A Research Use Only Validation Report for

HTG EdgeSeq Immune Response Panel

HTG Molecular Diagnostics, Inc.
3430 E Global Loop, Tucson, AZ 85706
www.htgmolecular.com 877.289.2615

Table of Contents

List of Tables

List of Figures

List of Definitions

CI = Confidence interval
CPM = counts per million
DNA = Deoxyribonucleic acid
EDTA = Ethylenediaminetetraacetic acid
ERCC = External RNA Controls Consortium
FFPE = formalin-fixed, paraffin-embedded
ICC = Intra-class correlation coefficient
Lc = Lin's correlation concordance coefficient
LBO = Lysis Buffer Only
MiSeq = Illumina MiSeq Sequencer
mRNA = messenger RNA
MTL = Multi-Tissue Lysate
MCL = Multi-Cell Line
NextSeq = Illumina NextSeq Sequencer
NPP = Nuclease Protection Probe
POS = Positive Process Control
QC = Quality Control
qNPA = Quantitative Nuclease Protection Assay
qNPS = HTG EdgeSeq qNPA
RSD = Relative standard deviation

RUO = Research Use Only
S1 = S1 nuclease

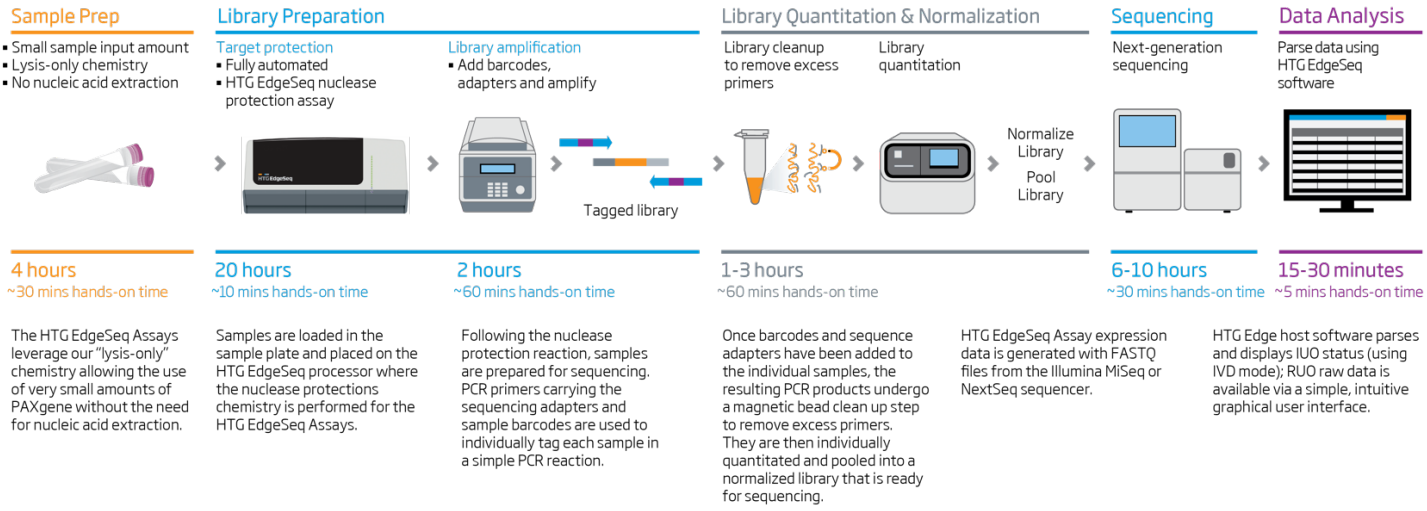
Validation Report Purpose

The purpose of this report is to describe the five studies performed to establish and validate the performance of the HTG EdgeSeq Immune Response Panel prior to release. These five studies were: Probe Design Quality Control, Probe Linearity and Sample Input, Quality Control Metrics, Precision, and Illumina Sequencer and Kit Configuration Equivalency. This report also provides the results of each study that demonstrate the HTG EdgeSeq Immune Response Panel is a robust and repeatable assay for the measurement of gene expression in PAXgene samples.

Product Overview

The HTG EdgeSeq Immune Response Panel is a Research Use Only (RUO) assay that was developed to measure mRNA expression levels of genes associated with the body's immune response to pathogenesis and disorders. The HTG EdgeSeq system combines HTG's proprietary quantitative nuclease protection assay (qNPA) chemistry with a next-generation sequencing (NGS)-based platform to enable the semi-quantitative analysis of a panel of targeted genes in a single assay (Figure 1).

Figure 1. Overview of the HTG EdgeSeq workflow



Functional DNA nuclease protection probes (NPPs) are hybridized to target mRNAs. S1 nuclease is added to digest excess non-hybridized DNA probes and non-hybridized mRNA, leaving only NPPs hybridized to mRNA fully intact and able to be amplified and barcoded. This produces essentially a 1:1 ratio of DNA detection probes to the mRNA initially targeted in the sample. RNA is digested and the NPPs are quantified by NGS.

The HTG EdgeSeq Immune Response Panel contains 2020 NPPs that measure gene expression counts of target genes, specifically 2,002 immune response-associated genes, ten external RNA Controls Consortium (ERCC) probes, four internal positive controls (POS), and four internal negative controls (ANTs). These genes are categorized into several functional groups, e.g., angiogenesis and endothelial activation, chemokine signaling, and general immune response genes. A fully annotated gene list for this assay can be accessed on the support section of our website at https://www.htgmolecular.com/support.

Probe Design Quality Control

NPPs were designed using HTG's proprietary design software as 50-mer DNA sequences. These sequences are targeted to gene transcripts and their specificity assessed *in silico* by blasting the probe sequences against the entire human transcriptome. Subsequently, the probes were synthesized, pooled, purified, and a quality control study was executed using the final probe mix. In this study, the HTG EdgeSeq Immune Response Panel probes were assessed for: the presence of all NPPs; correct sequences of all NPPs and alignment; relative concentrations of all NPPs; ability of all NPPs to undergo S1 nuclease digestion; and assessment of highly

expressed probes. Probe performance was also assessed within a titration series of a multi-cell line lysate (MCL), multi-tissue lysate (MTL), and the intended use sample of PAXgene. For PAXgene sample preparation, 500 µl of PAXgene solution is required. The cells are spun down and washed, then resuspended in 400 µl of HTG Lysis Buffer. From that lysate, 35 µl are loaded per well, representing 43.75 µl of the initial PAXgene solution. Samples utilized during this study are listed in Table 1.

Table 1. Samples Used for Probe Quality Control

Sample Type	Sample *n*	Sample Input Amount	qNPS replicates
PAXgene (multiple sclerosis, rheumatoid arthritis)	3	43.75 µL whole blood / 35 µL lysis buffer	3
Lysis buffer only (LBO)	1	N/A	28
MCL	1	625 – 10,000 cells / 35 µL lysis buffer	3
MTL	1	0.125 – 12 mm^2 / 35 µL lysis buffer	3

Results for each of the five assessments listed above are as follows:

1. **Probe Presence**
   All 2,020 probes were detected by the sequencer and were determined to be present in the probe mix.

2. **Probe Alignment Percentage**
   A pool of all NPPs was evaluated by aligning counts across the 50-mer NPP sequence, allowing for either zero or two base mismatches. The average relative alignment percentage for the 2,020 probes in the "all probes" pool ranged between 73% and 90% (median of was 87%). These results demonstrate the high sequencing fidelity of the NPPs in this assay.

3. **Relative NPP Concentrations**
   The relative amount of the probes in their respective 6X probe solution was estimated via the fitting of a Gaussian mixture model, which revealed a well-balanced probe pool with expected probe concentration distribution.

4. **S1 Sensitivity of All NPP**
   The resistance of each probe to S1 enzyme digestion was assessed using LBO samples. Figure 2 provides a graphical representation of the attenuated S1 sensitivity results for the target probes. The histogram depicts the results from the mixture model analysis. Probes clustered to the left of the histogram indicate low/no expression levels in samples that do not contain any mRNA (i.e., LBO). There are no probes within the HTG EdgeSeq Immune Response Panel that display significant resistance to S1 nuclease digestion.

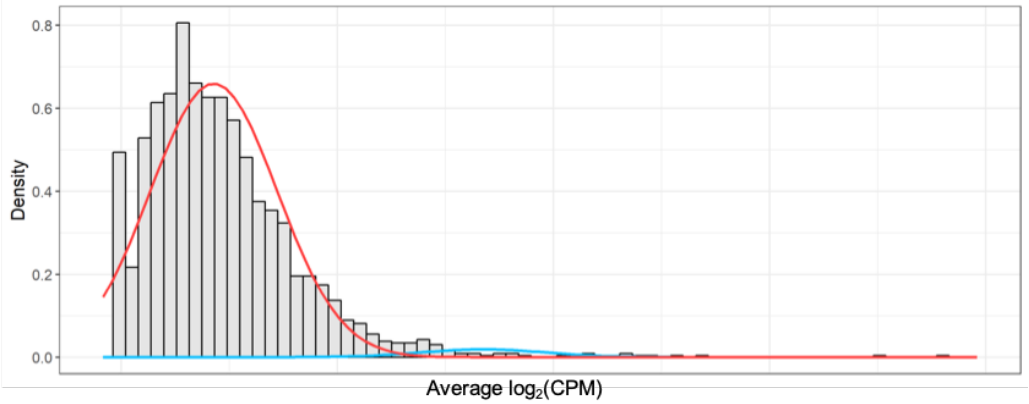


Figure 2. Attenuated S1 sensitivity of probes tested in a no-template sample (LBO).

Highly Expressed NPPs

The percentage of reads allocated to the highest expressing probes in each PAXgene sample were assessed to check for highly expressed probes. Table 2 lists the top five probes by expression for each biological sample. No NPPs were expressed at levels that were excessively high and would negatively impact the detection of low-expressing genes by taking up a significant percentage of sequencing reads.

Table 2. Highly Expressed NPPs					
PAXgene1		PAXgene2		PAXgene3	
Probe	%	Probe	%	Probe	%
HLA-B_HLA-C	4.6	HLA-B_HLA-C	4.9	HLA-B_HLA-C	6.8
B2M	3.6	B2M	3.8	S100A9	4.7
S100A9	2.8	S100A9	2.9	B2M	3.5
HLA-A	1.7	LYZ	2.4	FBXO7	2.1
NEAT1	1.7	NEAT1	2.2	ALAS2	1.9

Collectively, the results from this study demonstrate that: 1) all probes are present in the probe solution; 2) that each NPP is present in the probe solution in an equimolar ratio; and 3) all NPPs generate high sequencing fidelity. No probes were identified as resistant to S1 nuclease digestion. Lastly, no probes were highly expressed at a level which would negatively impact the ability to detect low-expressing genes.

Overall, the results of this Probe Design QC study show that the design and performance of the probes in the HTG EdgeSeq Immune Response Panel are robust and suitable for use in the RUO assay.

Probe Linearity and Sample Input Feasibility

Nuclease protection probe response in relation to sample input titration was assessed to confirm probe response linearity. Parsed data from a pooled PAXgene titration series (Table 3) were visualized by pairwise scatterplots of log_2-transformed probe counts versus log_2-transformed sample input. Based on the visualizations, the linear range for the probes was determined to be between sample inputs 2.73 – 43.75 µL / 35 µL for PAXgene.

All probes, except for a few weakly expressed probes, displayed expected linear performance with decreasing sample input. The probe linearity results suggest that the design of the HTG EdgeSeq Immune Response Panel probes is robust.

Table 3. Samples Used for Probe Linearity and Sample Input				
Sample Type	Sample Size (N)	Indications in Pool	Sample Input per qNPS Reaction	Replicates
PAXgene Pool	4 PAXgene / Pool	Rheumatoid Arthritis, Lupus, Normal, Multiple Sclerosis	0.68 – 175 µL / 35 µL	4 for each of 9 Titration Points

Determination of the recommended PAXgene sample input for the HTG EdgeSeq Immune Response Panel also used lysates from the samples prepared in Table 3. The recommended sample input for PAXgene was identified by evaluating the combination of technical replicate repeatability by Lin's Concordance Correlation Coefficient (Lc) and the relationship between mean probe expression and standard deviation of probe expression. Generally, PAXgene Lc values were consistently above 0.9 from 5.43 to 175 µL / 35 µL. A summary of Lc values is shown in Table 4.

Table 4. Summary of Lc Values.				
Sample Type	**Input (µL PAXgene / 35 µL)**	**Mean**	**Min**	**Max**
PAXgene	1.37	0.81	0.68	0.9
	2.73	0.88	0.82	0.92
	5.47	0.96	0.94	0.97
	10.94	0.91	0.74	0.98
	21.88	0.99	0.98	0.99
	43.75	0.99	0.99	1
	87.5	0.95	0.82	1
	175	1	1	1

The relationship between mean probe expression and the standard deviation of probe expression was compared for each sample input (Figure 3). Similar performance was exhibited from 21.88, 43.75, and 175 µL / 35µL for PAXgene.

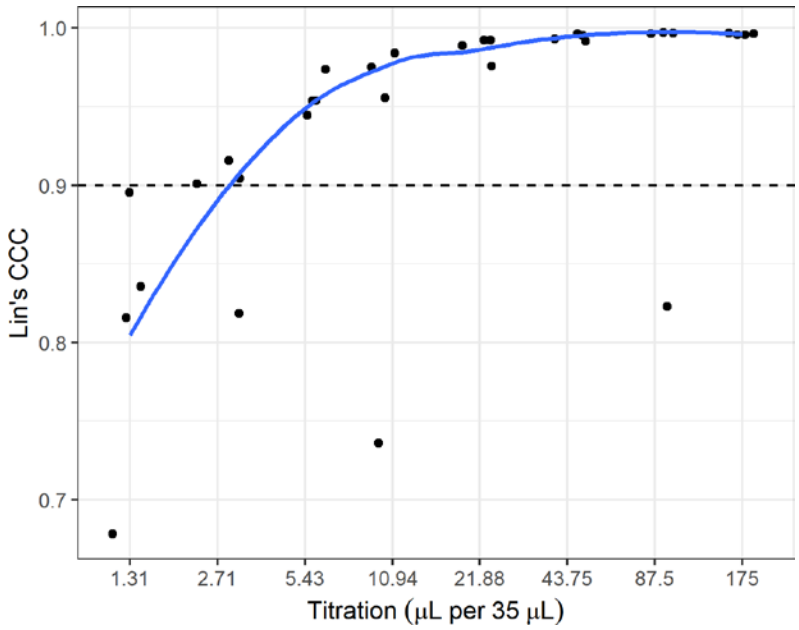


Figure 3. Lin's concordance between mean expression at each input concentration.

The result of this analysis is a recommended sample input of 43.75 µL / 35 µL for PAXgene samples to be used with the HTG EdgeSeq Immune Response Panel. The minimum required volume of PAXgene samples is 500 µL.

Quality Control Study

Post-sequencing quality control (QC) metrics for the HTG EdgeSeq Immune Response Panel were generated to identify PAXgene samples that should be removed from subsequent analysis due to one or a combination of factors. Three QC metrics that address specific post- sequencing failure modes were established. QC0, detects poor quality RNA within the sample (degraded or low concentration).
Samples with poor quality RNA are expected to have higher relative read depth for the positive (POS) control probes. The POS read depth is inversely proportional to the quality/quantity of mRNA contained within a sample, i.e., samples with high sample quality result in low read depth for POS and vice versa.

QC1, detects samples with insufficient read depth, given that sample total counts vary according to the sample size in each plate. Samples that failed to meet the established read depth requirement are excluded from further analysis.

QC2, uses relative standard deviation (RSD) of log_2-transformed counts of all genes of the Immune Response Panel minus the control probes to detect samples with low variation of counts across probes not reflective of biological expression variability.

To assess QC0, lysis buffer only (LBO), which serves as an experimental insufficient RNA/low quality sample, was added to a number of wells of an HTG EdgeSeq sample plate. These wells have no genetic material, and thus have most reads directed to the POS probes.

Assessment of QC1 is performed by down-sampling of the post-sequencing parsed data, where the number of reads is computationally reduced to simulate decreasing read depth for a given sample. Thirdly, the QC2 metric is established by inducing insufficient biological variability through the use of S1 nuclease inhibition by the addition of EDTA. The samples used during this study are provided in Table 5.

Table 5. QC Metrics Samples

Sample	Sample N	Sample Input	Replicates	QC Metric Analysis
LBO (no template control)	1	NA	48	QC0
Five-point dilution of PAXgene	3	2.74 - 43.75 µL / 35 µL	3 per dilution	QC0
PAXgene spiked with EDTA	3	43.75 µL / 35 µL + 0 – 6 mM EDTA	3 per concentration	QC2
Representative PAXgene samples	16	43.75 µL / 35 µL	3 per sample	QC1, evaluation of all metrics

1. QC0 Establishment

The QC0 metric was developed to detect poor sample quality (degraded or low concentration of RNA). Poor quality samples tend to have little or no real biological signal coming from sample mRNA. Each sample was evaluated by the percentage of counts allocated to the four positive controls probes (POS) within the sample. Figure 4 addresses percent POS for all PAXgene samples run at standard sample input and input titrations along with the LBOs. There is greater spread in percent POS in the LBOs due to higher background in those wells, resulting in more reads being distributed to all genes rather than just to the POS probes.

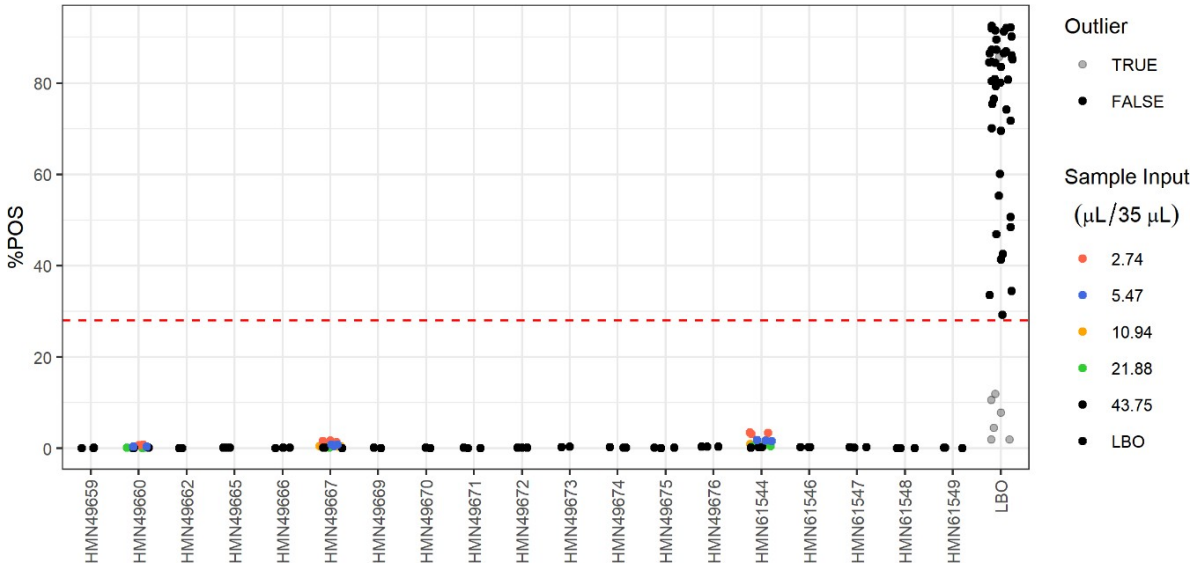


Figure 4. Percent POS by Sample. The red dashed line represents the QC0 cut-off of 28% POS.

Based on these results, a QC0 cut-off of 28% is recommended. All LBO replicates below the cutoff had elevated background and were rejected by QC2. This led to 100% of all LBO replicates being rejected by QC metrics.

2. QC1 Establishment

The QC1 metric was developed to detect samples with insufficient sequencing read depth. These failures would typically be caused by dilution or pooling errors. FASTQ files from five PAXgene samples run in triplicate at standard sample input were down sampled such that they contained a random subset of fewer reads. The result of this down sampling is that a sample that can be viewed at different read depths. A range of 50,000 to 2,000,00 reads with 40 down sampling points was used.

After the down-sampling was performed, the resulting FASTQ files were parsed. A $\log_2$CPM of the parsed data was used to investigate the effect of the read depth on sample repeatability, and the QC1 cut-off was chosen at the read depth for which acceptable repeatability was observed for probes measuring low expressed (bottom of the range) genes. The QC1 cutoff was chosen as the read depth below which unstable repeatability was observed for probes with low expression. Sample repeatability was measured by intra-class correlation coefficient (ICC) through a mixed effect model. Figure 5 illustrates the ICC values for down sampling results of all probes as well as the bottom 30%, 40%, and 50% of probes according to expression percentile for the 10 samples. The point at which ICC becomes unstable can be seen as a drop in the ICC values. While the drop is located at a different read depth in different samples, a read depth of greater than 750,000 counts ensures little difference in ICC across a variety of samples.

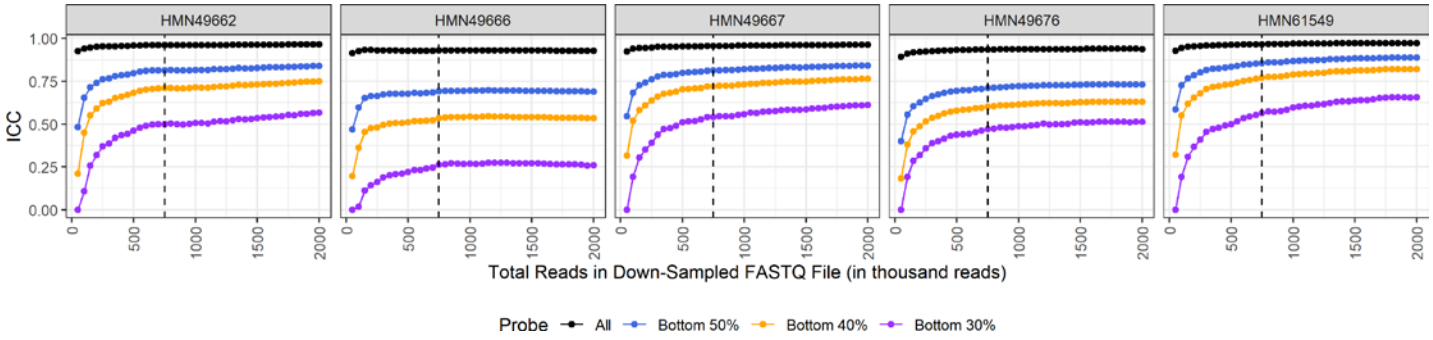


Figure 5. ICC values for sample repeatability versus read depth.

Based on these results, a QC1 cutoff of 750,000 counts has been selected. Any sample well with less than or equal to 750,000 counts are flagged as a QC1 failure.

3. QC2 Establishment

The QC2 metric detects samples with minimal expression variability across probes. These failures are typically caused by S1 nuclease inhibition, processing errors, and under some circumstances, may point to sample-related issues. Relative standard deviation (RSD) was used to assess the level of variation in expression throughout a sample. Less variation corresponds to a smaller RSD.

The addition of EDTA above 3 mM results in a bimodal RSD distribution, with lower EDTA concentrations and samples without EDTA having higher RSD values. This is caused by differences in signal variability, as samples with the lower EDTA concentrations demonstrate a level of signal variability greater than that of samples with an EDTA concentration of 6 mM due to full inhibition of the S1 nuclease activity, which leads to a loss of signal variability.

As shown in Figure 6, RSD values for samples with EDTA of 6 mM are significantly lower and less variable than samples with either 0 mM or 3 mM EDTA. Using the PAXgene observations run at an EDTA concentration of 6 mM, an RSD cutoff was set at 0.094. This cutoff detects the remaining LBOs which have initially passed QC0 and QC1. Any observations with an RSD value less than 0.094 are flagged as a QC2 failure.

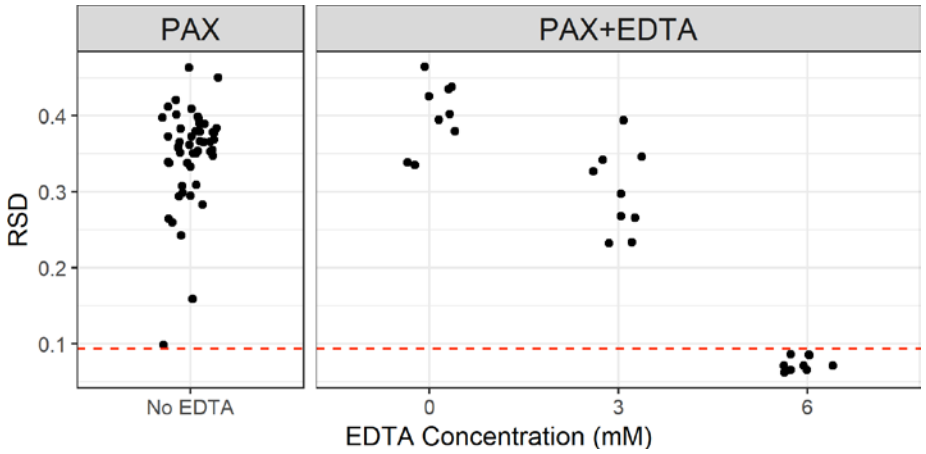


Figure 6. RSD for untreated PAXgene samples as well as samples at different EDTA concentrations.

The three QC metrics and their corresponding cutoffs are summarized in Table 6.

Table 6. QC Metrics Summary

Metric	Corresponding Failure Mode	QC Failure by Cutoff
QC0	Insufficient RNA	%POS ≥ 28
QC1	Insufficient Read Depth	Total counts ≤ 750,000
QC2	Lack of Variability	RSD ≤ 0.094

Precision Study

This study was performed to assess the precision of multiple lots of the HTG EdgeSeq Immune Response Panel, multiple operators and HTG EdgeSeq processors, multiple days, and sample replicates. Five different plates were run by three different operators, three different HTG EdgeSeq processors, and three manufacturing lots of Immune Response Panel probe mix to evaluate the assays precision (Table 7).

Table 7. Study Design

Plate/Run	Operator	Processor	Lot
1	1	1	1
2	2	2	1
3	3	3	1
4	1	1	2
5	1	1	3

Lin's concordance coefficient (Lc) values from eight separate PAXgene samples (Table 8) spanning four unique indications were analyzed with a focus on overall precision, intra-plate variability, inter-lot variability, and inter-processor/inter-operator variability and inter-replicate/within run variability (repeatability).

Table 8. PAXgene Samples

Sample Type	Sample Size (N)	Indications	Sample Input per qNPS Reaction	Replicates/Plate	Total Replicates/Sample
PAXgene	8	Rheumatoid Arthritis, Lupus, Multiple Sclerosis, Normal	43.75 µl / 35µl	5	25

Mean overall Lc values for PAXgene was 0.963 (Table 9).

Table 9. Summary of Lc Values for Overall Precision			
Sample Type	**Mean**	**Minimum**	**Maximum**
PAXgene	0.963	0.427	0.996

The five qNPS plates were processed using the same sample layout. Fresh lysates were prepared for each of the plates. Intra-plate variability displayed a high level of concordance, with average mean Lc values 0.956 (Table 10).

Table 10. Summary of Lc Values for Intra-Plate Variability			
Plate	**Mean**	**Minimum**	**Maximum**
Plate1	0.955	0.527	0.998
Plate2	0.942	0.431	0.997
Plate3	0.964	0.754	0.998
Plate4	0.961	0.686	0.998
Plate5	0.959	0.629	0.998
Mean = 0.956. 95% CI around the overall mean Lc: [0.948, 0.964]			

Three formulation lots of HTG EdgeSeq Immune Response Panel probes were manufactured to measure the variability across multiple lots. Table 11 displays that again, there is high concordance and multiple formulation lots produce equivalent results.

Table 11. Summary of Lc Values for Inter-lot Variability			
Lot Pair	**Mean**	**Minimum**	**Maximum**
Lot1 vs Lot2	0.974	0.829	0.998
Lot1 vs Lot3	0.976	0.839	0.997
Lot2 vs Lot3	0.979	0.839	0.998
95% CI around the overall mean Lc: [0.964, 0.989]			

Inter-processor and inter-operator variability were analyzed together to assess their contribution to the precision to the assay. Three operators and three processors were utilized, and the results are displayed in Table 12.

Table 12. Summary of Lc Values for Inter-Processor/ Inter-Operator Variability			
Proc Pair	**Mean**	**Minimum**	**Maximum**
Proc1 vs Proc2	0.964	0.836	0.993
Proc1 vs Proc3	0.931	0.523	0.993
Proc2 vs Proc3	0.924	0.500	0.996
95% CI around the overall mean Lc: [0.909, 0.971]			

The sum of the results of the Precision study demonstrate that there is high concordance (0.963) between assay results generated across multiple plates, lots, and operator/processor.

Illumina Sequencer and Kit Configuration Study

The HTG EdgeSeq Immune Response Panel is available in 8-, 24-, and 96-sample configurations. Each kit configuration was validated to be run on either an Illumina MiSeq or NextSeq sequencer depending on the number of samples and meeting the per sample read depth of 750,000 (Table 13).

Table 13. Recommended Illumina Sequencer

Kit Configuration	Sequencer
8-Well	MiSeq
24-Well	NextSeq
96-Well	NextSeq

To assess the equivalency of results obtained from each configuration and sequencer, libraries generated during the Limited Precision study (above) with PAXgene and multiple tissue lysate (MTL) were pooled and sequenced. Results were compared between each configuration and sequencer. Due to read depth limitations, a maximum of eight libraries can be sequenced with the Illumina MiSeq.

Inter-well comparisons between three libraries run on the Illumina MiSeq sequencer in 8-well configurations (for a total of 24 samples) were compared to the same libraries run on the Illumina NextSeq sequencer in a 96-well configuration. The 20 representative plots in Figure 7 contain the $\log_2$(CPM) values of all 2,020 probes. Lc values for each comparison are presented in each plot and range from 0.95 to 0.99 in this comparison.

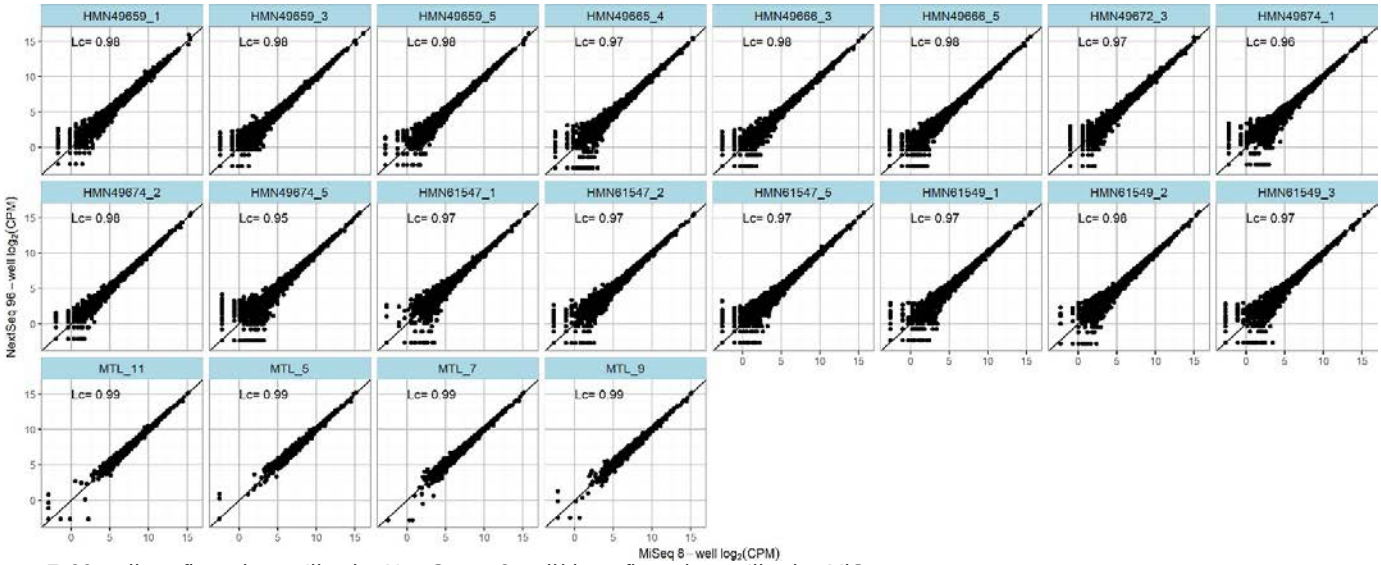


Figure 7. 96-well configuration on Illumina NextSeq vs 8-well kit configuration on Illumina MiSeq.

The inter-well comparisons of the 24-well and 96-well kit configurations sequenced on the Illumina NextSeq sequencer with 17 representative plots displayed in Figure 8. Lc values ranged from 0.92 to 0.99.

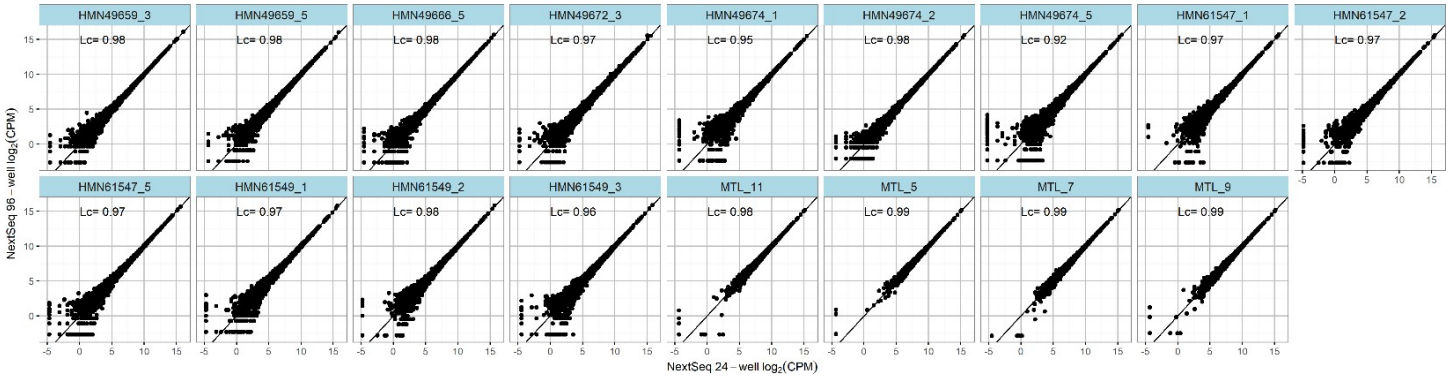


Figure 8. 96-well configuration vs 24-well kit configuration both sequenced on Illumina NextSeq.

Table 14 summarizes the results of the Illumina Sequencer and Kit Configuration Study.

Table 14. Summary of Lc Values

	Mean	Min.	Max.
PAXgene	0.97	0.92	0.98
MTL	0.99	0.98	0.99
Overall	**0.97**	**0.92**	**0.99**

The results of this study support the claim that comparable results are obtained regardless of the kit configuration and appropriate Illumina sequencer used for PAXgene samples.

HTG EdgeSeq Immune Response Panel Performance Summary

Development and validation of the HTG EdgeSeq Immune Response Panel consisted of five studies to address NPP performance, recommend sample input, establish quality control metrics, measure assay precision, and validate kit configuration equivalency. The data presented in this assay validation report confirm that the design and performance of the HTG EdgeSeq Immune Response Panel is robust and can reproducibly assess gene expression levels for 2002 immune response-related genes in blood samples that have been collected in PAXgene blood collection tubes and processed at the recommended sample input of 43.75 µl PAXgene per well.