Cytomegalovirus Standard Reference Material Development for Nucleic Acid Amplification Technologies

Ross J. Haynes, Marcia J. Holden, Margaret C. Kline, and John M. Butler

Ross.Haynes@nist.gov
301-975-4469

NIST
National Institute of Standards and Technology
Technology Administration
U.S. Department of Commerce

ID – 20

Cytomegalovirus Reference Material

•**Disease:** Cytomegalovirus (CMV) causes life-threatening infections in immunocompromised patients and in congenital transmission to infants, though commonly found and usually latent in the general population

•**Standard needs:** Calibration and quality control of quantitative real-time PCR assays of blood and other body fluids

•**Customers:** Producers of secondary standards or clinical laboratory using their own calibrant

•***Materials***: Pure DNA from Towne strain in a bacterial artificial chromosome containing all of the viral genome except for regions IRS1, US1-15, and UL147; viral DNA to be provided in a buffer for dilution into a user's matrix of choice

•***Certification***:

1. DNA sequence of genes that are targets for qPCR
2. Copy number via digital PCR measurements of multiple CMV specific qPCR targets

•***Additional Information***:

- •Restriction digest and microarray validation of BAC
- •Testing of various published PCR assays
- •Homogeneity and stability

Sequencing

•Sequence via Big Dye Terminator v3.1 (Applied Biosystems)

•Analyze by capillary electrophoresis using Applied Biosystems 3130xl with POP7 and an 80cm capillary

•Sequence checked against the Towne strain GenBank accession AY315197

•BAC has the same sequence as Towne strain (AY315197) in regions already sequenced – except for one ambiguous base.

•Regions sequenced: most or all of UL54 (polymerase), UL55-56 (glycoprotein B), UL83 (phosphoprotein 65), UL122-126 (major immediate early area and nearby sequence), and US17

•Regions to be sequenced: UL34 and UL80

Copy Number Determination

•Digital PCR requires multiple replicates at a concentration so that some, but not all, wells have target DNA

•Based on probability: number of wells with target translates to number of copies of target per volume.

•Multiple qPCR assays – spread throughout the genome –will be used to perform digital PCR on the SRM

•Digital PCR was performed using two qPCR assays, IE and CP1, which target the major immediate early region and the polymerase gene respectively

•Slight between group difference is explained by slight difference in amplification efficiency which should be corrected by increasing the number of amplification cycles

Molecules/uL		Anova: Single Factor						
IE	CP1							
4394	3452	SUMMARY						
3631	3537	*Groups*	*Count*	*Sum*	*Average*	*Variance*		
4192	3600	IE	6	24055	4009	71496		
3972	3653	CP1	6	21781	3630	101496		
3833	3308							
4031	4232							
X_c	3820	ANOVA						
SD_c	189	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
RSD_c	5%	Between Groups	430658	1	430658	4.97891	0.04973	4.965
$U_{k=2}$	10%	Within Groups	864965	10	86496			
		Total	1295623	11				

Fluidigm Technology

•Digital PCR platform with integrated fluidic circuit

•Automatically partitions sample into 765 reactions

•Each chip has 12 panels = 12 samples

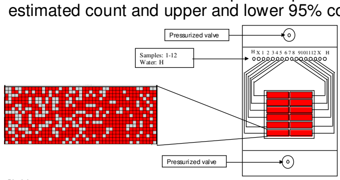
•Each panel uses 4.6 uL of master mix

- •765 partitions x 6 nL/partition = 4.6 uL

•Chips are pressure-filled; therefore internal geometries determine the volume of sample

- –Microscopy (height & width) and interferometry (depth) can be used to measure the reaction volumes

•Software detects and counts positive partitions and calculates estimated count and upper and lower 95% confidence intervals



Abstract

Viral load measurements are important for the treatment of cytomegalovirus (CMV) infections resulting from activated latent or new infections in immune-compromised patients. Viral load testing is primarily accomplished using quantitative real-time PCR (qPCR). Unfortunately, there is a lack of comparability between laboratories when performing qPCR testing for CMV viral load—in part because of the current unavailability of reference materials to enable measurement calibration. Variability among testing laboratories on CMV results has been noted in proficiency testing (PT) programs, such as the Quality Control for Molecular Diagnostics (QCMD) programme. While improvement has been seen over sequential PT programs, many laboratories still fall out of the acceptable range. Contributing to this variability is the fact that a variety of methods are used including both commercial and "in-house" assays.
The Applied Genetics Group at the National Institute of Standards and Technology (NIST) has made progress in the development of a certified reference material for CMV. This NIST Standard Reference Material (SRM) will consist of pure DNA isolated from the Towne strain (produced in bulk from bacterial artificial chromosome production). The goal is to certify this CMV DNA for quantity as well as sequence information across regions where published qPCR assays have been reported, such as UL54 (DNA Polymerase), UL55 (Glycoprotein B), UL83 (Phosphoprotein 65), UL123 (Major Immediate Early), UL32 (Phosphoprotein 32) and the US17 region. As part of the development of this CMV SRM, comparisons have been made of published qPCR primers and probes for various CMV detection assays to look for potential mutations across common strains of CMV. In addition, homogeneity and stability of the CMV DNA samples are being monitored and progress in our CMV SRM development thus far will be presented. Experiments performed with digital PCR to measure the number of copies of CMV DNA will also be described. We believe that the availability of certified reference materials for CMV will enable traceability for manufacturers of calibrant materials that should aid consistency in measurements between laboratories.

Pilot Materials

1) BAC DNA in buffer

2) BAC DNA in buffer with poly-thymine carrier DNA

•Distributed into tubes for stability and homogeneity testing

•Concentration determined with digital PCR

•No difference was found between in the two materials: therefore data was combined.

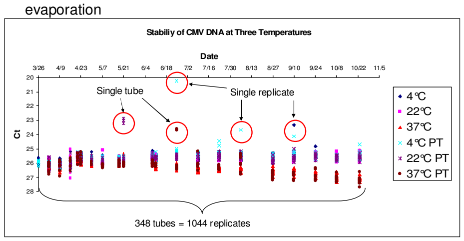
•Materials were stored at three temperatures 4℃, 22℃ and 37℃ as an accelerated aging model

Homogeneity

•2 out of 348 tubes have been discovered with divergent concentrations

•Several single replicates have shown divergent Cts, but have not been repeatable

- –We hypothesize that these wells did not seal well during PCR and the decrease in Ct is due to increase in concentration via evaporation

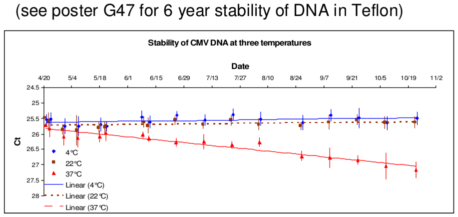


Stability

•Empirical data shows DNA stored at 4℃ and 22℃ appear stable, but signs of degradation appeared after 8 weeks in DNA stored at 37℃

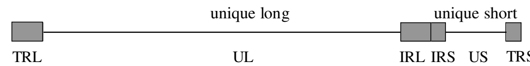
•Accelerated aging: for every 10℃ over room temperature aging is doubled [1]. Considering when the DNA stored at 37℃ showed signs of degradation the DNA stored at 22℃ should have started degrading already.

•SRM will be packaged in Teflon microcentrifuge tubes (see poster G47 for 6 year stability of DNA in Teflon)



Measurements taken between 3/26 and 4/16/09 used ABI 7000. ABI 7500 has been used from 4/21/09 on (only data from ABI 7500 is shown)

Genome Structure



Schematic map of the CMV genome. The CMV genome is organized as two regions of unique sequences, unique long (UL) and unique short (US), flanked by two sets of inverted repeats (TRL/IRL) and (IRS/TRS) (light shaded boxes) [2]

The regions to be sequence and qPCR assays which will be used to determine copy number represent a very small area (~14kb of 240kb)

Clinical DNA Informational Resource

http://www.nist.gov/cstl/biochemical/genetics/clinical_dna.cfm

"Although primer mismatching does not necessarily prevent PCR amplification, use of primers having a known high degree of sequence conservation…should reduce assay variability in clinical practice" (Chou 1992J of Clin Micro 30(9): 2307-2310).

•Primer size and location of mismatch affect amplification efficiency which can effect clinical sensitivity and make defining a clinical threshold for treatment difficult.

•Clinical DNA Informational Resource will contain:

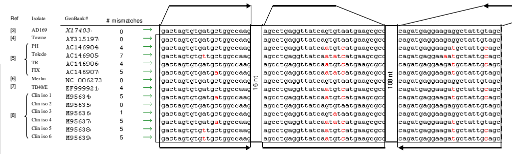
- •Sequence alignments to aid researchers designing primers and probes
- •Primer and probe sequences with references

Sequence Alignments

•The core of our sequence alignments consist of eight CMV strains with complete genome submissions found in Genbank and partial sequences in will be added to alignments when informational value is added

•Files will be available in SeqMan (Lasergene software) and text (.txt)

•Alignments will allow researchers to search for primers to quickly determine degree of sequence conservation



The above qPCR set comes from Tanaka 2000 Journal of Medical Virology 60:455–462, targets the UL122 region, and has a false negative rate of 24%

•There are 8 unique locations where a mismatch occurs 7 of which are found in the Toledo strain (AC146905)

•Only AD169, Towne, Merlin and one clinical isolate (4/14) do not contain a mismatch with this qPCR assay.

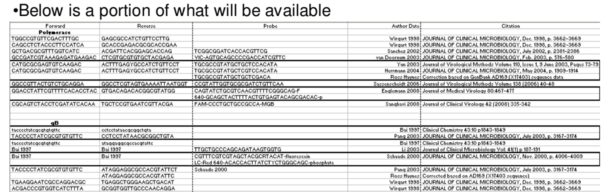
•Future researchers can avoid spending time and money on assays which have a high degree of sequence heterogeneity.

Primer Citations

•Information collected:

- –Primers and probes
- –Original paper and other papers which describe the primers and probe or alternate probes
- –Typos found based on Genbank sequence information

•Below is a portion of what will be available



QCMD CMV DNA External Quality Assessment 2009

NIST was invited to submit a sample of the SRM candidate material for analysis by the participants in the 2009 CMV DNA EQA.

▪The concentration of the packaged DNA was 3.8 M copies / mL. determined by digital PCR (Expanded uncertainty 3.4 M to 4.2 M copies/mL). Log10 = 6.5

▪Tubes were sent to QCMD for labeling and shipping with the QCMD test samples.

▪Participants were asked to add the DNA solution directly to their assay and run the assay in triplicate.

▪155 data sets were returned

▪The handling of the sample by participants was different from the standard samples – virus in plasma

- ▪Quality of data submitted is questionable

Summary NIST sample from QCMD EQA 2009

•Data broadly spread SD > 0.5 log (10)

•In general, no correlation was seen between reported values in any of the four groupings and:

- – the use of an external calibrant,
- – the use of any particular commercial assay
- – or repeatability within a lab.

•Those using lab developed assays reported values in all groups, but more in low range than those using commercial assays

Data summary

Technology Group	Mean	Log10	Std Dev	Range for	
	Copies/mL	Copies/mL	log 10	1 Std Dev	# data sets
RT Commercial	2.96 M	6.47	0.67	0.63 M - 14.03 M	74
RT In-house	1.44 M	6.16	0.89	0.18 M - 11.12 M	74
Conven. Commercial	1.49 M	6.17	0.52	0.45 M - 4.90 M	5
Maximum Likelihood					
Gaussian Mix two main peaks	1.45 M	6.16	0.21	0.89 M - 2.34 M	49
	6.16 M	6.79	0.34	2.82 M - 12.39 M	52
			U log 10		
Digital PCR	3.8 M	6.58	0.04	3.42M - 4.18 M	

Distribution of datasets within four groupings

		CutPoints				%					
Code	#	<5.733	5.738 -6.436	6.436-7.124	> 7.124	Low	Gp1	Gp2	High	%Out	
In-house	74	19	24	26	5	26	32	35	7	33	RTIH
QIAGEN	28	1	16	9	2	4	57	32	7	11	
Nanogen	15	2	3	7	3	13	20	47	20	33	
Argene	14	1	1	10	2	7	7	71	14	21	
Roche-LightCycler	8	0	7	1	0	0	88	13	0	0	
Roche-COBAS	5	1	2	2	0	20	40	40	0	20	
Cepheid	4	0	4	0	0	0	100	0	0	0	
astra	1	0	1	0	0	0	100	0	0	0	
ATQ	1	0	1	0	0	0	100	0	0	0	
Diagenode	1	1	0	0	0	100	0	0	0	100	
GeneProof	1	0	0	0	1	0	0	0	100	100	
LightUp	1	0	0	1	0	0	0	100	0	0	
Sum	153	25	59	55	13	17	39	36	8	25	All
										19	RTC

References:
[1] Hemmerich 1998 www.deviselink.com/mpb/archive/98/07/002.html
[2] Kotenko et al. PNAS February 15, 2000 vol 97 no. 4 pp. 1695-1700
[3] X17403 "AD169" Chee et al. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. 154, 125-169 (1990). Sequence revised Nov 14, 2006
[4] AY315197 "Towne" Dunn et al. Functional profiling of a human cytomegalovirus genome. Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14223-14228 (2003). Sequence revised Jan 31, 2007
[5] AC146904 - AC146907 "PH-BAC", "Toledo BAC", "TR-BAC", "FIX-BAC" Murphy et al. Coding potential of laboratory and clinical strains of human cytomegalovirus. Proc. Natl. Acad. Sci. U.S.A. 100 (25), 14976-14981 (2003). Sequence added Dec 10, 2003
[6] NC_006273 Dolan et al. Genetic content of wild-type human cytomegalovirus. J. Gen. Virol. 85 (PT 5), 1301-1312 (2004). Sequence revised Aug 28, 2009
[7] EF999921 "TB40-BAC4" Sinzger et al. Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. J. Gen. Virol. 89 (PT 2), 359-368 (2008). Sequence added Jan 18, 2008
[8] M95634 - M95639 Chou. Effect of interstrain variation on diagnostic DNA amplification of the cytomegalovirus major immediate-early gene region. J. Clin. Microbiol. 30 (9), 2307-2310 (1992). Sequence added Aug 2, 1993