Overview of NIST
U.S. National Institute of Standards and Technology

- National Metrological Institute (NMI)
- Founded in 1901
- ~3000 employees across multiple campuses

Gaithersburg, MD campus

- Maintains time measurement for the U.S. (atomic clock)
- Four Nobel prize winners
- NIST supplies over 1,300 Standard Reference Materials (SRMs) for industry, academia, and government use in calibration of measurements
Metrological Traceability

Unit Definition

Primary Standard

Secondary Standards
Maintained by NMI

Working Standards
Calibrated against the secondary standard

Laboratory Measurements
Users: industry, medical labs, academia, government

Traceability
Cost

Increasing measurement uncertainty
SRM intended use: higher order standard

- SRM 2366 (CMV DNA) sales:
  - majority of customers are manufacturers
Levels of Confidence
Information contained in SRM certificate of analysis

Certified

- A NIST certified value is a value for which NIST has the highest confidence in its accuracy in that all known or suspected sources of bias have been investigated or taken into account.

Reference

- A NIST Reference Value is a best estimate of the true value provided by NIST where all known or suspected sources of bias have not been fully investigated by NIST.

Information

- An information value is considered to be a value that will be of interest and use to the SRM user, but for which insufficient information is available to assess adequately the uncertainty associated with the value, or a value derived from a limited number of analyses.

NIST Efforts in DNA-based Clinical Standards

**Applied Genetics Group**, led by Peter Vallone

- Margaret Kline, Jo Lynne Harenza, Ross Haynes, **Marcia Holden** (retired)

- Forensic DNA standards, Huntington’s Disease, CMV, BKV

**Bioassay Methods Group**, led by Steven Choquette and Ken Cole

- Currently working on a *Her2* DNA copy number and pathlength nanoliter spectrophotometer standards

**Genome Scale Measurements Group**, led by Marc Salit

- Genome in a bottle (NA12878), RNA external controls
Nature of our DNA-based SRMs

- Extracted DNA (source: viral, human, synthetic, recombinant, cell lines)

- Highly characterized and certified for specific properties
  - Sequence (Sanger → NGS)
  - Copy number or concentration (UV → dPCR)
<table>
<thead>
<tr>
<th>SRM</th>
<th>NIST DNA-based SRMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2366</td>
<td>Cytomegalovirus (CMV) for DNA Measurements²</td>
</tr>
<tr>
<td>2393</td>
<td>CAG Repeat Length Mutation in Huntington's Disease¹</td>
</tr>
<tr>
<td>2374</td>
<td>DNA Sequence Library for External RNA Controls³</td>
</tr>
<tr>
<td>2399</td>
<td>Fragile X Human DNA Triplet Repeat Standard⁴</td>
</tr>
<tr>
<td>2372</td>
<td>Human DNA Quantitation Standard¹</td>
</tr>
<tr>
<td>2391c</td>
<td>PCR Based DNA Profiling Standard¹,⁵</td>
</tr>
<tr>
<td>2392, 2392-I</td>
<td>Mitochondrial DNA Sequencing¹</td>
</tr>
<tr>
<td>2394</td>
<td>Heteroplastic Mitochondrial DNA Mutation Detection Std⁴</td>
</tr>
<tr>
<td>2395</td>
<td>Human Y-Chromosome DNA Profiling Standard¹</td>
</tr>
<tr>
<td></td>
<td>Candidates currently under characterization</td>
</tr>
<tr>
<td>BK Virus³</td>
<td></td>
</tr>
<tr>
<td>HER2 Copy Number Measurement¹</td>
<td></td>
</tr>
<tr>
<td>Pathlength Standard for Nanoliter Spectrophotometers⁶</td>
<td></td>
</tr>
<tr>
<td>Genome in a Bottle (NA 12878)¹</td>
<td></td>
</tr>
</tbody>
</table>

¹extracted genomic DNA (human); ²extracted genomic DNA (viral in BAC); ³extracted DNA (plasmid) ⁴PCR products; ⁵cell lines on paper substrate; ⁶Uracil and Tryptophan solutions
Utility of Extracted DNA Materials

- Purified materials allow for high degree of characterization (copy number, sequence content)

- Do not exhibit the heterogeneity of the practical analyte: authentic clinical samples (matrix, cell, tumor, FFPE, inhibitors)
  - Characterized with a high degree of certainty, but are they of practical utility?

- Useful for understanding: extraction efficiency, assay bias due to sequence variants, the performance of an assay (LOD), post extraction characterization

- The certified properties of a material should not vary – stable and homogeneous – allows for traceability
Understanding variation in a process: Cause and Effect Analysis (General example)

- **Sample handling**
- **Dilution**
- **Losses to the tube**
- **Pipetting**
- **Homogeneity**
- **Spinning down/prep process**

- **Sampling**
  - **Dilution**
  - **Losses to the tube**
  - **Pipetting**
  - **Homogeneity**
  - **Spinning down/prep process**

- **DNA quantification**
  - **Fluorometry (Qubit)**
  - **Stability**
  - **Reproducibility**

- **Operator**
  - **Following the protocol**
  - **User experience**
  - **Data interpretation**

- **Use of NIST materials**

- **Sample preservation**
- **Homogeneity**
- **Dilution**
- **Counting**

- **Extraction methods**
- **Cell lysis**
- **DNA extraction**

- **Primes/probes**
- **Platform**
- **Software analysis**

- **Pipetting**
- **Setup**
  - **Workspace**
  - **Storage stocks**
  - **Reagents**

- **qPCR**

(Sandra deSilva – NIST)
Human Cytomegalovirus (HCMV)

- ~230kb herpesvirus, aka human herpesvirus 5 (HHV-5)
- Infects 50-80% of U.S. adults (>90% of adults worldwide)*
- Life-threatening for the immunocompromised: HIV-infected, organ transplant recipients, newborns
  - Between 0.2 - 2% of newborns infected**
  - Congenital HCMV infection = leading cause of deafness, learning disabilities, and mental retardation in children**

CMV SRM 2366 Status Update

- CMV Towne$\Delta_{147}$ Strain in a BAC (obtained from Hua Zhu, NJ Medical School): $\sim222,047$bp construct
  - US1-15 genes ($\sim$9kb) removed and replaced with 8kb BAC DNA
  - UL147 gene removed (477bp) and replaced with a 1938bp high-copy origin of DNA replication and kanamycin resistance gene cassette: ori$N$/kan$^R$

- E. coli DNA maintained as carrier DNA
  - $\sim$96% E. coli and 4% CMV/BAC DNA

<table>
<thead>
<tr>
<th>Component</th>
<th>Genome Copies/uL</th>
<th>Expanded Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>420</td>
<td>301-523</td>
</tr>
<tr>
<td>B</td>
<td>1,702</td>
<td>1,446-1,959</td>
</tr>
<tr>
<td>C</td>
<td>19,641</td>
<td>18,924-20,359</td>
</tr>
</tbody>
</table>
CMV SRM 2366 Status Update

- Out of stock since Summer 2013 due to a significant loss of concentration

- **Component C:**

![Graph showing mean concentration of CMV SRM 2366 over different seasons with logarithmic values.

\[ \log_{10} = 4.29 \quad 4.10 \quad 4.07 \]
CMV SRM 2366 Status Update

- Re-extracted bulk material and performed mini-stability study:

![Graph showing mean concentration over time for different reprocesses (None, PCI, ProK) at different temperatures (04C, 22C, 37C).]
CMV SRM 2366 Status Update

- Re-bottled newly extracted material in March 2014
- Diluted CMV DNA to \(~5\) ng/\(\mu\)L
- 1 component = 2366a
- Homogeneity performed in April 2014
- Stability study started mid-May 2014
- Next generation sequencing performed May 2014
2366a Homogeneity Results

Normalized Concentration vs. Sub-Lot

Normalized Concentration range: 0.8 to 1.3

Sub-Lot range: 6.4 to 5.1
CMV Sequencing Data

- Sanger sequence information for SRM 2366: ~17Kb

Table 2  Sequenced Regions of the CMV BAC Towne\textsubscript{Δ147}

<table>
<thead>
<tr>
<th>Region</th>
<th>Nucleotide range</th>
<th>Bases sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>UL34</td>
<td>43202–44971</td>
<td>1770</td>
</tr>
<tr>
<td>UL54</td>
<td>77695–79992</td>
<td>2298</td>
</tr>
<tr>
<td>UL55 to UL56</td>
<td>80848–82731</td>
<td>1884</td>
</tr>
<tr>
<td>UL80</td>
<td>114401–116793</td>
<td>2393</td>
</tr>
<tr>
<td>UL83</td>
<td>118890–119937</td>
<td>1048</td>
</tr>
<tr>
<td>UL97</td>
<td>140784–142090</td>
<td>1307</td>
</tr>
<tr>
<td>UL122 to UL126</td>
<td>170525–173182</td>
<td>2658</td>
</tr>
<tr>
<td>UL132</td>
<td>176380–177192</td>
<td>813</td>
</tr>
<tr>
<td>US17</td>
<td>198929–199312</td>
<td>384</td>
</tr>
</tbody>
</table>

Nucleotide ranges are based on GenBank accession number AY315197.2, which is the Towne strain. (Another CMV Towne strain submission, FJ616285.1, has slightly different numbering, and there is also a Towne-BAC sequence with the accession number AC146851.1.)

Next-generation sequencing of CMV material, 2366a

• Deep sequenced one vial of the reprocessed material, 2366a

• Nextera XT library prep (only requires 1 ng DNA input)

• Illumina MiSeq 600-cycle kit, v3
  - 301-cycle paired-end reads

• Yield: > 3 million mapped reads and > 7,000 x coverage
CMV Sequencing Data

- Next-generation sequencing information for SRM
- Mapped to CMV Towne Strain, Accession No. AY315197.2

3,471,174 mapped reads
BK Virus Background

- ~5kb polyomavirus; shares 75% DNA sequence similarity with JC virus

- 80% of population infected with the virus, which disseminates into the urinary tract and kidneys, usually latent

- Severe reactions in the immunocompromised: kidney or multiple organ transplantation recipients*
  
  - Presents as renal dysfunction and abnormal urinalysis
  
  - **BK-associated nephropathy (BKVAN)** results when BKV replicates within the graft due to high doses of immunosuppressants administered (1-10% of kidney transplant patients)**

  - Up to 80% of these patients lose their grafts


• Obtained clinical isolates for 6 genotypes (Ia, Ic, III, IV, V, VI) from Linda Cook, Univ Washington

• First cloning attempts proved difficult
  - Ia cloned into pACYC177 only following deletion of non-coding control region (NCCR)

• Requests for multiple BKV genotypes prompted SRM experimental redesign
  - Limited DNA concentration of samples presented a challenge for amplification - whole genome approach
Choice of reference material affects BKV DNA quantitation by qPCR

- MPS (mixed patient sample) — pooled urine samples of 30 patients and extracted BKV DNA
- Dunlop strain (Ia genotype) — obtained from ATCC

Whole Genome Amplification of BKV genotypes

(a) DNA from clinical sample

(b) RE site

(c) Genome units

(d) gel analysis

ϕ29 polymerase

RCA (random primers)
All genotypes have been amplified and were recently fully sequenced

- Agilent Bioanalyzer confirms amplification of all 6 BKV clinical isolates (~5kb following digestion)

Illumina MiSeq (Nextera XT)
BK NGS Coverage Data

Dunlop (1a)

Ia

Ic

III

IV

V

VI
Plans for BKV Reference Material

- Use sequence data to design cloning experiments
- Clone each full genotype sequence into a plasmid, transform and propagate in *E. coli*, purify, and characterize
BKV Ia material is homogeneous

- DNA linearized and bottled in March 2013
  - Ia cloned into pACYC177 (deletion of non-coding control region)
- Material to be used for NIBSC interlaboratory study in collaboration with Sheila Govind
• Pending the results of the stability study SRM 2366a is projected to be available the fall of 2014

• Continue work on BKV – open to input on importance of multiple genotypes in the SRM
  - Will provide type Ia for NIBSC interlaboratory study
  - Evaluating synthetically made constructs

• Open to feedback from the community on future candidate reference materials
Acknowledgements

Applied Genetics Group, NIST

Jo Lynne Harenza, Ph.D.
Ross Haynes, M.S.
Margaret Kline, M.S.

Chemical Sciences Division, NIST
David Duewer, Ph.D.

peter.vallone@nist.gov