NIST Candidate Standard Reference Material (SRM): Cytomegalovirus DNA

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

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

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

Poster #ID19 at Association for Molecular Pathology 2010 Annual Meeting, San Jose, CA. November 18-20, 2010

Cytomegalovirus Reference Material

- **Disease:** Cytomegalovirus (CMV) causes life-threatening infections in immune-compromised 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***: DNA from Towne strain in a bacterial artificial chromosome (BAC) 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. BAC provided by Dr. Hua Zhu at the University of Medicine and Dentistry of New Jersey
- ***Certification***:
  1. DNA sequence of genes that are targets for qPCR
  2. Concentration determined via digital PCR
- ***Additional Information***:
  - Testing of various published PCR assays
  - Homogeneity and stability

Abstract

Introduction: Cytomegalovirus (CMV) can lead to severe and potentially life threatening disease in immune-compromised patients. Measurements of viral load are used to make decision on when, how and for how long to treat. The most important tool for viral load measurements is quantitative real-time PCR. About 50% of clinical laboratories are using commercial reagents and 50% using laboratory- developed assays. The wide variety of methods used in laboratories has contributed to inter-laboratory variability as shown in proficiency testing programs. A second contributor to testing variability is the lack of reference materials for quantification. To address this need NIST is certifying a traceable Standard Reference Material (SRM) for CMV.
Materials and Methods: CMV Towne strain DNA in a bacterial artificial chromosome was propagated in E. coli and purified. The DNA was packaged at three different concentrations in 0.5 mL polyfluoroalkoxy (PFA) tubes. The DNA sequence of relevant regions was determined using Sanger sequencing methodology in both directions. Digital PCR is the approach used for determination of the concentration (copies/mL) for each component of the SRM and will use multiple assays targeted to different parts of the CMV genome. Stability studies of the SRM have been conducted for over a year using PCR assays.
Results: The pure CMV DNA preparations have proved to be very stable at 4 °C based on quantitative PCR (qPCR) assays. Almost 17,000 bases of the CMV genome have been sequenced with multiple coverage in both directions. The regions sequenced are targets for qPCR – both proprietary and published – and a site where mutations lead to drug resistance; they are: UL34, UL54, UL55 to UL56, UL80, UL83, UL97, UL122 to UL126, UL132, and US17. The regions sequenced exactly match the Towne strain (GenBank accession AY315197.2) except for one ambiguous base in the UL55 to UL56 region. The determination of the concentrations with expanded uncertainties of each SRM component is in progress with approximate concentrations of 20,000, 2000, and 1000 copies/mL.
Conclusions: When released, the CMV SRM can be used to provide traceability for commercial and lab-developed calibration reagents, quality control materials and assay reagents.

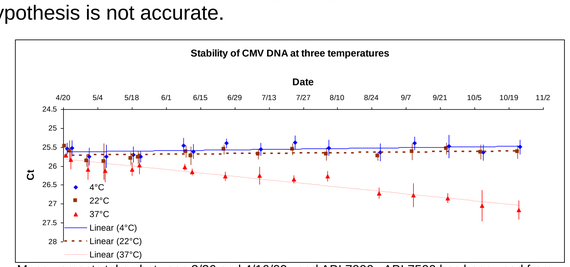
Candidate Materials

Candidate SRM was packaged in PFA screw cap tubes

- Three concentrations were prepared ~20,000, ~2,000, and ~500 copies/υL
- One concentration (~2,000 copies/μL) was included in the QCMD 2010 CMV EQA
- Sequencing is complete
- Copy number determination is still underway.

Stability

- Previously prepared "Pilot Material" – prepared from the same concentrated stock which the Candidate Material was made – appears stable at 4 C and 22 C.
- 2 of 348 tubes tested have divergent concentrations
- Several single replicates have divergent Cts, but results are not repeatable
  - –We hypothesize that these wells did not seal properly during PCR and the decrease in Ct is due to increase in concentration via evaporation
- DNA stored at 4 C and 22 C appear stable, but signs of degradation appeared after 8 weeks in DNA stored at 37 C
- Simple accelerated aging: for every 10 C over room temperature aging is doubled [1]. According to this hypothesis, the DNA stored at 22 C should have started degrading already; therefore hypothesis is not accurate.



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)

QCMD CMV DNA External Quality Assessment (EQA) 2010

NIST collaborated with Quality Control for Molecular Diagnostics (QCMD) for an inter-laboratory study. NIST provided aliquots of CMV DNA, component B of the candidate CMV SRM, to QCMD. For this year's External Quality Assessment program for CMV, a vial of DNA was sent out to each participating laboratory along with the 10 QCMD samples (lyophilized virus in plasma or VTM) While the QCMD samples required extraction, the NIST DNA was to be added directly to the assay. Participants were asked to run the assay in triplicate and report results in copies/mL. 181 data sets were submitted.

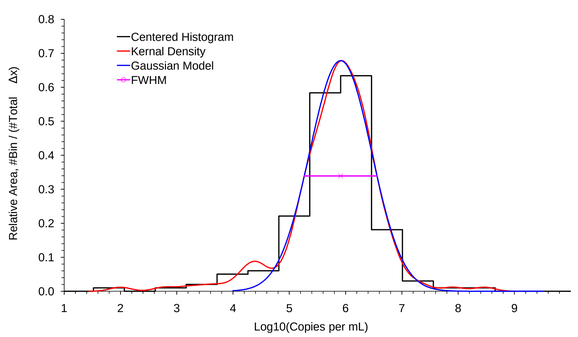
The handling of the sample by participants was different from the standard samples – no was extraction required; therefore, some data may be biased because of miss-use.

Distribution of Detection Chemistries

Each participating lab determined the concentration of the candidate SRM by using a commercial or laboratory developed quantitative PCR assay. There were a few labs which returned non-real-time quantitative data.
Data were analyzed based on assay type as shown in the table. For comparison purposes the assigned value for the CMV DNA was 6.200 Log_{10} copies/mL and the standard deviations of the other QCMD panel components ranged from 0.4 to 0.5 Log_{10} copies/mL

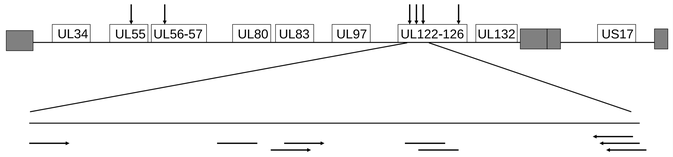
QCMD CMV EQA - Assays	# Data sets	Log_{10} copies/mL Median	MADe
Total Datasets	181	5.900	0.486
Conventional Commercial	5	5.854	0.872
Real-Time Laboratory developed - Total	78	6.002	0.650
Real-Time Commercial - Total	96	5.826	0.451
Argene CMV HHV6,7,8 R-gene	6	5.864	0.150
Argene CMV R-gene	15	6.205	0.332
Nanogen Q-CMV Real time Complete Kit	21	5.733	0.794
QIAGEN artus CMV PCR Kit (RG,LC,TM)	28	5.821	0.326
Roche LightCycler CMV Quant Kit	12	5.776	0.298

Distribution of Results



Each participating lab determined the concentration of the candidate SRM. Individual data points which were suspected to was removed by QCMD. The remaining data fits to a Gaussian Model, though there are low outliers. Robust statistics were used, giving a median of 5.900 Log_{10} copies/mL and a MADe (SD of the median) of 0.486 Log_{10}. The mid-point of the Kernel Density was 5.912 Log_{10} copies/mL.

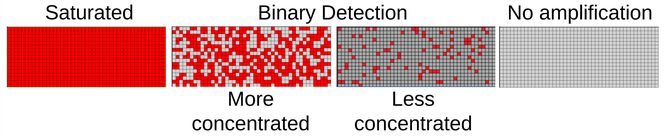
CMV Genome: dPCR targets and locations sequenced



- The CMV genome is ~ 240 kb long. There are two sets of inverted repeats (shaded boxes) that divide the genome into two unique regions designated Unique Long (UL) and Unique Short (US). The open reading frames (ORF) of each unique region is numbered sequentially, so UL97 is the 97th ORF in the UL region.
- Locations sequenced (~17 kb) are indicated by labeled boxes.
- The downward facing arrows indicate the location of dPCR targets, which are TaqMan assays. The expanded location shows overlapping targets that were chosen to test for measurement bias due to PCR amplicon length.

Concentration via Digital PCR (dPCR)

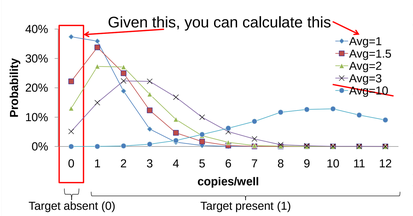
- dPCR can calculate absolute concentration of DNA via PCR without use of an external calibrant
- Digital PCR = Binary (011001…) PCR
- Binary detection – some replicate wells will not have target
- Calculate concentration using Poisson statistics



Poisson Statistics

Stats of counting events (0, 1, 2…∞)

As average copies/well increases, probability of well with Target Absent (0) decreses.



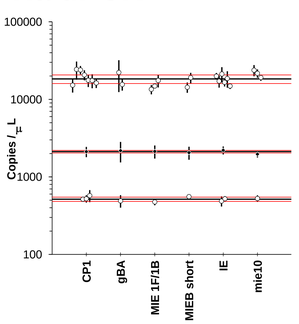
- If you know the proportion of wells without target – in the range of binary detection – then, you can calculate the average concentration using Poisson statistics.
- If the wells are saturated or there is no amplification, then you can glean very little information.

Concentration via dPCR

Digital PCR was used to analyze the three concentrations of the candidate SRM. The concentration is in copies/μL

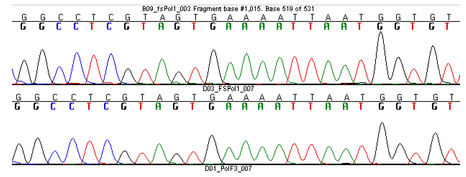
A: 515 (95%CI: 483 to 548)
B: 2118 (95%CI: 2037 to 2199)
C: 18366 (95%CI: 16026 to 20646)

The PCR assays are in the same order as in the genome that is mapped above.



Sequencing

- Sequence via Big Dye Terminator v3.1 (Applied Biosystems)
- Analyze by capillary electrophoresis using Applied Biosystems 3130xl with POP7 and an 80 cm capillary
- Towne strain sequence GenBank AY315197 used as reference
- BAC has the same sequence as Towne strain (AY315197) in regions already sequenced – except for one ambiguous base.
- Regions sequenced: most or all of UL34, UL54, UL55-56 (polymerase & glycoprotein B), UL80, UL83 (phosphoprotein 65), UL97, UL122-126 (major immediate early area and nearby sequence), UL132, and US17
- Targets for commercial reagents: UL34, UL80, UL132
- Targets for published assays: UL54, UL55-56, UL83, UL122-126, US17
- Mutations in UL97 confer resistance to ganciclovir
- Example: sequence from UL55 of CP1 reverse primer
  - –Sequence matches Towne strain and published primer sequence with 5 times coverage (2 x coverage shown)



Clinical DNA Informational Resource

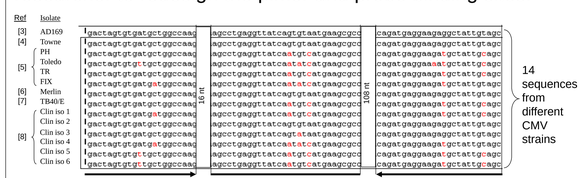
http://www.nist.gov/mml/biochemical/genetics/clinical_dna.cfm

Clinical DNA Informational Resource will contain:

- Recent publications and presentations – including this one
- Information DNA-based Standard Reference Materials (SRMs) and candidate SRMs
- Information to aid scientists in their research

Sequence alignments

The qPCR assay below [2] has a large number of mismatches that cause a false negative rate of 24%.
A little more time in the library, probably would have lead this researcher to redesign the pimer and probe binding sites.



Published Assays

- 73 unique qPCR assays have been collected, the primers and probes have been documented and put into one spreadsheet.
- Errors have been corrected and traced to:
  1. Transcribing from paper to paper and
  2. Changes in GenBank sequences.
- Each assay has a key to which strains of CMV [3] – [7] match that assay, which is a rough indicator of the sensitivity of the assay.

References:
[1] Hemmerich 1998 www.devioslink.com/mpb/archive/98/07/002.html
[2] Tanaka 2000 Journal of Medical Virology 60:455–462
[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

Poster available for download from NIST Applied Genetics website: http://www.nist.gov/mml/biochemical/genetics/index.cfm