



Evaluation of a Next-Generation Sequencing Metagenomics Assay to Detect and Quantify DNA Viruses in Plasma from Transplant Recipients

Soya S. Sam,^{*†} Ralph Rogers,[†] Fizza S. Gillani,^{*†} Gregory J. Tsongalis,[‡] Colleen S. Kraft,[§] and Angela M. Caliendo[†]

From the Division of Infectious Diseases,* The Miriam Hospital, Providence, Rhode Island; the Division of Infectious Diseases,[†] Warren Alpert School of Medicine, Brown University, Providence, Rhode Island; the Department of Pathology,[‡] Dartmouth-Hitchcock Medical Center, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire; and the Department of Pathology and Laboratory Medicine,[§] Emory University School of Medicine, Atlanta, Georgia

Accepted for publication
February 25, 2021.

Address correspondence to
Soya S. Sam, Ph.D., Division of
Infectious Diseases, The Mir-
iam Hospital, 14 Third St.,
Providence, RI 02906. E-mail:
ssam2@lifespan.org.

Viral infections are major causes of morbidity and mortality in solid-organ and hematopoietic stem cell transplant recipients. This study evaluated the performance of the Galileo Pathogen Solution metagenomics Next-Generation sequencing assay to detect and quantify 11 DNA viruses (cytomegalovirus, Epstein–Barr virus, BK virus, human adenovirus, JC virus, herpes simplex virus 1 and 2, varicella zoster virus, human herpesvirus 6A and 6B, and parvovirus B19) and to qualitatively detect torque teno virus. DNA extracted from 47 plasma samples of viremic transplant recipients were subjected to DNA library preparation with pathogen enrichment/human background depletion, sequencing, and automated data analysis. The viral loads were determined with the Galileo assay using a standard curve generated from a calibration panel. All of the samples tested had a 100% agreement with the real-time quantitative PCR (qPCR) assays in detecting the primary virus targets and the majority of the quantified samples had a viral load difference within 0.46 log₁₀ IU/mL or copies/mL. The mean difference for cytomegalovirus between the Galileo and qPCR assays was 0.21 log₁₀ IU/mL (SD, ±0.43 log₁₀ IU/mL). The mean difference for BK virus between the Galileo and qPCR assays was 0.17 log₁₀ cp/mL (SD, ±0.67 log₁₀ cp/mL). Additionally, 75 co-infections were detected in 31 samples by the Galileo assay. The study findings show that the Galileo assay can simultaneously detect and quantify multiple viruses in transplant recipients with results that are comparable with standard-of-care qPCR assays. (*J Mol Diagn* 2021, 23: 719–731; <https://doi.org/10.1016/j.jmoldx.2021.02.008>)

Opportunistic infections are a major cause of morbidity and mortality in solid-organ transplant and hematopoietic stem cell transplant recipients and occur mostly between 1 and 12 months after transplantation.¹ Viral infections that affect these immunocompromised individuals can be particularly challenging to efficiently diagnose and treat. Cytomegalovirus (CMV), Epstein–Barr virus (EBV), and BK virus (BKV) are among the most common viral infections after allogeneic transplantation. CMV infection occurs most often via reactivation of a latent infection permitted by impaired T-cell–mediated immunity and phagocytic function. Clinical manifestations may vary from mild to severe and life-threatening disease, and may present as a nonspecific systemic viral illness or a localized tissue-invasive infection, most commonly including pneumonitis, hepatitis, colitis,

and retinitis.^{2,3} EBV is associated with post-transplant lymphoproliferative disease after both solid-organ transplant and hematopoietic stem cell transplant.⁴ BKV infection has been an obstacle to improved outcomes in transplant recipients with polyomavirus-associated nephropathy and polyomavirus-associated hemorrhagic cystitis, the major complications reported in kidney transplant recipients and hematopoietic stem cell transplant recipients, respectively.^{5,6}

Other clinically relevant viral infections after transplantation include varicella zoster virus (VZV), herpes

Supported in part by Lifespan, Brown, Boston Medical Center, Boston University Center for AIDS Research grant P30AI042853, and Arc Bio, LLC.

Disclosures: None declared.

simplex virus (HSV) types 1 and 2, human herpesvirus 6 (HHV-6A and HHV-6B), polyomaviruses such as JC virus (JCV), as well as adenovirus and parvovirus B19.^{1,7,8} Beyond the direct pathogenic effects of all of these viral infections, the immunomodulatory effects of herpesviruses in particular also indirectly may impact graft outcomes as well as an individual's susceptibility to other infections and inflammatory disorders.⁹

Torque teno virus (TTV) has been implicated as a biomarker of immune status in individuals receiving immunosuppression after solid-organ transplant.^{10–12} One study reported an association between high levels of TTV with increased immunosuppression in kidney transplant recipients, and showed that quantified TTV titers assisted in identifying individuals at risk for infection approximately 3 months before infections manifested clinically.¹³ Therefore, detection and quantification of TTV might be a promising biomarker to reduce infections by allowing for personalized tailoring of immunosuppressive therapy.

Next-generation sequencing technologies are in widespread use for the diagnosis and monitoring of infectious diseases because of their manifold applications in pathogen identification. In contrast to targeted next-generation sequencing assays, metagenomics or shotgun next-generation sequencing (mNGS), allows for an unbiased, hypothesis-free testing approach to identify a broad array of infections noninvasively.¹⁴ Furthermore, beyond pathogen identification, mNGS has the potential to simultaneously provide information regarding virulence and antimicrobial resistance genes.¹⁵ The utility of mNGS in clinical practice has been evaluated in several studies using a variety of sample types to detect clinically relevant pathogens including bacteria, virus, and fungi.^{16–18} The Galileo Pathogen Solution (Galileo) (Arc Bio, LLC, Scotts Valley, CA) is a Research Use Only sample to resulting in an unbiased (shotgun) metagenomics assay designed to simultaneously detect and quantify 11 DNA viruses [HSV-1, HSV-2, VZV, CMV, EBV, HHV-6A, HHV-6B, BKV, JCV, human adenovirus (HAdV), and parvovirus B19] and to qualitatively detect TTV. The objective of this study was to evaluate the clinical performance characteristics of the precommercial version of the Galileo assay compared with standard-of-care real-time quantitative PCR (qPCR) assays using residual plasma from viremic transplant recipients.

Materials and Methods

Clinical Samples

EDTA samples chosen for evaluation included those that tested positive by qPCR assays in routine clinical testing for

one of the transplant-related viruses (CMV, BKV, EBV, HAdV, HSV-1, HSV-2, VZV, HHV-6A, HHV-6B, JCV, and parvovirus B19) and had sufficient residual volume to test for both the Galileo assay and to perform qPCR for additional viruses. A total of 50 residual retrospective and prospective plasma samples were evaluated in this study using Galileo precommercial library preparation (Galileo Viral Panel, beta version) and analysis software (Galileo Analytics, beta version). All samples were refrigerated for up to a week before aliquoting and samples were kept frozen at -80°C until testing.

The samples for the study were obtained from multiple institutions. The CMV samples were quantified originally using either Qiagen Artus (Valencia, CA) or Abbott Realtime assays, (Abbott Molecular Inc., Des Plaines, IL) and BKV and HSV-1 were quantified using laboratory-developed tests. The study was approved by the Institutional Review Board of The Miriam Hospital.

DNA Extraction, Library Preparation, and Sequencing

DNA was extracted from 400- μL samples/Galileo controls on the EZ1 Advanced XL system using the EZ1 DSP Virus Kit (Qiagen, Inc. Valencia, CA) and eluted into 60 μL . Sequencing libraries were prepared according to the manufacturer's protocols (Arc Bio, LLC). After DNA extraction, the eluates were concentrated using magnetic beads (KAPA Pure Beads; Roche; Roche, Wilmington, MA). These samples then were subjected to enzymatic fragmentation and end repair (Arc Bio, LLC) at 37°C for 5 minutes followed by 65°C for 30 minutes. DNA fragments then were adapter-ligated using dual-index sequencing adapters (Arc Bio, LLC) at 20°C for 15 minutes and purified using magnetic beads. This was followed by pathogen enrichment and human DNA depletion using depletion reagents (Arc Bio, LLC) at 45°C for 2 hours, followed by 70°C for 15 minutes, and purified using magnetic beads. Pathogen-enriched libraries then were amplified using amplification reagents (Arc Bio, LLC) with the following PCR cycling conditions: initial denaturation at 98°C for 30 seconds, followed by 14 cycles of denaturation at 98°C for 10 seconds, annealing at 65°C for 75 seconds, and extension at 65°C for 5 minutes. Fragmentation, adapter ligation, human DNA depletion, and amplification reactions were performed on the Mastercycler Pro S platform (Eppendorf, Enfield, CT). The assessment of amplicons was performed using a 2% eGel (Thermo Fisher Scientific, Waltham, MA). Appearance of faint smears ranging from approximately 200 to 900 bp confirmed the presence of amplified PCR products and these then were purified using magnetic beads. Libraries were quantified using the

Qubit fluorimeter and the high-sensitivity assay kit (Thermo Fisher Scientific). Quality was assessed using the Agilent 2100 Bioanalyzer and the Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA). Those libraries with adapter-dimer concentration visualized above 5% were subjected to an additional purification step before pooling to sequence. Based on the concentration and size distribution of each library, libraries were pooled using the Pooling Calculator available within the Galileo Analytics software (Arc Bio, LLC). The prepared pool then was quantified on an ABI ViiA 7 Real Time PCR instrument (Thermo Fisher Scientific) using a KAPA low Rox library quantification kit

(Roche) at three different dilutions and each dilution was performed in triplicate. Libraries then were sequenced on the NextSeq550 sequencer (Illumina, San Diego, California) using the paired-end, high-output v2.5 kit. The workflow of the Galileo assay is shown in [Figure 1](#).

Galileo Calibration Panel and Controls

Before clinical sample testing, an initial calibration run was performed using a multianalyte panel of whole-virus particles (Arc Bio, LLC) to generate a standard curve and estimate the Galileo viral load in the clinical samples. The panel consisted of 11 viruses (CMV, EBV, BKV, HAAdV, JCV, HSV-1, HSV-

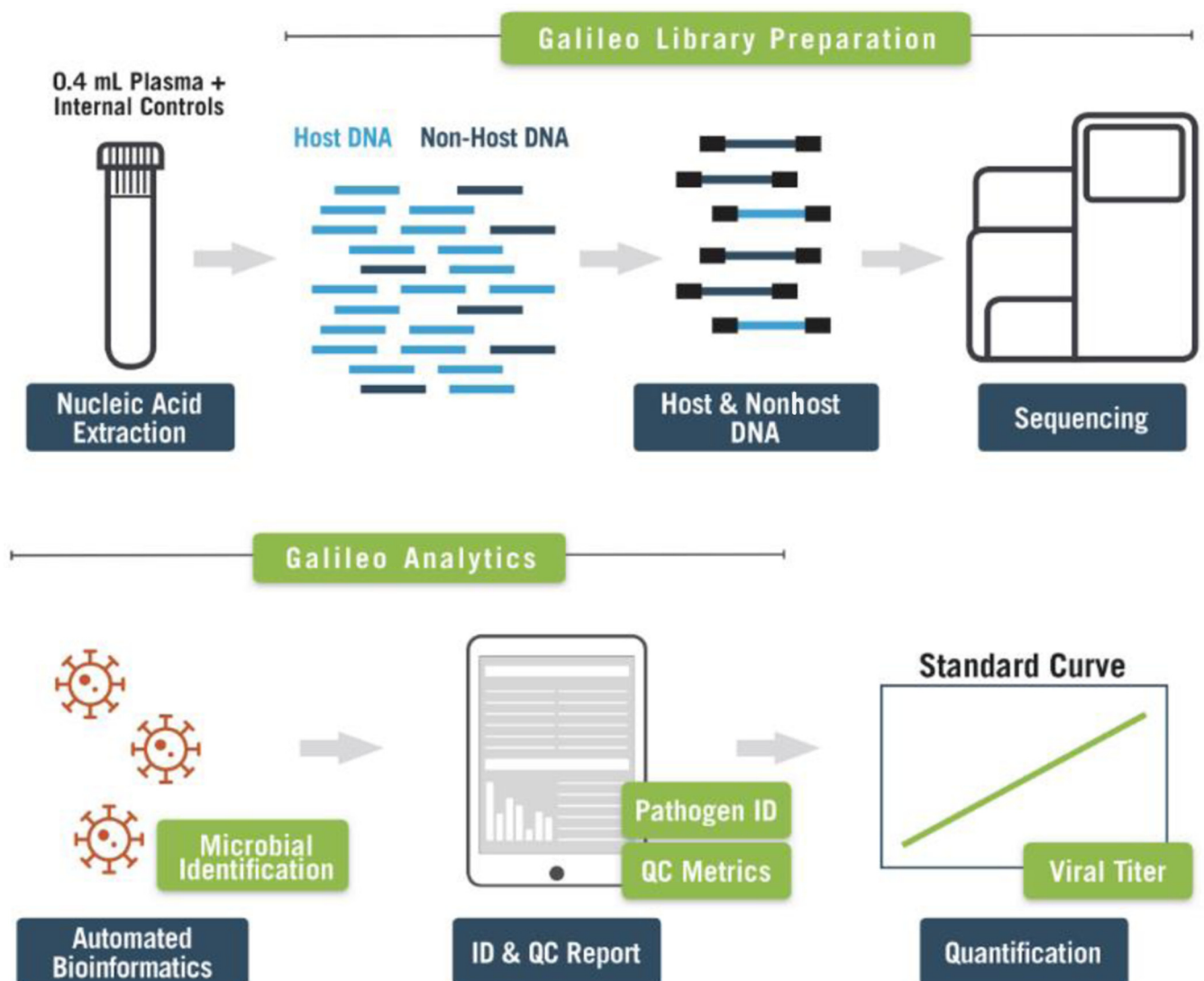


Figure 1 Galileo Pathogen Solution (GPS) workflow. The DNA extracted is converted into next-generation sequencing libraries using the Galileo kit (Arc Bio, LLC). After sequencing, the Galileo Analytics informatics pipeline produces quality control (QC) and pathogen identification (ID) reports from demultiplexed fastq files.

Table 1 Detection and Quantification of Samples by Galileo and qPCR Assays with Additional Viruses Detected: CMV (ES1-SOS65), BKV (SOS43-SOS66), and HSV-1 (SOS64)

Sample identification/virus	Galileo assay, log ₁₀ cp/mL or IU/mL	qPCR assay, log ₁₀ cp/mL or IU/mL	Log ₁₀ difference	Additional viruses detected			
				Galileo assay, log ₁₀ cp/mL or IU/mL	qPCR assay, log ₁₀ cp/mL or IU/mL	Log ₁₀ difference	
ES1/CMV	4.61	4.4	0.21				
ES2/CMV	4.53	4.54	0.01	EBV	2.89	4.1	1.21
ES3/CMV	6.82	6.52	0.3	EBV	2.74	3.7	0.96
				HHV-6A	2.54	ND	
				HHV-6B	2.83	ND	
ES4/CMV	3.62	3.34	0.28				
ES5/CMV	5.24	4.94	0.3	EBV	2.05	1.51	0.54
ES6/CMV	4.1	3.85	0.25	EBV	2.22	4.56	2.34
ES7/CMV	3.9	3.85	0.05	EBV	2.06	4.46	2.4
ES8/CMV	3.91	3.66	0.25	EBV	2.86	5.13	2.27
				BKV	3.11	3.01	0.1
ES9/CMV	3.8	3.62	0.18				
ES10/CMV	3.64	3.45	0.19				
SOS15/CMV	4.34	4.46	0.11				
SOS16/CMV	5.08	5.29	0.21	ADV	1.66	ND	
SOS17/CMV	2.19	3.32	1.13	HHV-6A	2.23	ND	
				HHV-6B	2.2	0.77	1.43
SOS18/CMV	3.86	4	0.14	EBV	2.02	3.87	1.85
SOS19/CMV	5.49	5.49	0				
SOS20/CMV	5.42	5.36	0.06				
SOS21/CMV	4.92	4.92	0	HHV-6A	0.7	ND	
				HHV-6B	1.15	ND	
SOS22/CMV	4.12	4.16	0.04				
SOS23/CMV	4.67	4.55	0.12				
SOS24/CMV	4.5	4.67	0.17	HHV-6A	1.23	ND	
				HHV-6B	1.15	ND	
SOS29/CMV	3.03	2.4	0.63				
SOS30/CMV	2.27	2.42	0.15	BKV	4.66	5.28	0.62
				JCV	2.88	0.6	2.28
SOS31/CMV	6.12	5.85	0.27	EBV	2.77	2.83	0.06
				VZV	1.59	ND	
SOS33/CMV	4.16	3.48	0.68	EBV	2.82	3.18	0.36
				VZV	3.04	3.26	0.22
				B-19	5.18	3.01	2.17
SOS34/CMV	3.88	3.15	0.73	BKV	5.19	2.93	2.26
				JCV	2.82	ND	
SOS35/CMV	3.35	2.8	0.55				
SOS37/CMV	3.71	3.27	0.44	HHV-6A	2.4	ND	
				HHV-6B		1.76	
				VZV	2.14	2.44	0.3
				HSV-1	2.54	ND	
SOS38/CMV	3.64	2.86	0.78				
SOS65/CMV	5	4.97	0.03	JCV	2.69	ND	
SOS43/BKV	2.82	2.18	0.64	CMV	1.57	ND	
SOS44/BKV	2.93	3.13	0.2				
SOS45/BKV	3.37	2.66	0.71	ADV	1.52	ND	
SOS46/BKV	3.58	3.12	0.46				
SOS47/BKV	4.56	3.9	0.66	JCV	0.48	ND	
				CMV	1.48	INSV	
				ADV	1.69	INSV	
SOS48/BKV	4.8	4.53	0.27	JCV	1.54	ND	
SOS49/BKV	5.07	4.77	0.3	JCV	1.93	ND	

(table continues)

Table 1 (continued)

Sample identification/virus	Galileo assay, log ₁₀ cp/mL or IU/mL	qPCR assay, log ₁₀ cp/mL or IU/mL	Log ₁₀ difference	Additional viruses detected			
				Galileo assay, log ₁₀ cp/mL or IU/mL	qPCR assay, log ₁₀ cp/mL or IU/mL	Log ₁₀ difference	
S0S50/BKV	5.6	5.26	0.34	JCV	2.87	ND	
S0S51/BKV	6.13	5.76	0.37	JCV	4.35	1.72	2.63
S0S52/BKV	6.26	5.8	0.46	JCV	4.31	1.76	2.55
S0S57/BKV	5.23	4.4	0.83	JCV	1.73	ND	
S0S58/BKV	4.93	4.31	0.62	JCV	0.6	ND	
S0S59/BKV	6.07	5.87	0.2	JCV	4.24	1.77	2.47
S0S60/BKV	4.48	3.62	0.86	JCV	0.7	ND	
S0S61/BKV	5.15	4.79	0.36	JCV	2.41	0.3	2.11
S0S63/BKV	2.15	2.12	0.03				
S0S66/BKV	2.97	2.16	0.81				
S0S64/HSV-1	2.7	1.79	0.91	HSV-2	0.48	ND	

ADV, adenovirus; BKV, BK virus; CMV, cytomegalovirus; EBV, Epstein–Barr virus; HHV, human herpesvirus; HSV, herpes simplex virus; INSV, insufficient volume; JCV, JC virus; ND, not detected; qPCR, real-time quantitative PCR; VZV, varicella zoster virus.

2, VZV, HHV-6A, HHV-6B, and parvovirus B19) in a pooled plasma background matrix at viral loads of 0, 1000, 5000, 10,000, and 100,000 copies/mL or IU/mL plasma and was processed in quadruplicate. Additional full-process positive controls that contain whole viruses of all viruses were included with the Galileo assay; a positive control (10,000 copies/mL or IU/mL multianalyte mix in plasma), a negative control (pooled plasma), and high-run controls (100,000 copies/mL or IU/mL multianalyte mix in plasma) and low-run controls (5000 copies/mL or IU/mL multianalyte mix in plasma). Spikes in internal controls were added to all samples/full-process controls before extraction. These full-process controls were processed alongside each run of 10 clinical samples per NextSeq batch (five batches of 10 samples plus four controls were tested in total). Twenty calibration and 50 clinical sample libraries were sequenced per high-output NextSeq run. Viral loads were quantified using the Galileo standard curve and compared with the respective qPCR viral load results.

qPCR Assays

Any additional viruses detected with the Galileo assay subsequently were confirmed using singleplex qPCR assays after DNA extraction of plasma samples on either the EZ1 Advanced XL or QIA Symphony (QIAGEN, Inc.) extraction platforms. The Artus EBV RG-PCR (Qiagen, Inc.), Artus BKV RG-PCR (Qiagen, Inc.), Artus CMV RGQ-MDx (Qiagen, Inc.), RealStar JCV PCR (Altona Diagnostics, Plain City, OH), RealStar HHV-6 A/B PCR 1.0 (Altona Diagnostics), HAdV

PCR (Diasorin Molecular, Cypress, CA), and TTV R-GENE PCR (bioMérieux, Chicago, IL) assays were performed following the manufacturer's protocol. Samples that detected positive for VZV, parvovirus B19, and HSV-1 and HSV-2 with the Galileo assay were referred for testing at the Viracor Eurofins reference laboratory (Lee's Summit, MO).

Bioinformatics Analysis

System-level NextSeq quality metrics, including error rate, cluster density, and cluster passing filter, were evaluated according to the manufacturer's recommendations (Illumina, Inc.). The sample sheet was downloaded from Galileo Analytics (Arc Bio, LLC), and demultiplexing then was performed using bcl2fastq 2.20 with default parameters and no lane splitting. The resulting FASTQ files were uploaded and analyzed to the Galileo Analytics cloud-based software, which automatically processes uploaded FASTQ files from both samples and controls and produces a quality control (QC) report and a microbial identification report for each library.

Galileo Analytics uses an alignment module with raw data from the uploaded FASTQ files transformed into a proprietary signal value, taking into account complexity, unique placement, and alignability of mapped reads. This signal value normalizes read counts across libraries, for differing genome lengths, and for technical bias via the synthetic spiked-in normalization controls. The final result is a reported signal, or evidence value, related to genomic depth and the likelihood of observing the nucleic acid of the

viruses in the sample, including nucleic acid belonging to nonconfounding genomic regions. The signal value enables quantitative evaluation of viral load via a standard calibration curve and the ability to compare results across different libraries and different runs.

Run-level quality control criteria were defined using the negative matrix and positive external controls. The negative matrix control was expected to yield no signal for each of the target viruses. The external positive control (10,000 IU or copies per mL) was expected to yield signal values within predefined ranges based on the manufacturer's internal QC data (Arc Bio, LLC). In addition, library-level QC metrics were reported in the QC report. All libraries, including the run-level controls, were recommended to be sequenced to a minimum of 30 million total reads and a minimum of 250,000 nonhuman reads, with 80% of bases having a Q score of 30 or greater and 85% of bases having a Q score of 20 or greater, according to the Illumina NextSeq 500 system specifications. GC-content was expected to be 35% to 50% because the majority of the DNA was of human origin. In addition, the synthetic normalization controls were expected to yield signal values in a predefined range based on the manufacturer's internal QC data (Arc Bio, LLC). For evaluation of the clinical specimens, a minimum of 250,000 nonhuman reads or at least 30 million reads per library were required for subsequent analysis.

Statistical Analysis

Data analysis was performed using SAS software, version 9.4 (SAS Institute, Cary, NC). The correlation between quantitative results was evaluated using linear regression analysis. The Bland–Altman¹⁹ analysis and plots were used to assess the relative agreement between two analytical methods. The difference between two assay results were plotted with additional reference lines (0 bias line, 95% upper, and 95% lower confidence levels).

Results

Among the 50 samples chosen for the evaluation, three samples failed QC and were removed from the study: two CMV samples did not have adequate DNA concentrations and therefore failed to produce an expected peak when fragments were analyzed. The third sample, a prospective BKV sample that was detected at a low concentration with the Galileo assay, was undetected by the qPCR and therefore was removed from the analysis. The remaining 47 samples analyzed generated an average of 49.5 million (range, 22.3 to 91.1 million) paired-end reads per sample. The average nonhuman reads generated was 2.07 million (range, 330,681 to 7.92 million) paired-end reads per

Table 2 TTV Viral Signal Score from Galileo Assay and qPCR Viral Load

Sample identification	qPCR assay log ₁₀ cp/mL	Galileo signal score
ES1	5.69	5.29
ES2	6.01	5.67
ES3	8.26	7.76
ES5	5.23	5.14
ES9	7.16	6.97
ES10	7.29	7.01
S0S15	7.43	7.06
S0S16	7.10	6.85
S0S18	5.95	6.40
S0S22	7.60	7.04
S0S23	7.58	7.21
S0S24	5.69	6.42
S0S29	6.98	6.51
S0S30	5.67	5.63
S0S31	5.30	4.92
S0S34	7.02	6.43
S0S35	8.03	7.04
S0S37	8.03	7.31
S0S43	4.39	3.81
S0S45	5.91	5.50
S0S46	5.27	7.01
S0S48	7.62	7.63
S0S49	5.59	5.35
S0S50	6.13	6.06
S0S57	3.81	3.70
S0S58	7.75	7.24
S0S60	5.56	4.85
S0S63	4.80	3.70
S0S65	8.24	7.47

qPCR, real-time quantitative PCR.

sample. All libraries contained an insert size of approximately 310 bp.

The calibration panel tested at different concentrations (100,000 IU or cp/mL, 10,000 IU or cp/mL, 5000 IU or cp/mL, or 1000 IU or cp/mL) was found to be linear with 100% sensitivity with the exception of parvovirus B19. Parvovirus B19 is a single-stranded DNA virus and as such only double-stranded DNA (dsDNA) replicates were sequenced using the Galileo assay, resulting in a higher limit of detection. After this study, parvovirus B19 was removed from the final calibration panel by Arc Bio, LLC. In addition, low-positive signals for HAdV and BKV in negative process controls were observed in two of the eight runs (<10 IU or cp/mL). But the observed titers were below the limit of detection when analyzed with a standard curve.

Forty-seven DNA samples (29 CMV, 17 BKV, and 1 HSV-1) from patients who had undergone either hematopoietic stem cell transplant or solid-organ transplant

were subjected to metagenomic sequencing using the Galileo assay. All of the samples tested had a 100% agreement with the qPCR assays in accurately detecting the primary virus targets. Furthermore, all samples were quantified by the Galileo assay with 70% of samples having a log difference within 0.46 log₁₀ IU/mL or copies/mL and the log₁₀ difference for the rest of the samples ranged from 0.55 to 1.13 log₁₀ IU/mL or copies/mL.

In addition, 75 co-infections were detected in 31 of the 47 samples by the Galileo assay and these included all 12 virus targets that the assay can detect and/or quantify (Tables 1 and 2). There were 29 TTV co-infections identified by the Galileo assay on the basis of the viral signal values (Table 2). Forty-six of the 75 co-infections had a virus other than TTV detected by the Galileo assay. Of the 46 co-infections, 21 had detectable viral loads with both Galileo and the respective qPCR (Table 1), 5 had a viral load within 0.5 log₁₀ copies/mL or IU/mL, 3 had a viral load within 1.0 log₁₀ copies/mL or IU/mL, and the remainder of the co-infections had titers ranging from 1.21 to 2.63 log₁₀ copies/mL or IU/mL. Twenty-two co-infections (JCV-9, HHV-6A-4, HHV-6B-3, HSV-1-1 and HSV-2-1, CMV-1, HAdV-2, and VZV-1) that had low viral titer with the Galileo assay were not detected by the respective qPCR assays. Two additional targets (CMV and adenovirus) detected in sample SOS47 could not be confirmed owing to insufficient sample material. Sample SOS37, which had HHV6-A detected by the Galileo assay, was identified as HHV6-B by the qPCR assay. The low specificity observed with BKV and JCV samples was analyzed further. Based on secondary analyses performed by the Arc Bio, LLC, team, it was found that seven JCV samples (SOS34, SOS47, SOS48, SOS49, SOS57, SOS58, and SOS60) that were reported as positive with Galileo reverted to negative, thus agreeing with the qPCR results when only unique regions were taken into account. However, the result of SOS50 remained positive with the secondary analyses.

Of the 9 retrospective EBV samples, 7 EBV samples were quantified lower by the Galileo than with the Qiagen Artus assay, with a mean log₁₀ difference of 1.65 log₁₀ IU/mL (range, 0.54 to 2.34 log IU/mL), whereas the mean log₁₀ difference of two prospective samples was 0.21 log₁₀ IU/mL. It was speculated that the observed disparity between the retrospective and prospectively collected samples could be attributed to the age of the samples, although all were stored at -80°C before testing. To evaluate the hypothesis, the initial troubleshooting was performed with the known Galileo assay full-process controls at three different concentrations (100,000 IU/mL, 10,000 IU/mL, and 1000 IU/mL). These were DNA-extracted and tested in triplicate using

Table 3 Re-Analysis of EBV Using Prospective Samples to Investigate Viral Load Disparity Seen with Galileo and qPCR Assays

Sample identification	Galileo assay, log ₁₀ IU/mL	qPCR assay, log ₁₀ IU/mL	Log difference
EBV 1	ND	3.70	
EBV 2	3.54	3.62	0.08
EBV 3	3.54	4.14	0.60
EBV 4	4.82	3.46	1.36
EBV 5	ND	3.38	
EBV 6	ND	3.38	
EBV 7	2.77	3.43	0.66
EBV 8	3.87	3.96	0.09
EBV 9	3.5	3.66	0.16
EBV 10	3.72	4.56	0.84
EBV 11	4.75	4.52	0.23
EBV 12	3.79	3.41	0.38
EBV 13	3.65	3.70	0.05
EBV 14	ND	3.65	
EBV 15	2.51	3.40	0.89
EBV 16	2.94	3.56	0.62
EBV 17	2.42	3.23	0.81
EBV 18	2.75	3.20	0.45
EBV 19	1.82	3.18	1.36
EBV 20	ND	3.18	

EBV, Epstein–Barr virus; ND, not determined; qPCR, real-time quantitative PCR.

the same Qiagen Artus qPCR assay. All results obtained were within 0.5 log₁₀ IU/mL of expected values. The relevant EBV FASTQ files were further analyzed with third-party BioX analysis tools that do not use any filters to rule out the Galileo Analytics filters being set too conservatively, and to see if third-party BioX tools would recover roughly the same proportion of useable reads. Kraken, Diamond, and BWA^{20–22} were run to compare against the qPCR/Galileo viral load ratio and obtained similar results to Galileo Analytics (data not shown).

To verify the disparity in viral loads of the seven retrospective EBV samples, 20 prospectively collected EBV samples (stored at -80°C for a maximum of 6 months) were sequenced with the Galileo assay and compared the viral loads with the original results from the Qiagen Artus qPCR assay. In 15 of 20 samples, a mean log₁₀ difference of 0.57 log₁₀ IU/mL was observed between the Galileo and qPCR assays. The remaining five samples were not detected by the Galileo assay, the viral load values in the Qiagen Artus qPCR test ranged from 3.18 to 3.70 log₁₀ IU/mL. The results of this testing support the hypothesis that the initial significant discrepancies between the qPCR and Galileo assays was the result of the age of the specimens (Table 3).

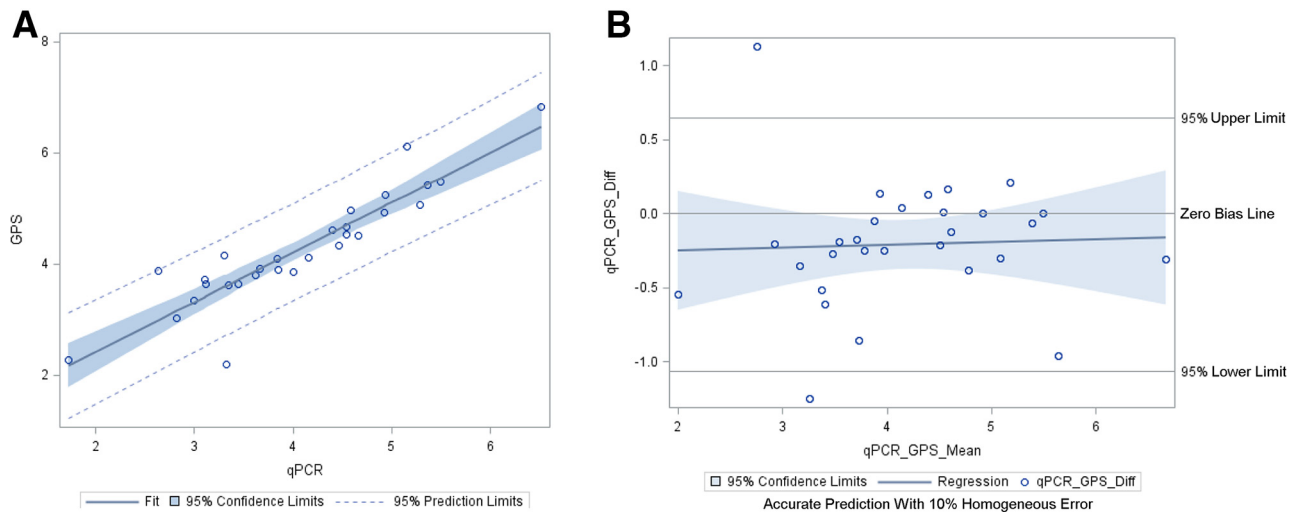


Figure 2 \log_{10} -transformed quantitative viral load agreement of cytomegalovirus (CMV) determined by both the Galileo Pathogen Solution (GPS) and real-time quantitative PCR (qPCR) assays. The linear regression line was calculated as follows: $y = 0.89x + 0.63$ and $R^2 = 0.83$. The Bland–Altman plot showed a mean difference of $0.21 \log_{10}$ IU/mL. **A:** CMV linear regression. **B:** CMV Bland–Altman plot. Diff, difference.

The analysis of all 29 CMV samples (all primary samples) resulted in a mean \log_{10} difference of $0.21 \log_{10}$ IU/mL (SD, $\pm 0.43 \log_{10}$ IU/mL) between the Galileo and Qiagen Artus assays. Linear regression analysis and Bland–Altman plots are shown in Figure 2, A and B, respectively. The analysis of 20 BKV (17 primary and 3 co-infections detected) samples resulted in a mean difference of $0.17 \log_{10}$ cp/mL (SD, $\pm 0.67 \log_{10}$ cp/mL) between the Galileo and Qiagen Artus assay. Linear regression analysis and Bland–Altman plots are shown in Figure 3, A and B, respectively. An analysis was performed with 68 samples (47 primary samples and 21 additional virus targets detected and quantified by Galileo and confirmed by respective qPCR assays) (CMV-29, BKV-20, HSV-1-1, JCV-5, HHV-6B-1, VZV-2, EBV-9, and parvovirus B19-1). The linear regression and Bland–Altman plots are shown in Figure 4, A and B, respectively. The \log_{10} difference for the one HSV-1 sample was $0.91 \log_{10}$ copies/mL.

Finally, the viral signal score obtained for TTV with the Galileo assay was compared against viral loads obtained with the bioMérieux TTV qPCR assay (Table 2, Figure 5). A direct comparison was not possible between signal score and the viral load because of the lack of whole virus quantified calibration material; however, a proportional trend was observed when signal score versus \log_{10} cp/mL was plotted as shown in Figure 5. The results of all of the 29 samples detected by Galileo agreed with the qPCR assay. In addition, the qPCR assay detected TTV in 16 more samples, which did not give a signal score with the Galileo assay. However, these samples had lower viral titers, with a mean viral titer of $3.16 \log_{10}$ cp/mL (range,

1.34 to $5.14 \log_{10}$ cp/mL) compared with those samples with a signal score, with a mean viral titer of $6.45 \log_{10}$ cp/mL (range, 3.81 to $8.26 \log_{10}$ cp/mL). Because the Galileo assay generates sequencing libraries from dsDNA, a poorer sensitivity is expected with single-stranded DNA viruses such as TTV.

Discussion

mNGS allows for the detection of a diversity of known or novel pathogens because of its agnostic approach. This study evaluated the performance and diagnostic utility of the Galileo assay in comparison with qPCR assays and found that the assay is capable of reliably detecting and quantifying an array of transplant-related viruses directly from a single sample. Viral titer quantification, a feature unique to this mNGS assay, quantifies viruses using a standard curve generated from a prior calibration run, which includes 11 viruses of interest in transplant diagnostics. A 100% sensitivity was obtained for detection of all viruses in the calibration analysis except for parvovirus B19. Therefore, B19 was removed from the final commercial calibration panel after this beta study because of its ssDNA genome and lower sensitivity with a dsDNA library preparation workflow. Both B19 and TTV are qualitative because of this reason.

In a separate study that investigated the analytical and clinical performance characteristics of the Galileo assay, it was observed that the assay had qualitative and quantitative performances comparable with qPCR single-target assays

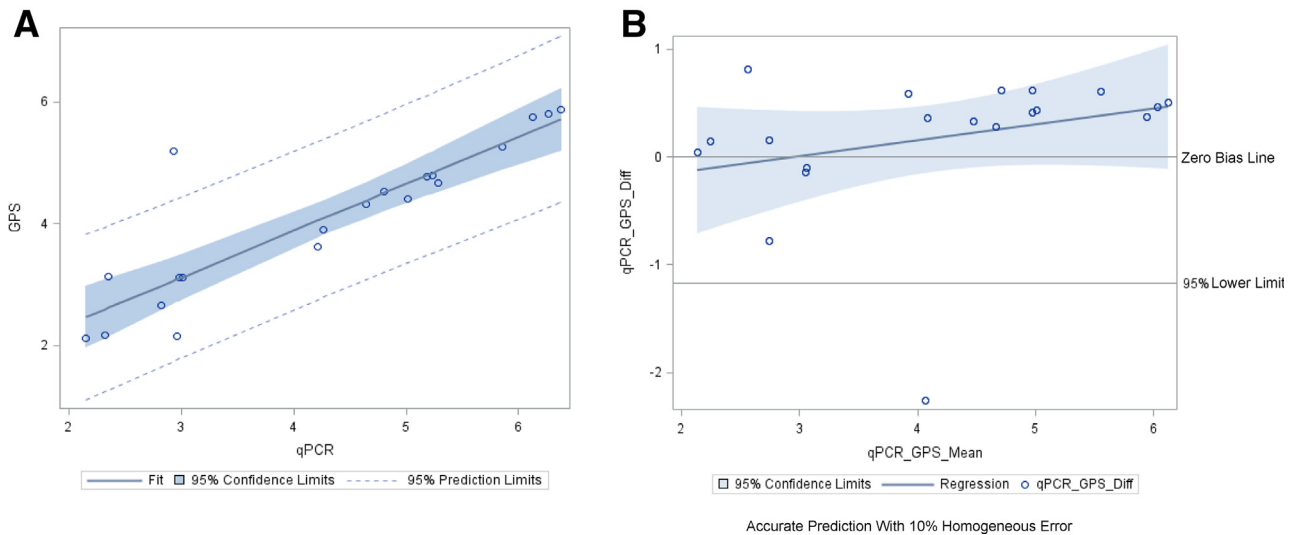


Figure 3 \log_{10} -transformed quantitative viral load agreement of BK virus (BKV) determined by both the Galileo Pathogen Solution (GPS) and real-time quantitative PCR (qPCR) assays. The linear regression line was calculated as follows: $y = 0.77x + 0.81$ and $R^2 = 0.78$. The Bland–Altman plot showed a mean difference of $0.17 \log_{10}$ copies/mL. **A:** BKV linear regression. **B:** BKV Bland–Altman plot. Diff, difference.

with 10 dsDNA transplant-related viruses included in the study. The 95% limit of detection ranged from 14 copies/mL (HHV-6) to 191 copies/mL (BKV), and the lower limit of quantitation ranged from 442 IU/mL (EBV) to 661 copies/mL (VZV).²³ The study also reported the presence of co-infections in clinical samples analyzed, with the exception of VZV. In this study, co-infections were detected in 31 of 47 samples tested and included all virus targets that the Galileo assay is capable of detecting and/or quantitating.

Although this study showed that mNGS has the potential to revolutionize microbial diagnostics with high coverage for species-level detection and the ability to generate quantitative data from multiple pathogens within a single sample, it is labor intensive and time consuming compared with standard-of-care qPCR assays. The turnaround time for the Galileo assay with a high-output kit is approximately 48 hours from DNA extraction to final reports using Galileo Analytics, whereas the turnaround time for a standard qPCR assay is approximately 3 to 5 hours to obtain final results depending on the assay. Although standard qPCR assays provide a more rapid turnaround time, they require knowledge of the target to be identified and are not capable of the level of multiplexing needed for this application. The diagnostic utility of the Galileo assay lies in its ability to detect and simultaneously quantify multiple viruses from a sample with comparable sensitivity, and turnaround time can be reduced by approximately 8 hours by using single-end sequencing. Unlike other mNGS assays, the Galileo assay alleviates the need for computational and

bioinformatics expertise because the analyses are performed using the Galileo Analytics software, which is relatively fast and adaptable. In addition, because the Galileo test is fully deployable into local laboratories, the time and resources needed to develop such an mNGS pipeline is removed.

One major limitation of the mNGS assay in general is that the human host DNA background dominates compared with the pathogen DNA in most patient samples.²⁴ In this study, the nonhuman reads accounted for approximately 4.2% of total paired-end reads on average. This is significant because the libraries were enriched for viral DNA by reducing human DNA background during the assay library preparation and computational human host subtraction performed during the bioinformatics analysis stages. Without the former method the cost of sequencing would be higher, with fewer samples able to be pooled and sequenced in one batch. Without the latter method the analysis times would be much longer.

A low specificity with closely related viruses was observed in the study. The most affected were BKV and JCV owing to both viruses sharing 75% homology of their genomes.²⁵ High-positive BKV samples showed low JCV signal owing to the highly conserved regions present in the two viruses (homologous mapping). The reference sequences used for BKV have a 92% overlap with JCV, with a percent identity of 77.96% (the difference representing small single-nucleotide polymorphisms and INDELs). There are enough divergent regions to enable differentiation between the two viruses, however, the beta version of

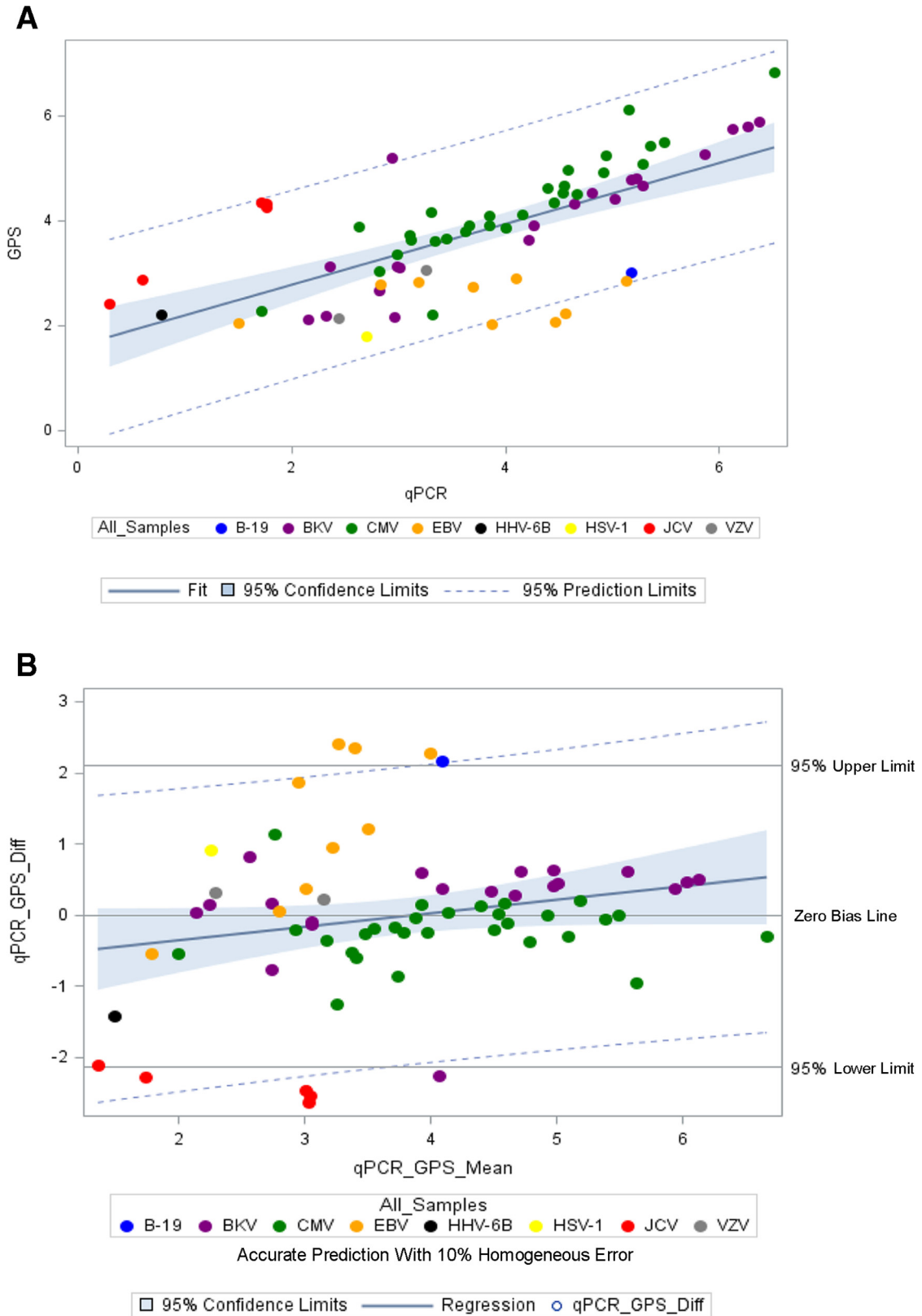


Figure 4 Log₁₀-transformed quantitative viral load agreement for all samples determined by both the Galileo Pathogen Solution (GPS) and real-time quantitative PCR (qPCR) assays. The linear regression line was calculated as follows: $y = 0.58x + 1.60$ and $R^2 = 0.46$. The Bland–Altman plot showed a mean difference of 0.91 log₁₀ copies/mL or IU/mL. **A:** Linear regression plot for all samples. **B:** Bland–Altman plot for all samples. B19, parvovirus B19; BKV, BK virus; CMV, cytomegalovirus; EBV, Epstein–Barr virus; HHV, human herpesvirus; HSV, human herpesvirus; JCV, JC virus; VZV, varicella zoster virus; Diff, difference.

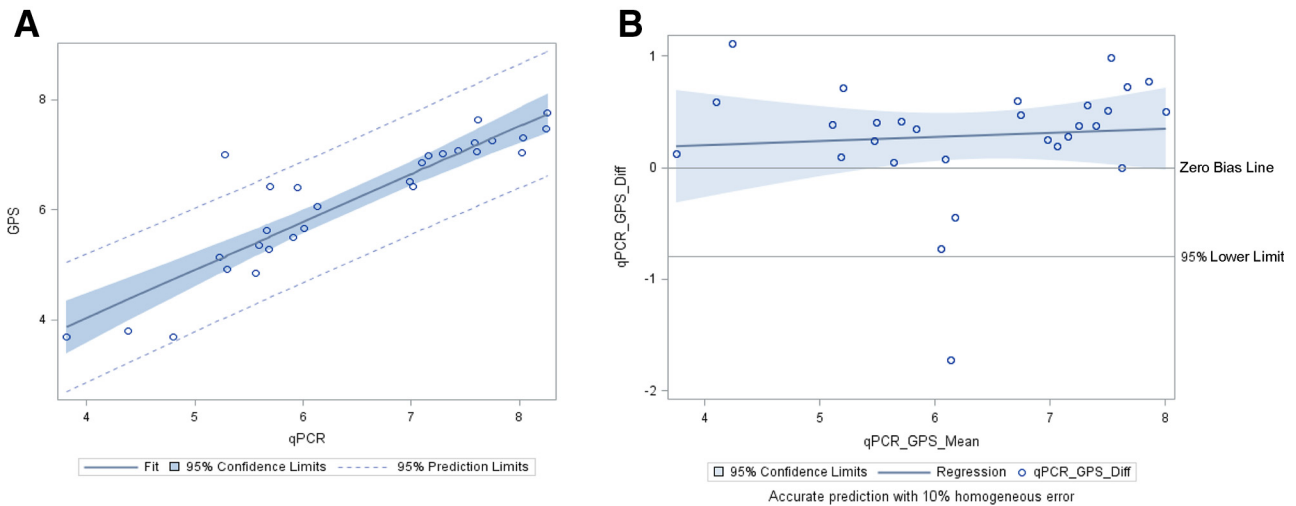


Figure 5 Log_{10} -transformed viral signal score by Galileo Pathogen Solution (GPS) and viral load measurements determined by real-time quantitative PCR (qPCR) assays. The linear regression line was calculated as follows: $y = 0.87x + 0.55$ and $R^2 = 0.82$. **A:** Torque teno virus (TTV) linear regression plot. **B:** TTV Bland–Altman plot. Diff, difference.

Galileo Analytics used in the study did not support this and the observed signal for all sequences was reported. When a secondary analysis was performed by the Arc Bio, LLC, team, the majority of the JCV-positive samples were reverted to negative when only unique regions were taken into account. Future versions will include confidence/reliability scores to help users determine the true-positive result.

The discrepancy of viral titers seen with the retrospective EBV samples was investigated further to determine the fundamental cause. The seven samples that deviated from the correlation metrics were approximately 5 years old and were stored at -80°C , although the CMV titers in the same samples showed excellent correlation. The initial troubleshooting with the known Galileo assay full-process controls using the Qiagen Artus qPCR assay provided comparable viral titers. The further analysis of EBV FASTQ files with third-party BioX analysis tools [Kraken (kmer based), Diamond (protein based), and BWA (alignment based)] determined that the problem was not related to analysis of the sequencing data. This raised the possibility that EBV DNA degraded differently than CMV DNA. The improved agreement between the qPCR results and the Galileo assay with the samples that were stored for fewer than 6 months supported this conclusion.

It was speculated that the EBV DNA in the older samples somehow had become degraded in a way that adversely affected library sequencing construction. The EBV DNA regions may have become single-stranded, which would not affect the qPCR result but would affect library sequencing

construction because dsDNA molecules with no internal nicks are required. In addition, the end-repair step that uses 5' exonuclease activity retains mainly dsDNA.²⁶ EBV, unlike CMV, can form episomes during latency that, because of their circular nature, could be more resistant to complete degradation. Thus, in old samples the circular EBV might lose only one strand of DNA to exonuclease activity, which would preserve most of the qPCR signal while destroying most of the library sequencing signal. In contrast, because CMV is never circular during its lifecycle, if it is degraded by exonucleases it would lose both the qPCR and sequencing-based signals and therefore false negatives/underquantification would have been observed with both technologies.

The presence of TTV in 95.7% of the samples included in this study were identified using qPCR. Those samples with a Galileo signal score also were qPCR positive, but the samples that had a lower qPCR viral titer were not detected by the Galileo assay. TTV is reported to be highly prevalent in humans, particularly with low-level viremia in 90% of healthy carriers, and therefore may have utility as a potential immune marker.^{13,27}

The study observed positive signals for HAdV and BKV in negative process controls in a couple of runs. When analyzed with the standard curve, these titers were below the limit of detections for the respective viruses. Because the negative controls are made up of pooled human plasma, these viruses are likely to be present at low backgrounds in human plasma pools. This is one of

the challenges with mNGS technologies owing to the presence of latent viruses or commensal, environmental, and reagent microbial populations in a sequencing library, which can complicate the analysis and interpretation of results. Therefore, a change in terminology was proposed from negative to background control because of the inherent contamination expected because of the unbiased nature of mNGS.

The potential advantages of mNGS assays include simultaneous detection and quantification from a single blood draw, the ability to use the metagenomics data for antiviral resistance determination, and the presence of more than 350 strains in the bioinformatics pipeline, which may circumvent challenges seen with qPCR and diverse viruses such as HAAdV. However, there is a need for personnel with a skillset in clinical laboratories to successfully implement mNGS diagnostics. An important limitation of this assay design is that detection of all of these viruses, although warranted in some clinical situations, is not necessary on a routine basis for most patients.

In conclusion, the study findings show that the novel Galileo assay was able to simultaneously detect and quantify multiple viruses from plasma samples of transplant recipients with results that are comparable with standard-of-care qPCR assays, both Food and Drug Administration approved and ASRs. Future studies are required to evaluate the test on a larger number of clinical samples to better assess its use as a diagnostic tool for transplant recipients.

Acknowledgment

We thank the Viracor Eurofins Reference Laboratory (Lee's Summit, MO) for assistance with varicella zoster virus, parvovirus B19, and herpes simplex viruses 1 and 2 quantitative testing.

References

- Kumar R, Ison MG: Opportunistic infections in transplant patients. *Infect Dis Clin North Am* 2019, 33:1143–1157
- Camargo JF, Komanduri KV: Emerging concepts in cytomegalovirus infection following hematopoietic stem cell transplantation. *Hematol Oncol Stem Cell Ther* 2017, 10:233–238
- Haidar G, Singh N: Viral infections in solid organ transplant recipients: novel updates and a review of the classics. *Curr Opin Infect Dis* 2017, 30:579–588
- Lau JSY, Low ZM, Abbott I, Shochet L, Kanellis J, Kitching AR, Korman TM: Epstein-Barr virus encephalitis in solid organ transplantation. *New Microbiol* 2017, 40:212–217
- Hirsch HH, Randhawa P; AST Infectious Diseases Community of Practice: BK polyomavirus in solid organ transplantation. *Am J Transplant* 2013, 13(Suppl 4):179–188
- Imlay H, Xie H, Leisenring WM, Duke ER, Kimball LE, Huang M, Pergam SA, Hill JA, Jerome KR, Milano F, Nichols WG, Pang PS, Hirsch HH, Limaye AP, Boeckh M: Presentation of BK polyomavirus-associated hemorrhagic cystitis after allogeneic hematopoietic cell transplantation. *Blood Adv* 2020, 4:617–628
- Shiley K, Blumberg E: Herpes viruses in transplant recipients: HSV, VZV, human herpes viruses, and EBV. *Infect Dis Clin North Am* 2010, 24:373–393
- Obeid KM: Infections with DNA viruses, adenovirus, polyomaviruses, and parvovirus B19 in hematopoietic stem cell transplant recipients and patients with hematologic malignancies. *Infect Dis Clin North Am* 2019, 33:501–521
- White DW, Beard RS, Barton ES: Immune modulation during latent herpesvirus infection. *Immunol Rev* 2012, 245:189–208
- Rezahosseini O, Drabe CH, Sorensen SS, Rasmussen A, Perch M, Ostrowski SR, Nielsen SD: Torque-teno virus viral load as a potential endogenous marker of immune function in solid organ transplantation. *Transpl Rev (Orlando)* 2019, 33:137–144
- Focosi D, Antonelli G, Pistello M, Maggi F: Torquetenovirus: the human virome from bench to bedside. *Clin Microbiol Infect* 2016, 22:589–593
- Gilles R, Herling M, Holtick U, Heger E, Awerkiew S, Fish I, Holler K, Sierra S, Knops E, Kaiser R, Scheid C, Cristanziano VD: Dynamics of torque teno virus viremia could predict risk of complications after allogeneic hematopoietic stem cell transplantation. *Med Microbiol Immunol* 2017, 206:355–362
- Strassl R, Schiemann M, Doberer K, Görzer I, Puchhammer-Stöckl E, Eskandary F, Kikic Z, Gualdoni GA, Vossen MG, Rasoul-Rockenschaub S, Herkner H, Bohmig GA, Bond G: Quantification of torque Teno virus viremia as a prospective biomarker for infectious disease in kidney allograft recipients. *J Infect Dis* 2018, 218:1191–1199
- Gu W, Miller S, Chiu CY: Clinical metagenomic next-generation sequencing for pathogen detection. *Annu Rev Pathol* 2019, 14:319–338
- Chiang AD, Dekker JP: From the pipeline to the bedside: advances and challenges in clinical metagenomics. *J Infect Dis* 2020, 221:(Supplement 3):S331–S340
- Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, Stryke D, Pham E, Fung B, Bolosky WJ, Ingebrigtsen D, Lorzio W, Paff SM, Leake JA, Pesano R, DeBiasi R, Dominguez S, Chiu CY: Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res* 2019, 29:831–842
- Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, Bourcy Cde, Egger R, Dimitrov B, Juan Y, Sheu J, Wang J, Kalantar K, Langelier C, Ruel T, Mpimbaza A, Wilson MR, Rosenthal PJ, DeRisi JL: Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS One* 2019, 14:e0218318
- Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, Kawli T, Christians FC, Venkatasubrahmanyam S, Wall GD, Cheung A, Rogers ZN, Meshulam-Simon G, Huijse L, Balakrishnan S, Quinn JV, Holleman D, Hong DK, Vaughn ML, Kertesz M, Bercovici S, Wilber JC, Yang S: Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat Microbiol* 2019, 4:663–674
- Bland JM, Altman DG: Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999, 8:135–160
- Wood DE, Salzberg SL: Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014, 15:R46
- Buchfink B, Xie C, Huson DH: Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015, 12:59–60
- Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009, 25:1754–1760
- Carpenter ML, Tan SK, Watson T, Bacher R, Nagesh V, Watts A, Bentley G, Weber J, Huang C, Sahoo MK, Hinterwirth A, Doan T,

- Carter T, Dong Q, Gourguechon S, Harness E, Kermes S, Radhakrishnan S, Wang G, Quiroz-Zárate A, Ching J, Pinsky BA: Metagenomic next-generation sequencing for identification and quantitation of transplant-related DNA viruses. *J Clin Microbiol* 2019, 57: e01113–e01119
24. Thoendel M, Jeraldo PR, Greenwood-Quaintance KE, Yao JZ, Chia N, Hanssen AD, Abdel MP, Patel R: Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J Microbiol Methods* 2016, 127:141–145
25. Boothpur R, Brennan DC: Human polyoma viruses and disease with emphasis on clinical BK and JC. *J Clin Virol* 2010, 47:306–312
26. Tin MM, Economo EP, Mikheyev AS: Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. *PLoS One* 2014, 9:e96793
27. Haloschan M, Bettesch R, Gorzer I, Weseslindtner L, Kundi M, Puchhammer-Stockl E: TTV DNA plasma load and its association with age, gender, and HCMV IgG serostatus in healthy adults. *Age (Dordr)* 2014, 36:9716