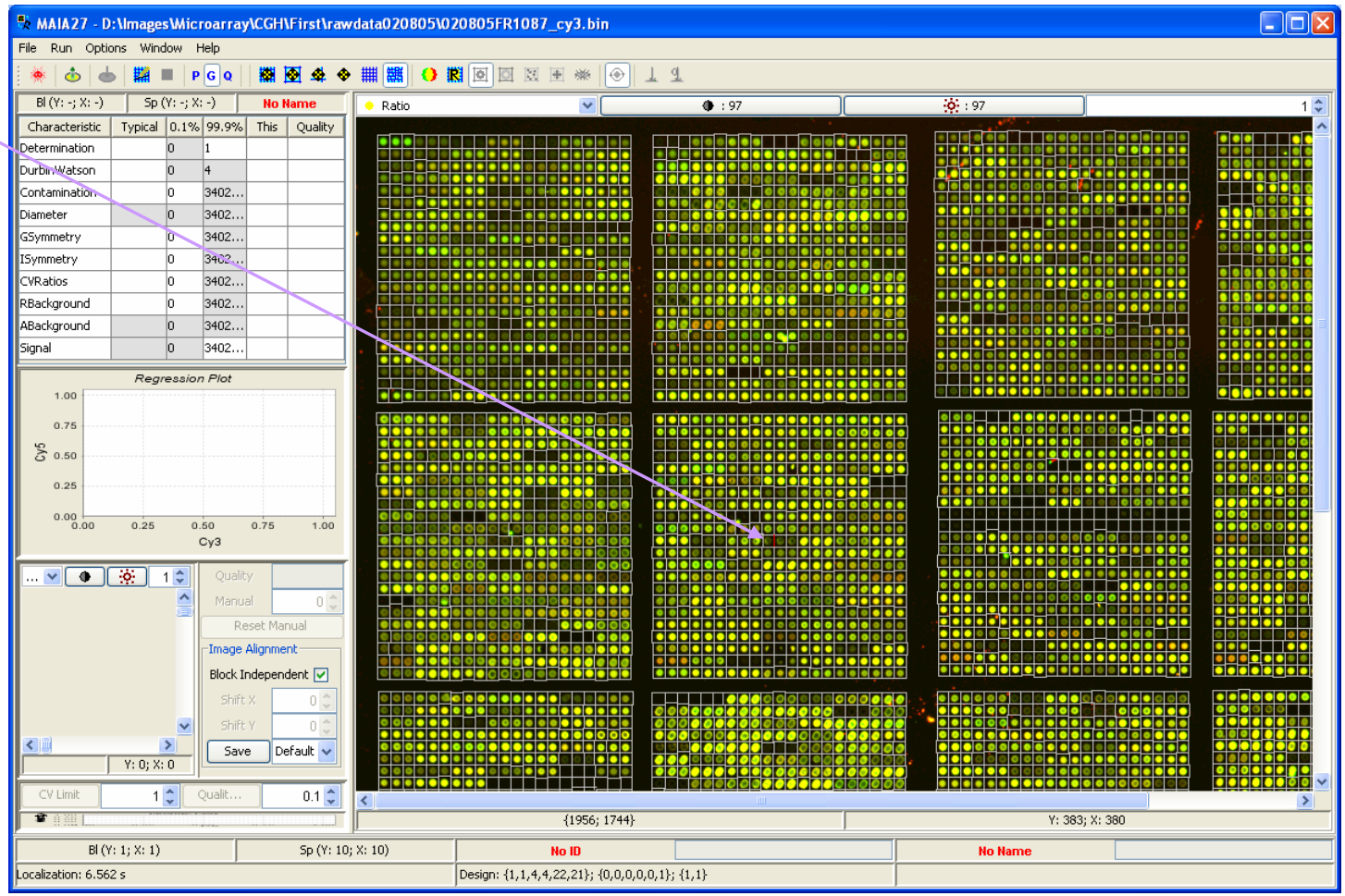
Move Selection by Pixel:

Alt + Drag
Alt + {↑,↓,→,←}

Clear Selection:

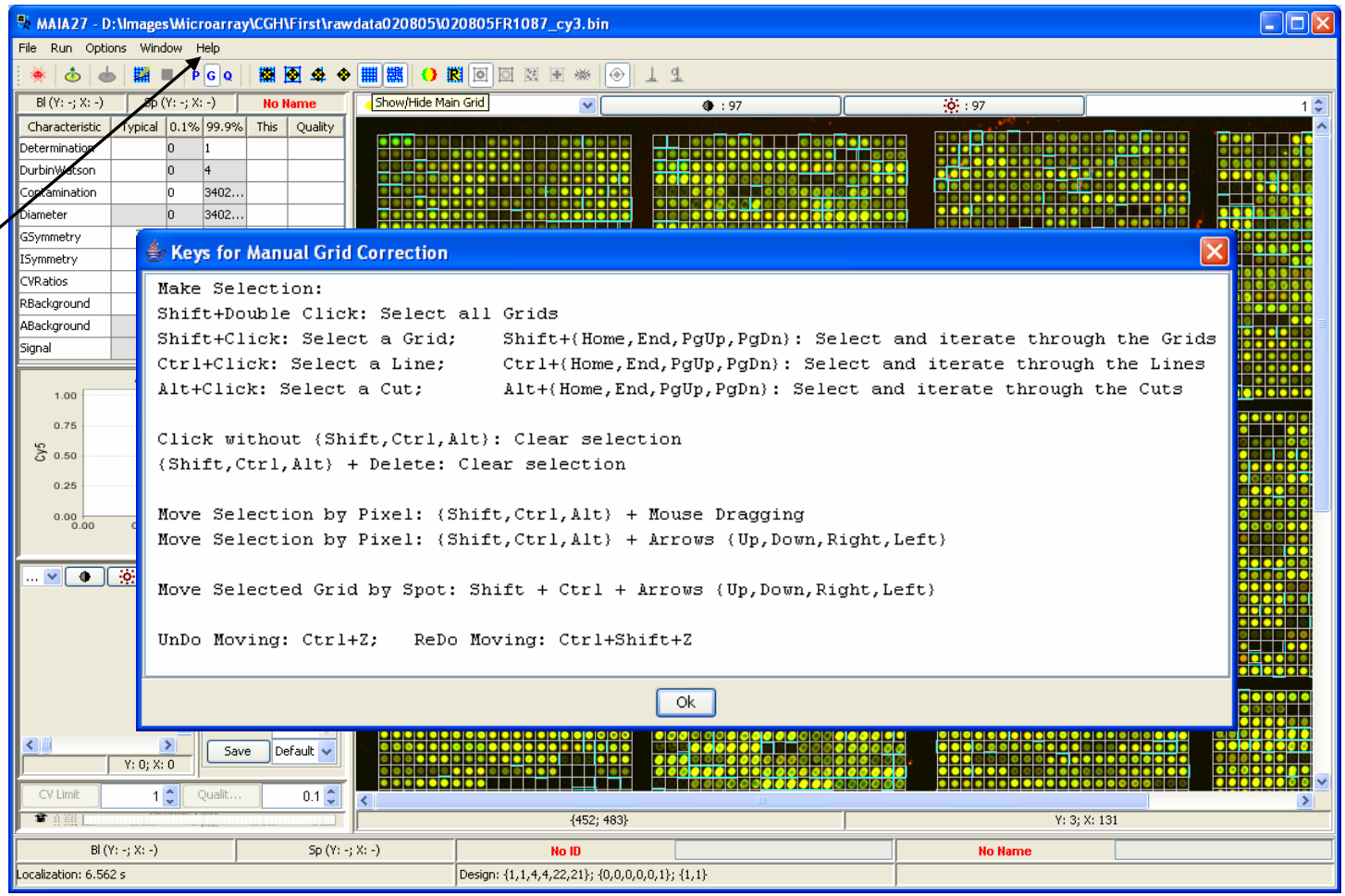
Mouse Click
Ctrl + Del

Upon selection the cut changes the color.



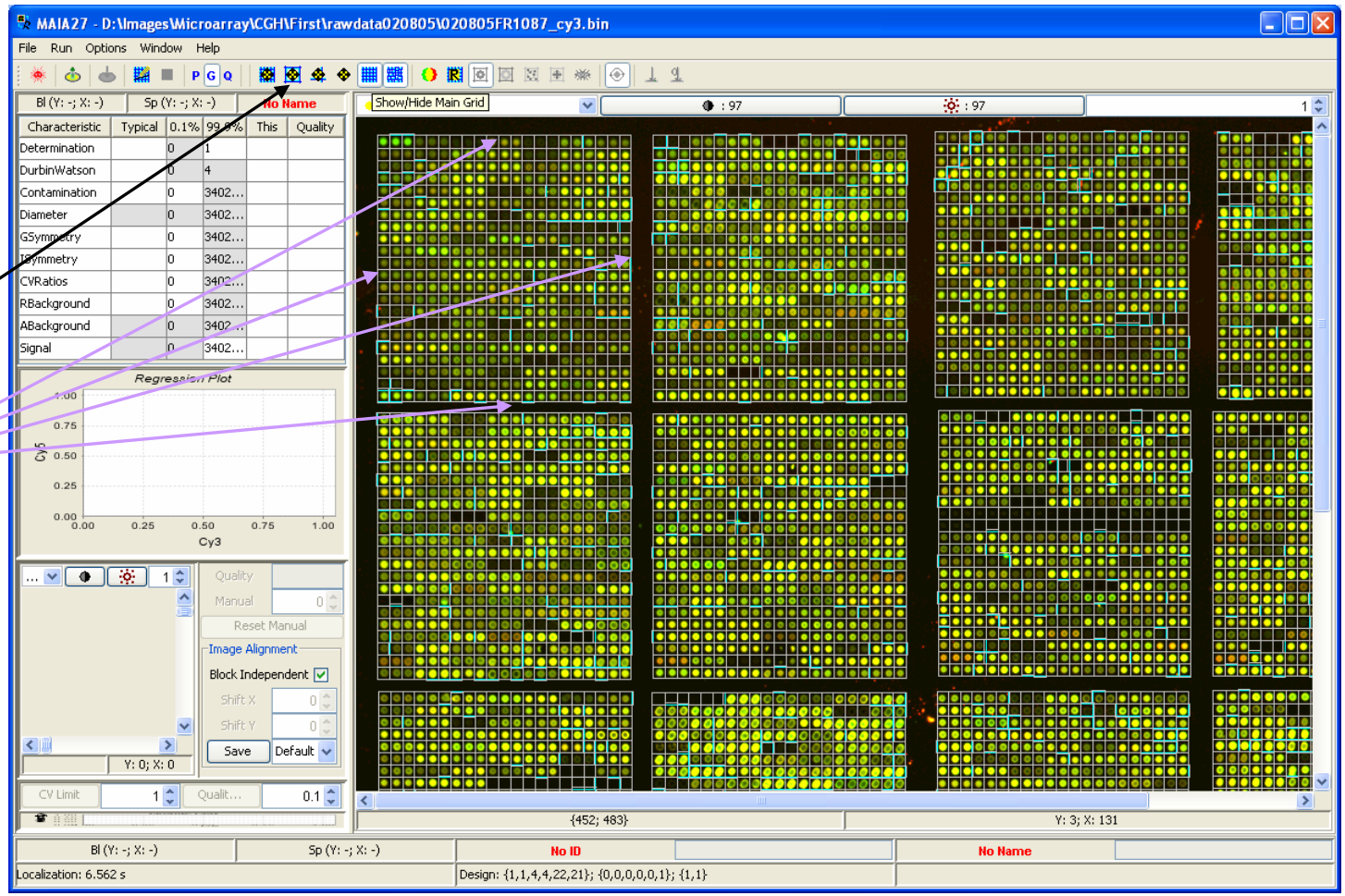
Brief Help on Manual Correction

Brief help on the manual correction possibilities is available at the Menu Item “Help|Manual Grid Info”.



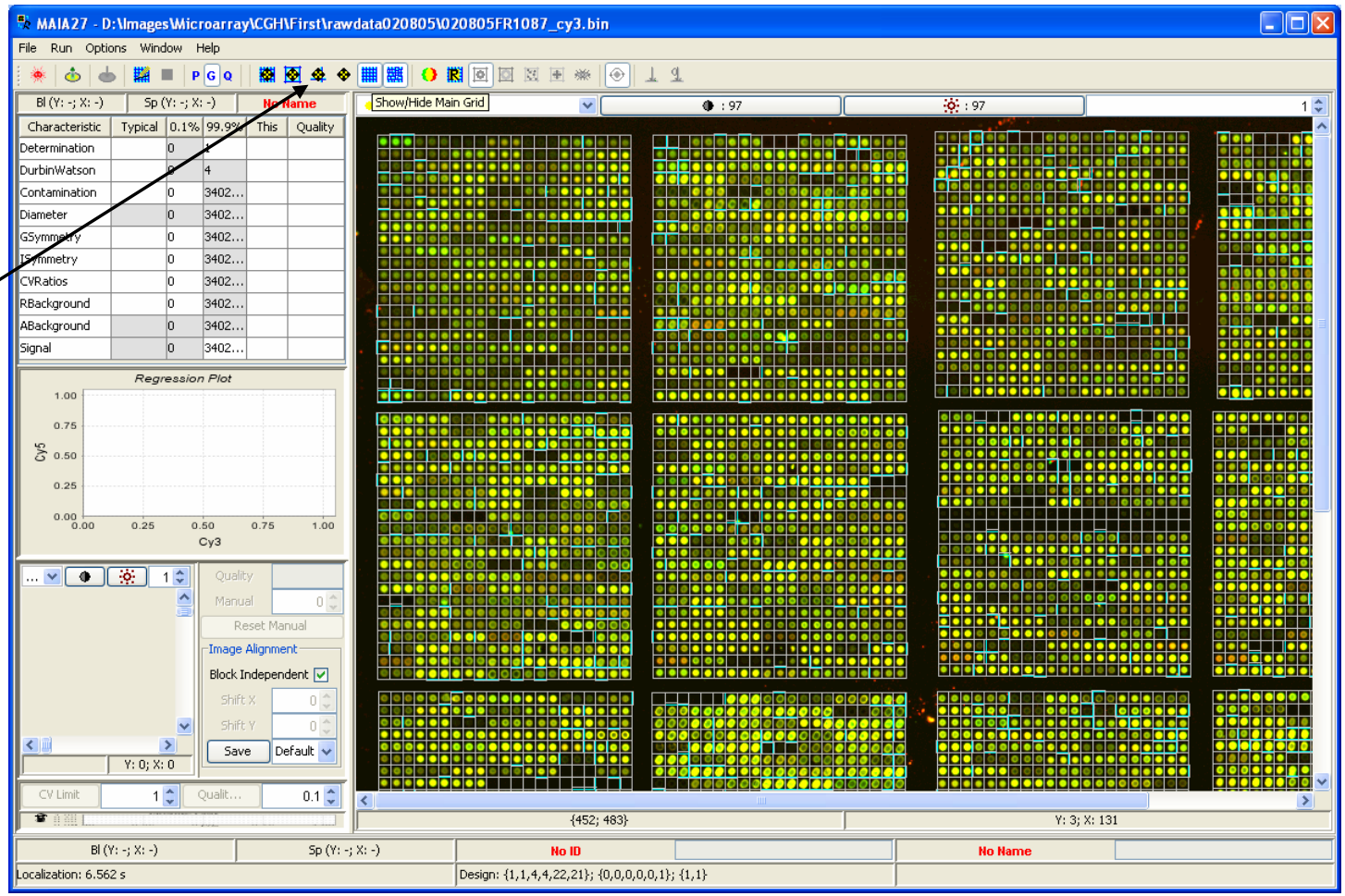
Main Grid Refinement: Find Grids in Blocks

Manual correction can be done only for the borders of the blocks. The other "internal lines" of the grids are found automatically using the "Grids in Blocks" button from the Toolbar or the Menu Item "Run|Grids in Blocks".



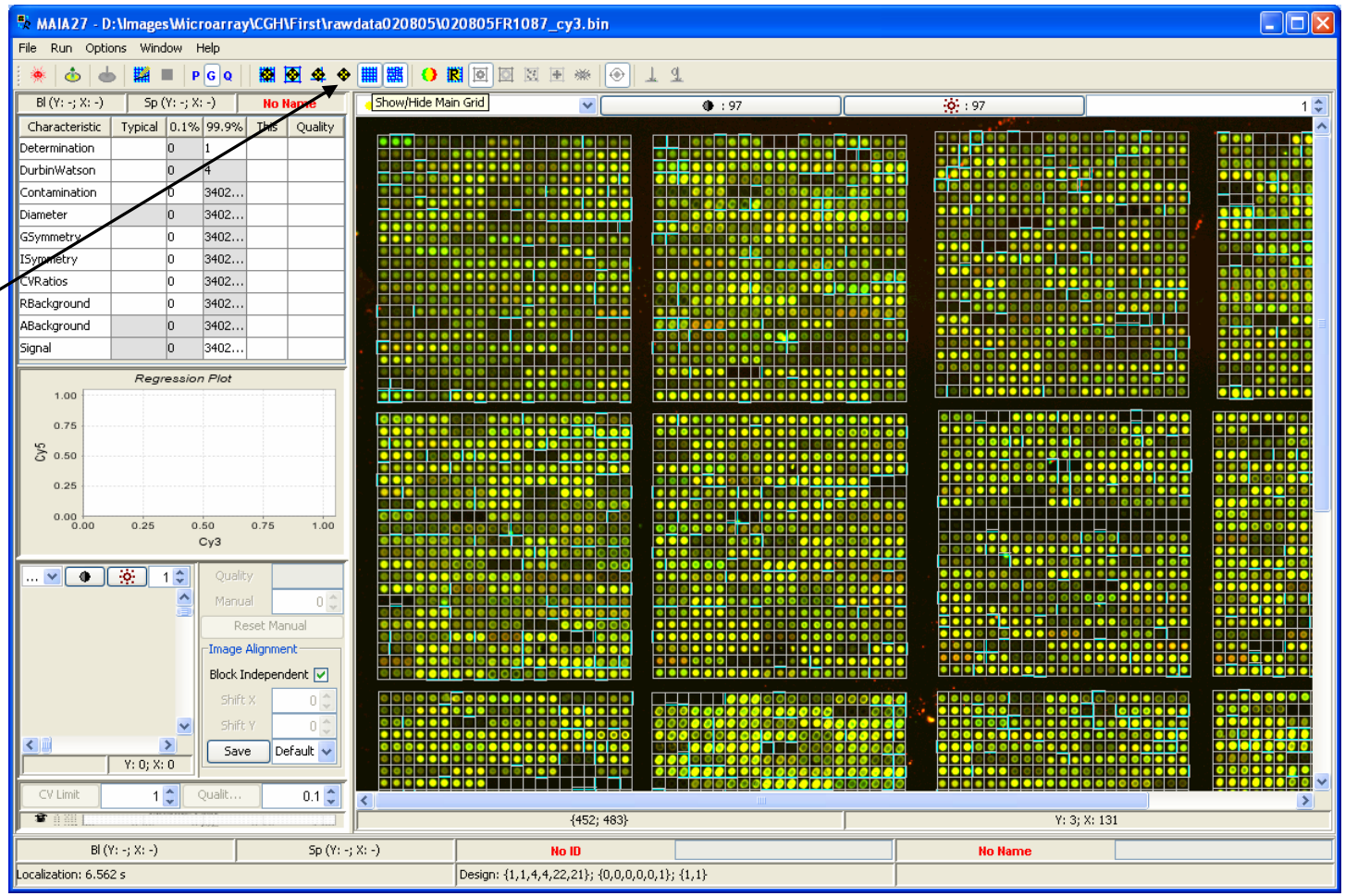
Main Grid Refinement: Lines Refinement

When the main grid is “almost” good, further refinement procedure will try to place the grid lines more precisely: use the “Lines Refinement” button from the Toolbar or the Menu Item “Run|Lines Refinement”.



Adjusted Grid Refinement: Cuts Refinement

When the adjusted grid is "almost" good, further refinement procedure will try to place the separators (cuts) between neighboring spots more precisely: use the "Cuts Refinement" button from the Toolbar or the Menu Item "Run|Cuts Refinement".

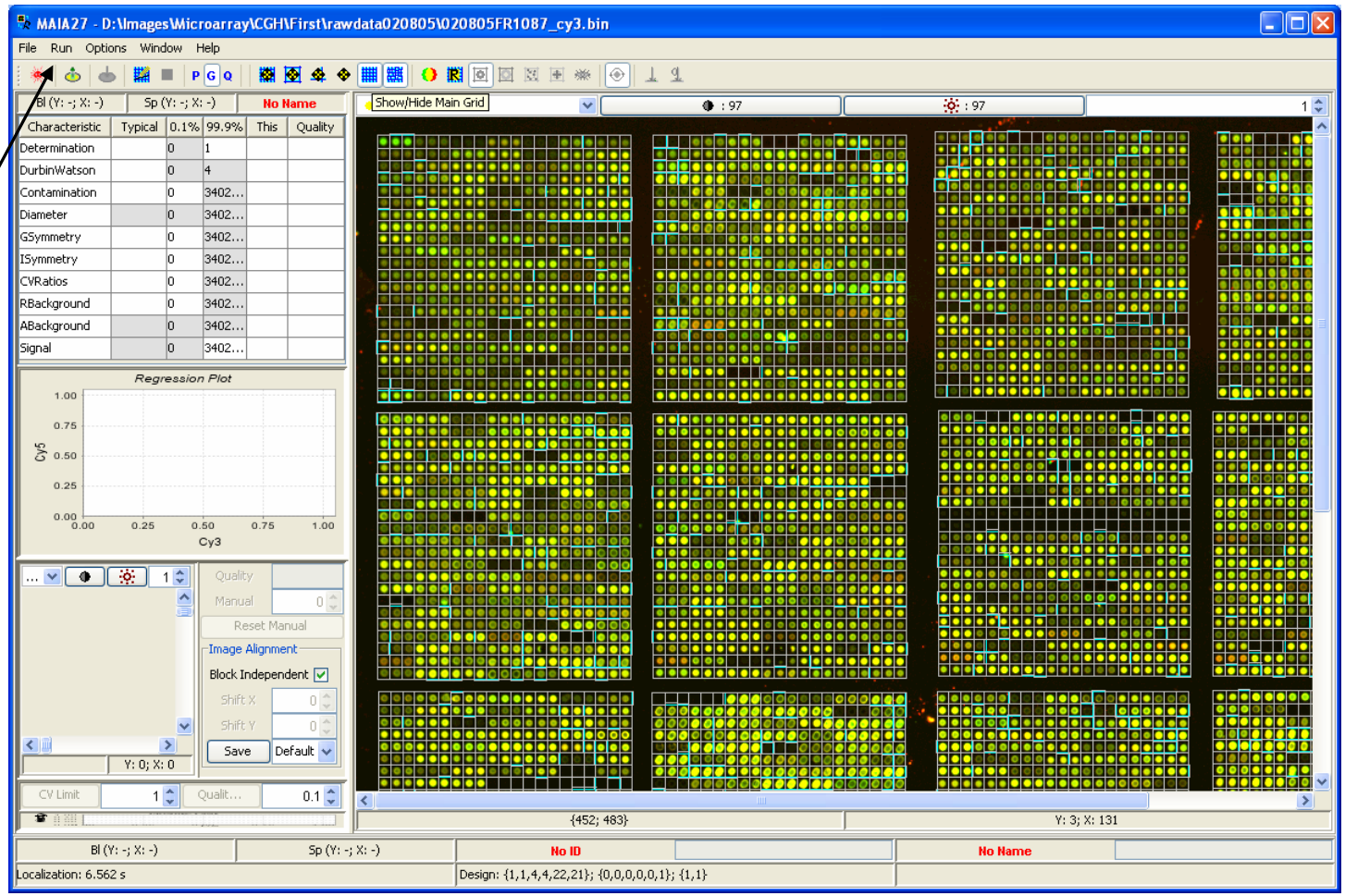


Manual Correction Hints

- If spot localization is satisfactory, there is no need to perform “grids in block”, “lines refinement” and “cuts refinement”. This is already done by the spot localization procedure.
- However, if grids were misplaced and manual correction has been performed, then either of “grids in block” or “lines refinement” or “cuts refinement” may be necessary. It depends on the manual correction.
- If the main grid is misplaced, only the external lines of the grid (i.e. the first and last lines of the main grid) can be adjusted and the “grids in block” will put all the other internal grid lines in-between the external grid lines.
- If internal lines of the main grid are misplaced, then only these lines can be corrected and the “lines refinement” puts them in the refined positions.
- Finally, if cuts of the adjusted grid are wrong, then, after their manual correction, the “cuts refinement” can be performed.

GAL Grid Generation

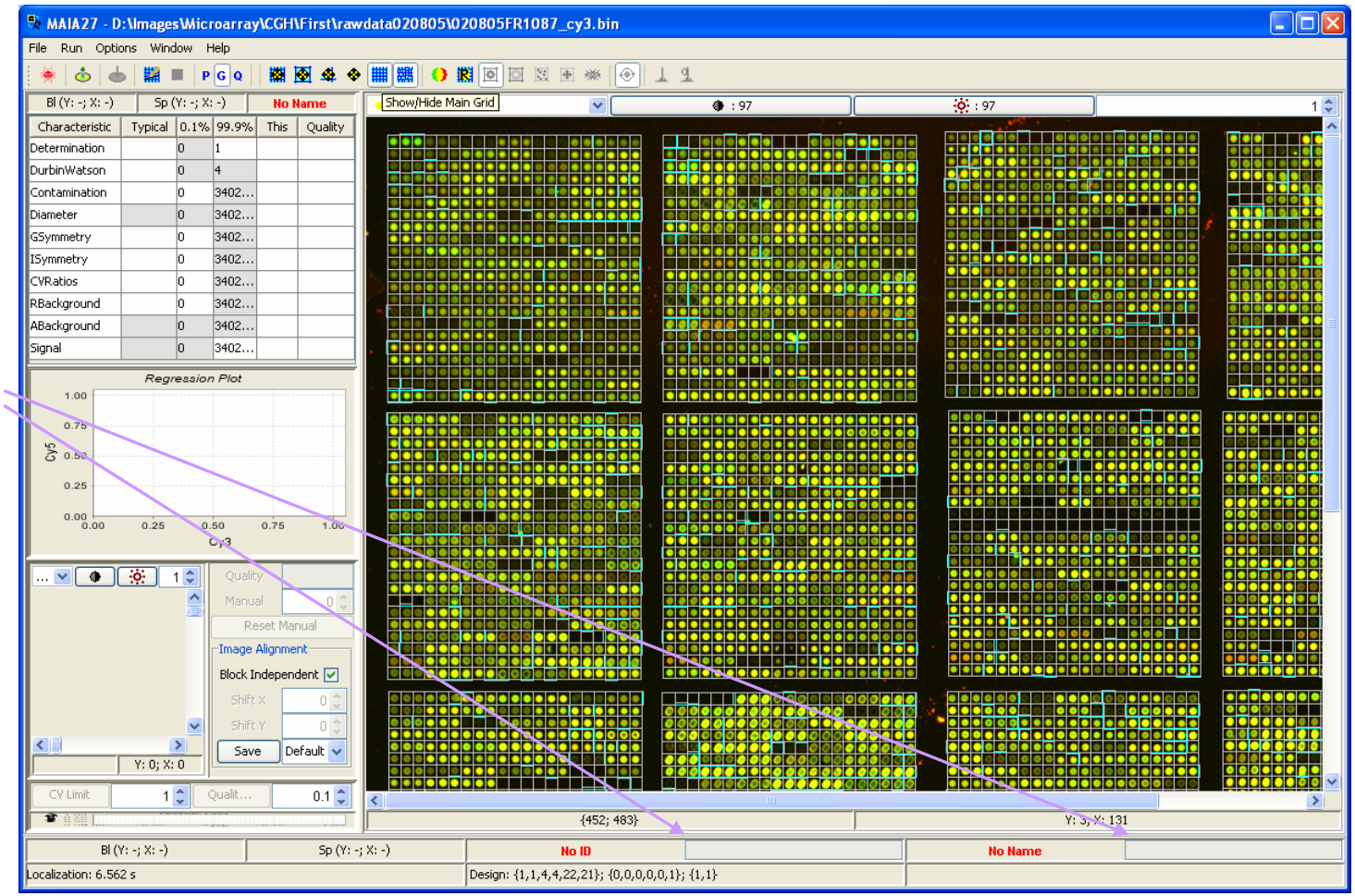
If Array Design is specified using GAL file, spot localization grid can be generated from this file: use the Menu Item “Run|GAL Grid”.



Find ID

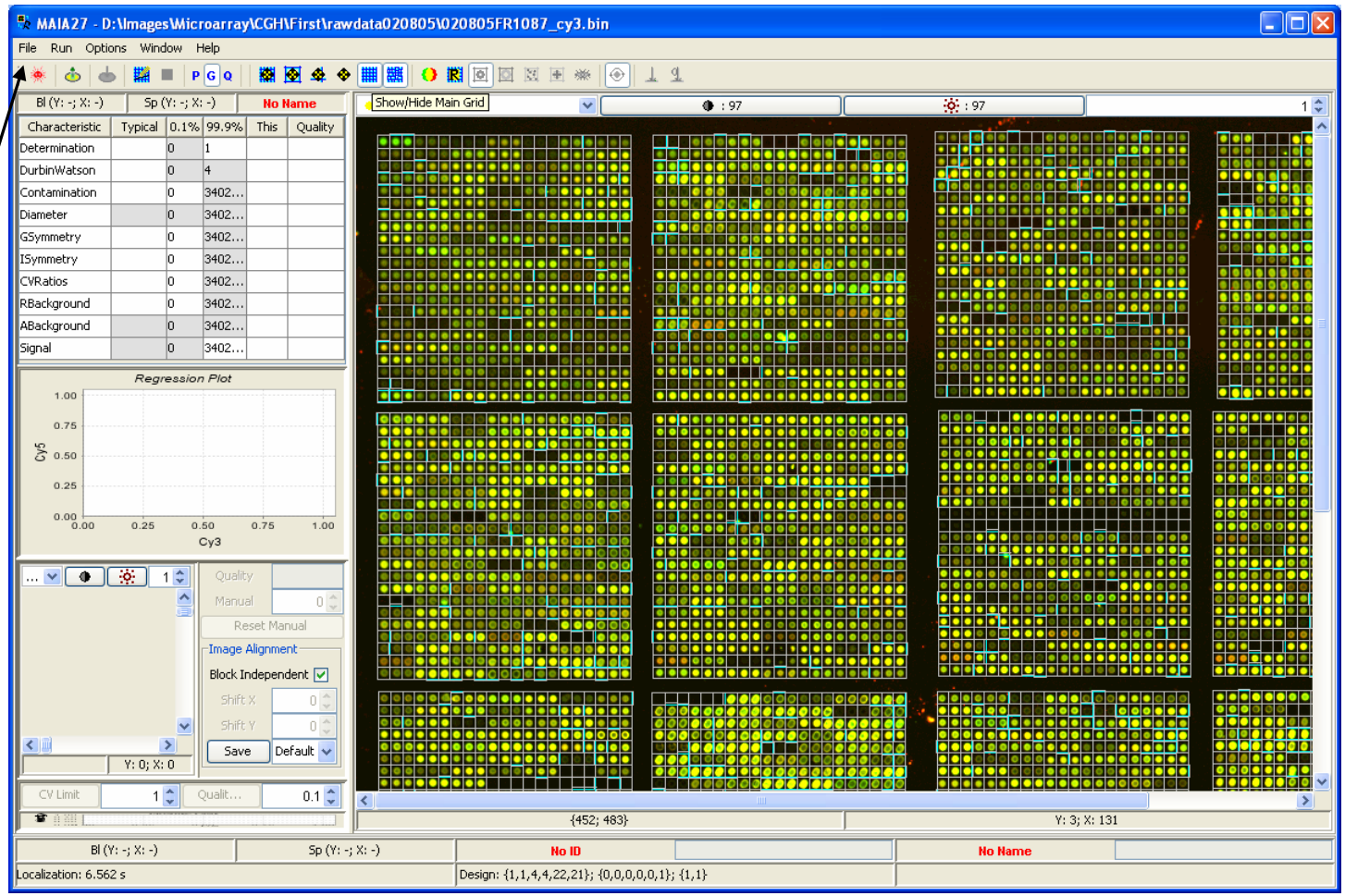
If GAL file contains the IDs and Names for the spotted clones, all spots representing interesting clones can be found: use the “Find ID” or “Find Name” text fields to search for the clones. Found spots will be highlighted.

The searching procedure supports regular expressions. The upper and lower case letters are distinguished. If spots are not found, the search field is highlighted by red.



Save/Restore Grids

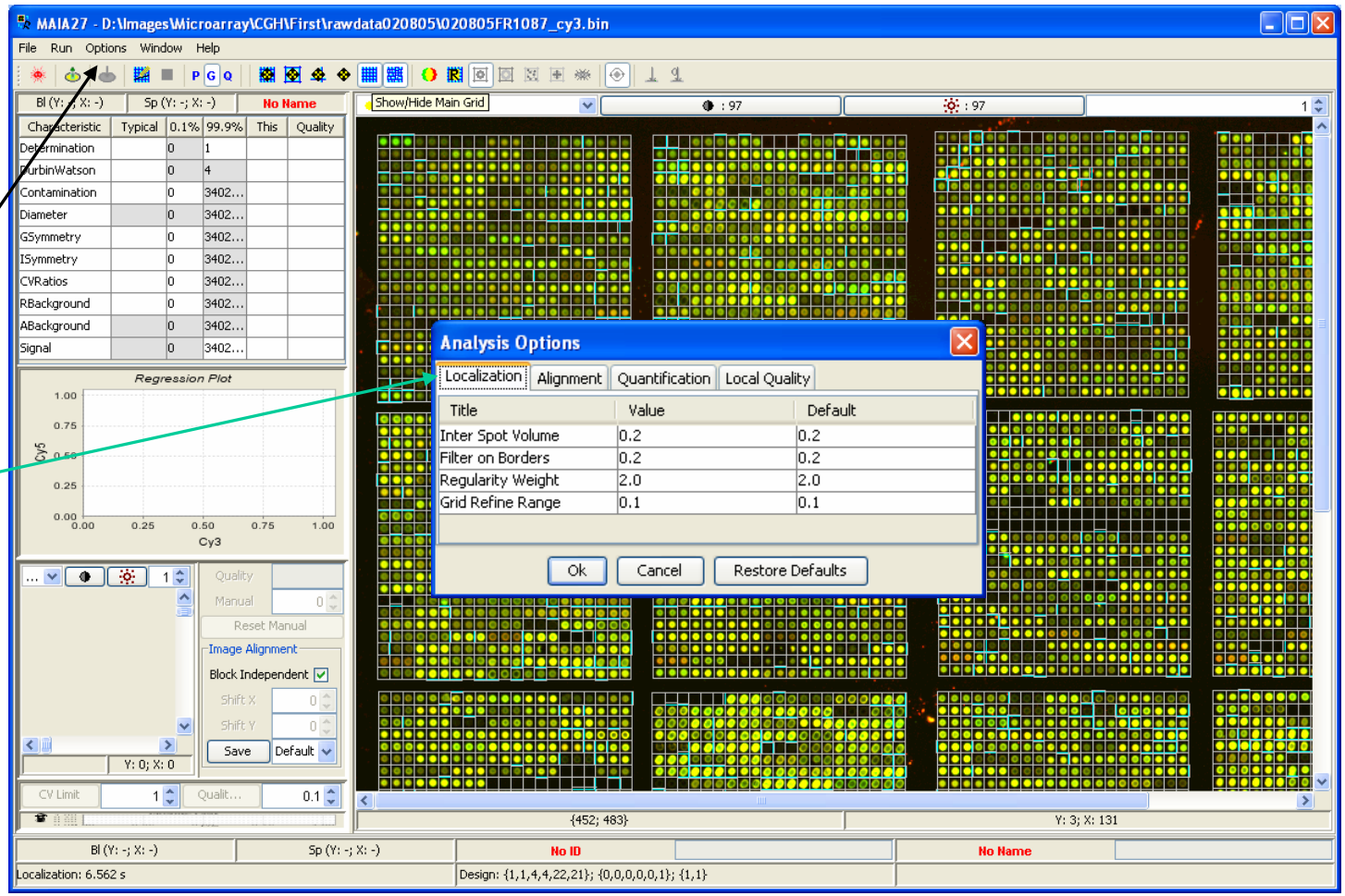
The generated grid can be saved on the disk (using the Menu Item "File|Save|Grid …") to be able to apply it (using the Menu Item "File|Load|Grid …") to analyze other images with the similar design.



Localization Settings

Several settings that may influence the localization procedure are available at the Menu Item “Options|Analysis Options” (Alt+O), tab “Localization”.

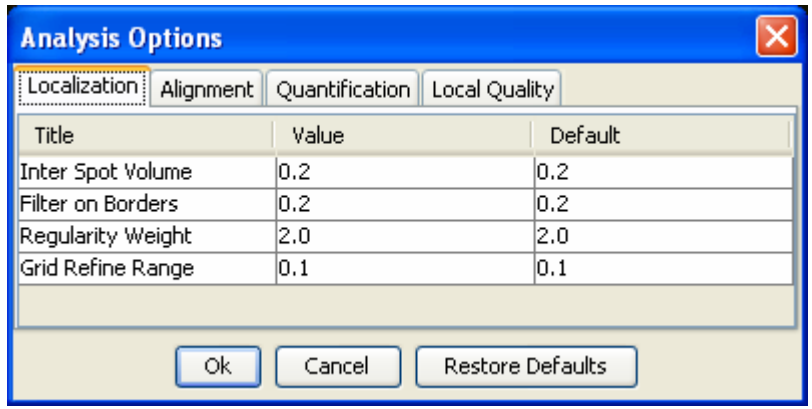
See next page for details.



Localization Settings in Detail

Inter Spot Volume represents (roughly) the ratio of the inter-spot gap to the inter-spot distance.

Regularity Weight controls contribution of the regularity components with respect to the intensity component in the regularity parameter. With the weight equals to 0 the regularity components will be ignored.



Filter on Borders defines filtering properties at the edges of the array. Higher this value, less sensitive the algorithm to the bright regions at the edges of the array.

Grid Refine Range defines the range (related to the inter-spot distance) for the final grid lines adjustments.

The default values of these parameters are suitable for a broad variety of experimental designs.

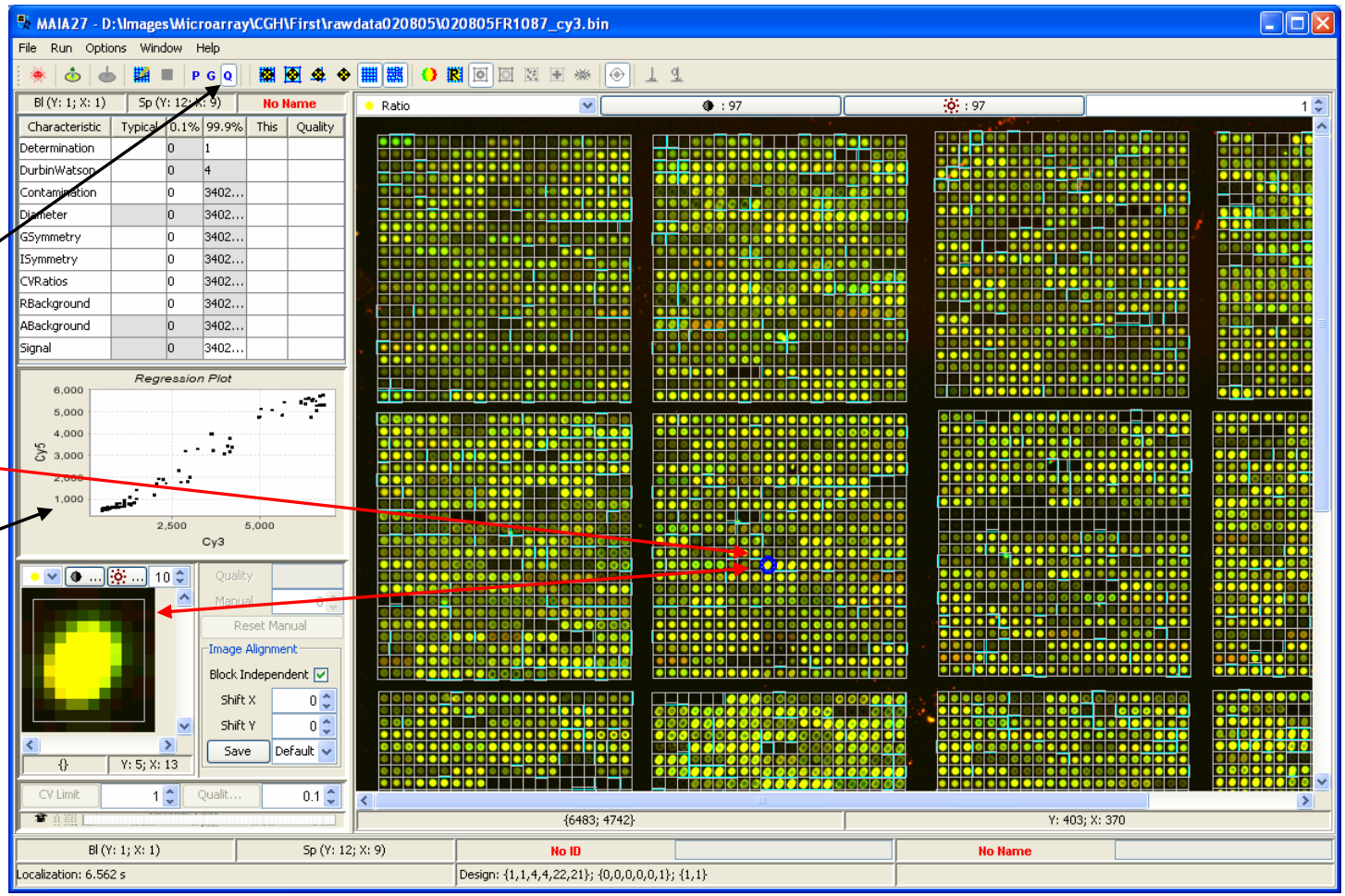
E. Novikov and E. Barillot, A noise-resistant algorithm for grid finding in microarray image analysis. *Machine Vision and Applications*, 2006, 17, 337-345.

Spot Selection

Select the toggle button "Manual Quality Control".

Left Click selects the spot.

Cy5 *vs.* Cy3 intensity plot for the selected spot – *Regression Plot* (*see page* Ratio Estimation).



Selected Spot

Spot can be zoomed using either the “Zoom” spinner box or the mouse wheel.

Brightness and contrast are copied from the whole image window, so that the spot appearances are consistent.
“Contrast” and “Brightness” controls can be used to further adjust brightness and contrast of the selected spot.

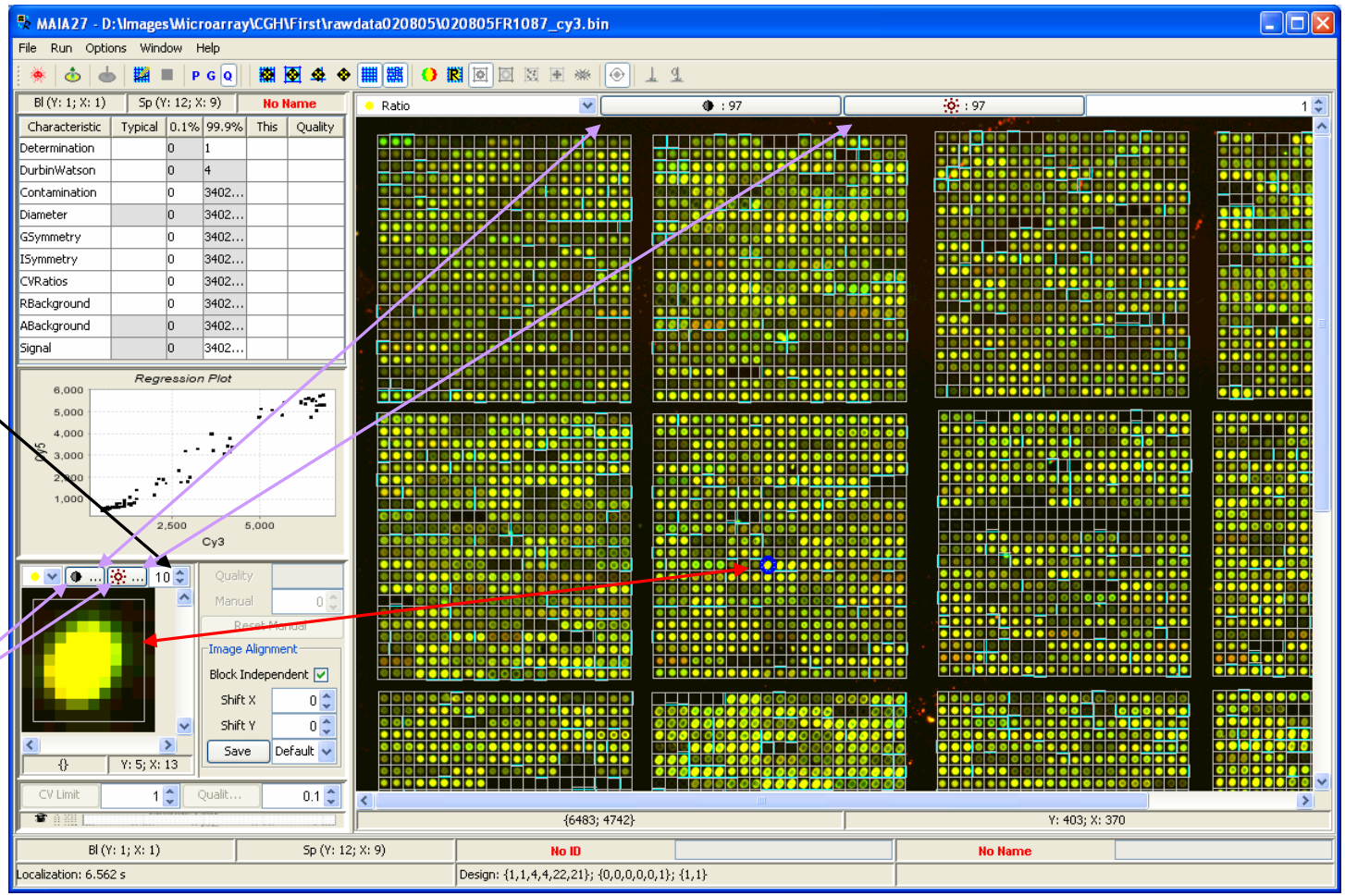


Image Alignment

There may be relative shift between the Cy3 and Cy5 images. The performance of the quantification procedures can be increased, if the two images are aligned.
Use the "Image Alignment" button from the Toolbar or the Menu Item "Run|Image Alignment" (Ctrl+F7) to align images.

The shift value may be the same for all blocks on the array ("Block Independent" is on) or specific for each block ("Block Independent" is off).

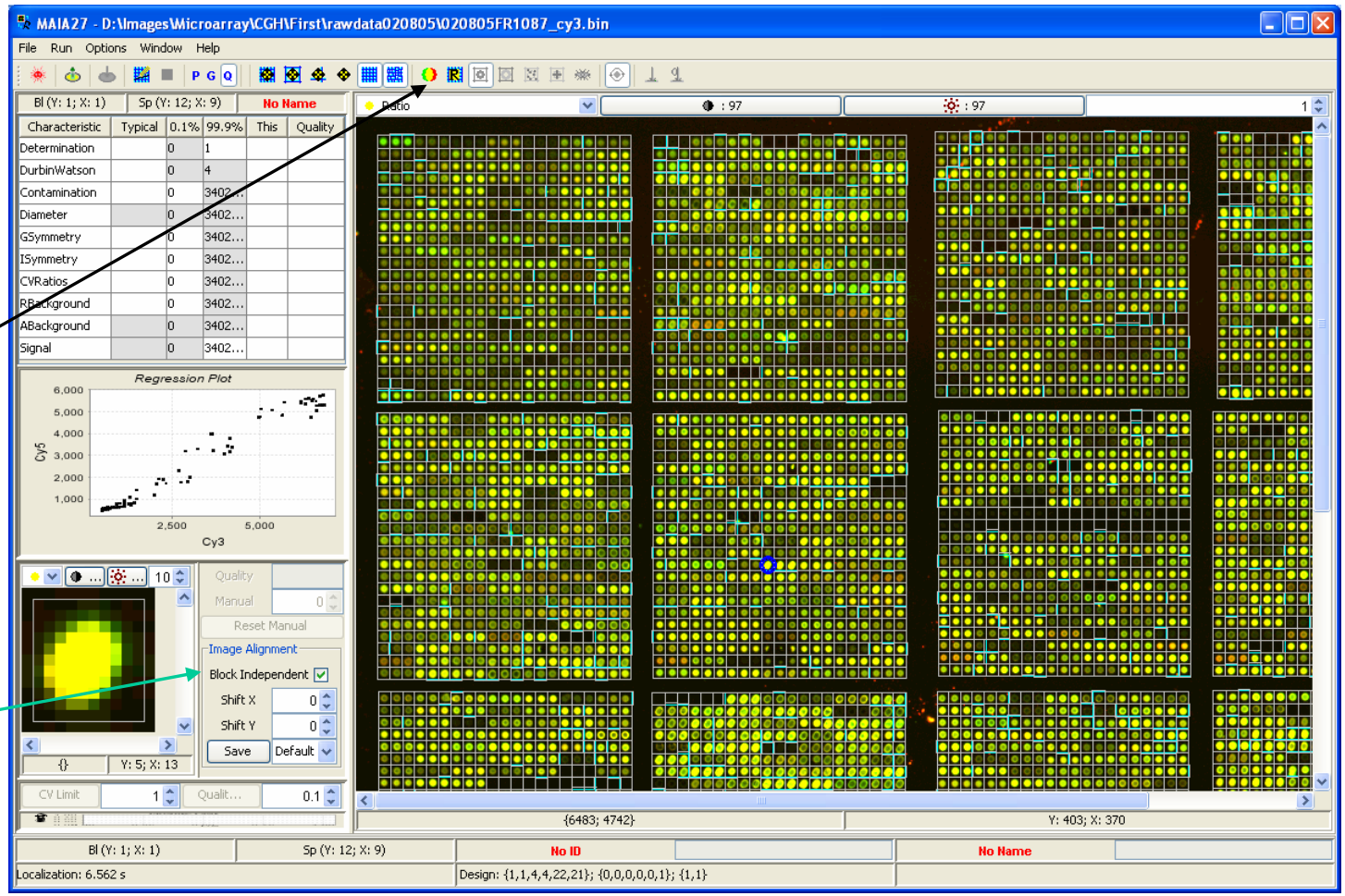
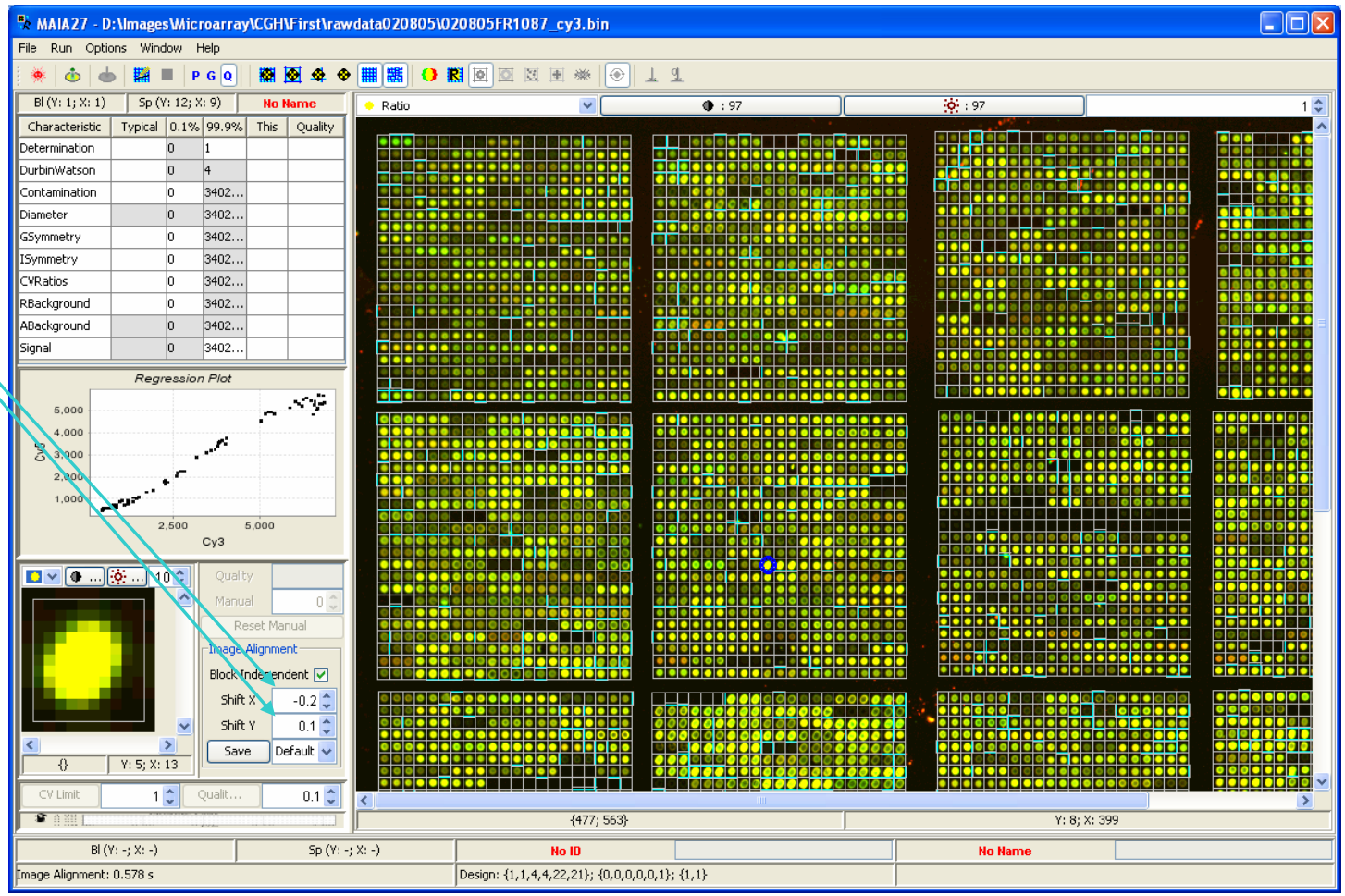


Image Alignment Output

Relative shift (in pixels) in the horizontal (X) and vertical (Y) directions between the Cy3 and Cy5 images.

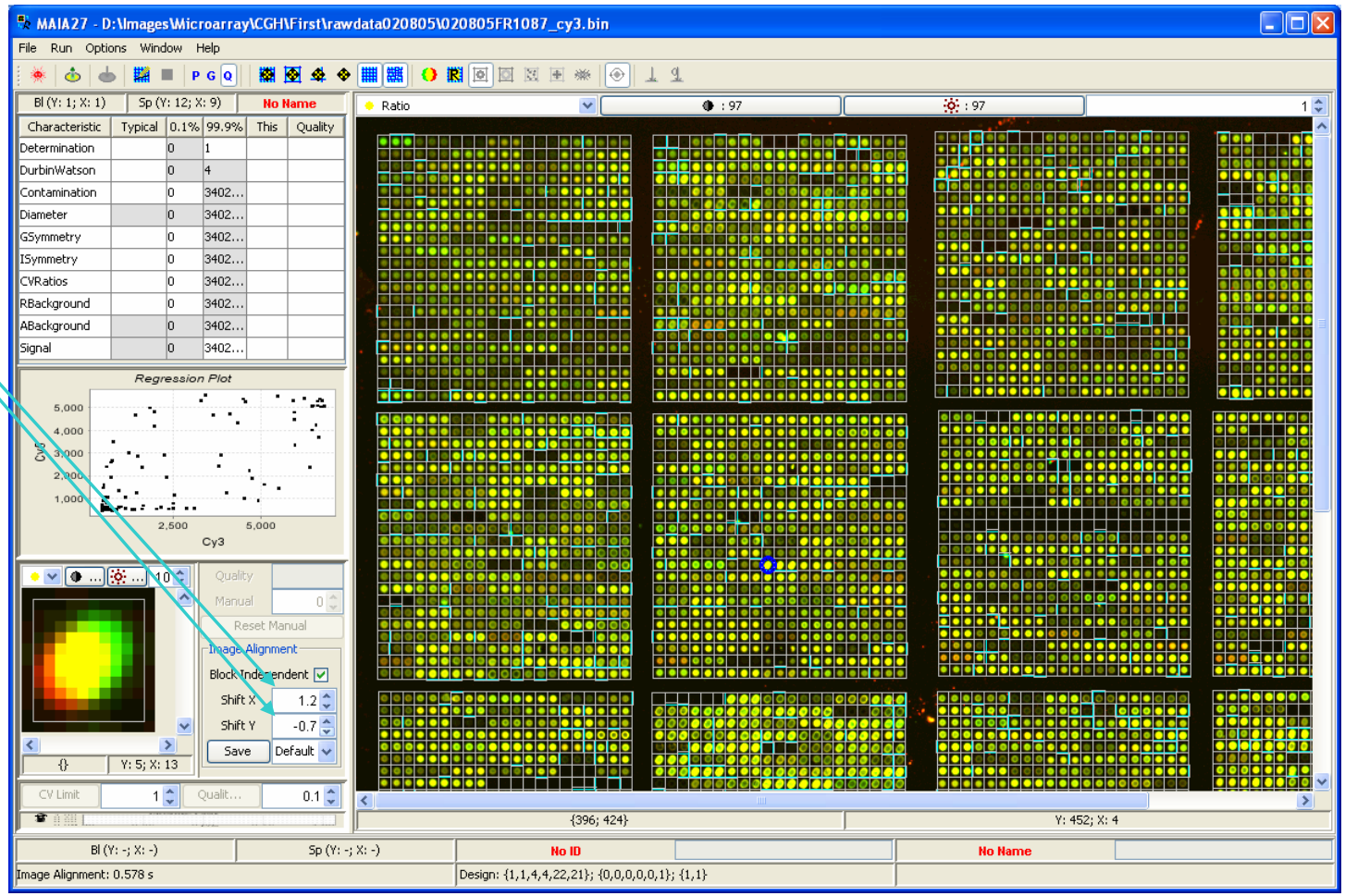
This shift is visualized only for the selected spot and not for the whole image.



Manual Adjustment of the Image Alignment

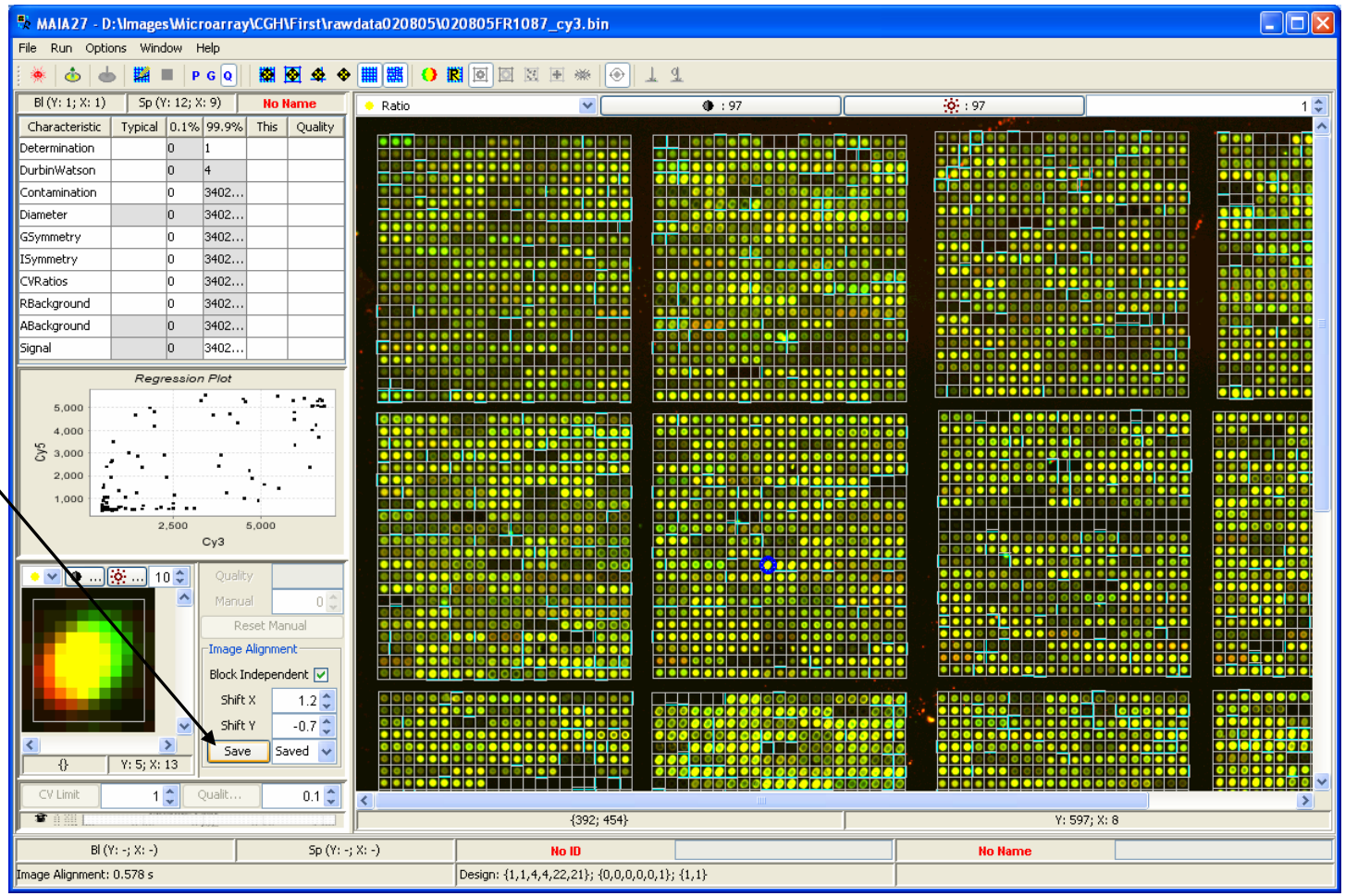
Using the “Shift” spinners one can adjust, if necessary, the shift.

The new values will be valid for all spots from the given block (“Block Independent” is off) or for all spots from the image (“Block Independent” is on).



Compare Different Image Shifts

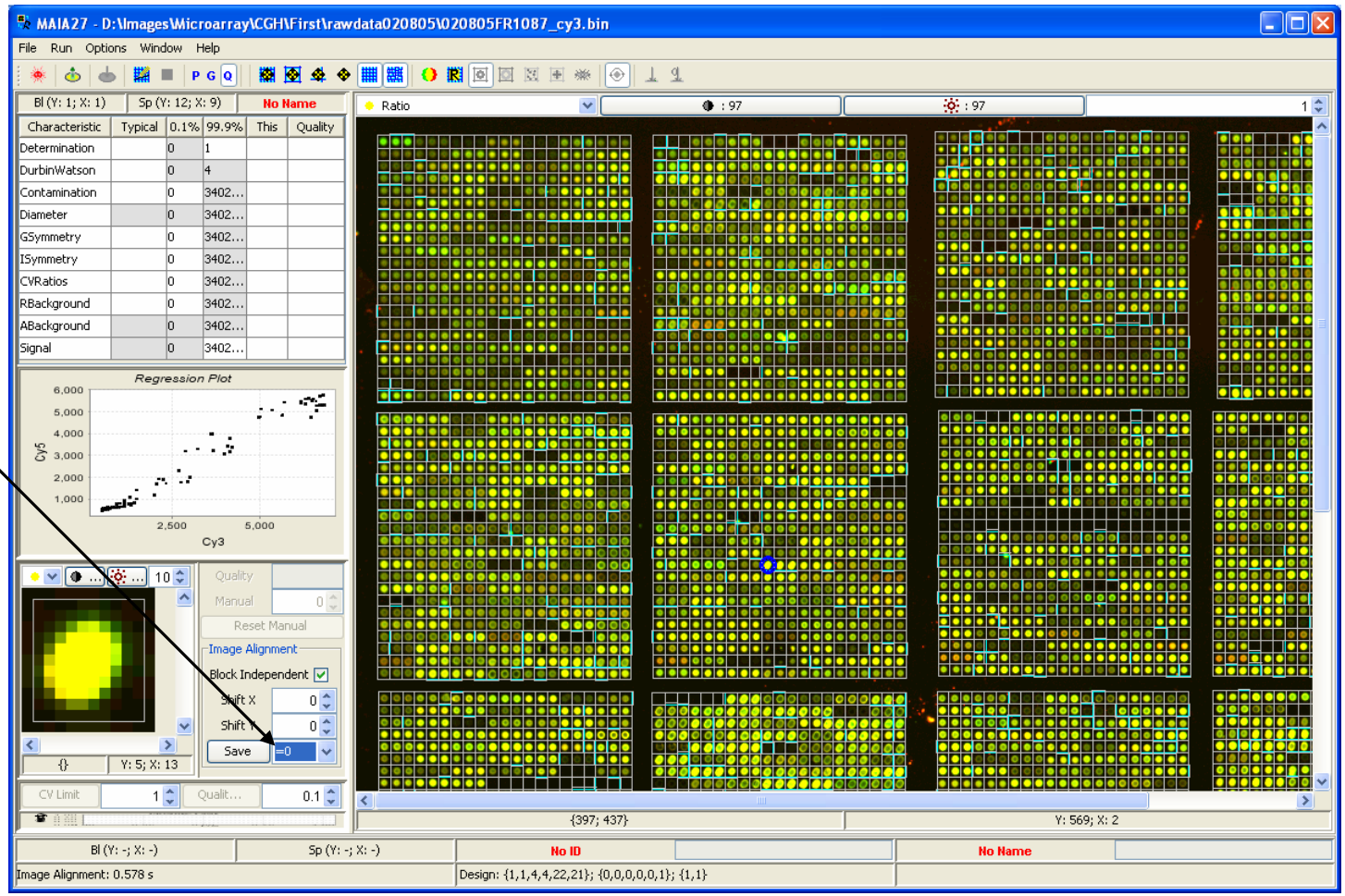
The new values of the shift can be saved (using the button “Save”) and used for comparison with the automatically generated (Default) and Zero (=0) shift values.



Zero Shift

The “Shift” combo box is used to switch between different shift values.

Zero shift is selected.

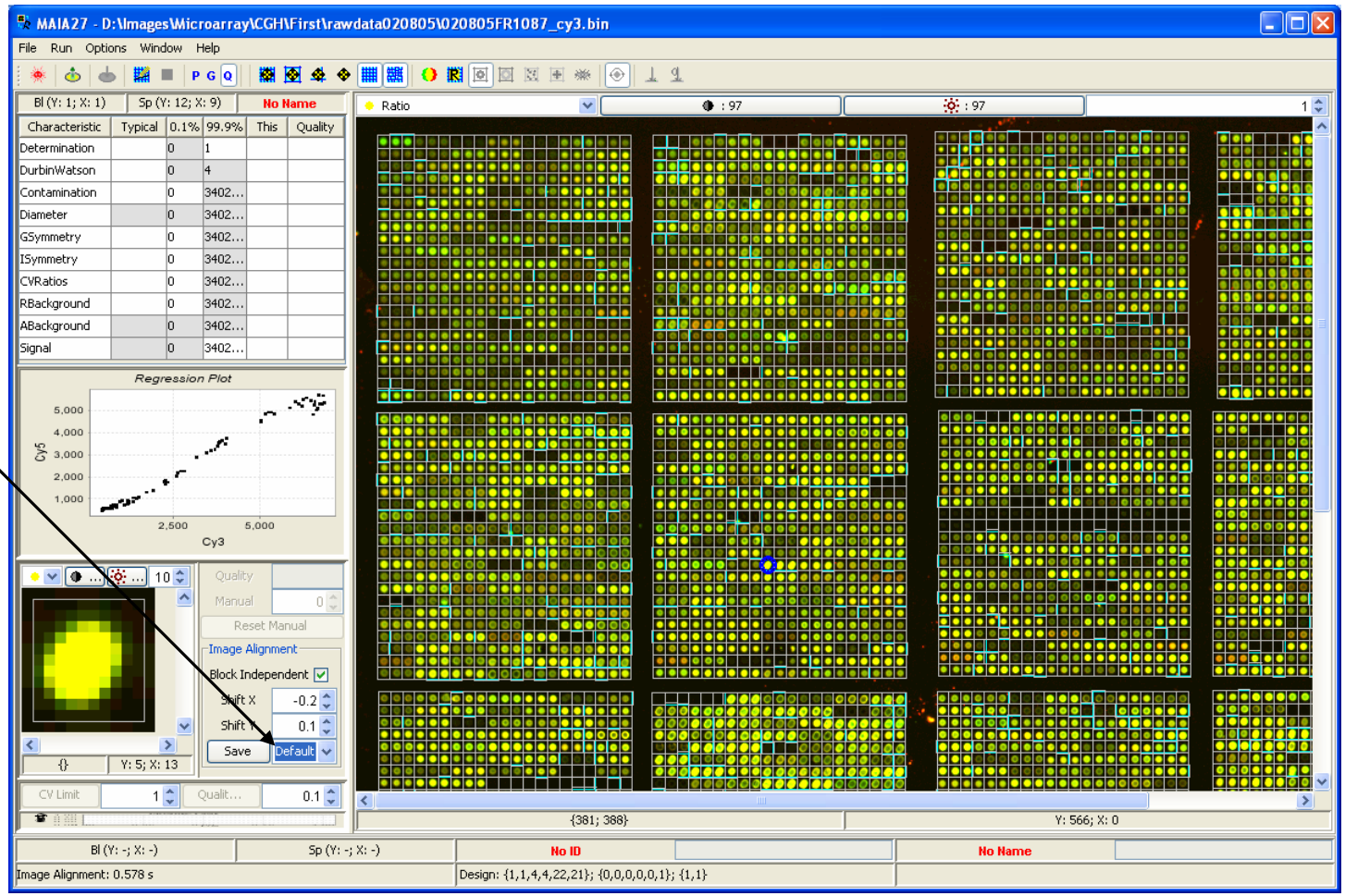


Default Shift

The “Shift” combo box is used to switch between different shift values.

Default shift is selected.

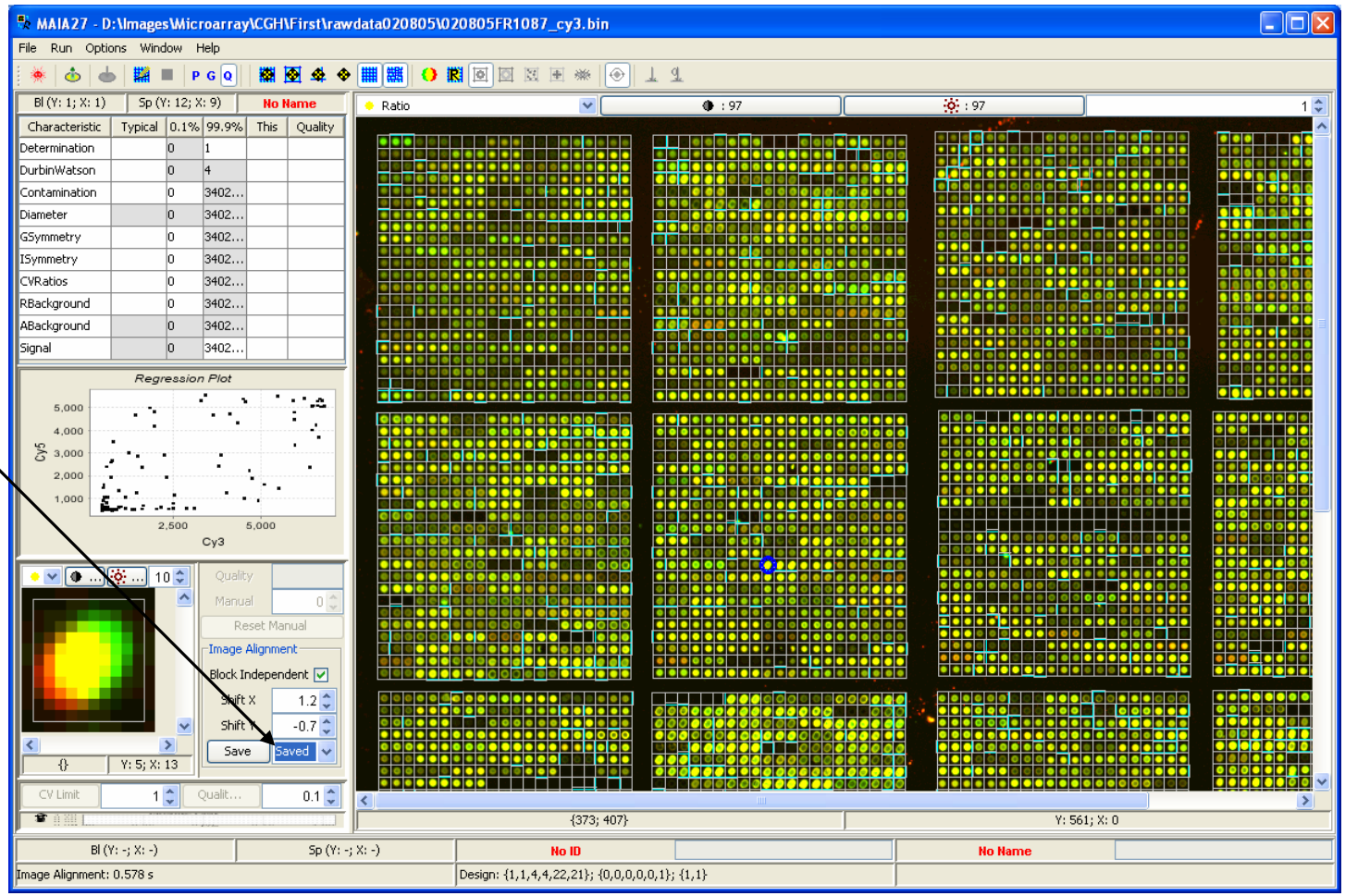
Note the difference in the linear regression plot as compared to the Zero shift.



Saved Shift

The “Shift” combo box is used to switch between different shift values.

Saved shift is selected.



Follow-Up Grid Refinement

Once the two images are aligned, additional grid refinement may be needed: image alignment slightly shifts the spots, so that the border between the spots may not be correct any more.

Use the "Lines Refinement" button from the Toolbar or the Menu Item "Run|Lines Refinement".

Image alignment is important in order to increase the efficiency of the linear regression filtering.

Removal of the shift enhances the correlation between the two color channels thus making uncorrelated pixels easier detectable.

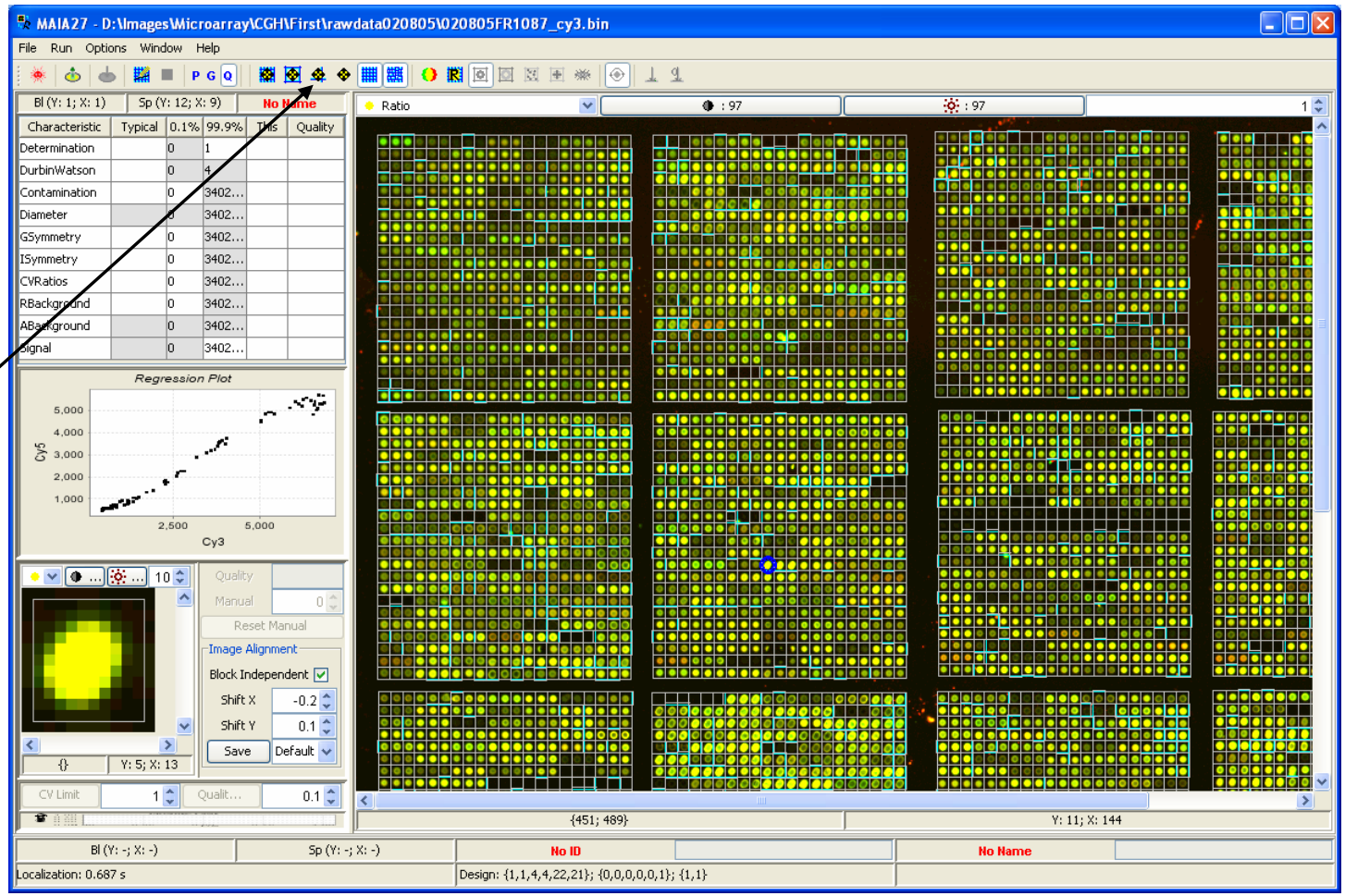


Image Alignment Settings

Several settings that may influence the image alignment procedure are available from the Menu Item “Options|Analysis Options” (Alt+O), tab “Alignment”.

See next page for details.

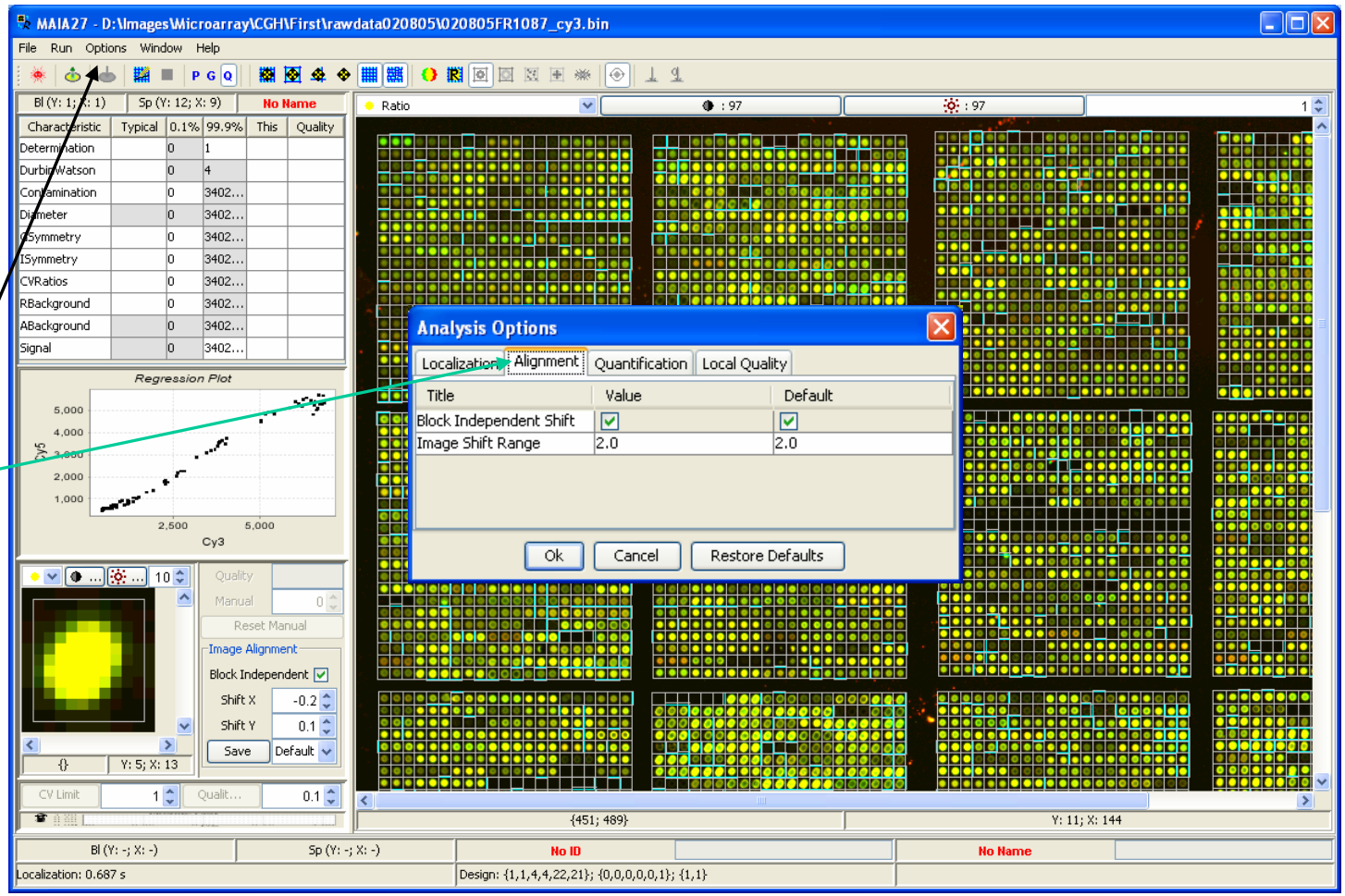


Image Alignment Settings in Detail

Block Independent Shift defines whether the shift is the same for all blocks on the array (on) or it is specific for each block (off).

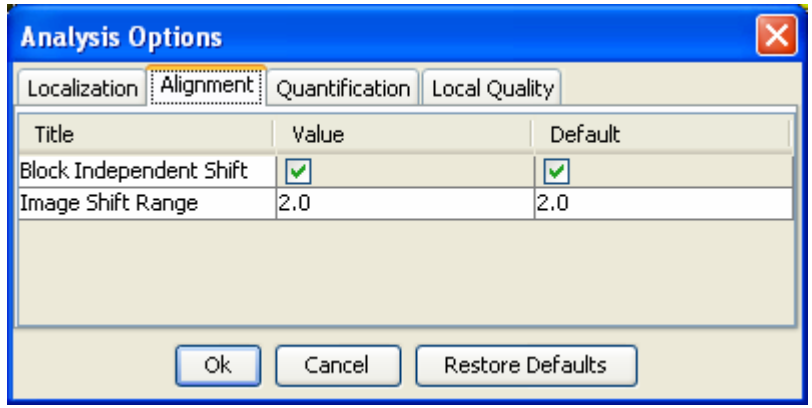
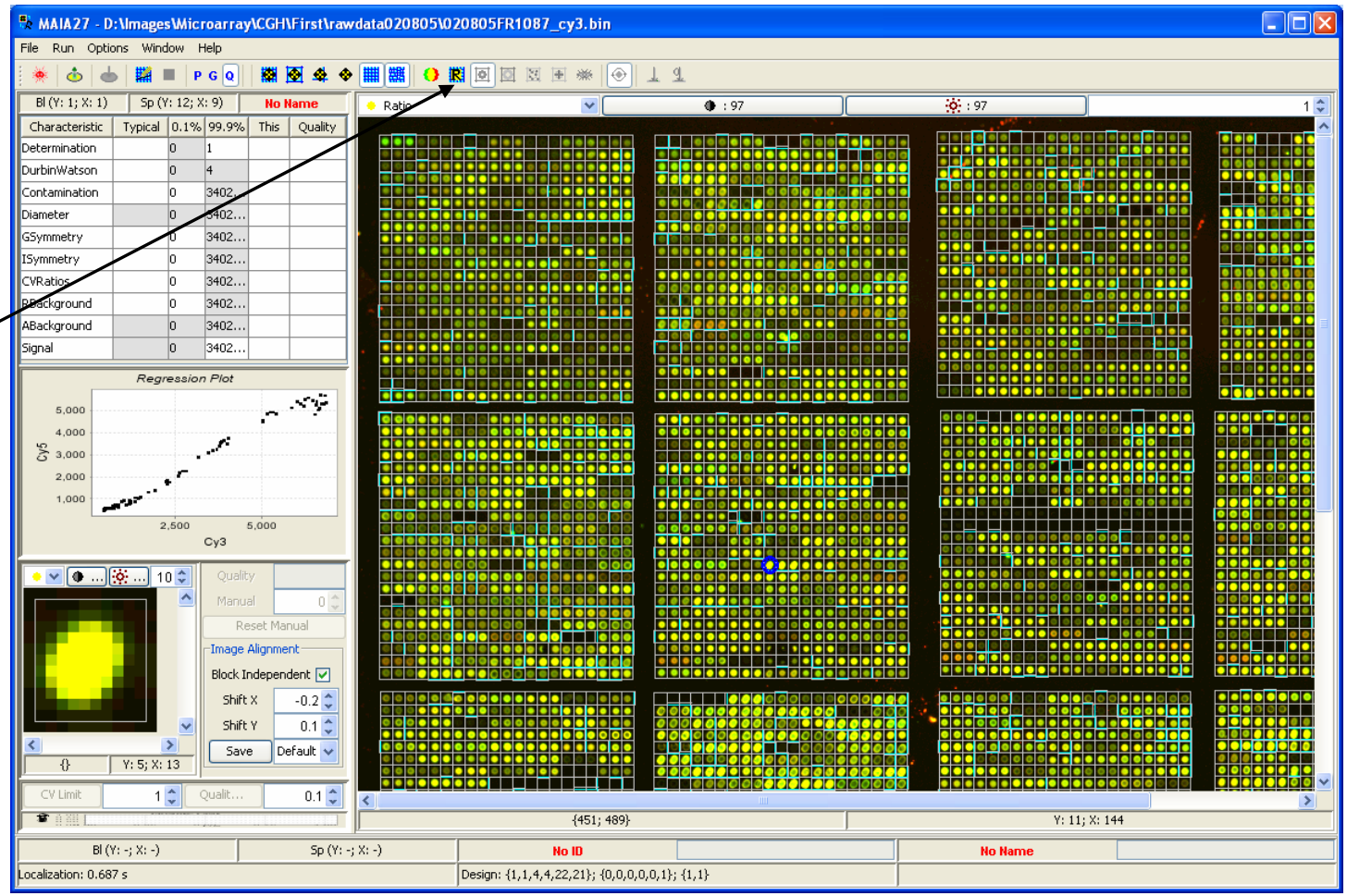


Image Shift Range establishes the boundaries (in pixels) for the relative shift between the two images (2 pixels, by default).

Spot Quantification

To start Spot Quantification use the "Spot Quantification" button from the Toolbar or the Menu Item "Run|Spot Quantification" (Ctrl+F8).



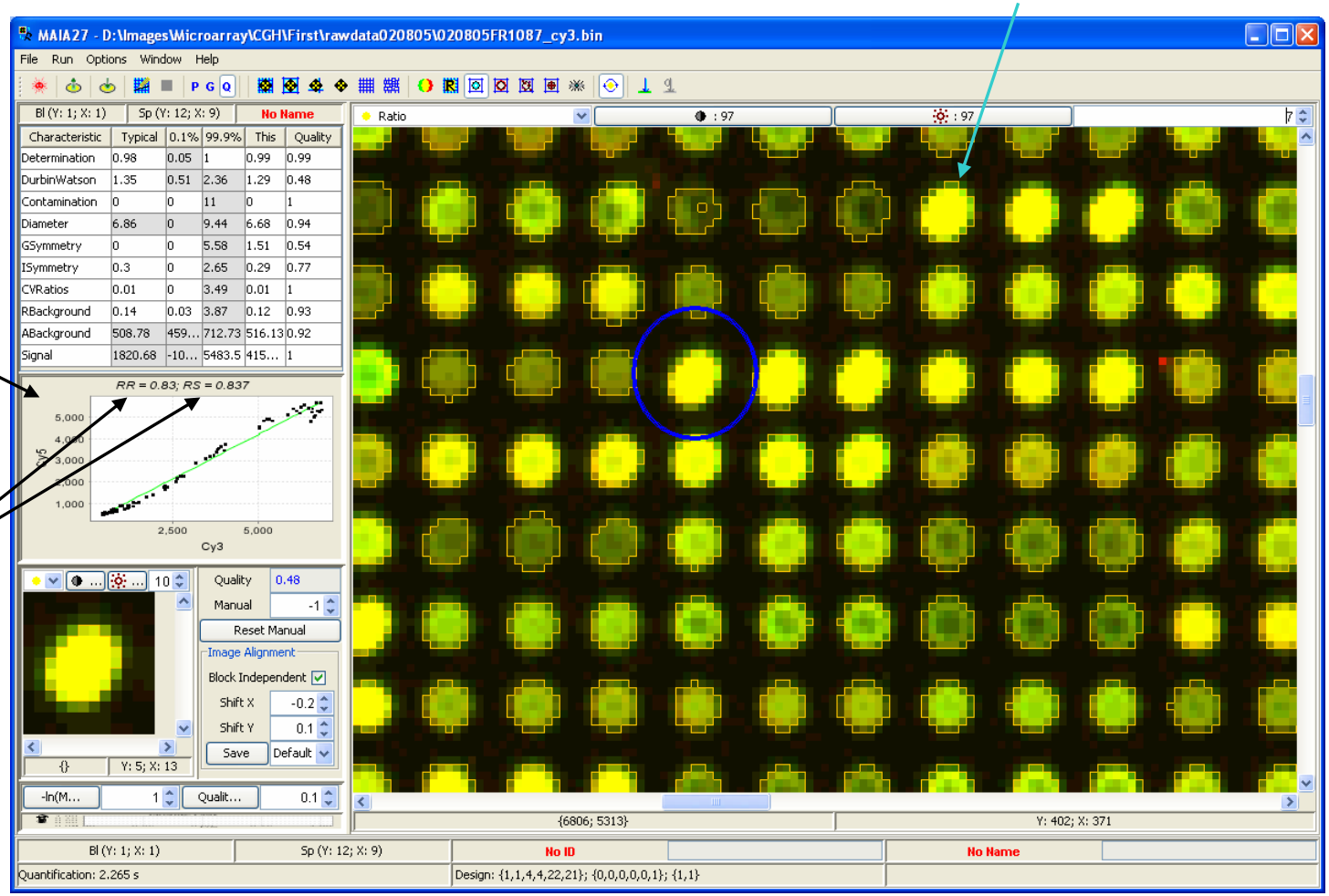
Ratio Estimation

Spots are contoured.

Linear regression plot for the selected spot.

Two ratio estimates:
RR is based on the slope of the linear regression;
RS is based on the segmentation of the spot area

See next page for details.



*Ratio Estimation in Detail**

Segmentation Ratio. This approach is based on isolation of the spot pixels from the background pixels surrounding the spot. Once this is done, the quantification procedure is fairly straightforward: one can compose the following ratio:

$$R=\frac{S_{Cy5}-B_{Cy5}}{S_{Cy3}-B_{Cy3}}$$

where S_{Cy5}(S_{Cy3}) is the mean estimate of the intensity within the contoured spot in the Cy5(Cy3) channel, and B_{Cy5}(B_{Cy3}) is the mean estimate of the background level in the Cy5(Cy3) channel. Mean estimates are known to be more precise, but they can be very much affected by the outliers. Since regression filtering eliminates outliers, we can safely use mean estimates for the spots.

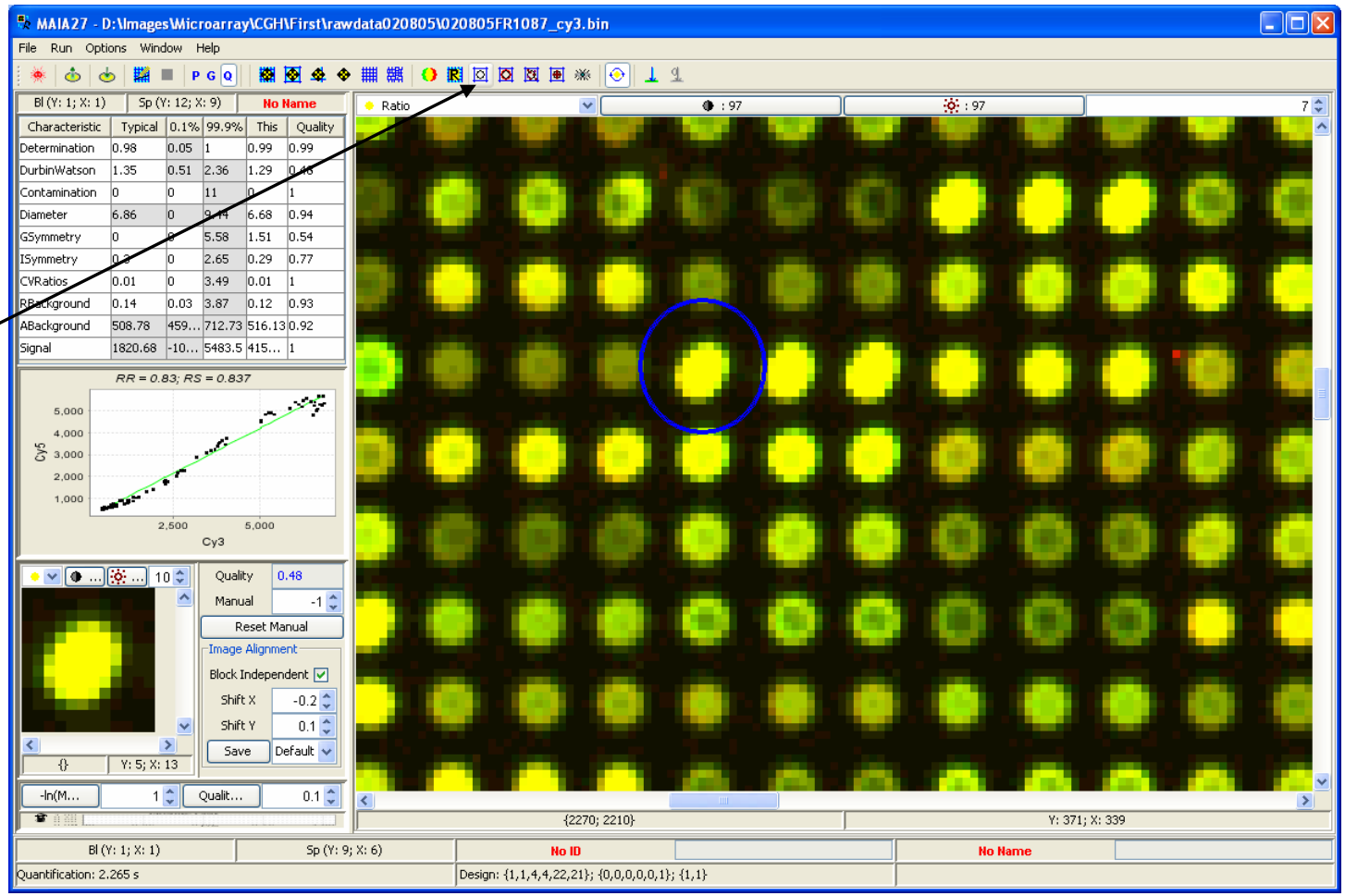
Regression Ratio. In this approach a ratio can be represented as a slope of the linear regression line of the pixel intensities in, say, Cy5 channel versus Cy3 channel. The main advantage of this method is that the obtained ratio is directly delivered from the regression analysis, thus making the procedure of spot segmentation unnecessary. Background pixels are concentrated at the initial part of the linear regression and do not influence the slope of the regression line. However the linear regression approach suffers from the presence of the outlier or aberrant pixels within the spot cells. These pixels, occurring even in small quantities, can distract the regression line and strongly bias the regression ratio. With the aim to fully exploit the advantages of the linear regression approach we have reinforced this procedure by systematical filtering out aberrant pixels

See page Pixel Regression Outliers.

*) E. Novikov and E. Barillot, A robust algorithm for ratio estimation in two-color microarray experiments. *Journal of Bioinformatics and Computational Biology*, 2005, 3, 1411-1428.

Spot Contours

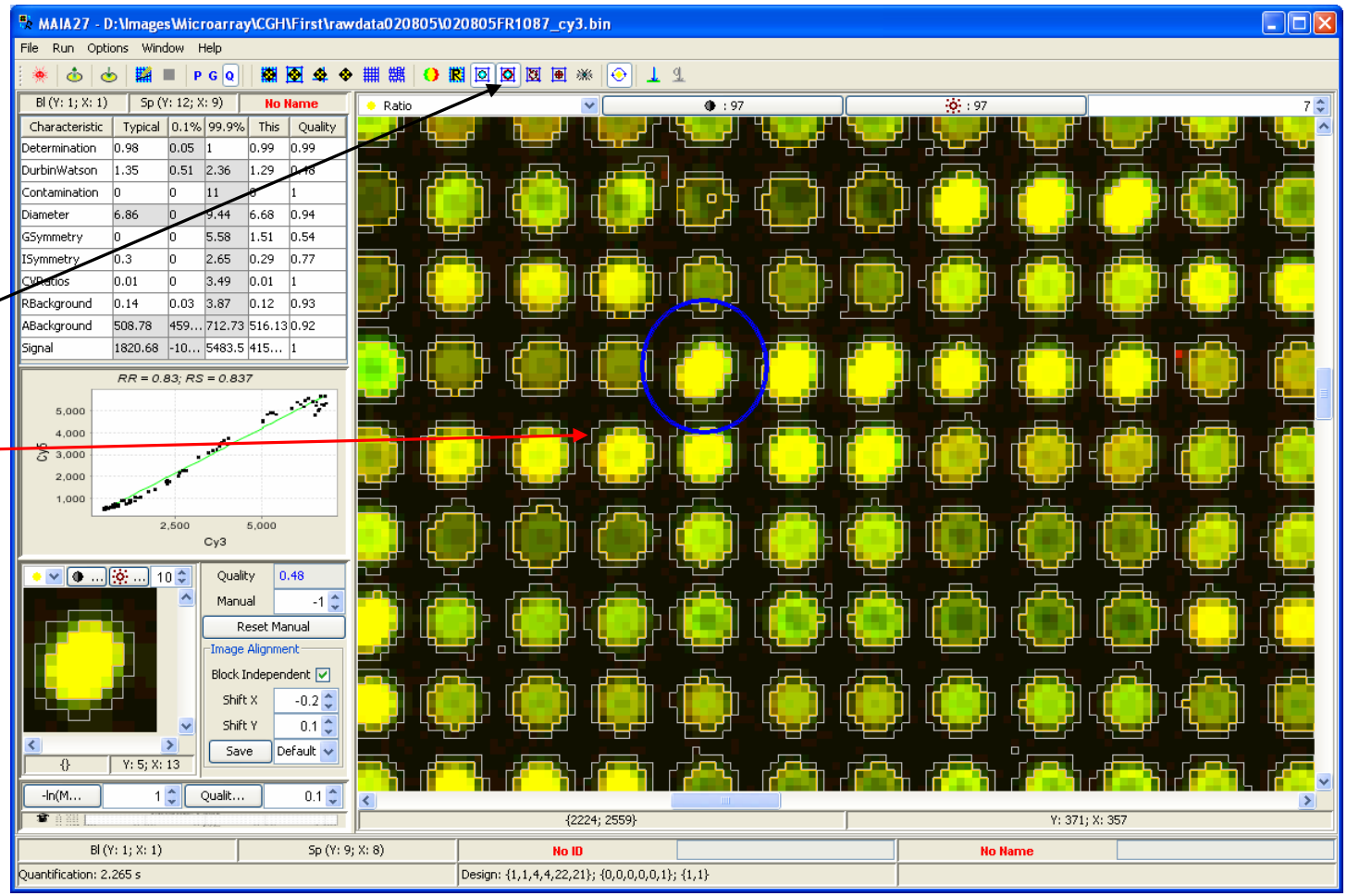
Using the Toolbar button "Show/Hide Spot Contour" one can control whether the spot contours are visible.



Background Contours

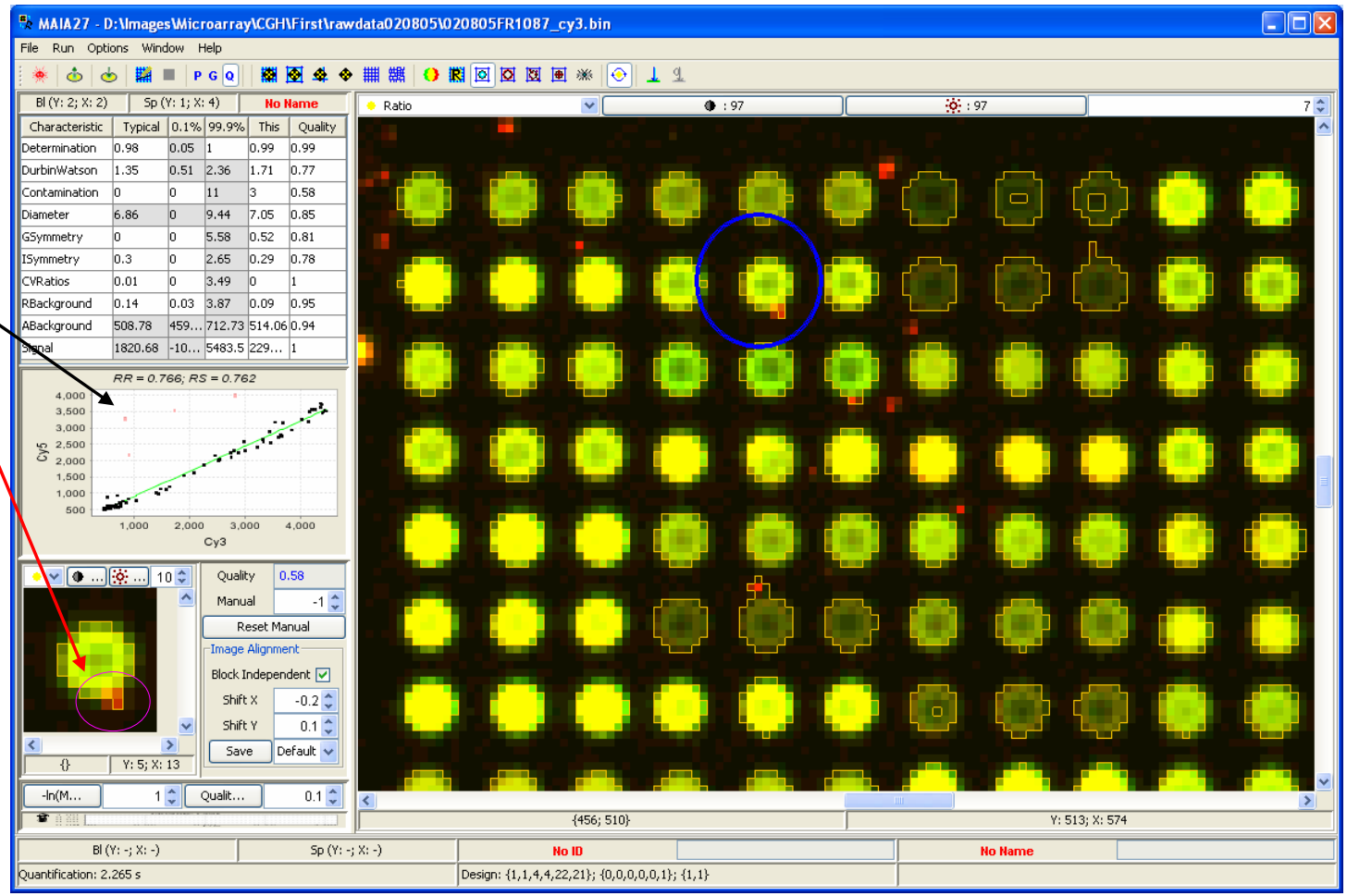
Using the Toolbar button "Show/Hide Background Contour" one can control whether the background contours are visible.

Segmentation procedure creates two contours: pixels within the spot contours are used to estimate the signal $S_{Cy5}(S_{Cy3})$, pixel outside the background contours are used to estimate the background $B_{Cy5}(B_{Cy3})$ and pixels that are between the two contours are ignored.



Pixel Regression Outliers

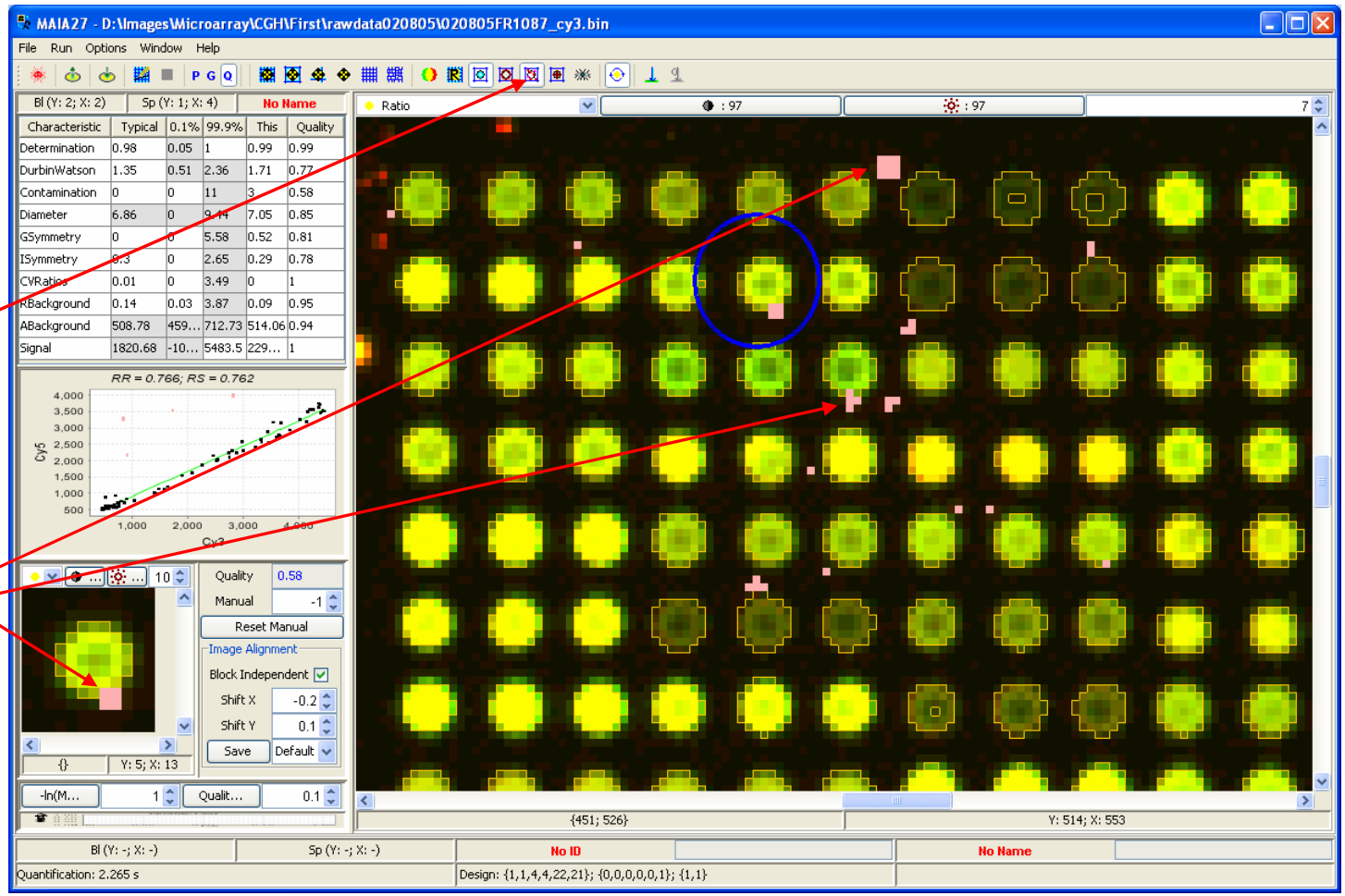
Pixel regression outliers.



All Pixel Regression Outliers

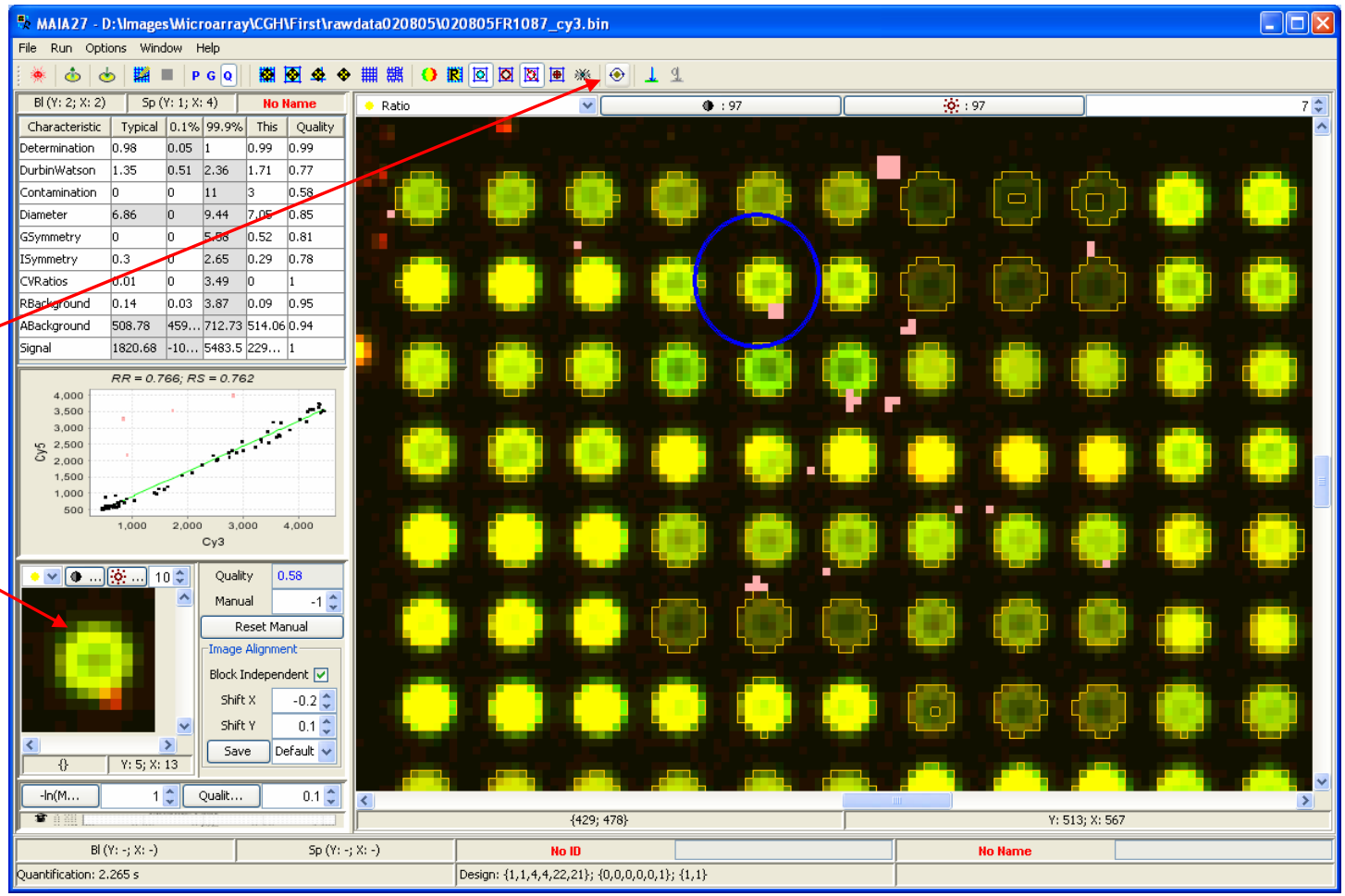
Using the Toolbar button "Show/Hide Outlier Pixels" one can visualize all pixel regression outliers.

Pixel regression outliers.



Decorations for the Selected Spot

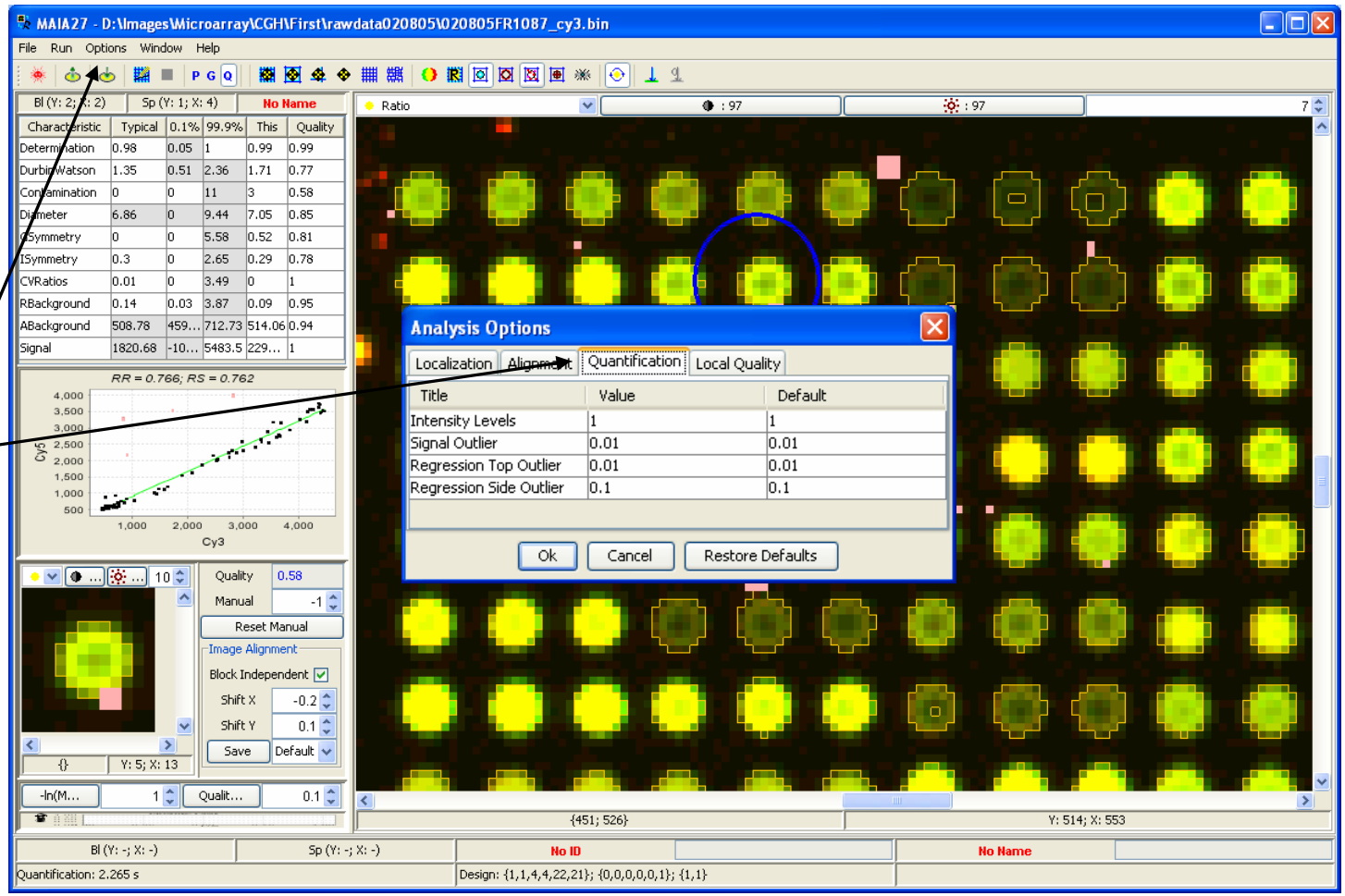
Using the Toolbar button "Show/Hide Inset Decorations" one can control whether the contour, pixel outliers and grid lines are visible for the selected spot.



Quantification Settings

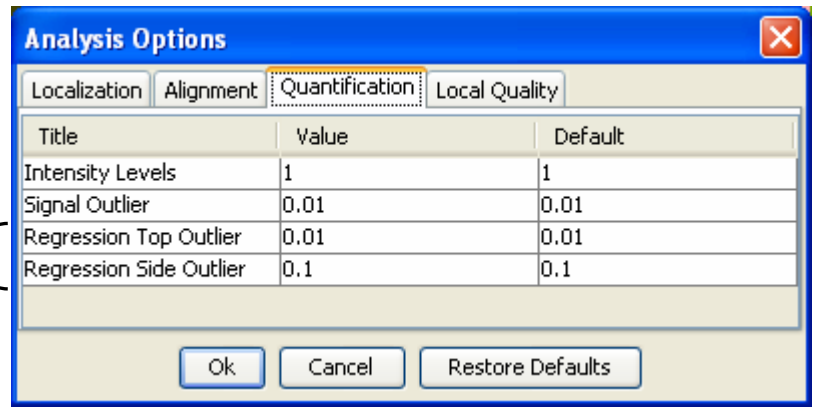
Several settings that may influence the quantification procedure are available from the Menu Item "Options|Analysis Options" (Alt+A), tab "Quantification".

See next page for details.



Quantification Settings in Detail

Visible spots may have several more or less well defined intensity levels. *Intensity Levels* specifies how many such levels should be identified at the spot. Spots will be segmented at the highest level of intensity.



Outlier Limit Top/Sides defines critical *p*-values of the *F*-statistics in the detection of the pixel outliers selected from the top of the intensity ranges and from the sides of the linear regression fit.

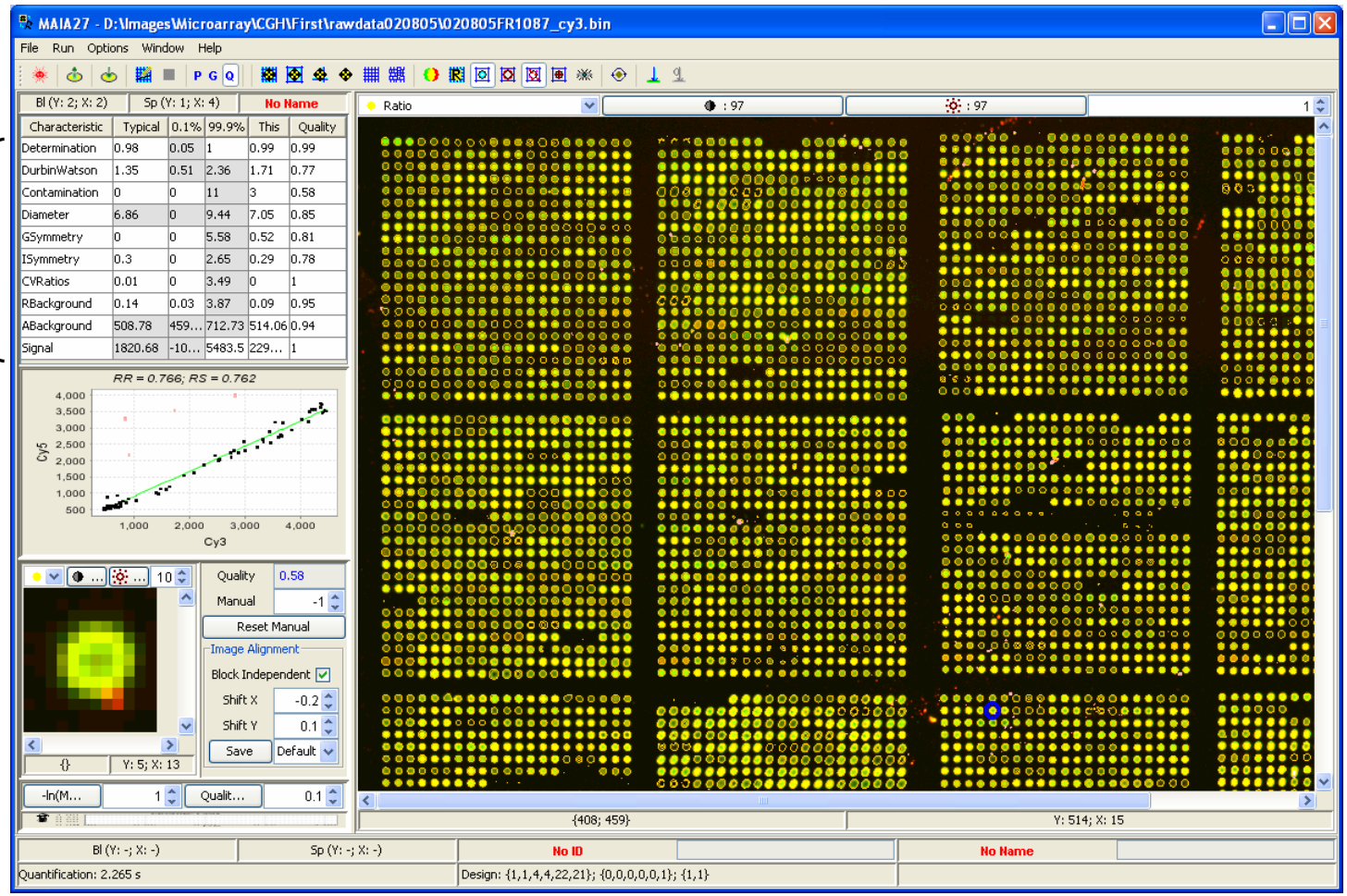
Spot pixels with excessively high or low intensity with respect to majority of the spot pixels are discarded. The admissible range is defined as "median of spots pixels" ± *n**"inter-quartile distance of the spot pixels"/1.35, where *n* = 1/*p*½, and *p* is a user-defined *Signal Outlier* confidence limit. This filtering procedure is appropriate for the spots with large amount of pixels.

The default values of these parameters are suitable for a broad variety of experimental designs.

Quality Characteristics

Quality characteristics of the spots.

See next page for details.



E. Novikov and E. Barillot, An algorithm for automatic evaluation of the spot quality in two-color DNA microarray experiments. *BMC Bioinformatics*, 2005, 6:293.

Quality Characteristics in Detail

Coefficient of determination (*CD*) of the linear regression indicates the degree of linear relationship between the intensities in Cy3 and Cy5 channels. For higher quality spots relatively high values of determination coefficient ($\rightarrow 1$) are expected. Much lower values would point on either strong contribution of statistical noise, which normally characterizes low-level (or absent) spots, or presence of a relatively bright but non-correlated contamination. $q(CD) = CD^*$.

Durbin-Watson statistic (*DWS*) controls the presence of first-order autocorrelation in the residuals of the linear regression fit. It ranges from 0 to 4, 0 meaning positive correlation and 4 – negative correlation. $DWS \cong 2$ leads to the conclusion that the residuals are uncorrelated and the model is appropriate. Large departures from 2 suggests that this spot can not be modeled in terms of simple linear regression. $q(DWS) = 1-|DWS-2|/2^*$.

Spot contamination is a number of aberrant pixels (within the spot contours) flagged out by the filtering procedure (*N*). $q(N) = 1-N/S$, where *S* is the size of the correspondent spot, i.e. the number of pixels within the spot contour*.

Diameter of the spot: $D = 2(S/\pi)^{1/2}$. Since it is hard to impose *a priory* an exact ideal value for the diameter, the median diameter over all spots on the array is taken as a typical one. Spots with exceptionally small or large diameters should normally be penalized. $q(D) = exp\{T_D-D\}$, if $D>T_D$ and $q(D) = exp\{T_D-D\}$, if $D<T_D$ where T_D is the typical diameter*.

Geometrical symmetry parameter measures deviation of the contoured spot from the ideal circle. Both the real spot and the ideal circle are divided into 8 sectors (pie slices defined as $[k\pi/4;(k+1)\pi/4]$, $k = 0,\ldots,7$) and for each sector the number of pixels belonging to the spot (N_{si}, $i = 1,\ldots,8$) and to the circle (N_{ci}, $i = 1,\ldots,8$) is counted. Then the quality characteristic is defined as $GS = \sum|N_{si}- N_{ci}|/ N_{ci}$. For ideal circular spots *GS* must approach 0, whereas highly un-circular spots should give relatively high *GS* values. $q(GS) = exp(-GS)^*$.

Intensity symmetry of the spot is defined as $IS = \sum|I_i-I|/I$, where I_i, $i = 1,\ldots,8$ are the mean intensities for the same 8 sectors and *I* is the mean intensity for the whole spot A spot may have perfect circular shape, but within this circle very bright (or dark) and highly concentrated groups of pixels originated from the pieces of dust or other contamination may occur. $q(IS) = exp(-IS)^*$.

Coefficient of variation of two ratio estimates: $CVR = 2^{1/2}|RR-RS|/(RR+RS)$. Despite the differences in the estimation, the variation between the two obtained ratios *RS* and *RR* should be as small as possible. Large variation would indicate a problematic spot. $q(CVR) = exp(-CVR)^*$.

Uniformity of the background around the spot, i.e. along the grid lines separating neighborhood spots, is defined as $UB = \sum|B_i-B|/B$, where B_i, $i = 1,\ldots,8$ are the mean intensities in 8 sectors of the grid line around the spot, and *B* is the mean intensity for the whole grid line around the spot. Extremely small values may be due to relatively bright contamination around the spot, large variability in the background or merged neighborhood spots. $q(UB) = exp(-UB)^*$.

Absolute level of background (*AB*) calculated in the proximity of each particular spot ($AB = \max(B_{Cy5},B_{Cy3})$) is compared to the typical level of the local background for a given array. Large deviations from the typical state may indicate the presence of the contamination areas, which are larger than the size of the spot. $q(AB) = exp(1-AB/T_{AB})$, if $AB>T_{AB}$ and $q(AB) = exp(AB/T_{AB}-1)$, if $AB<T_{AB}$. where T_{AB} is the typical background level*.

Signal (*S*) is defined as $S = \min(S_{Cy5} - B_{Cy5},S_{Cy3} - B_{Cy3})$, where $S_{Cy5}(S_{Cy3})$ is the mean estimate of the intensity within the contoured spot in the Cy5(Cy3) channel, and $B_{Cy5}(B_{Cy3})$ is the mean estimate of the background level in the Cy5(Cy3) channel. $q(S) = 1$, if $S>T_S$ and $q(S) = exp(S/T_S-1)$, if $S<T_S$. where T_S is the typical signal*.

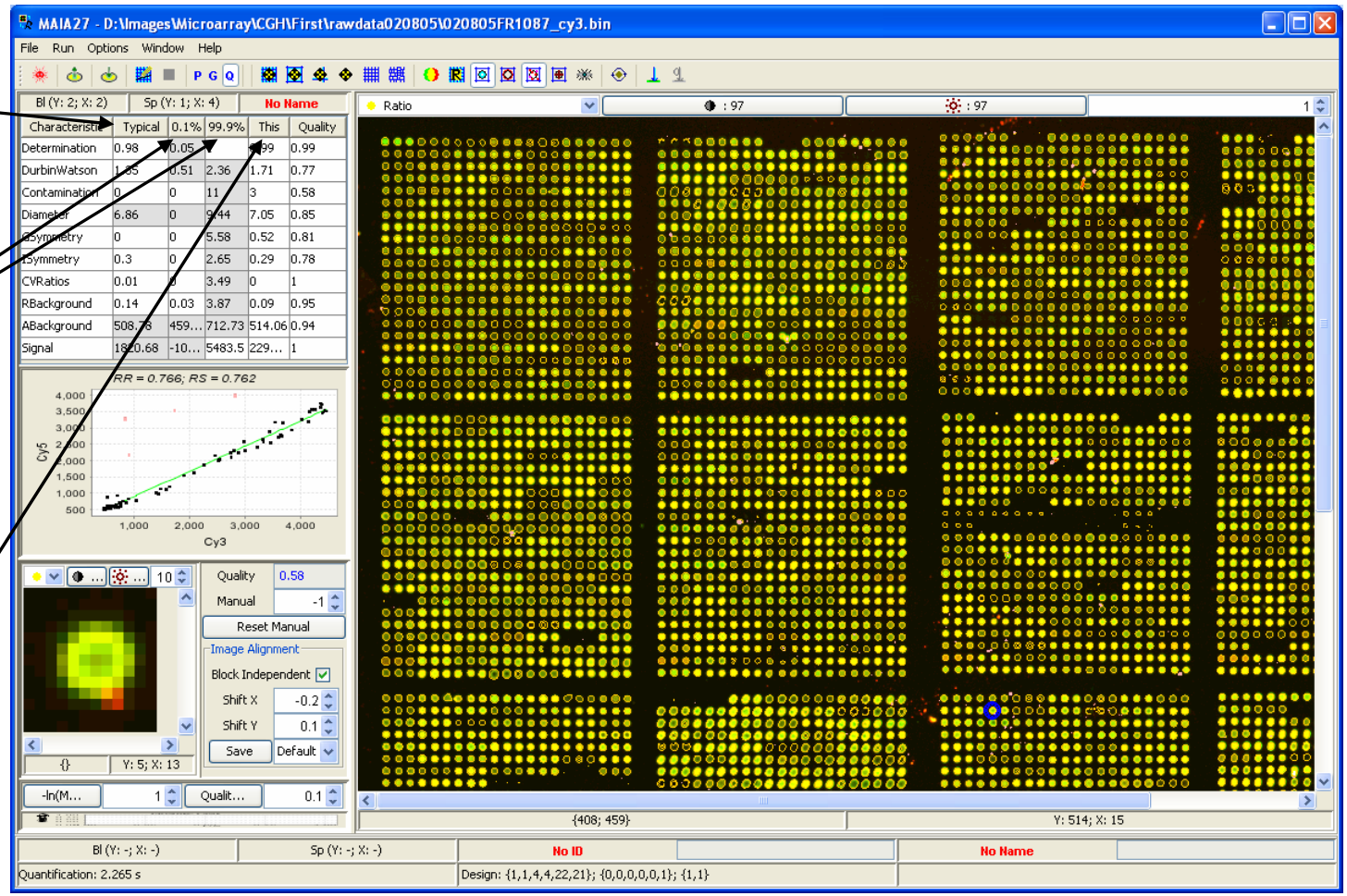
**For the purposes of further quality analysis, functions q rescale quality characteristics to fit the range between 0 ("bad" spot) and 1 ("good" spot).*

Quality Table

Typical (median) value for each characteristic over all spots on the current array.

0.1(%) and 99.9(%) percentiles for each characteristic over all spots on the current array. *The percentiles can be modified directly in the table header.*

Quality characteristics of the selected spot.



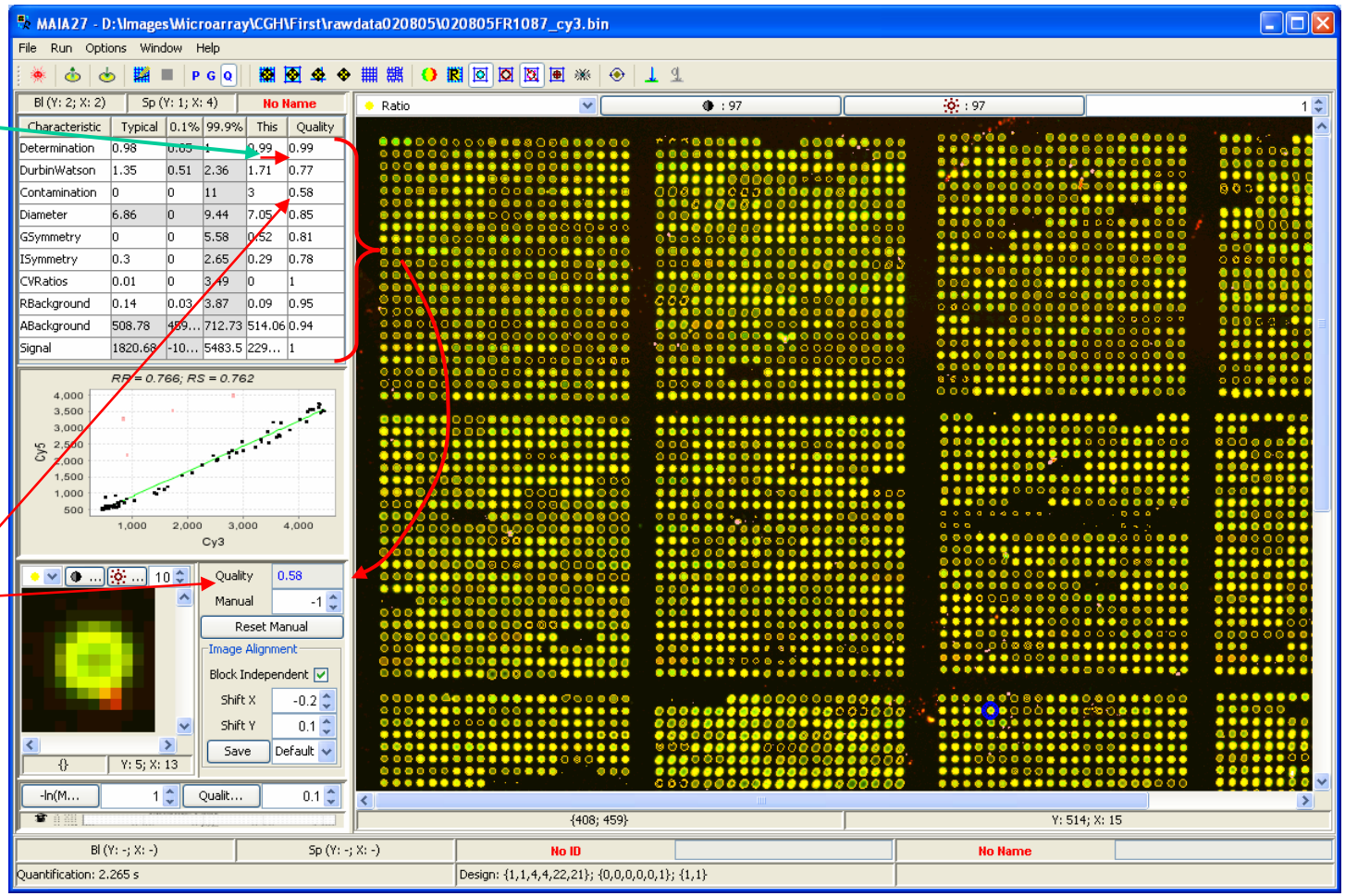
Quality Parameter

Each quality characteristic is rescaled into the corresponding marginal quality parameter ∈ [0;1].

See page Quality Characteristics.

The minimal quality value from a set of marginal quality parameters is taken as an overall quality value.

See next page for details.



*Quality Parameter in Detail**

The overall quality value is defined as:

$Q = min_i\{q_i^{w_i}\}$, (1)

where $q_i = q_i(x_i) \in [0;1]$, are the marginal scaled quality parameters defined on page Quality Characteristics for $x = \{CD, DWS, N, D, GS, IS, CVR, UB, AB, S\}$ and w_i are the weights that control the input of the correspondent quality components into the overall quality value. For the user-provided overall quality threshold $Q^{lim} \in [0;1]$, one can establish a link between the weight w_i and the critical value x_i^{lim} for each quality characteristic:

$w_i = log\{Q^{lim}\}/log\{q_i(x_i^{lim})\}$, or $x_i^{lim} = q_i^{-1}(\{Q^{lim}\}^{1/w_i})$, (2)

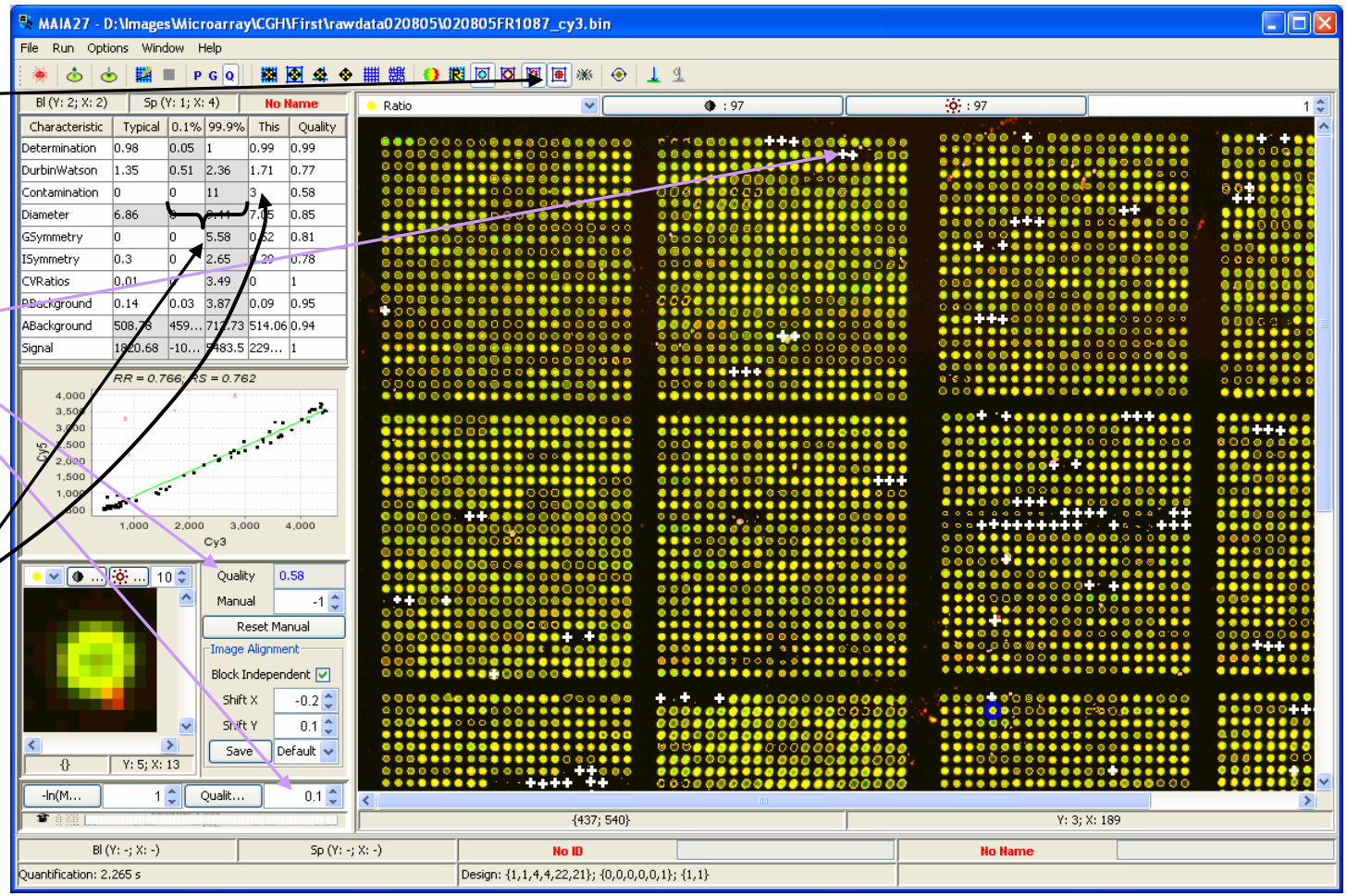
where $q_i(x_i^{lim})$ is the scaled quality parameter calculated for x_i^{lim}. The critical value x_i^{lim} sets up the limit such that if a certain characteristic i exceeds this limit, the correspondent quality parameter $q_i(x_i^{lim})$ will become lower than Q^{lim}.
The experimental quality parameters q_i are obtained from the quantification procedure, whereas the weights w_i (or the critical values x_i^{lim}) are yet unknown. The problem of spot quality analysis is therefore converted into the problem of weights (w_i) estimation, which can be solved only if additional information is provided, for example, from the replicated spots or user expertise.

*) E. Novikov and E. Barillot, An algorithm for automatic evaluation of the spot quality in two-color DNA microarray experiments. *BMC Bioinformatics*, 2005, 6:293

“Bad” Spots

Switch on the toolbar button “Show/Hide Quality Markers”.

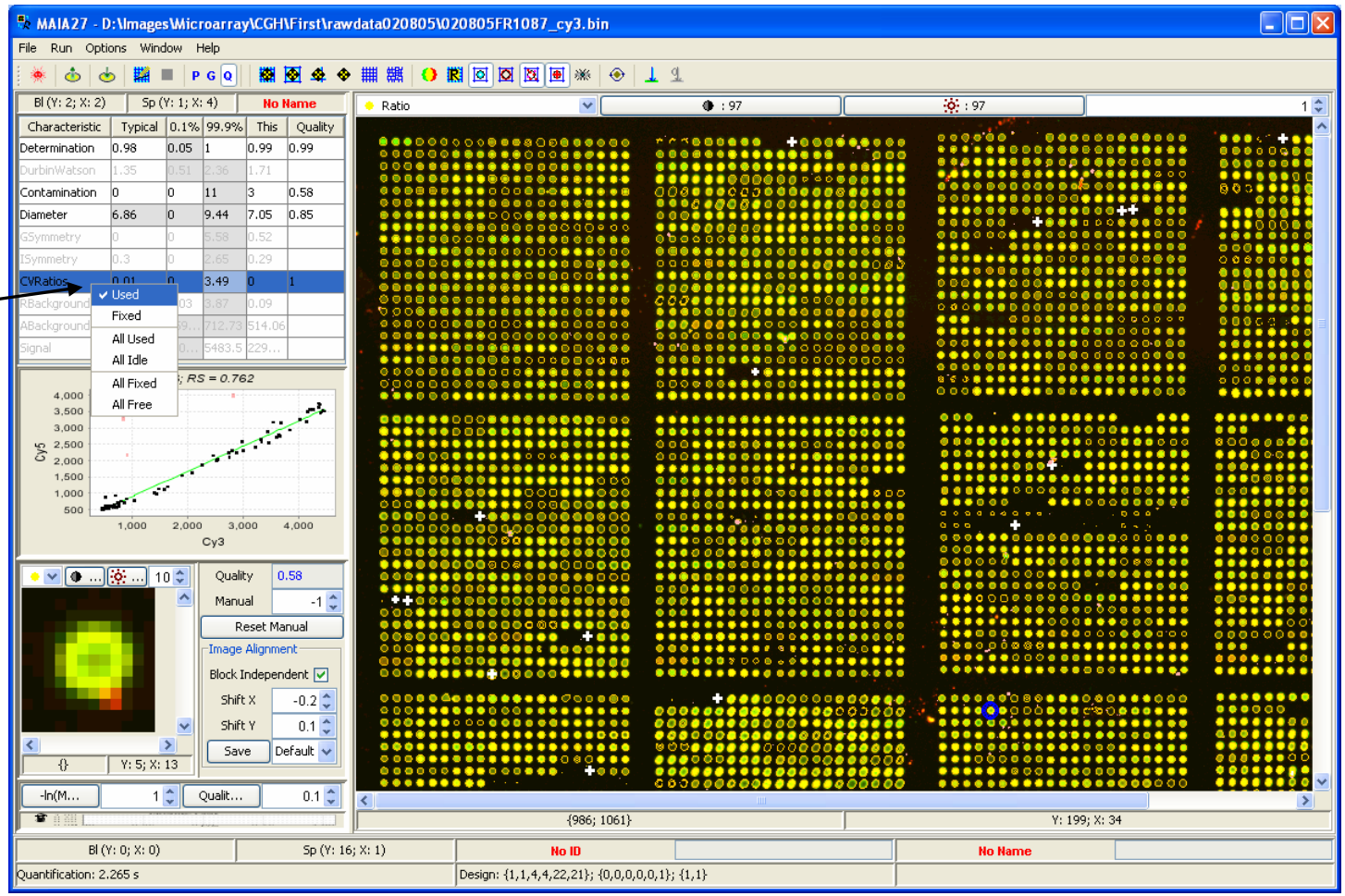
White crosses indicate “bad” spots, i.e. spots whose overall quality value is below the *Q Limit* as defined by the “Quality Limit” spinner …
or, equivalently, if one of the quality characteristics of a spot exceeds the admissible limits, defined by the corresponding percentiles.



Used Quality Characteristics

Using the right-button popup menu in the quality table select a set of quality characteristics, which can be relevant for this image.

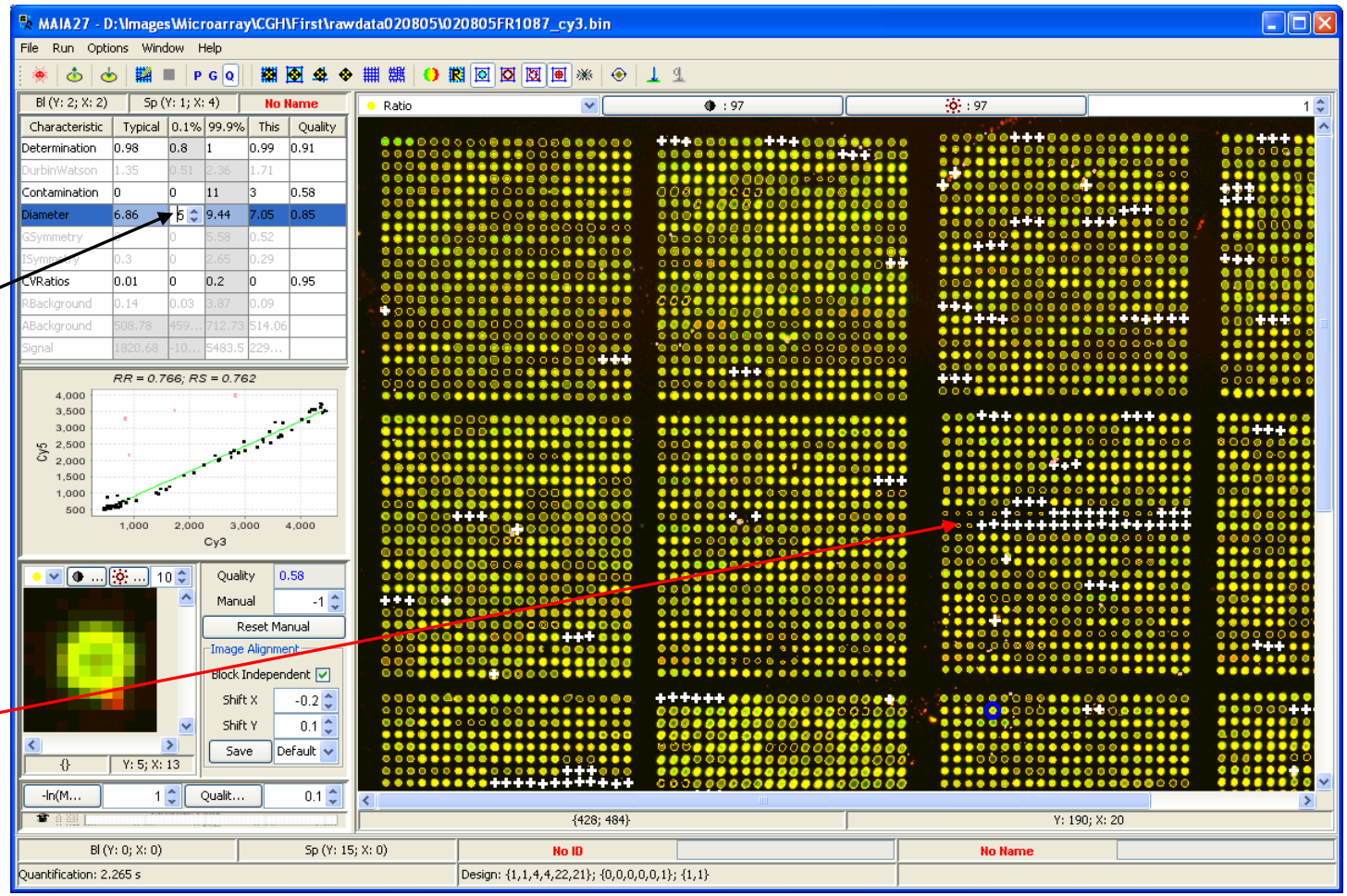
Idle characteristics are shown in gray.



Manual Limits Adjustment

For each used quality characteristic the limits can be adjusted.
The gray fields in the quality table are user-modifiable. Certain characteristics allow for changing both limits (DWS,D,AB) and/or typical value (D,AB,S).

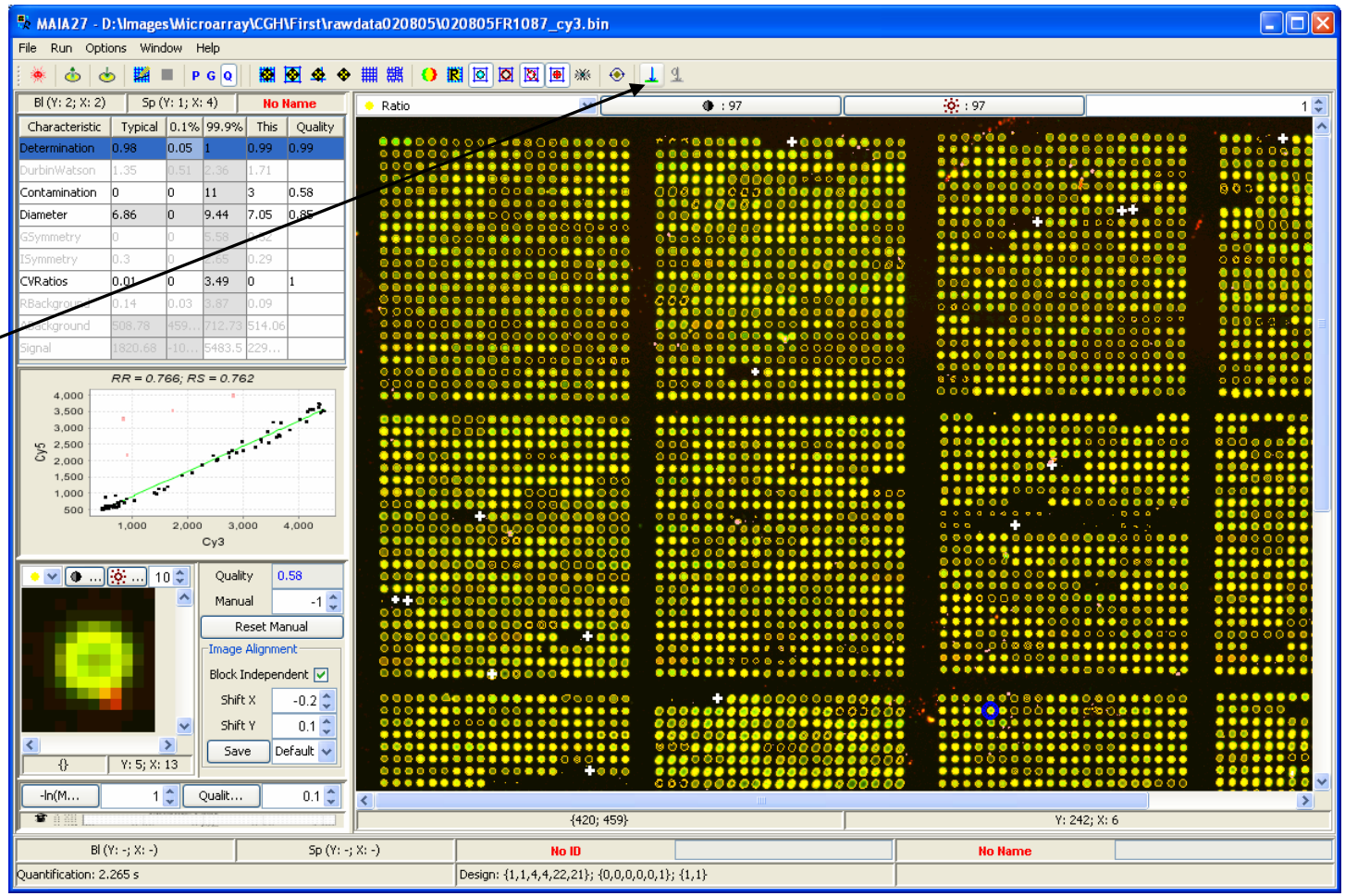
Limit adjustment should be continued until all spots, visually classified as "bad" spots, are flagged out.



Default Limits

The default limits can be restored using the “Init Limits” button from the Toolbar or the Menu Item “Run|Init Limits” (Ctrl+F9).

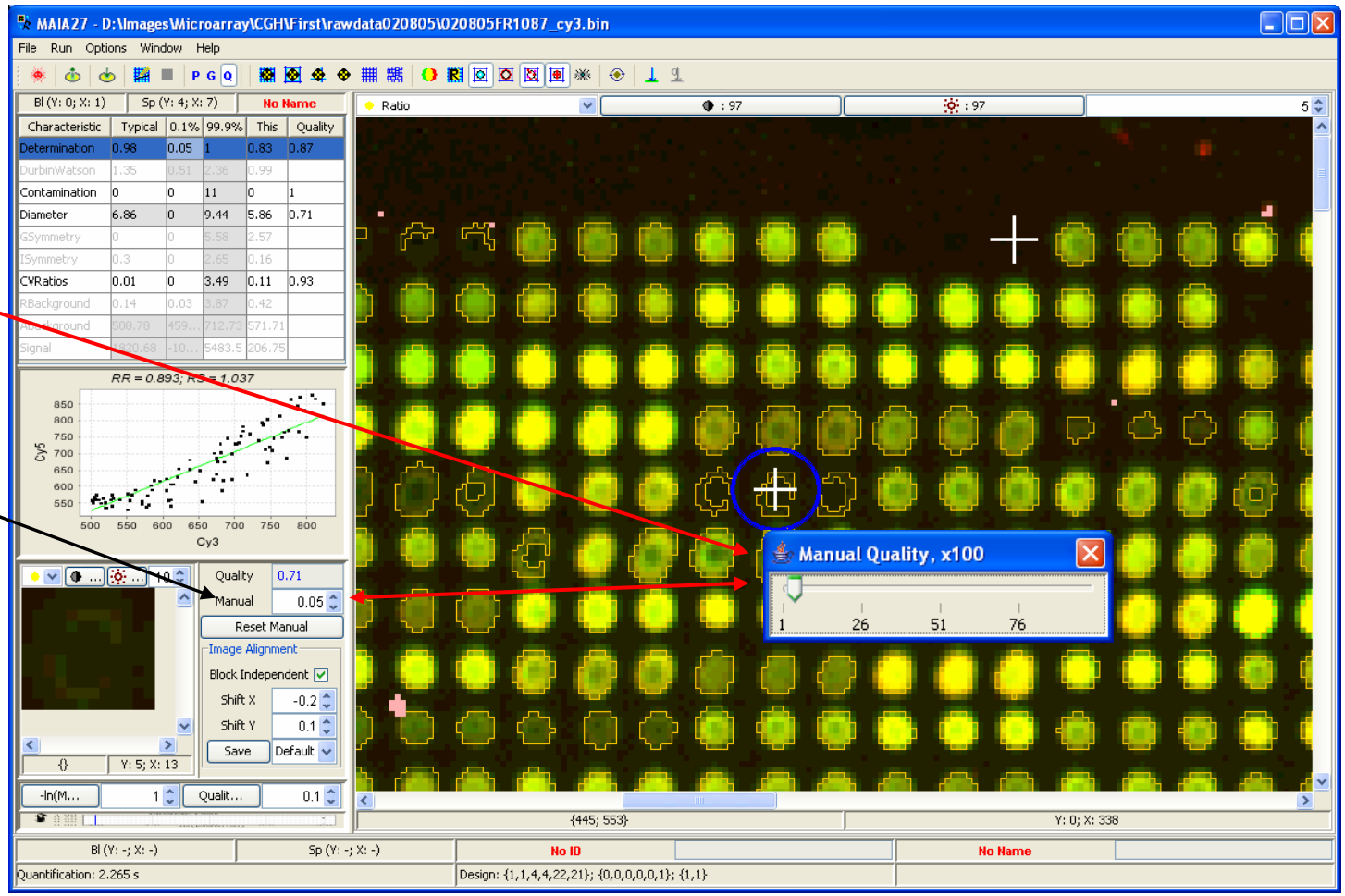
The default limits for each quality characteristic are the corresponding percentiles over all spots on the array. The percentage is defined in the table header.



Manual Qualification of the Selected Spot

Using the mouse right button or the spinner "Manual", any spot can be assigned a certain value from the interval [0;1], which can further be used as an additional parameter of quality.

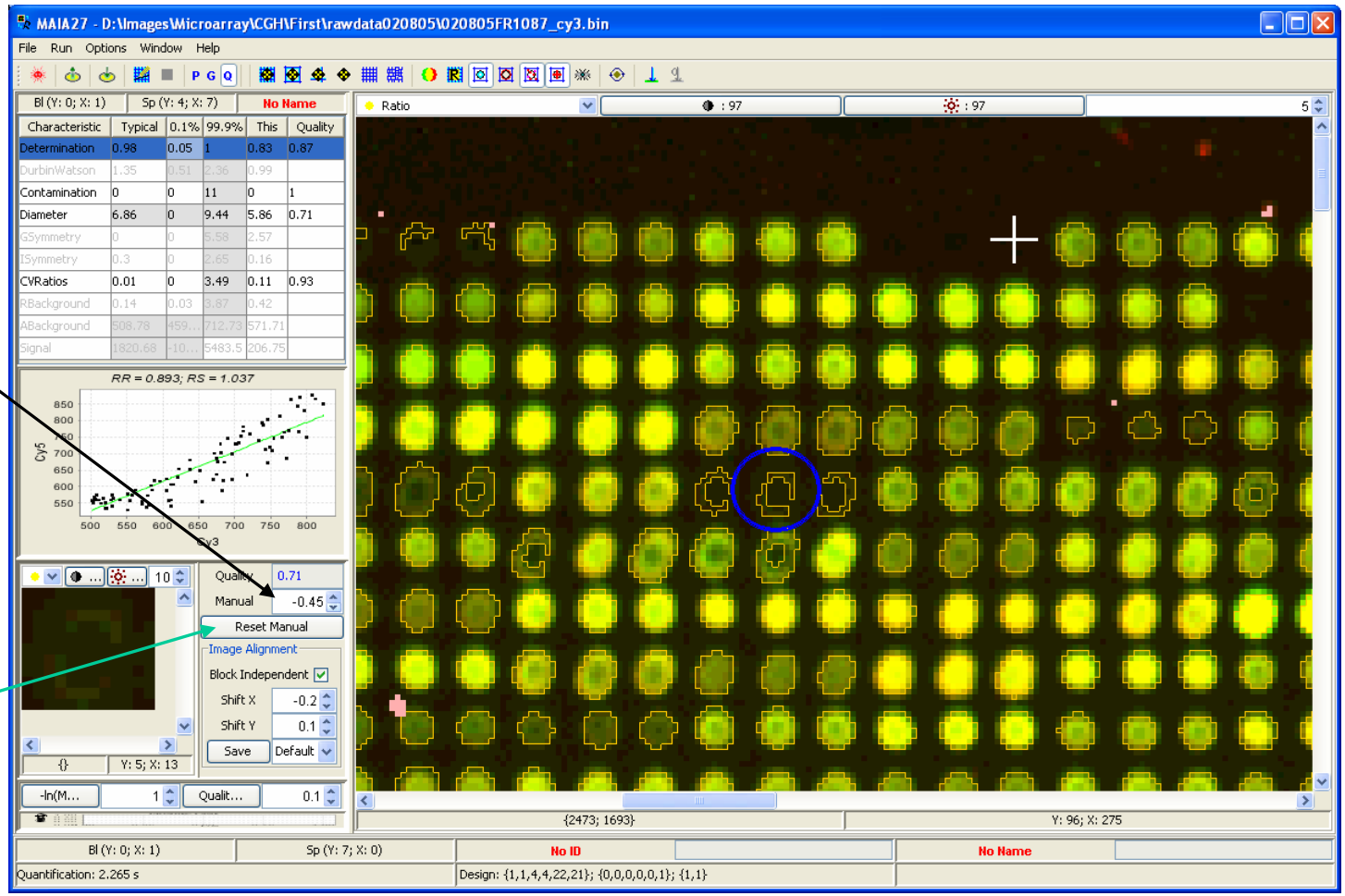
If the user-defined quality value is below the Quality Limit, the corresponding spot will be crossed.



Manual Spot Characterization

Negative values available in the “Manual” spinner are not considered as quality values and can be used for additional spot characterization.

The “Reset Manual” button sets the manual parameter for all spots on the array in “-1”.

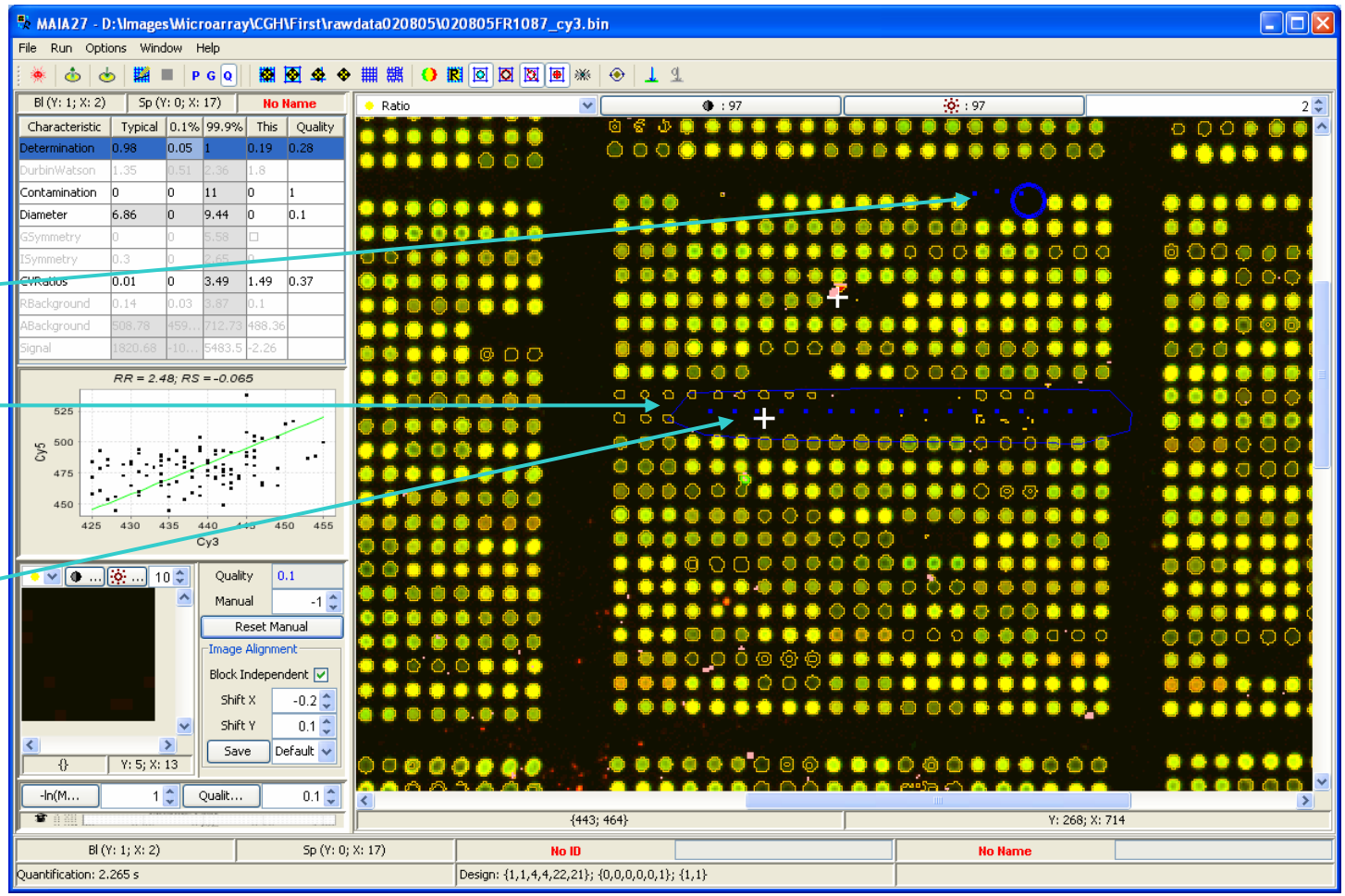


Groups of Spots for Manual Qualification

Groups of spots can be selected for manual qualification.

Spots can be added into the group one by one (Shift+Left Click), or several at once: Ctrl+Left Clicks followed by Ctrl+Right Click create the contour of the selected spots.

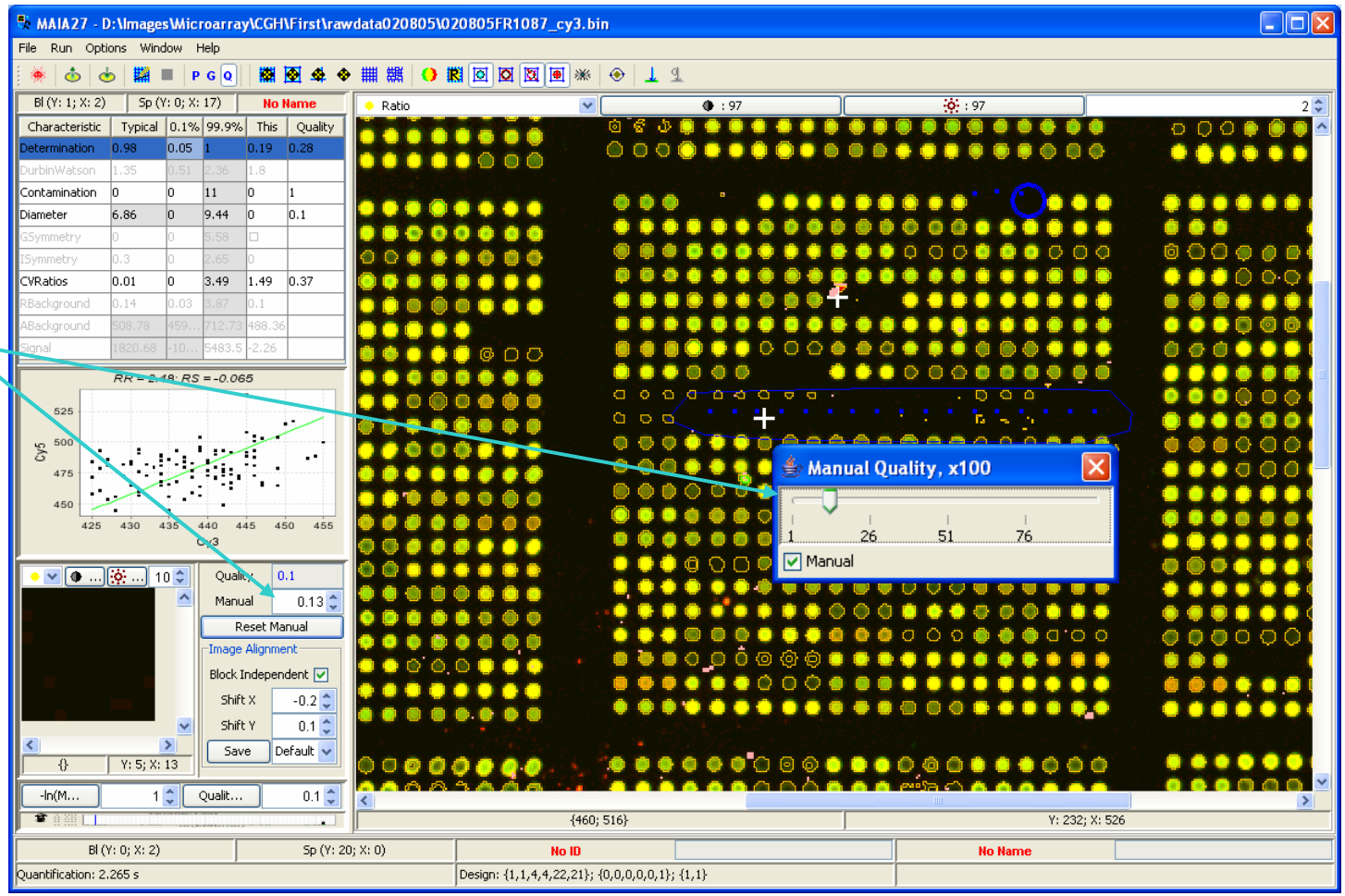
Selected spots are marked by the dots in the left upper corner of the spot area.



Manual Qualification of the Selected Group

All spots from the selection can be assigned the same quality value.

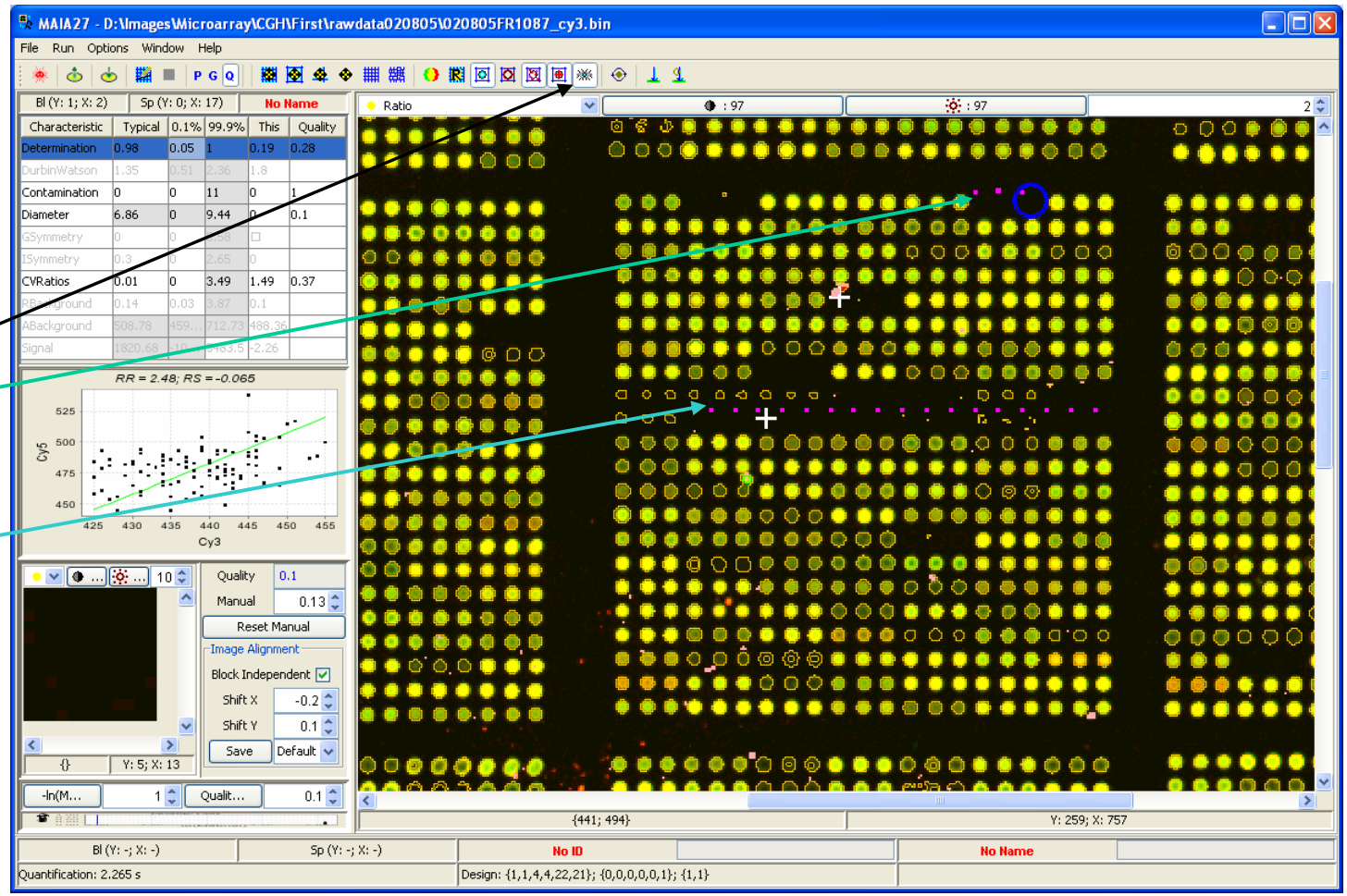
Ctrl+Left Click within a contour effaces this contour. Shift+Left Click inverts the selection of the spot. Double click on the image effaces all contours.



Manual Selection Markers

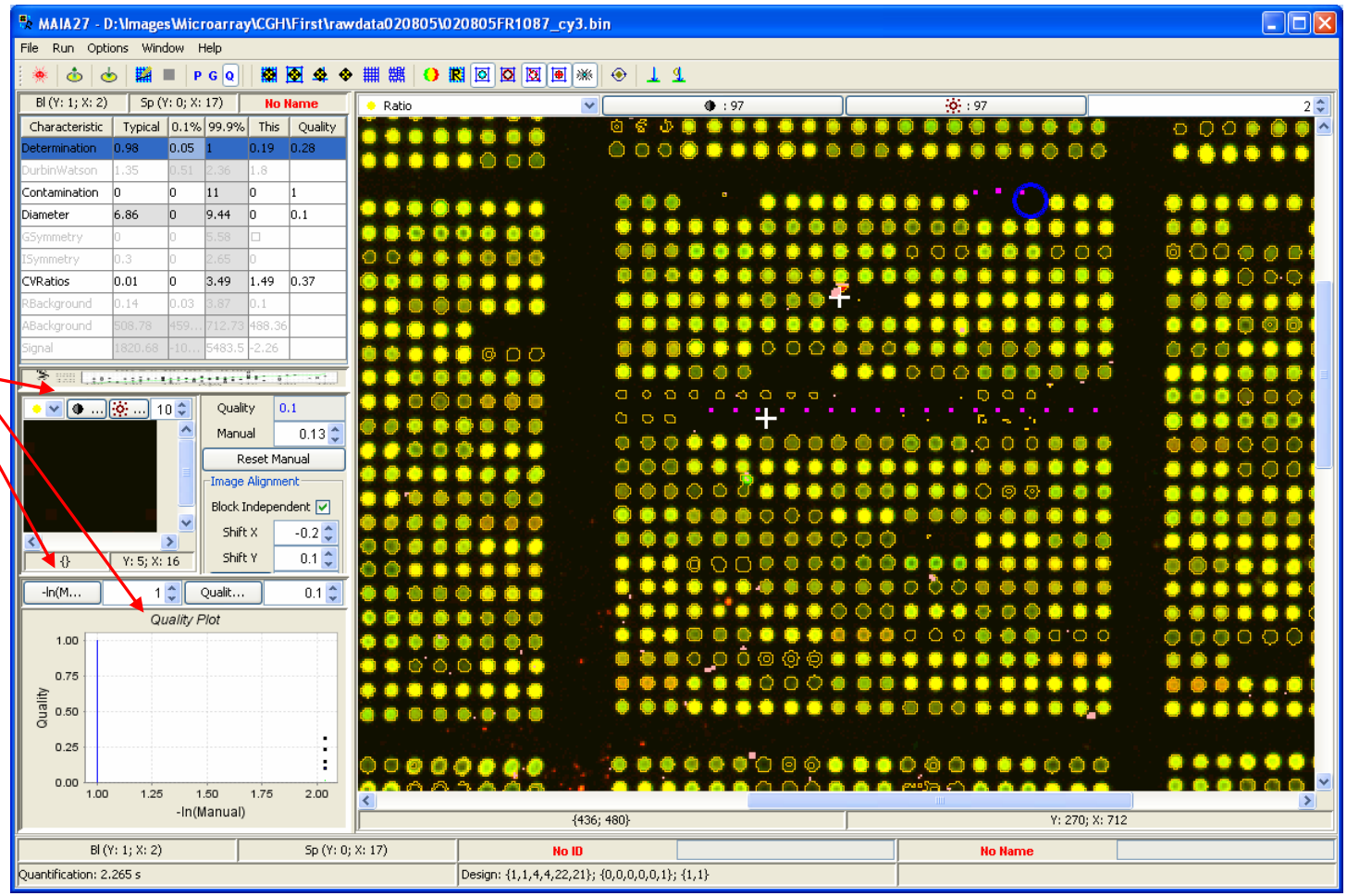
Using the Toolbar button "Show/Hide Manual Selection" user qualified spots can be visualized.

These spots are signed by a dot in the left upper corner of the spot area.



Quality Plot

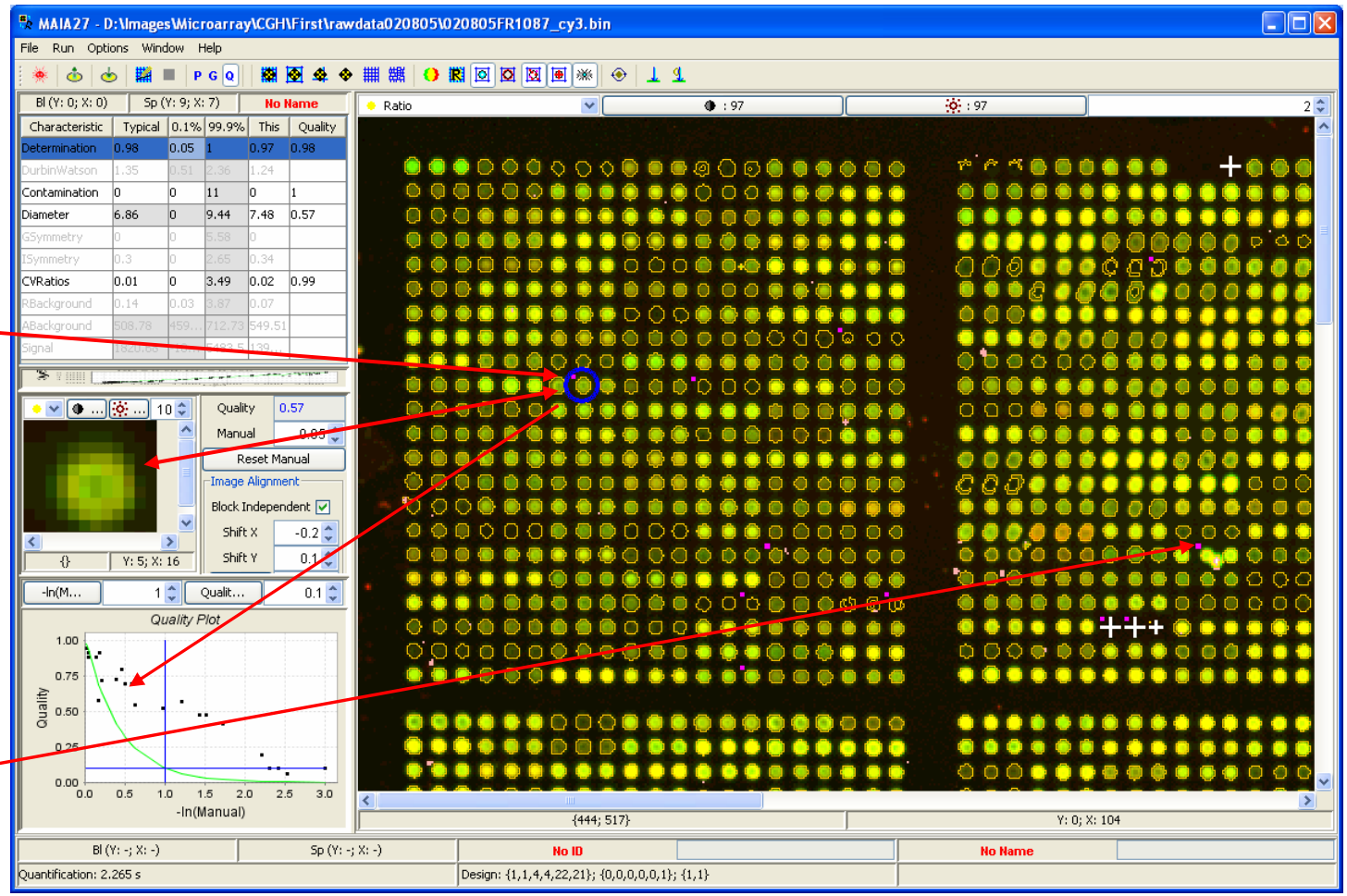
Slide up the bars separating the panels and open up the quality plot.



Selected Spots for Quality Analysis

User assigns quality values ∈[0;1] to some representative spots. These values (z) are converted as –ln(z) to create the *x* axis of the quality plot.
y-axis: the overall quality parameter.

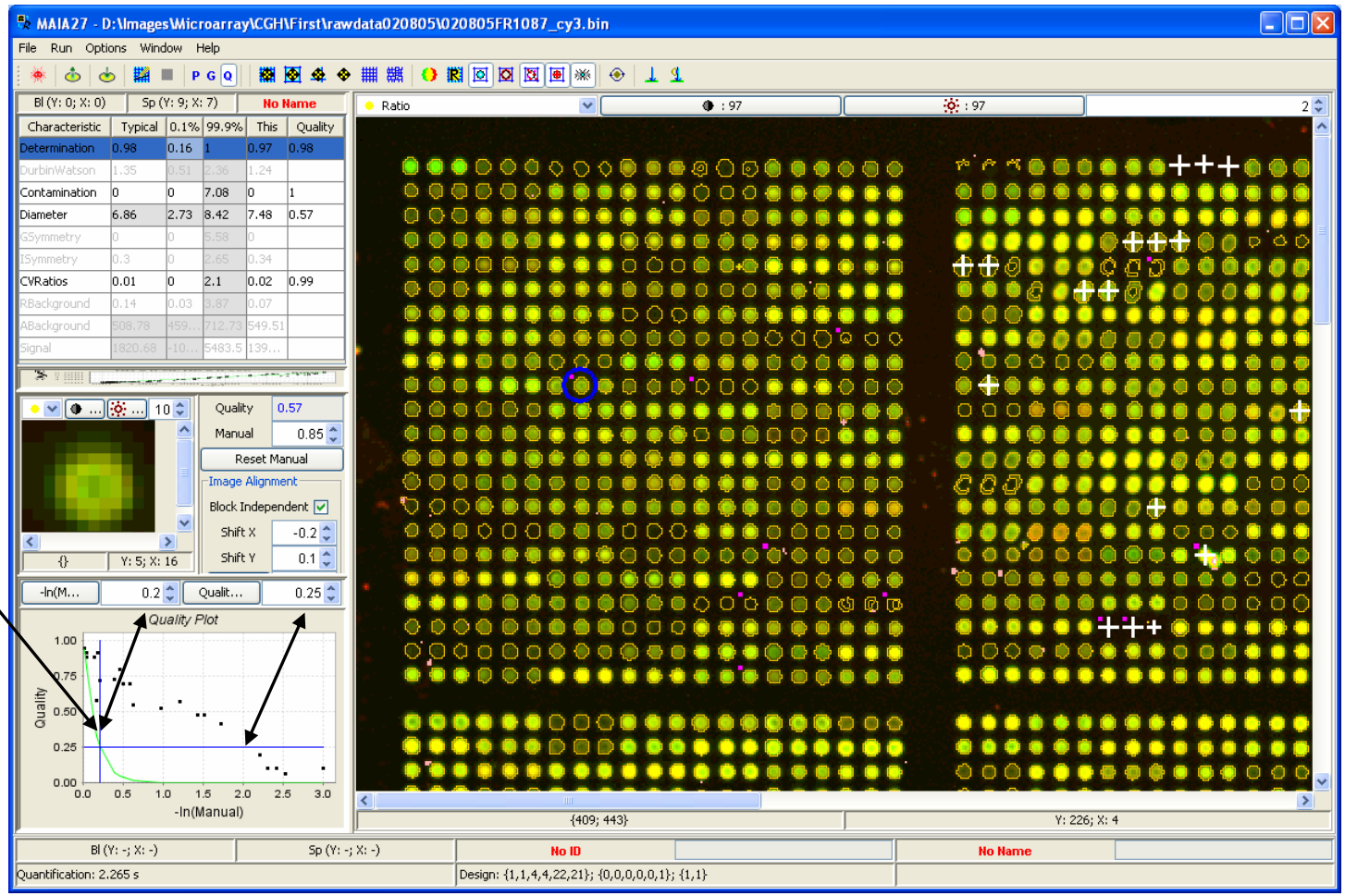
Visually qualified spots are marked by a dot in the upper left corner of the spot area.



Quality Curve

Use the mouse pointer or the spinners “Ratio CV Limit” and “Quality Limit” to define the quality curve (green line).

Quality curve defines how fast the overall quality must decrease with the decrease of the manually assigned quality. The user-defined quality curve is an exponent with the predefined decay constant.

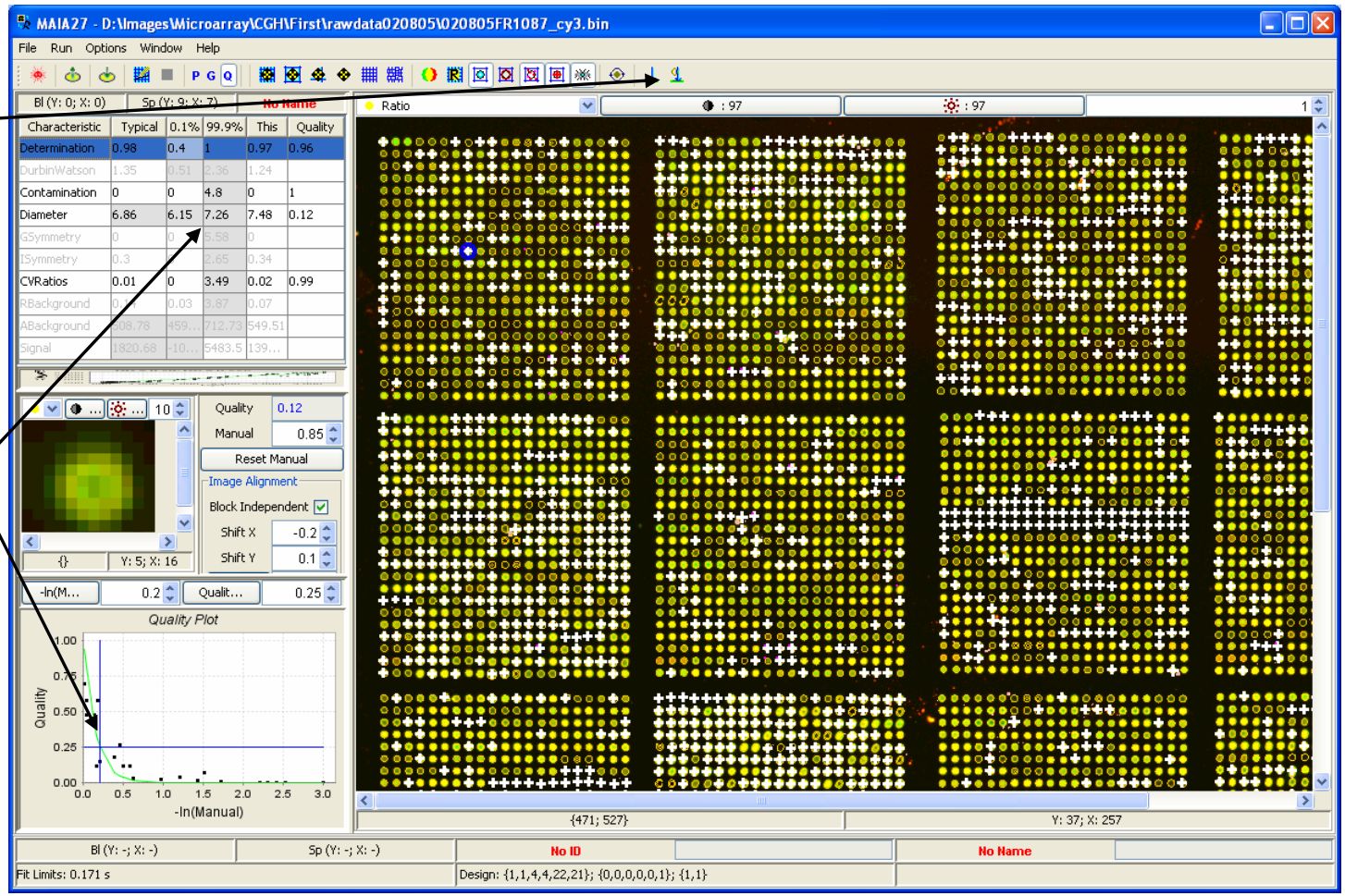


Fit the Limits

Fit the quality limits by the “Fit Limits” button from the Toolbar or by the Menu Item “Run|Fit Limits” (Ctrl+F10).

Quality fit estimates the limits of the quality characteristics such that the spot overall quality is aligned along the user-defined quality curve.

Before fitting it is advisable to restore the default limits (the “Init Limits” button from the Toolbar or the Menu Item “Run|Init Limits” (Ctrl+F9)).



Fix the Limits

Certain limits can be fixed, so that they are not changed by the fit.

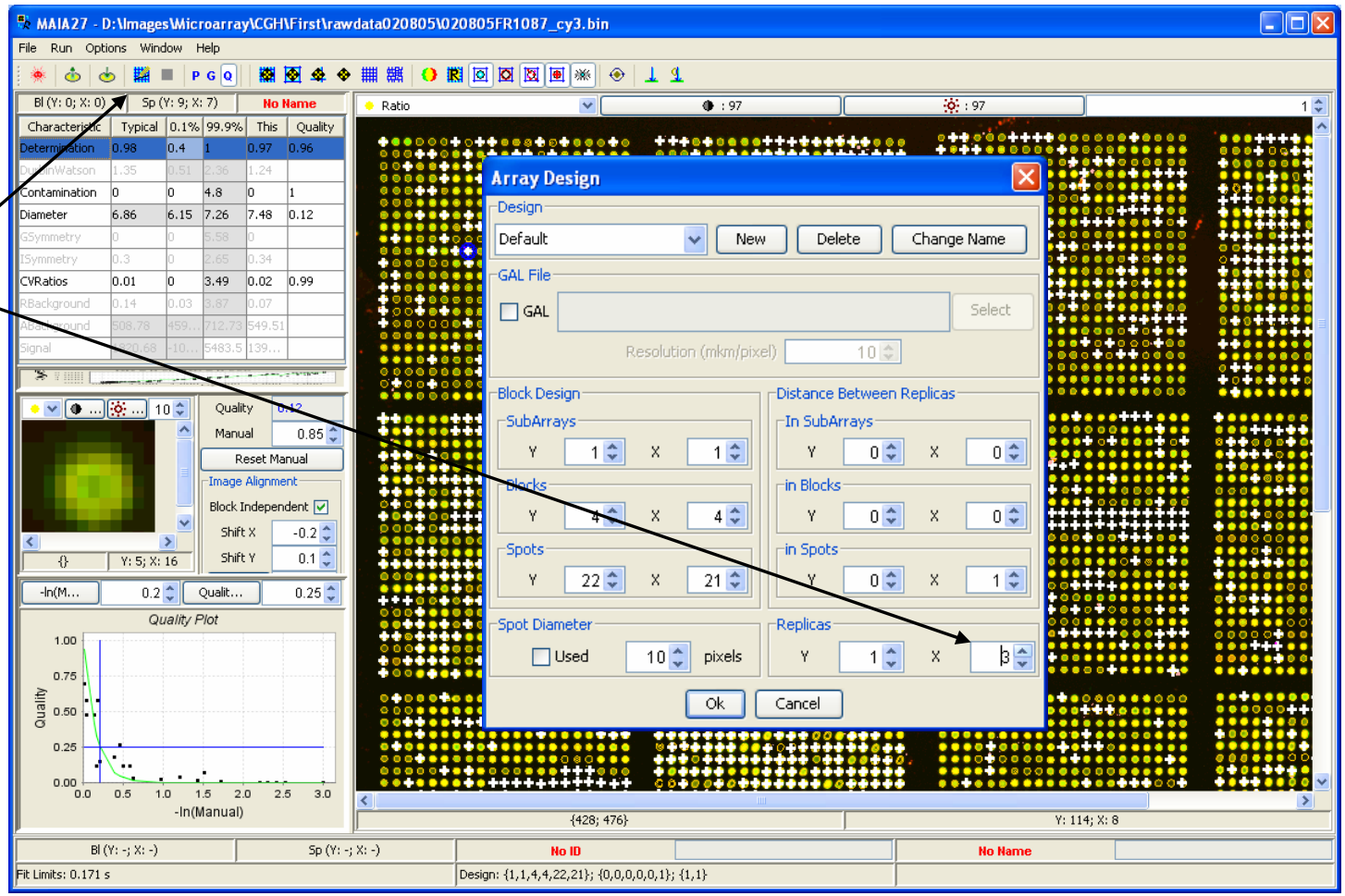
The fixed quality characteristics are shown in *italics*.

Quality Analysis Using Replicated Spots

On this image, three replicated spots are placed as neighbors in a row.

This is defined by the Array Design dialog (click the "Array Design" button from the Toolbar or select the Menu Item "Options|Array Design" (Alt+A)).

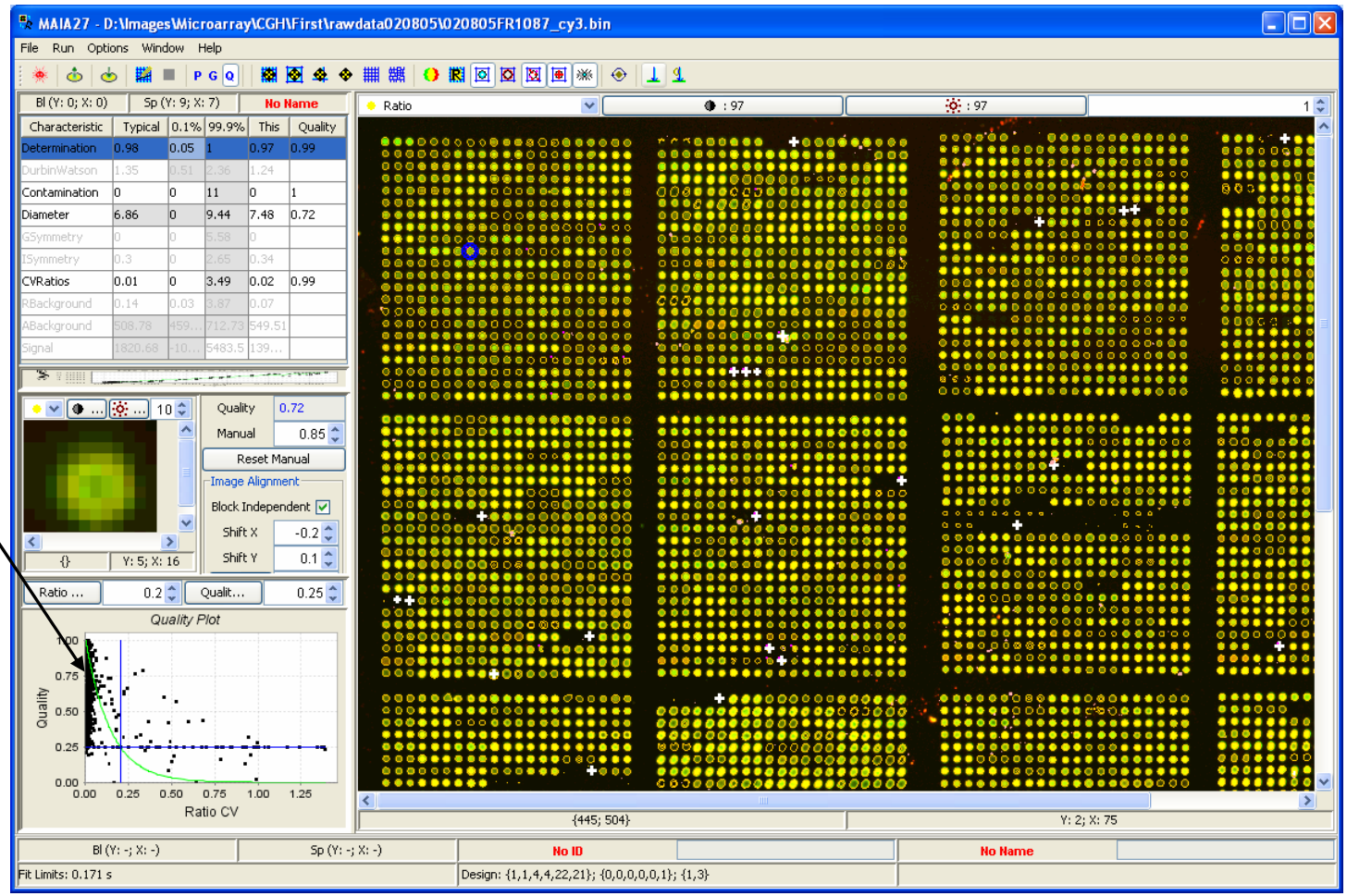
After changing the design, the "Init Limits" button from the Toolbar or the Menu Item "Run|Init Limits" (Ctrl+F9) can be used to restore the default limits.



Quality Plot with Replicated Spots

Each dot represents a replicate with the overall quality value at *y*-axis and ratio variation coefficient (CV) of the replicates at *x*-axis.

See next page for details.



*Spot Quality Fit**

The weights w_i for the overall quality parameter Q (see page Quality Parameter) can be estimated using replicated spots on the same array or over a set of replicated arrays. The high-quality spots belonging to the same replicate are expected to demonstrate very close to each other ratio value. Relatively big difference between the observed ratios in the same replicate will signal that some of the spots from this replicate are irregular. To formalize this approach, we first define the quality value for the replicate:

$$Q_k = min_{j=1...n}\{Q_{kj}\}, \quad (1)$$

where k enumerates the replicates, n is the number of spots in a replicate, and Q_{kj} is a spot quality value given by Eq. (1, page Quality Parameter). Substituting Eq. (1, page Quality Parameter) into (3) yields

$$Q_k = min_{j=1...n}\{min_{i=1,...,10}\{q_{kji}^{wi}\}\} \quad (4)$$

where q_{kji} is the i-th scaled quality parameter of the j-th replicated spot in the k-th replicate.
The weights w_i can be determined as the parameters ensuring the best fit of the obtained experimental quality values (Q_k versus V_k) to the user-defined (ideal) quality curve $f(V_k)$, where V_k is the ratio variation coefficient in the k-th replicate. $f(V_k)$ defines how fast the overall quality of the replicates must decrease with the increase of the ratio variation. The shape of the user-defined quality curve $f(V_k)$ should demonstrate monotonic decay. We always use the exponential function $f(V_k) = exp\{-V_k/V\}$, and in this case only the expected (typical) ratio variation coefficient V must be predefined.

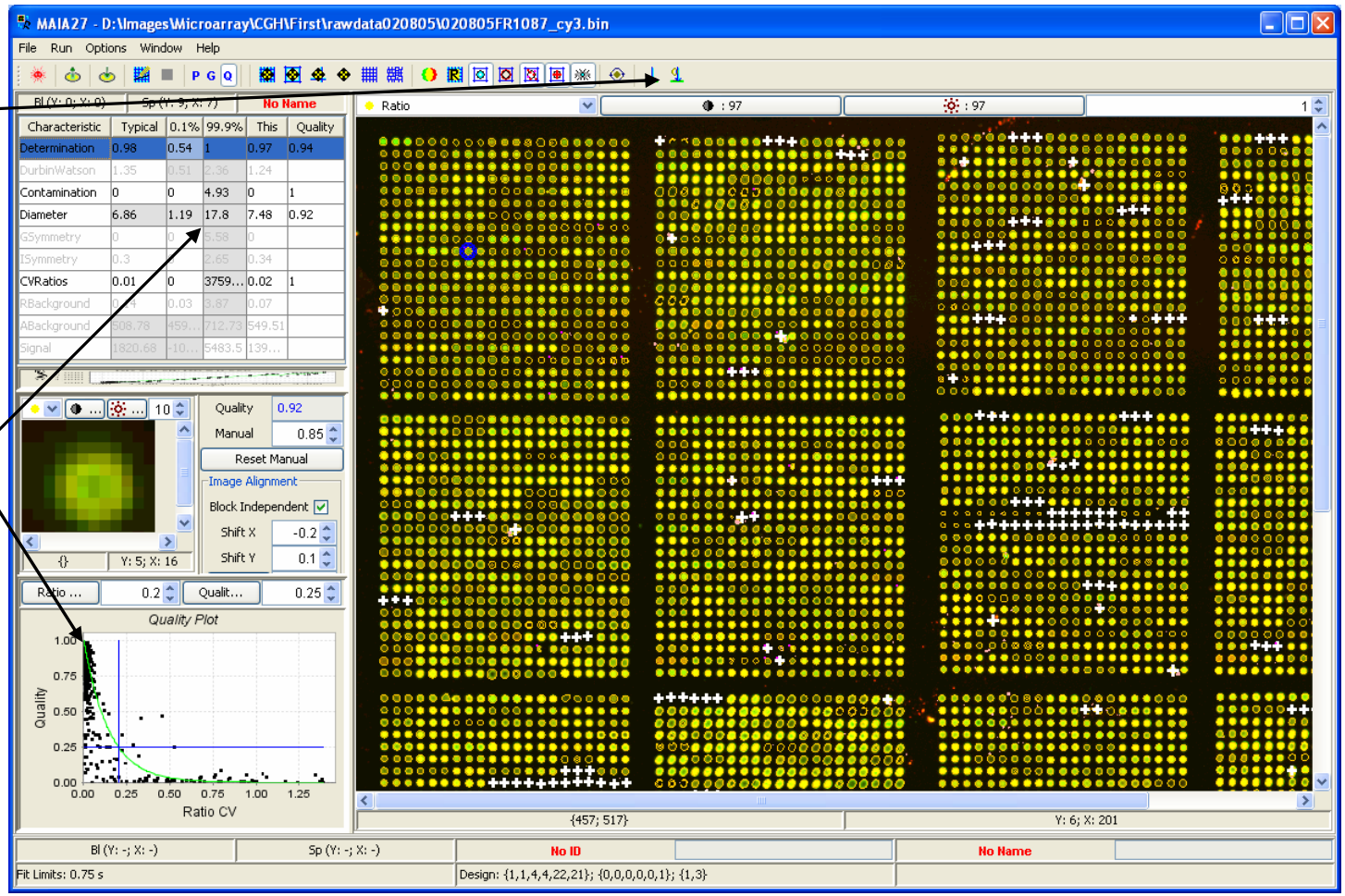
*) E. Novikov and E. Barillot, An algorithm for automatic evaluation of the spot quality in two-color DNA microarray experiments. *BMC Bioinformatics*, 2005, 6:293

Fit the Limits

The quality limits are fitted using the “Fit Limits” button from the Toolbar or the Menu Item “Run|Fit Limits” (Ctrl+F10).

Quality fit estimates the limits of the quality characteristics such that the spot overall quality is aligned along the user-defined quality curve.

Before fitting it is advisable to restore the default limits (the “Init Limits” button from the Toolbar or the Menu Item “Run|Init Limits” (Ctrl+F9)).

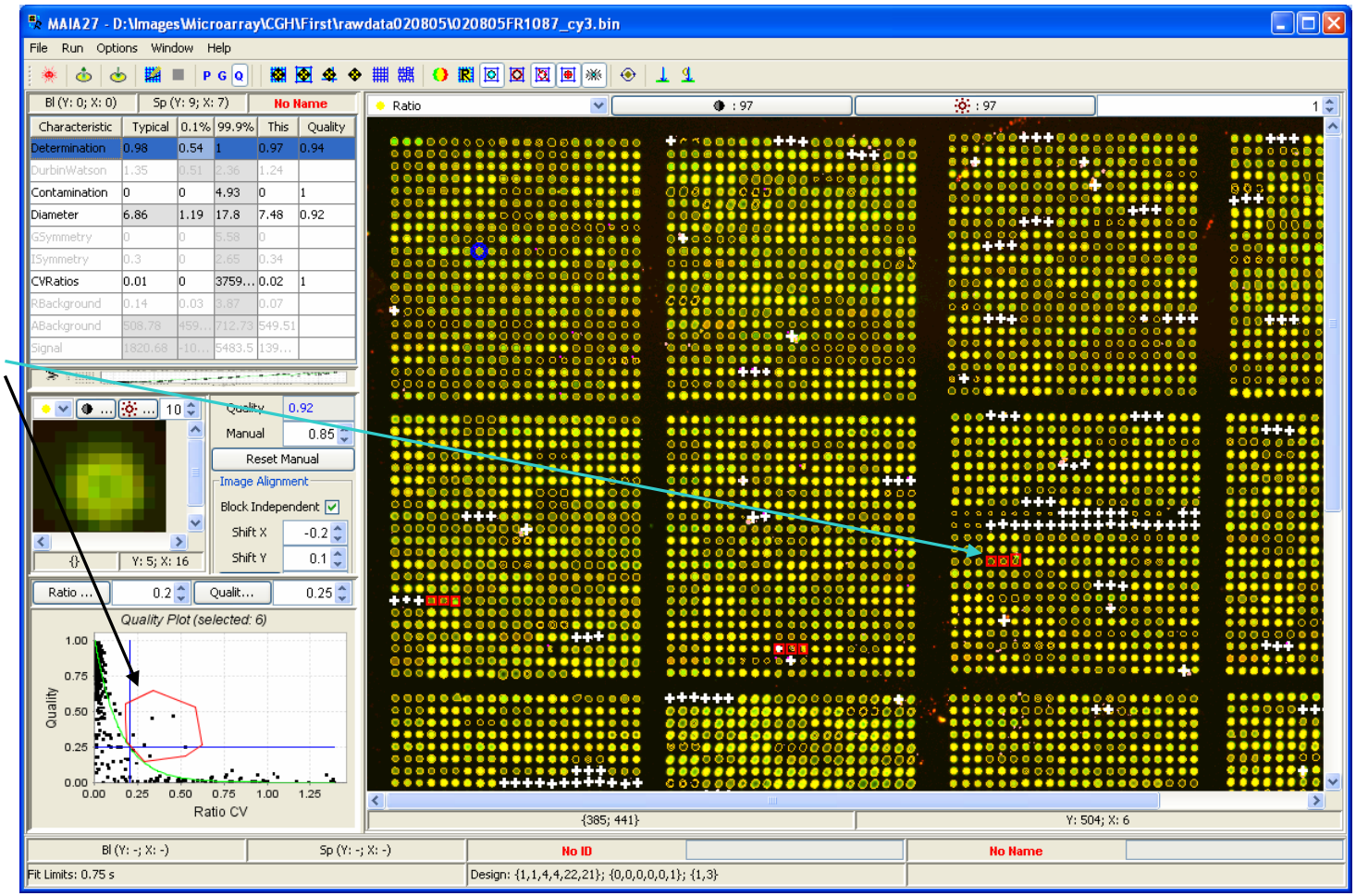


"Bad" Replicates

Ctrl+Left Clicks followed by Ctrl+Right Click create the contour on the Quality plot. The replicates that are within the contour are highlighted on the image.

Several contours (in different parts of the graph) can be created.

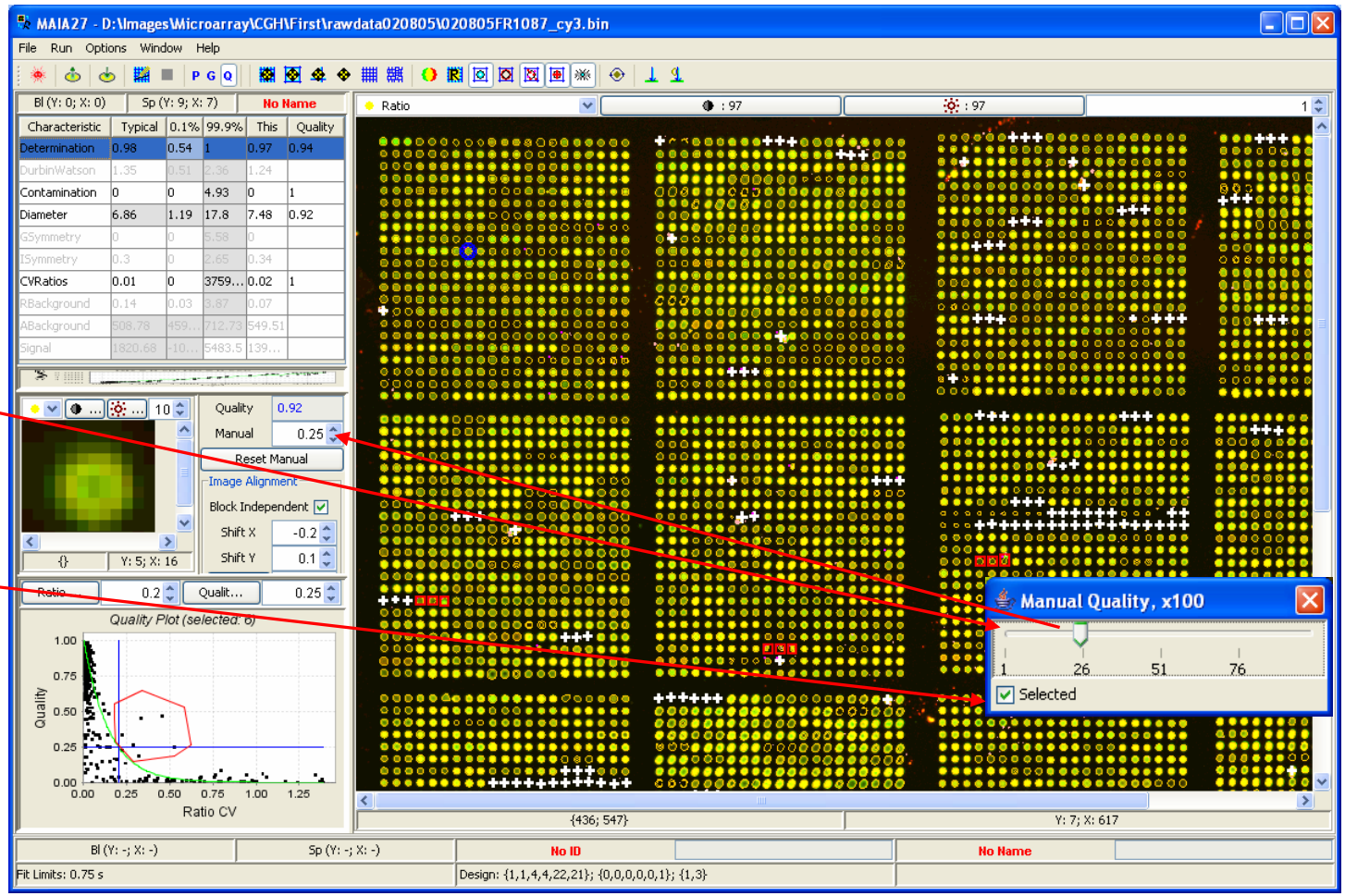
To efface contours, click of the graph.



Manual Qualification of the Selected Spots

The same manual quality value can be assigned to all selected spots using either the right button popup slider or the spinner “Manual”.

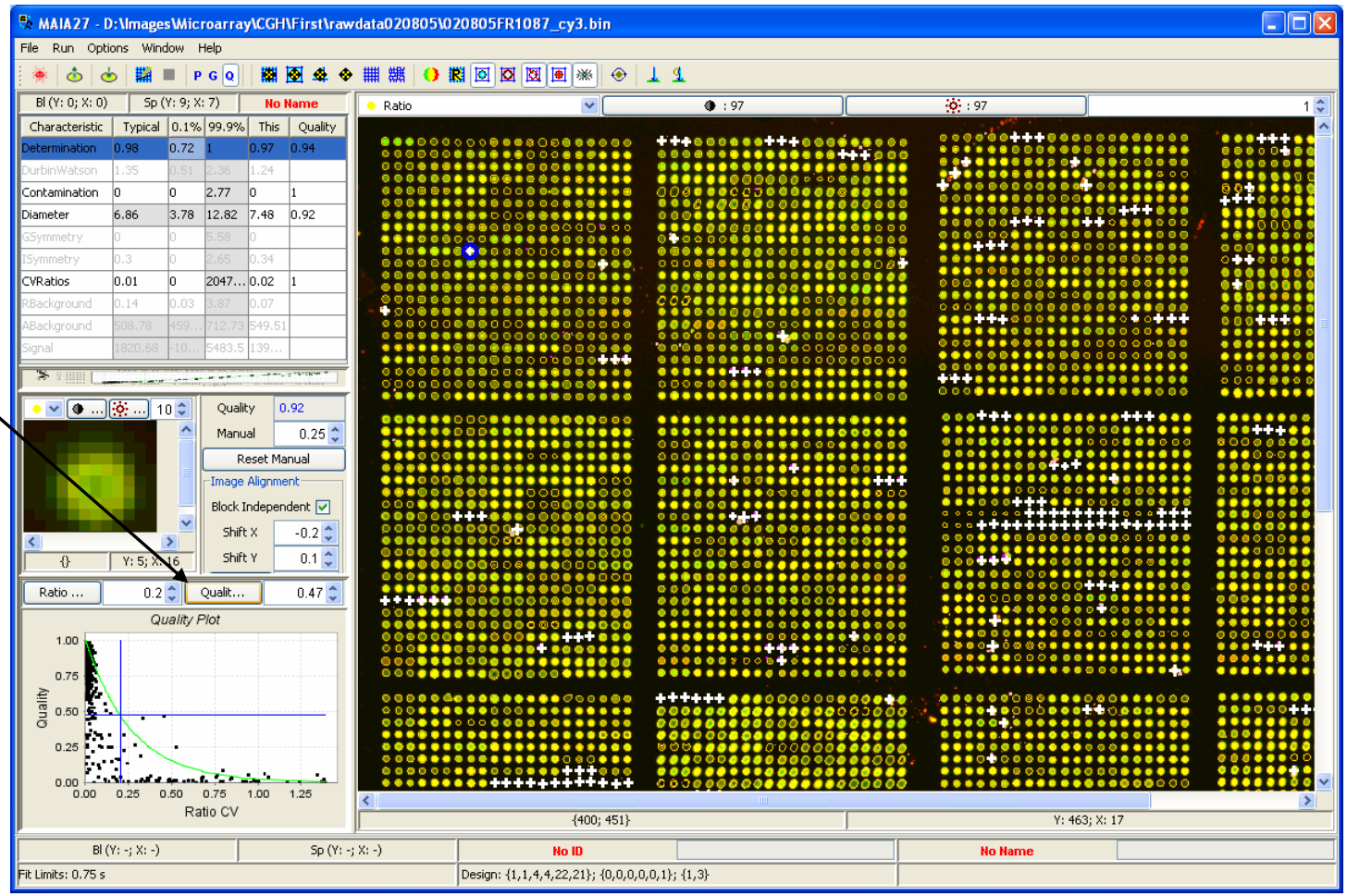
The checkbox “Selected” should be on.



Optimize the Quality Limit

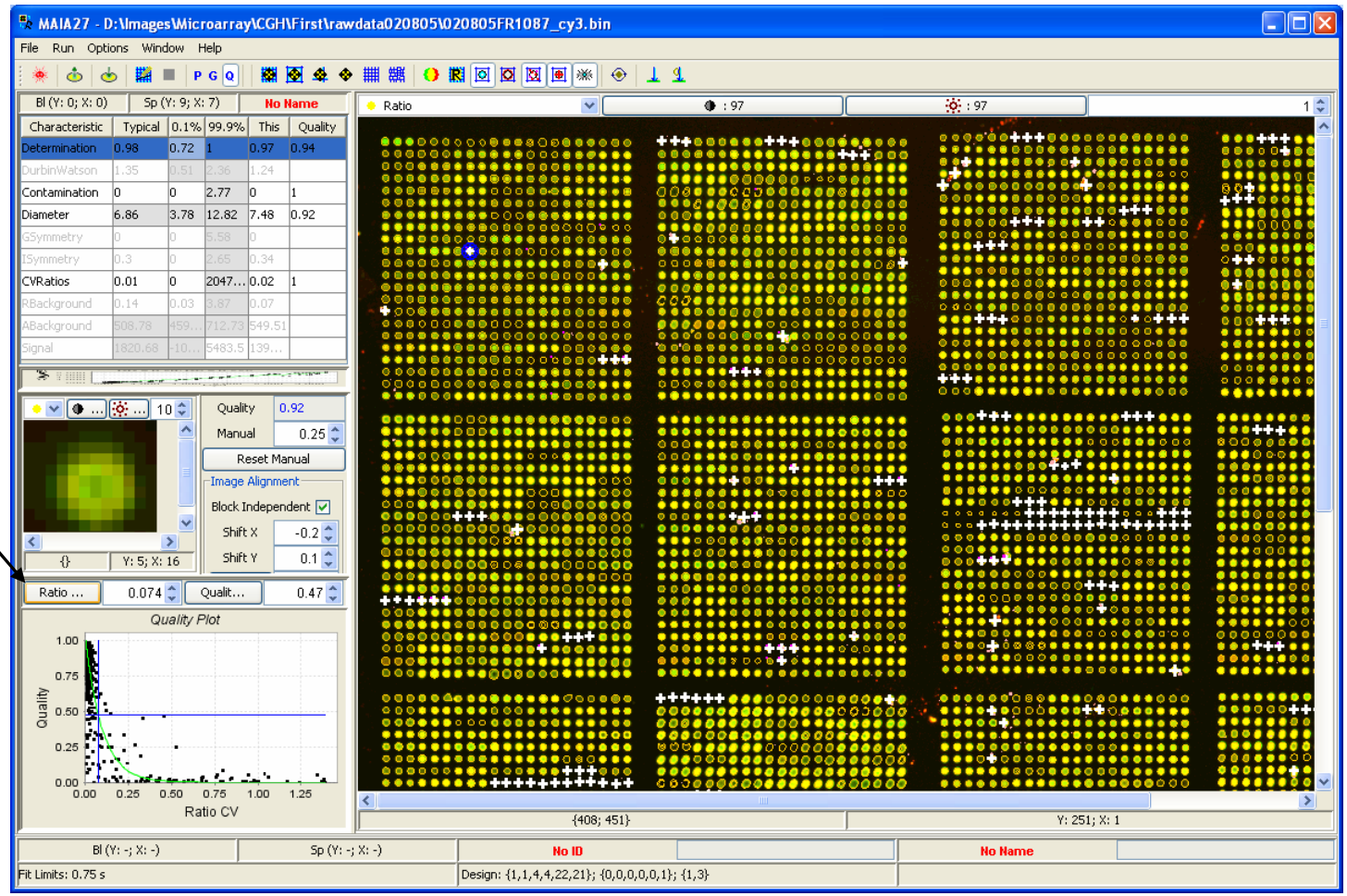
To optimize the position of the Quality Limit press the button “Quality Limit”.

A special procedure searches for the limit value such that the number of replicates in the “Bottom-Left + Top-Right” quadrants of the quality plot should be as small as possible, whereas in the “Bottom-Right+Top-Left” quadrants – as big as possible.



Optimize the Ratio CV Limit

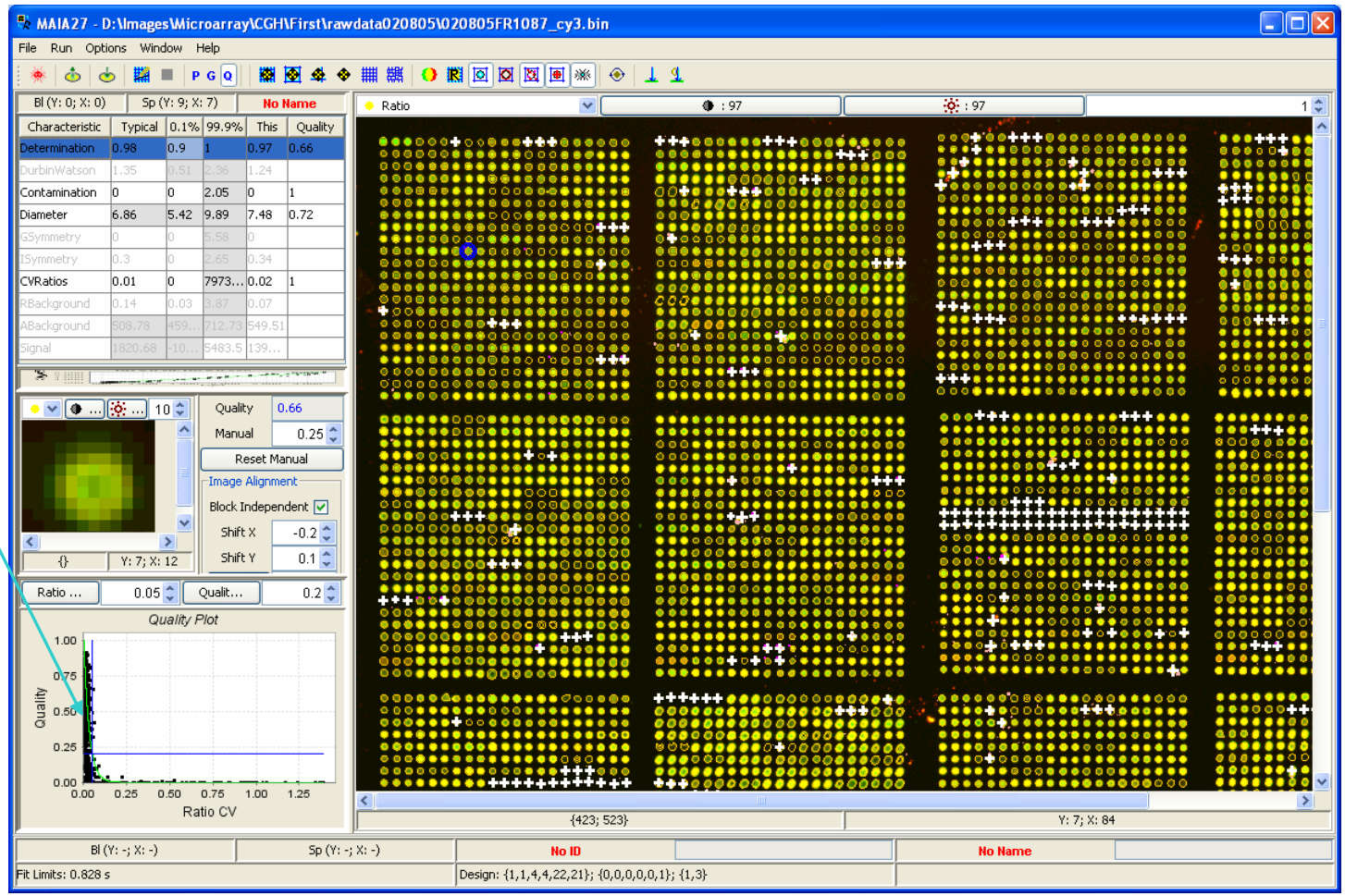
Using the button “Ratio CV Limit” the correspondent limit is set into a value ensuring the best exponential approximation for the “cloud” of replicates (black dots).



New Quality Plot

A somewhat more stringent quality curve is applied.

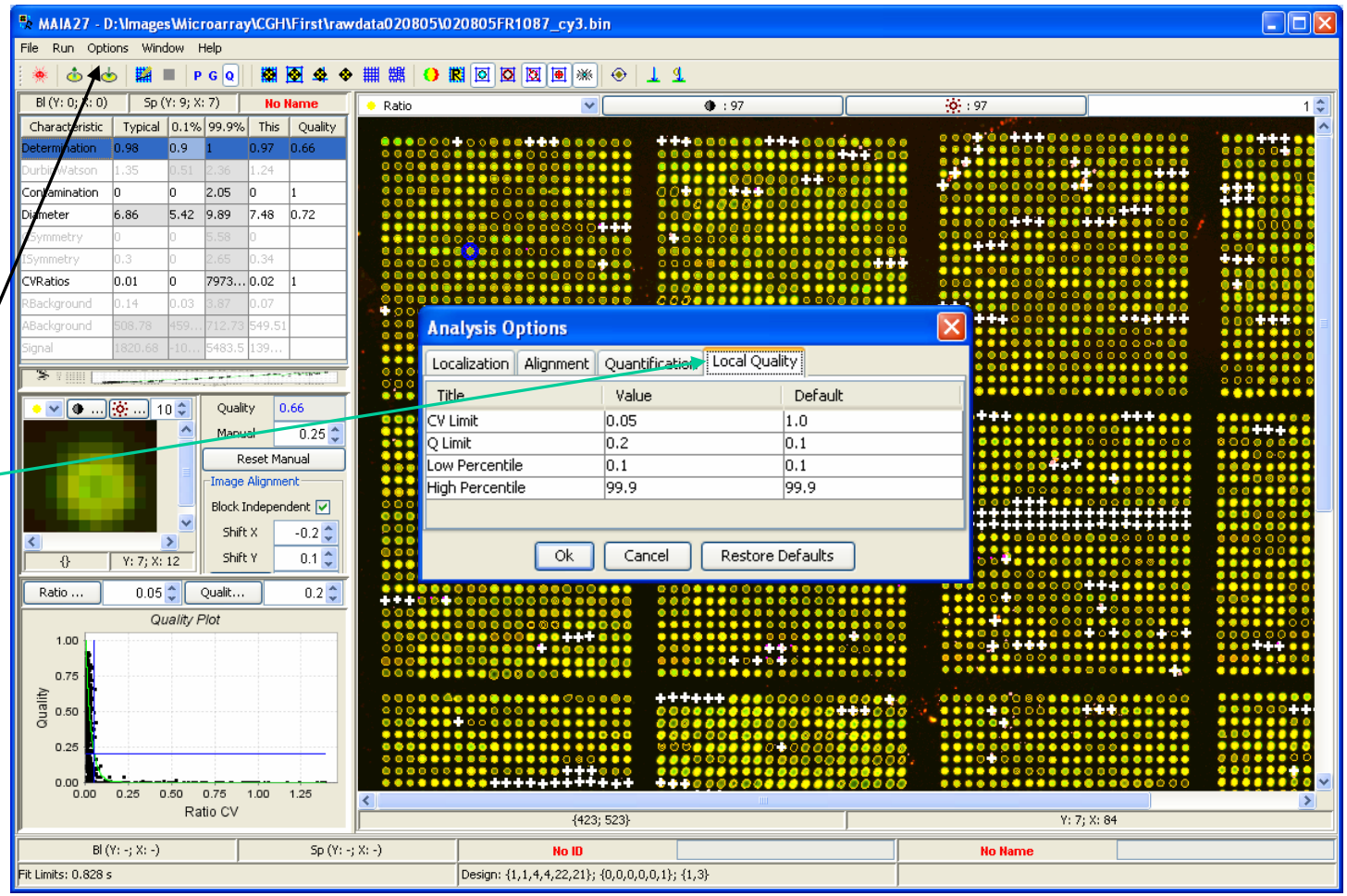
"Ratio CV Limit" and "Quality Limit" are used to generate the "ideal" quality curve. The decreasing rate of this curve characterizes how we are strict with respect to the spots quality. If this curve decays rapidly, one can expect that a lot of spots will be flagged out. This is a user decision, which depends on the image and user demands.



Quality Settings

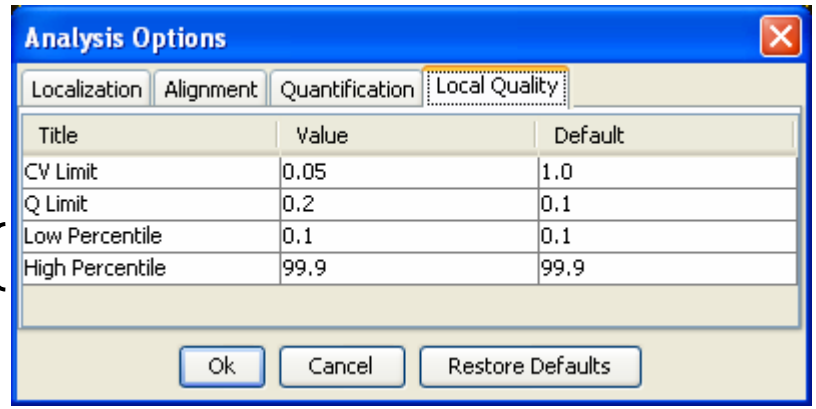
Several settings that may influence the quality analysis are available through the Menu Item “Options|Analysis Options” (Alt+O), tab “Local Quality”.

See next page for details.



Quality Settings in Detail

CV Limit is a characteristic value of the user-defined (ideal) quality curve.

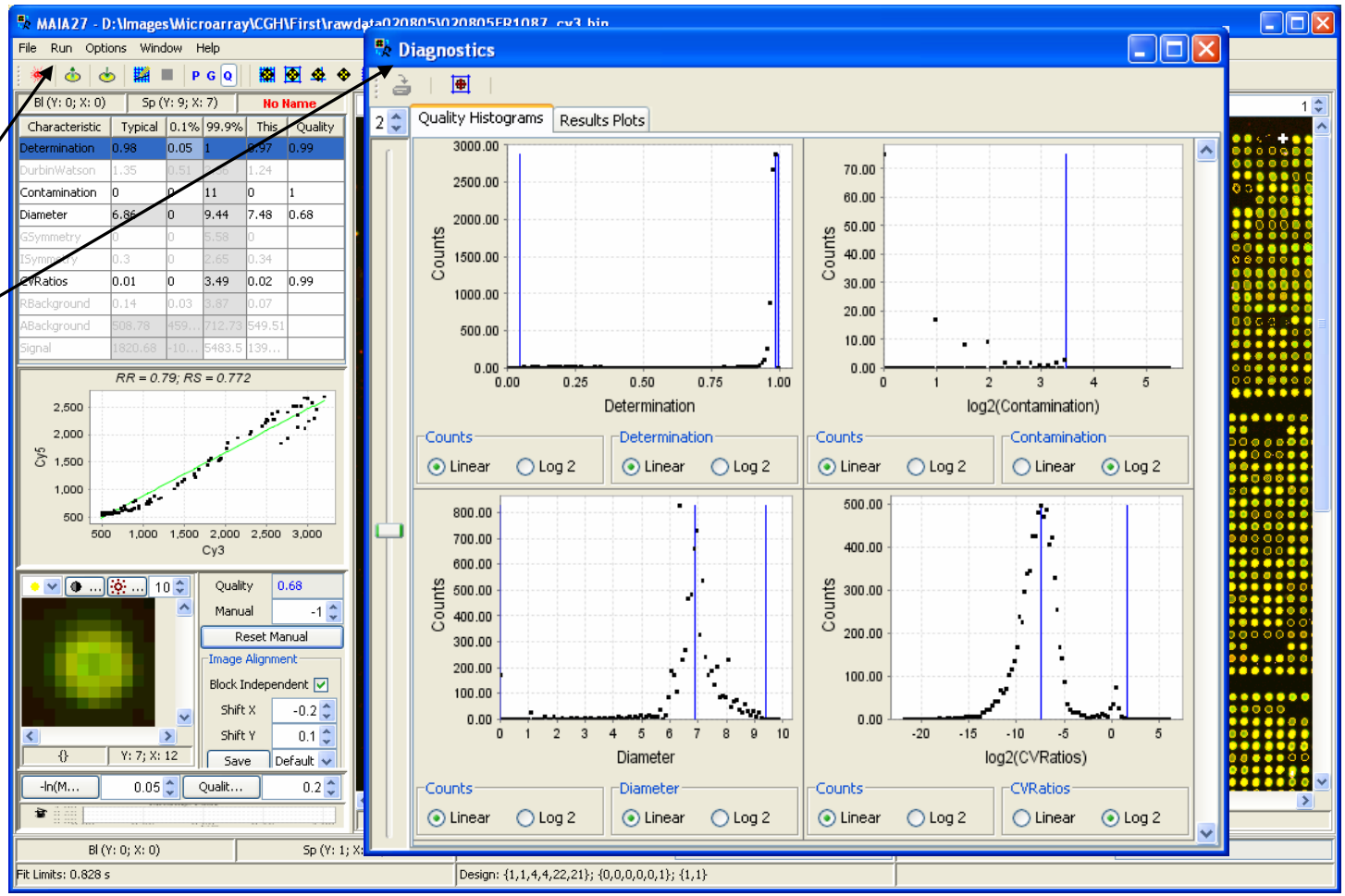


Q Limit is the limit such that the spots with the overall quality values below this limit will be indicated by a cross.

Low and High Quality Percentile establishes the values of the quality characteristics in the sorted lists of the quality characteristics (built up based on the results for all spots from the array) that will be displayed in the corresponding fields of the quality table.

Diagnostic Plots

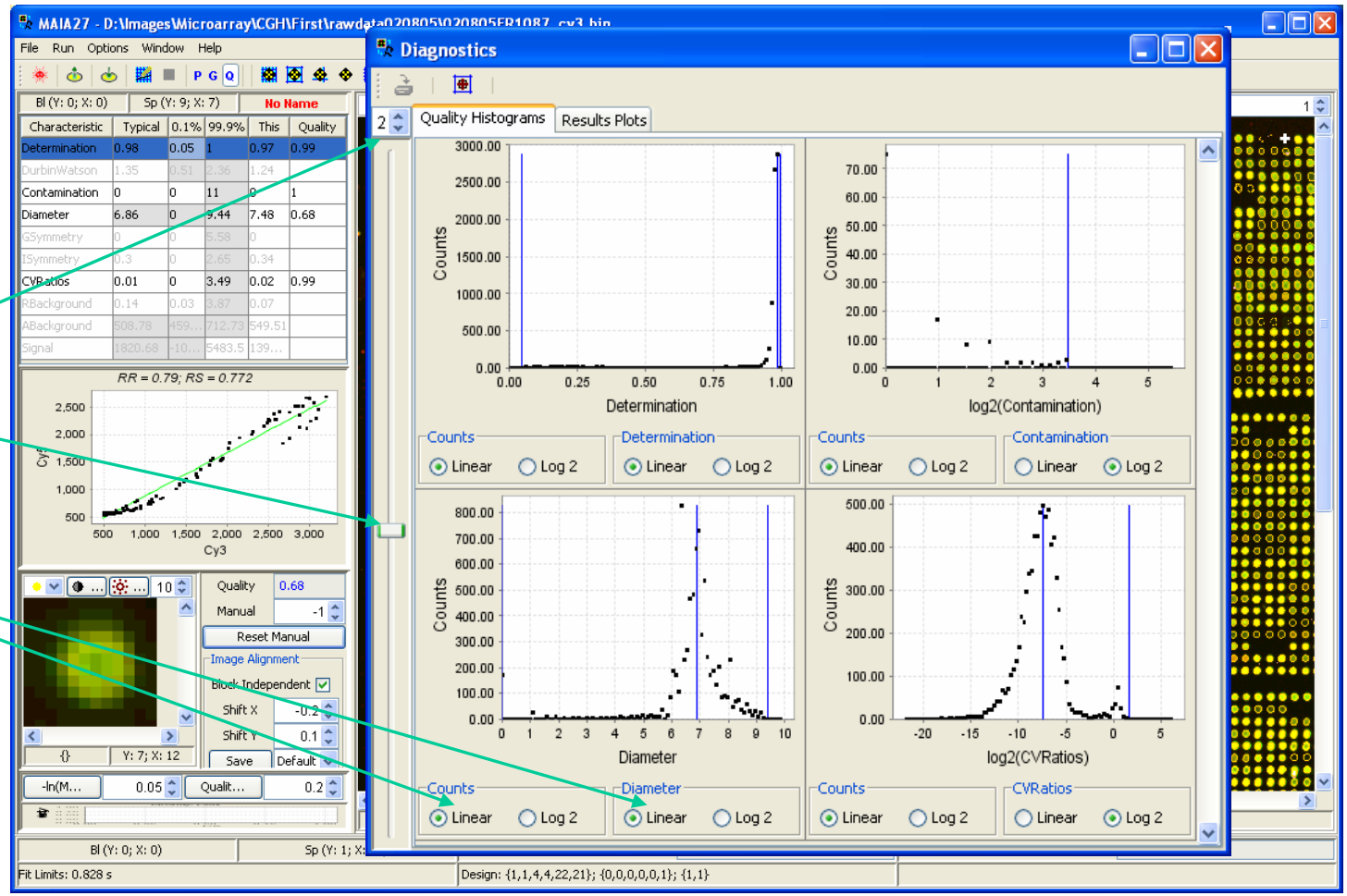
The Menu Item "Run|Diagnostics" (F3) opens up the window with different diagnostic plots.



Diagnostic Plots Layout

The “Columns” spinner defines number of graphs in the rows, and the “Row height” slider defines the height of the graphs panels.

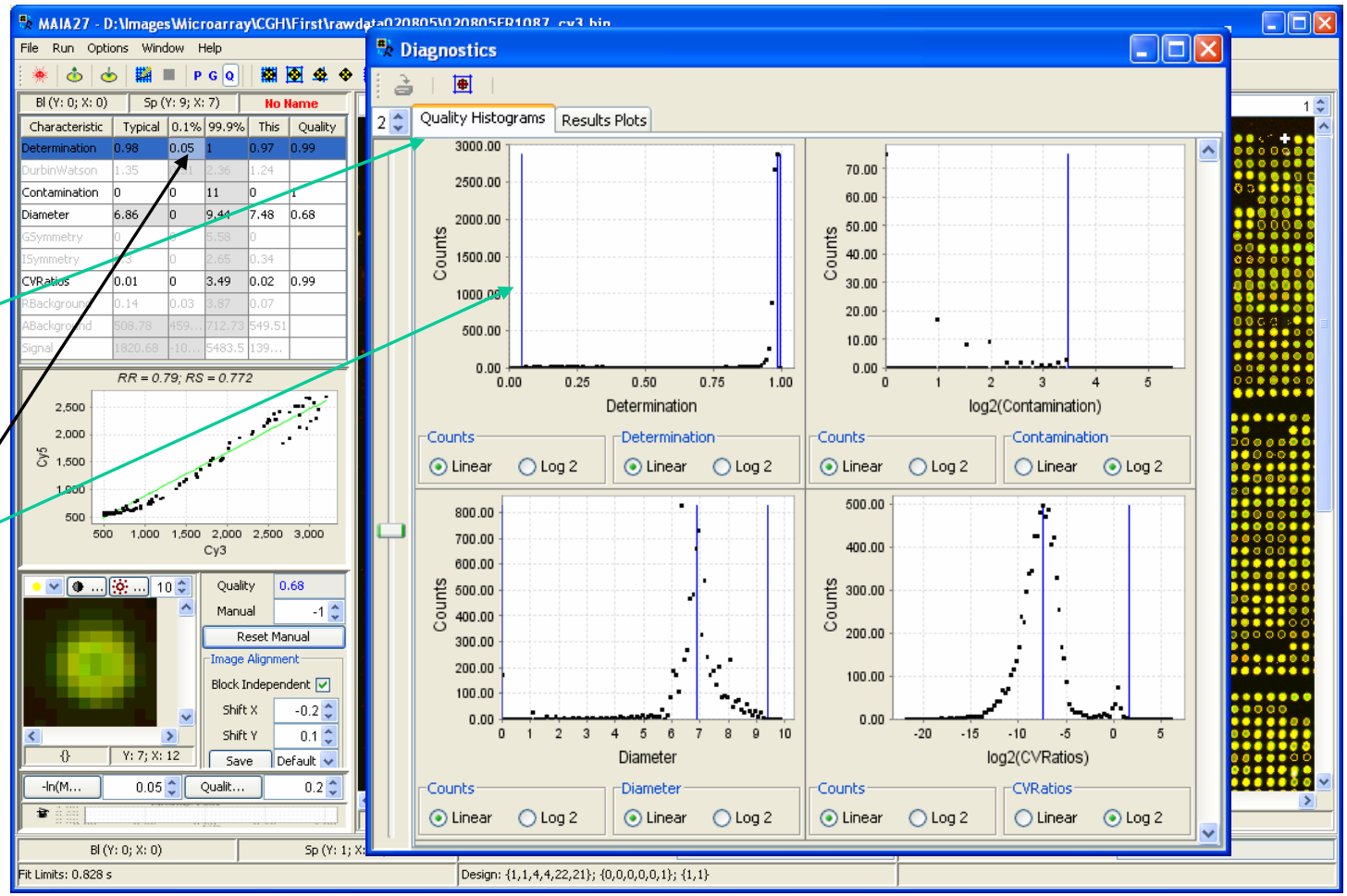
Graphs can be shown in linear or log scales.



Quality Histograms

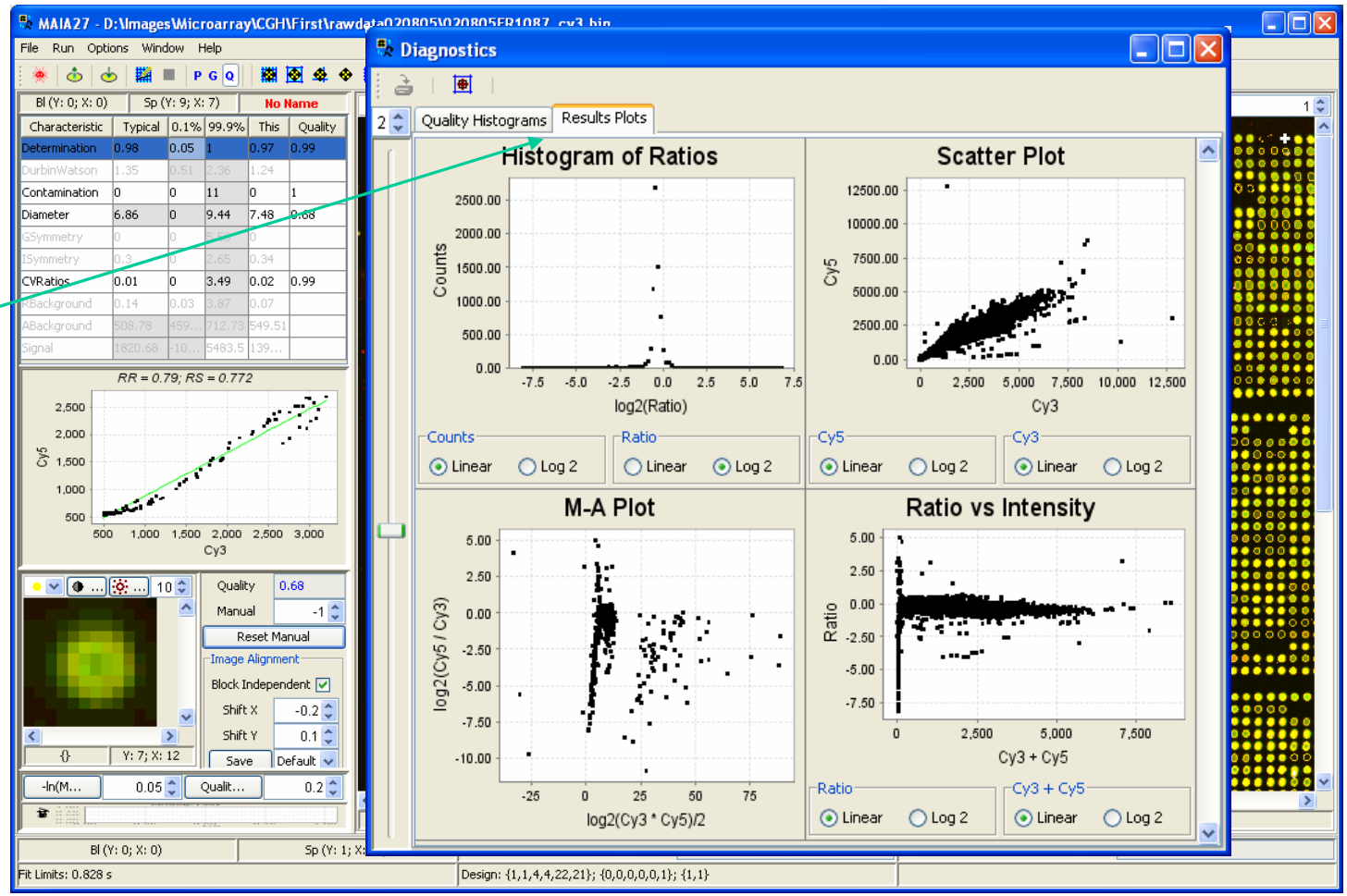
Quality Histograms panel contain histograms of the used quality characteristics.

Blue vertical lines correspond to the typical value and (0.1 and 99.9) percentiles from the quality table.



Results Plots

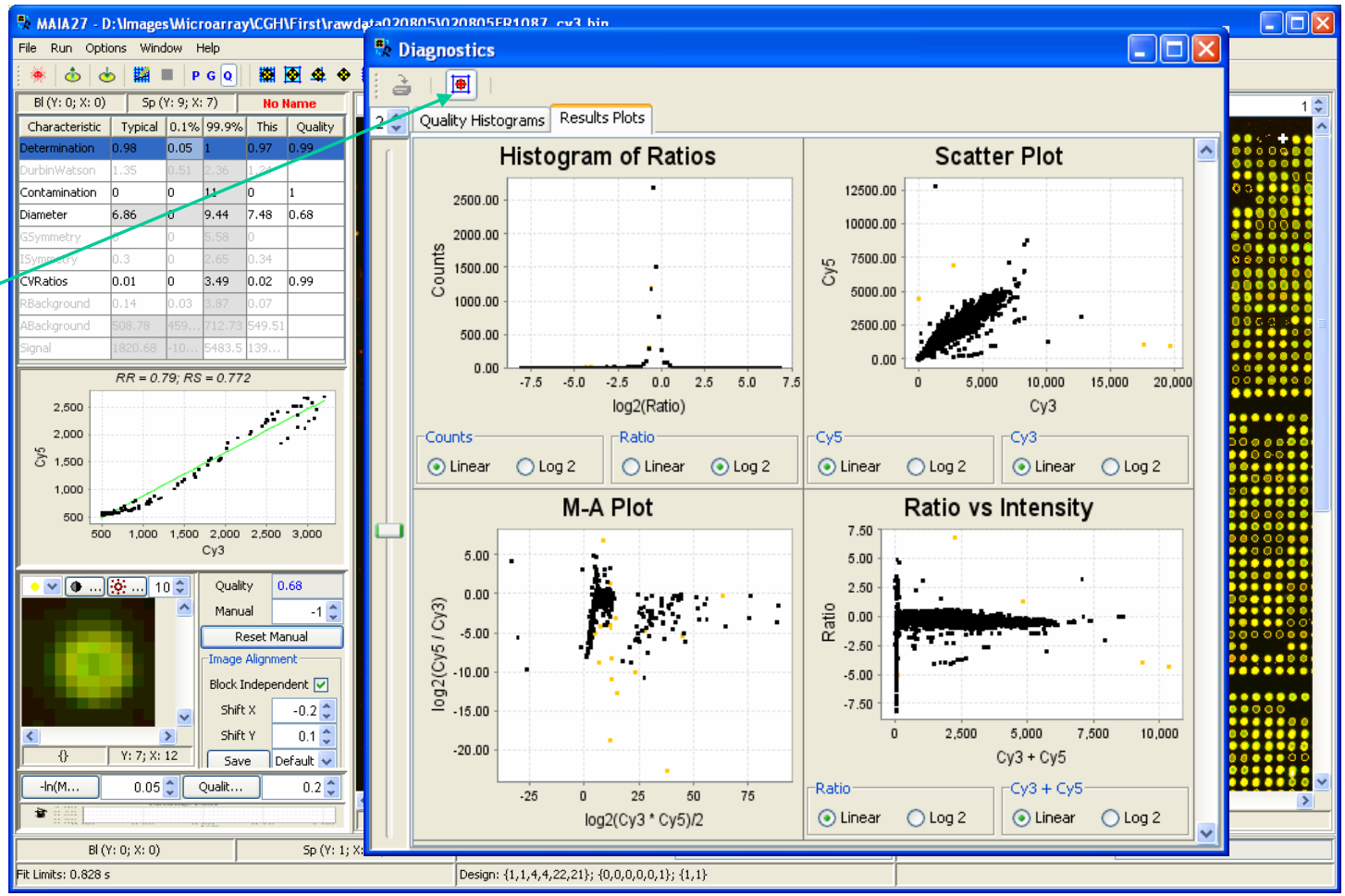
Results Plots panel contain:
Histogram of ratios;
Scatter plot;
M-A plot;
Ratio *vs* Intensity plot.



Diagnostic Plots with All Spots

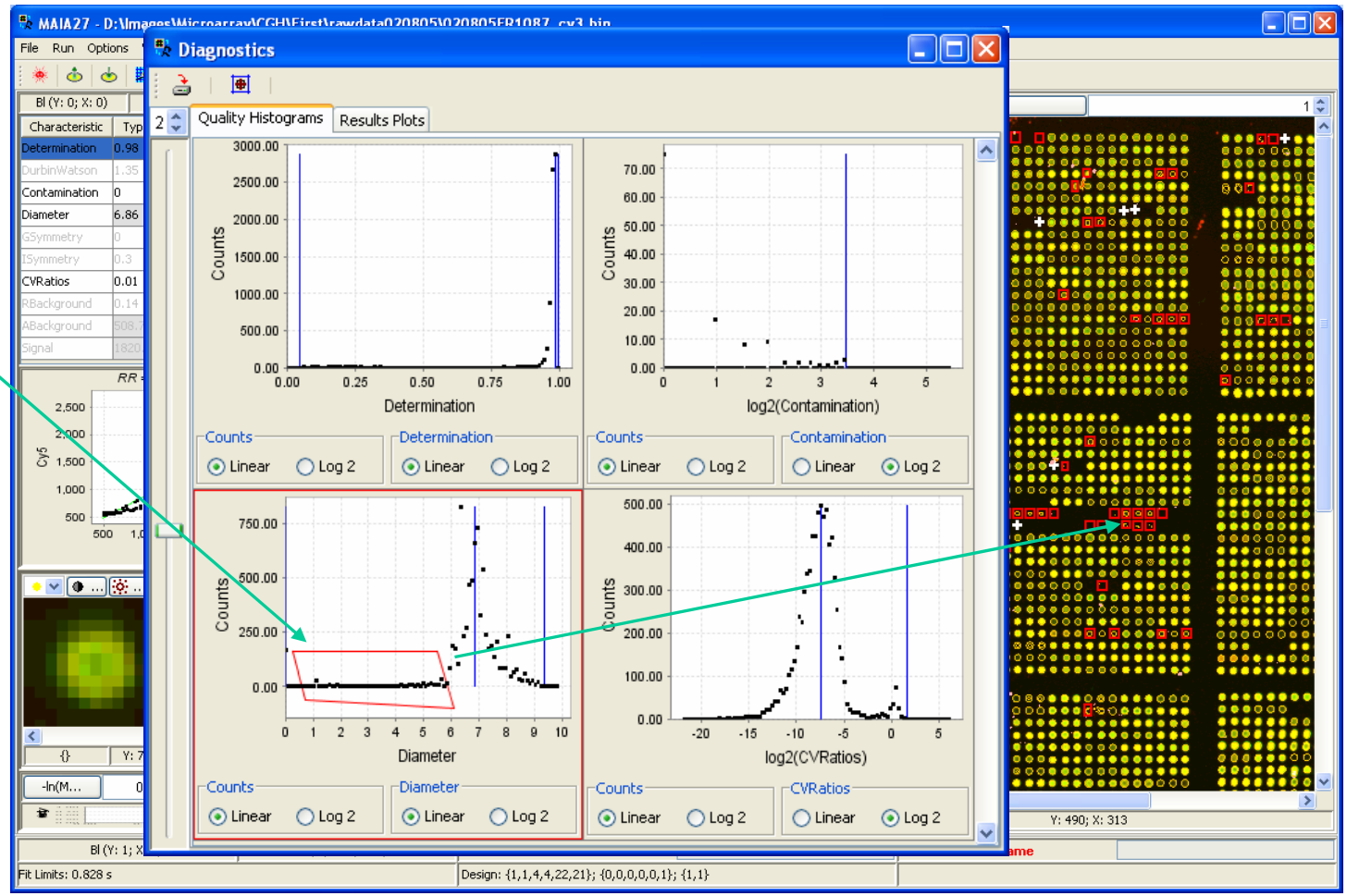
The Toolbar button "Show/Hide "Bad" Spots" allows one to show/hide "bad" spots on the diagnostic plots.

If the button is on, all spots are used to build up the diagnostic plots, and "bad" spots are indicated in orange.
If the button is off, only "good" spots are used to build up the diagnostic plots.



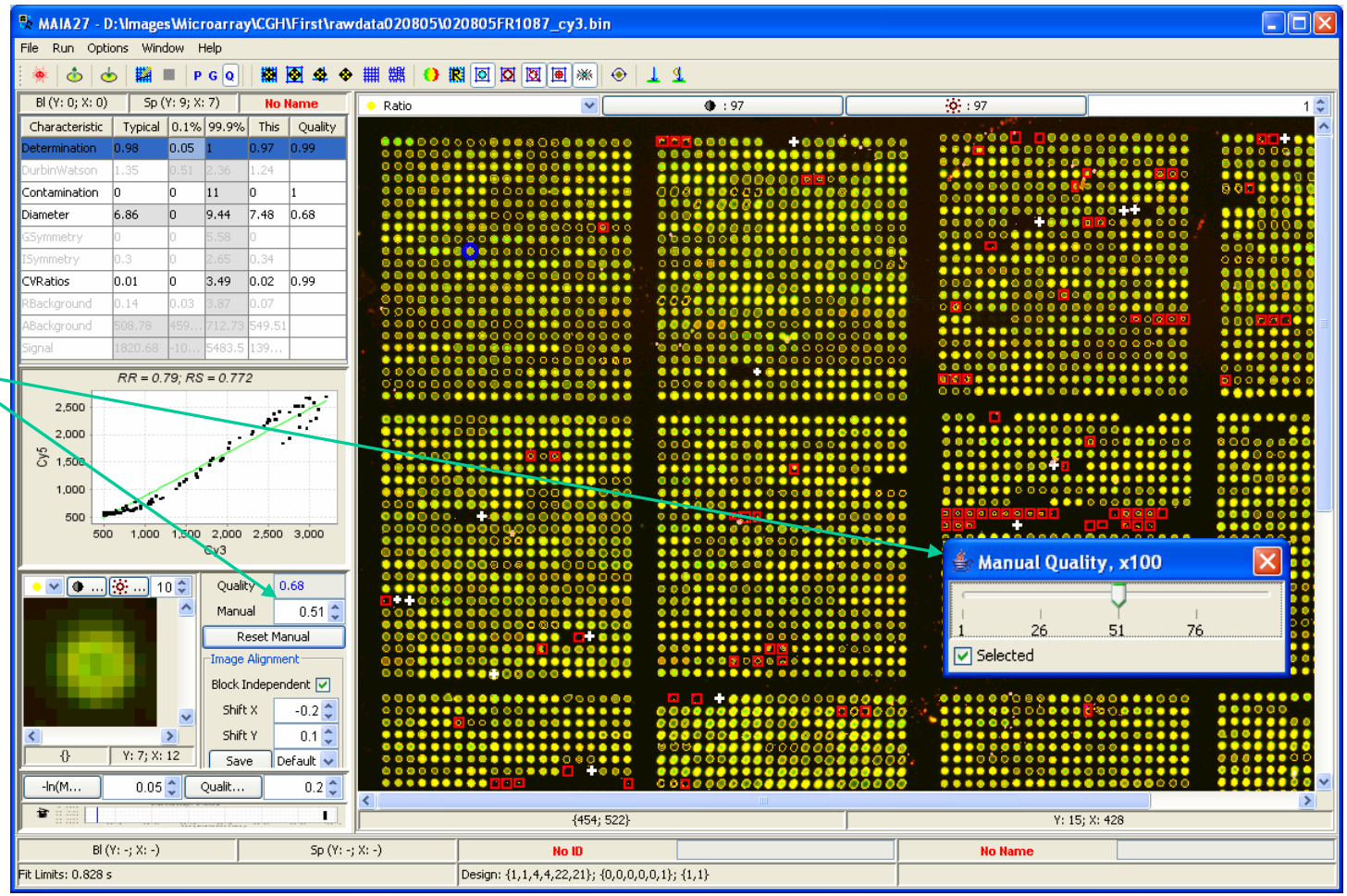
Spot Selection on Diagnostic Plots

Ctrl+Left Clicks followed by Ctrl+Right Click on the diagnostic plots select the spots to be highlighted on the image.



Manual Qualification of the Selected Spots

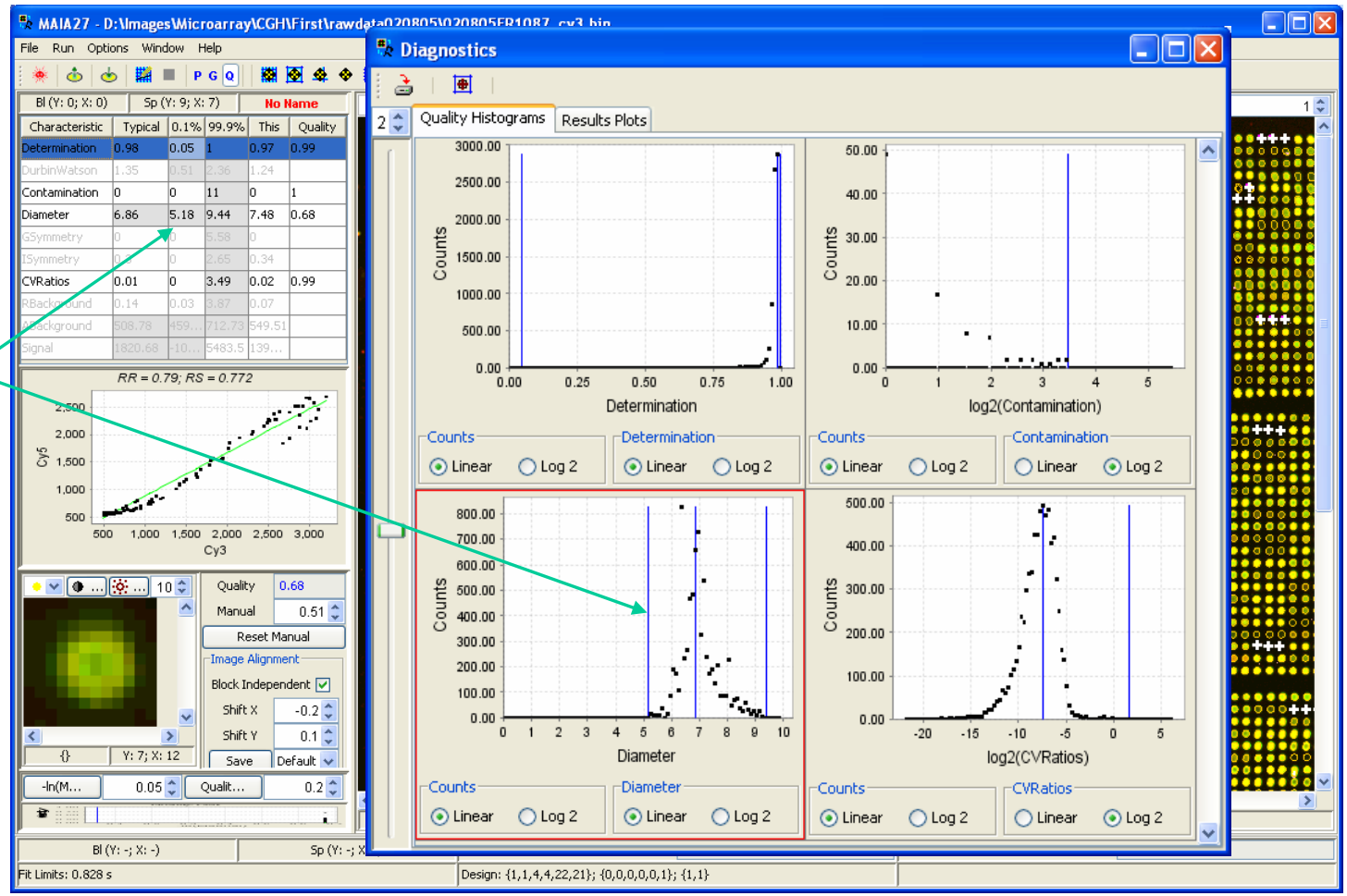
Using the spinner "Manual" one can assign a quality value to the selected spots.



Change the Quality Limits

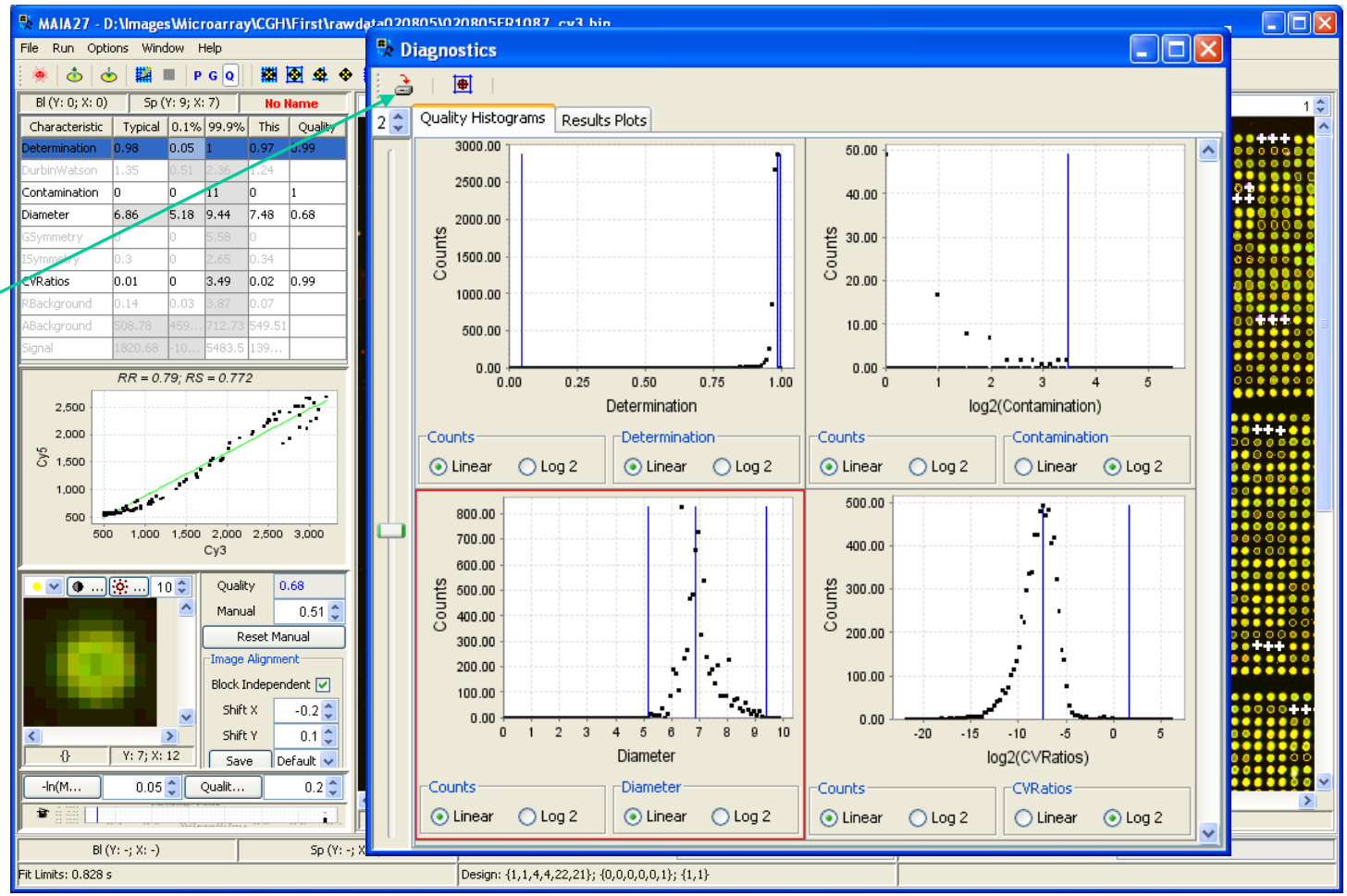
For each used quality characteristic the limits can be adjusted in the Quality Histograms.

Shift+Left Click – Move – Shift+Right Click transfers a typical value or a limit into a new location of the quality histogram.



Save the Selected Plot

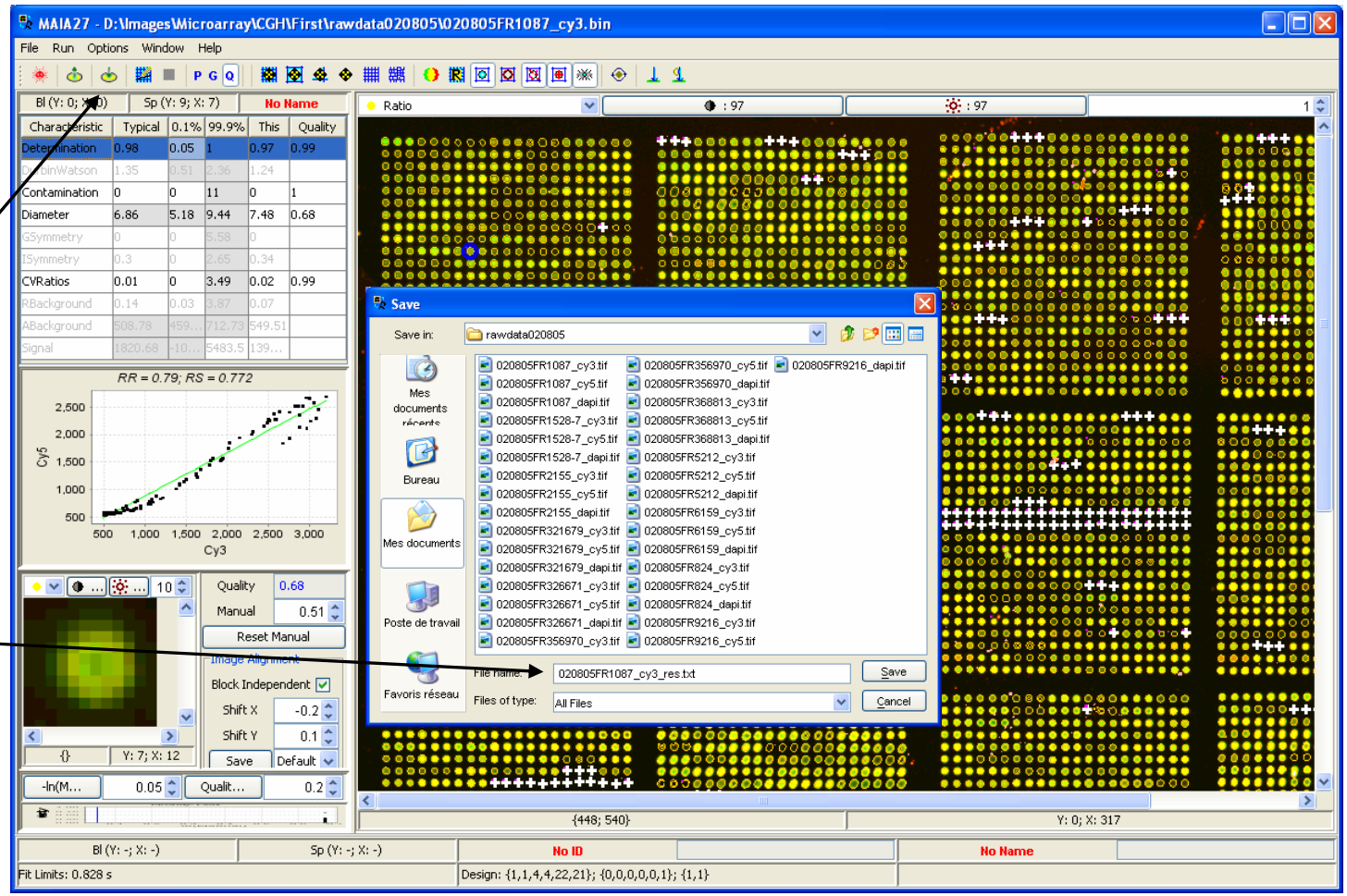
Any diagnostic plot can be saved as an image file (tif/jpg/gif/bmp formats).



Save the Results

To save the results of quantification and quality analysis use the “Save Analysis ...” button from the Toolbar or the Menu Item “File|Save|Analysis …” (Ctrl+S).

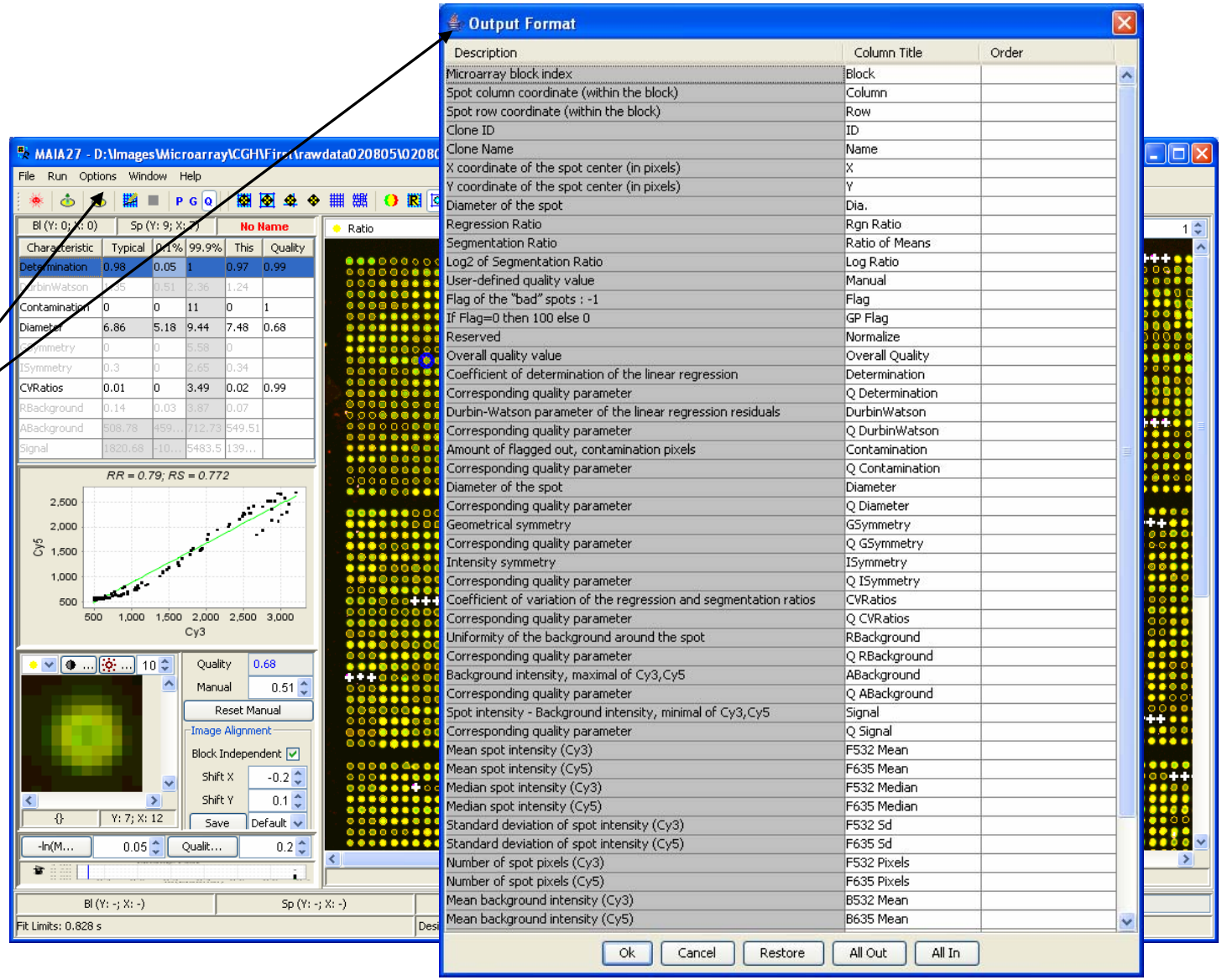
The results are saved as a table in the text file (importable into Microsoft Excel).



Output File Format

User can define which fields and in which order should be presented in the output file: select the Menu Item "Options|Output Format" (Alt+F).

See next page for details.



Output Table Format in Detail

Description of the field (non-editable).

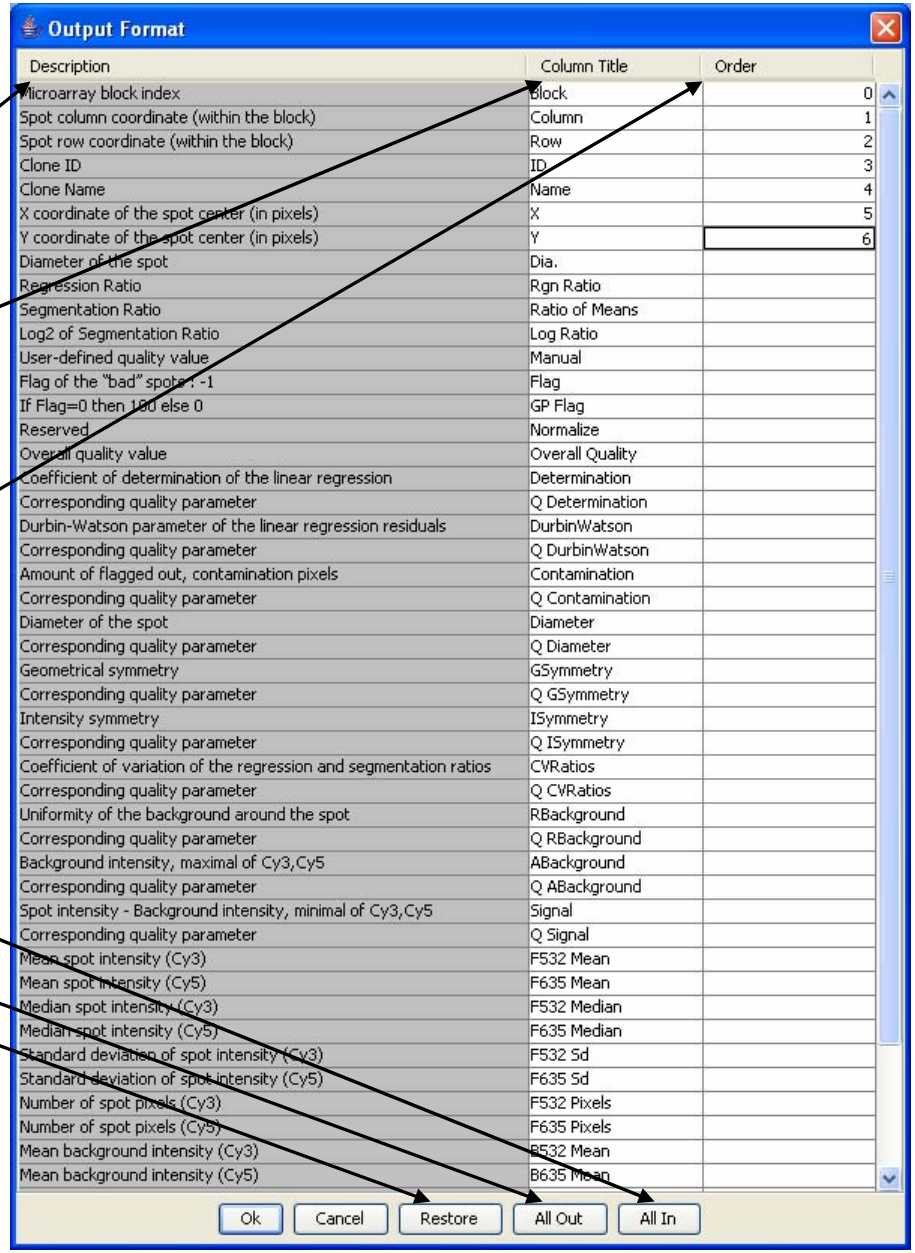
Editable name of the field to be appeared in the output file.

Order specifies the sequence of the fields. If this field is empty, the corresponding field is not included in the output file.

Include all fields.

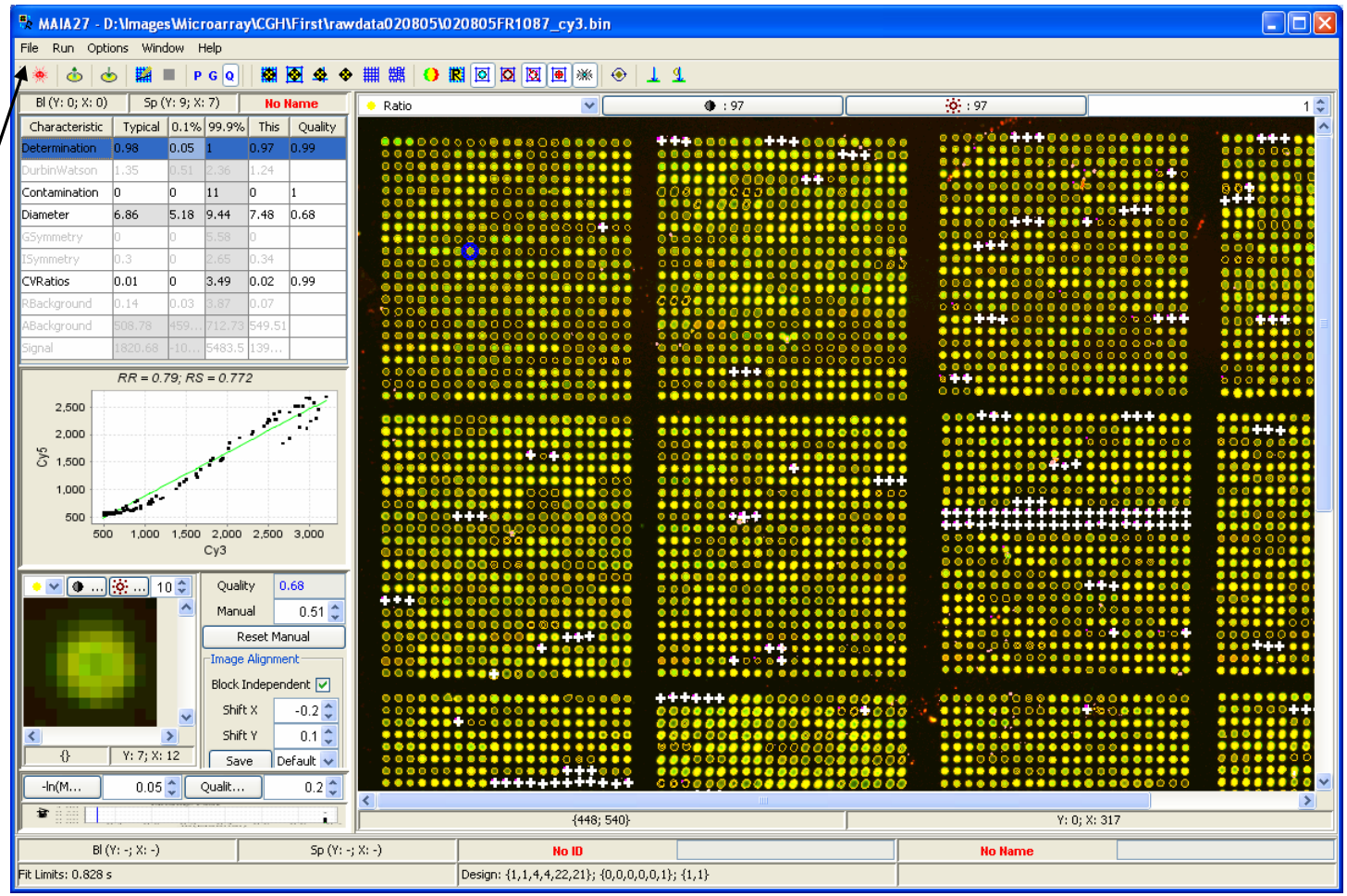
Exclude all fields.

Restore previous set of fields.



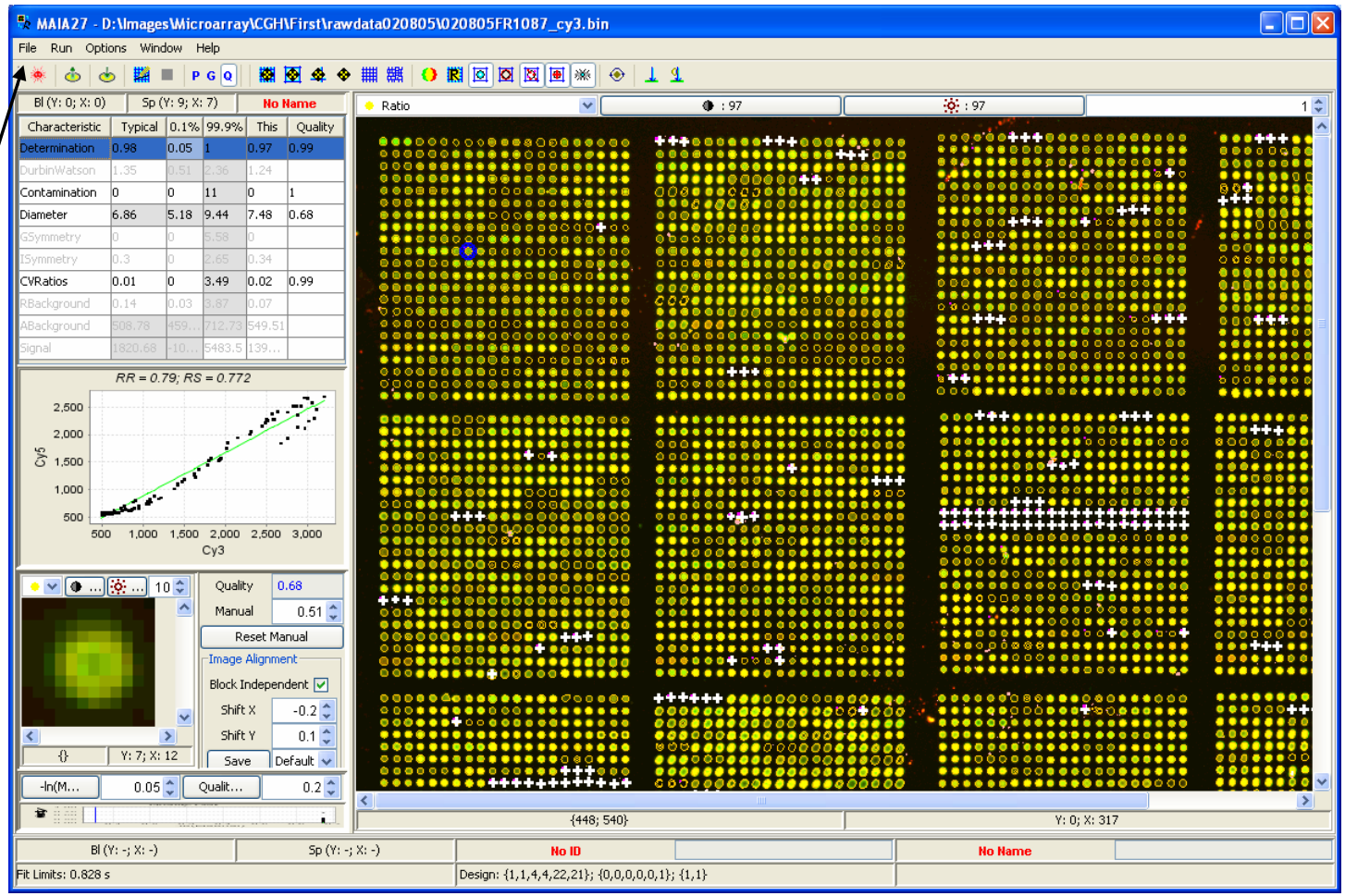
Save the Experiment: Experiment File

The whole experiment (results, parameters, grid, and other settings) can be saved on the disk (using the Menu Item "File|Save|Experiment …" (Ctrl+W)) in the internal (binary) format to be able to restore it (using the Menu Item "File|Load|Experiment …" (Ctrl+R)) to reanalyze the data.



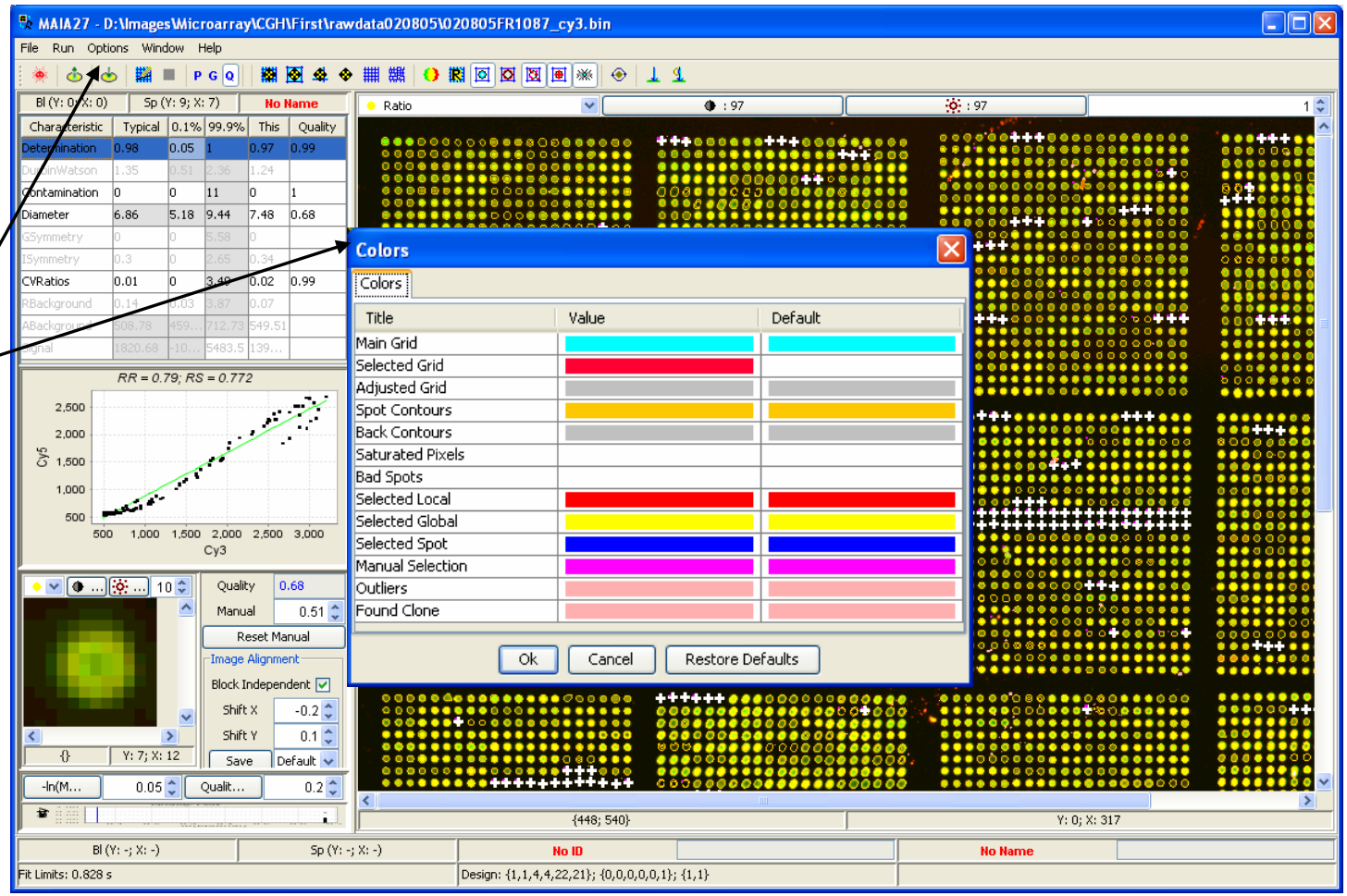
Set Batch Options

Using the Menu Item "File|Set Batch Options", all settings from the *Main Processing Window* can be sent to the *Batch Processing Window* to be applied to the other images from the same batch.



Colors

To change the color of some elements of the localization and quantification outputs use the Menu Item “Options|Colors”.

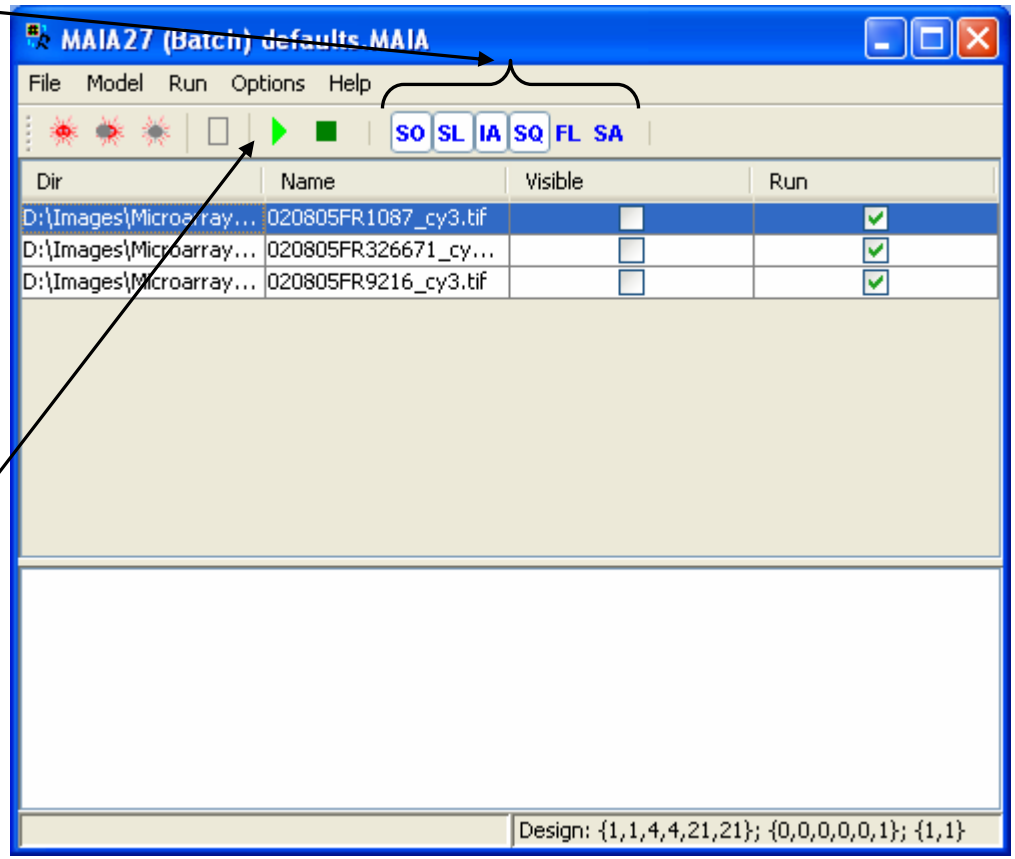


Batch Processing

To run batch processing a list of actions to be applied to files in the batch should be defined:

SO – Set Options;
SL – Spot Localization;
IA – Image Alignment;
SQ – Spot Quantification;
FL – Fit Limits;
SA – Save Analysis;

The batch processing can be started using the "Run Batch" button from the Toolbar or the Menu Item "Run| Run Batch" (F5).



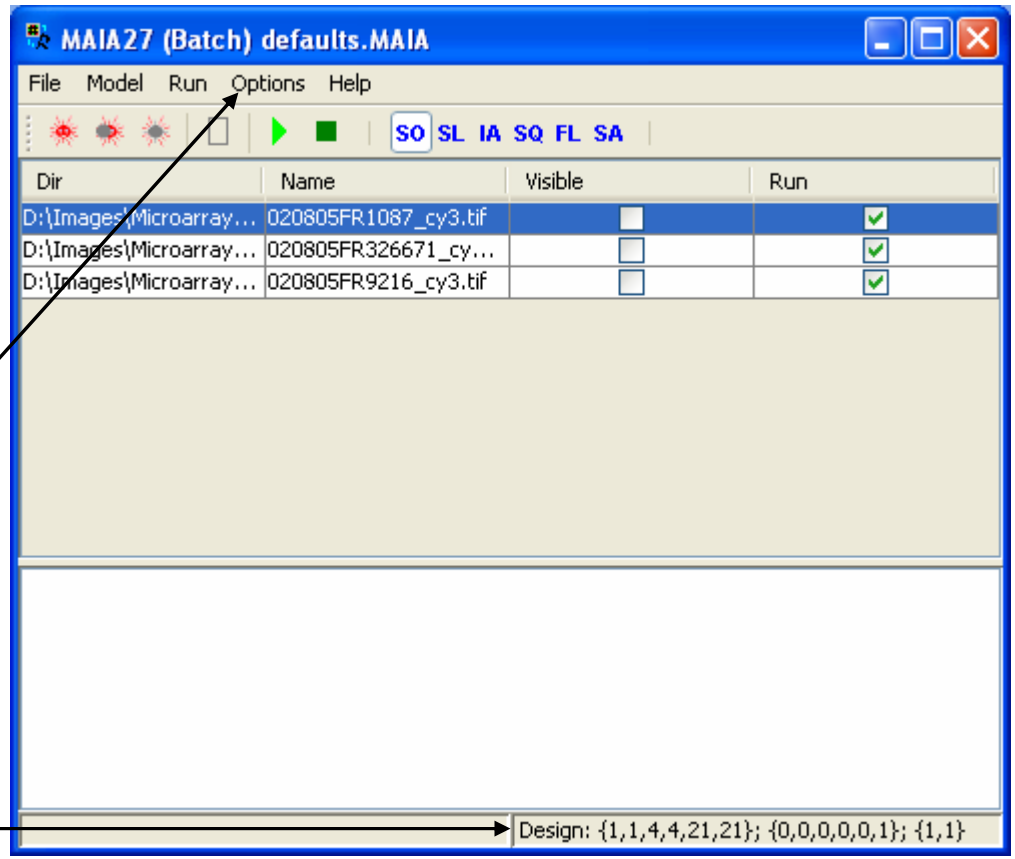
After the first processing, images with the obtained results (grid, parameters, settings, etc) are saved on the disk in the internal (binary) format (experiment files). If the program is unable to find such a file, it opens up the original image and applies the default settings (which can be defined via different items of the Menu "Options": "Data Options", "Analysis Options", "Colors", "Array Design" and "Output Format").

Modify Batch Settings

Typically all arrays from the batch are of the same array design, and have the same settings.
One may want to define/modify these settings before further processing.

This can be done using the items of the Menu "Options": "Data Options", "Analysis Options", "Colors", "Array Design" and "Output Format".

Description of the current Array Design

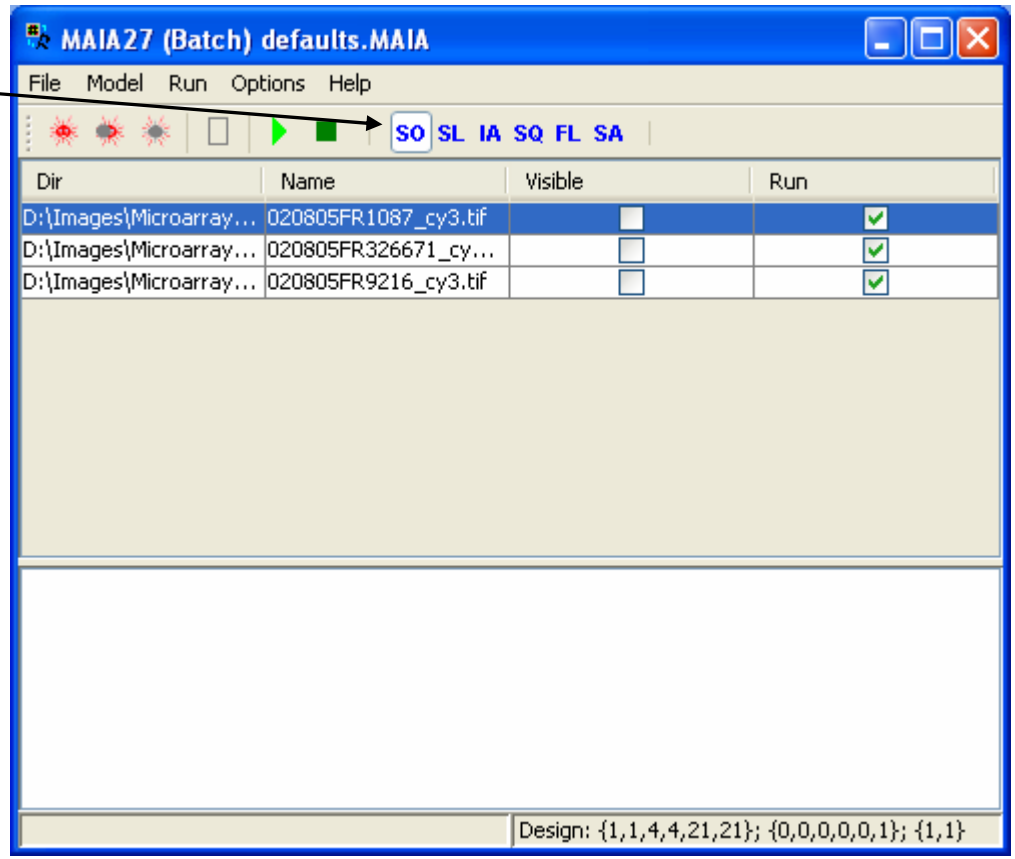


Yet another possibility to modify settings is to open (visualize) one of the images and modify settings for that image. Then the Menu Item "File|Set Batch Options" of the *Main Processing Window* will send the new settings into the *Batch Processing Window*.

See page Set Batch Options.

Apply Setting to the Batch

To send the modified settings to all images of the batch one needs to run the batch with the task “Set Options” (The toggle button “SO” is pressed).

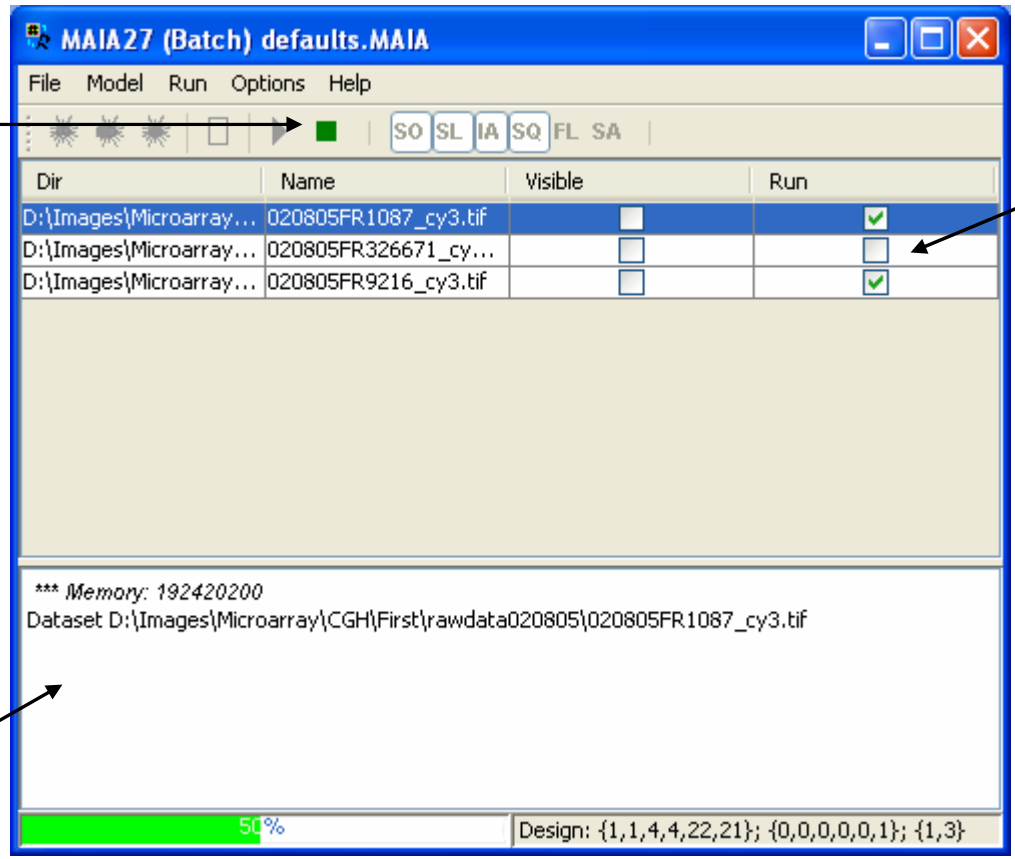


This is required only if the binary files have already been generated.

Otherwise new settings will be applied as defaults in the processing of each new image from the batch.

Run Batch

Batch processing can be stopped by pressing the “Stop” button” on the Toolbar or selecting the Menu Item “Run|Stop” (Ctrl+F5).

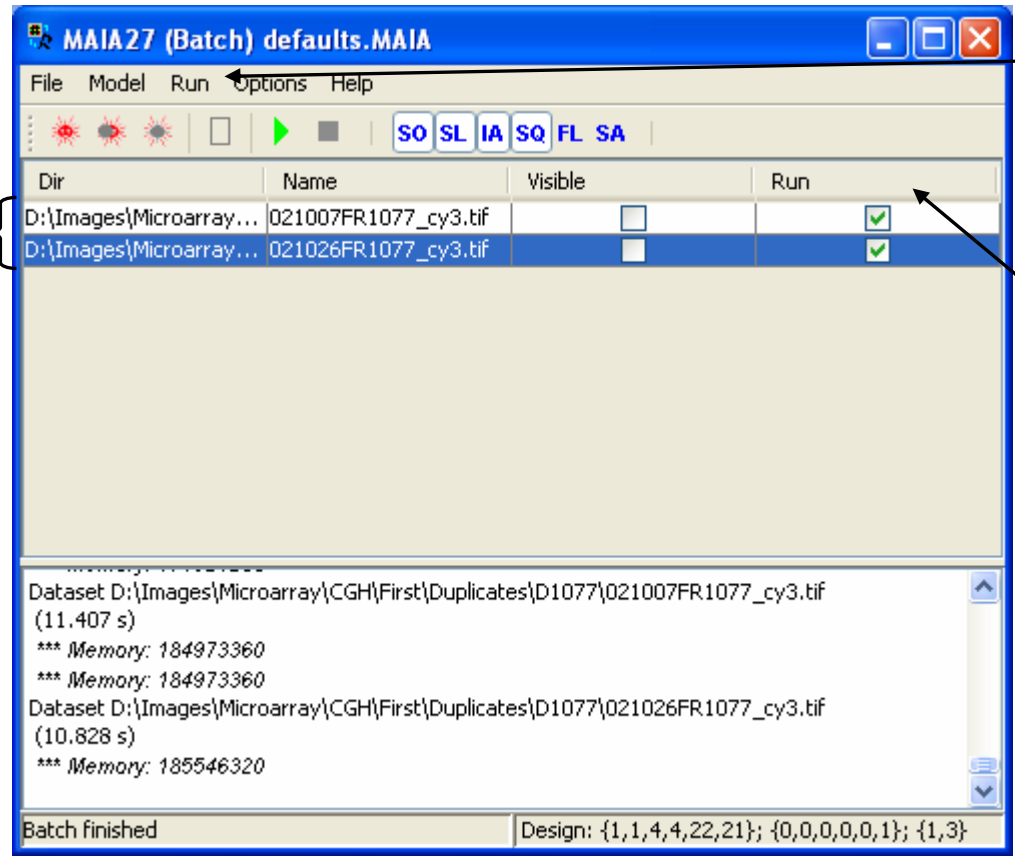


Using the field “Run” one may exclude (include) certain files from (in) the Batch processing.

Protocol of the Batch processing.

Global Quality Analysis

To start global quality analysis two, or more, arrays have to be selected and quantified.

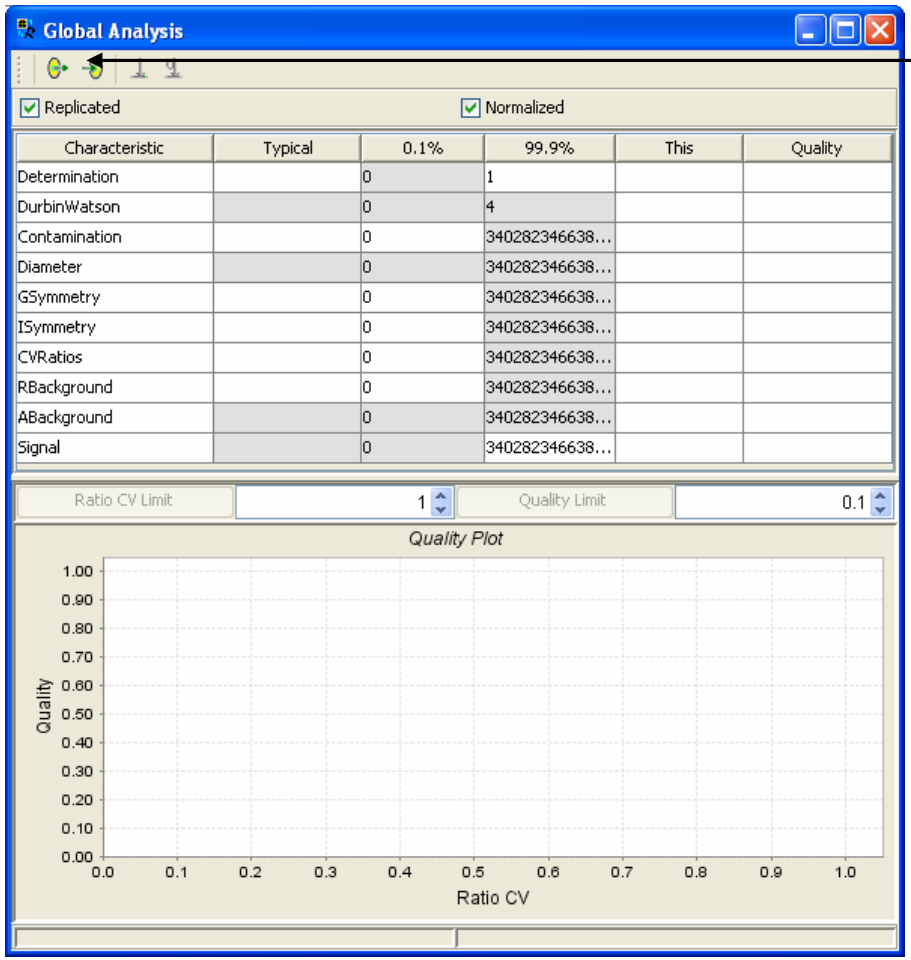


Select the Menu Item "Run|Global Analysis" to open the window for identification of the global *Quality Limits*.

Check the field "Run" to specify which arrays will be used for global quality analysis.

Global Quality Analysis: Main Window

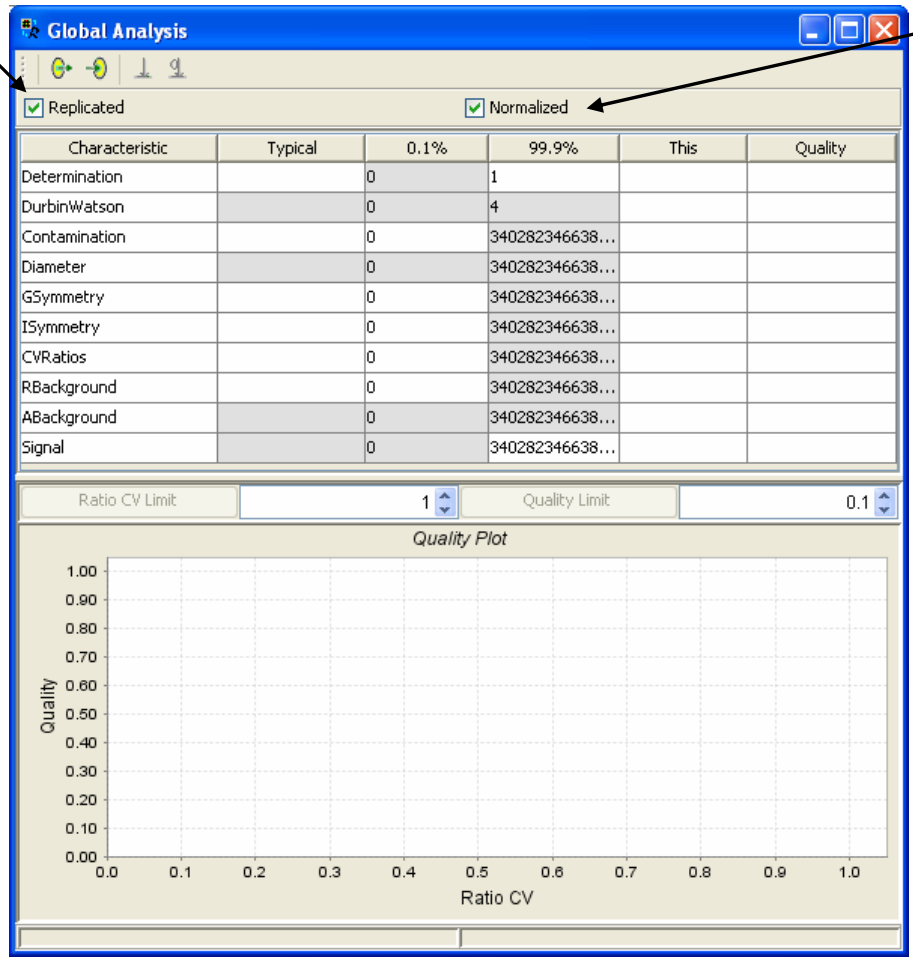
Global quality analysis panel shows up with the same set of quality characteristics as for each particular image.



Press the Toolbar button "Get Experiments" to copy quantification results from all selected arrays into the global quality analysis window.

Global quality analysis can be performed assuming that the selected arrays are either replicates or not.
If they are replicates, then all locally replicated spots from different arrays are combined, and a unique overall quality value and a unique ratio CV are calculated for each replicated clone.
If the selected arrays are not replicates, then local spot replicates* from different arrays are treated independently in the overall quality plot.

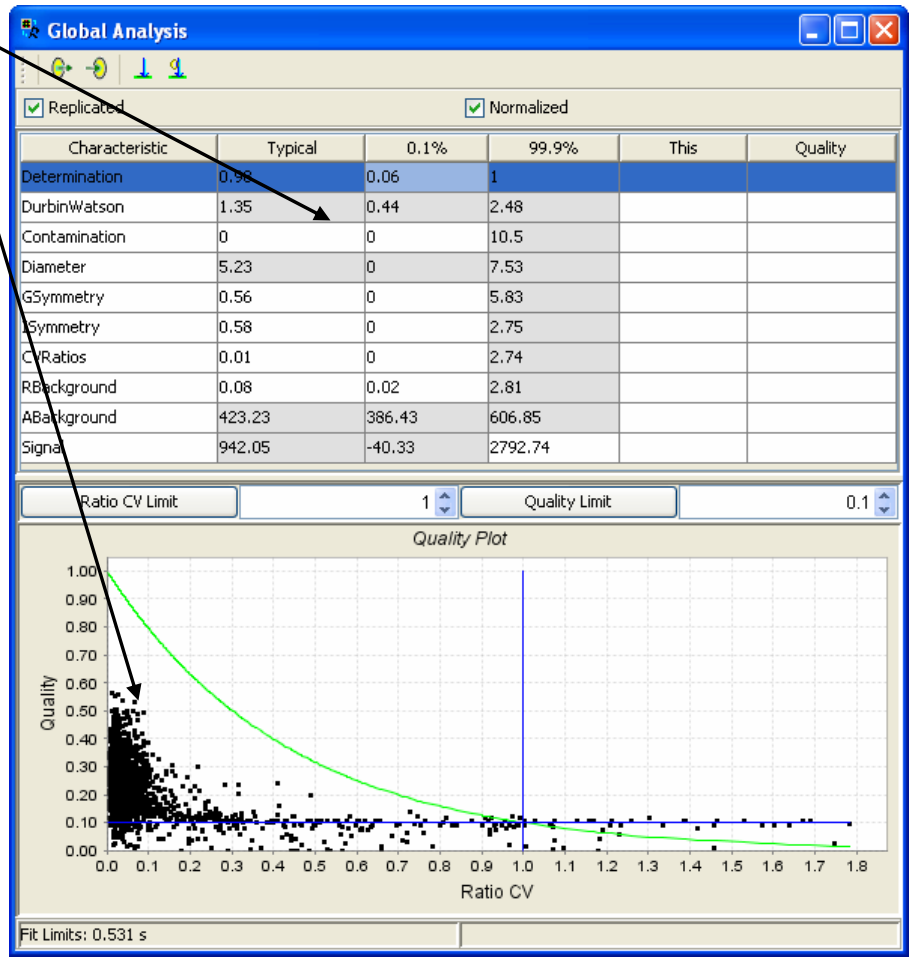
*) In this case, to have local spot replicates is essential for quality analysis.



If the selected arrays are replicates, then before combining locally replicated spots from different arrays into a unique overall quality value and a unique ratio CV, one may want to align arrays, so that the averaged log ratio is equivalent for all arrays in the selection.

Results Downloaded

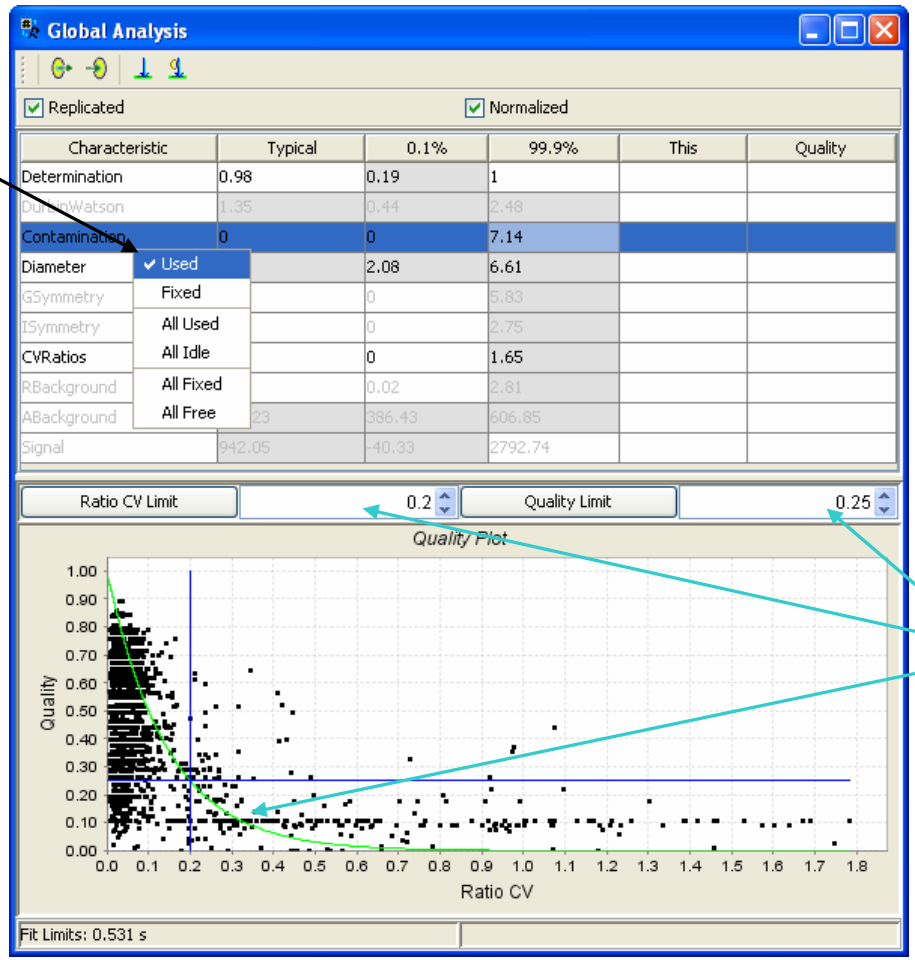
The quantification results have been downloaded.



The following quality analysis procedure is equivalent to the quality analysis performed for each particular image.

Global Quality Plot

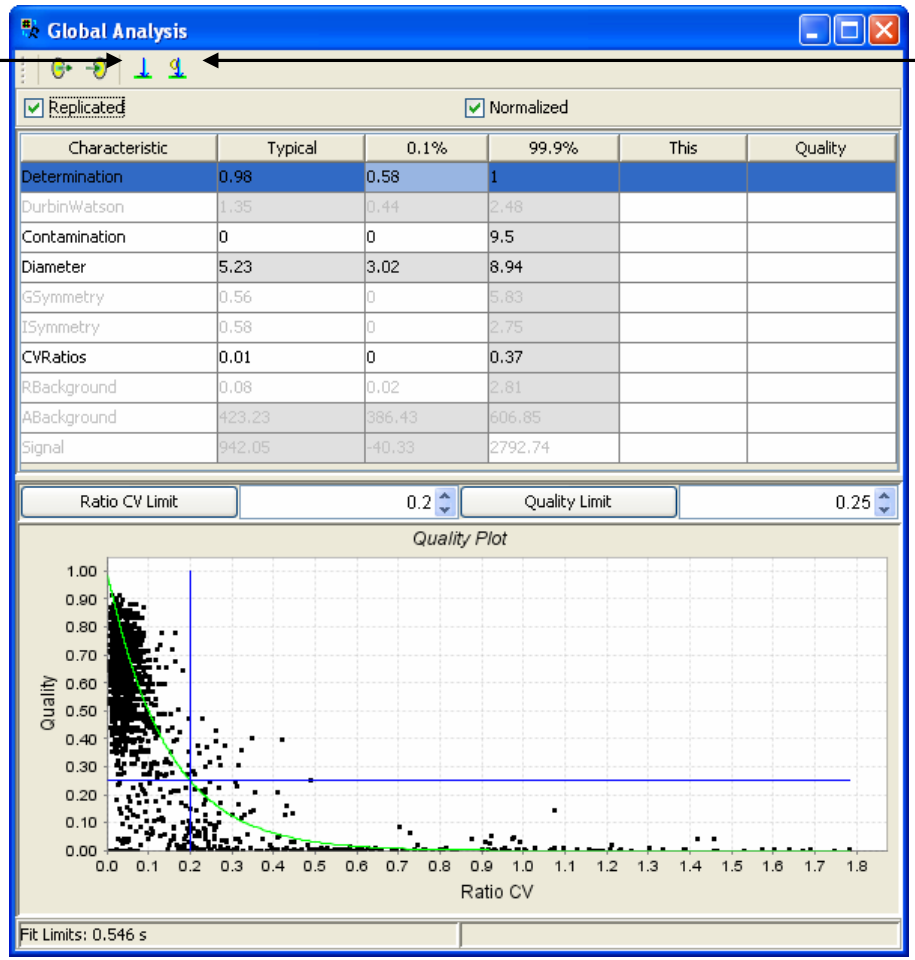
A set of used quality characteristics can be defined.



To identify the shape of the quality curve one can use the same tools as for each particular image.

Fit the Limits

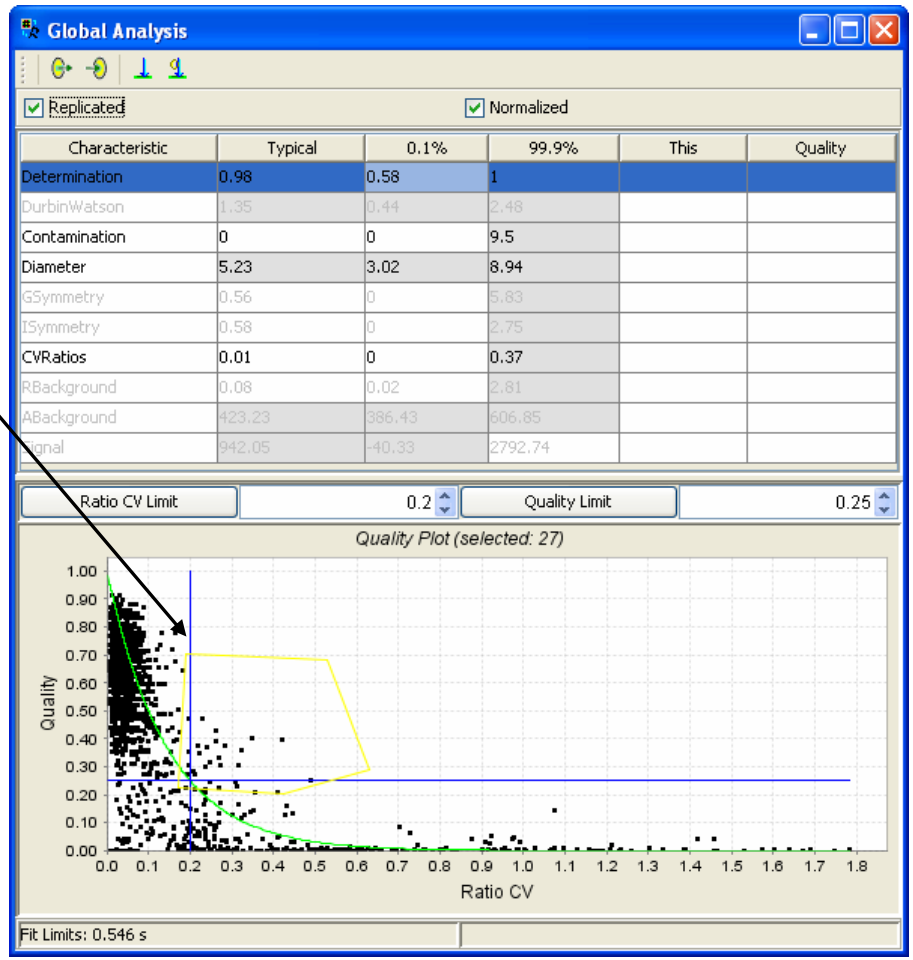
To initialize the Limits use the “Init Limits” button from the Toolbar.



To run fitting procedure use the “Fit Limits” button from the Toolbar.

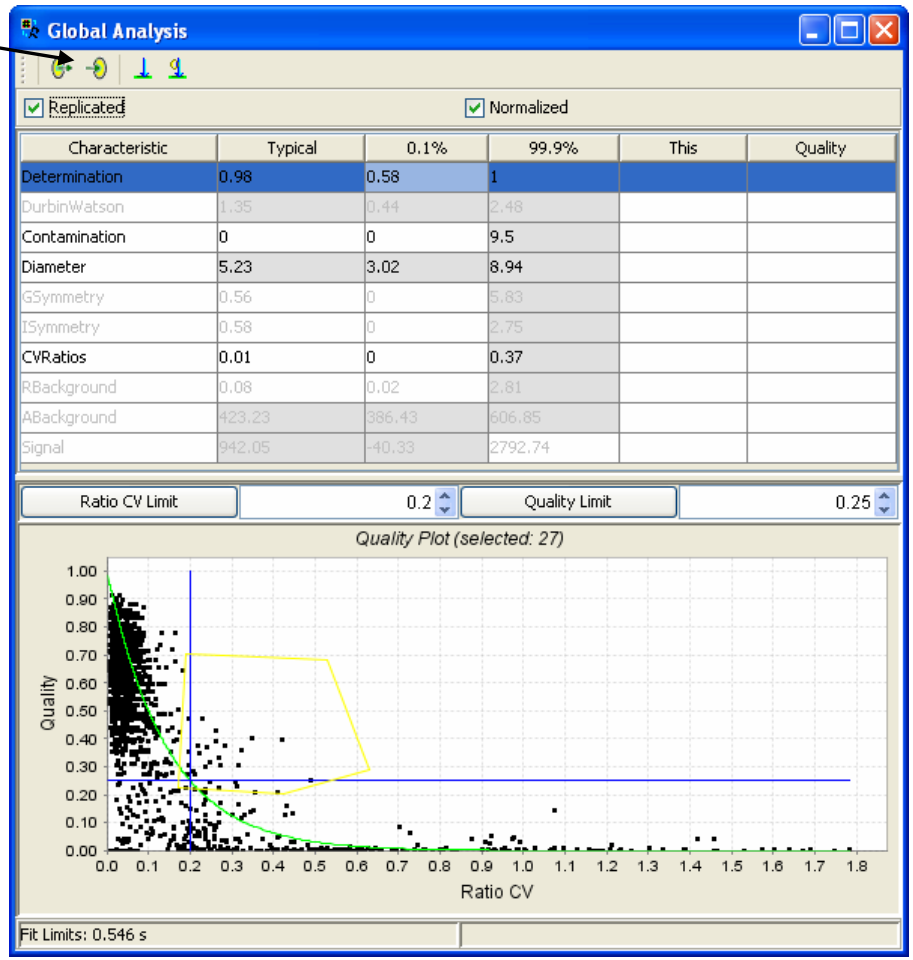
Select "Bad" Replicates

Ctrl+Left Clicks followed by Ctrl+Right Click create the contour on the Quality plot. This contour selects the replicates to be able to find them on the arrays from the globally analyzed selection of arrays.



Export Quality Limits

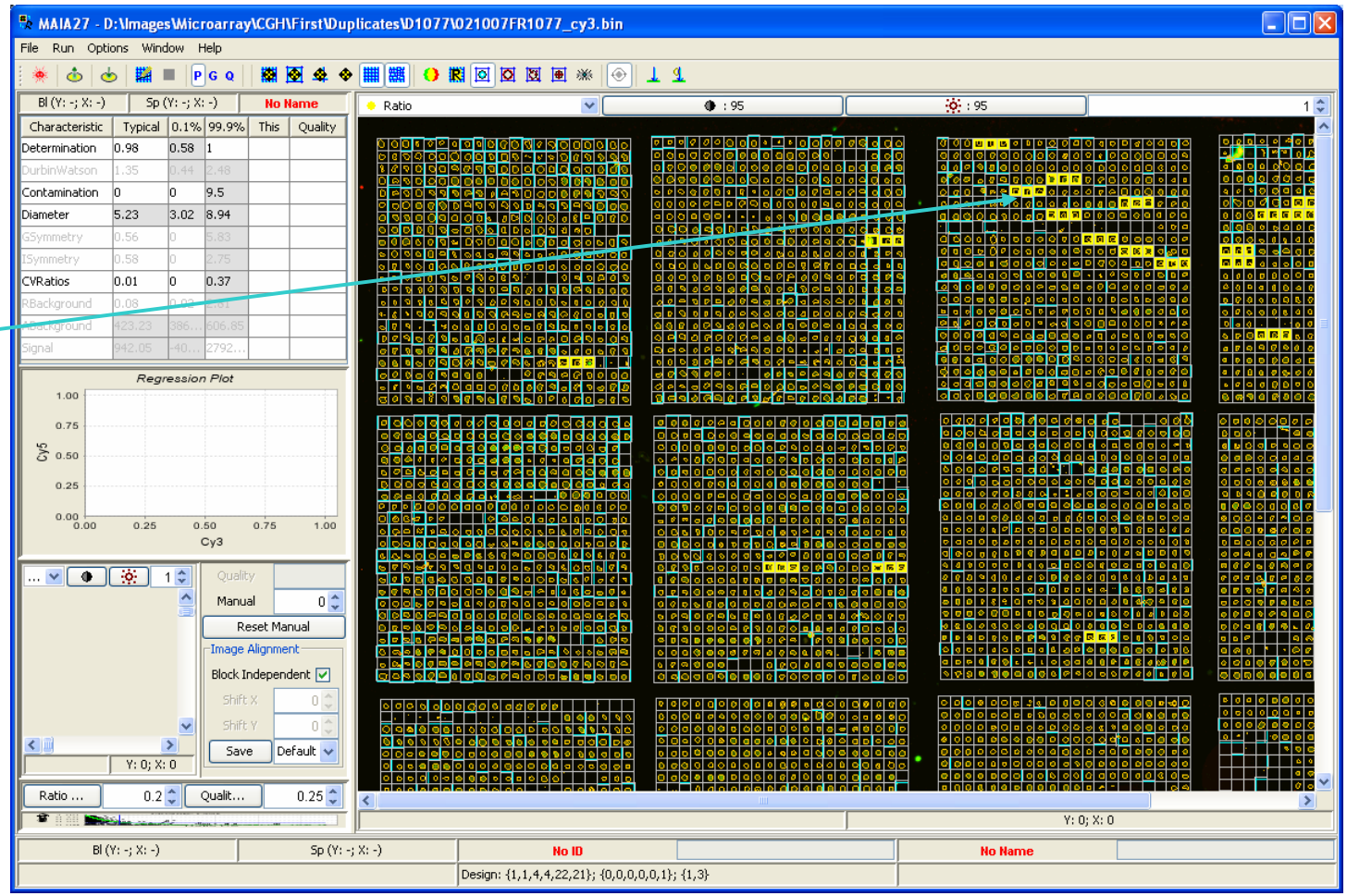
To send the obtained quality limits and selected replicates to each array file from the given selection press the Toolbar button “Set Limits”.



The Selected "Bad" Replicates (I)

The selected replicates are highlighted on both arrays.

The first array "021007".



The Selected "Bad" Replicates (II)

The selected replicates are highlighted on both arrays.

The first array "021026".

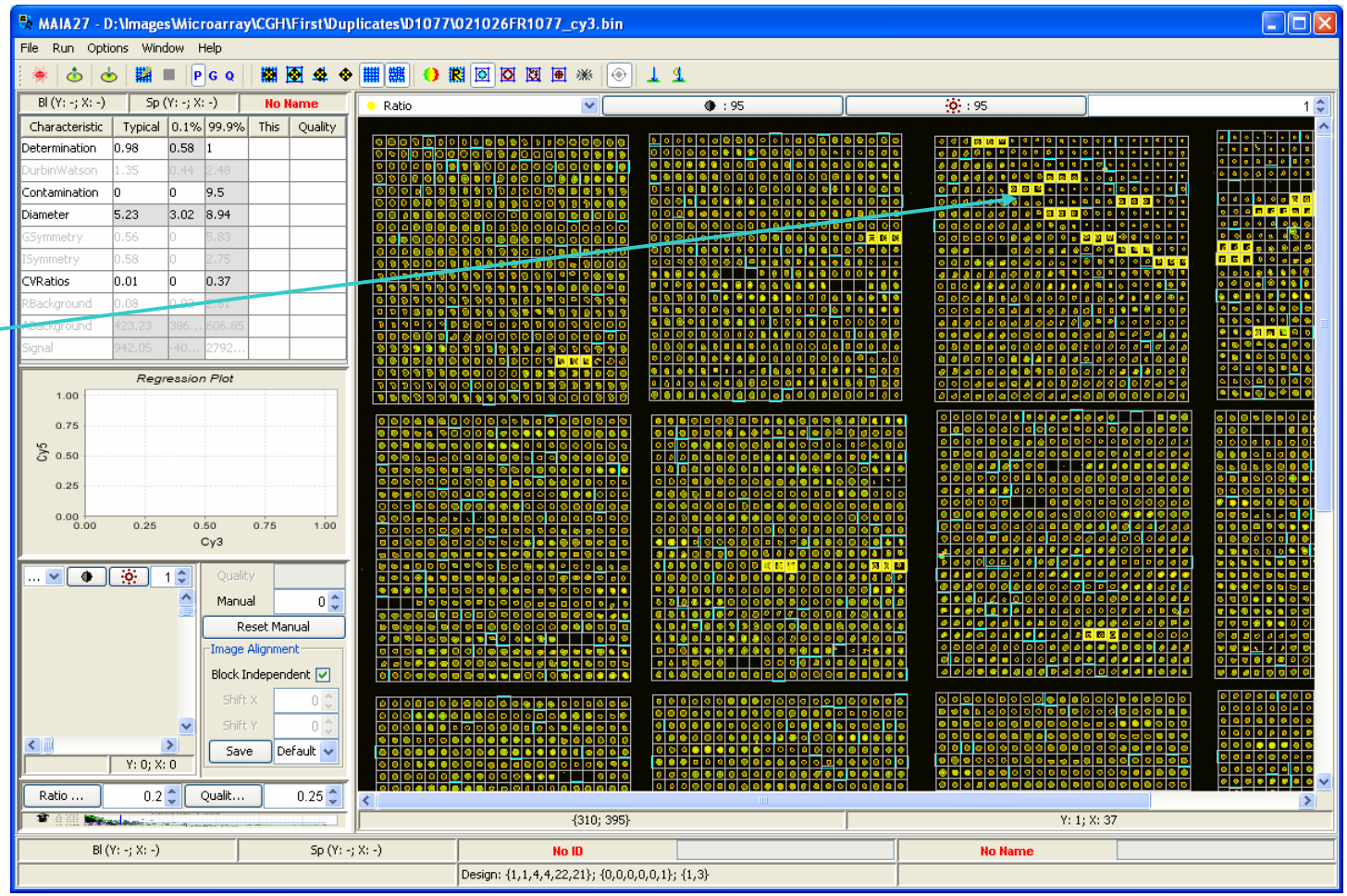
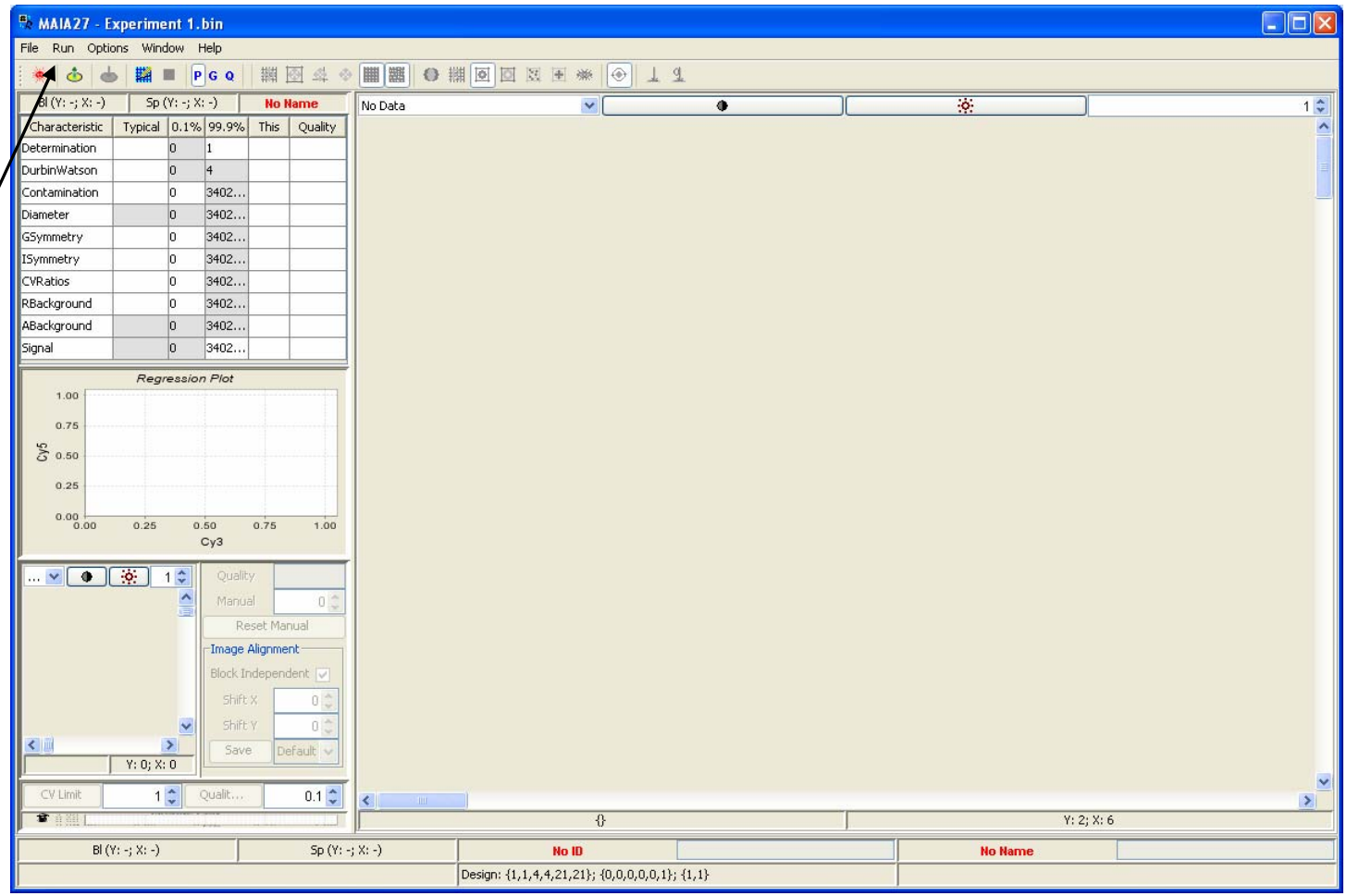


Image Simulator

To open Image Simulator Window select the Menu Item “Run|Simulator”.

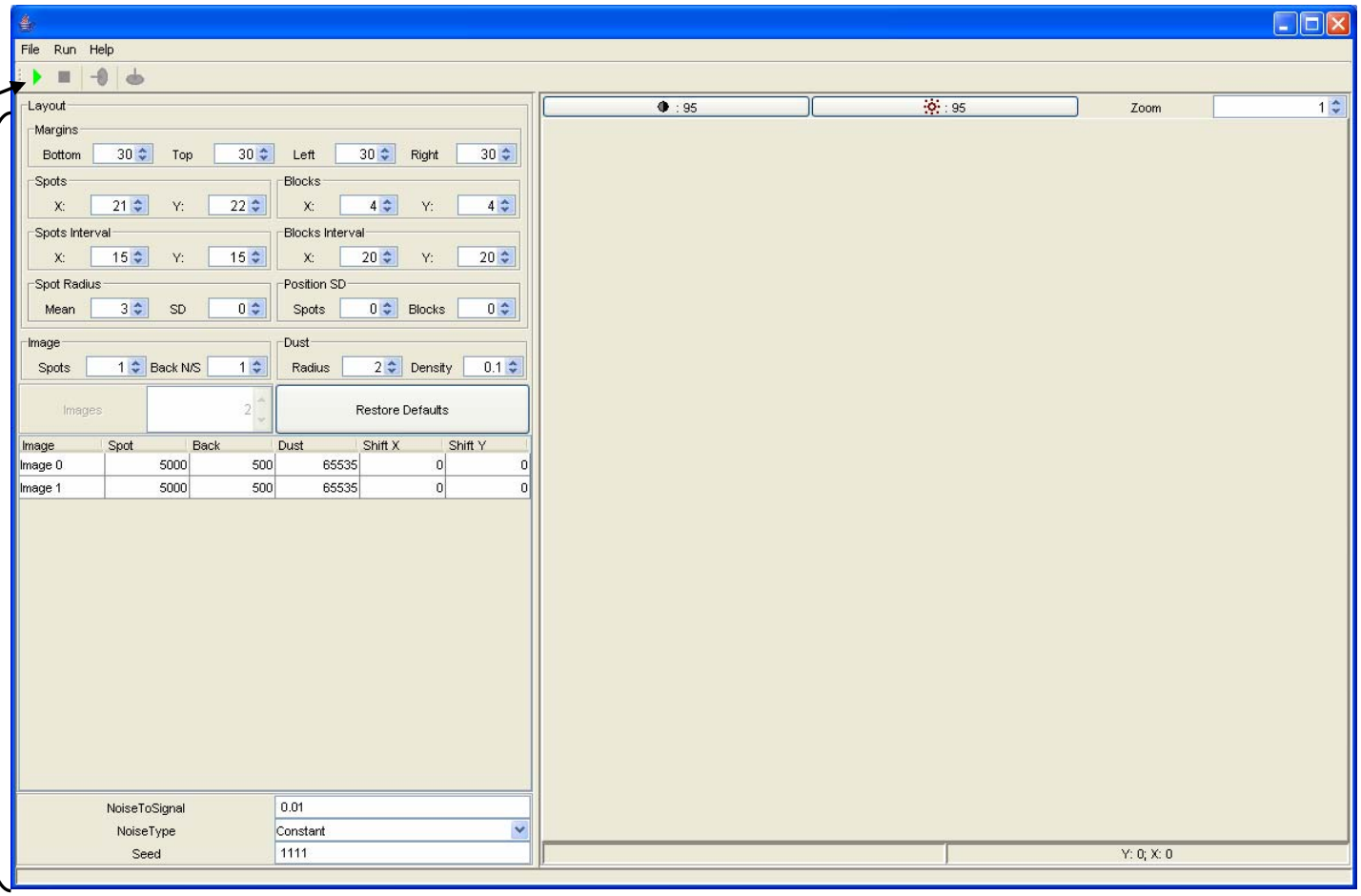


Main Simulator Window

To start simulations press the “Run Simulations” button from the Toolbar or select the Menu Item “File|Run Simulations” (F5).

To simulate an image the following parameters should be defined.

Image 0 stands for Cy3 image.
Image 1 stands for Cy5 image.



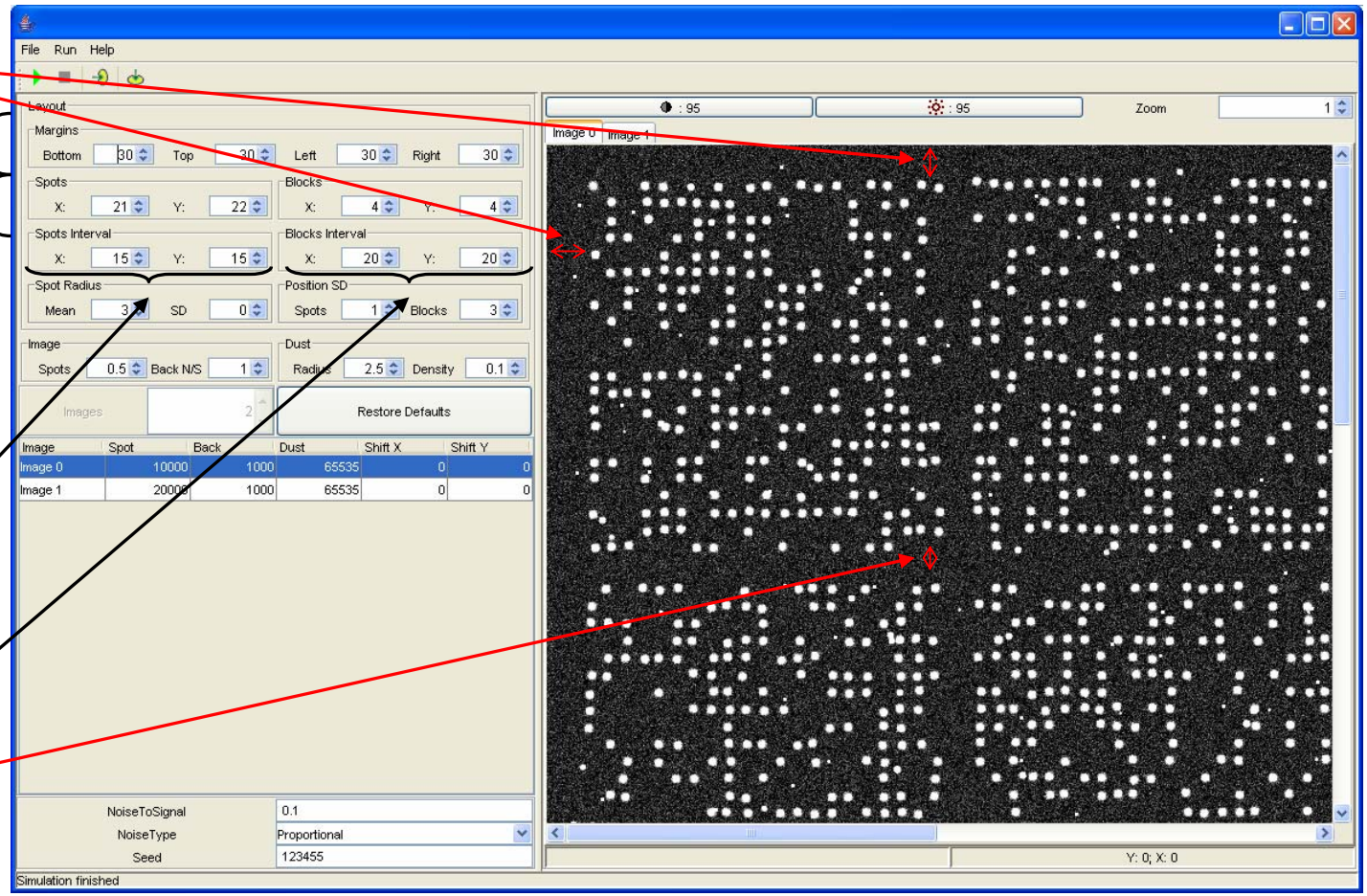
Array Layout (I)

The distance in pixels from the edges of the array to the spotting area.

Amount of spots (per block) and amount of blocks on the array.

Distance in pixels between the spots in the blocks.

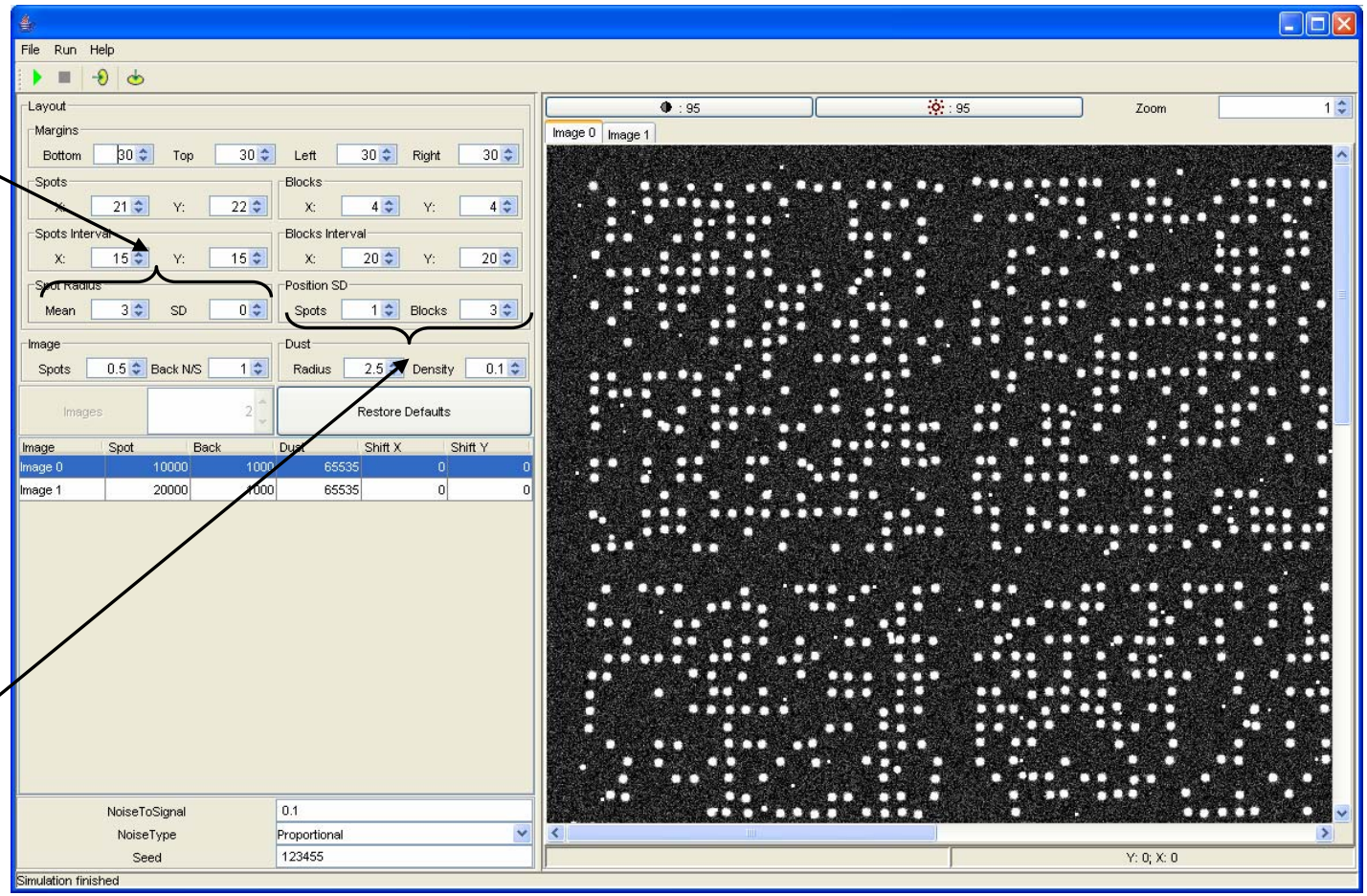
Distance in pixels between the blocks.



Array Layout (II)

Mean and standard deviation of the Spot Radius.

If *SD*>0, spots will be generated with randomly selected (around Mean) radius.



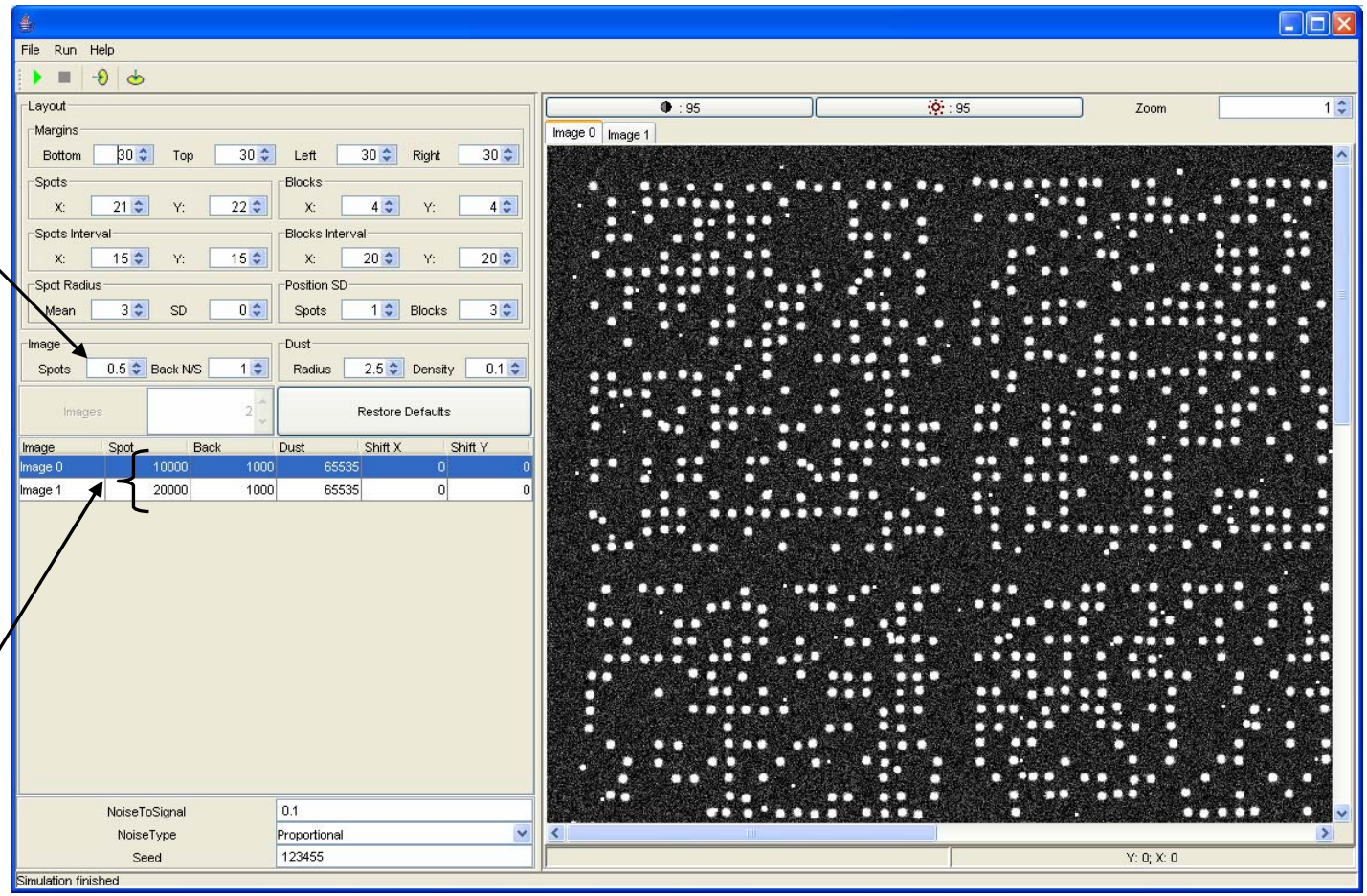
Standard deviation of the positions of the spots and blocks with respect to the ideal alignment.

Larger *SD* value, larger deviation of the positions of the spots/blocks from the ideal spot/block alignment.

Spot Characteristics

Rate of the bright (visible) spots on the array: 0 – no visible spots are generated, 1 – all spots are visible.

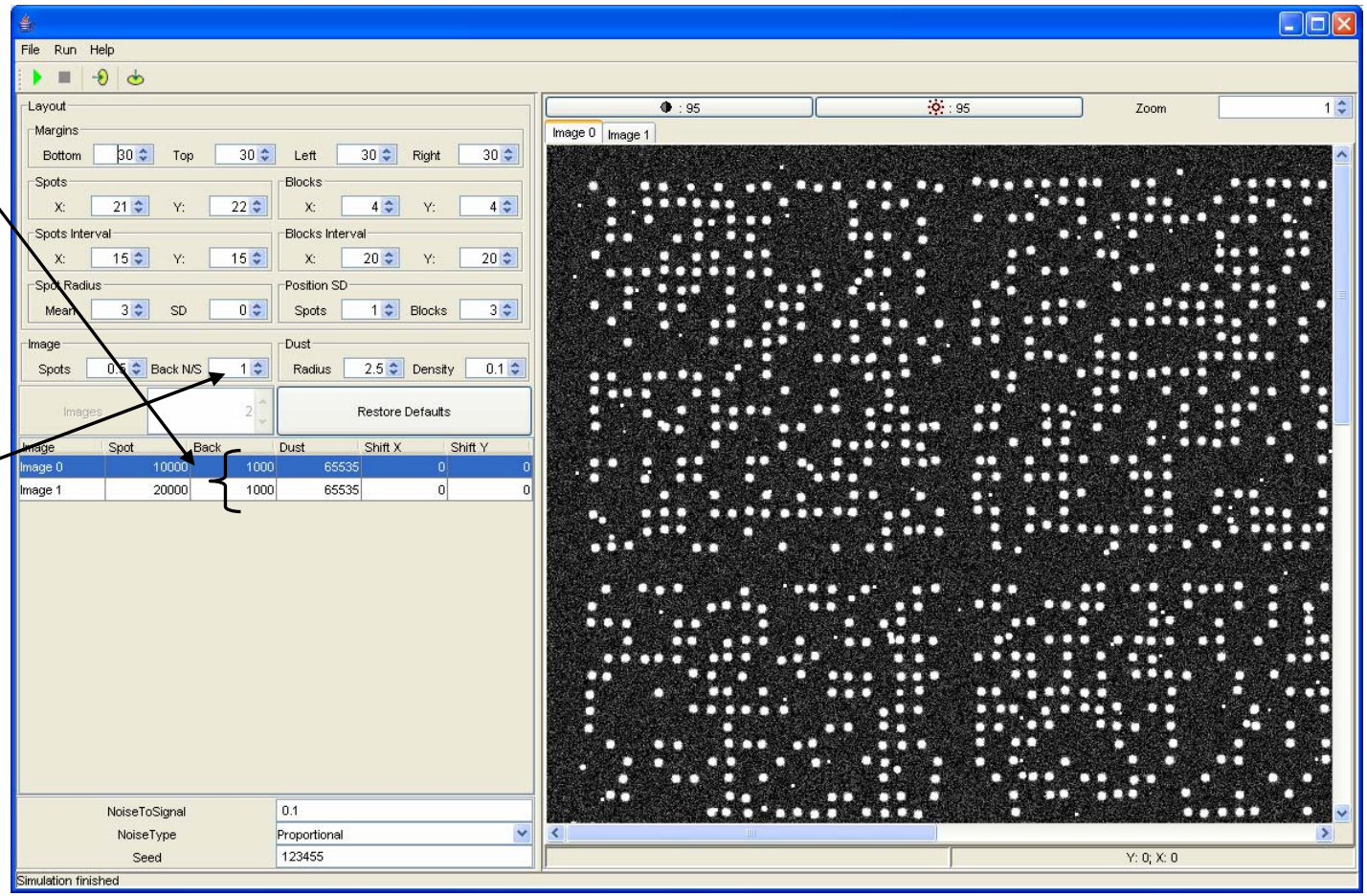
Maximal spot intensity in the Cy3 and Cy5 color channels (i.e. fluorescence intensity in the center of the spot).



Non-Specific Hybridization

Average intensity of non-specific hybridization in the Cy3 and Cy5 color channels.

Noise to signal ratio for non-specific hybridization for both color channels.



Dust

Density of dust is defined with respect to the number of “good” spots on the array:
0 – no dust spots, 1 – the number of dust spots equals to the number of “good” spots.

Maximal dust radius.
The radius of the dust spot is randomly chosen from the interval from 0 to the given value.

Maximal intensity* of dust in the Cy3 and Cy5 color channels.

* Real intensity will be randomly chosen from the interval from 0 to the given value.

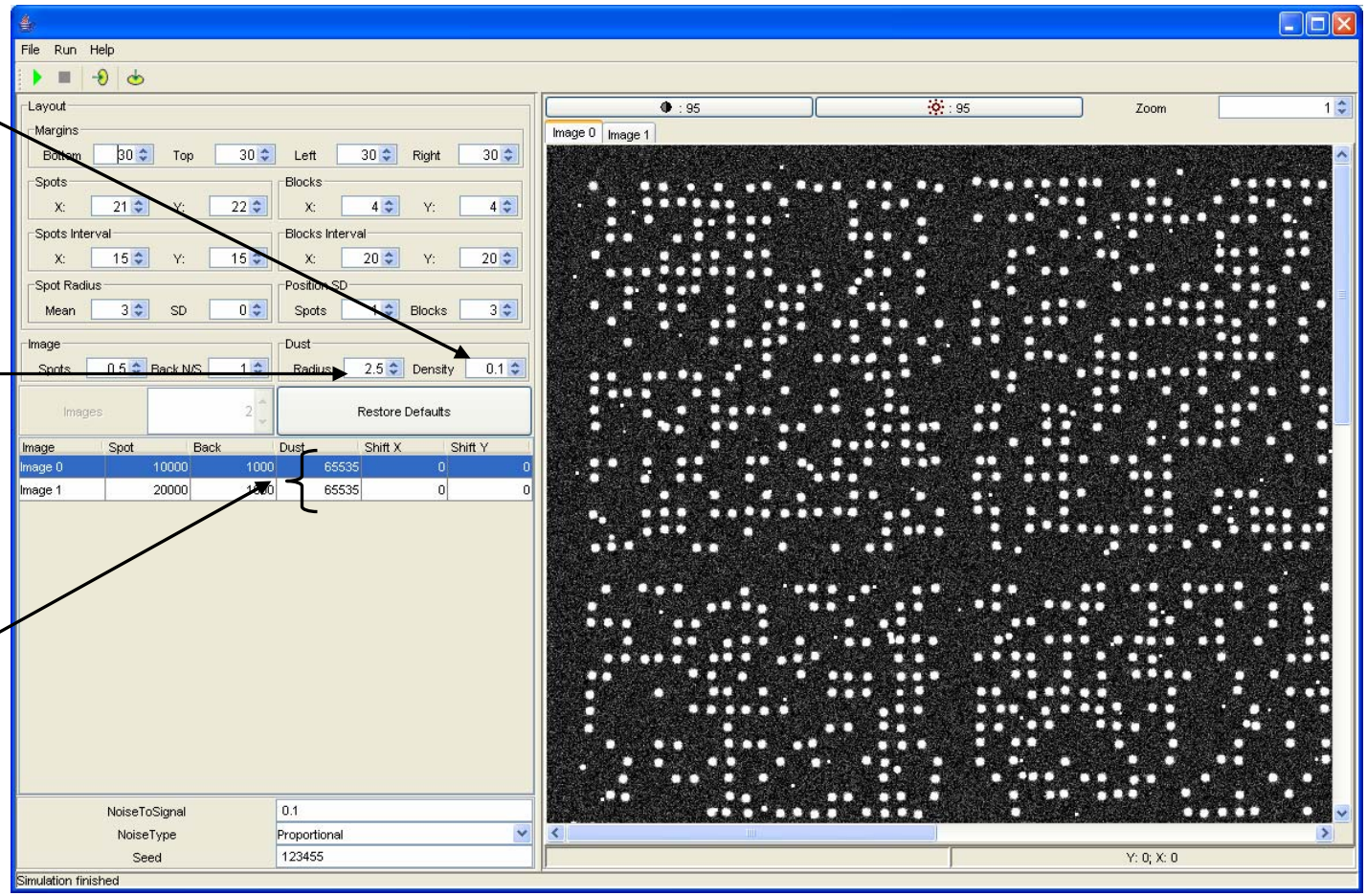
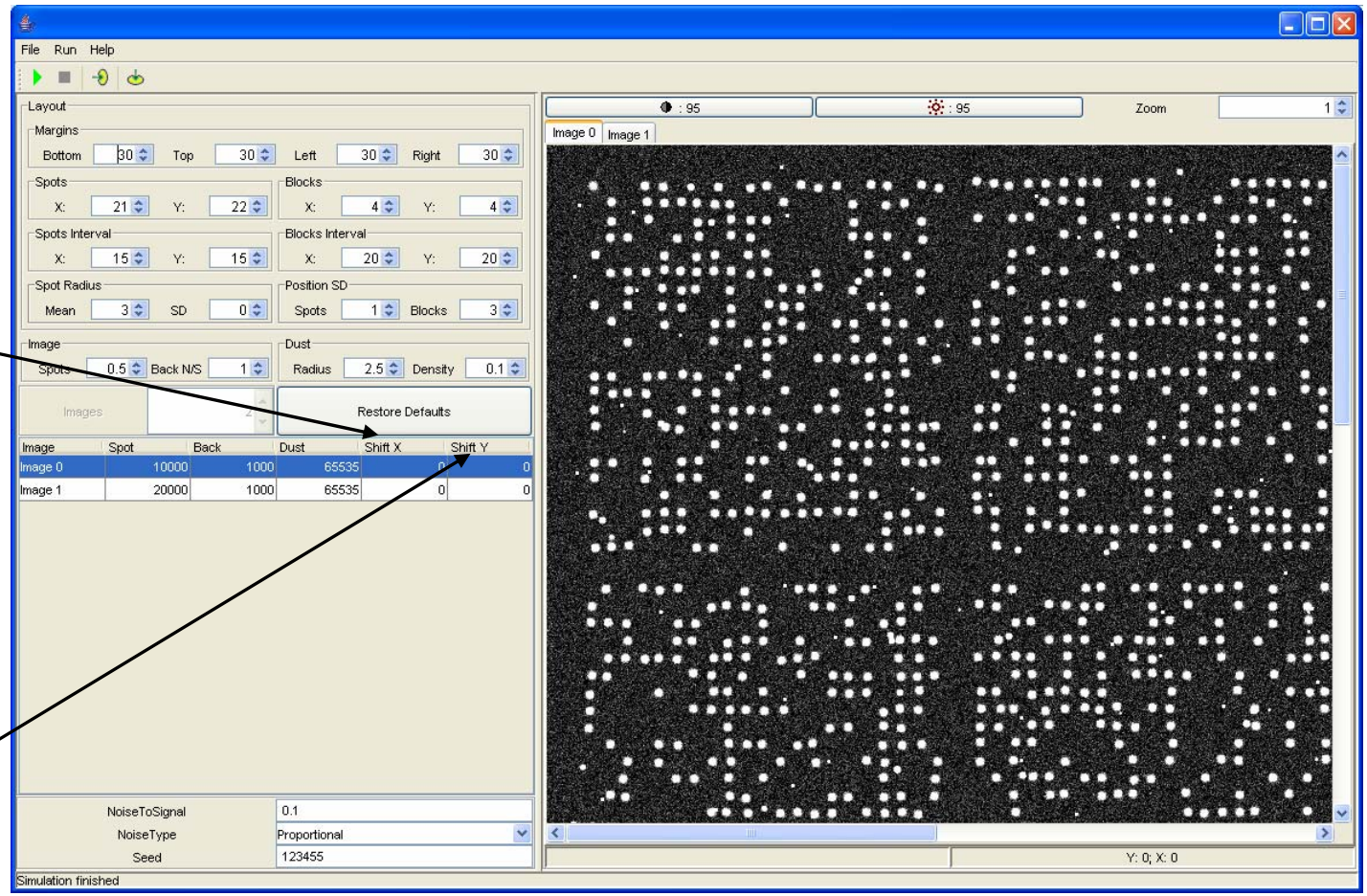


Image Shift

Image shift in pixels in horizontal direction.

Image shift in pixels in vertical direction.



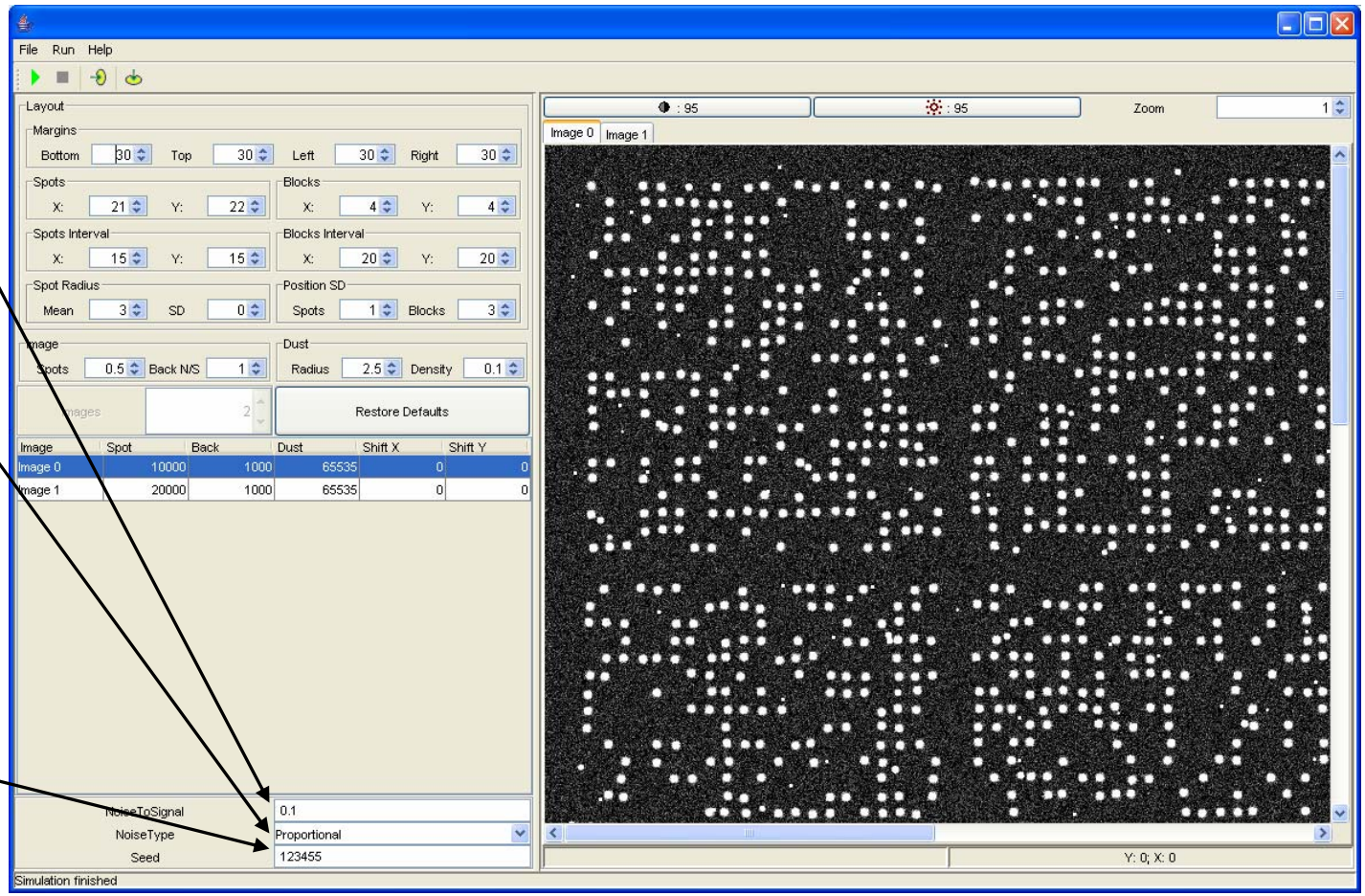
Non-integer pixel shifts are possible.

Additive Statistical Noise

Noise to signal level for the additive statistical noise. This noise is finally added to each pixel of the array.

Model for the standard deviation of the additive noise. It can be constant, proportional to signal, or proportional to the square root of signal.

Seed for random number generator (selection –1 as a seed will initiate the random generator with automatically (or randomly) chosen seed).

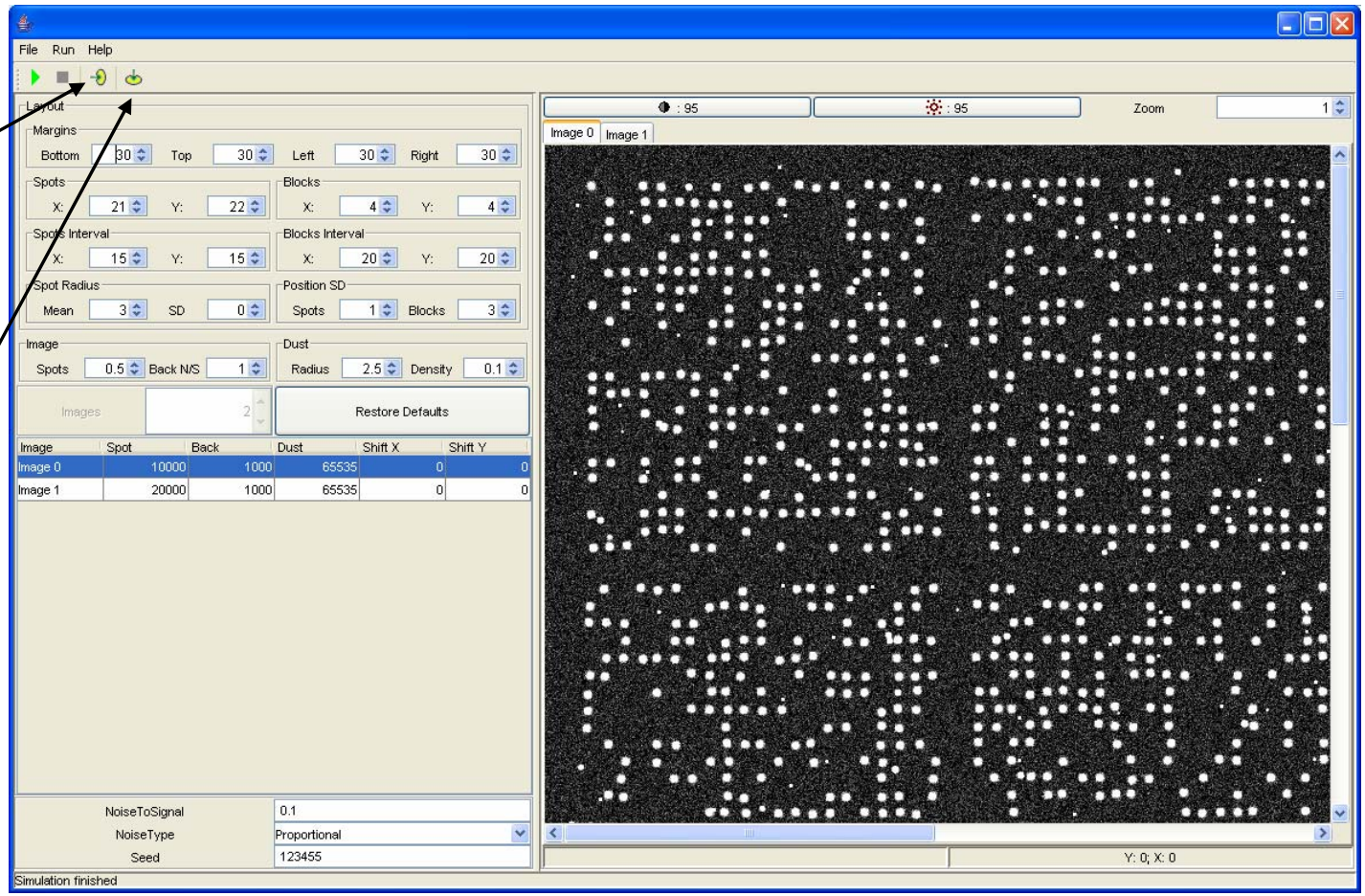


Export of the Generated Image

To send the generated images in the Main Processing Window, use the “Send Data” button from the Toolbar or the Menu Item “File|Send Data” (Alt+→).

To save the generated images in the TIFF files use the “Export Image” button from the Toolbar or the Menu Item “File|Export Image” (Alt+↓).

Only single-page TIFF files are currently supported.



*Artificial Images**

Model for a spot. The generated spots must have more or less circular contours (in the horizontal projection) and relatively sharp edges (in the vertical projection):

$$f_{Cy3}(x,y)=I\exp\left(-\left\{\left[\frac{x-x_c}{r}\right]^4-\left[\frac{y-y_c}{r}\right]^4-\left[\frac{x-x_c}{r}\right]^2\left[\frac{y-y_c}{r}\right]^2\right\}\Big/2\right)$$

where x_c and y_c are the coordinates of the center of the spot, r is its approximate radius and I is the fluorescence intensity in the center of the spot in the Cy3 color channel. Fluorescence intensity in the Cy5 color channel is defined as:

$$f_{Cy5}(x,y)=Rf_{Cy3}(x,y)$$

where R is the ratio of the test and control samples for each spot. The coordinates x_c and y_c, the radius r and the ranges for x and y for each spot cell are defined from the user-established array layout. The intensity parameters I and R should also be provided by the user.

Nonspecific hybridization results in an additional component (B_i) in the detected fluorescence intensity:

$$f_i^B(x,y)=f_i(x,y)+B_i$$

The number of non-specific molecules contributing into each scanned fluorescence pixel is a random value:

$$B_i=B_i^*+\sigma_{Bi}B_i^*G$$

where B_i^* and σ_{Bi} are the user-defined average and noise-to-signal ratio of nonspecific fluorescence intensity in the color channel i, and G is a gaussian random variable with zero mean and unit standard deviation.

Dust is represented by randomly distributed over the array more or less bright clusters of pixels, which can hardly be distinguished from the spots. We apply the same profile for the dust clusters as for the spots:

$$d_i(x,y)=I_d\exp\left(-\left\{\left[\frac{x-x_{cd}}{r_d}\right]^4-\left[\frac{y-y_{cd}}{r_d}\right]^4-\left[\frac{x-x_{cd}}{r_d}\right]^2\left[\frac{y-y_{cd}}{r_d}\right]^2\right\}\Big/2\right)$$

where x_{cd} and y_{cd} are the coordinates of the center of a dust cluster, r_d is its approximate radius and I_d is the intensity in the center of the cluster. All these parameters are random variables. We use uniform distributions for r_d (in the interval $[0;r_m]$) and I_d (in the interval $[0;I_m]$), where r_m and I_m are user-provided maximal dust cluster radius and maximal dust intensity, respectively. We also assume that the coordinates of the centers of dust clusters x_{cd} and y_{cd} are uniformly distributed over the array. Statistical laws of the dust characteristics can generally be different for two channels (i = Cy3, Cy5). Finally one has to define the number or density of the dust clusters on the array.

The general model for the microarray image takes the form:

$$\bar{f}_i(x,y)=\sum_{k=1}^{N}f_{ik}(x,y)+B_i+\sum_{k=1}^{M}d_{ik}(x,y)$$

where N is the number of spots and M is the number of dust clusters.

Statistical noise is finally added to each pixel of the image:

$$\tilde{f}_i(x,y)=\bar{f}_i(x,y)+\sigma(x,y)G$$

where $\sigma(x,y)$ is the standard deviation of the pixel noise: $\sigma(x,y)$ can be (i) constant, (ii) proportional to signal, or (iii) proportional to the square root of signal. The type of statistical noise as well as its quantitative characteristics is defined by the user.

*) E. Novikov and E. Barillot, A robust algorithm for ratio estimation in two-color microarray experiments. *Journal of Bioinformatics and Computational Biology*, 2005, 3, 1411-1428.

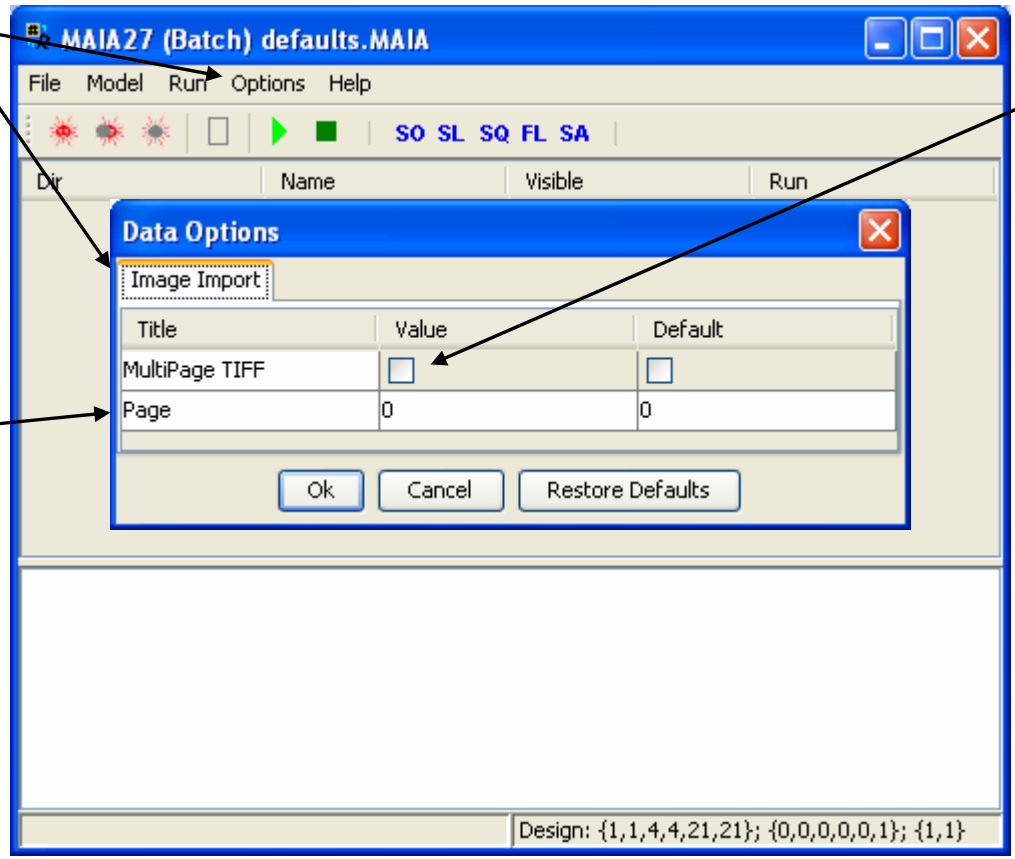
One Color Image Analysis

Very much similar to *Two Color Image Analysis*

Data Import Settings

To define the format of the microarray image files select the Menu Item “Options|Data Options” (Alt+D).

For multi-page TIFF, specify the page for the image to be analyzed.

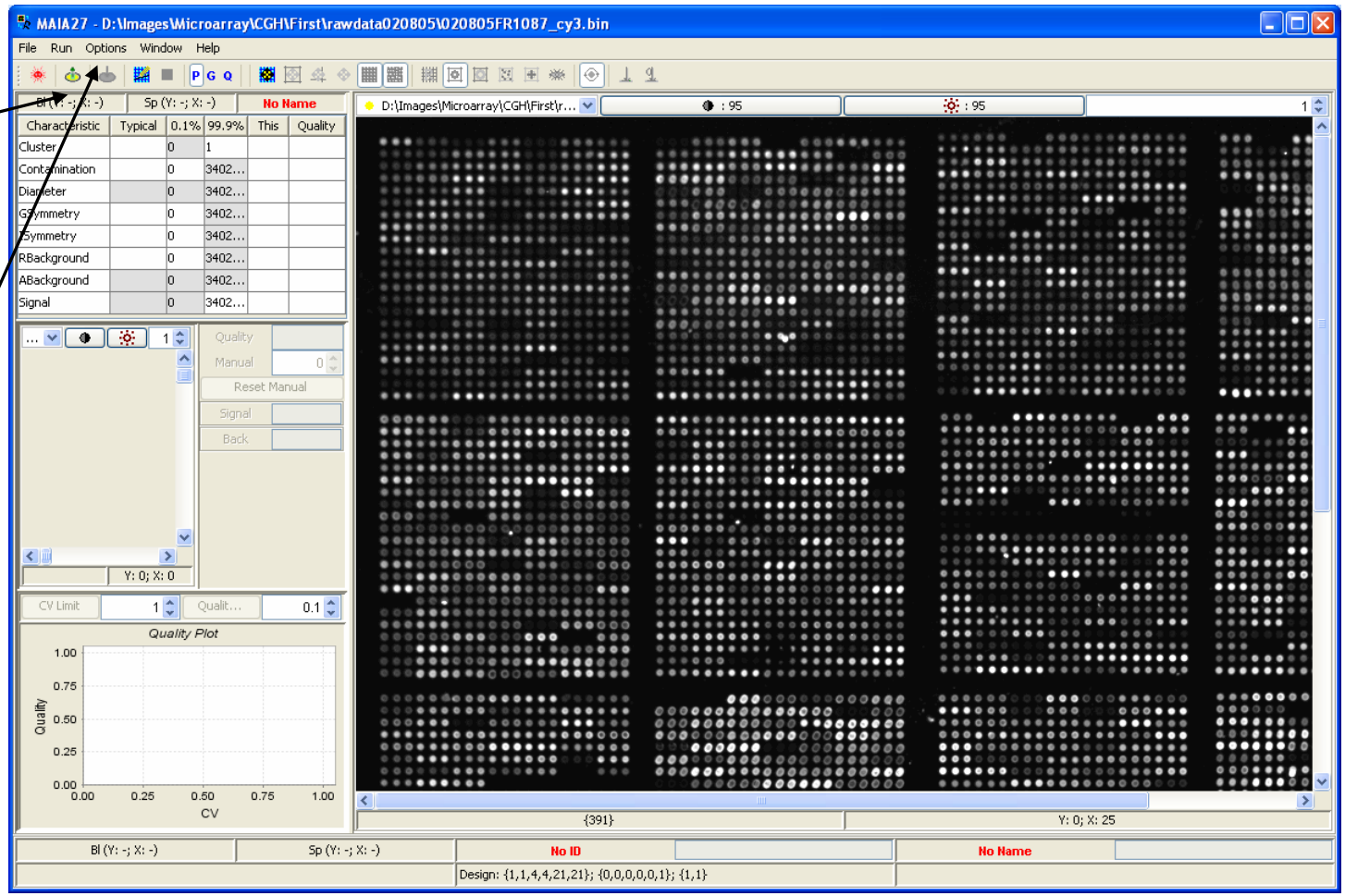


Two options are available:
(i) TIFF images are packed into one multi-page TIFF file (checked);
(ii) TIFF images are stored in separate files (unchecked).

Main Processing Window

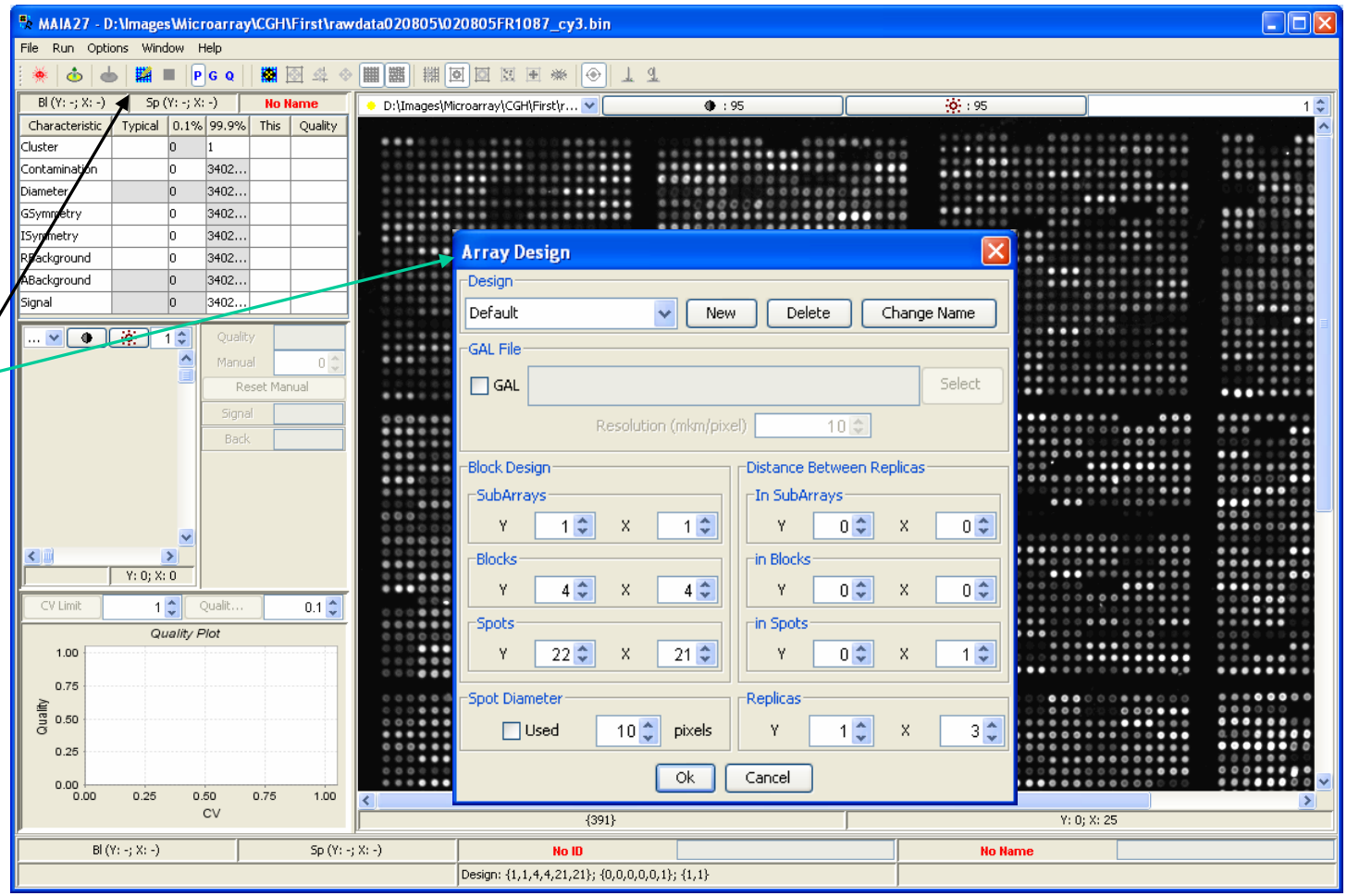
Another image can be downloaded using the "Load Data …" button from the Toolbar or the Menu Item "File|Load|Data …" (Ctrl+O).

For the new images, image file format (i.e. multi-page TIFF versus single-page TIFF) can be changed using the Menu Item "Options|Data Options" (Alt+D).



Array Design

Use the “Array Design” button from the Toolbar or select the Menu Item “Options|Array Design” (Alt+A).

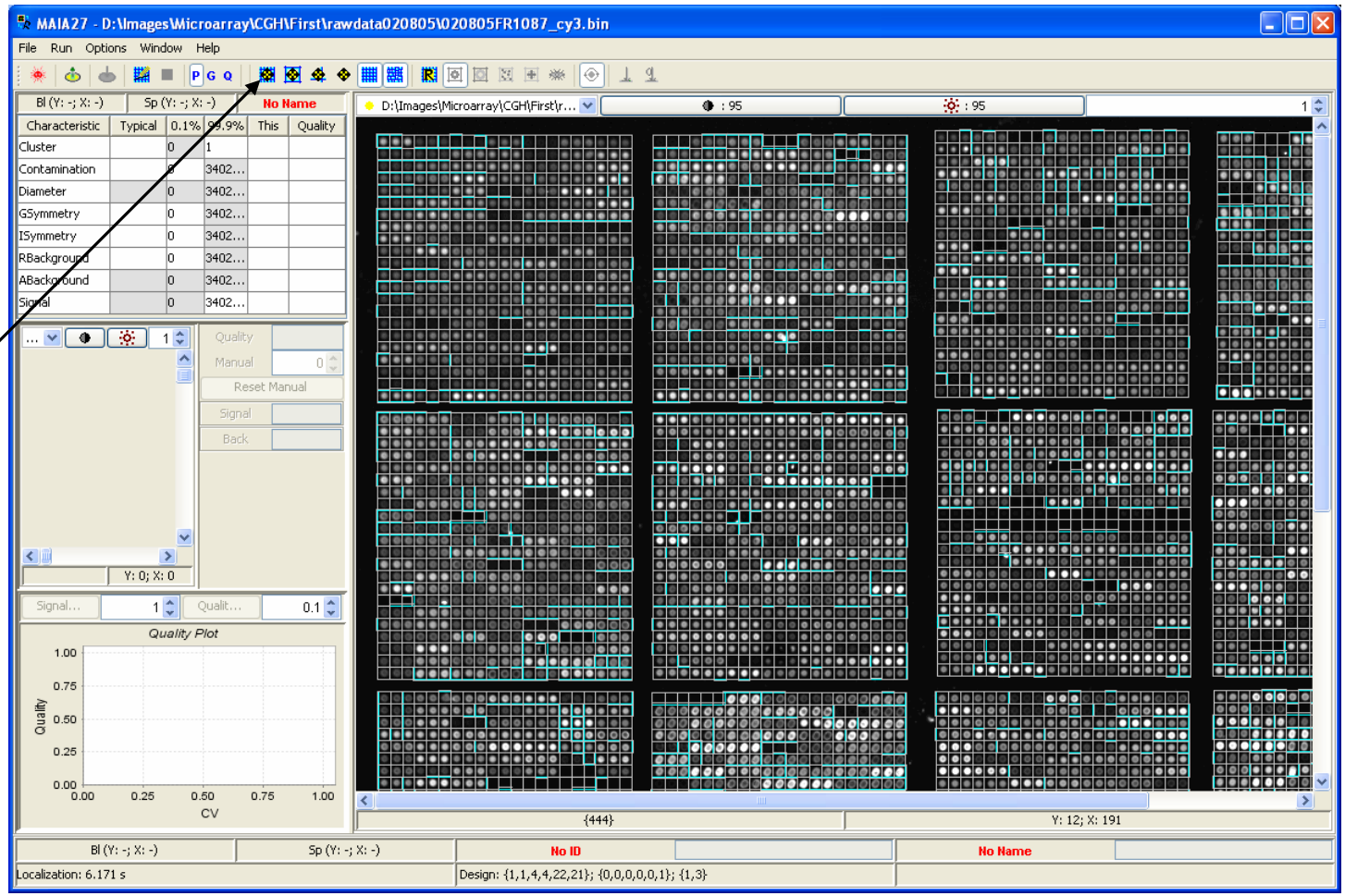


Array Design is equivalent to the *Two Color Image Case.*

Spot Localization

To start Spot Localization (or grid finding) use the "Spot Localization" button from the Toolbar or select the Menu Item "Run|Spot Localization" (Ctrl+F6).

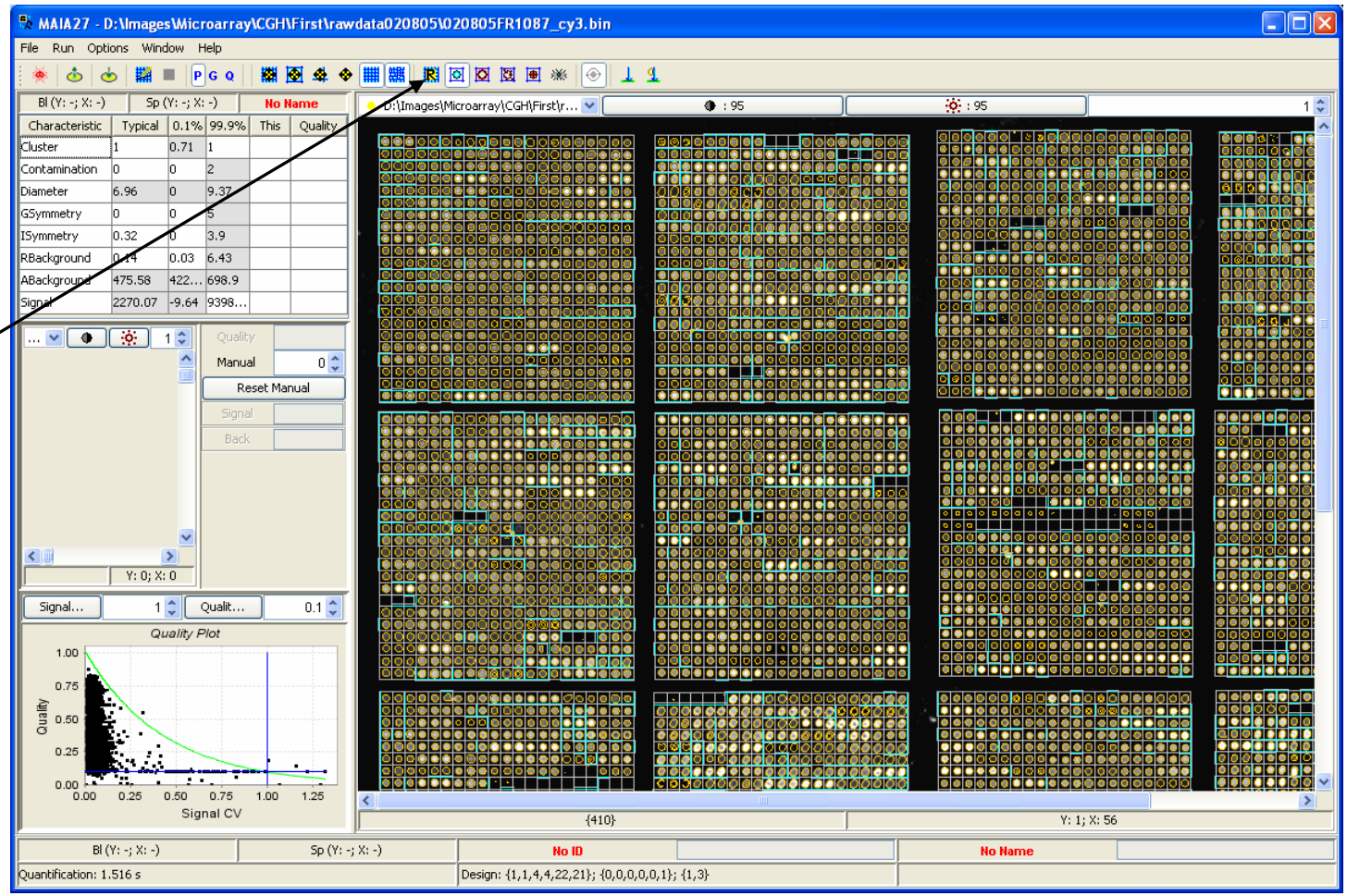
All possibilities for grid management are equivalent to the *Two Color Image Case*.



Spot Quantification

To start Spot Quantification use the "Spot Quantification" button from the Toolbar or the Menu Item "Run|Spot Quantification" (Ctrl+F8).

Note that the "Image Alignment" button from the Toolbar as well as the Menu Item "Run|Image Alignment" (Ctrl+F7) do not show up for One Color Image Analysis.



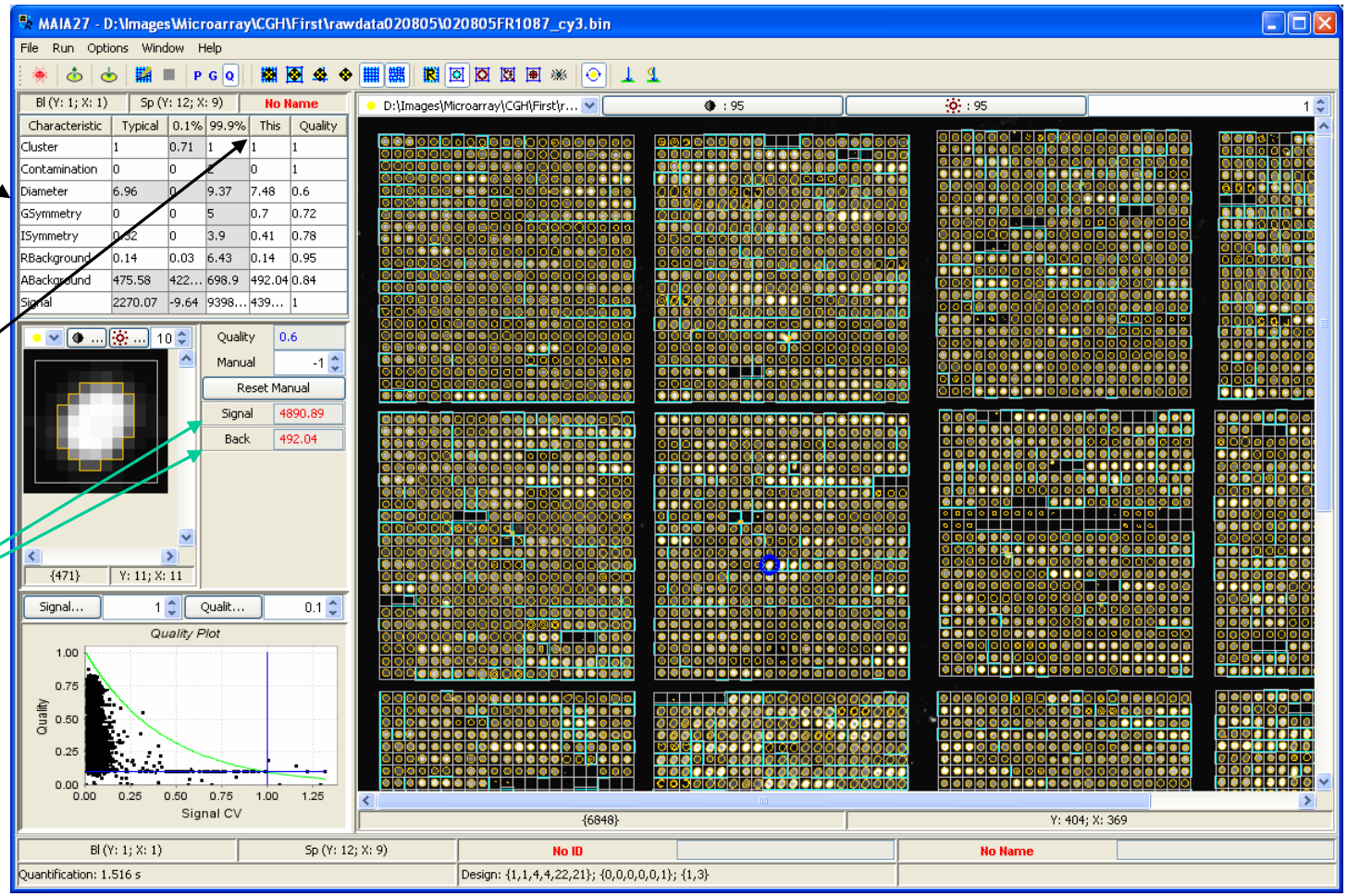
Spot Quantification Output

Quality characteristics of the spots.

See next page for details.

Quality characteristics of the selected spot.

Estimation of the signal and background.



Quality Characteristics

Cluster (*C*) is the ratio of the diameter of the largest cluster of bright pixels of the spot to the diameter of the spot. For low intensity spots, segmentation procedure may identify many non-intersecting pixel clusters with the average intensity somewhat higher than the background level. The parameter *C* is expected to be low for such spots. $q(C) = C^*$.

Spot contamination is a number of out-ranged pixels (with the intensity equal to 2^{16}-1) (*N*). $q(N) = 1-N/S$, where *S* is the size of the correspondent spot, i.e. the number of pixels within the spot contour*.

Diameter of the spot: $D = 2(S/\pi)^{1/2}$. Since it is hard to impose *a priory* an exact ideal value for the diameter, the median diameter over all spots on the array is taken as a typical one. Spots with exceptionally small or large diameters should normally be penalized. $q(D) = exp\{T_D-D\}$, if $D>T_D$ and $q(D) = exp\{T_D-D\}$, if $D<T_D$ where T_D is the typical diameter*.

Geometrical symmetry parameter measures deviation of the contoured spot from the ideal circle. Both the real spot and the ideal circle are divided into 8 sectors (pie slices defined as $[k\pi/4;(k+1)\pi/4]$, $k = 0,\ldots,7$) and for each sector the number of pixels belonging to the spot (N_{si}, $i = 1,\ldots,8$) and to the circle (N_{ci}, $i = 1,\ldots,8$) is counted. Then the quality characteristic is defined as $GS = \sum|N_{si}- N_{ci}|/ N_{ci}$. For ideal circular spots *GS* must approach 0, whereas highly un-circular spots should give relatively high *GS* values. $q(GS) = exp(-GS)^*$.

Intensity symmetry of the spot is defined as $IS = \sum|I_i-I|/I$, where I_i, $i = 1,\ldots,8$ are the mean intensities for the same 8 sectors and *I* is the mean intensity for the whole spot A spot may have perfect circular shape, but within this circle very bright (or dark) and highly concentrated groups of pixels originated from the pieces of dust or other contamination may occur. $q(IS) = exp(-IS)^*$.

Uniformity of the background around the spot, i.e. along the grid lines separating neighborhood spots, is defined as $UB = \sum|B_i-B|/B$, where B_i, $i = 1,\ldots,8$ are the mean intensities in 8 sectors of the grid line around the spot, and *B* is the mean intensity for the whole grid line around the spot. Extremely small values may be due to relatively bright contamination around the spot, large variability in the background or merged neighborhood spots. $q(UB) = exp(-UB)^*$.

Absolute level of background (*AB*) calculated in the proximity of each particular spot is compared to the typical level of the local background for a given array. Large deviations from the typical state may indicate the presence of the contamination areas, which are larger than the size of the spot. $q(AB) = exp(1-AB/T_{AB})$, if $AB>T_{AB}$ and $q(AB) = exp(AB/T_{AB}-1)$, if $AB<T_{AB}$. where T_{AB} is the typical background level*.

Signal (*S*) is defined as a difference between the mean estimate of the intensity within the contoured spot and the mean estimate of the background level. $q(S) = 1$, if $S>T_S$ and $q(S) = exp(S/T_S-1)$, if $S<T_S$. where T_S is the typical signal*.

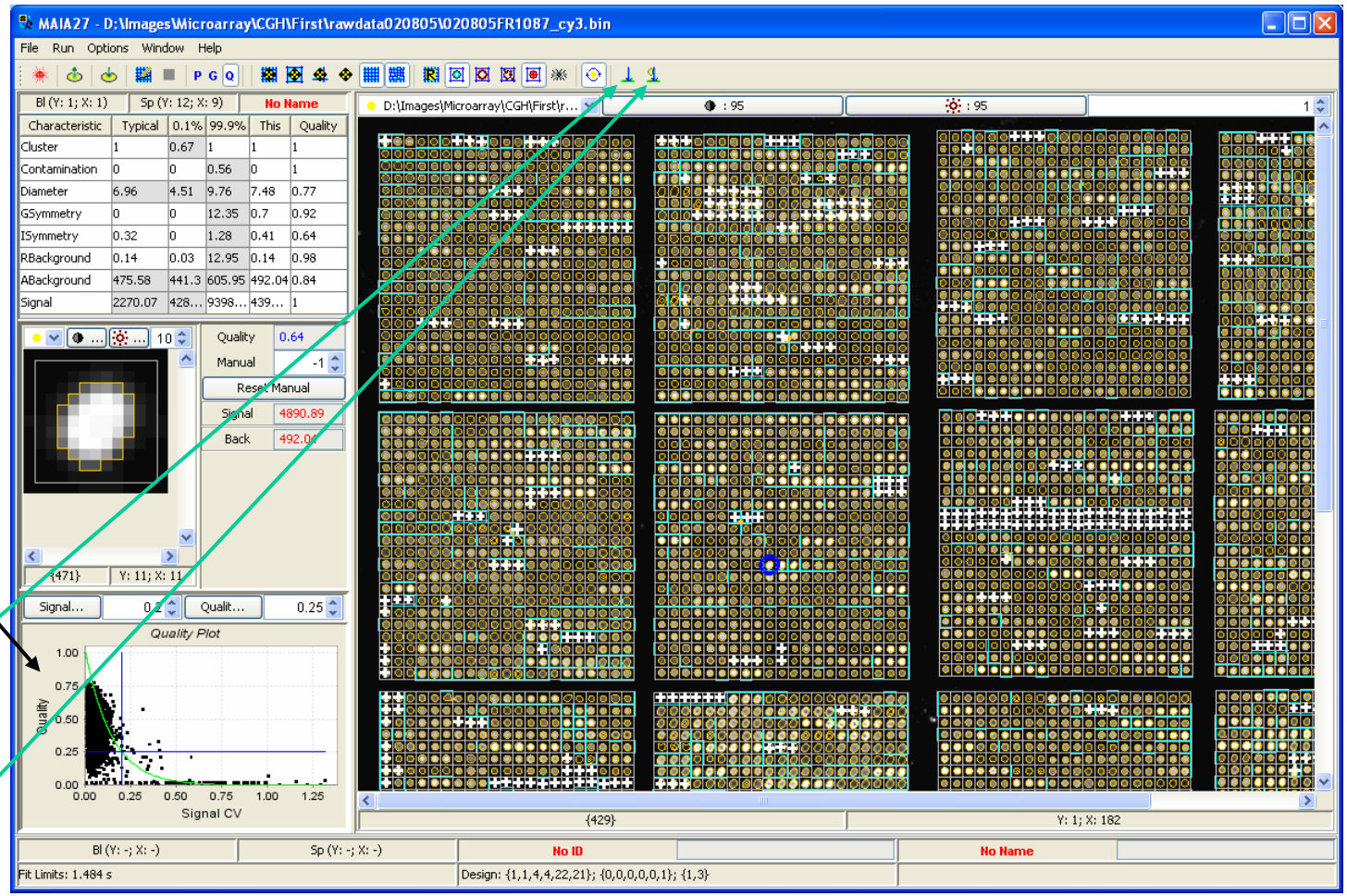
**For the purposes of further quality analysis, functions q, rescale quality characteristics to fit the range between 0 (“bad” spot) and 1 (“good” spot).*

Quality Analysis

Quality analysis is equivalent to the *Two Color Image Case.*

Quality plot:
y-axis is the overall quality value;
x-axis is the coefficient of variation (CV) of the **signal** of the replicates.

The quality limits are initialized using the "Init Limits" button from the Toolbar or the Menu Item "Run|Init Limits" (Ctrl+F9). Then they are fitted using the "Fit Limits" button from the Toolbar or the Menu Item "Run|Fit Limits" (Ctrl+F10).



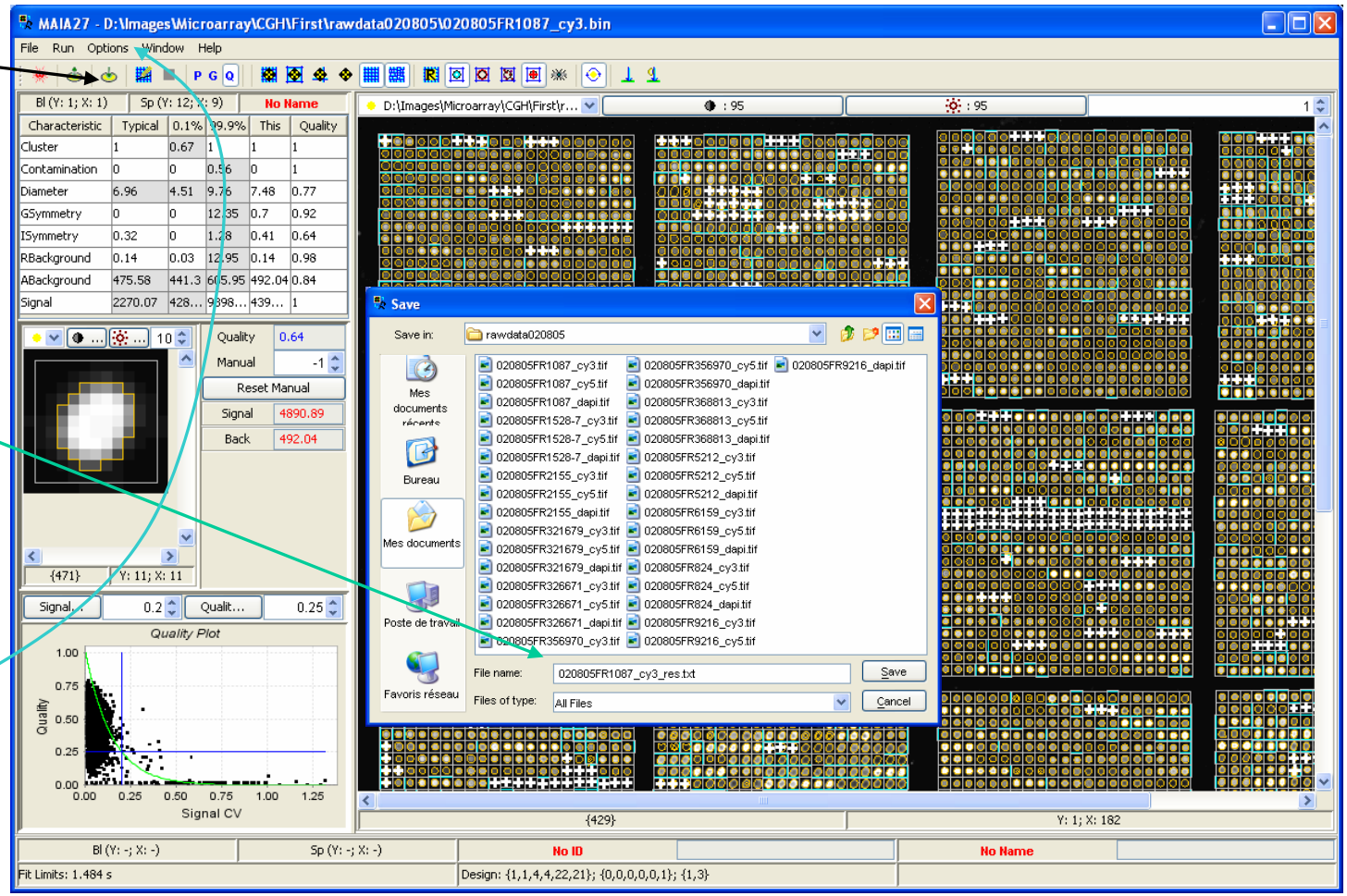
Save the Results

To save the results of quantification and quality analysis use the “Save Analysis ...” button from the Toolbar or the Menu Item “File|Save|Analysis …” (Ctrl+S).

The results are saved as a table in the text file (importable into Microsoft Excel).

User can define which fields and in which order should be presented in the output file: select the Menu Item “Options|Output Format” (Alt+F).

See next page for details.



Output Table Format

Description of the field (non-editable).

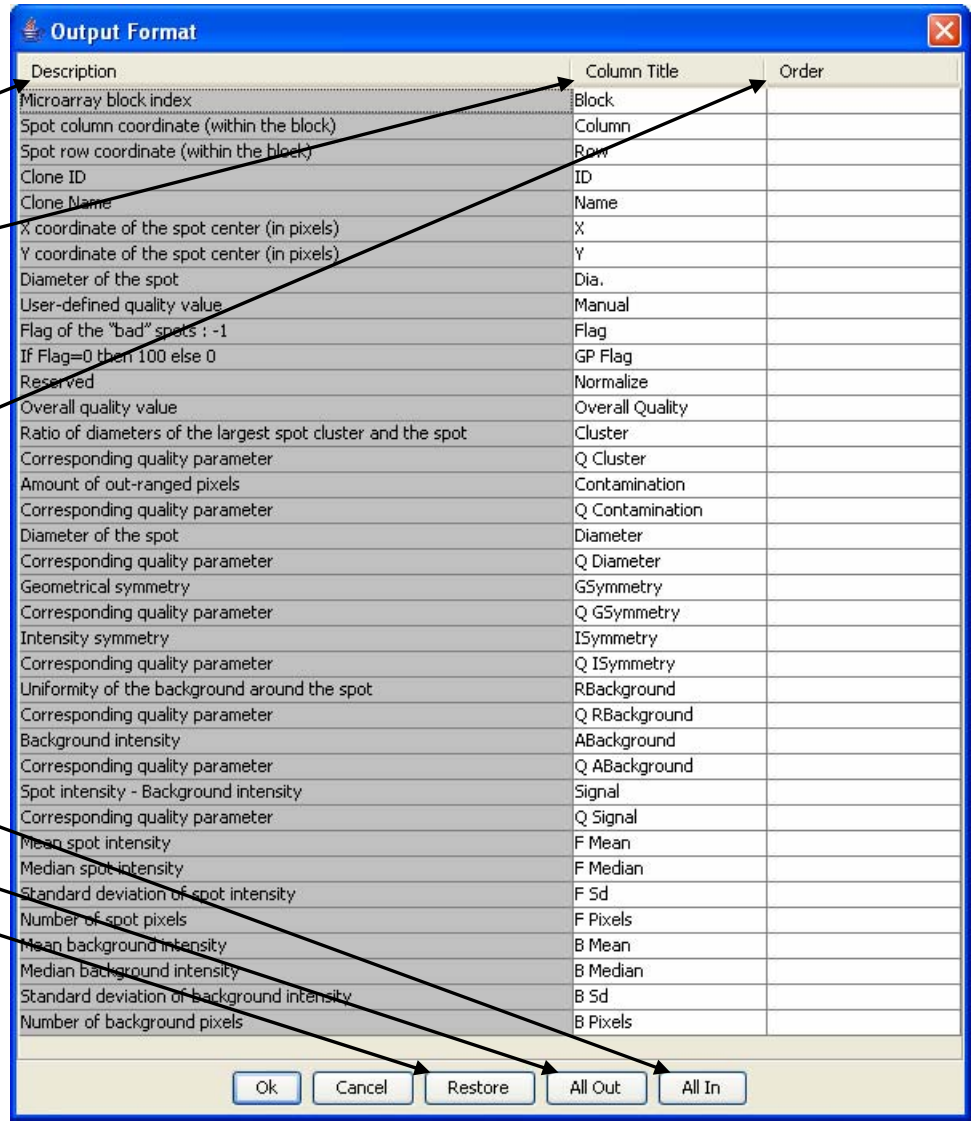
Editable name of the field to be appeared in the output file.

Order specifies the sequence of the fields. If this field is empty, the corresponding field is not included in the output file.

Include all fields.

Exclude all fields.

Restore previous set of fields.

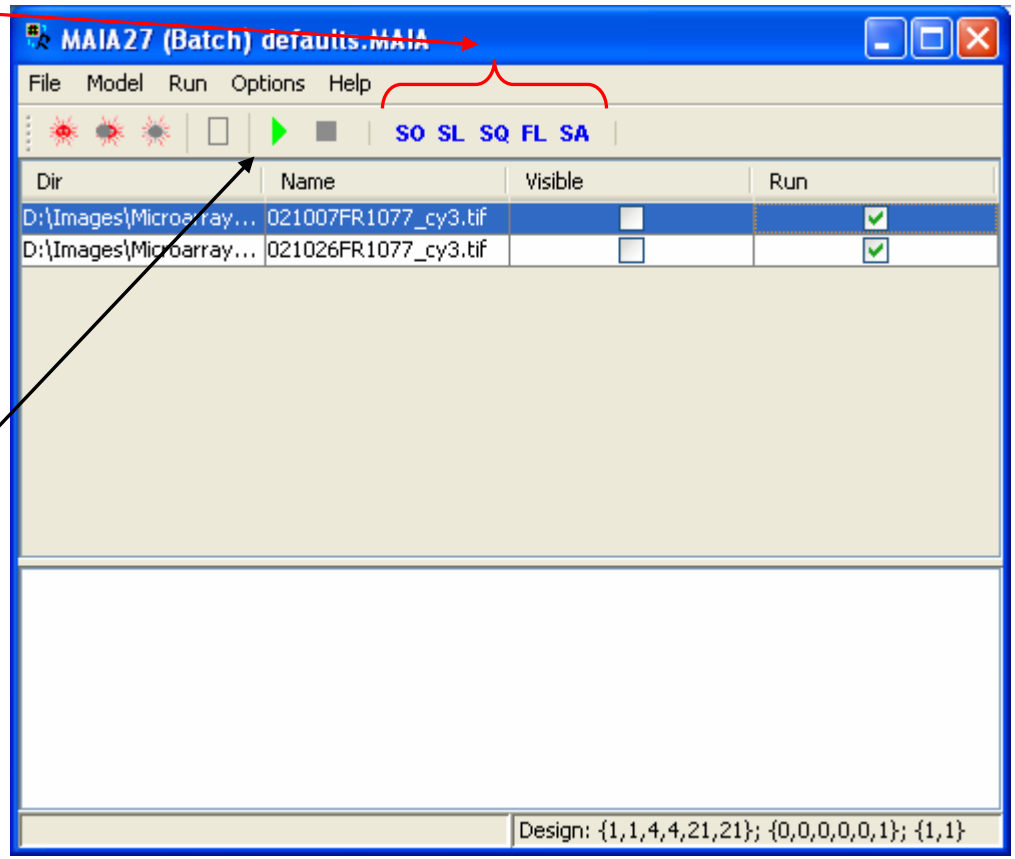


Batch Processing and Global Quality Analysis

To run batch processing one have to define actions to be applied to all files in the batch:

SO – Set Options;
SL – Spot Localization;
SQ – Spot Quantification;
FL – Fit Limits;
SA – Save Analysis;

The batch processing can be started using the "Run Batch" button from the Toolbar or the Menu Item "Run| Run Batch" (F5).



Batch Processing and Global Quality Analysis are equivalent to the *Two Color Image Case.*

Image Simulator

The parameters of the Image Simulator is equivalent to the *Two Color Image Case.*

One image is simulated.