

Copy Number Variation Analysis Module

USER GUIDE

for use with Diomni™ Design and Analysis (RUO) Software 3

Publication Number MAN0030169

Revision A



Life Technologies Holdings Pte Ltd | Block 33 | Marsiling Industrial Estate Road 3 | #07-06, Singapore 739256
For descriptions of symbols on product labels or product documents, go to thermofisher.com/symbols-definition.

Revision history: MAN0030169 A (English)

Revision	Date	Description
A	18 February 2025	New document for Copy Number Variation Analysis Module v1.0 (Dionni™ Design and Analysis (RUO) Software 3).

The information in this guide is subject to change without notice.

DISCLAIMER: TO THE EXTENT ALLOWED BY LAW, THERMO FISHER SCIENTIFIC INC. AND/OR ITS AFFILIATE(S) WILL NOT BE LIABLE FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING YOUR USE OF IT.

NOTICE TO PURCHASER: DISCLAIMER OF LICENSE: Purchase of this software product alone does not imply any license under any process, instrument or other apparatus, system, composition, reagent or kit rights under patent claims owned or otherwise controlled by Thermo Fisher Scientific, either expressly, or by estoppel.

Trademarks: All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. TaqMan is a trademark of Roche Molecular Systems, Inc., used under permission and license.

©2024-2025 Thermo Fisher Scientific Inc. All rights reserved.

Contents

■ CHAPTER 1	Copy Number Variation Analysis Module	5
■ CHAPTER 2	Workflow: Copy number variation analysis	6
■ CHAPTER 3	Set up a plate file	7
	Open the plate setup for a project	7
	Confirm or edit the plate setup	7
	Add samples and assign to wells	7
	Add CNV assays and assign to wells	8
	Convert the targets or SNPs to CNV assays	9
	Edit reagent information	12
	Select a passive reference	13
■ CHAPTER 4	Perform copy number variation analysis	14
	Analysis module	14
	Update the page layout	14
	Review results in the Amplification Plot	15
	Edit copy number variation analysis settings	15
	Edit the general settings	15
	Edit controls	16
	Edit the quality criteria	17
	Edit the advanced settings	17
	Use a two-copy control	18
	Review copy number variation results	19
	Select data files or CNV assays to review	19
	Review the copy number vs sample graph	20
	Review the results	21
	Overview of the result issues	22
	Analysis settings with low accuracy	23
	Export the analysis units	23
	Export the results	24

■	CHAPTER 5 Overview of copy number analysis	25
	Overview of the functions	25
	Overview of the analysis	25
	Overview of multiplate analysis	26
	Calibrator sample	27
	Overview of copy number bins	27
	Overview copy number sample graph	27
	Overview of the copy number results	28
	Data analysis workflow	29
	About the theoretical model	31
	Biochemical basis of the model	31
	About the statistical model	32
	Data preprocessing	34
	Data validation	34
	ΔC_t calculations	34
	Outlier detection	35
	Copy number assignment	35
	Copy number assignment with a calibrator sample	36
	Copy number assignment without a calibrator sample	36
	About the quality metrics	37
	Confidence estimate	37
	Deviation z-score estimate	39
	ΔC_t variability estimate	39
	Additional background information	42
	Recommended minimum data requirements	42
	Interaction of copy number and the confidence metric	44
■	APPENDIX A Documentation and support	45
	Related documentation	45
	Customer and technical support	45
	Limited product warranty	45



Copy Number Variation Analysis Module

The Copy Number Variation Analysis Module for Diomni™ Design and Analysis (RUO) Software 3 is used to perform relative quantitation analysis of genomic DNA targets using the real-time PCR data from copy number assays. The software and associated copy number assays can be used to detect and measure copy number variation of specific sequences in genomes.

For more information about copy number variation analysis, see Chapter 5, “Overview of copy number analysis”.



Workflow: Copy number variation analysis

Copy number variation analysis

Open a project and add run files to the project

See the primary user guides for the software.

Open the plate setup for a project (page 7)

Confirm or edit the plate setup (page 7)

Convert the targets or SNPs to CNV assays (page 9)

Review results in the Amplification Plot (page 15)

Edit copy number variation analysis settings (page 15)

Review copy number variation results

- Select data files or CNV assays to review (page 19)
- Review the copy number vs sample graph (page 20)
- Review the results (page 21)



Set up a plate file

For detailed instructions about setting up a plate file, see the primary user guides for the software. See “Related documentation” on page 45.

Open the plate setup for a project

1. Open a project file.
 - If the project file is in the software, navigate to the **Projects** page, then select the **Recents** tab.
 - If the project file is not in the software, navigate to the **Projects** page, then click **Actions ▶ Open File**.
 - If the project file is not in the software, navigate to the **Dashboard** page, then click **Open File**.
2. Select the **Plate Setup** tab of the project.
3. Select a data file to view.
 - Select a run file from the **Run file** dropdown list.
 - Use the arrows beside the **Run file** dropdown list to scroll through the run files.

Confirm or edit the plate setup

For detailed instructions about plate setup, see the primary user guides for the software.

Add samples and assign to wells

For detailed instructions about plate setup, see the primary user guides for the software.

1. In the **Plate Setup** tab, add samples and assign to wells using the following options.
 - Import a plate setup file.
 - Manually add samples to the **Samples** table.
To assign a sample from the **Samples** table to the well, select a well in the plate layout, then select the checkbox associated with the sample in the **Samples** table.
 - Manually add samples to wells in the plate layout.
The sample is added to the **Samples** table after it is added to the plate layout.

For the OpenArray™ Plate format, a sample layout can be imported. Samples cannot be added manually. The other formats allow a plate setup file to be imported, including targets and assays.

2. Confirm or edit sample information in the **Samples** table.

- Name
- Color
- Type

Controls that are defined in the copy number variation analysis settings overwrite any controls that are defined in the plate setup. For more information, see “Edit controls” on page 16.

3. Confirm or edit sample well assignments in the plate layout.

Add CNV assays and assign to wells

For detailed instructions about plate setup, see the primary user guides for the software (see “Related documentation” on page 45).

CNV assays cannot be imported or assigned for the OpenArray™ Plate format.

IMPORTANT! If the data are exported for import into AlleleTyper™ Software, the CNV name must end with `_cn`.

1. In the upper-left corner of the **Plate Setup** tab, click **CNV**.

2. Add CNV assays and assign to wells using the following options.

- Use the CNV assay converter for runs that were set up with SNPs or targets if the instrument software did not include CNVs.
- Import an Assay Information File (AIF).
- Import a plate setup file.
- Manually add CNV assays to the **CNV Assays** table.
To assign a CNV assay from the **CNV Assays** table to the well, select a well in the plate layout, then select the checkbox associated with the CNV assay in the **CNV Assays** table.
- Manually add CNV assays to wells in the plate layout.
The CNV assay is added to the **CNV Assays** table after it is added to the plate layout.
- Import TaqMan™ assay plate and card files.


If an AIF or a TaqMan™ assay file is imported, the SNPs or targets must be converted to CNV assays.

3. Confirm or edit target information in the **CNV Assays** table.

Column	Description
Name	CNV assay name
Color	CNV assay color
Reference reporter	Fluorescent reporter dye for the reference assay
Reference quencher	Quencher for the reference assay

(continued)

Column	Description
Test reporter	Fluorescent reporter dye for the test sequence
Test quencher	Quencher for test sequence

4. Click  (**Edit**) to open the **Edit CNV Assay** dialog box.
The **Edit CNV Assay** dialog box contains the information that was available in step 3.
5. Edit the following items in the **Edit CNV Assay** dialog box.
 - Enter an assay ID in the **Assay ID** field
 - Enter a name or base in the **Reference** field
 - Enter a name or base in the **Test** field
 - Select a reference color
 - Select a test color
 - Add comments to the **Comments** field
6. Confirm or edit CNV assay well assignments in the plate layout.

Convert the targets or SNPs to CNV assays

Overview of the CNV assay conversion

Many instrument runs are set up with CNV assays defined as SNP assays or targets. This enables setup if the instrument software does not include CNV assays.

The conversion allows data files that were set up with targets or SNP assays to be converted to CNV assays. This allows for CNV assay analysis in the Diomni™ Design and Analysis (RUO) Software 3.

The conversion can be set up for the project template or the project (see the primary user guides for the software, “Related documentation” on page 45).

If the conversion is applied to a project template, the conversion is applied to any projects that are created from the project template.

If the conversion is applied to a project, the conversion is applied to any data files that are added to the project.

The conversion can be added to an individual data file in a project (see “Add the CNV assay conversion to a plate” on page 10). When the CNV assay conversion is added to an individual data file within a project, the conversion rules are applied to all of the data files in the project.

The rules for CNV assay conversion can be exported. The rules for CNV assay conversion can be imported for other project templates, projects, or data files.

If the rules for CNV assay conversion were imported, they can be edited within the software for the specific project template, project, or data file.

Add the CNV assay conversion to a plate

For information about the use of the CNV assay converter in a project template or a project, see the primary user guides for the software (“Related documentation” on page 45).

When the CNV assay conversion is added to an individual plate (data file), the conversion rules are applied to all of the data files in the project.

Import a run file into the project.

1. In the **Plate Setup** tab of an open project, select the file from the **Run file** dropdown list.
2. In the plate layout pane, click **⋮ (More Options) ▶ CNV Assay Converter**.
3. In the **CNV Assay Converter** dialog box, click **+ Add rule**.
A new row is displayed in the **CNV Assay Converter** dialog box.
4. Enter the name of the target or SNP assay in the **Target / SNP Assay** field.
The following wildcard characters are accepted:
 - The * character for any number of wildcard characters
 - The ? character for a single wildcard character
 A wildcard character can be used only at the end of the name.
5. Enter the name of the reference assay in the **CNV Reference Name** field.
6. Enter the name of the copy number assay in the **CNV Assay Name** field.
The content of the fields in each row cannot be duplicated.
7. (Optional) Import rules.
See “Import a CNV assay conversion file for a plate” on page 11.
8. Click **⊗ (Delete)** in a row to delete a single rule.
9. Click **⊗ (Delete)** in the table header to delete all of the rules.
10. Click **Apply**.

The targets or the SNP assays are converted to CNV assays.

In the plate layout pane, click the **CNV** tab. The CNV assays are defined in dark text. The SNP assays are defined in light text when viewing the plate layout pane with the **CNV** tab selected.

The right pane displays the CNV assays.

Import a CNV assay conversion file for a plate

The file must be in CSV format.

The file must contain the following headers:

- **Target / SNP Assay**
- **CNV Reference Name**
- **CNV Assay Name**

The conversion does not overwrite any of the CNVs that were previously defined for the plate. The CNVs that are defined in the CNV assay conversion file are added to the list of CNVs that are defined in the **CNVs** tab.

If a CNV with the same name was previously defined in the project, the CNV from the CNV assay conversion file is not added to the list of CNVs that are defined in the **CNVs** tab.

1. In the **Plate Setup** tab of an open project, select the file from the **Run file** dropdown list.
2. In the plate layout pane, click **... (More Options) ▶ CNV Assay Converter**.
3. At the top-right corner of the grid view or the table view of the plate layout, click **... (More Options) ▶ CNV Assay Converter**.
4. In the **CNV Assay Converter** dialog box, click **Import**.
5. In the **Open** dialog box, navigate to the location of the file, then click **Open**.

The **CNV Assay Converter** is populated with the rules for conversion.

Add or remove rules from the rules that were imported, if necessary.

Click **Apply** to apply the imported rules.

Export a CNV assay conversion file from a plate

The file is in CSV format.

The exported file can be imported to other project templates, projects, or plate setup files.

1. In the **Plate Setup** tab of an open project, at the top-right corner of the grid view or the table view of the plate layout, click **... (More Options) ▶ CNV Assay Converter**.
2. In the **CNV Assay Converter** dialog box, click **Export**.
3. Navigate to the location to save the file, then edit the file name, if necessary.
4. Click **Save**.

Edit reagent information

Reagents can only be edited for the TaqMan™ Array Card format and the OpenArray™ Plate format. Reagents cannot be added.

In the on-premise configuration, adding, editing, or deleting a reagent is a controlled function. The user role must have the permission of **Add/Edit/Delete Reagent**.

1. In the **Plate Setup** tab, in the **Targets/SNP Assays** table pane, click **Reagents**.
2. In the **Reagents** table, perform one of the following actions.
 - Click **+** (**Add**).
 - Click **...** (**More Options**) ▶ **Export Reagents** to export reagents.
 - Click **...** (**More Options**) ▶ **Import Reagents** to import reagents.
 - Click **...** (**More Options**) ▶ **Scan Reagents** to scan reagents.
3. If you are manually adding reagents or editing reagents, enter the following information in the table.

• Name	• Part Number
• Type	• Lot Number
• Barcode	• Expiration Date

Note: If the master mix that you enter is not compatible with the current run method, you have the option to apply the recommended run method for your master mix, instrument, block, and run mode.

4. If you are scanning the reagent barcode, in the **Scan Reagent** dialog box, select or deselect the **Enable automatic parsing** checkbox.
5. If you are scanning the reagent barcode, when the **Scan Reagent** dialog box is displayed, use a barcode scanner to scan the reagent label.

Note: If the master mix that you enter is not compatible with the current run method, you have the option to apply the recommended run method for your master mix, instrument, block, and run mode.

The fields in the **Scan Reagent** dialog box are populated.

6. In the **Scan Reagent** dialog box, click **Add**.
7. (Optional) Click **✕** (**Remove**) in the row of a reagent to delete it from the table.

Select a passive reference

In the on-premise configuration, editing the passive reference is a controlled function. The user role must have the permission of **Edit Passive Reference**.

The passive reference is set for the plate. The default passive reference is ROX™ dye.

1. In the upper-left corner of the **Plate Setup** tab, select a passive reference from the dropdown list.
2. (Optional) Save the plate file or data file.



Perform copy number variation analysis

Analysis module

The analysis module for a project is defined in the project settings. The analysis module is applied to the run file when the run file is added to the project. The analysis module is defined based on the block type of the run file.

For more information, see the primary user guides for the software. See “Related documentation” on page 45.

Update the page layout

The **Customize Page Layout** dialog box has two sides. The left side displays the options that are available. The right side displays the layout.

In a project, the top pane of each analysis module tab is fixed. The top pane cannot be removed from the page layout and cannot moved on the page layout.

1. In the **Quality Check** tab or analysis module tab of an open run file or open project file, click **Actions ▶ Page Layout Setting**.
2. In the **Customize Page Layout** dialog box, drag an item from the left side to the right side in order to display the item.
3. Click and drag an item on the right side to arrange the display.
4. (Optional) Click **Reset to Default** to display the items according to the default.
5. Click **Save**.

Review results in the Amplification Plot

For detailed instructions about reviewing results in the **Amplification Plot**, see the primary user guides for the software.

If no data are displayed in the **Quality Check** tab, or if reanalysis is required, click **Analyze**.

1. In the **Quality Check** tab, in the plot pane, select **Amplification Plot** from the dropdown list.
2. Review the amplification status for each well.
3. Review or edit threshold settings.
4. Review or edit baseline settings.

Edit copy number variation analysis settings

Edit the general settings

1. In an open project, click **Actions** ▶ **Copy Number Variation Analysis Setting**.
2. In the **Copy Number Variation Analysis Setting** dialog box, select the **General** tab.
3. Select or deselect the **Use Default** checkbox for each CNV assay.
The default settings are defined in the first row. Deselecting the **Use Default** checkbox enables the settings to be updated for each individual CNV assay.
If the default settings are updated, these settings are applied to each CNV assay.
4. Select or deselect the **Multiplate** checkbox.
The **Multiplate** checkbox is selected by default.
If multiplate analysis is selected and the same samples are included on multiple plates, they are analyzed as separate items.
For more information, see “Overview of multiplate analysis” on page 26.
5. Select a value from the **Calibration** dropdown list.
 - **Auto calibration**
 - **Median ΔC_t**
 - A sample calibrator

A sample calibrator cannot be selected as the default setting. The **Use default** checkbox must be deselected for the CNV assay to be able to select the calibration method for the CNV assay.

A sample calibrator can be selected only for a CNV assay. The sample calibrator is specific to the CNV assay.

Each sample from the project that is assigned to a CNV assay can be available in the **Calibration** dropdown list. The sample is available if it meets the criteria that are required for it to be a calibrator sample.

For more information, see “Calibrator sample” on page 27.

6. Enter or edit the value in the **Copy Number** field.
 The number is the known copy number of the target in the calibrator sample.
 The **Copy Number** field is editable only if median ΔC_t or a sample calibrator was selected as the calibration method in step 5.
7. Select a value in the **CN Start** dropdown list and the **CN End** dropdown list.
 The values in the dropdown lists can be edited only if auto calibration was selected as the calibration method in step 5.
 The range for each field is between **1** and **9**. The default value for the **CN Start** dropdown list is **2** and the default value for the **CN End** dropdown list is **3**.
 The value in the **CN Start** dropdown list must be less than the value in the **CN End** dropdown list.
8. Enter or edit the value in the **Ref Cq threshold** field.
 The reference Cq threshold value defines the filtering threshold used by the software to remove poor quality data before the analysis. The default value is 32 for 384-well plates and 30 for an OpenArray™ Plate.
 If a replicate has a C_t value greater than the threshold value, the software excludes the well from the analysis.
 For more information, see “Data validation” on page 34.
9. Enter or edit the value in the **Zero copy ΔC_t threshold** field.
 Defines the filtering threshold used by the software to categorize samples with relatively weak, background amplification as having zero copies of the target. The default value is 4.0.
10. Click **Apply** or select the **Controls** tab in the dialog box to continue to edit the settings.

Edit controls

1. In an open project, click **Actions** ▶ **Copy Number Variation Analysis Setting**.
2. In the **Copy Number Variation Analysis Setting** dialog box, select the **Controls** tab.
3. Select or deselect the **Override Task Assignments from Plate Setup** checkbox.
 The **Override Task Assignments from Plate Setup** checkbox is selected by default.
 When the checkbox is selected, the controls defined in the **Copy Number Variation Analysis Setting** dialog box overwrite the controls that are defined in the plate setup.
4. Enter a value or multiple values in the **Negative Control** field.
 The default values are **NTC** and **NEC**.
5. In the **Positive Control** pane, click **+ Add Control**.
6. Select an assay from the **Assay name** dropdown list.
7. Enter a sample name or multiple sample names in the **Sample name** field.
 Each sample name can include wildcard characters at the end of the name. An asterisk (*) can represent any number of characters. A question mark (?) represents a single character.

8. Enter a value in the **Known Copy Number** field.
A value of 0 is permitted for a positive control.
The default value is 2.
9. Click **Apply** or select the **Quality Criteria** tab in the dialog box to continue to edit the settings.

Edit the quality criteria

If a sample fails the quality criteria that have been selected, the failure is noted in the results table, in the **Result issues** column. The following results are undetermined:

- Predicted copy number
- Difference in calculated copy number and predicted copy number

The software reports the calculated copy number for all of the samples, including samples with failed quality criteria.

1. In an open project, click **Actions** ▶ **Copy Number Variation Analysis Setting**.
2. In the **Copy Number Variation Analysis Setting** dialog box, select the **Quality Criteria** tab.
3. Select the checkboxes for the criteria to apply to the project.
 - **|Z-Score| threshold** checkbox
 - **Minimum % of analyzed replicates** checkbox
 - **Minimum confidence** checkbox
 - **Confidence interval QC** checkboxNo checkboxes are selected by default.
4. Enter a value in the fields associated with the criteria that were selected in step 3.
5. Click **Apply** or select the **Advanced** tab in the dialog box to continue to edit the settings.

Edit the advanced settings


For more information about bins, see “Overview of copy number bins” on page 27.

1. In an open project, click **Actions** ▶ **Copy Number Variation Analysis Setting**.
2. In the **Copy Number Variation Analysis Setting** dialog box, select the **Advanced** tab.
3. Select or deselect the **Use copy number bins** checkbox.
4. Click **+ Add Bin** to add a bin.
The default is 3 bins.
The maximum number of bins is 10.
5. For each bin, enter a value in the **Starting copy number** field and the **Ending copy number** field.
The default values are the same number in the **Starting copy number** field and the **Ending copy number** field.

The default value for a bin is one integer greater than the value in the **Ending copy number** field for the previous bin.


The values in each field must be sequential.

The value in the **Ending copy number** field of the last bin is greater than or equal to the value in the **Starting copy number** field. For example, a value of 3 in the **Starting copy number** field produces a value of ≥ 3 in the **Ending copy number** field.

6. Click  **Remove Bin** to remove a bin.
 Two bins are required. A bin cannot be removed if there are only two bins.
 The last bin is removed to preserve the sequence of the values in the **Starting copy number** field and the **Ending copy number** field.
7. Click **Apply** or select a different tab in the dialog box to continue to edit the settings.

Use a two-copy control

In order to use a two-copy control, the plate setup requires a two-copy positive control.

1. In an open project, click **Actions** ▶ **Copy Number Variation Analysis Setting**.
2. In the **Copy Number Variation Analysis Setting** dialog box, select the **General** tab.
3. In the **Calibration** dropdown list, select **Auto Calibration**.
4. Select the **Controls** tab.
5. In the **Positive Control** pane, click  **Add Control**.
6. In the **Assay name** dropdown list, select an assay.
7. Enter a name in the **Sample name** field.
 The sample name can include wildcard characters at the end of the name. An asterisk (*) can represent any number of characters. A question mark (?) represents a single character.
8. Enter 2 in the **Known Copy Number** field.
9. Click **Apply**.
10. Analyze the data.

Review copy number variation results

Select data files or CNV assays to review

In the **Copy Number Variation** tab of an open project, when the **Single plate** tab is selected, the top pane displays a list of the data files that are included in the project.

The bottom pane displays the results of the data files that are selected.

Multiplate analysis is the default. It is selected in the copy number variation analysis settings (see “Edit the general settings” on page 15).

When the **Multiplate** tab is selected, the top pane displays a list of the CNV assays that are included in the project.

If multiplate analysis is selected and the same samples are included on multiplate plates, they are analyzed as separate items.

1. In the top pane, select the tab that is associated with the data to analyze.

- **Single plate** tab
- **Multiplate** tab

The copy number variation analysis settings must define multiplate analysis in order for data to be displayed in the **Multiplate** tab.

Multiplate analysis can be selected for all of the CNV assays. Multiplate analysis can also be selected for one or more specific CNV assays.

Only the CNV assays that have multiplate analysis selected are displayed in the **Multiplate** tab. See “Edit the general settings” on page 15.

2. Click **View**, then select the items to display in the table.
3. Click the row that is associated with the data file for copy number variation analysis.
Use control-click or shift-click to select multiple rows.
4. (Optional) Click **Select all** to select all of the rows.

The bottom pane is updated to display the results for the data files that were selected. The bottom pane can take several minutes to update if multiple data files are selected.

Review the copy number vs sample graph


For information about the graph, see “Overview copy number sample graph” on page 27 and “About the quality metrics” on page 37.

The graph is not displayed by default. Update the page layout to view the graph (see “Update the page layout” on page 14). The plot must be displayed on the left side or the right side of the page. The plot cannot be displayed at the top or the bottom of the page.

In the top pane, select the data files or the CNV assays to include in the results that are displayed in the graph (see “Select data files or CNV assays to review” on page 19).

1. View the data in the graph.

The title of the graph depends on the plot type that is select in step 4.

2. At the top right of the graph, click  (**Settings**).
3. In the **General** tab, enter a title in the **Plot Title** field.
4. Select an item in the **Plot type** dropdown list.
 - **Calculated copy number**
 - **Predicted copy number**
5. Select an item in the **Sort by** dropdown list.
 - **Sample name**
 - **Assay**
6. Select the items to show in the graph.
 - Legend
 - Error bar
 - Grid
7. In the **X Axis** tab, enter a label in the **Label** field.

8. Enter a value in the **Rotate** field.

The value represents the degrees that the text is rotated.

A value of **0** does not rotate the text and the text is horizontal. A value of **90** rotates the text and the text is vertical. Values in between **0** and **90** can be added to display the text at an angle.

The default value is **25**.

9. In the **Y Axis** tab, enter a label in the **Label** field.

10. Review the following items.

- Samples should have calculated copy number values close to integers and small range bars.
- The presence of intermediate calculated copy number values, such as 1.5, can indicate that the calibrator or copy number was specified incorrectly or that a problem exists with the associated test sample or calibrator sample.
- Large range bars might indicate that the technical replicates of the associated sample exhibit a broad range of ΔC_t values. This could be due to suboptimal sample data quality. To examine the individual ΔC_t values for the technical replicates, see “Review the results” on page 21.

Note: The copy number range of replicates is frequently larger for samples that have high target copy numbers (>3) as a result of their smaller ΔC_t values.

Save the image of the copy number vs sample graph

1. In the **Copy Number Variation** tab, in the graph pane, click **⋮ (More Options) ▶ Save Image**.
2. In the **Save Image** dialog box, enter or select the following information.
 - Enter a file name in the **Filename** field.
 - Select a file format radio button.
 - Select a size from the **Size** dropdown list.
3. Click **Save Image**.
4. Navigate to the location to save the image, then click **Save**.

Review the results

In the top pane, select the data files to include in the results that are displayed in the bottom pane (see “Select data files or CNV assays to review” on page 19).

For more information about the information that is displayed in the results table, see “Overview of the copy number results” on page 28.

1. View the copy number variation data that are displayed in the bottom right pane.
2. At the top-right corner of the table, select the view.
 - **Sample table**
 - **Well table**
3. Click **View**, then select the items to display in the table.
4. (Optional) Use the search field at the top-left corner of the table to search for the assay, the reference, or the sample.
5. In the well table, select the **Omit** checkbox to omit a well from the analysis.
Ensure that the well table is selected. Omitting data is not available in the sample table.
6. Click **Analyze** to analyze the data.

Overview of the result issues

Issue	Description
Analysis unit result issues	
Not Enough Sample Conf (<7)	There are less than 7 samples of the same copy number. Confidence metrics and Z-score metrics cannot be generated.
Inaccurate Positive Control (displayed at both the analysis unit level and sample level)	The defined positive control does not match the expected predicted copy number.
Invalid Positive Control (displayed at both the analysis unit level and sample level)	The defined positive control has a sample QC failure.
Calibrator Invalid	The specified sample calibrator has sample QC failure.
No Positive Control	The specified positive control sample is not found in the analysis unit or experiment. This issue is displayed only if a positive control is defined.
Sample result issues	
Sample Failed QC	The sample has one or more QC failures.
Poor Quality Calibrator	The sample calibrator that was specific is a sample of low quality.
CN > Threshold Before Bin	The predicted copy number is greater than the expected copy number threshold (20).
Inaccurate Positive Control (displayed at both the analysis unit level and sample level)	The defined positive control does not match the expected predicted copy number.
Invalid Positive Control (displayed at both the analysis unit level and sample level)	The defined positive control has a sample QC failure.
Calibrator Invalid (displayed at both the analysis unit level and sample level)	The specified sample calibrator has sample QC failure.
Well results issues	
Delta Ct Exceeds Threshold	The difference Cq(Ct) value between target of interest and the reference assay is greater than the specific threshold.
Outlier Exceeds Confidence	The outlier exceeds the confidence value.
Undetermined Non Zero Copy	There is an undetermined non-zero copy. This is associated with a replicate group where half of the replicates are zero-copy and half of the replicates have non-zero copy number values.
Reference Dye Exceeds Max Ct	The reference dye exceeds the maximum ΔCt value.
Outlier is zero copy	An outlier is defined as zero-copy.

Analysis settings with low accuracy

We do not recommend the following analysis settings. The combination of the following analysis settings can result in results with low accuracy.

- Auto calibration
- A copy number start value of 1
- A copy number end value greater than 1
- No two-copy positive control

See “Edit the general settings” on page 15 and “Use a two-copy control” on page 18.

Export the analysis units

The analysis unit file contains the following information:

- Run file name
- Barcode
- CNV assay name
- Reference name

The analysis unit can be exported from the **Single plate** tab or the **Multiplate** tab.

In the on-premise configuration, the location to save the file is defined in the export settings. For users with the permission of **Edit Export Destination**, the location can be selected for a file download. For users without the permission of **Edit Export Destination**, the location cannot be selected.

In the Thermo Fisher™ Connect Platform, the results are exported to the downloads folder of the computer. The file name and location cannot be selected.

For more information about setting up the export location, see the primary user guides for the software.

Select the analysis units to export (see “Select data files or CNV assays to review” on page 19).

1. In the top pane of the **Copy Number Variation** tab of an open project, click ⋮ (**More Options**) ▶ **Export**.

In the Thermo Fisher™ Connect Platform, the file is downloaded.

2. In the **Export CSV** dialog box, edit the file name in the **File Name** field.

The **File Name** field is populated with a default file name.

3. Click **Browse** to select a location to save the file.

The **Browse** button is not available in the on-premise configuration.

4. Click one of the following options.

- Click **Download**, then select a location for the file download. This option is available for the on-premise configuration, for users who have the permission of **Edit Export Destination**.
- Click **Save**. This option is available for the on-premise configuration, for users who do not have the permission of **Edit Export Destination**.
- Click **Export**. This option is available in the desktop configuration.

Export the results

In the on-premise configuration, the location to save the file is defined in the export settings. For users with the permission of **Edit Export Destination**, the location can be selected for a file download. For users without the permission of **Edit Export Destination**, the location cannot be selected.

For more information about setting up the export location, see the primary user guides for the software.

To export a project, see the primary user guides for the software (see “Related documentation” on page 45).

1. In the table pane of the **Copy Number Variation** tab, click **⋮ (More Options) ▶ Export**.
In the Thermo Fisher™ Connect Platform, the file is downloaded.
2. (Optional) In the **Export CSV** dialog box, edit the file name in the **File Name** field.
The **File Name** field is populated with a default file name.
3. Click **Browse** to select a location to save the file.
The **Browse** button is not available in the on-premise configuration.
4. Click one of the following options.
 - Click **Download**, then select a location for the file download. This option is available for the on-premise configuration, for users who have the permission of **Edit Export Destination**.
 - Click **Save**. This option is available for the on-premise configuration, for users who do not have the permission of **Edit Export Destination**.
 - Click **Export**. This option is available in the desktop configuration.



Overview of copy number analysis

Overview of the functions

The Copy Number Variation Analysis Module can perform the following functions:

- Convert data files set up with targets or SNP assays to CNV assays
- Perform copy number analyses with or without a known calibrator sample
- Report confidence in individual or grouped copy number calls
- Display the calculated sample copy number data in both table and graphic formats
- Display detailed data analysis information
- Change analysis parameters and reanalyze the data
- Analyze and view data from multiple experiments simultaneously
- Save or print analyzed data for further analysis or for sharing with other laboratories that use Diomni™ Design and Analysis (RUO) Software 3.

Overview of the analysis

The copy number assays are run simultaneously with reference assays in a duplex real-time PCR. Each copy number assay detects a target gene or genomic sequence of interest.

The target assays is used to detect and quantify the genomic region of interest. The reference assay is used to normalize the genomic DNA input.

The reference assay detects a sequence known to exist in two copies in a diploid genome, for example, the human RNase P H1 RNA gene. The experiments are run with absolute quantitation settings.

After amplification, the experiment results table contains C_t values for the copy number and reference assay for each well.

The Copy Number Variation Analysis Module performs a comparative C_t ($\Delta\Delta C_t$) relative quantitation analysis of the real-time data. The analysis determines the number of copies of the target sequence in each test genomic DNA sample. The comparative C_t ($\Delta\Delta C_t$) method first calculates the difference (ΔC_t) between the threshold cycles of the target and reference assay sequences. It then compares the ΔC_t values of the test samples to a calibrator sample that contains a known number of copies of the target sequence. The analysis can be performed without the use of a calibrator sample.

Overview of multiplate analysis

Multiple plates containing samples run with the same assay can be combined for copy number analysis.

We recommend single plate copy number analysis when possible. The highest quality copy number data are generated when measurement variation is very low. When data from multiple plates are combined for analysis, plate-to-plate experimental variation might be introduced and might impact the quality of the copy number analysis results.

Copy number quantitation requires high quality real-time PCR data for the best results. Measurement variation can be controlled by use of high quality samples, avoiding process errors, and accurately following the recommended procedure to achieve the highest confidence results. Generally, we highly recommend that all samples to be analyzed together are run on the same plate to avoid experiment variation introduced by plate-to-plate variation. However, in specific instances, such as high throughput analysis of many samples or assays, it may be desirable to combine data from multiple plates for copy number analysis. In some situations, combining data might improve the call rate for samples falling into copy number groups that are otherwise under-represented on single plates. The following information provides basic guidelines for multiplate analysis.

First, the data for each individual plate in your study should be analyzed and the data quality evaluated. If the data are high quality, multiplate analysis can be selected in the software.

Each plate in a study should contain enough data points for copy number analysis and evaluation of data quality evaluation. Only high quality plate data is suitable for multiplate analysis.

- Minimize plate-to-plate ΔC_t variation differences by running each copy number reaction plate on the same real-time PCR instrument using the same experimental process. For example, use high quality gDNA samples prepared by the same method. Use the same amount of gDNA, the same master mix lot, and, optimally, the same copy number and reference assay lots.
- Run enough samples on each plate to enable copy number analysis for each assay used. All sample replicates *must* be run on the *same* plate. Replicates run on different plates cannot be combined in multiplate analysis. Seven samples of the same copy number group must be present to calculate confidence, absolute z-score, and ΔC_t variability. If fewer samples are run, we recommend including a well-characterized calibrator sample for copy number analysis on each plate.
- Analyze each assay by single plate and review the copy number analysis results. The copy number variation assay ΔC_t subdistributions should be distinct, with clear boundaries between distributions. Confidence values should be high for most samples, and the ΔC_t variability should generally be ≤ 0.10 for each plate to be combined in multiplate analysis. If data from plates with high subdistribution and ΔC_t variability are combined, the copy number group ΔC_t distribution boundaries might be ill-defined and overlap between groups, resulting in poor quality copy number analysis results.

IMPORTANT! High quality samples and a highly reproducible experimental process are required to generate the high quality data with low ΔC_t subdistribution and ΔC_t variability for each plate to be combined for multiplate analysis. If those requirements cannot be met, use the copy number analysis results from single plate analysis.

Note: All plates combined in multiplate analysis must have the same target name (copy number assay) applied to each sample well for the plate data to be analyzed together.

Note: Samples having the same sample name on different plates are analyzed on a per-plate basis, so the ΔC_t values for wells having the same sample name are *not* combined between plates in multiplate analysis. If the same calibrator sample is run on multiple plates in the analysis and the calibrator sample name is the same on each plate, the ΔC_t value of the calibrator from the first selected assay plate is used as the calibrator for the multiplate analysis.

Calibrator sample

The Copy Number Variation Analysis Module can analyze copy number experiment data with or without a calibrator sample.

If the data are analyzed with a calibrator sample, the software calculates copy number values using the comparative C_t method of relative quantitation.

If the data are analyzed without a calibrator sample, the software calculates copy number values using a maximum-likelihood algorithm.

Overview of copy number bins

Copy number bins allow grouping of copy numbers into sets or bins. This helps to estimate the confidence in the predicted copy number calls.

A bin can contain a single copy number or a range of copy numbers. The numbers must be sequential. There cannot be gaps in the numerical values between the bins. The numbers cannot overlap.

If copy number bins are used, a |z-score| value is not provided because copy number bins are not normal distributions.

A maximum of 10 bins can be used.

For more information, see “About the quality metrics” on page 37.

Overview copy number sample graph

The graph displays a bar graph of the data in the **CN Calculated** column of the results table or the **CN Predicted** column of the results table. To select the graph type, see “Review the copy number vs sample graph” on page 20.

The software displays one bar for each analyzed sample.

Each bar displays an error bar that represents the minimum and maximum copy number calculated for the sample replicate group.

The graph does not display data for undetermined samples, where the copy numbers for the associated samples cannot be calculated. For example, a sample can be undetermined if it shows no amplification for the reference assay (a C_t greater than the threshold specified). The threshold is defined in the **Ref Cq threshold** field of the **Copy Number Variation Analysis Setting** dialog box. See “Edit the general settings” on page 15.

Overview of the copy number results

Column	Description
CN Calculated column	The copy number for the associated sample calculated by the relative quantitation method as: $cn_{sample} = RQ_{sample} \times cn_{calibrator}$ where RQ = relative quantity of the associated sample, and <i>cn</i> = copy number
CN Predicted column	The predicted copy number of the associated sample. A range is displayed if bins are enabled with a range of greater than one. See “Edit the advanced settings” on page 17.
Difference in Calculated and Predicted column	The mathematical calculation of the difference between the calculated copy number variation and the predicted copy number variation. The value is always presented as a positive number.
Result issues column	The issues that are determined based on the quality criteria. See “Edit the quality criteria” on page 17.
Confidence column	The confidence values of the associated predicted copy number. See “Confidence estimate” on page 37.
Min Conf column and Max Conf column	The confidence interval minimum and the confidence interval maximum, respectively. The values are represented as the error bars in the bar graph.
 Z-Score column	The absolute z-score value of the associated copy number integer. See “Deviation z-score estimate” on page 39.
Replicate count column and Replicates analyzed column	The number of technical replicates for the sample and the number of replicates included in the analysis. The two values differ when replicate well data are omitted from the analysis.
Cq Mean column	The arithmetic mean of the C_t calculated for the technical replicates of the sample.
Ref Cq Mean column	The arithmetic mean of the C_t calculated for the technical replicates of the sample.
ΔCq mean column	The arithmetic mean of the ΔC_t of the technical replicates for the sample.
Sigma (ΔCq) column	The standard deviation of the ΔC_t of the technical replicate population for the sample.
$\Delta \Delta Cq$ column	The $\Delta \Delta C_t$ value for the associated sample, which is calculated as: $(\Delta \Delta C_t)_{s,t} = \mu(\Delta C_t)_{s,t} - \mu(\Delta C_t)_{calibrator}$ where <i>s</i> = sample, and <i>t</i> = target copy number assay

(continued)

Column	Description
Rq column	The relative quantity for the associated sample, which is calculated as: $RQ_{(s,t)} = 2^{-(\Delta\Delta C_t)_{s,t}}$ where s = sample, and t = target copy number assay
Minimum copy number column and Maximum copy number column	The minimum copy number and the maximum copy number, respectively. The limits for the calculated copy numbers of the technical replicates for the sample.

Data analysis workflow

The Copy Number Variation Analysis Module calculates copy number assignments from the data of copy number experiments using the following series of steps. An explanation of the stages follows the diagram.

(1) Preprocessing—The software detects invalid data, removes outliers, and calculates values for independent variables such as ΔC_t values.

(2) Sample copy number assignment—The software calculates copy numbers using one of two algorithms. If a calibrator sample is present, the software applies a relative quantitation algorithm similar to the algorithm used by the real-time PCR system software (see “Copy number assignment with a calibrator sample” on page 36). If a calibrator sample is not present, the software applies an algorithm that calculates copy number assignments that maximize the likelihood of the observed data with respect to a theoretical model (see “Copy number assignment without a calibrator sample” on page 36). For more information, see “About the theoretical model” on page 31.

(3) Quality metric assignment—For each sample, the software calculates two quality metrics.

- **Confidence metric:** An estimate of the confidence that the assigned copy number is the true copy number relative to any of the other possible assigned copy numbers for the group of analyzed samples. For example, a confidence value of 90% indicates a 10% chance that the actual copy number differs from the called copy number.
- **Absolute z-score metric:** An estimate of how much the ΔC_t value of a sample deviates from the mean ΔC_t value of the assigned copy number category. For example, a copy number of 2 and an absolute z-score of 1 indicate that the ΔC_t for the sample is ± 1 standard deviation from the mean ΔC_t for copy number 2.

①

Data preprocessing

Start with imported data

Data validation

Exclude failed PCR data

ΔC_t calculations

Outlier detection

②

Sample copy number assignment with calibrator present

Calculate $\Delta\Delta C_t$

Calculate copy number with calibrator

Calculate quality metric

See “About the quality metrics” on page 37.

③

Sample copy number assignment without calibrator present

Estimate parameters

Calculate copy number without calibrator

Calculate quality metric

See “About the quality metrics” on page 37.

About the theoretical model

The analysis module determines copy numbers from real-time PCR data and related confidence measures using a theoretical model that adequately describes the data.

Biochemical basis of the model

The data generated by the copy number assay experiment is a ΔC_t value. The value is the difference in cycle threshold, C_t , values between a genomic target sequence of unknown copy number and an endogenous reference sequence of known and invariant copy number. A C_t value is associated with the PCR cycle at which the concentration of free emitter dye of the TaqMan™ probe is high enough that its fluorescence can be detected by the real-time PCR instrument.

Samples with more copies of the target sequence achieve C_t at lower cycle numbers because the PCR has more genetic material. The endogenous reference gene is a gene observed to have little, if any, variation in copy number. Therefore, if an unknown target C_t value is the same as the endogenous reference C_t , then the unknown target has the same copy number. The PCR process doubles the amount of genetic material at each cycle, so if the unknown C_t is one cycle higher, the unknown must have half the number of copies of the endogenous reference. If the unknown C_t is one cycle lower, it must have double the number of copies of the endogenous reference.

In practice, the relationship between C_t values and target copy number can be modified because the PCR process does not exactly double the concentration of genetic material at each cycle. The unknown and endogenous reference genes are monitored by different dyes that might be differentially detectable by the real-time PCR instrument. The sequence context of the unknown and endogenous reference genes is generally not the same, which might differentially influence the PCR process. Mutations in the unknown and endogenous reference genes might also influence the PCR process. The net result is that the actual relationship between C_t values and gene copy number deviates from the ideal described above. A statistical model that can account for this deviation is described in the next section.

About the statistical model

The exponential nature of the PCR process determines the main feature of the model: as the number of gene copies increases, the difference between successive ΔC_t values decreases. The following equation expresses this behavior:

$$\Delta C_t = K - \log_{1+E} cn$$

where K is a constant, E is the PCR efficiency of the assay of interest, and cn is copy number with the range $[1, \infty)$.

If the biochemical processes and instrumentation behave perfectly, then $K = 0$ and $E = 1$. In practice, however, $K \neq 0$ and $E < 1$. For more information about possible causes for this deviation from the ideal, see “Biochemical basis of the model” on page 31. There are many other causes, including:

- Thermal fluctuations
- Binding behaviors of PCR primers and TaqMan™ probes
- The starting concentration of a sample

The endogenous reference can normalize for the contribution of starting concentration effects, but might not completely account for all sources of variation. Therefore, even if all samples have the same copy number, the ΔC_t values can differ. To model this variation, it is assumed that ΔC_t values for samples of a given copy number are normally distributed around a mean that is calculated by the equation above. An additional simplifying assumption is that the variability of the normal distribution (σ^2) for a given copy number is the same as that for all other copy numbers. Different copy numbers have different probabilities of occurring within the sampled population, and these probabilities are determined by the characteristics of the genes in question. So the total probability covered by the normal distribution for a given copy number is modulated by the probability of that copy number within the sampled population. This completes the statistical model relating ΔC_t values to cn .

The following figures show an example model in which 20% of the samples have a $cn = 1$, 54% have a $cn = 2$, 13% have a $cn = 3$, and 13% are distributed among $cn > 3$. K is set at 1, E is set at 0.9, and σ is set at 0.1.

In Figure 1 on page 33, subdistributions for each copy number are displayed separately to make it easier to see the component distributions.

In Figure 2 on page 33, the same model is displayed with the subdistributions combined to show the actual probability density function.

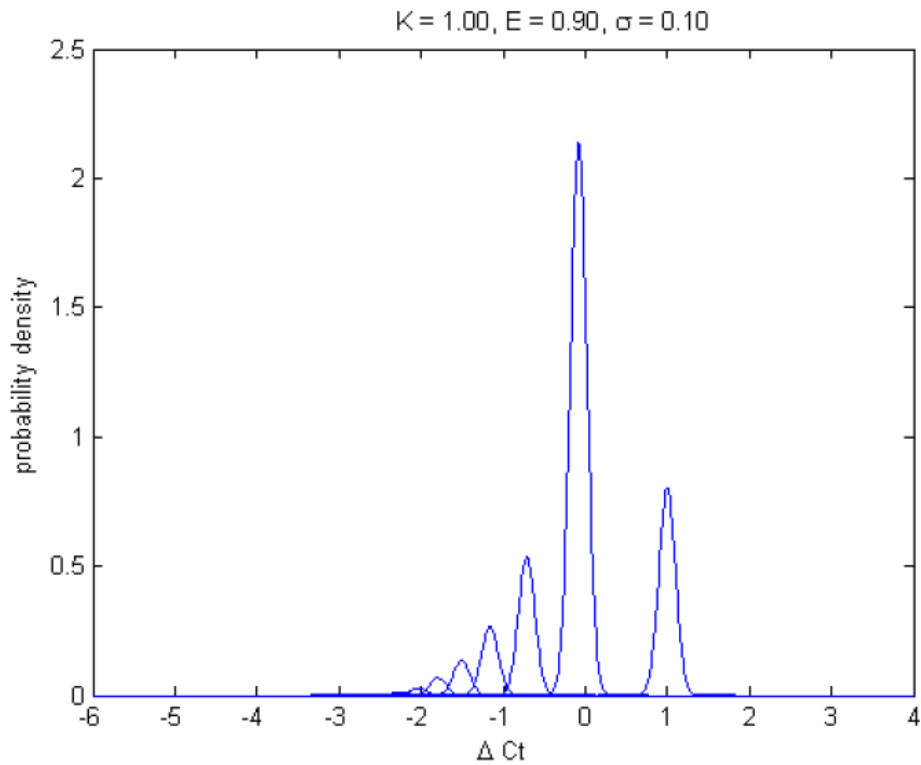


Figure 1 Separate subdistributions for each copy number

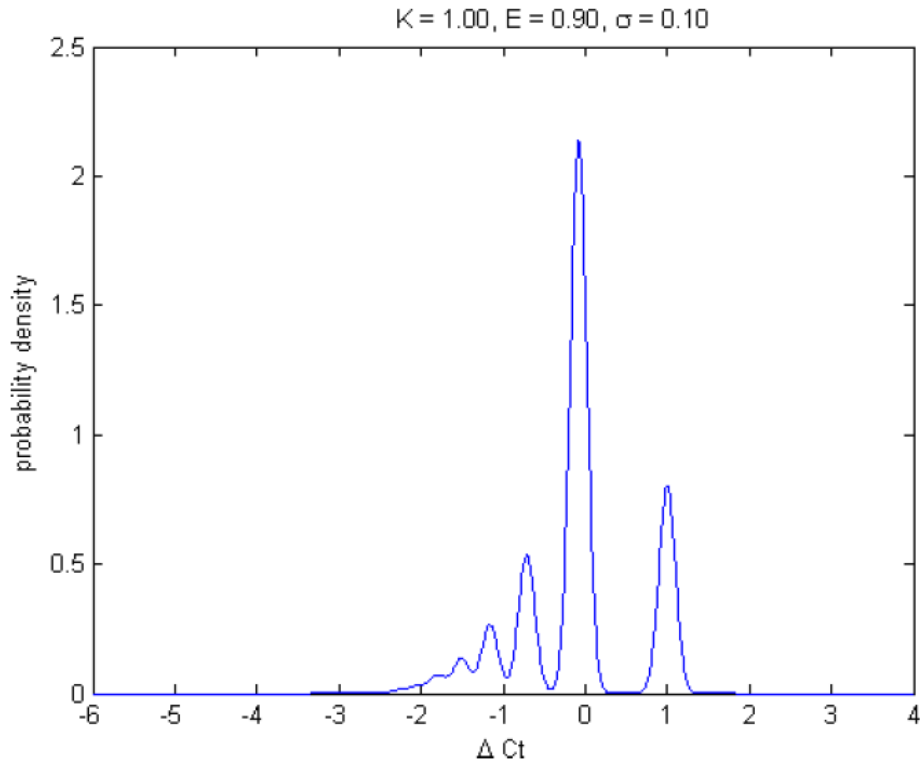


Figure 2 Combined subdistributions

Data preprocessing

Data validation

Before analyzing the data to assign copy number calls, the Copy Number Variation Analysis Module excludes wells in which the PCR reaction has failed. The software omits data from wells with Cq values greater than the threshold defined in the copy number variation analysis settings. The threshold is defined in the **Ref Cq threshold** field of the **Copy Number Variation Analysis Setting** dialog box. See “Edit the general settings” on page 15. If all replicates of a sample have Cq values greater than the threshold, the analysis module does not generate a copy number call for the associated sample.

Note: In general, wells with a C_T greater than 32 (default) are defined as having failed amplification, possibly due to an insufficient quantity of gDNA. The default value is 30 for an OpenArray™ Plate.

The analysis module analyzes zero-copy samples differently from nonzero-copy samples. Wells in which sufficient gDNA is present (where the well has a VIC™ Cq ≤ 32 , or ≤ 30 for an OpenArray™ Plate) but in which no FAM™ target is detected (FAM™ Cq is undetermined) might indicate either a failed copy number assays reaction or a zero-copy sample.

If all replicates of a particular sample have a VIC™ Cq ≤ 32 , or ≤ 30 for an OpenArray™ Plate, and undetermined FAM™ C_T , the sample is assumed to be a zero-copy sample.

Samples having extremely weak FAM™ signal relative to the VIC™ signal ($\Delta Cq > 4.0$) are also predicted to be zero-copy for the target sequence. The weak signal is assumed to be nonspecific.

ΔC_t calculations

After filtering the well data, the Copy Number Variation Analysis Module performs the following functions.

1. Calculates $(\Delta C_t)_w$, the well-level ΔC_t , for each target/reference combination in each well. The software begins by normalizing the C_t of each target assay in each well with the C_t of the corresponding reference assay in that well:

$$(\Delta C_t)_w = (C_t)_{t,w} - (C_t)_{r,w}$$

where: t = target assay; w = well; r = reference assay.

Note: This step helps negate variations in volume and pipetting inconsistencies in the experiment.

2. Calculates $\mu(\Delta C_t)$, the sample-level ΔC_t arithmetic mean of the replicates, for each target/reference combination in the analysis.

Note: This gives an estimate of the actual ΔC_t values and can help reduce experimental error.

For each replicate group of n wells representing a specific target/reference combination, a $\mu(\Delta C_t)$ value is calculated as:

$$\mu(\Delta C_T) = \frac{\sum_{w=1}^n (\Delta C_T)_w}{n}$$

Outlier detection

After calculating ΔC_t values, the Copy Number Variation Analysis Module removes from the analysis data outlier wells that produce ΔC_t values that differ significantly from the mean of all applicable wells, such as prefiltered wells). The differences in the ΔC_t values of outlier wells is commonly attributed to samples that amplify insufficiently or not at all due to experimental error such as pipetting inaccuracy or contamination.

To identify outliers among the replicate groups, the analysis module performs the following functions:

1. Estimates the plate level standard deviation.
2. Excludes data points greater than four standard deviations from the mean ΔC_t of the replicate group.

To estimate the standard deviation, the software excludes replicates that deviate from the replicate median by more than the distance between the one- and two-copy distributions, using a conservative value of 0.8 for PCR efficiency.

Note: We recommend against a more aggressive approach to outlier removal. The validity of the statistical analyses for copy number determination and sample quality metrics assume that the data are distributed as described in “About the theoretical model” on page 31.

Copy number assignment

The way the Copy Number Variation Analysis Module calculates copy number varies depending on the availability of a calibrator sample that has a known copy number for the target of interest.

Copy number assignment with a calibrator sample

When a calibrator sample that has a known copy number is available, the Copy Number Variation Analysis Module performs the following steps:

1. Calculates $\Delta\Delta C_t$ for each sample/target combination in the analysis.
The software calculates $\Delta\Delta C_t$ values for each target assay and each sample by subtracting the $\mu(\Delta C_t)$ of the calibrator sample from the corresponding $\mu(\Delta C_t)$ of each test sample (including the calibrator sample) for each target assay. This is the equation:
$$(\Delta\Delta C_t)_{s,t} = \mu(\Delta C_t)_{s,t} - \mu(\Delta C_t)_{c,t}$$
where: s = sample; t = target assay; c = calibrator.
2. Calculates copy number for each sample/target combination in the analysis.
When a calibrator sample with a known copy number is available, the analysis module calculates copy number by rounding the product of the following equation to the nearest integer:
$$\text{Copy number} = cn_c 2^{-\Delta\Delta C_t}$$
where cn_c is the copy number of the target sequence in the calibrator sample, and $\Delta\Delta C_t$ is the difference between ΔC_t for the unknown sample and the calibrator sample.

The calibrator sample is defined in the **Calibration** dropdown list of the **Copy Number Variation Analysis Setting** dialog box. See “Edit the general settings” on page 15.

Copy number assignment without a calibrator sample

When a calibrator sample is not specified for analysis, the Copy Number Variation Analysis Module determines the parameter values using the model described in “About the theoretical model” on page 31. The software uses the model to estimate K , which is an estimate of the mean ΔC_t for copy number 1. Using the calculated estimate, the software then employs the same copy number assignment described in “Copy number assignment with a calibrator sample” on page 36.

The software calculates the parameter values for the model as follows:

- Efficiency is assumed to be 100%.
- K is determined by maximizing the likelihood of the observed ΔC_t replicate means as a function of this parameter. The likelihood of the observations is taken to be the sum of probability density across all samples for a given model.

For each candidate value for K , the following initial model is assumed, then refined:

- Probability of a copy number—Calculated as the range of copy numbers with non-zero probability, where the limits are from 1 copy to the copy number that could be assigned to the sample with the lowest value for the replicate mean ΔC_t , given the candidate value for K . Initially, the software assumes that all copy numbers with non-zero probability have the same probability.
- Standard deviation of the subdistributions—The software assigns an initial value.

Note: When using the “Without calibrator sample” likelihood analysis method, the asymmetry of the statistical model around the mean ΔC_t value for each copy number might cause the predicted copy number value for a sample to be higher or lower than expected if the calculated copy number was rounded up. For example, a calculated value of 2.4 might correspond to a predicted value of 3.

The analysis module uses the initial model to calculate the proximity of the samples to the means of each copy number subdistribution. The proximity measure for a given sample is the probability density at the replicate mean ΔC_t value divided by the total probability density for the sample. Using this proximity measure, the software updates the model as follows:

- Probability of a copy number—The software calculates the sum of the proximity measures over samples as an estimate of the number of samples associated with a given copy number. The software then divides the sum by the total number of samples to calculate a revised estimate of the copy number probability.
- Standard deviation of the subdistributions—Using the proximity measures as the weightings, the software calculates a weighted sum of squared deviations between the replicate means and the means of the copy number sub-distributions. The software calculates the square root of the average value of the weighted sum over the samples as a revised estimate of the standard deviation of the subdistributions.

The software applies the revised model to calculate likelihood for a given value of K .

About the quality metrics

The Copy Number Variation Analysis Module generates two quality metrics for each analyzed sample:

- A confidence estimate
- A deviation estimate

The analysis module also generates a plate-level ΔC_t variability estimate.

Confidence estimate

The confidence estimate for a given sample is the probability that the calculated copy number is the correct assignment compared to other copy numbers that have nonzero probability of occurring. The Copy Number Variation Analysis Module uses the copy number and observed ΔC_t values of a sample to calculate the confidence estimate for the sample relative to the theoretical model. This is described in “About the theoretical model” on page 31.

The confidence estimate model is identical to that used for estimating copy number, except:

- Amplification efficiency is not assumed to be 100%.
- Parameter values are estimated differently.

The analysis module calculates the model parameters as follows:

- *K* and *E*—It uses the two most frequently assigned copy numbers and the equation describing means of the copy number subdistributions as a function of copy number. (See “About the theoretical model” on page 31 for details.) The analysis module forms two linearly independent equations that can be solved for *K* and *E*. The maximum likelihood estimate for the means of the copy number subdistributions completes the estimate for the parameters. When all samples are assigned the same copy number, *E* is assumed to be 1.
- Standard deviation of the subdistributions—It uses *K*, *E*, and copy numbers to estimate means of the copy-number subdistributions. The analysis module calculates the maximum likelihood estimate for this parameter for all samples with nonzero assigned copy numbers. The calculation reduces the degrees of freedom by two because *K* and *E* are estimated from the data.
- Probability of a copy number—For each nonzero copy number, the analysis module estimates the probability by dividing the number of samples assigned the copy number by the total number of samples. It modifies the estimate slightly by assigning a small amount of probability to copy numbers for which there are no samples, and then normalizing the total probability to 1.

The estimate of confidence as a function of assigned copy number and replicate mean is calculated as:

$$Confidence(\mu_r, cn_{assigned}) = \left[1 + \sum_{cn \neq cn_a} \frac{\Pi_{cn}}{\Pi_{cn_a}} e^{-\Omega} \right]^{-1}$$

where a = assigned, μ_r = replicate mean for the sample, $cn_{assigned}$ = copy number assigned to the sample, Π_{cn} = probability of copy number cn , Ω is calculated as:

$$\Omega = \frac{1}{\sigma^2 \log(1 + E)} \log\left(\frac{cn}{cn_a}\right) \left((\hat{\mu}_r - K) + \frac{\log(cn_a cn)}{2 \log(1 + E)} \right)$$

where σ = standard deviation of the subdistributions, *E* = PCR efficiency of the target assay, *K* = constant in the function relating the subdistribution mean (μ_{cn}) to copy number (cn) calculated as:

$$\mu_{cn} = K - \frac{1}{\log(1 + E)} \log(cn)$$

Using bootstrap techniques, the analysis module calculates an estimate of a 5% lower bound for confidence. For example, the analysis module estimates the value for which there is only a 5% chance that the actual value for confidence is lower than that reported. A consequence of this approach is that there will be some variation in the values reported.

Deviation |z-score| estimate

The deviation estimate is the absolute z-score, that is, the absolute value of the number of standard deviations separating the replicate mean ΔC_t , μ_r of a sample from the mean subdistribution, μ_{cn_r} , of the assigned copy number.

The deviation estimate uses the model estimated for the confidence metric and is calculated as:

$$absoluteZscore(\mu_r, cn_{assigned}) = \left| \frac{\mu_r - \mu_{cn_a}}{\sigma} \right|$$

Using bootstrap techniques, the analysis module calculates an estimate of a 5% lower bound for the deviation. A consequence of this approach is that there will be some variation in the values reported.

Note: If copy number bins are used to estimate confidence values, |z-score| estimates cannot be calculated because copy number bins are not normal distributions.

ΔC_t variability estimate

The plate-level ΔC_t variability is a standard deviation value. It is an estimate of the variability in the ΔC_t measurements for an assay across all samples in an analysis. It is derived by fitting the sample ΔC_t values to a model that is essentially a superimposition of the copy number group normal distributions with their centers at the average ΔC_t of each copy number group. The variability is assumed to be the same in all copy number groups—that is, the set of normal distributions has the same standard deviation. This standard deviation value estimated for the model is reported as the estimate for the plate-level variability.

The following figures show the distribution of plate-level variability seen in copy-number assay analyses. Figure 3 on page 40 shows a sample population of copy numbers 1 and 2 only.

Figure 4 on page 41 shows a sample population with copy numbers 1, 2, 3 and above.

The line denotes the mean standard deviation for these normal distributions (0.10).

In general, ΔC_t variability values below this limit might be considered to have low plate ΔC_t measurement variability.

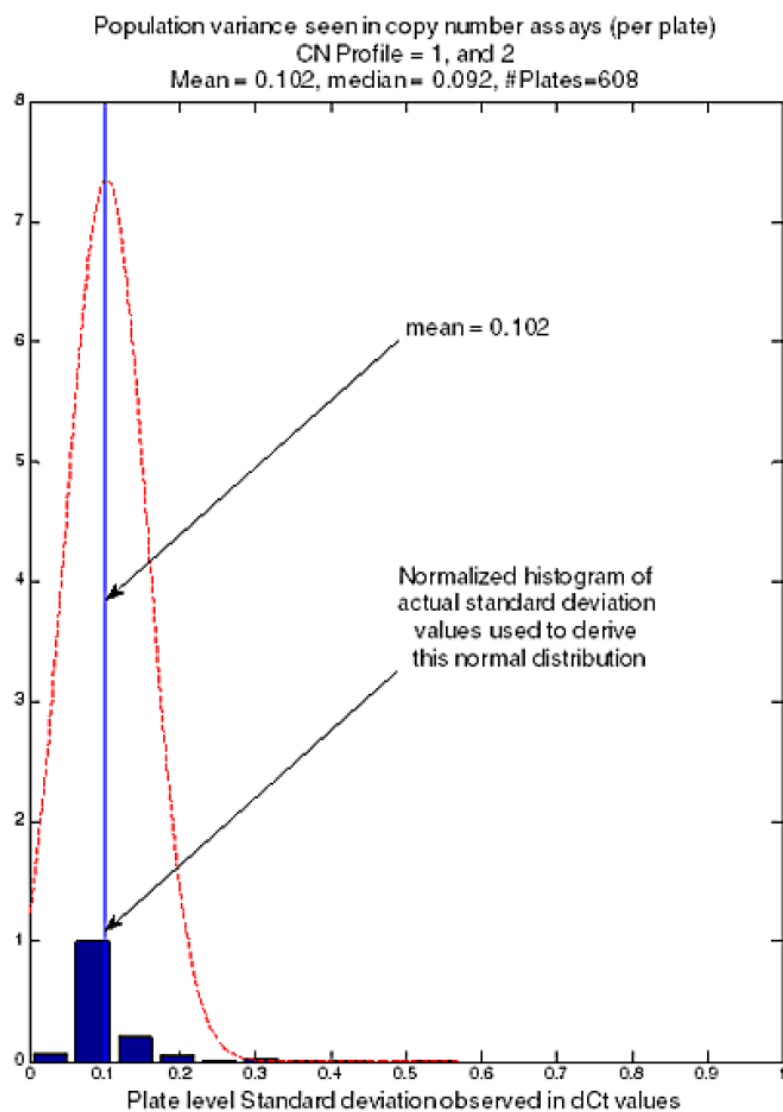


Figure 3 Sample population of copy numbers 1 and 2

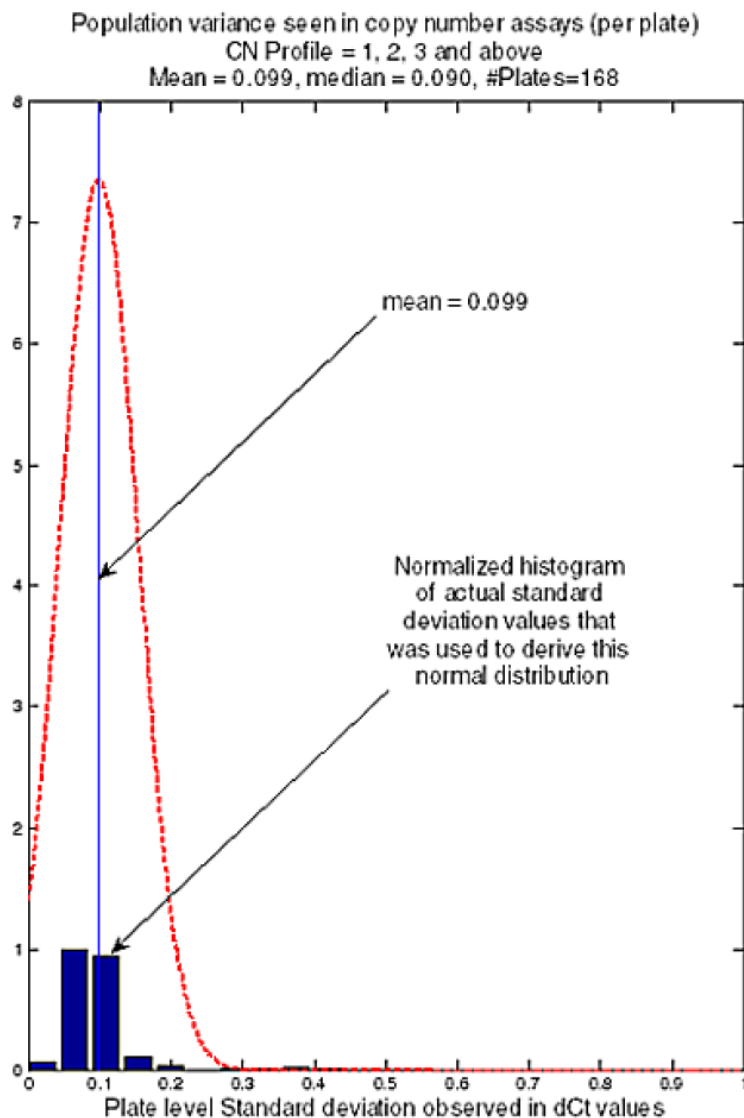


Figure 4 Sample population of copy numbers 1, 2, 3, and above

For multi-plate analysis, plate ΔC_t variability values can aid evaluation of whether or not plate data should be combined for copy-number analysis. If the plate ΔC_t variability is higher than the empirically derived threshold limit of 0.10 for an assay, combining data from different plates might not improve the sample copy number call rate. This is because the additional plate-to-plate variation might lead to spreading of the data points and to loss or overlap of the copy number group subdistributions.

Additional background information

Recommended minimum data requirements

The table below illustrates the relationship between the number of data points used in an analysis and the accuracy of the analysis. The table shows the accuracy of estimates of the mean and standard deviation of a normal random variate as a function of the number of points used to make the estimate. For estimates of the mean and standard deviation, the table shows “half point” percent deviations from actual values as a function of the number of points used to make the estimates.

Note: The term “half point” describes the deviation at which half the estimates would have lower deviations from actual values and half would have higher deviations. For the mean, the percentage is relative to the standard deviation of the underlying distribution.

The values in the table show that, for the number of replicates or samples typically used in copy number studies, estimates of mean and variability deviate considerably from actual values. Note that the gains in accuracy of the estimated mean and standard deviation decrease rapidly as the number of data points increases.

Number of points used in estimation	Deviation between estimate and actual (%)	
	Mean ^[1]	Standard deviation ^[2]
2	57.74	36.42
3	44.16	28.98
4	37.03	24.78
5	32.5	21.99
6	29.29	19.97
7	26.88	18.42
8	24.97	17.19
9	23.42	16.17
10	22.13	15.31
11	21.03	14.58
12	20.08	13.94
23	14.29	10.01
24	13.98	9.80
46	10.02	7.06
47	9.92	6.98
91	7.1	5.01

(continued)

Number of points used in estimation	Deviation between estimate and actual (%)	
	Mean ^[1]	Standard deviation ^[2]
92	7.06	4.98
183	5.00	3.53

^[1] Percent deviation from the actual mean (relative to the standard deviation of the distribution) exceeds the listed value for 50% of the estimates.

^[2] Percent deviation from the actual standard deviation exceeds the listed value for 50% of the estimates.

The number of replicates

To select the number of replicates to use in a copy-number study, one method uses the number at which the benefit of adding one more replicate falls below a target percentage. For example, if you choose 5% as the minimum percentage of improvement, the optimum number of replicates is:

- Four data points to estimate standard deviation. (The difference in the estimate of standard deviation between 4 and 5 replicates is $37.03 - 32.5 = 4.53$.)
or
- Three data points to estimate the mean. (The difference in the estimate of the mean between 3 and 4 replicates is $28.98 - 24.78 = 4.2$.)

In the example above, because the copy-number problem involves estimating many types of parameters, using four replicates might be preferable.

For the copy-number problem, the replicate mean for each sample must be estimated. Also, four more parameters must be estimated from the data when there are two unique copy numbers among the samples. One parameter must be added for each additional unique copy number beyond the first two. In the example above, where four replicates are used per parameter, four replicates are needed to estimate the replicate means, and $4 \times (2 + N_{cn})$ samples are needed to estimate the other parameters, where N_{cn} is the number of unique copy numbers among the samples.

Interaction of copy number and the confidence metric

The tables below show that for the theoretical models described in “About the theoretical model” on page 31, the ability to distinguish between adjacent copy numbers decreases as copy number increases. Consequently, the confidence metric is generally lower at high copy numbers. The interaction of several parameters in the model influences the severity of the effect. For this reason, several models, selected by examining the variation in data over a database of 282 assays, are used to show the effect.

The table below lists models used to show the relationship between copy number and confidence metric. For this table, standard deviations of ΔC_t values were obtained from an analysis of 282 assays, each applied to 91 samples.

Model ^[1]	Copy number subdistribution standard deviation
Optimal	0.05
Typical	0.09
Poor	0.2

^[1] Rated by ability to resolve copy numbers.

The table below shows the confidence metric value at which 95% of the probability for a given copy number is captured. Consequently, 95% of the samples with the given copy number should have a confidence value of that shown in the table or greater.

Model	Copy number									
	1	2	3	4	5	6	7	8	9	10
Optimal	1.0000	1.0000	1.0000	1.0000	0.9998	0.9903	0.8976	0.6619	0.4324	0.2934
Typical	1.0000	1.0000	0.9916	0.7468	0.2833	0.1938	0.1477	0.1299	0.1196	0.0856
Poor	0.9570	0.6251	0.1092	0.0897	0.0769	0.0760	0.0747	0.0682	0.0527	0.0306



Documentation and support

Related documentation

Document	Publication number
<i>Diomni™ Design and Analysis (RUO) Software 3 (Desktop) User Guide</i>	MAN0030162
<i>Diomni™ Design and Analysis (RUO) 3 User Guide (Thermo Fisher™ Connect Platform)</i>	MAN0030163
<i>Diomni™ Design and Analysis (RUO) Software 3 (On-Premise) User Guide</i>	MAN1000091

Customer and technical support

Visit [thermofisher.com/support](https://www.thermofisher.com/support) for the latest service and support information.

- Worldwide contact telephone numbers
- Product support information
 - Product FAQs
 - Software, patches, and updates
 - Training for many applications and instruments
- Order and web support
- Product documentation
 - User guides, manuals, and protocols
 - Certificates of Analysis
 - Safety Data Sheets (SDSs; also known as MSDSs)

Note: For SDSs for reagents and chemicals from other manufacturers, contact the manufacturer.

Limited product warranty

Life Technologies Corporation and its affiliates warrant their products as set forth in the Life Technologies' General Terms and Conditions of Sale at www.thermofisher.com/us/en/home/global/terms-and-conditions.html. If you have questions, contact Life Technologies at www.thermofisher.com/support.

