

EmbryoMap QC Guidance

This technical note provides guidance ranges for key metrics of the EmbryoMap solution. It also outlines the various factors affecting the key metrics and includes steps for troubleshooting

Introduction

The EmbryoMap Sample Prep system simultaneously amplifies DNA from a group of cells (3-10) and prepares sequencing ready libraries for Illumina platforms. This document is intended to assist in the assessment of whether the prepared libraries and the sequencing run are within the laboratory validated sample metric ranges (Guidance ranges are shown in Table 1) when laboratories are validating their sample collection, transportation, amplification/library preparation and sequencing processes. Refer to the **EmbryoMap Sample Prep Guide** and the **EmbryoMap eMap Software Guide** for more information on how to perform the test and analyse the results.

Sample input recommendations

All biopsy samples for EmbryoMap processing must be suspended in molecular grade 1× PBS (Phosphate Buffered Saline) and, if used, with a maximum concentration of 0.5% PVP (Polyvinylpyrrolidone) in a total volume of 2.5 µl. Biopsy stability under storage and shipment conditions used (including any alternative sample collection volume, wash or collection buffers) must be fully validated by the user before use and may impact the quality and integrity of the biopsy samples, leading to suboptimal library quality that may compromise the test results.

EmbryoMap amplification

Incorporation of extraneous material containing DNA prior to amplification can lead to misleading results. This is a concern with any DNA amplification technology. Pre-PCR environmental and sampling control measures (i.e., single use of molecular grade reagents and consumables) are required to reduce the risk of sample contamination prior to input into the EmbryoMap amplification reactions. Failure to follow Pre-PCR best practices during the reaction setup may lead to contamination of the PCR products, affecting subsequent analysis and/or compromise the test results.

Human DNA contamination

Contamination with human DNA can obscure the aneuploidy status of the tested sample or lead to shifts in sex chromosome calling. The sequencing reads of human DNA contamination will also map to the human genome and the system cannot distinguish it from the DNA of the samples. The extent to which the test results are compromised depends on the amount of human DNA contamination present in the samples.

Non-human DNA contamination

Incorporation Microorganism contamination can occur during the embryo biopsy process, from non-sterile reagents or equipment, or even airborne DNA. These extraneous materials will produce PCR products that will generate sequenceable libraries.

Technote

The resulting sequencing reads will not map against the human genome database and will be filtered by the system. When the number of reads generated by the contaminant(s) is elevated, it may decrease the reads of the sample below the validated sample metric ranges (Guidance ranges are described in Table 1). This may compromise the test results due to the reduction in usable reads.

Identification of contaminants

When contamination is suspected, sequencing can be used to confirm the presence or absence of extraneous DNA. For the investigation and identification of the contaminants, a longer sequencing run is recommended.

Metagenomic apps can be used on Illumina BaseSpace Sequence Hub for the analysis of contaminants. BaseSpace Sequence Hub offers a wide variety of next generation sequencing (NGS) data analysis apps that are developed or optimised by Illumina and a growing ecosystem of third-party app providers.

EmbryoMap amplification products

EmbryoMap amplification of genomic DNA generates PCR products between approximately 0.2 to 2.5 kb. After Size Selection library fragments are optimised for Illumina sequencing, with a median fragment length of 550 bp. This can be seen in Figure 1.

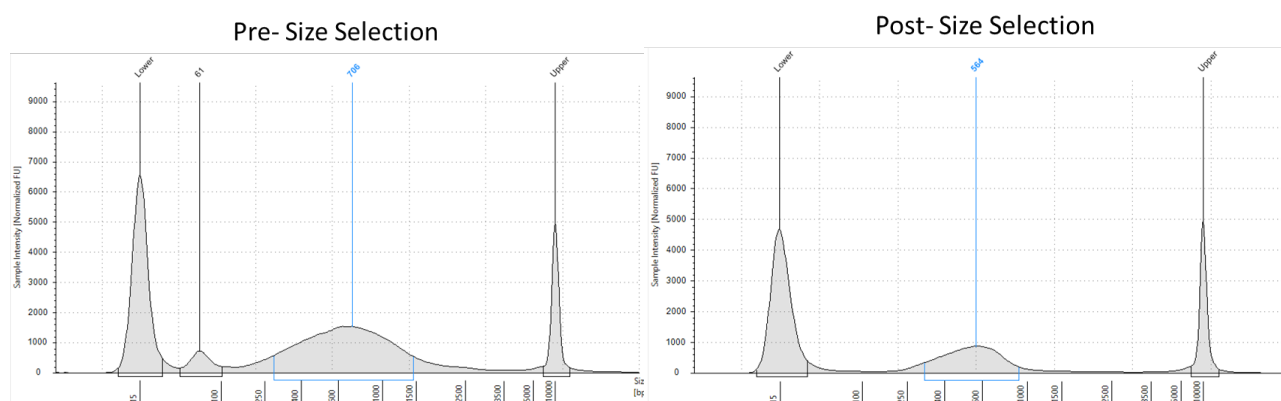


Figure 1: Fragment analysis of EmbryoMap Library Amplification products Pre- and Post- Size Selection.

Fluorometric quantification of **Size Selected** library duplex DNA shall result in positive amplification yield for an input mass of 60 pg or 5-cell isolates that is clearly separable from the yield of 'no template control' (NTC) samples. NTC or suspected amplification failures shall produce fluorescence that is equivalent to the background of the quantification kit.

NTC or suspected amplification failures may still be sequenced in an EmbryoMap run. During preparation of the 'Pooling Guide', select the '**Default**' pooling volume checkbox to include a nominal 5 µl of these samples into the final library pool. Similarly, if contamination is suspected, the NTC sample should be sequenced and reads associated with the sample's unique index should be interrogated to identify the source of contamination.

Technote

Library handling and sequencing

The eMap Software can help to assess the quality of data on two different levels, at the sequencing run (flow cell) level (Figure 2) or at a sample level (Figure 3).

Sequencing run-level metrics

The percentage of successful clusters is shown by Clusters Passing Filter (%). A low percentage of clusters PF might indicate that the amount of DNA loaded onto the instrument requires optimisation for subsequent sequencing runs.

The number of successful reads is shown by Total reads passing filter. Low reads passing filter might indicate low library quality or the presence of non-human DNA contamination.



HOME

PLANNING ▾

DATA ▾

ACCOUNT ▾

Global search

QC METRICS

Data by Run

Sequencing run information.

ID:210330_M70817_0105_000000000-JJDND

Instrument:M70817

Flowcell ID:JJDND

Workflow Version:EmbryoMap MiSeq v1.0

Reads:

Read1	Index1	Index2	Read2
36	8	8	36

Summary metrics

Cluster Density (K/mm^2):1,418 ± 36

Clusters Passing Filter (%):89.2 ± 3.9

Raw Cluster Count (M):34

Total yield (Gbp):2.6

PhiX aligned (%):0.0

PhiX error rate (%):0.00

%≥Q30 total:96.48

Indexing QC

Total reads:68,203,286

Total reads passing filter:60,830,110

Reads passing filter identified (%):98.98

Indexing CV (%):0.18

Minimum index representation (%):0.00

Maximum index representation (%):2.80

Figure 2: Example of eMap Software sequencing measurements.

Sample-level metrics

The Sample QC tab provided by the eMap Software displays metrics describing the quality of the data for each sample as it is processed from initial base calling to alignment, filtering, and copy-number calling. Reads that pass filters in the copy-number calling algorithm are indicated by 'Number of reads after filtering' in the eMap Software. The key sample-level metrics (colour shaded) are automatically checked against user-specified acceptance criteria specified in the eMap account Settings page (Figure 4).

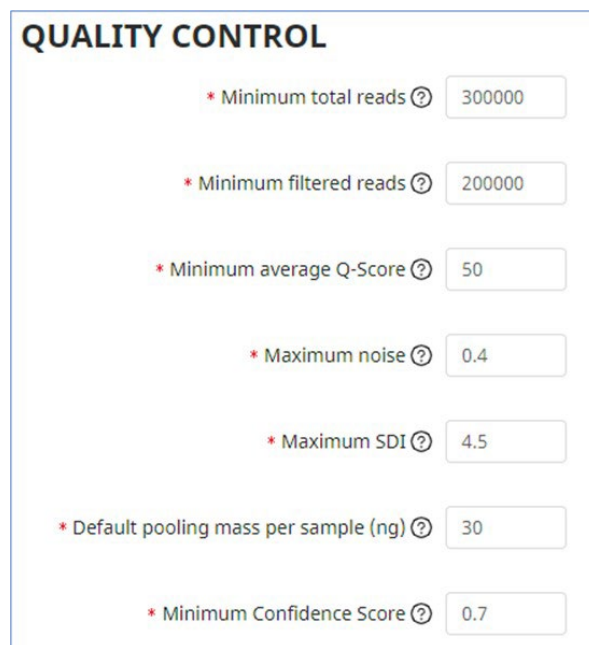
Idiogram	Sample QC	Experimental details
SAMPLE ID ?	Sample-01	
QC STATUS ?	PASS	
AVERAGE ALIGNMENT SCORE ?	57.1	
NUMBER OF TOTAL READS ?	912,379	
NUMBER OF MAPPED READS ?	897,886	
NUMBER OF READS AFTER FILTERING ?	771,261	
OVERALL NOISE ?	0.19	
SD ROBUST ?	0.21	
SDI ?	0.83	

Figure 3: eMap sample-level metrics. Coloured shading indicates where metrics have been compared to the thresholds specified.

Technote

QC status

The overall QC status is set to PASS or FAIL by default according to the sample-level QC criteria specified in the eMap account Settings. The QC Status can be manually changed by the user during the sample sign off procedure.



The screenshot shows a 'QUALITY CONTROL' settings panel. It contains seven rows, each with a red asterisk, a label, a help icon (question mark in a circle), and a text input field. The settings are: Minimum total reads (300000), Minimum filtered reads (200000), Minimum average Q-Score (50), Maximum noise (0.4), Maximum SDI (4.5), Default pooling mass per sample (ng) (30), and Minimum Confidence Score (0.7).

Setting	Value
* Minimum total reads ?	300000
* Minimum filtered reads ?	200000
* Minimum average Q-Score ?	50
* Maximum noise ?	0.4
* Maximum SDI ?	4.5
* Default pooling mass per sample (ng) ?	30
* Minimum Confidence Score ?	0.7

Figure 4: eMap Software sample-level quality control criteria may be customised by the user in account Settings. Quality criteria are used for automatic sample-level QC flagging and status.

Average Alignment Score

The alignment software calculates this score and considers the base quality score (Q-score), read length, and the number of alignments for every read. The Q-score is a measure of confidence in the base that was called at a given position. The Q-score is calculated as mean (mean [Phred33-based base quality] per read) for all reads per sample.

Total Number of Reads (Mapped Reads)

The run-level metrics described above can be used to check whether the libraries have been pooled in relatively equal proportions for sequencing. The typical total number of reads per sample range from 1.5% to 4.0% of reads passing filter for 48-plex sequencing runs. The guidance range of the total reads and reads after filtering is described in Table 3.

Low values of Mapped (aligned) reads and Reads after filtering might indicate the presence of non-human reads in your DNA source, or an elevated proportion of PCR artefacts caused by suboptimal amplification and library preparations. It is expected that out of ~50 million paired-end reads per sequencing run, several hundred reads may randomly align against random sequences.

Technote

Overall Noise (DLR-Derivative Log Ratio)

The overall sample noise measures the spread of the difference in copy number values between all bins of a chromosome. High DLR values might indicate low quality sample DNA or problems during the amplification or library preparation steps.

SDI Noise

The Summed Deviation to nearest Integer metric describes the step-wise noise across all chromosomes in a sample. CNV charts with high SDI values are more challenging to interpret for chromosome-level copy-number.

Sequencing issues

You can perform a more detailed analysis of your data using Illumina's Sequence Analysis Viewer (SAV) to troubleshoot the most common sequencing problems. If your run does not meet your expected run metrics, you should monitor the following sequencing metrics:

- For hardware and software performance of the instrument, a PhiX validation run needs to be performed in accordance with the System Guidance documentation.
- MiSeq System guide (part# 15039740) and MiSeq System User Guide (Part # 15027617).
- MiSeq i100 System documentation (<https://support-docs.illumina.com/IN/MiSeqi100Series/Content/IN/FrontPages/MiSeqi100.htm>)
- See the Illumina bulletins [What is the PhiX Control v3 Library and what is its function in Illumina Next Generation Sequencing?](#) and [How to set up the sample sheet for a PhiX validation run on the MiSeq system using Illumina Experiment Manager](#) for additional information for the MiSeq and [How to set up a PhiX validation run on MiSeq i100 Series](#) for MiSeq i100 series.
- The EmbryoMap recipe is a custom recipe designed for paired-end (2×36 bp), dual-index read sequencing of libraries on the MiSeq system prepared with the EmbryoMap Sample Prep kit only. This recipe is not suitable to assess the MiSeq instrument hardware and software performance.
- The EmbryoMap custom recipe is not required for sequencing on the MiSeq i100 series systems.
- Note that the EmbryoMap Q30 metrics vary based on the library pool quality, cluster density, and the percentage of clusters passing filter.
- To assess the sample level performance, refer to your validated sample metric ranges (See Table 1 for guidance ranges).

Technote

MiSeq Cluster Density

When assessing the quality of your EmbryoMap run, first check the cluster density against your validated sample metric ranges (Guidance ranges are described in Table 1). If optimisation is required, please refer to the Cluster Optimization guide (Illumina Part #1000000071511).

When performing the Template Line Wash, it is important that the correct concentration (0.01%) and volume (1 mL) of NaOCl is used. If the concentration or volume used is too high, it can impact subsequent runs by impeding cluster generation or even complete failure to generate clusters.

MiSeq i100 Patterned flow cells and Clusters passing filter

The MiSeq i100 series systems utilise patterned flow cells which contain fixed nanowells to allow efficient size and spacing of clusters, allowing a high cluster density to be achieved. However, this technology results in a fixed cluster density, therefore, other QC metrics must be used to determine sequencing run success. The percentage of clusters passing filter reliably informs the final count and number of reads generated for the run. In addition, it provides insight into the loading efficiency of the flow cell. For EmbryoMap libraries, a minimum clusters passing filter % of $\geq 63\%$ is expected which corresponds to a final cluster count of 25M and 50M paired-end reads for the run. Patterned flow cells often report lower percentages of cluster passing filter than random flow cells because clusters have defined locations and are all included in the final calculation. Whereas random flow cells require a template generation step to identify the clusters, this results in many dim or low-quality clusters being removed from the initial cluster count before any calculations are performed (Illumina #770-2014-043).

MiSeq i100 Optimal library loading and % Occupied

Loading libraries at a concentration that is too high or too low may lead to lower sequencing quality and yield, and, possibly, run failures in extreme conditions. Underloading may result in a low percentage of nanowell occupancy (% Occupied) and higher duplicate reads, which then requires more reads to achieve target coverage. In contrast, overloading may result in a low percentage of clusters passing filter (PF). To determine the optimal loading concentrations on the MiSeq i100 Series, the % Occupied and % PF metrics can be plotted in Sequencing Analysis Viewer to determine if a run was underloaded, optimally loaded, or overloaded (Illumina M-GI-03322). The optimal loading concentration for EmbryoMap libraries on MiSeq i100 series systems is 70-80 pM. An optimal EmbryoMap run should provide a % clusters passing filter of $\geq 75\%$ and a % occupied of $\geq 85\%$. These metrics may be reduced in runs which are underloaded or overloaded. Please refer to the MiSeq i100 Series Flow Cell Guidance Ranges in table 2.

Technote

Flow Check Error

If the MiSeq fails to measure the flow rate while starting a run, inspect the gasket of the flow cell for any damage that may have occurred during cleaning. Please see the Illumina bulletin [What should I do if my MiSeq fails to measure flow rate when starting a run?](#) for additional information.

Clean the flow cell as described in the MiSeq System User Guide. Not adhering to the flow cell best cleaning practices might raise or distort the gasket, causing flow check errors

MiSeq Flow Cell Guidance Ranges

Flow Cell	Parameters	Guidance Range	Sources of Values
MiSeq Flow cell level metrics	Density (K/mm ²)	1,100-1600	MiSeq instrument screen Display, eMap Run QC page, Illumina SAV software
	Total number of reads	50,000,000	
	Total number of reads (PF)	38,000,000	
	% of clusters passing filter	≥75%	

Table 1: EmbryoMap MiSeq Flow Cell guidance QC metrics

Note: The above metrics are guidance ranges only.

MiSeq i100 Series Flow Cell Guidance Ranges

Flow Cell	Parameters	Guidance Range	Sources of Values
MiSeq i100 Series 25M & 50M	% of clusters passing filter	≥63%	MiSeq i100 series instrument screen Display, eMap Run QC page, Illumina SAV software
	% Bases ≥ Q30	≥90%	
	% Occupied	≥80%	
MiSeq i100 Series 25M	Clusters (PF)	25,000,000	
	Total number of reads (PF)	50,000,000	
MiSeq i100 Series 50M	Clusters (PF)	50,000,000	
	Total number of reads (PF)	100,000,000	

Table 2: EmbryoMap MiSeq i100 Series Flow Cell guidance QC metrics

Note: The above metrics are guidance ranges only.

Technote

Sample Level QC Guidance Ranges

Metrics	Parameters	Guidance Range	Sources of Values
Amplification Test	Concentration of dsDNA ng/μl (5-cells)	5-15	Final library quantification after Size Selection step
Sample level metrics	Total number of reads	≥300,000	Downstream analysis in eMap software
	Number of reads after filtering	≥200,000	
	Average quality score	≥50	
	Overall noise (DLR)	≤0.4	
	SDI noise	≤4.5	
Chromosome-level metric	Region confidence (value in regions tab)	≥0.70	

Table 3: EmbryoMap Sample guidance QC metrics

Note: The above metrics are guidance ranges. Individual laboratories must validate their own sample metric ranges in accordance with their quality management system to ensure that they meet their performance and specification requirements.