VeriSeq PGS Guidance Note

This technical note provides guidance ranges for key metrics of the VeriSeq PGS application. It also summarizes best practice tips and techniques to achieve best results.

Introduction

VeriSeq™ PGS uses the SurePlex Amplification System (part #15043067) to increase the quantity of DNA in a single cell or group of cells. This amplification generates enough DNA template for next-generation sequencing (NGS) libraries sequenceable on the MiSeq Instrument to perform pre- implantation genetic screening (PGS) of all 24 chromosomes.

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 VeriSeq PGS Library Preparation Guide (part # 15052877) and the BlueFuse Multi Reference Guide (part # 15053620) for more information on how to perform the test and the analysis of the results.

Sample Input Recommendations

All biopsy samples for VeriSeq PGS processing must be suspended in molecular grade 1x PBS (Phosphate Buffered Saline) and, if used, with a maximum concentration of 0.5% PVP (Polyvinylpyrrolidone) in a 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 before use and may impact the quality and integrity of the biopsy samples, leading to suboptimal library quality that may compromise your test results.

SurePlex 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 SurePlex amplication reaction. 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 your 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.

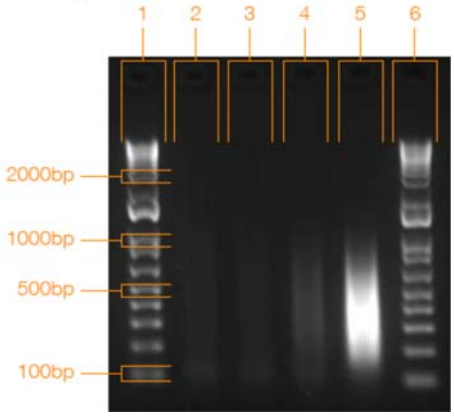
Non-human DNA Contamination

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. 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 your validated sample metric ranges (Guidance ranges are described in Table 1). This may compromise your test results due to the reduction in usable reads.

SurePlex Amplification Products

SurePlex amplification of genomic DNA generates PCR products between approximately 0.2 to 2 kb. This can be seen in Figure 1.

Figure 1. *Example of SurePlex Reactions analyzed by Gel Electrophoresis*



1.5% gel electrophoresis, 5µl of PCR products and 5µl of gel loading buffer, exposure time: 0.1ms, 20 minutes at 120 volts.

- Lane 1 and 6: 1Kb Plus DNA ladder
- Lane 2 and 3: NTC (no template control), neat dsDNA concentration <5.0 ng/µl)
- Lane 4: NTC with smear of PCR artefacts (neat dsDNA concentration = 8.4 ng/µl)
- Lane 5: Positive controls (neat dsDNA concentration = 32 ng/µl)

Samples with suspected amplification failure should not be processed and incorporated into the VeriSeq PGS library pool for sequencing.

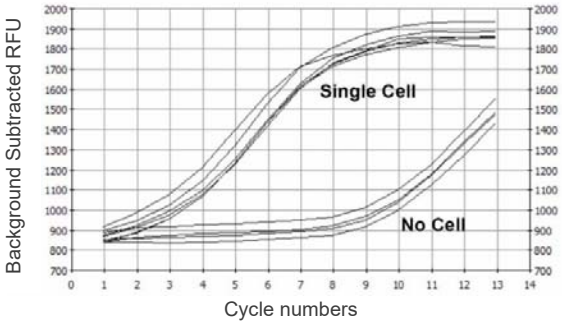
In the absence of DNA template, it is possible that PCR artefacts may be generated in any No Template Control (NTC) reactions that have been set up as batch controls using 1x PBS as input. If PCR artefacts have been generated, distinct bands should not be present. However, faint smears may be visible if assessed by gel electrophoresis.

When the NTC are quantified, the PCR artefacts are expected to be at a lower concentration than the amplified DNA present in sample or positive control reactions. Figure 1 shows an example of positive samples and NTC both with and without PCR artefacts present which have been run on a 1.5% agarose gel.

These PCR artefacts may generate sequencing reads that do not map against the human genome and are filtered out by the system.

Figure 2 shows an example of background-subtracted RFU amplification curves for replicate single-cell and control no-cell WGA reactions that were monitored on a Bio-Rad I-Cycler iQ.

Figure 2. *Real-Time PCR Detection*



- Amplification curves will have a similar appearance for all single-cell reactions, with a sloping phase followed by a “plateau” phase.
- No-cell control amplification curves are delayed (right-shifted) compared to single-cell amplification curves.

Library Handling and Sequencing

BlueFuse Multi v4.5 or later can help to assess the quality of data on two different levels, at the sequencing run (flow cell) level (Figure 3) or at a sample level (Figure 4).

Figure 3. *Example of BlueFuse Multi sequencing measurements*

Sequencing QC Measures

Sequencing Lane

Lane ID:	
Aligned Reads (PF) %:	

Sequencing Run

Cluster Number:	
Cluster Number Passing Filter:	

Sequencing Lane

- The percentage of successful reads is shown by Aligned Read Passing Filter (PF).
- Low values might indicate low library quality or the presence of non-human DNA contamination.

Sequencing Run

- The total number of clusters (or reads for the VeriSeq PGS application) and total number of PF reads.
- A low number of percentage of clusters PF might indicate that the amount of DNA added requires optimization for subsequent sequencing runs.

Sample Issues

The Sample Report provided by BlueFuse Multi displays metrics describing the quality of the data from a specific sequenced sample as it is processed from initial base calling, to alignment, filtering, and copy-number calling.

Figure 4. *Example of BlueFuse Multi metrics*

Sample QC measures

		Acceptance criteria:
Overall noise:	0.15	<0.4
Number of total reads:	974054	
Number of mapped reads:	832818	
Number of reads after filtering:	659568	>250000
Average quality score:	36.3	>30
Average alignment score:	35.4	>30

QC Status

- The overall QC status can be manually set by the operator to either Pass or Fail after interpretation of the QC performance metrics. ‘Not set’ is displayed by default.

Average Quality Score

- The base quality score (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.

Average Alignment Score

- The alignment software calculates this score and considers the Q-score, read length, and the number of alignments for every read.

Number of Reads and Mapped Reads

- Illumina's Sequence Analysis Viewer (SAV) 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 2.5% to 8.0% of reads passing filter.
- The guidance range of the total reads, mapped reads and filtered reads is described in Table 1.
- Low values of Aligned Reads (PF) and filtered reads might indicate the presence of non-human reads in your DNA source, or elevated proportion of PCR artefacts caused by suboptimal amplification and library preparations.
- It is expected that out of 25 million reads per sequencing run, a few hundred reads may randomly align against random sequences.

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.

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 to MiSeq System guide (part# 15039740) and MiSeq System User Guide (Part # 15027617).
- See the 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.
- The VeriSeq PGS recipe is a fast short-read recipe designed for 36 cycle single- and dual-index read sequencing run only. This recipe is not suitable to assess the MiSeq instrument hardware and software performance.

- The phasing and pre-phasing values of the VeriSeq PGS recipe are not suitable metrics to assess the sample level performance.
- The sequencing metrics of the VeriSeq PGS workflow differ with the PhiX Library QC workflow. When using the PhiX Control v3 Library, the quality score of the MiSeq Reagent Kit v3 is > 70% bases higher than Q30 at 2×300 bp. The Q30 of the fast short-read recipe VeriSeq PGS recipe is about 75%. Note that the VeriSeq PGS Q30 metrics vary based on the library pool quality, cluster density, and the percentage of clusters passing filter.
- To assess the sample level performances, refer to your validated sample metric ranges (See Table 1 for guidance ranges).

Cluster Density

- When assessing the quality of your VeriSeq PGS run, first check the cluster density against your validated sample metric ranges (Guidance ranges are described in Table 1). If optimization is required, please refer to the Cluster Optimization guide (Part # 1000000071511).
- When performing the Template Line Wash, it is important that the correct concentration 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.

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 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 (part # 15027617). Not adhering to the flow cell best cleaning practices might raise or distort the gasket, causing flow check errors.

Tips and Techniques

Strict adherence to laboratory best practices is critical to achieving successful assay results. Review the following recommendations before starting the SurePlex amplification and VeriSeq PGS library preparation.

Protocol Compliance

- Always use the current versions of the Reference Guides, found on the Illumina Support website.
- Review Best Practices from the VeriSeq PGS and MiSeq Instrument on Illumina support pages.
- Follow the protocols in the order shown, using the specified volumes and incubation parameters.
- Incubation times, incubation temperatures and pipetting volumes other than those specified or skipping any centrifuge or mixing steps will lead to suboptimal DNA sequencing libraries and compromise your test results.
- The integrity and shelf-life of the SurePlex PCR amplification products and libraries are affected by suboptimal storage conditions, the quality and grade of the plastic consumable used, and the nuclease-free control measure implemented.

Best Practices

- Thoroughly clean containment cabinets before starting the procedure.
- If recommended by the cabinet manufacturer, incubate tips, tubes, and pipettes with UV irradiation in the containment cabinet before starting the procedure.
- Use a clean set of calibrated and dedicated pipettes for both the Pre-PCR and Post-PCR activities.
- Briefly centrifuge all samples and reagents before use, confirming that their contents are at the bottom of the tube.
- Do not use kit components beyond the expiration date printed on labels.
- NTC should not be incorporated into the VeriSeq PGS library pool for sequencing.

Avoiding Cross-Contamination

- Avoid contamination of samples and reagents during the SurePlex and amplification library preparation steps. Contamination of reagents will compromise your test results.
- Conduct all pre-amplification activities (cell lysis, pre-amplification, and amplification setup procedures) in a dedicated environment (Pre-PCR) physically separated from amplified genetic material.
- When adding or transferring samples or primers, change tips between each sample.
- Remove unused index adapter tubes from the working area. Do not reuse index adapter tube caps. Use the provided index adapter replacement caps.

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 2x75 cycle or 2x100 cycle sequencing run can be performed on the libraries in question with the MiSeq Reagent Kit v3. Set up a sequencing run using the Library QC workflow from the Illumina Experiment Manager, as described in the MiSeq System User Guide (Part # 15027617).

Metagenomic apps can be used on 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 optimized by Illumina and a growing ecosystem of third-party app providers.

For example, the Kraken Metagenomics app analyzes sequencing data to identify the distribution of microorganisms present in a sample. The app aligns the sequence reads to a taxonomic database and classifies the reads in order to generate a report reflecting those reads originating from any contaminant DNA material.

Warnings and Precautions

- This kit is intended for use by qualified laboratory staff only. All laboratories are expected to follow Good Laboratory Practices (GLP) and have appropriate safety control measures in place.
- Check the documentation and safety data sheets (SDS) for guidelines on handling and executing the protocols.

Table 1. Flow cell and sample guidance metrics

Metrics	Parameters	Guidance Range	Sources of Values
Amplification Test	Concentration dsDNA ng/µl (single cell)	25-35	Unpurified dsDNA quantification
Flow cell level metrics	Density (k/mm2)	1,100 – 1,600	MiSeq instrument screen display or/and Illumina's Sequence Analysis Viewer
	Total number of reads	25,000,000	
	Total number of reads (PF)	19,000,000	
	% of clusters passing filter (PF)	≥ 75 %	
Sample level metrics	Number of total reads	≥ 700,000	Downstream analysis from BlueFuse Multi
	Number of mapped reads	≥ 500,000	
	Number of reads after filtering	≥ 250,000	
	% of total reads after filtering	≥ 35%	
	Average quality score	≥ 30	
	Average alignment score	≥ 30	
	Sample noise (DLR)	≤ 0.40	
Chromosome-based scores	Region Confidence (value in CNV Table)	≥ 0.70	

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

For Research Use Only. Not for use in diagnostic procedures.

Document # 1000000138837 v00

illumina