



Coffalyser digitalMLPA User Manual

for use with digitalMLPA products

Instructions For Use

Coffalyser digitalMLPA analysis software is for Research Use Only (RUO).
Not for use in diagnostic procedures.

Intended purpose

Coffalyser digitalMLPA is Research Use Only (RUO) software designed for the analysis of digitalMLPA data generated using a SALSA® digitalMLPA probemix* as described in the digitalMLPA General Protocol.

* To be used in combination with a SALSA® digitalMLPA Reagent Kit and SALSA® Barcode Solutions.



Manufacturer: MRC Holland bv, Willem Schoutenstraat 1, 1057 DL Amsterdam, The Netherlands

Website: www.mrcholland.com; www.mrcholland.eu; Phone: +31 888 657 200

E-mail: info@mrcholland.com (information & technical questions); order@mrcholland.com (for orders).

TABLE OF CONTENTS

IMPORTANT NOTIFICATION	3
1. GLOSSARY.....	4
2. BASIC CONCEPTS OF COFFALYSER digitalMLPA	6
2.1 SYSTEM REQUIREMENTS	6
3. GETTING STARTED.....	7
3.1 OBTAIN A COPY OF COFFALYSER digitalMLPA.....	7
3.2 ANALYSE YOUR DATA.....	7
3.2.1 CREATE A NEW EXPERIMENT.....	8
3.2.2 REANALYSE EXISTING EXPERIMENTS.....	10
4. COFFALYSER digitalMLPA REPORTS.....	13
4.1 [EXPERIMENT NAME] – OVERVIEW (PDF REPORT).....	13
4.2 [EXPERIMENT NAME] – RATIOS (EXCEL FILE).....	15
4.3 [BARCODE – SAMPLE NAME] – SAMPLE RESULTS 1 (PDF REPORT).....	17
4.3.1 DATES.....	17
4.3.2 SAMPLE QUALITY.....	17
4.3.3 GENERAL INFORMATION - SAMPLE.....	17
4.3.4 GENERAL INFORMATION – EXPERIMENT SETTINGS	18
4.3.5 GENERAL INFORMATION – ANALYSIS DETAILS.....	18
4.3.6 GENERAL INFORMATION – READS.....	18
4.3.7 GENERAL INFORMATION – PROBES.....	19
4.3.8 RESULTS SUMMARY – RATIO GRAPHS.....	19
4.3.9 RESULTS SUMMARY – MUTATION-SPECIFIC PROBES.....	20
4.3.10 RESULTS SUMMARY – ABERRANT RESULTS TABLE	20
4.3.11 QUALITY DETAILS – QUALITY TABLE	22
4.4 [BARCODE – SAMPLE NAME] – GENERAL DETAILS (PDF REPORT).....	24
4.4.1 ANALYSIS SETTINGS.....	24
4.4.2 ERRORS.....	24
5. ANALYSIS OF RESULTS.....	25
5.1 DISPLAY OF RESULTS	25
5.1.1 COMPARISON TO ARBITRARY BORDERS.....	25
5.1.2 COMPARISON TO REFERENCE POPULATION.....	25
5.2 INTERPRETATION.....	25
APPENDIX I – NORMALISATION	27
STEP 1.....	27
STEP 2.....	27
NO REFERENCE SAMPLES DEFINED IN ANALYSIS.....	27
APPENDIX II – QUALITY TABLE.....	28
APPENDIX III – TROUBLESHOOTING	31
GENERAL.....	31
EXPERIMENT DEFINITION	31
DATA ANALYSIS (COFFALYSER digitalMLPA – RUN WIZARD)	32
SAMPLE RESULTS 1 REPORT.....	33
GENERAL DETAILS REPORT	34
APPENDIX IV – COMMAND LINE INTERFACE (CLI).....	36
CLI ANALYSIS WITH COFFALYSER DEFINITION FILE	36
COFFALYSER DEFINITION FILE HEADER.....	36
COFFALYSER DEFINITION FILE EXPERIMENT DEFINITION	37
RUNNING COFFALYSER DEFINITION FILE IN CLI.....	38
CLI ANALYSIS WITHOUT A COFFALYSER DEFINITION FILE.....	38

IMPORTANT NOTIFICATION

Probe-specific information concerning certain genetic characteristics (e.g. whether a probe is located in a promoter region) and experimental sensitivity (e.g. sensitivity to salt) are not included in the Coffalyser digitalMLPA reports. When interpreting the data, it is therefore essential to consult the relevant product files provided by MRC Holland: the Product Description, and Probe Information File (and other product-specific files, if available).

The Coffalyser digitalMLPA software has an expiration date. This date is displayed in the bottom of the initial Coffalyser digitalMLPA window when opened. A new version will be made available by MRC Holland before the expiration date has passed. Changes between versions are described in the Coffalyser digitalMLPA Release Notes, present in the “_Documentation” folder in the downloaded software package.

In case of alerts, warnings and important information on updates that should be implemented immediately, MRC Holland will contact you via email. It is therefore essential for us to have your current contact information. Your contact details can be edited in your MRC Holland account (www.mrcholland.com).

This manual is valid for program versions 2020-05-12 and newer (software version 200609.0921 and newer). Software version is displayed in the Coffalyser digitalMLPA Release Notes present in the “_Documentation” folder in the downloaded software package.

1. GLOSSARY

Barcode	A Barcode Solutions plate is a microtiter plate with 96 different barcode solutions. A barcode solution is a digitalMLPA reagent that is added in the first step of digitalMLPA and that should be unique for each sample within a sequencing run. Each barcode solution contains an oligonucleotide (barcode oligo) that is incorporated in each digitalMLPA amplicon in order to distinguish digitalMLPA data from other sequencing data and to assign amplicons to the corresponding sample. The unique sequence part of the barcode oligo is called the barcode.
Coffalyser Definition File	A file (*.cdf) containing the information required for the analysis of an experiment with Coffalyser digitalMLPA. This includes analysis settings, file locations, sample sheets and other data such as sample names, sample types and genders.
Coffa file	A file (*.coffa) containing the sequencing output and several analysis properties of a single digitalMLPA reaction. It is generated from a FASTQ file by Coffalyser digitalMLPA.
FASTQ	A file containing the sequencing data obtained on an Illumina Next Generation Sequencer.
Intra ratio	Ratio obtained after the first step of normalisation (intra-normalisation) in which read counts are converted to relative values by normalising target probe read counts against the read counts of reference probes within a sample. See APPENDIX I – NORMALISATION for more information.
Inter ratio	Ratio obtained after the second step of normalisation (inter-normalisation) in which relative probe values of each sample are compared to those of the reference samples. See APPENDIX I – NORMALISATION for more information.
No DNA sample	A sample from a No DNA control reaction.
Positive reference sample	A sample with a known mutation and/or CNV.
Product Sheet	A file in the Coffalyser digitalMLPA package that contains information for the software on how to analyse the data obtained using a specific digitalMLPA probemix.
Reference probe	A probe that detects a sequence that is expected to have a normal copy number in (almost) all samples. Reference probes are used in the first part of data normalisation to normalise other probes against.
Reference sample	A sample from a healthy individual that is used to normalise other samples against.
Sample Track Probe	NOT YET AVAILABLE. An oligo that is added to a sample before or during DNA extraction. The resulting probe is recognised by Coffalyser digitalMLPA thereby aiding in sample identification and potential sample swaps.

SD	Sample DNA (SD) is a DNA sample with a specific genotype provided by MRC Holland for a specific application, e.g. an artificial mutation-positive DNA sample.
SNP code	<p>Most digitalMLPA probemixes contain a set of SNP-specific probes that generate a SNP code, also referred to as a <i>sample identifier</i> since DNA samples from the same individual will have the same identifier. This SNP code consists of 39 characters each corresponding to one specific SNP that are always in the same order. Every position can have four different options: 0, 1, 2 or ?, where 0 and 2 are SNPs that are homozygous for allele A or B respectively, while 1 represents a heterozygous SNP. If one of these options cannot be determined, this results in a "?". The presence of multiple "?" signs can indicate sample DNA contamination. This will be mentioned in the quality sample overview.</p> <p>The SNP code ends with a version number, e.g. v1 for version 1. More information on these SNP probes is described in the digitalMLPA general protocol.</p>
Test sample	The sample being investigated.

2. BASIC CONCEPTS OF COFFALYSER digitalMLPA

digitalMLPA products are currently only available for use with Illumina sequencers. These sequencers generate one or more large FASTQ files.

The digitalMLPA analysis flow consists of the following steps:

- **Data conversion:** Reads from a FASTQ file that are recognised as digitalMLPA reads are counted and assigned to the digitalMLPA probes. These reads are saved per digitalMLPA reaction in Coffa files (*.coffa, one file per digitalMLPA barcode). These files are typically less than 1 MB in size.
- **Analysis:** The obtained digitalMLPA probe read counts are normalised, see APPENDIX I – NORMALISATION for details, and evaluated for quality and the presence of aberrant ratios.
- **Reporting:** Reports are generated per experiment (Overview.pdf and Ratios.xls) and per sample (General Details.pdf, Sample Results 1.pdf and/or No Dna.pdf).

The Coffalyser digitalMLPA program requires a Coffalyser Definition File (*.cdf) to perform data analysis. This file specifies analysis settings, file locations, product sheets and other data such as sample names, sample types and gender. This file will be created when analysing your data using MRC.DataAnalysis.DigitalMlpa.Client.exe. In addition, the Coffalyser digitalMLPA software contains a Definition File Editor (MRC.DataAnalysis.DigitalMlpa.Definitions.Editor.exe) for easy creation and editing of Coffalyser Definition Files in case you want to prepare a file but not yet perform the analysis. A standard text editor such as Notepad can also be used to edit Coffalyser Definition Files.

It is possible to generate Coffa files as a separate step via a command line interface. An NGS service provider who pools Illumina sequencer experiments from several customers in a single Illumina run can perform the data conversion to Coffa files and provide each customer with the Coffa files of their experiment. Coffa files contain all relevant information for data analysis. Each customer can subsequently run the analysis and reporting steps on their own computer with Coffalyser digitalMLPA. For detailed information on command line interface, see APPENDIX IV – COMMAND LINE INTERFACE (CLI).

2.1 SYSTEM REQUIREMENTS

Coffalyser digitalMLPA does not require installation. For successful operation of the software, the following requirements should be met:

- A genuine copy of Windows 10 (32- or 64-bit) with the latest updates installed.
- An up-to-date .NET Framework (at least version 4.7).
- An Intel i5 processor or equivalent.
- 4 GB of available RAM.
- Sufficient free disk space with a minimum of the size of the FASTQ files to be processed.

For optimal operation, we recommend:

- An Intel i7 processor or equivalent.
- 8 GB of available RAM.
- A solid state drive (SSD) with sufficient free disk space.

Note: (1) The use of a USB stick to store files may greatly reduce the speed of the analysis. (2) An internet connection is not required to run Coffalyser digitalMLPA.

3. GETTING STARTED

3.1 OBTAIN A COPY OF COFFALYSER digitalMLPA

1. Log in to your MRC Holland account on www.mrcholland.com.
2. Navigate to software in the account menu and select Coffalyser digitalMLPA.
3. Accept the End User License Agreement (EULA) and download the Coffalyser digitalMLPA – software package.
4. Save the package on the desired location.
5. Unzip the file.
6. Optional: Create a shortcut for the MRC.DataAnalysis.DigitalMlpa.Client and place it on the desktop.



Notes:

- You may unpack and run Coffalyser digitalMLPA from a local hard-disk location or from a network location. A network location allows multiple users to use the same version, and central version management.
- Please note that network security settings or virus scanners may prevent Coffalyser digitalMLPA from running from network locations.

3.2 ANALYSE YOUR DATA

The time required for analysis of FASTQ files is highly dependent on the location of the FASTQ file and the location chosen for the generated Coffa files and the temporary files.

Coffalyser digitalMLPA can perform an analysis using zipped FASTQ files. Unzipping FASTQ files before analysis or copying a FASTQ file to a local disk is not necessary.

3.2.1 CREATE A NEW EXPERIMENT

1. Double click on MRC.DataAnalysis.DigitalMlpa.Client.exe or the shortcut you created to start Coffalyser digitalMLPA.



2. Under FastQ file(s) select *Create new definition file* and click **Next >**.

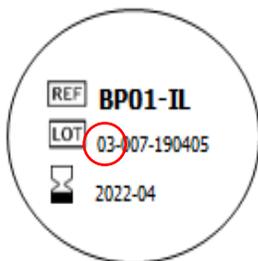


3. Navigate to the location where you want to save the Coffalyser Definition File (*.cdf).
4. Enter a name in the field *File name:* and click **Save**.

 The Coffalyser digitalMLPA - Data Analysis – General Settings window will open.

5. Leave the analysis settings on *Default profile*.
6. Select the appropriate barcode lot from the drop-down menu.

The first two digits of the barcode lot number indicate the barcode version and should match between the software and the Barcode Solutions plate. You can find the lot number on the label of the Barcode Solutions plate.



The “Illumina standard” setting should be used when reads are generated with the universal Illumina Rd1 primer and are ≥ 110 nt in length, either as single reads, or as the first read of paired end reads.

The “Illumina demultiplexed FastQ” setting should be used when sequencing reads are generated with a NextSeq 500/550, where a custom read 1 primer is used and a separate index read is generated resulting in demultiplexed FASTQ files. Email info@mrcholland.com for more information.

7. Click **Configure** to set up your experiment^a.

 The Coffalyser digitalMLPA – Samples window will open.

8. Select all barcodes and/or well positions on the Barcode Solutions plate that were used in your experiment.
9. Right click on any of the selected barcodes and select *Create New Experiment (Manual)*.

^aIn case an experiment is not (completely) configured, Coffalyser digitalMLPA will run using the default settings. Please be aware that these may not be correct for your experiment! MRC Holland recommends to always complete the experiment details to increase the likelihood of detecting mistakes.

 The Coffalyser digitalMLPA – Experiment Overview window will open.

10. Enter the experiment name in the designated text field.
11. Select the experiment type from the drop-down menu.
12. Select the digitalMLPA probemix used from the *product* drop-down menu.
13. Select the lot number of the digitalMLPA probemix used from the *sheet* drop-down menu.
14. Enter the name for each sample in the column *sample* or leave it as the barcode ID.

You can copy a column of sample names from another file. Right click on the first row of the range of samples and select Paste Sample Details From Clipboard.

15. Select the sample type for each sample from the drop-down menu in the column *type*.

You can also do this batch-wise: select samples while holding the Shift and/or Ctrl key and select the sample type from the drop-down menu of the last selected sample.

The various sample type options are explained in the Glossary. If sample type is left Undefined, it will be analysed as Test sample. In case no samples are indicated as Reference, all samples with type Test and Undefined will be used as Reference. For correct experiment set-up, please check the digitalMLPA General Protocol and the probemix-specific Product Description.

16. For samples that contain a mixture of DNA samples, select Pooled DNA source from the drop-down menu in the column *options*. For samples of a single source, leave this on *default*.
17. Optional: select the gender of each sample from the drop-down menu in the column *gender*.

You can also do this batch-wise: select samples while holding the Shift and/or Ctrl key, then select the gender from the drop-down menu of the last selected sample.

-  IF YOU ADDED SAMPLE TRACK PROBES^b TO YOUR SAMPLES, CONTINUE WITH STEP 18. OTHERWISE PROCEED WITH STEP 21.

18. Click in the column *tracker(s)*.

 The Coffalyser digitalMLPA – Select Sample Tracking Probes window will open.

19. Select the Sample Track Probe used and click **OK**.
20. Repeat steps 18 and 19 for other samples.
21. Click **OK** to save and exit the *Coffalyser digitalMLPA – Experiment Overview* window.

-  TO SET UP ANOTHER EXPERIMENT BEFORE STARTING THE ANALYSIS, FOLLOW THE PROCEDURE FROM STEP 8 ONWARDS.

22. Click **OK** to save the experiment(s) and close the *Coffalyser digitalMLPA – Samples* window.

^b Sample Track Probes are not yet available.

23. To select the output directory click the upper **Select** button.

24. Browse to the desired folder, select it and click **OK**.

For each run, Coffalyser digitalMLPA will create an output sub-directory in your selected location, named using the date and time of the analysis. For example: "C:\Coffalyser digitalMLPA Analysis Run01\20200322_110143".

Please note that the parent folder should not contain more than 100 subfolders.

25. Optional: click the lower **Select** button to change the temp directory.

The temp directory is by default the same as the output directory selected in the previous step.

Note: the speed of the analysis is influenced by the location of the temp directory. Preferably, the temp directory should be located on a local disk. A network drive is not recommended for this purpose.

26. Click **Add File**.

27. Browse to the location of the FASTQ files, select it and click **Open**.

28. Click **OK** to start data analysis.

 A progress bar will appear in the Coffalyser digitalMLPA – Run Wizard.

29. Click **Open Results Folder** to access the experiment subfolder in the output directory.

In this folder a subfolder is created for each experiment that was defined. This contains the Coffa files and all generated reports.

In addition, a subfolder called Undefined Experiment might have been created, which contains files for all found barcodes that were not defined in your experiment(s).

3.2.2 REANALYSE EXISTING EXPERIMENTS

Coffalyser digitalMLPA allows you to reanalyse existing experiments. This can be done by using an existing Coffalyser Definition File (*.cdf) and FASTQ files as input, or by using Coffa files (*.coffa). Reanalysis using Coffa files is much faster, but cannot be used in the following cases:

- The first analysis was done with an incorrectly selected probemix or lot number.
- A barcode used in the experiment was not selected in the initial set up.
- You want to change the output directory and/or temp directory.

When the name of the experiment has not been changed for the reanalysis, the results of the first analysis will be placed in a zipped folder in the experiment subfolder in the output directory.

3.2.2.1 REANALYSE AN EXPERIMENT USING EXISTING COFFALYSER DEFINITION FILE

1. Double click on MRC.DataAnalysis.DigitalMlpa.Client.exe or the shortcut you created to start Coffalyser digitalMLPA.



2. Under FastQ file(s) select *Use existing definition file* and click **Next >**.



3. Navigate to the location of the Coffalyser Definition File, select the file and click **Open**.

 The Coffalyser digitalMLPA - Data Conversion – General Settings window will open.

-  IF YOU WISH TO EDIT YOUR EXPERIMENT, CONTINUE WITH STEP 4. OTHERWISE PROCEED WITH STEP 0.

4. Click **Configure** to edit your experiment(s).

 The Coffalyser digitalMLPA – Samples window will open.

5. With the *Samples* window open begin editing. For example, to select barcodes that were missing in the initial experiment setup, right click on the selected barcodes and select *Add To Experiment > Your experiment*.

6. Right click and select *Edit Experiment*, if adjustments are needed.

 The Coffalyser digitalMLPA – Experiment Overview window will open.

7. Make the desired adjustments to the experiment (e.g. experiment name, sample type, sample name).

8. Click **OK** to save and exit the *Coffalyser digitalMLPA – Experiment Overview* window.

9. Click **OK** to save the experiment(s) and close the *Coffalyser digitalMLPA – Samples* window.

10. When applicable, change the output directory by clicking the upper **Select** button. Browse to the desired folder, select it and click **OK**.

11. When applicable, change the Temp directory by clicking the lower **Select** button. Browse to the desired folder, select it and click **OK**.

12. When applicable, select another FASTQ file by clicking **Add File**. Browse to the location of the FASTQ file, select it and click **Open**. To remove the FASTQ file that was already imported, select it and click **Remove File**.

13. Click **OK** to start data analysis.

 A progress bar will appear in the Coffalyser digitalMLPA – Run Wizard.

14. Click **Open Results Folder** to access the experiment subfolder in the output directory.

3.2.2.2 REANALYSE AN EXPERIMENT USING COFFA FILES

1. Double click on MRC.DataAnalysis.DigitalMlpa.Client.exe or the shortcut you created to start Coffalyser digitalMLPA.



2. Under Coffa data files select *Analyse* and click **Next >**.



3. Browse to the folder that contains the Coffa files, select it and click **OK**.

 The Coffalyser digitalMLPA – Data Analysis – General Settings window will open.

 IF YOU WISH TO EDIT YOUR EXPERIMENT, CONTINUE WITH STEP 4. OTHERWISE PROCEED WITH STEP 0.

4. Click **Configure** to edit your experiment(s).

 The Coffalyser digitalMLPA – Samples window will open.

5. Right click on any cell and select *Edit Experiment*.

 The Coffalyser digitalMLPA – Experiment Overview window will open.

6. Make the desired adjustments to the experiment (e.g. experiment name, product, sheet, sample type, sample name).

7. Click **OK** to save and exit the *Coffalyser digitalMLPA – Experiment Overview* window.

8. Click **OK** to save the experiment(s) and close the *Coffalyser digitalMLPA – Samples* window.

9. Click **OK** to start data analysis.

 A progress bar will appear in the Coffalyser digitalMLPA – Run Wizard.

10. Click **Open Results Folder** to access the experiment subfolder in the output directory.

4. COFFALYSER digitalMLPA REPORTS

Coffalyser digitalMLPA generates four different types of reports after data analysis.

Per experiment there are two types of reports:

- **Overview (PDF report):** This report gives a short overview of the general information of all samples included in an experiment. This includes sample name, barcode (well), sample type, errors encountered during analysis and quality of the sample (see chapter 4.1). We recommend first checking the overview report for a general impression of the quality of the experiment.
- **Ratios (Excel file):** This report provides an overview of the inter ratios, read counts and intra ratios for all relevant samples and probes in an experiment (see chapter 4.2).

Per sample, there are two types of reports:

- **Sample Results 1 (PDF report):** This report includes all information of a single sample including: the general details of a sample (e.g. sample name and gender), the experiment settings, quality details, results summary plot and aberrant results that were found (see chapter 0).
- **General Details (PDF report):** This report includes the general details information of a single sample, including the sample name, analysis settings, quality details and more information on errors if applicable (see chapter 0). It is useful for getting the first impression if the data from a sample is of sufficient quality.

Note: For a No DNA sample only a single No Dna PDF report will be generated.

When printing reports from an experiment, select Fit in the print settings to fit all columns of the tables on the printed page(s) (see Figure 1).

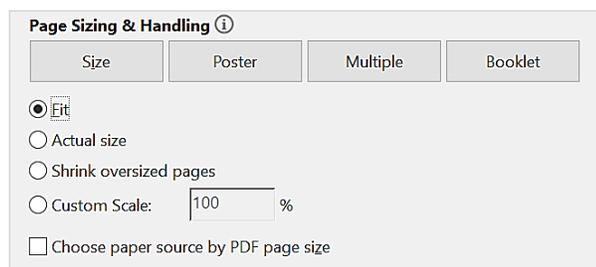


Figure 1: Print settings capture

4.1[EXPERIMENT NAME] – OVERVIEW (PDF REPORT)

This report displays an overview of all samples within an experiment. It can be used to get a first impression of the quality of an experiment, such as if the quality of your samples was sufficient or if there were any errors during the analysis. The Sample Results 1.pdf report and/or General Details.pdf report will show more information per sample.

Experiment Overview - Experiment 200414


Sample	Barcode (well)	Type	Options	Errors	Gender	Sample track probes	Quality
BP01-01	BP01-01 (A01)	Reference	Default		Female	(none)	Passed
BP01-02	BP01-02 (B01)	Test	Default		Female	(none)	Passed
BP01-03	BP01-03 (C01)	Reference	Default		Male	(none)	Passed
BP01-04	BP01-04 (D01)	Test	Default		Male	(none)	Passed
BP01-05	BP01-05 (E01)	Reference	Default		Male	(none)	Passed
BP01-06	BP01-06 (F01)	Test	Default		Male	(none)	Passed
BP01-07	BP01-07 (G01)	Test	Default		Male	(none)	Passed
BP01-08	BP01-08 (H01)	Test	Default		Female	(none)	Passed

Column: “Sample”

Displays the name that was specified for a sample in the Coffalyser Definition File. If no name was specified, the name of the barcode is displayed.

Column: “Barcode (well)”

Displays the name of the barcode used for a sample. This includes the Barcode Solutions plate reference (e.g. BP01 or BP02), the barcode number (between 01 and 96), and the well position of a given barcode on the Barcode Solutions plate (listed between brackets). E.g. well A1 of BP01 that contains barcode 1 = BP01-01(A01), and well H12 of BP01 that contains barcode 96 = BP01-96(H12).

Column: “Type”

Displays the type of sample (e.g. Reference, Test, NoDNA) as specified in the Coffalyser Definition File.

Column: “Options”

Displays the word default unless a sample was indicated as being from a “pooled DNA source” in the Coffalyser Definition File.

Column: “Errors”

Displays if an error occurred during the analysis of a sample. If an error is displayed this often indicates that no reaction analysis and/or comparative analysis was performed due to the error (meaning no Sample Results 1 report could be generated). A detailed explanation of the error can be found in the General Details report.

Column: “Gender”

Displays the detected gender of a sample. When Coffalyser digitalMLPA cannot reliably predict the gender (e.g. only one Y-control probe is found while three Y-control probes are expected in male samples), it will display the term "Ambiguous". If you have specified the gender of a sample in the Coffalyser Definition File and there is a discrepancy between the findings of Coffalyser digitalMLPA, this will be indicated in this column with an orange colour as well as a reference to the gender specified (see the example in section 4.3.3).

Column: “Sample track probes”

Displays the detected Sample Track Probe^c. A warning is given if a different Sample Track Probe is detected than the one specified in the Coffalyser Definition File.

Column: “Quality”

Displays if the reaction passed or failed the digitalMLPA quality criteria. More specific information on quality is listed in the Sample Results 1 report and the General Details report.

^c Sample Track Probes are not yet available.

4.2 [EXPERIMENT NAME] – RATIOS (EXCEL FILE)

Note: This file should be used with caution. Warnings on the quality of individual samples should be carefully checked in the pdf reports.

The Ratios Excel file can be used to investigate all samples in one experiment. It consists of four Excel sheets:

- **Inter ratios:** This sheet contains the inter ratios of all relevant reference, test and control probes in every sample. See chapter 5 and APPENDIX I – NORMALISATION for more information.
- **Read counts:** This sheet contains the total read counts of all relevant reference, test and control probes in every sample.
- **Intra ratios:** This sheet contains the intra ratios, these are the ratios obtained after intra-normalisation calculations (reaction analysis). More information on intra-normalisation is described in APPENDIX I – NORMALISATION.
- **Software version:** This sheet contains the information about the version of each software module that was used for analysis of the experiment.

Please note that use of Conditional Formatting in Excel can aid in identifying deviating probe values. In addition, the Freeze Panes option in Excel allows for an area of a worksheet to remain visible while you scroll to another area of the worksheet, creating an easy overview of the most important information.

	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Note: Please refer to the 'Probe Information File' for additional probe information which may be crucial for correct interpretation of results.													
2	Probe order	Probe number	Gene	Exon	Mapview (hg38)	Chromosomal band (hg38)	Normal copy number	Probe type	Reference probe	Additional information			BP01-06 - NA07891	BP01-07 - NA08618
3												Median total reads	872	824
4												Sample type	Test	Test
5												Estimated DNA quantity	29.3 ng	> 30.0 ng
6												SNP code	02102.21010.12111.20211.02 121.01121.12121.0020.v1	01122.20011.11102.21022.01 012.01211.20101.1021.v1
7												Mutation probes detected	0	0
8												Reference probe quality	Ok	Ok
9												Quality checks	Passed	Passed
10	6	S017541	MUTYH	16	01-045.329	1p34.1	2	CN	Yes				0.99	0.94
11	7	S010463	MUTYH	15	01-045.331	1p34.1	2	CN	No				1.03	1.03

Column: “Probe order”

Displays the order number of the probes as specified in the Product Sheet.

Column: “Probe number”

Displays the probe number of the digitalMLPA probe as designated by MRC Holland.

Column: “Gene”

Displays the name of the gene that the digitalMLPA probe targets.

Column: “Exon”

Displays the (nearest) exon number of the gene listed in the column “Gene” that the digitalMLPA probe targets.

Column: “Mapview (hg38)”

Displays the mapview location of the probe’s target sequence, based on genome build hg38.

Column: “Chromosomal band (hg38)”

Displays the chromosomal band location of the probe’s target sequence, based on genome build hg38.

Column: “Normal copy number”

Displays the copy number expected in samples from healthy individuals.

Column: “Probe type”

Displays the probe type which will be either: CN: copy number probe; CTRL: control probe (with the function specified in column “Gene”); MUT: mutation-specific probe; or WT: wild-type sequence specific probe at the location of a mutation.

Column: “Reference probe”

Displays if the probe is used as reference probe during normalisation. This is indicated by a “Yes” in this column.

Column: “Additional information”

Displays important additional information (if available) that may help with data interpretation.

Row: “Median total reads”

Displays the median number of reads obtained for reference probes in each sample.

Row: “Sample type”

Displays the sample type as defined in the Coffalyser Definition File.

Row: “Estimated DNA quantity”

Displays the estimated DNA amount of the sample used in the digitalMLPA reaction.

Row: “SNP code”

Displays the SNP code, generated based on a set of SNP-specific probes. See GLOSSARY for details.

Row: “Mutation probes detected”

Displays the number of mutation-specific probes that gave a signal in the sample. Not all probemixes contain mutation-specific probes.

Row: “Reference probe quality”

Displays the estimated quality of the normalisation in a sample. A warning highlighted in orange is displayed when the quality was found to be suboptimal (see APPENDIX II – QUALITY TABLE for more information).

Row: “Quality checks”

Displays the overall outcome of the quality evaluation of a sample. The quality checks will either be “passed” or “failed”. The quality checks are further specified in the Sample Results 1 and General details reports.

4.3 [BARCODE – SAMPLE NAME] – SAMPLE RESULTS 1 (PDF REPORT)

For every sample in an experiment a Sample Results 1 report is generated. This report consists of several sections that contain information about the sample, experiment settings, analysis details, quality details and aberrant results that were found.

4.3.1 DATES

This table indicates when each step of the data analysis was executed and which version of the Coffalyser digitalMLPA software module was used in this particular step. When a step has not been completed (e.g. due to an error), it is also shown here.

- **Configuration:** This is part of the software package and does not have to be executed.
- **Data conversion:** This is the conversion of digitalMLPA data from FASTQ to Coffa files.
- **Fragment analysis:** This is the first step of data analysis. In the fragment analysis step, each individual sample is analysed (reaction analysis).
- **Comparative analysis:** This is the second step of data analysis. In the comparative analysis step, test samples are compared to reference samples.
- **Extended analysis:** This is only applicable for certain probemixes that require an extended analysis.
- **Reporting:** This is the step where reports are generated.

4.3.2 SAMPLE QUALITY



These boxes indicate the overall outcome of the sample quality evaluation. It can be either "Failed" or "Passed". When the sample quality is failed, the results are not reliable.

More detailed information on the sample quality is located on the last page of the report. On this page the quality is divided in different components. More information on this part of the report can be found in the section 4.3.11. **Error! Reference source not found.**

4.3.3 GENERAL INFORMATION - SAMPLE

The table *Sample* contains the following information:

- **Sample name (type):** This displays the name and type that was specified for the sample in the Coffalyser Definition File. If no name was specified, the name of the barcode will be displayed.
- **Barcode – well (lot):** This displays the name of the barcode used for the sample, followed by the well position of that barcode on the Barcode Solutions plate and the detected barcode lot. The barcode name includes the Barcode Solutions plate reference (e.g. BP01 or BP02) and the barcode number (between 01 and 96). E.g. well A1 of BP01 that contains barcode 01 will be displayed as BP01-01 – A01, and well H12 of BP01 that contains barcode 96 will be displayed as BP01-96 – H12. The barcode lot is detected automatically and should match the lot number on the Barcode Solutions plate used. This is displayed as e.g. 03-007-yyymmdd.
- **SNP code:** This displays the SNP code, or sample identifier, of a sample and the version number of the SNP probe collection. Additional information is described in the Glossary and in the digitalMLPA general protocol.
- **Gender:** This displays the detected gender of the sample. When Coffalyser digitalMLPA cannot reliably determine the gender (e.g. only one Y-control probe is found while three Y-control probes are expected in male samples), it will display the term "Ambiguous". If you

have specified the gender in the Coffalyser Definition File and there is a discrepancy between the findings of Coffalyser digitalMLPA, this will be indicated here. See Figure 2 for an example.

Gender	Female (Male specified)
--------	-------------------------

Figure 2: Image from a report where there was a discrepancy between the gender identified by Coffalyser digitalMLPA and the gender indicated in the Coffalyser Definition File.

- **Estimated DNA quantity:** This displays the estimated sample DNA amount used in the digitalMLPA reaction.
- **Sample track probes:** This displays the detected Sample Track Probe^d. A warning is given if a different Sample Track Probe is detected than the one specified in the Coffalyser Definition File.

4.3.4 GENERAL INFORMATION – EXPERIMENT SETTINGS

The table *Experiment settings* contains the following information:

- **Experiment name (type):** This displays the name of the experiment. This can be either the name that was indicated in the Coffalyser Definition File, or if this was not defined, it will display Undefined Experiment 01 etc. Type refers to the experiment type as indicated in the Coffalyser Definition File (e.g. digitalMLPA).
- **Sheet name:** This displays the Product Sheet used for the data analysis. If a Product Sheet is indicated in the Coffalyser Definition File, but this is not the Product Sheet that is detected by Coffalyser digitalMLPA, the conflict will be indicated here. See Figure 3 for an example.

Sheet name	D001-B1-1119 [default sheet]. Detection based on ligation dependent identifiers, Has detected CONFLICT between used and detected product, Specified in definition file
------------	--

Figure 3: Image from a report where there was a discrepancy between the sheet identified by Coffalyser digitalMLPA and the sheet indicated in the Coffalyser Definition File.

- **Digital signatures:** This displays whether all Coffalyser digitalMLPA files in the package on your local computer used for the data analysis are identical to those provided by MRC Holland. If it states “All signatures valid”, this means these files are in their original form.
- **Probe catalogue:** This displays the Probe Catalogue version that was used for the analysis, including the release date and time. The Probe Catalogue contains (a reference to) information on all digitalMLPA probes.
- **Source files:** This displays the location and name of the FASTQ file used.

4.3.5 GENERAL INFORMATION – ANALYSIS DETAILS

The table *Analysis details* contains the following information:

- **Normalisation method:** This displays the normalisation method used.
- **Coffa file location:** This displays the location where the Coffa file was saved when the report was generated.
- **Reference samples:** This displays the reference samples used in the analysis of the sample.

4.3.6 GENERAL INFORMATION – READS

The table *Reads* provides more information on the reads associated with a sample.

^d Sample Track Probes are not yet available.

- **Recognised:** The column *Items* displays the number of unique sequences (items) that were recognised in this sample as belonging to a specific digitalMLPA probe present in the probemix used. This includes all target probes, reference probes and control probes. Note: this number does not have to be equal to the number of probes present in the probemix (e.g. reads for certain SNP-specific probes may not have been found). The column *Reads* displays the total number of reads that have the sample barcode and are recognised as digitalMLPA probes present in the probemix used.
- **Unrecognised:** This displays the total number of reads that have the sample barcode, but do not map to a digitalMLPA probe present in the probemix used.
- **Median total reads:** This displays the median number of recognised reads mapping to the reference probes in a sample. This is a measure for the read depth in a sample.
- **Median distinct reads:** This displays the median number of distinct reads of the reference probes in a sample. Distinct reads are based on the number of independent ligation events (recognised by unique molecular identifiers; UMIs).

4.3.7 GENERAL INFORMATION – PROBES

The table *Probes* contains information on different types of probes in the probemix. The number of *Expected* (as defined in the Product Sheet) and *Detected* probes are displayed. If there are more probes detected than expected, this number will be displayed in the column *Unexpected*.

- **Test probes:** This displays the number of probes that were inter-normalised. This includes both target probes and reference probes in a probemix. In reference samples, the number of expected and detected test probes must be the same. In test samples, the number of detected test probes may be lower than the number of expected test probes, for instance due to a homozygous deletion.
- **Reference probes:** This displays the number of probes used as reference probes during data normalisation. Reference probes are expected to have a normal copy number in all samples, and therefore should show the same number of expected and detected probes. However, in some samples, differences may occur in the number of expected and detected reference probes if copy number changes exist in the region where the reference probe is located. In case too many reference probe signals are missing a warning or error will appear in the quality table (see 4.3.11) as inter-normalisation may not be reliable anymore.
- **Y-control probes:** This displays the number of probes that target the Y-chromosome and are used to determine the gender of a sample. For female samples both the expected and detected number of Y probes should be 0.
- **Mutation probes:** This displays the number of mutation-specific probes in the probemix. The expected number of mutations is written as 0-n, with n being the number of mutation-specific probes in the probemix. If a mutation-specific probe generated a signal (because the mutation is present in the sample), it is displayed in the column *Detected*. More information on the mutation-specific probes (if present in a probemix) can be found in the table *Mutation specific probes* further down in the Sample Results 1 report (see section 4.3.9).

4.3.8 RESULTS SUMMARY – RATIO GRAPHS

In the *Inter ratios* plot, every dot represents a test or reference probe. They are sorted by gene or chromosomal location (X-axis), depending on the probemix. On the Y-axis the inter ratio of the probes is shown. This plot displays an overview of the variation of a sample.

The *Inter ratios aberrant region(s)* plot shows regions containing aberrant probes in more detail.

Please note, for detailed analysis refer to the *Aberrant results* table (see section 4.3.10) as this graph does not display the results of the statistical analyses.

4.3.9 RESULTS SUMMARY – MUTATION-SPECIFIC PROBES

The table *Mutation specific probes* lists all digitalMLPA mutation-specific probes in the probemix.

- **Gene:** This displays the gene name that the probe targets.
- **Mutation:** This displays the mutation that the probe targets.
- **Probe:** This displays the probe number as designated by MRC Holland.
- **Mutation Status:** This displays if the mutation is present or absent in a sample. The mutation is indicated as present when the number of reads found for the mutation-specific probe is above the defined background threshold determined by MRC Holland.
- **Additional information:** This displays any additional information about the probe or mutation.

4.3.10 RESULTS SUMMARY – ABERRANT RESULTS TABLE

The *Aberrant Results Table* displays regions with aberrant probe inter ratios. Aberrant is defined as deviating from what is expected (*within normal range* column, see Figure 4); it does not necessarily indicate that the sample has a true deletion or duplication! Deviation of a probe may be due to a SNP or experimental factor causing variation. Results must be confirmed by another probemix or technique, especially if only one probe is deviating! If all probes in the probemix are within the normal range, no table is displayed.

Note: If (many) reference probes are listed in the aberrant results table of the Sample Results 1 report, indicating deviating inter ratios, this may be an indication that the reference sample collection was not optimal. This will also result in a warning for reference sample quality on the last page of the Sample Results 1 report. Which probes are used as reference probes can be found in the product-specific Product Information File and in the Ratios Excel file.

Region information						Probe information									
Region	Chromosomal band (Hg38)	Region size	Region ratio	Stdev	Probes in region	Probe number	Probe function	Mapview (Hg38)	Normal copy number	Gene	Exon	Inter ratio	Within normal range	Different from ref population	Ambiguous
POLE	12q24.33	50 b	1.54		1	S010663	WT	12-132.674	2	POLE	13	1.54	TooHigh	Increased2StdDev	False
SMAD4	18q21.2	117,22 Kb	0.50	0.03	17	S011008	CN	18-050.968	2	SMAD4	up	0.47	TooLow	Decreased2StdDev	False
						S011009	CN	18-050.968	2	SMAD4	up	0.53	TooLow	Decreased2StdDev	False
						S010319	CN	18-051.029	2	SMAD4	up	0.54	TooLow	Decreased2StdDev	False
						S017495	CN	18-051.030	2	SMAD4	01	0.51	TooLow	Decreased2StdDev	False
						S010322	CN	18-051.047	2	SMAD4	02	0.53	TooLow	Decreased2StdDev	False
						S010323	CN	18-051.049	2	SMAD4	03	0.46	TooLow	Decreased2StdDev	False
						S017496	CN	18-051.049	2	SMAD4	04	0.55	TooLow	Decreased2StdDev	False
						S010325	CN	18-051.055	2	SMAD4	05	0.48	TooLow	Decreased2StdDev	False
						S017497	CN	18-051.058	2	SMAD4	06	0.50	TooLow	Decreased2StdDev	False
						S010327	CN	18-051.058	2	SMAD4	07	0.48	TooLow	Decreased2StdDev	False
						S010328	CN	18-051.060	2	SMAD4	08	0.59	TooLow	Decreased2StdDev	False
						S017498	CN	18-051.065	2	SMAD4	09	0.49	TooLow	Decreased2StdDev	False
						S010330	CN	18-051.067	2	SMAD4	10	0.51	TooLow	Decreased2StdDev	False
						S017499	CN	18-051.077	2	SMAD4	11	0.48	TooLow	Decreased2StdDev	False
						S010332	CN	18-051.078	2	SMAD4	12	0.50	TooLow	Decreased2StdDev	False
						S017500	CN	18-051.082	2	SMAD4	12	0.50	TooLow	Decreased2StdDev	False
						S010334	CN	18-051.085	2	SMAD4	12	0.49	TooLow	Decreased2StdDev	False
BRCA2	13q13.1	15,71 Kb	1.02	0.02	11	S010170	CN	13-032.314	2	BRCA2	up	1.04	AsExpected	Equal	False
						S010171	CN	13-032.315	2	BRCA2	01	1.03	AsExpected	Equal	False
						S017445	CN	13-032.316	2	BRCA2	01	1.01	AsExpected	Equal	False
						S010173	CN	13-032.316	2	BRCA2	02	1.04	AsExpected	Equal	False
						S010174	WT	13-032.319	2	BRCA2	03	0.95	AsExpected	Equal	False
						S017446	CN	13-032.319	2	BRCA2	03	1.00	AsExpected	Equal	False
						S010176	CN	13-032.325	2	BRCA2	04	1.03	AsExpected	Equal	False
						S010177	CN	13-032.326	2	BRCA2	05	1.02	AsExpected	Equal	False
	13q13.1	6,41 Kb	0.51	0.01	7	S017447	CN	13-032.326	2	BRCA2	06	1.03	AsExpected	Equal	False
						S010179	CN	13-032.327	2	BRCA2	07	0.95	AsExpected	Equal	False
						S010180	CN	13-032.329	2	BRCA2	08	0.99	AsExpected	Equal	False
						S017448	CN	13-032.331	2	BRCA2	09	0.51	TooLow	Decreased2StdDev	False
						S010182	CN	13-032.333	2	BRCA2	10	0.50	TooLow	Decreased2StdDev	False
						S010183	CN	13-032.333	2	BRCA2	10	0.52	TooLow	Decreased2StdDev	False
						S010184	CN	13-032.333	2	BRCA2	10	0.48	TooLow	Decreased2StdDev	False
	13q13.1	61,45 Kb	1.02	0.04	24	S017449	CN	13-032.336	2	BRCA2	11	0.53	TooLow	Decreased2StdDev	False
						S010186	CN	13-032.337	2	BRCA2	11	0.51	TooLow	Decreased2StdDev	False
						S017450	CN	13-032.337	2	BRCA2	11	0.51	TooLow	Decreased2StdDev	False
						S010659	CN	13-032.338	2	BRCA2	11	1.05	AsExpected	Equal	False
						S010189	CN	13-032.339	2	BRCA2	11	1.00	AsExpected	Equal	False
						S017451	CN	13-032.339	2	BRCA2	11	1.03	AsExpected	Equal	False
						S010191	CN	13-032.340	2	BRCA2	11	1.02	AsExpected	Equal	False
						S017452	CN	13-032.341	2	BRCA2	11	0.96	AsExpected	Equal	False
						S010193	CN	13-032.341	2	BRCA2	11	0.95	AsExpected	Equal	False
S017453	CN	13-032.345	2	BRCA2	12	1.05	AsExpected	Equal	False						
S010195	CN	13-032.347	2	BRCA2	13	0.96	AsExpected	Equal	False						
S017454	CN	13-032.355	2	BRCA2	14	1.03	AsExpected	Equal	False						
S010197	CN	13-032.357	2	BRCA2	15	1.02	AsExpected	Equal	False						
S010198	CN	13-032.358	2	BRCA2	16	1.00	AsExpected	Equal	False						
S017455	CN	13-032.363	2	BRCA2	17	1.06	AsExpected	Equal	False						
S010200	CN	13-032.363	2	BRCA2	18	1.02	AsExpected	Equal	False						
S010201	CN	13-032.370	2	BRCA2	19	0.99	AsExpected	Equal	False						
S017456	CN	13-032.371	2	BRCA2	20	1.00	AsExpected	Equal	False						

Aberrant (whole) region

Aberrant sub-region

Figure 4: Aberrant results table

4.3.10.1 REGION INFORMATION

Probes are categorised by region; this can be a chromosomal region or a gene region, depending on the probemix. The first six columns give information on the region.

- **Column Region:** This displays the name of the region (e.g. (part of) a chromosome or gene).

i If only a part of a region is aberrant (=aberrant sub-region), the region is divided. This will be visible from column 2 – *Chromosomal band* to column 6 – *Probes in region*.

- **Column Chromosomal band (hg38):** This displays the band on the chromosome where the aberrant (sub-) region is located, based on genome build hg38.
- **Column Region size:** This displays the distance between the first and last probe in this region.

- **Column *Region ratio*:** This displays the average over the inter ratios of the probes in the aberrant (sub-) region.
- **Column *Stdev*:** This displays the standard deviation calculated over the inter ratios of the probes in the aberrant (sub-) region.
- **Column *Probes in region*:** This displays the number of probes present in the aberrant (sub-) region.

4.3.10.2 PROBE INFORMATION

The last ten columns of the *Aberrant results* table display information per probe in the region.

- **Column *Probe number*:** This displays the probe number of the digitalMLPA probe as designated by MRC Holland.
- **Column *Probe function*:** This displays the probe type. CN: copy number probe; WT: wild-type sequence specific probe.
- **Column *Mapview*:** This displays the mapview location of the probe's target sequence, based on genome build hg38.
- **Column *Normal copy number*:** This displays the expected copy number of the target sequence of a probe in samples from healthy individuals.
- **Column *Gene*:** This displays the name of the gene that the digitalMLPA probe targets.
- **Column *Exon*:** This displays the (nearest) exon number of the gene listed in "Gene" that the digitalMLPA probe targets.
- **Column *Inter ratio*:** This displays the inter ratio of the probe.
- **Column *Within normal range*:** This displays the category that the inter ratio of a probe falls into (see chapter 5 for more information).

 Note: Probes with inter ratios in the normal range are only displayed in the aberrant results table when another probe in the same region has an aberrant inter ratio.

- **Column *Different from ref population*:** This displays if the inter ratio of the probe in a sample is different compared to the inter ratios of the same probe in the reference samples. This can state equal or be listed as increased or decreased by one or two standard deviations. See chapter 5 for more information.

 Note: If only one or two samples are specified as reference samples, the analysis cannot be reliably performed and this field will be left empty.

- **Column *Ambiguous*:** This displays if the calculated inter ratio of a probe is found to be ambiguous or not (indicated by 'True' or 'False', respectively). See chapter 5 for more information.

4.3.11 QUALITY DETAILS – QUALITY TABLE

Quality assessment is an essential part of digitalMLPA analysis: based on this evaluation it is determined if the data is of sufficient quality for reliable copy number determination. It may also help to identify possible problems with the sample and/or experiment.

Coffalyser digitalMLPA assesses the quality of each individual sample (*Reaction analysis*), and the quality of each sample compared to the reference samples (*Comparative analysis*). It displays the results of these evaluations in the Quality table.

digitalMLPA probemixes contain a large number of control probes. This allows for a more thorough quality evaluation as compared to conventional MLPA probemixes.

For each category of control probes (e.g. denaturation control) two to four different probes are included in every digitalMLPA probemix. The software checks if the control probes meet the set criteria. In the *Reaction analysis*, the intra ratios of the control probes are compared to what is expected based on internal testing data generated at MRC Holland. During *Comparative analysis*, the inter ratios of the control probes are compared between the test sample and reference samples. Control probe ratios should be similar in test and reference samples, which is necessary for reliable data interpretation.

When there are significant differences (outside of the “normal” range) in the evaluation of results of the control probes in the *Reaction analysis* and/or *Comparative analysis*, a warning or error is displayed in the reports. If warnings or errors are present, analysis results should be examined and interpreted with extra care.

For more information on the functions of the different quality checks present in the table see APPENDIX II – QUALITY TABLE.

4.4 [BARCODE – SAMPLE NAME] – GENERAL DETAILS (PDF REPORT)

The General details report displays an overview of the general details and the quality of a sample. In addition, errors mentioned in the Overview report are further explained in this report. Several tables in this report are similar to the tables in the Sample Results 1 report. Please refer to chapter 0 for information on the Sample, Reads, Probes, Dates and Quality details displayed in this report. The sections Analysis settings and Errors in the report are discussed below.

4.4.1 ANALYSIS SETTINGS

This table provides more information on the analysis settings of the experiment.

- **Analysis settings:** This displays the analysis mode used.
- **Experiment name (type):** This displays the name of the experiment. This can be either the name that was indicated in the Coffalyser Definition File, or if this was not defined, it displays Undefined Experiment followed by a number (e.g. 01 or 02 etc.). Type refers to the experiment type as indicated in the Coffalyser Definition File (e.g. digitalMLPA).
- **Sample file:** This displays the location where the Coffa file for the sample was stored.
- **Sample file id:** This displays the unique Coffa file identification code.
- **Product name:** This displays the name of the Product Sheet used for the analysis.
- **Product sheet / graph id:** This displays both the unique identifiers of the Product Sheet and the search graph that were used for the analysis of the data.
- **Product selection:** This displays the method used by Coffalyser digitalMLPA for identification of the digitalMLPA product lot and its corresponding product sheet. The software will detect the product lot that was used based on ligation dependent or ligation independent probemix lot identifiers, depending on the amount of DNA input used (sufficient DNA or low DNA respectively).

If a product and product sheet are selected by the user in the Coffalyser Definition File, but this does not correspond to the product lot that is detected by Coffalyser digitalMLPA, the software will indicate this conflict. Please note that the product sheet selected in the Coffalyser Definition File will be used for analysis.

- **Extended module:** This displays whether an extended analysis was performed.
- **Probe catalogue:** This displays the Probe catalogue version, including the release date and time that was used for the analysis. The Probe Catalogue contains (a reference to) information on all digitalMLPA probes.
- **Normalisation method:** This displays the normalisation method that was used.
- **Digital signatures:** This displays whether all Coffalyser digitalMLPA files in the package on your local computer used for the data analysis are identical to those provided by MRC Holland. If it states “All signatures valid”, this means these files are in their original form.
- **Read count type:** This displays which reads were used for calculations. In most probemixes this is the total read count (indicated by *Total*).
- **Source files:** This displays the location and name of the FASTQ file that was used for analysis.
- **Reference samples:** This displays the reference samples used in the analysis of the sample.

4.4.2 ERRORS

Errors are shown in the digitalMLPA Run Wizard directly after data analysis and are described in more detail in the General Details report. More information about interpretation of these error messages can be found in **Error! Reference source not found.**

5. ANALYSIS OF RESULTS

5.1 DISPLAY OF RESULTS

5.1.1 COMPARISON TO ARBITRARY BORDERS

Inter ratios are categorized based on the boundaries set for each probe in a specific probemix, also known as arbitrary borders. Based on the category, the software applies a colour highlight to the probe result in the column *Within normal range* of the Aberrant results table in the Sample Results 1 report and in the sheet *Inter ratios* of the Ratios Excel file.

Table 1 shows an example of the possible categories, the corresponding information, and the highlight colours of the column *Within normal range* and the sheet *inter ratios*. Refer to the digitalMLPA probemix-specific Product Description for interpretation of results.

Table 1: Categories of inter ratios and the corresponding information and background colour.

Category	Boundaries (example)	Within normal range column (Sample Results 1 PDF report)	Inter ratios sheet (Ratios Excel file)
1.	>1.3	Too high ; Orange	Dark blue
2.	1.2 – 1.3	High ; Yellow	Light blue
3.	0.8 – 1.2	As expected ; None ^e	None
4.	0.7 – 0.8	Low ; Yellow	Light orange
5.	<0.7	Too low ; Orange	Dark orange

5.1.2 COMPARISON TO REFERENCE POPULATION

In addition to calculating inter ratios, Coffalyser digitalMLPA also estimates the difference in the inter ratio of a sample compared to that of the reference sample population per probe. This deviation between the inter ratios of a probe can be either: Equal (negligible difference between the ratios), differ by 1 standard deviation decrease/increase, or differ by 2 standard deviations decrease/increase.

The results of the estimate of deviation are displayed in the column *Different from ref population* of the Aberrant results table in the Sample Results 1 report and in the sheet *Inter ratios* of the Ratios Excel file. Table 2: Background colour and formatting based on deviation from reference samples. shows the possibilities, the highlight colours and formatting of the column *Different from ref population* and the sheet *inter ratios*.

Table 2: Background colour and formatting based on deviation from reference samples.

Deviation from the reference samples	Different from ref population column (Sample Results 1 PDF report)	Inter ratios sheet (Ratios Excel file)
Equal	Equal ; None	None
Decreased or Increased by 1 Standard Deviation	Decreased1StdDev / Increased1StdDev ; Yellow	Bold + italic
Decreased or Increased by 2 Standard Deviation	Decreased2StdDev / Increased2StdDev ; Orange	Bold

5.2 INTERPRETATION

In addition to this manual please consult the relevant Product Description, Probe Information File and, if available, other product-specific files provided by MRC Holland for interpretation of results.

^e Probes with inter ratios in the normal range are only displayed in the aberrant results table of the Sample Results 1 report when another probe in the same region has an aberrant probe ratio.

For result interpretation consult all three rightmost columns (*Within normal range*, *Different from ref population* and *Ambiguous*) of the Aberrant results table in the Sample Results 1 report. Table 3: Overview of combination of calls and highlights in the Sample Results 1 report. provides an overview of the combination of calls in the columns *Within normal range*, *Different from ref population* and *Ambiguous* that may aid in result interpretation.

Table 3: Overview of combination of calls and highlights in the Sample Results 1 report.

Within normal range	Different from ref population	Ambiguous	Interpretation
As expected	Equal	False	Sample probe result is within the set boundaries for this probe and is similar to results of the reference samples.
As expected	Decreased 1 StdDev or Increased 1 StdDev	False	Sample probe result is within the set boundaries for this probe and is not significantly different from the reference samples ^f .
As expected	Decreased 2 StdDev or Increased 2 StdDev	False	Sample probe result is within the set boundaries for this probe and is significantly different from the reference samples ^f .
Low or High	Equal	False	Sample probe result is outside the set inner boundaries for this probe and is similar to results of the reference samples ^f .
Low or High	Decreased 1 StdDev or Increased 1 StdDev	False	Sample probe result is outside the set inner boundaries for this probe and is not significantly different from the reference samples ^f .
Low or High	Decreased 2 StdDev or Increased 2 StdDev	False	Sample probe result is outside the set inner boundaries for this probe and is significantly different from the reference samples ^f .
Too low or Too high	Equal	True	Sample probe result is outside the set boundaries for this probe and is similar to results of the reference samples ^f .
Too low or Too high	Decreased 1 StdDev or Increased 1 StdDev	False	Sample probe result is outside the set boundaries for this probe and is not significantly different from the reference samples ^f .
Too low or Too high	Decreased 2 StdDev or Increased 2 StdDev	False	Sample probe result is outside the set boundaries for this probe and is significantly different from the reference samples.

If the inter ratio of a probe falls into the highest or lowest category (resulting in 'Too high' or 'Too low' in the column *Within normal range*), it does not necessarily mean that the probe's target sequence is indeed duplicated or deleted. It is possible that the inter ratio of a probe in a sample is not significantly different from the reference samples used in the analysis. For instance, if a probe is highly variable in the reference samples used, the column *Ambiguous* will state 'True' in this case. Results for probes showing this warning should be interpreted with caution!

When a probe deviates by 2 standard deviations from the reference sample population AND multiple adjacent probes show the same result the likelihood that the observed result is due to a copy number change is increased.

^f Proceed with caution when interpreting this data.

APPENDIX I – NORMALISATION

Coffalyser digitalMLPA uses a series of normalisation steps and calculations to compute the inter ratios of digitalMLPA probes.

In the first step, called intra-normalisation, Coffalyser digitalMLPA converts absolute read counts of a sample into relative values (intra ratios) by normalising target probe read counts against the read counts of every reference probe in the sample. This is done for every probe and every sample. In the next step, called inter-normalisation, Coffalyser digitalMLPA compares the relative probe values of each sample to those of the reference samples, resulting in inter ratios.

A simplified version of the normalisation process is as follows:

STEP 1

The read count of target probe 1 (Tp1) is divided by the read count of reference probe 1 (Rp1) in sample 1 to give a relative value for the target probe. The same is done for reference sample 1. The relative target probe value in the sample is then divided by the relative target probe value in the reference sample resulting in an intermediate ratio.

This calculation is then repeated using every **reference probe** included in the probemix for the target probe. This results in the same number of intermediate ratios for target probe 1 as there are reference probes in the probemix. Next, the median value of these intermediate ratios is determined. See the equation below.

$$\text{Median} \left(\frac{(Tp1 \text{ in sample 1} / Rp1 \text{ in sample 1})}{(Tp1 \text{ in reference sample 1} / Rp1 \text{ in reference sample 1})}, \dots, \frac{(Tp1 \text{ in sample 1} / Rpn \text{ in sample 1})}{(Tp1 \text{ in reference sample 1} / Rpn \text{ in reference sample 1})} \right)$$

STEP 2

Step 1 is repeated using every **reference sample** included in the analysis. This results in as many median values for target probe 1 in sample 1 as there are reference samples in the analysis.

Coffalyser digitalMLPA then calculates the average value over these median values. This result is the inter ratio of target probe 1 in sample 1.

These calculations are repeated for all probes in the probemix and all samples in the experiment.

NO REFERENCE SAMPLES DEFINED IN ANALYSIS

When no reference samples are defined in an analysis, all test and undefined samples are used for normalisation. This will result in the same number of median values for a probe as there are samples in the experiment. The median of these values is the inter ratio for a probe in a sample.

APPENDIX II – QUALITY TABLE

In two Coffalyser digitalMLPA reports, General Details.pdf and Sample Results 1.pdf, a table is present outlining the detailed quality checks for the sample. This table is divided into two categories: (1) Reaction analysis, this focuses on the quality checks for each individual sample, and (2) Comparative analysis, this focuses on the quality checks from the comparison of a sample to the reference samples. The various quality checks are explained in Table 4 below.

Quality checks can have three different classifications based on the presence and/or severity of the issue: OK, Warning and Error. Note: some quality checks are only displayed when a warning or error is identified (see Table 4).

The overall quality of the sample, as depicted by the Quality checks “Passed” or “Failed” boxes in green or red respectively (see **Error! Reference source not found.**), is based on the cumulative observations from all quality checks described in Table 4. Please note that the overall quality of a sample is based on a weighted scoring system that takes the number of control probes affected per category into account. This quality scoring is not visible in reports or completely reflected in the warning and error messages. Therefore, it is possible that samples with identical warning and error messages displayed on a report will have a different overall quality assessment. For more information on the control probes affected, check the intra and inter ratios of the control probes in the Ratios Excel file.

Please refer to the digitalMLPA General Protocol for more information on the various quality checks and troubleshooting.

Table 4: Overview of quality checks which may appear in the Quality table.

Description	Category	Explanation
Disqualified lot identification probes	Reaction analysis	This quality classification is based on lot identification probes that are present in every probemix and are used for automatic Product Sheet selection or to check if the Product Sheet selection present in the Coffalyser Definition File was correct. If lot identification probes are found that do not belong to the Product Sheet selected for the sample, e.g. due to small amounts of another probemix lot being present, a warning or error is displayed. Please note that in case lot identification probes not belonging to a certain Product Sheet are seen in larger amounts, a conflict in sheet detection will appear.
DNA contamination ^g	Reaction analysis	This quality classification is based on the set of 39 SNP-specific probes that generate the SNP code. When there are three or more "?" signs, a warning or error is displayed.
Missing reference probes ^g	Reaction analysis	This quality classification is based on the identified read counts for reference probes. Data analysis may not be reliable when too many reference probes have insufficient read numbers. In case an error is displayed for this quality check, comparative analysis will not be performed and the Sample Results 1 report will not be available.
Read depth	Reaction analysis	This quality classification is based on the median number of total reads of the reference probes for a sample, displayed in the Reads table as “Median total reads”. When read depth is too low, an error is displayed as data analysis may not be reliable.
Read depth (qualified) ^g	Reaction analysis	Highly similar to Read depth, but based on the qualified (above qualification threshold) reference probe reads.

Description	Category	Explanation
Sample uniformity ^g	Reaction analysis	This quality classification is based on reads being evenly spread over the samples within one experiment. If a significantly different number of reads is assigned to a barcode as compared to the rest of the experiment, a warning or error is displayed.
Sequence quality	Reaction analysis	This quality classification is based on the quality of the sequencing reads. When many reads contain errors, a warning or error is displayed.
Unique reads ^g	Reaction analysis	This quality classification is based on the median number of distinct reads of the reference probes for a sample, displayed in the Reads table as “Median distinct reads”. When the median number of distinct reads is too low, an error is displayed.
Unrecognised reads	Reaction analysis	This quality classification is based on the amount of reads that cannot be reliably assigned to a probe. When many digitalMLPA reads are not reliably assigned to one of the probes present in the detected or selected Product Sheet, there is a warning or error displayed.
Hybridisation completeness	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that check the extent of the hybridisation reaction. If a warning or error is displayed in the Reaction analysis category this indicates that hybridisation was incomplete. If a warning or error is displayed in the Comparative analysis category this indicates that the level of hybridisation differs between the test and the reference samples.
Hybridisation T _m	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that are dependent on the temperature during the overnight hybridisation.
Ligase activity	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that check the ligase activity. If ligase activity is outside of the designated margins, a warning or error is displayed.
Ligase inactivation	Reaction analysis	This quality classification is based on the set of control probes in a probemix that detect when there is simultaneous ligase and polymerase activity, for example when there is incomplete ligase inactivation. When there is simultaneous activity of these enzymes a warning or error is displayed.
Ligation start temperature	Reaction analysis	This quality classification is based on the set of control probes in a probemix that indicate if the ligation reaction was started at room temperature. When deviation in these probes is detected a warning or error is displayed.
Polymerase activity	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that indicate if polymerase activity was sufficient. If polymerase activity was outside the expected margins a warning or error is displayed.

Description	Category	Explanation
Sample DNA denaturation	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that indicate if sample DNA denaturation was complete. A warning or error in the Reaction analysis category indicates that sample DNA denaturation was incomplete. A warning or error in the Comparative analysis category indicates that the test sample is not denatured to a similar extent as the reference samples.
Sample DNA depurination ⁹	Reaction analysis & comparative analysis	This quality classification is based on the set of control probes in a probemix that indicate the extent of sample DNA depurination. A warning or error in the Reaction analysis category indicates that the sample DNA is depurinated. A warning or error in the Comparative analysis category indicates that the test and reference samples do not have the same level of depurination.
X-presence control ⁹	Reaction analysis	This quality classification is based on the set of control probes in a probemix that indicate the presence of X-chromosome sequences. If one or more of the X-chromosome control probes do not have reads, a warning or error is displayed.
Y-presence control ⁹	Reaction analysis	This quality classification is based on the set of control probes in a probemix that indicate the presence of Y-chromosome sequences. If more than one but not all Y-chromosome control probes have reads visible in the <i>Probes</i> table (see section 4.3.7), a warning or error is displayed.
Sample DNA fragment length	Comparative analysis	This quality classification is based on the set of control probes in a probemix that indicate the sample DNA fragment length. When sample DNA fragmentation differs considerably between different samples a warning or error is displayed.
Reference probe quality	Comparative analysis	This quality classification is based on the estimated quality of the data normalisation within a sample based on the variation between intermediate ratios with different reference probes (see APPENDIX I – NORMALISATION). If a large amount of variation is present between intermediate ratios a warning or error is displayed. This typically indicates that reference probes are not stable in the sample.
Reference sample quality	Comparative analysis	This quality classification is based on the estimated quality of the data normalisation over the different reference samples as indicated by the variation between the median values of probes in a sample. If a large amount of variation is present between median values a warning or error is displayed. This typically indicates a sub-optimal choice of reference samples.

⁹ These quality checks will only be displayed in the Quality table in the event of a warning or error.

APPENDIX III – TROUBLESHOOTING

GENERAL

Problem	Cause and solution
Program “MRC.DataAnalysis.DigitalMLpa.Client.exe” is not able to start.	Virus scanners can prevent the program from running from a network drive. If you unpacked your files on a network drive, try to unpack the files on a local drive. If the problem persists, please contact MRC Holland at info@mrcholland.com .
.NET Framework is not up-to-date (older than version 4.7).	The latest .NET framework can be downloaded from https://dotnet.microsoft.com/download/dotnet-framework .

EXPERIMENT DEFINITION

Problem	Cause and solution
The following warning is shown when reusing an old Coffalyser Definition File: <i>The specified analysis settings xxx cannot be found. Do you want to continue?</i>	This message indicates that the configuration of Coffalyser digitalMLPA has changed compared to when the Coffalyser Definition File was made. The Coffalyser digitalMLPA Release Notes, available in the “_Documentation” folder of the software package, contain information about the changes between software versions. When you choose to continue, the Coffalyser Definition File will be updated using the new settings.
The following error message is shown when pasting sample details from the clipboard into the Coffalyser Definition File Editor: <i>Failed to paste the sample details: the samples on your clipboard does not match the number of selected rows</i>	This message indicates that the number of copied sample names being pasted into Coffalyser digitalMLPA is not equal to the number of selected samples in the Coffalyser Definition File Editor. Please ensure the correct number of samples is selected in Coffalyser digitalMLPA before pasting the sample names.
The following error message is shown when pasting sample details from the clipboard into the Coffalyser Definition File Editor: <i>Failed to paste the sample details: the samples on your clipboard do not fit in the remaining number of rows given your current start cell.</i>	This message indicates that the number of copied sample names being pasted into Coffalyser digitalMLPA exceeds the number of remaining samples in the range. Please ensure that the number of sample names you want to copy is smaller than or equal to the range of samples in the experiment in Coffalyser digitalMLPA.

DATA ANALYSIS (COFFALYSER digitalMLPA – RUN WIZARD)

Problem	Cause and solution
When using an old barcode lot (plate identifier has format “yymmdd”, instead of e.g. 03-007-yymmdd) an error for barcode BP1-33 is shown after data analysis.	In the older Barcode Solutions plates, two versions of BP1-33 exist, one called “BP1-33” and the other one called “Deprecated_BP1-33” (at the bottom of the barcode list). In most cases, BP1-33 is the correct one, but if you receive an error for barcode BP1-33 after data analysis, please reanalyse your experiment with Deprecated_BP1-33.
My run was completed but error messages are shown in the summary window.	Click “Open Results Folder” and review the Overview.pdf and General Details.pdf reports for more information.
Output files of some of the samples/barcodes ended up in the “Undefined Experiment {nn}” folder.	Ensure the barcodes were part of your experiment and included in the Coffalyser Definition File. Reanalysis can be done using all Coffa files belonging to your experiment (see section 3.2.2.1) or by using an updated Coffalyser Definition File (see section 3.2.2.1).
I received the following error message: <i>An unhandled exception occurred.</i> <i>The Coffalyser digitalMLPA encountered an error and needs to close.</i>	<ol style="list-style-type: none"> <li data-bbox="764 787 1476 913">1. This can happen when Coffa files are reanalysed, while reports of these Coffa files are still open. Close the reports and start the reanalysis again. <li data-bbox="764 913 1476 1186">2. This can happen when no samples in an experiment can be analysed e.g. only no DNA reactions are present in an (undefined) experiment, or all reactions have very low read counts and end up in the “Undefined Experiment {nn}” folder. Please check whether all your samples are correctly defined in the Coffalyser Definition File. <p data-bbox="764 1186 1476 1339">If you encounter other unhandled exceptions/errors, please contact MRC Holland at info@mrcholland.com and include a screenshot of the error.</p>
I received the following error message: <i>The parent folder contains more than 100 subfolders, which will prevent Microsoft Windows from showing the full list in the following form.</i>	Restructure the parent folder (e.g. combine output folders per week or month) so less than 100 subfolders are present in the output folder specified (see step 24 of the procedure in section 3.2.1).
I received the following error message: <i>The average read length (xx) for the first 500000 reads is too short (minimum required xx).</i>	The read length in the FASTQ file is too short for reliable analysis. Please check the instructions in the digitalMLPA General Protocol and the probemix-specific Product Description for the minimum required read length and rerun the digitalMLPA samples on the sequencer (ensuring read length is sufficient).

SAMPLE RESULTS 1 REPORT

Problem	Cause and solution
Sample Results 1 report is not generated for a specific sample.	The cause of the error is described in the General Details Report of this particular sample.
A warning is displayed under "Sheet name" in the Experiment settings table, stating the following: <i>"has detected CONFLICT between used and detected product"</i>	If Coffalyser digitalMLPA detected another product than the product indicated in the Coffalyser Definition File (i.e. the selected Product Sheet), the software issues this warning, see section 4.3.4. To correct this please ensure that the correct Product Sheet is selected.
A different gender than specified in the Coffalyser Definition File was detected by the Coffalyser digitalMLPA (as displayed in the Overview file or the Sample Results 1 report).	This can occur in case of a sample swap or if a sample was incorrectly specified in the Coffalyser Definition File.
Sequences for only two out of three control probes targeting the Y chromosome are detected in a male sample, as displayed in the <i>Probes</i> table under "Y-control probes".	This can be caused by a SNP at or around the ligation site of one of the control probes targeting the Y chromosome.
Sequences for one of the control probes targeting the Y chromosome are detected in a female sample, as displayed in the <i>Probes</i> table under "Y-control probes".	This could be due to a translocation of part of the Y-chromosome onto an autosomal or X-chromosome. Another possible explanation is sample contamination.
Sequences for only one of the control probes targeting the Y chromosome are detected in an SD sample provided by MRC Holland.	Some SD samples contain only a part of the Y-chromosome as additional DNA is added to female DNA. Details of the SD can be found in the SD-specific Product Description.
I have a warning for one control probe type (e.g. denaturation probes) in the comparative analysis, but not in the reaction analysis.	There is a difference (in quality, in the amount of contaminants present etc.) between this sample and the reference samples. This can occur for example if another DNA extraction method was used between test samples and reference samples. For more information on the control probe warnings, please see APPENDIX II – QUALITY TABLE and the digitalMLPA General Protocol.
I have a warning for one control probe type (e.g. denaturation probes) in the reaction analysis, but not in the comparative analysis.	This indicates that the quality of both the sample and the reference samples may be compromised. However, there is no difference (in quality, in the amount of contaminants present etc.) between this sample and the reference samples. For more information on the control probe warnings, please see APPENDIX II – QUALITY TABLE and the digitalMLPA General Protocol.

Problem	Cause and solution
<p>A warning is given for reference sample quality. Why do I receive this warning?</p>	<p>When reference samples have been designated:</p> <ol style="list-style-type: none"> 1. This can be caused by copy number variations in one or more of the reference samples. Check the table <i>Probes</i> in the Sample Results 1 report of the reference samples to see if all expected probes have been detected and that the reference samples do not contain any copy number aberrations. 2. This can be caused by a high standard deviation of one or multiple probes over the reference samples. To check this, go to the Ratios Excel file (chapter 4.2) and calculate the Standard deviation of all probes over the reference samples using the Excel function Stdev.P. Check if any of the values is higher than 0.12. It may be that one single reference sample is aberrant; in that case you might consider re-analysing the experiment without that specific sample being designated as a reference sample (provided that 3 or more reference samples remain). <p>If no samples have been designated as reference samples:</p> <ol style="list-style-type: none"> 1. Calculate the median absolute deviation for each probe over all samples. <ul style="list-style-type: none"> • Determine the median inter ratio per probe. • Determine the absolute difference between the inter ratio per sample and the median. • Determine the median of the absolute differences. 2. Sort the median absolute deviations from highest to lowest. 3. If the highest median absolute deviations is >0.08, this would be the most likely cause of the warning or error.

GENERAL DETAILS REPORT

Problem	Cause and solution
<p>A warning is displayed under “Product selection” in the Analysis settings table, stating the following:</p> <p><i>“has detected CONFLICT between used and detected product”</i></p>	<p>If Coffalyser digitalMLPA detected another product than the product specified in the Coffalyser Definition File (i.e. the selected Product Sheet), the software issues this warning, see section 4.3.4.</p> <p>To correct this please ensure that the correct Product Sheet is selected.</p>

Problem	Cause and solution				
<p>The following error was reported:</p> <p><i>Error source: comparative analysis. Error message: sample x missing too many reference probes.</i></p> <div style="border: 1px solid black; padding: 2px; margin-top: 10px;"> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="background-color: #f00; color: white; padding: 2px;">error source</td> <td style="background-color: #f00; color: white; padding: 2px;">error message</td> </tr> <tr> <td style="padding: 2px;">comparative analysis</td> <td style="padding: 2px;">sample ('BP01-08') missing too many reference probes</td> </tr> </table> </div>	error source	error message	comparative analysis	sample ('BP01-08') missing too many reference probes	<p>This error is given when no reads have been detected for a large number of reference probes, this does not allow for reliable comparative analysis. This may happen when for example a no DNA sample was not specified as such in the Coffalyser Definition File.</p> <p>Please select “No DNA” as sample type for this sample, if it is a no DNA sample, and reanalyse the experiment.</p>
error source	error message				
comparative analysis	sample ('BP01-08') missing too many reference probes				
<p>The following error was reported:</p> <p><i>Error source: comparative analysis. Error message: the experiment does not have any analysable samples.</i></p>	<p>Please ensure that the correct Product Sheet is selected.</p> <p>Check product selection info to see if there is a conflict between used and detected product.</p>				

APPENDIX IV – COMMAND LINE INTERFACE (CLI)

For easy integration in your sequencing pipeline, Coffalyser digitalMLPA modules can be accessed through the Command Line Interface (CLI): `MRC.DataAnalysis.DigitalMlpa.{module name}.ClientCli.exe` using the Windows Command Prompt. Executing any of these CLI console applications without any parameters invokes the help function and displays an example call and a brief overview of the parameters (see screen capture below).

```

*** EXAMPLE USAGE ***

MRC.DataAnalysis.DigitalMlpa.Analyses.ClientCli.exe [-Definition] "C:\Path\" [-AnalysisType] (FragmentAnalysis|ComparativeAnalysis|ExtendedAnalysis|All)

MANDATORY PARAMETERS:

-Definition      : path to the definition directory or file;
-AnalysisType   : the analysis type (FragmentAnalysis and/or ComparativeAnalysis and/or ExtendedAnalysis ... or just All);
  
```

CLI ANALYSIS WITH COFFALYSER DEFINITION FILE

Coffalyser digitalMLPA may be called with a Coffalyser Definition File (*.cdf) where parameters have been specified. Coffalyser digitalMLPA comes with a Definition File Editor for convenient creation of *.cdf files: `MRC.DataAnalysis.DigitalMlpa.Definitions.Editor.exe`. You can either generate a Coffalyser Definition File for each analysis with the Definition File Editor, or use the Definition File Editor generated Coffalyser Definition Files as examples/templates and generate Coffalyser Definition Files from your experiment input files in your own pipeline.

COFFALYSER DEFINITION FILE HEADER

The Coffalyser Definition File header layout is as follows:

```

DefinitionVersion → → → {version.number.(int)} CRLE
AnalysisSettings → → → {file.id.(guid)} CRLE
SampleBarcodes → → → {file.id.(guid)} CRLE
PathData → → → → {path.to.output.directory, for.example: "N:\NGS.output\CoffalyserNet\"} CRLE
PathTemp → → → → {path.to.directory.for.temp.files, for.example: "C:\temp"} CRLE
PathsInputConversion → {path.to.input.FASTQ(.gz).file(s)} CRLE
  
```

Tabs are represented as yellow arrows, spaces are represented as yellow periods, **CRLE** represents "enter" (carriage return, line feed), and text in {curly brackets} indicate an explanation of the field.

An example of a *.cdf file header:

```

DefinitionVersion → → → 1 CRLE
AnalysisSettings → → → 518d1a77-1d51-4c59-86b6-f95a6c1be30a CRLE
SampleBarcodes → → → c467ebc7-99bc-80ed-25f5-109097d77b97 CRLE
PathData → → → → C:\Coffalyser.NetOutput CRLE
PathTemp → → → → C:\temp CRLE
PathsInputConversion → \\networklocation\FASTQs\MiniSeq-RUN093.fastq.gz CRLE
  
```

The file IDs of the AnalysisSettings and the SampleBarcodes may be found in the header of the AnalysisSettings and the SampleBarcodes files, which are located in the "_Configuration" sub-directory of your Coffalyser package. These Global Unique Identifiers (GUID) are filled in automatically when you create a Coffalyser Definition File in the editor. The PathData parameter specifies the location where Coffalyser digitalMLPA will create the output directory. This header line is optional; if the PathData is not specified, Coffalyser digitalMLPA will use the location of this Coffalyser Definition File as the output directory. The PathTemp parameter specifies a location for temporary working files

(preferably on a local fast drive). The PathTemp line is optional; if not specified, Coffalyser digitalMLPA will use the PathData directory as the location for the temporary working files.

An analysis may be run with no further details in the Coffalyser Definition File. In that case, Coffalyser digitalMLPA will autodetect the digitalMLPA product used in each sample and will create experiments by grouping all samples tested with the same digitalMLPA probemix. All samples will be treated as "Test samples". In case your experiment contains samples of any other type (e.g. "No DNA", "Positive Reference", "Reference sample", "SD", "Undefined") you must define these in the Coffalyser Definition File. MRC Holland recommends always explicitly defining experiment type, sample type, sample names, gender and if the sample is from a pooled DNA source to raise the likelihood of detecting mistakes.

COFFALYSER DEFINITION FILE EXPERIMENT DEFINITION

Underneath the header in the Coffalyser Definition File the experiment definition is located. Each experiment section starts with a blank line, followed by an experiment header:

```
CRLE
Experiment → {Experiment.name} → {Product.Sheet.id.{guid}. {experiment.type} CRLE
```

For example:

```
CRLE
Experiment → My.DigitalMLPA.Experiment → 00000000-0000-0000-0000-000000000000 → DigitalMlpaDna CRLE
```

The experiment header line is single tab separated. The Product Sheet ID may be found in the header of the Product sheet in the "_Configuration" sub-directory of your Coffalyser digitalMLPA package. If the Product Sheet ID is left to "00000000-0000-0000-0000-000000000000", Coffalyser digitalMLPA will attempt to autodetect which Product Sheet to use. The Product Sheet id {GUID} can be found in the header of the Product Sheet .txt files which are located in the "_Configuration\Product Sheet" sub-directory of your Coffalyser digitalMLPA package.

The experiment header is followed by a line for each barcode in the experiment:

```
Sample → {barcode.id.{guid}} → {function.of.sample} → {Sample.name} → {Sample.gender} → {Sample.options} → CRLE
```

For example:

```
Sample → ba626105-9a0a-4b10-9a30-c56382d140be → SampleTest → BP01-01.REF.sample.1.PROMEGA → Male → PooledDnaSource CRLE
```

You may find the barcode IDs (GUID) in the SampleBarcodes file (for example "C:\Coffalyser\Coffalyser Package [20200525-145849]_Configuration\Sample Barcodes\03-xxx-yymmdd ; Illumina standard.txt").

The options for {function of sample} are: "NoDna", "SampleReference", "SampleTest", "SampleReferencePositive", "SampleDna" and "Undefined".

The options for {Sample gender} are: "Ambiguous", "Female" and "Male".

The options for {Sample options} are: "PooledDnaSource" and "Default".

RUNNING COFFALYSER DEFINITION FILE IN CLI

Once you have created your *.cdf file, you may call the DataConversion in the CLI using the following lines:

```
MRC.DataAnalysis.DigitalMlpa.DataConversion.ClientCli.exe -Definition
"C:\Path\MyDefinitionFile.cdf"
```

To call the Sample Analysis (including FragmentAnalysis for quality analysis of each sample and ComparativeAnalysis for comparison with the references) use the following command in the CLI:

```
MRC.DataAnalysis.DigitalMlpa.Analyses.ClientCli.exe -Definition
"C:\Path\MyDefinitionFile.cdf" -AnalysisType FragmentAnalysis,ComparativeAnalysis
```

And finally call Reporting (to generate the .PDF reports as well as the Excel file) using the following command in the CLI:

```
MRC.DataAnalysis.DigitalMlpa.Reporting.ClientCli.exe -Definition
"C:\Path\MyDefinitionFile.cdf"
```

CLI ANALYSIS WITHOUT A COFFALYSER DEFINITION FILE

The DataConversion module may be called without first making a Coffalyser Definition File. To do that, type into the CLI the following:

```
MRC.DataAnalysis.DigitalMlpa.DataConversion.ClientCli.exe -PathsInputConversion
"C:\Path1\File1.fastq.gz" -PathData "C:\Path" -PathTemp "C:\Path" -AnalysisSettings
"Profile Name" -SampleBarcodes "Barcode Lot Name"
```

The `-PathsInputConversion` parameter points to the FASTQ file(s) (multiple FastQ files can be selected simultaneously by listing them space separated).

The `-PathData` parameter points to the output directory.

The `-PathTemp` parameter points to a directory for the temporary files (optional, if not specified Coffalyser will use the PathData output directory).

The `-AnalysisSettings` parameter specifies which AnalysisSetting file to use (optional; if not specified, Coffalyser digitalMLPA will use the default from the configuration). The name of the AnalysisSettings can be found in the header of the files in the "Configuration\Analysis Settings" sub-directory of your Coffalyser digitalMLPA package.

The `-SampleBarcodes` parameter specifies which barcode collection file to use (optional; if not specified Coffalyser digitalMLPA will use the default from the configuration). The name of the SampleBarcodes can be found in the header of the files in the "_Configuration\Sample Barcodes" sub-directory of your Coffalyser digitalMLPA package.

The output of this conversion are Coffa files (*.coffa). The Coffa files are written to the output directory defined by the `-PathData` parameter. From here the Coffa files may be split up per experiment and distributed to the experiment owners for analysis.

The Coffa files may be analysed as they are. In this case, Coffalyser digitalMLPA will automatically create experiments by grouping all samples with the same digitalMLPA product together. All samples will be treated as test samples. In case your experiment contains samples of any other type (e.g. "No Dna", "Positive Reference", "Reference sample", "SD", "Undefined") you must define these in a Coffalyser Definition File. MRC Holland recommends always explicitly defining experiment type,

sample type, sample names, gender and if the sample is from a pooled DNA source to raise the likelihood of detecting mistakes. You can set the sample type in the Coffalyser digitalMLPA configuration editor ("MRC.DataAnalysis.DigitalMlpa.Definitions.Editor.exe") and by editing the *.cdf files.

If manual grouping of Coffa files is preferred, group Coffa files together in individual subfolders per experiment. Subsequently, the Sample Analysis may be called using the following command in the CLI:

```
MRC.DataAnalysis.DigitalMlpa.Analyses.ClientCli.exe -Definition "C:\Path1\" -  
AnalysisType FragmentAnalysis,ComparativeAnalysis
```

The -Definition parameter should point to an experiment directory with Coffa files.

After finishing the analysis, reports can be generated. The Reporting module may be called without a Coffalyser Definition File using the following command in the CLI:

```
MRC.DataAnalysis.DigitalMlpa.Reporting.ClientCli.exe -Definition "C:\Path1\"
```

The -Definition parameter should point to a directory with Coffa files that have already undergone Sample Analysis.

The .pdf Reports and the Excel ratio files are automatically generated in the directory containing the Coffa files.

Coffalyser digitalMLPA User Manual – Document History

Version 004 (13 July 2020)

- Manual completely restructured and rewritten.
- Name updated from Coffalyser digitalMLPA Manual to Coffalyser digitalMLPA User Manual.
- Screenshots updated where needed.
- Appendices on normalisation (I) and quality table (II) added.
- Layout changed.

Version 003 (25 November 2019)

- Added: this manual is valid from program versions.... In section “important notification”.
- Several minor textual changes.

Version 002 (29 October 2019)

- “Coffalyser.Net – digitalMLPA” replaced with “Coffalyser digitalMLPA” throughout document.
- Restructured setting up and running the data analysis into tables instead of text.
- Updated information where needed in compliance with changes done to Coffalyser digitalMLPA.
- All reports combined into one chapter (chapter 5).
- Relevant information moved to new section Data interpretation (chapter 6).
- Relevant information moved to new section Troubleshooting (chapter 7).
- Screenshots updated where needed.
- Several textual changes throughout the document.