

CytoSure™ Interpret - Software update guide

**Includes update guides 4.0 – 19.0
For research use only**



CytoSure™ Interpret Software update – version 19.0 January 2019

The following new features are available in CytoSure Interpret Software version 4.11:

Case Status	3
Usage	3
Classification History	4
Usage	5
Probe Filter Groups	6
Creating Probe Filter Groups	7
Analysing Cases	7
Updating Annotation tracks	8
Online Track Update	8
Update Tracks during Software Update	9
Workflow	9
Loading cases into the workflow	9
Analysing and reporting cases	10
Batch Processing Experiment Association	10
Usage	10
Miscellaneous Additions to v4.11	11



Case Status

CytoSure Interpret Software version 4.11 now includes a **Case Status** field for samples to enable easier tracking of the progress of a case.

Usage

There are four available statuses that can be used to track the progress of a sample - **To Do**, **In Progress**, **In Review**, and **Completed**. The software also logs the user who updated the status of a sample and when the status was updated. The case status can be used to search for samples in **Database Management** by selecting the status on which to filter from the **Quick Search** panel.

Sample ID	Gender	Case Status
<input type="checkbox"/> Test 036908	Male	In Progress
<input type="checkbox"/> 3f7733c1-4930-4bb...	Male	In Progress
<input type="checkbox"/> 253086411269_1_1	Male	In Progress

Figure 1 Searching for samples

The status for a sample is initially set to **To Do** and will then be automatically updated during analysis of a case. Once segmentation has been run on the sample and the aberrations identified, the status is updated to **In Progress**. When analysing cases using the workflow, selecting **Finish** to complete analysis of the opened cases will provide the option to transition the status to the next stage. If the current status is **In Progress**, then the sample can be set to **In Review** and if the status has already been set to **In Review** then the option is provided to set the status to **Completed**.

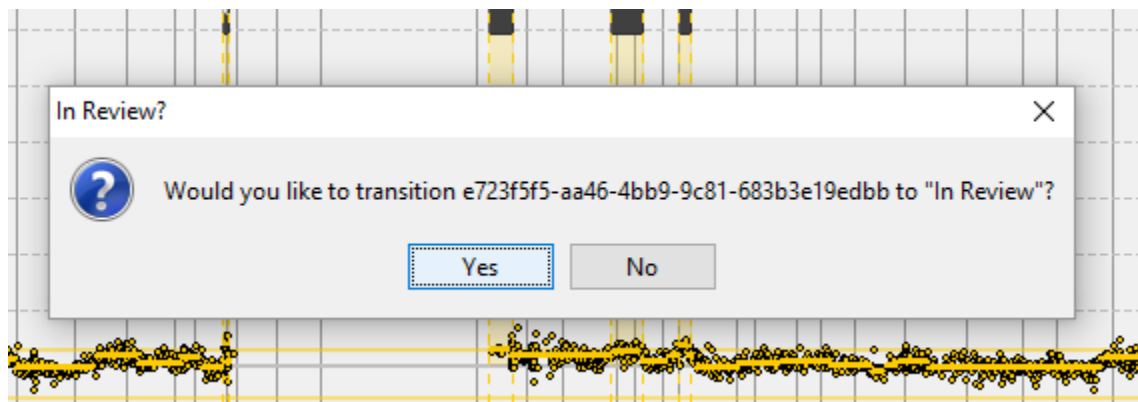


Figure 2 The cases status can be automatically updated

Alternatively, the status of a Sample can be changed via **Database Management** by selecting a sample, going to **Sample Data** → **Basic Information** → **Case Status** and selecting the required status, or via the **Sample Details** window.

Once a case has been marked as completed, the sample is locked and users are prevented from editing the sample or experiment details in **Database Management**. It is still possible to view and generate reports for completed cases but changes cannot be saved to the database and the .cgh file associated with a completed case cannot be modified by the software.

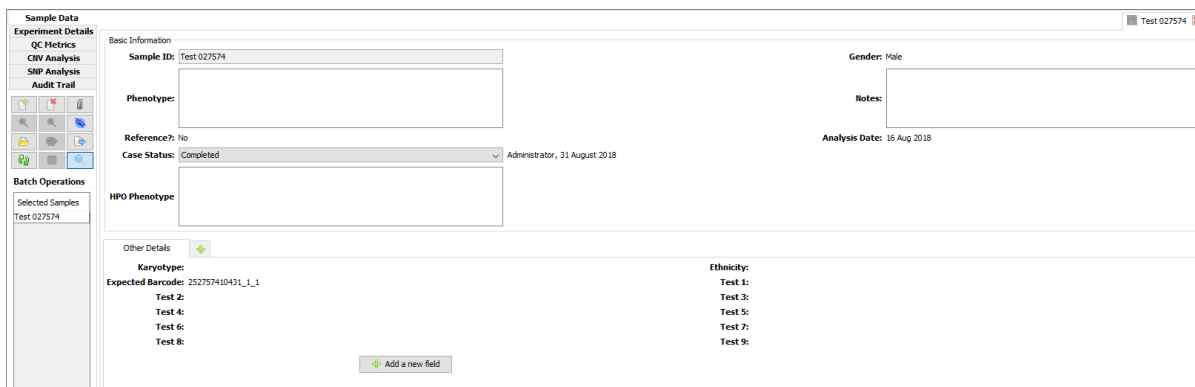



Figure 3 A completed case

 If necessary, a completed case can be re-opened by the user who completed the case or by users above them in the hierarchy.

Classification History

The classification history provides a more flexible approach to classifying CNVs and LOH regions, allowing users to modify the classification of an aberration. Each change to the classification is logged, storing the details of the user and time the aberration was classified.



i The classification history is intended to supersede the existing initial and final classification system.

Usage

To assist switching to using the classification history, it is possible to transfer the initial and final classifications for all CNVs and LOH regions already in the database to the classification history. When an admin user first logs into CytoSure Interpret v4.11, a window is displayed asking whether to transfer the existing classifications across. This window also allows you to choose to hide all references to the initial and final classifications. Selecting to do so prevents users from setting the initial or final classifications for CNVs or LOH regions and will prevent displaying or reporting these classifications. Instead, users will only be able to update the classification and classification history.

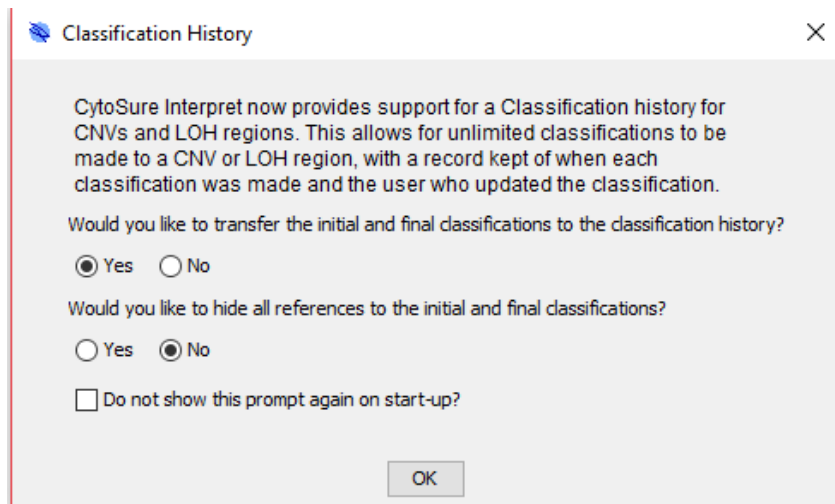


Figure 4 Classification History window

To continue using the initial and final classifications, select "No" in both cases and select "Do not show this prompt again on start-up?". The Classification History window is also available via **Tools** → **Options...** → **Classification** → **Classification History** → **Settings**.
 Classifying a CNV or LOH region works similarly to setting the initial or final classification - Right-click on the CNV or LOH Region to which you wish to assign a classification in the Genomic view or in the table, then select **Classify** and click on the classification you wish to set.

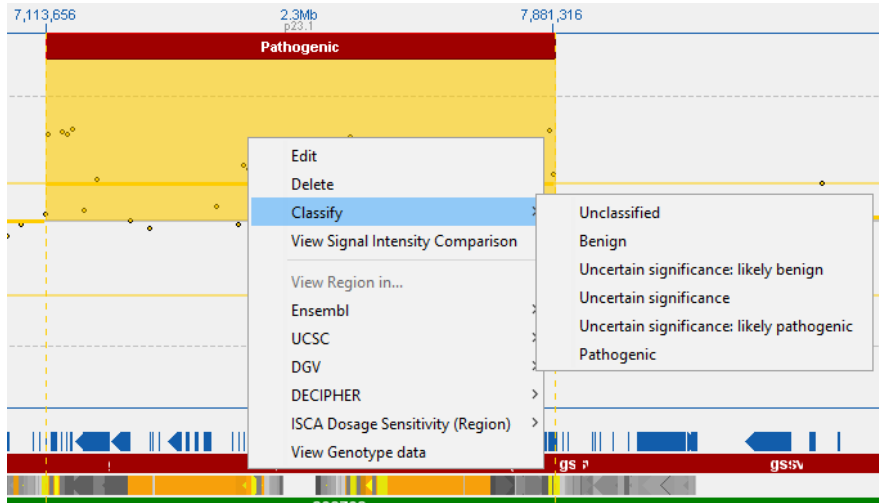


Figure 5 Classifying a CNV or LOH region

The CNVs and LOH regions will be colour-coded according to the current (most recent) classification. The tables also provide a column, **Classification**, which can be used to see the current classification of a CNV or LOH Region. As with the other columns, this column can be included in any reports via **Tools** → **Options...** → **Report Generation** → **Report Content** by adding the column to the **Included** section of the **Aberration Info** and the **LOH Info** tabs as appropriate. The full classification history can be viewed in the **Aberration Editor** window. An equivalent view is now also available for LOH Regions by right-clicking on a region in the table and selecting **Edit LOH Details**.

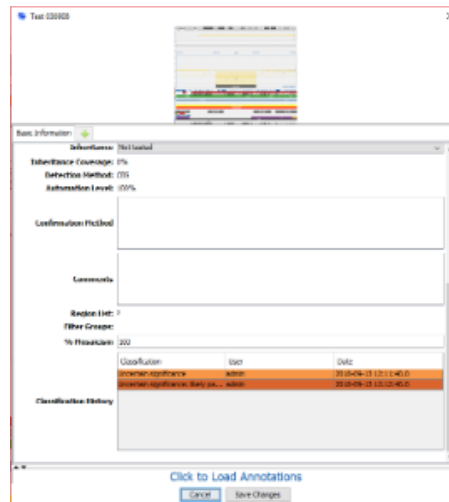


Figure 6 Aberration Editor

Probe Filter Groups

In some cases, applying additional probe filters, such as filtering out probes with a high GC content, and re-running segmentation can result in more accurate analysis of the sample by identifying additional CNVs or indicating CNVs which are not identified when different filters are applied. To



provide support for this, CytoSure Interpret v4.11 includes the ability to create groups of probe filters which can then be used during analysis of samples.

Creating Probe Filter Groups

New probe filter groups can be created via **Tools** → **Options** → **Probe Filtering** → **Probe Filter Groups** → **Create new Probe Filter group**. This will open a new window for defining the new group. Enter a suitable name and then set the filters for the group before clicking **Create**. Once created, probe filter groups can then be selected and added to the protocol, ensuring the associated filters are used during segmentation in addition to the default filters show in the main **Probe Filtering** tab.

Figure 7 Creating probe filter groups

It is also possible to view and edit the details of existing probe filter groups via **Tools** → **Options...** → **Probe Filtering** → **Probe Filter Groups** → **Modify Probe Filter Groups** to open the **Probe Filter Groups** window. For groups that are not part of a protocol that has been used in an analysis, it is possible to edit or delete the group via the **Edit** and **Delete** buttons. If the probe filter group has been used, then modifying it is not allowed. However, the details of the associated filters can still be seen by selecting **View** from the **Probe Filter Groups** window.

Analysing Cases

If the **Apply probe filters during Aberration Detection?** option is selected, after running segmentation with the main filters applied, the software will then go through each of the selected probe filter groups and re-run CBS and identify aberrations with the filters specified in each probe filter group.



As segmentation has to be re-run for each probe filter group selected, aberration detection will take longer.



The software identifies aberrations that have not been detected with all of the filter groups and highlights them in red in the **CNV Regions** table. This enables users to easily identify these aberrations and confirm whether or not they are genuine. These CNVs are also displayed with a question mark in the **Genomic View**. Right-clicking on a CNV allows users to validate the CNV, either confirming the CNV's existence or deleting the CNV. To assist with validating the CNV, the filter groups with which the CNV was detected can be viewed by hovering over a CNV in the **Genomic View**. The probe group information can also be viewed in the **CNV Regions** table or added to reports with the **Filter Groups** column.

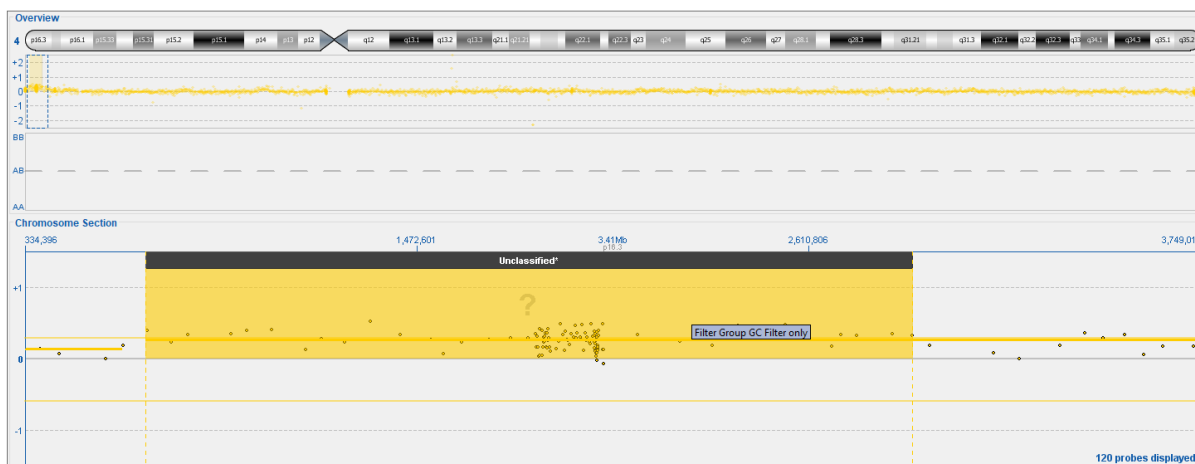


Figure 8 The software highlights CNVs not detected with all filter groups

Before submitting a case to the database, all CNVs that haven't been detected with all filter groups must be validated.

Note that the segmentation and probes will only be shown for one of the probe filter groups. By default, this is the main set of probe filters (those not used as part of a filter group). However, users can choose to view this information for other probe filter groups by right-clicking within the **Chromosome Section** of the **Genomic View**, going to **Display probes for Probe Filter Group** and then selecting the required probe filter group.

Updating Annotation tracks

To ensure the annotation tracks are as up-to-date as possible, CytoSure Interpret now supports automatically checking for and updating outdated tracks via the OGT FTP site.

Online Track Update

When new versions of annotation tracks are created, they will be added to the OGT FTP site, and can then be used to update the tracks in the software. To check whether there are any new versions of annotation tracks go to **Tools** → **Options...** → **Annotation** → **Annotation Tracks** → **Check for updates**. A window will appear asking for the connection details, to enable the software to connect to OGT's FTP site.



i Contact your OGT account manager for help connecting to the OGT FTP site if there are problems with your connection.

The software will then check for any available updates before asking for confirmation of which of the available tracks to update. By default, the update process will overwrite the existing track with the new version. Select **Create copies of the tracks?** to create a new track with the updated information instead. The software will then download the files for the annotation tracks and carry out the update process.

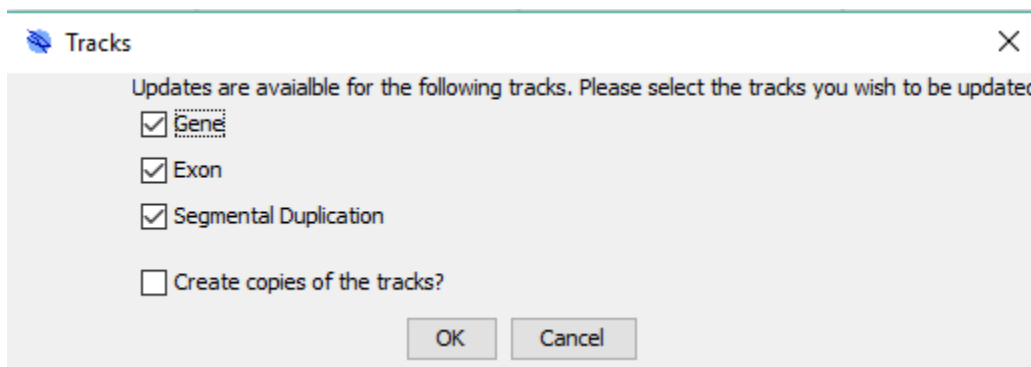


Figure 9 Confirming annotation tracks to update

Annotation tracks now also specify the current version of the track and the date the track information was last updated.

Update Tracks during Software Update

It is now possible to select tracks to update whilst updating the rest of the software. If the version of the track included with the software update is newer than the version of the track installed, then the software will update the track. Note that updating the tracks in this manner will overwrite the existing track, so if more control is required the aforementioned online update mechanism may be preferable.

Workflow


In CytoSure Interpret v4.11 the workflow has been simplified to provide a more streamlined process for analysing cases.



Figure 10 Workflow

Loading cases into the workflow



Previously analysed cases can be loaded into the workflow via **Load**. The option is provided to select cases either from the file system or from the database. If the option is chosen to open cases from the database, the software will open **Database Management**. Select the cases for analysis and then open the associated .cgh file via  , found in the central grid of buttons. Choosing to load cases from the file system will allow browsing for and selecting the batch processed .cgh files associated with the cases.



Feature extracted files can no longer be loaded directly into the workflow. Instead they should be processed via **Tools** → **Batch Processing** and then, once processing has completed, the cases can be loaded into the workflow for analysis.

Once the cases have been selected, they will be loaded into the software and, if necessary, aberration detection will be run.

Analysing and reporting cases

Once cases have been loaded, the workflow will progress to the **Classify** stage, allowing users to review and classify the detected CNVs as normal. Once all cases have been analysed and the CNVs suitably classified, select **Report**. This will save the cases, submit them to the database, and generate any reports according to the settings specified in the protocol being used for the current workflow. Finally, once analysis and reporting for the cases is complete, select **Finish** to remove all open cases from the workflow.

Batch Processing Experiment Association

In some cases, batch processing fails to correctly identify the sample to associate with a file, either because there is no sample with the expected barcode or because there are duplicate samples with the expected barcode. CytoSure Interpret v4.11 now provides better support for identifying these samples and selecting the correct sample.

Usage

When using batch processing via **Tools** → **Batch Processing**, after selecting **Match barcodes to existing sample information** and providing the fields to use for identifying the sample, the software will attempt to match the FE files with samples. If all files are successfully associated with a unique sample, batch processing will continue as normal. Otherwise, the option will be provided to review the files and barcodes which could not be resolved. Select **Yes** to open the **Review barcodes** window and resolve these issues.

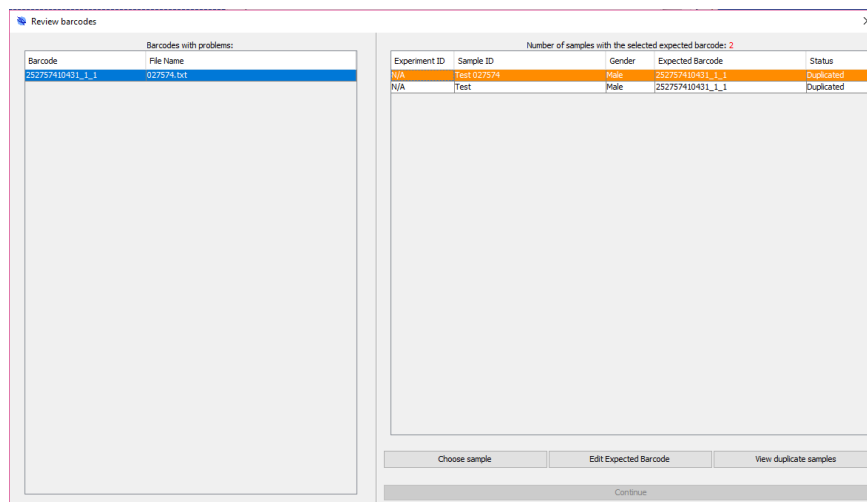


Figure 11 Review barcodes window

Begin by selecting a file from the table on the left. If no samples were found with the expected barcode, users can choose an existing sample or create a new one. Enter the **Sample ID** and **Reference ID** in the **Sample Details** window and then select the sample gender. If multiple matching samples have been identified, these will be displayed in the table on the right. For each sample either select **Choose sample** to associate the selected sample with the file or select **Edit Expected Barcode** and enter the barcode in the **New Barcode** window. Once a sample has been chosen, click **Continue** to save any changes made to the barcode to the database. Users will be asked to confirm the changes and then batch processing will continue as normal.

Miscellaneous Additions to v4.11

1. The software now opens in the **Database Management** tab for added convenience for users who open cases this tab.
2. The annotation tracks displayed in the Genomic View have been modified such that it is now possible to set a fixed height/width (depending on orientation) for the tracks and provide scroll bars when applicable. This is to provide more flexibility to the user regarding the space occupied by the tracks.
3. Enabled hiding the family tree when viewing the **Sample Details** via **Database Management**.
4. Enabled switching to the Variant Regions view for filtered aberration tracks as well as the main Aberration track.
5. Allowed deleting Protocols that haven't been used in an analysis via **Tools** → **Options...** → **Protocols** → **Delete Protocol**.
6. Added detailed description of LOH calls to reports, similar to the detailed report of CNV calls, including a screenshot of the region. The information to include can be specified via **Tools** → **Options...** → **Report Generation** → **Report Content** → **LOH Info** by adding the required fields to the **Included** section.
7. Enabled adding comments to LOH calls as well as CNV calls.



8. Various improvements to Batch Processing:
 - To simplify processing of samples, **Batch Processing** will automatically remember the folder previously selected for auto-detection of FE files and will only list and process the files generated since the previous time auto-detection was run on the folder.
 - Defaulted to previous barcode matching settings to save time setting up batch processing.
9. Enabled modification of the thickness of the CNV detection threshold lines via **Tools** → **Options...** → **Display** → **Threshold Line Thickness**
10. Colour-coded Spike-in "Match"/"No Match" text
11. Added a warning before switching to a different protocol when a case is loaded.
12. Removed "Invert" check box in Genomic View
13. Removed "Submit to database" button from individual case tabs in Genomic View to encourage synchronisation of database with .cgh file.
14. The collapsed/expanded status of annotation tracks is now remembered.
15. ISCN Notation modifications:
 - a. "Normal" ISCN now generated for cases with no reported CNVs or LOH regions.
 - b. CNV and LOH ISCN now in genomic order (CNVs followed by LOH regions).



CytoSure™ Interpret Software update — version 18.0 May 2017

The following new features are available in CytoSure Interpret Software version 4.10:

Target Regions	14
Specifying target regions for a case	14
Locking Target Regions	16
Using Target Regions in Analysis	16
Custom CNV Fields	18
Creating Custom CNV Fields	18
Editing and Viewing the values for a CNV	20
GRCh Notation	21
Usage	21
Indication of hidden probes	22
Usage	22
Named Database Connections	23
Usage	23
Template-based Report Customisation	24
Usage	24
Managing Report Templates	24
Miscellaneous Additions to v4.10	26



Target Regions

Specification of target regions (e.g. a panel of genes) for a particular case provides the following benefits to the user:

16. The ability to indicate to an analyst, in advance, on which regions to focus their analysis.
17. A permanent record of the targeted regions for each case in the database, for audit purposes.
18. The ability to remove all probes outside of the target regions during Batch Processing to prevent incidental findings in non-targeted regions during interpretation (optional).



This feature is intended to supersede the existing Data Masking and Gene Panel Scrolling features, although these will remain available in the software.

Specifying target regions for a case

To associate a target region with a case, the region must first be added to a Region List in the database. There are a number of ways to create and add to a region list, including:

- Importing a text file containing region co-ordinates via **Tools -> Options... -> Region Lists**.
- Right-clicking on an annotation (e.g. a gene) in track in the **Genomic View** and selecting **Add to region list...**
- Via the **Create Region List** button in the **Gene Panel Scrolling** plug-in (**Tools -> Gene Panel Scrolling**).

Once in the system, target regions can be added to cases in 3 ways:

2. Automatically importing them using a File Parser.
3. Manually adding them via the **Experiment Details** tab in **Database Management** before the case is opened.
4. Manually adding them via the **Genomic View** once the case has been opened.

Using a File Parser

The names of target regions to be associated with a case should be formatted in a text file to be read by a File Parser as follows:

- To add an individual target region, whose name is unique amongst all regions in all region lists in the system, simply include the name of the target region (e.g. "BRCA1").
- To add an individual target region within a particular region list, provide the name of the region list followed by a colon (:) and the name of the region (e.g. "Panel 1:BRCA1")
- To add all regions within a particular region list, provide the name of the region list followed by a colon and an asterisk (:*) (e.g. "Panel 1:*")
- Separate multiple region definitions with commas (,) (e.g. "Panel 1:*,Panel 2:TP53,Panel 3:CHEK2")



The column containing the target region specification should then be linked to the "Target Region(s)" field in the File Parser before being imported.

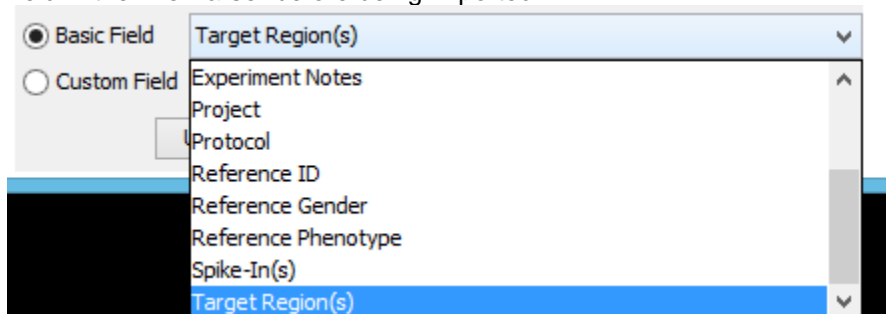


Figure 12 Linking a column in a text file to the Target Region(s) field.

Alternatively, a File Parser can be used to create target regions from the CNVs identified in a previously analysed sample and associate these target regions with the imported case. This can be useful during trio analysis to restrict analysis of the parents to CNVs found in the child. To create target regions in this way, simply include the Sample ID for the case from which to create the target regions in the text file to be read by the File Parser. The column containing the Sample ID should then be linked to the "Target Regions Sample ID" field in the File parser before being imported. This will ensure there is a region list (named *Sample ID* CNVs) containing regions matching the CNVs in the Sample, creating the list and regions if required, and associate all the regions in the list with the case.

Manually, via Database Management

To add target regions to a case in advance of loading its data for analysis, select the **Database Management** tab, select the sample in the table and then open the **Experiment Details** tab. The **Target Regions** section lists the regions currently associated with the case, and these can be modified by clicking on the **Edit** button to display the **Edit Target Regions** window.

Manually, in the Genomic View

Target regions can also be added to a case once it has been opened for analysis via the **Target Regions** tab in the bottom-right corner of the **Genomic View**. To modify the list of target regions, click on the **Edit** button to display the **Edit Target Regions** window.

The Edit Target Regions Window

The **Edit Target Regions** window enables the user to specify target regions for a case by selecting regions from existing region lists. To add regions to the case:

1. Click on the **Add Target(s)** button to open the **Select Regions** window,
2. Select one or more region lists from the list on the left-hand side.
3. Select one or more regions from the list on the right-hand side.
4. Click **OK**.

To remove regions from the case, select the regions to be removed and click the **Remove Target(s)** button.



#	Name	Chr	Start	Stop	
1	PTEN	10	89621195	89730532	↑
2	ATM	11	108091559	108241829	
3	BRCA2	13	32887617	32975809	
4	PALB2	16	23612481	23654678	
5	CDH1	16	68769128	68871445	
6	TP53	17	7569720	7592868	
7	RAD51D	17	33424811	33448888	
8	BRCA1	17	41194312	41279500	
9	RAD51C	17	56767934	56813703	
10	BRIP1	17	59754547	59942920	
11	STK11	19	1203798	1230434	

Buttons: Add Target(s) Remove Target(s) Lock

Figure 13 The Edit Target Regions window.

Locking Target Regions

The system also enables users to restrict modification of target regions for cases. Clicking the **Lock** button in the **Edit Target Regions** window prevents modification of the regions until the list is unlocked by a user with sufficient permissions. Only the user who locked the list and any user above them in the user hierarchy are permitted to unlock it. If the target regions have been added to the case via a File Parser, the list is automatically locked.

Using Target Regions in Analysis

Once the target regions for a case have been specified, the obvious benefit is that they can be used as a guide to the analyst to focus their attention on particular regions, but they can also be used to prevent detection of CNVs outside of these regions (incidental findings).

To guide the analyst

Once a case has been loaded, its target regions are displayed in the **Genomic View** in the **Target Regions** tab. Selecting a region in the table results in the display of that region in the scatter plot.



CNV Regions		LOH Regions		Target Regions	
#	Name	Chr	Start	Stop	
1	PTEN	10	89621195	89730532	^
2	ATM	11	108091559	108241829	
3	BRCA2	13	32887617	32975809	
4	PALB2	16	23612481	23654678	
5	CDH1	16	68769128	68871445	
6	TP53	17	7569720	7592868	
7	RAD51D	17	33424811	33448888	
8	BRCA1	17	41194312	41279500	
9	RAD51C	17	56767934	56813703	
10	BRIP1	17	59754547	59942920	
11	STK11	19	1203798	1230434	

Edit

Figure 14 The Target Regions table in the Genomic View.

To prevent incidental findings

In order to restrict analysis to only the specified target regions:

1. The feature extracted file for the case must be processed using **Batch Processing**. It is not possible to restrict analysis in this way on an ad hoc basis after Batch Processing.
2. The target regions must be specified before the case is processed using **Batch Processing**.
3. The **Remove probes outside target regions during Batch Processing?** option must be selected in **Tools -> Options... -> Probe Filtering -> Target Regions**.

If these conditions are satisfied, the feature extracted file will be processed as follows:

1. Normalisation and segmentation is carried out on all probes based on the specified Analysis Protocol.
2. All probes outside of the specified target regions are removed from the analysis.
3. Each target region is run through the segmentation process individually, and no subsequent normalisation takes place.
4. As a result, CNVs will only be detected in the specified target regions.



If no target regions have been specified this option will have no effect on the analysis of the case.




Custom CNV Fields

To provide additional flexibility in the annotation of detected CNVs, it is now possible to create custom fields for CNVs (in addition to Samples and Experiments). The custom information, as well as the basic information, can be viewed within the CytoSure Interpret Software and included in any reports that are generated.

Creating Custom CNV Fields

New custom CNV fields can be added via the **Edit Aberration Details** window. This can be accessed from **Database Management -> CNV Analysis** for an analysed and saved case by double-clicking on a CNV in the table. Alternatively, if a case is currently opened and being analysed then the **Edit Aberration Details** window can be opened by double-clicking on a CNV in the table in the **CNVs/SNPs** tab or the **CNV Regions** tab in the **Genomic View**, or right-clicking on a CNV in the **Genomic View** and selecting **Edit Aberration Details**. The default information for aberrations is still available under **Basic Information**, but new fields can be added and grouped into other categories.

To add a new category, click on the green plus () and enter a unique name for the category in the pop-up window. Click **OK** and the new category will be created. Fields can then be added to this category by selecting **Add a new field**, and again entering the name for this field in the pop-up. Click **OK** and then select the required data type from the drop-down before clicking **OK**.

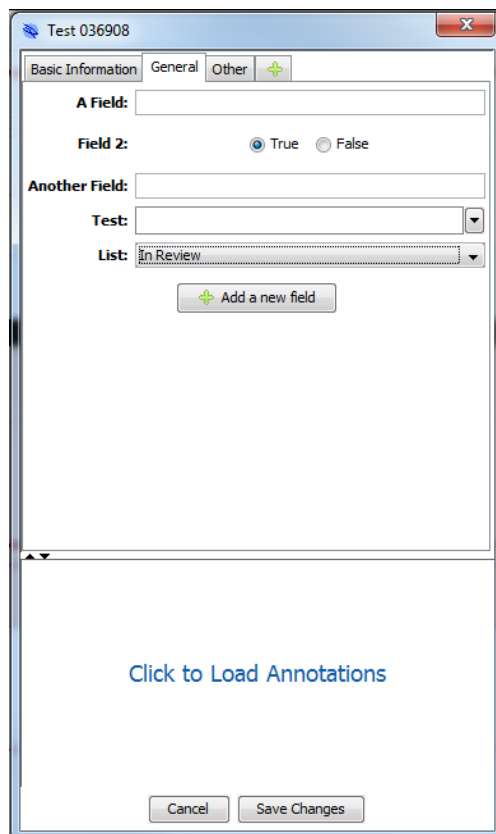


Figure 15 Example of custom fields

Fields using Lists

If the data type selected is **List** it is necessary to select the required list to finish creating the new field. This can be either an existing list or a new one can be created. To use an existing list, select **No** in the pop-up that will open, then select the list to be used from the drop-down before clicking **OK**.

Otherwise, select **Yes** in the window, then enter the name of the list and select **OK**. A window will open for adding the new items. Double click on the empty row to edit, then enter the first item and press enter. Add all the items to the list in the same way and then select **Confirm Changes** to create the new field and associated list.

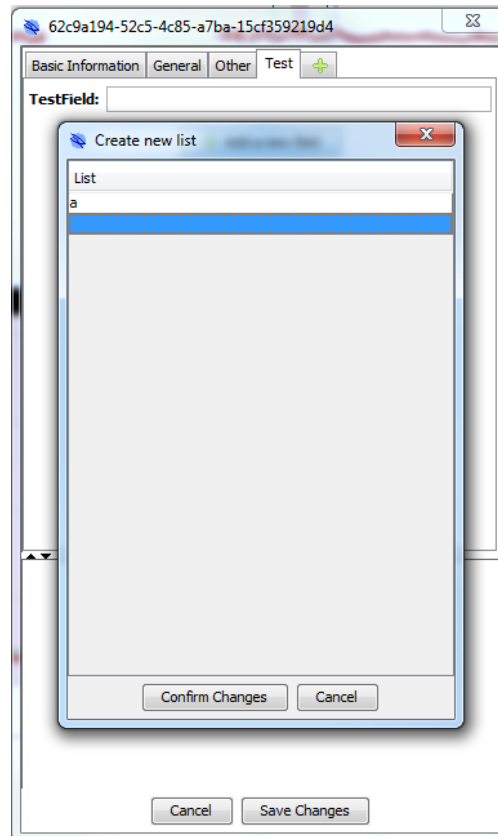


Figure 16 Creating a new list

Editing and Viewing the values for a CNV

The custom fields can be viewed similarly to the default fields. To do so, right-click on the CNV Regions table either for a currently open case in the **Genomic View** or **CNVs/SNPs** tabs, or, for a historical case, in **Database Management -> CNV Analysis** to open the column selection for that table as normal. Then, open the new sub-menu name **Custom Fields**, which can be found at the bottom, and select the sub-menu for the category. This will open a menu displaying the fields for that category. Select one to add it to the columns being displayed in the table.

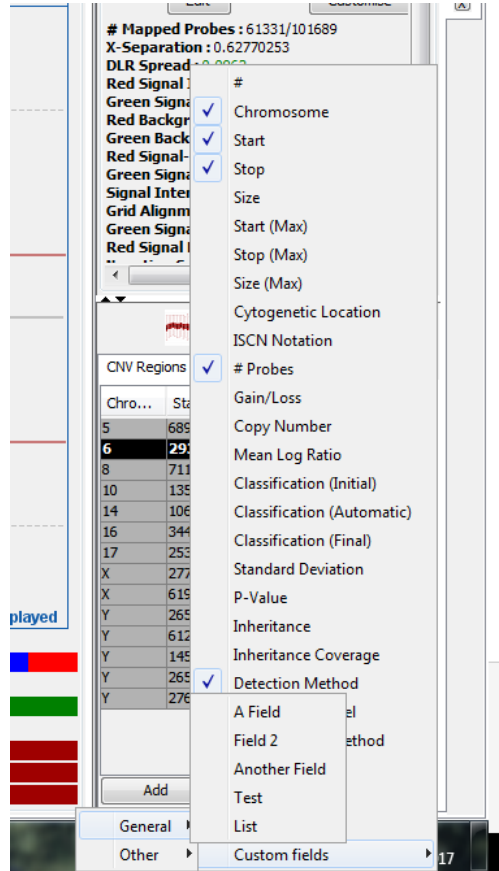


Figure 17 Selecting custom fields to view in the CNV Regions table for the open case

The values of the fields can be edited via the **Edit Aberration Details** window used to add the fields. Open the window, then select the correct category and enter the required information for the field in the associated text box before clicking **Save Changes**.

Custom fields can also be added to the reports on a case, just as for the default fields, via **Tools -> Options... -> Report Generation -> Report Content -> Aberration Info**. Any additional custom fields will be listed in the **Excluded** table underneath all the default fields and can be moved to the **Included** table as required.

GRCh Notation

CytoSure Interpret Software Version 4.10 allows the user to choose whether to use "GRCh" notation or "hg" notation for the genome build in the ISCN notation and when displaying the genome build throughout the software.

Usage

CytoSure Interpret defaults to displaying the Genome Build using "hg" notation (e.g. hg19) to maintain consistency with previous versions of the software. However, users may now choose to have the Genome Build displayed using "GRCh" notation (e.g. GRCh37) via **Tools -> Options... -**



> **Display -> Scatter Plot Display** and selecting **Use GRCh notation for Genome Build**. This will then ensure that "GRCh" notation is used in all places that the Genome Build is displayed, including using the GRCh notation whenever the ISCN notation is shown as required by ISCN 2016.

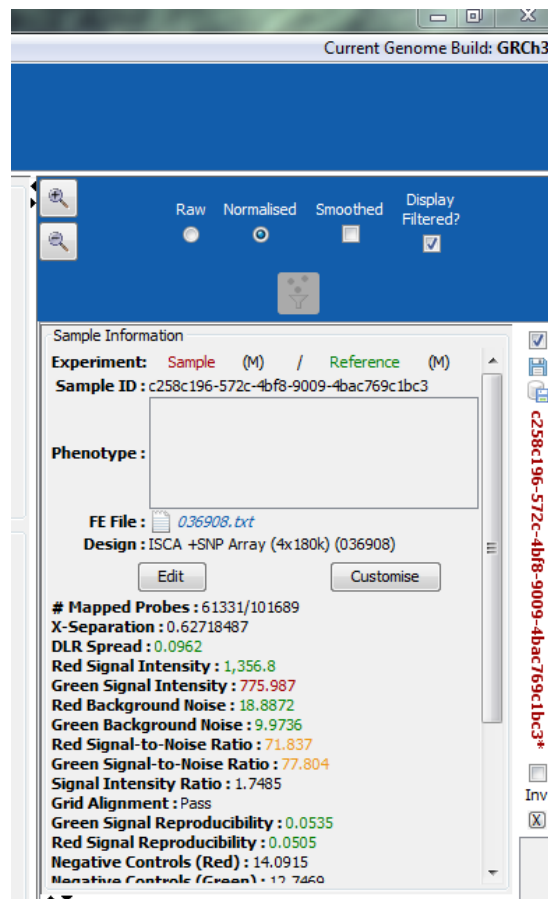


Figure 18 Example of software with the Genome Build displayed using "GRCh" notation

Indication of hidden probes

To help with analysing cases, CytoSure Interpret Software v4.10 visually indicates if there are probes within the selected genomic region with log ratio values that are outside the fixed log ratio axis scale and therefore not displayed in the plot.

Usage

If the signal axis scale has been fixed (**Tools -> Options... -> Display -> Automatic Signal Axis Rescaling** is not selected) it is possible to have probes with a very high or low log₂ ratio which are not visible in the **Genomic View**. However, these probes are still used to call aberrations in the case. To assist in interpreting any aberrations containing such probes, the software provides a visual indication to warn users that there are probes outside the current signal axis range. This is shown via a red circle with an exclamation mark in the top left-hand corner of the **Chromosome Section**.

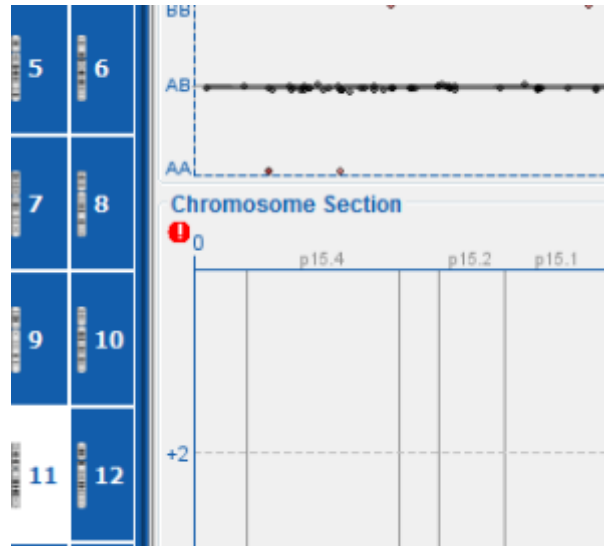


Figure 19 Indication of probes outside the signal axis range

To view all the probes in the currently selected chromosome section click on the exclamation mark. This will cause the signal axis to zoom out and display all probes until a different section of the view is navigated to.

Named Database Connections

To enable simple switching between different databases, CytoSure Interpret Software now includes the ability to save the settings for the different database connections. Users can then select the database to which they would like to connect from a drop-down of saved connections rather than having to re-enter all the details each time they wish to switch databases.

Usage

To save the details of a database connection, select **Switch Database** from the **Login** window to open the **Database Connection** window and then enter the correct details for the database connection you wish to save. Once the details have been entered press **Save**, enter the name under which you wish to save the connection in the text box provided and click **OK**. To update an existing saved connection, simply provide the name of the connection you wish to modify at this point, otherwise enter the new name.

The saved connection will then appear in the **Named Connection** drop-down when users open the **Database Connection** window. The required database can be selected and the remaining fields will be populated with the details saved for that connection. Click **Connect** to use this database.

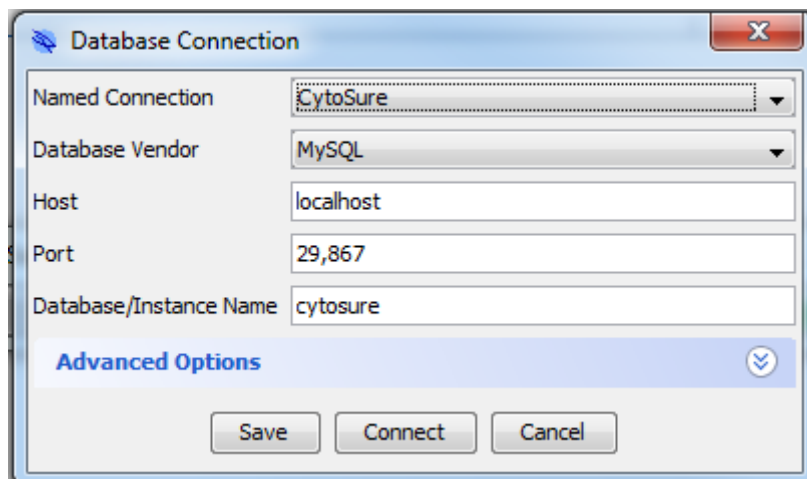


Figure 20 Saved database connections can be selected for quick and simple switching between different databases

Template-based Report Customisation

CytoSure Interpret Software now uses templates when generating the reports for a case. This makes it simpler to modify existing reports and add new ones, enabling users to update the templates used in the Software.

Usage

Reports are still generated from the software via **Tools -> Export Aberration Report...** or **CNVs/SNPs -> Export: -> Aberration Report**, or for previously analysed cases **Database Management -> Regenerate Report**, then selecting the report to be generated. However, the basic reports are now generated using templates to define the layout which can be managed by the users. Templates defining the current basic reports that can be generated (HTML, PDF, and Tab-Delimited Text) will automatically be added to the software to ensure these reports are still available to users.



Many modifications to the reports can now be done via these templates. However, for more complex, customised reports, plug-ins can still be provided and all existing report plug-ins will continue to work.

All available reports, including those generated using plug-ins, can be viewed via **Tools -> Options... -> Report Generation -> Report Type**, and the default report selected.

Managing Report Templates

Templates for the basic reports can be managed via **Tools -> Options... -> Report Generation -> Manage Templates**. This will open the **Manage Report Templates** window, which list all reports



that are generated from a template along with the file extension of the generated report (pdf, txt, html, etc.).

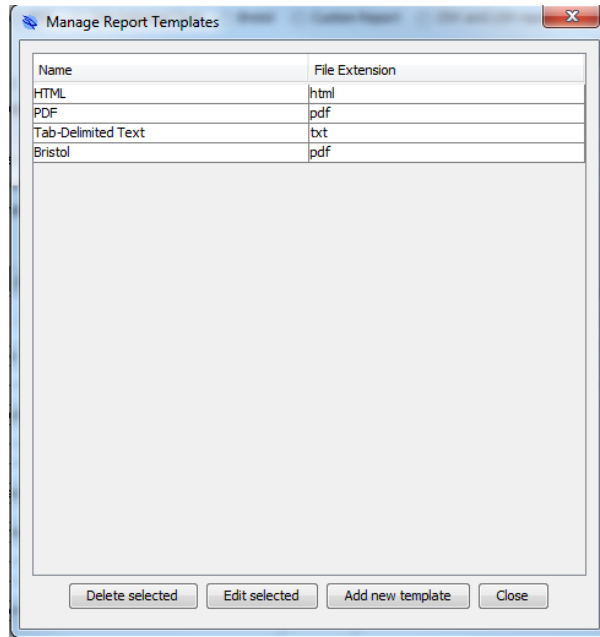


Figure 21 The window for viewing and managing the basic reports

Reports can be deleted by selecting the report to delete and then clicking **Delete selected**. To edit a report, select the report from the table and then click **Edit selected** to open the **Update Report** window. Make any required changes to the name or file extension. To update the template, click **Select template file** and browse to the file to be used as the new template. Once all required changes have been made select **Update report** and the modified report will be saved.

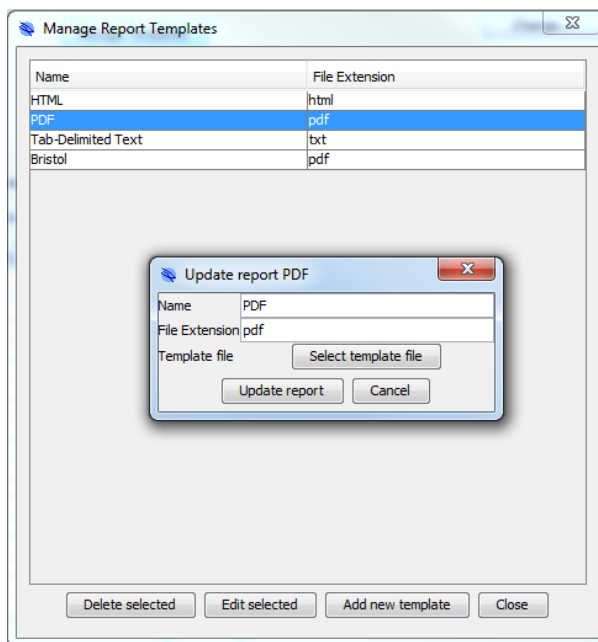


Figure 22 Update report window

New reports can be added by selecting **Add new template**, then entering the name and file extension and uploading a report as for editing reports before clicking **Create report** to save the new report and have it available for reporting on future cases.

i Customising Report Templates

It is not currently the intention to provide an API for users to develop their own templates in order to customise reports. Rather, users requiring modification to the report format are encouraged to contact their OGT account manager, describing the changes, to request a modified reporting template.

Miscellaneous Additions to v4.10

- Added table showing the copy number status of each chromosome for each open case to the Aneuploidy Summary (**Genomic View -> Aneuploidy Summary**) .
- Enabled specifying whether the classification should be assigned as the Initial Classification or the Final Classification of the CNV when the user has chosen to assign a classification to any CNV regions identified as artefacts during Aberration Detection.



- Enabled specifying the order in which fields should be included in a report (**Tools -> Options... -> Report Generation -> Report Content**) using **Up/Down** buttons to re-order the included fields.
- Allowed specifying the LOH columns in the report via **Tools -> Options... -> Report Generation -> Report Content -> LOH Columns** and adding the required fields to the **Included** section. By default, the information included for LOH columns is the columns that were previously displayed along with the classification of the LOH region.
- Provide options to hide archived users from the audit trail (**Database Management -> Audit Trail -> Hide Archived**) and select/deselect all users (**Database Management -> Audit Trail -> Select All**) to simplify filtering the audit trail to the relevant users.
- Modified LOH track to display LOH Regions from all historical cases that have been saved in the database (rather than the LOH regions detected in currently loaded cases, which has been superseded by the B-Allele Frequency and Allele Status plots in the **Overview** section) to aid users in interpretation of LOH regions in the current cases.
- Database user name and password moved to an **Advance Options** section within the **Database Connection** window.
- Database connection passwords encrypted in configuration file.



CytoSure™ Interpret Software update — version 17.0 October 2016

The following new features are available in CytoSure Interpret Software version 4.9:

ISCN 2016	29
Usage	29
Exon-focused Segmentation Algorithm	29
Usage	30
Additional Information	30
Export to DECIPHER	31
Activation	31
Assigning CNV Inheritance Status	32
Mapping Classification to Pathogenicity	33
Generating the Spreadsheet	34
PGS Reference Sample Quality Check	34
Usage	34
Sequential Classification	35
Activation	35
Classification	36
Saving	37
Application of Separate CNV Gain and Loss Size Thresholds	37
Activation	37
Usage	37
Different Analysis Protocols During Batch Processing	38
Usage	38
Automatic Waviness Calculation	40
Usage	40
Miscellaneous Additions to v4.9	40



ISCN 2016

CytoSure Interpret Software v4.9 now provides the option to use both ISCN 2013 and ISCN 2016 notation for CNVs and LOH regions in tables displayed in the software and reports it generates. The software also enables reporting of the overall ISCN notation for the experiment.

Usage

CytoSure Interpret Software v4.9 defaults to ISCN 2013 to maintain consistency between previous versions. However, users may choose to use ISCN 2016 notation via **Tools -> Options... -> Report Generation -> Report Content -> ISCN Version**. The differences between the 2013 and 2016 notation that have been implemented in the software are as follows:

1. No thousands separator (previously a comma) is included in co-ordinates. (i.e. 110,862,474 becomes 110862474)
2. An underscore (_) is used to separate start and end co-ordinates instead of a hyphen. (i.e. 110862474-110964775 becomes 110862474_110964775)
3. CNVs identified using mosaicism detection have their mosaic fraction indicated. (i.e. a 75% mosaic deletion of 2q13(110862474_110964775) would be reported as 2q13(110862474_110964775)x1[0.75])
4. If selected via **Tools -> Options... -> Report Generation -> Report Content -> Include Maximum Co-ordinates**, the notation can include the maximum size of the CNV (e.g. arr[hg19] 2q13(110783310x2, 110862474_110964775x1, 111079006x2)).

Additionally, users can now report ISCN notation for the experiment as a whole, including all relevant CNV and LOH regions. Inclusion of CNVs and LOH regions in the experiment ISCN notation is based on their Classification as follows:

- If **Tools -> Options... -> Classification -> Reported Classifications -> Also remove these aberrations from the experiment ISCN Notation?** is selected, only CNVs and LOH regions whose classification has their **Report?** column checked will be included.
- Otherwise, all detected CNV and LOH regions will be included.

To include the experiment ISCN notation in the report, select **Tools -> Options... -> Report Generation**, and in the **Experiment Details** tab move **ISCN Notation** to the **Included** column.

ISCN Notation	<pre>arr[hg19] 15q14q15.1(34533380_42636858)x1, 2q35q37.1(219992599_231235872)x1, 2q33.2q35(204598942_219263558)x1, 20p13p11.21(225414_23609009)x2 hmz, 17p13.3p11.2(1135130_21047102)x2 hmz, 2q33.2q37.2(204838619_236025302)x2 hmz, 15q24.3q26.3(77422280_101941562)x2 hmz</pre>
----------------------	--

Figure: Example of experiment-level ISCN notation

Exon-focused Segmentation Algorithm

CNV detection within CytoSure Interpret Software is facilitated by processing the log ratio data using the Circular Binary Segmentation (CBS) algorithm in order to group probes into regions of equal copy number. The sensitivity of the CBS algorithm (i.e. how likely it is to split a group of probes into separate segments) is determined by a number of settings, which have been fixed in



CytoSure Interpret Software since its initial release. After segmentation, thresholds based on log ratio and region size are applied to determine which segments are significant and therefore represent a CNV.

Following the release of OGT's exon-focused products, such as the CytoSure Medical Research Exome and Constitutional v3 arrays, customer feedback indicated that a reduction in the sensitivity of the algorithm, in order to reduce the likelihood of detecting false positives, would be beneficial. As a result, an alternative segmentation option, optimised for exon-focused arrays, has been developed and is now available as standard in CytoSure Interpret Software.

Usage

CytoSure Interpret Software uses the original segmentation (referred to as **Built-in CBS**) by default, but users wishing to process data with the **Exon-focused Segmentation** may select this option via **Tools -> Options... -> Aberration Detection -> Advanced Options -> Alternative Segmentation**.



The **Alternative Segmentation** setting is an Analysis Protocol option, so a new protocol will need to be created if this setting is changed.



The way in which the Exon-focused Segmentation has been implemented has the additional benefit of being substantially faster than the Built-in CBS implementation.

The selected method will then be applied during Batch Processing or when triggered by the user after data has been loaded. As it has been specifically developed to be less sensitive, processing the same data with the **Exon-focused Segmentation** is likely to produce fewer CNV calls when compared to results of the **Built-in CBS**.

Additional Information

The differences between the **Built-in CBS** and **Exon-focused Segmentation** are as follows:

1. The "kmax" and "alpha" CBS settings have been reset to their default values
- <https://bioconductor.org/packages/release/bioc/manuals/DNAcopy/man/DNAcopy.pdf>
2. Outlier removal is carried out before segmentation via the "smooth" operation
- <https://bioconductor.org/packages/release/bioc/manuals/DNAcopy/man/DNAcopy.pdf>
3. The "MergeLevels" post-processing algorithm is applied in order to combine segments which are not significantly different. <http://bioinformatics.oxfordjournals.org/content/21/22/4084.long>
4. The data is processed using the R statistical package rather than within the Java Runtime Environment.



Export to DECIPHER

One method of submitting classified CNV calls to DECIPHER is by completing a pre-formatted Excel spreadsheet and uploading it to DECIPHER. The DECIPHER Export feature provides an interface for assigning inheritance status to detected CNVs (if desired by the user), mapping CytoSure Interpret classifications to DECIPHER Pathogenicity and generating the Excel spreadsheet ready for submission.

Activation

To access the feature, first load one or more cases into the Genomic View in the usual way, then select **Tools -> Export -> Export in DECIPHER Format**. When the feature is run for the first time, a number of custom fields, corresponding to column headings within the standard DECIPHER Excel template, will be created in a new **DECIPHER** category in the **Sample Data** tab in **Database Management**. It may be necessary to populate these fields for the sample before exporting the data (either manually or via a File Parser). The fields are as follows:

- **Chromosomal sex - Required.** If left empty, this column will contain "46XY" for male samples and "46XX" for female samples.

! Unknown Gender

If this field is left empty and the sample gender is set to "Unknown", this column will be left blank in the spreadsheet and the submission may be rejected by DECIPHER.

- **Open-access consent** - Optional. Select either "True" or "False" to include "Yes" or "No" respectively in the export. Defaults to "No" in the export if not set.
- **Age at initial presentation** - Optional.
- **Prenatal age in weeks** - Optional.
- **Responsible contact** - Optional. Should contain an email address if used.
- **Mother is** - Optional. Whether the mother is affected with a related or similar phenotype.
- **Father is** - Optional. Whether the father is affected with a related or similar phenotype.
- **Note** - Optional. Any notes relevant to the DECIPHER submission.

Chromosomal sex:	<input type="text"/>	Open-access consent:	<input checked="" type="radio"/> True <input type="radio"/> False
Age at initial presentation:	<input type="text" value="5"/>	Responsible contact:	<input type="text"/>
Mother is:	<input type="text" value="Affected with related or similar phenotype"/>	Father is:	<input type="text" value="Unaffected with related or similar phenotype"/>
Note:	<input type="text"/>		
			<input type="button" value="Add a new field"/>

Figure: The fields of the **DECIPHER** category in **Database Management -> Sample Data**



If at least one case is loaded, the **DECIPHER Export** window will appear containing a tab for each loaded case and enabling CNV inheritance status assignment and data export.

#	Chromosome	Start	Stop	Size	Gain/Loss	DECIPHER Inheritance
1	1	243095856	243218565	122.71Kb	Loss	?
2	6	162654003	163000271	346.27Kb	Gain	?
3	16	942285	2165266	1.22Mb	Loss	?
4	16	31934674	34173734	2.24Mb	Loss	?
5	22	21468352	21759580	291.23Kb	Loss	?
6	X	60701	162434	101.73Kb	Loss	?
7	Y	10701	112434	101.73Kb	Loss	?
8	Y	25276456	25307702	31.25Kb	Loss	?
9	Y	25381444	25424792	43.35Kb	Loss	?
10	Y	26928817	26943521	14.7Kb	Loss	?
11	Y	27006933	27045134	38.2Kb	Loss	?

Figure: The **DECIPHER Export** window

Assigning CNV Inheritance Status

An optional column in the DECIPHER spreadsheet is "Inheritance", the value of which must be one of 11 predefined status types. The **DECIPHER Export** window displays the table of detected CNVs with an additional column labelled **DECIPHER Inheritance**, which provides a drop-down list for CNV that enables the user to assign one of the predefined classifications to the selected CNV. Additionally, multiple CNVs can be assigned the same DECIPHER Inheritance by:

1. Selecting the group of CNVs from the table to which the inheritance should be assigned.
2. Selecting the appropriate inheritance from the drop-down list at the bottom of the window.
3. Clicking the **Bulk Assign Inheritance** button.

⚠ Confirmation Method Field Usage

The DECIPHER Inheritance is actually stored in the "Confirmation Method" field for the CNV on the assumption that it will not be used as designed. As such, separate modification of this field for the CNV in the software will result in the DECIPHER Inheritance being overwritten.



Mapping Classification to Pathogenicity

Each CNV included in the DECIPHER spreadsheet should be assigned one of 6 predefined Pathogenicity values. As the DECIPHER Pathogenicity is equivalent to the CytoSure Interpret Software Classification, to avoid the requirement to separately assign Pathogenicity values to each CNV before generating the spreadsheet, the DECIPHER Export plug-in provides an interface enabling the user to map Classification to Pathogenicity. To access the interface, click on the **Classification Mapping** button.

Classification	Pathogenicity
Unclassified	
Benign	Benign
Uncertain significance: likely benign	Likely benign
Uncertain significance	Uncertain
Uncertain significance: likely pathogenic	Probably pathogenic
Pathogenic	Definitely pathogenic
Artefact	

Figure: The **Classification Mapping** window

The Pathogenicity column in the DECIPHER spreadsheet will contain the value displayed in the Pathogenicity column in the **Classification Mapping** window corresponding to the CNV's overall classification ("Classification (Final)", or "Classification (Initial)" if "Classification (Final)" is set to "Unclassified"). To modify the Pathogenicity associated with a Classification, simply click on the corresponding cell in the Pathogenicity column and select the appropriate item from the drop-down list.



❗ Blank Pathogenicity

If the Pathogenicity for a Classification is left blank, CNVs with that classification will **not** be included in the DECIPHER spreadsheet.

i Mapping Storage

The mapping information is stored on the local machine under the user's Windows home directory. As such, it will be necessary to configure the classification mapping separately for any user intending to run the feature.

Generating the Spreadsheet

Once all interpretation and any required inheritance assignment is complete, click on the **Export in DECIPHER Format** button, select an appropriate location for the file and click **Save**. Once the spreadsheet is generated, the software will attempt to open it automatically.

✔ Sample Data

The Sample-level information included in the spreadsheet is requested from the database just before it is created. Therefore, if any Sample-level data is missing or incorrect, simply make the required changes in Database Management, save the sample and re-run the export.

PGS Reference Sample Quality Check

The accuracy of single-channel analysis of PGS data relies on the quality of the male and female reference samples against which each test sample is compared. In order to enable quick identification of reference sample quality issues before test samples are processed, the Excel Planner processing interface has been updated to include a review of reference QC metrics.

Usage

To access the reference metrics, simply import data from the worksheet as previously, via **Tools -> PGS Analysis -> With Worksheet (Updated)**. After confirming the sample selection, instead of immediately processing all data sets, a window will appear displaying the QC metrics of the male and female reference samples when analysed against each other.

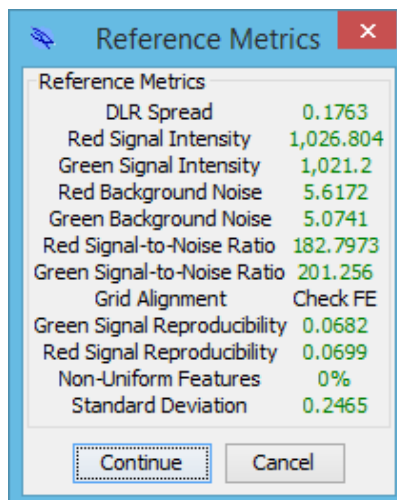


Figure: The metrics of the male and female reference samples for PGS data

If the metrics displayed in the **Reference Metrics** window are acceptable, click **Continue** to trigger processing of the test sample data. Otherwise, click **Cancel** to abort processing.

✔ Alternative Reference Data

If the reference data is not of sufficient quality, and it is deemed appropriate, it is possible to use reference data taken from a different slide. To do this, include the slide in the Excel Planner, add the reference sample names in the appropriate positions on the slide, and link the male and female references to those cells instead of the cells on the original slide.

Sequential Classification

The purpose of the Sequential Classification tool is to store the time at which classifications are made by each user who interprets a case. This information is stored in the .cgh file, so it is essential that, after classifications have been made, the .cgh file is updated by saving the case. Whether the case should also be submitted to the database will depend on whether the lab considers it appropriate for potentially temporary classifications for CNVs to be included in the Aberration database.

Activation

To activate the plug-in:

1. Open a case so that it is displayed in the **Genomic View** (the mechanism used is irrelevant and will depend on your preferred analysis workflow).
2. Select **Tools -> Sequential Classification**.



3. A new window will appear at the bottom of the screen labelled **CNVs** which includes a tab for each currently loaded case containing a table listing the CNVs detected in the case.

i The CNV table

The columns included in the CNV table match those chosen for inclusion in the **Detected CNVs/SNPs** table (when the **Table View** radio button is selected) in the **CNVs/SNPs** tab (rather than the **CNV Regions** table in the **Genomic View**). If these columns are updated in the **CNVs/SNPs** tab (by right-clicking on the column headings) it may be necessary to click the **Refresh Table** button to ensure consistency. In addition to the standard CNV information columns, the table contains one column for each user who has previously determined a classification for a CNV in the case, as well as a column for the current user (displayed in **bold** and always as the last column in the table).

Classification

There are a number of ways in which a classification can be made for one or more CNVs in the table:

1. To **classify an individual CNV**, select the appropriate classification by clicking on the cell in the last column of the table for that CNV and select the appropriate classification from the drop-down box.
2. To **agree with the classifications of one or more CNVs made by a particular user**, select the CNVs to which this applies and click the **Agree with User Name** button.
3. To **agree with the latest classification of one of more CNVs** (which is stored in the "Classification (Initial)" field), select the CNVs to which this applies and click the **Agree with Classification (Initial)** button.
4. To clear the classification of one or more CNVs made by the current user, select the CNVs to which this applies and click **Clear classification**.



Do not use any other (built-in) mechanism for assigning a classification to a CNV (e.g. by right-clicking on the CNV and selecting **Classify (Initial)**), as the software will not record the assigning user or time at which the classification was made.




i Important Note

The tool overwrites the **Comments** field for the CNV, so if you intend to use this field to store other information the use of the Sequential Classification function will not be appropriate.

Saving

Once classifications have been made for all CNVs, they will only be available to the next user who interprets the case if the .cgh file is saved before the case is closed:

- If the case is being analysed in Workflow Mode (it has been opened via the **Load** button), close the case by clicking the chequered flag icon and click **Yes** when asked whether to save the changes.
- Otherwise, use either **File -> Save** or the  icon.

✓ Re-opening Cases

If a case is re-opened and doesn't appear to contain the records of previous classifications by other users, ensure that it hasn't been saved in a different location by the previous user.

Application of Separate CNV Gain and Loss Size Thresholds

By default, the same size threshold, as specified in **Tools -> Options... -> Aberration Detection -> Aberration Calling Options -> Min. Size (Mb)/Min.Size (Mb) (Mosaicism)**, is applied to both gains and losses during CNV detection. However, from version 4.9, CytoSure Interpret Software can be configured to apply different size thresholds for CNV gains to those used for losses.

Activation

In order to activate this feature, select **Tools -> Options... -> Aberration Detection -> Advanced Options**, and enable **Apply Separate Gain and Loss Size Thresholds?**.

Usage

The **Separate Size Thresholds** section enables modification of the size thresholds for gains only. To modify the loss thresholds, use the **Min. Size (Mb)** and **Min. Size (Mb) (Mosaicism)** fields in the standard **Aberration Calling** options section, and these will be reflected in the appropriate fields in the **Separate Size Thresholds** section.



Once applied, subsequent aberration detection will filter segments based on the appropriate calling threshold depending on whether they represent a gain or a loss in the sample.

! As the settings provided by this functionality form part of the analysis protocol, a new protocol will need to be created for the feature to be used in workflow mode.

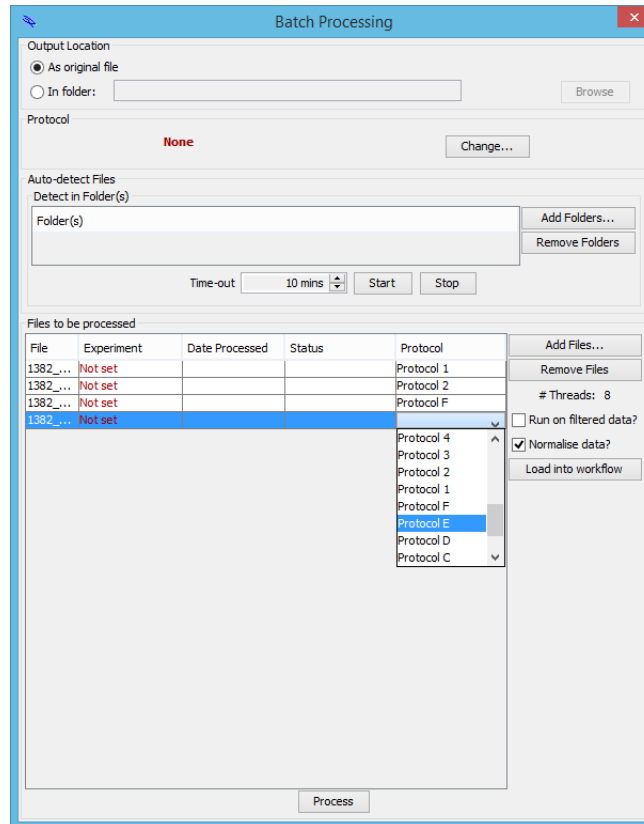
Different Analysis Protocols During Batch Processing

Previously, a Batch Processing run restricted users to selecting one analysis protocol for all files being processed. CytoSure Interpret Software v4.9 now provides the ability to specify analysis protocols on a per-file basis.

Usage

Individual protocols can be assigned to experiments during batch processing in two ways:

1. Once the data has been added to the list of files to be processed, the user can select the analysis protocol from the drop-down box for the file in the **Protocol**





- The protocol can be associated with the experiment before batch processing by including the protocol name as a column in the File Parser and linking it to the **Protocol Basic Field**.

Edit File Parser

Parser Name: **Start from line:**

Column separation character(s): **File Type:** One sample per LINE
 One sample per FILE

Tab
 Space
 Other

Automatically create samples and experiments if required?

A	B	C	D	E	F	G	H
257517510001	1_1	Expecte... 2575175...	Sample ID ABC_123	Sample G... Male	Referenc... Male Ref...	Referenc... Male	Protocol 1

H1

Basic Field Custom Field

- Sample ID
- Sample Label
- Experiment Notes
- Project
- Protocol

The Batch Processing module (**Tools -> Batch Processing**) now selects the analysis protocol used for each file in the following order:

- If option 2 above (association via a File Parser) has been carried out, and the file is automatically associated with the sample via the barcode as part of the process, the protocol loaded by the File Parser is used.
- If option 1 above has been used, and the file is not automatically associated with a sample via the barcode or the sample does not already have an associated experiment which has had its protocol set via a File Parser, the protocol selected in the drop-down list for the file is used.
- Otherwise, the current protocol (shown in the **Protocol** section of the **Batch Processing** window) is used.



Automatic Waviness Calculation

Previously, users wishing to view the Waviness metric were required to trigger its calculation once the case had been loaded into CytoSure Interpret Software for analysis (via the **Calculate** link in the **Sample Information** section for the sample on the right-hand side of the **Genomic View**). It is now possible to configure the software to automatically calculate waviness during **Batch Processing**.

Usage

To enable automatic waviness calculation during Batch Processing, select **Tools -> Options... -> QC Metrics** and ensure that **Automatically calculate waviness during Batch Processing** is enabled.

Miscellaneous Additions to v4.9

- Additional PGS Planner formats to provide a "Gasket Orientation" view (PGS_Planner_GPR_Gasket_Orientation.xlsx) and an alternative system of sample and barcode entry (PGS_Planner_Alternative.xlsx) which may be more convenient than the original planners for some users. The files are located in the software installation directory.
- Ability to submit multiple .cgh files to the database in bulk via **Tools -> Submit CGH Files**.
- **View -> Go To Position (bp)** accepts more flexible formatting of start and end co-ordinates.
- Various improvements to manual editing of CNV and LOH region boundaries:
 - Prevented accidental creation of negative-sized regions.
 - Confirmation required before size adjustment.
 - Provided option to "snap" to first and last probes within the adjusted region.
 - Maximum size of region adjusted accordingly.
- Enabled restriction of reported LOH regions by minimum size via **CNVs/SNPs -> Detected CNVs/SNPs** (with **Ideogram View** selected) -> **Options... -> LOH Calling -> Minimum LOH region size (Mb)**.
- Regions ordered alphabetically in the data masking interface, which also provides the option to deselect all regions.
- Chromosome boundaries highlighted in **Whole Genome View** in the **PGS Analysis** tab.
- Advanced filters for HPO terms, custom fields and projects now applied correctly.
- Bulk sample deletion via **Database Management -> Batch Operations -> Samples -> Delete** improved
- QC Metric display and reporting, and metric threshold set management is now also available via **Tools -> Options... -> QC Metrics**.

CytoSure Interpret - Software update guide



- Annotation track updates, including:
 - DECIPHER - also including gain/loss information.
 - Copy Number Variation (DGV) - "Gold Standard" track now provided instead of all variants.
 - Genes and Exons - Canonical Transcript of genes with associated exons now provided.
 - Users are now able to install updates as separate annotation tracks to enable review before transitioning to new content.



CytoSure™ Interpret Software update — version 16.0 May 2015

The following new features are available in CytoSure Interpret Software version 4.8:

Streamlining of Targeted Panel Analysis	43
LOH Analysis Improvements	44
LOH Region Classification	44
SNP Data Zooming	46
Manual LOH Region Addition	46
LOH Region Adjustment	47
PGS Analysis Improvements	48
Excel Planner for Batch Analysis	48
Flexible Sample Selection	52
Dual Reference Analysis	52
Annotation Reporting Updates	53
Custom Annotation Reporting	53
Probe-Annotation Overlap Reporting	56
Custom Annotation Information Display	57
Miscellaneous	58



Streamlining of Targeted Panel Analysis

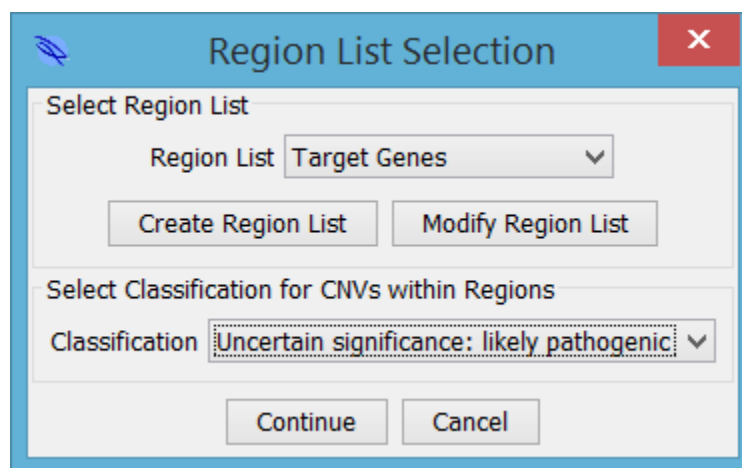
As an addition to the features supporting more streamlined analysis of targeted designs such as the CytoSure Medical Research Exome Array, an interface allowing users to easily scroll through customisable “panels” of target regions has been created. The **Gene Panel Scrolling** plug-in is available via the **Tools** menu, and enables users to:

1. Create multiple gene/targeted region panels.
2. Select a specific panel.
3. Assign a specific classification to all CNVs detected in the panel regions.
4. Scroll through all genes in the panel to easily visualise the results in the panel regions.

In combination with the selective reporting of CNVs based on their classification (**Tools -> Options... Classification -> Existing classifications -> Report?**), the plug-in also provides a quick method for only reporting CNVs within the panel regions.

To use the plug-in:

1. Load a case and ensure that segmentation has been carried so that CNVs have been automatically detected.
2. Select **Tools -> Gene Panel Scrolling** to display the **Region List Selection** window.



3. If the panel has not yet been created, click the **Create Region List** button, enter a name for the panel in the **Region List Name** field and click **Import annotations from file**. The selected file should be a plain text file containing a list of gene names separated by a line break. After clicking **Open**, the plug-in will search the annotation database for annotations matching the names in the file and create a region list using the associated co-ordinates. The new panel will then be available from the **Region List** drop-down box.
4. Select the appropriate panel from the **Region List** drop-down box.
5. Select the classification to be applied to all CNVs detected within the selected panel region using the **Classification** drop-down box.
6. Click **Continue**. This will result in the following:
 - a. The Initial Classification of all CNVs overlapping a region of the selected panel will be set to the selected classification.



- b. A table displaying the regions contained in the selected panel will appear. Selecting a row of the table will result in the display of the associated panel region in the **Genomic View**.

Name	Chr	Start	Stop
TPM3	1	154127784	154167124
NRXN1	2	50145643	51259674
NEB	2	152341850	152591001
ELN	7	73442119	73484237
LIMK1	7	73497263	73536855
HGSNAT	8	42995556	43057998
EDEM2	20	33703160	33865928
UQCC	20	33890369	33999944
CEP250	20	34042985	34099804
FER1L4	20	34146507	34195484
PHF20	20	34359896	34538303
EPB41L1	20	34679426	34820721
DLGAP4	20	34894258	35157040
RP5-977B1.11	20	35177530	35232932
NDRG3	20	35280169	35374481
SAMHD1	20	35518632	35580246
RBL1	20	35624752	35724398
BLCAP	20	36130816	36156333
CTNBL1	20	36322408	36500531
BPI	20	36888551	36965907
RALGAPB	20	37101459	37207504
PPP1R16B	20	37434348	37551667
DHX35	20	37590942	37668366

To scroll through the CNVs located within the panel regions, use the **CNV Regions** table in the **Genomic View** and select those regions given the classification selected in step 5.

LOH Analysis Improvements

LOH interpretation has been improved by adding functionality in line with that available for CNVs, as described in the following sections.

LOH Region Classification

Users are now able to classify LOH regions detected using CytoSure arrays with SNP probe content in the same manner as CNV regions, by either:

1. Right-clicking on the LOH region in the **B-Allele Frequency** or **Allele Status** plot in the **Genomic View**.

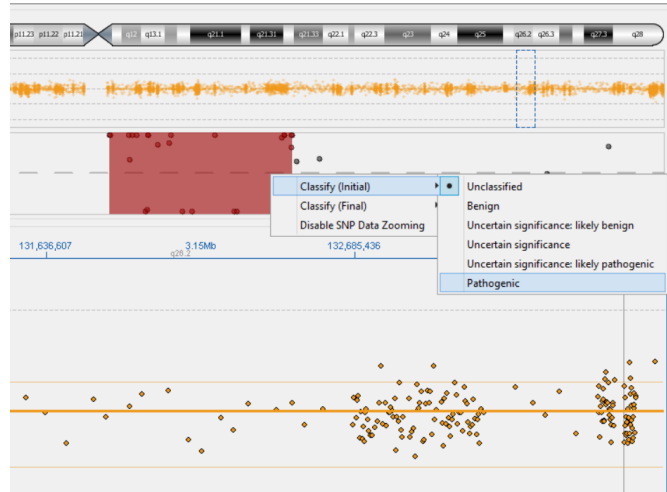


Figure: Classifying an LOH region via the Genomic View B-Allele Frequency or Allele Status plots.

2. Right-clicking on the corresponding row of the **LOH Regions** table in the **Genomic View**.

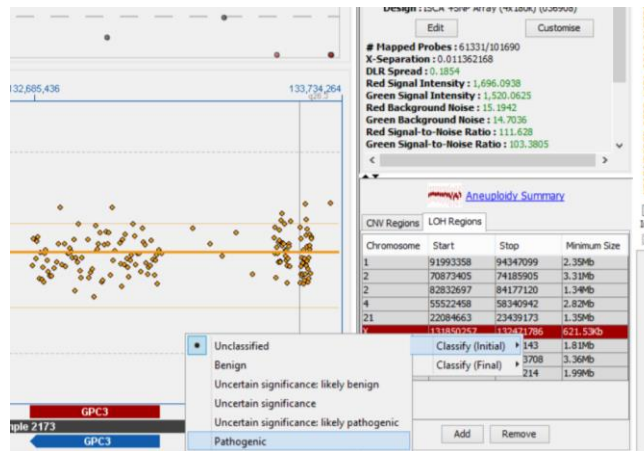


Figure: Classifying an LOH region via the LOH regions table.

3. If an LOH region has been submitted to the database, right-clicking on the corresponding row of the **LOH Regions** table in the **SNP Analysis** section of the **Database Management** tab.

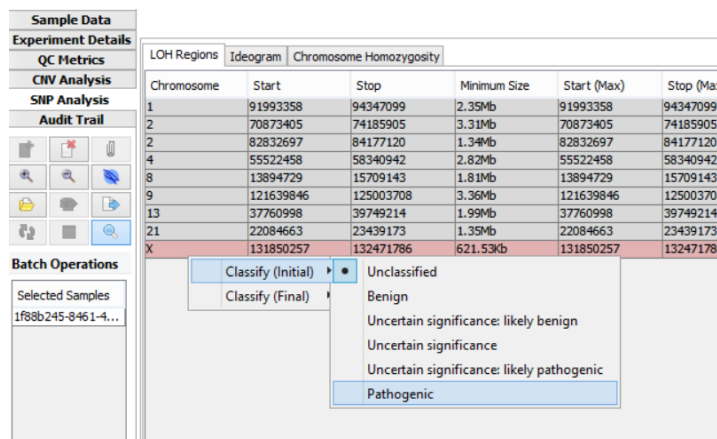


Figure: Classifying an LOH region via the Database Management tab.

As with CNV regions, when an LOH region is given a classification it is colour-coded according to the classification type with the final classification taking precedence over the initial classification.

SNP Data Zooming

It is now possible to zoom in on SNP data displayed in the **B-Allele Frequency** or **Allele Status** plots in the **Genomic View** by right-clicking within the **Overview** section and selecting **Enable SNP Data Zooming**.

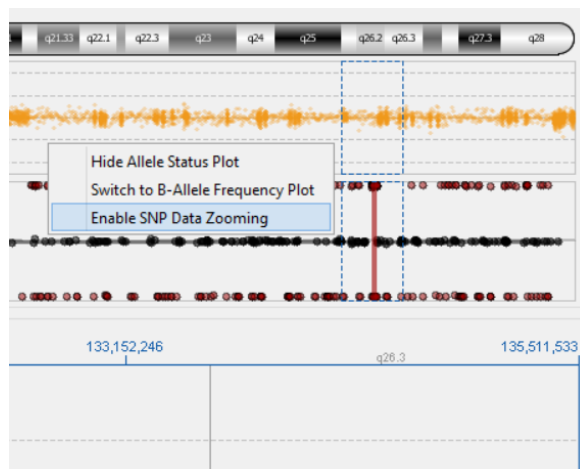


Figure: Enabling SNP Data Zooming


When enabled, the **B-Allele Frequency** and **Allele Status** plots align with the data displayed in the **Chromosome Section**, making it easier to visually identify, for example, which genes are covered by a particular LOH region. Zooming can be disabled in the same way, by right-clicking within the **Overview** section and selecting **Disable SNP Data Zooming**. This will result in the plots constantly displaying SNP data from the whole chromosome (or whole genome).

Manual LOH Region Addition

In conjunction with **SNP Data Zooming**, it is also now possible to manually add LOH regions within the **Genomic View**. This can be achieved by enabling **SNP Data Zooming**, selecting only the



region of LOH in the **Allele Status** or **B-Allele Frequency** plot and clicking the **Add** button in the **LOH Regions** tab.

 [Aneuploidy Summary](#)

CNV Regions		LOH Regions	
Chromosome	Start	Stop	Minimum Size
1	91993358	94347099	2.35Mb
2	70873405	74185905	3.31Mb
2	82832697	84177120	1.34Mb
4	55522458	58340942	2.82Mb
21	22084663	23439173	1.35Mb
X	131850257	132471786	621.53kb
8	13894729	15709143	1.81Mb
9	121639846	125003708	3.36Mb
13	37760998	39749214	1.99Mb

Figure: Manually adding and removing LOH regions.

It is also possible to manually remove LOH regions using the **Remove** button in the same section. Please note, however, that any change to LOH detection settings that trigger recalculation of LOH regions may result in a deleted region re-appearing.

LOH Region Adjustment

Further flexibility in LOH region detection has also been added by enabling manual adjustment of LOH region boundaries via the **Allele Status** and **B-Allele Frequency** plots in the **Genomic View**. To make such adjustments, simply hover the mouse over the edge of an LOH region so that the default cursor is replaced with the resize cursor and drag the LOH region to the required start/end location.

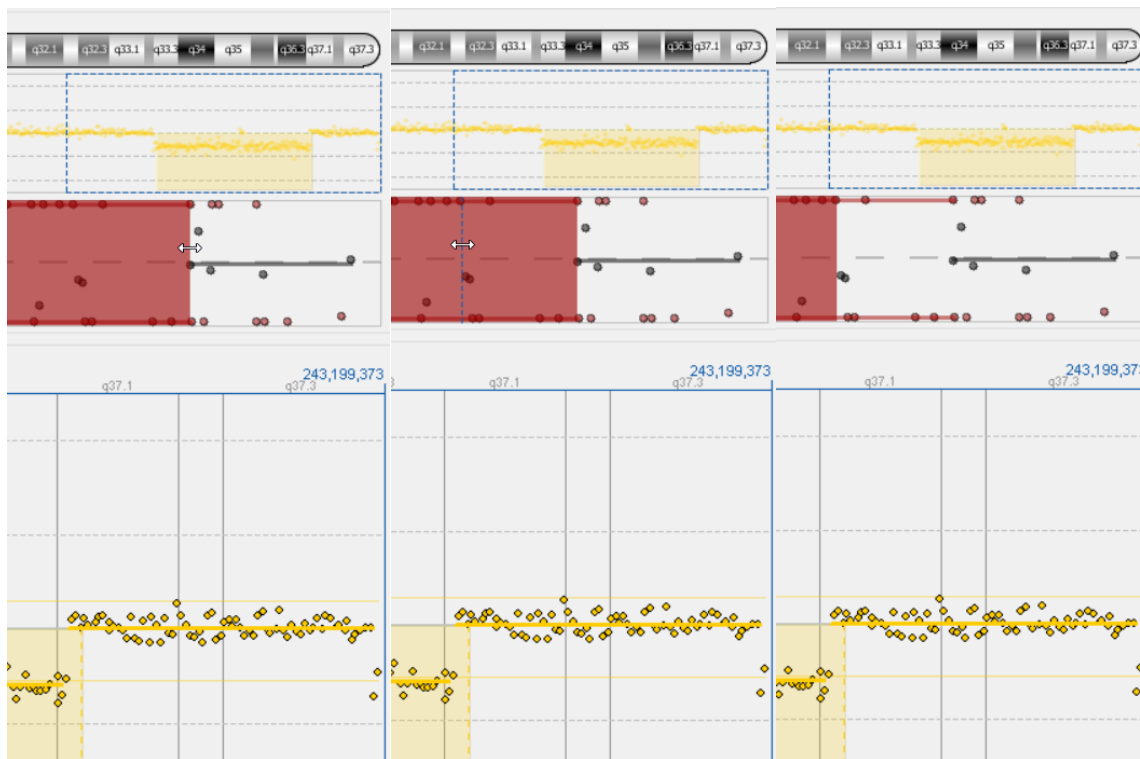


Figure: Manual adjustment of LOH region boundaries.

Please note that this functionality is also now available for CNV regions in the **Chromosome Section** view.

PGS Analysis Improvements

In order to support streamlined analysis of multiple CytoSure Embryo Screen Arrays, a new workflow has been developed and incorporated into the latest release. The workflow involves loading a pre-completed spreadsheet containing sample and reference information into CytoSure Interpret Software for batch processing, and subsequently enables convenient selection of individual samples for interpretation. Furthermore, the workflow supports analysis and visualisation of each sample against both male and female references simultaneously, simplifying sex chromosome interpretation.

Excel Planner for Batch Analysis

The Microsoft Excel PGS Planner enables users to plan a run of multiple CytoSure Embryo Screen Arrays by indicating the location and label of each sample on each slide, as well as the location of the male and female reference samples to be used for analysis. A template of the planner, **PGS_Planner.xlsx**, is located in the software installation directory (usually C:\Program Files\CytoSure Analysis Software). **Users may wish to copy the template file to a more convenient location, and make a copy for each batch of slides being processed.**

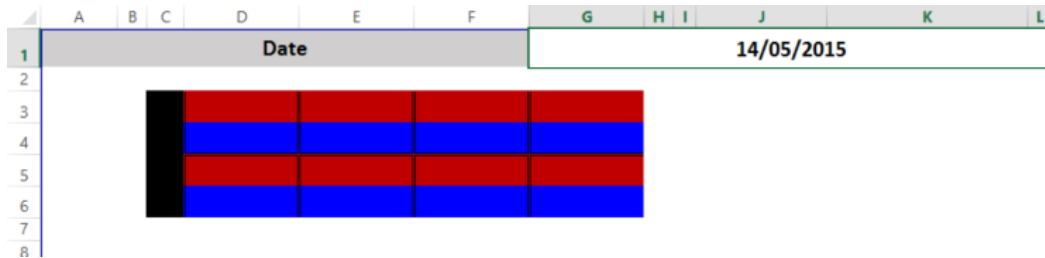
Please note the following before using the planner:



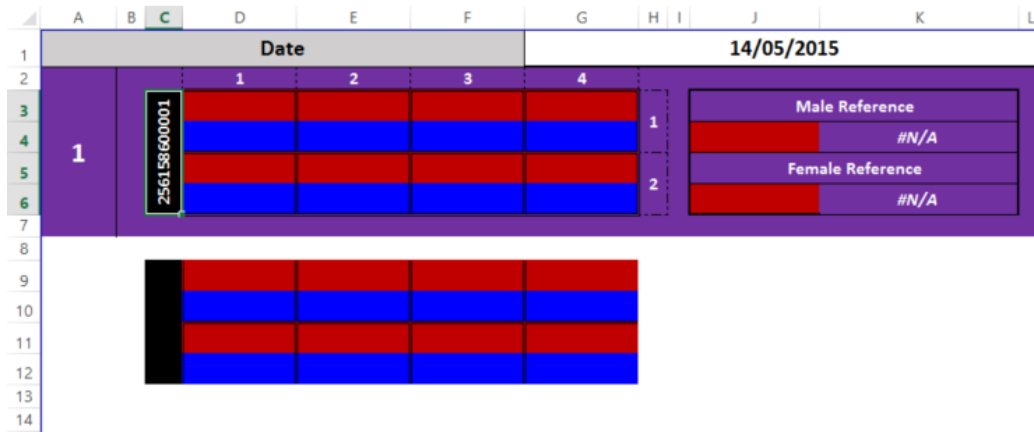
1. It is only compatible with Microsoft Excel 2013 and above.
2. The assumption is that at least one male reference and one female reference is present in the batch of slides being processed.
3. It is recommended that each slide contain one male and one female reference, but it is possible to use reference samples from one slide to analyse test samples from other slides in the planner if necessary.
4. All non-reference samples should have unique identifiers in the context of the planner.

Users should complete the planner as follows for each slide in the batch (up to a maximum of 6 slides):

1. Enter the date of processing in the corresponding cell.



2. Enter the slide barcode in the **black** cell.



3. In the **red** cells, enter the identifier of the sample labelled in **Cy3** in the corresponding array of the slide.
4. In the **blue** cells, enter the identifier of the sample labelled in **Cy5** in the corresponding array of the slide.



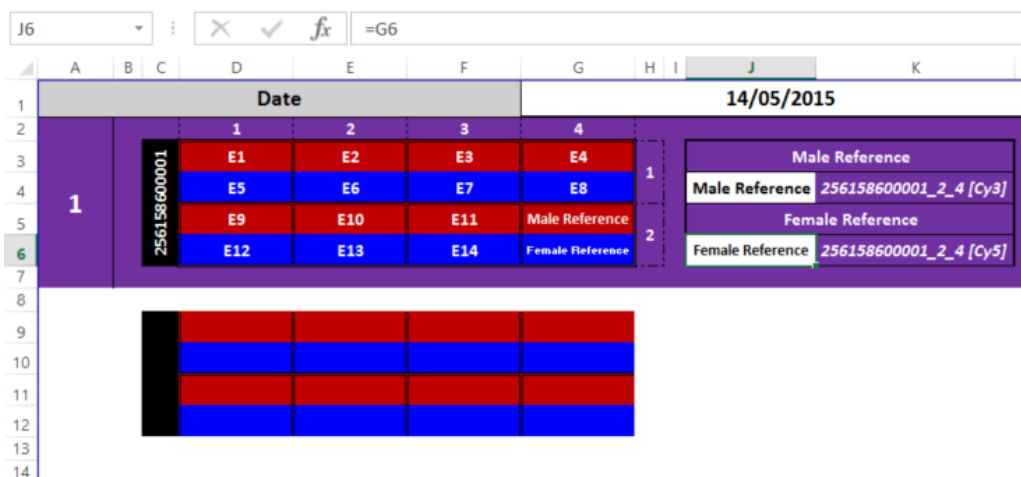
Date				14/05/2015			
1	256158600001	1	2	3	4	1	Male Reference
		E1	E2	E3	E4		#N/A
		E5	E6	E7	E8	2	Female Reference
		E9	E10	E11	Male Reference		#N/A
		E12	E13	E14	Female Reference		

5. Select the empty cell directly beneath **Male Reference**, type "=", select the cell within the slide containing the sample to be used as the male reference and press Enter.

J4 : =G5

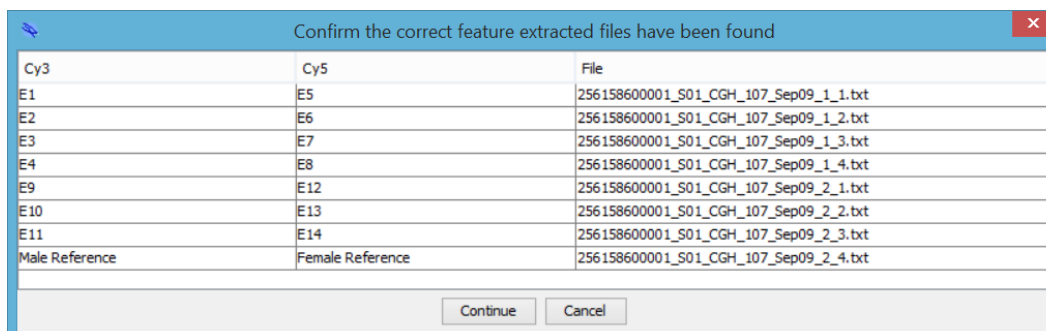
Date				14/05/2015			
1	256158600001	1	2	3	4	1	Male Reference
		E1	E2	E3	E4		Male Reference 256158600001_2_4 [Cy3]
		E5	E6	E7	E8	2	Female Reference
		E9	E10	E11	Male Reference		#N/A
		E12	E13	E14	Female Reference		

6. Select the empty cell directly beneath **Female Reference**, type "=", select the cell within the slide containing the sample to be used as the male reference and press Enter.



When this has been completed for all slides, save the file and make a note of the file location. The planner can now be used to guide lab work and, once the corresponding feature extracted files have been generated, automate processing of the data in CytoSure Interpret Software as follows:

1. Select **Tools -> PGS analysis -> With worksheet.**
2. Select the PGS Planner from the file system and click **Open.**
3. Select the folder containing the feature extracted files associated with the slides described in the PGS Planner and click **Open.** Please note that the software will recursively search subfolders for the files if necessary.
4. In the window that subsequently appears, check that all samples have been associated with a feature extracted file.
 - a. If any samples are missing, check that their corresponding feature extracted files have been generated and are located within the selected folder. If necessary, click **Cancel.**
 - b. If all samples are present and associated with a feature extracted file according to the table, click **Continue.**



5. The software will now process each sample according to the information contained in the selected PGS planner as follows:
 - a. A file containing the sample and male reference data is generated.
 - b. A file containing the sample and female reference data is generated.



- c. Both files are processed and saved as one CytoSure Interpret .cgh file containing both data sets.
- d. The sample database is updated with the location of this combined, processed file ready for subsequent interpretation.

The user is notified when processing is complete on all samples and advised on how to complete the interpretation process. This process is also described in the following section.

Flexible Sample Selection

In order to allow flexible interpretation, samples processed from a PGS Planner are stored in the sample database rather than being immediately loaded into CytoSure Interpret Software. They are then available for further interpretation via the **Database Management** tab, where the user can select a subset of the processed samples for interpretation. To analyse a set of samples, select the corresponding check boxes in the Sample table and select **Batch Operations -> PGS -> Open PGS Data**, and click **Yes** when prompted.

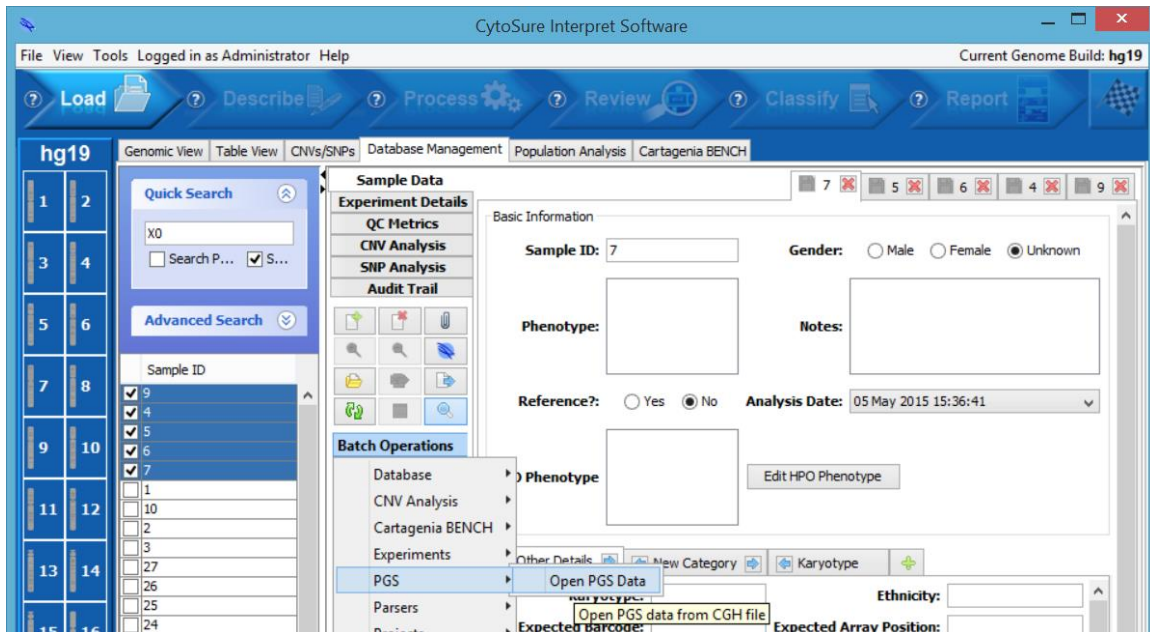


Figure: Importing PGS data for interpretation from the Database Management tab

The selected data sets will be loaded into the software ready for interpretation in the **PGS Analysis** tab.

Dual Reference Analysis

In order to simplify sex chromosome interpretation, the software now supports simultaneous visualisation of PGS data against both male and female reference samples. By default, if the data has been generated by processing a PGS Planner, both the male and female reference data will be displayed in the **Histogram, Whole Genome View**, segmentation panel and CNV table panel.

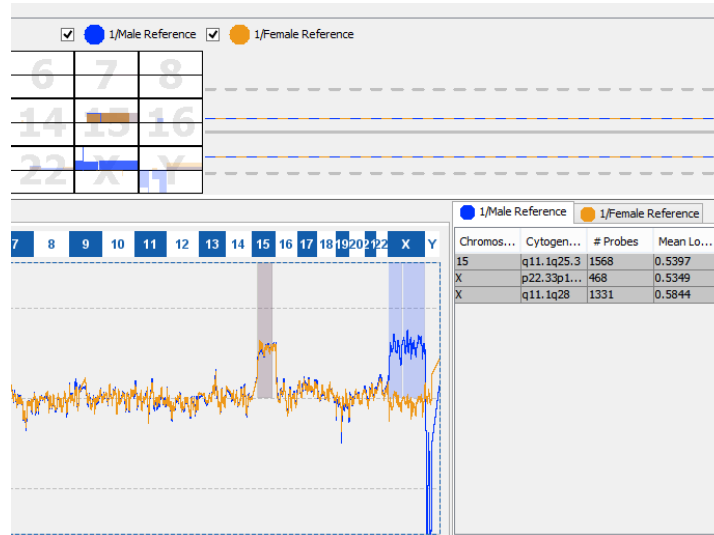


Figure: A sample visualised against both male and female references

Users can toggle the display of individual reference samples using the check boxes provided in the middle panel. Please note that the interpretations displayed in the sample table are for the male reference data only, but users can manually adjust individual chromosome interpretations for each sample as described [here](#).

Annotation Reporting Updates

A number of features have been added to enable more flexible display and reporting of annotation track content.

Custom Annotation Reporting

Aberration reports can be configured to include the names of any annotations overlapping detected CNV regions on a track-by-track basis. However, it is now possible to customise the text in the report to include other information related to the annotation as well as, or instead of, its displayed name. For example, as all annotations in the default **Gene** track (taken from Ensembl) include a “Gene Biotype” custom field (all fields can be displayed by hovering over an annotation), CytoSure Interpret Software can be configured so that both pieces of information are included when the **Genes** field is part of the report.

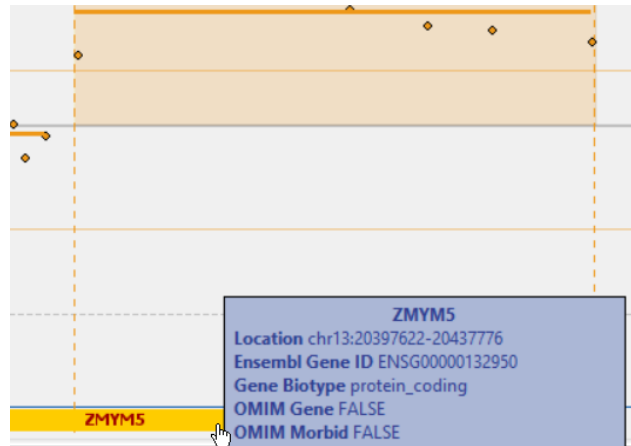


Figure: Gene information available for inclusion in reports.

To do this:

1. Enable reporting of the annotation name by selecting **Tools -> Options... -> Annotation -> CNV Reporting** and ensure that the **Report** column of the annotation track is checked.
2. Right-click in the same cell of the table as the checkbox and select **Advanced Options**.

Name	Initial	Colour	Report	Report #	Report %
Gene	G	[Blue]	<input checked="" type="checkbox"/>		<input type="checkbox"/>
Exon	E	[Purple]	<input type="checkbox"/>		<input type="checkbox"/>
Copy Number Vari...	V	[Red]	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
Segmental Duplica...	S	[Orange]	<input type="checkbox"/>		<input type="checkbox"/>
DECIPHER Patient	D	[Green]	<input type="checkbox"/>		<input type="checkbox"/>
ISCA Variation	I	[Light Blue]	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Figure: Accessing advanced CNV Reporting options for annotation tracks.

3. In the **Customise annotation information** section of the **Advanced annotation options** window, enter the format to be used to report overlapping annotations in the selected track using the same dynamic field codes described [here](#). For example, to report the name of the annotation and the value of the “Gene Biotype” field, separated by a hyphen, set this value to “{#name} – {Gene Biotype}”.

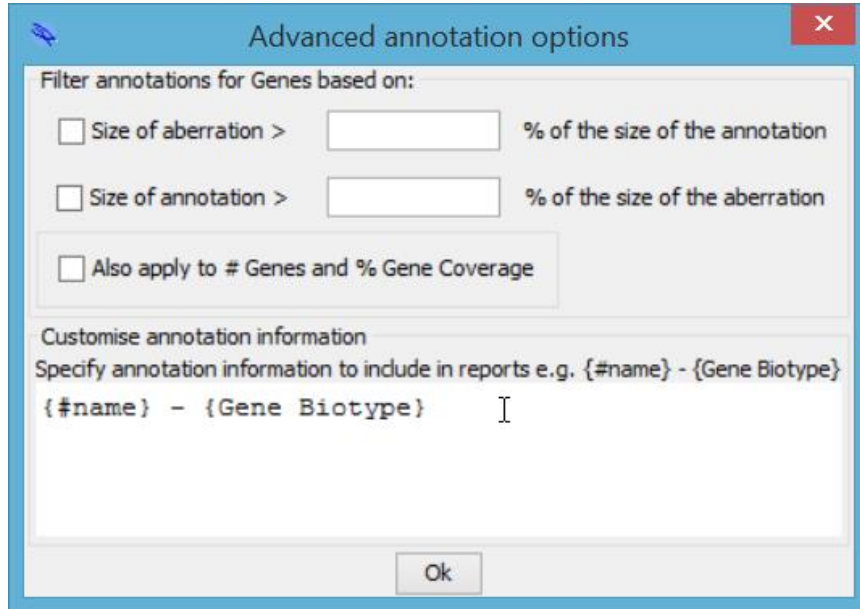


Figure: Setting the annotation reporting format

When reports are subsequently generated that have been configured to include the names of overlapping annotations in the track (either in the **Aberration Info** or **Summary Columns** section of the **Report Generation** tab in the **Options** window), they should now be displayed in the specified format.

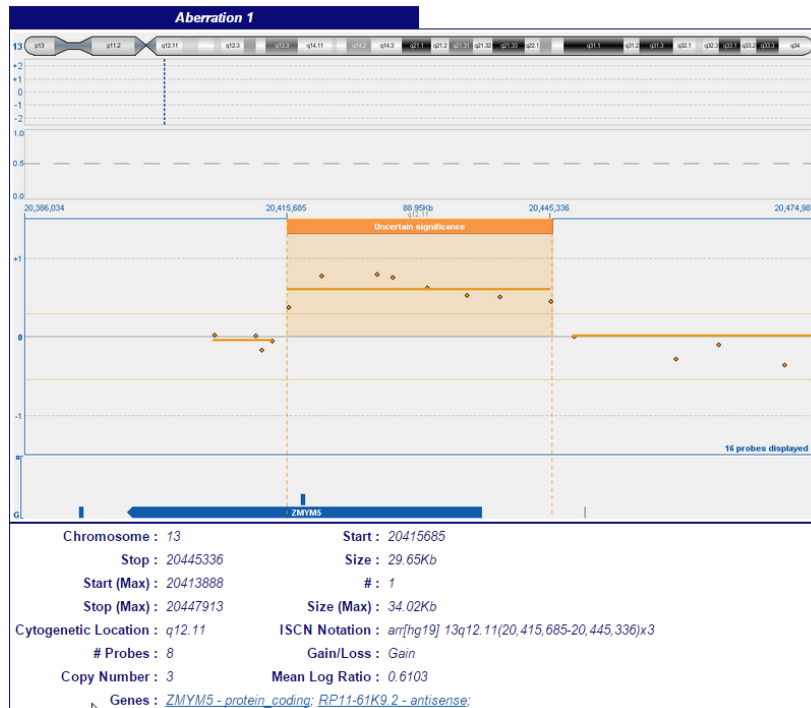




Figure: Custom formatting of overlapping annotation information.

Probe-Annotation Overlap Reporting

The **Table View** tab contains information for all probes currently displayed in the **Genomic View**, including probe name, location, signal intensities and GC content. This table can now be customised to also include the names of any annotations which over the probe (e.g. genes). Configuration can be carried out as follows:

1. Enable reporting of the annotation name by selecting **Tools -> Options... -> Annotation -> CNV Reporting** and ensure that the **Report** column of the required annotation track is checked.
2. In the **Table View** tab, right-click on the table header and select the required column from the pop-up menu displayed.

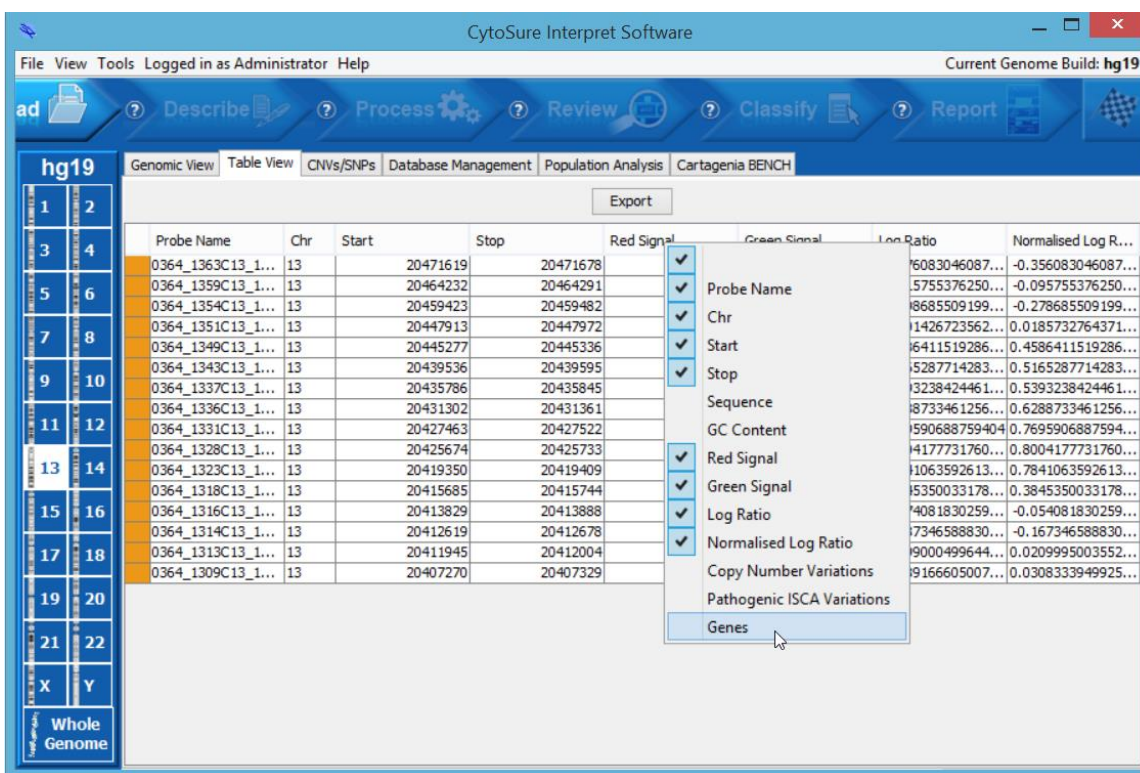


Figure: Adding annotation overlap information columns to the Table View.

If a large number of probes are displayed in the **Genomic View**, a large number of calculations will need to be made to determine which annotations overlap with the probes in the table. In order to reduce the time for these calculations to be made, load the required annotation tracks into memory by right-clicking on the track in the **Genomic View** and selecting **Load into memory**.

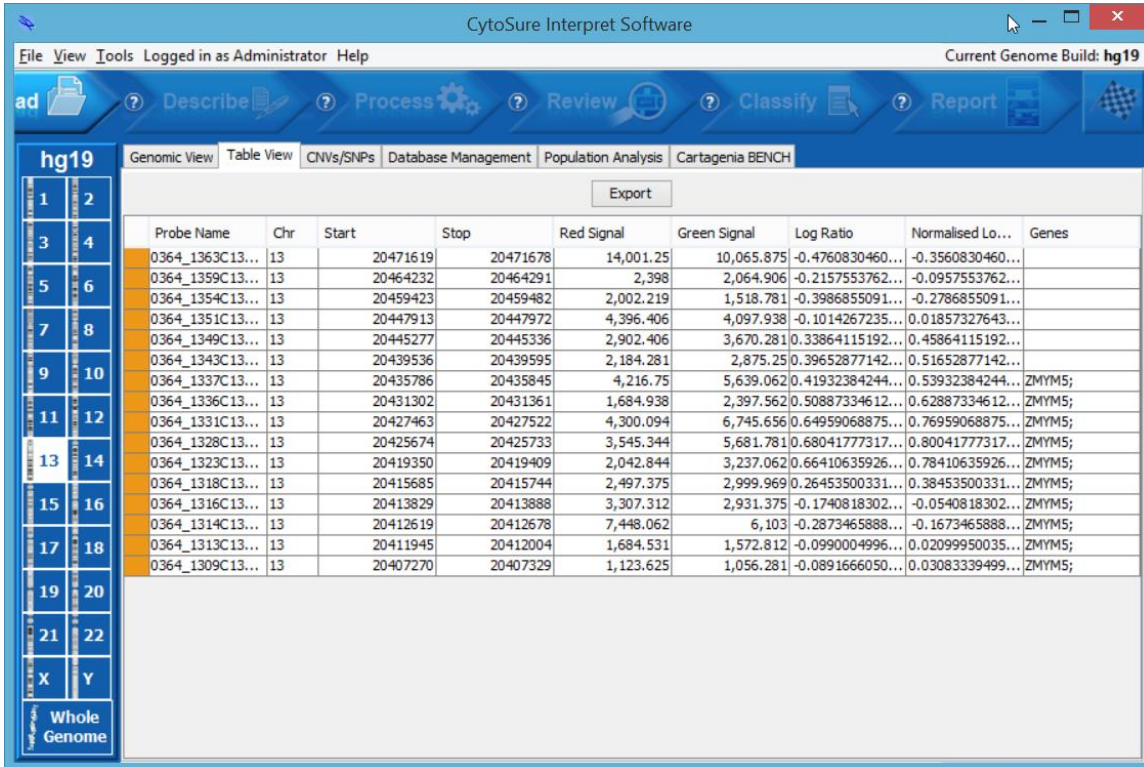


Figure: The Genes column added to the Table View.

Custom Annotation Information Display

The Aberration annotation track contains CNVs detected in previously analysed cases, and hovering over a CNV in the track results in the display of a number of pieces of information relating to the selected CNV. Previously, a specific set of information was displayed with no means to modify the included fields, but the latest version of CytoSure Interpret Software now enables customisation of the content of this “tooltip”.

Customisation can be carried out by right-clicking within the Aberration track and using the **Tooltip Fields** menu to individually select the information that should be displayed when the user hovers the mouse over an annotation in this track.

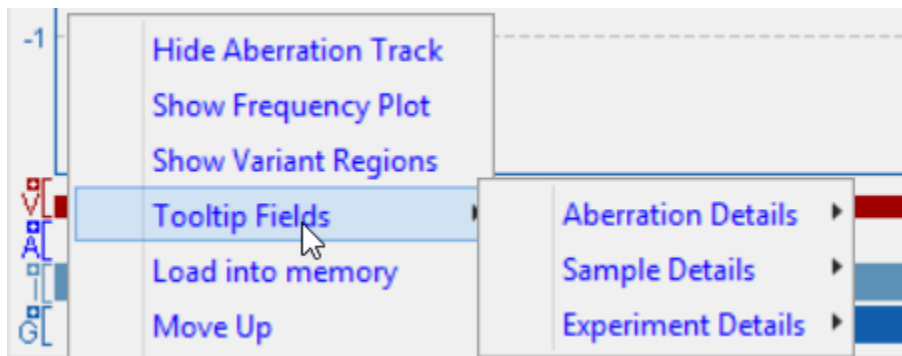




Figure: Aberration track tooltip customisation.

Fields are grouped into three categories – **Aberration Details**, **Sample Details** and **Experiment Details** – and users are able to select a number of fields for inclusion in the tooltip to make the view as useful as possible within their analysis workflow.

Miscellaneous

- Enabled complete removal of summary table from reports (remove all fields from **Tools -> Options... -> Report Generation -> Summary Columns**).
- Enabled use of alternative identifier for submission to Cartagenia Bench (**Cartagenia BENCH -> Accounts -> Sample Identifier**).
- Enabled hiding of all track in the **Genomic View** with one click (right-click followed by **Hide all Tracks**).
- Enabled omission of non-UPD chromosomes from the chromosome homozygosity section of the report (**CNVs/SNPs -> Detected CNVs/SNPs -> Options... -> LOH Calling -> Exclude non-UPD chromosomes from reports?**).
- Updated all standard annotation tracks with latest information and hg38 content.
- Enabled visualisation of the hierarchy of Auto-Classification classifiers (**Tools -> Options... -> Classification -> Auto-classification Settings -> Advanced Settings... -> right-click on classifier -> View**).
- Enabled “Unknown” sample gender selection.



CytoSure™ Interpret Software update — version 15.0 November 2014

The following new features are available in CytoSure Interpret Software version 4.7:

Enhancements for PGS analysis	60
Support for Single Channel PGS analysis	60
Simplified standard workflow	64
Customisable interpretation and reporting	64
The Analysis tab	65
The Reporting tab	68
Other Controls	69
Sample vs sample CNV detection	70
Classification archiving	70
Automatic security lockout	72
Miscellaneous updates	74



Enhancements for PGS analysis

A number of new features have been added to version 4.7 of CytoSure Interpret Software in order to improve support for streamlined analysis of data generated from OGT's CytoSure Embryo Screen array. Analysis of such data, which is often of lower quality when compared to constitutional samples, involves identifying large CNVs or aneuploidies across a number of samples in parallel. These requirements are supported by a dedicated PGS workflow in CytoSure Interpret Software which is described in the following sections.

Support for Single Channel PGS analysis

Up to fourteen embryos (plus two reference samples) may be screened on a single slide and automatically analysed in the software by combining reference sample data with the appropriate embryo data before carrying out analysis. The Single Channel analysis process works as follows:

1. Select **Tools -> Single Channel PGS analysis**.
2. In the **Select Feature Extracted Files** window, click the **Add** button and select the feature extracted files generated from the embryos and reference samples to be included in the analysis.

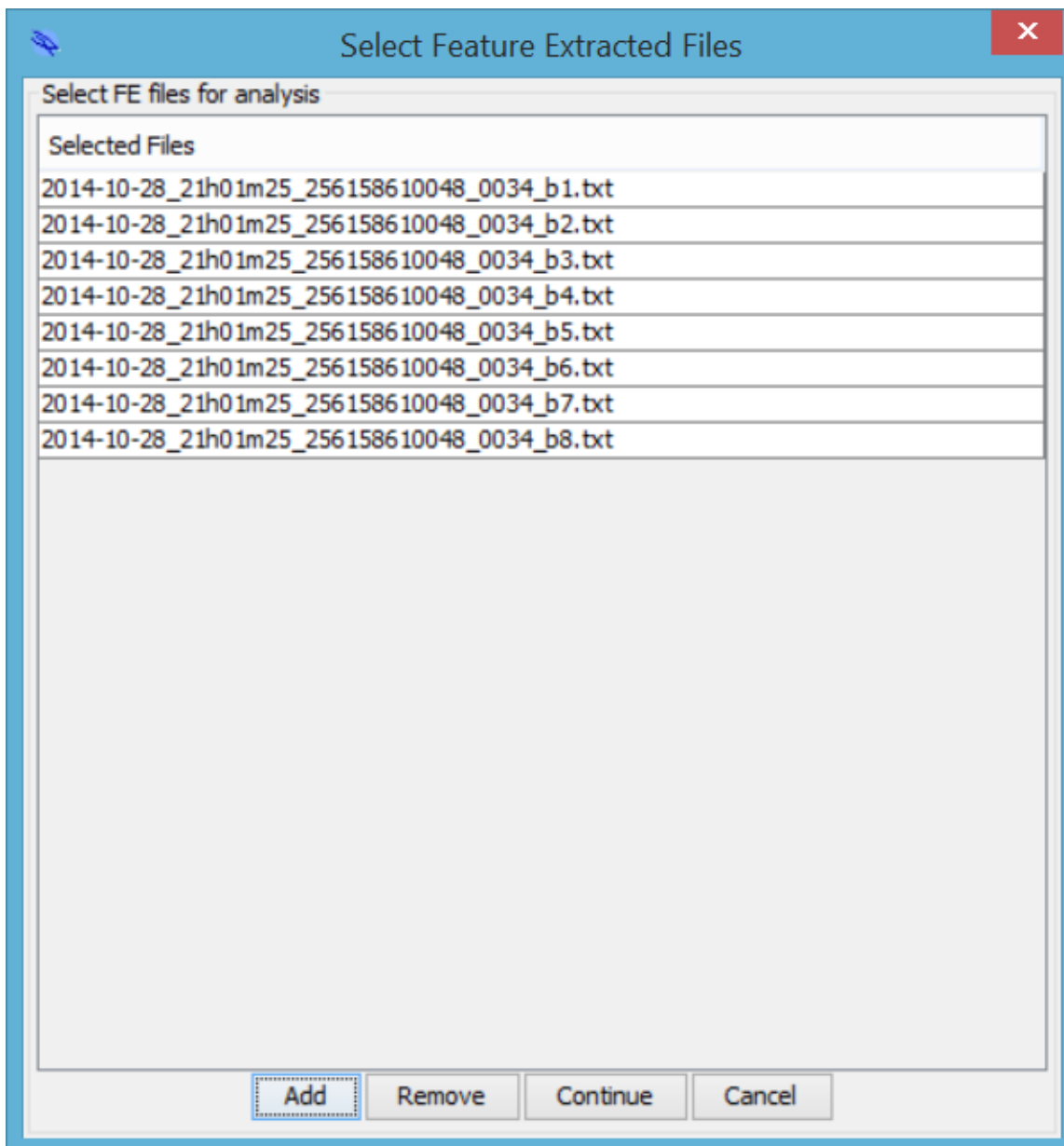


Figure: Selecting feature extracted files for Single Channel PGS analysis

3. When all required feature extracted files have been added to the list, click **Continue**.
4. In the **References** window, indicate which of the selected feature extracted files contains the reference sample for each label. Please note that there is no requirement to select both references from the same feature extracted file (i.e. the same array). If no reference data exists in the selected feature extracted files for a particular label, select the **Not Present** radio button for that label (embryos assigned the alternative label will not be analysed). When the selection is correct, click **Continue**.

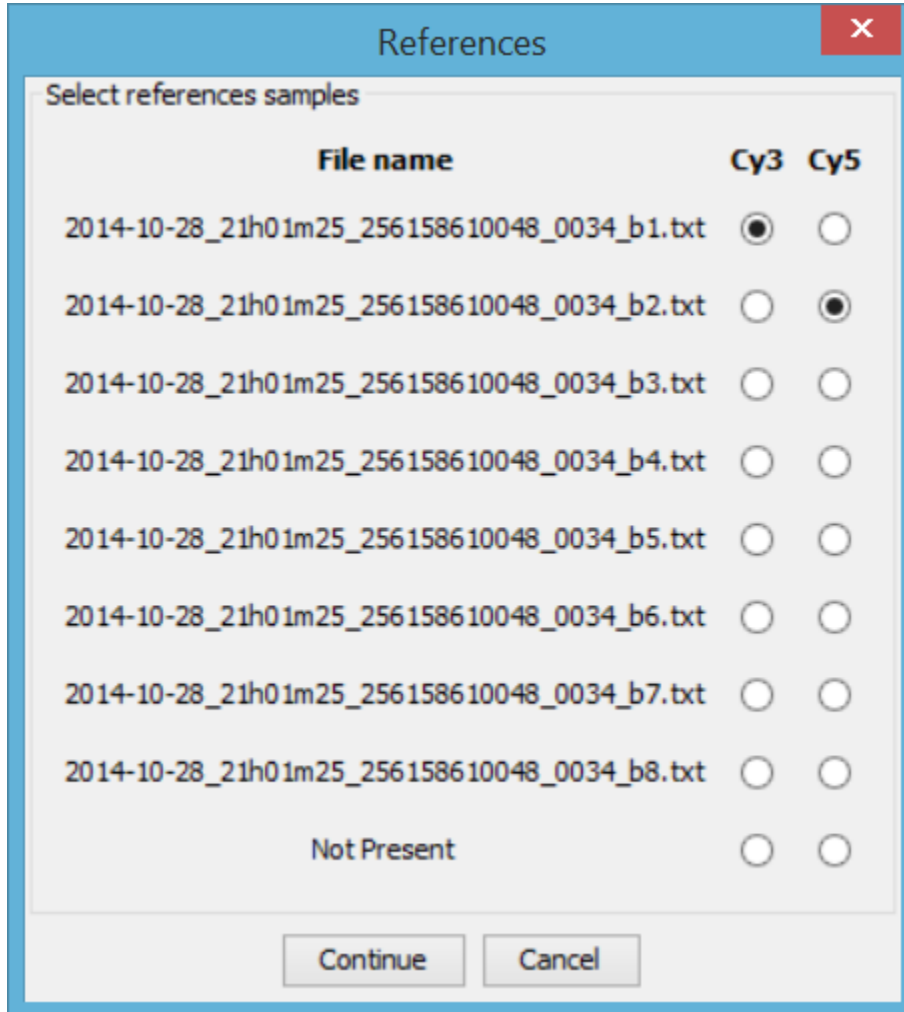


Figure: Indicating the location of the reference samples. In the above example, the Cy3 and Cy5 references are shown to be located on different arrays.

5. In the **Select Reference Samples** window, indicate the identifiers of the reference samples, creating new reference samples in the database if necessary, then click **Continue**.

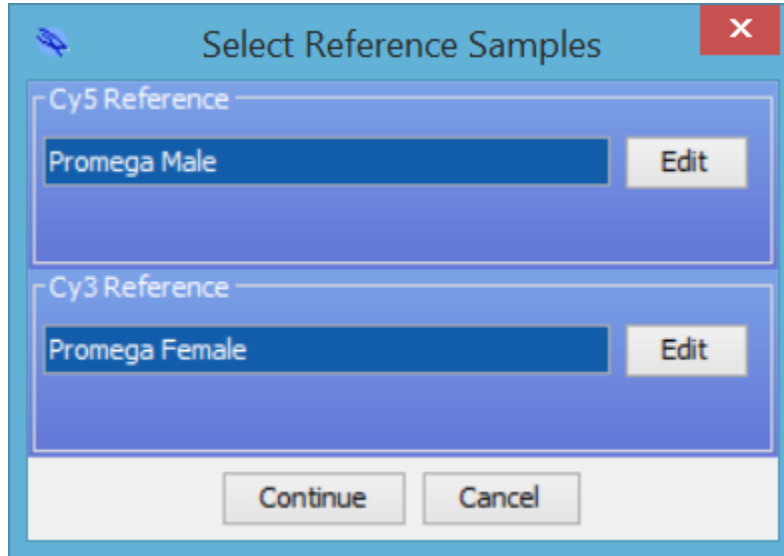


Figure: Indicating the reference sample identifiers.

6. In the **Test samples** window, indicate the identifiers of each embryo to be analysed (data to which no identifier has been allocated will not be analysed), creating new identifiers in the database if required. When complete, click **Continue**.

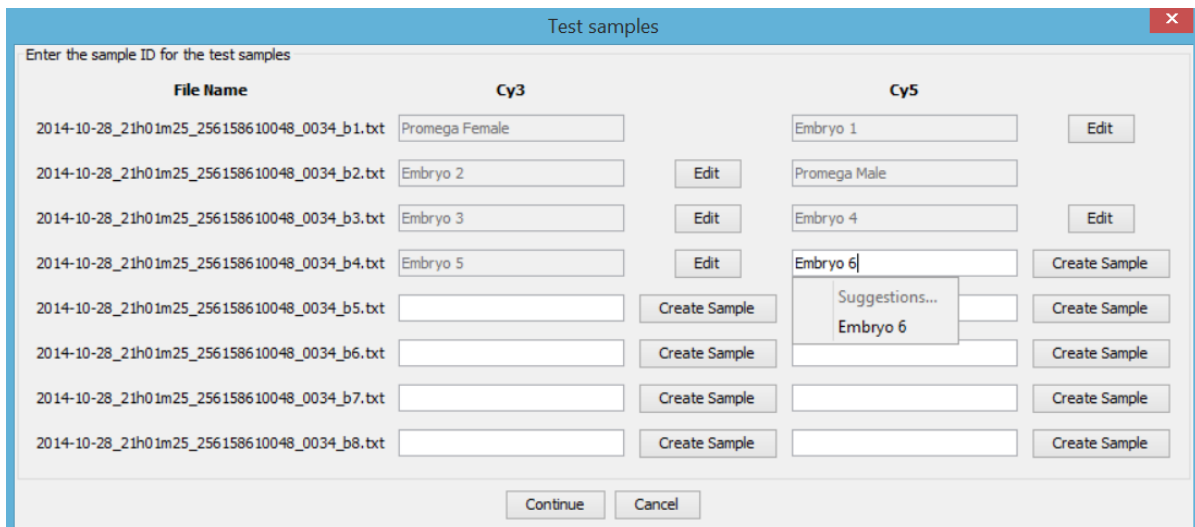


Figure: Indicating the embryo identifiers.

The software will then combine the embryo data with the appropriate reference data to create new feature extracted files, which are stored in the **Default Save Location (Tools -> Options... -> Files -> Default Save Location)**, and carry out the standard PGS analysis workflow on all selected embryos. If subsequent analysis is required on the same embryos, importing the combined feature extracted files and following the standard workflow will avoid a number of the steps described above.



Simplified standard workflow

The standard workflow for analysing multiple feature extracted files consisting of data from one embryo and one feature extracted file is described below. This process should be used following generation of data from a slide containing 8 embryos run against 8 reference samples, or when selecting feature extracted files that have been generated by running the Single Channel PGS Analysis to combine embryo data with reference sample data from different feature extracted files.

1. Select **File -> Import -> PGS Data**.
2. Select the feature extracted files containing the data to be analysed and click **Open**.
3. In the **Assign Sample IDs** window, assign identifiers to both the sample and the reference on each of the selected feature extracted files.

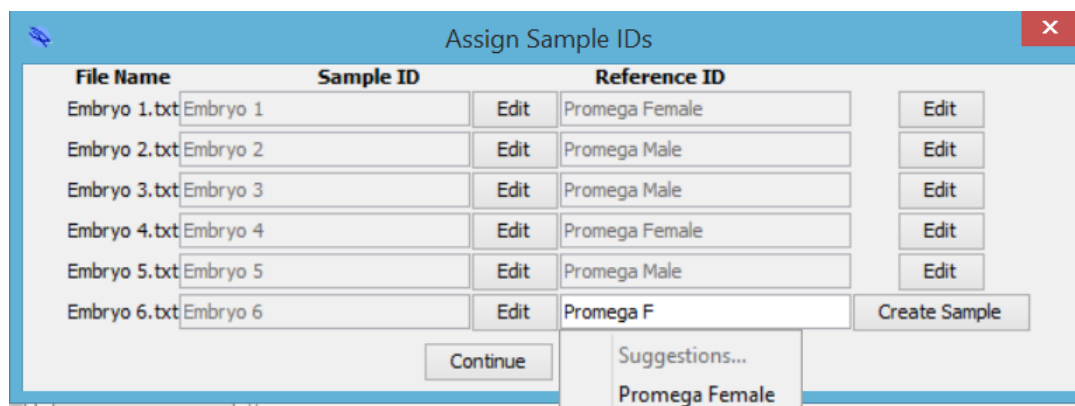


Figure: Assigning the sample and reference identifiers to selected feature extracted files generated from the Embryo Screen Array.

Segmentation and normalisation are then carried out on the data, and large CNVs and aneuploidies are automatically identified, finally resulting in the display of the **PGS Analysis** tab, within which all analysis and interpretation may be completed. The content of this tab is described in the next section.

Customisable interpretation and reporting

The **PGS Analysis** tab, displayed after data from the CytoSure Embryo Screen array is loaded into CytoSure Interpret Software, contains a number of intuitive visualisations of the data in order to aid streamlined interpretation of the results for all users. The sections of the interface provide information to help the user determine whether each embryo is normal or abnormal, and this information can then be used to generate a summary report.

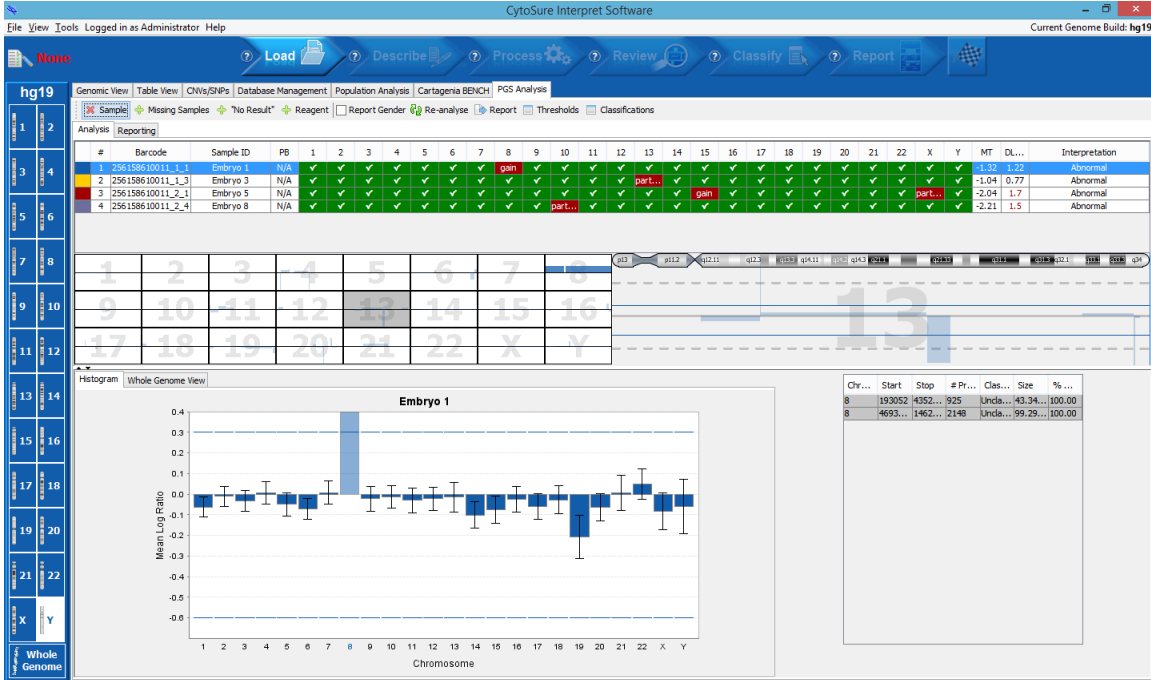


Figure: The PGS Analysis tab.

The interface is split into two tabs: **Analysis** and **Reporting**.

The Analysis tab

The **Analysis** tab itself contains a number of smaller sections that display the data in different ways to aid interpretation.

The sample table

The section at the top of the tab is the sample table, which contains the list of embryos being analysed along with the interpretation of each chromosome, QC metrics and an overall interpretation for the embryo.

#	Barcode	Sample ID	PB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	MT	D...	Interpretation	
1	256158610011_1_1	Embryo 1	N/A	✓	✓	✓	✓	✓	✓	✓	gain	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-1.32	1.22	Abnormal
2	256158610011_1_3	Embryo 3	N/A	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	pa...	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-1.04	0.77	Abnormal
3	256158610011_2_1	Embryo 5	N/A	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	gain	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-2.04	1.7	Abnormal
4	256158610011_2_4	Embryo 8	N/A	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	pa...	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-2.21	1.5	Abnormal

Figure: The sample table in the Analysis tab

Following segmentation and normalisation, CytoSure Interpret Software attempts to automatically indicate the status of a chromosome based on the mean log ratio of probes in the chromosome (to identify monosomy or trisomy) and the segmentation (to identify large CNVs). If no such abnormalities are detected, the chromosome cell displays a white tick (✓) on a green background. Otherwise, the chromosome cell will be given a red background and will display appropriate text describing the abnormality. This text can be customised by the user by clicking on the **Classifications** button at the top of the screen, but defaults to the following:



- Gain/Trisomy: “gain”
- Loss/Monosomy: “loss”
- Partial Gain: “part gain”
- Partial Loss: “part loss”

Although these classifications are automatically allocated by the software based on the data, the user is able to modify all chromosome classifications if necessary, based on the information provided in the other **Analysis** tab sections, by clicking on the relevant cell and selecting the appropriate option from the drop-down box. Modified classifications are indicated with a white border. In addition to the “Normal” classification and those described above, the user has the option to indicate that the interpretation is “Unknown” by select the “?” option.

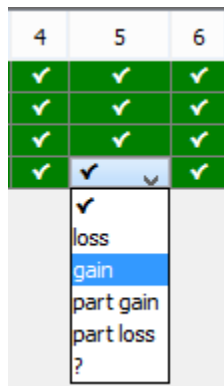


Figure: Manual chromosome classification via the drop-down box.

The **Interpretation** column indicates the overall interpretation for the embryo based on the individual chromosome classifications as follows:

- **Normal** – If all chromosomes are classified as normal (✓).
- **Abnormal** – If any chromosomes are classified as abnormal (**gain, loss, part gain, part loss**).
- **Unknown** – If no chromosomes are classified as abnormal and at least one is classified as unknown (?).

All embryos need an overall interpretation of either **Normal** or **Abnormal** for a report to be generated.

The segmentation grid

Beneath the sample table is a grid indicating the segmentation of the data for the embryo(s) selected in the table. Significant CNVs, as determined by the aberration detection thresholds in **Tools -> Options... -> Aberration Detection**, are more prominently displayed, and the section to the right of the grid displays a larger version of the selected chromosome.

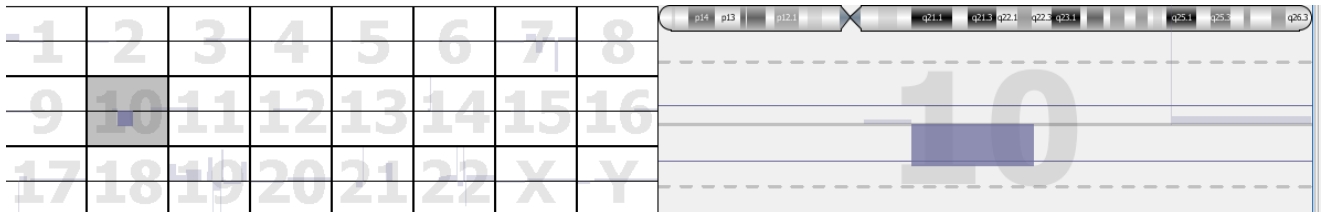


Figure: The segmentation grid, indicating the presence of a significant CNV in 10q.

Clicking on a segment in the larger view results in the display of the region in the **Genomic View** tab.

The Histogram

The **Histogram** tab at the bottom of the **Analysis** tab indicates the mean log ratio and log ratio standard error of the probes within each chromosome of the embryos selected in the sample table. Additionally, the gain/loss calling thresholds are indicated with lines at the appropriate positions, and any values exceeding these thresholds will be highlighted.

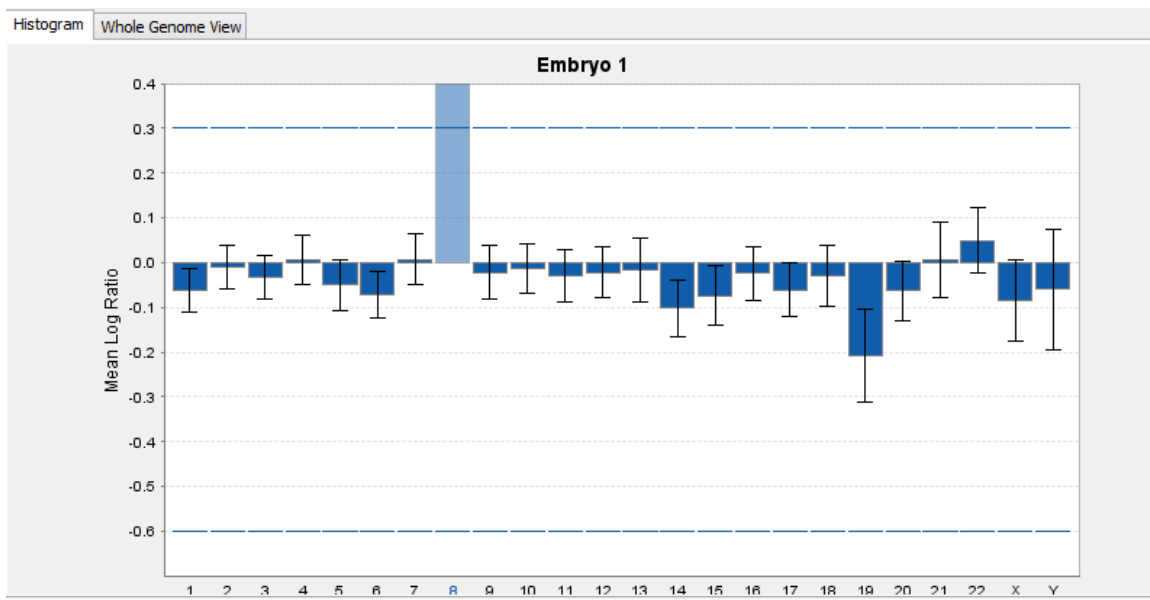


Figure: The Histogram tab, indicating trisomy 8.

This data corresponds directly with the trisomy/monosomy calls indicated in the sample table, but does not display partial gain or loss of any chromosome. If necessary, the thresholds for each chromosome can be adjusted via the **Thresholds** button at the top of the **PGS Analysis** tab.

The Whole Genome View

The **Whole Genome View** tab displays the moving average log ratio of all chromosomes and highlights the significant CNV regions in the data for the embryos selected in the sample table.

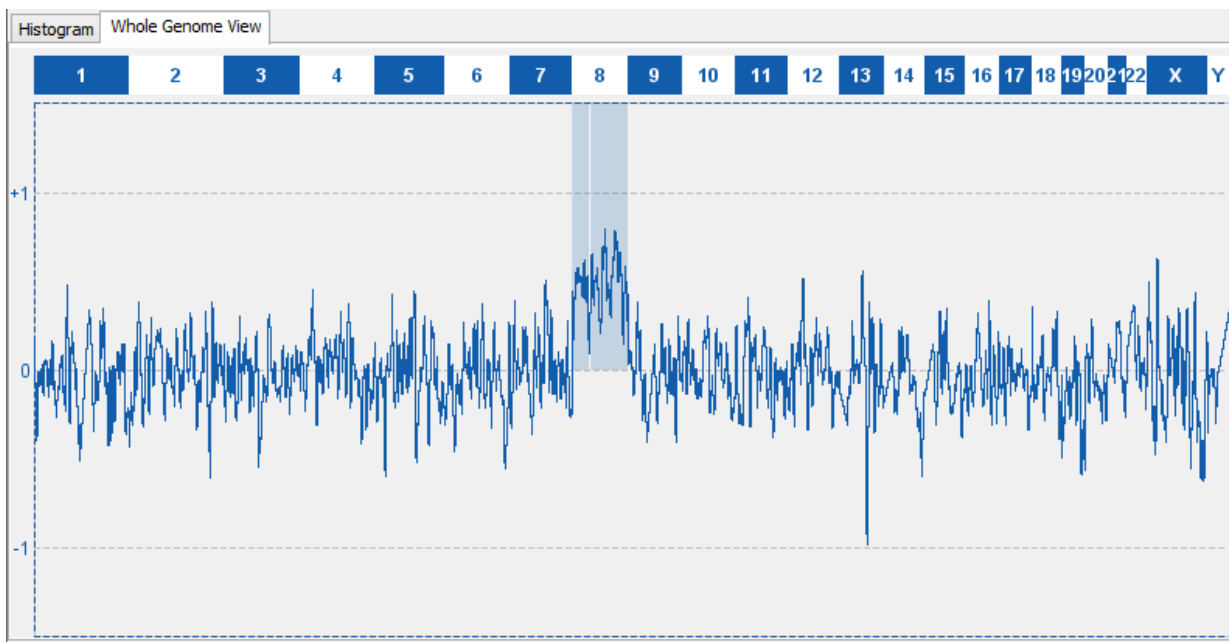


Figure: The Whole Genome View tab indicating gains on chromosome 8.

The CNV table

The table to the right of the **Histogram** and **Whole Genome View** tabs lists the significant CNVs detected in the first of the embryos selected in the samples table. The list directly corresponds with the partial gain/loss classifications in the sample table, but is superseded by a trisomy or monosomy classification. The table's columns can be modified by right-clicking on the table header and selecting or deselecting columns as required. Clicking on a CNV in the table results in the display of the CNV region in the **Genomic View**.

Chromo...	Start	Stop	Size	Gain/Loss
13	37563772	60086662	22.52Mb	Gain

Figure: The CNV table, listing a gain on chromosome 13.

The Reporting tab

Once the user has made their interpretation, a report may be generated by clicking on the **Report** button at the top of the **PGS Analysis** tab. The contents of this report can be configured via the **Reporting** tab.



Analysis Reporting

Report Content

Report Title: Preimplantation Genetic Screening of Numerical Chromosome Anomalies

Report Subtitle:

Sample Data

Patient ID: Patient Name: Procedure ID:

Reason For Referral: Referred By: Date Of Biopsy:

Specimen Type: Blastomere Date Of Birth: IVF Clinic:

Date Specimen Received: Centre Code: Billing Code:

Biopsy Performed By: Tubing Performed By: Amplification Performed By:

aCGH Performed By: Analyst: Witness:

Analysed By: Witnessed And Approved By:

Reagents

Cot1 DNA: DAPI: LSI buffer:

NP40: NT buffer: NT enzyme:

NT water: Polymerase: Primer:

SpectrumGreen: SpectrumRed: dATP:

dCTP: dGTP: dTTP:

Figure: The Reporting tab, enabling configuration of report content.

The Reporting tab is split into three sections:

1. **Report Content** – Enables the user to modify the **Report Title** and **Report Subtitle**. These settings will be remembered by the software for subsequent analyses.
2. **Sample Data** – Provides the means to enter various sample and experiment information. Should be completed for each new analysis.
3. **Reagents** – Customisable list of reagents used in processing the embryos. To add new types of reagent to the list, click on the **Reagents** button at the top of the **PGS Analysis** tab. The software will remember the previous lot numbers of all reagents.

All non-empty fields will be included in the report that can be generated via the **Report** button.

Other Controls

The following buttons at the top of the **PGS Analysis** tab provide additional functionality:

- **Remove Sample** – Removes the selected sample from the analysis.
- **Add Missing Samples** – If the **Specimen Type** is set to **First and Second Polar Bodies** in the **Reporting** tab, this button will add a “No Result” sample for any polar body missing its corresponding polar body.
- **Add “No Result”** – Add a “No Result” sample. This can be used to indicate the existence of other embryos that weren’t deemed to be of sufficient quality to be analysed by the CytoSure Embryo Screen array.
- **Add Reagent** – Add a new reagent type to the list in the **Reporting** tab.
- **Report Gender** – Check this box to indicate that the classification of the X and Y chromosomes should be reported together.
- **Re-analyse** – If significant parameters of the analysis are modified (e.g. aberration detection settings, sample gender or sample label) clicking this button will update the automatic chromosome classifications accordingly. All manual classification will be lost.



- **Report** – Generate a PDF report of the results, including sample, experiment and reagent information, if entered by the user.
- **Thresholds** – Modify the chromosome mean log ratio thresholds for detecting aneuploidy.
- **Classifications** – Modify the names of the aneuploidy and CNV classifications.

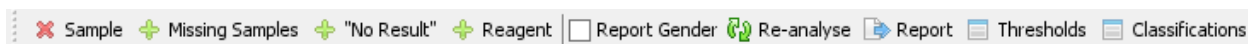


Figure: The PGS Analysis tab buttons panel

Sample vs sample CNV detection

CytoSure Interpret Software v4.7 includes additional functionality to support the correct association of CNVs with samples when no reference is used in the experimental design. In order to achieve this, the signal intensity of probes within each channel of all loaded datasets within the CNV region are displayed in a graph, helping the user to identify the sample producing dissimilar signal intensities in the region, and thus the source of the CNV.

The analysis process for such an experimental design works as follows:

1. Load multiple (at least 2) datasets from the same slide into the software and run segmentation and aberration detection if necessary.
2. Ensure that the **Sample ID** and **Reference ID** are correct as determined by the **Sample Label**.
3. For each CNV on each dataset:
 - a. Right-click on the aberration in the **Genomic View**, either in the **Chromosome Section** or the **CNV Regions** table, and select **View Signal Intensity Comparison**.
 - b. The **CNV Signal Intensity Comparison** window will appear containing a graph representing the signal intensities of each probe in each dataset for both Cy3 and Cy5. The graph for the current Sample Label and the sample to which the CNV is currently associated are highlighted.
 - c. Using the information displayed in the graphs, either **Confirm association** with the sample or **Remove** the CNV. Association with the sample is manifested by a significant difference in probe signal intensity for the sample in the current Sample Label when compared to the other samples. If there is little difference in the signal intensities, but one is present in the signal intensities of the “reference”, this is a good indication that the CNV is associated with the other sample on the array.
4. When all datasets have been analysed for the current Sample Label, carry out report generation, file saving and database submission as required.
5. To enable analysis of the CNVs on the other samples on each array, select **Tools -> Swap Samples and References**. For each dataset, the Sample ID and Reference ID will be swapped, the alternative **Sample Label** will be selected and aberration detection will be re-applied to generate a new CNV list.
6. To complete analysis, follow steps 1-4 above.

Classification archiving



In some cases, classifications used by a lab to label detected CNVs have changed as analysis and interpretation processes have developed, resulting in an unwieldy list of classification types, many of which are obsolete. Although it is possible to remove classification types from the system altogether, this requires re-assigning affected CNVs with one of the remaining classifications, which may not be appropriate.

In order to enable maintenance of a manageable list of classification types without the loss of historical information, classifications can now be archived. An archived classification will be retained by the system but will not be available for assignment to CNVs during the interpretation process.

To archive a classification:

1. Select **Tools -> Options... -> Classification**.
2. In the **Existing classifications** table, select the classification(s) to be archived.
3. Click the **Archive Selected** button.

To restore a classification so that it becomes available for assignment to CNVs:

1. Select **Tools -> Options... -> Classification**.
2. In the **Archived Classifications** table, select the classification(s) to be restored.
3. Click the **Restore Selected** button.

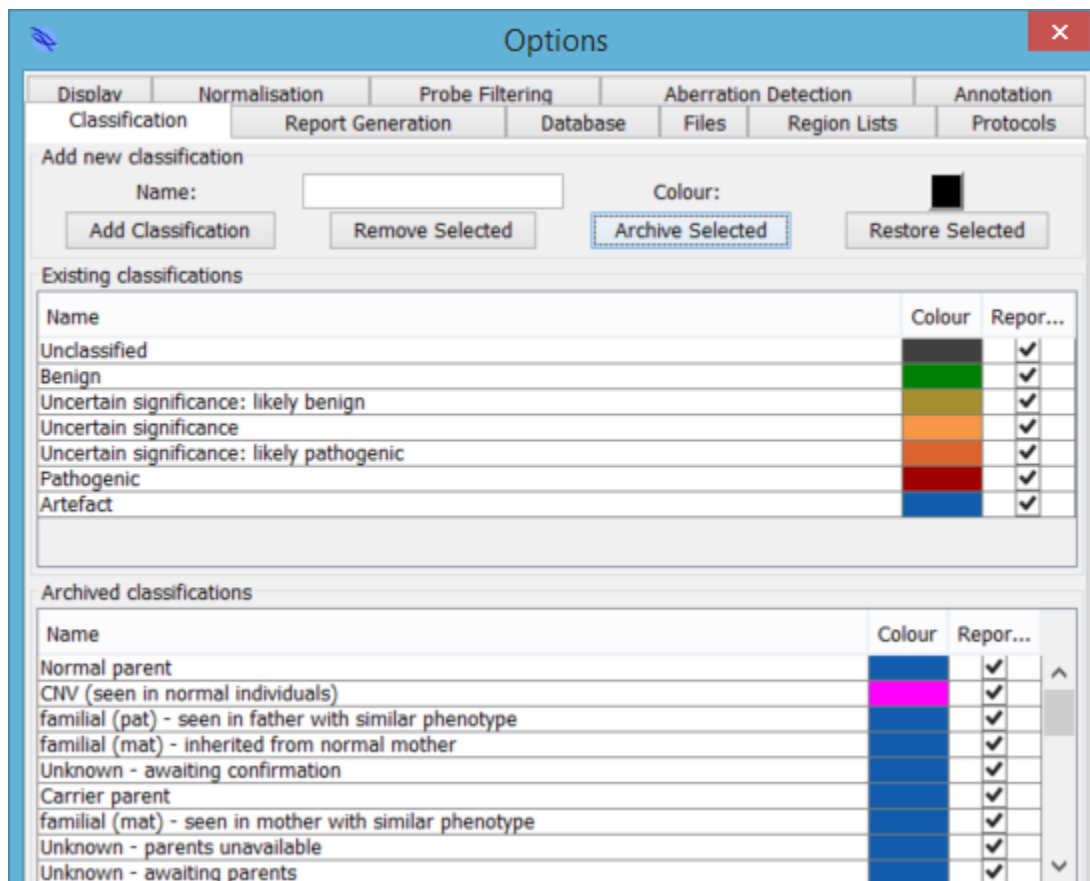


Figure: The Classification tab in the Options window, enabling archiving and restoration of CNV classifications.

Automatic security lockout

In order to protect any confidential information stored in the system while CytoSure Interpret Software is in use, the software now enables users to set a period of inactivity (the **Lockout time**) after which a user must re-enter their password to view or analyse data.

To enable automatic lockout, select **Logged in as... -> Manage Account**, set the **Lockout time (min)** to a value greater than 0 and click **Update Account**.

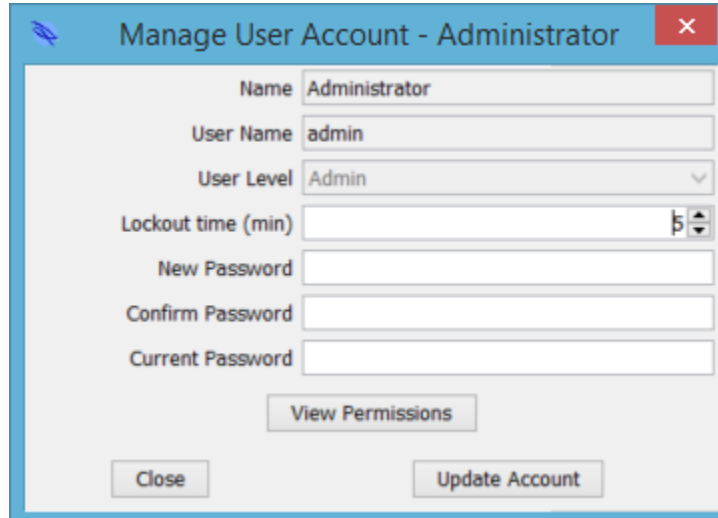


Figure: The Manage User Account window, now enabling the user to specify the Lockout time.

After the specified period of inactivity, the CytoSure Interpret Software window will go blank and the user will be required to re-enter their password before continuing.

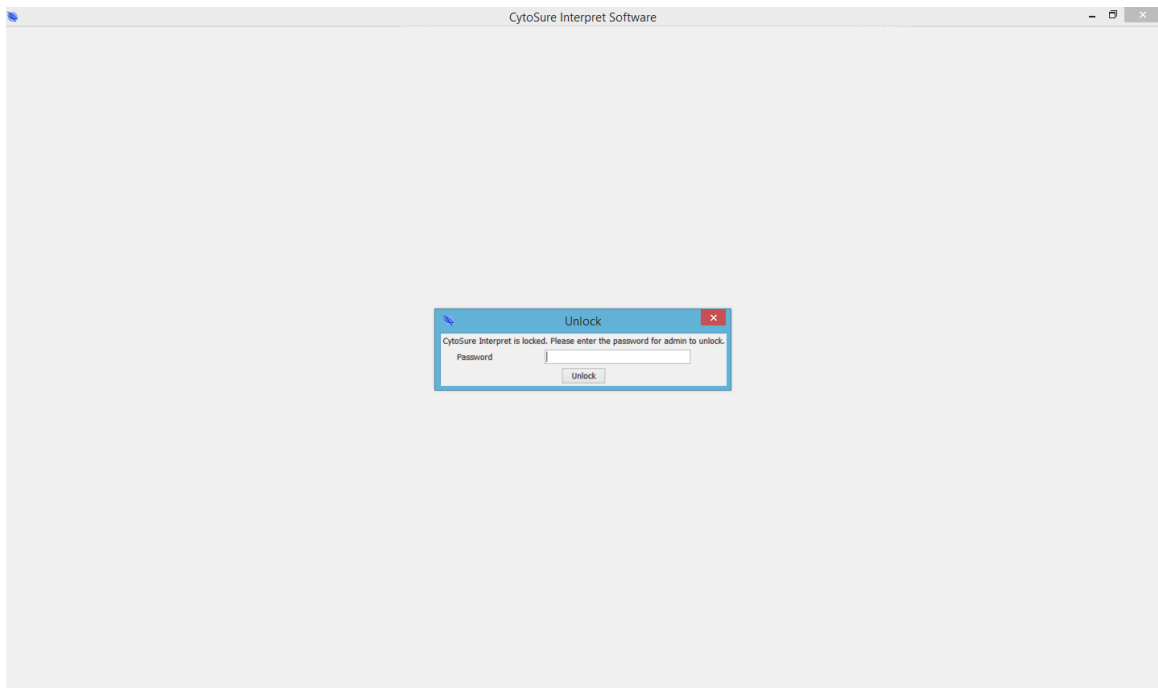


Figure: The Unlock screen, displayed after the specified period of inactivity.



Miscellaneous updates

- Ability to remove regions from region lists and to archive or delete old region lists (**Tools -> Options... -> Region Lists**).
- Ability to load the reference phenotype in a file parser (**Database Management -> Attach**).
- Ability to identify the slide barcode and the array position from two different fields when running batch processing.
- Ability to view all LOH regions from selected samples on one ideogram (**Database Management -> Batch Operations -> LOH Analysis -> Display Ideogram**).
- Ability to update annotation tracks in bulk (**Tools -> Options... -> Annotation**).
- Ability to merge segments after running CBS based on the number of probes or size of the gap between segments as specified by the segment merging thresholds (**Tools -> Options... -> Aberration Detection -> Advanced Options**).
- Ability to remove the Chromosome Homozygosity table from reports (**CNVs/SNPs -> Ideogram View -> Options... -> LOH Calling**).
- Ability to choose whether to include or to exclude the specified genes when displaying data after applying data masking.



CytoSure™ Interpret Software update — version 14.0 June 2014

The following new features are available in CytoSure Interpret Software version 4.6:

GRCh38 compatibility	76
Data Masking	76
Additional SNP quality metrics	81
SNP Call Rate	81
SNP Trough:Peak Ratio	81
Alternative Aberration track visualisation	82
Screenshot generation interface	83
Miscellaneous improvements	84



GRCh38 compatibility

Following the release of GRCh38/hg38 in December 2013 (see <http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/>), CytoSure Interpret Software now enables analysis of data using the latest genome build. For new database installations, hg38 will be installed by default, while users with existing database may add the new genome build via **Tools -> Install hg38**.

This process will add the genome build information to database and also add design information for any CytoSure designs usually available during the database installation process as well as RefSeq gene information in the Gene track. Further hg38 annotation will be made available in future releases via plug-ins when the relevant sources are updated with the new co-ordinates. If you require hg38 design files for other designs, please contact software@ogt.com.

To select hg38 as the current genome build, click **Current Genome Build:... -> Switch to hg38**. It is also possible to map existing CNVs in previous genome builds to hg38 by navigating to the **CNV Analysis** tab in **Database Management**, selecting a sample in the database with which CNVs are associated, right-clicking on a CNV in the table and selecting **Attempt mapping to hg38**.

Data Masking

The software now enables users to dynamically remove data points from regions of the genome to allow more focussed analysis of specific target regions and avoid incidental detection of CNVs. This process has been labelled “Data Masking”, and is available to the user when the **Enable Data Masking** option is selected in the **Region Lists** tab of the **Options** window.

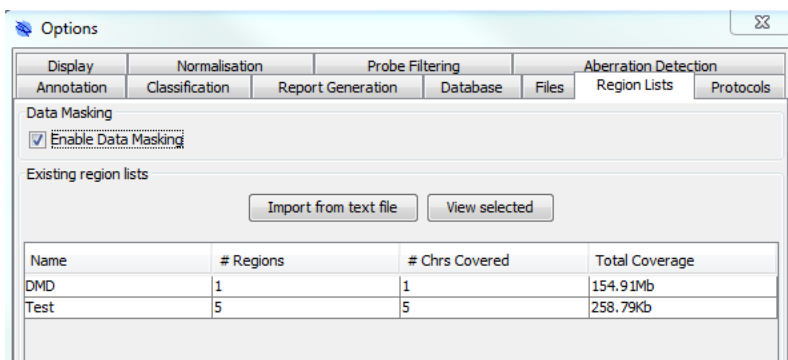


Figure: Enabling Data Masking via the Region Lists tab

Once this option is enabled, subsequent import of feature extracted or batch processed data results in the appearance of the **Select Regions to Display** window before any data is visualised. The window provides options enabling the user to select which **Region Lists** or individual regions to exclude from the analysis before continuing.

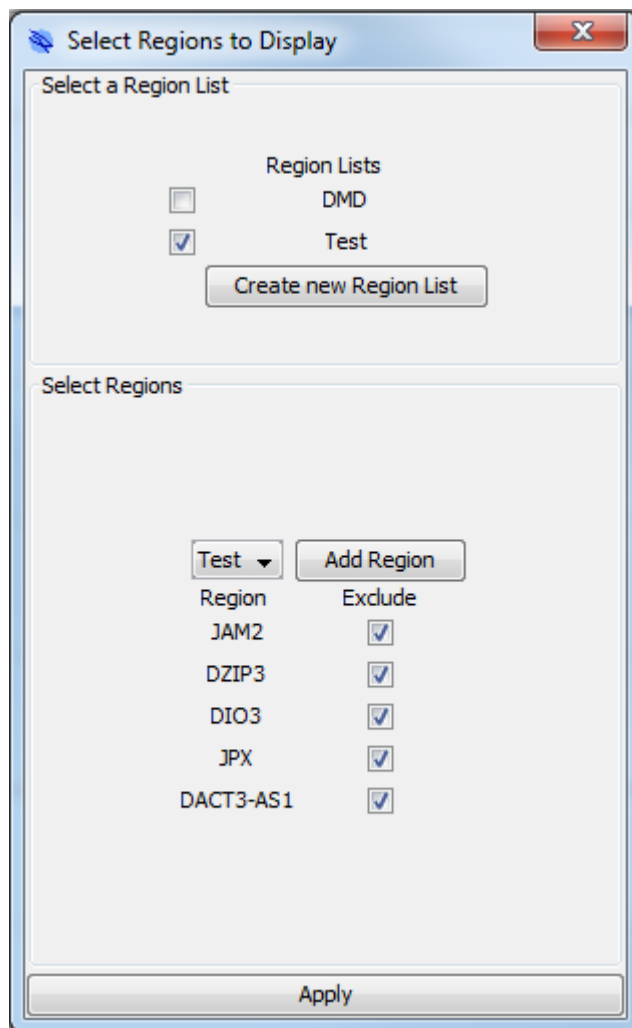


Figure: The Select Regions to Display window allowing users to set Data Masking options before visualisation

To create new Region Lists, click on the **Create new Region List** button and enter a name for the region list in the input field. There are three means by which individual regions may be added to the new region list:

1. By importing a list of regions in bulk from a text file via the **Import annotations from file** button. This process is similar to importing an annotation track, and is identical to the process followed by adding a Region List within **Tools -> Options... -> Region Lists -> Import from text file**.
2. By creating an empty region list and then adding individual regions to it by searching the existing database of annotation tracks. This can be achieved by ensuring that the new region list is selected in the **Select a Region List** panel, selecting it from the drop-down box in the **Select Regions** panel and then clicking the **Add Region** button. This will result in the display of the **Add Region** window enabling searching of regions in existing annotation tracks for addition to the region list via the **Add Region** button.

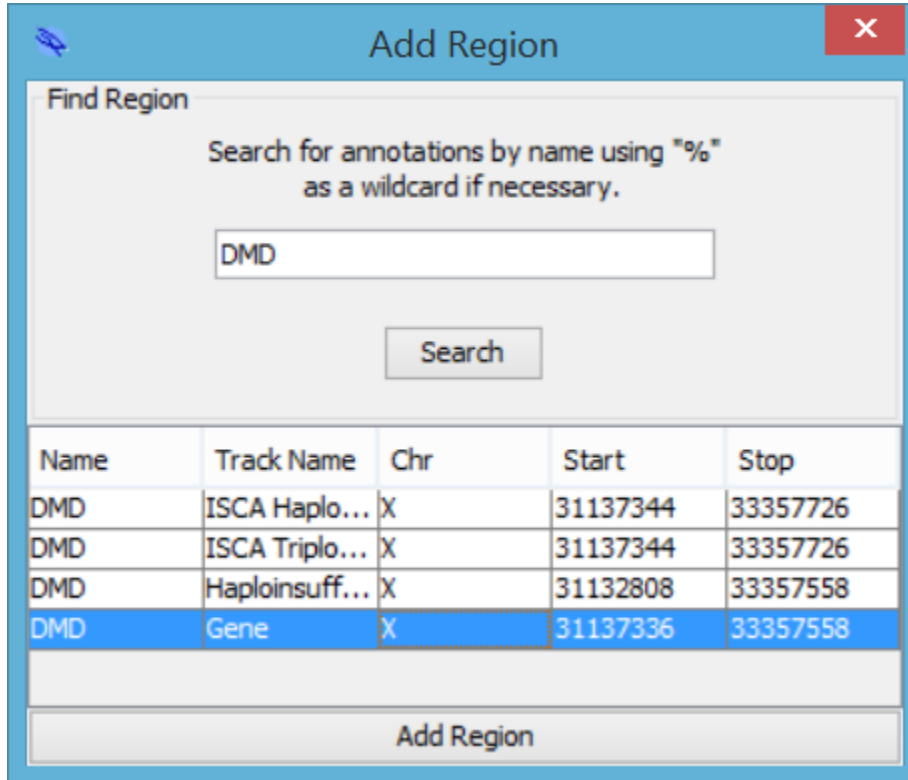


Figure: The Add Region window enabling addition of individual annotations to the selected region list

3. By adding individual regions from annotation tracks in the **Genomic View**. To do this, simply right-click on the annotation to be added to the region list, select **Add to region list** and click on the region list to which it should be added. It is also possible create a new region list via the same method by clicking the **New region list...**

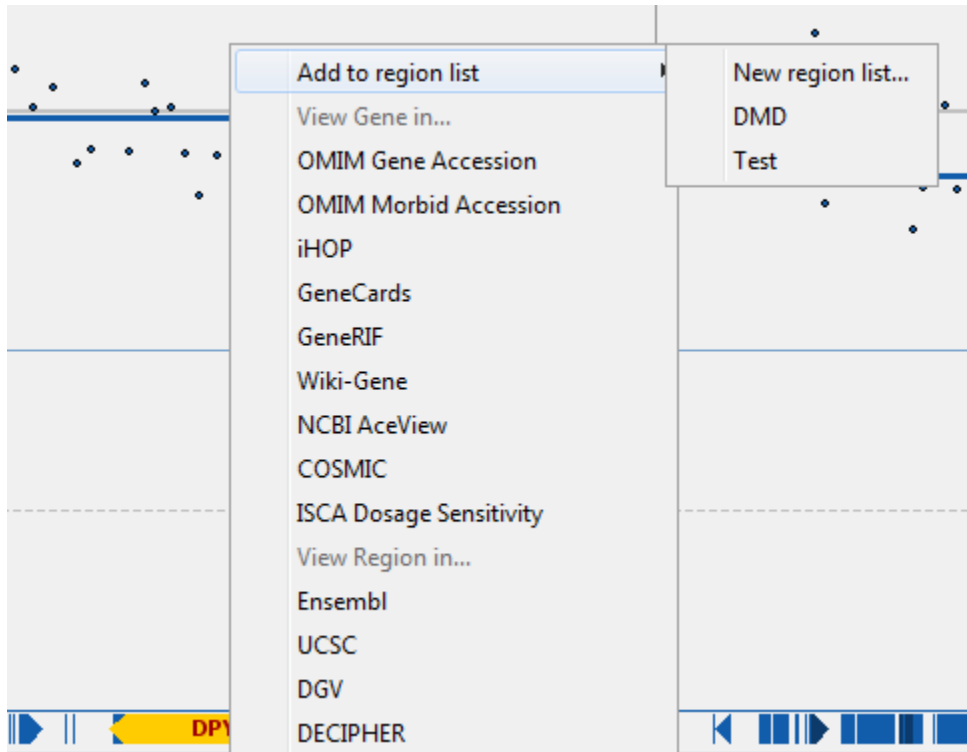


Figure: Adding an annotation to a region list via an annotation track in the Genomic View

It is possible to create multiple region lists in order to conveniently group regions to be masked, and once these region lists have been created, the user is able to select individual regions to be excluded from the data set as required.



Select Regions

Genes ▾ Add Region

Region	Exclude
AZFb+AZFc	<input checked="" type="checkbox"/>
ELN	<input type="checkbox"/>
DiGeorge Syndrome	<input checked="" type="checkbox"/>
NRXN1	<input type="checkbox"/>
TPM3	<input checked="" type="checkbox"/>
LIMK1	<input checked="" type="checkbox"/>
NEB	<input checked="" type="checkbox"/>
DMD	<input type="checkbox"/>

Apply

Figure: Selecting or deselecting regions to be excluded from the data

Once individual regions have been selected, clicking the **Apply** button will result in data from the selected regions being excluded from the analysis. The software will also remember the current settings for future analysis.

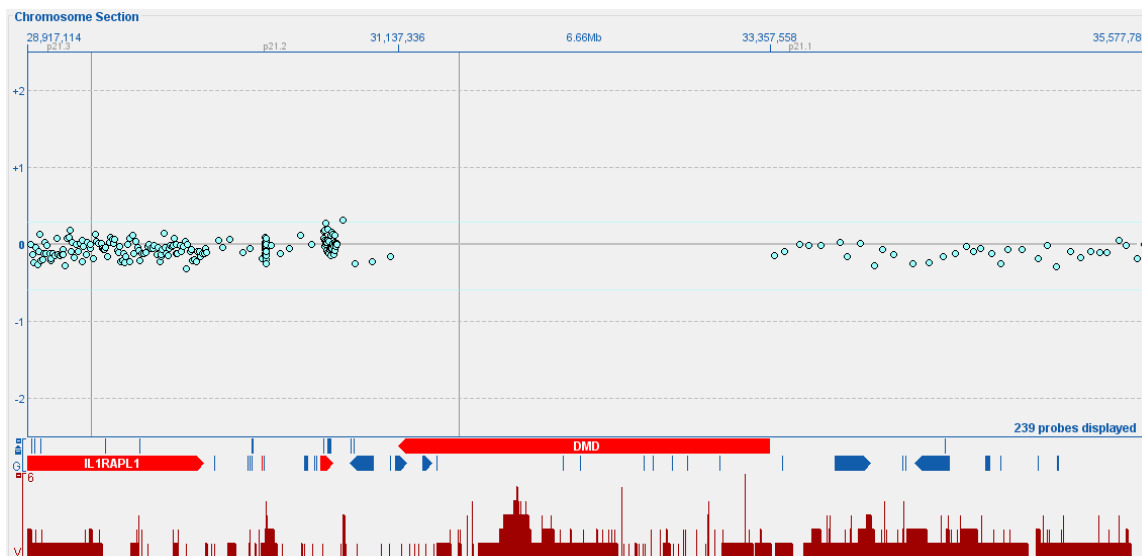


Figure: The result of applying data masking to the DMD gene region

Additional SNP quality metrics

In order to more accurately indicate the quality of SNP data and its ability to identify regions of LOH, CytoSure Interpret Software v4.6 includes two additional quality metrics.

SNP Call Rate

This metric measures the proportion of SNPs that have been successfully assigned a genotype, therefore good SNP data will give a high SNP Call Rate. The default thresholds for this metric are as follows:

- Excellent ≥ 0.988
- $0.988 > \text{satisfactory} \geq 0.983$
- Poor < 0.983

SNP Trough:Peak Ratio

High quality SNP data for most samples results in three peaks in the log ratio distribution plot, while low quality data may result in less well-defined peaks or even the complete loss of this expected distribution. As this may happen without affecting the SNP Call Rate, an additional metric to quantify the presence of this distribution has also been added — the SNP Trough:Peak Ratio. The ratio of the trough heights to the peak heights is low in high quality data, and the default thresholds for this metric are as follows:

- Excellent ≤ 0.6
- $0.6 < \text{satisfactory} \leq 0.8$
- Poor > 0.8



Alternative Aberration track visualisation

A new option for the visualisation of the Aberration track, containing previously detected CNVs, has been added in order to enable easier determination of the prevalence of CNVs in the selected region of the genome. When activated, CNVs in the Aberration track are automatically separated into gains and losses (and colour-coded accordingly). Additionally, the original **Supporting Variants** are consolidated into **Variant Regions**, and are then given a level of transparency to indicate the proportion of cases containing a CNV in the region and annotated with this proportion.

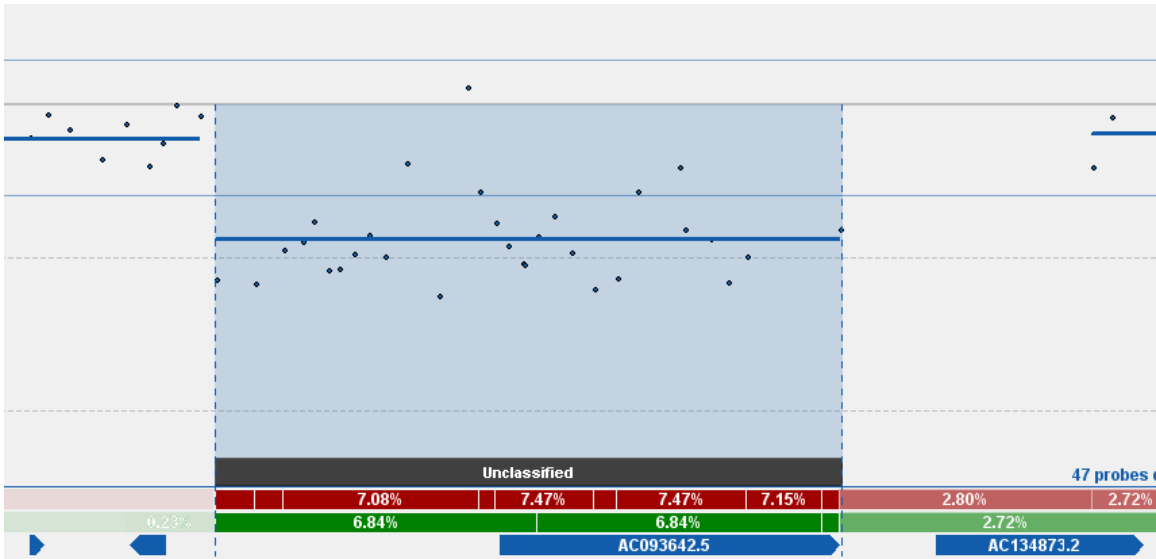


Figure: The Variant Region view of the aberration track

In order to activate the variant region view, right-click within the Aberration track and select **Show Variant Regions**. To revert to displaying the original supporting variants, right-click on the track and select **Show Supporting Variants**.

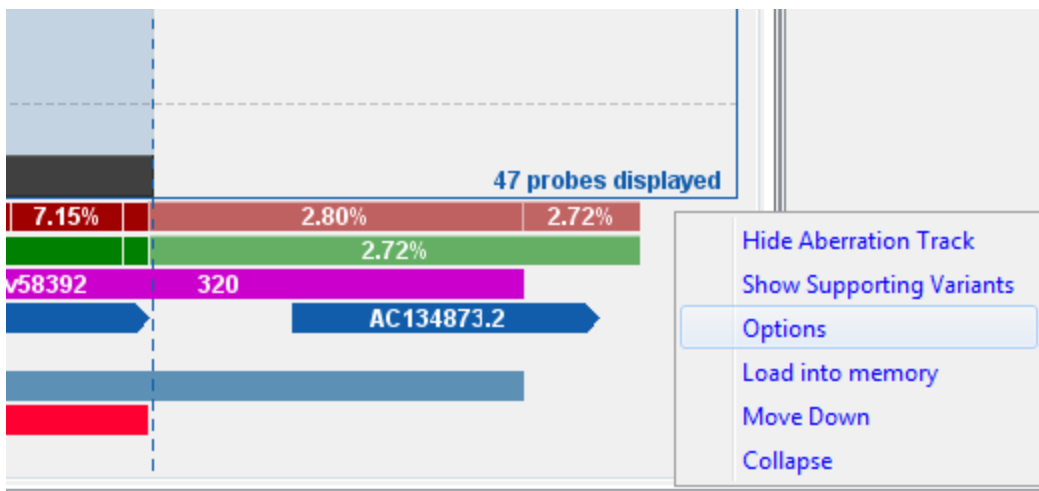


Figure: Switching between Aberration track views via the right-click menu



Users are also able to customise how different CNV penetrance levels correspond to transparency by clicking on the **Options** button in the same menu. This will open the **% CNV display** window, which provides controls for setting the **Proportion for opaque annotation** and the **Transparency at minimum** levels, with changes to these options being reflected in the **Current Settings** panel that indicates the transparency of variant regions that are present in varying proportions of samples.

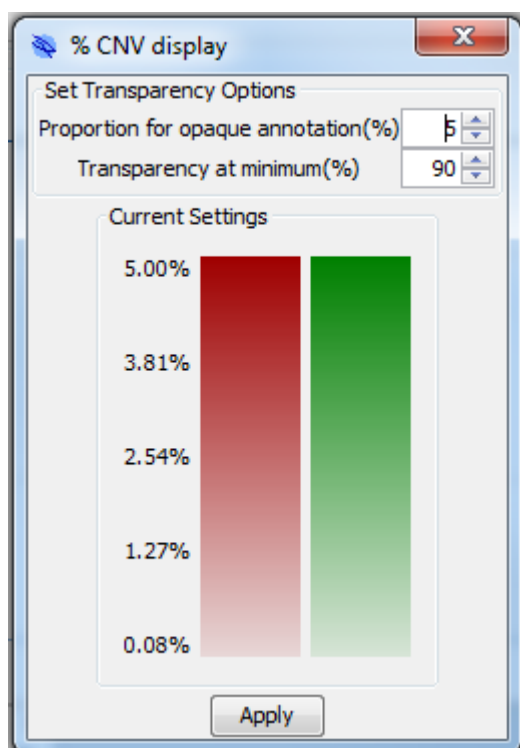


Figure: Modifying the transparency levels of variant region annotations based on their prevalence in the sample database.

Screenshot generation interface

In order to facilitate easy and accurate capture of user interface component images for sharing data visualisations with others, a new menu has been added to CytoSure Interpret Software that is available by right-clicking on most sections of the software.

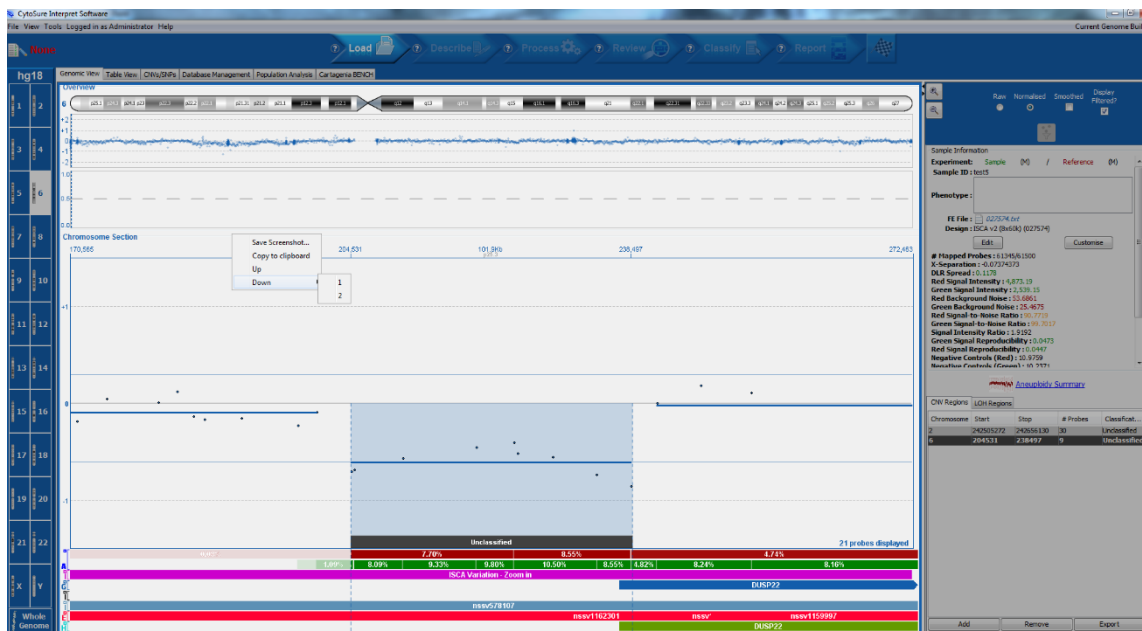


Figure: The screenshot capture menu shown following a right-click.

The screenshot capture system works by allowing the user to navigate the hierarchy of the CytoSure Interpret user interface in order to select only the component for which an image is required. Depending on where in the hierarchy the selected component falls, the menu will display the following options:

- **Save screenshot...** — Save the image of the selected component as a PNG.
- **Copy to clipboard** — Copy the image of the selected component to the clipboard so that it can be pasted into another application.
- **Up** — Move up the user interface hierarchy to select the component within which the selected component is contained, and display the menu for this component.
- **Down** — If the selected component only contains one sub-component, clicking the **Down** button will move down the user interface hierarchy to this component and display the menu for the component. If the selected component contains more than one sub-component, a sub-menu will appear containing a button for each sub-component to provide the same functionality.

Miscellaneous improvements

- Ability to share file parsers between databases by exporting from one database (**Database Management -> Batch Operations -> Parsers -> Export**) to create an XML file which can then be imported to another database (**Database Management -> Batch Operations -> Parsers -> Import**).
- Removal of any Loop calls generated from partially overlapping calls which contain fewer than the specified minimum number of probes and all CBS calls after Loop analysis (**Tools -> Analyse Loop Design**).
- Warning now provided if the genome build is automatically changed after a .cgh file is loaded with a different genome build.

CytoSure Interpret - Software update guide



- Addition of the sample gender to the tool tip that is displayed when the user hovers over a CNV in the Aberration track.
- Addition of new editable field to the Aberration Details showing the % Mosaicism for the CNV.
- Ability to create an ideogram which displays all CNVs from selected cases in the database (**Database Management -> Batch Operations -> CNV Analysis -> Display Ideogram**).



CytoSure™ Interpret Software update — version 13.0 November 2013

The following new features are available in CytoSure Interpret Software version 4.5:

Automatic Aberration Classification	87
Classifier Set creation	88
Classifier creation	89
Adding Classifiers to Classifier Sets	90
Applying the Classifier Set	91
Application options	91
Advanced settings	92
Resolving Conflicts	94
Standardised Phenotypes (HPO)	95
Inter-laboratory database sharing	97
Sharing cases	97
Importing shared cases	98
Allele Status Plot	98
Miscellaneous	99



Automatic Aberration Classification

CytoSure Interpret Software can now be configured to automatically classify detected CNVs based on criteria set by the user in order to accelerate and standardise the interpretation process. Configuration of this utility is available via **Tools -> Options... -> Classification**, with a number of related settings available in the **Automatic Aberration Classification** panel.

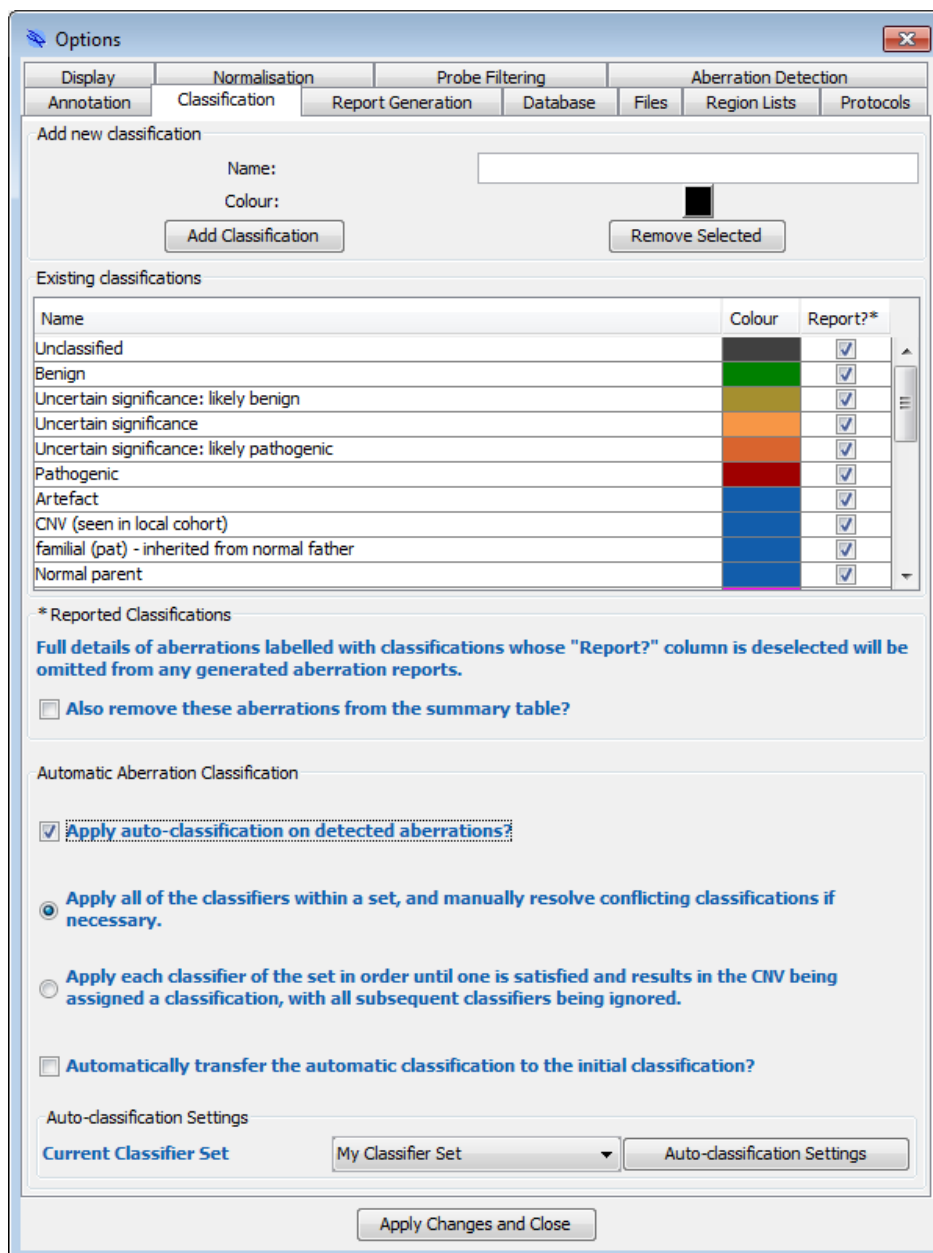


Figure: The Automatic Aberration Classification panel in the Classification tab of the Options window.



To enable automatic classification, users create a “Classifier Set” made up of multiple “Classifiers”. Each classifier describes the criterion against which a CNV will be tested and the classification that it will be assigned if it satisfies the criterion. For example, a user may create a classifier that checks whether a CNV is within an annotation in the Copy Number Variation track and, if so, assigns the “Benign” classification to that CNV automatically. Once the set of classifiers is constructed, the user may opt to apply the classifiers in order until one passes its criterion, or apply all classifiers and resolve conflicting classifications manually if necessary. Multiple classifier sets can be created for use in different types of analyses, and the selected classifier set can be linked to an analysis protocol in order to standardise interpretation when the protocol is in use.

The following sections describe the administration and application of the automatic classification functionality in full.

Classifier Set creation

To enable automatic CNV classification, the user must first create a classifier set via **Tools -> Options... -> Classification -> Automatic Aberration Classification -> Auto-classification Settings -> Auto-classification Settings**. This will open the **Automatic Classification Settings** window containing the controls to manage the classifier sets and individual classifiers that are available to the user.

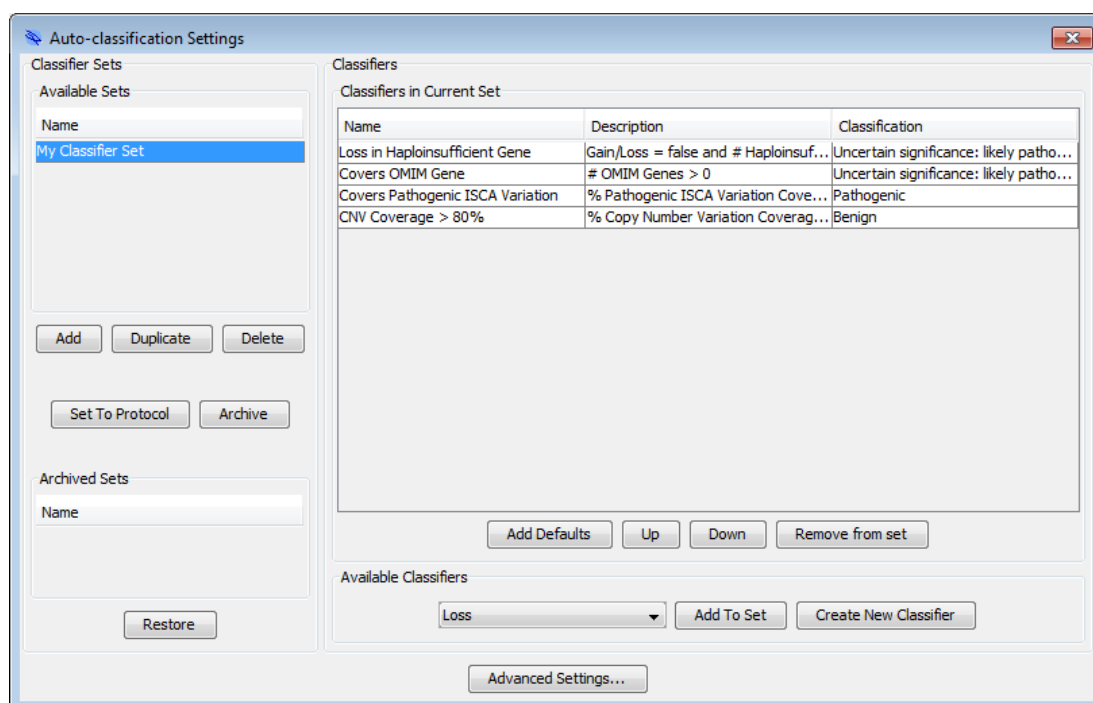


Figure: The Auto-classification Settings window providing functionality for configuration of classifier sets.



To create a new classifier set, click the **Add** button in the **Classifier Sets** section, enter a name for the new set and click **OK**. The name of the set will then appear in the **Available Sets** table.

Classifier creation

To create classifiers to add to the classifier set, click the **Create New Classifier** button in the **Available Classifiers** panel at the bottom of the screen to open the **Create Classifier** window.

Figure: The Create Classifier window providing the means to create new classifiers.

The software enables the user to create two types of classifier – **Basic** and **Advanced**.

Basic Classifiers test a standard attribute of each CNV against the given criteria. For example, they can determine whether the CNV is a duplication, or whether the mean log ratio of the CNV exceeds a certain threshold, and assign a classification accordingly. To create a Basic Classifier, in the **Basic** tab select the required **Field Name** and **Operator Type**, enter the value against which the CNV will be compared, select the **Classification Type** that will be assigned to a CNV that satisfies the criterion, enter a **Classifier Name** describing the Classifier and click the **Create Classifier** button.



Advanced Classifiers consider the annotations in a particular track within each CNV region, allowing the user to classify based on the names of these annotations, the number of annotations in the region, or the total proportion of the CNV region covered by the annotations (the “Overlap Statistic”). For example, users can create an Advanced Classifier to test whether a CNV covers an OMIM gene, or whether more than 80% of the CNV region is covered by annotations in the Copy Number Variation track, and assign a classification accordingly. To create an Advanced Classifier, in the **Advanced** tab select the **Annotation Type** and the **Overlap Statistic** to calculate for the CNV region in the selected track, then select the **Operator Type** and enter the appropriate value into the field(s) beneath in order to create the criterion. To complete the Classifier, select the **Classification Type** to which CNVs that satisfy the criterion should be assigned, enter a name for the Classifier in the **Classifier Name** field and click the **Create Classifier** button.

Figure: The Advanced classifier creation interface in the Create Classifier window.

Adding Classifiers to Classifier Sets

All classifiers that have been created by the user (that haven't been archived) are listed in the **Available Classifiers** drop-down box in the **Auto-classification Settings** window. To add a classifier to a set, simply select the set from the list of **Available Sets**, select the classifier from the **Available Classifiers** drop-down box and click the **Add To Set** button. The selected classifier, which will appear in the **Classifiers in Current Set** table, can be individually moved up or down



the list to determine the order in which they will be applied by using the **Up** and **Down** buttons and can be removed from the set by clicking the **Remove from set** button.

The **Auto-classification Settings** window also provides the means to add a list of example classifiers as a potential starting point for building a classifier set via the **Add Examples** button.

Applying the Classifier Set

When the classifier set has been created, close the **Auto-classification Settings** window to return to the **Classification** tab in the **Options** window. The new set will appear in the **Current Classifier Set** drop-down box, and selecting it will ensure that it is used whenever the software runs automatic aberration classification. CytoSure Interpret Software can be configured to run automatic aberration classification whenever a CNV is detected by selecting the **Apply auto-classification on detected aberrations?** checkbox, and the user can also manually trigger application of the classifier set by selecting **Tools -> Apply Auto-Classification**.

Please note that calculation of Advanced Classifiers may be accelerated by loading the relevant annotation tracks into memory by right-clicking on the track in the **Genomic View** and selecting **Load into Memory**.

The **Current Classifier Set** and the other settings in the **Automatic Aberration Classification** section described below are analysis protocol options, meaning that they are linked to the analysis protocol and their modification will result in the nullifying of the current protocol. This also means that it is possible to establish which classifier set (if any) was used in all experiments run in Workflow mode.

Application options

The user can choose to apply the classifier set in one of two ways:

1. **Apply all of the classifiers within a set, and manually resolve conflicting classifications if necessary.**
2. **Apply each classifier of the set in order until one is satisfied and results in the CNV being assigned a classification, with all subsequent classifiers being ignored.**

The user should select the appropriate option in the **Automatic Aberration Classification** section to suit their requirements in this respect.



Automatic Aberration Classification

Apply auto-classification on detected aberrations?

Apply all of the classifiers within a set, and manually resolve conflicting classifications if necessary.

Apply each classifier of the set in order until one is satisfied and results in the CNV being assigned a classification, with all subsequent classifiers being ignored.

Automatically transfer the automatic classification to the initial classification?

Auto-classification Settings

Current Classifier Set My Classifier Set Auto-classification Settings

Figure: A close-up of the Automatic Aberration Classification section in the Classification tab of the Options window.

When a classification is assigned to a CNV via automatic aberration classification, the **Classification (Automatic)** field is set by the software to suggest an initial classification to the user. If required, users can configure the software to transfer the automatic classification to the **Classification (Initial)** field as part of the automatic aberration classification process by ensuring that the **Automatically transfer the automatic classification to the initial classification?** checkbox is selected.

Advanced settings

Further customisation of Classifiers is available via the **Advanced Settings...** button in the **Auto-classification Settings** window, where existing classifiers may be archived or translated into new classifiers to create more complex criteria.

In order to be able to maintain a manageable list of classifiers, it is possible to archive classifiers that are no longer required so that they do not appear in the **Available Classifiers** drop-down box in the **Auto-classification Settings** window by selecting them in the **Available Classifiers** table and clicking on the **Archive** button. Archived classifiers will appear in the **Archived Classifiers** table and can be restored for subsequent use in a classifier set by clicking on the **Restore** button.

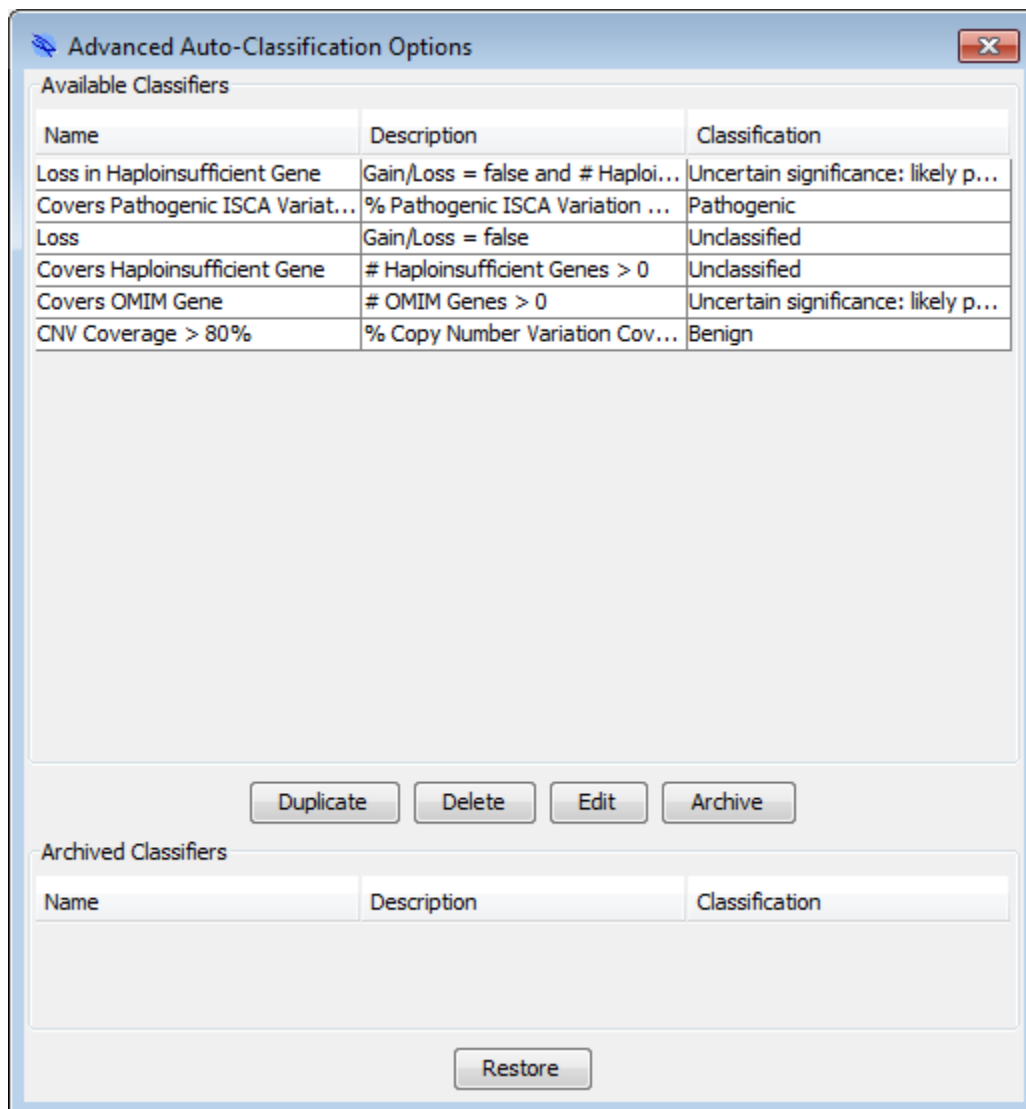


Figure: The Advanced Auto-Classification Options window providing further options to customise and manage individual classifiers.

It is also possible to translate existing classifiers into new classifiers to create more complex classification criteria in the following ways:

- **Classifier Negation** – Create a classifier that tests the opposite of a given classifier by right-clicking on the entry in the **Available Classifiers** table, selecting **Negate**, completing the **Classifier name** and **Classification type** fields and clicking the **Create** button.



- **Classifier Duplication** – Create a duplicate of an existing classifier, perhaps in order to make only a slight modification to an attribute, by selecting the entry in the **Available Classifiers** table and clicking on the **Duplicate** button. The new copy of the classifier will appear at the bottom of the table with the word “Copy” added to its name.
- **Classifier Combination** – Combine 2 classifiers to create a new classifier that checks whether either or both criteria are satisfied by a CNV by selecting exactly 2 classifiers from the **Available Classifiers** table, right-clicking on the table and selecting the **Combine** button. In the **Advanced Classifier** window that appears, enter a **Classifier name**, select the type of combination, choose the **Classification type** which will be assigned to any CNVs that satisfy the criteria and click **Create**.

The **Advanced Auto-Classification Options** window also allows users to **Edit** or **Delete** existing classifiers via the appropriate buttons.

Resolving Conflicts

If the user has chosen to apply all classifiers and to also automatically transfer the automatic classification to the initial classification, and application results in the attempted assignment of two or more different classifications, a window will appear prompting the user to resolve this conflict by selecting a single initial classification for each CNV affected.

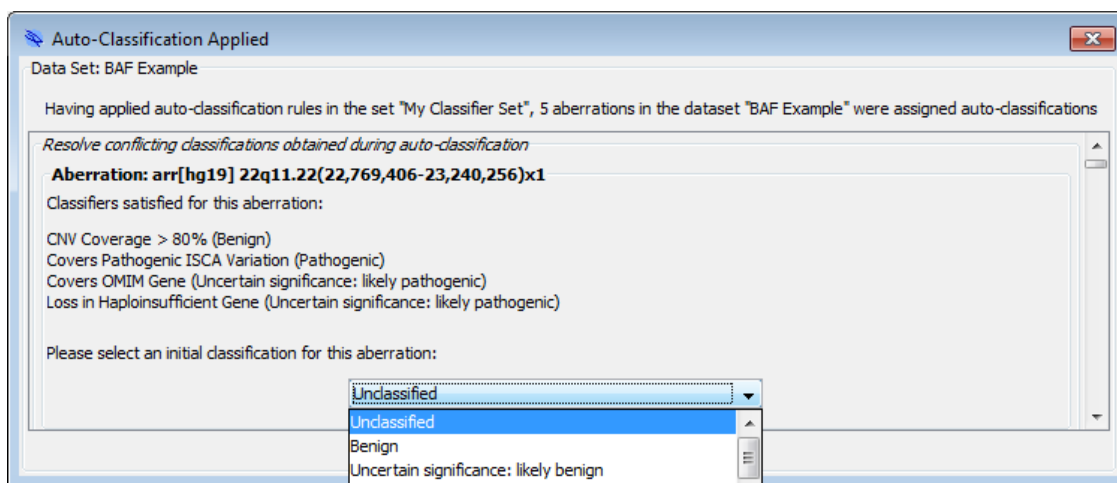


Figure: The interface enabling users to resolve conflicting automatic classifications by manually selecting a classification to be associated with the CNV.



Standardised Phenotypes (HPO)

In addition to the current free text Phenotype field, users are now able to construct standardised phenotypes for a sample using Human Phenotype Ontology (HPO) terms via the **Sample Data** section of the **Database Management** tab.

The screenshot shows the 'Sample Data' window with the 'HPO Example' sample selected. The 'Basic Information' section includes fields for 'Sample ID' (HPO Example), 'Gender' (Male selected), 'Phenotype' (empty), and 'Notes' (empty). There are radio buttons for 'Reference?' (Yes/No) and 'Analysis Date: N/A'. The 'HPO Phenotype' field is a list box containing terms like 'Abnormal facial shape', 'Central hypotonia', 'Macrocephaly', 'Gastroschisis', and 'Spasticity'. An 'Edit HPO Phenotype' button is located to the right of the list box.

Figure: The Sample Data tab in Database Management including the HPO Phenotype field.

HPO Phenotypes can be constructed by adding one or more HPO terms to a list of terms for the sample via the **Edit HPO Phenotype** button, which opens the **HPO Phenotype Selector** window. To add a term to the phenotype for the sample, enter appropriate text in the **Search** field and press Enter to filter the complete list of terms to those whose name, definition or synonym fields contain the search text.

The screenshot shows the 'HPO Phenotype Selector' window. It has a search field with 'height' entered. The 'HPO Terms' table lists terms like 'Abnormality of body height', 'Tall stature', etc. The 'Sample Phenotypes' table lists terms like 'Autosomal dominant inheritance', 'Bladder diverticula'. There are 'Associate with sample' and 'Remove' buttons at the bottom right, and a 'Close' button at the bottom center.

HPO ID	Name
HP:0000002	Abnormality of body height
HP:0000098	Tall stature
HP:0000179	Thick lower lip vermillion
HP:0000215	Thick upper lip vermillion
HP:0000218	High palate
HP:0000219	Thin upper lip vermillion
HP:0000276	Long face
HP:0000348	High forehead

HPO ID	Name
HP:0000006	Autosomal dominant inheritance
HP:0000015	Bladder diverticula

Figure: The HPO Phenotype Selector window providing the functionality to manage HPO Phenotypes for samples.

More information relating to an individual HPO term can be viewed by double-clicking on the term in the table to display its details window. This includes **Synonyms** and a **Definition** for the term, as well its **Super-classes** and **Sub-classes** which provide links to other terms that are respectively more general or more specific than the current term. A link to the Human Phenotype Ontology web



page for the term is also displayed at the bottom of the window which may provide more resources related to the term.

Figure: Full details of the “Tall stature” HPO term.

Once the user has identified the appropriate term from the filtered list in the **HPO Phenotype Selector** window, it can be linked to the sample by clicking the **Associate with sample** button. Once all required terms have been added, click the **Close** button and save the changes to the sample to confirm the update to the HPO Phenotype in the database.

Figure: Addition of the “Tall stature” term to the HPO Phenotype of the selected sample.

Users are able to remove terms from the HPO Phenotype by either right-clicking on the term in the list and selecting **Remove**, or clicking the **Edit HPO Phenotype** button, selecting the term in the **Sample Phenotypes** table and clicking the **Remove** button. Both methods require subsequent saving of the sample for the changes to be committed to the database.

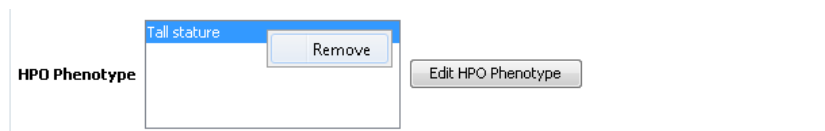


Figure: Removal of a term from the HPO Phenotype

HPO Phenotypes can also be included in aberration reports by adding the **HPO Phenotype** field to the **Included** list in the **Sample Data** tab of the **Report Generation** section of the **Options** dialog box (available via **Tools -> Options...**). In PDF and HTML reports, individual HPO terms will also link to their respective page on the Human Phenotype Ontology website.

File	
HPO Phenotype	Abnormal facial shape ; Central hypotonia ; Macrocephaly ; Gastroschisis ; Spasticity ;
Genome Build	hg19:GRCh37:Feb2009
Analysis Date	15 Nov 2012

Figure: The HPO Phenotype of a sample included in a report.

Inter-laboratory database sharing

CytoSure Interpret Software supports a number of common relational database management systems, enabling concurrent access to sample and experiment information across a network, providing the potential for labs on multiple sites to access a central database if the IT infrastructure permits. Where this is not possible, the software now provides functionality to allow a lab (the source) to export sample and experiment information so that it may be transferred to collaborators (the targets) and imported into their own database to facilitate sharing of case information and aberration calls.

Sharing cases

To export data to be shared with another lab, users should first, if necessary, filter the sample list in the **Database Management** tab to only include samples that they would like to share (either via **Quick Search** or **Advanced Search**), then select **Batch Operations -> Database -> Sharing -> Share Case Results**. After selecting a location and name for the results file, a prompt will appear asking the user to choose whether to export the original sample identifiers, add a prefix to the identifiers (perhaps to distinguish samples within the same identifiers in the target laboratory) or generate unique random identifiers for each sample (to anonymise the samples) .

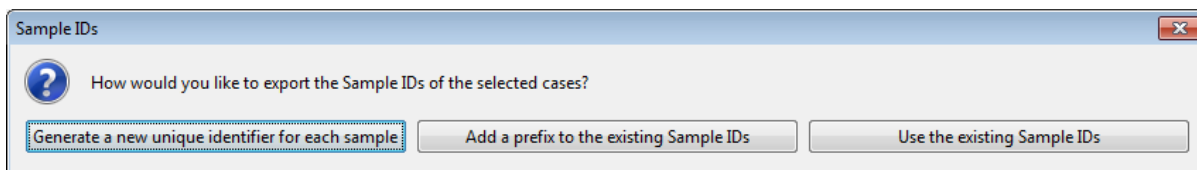


Figure: Options available to the user regarding the export of sample identifiers during database sharing.



The final step in the export process is to select the custom sample and experiment fields that the user would like to include in the export (and therefore share with the target laboratory). Once any fields that need to be omitted from the export are moved to the **Excluded** column, click the **Continue** button and the software will generate a zip file in the specified location that can be sent to the target lab.

Importing shared cases

Once a zip file containing shared cases has been generated by the source lab, the target lab can import the data by selecting **Database Management -> Batch Operations -> Database -> Sharing -> Import Shared Case Results**, choosing the zip file from the file system, clicking the **Open** button and selecting the appropriate option when asked to determine the software's behaviour should samples in the import file match those already present in the target database.

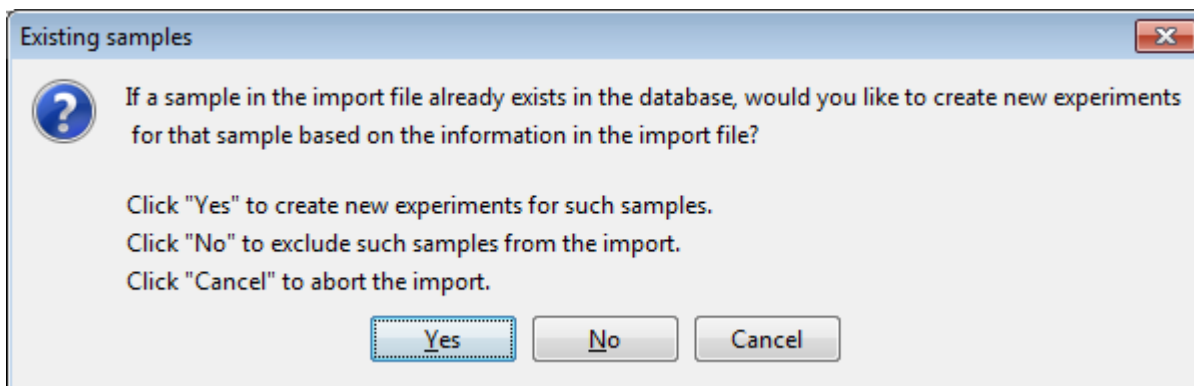


Figure: Options available to the user regarding import of information for samples that already exist in the database.

Allele Status Plot

As an alternative to the B-Allele Frequency Plot, CytoSure Interpret Software now includes an "Allele Status Plot" for the visualisation of SNP data. The plot displays an estimate of the allele status (AA, AB, BB or allelic imbalance) of each SNP probe on a CGH +SNP array based on its assigned genotype and the results of the segmentation of the B-Allele Frequency values, simplifying interpretation of the SNP data.

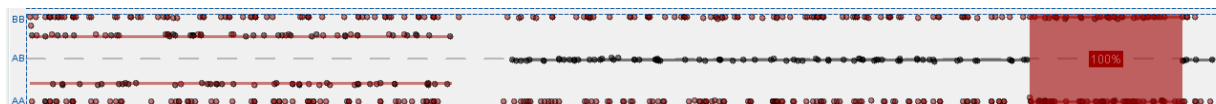


Figure: An example of an Allele Status Plot containing a duplication causing allelic imbalance of the p-arm and a deletion causing LOH in a region of the q-arm.

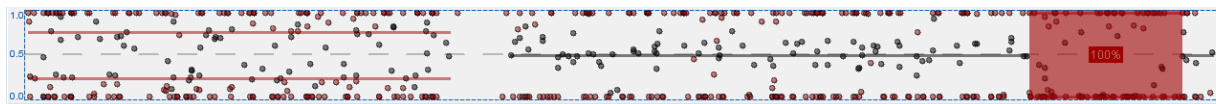




Figure: The B-Allele Frequency plot of the same data.

The plot is available in place of the B-Allele Frequency Plot in both the **Genomic View** tab and the **CNVs/SNPs** tab, and users can switch between views by right-clicking on the plot and selecting **Switch to B-Allele Frequency/Allele Status Plot**.

Miscellaneous

- New interface for annotation searching providing record of search history (**View -> Find Annotation...**).
- Ability to create filters based on the numeric values in custom annotation track fields in order to create sub-tracks. For example, it may be possible to create a filter of the Copy Number Variation track that only displays CNVs detected in studies with a sample size greater than a certain number (**Tools -> Options... Annotation -> Filters**).
- Additional advanced options for CNV artefact screening based on probe distribution (**Tools -> Options... -> Aberration Detection -> Advanced Options -> Artefact Screening**).
- Addition of genome build to ISCN notation in line with 2013 recommendations.
- Ability to filter database based on QC metric values (**Database Management -> Advanced Search -> Experiment Details**).
- Ability to select fields to be included in the database export feature (**Database Management -> Batch Operations -> Database -> Export**).
- Ability to search samples for details included in previous experiments.
- Ability to include custom sample and experiment fields in the sample table in the **Database Management** tab.
- Ability to filter probes based on GC content (**Tools -> Probe Filtering -> Probe Filters -> GC Content**).
- Ability to visualise CNV data in the **Genomic View** with each probe joined by a line (**Tools -> Options... -> Display -> Join Probes**).
- Ability to visually indicate which annotations overlap with CNVs in the **Genomic View** by extending a dashed line from the start and end positions of each CNV across the annotation tracks (**Tools -> Options... -> Display -> Indicate CNV/Annotation Overlap**).
- Ability to save individual cases to .cgh files and to the database from the **Genomic View** via the buttons  and  respectively.



- Ability to restrict annotations reported as overlapping detected CNVs based on both the minimum proportion of the annotation covered by the CNV and the minimum proportion of the CNV covered by the annotation. This function is available by right-clicking a selected overlap statistic (**Report, Report # or Report %**) in **Tools -> Options... -> Annotation -> CNV Reporting** and selecting **Advanced Options...**
- Easy access to log files via **Help -> Logs**.
- November 2013 updates to the following annotation tracks:
 - Copy Number Variation
 - Gene
 - Exon
 - ISCA Variation
 - ISCA Haploinsufficient Gene
 - ISCA Triplosensitive Gene

To update these tracks in an existing database, select **Tools -> Options... -> Annotation**, select the annotation track from the table and click the **Update Selected** button.



CytoSure™ Interpret Software update — version 12.0 March 2013

The following new features are available in CytoSure Interpret Software version 4.4:

Feature extraction	102
Segmentation-based normalisation	102
Custom SNP colouring	103
User level modification	105
User archiving	105
Miscellaneous	106



Feature extraction

CytoSure Interpret Software is now able to perform feature extraction on TIFF images generated from a number of different scanners. For full information on how to take advantage of this new functionality, please refer to the Feature Extraction User Guide (**Help -> Feature Extraction Guide**).

Segmentation-based normalisation

Although standard methods for dye-bias correction, such as median normalisation and LOWESS normalisation, are effective for largely diploid genomes, sub-optimal adjustment may occur in samples containing many whole-chromosome aberrations.

In order to determine a more accurate correction for such samples, a normalisation method based on the results of Circular Binary Segmentation (CBS) has been implemented and can be seamlessly integrated into the analysis workflow. The method attempts to determine the log ratio of the raw data most likely to represent 0 by scanning log ratios and applying a weighting to any segments that are within a certain range based on their proximity to the log ratio being tested and the number of probes in the segment to calculate a score. Once all potential log ratio values have been tested, the log ratio with the highest score is assumed to represent 0, and all segments and probe data points are adjusted accordingly.

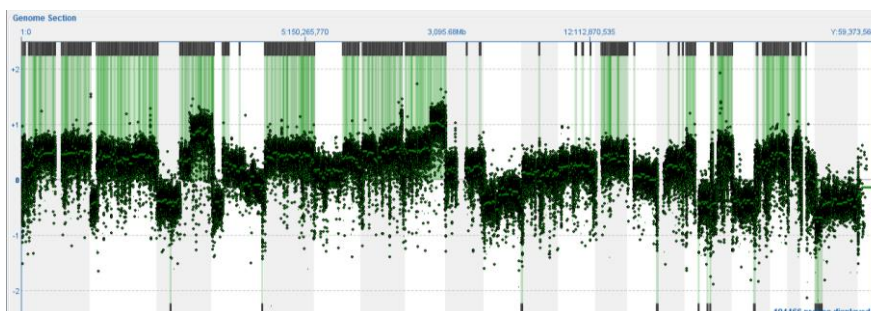


Figure: Incorrect normalisation of a highly aberrant sample.

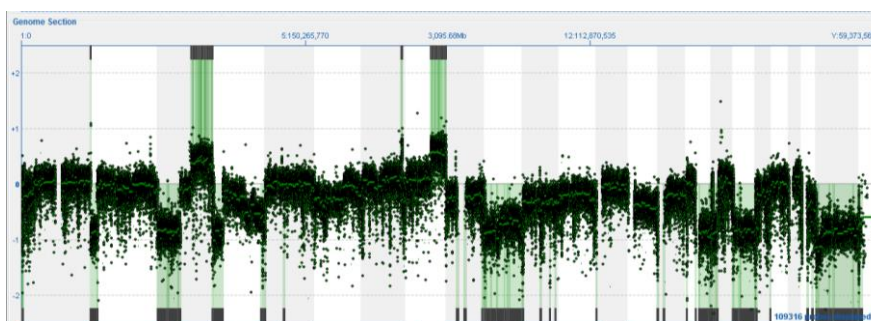


Figure: The same sample after segmentation-based normalisation.

Segmentation-based normalisation can be applied in two ways:

- Manually – By selecting **Tools -> Apply Segmentation-based normalisation**.



- Automatically – By selecting **Tools -> Options... -> Normalisation** and ensuring that **Automatically carry out Segmentation-based normalisation?** is selected. This will result in the normalisation being applied after a sample has been processed with CBS, including during batch processing.

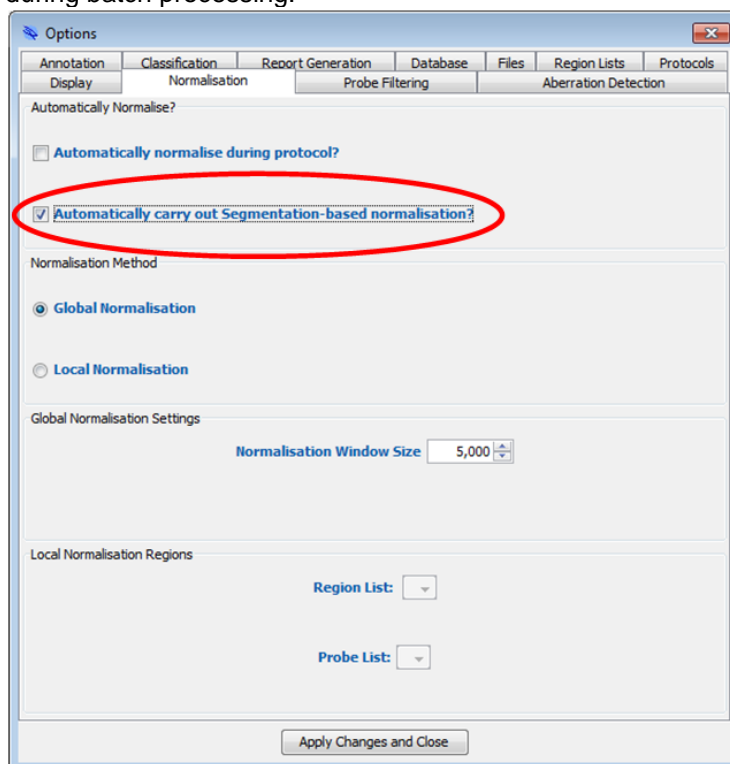


Figure: Setting up CytoSure Interpret Software to automatically carry out segmentation-based normalisation.

Please note that it is important to ensure that automatic segmentation-based normalisation is applied if TIFF images are feature extracted within CytoSure Interpret Software, as no normalisation is carried during the feature extraction process.

Custom SNP colouring

It is now possible for the user to associate custom colours to the homozygous and heterozygous SNPs for SNP/UPD analysis. The colours can be chosen by selecting **CNVs/SNPs -> Ideogram View -> Options... -> LOH Calling** and clicking on the relevant coloured button.

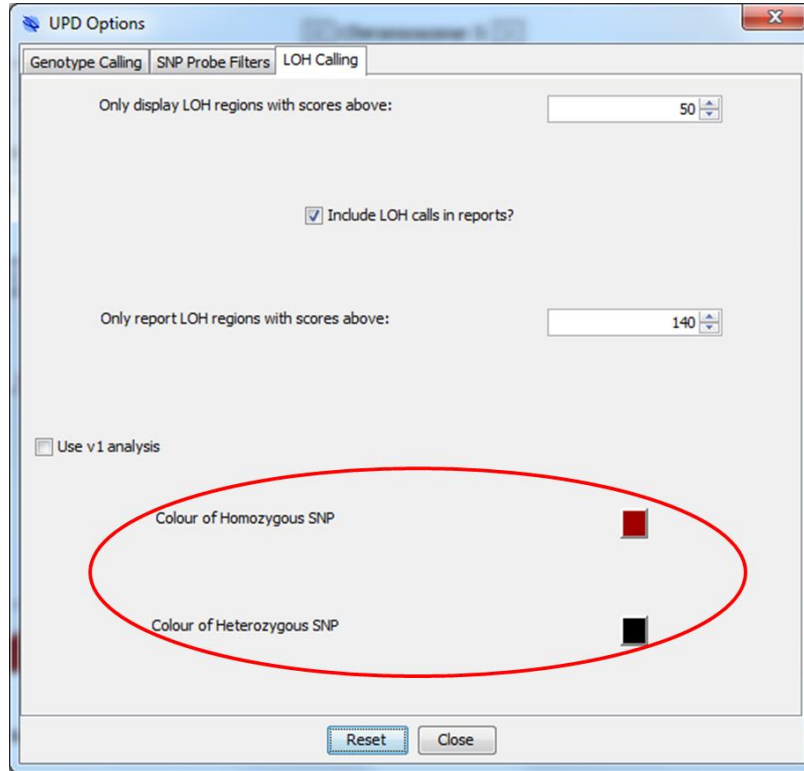


Figure: Changing the SNP colours via the UPD Options window.

Changing the SNP colours affects the B-Allele frequency plots, the distribution plot and the ideogram view of the SNPs.

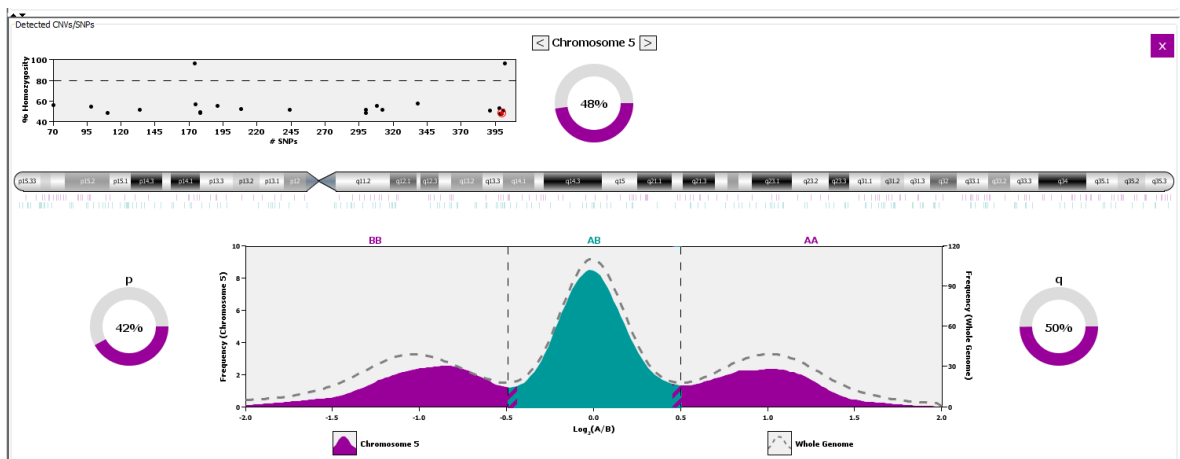


Figure: The SNP frequency plot with alternative homozygous and heterozygous colours.

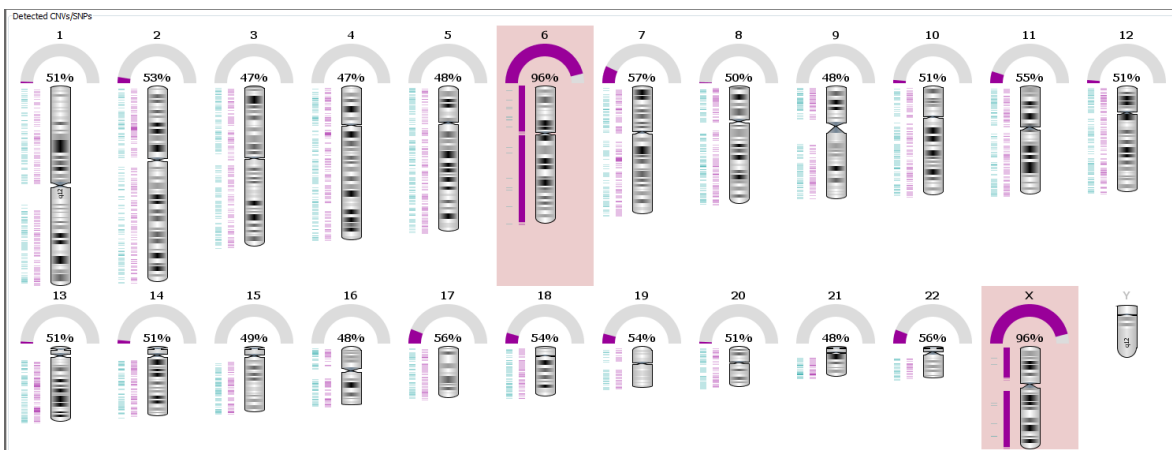


Figure: The Ideogram View with alternative homozygous and heterozygous colours.

User level modification

Users with permission to manage other accounts are now able to modify the **User Level** of anyone they “manage” up to the level immediately below their own. The user levels can be changed by selecting **Logged in as [user name] -> User List**, selecting the user to be modified, clicking the **Manage User** button, editing the **User Level** drop-down box and clicking the **Update Account** button.

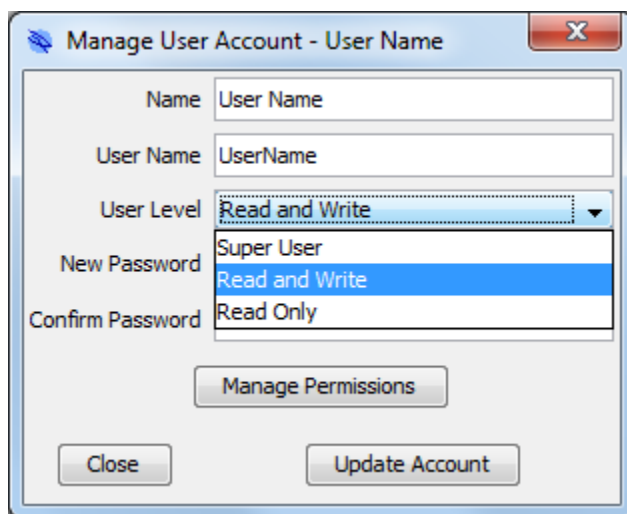


Figure: User level modification via the Manage User Account window.

User archiving

A user account can now be archived by users with permission to manage other accounts. This can be done by selecting **Logged in as [user name] -> User List**, selecting a user from the table and clicking the **Archive Account** button. Once archived, the account can no longer be used to log in to the software. However, the user name still appears on the submissions made by the user while the account was active.

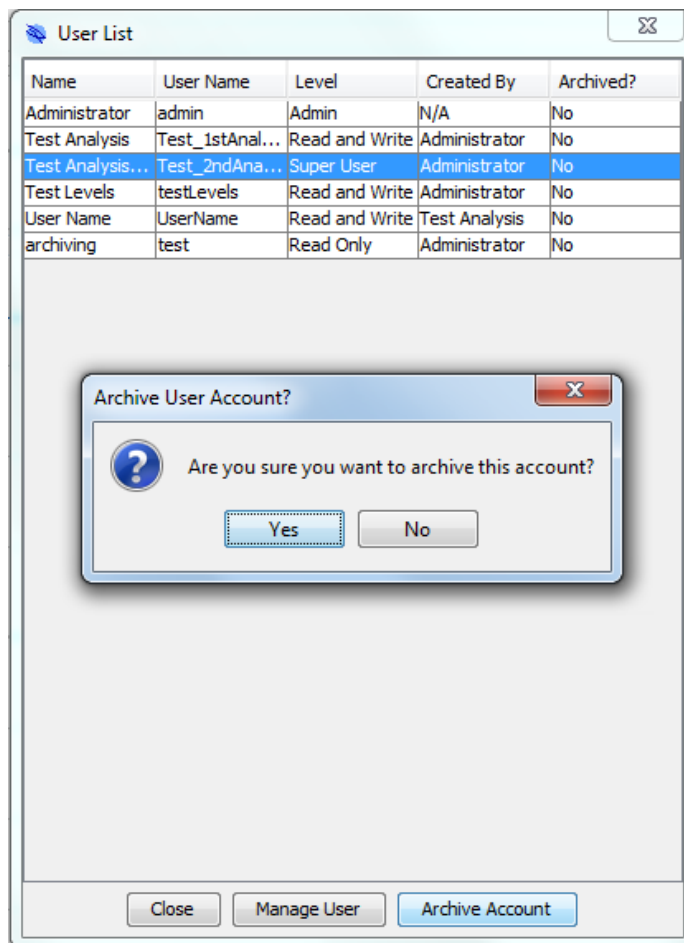


Figure: Archiving of a user account.

An archived account can be restored by a managing account by returning to the **User List** window, selecting the user from the list and clicking the **Restore Account** button.

Miscellaneous

Metric threshold set linked to protocol

Metric threshold sets can now be linked with analysis protocols to ensure consistent comparison of metric values with expected results across samples. To link a protocol with a particular metric threshold set, select **Tools -> Options... -> Protocols** and select the require threshold set from the **Metric Set** drop-down box. As the metric set is an analysis protocol option, a new protocol will also need to be created by entering a new name for the protocol and clicking **Create Protocol**.

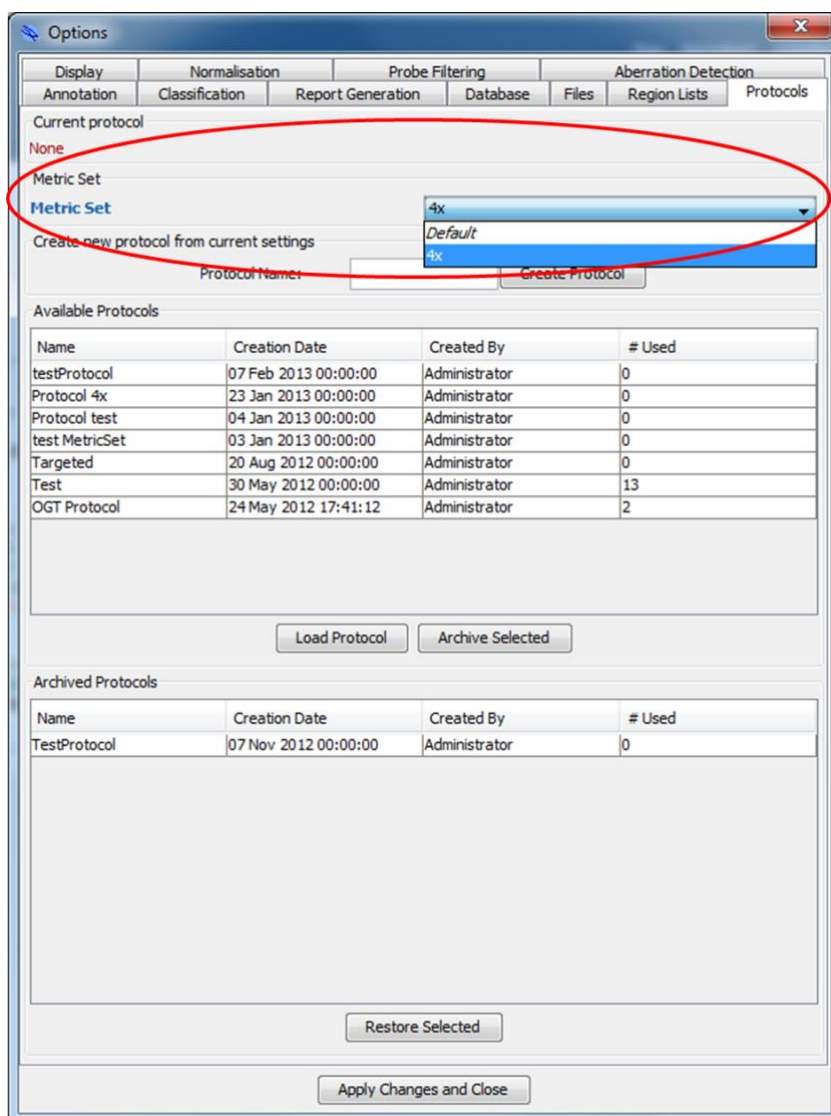


Figure: Modification of the metric threshold set via the Protocols tab of the Options window.

Percentage homozygosity

The software now calculates the total percentage homozygosity of each sample containing SNP data based on the method described by Sund *et al* (Genet Med 2013;15[1]:70–78) as a new metric named **Homozygosity**.

Expected barcode fields

For new installations, a field named **Expected Barcode** has been added by default in the Sample and Experiment sections in order to facilitate linking of sample and experiment information to



feature extracted data via the barcode. See Automatic Sample ID Detection in Batch Processing for more details.

Links to external files

Users can now create custom sample and experiment fields for linking to external files (e.g. PDF reports or image files). To do this, create a new custom field with the “File” data type and modify its value by clicking on “Unknown” and selecting the file from the file system. Left-clicking on the file name will show the location of the file in your file system, and right-clicking and selecting **Open** will open the file in the default program.

Sample Information panel customisation

It is now possible to customise the information displayed in the **Sample Information** panel in the **Genomic View** by clicking on the **Customise** button and selecting the sample and experiment fields to be displayed in this section.



CytoSure™ Interpret Software update — version 11.0 October 2012

The following new features are available in CytoSure Interpret Software version 4.3:

B-Allele Frequency display and analysis	110
CNV snapshot regeneration	111
Automated dye-swap analysis	111
Bulk upload of CytoSure ISCA Sample Tracking Spike-in information	112
Miscellaneous	113



B-Allele Frequency display and analysis

In order to facilitate more in-depth analysis of allelic imbalance within samples that have been assayed with one of OGT's CytoSure CNV +SNP arrays, CytoSure Interpret Software v4.3 includes the display and analysis of the B-Allele Frequency (BAF) of each SNP included in the design.

B-Allele Frequency

The B-Allele Frequency is a normalized measure of the allelic intensity ratio of two alleles (A and B), such that a BAF of 1 or 0 indicates the complete absence of one of the two alleles (e.g. AA or BB), and a BAF of 0.5 indicates the equal presence of both alleles (e.g. AB).

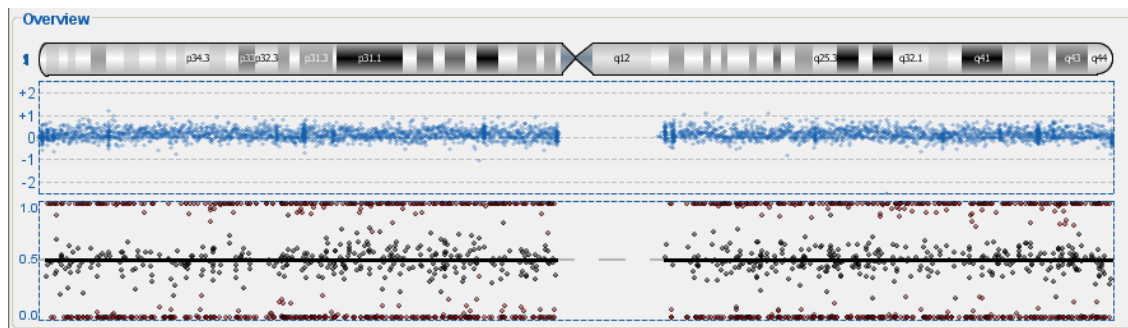


Figure: A screenshot of a section of the Overview section of the Genomic View showing, from top to bottom, the chromosome ideogram, the CNV probe log ratio scatter plot, and the SNP probe B-Allele Frequency plot.

Previous visualisation and analysis of SNP data within CytoSure Interpret Software enabled detection of regions of LOH (containing only BAF values around 0 or 1), but did not enable detection of allelic imbalances such as those caused by duplications (e.g. AAB/BBA) or mosaic deletions in the sample. Such imbalances can be identified on a BAF plot by the presence of SNPs at frequencies between 0.5 and 0 or 1. For example, the theoretical BAF values of triploid regions (AAA, AAB, ABB or BBB) are 0, 0.33, 0.66 and 1 respectively.

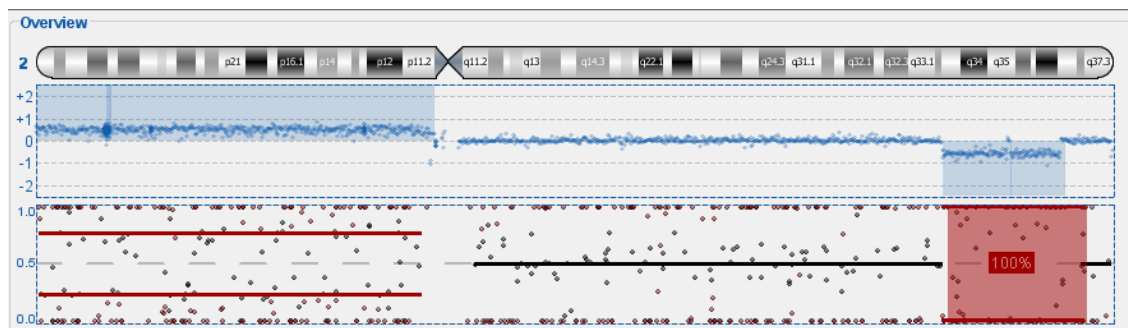


Figure: An example of a duplication of 2p resulting in BAF values of approx 0, 0.33, 0.66 and 1, and a hemizygous deletion on 2q resulting in BAF values of 0 and 1.



In order to automatically detect such regions, CytoSure Interpret Software segments the BAF data mirrored at 0.5 (see Staaf et al., Genome Biology 2008, **9**:R136/ <http://genomebiology.com/content/9/9/R136>) and highlights regions showing an absence of heterozygous SNPs (i.e. SNPs with BAF values of approx. 0.5). In the case of losses and copy neutral events, the software also calculates the theoretical fraction of cells showing the allelic imbalance and displays this value in the BAF plot.

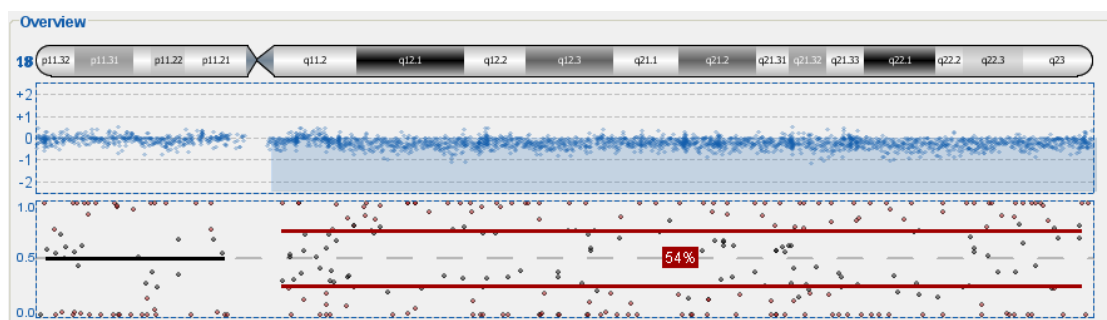


Figure: An example of a mosaic deletion of 18q resulting in BAF values of 0, 1 and approx. 0.32 and 0.68, indicating ~54% aberrant cells.

Please also note that areas shaded in red indicate regions of LOH.

Display options

The B-Allele Frequency plot can be displayed in both the **Genomic View** tab (next to the summary of the CNV probes for the selected chromosome) and the **CNVs/SNPs** tab (on selecting an individual chromosome from the ideogram view grid). To toggle the display of the B-Allele Frequency plot in the **Genomic View** tab, right-click on the CNV probe summary plot in the **Overview** section and select **Show/Hide B-Allele Frequency Plot**. In the **CNVs/SNPs** tab, ensure that the **Ideogram View** and the **Show CNVs?** checkbox are selected, select a chromosome containing SNP probes, then right-click anywhere in the **Detected CNVs/SNPs** section. This will provide the option to either **Show B-Allele Frequency Plot** or **Show Distribution Plot** depending on your current configuration.

CNV snapshot regeneration

It is now possible to regenerate the CNV snapshots that are included in HTML and PDF reports and stored in the database to enable, for example, the inclusion of additional annotation tracks in the snapshot after its initial generation following reporting or database submission. In order to generate these snapshots for a particular CNV or group of CNVs, select the appropriate row(s) of the table in either the **Aberrations** section of the **Genomic View** tab or the **Detected CNVs/SNPs** section of the **CNVs/SNPs** tab (ensure that **Table View** is selected), then right-click on the selection and click **(Re)Generate Snapshot(s)**. Subsequent report generation or database submission will include the latest snapshots.


Automated dye-swap analysis

The **Loop Analysis** module has now been extended to enable automatic detection of matching aberrations in dye swap experiments. In order to use this feature, load both sets of feature extracted data into CytoSure Interpret Software, ensure that, for each dataset, the test sample data is associated with correct dye (i.e. one is Cy3 and one is Cy5), then select **Tools -> Analyse Loop**



Design -> Analyse. All regions containing aberrations detected in both datasets will be added to the list of CNVs for the first dataset with the detection method set to **Loop Analysis**.

Bulk upload of CytoSure ISCA Sample Tracking Spike-in information

It is now possible to import spike-in information, along with other sample and experiment details, using the [Attach](#)  button (see page 144) of the **Database Management** tab. To set up the import, first ensure that the spike-in identifiers (uppercase letters from A-P) are included in a column in the file to be parsed, with multiple spike-ins on the same sample separated by commas. To configure a parser to import the spike-in information, simply select the column including the spike-in(s) and associate it with the new **Basic Field** named **Spike-In(s)**.

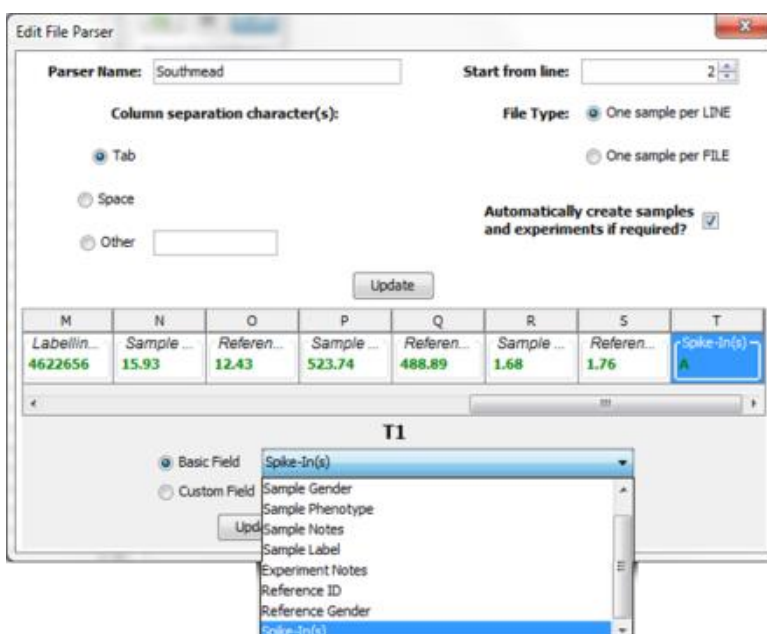


Figure: Importing spike-in information via the Attach button

Please note that spike-in ratio checking is only carried out once the user has chosen whether to apply dye bias normalisation, and therefore spike-in ratios will appear as ? in the **Spike-in Selection** screen (available by clicking on the **Edit** button next to the **Spike-ins** field in the **Sample Details** dialog box) before this decision has been made.

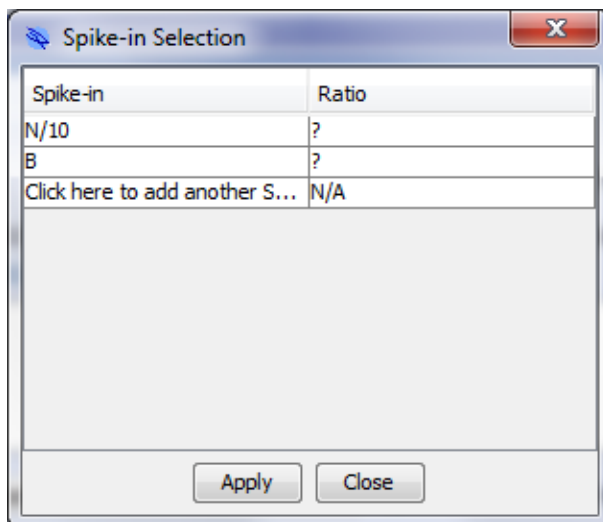


Figure: Spike-in ratios are unknown until the user has decided to whether to apply dye bias normalisation, and therefore appear as “?”.

Miscellaneous

Annotation track collapse/expansion

Users can now collectively expand or collapse all of the currently activated annotation tracks by right-clicking anywhere on the scatter plot in the **Chromosome Section** of the **Genomic View** tab and selecting **Expand all tracks** or **Collapse all tracks**.

Inclusion of custom sample fields in Sample Details dialog box

CytoSure Interpret Software now allows users to view the custom fields associated with the sample and the reference in the **Sample Details** dialog box. When a feature extracted file is imported into the software and the **Sample ID** is selected or created for the sample or reference, the custom details associated with the samples will appear under the basic details, as shown in the figure below. Please note, however, that these custom details are only editable via the **Database Management** section.

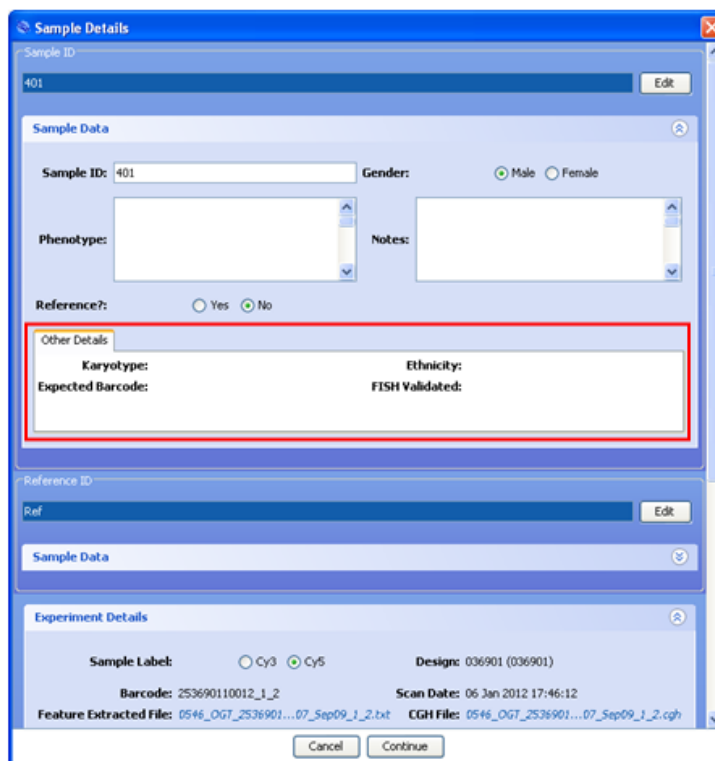


Figure: The Sample Details dialog box with the custom sample fields displayed.

Batch processed file naming

Files generated by the batch processing feature of the software are now named consistently with the file naming options specified via **Tools -> Options -> Files**. As such, these files are automatically saved in the specified location using the **Raw data file name**, the **Sample ID** or the **Barcode**, depending on the selection made by the user.

Project deletion/renaming

Further options have been added to the **Batch Operations** utility within the **Database Management** tab. The projects in the database can now be deleted as well as renamed using the corresponding option via the **Projects** sub-menu.

Parser deletion/renaming

Similarly, file parsers used in the loading of sample and experiment information can now be deleted or renamed via **Database Management -> Batch Operations -> Parsers -> Rename/Delete**.



CytoSure™ Interpret Software update — version 10.0 August 2012

The following new features are available in CytoSure Interpret Software version 4.2:

Automatic Sample ID Detection in Batch Processing	116
BENCH data submission upgrades	117
Additional logging to facilitate HIPAA compliance	118
LOH region annotation reporting	118
Probe list generation via the Genomic View	118
Miscellaneous updates	119



Automatic Sample ID Detection in Batch Processing

By utilising a combination of two of CytoSure Interpret Software's features, it is now possible for the software to automatically associate feature extracted data with samples in the CytoSure Interpret Software database. The process involves two phases:

1. Loading sample and barcode information into the database via the parser utility in the **Database Management** tab so that each sample or sample experiment is annotated with the barcode of the array on which it is located.
2. Identifying the field containing the barcode before commencing **Batch Processing** so that the software can automatically determine which sample to associate with the feature extracted file using its barcode information.

Initial Setup

Creating an "Expected Barcode" field

In order to enable automatic association with feature extracted data, a sample must already be annotated with its array barcode. To facilitate this, a custom Sample or Experiment field must be created in which the array barcodes will be entered. To create the custom field, select the **Sample Data** or **Experiment Details** tab in the **Database Management** section, select a category in the custom fields section (the tabbed areas below the **Basic Information** section), or create a new category, and select the **Add new field** button. Enter "Expected Barcode" as the name of the field and click **OK**, then select **Text – Single Line** as the data type for the new field and click **OK**. The field is now ready to be populated with the expected barcode for each sample to be batch processed.

Creating a parser to load barcode information for each sample from a text file

It is possible to manually enter the expected barcode of each sample by selecting the sample in the **Database Management** section, selecting the previously-created Expected Barcode field, entering the expected barcode and finally saving the changes. However, creating a parser to automatically extract this information from a text file (perhaps generated from a LIMS system or existing spreadsheet) reduces the possibility of incorrect barcode entry or associating a sample with the wrong barcode. The procedure assumes that it is possible to generate a text file containing, at a minimum, the sample ID and the barcode of the array on which the sample was located. If applicable, the barcode should be suffixed with the row and column location of the array, separated by underscores (e.g. 252757400001_2_3 for the array in column 3 of row 2 of the slide with barcode 252757400001).




	A	B
1	Sample ID	Expected Barcode
2	ABC-123	252757400001_1_1
3	EFG-456	252757400001_1_2
4	HIJ-789	252757400001_1_3
5	KLM-012	252757400001_1_4
6		

Figure: An example of a spreadsheet containing Sample ID and barcode information

Once the text file has been created follow the process of creating a file parser in the [Attach section](#) on page 144.

Routine Use

Once the custom field and file parser have been created, for each set of feature extracted files to be batch processed:

1. Load the data associated with the files into the database using the file parser via the  button in the **Database Management** tab.
2. Select **Tools -> Batch Processing** and import the set of feature extracted files so that they are listed in the **Files to be processed** table (either via the **Add Files...** button or by dragging and dropping from Windows Explorer).
3. Click the **Process** button.
4. When presented with the **Missing details** dialog box, select **Match barcodes to existing sample information**.
5. Select the **Expected Barcode** field from the list presented and **OK** to begin batch processing.

If the software is not able to detect the sample associated with the barcode of a feature extracted file, it will display an error in the **Status** column.

BENCH data submission upgrades

Additional Information

The following additional information is now transmitted during a BENCH submission:

- LOH region information from SNP data
- Inheritance classification for CNVs
- Maximum bounds of CNVs and LOH regions



- Confirmation method

Bulk Submission

It is now possible to submit multiple cases to Cartagenia BENCH via the **Database Management** tab by carrying out the following process:

1. Ensure that all cases to be submitted to Cartagenia BENCH have already been analysed and submitted to the CytoSure Interpret Software database.
2. Create and apply a database filter so that only the cases that should be submitted to BENCH are listed in the Samples table in the **Database Management** tab. The filter created may vary depending the cases have been annotated, but the following process describes how to restrict the list of cases to those that have been submitted to the CytoSure Interpret Software database within a specified date range:
 - a. Expand the **Advanced Search** pane.
 - b. Select the **Experiment Details** radio button.
 - c. Under **Field Name**, select **Analysis Date**.
 - d. Under **Filter Type**, select **between**.
 - e. Using the date selector, choose a start and end date.
 - f. Enter a name for the filter.
 - g. Click the **Create Filter** button – the filter should appear in the **Existing filters** table.
 - h. Select the newly-created filter in the table and click the **Apply** button.
3. When only those cases that need to be submitted to BENCH are listed, select **Batch Operations** -> **Cartagenia BENCH** -> **Submit**. A dialog box will appear confirming the number of samples that have passed the filter and whether to submit the data to Cartagenia BENCH.
4. Click **Yes** to carry out the submission.

Please note that the above process assumes that the details of a BENCH account have already been entered via the **Cartagenia BENCH** tab.

Additional logging to facilitate HIPAA compliance

In addition to the audit trail created by users when creating, deleting or editing sample and experiment information, CytoSure Interpret Software now also logs instances when a user views the **Sample Data** or **Experiment Details** sections of the **Database Management** tab for a particular sample. This feature has been added so that access to Electronic Public Health Information (e-PHI) that may potentially be stored in these two areas of the database is logged by the system to help users to comply with HIPAA regulations.

LOH region annotation reporting

It is now possible to customise reporting of annotation information based on the location of LOH regions in the same manner as CNV regions via **Tools** -> **Options...** -> **Annotation** -> **LOH Reporting**.

Probe list generation via the Genomic View

Probe lists can now be generated by either:


1. Right-clicking in the **Genomic View** scatter plot and selecting **Add all probes to probe list**.



2. Right-clicking on an individual probe within the **Genomic View** scatter plot and selecting **Add to probe list**.

In both cases, the option to add the probe or probes to an existing probe list or to create a new probe list is provided.

Miscellaneous updates

- **Annotation colour change options** – Users now have the option of controlling how the software changes the colour of annotations according to their copy number status (red for losses; blue/green for gains) when the mouse hovers over an annotation track. The **Display** tab in the **Options** dialog box (accessible via **Tools -> Options...**) provides two controls relating to this function:
 1. **Enable On-Hover Annotation Colour Changes** – If this checkbox is deselected, the above functionality will be disabled and no annotation colour changes will occur.
 2. **Invert On-Hover Annotation Colour Changes** – If this checkbox is selected, annotation tracks will be coloured according to their copy number status by default (of applicable).
- **Opening .cgh files in workflow mode via database management** – Clicking on the  icon in the **Database Management** tab now automatically attempts to load the case into Workflow mode if possible.
- **Active Directory Authentication for MS SQL Server** – It is now possible for CytoSure Interpret Software to connect to Microsoft SQL Server databases using “Active Directory Authentication”, whereby the system provides access using the user’s Windows account, circumventing the need for a dedicated user name and password to connect.



CytoSure™ Interpret Software update — version 9.0 January 2012

The following new features are available in CytoSure Interpret Software version 4:

Contents

• Introduction	49
• Loading data	49
• Experimental details	51
• Understanding and managing user permissions	52
• Changing the genome build	52
• Protocols	53
• Updates to genomic view	56
• Annotation tracks	57
• Database management	63
○ XML database conversion	63
○ Sample data	68
○ QC metrics	68
○ CNV analysis	69
○ SNP analysis	70
○ Audit trail	71
○ Menu buttons	74
○ Batch operations	75



Introduction

This latest version of CytoSure Interpret Software contains a major upgrade to the Database Management module to provide scalability, flexibility and improved performance, as well as support for analysis with multiple genome builds. The software has changed from using a text-file based database to a robust relational database management system, which is ideally suited to concurrent access by multiple users across a network, and will further increase the usability of the software.

Loading data

To maximise the utility of the new database structure it is necessary to link each case to be analysed to both a sample and a reference. Making this link reduces hands-on analysis time by increasing the number of processes in the software that can be automated.

On loading a new array design into CytoSure Interpret Software a prompt will appear requesting the design file (.xml) to be loaded. This will only need to be done once. During installation it is possible to select design files from a list of catalogue products. For older designs or custom arrays, the design file can be found on the CD that is supplied with the arrays.

Loading samples — feature extracted files

When loading feature extracted files, after selecting the file to be loaded, the **'Sample Details'** dialog box appears. It is necessary to identify what sample has been run on the array. Type the name of the sample in the Sample ID box and click **'Create New Sample'** (Figure 1). It is also necessary to type in the reference details.



Figure 1: Linking samples and references to experiments when loading feature extracted files

Loading samples – batch processed files

Samples can also be linked to data sets when loading using the batch processing function, found under the **'Tools'** drop down menu. After selecting **'Batch Processing'**, a table appears which allows you to select the files to be analysed (Figure 2). Once files have been added, double-clicking on an individual file name in the **'Files to be processed'** table will automatically open up the **'Sample Details'** dialog box giving the option to link the dataset to a previously-created or new sample and reference (Figure 2).

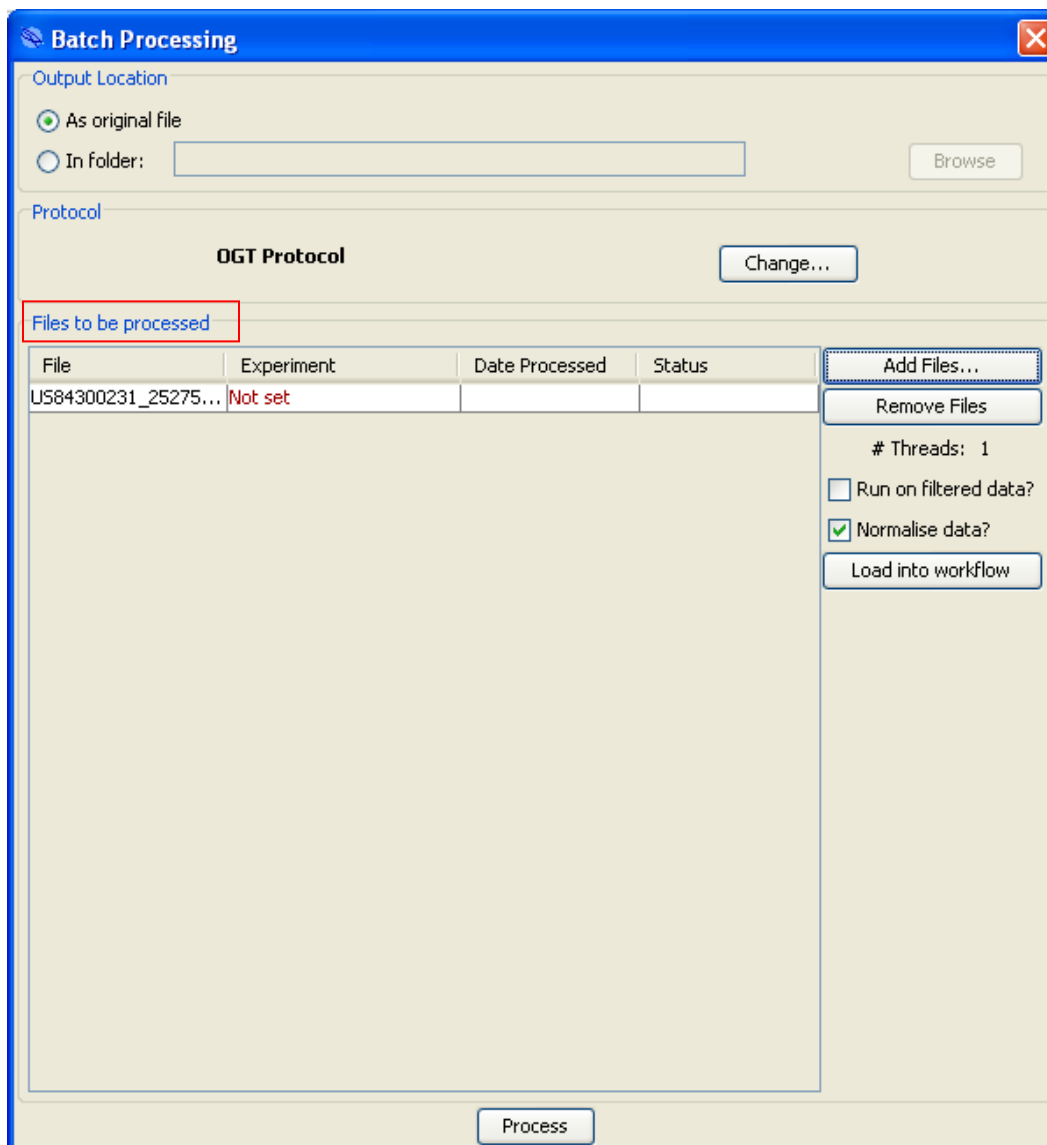


Figure 2: Linking samples and references to experiments through batch processing

Experimental details

On loading feature extracted data it is possible to add in experimental details, for example details about how the array was processed or the sample type used (Figure 1). It is also possible to input or edit this information from the **'Database Management'** tab. On loading batch processed files experimental information can only be added under the **'Database Management'** tab in the **'Basic Information'** window as described below.



Understanding and managing user permissions

It is now possible for each user to be assigned specific permissions or privileges within CytoSure Interpret Software. In addition, the associated user log-in provides an audit trail allowing all workflow changes to be tracked. These features ensure data integrity and that system settings are maintained. The administrator account is automatically created during installation.

To access the Administrator functions select '**Logged in as Administrator**' from the menu bar (Figure 3).

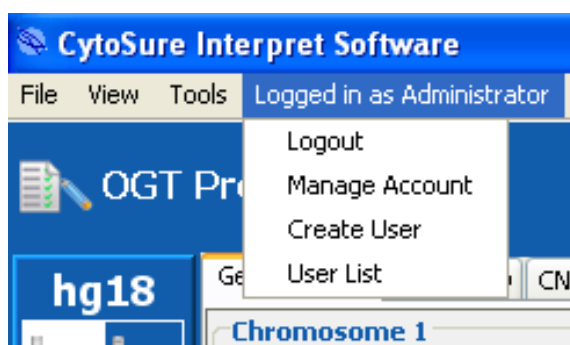


Figure 3: Administrator functions

The following user types can be created:

- Administrator Can create Super User, Read & Write and Read Only Accounts
- Super User Can create Read & Write, Read Only Accounts
- Read & Write Can create Read Only Accounts
- Read Only Cannot create accounts

Read and write permissions allow a user to edit the protocol settings and to change the classifications of an aberration that has been submitted to the database. Read only accounts are not able to make any changes to the protocol settings or the classification of aberrations.

To change the permissions or password associated with a particular account or user, select 'Logged in as **[User Name]**' from the tool bar. Then select the desired user from the user list and select the '**Manage User**' option.

Changing the genome build

It is now possible to easily switch between genome builds (hg18 and hg19) without repeating the feature extraction step or duplicating the analysis. This is achieved by storing the probe co-ordinates for each array design in the database rather than using the co-ordinates contained in feature extracted files (as these may not be of the desired genome build). For new array designs or new genome builds it may be necessary to load the design file (.xml) into the database. On installing CytoSure Interpret Software it is possible to select from a list of design files for the most commonly used catalogue products. If the required design file is not present in the installation list



it is possible to load a design file after installation. All arrays are shipped with a CD containing the appropriate design file.

Array data from hg19/GRCh37 as well as hg18/NCBI36 can be analysed. To switch between different builds, left click on the '**Current Genome Build**' button in the top-right corner of the software and select the desired build (Figure 4). It is possible to add new genome builds to CytoSure Interpret Software, for example the screenshot below shows the mouse genome build mm9 for users analysing mouse data.

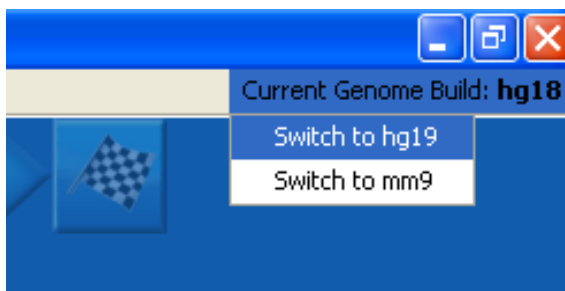


Figure 4: Changing genome build

Protocols

When working in the Options dialog box, found under '**Tools**' on the main tool bar, any changes made to fields with bold blue text represent a change in the analysis protocol and will need to be saved under a different name within the '**Protocols**' tab (Figure 5). Standard text options can be altered without the need for the protocol to be saved.

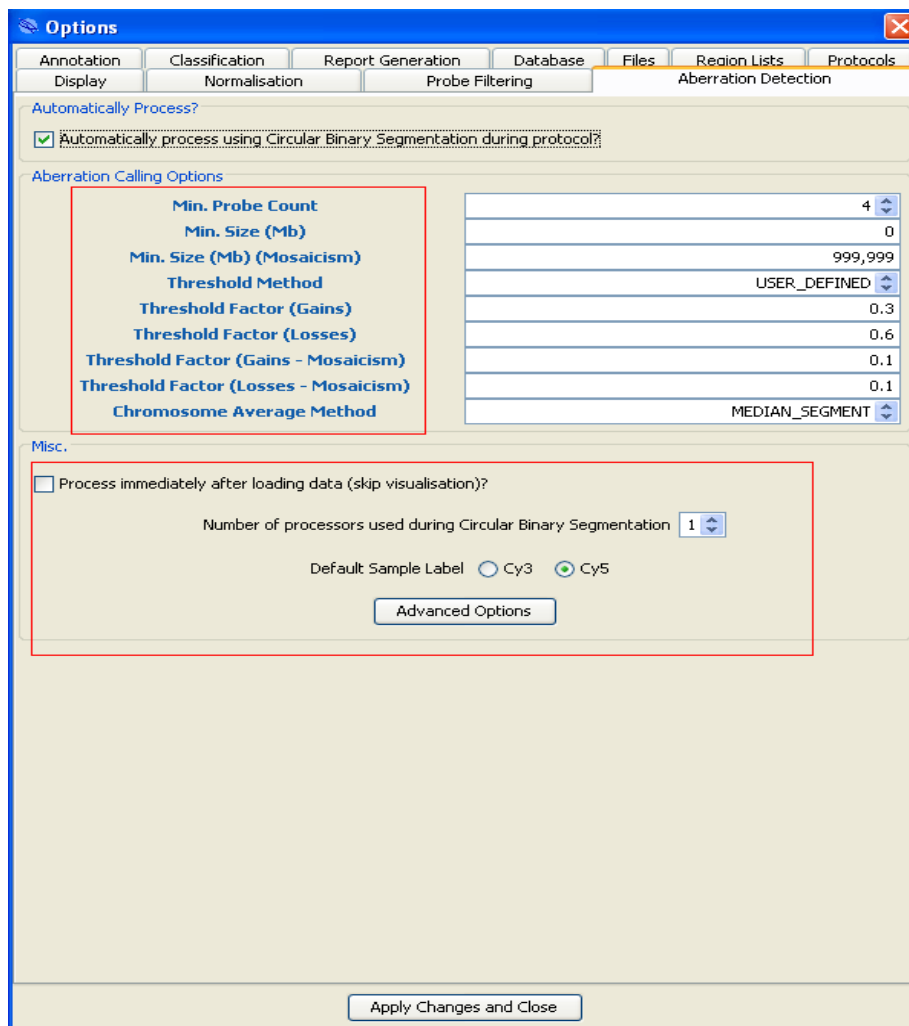


Figure 5: Example of fields which, if changed, affect the data analysis output. Blue text will change how aberrations are detected. Black text will change the timing of data analysis and how the data is processed.



Under **'Tools', 'Options', 'Protocols'** old or unused protocols can be archived by selecting a protocol and selecting **'Archive Selected'** (Figure 6). The protocol in use will be shown in bold text.

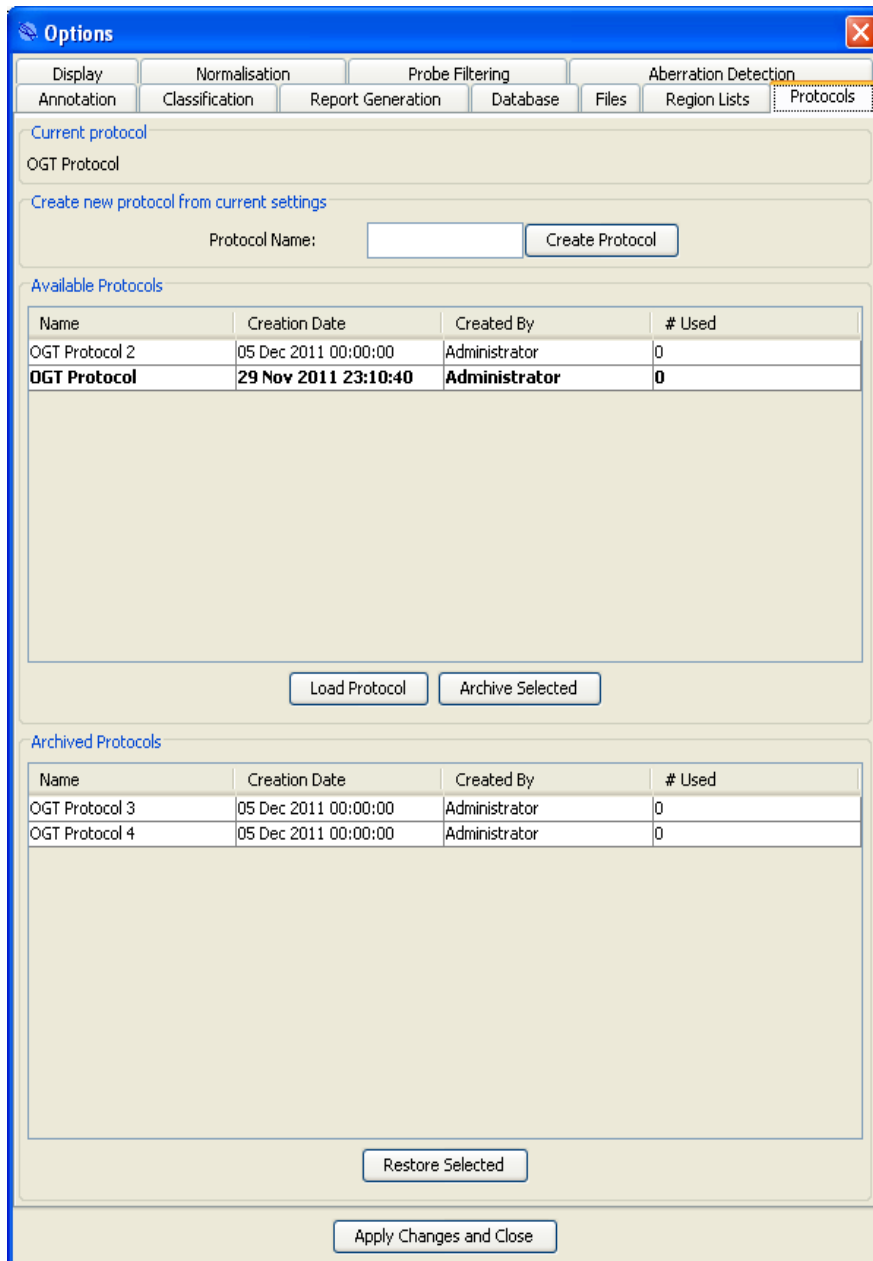


Figure 6: Protocol archiving

To restore a protocol, highlight the desired protocol and click on **'Restore Selected'**.
Updates to genomic view



Within 'Genomic View' (Figure 7) it is possible to alter how the data is displayed.

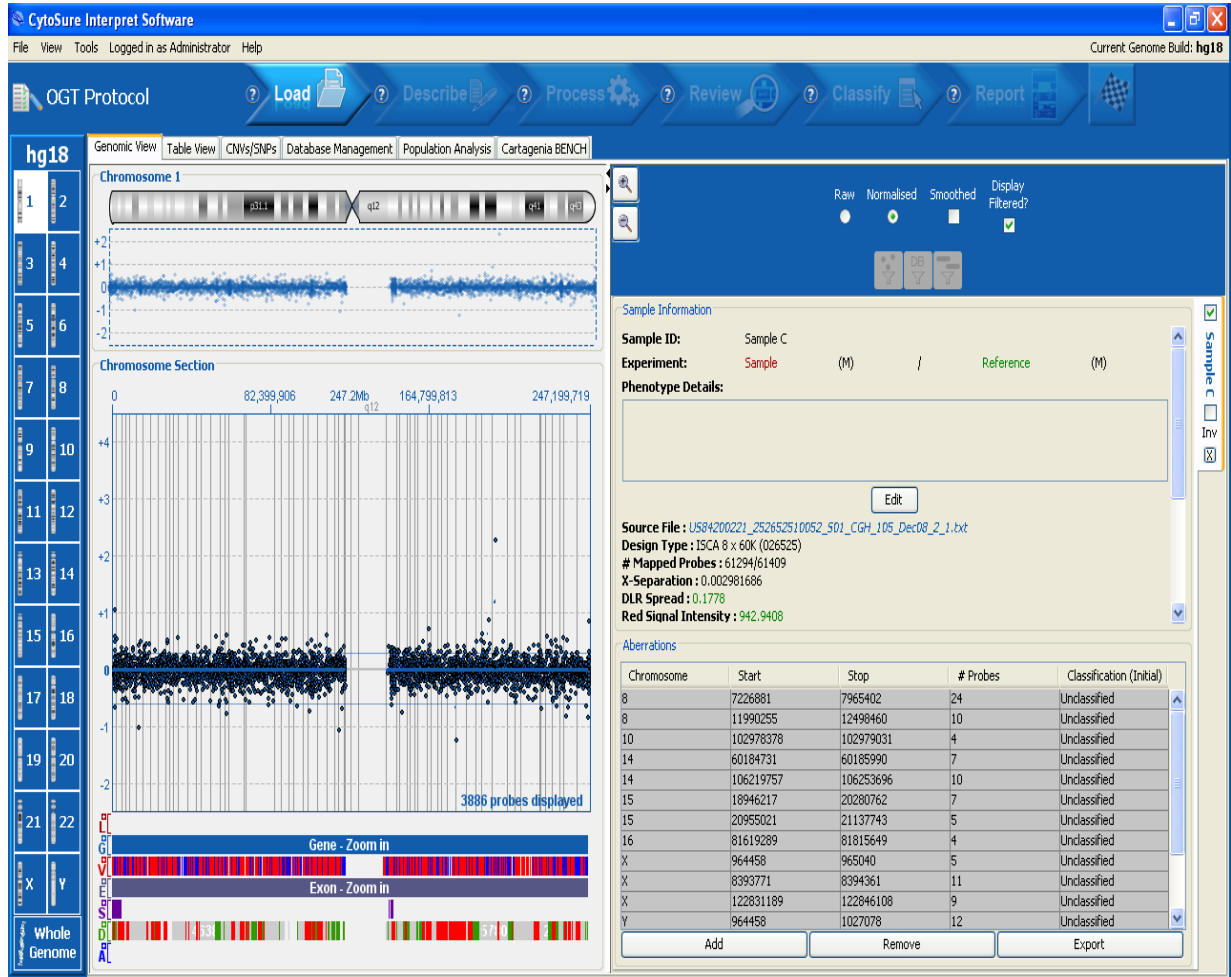


Figure 7: Customising the 'Genomic View'

The size of the windows can be altered by clicking on the pane divider and dragging the divider to the desired position. It is also possible to customise the aberration table by right-clicking on any of the columns. A list appears allowing the columns to be selected or de-selected.



Annotation tracks

The following annotation tracks are supplied with the software:

	Constitutional	Cancer	Molecular
Genes (Ensembl)	x	x	x
Exons (Ensembl)	x	x	x
Syndromes (DECIPHER)	x		x
Copy Number Variations (Database of Genomic Variants)	x	x	x
ISCA Variations (ISCA dbVar)	x		
CLL Regions (Mitelman Database)		x	
Cancer Gene Census Genes (COSMIC)		x	
DECIPHER Patients (DECIPHER)	x		
Segmental Duplications (Segmental Duplication DB)	x		x
Haematology Regions (Atlas of Genetics and Cytogenetics in Oncology and Haematology)		x	
Confirmation Tests (Various)	x	x	x
Haploinsufficient Genes (Hurles <i>et al.</i>, 2010)	x	x	x
Eichler Variations (Eichler dbVar)	x		

Table 1: Tracks available in CytoSure Interpret Software grouped by application type.

The ISCA track enables easy links to the information stored in the ISCA database. The haploinsufficiency track shows the locations of genes predicted to be deleterious if only a single functional copy is present, please note the scores are theoretical and are based on a model described in detail in: <http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1001154>

(Hurles, M. *et al* [2010] Characterising and predicting haploinsufficiency in the human genome. PLoS Genetics 6, 10, e1001154 1-11).

It is possible to install all of the available tracks or to just select a group of tracks. Tracks have been grouped by research application as shown in Table 1.

It is also possible to create custom Annotation tracks in addition to the available Annotation tracks

Track width

Within the **'Display'** tab under **'Options'** found under **'Tools'** on the main toolbar there is an option to change the width of the annotation tracks, allowing you to customise your screen view. The default setting is 17.



Track settings

All annotation tracks within CytoSure Interpret Software are now customisable. In the ‘**Annotation**’ tab under ‘**Options**’ found under ‘**Tools**’ on the main toolbar. Individual tracks can be edited or created (Figure 8). To simplify the creation process a ‘Duplicate’ button has been added. For example, to create the two versions of the NCBI AceView track it is easy to create one version, duplicate and then edit the link information and build descriptions.

The screenshot shows the 'Options' dialog box with the 'Annotation' tab selected. It contains two main sections: 'Annotation Tracks' and 'Annotation Links'.

Annotation Tracks

Name	Initial	Colour	Report	Report #	Report %
Gene	G	Blue	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Copy Number Va...	V	Red	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Exon	E	Purple	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Syndrome	S	Dark Purple	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
CHOP	C	Cyan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Example Track	E	Light Green	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
ISCA	I	Magenta	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Aberration	A	Dark Blue	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Annotation Links

Name	URL	Build	Track
PubMed	http://www.ncbi.nlm.nih.gov/pubme...	All	Copy Number Varia...
Ensembl	http://www.ensembl.org/{#organis...	hg19 only	All
UCSC	http://genome.ucsc.edu/cgi-bin/hgTr...	All	All
DGV	http://projects.tcag.ca/cgi-bin/vari...	All	All
DECIPHER	https://decipher.sanger.ac.uk/searc...	hg19 only	All
OMIM	http://omim.org/search?index=entry...	All	Gene only
iHOP	http://www.ihop-net.org/UniPub/iHO...	All	Gene only
GeneCards	http://www.genecards.org/cgi-bin/c...	All	Gene only
GeneRIF	http://www.ncbi.nlm.nih.gov/sites/e...	All	Gene only
Wiki-Gene	http://andromeda.gsf.de/wiki/index....	All	Gene only
NCBI AceView	http://www.ncbi.nlm.nih.gov/IEB/Re...	hg18 only	Gene only
NCBI AceView	http://www.ncbi.nlm.nih.gov/IEB/Re...	hg19 only	Gene only

Buttons: Import new track, Remove selected, Edit selected, Create, Duplicate, Edit, Delete, Apply Changes and Close.

Figure 8: Annotation track settings



There are three new descriptive options within **'Import New Track'**

- Gain/Loss
- Strand
- Colour

The following information can be added to the track to ensure the data is correctly displayed within CytoSure Interpret Software (Figure 8):

Gain/Loss

- Gain (+), duplication, dup, gain
- Loss (-), deletion, del, loss

Strand

- Positive strand (+), 1
- Negative strand (-), -1

Colours

Red, Green, Blue etc. or Hexadecimal colours

It is possible in Excel® to create a file which automatically links the gain or loss column with the correct colour.

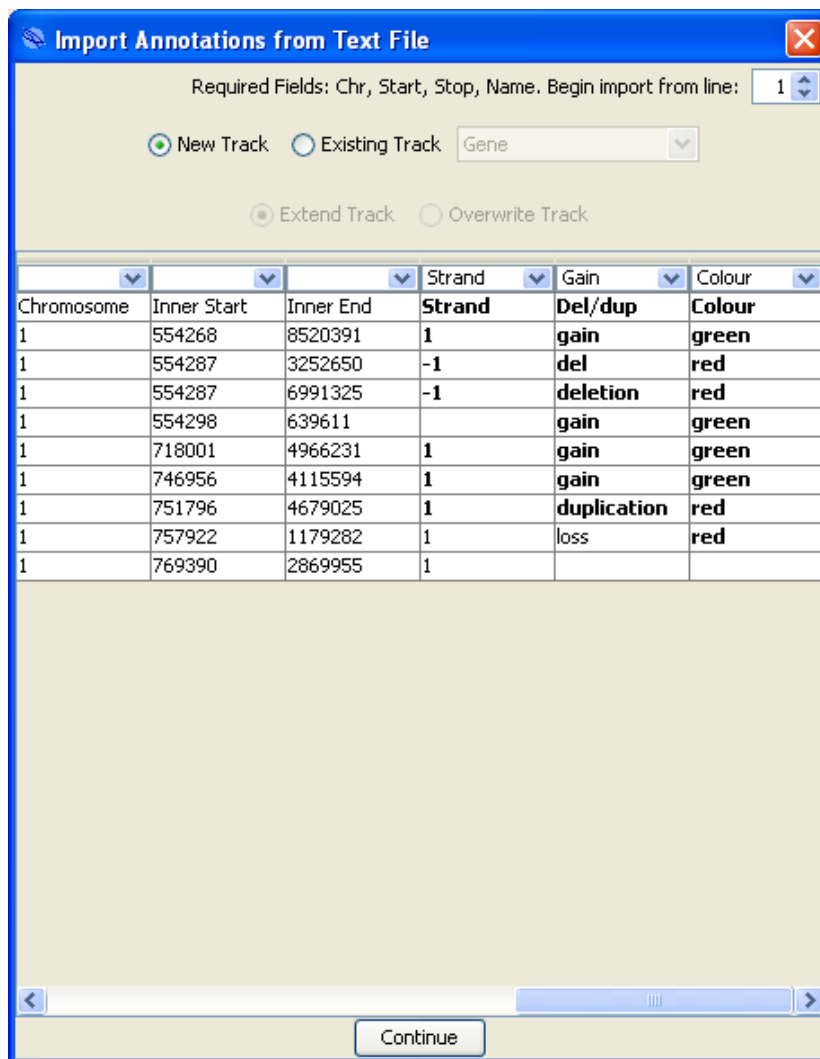


Figure 9: Importing tracks — additional annotation information

Under ‘Options’, ‘Tools’ and ‘Annotations’ selecting the ‘Reporting’ tab allows the user to define what tracks, and what information from each track, should be included in reports. As well as allowing customisation of reports, the information selected will also be available as a column in the aberration tables in the ‘Genomic View’ tab, ‘CNVs/SNPs’ tab and ‘Table’ tab of the ‘CNV Analysis’ section of the ‘Database Management’ tab. The following categories can be selected:

- **Report** Includes regions of the annotation track which are contained in the aberrant region
- **Report #** Reports number of regions within the track within the aberrant region
- **Report %** Includes the percentage of the aberrant region covered by the annotation track



Note: For CNV tracks and aberration tracks, only relevant information will be reported. For example, if a deletion has been detected, only track information matching that deletion will be reported.

Editing tracks

Highlighting a track and selecting **'Edit selected'** allows a track to be edited (Figure 10).

Figure 10: Editing annotation track details

The editable options include:

- **URL Composition** Gives a direct link (via left click on the track) to a web page or database
- **Maximum Track Width** Limits the number of entries within a track that are displayed stacked on top of each other. If the width limit is exceeded, a scroll bar can be used to view information.
- **Maximum # Displayed** Will only show the desired number of annotations (e.g. 300 will only display track information when fewer than 300 annotations are within the region. Where there are greater than 300 annotations it will be necessary to zoom in to display the track information).



Annotation links

Right click on a track or aberration and select '**Create**' to add a direct link to a specific web site or database (Figure 11).

Figure 11: Creation of annotation links

The following information can be added:

- **Name** For track identification — will be displayed in the main window
- **URL** The URL which will be displayed when this link is clicked. A number of variables can be included in this field to enable the software to dynamically generate URLs based on information relating to the selected annotation track (e.g. annotation name, chromosome, start position, stop position etc.). For example, to create a link that searches the OMIM database for entries that match the name of the selected annotation, the URL could be set to "http://omim.org/search?search={#name}". The text "{#name}" will be replaced with the name of the selected annotation when the user clicks on the link to create, for example, "http://omim.org/search?search=**DMD**". Available variables include:
 - #name – Name of the annotation.
 - #chrName – Chromosome name (e.g. X)
 - #chrInt – Chromosome index (e.g. 23)
 - #start – Base-pair minimum span start position
 - #stop – Base-pair minimum span stop position
 - #maxStart – Base-pair maximum span start position
 - #maxStop – Base-pair maximum span stop position
 - #build – Genome build (e.g. hg18)



Users can also include the values of extra annotation information (i.e. Columns labelled as “Other” during annotation import) in URLs. An example of this is the “PubMed” annotation link for the Copy Number Variation track, whose link is “http://www.ncbi.nlm.nih.gov/pubmed/{PubMedID}”. The same technique can be used to complete the “URL Composition” field of an annotation track.

- **Build** Display link in only in a specific genome build or all genome builds
- **Track** The link can be displayed in the pop-up menu when right clicking on any annotation track or a specific annotation track

Loading tracks into the memory

When navigating to a new location within the genome, annotation track information for the region has to be loaded from the database before being displayed on the screen. This can take a few seconds if many tracks or many annotations need to be displayed. To have particular tracks appear instantly, right-click on the track name (Figure 12) and select ‘**Load into Memory**’.

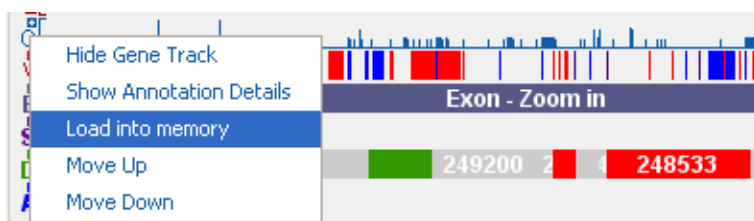


Figure 12: Loading tracks into memory

Database management

The ‘**Database Management**’ tab has been redesigned and now includes additional functionality. There are six sub-tabs within this tab (Figure 13) the functions of which are described in the following sections.

XML database conversion

Databases from previous versions of CytoSure Interpret Software can be converted and loaded into the new database by selecting ‘**Tools**’ and then ‘**Import XML database**’ from the tool bar. A browser window will appear allowing selection of the previous database (.xml file).

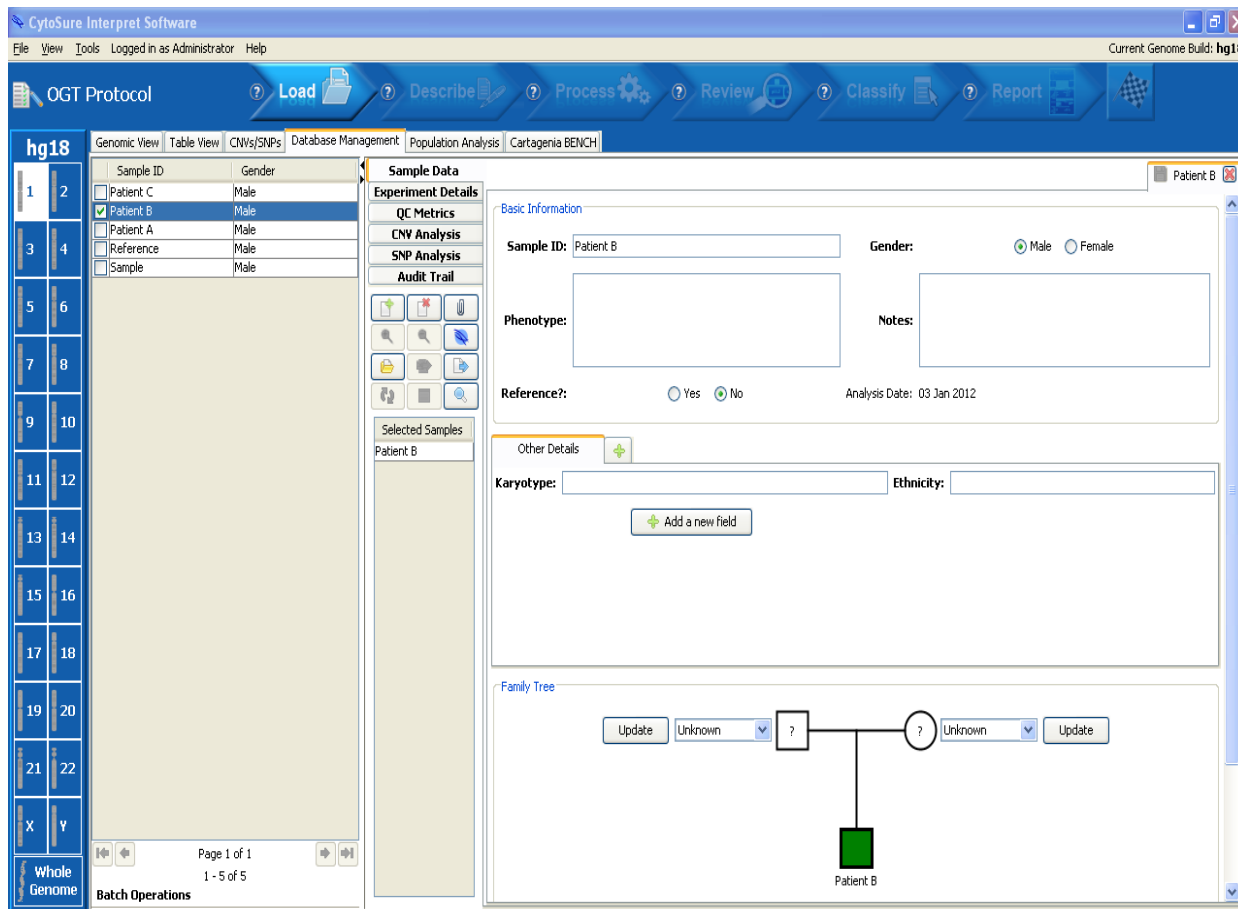
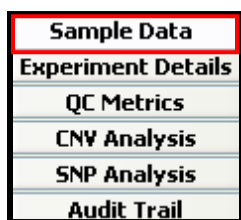


Figure 13: The new Database Management window



The window is divided into two sections; the left-hand pane is a customisable table giving an overview of samples loaded into the database and the right-hand pane displays information about each sample.



Sample table

Sample ID	Gender	Design	#
<input type="checkbox"/> b5cf029ef829...	Male	ISCA v2 (8x60K)	68
<input type="checkbox"/> 9d16d9422d9...	Male	ISCA v2 (8x60K)	70
<input type="checkbox"/> 37b69593c00...	Male	ISCA v2 (8x60K)	73
<input type="checkbox"/> 367580a0a83...	Male	ISCA v2 (8x60K)	81
<input type="checkbox"/> 3ac75d9c1f1...	Male	ISCA v2 (8x60K)	85
<input type="checkbox"/> fdc1b9a0b8c...	Male	ISCA v2 (8x60K)	95
<input type="checkbox"/> 1ab7332b0eb...	Male	ISCA v2 (8x60K)	96
<input type="checkbox"/> 69621668318...	Male	ISCA v2 (8x60K)	113
<input type="checkbox"/> 7e8503b1a78...	Male	ISCA v2 (8x60K)	114
<input type="checkbox"/> be387b0ccb2...	Male	ISCA v2 (8x60K)	115
<input type="checkbox"/> a04fb38719c...	Male	ISCA v2 (8x60K)	124
<input type="checkbox"/> 698b186d0b1...	Male	ISCA v2 (8x60K)	125
<input type="checkbox"/> 4ae95b10907...	Male	ISCA v2 (8x60K)	140
<input type="checkbox"/> b2d86309c30...	Male	ISCA v2 (8x60K)	144
<input type="checkbox"/> 4c0cb5a8d4e...	Male	ISCA v2 (8x60K)	146
<input type="checkbox"/> c626de47439...	Male	ISCA v2 (8x60K)	154
<input type="checkbox"/> ab244d4a28c...	Male	ISCA v2 (8x60K)	157
<input type="checkbox"/> d9868bde342...	Male	ISCA v2 (8x60K)	168
<input type="checkbox"/> 28eddf98dcf...	Male	ISCA v2 (8x60K)	177
<input type="checkbox"/> b8490979341...	Male	ISCA v2 (8x60K)	179
<input type="checkbox"/> cc3e3d74696...	Male	ISCA v2 (8x60K)	180
<input type="checkbox"/> 6d870a90c4b...	Male	ISCA v2 (8x60K)	188

Sample Data

- Sample ID
- Gender
- Phenotype
- Mother
- Father
- Reference?
- #
- Design
- Barcode
- Analysis Date
- Analysed By
- Scan Date
- Analysis Protocol
- Software Version

Figure 14: Sample table with customisable column headings

The composition of the sample table (Figure 14) can be customised by:

- Adding or removing available columns by right-clicking on the column heading
- Changing the order of selected columns by dragging the column headings to the required position.

The software will remember the user's preference.

Left-clicking on any header will also sort the column in an ascending or descending order. Fifty samples per page will be displayed on the left side of the window.

To display information about a specific sample or a group of samples, tick the check box next to the sample(s) of interest.



Sample searching

To refine the list of samples displayed in the sample table there are two search options. Selecting the sample search icon will display the options, '**Quick Search**' and '**Advanced Search**'. Quick search gives you the option to type anything alphanumeric, and will search through the database for matches in the '**Sample ID**', '**Notes**' and '**Phenotype**' fields. For example, it would be possible to search for samples labelled with a particular batch of labelling kits or to search for a group of samples with the same phenotype. '**Advanced Search**' can be used to construct complex queries in order to filter the database contents. Each generated filter can be displayed as a separate annotation track in the '**Genomic**' or '**Population Analysis**' views. The filters are mirrored as '**Groups**' in the '**Population Analysis**' section. It is possible to combine filters in order to build up queries by selecting multiple filters in the table and clicking the '**Combine**' button and selecting a combination method either '**Union**' (Or) or '**Intersection**' (And).

Sample data window – basic information

The right-hand side of the '**Database Management**' window contains details related to the samples selected in the sample table. If multiple samples have been selected, these are displayed on separate tabs in the right-hand pane.

Basic information fields such as '**Sample ID**', '**Gender**', '**Phenotype**' and '**Notes**' are automatically displayed and can be entered or amended.

In the '**Other Details**' pane it is possible to create custom fields. This is done by clicking on the '**Add a New Field**' icon.

The following functionality is available:

- Left-clicking on the blue arrows within the tab headers will sort the tabs into a specific order.
- Right-clicking directly on a sample tab enables the user to '**Edit**' or '**Delete**' the chosen tab.
- Right-clicking on any created field text gives the user the option to move the field within the category, delete, edit or move to a different category or link the field to a chosen or custom legacy field. When data is entered into any field, this can be saved by clicking the '**Save**' icon on the sample data tab. The saved icon will be greyed out if there are no new changes to save.

Other details

Various data types can populate the custom fields. After typing the name of the new field these are displayed as a drop-down list. For example, it would be possible to create a custom field containing information regarding how and when the samples had been collected or whether any follow-up studies had been done.

- Floating Point Number e.g. (3.3 or –115.5)
- Text – Single Line
- Text – Multiple Lines
- Integer (whole number)
- True or False
- Date
- Time



- Date & Time
- List

Choosing **'List'** gives the option to create lists which can be used to simplify the creation of new fields in the future.

Family tree (pedigree analysis)

This provides the option to link samples together, generating a pedigree chart. It allows the software to automatically determine the inheritance status of aberrations if parent aberrations are already in the database. To create a pedigree, the parental and patient samples need to be open. Select the appropriate samples in the drop-down menu and click on **'Update'** (Figure 15). The pedigree analysis will update automatically when the patient sample is analysed if the parental samples have been analysed previously. If the parental samples have not been analysed before the patient sample, once parental samples have been analysed and linked with the patient sample, select **'Batch Operations', 'CNV Analysis', 'Update Inheritance'**. This action should update the **'Inheritance'** and **'Inheritance Coverage'** fields of the patient.

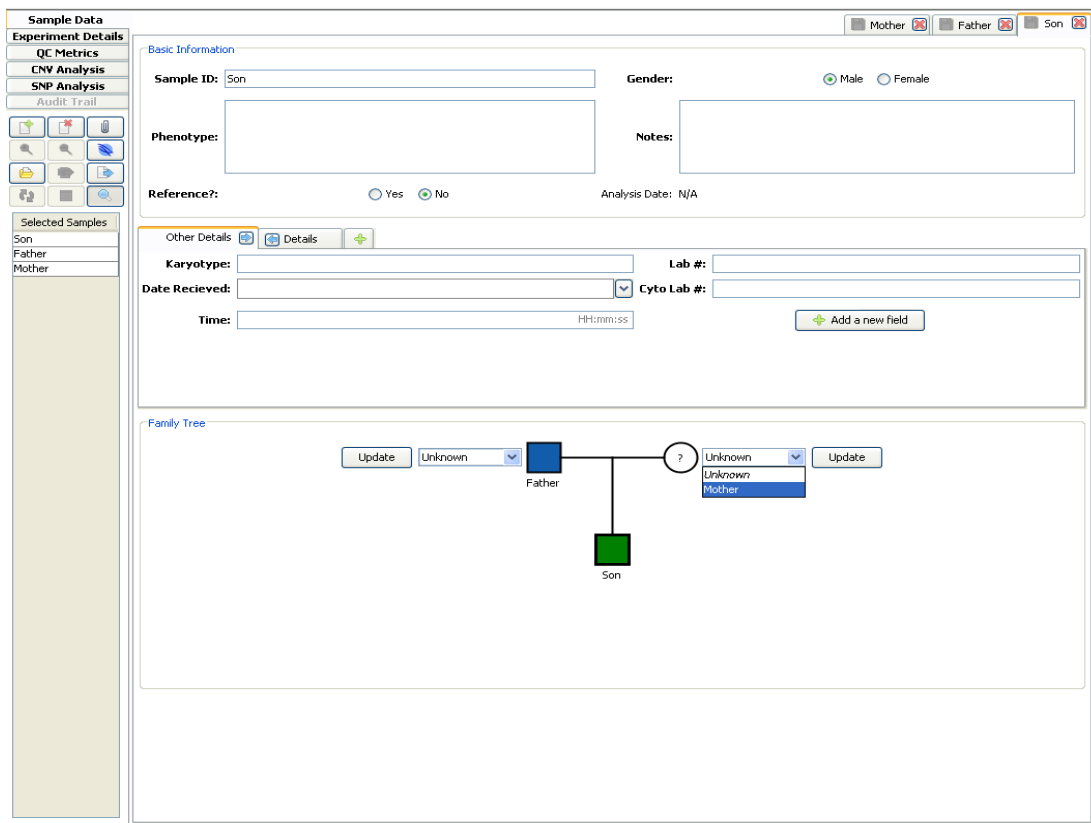
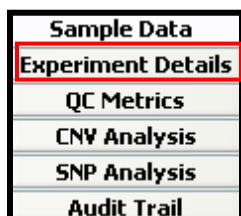
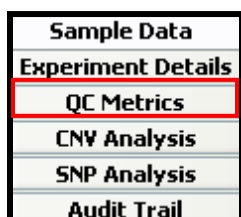


Figure 15: Linking samples within a family tree
 Three generations will be displayed within one field of view. Left-clicking on any of the linked samples will open the related sample.



Basic information regarding the experiment will appear in this window. If the feature extraction file or the CGH file is not shown, it is possible to link the correct files to the experiment by left-clicking and browsing to the file location. In this view it is also possible to select an existing QC metric set from the **'Metric Set'** drop-down menu. New fields to record custom information relating to the experiment can be created using the same method as described in the 'Sample Data' section.



To display sample specific QC metrics, select the **'QC Metrics'** tab. The metrics of the chosen samples will be displayed, including colour-coded thresholds chosen for the experiment. Selecting the check box next to a metric will plot a graph of that metric for all the samples in an experiment at the bottom of screen. It is possible to highlight more than one metric. To view the current metric settings or to create new metric sets, click on the **'Settings'** icon. The default setting is to display the QC metrics expected when the CytoSure Genomic DNA Labelling Kit is used in conjunction with CytoSure 8x60k format arrays. It is possible to display QC metrics for multiple samples (e.g. a list of samples generated after filtering or querying the database). This is described in the **'Batch Operations'** section.

The QC Metrics settings button displays the metric sets currently saved within CytoSure Interpret Software. Selecting different sets will display the thresholds within the right-hand side pane. A double-click in any of the threshold boxes will enable the user to change the value of the metric, and selecting **'Add'** will create a new metric. De-selecting the check boxes under **'Report'** will omit the chosen metric from the report. Similarly, de-selecting the check box from under the **'Display'** column will prevent the metric from being displayed within the sample information window in **'Genomic View'** (Figure 16).



Sample ID: Example

Experiment: Sample (M) / Reference (M)

Phenotype Details:

Source File : [ISCA 4x180k.txt](#)

Design Type : Unknown (029002)

Mapped Probes : 49431/59646

X-Separation : -0.7842877

DLR Spread : 0.1795

Red Signal Intensity : 1,408.46

Green Signal Intensity : 1,222.56

Red Signal-to-Noise Ratio : 136.082

Green Signal-to-Noise Ratio : 146.395

Non-Uniform Features : 0.01%

Waviness : [Calculate](#)

Figure 16: QC metric display in the 'Genomic View'

Sample Data
Experiment Details
QC Metrics
CNV Analysis
SNP Analysis
Audit Trail

Table view

The data displayed within the table can be sorted in ascending or descending order by left-clicking on the chosen column header. Dragging the column header will change the location of the column header. Right-clicking on the column header will give the option to select different CNV information for viewing.

Ideogram view

An ideogram of the selected dataset is displayed, left-click on a chromosome to zoom in.



Genomic view

Below both the table and the ideogram, individual aberrations are displayed along with the annotation tracks. Selecting aberrations from the annotation track from a different sample will open that sample allowing comparison of aberrations (Figure 17). Aberrations are coloured according to user classification.

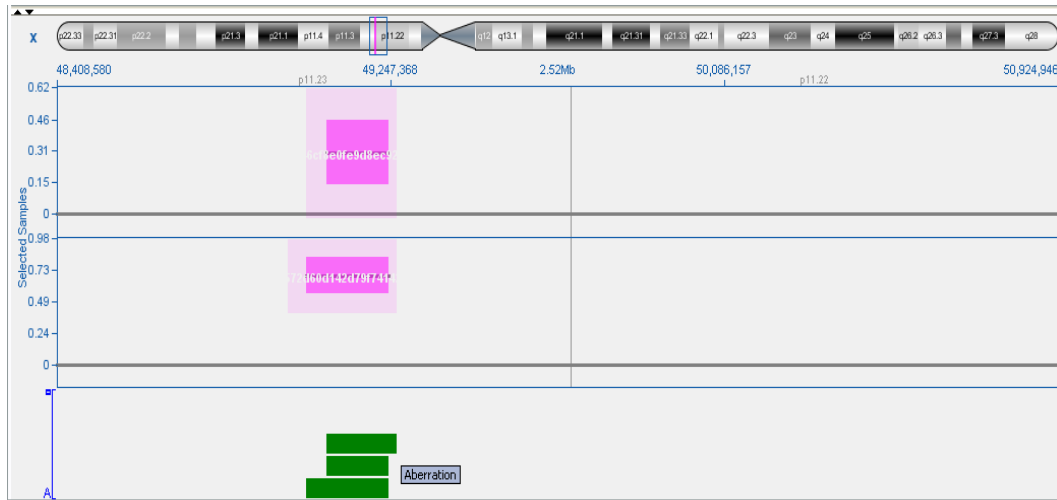


Figure 17: 'Genomic View' multiple sample aberration representation

Sample Data
Experiment Details
QC Metrics
CNV Analysis
SNP Analysis
Audit Trail

SNP information is displayed under the loss of heterozygosity (LOH) regions tab. As with CNV analysis (see section CNV Analysis), the column order can be customised and sorted. The '**Ideogram View**' displays the regions of LOH within each chromosome. The '**Chromosome Homozygosity**' view displays the number (#) of homozygous SNPs and also displays the percentage of homozygous, heterozygous and unknown SNPs for each chromosome.



Sample Data
Experiment Details
QC Metrics
CNV Analysis
SNP Analysis
Audit Trail

The software records certain user actions in order to provide an audit trail of changes made to sample, experiment or analysis information. This information is displayed in the **'Audit Trail'** tab. The tab displays a table of modifications with the following column headings:

- **Sample ID** The identifier of the sample affected by the modification
- **Modified Entity** The type of information that has been modified. Can be one of the following:
 - Sample The user has modified basic, custom or pedigree information related to the sample
 - Experiment The user has modified basic or customer experiment information
 - Aberration The user has modified editable CNV information (e.g. Final Classification)
- **User** The user who carried out the modification
- **Modification Date** The date and time of the modification
- **Modification Details** The details of the modification. The larger text window at the bottom of the panel also displays the details of the selected modification.

Users are able to navigate to affected entities (if applicable) by double-clicking on the relevant row of this table. It is also possible to restrict the list to modifications made by particular users by selecting or deselecting users from the **'Users'** section on the left-hand side of this panel (Figure 18).



Sample ID	Modified Entity	User	Modification Date	Modification Det...
80c56b10-d41d-4...	Experiment	Administrator	05 Jan 2012 14:3...	Experiment created
94810143-0831-4...	Sample	Administrator	05 Jan 2012 14:3...	Sample created
80c56b10-d41d-4...	Sample	Administrator	05 Jan 2012 14:3...	Sample created

Experiment created

Figure 18: Viewing the audit trail

Menu buttons

A simple menu has been created to help provide shortcuts to certain functions within CytoSure Interpret Software and its database (Figure 19).

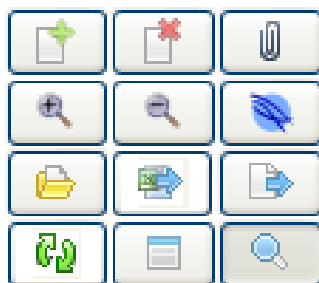


Figure 19: Menu buttons



'Add/Create' will add or create a new experiment or a sample depending on whether **'Sample Data'** or **'Experiment Details'** is selected.



'Delete' will delete an experiment or sample.



'Attach' provides the ability to bulk import information for samples and/or experiments from a text file. It is an ideal tool for uploading experimental data such as the output from a Nanodrop™. After selecting the information file it is necessary to create and name a parser, the parser will take the data from the text file and load it directly into the database. It is also necessary to identify at what line within the file the import should begin (Figure 20).



Create File Parser

Parser Name: Start from line:

Column separation character(s): Tab Space Other

File Type: One sample per LINE One sample per FILE

Automatically create samples and experiments if required?

	A	B	C	D	E	F	G
1	Sample ID	DOB	Notes	Sex	Sample label	Karyotype	Phenotype
2	Sample ID B11-109...	Experiment Date 01/01/1999	Lab Notes	Sample Gender Male	Sample Label Cy3	Karyotype "47,XXY"	Sample Phenotype Hypogonadi...
3	A11-12776	01/05/1985	Mother: A09...	f	green	"46,XX"	Autism
4	B11-00432	25/12/2001		Male	Cy5	NYK	MR
5	B10-00543		Partner: B0...	Female	Cy3	NYK	RA
6	B10-00223			Female	Cy3	"46,XX"	Devdel
7	R11-0456		Twin: R11.0...	Male	Cy3	NYK	MCA

A2

Basic Field

Custom Field

Figure 20: Importing information — creating a parser

The type of separator used needs to be identified, select either 'Tab' or 'Space' via the radio buttons. If a different separator has been used, this can be identified by selecting 'Other' and inputting the chosen separator, for example a semi-colon (;). It is also necessary to select 'One sample per line' or 'One sample per file'. The explanation of which should be used is given below:

- **One sample per file** Information can be taken from any cell within the worksheet for one sample
- **One sample per line** Will take information from a line or multiple lines within the worksheet for multiple samples (Figure 21)

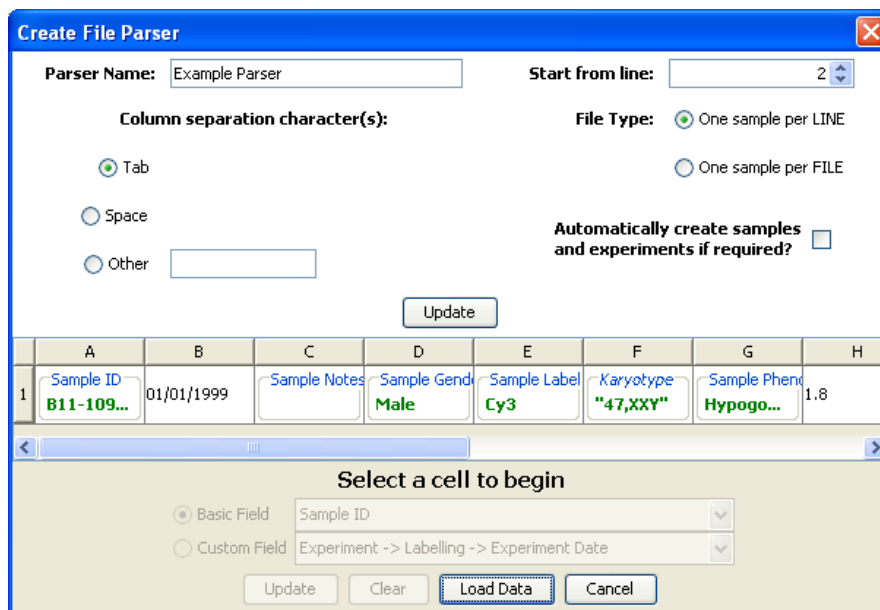


Figure 21: Importing information — one sample per line

It is possible to link uploaded data to existing data fields in the software, as well as adding additional information.

To read the Sample ID from the first case listed in the file:

- Select cell **'A1'** in the table.
- Select the **'Basic Field'** radio button.
- Select **'Sample ID'** from the drop-down box.
- Click **'Update'**.

Some existing fields require data to be in a specific format in order to be loaded. For example, **'Sample Gender'** and **'Reference Gender'** may only be linked to cells reading either **'M'**, **'F'**, **'Male'** or **'Female'** (case-insensitive). If the data is not in the correct format, the text in the cell will be red. Text within a cell may be altered, for example, **'M'** is converted to the word **'Male'** when linked to **'Sample Gender'** or **'Reference Gender'**.

To remove a link between a cell and a field, highlight the cell and select **'Clear'**.

A parser can be created and saved(e.g. for loading Nanodrop data) and used for subsequent imports.



'Zoom in/Zoom out' will zoom in or out whilst in **'CNV Analysis'** tab of the selected aberration.



'Open CytoSure Interpret file' will open a browser window to navigate to the location of a saved Interpret file (.cgh).



'Open Feature Extracted file' will open a browser window to navigate to the location of a saved Feature Extracted file (.txt).



'Regenerate Report' will populate a report for the selected sample using the current report settings within the protocol.



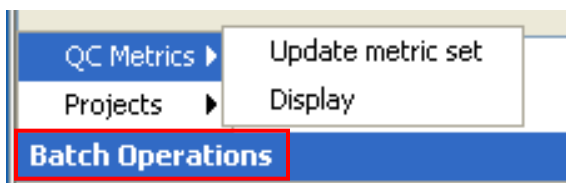
'Refresh Annotation Overlap Calculations' will recalculate the information related to the overlap of annotations from selected tracks for each aberration in the selected sample.



'Settings' displays the QC metric settings options.



'Search' toggles the display of the **'Quick Search'** and **'Advanced Search'** fields.



The **'Batch Operations'** button (located beneath the menu buttons described above) provides a number of tools to summarise or modify the information related to all samples listed in the table on the left-hand side of the **'Database Management'** tab (i.e. all samples that have passed any filters that have been applied by the user). For example, under **'QC Metrics'**, the **'Display'** option will create graphs of the QC metrics for all samples in the list, enabling the user to compare subsets of samples for differences in QC metrics, perhaps by filtering based on consumable batch numbers. Alternatively, it also provides a way of summarising any QC trends over a particular time period. An example of a function to modify, rather than summarise, the data related to the list of samples is the **'Update Metric Set'** function in the same menu. The function provides the means to change the metric threshold set applied to the default experiment of the listed samples.

The other 'Bulk Operations' provided by the software are as follows:

- **CNV Analysis**
 - **Update inheritance** – See *Family Tree* section.
- **Projects**
 - **Create new project** – Creates a new project.
 - **Add to project** - Adds the default experiment of all samples in the sample list to the selected project.
 - **Remove from project** – Removes the default experiment of all samples in the sample list from the selected project.



CytoSure™ Interpret Software update — version 8.0 January 2011

The following new features are available in CytoSure Interpret Software version 3.4.2:

Additional sample tracking spike-in-related functionality

77

Mosaicism detection

79

Bulk import of sample information

80

Miscellaneous

87

Bug fixes

88



Additional sample tracking spike-in-related functionality

In addition to the **Sample Identification** interface (see Update Guide v6.0 [page 29]), users are now able to specify and verify the presence of, CytoSure ISCA Sample Tracking Spike-ins added to their sample within the CytoSure Interpret Software. CytoSure ISCA Sample Tracking Spike-ins ensure that any potential sample mix-up during array processing can be easily identified, safeguarding against mistakes and maximising data confidence. To facilitate this, the **Sample Details** dialog box now contains a **Spike-ins** section in the **Other** tab. When the user selects one or more CytoSure ISCA Sample Tracking Spike-ins from the drop-down box displayed after clicking on **Click here to add a Spike-in**, the signal intensities of the selected spike-in are compared in order to determine whether it has indeed been added to the sample.

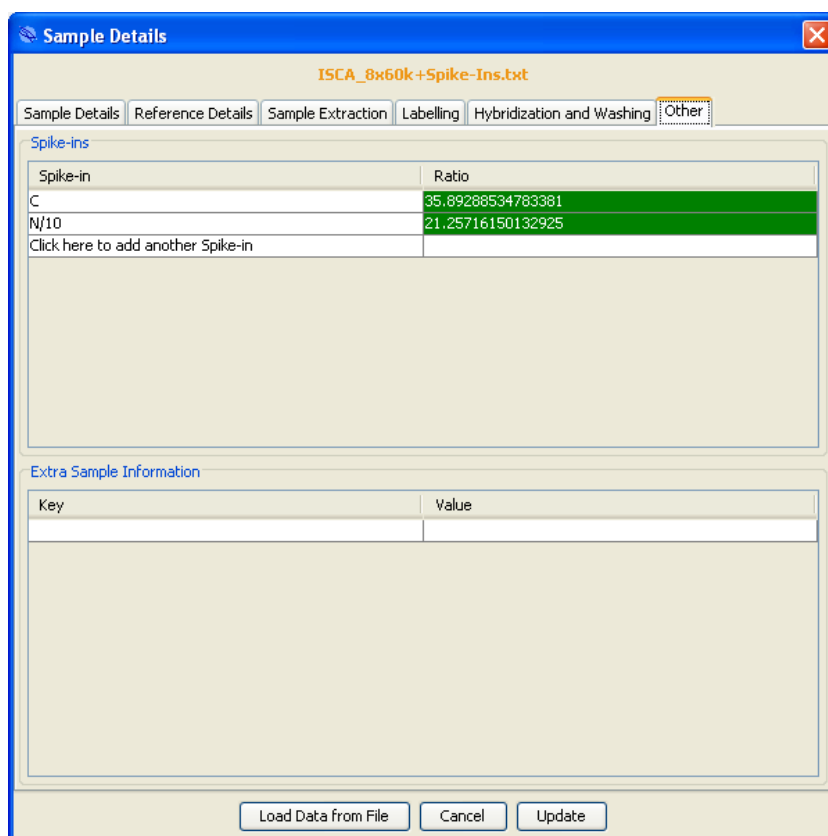


Figure: Spike-ins C and N selected in the **Spike-ins** section of the **Other** tab in the **Sample Details** dialog box.

A **green** background in the **Ratio** column, which displays the ratio of the sample signal to that of the reference for the selected spike-in probe, indicates that the selected spike-in is present, thereby



confirming sample identity. A **red** background indicates the absence of the selected spike-in and identifies potential sample mix-up. The ratio threshold for identifying a spike-in as present can be modified in the **Sample Identification** dialog box, available by clicking **Tools -> Sample Identification**. This dialog box can also be used to identify which spike-ins have been identified as present by the software.

If one or more spike-ins have been selected, an additional QC metric, named **Spike-ins**, will be displayed in the **Sample Information** section on the right-hand side of the **Genomic View** tab. If the software has determined that **all** selected spike-ins are present, **Match** will be displayed, otherwise **No Match** will be displayed. This information will also be included in any reports generated for the sample, if applicable.

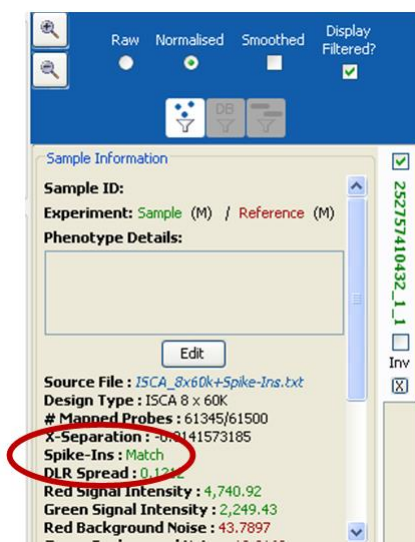


Figure: The **Spike-ins** QC metric.

Some spike-ins may bind to other CNV probes as well as the spike-in probe to which it is targeted, and as such it may be beneficial to exclude these probes when attempting to detect copy number changes. In order to make this easier for the user, CytoSure Interpret Software automatically applies a filter to relevant probes when spike-ins are selected by the user. For example, in the figure below, the CNV probe has cross-hybridised with the spike-in (CytoSure ISCA Sample Tracking Spike-in C) and as the user has correctly selected that this spike-in is present, both the spike-in probe and this CNV probe have been filtered out automatically. To also apply this filtration during aberration detection (CBS processing), the user needs to ensure that the **Apply probe filters during Aberration Detection?** checkbox in the **Probe Filtering** tab of the **Options** dialog box is selected.

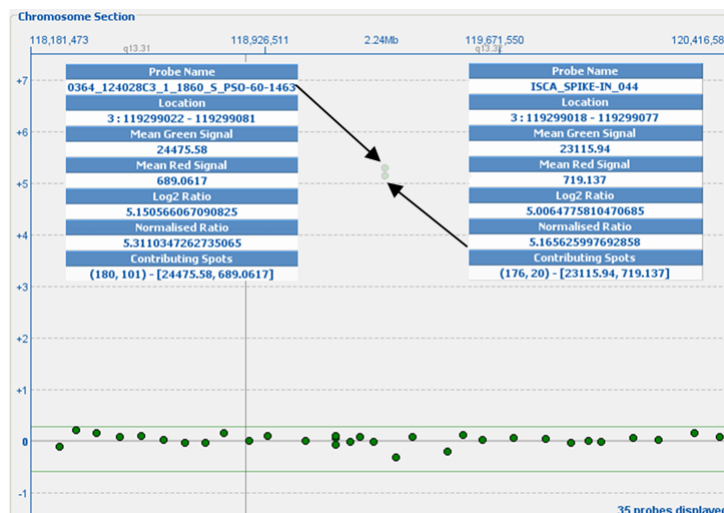


Figure: The **Genomic View** displaying the automatic filtration of a spike-in probe and cross-hybridizing CNV probe.

Mosaicism detection

In order to detect large aberrations with small log ratio shifts, potentially caused by mosaicism, extra options have been added to the **Aberration Calling Options** section of the **Aberration Detection** tab of the **Options** dialog box, namely:

- **Min. Size (Mb) (Mosaicism)** — The minimum size of significant mosaic segments.
- **Threshold Factor (Gains – Mosaicism)** — The factor by which any threshold statistic should be multiplied to generate the threshold value for mosaic gains.
- **Threshold Factor (Losses – Mosaicism)** — The factor by which any threshold statistic should be multiplied to generate the threshold value for mosaic losses.

The default value of 999,999 for **Min. Size (Mb) (Mosaicism)** effectively disables this feature, so users must reduce this to an appropriate value in order to call this type of aberration.

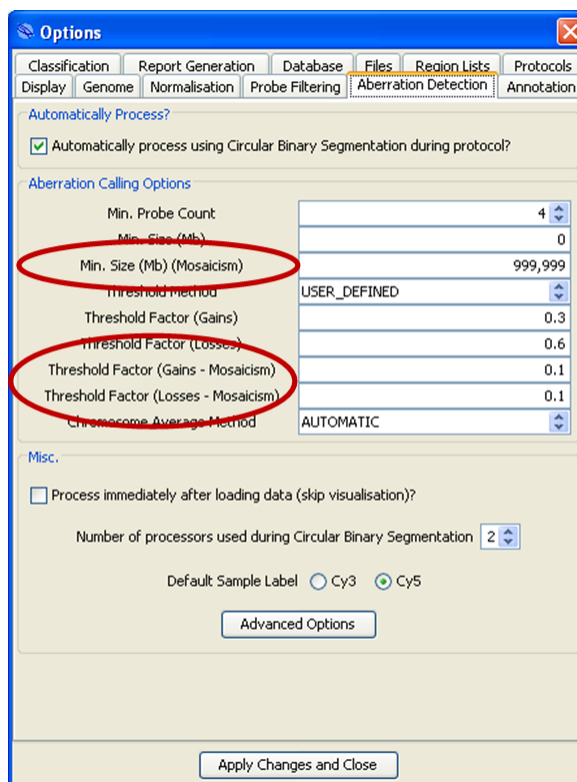


Figure: New mosaicism aberration calling settings in the **Aberration Detection** tab of the **Options** dialog box.

The mosaicism-related options are applied after the standard options and any extra aberrations called using these options are added to the list of aberrations with their **Region List** field annotated with the word “Mosaic”.

Bulk import of sample information

The ability to load information for an individual sample from a text file has been added to the **Sample Details** dialog box in the form of the **Load Data from File** button. Once clicked, the user is prompted to select the file containing their sample data (which should be a delimited, plain text file such as a Microsoft Excel spreadsheet saved as a text file), which is subsequently displayed in a table within the software.

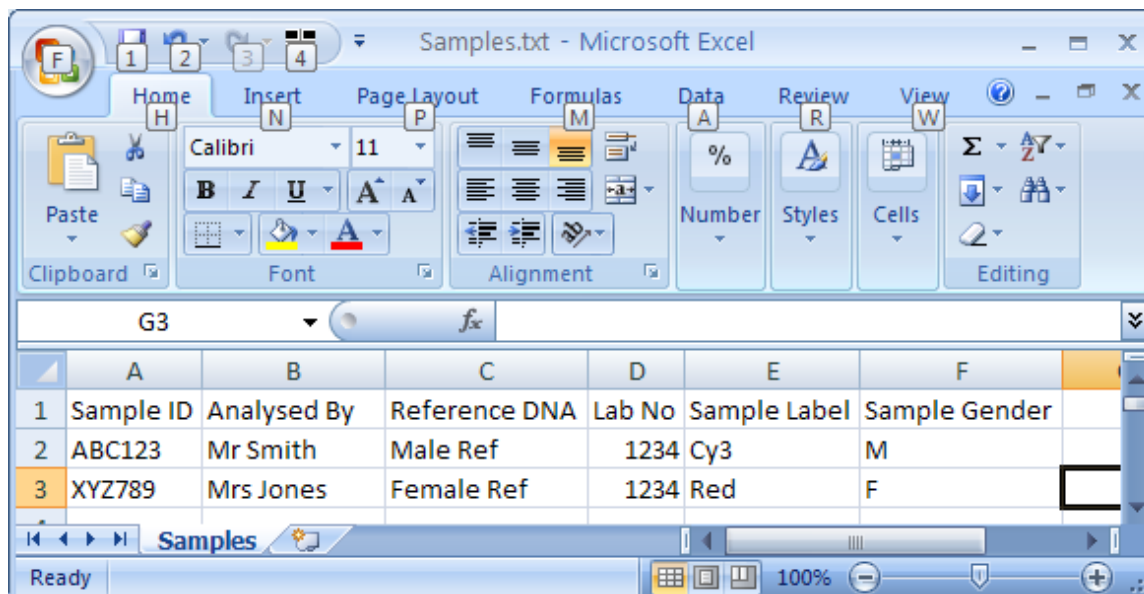


Figure: An example text file containing sample data displayed in Microsoft Excel.

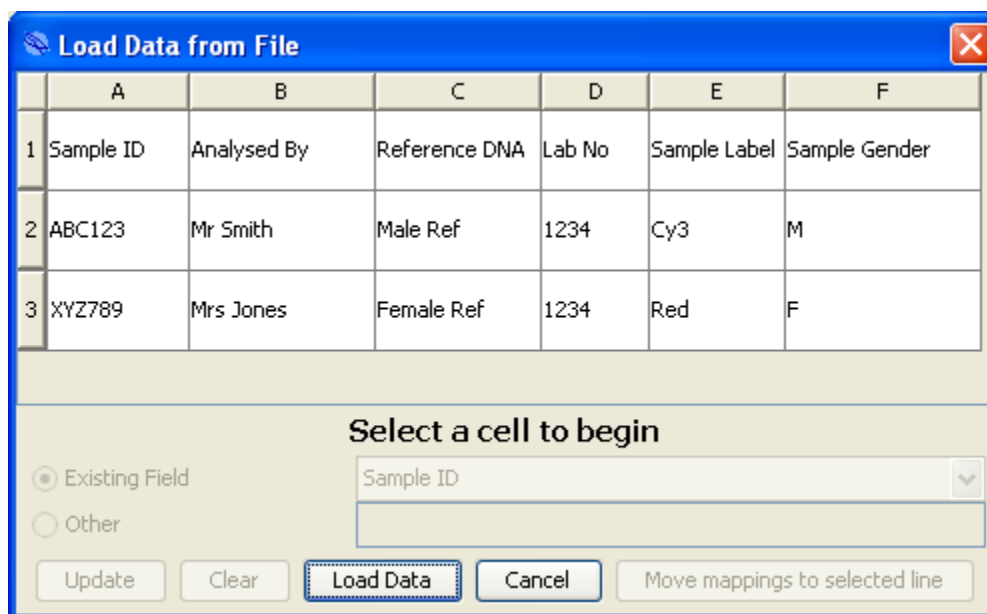


Figure: The same text file after being loaded into the software.

The **Load Data from File** dialog box provides controls enabling the user to link cells of the table with existing data fields in the software, and also to add extra information not already stored. For example, in order to read the Sample ID from the first case listed in the file, the user would:

- Select cell A2 in the table.



- Select the **Existing Field** radio button.
- Select **Sample ID** from the drop-down box.
- Click **Update**.

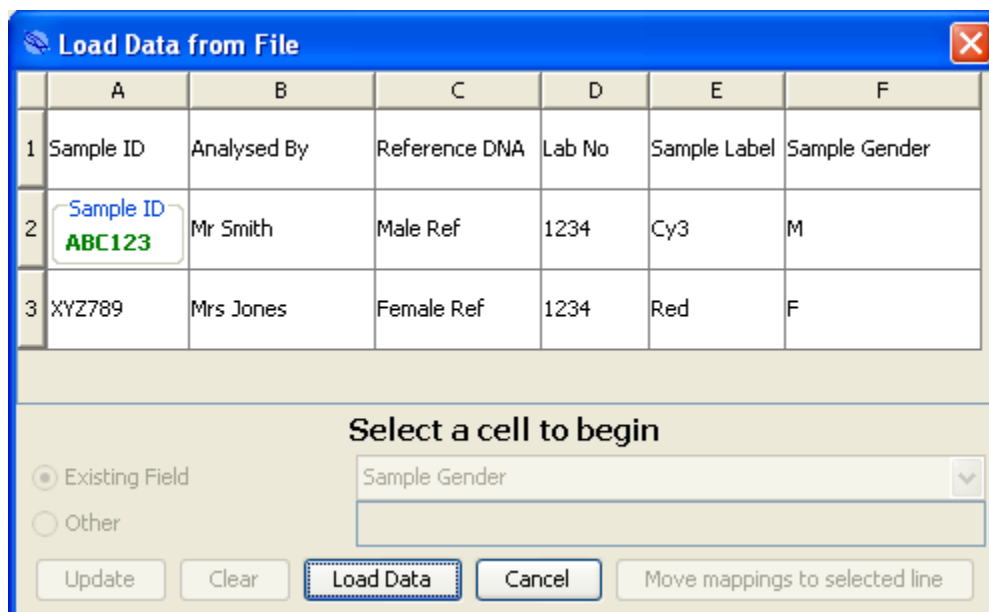


Figure: The result of linking cell A2 with the Sample ID field.

Some existing fields require data to be in a specific format in order to be loaded. For example, Sample Gender and Reference Gender may only be linked to cells reading either “M”, “F”, “Male” or “Female” (case-insensitively). If the data is not in the correct format, the cell will display its text in red as show below.

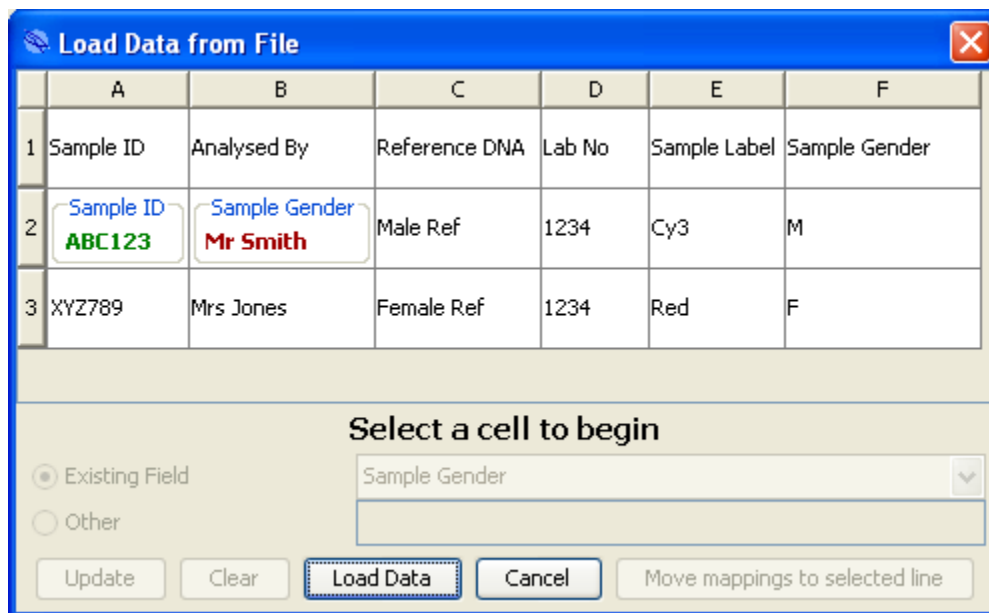


Figure: Incorrect data format for the **Sample Gender** data field.

To remove a link between a cell and a field, select the cell and click the **Clear** button. If the software has been able to correctly read the input data, it may alter the text within the cell in question to the form in which it is displayed within the software. For example, “M” is converted to the word “Male” when linked to Sample Gender or Reference Gender.

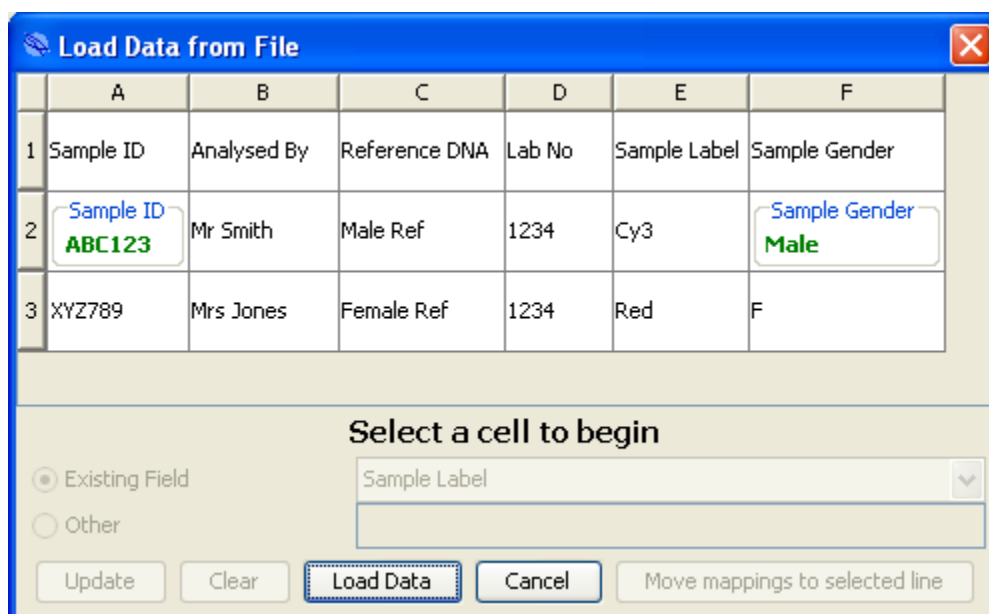


Figure: Modification of cell text following successful data recognition.



It is also possible to associate new information with the sample in the case that no existing field is appropriate for a column contained in the data file. For example, there is no suitable standard field that can be used to represent “Lab No” in the software. To add this field, the user should select the appropriate cell (D2, for example), select the **Other** radio button, enter “Lab No” in the input field and click **Update**. New field names are italicized above the data text, which is left black

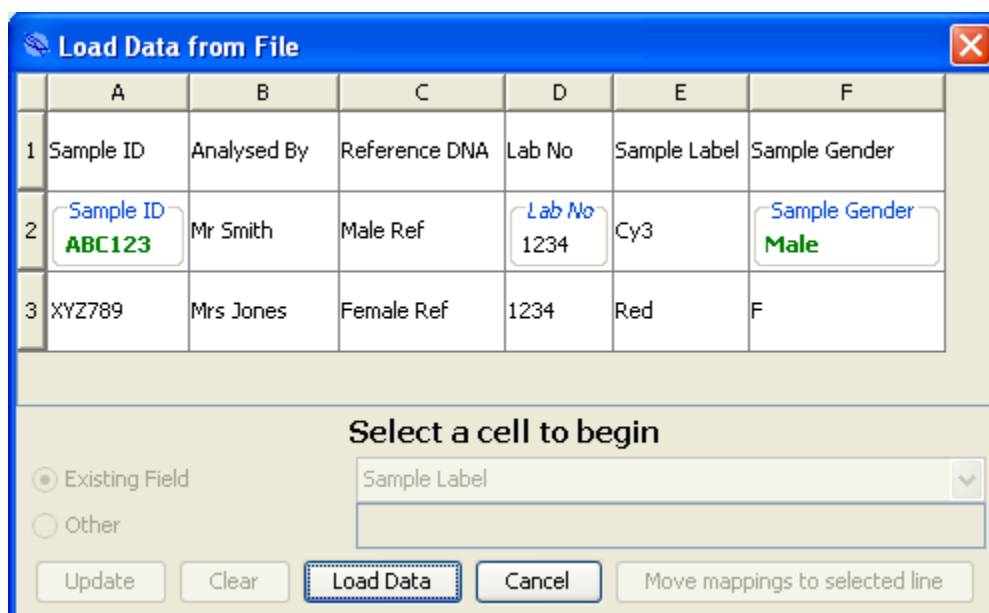


Figure: Creating new table fields.

Once all required cells have been linked with existing or new fields, clicking **Load Data** results in the information being associated with the sample and the **Sample Details** dialog box is updated accordingly.



Load Data from File

	A	B	C	D	E	F
1	Sample ID	Analysed By	Reference DNA	Lab No	Sample Label	Sample Gender
2	Sample ID ABC123	Lab User Mr Smith	Reference DNA Male Ref	Lab No 1234	Sample Label Cy3	Sample Gender Male
3	XYZ789	Mrs Jones	Female Ref	1234	Red	F

Select a cell to begin

Existing Field Lab User

Other

Update Clear Load Data Cancel Move mappings to selected line

Figure: All required cells have been linked with new or existing fields.

Sample Details

Sample Details | Reference Details | Sample Extraction | Labelling | Hybridization and Washing | Other

Sample ID: ABC123

Sample Label: Cy3 Cy5

Sample Gender: Male Female

Ethnicity: _____

Assigned Technologist: _____

Case Status: In Progress In Review Completed Archived

Sample Details: _____

Phenotype Details: _____

Load Data from File Cancel Continue

Figure: The **Sample Details** dialog box is updated with the new information.



New field information is displayed in the **Extra Sample Information** section of the **Other** tab. This extra sample information can also be added manually by selecting a cell in the bottom line of the table and entering a new key-value pair.

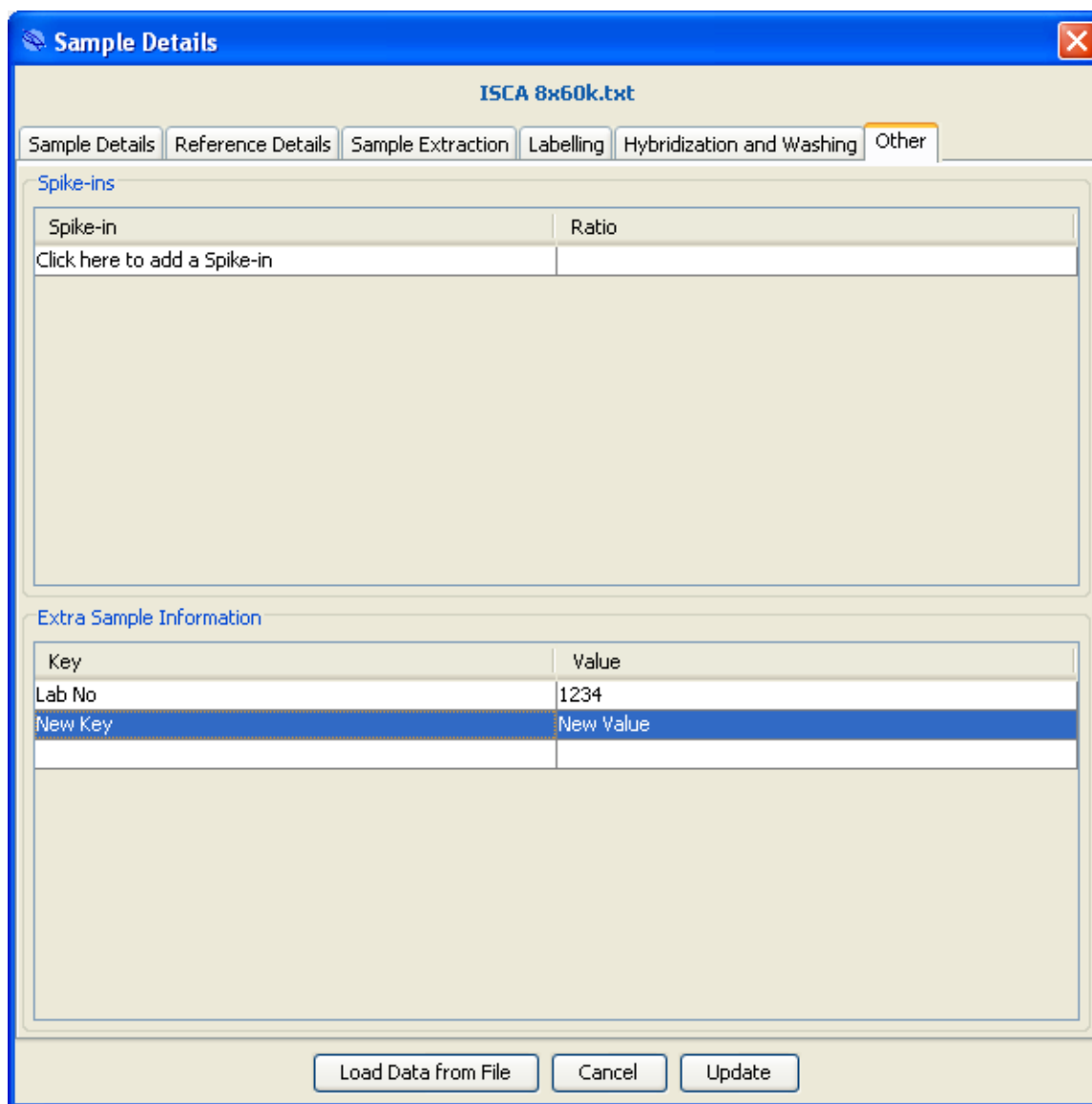


Figure: The **Extra Sample Information** table in the **Options** tab of the **Sample Details** dialog box where new key-value pairs are stored.

The link between cells in the text file and the new and existing sample fields has now been stored by the software, enabling the user to bypass the linking step when importing subsequent sample data. For example, in the case of the type of file shown in this guide, where the sample data file



contains information relating to many different samples in subsequent rows, the user simply needs to move the current mappings to the line containing the correct sample information using the **Move mappings to selected line** button.

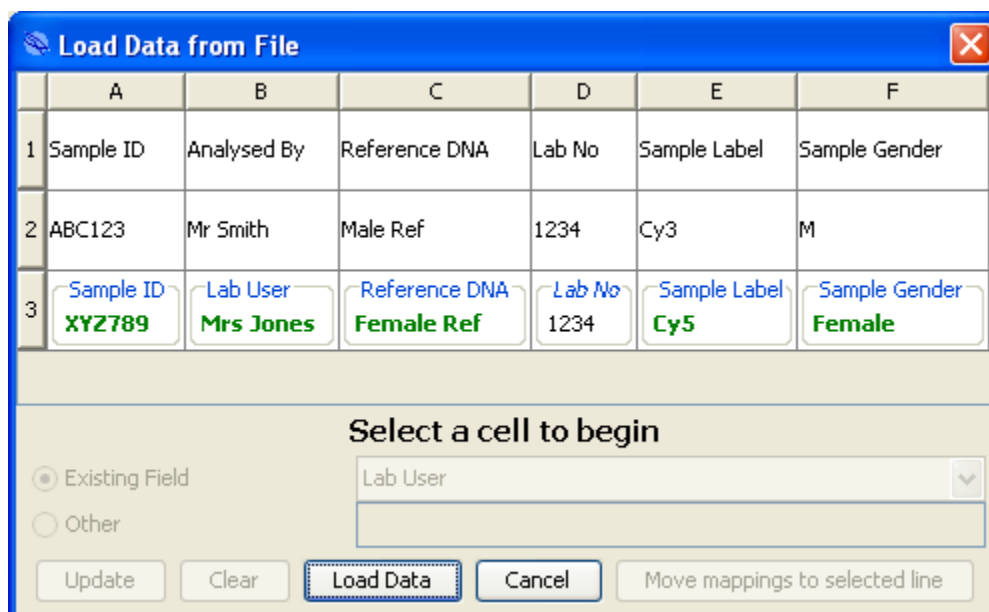


Figure: Stored links between cells and sample fields after being moved to row 3 by clicking **Move mappings to selected line**.

If the user has a single text file for each sample all in the sample format, simply selecting the new file will result in the correct cells being linked to the correct data fields within the software.

Miscellaneous

- New loss of heterozygosity (LOH) detection algorithm (BETA version).
- Option to overwrite database submissions if the sample ID matches one or more already present in the database.
- Report names default to the naming scheme defined in the protocol options.
- Number of annotations displayed per track limited to 5000 to optimise genomic view display.
- dbGaP submission file folders grouped by month.
- “Lab Code” input field added to dbGaP submission options.
- “Important Genes” renamed to “OMIM Genes”.
- Added columns for total number of genes in “All Genes” and “OMIM Genes” fields.



- Size of selected region displayed in **Genomic View** and genomic location axis labels reformatted.
- Visualisation of sample details now available via right-click in the **Submissions** table of **Database Management** tab.
- **Go To Position...** function now takes both start and stop positions.
- Feedback given to user following problems loading or modifying a database.
- p and q arm homozygosity percentage add to whole chromosome UPD ideogram view.
- All components modified to take sample dye selection into account when displaying log ratios.
- DECIPHER and DGV annotation updated.

Bug fixes

- Reproducibility metric calculation fixed for non-Agilent feature extracted files.
- Reproducibility metric thresholds aligned with updated Agilent thresholds.
- Aberration calls recalculated following reference gender selection.



CytoSure Interpret Software update — version 7.0 September 2010

The following new features are available in CytoSure Interpret Software version 3.3.2 onwards:

UPD analysis interface

90

- SNP data representation
- Analysis options
- Chromosome overview
- Reporting
- Database functionality

Trio analysis

98

Extra aberration annotation

100



UPD analysis interface

A new analysis interface, with accompanying reporting and database functionality, has been added to enable analysis of SNP data on CytoSure arrays. The following explains the use of these new features.

SNP data representation

Genotyping data contained on CytoSure arrays can be analysed using the new UPD analysis interface, located in the **Ideogram View** in the **Aberrations** tab (or by clicking **Tools -> SNP/UPD Data**).

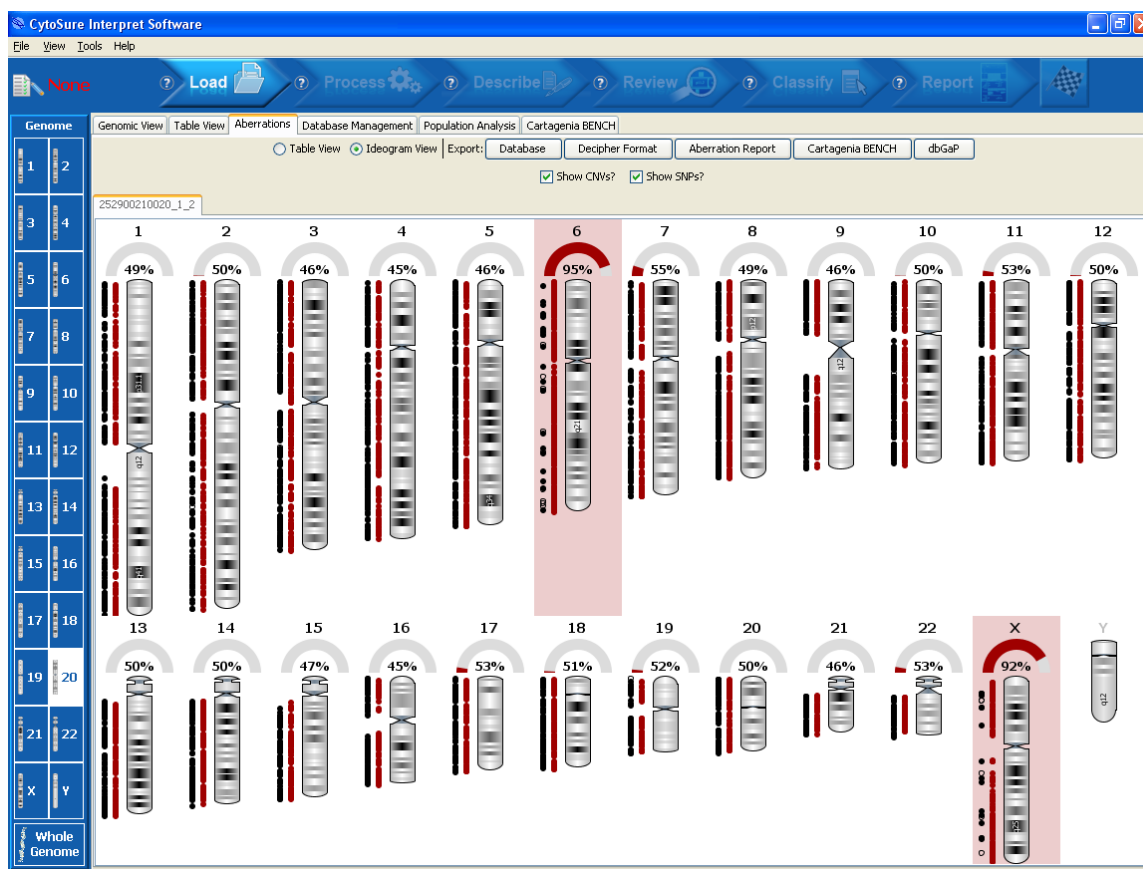


Figure: Ideogram view with SNP data.

In addition to displaying the locations of gains and losses in the data, each chromosome ideogram will also display SNP locations and colour-code them according to genotype, with heterozygotes coloured black and homozygotes coloured red. There are two views of the distribution of



homozygous and heterozygous SNPs available, and these can be toggled by clicking on the area covered by the SNP data:

1. Each SNP is represented by a circle.

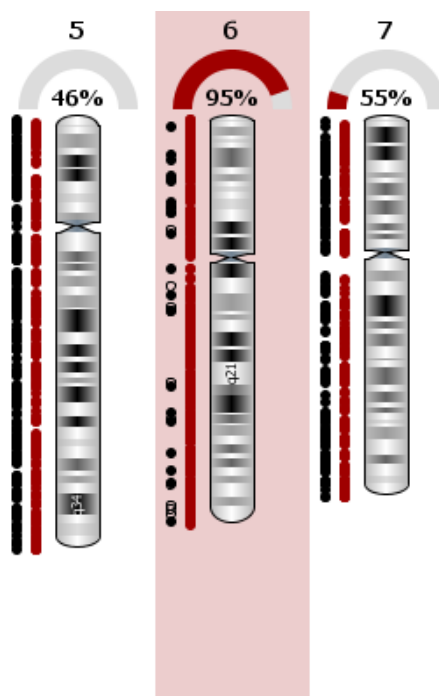


Figure: Circle SNP representation.

2. Each contiguous stretch of homozygosity/heterozygosity is represented by a rectangle covering the region between the first and last SNP in the stretch. Stretches containing more than 10 probes are filled in.

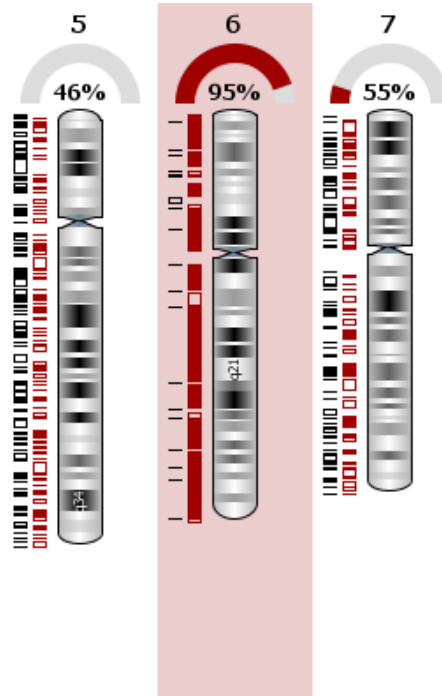


Figure: Rectangle SNP representation.

The dials above the ideograms display the percentage of homozygous SNPs for the chromosome.

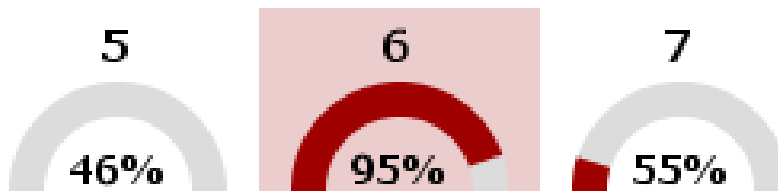
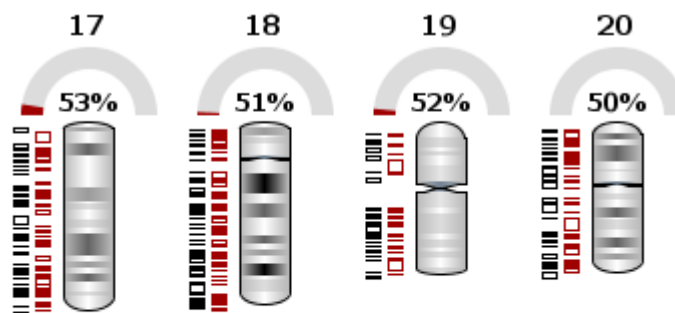


Figure: % Homozygosity dials.

Ideograms of chromosomes with a homozygosity level above a user-definable threshold will be displayed with a red background.

Analysis options

It is possible to modify this threshold and other analysis settings by clicking on the **Options...** button which is displayed at the bottom of the screen when the mouse hovers over the ideogram view.



Options...

Figure: Location of the **Options...** button.

Clicking on the **Options...** button will display the **UPD Options** dialog box. The dialog box contains two tabs: **Genotype Calling** and **SNP Probe Filters**. The **Genotype Calling** tab contains the following:

- **Homozygosity Threshold** slider – Ideograms of chromosomes whose percentage homozygosity exceeds this threshold will be given a red background.
- **Limits for Genotype Assignment** – This graph displays the distribution of the \log_2 signal intensity ratio for all pairs of SNPs on the array. Two thresholds can be modified in order to determine the genotype of each SNP:
 - **Lower Bound** – SNPs whose absolute \log_2 signal intensity ratio falls **below** this value will be called as **heterozygous**.
 - **Upper Bound** – SNPs whose absolute \log_2 signal intensity ratio falls **above** this value will be called as **homozygous**.
 - SNPs whose absolute \log_2 signal intensity ratio falls **between** these two thresholds **will not be included in the analysis** and will be visualised as an empty black circle in the ideogram view.

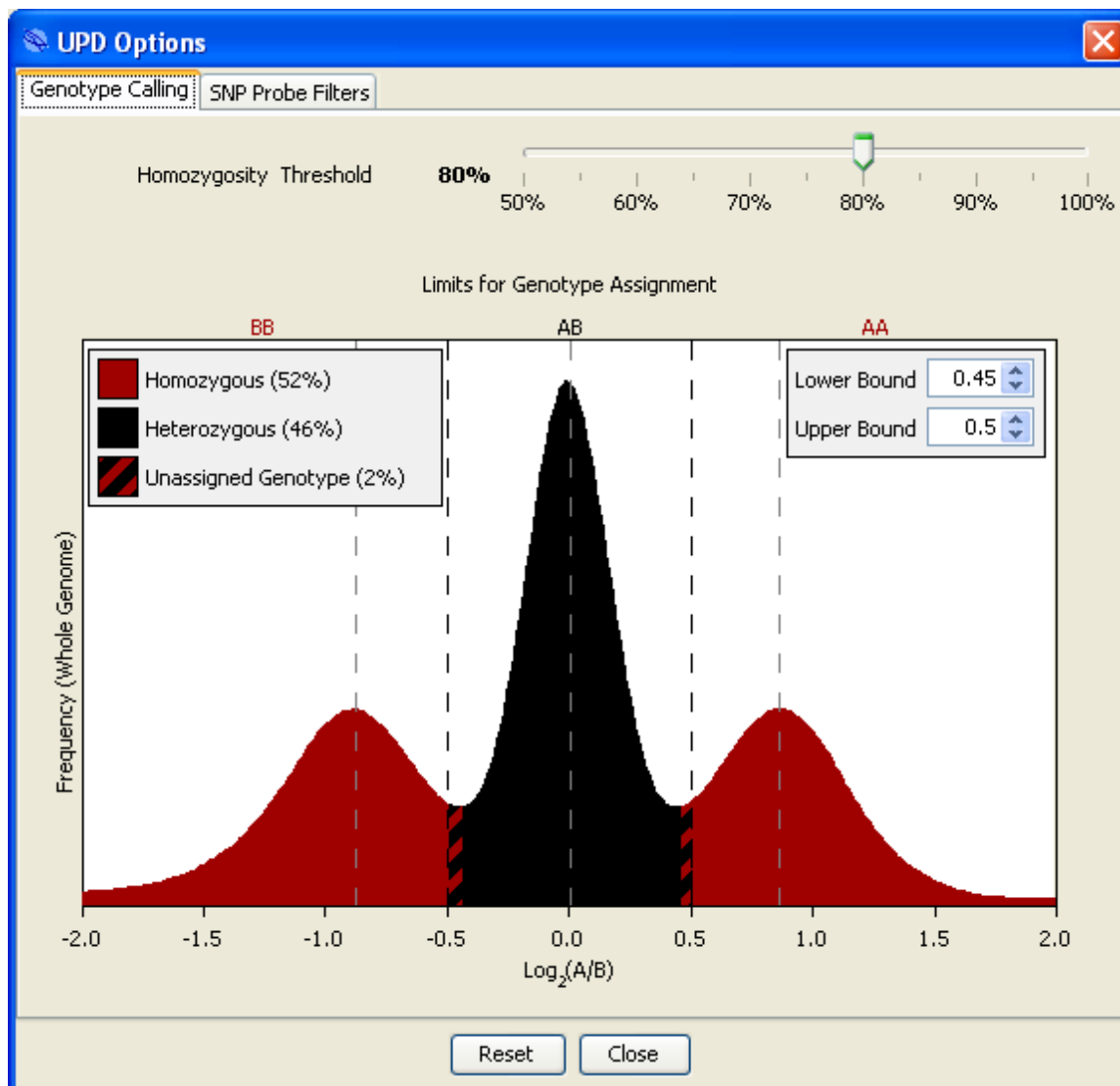


Figure: **UPD Options** dialog box with **Genotype Calling** tab displayed.

The **SNP Probe Filters** tab contains the following:

- **Apply SNP Probe Filters** check box – Determines whether the proceeding probe filters will be applied if selected.
- **Exclude Non-Uniform Outliers** check box – If selected, SNPs with 1 or more non-uniform probes will be excluded from the analysis.



- **Exclude SNPs within aberrant regions?** check box – If selected, SNPs within regions covered by gains or losses already identified in the CNV data will be excluded from the analysis.
- **Exclude SNPs Based on...** graphs – SNPs whose **Signal Intensity** or **Reproducibility** values fall within the grey area are excluded from the analysis.

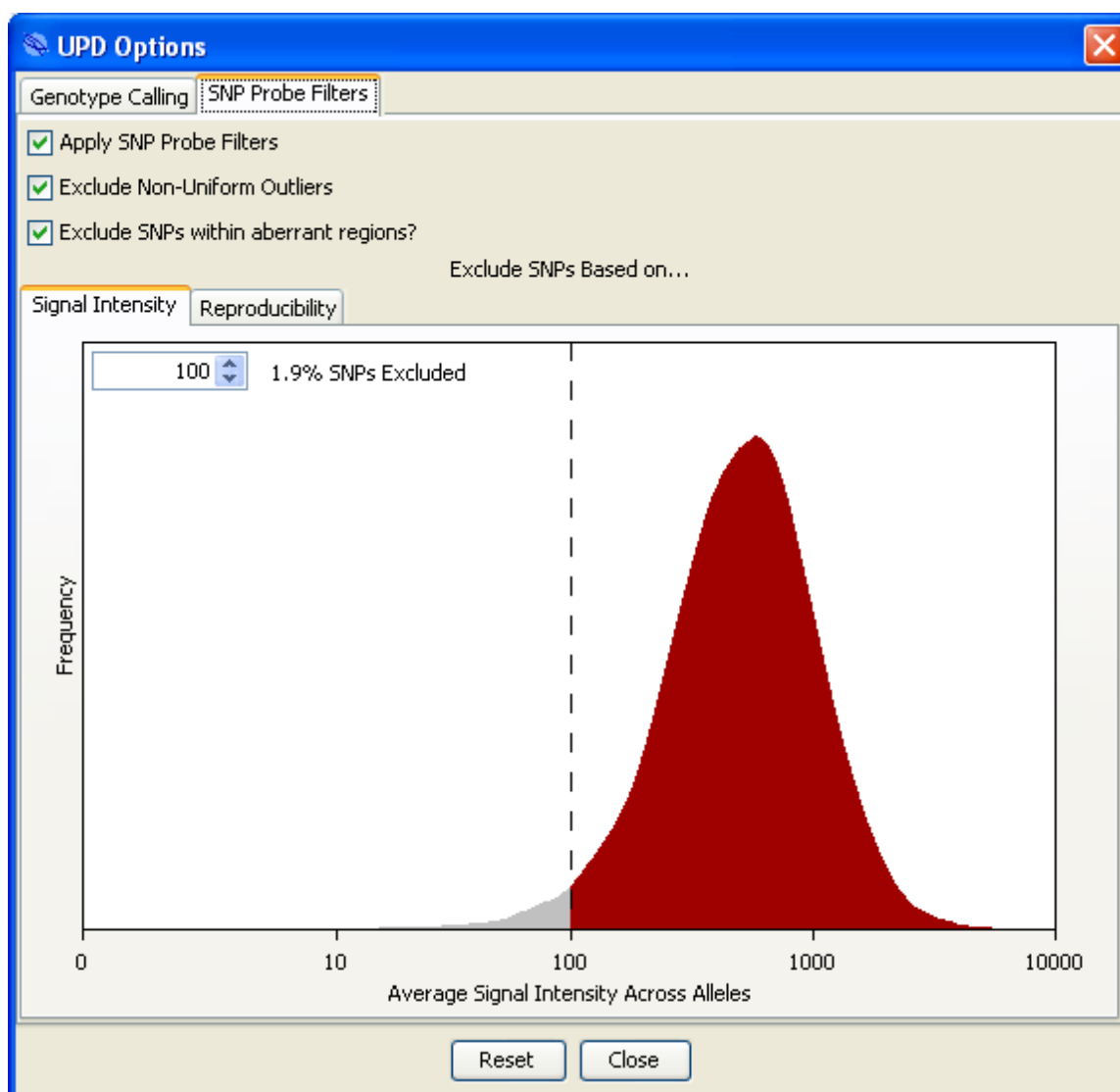


Figure: **UPD Options** dialog box with **SNP Probe Filters** tab displayed.

Chromosome overview



More detailed information relating to an individual chromosome is available by clicking on its ideogram. This information includes:

- Larger, horizontal ideogram.
- Log₂ signal intensity ratio distribution plot for SNPs in the chromosome compared to the whole genome.
- Table of largest (most SNPs) regions of LOH in the chromosome.
- Graph of % Homozygosity against # SNPs in each chromosome, with the selected chromosome highlighted in red.

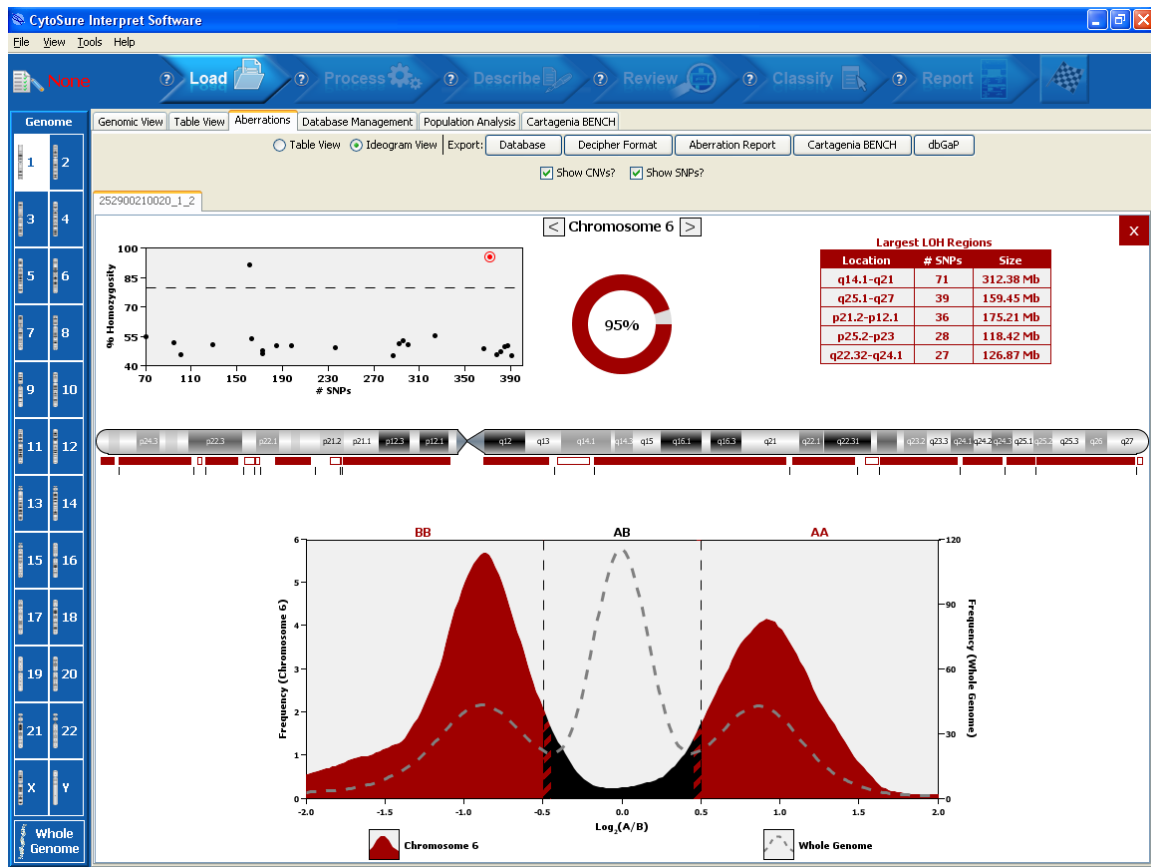


Figure: Chromosome overview including signal ratio distribution graph and table of largest regions of loss of heterozygosity.

Reporting

HTML and PDF reports generated from data containing genotyping information will include a table of percentage homozygosity for each chromosome.



report.pdf - Adobe Reader

File Edit View Document Tools Window Help

Chromosome	# SNPs	Homozygous	Heterozygous	Unknown Genotype	% Homozygosity
6	381	354	18	9	95%
X	168	147	14	7	91%
7	327	178	146	3	55%
22	71	38	32	1	54%
17	166	87	76	3	53%
11	300	155	141	4	52%
19	97	49	46	2	52%
10	295	148	144	3	51%
20	130	65	64	1	50%
12	303	151	149	3	50%
2	392	194	193	5	50%
14	201	99	99	3	50%
18	187	92	93	2	50%
1	388	190	195	3	49%
13	239	116	120	3	49%
8	372	178	188	6	49%
15	175	82	91	2	47%
5	384	178	203	3	47%
16	178	79	94	5	46%
21	103	46	55	2	46%
4	387	171	207	9	45%
3	395	175	216	4	45%
9	294	128	159	7	45%

Figure: % Homozygosity table included in a PDF report.

Database functionality

If sample information is submitted to the database, the table of homozygosity will be available by clicking on the icon in the **UPD** column of the **Submissions** table of the **Database Management** tab. If the percentage homozygosity of any of the autosomes of a particular sample exceeds the current threshold, the icon will be displayed in red; otherwise, it will be displayed in blue.



Chromosome	# Homozygous	# Heterozygous	Total	% Homozygosity
6	354	18	372	95%
X	147	14	161	91%
7	178	146	324	55%
22	38	32	70	54%
17	87	76	163	53%
11	155	141	296	52%
19	49	46	95	52%
10	148	144	292	51%
20	65	64	129	50%
12	151	149	300	50%
2	194	193	387	50%
14	99	99	198	50%
18	92	93	185	50%
1	190	195	385	49%
13	116	120	236	49%
8	178	188	366	49%
15	82	91	173	47%
5	178	203	381	47%
16	79	94	173	46%
21	46	55	101	46%
4	171	207	378	45%
3	175	216	391	45%
9	128	159	287	45%

Figure: % Homozygosity table dialog box in **Database Management** tab.

Trio analysis

If parental data is available, it may now be loaded into the software and linked to the sample, providing inheritance information for detected aberrations. To link the sample to parental data, click **File -> Properties**, select the relevant tab and choose the appropriate datasets in the **Mother** and **Father** drop-down boxes.

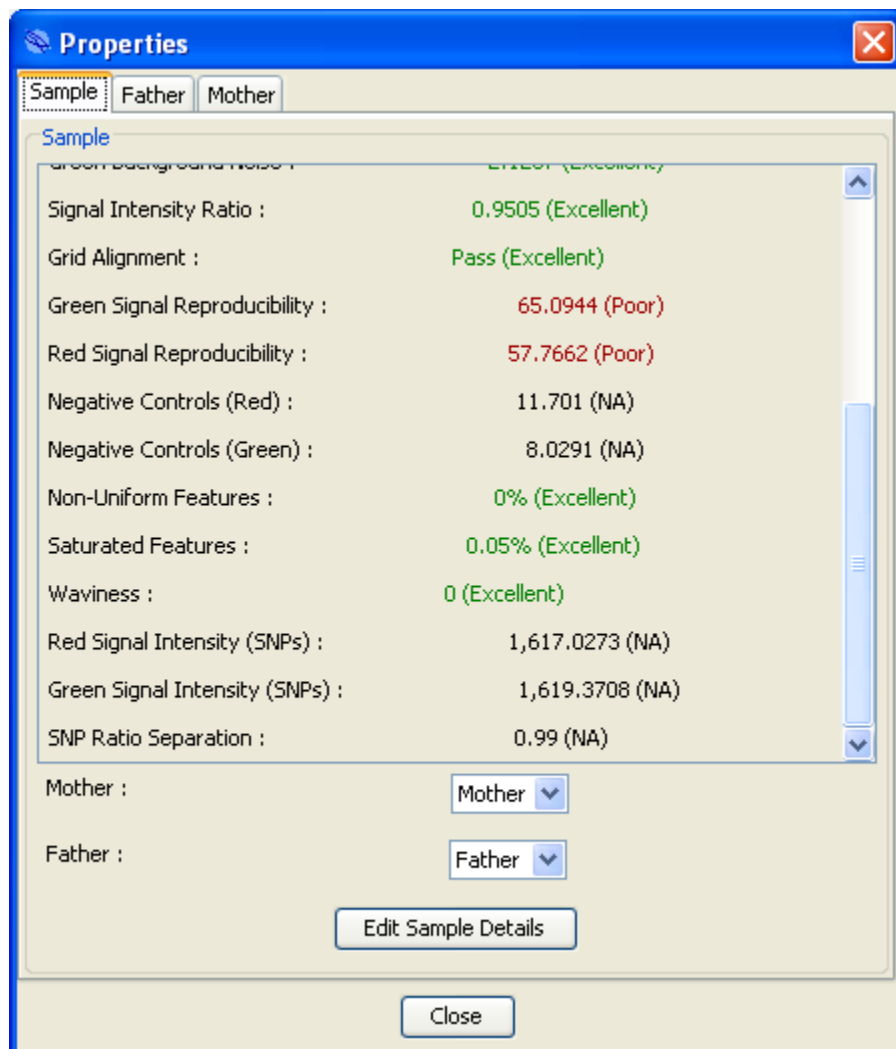


Figure: **Properties** dialog box containing new **Mother** and **Father** sample drop-down boxes.

If aberrations have been detected in the datasets, the software will attempt to attribute gains and losses in the sample to the mother and/or father samples on the following basis:

- If there are no maternal or paternal aberrations overlapping the aberration region, **Inheritance** will be set to “De Novo” and **Inheritance %** will be set to 100%.
- If only one parent has an aberration overlapping the aberration, **Inheritance** will be set to either “Maternal” or “Paternal” as appropriate, and **Inheritance %** will be set to the percentage of the aberration region which is covered by the overlapping parent aberrations.



- If both parents have aberrations overlapping the aberration region, **Inheritance** will be set to “Biparental” and **Inheritance %** will be set to the percentage of the aberration region which is covered by either a maternal or paternal aberration.

Extra aberration annotation

Four additional fields for aberration-specific data have been added in order to provide more comprehensive details about aberration location and interpretation. These fields are:

- **Start (Max)** – The molecular position of the last normal probe before the aberration region.
- **Stop (Max)** – The molecular position of the first normal probe after the aberration region.
- **Max Size (Mb)** – The size of the maximal span of the aberration region (the distance between the last normal probe before the aberration and the first normal probe after the aberration).
- **Classification (Final)** – An additional classification field has been added for users whose workflow involves the recording of an initial classification and a final classification. If a final classification is selected, the aberration will be colour-coded according to the final classification; otherwise, it will be colour-coded according to the initial classification.



CytoSure Interpret Software update — version 6.0 June 2010

The following new features are available in CytoSure Interpret Software version 3.2.4 onwards:

Ability to save status of samples during workflow processing	102
Population Analysis gene list reporting	102
Sample Identification interface	102
dbGaP submission file generation (ISCA)	102
“View region in DGV” link	104
Ability to filter and normalise using a probe list	106
Waviness Metric	106

CytoSure Interpret - Software update guide



	110
Miscellaneous	
	110
Bug fixes	110



Ability to save status of samples during workflow processing

It is now possible to save samples loaded into the workflow mode before analysis has been completed. Clicking on the chequered flag icon will now also result in the user being prompted to either save or discard the changes made to the loaded samples. If the user decides to save the changes, it will be possible to subsequently load the generated .cgh files into the workflow mode provided the same protocol is selected. Please note that .cgh files generated by clicking **File -> Save** or **File -> Save As...** will not load into workflow mode.

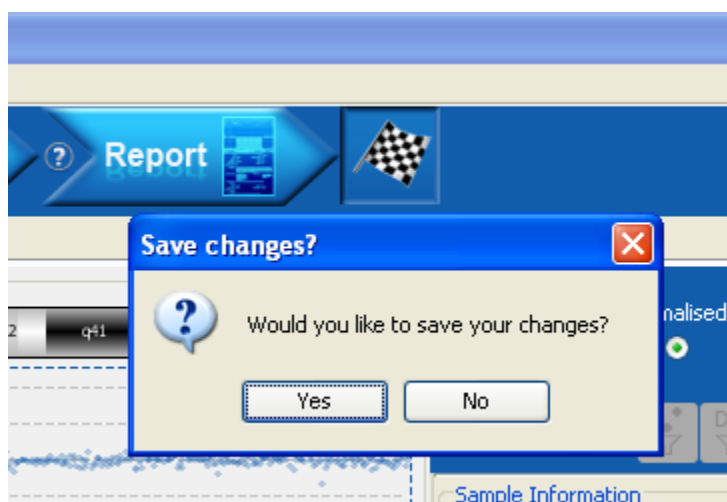


Figure: Prompt to save changes after chequered flag icon is clicked.

Population Analysis gene list reporting

Generating reports in the **Population Analysis** tab will now result in a list of all genes within each region of interest being included in an extra column in the report.

Sample Identification interface

Some CytoSure array designs now include spike-in probes for sample tracking and, if tracking PCR spikes are included in the labelling reaction, these can now be visualised in the new **Sample Identification** interface, available from the **Tools** menu.



Sample Identification	
Signal Ratio Threshold:	5 <input type="button" value="Update"/> <input checked="" type="checkbox"/> Pool Ratios?
Chr	252757510002_1_4
1	56.7
2	43.9
3	46.9
4	43.8
5	48.8
6	27.4
8	46.8
9	63.3
10	46
11	55.5
14	67.7
15	78.3
16	34.7
18	45.9
19	31.6
20	65.9
21	33.8

Figure: **Sample Identification** interface

The interface displays, by default, the average signal ratio of the spike-in probes in each chromosome, with cells colour-coded as follows:

- **Green** if **all** spike-ins within the chromosome have a signal ratio above the specified **Signal Ratio Threshold**.
- **Orange** if **at least one, but not all**, of the spike-ins within the chromosome have a signal ratio above the specified **Signal Ratio Threshold**.
- **Red** if **none** of the spike-ins within the chromosome have a signal ratio above the specified **Signal Ratio Threshold**.

The **Signal Ratio Threshold** can be modified by updating the associated input field and clicking the **Update** button. If the **Pool Ratios?** checkbox is deselected, the ratios for all spike-in probes will be displayed.



It is also possible to include the pooled summary of the spike-in probes in HTML or PDF reports by selecting the **Report Sample Identification information if available?** checkbox in the **Report Generation** tab of the **Options** dialog box.

Sample Extraction	
• Specimen Type: Blood	
Labelling	
• Labelling Protocol: Agilent	
• Columns Used: OGT	
Sample Identification	
Chr	1 2 3 4 5 6 8 9 10 11 14 15 16 18 19 20 21
Ratio	56.7 43.9 46.9 43.8 48.8 27.4 46.8 63.3 46 55.5 67.7 78.3 34.7 45.9 31.6 65.9 33.8

Figure: Sample Identification information displayed in a report.

dbGaP submission file generation (ISCA)

Users are now able to automatically generate the files required for submission of sample data to dbGaP by clicking on the **dbGaP** button in the **Aberrations** tab. When clicked, the user will be presented with the **dbGaP Submission** dialog box, allowing the completion of input fields to specify the data required to be included in the *ISCA_submission_template_V8* Excel spreadsheet.



Figure: **dbGaP Submission** dialog box.

When the **Export** button is clicked, the software will generate:

- A de-identified feature extracted file for each sample loaded.
- A normalised data file for each sample loaded.
- A completed *ISCA_submission_template_V8* spreadsheet **for each different array design** loaded.

All files will be placed in the **dbGaP** folder of the software installation directory within a folder named according to the date and time at which the files were generated (e.g. 04-Jun-2010_103523786).

The values entered into the **Methods** section of the **dbGaP Submission** dialog box will be remembered by the software for future submissions. When sample data is saved (by either clicking **File -> Save/Save As...** or the **Report** workflow button), values entered into the **Samples** section will be saved with the sample data and be available in the **dbGaP Submission** dialog box when



the sample is subsequently loaded in to the software. Please note that the **Methods** section information forms part of the protocol.

“View region in DGV” link

In addition to the UCSC and Ensembl genome browsers, users can also view the selected region in the **Genomic View** in the Database of Genomic Variants browser by clicking **View -> Selected region in DGV**.

Ability to filter and normalise using a probe list

In addition to using **Region Lists**, users can now carry out normalisation and filtering based on a list of probe names by using the **Probe Lists** section of the **Region Lists** tab of the **Options** dialog box. To create a new probe list, complete the **Name** input field, select a file containing the list of probes (each on a separate line) by clicking on the **Browse** button and click **Add**.

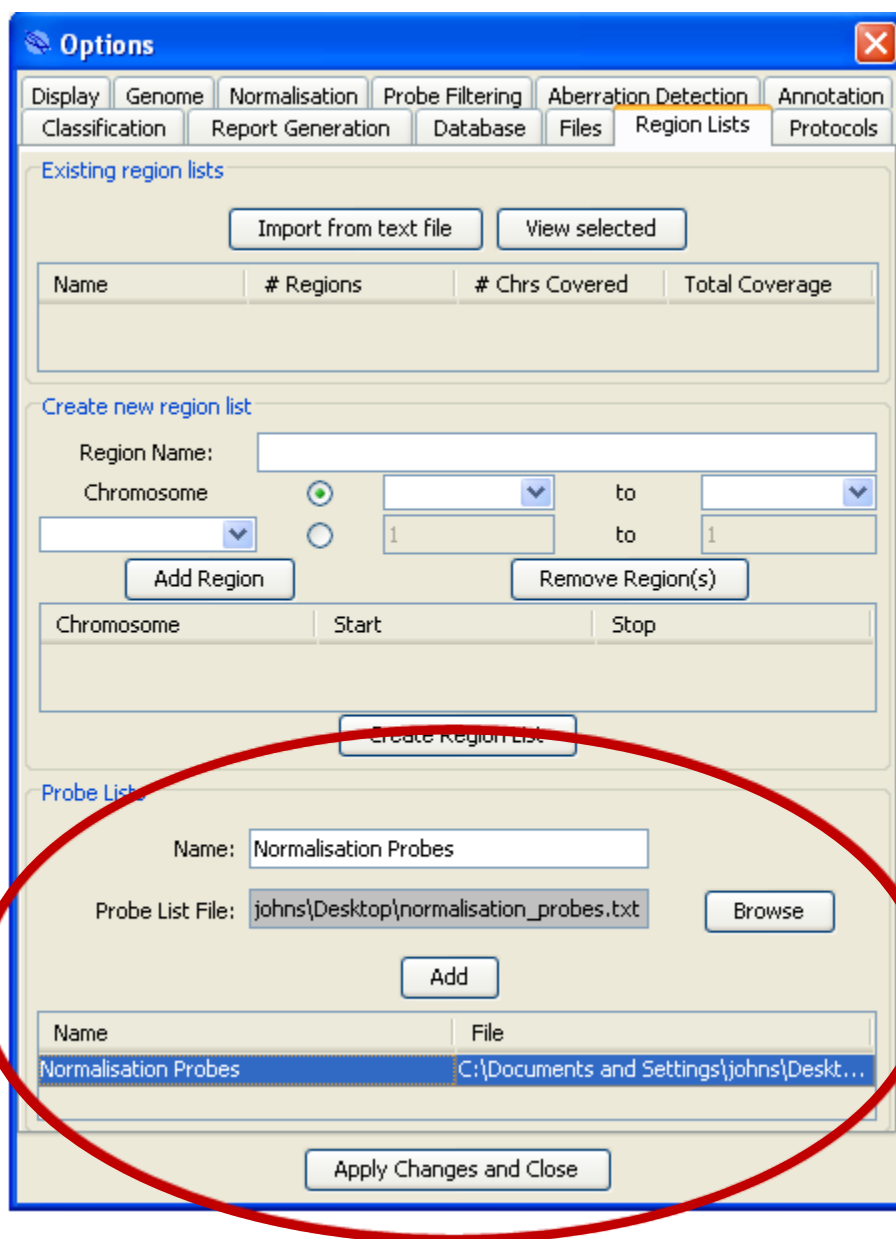


Figure: **Probe List** creation interface in the **Region Lists** tab of the **Options** dialog box.

Generated probe lists can be applied to both normalisation and probe filtering. To filter out all probes contained in a specific probe list, select the probe list from the **Exclusion Probe List** drop-down box in the **Probe Filtering** tab of the **Options** dialog box. To filter out all probes not contained in a specific probe list, select the probe list from the **Inclusion Probe List** drop-down box.

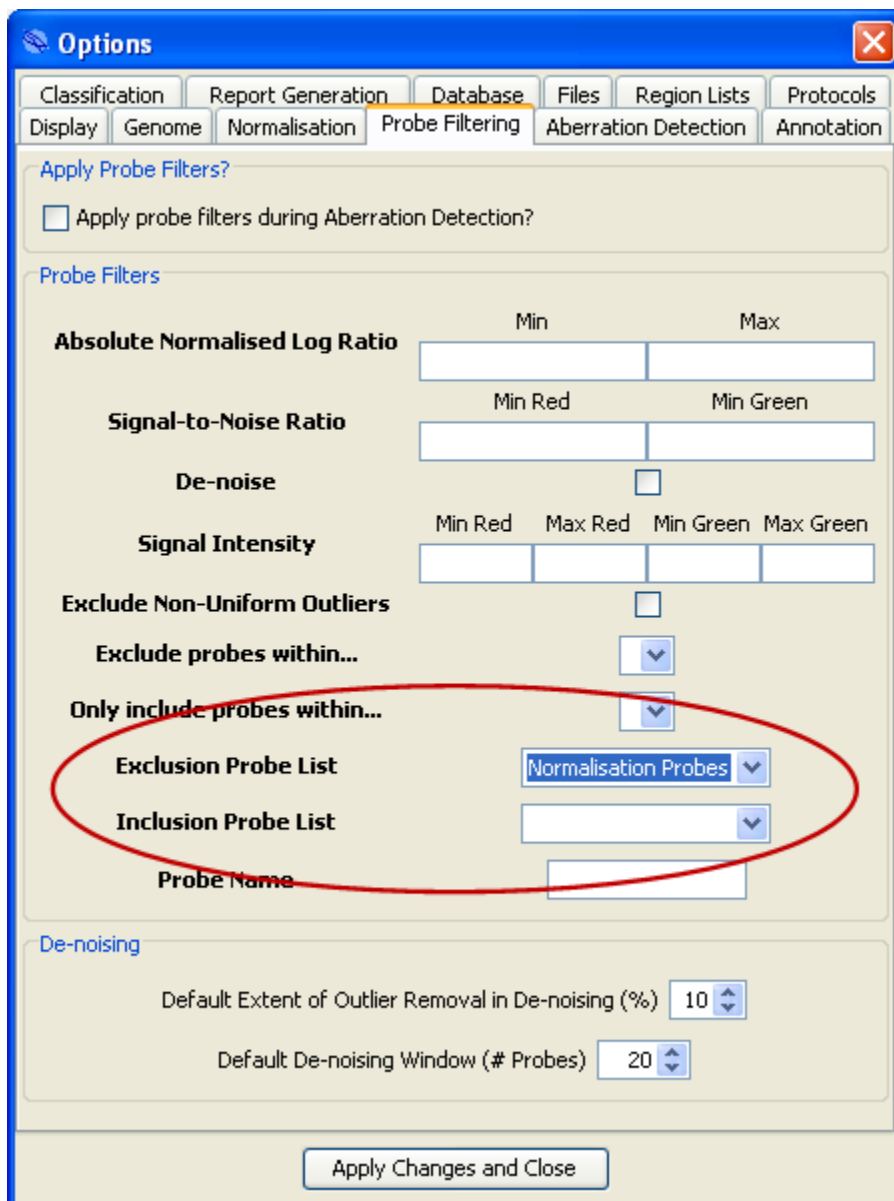
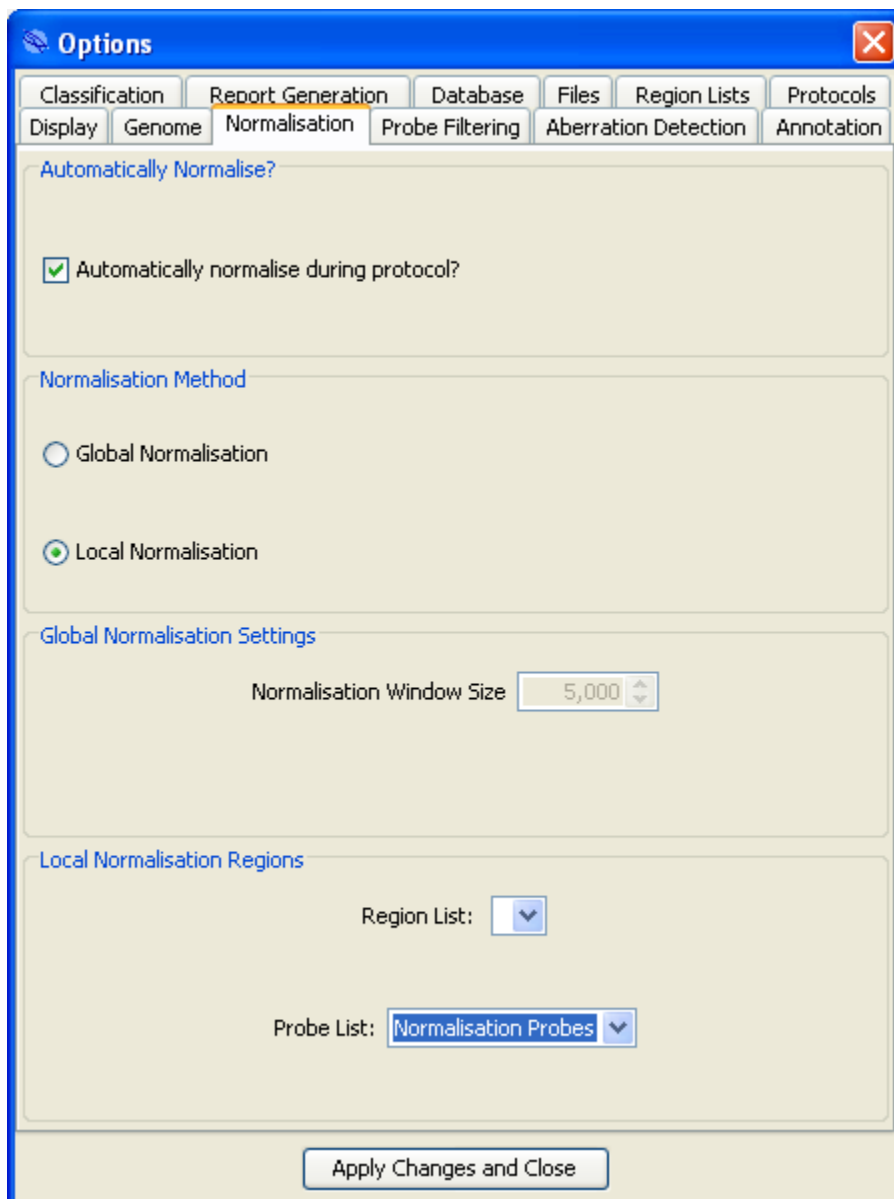


Figure: Probe List drop-down boxes in the **Probe Filtering** tab.

To normalise data using only the probes contained in a specific probe list, select the **Local Normalisation** radio button in the **Normalisation** tab of the **Options** dialog box and select probe list from the **Probe List** drop-down box. Samples subsequently loaded into the software from feature extracted files will be normalised using only the probes whose names are included in the selected probe list.



Users may either locally normalise using a probe list or a region list. Selection of a probe list will result in de-selection of a region list (and vice versa) if necessary. Please note: Modification of the file referenced by a probe list will be reflected in subsequent normalisation and probe filtering. Users wishing to ensure consistency in these applications should ensure that probe list files are “read-only” to avoid modification.

Waviness Metric



Some poor quality samples present an undulating log ratio profile along the chromosome making it more difficult to identify genuine aberrations in the data. However, the presence of these “waves” is not usually highlighted by standard quality control metrics, so a new metric has been added in order to indicate their presence and magnitude. The metric is calculated as follows:

- For each autosome:
 - Calculate the moving average of the normalised log ratios using a 2Mb window size.
 - For each band in the chromosome, calculate the standard deviation of the moving average for each probe in the band.
 - Calculate the mean of all of the standard deviations.
- Return the mean of this value across each autosome.

A value of greater than 0.025 is considered poor. To calculate this metric, scroll to the bottom of the **Sample Information** panel in the **Genomic View** and click **Calculate** next to “Waviness : ”.

Miscellaneous

- The Copy Number Variation (DGV) and Decipher annotation tracks have been updated.
- Additional text files containing cancer genes, imprinted genes and OMIM Morbid Map genes and confirmation probes that may be used to create custom annotation tracks are now included with the software in the extra_annotation folder of the installation directory.
- QC Metrics available in Agilent Feature Extracted files are now read by the software, rather than being recalculated.
- “Default Protocol” has now been renamed to “OGT Protocol”.
- BENCH server URLs are automatically converted from HTTP to HTTPS if necessary.
- The splash screen can be pushed to the background if required.

Bug fixes

- Memory leak and workflow database submission from Genomic View error fixed.
- BENCH loop design submission error fixed.
- File generation error caused by the presence of ampersands fixed.
- Classification and aberration annotation is now retained if files are reloaded into workflow mode.

CytoSure Interpret Software update — version 5.0 February 2010

CytoSure Interpret - Software update guide



The following new features are available in CytoSure Interpret Software version 3.1.7 onwards:

Ability to automatically display sample/reference	
Probes coloured according to thresholds	112
Ability to remove aberrations from the summary table based on classification	112
Protein coding genes highlighted in the report	114
Ability to filter the Gene annotation track	115
Segmentation and probe filtration consistency warning	115
Filtration status indicators	116
Delete button added to Database Management	117
	118

CytoSure Interpret - Software update guide



Ability to filter probes by name	118
File naming convention applied during manual saving	118
Bug fixes	118



Ability to automatically display sample/reference

Previously, the scatter plot in the **Genomic View** would always default to displaying the \log_2 ratio of the Cy5 signal divided by the Cy3 signal, unless the individual dataset was inverted by the user. Users can now default to displaying the \log_2 ratio of the sample signal divided by the reference signal by selecting the “Always display sample/reference by default (rather than Cy5/Cy3)” check box in the **Display** tab of the **Options** dialog box.

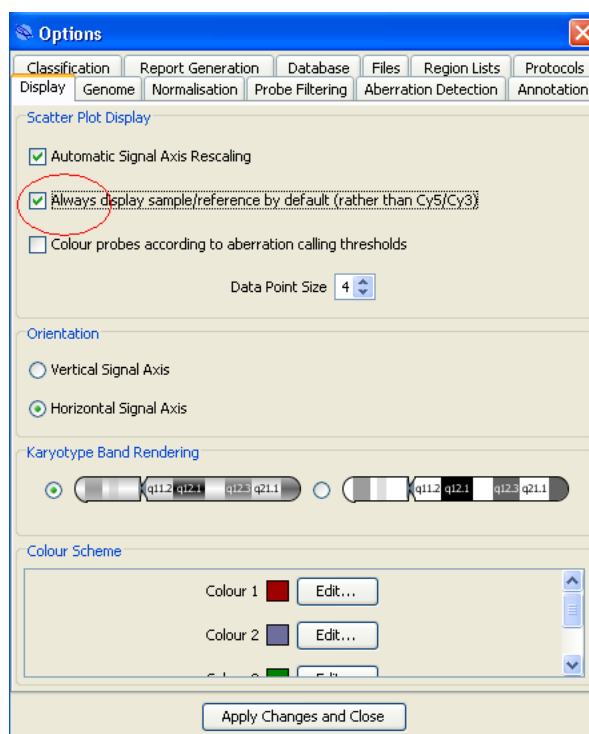


Figure: New option in the **Display** tab to enable automatic inversion of the scatter plot.

Probes coloured according to thresholds

It is now possible to colour the probes of an individual dataset according to the aberration calling thresholds set in the **Aberration Detection** tab of the **Options** dialog box. When the “Colour probes according to aberration calling thresholds” checkbox is selected in the **Display** tab of the **Options** dialog box, all probes above the gain threshold will be coloured green, all probes below the loss threshold will be coloured red, and all probes in between will be coloured black. For this view to be shown, the software must be able to calculate the threshold (it may be necessary to carry out



Circular Binary Segmentation first, depending on the threshold settings) and there must be only one dataset visible.

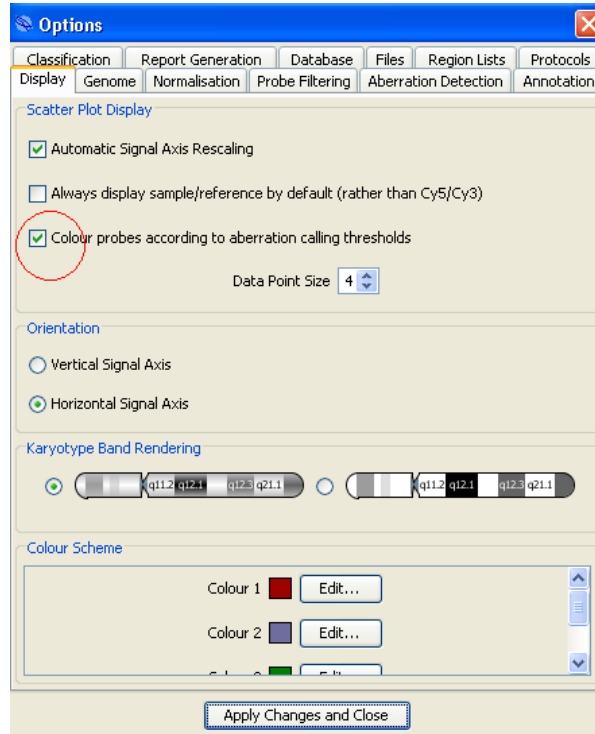


Figure: The “Colour probes according to aberration calling thresholds” check box in the **Display** tab of the **Options** dialog box.



The figure below shows an example with probes below the gain threshold displayed in red and other probes displayed in black.

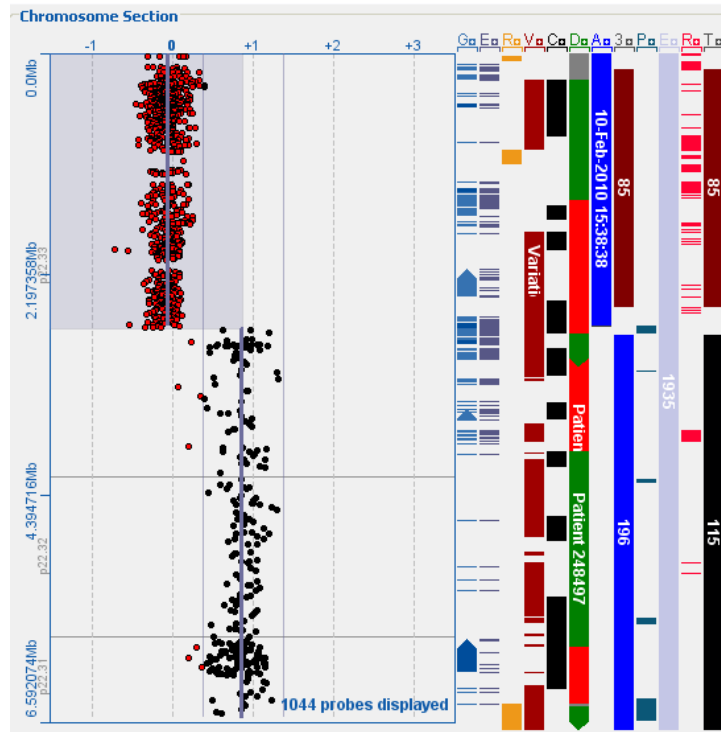


Figure: Probes coloured according to aberration calling thresholds in the **Genomic View**.

Ability to remove aberrations from the summary table based on classification

It is possible to remove aberrations from the report based on classification by modifying the "Report?" column of the table in the **Classification** tab of the **Options** dialog box. By default, the excluded aberrations will be included in the summary table at the beginning of the report, but it is now possible to remove these aberrations from the summary table by selecting the "Also remove these aberrations from the summary table?" check box in the ***Reported Classifications** section.

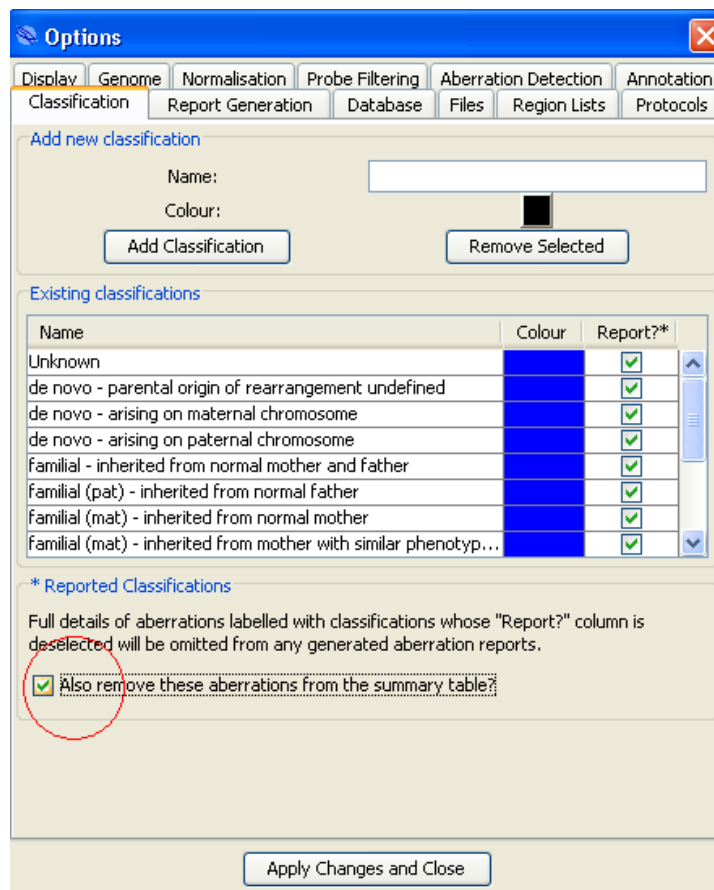


Figure: Removing aberrations from the report summary table.

Protein coding genes highlighted in the report

The open reading frames annotated by Ensembl as protein coding will now be displayed in bold in the report.

Ability to filter the Gene annotation track

It is now possible to filter the gene annotation track by gene type. For example, to only display the protein coding genes in the Gene annotation track, complete the **Annotation Filters** form of the **Annotation** tab in the **Options** dialog box as shown in the figure below, and click **Create Filter**.

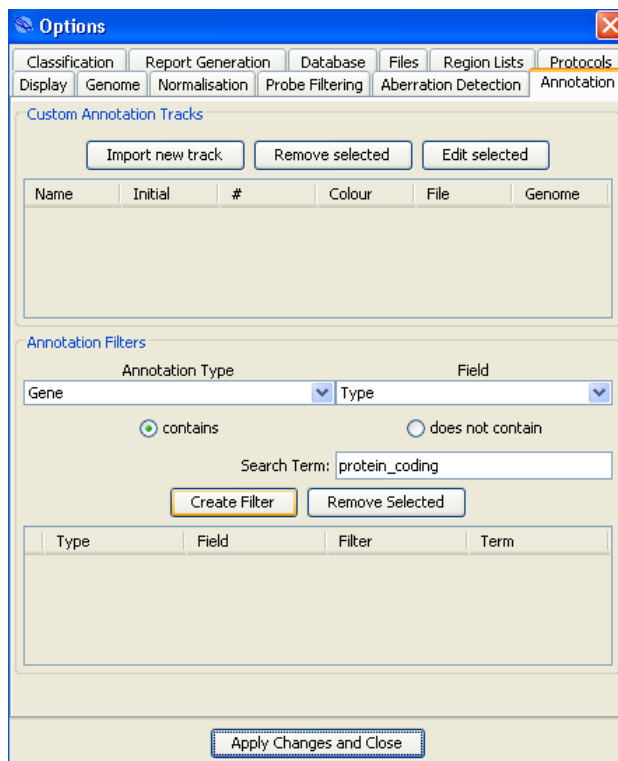


Figure: Filtering the Gene annotation track

To identify possible search terms, hover the mouse over the annotations to display their tooltips with extra information about the annotation.

Segmentation and probe filtration consistency warning

It is possible to apply probe filters during Circular Binary Segmentation (CBS) to exclude poorly-performing probes. If the probe filtration settings are subsequently altered during analysis, the software will now warn the user that the settings are inconsistent with the segmentation, and advise re-segmentation if analysis is to continue with the new probe filters. Please note that probe filtration settings will only be applied during CBS and aberration calling if the “Apply probe filters during Aberration Detection?” check box in the **Probe Filtering** tab of the **Options** dialog box is selected, and so if this setting is also subsequently changed, the segmentation may be inconsistent. For example, if this option is selected and the “Exclude Non-Uniform Outliers” filter is turned on before segmentation, the software will ignore non-uniform probes when carrying out segmentation. If either of these settings is subsequently altered, aberration calling will be carried out on the full dataset and the filtration settings will be inconsistent with the segmentation. In this case, the software will display the following warning:

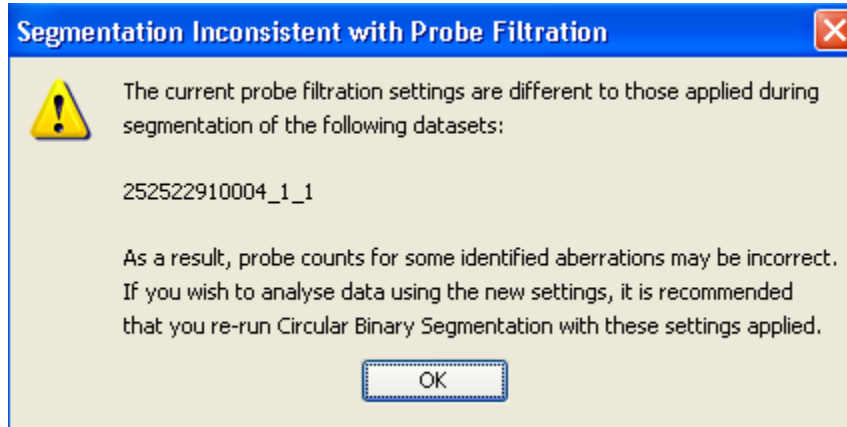


Figure: Warning following modification of probe filtration settings after segmentation.

Filtration status indicators

To highlight the status of probe, database and annotation filtration, three new icons have been added to the controls section in the top-right corner of the **Genomic View**. The icons will initially appear disabled and become live when the user applies filters to the probes, annotations or database aberrations. Clicking on a live filtration status indicator will result in the display of a summary of the current filtration settings for the selected indicator.

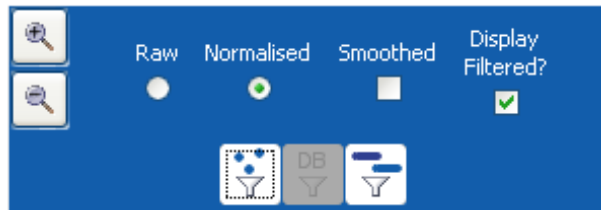


Figure: Probe, database and annotation filtration status indicators.

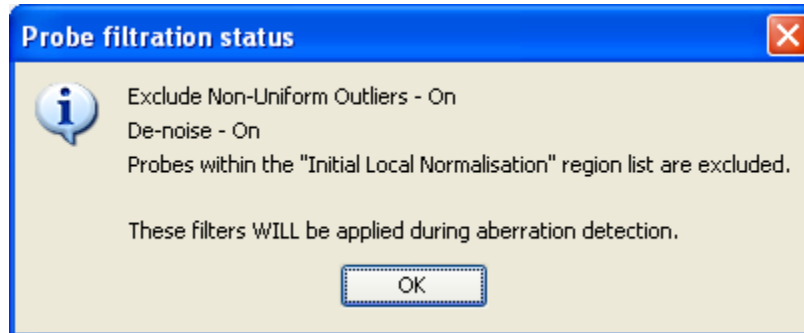


Figure: Example probe filtration status displayed after clicking on the indicator.



Delete button added to Database Management

A new “Delete” button has been added to the **Database Management** tab to enable deletion of the selected submission or aberration with the mouse. Deletion using the keyboard delete key has been disabled.

Ability to filter probes by name

It is now possible to filter probes by name by entering a search term in the “Probe Name” input field in the **Probe Filtering** tab of the **Options** dialog box. This can be applied to separate analysis of some custom array designs generated by combining existing designs. Contact OGT for help with the appropriate term to add for custom arrays.

File naming convention applied during manual saving

If exactly one dataset is loaded, the file naming convention selected in the **Files** tab of the **Options** dialog box will be suggested automatically when the user clicks **File -> Save As....**

Bug fixes

- ISCA array design names now recognised.
- Aberration size now rounded rather than truncated.
- Low resolution annotation placement corrected.
- Log ratio scale now maintained between chromosomes when automatic rescaling is turned off.



CytoSure Interpret Software update — version 4.0 December 09

The following new features are available in CytoSure Interpret software version 3.1.4 onwards:

Improved “Find Annotation” function

120

Two thresholds for aberration calling

120

Threshold lines visible in the genomic view

121

Ability to add region lists

121

For filtration

For differential aberration calling

ISCN notation in reports

125

Expanded QC metrics with possible reasons for poor metrics

126

Custom reporting on the basis of classification

CytoSure Interpret - Software update guide



	126
Cartagenia BENCH compatibility	
	127
Ability to add custom annotation tracks using banding information	129
Updated annotation	129
Various bug fixes	129



Improved “Find Annotation” function

The **Find Annotation** function has been improved. The function can be used to search for annotations, including genes, syndromes and CNVs, using the name or part of the name or the annotation. Users can also now search with wildcards by using “*” in place of any unknown text. For example, entering “BRCA*” would return both BRCA1 and BRCA2. The **Find Annotation** function is available in **View** menu in the toolbar.

Two thresholds for aberration calling

It is now possible to use two thresholds for calling aberrations following Circular Binary Segmentation (CBS). One threshold can be set for losses and the second threshold for gains. Typically the loss threshold (e.g. 0.8) will be higher than the gain threshold (e.g. 0.4). There are now two threshold input fields in the **Aberration Detection** tab of the **Options** dialog box to facilitate this.

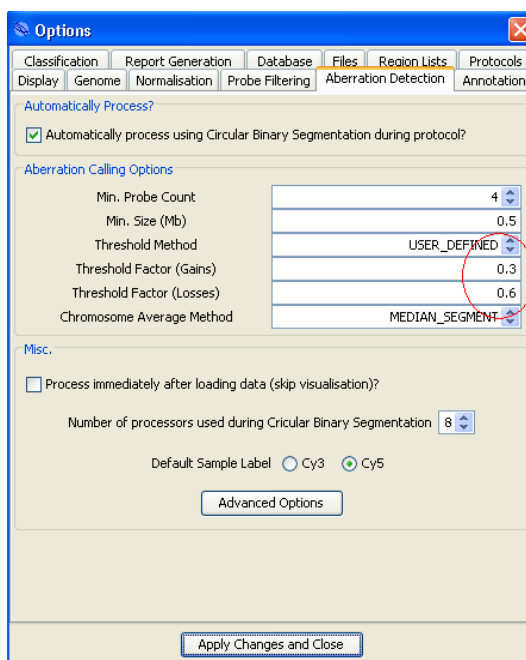


Figure: The two threshold factor input boxes in the **Aberration Detection** tab in the **Options** dialog box.



Threshold lines visible in the genomic view

The position of the thresholds for both loss and gain are now visible in the genomic view.

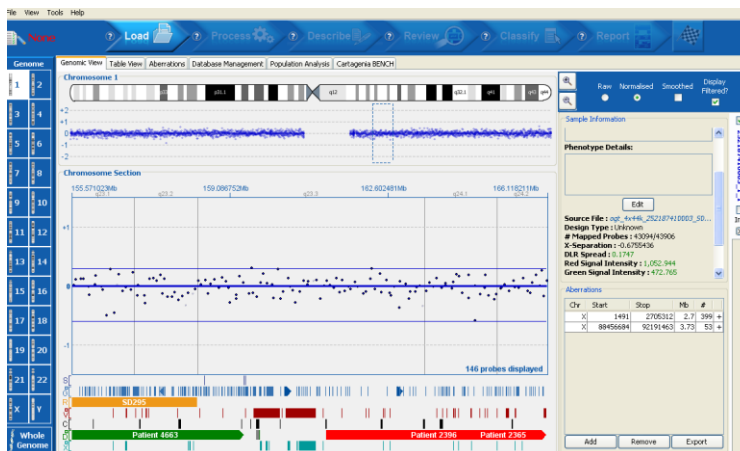


Figure: Image of the **Genomic View** showing the two threshold lines in addition to the segmentation line

Ability to add region lists

By adding genomic region lists to the software the user can then do the following:

Filter out or only include certain regions of the genome

This could be useful, for example, if the user wishes to remove certain commonly occurring CNVs from the analysis.

Set different thresholds for aberration calling at different regions

The user can use different thresholds in different genomic regions. So, for example, it is possible to use a different threshold in syndromic regions compared to backbone regions.

Use regions to do local normalisation

The regions can be used for a local normalisation. Local normalisation is important for analysing data from very high resolution arrays where an aberration covers a large percentage of the probes on the array.

Adding regions

To add regions to the software, click **Tools -> Options -> Regions**.

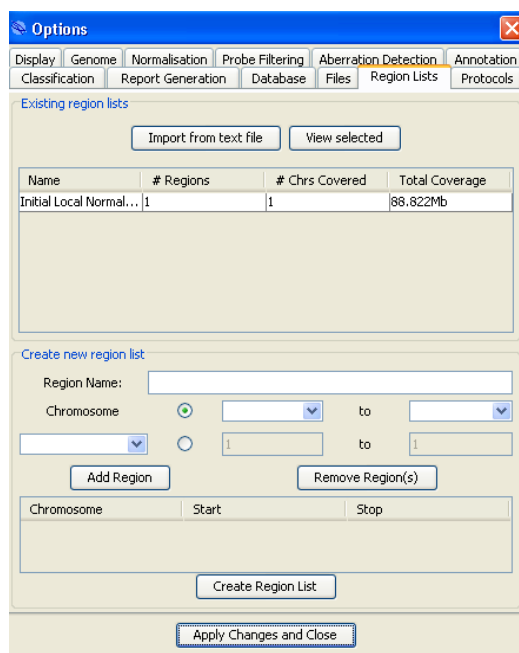


Figure: The **Region Lists** tab in the **Options** dialog box for importing region lists

There are two methods for importing the regions. First, a .txt file can be created using Excel which has the columns “Chr”, “Start (basepair)”, “Stop (basepair)”. Then, in the software, click **Import from text file** and select the relevant file and the regions will be loaded.

The second method is to create a list directly within the software. First, type in the **Region Name**, select a chromosome and the region either as a band or by entering the basepair position, then click **Add Region**. Multiple regions can then be added. When finished, click **Create Region List**.

Use of the region lists

Filter on regions (e.g. remove common CNVs)

Once a region list has been created and named, this can be used to filter out probes from the data analysis. Click **Tools -> Options -> Probe Filtering**

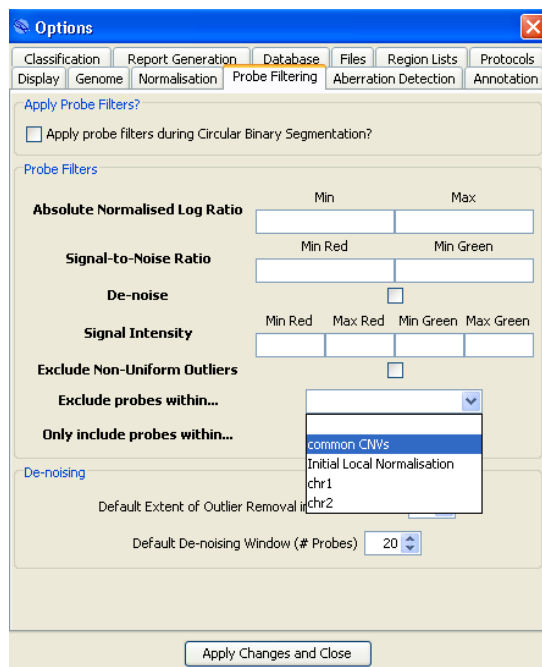


Figure: Filtering on a region list

To exclude probes in a gene list, select the relevant region list from the **Exclude probes within...** drop down menu. Alternatively, to only include probes within these regions, select the relevant region list from the **Only include probes within...** drop down menu.

Set different thresholds for aberration calling at different regions

Once a region list has been established, different thresholds in different regions can be set. Select the **Aberration Detection** tab in the Options dialog and click **Advanced Options**.

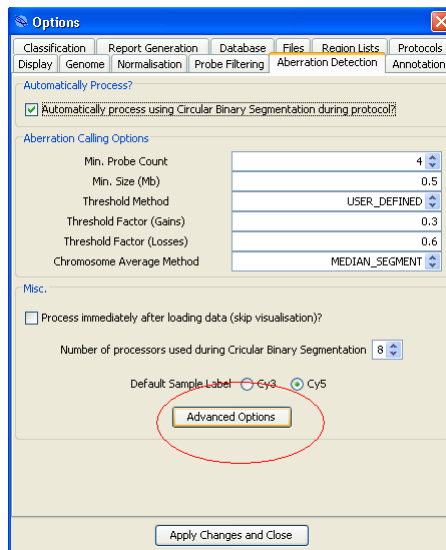
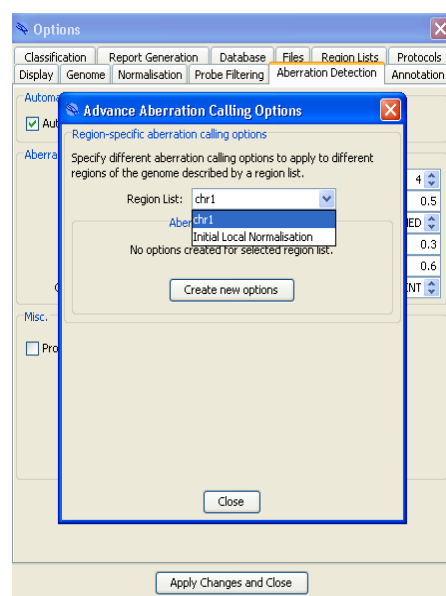
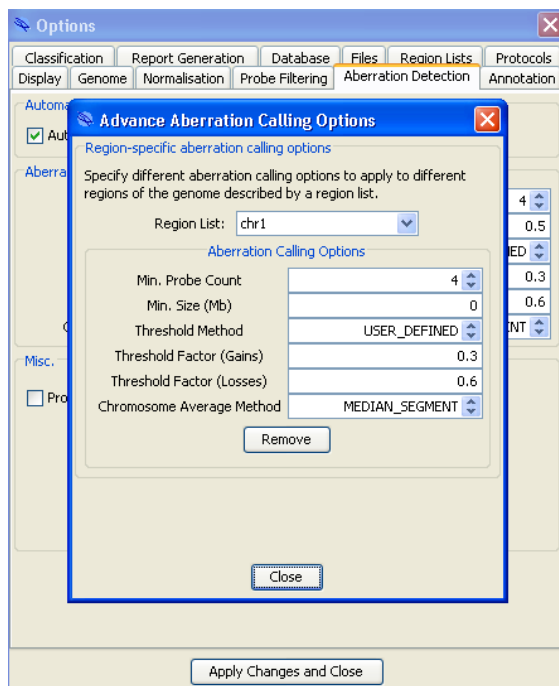


Figure: Select **Advanced Options**



Then, select the relevant region list in the drop down menu and click **Create new options**. A dialog box opens that allows the user to select alternative aberration calling parameters for the particular region list. So, in this case, we can select different parameters for chromosome 1.



If a particular segment is partly over a region list (e.g. a syndromic region) and a non-region list then the aberration will be called using the least 'stringent' method. The aberration will be annotated as being partly in a region list and partly outside the region list.

ISCN notation in reports

The ISCN notation is now included as a new field in the report. The software will calculate the copy number of the aberration by using the thresholds values that have been set by the user. So, for the autosomes, a x1 aberration would have a mean segment value of the loss threshold value, whilst the mean of a x4 aberration will be twice the gain threshold value.

Expanded QC metrics and possible reasons for poor metrics

Many of these metrics are available in the Agilent quality report PDF file. As such, this feature is principally for users of GenePix feature extraction software.

The extra metrics are:

Grid Alignment

Uses the positive controls in the corner of the array to detect if the feature extraction has been carried out correctly.



Negative Controls

Reports the mean value of the negative control probes. If this figure is high then it suggests that either the washing or hybridisation was not optimal.

Signal Intensity Ratio

The ratio of the mean intensity of Cy5 signal to mean Cy3 signal intensity is calculated.

If the value is very low, then this would suggest an Ozone issue as it degrades the Cy5 dyes. OGT encloses scanners in boxes with Ozone scrubbers fitted.

Non-Uniform Features

Reports the number of flagged probes from the Agilent or GenePix feature extraction. If high, the outlier probes can be filtered out from the analysis using **Tools -> Options -> Probe Filtering**.

Signal Reproducibility

Gives a metric on the evenness of the hybridisation across the slide. This will only be given if the appropriate controls are present on the array. These are on OGT's recent arrays.

The QC metrics are highlighted on opening of the data if there is a problem. The metrics are also available in the **Sample Information** section of the **Genomic View** and by selecting **File -> Properties**.

In addition to these QC metrics, an extra field containing the Standard Deviation of the probes within called segments has been added to the aberration table.

Custom reporting on the basis of classification

It is now possible to specify which aberrations to report on the basis of their classification. This could be useful if, for example, the user does not want CNVs included in the report. Click on **Tools -> Options -> Classification** and click on the **Report?** tick box to exclude aberrations classified with the selected classification from any generated reports. Please note that this does not affect the submission of aberrations to the database.

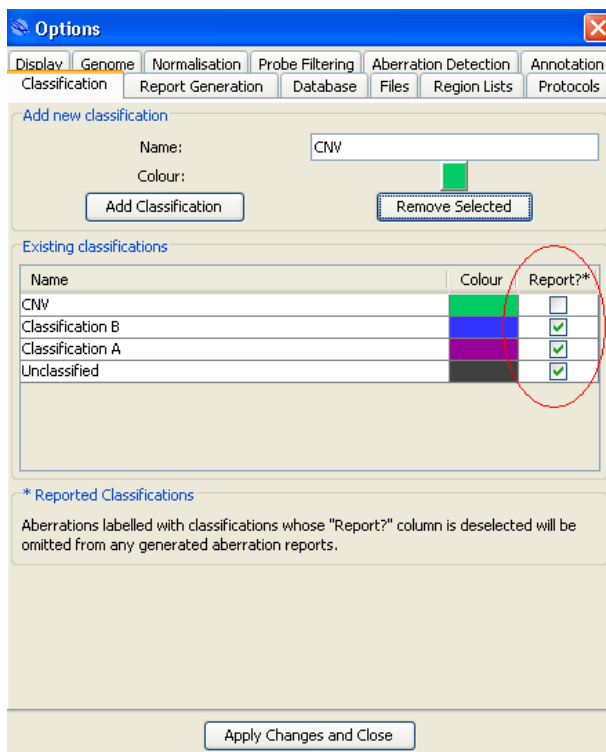


Figure: Selective reporting on the basis of classification

Cartagenia BENCH compatibility

The Cartagenia software is a sophisticated tool that includes interpretation and data sharing capabilities. The user needs to acquire a licence for the Cartagenia BENCH software in order to use this feature. The BENCH software allows users to manage copy number data within one lab as well as enabling consortium members to share aberration information. The aberration data can be uploaded to BENCH with the classification attached. Conversely, data that is stored in BENCH can be downloaded onto a custom annotation track into CytoSure.

The user can:

- Specify login details for multiple BENCH accounts
- Download the latest annotations from those accounts
- Specify a “default” BENCH account so that they can...
- Specify whether their data should be automatically uploaded to the “default” account when processed in workflow mode

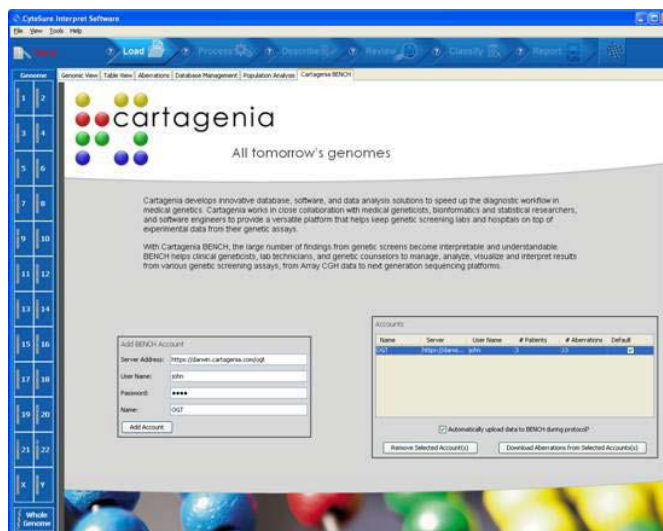


Figure: The **Cartagenia BENCH** tab

An additional annotation track for each BENCH account, with all the usual features of the annotation tracks, such as linking to a related webpage and display of various information when hovering over each annotation, is available in the **Genomic View** and **Population Analysis** tabs.

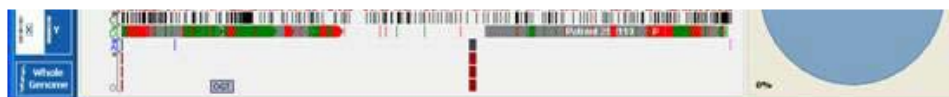


Figure: The BENCH annotation track

An extra button in the Aberrations tab labelled “Export to Cartagenia BENCH” has been added to facilitate manual submission of detected aberrations to the default BENCH account.

The information that is downloaded from BENCH is stored locally, so the user doesn't have to be connected to the internet all the time to see their aberrations. The data is managed within the BENCH interface and any changes to aberrations (extra comments, changes in classification etc.) are logged so that the next time the user clicks the **Download New/Updated Aberrations** button, these changes will be mirrored in the CytoSure annotation track(s).



Ability to add custom annotation tracks using banding information

The ability to enter custom annotation tracks using band names, as opposed to molecular location, has been added. If banding information is used, there are two important points to note: Firstly, the band names need to include the chromosome; Secondly, the band names must include sub-band numbers, if applicable.

	A	B	C	D
1	chr	start	stop	name
2	1	1p36.33	1p36.13	345345
3	9	9p24.1	9p23	6575676
4				
5				
6				
7				
8				

Figure: The required format for the .txt file for creating a custom annotation track using banding

Updated annotation

The Syndrome, DGV, Decipher annotation has all been updated.

Various bug fixes

- Inverted images shown correctly in the report.
- Database annotation bug - with the previous version, database fields may be incomplete. To regenerate the missing information, click **Tools -> Update Database Annotations**.
- Faster loading of previously saved .cgh files.
- Loop issue bug fix.
- Bug resulting in inability to open software if report logo image file is moved fixed.
- Genes in aberration table fixed.

CytoSure Interpret - Software update guide



Contact us

Oxford Gene Technology,
Begbroke Hill,
Woodstock Road,
Begbroke,
Oxfordshire
OX51PF,UK

T: +44(0)1865 856826; (US: 914-467-5285)
E: contact@ogt.com; W: www.ogt.com

Technical support : E : contact@ogt.com
 T: +44(0)1865 856826

Trademarks

Trademarks: OGT™, CytoSure™, Excel®, Nanodrop™

For research use only

This product is provided under an agreement between Agilent Technologies, Inc. and OGT. The manufacture, use, sale or import of this product may be subject to one or more of U.S. patents, pending applications, and corresponding international equivalents, owned by Agilent Technologies, Inc. The purchaser has the non-transferable right to use and consume the product for RESEARCH USE ONLY AND NOT for DIAGNOSTICS PROCEDURES. It is not intended for use, and should not be used, for the diagnosis, prevention, monitoring, treatment or alleviation of any disease or condition, or for the investigation of any physiological process, in any identifiable human, or for any other medical purpose.

The information and content included in this update guide are not intended to address any particular technical requirements; they do not constitute any form of technical advice or recommendation by Oxford Gene Technology, and are not intended to be relied upon when making (or refraining from making) any decision. Oxford Gene Technology accepts no liability for any loss or damage caused by any reliance on information or content included in this update guide.

Oxford Gene Technology (Operations) Ltd. Registered in England No: 03845432 Begbroke Science Park, Begbroke Hill, Woodstock Road, Begbroke, Oxfordshire, OX5 1PF, UK.

This document and its contents are © Oxford Gene Technology (Operations) Ltd. – All rights reserved. OGT™ and CytoSure™ are trademarks of Oxford Gene Technology (Operations) Ltd .