An update on CMV and BKV DNA reference materials

Standardisation of Genome Amplification Technologies Meeting
28 May 2014

Peter M. Vallone, Ph.D.
Applied Genetics Group
Biomolecular Measurement Division

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

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

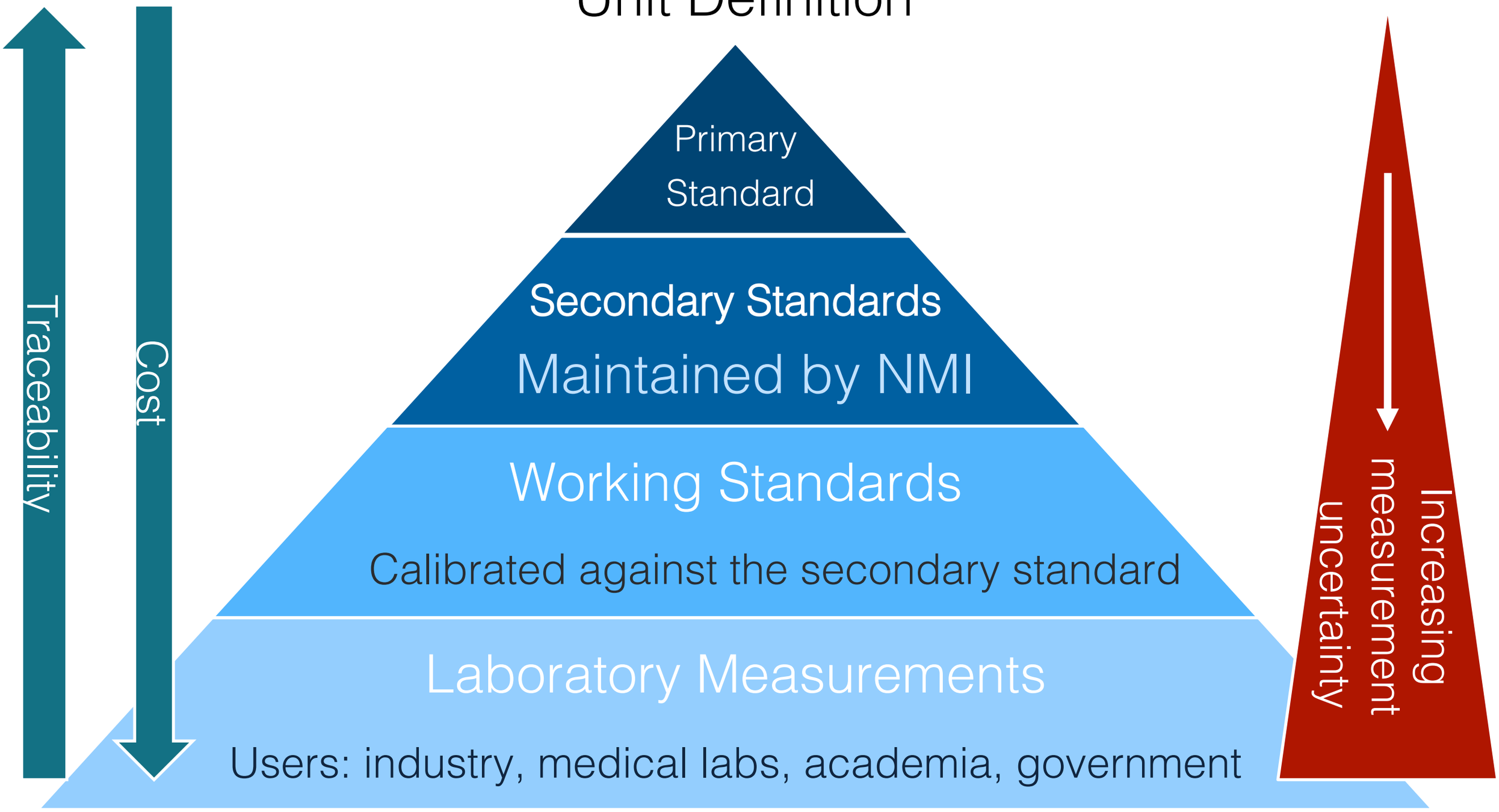
- National Metrological Institute (NMI)
- Founded in 1901
- ~3000 employees across multiple campuses

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

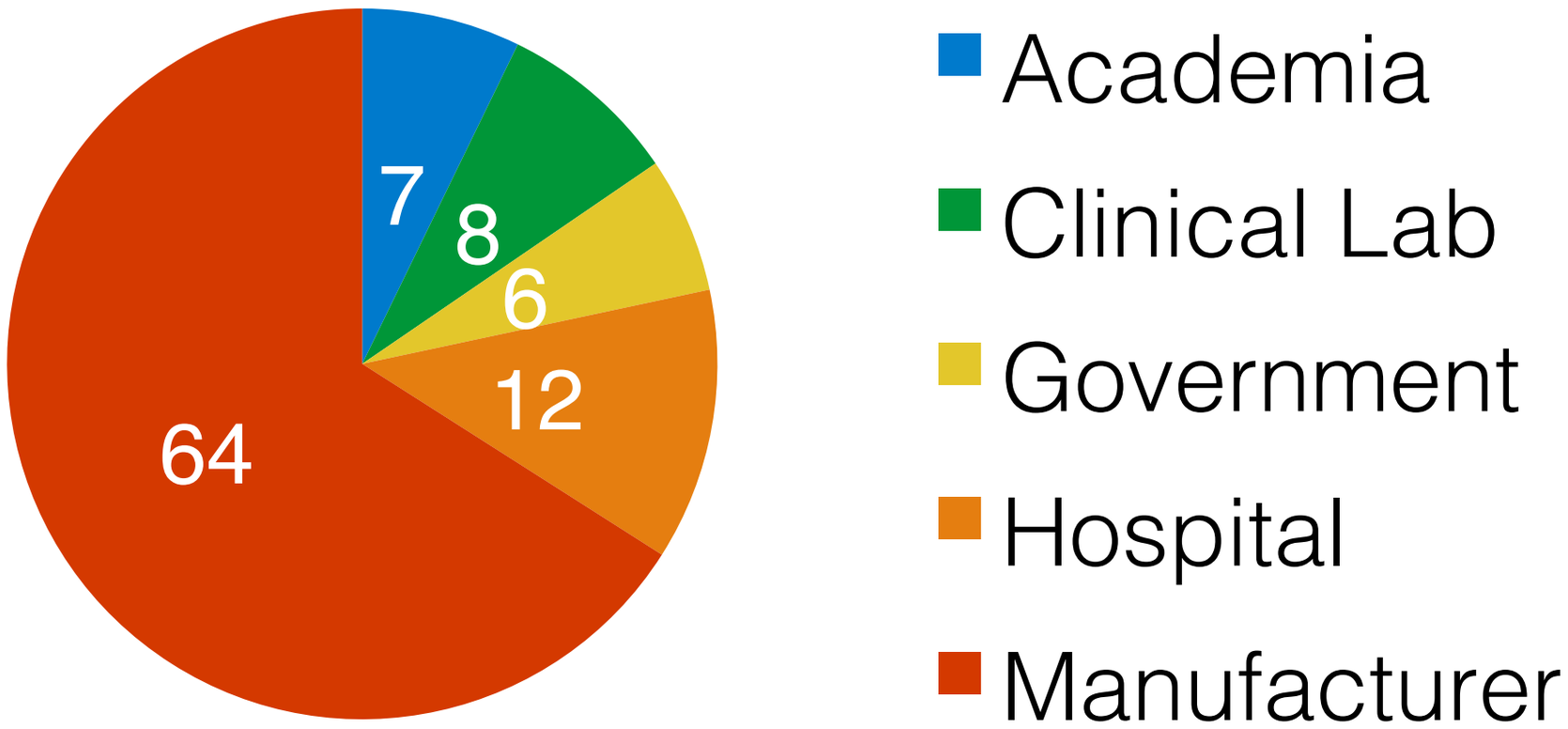
Metrological Traceability

Unit Definition



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.

May, W.E.; Gills, T.E.; Parris, R.; Beck, II, C.M.; Fassett, J.D.; Gettings, R.J.; Greenberg, R.R.; Guenther, F.R.; Kramer, G.; MacDonald, B.S.; Wise, S.A.; Definitions of Terms and Modes Used at NIST for Value-Assignment of Reference Materials for Chemical Measurements; NIST Special Publication 260-136 (2000); available at http://ts.nist.gov/MeasurementServices/ReferenceMaterials/PUBLICATIONS.cfm

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)

Clinical

DNA Typing

Forensic

SRM	NIST DNA-based SRMs
2366	Cytomegalovirus (CMV) for DNA Measurements[2]
2393	CAG Repeat Length Mutation in Huntington's Disease[1]
2374	DNA Sequence Library for External RNA Controls[3]
2399	Fragile X Human DNA Triplet Repeat Standard[4]
2372	Human DNA Quantitation Standard[1]
2391c	PCR Based DNA Profiling Standard[1,5]
2392, 2392-I	Mitochondrial DNA Sequencing[1]
2394	Heteroplasmic Mitochondrial DNA Mutation Detection Std[4]
2395	Human Y-Chromosome DNA Profiling Standard[1]
	Candidates currently under characterization
	BK Virus[3]
	HER2 Copy Number Measurement[1]
	Pathlength Standard for Nanoliter Spectrophotometers[6]
	Genome in a Bottle (NA 12878)[1]

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

Utility of Extracted DNA Materials

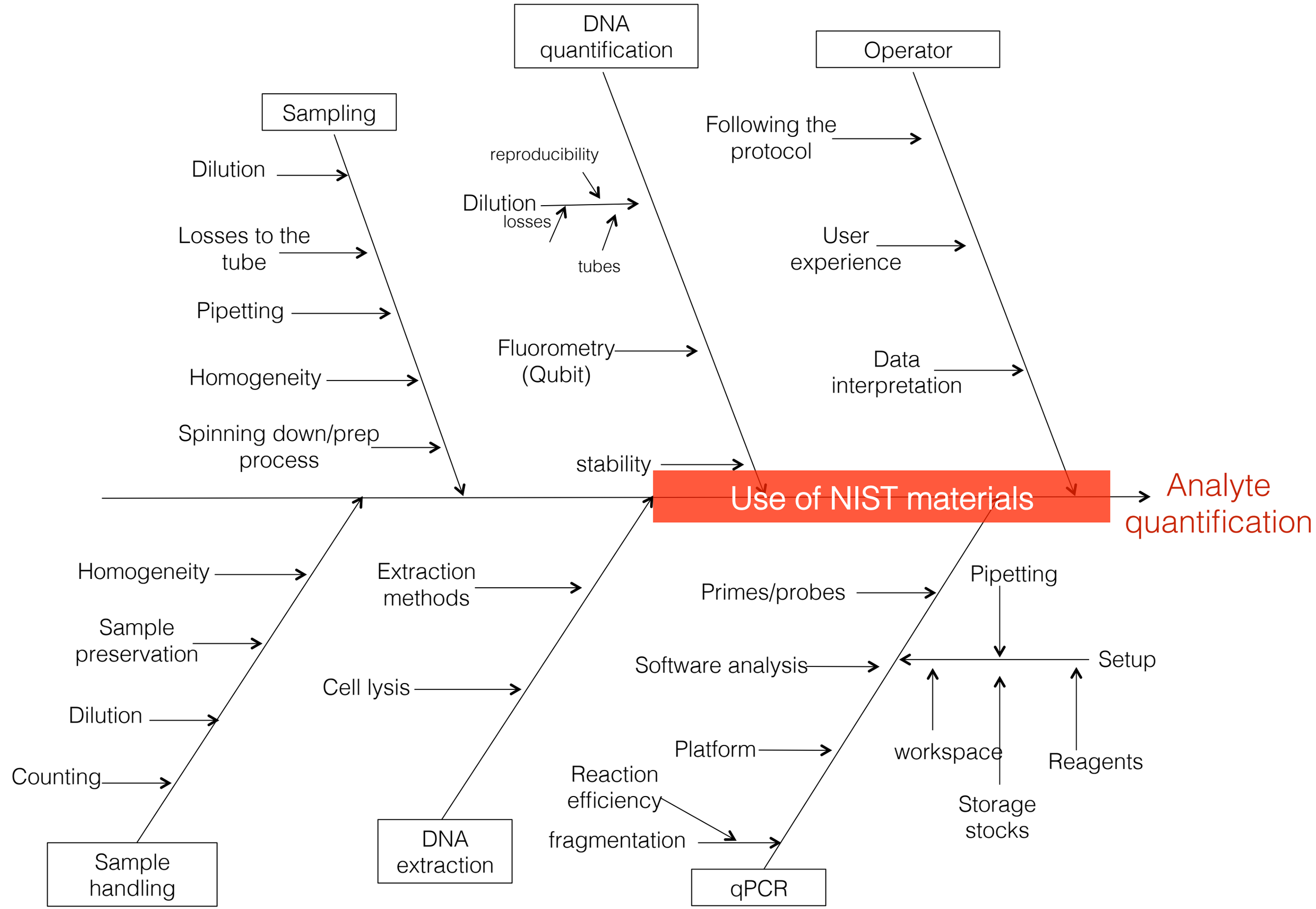
- Purified materials allow for high degree of characterization (copy number, sequence content)

- 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)



(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**

*Mocarski, et al, 2013. *Fields Virology* (6th ed.), pp.1960-2014.
Damanot & Winnen, 2006. “Cytomegalovirus infection: perinatal implications.” *J Obstet Gynecol Neonatal Nurs* **31 (1): 86-92

CMV SRM 2366 Status Update

- CMV Towne$_{\Delta 147}$ Strain in a BAC (obtained from Hua Zhu, NJ Medical School): ~**222,047bp** construct

  - US1-15 genes (~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 casette: *oriV*/*kan*^{R}

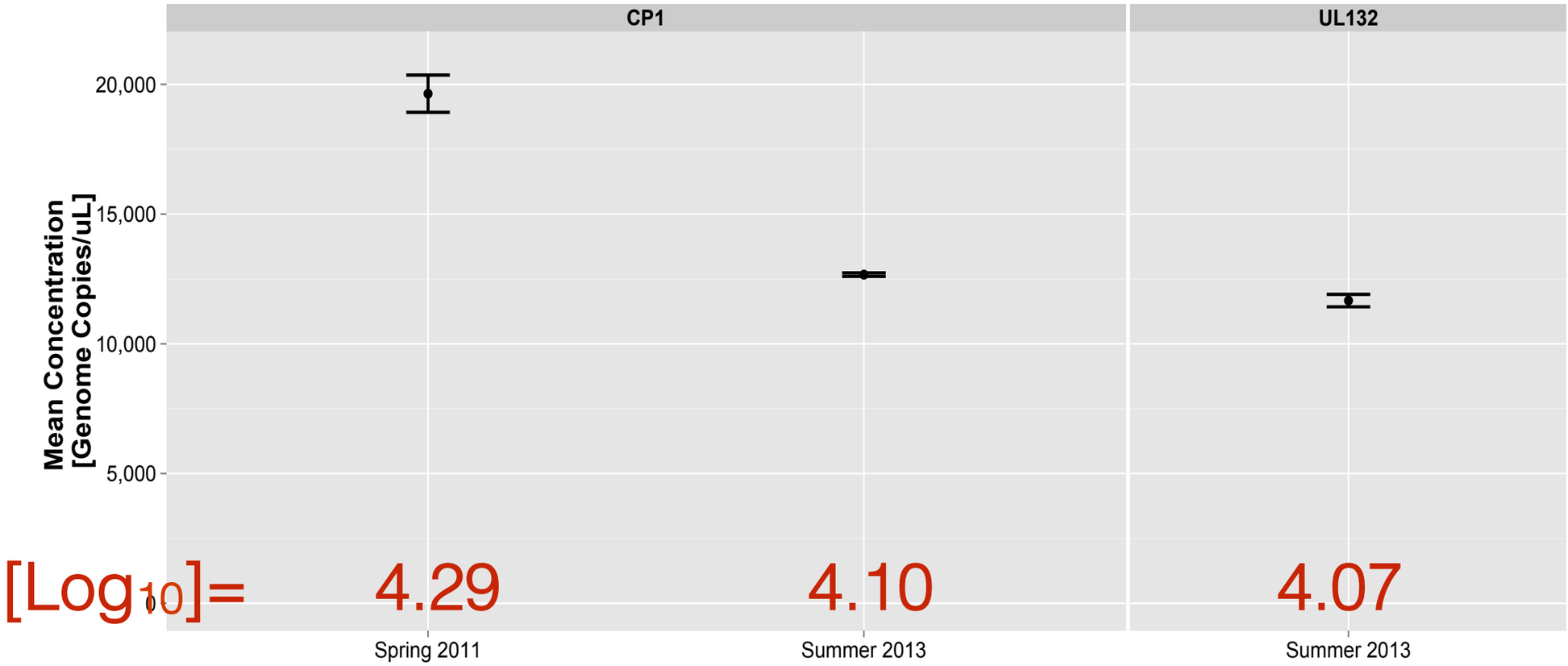
- *E. coli* DNA maintained as carrier DNA

  - ~96% *E. coli* and 4% CMV/BAC DNA

Component	Genome Copies/uL	Expanded Uncertainty
A	420	301-523
B	1,702	1,446-1,959
C	19,641	18,924-20,359

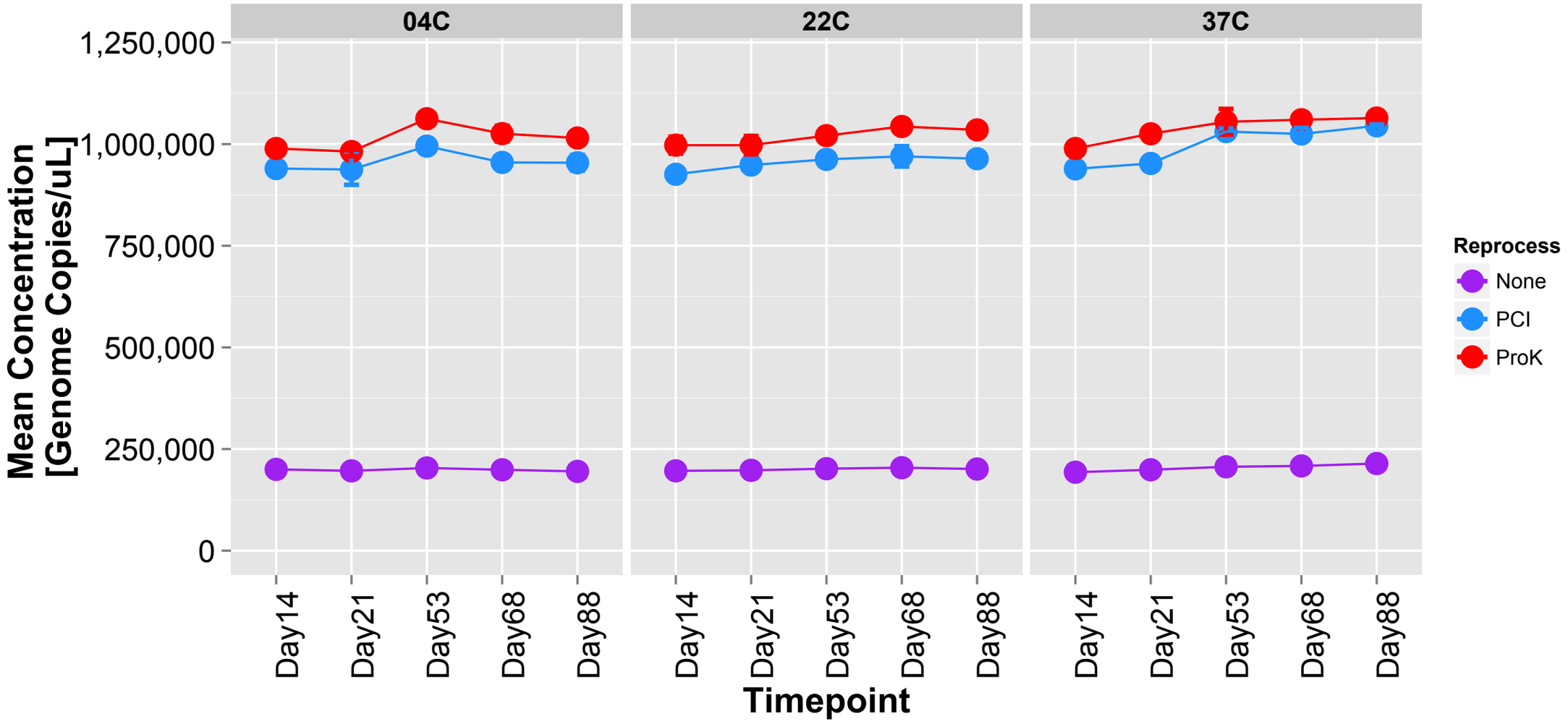
CMV SRM 2366 Status Update

- Out of stock since Summer 2013 due to a significant loss of concentration
- **Component C**:



CMV SRM 2366 Status Update

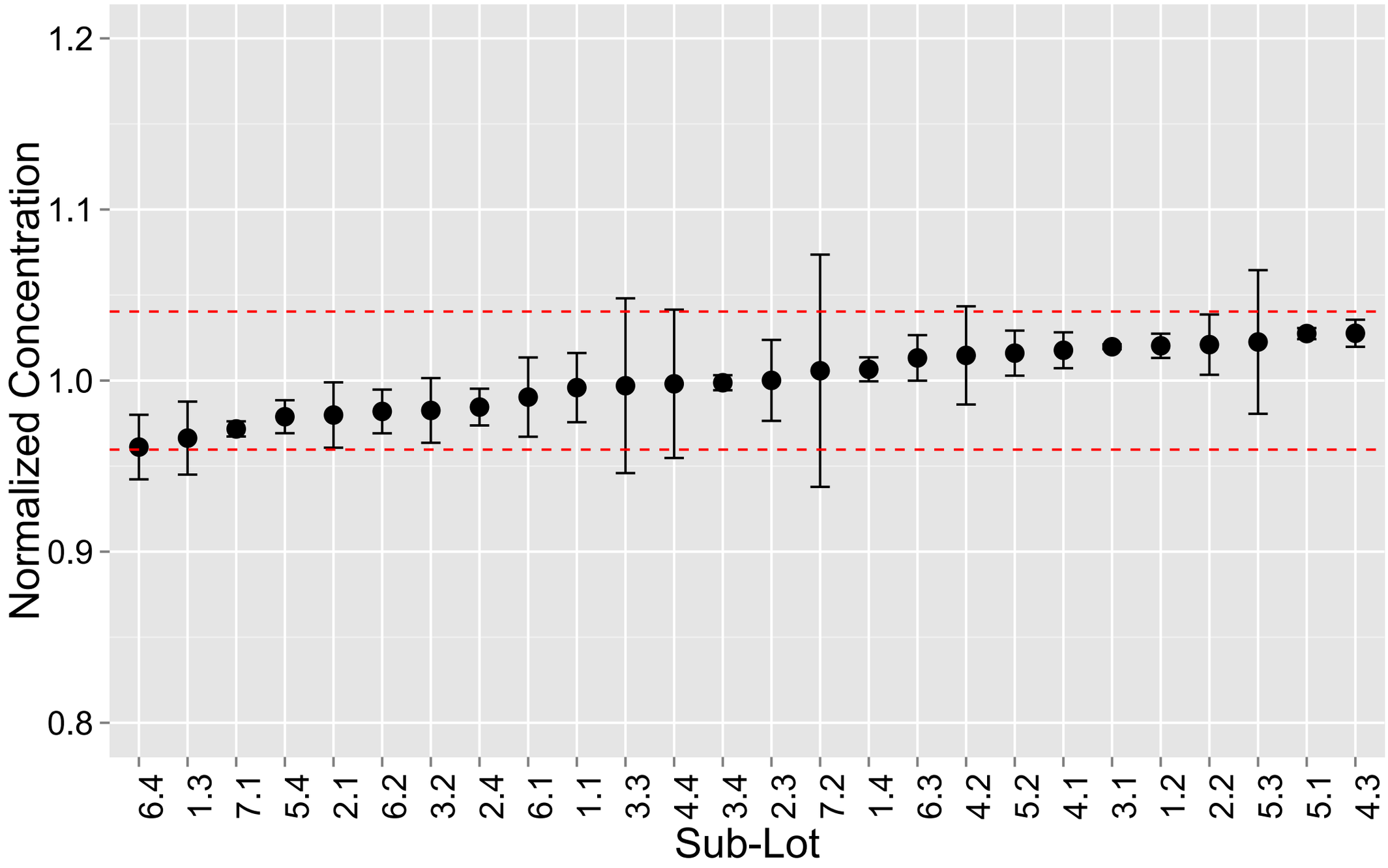
- Re-extracted bulk material and performed mini-stability study:



CMV SRM 2366 Status Update

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

2366a Homogeneity Results



CMV Sequencing Data

- Sanger sequence information for SRM 2366: ~17Kb

Table 2 Sequenced Regions of the CMV BAC Towne$_{\Delta 147}$

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

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.)

Haynes, et al. (2013). TECHNICAL ADVANCE. *The Journal of Molecular Diagnostics*, **15**(2), 177–185.

Next-generation sequencing of CMV material, 2366a

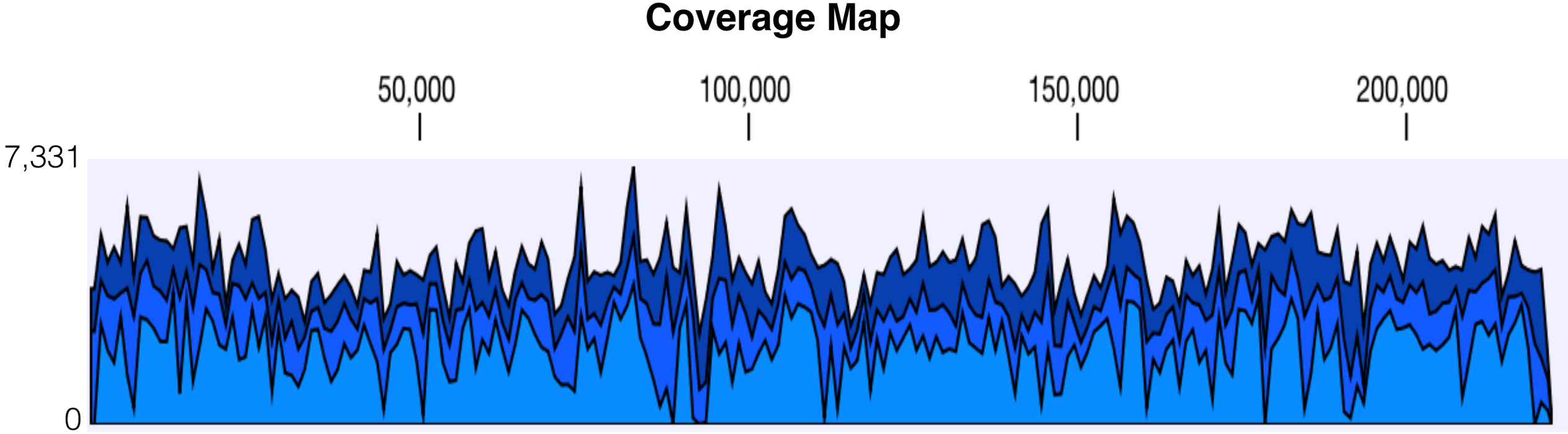
- Deep sequenced one vial of the reprocessed material, 2366a
- Nextera XT library prep (only requires 1 ng DNA input)
- Illumina MiSeq 600-cycle kit, v3
  - 301-cycle paired-end reads
- Yield: > 3 million mapped reads and > 7,000 x coverage

Illumina MiSeq

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

*Gupta, et al, 2006. “Low Incidence of BK Virus nephropathy after simultaneous kidney pancreas transplantation”. *Transplantation.* **82** (3): 382-8.

Fishman, 2002. “BK Virus Nephropathy - Polyomavirus Adding to Injury”. *New Eng J Med.* **347 (7) 527-530.

BKV DNA Reference Material

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

- First cloning attempts proved difficult

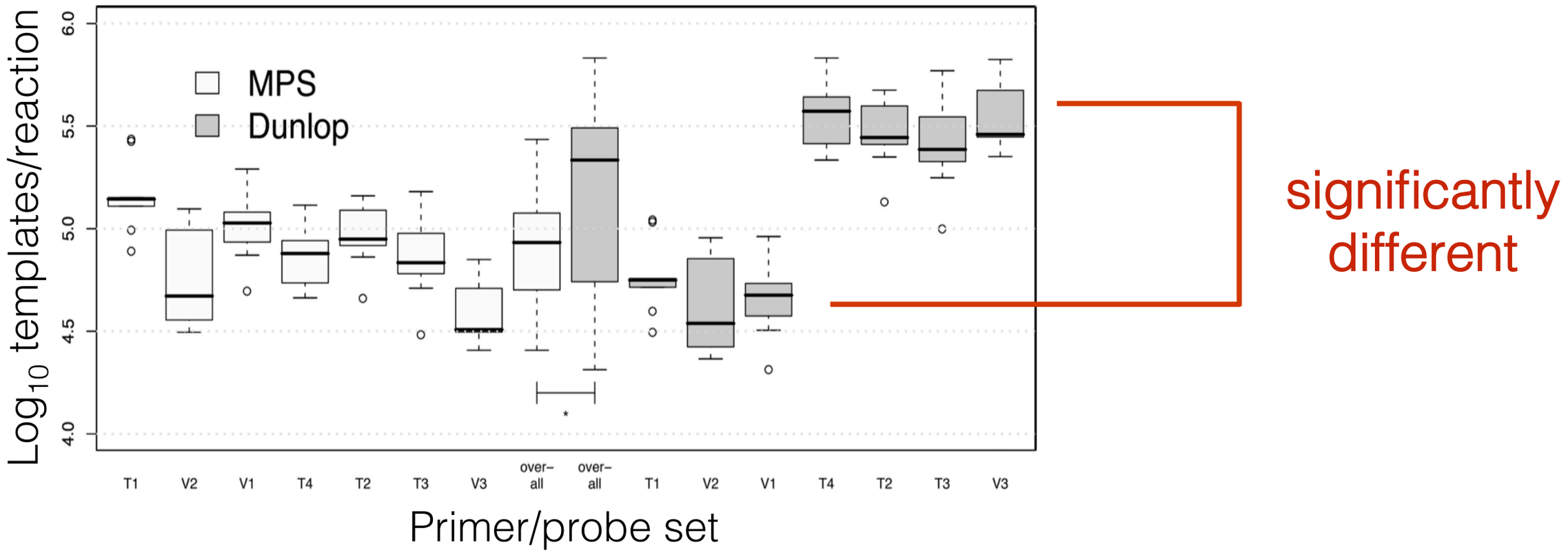
  - Ia cloned into pACYC177 only following deletion of non-coding control region (NCCR)

- Requests for multiple BKV genotypes prompted SRM experimental redesign

  - Limited DNA concentration of samples presented a challenge for amplification - whole genome approach

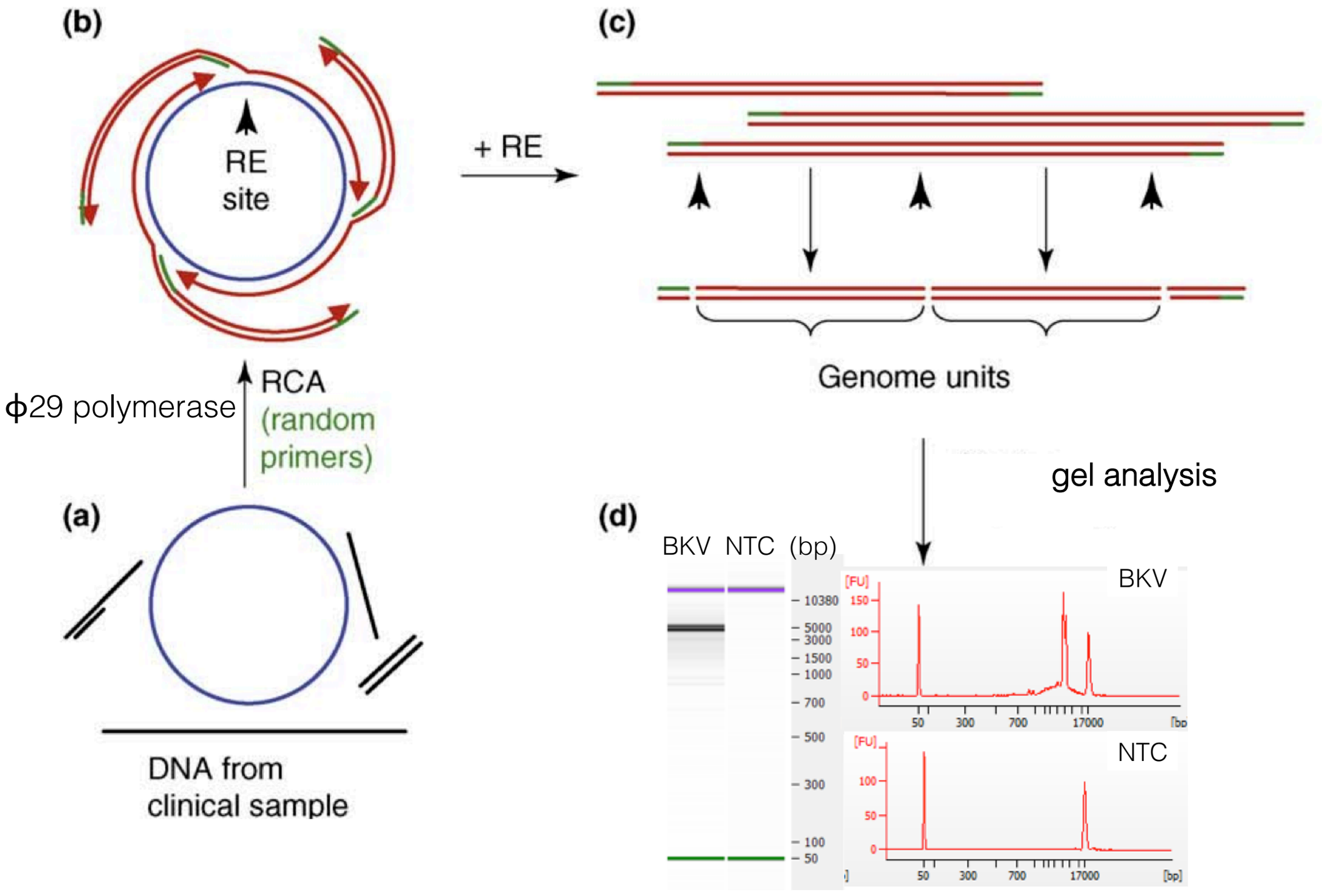
Choice of reference material affects BKV DNA quantitation by qPCR

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



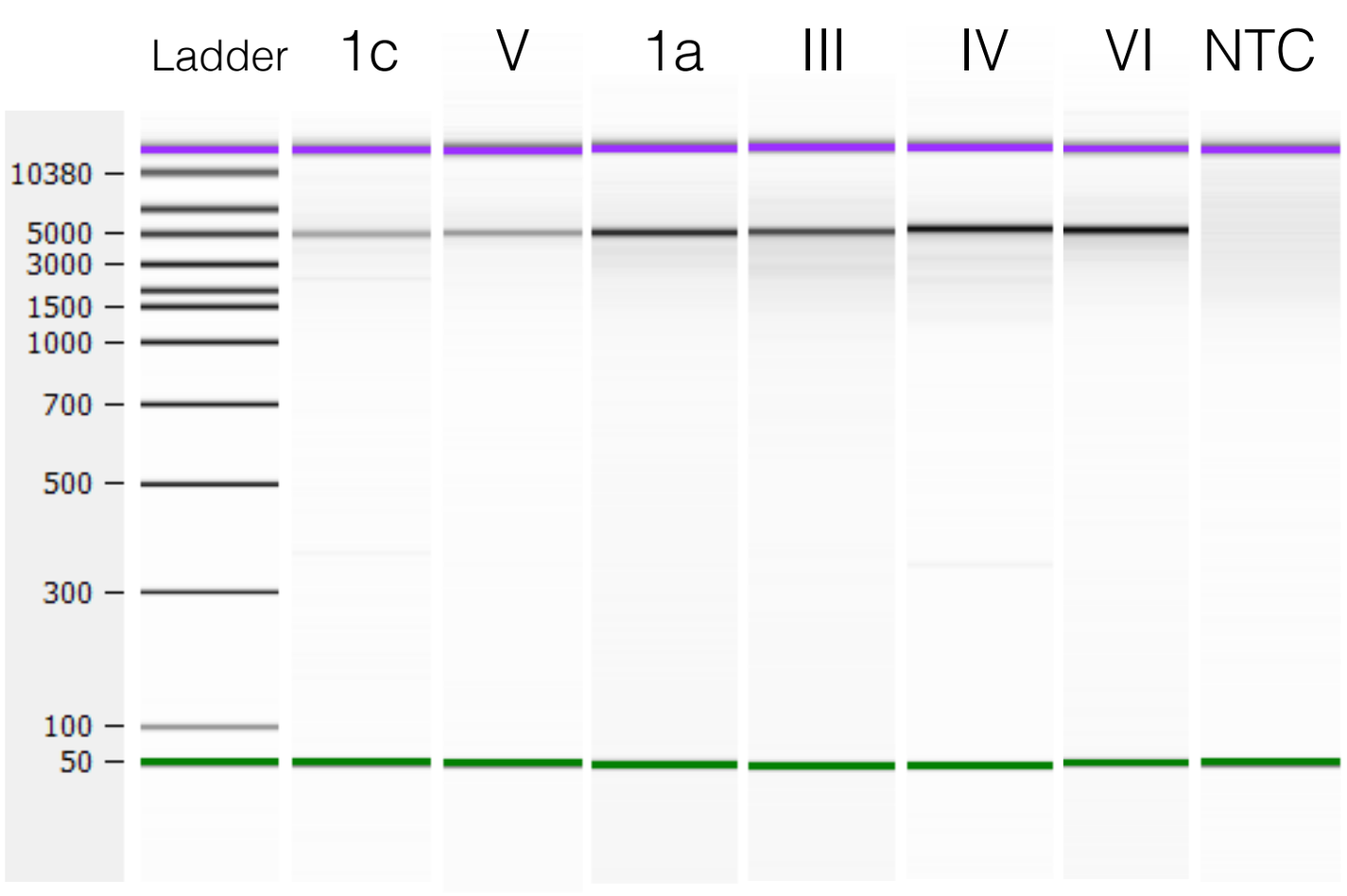
Hoffman, et al, 2008. Marked Variability of BK Virus Load Measurement Using Quantitative Real-Time PCR among Commonly Used Assays. *Journal of Clinical Microbiology,* **46**(8), 2671–2680.

Whole Genome Amplification of BKV genotypes



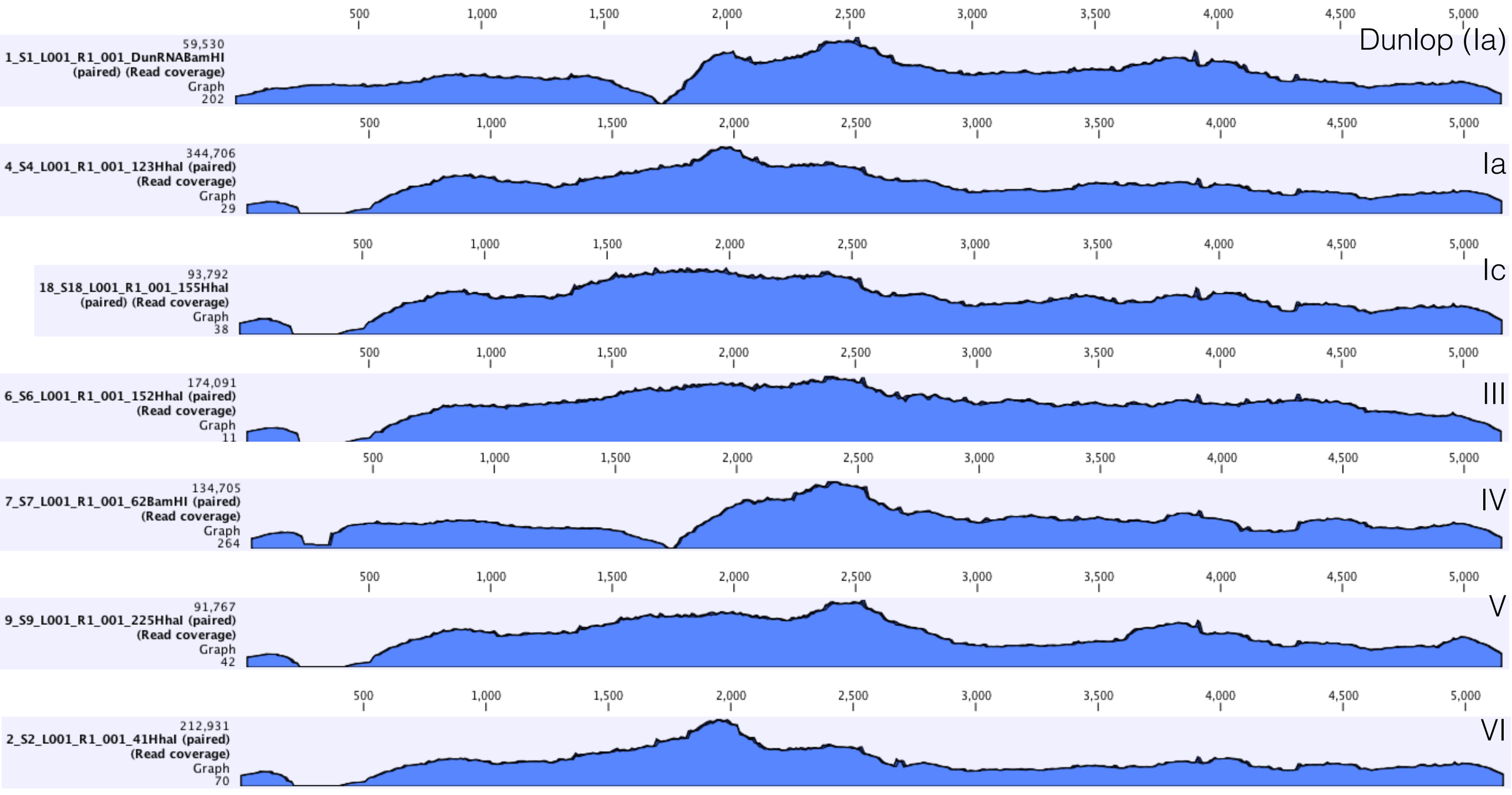
All genotypes have been amplified and were recently fully sequenced

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



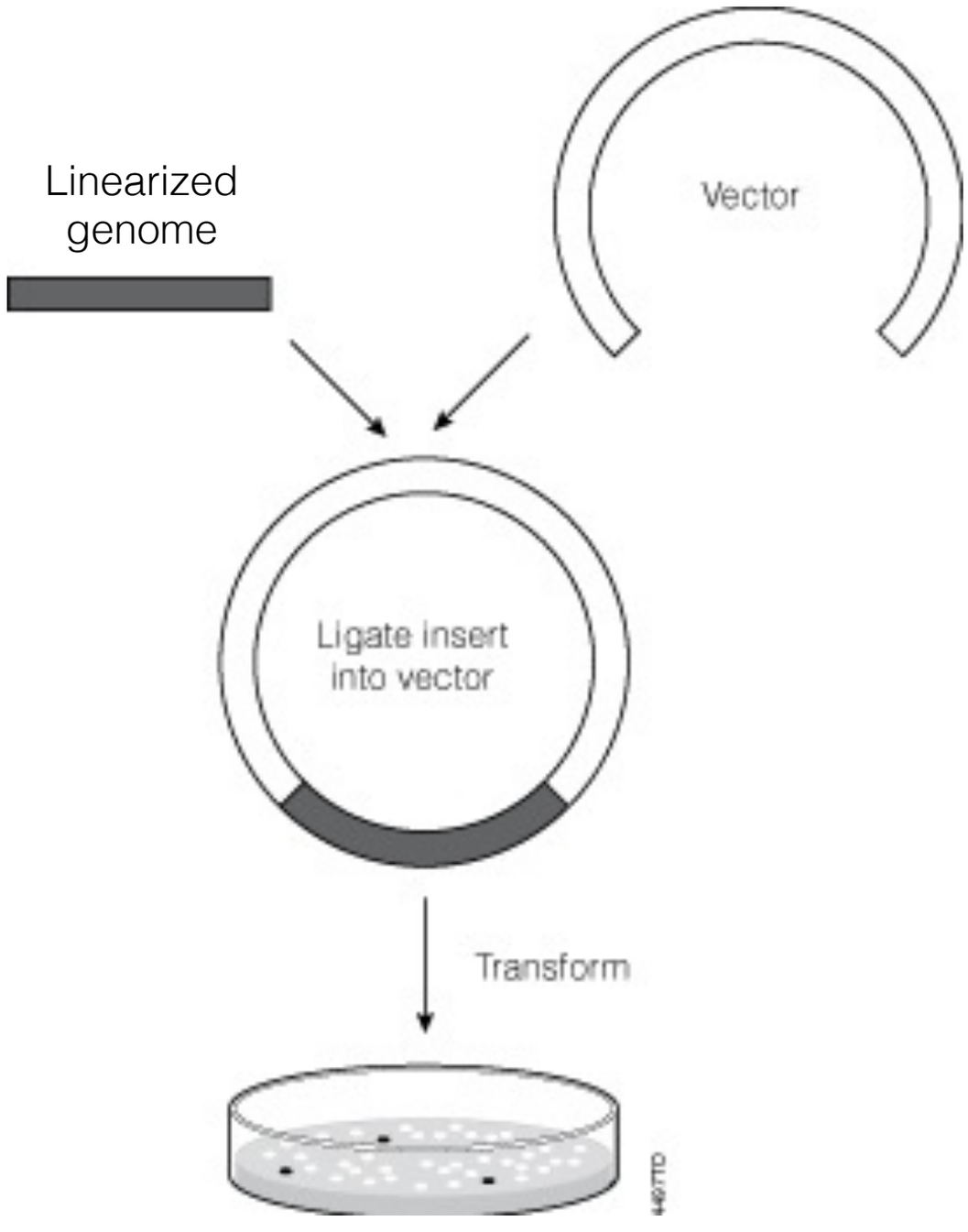
Illumina MiSeq (Nextera XT)

BK NGS Coverage Data



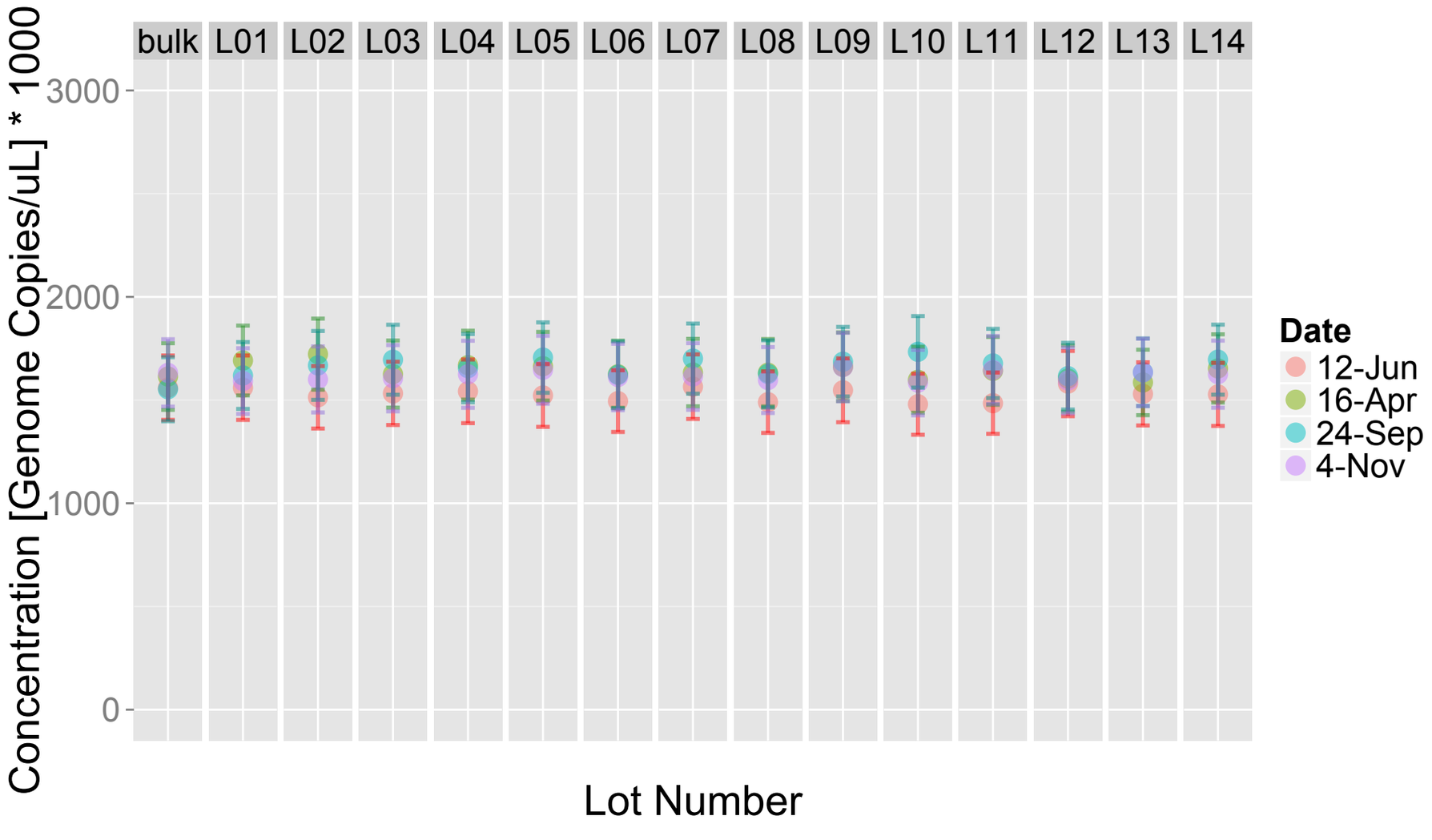
Plans for BKV Reference Material

- 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



Summary

- 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