Coffalyser digitalMLPA User Manual

For version 2.6.x

Instructions For Use

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

General information
This manual can be used for Coffalyser digitalMLPA version 2.6.0 and subsequent patches (2.6.1 etc.).

Coffalyser digitalMLPA™ is Research Use Only (RUO) software designed for the analysis of digitalMLPA™ data generated using a NXtec™ probemix* as described in the digitalMLPA NXtec Protocol and NXtec-MS Protocol.
* To be used in combination with a NXtec Reagent Kit, one or multiple NXtec Barcode Plates, NXtec-MS Hhal and SDXXX as specified in the Product Description of the application-specific probemix.

Latest version of this document can be found on www.mrcholland.com.

Manufacturer: MRC Holland BV, Willem Schoutenstraat 1, 1057 DL Amsterdam, The Netherlands
Website: www.mrcholland.com; 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 7
3.2.2 Reanalyse existing experiments 10
3.3 PREPARING COFFALYSER DEFINITION FILES 11
4. COFFALYSER digitalMLPA OUTPUT 12
4.1 OUTPUT FILES 12
4.2 INFORMATION IN REPORTS 13
4.2.1 Sample quality 13
4.2.2 Colour coding in reports 13
4.2.3 Shared information in reports 13
4.3 digitalMLPA RESULTS IN RESULTS PDF REPORT 21
4.3.1 Panels 21
4.3.2 Results section – Descriptive statistics row 21
4.3.3 Results section – Chart 22
4.3.4 Results section – Probe information 22
4.3.5 Results section – Comparison reference samples 23
4.3.6 Results section – Warning / error column 24
4.3.7 Results section – Aberrant column 27
4.3.8 Results section – Result column 28
4.3.9 Reference Sample Quality Issues table 28
4.4 NO-DNA PDF REPORT 30
4.5 EXCEL REPORT 31
4.6 XML REPORT 32
APPENDIX I – DATA ANALYSIS 33
NORMALISATION 33
Normalisation calculations 33
COMPARISON 34
APPENDIX II – ROOT CAUSE 35
APPENDIX III – TROUBLESHOOTING 38
GENERAL 38
EXPERIMENT DEFINITION 40
DATA ANALYSIS (COFFALYSER digitalMLPA – RUN WIZARD) 40
REPORTS 42
APPENDIX IV – COMMAND LINE INTERFACE (CLI) 44
CLI ANALYSIS WITH COFFALYSER DEFINITION FILE 44
Coffalyser Definition File header 44
Coffalyser Definition File experiment definition 45
Running Coffalyser Definition File in CLI 46
CLI ANALYSIS WITHOUT A COFFALYSER DEFINITION FILE 47
Data conversion without Coffalyser Definition File 47
Data analysis and reporting without Coffalyser Definition File 48

IMPORTANT NOTIFICATION

Always consult the most recent Product Description, Probe Information File and other probemix-specific files (if available) AND the digitalMLPA NXtec/NXtec-MS Protocol before use. These are available online at www.mrcholland.com. It is the responsibility of the user to be aware of the latest scientific knowledge of the application.

Coffalyser digitalMLPA has an expiration date. This date is displayed at 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 and via https://www.mrcholland.com/r/coffalyser-digitalmlpa-release-notes on the MRC Holland support portal.

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 to keep the contact information in your MRC Holland account up-to-date (www.mrcholland.com).

Software version is displayed on the About page and in the Coffalyser digitalMLPA Release Notes present in the "_Documentation" folder of the downloaded software package. The latest version of this manual is available for download on www.mrcholland.com. The manual is also included in the "_Documentation" folder in the downloaded software package. Please ensure you always use the latest version of the manual corresponding with your software version.

Please note that all data used and generated by Coffalyser digitalMLPA, like FASTQ files, PDF and Excel reports, Coffalyser Definition Files and Coffa files, may contain sensitive information. It is the responsibility of the user to ensure data is stored in secure locations to prevent unauthorized access to these data.

MRC Holland, SALSA, MLPA, digitalMLPA, NXtec, Coffalyser.Net, Coffalyser digitalMLPA, and their logos are trademarks or registered trademarks of MRC Holland BV. All other brands and names herein are the property of their respective owners.

1. GLOSSARY

Barcode	A barcode sequence is incorporated in digitalMLPA amplicons. The barcode should be unique for each sample within a sequencing run to allow assigning amplicons to the corresponding sample. Furthermore, barcodes ensure digitalMLPA data can be distinguished from other sequencing data.
Coffalyser Definition File	A file (*.cdf) containing the information required for the analysis of an experiment with Coffalyser digitalMLPA. This file specifies analysis settings, file locations, Product Sheets and other data such as sample names and sample types. It will be created when analysing your data using Coffalyser digitalMLPA via MRC.DataAnalysis.Client.exe. A *.cdf can also be created using the Definition File Editor (MRC.DataAnalysis.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.
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. These files can be used for reanalysis of samples.
Expected range	The range in which the result of a probe is expected to fall when no aberration is present. If the result is outside the expected range, it is considered aberrant.
FASTQ	A file containing the sequencing data obtained on a supported 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 – Data analysis 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 population. See Appendix I – Data analysis more information.
MIRR	Minimum Intra Ratio References. This is a threshold above which the intra ratios of a probe in all reference samples should be. In case no reference samples are defined, the median of the intra ratios of a probe over all samples should be above this threshold.
no-DNA sample	A sample from a no-DNA control reaction.
Panel	A subset of related probes in a digitalMLPA probemix. Including panels in the analysis will make their results visible in the reports. When panels are excluded from the analysis, their results will be masked and will not appear in the reports. This can be set in the Coffalyser Definition File. Note, some panels are mandatory and cannot be masked. Furthermore, certain panels in a probemix may only be shown for samples with specific characteristics (e.g. depending on sample sex).
Pooled DNA source	A sample that consists of a mixture of DNA samples.
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 lot.
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 step of normalisation to normalise other probes against (intra-normalisation). For most probes, the default set of reference probes that is present in the probemix is used for normalisation. However, some probes require a tailored normalisation to ensure proper analysis. For these probes a specific set of

	reference probes is used. There are various types of these specific reference probes, each designed for a particular analytical purpose. For example, most methylation-specific probes are normalised against a set of reference probes that are in close proximity to be able to give a methylation result that is not influenced by the copy number of the region. Please check the probemix-specific Product Description and Probe Information File for details.
Reference sample	A sample from an individual expected to have a normal copy number across all probe-targeted regions, please check the probemix-specific Product Description for more information. In most applications, reference samples are used in the second part of data normalisation to normalise other samples against (inter-normalisation).
Result	The result of a probe. What is displayed as the result depends on the probe type and application. Most copy number probes will have the inter ratio displayed as the result. For most methylation-specific probes, the result is an intra or inter ratio (dependent on the probemix application) that may be displayed as percentage. Mutation-specific probe results are commonly displayed as present/absent. For certain applications, a result is a combination of multiple probes results, e.g. the median. Please check the probemix-specific Product Description and Probe Information File for details.
Sample Track Probe	NOT CURRENTLY 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 detection of 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. Reference DNA.
Segment	A group of consecutive probes with similar results. A segment can also consist of a single probe.
SNP id code	Most digitalMLPA probemixes contain a set of SNP-specific probes that generate a SNP id code, also referred to as a *sample identifier* since DNA samples from the same individual will have the same identifier. This SNP id code consists of 39 characters, each corresponding to one specific SNP, that are always in the same order. Every position has four different options[a]: 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 the zygosity of the SNP is unclear, this results in a '*?*'. The presence of multiple '*?*' signs can indicate contamination of the sample DNA. This will be mentioned in the Quality table. Please note that a SNP id code is not listed, and the associated contamination check is not executed, when a sample is defined as '*pooled DNA source*'. Furthermore, for some probemixes the contamination check is not performed, as detection of sample DNA contamination is not always possible, e.g. in tumour derived samples. Please consult the probemix-specific Product Description for more information. The SNP id code ends with a version number, e.g. v1 for version 1. More information on these SNP probes is described in the digitalMLPA NXtec/NXtec-MS Protocol.
Test sample	A sample being investigated. For some applications test samples can be used in the second part of data normalisation (inter-normalisation) instead of defined reference samples. Please check the digitalMLPA NXtec/NXtec-MS Protocol and probemix-specific Product Description for details.
Undefined sample	A sample that is not specified in the experiment. An undefined sample will be analysed as a Test sample.

[a] Please note, '*#*' is printed in case there is no result for that position in the SNP id code.

2. BASIC CONCEPTS OF COFFALYSER digitalMLPA

The digitalMLPA data analysis flow consists of the following steps:

- **Data conversion:** Reads from a FASTQ file that are recognised as digitalMLPA reads are assigned to the digitalMLPA probes and counted. These reads are saved per digitalMLPA reaction in Coffa files (*.coffa, one file per digitalMLPA barcode).
- **Analysis:** The obtained digitalMLPA probe read counts are subjected to normalisation and comparison, see Appendix I – Data analysis for details.
- **Reporting:** Reports are generated per experiment and per sample (see chapter 4 Coffalyser digitalMLPA output for more information).

It is possible to generate Coffa files as a separate step via a command line interface. 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 (64-bit, x64) that is still supported by Microsoft, with the latest updates installed.
- .NET Framework version 4.8.
- 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.
- Read and write access to the directory the software is run from.

For optimal operation, we recommend:

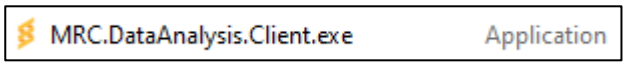
- An Intel i7 processor or equivalent.
- 8 GB of available RAM.
- A solid state drive (SSD) with sufficient free disk space.

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. ANALYSIS OF VERY LARGE FASTQ FILES MAY TAKE A CONSIDERABLE AMOUNT OF TIME. THIS TIME CAN BE SHORTENED BY USING A MORE POWERFUL COMPUTER.
4. WHEN COFFALYSER digitalMLPA IS STORED ON A LOCATION WITH A VERY LONG PATH, THIS MAY CAUSE ISSUES RUNNING THE SOFTWARE. AVOID PLACING THE SOFTWARE IN DEEPLY NESTED FOLDERS.
5. COFFALYSER digitalMLPA CAN BE RUN FROM A NETWORK LOCATION. THIS ALLOWS MULTIPLE USERS TO USE THE SAME VERSION, AND CENTRAL VERSION MANAGEMENT. HOWEVER, NETWORK SECURITY SETTINGS OR VIRUS SCANNERS MAY PREVENT COFFALYSER digitalMLPA FROM RUNNING FROM NETWORK LOCATIONS. FURTHERMORE, RUNNING FROM NETWORK LOCATION MAY REDUCE THE SPEED OF THE ANALYSIS.
6. DUE TO SECURITY SETTINGS IN WINDOWS 11, COFFALYSER digitalMLPA EXECUTABLES COULD BE BLOCKED. THIS CAN BE RESOLVED BY RIGHT CLICKING ON THE .EXE, SELECTING PROPERTIES AND TICKING 'UNBLOCK'.

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 in the desired location.
5. Unzip the file (7-Zip is recommended).
6. Optional: Create a shortcut for the MRC.DataAnalysis.Client.exe and place it on the desktop.



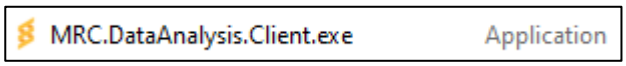
3.2 ANALYSE YOUR DATA

⚠ PLEASE NOTE THAT DATA ANALYSIS BY COFFALYSER digitalMLPA MAY RANGE FROM SEVERAL MINUTES TO A FEW HOURS. THIS DEPENDS ON VARIOUS FACTORS, INCLUDING COMPUTER PROPERTIES, NETWORK STABILITY AND SPEED, FASTQ FILE SIZE, EXPERIMENT SIZE AND NUMBER OF UNRECOGNISED READS.
WE RECOMMEND USING A LOCAL DRIVE AS TEMP DIRECTORY (E.G. C:\) AND PREFERABLY ALSO FOR THE STORAGE OF THE SOFTWARE PACKAGE AND OUTPUT DIRECTORY. THE FASTQ FILE CAN BE STORED ON A NETWORK DRIVE, BUT IF YOU RUN INTO ISSUES WITH NETWORK STABILITY OR SPEED, PLEASE CONSULT YOUR NETWORK ADMINISTRATOR.

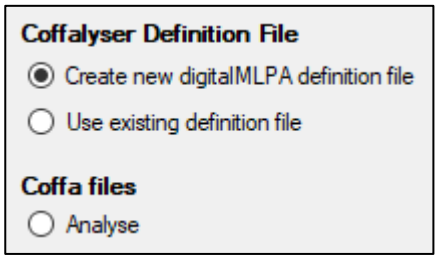
Coffalyser digitalMLPA can perform an analysis using zipped FASTQ files. Unzipping FASTQ files before analysis or copying a FASTQ file to a local disk is not necessary.

3.2.1 Create a new experiment

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



2. Under Coffalyser Definition File, select '*Create new digitalMLPA 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**.

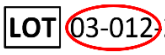
The Coffalyser - Data Conversion window will open.

3.2.1.1 Configure general settings

1. Leave the settings profile on '*Default profile*'.

2. Select the appropriate barcode collection from the '*barcode collection*' drop-down menu. In most cases, the barcode collection '*From lot 03-009-yymmdd onwards (Default)*' is applicable. Only select a different barcode collection if either of the following applies:
   - You are analysing data that includes barcode plate lot 03-008-yymmdd. In that case, select '*Up to and including lot 03-012-yymmdd (not compatible with XLEAP-SBS)*'.
   - Your sequencing run uses a custom read 1 primer, resulting in demultiplexed FASTQ files (more information on www.mrcholland.com). In that case, select a barcode collection ending with '*-demultiplexed*'.

   The selected barcode collection should include all barcode plate lots used in your experiment. The lot number is printed on the barcode plate label.

3. Click **Configure** to set up your experiment.[b]

   The Coffalyser - Samples window will open.

4. Sample selection:
   a) Add samples to a new experiment:
      Select all barcodes used in your experiment. Right click on any of the selected barcodes and select '*Create New Experiment (Manual)*' and continue with step 6.
   b) Add samples to an existing experiment:
      Select the barcodes not yet included in the experiment(s). Right click on any of the selected barcodes and select '*Add To Experiment*'. Select the applicable experiment from the list and continue with step 5.
   c) Remove samples from an existing experiment:
      Select the barcodes to be removed. Right click on any of the selected barcodes and select '*Remove From Experiment*'. If you want to change experiment or sample settings, continue with step 5, otherwise continue with step 20.

5. Right click on a sample included in the experiment you want to edit and select '*Edit Experiment*'.

   The Coffalyser - Experiments window will open.

6. Enter or adjust the experiment name in the designated text field.

7. Select the applicable digitalMLPA probemix from the '*product*' drop-down menu.

8. Select the applicable version and lot number of the digitalMLPA probemix from the '*sheet*' drop-down menu.

9. When desired, enter the name for each sample in the column '*sample*'.

   You can paste sample names (and sample type and sample sex) from clipboard (e.g. from Excel or Notepad source), where source file should contain a list of desired sample names and optionally columns or tab separated data on sample type and sex. Right click on the first row of the range of samples and select '*Paste Sample Details From Clipboard*'.

10. Select or adjust the sample type for each sample from the drop-down menu in the column '*type*'.

    You can also do this batch-wise: select the desired samples and type '*n*' for no-DNA, '*r*' for Reference, '*s*' for SD, '*t*' for Test or '*u*' for Undefined.

[b] In 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.

11. For samples that consist of a mixture of DNA samples, check the selection box in the column '*pooled DNA*'. For samples from a single source, leave the selection box unchecked (default setting).

    You can also check or uncheck the pooled DNA box per sample type: right click on the cell and select '*Mark Samples As*' and choose to set all Reference, Test or SD samples as '*Pooled DNA*' or '*Not Pooled DNA*'.

12. Check the selection box in the column '*digested*' for all samples to which the HhaI restriction enzyme has been added.

    You can also check or uncheck the digested box per sample type: right click on the cell and select '*Mark Samples As*' and choose to set all Reference, Test or SD samples as '*Digested*' or '*Not Digested*'.

13. Optional: select the sex of each sample from the drop-down menu in the column *sex*.

    You can also do this batch-wise: select the desired samples and type '*f*' for Female or '*m*' for Male.

TO USE PANELS, CONTINUE WITH STEP 14. OTHERWISE PROCEED WITH STEP 20 TO CONTINUE WITH THE ANALYSIS.

14. Right click in the column *panel(s)* and select '*Select Panels*'.

15. Select a panel set from the list and continue with step 19, or click '*Create New Panel Selection*'.

    The Coffalyser – Panel Set window will open.

16. Enter a name for the panel set in the designated text field.

17. Deselect the panels that are not of interest.

    Note, some panels are mandatory and cannot be deselected. Furthermore, certain panels in a probemix may be masked automatically dependent on specific sample characteristics (e.g. depending on sample sex). Details on panel content can be found in the probemix-specific Probe Information File.

18. Click **OK** to save and exit the '*Coffalyser – Panel Set*' window.

19. Repeat steps 14 to 18 for other samples, if applicable. Otherwise click **OK** to save and exit the '*Coffalyser – Experiments window*'.

REPEAT STEPS 4-19 FOR YOUR ADDITIONAL EXPERIMENT(S).

20. Click **OK** to save and exit the '*Coffalyser – Samples* window'.

21. From the '*undefined samples*' drop-down menu choose one of these options:
    - '*Detect and include*' to analyse reads from samples (barcodes) that are not defined in the steps above (their output will be collected in an Undefined Experiment folder in the Results folder).
    - '*Ignore and exclude*' to only analyse the reads of the defined samples.

22. Select or change the desired output files from the '*output types*' drop-down menu.

    For details on the output files that can be generated, see section 4.1.

23. Click the upper **Select** button to select or change the '*output directory*'. Browse to the desired folder, select it and click **OK**. Coffalyser digitalMLPA will store the output in a sub-directory named using the date and time of the analysis.

24. Click the lower **Select** button to select or change the '*temp directory*'. Browse to the desired folder, select it and click **OK**.

THE SPEED OF THE ANALYSIS IS GREATLY INFLUENCED BY THE LOCATION OF THE TEMP DIRECTORY. A NETWORK DRIVE IS NOT RECOMMENDED FOR THIS PURPOSE.

25. Click **Add File**, browse to the location of the FASTQ file(s), select it and click **Open**.

    When selecting multiple FASTQ files, ensure that barcodes are unique or identical barcodes belong to the same experiment and sample.

TO EXCLUDE A FASTQ FILE FROM THE ANALYSIS, SELECT IT IN THE *INPUT FILE(S)* FIELD AND CLICK **REMOVE FILE**.

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

    A progress bar will appear in the Coffalyser - Run Wizard.

27. After analysis has finished, click **Open Results Folder** to access the experiment subfolder in the output directory.

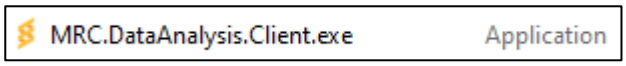
3.2.2 Reanalyse existing experiments

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

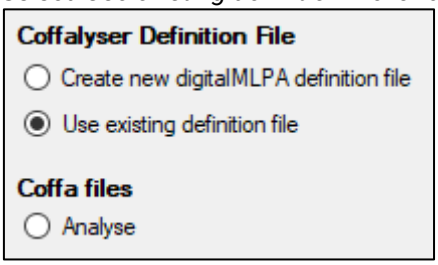
- When the wrong probemix or lot number was selected in the first analysis.
- When the wrong barcode collection was selected in the first analysis.
- When you want to change the output directory and/or temp directory.
- When Coffa files were generated using a different version of the Coffalyser digitalMLPA software that is not compatible with the version used for reanalysis.

PLEASE DO NOT TO USE COFFA FILES FOR LONG-TERM DATA STORAGE, AS THEIR COMPATIBILITY BETWEEN SOFTWARE VERSIONS IS NOT GUARANTEED. THE FASTQ FILE SHOULD BE KEPT AS YOUR RAW DATA SOURCE.

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



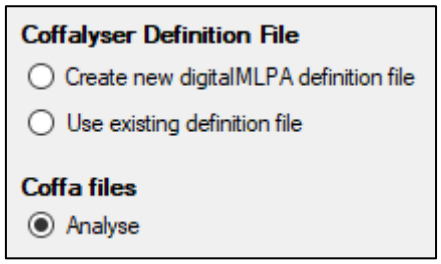
2. Reanalysing data.
   a) To start from an existing Coffalyser Definition File:
      i. Select '*Use existing definition file*' under Coffalyser Definition File and click **Next >**.



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

      The Coffalyser - Data Conversion window will open.

   b) To start from Coffa files:
      i. Under Coffa data files select *Analyse* and click **Next >**.



ii. Browse to the folder that contains the Coffa files you wish to reanalyse, select it and click **OK**.

The Coffalyser - Data Analysis window will open.

3. Configure the general settings and experiment(s) as described in chapter 3.2.1.1 Configure general settings.

   Please note that when reanalysing Coffa files, only the settings that are allowed to be changed are displayed.

3.3 PREPARING COFFALYSER DEFINITION FILES

Coffalyser digitalMLPA also allows to make or edit a Coffalyser Definition File without analysing it immediately, using the *MRC.DataAnalysis.Definitions.Editor.exe*. The Coffalyser Definition File can be used at a later time point as explained in section 3.2.2, or as input when using the Command Line Interface (see Appendix IV – Command Line Interface (CLI)).

For instructions on how to make or edit the Coffalyser Definition File via *MRC.DataAnalysis.Definitions.Editor.exe*, see section 3.2.

4. COFFALYSER digitalMLPA OUTPUT

4.1 OUTPUT FILES

Coffalyser digitalMLPA can generate different types of output files after data analysis. Which files are generated depends on the option selected in the drop-down menu *output types* during configuration of the experiment(s).

Types of output per experiment:

- **Quality Report (PDF report)**
  File name: [Experiment name]_Quality Report.pdf
  This file provides an overview of the samples included in the experiment as well as extensive information on the quality assessment of the samples and the analysis. More detailed information about the samples in which issues were detected are displayed in the bottom part of the report. See section 4.2 for details.

⚠ IT IS RECOMMENDED TO OPEN THE QUALITY REPORT.PDF FILE FIRST TO ASSESS THE QUALITY OF THE ANALYSIS. HEREAFTER, THE RESULTS.PDF FILES PER SAMPLE CAN BE VIEWED TO EXAMINE AND INTERPRET THE ANALYSIS RESULTS.

- **Excel Report (Excel file)**
  File name: [Experiment name]_Excel Report.xlsx
  This file provides an overview of the analysis results, read counts and quality assessment of all samples included in the experiment. See section 4.2 and 4.5 for details.

⚠ PLEASE BE AWARE THAT THE EXCEL REPORT DOES NOT CONTAIN DETAILS ON THE STATISTICAL SIGNIFICANCE OF RESULTS. IT IS HIGHLY RECOMMENDED TO CONSULT THE SAMPLE-SPECIFIC PDF REPORTS AS WELL.

Types of output per sample:

- **Results (PDF report)**
  File name: [Sample name]_Results.pdf
  This file contains the analysis results of a sample as well as a summary of the quality assessment, see sections 4.2 and 4.3 for details.
- **no-DNA (PDF report)**
  File name: [Sample name]_no-DNA.pdf
  This file is generated for no-DNA samples and provides general information, see section 4.4 for details.
- **Coffa file**
  File name: [Sample name].Coffa
  This is a Coffalyser digitalMLPA specific output file that can be used for quick reanalysis of digitalMLPA data within the same software version, see section 3.2.2 for details.
- **XML file**
  File name: [Sample name].xml
  This file contains sample information, analysis results and quality assessment. It is intended to be used in a custom analysis pipeline. See section 4.6 for details.

4.2 INFORMATION IN REPORTS

4.2.1 Sample quality

The overall outcome of the sample quality evaluation is displayed as either '*passed*' or '*failed*'. When the sample quality is failed, the results are considered unreliable. If a sample is passed but there are other warnings or errors, data should be interpreted with caution.

This overall sample quality evaluation is based on the cumulative observations from all quality checks (☑) as described in Table 1 below. More information on the various quality checks and troubleshooting can be found via https://www.mrcholland.com/r/ifu/digitalmlpa-troubleshooting on the MRC Holland support portal.

4.2.2 Colour coding in reports

Throughout the different reports, the following colour coding is used:

- **Red** font: Errors
- **Orange** font: Aberrant results and warnings
- Grey cells: Not applicable or not performed.

4.2.3 Shared information in reports

Information displayed on the various reports is tailored to the envisioned purpose of the report. Some information is shared between reports, while other information is present in a certain report, and only summarised or absent in other reports. The information in the different reports is displayed in Table 1 below.

Symbols used in Table 1:

☑	Check mark:	This feature is (also) functioning as a quality check for digitalMLPA data and contributing to the overall sample quality evaluation (passed/failed) that determines whether the data is of sufficient quality for interpretation.
●	Solid bullet:	Information is present in report.
○	Open bullet:	Summary of information is present in report.
-	Dash:	Information is absent in report.

Table 1. Information shared between reports.

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Experiment information					
Experiment name	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 *Nameless* - or *Undefined Experiment 01* etc.	●	-	●	●
Product name	This displays the name of the Product Sheet used for the analysis.	●	-	●	●
Normalisation options	This displays the normalisation method used.	●	-	●[c]	-

[c] *Normalisation options* and *Barcode collection* are present in section *Other* at bottom of Sample PDF report.

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Barcode collection	This displays the barcode collection number (or version). This should match with the first two digits of the barcode plate lot number that is printed on the barcode plates used in the experiment.	●	-	●[c]	-
Analysed panels	This displays the panels that were analysed. The panel selection can be made during experiment definition.	-	-	●	-
Masked panels	This displays the panels for which the results are hidden. The panel selection can be made during experiment definition.[d] ⚠ PLEASE NOTE THAT RESULTS OF THE PANELS LISTED HERE ARE NOT REPORTED. ENSURE THE PANEL SELECTION IS AS INTENDED.	-	-	●	-
Software versions					
Software version	This displays the Coffalyser digitalMLPA version (major.minor.patch) and the versions of the individual software modules used in the analysis. In the Coffalyser digitalMLPA Release Notes, changes between versions are explained.	●	-	●	●
Executed	This displays the execution time of the Coffalyser digitalMLPA software modules.	-	-	●	-
Sample information					
Barcode	This displays the name of the barcode used for the sample.	●	●	●	●
Sample name	This displays the name that was specified for the sample in the Coffalyser Definition File. If no name was specified, the name of the barcode will be displayed.	●	●	●	●
Sample type	This displays the sample type based on what was specified in the Coffalyser Definition File. • *Test*: Test sample • *Test [D]*: Digested test sample • *Reference*: Reference sample • *Reference [D]*: Digested reference sample • *SD*: Sample DNA • *no-DNA*: no-DNA sample • *Custom*: Other sample types, e.g. when sample type is '*Undefined*' • *Pooled DNA*: Displayed if the sample is specified as *pooled DNA* during experiment definition	○[e]	●	●	●

[d] A panel can be set as mandatory for a probemix lot and may thus not be masked by the user. Furthermore, certain panels in a probemix may be masked automatically depending on specific sample characteristics (e.g. sample sex).

[e] Only *Pooled DNA* is not present in the top overview part of the Experiment Quality PDF report.

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Sex	☑ This displays the sex of the sample used for analysis. The sex can be defined by the user in the Coffalyser Definition File as '*Male*' or '*Female*', or be left as '*Undefined*'. In addition, Coffalyser digitalMLPA will attempt to detect the sex of the sample based on the read counts of specific control probes[f]. If Coffalyser digitalMLPA cannot reliably determine the sex (e.g. only one Y-control probe is found while three Y-control probes are expected in male samples), the detected sex will be '*Ambiguous*'. If the sex is defined in the Coffalyser Definition File and there is a discrepancy with the sex detected by Coffalyser digitalMLPA, this will be indicated here (note that the defined sex will be used for analysis). If sex is neither defined nor detected, analysis of the sample cannot commence, and an exception will be shown under program errors in the Experiment Quality and Sample PDF reports. The Excel Report shows the sex that is defined by the user, detected by Coffalyser digitalMLPA and used for analysis. Note that certain probe results may or may not be shown depending on the sample sex used for analysis.	○	●	●	●
Id	This displays the unique sample identification code to aid in traceability of samples between the Experiment Quality and the Sample PDF reports.	○	●	●[g]	-
Overall quality					
Quality checks / QC	☑ This displays the overall outcome of the sample quality evaluation. The result can be either '*failed*' or '*passed*'. When the sample quality is failed, the results are not reliable. If a sample is passed but there are other warnings or errors, data should be interpreted with caution.	●	●	●	●
Plausible root cause	This displays the possible reason for issue(s) detected in a sample. If multiple issues were detected, only the cause with highest importance is displayed. See Appendix II – Root cause for more information on the plausible root cause.	-	●	-	●
Program information					
Invalid signatures	☑ This displays whether all Coffalyser digitalMLPA files in the package used for the data analysis are identical to those provided by MRC Holland. If it states '*None*', this means these files are in their original form.	●[h]	●	○	●

[f] Sex detection is only based on X- and Y-control probe signals being present or absent. Sex-chromosome aneuploidies (e.g. Turner syndrome (X0), Klinefelter syndrome (XXY) or Trisomy X (XXX)) are not detected here.

[g] *Sample id* is present in section *Other* at bottom of Sample PDF report.

[h] *Invalid signatures* is present in section *Experiment* in Experiment Quality PDF report.

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Program errors	☑ This displays whether errors occurred while running the analysis software. When program errors occurred during analysis, the number of errors as well as the first error are displayed in the samples section in the Experiment Quality PDF report. The complete list of errors is displayed in the Results PDF report.	○	●	●	-
Sequence run					
Minimum reads	☑ This is a measure for the minimum number of reads assigned to a barcode. When the minimum number is not reached, an error will appear. Ensure the experiment has been performed according to the instructions in the digitalMLPA NXtec/NXtec-MS Protocol and check whether all your samples are correctly defined in the Coffalyser Definition File.	●	●	○	●
MTRQ & MDRQ	☑ These are measures for the read depth and number of independent ligation events. MTRQ is the median number of total reads assigned to the reference probes in a sample. This is a measure for the read depth in a sample. MDRQ is the median number of distinct reads assigned to the reference probes in a sample. Distinct reads are based on the number of independent ligation events (recognised by unique molecular identifiers; UMIs). When MTRQ and/or MDRQ are too low, an error is displayed as data analysis may not be reliable.	○	●	○	●
Unrecognised reads	☑ 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 Product Sheet, a warning or error is displayed.	●	●	○	●
Sequence quality	☑ This quality classification is based on the quality of the sequencing reads. When many reads contain errors or when read lengths in the FASTQ file are not identical, a warning or error is shown.	●	●	○	●
Sample uniformity	☑ 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 median of the other samples in the experiment, a warning or error is displayed.	●	●	○	●
Identification					
SNP id code	☑ This displays the detected SNP id code, or sample identifier, of a sample and the version number of the SNP probe collection. See Glossary for details. The DNA contamination check (see below) is based on this code.	○	●	●	●[i]

[i] *SNP id code* is present in section *Sample details* in the Excel Report.

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report

Probemix lot	☑ This displays whether probemix identity detection was successful. The software will detect the probemix lot that was used based on a set of identifiers of which the combination is unique for a probemix lot and corresponding Product Sheet. A warning or error can be triggered.
	• *Identity detected:* Product Sheet was successfully detected by the software and corresponds to the Product Sheet selected in the Coffalyser Definition File (if one was specified).
	• *Identity detected (≥1 disqualified identifiers):* Product Sheet detection by the software was successful, but identifier(s) with low read counts are observed.
	  ○ In case of a warning (orange): In rare cases identifier probes may give lower read counts than expected. Ensure the experiment is performed according to protocol and contact MRC Holland at info@mrcholland.com.
	  ○ In case of an error (red): Ensure there is no contamination during your experiment, e.g. from another probemix lot.
	• *Identity NOT detected (0 or ≥1 disqualified identifiers):* The detected identifiers do not correspond to a Product Sheet (probemix lot) that is known by the software. If there are disqualified identifier(s), there are identifier(s) observed with low read counts. Ensure you are using the latest software version and ensure there is no contamination during your experiment, e.g. from another probemix lot. In rare cases identifier probes may give lower read counts than expected. Ensure the experiment is performed according to protocol and contact MRC Holland at info@mrcholland.com.
	• *CONFLICTING identity detected (0 or ≥1 disqualified identifiers):* The Product Sheet (probemix lot) selected in the Coffalyser Definition File does not correspond to the probemix lot that is detected by the software. If there are disqualified identifier(s), there are identifier(s) observed with low read counts. Ensure the correct Product Sheet is selected in the Coffalyser Definition File and that there is no contamination during your experiment, e.g. from another probemix lot.
		○	●	●	○

Barcode lot	☑ This displays the detected barcode lot based on an identifier probe. This number corresponds to the information on the barcode plate. Compatibility of the barcode lot with the probemix is checked: • The error *known as incompatible* is triggered in case the used barcode plate and/or lot is incompatible with the probemix lot. • A warning *not known as compatible* is triggered in case the compatibility between the barcode plate and/or lot with the probemix lot is unknown. Please consult the probemix-specific Product Description to check which barcode plates lots are compatible. And ensure to use the latest version of Coffalyser digitalMLPA. In case multiple barcode lots are detected, the error *too many* is triggered. The three-digit number between dashes (e.g. -009-) indicates the barcode plate lot. Ensure you are using one barcode plate lot per sample.	○	●	●	●

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Sample tracker	☑ This displays the detected Sample Track Probe(s). An error is given in case the Sample Track Probe(s) that are detected differ from what is specified in the Coffalyser Definition File. Note: Sample Track Probes are not currently available.	○	●	●	●
Reaction					
Intra & inter	In the sample details part of the Experiment Quality PDF report, the reaction checks are split into *intra* and *inter*. The intra checks focus on the quality of each individual sample. A warning or error is triggered in case the tested condition is outside the set boundaries. E.g. an intra error on denaturation indicates that sample DNA denaturation was incomplete. The inter checks focus on the comparison of the sample to the reference population. A warning or error is triggered in case the tested condition differs between the sample and the reference population. E.g. an inter error on denaturation indicates that the sample DNA is not denatured to a similar extent as the DNA of the reference population.	-	●	-	●
DNA quantity	☑ For a DNA sample, this is an indication whether sufficient input DNA was used in the digitalMLPA experiment. If the MDRQ is too low (see MTRQ & MDRQ above), no estimation of DNA quantity can take place and an error will be shown here. For a no-DNA sample, this summarises whether the no-DNA quality evaluation was within acceptable thresholds.	○	●	○	●
DNA contamination	☑ This quality classification is based on the SNP id code and checks if the DNA sample is of a single source. Please note that this contamination check is not performed when a sample is defined as '*pooled DNA*'. Furthermore, for some probemixes the contamination check is not performed, as detection of sample DNA contamination is not always possible, e.g. in tumour-derived samples. Please consult the probemix-specific Product Description for more information.	○	●	○	●
DNA denaturation	☑ This quality classification is based on the set of control probes in a probemix that are sensitive to sample DNA denaturation.	○	●	○	●
DNA depurination	☑ This quality classification is based on the set of control probes in a probemix that are sensitive to sample DNA depurination.	○	●	○	●
DNA fragmentation	☑ This quality classification is based on the set of control probes in a probemix that are sensitive to sample DNA fragmentation.	○	●	○	●
DNA digestion	☑ This quality classification is based on the set of control probes in a probemix that are sensitive to digestion. Undigested samples are checked to really be undigested, and digested samples are checked to really be digested.	○	●	○	●

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
Hybridisation temperature	☑ This quality classification is based on the set of control probes in a probemix that check on the overnight hybridisation condition.	○	●	○	●
Hybridisation completeness	☑ This quality classification is based on the set of control probes in a probemix that check the extent of the hybridisation reaction.	○	●	○	●
Ligase start temperature	☑ 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 instead of 48°C.	○	●	○	●
Ligase activity	☑ This quality classification is based on the set of control probes in a probemix that check the ligase activity.	○	●	○	●
Ligase inactivation	☑ This quality classification is based on the set of control probes in a probemix that detect incomplete ligase inactivation.	○	●	○	●
Polymerase activity	☑ This quality classification is based on the set of control probes in a probemix that check the polymerase activity.	○	●	○	●
Data analysis					
Normalisation & comparison	In the sample details part of the Experiment Quality PDF report, the data analysis checks are split into *normalisation* and *comparison*. The normalisation checks focus on the normalisation part of data analysis and the comparison checks focus on the comparison part, see Appendix I – Data analysis.	-	●	-	●
Reference probe quality	☑ This quality classification is based on the result of the reference probes that are used for normalising the test probe. The reference probe quality (RPQ) warning or error on experiment level is based on the RPQ detected for individual probes, see section 4.3.6. If there is an RPQ warning or error related to the default set of reference probes, a warning or error will be triggered on experiment level respectively. If there is an RPQ warning or error related to a specific set of reference probes (e.g. local references used for normalising most methylation-specific probes), a warning will be triggered on experiment level.	○	●	●	●
Reference sample quality	☑ This quality classification is based on the reference population used for normalisation and/or comparison. The reference sample quality (RSQ) warning on experiment level is based on the RSQ (excl), RSQ (MIRR) and RSQ (range) detected for individual probes, see section 4.3.6.	○	●	●	●
Reference sample quantity	☑ This quality classification is based on the size of the reference population used for normalisation and/or comparison. The reference sample quantity warning or error on experiment level is based on the RSQ (qty) detected for individual probes, see section 4.3.6.	○	●	●	●

Description	Explanation	Experiment Quality PDF: Overview table	Experiment Quality PDF: Sample details	Results PDF	Excel Report
no-DNA control(s)	☑ If no no-DNA control is included in the experiment, the cell is grey. If one or more no-DNA controls are included in the experiment, the quality evaluation of the no-DNA checks is summarised in the cell. If '*Ok*' is displayed, the no-DNA controls are within acceptable thresholds. If '*Error*' is displayed, please open the no-DNA.pdf report for more information.	●[j]	●	-	●
File information					
File name	This displays the name of the Coffa file.	○	●	●[k]	-
Source file(s)	This displays the location and name(s) of the FASTQ file(s) used.	○	●	●[k]	-

[j] *No-DNA control(s)* is present in section *Experiment* in Experiment Quality PDF report.

[k] *File name* and *Source file(s)* are present in section *Other* at bottom of Sample PDF report.

4.3 digitalMLPA RESULTS IN RESULTS PDF REPORT

The digitalMLPA results are reported in the Results PDF report. This report is tailored to the envisioned purpose of the probemix. Therefore, the content may be slightly different between reports from different probemixes.

4.3.1 Panels

The results shown in the Results PDF report are organised in panels (see Table 2). A panel contains related probes grouped per probe type, target gene or related disorder, as listed in the probemix-specific Probe Information File. During experiment definition the user can specify to have a panel analysed or have a panel masked.[I] Results of a masked panel are not shown on the reports. Furthermore, panels are set up in a manner that aids result interpretation, and only results that are considered relevant are shown on the report. This may for example mean that non-aberrant probe results are not shown on the Results PDF report.

Table 2. Lay-out of the results section in the Results PDF report.

Description	Explanation
Panel name	The user can specify to analyse or mask a panel during experiment definition. If there are subpanels within the panel that contain aberrant probe results, this is displayed in orange text '*[n aberrant subpanel(s)]*'.
Subpanel name	A panel is divided in one or more subpanels. If there are aberrant probe results within the subpanel, these may be grouped together in segments and the number of aberrant segments is displayed in orange text '*[n aberrant segment(s)]*'.
Segment	The probe results within the subpanel may be displayed in one or more segments. If multiple segments are shown, these are based on whether the result is considered aberrant or not. If applicable, the number of aberrant probes within a segment is displayed in orange text '*[n aberrant probes]*'.
Chart	For certain panels, subpanels or segments, a chart visualising the probe results may be displayed.
Probe results	Table of results per probe that are part of the panel, subpanel or segment.

4.3.2 Results section – Descriptive statistics row

When considered relevant for the subpanel, there is a row with descriptive statistics shown above the probe results. The information in this row is explained in Table 3.

Table 3. Descriptive statistics of subpanel or segment.

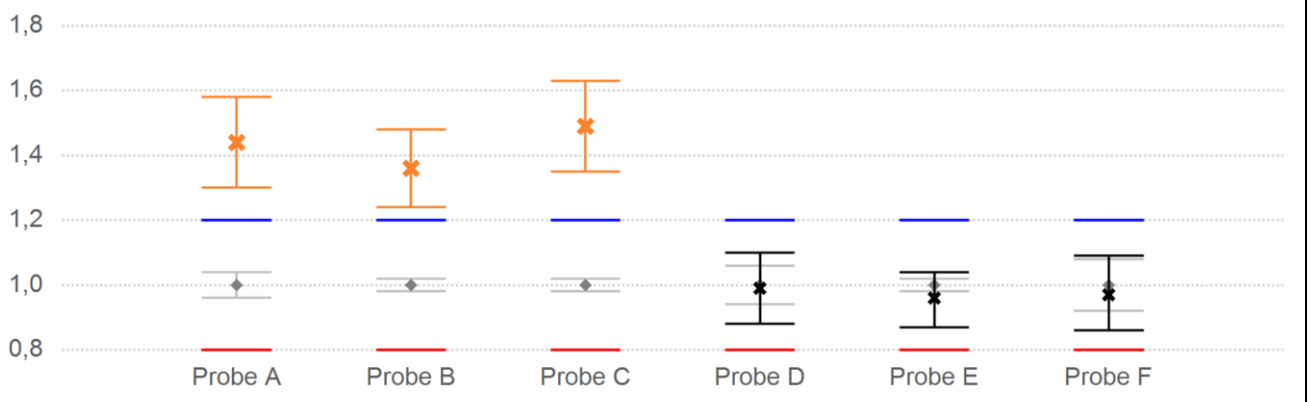
Description	Explanation
Size	This displays the genome size of the subpanel or segment when all probes listed are located on one chromosome.
Median	Median of the results for probes present in the subpanel or segment.
Mean	Mean of the results for probes present in the subpanel or segment.
St. dev.	Standard deviation of the results for probes present in the subpanel or segment.
Minimum	Lowest result of the probes present in the subpanel or segment.
Maximum	Highest result of the probes present in the subpanel or segment.
Range	Range (difference between minimum and maximum) of the results for probes present in the subpanel or segment.

[I] A panel can be set as mandatory for a probemix lot and may thus not be masked by the user. Furthermore, certain panels in a probemix may be masked automatically depending on specific sample characteristics (e.g. sample sex).

4.3.3 Results section – Chart

Probe results are visualised in a chart when considered relevant for the probemix. The chart characteristics are shown in Table 4. Please note that not all charts may contain all elements.

Table 4. Chart lay-out in Results PDF report.



Description	Explanation
X-axis	The probes are represented on the X-axis.
Y-axis	The probe result is displayed on the Y-axis.
– Blue line:	The upper boundary of the expected range.
– Red line:	The lower boundary of the expected range.
x Cross with error bars:	The cross represents the probe result, and the error bars represent the 95% confidence range of this result.[m] This is displayed in orange in case the probe result is outside the expected range (probes A-C above), and **black** in case it lies within the expected range (probes D-F above).
◆ Diamond with error bars:	The diamond represents the mean (median in case there are no defined reference samples) probe results in the reference population used during comparison. The error bars represent the 95% confidence range.

4.3.4 Results section – Probe information

In Table 5 the probe information displayed in the Results PDF report is explained.

Table 5. Probe information in Results PDF report.

Description	Explanation
Number	This displays the probe number of the digitalMLPA probe as designated by MRC Holland.

[m] For some probe results, the 95% confidence range may not be available.

Description	Explanation
Function	This displays the probe function. • *CN*: Copy number probe • *CN/WT*: Wildtype probe. This probe detects the normal allele (wildtype sequence) of a mutation. • *ME*: Methylation-specific probe. This probe contains an HhaI site and will not generate reads when this HhaI site is unmethylated. • *MUT*: Mutation-specific probe. This probe will only generate reads when the mutation is present. • *SNP*: SNP probe. This probe detects the presence of one of the two or more possible alleles of the single nucleotide polymorphism.
Location (kb)	This displays the mapview location of the probe's target sequence, based on genome build hg38.
Chr. band	This displays the chromosomal band location of the probe's target sequence, based on genome build hg38.
Gene	This displays the name of the gene that the digitalMLPA probe targets.
Exon	This displays the (nearest) exon number of the gene listed in column Gene that the digitalMLPA probe targets.

4.3.5 Results section – Comparison reference samples

Table 6 explains the information about the reference population used during comparison that is displayed in the Results PDF report. For more information on normalisation and comparison, see Appendix I – Data analysis.

Table 6. Comparison reference samples in Results PDF report.

Description	Explanation

Type	This displays the sample type population to which the probe is compared.
	• *Reference*:
	  ○ If '*Reference*' is displayed and there are defined reference samples in the experiment, these samples are used for comparison of this probe.
	  ○ If '*Reference*' is displayed and there are no reference samples defined, the test samples are used for comparison of this probe (if this functionality is enabled for the probemix).
	  ○ If '*Reference*' is displayed for methylation-specific digitalMLPA, both digested and undigested reference samples are used for comparison of this probe.
	• *Reference [D]*: Digested reference samples are used for comparison of this probe.
	• *Reference [U]*: Undigested reference samples are used for comparison of this probe.
	• *SD*: The SD is used for comparison of this probe.
	• *Custom*: Check the probemix-specific Product Description and Probe Information File to see what samples are used for comparison of this probe.

CN exp.	This displays the expected copy number of the target sequence of a probe in samples from unaffected individuals.

Description	Explanation
Range exp.	This displays the expected range of the probe result in unaffected individuals. A circumflex (^) next to a border of the expected range indicates that this border was widened according to the results obtained in the experiment. Such dynamic borders, allowing the widening of the expected range, are typically used for certain methylation-specific probes. Consult the probemix-specific Product Description to find whether widening of the expected range is enabled. When widening is enabled, and the 95% (confidence) range of the probe signal in reference samples extends beyond the initial expected range, one or both of the borders are extended to include the full 95% (confidence) range.
Range 95%	This displays the 95% confidence range of the reference population used for comparison in the experiment.

4.3.6 Results section – Warning / error column

Checking data quality is essential prior to result interpretation. Therefore, the most important error or warning related to the probe result is displayed (if applicable) in the *warning / error* column of the results section of the Results report. There can be more errors related to one probe result, but the one with highest priority is displayed. In the *Reference Sample Quality Issues* table (see section 4.3.9) at the bottom of the Results report, all Reference Sample Quality (RSQ) errors are shown per probe. Please note that not all probes can trigger all error and warning types. This is dependent on the probemix and probe.

The errors and warnings that can occur, and their priority, are shown in Table 7. In case an error or warning is present, the data in columns *aberrant* and *result* are displayed in grey.

Table 7. Error types in results section.

Priority	Error type	Explanation [n]	How to proceed
1	**nRSQ (excl)**	*Normalisation Reference Sample Quality Excluded error.* Error during normalisation. A defined reference sample is not used in the normalisation, for example because probe signal was too low. This is unexpected for a reference sample. Information and statistics of the reference population that was used during normalisation is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	You can find the reference sample that causes this nRSQ (excl) by assessing the probe result in all reference samples. Most likely, the probe that had the nRSQ (excl) error has hardly any signal in a reference sample. If you have ample reference samples in your experiment, you could consider reanalysing your experiment while omitting the affected reference sample. If sufficient reference samples were available during normalisation (check the *Reference Sample Quality Issues* table in the Results PDF report), proceed with caution when interpreting this probe result. Otherwise interpretation of the probe result is not recommended.

[n] For details about the data analysis, including normalisation and comparison, see Appendix I – Data analysis.

Priority	Error type	Explanation [n]	How to proceed
2	**nRSQ (qty)**	*Normalisation Reference Sample Quality Quantity error.* Error during normalisation. The size of the reference population used during normalisation is too small. Information and statistics of the reference population that was used during normalisation is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	Ensure you have included sufficient reference samples in the experiment. For some probemixes in certain applications it is allowed to use test samples for normalisation and not include reference samples. In that case ensure the test samples are of sufficient quality using the Quality Report PDF. Please check the instructions related to reference samples in the digitalMLPA NXtec/NXtec-MS Protocol and in the probemix-specific Product Description. Interpretation of the probe result is not recommended.
3	**cRSQ (qty)**	*Comparison Reference Sample Quality Quantity error.* Error during comparison. The size of the reference population used during comparison is too small. Information and statistics of the reference population that was used during comparison is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	Ensure you have included sufficient reference samples in the experiment. Please check the instructions related to reference samples in the digitalMLPA NXtec/NXtec-MS Protocol and in the probemix-specific Product Description. Interpretation of the probe result is not recommended.
4	**nRSQ (MIRR)**	*Normalisation Reference Sample Quality Minimum (or median) Intra Ratio References error.* Error during normalisation. The minimum (or median in case no reference samples are defined) intra ratio of reference samples (MIRR) used during normalisation is too low. Information and statistics of the intra ratios observed in the reference population used during normalisation is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	The probe signal is low in the reference population used for normalisation. This can for example be caused by a sequencing bias or a deletion or SNV in a defined reference sample. This low signal negatively affects the reliability of the result. Proceed with caution when interpreting this probe result. It is recommended to check the *statistics* column in the *Reference Sample Quality Issues* table (details in section 4.3.9): • If the displayed maximum[o] appears low, there may be a sequencing issue related to this probe. Please check the read counts for this probe in the Excel Report to see if this is remarkably lower as compared to other probes. • If the displayed maximum[o] appears OK, check the probe result in reference samples. If a defined reference sample has a low signal, this may be the cause of the nRSQ (MIRR).

[o] Please note this is an intra ratio for which no expected range can be given.

Priority	Error type	Explanation [n]	How to proceed
5	**cRSQ (MIRR)**	*Comparison Reference Sample Quality Minimum (or median) Intra Ratio References error.* Error during comparison. The minimum (or median in case no reference samples are defined) intra ratio of reference population (MIRR) used during comparison is too low. Information and statistics of the intra ratios observed in the reference population used during comparison is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	The probe signal is low in the reference population used for comparison. This can for example be caused by a sequencing bias or a deletion or SNV in a defined reference sample. This low signal negatively affects the reliability of the result. Proceed with caution when interpreting this probe result. It is recommended to check the *statistics* column in the *Reference Sample Quality Issues* table (details in section 4.3.9): • If the displayed maximum[o] appears low, there may be a sequencing issue related to this probe. Please check the read counts for this probe in the Excel Report to see if this is remarkably lower as compared to other probes. • If the displayed maximum[o] appears OK, check the probe result in reference samples. If a defined reference sample has a low signal, this may be the cause of the nRSQ (MIRR).
6	**nRSQ (range)**	*Normalisation Reference Sample Quality 95% confidence range error.* Error during normalisation. The 95% confidence range of the results from the reference population used during normalisation lies outside the expected range. Information and statistics of the reference population that was used during normalisation is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	An nRSQ (range) may indicate a sub-optimal choice of reference samples or, in case no references are defined, a suboptimal test population (e.g. too small, too many samples with the same aberration or too much variation). In case defined reference samples are used during normalisation, you can assess the variability of the probe result in all reference samples. Dedicated reference samples should not have any copy number variation in the genes tested. Interpretation of the probe result is not recommended.
7	**cRSQ (range)**	*Comparison Reference Sample Quality 95% confidence range error.* Error during comparison. The 95% confidence range of the results from the reference population used during comparison lies outside the expected range. Information and statistics of the reference population that was used during comparison is displayed in the *Reference Sample Quality Issues* table in the Results PDF report.	A cRSQ (range) may indicate a sub-optimal choice of reference samples or, in case no references are defined, a suboptimal test population (e.g. too small, too many samples with the same aberration or too much variation). In case defined reference samples are used during comparison, you can assess the variability of the probe result in all reference samples. Dedicated reference samples should not have any copy number variation in the genes tested. Interpretation of the probe result is not recommended.

Priority	Error type	Explanation [n]	How to proceed
8	**RPQ**	*Reference Probe Quality error* Error during normalisation. Too many reference probes that are used during normalisation have a result outside the expected range.	An RPQ error typically indicates that reference probes are not stable in the sample, for example by having too many aberrations in the sample or using a sample type that the probemix is not intended for. Most probes are normalised using the default reference probes in the probemix, but for certain probes a defined set of reference probes may have been selected by MRC Holland to be used for normalisation. Information about this can be found in the probemix-specific Probe Information File. Interpretation of the probe result is not recommended.
9	**Ambiguous**	*Ambiguous result warning* The probe result lies in the ambiguous range based on the set boundaries for this probe.	Check the probemix-specific Product Description for interpretation of results. Proceed with caution when interpreting this probe result.
10	**Lower conf**	*Lower confidence result warning* The probe result is aberrant (outside the expected range of this probe), but the 95% confidence range of the test result[p] overlaps with the 95% confidence range of the results from the reference population used for comparison.	Proceed with caution when interpreting this probe result. The following is recommended: • Check the results of neighbouring probes. • Check whether the result is outside the 95% confidence range of the comparison reference population. • When available, check the chart for visualisation of the 95% confidence ranges.
11	**RPQ**	*Reference Probe Quality warning* Many reference probes that are used during normalisation have a result outside the expected range.	See information on *RPQ error* above. This RPQ warning is less severe than the RPQ error, but proceed with caution when interpreting this probe result.
12	-	There is no error or warning related to this probe result.	Ensure overall sample quality is 'passed' and suitable for interpretation of the probe result.

4.3.7 Results section – Aberrant column

The call whether a probe result is considered aberrant (outside the expected range) and its significance (related to the reference population) is displayed in the *aberrant* column of the results section of the Results report, see Table 8.

Text formatting:
- Grey: There is an error or warning related to this probe. Please check this before proceeding.
- **Orange and bold**: Result is outside the expected range for this probe, i.e. aberrant.
- Black: Result is within the expected range for this probe, i.e. not aberrant.

[p] For some probe results, the 95% confidence range may not be available.

Table 8. Aberrant call in results section.

Aberrant	Explanation
YES*	The result is outside the expected range for this probe and significantly different from the reference population.
No	The result is within the expected range for this probe.
The following cases all indicate that a (statistical) warning or error is issued. It is advised to review these results carefully.	
Yes	The result is outside the expected range for this probe (or no result is available) and no significance information is available.
Yes NS	The result is outside the expected range for this probe and the 95% confidence range of the test result[p] overlaps with the 95% confidence range of the results from the reference population, but the *Inconclusive* error is turned off for this probe. Please consider the follow-up steps as described for the *Inconclusive* error in Table 7.
Yes?	There is an error; please check the error prior to proceeding with this probe result. The result is outside the expected range for this probe (or no result is available) and no significance information is available.
Yes? *	There is an error; please check the error prior to proceeding with this probe result. The result is outside the expected range for this probe (or no result is available) and the result is significantly different from the reference population.
Yes? NS	There is an error; please check the error prior to proceeding with this probe result. The result is outside the expected range for this probe (or no result is available) and the 95% confidence range of the test result[p] overlaps with the 95% confidence range of the results from the reference population.
No?	There is an error; please check the error prior to proceeding with this probe result. The result is within the expected range for this probe.

4.3.8 Results section – Result column

The probe result is displayed in the *result* column of the results section of the Results report. The exact calculations used to produce the probe result are dependent on the probemix, probe and experimental setup. Probe-specific information is available in the probemix-specific Product Description and Probe Information File. In general, the data is displayed as represented in Table 9.

⚠ FOR RELIABLE RESULT INTERPRETATION, PLEASE ENSURE THE OVERALL SAMPLE QUALITY IS PASSED AND CAREFULLY EVALUATE ANY ERRORS AND WARNINGS.

Table 9. Result types in results section.

Result type	Explanation
Inter ratio	Common for copy number probes.
Percentage	Used for some methylation-specific probes. Whether the percentage is based on an intra or inter ratio depends on the application.
Present/absent	Common for mutation-specific probes. Depending on the probe ratio being above or below the threshold, the mutation is considered present or absent, respectively.

4.3.9 Reference Sample Quality Issues table

The *Reference Sample Quality Issues* table is only displayed in the Results PDF report in case there are Reference Sample Quality (RSQ) errors for one or more probe results. Details displayed on the report are described in Table 10. For more information on the different RSQ errors and how to proceed, see Table 7.

Table 10. *Information displayed in the Reference Sample Quality Issues table.*

Description	Explanation
Probe information	
See Table 5 for information.	
Reference samples	

Type	This displays the sample type population to which the probe is normalised or compared (see column *role* whether this is applicable for normalisation or comparison).
	• *Reference*:
	  ○ In case there are defined reference samples in the experiment these are used for normalisation or comparison.
	  ○ In case there are no reference samples defined, the test samples are used for normalisation or comparison.
	  ○ In case of methylation-specific digitalMLPA, both digested and undigested reference samples are used for normalisation or comparison.
	• *Reference [D]*: Digested reference samples are used for normalisation or comparison.
	• *Reference [U]*: Undigested reference samples are used for normalisation or comparison.
	• *SD*: The SD is used for normalisation or comparison.
	• *Custom*: Check the probemix-specific Product Description and Probe Information File to see what samples are used for normalisation or comparison.

#refs	This displays the size of the reference population used during normalisation or comparison (see column *role* whether this is applicable for normalisation or comparison).
Role	This displays in what process the error occurred: • *Normalisation*: Quantity, range or excluded error during normalisation (nRSQ qty, range or excl). • *Norm. MIRR*: MIRR error during normalisation (nRSQ MIRR). • *Comparison*: Quantity or range error during comparison (cRSQ qty or range). • *Comp. MIRR*: MIRR error during comparison (cRSQ MIRR).
Range exp.	This displays the expected range of the reference population (see column *type*) during normalisation or comparison (see column *role*). For role *MIRR*, this displays the expected intra ratio for the reference population.
Statistics	
Minimum	This displays the minimum result observed in the reference population (see column *type*) during normalisation or comparison (see column *role*). For role *MIRR*, this displays the minimum intra ratio for the reference population.
Range 95%	This displays the 95% confidence range of the reference population (see column *type*) during normalisation or comparison (see column *role*). For role *MIRR*, this displays the 95% confidence range of the intra ratio for the reference population.
Maximum	This displays the maximum result observed in the reference population (see column *type*) during normalisation or comparison (see column *role*). For role *MIRR*, this displays the maximum intra ratio for the reference population.
Errors	
Quantity	A red cross is displayed in case the probe result has a quantity error: nRSQ (qty) or cRSQ (qty) for normalisation or comparison respectively. The size of the reference population (see column *#refs*) is lower than what is required.

Description	Explanation
Range	A red cross is displayed in case the probe result has a range error: nRSQ (range) or cRSQ (range) for normalisation or comparison respectively. The 95% confidence range of the reference population (see column *range 95%*) is wider than what is expected (see column *range exp.*).
Excluded	A red cross is displayed in case the probe result has an excluded error: nRSQ (excl). At least one defined reference sample is excluded (the number of reference samples used for normalisation is displayed in column *#refs*).
MIRR	A red cross is displayed in case the probe result has a MIRR error: nRSQ (MIRR) or cRSQ (MIRR) for *roles norm. MIRR* or *comp. MIRR* respectively. The minimum (in case of defined reference samples; see column *minimum*) or median (in case no reference samples are defined) intra ratio is lower than what is expected (see column *range exp.*).

4.4 NO-DNA PDF REPORT

A no-DNA PDF report is generated for no-DNA samples. The information printed in the tables is similar to the Results PDF report (see Table 1 for explanations), except that some cells are greyed out as this information is not available for a no-DNA sample.

No-DNA samples are subjected to quality evaluation and the result can be either '*passed*' or '*failed*'. In case a no-DNA sample is failed, a list of the 10 most frequently recognised probes is displayed in the no-DNA report, see Table 11 for more information. Please refer to the Excel Report for a full list of the total number of reads assigned per probe.

Table 11. Table with most frequently recognised probes in no-DNA PDF report.

Description	Explanation
Probe information	
See Table 5 for information.	
Read counts	
Total	This displays the total number of reads assigned to the probe in the no-DNA sample.
Distinct	This displays the number of distinct reads assigned to the probe in the no-DNA sample. Distinct reads are based on the number of independent ligation events.

4.5 EXCEL REPORT

PLEASE BE AWARE THAT THE EXCEL REPORT DOES NOT CONTAIN DETAILS ON THE STATISTICAL SIGNIFICANCE OF RESULTS. IT IS HIGHLY RECOMMENDED TO CONSULT THE SAMPLE-SPECIFIC PDF REPORTS AS WELL.

Alike the PDF reports, the Excel Report contains the quality assessment of each sample and probe results. The Excel Report also provides probe read counts and probe details (as also available in the probemix-specific Probe Information File), but lacks information on the significance of the result. For an overview of the different tabs in the Excel Reports, see Table 12. Please note that this Excel Report is tailored to the envisioned purpose of the probemix. Therefore, the content may be slightly different between reports from different probemixes.

Table 12. Information in the tabs of the Excel Report.

Tab	**Explanation**
Welcome Page	
This tab contains background information on how to use the Excel Report.	
Results	
This tab contains the results of all relevant probes for every DNA sample included in the experiment. Please check the legend on the *Welcome Page tab* for information on the colouring of the cells.	
Description	*Explanation*
Probe number	See Table 5.
Probe function	*CTRL*: Quality control probe. This probe is part of a quality check. For other probe functions, see Table 5.
Location (kb)	See Table 5.
Chr band	
Gene	
Exon	
Probe details	This displays probe details when considered informative, e.g. mutation details. See also the probemix-specific Probe Information File.
Probe info	This displays additional information when considered informative. See also the probemix-specific Probe Information File.
CN expected	See Table 6.
Probe warning	This displays known issues related to the probe. See also the probemix-specific Probe Information File.
Reference probe	This displays whether the probe is selected as default reference probe. See Appendix I – Data analysis for more information.
Barcode	See Table 1.
Sample name	
Sample type	
Sample sex	

DNA read counts
This tab contains the total read counts of all relevant probes in each DNA sample. For an explanation of the terms used in this tab, see info at *Results tab* above.
No-DNA read counts
This tab contains the total read counts of all relevant probes in no-DNA samples. For an explanation of the terms used in this tab, see info at *Results tab* above.
Sample quality
This tab contains information on the quality checks performed in every sample. For an explanation of the terms used in this tab, see Table 1.
Software version
This tab displays the Coffalyser digitalMLPA version (major.minor.patch) and the versions of the individual software modules used in the analysis.

4.6 XML REPORT

Coffalyser digitalMLPA can generate XML output files per sample, that can, for example, be used in a custom analysis pipeline. The content of the XML file is highly similar to the Results PDF report. Please consult the preceding sections for an explanation of the data presented. Please contact MRC Holland at info@mrcholland.com for support.

APPENDIX I – DATA ANALYSIS

Coffalyser digitalMLPA uses a series of normalisation steps and calculations to compute the digitalMLPA probe results. In the subsequent comparison stage the result is compared against the expected range for the probe and the reference population to determine whether the result is considered aberrant and assess its significance. In this whole process and the associated statistics, it is assumed that digitalMLPA data follows normal distribution.

⚠ PLEASE NOTE THAT DATA ANALYSIS IS TAILORED FOR EACH PROBE AND DEPENDS ON THE PRODUCT, APPLICATION AND EXPERIMENTAL SETUP. GENERAL GUIDELINES ARE LISTED IN THE SECTION BELOW, BUT PLEASE CHECK THE PROBEMIX-SPECIFIC PRODUCT DESCRIPTION AND/OR PROBE INFORMATION FILE FOR MORE DETAILS.

NORMALISATION

In a process 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 reference probes in the sample. For most probes, the default set of reference probes that is present in the probemix is used, but for some probes (e.g. most methylation-specific probes) a specific set of reference probes is used for intra-normalisation. Intra-normalisation is performed for every probe and every sample.

During inter-normalisation, Coffalyser digitalMLPA normalises the relative probe values of each sample to those of the reference population, resulting in inter ratios. The reference population that is used during inter-normalisation depends on the probe, the probemix and experimental setup.

Normalisation calculations

A simplified version of the complete 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.

$$Median\left(\frac{(Tp1\ in\ sample\ 1/Rp1\ in\ sample\ 1)}{(Tp1\ in\ reference\ sample\ 1/Rp1 in\ reference\ sample\ 1)},\ldots,\frac{(Tp1\ in\ sample\ 1/Rp_n\ in\ sample\ 1)}{(Tp1\ in\ reference\ sample\ 1/Rp_n\ in\ reference\ sample\ 1)}\right)$$

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 results in 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.

Samples used as reference for normalisation

Most frequently, probes are inter-normalised against the defined reference samples present in the experiment. If no reference samples are defined in an analysis and this functionality is enabled for the probemix, every sample will be inter-normalised against the other test and undefined samples that are considered of sufficient quality after intra-normalisation. In this case, the calculations slightly differ from what is outlined above: in step 2, instead of taking the average over the previously calculated median values, Coffalyser digitalMLPA calculates the inter ratio for a probe by taking the median of these medians. Furthermore, an additional round

of normalisation is performed to exclude samples with signals in the outlier range from the population of samples used for normalisation.

Additional notes:

- In methylation-specific digitalMLPA, depending on the probemix, certain methylation-specific probes in digested test samples are inter-normalised using undigested reference samples. Copy number probes in a methylation-specific digitalMLPA probemix are generally inter-normalised using all reference samples irrespective of the digestion-state.
- For normalisation of copy number probes targeting non-PAR locations on the sex chromosomes, the reference population should consist of samples of the same sex as the test sample. If insufficient reference samples of the same sex are present, the reference population may be extended using converted probe signals of reference samples of the opposite sex.
- Please note that some probes do not undergo inter-normalisation. In general, if no signal is expected in reference samples, e.g. for mutation-specific probes and methylation-specific probes in tumour analysis, the probe is likely to be subjected to intra-normalisation only.

COMPARISON

In a process called comparison, Coffalyser digitalMLPA compares the intra- or inter-normalised result (calculated above) of the probe in the test sample to the expected range to determine whether the result is considered aberrant. Furthermore, the significance of the result is determined by comparing the 95% confidence range of the test result, that is projected during normalisation[q], to the 95% confidence range of the reference population, that is based on the variation of the results within that population. The reference population that is used depends on the probe, the probemix and experimental setup.

Samples used as reference for comparison

In general, a copy number probe in a digitalMLPA probemix is compared to the defined reference samples present in the experiment. If no reference samples are defined in an analysis and this functionality is enabled for the probemix, every sample will be inter-normalised against the other test and undefined samples that are considered of sufficient quality after intra-normalisation. Note that in this case, the population of samples used for comparison is different from that used for normalisation, as samples with outlier signals as well as the sample under analysis itself are excluded for normalisation, but are included for comparison.

Additional notes:

- In methylation-specific digitalMLPA, most methylation-specific probes in digested test samples are compared to digested reference samples. Note that for these probes, the population of samples used for comparison is different from that used for normalisation, where undigested rather than digested samples are used. For some methylation-specific probes, the expected range that is used to determine whether a result is considered aberrant is dynamic and can grow wider based on the confidence range observed in the reference population. Other probes in a methylation-specific digitalMLPA probemix are generally compared using all reference samples irrespective of the digestion-state.
- For most mutation-specific probes, the result in the test sample is compared to what is expected for the reference population. For some mutation-specific probes, the expected range that is used to determine whether a result is considered aberrant is dynamic and can grow wider based on the confidence range observed in the reference population.

[q] For some probe results, the 95% confidence range may not be available.

APPENDIX II – ROOT CAUSE

Coffalyser digitalMLPA will determine a plausible root cause if quality issues are detected in the sample. Please note that the root cause is considered the most important issue to address, but multiple other causes could be at play. The root causes that can be displayed, and their priority, are shown in Table 13.

Information on the related quality checks can be found in Table 1. More information for troubleshooting can be found via https://www.mrcholland.com/r/ifu/digitalmlpa-troubleshooting on the MRC Holland support portal.

Table 13. Plausible root causes.

Priority	Plausible root cause	Explanation
1	Undetermined	Root cause determination is not performed, as data analysis was not completed. Please check the *program errors* that occurred.
2	Not applicable	Data quality is OK.
3	Reference sample quantity	In case of defined reference samples in the analysis, an insufficient number of (suitable) reference samples is present. If no reference samples are defined in the analysis, several issues can trigger this plausible root cause: - The analysis of the specific probemix requires defined reference samples. Check the probemix-specific product description for details. - (Almost) all samples have quality issues that lead to sample quality failure and none of these samples can be used for reliable inter-normalisation. To troubleshoot the sample quality issue, check whether there are errors in the *sequence run*, *identification* and *reaction - intra* sections (see Table 1). Or reanalyse the data with defined reference samples and check the plausible root cause again. *Related quality checks: reference sample quantity & on probe level: RSQ (qty)*
4	Too many sequence errors	Many reads that are assigned to probes contain errors. *Related quality check: sequence quality*
5 [r]	DNA quantity	The amount of input DNA is considered too low. For no-DNA samples, this cause is triggered if the no-DNA quality evaluation was not within criteria. *Related quality check: DNA quantity*
6	Possible incorrect sheet	Probemix lot detection was not successful, possibly caused by selecting a wrong Product Sheet (probemix lot) in the Coffalyser Definition File. *Related quality checks: probemix lot & unrecognised reads*
7	Variable read length	The reads in the FASTQ file are not of identical length. *Related quality check: sequence quality.*
8	Probe read depth too low	The read depth is too low. *Related quality check: MTRQ & MDRQ.*

[r] Plausible root cause *DNA quantity* can be triggered via multiple routes with different priority. The highest priority is listed here.

Priority	Plausible root cause	Explanation
9 [s]	Barcode lot detection	The barcode lot is incompatible or not known as compatible with the probemix. Or multiple barcode lots are detected in one sample. *Related quality check: barcode lot.*
10 [t]	Probemix lot detection	Probemix lot detection was not successful. *Related quality check: probemix lot*
11	Sample tracker detection	The Sample Track Probe(s) that are detected differ from what is specified in the Coffalyser Definition File. *Related quality check: sample tracker* Note: Sample Track Probes are not currently available.
12	Sex detection	There is an issue with sex detection. The detected sex of the sample differs from what is specified in the Coffalyser Definition File. Or the software cannot reliably determine the sex. *Related quality check: sex*
13	DNA contamination	The DNA sample appears not of a single source. *Related quality check: DNA contamination*
14	DNA denaturation	There is an issue with sample DNA denaturation. *Related quality check: DNA denaturation*
15	DNA depurination	There is an issue with sample DNA depurination. *Related quality check: DNA depurination*
16	DNA fragmentation	There is an issue with sample DNA fragmentation. *Related quality check: DNA fragmentation*
17	Digestion	There is an issue with sample digestion. *HhaI* digestion is detected in a sample that is not specified as digested, or the extent of *HhaI* digestion in a digested sample is not as expected. *Related quality check: DNA digestion*
18	Hybridisation condition	The conditions during the overnight hybridisation reaction were not optimal. This could be due to evaporation during the hybridisation reaction, which can be detected by elevated ratios of both the *hybridisation completeness* and *hybridisation temperature* control probes (ratios are available in the Excel Report). *Related quality checks: hybridisation temperature & hybridisation completeness*
19	Ligation temperature	There is an issue with the temperature during the ligation reaction. *Related quality check: ligase start temperature*
20	Ligation activity	There is an issue with the ligase activity. *Related quality check: ligase activity*
21	Ligation inactivation	There is an issue with ligase inactivation. *Related quality check: ligase inactivation*

[s] Plausible root cause *Barcode lot detection* can be triggered via multiple routes with different priority. The highest priority is listed here.
[t] Plausible root cause *Probemix lot detection* can be triggered via multiple routes with different priority. The highest priority is listed here.

Priority	Plausible root cause	Explanation
22	Polymerase activity	There is an issue with the polymerase activity. *Related quality check: polymerase activity*
23	Unrecognised reads	There are many unrecognised reads; reads that cannot be reliably assigned to a probe. *Related quality check: unrecognised reads*
24	Unbalanced sample read depths	The reads are not evenly spread over the samples within one experiment. *Related quality check: sample uniformity.*
25	no-DNA control(s)	There is an issue in the no-DNA control(s) of this experiment. *Related quality check: no-DNA control(s)*
26	Reference sample population	There is an issue with the reference sample population. *Related quality checks: reference sample quality, reference sample quantity & on probe level: RSQ (excl), RSQ (qty), RSQ (MIRR) and RSQ (range)*
27	Reference probe quality	There is an issue with the reference probe quality. *Related quality checks: reference probe quality & on probe level: RPQ*

APPENDIX III – TROUBLESHOOTING

The tables below list common errors encountered during data analysis with Coffalyser digitalMLPA. If you encounter problems or software errors not described below or require assistance, please contact info@mrcholland.com and include the Coffalyser Crash Log.txt file and a screenshot of the error if applicable.

⚠ ONLY ISSUES AND ERRORS RELATED TO THE SOFTWARE ARE LISTED HERE. MORE INFORMATION FOR TROUBLESHOOTING digitalMLPA DATA CAN BE FOUND ON THE MRC HOLLAND SUPPORT PORTAL.

GENERAL

Problem	Cause
Program MRC.DataAnalysis.Client.exe is not able to start.	Security settings in Windows 11 can block Coffalyser digitalMLPA executables. This can be resolved by right clicking on the .exe, selecting properties and ticking 'Unblock'. Or unblock the whole package using Windows PowerShell with the following prompt: '*Get-ChildItem "path to your Coffalyser digitalMLPA package" -Recurse | Unblock-File*'. We recommend using 7-Zip to unpack the Coffalyser digitalMLPA package. 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.
The following message is shown upon double clicking on the MRC.DataAnalysis.Client.exe: *The app you're trying to install isn't a Microsoft-verified app*	
The following error occurred: *The program encountered an unexpected exception:* *An attempt was made to load an assembly from a network location which would have caused the assembly to be sandboxed in previous versions of the .NET Framework. This release of the .NET Framework does not enable CAS policy by default, so this load may be dangerous. If this load is not intended to sandbox the assembly, please enable the loadFromRemoteSources switch.*	
.NET Framework is not up-to-date (older than version 4.8).	The latest .NET framework can be downloaded from https://dotnet.microsoft.com/download/dotnet-framework.
The following error message is shown when starting Coffalyser digitalMLPA: *No valid Coffalyser digitalMLPA license found in the configuration.*	It is necessary to renew the free license for Coffalyser digitalMLPA. To renew the license, download a new version of Coffalyser digitalMLPA from your account on www.mrcholland.com. Remove or archive the old version, and extract the new package. No further installation is required.
The following message is shown when starting Coffalyser digitalMLPA: *The current license will expire in xx days!* *Update your configuration and/or software to continue using this product after the expiration.*	The license expiration date is approaching. To renew the license, download a new version of Coffalyser digitalMLPA from your account on www.mrcholland.com. Remove or archive the old version, and extract the new package. No further installation is required.
Software does not start and a fatal error is shown: *CLR error: 8007007a. The program will now terminate.*	This can happen when the full path name of the Coffalyser digitalMLPA location is too long, or when the path name contains special characters.

Problem	Cause
The following error is shown after the software has started: *The program encountered an unexpected exception:* *The configuration directory ('…') does not exist*	This can be solved by shortening the path name by renaming (sub)folders, or by moving Coffalyser digitalMLPA to a higher-level (less deeply nested) directory, and by eliminating special characters from the path name.
The following error is shown after the software has started: *The program encountered an unexpected exception:* *Could not find a part of the path '…'.*	
Encountered problems with the Next Generation Sequencing data quality, e.g. overclustering.	To optimize your sequencing run for digitalMLPA, follow the guidelines provided in the digitalMLPA NXtec/NXtec-MS Protocol or check platform-specific application notes when available, and the recommendations of the sequencing platform manufacturer. We also recommend reviewing sequencing quality metrics using the software provided for your sequencing platform prior to analysis with Coffalyser digitalMLPA.
Software cannot finish the analysis and no files with analysis results are generated.	Data could not be analysed by the software. The cause of the error is described in a Coffalyser Crash Log.txt file which is located in the defined output folder.
An unhandled exception occurred: *Exception of type 'System.OutOfMemoryException' was thrown.*	There is an issue with memory usage. Please check whether the system requirements as outlined in section 2.1 are met. If the source is Microsoft.ReportViewer.Common, please consider splitting Coffa files in multiple folders and use the command line interface to generate reports (see APPENDIX IV – COMMAND LINE INTERFACE (CLI). If other out of memory errors occur or the issue persists, please contact MRC Holland at info@mrcholland.com.
An exception occurred: *The program encountered an unexpected exception:* *required data not present in stream (truncated file?)*	There is a problem with the FASTQ input file. We recommend regenerating the FASTQ file using the software provided for the sequencing platform and ensuring that sufficient disk space is available during FASTQ generation and storage.
The text in the Coffalyser digitalMLPA dialog screens does not properly fit.	The display settings of your computer may influence this. Ensure scaling is set to 100% or 125% or higher. Custom scaling between 101 and 124% may negatively influence readability.

EXPERIMENT DEFINITION

Problem	Cause
The following warning is shown when reusing an old Coffalyser Definition File: *The specified analysis settings xxx cannot be found. Do you want to continue?*	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 and on our website, 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: *Failed to paste the sample details:* *the samples on your clipboard does not match the number of selected rows*	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. 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: *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.*	This message indicates that the number of copied sample names being pasted into Coffalyser digitalMLPA exceeds the number of remaining samples in the range. 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.
The following error message is shown: *Please correct the following error before continuing:* *- Output/temporary directory name is too long (detected = 'xx', available = 'xx')*	This can happen when the path of the output and/or temporary directory is too long. This can be solved by: • Selecting an output/temporary directory with a shorter name. • Shortening the experiment name. • Shortening sample names.

DATA ANALYSIS (COFFALYSER digitalMLPA – RUN WIZARD)

Problem	Cause
The run is completed but error messages are shown in the summary window.	Click '*Open Results Folder*' and review the Quality Report PDF file for more information.
Output files of some of the samples/barcodes ended up in the '*Undefined Experiment {nn}*' folder.	Not all barcodes used in the experiment were included in the Coffalyser Definition file. Ensure that all the barcodes that were part of your experiment are included in the Coffalyser Definition File and reanalyse the experiment. Reanalysis can also be done by using all Coffa files belonging to the experiment.

Problem	Cause
The following error message is shown: *sample (xxx) does not have enough reads for analysis.*	This can happen if no reads are detected for a barcode that was included in the Coffalyser Definition File. When defining experiments in the Coffalyser Definition File, ensure you only select the barcodes that are present in the data. Also, ensure the experiment has been performed according to the instructions in the digitalMLPA NXtec/NXtec-MS Protocol, and both barcode solution and probemix are added to a sample.
The following error message is shown: *not a single experiment qualified for analysis (detected 01 exceptions: [01] 'the experiment does not have a valid product sheet ('Undefined Experiment 01')')*	This error can occur when the product and sheet are not defined in the Coffalyser Definition File, and the Product Sheet cannot be automatically detected by the software. Ensure you are using the latest software version and ensure there is no contamination during your experiment, e.g. from another probemix lot. Please define the product and sheet that should be used for analysis in the Coffalyser Definition File.
The following error message is shown: *The experiment (xxx) does not have a sheet.*	This can happen when the Product Sheet required for analysis of the digitalMLPA results is not available in the version of Coffalyser digitalMLPA used. Ensure you use the latest version of Coffalyser digitalMLPA. The latest version can be downloaded from your account on www.mrcholland.com.
The following error message is shown: *Sample (xxx) missing too many reference probes.*	This error is given when not enough reads have been detected for a large number of reference probes. Ensure the experiment has been performed according to the instructions in the digitalMLPA NXtec/NXtec-MS Protocol in order to obtain sufficient reads for all probes (e.g. DNA input amount should be sufficient, ligation should be successful and read depth adequate). Also, ensure the correct Product Sheet has been selected while making the Coffalyser Definition File.
The following error message is shown: *The experiment does not have any analysable samples.*	This can happen when none of the samples in an experiment can be analysed, for example when only no-DNA reactions are present in an (undefined) experiment. Ensure the experiment has been performed according to the instructions in the digitalMLPA NXtec/NXtec-MS Protocol and check whether all your samples are correctly defined in the Coffalyser Definition File.
The following error message is shown: *Not a single experiment defined in the experiment definition file qualified for analysis.*	This error message is often accompanied by other error messages that explain what has gone wrong that prevented analysis of the digitalMLPA experiment(s). Ensure the experiment has been performed according to the instruction in the digitalMLPA NXtec/NXtec-MS Protocol and check whether all your samples are correctly defined in the Coffalyser Definition File.
The following error message is shown: *The process cannot access the file xxx because it is being used by another process.*	This can happen when Coffa files are reanalysed, while Coffalyser digitalMLPA output files generated based on these Coffa files are still open. Ensure that the Coffalyser digitalMLPA output files are closed and start the reanalysis again.

Problem	Cause
The following error message is shown: *The parent folder contains more than 100 subfolders, which will prevent Microsoft Windows from showing the full list in the following form.*	The folder where the output or temporary directory are placed, or where the FASTQ file is stored, contains more than 100 subfolders. To solve this, restructure the folder where the output directory, temporary directory or FASTQ file are located (e.g. combine folders per week or month) so less than 100 subfolders are present.
The following error message is shown: *The average read length (xx) for the first 50000 reads is too short (minimum required xx).*	The read length in the FASTQ file is too short for reliable analysis. Please check the instructions in the digitalMLPA NXtec/NXtec-MS 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). Also check for sequencing issues that could have affected the read length. Please note that the read length mentioned in the Coffalyser digitalMLPA error message is not the complete read length present in the FASTQ file. The length in the message corresponds to the part of the read that is required for read assignment to a probe sequence.
The following error message is shown: *The following errors have been detected while initialising the definition:* *- failed to retrieve the search graph blob (xxx) specified in the first encountered value (xxx).*	This can happen when Coffa files are analysed that were initially generated using another version of Coffalyser digitalMLPA. To solve this, the digitalMLPA data can be reanalysed by starting with the FASTQ file as input.

REPORTS

Problem	Cause
There is a '*(no consensus)*' error on barcode collection and/or no-DNA control(s).	This can occur when Coffa files that originate from different analyses are reanalysed together in one experiment. If the error is on barcode collection, the original Coffa files have a difference in barcode collection. Please note that all samples in an experiment should have the same barcode collection. If the error is on no-DNA control(s), the no-DNA control check was different in each originally analysed experiment (e.g. Ok vs Error, or Ok vs grey (thus no no-DNA sample included)).
A different sex was detected by Coffalyser digitalMLPA than specified in the Coffalyser Definition File.	This can occur in case of a sample swap or if the sex was incorrectly defined in the Coffalyser Definition File. It may also have a biological cause. For example, in older men, cells may lose the Y chromosome causing the sample to be detected as female.

Problem	Cause
The Results.pdf file shows the error message: *sample ('xxx') has an invalid sex (specified = 'Undefined', detected = 'Ambiguous')*. In addition, the quality summary table shows QC checks FAILED and the other checks are grey.	This can occur when the sample sex is not defined in the Coffalyser Definition File, and the sex cannot be determined by Coffalyser digitalMLPA. Please define the sex of the sample in the Coffalyser Definition File. This can also occur when a no-DNA sample is defined as Test sample. Change the sample type in the Coffalyser Definition File to '*no-DNA*' and reanalyse the experiment.
A warning is given for one control probe type (e.g. denaturation probes) in the column *inter* in the table *reaction* in the Quality Report.pdf file, but not in the column *intra*.	There is a difference (in quality, in the amount of contaminants present etc.) between this sample and the reference population. This can occur for example if another DNA extraction method was used for test samples compared to reference population.
A warning is given for one control probe type (e.g. denaturation probes) in the column *intra* in the table *reaction* in the Quality Report.pdf file, but not in the column *inter*.	This indicates that the quality of both the test sample and the reference population may be compromised. However, there is no difference (in quality, in the amount of contaminants present etc.) between this sample and the reference population.
There is an nRSQ (excl) or nRSQ (qty) error for Y-probes.	Ensure that a sufficient number of suitable male reference samples is included in your experiment. Female reference samples do not have a signal for Y-probes and can therefore not function as reference sample for Y-probes. Please check the instructions related to reference samples in the digitalMLPA NXtec/NXtec-MS Protocol and in the probemix-specific Product Description.
In the Results report, different decimal separators are used in the *comparison reference sample* and *results* columns.	In some cases there is an issue with number formatting. We recommend using the language of your Windows installation as regional format without overwriting the number format defaults for that language for best performance.

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.{module name}.Client.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 for an example).

```
*** EXAMPLE USAGE ***

       MRC.DataAnalysis.Analyses.Client.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.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)}CRLF
Technology ⟶ ⟶ ⟶DigitalMlpaCRLF
SettingsProfile ⟶ ⟶{file id (guid)}CRLF
SampleBarcodeCollection ⟶{file id (guid)}CRLF
PathData ⟶ ⟶ ⟶{path to output directory, e.g.: "C:\digitalMLPA output"}CRLF
PathTemp ⟶ ⟶ ⟶{path to directory for temp files, e.g.: "C:\temp"}CRLF
PathsInputConversion ⟶{path to input FASTQ(.gz) file(s)}CRLF
UndefinedSampleAction ⟶{setting regarding undefined samples, e.g.: "Exclude"}CRLF
OutputTypes ⟶ ⟶ ⟶{setting regarding output files to be generated, e.g. "ReportsAndExcel" or "ReportsAndXml"}
```

Tabs are represented as yellow arrows, spaces are represented as yellow periods, CRLF 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 ⟶ ⟶5CRLF
Technology ⟶ ⟶ ⟶ ⟶DigitalMlpaCRLF
SettingsProfile ⟶ ⟶ ⟶518d1a77-1d51-4c59-86b6-f95a6c1be30aCRLF
SampleBarcodeCollection ⟶bedbe68c-7b74-a02c-9fc0-af3e22142b82CRLF
PathData ⟶ ⟶ ⟶ ⟶C:\digitalMLPA output CRLF
PathTemp ⟶ ⟶ ⟶ ⟶C:\tempCRLF
PathsInputConversion ⟶\\networklocation\FASTQs\NextSeq1-RUN121_S1_R1_001.fastq.gzCRLF
UndefinedSampleAction ⟶ExcludeCRLF
OutputTypes ⟶ ⟶ ⟶ReportsAndExcel
```

The file IDs (or Global Unique Identifiers (GUID)) of the *SettingsProfile* and the *SampleBarcodeCollection* are filled in automatically when you create a Coffalyser Definition File in the editor. Alternatively, they may be found in the header of the Profile Default.txt and the Barcodes .txt files (e.g. Collection 3.txt) located in the '*_Configuration*' sub-directory of your Coffalyser digitalMLPA package (for barcodes: …_Configuration\dMLPA\Barcodes).

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.

The *PathsInputConversion* parameter points to the FASTQ file(s). Multiple FASTQ files can be selected simultaneously (provided that barcodes are unique or identical barcodes belong to the same experiment and sample) by adding the *PathsInputConversion* line multiple times.

The *UndefinedSampleAction* line specifies how to treat any barcodes (samples) that were not defined in the Coffalyser Definition File. The options are '*Include*' and '*Exclude*', to respectively detect and include undefined samples in the analysis or ignore and exclude undefined samples from the analysis.

An analysis may be run with no further details in the Coffalyser Definition File. In that case, Coffalyser digitalMLPA will autodetect the digitalMLPA probemix used in each sample and will create experiments by grouping all samples tested with the same digitalMLPA probemix, provided that *UndefinedSampleAction* is set to *Include*. All samples will be treated as *Test samples*. In case your experiment contains samples of any other type (e.g. reference samples, SD or no-DNA) or that require specific details for correct analysis (e.g. digested samples or pooled DNA source) you must define these in the Coffalyser Definition File.



MRC HOLLAND RECOMMENDS ALWAYS EXPLICITLY DEFINING THE EXPERIMENT AND SAMPLE DETAILS 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:

```
CR LF
Experiment →{Experiment name} ⟶{Product Sheet id (guid)} ⟶{panels to be analysed}› CR LF
```

For example:

```
CR LF
Experiment →My digitalMLPA experiment ⟶00000000-0000-0000-0000-000000000000 ⟶all_available ⟶ CR LF
```

The experiment header line is single tab separated. The Product Sheet ID is filled in automatically when you select a probemix lot while creating a Coffalyser Definition File in the editor. Alternatively, it may be found in the header of the Product Sheet in the '*_Configuration*' sub-directory of your Coffalyser digitalMLPA package (…_Configuration\dMLPA\Products). 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 default panel set that should be analysed can be specified, this will only be used in case the panel set is not defined for the sample. If set to '*all_available*', all panels will be analysed. If a selection of the panels is desired, write down the panel names (comma separated).

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

```
Sample →{barcode id (guid)}›{sample definition}›{sample name} ⟶{sample sex} ⟶{panels to be analysed}› CR LF
```

For example:

```
Sample →ba626105-9a0a-4b10-9a30-c56382d140be ⟶SampleTest →BP01-01 Test sample 1 →Male ⟶all_available ⟶ CR LF
```

The information in this line is explained in Table 14 below.

Table 14. Information in sample line in Coffalyser Definition File.

Description	Info	In *cdf
{Barcode id (guid)}	The barcode IDs (GUID) are filled in automatically when you create a Coffalyser Definition File in the editor. Alternatively, they may be found in the Barcodes .txt files (e.g. Collection 3.txt) located in the '*_Configuration*' sub-directory of your Coffalyser digitalMLPA package (…_Configuration\dMLPA\Barcodes).	{guid}
{Sample definition}	Test sample (undigested)	*SampleTest*
	Digested test sample	*SampleTestDigested*
	Reference sample (undigested)	*SampleReference*
	Digested reference sample	*SampleReferenceDigested*
	Reference sample from pooled DNA source	*SampleReference, _TargetDnaIsPooled*
	SD sample	*SampleReferenceSpikedWithSd*
	no-DNA sample	*SampleNoDna* [u]
	Undefined	*Default*
{Sample name}	Provide a sample name	{text}
{Sample sex}	Male sample	*Male*
	Female sample	*Female*
	Leave sample sex undefined	*Undefined*
{Panels to be analysed}	All available panels	*all_available*
	Specific panel set The panel names are filled in automatically when you create a Coffalyser Definition File in the editor with a specific panel set.	{List of panel names, comma separated}

Running Coffalyser Definition File in CLI

Once you have created your *.cdf file, you can analyse your data and generate reports using the commands in Table 15 below.

Table 15. Commands for Coffalyser digitalMLPA when running with Coffalyser Definition File.

Module	Command
Data Conversion	`MRC.DataAnalysis.DataConversion.Client.exe -Definition "C:\Path\MyDefinitionFile.cdf"` The *-Definition* parameter should point to the directory of the Coffalyser Definition File.

[u] After Coffalyser digitalMLPA has processed the Coffalyser Definition File, 'None, _PopulationReference' will present for a no-DNA sample.

Module	Command
Sample Analysis	`MRC.DataAnalysis.Analyses.Client.exe -Definition "C:\Path\MyDefinitionFile.cdf" -AnalysisType All` The *-Definition* parameter should point to the directory of the Coffalyser Definition File. The *-AnalysisType* parameter can be set to '*All*' to run all analysis modules at once, or to '*FragmentAnalysis*', '*ComparativeAnalysis*' and/or '*ResultAnalysis*' to run the analysis modules separately.
Reporting (to generate PDF reports)	`MRC.DataAnalysis.Reporting.Client.exe -Definition "C:\Path\MyDefinitionFile.cdf"` The *-Definition* parameter should point to the directory of the Coffalyser Definition File.
Excel Writer (to generate Excel reports)	`MRC.DataAnalysis.ExcelWriter.Client.exe -Definition "C:\Path\MyDefinitionFile.cdf"` The *-Definition* parameter should point to the directory of the Coffalyser Definition File.
XML Writer (to generate XML reports)	`MRC.DataAnalysis.XmlWriter.Client.exe -Definition "C:\Path\MyDefinitionFile.cdf" -XmlLayout All` The *-Definition* parameter should point to the directory of the Coffalyser Definition File. The *-XmlLayout* parameter can be set to '*All*' to generate XML reports for both the experiment and samples. The XML reports can also be generated for only the experiment or samples, writing '*PerExperiment*' and/or '*PerSample*'.

CLI ANALYSIS WITHOUT A COFFALYSER DEFINITION FILE

⚠ RUNNING WITHOUT A COFFALYSER DEFINITION FILE IS ONLY POSSIBLE WHEN NO SAMPLE DEFINITION IS REQUIRED. FOR EXAMPLE, WHEN THE EXPERIMENT CONTAINS REFERENCE SAMPLES, DIGESTED SAMPLES, SAMPLES OF POOLED DNA SOURCE OR NO-DNA SAMPLES, A COFFALYSER DEFINITION FILE IS REQUIRED. FURTHERMORE, MRC HOLLAND RECOMMENDS ALWAYS EXPLICITLY DEFINING THE EXPERIMENT AND SAMPLE DETAILS TO RAISE THE LIKELIHOOD OF DETECTING MISTAKES.

For some applications, a Coffalyser Definition File is not required. Commands are highly similar to the section above, but for initialising the Data Conversion module, there are some additional parameters.

Data conversion without Coffalyser Definition File

To start data conversion use the following command:

```
MRC.DataAnalysis.DataConversion.Client.exe -Technology DigitalMlpa -PathsInputConversion "C:\Path1\File1.fastq.gz" -PathData "C:\Path\Out" [-PathTemp "C:\Path\Temp"] [-SettingsProfile "Profile Name"] [-SampleBarcodeCollection "Collection Name"] [-UndefinedSampleAction Include]
```

Mandatory parameters:

- *-Technology* parameter must be specified as '*digitalMLPA*' for digitalMLPA experiments.
- *-PathsInputConversion* parameter points to the FASTQ file(s). Multiple FASTQ files can be selected simultaneously (provided that barcodes are unique or identical barcodes belong to the same experiment and sample) by listing them space separated.
- *-PathData* parameter points to the output directory.

Optional parameters (between [ ] brackets in command above)

- *-PathTemp* parameter to a directory for the temporary files. If not specified the *PathData* directory will be used.
- *-SettingsProfile* parameter specifies with settings profile is used. If not specified, the default from the configuration will be used. The file ID (or GUID) of the *SettingsProfile* can be found in the header of the Profile Default.txt located in the '*_Configuration*' sub-directory of your Coffalyser digitalMLPA package.
- *-SampleBarcodeCollection* parameter specifies which barcode collection file to use. If not specified, the default from the configuration will be used. The file ID (or GUID) of the *SampleBarcodeCollection* can be found in the header of the Barcode .txt files (e.g. Collection 3.txt), located in the "_Configuration" sub-directory of your Coffalyser digitalMLPA package (…_Configuration\dMLPA\Barcodes).
- *-UndefinedSampleAction* parameter defines how to treat any barcodes (samples) that were not defined in the Coffalyser Definition File. The options are '*Include*' and '*Exclude*', to respectively detect and include undefined samples in the analysis or ignore and exclude undefined samples from the analysis.

The output of this conversion are Coffa files (*.coffa). The Coffa files are written to the output directory defined by the *–PathData* parameter.

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 probemix together. If manual grouping of Coffa files is preferred, group Coffa files together in individual subfolders per experiment.

All samples will be treated as (undigested) test samples. If this is not considered appropriate for your experiment, please specify your experiment in a Coffalyser Definition File.

Data analysis and reporting without Coffalyser Definition File

The previously generated Coffa files are used as input for the subsequent modules, see Table 16 below.

Table 16. Commands for Coffalyser digitalMLPA when running without Coffalyser Definition File.

Module	Command
Sample Analysis	`MRC.DataAnalysis.Analyses.Client.exe -Definition "C:\Path\MyCoffaFiles" -AnalysisType All` The *-Definition* parameter should point to an experiment directory with Coffa files. The *-AnalysisType* parameter can be set to '*All*' to run all analysis modules at once, or to '*FragmentAnalysis*', '*ComparativeAnalysis*' and/or '*ResultAnalysis*' to run the analysis modules separately.
Reporting (to generate PDF reports)	`MRC.DataAnalysis.Reporting.Client.exe -Definition "C:\Path\MyCoffaFiles"` The *-Definition* parameter should point to a directory with Coffa files that have already undergone Sample Analysis.
Excel Writer (to generate Excel reports)	`MRC.DataAnalysis.ExcelWriter.Client.exe -Definition "C:\Path\MyCoffaFiles"` The *-Definition* parameter should point to a directory with Coffa files that have already undergone Sample Analysis.
XML Writer (to generate XML reports)	`MRC.DataAnalysis.XmlWriter.Client.exe -Definition "C:\Path\MyCoffaFiles" -XmlLayout All` The *-Definition* parameter should point to a directory with Coffa files that have already undergone Sample Analysis. The *-XmlLayout* parameter can be set to '*All*' to generate XML reports for both the experiment and samples. The XML reports can also be generated for only the experiment or samples, writing '*PerExperiment*' and/or '*PerSample*'.

Coffalyser digitalMLPA User Manual – Document History
Version 002 (2 November 2026) - Clarification added for reporting of the sample sex. - Phrasing referring to next generation sequencing platforms adjusted to align with updated knowledge on platform compatibility. - Additional information added regarding DNA quantity error, noting that it can also be triggered when the MDRQ is too low and no estimation of DNA quantity can take place. - Correction of a typo in the error message in Appendix III - Troubleshooting regarding a too short read length. - Additional reference to and information added in Appendix I - Data Analysis, noting that the population of samples used for normalisation and comparison may differ from each other for certain types of probes or methods of analysis. *Version 001 (10 June 2026)* - First version.