Journal of Clinical Virology Plus 2 (2022) 100073

Contents lists available at ScienceDirect

Journal of Clinical Virology Plus

journal homepage: www.elsevier.com/locate/jcvp

Use of a sample-to-result shotgun metagenomics platform for the detection and quantification of viral pathogens in paediatric immunocompromised patients

Divya Shah[a], Julianne R. Brown[a], Jack C.D. Lee[a], Meredith L. Carpenter[b,*], Gavin Wall[c], Judith Breuer[a,d]

[a] Microbiology, Virology and Infection Prevention and Control, Great Ormond Street Hospital for Children NHS Foundation Trust, UK
[b] Arc Bio, LLC, Scotts Valley, CA and Cambridge, MA, USA
[c] QIAGEN, Hilden, Germany
[d] Division of Infection and Immunity, University College London, UK

ARTICLE INFO

Keywords:
Metagenomics
Transplant viruses
Next-generation sequencing
Paediatric viral infection
qPCR

ABSTRACT

Background: Infections by several DNA viruses can severely impact outcomes in paediatric immunocompromised patients. Current testing, which is generally limited to singleplex qPCR assays, can miss both common and rarer viruses if they are not targeted.
Objectives: To evaluate the performance of the Galileo Viral Panel (Galileo), a sample-to-result shotgun metagenomics platform for the detection and quantification of 12 DNA viruses, compared to standard of care qPCR assays.
Study design: A clinical performance evaluation was carried out using 43 prospectively collected EDTA plasma samples positive for one or more DNA viruses. Agreement between assays was assessed by overall, positive, and negative percent agreement, as well as quantitative agreement by linear regression and Bland-Altman analysis.
Results: Overall positive percent agreement was 84% (95% CI: 76%-90%), and negative percent agreement was 95% (95% CI: 92%-97%). There was a high correlation between Galileo and qPCR for ADV, CMV, EBV, and VZV (R^2 = 0.91) and a mean difference by Bland Altman of -0.43 $\log_{10}$ IU or cp/ml (95% limits of agreement, -1.37 to 0.51). In addition, there was a high correlation between Galileo Signal Score and qPCR for TTV (R^2 = 0.85).
Conclusion: We observed high qualitative and quantitative agreement between qPCR and Galileo. Galileo identified additional viruses that were not tested with routine qPCR and could impact clinical outcomes.

1. Introduction

Infections by DNA viruses are a common cause of morbidity and mortality in both adult and paediatric immunocompromised transplant recipients [1,2]. However, children are at an increased risk of some complications of these viruses due to their immune status and history of infection [2]. For example, children are more likely than adults to be Epstein-Barr virus (EBV) seronegative before transplant, putting them at an increased risk of developing post-transplant lymphoproliferative disorders [2]. Primary CMV infections are also seen more frequently in immunocompromised paediatric patient, and CMV viral load has been shown to predict CMV disease and death [3–6]. BK virus (BKV), varicella zoster virus (VZV), and adenovirus (ADV) are also of clinical importance, particularly in paediatric bone marrow transplant patients [2].

Standard of care testing for these viruses relies on an array of singleplex quantitative PCR (qPCR) assays. Use of these targeted assays can miss less-common viruses, as well as putative biomarkers of immune status such as torque teno virus (TTV) [7] . In contrast to targeted PCR assays, metagenomic next-generation sequencing (mNGS) is a hypothesis-free testing approach that can be used to identify and quantify a potentially unlimited array of microbes. The Galileo Viral Panel (Galileo) (Arc Bio, LLC, Scotts Valley, CA) is a Research Use Only sample-to-result mNGS platform designed to simultaneously detect and quantify 10 DNA viruses [herpes simplex virus 1 and 2 (HSV1/2), human herpesvirus 6A and B (HHV-6A/B), VZV, CMV, EBV, JC virus (JCV), BKV, and ADV] and to semi-quantitatively detect Parvovirus B19 (B19) and TTV [8,9]. Galileo includes all reagents, controls, and software required to perform mNGS on plasma samples in under 48 h. The objective of this study was to evaluate the clinical performance characteristics of a pre-commercial version of the Galileo platform compared with standard-of-care qPCR

* Corresponding author.
E-mail address: meredith@arcbio.com (M.L. Carpenter).

https://doi.org/10.1016/j.jcvp.2022.100073
Received 8 December 2021; Received in revised form 11 February 2022; Accepted 17 March 2022
2667-0380/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Table 1
Performance characteristics of PCR assays used in this study.

PCR Target	Limit of Detection	Linear Range (for quantitative assays)
VZV	2000 cp/ml	2000–20 M cp/ml
CMV	50 IU/ml or 200 cp/ml	200–20 M cp/ml
EBV	18 IU/ml or 200 cp/ml	200–20 M cp/ml
HHV-6	1000 cp/ml	N/A
JCV	100 cp/ml	N/A
BKV	100 cp/ml	N/A
hADV	200 cp/ml	200–20 M cp/ml

assays using prospectively collected residual plasma from viremic paediatric immunocompromised patients.

2. Materials and methods

2.1. Patient consent statement

This study utilized existing ethics approval for use of residual, blinded samples for method evaluation (National Research Ethics Service (NRES) Committee London – Fulham (REC reference: 17/LO/1530)).

2.2. Clinical samples

Samples chosen for evaluation included those from immunocompromised patients that tested positive by qPCR in whole blood for at least one or more of ten transplant-related viruses used in routine clinical testing (CMV, BKV, EBV, ADV, VZV, HHV-6, HSV1/2, and JCV). A total of 50 whole blood samples were refrigerated for up to 3 weeks before separation of EDTA plasma for this study. The whole blood samples were refrigerated to ensure availability of samples for additional routine clinical testing. After exclusion of samples without detectable viral DNA in plasma (due to differences in viral detection in whole blood vs. plasma compartments) or sufficient residual volume for testing, a total of 43 samples from 26 patients were selected for further processing with the Galileo workflow. All patients were between the ages of 8 months – 14 years with the exception of one (62 years), and the median age was 6 years.

2.3. DNA extraction, library preparation, and sequencing

DNA was extracted from 400 μL samples/Galileo controls using the EZ1 Virus Mini Kit v2 (Qiagen, Ltd.) and eluted into 90 μL. Galileo internal controls were added to all samples before extraction. Negative and positive full-process controls (Arc Bio, LLC) were processed alongside each run of 10 clinical samples. Sequencing libraries were prepared according to the manufacturer's protocols (Arc Bio, LLC) and as described previously [9]. Libraries were sequenced on the NextSeq 500 sequencer (Illumina) using the paired-end, high-output 150-cycle v2.5 kit. Before clinical sample testing, an initial calibration run was performed using a multianalyte panel of whole-virus controls spiked into plasma (Arc Bio, LLC) to generate a standard curve and estimate the viral load as previously described [8,9].

2.4. PCR assays

Initial qPCR testing of plasma samples was performed using validated laboratory-developed tests [10–13] on the ABI 7500 (ThermoFisher). The limits of detection and linear ranges of these assays are shown in Table 1. TTV was tested using a commercially available qPCR kit (RUO, TTV-R-GENE, Biomerieux).

Table 2
Demographic characteristics of the 26 patients included in the study.

Detail	Number (%/range)
Median Age	6 years (8 months–62 years)
Sample Date	07/11/18 – 01/04/19
Male	13 (50%)
Female	13 (50%)
Underlying Condition	
Post-HSCT	12 (46.2%)
Malignancy	5/12
Congenital immune disorders	4/12
Congenital metabolic disorders	3/12
Post-SOT	2 (7.7%)
No transplant	10 (38.5%)
Congenital immune disorders	6/10
Congenital Metabolic disorder	1/10
Malignancy	1/10
General medical admission	2/10
Unknown	2 (7.7%)

2.5. Bioinformatics analysis

System-level NextSeq quality metrics were evaluated according to the manufacturer's recommendations (Illumina, Inc.). The sample sheet was downloaded from Galileo Analytics (Arc Bio, LLC), and demultiplexing 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. QC criteria were assessed as recommended by the manufacturer and as previously described [8,9].

2.6. Statistical analysis

Statistical analysis (including Bland-Altman analysis) and graphing was performed in R version 4.0.4 [14], using the dplyr (v1.0.6) [15] and ggplot2 (v3.3.3) [16] packages.

3. Results

Demographic characteristics of the patient cohort included in this study are shown in Table 2. An average of 50 M total reads (range 25,599,402–98,222,454) were sequenced per library. Of the 43 libraries, one was negative for one internal control, and two runs had low-level viral DNA detected in the negative controls (pooled human plasma).

From the 43 plasma samples, 56 viruses were initially detected by routine qPCR. Galileo identified an additional 51 viruses for 107 viruses in total (Table 3, Supplemental Table 1). Any additional viruses detected by Galileo were then assessed by plasma qPCR. Positive percent agreement was 84% (95% CI: 76–90%), and negative percent agreement was 95% (95% CI: 92–97%). There were 11 viruses detected by Galileo only and 16 viruses detected by qPCR only. The viruses detected by Galileo only were 1 CMV (154 IU/ml), 1 EBV (477 IU/ml), 2 HHV-6 (333 and 2153 cp/ml), and 7 JCV (20–13,372 IU/ml); all JCV samples were BKV positive, so these were likely false positives due to genome homology between BKV and JCV, as previously reported [8,9]. The viruses detected by qPCR only were 5 EBV (1689–12,545 cp/ml), 3 CMV (1263 - 2054 cp/ml), 1 JCV (Ct 37), 1 HHV-6 (Ct 38), and 6 TTV (2000–248,000 cp/ml); all had Ct values 35–39, except TTV with Ct values 24–41. Galileo differentiates between HHV-6A and -6B; however, because the qPCR assay does not, these results were grouped together as HHV-6. See Supplemental Table 1 for the full qPCR and Galileo results for each patient.

For those viruses with quantatitive PCR data, 40 samples with quantitative results by both methods were compared (Fig. 1). There was high quantitative correlation between the assays for all viruses (R^2 = 0.91) (Fig. 1A). Quantitative agreement by Bland-Altman was –0.43 $\log_{10}$

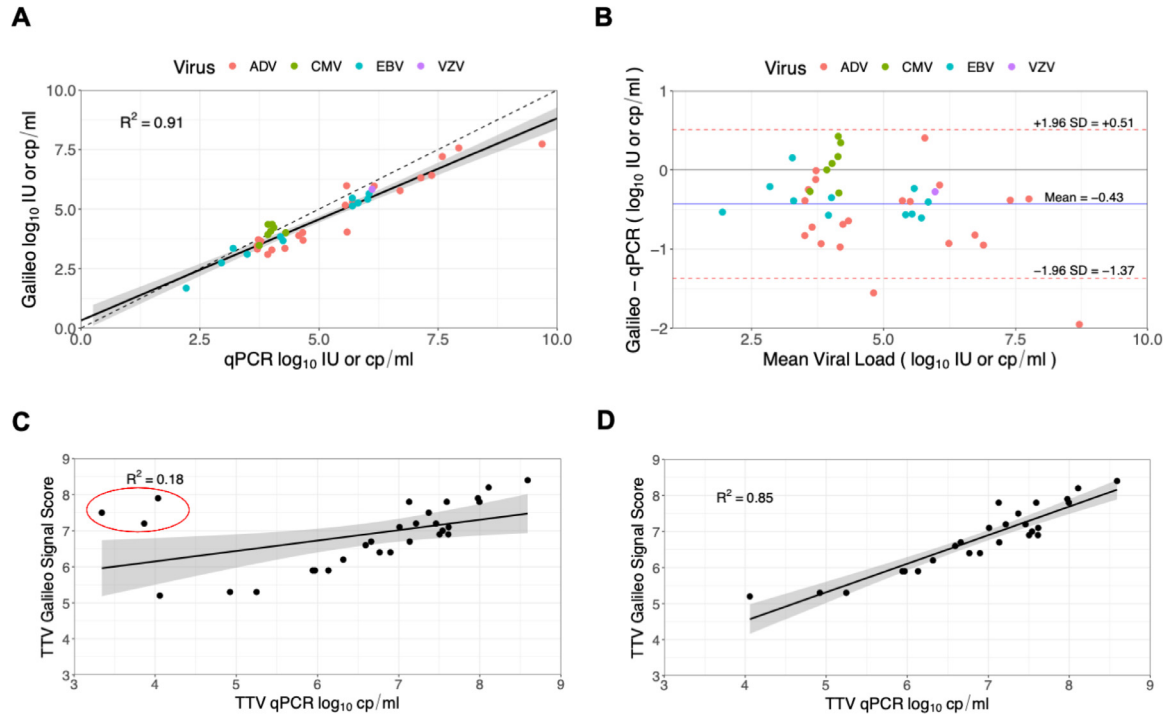

Fig. 1. Quantitative agreement between Galileo mNGS and qPCR. (A) Linear regression of qPCR titer in $\log_{10}$ IU or cp/ml on the *X*-axis and Galileo mNGS $\log_{10}$ IU or cp/ml on the *Y*-axis for the four viruses where qPCR results were available, and where viruses were detected by both assays. The regression line (solid line), line of identity (dotted line), and 95% confidence intervals (gray shaded areas) are shown. (B) Quantitative agreement by Bland-Altman plot with mean (blue line), 95% limits of agreement (dashed lines), and zero line (gray line) shown. (C) Linear regression of qPCR vs Galileo Signal Score for TTV, where whole-virus calibration material is not available to convert GSS to cp/ml. Three outliers from the same patient are indicated (red circle). 95% confidence intervals are shown in gray. (D) Linear regression of qPCR titer and Galileo Signal Score with three outliers removed. 95% confidence intervals are shown in gray (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 3
Summary of viral detection by Galileo mNGS and qPCR across the sample set. Positive percent agreement (PPA) and negative percent agreement (NPA) are also shown.

		qPCR		PPA	NPA
		+	–		
Galileo	**All viruses (+)**	87	11	84%	95%
	All viruses (-)	16	230		
	HHV-6 (+)	8	2	89%	94%
	HHV-6 (–)	1	32		
	BKV (+)	9	0	100%	100%
	BKV (–)	0	34		
	JCV (+)	1	7	50%	83%
	JCV (–)	1	34		
	VZV (+)	1	0	100%	100%
	VZV (–)	0	42		
	ADV (+)	21	0	100%	100%
	ADV (–)	0	22		
	CMV (+)	7	1	70%	97%
	CMV (–)	3	32		
	EBV (+)	11	1	69%	96%
	EBV (–)	5	26		
	TTV (+)	29	0	83%	100%
	TTV (-)	6	8		

IU or cp/ml mean difference (95% limits of agreement, -1.37 to 0.51) (Fig. 1B). Focusing on TTV (which was quantified by Galileo's proprietary Signal Score metric due to the lack of a commercially available whole-virus calibration material for TTV), results from the 29 samples from 15 patients positive by both assays showed high correlation between Galileo and qPCR, except for three outliers, all from the same patient, that showed lower quantification by qPCR (Fig. 1C). When these were removed, the linear regression R^2 was 0.85 (Fig. 1D).

4. Discussion

This study demonstrates overall high levels of agreement between single-target qPCR assays and the Galileo Viral Panel shotgun mNGS platform. While both assays missed several low titer/late Ct samples, Galileo detected an additional 51 viruses that had not initially been tested for by standard of care PCR, all of which would be important if encountered during patient care at clinically significant levels. While the low-level signals in both assays may represent noise (as is likely the case for the detection of JCV in samples positive for BKV), they may also indicate early viral replication or latent viral genomes, which are both of uncertain clinical significance. Quantitatively, Galileo slightly underquantified compared with qPCR assays, suggesting that only one assay should be used to monitor a patient.

Interestingly, three samples from the same patient showed lower quantification of TTV by qPCR compared with Galileo. This discrepancy may be due to sequence differences in the primer binding region related to the high genetic variation in TTV [17]. In contrast, Galileo can utilize the entire TTV genome to detect and quantify.

Although Galileo can detect many viruses simultaneously, including viruses that might be missed in standard care singleplex qPCR testing, qPCR remains less costly, less laborious, and faster than mNGS. The

Galileo workflow takes approximately 48 h (of which 20 h is sequencing), while qPCR requires 4 to 6 h, including extraction, reaction setup, PCR, and analysis. However, the expansion of the Galileo assay to report additional pathogens (e.g., bacteria and fungi), coupled with reductions in sequencing time, could increase the utility and diagnostic yield of mNGS, making it more amenable to routine clinical use.

Limitations of this study include the low numbers of certain viruses (1 VZV sample, no HSV1/2 or B19 included); extended sample storage time prior to plasma analysis; low total number of samples; and its retrospective nature for the purposes of method comparison rather than analysis of clinical outcomes. In conclusion, the Galileo mNGS platform is a promising method for comprehensive viral monitoring in immunocompromised paediatric patients.

Declaration of Competing Interest

MLC and GW are current or former employees of Arc Bio, LLC. JB is a clinical advisor for Arc Bio.

Acknowledgments

The authors would like to thank Jannine Forst for assistance with graph generation. Arc Bio provided a precommercial version of Galileo Viral Panel and access to the Galileo Analytics software used in this study. JB receives funding from the NIHR UCL/UCLH Biomedical Research Centre. JL is funded by an NIHR award NIHR200652 to JB.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jcvp.2022.100073.

References

[1] J.A. Fishman, Infection in organ transplantation, Am. J. Transplant. 17 (2017) 856–879.
[2] M.G. Michaels, M. Green, Infections in pediatric transplant recipients: not just small adults, Hematol. Oncol. Clin. N. Am. 25 (2011) 139–150.
[3] A. Stern, et al., CMV viral load kinetics predict CMV end-organ disease and mortality after hematopoietic cell transplant (HCT), J. Infect. Dis. 224 (2021) 620–631.
[4] W. Lu, H.P. Chen, Y.J. Chan, F.D. Wang, Clinical significance of post-treatment viral load of cytomegalovirus in patients with hematologic malignancies, J. Microbiol. Immunol. Infect. 54 (2021) 245–252.
[5] E.R. Duke, et al., CMV viral load kinetics as surrogate endpoints after allogeneic transplantation, J. Clin. Invest. 131 (1) (2021) e133960–e133974.
[6] S.B. Vora, J.A. Englund, Cytomegalovirus in immunocompromised children, Curr. Opin. Infect. Dis. 28 (2015) 323–329.
[7] W. Mouton, et al., Torque teno virus viral load as a marker of immune function in allogeneic haematopoietic stem cell transplantation recipients, Viruses 12 (11) (2020) 1292–1302.
[8] M.L. Carpenter, et al., Metagenomic next-generation sequencing for identification and quantitation of transplant-related DNA viruses, J. Clin. Microbiol. 57 (12) (2019) e01113–e01119.
[9] S.S. Sam, et al., Evaluation of a next-generation sequencing metagenomics assay to detect and quantify DNA viruses in plasma from transplant recipients, J. Mol. Diagn. 23 (2021) 719–731.
[10] M.L. Quinlivan, et al., Effect of viral load on the outcome of herpes zoster, J. Clin. Microbiol. 45 (2007) 3909–3914.
[11] W.J. Jabs, et al., Normalized quantification by real-time PCR of epstein-barr virus load in patients at risk for posttransplant lymphoproliferative disorders, J. Clin. Microbiol. 39 (2001) 564–569.
[12] A. Heim, C. Ebnet, G. Harste, P. Pring-Akerblom, Rapid and quantitative detection of human adenovirus DNA by real-time PCR, J. Med. Virol. 70 (2003) 228–239.
[13] T.J. Dumonceaux, C. Mesa, A. Severini, Internally controlled triplex quantitative PCR assay for human polyomaviruses JC and BK, J. Clin. Microbiol. 46 (2008) 2829–2836.
[14] R. Team, R: a language and environment for statistical computing. http://www.r-project.org (2020).
[15] R.F. Hadley Wickham, L. Henry, K. Müller, dplyr: a grammar of data manipulation. https://dplyr.tidyverse.org/(2021).
[16] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2016.
[17] S. Hino, H. Miyata, Torque teno virus (TTV): current status, Rev. Med. Virol. 17 (2007) 45–57.